CEG1 DEPLETION REVEALS MECHANISMS GOVERNING DEGRADATION OF NON-CAPPED RNAs

Onofrio Zanin[1], Daniel Hebenstreit[2] and Pawel Grzechnik[3*]

[1]School of Biosciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

[2]School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK

[3]School of Biological Sciences, University of Manchester, Oxford Road, Manchester M13 9PT, UK

*Correspondence: pawel.grzechnik@manchester.ac.uk

ABSTRACT

All eukaryotic mRNAs contain a 7-methylguanosine (m7G) cap which serves as a platform that recruits proteins to support essential biological functions such as mRNA processing, nuclear export and cap-dependent translation. Although the capping is one of the first steps of transcription and uncapped mRNA is not functional, the fate and turnover of uncapped transcripts have not been studied extensively. Here, we employed fast nuclear depletion of the capping enzymes in *Saccharomyces cerevisiae* to uncover the turnover of transcripts that failed to be capped. We found that the degradation of non-capped mRNA is mainly performed by Xrn1, the exonuclease which is predominantly localized in the cytoplasm, and the fate of such transcripts is determined principally by the abundance of their synthesis. Nuclear depletion of capping enzymes increased the levels of lowly expressed mRNAs as well as non-coding RNAs and did not affect the distribution of RNA Polymerase II on chromatin. Overall, our data indicate that mRNAs that failed to be capped are not directed to specific quality-control pathway and are stochastically degraded.

INTRODUCTION

The m7G 5′ cap, N7-methylated guanosine linked to the first nucleotide of the RNA molecule, is a hallmark structure of eukaryotic transcripts generated by RNA Polymerase II (Pol II). The cap is essential for cell viability as it protects RNA molecules from 5' to 3' exonucleolytic degradation as well as is necessary for various steps of RNA processing, RNA transport and translation (Schwer, Mao and Shuman, 1998; Furuichi, 2015). Many of these biological functions are related to the cap-binding complex (CBC) which consists of Cbc80 and Cbc20 and binds to the m7G caps of nascent RNA (Izaurralde et al., 1995; Izaurralde et al., 1994). Although the CBC is involved in multiple processes including transcription elongation, mRNA splicing, polyadenylation, nuclear export and translation (Gonatopoulos-Pournatzis and Cowling, 2014), the complex is not essential in *Saccharomyces cerevisiae*.

The biosynthesis of the m7G cap requires three enzymes: RNA triphosphatase (TPase) which removes the γ-phosphate from the RNA 5′ end; RNA guanylyltransferase (GTase) which transfers a GMP group to the 5′ end; and guanine-N7 methyltransferase adding a methyl group to the N7 amine of the guanine cap (Decroly *et al.*, 2012). This synthesis pathway is conserved in eukaryotes, however, the enzymes catalysing capping vary between the organisms (Ramanathan, Robb and Chan, 2016). In humans and other higher eukaryotes, a bifunctional protein RNGTT acts as RNA guanylyltransferase and 5′-triphosphatase while methylation is mediated by a separate protein RNMT (Varshney *et al.*, 2018). In *S. cerevisiae* there are three proteins responsible for capping: guanylyltransferase Ceg1 and 5′-triphosphatase Cet1 work together as a stable heterodimer or heterotetramer, while the methyltransferase Abd1 acts independently, altogether resembling the configuration in metazoans (Schroeder *et al.*, 2000; Hausmann, Pei and Shuman, 2003; Martinez-

Rucobo *et al.*, 2015). The capping enzymes are recruited to Pol II at the early stage of transcription. In *S. cerevisiae*, Cet1 and Ceg1 bind the serine 5 phosphorylated C-terminal domain (CTD) of Pol II when the transcribing complex is near the transcription start site (TSS) (Schroeder *et al.*, 2000; Lidschreiber, Leike and Cramer, 2013). When the nascent RNA reaches a length of ~17 nt, Ceg1-Cet1 complex docks to the Pol II surface close to the RNA exit tunnel (Martinez-Rucobo *et al.*, 2015) and adds the terminal guanosine as soon as the first 25–30 nt of the nascent transcript extrude from the transcribing complex (Shatkin and Manley, 2000). Changes in phosphorylation of Pol II CTD promoting transcription elongation results in rapid dissociation of Ceg1-Cet1 (Martinez-Rucobo *et al.*, 2015). This and the increasing phosphorylation of CTD Ser2 promote the binding of Abd1, which is recruited about 110 nt downstream of the TSS (Lidschreiber, Leike and Cramer, 2013) and completes the formation of the m7G cap.

If the methylation of the added 5' guanosine residue fails, the defective transcripts are detected by a surveillance mechanism that triggers decapping by Rai1 and subsequent RNA degradation by its co-factor, the 5'-3' exonuclease Rat1 (Jiao et al., 2010). The complete lack of the cap structure on the 5' end may also lead to co-transcriptional degradation of the nascent transcript. This process naturally occurs on the 3' ends of genes and is essential for transcription termination. The cleavage over the poly(A) signal in the RNA creates an unprotected 5' monophosphate RNA end in the nascent RNA. This serves as an entry point for Rat1 which co-transcriptionally degrades RNA pursuing Pol II transcribing downstream of the gene (Kim *et al.*, 2004; West, Gromak and Proudfoot, 2004; Eaton *et al.*, 2020). If the degradation rate exceeds the RNA synthesis rate, Rat1 reaches the transcribing complex and Pol II is displaced from DNA (Cortazar et al., 2019; Eaton, Francis, Davidson, & West, 2020;

Kim et al., 2004; West, Gromak, & Proudfoot, 2004). A similar mechanism has been suggested to be active upstream to the transcription end site (TES) when the capping fails at the 5' ends of the genes. Analysis of the *FMP27* gene in *S. cerevisiae* revealed that transcription of uncapped pre-mRNA is prematurely terminated by Rat1 within the gene body (Jimeno-González 2010). Moreover, analysis of the temperature-sensitive mutants revealed that inactivation of Ceg1 affects the accumulation of heat-shock induced *SSA1* and *SSA4* mRNAs. However, their levels were restored by the deletion of the predominantly cytoplasmic 5'-3' exonuclease Xrn1 (Schwer, Mao and Shuman, 1998).

Here, we investigated the consequences of defective capping on RNA accumulation and transcription. We took advantage of the yeast model where, in contrast to metazoans, the function of the CBC is not essential (Fortes *et al.*, 1999) which presents a unique opportunity to distinguish the functions of m7G cap and CBC. We employed rapid nuclear depletion of Ceg1 and Cet1 to investigate the direct effects of the lack of the capping on the transcriptome. We found that the absence of the capping complex resulted in differential accumulation of RNA classes. The mRNA levels of highly expressed genes decreased, while on the contrary, mRNA of genes expressed at low levels increased upon the nuclear depletion of the capping machinery. Our data indicate that the degradation of non-capped mRNA is not determined by quality-control mechanisms in the nucleus. The lack of 5' end capping had a minor effect on the transcription and distribution of Pol II on the chromatin confirming that co-transcriptional RNA degradation must be accompanied by additional signals enforcing premature transcription termination.

RESULTS

Differential accumulation of non-capped mRNAs

The impact of defective capping on RNA accumulation has never been studied genome-wide before. In our experimental approach, we employed the anchor-away (AA) system to deplete the capping enzymes from the nucleus to rapidly affect the formation of the cap. FRB tag fused with Cet1, Ceg1 or Abd1 (*cet1-AA*, *ceg1-AA* and *abd1-AA* strains) in the presence of rapamycin heterodimerizes with FKB12 attached to the ribosomal protein Rpl13a, which is exported to the cytoplasm during ribosome biogenesis, and thus removes the capping enzymes from the nucleus (Haruki, Nishikawa and Laemmli, 2008). Nuclear depletion of Cet1 and Ceg1, was lethal (Fig. 1A) indicating the efficacy of the method as cap formation is essential for cell viability (Tsukamoto *et al.*, 1997; Chu and Shatkin, 2008). The inactivation of Abd1 activity has been shown to be lethal (Mao, Schwer and Shuman, 1996), however, the nuclear depletion of Abd1 had no impact on cell survival and only slightly affected the abundance of *ADH1* mRNA (Fig. S1A). This indicates that Abd1 may be still functional after the relocation to the cytoplasm by the anchor away system. The nuclear depletion of either Cbp20 or Cbp80 was viable, consistent with the fact that CBC is not essential in *S. cerevisiae* (Fig. 1A).

In our functional experiments, the same yeast culture was split into a medium containing either DMSO or rapamycin and represented WT (DMSO) or nuclear depletion (Rap), respectively (Fig. 1B) in the subsequent analyses. First, we confirmed that RNA capping was indeed affected by the nuclear depletion of Ceg1 (Fig. 1C). We co-immunoprecipitated RNA isolated from *ceg1-AA* strain growing on either DMSO or Rap with anti m7G cap antibody (H-20). The amount of *ADH1* mRNA (normalised to human *GAPDH* spike) was reduced by 80% in *ceg1-AA* (Rap) when compared to the

control cells grown on DMSO. The nuclear depletion of either Ceg1 or Cet1 resulted in a similar decrease of *ADH1* mRNA (Fig. 1D and Fig. S1B) and therefore, in the subsequent analyses, we focused mainly on the effects of Ceg1 removal. We also confirmed that the nuclear depletion of the CBC did not cause the decrease of *ADH1* mRNA (Fig. 1D and Fig. S1B).

Next, we performed RNA sequencing to investigate mRNA levels genome-wide after Ceg1 nuclear depletion. We chose to deplete Ceg1 for 45 minutes (min) as this time point was already sufficient to decrease the level of *ADH1* (Fig. S1B, 1D and 1F). Our RNA-seq analysis identified 1205 mRNAs affected by the nuclear depletion of Ceg1 (displaying log2 fold change < -1 and > 1) (Fig. 1E). 736 (61%) of these mRNAs decreased when compared to the control, reflecting the fraction of non-capped transcripts that were subjected to degradation. Unexpectedly, the levels of 469 genes (39%) increased (Fig. 1E). Both effects were on average similar to each other (log2 fold change -1.4 and +1.5, respectively) (Fig. S1C). We also tested how the levels of mRNA changed over time following nuclear depletion of Ceg1 (Fig. 1F). We selected *ADH1*, *PYK2* and *PMA1* genes whose levels decreased upon Ceg1 depletion in the RNA-seq data. RT-qPCR analysis revealed that their mRNA decreased even further after 90 min and stabilised after 120 min. We also tested the levels of *PML39*, *STP2* and *CTA1* which increased upon Ceg1 depletion in the RNA-seq data. *PML39* and *STP2* mRNAs peaked after 90 min, while the accumulation of *CTA1,* displayed a constant increase through time. Next, we tested if the differential levels of mRNAs in *ceg1-AA* (Rap) could be rescued by the deletion of Xrn1 exonuclease. Xrn1 is mainly associated with the cytoplasm; however, the enzyme shuttles between the cytoplasm and the nucleus and the Xrn1 mutants were shown to rescue the levels of heat-shock induced uncapped mRNAs in *CEG1* temperature-sensitive mutants (Schwer, Mao and

Shuman, 1998). Thus, we introduced an *XRN1* deletion into the *ceg1-AA* strain and performed the RNA-seq analysis on the cells growing on either DMSO or Rap for 45 min. The *XRN1* deletion rescued the decreased mRNA levels (Fig. 1G, as we detected only 18 depleted mRNAs, 13 of which were reduced in *ceg1-AA* (Rap). Moreover, the deletion of Xrn1 exonuclease restored the levels of mRNAs that increased in *ceg1-AA* (Rap); we detected only 23 increased mRNAs 10 of which were upregulated in the *ceg1-AA* (Rap). Notably, the *ceg1-AA/xrn1Δ* strain is non-viable on Rap-containing medium (Fig. S1D), underlining the importance of the m7G cap for the mRNA functionality.

