Review

Molecular Insights into mRNA Polyadenylation and Deadenylation

Junjie Liu †, Xubing Lu †, Siyu Zhang, Ling Yuan and Yadong Sun *

School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China
* Correspondence: sunyd1@shanghaitech.edu.cn
† These authors contributed equally to this work.

Abstract: Poly(A) tails are present on almost all eukaryotic mRNAs, and play critical roles in mRNA stability, nuclear export, and translation efficiency. The biosynthesis and shortening of a poly(A) tail are regulated by large multiprotein complexes. However, the molecular mechanisms of these protein machineries still remain unclear. Recent studies regarding the structural and biochemical characteristics of those protein complexes have shed light on the potential mechanisms of polyadenylation and deadenylation. This review summarizes the recent structural studies on pre-mRNA 3′-end processing complexes that initiate the polyadenylation and discusses the similarities and differences between yeast and human machineries. Specifically, we highlight recent biochemical efforts in the reconstitution of the active human canonical pre-mRNA 3′-end processing systems, as well as the roles of RBBP6/Mpe1 in activating the entire machinery. We also describe how poly(A) tails are removed by the PAN2-PAN3 and CCR4-NOT deadenylation complexes and discuss the emerging role of the cytoplasmic poly(A)-binding protein (PABPC) in promoting deadenylation. Together, these recent discoveries show that the dynamic features of these machineries play important roles in regulating polyadenylation and deadenylation.

Keywords: poly(A) tail; polyadenylation; deadenylation; 3′-end processing; RBBP6/Mpe1; PAN2-PAN3; CCR4-NOT; PABPC; RNA-binding protein

1. Introduction

Poly(A) tails, as a major modification of mRNA, are present on almost all eukaryotic mRNAs. The accurate formation of 3′ end on newly synthesized mRNA, known as polyadenylation, is critical for mRNA maturation, mRNA stability, and mRNA export from the nucleus into the cytoplasm. Poly(A) tails are bound by the poly(A)-binding protein (PAB or PABP). The nuclear isoform PABPN1 is essential for the efficient polymerization of poly(A) tails [1–3] and is substantially correlated with poly(A) tail length [4]. In the cytoplasm, PABPN1 is replaced by cytosolic poly(A)-binding protein (PABPC), which protects mRNA from non-specific degradation and promotes mRNA translation. Deadenylation, a process that can release PABPC and initiate mRNA decay, is controlled by a variety of deadenylases to dynamically modulate poly(A) tail length. Poly(A) tails of various lengths were discovered on mRNAs in the early 1970s [5,6]. The length of the poly(A) tail has been found to be closely related to almost every stage in the mRNA life cycle, including transcription termination, mRNA subcellular localization, mRNA quality control, translation efficiency, and mRNA decay [7–9]. Poly(A) tails in mammals can be up to 200–250 nt [6] and 90 nt long in yeast [10]. Mutations and defects in these two key processes, polyadenylation and deadenylation, will result in human diseases and viral infections [11–13].

The nascent pre-mRNA is polyadenylated in the following two steps: cleavage and polyadenylation (Figure 1). The molecular insight into poly(A) tail biosynthesis has been revealed by the earlier crystal structure of yeast Pap1 in complex with poly(A) RNA [14]. In charge of mRNA cleavage is the conserved endonuclease CPSF73 [15], which has recently been identified as a potential therapeutic target for several cancers, inflammatory diseases, and protozoan infections [16–19]. Deadenylation, on the other hand, contributes to the

initiation of mRNA degradation and serves as a checkpoint for gene expression [20]. Both polyadenylation and deadenylation are carried out by large protein machineries, consisting of a great number of conserved and highly regulated multiprotein sub-complexes. How these sub-complexes interact with each other in a cooperative manner and, consequently, safeguard the accuracy of these two important processes remains elusive. With the aid of recent breakthroughs in sequencing methods and cryo-electron microscopy (cryo-EM) techniques, our knowledge of molecular insights into polyadenylation and deadenylation has been expanded.
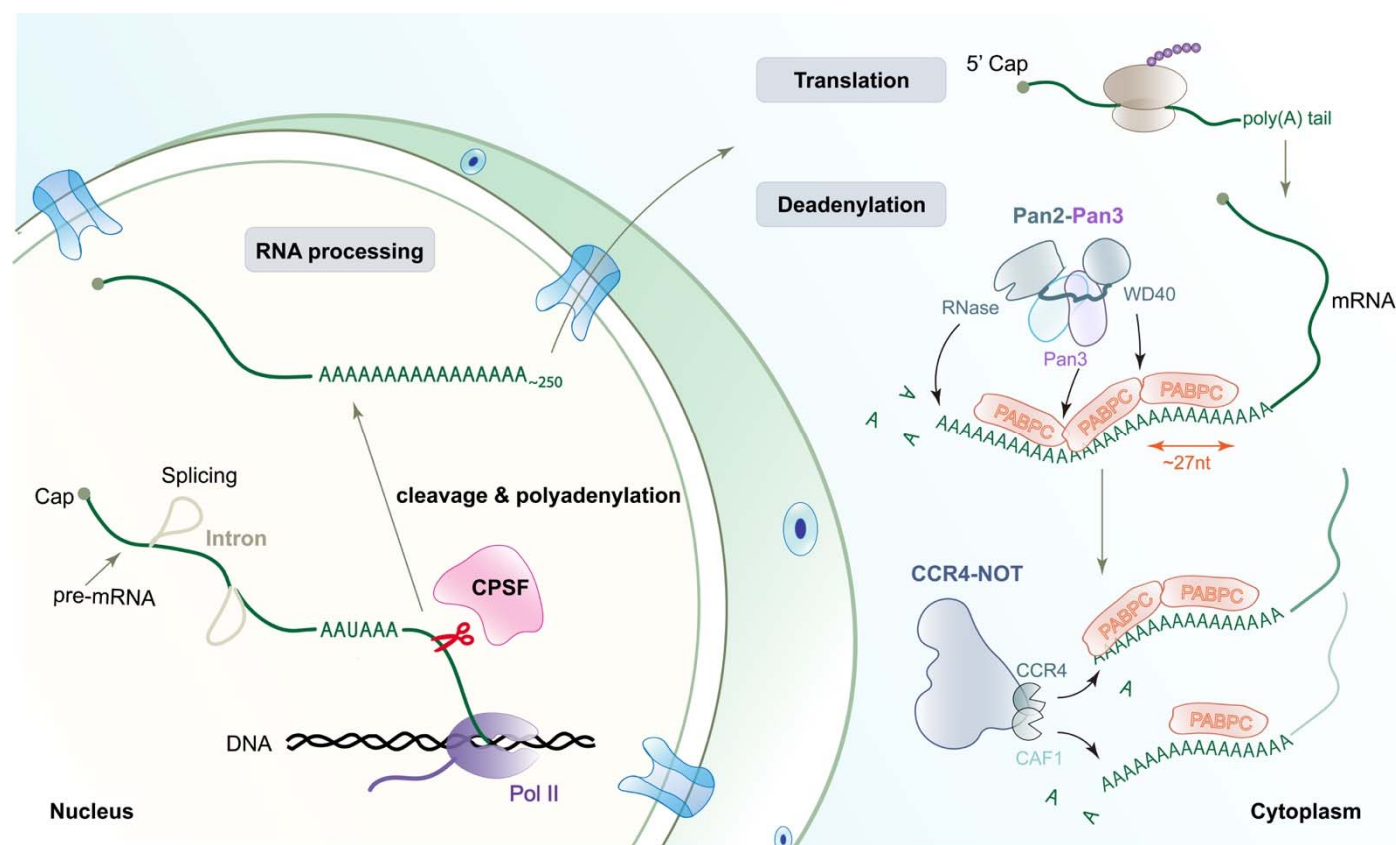

Figure 1. Schematic representation of polyadenylation and deadenylation. In the eukaryotic nucleus, primary RNA transcripts undergo extensive modifications to generate mature functional mRNA; this process is named RNA processing, which includes 5′ capping, splicing, and 3′ polyadenylation. It must be noted that 3′-end processing includes two steps, cleavage at a specific site (the scissors, indicating the site of cleavage by CPSF) and addition of a poly(A) tail at the same site. Then, the mature mRNAs with long poly(A) tails will be exported to the cytoplasm and translated into protein. PABPC covers the poly(A) tail, protecting mRNA from non-specific degradation and meanwhile promoting poly(A) tail removal or shortening via interactions with PAN2-PAN3 and CCR4-NOT complexes (deadenylation). Pan3 and the WD40 domain in Pan2 recognize the interface between different PABP molecules. Deadenylase CCR4 works on PABPC-bound A tails, whereas CAF1 only trims naked poly(A) RNA. CPSF, cleavage and polyadenylation specificity factor.

In this review, we outline current structural investigations into the human and yeast pre-mRNA 3′-end processing complexes, focusing on their recognition of the polyadenylation signal (PAS), machinery organization, and cleavage activation. In particular, we highlight the reconstitution of the active human canonical pre-mRNA 3′-end processing systems, which clarifies, for the first time, the minimal composition of human machineries, and address the functions of RBBP6/Mpe1 in activating the entire machinery. We also describe recent molecular insights into the PAN2-PAN3 and CCR4-NOT deadenylation

complexes and discuss the role of cytoplasmic PABPC in promoting deadenylation. Finally, we offer an outlook for future research.

2. The Cleavage Step in Polyadenylation

Pre-mRNA 3′-end processing includes endonuclease-mediated cleavage of the pre-mRNA at a specific site and poly (A) polymerase-mediated insertion of a poly(A) tail. The only known exception is replication-dependent histone pre-mRNAs in metazoan, which only involves the cleavage step [21,22]. The cleavage step in polyadenylation is carried out by a series of conserved multiprotein complexes, which were first accomplished in cell extracts in the 1980s [23,24]. During nearly four decades of research, several groups have made tremendous efforts to reconstitute this biological process with recombinant proteins in vitro, which not only simplifies the biochemical reaction for studying the mechanism of the whole machinery and its regulatory factors, but also provides an easily accessible system for identifying new small-molecule inhibitors against CPSF73 [25,26]. Recent studies have defined a minimal list of essential components for this process through the successful reconstitution of pre-mRNA 3′-end processing machinery in vitro [27–30]. By combining the new biochemical and structural studies, the composition and molecular mechanism for pre-mRNA 3′-end cleavage are highly conserved between yeast and humans.

