Investigation of the PABPC1-BTG2 interaction and its effects on mRNA deadenylation by CNOT7/CAF1.

Aalam Mohammed Ameerul Moominin

A thesis submitted to the University of Nottingham for the
degree of Doctor of Philosophy.

September 2022

Abstract

Polyadenylation is the process of the addition of a poly(A) tail to the 3′ end of mRNA. The binding of human poly(A) binding protein to the poly(A) tail can either initiate or repress translation or regulate mRNA degradation. Cytoplasmic mRNA degradation in eukaryotes involves shortening of mRNA poly(A) tail by poly(A)-selective ribonuclease (deadenylase) enzymes. A recent study reported that PABPC1 could stimulate deadenylation by CNOT7/CAF1, a catalytic subunit of the CCR4-NOT complex, in the presence of BTG2. BTG2 is antiproliferative, regulating cell cycle progression in various cell types. BTG2 bridges PABPC1 and CNOT7/CAF1 and regulates deadenylation. This study aims to understand how PABPC1 and BTG2 interact, starting by identifying the amino acid residues of PABPC1 and BTG2 that contribute to the interaction.

The surface conserved PABPC1 residues predicted to interface with BTG2 were identified. Since PABPC1 has RNA bound to its structure, the expected PABPC1 interaction surface residues were compared with known PABPC1 residues interfacing with poly(A) RNA. Comparing residues led to identifying PABPC1 candidate residues free from RNA interfaces and residues required for interaction with BTG2. Subsequently, from the identified candidate residues, data-driven docking was used to generate a computational model depicting PABPC1-BTG2 interaction. The generated PABPC1-BTG2 model was used to create a quaternary model for poly(A) RNA-PABPC1-BTG2-CNOT7/CAF1. The quaternary model revealed that the 3′ end of the poly(A) RNA is accessible for the active site of CNOT7/CAF1. Site-directed mutagenesis studies were consequently conducted by introducing point mutations to candidate residues of PABPC1 and residues in Box C of BTG2. Subsequently, pull-down assays were performed to determine several amino acids residues of PABPC1 required for interaction with BTG2.

Further investigation of RNA binding for PABPC1/variants revealed that the PABPC1 residues necessary for interaction with BTG2 do not interfere with poly(A) RNA binding. The quaternary poly(A)-PABPC1-BTG2-CAF1 model created was validated by performing a functional assay. The functional assay revealed poly(A) RNA 3' end situated towards CNOT7/CAF1 catalytic centre suggests a rationale for enhanced deadenylation by CNOT7/CAF1 in the presence of BTG2 and PABPC.

Acknowledgement

Table of Contents

List of figures

List of Tables

Abbreviations

APS	Ammonium persulphate
AREs	Adenylate-uridylate-rich elements
AGO2	Argonaute protein 2
BTG	B-cell translocation gene
BL	Burkitt lymphoma
BRF	Butyrate response factors
CCR4-NOT	Catalytic module of carbon catabolite repression–negative on TATA-less
COSMIC	Catalogue of Somatic Mutations in Cancer
CIF	CCR4-NOT interacting motifs
CPSF	Cleavage and polyadenylation specific factors
CFIIm	Cleavage factor IIm
CFIm	Cleavage factor Im
CF IA	Cleavage factors IA
CF IIB	cleavage factors IIB
CPF	Cleavage polyadenylation factor
CSTF	Cleavage stimulation factors
CPORT	Consensus prediction of interface residues in transient complex
DCP1-2	Decapping protein 1 and 2
DLBCL	Diffuse large B-cell lymphoma
EDTA	EDTA
eDCP	Decapping enhancers
eIF4F	Eukaryotic initiation factor
EEP	Exonuclease-endonuclease-phosphatase
XRN 1	Exoribonuclease
FW	Forward primer
FL	Follicular lymphoma
GST	Glutathione S-transferase
GAP	GTPase activating protein
GTP	Guanine triphosphate
IPTG	Isopropyl β-2-1-thiogalactopyranoside (
LRR	Leucine-rich repeats
LiAc	Lithium acetate
LB	Lysogeny broth
mPAP	Mammalian poly(A) polymerase
m7G	Methylated cap structure - m7G
miRISC	Micro RNA induced silencing complex
MIF4G	Middle domain initiation factor 4G
NCOR	Nuclear hormone receptor co-repressor

NHL	Non-Hodgkin lymphoma
ER α	Oestrogen receptor α
OD	Optical density
PAM	PABP interacting, motifs
PCR	Polymerase Chain Reaction
PC3	Pheochromocytoma cell 3
PABP	Poly (A) binding proteins
PARN	Poly (A) specific ribonucleases
PABPN1	poly(A) binding protein nuclear
PAM	Poly(A) motifs
PAN	Poly(A) nucleases
PAS	Polyadenylation signal
PMR1	Polysome ribonuclease 1
PIC	Preinitiation complex
PRMT1	Protein arginine methyltransferase
PISA	Protein, Interfaces, Structure and Assemblies
RQCD1	Required for cell differentiation 1
RAR α	Retinoic acid receptor α
RNP1-2	Ribonucleic protein1-2
RV	Reverse primer
SDS	Sodium Dodecyl Sulfate
SMRT	Silencing mediator of retinoid and thyroid hormone receptor
SHM	Somatic hypermutation
TEMED	Tetramethyl-ethylenediamine
TC	Tertiary complex
Tis21	Tetradecanoyl phorbol acetate- inducible sequence 21
TSS	Transcription start site
TNRC	Trinucleotide repeat-containing
TBE	Tris-borate-EDTA
UCH	Ubiquitin C-terminal hydrolase
UTR	Untranslated region
YPD	Yeast extract peptone dextrose
YSC	Yeast selective complete

Chapter 1 – Introduction

1.1. Synthesis of messenger RNA

Gene expression is a process in which a gene synthesises functional molecules such as protein or non-coding RNA. In order to do this, the nucleotide sequence in a gene is interpreted by the regulatory mechanism in cells, and one of the 20 amino acids is added for each group of three sequence codes that are basic units for building a protein. The gene expression involves transcription, in which an RNA is copied from a DNA, and the translation process involves the synthesis of proteins from the molecular mRNA. In eukaryotes, messenger RNA (mRNA) and several small nuclear RNAs are transcribed by RNA polymerase II. RNA polymerase II cannot function independently; it requires transcription factors (TF). RNA polymerase II and several TFs binds to the transcription start sites of mRNA and small nuclear RNAs, located 25 base pairs upstream (Figure 1-1A, B).

Eukaryotic transcription involves three steps: initiation, elongation, and termination. The initiation process requires TFs assembling on the DNA with polymerase II (Butler & Kadonaga, 2002). The helicases in the TFs unwind double-stranded DNA, while RNA is synthesised by RNA polymerase II starting at the transcription start site (TSS; see Figure 1-1C; Svejstrup et al., 1996). RNA polymerase II dissociates from the promoter after synthesising a few bases, while the preinitiation complex disassembles.

Consequently, a second RNA polymerase II is recruited. The second RNA polymerase II is highly processive; it elongates RNA (Figure 1-1D) and generates many base pairs without dissociating (De Klerk & T Hoen, 2015). The termination process occurs when the polymerase reaches the end of the marked nucleic acid sequence in a gene; RNA polymerase II and the DNA template disassemble the elongation factors and initiate transcription termination by binding to the transcribed RNA (Figure 1-1E; Richard & Manley, 2009).

Following the termination, the cleavage and polyadenylation specific factors (CPSF) and cleavage stimulation factors (CSTF) assemble on the transcribed RNA surrounding the poly(A) tail signal. The transcript undergoes intron splicing before mRNA is processed and exported to the cytoplasm. Introns are non-coding segments of DNA located between two exons, and the spliceosome removes intron during RNA processing by RNA splicing. The 5′ end of the intron is the splice donor site, with an invariant sequence GU, within a less highly conserved region. In comparison, the 3′ end of the intron is the acceptor site that terminates the intron, containing an invariant AG sequence. Upstream from this consistent AG sequence, there is a region rich

with pyrimidine or polypyrimidine tracts and adenosine nucleotides involved in lariat formation
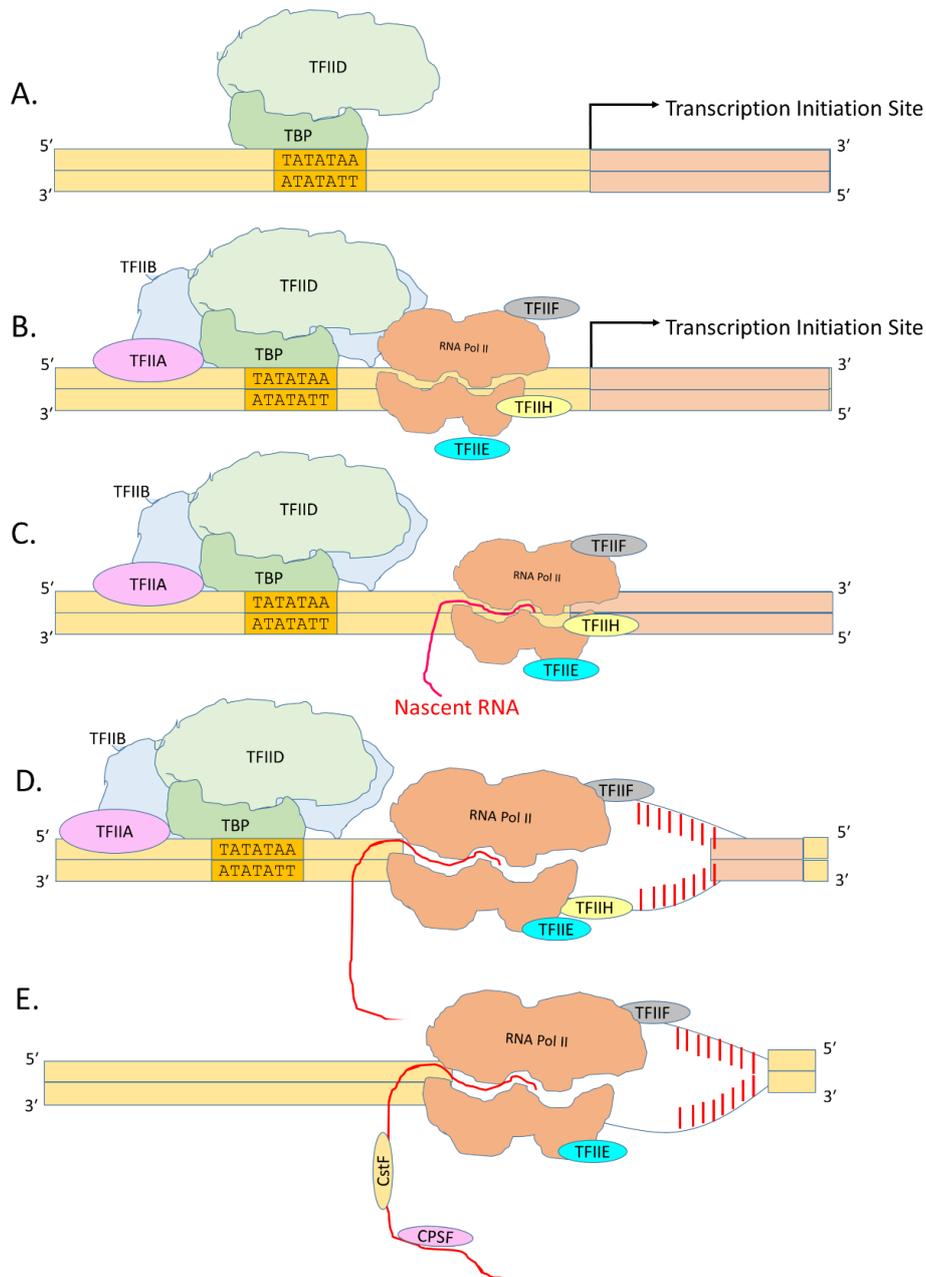
Figure 1-1. Factors regulating mRNA transcription. A schematic representation of transcription factors regulating transcription initiation, elongation, and termination. A) TFIID recognises the TATA site and binds to it via TBP protein. B) TFIID serves as the scaffold protein and initiates the assembly of TFs (TFIIB, TFIIA, TFIIF, TFIIH, TFIIE) along with RNA polymerase II. C) The helicase in TFIIH unwinds the DNA, the TSS in RNA polymerase II initiates transcription from the transcription activation site, synthesising a nascent RNA. D) RNA polymerase binds to the nascent RNA and initiates the elongation process. E) Transcription termination by RNA polymerase is initiated by polyadenylation cleaving factors, such as CPSF and CSTF. The binding of CPSF and CSTF to the transcribed RNA initiates transcription termination.

1.1.1. Characteristic traits of poly(A) binding proteins

Poly(A) binding proteins (PABPs) are classified based on whether they are cytoplasmic (PABPC1, PABPC3, PABPC4, PABPC4L), embryonic (ePABP), nuclear (PABPN1), or X-chromosome encoded proteins (PABPC5). Among PABP, PABPC1 is well known for its role in the post-translational modification of mRNA in the cytoplasm, namely in its translation, repression, and decay (Svitkin & Sonenberg, 2006, Meijer et al., 2019).

In general, cytoplasmic PABP contains four RNA recognition motifs (RRMs) in the N-terminal. Each RRM comprises four-stranded antiparallel β-strands backed by two α-helices (Figure 1-2A; Safaee et al., 2012). The β-sheet in RRM1-2 (Figure 1-2B) regulates poly(A) binding using two sequence motifs, such as ribonucleic protein1-2 (RNP1-2). These motifs are responsible for RNA binding. RNP1 and RNP2 are highly conserved in the β-sheets of RRM1-2 and consist of hydrophobic and positively charged residues for base recognition and negatively charged nucleic acid backbones (Safaee et al., 2012). In contrast, the C-terminal of PABPC1 has a PABP or MLLE domain with five α-helices (Figure 1-2C). MLLE (Met-Leu-Leu-Glu) was named after the highly conserved amino acids in the domain, which are linked with the RRM motif via a disordered Pro/Met-rich linker (Kozlov et al., 2010).

The nuclear PABPN is different from cytoplasmic PABP, as it only has one RRM, unlike PABP, which has four. The RRM of PABPN is canonical like that of PABPC1, with a four-stranded antiparallel β-sheet where the conserved motifs RNP1 and RNP2 are juxtaposed in the central strand (Ge et al., 2008). In contrast, *Xenopus* oocytes eggs and embryos express embryonic PABP. There are two ePABPs, the first of which contains only a single RRM with a cluster of negatively charged amino acids in the N-terminus and an Arg-rich sequence in the short C-terminus (Song et al., 2008). The second ePABP, which has four RRMs, and better researched for ePABP phosphorylation required for *Xenopus* oocyte maturation (Friend et al., 2012)

PABPs differ in length, domain organisation, and the homology of their amino acid. Sequence comparison of PABPC1, PABPN1, and ePABP revealed that PABPN1 and ePABP are nearly identical compared to PABPC1 (Figure1-3; Eliseeva et al., 2013). The first interaction between PABPC1 and mRNA occurs in the nucleus, prior to PABPC1 binding to processed mRNA in the cytoplasm. Along with PABPN, PABPC1 associates with the poly(A) tail of newly synthesised transcripts in the nucleus and regulates the first round of translation via the cap-binding protein (CBP; Hosoda et al., 2006). PABPC1 plays a crucial role in regulating

translation by binding to the processed mRNA's poly(A) tail in the cytoplasm. The RRM2 domain of PABPC1 serves as the binding site for the eukaryotic initiation factor (eIF4G), a translation initiation factor that mediates mRNA circularisation with other eIFs to regulate translation (Safaee et al., 2012, Gruner et al., 2016).

PABPC1 protects mature mRNA from deadenylation via CNOT7/CAF1, a poly(A) nuclease and a component of the CCR4-NOT complex (Yi et al., 2018). Furthermore, PABPC1 regulates mRNA decay by promoting deadenylation via poly(A) nucleases through poly(A) motifs (PAMs; Kozlov et al., 2010, Schafer et al., 2014). In addition to recognition by PAM motifs, PABPC1 also interacts with either CNOT6/CCR4a or CNOT6L/CCR4b to regulate deadenylation mediated by the catalytic module of the carbon catabolite repression–negative on TATA-less (CCR4-NOT) complex (Yi et al., 2018). PABPC1 also mediates interactions with the B-cell translocation gene (BTG) family of proteins to regulate deadenylation by either CNOT7/CAF1 or CNOT8/CAF1b, an alternative catalytic module of the CCR4-NOT complex (Stupfler et al., 2016; Ezzeddine et al., 2007).
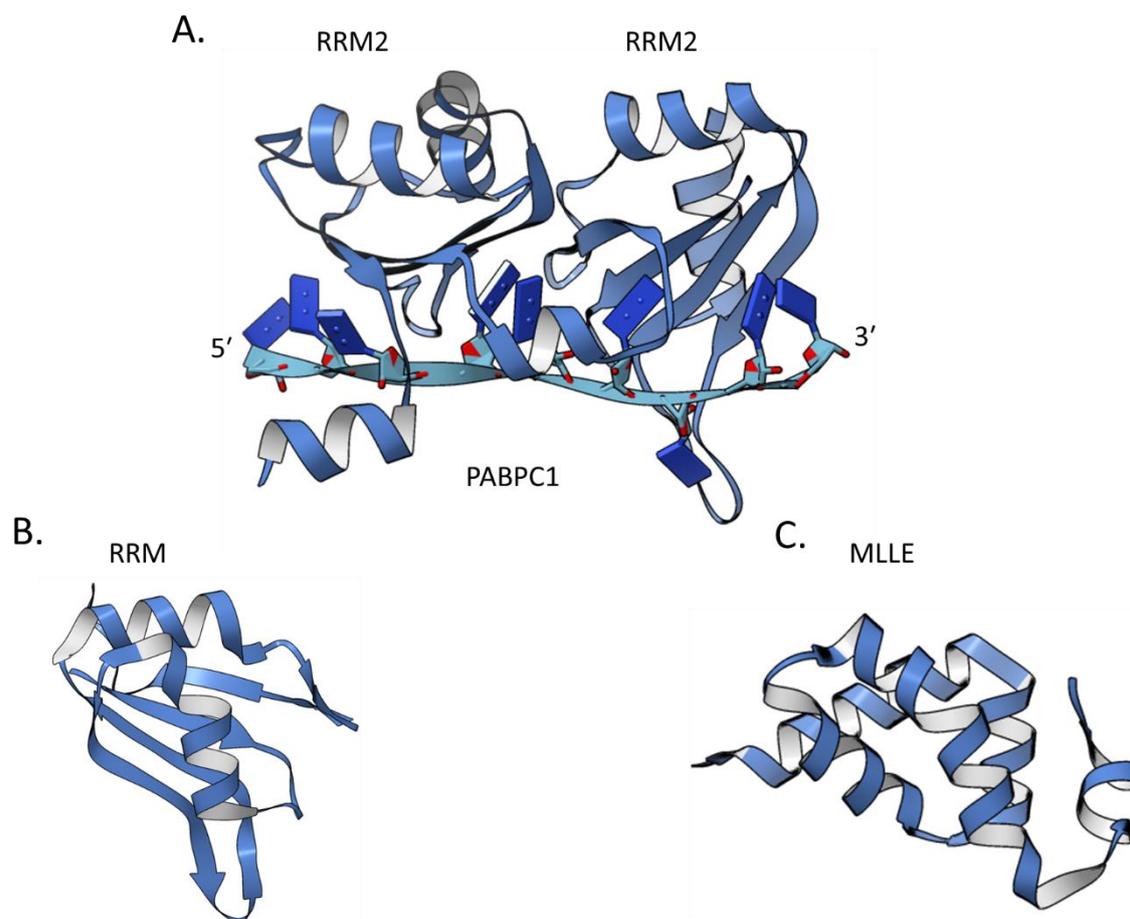
Figure 1-2. Crystal structure of PABPC1 with RRM and MLLE domains. The structure of PABPC1 contains the RRMs and the MLLE domains. A) The structure of PABPC1 (PDB 4F02) with RNA wound around the RRM1 and RRM2 (Safaee et al., 2012). B) The RRM domain is present in the N-terminal of PABPC1, independent of RNA. C) The MLLE domain in the C-terminal of PABPC1 (Kozlov et al., 2010).

Figure 1-3. Sequence analysis for PABP. A) Sequence analysis of the PABP family of proteins compared to PABPC1 in ENDscript (Robert & Gouet, 2014). Amino acids highlighted in red are identical residues found in the PABPC1, PABPN, and ePABP proteins. The regions highlighted in yellow are amino acids that differ in the PABP family of proteins.

1.1.2. Synthesis of mRNA poly(A) tail in the nucleus

Polyadenylation is the process of adding adenosine nucleotides to the 3′ end of mature mRNA. The term poly(A) tail refers to the stretch of added adenosine nucleotides in the 3′ end mature mRNA. The poly(A) tail plays a vital role in mRNA's transcriptional and post-transcriptional modification. CPSF regulates the synthesis of poly(A) tails with adenosine nucleotides at the 3′ end of fully processed eukaryotic mRNA (Tian et al., 2005).

In higher eukaryotes, polyadenylation is regulated by factors such as Cleavage factor Im (CFIm), cleavage factor IIm (CFIIm), CSTF, and CPSF. Whereas in the yeast *Saccharomyces cerevisiae* Cleavage factor IA (CF IA), cleavage factor IIB (CF II B), and cleavage polyadenylation factor (CPF) regulate both cleavage and polyadenylation (Kaufmann et al., 2004, Viphakone et al., 2008, Preker et al., 1997).

During the initial stage of poly adenylation, CPSF binds to the polyadenylation signal (PAS), an AAUAAA hexamer sequence located between 10–30 nucleotides upstream of the terminal sequence in mammalian mRNA (Figure 1-4 A). CSTF binds 20–40 nucleotides downstream of the cleavage site that is rich with U or U/G elements (Figure 1-4 B; Bienroth et al., 1993, Tian et al., 2005). CPSF and CSTF tether the mRNA to mammalian poly(A) polymerase (mPAP) to process the addition of adenosine bases to the 3′ end of the mRNA (Figure 1-4 C; Bienroth et al., 1993).

The mRNA tethered to mPAP is unstable due to its low affinity; since the mRNA is unstable, the mPAP is in distributive mode generating 11 adenosine residues (Figure 1-4 C). The mRNA stabilises with the tethered mPAP upon PABPN1 binding. PABPN1 binds to the first 11 polymerised adenosine residues (Kerwitz et al., 2003). The PABPN1 binding to polymerised adenosine residues causes the mPAP to switch from the distributive to processive mode, thus stimulating the synthesis of poly(A) tail (Figure 1-4 D; Wahle, 1995, Kerwitz et al., 2003) when the poly(A) tail reaches an approximate value of 250 adenosine residues, mPAP switches back to its distributive mode (Wahle, 1995).
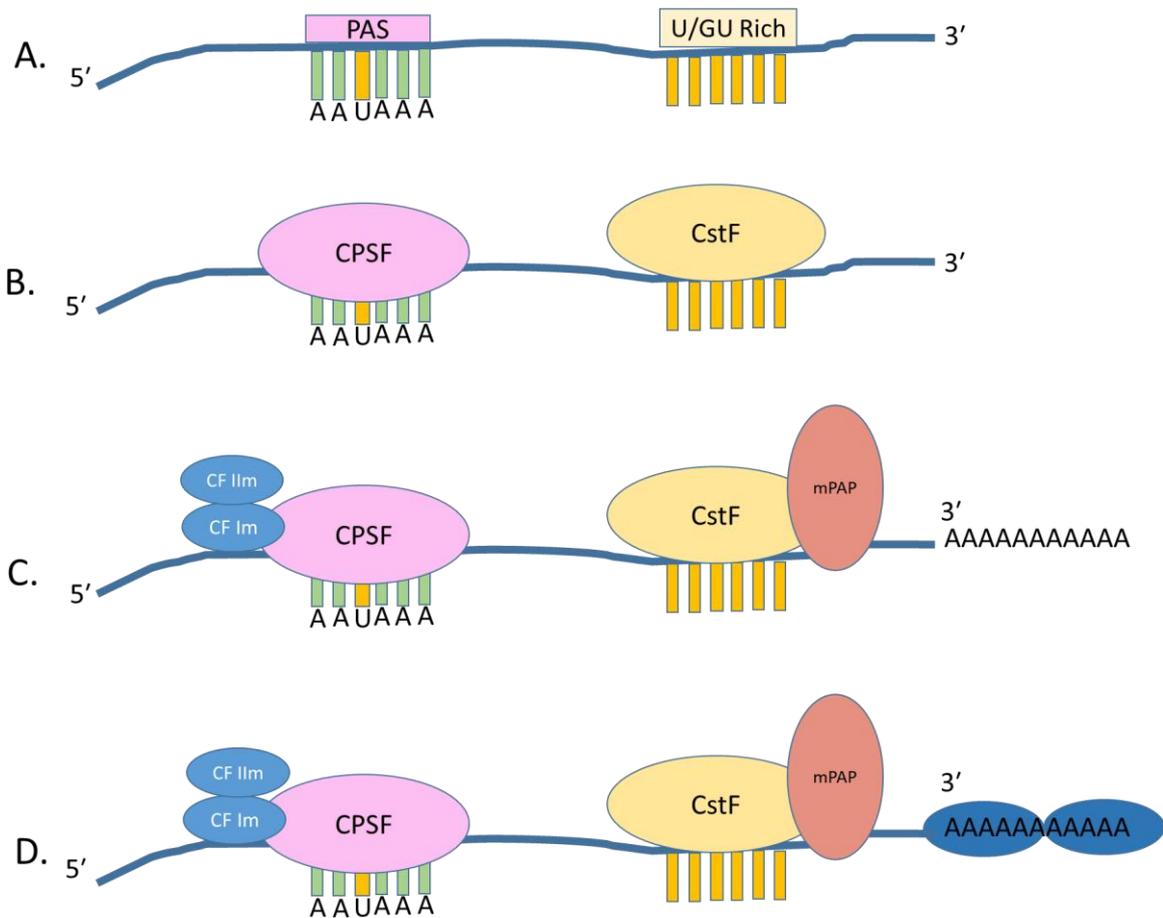
Figure 1-4. Factors regulating the addition of poly(A) tail to the 3′ end of mRNA. A schematic representation of the factors involved in adding adenosine nucleotides to the 3′ end of mRNA. **A)** The polyadenylation signal (PAS) located upstream of mRNA with 10–30 nucleotides of a consensus AAUAAA hexamer, followed by a downstream U/GU-rich region of 20–40 nucleotides. Together, these initiate the polyadenylation reaction. **B)** Binding of CPSF to the upstream PAS region and the binding of CSTF to the downstream U/GU-rich region**. C)** CF Im and CF IIm associate with CPSF, followed by mPAP interaction with CSTF. **D)** PABPN1 interacts with 11 polymerised adenosine residues and stabilises mPAP for processing the poly(A) tail.

1.1.3. Translation factors regulating mRNA closed-loop structure

In eukaryotes, translation is regulated by eukaryotic initiation factors 1 (eIF1), 2 (eIF2), 3 (eIF3), and 4 (eIF4). Eukaryotic translation machinery identifies initiation codons by forming a ternary complex (TC) comprised of methionyl-tRNA (Met-tRNA) and guanine triphosphate (GTP)-bound eIF2, which binds to the small 40S ribosomal subunit. The binding of the TC to the 40S subunit is promoted by eIF1, eIF1A, eIF5, and eIF3, forming the 43S preinitiation complex (Marintchev & Wagner, 2004). Multiple copies of PABPC1 bind to processed mRNA's poly(A) tracts in the cytoplasm. The poly(A)-bound PABPC1 yields a repeating pattern of 27 nucleotides, stabilising the mRNA for translation (Figure 1-5A; (Webster et al., 2018). Cap-binding translation is mediated by eIF4F by recognising the methylated cap structure (m^7G) at 5' of eukaryotic mRNA (Svitkin & Sonenberg, 2006).

The protein complex of eIF4F is a tripartite subunit, comprised of eIF4G, eIF4E (a cap-binding protein), and eIF4A (an ATP-dependent RNA helicase; Figure 1-5B). The protein eIF4G is a scaffold protein containing the binding domain for PABPC1, eIF4E, eIF4A, and eIF3 (Svitkin & Sonenberg, 2006). The protein eIF4G interacts with PABC1 bound to poly(A) tracts in the 3′ end of mRNA (Safaee et al., 2012). The RNA binding property of eIF4G enables it to co-ordinate with mRNA via cap-bound eIF4E, poly(A) tails, and sequences in the mRNA to assemble a stable circular messenger ribonucleoprotein (mRNP) in a closed-loop structure (Figure 1-5C; Gruner et al., 2016).

The protein eIF3 promotes the recruitment of a ternary complex to the 40S subunits and mRNA to the PIC (Figure 1-5D). The interaction between eIF4G and eIF3 forms a protein bridge between the activated mRNP and 43S PIC to stimulate the attachment of 43S to the mRNA (Figure 1-5E; Chu et al., 2016). The helicase activity of eIF4A generates a single-stranded landing pad in the mRNA to which 43S PIC can load. The 43S PIC bound to m^7G scans the mRNA for the AUG start codon; AUG recognition arrests PIC and converts the eIF2-GTP complex in the TC to eIF2-GDP (Chu et al., 2016, Hinnebusch & Lorsch, 2012).

The protein eIF2 hydrolyses the GTP in the complex but requires an additional factor, such as eIF5, a GTPase activating protein (GAP); further initiation codon AUG subsequently pairs with the Met-tRNA anticodon to induce GTP hydrolysis (Kapp & Lorsch, 2004). As a result, eIF2-GTP is converted to eIF2-GDP complex, and eIFs present in the PIC is released to join the 60S subunit. Both eIF2-GDP complex and eIFs are then catalysed by eIF5B to produce an 80S

initiation complex containing Met-tRNA base-paired to AUG in the P site. Following this, eIFs are ready to start the elongation phase of protein synthesis (Kapp & Lorsch, 2004).
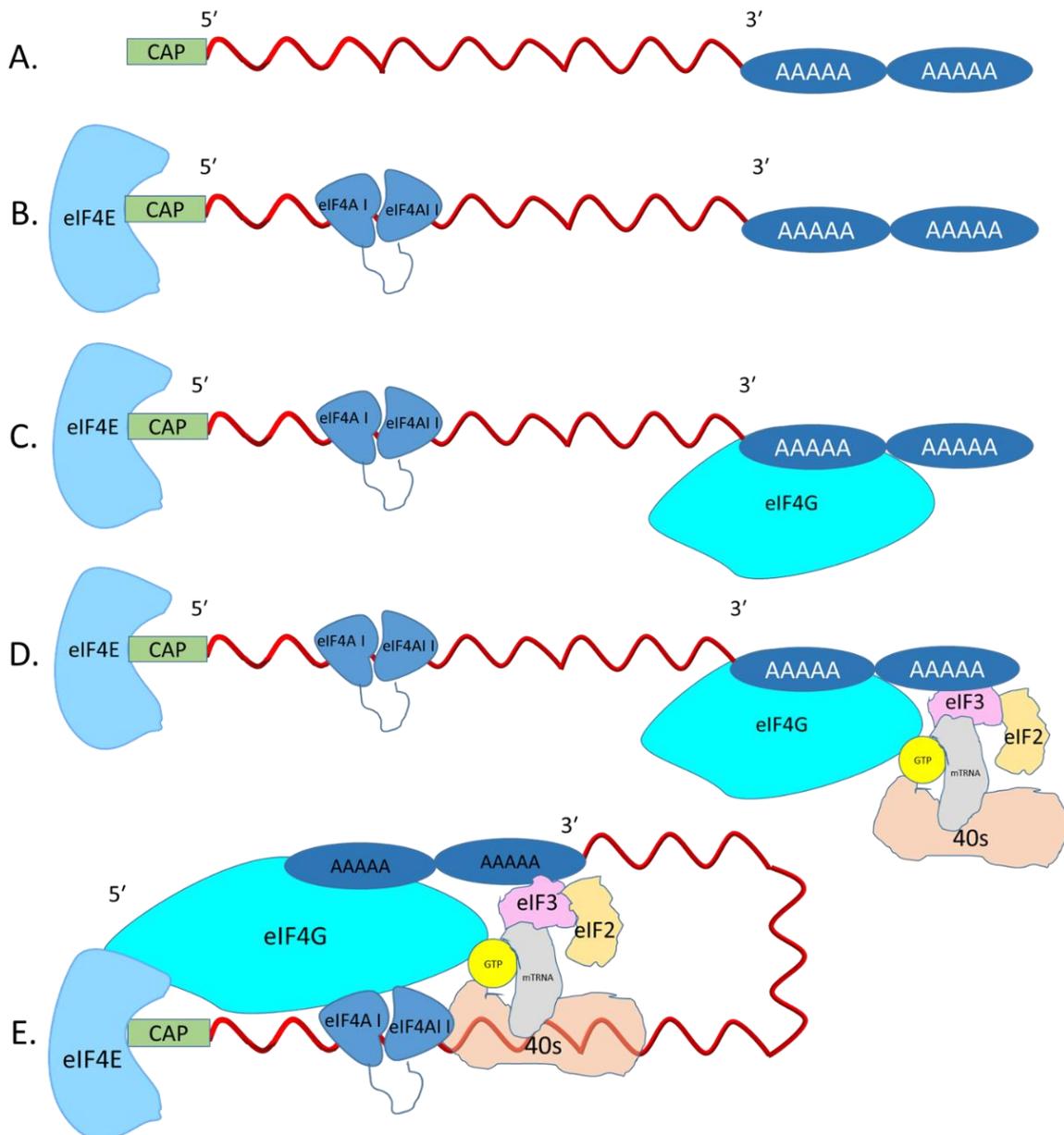
Figure 1-5. Factors involved in translation initiation. A schematic representation of factors regulating translation. **A)** poly(A) binding protein bound to the poly(A) tail at the 3′ end yielding at a repeating pattern of 27 nucleotides. **B)** eIF4E binds to m7G at the 5′ end, followed by the binding of ATP-dependent RNA helicases eIF4AI and eIF4A II. **C)** eIF4G is bound to the RRM2 of PABPC1. **D)** eIF3 interacts with PABPC1 along with eIF2, Met-tRNA, and 40S ribosomes. **E)** eIF4G bound to the RRM2 of PABPC1 circularises the mRNA to initiate translation by associating the 40S ribosome subunit with eIF4F.

1.2. Proteins regulating deadenylation and mRNA degradation

The length of the poly(A) tail can initiate or repress translation or deadenylation. In other words, the poly(A) tail can determine the fate of mRNA in the cytoplasm (Meijer et al., 2019). The length of the poly(A) tail at the 3′ end of the mRNA depends upon the organism, for instance, the length is initially 250 nucleotides in higher eukaryotes and around 70–80 nucleotides in yeast (Morrissey et al., 1999, Brown & Sachs, 1998). The degradation of mRNA plays a vital role in post-transcriptional gene regulation platforms. Typically, mRNA decay occurs via one of three pathways: (a) deadenylation-dependent decay, (b) endonucleolytic cleavage-dependent decay, or (c) adenylation-independent decay. The deadenylation-dependent decay pathway is a two-stage process involving poly(A) tail shortening followed by 5′ cap removal, which results in the subsequent degradation of mRNA (Parker & Song, 2004, Nagarajan et al., 2013).

The initial stage in mRNA degradation is deadenylation by poly(A) nucleases, which regulates the removal of poly(A) tails from the 3′ end of processed mRNA (Brown & Sachs, 1998). Deadenylation occurs in both the nucleus and cytoplasm; however, nuclear deadenylation is not well understood in eukaryotes. In yeast, Pan2-Pan3 deadenylates some of the adenosine residues in the poly(A) tail, with the remaining residues removed by the Ccr4-Not complex (Yamashita et al., 2005). The following sections explain the classification of poly(A) nucleases, deadenylation of poly(A) tails in both the nucleus and cytoplasm and mRNA degradation.

1.2.1. Classification of poly(A) nucleases

The human genome encodes 11 deadenylases into two categories (Table 1-1). The first type is the Rnase D superfamily, characterised by DEDD amino acids (Asp-Asp-Glu-Asp) in the catalytic domain (Siwaszek et al., 2014). DEDD amino acids coordinate with divalent metal ions in the active site and regulate enzyme catalytic activity (Petit et al., 2012, Jonstrup et al., 2007, Andersen et al., 2009). The DEDD group of deadenylases include CNOT7/CAF1 (Pop2), PAN2-PAN3, poly(A) specific ribonucleases (PARNs), and CAF1Z. The second type of deadenylases are categorised as the exonuclease-endonuclease-phosphatase (EEP) superfamily and are characterised by Asp and His amino acids in the catalytic domain (Siwaszek et al., 2014). Conserved catalytic residue Asp and His determine the enzyme's active site, and deadenylase enzymes from this type include CCR4, NOCTURN, ANGEL, and 2′PDE (Bartlam & Yamamoto, 2010).

Gene	Synonym	Catalytic Domain	Other Domain	kDa
Poly (A) specific	PARN			
ribonucleases	DAN	DEDD	Cap binding	73.4
Poly (A)	PAN2			135
nucleases	PAN3	DEDD	PAM2	
				95.6
CNOT7/CAF1	CAF1	DEDD	NOT	32.7
CNOT8/CAF1b	CALIF	DEDD	NOT	33.5
	POP2			
Caf1z		DEDD	NOT	56.5
CNOT6/CCR4a	CCR4a	EEP	NOT	63.5
CNOT6L/CCR4b	CCR4b	EEP	NOT	63
	Protein angel			
Angel2	homolog 2	EEP	-	31
	Protein angel			
Angel	homolog 1	EEP	-	75.3
	2',5'-			
PDE12	phosphodiesterase	EEP	-	67.3
	12			
CCR4NL	Nocturnin	EEP	-	48.1
CNOT4	NOT4		RING RRM	63.5
CNOT9	RQCD1		Armadillo repeat	33.6
(RCCD1)	RCC domain 1			82.3
CNOT10				

Table 1-1. Classification of deadenylase enzymes. Poly(A) nucleases are classified based on DEDD or EEP residues in the catalytic domain. An alternative domain to which molecules are associated and deadenylases molecular weight are indicated.

1.2.2. Structure and characteristic traits of PAN2/Pan2-PAN3/Pan3

PAN/Pan was the first deadenylation enzyme to be discovered. The protein is a complex of PAN2/Pan2p-PAN3/Pan3p. PAN proteins are present in mammals (PAN2/PAN3) and yeast (Pan2p/Pan3p). Orthologues include *Drosophila melanogaster* (*D. Melanogaster*) and *Caenorhabditis elegans* (*C. Elegans*; Schafer et al., 2014; Figure 1-6A). The PAN2/Pan2p-PAN3/Pan3p complex plays a vital role in poly(A) tail removal from the 3′ end of processed mRNA, which it does by binding to PABPC1 (Sachs and Deardorff, 1992).

Pan2 of *S. Cerevisiae* comprises 1,115 residues, a WD40 domain, and a short connecting segment named the ubiquitin C-terminal hydrolase (UCH) in the N-terminus. A linker region in Pan2 separates the WD40 domain from the UCH domain. The C-terminal comprises the DEDD family's RNase domain (Figure 1-6B; Wolf et al., 2014, Schafer et al., 2014). The C-terminal region in Pan3 of *S. Cerevisiae* has 679 residues, a pseudokinase domain, a coiled-coil domain, and a C-terminal knob domain (Schafer et al., 2014, Wolf et al., 2014). The N-terminal contains CCCH-type zinc fingers held together by zinc ions. The zinc co-ordinating residues Cys14, Cys23, Cys30, and His34 are conserved in the zinc fingers of Pan3 and stimulate interaction with poly(A) RNA (Wolf et al., 2014). An additional 250 residues in the N-terminal, along with PAMs, mediate interaction with PABPC1 (Figure 1-6B; Schafer et al., 2014, Wolf et al., 2014).

PAN2/Pan2 catalytic activity is dependent on the coordination of the conserved DEDD residues with metal ions (Wolf & Passmore, 2014). PABPC1 assists in regulating the catalytic activity of PAN/Pan molecules. PAN3/Pan3 mediates PABPC1-dependent activity by binding to the MLLE domain of PABPC1 through a PABP-interacting motif located in the N-terminal region (Uchida et al., 2004, Siddiqui et al., 2007).

A.
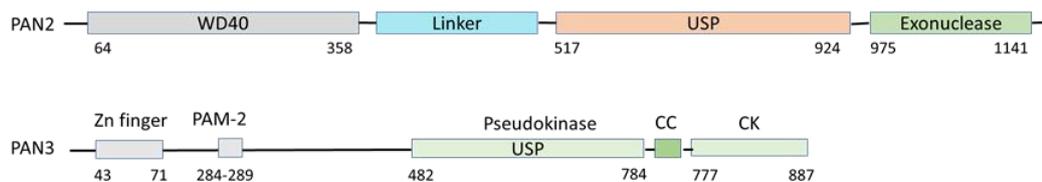
B.
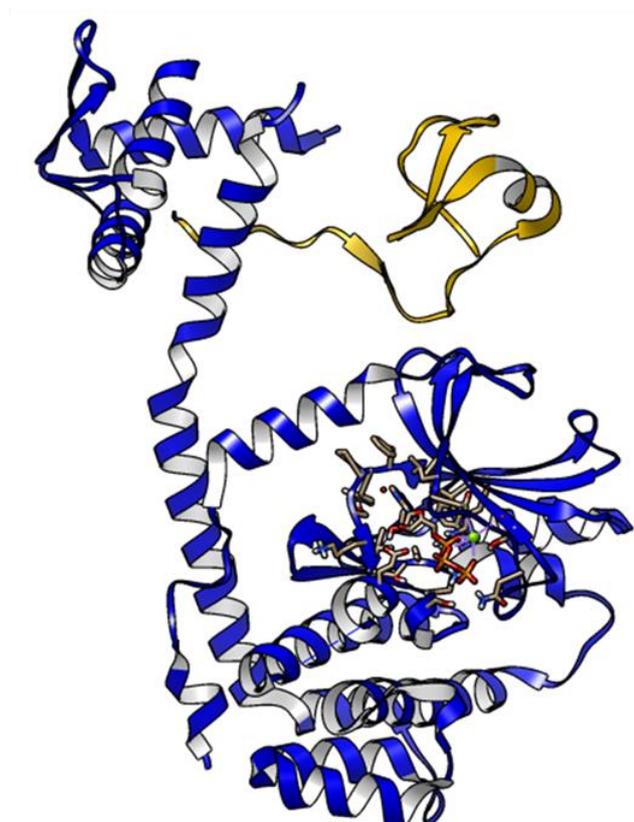
Figure 1-6. Structural characterisation of PAN molecules. A) Schematic representation of PAN2/Pan2p and PAN3/Pan3p molecules, indicated are the WD40 domain, the linker region connecting the WD40 and USP domains, and the RNAse domain of PAN2/Pan2p. At the same time, the PAN3/Pan3p represents the Zn finger domain, PAM 2, pseudokinase domain, CC and CK domains in the C-terminal. **B)** The crystal structure of the PAN2-PAN3 molecule. The ribbon in blue is the PAN3 molecule; the protruding sticks are the conserved residues in coordination with Mg^{2+} ions, represented as a green ball. The ribbon in gold is the PAN2 molecule (Wolf et al., 2014).

1.2.2.1. Deadenylation by Pan2-Pan3 proteins in the nucleus

In yeast, Pab1p is critical for regulating deadenylation in the nucleus, which it does by binding to PAM motifs in the N-terminal of Pan3 (Schafer et al., 2014, Wolf et al., 2014, Brown & Sachs, 1998). A study by Brown & Sachs (1998) discovered that the Pab1p variant negatively interacts with Pan3, which affected deadenylation by Pan2-Pan3 and resulted in a longer poly(A) tail.

The poly(A) tail of *S. Cerevisiae* is initially 70–80 nucleotides long before it is exported to the cytoplasm and it undergoes the first stage of deadenylation in the nucleus. The action of Pan2-Pan3 results in the removal of 10 nucleotides (Brown & Sachs, 1998). The deadenylated mRNA poly(A) tail consequently has a length of 60–70 nucleotides when exported to the cytoplasm, which causes fewer Pabp1 to bind, thus affecting its interaction with Pan2-Pan3 and resulting in no further deadenylation. The reduced ability of Pan2-Pan3 to bind to Pabp1 affects poly(A) tail removal, which in turn activates deadenylation by CCR4-NOT to remove the remaining nucleotides in the poly(A) tail (Brown & Sachs, 1998). The process of nuclear deadenylation in eukaryotes is unclear; however, cytoplasmic deadenylation occurs biphasically (Yamashita et al., 2005).

1.2.2.2. Biphasic deadenylation regulated by PAN2-PAN3 in the cytoplasm

In higher eukaryotes, mRNA poly(A) tails of 200 nucleotides in length go through biphasic deadenylation (Figure 1-7). The PABPC1 bound to mRNA protects the poly(A) tail from deadenylation by CNOT7/CAF1 (Yi et al., 2018). Therefore, the removal of poly(A) tail by PAN proteins is essential during the initial stage of deadenylation. When a steady state mRNA poly(A) tail of 200 nucleotides arrives in the cytoplasm, multiple PABPs associate with the poly(A) tract, and this provides sufficient interaction with PAN2-PAN3 to result in the shortening of the poly(A) tail to 110 nucleotides (Yamashita et al., 2005, Wolf et al., 2014, Wolf & Passmore, 2014). The poly(A) tail, after being deadenylated to 110 nucleotides, associate with fewer PABPC1s, which reduces the ability of PAN molecules to bind with PABPC1, thereby diminishing PAN2-PAN3 and increasing CCR4-NOT activity. The reduced poly(A) tail affects PABPC1 binding, making the 3′ end of the poly(A) tract accessible for deadenylation by the CCR4-NOT complex following decapping (Yamashita et al., 2005, Wolf et al., 2014, Wahle & Winkler, 2013).
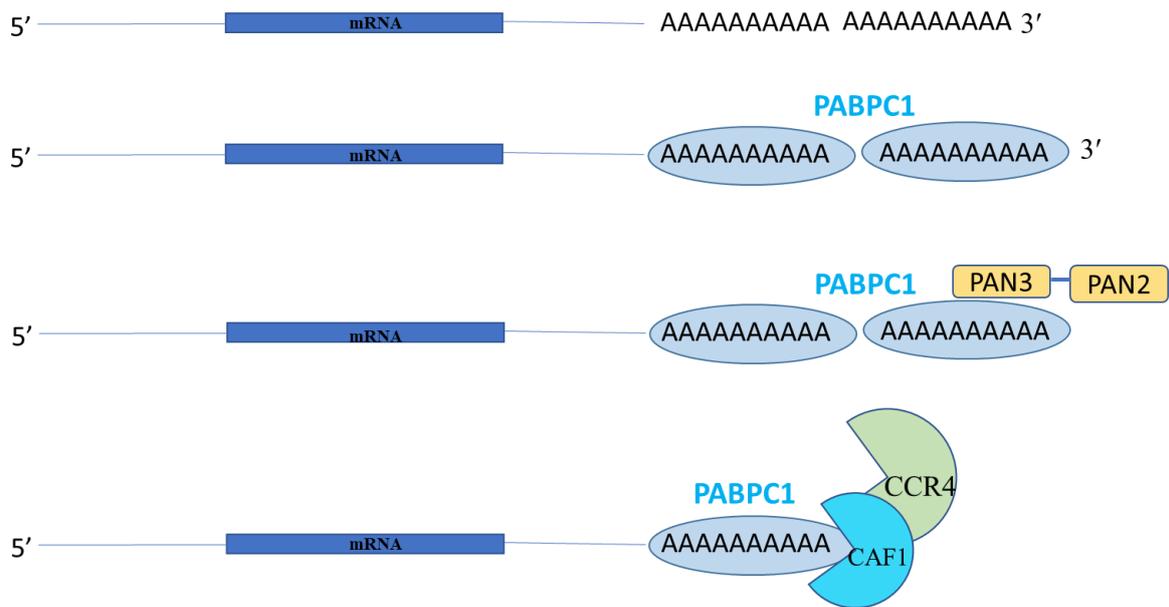
Figure 1-7. Biphasic deadenylation regulated by PAN2-PAN3 and catalytic subunits of the CCR4-NOT complex. A pictorial representation of the molecules of PABPC1 bound to the mRNA's poly(A) tail, which, in turn, recruits PAN molecules to regulate deadenylation. Upon removing 100 nucleotides from the poly(A) tail, PAN dissociates, and the CCR4-NOT complex deadenylates the remaining nucleotides.

1.2.3. The role of the CCR4-NOT complex in mRNA decay

CCR4-NOT is a multi-subunit complex that regulates deadenylation via catalytic subunits and translation repression via NOT subunits (Meijer et al., 2019, Bartlam & Yamamoto, 2010). The structure of the CCR4-NOT complex includes an N-terminal region, MIF4G domain, nuclease domain required for cell differentiation 1 (RQCD1), and a NOT module containing eight subunits with two catalytic and six non-catalytic subunits (Figure 1-8; Xu et al., 2014, Bartlam and Yamamoto, 2010).

The MIF4G domain in CNOT1 of the CCR4-NOT complex acts as a scaffold subunit in the assembly of catalytic and non-catalytic subunits (Boland et al., 2013). The catalytic subunit regulates deadenylation by the CCR4-NOT complex. It comprises yeast Ccr4 and human CNOT7/CAF1, CNOT8/CAF1b orthologues, CNOT6/CCR4a, and CNOT6L/CCR4b orthologues (Figure 1-9A; Bartlam & Yamamoto, 2010). Deadenylation is a two-stage process in which the CCR4-NOT complex plays a critical role by removing the poly(A) tail in the second stage after PAN2/Pan2-PAN3/Pan3 can no longer deadenylate (Wahle & Winkler, 2013, Wolf & Passmore, 2014). As the PAN/Pan proteins shorten the poly(A) tail, fewer PABPC1s can bind, leaving the 3' end of the poly(A) tail exposed. Following this, the CCR4-NOT complex can remove the remaining nucleotides (Yamashita et al., 2005).