Accumulation of non-capped mRNAs depends on expression level

To understand the differential accumulation of the mRNAs upon depletion of the capping machinery we analysed the features of the identified affected mRNAs. We split all of 1205 protein-coding genes identified as differentially expressed in *ceg1-AA* into three bins according to their expression levels in WT cells *ceg1-AA* (DMSO): the top 25% were classified as highly expressed (High), the bottom 25% as lowly expressed (Low) and the rest 50% as medially expressed (Mid). The majority of lowly expressed genes increased their mRNA levels (291 out of 301), while most of highly expressed mRNAs were decreased (300 out of 301) after the Ceg1 depletion. Mid-range expressed genes were the least affected by the depletion; however, they displayed the tendency to decrease the levels in *ceg1-AA* (Rap) (426 decreased versus 176 increased), correlating with increasing expression (Fig. 2B-C). Overall, we observed a progressive decrease of the log2 fold change with the increase in the basal level of gene expression. The correlation between the expression levels and RNA accumulation in capping defective mutant was also reflected on a single gene level as

shown for *ZRT1*, *STB4* and *ADH1* (Fig. 2D). The mRNAs which accumulated in *ceg1-AA* (Rap) were also longer than those from the decreased set (average gene length 1430 vs 1113 nt) (Fig. 2E and S2A) consistently with the fact that shorter genes are generally associated with higher expression (Urrutia and Hurst, 2003). We also employed available datasets estimating *S. cerevisiae* mRNA half-life (Geisberg et al., 2015), transcription rates (TR) and Pol II density (Pelechano, Wei and Steinmetz, 2013) and applied them to our sets of affected mRNAs identified in the *ceg1-AA* (Rap) strain. These analyses revealed that mRNAs which increased upon Ceg1 nuclear depletion are generally more stable and display longer half-lives (average half-life 41.9 vs 30.7 min) (Fig. 2F and S2B), they also have lower transcription rates (Fig. 2G and S2C) and Pol II density (Fig. S2D) than the genes whose mRNAs decreased in *ceg1-AA* (Rap). Next, we investigated if decreasing RNA synthesis would alleviate the reduction of *ADH1* mRNA observed upon Ceg1 nuclear depletion. We treated cells for 30 min with the transcription elongation inhibitor 6-azouracil (6-AU), followed by the addition of Rap to the medium. We measured the levels of *ADH1* in *ceg1-AA* (Rap) growing on 6-AU or the control medium and compared them to levels in *ceg1-AA* (DMSO). This analysis revealed that slower synthesis stabilised *ADH1* mRNA levels (Fig. 2H).

Previous reports suggested that the decrease of mRNA caused by capping defects can be restored by mutation of Rat1 exonuclease (Jimeno-González *et al.*, 2010). Thus, we employed *ceg1-63* and *ceg1-63/rat1-1* temperature sensitive mutants to assess the levels of highly expressed *ADH1* and lowly expressed *FMP27* previously used to study capping-dependent RNA degradation (Jimeno-González *et al.*, 2010). However, we did not observe that the Rat1 inactivation rescued the mRNA levels in our experimental setup (Fig. 2I). *FMP27* mRNA, similarly to other lowly expressed

genes, did not decrease but remains stable in *ceg1-AA* (Rap). These variations between datasets may be caused by the difference in using ts mutants versus rapid nuclear depletion as in the latter it is more likely to observe direct effects. We also noticed that *FMP27* level was affected already by shifting WT cells to 37° C which may contribute to the effects observed in *ceg1-63*. However, we cannot exclude that *FMP27* mRNA decreases in *ceg1-63* and is rescued by Rat1 inactivation in sub-permissive temperature (34° C) as previously described (Jimeno-González *et al.*, 2010).

Finally, we performed gene ontology (GO) analysis to identify the functions of the genes affected by Ceg1 depletion. Interestingly, some upregulated mRNAs (45 out of the 469 tested) were associated with genes coding for DNA binding proteins directly associated with transcription (Fig. S3A), along with the zinc finger transcriptional factor *PPR1,* the regulator of stress response *STB5* and also *PCF11,* a component of the 3' end processing factor (Fig. S3B). GO terms analysis did not reveal consistent terms emerging from the genes whose mRNAs decreased in *ceg1-AA* (Rap) (Fig. S3C).

Non-coding RNAs highly accumulate in capping-deficient cells

Next, we investigated Pol II-transcribed non-coding RNAs (ncRNAs), called cryptic unstable transcripts (CUTs) and stable unannotated transcripts (SUTs), which in normal conditions are expressed at a very low level in cells (Wyers *et al.*, 2005; Marquardt, Hazelbaker and Buratowski, 2011). Most of these transcripts are detectable only upon inactivation of the 3'-5' exonuclease Rrp6 subunit of the nuclear exosome (Xu *et al.*, 2009) and, probably for this reason, only a handful of CUTs and SUTs were detected in our RNA-seq analysis of the *ceg1-AA* (Rap) strain. Thus, we focused our investigation on the well-studied ncRNA *CUT542* (*NEL025c*). The nuclear

depletion of Ceg1 increased the levels of *NEL025c* (Fig. 3A) up to 10 times after 120 min of rapamycin treatment, consistent with the fact that lowly expressed Pol II transcripts accumulate when capping enzymes are not functional and in place. A similar effect was also observed for *SUT238* (Fig. 3A) which suggests that this may be a general feature for both CUTs and SUTs. The excessive accumulation of *NEL025c* in *ceg1-AA* (Rap) was tempered by the deletion of Rrp6 exonuclease (Fig. 3B) which substantially elevated the ncRNA basal levels (~7x) compared to WT (Fig. 3C). The deletion of Xrn1 also suppressed the accumulation (Fig. 3B), although the overall level of *NEL025c* in *ceg1-AA/ xrn1Δ* (DMSO) was twice higher than the one observed in the WT (*ceg1-AA* (DMSO)) (Fig. 3B-C). This is consistent with the pattern observed for the degradation of uncapped mRNAs, further suggesting that this process may be non-specific and relies on the efficiency and availability of nucleases. Since *NEL025c* ncRNA is strictly nuclear, we tested if the inactivation of nuclear 5'-3' exonuclease Rat1 would exacerbate its accumulation upon inactivation of cap synthesis. We employed a temperature-sensitive mutant *ceg1-63/rat1-1* and compared the accumulation of *NEL025c* to *ceg1-63* and the parental WT strain shifted to non-permissive temperature (37° C) up to 120 min. The *rat1-1* mutation did not aggravate the accumulation of *NEL025c* which indicates that Rat1 was not involved in the degradation of this transcript (Fig. 3D). Next, we investigated if a reduction of RNA synthesis would increase the accumulation of *NEL025c* upon Ceg1 nuclear depletion. As described above, we treated cells with 6-azouracil (6-AU) for 30 min followed by the addition of Rap to the medium. We measured the levels of *NEL025c* in *ceg1-AA* (Rap) growing on 6-AU or the control medium and compared them to levels in *ceg1-AA* (DMSO). However, the decrease in transcription did not affect the accumulation of *NEL025c* (Fig. 3E) which suggests that the low expression of this gene reached the

threshold where further decrease had minimal impact on its accumulation. Overall, we concluded that both nuclear localization which makes ncRNAs less available for Xrn1 as well as low expression of ncRNAs contribute to the accumulation of these transcripts in cap-deficient cells.

Non-capped RNAs are not quality controlled in the nucleus

The accumulation of lowly expressed mRNA in *ceg1-AA* (Rap) suggested that Pol II transcripts may not be quality controlled for the presence of the cap at the 5' end. Since most highly expressed mRNAs are quickly degraded we tested the levels of small nucleolar RNAs (snoRNAs) which are also highly expressed and transcribed by Pol II (Dieci, Preti and Montanini, 2009). SnoRNAs form secondary structures which prevent from exonucleolytic degradation. Pre-snoRNA are initially capped, however the cap is removed in the later stages which is required for their correct processing and localization (Grzechnik *et al.*, 2018). We did not detect any significant differences in snoRNA accumulation upon Ceg1 nuclear depletion (Fig. 4A). This data indicates that highly transcribed RNAs, despite being uncapped, may accumulate normally if protected from degradation. Therefore, the importance of the m7G cap may lay in marking mRNAs for nuclear export and translation rather than protection from immediate degradation in the nucleus.

Previous studies demonstrated co-transcriptional mechanisms testing the quality of the m7G cap where non-maturated caps are removed and decapped RNA degraded. This pathway is mediated by decapping and exonucleolytic activities of the Rai1-Rat1 complex (Xiang *et al.*, 2009; Jiao *et al.*, 2010). Although Rat1 seems not to be involved in the degradation of steady state mRNA we could not rule out that it is engaged in co-transcriptional degradation. Thus, we tested if the degradation of nascent/newly

synthesised RNAs was enhanced in *ceg1-AA* (Rap) which would reflect the increased activity of co-transcriptionally associated Rat1. We employed *in vivo* RNA metabolic labelling with 4-thiouracil (4tU) to assess the levels of nascent/newly synthesised *ADH1*, *ACT1* and *NEL025C* RNAs. 4tU was added to the medium for 5 min after 40 min of Ceg1 nuclear depletion. For both abundant mRNAs (*ADH1* and *ACT1*) the decrease of 4tU labelled mRNAs was matched to the decrease in steady-state level, while the levels of the nascent/newly synthesised *NEL025c* mRNA was unaffected in the *ceg1-AA* (Rap) strain (Fig. 4B). This result indicates that either RNA synthesis rates decreased or co-transcriptional degradation depends on the availability of the transcripts as seen for degradation of steady-state RNAs. The observation that the levels of nascent/newly synthesised *NEL025C* did not change in *ceg1-AA* (Rap) suggests that co-transcriptional degradation was not enhanced for this uncapped ncRNA.

Finally, we tested if the presence of the cap at the early stages of transcription could trigger the degradation of lowly expressed mRNAs if the cap is removed later during transcription in wild-type cells. To create such an uncapped mRNA, we introduced the self-cleaving hepatitis δ ribozyme sequence into the 5' UTR of *FMP27* gene (*fmp27-RZ).* As a control we used the catalytically inactive ribozyme mutant harbouring the substitution C76:U (*fmp27-C76U*) (Fig. 4C) (Perrotta, Shih and Been, 1999; Bird *et al.*, 2005). *FMP27* levels did not change upon Ceg1 nuclear depletion suggesting that this mRNA is not degraded when it is non-capped (Fig. 4D). In the WT cells, the introduction of the ribozyme decreased the level *of fmp27-RZ* by 64% compared to the control *fmp27-C76U* containing the catalytically inactive ribozyme (Fig. 4E). This confirms that the presence of cap or the capping enzymes may be required for the quality control mechanisms. Moreover, unlike triphosphate 5' RNA ends (pppRNAs)

generated by Pol II, hydroxyl 5' RNA (OH-RNAs) ends generated by the ribozyme cleavage are substrate for the 5'-3' exonucleases Rat1 and Xrn1 (Doamekpor et al., 2020; Jinek, Coyle, & Doudna, 2014). Thus, the reduced levels of *fmp27-RZ* indicate that the generation of the correct substrate for exonucleases was required for the degradation and that the 5' end of non-capped RNAs may not be efficiently converted to become suitable for 5'-3' exonucleases in *ceg1-AA* (Rap).

Non-capped mRNAs are not prematurely terminated

Although efficient co-transcriptional degradation is required for termination, previous studies reported that shifting the temperature-sensitive mutant *ceg1-63* to sub-permissive temperature resulted in premature transcription termination of the 8 kb-long gene *FMP27* (Jimeno-González *et al.*, 2010). This effect was not fully recapitulated in the *ceg1-AA* strain. Our ChIP analysis using an antibody recognizing Pol II CTD revealed only a slight reduction of Pol II occupancy 0.5 kb from the transcription start site but not further downstream of the gene (Fig. 5A). Since the average gene length in yeast is 1.4 kb (Hurowitz and Brown, 2003)(Krebs, Goldstein & Kilpatrick, Lewin's Genes XI, 2014, Jones & Bartlett Learning), we investigated how transcription was affected by Ceg1 nuclear depletion on shorter genes (Fig 5B). We depleted Ceg1 for 45 min and analysed the distribution of Pol II phosphoisoforms on highly and lowly expressed *ADH1* and *PML39*, respectively (Fig. 5C-D). ChIP analysis revealed that total Pol II was partially depleted mainly on the 5' ends for both highly and lowly expressed genes. We used 8WG16 antibody recognising non-modified heptapeptides of the CTD to analyse total Pol II distribution. Thus, we speculate that this decrease of the Pol II signal detected by this antibody may reflect prematurely terminated unphosphorylated or hypophosphorylated Pol II. Indeed, both ser5-P and

ser2-P Pol II phosphoisoforms, that represent actively transcribing Pol II, were only marginally affected by Ceg1 depletion (Fig. 5C-D).

A similar but less pronounced effect was observed for *ADH1* gene in *cet1-AA* (Rap) (Fig. S4A) as both enzymes are essential for cap formation (Martinez-Rucobo *et al.*, 2015). Nuclear depletion of Abd1 did not affect the distribution of Pol II over *ADH1* (Fig. S4B). We also investigated if longer depletion of Ceg1 will result in increased premature transcription termination. We shifted cells to the medium containing Rap for up to 120 min and tested the levels of Pol II over the 3' end of high and mid-range expressed *ADH1*, *PYK1* and *PMA1* genes (Fig. 5E). Consequently, we observed only ~20% decrease in Pol II signal at 45 min which was not exacerbated by longer nuclear depletion of Ceg1.