2.1. Essential Components in Canonical Pre-mRNA 3′-End Processing Machinery

The essential components of human canonical machinery include CPSF (cleavage and polyadenylation specificity factor), CstF (cleavage stimulation factor), CF IIm (cleavage factor IIm), PAP(poly (A) polymerase) and Rbbp6 (retinoblastoma-binding protein 6) [27,29]. Surprisingly, with the exception of symplekin, the yeast homologs of all the essential factors involved in human canonical machinery are necessary for yeast pre-mRNA 3′-end processing machinery as well [31], illustrating that 3′-end processing is highly conserved in eukaryotes. The detailed composition of each factor in humans and yeast is shown in Table 1. Specifically, earlier studies indicated that human RBBP6 is involved in human pre-mRNA 3′-end processing, but omitted or underestimated in previous reconstitution and structural studies, due to the fact that RBBP6 is not connected with human CPSF in a stable manner [32,33]. In contrast, yeast Mpe1 is a stable subunit of CPF and binds to Ysh1 (human homolog CPSF73) [28]. The most recent structural and biochemical studies demonstrate the essential role of RBBP6/Mpe1 in regulating 3′-end cleavage, polyadenylation, and transcription termination [27,29,34]. CF Im (cleavage factor Im), a sub-complex only found in multicellular eukaryotes, is dispensable in pre-mRNA 3′-end processing. However, CF Im plays a regulatory role in alternative polyadenylation (APA) by specifically binding to an upstream UGUA motif, altering the efficiency of different poly(A) sites and generating distinct mRNA 3′UTR, which is particularly important for gene regulation [35,36]. The essential components of human histone pre-mRNA 3′-end processing have also recently been defined by the reconstitution and structural studies, which include the histone pre-mRNA cleavage complex (HCC), the stem loop binding protein (SLBP), the U7 small nuclear ribonucleoprotein (snRNP), and FLASH [30]. HCC contains CPSF73, CPSF100, symplekin, and CstF64. The molecular insights into HCC may be similar to its equivalent mCF (mammalian cleavage factor, consisting of CPSF73, CPSF100, and symplekin) in canonical macheniry, as discussed in more detail below.

Table 1. Human and yeast mRNA 3′-end processing machinery.

Human Complex	Subunit	Role	Yeast Complex	Subunit
CPSF	**mPSF [1] including**	**PAS Recognition**	**CPF**	
	CPSF160	Scaffold		Cft1
	WDR33	Scaffold, RNA binding		Pfs2
	CPSF30	Recruits hFip1, RNA binding		Yth1
	hFip1	Binds PAP		Fip1
	mCF [2] including	**Cleavage**		
	CPSF73	Endonuclease		Ysh1
	CPSF100	(Pseudo-)endonuclease, assembles		Cft2
	Symplekin	Scaffold		Pta1 *
	RBBP6	Activates cleavage, RNA binding		Mpe1
	PAP	**Poly(A) polymerase**		Pap1
CstF	CstF50	Scaffold	**CFIA**	n.d.
	CstF64	RNA binding		Rna15
	CstF77	Scaffold, assembles		Rna14
CF IIm	hPcf11	Binds Pol II		Pcf11
	hClp1	RNA kinase		Clp1
CF Im *	CFIm68	RNA binding		n.d.
	CFIm25	RNA binding		n.d.

n.d., none detected. *, dispensable for the cleavage of mRNA 3′-end processing machinery. [1], mPSF is short for mammalian polyadenylation specificity factor. [2], mCF is short for mammalian cleavage factor. The essential complexes or subunits in human canonical machinery are shown in bold.

2.2. CPSF160-WDR33 Complex Acts as a Rigid Interaction Platform

CPSF160, as a scaffold protein, contains three β-propellers (BPA, BPB, and BPC) and their relative positions are stabilized by the WD40 domain in WDR33. As a result, CPSF160 and WDR33 form a tight heterodimeric subcomplex that functions as a rigid interaction platform of the canonical machinery and recruits other factors to its distinct areas (Figure 2) [37]. The molecular architecture of human 3′-end processing machinery shows that CPSF30, CPSF100, and CstF77, respectively, bind to the BPC, CTD, and BPA regions of CPSF160, making contact with both CPSF160 and WDR33 (Figure 2) [37]. CPSF30 takes part in RNA recognition and recruits hFip1, and thereby PAP. CPSF100 forms a heterodimer with CPSF73 in mCF for pre-mRNA cleavage. CstF77 is in a complex with CstF50 and CstF64 is used for the recognition of downstream elements (DSE). CPSF30-hFip1-PAP, mCF, and CstF are all highly dynamic [37,38]. They assemble with the CPSF160-WDR33 core, waiting for additional factors and pre-mRNA substrates to initiate the RNA cleavage activity. These are described in further depth in the following paragraphs. Similar interactions were also observed in the yeast homologs, Yth1 (CPSF30) and Cft2 (CPSF100) to Pfs2 (WDR33) and Cft1 (CPSF160). Unexpectedly, recent structural and biochemical studies have shown that Mpe1 makes multiple contacts with Pfs2 and Cft1 as well [34]. Interestingly, the Cft2 protein itself somehow interferes with the binding of Mpe1 to the yeast polymerase module [34], indicating that machinery assembly is a highly regulated event that may be a kind of safeguard mechanism. More research is needed to figure out the exact mechanism. On the other hand, the overall structure of the CPSF160-WDR33 complex is similar to that of the DDB1-DDB2 DNA repair complex [39]. All three β-propellers that are present in DDB1 offer a surface for interactions with other factors [40], while the binding partner for the BPB domain of CPSF160 remains unknown.
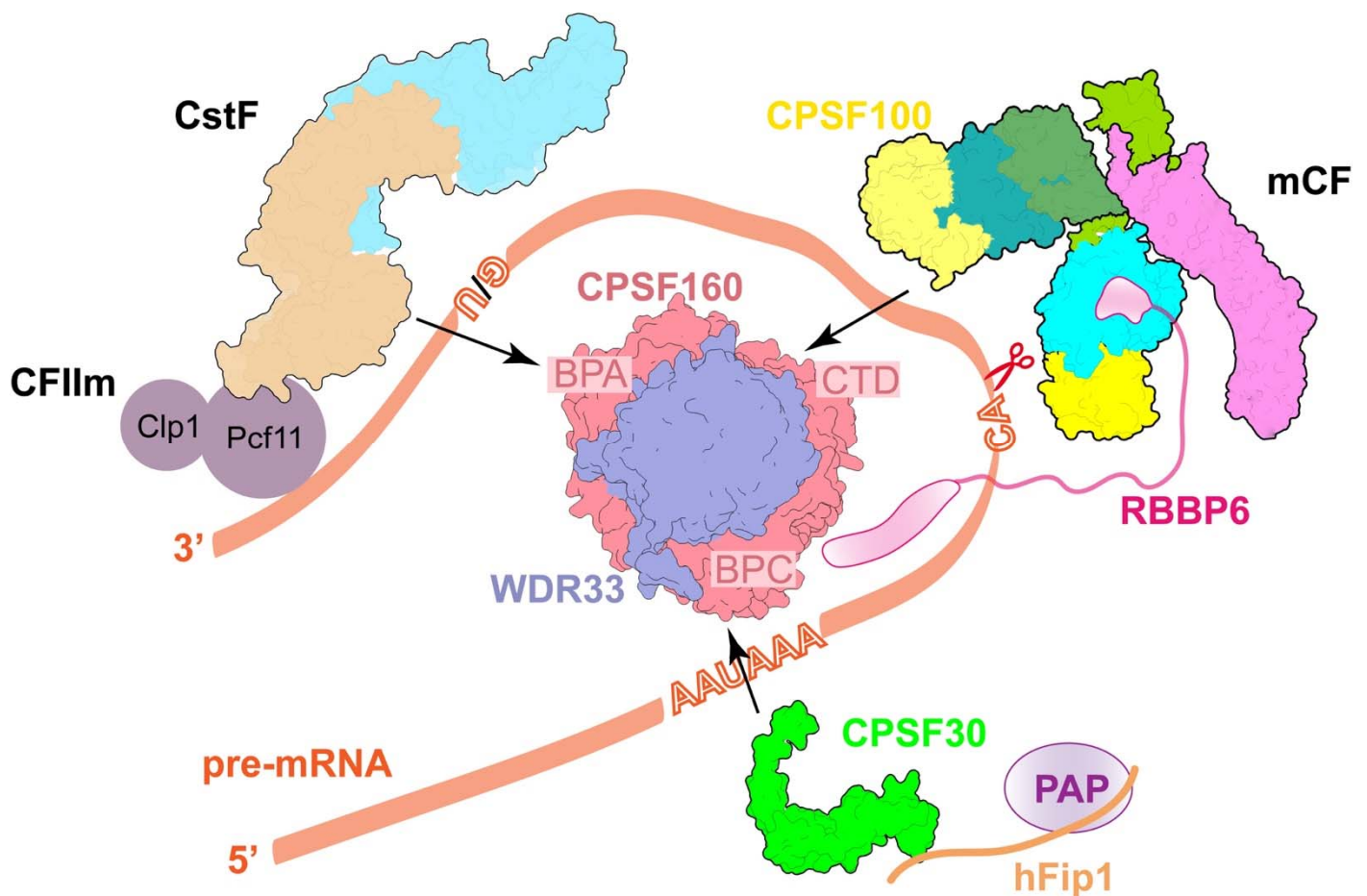
Figure 2. Schematic drawing of the minimal mammalian canonical 3′-end processing machinery. The organizations of the CPSF160-WDR33, mCF, CPSF30, and CstF complex are based on recent structural studies [34,37,41–43]. This view of the CPSF160-WDR33 complex is from the bottom, showing that the CPSF160-WDR33 complex functions as a core of the machinery to recruit CPSF30, CPSF100, and CstF77 to the BPC, CTD, and BPA regions of CPSF160, respectively. All of these bindings require both CPSF160 and WDR33. CPSF30 and WDR33 take part in PAS signal (AAUAAA) recognition. PAP is recruited by the CPSF30-hFip1 complex. Based on the high sequence conservation that exists between RBBP6 and yeast Mpe1, it is possible that RBBP6 binds to the metallo-β-lactamase domain of CPSF73 and meanwhile connects to the mPSF module. CTD, C-terminal domain.

