

Mechanistic insights into RNA surveillance by the canonical poly(A) polymerase Pla1 of the MTREC complex

Received: 30 June 2022

Accepted: 31 January 2023

Published online: 11 February 2023

 Check for updates

Komal Soni ^{1,4}, Anusree Sivadas^{2,4}, Attila Horvath ², Nikolay Dobrev ¹, Rippei Hayashi ², Leo Kiss ¹, Bernd Simon ³, Klemens Wild ¹, Irmgard Sinning ¹  & Tamás Fischer ² 

The *S. pombe* orthologue of the human PAXT connection, Mtl1-Red1 Core (MTREC), is an eleven-subunit complex that targets cryptic unstable transcripts (CUTs) to the nuclear RNA exosome for degradation. It encompasses the canonical poly(A) polymerase Pla1, responsible for polyadenylation of nascent RNA transcripts as part of the cleavage and polyadenylation factor (CPF/CPSF). In this study we identify and characterise the interaction between Pla1 and the MTREC complex core component Red1 and analyse the functional relevance of this interaction in vivo. Our crystal structure of the Pla1-Red1 complex shows that a 58-residue fragment in Red1 binds to the RNA recognition motif domain of Pla1 and tethers it to the MTREC complex. Structure-based Pla1-Red1 interaction mutations show that Pla1, as part of MTREC complex, hyper-adenylates CUTs for their efficient degradation. Interestingly, the Red1-Pla1 interaction is also required for the efficient assembly of the fission yeast facultative heterochromatic islands. Together, our data suggest a complex interplay between the RNA surveillance and 3'-end processing machineries.

Nuclear RNA surveillance is carried out primarily by the exosome which is involved in 3'-5' RNA degradation and maturation processes¹⁻³. While export-competent mature RNAs are packaged into mRNP particles, transcripts that fail to mature properly are rapidly degraded by the nuclear RNA surveillance machinery. In addition, pervasive transcription, occurring widely in eukaryotes largely due to bidirectional promoters, produces tremendous amounts of non-coding RNA transcripts that are degraded by the RNA surveillance machinery⁴⁻⁷. These pervasive transcripts, also known as cryptic unstable transcripts (CUTs) in yeast, or promoter upstream transcripts (PROMPTs) and upstream antisense RNAs (uaRNAs) in humans, are barely detectable under steady-state conditions due to their rapid degradation^{8,9}. Specific recognition of the wide variety of defective RNAs degraded by the exosome is done by its so-called adaptor

complexes¹⁰. The nuclear exosome relies on the helicase activity of the DExH-box RNA helicase Mtr4 to thread single-stranded RNAs into its core, thereby making it a bona fide member of all adaptor complexes¹¹⁻¹⁶. The best characterised adaptor complex is the TRAMP complex¹⁷, which was identified as the major exosome targeting factor for CUTs in *Saccharomyces cerevisiae*¹⁸.

Polyadenylation is a post-transcriptional modification of RNA involving the addition of poly(A) tails to the 3'-ends of RNA substrates^{19,20}. Following the discovery of the yeast TRAMP complex, the connection between polyadenylation-mediated destabilisation of RNAs as a signal for exosome degradation in eukaryotes was established^{21,22}. In humans, two recently identified adaptors known as the nuclear exosome targeting (NEXT)²³ complex and the poly(A) tail exosome targeting (PAXT) connection²⁴ are responsible for targeting

¹Heidelberg University Biochemistry Center (BZH), INF 328, D-69120 Heidelberg, Germany. ²The John Curtin School of Medical Research, The Australian National University, Canberra ACT 2601, Australia. ³European Molecular Biology Laboratory (EMBL), Meyerhofstr. 1, D-69117 Heidelberg, Germany. ⁴These authors contributed equally: Komal Soni, Anusree Sivadas. ✉ e-mail: irmi.sinning@bzh.uni-heidelberg.de; tamas.fischer@anu.edu.au

early, unprocessed RNAs, PROMPTs and enhancer RNAs or transcripts with poly(A) tails, respectively.

In *Schizosaccharomyces pombe*, a highly conserved MTREC (MtlI-Red1 core)^{25,26} or NURS (nuclear RNA silencing) complex²⁷ is responsible for degradation of CUTs and meiotic mRNAs. Notably, the MTREC complex was identified in studies of the Mtr4-like protein 1 (MtlI) and the zinc-finger protein Red1, which in turn contributed to the definition of the PAXT connection in humans²⁸. MTREC is a multi-subunit complex comprising the CAP binding complex Cbc1-Cbc2-Ars2 (CBCA), Iss10-Mmi1, Red5-Pab2-Rmn1 and the canonical poly(A) polymerase Pla1, in addition to the core MtlI and Red1 proteins²⁶. We have recently described Red1 as the central scaffolding protein of MTREC. Red1 directly interacts with each submodule of the MTREC complex, including Pla1, and connects them to the MtlI helicase to deliver the associated RNA cargo to the exosome (ref.^{29,30}, Supplementary Fig. 1A). Importantly, Pla1 is usually part of the highly conserved cleavage and polyadenylation factor (CPF or CPSF in mammals), where it functions to add poly(A) tails during the 3'-end processing of transcripts^{31–33}. Co-purification of Pla1 with the MTREC complex and the extended poly(A) tails in meiotic mRNAs and CUTs that are necessary for their efficient degradation^{34–36} therefore appear to be correlated. Consistently, Mmi1, along with Pla1 and poly(A) binding protein Pab2 have been implicated in promoting hyperadenylation of CUTs and meiotic RNAs in fission yeast^{35,37}. In addition, Pla1 co-localised and immunoprecipitated with Red1 in an RNA-independent fashion, suggesting a physical interaction between them³⁶. Furthermore, it was shown that Red1 and Mmi1 facilitate the hyper-polyadenylation of DSR-containing mRNAs by Pla1³⁶ and that this hyper-adenylation is required for their efficient degradation.

Studies of Pla1 homologues from mammals (*h*sPAP) and *S. cerevisiae* (*sc*Pap1) show that the polymerase adopts a tri-partite domain architecture^{38,39}. It consists of three globular domains, an N-terminal nucleotide binding domain (NTD), a middle domain (MD) and a C-terminal RNA Recognition Motif (RRM)⁴⁰, where the NTD and MD form the catalytic centre (Fig. 1A). Detailed kinetic and thermodynamic analyses of *sc*Pap1 suggest that large-scale domain movements in the protein are required for substrate recognition and catalysis^{41–43} and the enzyme is stabilised in a closed conformation with extensive contacts between the NTD and RRM domain upon substrate binding⁴¹. Sequence alignment of Pla1 from different species shows that the NTD and MD are highly conserved between yeast and mammals (66.6% similarity), while the RRM domain is quite diverse (19.5% similarity; Supplementary Fig. 1B). Interestingly, the C-terminal RRM domain lacks the consensus RNA binding residues (reviewed in ref.⁴⁴) and does not interact with the RNA substrate via its canonical RNA binding β -sheet interface, but instead binds via its opposite interface⁴¹. In *S. cerevisiae*, the RRM domain of *sc*Pap1 is also responsible for binding to the cleavage and polyadenylation subunit *sc*Fip1 (homologous to Iss1 in *S. pombe*)^{45,46}, tethering it to the CPF. Binding of *sc*Fip1 leads to *sc*Pap1 inhibition in yeast, whereas activation has been observed in mammals and plants^{45,47,48}.

While Pla1-dependent excessive polyadenylation of meiotic mRNAs that are destined for degradation³⁴ and hyperadenylation of CUTs has been reported in *S. pombe*²⁶ (compared to *S. cerevisiae* where MTREC is absent), the functional importance of Pla1 in MTREC has not been studied. In this study, we have used a combination of structural, biochemical and in vivo experiments to understand the specific role of Pla1 as an associated factor of the MTREC complex. Together with nuclear magnetic resonance spectroscopy (NMR) and isothermal titration calorimetry (ITC), we show that the C-terminal RRM domain of Pla1 specifically recognises a 58-residue fragment of Red1. We report crystal structures of Pla1 in its apo form and in complex with Red1. Specific point mutations disrupting the Pla1-Red1 interaction and deletion of the complete Pla1 interacting region in Red1 lead to the widespread accumulation of PROMPTs which also

harbour shorter poly(A) tails compared to wild-type cells. Interestingly, the Pla1 “truncated” MTREC complex is also defective in the assembly of facultative heterochromatic islands at meiotic genes. Taken together, our data reveal the molecular details of recruitment of Pla1 by Red1 which leads to hyperadenylation of CUTs, serving as a signal for exosome-mediated target degradation. In addition, in vitro competition experiments between Red1 and Iss1, and additional interactions between MTREC and the CPF-associated factor Msi2 (homologous to CF IB or *sc*Hrp1 in *S. cerevisiae*), indicate the existence of an intricate interaction network between CPF and the MTREC complex to control the processing and surveillance of nascent RNA transcripts.

Results

The C-terminal RRM domain of Pla1 mediates its interaction with Red1

We have previously shown that the Red1 fragment comprising residues 240–345 (Red1_{240–345}) mediates interaction of Pla1 with the MTREC complex (ref.²⁹, Supplementary Fig. 1A), but the exact interaction region within Pla1 was not identified. Therefore, we set out to map the minimal interaction region of Pla1 with Red1 using yeast-two hybrid experiments (Y2H). Y2H assays were performed with constructs comprising the NTD and MD of Pla1 (residues 1–352, NTD-MD) or RRM domain of Pla1 (residues 352–566) (Fig. 1B). We found that the Pla1 RRM domain (Pla1_{RRM}) is necessary and sufficient for interaction with Red1, consistent with previous reports that the RRM domain of *sc*Pap1 is responsible for mediating protein-protein interaction between *sc*Pap1 and its co-factor, *sc*Fip1^{45,46}.

To further validate our Y2H results, we decided to reconstitute the Pla1-Red1 complex in vitro. To this end, we used bacterial cell lysates co-expressing His₆-MBP fused to various Red1 fragments together with untagged Pla1_{RRM} and assessed the ability of the Red1 constructs to pull down Pla1_{RRM} (Fig. 1C). Red1 is predicted to be rather unstructured in the originally identified Pla1 interacting fragment comprising residues 240–345 (Red1_{240–345}) (Supplementary Fig. 1C). Therefore, we used partially overlapping, short fragments of Red1, comprising residues 259–288 (Red1_{259–288}), 288–345 (Red1_{288–345}) and 288–322 (Red1_{288–322}) in addition to Red1_{240–345}. While Red1_{240–345}, Red1_{288–345} and Red1_{288–322} co-purified Pla1_{RRM}, Red1_{259–288} did not, suggesting that residues 288–322 of Red1 contain the Pla1 binding site. We further probed the Pla1-Red1 binding interface using NMR. A two dimensional ¹H, ¹⁵N-HSQC spectrum of Red1_{288–322} shows that this region of Red1 is largely unstructured, as inferred from the low chemical shift dispersion in the ¹H-dimension (Fig. 1D) and analysis of the ¹³C α and ¹³C β secondary chemical shifts (Supplementary Fig. 2A, B). Subsequently, we performed NMR titrations of unlabelled Pla1_{RRM} into ¹⁵N-labelled Red1_{288–322} and monitored the chemical shifts using two dimensional ¹H, ¹⁵N-HSQC spectra (Fig. 1D). We observed that a majority of the backbone amide resonances of the Red1 peptide show severe line broadening while a handful show chemical shift perturbations (CSPs), both indicative of a clear binding event (Supplementary Fig. S2C, D). To confirm whether Red1_{288–322} contains the complete Pla1 binding site, we performed NMR titrations of unlabelled Pla1_{RRM} into ¹⁵N-labelled Red1_{288–345} (Supplementary Fig. 2E). Surprisingly, we observed additional residues in Red1_{288–345} which show CSPs compared to that of Red1_{288–322}. To understand the affinity contribution of these additional residues in Red1_{288–345} and define the exact region of Red1 necessary and sufficient for binding to Pla1, we performed isothermal titration calorimetry. We found that while Red1_{288–322} binds to Pla1 with a $K_D = 7.9 \mu\text{M}$ (Fig. 1E and Supplementary Table 1), the addition of residues 323–345 in Red1_{288–345} leads to a -5.6-fold increase in binding affinity ($K_D = 1.4 \mu\text{M}$, Fig. 1F and Supplementary Table 1). Therefore, from the fragments tested in this study, the fragment comprising residues 288–345 binds best to Pla1, and is referred to as the Pla1 interaction region.

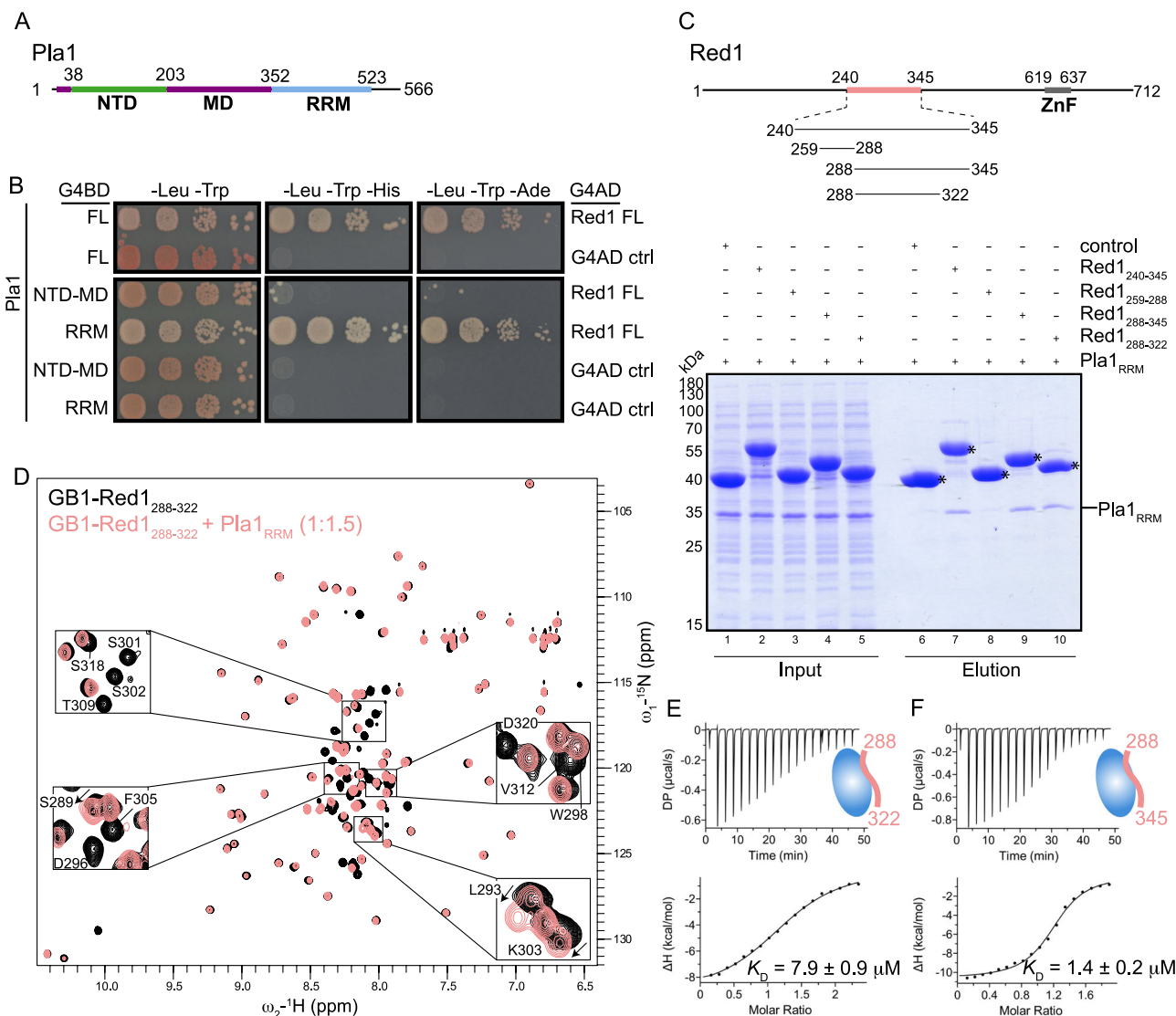


Fig. 1 | Interaction mapping between Pla1 and Red1. **A** Domain organisation of Pla1, consisting of the NTD, MD and RRM domains shown in green, purple and blue, respectively. **B** Y2H experiments show that the Pla1 RRM domain is responsible for the interaction with Red1. Full-length (FL) Pla1, Pla1_{NTD-MD} or Pla1_{RRM} constructs were fused to the Gal4 DNA binding domain (G4BD) while full-length Red1 was fused to the Gal4 activation domain (G4AD). Auto-activation controls (ctrl) for FL-Pla1, Pla1_{NTD-MD} and Pla1_{RRM} are provided. Serial dilutions of equivalent amounts of yeast were plated on double (-Leu-Trp) and triple dropout media (-Leu-Trp-His, -Leu-Trp-Ade), with growth on triple dropout media indicating an interaction between the tested proteins. **C** In vitro pull-down assays. Untagged Pla1_{RRM} was co-

expressed with His₆-MBP fusion constructs of Red1 or His₆-MBP (control). 0.01% of input (lanes 1–5) and 25% of elution fractions (lanes 6–10) were separated on 12% SDS-PAGE gel. Asterisks mark the different His₆-MBP fusion constructs. A representative gel from two independent runs is shown. Source data are provided as a Source Data file. **D** Overlay of ^1H , ^{15}N -HSQC NMR spectra of GB1-tagged Red1₂₈₈₋₃₂₂ in the absence and presence of Pla1_{RRM} are shown in black and salmon, respectively. Zoom-in views of Red1 residues showing chemical shift perturbations or line broadening are shown. **E**, **F** ITC experiments with a serial titration of Red1₂₈₈₋₃₂₂ (**E**) or Red1₂₈₈₋₃₄₅ (**F**) into Pla1_{RRM}. The calculated dissociation constants (K_D) from an average of two or three independent measurements are shown.