DISCUSSION

Capping is the first co-transcriptional RNA modification defining mRNAs. The m7G cap is essential for the functionality of mRNAs and therefore, mutations affecting the cap formation are non-viable. Previous studies suggest that RNAs with defective caps (e.g. lacking m7G or having non-methylated caps) are rapidly degraded by exonucleases, however these analyses were performed only on few selected genes (Lewis and Izaurralde, 1997; Schwer, Mao and Shuman, 1998; Cougot *et al.*, 2004; Topisirovic *et al.*, 2011). Now, we performed a global analysis of the capping defective transcriptome. We show that although the overall mRNA decrease is a dominant effect following nuclear depletion of the capping enzymes, the levels of many mRNAs increase when they are not capped. Our data show clear dependence between the amount of RNA synthesised and the degradation of uncapped transcripts. We conclude that non-capped RNAs transcribed by Pol II are degraded by non-specific

pathways. Therefore, the RNAs displaying the highest levels are most likely to be reduced, probably in the cytoplasm. This is suggested by the strong increase of nuclear ncRNAs levels upon nuclear depletion of capping enzymes. The lack of a specific pathway targeting RNAs that are synthesised without the cap may aid the production of other essential RNA classes such as ribosomal RNA (rRNA) or transfer RNA (tRNA) produced by the RNA Pol I and III, respectively, that are not capped but stable in the nucleus and cytoplasm (Ramanathan, Robb and Chan, 2016; Akiyama, Eiler and Kieft, 2017). In these cases, the synthesis would not be affected by surveillance mechanisms degrading non-capped RNAs in the nucleus. However, releasing and maintaining non-functional mRNAs may have deleterious effects on the cell. Thus, we speculate that the binding of the capping enzymes to the transcriptional machinery or the synthesis of m7G cap itself, is the condition required for the cap quality control pathway to be functional and in place. Indeed, as we observed for the *fmp27-RZ* transcript, if mRNA is capped and the cap is removed during transcription, for example by a self-cleaving ribozyme, the mRNA is efficiently degraded, even if it is expressed at low level.

The Rai1-Rat1 complex was shown to be responsible for the degradation of the RNAs with capping defects. Rai1 shows both decapping and pyrophosphohydrolase activity generating in both cases a monophosphate RNA 5' end (pRNA) substrate of the cellular 5'-3' ribonucleases Rat1 and Xrn1 (Xiang *et al.*, 2009; Jiao *et al.*, 2010). The nuclear depletion of the Ceg1-Cet1 capping complex, and thus the complete abrogation of the m7G cap formation, may mark the nascent transcript as non-mRNA. A similar mechanism where the presence of the cap at the 5' end decides the biology of the transcript, has been observed for yeast box C/D snoRNAs which do not possess the m7G cap as mature species. However, when their 5' end processing is affected

and the cap retained at the 5' end, these snoRNAs are subjected to mRNA processing pathways, become polyadenylated at the 3' end as well as exported to the cytoplasm (Grzechnik *et al.*, 2018). Thus, the lack of capping in the mRNA synthesis may also determine the further fate of the transcript. This in turn may affect the activity or recruitment of the Rai1-Rat1 to the Pol II complex and therefore the efficient conversion of the RNA 5' end into the 5'-3' exonuclease substrate.

We did not observe a significant increase in premature transcription termination upon depletion of the capping enzymes. This contrasts with previous models suggesting that such RNAs are degraded co-transcriptionally by Rat1 eliciting the removal of Pol II from DNA by the "torpedo termination" mechanism, where 5'-3' exonuclease Rat1 attacks the 5' unprotected end, chases the Pol II by degrading the nascent RNA, and ultimately displaces the transcribing polymerase from the DNA template. However, the absence of premature transcription termination of non-capped RNAs is consistent with the possible mode of random degradation of these transcripts by Xrn1 proposed by this study. Our ChIP analysis revealed that only the fraction of Pol II tested using 8WG16 antibody was slightly depleted in *ceg1-AA* (Rap). This antibody recognises unphosphorylated heptapeptides in the CTD and therefore the precipitated fraction may be enriched for hypophosphorylated Pol II. Such Pol II may not efficiently transit to the elongation phase, pause or transcribe at a low speed (Bartkowiak and Greenleaf, 2011; Harlen *et al.*, 2016; Harlen and Churchman, 2017). All these conditions facilitate transcription termination by a torpedo mechanism, which may contribute to the lowered levels of Pol II observed in the ChIP analyses employing 8WG16 antibody. Consistently we did not observe any significant change in Ser5 and Ser2 phosphorylated form of Pol II. The previous studies reported premature transcription termination for the 8 kb long *FMP27* gene (Jimeno-González *et al.*, 2010).

Although we did not observe a similar effect in *ceg1-AA* (Rap) we hypothesise that for extremally long yeast genes like *FMP27*, the opportunity window to recruit Rai1-Rat1 to Pol II is wider and therefore premature transcription termination may occur under specific conditions. We speculate that for average length genes (1.4 kb) the window in which Rai-Rat1 can be recruited to the transcribing complex and thus catch normally transcribing Pol II is too narrow to trigger transcription termination upstream of the canonical transcription end site. Premature transcription termination is observed when nascent RNAs are decapped which triggers Pol II pausing and consequently, termination by Rat1 (Brannan *et al.*, 2012; Contreras, Benkirane and Kiernan, 2013; Eaton and West, 2020; Eaton *et al.*, 2020). We speculate that the capping process must be at least initiated and the capping enzymes present on the transcribing complex to permit the quality control mechanism.

Overall, we found that Pol II-generated transcripts are not specifically degraded when synthesised without the m7G cap. Instead, they undergo degradation pathways by available enzymes resulting in rapid depletion of the abundant transcripts in the first instance. The lack of dedicated pathways degrading non-capped RNA may have evolved to facilitate synthesis and export of all RNAs generated in the nucleus.

ACKNOWLEDGMENTS

This work was funded by Sir Henry Dale Fellowship from the Wellcome Trust and the Royal Society (218537/Z/19/Z). O.Z. was supported by Midlands Doctoral Training Partnership by a studentship to O.Z. funded by BBSRC (grant number: BB/M01116X/1). DH is funded by EPSRC (grant EP/T002794/1). We thank Kinga Winczura and Lauren Garbett for critical reading of the manuscript and Genomics Birmingham at the University of Birmingham for assistance with NGS.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

AUTHOR CONTRIBUTION

O.Z. performed all experiments, bioinformatics and analysed data. D.H. supported bioinformatics. P.G. and O.Z. designed experimental approaches and wrote the manuscript.

MATERIALS AND METHODS

Data availability

The RNA-seq data for the *ceg1-AA* strain upon 45 min of treatment with DMSO or rapamycin are available in Gene Expression Omnibus (GEO) database with the accession number GSE213942 .

Yeast strain and Media

Strain	Genotype	Source
HHY168	MATalpha tor1-1 fpr1::NAT RPL13AxFKB12::TRP	Euroscarf
ceg1-AA	MATalpha tor1-1 fpr1::NAT RPL13AxFKB12::TRP CEG1-FRB::HIS	This work
cet1-AA	MATalpha tor1-1 fpr1::NAT RPL13AxFKB12::TRP CET1-FRB::HIS	This work
abd1-AA	MATalpha tor1-1 fpr1::NAT RPL13AxFKB12::TRP ABD1-FRB::HIS	This work
cbp20-AA	MATalpha tor1-1 fpr1::NAT RPL13AxFKB12::TRP CBP20-FRB::HIS	This work
cbp80-AA	MATalpha tor1-1 fpr1::NAT RPL13AxFKB12::TRP CBP80-FRB::HIS	This work
ceg1-AA/xrn1Δ	MATalpha tor1-1 fpr1::NAT RPL13AxFKB12::TRP CEG1-FRB::HIS KAN::XRN1	This work
ceg1-AA/rrp6Δ	MATalpha tor1-1 fpr1::NAT RPL13AxFKB12::TRP CEG1-FRB::HIS KAN::RRP6	This work
ceg1-63	MATa ceg1Δ::HIS3 PGAL1-YLR454::URA3 ura3Δ [pRS315-ceg1-63]	(Jimeno-González *et al.*, 2010)

ceg1-63/rat1-1	MATa rat1-1 ceg1Δ::HIS3 PGAL1-YLR454::URA3 ura3Δ [pRS315-ceg1-63]	(Jimeno-González *et al.*, 2010)
fmp27-C76U	MATalpha tor1-1 fpr1::NAT RPL13AxFKB12::TRP CEG1-FRB::HIS URA3::RZc76u::YLR454	This work
fmp27-RZ	MATalpha tor1-1 fpr1::NAT RPL13AxFKB12::TRP CEG1-FRB::HIS URA3::RZ::YLR454	This work

Yeast strains were grown at 30° C on YPD (1% yeast extract, 2% peptone, 2% dextrose/glucose) media enriched in DMSO or rapamycin to the final concentration of 1 µg/mL where indicated. For 6-AU treatment, the yeast cells were grown in minimal media (yeast nitrogen base without amino acids) with synthetic dropout ura-. For solid media, agar was added to the final concentration of 2%.

Growth spot assay

Yeast cells were grown to OD600 = 0.4. Serial dilutions 1 :25 were performed and 5 µL of each dilution spotted onto YPD (2% agar) plates.

Yeast cell transformation

Cells were cultured to OD600 = 1 and pelleted by centrifugation for 5 min at 2 400 x g at room temperature (RT). The pellet was resuspended in sterile 10 mM Tris-HCl pH 7.5 or water and centrifuged again as before. The pellet was resuspended in filtered LiT (10 mM Tris-HCl pH 7.5;100 mM Lithium acetate), 1 M DTT and incubated at RT for 40 min with gentle shaking. After incubation the cells were pelleted again (as before) and resuspended in LiT and 1 M DTT. The suspension was added with LiT; and dsDNA (10 mg/mL); transforming DNA (0.1 to 1 ug or more) and incubated at RT

for 10 min. Next, PEG solution (1:1 PEG4000 : LiT) was added to the suspension and incubated at RT for 10 min. DMSO was then added, and the suspension incubated at 42° C for 15 min. After the incubation the suspension was pelleted and resuspended in 1 mL of YPD (1% Bacto yeast extract, 2% Bacto peptone, 2% glucose/dextrose). Then the suspension was incubated at 30° C for 1 hour, pelleted with 10 seconds (sec) spin, max speed and inoculated on selective plates.

RNA extraction

RNA was purified by phenol/chloroform method. The yeasts were cultured to the final concentration OD600 = 1, pelleted in 50 mL falcon tubes by centrifugation for 2 min at 1000 rpm at room temperature (RT). The pellet was frozen at -80° C before further manipulation. Next, the pellet was resuspended in 1 mL of ice-cold RNase free water, moved to a 1.5 mL tube, centrifuged for 10 sec at 4° C max speed. Then the supernatant (water) was removed, and the pellet resuspended in 400 µL of filtered AE buffer (50mM Sodium acetate pH5.3; 10mM EDTA), 40 µL of 10% Sodium dodecyl sulphate (SDS) and 400 µL of Acid Phenol (pH 4). The suspension was vortexed for 20 sec and incubated at 65° C for 10 min, then incubated for further 10 min at -80° C. The defrosted sample was then centrifuged for 5 min, 13 000 rpm at RT. 400 µL of 1:1 phenol:chloroform solution were added to the supernatant and vortexed for 30 sec. Then the sample was centrifuged 10 min, 13 000 rpm at RT. The upper phase was transferred and mixed with 400 µL of Chloroform, vortexed and centrifuged for 5 min, 13 000 rpm at RT. The RNA was precipitated by adding to the upper phase 1 mL of 100% ethanol and 40 µL of 7.5 M Ammonium acetate. Left at -80° C for 2 hours (hrs) then centrifuged for 20 min, 13 000 at 4° C. The pellet was washed with 70% ethanol (in water) and resuspended in RNase free water.

RT-qPCR

Reverse transcription was performed with Super Script III Reverse transcriptase (Thermo Fisher, 18080044) following manufacturer's instruction using random hexamers (Thermo Fisher, N8080127) for cDNA synthesis. qPCR data were analysed with the ΔΔCt method normalised to 25S ribosomal RNA (rRNA) and control condition.