Alternative proteins, such as BTG1/BTG2 and TOB, regulate deadenylation by bridging CNOT7/CAF1 and PABPC1 (Figure 1-9B; Stupfler et al., 2016). BTG and TOB bind to CNOT7/CAF1 with their similar BTG domains. BTG1/BTG2 interact with PABPC1 through residues in the Box C region, while TOB interacts with PAM motifs in the extra-long C-terminal (Stupfler et al., 2016, Ezzeddine et al., 2012).

The DEAD-box helicases, such as eIF4A2 and DDX6, compete to bind CNOT1 to regulate CNOT7/CAF1 activity. The binding of DDX6 to CNOT1 stimulates the deadenylation activity of CNOT7/CAF1. By contrast, the binding of eIF4A2 to the MIF domain of CNOT1 represses deadenylation by CNOT7/CAF1. An additional function of the DEAD-box helicases is to regulate translation repression. DDX6 and eIF4A2 bind to the MIF domain of CNOT1 to regulate translation repression by CNOT7/CAF1 (Figure 1-9C; Meijer et al., 2019).

The CCR4-NOT complex can bind to GW182 to repress translation, independent of CNOT7/CAF1. The molecule's interaction with GW182 displaces PABPC1 from the target mRNA's poly(A) tail. This displacement disrupts the structure formed by the interaction between eIF4G and PABPC1, resulting in translation repression (Zekri et al., 2013). The CCR4-NOT complex is also known to interact with proteins essential for the micro-RNA-induced silencing complex (miRISC). Argonaute proteins recruit GW182 to miRNA that targets mRNA for silencing, which in turn recruits CCR4-NOT in a PABP-independent manner using the WG motif (Fabian et al., 2011).
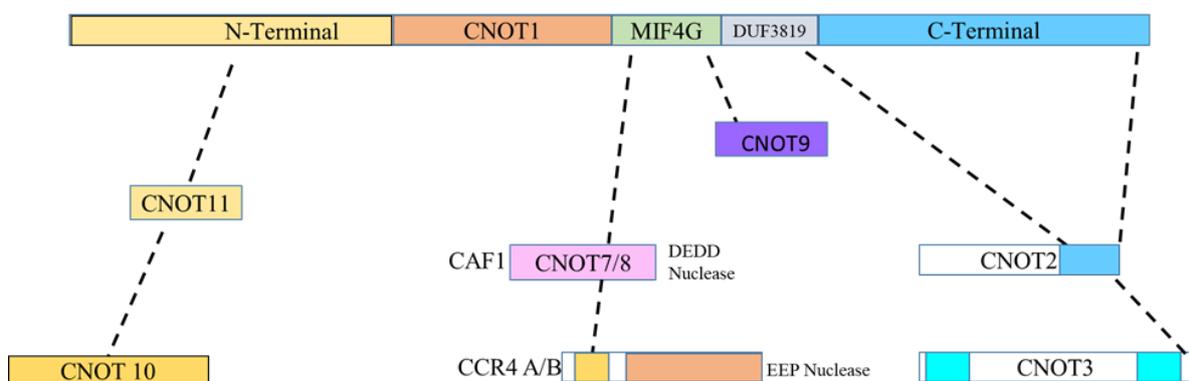
Figure 1-8. A schematic map of the CCR4-NOT complex. The map represents a domain in the N-terminal with CNOT11/C2orf29 and CNOT10. At the same time, the MIF4G domain of the CNOT1 harbours CNOT7/CAF1, CNOT8/CAF1b, and CNOT6/CCR4a CNOT6L/CCR4b. A DUF3819 domain binds CNOT9/RQCD1 and the C-terminal with CNOT2 and CNOT3.
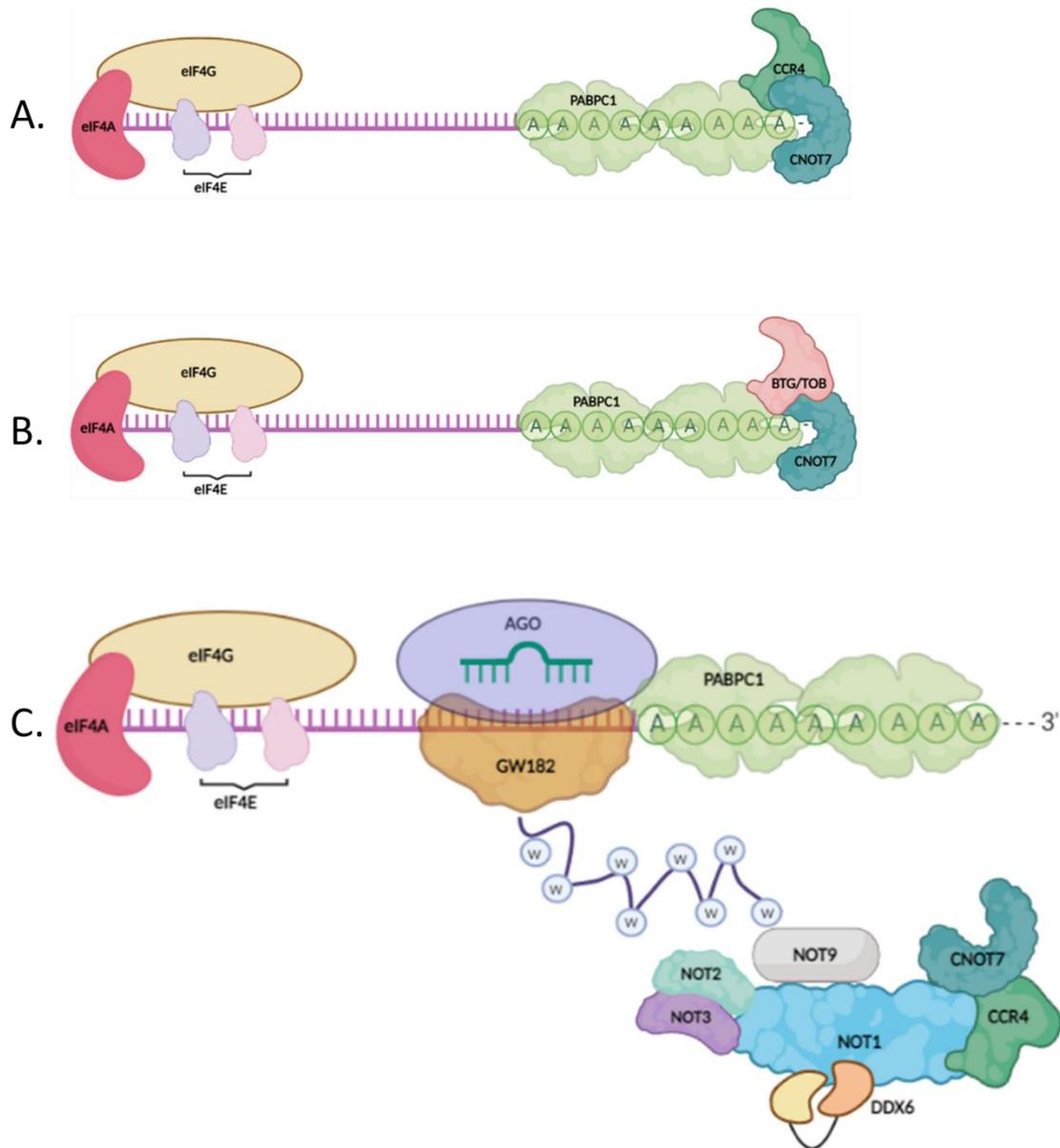
Figure 1-9. Factors regulating deadenylation by catalytic subunits of the CCR4-NOT complex. A) The figure depicts deadenylation by CNOT7/CAF1 with CNOT6/CCR4a bound to PABPC1. **B)** A schematic representation of BTG2 bridging PABPC1 and CNOT7/CAF1 to regulate deadenylation. **C)** The image depicts the interaction of the CCR4-NOT complex with GW182 of the RISC complex of DEAD-box helicases bound to the CNOT1 subunit of the CCR4-NOT complex.

1.2.4. The catalytic module of the CCR4-NOT complex

The catalytic subunit of CCR4-NOT mediates deadenylation activity and comprises CNOT7/CAF1, CNOT8/CAF1b, CNOT6/CCR4a, and CNOT6L/CCR4b. The MIF4G domain in CNOT1 forms a tertiary structure with CNOT6L/CCR4a and CNOT7/CAF1 at the N-terminal (Figure 1-10A; Xu et al., 2014). The concave surface in the B-helices of the MIF4G consists of 1–5 HEAT repeat motifs, which contain conserved residues to mediate interaction with CNOT7/CAF1 (Figure 1-10B; Petit et al., 2012). The leucine-rich repeats (LRR) of CNOT6L/CCR4a interact with CNOT7/CAF1, which functions as the bridge to connect CCR4 to the CCR4-NOT complex (Wang et al., 2010, Malvar et al., 1992).

CNOT7/CAF1 and CNOT8/CAF1b belong to the RNaseD superfamily in mammals and humans, while Caf1/Pop2 in yeast *S. cerevisiae* also belongs to this superfamily. The conserved residues Asp and Glu are present in the 3′ exonuclease motif of the protein (Petit et al., 2012, Jonstrup et al., 2007, Andersen et al., 2009). CNOT6/CCR4a is a subfamily of Rnases; it belongs to the EEP family of proteins and has an exonuclease domain. Moreover, its human orthologue, CNOT6L/CCR4b, is known to control poly(A) removal and regulate mRNA degradation (Wang et al., 2010).

While Caf1/Pop2 and Ccr4 subunits are found in yeast, humans have CNOT7/CAF1 and CNOT8/CAF1b as Caf1/Pop2 orthologues and CNOT6/CCR4a and CNOT6L/CCR4b as Ccr4 orthologues. Though all four deadenylases are associated with the CCR4-NOT complex in humans, only one DEDD and one EEP are present in each subunit (Bartlam & Yamamoto, 2010).
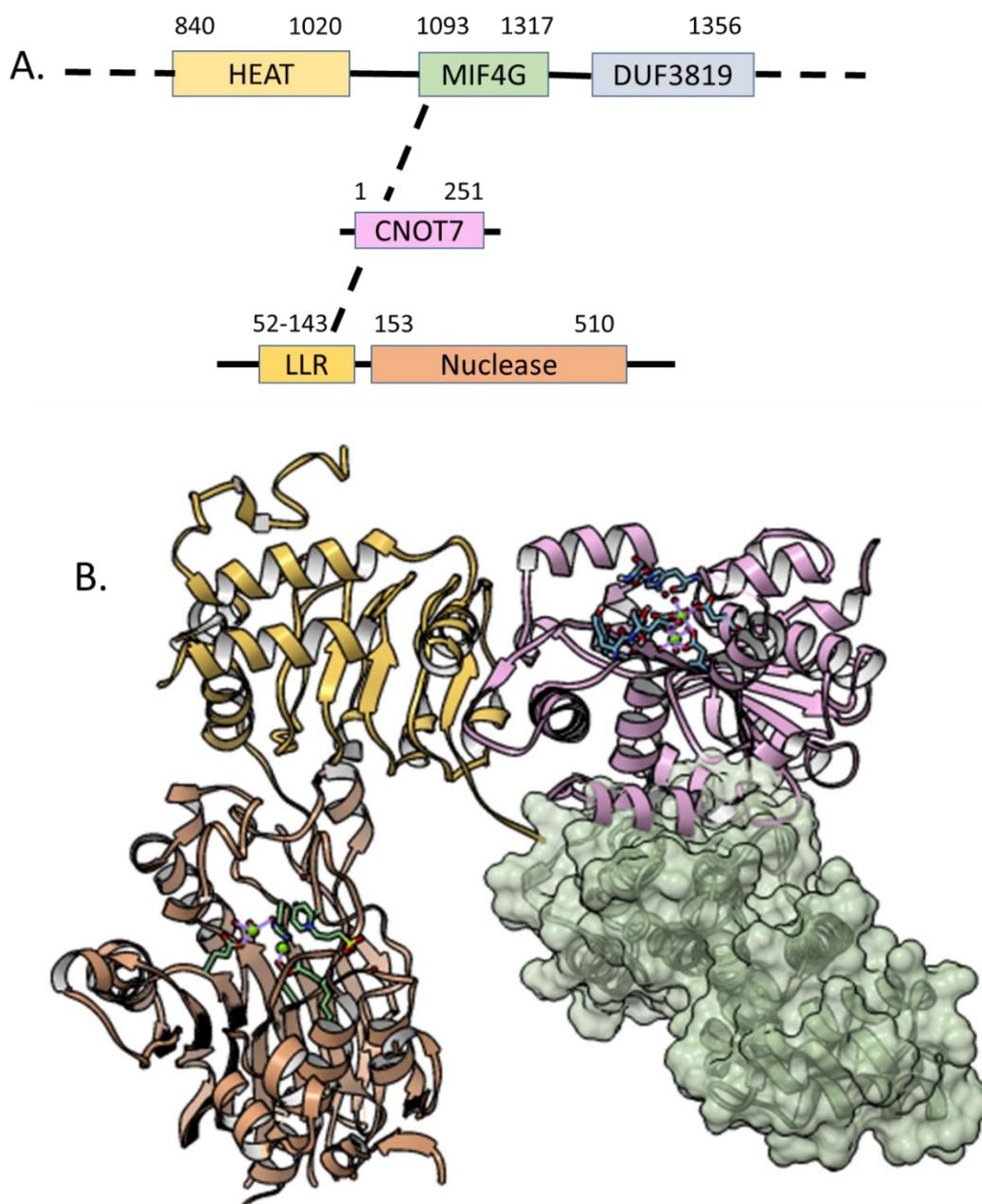
Figure 1-10. A schematic map and crystal structure representing the interaction of catalytic and non-catalytic modules of the CCR-4NOT complex. The structure of CNOT1 (MIF4G), CNOT6/CCR4a, and CNOT7/CAF1. **A)** A map representation of CNOT7/CAF1 interacting with MIF4G of CNOT1 and with CNOT6/CCR4a. **B)** Crystal structures representing the MIF4G domain (PDB 4GMJ) of CNOT1 (Petit et al., 2012) is the green structure. The purple structure is the interaction of CNOT7/CAF1 (4GMJ) with the MIF4G domain of CNOT1 (Petit et al., 2012). The green balls in the structure of CNOT7/CAF1 are Mg^{+2} ions conserved in the catalytic region around active residues. The structure in gold is the leucine-rich repeats of CNOT6/CCR4a (7AX1) (Wang et al., 2010), which regulate the interaction of CNOT7/CAF1 (7AX1) (Chen et al., 2021). The brown structure coloured is the nuclease module of CNOT6/CCR4a (3NGO) with Mg^{+2} ions conserved in the catalytic region around active residues (Wang et al., 2010).

1.2.4.1. The structure of CNOT7/CAF1

CNOT7/CAF1 is a catalytic subunit of the CCR4-NOT complex, which regulate poly(A) tail removal from the 3′ end of the mature mRNA. The protein is kidney-shaped with 13α-helices and six β-strands. Structural analysis for hCAF1 revealed similarities with Caf1/Pop2 of *S. Cerevisiae* and *S. Pombe*, with Mg^{2+} metal ions conserved in the active site around DEDD residues (Andersen et al., 2009, Thore et al., 2003, Horiuchi et al., 2009b). The structure of CNOT7/CAF1 was studied using Pop2 in *Saccharomyces cerevisiae* (PDB 1UOC; Figure 1-11A; Thore et al., 2003) and later in *Schizosaccharomyces pombe* (PDB 3G10; Figure 1-11B; Andersen et al., 2009).

The term DEDD is derived from the coordination of divalent metal ions with Asp and Glu residues (Asp40, Glu42, Asp161, and Asp230) in active sites (Figure 1-11C; Petit et al., 2012). The coordination of Mg^{2+} metal ions with DEDD residues regulate the catalytic activity of CNOT7/CAF1 (Petit et al., 2012, Wahle & Winkler, 2013). The Pop2 protein of *S. Pombe* comprises Mg^{2+} and Zn^{2+} ions in the A site and Mn^{2+} ions in the B site; although Mg^{2+} is still present in high concentration, Zn^{2+} and Mn^{2+} ions regulate catalytic activity by co-ordinating with DEDD residues (Andersen et al., 2009).

A study by Petit et al. (2012) revealed a CNOT7/CAF1 interaction site with the MIF4G domain of CNOT1. The study discloses a structure of CNOT7/CAF1-MIF4G, revealing that CNOT7/CAF1 interaction with MIF4G occurs away from the active site, shedding light on the coordination of Asp and Glu residues with metal ions in the conserved region (Figure 1-11D).

1.2.4.2. The structure of CNOT6/CCR4a and CNOT6L/CCR4b

CNOT6/CCR4a and CNOT6L/CCR4b are the second catalytic subunits of the CCR4-NOT complex. Yeast Ccr4p comprises three major functional domains: the main domain, the activation domain, and the deadenylase domain. The primary domain consists of tandem LRRs that connect CCR4 to the CCR4-NOT complex. The activation domain in the N-terminal region is rich in glutamine/asparagine residues and contains transcriptional regulators that interact with the transcriptional machinery (Winkler & Balacco, 2013).

The C-terminal region of the deadenylase domain contains Asp and His residues that are conserved in the catalytic region (Wang et al., 2010, Dupressoir et al., 2001). Structural

analysis revealed that all five residues are close to each other and clustered around the metal ions within the negatively charged pockets (Wang et al., 2010). The EEP superfamily of nucleases includes *Sphingomyelin phosphodiesterase* (PDB 2DDS; Figure 1-12B; Ago et al., 2006), the L1-EN reverse transcriptase domain (PDB 1VYB; Figure 1-12C; Weichenrieder et al., 2004), DNA apurinic or apyrimidinic site lyase (PDB 2ISI; Figure 1-12D), human APE1 (PDB 1DEW; Figure 1-12E Mol et al., 2000), and *E. coli*. Exo III (PDB 1AKO; Figure 1-12F; Mol et al., 1995). The similarity of human CNOT6/CCR4a (PDB 3NGO; Figure 1-12A) with other EEP families of nucleases means that CCR4 is categorised under the endonuclease exonucleases family of deadenylases.

A. *S. Cerevisiae*

B. *S. Pombe*

C. CNOT7/CAF1a

D. CNOT7/CAF1a − MIF4G
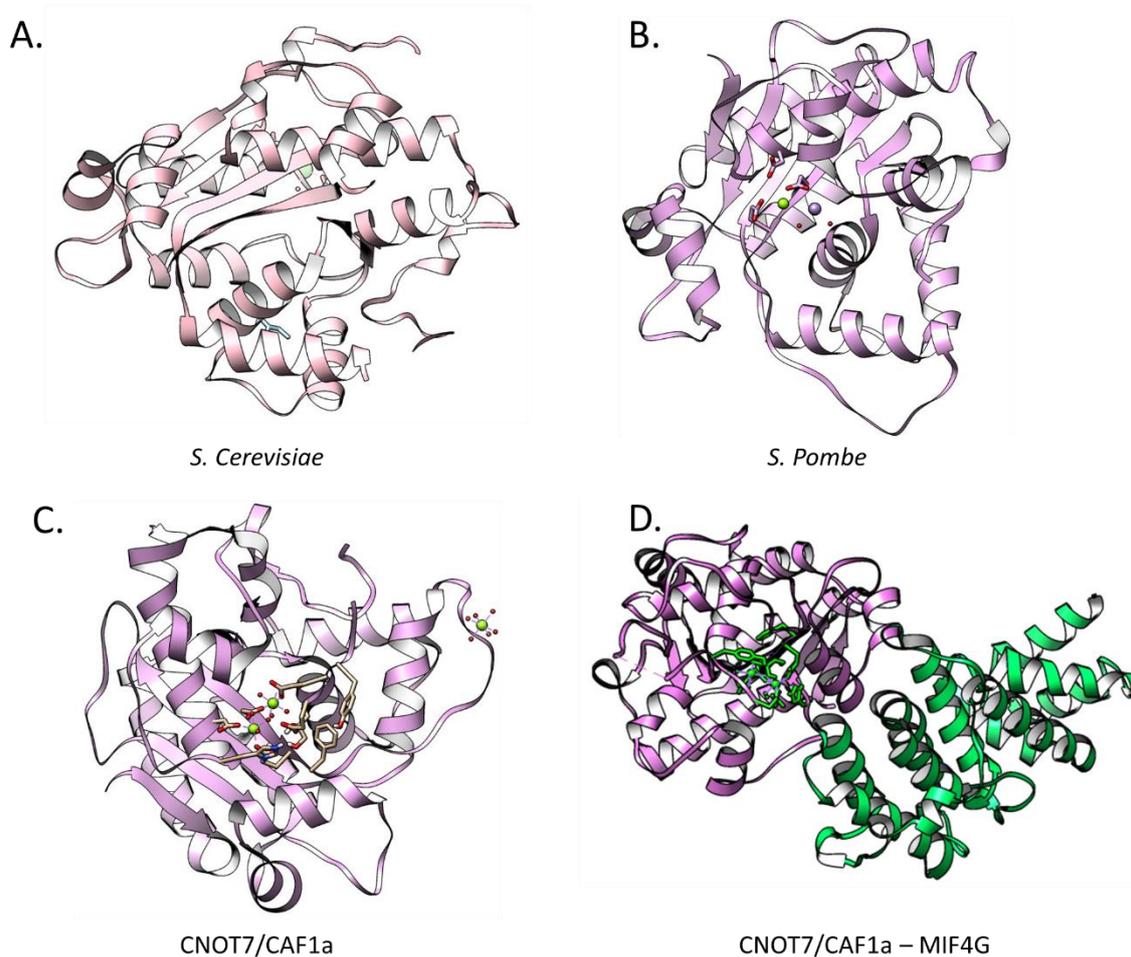
Figure 1-11.Crystal structure of CAF1 and its orthologues. Comparison of hCAF1 with its orthologues reveals the coordination of metal ions with active residue sites. **A)** Crystal structure of Pop2 (1UOC) with Ca^{+2} ions in the conserved region (Thore et al., 2003). **B)** Crystal structure of Pop2p (PDB 3G10) with Mn^{+2} and Mg^{+2} ions in coordination with residues in the active site of the catalytic region (Andersen et al., 2009). **C)** Crystal structure of CAF1 (PDB 4GMJ) with Mg^{+2} metal ions in coordination with residues in the active site (Petit et al., 2012). **D)** Crystal structure of CNOT7/CAF1 interacting with the MIF4G domain of CNOT1 (PDB 4GMJ) reveals MIF4G binding to CAF1 occurs away from the active site of CNOT7/CAF1 (Petit et al., 2012).
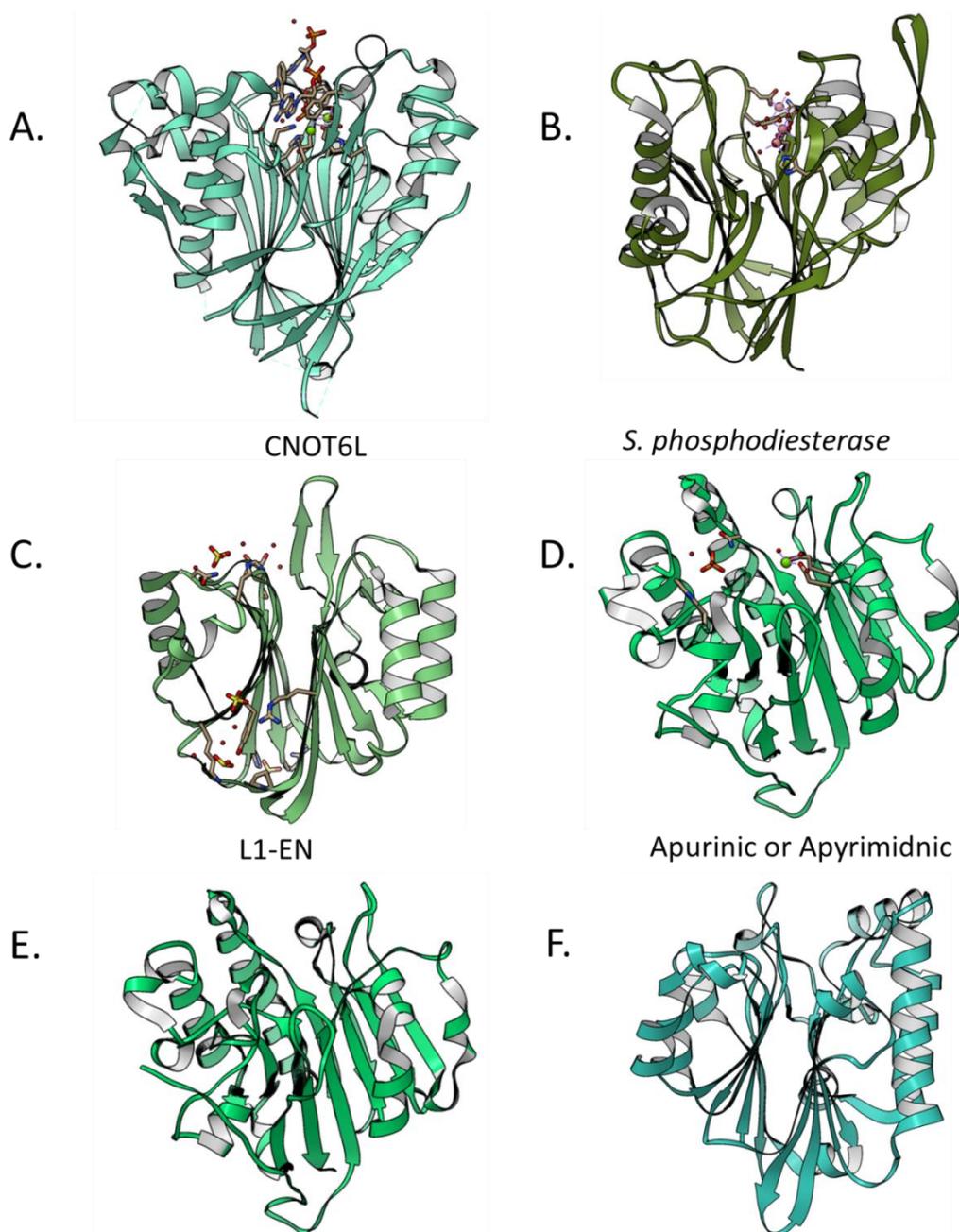
Figure 1-12. The structure of CNOT6/CCR4a and its orthologues. Structure of CNOT6/CCR4a with its orthologues revealing coordination between metal ions and residues in active sites. **A)** Crystal structure of CNOT6L/CCR4b (PDB 3NGO) with Mg^{2+} metal ions conserved in the catalytic region that co-ordinate with conserved residues (Wang et al., 2010). **B)** The crystal structure of *S. Phosphodiesterase* (PDB 2DDS) with Ca^{+2} ions conserved in the catalytic region in coordination with active residues (Ago et al., 2006). **C)** Crystal structure of L1-EN 1 (PDB 1VYB) with sulfate ions around its reverse transcriptase domain (Weichenrieder et al., 2004). **D)** Apurinic or apyramidic enzyme (PDB 2ISI) structure with Mg+2 ions conserved in the catalytic region in coordination with residues in active sites. **E)** Crystal structure of APE I (PDB 1DEW, Mol et al., 2000). **F)** *E. coli* Exo III (PDB 1AKO, Mol et al., 1995).

1.2.5. The NOT module of the CCR4-NOT complex

The NOT module of the CCR4-NOT complex in humans comprises CNOT1, CNOT2, CNOT3, and CNOT4, while the yeast comprises Not1, Not2, Not3 or Not5, and Not4p (Figure 1-13A). In the yeast complex, Not2 is an orthologue of CNOT2, whereas Not3 and Not5 are orthologues of CNOT3 (Boland et al., 2013). CNOT1 (PDB 4C0E) is composed of HEAT repeats in the C-terminal region, and it is defined by helix A-turn-helix B-motif sequences (Figure 1-13B). The 13 HEAT repeats in the C-terminal of CNOT1 are further differentiated as HEAT motifs 1–4, 5–9, and 10–13 (Boland et al., 2013).

CNOT1 functions as a scaffold protein to assemble catalytic and non-catalytic subunits (Boland et al., 2013, Bhaskar et al., 2013). The *Drosophila melanogaster* model was used to study the interactions of the individual subunits of the CCR4-NOT complex, revealing the three regions of CNOT1 and its role as a scaffold protein in the assembly of catalytic and non-catalytic subunits (Bawankar et al., 2013). The C-terminal domain of CNOT2 (PDB 4BY6; Figure 1-13C) and CNOT3 in humans and Not5 in yeast (PDB 4BY6; Figure 1-13D) consist of NOT boxes with α-helices in the N-terminal and β-sheets towards the C-terminal, formed by four or five β-strands (Kridel et al., Bhaskar et al., 2013). CNOT2 and CNOT3 of the CCR4-NOT complex contain identical NOT boxes localised in their respective C-terminals (Boland et al., 2013).

The C-terminal region of CNOT1 harbours CNOT2; there is no direct interaction of CNOT3 with CNOT1. CNOT3 interacts with CNOT1 by homo-dimerising with CNOT2 through NOT boxes, the N-terminus of CNOT2 and CNOT3 mediates dimerisation (Figure 1-13E; Xu et al., 2014, Wang et al., 2010, Petit et al., 2012, Boland et al., 2013). CNOT2 and CNOT3 have a close dimerisation interface. Therefore, the interaction of these subunits results in the formation of a heterodimer (Boland et al., 2013).

An unstructured extended peptide region termed the asymmetrical lobe anchors CNOT2 and CNOT3 to CNOT1 (Bhaskar et al., 2013). CNOT3 interacts via hydrophobic polar residues. It anchors to HEAT motifs 1–5 and extends from HEAT motifs 9–10 (Figure 1-13E). CNOT2 staggers towards HEAT motifs 4–6 and harbours on the surface of CNOT1 (Boland et al., 2013, Xu et al., 2014). The CNOT4 subunit in yeast and humans share a similar N-terminus but a divergent C-terminus. The structure comprises three long loops, L1, L2, and L3, with an α-helix between the second and third loops. CNOT4 subunit interacts with CNOT1/Not1 via the

C-terminus (Hanzawa et al., 2001). CNOT1 play multiple roles in gene regulation, such as in the repression of ligand-dependent transcription by oestrogen receptor (ER) α, which it does by binding to the ligand-binding domain in ER α (Winkler et al., 2006). Similarly, CNOT1 represses translation by binding to DEAD-box helicases eIFA2 and DDX. DDX and eIFA2 compete to bind to the MIF domain of CNOT1 to regulate translation repression (Meijer et al., 2019).

Like CNOT1, CNOT2 plays a crucial role in gene regulation. CNOT2 represses transcription by binding to the silencing mediator of the retinoid and thyroid hormone receptor (SMRT)/nuclear hormone receptor co-repressor (NCOR) HDAC3 complex. SMRT/NCOR are co-repressor proteins that are part of HADC3. CNOT2 represses transcription by interacting with HDAC3 through a NOT box (Jayne et al., 2006).

1.2.6. The DUF3819 module of the CCR4-NOT complex

CNOT9/CAF40 is a non-catalytic subunit of the CCR4-NOT complex. It is known as the RQCD1 domain (Figure 1-14). CNOT9/CAF40, along with CNOT1, regulates the CCR4-NOT complex (Pavanello et al., 2018). CNOT9/CAF40 comprises six armadillo repeats; each ARM repeat has three helices (H1, H2, and H3). Together, the three helices form a crescent-shaped molecule with a convex surface lined by the H1 helices and a concave surface lined by helices H2 and H3 (Mathys et al., 2014). The convex surface in ARM2 and ARM3 allows CNOT9/CAF40 to form homodimers and facilitates binding to the domain of unknown function 3819 (DUF3819), which is in the middle region of CNOT1 (Figure 1-14). The binding of CNOT9/CAF40 to DUF3819 of CNOT1 causes structural rearrangement and exposes tryptophan pockets, which, in turn, mediate interaction with GW182 and regulate miRNA-mediated decay (Mathys et al., 2014).

1.2.7. The N-terminal module of the CCR4-NOT complex

CNOT10 and CNOT11/C2orf29 are subunits of the CCR4-NOT complex; however, their functions are unclear (Bawankar et al., 2013, Boland et al., 2013). The CNOT11 subunit is identified in eukaryotes as C2orf29 and is conserved in *D. Melanogaster* as CG13467. The N-terminal of CNOT1 is responsible for mediating the interaction with CNOT11 and CNOT10. CNOT10 does not directly bind to the N-terminal. However, it forms a complex with CNOT11, which, in turn, interacts with the N-terminal of CNOT1 (Mauxion et al., 2013).
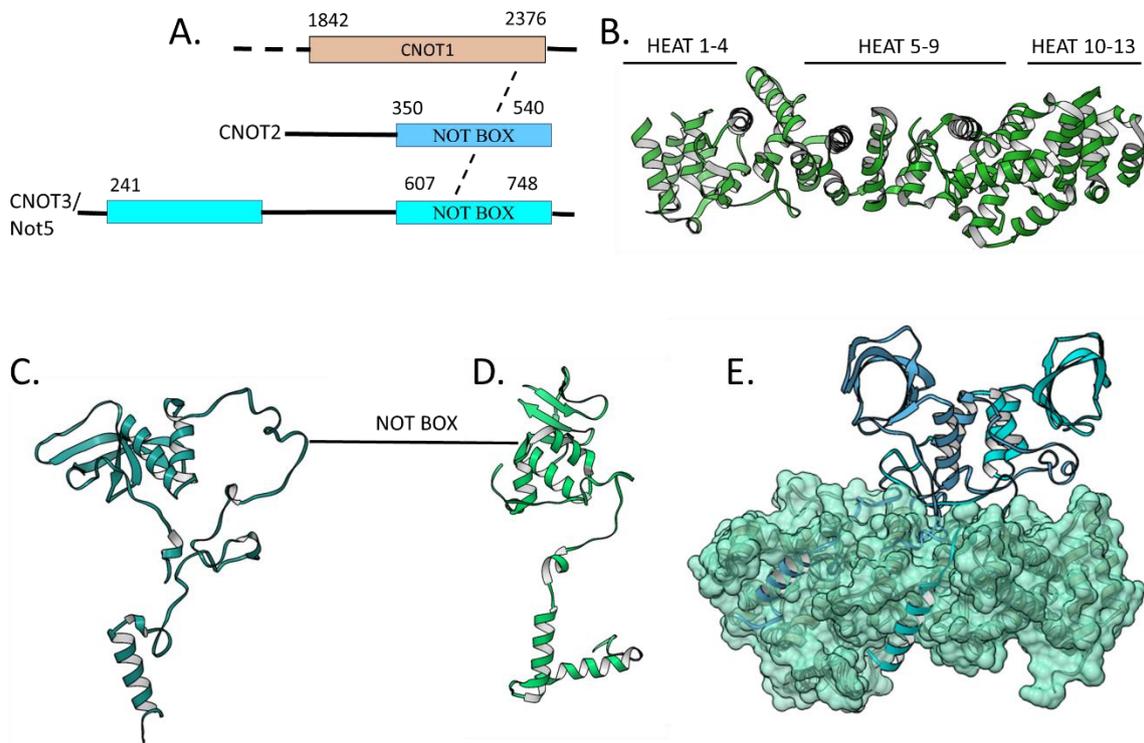
Figure 1-13. Association of CNOT2 and CNOT3 to CNOT1. MIF4G domain functions as a protein scaffold to interact with NOT modules. **A)** A schematic representation of CNOT2 and CNOT3 interacting with CNOT1 through NOT boxes. **B)** The structure of the C-terminal region of CNOT1 (green), depicting helix A-turn-helix B-motifs with thirteen heat repeats (PDB 4C0E, Boland et al., 2013). **C)** The structure in light blue is CNOT2 (PDB 4BY6, Bhaskar et al., 2013) **D)** The structure in green is CNOT3 (PDB 4BY6). NOT boxes are visible in the C-terminal of CNOT2 and CNOT3 (Bhaskar et al., 2013). **E)** The aqua structure is CNOT1 (Bhaskar et al., 2013) with helix A-turn-helix B-motifs with CNOT2 (light blue) (Boland et al., 2013) and CNOT3 (Cyan) anchored to the HEAT repeat motifs of CNOT1 (Boland et al., 2013).
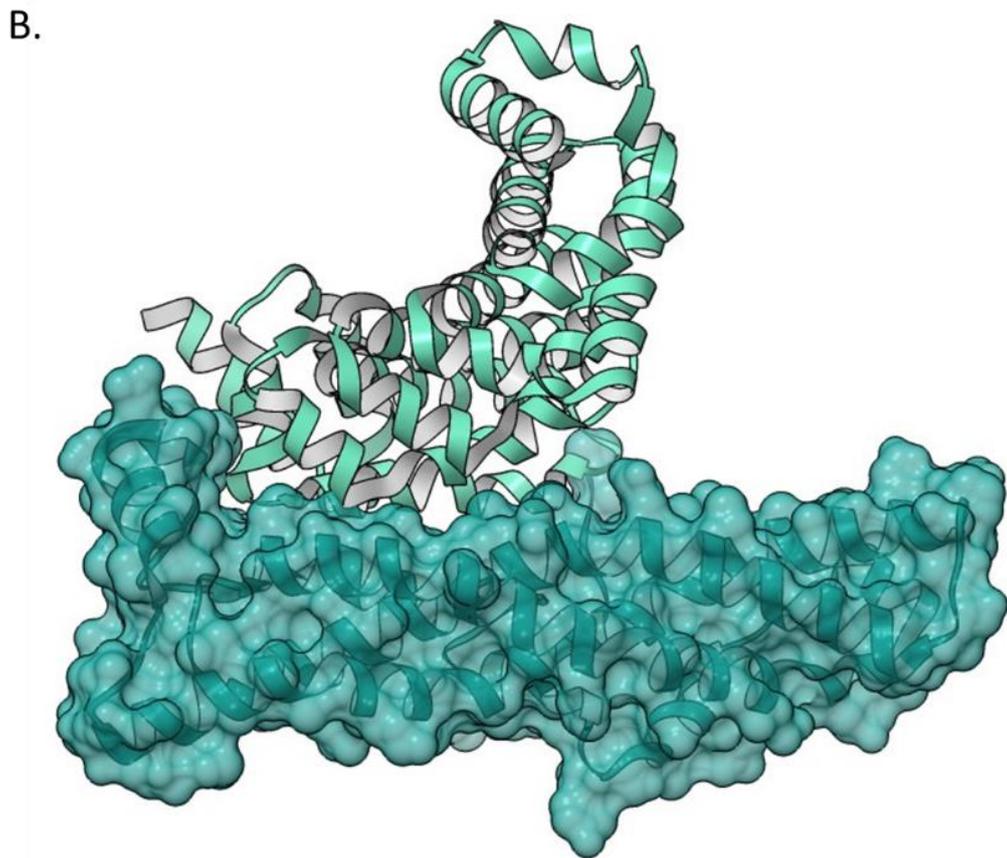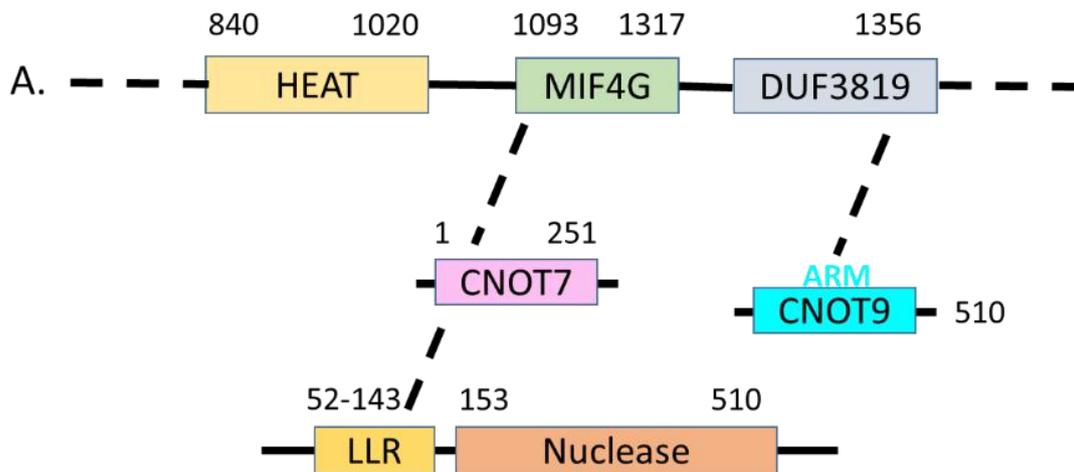
Figure 1-14. Interaction of CNOT9 with the DUF3819 region in the CNOT1 subunit. A) Schematic representation of the DUF3819 region of CNOT1 associating with CNOT9, and MIF4G with CNOT7/CAF1, which binds to the LLR region of CNOT6/CCR4a B) The aqua surfaced structure is the DUF3819 region of CNOT1, the cyan ribbon bound to the DUF3819 region in CNOT1 is CNOT9.

1.2.8. Deadenylation-dependent mRNA degradation

Upon completion of poly(A) tail removal, LSM (1–7), in a complex with PAT1 (LSM-PAT1), binds to the 3′ end of the mRNA to regulate 5′→3′ degradation (Figure 1–15; Bouveret et al., 2000, Schoenberg & Maquat, 2012). The LSM-PAT1 complex binds to the 3′ end and subsequently recruit decapping proteins 1 and 2 (DCP1-2). The DCP1-DCP2 recruited is activated by connecting to the 5′ end. The DCP2 includes a Nudix motif for catalytic activity and is, therefore, a member of the Nudix hydrolase superfamily (Wurm & Sprangers, 2019).

The catalytic site in the Nudix motif of DCP2 comprises glutamate residues, which are known to co-ordinate with divalent cations, which is essential for cap hydrolysis (Wurm & Sprangers, 2019, Nagarajan et al., 2013). DCP1-DCP2, along with DCP enhancers (eDCP), regulates the removal of the 5′ cap structure from mRNA. After decapping, exoribonuclease (XRN 1) degrades mRNA from 5′→3′ (Figure 1-15; Garneau et al., 2007, Parker & Song, 2004).

Exosomes regulate mRNA degradation at the 3′end. Exosomes consist of multi-subunit complexes in their central core, and the subunits consist of six 3′→5′ exoribonucleases that are catalytically inactive and arranged in a ring. There are three extensions known as S1, KH, and RNA binding domains (Schmid & Jensen, 2008). The degradation of 3′→5′ mRNA by exosome is similar to 5′→3′ degradation, which involves the hydrolysis of the 5′ cap by scavenger DCPs followed by degradation by exosomes.

1.2.8.1. mRNA degradation independent of deadenylation

The degradation of mRNA independent of deadenylation is a process that does not require poly(A) tail removal by poly(A) nucleases. An example of deadenylation independent mRNA degradation is Ribosomal protein S28B (RPS28B), which regulates deadenylation-independent mRNA degradation by surpassing the deadenylation step. RPS28B/Rps28B in *S. cerevisiae* adapts an autoregulatory mechanism in recruiting Edc3 enzymes by binding directly to the stem-loop structure in the 3′ untranslated region. The Edc3 recruited by Rps28B regulates decapping via Dcp1 and Dcp2. The decapping is followed by mRNA degradation of transcripts from 5′→3′ by XRN1 (Figure 1-16; Badis et al., 2004, Garneau et al., 2007). An alternative example for deadenylation independent degradation was observed in yeast; when an aberrant yeast transcript lacks its stop codon, it becomes a non-stop mRNA and is degraded by exosomes at 3′→5′ (Frischmeyer et al., 2002, van Hoof et al., 2002). Additionally, non-sense

codons in the yeast trigger Dcp1-Dcp2 mediated decapping followed by 5′→3′ Xrn1 digestion of the mRNA before the poly(A) tail is removed (Muhlrad & Parker, 1994).

1.2.8.2. mRNA degradation by endonucleolytic cleavage

Degradation via endonucleolytic cleavage involves the internal cleavage of mRNA by endonucleases. The enzymes which cleave the mRNA generate two unprotected fragments. The unprotected fragments are susceptible to degradation by XRN1 at 5′→3′ and by exosomes at 3′→5′ (Figure 1-17). An example of a cellular endonuclease known for targeting mRNA is polysome ribonuclease 1 (PMR1). PMR1 is activated by oestrogen stimulation in Xenopus hepatocytes to degrade mRNA coding for serum protein (Garneau et al., 2007). The polysome targeting domains in PMR1 regulate endonucleolytic cleavage and target translated mRNA for degradation (Yang & Schoenberg, 2004). An alternative example of mRNA degradation by endonucleolytic cleavage is through short interfering RNA (siRNA), which incorporates the ribosome induced silencing complex (RISC) and argonaute protein 2 (AGO2). A component of RISC regulates mRNA decay by endonucleolytic cleavage (Houseley & Tollervey, 2009).
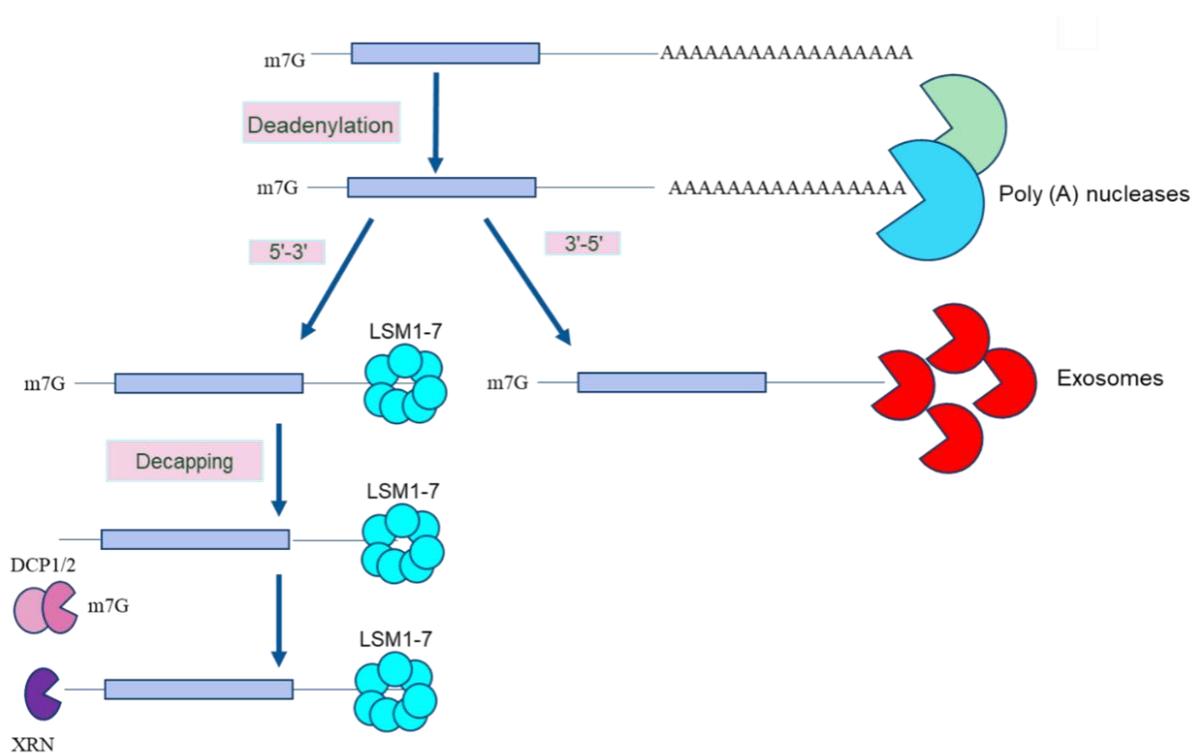
Figure 1-15 Decapping and mRNA degradation. Poly(A) nucleases degrade the addition of the poly(A) tail to the 3′ end of mRNA. LSM 1-7 complex binds to the 5′ end of the mRNA to initiate decapping by DCP1-DCP2. After decapping the methylated cap structure, mRNA is degraded from 5′-3′ by XRN1. 3′-5′ degradation is similar to 5′-3′ in which the decapping is done by scavengers DCPs and exosomes degrade the mRNA.
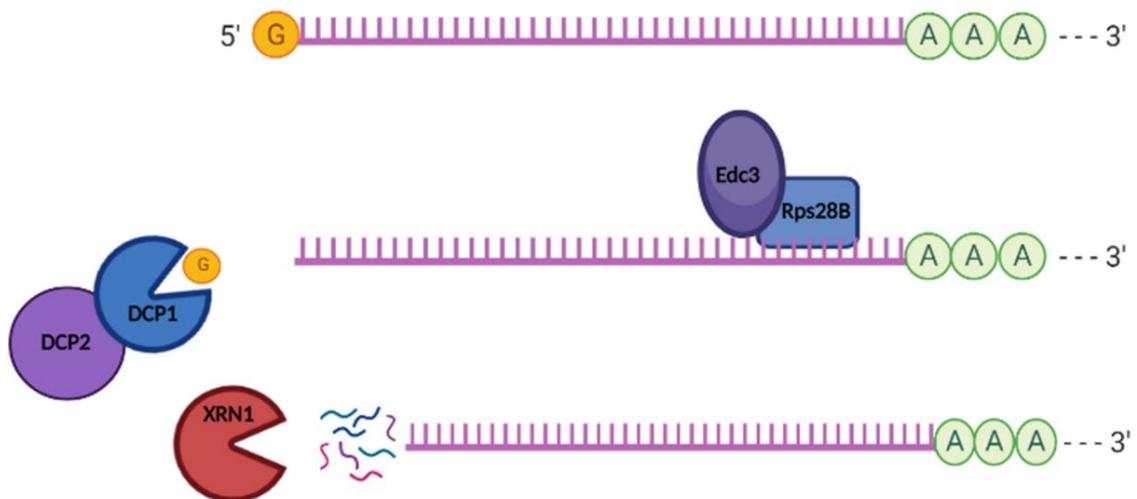
Figure 1-16. Schematic representation of deadenylation-independent mRNA degradation. Rps28B recruits Edc3 by binding to the stem-loop structure in the 3′ untranslated region, which regulates decapping by Dcp1 and Dcp2, followed by mRNA degradation by XRN1.
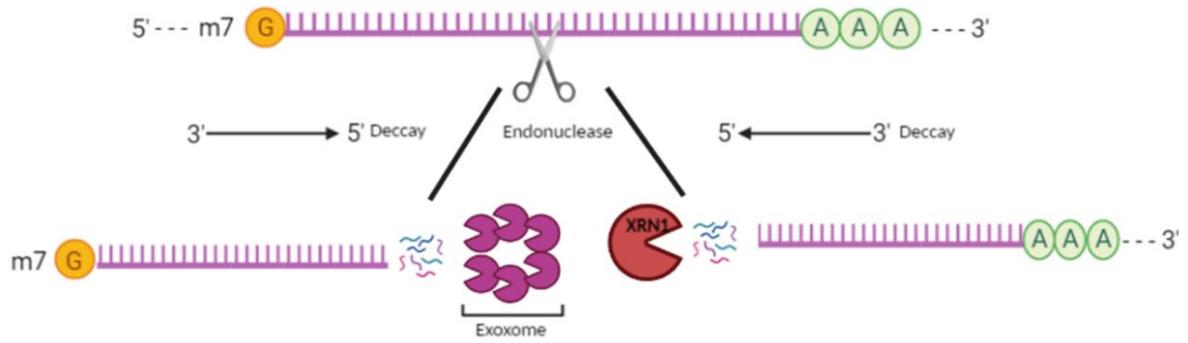
Figure 1-17. mRNA degradation by endonucleolytic cleavage. The middle region of the processed mRNA is cleaved by endonucleases in the cytoplasm. The endonucleases digest the mRNA, cleaving it into two unprotected fragments susceptible to degradation by exosomes and XRN1.

