The transcription factor Sfp1 imprints specific classes of mRNAs and links their synthesis to cytoplasmic decay

Kelbert, M.[1]*, Jordán-Pla, A.[2]*, de-Miguel-Jiménez, L.[3]*, García-Martínez, J [2]., Selitrennik, M.[1], Guterman, A.[1], Henig, N[1], Granneman, S[4]., Pérez-Ortín, J.E. [2#], Chávez, S.[3#], and Choder, M. [1#]

1. Department of Molecular Microbiology, Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel
2. Departamento de Bioquímica y Biología Molecular, Facultad de Biológicas and Instituto Biotecmed, Universitat de València, Dr Moliner 50, E-46100 Burjassot, Valencia, Spain
3. Instituto de Biomedicina de Sevilla, Universidad de Sevilla-CSIC-Hospital Universitario V. del Rocío, and Departamento de Genética, Facultad de Biología, Universidad de Sevilla, Seville, E-41012, Spain.
4. Centre for Engineering Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom

* Equal contribution

Corresponding authors

Abstract

To function effectively as an integrated system, the transcriptional and post-transcriptional machineries must communicate through mechanisms that are still poorly understood. Here we focus on Sfp1, known to regulate transcription of proliferation-related genes. We show that Sfp1 can regulate transcription either by binding to promoters, like most known transcription activators, or by binding to the transcribed regions (gene bodies), probably via RNA polymerase II (Pol II). We further studied the first mode of Sfp1 activity and found that, following promoter binding, Sfp1 binds to gene bodies and affects Pol II configuration, manifested by dissociation or conformational change of its Rpb4 subunit and increased backtracking. Surprisingly, Sfp1 binds to a subset of mRNAs co-transcriptionally and stabilizes them. The interaction between Sfp1 and its client mRNAs is controlled by their respective promoters and coincides with Sfp1's dissociation from chromatin. . Intriguingly, Sfp1 dissociation from the chromatin correlates with the extent of the backtracked Pol II. We propose that, following promoter recruitment, Sfp1 accompanies Pol II and regulates backtracking. The backtracked Pol II is more compatible with Sfp1's relocation to the nascent transcripts, whereupon Sfp1 accompanies these mRNAs to the cytoplasm and regulates their stability. Thus, Sfp1's co-transcriptional binding imprints the mRNA fate, serving as a paradigm for the cross-talk between the synthesis and decay of specific mRNAs. The interplay between Sfp1's two modes of transcription regulation remains to be examined.

INTRODUCTION

In the last decade, investigators have accumulated evidence for a direct cross-talk between the transcriptional and post-transcriptional stages of gene expression. The dialogue between transcription and mRNA decay has been reported by several investigators, including us (Begley et al., 2019; Blasco-Moreno et al., 2019; Bregman et al., 2011; Bryll and Peterson, 2023; Buratti and Baralle, 2012; Chattopadhyay et al., 2022; El-Brolosy et al., 2019; Gilbertson et al., 2018; Goler-Baron et al., 2008; Haimovich et al., 2013b; Harel-Sharvit et al., 2010; Hartenian and Glaunsinger, 2019; Ma et al., 2019; Sun et al., 2012; Timmers and Tora, 2018; Trcek et al., 2011; Vera et al., 2014; Zid and O'Shea, 2014). Often, this dialog results in "mRNA buffering", which maintains an approximately constant mRNA level despite transient changes in mRNA synthesis or decay rates (Haimovich et al., 2013b; Sun et al., 2012). In particular, communication pathways have been discovered between promoters and mRNA translation (Chen et al., 2022; Vera et al., 2014; Zid and O'Shea, 2014) or decay (Bregman et al., 2011; Trcek et al., 2011). However, the underlying mechanism remains elusive.

During transcription by RNA polymerase II (Pol II), several RNA-binding proteins (RBPs) bind to nascent transcripts, thus modulating the transcription cycle (Battaglia et al., 2017). Importantly, in some cases, RBPs remain bound to the mRNA, accompanying it to the cytoplasm and regulating post-transcriptional stages. For example, the Pol II subunits Rpb4 and Rpb7 bind to mRNAs and regulate their translation and decay (Goler-Baron et al., 2008; Harel-Sharvit et al., 2010). We named the latter cases "mRNA imprinting", because this binding imprints the mRNA fate (Choder, 2011; Dahan and Choder, 2013). Although Rpb4 is a quintessential imprinted factor, it functions as a general factor that binds to numerous mRNAs (Garrido-Godino et al., 2021). Imprinting by a "classical" transcription factor (TF) that binds to specific promoters, thus regulating transcription by modulating the assembly of general transcription factors, including Pol II, is currently unknown. Such imprinting might broaden our understanding of the functions of TFs – not just transcription regulators but regulators of both the nuclear transcription and cytoplasmic post-transcription stages of gene expression.

We previously reported a special case of cross-talk between promoter elements and mRNA decay which is mediated by Rap1 binding sites (RapBSs) and by Rap1 protein (Bregman et al., 2011). RapBSs (12-14 bp) are found in ~5% of the promoters, including ~90% of ribosomal protein (RP) and a small portion of Ribosome Biogenesis (RiBi) promoters (Lascaris et al., 1999; Lieb et

al., 2001; Shore et al., 2021; Warner, 1999). Rap1 is bound to essentially all such RP promoters *in vivo* (Lieb et al., 2001; Schawalder et al., 2004) and is involved in their transcription activation. Rap1 functions as a pioneering transcription factor (TF) required for the binding of a pair of RP-specific TFs called Fhl1 and Ifh1(Lieb et al., 2001; Shore et al., 2021). An additional TF recruited by Rap1 is Sfp1 (see next paragraph). Previously, we demonstrated that Rap1 plays a dual role in maintaining the level of specific mRNAs – stimulating both mRNA synthesis and decay. We proposed that Rap1 represents a class of factors, synthegradases, whose recruitment to promoters stimulates (or represses) both mRNA synthesis and decay (Bregman et al., 2011). How Rap1 stimulates mRNA decay is unknown.

The split-finger protein 1 (Sfp1) is a TF, the deletion or overexpression of which has been shown to affect the recruitment of the TATA-binding protein (TBP) and Pol II to RiBi and RiBi-like genes, and to affect the transcription of RP and snoRNA genes (Albert et al., 2019; Fingerman et al., 2003; Jorgensen et al., 2004; Marion et al., 2004). It also impacts many G1/S ("START") genes, where it appears to act as a repressor (Albert et al., 2019). Sfp1 was therefore viewed as a classical TF. Shore's lab has examined the Sfp1 configuration within promoters, using either chromatin immunoprecipitation (ChIP-seq) or chromatin endogenous cleavage (ChEC-seq) and found that the two approaches exhibit different sensitivity to different configurations. While Sfp1 binding sites detected by ChEC are enriched for the motif gAAAATTTTc, binding identified by ChIP is dependent on Ifh1, hence on Rap1 as well (Albert et al., 2019). Thus, Sfp1 can be viewed as a "classical" zinc-finger containing TF. Zinc-finger domains have also been described to function as RNA-binding domains (Font and MacKay, 2010). In some proteins, such as TFIIIA, the same zinc-finger domain can recognize both a specific double stranded DNA sequence and another specific sequence in single stranded RNA (discussed in (Font and MacKay, 2010)).

In response to misincorporation of nucleotides, damage on the DNA, nucleosomes, hairpins or certain DNA sequences, elongating Pol II transiently pauses. During pausing, Pol II remains catalytically active and can resume transcription if the obstacle is removed (Bar-Nahum et al., 2005); (Gómez-Herreros et al., 2012a). However, these pauses increase the probability of Pol II backward movement along the DNA template, a process named "backtracking". During backtracking, the 3' end of the nascent RNA extrudes from the active site, traps the trigger loop of the enzyme and stably blocks mRNA elongation (Cheung and Cramer, 2011). Once arrested, Pol II can reactivate itself by cleaving its own synthetized RNA at the active site (Sigurdsson et al.,

2010). Cleavage is substantially enhanced by TFIIS (Churchman and Weissman, 2011), which becomes essential under NTP scarcity (Archambault et al., 1992). Important for this paper, Pol II backtracking is highly frequent at RP and RiBi genes, compared to other gene regulons, and is highly dependent on Rap1 (Pelechano et al., 2009). Our previous work has also shown that several RP regulators, including Rap1 and Sfp1, make transcription of RP genes highly dependent on TFIIS, especially under conditions of transcriptional stress due to NTP depletion (Gómez-Herreros et al., 2012b). Thus, Sfp1 can be considered as a backtracking regulator. Regulators that antagonize Pol II backtracking are Xrn1 (Begley et al., 2021, 2019; Fischer et al., 2020), Rpb4 (Fischer et al., 2020) and the-Ccr4-Not complex (Collart, 2016), which prevents backtracking in a manner that depends on its Rpb4 and Rpb7 subunits (Babbarwal et al., 2014; Kruk et al., 2011a). Interestingly, backtracking and TFIIS recruitment in RP genes are not influenced by Xrn1, but are strongly affected by Ccr4 (Begley et al., 2019).