RNA sequencing

Before library preparation, ribosomal RNA (rRNA) depletion was performed on 5 µg of total yeast RNA using 200 pmol of 3' biotinylated probes (IDT) designed in house following the RiboPOP method (Thompson, Kiourlappou, and Davis 2020). The libraries preparation was performed with Ultra II RNA Library Prep Kit for Illumina (NEB, E7770) and Multiplex Oligos for Illumina (NEB, E7335) according to manufacturer's instructions. The RNA sequencing was performed from the Genomics Facility at the University of Birmingham on Illumina NEXTseq apparatus.

RNA-seq data analysis

The quality of the data was checked using FastQC v0.11.5 (Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc).

Reads were aligned to the yeast genome R64 (sacCer3) using HISAT2 v2.1.1 and sorted and indexed using SAMtools v1.15.1 (Wysoker et al. 2009). The read count per gene were calculated with LiBiNorm v.2.5 (Dyer, Shahrezaei and Hebenstreit, 2019) with the following parameters count -u name-root -f --order=pos --minaqual=10 --mode=intersection-strict --idattr=gene_id --type=exon. Differential expression analysis

was performed using DESeq2 package. GO analysis was generated using g:profiler (https://biit.cs.ut.ee/gprofiler).

Cap Immunoprecipitation

Total RNA from cells treated with DMSO or rapamycin was extracted and use for cap pull-down. 5 µg of total yeast RNA was spiked with 1 µg of total human RNA as immunoprecipitation efficiency control. Capped RNA species were immunoprecipitated using 5 µg of anti m7G cap antibody (H-20) as previously described (Jimeno-González et al. 2010).

6-azauracil (6-AU) treatment

Overnight culture was prepared inoculating a single colony on minimal media as described above. The next day the culture was diluted to OD600 = 0.2. The diluted culture was grown with agitation at 30° C to OD600 = 1 and split in 4. Two cultures were added of 6-azauracil (6-AU) and the other two with ammonium hydroxide (NH4OH, control) at the final concentration of 50 µg/mL for 30 min. Next, DMSO or rapamycin (Rap) at the final concentration of 1 µg/mL were added according to the following scheme: 6-AU+DMSO, 6-AU+Rap, NH4OH+DMSO, NH4OH+Rap. 10 mL of each suspension were collected a time 0, 45, 90 and 120 min after the addition of DMSO or Rap and prepared for RNA purification.

4-Thiouracil (4tU) labelling and purification

Liquid yeast culture traded for 40 min with rapamycin (1 µg/mL) or DMSO were added of 4tU 500 µM final concentration (Sigma-Aldrich, 440736) to final concentration 5 mM for 5 min and subsequently flash frozen in cold ethanol (-80° C). The suspension was

pelleted RNA purified by Phenol/Chloroform extraction. 200 µg of total RNA from *S. cerevisiae* was spiked with *S. pombe* total RNA (4:1 ratio *S. cereviasiae* : *S. pombe*) treated with 4-tU 500 µM final concentration (Sigma-Aldrich, 440736). The biotinylation was performed with 200ul EZ-link-HPDP-biotin (Thermo Fisher, 21341) for 30 min at 65° C. The RNA was then incubated with 5M sodium chloride and isopropanol and incubated with Dynabeads M-270 Streptavidin (Thermo Fisher, 65305) for 1 hour at room temperature. Beads were washed five times with Binding & Washing buffer (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA and 1 M NaCl) and 4tU labelled-RNA was eluted with 100 µM DTT, precipitated with 5M NaCl, isopropanol and glycogen, and resuspended DNase/RNase free water. Enrichment of nascent RNA was confirmed with qPCR using exon–intron primers of ACT1 gene.

Chromatin immunoprecipitation (ChIP)

The specific yeast strain was grown over-night to OD600 = 1. Then 37% formaldehyde was added to the media to final concentration 1% dropwise with gentle shaking for 15 min. Next, 15 mL of 2.5 M glycine was added for further 5 min keeping the shaking. The cells were pelleted by centrifugation at 3 500 rpm for 3 min at 4° C. The pellet was resuspended in 40 mL of ice-cold 1x PBS and centrifuged at 3 500 rpm for 5 min at 4° C, twice. The pellet obtained was then transferred and prepared for lysis. The pellet was resuspended in 1 mL of ice-cold FA1 (50 mM HEPES-KOH, pH7.5; 150 mM NaCl; 1 mM EDTA; 1% triton-x 100; 0.1% Sodium deoxycholate) added with proteinase inhibitor (Roche Mini-EDTA free) and about 400 µL zirconia beads. The cells were lysed in MagnaLyser (30 sec 7000 rpm, 3 cycles with 5 min incubation on ice in between). The lysate was drained out of the beads by centrifugation 1 000 rpm for 1 min at 4° C and SDS was added to a final concentration of 0.05%. The lysate was

sonicated as follow: 25 cycles 15 sec ON, 15 sec OFF on maximum power. Next, it was centrifuged for 20 min at 12 500 rpm at 4° C. 200 µL of lysate was saved to check the fragmentation. SDS was added to the lysate with to final concentration of 1% and 0.05 mg of Proteinase K and incubated at 42° C for 1 hour. Next it was moved to 65° C for 4 hrs. The DNA de-crosslinked was precipitated by Phenol/Chloroform, the contaminant RNA removed by adding 2 µL of RNase A incubating for 1 hour at 37° C and checked on agarose gel. The rest of the lysate was centrifuged max speed for 5 min at 4° C. 150 µg of chromatin were used for immunoprecipitation. FA1 enriched with proteinase inhibitor was added to the DNA to a final volume of 700 µL. 20 µL of such mix was saved as input for the further ChIP analysis. 2 µg of the specific antibody was added to the sample and incubated on rotation over-night at 4° C.

15 µL of protein A (Thermo Fisher, 10002D) and 15 µL of protein G-coated beads (Thermo Fisher, 10004D) were washed with 1 mL of FA1 and resuspended in 100 µL of FA1. The beads were added to the sample and incubated by rotation for 1 hour at 4° C. After incubation the beads were washed 6 times with FA1, 1 time with FA2 (50 mM HEPES-KOH, pH 7.5; 500 mM NaCl; 1 mM EDTA; 1% Triton-X 100; 0.1% Sodium deoxycholate), 1 time with FA3 (20 mM Tris pH 8.0; 250 mM LiCl; 0.5% NP-40; 0.5% Sodium deoxycholate; 1 mM EDTA) and TE buffer pH 8 (100 mM Tris-HCl; 10 mM EDTA). The beads were then resuspended in 195 µL of Elution buffer (50 mM Tris pH 7.5; 10 mM EDTA; 1% SDS) and incubated for 10 min at 65° C and 0.05 mg of Proteinase K added. The Chromatin was then de-crosslinked by incubating the sample at 42 ° C for 1 hour and moved at 65° C for 4 hrs. The DNA was finally purified using Quiagen cleanup (28104) kit following the manufacturer's instruction.

FIGURES AND SUPPLEMENTAL FIGURES

Figure 1

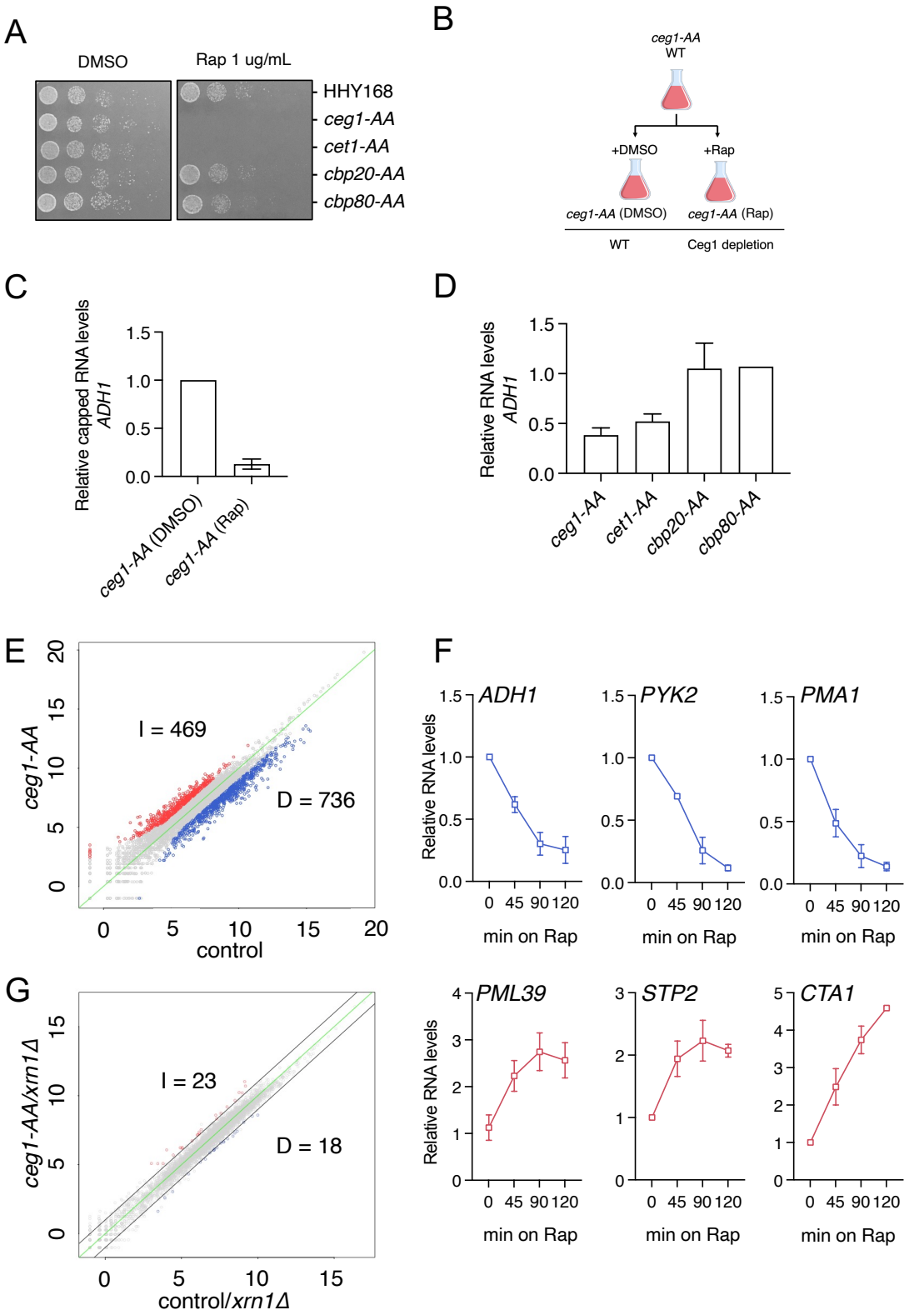


Figure 1. Nuclear depletion of capping enzymes results in differential accumulation of mRNAs.

A. Spot test showing the growth of the AA strains. Serial dilution of each strain was spotted on solid YPD media in the presence of DMSO or rapamycin at the final concentration of 1 ug/mL. **B.** The outline of the experimental approach. Liquid cultures of cells at the exponential phase were split and supplemented with DMSO (control sample) or rapamycin (RAP; nuclear depletion sample) shown on *ceg1-AA* example. **C.** Nuclear depletion of Ceg1 affects formation of the m7G cap. Quantification of the capped RNA by RT-qPCR after 45 min of rapamycin treatment compared to control (DMSO). The RNA isolated from yeast was spiked-in with human RNA and immunoprecipitated by the anti-m7G cap antibody (clone H20). The RNA levels were normalised to human *GAPDH* mRNA. The error bars show the standard deviation of three independent experiments. **D.** Nuclear depletion of the Cap Binding Complex does not affect mRNA levels. RT-qPCR analysis of the *ADH1* mRNA levels after 45 min of rapamycin treatment compared to the control. The error bars show the standard deviation of two (*cbp20-AA* and *cbp-80-AA*) or three independent experiments **E.** RNA-seq analysis showing differential accumulation of mRNA species in *ceg1-AA* (Rap) strain after 45 min of rapamycin treatment compared to the control condition. Decreased (D) RNA species showing log2 Fold change <1 are in blue. Increased (I) RNA species with log2 Fold change >1 are labelled in red. The zero-change line is in green. **F.** The changes in mRNA levels over time during nuclear depletion of Ceg1. RT-qPCR analysis of RNA levels for different genes measured at 0, 45, 90 and 120 min of rapamycin treatment relative to control. The error bars show the standard deviation of three independent experiments. **G.** *XRN1* deletion rescues the mRNA levels in *ceg1-AA* (Rap). Differential accumulation of mRNAs after 45 min of rapamycin

treatment in *ceg1-AA/xrn1Δ mutant* compared to the control condition. Decreased (D) RNA species showing log2 Fold change <1 are in blue. Increased (I) RNA species with log2 Fold change >1 are labelled in red. The zero-change line is in green. The black lines represent the ±1 log2 Fold change threshold.