1.2.9. miRNA-mediated degradation

Studies examining miRNA have found it to play various roles, such as in cell differentiation, stem cells, tumour genesis, apoptosis, metabolism longevity, and viral infection (Carthew & Sontheimer, 2009). Moreover, miRNA is endogenous and made up of 22 nucleotides (Carthew & Sontheimer, 2009), which incorporate into RISC after being transcribed and processed (Gavrilov & Saltzman, 2012).

RISC is a multiprotein complex that regulates gene silencing by incorporating miRNA or double-stranded small interfering RNA (siRNA). The dsRNA is a critical tool in gene regulation as it acts as a template for RISC. The dsRNA incorporated into RISC retains one strand and expels the other to recognise complementary mRNA transcripts. AGO proteins in RISC cleave the mRNA via a process called RNA interference (RNAi) and, consequently, silence the mRNA (Gavrilov & Saltzman, 2012). There are four AGO proteins in humans (AGO1-4) that are the primary and essential components of mRNA silencing, and miRNA incorporated into the RISC complex mediates AGO proteins to silence the gene of interest (Song et al., 2004).

The AGO component in RISC interacts with GW182 through the N-terminus (Ding & Han, 2007). There are three paralogues in mammals and humans, termed the trinucleotide repeat-containing (TNRC)-6A, 6B, and 6C (Chen et al., 2009). The first part of the term, GW, refers to several glycine-tryptophan repeats, while the designation 182 refers to the protein's molecular mass (Ding & Han, 2007). Two additional domains within GW182 are an associated ubiquitin domain and non-canonical RRMs. The GW182 N-terminus binds to AGO, and the silencing domain is in the C-terminal (Fabian et al., 2011). GW182 interacts with members of the AGO subfamily and regulates miRNA-mediated mRNA silencing in mammalian cells (Chen et al., 2009). Once the miRNA targeting transcript is bound to AGO1 in the RISC complex, the recruitment of GW182 occurs. The silencing domain in the C-terminal interacts with CNOT1 by two distinct CCR4-NOT interacting motifs (CIFs), which serve as the link between miRNA regulation and mRNA degradation (Fabian et al., 2011). GW182 bound to CNOT1 regulates mRNA degradation by activating the CCR4-NOT complex.

1.2.10. Protein-mediated recruitment for mRNA degradation

An alternative recruitment mechanism for the CCR4-NOT complex to mRNA is via RNA binding proteins. RNA binding proteins, such as tristetraprolin (TTP) and TTP-related butyrate response factors proteins (BRF) 1 and 2, bind to adenylate-uridylate-rich elements (AREs), which are present in most mRNA (Fabian et al., 2013). TTP are vital for controlling the inflammatory response; therefore, they are considered tumour suppressors in lymphomas (Carballo et al., 1998, Rounbehler et al., 2012). TTP, along with BRF1 and BRF2, recruits the CCR4-NOT complex to the ARE in the 3′ untranslated region (UTR), promoting deadenylation, thus resulting in the degradation of the target mRNA (Blackshear, 2002, Lai et al., 2000).

Nanos protein regulates mRNA degradation by interacting with the CCR4-NOT complex. The Nanos protein consists of a conserved CCHC zinc finger domain and divergent N and C-terminals that are unstructured with variable lengths and low sequence complexity. The zinc finger domain mediates RNA binding (Curtis et al., 1997, Arrizabalaga & Lehmann, 1999). The regions in the N-terminus of Murine Nanos2 interact with CNOT1, which recruits the CCR4-NOT complex to trigger the degradation of specific RNAs.

The degradation of mRNA can also be achieved by roquin 1 and 2 through their conserved ROQ domains. The ROQ domain recognises a stem-loop structure in the 3′ UTR of mRNA (Schlundt et al., 2014). The highly conserved N-terminal region of roquin comprises a ring finger with a potential E3 ubiquitin ligase function. A conserved ROQ domain, alongside a CCCH zinc finger, is involved in RNA recognition (Schlundt et al., 2014). Roquin recruits the CCR4-NOT complex with its non-conserved C-terminal region. The recruitment of the CCR4-NOT complex in *D. Melanogaster* explains that the recognition of the conserved motif in CNOT9/CAF40 by roquin helps to regulate mRNA degradation (Sgromo et al., 2017)

1.3. Anti-proliferative proteins encoded by B-Cell translocation gene

The BTG family of proteins BTG1, BTG2/PC3/Tis21, BTG3/ANA, BTG4/PC3B, TOB1, and TOB2 are encoded by the B-cell translocation gene (BTG). BTG2 proteins are homologues of the pheochromocytoma cell 3 (PC3) protein in rats, induced by nerve growth factor (NGF). PC3 mRNA is expressed, with high specificity, in the placenta, heart, spleen, and skeletal muscle to a reduced extent (Bradbury et al., 1991). An alternative homologue of the BTG proteins is tetradecanoyl phorbol acetate- inducible sequence 21 (Tis21). Tetradecanoyl phorbol acetate (TPA) induces Tis21 in the Swiss 3T3 cells of mice (Fletcher et al., 1991, Matsuda et al., 2001).

The BTG family of proteins and their homologues are antiproliferative proteins that regulate cell cycle progression in various cell types. The first evidence of BTG1's role in regulating cell growth was observed in quiescence cells. Overexpression of BTG1 mRNA in the cell cycle during the G0/G1 phase resulted in a drastic reduction in cell growth during the G1/S phase transition (Rouault et al., 1996, Rouault et al., 1992). Similarly, BTG2 mRNA levels are high in quiescent cells, and overexpression of the BTG2 gene results in suppressing growth (Montagnoli et al., 1996, Lim et al., 1998).

BTG1 and BTG2 function as effectors of signalling pathways by regulating cellular differentiation and apoptosis. Since BTG1 and BTG2 are negative regulators of the cell cycle, overexpression of BTG1 and BTG2 leads to apoptosis in various cell types, such as human breast cancer cells, microglia, and murine fibroblasts. In fibroblasts, cellular senescence is caused by expressing BTG2, which antagonises the Pin1 cell cycle regulator (Wheaton et al., 2010). The expression of BTG2 in breast tissue induces apoptosis by repressing the anti-apoptotic protein BCL2 (Nahta et al., 2006).

Downregulation of BTG2 in breast carcinoma was found to result in the increased expression of cyclin D1 and a high level of AKT phosphorylation, resulting in a decreased rate of survival, increased disease progression, and a rise in tumour grade (Kawakubo et al., 2006, Takahashi et al., 2011). A low level of BTG2 in liver cancer also results in an elevated level of cyclinD1/cyclinE, leading to an increase in tumour grade (Zhang et al., 2011).

1.3.1. The role of conserved BTG or APRO domain in BTG/TOB

The BTG/TOB protein structure comprises five α-helices and four β-strands that form two antiparallel β-sheets (Yang et al., 2008). The N-terminal of BTG1/BTG2 and TOB1/TOB2 have similar BTG domains, which are known as the antiproliferative (APRO) domain with Box A (Y50–N71) region and Box B (L97– E115) region (Figure 1-18). These are identical in TIS21 and PC3 (Rouault et al., 1992). Sequence comparison of the BTG/TOB family of proteins reveals that BTG1 and BTG2 are highly similar, and 40% of BTG1 and BTG2 are identical to TOB. However, BTG3 and BTG4 are only remotely related to other family members (Figure 1-19; Winkler, 2010).

BTG domains function as growth regulators by binding to several molecular targets. BTG1/BTG2 regulate transcription in Hoxb9 by forming a ternary structure with the homeodomain binding site II of the neural cell adhesion molecule (Prevot et al., 2000, Yang et al., 2008). In comparison, BTG3 regulates cell cycle progression by interacting with E2F1, a transcription factor essential for S phase entry. The BTG3 residues involved in E2F1 interactions have been traced to the N-terminus and Box A region of BTG3 (Winkler, 2010, Ou et al., 2007).

1.3.2. The role of the C-terminal region of BTG/TOB

BTG1 and BTG2 possess an additional Box C (D116-A126) region in the C-terminus. The Box C region of BTG1/BTG2 interacts with protein arginine methyltransferase (PRMT1; Yang et al., 2008). BTG2-PRMT1 interaction occurs during gene regulation by the retinoic acid receptor (RAR) α. BTG2-PRMT1 proteins in RAR complexes are recruited into RARβ promoters without retinoic acid to induce demethylation at the arginine three residues of histone H4. By contrast, histone H4 acetylation level increases due to displacement of BTG2 and PRMT1 from the RARβ promoter upon the addition of retinoic acid. BTG2 is well known to associate with Smad I and Smad 8, and overexpression of BTG2 results in an increase in the BMP-dependent synthetic reporter (Park et al., 2004).

Though TOB proteins have similar BTG domains, they are divergent in their C-terminus. TOB proteins have an extra-long C-terminal that contains a PAM2 motif. The C-terminus of TOB1 can bind with Smad transcription factors and influence their DNA binding activity. TOB1 increases the DNA binding ability of the Smad 4 transcription factor to its consensus DNA sequence (Winkler, 2010, Yoshida et al., 2000).

TOB1 is also known to interact with kinases, including transmembrane receptor-type tyrosine kinase ErbB2, p90/Rsk1, Jnk2, and Erk1/Erk2. The residues, such as serine and threonine, in the C-terminal region of TOB1 are subject to phosphorylation (Suzuki et al., 2002). NIH3T3 cells stimulated with mitogenic growth factors result in Ras/MAP kinase signalling and the phosphorylation of TOB1. The MAP kinase signalling component ErkI/Erk2 phosphorylates TOB1 on residues Ser-147 and Ser-149 in the C-terminal region. Interestingly, phosphorylation of TOB1 results in its inactivation, which is essential for Ras-mediated cell transformation (Suzuki et al., 2002).
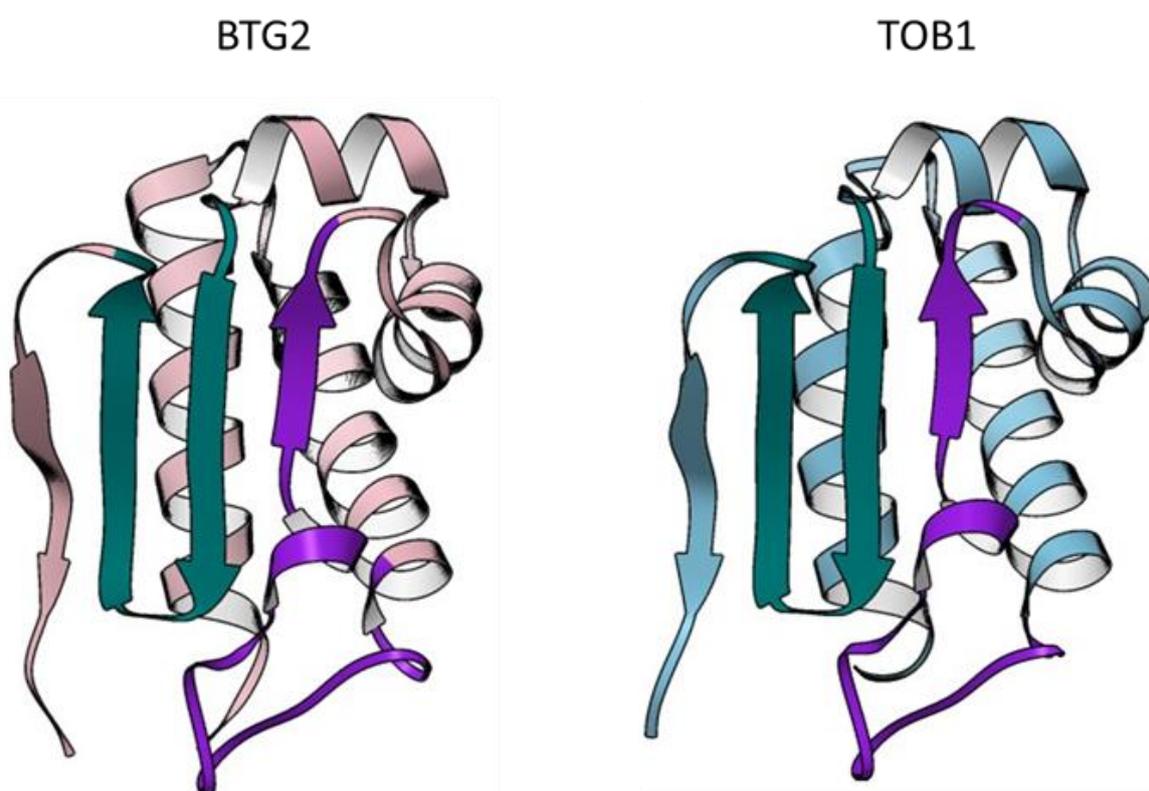
Figure 1-18. Structure of BTG/TOB. BTG/TOB proteins have an identical Box A and Box B, known as the BTG domain. The crystal structure of BTG2 (PDB 3DJU) is pink (Yang et al., 2008). The crystal structure of TOB1 (2D5R) is in blue (Horiuchi et al., 2009b). The purple colour represents the Box A region in the β-sheet of BTG/TOB, and the Box B region is highlighted in the green of proteins.
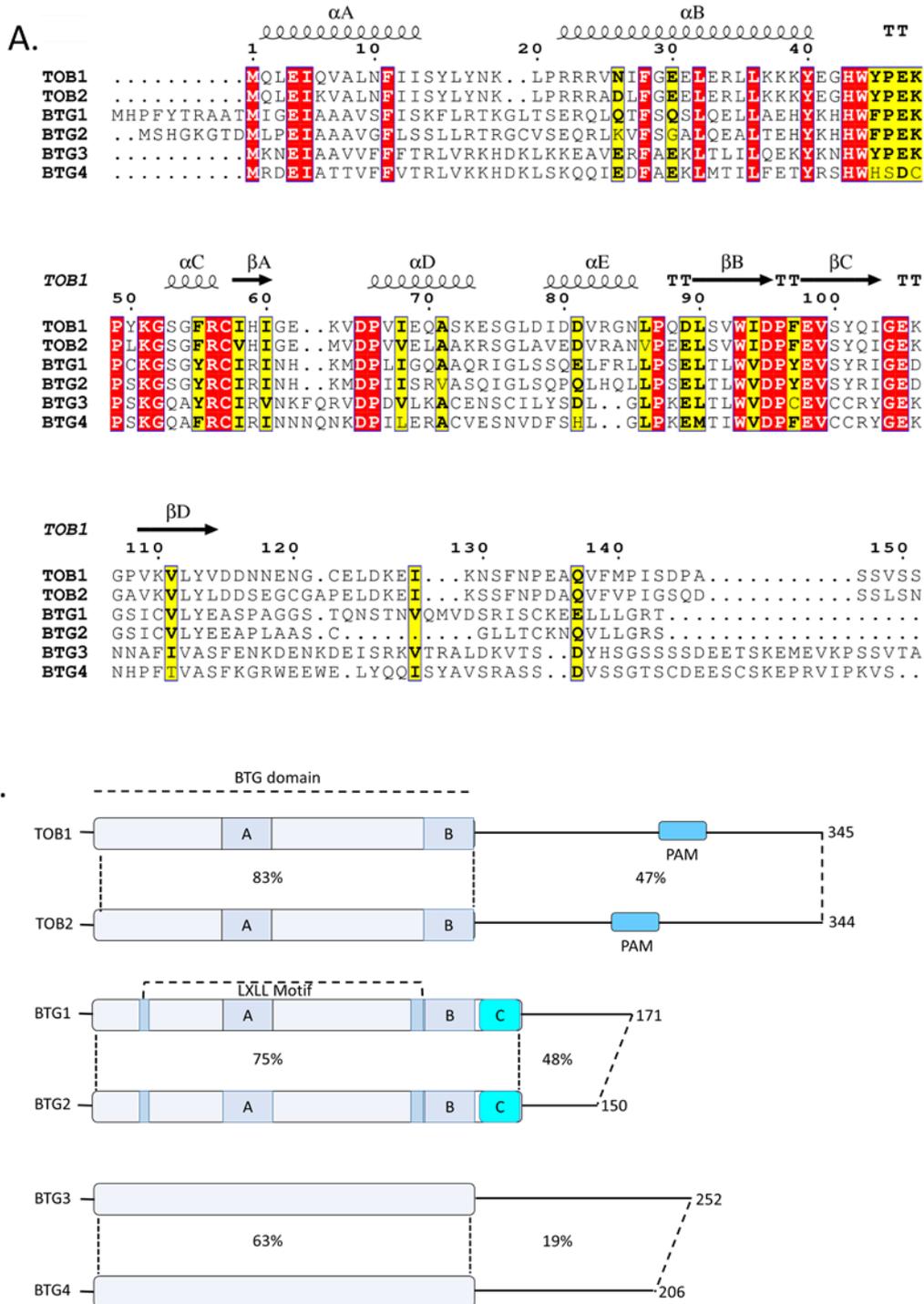
Figure 1-19 Sequence analysis for the BTG/TOB family of proteins. A) Sequence analysis of the BTG family of proteins with TOB 1 and TOB 2 in ENDscript (Robert & Gouet, 2014). Amino acids highlighted in red are identical residues found in the BTG and TOB family of proteins. The regions highlighted in yellow are the amino acids that differ between the two families. **B)** A schematic representation of the relationship between the members of the human BTG/TOB family of proteins. The figure represents the conserved BTG domain (blue), the conserved TOB1/TOB2 PAM motif (dark blue) and Box C of BTG1/BTG2 (turquoise).

1.3.3. The role of BTG1/BTG2 and TOB1 in regulating deadenylation

BTG/TOB proteins engage in mRNA turnover by regulating deadenylation through bridging CNOT7/CAF1 and PABPC1 (Stupfler et al., 2016). Except for BTG4, all BTG/TOB protein family members interact with CNOT7/CAF1. Since 40% of BTG2 and TOB1 are identical to each other, superposing BTG2 (PDB 3DJU) on the structure of TOB-CAF1 (PDB 2D5R) reveals the CNOT7/CAF1 binding site (Figure 1-20). The residues Tyr-65 in Box A and Trp-103 and Asp-105 in Box B of TOB1 mediate interaction with CNOT7/CAF1 (Horiuchi et al., 2009b).

Another interacting partner of BTG/TOB is PABPC1. The residues that interface with PABPC1 are in the C-terminus of BTG/TOB. Since the C-terminus of BTG/TOB are divergent, TOB mediates its interaction with PABPC1 via PAM2 motifs, while the BTG proteins interact with PABPC1 through residues in the Box C region. A previous study by Stupfler et al. (2016) revealed that the residues in the Box C region of BTG2 are critical for interfacing with the RRM1-2 of PABPC1. The study suggested that BTG2 bridges PABPC1 and CNOT7/CAF1 and regulates deadenylation by making the 3′ end of the mRNA accessible for the catalytic site of CNOT7/CAF1. Similarly, a study by (Ezzeddine et al., 2012)) suggested that TOB1 uses the PAM motif in its C-terminal to bind PABPC1 and regulate deadenylation by bridging CNOT7/CAF1 with its BTG domain.

1.3.4. The role of BTG1 and BTG2 in B and T cells

The BTG1 mutations identified in lymphoma are due to erroneous somatic hypermutation SHM being orchestrated by activation-induced deaminase during the development of B-cells. AID promotes antibody diversity and maturation. The off-target activity of AID can also result in mutation to adjacent oncogene or tumour suppressors and contribute to lymphomagenesis (Jiang et al., 2012). BTG1 and BTG2 play a key role in maintaining T-cell quiescence. A conditional knockout study of BTG1 and BTG2 revealed this role. No significant changes were revealed in a single knockout study in mice, suggesting that BTG1 and BTG2 are functionally redundant. However, the double knockout of BTG1 and BTG2 revealed a decreased number and frequency of naïve CD4 and CD8 T-cells compared to WT. BTG1 and BTG2 double knockout resulted in increased cell size and reduced cell proportion, indicating that the absence of BTG1 and BTG2 compromises T cell quiescence (Hwang et al., 2020).

Genome-wide profiling observed genetic aberration in BTG1 and BTG2 in B-cell malignancies. BTG1 and BTG2 identified in diffuse large B-cell lymphoma (DLBCL) were affected by somatic missense mutations (Lohr et al., 2012, Reddy et al., 2017). Poor survival of BTG1 mutations is associated with the subtype of DLBCL in patients. Thus, BTG1 mutations in follicular lymphoma (FL) correlates with disease prognosis (Kridel et al., 2016, Reddy et al., 2017). In Burkitt lymphoma (BL) subtypes, loss of growth control is due to suppressing BTG1 expression (De Falco et al., 2007). In lymphoplasmacytic lymphoma, a high alteration of BTG1 has been detected, revealing that the loss of BTG1 can contribute to the non-Hodgkin lymphoma subtype (Hunter et al., 2014).
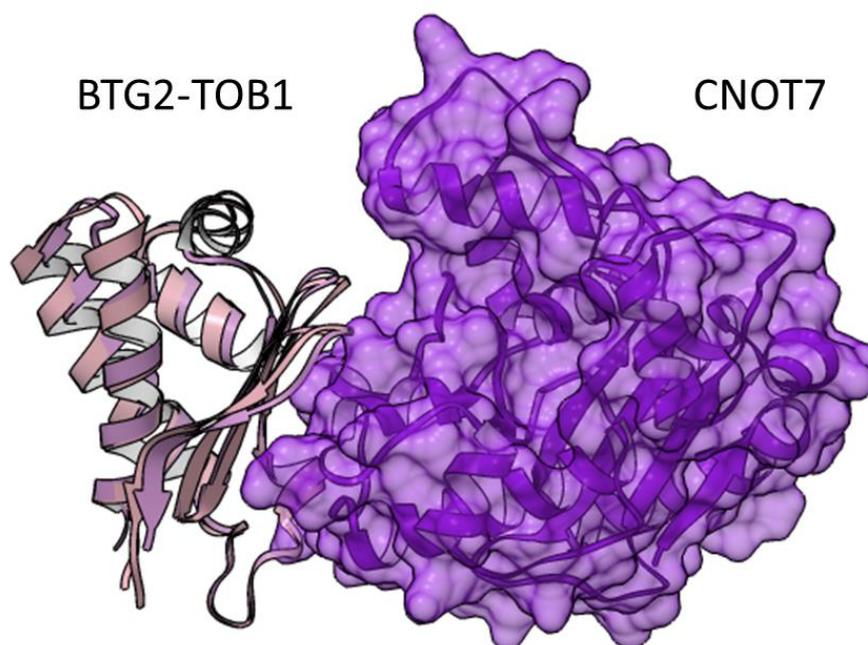
Figure 1-20 Structure of TOB1-CNOT7/CAF1 interaction. Crystal structure of BTG2 (PDB 3DJU, Yang et al., 2008) and TOB1-CNOT7/CAF1 (PDB 2D5R, Horiuchi et al., 2009b) interaction. The image represents superposing BTG2 (plum) to TOB 1 (pink) that is bound to CNOT7/CAF1 (purple). Sperposing BTG2 to TOB1-CNOT7/CAF1 demonstrates the identical nature of Box A and Box B in BTG2 and TOB1.

1.3.5. Study aim

This study aimed to obtain further insight into the mechanism by which PABPC1 stimulates deadenylation by CNOT7/CAF1 in the presence of BTG2. A study conducted by Stupfler et al. (2016) revealed that the residues in the Box C region of BTG2 are crucial for mediating the interaction with PABPC1 and regulating the enhanced rate of deadenylation by CNOT7/CAF1. Although the authors discovered that residues in Box C of BTG2 are responsible for interacting with the N-terminal RRM motif of PABPC1 (Stupfler et al., 2016), no studies have yet identified the specific PABPC1 residues that are responsible for interfacing with BTG2. Since no crystal structure has revealed PABPC1-BTG2 interactions, the present study aimed to identify the PABPC1 residues responsible for interacting with BTG2.

PABPC1 is an RNA binding protein that mediates interaction with the 3′ end of processed mRNA poly(A) tail through its RRM motif. The binding of PABPC1 to mRNA determines its fate in post-transcriptional modification. One of the critical roles of PABPC1 is to regulate poly(A) tail removal by interacting with poly(A) nucleases, such as PAN2-PAN3, PARN, and the CCR4-NOT complex. CNOT7/CAF1 and its homologue CNOT8/CAF1b are the catalytic subunits of the CCR4-NOT complex and regulate deadenylation. Deadenylation by CNOT7/CAF1 and CNOT8/CAF1b on its own does not reveal intense catalytic activity. This is due to PABPC1 being bound to the mRNA poly(A) tail, protecting the nucleotides from degradation.

CNOT7/CAF1 and CNOT8/CAF1b exhibit intense catalytic activity with their binding partners. A study conducted by Pavanello et al. (2018) suggested an enhanced rate of deadenylation was observed with CNOT7/CAF1 in combination with BTG2/TOB1 compared to CNOT7/CAF1 on its own. Another study revealed an enhanced rate of deadenylation by CNOT7/CAF1 in combination with CNOT6/CCR4a, an alternative catalytic subunit of the CCR4-NOT complex (Yi et al., 2018).

To achieve the aim, three specific objectives were defined:

1. To use a computational method to predict which PABPC1 residues are involved in the interaction with BTG2. This approach will be validated using known structures of PABPC1 in complex with poly(A) RNA (Safaee et al., 2012), and putative residues required for the interaction will be identified. This approach will also be used to generate computational models of BTG2-PABPC1 using data-driven docking, which will be used to select PABPC1 residues for experimental validation.

2. To experimentally validate the role of PABPC1 residues in PABPC1-BTG2 interactions by protein-protein interaction assays. To achieve this, point mutations will be introduced into PABPC1 using site-directed mutagenesis. The ability of PABPC1 variants to support interactions with BTG2 will be evaluated using several protein-protein interaction methods, including the yeast two-hybrid system and GST and His-pulldown assays, using proteins expressed in bacteria.

3. Finally, selected PABPC1 variants will be functionally characterised. This will be done by evaluating their capacity to bind poly(A) RNA using electrophoretic mobility assays (EMSAs) and their ability to stimulate deadenylation by CNOT7/CAF1 in the presence of BTG2.

Together, these experiments will provide more detailed insight into the molecular mechanism by which PABPC1 can stimulate deadenylation by CNOT7/CAF1 in the presence of BTG.

Chapter 2 - Materials and Methods

2.1 Computational analysis.

For predicting human PABPC1 interacting surfaces, the consensus prediction of interface residues in transient complex (CPORT) web server (https://milou.science.uu.nl/services/CPORT/) was used (de Vries and Bonvin, 2011a), which predicts the interfacing residues in combination with six algorithms: WHISCY (de Vries et al., 2006), ProMate (Neuvirth et al., 2004), SPPIDER (Porollo and Meller, 2007), PINUP (Liang et al., 2006), PIER (Kufareva et al., 2007) and cons-PPISP (Chen and Zhou, 2005). The "very sensitive" threshold setting was adapted by exploiting the PDB structure 4F02 (chain A) of human PABPC1 in complex with poly(A) and eIF4G and the corresponding HSSP alignment (Safaee et al., 2012).

To identify residues involved in protein-protein or protein-RNA interactions, the Protein, Interfaces, Structure and Assemblies (PISA) web server (https://www.ebi.ac.uk/pdbe/pisa/) as utilised using default settings (Krissinel and Henrick, 2007). Protein docking was conducted using the HADDOCK 2.2 webserver (https://milou.science.uu.nl/services/HADDOCK2.2/haddock.php. van Zundert et al., 2016). 'Active' residues were obtained from the PABPC1 structure (PDB 4F02) by CPORT, which was used as an input in docking; in addition to predicting active residues, CPORT also revealed surrounding surface residues, termed as 'Passive' residues. For BTG2 (PDB 3DJU), residues in Box C of BTG2 were assigned as an input.

The UCSF Chimera package was used to generate and visualise the structural model of the CNOT7-BTG2-PABPC1-poly(A) complex (Pettersen et al., 2004) based on the structures of CNOT7-TOB1 (PDB 2D5R, Horiuchi et al., 2009a), human BTG2 (PDB 3DJU, Yang et al., 2008), and the complex of PABPC1-poly(A)-eIF4G (PDB 4F02, Safaee et al., 2012).

2.2. Bacterial growth and transformation.

2.2.1. Reagents, stock solutions, and buffers.

Lysogeny broth (LB) media and agar: 10g/L tryptone, 5g/L yeast extract, 5g/L NaCl, pH 7.2 (NaOH), for LB agar include 15g/L bacteriological agar. The medium was sterilised by heat using a Prestige medical 2100 classic bench-top autoclave and stored at room temperature.

Ampicillin: 1000x stock solution: 100mg/ml in 50% ethanol/water. Sterilised by filtration (0.22 μm pore size). Stored at -20°C.

1M IPTG: Isopropyl β-2-1-thiogalactopyranoside (IPTG), in H_2O, was sterilised by filtration (0.22 μm pore size) stored at -20°C.

Glycerol 50% solution: 50 ml of glycerol mixed with 50 ml ddH2O, autoclaved with a Prestige medical 2100 classic bench-top autoclave and stored at 4°C.

2.2.2. Culture of *E. coli* DH5α/BL21 (DE3).

The bacterial strain *E. coli* DH5α was used for the manipulation of DNA. The proteins were expressed in the *E. coli* BL21 DE3 strain. Colonies were obtained by streaking the bacterial strain on the agar plate and incubating overnight at 37°C. Colonies formed in the plate were stored at 4°C for several weeks. A single colony from the plate was inoculated to liquid cultures and grown in LB medium incubated at 37°C, 200rpm for 12-24 h.

2.2.3. Preparation of *E. coli* competent cells.

The *E. coli* strains DH5α or BL21 (DE3) were made competent for transformation using calcium chloride. A disposable plastic inoculation loop was used to pick a single colony from the LB plate and inoculated in a 2ml LB medium using a universal tube. The inoculated culture was grown overnight at 37°C in a shaker at 200rpm. The following morning, 0.5ml of the overnight grown culture was added to 50 ml of LB medium free of antibiotics and shaken vigorously at 37°C to an optical density (OD_{600}) of 0.3.

The culture was kept on ice for 20-30 min in 50 ml falcon tubes and harvested by centrifugation at 1100 x g at 4°C for 10 min (Eppendorf 5810R) pellet was suspended with 25ml of ice-cold 0.1 M $CaCl_2$. The suspended solution was centrifuged (Eppendorf 5810R) again at 1100g, at 4°C for 10 min. The cell pellet was then gently re-suspended in 4 ml of ice-cold 0.1 M $CaCl_2$, containing 25% of glycerol. The re-suspended cells were distributed into 200µl aliquots in a microfuge and stored at -80°C.

2.2.4. Bacterial transformation of plasmids

Bacterial strains BL21 (DE3) or DH5α competent cells transformed plasmids. Before plasmid transformation, the competent cells were thawed on ice and 10ng of the designated plasmid was added to 50µl competent cells in an Eppendorf tube and mixed carefully. The Eppendorf tube was placed in ice and transferred to a water bath at 42°C for 90 secs to achieve transformation. Transformed DH5α or BL21 (DE3) cells were placed on ice for 5 minutes to cool. 1000µl of LB medium was added to the transformed cells. The mixture in the Eppendorf was mixed carefully, and 100µl of LB medium mixed to the transformed plasmid was spread on an LB plate containing appropriate antibiotics and incubated overnight in an incubator at 37°C.

2.3. Molecular Biology

2.3.1. Reagents, stock solutions and buffers

10x Tris-borate-EDTA (TBE): 40 mM Tris base, 40 mM boric acid, 1 mM EDTA (Ethylene Diamine Tetraacetic Acid), pH 7.9, stored at room temperature.

Oligonucleotides: Dry primers were dissolved in H_2O to 100 μM (final concentration) and stored at -20°C.

dNTP mix (10 mM): 2.5 mM each of dATP, dGTP, dCTP and dTTP in Tris-HCl pH 8.0

6× Gel Loading Dye Purple: 60 mM Tris-HCl pH 7.9, 60 mM EDTA, 60% glycerol; 0.075 g Orange G; (0.15% w/v). Stored at -20°C.

2.3.2. Small scale plasmid DNA preparation

Plasmid DNA was prepared on a small scale using Sigma Aldrich GenElute plasmid miniprep kits (PLN35-1KT). In brief, a single bacterial colony was inoculated to 5ml LB containing the appropriate antibiotic. The inoculated culture was grown overnight in an orbital shaker at 37°C, 200rpm. The overnight cultures were centrifuged in a microfuge tube at 4000g for 5 min, and the pellet was re-suspended in 200µl of resuspension buffer (Sigma Aldrich). The re-suspended pellet was lysed by 200µl of lysis reaction (Sigma Aldrich), and the mixture was neutralised by adding a 350µl binding solution.

The mixture was centrifuged at 12,000 x g for 10 minutes. Before adding the clear lysate to the column, the column was prepared using a 500µl column prep solution (Sigma Aldrich) and centrifuged at 12,000 x g for 30sec. The clear lysate was added to the column and centrifuged at 12,000 x g for 30 sec. The flow-through was discarded and followed by adding 750µl of wash solution (Sigma Aldrich), and further centrifugation step was followed at 12,000 x g for 30 sec, and the flow-through was discarded. The column was centrifuged to remove excess wash solution prior elution step. The plasmid DNA was then eluted in 30µl of elution buffer by centrifuging at 12,000 x g for 30sec.

2.3.3. Determination of DNA concentration

The concentration of DNA prepared by miniprep was quantified by spectrophotometry in Nanodrop ND-1000 UV-Vis Spectrophotometer (Nanodrop Technologies). The sample purity was estimated by calculating A260/A280 ratio value.

2.3.4 Site-directed mutagenesis (Inverse PCR)

Inverse PCR (Polymerase Chain Reaction) method was approached to delete a chain of amino acids from BTG2 and introduce point mutation to PABPC1 (1-190) at position Arg-49. Inverse PCR primers were designed to extend away from each other rather than move towards each other in contrast to regular primers (Figure 2-2.). The SnapGene program was used for designing primers to delete BTG2 residues (Table 2-3.) and to introduce point mutation to PABPC1 (1-190) residues at position Arg49 (Table 2-3.). The PCR reaction was carried out in 20 µl (final volume) following the manufacturer's protocol using the Phusion polymerase (Table 2-4.). The reaction was set up in a thin-wall PCR tube containing 4µl of 5x HF Phusion reaction buffer, 0.4 µl dNTP (10mM), 1µl FW primer (10µM), 1µl RV primer (10µM), 5ng of template DNA and 0.5 µl Phusion enzyme mixed carefully. The settings in a thermal cycler (primus 96) were set to initial denaturation 95°C/30sec; followed by 30 cycles consisting of denaturation 98°C/15 secs, annealing 55°C/30 sec, extension 72°C/5 min and a final extension at 72°C for 10 min.

2.3.5 Site-directed mutagenesis (Quick-Change)

The primers for site-directed mutagenesis of plasmid pQE80L-PABPC1 (1-190, Table.2-1) and pGEX4T1-BTG2 (Table.2-2) were designed using the Quick Change[®] primer design program (https://www.agilent.com/store/primerDesignProgram.jsp). The primers designed by the Quick Change® program minimises the free energy of the mismatched primer-template duplex. The energy of the nearest neighbour calculates the free energy for nucleic acid duplexes. The program is designed to place the mismatches close to the middle region of the primer-template; this enables the more stable region of the primer template to flank over the less stable region of the primer (Figure.2-1 Alexey Novoradovsky et al., 2005)

The mutagenesis reaction was carried out in 40µl (final volume) following a modified Quikchange protocol (Agilent) using Phusion polymerase (Table.2-4). The reaction was set up in thin-wall PCR tubes, 8µl of 5x HF Phusion reaction, followed by 0.8µl dNTP (10mM), 2µl forward primer (FW 10µM), 2µl 10µM reverse primer (RV 10µM)), 250ng of template DNA and 0.5 µl Phusion enzyme. After combining all components, the reaction components were mixed carefully. The mutagenesis reaction was carried out using a thermal cycler (primus 96) and the following conditions: initial denaturation 98°C/30sec, followed by 30 cycles composed of denaturation 98°C/15 secs, annealing 55°C/30 secs, extension 72°C/15 mins. A final extension for 72°C/10 min was carried out.

Step 1
Plasmid preparation

Target site for mutation

Step 2
Temperature cycling

Plasmids denatured Oligonucleotide
Primers with desired mutations are annealed

Phusion polymerase extend and incorporate
Mutagenic primers resulting in circular strand.

Step 3
Digestion

Dpn I digestion of methylated and
non mutated parental DNA.

Step 4
Transformation

Transform circular nicked DNA in to DH5α.

Figure 2-1. A schematic representation of Quik change protocol for site-directed mutagenesis. The Agilent Quick change program designed the primers for site-directed mutagenesis based on Quik-change protocol. The primers with target mutations are subjected to a temperature cycle in PCR, followed by Dpn I digestion and transformation of nicked DNA into DH5α. The publication on computational principles of primer design for site-directed mutagenesis by (Alexey Novoradovsky et al., 2005) was used to make the figure.

2.3.6. DpnI digestion

To remove template DNA from the mutagenesis reaction, 2 µl of DpnI enzyme (20,000units/ml. NEB) was added to 30µl of the quick-change mutagenesis reaction and 20µl to inverse PCR reaction. The reactions were incubated at 37°C in a heat block for 2 hours. 5µl of quick-change mutagenesis product and 20 µl of inverse PCR product digested by DpnI was analysed by agarose gel electrophoresis mentioned in section 2.3.7 to verify the removal of the DNA template. Following DpnI digestion, 5 µl of the quick-change reaction products were transformed into DH5α and incubated overnight at 37°C. For inverse PCR mutagenesis, the product was subjected to phosphorylation and ligation mentioned in section 2.3.9.

2.3.7 Agarose gel electrophoresis

Agarose gel electrophoresis was used for analysing the products of the site-directed mutagenesis reactions. 0.8% - 1% gels were cast using TBE according to the size of the DNA fragments resolved. Before casting the gel, ethidium bromide (0.5μg/ml) was added to a 1x TBE buffer containing agarose. 20µl of inverse PCR products and 5µl of the quick-change product was mixed with 1x loading dye and loaded in the gel wells along with the undigested template. DNA gel with digested samples and undigested template was run parallel with 1x TBE buffer at 80-90 V for 45 mins. The difference in bands from digested samples and undigested templates in gels were visualised by trans-illumination with UV light and imaged using a Gel-Doc 2000 (Bio-Rad) and Quantity One software (Bio-Rad).

2.3.8. DNA extraction from agarose gel

Following digestion by Dpn I, the inverse PCR mutagenesis reactions were purified by agarose gel electrophoresis. Following electrophoresis, the DNA band was excised using a surgical scalpel and purification was performed by Monarch gel extraction kit (Hinnebusch and LorschT1020l) following the manufacturer's protocol. To excise the DNA fragment from the agaroses gel, a gel dissolving buffer was added at a ratio of 100mg of agarose/400 µl buffer (1:4).

The sample was incubated for 5 min at 37°C on a heat block; the samples were vortexed periodically to completely dissolve the gel slice in the buffer. A column was inserted into the collection tube, and the samples were loaded onto the column. The column was spun down at 13,000 x RPM for 1 min, and the flow-through was discarded after the centrifugation step. The column was re-inserted in the collection tube, 200µl of wash buffer was added and spun down at 13,000 x RPM for 1 min. The wash step was repeated once, and the column was transferred to an Eppendorf tube. DNA was eluted by the addition of 10µl of elution buffer followed by centrifugation at 13,000 x RPM for 1 min.

2.3.9. Phosphorylation and Ligation

Since the inverse PCR products (section 2.3.4) Following phosphorylation, ligation reactions were set up to generate plasmids pQE80L-PABPC1 (1-190) and pGEX4T1-BTG2 with mismatched and deleted amino acid residues. The ligase reaction was set up by adding 1µl of T4 DNA ligase in a 1x T4 DNA ligase buffer bringing it to the final reaction of 10µl. The reactions were incubated overnight at room temperature, and the reaction was transformed to DH5α competent cells and analysed the next day for colonies.
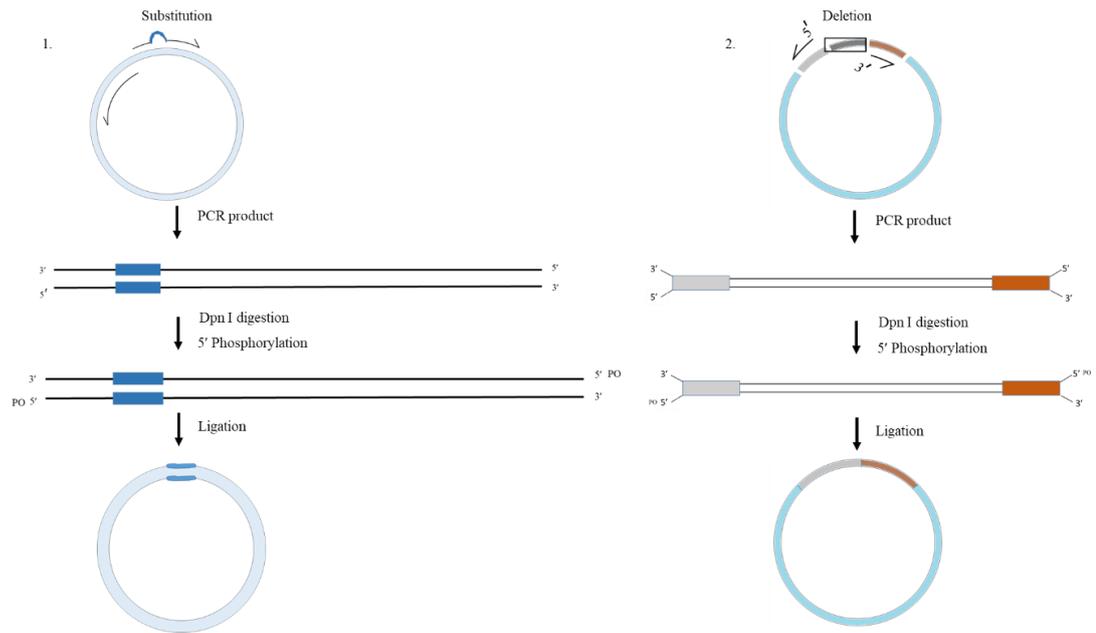
Figure 2-2. A schematic representation of site-directed mutagenesis by inverse PCR. The primers for site-directed mutagenesis by inverse PCR was designed in SnapGene. The primers with amino acids of interest for substitution and deletion are subjected to inverse PCR. The amplicons obtained from inverse PCR reaction are subjected to Dpn I digestion, phosphorylation, and ligation. The Integrated DNA Technologies website link was used as a reference for creating the image. (https://eu.idtdna.com/pages/education/decoded/article/methods-for-sitedirected mutagenesis)

Name	Sequence
D21A_FW	3'GGAGGTGGGGC**G**GCACTGGCTCC5'
D21A_RV	5'CCTCCACCCCG**C**CGTGACCGAGG3'
E24A_FW	3'GGGCTGCACTGGC**G**CCGCTACGAGATG5'
E24A_RV	5' CCCGACGTGACCG**C**GGCGATGCTCTAC3'
E29A_FW	3' CCGCTACGAGATGC**G**CTTCAAGTCGGGCC5'
E29A_RV	5' GGCGATGCTCTACG**C**GAAGTTCAGCCCGG3'
I40A_FW	3'CGGGTAGGAGAGG**CG**GGCCCAGACGTCC5'
I40A_RV	5'GCCCATCCTCTCC**GC**CCGGGTCTGCAGG3'
R44A_FW	3'GGTAGGCCCAGACG**CG**CCTGTACTAGTGGG5'
R44A_RV	5'CCATCCGGGTCTGC**GC**GGACATGATCACCC3'
M46A_FW	3'CCAGACGTCCCTG**CG**CTAGTGGGCGGCG5'
M46A_RV	5'GGTCTGCAGGGAC**GC**GATCACCCGCCGC3'
T48A_FW	3'GTCCCTGTACTAG**C**GGGCGGCGAGG5'
T48A_RV	5'CAGGGACATGATC**G**CCCGCCGCTCC3'
T48P_FW	3'GTCCCTGTACTAG**G**GGGCGGCGAGG5'
T48P_RV	5'CAGGGACATGATC**C**CCCGCCGCTCC3'
R50A_FW	3'TGTACTAGTGGGCG**CG**GAGGAACCCGATGC5'
R50A_RV	5'ACATGATCACCCGC**GC**CTCCTTGGGCTACG3'

Table 2-1. PABPC1 primers for site-directed mutagenesis. Primers designed for site-directed mutagenesis of codons encoding residues in RRM1 of PABPC1. The mismatches are indicated in bold with larger font size.

Name	Sequence
S118A_FW	3'CCCCTCCTGCCG**C**GGTAGACGCAGA5'
S118A_RV	5'GGGGAGGACGGC**G**CCATCTGCGTCT3'
I119V_FW	3'CCCTCCTGCCGAGG**C**AGACGCAGAACATG5'
I119V_RV	5'GGGAGGACGGCTCC**G**TCTGCGTCTTGTAC3'
E124A_FW	3'ACGCAGAACATGC**G**CCTCCGGGGTGAC5'
E124A_RV	5'TGCGTCTTGTACG**C**GGAGGCCCCACTG3'
E125A_FW	3'CAGAACATGCTCC**G**CCGGGGTGACCGG5'
E125A_RV	5'GTCTTGTACGAGG**C**GGCCCCACTGGCC3'
S118A_I119V_FW	3'AACCCCTCCTGCCG**C**GG**C**AGACGCAGAACATG5'
S118A_I119V_RV	5'TTGGGGAGGACGGC**G**CC**G**TCTGCGTCTTGTAC3'
D124L_E125D_A12 6D_FW	3'AGGTAGACGCAGAACATG**GAT**CT**A**CTGGGTGACCG GCGGAGGAC5'
D124L_E125D_A12 6D_RV	5'TCCATCTGCGTCTTGTAC**CTA**GA**T**GACCCACTGGC CGCCTCCTG3'

Table 2-2. BTG2 primers for site-directed mutagenesis. Primers designed for site-directed mutagenesis of codons encoding residues in Box C of BTG2. The mismatches are indicated in bold with larger font size.

Name	Sequence
R49A_FW	**GCG**CGCTCCTTGGGCTACGCGTATG
R49A_RV	GGGTCTGCAGGGACATGATCACCCGC
R49E_FW	**GAA**CGCTCCTTGGGCTACGCGTATG
R49E_RV	GGGTCTGCAGGGACATGATCACCCGC
BTG2_FW_Del_After_Box-C	CTCCCCAATGCGGTAGGACAC
BTG2_RV_Del_After_Box-C	CCACTGGCCGCCTCCTGT
BTG2_FW_Del_Box-C	GGCCTCCTCGTACAAGACG
BTG2_RV_ Del_Box-C	TAGCGGCCGCATCGTGAC

Table 2-3. PABPC1 and BTG2 primers for inverse PCR. Primers designed for site-directed mutagenesis using inverse PCR. Mutations were introduced to the 49th position of PABPC1. This approach was also used to delete regions of BTG2.

Component	20µl reaction	40µl reaction	Final concentration
Nuclease free water	20 µl	40 µl	
5 x Phusion HF	4 µl	8 µl	1x
10 mM dNTPs	0.4 µl	0.8 µl	200 µM
10 µM FW primer	1 µl	2 µl	0.5 µM
10 µM RV primer	1 µl	2 µl	0.5 µM
Template DNA	variable	variable	< 250 ng (Quick-change)
			5ng (Inverse PCR)
Phusion DNA polymerase	0.2 µl	0.2 µl	1.0 units/50 µl PCR

Table 2-4. Site-directed mutagenesis reactions. The table represents the reaction set-up for the site-directed mutagenesis experiment. 20µl reaction was set up for mutagenesis experiment using inverse PCR, while 40µl reaction was set up for method adapted by quick-change protocol.

2.4. Yeast two-hybrid analysis

2.4.1. Reagents, stock solutions, and buffers used yeast culture

Yeast extract peptone dextrose (YPD) medium: 4g/500ml yeast extract, 8g/L glucose, 8g/L bacteriological peptone. For YPD agar, 8g/L bacteriological agar was added and filled with 1L dH$_2$O. Prestige medical 2100 classic bench-top autoclave was used to sterilise the medium. The medium was sterilised and stored at 4°C.

Yeast selective complete (YSC) Glucose medium: 2.67g/L yeast nitrogen base, 10g/L glucose, 0.4g/500L amino acid drop out mix (minus histidine, uracil, tryptophan, and leucine). For selective agar plates, 10g/L bacteriological agar was added. Prestige medical 2100 classic bench-top autoclave was used to sterilise the medium. The medium was sterilised and stored at 4°C.

Galactose & Sucrose: 20% of Galactose was made in 100ml of dH$_2$O, and 10% of Sucrose was made in 100ml dH$_2$O.