Here we assigned a much broader function to a TF in gene expression. We discovered that Sfp1 regulates both transcription and mRNA decay through its capacity to bind RNAs co-transcriptionally. The Sfp1-mediated cross-talk between mRNA synthesis and decay results in cooperation of these opposing processes to increase mRNA abundance, by stimulating the former and repressing the latter. This cooperation is limited to a subgroup of Sfp1 targets, most of which are also Rap1 targets. Our results also unveil a role for some promoters as mediators between RPB and its interacting RNAs.

RESULTS

Sfp1 interacts with Rpb4

Previously, we discovered that the yeast Pol II subunits, Rpb4 and Rpb7, which form a heterodimer Rpb4/7, have post-transcriptional roles. Rpb4/7 is recruited onto mRNAs co-transcriptionally and is directly involved in all major post-transcriptional stages of the mRNA lifecycle such as RNA export (Farago et al., 2003) translation (Harel-Sharvit et al., 2010) and decay (Goler-Baron et al., 2008; Lotan et al., 2007, 2005). Thus, nascent mRNA emerges from the nucleus with “imprinted” information that serves to regulate post-transcriptional stages of gene expression (Dahan and Choder, 2012). To identify novel factors that may be involved in “mRNA imprinting”, we searched for Rpb4-interacting proteins, using a yeast two-hybrid screen. Among several interacting partners,

SFP1 was singled out, as it is a known transcription factor of genes encoding ribosomal proteins (RPs). RP genes represent a subclass of genes known to be preferentially regulated by Rpb4 at the level of mRNA degradation (Lotan et al., 2005). The interaction of Rpb4 with Sfp1 was corroborated by pairwise two-hybrid analysis and co-immunoprecipitation experiment (Fig. S1A and results not shown), as well as by an imaging approach (see later). Note that no interaction was observed with Fhl1, Ifh1 and Abf1, which encode additional transcription factors that often function in concert with Sfp1 (see Introduction). It is important to note that our two hybrid, immunoprecipitation and imaging results do not differentiate between a direct or indirect interactions.

Transcription-dependent export of Sfp1 from the nucleus to the cytoplasm

As a first step in determining whether Sfp1 plays a post-transcriptional role outside the nucleus, we examined its ability to shuttle between the nucleus and cytoplasm. GFP-Sfp1 was expressed in WT cells and in cells carrying the temperature-sensitive (ts) allele of *NUP49* (*nup49-313*), which encodes a nucleoporin used in a standard shuttling assay (Selitrennik et al., 2006, and references therein). Following a temperature shift up to 37°C to inactivate Nup49-dependent import, GFP-Sfp1 accumulated in the cytoplasm (Fig. 1A, "*nup49-313*"), indicating that this protein normally shuttles between the nucleus and cytoplasm, at least at 37°C. We next simultaneously disabled both protein import and transcription and by using cells with a double ts mutation in both *NUP49* and in Pol II subunit gene, *RPB1* (referred to as *nup49-313(ts), rpb1-1* double mutant cells). Simultaneous temperature inactivation of both import and transcription, using these double mutant cells, abolished Sfp1-GFP export (Fig. 1A, compare "*nup49-313*" and "*nup49-313, rpb1-1*"). This suggests that Sfp1 export to the cytoplasm is transcription-dependent, similar to Rpb4 (Selitrennik et al., 2006). Previously, it was observed that Sfp1 accumulates in the cytoplasm of cells when they are starved (Marion et al., 2004). In our experiment, we provided glucose to starved cells and monitored the import of GFP-Sfp1 in both normal (WT) and cells lacking the Rpb4 protein (*rpb4Δ*). We found that the import of GFP-Sfp1was slower in cells lacking Rpb4, which supports the idea that there is a connection between Sfp1 and Rpb4 (Fig. S1B-C).

Sfp1 is localized to P bodies

During starvation, Sfp1 localizes to discrete cytoplasmic foci. These foci represent P bodies - phase separated droplets containing RNAs, mRNA decay factors and other proteins (Luo et al., 2018) - as they contain the P body markers Dcp2 and Lsm1, as well as Rpb4 (Fig. 1B), and decrease in number if cells are pre-treated with cycloheximide before starvation (Luo et al., 2018) (Fig. 1C).

Sfp1p binds to a specific set of mRNAs

The observation that Sfp1 is exported in a transcription-dependent manner (Fig. 1A) raised the possibility that Sfp1 is exported in association with mRNAs, similar to Rpb4 (Duek et al., 2018; Goler-Baron et al., 2008). To examine this possibility, we performed a UV crosslinking and analysis of cDNA (CRAC) (Granneman et al., 2009). This approach is both highly reliable and can map the position within the RNA where the UV crosslinking occurs at nearly nucleotide resolution (Bohnsack et al., 2012; Haag et al., 2017). The endogenous *SFP1* gene was surgically tagged with an N-terminal His6-TEV-Protein A (HTP) tag, retaining intact 5' and 3' non-coding regions. Disruption of *SFP1* results in slow growth and a small cell size (Jorgensen et al., 2004). N-terminal tagging permitted normal cell size and wild-type growth (Fig. S2A, results not shown), indicating that the fusion protein is functional. The CRAC sequencing results were processed by a pipeline that Granneman's group has developed previously (van Nues et al., 2017). Metagene analysis of two biological replicates demonstrated a sharp peak toward the 3' end of the mRNAs, at a discrete relative distance (percentagewise) from polyadenylation sites (pAs) (Fig. 2A). Since the two replicates were very similar, we combined them for further analyses.

To identify mRNA targets that Sfp1 binds to, we took the top 600 mRNAs with the highest number of CRAC reads and performed k-means clustering of the signal around the pAs (Fig. 2B). We found four distinct clusters, three of which (C1-3) together accounted for 97.9% of the reads, distributed across 264 genes, spanning a region of 250 bp upstream of the pAs. The remaining mRNAs in the heatmap (C4), representing only 2.1% of the mapped CRAC reads, showed no clear accumulation of CRAC reads. The above 264 genes were named herein "CRAC positive (CRAC+)" genes (Table S2). We next assigned a relative value, named CRAC index, to each CRAC+ mRNA. This value reflects the number of CRAC RPKM normalized to its mRNA level (Table S2). Normalizing it to ChIP-seq signal yielded similar results (see sheet 2 in Table S2). We assume that these values reflect the propensity of Sfp1 to co-transcriptionally bind Pol II

transcripts. GO term enrichment analysis indicates that many CRAC+ genes encode proteins related to protein biosynthesis (Fig. S2B), including genes encoding translation factors, 70 ribosomal proteins (RP) ($p=7.7e^{-58}$) and 25 ribosome biogenesis (RiBi+RiBi-like) proteins ($p=7.2e^{-11}$). "RiBi-like" genes were previously defined as being activated by Sfp1, similarly to canonical RiBi. Specifically, Sfp1 binds to the upstream regions of both RiBI and Ribi-like genes and functions as an activator - as determined by anchor-away depletion followed by Pol II ChIP-seq (Albert et al., 2019). To obtain a more quantitative view, we crossed the list of CRAC+ genes with a previously published list of genes whose expression is altered by either Sfp1 depletion or overexpression (Albert et al., 2019), and found that 42% of the CRAC+ genes belong to a group whose transcription is upregulated by overexpressing Sfp1 ($p<2.2e^{-36}$) (Fig. S2C). The significant overlap between mRNAs that are both transcriptionally regulated and physically bound by Sfp1 suggest a link between these two Sfp1 activities.

Next, we performed motif analysis using DRIMust and MEME tools (see M&M). We focused on C1-C2 clusters that show a clear peak upstream of the pA sites. Out of 113 mRNAs in these clusters, we found 73 mRNAs with a short motif (GCTGCT, some with more than two GCT repeats) ($p < 7.925e^{-08}$), at an average distance of 150 bases upstream of the pA (Fig. S2F). Metagene analysis showed that Sfp1 tends to bind mRNAs in close proximity to the motif (Fig. 2C). The C3 cluster also contains many genes having similar motifs but not located at a distinct distance from the pA. Screen-shots of two representative genes with the GCTGCT motif are shown in Fig. S2G.

To corroborate the interaction of Sfp1 with CRAC+ mRNAs, we performed an RNA immunoprecipitation (RIP) analysis. Four out of the five mRNAs tested were co-immunoprecipitated with Sfp1 (Fig. 2D). Interestingly, despite being CRAC+ (albeit with low CRAC index), interaction of Sfp1 with *GPP1* mRNAs could not be detected. As a positive control, we performed Rpb4-RIP and found that all five pre-mRNAs were co-purified with Rpb4. Thus, either *GPP1* mRNAs does not bind Sfp1 and our CRAC analysis is not free of false positives, or its binding is relatively inefficient, hence its small CRAC index. Importantly, Sfp1 pulled down intron-containing *RPL30* pre-mRNA (Fig. 2D). Note that the ratio between pre-mRNA and mature mRNA is higher in the IPed lane than in the input one. Since introns are almost fully spliced out during transcription (Wan et al., 2020), these results strongly suggest that Sfp1 binds nascent RNA co-transcriptionally.