Figure 2

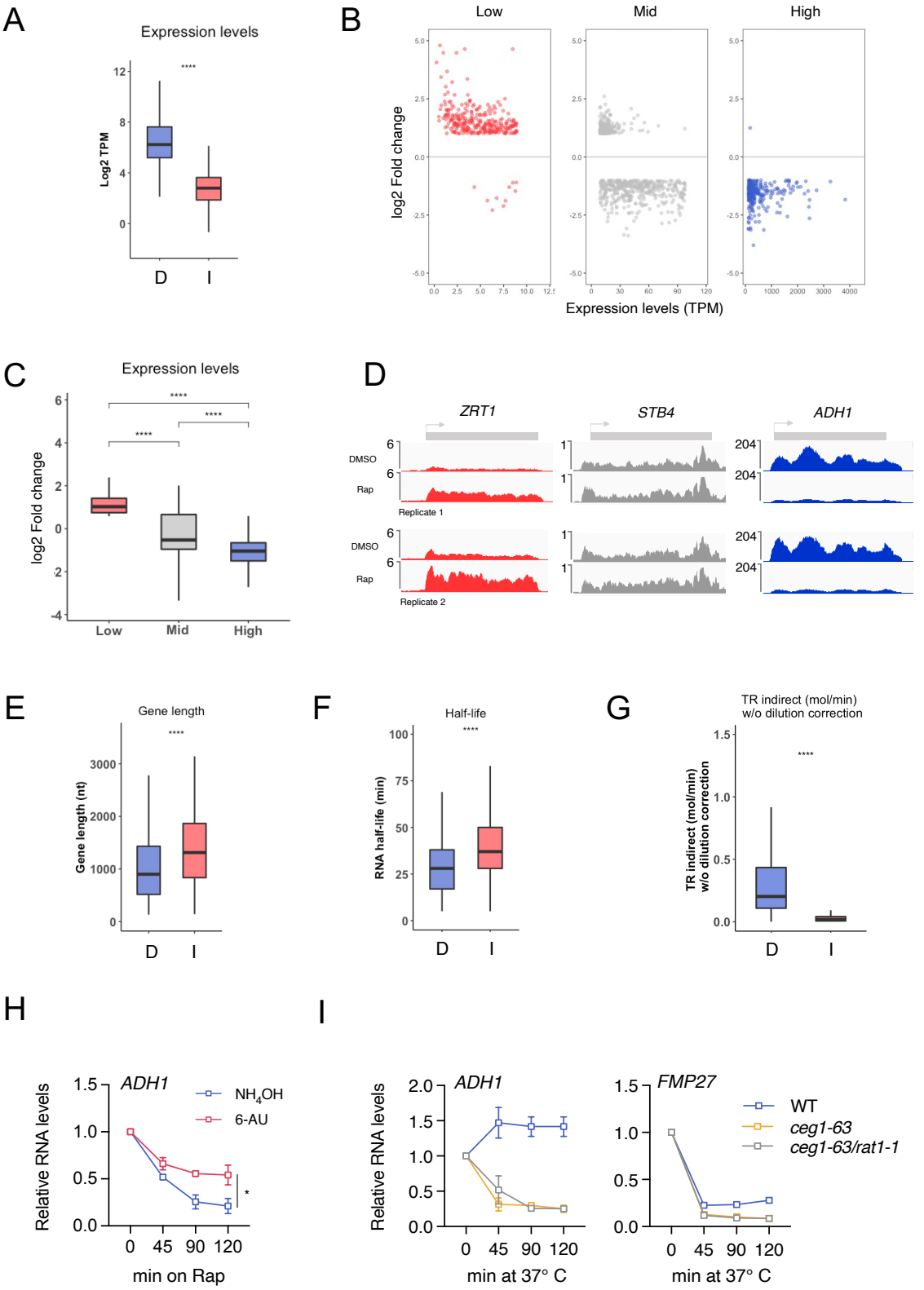

A
Expression levels
Log2 TPM
D
I
B
Low
Mid
High
log2 Fold change
Expression levels (TPM)
C
Expression levels
log2 Fold change
Low
Mid
High
D
ZRT1
STB4
ADH1
DMSO
Rap
Replicate 1
Replicate 2
E
Gene length
Gene length (nt)
F
Half-life
RNA half-life (min)
G
TR indirect (mol/min)
w/o dilution correction
H
ADH1
NH4OH
6-AU
Relative RNA levels
min on Rap
I
ADH1
FMP27
WT
ceg1-63
ceg1-63/rat1-1
min at 37° C

Figure 2. Factors defining the levels of non-capped mRNAs

A. Differentially accumulated mRNAs in *ceg1-AA* (Rap) vs *ceg1-AA* (DMSO) sorted by their expression levels: Increased (I) or decreased (D) mRNA in *ceg1-AA* (Rap) are shown. The expression levels were calculated by Transcript Per Million (TPM) value in *ceg1-AA* (DMSO). **B.** Plots showing the distribution of differentially accumulated mRNAs in *ceg1-AA* (Rap) depending on their expression levels. **C.** A box plot showing overall fold change for differentially accumulated mRNAs in *ceg1-AA* (Rap) vs *ceg1-AA* (DMSO) ranked accordingly to their expression levels. **D.** Genome viewer tracks showing RNA-seq reads in *ceg1-AA* (DMSO) and *ceg1-AA* (Rap) for highly, medial and lowly expressed genes. **E-G.** Box plots showing length (E), half-life (F) and transcription rates (G) of differentially accumulated mRNAs in *ceg1-AA* (Rap) vs *ceg1-AA* (DMSO). **H.** RT-qPCR analysis showing the levels of *ADH1* mRNA upon 6AU (red) and NH_4OH (blue) treatments in Ceg1-depleted cells. The error bars show the standard deviation of three independent experiments. Statistical analysis is shown for the last timepoint (120 min). **I.** RNA levels of *ADH1* (left) and *FMP27* (right) in WT and temperature-sensitive mutants (*ceg1-63* and *ceg1-63/rat1-1)* after 0, 45, 90 and 120 min at non-permissive temperature (37° C) compared to control (25° C). The error bars show the standard deviation of three independent experiments. **A, C, E-G.** ns=P > 0.05; *=P ≤ 0.05; **=P ≤ 0.01; ***= P ≤ 0.001; ****= P ≤ 0.0001.

Figure 3

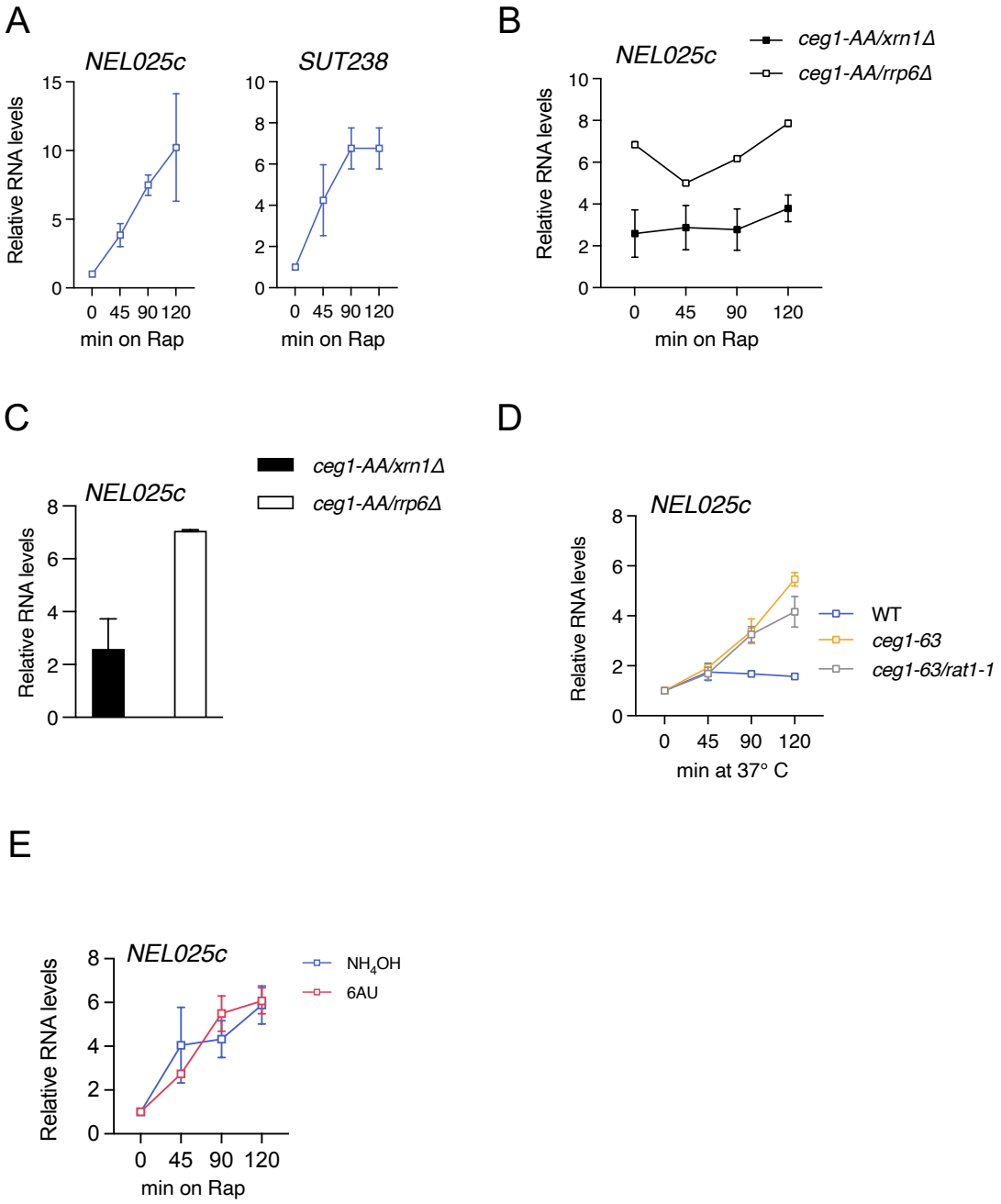
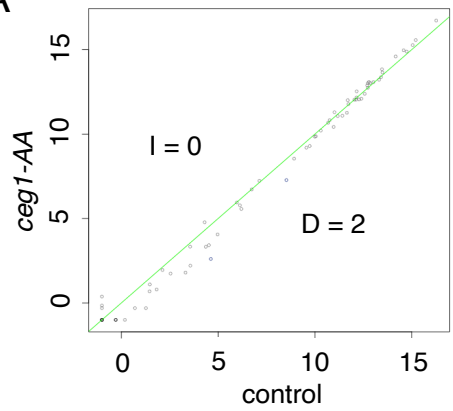

A
NEL025c
SUT238
Relative RNA levels
min on Rap
B
NEL025c
ceg1-AA/xrn1Δ
ceg1-AA/rrp6Δ
Relative RNA levels
min on Rap
C
NEL025c
ceg1-AA/xrn1Δ
ceg1-AA/rrp6Δ
Relative RNA levels
D
NEL025c
WT
ceg1-63
ceg1-63/rat1-1
Relative RNA levels
min at 37° C
E
NEL025c
NH4OH
6AU
Relative RNA levels
min on Rap

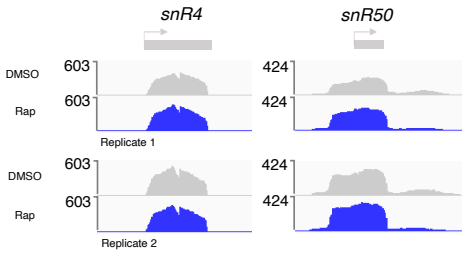
Figure 3. The stability of non-capped non-coding RNAs.

A. The accumulation of the *NEL025c* (left) and *SUT238* (right) RNAs in *ceg1-AA* strain upon 0, 45, 90 and 120 min of rapamycin treatment compared to control (DMSO). **B.** RNA levels of *NEL025c* in *ceg1-AA/rrp6Δ* and *ceg1-AA/xrn1Δ* strains after 0, 45, 90 and 120 min of rapamycin treatment compared to control (DMSO). **C.** Plot showing the increased basal RNA levels of *NEL025c* in *ceg1-AA/rrp6Δ* and *ceg1-AA/xrn1Δ* compared to *ceg1-AA* (all DMSO**D.** RNA levels of *NEL025c* in WT and temperature-sensitive mutant *ceg1-63* after 0, 45, 90 and 120 min at non-permissive temperature (37° C) compared to control (25° C). **E.** RT-qPCR showing the levels of *NEL025c* RNA upon 6AU (red) and NH_4OH (blue) treatments in Ceg1-depleted cells. A-E. are RT-PCR analyses. Error bars show the standard deviation of three independent experiments.