2.3. The Recognition of PAS

The most common PAS sequence in mammalian pre-mRNAs is the highly conserved AAUAAA hexamer, which is present in over 60% of pre-mRNAs [44–46] and usually located 10~30 nucleotides upstream of the cleavage site. The molecular mechanism of PAS RNA recognition is well understood by the recent cryo-EM studies [41,42]. Structures of human mPSF with AAUAAA-containing RNA show that both CPSF30 and WDR33 are essential for high-affinity binding of PAS RNA (Figure 3A–C). The S-shaped sugar-phosphate backbone of the AAUAAA motif bends the six bases into three pairs [42]. The A1-A2 pair is specifically recognized by zinc finger 2 (ZF) of CPSF30 through hydrogen bonding interactions, π stacking, and van der Waals interactions, and ZF3 binds to the A4–A5 pair in the same manner. The U3-A6 forms a Hoogsteen base pair and interacts with two conserved phenylalanine residues in WDR33. The whole PAS is covered and stabilized by the N-terminal loop of WDR33 (Figure 3C). More details have been summarized in several review articles [47–49].
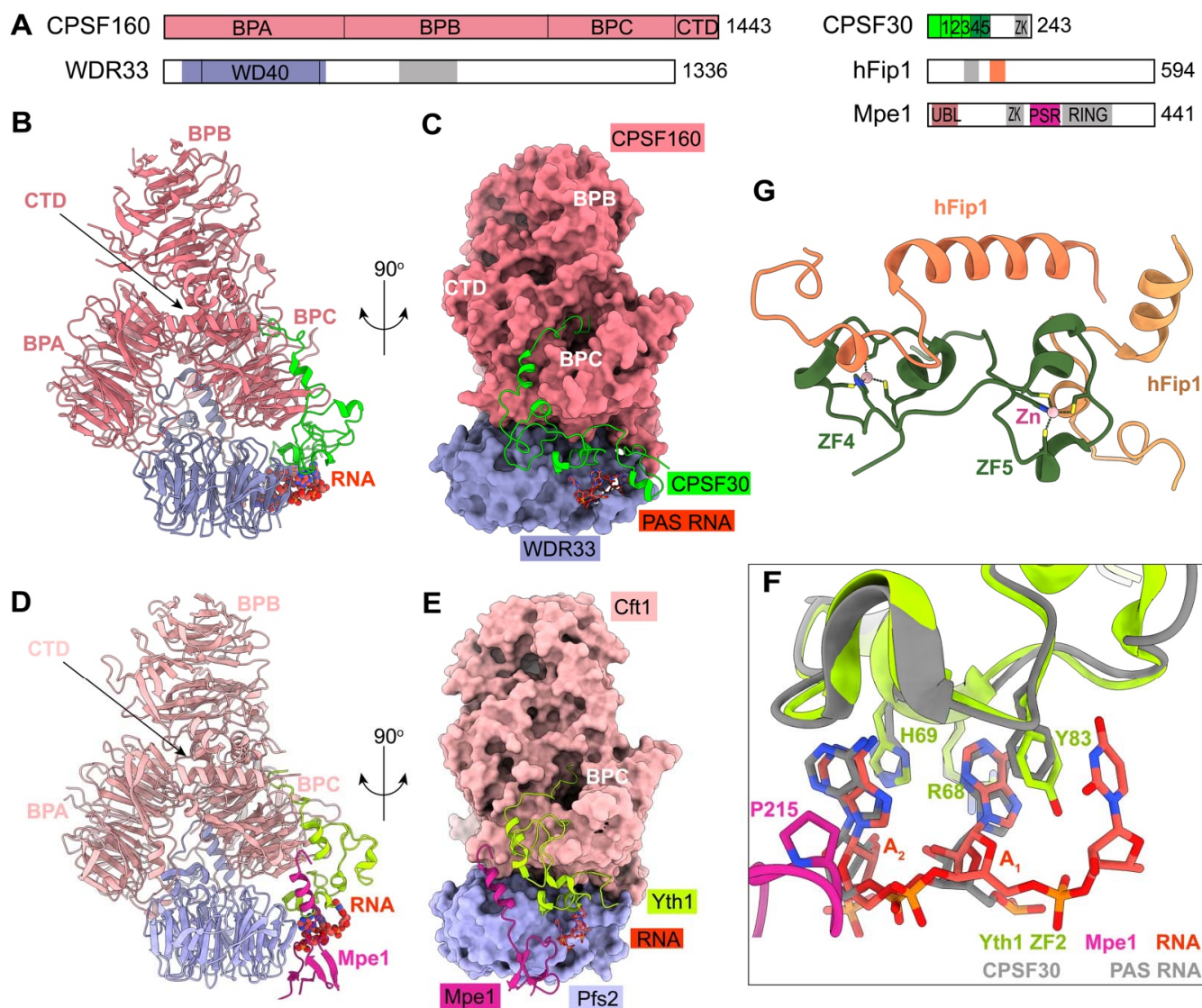
Figure 3. Molecular architecture of human mPSF and the polymerase module of yeast CPF. (**A**) Domain organizations of the subunits of mPSF, and Mpe1. The collagen-like domain in WDR33 is in gray. The zinc fingers (ZFs) of CPSF30 are in green and dark green, and the zinc knuckle in gray. The PAP-binding site of hFip1 is in gray, and the CPSF30-binding site in coral/sandy brown. The linker between CPSF30 and PAP binding sites is the central low-complexity region (LCR) in hFip1. The zinc knuckle and ring domain of yeast Mpe1 are in gray. ZK, zinc knuckle; UBL, ubiquitin-like; PSR, pre-mRNA-sensing region. (**B,C**) Structure of the human CPSF160-WDR33-CPSF30-PAS RNA complex (PDB 6DNH). The arrow points to the location of the CTD of CPSF160. The PAS RNA is shown in red orange and ball-and-stick model. The structural model in panel C is viewed after a 90° rotation around the vertical axis. The CPSF160 and WDR33 are shown as molecular surfaces. (**D,E**) Structure of the yeast Cft1-Pfs2-Yth1-Mpe1-RNA complex (PDB 7ZGP). The domains in the yeast polymerase module are given lighter colors compared to their homologs in human mPSF. The PSR domain of Mpe1 interacts directly with the polymerase module. (**F**) Comparison of the recognition mechanism of A1A2 in the human mPSF (in gray) and yeast polymerase module-Mpe1 complex (in color). The residue P215 of yeast Mpe1 directly contacts with A2. (**G**) Structure of the human CPSF30-hFip1 complex. The ZF4 and ZF5 of CPSF30 are in dark green. The two hFip1 molecules are in coral and sandy brown. Structure figures are produced with ChimeraX [50].

In contrast, PAS in yeast, also known as the positioning element, is highly degenerate and less well defined with A-rich sequences [51]. The molecular details of recognition of

the yeast positioning element AAGAA have been recently reported by structural biological research [34]. The polymerase module of yeast CPF shares structural similarities with human mPSF (Figure 3D,E). Although the polymerase module is in complex with 42-nt long CYC1 RNA, only three nucleotides of PAS can be assigned with well-ordered electron density. Particularly, the A1-A2 is recognized by ZF2 of Yth1 (CPSF30 homolog in yeast) in a way that is very comparable to that of CPSF30, suggesting that the first A-A dinucleotide of PAS is recognized by a conserved mechanism in eukaryotes (Figure 3F).

Surprisingly, besides Yth1/CPSF30 involved in PAS recognition, the structure shows that yeast Mpe1 directly contacts A2 via residue P215 in the pre-mRNA-sensing region (PSR) (Figure 3F) [34]. However, based on the biochemical evidence, Mpe1 PSR does not have any significant effect on the RNA binding affinity of the polymerase module. In contrast, the binding between Mpe1 and the polymerase module is improved by RNA binding. In addition, RNA binding induces conformational changes in Mpe1, resulting in a more compact structure [34]. All of the evidence presented above points to the possibility that Mpe1 may perform the role of a PAS motif sensor to verify the substrate entering the whole machinery. The residue P215 of Mpe1 and its surrounding residues are conserved in Mpe1 orthologs [34], suggesting that human RBBP6 may function similarly to Mpe1 in human canonical pre-mRNA 3′-end processing. Meanwhile, compared to the structure of mPSF-RNA, the helix in Mpe1 PSR overlaps with CPSF30 residues 22–34, indicating that the equivalent region in RBBP6 PSR may attach to a different position on mPSF [34]. In further research, it should be determined whether or not RBBP6 is able to sense RNA binding in the same way.

2.4. The Role of hFip1

Fip1 recruits PAP onto the machinery via binding to CPSF30. However, hFip1 and the C-terminal region of CPSF30, including ZF4-ZF5 and the zinc knuckle, are disordered in all of the existing structures, due to their flexible features relative to the mPSF core. The recent crystal structure reports that both ZF4 and ZF5 can bind to hFip1 (Figure 3G). Furthermore, the CPSF30-hFip1 complex is unexpectedly formed with 1:2 stoichiometry and further recruits two copies of PAP in vitro [43]. This finding is compatible with earlier mass spectrometry studies, showing that there are two copies of Fip1 and Pap1 in the yeast machinery [52]. The residues at the interface between CPSF30 and hFip1 are highly conserved from yeast to human, suggesting that the binding mode might be conserved in most species [43]. Indeed, the essential role of yeast Fip1 in connecting Yth1 and Pap1 has been illuminated in depth as well [38]. In a recent study, it was found that yeast Fip1 is able to stabilize the Yth1 zinc-finger fold by directly interacting with Yth1, and the region for Yth1 binding is conserved and is the only sequence that has low disorder propensity in Fip1 [38]. The rest of the part in Fip1 is largely unstructured. The linker (the central low-complexity region (LCR) in Fip1) between Yth1 and Pap1 binding sites, in particular, remains disordered, even after being assembled into the entire 3′-end processing machinery [38]. Moreover, the dynamics of the central LCR in Fip1 are important for the high efficiency of pre-mRNA 3′-end cleavage [38]. However, the exact mechanism remains unclear, which needs further investigation.

2.5. Molecular Architecture of CPSF

CPSF consists of the following two modules: mPSF that recognizes PAS and recruits PAP to catalyze the polyadenylation as discussed above, and mCF that catalyzes the cleavage reaction [53,54]. mCF contains the endonuclease CPSF73, its homolog CPSF100, and a scaffold symplekin. In contrast to mPSF, the structure of mCF is highly dynamic. CPSF73 and CPSF100 bind to sympekin via their CTDs, generating a trilobal structure, but the relative positions of the three lobes are highly variable [37]. mCF interacts with mPSF, leading to the formation of CPSF. However, CPSF shows structural variability as well. Cryo-EM structures of CPSF revealed that mCF and mPSF tether together via a small peptide (residues 460–486) in CPSF100, named PIM (the PSF interaction motif),

which forms extensive hydrophobic interactions within both CPSF160 (BPA and CTD) and WDR33 (Figure 4A) [37]. The PIM is located in a disordered, extremely hydrophilic, and long fragment (~100 residues) inside the β-CASP domain of CPSF100, which explains the highly dynamic nature within CPSF. The sequence in the PIM is highly conserved in CPSF100 homologs, and a similar peptide in yeast Ctf2 (residues 525–562, named yPIM) was also observed by the most recent cryo-EM study [34]. The binding mode of yPIM with Ctf1 and Pfs2 is similar to that with human mPSF (Figure 4B), indicating the conserved role of Ctf2 in tethering the nuclease module to the polymerase module. Interestingly, as previously mentioned, Ctf2 blocks Mpe1 PSR binding to the polymerase module, which can be recovered by adding RNA. However, biochemical data show that Ctf2 and yPIM alone have different effects on the recovery in the presence of RNA, with substoichiometric levels of Mpe1 if Ctf2 is present in the complex [34], suggesting the location of Ctf2 might be regulated in 3′-end processing activation.
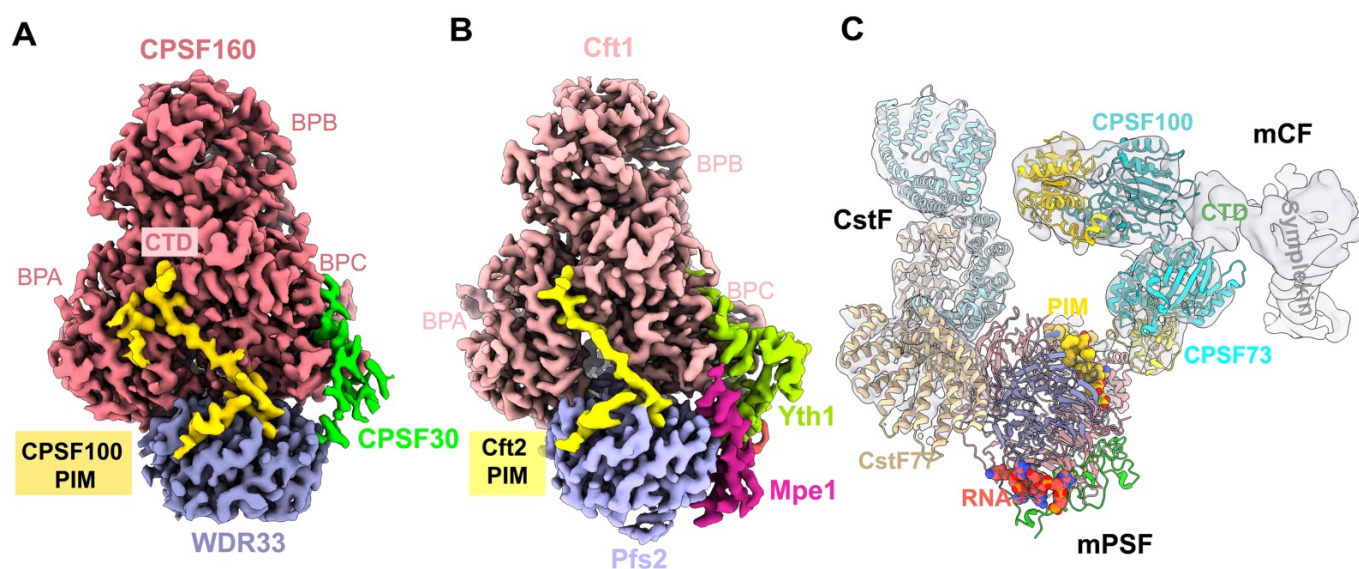