Yeast galactose medium: 0.85g/L yeast nitrogen base, 2.5g/L ammonium sulphate, 0.4g/L amino acid dropout mix (minus histidine, uracil, tryptophan, and leucine) and 10g/L bacterial agar were added along with 50ml of made galactose and sucrose.

Leucine: 120g/20ml was made and poured into glucose and galactose medium with agar. The recipe for all three mediums was made in 1 litre and divided into two in which 500 ml of medium contained a bacteriological agar and ingredients as mentioned earlier and the remaining 500 ml with no agar containing all ingredients as mentioned above. The 500ml of medium containing agar was poured into a petri dish and stored at 4°.

Single-stranded DNA: Herring Sperm single-stranded DNA (2 mg/ml) was heated to 98°C for 5min, placed on ice before use, and stored at -20°C.

Lithium acetate solutions: (LiAc, 1M and 0.1M; Sigma) were prepared and sterilised by filtration (0.22 μm pore size). Stored at room temperature. PEG-3350 solution: a solution containing 50% (w/v) polyethylene glycol (PEG-3350; Sigma) was sterilisation by filtration (0.22 µm pore size). The solution was stored at room temperature.

2.4.2. Culture of Saccharomyces cerevisiae strain EGY48

The *S. cerevisiae* yeast strain EGY48 was used for conducting the yeast two-hybrid experiment (Golemis et al., 2001b, Gyuris et al., 1993). The yeast strain was stored at -80°C in a YPD medium containing 50% glycerol. EGY48 strain from -80° was streaked on YPD plates, and the yeast was grown incubating at 30°C for 2-4 days. The yeast colonies formed in the plates were stored at 4° for several weeks. The YPD plate with yeast colonies is stored at 4°C to use for transformation the day of the experiment transformation. A single colony from the plate was inoculated in liquid culture containing YPD medium and grown in universal tubes. The culture was incubated at 30°C at 200rpm for 12-24 h.

2.4.3. Transformation of EGY48 cells

The plasmids pEG-202, pJG4-5 and reporter plasmid pSH138(Origene) with the gene of interest were used for transformation. Yeast strain EGY48 cells were made competent on the day of the transformation by heat shock treatment. A single yeast colony incubated for 48hrs was used to inoculate 10ml of YPD medium and grown overnight at 30°C at 200rpm. Two aliquots were made by inoculating 1ml of overnight grown culture in 50 ml YPD medium, and the culture was grown to give an OD_{600} of 0.2. The inoculated culture was further incubated at 30°C, to attain an OD_{600} of 0.8 (3-5 h). After attaining the required OD_{600}, the culture was centrifuged at 3000 x g for 5 mins, decanting supernatant and re-suspending the pellet with 25ml of sterile water.

The resuspension was further centrifuged at 3000 x g for 5 mins, the supernatant was decanted, and the yeast was resuspended with 1ml of 0.1M lithium acetate. The pellet re-suspended with 0.1M lithium acetate was transferred to a 1.5ml Eppendorf and further centrifuged at 4000 x g for 5 mins. The supernatant was decanted, and the pellet was re-suspended with 500µl of 0.1M lithium acetate. 50µl of the resuspended pellet with 0.1M lithium acetate was pipetted to five 1.5ml tubes.

The aliquots were centrifuged at 12,000 x rpm for 30 secs. Before adding transformation components, 50% of PEG solutions were added to each aliquot and mixed with transformation components (Table.2-5). Further, the plasmids containing the desired protein of interest was added to yeast cells (Table.2-5). The mix was vortexed after adding all transformation components until the pellet was resuspended entirely. Once the pellet was resuspended, the mix was incubated for 30 min at 30°C. The mix was incubated further for 20 min at 42°C in a water bath. After centrifuging the samples at 4000 x g for 15 sec, 500µl of sterile H_2O was used to re-suspend the yeast pellet, and 100µl was spread on the SC selective plates with leucine. The plates were incubated at 30°C until colonies formed.

pEG202(LexA)	pEJG4-5 (B42)	Transformation Component	Quantity
Empty Vector	Empty vector	-	250ng
BTG1	Empty vector	-	250ng
CNOT7	PABPC1 (1-140)	-	250ng
BTG1	PABPC1 (1-140)	-	250ng
-	-	ssDNA	25µl
-	-	LiAC 1mM	36 µl
-	-	DH_2O	48 µl
-	-	pSH138 (reporter)	250ng

Table 2-5. Plasmids and master mix for yeast two-hybrid experiments. The sample preparation for the experiments and all pEG-202 plasmids contains the Lex-A domain, whereas pEJG4-5 contained B42. Master Mix prepared for yeast two-hybrid to mix with samples for transformation.

2.4.4. Yeast two-hybrid β-galactosidase assay

For each successful transformation, three independent colonies were tested. Galactose medium agar plates were used to streak colonies in a zig-zag manner. The streaked colonies are incubated at 30°C for 48-72 hrs. Following the duration of incubation, the yeast grown was inoculated to 2 ml of YPD medium and incubated at 30°C, 200 rpm for 24h. A 1:9 dilution from the overnight grown culture was made to obtain OD_{600}. The β-galactosidase assay was conducted by adding 25µl of the diluted culture in a white flat bottomed 96 well plates which are followed by the addition of 25µl of β-Glo reagent (Promega). The plate was incubated in the dark for 30min at room temperature. Berthold Orion microplate luminometer was used to measure the β-Galactosidase 1 sec using Simplicity 4.02 software. The β-galactosidase activities for each culture were tested in duplicates and normalised to OD_{600} values.

2.5. Protein interaction analysis by His/GST pulldown and SDS-PAGE analysis

2.5.1. Reagents, stock solutions and buffers for His/GST pulldown.

1M IPTG: Isopropyl β-2-1-thiogalactopyranoside (IPTG), in H2O. It is sterilised by filtration (0.22 µm pore size) and stored at -20°C.

Lysis buffer (His-tagged): 20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 40mM imidazole and 0.1% NP-40.

Lysis buffer (GST-tagged): 20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol.

Wash buffer (His-tagged): 20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 0.1% NP-40, 40mM imidazole.

Wash buffer (GST-tagged): 20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol.

Elution buffer (His-tagged): 20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 0.1% NP-40, 200mM imidazole.

Elution buffer (GST-tagged): 20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol. 10% Glutathione was made on the day of purification for eluting GST tagged proteins.

Upper buffer (4x): 0.5M Tris base, 0.4% SDS, pH6.8; stored at room temperature.

Lower buffer (4x): 1.5M Tris base, 0.4% SDS, pH8.8; stored at room temperature.

Running buffer (10x): 0.25M Tris Base, 1.0% SDS, 1.92M glycine; stored at room temperature.

SDS (Sodium Dodecyl Sulfate) loading buffer (4x): 50% upper buffer, 40% glycerol, 1% β-mercaptoethanol, 1.25% Bromophenol blue (0.1%), 0.25% H2O; stored stock at -20°C. Small aliquots were stored short-term at 4°C.

10% APS: 10% ammonium persulphate (APS) in H2O

40% acrylamide: bisacrylamide (29:1) solution

2.5.1. Protein expression of His/GST tagged proteins

A single colony from the transformed plasmid was inoculated to a 20 ml LB medium containing Ampicillin (1:1000) and was grown overnight in an orbital incubator (Stuart-S1500) 190 rpm at 37°C. The overnight grown culture was transferred to a fresh stock of 200 ml LB containing ampicillin and grown in a shaking incubator (Lab net-311DS) 190 rpm at 37°C until the OD_{600} reached 0.6 – 0.8. The culture was induced with 0.2 M IPTG for 3hrs at 30°C in the orbital incubator. After 3hrs of induction, 1000µl of induced culture was pipetted and centrifuged at 12,000 rpm for one minute. The pellet was re-suspended with 1000µl of dH_2O, and 10µl of the re-suspended pellet was analysed in 14% SDS PAGE for expression.

2.5.2. Sonication

The induced culture for His-PABPC1/Variants was centrifuged in 50ml tubes (Eppendorf-5810R) at 4000 rpm in 4° for 10 minutes and re-suspended in His-PABPC1 lysis buffer. The cultures re-suspended with lysis buffer were sonicated in (Qsonica XL2000) sonicator for five cycles at 60% amplitude and power being 20sec on/20sec off. The crude lysates were spun in (ALC PK131R) centrifuge in (A-M10) rotor at 9000 rpm at 4ºC for 30 minutes. The clear lysates were stored at -80° in aliquots of 500µl for future use.

2.5.3. Calibration of His/GST beads

The pulldown experiments were performed with four types of beads, as detailed in the following sections.

2.5.3.1. Washing and equilibrating His- agarose beads

500µl of Ni-NTA agarose beads (Qiagen Catalog number 30210) in a microfuge tube was centrifuged at 2500 x rpm for 2 mins. The supernatant was decanted; the beads were washed with 1000µl (1:2) of wash buffer. The washed beads were centrifuged; the supernatant was decanted, and the beads were equilibrated with 2000µl (1:4) resuspension buffer (20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 0.1% NP-40, 40mM imidazole). 50µl of equilibrated beads were used for pull-down analysis, the resuspension buffer in beads was centrifuged at 2500 x g for 5 mins, and the supernatant was decanted before pull-down.

2.5.3.2. Washing and equilibrating Co^{2+}-agarose bead

20µl of Co^{2+}-agarose beads (HisPur, Thermo Scientific, Catalog Number 89964) aliquoted in microfuge tubes were centrifuged at 10,000 rpm for 1 min, the supernatant was decanted, and the beads were washed with 40µl (1:2) of resuspension buffer. The washed beads were centrifuged; further, the supernatants were decanted, and the beads were equilibrated with 80µl (1:4) of resuspension buffer (20mM, Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 0.1% NP-40, 40mM imidazole). The equilibrated beads with resuspension buffer were centrifuged at 10,000 rpm for 1 min, and then equilibrated beads were used for pulldown analysis after decanting the supernatant.

2.5.3.3. Washing and equilibrating Glutathione agarose magnetic beads

50µl of beads (Pierce, Thermo Scientific, Catalog Number 78602) aliquoted in microfuge tubes were placed on magnetic stands (Dyna Mag), the supernatant was pipetted, and the beads were washed with 100µl (1:2) of resuspension buffer vortexed for 30 secs and placed in the magnetic stand. The supernatant was pipetted, and the beads were equilibrated with 200µl (1:4) of resuspension buffer further, the Eppendorf was placed in the magnetic stand, and the supernatant was decanted for pulldown analysis.

2.5.3.4. Washing and equilibrating nickel-nitrilotriacetic acid (Ni²⁺-NTA) magnetic beads

10µl of beads (HisPur, Thermo Scientific, Catalog Number 88832) aliquoted in microfuge tubes were placed on magnetic stands (Dyna Mag), the supernatant was pipetted, and the beads were washed with 20µl (1:2) of resuspension buffer vortexed for 30 secs and placed in the magnetic stand. The supernatant was pipetted, and the beads were equilibrated with 40µl (1:4) of resuspension buffer; further, the Eppendorf was placed in a magnetic stand, and the supernatant was decanted for pulldown analysis.

2.5.4. Protein interaction by His/GST pulldown

2.5.4.1. His-Pulldown using agarose beads

His-pulldown experiments were done to study the interaction of His-PABPC1/Variants (1-190) with GST-BTG2 using His-agarose beads. The analysis was performed by mixing the lysates of PABPC1 and BTG2 proteins with equilibrated beads and incubating for an hour at 4°C by mounting Eppendorf's on the microfuge tube rotator. The agarose beads (Qiagen Catalog number 30210) are washed thrice with 500µl of wash buffer (20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 0.1% NP-40, 40mM imidazole) by centrifugation at 3000g for 5mins. 25µl of elution buffer (20mM, Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 0.1% NP-40, and 200mM imidazole) was added to the samples and centrifuge at 5000g for 5 mins. The samples were eluted using gel loading tips, and the eluted samples were added to a 1x SDS loading buffer. The samples were heated for 5 mins at 95°C and analysed through the SDS page and staining with coomassie blue.

2.5.4.2. His-pulldown using Co^{2+}-agarose beads

His-pulldown experiments were done to study the interaction of His-PABPC1/Variants (1-190) with GST-BTG2 using (HisPur, Thermo Scientific, Catalog Number 89964). The analysis was performed by mixing the lysates of PABPC1 and BTG2 proteins with equilibrated beads in a microfuge tube. The mixture was incubated for an hour at 4°C by mounting the microfuge tube on the rotator. The agarose beads are washed thrice with 500µl of wash buffer (20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 0.1% NP-40, 40mM imidazole) by centrifugation at 3000g for 5mins. 50µl of elution buffer (20mM, Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 0.1% NP-40, 200mM imidazole) was added to the samples and centrifuge at 5000g for 5 mins, and the samples were eluted using gel loading tips, the eluted samples were added to 1x SDS loading buffer, and the samples were heated for 5mins at 95°C and analysed through SDS-PAGE and staining with coomassie blue.

2.5.4.3. Glutathione S-transferase (GST) pulldown by glutathione agarose magnetic beads

GST-pulldown experiments were done to study the interaction of GST-BTG2/Variants with His-PABPC1 (1-190) using GST agarose beads (Pierce, Thermo Scientific, Catalog Number 78602). The analysis was performed by mixing GST-BTG2 lysates with equilibrated beads in a microfuge tube. The mixture was incubated for an hour at 4°C on a microfuge rotator. The beads incubated with GST-BTG2 for one hour was mixed with lysate of His-PABPC1 and further incubated for one hour at 4°C by mounting the Eppendorf on the microfuge tube rotator.

The agarose beads are washed thrice with 500µl of wash buffer (20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol). The Eppendorf was placed on a magnetic stand for each wash, and the supernatant was pipetted gently. Following the wash, samples were eluted in 50µl elution buffer by placing the Eppendorf on a magnetic stand and subjected for SDS page analysis. Eluted samples were added to 1x SDS loading buffer, and the samples were heated for 5mins at 95°C and analysed through SDS-PAGE and staining with coomassie blue.

2.5.4.4. His-pulldown by nickel-nitrilotriacetic acid (Ni^{2+}-NTA) magnetic beads

His-pulldown experiments were done to study the interaction of His-PABPC1/Variants (1-190) with GST-BTG2 using Ni2^{+}-NTA beads (HisPur, Thermo Scientific, Catalog Number 88832). The analysis was performed by mixing lysate of His-PABPC1 (1-190) proteins with equilibrated beads in microfuge tubes and incubating for an hour at 4°C on a microfuge rotator. The beads incubated with His-PABPC1 (1-190) for one hour was mixed with lysate of GST-BTG2 and further incubated for one hour at 4°C by mounting the Eppendorf s on the microfuge tube rotator. The agarose beads are washed thrice with 250µl of wash buffer (20mM, Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 0.1% NP-40, and 200mM imidazole). The Eppendorf was placed on a magnetic stand for each wash decanted the supernatant. The samples were eluted in 50µl elution buffer (20mM, Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 0.1% NP-40, 200mM imidazole) by placing the Eppendorf's on a magnetic stand and subjected for SDS page analysis. 1x SDS loading buffer and the samples were heated for 5mins at 95°C and analysed through SDS page and staining with Coomassie blue.

2.5.5. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

His-Pulldown samples were analysed using SDS-PAGE with the X-Cell Sure Lock Mini Cell system (Invitrogen). SDS-PAGE gels were prepared to the appropriate percentage of resolving gel based on the molecular mass of the protein of interest. The gels were cast according to the manufacturer's Protocol. The wells were washed with 1x running buffer before use, and protein samples were denatured by boiling for 5 minutes in 1x SDS loading buffer immediately before loading. Gels were run in 1x running buffer at 180 V for 1-1.5 h according to protein size and percentage of resolving gel. The protein gel separated by SDS-PAGE was washed three times for 5 mins on an orbital shaker (SSM1) and stained with coomassie for an hour. Following coomassie staining, the gel was washed twice with dH_2O for an hour and analysed using a LAS-4000 imager (Fujifilm).

2.6. Protein Purification

2.6.1. Reagents, stock solutions and buffers for use in protein analysis

Wash Buffer: 20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 20mM imidazole.

Elution Buffer: 20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 200mM imidazole.

Desalting Buffer: 20mM, Tris-HCl pH7.8, 250mM NaCl, 5% glycerol, 1mM β-mercaptoethanol.

2.6.2. Purification of His-tagged proteins.

The conditions for expression were followed as mentioned in section 2.5.1, except the proteins were induced and expressed in 1L culture. The cultures re-suspended with lysis buffer were sonicated in (Qsonica XL2000) sonicator for five cycles at an amplitude of 40% and power being 30sec on/30sec off, followed by centrifugation step mentioned in section 2.5.2. The clear lysate obtained was purified immediately, or the lysate was stored at -80°C for future use.

His-tagged proteins were purified using a 1 ml bed volume of His Trap FF crude column (GE Healthcare). The column was washed with 5-10 column volumes of ddH2O, then equilibrated with 5-10 column volumes of wash buffer. The soluble lysate was then loaded on the column with a 0.5 ml/min flow rate. The column was washed with 20 column volumes of buffer A (Wash buffer), and proteins were eluted with buffer B (elution buffer). Five fractions of eluted proteins in 1 ml volume were collected. SDS-PAGE and coomassie staining was analysed for elution fractions.

2.6.3. Buffer exchange.

To reduce the salt concentration and remove imidazole from the purified proteins, the elution buffer (20mM Tris-HCl pH7.8, 300mM NaCl, 5% glycerol, 1mM β-mercaptoethanol, 200mM imidazole) was exchanged with desalting buffer (20mM Tris-HCl pH7.8, 250mM NaCl, 5% glycerol, 1mM β-mercaptoethanol) using PD–10 columns (GE Healthcare). The column was equilibrated with 25 ml of desalting buffer, and the flow-through was discarded. 2.5 ml of the purified protein was added to the equilibrated column. The proteins in the column are eluted by adding 3.5ml of desalting buffer (20mM Tris-HCl pH7.8, 250mM NaCl, 5% glycerol, 1mM β-mercaptoethanol). The eluted proteins were stored in aliquots (1000 µl) at -80°C.

2.6.4. Bradford assay.

Bradford assay was done to determine the concentration of purified proteins. Before use, the Bio-Rad protein reagent (Bio-Rad) was diluted at 1:5 in ddH2O. The protein standard curve was measured in the photometer (Eppendorf) by dissolving BSA in ddH$_2$O (Table.2-6). Protein samples were diluted with Bradford reagent, vortexed and incubated for 5 min at room temperature. Absorbance values were read at 600 nm, and the standard curve was obtained by linear regression analysis using GraphPad Prism 7.0 and used to interpolate the protein concentration of the samples.

Concentration (µg)	0	2	4	6	8	10
ddH2O (µl)	800	798	796	794	792	790
Elution Buffer (µl)	2	2	2	2	2	2
BSA (1mg/ml) (µl)	0	2	4	6	8	10
Bradford reagent (µl)	200	200	200	200	200	200

Table 2-6. Reagents required for Bradford assay. Regents were made at different concentrations and measured in a photometer to obtain a standard curve in measuring protein concentration.

2.7. RNA binding and deadenylation.

2.7.1. Reagents, stock solutions and buffers

5× reaction buffer: 100 mM Tris-HCl pH 7.8, 25% glycerol, 5 mM β-mercaptoethanol, 20 mM $MgCl_2$.

Fluorescent RNA substrate: (5′-FITC-AAAAAAAAAAAAAAAAAAAA-3′ Eurogentec)

RNA native dye: 0.05% bromophenol blue, 0.05% xylene cyanol, 25% glycerol and 1X TBE in ddH2O.

RNA loading buffer: 95% formamide, 0.025% SDS, 0.025% bromophenol blue, 0.025% xylene cyanol FF, 0.025% ethidium bromide, 0.5 mM EDTA (Deadenylation assay)

Stock solutions:

- 40% acrylamide: bisacrylamide (19:1) solution
- 10% ammonium persulphate (APS) in ddH2O.
- TEMED (Tetramethyl-ethylenediamine)

2.7.2. Electrophoretic mobility shift assay.

EMSA (Electrophoretic Mobility Shift Assay) was set up to confirm the binding of PABPC1/Variants 1-190 to synthesised RNA oligonucleotides (FLC-A_{20}) and to find the lowest excess of PABPC1/Variants needed to bind most of the RNA substrate in assay conditions. The reactions were set up in a six-microfuge tube to a final volume of 10 μl in each tube (Table.2-7). The concentration of proteins needed for the reaction was used in a four-fold increasing order multiplied by the concentration of RNA used for the reaction. 0.2μM of RNA was added to the reaction, and the proteins were added at an increasing rate of 0-20x, followed by the reaction buffer (100 mM Tris-HCl pH 7.8, 25% glycerol, 5 mM β-mercaptoethanol).

No non-specific RNA carrier was added. The complete shift of PABPC1/Variants in EMSA assay was achieved at room temperature of 21°C. The reaction was mixed gently and incubated at room temperature for one hour. After one hour of incubating PABPC1/Variants at room temperature with RNA, the samples were examined in RNA native gel.

Reagent	0	4x	8x	12x	16x	20x
5x Rxn Buffer	1	1	1	1	1	1
FLC-A$_{20}$ (1µM)	2	2	2	2	2	2
PABPC1/Variants(l00µM)	0	0.8	1.6	2.4	3.2	4
dH$_2$O	7	6.2	5.4	4.6	3.8	2

Table 2-7. EMSA Reaction. The table represents the reaction setup for the RNA binding assay. The reagent columns highlight the increased protein required to bind fixed RNA concentration.

2.7.2.1. RNA binding analysis.

Before running the samples on RNA native gel, 2µl of native loading dye (0.05% bromophenol blue, 0.05% xylene cyanol, 25% glycerol and 1X TBE in ddH2O) was added to 6µl of samples. 8% of RNA native gel was cast by adding 5ml of 40% bis-acrylamide (19:1), 1ml of 10x TBE, 10µl of TEMED and 100µl of APS (Table.2-8). The gels were pre-run at 4° to remove excess acrylamide from the well at 120V for 50 – 60 min. Following the pre-run, 4µl of samples are added to the wells and allowed to run for one hour at 120V. The binding reaction was set up at room temperature because the room temperature was adequate to attain a shift of PABPC1/Variants to RNA. At the same time, the gel is run at 4°C to minimise overheating. A Fujifilm LAS-4000 imager was used to analyse the gel using Epi-Blue illumination. Image analysis was carried out using ImageJ.

Reagent	Volume or weight
40% acrylamide: bisacrylamide	4 ml
10xTBE	1 ml
TEMED	10 µl
APS	100 µl

Table 2-8. EMSA Native gel. 15% RNA native gel composition for RNA binding assay.

2.7.3. Deadenylation assay with PABPC1 variants.

Deadenylation assay was conducted to study the poly(A) tail removal rate by CNO7/CAF1, with or independent of BTG2. On RNA bound PABPC1 variants. Before the deadenylation assay, RNA binding reaction was set up for PABPC1 variants, which was used as a substrate to study the rate of CNOT7/CAF1-BTG2 complex degradation compared to CNOT7/CAF1 on its own at various time intervals. RNA binding reaction was set up in the 500µl microfuge tube by incubating PABPC1/Variants with RNA. The final concentration of RNA was 0.2µM; the concentration of PABPC1 variants was 1.6µM (8-fold excess compared to the concentration of RNA in the reaction). Both RNA and PABPC1 were added to a reaction buffer (100 mM Tris-HCl pH 7.8, 25% glycerol, 5 mM β-mercaptoethanol, 20 mM MgCl2). The PABPC1-RNA were incubated at room temperature (21°C) for 40 mins simultaneously reactions were set up for deadenylation assay (Table.2-9). 1µl of RNA-PABP substrate from the binding reaction was added to 20mM $MgCl_2$, 0.3 µM enzymes and 1x reaction buffer at various time intervals (0, 10, 20, 40, 60 minutes) incubated at 30°C. Upon completion of each time course, the reactions were stopped by adding RNA loading dye and placed on ice blocks. The samples were heated for 3mins at 85°C and subjected to denaturing gel electrophoresis.

Reagent	0mins	10mins	20mins	40mins	60mins
1x Rxn Buffer	7	7	7	7	7
MgCl$_2$ (20mM)	1	1	1	1	1
RNA binding	1	1	1	1	1
CNOT7/CAF1 (3µM)	1	1	1	1	1
CNOT7/CAF1-BTG2 (3µM)	1	1	1	1	1

Table 2-9. The deadenylation assay reaction. The reaction was set up for deadenylation assay at a five-time interval.

2.7.3.1. Denaturing polyacrylamide gel electrophoresis.

Before running the samples on urea gel, 20% of denaturing gel was cast by adding 5ml of 40% bis-acrylamide (19:1), 1ml of 10x TBE, 5g of urea, 10µl of TEMED, 100µl of APS and filled with dH$_2$O to a final volume 10 ml (Table 2-10). The gel was pre-run in an X-Cell Sure Lock Mini Cell system (Invitrogen) at 300 V for 30 min. After the pre-run, the wells were thoroughly washed with 1x TBE to remove excess acrylamide and urea settled in the bottom of the well. The gel was run for 20-25 min at 300V. The gel was analysed by Fujifilm LAS-4000 using the Epi-Blue Illuminator. ImageJ was used for image analysis. Densitometry analysis was performed using the 'Multiplot' function of the ROI manager. Data were exported to GraphPad Prism 7.0 for graphical analysis.

Reagent	Volume or weight
40% acrylamide: bisacrylamide	5 ml
10xTBE	2 ml
UREA	5 g
TEMED	10 µl
APS	100 µl

Table 2-10. Urea gel set up for deadenylation assay. The 20% denaturing gel composition for deadenylation assays.

Chapter 3 – Computational modelling and structure analysis

3.1. PABPC1-BTG2 model dimer and RNA-PABPC1-BTG2-CNOT7/CAF1 quaternary complex

Poly(A) specific nucleases regulate poly(A) tail removal at the 3′ end of the mRNA by binding to PABPC1. Among the binding partners of PABPC1 are the BTG/TOB family of proteins. BTG/TOB are anti-proliferative proteins that regulate cell cycle progression in various cell types. In the BTG/TOB family of proteins, BTG1 and BTG2 resemble each other, specifically in three regions: Box A, Box B, and Box C. Box C of BTG2 are next to the BTG domain, which mediates interaction with RRM1 of PABPC1 (Yang et al., 2008, Stupfler et al., 2016). Box C is not conserved in TOB1 and TOB2, and, by contrast, PAM motifs in the long C-terminal regions of TOB1/TOB2 mediates interactions with PABPC1 (Winkler, 2010, Ezzeddine et al., 2012, Ezzeddine et al., 2007).

A study revealed that BTG2 binding sites for CNOT7/CAF1 are in the Box A (Y50–N71) and Box B (L96 – E115) region, while Box C (117-126) residues mediate BTG2 interaction with PABPC1 RRM1 (Yang et al., 2008, Doidge et al., 2012, Prevot et al., 2001). The function of BTG2 is to mediate deadenylation by bridging PABPC1 and CNOT7/CAF1, thus exposing the 3′ end of the mRNA poly(A) tail to the catalytic site in CNOT7/CAF1 (Stupfler et al., 2016).

A recent study (Stupfler et al., 2016) revealed Box C of BTG2 interacting with PABPC1 RRM1 (1-99), but the residues in RRM1 of PABPC1, responsible for interfacing residues in Box C of BTG1/BTG2, are yet to be discovered. Since no crystal structure reveals the details of the PABPC1-BTG2 interaction, it is challenging to identify the residues in RRM1 of PABPC1 interfacing with residues in Box C of BTG2.

The experiments described in this chapter aimed to identify possible amino acid residues in PABPC1, which are responsible for interfacing with residues in Box C of BTG2. Therefore, a bioinformatics approach was used based on structural information of PABPC1 and BTG2: (a) the crystal structure of RRM1 and RRM2 of human PABPC1 (Figure 3-1A) bound to poly(A) RNA and a peptide of translation initiation factor eIF4G (Safaee et al., 2012); and (b) the crystal structure of human BTG2 (Figure 3-1B) including the BTG domain and the Box C region (Yang et al., 2008).

The algorithm used predicted PABPC1 amino acid residues; the predicted residues interfacing RNA and translation factor bound to RRM2 of PABPC1 were neglected by analysing the Protein Interfacing Structural Assembly (PISA) tool. Finally, a possible model of the PABPC1-BTG2 complex was created in HADDOCK, a docking algorithm using CPORT that predicted PABPC1 residues free of RNA interfaces and residues in Box C of BTG2 as data.
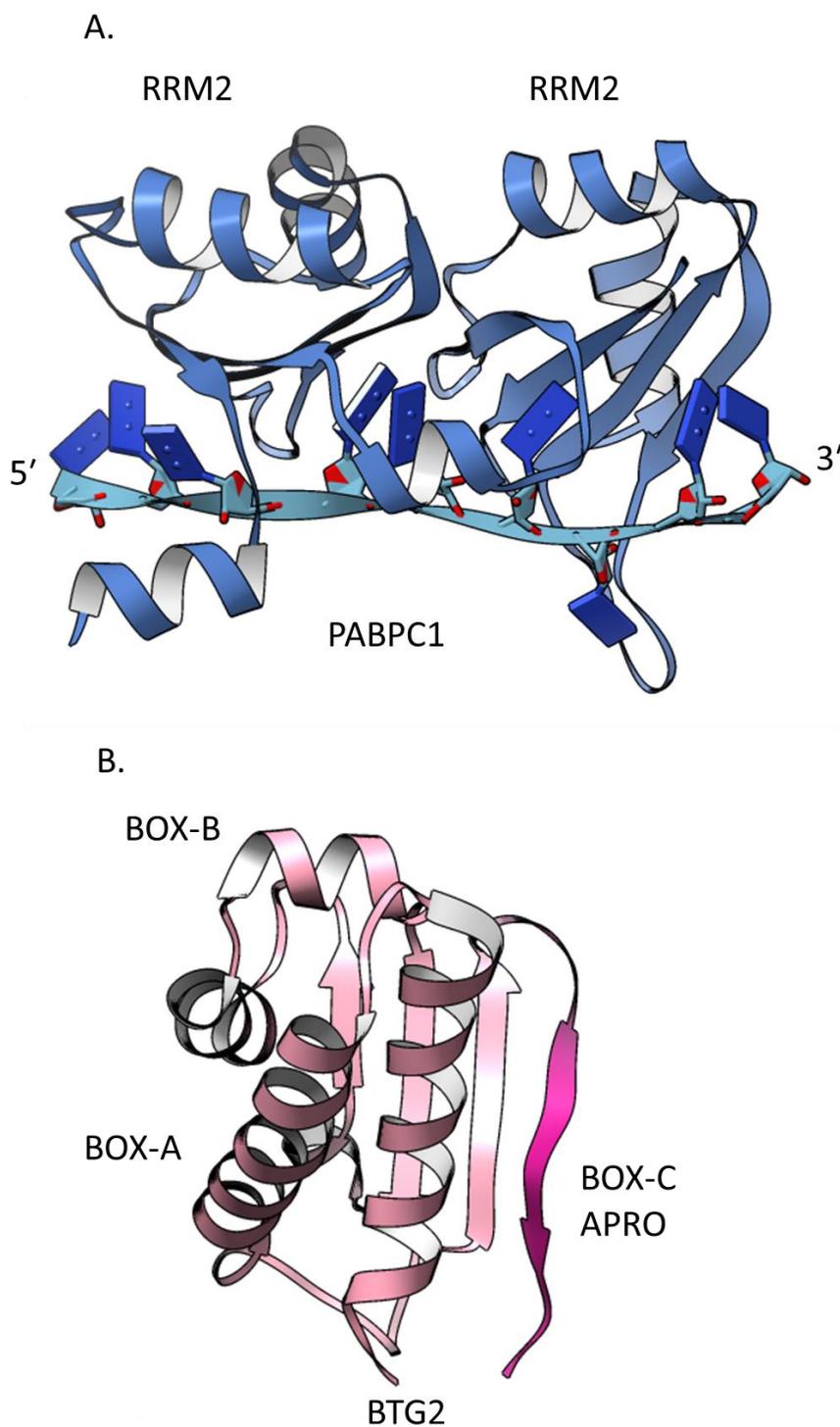
Figure 3-1. Structures of PABPC1 and BTG2. (**A**) Structure of RRM1 and RRM2 of PABPC1 with bound poly(A) RNA and translation initiation factor eIF4G bound to RRM2 (PDB 4F02, Safaee et al., 2012). (**B**) Structure of BTG2 (PDB 3DJU) with Box A (Purple), Box B (Deep Pink), and Box C (Orchid, Yang et al., 2008).

3.1.1. Prediction of macromolecular interface residues of PABPC1

The residues mediating the macromolecular interface was predicted by the CPORT method, which was applied to identify conserved surface residues of PABPC1 that are predicted to be at macromolecular interfaces. The process predicts residues by combining six predictors of macromolecular interfaces (de Vries and Bonvin, 2011b). These algorithms predict interfacing residues; for instance, the ProMate prediction process is based on three stages. The first stage involves analysing the proteins to produce interface and surface histograms relevant to structural and chemical properties. The second stage requires the estimated interface probability of each residue of the protein, and the third stage involves neighbouring residues grouped in predicted interfaces (Neuvirth et al., 2004). SPPIDER, residues prediction, focuses on relative solvent accessibility of the amino acid (Porollo and Meller, 2007). The projections by PINUP are based on empirical score function by a linear combination of energy score, the propensity of the interface, and the residue conservation score (Liang et al., 2006). The WHISCY prediction targets the conservation of amino acids in proteins (de Vries et al., 2006). The algorithm was run on the CPORT web server (https://alcazar.science.uu.nl/services/CPORT/) using the 'Very sensitive' threshold setting and associated structure HSSP sequence alignment of PABPC1 as input (4F02, chain A). The conserved residues predicted by each predictor are shown in (Table 3-1)

A graphical representation of the predicted interfacing residues mapped onto the structure of PABPC1, based on the six interface prediction algorithms combined by CPORT, is shown in Figure 3-2A; ProMate is shown in Figure 3-2B, SPPIDER in Figure 3-2C, PINUP in Figure 3-2D, and WHISCY in Figure 3-2E. As shown in Figure 3-2, CPORT was able to identify the interface with poly(A) even though the individual prediction algorithms were unable to identify the complete RNA interface. In addition, CPORT identified another region at the tip of RRM1 with potential involvement in macromolecular interactions. The known binding region to eIF4G was not identified by any algorithm, thus indicating the limitation of the approach.

Residue	Position	WHISCY	Pro Mate	SPPIDER	PINUP	CPORT
Tyr	14	-	x	-	x	x
Asp	17	-	-	-	x	x
His	19	-	-	-	x	x
Asp	21	-	-	-	x	x
Thr	23	-	-	-	x	x
Glu	24	-	-	-	x	x
Ile	40	x	-	-	x	x
Arg	41	x	-	-	x	x
Cys	43	x	-	-	x	x
Arg	44	x	-	x	x	x
Asp	45	x	-	x	x	x
Met	46	x	-	x	x	x
Ile	47	x	-	x	x	x
Thr	48	x	-	x	x	x
Arg	49	x	-	-	x	x
Arg	50	x	-	-	-	x
Leu	52	x	-	-	x	x
Gly	53	x	-	-	x	x
Tyr	54	x	-	-	x	x
Tyr	56	x	x	-	x	x
Leu	69	-	x	-	-	x
Arg	83	-	x	-	-	x
Trp	86	-	x	-	-	x
Ser	87	-	x	-	x	x
Arg	89	-	-	-	x	x
Asn	100	-	x	-	-	x
Phe	102	-	x	-	-	x
Lys	104	-	x	-	-	x
Tyr	116	-	x	-	-	x
Ile	125	-	x	-	-	x
Leu	126	-	x	-	-	x
Ser	127	-	x	-	-	x
Lys	129	-	x	-	-	x
Val	131	-	x	-	-	x
Gly	140	-	x	-	-	x
Val	142	-	x	-	-	x
Gly	171	-	x	-	-	x
Phe	173	-	x	-	-	x

Table 3-1 Predicted interfacing residues in RRM1-RRM2 of PABPC1. Indicated are interfacing residues of RRM1-RRM2 of PABPC1 predicted by WHISCY (de Vries et al., 2006), ProMate (Neuvirth et al., 2004), SSPIDER (Porollo and Meller, 2007), and PINUP (Liang et al., 2006). The residues determined by CPORT were obtained using the 'very sensitive' setting (de Vries and Bonvin, 2011b). The PABPC1 structure with RNA bound to RRM1-RRM2 and translation initiation factor eIF4G bound to RRM2 (PDB 4F02) was used to predict residues interfacing with BTG2 (Stupfler et al., 2016). The residues predicted by predictors are marked as (X), the position of residues and reference to their three-letter code of amino acids are indicated.
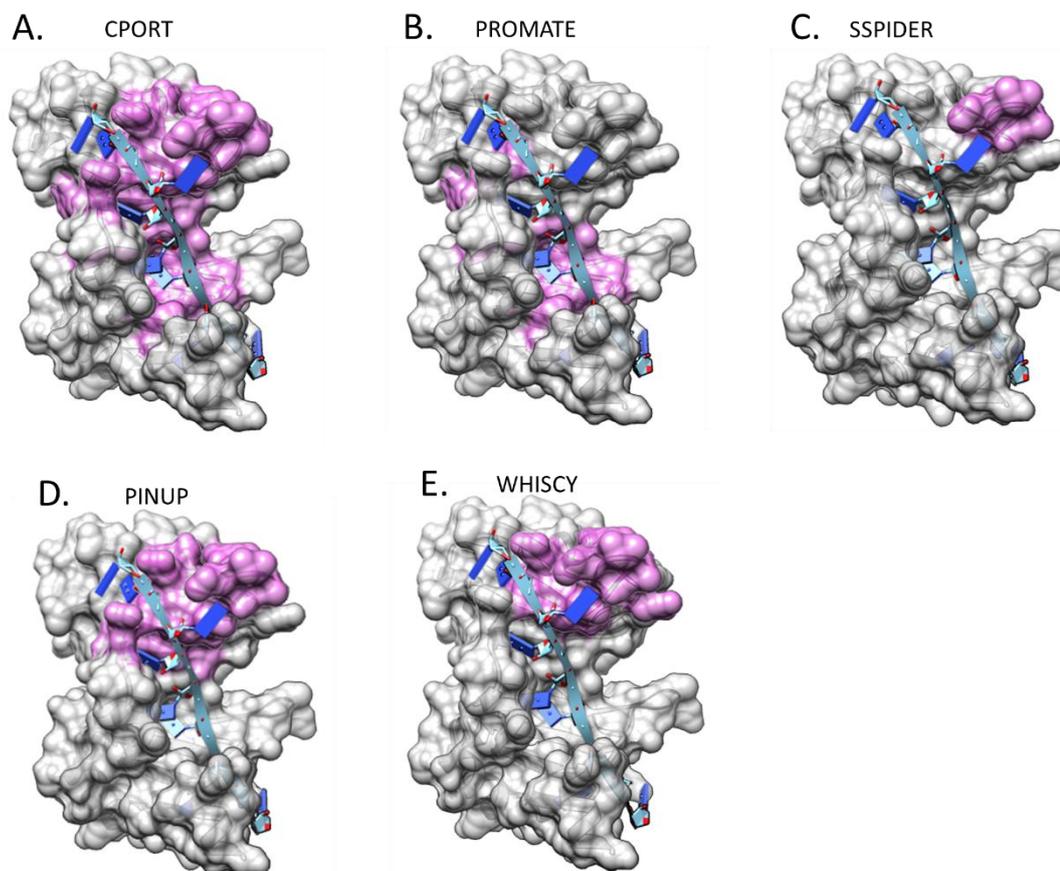
Figure 3-2. Prediction of interfacing residues of PABPC1. The crystal structure of PABPC1 (PDB 4F02, Safaee et al., 2012) from PDB was used in the CPORT algorithm to predict conserved surface residue of PABPC1 that are part of macromolecular interfaces. The conserved surface residues identified by CPORT and other predictors are highlighted (plum). **A)** Surface conserved residues identified by CPORT. **B)** Residues predicted by ProMate. **C)** Prediction by SPPIDER based on the relative solvent accessibility, **D)** Prediction by PINUP based on the linear combination of energy score, and the propensity of the interface **E)** WHISCY Prediction by conservation of amino acids and structural information of proteins.

3.1.2. Identification of residues of PABPC1 interfacing RNA

The region encompassing RRM1 and RRM2 of PABPC1 is bound to poly(A) RNA, and translation factor eIF4G is bound to RRM2 of PABPC1. Based on the model proposed (Stupfler et al., 2016), PABPC1, BTG2, and poly(A) RNA can form a ternary complex. In addition, the same study revealed that residues in RRM1 of PABPC1 mediate interactions with BTG2.

The binding site for eIF4G to PABPC1 is in RRM2 of PABPC1 (Safaee et al., 2012). Because RRM1 of PABPC1 is sufficient for binding BTG2 (Stupfler et al., 2016), it is unlikely that binding of eIF4G to RRM2 of PABPC1 interferes with the binding of BTG2 to RRM1 of PABPC1. To identify the specific amino acid residues of PABPC1 that interact with poly(A) RNA and eIF4G, the PISA tool was used, which was accessed through the webserver of the European Bioinformatics Institute (https://www.ebi.ac.uk/msd-srv/prot_int/pistart.html). A summary of the analysis is shown in Table 3-2.

A comparison of the predicted and calculated interfacing residues showed that most RNA interfacing residues from RRM1 of PABPC1 were identified, while the residues interfacing with eIF4G were not predicted by CPORT. In addition to residues in RRM1 (amino acids 1–99) of PABPC1 (Figure 3-3A, Yellow), a limited number of amino acids were predicted to be interfacing residues (Figure 3-3B, Green). No interaction partner is known for this set of residues, including His19, Asp21, Thr23, Glu24, Ile47, Thr48, Arg49, Arg50, Tyr56, and Leu69 from RRM1 of PABPC1. Thus, these residues could mediate the interaction with the Box C region of BTG2. Therefore, these residues were selected for information-based docking in HADDOCK to obtain a possible computational model of PABPC1-BTG2. In addition to identifying residues for docking in HADDOCK, few residues interfacing RNA were missed by CPORT but were identified in PISA (Table 3-3, Figure. 3-3C, Orange).

Residues	Position	CPORT Prediction	RNA Interfaces	Residues for Docking
Tyr	14	x	x	-
Asp	17	x	x	-
His	19	x	-	x
Asp	21	x	-	x
Thr	23	x	-	x
Glu	24	x	-	x
Ile	40	x	x	-
Arg	41	x	x	-
Cys	43	x	x	-
Arg	44	x	x	-
Asp	45	x	x	-
Met	46	x	x	-
Ile	47	x	-	x
Thr	48	x	-	x
Arg	49	x	-	x
Arg	50	x	-	x
Leu	52	x	x	-
Gly	53	x	x	-
Tyr	54	x	x	-
Tyr	56	x	-	x
Leu	69	x	-	x
Arg	83	x	x	-
Trp	86	x	x	-
Ser	87	x	x	-
Arg	89	x	x	-
Asn	100	x	x	-
Phe	102	x	x	-
Lys	104	x	x	-
Tyr	116	x	-	-
Ile	125	x	-	-
Leu	126	x	x	-
Ser	127	x	x	-
Lys	129	x	x	-
Val	131	x	-	-
Gly	140	x	x	-
Val	142	x	-	-
Gly	171	x	-	-
Phe	173	x	-	-

Table 3-2. PISA analysis for RNA interfacing residues. The table contains a list of PABPC1 residues predicted by CPORT to interface BTG2 (x), followed by predicted residues interfacing RNA (x), and residues that were free of RNA interfaces were selected for docking (x).

Residues	Position	CPORT Prediction	Residues for Docking	RNA interfaces not predicted
Ser	12	-	-	x
Ser	39	-	-	x
Asn	58	-	-	x
Met	85	-	-	x
Gln	88	-	-	x
Pro	91	-	-	x
Arg	94	-	-	x
Asn	105	-	-	x
Cys	128	-	-	x
Lys	138	-	-	x
Gly	139	-	-	x
His	144	-	-	x
Phe	169	-	-	x
Arg	172	-	-	x
Lys	174	-	-	x
Ser	175	-	-	x
Arg	176	-	-	x

Table 3-3. Residue interfacing RNA not predicted by CPORT. The table contains a list of PABPC1 residues interfacing RNA but was not predicted by CPORT analysis (x).
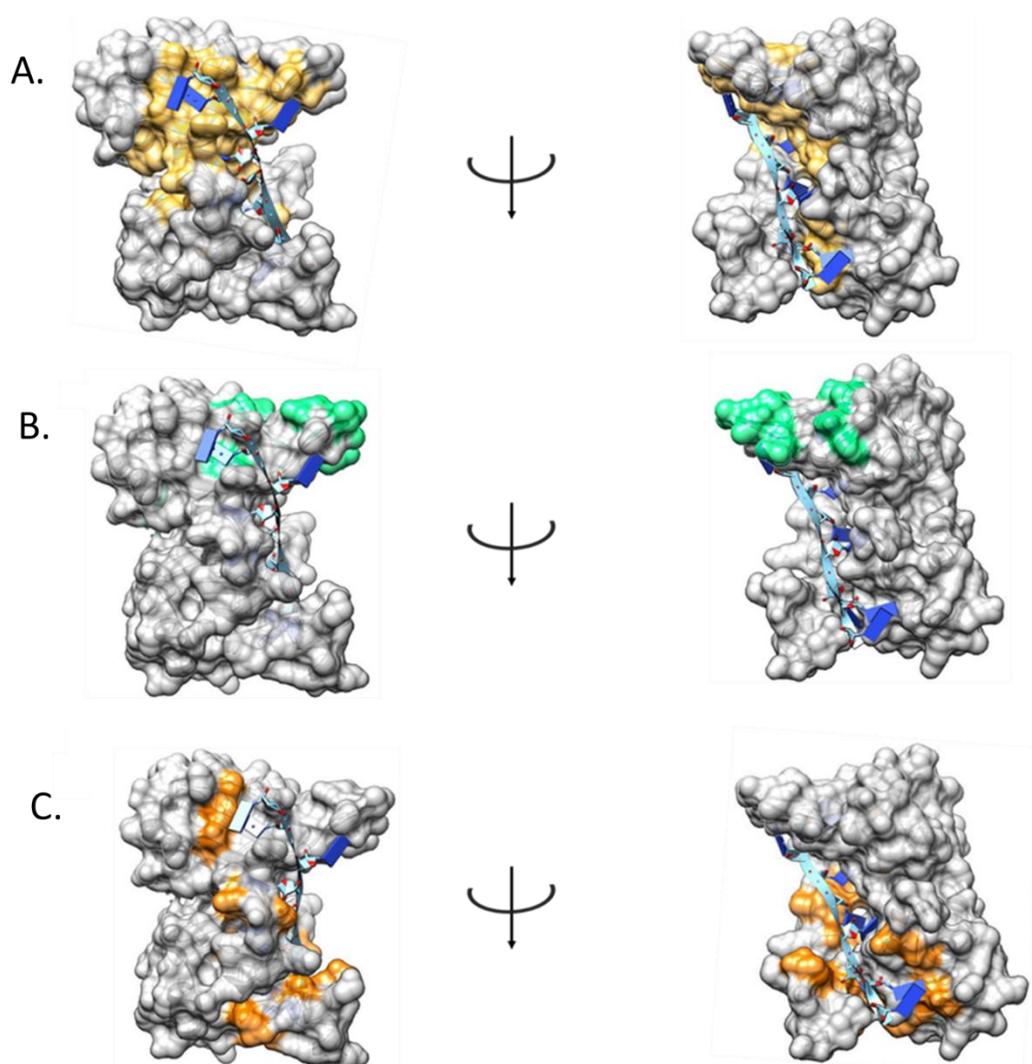
Figure 3-3. Identification of possible residues of PABPC1 interfacing with BTG2. Predicted interfacing residues of PABPC1 compared to residue involved in interactions with RNA. **(A)** Predicted and calculated RNA interfacing residues (yellow) of PABPC1 RRM1. **(B)** Identified residues in RRM1 of PABPC1 that were away from known macromolecular interfaces were selected for docking. **(C)** Residues in RRM1 and RRM2 interfacing with RNA were not predicted by CPORT but identified by PISA (Orange).

3.1.3. Generation of a computational model of the PABPC1-BTG2 complex

The ten predicted interfacing residues, His19, Asp21, Thr23, Glu24, Ile47, Thr48, Arg49, Arg50, Tyr56, and Leu69, fulfil the criteria for interacting Box C of BTG2 since they are (a) located in RRM1 of PABPC1, and (b) not overlapping with residues of PABPC1 involved in binding poly(A) RNA. To extend the computational approach, the predicted interaction residues of PABPC1 and the Box C residues of BTG2 (116–127) were used as restraints to model the PABPC1-BTG2 complex using the HADDOCK docking algorithm.