Sfp1 RNA-binding capacity is regulated by a promoter-located Rap1-binding sites

There is increasing evidence that promoters can affect mRNA stability, but the underlying mechanisms are unclear (Dahan and Choder, 2013; Haimovich et al., 2013a). Sfp1 is recruited to its cognate promoter by two alternative modes: either (i) directly through an A/T rich binding site or by Rap1 - directly or via Ifh1 (Albert et al., 2019; Reja et al., 2015). Previously, we reported that Rap1 also regulates decay of its transcripts (Bregman et al., 2011), raising the possibility that this feature is mediated by Rap1 capacity to recruit mRNA decay factor(s) (Sfp1?). To test the possibility that Rap1 function in mRNA decay is mediated by Sfp1 recruitment, we examined whether a Rap1 binding site (RapBS), coincided with Sfp1 DNA binding in many promoters (Reja et al., 2015 and references therein), is required for Sfp1 RNA binding. To this end, we compared the binding of Sfp1 to mRNAs derived from two similar plasmids, which transcribe identical mRNAs (Bregman et al., 2011). Each construct contains the *RPL30* transcription unit, including the 3' noncoding region. *RPL30* mRNA is one of the CRAC+ mRNAs that bind to Sfp1 (Table S2 and Fig. 2D). Transcription from both constructs is governed by the same TATA box (Fig. 3A). One of our constructs (''construct A'') contains an upstream activating sequence (UAS) that naturally lacks a RapBS *(ACT1*p), and the other (''construct B'') contains a UAS that naturally contains 2 RapBSs (*RPL30*p). To differentiate plasmid-borne from endogenous *RPL30* transcripts, we introduced a tract of oligo(G)18 into the 3' noncoding sequence of the plasmid-derived *RPL30* genes. An RNA immunoprecipitation (RIP) assay demonstrated that, despite being identical mRNAs (see (Bregman et al., 2011)), only the mRNA transcribed by construct B bound Sfp1 (Fig. 3B). To determine whether RapBS is responsible for Sfp1 RNA binding, we first introduced RapBS into construct A, creating "construct E", and found that the mRNA gained Sfp1-binding capacity (Fig. 3B, construct E). We next deleted the two RapBS from their natural position in the *RPL30* promoter, creating "construct F", and found that the mRNA could no longer bind to Sfp1 (Fig. 3B, construct F). Collectively, these results indicate that the capacity of Sfp1 to bind to RNA is mediated by promoter elements. RapBS is necessary and sufficient to mediate the interaction of Sfp1 with the gene transcript.

To obtain a more general perspective, we crossed the lists of RiBi (+ RiBi-like) mRNAs that bind to Sfp1 (RiBi CRAC+) (25 genes), non-CRAC+ RiBi (412 genes), and binding of Rap1 to their promoters as demonstrated by ChIP-seq (Lieb et at 2001) (Fig. 3C). While most RiBi

CRAC- promoters do not bind to Rap1, most RiBi CRAC^{+} do bind to Rap1 (p-value=1.9x10^{-22}). Moreover, RiBi non-CRAC+ exclude genes having Rap1 binding (Fig. 3C; p-value of exclusion is 1.5x10^{-5}). In contrast, RP genes show no statistical relationship with Rap1 binding because all of them (except for 10 genes) contain RapBS (Reja et al., 2015). Thus, the correlation between the presence of promoter RapBS and the capacity of Sfp1 to bind to mRNA is highly significant.

Sfp1 interacted also with endogenous RP mRNAs (Fig. 3B, lower panels). We also examined TAP-tagged Fhl1 and Ifh1 by RIP analysis. Unlike Sfp1, no interaction was detected between these two transcription factors and the examined mRNAs (Fig. 3B, lower panels).

Taken together, these results are consistent with the possibility that recruitment of Sfp1 to its cognate promoters, mediated by Rap1, is necessary for its capacity to co-transcriptionally interact with the transcript.

Sfp1 affects mRNA decay of CRAC+ genes

The interaction of Sfp1 with mRNA and with Rpb4, as well as its localization to P-bodies, suggests that Sfp1 plays a role in mRNA decay. To test the involvement of Sfp1 in mRNA synthesis and its possible role in mRNA decay, we performed a Genomic Run-On (GRO) assay (Garcia-Martinez et al., 2004) in WT, *sfp1*Δ and Sfp1-depleted cells. As expected, deletion of *SFP1* affected the mRNA synthesis rates (SR) and abundances (RA) of the ribosomal biosynthetic (RiBi) and ribosomal (RP) genes (Fig. 4A, see “SR” and “RA” panels). Calculating the mRNA half-lives (HL) revealed, unexpectedly, that deletion of Sfp1 led to reduced stability of CRAC+ mRNAs (Fig. 4A, “HL” panel) including RP genes and RiBi CRAC+. Depletion of Sfp1 by the auxin-induced degron system (AID) (Nishimura et al., 2009) for 20 min resulted in a decrease in synthesis rate (SR) of CRAC+ genes (Fig. 4B left, see shift of green spots distribution to the left), consistent with previous results (Albert et al., 2019) and references therein). This resulted in a proportional decrease in mRNA abundance (RA) ratios (RA in auxin treated / RA before treatment) (Spearman coefficient = 0.59). Forty min later, we observed a larger decrease in SR and RA of many genes (Fig. 4B right). RA of CRAC+ mRNAs decreased more than expected from just an effect on transcription (Fig 4B right, the green spots are scattered below the correlation line), suggesting that they were degraded faster than the rest of the transcriptome. This conclusion was corroborated by examining degradation rates of single CRAC+ mRNAs, either by comparing mRNA decay kinetics in *Δsfp1* and in WT (Fig. 4C, "endogenous genes"), or at 1 h after auxin

addition (Fig. S3A). As a control, we examined the degradation kinetics of mRNA with one of the lowest CRAC index (*MFA2*) and observed little effect of Sfp1. To further corroborate the high-throughput results, we took an additional single-gene approach. In this approach, we measured HL by naturally blocking transcription after shifting cells, which had proliferated at 30°C, to 42°C, and determined the rate of decrease in mRNA as a function of time post transcription arrest. We monitored the HL of two mRNAs with high CRAC index, *RPL28* and *RPL29* mRNAs. We chose *GPP1* and *PMA1* to represent mRNAs with low or no CRAC index respectively. Consistently with the other methods we used to determine HL, also in this case, we observed a substantial shortening of *RPL28* and *RPL29* mRNA HLs due to *SFP1* deletion (Fig. S3B, *RPL28* and *RPL29*). In contrast, the HLs of *PMA1* mRNAs was not affected by *SFP1* deletion, and that of GPP1 was little affected (Fig. S3A, *PMA1* and *GPP1*).

Consistent with Sfp1-binding as a prerequisite feature for its capacity to stabilize mRNAs, the calculated HLs of RP mRNAs and the 25 RiBi CRAC+, but not RiBi non-CRAC+ mRNAs and "Rest" (the bulk of the mRNAs), decreased more than those of the other genes due to *SFP1* deletion (Fig. 4A, "HL" panel). Remarkably, although Sfp1 binds RiBi promoters and stimulates their transcription (Albert et al., 2019), it only affects the stability of a small subset of them – the CRAC+ ones (see Fig. 4A HL panel "RiBi CRAC+"). Taken together, our data indicate that Sfp1 specifically stabilizes CRAC+ mRNAs. I think we want to check ChIP-exo signal and ChIP-seq and CheC-seq signals of Ribi non-CRAC+ genes.

Since RapBS is required for binding of Sfp1 to some mRNAs (Fig. 3B), we examined whether RapBS was also required for Sfp1-mediated mRNA stability. To this end, we analyzed the effect of deleting *SFP1* on the stability of these identical plasmid borne mRNAs. Sfp1 did not affect the stability of an mRNA whose synthesis was driven by a promoter that lacks RapBS (Fig. 4C, left panel); yet, introducing just RapBS to the same promoter resulted in a transcript whose stability was dependent on Sfp1 (Fig. 4C, right panel).

The first step in mRNA decay is shortening of the mRNA poly(A) tail (Parker, 2012). Polyacrylamide gel electrophoresis northern (PAGEN) has been used to determine deadenylation rates (Lotan et al., 2005; Sachs and Davis, 1989). Using this technique, we found that Sfp1 slows down the mRNA deadenylation step of a few CRAC+ mRNAs, as well as their subsequent degradation (Fig. 4D). These findings suggest that Sfp1 plays a role in deadenylation and decay of CRAC+ mRNAs.

Collectively, we found that binding of Sfp1 to mRNAs slows down their deadenylation and stabilizes them. Promoter binding of Sfp1 via Rap1 is necessary for its mRNA binding (Fig. 3) and mRNA stability (Fig. 4C); however, since Sfp1 binds most RiBi promoters, not via Rap1 (Fig. 3C), but not RiBi transcripts (most RiBi are not CRAC+, see Fig. 3C) and does not affect HLs of most RiBi mRNA (Fig 4A), promoter binding is not sufficient for its mRNA binding and mRNA degradation activities. Perhaps, promoter binding of Sfp1 via Rap1 is necessary and sufficient for its RNA binding.