Figure 4. Overall architecture of the human canonical pre-mRNA 3′-end processing machinery. (**A**) Cryo-EM reconstruction of the human mPSF in complex with the CPSF100 PIM (in gold). (**B**) Cryo-EM reconstruction of the yeast polymerase module in complex with the Cft2 PIM (in yellow). (**C**) A model for the mPSF-mCF-CstF complex of the human canonical machinery, which is produced by overlaying the structures of the mPSF-mCF and mPSF-CstF complexes. The view is similar to that in Figure 2. mPSF module is colored as in panel A, and its view is from the bottom. The PAS RNA and PIM are shown as ball-and-stick models. The region of symplekin for CstF64 binding is disordered. The symplekin CTD is labeled in panel C. CstF77 is in blue and beige.

Due to the dynamic nature of CPSF and mCF alone, the structure of mCF is limited to a low resolution at 7.4 Å (Figure 4C) [37]. The structures of the catalytic modules of CPSF73 and CPSF100 fit well into two of the three lobes, with the only contacts by their C-terminal domain, suggesting that CPSF73 and CPSF100 form a heterodimer by their CTDs (Figure 4C). However, CPSF73 is still in a close state, since there is no room for RNA to go through its active site, indicating that conformational changes in both mCF and CPSF73 are required for the machinery activation [37]. Symplekin CTD has an elongated shape in the structure, whereas the NTD for CstF64 binding is disordered (Figure 4C) [37].

2.6. Accessory Factors CstF and CF IIm

CstF is essential for pre-mRNA 3′-end cleavage as well, but dispensable for the polyadenylation step. It includes three subunits, CstF50, CstF64, and CstF77, but their relative positions are highly variable [37]. Cryo-EM structure analyses show that the bow-shaped HAT-C dimer of CstF77 is recruited to mPSF by interacting with both the BPA domain of CPSF160 and WDR33, and there is no density for CstF50 and CstF64 (Figure 4C), which is consistent with the dynamic nature of CstF and the yeast Rna14-Rna15 complex (homologs of CstF77 and CstF64) [37,55]. The N-terminal RRM of CstF64 is responsible for the recognition of G/U-rich DSE. The C-terminus of CstF77 binds to the CstF64 hinge domain and plays an important role in boosting the RNA binding affinity of RRM [56]. CstF50 only exists in multicellular eukaryotes and has no homolog in yeast. The crystal structure reveals that CstF77 is in a complex with the CstF50 WD40 domain via a segment between its HAT domain and the CstF64 binding site [56]. However, the exact function of CstF50 is still a mystery, although some studies reveal that CstF50 may have a potential role in efficient DSE selection [56].

CF IIm that contains hClp1 and hPcf11 is also essential for cleavage activity, which has been proved by recent biochemical research [57]. Specifically, the 30 FEGP repeats of hPcf11 are critical for RNA cleavage with an unknown function. In yeast, yClp1 and yPcf11, associated with Rna14 and Rna15, together form the factor named CF IA (Table 1), which is essential for yeast machinery as well. However, the FEGP repeats are not conserved in yPcf11. Further structural research is required to figure out the interaction network of CF IIm and clarify its function in eukaryotic pre-mRNA 3′-end processing.

2.7. Activation of the Pre-mRNA 3′-End Cleavage

The overall organization of the human pre-mRNA 3′-end processing machinery shows the highly dynamic nature, which is also observed in yeast machinery and histone pre-mRNA 3′-end processing machinery. The extensive rearrangements of mCF/HCC and CPSF73 for catalyzing the processing are demonstrated by a recent cryo-EM study of histone recombinant machinery (Figure 5A–C), which is the only structure so far that shows the endonuclease CPSF73 in an open, active state, in which CPSF73 captures the pre-mRNA in its active site, ready to be cleaved (Figure 5D) [30]. The adenine base at the cleavage site forms hydrogen bonds with CPSF73 and the phosphate after this adenine is located at the CPSF73 active site, explaining why CPSF73 prefers an adenine as the cleavage site (Figure 5D). The binding mode clearly shows the molecular mechanism for the cleavage reaction. Compared with CPSF73 in a close state, the β-CASP domain makes a 17° rotation relative to the metallo-β-lactamase domain, thus creating a narrow canyon between the two domains, which can only accommodate single-stranded RNA (Figure 5E) [30]. Most importantly, the rearrangement of HCC is induced by their recognition of RNA duplex, which makes symplekin-NTD, CPSF73, and CPSF100 more compact (Figure 5A,C). A new interface is generated by the formation of a CPSF73-CPSF100 pseudo-dimer, which fixes the catalytic module of CPSF73 against the Lsm10 subunit of the U7snRNP. Lsm10 pushes CPSF73 open (Figure 5A). The targeted single-strand RNA can be guided into the canyon by the long handle and further sensed by Lsm10 (Figure 5A). As previously noted, the activation of CPSF73 in histone machinery requires the coordinated assembly of HCC with U7 snRNP and pre-mRNA.
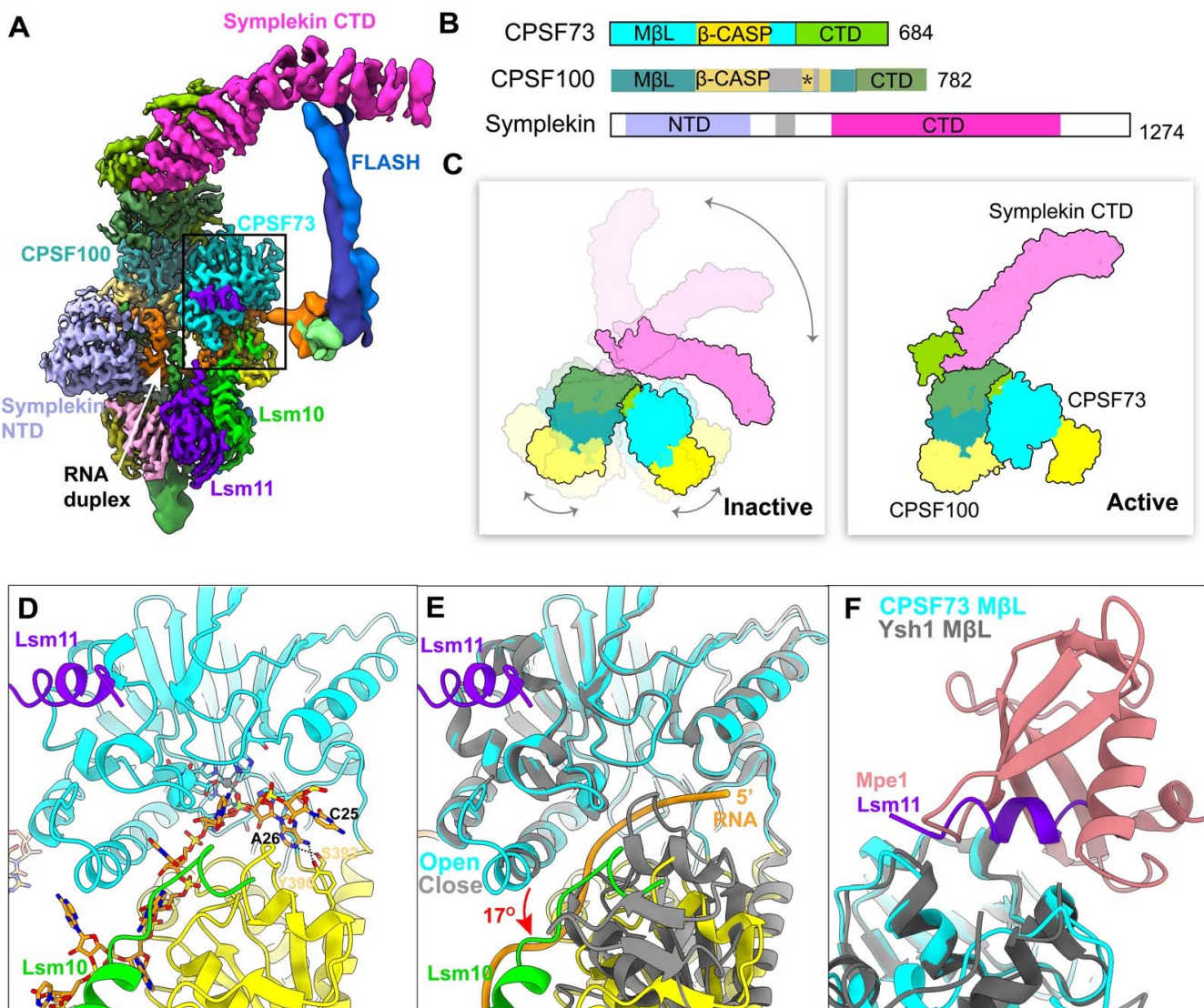
Figure 5. The activation of mCF/HCC and CPSF73. (**A**) Composite cryo-EM density map for the histone pre-mRNA 3′-end processing machinery (reproduced from [30]). The white arrow points to the RNA duplex, showing pre-mRNA in orange and U7 snRNA in green. The positions of CPSF73, CPSF100, symplekin NTD, symplekin CTD, Lsm10, Lsm11, and FLASH are labeled. Specifically, the active CPSF73 is in the black box, and additional details are shown in panels (**D**–**F**). (**B**) Domain organizations of the subunits of mCF. The disordered and highly hydrophilic segment in the β-CASP domain of CPSF100 is in gray, and the position of PIM is labeled as '*'(star). (**C**) A schematic drawing of inactive and active mCF. The conformational changes are observed in recent structural studies [30,37]. The proteins are colored as in panel B. After rearrangement, the subunits of mCF become more compact. (**D**) Recognition of the pre-mRNA substrate by CPSF73 (PDB 6V4X). The hydrogen-bond interactions are formed between A26 and the β-CASP domain of CPSF73. (**E**) Overlay of the structure of CPSF73 in the active state (in color, PDB 6V4X) with that of CPSF73 in the inactive state (in gray, PDB 2I7V). The MβL domain was used for the overlay. The β-CASP domain (yellow) makes a rotation of 17° relative to MβL domain (cyan). The positions of Lsm10 and Lsm11 are labeled. RNA is shown by orange ribbons. (**F**) Overlay of the structure of the Lsm11-CPSF73 with Mpe1-Ysh1 complex (PDB 6I1D). The MβL domain was used for the overlay, which is in cyan for CPSF73 and in gray for Ysh1. The binding sites of Mpe1 and Lsm11 on the MβL domain are overlapped. MβL, metallo-β-lactamase.