Overall, HADDOCK generated 178 solutions into nine clusters. A graph was plotted based on the score and size of solutions in clusters using default settings (Figure 3-4). After discarding clusters with fewer than ten solutions, the highest-scoring solutions of four clusters—cluster 1 (Figure 3-5A), cluster 2 (Figure 3-5B), cluster 3 (Figure 3-5C), and cluster 4 (Figure 3-5D)—were chosen for further structural analysis.
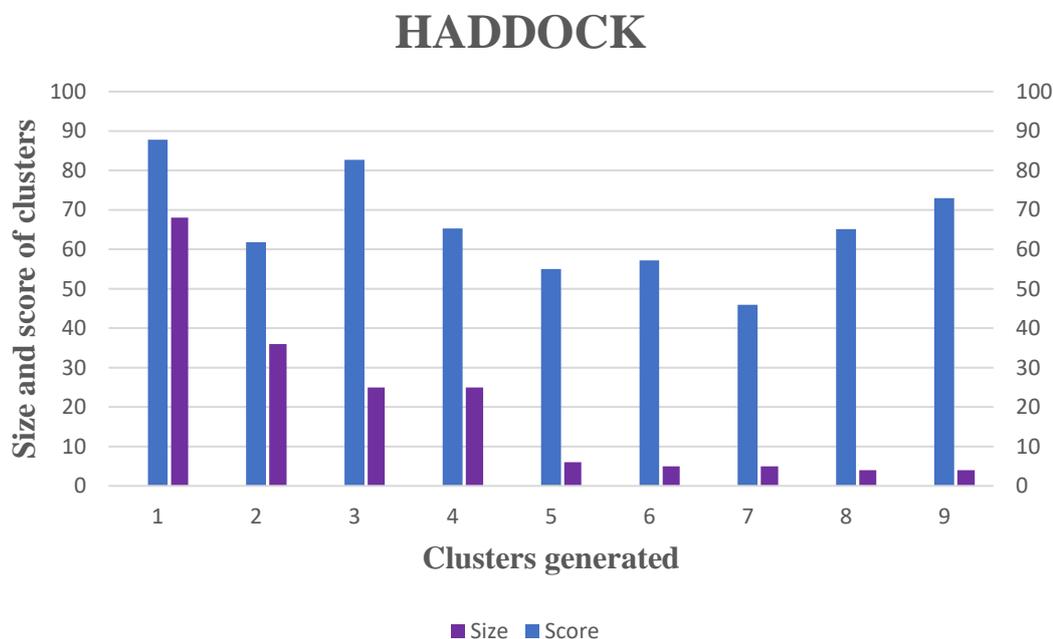
Figure 3-4. Overview of HADDOCK using the predicted interfacing residues of PABPC1 and BTG2. In total, 178 solutions formed 9 clusters generated by HADDOCK; the clusters with the best score and size are clusters 1-4.
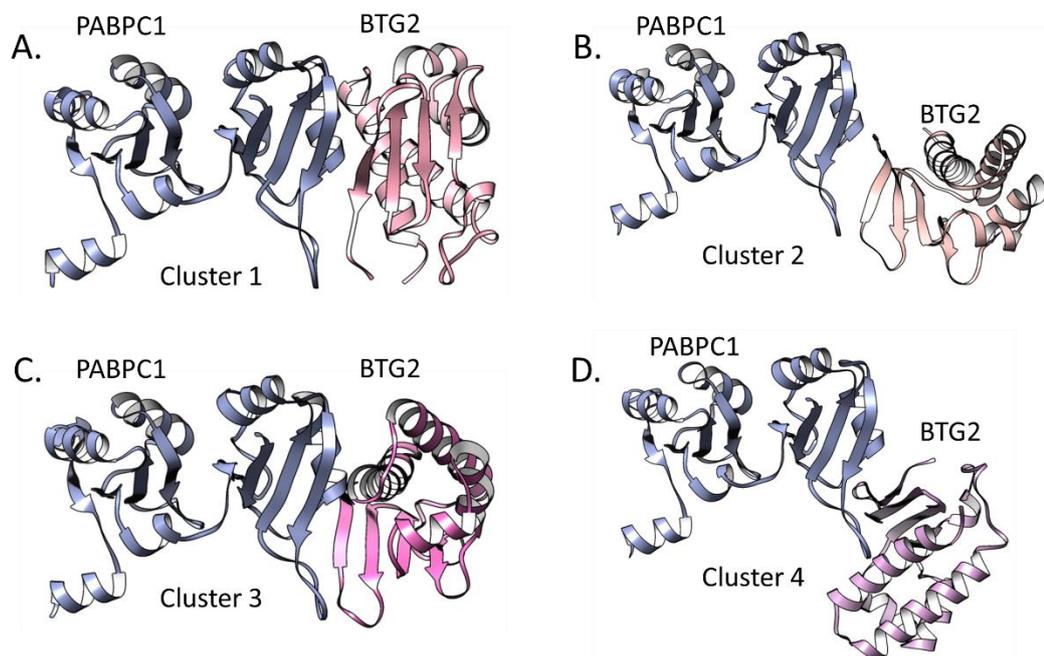
Figure 3-5. Representation of highest scoring solutions of clusters 1-4. Shown are PABPC1 RRM1-RRM2 (blue) and BTG2 (pink) interactions. A computational model generated by docking in HADDOCK. **A)** Cluster1, **B)** Cluster2, **C)** Cluster3 and **D)** Cluster 4.

3.1.4. Generation of models for quaternary poly(A)-PABPC1-BTG2-CNOT7/CAF1 complexes

Since it is proposed that the BTG2 protein can bridge the interaction between PABPC1 and CNOT7/CAF1, thereby promoting the deadenylation of poly(A) RNA, possible models of the quaternary poly(A)-PABPC1-BTG2-CNOT7/CAF1 complex were built next. These models of the quaternary complex may result in a better interpretation of RNA accessibility to the CNOT7/CAF1 and its distance from the active site. The quaternary complex was built from PDB structures of PABPC1 (PDB 4F02, Safaee et al., 2012), BTG2 (PDB 3DJU, Yang et al., 2008), CNOT7/CAF1-TOB1 (PDB 2D5R, Horiuchi et al., 2009b), and the computational model of the PABPC1-BTG2 complex obtained by HADDOCK.

The highest scoring solution of clusters 1–4 was used to initiate the models for the quaternary complexes. First, the structure of PABPC1 (RRM1-RRM2) and poly(A) RNA comprising nine adenosine residues (PDB 4F02) were superimposed on the highest scoring solution of cluster 1. Next, the Tob-CNOT7/CAF1 2D5R) structure was superimposed, taking advantage of the high structural similarity between BTG2 and TOB1 (Figure 3-6A). The same steps were carried out using the highest scoring solutions of clusters 2–4 (Figure 3-6B-D).

The RNA-PABPC1-BTG2-CNOT7/CAF1 quaternary complex created from cluster 1 revealed that BTG2 bridges PABPC1 and CNOT7/CAF1, thereby making the 3′ end (A) RNA accessible to the active site of CNOT7/CAF1. The quaternary complex created for cluster 2 and cluster 3 revealed that the RNA bound to PABPC1 was away from the CNOT7/CAF1 active site. Further analysis of the created quaternary complex for cluster 4 revealed that the active site of CNOT7/CAF1 was pointing in the opposite direction of the 3' end of the poly(A) RNA, thereby making the RNA inaccessible to the active site of CNOT7/CAF1. In contrast to clusters 2, 3 and 4, the quaternary complex created from cluster 1 provided a rationale for the stimulatory effect of deadenylation by CNOT7/CAF1 in the presence of both BTG2 and PABPC1.
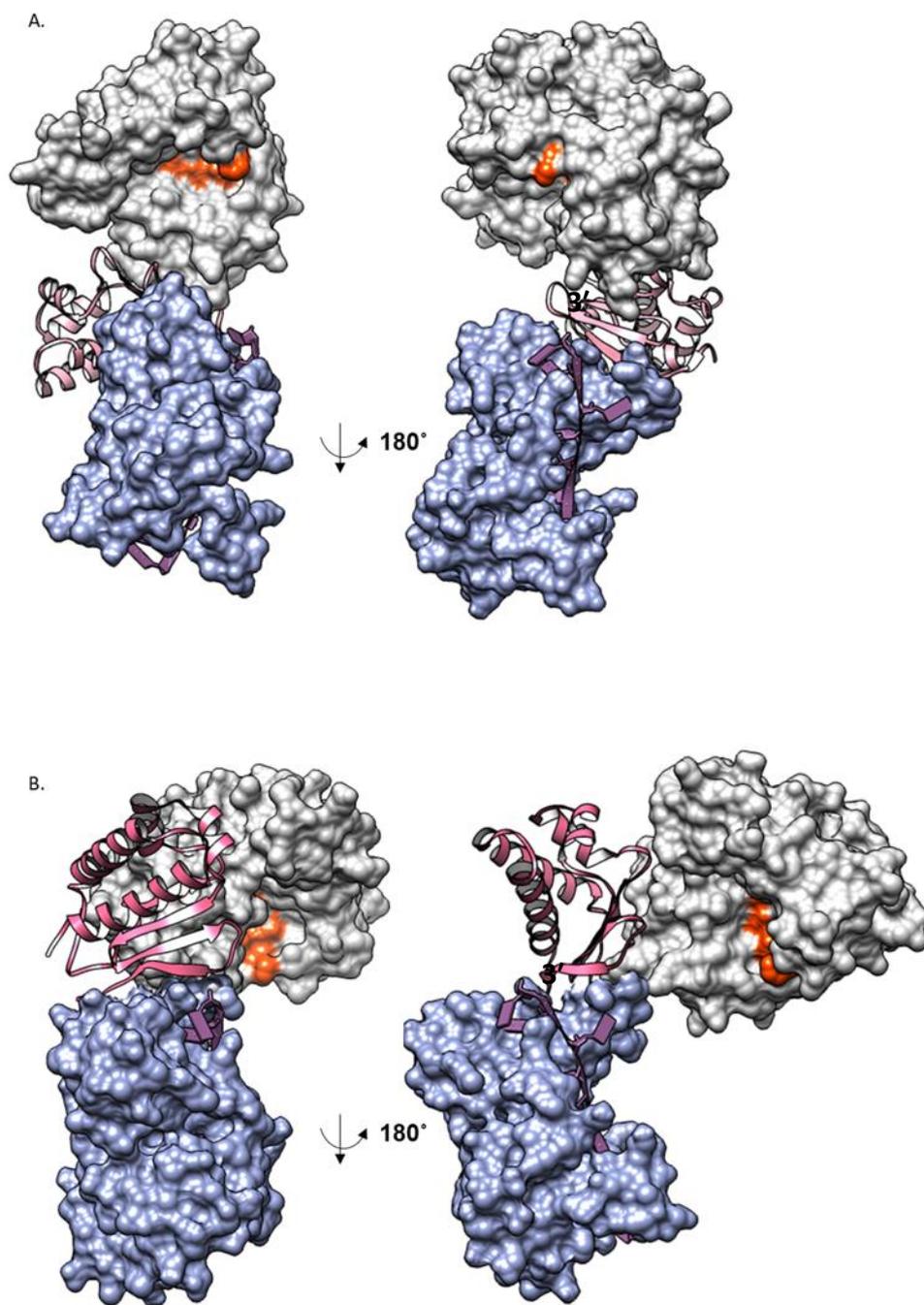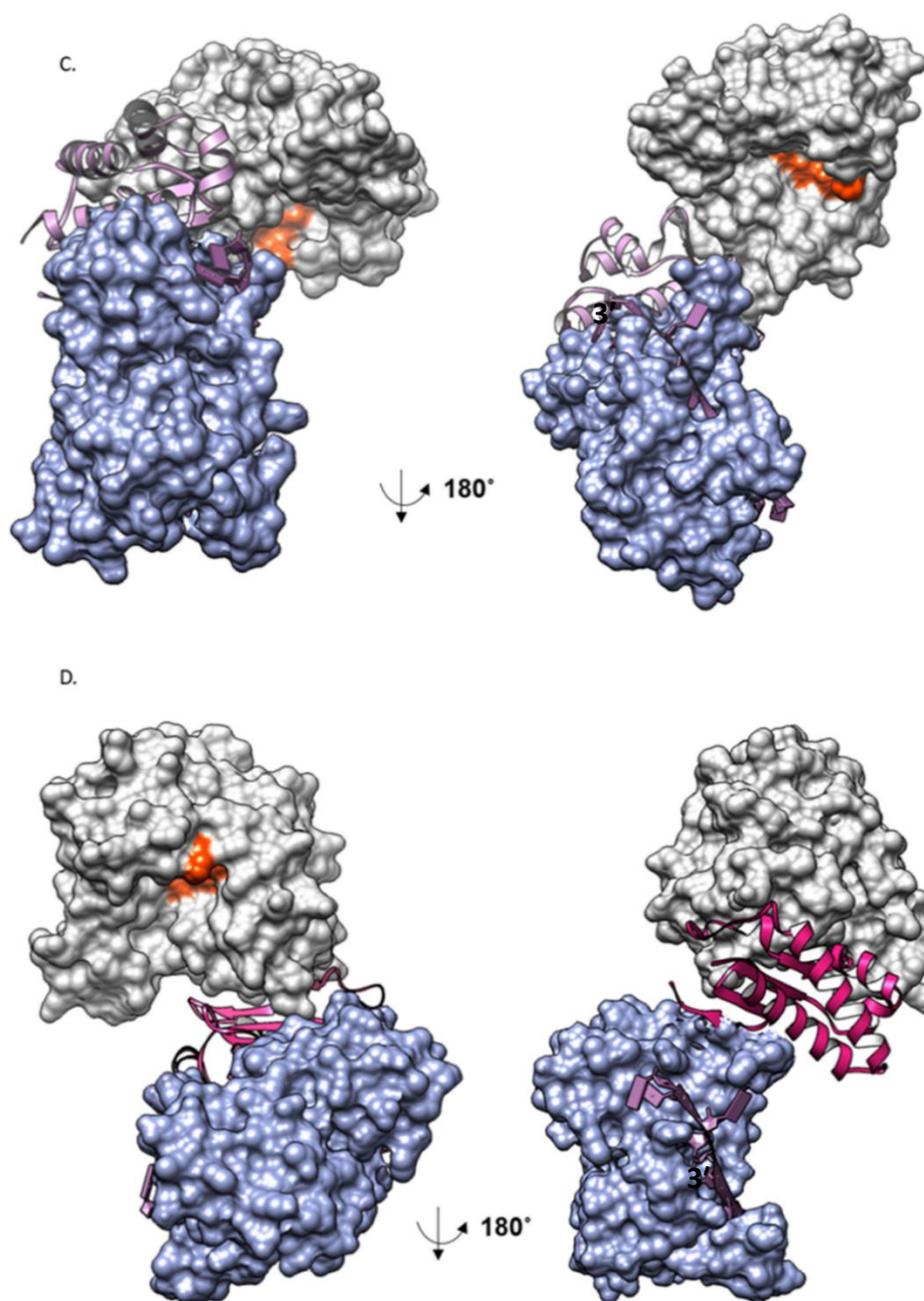
Figure 3-6. Models of quaternary poly(A)-PABPC1-BTG2-CNOT7/CAF1 complexes. A trimeric complex was created with the PABPC1-BTG2 model. **A)** Cluster 1, **B)** Cluster 2, **C)** Cluster 3 and **D)** Cluster 4. The created model depicts BTG2 (Pink ribbon) bridging between PABPC1 (Blue surface, CNOT7/CAF1 (Grey surface) and reveals RNA (Violet ribbon) accessibility to the active site of CNOT7/CAF1 (Orange). The model was superposed to the previously obtained structure of PABPC1 (4F02), BTG2 (3DJU), TOB-CNOT7/CAF1 (2D5R) and NOT1-CNOT7/CAF1 (4GMJ).

3.1.5. Selecting interfacing residues in RRM1 of PABPC1 for site-directed mutagenesis

Based on the observation that residues at macromolecular interfaces can be predicted based on docking solutions (Fernandez-Recio et al., 2004), interfacing residues in the highest-scoring solutions of clusters 1–4 were analysed next. This was achieved using PISA, which was accessed through the webserver of the European Bioinformatics Institute (https://www.ebi.ac.uk/msd-srv/prot_int/pistart.html).

Analysing the solutions of clusters in PISA, the predicted PABPC1 residues interfacing with BTG2 are shown in Table 3-3. Examining the solutions of cluster 1 in PISA revealed a possible role for residues Glu-24, Arg-44, Arg-49, Ile-40, and Arg-50; cluster 2 revealed Ala-26 and Met-46; cluster 3 revealed Asp-21, Met-26, Thr-28, Glu-29, Ile-47, and Thr-48; and cluster 4 revealed Arg-41 of PABPC1 as a possible residue interfacing BTG2 (Table 3.3).

Moreover, analysing the solutions revealed residues in Box C of BTG2, such as Ala-116, Ile-119, Cys-120, Glu-124, and Glu-125 from solutions cluster 1; cluster 3 revealed Gly-117, Ser-118, Cys-120, Val-121, Leu-122, and Tyr-123. The residues revealed by cluster 2 and cluster 4 are found to be identical in other clusters. Based on this analysis, the PABPC1 residues identified in two or more clusters are Asp-21, Glu-24, Ile-40, Arg-44, Arg-49, Arg-50, which were prioritised for site-directed mutagenesis. In addition, Glu-29 and Met-46 were also selected. The structure (Figure 3-7) shows a graphical representation of the selected residues mapped onto the structure of RRM1-RRM2 of PABPC1 (Figure 3-7A) and residues from the Box C region of BTG2 (Figure 3-7B).

#	Amino acid (PABPC1)	Clusters				Amino acid (BTG2)	Clusters			
		1	2	3	4		1	2	3	4
1	Asp-21			x	x	Ala-116	x			
2	Glu-24	x	x	x	x	Gly-117			x	
3	Ala-25		x			Ser-118			x	
4	Met-26			x		Ile-119	x	x		x
5	Thr-28			x		Cys-120	x		x	x
6	Glu-29			x		Val-121			x	x
7	Ile-40	x		x		Leu-122			x	
8	Arg-41				x	Tyr-123			x	
9	Arg-44	x	x	x	x	Glu-124	x	x	x	x
10	Met 46		x			Glu-125	x			x
11	Ile-47			x		Ala-126				x
12	Thr-48			x		Pro-127				
13	Arg-49	x	x		x					
14	Arg-50	x		x						
15	Glu-60			x						
16	Leu-78			x						

Table 3-3. PABPC1 residues predicted to interface with BTG2 by analysing predicted interfacing residues. Predicted PABPC1 residues mapped to interface with BTG2 (Left) and residues in Box C of BTG2 identified to interface with RRM1 of PABPC1 (Right). The PABPC1 residues identified in clusters are marked (x) in the left column, whereas BTG2 domain residues in the Box C region of solutions of clusters are marked (x) in the right column. The residues were identified from analysing the solutions of clusters, and residues were selected based on analysis of interfacing residues of the highest-scoring solution of clusters 1-4 using the PISA tool.

Data driven docking
PABPC1-BTG2 model

Clusters 1-4
Solutions of Clusters

PISA
Examining the solutions

Mapping solutions in UCSF CHIMERA

Site-directed mutagenesis
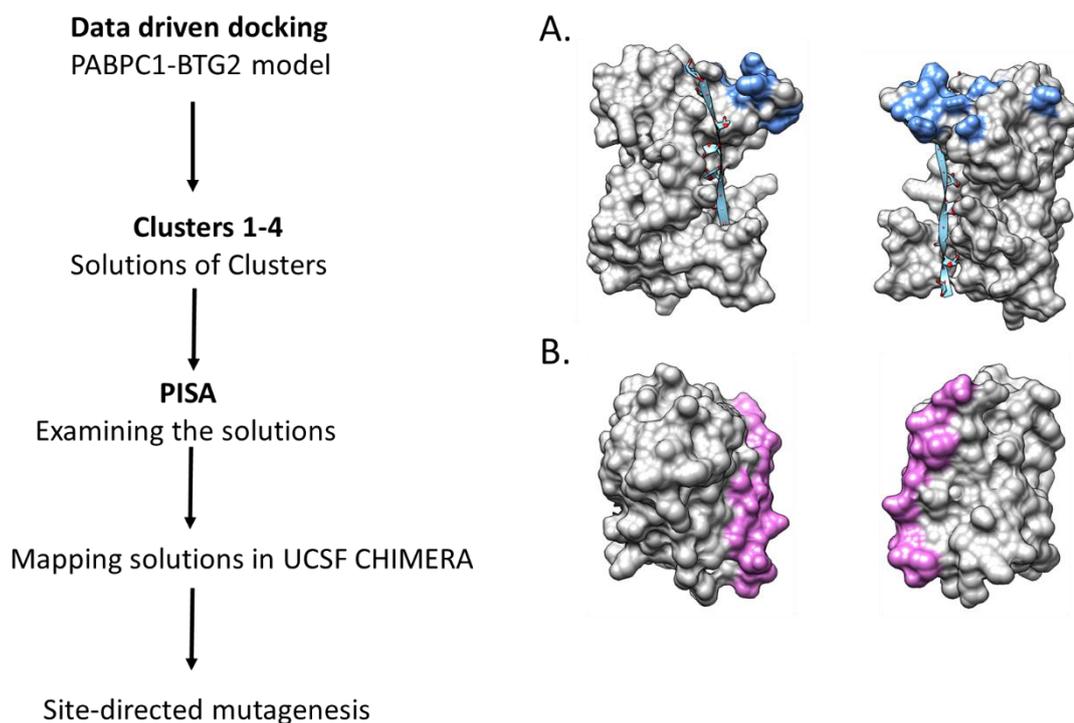
Figure 3-7 Selection of residues for mutagenesis. The residues were selected based on solutions examined in PISA for PABPC1 residue interfaces with BTG2. **A)** The residues of PABPC1 predicted to interface with Box C of BTG2 are highlighted in Blue. **B)** The surface highlighted in the Orchid contains residues in the Box C region of BTG2.

3.2. Discussion

3.2.1. Prediction of amino acid residues of PABPC1 involved in macromolecular interactions

Studies reveal residues in Box C of BTG2 interface with residues in RRM1 of PABPC1 (Stupfler et al., 2016). However, since no structure depicting PABPC1-BTG2 interactions is available, it is a challenge to identify PABPC1 residues that direct these interfaces. Therefore, CPORT was used, which is a bioinformatics tool that predicts surface conserved residues responsible for macromolecular interfaces. CPORT predicts interfacing residues by combining the input of four residue interface predictors: WHISCY (de Vries et al., 2006), ProMate (Neuvirth et al., 2004), SPPIDER (Porollo and Meller, 2007), and PINUP (Liang et al., 2006).

Though CPORT disclosed residues by combining the input of four residue interfaces, each interface uses a different approach to predict residues. The residues disclosed by an individual interface overlap with each other; for example, the SSPIDER interface predicts very few residues, but it largely overlaps with WHISCY and PINUP. While the residues predicted by WHISCY mostly overlap with PINUP, the residues predicted by PROMATE are largely different and primarily out of the target region. However, all four predictors revealed PABPC1 residues interfacing with RNA, but the residues interfacing with eIF4G were not disclosed. Selecting residues from one among the four predictors for docking could affect the PABPC1-BTG2 structure because the critical residues regulating interaction with BTG2 might be overlooked. All predictors revealing PABPC1 residues that can interface with BTG2 contribute to an analysis by CPORT. Therefore, CPORT integrates the residues revealed by individual predictors. The integrated residues in CPORT are further analysed for RNA interfaces, and the residues free of RNA interfaces are selected for docking.

3.2.2. Identifying residues of PABPC1 involved in BTG2 interaction

The PABPC1 residues predicted by CPORT revealed that RRM1 and RRM2 are involved in possible macromolecular interactions. A study (Stupfler et al., 2016) has revealed that RRM1 of PABPC1 mediates interaction with Box C of BTG2. Since the residues mediate the macromolecular interfaces between PABPC1 and BTG2 in RRM1 of PABPC1 within Box C of BTG2, the objective was to identify possible residues in RRM1, which are responsible for mediating interfaces. Therefore, residues in RRM2 of PABPC1 were excluded. The PABPC1 (PDB 4F02) structure was used to predict residues involved in macromolecular interfaces. The structure depicts two RRM domains with translation initiation factor (eIF4G) bound to RRM2. In addition to eIF4G being bound to RRM2, the structure also discloses that RNA is bound to RRM1 and RRM2. Because PABPC1 can form a complex with both RNA and BTG2 (Stupfler et al., 2016), PABPC1 residues involved in the interaction with BTG2 should not overlap with residues interfacing RNA. Thus, predicted interfacing PABPC1 residues in RRM1 that were not located on the RNA interface were chosen for information-driven docking.

3.2.3. Generation of computational models of the PABPC1-BTG2 complex

CPORT-predicted residues were selected as information for the HADDOCK docking algorithm. The information for HADDOCK was entered in the form of active and passive residues. The active residues in RRM1 of PABPC1 were predicted to interface with the Box C region of BTG2 by the CPORT algorithm and the HADDOCK program assigned passive residues. The selected active residues in RRM1 did not overlap with residues interacting with RNA. The HADDOCK protocol involves three stages: a rigid-body energy minimisation, a semi-flexible refinement in torsion angle space, and a final refinement in explicit solvent.

After each stage, the PABPC1-BTG2 model was obtained, generating nine clusters with 178 solutions. The structures were scored and ranked. The clusters with the best score and rank were used for further analysis; therefore, clusters 1–4 were high in score and size and were used for developing a quaternary complex of PABPC1 BTG2-CNOT7/CAF1.

3.2.4. Creation of a quaternary model of PABPC1-RNA-BTG2-CNOT7/CAF1

Docking by HADDOCK revealed a computational model of PABPC1-BTG2, and the generated model revealed the residues in RRM1 of PABPC1, guiding the interaction with Box C of BTG2. The structure of PABPC1 used for docking has RNA bound to RRM1-RRM2 and eIF4G in RRM2. Since the aim was only to study macromolecular interfaces between PABPC1 (RRM1) with Box C of BTG2, the eIF4G and RNA in the PABPC1 (4F02) structure were excluded before docking. Because the generated PABPC1-BTG2 model lacked RNA during docking, the quaternary complex of RNA-PABPC1-BTG2-CNOT7/CAF1 was created. This was done using UCSF Chimera (Pettersen et al., 2004) to study accessibility of the 3' end of RNA to the active site of CNOT7/CAF1 and the role of BTG2 in regulating the rate of deadenylation.

The quaternary complex was created by superposing the structural models of PABCP1-RNA and TOB-CNOT7/CAF1 complexes onto the PABPC1-BTG2 computational model. The quaternary complex that was created assisted in understanding BTG2 with the purpose of bridging PABPC1 and CNOT7/CAF1 as well as understanding its role in stimulating deadenylation by CNOT7/CAF1 in the presence of PABPC1 in a computational model.

Though the created quaternary structure using the PABPC1-BTG2 computational model disclosed the role of BTG2 in stimulating deadenylation by bridging CNOT7/CAF1 and PABPC1, it still requires experimental validation. However, the central aim pursued in this chapter is to identify possible amino acid residues in RRM1 of PABPC1 that is responsible for interfacing with residues in Box C of BTG2. The identified residues that were free of RNA interfaces and residues in Box C of BTG2 were selected for docking in HADDOCK, thereby creating a possible model of the PABPC1-BTG2 complex.

Chapter 4 – Site-directed Mutagenesis and Interaction studies

4.1. Identification of residues required for the interaction between BTG1/BTG2 and PABP

A study by (Stupfler et al., 2016) revealed that RRM1 of PABPC1 regulates interactions with residues in Box C of BTG1 and BTG2. Though the study showed that the RRM1 domain of PABPC1 is responsible for binding BTG2, the residues in RRM1 of PABPC1 that are responsible for interfacing with BTG2 were not determined. The experiments described in this chapter aimed to validate predicted interfacing residues. To this end, selected residues in PABPC1 were subjected to site-directed mutagenesis by substituting with Ala amino acid. In addition to amino acid substitutions in RRM1 of PABPC1, variants for BTG2 were created by deleting residues in Box C (ΔBox C; deletion of residues 116–127), as well as a variant containing a stop codon following Box C (ΔC-term, amino acids 1–127). Following the study by (Stupfler et al., 2016), the mutations introduced to Box C were based on substituting residues in Box C with TOB1 in the corresponding region.

To study PABPC1-BTG1 binding, a yeast two-hybrid system was initially used for detecting protein interactions. A LexA-based yeast two-hybrid assay performed by (Stupfler et al., 2016) revealed that RRM1 of PABPC1 interacted with the Box C region of BTG1 by monitoring β-galactosidase production. The yeast two-hybrid system is based on the principle that the transcriptional activator protein GAL4 has separate DNA binding and transcription activation domains (Fields and Song, 1989).

The LexA-based system that was used is based on the same principle, except that the DNA binding protein is derived from the *E. coli* LexA protein, and the activation domain is the acid blob domain (B42). Neither LexA nor B42 can activate transcription independently (Figure 4-1A). The yeast two-hybrid experiment requires protein of interest bound to both binding domain and activation domain to activate transcription; thus, the reporter can only detect gene expression by fusing the binding domain and activation domain with interacting proteins (Figure 4-1D; Golemis et al., 2001a). Therefore, transcription in yeast two-hybrid experiment cannot be activated with protein-bound only to the binding or activation domain (Figure 4-1B-C).

In addition to the yeast two-hybrid system, His-pulldown assays were performed using His tagged PABPC1 (amino acid residues 1–190) and GST-BTG2. Pulldown experiments were conducted using nickel-nitrilotriacetic acid (Ni^{2+}-NTA) beads, which bind to histidine residues through immobilised metal ions with a high binding affinity (Bornhorst & Falke, 2000). Pulldown experiments were used to identify the interfaces of predicted residues in RRM1 of PABPC1 with BTG2. The requirement for Box C residues of BTG2 was also defined in more detail. Furthermore, a cancer mutation T48P of PABPC1 was investigated from the catalogue of somatic mutations in cancer (COSMIC; Tate et al., 2019). T48P is a rare mutation of PABPC1, and it is the most common mutation in RRM1 of PABPC1. The amino acid substitution at position T48P is recurring in 26 thyroids from head and neck cancer samples (Martin et al., 2014). Since residue T48 is one among the residues predicted to interface with BTG2, PABPC1 variants T48A and T48P were analysed further.
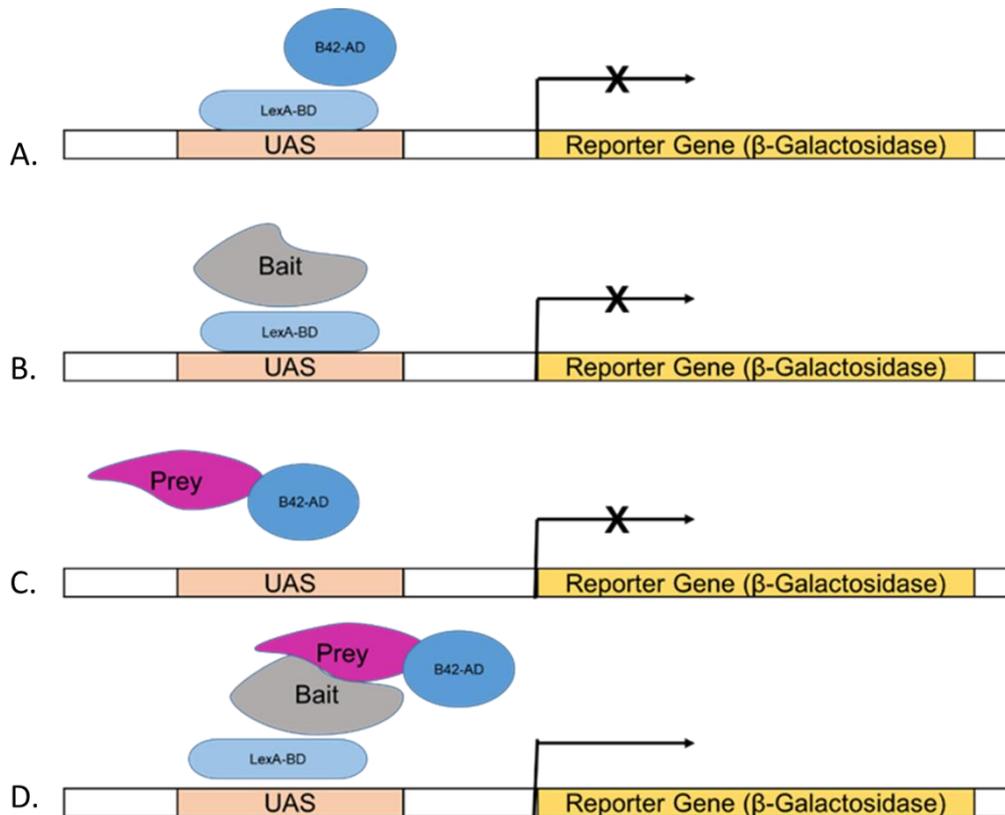
Figure 4-1. A schematic representation of the LexA/B42 yeast two-hybrid system. The model uses the EGY48 yeast strain with LexA binding domain and B42 activation domain. **A)** Transcription can neither be activated by binding nor activation on its own. **B)** Transcription cannot be activated with protein-bound only to the binding domain. **C)** Transcription cannot be activated by protein-bound only to the Activation domain. **D)** reporter detects gene expression only when binding domain and activation domain fused with interacting protein is in contact.

4.1.1. Evaluation of the yeast two-hybrid assay to analyse the PABPC1-BTG1 interaction

The yeast two-hybrid system was first evaluated to investigate the interactions between BTG1/BTG2 with PABPC1. The following plasmids were used: (a) BTG1 cloned into vector pEG202 (Figure 4-2A), expressing BTG1 as a fusion protein with the LexA DNA binding domain; (b) PABPC1 (1–146) and CNOT7/CAF1 cloned into vector pJG4-5 (Figure 4-2B), containing the B42 activation domain; and (c) plasmid pSH18-34, containing a β-galactosidase reporter under the control of LexA binding sites. Plasmids were transformed into yeast strain EGY48 using the LiAc method. The transformed yeast cells were grown in selective glucose media without uracil, tryptophan, and histidine.

Three independent yeast colonies were tested for β-galactosidase activity after induction of the B42 fusion proteins in media containing galactose. The β-galactosidase activity was tested for each yeast culture in duplicate and normalised using culture density using the OD_{600} value. The two negative controls used in the assay were: (a) empty vectors with no protein of interest fused to the binding and activation domain; and (b) BTG1 fused to the LexA DNA binding domain with no protein of interest fused to the activation domain. In both cases, no β-galactosidase activity was detected (Figure 4-3). As a positive control, BTG1 fused to the LexA DNA-binding domain was combined with CNOT7/CAF1 to the B42 activation domain induced β-galactosidase activity (Figure 4-3). However, when plasmids encoding LexA-BTG1 and B42-PABPC1 (1–146) were co-transformed with the reporter plasmid, no β-galactosidase activity was observed (Figure 4-3).

Co-transforming plasmids encoding LexA-BTG1 and B42-PABPC1 (1–146) with the reporter plasmid revealed colonies; however, no β-galactosidase activity was detected. By contrast, BTG1-CNOT7/CAF1 binding was detected, which was used as a positive control. Even though B42-PABPC1 RRM1-RRM2 presence was not confirmed by western blotting, the results suggest that selection occurred against the expression of B42-PABPC1 RRM1-RRM2. This may be related to the repression of endogenous PABP1, such as that reported for mammalian PABPC1, which contains an A-rich sequence in the 5' UTR (de Melo Neto et al., 1995).

The experiment was repeated three times, revealing the same results. Therefore, a pull-down assay was evaluated as an alternative method for studying the BTG1/BTG2-PABPC1 interactions.
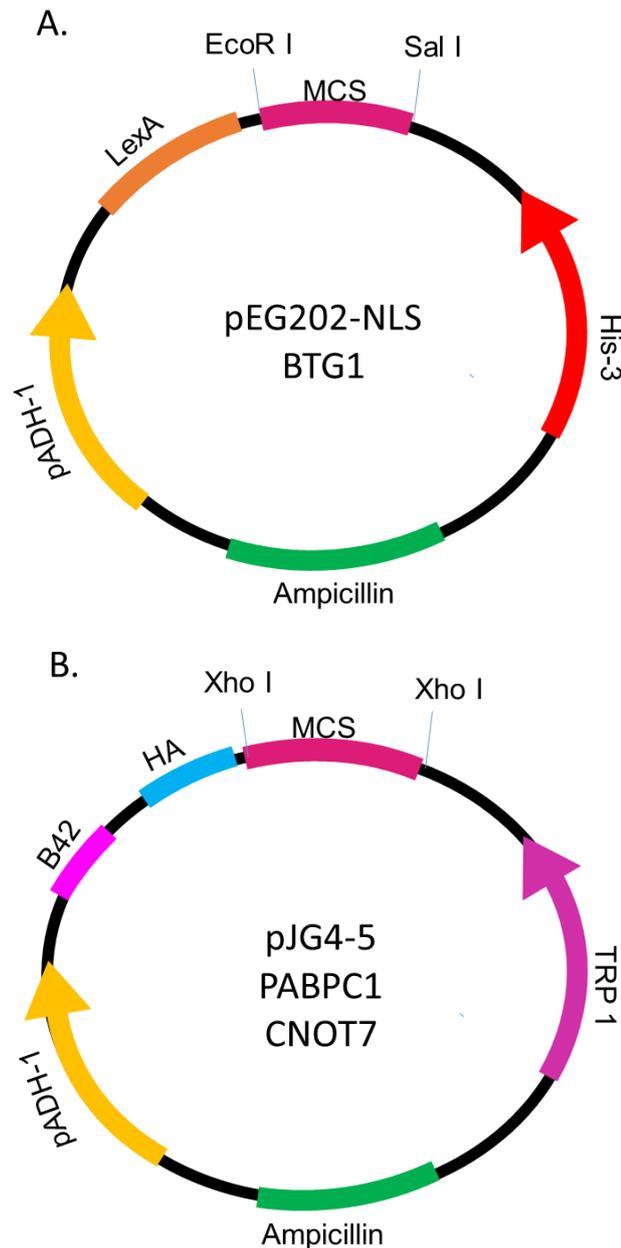
Figure 4-2. Schematic representations of the vectors used for cloning BTG1 and PABPC1 (1-146) by Hibah Almasoum for yeast two-hybrid analysis. A) Human BTG1 cDNA was cloned in-frame with the LexA DNA binding domain expression cassette of plasmid pEG202-NLS. **B)** Human cDNAs encoding PABPC1 (1-146) and CNOT7/CAF1 were cloned in frame with the B42 activation domain cassette of vector pJG4-5.
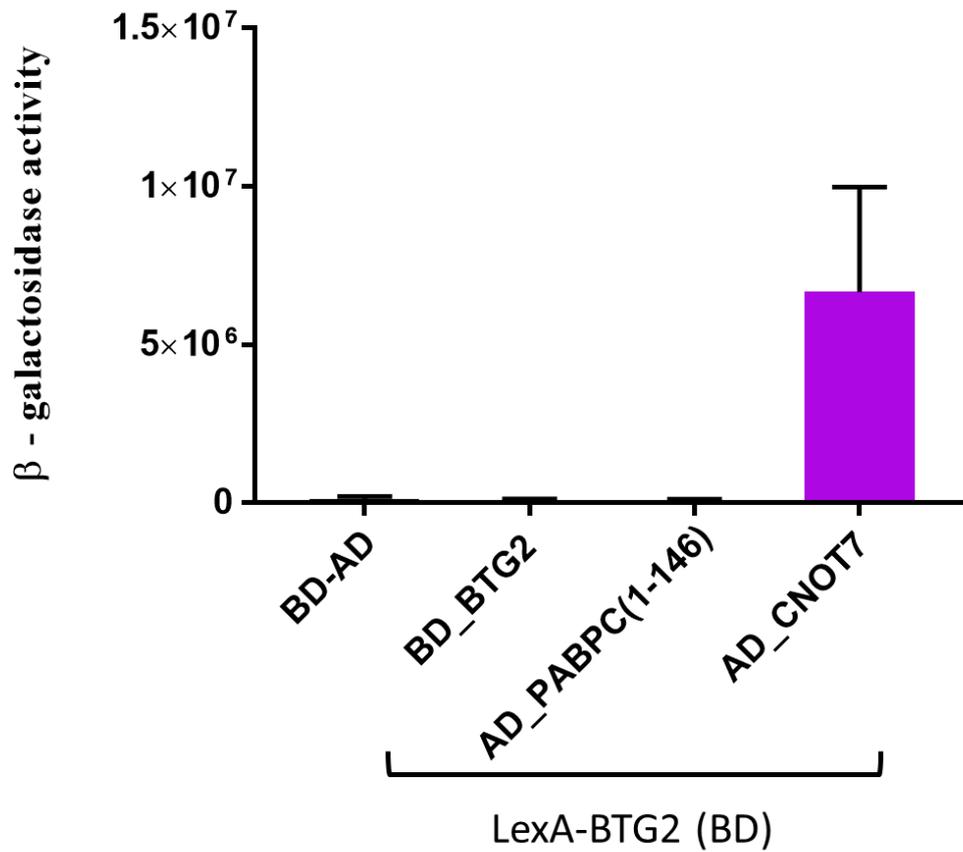
Figure 4-3. LexA-based yeast two-hybrid analysis to study PABPC1-BTG1 interactions. Three yeast colonies were tested for transformation. Yeast colonies were streaked on galactose (SC) medium agar plates. After overnight incubation, yeast cells were inoculated in a galactose medium. The β-galactosidase activity was determined and normalised to OD_{600} values. Error bars represent the standard error of the standard deviation ($n = 3$).

4.1.2. Pulldown analysis of PABPC1 (1–190)-BTG1/BTG2 using His-agarose beads

As an alternative to the yeast two-hybrid system, pulldown assays were evaluated subsequently using recombinant proteins. To this end, His-pQE80L-PABPC1 (1–190) and GST-pGEX4T1-BTG1/BTG2 were expressed in *E. coli* BL21. Expressing His-PABPC1 (1–190) wild type, GST-BTG1/BTG2 wild type, and GST wild type in *E. coli* was accomplished by inoculating a colony in a selective medium that was incubating overnight. SDS-PAGE analysis was performed to confirm the expression of His-PABPC1 (1–190, 22.4 kDa), GST (25.69 kDa), GST-BTG1 (44.89 kDa), and GST-BTG2 (43.19 kDa; Figure 4-4A).

After preparing soluble lysates of bacteria expressing wild type His-PABPC1 (1–190), GST-BTG1/BTG2, and GST, Ni-NTA agarose beads were equilibrated in wash buffer. After lysis by sonication and the separating of soluble and insoluble fractions by centrifuging, the soluble lysate containing wild type His-PABPC1 (1–190) was mixed with lysates containing GST-BTG1 or GST-BTG2 and added to equilibrated agarose beads. The mixtures were incubated at 4°C for one hour and eluted in the elution buffer.

Analysis of the eluted proteins by SDS-PAGE showed that incubating His-PABPC1 (1–190) with beads revealed efficient binding of His-PABPC1 (1–190) to the beads (Figure 4-4B, lane 1). His-PABPC1 (1–190) incubated with GST did not result in pulldown of GST (Figure 4-4B, lane 2). However, GST-BTG1 was pulled down with His-PABPC1 (1–190), confirming that His-PABPC1 (1–190) can bind to GST-BTG1 (Figure 4-4B, lane 3). Similarly, GST-BTG2 was pulled down with His-PABPC1 (1–190), showing a robust interaction (Figure 4-4B, lane 4). The pull-down result revealed BTG1 and BTG2 interacted with PABPC1 (1–190); however, the intensity of bands observed in the gel suggests the BTG2 band was more substantial than BTG.
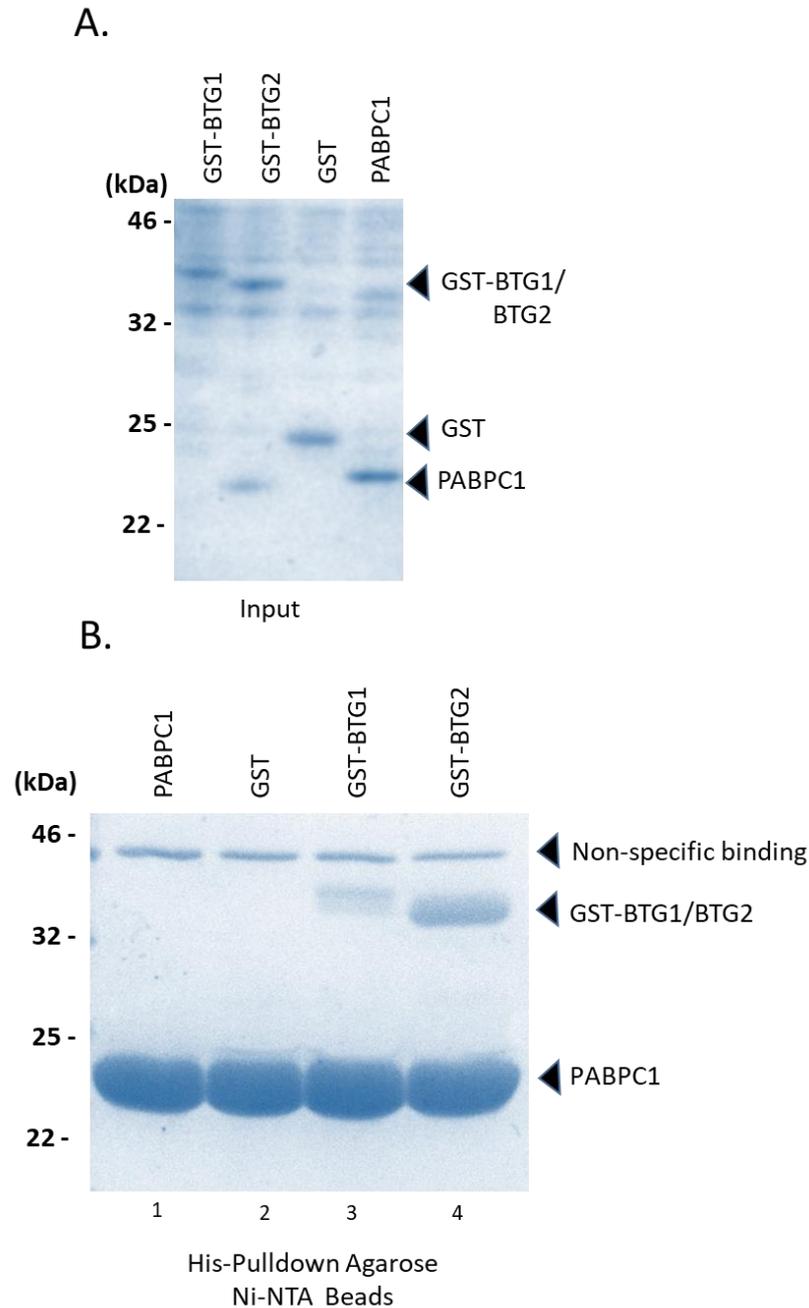
Figure 4-4. His-pulldown determining PABPC1-BTG1/BTG2 interactions by Ni-NTA agarose beads. Pull-down was performed using Ni-NTA agarose beads to study GST-BTG1/BTG2 interactions with PABPC1 (1-190) wild type. **A)** Soluble lysates of *E. coli* Bl21 (DE3) expressing GST-BTG1, GST-BTG2, GST, and His-PABPC1 (1-190). Proteins were separated by 14% SDS-PAGE and stained with Coomassie brilliant blue. **B)** His-pulldown analysis of His-PABPC1 with GST-BTG1 and GST-BTG2. Soluble lysates containing His-PABPC1 (1-190) wild type were mixed with lysates containing GST, GST-BTG1 and GST- BTG2 and incubated with Ni-NTA beads for 1 one hour at 4°C. After three washes, eluted proteins were analysed by 14% SDS PAGE.

4.2. Sited-directed mutagenesis of PABPC1

Having established a method to investigate interactions between His-PABPC1 (1–190) and GST-BTG2, the predicted interfacing residues were subjected to site-directed mutagenesis. A modified QuikChange site-directed mutagenesis method introduced point mutations using Phusion High-Fidelity DNA Polymerase. The procedure used plasmids pQE80L-PABPC1 (1–190) and pGEX4T1-BTG2 as templates and two synthetic oligonucleotide primers containing the mutation of interest. The oligonucleotides primers, complementing the opposite vector strands, are extended during PCR cycling by Phusion High-Fidelity DNA Polymerase.

A mutated plasmid containing staggered nicks is generated by incorporating oligonucleotide primers. The variants were transformed into competent DH5α; a colony was then extracted for plasmid isolation. The QuikChange method was used to substitute residues of PABPC1 (1–190) at positions Asp-21, Glu-24, Glu-29, Ile-40, Arg-44, Met-46, Thr-48, and Arg-50 with Ala. A previously cloned PABPC1 (1–190) to pQE80L vector (Figure 4-5; Pavanello et al., 2018) was used as a template. Therefore, PABPC1 (1–190) variants D21A, E24A, E29A, I40A, R44A, M46A, T48A, and R50A were obtained.

Inverse PCR was an alternative method for substituting residues with Phusion High-Fidelity DNA Polymerase. The technique relies on a gene of interest in a small vector to amplify the entire plasmid by PCR. Site-directed mutagenesis by inverse PCR involves a pair of complementary primers that anneals in the same region as the opposing DNA strand of the construct. One of the primers designed has a mutation relative to the wild type. The entire plasmid is amplified with a mutation introduced by the oligonucleotide(s). The linear product is digested with Dpn I that is phosphorylated and circularised by ligation following amplification. The variants were transformed into competent DH5α, and a single colony was used for plasmid isolation.

The inverse PCR method was used for substituting Arg-49 of PABPC1 (1–190) with Ala; Arg-49 of PABPC1 is among the predicted residues involved in mediating interaction with BTG2. Since Arg-49 is a positively charged residue, it was chosen to substitute with a negatively charged Glu acid and Ala.

During the preparation of the R49E substitution, sequence analysis revealed that a 36-nucleotide sequence corresponding to 12 amino acids (37–49) were missing and therefore named Δ37-49 (Figure 4-6), obtained incidentally. Therefore, R49E was re-cloned to the pQE80L-PABPC1 (1–190) vector by inverse PCR, and the presence of mutations in PABPC1 (1–190) was confirmed by Sanger DNA sequencing
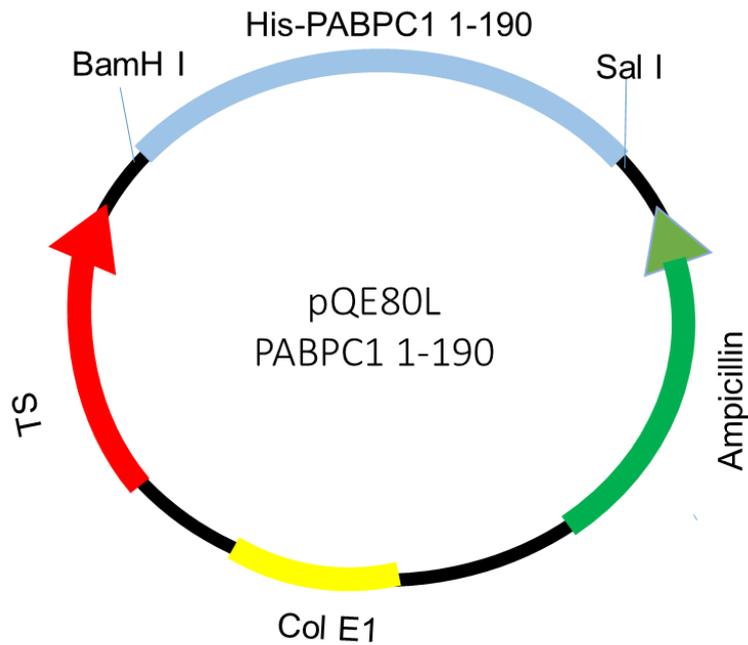
Figure 4-5. Plasmids used as templates for site-directed mutagenesis. Bacterial expression vector pQE80L-PABPC1 (1-190) containing His-tagged wild type PABPC1 (amino acids 1-190).