Crystal structures of the canonical poly(A) polymerase Pla1 in its apo form and in complex with Red1

To first gain insights into the inter-domain arrangement and interactions between the NTD, MD and RRM domains of Pla1, we determined the crystal structure of the protein in its apo form. We expressed, purified and performed crystallisation trials with the full-length (Pla1_{FL}) and a C-terminal truncation (Pla1_{Δ14}) of Pla1. Both constructs readily crystallised under a variety of conditions within 2–4 days. Crystals of Pla1_{FL} diffracted to substantially higher 1.9 Å resolution while Pla1_{Δ14} diffracted to 2.6 Å resolution (Table 1).

The crystal structure of the full-length protein shows a tripartite domain architecture of Pla1 (Fig. 2A). The overall topology of the three domains is largely similar to that of the Pla1 homologues (Supplementary Fig. 3A). Briefly, the N-terminal catalytic domain is

homologous to the catalytic domain of other nucleotidyl-transferases forming a five stranded mixed β -sheet along with four α -helices^{49,50}. The middle domain is formed by a four-helix bundle which is capped by the N-terminus of Pla1. The C-terminal RRM domain contains additional secondary structure elements extending the canonical $\beta\alpha\beta\alpha\beta$ topology of RRM⁴⁴. It comprises additional helices α 14, α 15 and β -strands β 8, β 9 in the extended loop L1 connecting the canonical RRM strands β 7 and β 10 compared to its homologues (Supplementary Fig. 3B). In our crystal structure, 18 residues of loop L1, loop L2 connecting β 10 and α 16 and 25 residues at the C-terminus are disordered. Interestingly, residues 415–434, comprising helix α 14 and β -strand β 9, partially block the canonical RNA binding interface of the RRM domain on one side⁴⁰, while the other side is blocked by residues 523–532 belonging to the C-terminus of the protein. The helix α 15 of the RRM

Table 1 | X-ray crystallography data collection and refinement statistics

	Pla _{FL}	Pla _{Δ14}	Pla _{Red1}
Data collection			
Beamline	PETRA III, P13 DESY	PETRA III, P13 DESY	PETRA III, P14 DESY
Wavelength (Å)	0.9762	0.9762	0.9801
Resolution range (Å)	77.68–1.9 (1.968–1.9)	57.35–2.599 (2.692–2.599)	98.02–2.805 (2.905–2.805) ^a
Space group	C121	P1	P2 ₁ 2 ₁ 2
Unit cell dimensions			
a, b, c (Å)	122.06 66.75 88.12	71.79 72.48 73.21	128.64 151.34 65.41
α, β, γ (°)	90 118.18 90	99.96 105.70 118.83	90 90 90
Total reflections	333,368 (32,959)	119,702 (11,522)	141,478 (12,446) 90,056 (3985) ^a
Unique reflections	49,226 (4854)	34,690 (3409)	31,789 (654) 20,986 (1049) ^a
Multiplicity	6.8 (6.8)	3.5 (3.4)	4.3 (3.8) ^a
Completeness (%)	99.73 (99.02)	97.04 (95.57)	63.95 (20.67) 92.4 (96.7) ^a
Mean I/sigma(I)	11.87 (1.53)	7.26 (0.89)	9.91 (0.64) 12.8 (2.6) ^a
R _{merge}	0.1145 (1.18)	0.1545 (1.402)	0.065 (0.479) ^a
R _{meas}	0.1242 (1.279)	0.1837 (1.67)	0.074 (0.555) ^a
R _{pim}	0.04764 (0.4875)	0.09824 (0.8945)	0.036 (0.275) ^a
CC _{1/2}	0.998 (0.663)	0.99 (0.298)	0.999(0.868) ^a
Refinement			
Reflections used in refinement	49,223 (4854)	34,678 (3409)	20,527 (654)
Reflections used for R _{free}	2462 (243)	1733 (171)	1999 (63)
R _{work}	0.1690 (0.3095)	0.2161 (0.3710)	0.2413 (0.3401)
R _{free}	0.1923 (0.3453)	0.2575 (0.4146)	0.2840 (0.4666)
Number of non-hydrogen atoms	4718	8199	8531
Macromolecules	4185	8113	8528
Ligands	6	–	–
Solvent	527	86	3
R.M.S deviations			
Bond lengths (Å)	0.003	0.002	0.002
Bond angles (°)	0.56	0.51	0.49
Ramachandran plot Most favoured (%)	97.44	97.60	93.71
Allowed (%)	2.37	2.30	5.71
Outliers (%)	0.20	0.10	0.57
Rotamer outliers (%)	0.65	1.25	1.71
Clashscore	6.52	3.80	4.44
Average B-factor	32.65	64.81	47.93
Macromolecules	31.75	64.91	47.93
Ligands	36.74	–	–
Solvent	39.69	55.47	41.78

Statistics for the highest-resolution shell are shown in parentheses.

^aStatistics reported for ellipsoidal diffraction cut-off.

domain is positioned away from the β-sheet interface facing the back side of the RRM, to form an interface with the Pla_{NTD} (Fig. 2A and Supplementary Fig. 4A).

A large interface is formed by a network of interactions between the three domains of the protein burying a total surface area of ~2982 Å². The tripartite domain architecture of Pla1 is held together by

hydrogen bonds between the interface residues of the domains (Supplementary Fig. 4A). A comparison of our crystal structures of Pla_{FL} and Pla_{Δ14} shows that the individual domains have undergone significant movements. An alignment based on the middle domains illustrates that the N-terminal has a modest rotation of 2.95° with a 1.13 Å displacement, while the C-terminal RRM domain has a more pronounced rotation of 9.8° and 2.26 Å displacement (Supplementary Fig. 4B). Pla1 is also known to undergo large-scale domain motions around the defined hinges connecting the NTD and MD, and the MD and RRM domains^{41,43}, which are deemed necessary for proper functioning of the enzyme.

To understand the molecular details of the Red1-Pla1 interaction, we determined the crystal structure of the complex (Table 1). Since Red1 binds to Pla_{RRM} with low micromolar affinity, we reasoned that the complex should be stably purified using size exclusion chromatography. Indeed, we could reconstitute a highly pure complex which was subjected to co-crystallisation trials (Supplementary Fig. 5A). Co-crystals of the complex appeared within three weeks and diffracted to 2.81 Å resolution. The asymmetric unit contains two molecules of the Pla1-Red1 heterodimer which exhibit small differences between them, as indicated by a root mean square deviation (RMSD) of 0.96 Å. Although the Pla1-Red1 interaction interface is conserved between the two molecules, the overall density for the second heterodimer is much weaker and therefore our structural analyses are based on the first molecule. As in the apo structure, part of loop L1, the complete loop L2 and the C-termini of Pla_{RRM} are disordered. RRM domain α15, which forms an interface with the NTD in our apo structure, is also partly disordered. In addition to these, the Red1 C-terminus, spanning residues 323–345, is flexible and therefore not observed in the crystal structure of the complex. A superposition of the apo Pla_{FL} and the Pla1-Red1 complex shows that the Pla_{NTD} has a rotation of 4.98° with a 1.01 Å displacement, while the Pla_{RRM} domain has an angular rotation of 8.06° with a 1.63 Å displacement, again indicative of the flexibility between the domains (Supplementary Fig. 5B).

Our crystal structure of the Pla1-Red1 complex shows that Red1 contacts the C-terminal RRM domain of Pla1 (Fig. 2B). The interaction between Pla1 and Red1 buries a total solvent accessible area of 1338.5 Å². The interaction interface of Red1 can be largely divided into two regions. The first half comprises a short α-helix at the N-terminus (α0) that helps orient Red1 onto Pla1 (Fig. 2C). A network of aromatic stacking interactions occurs between Pla_{RRM} Phe468 and Red1 Trp298 and Phe305 (Fig. 2C and Supplementary Fig. 5C). This cluster is further stabilised by hydrophobic stacking of the aliphatic sidechain of Lys368, a hydrogen bond between its side chain and backbone carbonyl of Trp298, a salt bridge formed between Glu416 and Lys294 and packing of Asp513 against Phe305. Importantly, Lys368 positioned close to the C-terminus of Red1 helix α0 reads out the negative dipole of α0 and therefore further stabilises it. The second half of Red1 is largely involved in mainchain-sidechain interactions between Red1 residues Gly306, Ser308, Asn311 and Val312 that pack against Pla1 residues Gln420, Val518, Ile520 and Asn522 (Fig. 2D). In addition, hydrogen bonding interactions occur between Red1 Ser307 and Ser308 side chains with Pla1 Ser424 and Lys506. Interestingly, Red1 Val312-Ile314 form a β-strand (β0) that packs against Pla1 Arg491-Asp493 that also form a short β-strand, leading to the formation of an antiparallel β-zipper⁵¹ (Fig. 2D). Importantly, while our NMR results show that Red1 residues 312–314 have a propensity to form β0 in apo (Supplementary Fig. 2A, B), the Pla1 β-strand is not formed in our Pla1 apo crystal structures. Therefore, this β-strand addition by means of a β-zipper formation stabilises and holds the second half of Red1 in position.

To provide support for the structure of Pla1-Red1 complex and validate its conformation in solution, we recorded SAXS experiments (Supplementary Fig. 5D, Supplementary Table 2). The Kratky plot of the Pla1-Red1 complex shows a bell-shaped curve characteristic of well-

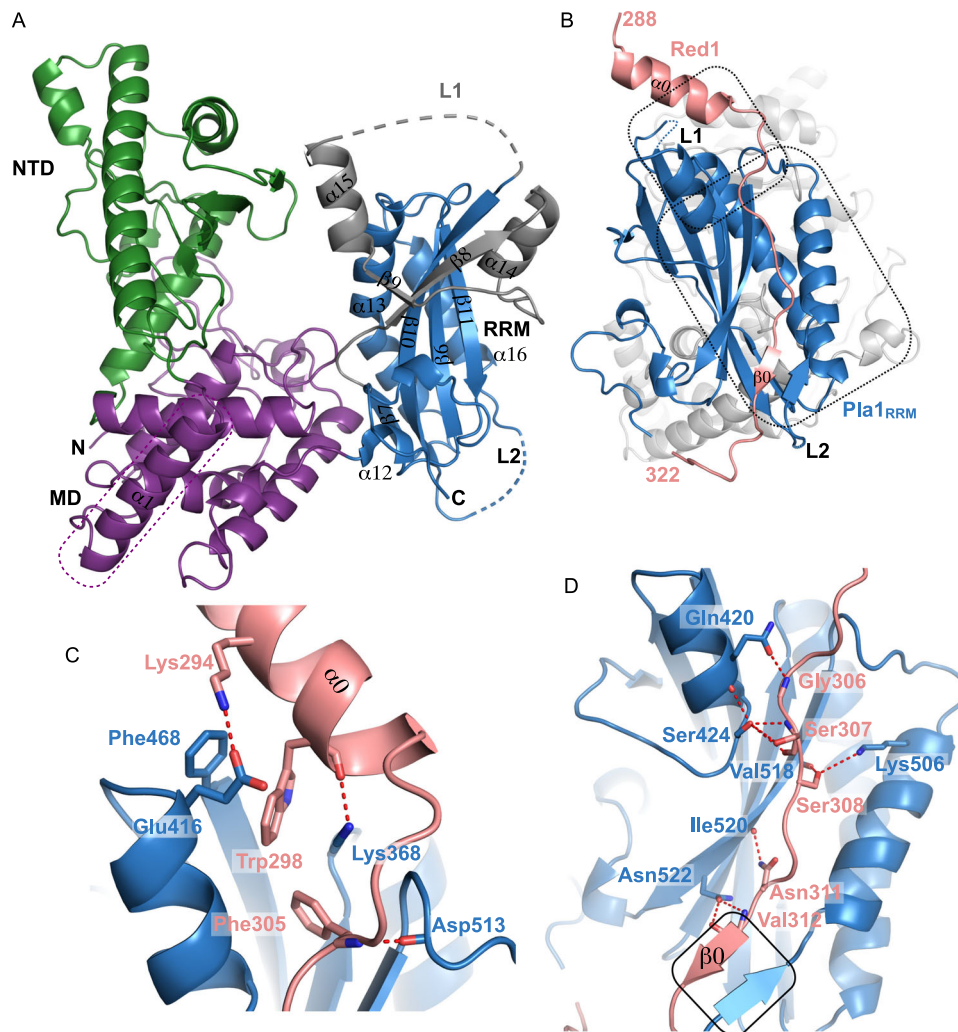


Fig. 2 | Crystal structures of Pla1 in its apo form and in complex with Red1.

A The overall architecture of Pla1 in its apo form is shown. Colour scheme for Pla1 domains are the same as Fig. 1A. The N- and C-termini of the protein and loops L1 (shown in grey) and L2 of RRM domain are marked. The Pla1 N-terminal $\alpha 1$, which forms a part of the MD, is boxed in purple. The secondary structure elements of the RRM domain comprising helices $\alpha 12$ - $\alpha 16$ and β -strands $\beta 6$ - $\beta 11$ are marked. **B** Crystal structure of Pla1-Red1 complex. Red1 and Pla1_{RRM} are shown in salmon and blue,

respectively, while the NTD and MD are coloured in grey. The Red1 N- and C-termini, its secondary structure elements $\alpha 0$ and $\beta 0$, along with the Pla1_{RRM} domain loops L1 and L2 are marked. The two Red1 interaction interfaces are boxed. **C, D** Zoom-in views of residues involved in the Pla1-Red1 interaction. Polar contacts are shown by red dotted lines. The β -zipper formed between Red1 $\beta 0$ and Pla1 residues Arg491-Asp493 is boxed in black.

folded globular molecules (Supplementary Fig. 5E). The pairwise distribution curve also shows that the Pla1-Red1 complex has a globular architecture with a D_{\max} of 10.7 nm (Supplementary Fig. 5F). Since the Pla1_{RRM} loops L1 and L2, and both Pla1 and Red1 C-termini are missing from the crystal structure, we used CORAL⁵² to first model these regions based on the SAXS data (Supplementary Fig. 5G). Subsequent fitting of the experimental SAXS scattering profile of the complex with the crystal structure show that the data are in good agreement with a χ^2 fitting-value of 1.23 (Supplementary Fig. 5H).

Structure-based mutation analysis in vitro

Based on the NMR data and our crystal structure of the Pla1-Red1 complex, we performed mutational analyses and evaluated the importance of the affected contacts in ITC measurements using constructs comprising only the Pla1_{RRM} and Red1 residues 288-345 (Supplementary Table 1). We first mutated the aromatic residues in the Red1 N-terminal helix $\alpha 0$ by replacing them with alanine residues to create the double mutant W298/F305A. This double mutant led to a complete disruption of Pla1-Red1 interaction (Fig. 3A). This is not surprising as stacking interactions of Red1 helix $\alpha 0$ residues Trp298

and Phe305 with Pla1_{RRM} Phe468 and Lys368 creates a hydrophobic core which helps orient the Red1 peptide onto Pla1 (Fig. 2C). A charge reversal mutation of Lys368 (K368E) which reads the negative dipole moment of Red1 helix $\alpha 0$ also leads to a -13-fold loss in affinity (Fig. 3B). Furthermore, mutation of Red1 Val313, which is part of $\beta 0$ -strand, to an arginine residue also leads to -13-fold loss in affinity (Fig. 3C), signifying the affinity contribution of the β -zipper. In contrast, alanine replacements of Ser308 and Asn311 did not affect the interaction (Supplementary Fig. 6A, B).