Sfp1-bound mRNAs are transcribed by Sfp1-bound genes in a manner that suggest a mechanistic linkage

Sfp1 is a well-known transcription factor that binds to promoters of its cognate genes (see Introduction). Our discovery that Sfp1 also binds to CRAC+ mRNAs, which are only encoded by a fraction of the Sfp1 target genes (see previous sections), prompted us to investigate whether the chromatin binding feature of Sfp1 with CRAC+ genes is different from its binding with other gene targets. To do so, we leveraged published ChIP-exo datasets (Reja et al., 2015) and found that Sfp1 binds not only to promoters but also to gene bodies of CRAC+ genes (Fig. 5A). In contrast, Rap1 binds almost exclusively to promoters of CRAC+ genes (Fig. 5A and Fig. S4A). ChIP-exo data of a random subset of non-CRAC+ genes showed that Sfp1 binds weakly to both promoters and gene bodies (Fig. 5A, “Sfp1 control”). However, because promoter binding is not higher than binding to other chromatin regions, it is unclear whether this binding is specific (but see below).

Importantly, the intensity of Sfp1 binding to the bodies of CRAC+ genes correlates with the number of transcriptionally active Pol II per gene (this is what BioGRO-seq measures) (Fig. 5B, left), suggesting that the binding of Sfp1 to gene bodies occurs through the binding of active Pol II. However, in addition to active Pol II, the chromatin accommodates transcriptionally arrested Pol II, most of which is presumed to be in a backtracked configuration. If, in addition to its binding to active Pol II, Sfp1 remains bound to backtracked Pol II, it is expected that the correlation of Sfp1 occupancy with the occupancy of all Pol II molecules would be even higher than it is with just active Pol II. Comparing the left and right panels of Fig. 5B indicates that this is the case. Interestingly, we found a modest correlation between Sfp1 binding to the bodies of all genes (excluding CRAC+) and their BioGRO-seq signal (Fig. 5B, left panel). This weak binding could indicate non-specific attachment to accessible chromatin, which characterizes

transcriptionally active regions. However, because over-expression of Sfp1 affects transcription of ~30% of the genes (Albert et al., 2019), we suspect that the weak binding of Sfp1 to the bodies of these genes is also through Pol II interaction.

Albert et al. (2019) discovered that the complete set of ~500 gene promoters bound by Sfp1 could be revealed only by a combination of ChIP-seq and ChEC-seq ("chromatin endogenous cleavage", using micrococcal nuclease [MNase] fused to Sfp1) methods. These two methods revealed the two distinct promoter binding modes of Sfp1, discussed earlier. We compared the Sfp1 ChIP-exo and ChEC-seq metagene profiles, which were obtained in different laboratories, and found, in both methods, substantial differences between Sfp1 binding to CRAC+ and non-CRAC+ genes (Fig. S4B). Specifically, in both methods there is a substantial difference between CRAC+ and non-CRAC+. In ChIP-exo the difference is considerable. Focusing on ChEC-seq profiles, there seems to be a difference in Sfp1 positions in the CRAC+ promoters relative to the TSS with that of non-CRAC+ (Fig. S4B, see the centers of the orange and the red peaks). Taken together, these results suggest that the binding of Sfp1 to CRAC+ promoters is different from that to non-CRAC+ promoters. This difference is in addition to that found in gene bodies, observed by ChIP-exo (Figs 5A and S4B).

The presence of Sfp1 along gene bodies of CRAC+ genes raised a possibility that this feature could be related to its capacity to bind their transcripts. To explore this possibility, we examined Sfp1 ChIP-exo signal of CRAC+ genes from -500 bp up to the pA sites. We found a decline in the signal that coincided with the position of the GCTGCT (Fig. 5C), which is mapped ~160 bp upstream of the pA sites (Fig. S2F). To obtain a wider view, we examined ChIP-exo signal at ± 500 bp around the GCTGCT motif and observed similar drop in the signal around the motif (Fig. S2D). We also examined non-CRAC+ genes and found a decline of Sfp1 ChIP-exo signal around the the pAs. However, this decline was milder than that of CRAC+ and occurs closer to the pAs (Fig. S2E). These results, in combination with the dependence of Sfp1-mRNA binding on RapBS (Fig. 3B-C), strongly suggest that Sfp1 dissociates from the chromatin and binds the RNA co-transcriptionally. Other observations that are consistent with co-transcriptional binding are outlined in the Discussion.

Genes transcribing Sfp1-bound ("CRAC+") mRNAs exhibit higher levels of Pol II backtracking

Sfp1 negatively affects transcription elongation, as deletion of *SFP1* results in higher Pol II elongation rate (Begley et al., 2019). Consistently, Sfp1 seems to enhance Pol II backtracking as a lack of *SFP1* suppresses the effect of TFIIS deletion on transcription (Gómez-Herreros et al., 2012a). A characteristic feature of backtracked Pol II is the displacement of nascent RNA 3' end from the active site (Cheung and Cramer, 2011). Consequently, backtracked Pol II cannot elongate transcription in GRO assays (Jordán-Pla et al., 2015; Pelechano et al., 2009). Therefore, a Pol II backtracking index (BI) can be determined by comparing Pol II ChIP (total Pol II) and GRO (actively elongating Pol II – not backtracked) signals. To determine the impact of Sfp1 on backtracking, we performed Rpb3 ChIP and GRO analyses in WT and *sfp1*Δ cells. As expected from a TF that stimulates transcription initiation, Rpb3 ChIP signals were higher in WT than in *sfp1*Δ cells (y=0.578x) (Fig. 6A, left). However, despite different Pol II occupancy, GRO signals were similar in WT and *sfp1*Δ strains (y=0.844x) (Fig. 6A, right), consistent with Sfp1-mediated stimulation of Pol II backtracking (in WT cells). This effect was particularly intense in genes whose mRNAs bind Sfp1 (CRAC+) (green dots in Fig. 6A). Consequently, BI values of CRAC+ genes were higher than the average (Fig. 6B), suggesting a linkage between backtracking and Sfp1 imprinting (see below). Perhaps, the backtracked configuration is compatible with the movement of Sfp1 from Pol II to the nascent transcript (see Discussion).

We note that CRAC+ genes cover a wide range of transcription levels (see Fig 6A), indicating that BI and transcription level are unrelated. Moreover, among CRAC+ genes, BI of RiBi genes whose transcripts bind Sfp1 (RiBi CRAC+) exhibited high BI with strong dependence on Sfp1, whereas RiBi genes that encode mRNAs that are not bound by Sfp1 ("RiBi non-CRAC+") did not (Fig. 6B).

Sfp1 disruption affects both mRNA HL (Figs. 4 and 6C) and BI (Fig. 6B and C). In both cases, these effects were significantly stronger for the CRAC+ group. We were intrigued by a possible link between these two effects, as it raises a possible role for backtracking in imprinting-mediated mRNA decay. In favor of this idea, we found a modest correlation between the effect of Sfp1 on HLs and on BIs (Fig. 6D; r=0.366), which was higher for the CRAC+ genes (Fig. 6D: green spots, r=0.423). This suggests that the interaction of Sfp1 with chromatin has an effect on both Pol II elongation and mRNA HLs, suggesting a mechanistic linkage.

Sfp1 locally changes Rpb4 stoichiometry or configuration during transcription elongation

The yeast Rpb4 is a sub-stoichiometric Pol II subunit (Choder, 2004; Choder and Young, 1993), which enhances Pol II polymerization activity (Fischer et al., 2020; Rosenheck and Choder, 1998) and promotes mRNA instability due to co-transcriptional imprinting (Goler-Baron et al., 2008; Lotan et al., 2007, 2005). Since Sfp1, which physically interacts with Rpb4 (Fig. S1A), also affects mRNAs stability via imprinting (Figs. 2D and 4), we examined whether Sfp1 and Rpb4 activities are interrelated. As a proxy for Rpb4 stoichiometry, we calculated the Rpb4-ChIP/Rpb3-ChIP ratio and observed a gradual decrease as Pol II traversed past the TSS, decreasing even further towards transcript pAs (Fig. 7A, WT). These data suggest that Rpb4 gradually dissociates from Pol II during elongation (possibly concomitantly with transcript binding). However, this result could also be attributed to changes during transcription elongation in the capability of Rpb4 to contact DNA, or to changes in ChIP antibody accessibility that occur during elongation (referred to herein as "Rpb4 configuration"). Global changes in Rpb4-ChIP/Rpb3-ChIP ratio were stronger in the highly transcribed genes in WT (Fig. 7A cf "Q1" and "Q4"), suggesting a linkage to Pol II activity; i.e., the more active Pol II is the more likely it is to gradually lose Rpb4. Interestingly, the CRAC+ genes exhibited an even stronger change in Rpb4 stoichiometry/configuration (Fig. 7A, CRAC+) despite not being highly transcribed in most cases (Fig 6A), suggesting that this Rpb4 feature is modulated by Sfp1 molecules that are capable of binding CRAC+ mRNA. Indeed, these changes in Rpb4 stoichiometry/configuration were highly dependent on Sfp1, as they were abolished in *sfp1*Δ for all gene sets analyzed (Fig 7A, *sfp1*Δ). Again, the effect of *sfp1*Δ was more pronounced in CRAC+ than in non-CRAC+ genes (Fig 7A, all panels). In summary, we detected a general Sfp1-dependent alteration in Rpb4 stoichiometry/configuration with Pol II during elongation, which was more pronounced in CRAC+ genes.