Figure 4-6. Sequence analysis of PABPC1 (1-190) mutant (Δ37-49). DNA sequence analysis PABPC1 (1-190) wild type with Δ37-49E. 36 nucleotides were deleted, resulting in a deletion of amino acids 37-40 and insertion of a Glu.

4.2.1. His-pulldown analysis for PABPC1 mutants R49A and Δ37-48 with BTG2 using cobalt beads (Co^{2+}-agarose)

To study the R49A and Δ37-48 interaction with BTG2, a His-pulldown analysis using Co^{2+}-agarose beads was evaluated. Plasmid His-pQE80L-PABPC1, R49A (1–190), Δ37-49 containing a 6-Histidine tagged at N-terminus, GST-pGEX4T1-BTG2 and pGEX4T1-GST tagged with GST at the N-terminus were expressed in *E. coli* BL21(DE3). The induced lysates containing His-PABPC1 (1–190; 22.4 kDa), R49A (1–190; 22.4 kDa), Δ37-49, GST (25.7 kDa), and GST-BTG2 (43.1 kDa) were analysed by SDS-PAGE (Figure 4-7A).

Co^{2+} agarose beads were used subsequently because the non-specific binding was identified with Ni-NTA agarose beads. The His-PABPC1, R49A (1–190), and Δ37-49 soluble lysates mixed with GST or GST-BTG2 lysates along with Co^{2+} beads were incubated at 4°C for one hour. The mixture was washed three times before eluting bound proteins in the elution buffer. Analysis of eluted proteins by SDS-PAGE showed that His-PABPC1 wild type, R49A (1–190) and Δ37-49 all bind with similar efficiency to cobalt beads (Figure 4-7B, lanes 1, 2, 3). Incubation of His-PABPC1 wild type, R49A (1–190), and Δ37-49 with GST did not result in pulldown of GST (Figure 4-7B, lanes 4, 5, 6). However, a robust interaction was observed after incubating wild type His-PABPC1 (1–190) and GST-BTG2, whereas pulldown for R49A resulted in low affinity for GST-BTG2. No interaction was observed with Δ37-49 (Figure 4-10B, lanes 7, 8, 9). When GST-BTG2 was incubated without His-PABPC1, some non-specific binding of GST-BTG2 was observed (Figure 4-7B, lane 10).

The pulldown revealed that the interaction observed between His-PABPC1 (1–190) and GST-BTG2 was specific because no or reduced pulldown was observed in His-PABPC1 (1–190) R49A and the Δ37-49 variant. Non-specific binding of GST-BTG2 to the beads was suppressed in His-PABPC1 (1–190) R49A and Δ37-49. The experiment also showed that Arg-49 is an essential residue mediating interactions with GST-BTG2.
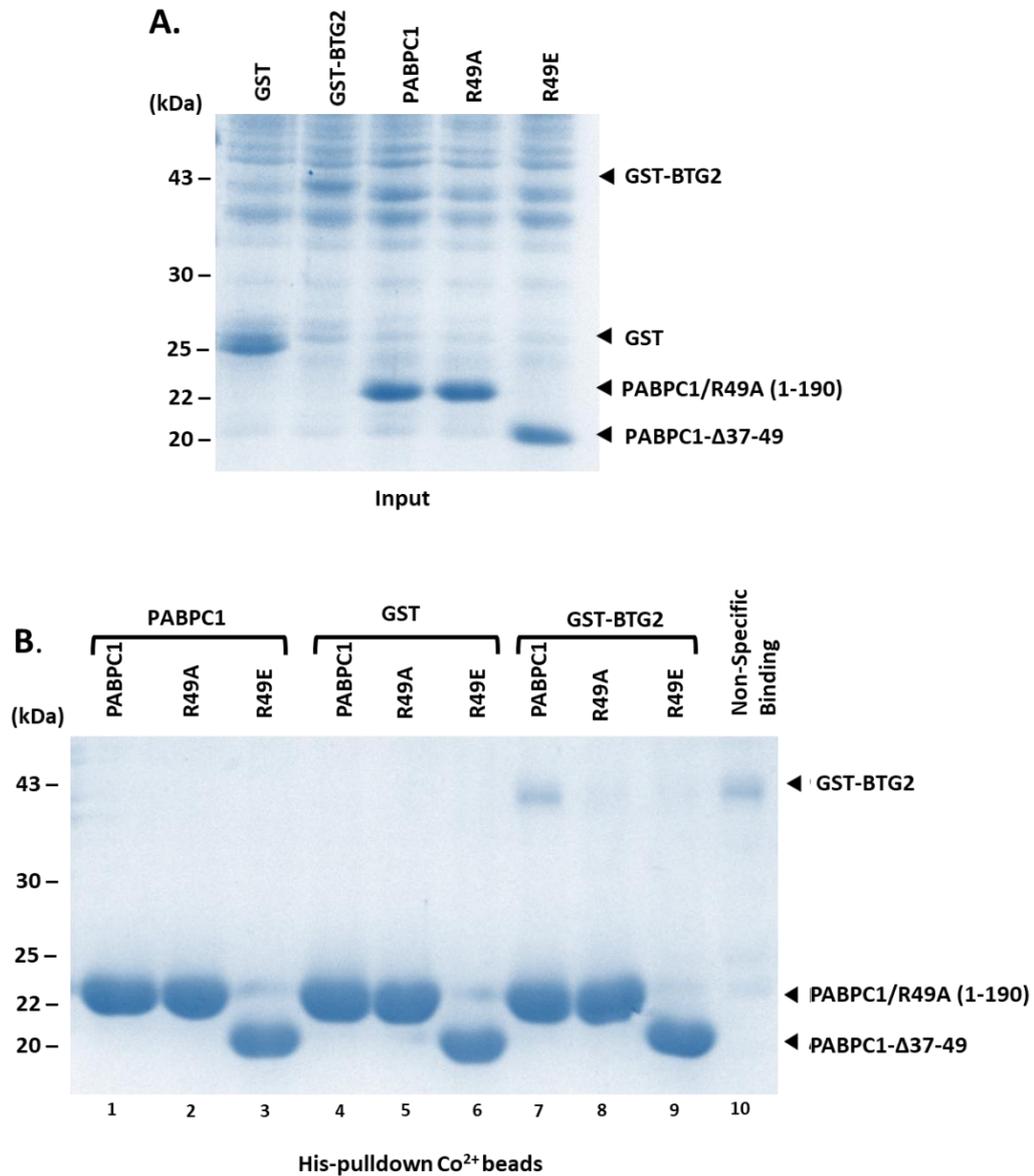
Figure 4-7. His-Pulldown using Co^{2+}-agarose beads. Using Co2+ -agarose beads, Pulldown was performed to study interactions between wild type His-PABPC (1-190), R49A, Δ37-49 and GST-BTG2. **A)** Lysates containing GST, GST-BTG2, His-PABPC1 wild type, R49A, and R49E (1-190) were examined by 14% SDS-PAGE **B)** Pulldown analysis of wild type His-PABPC1, R49A (1-190), Δ37-49 with GST, and GST-BTG2. Proteins were incubated with Co^{2+}-agarose beads for 1 one hour at 4°C. After three washes, the eluted proteins were analysed by 14% SDS PAGE and stained with Coomassie Brilliant Blue.

4.2.2. Studying His-PABPC1 interactions with GST-BTG2 by Glutathione Agarose Magnetic Beads

To further optimise the pulldown assay, GST-pulldown was subsequently evaluated to study interactions between His-PABPC1 and GST-BTG2. GST-pulldown was carried out using Glutathione Agarose Magnetic Beads. An additional two mutations, E24A and R50A, were included in these experiments to study its interactions with GST-BTG2 compared to PABPC1. His-PABPC1, E24A, R49A, R50A (1–190), and GST-BTG2 were expressed in *E. coli* BL21 (DE3). The induced lysates containing His-PABPC1 wild type, E24A, R49A, R50A (1–190), GST (25.69 kDa), and GST-BTG2 (43.1 kDa) were analysed by SDS-PAGE (Figure 4-8A).

The soluble lysate containing GST-BTG2 was mixed with Glutathione Agarose Beads, followed by the addition of lysates containing His-PABPC1 (1–190) and the E24A, R49A, and R50A variants. The mixture was incubated at 4°C for one hour and then washed three times with wash buffer. Bound proteins were eluted in elution buffer and analysed by SDS-PAGE. However, there is no pulldown of His-PABPC1, E24A, R49A, or R50A (1–190, Figure 4-8B, lanes 2, 3, 4 and 5) GST-BTG2 was observed. Since no PABPC1, E24A, or R50A bands were identified, no conclusive results were obtained. The use of GST pulldown to study interactions between His-PABPC1 and GST-BTG2 was not further evaluated, and further experiments focussed on pulldown experiments using His-pulldown of PABPC1 (1–190).
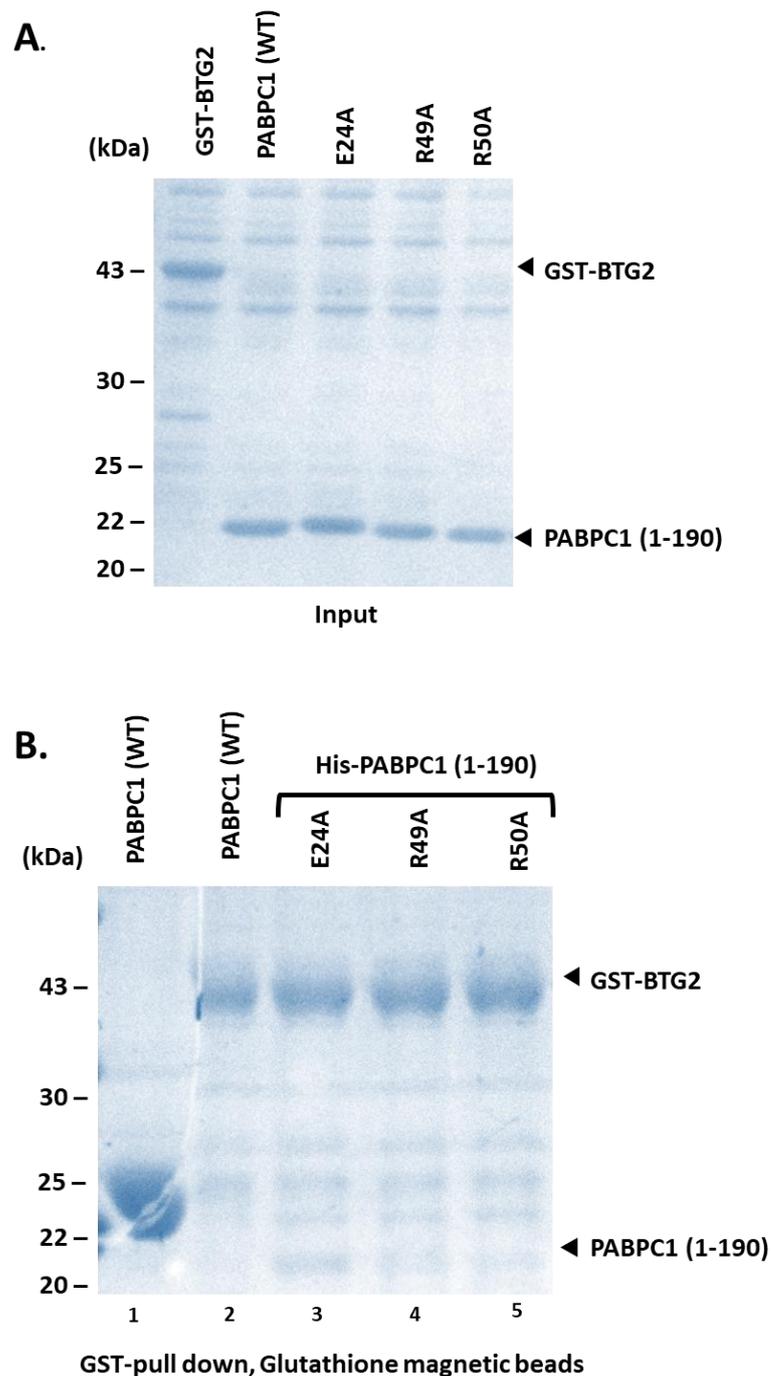
Figure 4-8. GST-Pulldown by magnetic glutathione beads. Pulldown performed using magnetic glutathione beads to study interactions between GST, GST-BTG2 and His-PABPC1 wild type, E24A, R49A, and R50A (1-190). **A)** Soluble lysates containing GST, GST-BTG2, His-PABPC1 wild type, E24A R49A, and R50A (1-190) were examined by 14% SDS-PAGE. **B)** Pulldown of His-PABPC1 wild type E24A, R49A, and R50A (1-190) by GST-BTG2. Lysates were mixed and incubated with magnetic glutathione beads for one hour at 4°C after three washes of the bead's bound protein was eluted and analysed by 14% SDS PAGE and stained with Coomassie Brilliant Blue.

4.2.3. Optimisation of His-pulldown experiments with nickel-nitrilotriacetic acid (Ni2+-NTA) magnetic beads

The His-pulldown experiment results in section 4.3.3 disclosed the non-specific binding of GST-tagged BTG2 when incubated with beads only (Figure 4.7B, Lane 10). The conditions for the His-pulldown experiments were optimised, and magnetic Ni^{2+}-NTA beads were evaluated next to avoid the non-specific binding of GST-BTG2 to His-affinity beads.

An overnight grown culture of pQE80L-His-PABPC1, R49A (1–190), pGEX4T1-BTG2, and pGEX4T1 diluted with the fresh medium was induced. The induced lysates containing His-PABPC1 (1–190), R49A, and GST-BTG2 were analysed by SDS-PAGE (Figure 4-9A). The gel was stained with Co-massie and showed that all proteins migrated as expected based on their calculated molecular weight: GST (25.6 kDa), GST-BTG2 (43 kDa), and PABPC1 (1–190; 22.4 kDa). The soluble lysates were obtained by centrifuging the induced culture and re-suspending bacteria expressing His-tagged proteins in lysis buffer. The Ni^{2+}-NTA magnetic beads were equilibrated with wash buffer. Subsequent parts of the experiment were conducted in three parallel batches based on the addition of beads to the lysates and length of incubation.

Batch 1

Soluble lysates containing His-PABPC1 (1–190) and His-PABPC1 R49A (1–190) were mixed with lysates containing GST or GST-BTG2 and incubated 4°C for one hour. Next, Ni^{2+}-NTA beads were added to the lysates and incubated for one hour at 4°C. The beads were washed three times with wash buffer, and the proteins were eluted in the elution buffer. The eluted proteins were analysed by SDS-PAGE, which revealed that His-PABPC1 (1–190) and His-PABPC1 R49A (1–190) did not pull-down GST as expected (Figure 4-9B, Batch 1). However, a robust interaction of GST-BTG2 with His-PABPC1 (1–190) was identified (Figure 4-9B, Batch 1); however, a very weak interaction was observed with His-PABPC1 R49A (1-190) when compared to the wild type His-PABPC1 (1–190; Figure 4-9B, Batch 1).

Batch 2

In this batch, lysates containing His-PABPC1 (1–190) and His-PABPC1 R49A (1–190) were incubated with Ni^{+2}-NTA beads at 4°C for one hour. Following the incubation, lysates containing GST and GST-BTG2 were added and incubated for an additional one hour at 4°C. After three washes, the eluted proteins were subjected to SDS-PAGE, which revealed that GST, as expected, was not pulled down with His-PABPC1 (1–190) or R49A (Figure 4-9B, Batch 2). However, a robust interaction was identified between His PABPC1 (1–190) and GST-BTG2 (Figure 4-9B, Batch 2). His-PABPC1 R49A (1–190) showed reduced affinity to GST-BTG2 when compared to interaction in Batch I (Figure 4-9B, Batch 2).

Batch 3

The final batch involved combining lysates containing His-PABPC1 (1–190), His-PABPC1 R49A (1–190) with GST or GST-BTG2, and Ni^{2+}-NTA beads and incubating the mixture at 4°C for one hour. The proteins were eluted in the elution buffer after washing the beads three times in the wash buffer. In this batch, no pulldown was observed when GST lysate was incubated with His-PABPC1 (1–190) or His-PABPC1 R49A (1–190; Figure 4-9B, Batch 3). By contrast, when His-PABPC1 (1–190) was incubated with GST-BTG2, pulldown of GST-BTG2 was identified (Figure 4-9B, Batch 3). A reduced pulldown of GST-BTG2 was observed in the presence of His-PABPC1 R49A (1–190, Figure 4-9B, Batch 3).

A comparison of the pulldown conditions showed that the conditions used for Batch 1 were best for investigating the interaction between His PABPC1 (1–190) and GST BTG2. In Batch 1, mixing soluble lysates containing His-PABPC1 (1–190) and GST-BTG2 prior to adding beads resulted in the robust pulldown, which was not observed when using His-PABPC1 R49A (1–190). While the methods used for Batch 2 and Batch 3 also resulted in the robust pulldown of GST-BTG2, these procedures led to a more significant pulldown of His-PABPC1 R49A (1–190). Therefore, the conditions used for Batch 1 were used subsequently to assess the interactions between the GST-BTG2 and His-PABPC1 (1–190) variants containing single amino acid substitutions.
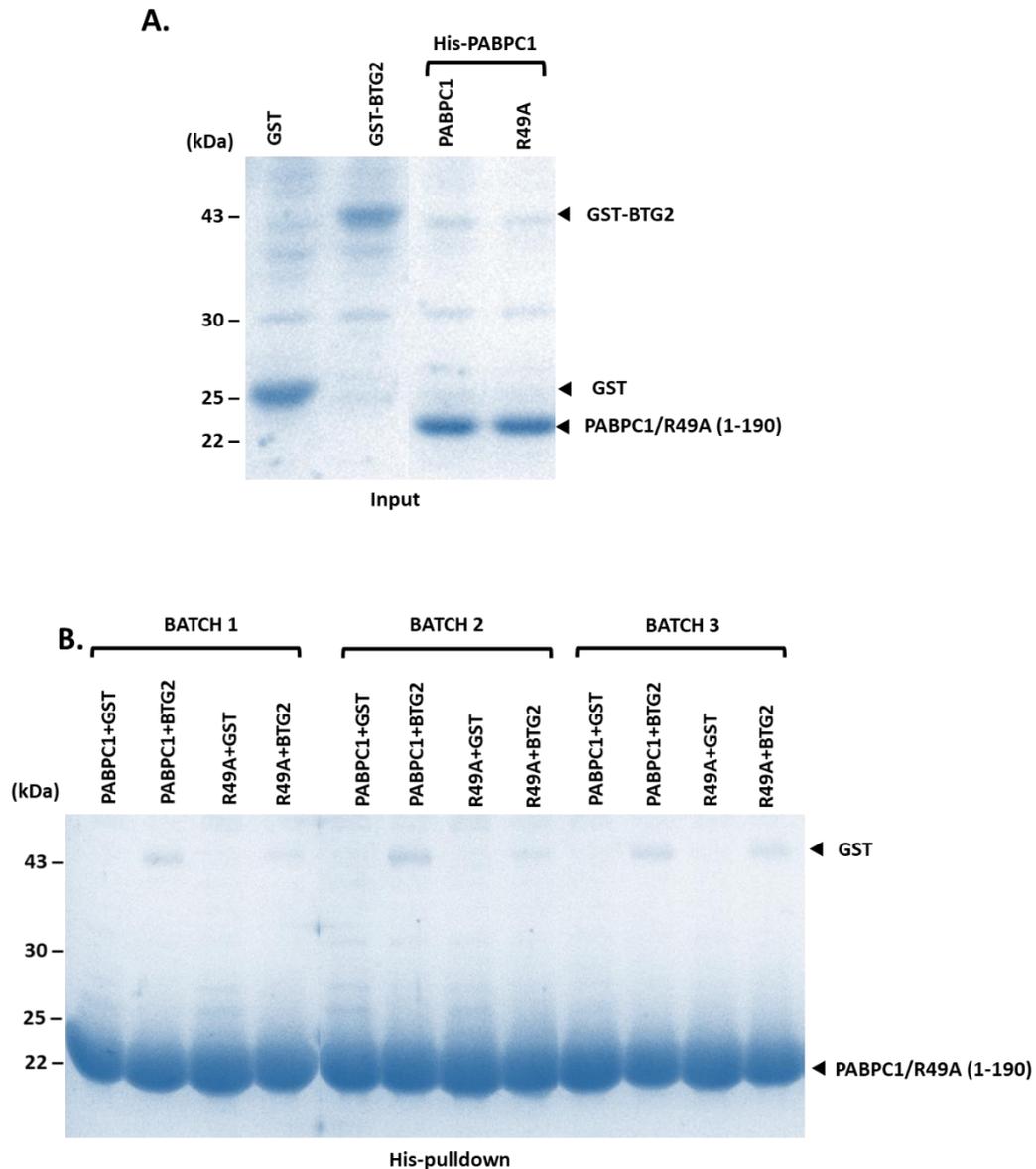
Figure 4-9. Optimisation of conditions for pulldown assays using magnetic Ni2+NTA beads. A) Soluble lysates containing His-PABPC1 (1-190), His-PABPC1 R49A (1-190), GST, and GST-BTG2. Proteins were analysed by SDS-PAGE and stained with Coomassie Brilliant Blue. **B)** Optimisation of conditions for pulldown experiments using lysates containing His-PABPC1(1-190), His-PABPC1 R49A (1-190), GST, and GST BTG2. The batch I: lysates were incubated before adding magnetic Ni^{2+}-NTA beads. Batch II: lysates containing His-PABPC1 (1-190) and His-PABPC1 R49A (1-190) were incubated with magnetic Ni^{2+}-NTA beads before adding lysates containing GST and GST-BTG2. Batch III: lysates containing His-PABPC1 (1-190), His-PABPC1 R49A (1-190), GST, and GST-BTG2 were combined and added directly to magnetic Ni^{2+}-NTA beads. All incubation steps were carried out at 4°C for one hour. Proteins were analysed by SDS-PAGE and stained with Coomassie Brilliant Blue.

4.2.4. Identification of residues in RRM1 of PABPC1 involved in the interaction with BTG2

Having optimised conditions for studying the interaction between His-PABPC1 (1–190) and GST-BTG2, the role of residues predicted to mediate this interaction as described in Chapter 3 were evaluated. Site-specific mutations were introduced to the predicted residues of PABPC1 (1–190) by substituting them with Ala and Glu acid. The variants D21A, E24A, E29A, I40A, R44A, M46A, R49A, R49E, and R50A were investigated to study their interaction with BTG2. Before pulldown experiments, wild type His-PABPC1 (1–190), D21A, E29A, I40A, R44A, M46A, R49A, R49E, R50A, and GST-BTG2 were expressed in *E. coli* BL21 (DE3). After lysis by sonication and separating the insoluble and soluble fraction by centrifugation, the soluble lysates were subjected to SDS-PAGE analysis and stained with Coomassie Brilliant Blue (Figure 4-10A).

After mixing lysates containing His-PABPC1 (1–190) variants and GST-BTG2, reactions were incubated at 4°C for one hour. Next, Ni^{2+}-NTA magnetic beads were added, and reactions were incubated at 4°C for an additional one hour. Bound proteins were eluted in elution buffer after washing the beads three times in wash buffer. The eluted proteins were analysed by SDS-PAGE (Figure 4-10B). In the presence of wild type His-PABPC1 (1–190), robust pulldown of GST-BTG2 was observed (Figure 4-10B). In the presence of His-PABPC1 (1–190) variants D21A and E29A, pulldown of GST-BTG2 was observed to a similar level as compared to wild type His-PABPC1 (1–190, Figure 4-10B). Pulldowns using variants I40A, R44A, R49A, and R50A resulted in reduced pulldown of GST-BTG2 (Figure 4-10B), whereas variants E24A, M46A, and R49E resulted in strongly reduced pulldown of GST-BTG2 (Figure 4-10B).

These results indicate that Glu-24, Met-46 and Arg-49 are key residues of RRM1 of PABPC1 that mediate interactions with BTG2. Residues Ile-40, Arg-44, and Arg-50 of PABPC1 RRM1 may make additional contacts with BTG2, while the side chains of PABPC1 residues Asp-21 and Glu-29 make a minor contribution or are not involved in binding to BTG2.
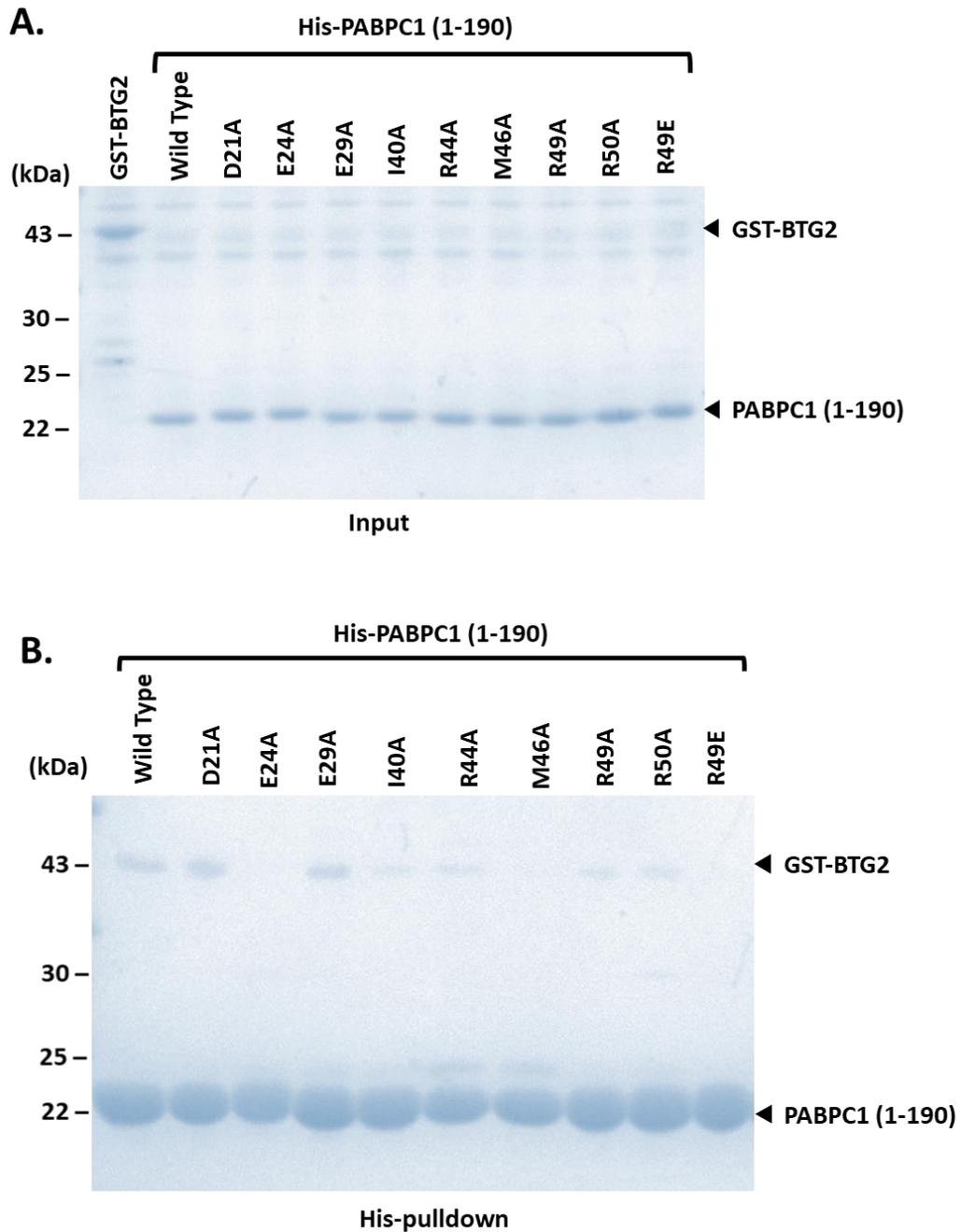
Figure 4-10. Identification of residues of RRM1 of PABPC1 involved in binding GST-BTG2. A) Soluble lysates containing wild type His-PABPC1 (1-190), single amino acid variants and GST-BTG2 were separated by 14% SDS-PAGE and stained with Coomassie Brilliant Blue. **B)** Pulldown analysis of interactions between wild type His-PABPC1 (1-190), His-PABPC1 variants containing single amino acid substitutions, and GST-BTG2. Lysates were combined and incubated at 4°C for one hour before the addition of magnetic Ni^{2+}-NTA beads and a further incubation at 4°C. Bound proteins were eluted, separated by 14% SDS-PAGE, and stained with Coomassie Brilliant Blue.

4.2.5. Expression and interaction of cancer-associated mutation T48P

Subsequently, the possible involvement of residue Thr-48 of PABPC1 in mediating the interaction with BTG2 was investigated. Thr-48 of PABPC1 was among the amino acids predicted to interface with BTG2 (Chapter 1). Interestingly, a recurring cancer mutation at this position that results in the single amino acid change T48P has been reported in the COSMIC database (Tate et al., 2019). To investigate the importance of this residue in mediating interactions with BTG2, wild type His-PABPC1 (1–190) and variants containing the single amino acid substitution T48A and T48P GST-BTG2 were expressed in *E. coli* BL21 (DE3). After lysis by sonication and separating the insoluble and soluble fraction by centrifugation, the soluble lysates were subjected to SDS-PAGE analysis and stained with Coomassie Brilliant Blue (Figure 4-11A).

Subsequently, the pulldown analysis was done using soluble lysates with Ni^{2+}-NTA magnetic beads. The lysates containing His-PABPC1 wild type, T48A, and T48P were mixed with lysates containing GST-BTG2, and the reactions were incubated at 4°C for one hour. Then, Ni^{2+}-NTA magnetic beads were added, and reactions were further incubated at 4°C for an additional one hour. Bound proteins were eluted in elution buffer after washing the beads three times in wash buffer. The eluted proteins were analysed by SDS PAGE (Figure 4-11B). The results revealed that the variant T48A could mediate interactions with GST-BTG2, such as wild type PABPC1 (Figure 4-11B). In comparison, the cancer-associated variant T48P resulted in reduced pulldown of GST-BTG2 (Figure 4-11B). These results suggest that the T48P cancer-associated PABPC1 variant found with low frequency in head and neck cancer (squamous cell carcinoma) and thyroid cancer (Martin et al., 2014) may contribute to the tumour phenotype.
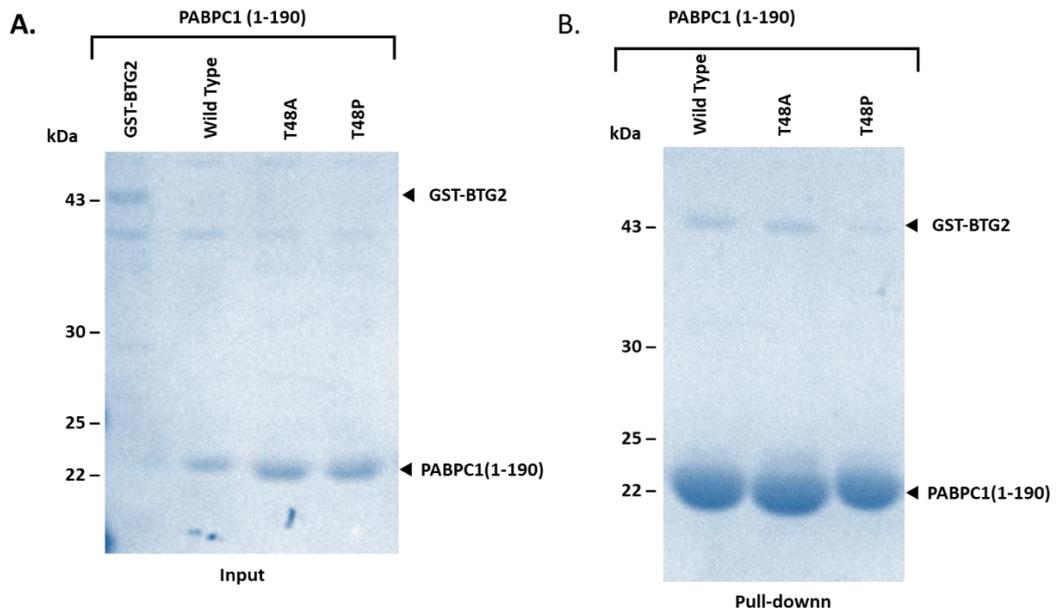
Figure 4-11. Interaction studies for T48P patient mutation compared to T48A and PABPC1 (1-190). A) Soluble lysates containing wild type His-PABPC1 (1-190) and single amino acid variants T48A, T48P, and GST-BTG2. **B)** Pulldown analysis of interactions between wild type His-PABPC1 and single amino acid substitution variants T48A and T48P with GST-BTG2. Lysates were incubated at 4°C for one hour before adding magnetic Ni^{2+}-NTA beads and a further incubation at 4°C for one hour. Bound proteins were eluted, separated by 14% SDS-PAGE, and stained with Coomassie Brilliant Blue.

4.2.6. Site-directed mutagenesis of BTG2

Point mutations were introduced into BTG2 to define in more detail the requirements of BTG2 residues involved in the interaction with PABPC1. The residues found in Box C of BTG2 at positions Ser-118, Ile-119, Glu-124, Glu-125, and Ala-126 were substituted with residues of TOB1 at positions Ala-118, Val-119, Leu-124, Asp-125, and Asp-126 (Figure 4-12). The QuikChange method replaced the residues in plasmid pGEX4T1-BTG2 (Figure 4-13). Moreover, in addition to the point mutations, mutations resulting in amino acid substitutions S118A-I119V, and the triple alteration E124L-E125D-A126D, were obtained using the QuikChange method. Inverse PCR was applied to obtain variants E124A, E125A, and a variant of BTG2 containing a deletion of Box C (residues 116-127; BTG2-ΔBox C) along with a variant containing a stop codon following Box C (residues 1-127; BTG2-ΔCterm, Figure 4-14).
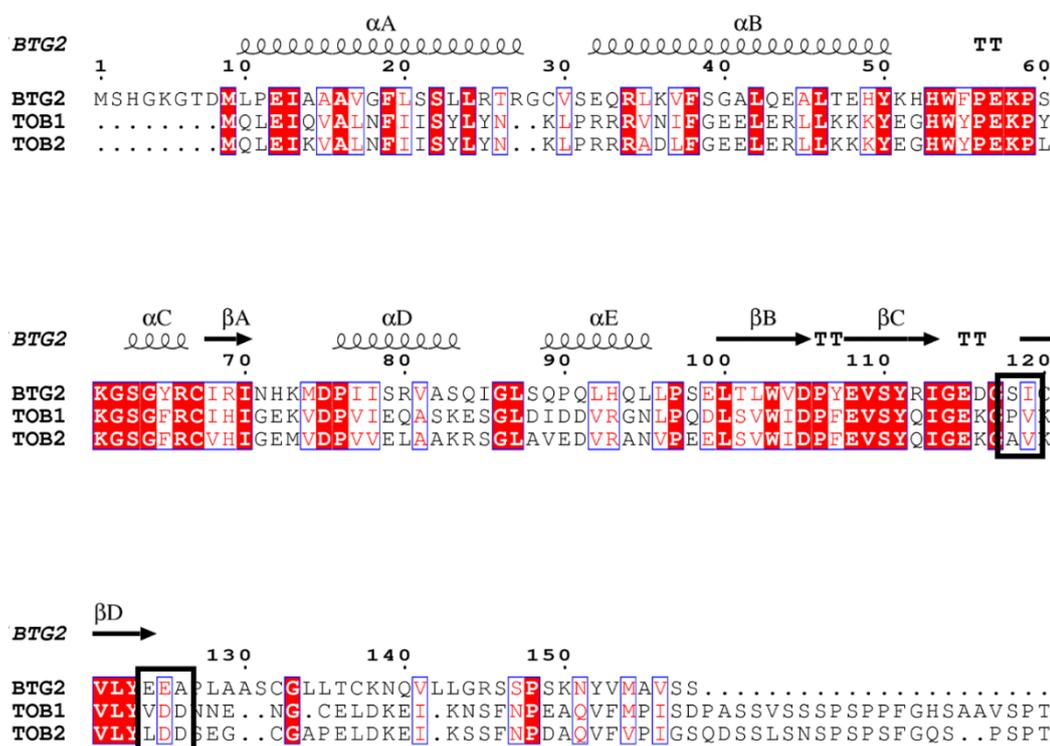
Figure 4-12. Sequence Comparison of BTG2 TOB1 and TOB2. The sequences were compared to show the similarity between BTG2 and TOB2. The highlighted amino acids in black represent the residues of TOB2 at positions 118, 119, 124, 125, and 126 substitutes with residues in the Box C region in the exact location. Clustal omega was used to perform sequence alignment for BTG2, TOB1 and TOB2 and formatted using ESPript 3 (Robert and Gouet, 2014). Dots represent the 10th amino acid residue in the sequence under *number*.
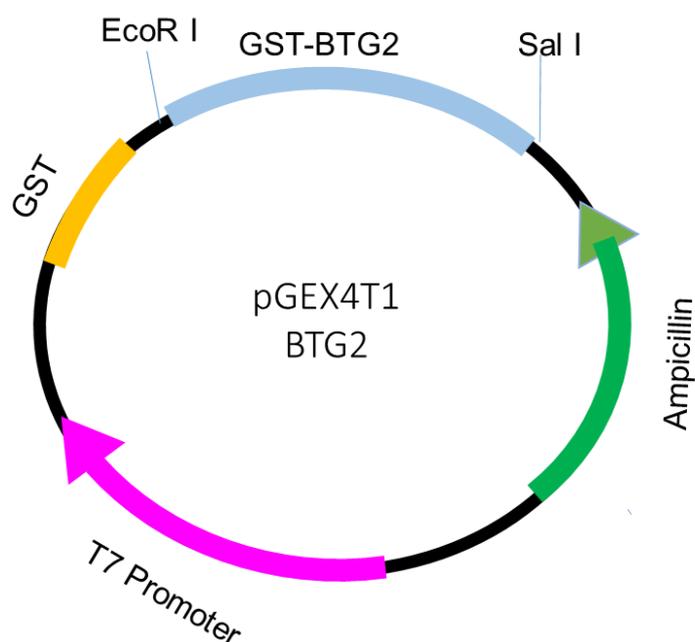
Figure 4-13. Plasmids are used as templates for site-directed mutagenesis. Bacterial expression vector pGEX4T1 containing GST tagged in the N-terminal of wild type BTG2.
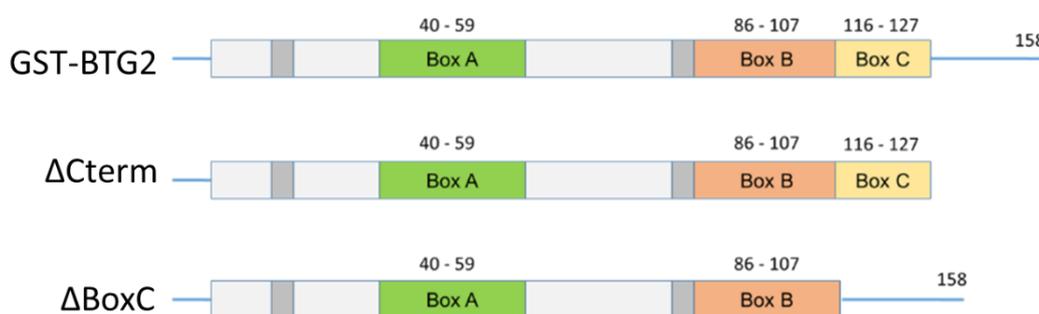
Figure 4-14. An image of ΔC-term and ΔBox C. Schematic representation of wild type BTG2 and BTG2 variants containing a deletion of C-terminal (ΔC-term) and a deletion of the residues in Box C of BTG2 (ΔBox C).

4.3.6. The role of Box C residues of BTG2 in interacting with RRM1 of PABPC1

Next, the pulldown assay was used to define in more detail the requirement of Box C residues for the interaction with PABPC1. To prepare GST-BTG2 variants for interaction analysis, a single colony of pGEX4T1-GST-BTG2, or derivatives encoding the variants S118A, I119V, E124A, E125A, S118A-I119V, or E124L-E125D-A126D, were grown overnight. The overnight cultures were diluted with the fresh medium and were induced. After lysis by sonication and separating the insoluble and soluble fraction by centrifugation, the soluble lysates were incubated with Magnetic Glutathione Beads at 4°C for one hour, washed three times in buffer. Analysis of the bound proteins by SDS-PAGE indicates that the soluble lysates contained comparable levels of all GST-BTG2 variants (Figure 4-15A).

Subsequently, lysates containing GST-BTG2 variants were mixed with soluble lysate containing wild type His-PABPC1 (1–190) and magnetic Ni^{2+}-NTA beads. The bound proteins were eluted and analysed by 14% SDS-PAGE (Figure 4-15B). Compared to the level of wild type GST-BTG2, a modest reduction was seen in the pulldown of GST-BTG2 containing S116A and I119V (Figure 4-15B). A more significant decrease was seen for the pulldown of GST-BTG2 containing the double amino acid substitution S116A-I119V or the triple amino acid substitution E124L-E125D-A126D (Figure 4-15B).

To investigate the role of the Box C and the C-terminal region of BTG2 in interacting with PABPC1, a pulldown experiment was carried out using GST-BTG2, GST-BTG2ΔBox C, or GST-BTG2ΔCterm and His-PABPC1 (1–190). To prepare the GST-BTG2 variants for interaction analysis, the transformed plasmid pGEX4T1-GST-BTG2 wild type, GST-BTG2Δ-Box C, or GST-BTG2Δ C-term, was grown overnight. The overnight culture was diluted with the fresh medium. The soluble lysates obtained from the culture were incubated with Magnetic Glutathione Beads for one hour at 4°C. The eluted proteins after three washes are subjected to 14% SDS PAGE analysis and stained with Coomassie Brilliant Blue (Figure 4-15C).

The His-pulldown assay was performed by incubating soluble lysates containing wild type GST-BTG2, GST-BTG2ΔBox C, or GST-BTG2ΔCterm with His-PABPC1 (1–190) and magnetic Ni^{2+}-NTA. The bound proteins were eluted and analysed by 14% SDS-PAGE (Figure 4-15D). The pulldown revealed a robust interaction of wild type GST-BTG2 and His-PABPC1 (1–190; Figure 4-15D). By contrast, incubating GST-BTG2ΔBox C with His-PABPC1 (1–190) resulted in the disruption of the interaction with His-PABPC1 (1–190). When GST-BTG2ΔCterm was incubated with His-PABPC1 (1–190), a robust pulldown of GST-BTG2ΔCterm was seen similar to that of wild type GST-BTG2 (Figure 4-15D). Considered together, these experiments confirm that Box C residues of BTG2 mediate interaction with PABPC1 RRM1 residues as proposed by (Stupfler et al., 2016) and that the C-terminal region of BTG2 does not make a significant contribution to the interaction with RRM1 of PABPC1.
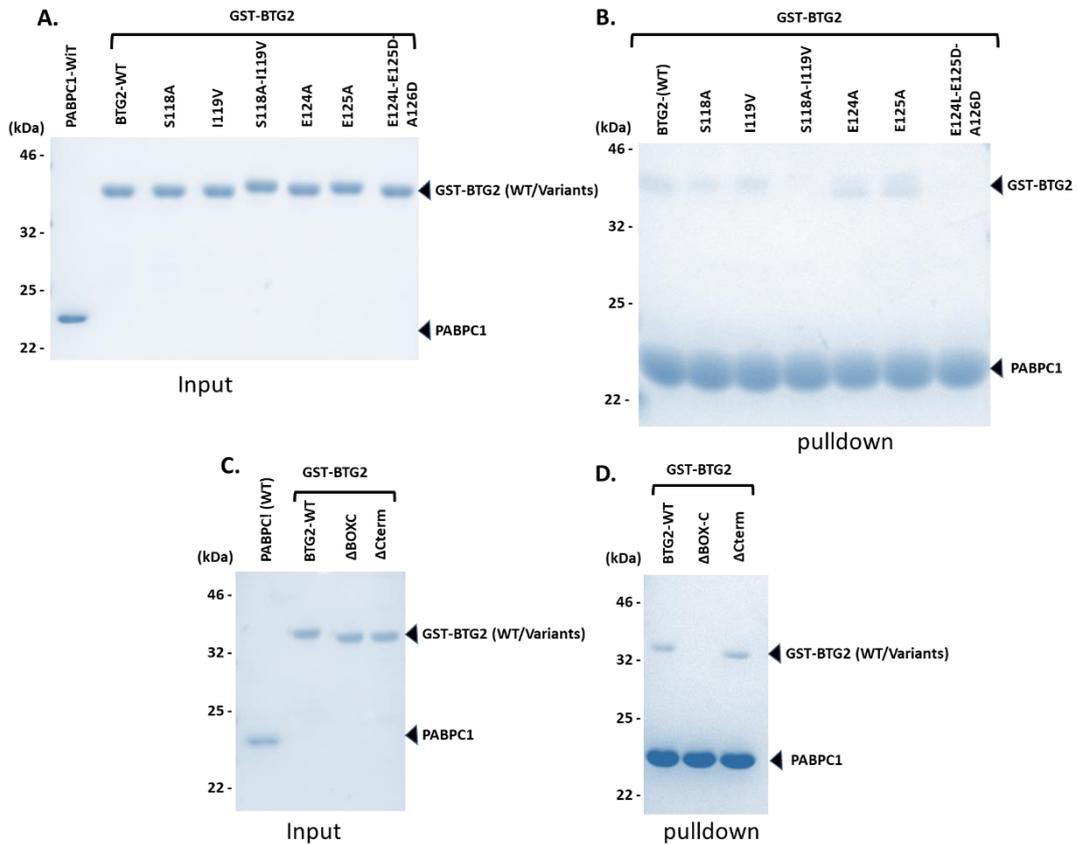
Figure 4-15. The Box C region of BTG2 is needed forPABPC1-BTG2 interactions. A) The input GST-BTG2 (wild type), S118A, I119V, S118A-I119V, E124A, E125A, E124L-E125D-A126D. **B)** His-pulldown using GST-BTG2 variants and His-PABPC1 (1-190). **C)** The input for BTG2 and variants ΔC-term, ΔBox C. Proteins used for pulldown analysis. **D)** His-pulldown using GST-BTG2, ΔC-term, and ΔBox C and His-PABPC1 (1-190). Proteins were separated by 14% SDS-PAGE and stained with Coomassie Brilliant Blue.

4.3. Discussion

4.3.1. The analyses for PABPC1 (1–190)-BTG1/BTG2 interaction

The yeast two-hybrid method used to perform the experiments has LexA binding domain fused to BTG1 and the B42 activation domain fused to PABPC1. The yeast strain EGY48 used in the experiment has two reporter genes, LEU2 and lacZ, that are responsive to transcription activated LexA operator sequences (Golemis et al., 2001a). Before studying PABPC1 (1–146) interaction with BTG1, the binding containing LexA-BTG1 was co-transformed with reporter plasmid pSH18-34 without transcription activator domain. The plasmids transformed to the selective medium containing leucine revealed colonies; however, independent testing colonies did not show β-galactosidase activity, indicating that BTG1 does not lead to activation in the absence of the B42 activation domain. However, co-transformation of plasmids expressing LexA-BTG1, B42-PABPC1, and the pSH18-34 reporter plasmid in the selective medium containing leucine showed colonies. However, testing three independent colonies did not show β-galactosidase activity, indicating any PABPC1-BTG1 interactions.

No western blot analysis was performed to determine the expression of PABPC1; however, no β-galactosidase activity was detected upon expression of LexA-BTG1 and B42-PABPC1; moreover, the β-galactosidase assay conducted to study BTG1-CNOT7/CAF1 interaction revealed strong activity. Therefore, it was concluded that both BTG1 and CNOT7/CAF1 were expressed while PABPC1 was repressed in the yeast genome. In rabbit reticulocyte lysates, PABPC1 expression is regulated by a feedback loop (de Melo Neto et al., 1995). The present study reveals that excess PABPC1 binds to the 5′ UTR upstream of the initiator codon AUG of PABPC1 mRNA, repressing translation. A similar mechanism in yeast may lead to downregulation of LexA-PABPC1.

4.3.2. Site-directed mutagenesis approach to substitute residues

Alanine scanning is a widely used mutagenesis approach for substituting residues. This is because it eliminates the side chain, does not alter the main-chain conformation, and does not impose extreme electrostatic or steric effects (Lefevre et al., 1997). Therefore, PABPC1 residues were substituted with alanine amino acids at Asp-21, Glu-24, Glu-29, Ile-40, Arg-44, Met-46, Thr-48, Arg-49, and Arg-50. In addition, arginine amino acids can pair with negatively charged glutamic acid by salt bridges and can act as an acceptor of hydrogen bonds (Moreira et al., 2007).

Since arginine is a positively charged amino acid, an alternative amino acid that was explored was glutamic acid. The positively charged residue Arg-49 of PABPC1 was substituted with negatively charged glutamic acid, under the hypothesis that R49E could cause electrostatic repulsion of PABPC1 interaction with BTG2 if residue Arg-49 of PABPC1 were involved in electrostatic interactions with an acidic residue of BTG2.

Mutations introduced to BTG2 were based on substituting BTG2 amino acid residues with the amino acids found in the equivalent positions in TOB1. BTG2 residues at positions Ser-118, Ile-119, Glu-124, Glu-125, and Ala-126 were substituted with TOB1 residues at positions Ala-118, Val-119, Leu-124, Ala-125, and Asp-126. Unlike other BTG family members, TOB1 and TOB2 contain an extended C-terminal domain with two PAM motifs that are responsible for interacting with PABPC1 (Ezzeddine et al., 2007), Therefore, substituting BTG2 amino acid with the equivalent residues of TOB1 resulted in reduced interactions with variant S118A and I119V, while the variants with two- or three-point mutations affected the ability of S118A-I119V and E124L-E125D-A126D to bind to PABPC1 (1–190).