Our NMR titrations had indicated that the negatively charged patch at the Red1₂₈₈₋₃₂₂ C-terminus (namely residues Asp317, Ser318, Asp319 and Asp320) also show line broadening upon addition of Pla1_{RRM} (Fig. 1D and Supplementary Fig. 2D). In our crystal structure, however, these residues do not show contacts with Pla1. To probe the importance of these residues for Pla1-Red1 interaction, we made charge reversal mutations of the aspartate residues to arginine (D317/S318/D319/D320 to RSRR, named Red1_{RSRR}). To our surprise this triple mutant also completely abrogated Pla1-Red1 binding (Fig. 3D). Since 25 residues at the Pla1 C-terminus are disordered in our crystal structure and the Pla1 and Red1 C-termini are in close proximity to each other

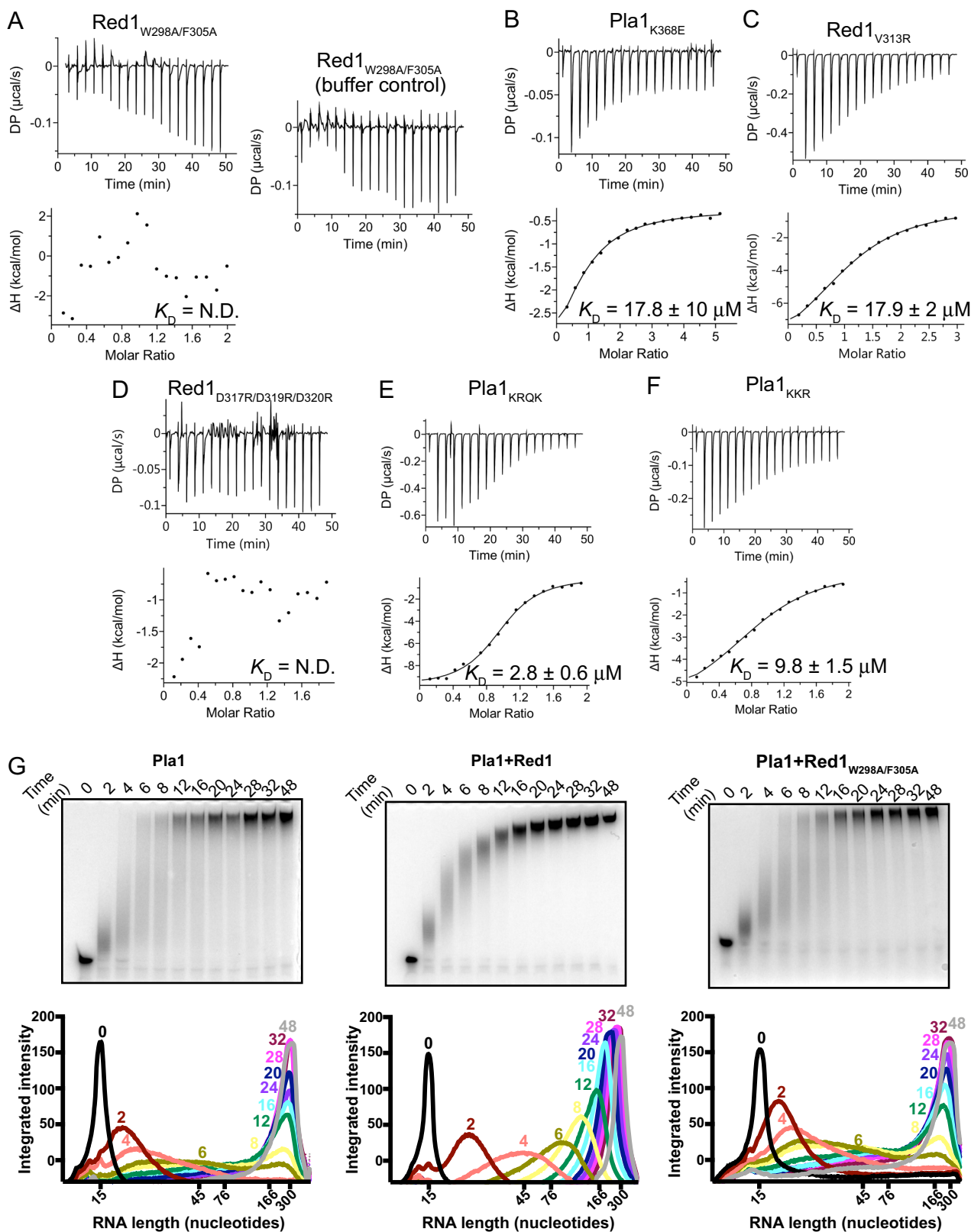


Fig. 3 | Mutational analyses of Pla1-Red1 interaction. Isothermal titration calorimetry experiments with specific Pla1 and Red1 point mutants that affect their interaction interface are shown in panels A–F. The calculated dissociation constant (K_D) from an average of two independent measurements is shown. **G** In vitro polyadenylation assay. Polyadenylation of a 5'-Cy3 labelled A₁₅ RNA primer by Pla1

alone (left panel), in the presence of Red1₂₈₈₋₃₄₅ (middle panel) and Red1_{W298A/F305A} (right panel), analysed by 14% denaturing urea PAGE at different time points. Densitometric analyses of the gels are plotted below, where the RNA length is marked based on an RNA ladder. Representative gels from two independent experiments are shown. Source data are provided as a Source Data file.

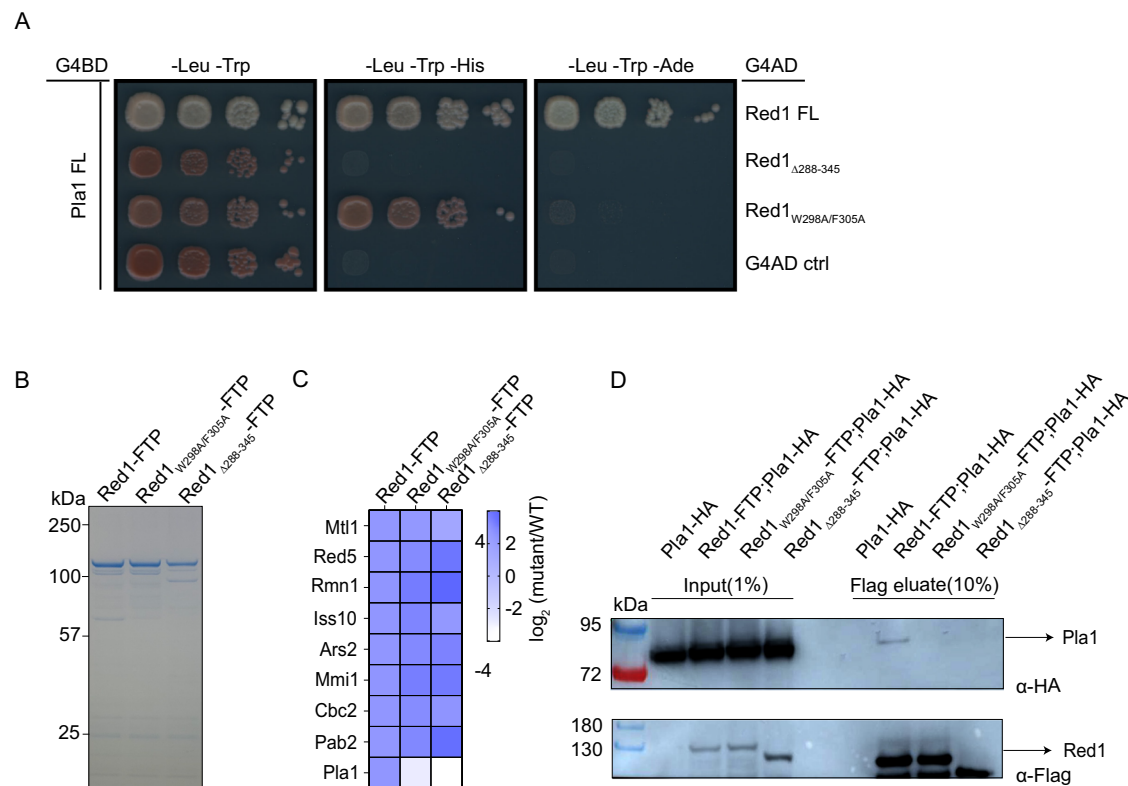


Fig. 4 | Disrupting Red1-Pla1 interaction in vivo leads to 'Pla1-truncated' MTREC complex. **A** Y2H experiments showing the effects of deletion of Red1 residues 288-345 or Red1 $_{W298A/F305A}$ double mutant on Pla1-Red1 interaction in the context of full-length proteins. **B** Coomassie blue stained SDS polyacrylamide gel of tandem affinity purified MTREC complex using Red1 as bait, from respective mutant strains (a representative gel from two independent experiments is shown); and **C** heat map representation of mass-spectrometry (MS) analysis of these purifications. Heat map shows the changes in the amount of co-purifying MTREC subunits (except the Cbc1 subunit which was not conclusively quantified), in indicated mutant strains compared to WT. Co-purifying protein amounts were normalised to the corresponding purified bait protein (WT or mutant Red1) amount. **D** Co-

immunoprecipitation (co-IP) of HA-tagged Pla1 (Pla1-HA) using Red1 as bait in WT and Red1 mutant strains. Inputs and final eluates of the tandem affinity purifications of the indicated strains were blotted to nitrocellulose membrane and cut around the 100 kDa marker. The bottom part of the membrane was probed with α -HA to detect co-purifying Pla1-HA (top panel) and the top part was probed with α -Flag to detect the bait protein Red1-FTP (bottom panel). Note that the Red1 bait protein in the Flag eluate runs ~20 kDa lower than in the input, due to the cleavage of the C-terminal Protein-A tag by the TEV enzyme during the elution step from the IgG column (see details in Methods section). A representative membrane from two independent experiments is shown. Source data are provided in Source Data file.

(Supplementary Fig. 6C), we reasoned that two positively charged patches present in the disordered Pla1 C-terminus could possibly be involved in this interaction. Therefore, we created two Pla1 $_{RRM}$ charge reversal mutants: K544/K545/R546 to EEE and K559/R560/Q561/K562 to EEQE, named Pla1 $_{EEE}$ and Pla1 $_{EEQE}$, respectively. While the Pla1 $_{EEQE}$ mutant led to a -2-fold loss in affinity (Fig. 3E), the Pla1 $_{EEE}$ mutant led to a -7-fold loss in affinity (Fig. 3F), indicating that these charged patches of Pla1 and Red1 C-termini are possibly in close proximity in solution and might be involved in electrostatic interactions. In accordance with these data, structure prediction using ColabFold⁵³ also shows that Pla1 residues K544/K545/R546 are spatially close to Red1 residues D317/S318/D319/D320 (Supplementary Fig. 6D).

To assess if recombinantly purified Pla1 is active in vitro and if its polyadenylation activity is affected by Red1, we used in vitro polyadenylation assays. Recombinantly purified Pla1 was incubated with 5'-Cy3 labelled A₁₅ RNA primer in the absence or presence of Red1 and the reaction was started by addition of ATP. As previously reported, Pla1 is active in in vitro polyadenylation assays⁵⁴ (Fig. 3G) compared to a control where the catalytic mutant of Pla1 (D153A) did not show any polyadenylation activity (Supplementary Fig. 6E). Remarkably, Red1 affects the processivity of Pla1 and the poly(A) tail synthesis becomes more distributive, as evidenced by the gradual increase in length of poly(A) of all products compared to a significant difference in the tail lengths of the different RNA species in the absence of Red1. This effect of Red1 on the Pla1 catalytic activity is similar to that observed for the *S.*

cerevisiae homologues scPap1 and scFip1^{45,55}. Of note, the distribution of poly(A) tail lengths of products was similar to wild-type Pla1 when the Pla1-Red1 interaction mutant (Red1 $_{W298A/F305A}$) was used (Fig. 3G), indicating that Red1 $_{W298A/F305A}$ is unable to bind and therefore influence the processivity of Pla1.

Disrupting Red1-Pla1 interaction in vivo leads to 'Pla1-truncated' MTREC complex

To assess the effects of Pla1-Red1 interaction mutants, we first performed Y2H experiments using the full-length Red1 containing the W298A/F305A double mutation, together with full-length Pla1 (Fig. 4A). Indeed, these point mutations strongly impaired the Red1-Pla1 interaction, as evidenced by the lack of colonies on -Leu -Trp -Ade media, although yeast colonies on the more sensitive -Leu -Trp -His media indicated a residual binding. However, this interaction was completely abrogated when the Red1 residues 288-345 were deleted (Red1 $_{\Delta 288-345}$; Fig. 4A).

To confirm that the MTREC complex lost its association with its Pla1 subunit in the *red1* $_{W298A/F305A}$ and *red1* $_{\Delta 288-345}$ strains in vivo, we replaced the genomic copy of Red1 in *S. pombe* cells with a C-terminally Flag-ProtA-tagged version of wild-type (WT) Red1, *red1* $_{W298A/F305A}$ double mutant and *red1* $_{\Delta 288-345}$ deletion mutant. We performed tandem affinity purifications with these strains (Fig. 4B) and analysed the final eluates using mass spectrometry with TMT 10plex mass tag labelling (Fig. 4C). These analyses confirmed that both Red1 $_{W298A/F305A}$

and Red1 $_{\Delta 288-345}$ proteins maintained their association with MTREC components, purifying similar amounts of MTREC subunits to WT Red1, with the exception of the Pla1 subunit, which was not detectable in the purification of Red1 mutants (Fig. 4C). To further confirm that the Pla1-Red1 interaction was disrupted in the *red1* $_{W298A/F305A}$ and *red1* $_{\Delta 288-345}$ mutants, we tagged the genomic copy of Pla1 with an HA-tag in WT and mutant Red1-FTP strain backgrounds and performed co-immunoprecipitation (co-IP) experiments. Western blot (WB) results showed that Pla1-HA signal could not be detected in the final eluate of our Red1-Pla1 interaction mutant strains (Fig. 4D). Taken together, these data confirm that *red1* $_{W298A/F305A}$ and *red1* $_{\Delta 288-345}$ mutants lead to a 'truncated MTREC', devoid of Pla1, both in vitro and in vivo.

Pla1 activity as part of MTREC complex is required for the efficient degradation of PROMPTs

To understand the role of Pla1 in the recognition and degradation of CUTs, we sequenced poly(A)⁺ RNA from WT, Red1 knock-out (*red1* Δ), *red1* $_{W298A/F305A}$ and *red1* $_{\Delta 288-345}$ strains. Metagene plots of sense and antisense RNA levels 500 bp upstream and downstream of all *S. pombe* genes (Supplementary Fig. 7A, B) and a subset of genes (2400 genes, ~40% of all genes) filtered for detectable levels of PROMPTs (Fig. 5A–C) in WT and red1 mutant cells confirmed the strong accumulation of PROMPTs, antisense (AS) RNAs and 3' intergenic transcripts (3'IGTs) in *red1* Δ cells compared to WT. Interestingly, *red1* $_{W298A/F305A}$ and *red1* $_{\Delta 288-345}$ cells showed only a moderate but highly reproducible accumulation in the levels of PROMPTs, while AS RNAs and 3' intergenic transcript levels were not affected. Meiotic genes and levels of intronic sequences were also unaffected in these mutants, compared to WT (Supplementary Fig. 7C–F). To exclude the possibility that changes in the Red1-Pla1 interaction mutants remained undetected due to the complete lack of poly(A) tails in stabilised transcripts, we repeated these analyses, using total RNA sequencing (Supplementary Fig. 8A, B). While the total RNA sequencing detected less antisense RNA transcripts in general, the difference between WT and Red1-Pla1 interaction mutants was nearly identical to the poly(A)⁺ transcriptome analysis, strongly suggesting that the role of Pla1 in the context of the MTREC complex is not the initial poly-adenylation but rather the extension of the poly(A) tails of MTREC substrates, leading to the well documented hyperadenylation of these transcripts^{34–36}.

Disrupting Red1-Pla1 interaction affects the poly(A) tail lengths of CUTs and meiotic mRNAs

To assess poly(A) tail length of CUTs and meiotic mRNAs in WT and Red1-Pla1 interaction mutants, we used Oxford Nanopore Technologies' direct RNA sequencing technology, combined with the *tailfindr* R package⁵⁶. This package can estimate poly(A) tail length of individual reads from the length of the monotonous low-variance raw signal, corresponding to poly(A) tails, at the beginning of each read and combines this information with the unique read ID. After mapping the individual reads to the *S. pombe* genome, we can assign the poly(A) tail length information to these mapped reads, determine mean poly(A) tail length of selected transcripts (e.g., all sequenced transcripts of a particular gene or a particular CUT), and display this information for selected population of transcripts (e.g., CUTs or meiotic mRNAs). Since CUTs and meiotic mRNAs are extremely low abundant in WT cells, and might represent a specific sub-population of these transcripts that escaped MTREC and exosome-mediated degradation, we decided to measure the poly(A) tail length of RNA transcripts associated with the MTREC complex. We have previously shown that RNA immunoprecipitation (RIP) of MTREC complex components strongly enrich CUTs and meiotic mRNAs²⁶, therefore, we purified Flag-ProtA-tagged WT Red1 as well as Red1 $_{W298A/F305A}$ and Red1 $_{\Delta 288-345}$ mutants, using conditions that preserve RNA-protein complexes and analysed the co-precipitating RNA transcripts by direct RNA sequencing.