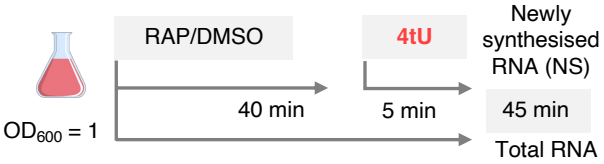
Figure 4

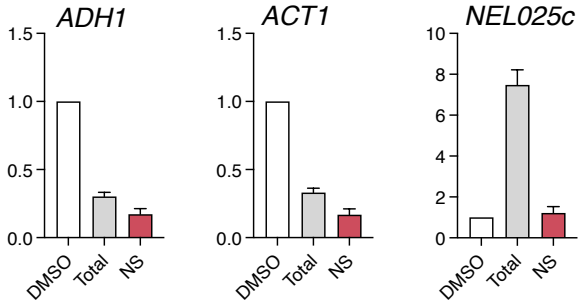
A



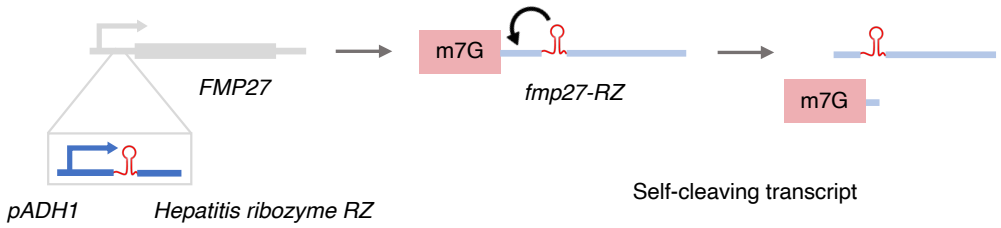


B

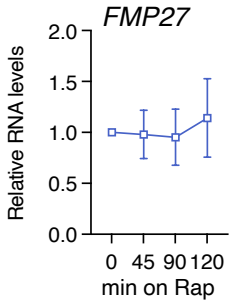




C



D



E

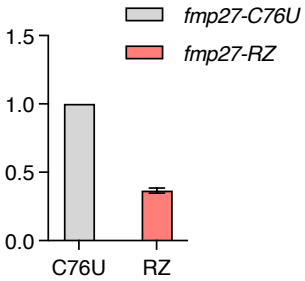


Figure 4. Cap removal is required for 5'-3' degradation.

A. RNA-seq analysis showing differential accumulation of snoRNA species in *ceg1-AA* (Rap) strain after 45 min of rapamycin treatment compared to the control condition. Decreased (D) RNA species showing log2 Fold change <1 are labelled in blue. The zero-change line is in green. The normalised RNA-seq reads for snR4 and snR50 snoRNA in *ceg1-AA* (Rap) and (DMSO) are shown on the right. **B.** Schematic representation of the experimental procedure for the Ceg1 depletion and metabolic 4-thiouracil (4tU) labelling of newly synthesised (NS) RNA (top). Chart showing total or newly synthesised RNA levels of *ADH1*, *ACT1* and *NEL025c* in control conditions (DMSO) and for 45 min of rapamycin treatment. The error bars show the standard deviation of three independent experiments (bottom). **C.** The experimental approach for the generation of the self-cleaving *fmp27-RZ* transcript. **D.** *FMP27* mRNA levels in *ceg1-AA* after 0, 45, 90 and 120 min of rapamycin treatment compared to control (DMSO). The error bars show the standard deviation of three independent experiments. **E.** A plot showing mRNA levels of *fmp27-RZ* and *fmp27-C76U* mutants growing in media containing DMSO. The error bars show the standard deviation of three independent RT-qPCR experiments.

Figure 5

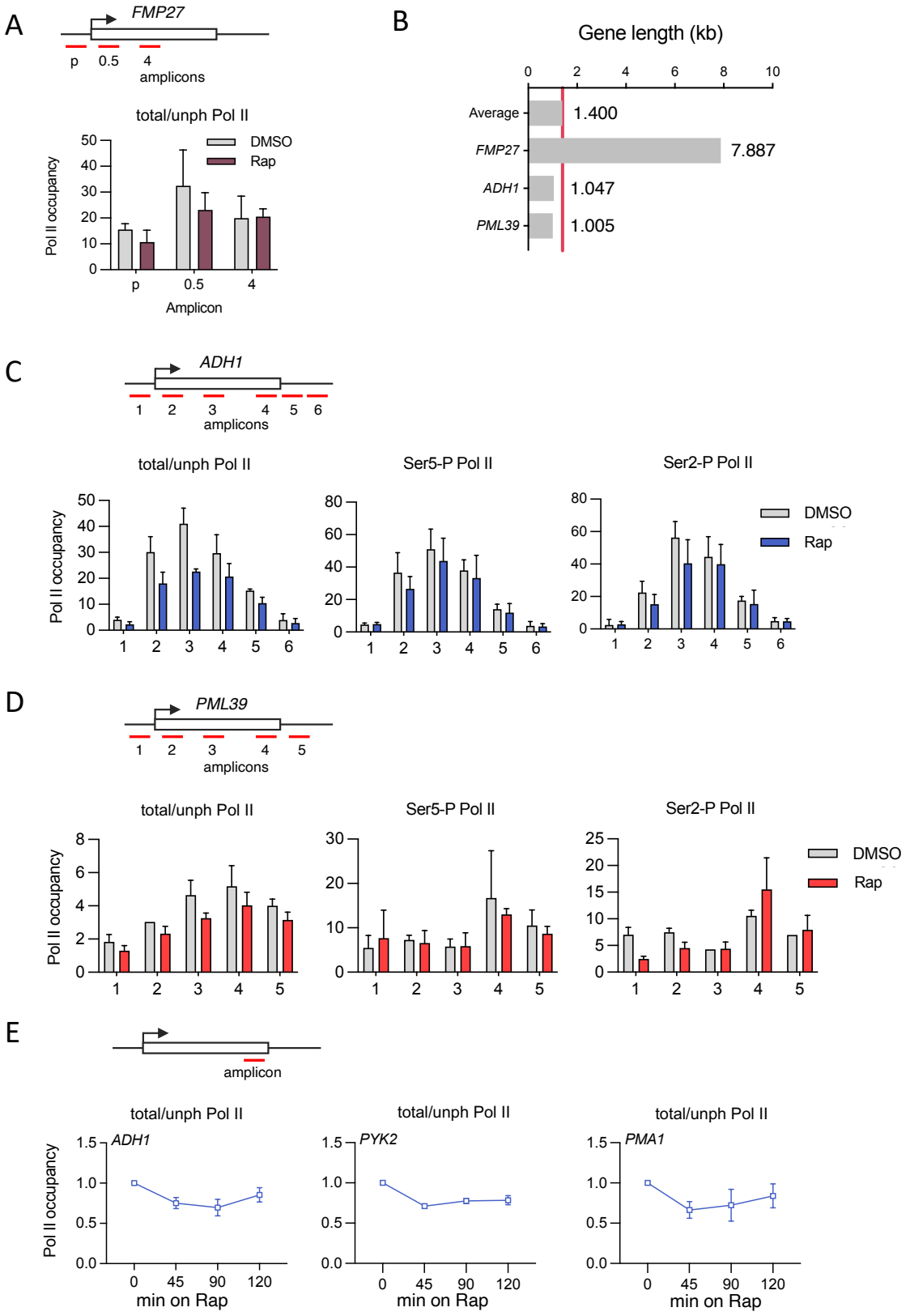

A
FMP27
p
0.5
4
amplicons
total/unph Pol II
DMSO
Rap
Pol II occupancy
Amplicon
B
Gene length (kb)
Average
FMP27
ADH1
PML39
1.400
7.887
1.047
1.005
C
ADH1
amplicons
total/unph Pol II
Ser5-P Pol II
Ser2-P Pol II
DMSO
Rap
Pol II occupancy
D
PML39
amplicons
total/unph Pol II
Ser5-P Pol II
Ser2-P Pol II
DMSO
Rap
Pol II occupancy
E
amplicon
total/unph Pol II
ADH1
PYK2
PMA1
Pol II occupancy
min on Rap

Figure 5. Transcription of non-capped mRNAs is not disrupted by pre-mature termination.

A. The distribution of total/unphosphorylated Pol II (total/unph Pol II) over *FMP27* in control (DMSO) and after 45 min of rapamycin treatment (Rap) in *ceg1-AA*. The location of the amplicons used for the ChIP-qPCR over *FMP27* is shown above the chart. The error bars show the standard deviation of three independent ChIP-qPCR experiments. **B.** Chart showing the length of FMP27, *ADH1* and *PML39* genes. The red line represents the average gene length in *S. cerevisiae* (1.4 kb). **C-D.** The distribution of total/unphosphorylated Pol II (total/unph Pol II) as well as serine 5-phosphorylated (Ser5-P Pol II) and serine 2-phosphorylated (Ser2-P Pol II) Pol II isoforms over *ADH1 (C)* and *PML39* (D) in control (DMSO) and after 45 min of rapamycin (Rap) in *ceg1-AA* strain. The location of the amplicons used for the ChIP-qPCR is shown above the charts. The error bars show the standard deviation of three independent ChIP-qPCR experiments. **E.** The occupancy of the total/unphosphorylated Pol II at the 3' end of *ADH1, PYK2* and *PMA.* Timepoints after 0, 45, 90 and 120 min on rapamycin compared to control (DMSO). The location of the amplicons used for the ChIP-qPCR is shown above the charts. The error bars show the standard deviation of three independent ChIP-qPCR experiments.

Figure S1

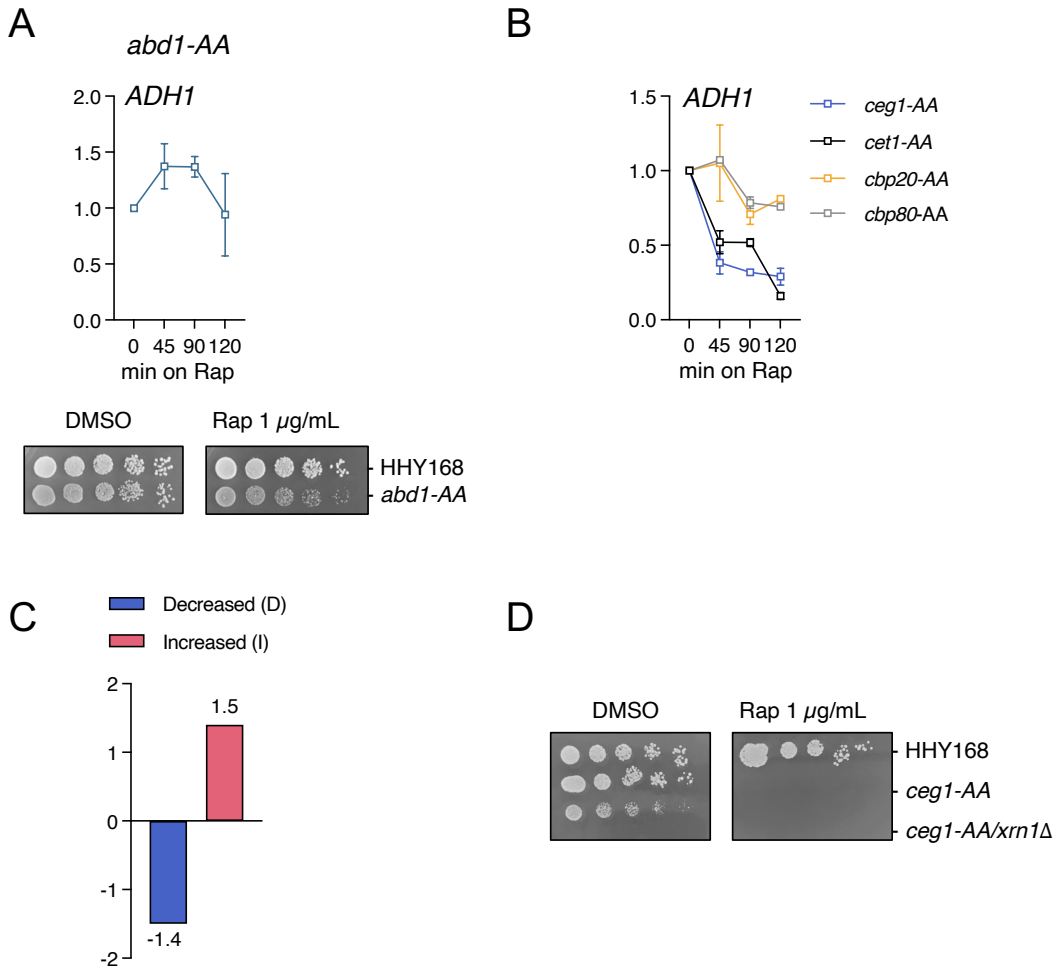


Figure S1.