We also found a general inverse correlation between the Rpb4-ChIP/Rpb3-ChIP ratio and BI in the WT, which was stronger in CRAC+ genes (Fig. 7B), suggesting that Rpb4 stoichiometry/configuration might affect Pol II backtracking, or vice versa –backtracking enhances changes in Rpb4 stoichiometry/configuration. This correlation was clearly reduced in *sfp1*Δ cells, which also showed almost no difference between CRAC+ and non-CRAC+ genes (Fig 7B cf left and right panels). Accordingly, the correlation between Sfp1 effects on BI and Rpb4-ChIP/Rpb3-ChIP ratio was tripled in CRAC+ than in non-CRAC+ genes (Suppl. Fig. S5A). These results suggest that Sfp1 activity during transcription elongation, Pol II backtracking, and Rpb4 stoichiometry/configuration are connected.

Finally, focusing on the CRAC+ genes in the WT strain, we found a modest correlation between Rpb4 ChIP/Rpb3 ChIP ratio and mRNA HL (Fig. 7C, r=0.488). Interestingly, this correlation was specific to CRAC+ genes (very little correlation was found for the rest of the genes marked by gray spots) and was dependent on Sfp1 (Fig. 7C, *sfp1*Δ). Accordingly, the correlation between Sfp1 effects on mRNA HL and Rpb4-ChIP/Rpb3-ChIP ratio was significant in CRAC+ but absent in non-CRAC+ genes (Suppl. Fig. S5B). Therefore, the link between Rpb4 stoichiometry/configuration and mRNAs HLs of CRAC+ genes is entirely dependent on Sfp1, whereas the correlation between the Rpb4-ChIP/Rpb3-ChIP ratio and BI is still present and significant in *sfp1*Δ (Fig. 7B). This suggests that the impact of Sfp1 on mRNA stability is preferentially mediated by Rpb4 association, rather than by Pol II backtracking.

Taken together, we propose that the effects of Sfp1 on mRNA stability and on Pol II elongation are linked. Our results are compatible with a model whereby Sfp1 impact on Rpb4 during elongation provokes mRNA imprinting, which affects stability, and concomitantly produces Pol II backtracking.

Discussion

Sfp1 has traditionally been regarded as a "classical" transcription factor (TF) known to bind specific promoters, either through interactions with other TFs or by directly binding to promoter DNA (Albert et al., 2019; Reja et al., 2015). Further analysis that we performed using published ChIP-exo and ChEC-seq data has revealed an additional aspect of Sfp1's binding behavior—it also interacts with gene bodies. Notably, the binding of Sfp1 to CRAC+ gene bodies correlates with the presence of transcriptionally active Pol II molecules (Fig. 5B), suggesting a physical association between Sfp1 and Pol II. Consequently, the binding of Sfp1 has an impact on Pol II configuration, stoichiometry, and backtracking, particularly in the context of CRAC+ genes (Figs. 6 and 7). This alteration in Pol II configuration is reflected in the increased backtracking frequency, as indicated by the BI values. Moreover, the binding of Sfp1 to gene bodies correlates with Pol II activity, particularly within CRAC+ genes (Fig. 5B). Thus, Sfp1's function in a subset of its target genes extends beyond transcription initiation and encompasses transcription elongation, as we proposed earlier (Gomez-Herreros et al, 2012b). Based on the following observations, we propose that Sfp1 binds to Pol II in proximity to both DNA and Rpb4, accompanying Pol II during elongation and influencing its configuration, thereby enhancing its propensity to backtrack (Fig. 8A): (I) Sfp1 physically interacts with Rpb4, either directly or indirectly (Fig. S1). (II) ChIP-exo results, dependent on Sfp1's cross-linking with DNA, indicate Sfp1's presence near DNA, both at promoters and gene bodies (Fig. 5A). (III) Sfp1 influences Pol II configuration in a manner that impacts the architecture or stoichiometry of Rpb4 within Pol II (Fig. 7). (IV) Sfp1 binding to chromatin better correlates with all Pol II (both backtracked and active) than with transcriptionally active Pol II, examined by BioGRO-seq (Fig. 5B). (V) The global reduction in BI is observed upon the deletion of *SFP1* (Figs. 6 and 7).

The effects of Sfp1 on backtracking occur in the context of Rpb4 (Fig. 7B). We note that backtracking is influenced by alteration of Rpb4 within the elongation complex, independently of Sfp1, as the correlation between Rpb4 ChIP/Rpb3 ChIP ratios and BI is still present, albeit more mildly, in *sfp1*Δ. Rpb4/7 was previously implicated in backtracking either by itself (Fischer et al., 2020) or in the context of the Ccr4-NOT complex (Babbarwal et al., 2014; Kruk et al., 2011a). It is possible that the effect of Sfp1 on backtracking is through changing Rpb4 stoichiometry within

Pol II or Rpb4 conformation. Backtracking is a conserved process across evolution that influences Pol II processivity and transcription rate (Bar-Nahum et al, 2005). We propose an additional potential function: the backtracking-prone configuration of Pol II might be more compatible with mRNA imprinting than the 'regular' configuration (Fig. 8A). This hypothesis is prompted by the following observations: (I) Sfp1 shows a positive general influence on mRNA stability that is maximal in those mRNAs that it physically binds (Fig. 4A “HL”, 4B and 6C). (II) Sfp1 physical interaction with mRNAs occurs cotranscriptionally (Fig. 5C and S2D). (III) The stronger the effect of Sfp1 on Pol II backtracking the stronger its effect on mRNA HL (Fig. 6D). (IV) Backtracking is substantially more frequent in those genes whose mRNAs are bound by Sfp1 compared to the genome average, and significantly more dependent on this factor (Fig. 6B and 6C). Taking all these observations together, it appears that backtracking is associated with co-transcriptional binding of Sfp1 to nascent RNA. Reassuringly, RiBi genes that exhibit Sfp1-dependent backtracking also imprint Sfp1, whereas those that do not also do not imprint (cf Fig. 6B, "RiBi CRAC+" and "RiBi non-CRAC+").

Unexpectedly, in addition to its chromatin binding feature, we find that Sfp1 binds a subpopulation of mRNAs whose transcription is stimulated by Sfp1 – “CRAC+ mRNAs”. CRAC+ mRNAs exhibit distinctive characteristics compared to other Sfp1-regulated genes (whose mRNAs do not bind Sfp1), or other non-Sfp1-regulated genes (Figs. 2A-C, 3C, 4A, 4B, 5A-C, 6B-D, 7A-B, S2D, S2F, S4A-B, S5A-B). Collectively, these findings demonstrate that CRAC+ genes and their transcripts constitute a unique group, distinguishable from others based on several criteria, only one of them is the ability of their transcripts to bind Sfp1. CRAC+ genes display a broad range of transcription rates (Fig. 6A) and behave differently from the most highly transcribed genes (Fig. 7A). Furthermore, while all RiBi mRNAs exhibit similar expression levels, only a subset of them binds to Sfp1. These observations collectively suggest that the 264 CRAC+ constitute a distinct set of genes that unveils a new paradigm: genes controlled by a common factor both transcriptionally and post-transcriptionally, via mRNA imprinting.

We observed that Rap1 binding site (RapBS)-mediated promoter binding is critical for Sfp1 capacity to bind RNA (Fig. 3). In RiBi genes, we found highly significant correspondence between the propensity of RiBi promoters to carry RapBS and binding of RiBi mRNAs to Sfp1 (Fig. 3C). We propose that RNA-binding occurs co-transcriptionally for the following reasons: (I) Sfp1 export from the nucleus to the cytoplasm is dependent on transcription (Fig. 1A), suggesting

that it is exported together with the Pol II transcripts. (II) Splicing occurs co-transcriptionally (e.g., (Churchman and Weissman, 2011). Consistent with co-transcriptional binding, Sfp1 binds intron-containing *RPL30* RNA (Fig. 2D; note that the ratio between intron-containing RNA and mature RNA is higher in the IPed lane than in the input one). (III) Binding of Sfp1 to mRNA is dependent on RapBS, suggesting that the same Sfp1 recruited to the promoter by Rap1 binds the transcript. (IV) The ChIP-exo signal drops past the GCTGCT motif in C1-C2 genes (for clusters' definition see Fig. 2B), a position where Sfp1 prefers to bind the motif containing CRAC+ transcripts (Fig. 5C). These observations suggest that, at least in some of their mRNA targets, Sfp1 is released from Rpb4-containing Pol II to the nascent transcripts co-transcriptionally as GCTGCT motif emerges from Pol II. However, many CRAC+ do not contain detectable GCTGCT motif or have it in a different position (Fig. 2B, C3 cluster). Our observation that RapBS is required to promote binding to *RPL30* mRNA sequence, a gene that contains the motif, (Fig. 3B) demonstrates that motif 1 is not sufficient to recruit Sfp1. Perhaps this motif is used to stabilize the interaction. In the absence of the motif, the movement from Rpb4-containing Pol II to the emerging transcript might be related to the process of polyadenylation (see Fig. 2A). Collectively, our results unveil a role for CRAC+ promoters as mediators between RPB and its interacting RNAs. Whether Rpb4 dissociates together with these proteins remains to be determined. Whether the probability of mRNA imprinting is affected by the time Pol II is engaged in backtracking, until its resolution by TFIIS, is another appealing hypothesis that remains to be examined.