Previous studies evaluated the poly(A) tail length of individual transcripts (individual meiotic mRNAs or individual CUTs^{26,34–36}) and concluded that these transcripts are hyper-adenylated, compared to normal mRNAs. Our genome-wide data (Fig. 5D, E) confirm these findings, estimating median poly(A) tail length of mRNAs to be around 31 nucleotides (nts) in WT *S. pombe* cells, while PROMPTs and meiotic mRNAs have nearly twice longer poly(A) tails.

The mean poly(A) tail length distribution of IP-ed mRNA transcripts (likely representing contaminating RNA transcripts in these RIP experiments) shows only a minor difference between WT and mutants (Fig. 5D and Supplementary Fig. 8C; median values are 31, 29 and 27 nts in WT, *red1* $_{W298A/F305A}$ and *red1* $_{\Delta 288-345}$ mutants, respectively), while PROMPTs show poly(A) tail length decreased by ~20 nts in *red1* $_{W298A/F305A}$ and *red1* $_{\Delta 288-345}$ mutants, compared to WT (Fig. 5E and Supplementary Fig. 8D; median 54 in WT to median 35 and 37 nts in *red1* $_{W298A/F305A}$ and *red1* $_{\Delta 288-345}$ mutants, respectively). Interestingly, while Red1-Pla1 interaction mutants did not affect the degradation of AS RNAs or meiotic mRNAs, the median poly(A) tail lengths of these transcripts were also reduced in the mutant strains (Supplementary Fig. 8E–H; meiotic mRNAs: median 56 nts in WT to 32 nts in *red1* $_{W298A/F305A}$ and 42 nts in *red1* $_{\Delta 288-345}$ mutants; AS RNAs: median 46 nts in WT to 33 nts in *red1* $_{W298A/F305A}$ and 40 nts in *red1* $_{\Delta 288-345}$ mutants). Overall, these experiments show that the poly(A) tail length of CUTs and meiotic mRNAs are reduced by ~20 nts in *red1* $_{W298A/F305A}$ and *red1* $_{\Delta 288-345}$ mutants, while the poly(A) tail length of mRNAs remains unaffected.

Interaction of MTREC complex with CPF

3'-end processing of pre-mRNA involves the complex assembly and action of a number of proteins, including the highly conserved CPF^{57–60}. CPF is involved in site-specific endonucleolytic cleavage of the pre-mRNA followed by addition of a poly(A) tail at its 3' end, which is required for nuclear export of mRNAs. In *S. cerevisiae*, the CPF is composed of three enzymatically active modules of which the polymerase module encompasses, among others, the scPap1⁶⁰ and scFip1 subunits, and scFip1 has been suggested to tether scPap1 to CPF^{45,46,61–63}. Consistently, the interaction between *hsFip1* and *hsPAP* is also conserved in humans⁴⁸. scFip1 is an intrinsically disordered protein with short regions directly contacting scPap1/scYth1^{61–63}. The crystal structure of scPap1 in complex with a 26 amino acid scFip1 peptide shows that the N-terminus of the short peptide forms a parallel β -ribbon with scPap1 and adds an antiparallel β -strand extending the β -sheet interface of scPap1⁶¹. Structure alignment of the scPap1-scFip1 complex with our *S. pombe* Pla1-Red1 structure shows that scFip1 and Red1 partially occupy a similar binding interface on the polymerase (Supplementary Fig. 9A). Interestingly, there have been suggestions that the MTREC complex and 3'-end processing machinery act in a coordinated fashion where the CPF complex cleaves the RNA destined for degradation, thereby creating an entry point for the MTREC complex^{64,65}.

Since the binding interfaces on the polymerase for scFip1 and Red1 are overlapping, we were wondering if Red1 and Iss1, the fission yeast ortholog of scFip1, might compete for binding to Pla1. Based on sequence alignment between scFip1 and Iss1 and the extended scPap1 binding site recently identified within scFip1⁶³ (Supplementary Fig. 9B), we tested the binding of a 47-amino acid construct of Iss1 (residues 30–76) to Pla1 using ITC. We found that Iss1_{30–76} binds to Pla1 with micromolar affinity ($K_D = 2.5 \mu\text{M}$) (Fig. 6A). In competition experiments, where Red1_{288–345} is titrated into a pre-formed Pla1_{RRM}-Iss1_{30–76} complex, we observed an ~8-fold decrease in affinity of Pla1 for Red1 ($K_D = 11 \mu\text{M}$; Fig. 6B). In contrast, a titration of Iss1 into Pla1_{RRM}-Red1_{288–345} complex is unable to displace Red1 from the complex (Fig. 6C) compared to a control where the double mutant W298A/F305A of Red1_{288–345} was used for complex formation (Supplementary Fig. 9C). These data show the existence of negative cooperativity between Iss1 and Red1 due to an

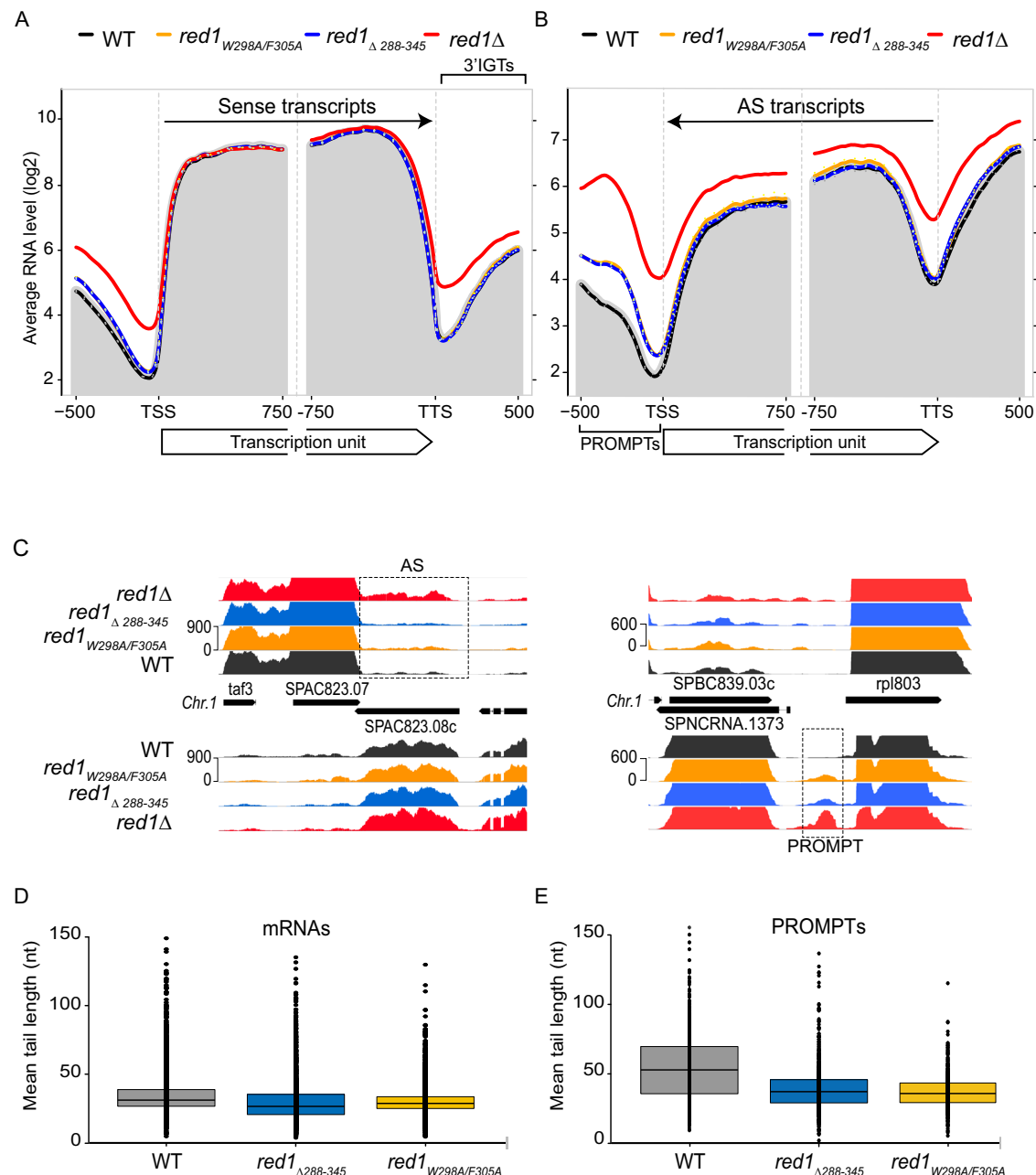


Fig. 5 | Pla1 in the context of the MTREC complex is responsible for degradation of PROMPTs. **A, B** Metagenome profile of sense (**A**) and antisense (**B**) RNA levels in the indicated strains for a subset of *S. pombe* genes with detectable levels of PROMPTs (2400 genes). The geometric average of RNA levels from 500 bp upstream to 750 bp downstream of the transcription start site (TSS) and 750 bp upstream to 500 bp downstream of the transcription termination site (TTS) are shown. Solid lines represent the average of two replicates for all indicated strains, with the exception of *red1*Δ which represent a single dataset. Dotted lines indicate the individual biological replicates. The grey shading represents the average RNA levels in the WT strain. **C** Strand-specific RNA-seq read coverage of a representative set of

genes in WT and mutant strains. Dashed boxes highlight two representative examples of CUTs: antisense RNA transcripts (AS) on the left panel and PROMPTs on the right panel. **D, E** Box-plot of mean poly(A) tail length distribution of MTREC-associated mRNAs (**D**) and PROMPTs (**E**) for the WT and indicated mutant strains (5135 mRNAs and 619 PROMPTs with FC > 2 are shown). Dots represent the mean poly(A) tail length of individual mRNAs/PROMPTs, boxes show the 25–75 percentile range and the horizontal lines represent the median values. Upper and lower whiskers represent 75th percentile plus 1.5 times the inter-quartile distance (IQR) and the 25th percentile minus 1.5IQR, respectively. Source data are provided in Source Data file.

overlapping binding interface on Pla1_{RRM}. The inability of Iss1 to out-compete Red1 from the Pla1-Red1 complex, even though it binds only ~1.8-fold weaker compared to Red1₂₈₈₋₃₄₅, prompted us to investigate the binding kinetics of the two proteins. We therefore used bio-layer interferometry (BLI) to measure possible differences in the K_{on}/K_{off} rates of biotinylated Red1₂₈₈₋₃₄₅ (Fig. 6D) or Iss1₃₀₋₇₆ (Fig. 6E) immobilised on streptavidin-coated biosensors to Pla1_{FL}. Compared to Iss1 ($k_{on} = 2.1 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, $k_{off} = 2.4 \times 10^{-2} \text{ s}^{-1}$), Red1₂₈₈₋₃₄₅ has a faster

k_{on} ($2.8 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$) and a slower k_{off} ($1.6 \times 10^{-2} \text{ s}^{-1}$; Supplementary Table 3).

Furthermore, we have previously observed that a member of accessory cleavage factors (CF) of CPF known as Msi2 (scHrp1 or CF IB in *S. cerevisiae*), co-purifies with MTREC²⁶, suggesting a possible association between these two machineries. We therefore wanted to understand and identify whether Msi2 could directly interact with individual components of MTREC. Indeed, we observed that Msi2

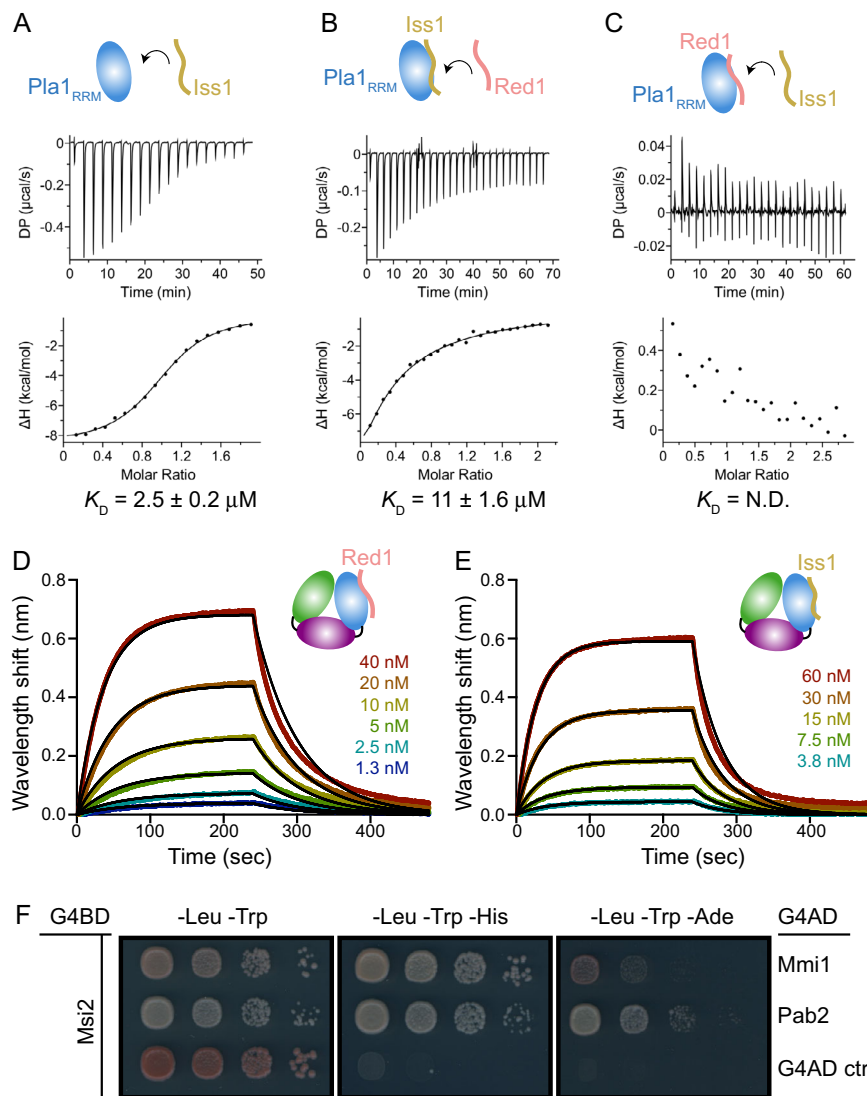


Fig. 6 | Interactions between MTREC and CPF. A ITC titration of Iss1 into Pla1_{RRM}. **B** ITC titration of Red1₂₈₈₋₃₄₅ into a pre-formed complex of Pla1_{RRM}-Iss1₃₀₋₇₆. **C** ITC titration of Iss1 into a pre-formed complex of Pla1_{RRM}-Red1₂₈₈₋₃₄₅. The calculated dissociation constants (K_D) from an average of two independent measurements are shown. BLI kinetic analyses of biotinylated Red1₂₈₈₋₃₄₅ and Iss1₃₀₋₇₆ with varying

concentrations of Pla1_{FL} are shown in panels **D** and **E**, respectively. Black lines represent fits to the experimental data obtained using a 1:1 binding global fitting model. **F** Y2H experiments show interaction between Msi2 fused to G4BD and MTREC components Mmi1 or Pab2 fused to G4AD compared to an auto-activation control (ctrl).

binds to two distinct members of the MTREC complex, Pab2 and Mmi1, in Y2H experiments (Fig. 6F and Supplementary Fig. 9D). In summary, these data show that multiple points of association exist between CPF and the MTREC complex and suggest a close functional cooperation between these two complexes in RNA surveillance.