A. *ADH1* mRNA levels time during nuclear depletion of Abd1. RT-qPCR RNA levels at 0, 45, 90 and 120 min of rapamycin treatment relative to control (top). The error bars show the standard deviation of three independent experiments. Spot test showing the growth of *abd1-AA* strain on DMSO and Rap (bottom) in comparison with to the parental strain HHY168. Serial dilution of each strain was spotted on YPD media in the presence of DMSO or rapamycin at the final concentration of 1 ug/mL. **B.** *ADH1* mRNA levels during nuclear depletion of Ceg1, Cet1, Cbp20 and Cbp80. RT-qPCR showing the RNA levels measured at 0, 45, 90 and 120 min of rapamycin treatment

relative to control (DMSO). The error bars show the standard deviation of three independent experiments. **C.** Bar chart showing the average fold change for decreased (D) and increased (I) mRNA species upon Ceg1 nuclear depletion. **D.** Spot test showing the growth of *ceg1-AA/xrn1Δ* strain on DMSO and rapamycin. Comparison with the growth of the parental strains *ceg1-AA* and HHY168. Serial dilution of each strain was spotted on solid YPD media in the presence of DMSO or rapamycin at the final concentration of 1 ug/mL.

Figure S2

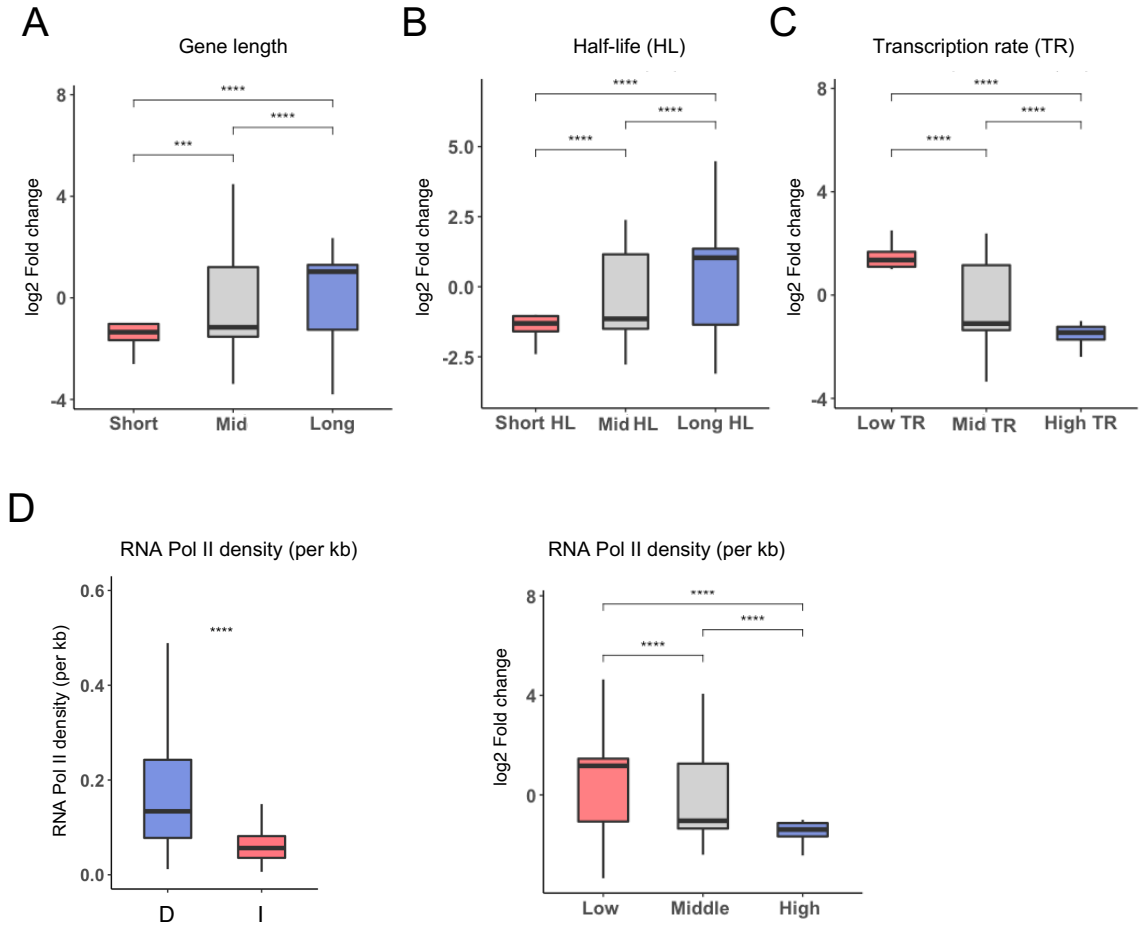
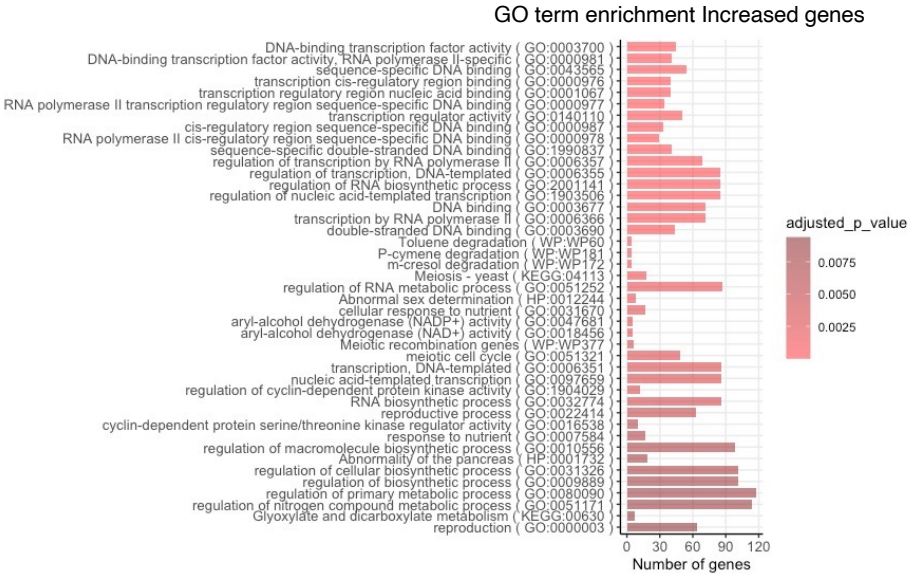


Figure S2.

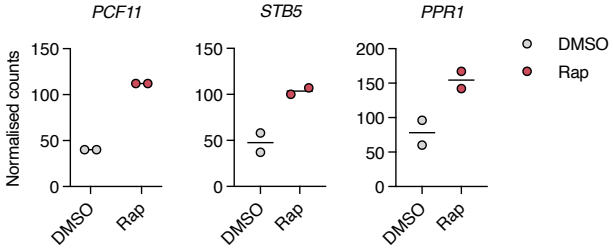
A-C. Fold change for differentially accumulated mRNAs in *ceg1-AA* (Rap) vs *ceg1-AA* (DMSO) plotted accordingly to their gene length (A), half-life (B) and transcription rates (C) and sorted by the expression level in *ceg1-AA* (DMSO). ns=P > 0.05; *=P ≤ 0.05; **=P ≤ 0.01; ***= P ≤ 0.001; ****= P ≤ 0.0001. **D.** Box plot showing the dependency of uncapped mRNA accumulation in Ceg1-depleted cells on RNA Pol II density (left). Fold change for differentially accumulated mRNAs in *ceg1-AA* (Rap) vs *ceg1-AA* (DMSO) plotted accordingly to their Pol II density and sorted by the expression level in *ceg1-AA* (DMSO). ns=P > 0.05; *=P ≤ 0.05; **=P ≤ 0.01; ***= P ≤ 0.001; ****= P ≤ 0.0001.

Figure S3

A



B



C

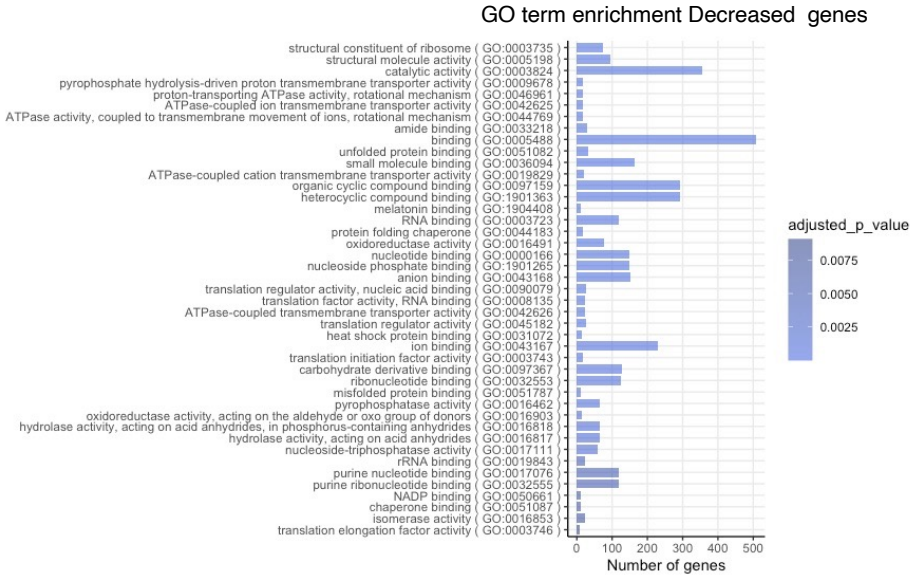


Figure S3.

A. GO terms associated with increased mRNAs in *ceg1-AA* (Rap). Charts showing the number of the genes belonging to the specific GO group. The classes are listed in a p-value increasing order. **B.** Plots showing the counts per gene for *PCF11*, *STB5* and *PPR1* obtained from the RNA-seq experiment. Charts showing the number of the genes belonging to the specific GO group. The classes are listed in a p-value increasing order. **C.** GO terms associated with decreased mRNAs in *ceg1-AA* (Rap). Charts showing the number of the genes belonging to the specific GO group. The classes are listed in a p-value increasing order.

Figure S4

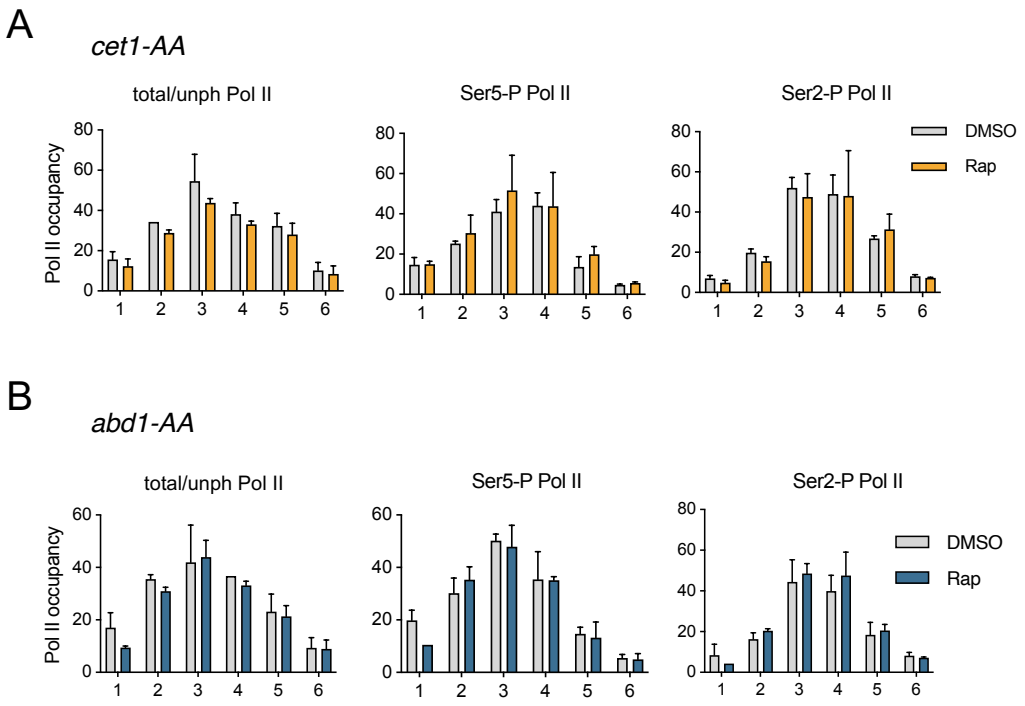


Figure S4.