4.3.3. Pulldown for variants R49A and Δ37-49

Examining His-pulldown results from Figure 4-10 with Co^{2+} beads revealed weak interactions of BTG2 with variant R49A. The residue Arg-49 can cause electrostatic interaction with negatively charged residues in Box C of BTG2. Substituting Arg-49 with alanine may have caused weak interaction by disrupting the electrostatic interfaces of Arg-49 with negatively charged residues in Box C of BTG2. By contrast, substituting Arg-49 with negatively charged glutamic acid could cause electrostatic repulsion and disrupt the interaction with BTG2 more strongly.

Analysing the pull-down result for R49E with BTG2 revealed disruption in the interaction (Fig 4-13). However, while constructing the expression plasmid containing a single amino acid, a deletion mutant is replacing amino acids 37-49 with glutamate was fortuitously obtained (designated Δ37-49). No pulldown of BTG2 was observed using the His-PABPC1 Δ37-49 variant, confirming the importance of this region for the interaction (Figure 4-7). Moreover, it was observed that in the presence of this variant, no detectable binding of GST-BTG2 to the beads was observed, further supporting the notion that the pulldown of GST-BTG2 by His-PABPC1 was specific.

4.3.4. Non-specific interaction of GST-BTG2

The beads used for the His pull-down experiment has a strong affinity for His-tagged protein, such as PABPC1(1–190). However, GST-BTG2 could not bind beads specific to His-tagged proteins, but the non-specific binding was shown when GST BTG2 was incubated with Co^{2+} beads (Figure 4-10). The proteins tagged with histidine have a high binding affinity for Co2+ and Ni2+-NTA beads; however, the problem with using poly-histidine affinity beads lies with the non-specific binding of untagged proteins. The cellular proteins with two or more histidine amino acids could bind to the beads due to their high-affinity property (Bornhorst & Falke, 2000). Non-specific binding of GST-BTG2 to the beads can result in false positives for protein interaction studies. Therefore, the conditions for the pulldown assay were optimised in three batches. The interaction of the variant His-PABPC1 R49A (1–190) with GST-BTG2 was used to control the interaction between His-PABPC1 (1–190) GST-BTG2.

The result from the previous pulldown experiment (Figure 4-7) exposed weak interaction of His-PABPC1 R49A (1–190) with GST-BTG2 (Fig. 4-10). In order to optimise the pulldown experiment and avoid non-specific binding observed when GST-BTG2 was incubated with beads in the absence of His-PABPC1, His-PABPC1 R49A (1–190) utilised a negative control. Optimisation of the His pulldown experiment was performed in three batches based on the order of addition and length of incubation. The results concluded that the method adopted in Batch 1 was the optimal condition for His-pulldown analysis, and mixing soluble lysates prior to adding beads prevented GST-BTG2 from binding beads and leading to non-specific binding.

4.3.5. PABPC1 residues regulating interaction with BTG2

Examining the results from His-pulldown using Ni^{2+}-NTA beads revealed that the predicted interfacing residues of PABPC1 Glu-24, Ile-40, Arg-44, Met-46, Arg-49, and Arg-50 contribute to the interaction with BTG2. It is possible that Arg-49—as a positively charged amino acid residue was able to interface with negatively charged glutamic acid in the Box C region of BTG2 because the interaction with BTG2 was further reduced in the presence of the PABPC1 R49E variant. This process, in turn, results in a charge reversal of the side chain, compared to the R49A variant, whose side chain has no charge. Apart from Arg-49, other residues which contribute to PABPC1 (1–190) to bind with BTG2 are Glu-24 and Met-46, which were substituted by Ala residues, resulting in weak interactions with GST-BTG2.

Other mutations explored were T48A and T48P. T48P is a cancer-associated mutation in RRM1 of PABPC1, identified from the COSMIC (Catalogue of Somatic Mutations in Cancer) database (Tate et al., 2019). In addition, residue Thr-48 was identified as one among several predicted residues of PABPC1 to interface with BTG2. Therefore, Thr-48 was substituted with Ala and Pro (T48A and T48P). Pulldown analysis conducted for T48A revealed a strong interaction with BTG2 comparable to wild type PABPC1, while T48P revealed reduced ability to interact with BTG2 compared to wild type PABPC1.

A particular characteristic of proline is its conformational rigidity compared to other amino acids. Proline residues can distort secondary elements, including α helices and sheets, located in their vicinity. Thr-48 is located in a PABPC1 loop region; therefore, T48P substitution may affect the loop's local structure or the β-sheets nearby. In this way, the T48P substitution may affect interactions with BTG2 even though it does not make direct contact involving its sidechain with BTG2.

4.3.6. Box C of BTG2 is critical for interacting with BTG2

The pulldown (Figure 4-7) revealed that both BTG1 and BTG2 interacted with RRM1 of PABPC1 (1–190); however, reduced intensity of BTG1 was observed compared to BTG2. While investigating the possible cause for the reduced intensity of BTG1, it was noted that two residues in Box C of BTG1 were not identical to those in BTG2 at positions 125 and 126. BTG2 has Glu-125 and Ala-126, whereas BTG1 has Ala-125 and Ser-126 (Figure 4-15). The dissimilarity in amino acids at the Box C region of BTG1/BTG2 domain may have caused an interaction of BTG1 with PABPC1 that is weaker than BTG2.

In addition, BTG2 residues in the Box C region were substituted by the corresponding TOB1 residues, which does not interact with PABPC1 via residues in this region. Sequence alignment of BTG2 and TOB1 by ClustalW2 revealed that 40% of TOB1 was identical to BTG1/BTG2 (Yang et al., 2008); however, residues in the C-terminus of TOB1 and BTG2 are different. While BTG2 interacts with RRM1 of PABPC1 with residues in Box C, which are located adjacent to the BTG domain, TOB1 has an extended C-terminal region with PAM motifs that can interact with the C-terminal domain of PABPC1 (Ezzeddine et al., 2007). Analysing the ability of BTG2 variants with single amino acid substitutions to interact with PABPC1 disclosed no disruption in interactions with BTG2. However, substituting BTG2 with S118A-I119V or E124L-E125D-A126D amino acids revealed residues required for the interaction.

Finally, it was confirmed that the residues in Box C of BTG2 are the key residues to interact with PABPC1 by performing pulldown using a BTG2 variant lacking Box C. Moreover, the introduction of a stop codon immediately after the BTG2 Box C region also led to a complete disruption in the interaction with PABPC1, indicating that the C-terminal residues of BTG2 do not contribute significantly to the interaction with PABPC1. The results described in this chapter define in more detail the BTG2 residues and PABPC1 residues required for the interaction between PABPC1 and BTG2.

Figure 4-16. Sequence comparison of BTG1 and BTG2. Sequence alignment was carried out using Clustal Omega. The highlighted position (box) reveals the dissimilarity of amino acids in BTG1 and BTG2 at positions 125 and 126.

Chapter 5 – RNA binding and deadenylation activity by PABPC1-BTG2-CNOT7/CAF1

5.1. RNA binding by PABPC1 (1-190)/variants and BTG2 mediated deadenylation by CNOT7/CAF1

CNOT7/CAF1 is a component of the CCR4-NOT complex, classified as a member of the Rnase D superfamily. The protein's active site comprises DEDD residues to coordinate two Mg^{2+} ions (Petit et al., 2012). The CNOT7/CAF1 associates with the CCR4-NOT complex by interacting with the MIF4G domain of the CNOT1 subunit, which functions as a scaffold protein to assemble the complex (Bawankar et al., 2013).

Recent studies have disclosed that the binding of an additional protein to CNOT7/CAF1 can enhance the rate of deadenylation. For example, BTG2 can increase the rate of deadenylation in the presence of PABPC1 (Stupfler et al., 2016). Alternatively, deadenylation by CNOT7/CAF1 was also stimulated by BTG2 in the absence of PABPC1 (Pavanello et al., 2018).

As described in this chapter, wild type PABPC1 (1–190) and variants E24A, I40A, R49A, and R49E were purified. In addition, a cancer-associated variant, PABPC1 T48P, was purified. The purified proteins were subjected to functional analysis using electrophoretic mobility shift assay (EMSA), which was conducted to study the binding ability of PABPC1 (1–190) and variants to poly(A)$_{20}$ RNA. Following EMSA analysis, deadenylation assays were conducted to assess PABPC1 (1–190) and variants E24A, I40A, R49A, and R49E activity in the presence of poly(A)$_{20}$, CNOT7/CAF1, and CNOT7/CAF1-BTG2.

5.1.1. Expression and purification of PABPC1 (RRM1-RRM2) variants

PABPC1 (1–190) and variants were expressed and purified to analyse the activity of PABPC1 variants compared to the wild type. A wild type PABPC1 (1–190) cDNA fragment was cloned into the prokaryotic expression vector pQE80L (Pavanello et al., 2018). This plasmid was used as a template for site-directed mutagenesis to obtain single amino acid variants E24A, I40A, R49A and R49E. Site-directed mutagenesis was confirmed by Sanger sequencing before using the plasmids for expression and purification (Figure 5-1A).

After transformation of *E. coli* BL21 (DE3), a single colony with pQE80L-His-PABPC1, E24A, I40A, R49A, and R49E (1–190) was used for expression. After the expression of the PABPC1 variants was induced by the addition of IPTG, lysates were prepared by sonication and analysed by 14% SDS-PAGE (Figure 5-1B).

The soluble lysate obtained after centrifugation was then loaded onto a HisTrap FF pre-packed column (GE) for affinity purification. After washing, bound proteins were collected in three fractions of 1 mL. After buffer exchange using PD-10 desalting columns (GE), the protein preparations were analysed by SDS-PAGE and staining by Coomassie Brilliant Blue (Figure 5-1C). The PABPC1 variants were purified to near homogeneity (> 95%) and used for functional assays.
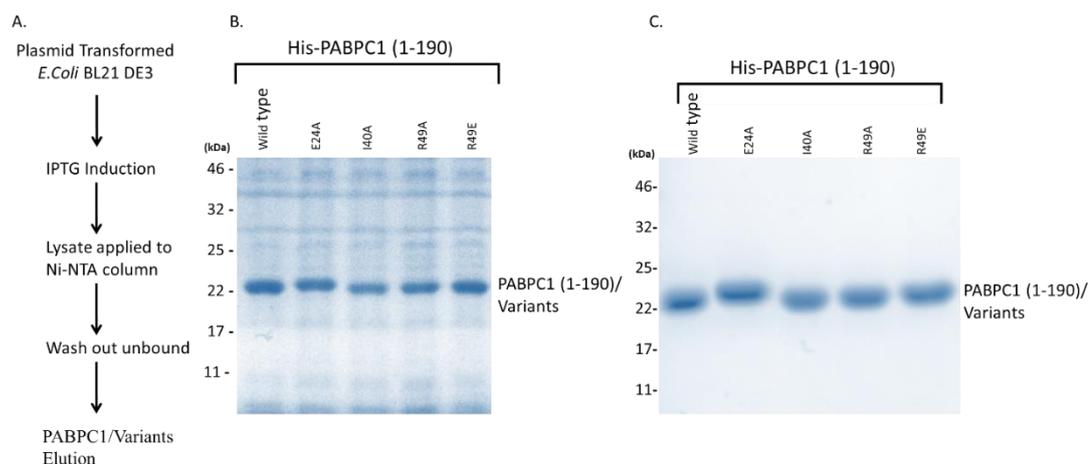
Figure 5-1. Expression and purification of His-PABPC1 (1-190) variants. Crude lysate of overnight grown culture purified by AKTA. **A)** An outline of a purification strategy to purify PABPC1 variants. **B)** Lysates obtained after expression of PABPC1 variants were analysed by 14% SDS-PAGE and stained with Coomassie Brilliant Blue. **C)** Purified PABPC1 (1-190) variants were analysed by 14% SDS-PAGE and stained with Coomassie Brilliant Blue.

5.1.2. Binding of PABPC1 variants to poly(A)$_{20}$ RNA

In the next step, EMSA assays were conducted to verify the binding ability of purified wild type PABPC1 (1–190) and the E24A, I40A, R49A, and R49E variants to poly(A)$_{20}$ RNA. A 5′-fluorescein-labelled synthetic oligonucleotide containing 20 adenosine residues poly(A)$_{20}$ was used as a substrate. Binding reactions were set up using 0.2 μM poly(A)$_{20}$ substrate and 0-4.0 μM PABPC1 (1–190). As shown in Figure 5-2A, limited binding of the poly(A)$_{20}$ RNA was observed when 0.8 μM PABPC1 (1–190) was added, while a complete shift of the unbound probe was observed with the highest concentrations. Similar observations were made using the E24A, I40A, R49A, and R49E variants (Figure 5-2B-E).

The percentage of bound RNA was plotted against the concentration of PABPC1 (1–190) to estimate the K_d of the PABPC1 variants for poly(A)$_{20}$ RNA (Figure 5-3). This was done by densitometric analysis of the gel images shown in Figure 5-2 and in replicated experiments. The data were fitted to a sigmoid function that made it possible to obtain an estimate of the K_d values for wild type PABPC1 (1–190) and each single amino acid substitution variant (Figure 5-3).

This analysis showed that the K_d value for wild type PABPC1 (1–190, $n = 3$, 1.9 μM; Figure 5-3A) was comparable to those of the E24A variant ($n = 2$, K_d 1.7 μM; Figure 5-3B), I40A ($n = 2$, K_d 1.6 μM; Figure 5-3C), R49A ($n = 3$, K_d 1.1 μM; Figure 5-3D) and R49E ($n = 3$, K_d 2.1 μM; Figure 5-3E). These results indicate that the designed PABPC1 variants containing single amino acid substitutions displayed RNA binding activity comparable to wild type PABPC1 but displayed a reduced affinity for BTG2.
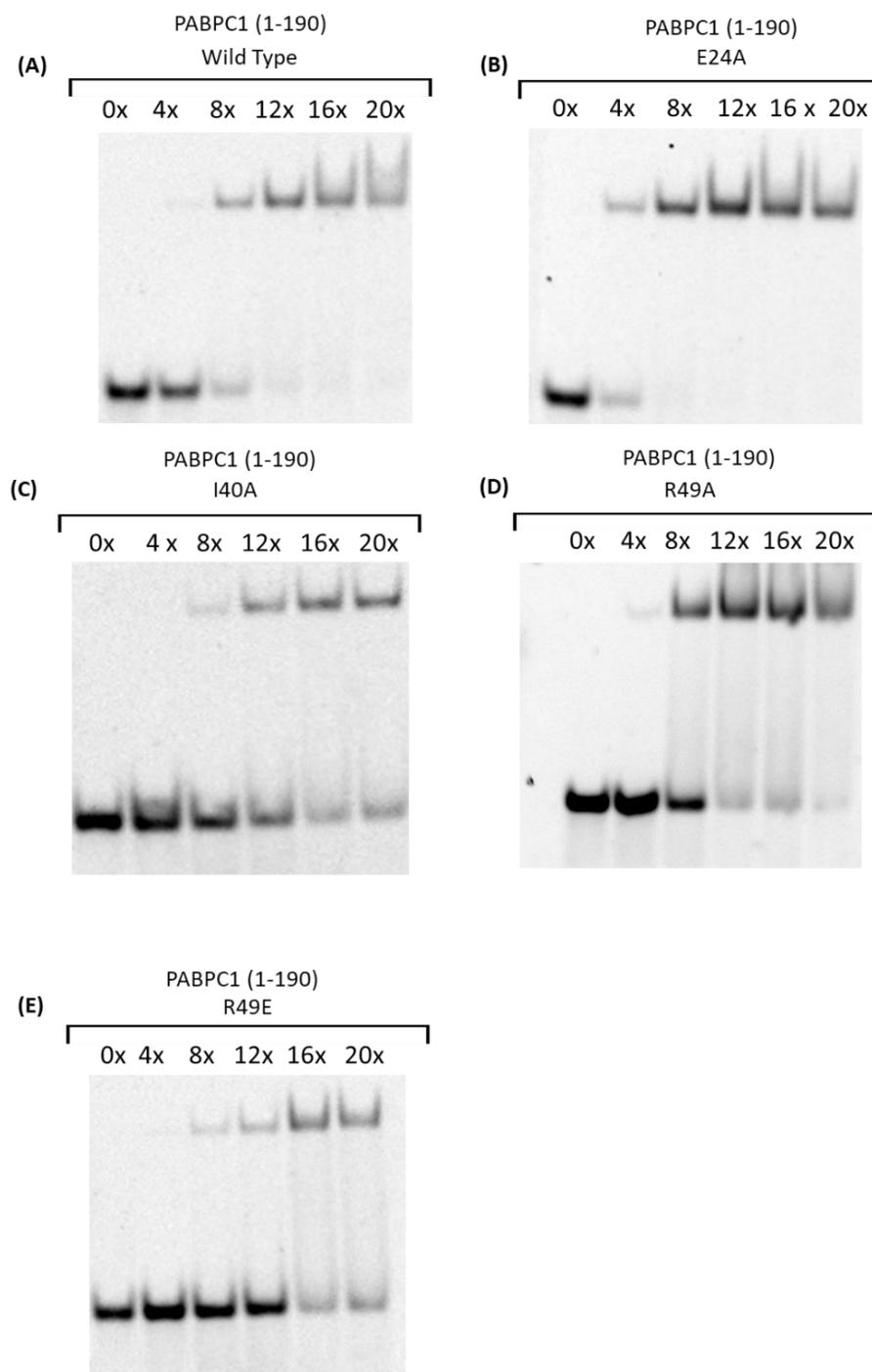
Figure 5-2. Binding of wild type PABPC1 (1-190) and PABPC1 variants to Flc-A_{20}. A) Analysis of PABPC1 (1-190) wild type binding to poly(A)$_{20}$. **B)** Analysis of PABPC1 (1-190) E24A binding to poly(A)$_{20}$. **C)** Analysis of PABPC1 (1-190) I40A binding to poly(A)$_{20}$. **D)** Analysis of PABPC1 (1-190) R49E binding to poly(A)$_{20}$. **E)** Analysis of PABPC1 (1-190) R49A binding to poly(A)$_{20}$. Excesses of PABPC1/variants (0-20x) were incubated with RNA substrate (200 nM) for 20 min at room temperature. Then, nondenaturing loading dye was added, and reactions were analysed by 8% PAGE.
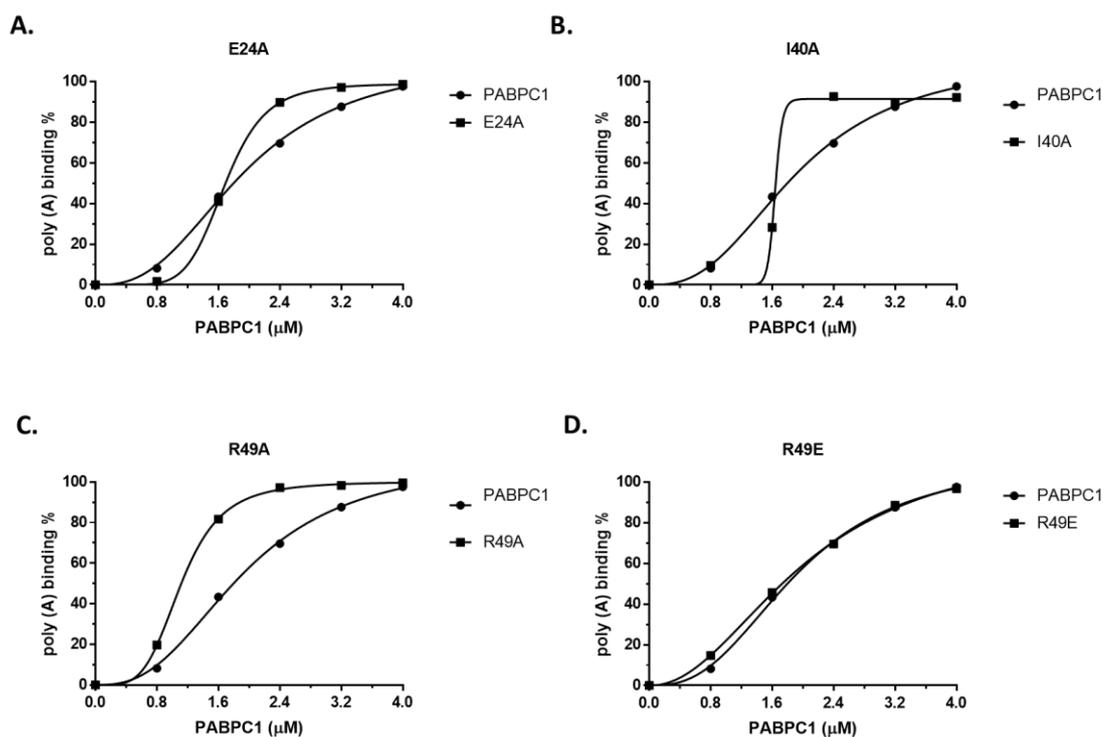
Figure 5-3. Estimated dissociation constant of binding of wild type PABPC1 (1-190) and PABPC1 variants and poly(A)$_{20}$ RNA. (A-D) The percentage of bound RNA was plotted against the concentration of PABPC1 (1-190) as determined using densitometric analysis of replicate experiments mentioned in section 5.2.2. Data were fitted to a sigmoid binding function using GraphPad Prism *9.0*.

5.1.3. Purification of CNOT7/CAF1 and CNOT7/CAF1-BTG2

Having identified PABPC1 residues responsible for interfacing with BTG2 but not interfering with poly(A)$_{20}$ RNA binding, the next aim was to investigate the deadenylation activity of purified CNOT7 and CNOT7/CAF1-BTG2 in the presence of wild type PABPC1 (1–190) and single amino acid substitution variants.

To this end, CNOT7/CAF1 and CNOT7/CAF1-BTG2 were purified by Dr Lorenzo Pavanello, the University of Nottingham, applying the steps shown in Figure 5-4A. Soluble lysates were prepared after expression in *E. coli* BL21 (DE3). The soluble lysate was loaded on a HisTrap FF pre-packed column (GE), and the peak fraction was used in subsequent steps. One microlitre of the sample containing purified CNOT7/CAF1 was analysed by SDS-PAGE (Figure 5-4B). The BTG2-CNOT7/CAF1 dimer was further purified by size-exclusion chromatography. Approximately 1 mL of the sample was applied to an SEC column (GE Sephacryl S-100 HR), and 2 mL fractions were collected. Typically, 4–5 peak fractions were pooled and re-concentrated. SDS-PAGE analysed one microlitre sample (Figure 5-4B). The purified CNOT7/CAF1 and CNOT7/CAF1-BTG2 dimer were nearly homogeneous (> 90%) and were used to perform deadenylation assay in conjunction with wild type PABPC1 and PABPC1 variants.
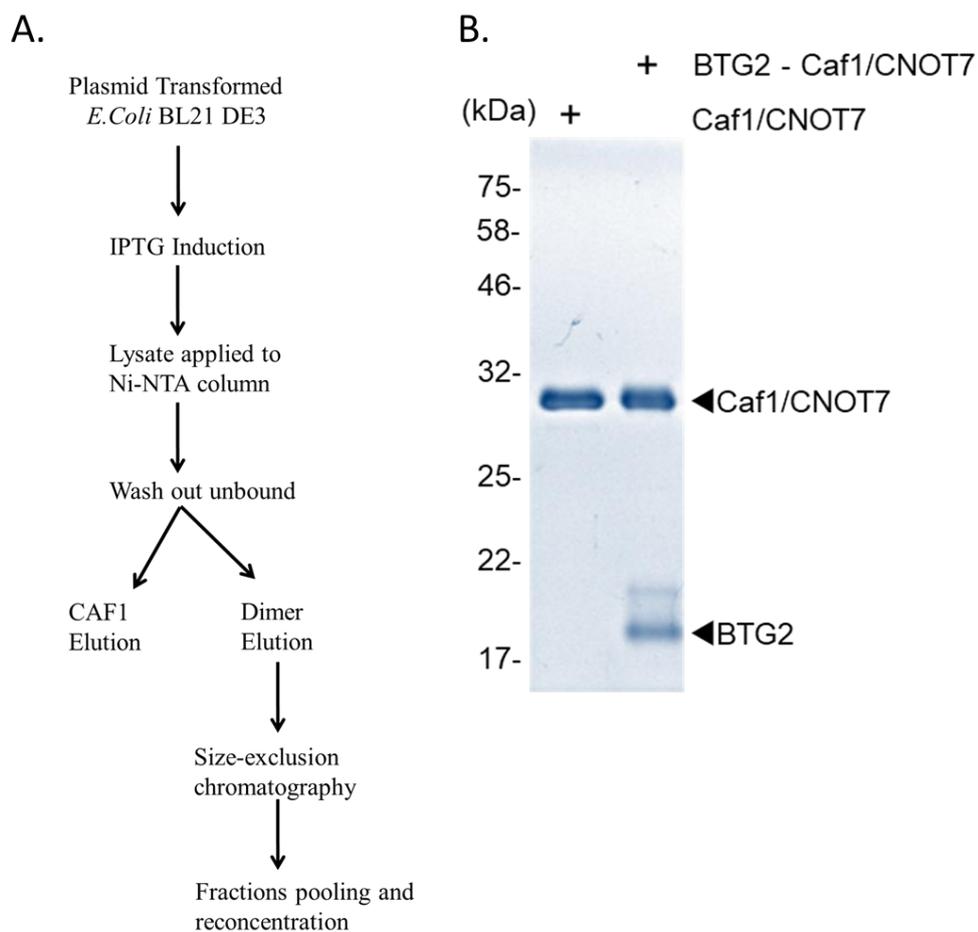
Figure 5-4. Purification of CNOT7/CAF1 and CNOT7/CAF1-BTG2 A) Outline of the purification strategy designed to express and purify CNOT7/CAF1 and the CNOT7/CAF1-BTG2 dimer. **B)** The purified deadenylase enzymes CNOT7/CAF1 and BTG2-CNOT7/CAF1 using 14% SDS-PAGE. Proteins were visualised by Coomassie staining.

5.1.4. Analysis of the catalytic activity of CNOT7/CAF1 and CNOT7/CAF1-BTG2 in the presence of poly(A)$_{20}$ and PABPC1 variants

Next, the wild type PABPC1 (1–190) and the variants PABPC1 E24A, I40A, R49A, and R49E were evaluated using deadenylation reactions. Thus, an excess of PABPC1 variants was incubated with poly(A)$_{20}$ before adding CNOT7/CAF1 or CNOT7/CAF1-BTG2 as described previously (Pavanello et al., 2018). Reactions were incubated for up to 60 minutes at 30°C. Aliquots were taken at specific time points, mixed with RNA denaturing dye, and subsequently analysed with PAGE (Figure 5-5).

The assay revealed that deadenylation activity by CNOT7/CAF1 in the absence of BTG2 was barely detected in the presence of poly(A)$_{20}$-PABPC1 (1–190) substrate (Figure 5-5A). In contrast, deadenylation by the BTG2-CNOT7/CAF1 complex was readily observed using poly(A)$_{20}$ PABPC1 (1–190), and the substrate was completely degraded after 20 minutes of incubation with CNOT7/CAF1-BTG2 (Figure 5-5A).

Using the poly(A)$_{20}$-PABPC1 E24A substrate, a barely detectable level of deadenylation was observed in the presence of CNOT7/CAF1; however, CNOT7/CAF1 completely degraded the substrate after 60 minutes of incubation (Figure 5-4B). In the presence of the poly(A)$_{20}$-PABPC1 E24A substrate, the rate of deadenylation by CNOT7/CAF1-BTG2 was similar to that observed using the poly(A)$_{20}$-wild type PABPC1 (1–190) substrate (Figure 5-5B).

Deadenylation by CNOT7/CAF1 was barely detected using the poly(A)$_{20}$-PABPC1 I40A substrate. However, after 60 minutes of incubation, activity was observed, and the substrate was not wholly degraded, which is similar to poly(A)$_{20}$ containing wild type PABPC1 (Figure 5-5C). In the presence of the CNOT7/CAF1-BTG2 complex, a slow activity rate was observed; however, the substrate was degraded after 60 minutes of incubation.

The activity was barely detected using the poly(A)$_{20}$-PABPC1 R49A substrate when treated with CNOT7/CAF1 (Figure 5-5D). While degradation was readily observed in RNA substrate bound by PABPC1 R49A in the presence of BTG2-CNOT7/CAF1, the substrate was completely degraded after 20 minutes of incubation (Figure 5-5D). Finally, when poly(A)$_{20}$-PABPC1 R49E was used as a substrate, activity was barely detected with CNOT7/CAF1, and a substantial decrease in the substrate by BTG2-CNOT7/CAF1 was observed as compared to the wild type complex (Figure 5-5E).

Taken together, these results suggest that the residual ability of the PABPC1 E24A, I40A, and R49A variants to bind BTG2 is sufficient to stimulate deadenylation by the BTG2-CNOT7/CAF1 dimer. By contrast, in the presence of PABPC1 R49E, a substantial decrease in deadenylation by BTG2-CNOT7/CAF1 was observed compared to the wild type complex (Figure 5-5A). This agrees with the observation that the PABPC1 R49E variant could not bind GST-BTG2 in the pulldown assay. Thus, amino acid Arg-49 of PABPC1 is located on the interface of PABPC1 and BTG2. In addition, the R49E substitution results in a PABPC1 variant that binds poly(A) RNA comparable to wild type but cannot bind BTG2 and stimulate deadenylation by BTG2-CNOT7/CAF1.
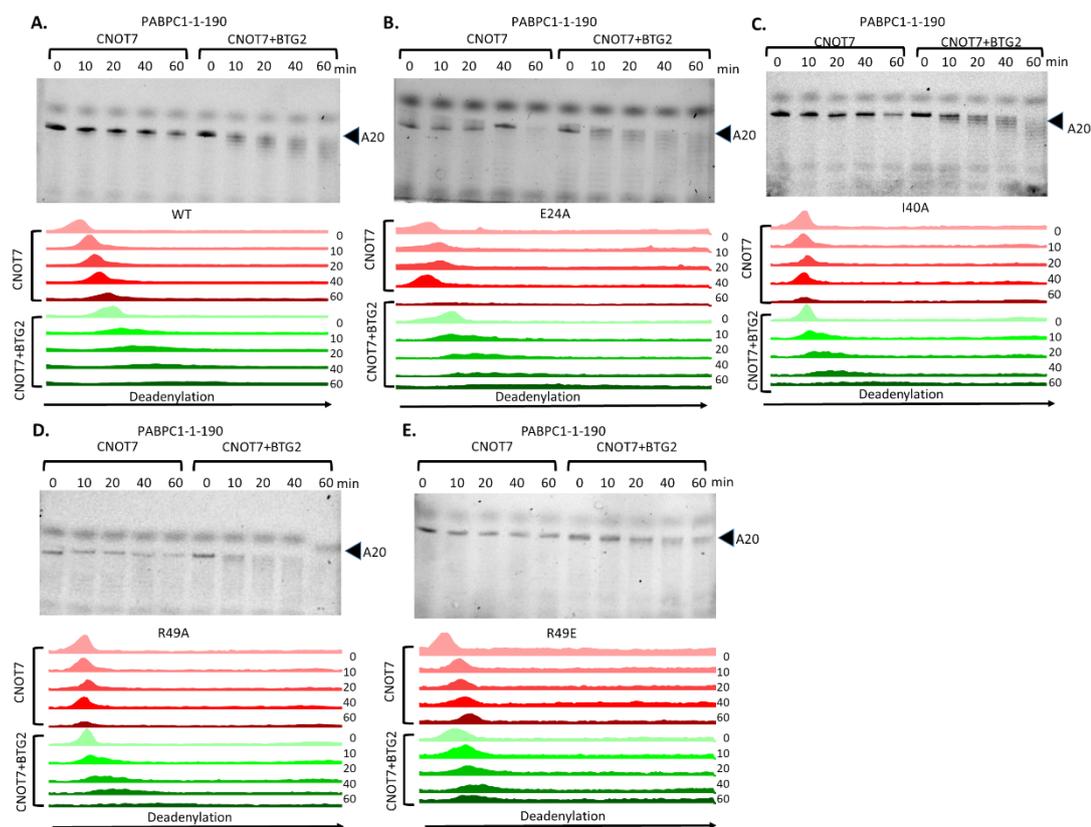
Figure 5-5. Reduced amount of deadenylation activity determined by PABPC1 R49E. The deadenylation activity of the CNOT7/CAF1 enzyme compared with the CNOT7/CAF1-BTG2 complex. Reactions were carried out in the presence of poly(A)$_{20}$ and **(A)** wild type PABPC1(1-190); **(B)** PABPC1 (1-190) E24A; **(C)** PABPC1 (1-190) I40A; **(D)** PABPC1 (1-190) R49A; **(E)** PABPC1(1-190) R49E. An eight-fold molar excess of purified PABPC1 proteins was incubated with fluorescein labelled poly(A)$_{20}$ (200nM) at 30° C before adding the enzyme preparations. Reaction products were separated by 50% urea/20% PAGE. Gel images were quantified using the ImageJ package.

5.1.5. Analysis for T48P cancer-associated mutation of PABPC1

Since the pulldown analysis using cancer associated PABPC1 variant T48P revealed a weak interaction with BTG2 (Chapter 4), the PABPC1 T48P variant was subsequently purified to determine its binding ability to poly(A)$_{20}$ RNA. Consequently, deadenylation assays were conducted to study the degradation rate for poly (A$_{20}$) substrate bound T48P. The grown cells in the LB medium were harvested by centrifugation, re-suspended in lysis buffer, and lysed by sonication. The soluble lysate was isolated by centrifugation and then loaded on a HisTrap FF pre-packed column (GE). Three elution fractions of 1 mL were collected using elution buffer, and protein buffer was exchanged to the final desalting buffer using PD-10 desalting columns, and the desalted proteins were analysed by SDS-PAGE (Figure 5-6A).

Following protein purification of PABPC1 (1–190) T48P using the procedure described in section 5.1.1 (Figure 5-6A), binding reactions for EMSA analysis were set up as described in Section 5.1.2. After a one-hour incubation at room temperature of Flc-poly(A)$_{20}$ with wild type PABPC1 and PABPC1 T48P, native loading dye was added, and the samples were analysed by 8% PAGE (Figure 5-6C, D). The results revealed that the PABPC1 T48P variant was able to shift poly(A)$_{20}$ RNA similar to wild type PABPC1 (1–190; Figure 5-6C and D).

Next, the gel images were subjected to densitometry analysis to determine the bound poly(A)$_{20}$ RNA plotted versus the concentration of PABPC1 (Figure 5-6E). The data were fitted to a sigmoid binding equation to estimate K_d values, which confirmed that wild type PABPC1 (1–190; K_d 1.9 μM) is bound with similar affinity to poly(A)$_{20}$ RNA as PABPC1 (1–190) T48P ($n = 2$, K_d 1.6 μM, Figure 5-6E).

Next, the activity of the PABPC1 (1–190) T48P variant in deadenylation assays was evaluated using purified CNOT7/CAF1 and CNOT7/CAF1-BTG2 (Figure 5-6B). First, an eight-fold molar excess of wild type PABPC1 (1–190) and PABPC1 (1–190) T48P was incubated with poly(A)$_{20}$ substrate. Then, CNOT7/CAF1 and CNOT7/CAF1-BTG2 were added, and aliquots were taken at specific time points. After mixing the samples with an RNA denaturing loading buffer, samples were subsequently analysed using denaturing gel electrophoresis (Figure 5-6F, G). The assays revealed that deadenylation by CNOT7/CAF1 in the absence of BTG2 was barely detectable (Figure 5-6F) as expected.

In contrast, enhanced deadenylation activity by BTG2-CNOT7/CAF1 is readily observed in the presence of wild type PABPC1 (1–190; Figure 5-6F). Observing the rate of deadenylation by CNOT7/CAF1 in the presence of PABPC1 (1–190), T48P showed no activity (Figure 5-6G); but deadenylation by CNOT7/CAF1-BTG2 was readily observed and similar to that observed in the presence of wild type PABPC1 (1–190; Figure 56G). Therefore, while PABPC1 T48P shows reduced binding to BTG2, PABPC1 T48P can bind poly(A) RNA with similar affinity as wild type PABPC1, and the residual ability of BTG2 to bind PABPC1 T48P is sufficient to stimulate deadenylation by CNOT7/CAF1.
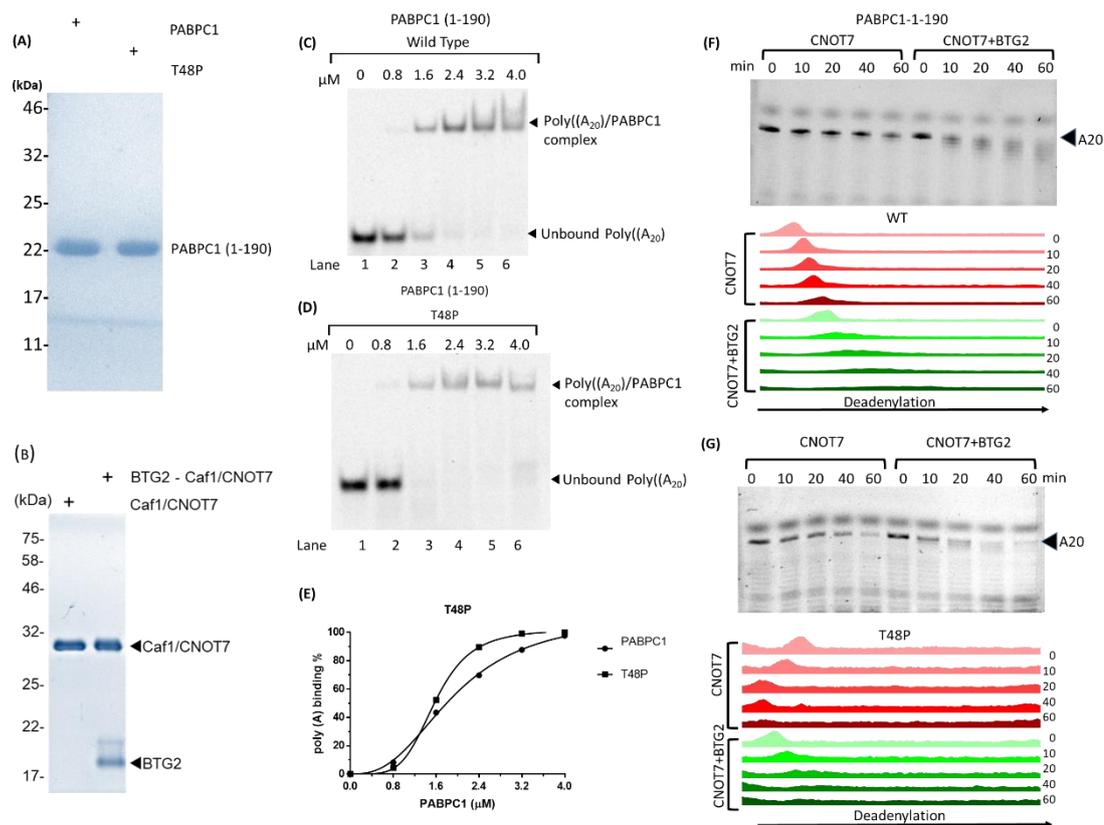
Figure 5-6. Analysis of cancer mutation T48P. A) PABPC1 (1-190) wild type and T48P variant was purified by affinity chromatography and analysed using 14% SDS-PAGE and staining by Coomassie Brilliant Blue. **B)** Purified CNOT7/CAF1 and BTG2-CNOT7/CAF1 Proteins were separated by 14% SDS-PAGE and stained with Coomassie Brilliant Blue. **C)** Analysis of PABPC1 (1-190) wild type binding to poly(A)$_{20}$. **D)** Analysis of PABPC1 T48P (1-190) binding to poly(A)$_{20}$. Excesses of PABPC1(0-20x) were incubated with RNA substrate (200 nM) for 20 min at room temperatures. Then, nondenaturing loading dye was added, and reactions were analysed by 8% PAGE. **E)** The percentage of bound RNA was plotted against the concentration of PABPC1 (1-190) as determined using densitometric analysis of replicate experiments mentioned in section 5.2.5. Data were fitted to a sigmoid binding function using GraphPad Prism *9.0*. **F)** Deadenylase activity of CNOT7/CAF1 and BTG2-CNOT7/CAF1 in the presence of poly(A)$_{20}$ bound PABPC1 and **G)** Deadenylase activity of CNOT7/CAF1 and BTG2-CNOT7/CAF1 complex in the presence of poly(A)$_{20}$ bound PABPC1 T48P. PABPC1 T48P The purified proteins were incubated with an equimolar amount of fluorescein labelled poly(A)$_{20}$ (200nm) at 30° C. The RNA was separated by 50% urea: The ImageJ package quantified the Gel images.

5.2. Discussion

5.2.1. Protein purification and binding of PABC1 variants to poly(A)$_{20}$

The N-terminal RRM1 and RRM2 domains of PABPC1 are associated with mRNA's poly(A) tail (Safaee et al., 2012). In addition, the RRM1 of PABPC1 is known to bind Box C of BTG2 (Stupfler et al., 2016), but the residue interfacing has not been reported. Several residues of PABPC1 that interact with BTG2 were discovered, as described in Chapter 4. The variants that displayed reduced and strongly reduced interactions with BTG2 in pulldown analysis were used as a criterion for purifying variants of PABPC1 to assess their activity in RNA binding and mRNA deadenylation. Therefore, wild type PABPC1, and the E24A, I40A, R49A, and R49E (1–190) variants were expressed in *E. coli* and purified using affinity chromatography for further analysis.

The binding of the selected PABPC1 variants to poly(A) RNA was investigated using EMSA analysis. These experiments revealed that the selected variants were efficiently bound to poly(A)$_{20}$, and, thus, the amino acid substitutions introduced at positions Glu-24, Ile-40, and Arg-49 were not located at the RNA interface. A 200nM concentration of poly(A)$_{20}$ was used for RNA binding assay; this agrees with the results from a publication by Pavanello et al. (2019). The densitometry analysis showed that K_d $_=$1.6µM of wild type PABPC1 (1–190) was sufficient to bind poly(A)$_{20}$ in assay conditions. A similar study (Safaee et al., 2012) that revealed eIF4G-PABPC1-RNA binding disclosed that Kd =0.6µM of PABPC1 (1–190) was adequate to bind 2 pM poly(A) RNA. Thus, the study revealed that a low concentration of PABPC1 and poly(A) was sufficient to bind each other, and eIF4G bound to RRM2 of PABPC1 enhances PABPC1 affinity with RNA.

In addition to studying the RNA binding property of PABPC1 variants, EMSA analysis was conducted for PABPC1 T48P, a cancer-associated mutation in RRM1. The EMSA performed with T48P revealed its ability to bind poly(A)$_{20}$, similar to wild type PABPC1. Moreover, it was shown that Thr48 is not located at the RNA interface.

5.2.2. Effect of PABPC1 variants on deadenylation by CNOT7/CAF1

BTG2 is vital in regulating deadenylation by bridging PABPC1 and CNOT7/CAF1, making the 3′ end of the poly(A) tail (Stupfler et al., 2016). BTG2 is associated with CNOT7/CAF1 by Box A and Box B (Winkler, 2010), and BTG2 interacts with PABPC1 through residues in Box C (Stupfler et al., 2016). In the absence of BTG2, deadenylation by CNOT7/CAF1 was very low and barely detectable with poly(A)$_{20}$ bound by wild type PABPC1 or PABPC1 variants. This agrees with Pavanello et al. (2019) and confirms that PABPC1 can inhibit the deadenylase activity of CNOT7/CAF1. In the presence of BTG2, however, an enhanced rate of deadenylation by CNOT7/CAF1 was observed, in agreement with the notion that BTG2 can recruit CNOT7/CAF1 to the 3′ end of the poly(A) tail bound by wild type PABPC1 (1–190) (Stupfler et al., 2016, Pavanello et al., 2018).

When single amino acid alanine substitutions at positions Glu-24, Ile-40, and Arg-49 of PABPC1 RRM1 (E24A, I40A and R49A) were analysed, deadenylation by BTG2-CNOT7/CAF1 was similar to that observed in the presence of wild type PABPC1. This was unexpected since BTG2 showed a reduced ability to bind these variants of PABPC1; instead, a reduced rate of deadenylation by BTG2-CNOT7/CAF1 had been expected. Thus, these results suggest that the residual ability of the BTG2 to bind selected PABPC1 variants is sufficient to stimulate deadenylation by CNOT7/CAF1 in the presence of BTG2.

An alternative variant was explored at position Arg-49, where this residue was substituted with Glu (R49E), inverting the charges in the sidechain. This variant had a very low affinity for BTG2 in the pulldown experiments (Chapter 4). This variant could bind poly(A) RNA similar to wild type. However, unlike wild type PABPC1, the R49E variant was unable to stimulate deadenylation by CNOT7/CAF1 in the presence of BTG2. This suggests that substituting Arg-49 with negatively charged Glu acid could have repulsed electrostatically and reduced deadenylation consistent with the strongly reduced ability of this variant to bind BTG2.

Cancer-associated PABPC1 variant T48P was also investigated. The characteristic feature of proline is confirmational rigidity; therefore, Thr-48 substituted with Pro could deform secondary elements, including α-helix and β-sheets. The EMSA performed for cancer mutation T48P disclosed that it could efficiently bind the poly(A)$_{20}$ substrate and not on the poly(A) interface. However, the variant was able to stimulate deadenylation by CNOT7/CAF1 in the presence of BTG2; therefore, the reduced affinity for BTG2 was sufficient to stimulate CNOT7/CAF1 activity under the conditions used.

Chapter 6 – Conclusion

6.1. Concluding remarks

Deadenylation is the process of the shortening and removal of poly(A) tail from the 3′ end of mRNA. Deadenylation is a rate-limiting step in mRNA decay in eukaryotes. The major components which play a crucial role in deadenylation are the CCR4-NOT complex, PABPC1, and the PAN2-PAN3 deadenylase in eukaryotes. While significant progress has been made, the mechanism of how enzymes target and regulate the removal of poly(A) tail is not well understood.

In the experimental work described in this thesis, I focused on three proteins involved in shortening the mRNA poly(A) tail. PABPC1 recognises RNA via four RNA-recognition motifs (RRMs). The structure of RRM 1 and RRM2 with poly(A) RNA has been described (Safaee et al., 2012). In addition to RNA, the N-terminal RRM domain of PABPC1 (RRM1) can also interact with BTG1/BTG2, which are members of the small BTG/TOB family of proteins (Winkler, 2010) Though BTG1/BTG2 are 40% identical to TOB1, the BTG1/BTG2 interaction with PABPC1 is regulated by residues in the Box C region (Stupfler et al., 2016), which is not conserved in TOB1 and other BTG/TOB proteins. Instead, a PAM motif in the extended C-terminal region of TOB1/TOB2 mediates interactions with the C-terminal domain of PABPC1 (Okochi et al., 2005, Ezzeddine et al., 2007).

A recent study revealed that BTG2 could stimulate deadenylation by CNOT7/CAF1 in the presence of PABPC1 (Stupfler et al., 2016). In addition, another recent study reported an increased rate of deadenylation by BTG2-CNOT7/CAF1 compared to CNOT7/CAF1 either in the presence or absence of PABPC1 (Pavanello et al., 2018). This suggests that a quaternary complex is formed by the BTG2, PABPC1, poly(A) RNA and CNOT7/CAF1 formations, resulting in enhanced deadenylation compared to a situation in which PABPC1 and BTG2 are absent. This study focused on three different proteins, PABPC1, BTG2, and CNOT7/CAF1 and their role in mRNA degradation. Though studies have disclosed that Box C of BTG2 is responsible for interacting with the PABPC1 RRM1 domain, the residues of PABPC1 involved in mediating BTG2 interaction were yet to be discovered. Therefore, the present work aimed to identify residues of PABPC1 that are important for the interaction with BTG2

6.1.1. Computational approaches to identify PABPC1 residues mediating interactions with BTG2

The study aimed to create a computational model of PABPC1-BTG2 from the available structure of PABPC1 (PDB 4F02) and BTG2 (3DJU) to identify the PABPC1 residues mediating interaction with BTG2. PABPC1 (PDB 4F02) structure with RNA bound to RRM1-RRM2 was utilised for CPORT analysis. The CPORT analysis for PABPC1 revealed conserved surface residues in a consensus manner.

Since PABPC1 (1–190) used for CPORT analysis had poly(A) RNA bound to its structure, it was essential to avoid PABPC1 residues interfacing with RNA before docking; therefore, the PABPC1 structure was gradually examined in PISA and residues interfacing with RNA were eliminated before docking in HADDOCK.

The candidate residues in PABPC1 void of RNA interfaces along with residues in Box C of BTG2 were used as data for docking in HADDOCK. Docking by HADDOCK generated a series of clusters, each depicting the PABPC1-BTG2 model and the clusters with the best score and size were chosen for further analysis. The obtained computational models created a quarternary structure of PABPC1-RNA-BTG2-CNOT7/CAF1, showing how BTG2 improves poly(A) accessibility to the active site of CNOT7/CAF1.

Though the PABPC1-BTG2 model was obtained computationally, the interaction was experimentally validated to trace putative residues responsible for interfacing with BTG2 based on the input for HADDOCK for interaction studies. The putative residues were traced by analysing the clusters in PISA, and the candidate residues for site-directed mutagenesis were chosen from cluster 1 and cluster 2. Site-directed mutagenesis was conducted for the candidate residues. It had experimentally validated that the residues responsible for interaction with BTG2 is crucial.