MTREC-Pla1 interaction is required for heterochromatic island formation at meiotic genes

Recent studies reported the involvement of the CPF, including Pla1, in the formation of small, facultative heterochromatic islands at meiotic genes^{64,65}. We wondered if the Red1-Pla1 interaction might play a direct role in this process. We carried out ChIPseq experiments to detect H3K9me2 modifications throughout the fission yeast genome in WT, *red1Δ* and the Red1-Pla1 interaction mutant strains (*red1*_{W298A/F305A} and *red1*_{Δ288-345}). We used 3 independent WT strains in this experiment, with slightly different genetic backgrounds (P1, P419, F3230; see Supplementary Table 4 for genotype information), to account for potential variabilities in the appearance of these facultative heterochromatic islands. We detected all previously described heterochromatic islands

within the *S. pombe* genome⁶⁶⁻⁶⁹, and their size and H3K9me2 enrichment levels were remarkably uniform between the three independent WT strains (Fig. 7A, B). We also confirmed that most of the facultative heterochromatic islands, mainly located at meiotic genes, are dependent on an intact MTREC complex, as evidenced by the complete absence of these islands in the *red1Δ* cells (Fig. 7B, C and refs.^{64,65}). Other islands (non-meiotic islands) are independent of the MTREC complex and they are unaffected in the *red1Δ* strain⁶⁶ (Fig. 7D). Interestingly, H3K9me2 enrichment levels at all MTREC-dependent facultative heterochromatic islands were strongly reduced in the Red1-Pla1 interaction mutants, while MTREC-independent islands and the major heterochromatic regions (pericentromeric regions, telomers, mating-type and rDNA-region) were unaffected (Fig. 7B–D). The reduction of the H3K9me2 levels at these islands were comparable to the effect that was reported in CPF mutant strains⁶⁴. This finding is remarkable, since *red1*_{W298A/F305A} and *red1*_{Δ288-345} mutants do not interfere with MTREC-mediated degradation of meiotic mRNAs. These results suggest that while the recruitment of MTREC complex to meiotic mRNAs does not require Pla1 interaction with the MTREC complex, efficient

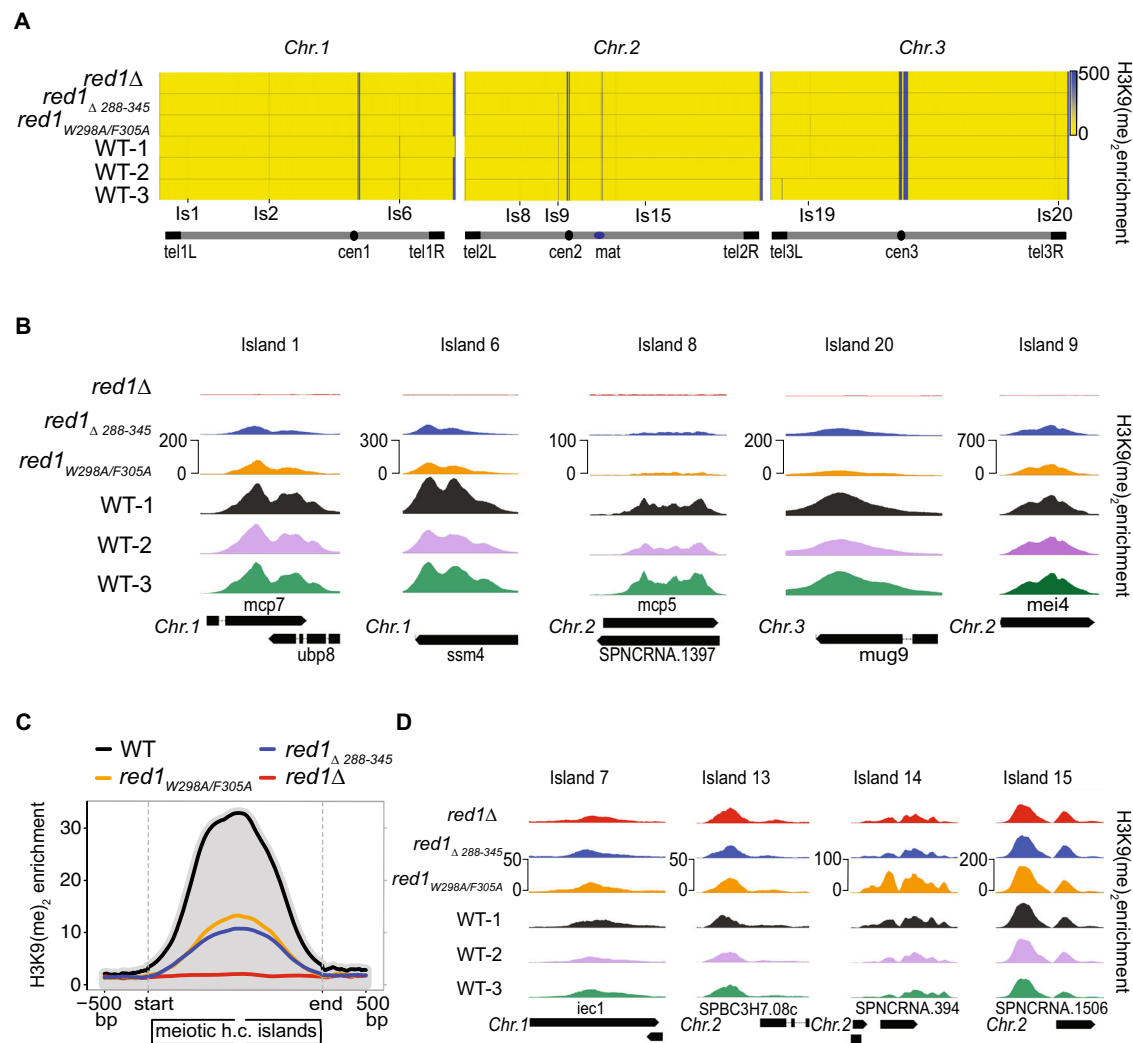


Fig. 7 | Pla1 – MTREC interaction is required for efficient assembly of facultative heterochromatic islands at meiotic genes. **A** Genome-wide heat-map representation of H3K9me2 ChIP-seq analysis of three independent WT strains (WT-1 is the isogenic WT strain for the mutants) and *red1* Δ , *red1*_{W298A/F305A} and *red1* $\Delta_{288-345}$ mutants, showing the H3K9me2 enrichment levels (yellow/blue scale of 0-500) for the 3 chromosomes of the *S. pombe* genome (length not to scale). Centromeres, Mating-type locus, telomeres and a selection of facultative heterochromatic islands are indicated. **B** Representative examples of *red1* Δ -sensitive meiotic heterochromatic islands in WT and mutant strains are shown in higher resolution (scales are indicated for individual islands). **C** Average of H3K9me2 enrichment over all

meiotic heterochromatic islands (12 islands: is1(mcp7); is1.5(tht2); is1.6(SPAC631.02), is2(mug8); is4(SPAC8C9.04/SPNCRNA.925); is5(vps29); is6(ssm4); is8(mcp5); is9(mei4); is16(mbx2/SPNCRNA.1626); is17(mug45); is20(mug9)) plotted on a linear scale. The plots represent the geometric average of enrichment values from 500 bp upstream to 500 bp downstream of the islands, with the island regions scaled to the same lengths for all islands (stretched or condensed to 2000 bp). The grey shading represents the average H3K9me2 enrichment levels in the WT strain. **D** Representative examples of non-meiotic (*red1* Δ -insensitive) heterochromatic islands in WT and mutant strains (scales are indicated for individual islands).

establishment and/or maintenance of the heterochromatic islands at these loci is dependent on the intact MTREC-Pla1 physical interaction.

Discussion

The multi-subunit MTREC complex serves as the exosome-adaptor complex required for degradation of CUTs in *S. pombe*. In this study, we identified and characterised the Pla1-Red1 interaction using Y2H, NMR and X-ray crystallography. We showed that the C-terminal RRM domain of Pla1 binds to a 58-residue region of the MTREC core component Red1 (residues 288–345), tethering it to the MTREC complex (Fig. 1). Our crystal structure of the Pla1-Red1 complex showed that Red1 is largely unstructured but comprises an α -helix and a β -strand at the N- and C-termini, which hold Red1 in position to form a tight binding interface with Pla1_{RRM} (Fig. 2B). Using structure-based mutational analyses, we identified a *Red1*_{W298A/F305A} double mutant in the N-terminal α -helix that severely compromises the Pla1-Red1 interaction in vitro (Fig. 3A) and in vivo (Fig. 4), while deletion of the entire

interaction surface in the *Red1* $\Delta_{288-345}$ mutant leads to no detectable interaction between the MTREC complex and Pla1. Using these Pla1-Red1 interaction mutants, we showed that the truncated MTREC, devoid of Pla1, is unable to hyper-adenylate CUTs and meiotic mRNAs, but interestingly, this only leads to the inefficient degradation of PROMPTS. In addition, the Pla1-Red1 interaction is required for the efficient establishment and/or maintenance of facultative heterochromatic islands around meiotic genes, as evidenced by the strongly impaired levels of H3K9me2 at these loci in the mutants.

Interestingly, similar to the MTREC complex where Pla1 is tethered to the complex via Red1, the *S. cerevisiae* homologue of Pla1 (scPap1) has been shown to be flexibly tethered to the CPF core machinery via the intrinsically disordered protein scFip1^{61–63}. Structural superimposition of Pla1-Red1 and the *S. cerevisiae* scPap1-scFip1 complexes show that scFip1 and Red1 occupy a similar binding interface on the respective polymerase (Supplementary Fig. 9A). Surprisingly, scFip1 binds to scPap1 with picomolar affinity⁶¹, while in our

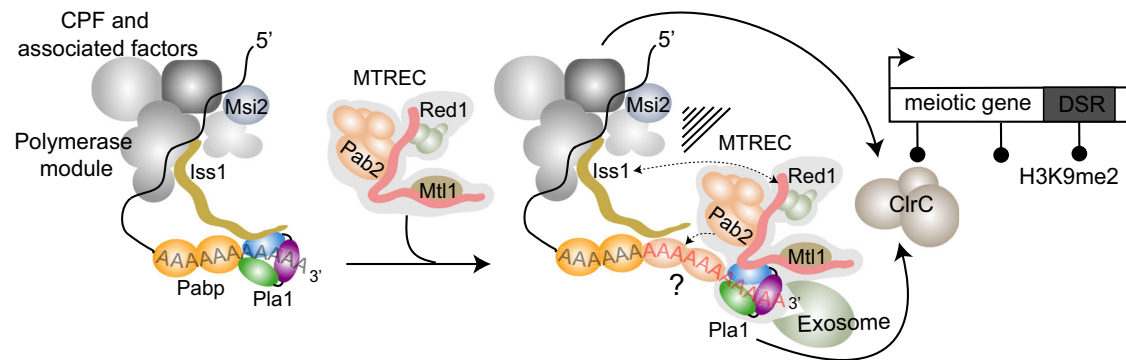


Fig. 8 | Model of the role of Pla1 in MTREC-mediated degradation of CUTs.

MTREC complex is recruited to CUTs and meiotic mRNAs during their transcription by a not well understood mechanism. Pla1, as part of the CPF, is responsible for the initial poly-adenylation of CUTs and the resulting poly(A) tail is likely bound by the canonical poly(A) binding protein Pabp. MTREC complex sequesters Pla1 from CPF via Red1, replacing Iss1 that anchors Pla1 to CPF. MTREC-bound Pla1 hyper-adenylates CUTs. We suggest that the non-canonical poly(A)-binding protein, Pab2,

might be preferentially loaded on these poly(A) tail extensions (marked with a "?" in the Fig.) to facilitate exosome-mediated degradation of CUTs. Both the CPF complex and the Red1-Pla1 interaction are required for the efficient recruitment of the CtrC complex to methylate histone H3K9 at meiotic heterochromatic islands, indicating a sophisticated functional interplay between CPF and MTREC complex during the transcriptional termination, end-processing and degradation of CUTs.

experiments the *S. pombe* homologue of scFip1 (Iss1) binds with a weaker, low micromolar affinity ($K_D = 2.5 \mu\text{M}$, Fig. 6A), which can be attributed to the low sequence conservation of scFip1/Iss1 in the polymerase binding region and different assay conditions (ref. ⁶³ and Supplementary Fig. 9B). We find that the binding affinities of Pla1 with Red1 and Iss1 derived using BLI ($K_D = 59.2 \text{ nM}$ for Red1, $K_D = 115.2 \text{ nM}$ for Iss1) are $\sim 22\text{--}24$ fold stronger than those obtained using ITC ($K_D = 1.4 \mu\text{M}$ for Red1, $K_D = 2.5 \mu\text{M}$ for Iss1), however the general trend with Red1 binding 2-fold stronger than Iss1 is maintained between the methods. We reason that immobilisation of Red1 or Iss1 on a surface (as done in BLI) decreases the degree of freedom for these proteins leading to an increase in apparent binding affinity as opposed to ITC where the components are mobile and free in solution. In our *in vitro* competition experiments, we observed that Red1 is able to out-compete Iss1 from its complex with Pla1 (Fig. 6B, C) owing to faster k_{on} and slower k_{off} rates of Red1 when compared to that of Iss1 (Fig. 6D, E and Supplementary Table 3). While these results might differ in the context of the complete CPF and MTREC complexes, there is no evidence to suggest additional interactions between scPap1/Pla1 and other components of CPF or MTREC, besides scFip1⁶³ and Red1²⁹, respectively. Such a negative cooperation between Iss1 and Red1 is suggestive of Pla1 sequestration from CPF via Red1 to hyper-adenylate CUTs as part of the MTREC complex, although the simultaneous existence of separate copies of Pla1 in these complexes cannot be ruled out (Fig. 8).

Our total RNA sequencing experiments in the Pla1-Red1 interaction mutants did not uncover stabilised transcripts without poly(A) tail, strongly suggesting that the role of Pla1 in the context of the MTREC complex is not the initial poly-adenylation of CUTs, but rather the extension of the poly(A) tail of these transcripts. We carried out poly(A) tail length analyses and showed that the median poly(A) tail length of mRNAs is ~ 31 nts in WT *S. pombe* cells, while PROMPTs and meiotic mRNAs harbour a poly(A) tail of about 20 nts longer. The median poly(A) tail length of PROMPTs and meiotic mRNAs were 54 and 56 nts, respectively, while AS RNAs showed a somewhat shorter median tail length of 46 nts. It is unclear if PROMPTs and meiotic mRNAs are indeed more extensively hyper-adenylated than other subclasses of CUTs, or whether this lower median value is a data analysis artefact. While PROMPTs and meiotic mRNAs can be annotated and bioinformatically captured relatively easily, intergenic and long antisense CUTs are harder to define and our analysis likely also includes stable ncRNAs which are not hyper-adenylated. A recent study⁷⁰ estimated that the median poly(A) tail length of mRNAs is

48.9 nts in *S. pombe* cells. The discrepancy between these numbers (~ 31 nts in our study versus ~ 49 nts in ref. ⁷⁰) might arise from the fact that we used RNA transcripts isolated from RIP experiments (RNA IPs using WT or mutant Red1-FTP) to enrich low abundant CUTs. Residual RNase activity during the IP procedure might shorten the poly(A) tails in our experimental conditions and we likely underestimate the lengths of the poly(A) tails for all transcripts (both mRNAs and CUTs). Nevertheless, our genome-wide poly(A) tail length analysis confirms that, in WT *S. pombe* cells, CUTs and meiotic mRNAs are hyper-adenylated, similar to previous reports that analysed individual CUTs or meiotic mRNAs^{26,34,35,71}. However, in the Pla1-Red1 interaction mutant strains, the poly(A) tail lengths of CUTs and meiotic mRNAs are decreased, while the poly(A) tails of mRNAs (other than meiotic mRNAs) are not affected. In these mutants, all RNA species, including CUTs and mRNAs, have a relatively uniform median poly(A) tail length of ~ 30 to 35 nts, likely representing the consistent action of Pla1 as part of CPF. These results further support our model that Pla1, as part of the MTREC complex, hyper-adenylates MTREC target RNAs, extending the existing poly(A) tail by ~ 20 nts (Fig. 8).

It is intriguing why the addition of 20–25 nt poly(A) tails to PROMPTs leads to their efficient degradation, while PROMPTs harbouring shorter poly(A) tails, due to decoupling of Pla1 from the MTREC complex, are inefficiently degraded by the nuclear exosome. Given that the core exosome channel leading to the Dis3 active site can accommodate $\sim 25\text{--}30$ nts⁷², the additional hyper-adenylation of CUTs by MTREC must serve another purpose. One possibility is that hyper-adenylated CUTs are instantaneously bound by the MTREC component poly(A) binding protein 2 (Pab2), making the extended poly(A) tail sterically unavailable for binding to Pabp/Pab1, the major poly(A) binding protein involved in nuclear export and poly(A) tail length control in mRNAs (reviewed in⁷³). While Pab2 does not directly influence the polyadenylation activity of Pla1⁵⁴, as opposed to its mammalian homologue PABPN1^{74,75}, it is required for efficient nuclear exosome-mediated degradation of CUTs, meiotic mRNAs and unspliced pre-mRNAs^{25,26,71,76,77}.

Determination of RNA fate is an extremely complex process requiring timely protein-protein interactions. Our data suggest a close functional interplay between 3'-end processing and the RNA surveillance machinery. However, we found that Msi2 was the only component of the 3'-end processing machinery co-purifying with MTREC complex when we use benzonase treatment, indicating direct protein-protein interactions³⁶. Indeed, in our Y2H experiments, Msi2 interacts with two subunits of the MTREC complex, Mmi1 and Pab2 (Fig. 6F).

Importantly in *S. cerevisiae*, scHrp1 (homologue of Msi2) has been reported to participate in surveillance of CUTs⁷⁸, scNrd1-dependent termination⁷⁹ and cytoplasmic nonsense-mediated decay⁸⁰, in addition to its canonical role in the 3'-end processing machinery⁸¹. Whether the interaction between Msi2 and Mmi1/Pab2 helps the recruitment of MTREC to CUTs, or MTREC-bound transcripts are cleaved and polyadenylated by a specialised form of CPF, remains a topic for future investigations.