Remarkably, HCC is recruited to histone machinery by two tethering contacts, FLASH with symplekin CTD and a segment of Lsm11 with the metallo-β-lactamase domain of CPSF73 (Figure 5A). The same surface on CPSF73 is also observed in the Ysh1-Mpe1 interaction, which is important for promoting cleavage and polyadenylation in yeast (Figure 5F) [28,34]. In addition to this, the PSR motif in Mpe1 is stably bound to Pfs2 through sensing the PAS RNA/positioning element, which leads to the rearrangement of Mpe1 (Figure 3F). Thus, Mpe1 connects both the nuclease module and the polymerase module through its ubiquitin-like domain (UBL) and PSR, which is conserved in RBBP6 from the sequence alignment analysis [28,34]. Recently, two independent studies discussed the critical role of human RBBP6 in activating mRNA processing. Similar to Mpe1, RBBP6 associates with CPSF as well, but in an RNA-dependent manner. The N-terminal domain of RBBP6, including UBL, zinc knuckle, and PSR motif, is sufficient to activate canonical pre-mRNA 3′-end cleavage [27,29]. The structural and biochemical studies support the argument that RBBP6/Mpe1 may act as an essential activator of eukaryotic canonical machinery by acting as an RNA sensor and, meanwhile, tethering the endonuclease CPSF73/Ysh1 (Figure 2).

Interestingly, symplekin, a component of HCC/mCF, plays multiple important roles in mammalian pre-mRNA 3′-end processing, whereas Pta1 (homolog of symplekin) is dispensable in yeast machinery. According to the analysis of the structure of histone machinery, symplekin NTD is responsible for RNA duplex recognition, leading to HCC rearrangement (Figure 5A). In contrast, symplekin NTD is dispensable in canonical machinery [27], suggesting the mechanism of mCF rearrangement and CPSF73 activation is different in histone and canonical machinery. However, there are some differences in the current findings about the necessity of PAP and ATP in pre-mRNA 3′-end cleavage, for which further studies are needed in order to clarify their criticality.

3. The Biphasic Model of Deadenylation

The dynamic metabolism of the poly(A) tail plays a key role in the regulation of gene expression, which in turn directly determines the composition of the cellular proteome and, as a result, affects most life activities in eukaryotes. In the cytoplasm, poly(A) tails are coated by PABPC (PABPC1 in mammals, Pab1 in yeast) and shortened in a 3′-to-5′ direction by two highly conserved multiprotein complexes, PAN2-PAN3 and CCR4-NOT, in a biphasic manner (Figure 1) [58–60]. In the initial phase, PAN2-PAN3 is responsible for the slower removal of the distal part of the poly(A) tail (200–110 nt in mammals, 90 nt in yeast), while in the second, fast phase, the CCR4-NOT complex acts mainly on the final 110 nt (in mammals) to a very short one, triggering mRNA decay [61]. The exonucleases (deadenylase) in these two complexes are PAN2, CNOT6/CNOT6L and CNOT7/CNOT8, respectively. Recent insights from biochemical reconstitution and structural biology have brought to light the dichotomous role of PABPC in mRNA life, the organizational principles of yeast Pan2-Pan3 on poly(A) RNP, and the important role that non-enzymatic components play in the CCR4-NOT complex (see below).

3.1. PABPC's Role in Deadenylation

PABPC contains four RNA recognition motif (RRM) domains, which are responsible for binding poly(A) tails, followed by a proline-rich linker and a C-terminal domain. An earlier crystal structure has revealed that the poly(A) tail in an extended conformation is specifically recognized by a long, narrow groove, created by the tandem antiparallel β sheets of RRM1 and RRM2 of human PABPC1 [62]. The RRM1 is located at the 3′ region of the RNA substrate. The sequences of RRM1 and RRM2 are very similar to that of RRM3 and RRM4. However, the binding mode of RRM3 and RRM4 to poly(A) is quite different from that of the tandem RRM1-RRM2 [63,64]. This is possibly due to the length variation in the loop regions between RRMs, which may tightly associate with the RNA-binding affinity and the range of RNA length that one PABPC molecule can protect. PABPC is highly conserved in eukaryotes and both PABPC1 and Pab1 have a RNA-binding footprint

of 27 nt [65,66]; thus, a longer poly(A) tail can attract multiple PABPC1 molecules (Figure 1), which are linearly arranged on the tail and form a worm-like structure, as observed in negative-stain studies [67]. PABPC plays a contradictory role in deadenylation. On the one hand, it prevents non-specific degradation of mRNA by specifically binding the poly(A) tail [68], and meanwhile stimulates translation [69]; on the other hand, as discussed below, it promotes poly(A) tail shortening and mRNA decay by interacting with PAN2-PAN3 and CCR4-NOT complexes [60,61,70].

3.2. The Architecture and Recognition Mechanism of Poly(A) RNP by Pan2-Pan3 Complex

PAN2 is a deadenylase, consisting of an N-terminal WD40 domain, an inactive ubiquitin C-terminal hydrolase (UCH), and an exonuclease domain at its C terminus (Figure 6A). PAN2, by forming a complex with PAN3 in a 1:2 stoichiometry (Figure 6A), contributes to the initial trimming of the distal part of the poly(A) tail [71], which requires two or more PABPC1/Pab1 molecules in this process. The recent cryo-EM reconstruction of the yeast Pan2-Pan3 in complex with a poly(A) RNP composed of 90 adenosines and three Pab1 has illuminated the molecular mechanism for why Pan2-Pan3 preferentially acts on a longer poly(A) RNP and how Pab1 stimulates this process (Figure 6) [60]. Based on comprehensively structural and biochemical analyses, it was found that the asymmetric homodimer formed by Pan3 functions as an important interaction hub, recruiting Pan2 via a long linker between the WD40 domain and UCH domain of Pan2, and forming different functional interfaces on opposite sides of the Pan3 homodimer (Figure 6A), which is critical for recognizing the Pab1-90A RNP [60,72]. The Pab1-90A RNP binds to one side of the Pan2-Pan3 complex (Figure 6B,C). The architecture of the poly(A) RNP in this structure interestingly presents repeated arches, which are shaped by RRM3, the Pab1-Pab1 oligomerization interface, and the interacting position with Pan2-Pan3 (Figure 6C–E). The Pab1-Pab1 oligomerization interface adopts a sharp "V" shape, which is formed by a long conserved α-helix (linker helix) from RRM4 and the RRM1 of the following Pab1 molecule (Figure 6D,E). The interface between the first and second Pab1 is recognized by the Pan3 pseudokinase domain (Figure 6D), while the second interface between the second and third Pab1 interacts with the WD40 domain of Pan2 (Figure 6E), indicating that the WD40 domain is the key checkpoint for the length of the poly(A) tail. Moreover, RRM1 of the first instance of Pab1 binding at the most 3′ end drives the poly(A) tail into the active site by interacting with the RNase domain of Pan2. Recent crystal structures show that the deadenylation specificity of Pan2 is determined by recognizing the A-form helical conformation formed by poly(A) RNA, rather than canonical base-specific contacts [73]. In summary, Pab1 stimulates deadenylation by providing the main structural features for the Pan2-Pan3 complex recognition, which is critical for its length specificity.

3.3. The Roles of Non-Enzymatic Components in CCR4-NOT Complex

The CCR4-NOT complex is a multi-subunit complex that is partially redundant with PAN2-PAN3, but mainly focuses on removing adenosines proximal to the 3′ UTR [59]. Cryo-EM structures of the yeast Ccr4-Not complex have been reported in the past several years [74]. However, the resolution is not sufficient to elucidate the detailed molecular mechanism of the complex assembly and deadenylation.
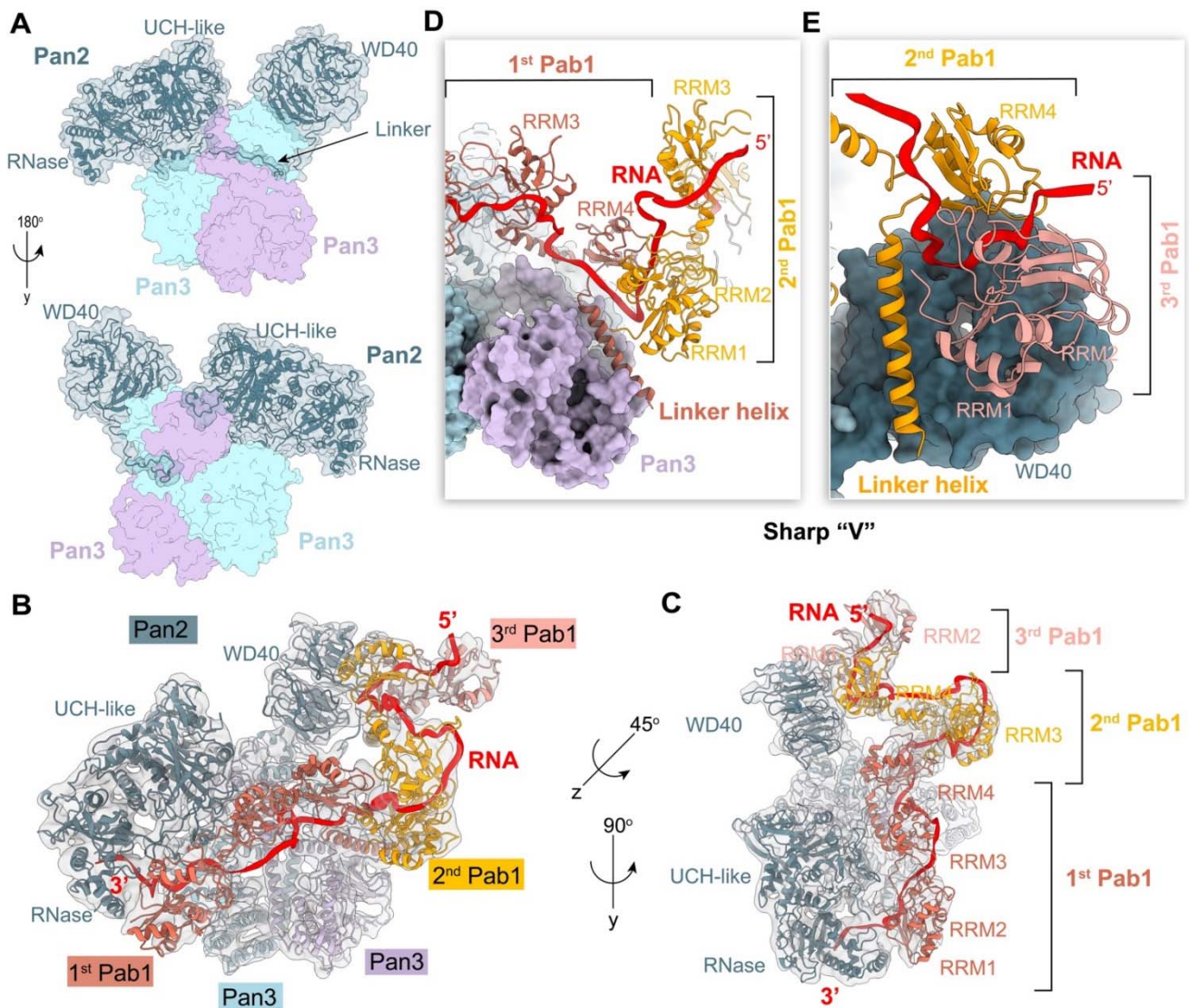
Figure 6. The architecture of the poly(A) RNP-Pan2-Pan3 complex. (**A**) Structure of the apo Pan2-Pan3 complex [60]. The two Pan3 molecules are in light blue and purple, respectively. The molecular surface of Pan2 is shown as a transparent surface. The black arrow indicates the important linker between the UCH domain and the WD40 domain of Pan2. (**B**,**C**) Cryo-EM density for the entire poly(A) RNP-Pan2-Pan3 complex (gray, semitransparent), in side view (**B**) and top view (**C**), fitted with the model of Pab1, Pan2, and Pan3. The poly(A) RNA is shown by a red ribbon. PDB 6R5K. (**D**,**E**) The interaction between the Pan2-Pan3 complex and the Pab1-Pab1 oligomerization interface. The pseudokinase domain of Pan3 (purple) and the WD40 domain of Pan2 are shown on the molecular surface. The linker helix in RRM4 is labeled. The RRM4 and the RRM1-RRM2 domains of the following Pab1 shape the substate RNA into a sharp "V" shape.