6.1.2. Experimental validation of residues of PABPC1 responsible for interfacing BTG2

Based on the computational analysis, residues Asp-21, Glu-24, Glu-29, Ile-40, Met-46, Arg-44, Arg-49A, and Arg-50 were predicted to interface with residues in BTG2 Box C. To validate the role of these amino acids; Ala substitutions were introduced at these positions using site-directed mutagenesis. In addition, the R49E single amino acid substitution was also analysed. The PABPC1-BTG2 interactions were validated by pulldown analysis. After optimisation of experimental conditions, including evaluating various types of beads (Ni^{2+}-NTA agarose, Co^{2+}-agarose, Ni^{2+}-NTA magnetic beads), and other factors such as temperature, duration of incubation and order of addition, a robust interaction assay was developed that minimised non-specific binding. The pulldown analysis was then performed to validate the interfacing residues of PABPC1 with BTG2 and confirmed a role for residues Glu-24, Ile-40, Arg-44, Arg-49, Arg-50, and Met-46 of PABPC1 to interface with BTG2. In addition, a substantial reduction in interaction with variant R49E was observed due to replacing the positive charge of the arginine sidechain with a negative charge of glutamate. This may suggest that the positive charge of Arg-49 could be interfacing with negatively charged residues in Box C of BTG2.

6.1.3. Functional analysis of single amino acid substitution variants of PABPC1

To confirm that the PABPC1 (1–190) variants were able to bind RNA, the variants such as PABPC1 (1–190) E24A, PABPC1 (1–190) I40A, PABPC1 (1–190) R49A, and PABPC1 (1–190) R49E, were selected for protein purification, which is consistent with the notion that PABPC1 can bind poly(A) and BTG2 at the same time.

The RNA binding assay performed with the PABPC1 (1–190) E24A, PABPC1 (1–190) I40A, PABPC1 (1–190) R49A, and PABPC1 (1–190) R49E revealed their ability to bind RNA. Further densitometric analysis of the gel images was done by plotting the percentage of bound RNA against the concentration of PABPC1 (1–190) to estimate the K_d of PABPC1 variants for $poly(A)_{20}$ RNA; notably, the variant PABPC1 (1–190) R49E binds to RNA as PABPC1 (1–190) wild type.

Further, to assess the importance of the PABPC1-BTG2 interaction in deadenylation by CNOT7/CAF1, deadenylation assays were conducted. The deadenylation assay was conducted with poly(A) bound PABPC1 (1–190) E24A, PABPC1 (1–190) I40A, and PABPC1 (1–190) R49A, stimulated deadenylation by CNOT7/CAF1-BTG2 as PABPC1. While the deadenylation was performed with PABPC1 (1–190), R49E did not stimulate deadenylation by CNOT7/CAF1-BTG2 compared to wild type PABPC1 (1–190). This states that Arg49 appeared to be free of RNA binding and is a critical residue for binding BTG2 and stimulating deadenylation by CNOT7/CAF1.

6.1.4 Analysis for cancer mutation PABPC1 (1-190) T48P

In addition to determining residues that regulated interaction with BTG2, a cancer mutation T48P in RRM1 of PABPC1 was analysed. This variant is reported in the COSMIC database (Tate et al., 2019). T48P is a variant identified in head, neck, and thyroid cancer (Martin et al., 2014). The pulldown analysis performed with PABPC1 (1–190) T48P revealed reduced ability to bind BTG2. Since PABPC1 (1–190) T48P showed reduced ability to bind to BTG2, the variant was purified to test its ability to bind to RNA compared to wild type PABPC1 (1–190). The RNA binding assay performed with cancer mutation PABPC1 (1–190) T48P revealed that it could effectively bind to RNA and is not located on the RNA interfacing region. Further densitometric analysis of the gel images was conducted by plotting the percentage of bound RNA against the concentration of PABPC1 (1–190) to estimate the K_d of PABPC1 (1–190) T48P for poly(A)$_{20}$ RNA. Notably, the variant PABPC1 (1–190) T48P binds to RNA as a PABPC1 (1–190) wild type. Therefore, the deadenylation assay conducted with poly(A) bound PABPC1 (1–190) T48P stimulated deadenylation by CNOT7/CAF1-BTG2 as PABPC1. The result from the deadenylation assay suggests that the PABPC1 (1–190) T48P reduced affinity for BTG2 was sufficient to stimulate CNOT7/CAF1 activity.

6.1.4. Future outlook

Protein interactions are either stable or transient in regulating biological functions, and it is rare for a protein to act alone. Characterising protein and its interacting partners is essential to understanding the biological role of proteins in a cell. The objectives set in this study to characterise proteins of interest and understand their ability to associate with their interaction partners were approached strategically.

This study exposed a binding model for PABPC1-BTG2 by computational analysis and identified its interfacing residues. The results obtained from the experiments conducted in characterising PABPC1-BTG2 interactions, identifying PABPC1 interfacing residues, and experimental validations to study interactions by introducing point mutations and functional assays could lead to a wide range of future studies. Though a broad spectrum is covered in this study, a few questions are yet to be answered. These questions can be strategically addressed in future studies by addressing analytical techniques such as crosslinker and x-ray crystallography, followed by mammalian cell-based protein interaction and activity.

The PABPC1 amino acid interfaces derived from the PABPC1-BTG2 computational model can be verified by performing a cross-linker assay. In-depth structural information can be obtained from cross-linking experiments combined with mass spectrometry-based on three principles: (a) mild conditions are preferred for cross-linking to maintain the conformation of the native protein and its complexes; (b) mixtures are analysed by mass spectrometry; and (c) the peptides cross-linked from mass spectra are identified, and the cross-linked amino acid residues are mapped from a protein of interest (Jones et al., 2019).

An alternative method to determine protein structure is X-ray crystallography—a method that uses X-ray to study the arrangement and position of atoms in crystals. Crystallography is a powerful tool to obtain structural data for biological molecules and aid in understanding protein recognition and binding of ligand molecules at the atomic level.

The advantages of using X-ray crystallography include a high yield of atomic resolution without being limited to a molecular weight of proteins. Crystallography is also suitable for predicting the structure of macromolecules, membrane proteins, and water-soluble proteins; however, the technique also has several disadvantages. For example, the sample must be crystallized first, and crystallizing protein samples with sizeable molecular weight can be difficult. Crystallizing membrane proteins are more challenging due to their large size and low solubilisation. In order to allow appropriate diffraction, only a single crystal must be obtained in an organised manner. Finally, the protein's three-dimensional structure is represented only in static form rather than a dynamic core (Acharya and Lloyd, 2005).

The process is divided into four steps: (a) The first step involves obtaining a high-quality single crystal of the protein of interest. (b) This is followed by diffraction experiments that immobilise crystal by an intense X-ray beam, producing diffraction patterns recorded as data. (c) Subsequently, the data generated by the diffraction pattern are combined by various methods of structural analysis and data fitting. (d) The last step involves the atomic arrangement of the model in the crystal-based electron density map (Acharya and Lloyd, 2005).

Moreover, PABPC1-BTG2 interaction can be studied in vitro in the mammalian cell line without being confined to structural analysis for future work. The ability of the lymphoma associated amino acid substitution in BTG1 to inhibit cell cycle progression has been studied in the past using the HEK293 cell line (Almasmoum et al., 2021); similarly, HEK293 cells were used for stable expression of CNOT7/CAF1 in determining PABPC1-BTG2-CNOT7/CAF1 interaction and deadenylation activity (Amine et al., 2021; Almasmoum et al., 2021).

Therefore, the HEK293 cell line can be used to study PABPC1-BTG2 interaction in the future. A disadvantage is that induced expression of PABPC1 mutants may result in repression of endogenous PABPC1; this can be overcome by a CRISPR or siRNA mediated knock-out of endogenous PABPC1. At the same time, the deadenylation activity can be studied by stable expression of CNOT7/CAF1 and co-transfection of PABPC1 mutations along with BTG2. A study by (Maryati et al., 2014) revealed

reduced CNOT7/CAF1 deadenylation activity when exposed to compounds with different compositions. Since BTG2 enhances and regulates deadenylation by bridging CNOT7/CAF1 and PABPC1, it would be interesting to test the inhibition properties of the compounds by treating them with the quaternary structure (RNA-PABPC1-BTG2-CNOT7/CAF1).

To this end, our research has found a crucial residue in PABPC1, which is responsible for mediating interaction with BTG2. This discovery widens the horizons to explore BTG1 and BTG2 mutations associated with non-Hodgkin lymphomas (NHLs) which are the cancers of B, T or killer lymphocytes. The two most common NHLs are DLBCL and FL. DLBCL are heterogeneous and aggressive, affecting 30,000 new patients in the United States every year (Abramson and Shipp, 2005). The DLBCL comprises two distinct subtypes: germinal centre B-cell (GCB) and activated B-cell (ABC). At the same time, FL is indolent, incurable, and characterised by clinical and genetic heterogeneity.

Single nucleotide variants of both BTG1 and BTG2 were identified in NHLs and its subtypes; however, their functional role has not been suspected (Lohr et al., 2012). A recent study by Almasmoum et al. (2021) has disclosed the binding properties of CNOT7/CAF1 with NHLs associated with BTG1 mutations in the Box A and Box B region. As published in the studies earlier, the binding sites for CNOT7/CAF1 are in the Box A and Box B region of BTG1/BTG2 (Horiuchi et al., 2009b), and the binding site for PABPC1 is in the Box C region (Stupfler et al., 2016). It is crucial to investigate if the NHLs mutations located in Box A and Box B region of BTG1/BTG2 affects the binding sites for CNOT7/CAF1 or residual binding of BTG1/BTG2 mutations with CNOT7/CAF1 via PABPC1 are sufficient to regulate deadenylation. These questions are yet to be explored and answered in regards to studying the activity of the molecules with NHLs.

References.

ABRAMSON, J. S. & SHIPP, M. A. 2005. Advances in the biology and therapy of diffuse large B-cell lymphoma: moving toward a molecularly targeted approach. *Blood,* 106**,** 1164-74.

ACHARYA, K. R. & LLOYD, M. D. 2005. The advantages and limitations of protein crystal structures. *Trends Pharmacol Sci,* 26**,** 10-4.

AGO, H., ODA, M., TAKAHASHI, M., TSUGE, H., OCHI, S., KATUNUMA, N., MIYANO, M. & SAKURAI, J. 2006. Structural basis of the sphingomyelin phosphodiesterase activity in neutral sphingomyelinase from Bacillus cereus. *J Biol Chem,* 281**,** 16157-67.

ALEXEY NOVORADOVSKY, VIVIAN ZHANG, MADHUSHREE GHOSH, HOLLY HOGREFE, JOSEPH A. SORGE & GAASTERLAND., T. 2005. Computational Principles of Primer Design for Site Directed Mutagenesis. *Technical Proceedings of 2005 NSTI Nanotechnology Conference and Trade Show, Anaheim, 2005,***,** 532-535.

ALMASMOUM, H. A., AIRHIHEN, B., SEEDHOUSE, C. & WINKLER, G. S. 2021. Frequent loss of BTG1 activity and impaired interactions with the Caf1 subunit of the Ccr4-Not deadenylase in non-Hodgkin lymphoma. *Leuk Lymphoma,* 62**,** 281-290.

AMINE, H., RIPIN, N., SHARMA, S., STOECKLIN, G., ALLAIN, F. H., SERAPHIN, B. & MAUXION, F. 2021. A conserved motif in human BTG1 and BTG2 proteins mediates interaction with the poly(A) binding protein PABPC1 to stimulate mRNA deadenylation. *RNA Biol,* 18**,** 2450-2465.

ANDERSEN, K. R., JONSTRUP, A. T., VAN, L. B. & 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**,** 850-61.

ARRIZABALAGA, G. & LEHMANN, R. 1999. A selective screen reveals discrete functional domains in Drosophila Nanos. *Genetics,* 153**,** 1825-38.

BADIS, G., SAVEANU, C., FROMONT-RACINE, M. & JACQUIER, A. 2004. Targeted mRNA degradation by deadenylation-independent decapping. *Mol Cell,* 15**,** 5-15.

BARTLAM, M. & YAMAMOTO, T. 2010. The structural basis for deadenylation by the CCR4-NOT complex. *Protein & Cell,* 1**,** 443-452.

BAWANKAR, P., LOH, B., WOHLBOLD, L., SCHMIDT, S. & IZAURRALDE, E. 2013. NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biol,* 10**,** 228-44.

BHASKAR, V., ROUDKO, V., BASQUIN, J., SHARMA, K., URLAUB, H., SERAPHIN, B. & 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**,** 1281-8.

BIENROTH, S., KELLER, W. & WAHLE, E. 1993. Assembly of a processive messenger RNA polyadenylation complex. *EMBO J,* 12**,** 585-94.

BLACKSHEAR, P. J. 2002. Tristetraprolin and other CCCH tandem zinc-finger proteins in the regulation of mRNA turnover. *Biochem Soc Trans,* 30**,** 945-52.

BOLAND, A., CHEN, Y., RAISCH, T., JONAS, S., KUZUOGLU-OZTURK, D., WOHLBOLD, L., WEICHENRIEDER, O. & IZAURRALDE, E. 2013. Structure and assembly of the NOT module of the human CCR4-NOT complex. *Nat Struct Mol Biol,* 20**,** 1289-97.

BORNHORST, J. A. & FALKE, J. J. 2000. Purification of proteins using polyhistidine affinity tags. *Methods Enzymol,* 326**,** 245-54.

BOUVERET, E., RIGAUT, G., SHEVCHENKO, A., WILM, M. & SERAPHIN, B. 2000. A Sm-like protein complex that participates in mRNA degradation. *EMBO J,* 19**,** 1661-71.

BRADBURY, A., POSSENTI, R., SHOOTER, E. M. & TIRONE, F. 1991. Molecular cloning of PC3, a putatively secreted protein whose mRNA is induced by nerve growth factor and depolarization. *Proc Natl Acad Sci U S A,* 88**,** 3353-7.

BROWN, C. E. & SACHS, A. B. 1998. Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation. *Mol Cell Biol,* 18**,** 6548-59.

BUTLER, J. E. & KADONAGA, J. T. 2002. The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev,* 16**,** 2583-92.

CARBALLO, E., LAI, W. S. & BLACKSHEAR, P. J. 1998. Feedback inhibition of macrophage tumor necrosis factor-alpha production by tristetraprolin. *Science,* 281**,** 1001-5.

CARTHEW, R. W. & SONTHEIMER, E. J. 2009. Origins and Mechanisms of miRNAs and siRNAs. *Cell,* 136**,** 642-55.

References

CHEN, C. Y., ZHENG, D., XIA, Z. & SHYU, A. B. 2009. Ago-TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. *Nat Struct Mol Biol,* 16**,** 1160-6.

CHEN, H. & ZHOU, H. X. 2005. Prediction of interface residues in protein-protein complexes by a consensus neural network method: test against NMR data. *Proteins,* 61**,** 21-35.

CHEN, Y., KHAZINA, E., IZAURRALDE, E. & WEICHENRIEDER, O. 2021. Crystal structure and functional properties of the human CCR4-CAF1 deadenylase complex. *Nucleic Acids Res,* 49**,** 6489-6510.

CHU, J., CARGNELLO, M., TOPISIROVIC, I. & PELLETIER, J. 2016. Translation Initiation Factors: Reprogramming Protein Synthesis in Cancer. *Trends Cell Biol,* 26**,** 918-933.

CURTIS, D., TREIBER, D. K., TAO, F., ZAMORE, P. D., WILLIAMSON, J. R. & LEHMANN, R. 1997. A CCHC metal-binding domain in Nanos is essential for translational regulation. *EMBO J,* 16**,** 834-43.

DE FALCO, G., LEUCCI, E., LENZE, D., PICCALUGA, P. P., CLAUDIO, P. P., ONNIS, A., CERINO, G., NYAGOL, J., MWANDA, W., BELLAN, C., HUMMEL, M., PILERI, S., TOSI, P., STEIN, H., GIORDANO, A. & LEONCINI, L. 2007. Gene-expression analysis identifies novel RBL2/p130 target genes in endemic Burkitt lymphoma cell lines and primary tumors. *Blood,* 110**,** 1301-7.

DE KLERK, E. & T HOEN, P. A. 2015. Alternative mRNA transcription, processing, and translation: insights from RNA sequencing. *Trends Genet,* 31**,** 128-39.

DE MELO NETO, O. P., STANDART, N. & MARTINS DE SA, C. 1995. Autoregulation of poly(A)-binding protein synthesis in vitro. *Nucleic Acids Res,* 23**,** 2198-205.

DE VRIES, S. J. & BONVIN, A. M. 2011a. CPORT: a consensus interface predictor and its performance in prediction-driven docking with HADDOCK. *PLoS One,* 6**,** e17695.

DE VRIES, S. J. & BONVIN, A. M. J. J. 2011b. CPORT: A Consensus Interface Predictor and Its Performance in Prediction-Driven Docking with HADDOCK. *Plos One,* 6.

DE VRIES, S. J., VAN DIJK, A. D. & BONVIN, A. M. 2006. WHISCY: what information does surface conservation yield? Application to data-driven docking. *Proteins,* 63**,** 479-89.

DING, L. & HAN, M. 2007. GW182 family proteins are crucial for microRNA-mediated gene silencing. *Trends Cell Biol,* 17**,** 411-6.

DOIDGE, R., MITTAL, S., ASLAM, A. & 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.

DUPRESSOIR, A., MOREL, A. P., BARBOT, W., LOIREAU, M. P., CORBO, L. & HEIDMANN, T. 2001. Identification of four families of yCCR4- and Mg2+-dependent endonuclease-related proteins in higher eukaryotes, and characterization of orthologs of yCCR4 with a conserved leucine-rich repeat essential for hCAF1/hPOP2 binding. *BMC Genomics,* 2**,** 9.

ELISEEVA, I. A., LYABIN, D. N. & OVCHINNIKOV, L. P. 2013. Poly(A)-binding proteins: structure, domain organization, and activity regulation. *Biochemistry (Mosc),* 78**,** 1377-91.

EZZEDDINE, N., CHANG, T. C., ZHU, W., YAMASHITA, A., CHEN, C. Y., ZHONG, Z., YAMASHITA, Y., ZHENG, D. & SHYU, A. B. 2007. Human TOB, an antiproliferative transcription factor, is a poly(A)-binding protein-dependent positive regulator of cytoplasmic mRNA deadenylation. *Mol Cell Biol,* 27**,** 7791-801.

EZZEDDINE, N., CHEN, C. Y. & SHYU, A. B. 2012. Evidence providing new insights into TOB-promoted deadenylation and supporting a link between TOB's deadenylation-enhancing and antiproliferative activities. *Mol Cell Biol,* 32**,** 1089-98.

FABIAN, M. R., CIEPLAK, M. K., FRANK, F., MORITA, M., GREEN, J., SRIKUMAR, T., NAGAR, B., YAMAMOTO, T., RAUGHT, B., DUCHAINE, T. F. & SONENBERG, N. 2011. miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. *Nat Struct Mol Biol,* 18**,** 1211-7.

FABIAN, M. R., FRANK, F., ROUYA, C., SIDDIQUI, N., LAI, W. S., KARETNIKOV, A., BLACKSHEAR, P. J., NAGAR, B. & SONENBERG, N. 2013. Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. *Nat Struct Mol Biol,* 20**,** 735-9.

FERNANDEZ-RECIO, J., TOTROV, M. & ABAGYAN, R. 2004. Identification of protein-protein interaction sites from docking energy landscapes. *J Mol Biol,* 335**,** 843-65.

FIELDS, S. & SONG, O. 1989. A novel genetic system to detect protein-protein interactions. *Nature,* 340**,** 245-6.

References

FLETCHER, B. S., LIM, R. W., VARNUM, B. C., KUJUBU, D. A., KOSKI, R. A. & HERSCHMAN, H. R. 1991. Structure and expression of TIS21, a primary response gene induced by growth factors and tumor promoters. *J Biol Chem,* 266**,** 14511-8.

FRIEND, K., BROOK, M., BEZIRCI, F. B., SHEETS, M. D., GRAY, N. K. & SELI, E. 2012. Embryonic poly(A)-binding protein (ePAB) phosphorylation is required for Xenopus oocyte maturation. *Biochem J,* 445**,** 93-100.

FRISCHMEYER, P. A., VAN HOOF, A., O'DONNELL, K., GUERRERIO, A. L., PARKER, R. & DIETZ, H. C. 2002. An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science,* 295**,** 2258-61.

GARNEAU, N. L., WILUSZ, J. & WILUSZ, C. J. 2007. The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol,* 8**,** 113-26.

GAVRILOV, K. & SALTZMAN, W. M. 2012. Therapeutic siRNA: principles, challenges, and strategies. *Yale J Biol Med,* 85**,** 187-200.

GE, H., ZHOU, D., TONG, S., GAO, Y., TENG, M. & NIU, L. 2008. Crystal structure and possible dimerization of the single RRM of human PABPN1. *Proteins,* 71**,** 1539-45.

GOLEMIS, E. A., SEREBRIISKII, I., FINLEY, R. L., JR., KOLONIN, M. G., GYURIS, J. & BRENT, R. 2001a. Interaction trap/two-hybrid system to identify interacting proteins. *Curr Protoc Cell Biol,* Chapter 17**,** Unit 17 3.

GOLEMIS, E. A., SEREBRIISKII, I., FINLEY, R. L., JR., KOLONIN, M. G., GYURIS, J. & BRENT, R. 2001b. Interaction trap/two-hybrid system to identify interacting proteins. *Curr Protoc Mol Biol,* Chapter 20**,** Unit 20 1.

GRUNER, S., PETER, D., WEBER, R., WOHLBOLD, L., CHUNG, M. Y., WEICHENRIEDER, O., VALKOV, E., IGREJA, C. & IZAURRALDE, E. 2016. The Structures of eIF4E-eIF4G Complexes Reveal an Extended Interface to Regulate Translation Initiation. *Mol Cell,* 64**,** 467-479.

GYURIS, J., GOLEMIS, E., CHERTKOV, H. & BRENT, R. 1993. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell,* 75**,** 791-803.

HANZAWA, H., DE RUWE, M. J., ALBERT, T. K., VAN DER VLIET, P. C., TIMMERS, H. T. & BOELENS, R. 2001. The structure of the C4C4 ring finger of human NOT4 reveals features distinct from those of C3HC4 RING fingers. *J Biol Chem,* 276**,** 10185-90.

HINNEBUSCH, A. G. & LORSCH, J. R. 2012. The mechanism of eukaryotic translation initiation: new insights and challenges. *Cold Spring Harb Perspect Biol,* 4.

HORIUCHI, M., TAKEUCHI, K., NODA, N., MUROYA, N., SUZUKI, T., NAKAMURA, T., KAWAMURA-TSUZUKU, J., TAKAHASI, K., YAMAMOTO, T. & INAGAKI, F. 2009a. Structural basis for the antiproliferative activity of the Tob-hCaf1 complex. *J Biol Chem,* 284**,** 13244-13255.

HORIUCHI, M., TAKEUCHI, K., NODA, N., MUROYA, N., SUZUKI, T., NAKAMURA, T., KAWAMURA-TSUZUKU, J., TAKAHASI, K., YAMAMOTO, T. & INAGAKI, F. 2009b. Structural basis for the antiproliferative activity of the Tob-hCaf1 complex. *J Biol Chem,* 284**,** 13244-55.

HOSODA, N., LEJEUNE, F. & MAQUAT, L. E. 2006. Evidence that poly(A) binding protein C1 binds nuclear pre-mRNA poly(A) tails. *Mol Cell Biol,* 26**,** 3085-97.

HOUSELEY, J. & TOLLERVEY, D. 2009. The many pathways of RNA degradation. *Cell,* 136**,** 763-76.

HUNTER, Z. R., XU, L., YANG, G., ZHOU, Y., LIU, X., CAO, Y., MANNING, R. J., TRIPSAS, C., PATTERSON, C. J., SHEEHY, P. & TREON, S. P. 2014. The genomic landscape of Waldenstrom macroglobulinemia is characterized by highly recurring MYD88 and WHIM-like CXCR4 mutations, and small somatic deletions associated with B-cell lymphomagenesis. *Blood,* 123**,** 1637-46.

HWANG, S. S., LIM, J., YU, Z., KONG, P., SEFIK, E., XU, H., HARMAN, C. C. D., KIM, L. K., LEE, G. R., LI, H. B. & FLAVELL, R. A. 2020. mRNA destabilization by BTG1 and BTG2 maintains T cell quiescence. *Science,* 367**,** 1255-1260.

JAYNE, S., ZWARTJES, C. G., VAN SCHAIK, F. M. & TIMMERS, H. T. 2006. Involvement of the SMRT/NCoR-HDAC3 complex in transcriptional repression by the CNOT2 subunit of the human Ccr4-Not complex. *Biochem J,* 398**,** 461-7.

JIANG, Y., SOONG, T. D., WANG, L., MELNICK, A. M. & ELEMENTO, O. 2012. Genome-wide detection of genes targeted by non-Ig somatic hypermutation in lymphoma. *PLoS One,* 7**,** e40332.

References

JONES, A. X., CAO, Y., TANG, Y. L., WANG, J. H., DING, Y. H., TAN, H., CHEN, Z. L., FANG, R. Q., YIN, J., CHEN, R. C., ZHU, X., SHE, Y., HUANG, N., SHAO, F., YE, K., SUN, R. X., HE, S. M., LEI, X. & DONG, M. Q. 2019. Improving mass spectrometry analysis of protein structures with arginine-selective chemical cross-linkers. *Nat Commun,* 10**,** 3911.

JONSTRUP, A. T., ANDERSEN, K. R., VAN, L. B. & 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**,** 3153-64.

KAPP, L. D. & LORSCH, J. R. 2004. GTP-dependent recognition of the methionine moiety on initiator tRNA by translation factor eIF2. *J Mol Biol,* 335**,** 923-36.

KAUFMANN, I., MARTIN, G., FRIEDLEIN, A., LANGEN, H. & KELLER, W. 2004. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J,* 23**,** 616-26.

KAWAKUBO, H., BRACHTEL, E., HAYASHIDA, T., YEO, G., KISH, J., MUZIKANSKY, A., WALDEN, P. D. & MAHESWARAN, S. 2006. Loss of B-cell translocation gene-2 in estrogen receptor-positive breast carcinoma is associated with tumor grade and overexpression of cyclin d1 protein. *Cancer Res,* 66**,** 7075-82.

KERWITZ, Y., KUHN, U., LILIE, H., KNOTH, A., SCHEUERMANN, T., FRIEDRICH, H., SCHWARZ, E. & WAHLE, E. 2003. Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *EMBO J,* 22**,** 3705-14.

KOZLOV, G., MENADE, M., ROSENAUER, A., NGUYEN, L. & GEHRING, K. 2010. Molecular determinants of PAM2 recognition by the MLLE domain of poly(A)-binding protein. *J Mol Biol,* 397**,** 397-407.

KRIDEL, R., CHAN, F. C., MOTTOK, A., BOYLE, M., FARINHA, P., TAN, K., MEISSNER, B., BASHASHATI, A., MCPHERSON, A., ROTH, A., SHUMANSKY, K., YAP, D., BEN-NERIAH, S., ROSNER, J., SMITH, M. A., NIELSEN, C., GINE, E., TELENIUS, A., ENNISHI, D., MUNGALL, A., MOORE, R., MORIN, R. D., JOHNSON, N. A., SEHN, L. H., TOUSSEYN, T., DOGAN, A., CONNORS, J. M., SCOTT, D. W., STEIDL, C., MARRA, M. A., GASCOYNE, R. D. & SHAH, S. P. 2016. Histological Transformation and Progression in Follicular Lymphoma: A Clonal Evolution Study. *PLoS Med,* 13**,** e1002197.

KRISSINEL, E. & HENRICK, K. 2007. Inference of macromolecular assemblies from crystalline state. *J Mol Biol,* 372**,** 774-97.

KUFAREVA, I., BUDAGYAN, L., RAUSH, E., TOTROV, M. & ABAGYAN, R. 2007. PIER: protein interface recognition for structural proteomics. *Proteins,* 67**,** 400-17.

LAI, W. S., CARBALLO, E., THORN, J. M., KENNINGTON, E. A. & BLACKSHEAR, P. J. 2000. Interactions of CCCH zinc finger proteins with mRNA. Binding of tristetraprolin-related zinc finger proteins to Au-rich elements and destabilization of mRNA. *J Biol Chem,* 275**,** 17827-37.

LEFEVRE, F., REMY, M. H. & MASSON, J. M. 1997. Alanine-stretch scanning mutagenesis: a simple and efficient method to probe protein structure and function. *Nucleic Acids Res,* 25**,** 447-8.

LIANG, S., ZHANG, C., LIU, S. & ZHOU, Y. 2006. Protein binding site prediction using an empirical scoring function. *Nucleic Acids Res,* 34**,** 3698-707.

LIM, I. K., LEE, M. S., RYU, M. S., PARK, T. J., FUJIKI, H., EGUCHI, H. & PAIK, W. K. 1998. Induction of growth inhibition of 293 cells by downregulation of the cyclin E and cyclin-dependent kinase 4 proteins due to overexpression of TIS21. *Mol Carcinog,* 23**,** 25-35.

LOHR, J. G., STOJANOV, P., LAWRENCE, M. S., AUCLAIR, D., CHAPUY, B., SOUGNEZ, C., CRUZ-GORDILLO, P., KNOECHEL, B., ASMANN, Y. W., SLAGER, S. L., NOVAK, A. J., DOGAN, A., ANSELL, S. M., LINK, B. K., ZOU, L., GOULD, J., SAKSENA, G., STRANSKY, N., RANGEL-ESCARENO, C., FERNANDEZ-LOPEZ, J. C., HIDALGO-MIRANDA, A., MELENDEZ-ZAJGLA, J., HERNANDEZ-LEMUS, E., SCHWARZ-CRUZ Y CELIS, A., IMAZ-ROSSHANDLER, I., OJESINA, A. I., JUNG, J., PEDAMALLU, C. S., LANDER, E. S., HABERMANN, T. M., CERHAN, J. R., SHIPP, M. A., GETZ, G. & GOLUB, T. R. 2012. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci U S A,* 109**,** 3879-84.

MALVAR, T., BIRON, R. W., KABACK, D. B. & DENIS, C. L. 1992. The CCR4 protein from Saccharomyces cerevisiae contains a leucine-rich repeat region which is required for its control of ADH2 gene expression. *Genetics,* 132**,** 951-62.

MARINTCHEV, A. & WAGNER, G. 2004. Translation initiation: structures, mechanisms and evolution. *Q Rev Biophys,* 37**,** 197-284.

MARTIN, D., ABBA, M. C., MOLINOLO, A. A., VITALE-CROSS, L., WANG, Z., ZAIDA, M., DELIC, N. C., SAMUELS, Y., LYONS, J. G. & GUTKIND, J. S. 2014. The head and neck cancer cell oncogenome: a platform for the development of precision molecular therapies. *Oncotarget,* 5**,** 8906-23.

MARYATI, M., KAUR, I., JADHAV, G. P., OLOTU-UMOREN, L., OVEH, B., HASHMI, L., FISCHER, P. M. & WINKLER, G. S. 2014. A fluorescence-based assay suitable for quantitative analysis of deadenylase enzyme activity. *Nucleic Acids Res,* 42**,** e30.

MATHYS, H., BASQUIN, J., OZGUR, S., CZARNOCKI-CIECIURA, M., BONNEAU, F., AARTSE, A., DZIEMBOWSKI, A., NOWOTNY, M., CONTI, E. & FILIPOWICZ, W. 2014. Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in microRNA repression. *Mol Cell,* 54**,** 751-65.

MATSUDA, S., ROUAULT, J., MAGAUD, J. & BERTHET, C. 2001. In search of a function for the TIS21/PC3/BTG1/TOB family. *FEBS Lett,* 497**,** 67-72.

MAUXION, F., PREVE, B. & SERAPHIN, B. 2013. C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4-NOT complex. *RNA Biol,* 10**,** 267-76.

MEIJER, H. A., SCHMIDT, T., GILLEN, S. L., LANGLAIS, C., JUKES-JONES, R., DE MOOR, C. H., CAIN, K., WILCZYNSKA, A. & BUSHELL, M. 2019. DEAD-box helicase eIF4A2 inhibits CNOT7 deadenylation activity. *Nucleic Acids Res,* 47**,** 8224-8238.

MOL, C. D., IZUMI, T., MITRA, S. & TAINER, J. A. 2000. DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected]. *Nature,* 403**,** 451-6.

MOL, C. D., KUO, C. F., THAYER, M. M., CUNNINGHAM, R. P. & TAINER, J. A. 1995. Structure and function of the multifunctional DNA-repair enzyme exonuclease III. *Nature,* 374**,** 381-6.

MONTAGNOLI, A., GUARDAVACCARO, D., STARACE, G. & TIRONE, F. 1996. Overexpression of the nerve growth factor-inducible PC3 immediate early gene is associated with growth inhibition. *Cell Growth Differ,* 7**,** 1327-36.

MOREIRA, I. S., FERNANDES, P. A. & RAMOS, M. J. 2007. Hot spots--a review of the protein-protein interface determinant amino-acid residues. *Proteins,* 68**,** 803-12.

MORRISSEY, J. P., DEARDORFF, J. A., HEBRON, C. & SACHS, A. B. 1999. Decapping of stabilized, polyadenylated mRNA in yeast pab1 mutants. *Yeast,* 15**,** 687-702.

MUHLRAD, D. & PARKER, R. 1994. Premature translational termination triggers mRNA decapping. *Nature,* 370**,** 578-81.

NAGARAJAN, V. K., JONES, C. I., NEWBURY, S. F. & GREEN, P. J. 2013. XRN 5'-->3' exoribonucleases: structure, mechanisms and functions. *Biochim Biophys Acta,* 1829**,** 590-603.

NAHTA, R., YUAN, L. X., FITERMAN, D. J., ZHANG, L., SYMMANS, W. F., UENO, N. T. & ESTEVA, F. J. 2006. B cell translocation gene 1 contributes to antisense Bcl-2-mediated apoptosis in breast cancer cells. *Mol Cancer Ther,* 5**,** 1593-601.

NEUVIRTH, H., RAZ, R. & SCHREIBER, G. 2004. ProMate: a structure based prediction program to identify the location of protein-protein binding sites. *J Mol Biol,* 338**,** 181-99.

OKOCHI, K., SUZUKI, T., INOUE, J., MATSUDA, S. & YAMAMOTO, T. 2005. Interaction of anti-proliferative protein Tob with poly(A)-binding protein and inducible poly(A)-binding protein: implication of Tob in translational control. *Genes Cells,* 10**,** 151-63.

OU, Y. H., CHUNG, P. H., HSU, F. F., SUN, T. P., CHANG, W. Y. & SHIEH, S. Y. 2007. The candidate tumor suppressor BTG3 is a transcriptional target of p53 that inhibits E2F1. *EMBO J,* 26**,** 3968-80.

PARK, S., LEE, Y. J., LEE, H. J., SEKI, T., HONG, K. H., PARK, J., BEPPU, H., LIM, I. K., YOON, J. W., LI, E., KIM, S. J. & OH, S. P. 2004. B-cell translocation gene 2 (Btg2) regulates vertebral patterning by modulating bone morphogenetic protein/smad signaling. *Mol Cell Biol,* 24**,** 10256-62.

PARKER, R. & SONG, H. 2004. The enzymes and control of eukaryotic mRNA turnover. *Nat Struct Mol Biol,* 11**,** 121-7.

References

PAVANELLO, L., HALL, B., AIRHIHEN, B. & WINKLER, G. S. 2018. The central region of CNOT1 and CNOT9 stimulates deadenylation by the Ccr4-Not nuclease module. *Biochem J,* 475**,** 3437-3450.

PETIT, A. P., WOHLBOLD, L., BAWANKAR, P., HUNTZINGER, E., SCHMIDT, S., IZAURRALDE, E. & WEICHENRIEDER, O. 2012. The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. *Nucleic Acids Res,* 40**,** 11058-72.

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem,* 25**,** 1605-12.

POROLLO, A. & MELLER, J. 2007. Prediction-based fingerprints of protein-protein interactions. *Proteins,* 66**,** 630-45.

PREKER, P. J., OHNACKER, M., MINVIELLE-SEBASTIA, L. & KELLER, W. 1997. A multisubunit 3' end processing factor from yeast containing poly(A) polymerase and homologues of the subunits of mammalian cleavage and polyadenylation specificity factor. *EMBO J,* 16**,** 4727-37.

PREVOT, D., MOREL, A. P., VOELTZEL, T., ROSTAN, M. C., RIMOKH, R., MAGAUD, J. P. & CORBO, L. 2001. Relationships of the antiproliferative proteins BTG1 and BTG2 with CAF1, the human homolog of a component of the yeast CCR4 transcriptional complex: involvement in estrogen receptor alpha signaling pathway. *J Biol Chem,* 276**,** 9640-8.

PREVOT, D., VOELTZEL, T., BIROT, A. M., MOREL, A. P., ROSTAN, M. C., MAGAUD, J. P. & CORBO, L. 2000. The leukemia-associated protein Btg1 and the p53-regulated protein Btg2 interact with the homeoprotein Hoxb9 and enhance its transcriptional activation. *Journal of Biological Chemistry,* 275**,** 147-153.

REDDY, A., ZHANG, J., DAVIS, N. S., MOFFITT, A. B., LOVE, C. L., WALDROP, A., LEPPA, S., PASANEN, A., MERIRANTA, L., KARJALAINEN-LINDSBERG, M. L., NORGAARD, P., PEDERSEN, M., GANG, A. O., HOGDALL, E., HEAVICAN, T. B., LONE, W., IQBAL, J., QIN, Q., LI, G., KIM, S. Y., HEALY, J., RICHARDS, K. L., FEDORIW, Y., BERNAL-MIZRACHI, L., KOFF, J. L., STATON, A. D., FLOWERS, C. R., PALTIEL, O., GOLDSCHMIDT, N., CALAMINICI, M., CLEAR, A., GRIBBEN, J., NGUYEN, E., CZADER, M. B., ONDREJKA, S. L., COLLIE, A., HSI, E. D., TSE, E., AU-YEUNG, R. K. H., KWONG, Y. L., SRIVASTAVA, G., CHOI, W. W. L., EVENS, A. M., PILICHOWSKA, M., SENGAR, M., REDDY, N., LI, S., CHADBURN, A., GORDON, L. I., JAFFE, E. S., LEVY, S., REMPEL, R., TZENG, T., HAPP, L. E., DAVE, T., RAJAGOPALAN, D., DATTA, J., DUNSON, D. B. & DAVE, S. S. 2017. Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. *Cell,* 171**,** 481-494 e15.

RICHARD, P. & MANLEY, J. L. 2009. Transcription termination by nuclear RNA polymerases. *Genes Dev,* 23**,** 1247-69.

ROBERT, X. & GOUET, P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res,* 42**,** W320-4.

ROUAULT, J. P., FALETTE, N., GUEHENNEUX, F., GUILLOT, C., RIMOKH, R., WANG, Q., BERTHET, C., MOYRET-LALLE, C., SAVATIER, P., PAIN, B., SHAW, P., BERGER, R., SAMARUT, J., MAGAUD, J. P., OZTURK, M., SAMARUT, C. & PUISIEUX, A. 1996. Identification of BTG2, an antiproliferative p53-dependent component of the DNA damage cellular response pathway. *Nat Genet,* 14**,** 482-6.

ROUAULT, J. P., RIMOKH, R., TESSA, C., PARANHOS, G., FFRENCH, M., DURET, L., GAROCCIO, M., GERMAIN, D., SAMARUT, J. & MAGAUD, J. P. 1992. BTG1, a member of a new family of antiproliferative genes. *EMBO J,* 11**,** 1663-70.

ROUNBEHLER, R. J., FALLAHI, M., YANG, C., STEEVES, M. A., LI, W., DOHERTY, J. R., SCHAUB, F. X., SANDUJA, S., DIXON, D. A., BLACKSHEAR, P. J. & CLEVELAND, J. L. 2012. Tristetraprolin impairs myc-induced lymphoma and abolishes the malignant state. *Cell,* 150**,** 563-74.

SACHS, A. B. & DEARDORFF, J. A. 1992. Translation initiation requires the PAB-dependent poly(A) ribonuclease in yeast. *Cell,* 70**,** 961-73.

SAFAEE, N., KOZLOV, G., NORONHA, A. M., XIE, J., WILDS, C. J. & GEHRING, K. 2012. Interdomain allostery promotes assembly of the poly(A) mRNA complex with PABP and eIF4G. *Mol Cell,* 48**,** 375-86.

SCHAFER, I. B., RODE, M., BONNEAU, F., SCHUSSLER, S. & CONTI, E. 2014. The structure of the Pan2-Pan3 core complex reveals cross-talk between deadenylase and pseudokinase. *Nat Struct Mol Biol,* 21**,** 591-8.

SCHLUNDT, A., HEINZ, G. A., JANOWSKI, R., GEERLOF, A., STEHLE, R., HEISSMEYER, V., NIESSING, D. & SATTLER, M. 2014. Structural basis for RNA recognition in roquin-mediated post-transcriptional gene regulation. *Nat Struct Mol Biol,* 21**,** 671-8.

SCHMID, M. & JENSEN, T. H. 2008. The exosome: a multipurpose RNA-decay machine. *Trends Biochem Sci,* 33**,** 501-10.

SCHOENBERG, D. R. & MAQUAT, L. E. 2012. Regulation of cytoplasmic mRNA decay. *Nat Rev Genet,* 13**,** 246-59.

SGROMO, A., RAISCH, T., BAWANKAR, P., BHANDARI, D., CHEN, Y., KUZUOGLU-OZTURK, D., WEICHENRIEDER, O. & IZAURRALDE, E. 2017. A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by Drosophila Roquin. *Nat Commun,* 8**,** 14307.

SIDDIQUI, N., MANGUS, D. A., CHANG, T. C., PALERMINO, J. M., SHYU, A. B. & GEHRING, K. 2007. Poly(A) nuclease interacts with the C-terminal domain of polyadenylate-binding protein domain from poly(A)-binding protein. *J Biol Chem,* 282**,** 25067-75.

SIWASZEK, A., UKLEJA, M. & DZIEMBOWSKI, A. 2014. Proteins involved in the degradation of cytoplasmic mRNA in the major eukaryotic model systems. *RNA Biol,* 11**,** 1122-36.

SONG, J., MCGIVERN, J. V., NICHOLS, K. W., MARKLEY, J. L. & SHEETS, M. D. 2008. Structural basis for RNA recognition by a type II poly(A)-binding protein. *Proc Natl Acad Sci U S A,* 105**,** 15317-22.

SONG, J. J., SMITH, S. K., HANNON, G. J. & JOSHUA-TOR, L. 2004. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science,* 305**,** 1434-7.

STUPFLER, B., BIRCK, C., SERAPHIN, B. & MAUXION, F. 2016. BTG2 bridges PABPC1 RNA-binding domains and CAF1 deadenylase to control cell proliferation. *Nat Commun,* 7**,** 10811.

SUZUKI, T., J, K. T., AJIMA, R., NAKAMURA, T., YOSHIDA, Y. & YAMAMOTO, T. 2002. Phosphorylation of three regulatory serines of Tob by Erk1 and Erk2 is required for Ras-mediated cell proliferation and transformation. *Genes Dev,* 16**,** 1356-70.

SVEJSTRUP, J. Q., VICHI, P. & EGLY, J. M. 1996. The multiple roles of transcription/repair factor TFIIH. *Trends Biochem Sci,* 21**,** 346-50.

SVITKIN, Y. V. & SONENBERG, N. 2006. Translational control by the poly(A) binding protein: A check for mRNA integrity. *Molecular Biology,* 40**,** 684-693.

TAKAHASHI, F., CHIBA, N., TAJIMA, K., HAYASHIDA, T., SHIMADA, T., TAKAHASHI, M., MORIYAMA, H., BRACHTEL, E., EDELMAN, E. J., RAMASWAMY, S. & MAHESWARAN, S. 2011. Breast tumor progression induced by loss of BTG2 expression is inhibited by targeted therapy with the ErbB/HER inhibitor lapatinib. *Oncogene,* 30**,** 3084-95.

TATE, J. G., BAMFORD, S., JUBB, H. C., SONDKA, Z., BEARE, D. M., BINDAL, N., BOUTSELAKIS, H., COLE, C. G., CREATORE, C., DAWSON, E., FISH, P., HARSHA, B., HATHAWAY, C., JUPE, S. C., KOK, C. Y., NOBLE, K., PONTING, L., RAMSHAW, C. C., RYE, C. E., SPEEDY, H. E., STEFANCSIK, R., THOMPSON, S. L., WANG, S., WARD, S., CAMPBELL, P. J. & FORBES, S. A. 2019. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res,* 47**,** D941-D947.

THORE, S., MAUXION, F., SERAPHIN, B. & SUCK, D. 2003. X-ray structure and activity of the yeast Pop2 protein: a nuclease subunit of the mRNA deadenylase complex. *EMBO Rep,* 4**,** 1150-5.

TIAN, B., HU, J., ZHANG, H. & LUTZ, C. S. 2005. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res,* 33**,** 201-12.

UCHIDA, N., HOSHINO, S. & KATADA, T. 2004. Identification of a human cytoplasmic poly(A) nuclease complex stimulated by poly(A)-binding protein. *J Biol Chem,* 279**,** 1383-91.

VAN HOOF, A., FRISCHMEYER, P. A., DIETZ, H. C. & PARKER, R. 2002. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science,* 295**,** 2262-4.

VAN ZUNDERT, G. C. P., RODRIGUES, J., TRELLET, M., SCHMITZ, C., KASTRITIS, P. L., KARACA, E., MELQUIOND, A. S. J., VAN DIJK, M., DE VRIES, S. J. & BONVIN, A. 2016. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol,* 428**,** 720-725.

VIPHAKONE, N., VOISINET-HAKIL, F. & MINVIELLE-SEBASTIA, L. 2008. Molecular dissection of mRNA poly(A) tail length control in yeast. *Nucleic Acids Res,* 36**,** 2418-33.

References

WAHLE, E. 1995. Poly(A) tail length control is caused by termination of processive synthesis. *J Biol Chem,* 270**,** 2800-8.

WAHLE, E. & WINKLER, G. S. 2013. RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim Biophys Acta,* 1829**,** 561-70.

WANG, H., MORITA, M., YANG, X., SUZUKI, T., YANG, W., WANG, J., ITO, K., WANG, Q., ZHAO, C., BARTLAM, M., YAMAMOTO, T. & RAO, Z. 2010. Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. *EMBO J,* 29**,** 2566-76.

WEBSTER, M. W., CHEN, Y. H., STOWELL, J. A. W., ALHUSAINI, N., SWEET, T., GRAVELEY, B. R., COLLER, J. & PASSMORE, L. A. 2018. mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. *Mol Cell,* 70**,** 1089-1100 e8.

WEICHENRIEDER, O., REPANAS, K. & PERRAKIS, A. 2004. Crystal structure of the targeting endonuclease of the human LINE-1 retrotransposon. *Structure,* 12**,** 975-86.

WHEATON, K., MUIR, J., MA, W. & BENCHIMOL, S. 2010. BTG2 antagonizes Pin1 in response to mitogens and telomere disruption during replicative senescence. *Aging Cell,* 9**,** 747-60.

WINKLER, G. S. 2010. The mammalian anti-proliferative BTG/Tob protein family. *J Cell Physiol,* 222**,** 66-72.

WINKLER, G. S. & BALACCO, D. L. 2013. Heterogeneity and complexity within the nuclease module of the Ccr4-Not complex. *Front Genet,* 4**,** 296.

WINKLER, G. S., MULDER, K. W., BARDWELL, V. J., KALKHOVEN, E. & TIMMERS, H. T. 2006. Human Ccr4-Not complex is a ligand-dependent repressor of nuclear receptor-mediated transcription. *EMBO J,* 25**,** 3089-99.

WOLF, J. & PASSMORE, L. A. 2014. mRNA deadenylation by Pan2-Pan3. *Biochem Soc Trans,* 42**,** 184-7.

WOLF, J., VALKOV, E., ALLEN, M. D., MEINEKE, B., GORDIYENKO, Y., MCLAUGHLIN, S. H., OLSEN, T. M., ROBINSON, C. V., BYCROFT, M., STEWART, M. & PASSMORE, L. A. 2014. Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation. *EMBO J,* 33**,** 1514-26.

WURM, J. P. & SPRANGERS, R. 2019. Dcp2: an mRNA decapping enzyme that adopts many different shapes and forms. *Curr Opin Struct Biol,* 59**,** 115-123.

XU, K., BAI, Y., ZHANG, A., ZHANG, Q. & BARTLAM, M. G. 2014. Insights into the structure and architecture of the CCR4-NOT complex. *Front Genet,* 5**,** 137.

YAMASHITA, A., CHANG, T. C., YAMASHITA, Y., ZHU, W., ZHONG, Z., CHEN, C. Y. & SHYU, A. B. 2005. Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat Struct Mol Biol,* 12**,** 1054-63.

YANG, F. & SCHOENBERG, D. R. 2004. Endonuclease-mediated mRNA decay involves the selective targeting of PMR1 to polyribosome-bound substrate mRNA. *Mol Cell,* 14**,** 435-45.

YANG, X., MORITA, M., WANG, H., SUZUKI, T., YANG, W., LUO, Y., ZHAO, C., YU, Y., BARTLAM, M., YAMAMOTO, T. & RAO, Z. 2008. Crystal structures of human BTG2 and mouse TIS21 involved in suppression of CAF1 deadenylase activity. *Nucleic Acids Res,* 36**,** 6872-81.

YI, H., PARK, J., HA, M., LIM, J., CHANG, H. & KIM, V. N. 2018. PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. *Mol Cell,* 70**,** 1081-1088 e5.

YOSHIDA, Y., TANAKA, S., UMEMORI, H., MINOWA, O., USUI, M., IKEMATSU, N., HOSODA, E., IMAMURA, T., KUNO, J., YAMASHITA, T., MIYAZONO, K., NODA, M., NODA, T. & YAMAMOTO, T. 2000. Negative regulation of BMP/Smad signaling by Tob in osteoblasts. *Cell,* 103**,** 1085-97.

ZEKRI, L., KUZUOGLU-OZTURK, D. & IZAURRALDE, E. 2013. GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. *EMBO J,* 32**,** 1052-65.

ZHANG, Z., CHEN, C., WANG, G., YANG, Z., SAN, J., ZHENG, J., LI, Q., LUO, X., HU, Q., LI, Z. & WANG, D. 2011. Aberrant expression of the p53-inducible antiproliferative gene BTG2 in hepatocellular carcinoma is associated with overexpression of the cell cycle-related proteins. *Cell Biochem Biophys,* 61**,** 83-91.