In recent years, evidence is emerging for a co-transcriptional interaction between CPF and MTREC components⁸². Both the MTREC complex and multiple subunits of the CPF are required for the establishment and/or maintenance of facultative heterochromatic islands at meiotic genes^{27,34,64,65,69,83}. The fact that the Pla1-Red1 interaction mutants strongly destabilise these heterochromatic islands, very similar to the deletion of various CPF subunits, suggests that Pla1, as part of the MTREC complex, must functionally closely interact with CPF, as opposed to having only a strictly downstream function in this process. The exact nature of these interactions and the functional interplay between MTREC, CPF and the ClrC complex that is recruited to these loci to methylate histone H3K9, is not yet understood. Interestingly, PAXT component ZFC3H1 (human orthologue of Red1) was shown to functionally interact with the PRC2 complex⁸⁴, a key factor in the establishment and maintenance of facultative heterochromatin in higher eukaryotes. *S. pombe* does not have orthologues of PRC complexes and H3K27 methylated heterochromatin domains. However, the facultative H3K9me2 heterochromatin islands in *S. pombe* are thought to be the functional predecessor of the PRC-mediated facultative heterochromatin domains in higher eukaryotes. Further studies will be required to delineate the mechanistic and functional details of this fascinating interaction between the RNA surveillance machinery and the epigenetic regulation of the genome.

Methods

PCR and cloning

The DNA sequence encoding Pla1_{FL} (residues 1-566), Pla1_{Δ14} (residues 1-542), Pla1_{Δ153A} (catalytic mutant, residues 1-566) were cloned and ligated into a modified pET24d vector containing non-cleavable N-terminal His₆ tag. The C-terminal RRM domain of Pla1 (Pla1_{RRM}, residues 352-566) was sub-cloned from Pla1_{FL} into a modified pET24d vector containing N-terminal His₆ tag with the Tobacco-Etch-Virus cleavage site present before the corresponding protein. The DNA sequence encoding Iss1 (residues 30-76) was cloned into a modified pET24d vector containing an N-terminal His-Thioredoxin tag upstream of a TEV site. All wild-type and mutant proteins of Red1 with different boundaries were cloned into a modified pET24a vector containing N-terminal GBI tag-TEV site before the corresponding protein sequence and a C-terminal non-cleavable His₆ tag. Site directed mutagenesis with QuickChange Lightning kit was used to introduce point mutations in Pla1 and Red1 proteins, or deletion of residues 288-345 in Red1 and the mutations were confirmed using DNA sequencing.

Protein expression and purification

The plasmids were transformed in BL21 (DE3) Rosetta2 chemically competent cells for Pla1 and Iss1, grown at 37 °C up to an OD of 1.4–1.6 in Luria broth or Terrific broth subsequently expressed at 20 °C for -16 hr after induction with 0.5 mM IPTG. For Red1, the plasmids were transformed in BL21 (DE3) chemically competent cells, grown at 37 °C up to an OD of 1 in auto-induction media⁸⁵ and subsequently expressed at 20 °C for -16 h. For isotope-labelled proteins, the bacteria were grown in M9 minimal media supplemented with ¹³C-glucose and/or ¹⁵NH₄Cl. Cells were lysed in IMAC buffer containing 20 mM Tris pH 7.5, 200 mM NaCl, 20 mM Imidazole and 2 mM β-mercaptoethanol. The proteins were purified over 1–2 ml His-Trap FF columns (GE Healthcare) with elution buffer containing 250 mM Imidazole. In case of Red1 or Pla1 CTD, overnight tag-cleavage using Tobacco Etch Virus protease

and simultaneous dialysis of the proteins into 20 mM HEPES pH 7.5, 150 mM NaCl, 2 mM BME was carried out. The Red1 proteins were further purified over a second IMAC column where the cleaved proteins (still containing C-terminal His₆ tag) were eluted using 250 mM Imidazole. For Iss1, purification over a second IMAC column was made to remove uncleaved protein and the fusion tag from the cleaved protein and the flow through was collected. Finally, the Red1 and Iss1 proteins were polished using gel filtration (Superdex 75 16/60) column (GE Healthcare) equilibrated with buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl and 1 mM DTT. For ITC experiments, the DTT was replaced by 2 mM BME and for NMR experiments the proteins were purified in buffer containing 20 mM sodium phosphate pH 6.5, 100 mM NaCl, 1 mM DTT. For Pla1 proteins, after the IMAC column the proteins were diluted to buffer containing 50 mM NaCl and further purified using 1 ml Resource Q cation exchange column (GE Healthcare), where they were eluted with a linear gradient of 50 mM NaCl to 1 M NaCl. Gel filtration (Superdex 200 16/60) column (GE Healthcare) was used as a final polishing step, where the protein was purified in buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl and 1 mM DTT. For crystallisation of the apo form, the salt concentration was adjusted to 200 mM.

X-ray crystallography

Pla1_{FL} crystallised at a concentration of 12.7 mg/ml in a drop containing 0.2 M lithium citrate, 20% PEG 3350 at 4 °C as needles within 1 week. Pla1_{Δ14} crystallised at a concentration of 6.3 mg/ml in a drop containing 0.2 M sodium formate, 20% PEG 3350 at 18 °C as thin plates within 12 days. The Pla1-Red1 complex was prepared by addition of 1.2-fold molar excess of Red1₂₈₈₋₃₄₅ over Pla1_{FL} and subsequent purification over a size exclusion column in buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl and 1 mM DTT, to remove excess Red1. The complex crystallised in a drop containing 0.2 M potassium formate, 20% PEG 3350 at 18 °C as thin plates within 3 weeks. Crystals were flash frozen in mother liquor supplemented with 20% glycerol. Several datasets for the crystals were collected at P13 and P14 beamlines at PETRA III, EMBL Hamburg using mxCuBE v2. Datasets from best diffracting crystals were then processed with XDS software package⁸⁶ and the structure was solved by molecular replacement using PDB ID: 2HHP (chain A). The missing residues of Pla1 were built using Coot model building software⁸⁷ with multiple rounds of model building and refinement with Phenix⁸⁸. The Pla1_{FL} apo structure was then used as a molecular replacement model for solving the structure of Pla1-Red1 complex. The data were processed with XDS software package⁸⁶. Since the crystal showed anisotropic diffraction (with resolution along the best diffracting axis being 2.81 Å and diffraction limits along the other axes being 3.2 Å and 3.8 Å), STARANISO server (<http://staraniso.globalphasing.org>)⁸⁹ was used to apply an anisotropic correction to the data. The anisotropic correction provided an improvement in the overall interpretability of the map especially for the Pla1-Red1 interface residues. Multiple rounds of model building were done in Coot⁸⁷ and refinement was performed with Phenix⁸⁸.

NMR spectroscopy

All spectra were recorded at 293 K on Avance III Bruker NMR spectrometers with proton Larmor frequencies of 600 MHz, 700 MHz or 800 MHz, equipped with cryogenic (600 MHz, 800 MHz) or room temperature (700 MHz) triple resonance gradient probes using Topspin v3.2. NMR experiments were performed with Red1 constructs (Red1₂₈₈₋₃₄₅ or Red1₂₈₈₋₃₂₂) harbouring a GBI-tag at the N-terminus and a His₆-tag at the C-terminus. For backbone assignment of Red1₂₈₈₋₃₂₂, data were recorded on 1 mM protein in buffer containing 25 mM sodium phosphate pH 6.5 and 100 mM NaCl, supplemented with 10% D₂O for lock. Protein backbone resonance assignments were obtained using 3D HNCA, HNCACB, CBCA(CO)NH and HNCO⁹⁰. NMR titrations were done by adding 1.5-fold excess of Pla1_{RRM} into 0.3 mM ¹⁵N-labelled

RedI constructs (RedI₂₈₈₋₃₄₅ or RedI₂₈₈₋₃₂₂) and recording ¹H,¹⁵N HSQC spectra. All experiments were performed at 293 K. Spectra were processed in NMRPipe/Draw⁹¹ and analysed in CCPN Analysis⁹².

Small angle X-ray scattering experiments

The PlaI-RedI complex was prepared by addition of 1.5-fold molar excess of RedI₂₈₈₋₃₄₅ over PlaI_{FL} and subsequent purification over a size exclusion column in buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl and 1 mM DTT, to remove excess RedI. Peak fractions corresponding to the complex were pooled and concentrated. Measurements were performed at 20 °C at EMBL P12 beamline, PETRA III (DESY, Hamburg, Germany) at concentrations ranging from 0.05–1.7 mg/ml. Forty successive frames with 0.195 s/frame were recorded using an X-ray wavelength of $\lambda = 0.124193$ nm. 1D scattering intensities of samples and buffers were expressed as a function of the modulus of the scattering vector $Q = (4\pi/\lambda)\sin\theta$ with 2θ being the scattering angle and λ the X-ray wavelength. Downstream processing after buffer subtraction was done with PRIMUS⁹³. R_g was determined using Guinier approximation and from $p(r)$ curve. Disordered regions in the crystal structure were modelled using CORAL⁵² and subsequently crystal structure validation was done using CRY SOL⁹⁴.

Isothermal titration calorimetry

ITC experiments were performed with MicroCal PEAQ-ITC calorimeter (Malvern) at 25 °C. Protein samples were dialysed overnight in buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl, 2 mM BME. The cell was filled completely with 40–50 μ M protein and the syringe was filled with concentrations in the range of 400–500 μ M of the respective ligand. A series of 19 or 26 injections of 2 μ l or 1.5 μ l titrants were made into the respective protein. For competition experiments, PlaI_{RRM} was pre-incubated with a 1.1-fold molar excess of either RedI₂₈₈₋₃₄₅, W298A/F305A double mutant of RedI₂₈₈₋₃₄₅ or IssI₃₀₋₇₆ before loading the sample into the cell. The data were processed with PEAQ-ITC Analysis software (PEAQ-ITC) and were fit to a one-binding site model.

In vitro polyadenylation assay

A 15-nt polyA chemically synthesised RNA primer with a 5' Cy3 fluorophore (purchased from iba Life Sciences) was used as a substrate for PlaI. Polyadenylation assays were performed in buffer containing 10 mM HEPES pH 7.5, 20 mM KCl, 5 mM MgCl₂, 0.01 mM EDTA, 5% glycerol and 1 mM DTT at 30 °C. 75 nM PlaI alone or in complex with RedI₂₈₈₋₃₄₅ or W298A/F305A double mutant of RedI₂₈₈₋₃₄₅ at a -10-fold molar excess was incubated with 200 nM RNA primer for 5 min at 30 °C before starting the reaction with the addition of 2 mM ATP. In all, 10 μ l fractions were collected at different time points and mixed with 10 μ l denaturing formamide loading dye to stop the reaction. Products were analysed on 14% denaturing urea-PAGE, run in 0.5x Tris-borate-EDTA buffer for 45 min at 35 mA and the gels were imaged using Amersham Imager 600 (GE Healthcare). Densitometric analyses for integration of gel lane intensity were performed using Fiji ImageJ software.

Yeast two hybrid

The respective sequences of the proteins were cloned into pGBKT7 and pGADT7 vectors from Clontech which harbour either the DNA binding domain or activation domain at the N-termini, respectively. The interaction pairs were analysed by co-transformation into PJ69-4A strain. After a 10-fold serial dilution, colonies were spotted on SDC (SDC-Leu-Trp), SDC-His (SDC-Leu-Trp-His) and SDC-Ade (SDC-Leu-Trp-Ade) plates, incubated at 30 °C and analysed after 3 days. The strength of the interaction was assessed by growth achieved on SDC-His and SDC-Ade as weak and strong, respectively.

Bio-layer Interferometry

BLI experiments were performed on Octet RED96e system (Fortébio) at 25 °C in buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl, 2 mM

BME and 0.01% Tween-20. The protein ligands (RedI₂₈₈₋₃₄₅ and IssI₃₀₋₇₆) were biotinylated using EZ-Link NHS-PEG4-Biotin (Thermo Fisher Scientific). Biotin-labelled proteins were immobilised on the streptavidin (SA) biosensors (Fortébio) and a 2-fold serial dilution of PlaI_{FL} was applied to the biosensors. Parallel experiments with a reference sensor where no analyte was added served as control and its signal was subtracted during data analysis. The association and dissociation periods were both set to 240 sec. Data measurements and analysis were performed by using the Data analysis HT 10.0 (Fortébio) software, with a global (full) 1:1 fitting model.

Tandem affinity purification followed by interaction analysis/ MS

Flag-TEV-protein A (FTP)-tagged bait proteins were harvested from 2 L YEA cultures of yeast strains grown to OD₆₀₀ 1.8–2.2. Cell pellets were snap-frozen in liquid nitrogen and ground into powder using the Cryo-mill MM-400 (Retsch machine). Cells were resuspended in purification buffer (50 mM HEPES, pH 7.0; 100 mM NaCl; 1.5 mM MgCl₂; 0.15% NP-40), supplemented with 1 mM dithiothreitol (Sigma), 1 mM phenylmethylsulphonyl fluoride (Sigma Aldrich, 78830) and protease inhibitor mix (SERVA, 39104). Cell extracts were centrifuged at 3500 g for 10 min at 4 °C, then the supernatants were further centrifuged at 27,000 \times g for 45 min at 4 °C. Clarified supernatants were then incubated with 150 L slurry of IgG beads (GE Healthcare, 17-0969-02) for 2 h at 4 °C on a turning wheel. After binding, the beads were washed with 2 \times 15 mL purification buffer and TEV cleavage was performed in purification buffer containing 20 units AcTEV protease (Invitrogen 12575015), 0.5 mM dithiothreitol and 50 units Benzonase (Sigma, E1014) for the removal of nucleic acids, for 2 h at 16 °C. The eluate was collected and incubated with 100 μ l slurry of anti-Flag beads (Sigma Aldrich, A2220) for 1 h at 4 °C. The protein-bound anti-Flag beads were washed with 2 \times 10 mL purification buffer and the proteins were subsequently eluted from the beads by competition with 200 μ l Flag peptide (Assay Matrix, A6001). The resulting eluate was collected and used for total protein isolation using Trichloroacetic acid- TCA (Sigma, T0699). The eluted proteins were then analysed by Coomassie staining using Brilliant Blue G colloidal concentrate (Sigma Aldrich, B2025) on 4–12% NuPAGE Bis-Tris gels (Life Technologies).

For interaction analysis, the protein samples separated by SDS-PAGE gel was transferred onto a nitrocellulose (Invitrogen, IB301002) membrane, the membrane was cut around 100 kDa marker and the two parts were probed with HRP anti-HA antibody (Abcam ab1190) or HRP anti-Flag M2 antibody (Sigma-Aldrich A8592) in 1:2000 dilution. For mass spectrometry analysis, the precipitated protein samples were subjected to an in-solution tryptic digest using a modified version of the Single-Pot Solid-Phase-enhanced Sample Preparation (SP3) protocol^{95,96}. In total three biological replicates were prepared including control, wild-type and mutant derived lysates (n = 3). Lysates were added to Sera-Mag Beads (Thermo Scientific, #4515-2105-050250, 6515-2105-050250) in 10 μ l 15% formic acid and 30 μ l of ethanol. Binding of proteins was achieved by shaking for 15 min at room temperature (RT). SDS was removed by 4 subsequent washes with 200 μ l of 70% ethanol. Proteins were digested overnight at room temperature with 0.4 μ g of sequencing grade modified trypsin (Promega, #V5111) in 40 μ l Hepes/NaOH, pH 8.4 in the presence of 1.25 mM TCEP and 5 mM chloroacetamide (Sigma-Aldrich, #C0267). Beads were separated, washed with 10 μ l of an aqueous solution of 2% DMSO and the combined eluates were dried down. Peptides were reconstituted in 10 μ l of H₂O and reacted for 1 h at room temperature with 80 μ g of TMT10plex (Thermo Scientific, #90111)⁹⁷ label reagent dissolved in 4 μ l of acetonitrile. Excess TMT reagent was quenched by the addition of 4 μ l of an aqueous 5% hydroxylamine solution (Sigma, 438227). Peptides were reconstituted in 0.1% formic acid, mixed to achieve a 1:1 ratio across all TMT-channels and purified by a reverse phase clean-up step (OASIS HLB 96-well μ Elution Plate, Waters #186001828BA).