The CCR4-NOT complex is evolutionarily conserved from yeast to human, but the composition shows species specificity and differs among humans, Drosophila, and yeast [75–78]. CCR4-NOT in humans contains at least eight subunits, including the following two deadenylases: CNOT6/CNOT6L/CCR4 (also known as Ccr4 in yeast) and CNOT7/CNOT8/CAF1 (also known as Ccr4-associated factor 1, Caf1 in yeast) [79]. Both deadenylases contribute to adenosine removal, but present different activities with PABPC; CCR4 performs its function on PABPC-protected A tails, whereas the activity of CAF1 is blocked by PABPC, resulting in only trimming naked poly(A) RNA and the generation of deadenylation cycles every 27 nt (Figure 1) [61]. However, the molecular basis of the

completely opposite effects generated by PABPC with the CCR4-NOT complex remains a mystery. CNOT1 is the largest subunit (>200 kDa) in the complex, acting as an essential scaffold to assemble the CCR4-NOT complex [76], which is roughly divided into the following four modules: C-terminal NOT module (binding to CNOT2 and CNOT3) [80]; catalytic module on the MIF4G domain of CNOT1 (recruiting the two deadenylases) [81]; ubiquitin ligase module next to the catalytic module (binding to CNOT9 and, in some conditions, an E3 ubiquitin ligase, CNOT4) [82,83]; N-terminal module (recruiting CNOT10 and CNOT11). The non-enzymatic components, CNOT10 and CNOT11, are conserved in eukaryotes, except for yeast [76,84,85]. Both CNOT9 and CNOT10:CNOT11 modules bind to RNA directly and stimulate deadenylation [86]. The non-enzymatic modules in the human CCR4-NOT complex play important roles in increasing deadenylation activity and sequence selectivity, probably due to their ability of providing more RNA-binding sites to cooperatively strengthen the interaction between the complex and the RNA substrate [86]. However, so far, only limited fragmentary information is available to explain how these subunits act together and form the CCR4-NOT complex. The recent crystal structures of human CCR4-CAF1 show a similar binding mode to that of *S. cerevisiae*. CCR4 binds to CAF1 via the CCR4 LRR domain, which seems to tether the CCR4 nuclease domain by a flexible linker [87]. Therefore, the distance between the two nucleases is probably regulated in the context of the CCR4-NOT complex, which is observed in the latest structural studies on the human catalytic module (CNOT1(MIF4G)-CAF1-CCR4) [88], illustrating that the position of the nuclease domain of CCR4 may be restricted by the non-enzymatic components in the CCR4-NOT complex. The biochemical data also indicate that CAF1 functions as a tunable enzyme, which is highly sensitive to experimental conditions, such as pH, magnesium, and zinc ions [87].

4. Conclusions and Future Perspectives

The catalytic activities of the endonuclease CPSF73/Yth1 in polyadenylation, and the deadenylases CAF1 and CCR4 in deadenylation are quite poor if they function as stand-alone enzymes [15,86]. To achieve full functional activities, they have to be assembled into large machineries, with multiple non-enzymatic components. These components are essential for promoting the subcomplex rearrangement and stabilizing the assembly of the machinery, as well as restricting the flexibility by providing additional RNA-binding sites or coordinating a series of interactions among different factors. As a result, these machineries possess significantly stronger enzymatic activities and improved substrate specificity. The overview of the recent progress in structural and biochemical studies of polyadenylation and deadenylation shows that the dynamic features of the machineries play important regulatory roles. However, many questions about the molecular mechanisms of the whole machinery and the role of each component are still pending. Recent structural studies on histone pre-mRNA processing have shown the binding mode of the active CPSF73 with the substrate RNA, suggesting that the distance from the HDE-U7 duplex is a key factor for defining the cleavage site. However, it remains unclear how the cleavage position in canonical machinery is determined. Based on successful reconstitution with recombinant proteins in vitro, the minimal list of essential components in canonical pre-mRNA 3′-end processing machinery has been defined, which indicates that CF IIm and RBBP6/Mpe1 play critical roles in the cleavage activation step [27,29,34,57]. Nevertheless, further studies are required to reveal more details at the molecular level.

In deadenylation, the PAN2-PAN3 and CCR4-NOT complex have different and partially redundant substrate preferences. Although structural analyses have unveiled why yeast Pan2-Pan3 works on longer poly(A) tails [60], the structural details of the CCR4-NOT complex and how PABPC affects CCR4-NOT at the molecular level remain to be determined. Furthermore, it was thought that long poly(A) tails give better stability to mRNAs, but several very stable transcripts with short poly(A) tails have been recently discovered [89,90]. Therefore, the exact relationship between the length of the poly(A) tail and mRNA stability needs further investigation. On the other hand, the poly(A) tail can be

elongated not only in the nucleus, but also in the cytoplasm. Cytoplasmic polyadenylation increases protein expression of dormant mRNAs with short poly(A) tails, thereby regulating the translation of specific mRNAs at various times and locations [91]. Notably, there is an alternative pathway for poly(A) tail 3′-end protection from deadenylation, in which the poly(A) tail forms RNA triple helices with cis-acting U-rich RNA elements [92]. Recent structural studies have revealed that a poly(A) 3′-end binding pocket is formed via a steric mechanism, allowing for unexpected kinds of RNA-RNA interaction [93,94]. Overall, the recent progress provides fresh molecular insights into poly(A) tail biology. More exciting research is needed to figure out how poly(A) tail regulation works at the molecular level.

References

1. Bienroth, S.; Keller, W.; Wahle, E. Assembly of a processive messenger RNA polyadenylation complex. *EMBO J.* **1993**, *12*, 585–594. [CrossRef] [PubMed]
2. Wahle, E. A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell* **1991**, *66*, 759–768. [CrossRef]
3. Kühn, U.; Buschmann, J.; Wahle, E. The nuclear poly(A) binding protein of mammals, but not of fission yeast, participates in mRNA polyadenylation. *RNA* **2017**, *23*, 473–482. [CrossRef]
4. Keller, R.W.; Kühn, U.; Aragón, M.; Bornikova, L.; Wahle, E.; Bear, D.G. The nuclear poly(A) binding protein, PABP2, forms an oligomeric particle covering the length of the poly(A) tail. *J. Mol. Biol.* **2000**, *297*, 569–583. [CrossRef] [PubMed]
5. Darnell, J.E.; Wall, R.; Tushinski, R.J. An Adenylic Acid-Rich Sequence in Messenger RNA of HeLa Cells and Its Possible Relationship to Reiterated Sites in DNA. *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1321–1325. [CrossRef]
6. Edmonds, M.; Vaughan, M.H., Jr.; Nakazato, H. Polyadenylic acid sequences in the heterogeneous nuclear RNA and rapidly-labeled polyribosomal RNA of HeLa cells: Possible evidence for a precursor relationship. *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1336–1340. [CrossRef]
7. Berkovits, B.D.; Mayr, C. Alternative 3′ UTRs act as scaffolds to regulate membrane protein localization. *Nature* **2015**, *522*, 363–367. [CrossRef]
8. Djebali, S.; Davis, C.A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F.; et al. Landscape of transcription in human cells. *Nature* **2012**, *489*, 101–108. [CrossRef] [PubMed]
9. Fu, Y.; Chen, L.; Chen, C.; Ge, Y.; Kang, M.; Song, Z.; Li, J.; Feng, Y.; Huo, Z.; He, G.; et al. Crosstalk between alternative polyadenylation and miRNAs in the regulation of protein translational efficiency. *Genome Res.* **2018**, *28*, 1656–1663. [CrossRef] [PubMed]
10. McLaughlin, C.S.; Warner, J.R.; Edmonds, M.; Nakazato, H.; Vaughan, M.H. Polyadenylic acid sequences in yeast messenger ribonucleic acid. *J. Biol. Chem.* **1973**, *248*, 508–519. [CrossRef]
11. Curinha, A.; Braz, S.O.; Pereira-Castro, I.; Cruz, A.; Moreira, A. Implications of polyadenylation in health and disease. *Nucleus* **2014**, *5*, 508–519. [CrossRef] [PubMed]
12. Ouyang, J.; Sun, W.; Xiao, X.; Li, S.; Jia, X.; Zhou, L.; Wang, P.; Zhang, Q. CPSF1 mutations are associated with early-onset high myopia and involved in retinal ganglion cell axon projection. *Hum. Mol. Genet.* **2019**, *28*, 1959–1970. [CrossRef] [PubMed]
13. Zhang, X.; Virtanen, A.; Kleiman, F.E. To polyadenylate or to deadenylate: That is the question. *Cell Cycle* **2010**, *9*, 4437–4449. [CrossRef]
14. Balbo, P.B.; Bohm, A. Mechanism of poly(A) polymerase: Structure of the enzyme-MgATP-RNA ternary complex and kinetic analysis. *Structure* **2007**, *15*, 1117–1131. [CrossRef] [PubMed]
15. Mandel, C.R.; Kaneko, S.; Zhang, H.; Gebauer, D.; Vethantham, V.; Manley, J.L.; Tong, L. Polyadenylation factor CPSF-73 is the pre-mRNA 3′-end-processing endonuclease. *Nature* **2006**, *444*, 953–956. [CrossRef]
16. Desterro, J.; Bak-Gordon, P.; Carmo-Fonseca, M. Targeting mRNA processing as an anticancer strategy. *Nat. Rev. Drug Discov.* **2019**, *19*, 112–129. [CrossRef]