A-B. The distribution of total/unphosphorylated (total/unph) Pol II as well as serine 5-phosphorylated (Ser5-P) and serine 2-phosphorylated (Ser2-P) Pol II isoforms over *ADH1* in control (DMSO) and after 45 min of rapamycin (Rap) in (A) *cet1-AA* and *abd1-AA* strain (B). The location of the amplicons used for the ChIP-qPCR is shown above the charts. The error bars show the standard deviation of three independent ChIP-qPCR experiments.

REFERENCES

Akiyama, B. M., Eiler, D. and Kieft, J. S. (2017) 'Structured RNAs that evade or confound exonucleases: function follows form', pp. 40–47. doi: 10.1016/j.sbi.2015.12.006.Structured.

Bartkowiak, B. and Greenleaf, A. L. (2011) 'Phosphorylation of RNAPII: To P-TEFB or not to P-TEFB?', *Transcription*, 2(3), pp. 115–119. doi: 10.4161/trns.2.3.15004.

Bird, G. *et al.* (2005) 'Ribozyme cleavage reveals connections between mRNA release from the site of transcription and pre-mRNA processing', *Molecular Cell*, 20(5), pp. 747–758. doi: 10.1016/j.molcel.2005.11.009.

Brannan, K. *et al.* (2012) 'mRNA Decapping Factors and the Exonuclease Xrn2 Function in Widespread Premature Termination of RNA Polymerase II Transcription', *Molecular Cell*. Elsevier Inc., 46(3), pp. 311–324. doi: 10.1016/j.molcel.2012.03.006.

Chu, C. and Shatkin, A. J. (2008) 'Apoptosis and autophagy induction in mammalian cells by small interfering RNA knockdown of mRNA capping enzymes.', *Molecular and cellular biology*, 28(19), pp. 5829–5836. doi: 10.1128/MCB.00021-08.

Contreras, X., Benkirane, M. and Kiernan, R. (2013) 'Premature termination of transcription by RNAP II: The beginning of the end', *Transcription*, 4(2). doi: 10.4161/trans.24148.

Cougot, N. *et al.* (2004) '"Cap-tabolism"', *Trends in Biochemical Sciences*, 29(8), pp. 436–444. doi: 10.1016/j.tibs.2004.06.008.

Decroly, E. *et al.* (2012) 'Conventional and unconventional mechanisms for capping viral mRNA', *Nature Reviews Microbiology*. Nature Publishing Group, 10(1), pp. 51–65. doi: 10.1038/nrmicro2675.

Dieci, G., Preti, M. and Montanini, B. (2009) 'Eukaryotic snoRNAs: A paradigm for gene expression flexibility', *Genomics*, 94(2), pp. 83–88. doi:

https://doi.org/10.1016/j.ygeno.2009.05.002.

Doamekpor, S. K. *et al.* (2020) 'A novel 5′-hydroxyl dinucleotide hydrolase activity for the DXO/Rai1 family of enzymes', *Nucleic Acids Research*. Oxford University Press, 48(1), pp. 349–358. doi: 10.1093/nar/gkz1107.

Dyer, N. P., Shahrezaei, V. and Hebenstreit, D. (2019) 'LiBiNorm: an htseq-count analogue with improved normalisation of Smart-seq2 data and library preparation diagnostics.', *PeerJ*, 7, p. e6222. doi: 10.7717/peerj.6222.

Eaton, J. D. *et al.* (2020) 'A unified allosteric / torpedo mechanism for transcriptional termination on human protein-coding genes', pp. 1–14. doi: 10.1101/gad.332833.119.

Eaton, J. D. and West, S. (2020) 'Termination of Transcription by RNA Polymerase II: BOOM!', *Trends in Genetics*. Elsevier Ltd, 36(9), pp. 664–675. doi: 10.1016/j.tig.2020.05.008.

Fortes, P. *et al.* (1999) 'Genetic and physical interactions involving the yeast nuclear cap-binding complex.', *Molecular and cellular biology*, 19(10), pp. 6543–6553. doi: 10.1128/MCB.19.10.6543.

Furuichi, Y. (2015) 'Discovery of m 7 G-cap in eukaryotic mRNAs', *Proceedings of the Japan Academy. Series B, Physical and biological sciences*, 91(8), pp. 394–409.

Gonatopoulos-Pournatzis, T. and Cowling, V. H. (2014) 'Cap-binding complex (CBC)', *Biochemical Journal*, 457(2), pp. 231–242. doi: 10.1042/BJ20131214.

Grzechnik, P. *et al.* (2018) 'Nuclear fate of yeast snoRNA is determined by co-transcriptional Rnt1 cleavage', *Nature Communications*. Springer US, 9(1). doi: 10.1038/s41467-018-04094-y.

Harlen, K. M. *et al.* (2016) 'Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue', *Cell Reports*. The

Authors, 15(10), pp. 2147–2158. doi: 10.1016/j.celrep.2016.05.010.

Harlen, K. M. and Churchman, L. S. (2017) 'The code and beyond: Transcription regulation by the RNA polymerase II carboxy-terminal domain', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 18(4), pp. 263–273. doi: 10.1038/nrm.2017.10.

Haruki, H., Nishikawa, J. and Laemmli, U. K. (2008) 'The Anchor-Away Technique: Rapid, Conditional Establishment of Yeast Mutant Phenotypes', *Molecular Cell*, 31(6), pp. 925–932. doi: 10.1016/j.molcel.2008.07.020.

Hausmann, S., Pei, Y. and Shuman, S. (2003) 'Homodimeric quaternary structure is required for the in vivo function and thermal stability of Saccharomyces cerevisiae and Schizosaccharomyces pombe RNA triphosphatases', *Journal of Biological Chemistry*, 278(33), pp. 30487–30496. doi: 10.1074/jbc.M303060200.

Hurowitz, E. H. and Brown, P. O. (2003) 'Genome-wide analysis of mRNA lengths in Saccharomyces cerevisiae.', *Genome biology*, 5(1), p. R2. doi: 10.1186/gb-2003-5-1-r2.

Izaurralde, E. *et al.* (1994) 'A nuclear cap binding protein complex involved in pre-mRNA splicing.', *Cell*. United States, 78(4), pp. 657–668. doi: 10.1016/0092-8674(94)90530-4.

Izaurralde, E. *et al.* (1995) 'A cap-binding protein complex mediating U snRNA export.', *Nature*. England, 376(6542), pp. 709–712. doi: 10.1038/376709a0.

Jiao, X. *et al.* (2010) 'Identification of a quality-control mechanism for mRNA 5′-end capping', *Nature*. Nature Publishing Group, 467(7315), pp. 608–611. doi: 10.1038/nature09338.

Jimeno-González, S. *et al.* (2010) 'The Yeast 5′-3′ Exonuclease Rat1p Functions during Transcription Elongation by RNA Polymerase II', *Molecular Cell*, 37(4), pp.

580–587. doi: 10.1016/j.molcel.2010.01.019.

Jinek, M., Scott M. Coyle and Doudna, and J. A. (2014) ‘Coupled 5′ nucleotide recognition and processivity in Xrn1- mediated mRNA decay’, *Bone*, 23(1), pp. 1–7. doi: 10.1016/j.molcel.2011.02.004.Coupled.

Kim, M. *et al.* (2004) ‘The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II’, *Nature*, 432(7016), pp. 517–522. doi: 10.1038/nature03041.

Lewis, J. D. and Izaurralde, E. (1997) ‘The role of the cap structure in RNA processing and nuclear export’, *European Journal of Biochemistry*, 247(2), pp. 461–469. doi: 10.1111/j.1432-1033.1997.00461.x.

Lidschreiber, M., Leike, K. and Cramer, P. (2013) ‘Cap Completion and C-Terminal Repeat Domain Kinase Recruitment Underlie the Initiation-Elongation Transition of RNA Polymerase II’, *Molecular and Cellular Biology*, 33(19), pp. 3805–3816. doi: 10.1128/MCB.00361-13.

Mao, X., Schwer, B. and Shuman, S. (1996) ‘Mutational analysis of the Saccharomyces cerevisiae ABD1 gene: cap methyltransferase activity is essential for cell growth’, *Molecular and Cellular Biology*, 16(2), pp. 475–480. doi: 10.1128/mcb.16.2.475.

Marquardt, S., Hazelbaker, D. Z. and Buratowski, S. (2011) ‘Distinct RNA degradation pathways and 3’ extensions of yeast non-coding RNA species’, *Transcription*, 2(3), pp. 145–154. doi: 10.4161/trns.2.3.16298.

Martinez-Rucobo, F. W. *et al.* (2015) ‘Molecular Basis of Transcription-Coupled Pre-mRNA Capping’, *Molecular Cell*. Elsevier Inc., 58(6), pp. 1079–1089. doi: 10.1016/j.molcel.2015.04.004.

Pelechano, V., Wei, W. and Steinmetz, L. M. (2013) ‘Extensive transcriptional heterogeneity revealed by isoform profiling’, *Nature*. doi: 10.1038/nature12121.

Perrotta, A. T., Shih, I. and Been, M. D. (1999) 'Imidazole Rescue of a Cytosine Mutation in a Self-Cleaving Ribozyme', *Science*. American Association for the Advancement of Science, 286(5437), pp. 123–126. doi: 10.1126/science.286.5437.123.

Ramanathan, A., Robb, G. B. and Chan, S. H. (2016) 'mRNA capping: Biological functions and applications', *Nucleic Acids Research*, 44(16), pp. 7511–7526. doi: 10.1093/nar/gkw551.

Schroeder, S. C. *et al.* (2000) 'Dynamic association of capping enzymes with transcribing RNA polymerase II', *Genes and Development*, 14(19), pp. 2435–2440. doi: 10.1101/gad.836300.

Schwer, B., Mao, X. and Shuman, S. (1998) 'Accelerated mRNA decay in conditional mutants of yeast mRNA capping enzyme', *Nucleic Acids Research*, 26(9), pp. 2050–2057. doi: 10.1093/nar/26.9.2050.

Shatkin, A. J. and Manley, J. L. (2000) 'The ends of the affair: Capping and polyadenylation', *Nature Structural Biology*, 7(10), pp. 838–842. doi: 10.1038/79583.

Topisirovic, I. *et al.* (2011) 'Cap and cap-binding proteins in the control of gene expression', *Wiley Interdisciplinary Reviews: RNA*, 2(2), pp. 277–298. doi: 10.1002/wrna.52.

Tsukamoto, T. *et al.* (1997) 'Isolation and characterization of the yeast mRNA capping enzyme beta subunit gene encoding RNA 5'-triphosphatase, which is essential for cell viability.', *Biochemical and biophysical research communications*. United States, 239(1), pp. 116–122. doi: 10.1006/bbrc.1997.7439.

Urrutia, A. O. and Hurst, L. D. (2003) 'The signature of selection mediated by expression on human genes', *Genome Research*, 13(10), pp. 2260–2264. doi: 10.1101/gr.641103.

Varshney, D. *et al.* (2018) ‘mRNA Cap Methyltransferase, RNMT-RAM, Promotes RNA Pol II-Dependent Transcription’, *Cell Reports*. ElsevierCompany., 23(5), pp. 1530–1542. doi: 10.1016/j.celrep.2018.04.004.

West, S., Gromak, N. and Proudfoot, N. J. (2004) ‘Human 5′ → 3′ exonuclease Xm2 promotes transcription termination at co-transcriptional cleavage sites’, *Nature*, 432(7016), pp. 522–525. doi: 10.1038/nature03035.

Wyers, F. *et al.* (2005) ‘Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase’, *Cell*, 121(5), pp. 725–737. doi: 10.1016/j.cell.2005.04.030.

Xiang, S. *et al.* (2009) ‘Structure and function of the 5′3′ exoribonuclease Rat1 and its activating partner Rai1’, *Nature*. Nature Publishing Group, 458(7239), pp. 784–788. doi: 10.1038/nature07731.

Xu, Z. *et al.* (2009) ‘Bidirectional promoters generate pervasive transcription in yeast’, *Nature*. doi: 10.1038/nature07728.