Peptides were subjected to an off-line fractionation under high pH conditions⁹⁶. The resulting 12 fractions were then analysed by LC-MS/MS on an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific) as previously described⁹⁸. To this end, peptides were separated using an Ultimate 3000 nano RSLC system (Dionex) equipped with a trapping cartridge (Precolumn C18 PepMap100, 5 mm, 300 µm i.d., 5 µm, 100 Å) and an analytical column (Acclaim PepMap 100. In all, 75 × 50 cm C18, 3 mm, 100 Å) connected to a nanospray-Flex ion source. The peptides were loaded onto the trap column at 30 µl per min using solvent A (0.1% formic acid) and eluted using a gradient from 2 to 40% Solvent B (0.1% formic acid in acetonitrile) over 2 h at 0.3 µl per min (all solvents were of LC-MS grade). The Orbitrap Fusion Lumos was operated in positive ion mode with a spray voltage of 2.4 kV and capillary temperature of 275 °C. Full scan MS spectra with a mass range of 375–1500 m/z were acquired in profile mode using a resolution of 120,000 (maximum fill time of 50 ms or a maximum of 4e5 ions (AGC) and a RF lens setting of 30%. Fragmentation was triggered for 3 s cycle time for peptide like features with charge states of 2–7 on the MS scan (data-dependent acquisition). Precursors were isolated using the quadrupole with a window of 0.7 m/z and fragmented with a normalised collision energy of 38. Fragment mass spectra were acquired in profile mode and a resolution of 30,000 in profile mode. Maximum fill time was set to 64 ms or an AGC target of 1e5 ions). The dynamic exclusion was set to 45 s.

Acquired data were analysed using IsobarQuant⁹⁹ and Mascot V2.4 (Matrix Science) using a reverse UniProt FASTA Schizosaccharomyces pombe database (UP000002485) including common contaminants. The following modifications were taken into account: Carbamidomethyl (C, fixed), TMT10plex (K, fixed), Acetyl (N-term, variable), Oxidation (M, variable) and TMT10plex (N-term, variable). The mass error tolerance for full scan MS spectra was set to 10 ppm and for MS/MS spectra to 0.02 Da. A maximum of 2 missed cleavages were allowed. A minimum of 2 unique peptides with a peptide length of at least seven amino acids and a false discovery rate below 0.01 were required on the peptide and protein level¹⁰⁰, resulting in 164 proteins. Raw TMT reporter ion intensities (signal_sum columns) were first cleaned for batch effects using limma¹⁰¹ and further normalised using vsn (variance stabilization normalisation¹⁰²). Proteins were tested for differential expression using the limma package. The replicate information was not added as a factor in the design matrix since some condition were measured with a single replicate only. A protein was annotated as a hit with a false discovery rate (fdr) smaller 5% and a fold-change of at least 100% and as a candidate with a fdr below 20% and a fold-change of at least 50%.

Total RNA isolation

Total RNAs were isolated using TriReagent (Sigma-Aldrich, T9424). Briefly, cell pellets measuring an OD₆₀₀ of 8 was resuspended in Tri Reagent and lysed, treated twice using 1-Bromo-3-chloropropane (Sigma-Aldrich, B9673) followed by RNA precipitation using 2-propanol (Sigma-Aldrich). The RNA pellet was washed in ice-cold 75% ethanol and solubilized in nuclease-free water. The RNA concentration was measured using NanoDrop (Thermo Scientific) and 10 µg of RNA was used for DNase treatment using NEB DNase I (M0303). The DNase treated samples were purified further using the RNA Clean and Concentrator Kit (Zymo Research, R1015/R1017) and stored at –80 °C after addition of ribonuclease inhibitor (Invitrogen, 10777019).

Illumina sequencing of poly(A) + RNA

Total RNA was isolated, DNase treated and purified from respective strains as mentioned above. The RNA quality was assessed using Bioanalyzer (Agilent 2100) and only those of high quality was proceeded to library preparation. Briefly, 1 µg was taken from individual samples, and ERCC-RNA spikeIn (Life technologies, 4456740) was added according to manufacturer's instructions. The RNA samples

were then subjected to OligodT purification of poly(A) RNA (NEB, E7490) following the manufacturer's instructions and the recovered poly(A) selected RNA was used for cDNA library preparation using NEBNext Ultra II directional RNA library kit (NEB, E7760) following library preparation protocol. The final cDNA libraries were amplified (cycle number of 8) using NEBNext Multiplex Oligos for Illumina (NEB, E7335) and purified using SPRIselect size selection beads (Beckman Coulter, B23317). Individual library quality was assessed on TapeStation (Agilent 4200) using a DNA D1000 High sensitivity tape. Indexed libraries were pooled and sequenced with ~20 M reads per sample on a NovaSeq sequencer with 150 bp paired-end reads.

Illumina sequencing of total RNA

Total RNA was isolated, DNase treated and purified from respective strains as mentioned above. The RNA quality was assessed using Bioanalyzer (Agilent 2100) and only those of high quality was proceeded to library preparation. Briefly, 100 ng was taken from individual samples, and an ERCC-RNA spikeIn (Life technologies, 4456740) amount equivalent to that for 1 µg of RNA input (equivalent to that of poly(A) + sample) was added. The RNA samples were directly proceeded for cDNA library preparation using NEBNext Ultra II directional RNA library kit (NEB, E7760) following the library preparation protocol. The final cDNA libraries were amplified (cycle number of 6) using NEBNext Multiplex Oligos for Illumina (NEB, E7335) and purified using SPRIselect size selection beads (Beckman Coulter, B23317). Individual library quality was assessed on TapeStation (Agilent 4200) using a DNA D1000 High sensitivity tape. Indexed libraries were pooled and subjected for deep sequencing with ~80 M reads per sample on a NovaSeq sequencer with 150 bp paired-end reads.

Nanopore sequencing of poly(A) + RNA

Total RNA was isolated, DNase treated and purified from respective strains as mentioned above. The RNA quality was assessed using Bioanalyzer (Agilent 2100) and only those of high quality was proceeded to library preparation. Briefly, 5 µg of RNA samples was subjected to OligodT purification of poly(A) RNA (NEB, E7490) following the manufacturer's instructions and 100 ng of recovered poly(A) + selected RNA was used for library preparation using Oxford Nanopore direct RNA sequencing library kit (SQK-RNA002; Version: DRS_9080_v2_revM_14Ag2019) protocol. The prepared libraries were run on individual flow cells following manufacturer guidelines.

RNA- Immunoprecipitation followed by sequencing (Illumina and Nanopore)

Tandem affinity purifications of FTP-tagged strains were performed as described above with minor modifications. Reagents were prepared under RNase-free conditions, in the presence of RNase inhibitor (Invitrogen, 10777019). Two-third of final flag elute was taken for DNase treatment using NEB DNase I (M0303) and RNA was purified using RNA Clean and Concentrator Kit (Zymo Research, R1015/R1017). The final RNA concentration was determined using Qubit RNA high Sensitivity Assay Kit and 100 ng of RIP-purified RNA was used for cDNA library preparation using NEBNext Ultra II directional RNA library kit (NEB, E7760) following library preparation protocol. The final cDNA libraries were amplified using NEBNext Multiplex Oligos for Illumina (NEB, E7335) and purified using SPRIselect size selection beads (Beckman Coulter, B23317). Individual library quality was assessed on TapeStation (Agilent 4200) using a DNA D1000 High sensitivity tape. Indexed libraries were pooled and sequenced on NovaSeq sequencer with 150 bp paired-end reads. Similarly, 200 ng of RIP-purified RNA was used for library preparation using Oxford Nanopore direct RNA sequencing library kit (SQK-RNA002; Version: DRS_9080_v2_revM_14Ag2019) protocol. The prepared libraries were run on individual flow cells following manufacturer guidelines.

RNA-seq analysis

Paired-end illumina reads were aligned with hisat 2.1.0¹⁰³ allowing introns with a maximum length of 2000 nt (‘-max-intronlen 2000’). Aligned bam files were then sorted and indexed using samtools 1.3.1¹⁰⁴. Nanopore reads were aligned with minimap (2.10) with long-read spliced alignment (with splice and -k7 parameters). Strand-specific bigwig tracks were generated using bamCoverage¹⁰⁵ to the pombe genome (ASM294v2) with the parameter ‘-binSize 1’. Bigwig tracks were normalised by the sum of the raw signals of chromosome I and II multiplied by 100 million. Meta plots were generated with in-house R scripts using GenomicRanges¹⁰⁶ packages. Integrative Genomics Viewer (IGV2.3, Broad Institute) was used for data browsing and creating representative snapshots.

Detection and quantification of CUTs and PROMPTs

Putative CUT regions were created by neighbouring signal islands closer than 25nts and showing FC > 1.2 relative to red1Δ (length ≥ 100 nts) using Rsamtools¹⁰⁷ and dplyr¹⁰⁸ separately on the forward and reverse bigwig tracks. PROMPT regions were defined as CUTs that are shorter than 1500 nts and in the + -250nts vicinity of an annotated pombe gene (ASM294v2) on the opposite strand. Intersections, subtractions and merging of the predicted regions were done with BedTools 2.28.0.

Quantification of polyA length from poly(A) + RNA Nanopore sequencing

polyA length for the whole transcriptome, meiotic genes, CUT and PROMPT regions was estimated using tailfinder⁵⁶ and boxplots of arithmetic mean of signals per regions were plotted with the ggplot2 R package.

Chromatin Immunoprecipitation-sequencing

S. pombe strains grown to an OD₆₀₀ of 0.5–0.8 were crosslinked using formaldehyde solution (Sigma, F1635) to a final concentration of 1% at RT for 15 min and quenched by addition of glycine (ThermoFisher, AJA1083) to a final concentration of 150 mM for 5 min at RT. The crosslinked culture was then washed and harvested twice using 1x PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.2 mM KH₂PO₄) by centrifugation at 3500 g for 2 min at 4 °C each and the final cell pellet was frozen in liquid nitrogen. The cell pellet was then resuspended in 300 μL FA-SDS buffer (50 mM HEPES-KOH pH 8, 2 mM EDTA pH 8, 150 mM NaCl, 1% TritonX-100, 0.1% Sodium Deoxycholate, 0.1% SDS) supplemented with 1x protease inhibitor mix (SERVA, 39104), 1 mM phenylmethylsulfonyl fluoride (Sigma Aldrich, 78830) was homogenised at 4 °C using 700 μL zirconia beads (BioSpec, 110791) in Precellys 24 homogenizer (Bertin Technologies, France) at 5500 rpm. The lysate was collected and transferred into the Covaris AFA Fiber&cap 12 × 12 mm millitube (Covaris, 520135) after bringing the final amount to 1 mL total. The chromatin was then sheared to a median size of ~300 bp by sonication using Covaris S2 (Adaptive Focused Acoustics™ (AFA) technology) at the given parameters: Duty cycle: 20%, Intensity: 5, Cycles/burst: 200, Temp: 7 °C, Time: 6 min. The sonicated lysate was collected, centrifuged at 4000 g for 10 minutes at 4 °C and the supernatant collected was diluted at 1:1 ratio using FA-lysis buffer (50 mM HEPES-KOH pH 7.5, 2 mM EDTA pH 7.5, 150 mM NaCl, 1% TritonX-100, 0.1% Sodium Deoxycholate) to dilute the SDS concentration to 0.05%. The samples were then incubated overnight at 4 °C with 2 μg of H3K9me2 antibody (mAbcam 1220). Following antibody binding, the samples were incubated at 4 °C for at least 1 h with ~40 μL of appropriate bead slurry (nProtein A Sepharose, GE Healthcare, 5280-01). Following incubation, the beads were collected by centrifugation at 4 °C at 96 g for 1 min. The beads were then washed twice with washing buffer I (50 mM HEPES-KOH pH 7.5, 1 mM EDTA pH 7.5, 140 mM NaCl, 1% TritonX-100, 0.1% Sodium Deoxycholate) and wash

buffer II (50 mM HEPES-KOH pH 7.5, 1 mM EDTA pH 7.5, 0.5 M NaCl, 1% TritonX-100, 0.1% Sodium Deoxycholate). Finally, the beads were washed once again with wash buffer III (10 mM Tris-Cl pH 8, 1 mM EDTA pH 7.5, 250 mM LiCl, 1% NP-40 Igepal, 0.5% Sodium Deoxycholate) and briefly with 1 mL Tris-EDTA pH 8 buffer. The bound protein and chromatin fractions were eluted twice by adding 50 μL of elution buffer (10 mM Tris-HCl pH 8, 10 mM EDTA pH 7.5, 2% SDS) at 65 °C for 10–15 min on a thermomixer. The samples were then de-crosslinked by incubating at 65 °C overnight after addition of 2 μL proteinase K. The DNA was purified using Phenol: Chloroform: Isomyl alcohol method following RNase treatment. The purified DNA pellet was eluted in 15–20 μL of nuclease-free water (Invitrogen, 10977015) or 0.1x TE and was used for library preparation using NEBNext Ultra II DNA library prep kit for Illumina (NEB, E7645) following manufacturer instructions. The final cDNA libraries were amplified using NEBNext Multiplex Oligos for illumina (NEB, E7335) and purified using SPRIselect size selection beads (Beckman Coulter, B23317). Individual library quality was assessed on TapeStation (Agilent 4200) using a DNA D1000 High sensitivity tape. Indexed libraries were pooled and sequenced on NovaSeq sequencer with 150 bp paired-end reads.

ChIP-seq analysis

Paired-end illumina reads were subject to a thorough Quality Control (QC). Shortly, FastQC 0.11.8¹⁰⁹ as used to generate QC reports. Low complexity regions were filtered out with prinseq-lite.pl from PRINSEQ-lite 0.20.4¹¹⁰ maximum dust was set to 3) and the remaining sequences were trimmed using Trimmomatic¹¹¹ (SLIDINGWINDOW: 4 nt; phred quality cut-off: 15, MINLEN = 36). QC filtered were aligned with BWA MEM -0.7.17¹¹². Aligned bam files were sorted and indexed using samtools 1.3.1¹⁰⁴. Potential PCR duplicates were removed with samtools using the rmdup option. Bigwig tracks were generated using bamCoverage¹⁰⁵ to the pombe genome (ASM294v2) with the parameter “-binSize 1”. Bigwig tracks were normalised by the sum of the raw signals of chromosome I and II multiplied by 100 million. Potential artifacts were defined as regions consisting of <15 edges and subsequently removed with a custom R script using rtracklayer¹¹³ and GenomicRanges¹⁰⁶ packages. Meta plots were generated using in-house R-based pipeline.

Generating the Pla1 interaction mutants

The *red1*_{W298A/F305A} and *red1*_{Δ288-345} mutants were generated using conventional primer-based cloning. Together with the wild-type Red1, these constructs were cloned into a pFA6a-Hygro plasmid in untagged and 3xFTP tagged versions. The primers used to generate the mutations are listed in Supplementary Table 5. The plasmids generated were confirmed for mutations using PCR and Sanger sequencing at the ACRF Biomolecular Resource Facility at JCSMR, ANU. The desired cassettes were then digested from the positive plasmid using SpeI restriction enzyme that cuts in the 5'UTR and 3'UTR sequence generating overlapping region for homologous recombination in the Red1 locus when used to transform into a *red1Δ* strain.

In-order to generate strains retaining the 3'UTR, 300 bp starting from the stop codon of Red1 gene was amplified along with a 500 bp homology region upstream. Similarly, 500 bp region following the 300 bp was amplified and Gibson assembled to PCR amplified product from pFA6a-NatNT2 using appropriate primer pairs. This overrode the Hygromycin resistance marker to a natNT2 resistance along with retaining 300 bp 3'UTR region of Red1. Hence all the untagged mutants used for RNA expression analysis were constructed retaining their 3'UTR region.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data that support this study are available from the corresponding authors upon reasonable request. NMR backbone chemical shifts for Red1 have been deposited to the BMRB under the accession code [50680](#). Coordinates and structure factors for the PlaI-Red1 complex, PlaI_{FL} and PlaI_{Δ14} have been deposited in the PDB with accession codes [7Q72](#), [7Q73](#) and [7Q74](#), respectively. SAXS data for PlaI-Red1 complex have been deposited to the SASBDB with accession code [SASDKE6](#). Genome-wide datasets are deposited in NCBI GEO under the reference number GEO: [GSE206106](#). Source data are provided with this paper.

Code availability

Custom R and Bash scripts that were used to analyse the data is available at <https://github.com/ahorvath/Soni-et-al.-Nat.-Commun.-2022>.