17. Sonoiki, E.; Ng, C.L.; Lee, M.C.S.; Guo, D.; Zhang, Y.-K.; Zhou, Y.; Alley, M.R.K.; Ahyong, V.; Sanz, L.M.; Lafuente-Monasterio, M.J.; et al. A potent antimalarial benzoxaborole targets a Plasmodium falciparum cleavage and polyadenylation specificity factor homologue. *Nat. Commun.* **2017**, *8*, 14574. [CrossRef]

18. Liu, H.; Moore, C.L. On the Cutting Edge: Regulation and Therapeutic Potential of the mRNA 3′ End Nuclease. *Trends Biochem. Sci.* **2021**, *46*, 772–784. [CrossRef]

19. Ross, N.T.; Lohmann, F.; Carbonneau, S.; Fazal, A.; Weihofen, W.A.; Gleim, S.; Salcius, M.; Sigoillot, F.; Henault, M.; Carl, S.H.; et al. CPSF3-dependent pre-mRNA processing as a druggable node in AML and Ewing's sarcoma. *Nat. Chem. Biol.* **2020**, *16*, 50–59. [CrossRef]

20. Murphy, M.R.; Doymaz, A.; Kleiman, F.E. Poly(A) tail dynamics: Measuring polyadenylation, deadenylation and poly(A) tail length. *Methods Enzymol.* **2021**, *655*, 265–290. [CrossRef]

21. Marzluff, W.F.; Wagner, E.J.; Duronio, R.J. Metabolism and regulation of canonical histone mRNAs: Life without a poly(A) tail. *Nat. Rev. Genet.* **2008**, *9*, 843–854. [CrossRef] [PubMed]

22. Sullivan, K.; Steiniger, M.; Marzluff, W.F. A Core Complex of CPSF73, CPSF100, and Symplekin May Form Two Different Cleavage Factors for Processing of Poly(A) and Histone mRNAs. *Mol. Cell* **2009**, *34*, 322–332. [CrossRef] [PubMed]

23. Butler, J.S.; Platt, T. RNA Processing Generates the Mature 3′ End of Yeast *CYC1* Messenger RNA in Vitro. *Science* **1988**, *242*, 1270–1274. [CrossRef] [PubMed]

24. Moore, C.L.; Sharp, P.A. Accurate cleavage and polyadenylation of exogenous RNA substrate. *Cell* **1985**, *41*, 845–855. [CrossRef]

25. Gutierrez, P.A.; Baughman, K.; Sun, Y.; Tong, L. A real-time fluorescence assay for CPSF73, the nuclease for pre-mRNA 3′-end processing. *RNA* **2021**, *27*, 1148–1154. [CrossRef]

26. Yang, X.-C.; Sun, Y.; Aik, W.S.; Marzluff, W.F.; Tong, L.; Dominski, Z. Studies with recombinant U7 snRNP demonstrate that CPSF73 is both an endonuclease and a 5′-3′ exonuclease. *RNA* **2020**, *26*, 1345–1359. [CrossRef]

27. Boreikaite, V.; Elliott, T.S.; Chin, J.W.; Passmore, L.A. RBBP6 activates the pre-mRNA 3′ end processing machinery in humans. *Genes Dev.* **2022**, *36*, 210–224. [CrossRef]

28. Hill, C.H.; Boreikaite, V.; Kumar, A.; Casañal, A.; Kubík, P.; Degliesposti, G.; Maslen, S.; Mariani, A.; von Loeffelholz, O.; Girbig, M.; et al. Activation of the Endonuclease that Defines mRNA 3′ Ends Requires Incorporation into an 8-Subunit Core Cleavage and Polyadenylation Factor Complex. *Mol. Cell* **2019**, *73*, 1217–1231.e11. [CrossRef] [PubMed]

29. Schmidt, M.; Kluge, F.; Sandmeir, F.; Kühn, U.; Schäfer, P.; Tüting, C.; Ihling, C.; Conti, E.; Wahle, E. Reconstitution of 3′ end processing of mammalian pre-mRNA reveals a central role of RBBP6. *Genes Dev.* **2022**, *36*, 195–209. [CrossRef] [PubMed]

30. Sun, Y.; Zhang, Y.; Aik, W.S.; Yang, X.C.; Marzluff, W.F.; Walz, T.; Dominski, Z.; Tong, L. Structure of an active human histone pre-mRNA 3′-end processing machinery. *Science* **2020**, *367*, 700–703. [CrossRef]

31. Yoon, Y.; Shi, Y. Human pre-mRNA 3′ end processing: Reconstituting is believing. *Genes Dev.* **2022**, *36*, 106–107. [CrossRef] [PubMed]

32. Di Giammartino, D.C.; Li, W.; Ogami, K.; Yashinskie, J.J.; Hoque, M.; Tian, B.; Manley, J.L. RBBP6 isoforms regulate the human polyadenylation machinery and modulate expression of mRNAs with AU-rich 3′ UTRs. *Genes Dev.* **2014**, *28*, 2248–2260. [CrossRef] [PubMed]

33. Shi, Y.; Di Giammartino, D.C.; Taylor, D.; Sarkeshik, A.; Rice, W.J.; Yates, J.R., 3rd; Frank, J.; Manley, J.L. Molecular architecture of the human pre-mRNA 3′ processing complex. *Mol. Cell* **2009**, *33*, 365–376. [CrossRef] [PubMed]

34. Rodriguez-Molina, J.B.; O'Reilly, F.J.; Fagarasan, H.; Sheekey, E.; Maslen, S.; Skehel, J.M.; Rappsilber, J.; Passmore, L.A. Mpe1 senses the binding of pre-mRNA and controls 3′ end processing by CPF. *Mol. Cell* **2022**, *82*, 2490–2504.e12. [CrossRef] [PubMed]

35. Zhu, Y.; Wang, X.; Forouzmand, E.; Jeong, J.; Qiao, F.; Sowd, G.A.; Engelman, A.N.; Xie, X.; Hertel, K.J.; Shi, Y. Molecular Mechanisms for CFIm-Mediated Regulation of mRNA Alternative Polyadenylation. *Mol. Cell* **2017**, *69*, 62–74.e4. [CrossRef]

36. Hardy, J.G.; Norbury, C.J. Cleavage factor Im (CFIm) as a regulator of alternative polyadenylation. *Biochem. Soc. Trans.* **2016**, *44*, 1051–1057. [CrossRef]

37. Zhang, Y.; Sun, Y.; Shi, Y.; Walz, T.; Tong, L. Structural Insights into the Human Pre-mRNA 3′-End Processing Machinery. *Mol. Cell* **2020**, *77*, 800–809.e6. [CrossRef]

38. Kumar, A.; Yu, C.W.; Rodríguez-Molina, J.B.; Li, X.-H.; Freund, S.M.; Passmore, L.A. Dynamics in Fip1 regulate eukaryotic mRNA 3′ end processing. *Genes Dev.* **2021**, *35*, 1510–1526. [CrossRef]

39. Scrima, A.; Koníčková, R.; Czyzewski, B.K.; Kawasaki, Y.; Jeffrey, P.D.; Groisman, R.; Nakatani, Y.; Iwai, S.; Pavletich, N.P.; Thomä, N.H. Structural Basis of UV DNA-Damage Recognition by the DDB1–DDB2 Complex. *Cell* **2008**, *135*, 1213–1223. [CrossRef]

40. Angers, S.; Li, T.; Yi, X.; MacCoss, M.J.; Moon, R.T.; Zheng, N. Molecular architecture and assembly of the DDB1–CUL4A ubiquitin ligase machinery. *Nature* **2006**, *443*, 590–593. [CrossRef]

41. Clerici, M.; Faini, M.; Muckenfuss, L.M.; Aebersold, R.; Jinek, M. Structural basis of AAUAAA polyadenylation signal recognition by the human CPSF complex. *Nat. Struct. Mol. Biol.* **2018**, *25*, 135–138. [CrossRef] [PubMed]

42. Sun, Y.; Zhang, Y.; Hamilton, K.; Manley, J.L.; Shi, Y.; Walz, T.; Tong, L. Molecular basis for the recognition of the human AAUAAA polyadenylation signal. *Proc. Natl. Acad. Sci.* **2017**, *115*, E1419–E1428. [CrossRef] [PubMed]

43. Hamilton, K.; Tong, L. Molecular mechanism for the interaction between human CPSF30 and hFip1. *Genes Dev.* **2020**, *34*, 1753–1761. [CrossRef] [PubMed]

44. Beaudoing, E.; Freier, S.; Wyatt, J.R.; Claverie, J.-M.; Gautheret, D. Patterns of Variant Polyadenylation Signal Usage in Human Genes. *Genome Res.* **2000**, *10*, 1001–1010. [CrossRef]

45. Tian, B.; Hu, J.; Zhang, H.; Lutz, C.S. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.* **2005**, *33*, 201–212. [CrossRef] [PubMed]

46. Hamilton, K.; Sun, Y.; Tong, L. Biophysical characterizations of the recognition of the AAUAAA polyadenylation signal. *RNA* **2019**, *25*, 1673–1680. [CrossRef] [PubMed]

47. Kumar, A.; Clerici, M.; Muckenfuss, L.M.; A Passmore, L.; Jinek, M. Mechanistic insights into mRNA 3′-end processing. *Curr. Opin. Struct. Biol.* **2019**, *59*, 143–150. [CrossRef] [PubMed]

48. Sun, Y.; Hamilton, K.; Tong, L. Recent molecular insights into canonical pre-mRNA 3′-end processing. *Transcription* **2020**, *11*, 83–96. [CrossRef]

49. Thore, S.; Fribourg, S. Structural insights into the 3′-end mRNA maturation machinery: Snapshot on polyadenylation signal recognition. *Biochimie* **2019**, *164*, 105–110. [CrossRef]

50. Goddard, T.D.; Huang, C.C.; Meng, E.C.; Pettersen, E.F.; Couch, G.S.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **2017**, *27*, 14–25. [CrossRef]

51. Russo, P.; Li, W.Z.; Guo, Z.; Sherman, F. Signals that produce 3′ termini in CYC1 mRNA of the yeast Saccharomyces cerevisiae. *Mol. Cell. Biol.* **1993**, *13*, 7836–7849. [CrossRef] [PubMed]

52. Clerici, M.; Faini, M.; Aebersold, R.; Jinek, M. Structural insights into the assembly and polyA signal recognition mechanism of the human CPSF complex. *eLife* **2017**, *6*, e33111. [CrossRef]

53. Chan, S.L.; Huppertz, I.; Yao, C.; Weng, L.; Moresco, J.J.; Yates, J.R.; Ule, J.; Manley, J.L.; Shi, Y. CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3′ processing. *Genes Dev.* **2014**, *28*, 2370–2380. [CrossRef]