Binding of Sfp1 to mRNAs regulates deadenylation-mediated mRNA decay, mainly by slowing down these processes (Fig. 4). Following mRNA decay, Sfp1 is imported back to the nucleus (Fig. S1B-C). Following import, Sfp1 binds to specific promoters and regulates transcription, closing the circle of gene expression regulation (Fig. 8B). In this way, the synthesis and decay of Sfp1-regulated mRNAs is coordinated in a manner that maintains proper mRNA levels of a specific subset of genes.

The capacity of Sfp1 to bind mature mRNA adds additional complexity to the expression of CRAC+ genes. Here we report that Sfp1 regulates the levels of these gene products by two mechanisms: by stimulation of synthesis and by repression of decay. This raises a possible new mode of regulating mRNA level that targets the Sfp1 imprinting machinery, the extent of which would determine the mRNA HL and consequently the mRNA level. Since Sfp1 enhances both mRNA synthesis and stability, it can serve as a signaling pathway target to rapidly regulate

expression of its clients. Indeed, expression of Sfp1 targets is highly responsive to the environment (Albert et al., 2019; Jorgensen et al., 2004).

Sfp1 was viewed as a classical TF that specifically modulates the transcription of a subset of a few hundred genes (see Introduction). We were therefore surprised to discover a genome-wide effects of Sfp1 depletion or deletion (Figs. 4B, 6A, 6D, 7B-C). These results, in combination with the weak interaction of Sfp1 with chromatin (Fig. 5A "Sfp1 Control"), are in accord with previous observation that expression of >30% of the genes, most are not considered to be direct targets of Sfp1, is affected by either Sfp1 depletion or its overexpression (Albert et al., 2019), and our unpublished NET-seq data). An important support for this notion is the correlation we found between the binding of Sfp1 to gene bodies and Pol II occupancy in all genes (Fig. 5B). The general capacity of Sfp1 to alter elongating Pol II in the context of Rpb4 and to affect backtracking (Fig. 7B, S5A) likely underlies its widespread effect on transcription.

As mentioned above, this widespread effect of Sfp1 on backtracking correlates with higher mRNA HL in many genes (not only in CRAC+; Fig. 6D). This suggests that the effect of Sfp1-dependent backtracking on mRNA HL does not always require Sfp1 binding to mRNA. Perhaps, this indirect effect of Sfp1 on mRNA HL is mediated by other factors, like Ccr4-Not, whose imprinting may also be affected by Pol II backtracking ((Kruk et al., 2011b; Villanyi et al., 2014; Begley et al, 2019). In contrast, the correlation between Rpb4/Rpb3 and mRNA HL is much higher in CRAC+ than in the rest of the genome, and is entirely dependent on Sfp1 (Fig. 7C, S5B). Interestingly, we previously demonstrated that transcription elongation of Rap1-controlled genes is exceptionally affected by backtracking (Pelechano et al., 2009). We propose that Rap1-dependent recruitment of Sfp1 modulates its interaction with Pol II in a manner that permits Sfp1 transfer to mRNA (Fig. 8).

Binding of Sfp1 to CRAC+ mRNAs, which stabilizes the bound mRNAs, is dependent on Rap1 binding sites. On the other hand, we previously found that depletion of Rap1 leads to mRNA stabilization (Bregman et al., 2011). If the effect of Rap1 on mRNA stability is mediated only through Sfp1, we would expect that its depletion should result in mRNA destabilization, contrary to our previous observation (Bregman et al., 2011). Therefore, we conclude that Rap1 mediates additional activity that enhances mRNA decay and override the destabilizing effect of Sfp1, which remains to be discovered.

The Rpb4 stoichiometry within yeast Pol II is <1 (Choder and Young, 1993). It was not clear whether the sub-stoichiometric feature of Rpb4 is constant across all Pol II molecules. Here we show that stoichiometry, as reflected by the Rpb4 ChIP/Rpb3 ChIP ratio, decreases gradually during transcription elongation. It is possible that there is a link between the Rpb4 capacity to modulate backtracking and the gradual decrease in its stoichiometry. Interestingly Sfp1 affects this stoichiometry drop, thus serving as a regulator of Rpb4 stoichiometry. Moreover, the effect of Sfp1 on mRNA stability is entirely linked to its action on Rpb4 stoichiometry (Fig. 7C and S5B). We note, however, that, although changes in the stoichiometry is the most likely interpretation of the change in Rpb4 ChIP/Rpb3 ChIP ratio, this change can merely reflect changes in Pol II configuration that compromise the capacity of Rpb4 to produce ChIP signal.

Interestingly, the Sfp1-mediated mechanism contrasts with that of mRNA buffering. The known model of mRNA buffering posits that a modulation of one process (e.g., transcription) is balanced by a reciprocal modulation of the other process (e.g., mRNA decay), thus maintaining the mRNA level constant, or nearly constant (Bryll and Peterson, 2023; Haimovich et al., 2013b, 2013a; Pérez-Ortín and Chávez, 2022; Timmers and Tora, 2018). In contrast, the functions of Sfp1 do not result in balancing, but the contrary: the two activities of Sfp1 cooperatively increase mRNA level – increase mRNA synthesis and decrease its degradation. Thus, Sfp1 is a paramount example of enhanced gene regulation by cooperation between mRNA decay and gene transcription that we previously modelled(García-Martínez et al., 2023). Moreover, we reported that mRNA buffering maintains a constant mRNA concentration regardless of the strain growth rate, except for growth-related genes (García-Martínez et al., 2016; Pérez-Ortín and Chávez, 2022). The expression of the latter genes increases with growth rate. The discovered effect of Sfp1 on its target genes, most of which are growth related genes, provides a plausible mechanism to explain how growth-related genes avoid buffering when their expression must be rapidly adjusted to the ever-changing environment.

In summary, we propose that the role of certain class-specific TFs, such as Sfp1, extends beyond merely controlling various stages of mRNA synthesis and processing in the nucleus; they also regulate post-transcriptional functions in the cytoplasm. We hypothesize that a single factor (or a complex of factors) has evolved the capacity to regulate both transcriptional and post-transcriptional functions to facilitate cross-talk between these two mechanisms. Sfp1 possesses two classical zinc fingers. Given that proteins with zinc fingers, characteristic of many TFs, are

known to be involved in DNA, RNA, and protein binding, it is conceivable that this zinc finger-containing protein is suitable for participating in the mRNA imprinting process, as proposed here (Fig. 8B). Given that RBPs have been proposed to perform multiple roles in RNA-based regulation of transcription at the level of chromatin and gene promoters in mammals (Xiao et al., 2019) and the high abundance of zin-finger TFs in humans (Lander et al., 2001), we anticipate that the case of Sfp1 as an imprinting factor will serve an example of this important type of gene regulators in other eukaryotes.

Materials and Methods

Yeast Strains and Plasmids construction

Yeast strains, and plasmids are listed in Supplementary Table S1. HISx6-TEV protease site-Protein A (HTP) tag was inserted into the chromosome by homologous recombination with PCR amplified fragment carrying *SFP1p::CaURA3::SFP1p*:: HTP::*SFP1* ORF (only 100 bp repeats of the promoter and 100 bp of the ORF). After integration into *SFP1* locus, the *Ca::URA3* was popped out by selection on 5-FOA, utilizing the two identical *SFP1p* repeats, recreating the 5' non-coding region. In this way, the tag was surgically introduced without adding other sequences. All strains were verified by both PCR and Sanger sequencing.

Yeast proliferation conditions, under normal and fluctuating temperatures, and during exit from starvation

Yeast cells were grown in synthetic complete (SC), Synthetic dropout (SD) or in YPD medium at 30°C unless otherwise indicated; for harvesting optimally proliferating cells, strains were grown for at least seven generations in logarithmic phase before harvesting.