References

- Kilchert, C., Wittmann, S. & Vasiljeva, L. The regulation and functions of the nuclear RNA exosome complex. *Nat. Rev. Mol. Cell Biol.* **17**, 227–239 (2016).
- Mitchell, P. Exosome substrate targeting: the long and short of it. *Biochem Soc. Trans.* **42**, 1129–1134 (2014).
- Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. & Tollervey, D. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases. *Cell* **91**, 457–466 (1997).
- Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845–1848 (2008).
- Xu, Z. et al. Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457**, 1033–1037 (2009).
- Neil, H. et al. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* **457**, 1038–1042 (2009).
- Porrua, O. & Libri, D. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat. Rev. Mol. Cell Biol.* **16**, 190–202 (2015).
- Davis, C. A. & Ares, M. Jr. Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **103**, 3262–3267 (2006).
- Houalla, R. et al. Microarray detection of novel nuclear RNA substrates for the exosome. *Yeast* **23**, 439–454 (2006).
- Zinder, J. C. & Lima, C. D. Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. *Genes Dev.* **31**, 88–100 (2017).
- Schneider, C. & Tollervey, D. Threading the barrel of the RNA exosome. *Trends Biochem. Sci.* **38**, 485–493 (2013).
- Jia, H. et al. The RNA helicase Mtr4p modulates polyadenylation in the TRAMP complex. *Cell* **145**, 890–901 (2011).
- Wlotzka, W., Kudla, G., Granneman, S. & Tollervey, D. The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *EMBO J.* **30**, 1790–1803 (2011).
- Jackson, R. N. et al. The crystal structure of Mtr4 reveals a novel arch domain required for rRNA processing. *EMBO J.* **29**, 2205–2216 (2010).
- Callahan, K. P. & Butler, J. S. TRAMP complex enhances RNA degradation by the nuclear exosome component Rrp6. *J. Biol. Chem.* **285**, 3540–3547 (2010).
- Ogami, K., Chen, Y. & Manley, J. L. RNA surveillance by the nuclear RNA exosome: mechanisms and significance. *Noncoding RNA* **4**, 8 (2018).
- LaCava, J. et al. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121**, 713–724 (2005).
- Wyers, F. et al. Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**, 725–737 (2005).
- Mohanty, B. K. & Kushner, S. R. Bacterial/archaeal/organelle polyadenylation. *Wiley Interdiscip. Rev. RNA* **2**, 256–276 (2011).
- Eckmann, C. R., Rammelt, C. & Wahle, E. Control of poly(A) tail length. *Wiley Interdiscip. Rev. RNA* **2**, 348–361 (2011).
- Anderson, J. T. & Wang, X. Nuclear RNA surveillance: no sign of substrates tailing off. *Crit. Rev. Biochem. Mol. Biol.* **44**, 16–24 (2009).
- Tudek, A., Lloret-Llinares, M. & Jensen, T. H. The multitasking polyA tail: nuclear RNA maturation, degradation and export. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **373**, 20180169 (2018).
- Lubas, M. et al. Interaction profiling identifies the human nuclear exosome targeting complex. *Mol. Cell* **43**, 624–637 (2011).
- Meola, N. et al. Identification of a nuclear exosome decay pathway for processed transcripts. *Mol. Cell* **64**, 520–533 (2016).
- Lee, N. N. et al. Mtr4-like protein coordinates nuclear RNA processing for heterochromatin assembly and for telomere maintenance. *Cell* **155**, 1061–1074 (2013).
- Zhou, Y. et al. The fission yeast MTREC complex targets CUTs and unspliced pre-mRNAs to the nuclear exosome. *Nat. Commun.* **6**, 7050 (2015).
- Egan, E. D., Braun, C. R., Gygi, S. P. & Moazed, D. Post-transcriptional regulation of meiotic genes by a nuclear RNA silencing complex. *RNA* **20**, 867–881 (2014).
- Silla, T., Karadoulama, E., Makosa, D., Lubas, M. & Jensen, T. H. The RNA exosome adaptor ZFC3H1 functionally competes with nuclear export activity to retain target transcripts. *Cell Rep.* **23**, 2199–2210 (2018).
- Dobrev, N. et al. The zinc-finger protein Red1 orchestrates MTREC submodules and binds the Mtl1 helicase arch domain. *Nat. Commun.* **12**, 3456 (2021).
- Foucher, A. E. et al. Structural analysis of Red1 as a conserved scaffold of the RNA-targeting MTREC/PAXT complex. *Nat. Commun.* **13**, 4969 (2022).
- Christofori, G. & Keller, W. 3' cleavage and polyadenylation of mRNA precursors in vitro requires a poly(A) polymerase, a cleavage factor, and a snRNP. *Cell* **54**, 875–889 (1988).
- Takagaki, Y., Ryner, L. C. & Manley, J. L. Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mRNA polyadenylation. *Cell* **52**, 731–742 (1988).
- Gilmartin, G. M. & Nevins, J. R. An ordered pathway of assembly of components required for polyadenylation site recognition and processing. *Genes Dev.* **3**, 2180–2190 (1989).
- Yamanaka, S., Yamashita, A., Harigaya, Y., Iwata, R. & Yamamoto, M. Importance of polyadenylation in the selective elimination of meiotic mRNAs in growing *S. pombe* cells. *EMBO J.* **29**, 2173–2181 (2010).
- Chen, H. M., Futcher, B. & Leatherwood, J. The fission yeast RNA binding protein Mmi1 regulates meiotic genes by controlling intron specific splicing and polyadenylation coupled RNA turnover. *PLoS ONE* **6**, e26804 (2011).
- Sugiyama, T. & Sugioka-Sugiyama, R. Red1 promotes the elimination of meiosis-specific mRNAs in vegetatively growing fission yeast. *EMBO J.* **30**, 1027–1039 (2011).
- Yamashita, A. et al. Hexanucleotide motifs mediate recruitment of the RNA elimination machinery to silent meiotic genes. *Open Biol.* **2**, 120014 (2012).
- Bard, J. et al. Structure of yeast poly(A) polymerase alone and in complex with 3'-dATP. *Science* **289**, 1346–1349 (2000).

39. Martin, G., Keller, W. & Doublié, S. Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. *EMBO J.* **19**, 4193–4203 (2000).
40. Daubner, G. M., Clery, A. & Allain, F. H. RRM-RNA recognition: NMR or crystallography...and new findings. *Curr. Opin. Struct. Biol.* **23**, 100–108 (2013).
41. Balbo, P. B. & Bohm, A. Mechanism of poly(A) polymerase: structure of the enzyme-MgATP-RNA ternary complex and kinetic analysis. *Structure* **15**, 1117–1131 (2007).
42. Balbo, P. B., Meinke, G. & Bohm, A. Kinetic studies of yeast poly(A) polymerase indicate an induced fit mechanism for nucleotide specificity. *Biochemistry* **44**, 7777–7786 (2005).
43. Balbo, P. B., Toth, J. & Bohm, A. X-ray crystallographic and steady state fluorescence characterization of the protein dynamics of yeast polyadenylate polymerase. *J. Mol. Biol.* **366**, 1401–1415 (2007).
44. Kielkopf, C. L., Lucke, S. & Green, M. R. U2AF homology motifs: protein recognition in the RRM world. *Genes Dev.* **18**, 1513–1526 (2004).
45. Helmling, S., Zhelkovsky, A. & Moore, C. L. Fip1 regulates the activity of Poly(A) polymerase through multiple interactions. *Mol. Cell Biol.* **21**, 2026–2037 (2001).
46. Preker, P. J., Lingner, J., Minvielle-Sebastia, L. & Keller, W. The FIP1 gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase. *Cell* **81**, 379–389 (1995).
47. Forbes, K. P., Addepalli, B. & Hunt, A. G. An Arabidopsis Fip1 homolog interacts with RNA and provides conceptual links with a number of other polyadenylation factor subunits. *J. Biol. Chem.* **281**, 176–186 (2006).
48. Kaufmann, I., Martin, G., Friedlein, A., Langen, H. & Keller, W. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J.* **23**, 616–626 (2004).
49. Pedersen, L. C., Benning, M. M. & Holden, H. M. Structural investigation of the antibiotic and ATP-binding sites in kanamycin nucleotidyltransferase. *Biochemistry* **34**, 13305–13311 (1995).
50. Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H. & Kraut, J. Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. *Science* **264**, 1891–1903 (1994).
51. Remaut, H. & Waksman, G. Protein-protein interaction through beta-strand addition. *Trends Biochem. Sci.* **31**, 436–444 (2006).
52. Petoukhov, M. V. et al. New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).
53. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
54. Kuhn, U., Buschmann, J. & Wahle, E. The nuclear poly(A) binding protein of mammals, but not of fission yeast, participates in mRNA polyadenylation. *RNA* **23**, 473–482 (2017).
55. Zhelkovsky, A., Helmling, S. & Moore, C. Processivity of the *Saccharomyces cerevisiae* poly(A) polymerase requires interactions at the carboxyl-terminal RNA binding domain. *Mol. Cell Biol.* **18**, 5942–5951 (1998).
56. Krause, M. et al. tailfindr: alignment-free poly(A) length measurement for Oxford Nanopore RNA and DNA sequencing. *RNA* **25**, 1229–1241 (2019).
57. Zhao, J., Kessler, M. M. & Moore, C. L. Cleavage factor II of *Saccharomyces cerevisiae* contains homologues to subunits of the mammalian Cleavage/ polyadenylation specificity factor and exhibits sequence-specific, ATP-dependent interaction with precursor RNA. *J. Biol. Chem.* **272**, 10831–10838 (1997).
58. Manley, J. L. A complex protein assembly catalyzes polyadenylation of mRNA precursors. *Curr. Opin. Genet. Dev.* **5**, 222–228 (1995).
59. Preker, P. J., Ohnacker, M., Minvielle-Sebastia, L. & Keller, W. 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**, 4727–4737 (1997).
60. Casanal, A. et al. Architecture of eukaryotic mRNA 3'-end processing machinery. *Science* **358**, 1056–1059 (2017).
61. Meinke, G. et al. Structure of yeast poly(A) polymerase in complex with a peptide from Fip1, an intrinsically disordered protein. *Biochemistry* **47**, 6859–6869 (2008).
62. Ezeokonkwo, C., Zhelkovsky, A., Lee, R., Bohm, A. & Moore, C. L. A flexible linker region in Fip1 is needed for efficient mRNA polyadenylation. *RNA* **17**, 652–664 (2011).
63. Kumar, A. et al. Dynamics in Fip1 regulate eukaryotic mRNA 3' end processing. *Genes Dev.* **35**, 1510–1526 (2021).
64. Vo, T. V. et al. CPF recruitment to non-canonical transcription termination sites triggers heterochromatin assembly and gene silencing. *Cell Rep.* **28**, 267–281.e5 (2019).
65. Lee, S. Y. et al. Dense transposon integration reveals essential cleavage and polyadenylation factors promote heterochromatin formation. *Cell Rep.* **30**, 2686–2698.e8 (2020).
66. Zofall, M. et al. RNA elimination machinery targeting meiotic mRNAs promotes facultative heterochromatin formation. *Science* **335**, 96–100 (2012).
67. Yamanaka, S. et al. RNAi triggered by specialized machinery silences developmental genes and retrotransposons. *Nature* **493**, 557–560 (2013).
68. Hiriart, E. et al. Mmi1 RNA surveillance machinery directs RNAi complex RITS to specific meiotic genes in fission yeast. *EMBO J.* **31**, 2296–2308 (2012).
69. Tashiro, S., Asano, T., Kanoh, J. & Ishikawa, F. Transcription-induced chromatin association of RNA surveillance factors mediates facultative heterochromatin formation in fission yeast. *Genes Cells* **18**, 327–339 (2013).
70. Montanes, J. C. et al. Native RNA sequencing in fission yeast reveals frequent alternative splicing isoforms. *Genome Res.* **32**, 1215–1227 (2022).
71. St-Andre, O. et al. Negative regulation of meiotic gene expression by the nuclear poly(a)-binding protein in fission yeast. *J. Biol. Chem.* **285**, 27859–27868 (2010).
72. Makino, D. L., Baumgartner, M. & Conti, E. Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. *Nature* **495**, 70–75 (2013).
73. Mangus, D. A., Evans, M. C. & Jacobson, A. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.* **4**, 223 (2003).
74. Kerwitz, Y. et al. Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *EMBO J.* **22**, 3705–3714 (2003).
75. Wahle, E. A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell* **66**, 759–768 (1991).
76. Lemieux, C. et al. A Pre-mRNA degradation pathway that selectively targets intron-containing genes requires the nuclear poly(A)-binding protein. *Mol. Cell* **44**, 108–119 (2011).
77. Bresson, S. M. & Conrad, N. K. The human nuclear poly(a)-binding protein promotes RNA hyperadenylation and decay. *PLoS Genet.* **9**, e1003893 (2013).
78. Tuck, A. C. & Tollervey, D. A transcriptome-wide atlas of RNP composition reveals diverse classes of mRNAs and lncRNAs. *Cell* **154**, 996–1009 (2013).
79. Kuehner, J. N. & Brow, D. A. Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation. *Mol. Cell* **31**, 201–211 (2008).

80. Gonzalez, C. I., Ruiz-Echevarria, M. J., Vasudevan, S., Henry, M. F. & Peltz, S. W. The yeast hnRNP-like protein Hrp1/Nab4 marks a transcript for nonsense-mediated mRNA decay. *Mol. Cell* **5**, 489–499 (2000).
81. Kessler, M. M. et al. Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast. *Genes Dev.* **11**, 2545–2556 (1997).
82. Touat-Todeschini, L. et al. Selective termination of lncRNA transcription promotes heterochromatin silencing and cell differentiation. *EMBO J.* **36**, 2626–2641 (2017).
83. Harigaya, Y. et al. Selective elimination of messenger RNA prevents an incidence of untimely meiosis. *Nature* **442**, 45–50 (2006).
84. Garland, W. et al. A functional link between nuclear RNA decay and transcriptional control mediated by the polycomb repressive complex 2. *Cell Rep.* **29**, 1800–1811.e6 (2019).
85. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
86. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
87. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
88. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
89. Tickle, I. J. F. et al. (<http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi>). (Global Phasing Ltd., 2018).
90. Sattler, M., Schleucher, J. & Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 93–158 (1999).
91. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
92. Vranken, W. F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
93. Konarev, P., Volkov, V., Sokolova, A., Koch, M. & Svergun, D. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282 (2003).
94. Franke, D. et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212–1225 (2017).
95. Moggridge, S., Sorensen, P. H., Morin, G. B. & Hughes, C. S. Extending the compatibility of the SP3 paramagnetic bead processing approach for proteomics. *J. Proteome Res.* **17**, 1730–1740 (2018).
96. Hughes, C. S. et al. Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* **10**, 757 (2014).
97. Werner, T. et al. Ion coalescence of neutron encoded TMT 10-plex reporter ions. *Anal. Chem.* **86**, 3594–3601 (2014).
98. Sridharan, S. et al. Proteome-wide solubility and thermal stability profiling reveals distinct regulatory roles for ATP. *Nat. Commun.* **10**, 1155 (2019).
99. Franken, H. et al. Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat. Protoc.* **10**, 1567–1593 (2015).
100. Savitski, M. M., Wilhelm, M., Hahne, H., Kuster, B. & Bantscheff, M. A scalable approach for protein false discovery rate estimation in large proteomic data sets. *Mol. Cell Proteom.* **14**, 2394–2404 (2015).
101. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
102. Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A. & Vingron, M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18**, S96–S104 (2002).
103. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
104. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
105. Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
106. Lawrence, M. et al. Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).
107. Morgan, M., Pagès, H., Obenchain, V. & Hayden, N. Rsamtools: Binary alignment (BAM), FASTA, variant call (BCF), and tabix file import (2022).
108. Wickham, H., François, R., Henry, L. & Müller, K. *dplyr: A Grammar of Data Manipulation. R package version* (2018).
109. Andrews, S. FastQC: A Quality Control Tool For High Throughput Sequence Data (2010).
110. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863–864 (2011).
111. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
112. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **1303**, 3997v2 (2013).
113. Lawrence, M., Gentleman, R. & Carey, V. tracklayer: an R package for interfacing with genome browsers. *Bioinformatics* **25**, 1841–1842 (2009).

Acknowledgements

We thank Jürgen Kopp and Claudia Siegmann from the BZH/Cluster of Excellence: CellNetworks crystallization platform and acknowledge access to the beamlines P12, P13 and P14 at PETRA III at DESY in Hamburg and the support of the beamline scientists. We thank Per Haberkant and Frank Stein at the Proteomics Core Facility (PCF), EMBL Heidelberg for mass spectrometry analysis. This work was supported by a DAAD fellowship to N.D., the Deutsche Forschungsgemeinschaft (DFG) through the Leibniz program (SI 586/6-1) and TRR 319 (Project-ID 439669440, TP B03) to I.S. and the Australian Research Council's Discovery Projects funding scheme to T.F. (project DP190100423).

Author contributions

K.S., A.S., T.F. and I.S. designed the study and interpreted the results. K.S. performed NMR, crystallographic structure determination, SAXS, in vitro polyadenylation assay, biophysical characterisation using pull-down assays, ITC, BLI and Y2H. N.D. and L.K. performed Y2H and NMR experiments. A.S. performed TAP purification followed by interaction analysis and mass spectrometry, Illumina sequencing of total, poly(A)⁺ and RIP-ed RNA, Direct RNA sequencing by Nanopore of poly(A)⁺ and RIP-ed RNA, ChIP seq. experiments. R.H. supported Nanopore sequencing experiments. A.H. and T.F. performed the bioinformatics analyses. B.S. recorded NMR experiments. K.W. supported crystallographic analysis. K.S., I.S. and T.F. wrote the manuscript. All authors contributed to the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-36402-6>.

Correspondence and requests for materials should be addressed to Irmgard Sinning or Tamás. Fischer.

Peer review information *Nature Communications* thanks the anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023