54. Schönemann, L.; Kühn, U.; Martin, G.; Schäfer, P.; Gruber, A.R.; Keller, W.; Zavolan, M.; Wahle, E. Reconstitution of CPSF active in polyadenylation: Recognition of the polyadenylation signal by WDR33. *Genes Dev.* **2014**, *28*, 2381–2393. [CrossRef]

55. Paulson, A.R.; Tong, L. Crystal structure of the Rna14–Rna15 complex. *RNA* **2012**, *18*, 1154–1162. [CrossRef] [PubMed]

56. Yang, W.; Hsu, P.L.; Yang, F.; Song, J.-E.; Varani, G. Reconstitution of the CstF complex unveils a regulatory role for CstF-50 in recognition of 3′-end processing signals. *Nucleic Acids Res.* **2017**, *46*, 493–503. [CrossRef] [PubMed]

57. Schäfer, P.; Tüting, C.; Schönemann, L.; Kühn, U.; Treiber, T.; Treiber, N.; Ihling, C.; Graber, A.; Keller, W.; Meister, G.; et al. Reconstitution of mammalian cleavage factor II involved in 3′ processing of mRNA precursors. *RNA* **2018**, *24*, 1721–1737. [CrossRef]

58. Jolles, B.; Jean-Jean, O. Poly(A) tail degradation in human cells: ATF4 mRNA as a model for biphasic deadenylation. *Biochimie* **2021**, *185*, 128–134. [CrossRef]

59. Passmore, L.A.; Coller, J. Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. *Nat. Rev. Mol. Cell Biol.* **2021**, *23*, 93–106. [CrossRef]

60. Schäfer, I.B.; Yamashita, M.; Schuller, J.M.; Schüssler, S.; Reichelt, P.; Strauss, M.; Conti, E. Molecular Basis for poly(A) RNP Architecture and Recognition by the Pan2-Pan3 Deadenylase. *Cell* **2019**, *177*, 1619–1631.e21. [CrossRef] [PubMed]

61. Yi, H.; Park, J.; Ha, M.; Lim, J.; Chang, H.; Kim, V.N. PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. *Mol. Cell* **2018**, *70*, 1081–1088.e5. [CrossRef] [PubMed]

62. Deo, R.C.; Bonanno, J.B.; Sonenberg, N.; Burley, S.K. Recognition of Polyadenylate RNA by the Poly(A)-Binding Protein. *Cell* **1999**, *98*, 835–845. [CrossRef]

63. Burd, C.G.; Matunis, E.L.; Dreyfuss, G. The multiple RNA-binding domains of the mRNA poly(A)-binding protein have different RNA-binding activities. *Mol. Cell. Biol.* **1991**, *11*, 3419–3424. [CrossRef] [PubMed]

64. Kühn, U.; Pieler, T. XenopusPoly(A) Binding Protein: Functional Domains in RNA Binding and Protein–Protein Interaction. *J. Mol. Biol.* **1996**, *256*, 20–30. [CrossRef] [PubMed]

65. Baer, B.W.; Kornberg, R.D. Repeating structure of cytoplasmic poly(A)-ribonucleoprotein. *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 1890–1892. [CrossRef]

66. Smith, B.L.; Gallie, D.R.; Leb, H.; Hansma, P.K. Visualization of Poly(A)-Binding Protein Complex Formation with Poly(A) RNA Using Atomic Force Microscopy. *J. Struct. Biol.* **1997**, *119*, 109–117. [CrossRef]

67. Sawazaki, R.; Imai, S.; Yokogawa, M.; Hosoda, N.; Hoshino, S.-I.; Mio, M.; Mio, K.; Shimada, I.; Osawa, M. Characterization of the multimeric structure of poly(A)-binding protein on a poly(A) tail. *Sci. Rep.* **2018**, *8*, 1455. [CrossRef]

68. Ford, L.P.; Bagga, P.S.; Wilusz, J. The poly(A) tail inhibits the assembly of a 3′-to-5′ exonuclease in an in vitro RNA stability system. *Mol. Cell. Biol.* **1997**, *17*, 398–406. [CrossRef]

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

70. Webster, M.W.; Chen, Y.-H.; Stowell, J.A.; Alhusaini, N.; Sweet, T.; Graveley, B.R.; Coller, J.; Passmore, L.A. mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. *Mol. Cell* **2018**, *70*, 1089–1100.e8. [CrossRef]

71. Uchida, N.; Hoshino, S.-I.; Katada, T. Identification of a Human Cytoplasmic Poly(A) Nuclease Complex Stimulated by Poly(A)-binding Protein. *J. Biol. Chem.* **2004**, *279*, 1383–1391. [CrossRef] [PubMed]

72. Mace, P.D.; Murphy, J.M. There's more to death than life: Noncatalytic functions in kinase and pseudokinase signaling. *J. Biol. Chem.* **2021**, *296*, 100705. [CrossRef] [PubMed]

73. Tang, T.T.L.; Stowell, J.A.W.; Hill, C.H.; Passmore, L.A. The intrinsic structure of poly(A) RNA determines the specificity of Pan2 and Caf1 deadenylases. *Nat. Struct. Mol. Biol.* **2019**, *26*, 433–442. [CrossRef] [PubMed]

74. Ukleja, M.; Cuéllar, J.; Siwaszek, A.; Kasprzak, J.M.; Czarnocki-Cieciura, M.; Bujnicki, J.M.; Dziembowski, A.; Valpuesta, J. The architecture of the Schizosaccharomyces pombe CCR4-NOT complex. *Nat. Commun.* **2016**, 7, 10433. [CrossRef]
75. Albert, T.K. Isolation and characterization of human orthologs of yeast CCR4-NOT complex subunits. *Nucleic Acids Res.* **2000**, 28, 809–817. [CrossRef]
76. Bawankar, P.; Loh, B.; Wohlbold, L.; Schmidt, S.; Izaurralde, E. NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biol.* **2013**, 10, 228–244. [CrossRef]
77. Chen, J.; Rappsilber, J.; Chiang, Y.-C.; Russell, P.; Mann, M.; Denis, C.L. Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex. *J. Mol. Biol.* **2001**, 314, 683–694. [CrossRef]
78. Temme, C.; Zhang, L.; Kremmer, E.; Ihling, C.; Chartier, A.; Sinz, A.; Simonelig, M.; Wahle, E. Subunits of the *Drosophila* CCR4-NOT complex and their roles in mRNA deadenylation. *RNA* **2010**, 16, 1356–1370. [CrossRef]
79. Mostafa, D.; Takahashi, A.; Yanagiya, A.; Yamaguchi, T.; Abe, T.; Kureha, T.; Kuba, K.; Kanegae, Y.; Furuta, Y.; Yamamoto, T.; et al. Essential functions of the CNOT7/8 catalytic subunits of the CCR4-NOT complex in mRNA regulation and cell viability. *RNA Biol.* **2020**, 17, 403–416. [CrossRef]
80. Boland, A.; Chen, Y.; Raisch, T.; Jonas, S.; Ozturk, D.K.; Wohlbold, L.; Weichenrieder, O.; Izaurralde, E. Structure and assembly of the NOT module of the human CCR4–NOT complex. *Nat. Struct. Mol. Biol.* **2013**, 20, 1289–1297. [CrossRef]
81. Basquin, J.; Roudko, V.V.; Rode, M.; Basquin, C.; Séraphin, B.; Conti, E. Architecture of the Nuclease Module of the Yeast Ccr4-Not Complex: The Not1-Caf1-Ccr4 Interaction. *Mol. Cell* **2012**, 48, 207–218. [CrossRef] [PubMed]
82. Keskeny, C.; Raisch, T.; Sgromo, A.; Igreja, C.; Bhandari, D.; Weichenrieder, O.; Izaurralde, E. A conserved CAF40-binding motif in metazoan NOT4 mediates association with the CCR4–NOT complex. *Genes Dev.* **2019**, 33, 236–252. [CrossRef] [PubMed]
83. Pavanello, L.; Hall, B.; Airhihen, B.; Winkler, G.S. The central region of CNOT1 and CNOT9 stimulates deadenylation by the Ccr4–Not nuclease module. *Biochem. J.* **2018**, 475, 3437–3450. [CrossRef] [PubMed]
84. Mauxion, F.; Preve, B.; Seraphin, B. C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4-NOT complex. *RNA Biol.* **2013**, 10, 267–276. [CrossRef] [PubMed]
85. Lau, N.-C.; Kolkman, A.; Van Schaik, F.M.A.; Mulder, K.W.; Pijnappel, W.W.M.P.; Heck, A.; Timmers, H.T.M. Human Ccr4–Not complexes contain variable deadenylase subunits. *Biochem. J.* **2009**, 422, 443–453. [CrossRef]
86. Raisch, T.; Chang, C.-T.; Levdansky, Y.; Muthukumar, S.; Raunser, S.; Valkov, E. Reconstitution of recombinant human CCR4-NOT reveals molecular insights into regulated deadenylation. *Nat. Commun.* **2019**, 10, 1–14. [CrossRef]
87. Chen, Y.; Khazina, E.; Izaurralde, E.; Weichenrieder, O. Crystal structure and functional properties of the human CCR4-CAF1 deadenylase complex. *Nucleic Acids Res.* **2021**, 49, 6489–6510. [CrossRef]
88. Zhang, Q.; Pavanello, L.; Potapov, A.; Bartlam, M.; Winkler, G.S. Structure of the human Ccr4-Not nuclease module using X-ray crystallography and electron paramagnetic resonance spectroscopy distance measurements. *Protein Sci.* **2021**, 31, 758–764. [CrossRef]
89. Nicholson, A.L.; Pasquinelli, A.E. Tales of Detailed Poly(A) Tails. *Trends Cell Biol.* **2018**, 29, 191–200. [CrossRef]
90. Eisen, T.J.; Eichhorn, S.W.; Subtelny, A.O.; Lin, K.S.; McGeary, S.E.; Gupta, S.; Bartel, D.P. The Dynamics of Cytoplasmic mRNA Metabolism. *Mol. Cell* **2020**, 77, 786–799.e10. [CrossRef]
91. Charlesworth, A.; Meijer, H.A.; De Moor, C.H. Specificity factors in cytoplasmic polyadenylation. *Wiley Interdiscip. Rev. RNA* **2013**, 4, 437–461. [CrossRef] [PubMed]
92. Conrad, N.K. The emerging role of triple helices in RNA biology. *Wiley Interdiscip. Rev. RNA* **2013**, 5, 15–29. [CrossRef] [PubMed]
93. Torabi, S.-F.; Chen, Y.-L.; Zhang, K.; Wang, J.; DeGregorio, S.J.; Vaidya, A.T.; Su, Z.; Pabit, S.A.; Chiu, W.; Pollack, L.; et al. Structural analyses of an RNA stability element interacting with poly(A). *Proc. Natl. Acad. Sci. USA* **2021**, 118, e2026656118. [CrossRef] [PubMed]
94. Torabi, S.-F.; Vaidya, A.T.; Tycowski, K.T.; DeGregorio, S.J.; Wang, J.; Shu, M.-D.; Steitz, T.A.; Steitz, J.A. RNA stabilization by a poly(A) tail 3′-end binding pocket and other modes of poly(A)-RNA interaction. *Science* **2021**, 371, eabe6523. [CrossRef]