Genomic run-on and degron procedures

Genomic run-on (GRO) was performed as described in (García-Martínez et al., 2004), as modified in Oliete-Calvo et al. (2018). Briefly, GRO detects by macroarray hybridization, genome-wide, active elongating RNA pol II, whose density per gene is taken as a measurement of its synthesis rate (SR). Total SR for a given yeast strain was calculated as the sum of all individual gene SRs. At the same time, the protocol allows the mRNA amounts (RA) for all the genes to be measured by means of the hybridization of labeled cDNA onto the same nylon filters. Total mRNA concentration in yeast cells was determined by quantifying polyA+ in total RNA samples by oligo-dT hybridization of a dot-blot following the protocol described in (García-Martínez et al., 2004). mRNA half-lives, in arbitrary units, are calculated as RA/SR by assuming steady-state conditions for the transcriptome. All the experiments were done in triplicate.

For the degron experiment, the auxin degron strain (AID-Sfp1) was inoculated in YPD medium until a OD_{600} of 0.4 and then an aliquot was taken, corresponding to t=0. After auxin

addition two more samples were taken at 20 and 60 minutes. At each time point (0, 20 and 60 minutes) Genomic Run-on (GRO) was performed, as described in the previous paragraph, and for SR, RA and HL data were obtained for each time point.

GEO accession numbers for the genomic data are: GSE57467 and GSE202748.

Performing and analyzing the UV-cross-linking and analysis of cDNA (CRAC)

The endogenous Sfp1 gene was surgically tagged with an N-terminal ProteinA-TEV-His6 (PTH)tag, retaining intact 5' and 3' non-coding regions, as described above. Two WT independent colonies and two *rpb4Δ* strains (yMC1019-1022) were subject to CRAC , utilizing a previously reported protocol (Haag et al., 2017). Briefly, yeast cells were allowed to proliferate in a synthetic complete medium at 30°C till $1x10^7$ cells/ml, UV irradiated and harvest and freeze in liquid nitrogen. Frozen cells were pulverized cryogenically using a mixer mill 400 (RETSCH). PTH-Sfp1 was isolated, RNA fragments cross-linked to Sfp1 were end labeled and purified, converted to a cDNA library and sequenced by next generation sequencing NGS 500 (all sequence data are available from GEO under accession number **GSE230761**).

The CRAC sequencing results were processed by the CRAC pipeline, using the pyCRAC package (PMCID: PMC4053934) for single-end reads that the Granneman lab previously developed (van Nues et al., 2017). Briefly, low quality sequences and adapter sequences were trimmed using Flexbar (version 3.5.0; PMID: 24832523). Reads were then collapsed using random barcode information provided in the in-read barcodes using pyFastqDupicateRemover.py from the pyCRAC package (PMCID: PMC4053934). Collapsed reads were aligned to the yeast reference genome (R64) using novoalign (www.novocraft.com) version 2.0.7. Reads that mapped to multiple genomic regions were randomly distributed over each possible location. PyReadCounters.py was then used, makes read count and fragments per kilobase transcript per million reads (FPKM) tables for each annotated genomic feature.

Bioinformatics analysis of CRAC, ChIP-exo and ChEC-seq datasets

The CRAC sequencing results were processed by a pipeline that we developed previously (van Nues et al., 2017). High quality reads were mapped against the Saccharomyces cerevisiae sacCer3 (R64) reference genome. ChIP-exo (Reja et al., 2015) and ChEC-seq (Albert et al., 2019) raw sequencing datasets were retrieved from accessions PRJNA245761 and PRJNA486090,

respectively. All fastq files were also aligned to the *Saccharomyces cerevisiae* R64 reference genome with Bowtie2, using default parameters. Alignment files were used to generate average metagene profiles with the ngs.plot R package (Shen, L., et al. 2014). The robust parameter -RB was used in every plot to filter genes with the 5 % most extreme coverage values. Ngs.plot was also used to generate CRAC signal heatmaps around pA sites, and the -GO km function was used to perform k-means clustering of genes with similar profiles. Venn diagrams were generated in R, and the overlap was statistically evaluated for overrepresentation or depletion using the hypergeometric test. For motif analysis, 160 bp sequences around pA sites were extracted in fasta format from the 113 CRAC+ genes in clusters 1 and 2 (Fig 2B) with the sequence tools function from the RSAT website (http://rsat.france-bioinformatique.fr/fungi/) and used with two sets of motif discovery tools: MEME (https://meme-suite.org/meme/tools/meme), and DRImust (http://drimust.technion.ac.il). For MEME we used the differential enrichment mode, establishing a control set of 262 sequences belonging to non-CRAC+ genes randomly selected but excluding RP and RiBi. We set up a minimum motif size of 6 nt, and a maximum of 50 nt. We also restricted the search to the same strand of the gene analyzed. For DRImust, we also used a target vs background sequences strategy, a strand-specific search mode and the same minimum and maximum motif lengths as MEME. Both tools encountered the same GCTGCT consensus motif, but in overlapping but slightly different sets of genes. We fused the lists of genes from both tools for downstream analysis. GO term enrichment analysis (biological process ontology) was done with the enrichment tool from Yeastmine (https://yeastmine.yeastgenome.org/yeastmine/begin.do), selecting only GO terms with Holm-Bonferroni-corrected p-values < 0.05. Spearman correlation values and statistical test result asterisks were inserted into ggplot2-generated plots by using the package ggpubr (https://rpkgs.datanovia.com/ggpubr).

ChIP on chip experiments

Chromatin immunoprecipitation was performed as previously described (Rodriguez-Gil et al, 2010) using ab81859 (Abcam) anti-Rpb3 and 9E10 (Santa Cruz Biotechnology) anti-C-Myc antibodies. After crosslinking reversal, the obtained fragments (300 bp approximately) of enriched DNA were amplified unspecifically and labelled following Affymetrix Chromatin Immunoprecipitation Assay Protocol P/N 702238. Genomic DNA controls were processed in

parallel. 10 µl of each sample were amplified using Sequenase ™. The reaction mix consisted of 10 µl purified DNA, 4µl 5X Sequenase ™ reaction buffer and 4µl Primer A (200µM) for each reaction. The cycle conditions for random priming were 95°C for 4 minutes, snap cool on ice and hold at 10°C. Next, 2.6µl of "first cocktail" (0.1µl 20mg mg/ml BSA), 1µl 0.1M DTT, 0.5µl 25 mM dNTPs and 1µl diluted Sequenase ™ 1/10 from 13U/µl stock) were added to each reaction and put back in the thermocycler for the following program: 10°C for 5 minutes, ramp from 10°C to 37°C over 9 minutes, 37°C for 8 minutes, 95°C for 4 minutes, snap cool on ice and 10°C hold. Then another 1µl of cocktail was added to each sample and these steps were repeated for two more cycles. The samples were kept at a 4°C hold. After PCR amplification with dUTP, the samples were purified using Qiagen QIAquick PCR Purification Kit (50) (Cat.No. 28104). About 56µl of first round purified DNA were collected for each reaction. The amplification PCR was performed as usual but using 20µl of first-round DNA from the previous step, 3.75µl of a 10mM dNTPs + dUTP mix and 4µl of 100µM Primer B. The cycling conditions were: 15 cycles consisting of 95°C for 30 seconds, 45°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 minute, and 15 cycles of 95°C for 30 seconds, 45°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 minute, adding 5 seconds for every subsequent cycle. DNA quality and quantity were checked in a 1% agarose gel and using a NanoDrop ND-1000 Spectrophotometer. The samples were purified using the QIAquick PCR purification kit (Qiagen). Then 0.5 µg of each were used to hybridize GeneChip *S. cerevisiae* Tiling 1.0R custom arrays. This step was carried out in the Multigenic Analysis Service of the University of Valencia. The obtained CEL archives were normalized by quantile normalization and the intensities of the signal were extracted using the TAS (Tiling Analysis Software) developed by Affymetrix. The resulting text files were read using R scripts to adjudicate probe intensities to genes. The $\log_2$ values of the median intensities of the chosen different group of genes were represented. In order to compare the data between different experiments the values were normalized by median and standard deviation (z-score $\log_2$).

mRNA decay assay

Assay was performed as described (Lotan et al., 2005). Briefly, optimally proliferating cells were treated with thiolutin, or shifted to 42°C, to stop transcription, and cells samples were harvested at various time point post transcription arrest. RNA was extracted and equal amounts of total RNA were loaded on an agarose gel for Northern blot hybridization, using radiolabeled DNA probes.

The same membrane was sequentially hybridized with the indicated probes. SCR1 RNA (Pol III transcript) and rRNAs were used for normalization. For quantification: The mRNAs levels were determined by PhosphoImager technology. mRNA level was normalized to that of the Pol III transcript *SCR1*. Band intensity at time 0, before transcription inhibition, was defined as 100% of initial mRNA level and the intensities at the other time points were calculated relative to time 0.

Polyacrylamide Northern (PAGEN) analysis

PAGEN was performed as described previously (Lotan et al., 2005; Sachs and Davis, 1989). Briefly, equal amounts of total RNA pellets were suspended in ~15µl formamide loading dye and loaded on a 20x20x0.1 cm gel of 6% polyacrylamide, 7M urea in 1xTBE (tris borate EDTA) buffer. Following electrophoresis, the gel was then subjected to electro-transfer onto nylon filter (GeneScreen Plus) in 0.5xTBE at 30 Volts for 7-15 hours at 4°C. Following blotting, the PAGEN filter was reacted with radioactive probes, as described (Lotan et al., 2005).

Acknowledgements

We thank David Shore for a gift of AID-SFP1 strain and for critically reading the manuscript. This work was supported by the Israel Science Foundation grant # 301/20 (to MC) and by grants [PID2020-112853GB-C31 funded by MCIN/AEI/10.13039/501100011033 to JEP-O and grant PID2020-112853GB-C32 funded by MCIN/AEI/10.13039/501100011033 to SC]. SG was supported by a Medical Research Council Non-Clinical Senior Research Fellowship [MR/R008205/1 to S.G.]

References

Albert B, Tomassetti S, Gloor Y, Dilg D, Mattarocci S, Kubik S, Hafner L, Shore D. 2019. Sfp1 regulates transcriptional networks driving cell growth and division through multiple promoter-binding modes. *Genes Dev* **33**:288–293. doi:10.1101/gad.322040.118

Archambault J, Lacroute O, Ruet A, Friesen JD. 1992. Genetic interaction between transcription elongation factor TFIIS and RNA polymerase II. *Mol Cell Biol* **12**:4142–4152. doi:10.1128/MCB.12.9.4142-4152.1992

Babbarwal V, Fu J, Reese JC. 2014. The Rpb4/7 module of RNA polymerase II is required for carbon catabolite repressor protein 4-negative on TATA (Ccr4-Not) complex to promote elongation. *Journal of Biological Chemistry* **289**:33125–33130. doi:10.1074/jbc.C114.601088

Bar-Nahum G, Epshtein V, Ruckenstein AE, Rafikov R, Mustaev A, Nudler E. 2005. A ratchet mechanism of transcription elongation and its control. *Cell* **120**:183–193. doi:10.1016/j.cell.2004.11.045

Battaglia S, Lidschreiber M, Baejen C, Torkler P, Vos SM, Cramer P. 2017. RNA-dependent chromatin association of transcription elongation factors and pol II CTD kinases. *Elife* **6**. doi:10.7554/eLife.25637

Begley V, Corzo D, Jordán-Pla A, Cuevas-Bermúdez A, Miguel-Jiménez L de, Pérez-Aguado D, Machuca-Ostos M, Navarro F, Chávez MJ, Pérez-Ortín JE, Chávez S. 2019. The mRNA degradation factor Xrn1 regulates transcription elongation in parallel to Ccr4. *Nucleic Acids Res* **47**:9524–9541. doi:10.1093/nar/gkz660

Begley V, Jordán-Pla A, Peñate X, Garrido-Godino AI, Challal D, Cuevas-Bermúdez A, Mitjavila A, Barucco M, Gutiérrez G, Singh A, Alepuz P, Navarro F, Libri D, Pérez-Ortín JE, Chávez S. 2021. Xrn1 influence on gene transcription results from the combination of general effects on elongating RNA pol II and gene-specific chromatin configuration. *RNA Biol* **18**:1310–1323. doi:10.1080/15476286.2020.1845504

Blasco-Moreno B, de Campos-Mata L, Böttcher R, García-Martínez J, Jungfleisch J, Nedialkova DD, Chattopadhyay S, Gas M-E, Oliva B, Pérez-Ortín JE, Leidel SA, Choder M, Díez J. 2019. The exonuclease Xrn1 activates transcription and translation of mRNAs encoding membrane proteins. *Nat Commun* **10**. doi:10.1038/s41467-019-09199-6

Bohnsack MT, Tollervey D, Granneman S. 2012. Identification of RNA helicase target sites by UV cross-linking and analysis of cDNAMethods in Enzymology. Academic Press Inc. pp. 275–288. doi:10.1016/B978-0-12-396546-2.00013-9

Bregman A, Avraham-Kelbert M, Barkai O, Duek L, Guterman A, Choder M. 2011. Promoter elements regulate cytoplasmic mRNA decay. *Cell* **147**. doi:10.1016/j.cell.2011.12.005

Bryll AR, Peterson CL. 2023. The circular logic of mRNA homeostasis. doi:10.1080/21541264.2023.2183684

Buratti E, Baralle FE. 2012. TDP-43: Gumming up neurons through protein-protein and protein-RNA interactions. *Trends Biochem Sci*. doi:10.1016/j.tibs.2012.03.003

Chattopadhyay S, Garcia-Martinez J, Haimovich G, Khwaja A, Barkai O, Ghosh A, Chuarzman SG, Schuldiner M, Elran R, Rosenberg M, Bohnsack K, Bohnsack M, Perez-Ortin JE, Choder M. 2022. RNA-controlled nucleocytoplasmic shuttling of mRNA decay factors regulates mRNA synthesis and initiates a novel mRNA decay pathway. *Nat Commun* **13**:7184–7201. doi:10.1101/2021.04.01.437949

Chen YS, Hou W, Tracy S, Harvey AT, Harjono V, Xu F, Moresco JJ, Yates JR, Zid BM. 2022. Rvb1/Rvb2 proteins couple transcription and translation during glucose starvation. *Elife* **11**. doi:10.7554/ELIFE.76965

Cheung ACM, Cramer P. 2011. Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature* **471**:249–253. doi:10.1038/NATURE09785

Choder M. 2011. mRNA imprinting. *Cell Logist* **1**:37–40. doi:10.4161/cl.1.1.14465

Choder M. 2004. Rpb4 and Rpb7: Subunits of RNA polymerase II and beyond, Trends in Biochemical Sciences. doi:10.1016/j.tibs.2004.10.007

Choder M, Young RA. 1993. A portion of RNA polymerase II molecules has a component essential for stress responses and stress survival. *Mol Cell Biol* **13**. doi:10.1128/MCB.13.11.6984

Churchman LS, Weissman JS. 2011. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469**:368–373. doi:10.1038/NATURE09652

Collart MA. 2016. The Ccr4-Not complex is a key regulator of eukaryotic gene expression. *Wiley Interdiscip Rev RNA*. doi:10.1002/wrna.1332

Dahan N, Choder M. 2013. The eukaryotic transcriptional machinery regulates mRNA translation and decay in the cytoplasm. *Biochim Biophys Acta Gene Regul Mech* **1829**. doi:10.1016/j.bbagrm.2012.08.004

Duek L, Barkai O, Elran R, Adawi I, Choder M. 2018. Dissociation of Rpb4 from RNA polymerase II is important for yeast functionality. *PLoS One* **13**. doi:10.1371/journal.pone.0206161

El-Brolosy MA, Kontarakis Z, Rossi A, Kuenne C, Günther S, Fukuda N, Kikhi K, Boezio GLM, Takacs CM, Lai S-L, Fukuda R, Gerri C, Giraldez AJ, Stainier DYR. 2019. Genetic compensation triggered by mutant mRNA degradation. *Nature 2019 568:7751* **568**:193–197. doi:10.1038/s41586-019-1064-z

Farago M, Nahari T, Hammel C, Cole CN, Choder M. 2003. Rpb4p, a subunit of RNA polymerase II, mediates mRNA export during stress. *Mol Biol Cell* **14**. doi:10.1091/mbc.E02-11-0740

Fingerman I, Nagaraj V, Norris D, Vershon A. 2003. Sfp1 plays a key role in yeast ribosome biogenesis. *Eukaryot Cell* **2**:1061–1068.

Fischer J, Song YS, Yosef N, di Iulio J, Churchman LS, Choder M. 2020. The yeast exoribonuclease Xrn1 and associated factors modulate RNA polymerase II processivity in 5’ and 3’ gene regions. *J Biol Chem* **295**:11435–11454. doi:10.1074/jbc.RA120.013426

Font J, MacKay JP. 2010. Beyond DNA: Zinc finger domains as RNA-binding modules. *Methods in Molecular Biology* **649**:479–491. doi:10.1007/978-1-60761-753-2_29/FIGURES/2

García-Martínez J, Aranda A, Pérez-Ortín JE. 2004. Genomic run-on evaluates transcription rates for all yeast genes and identifies gene regulatory mechanisms. *Mol Cell* **15**:303–313. doi:10.1016/j.molcel.2004.06.004

García-Martínez J, Singh A, Medina D, Chávez S, Pérez-Ortín JE. 2023. Enhanced gene regulation by cooperation between mRNA decay and gene transcription. *Biochim Biophys Acta Gene Regul Mech* **1866**. doi:10.1016/J.BBAGRM.2023.194910

García-Martínez J, Troulé K, Chávez S, Pérez-Ortín JE. 2016. Growth rate controls mRNA turnover in steady and non-steady states. *RNA Biol* **13**:1175–1181. doi:10.1080/15476286.2016.1236171

Garrido-Godino A, I G, F G-S, AB M-P, A A, LM S, JE P-O, V P, F N. 2021. Rpb4 and Puf3 imprint and post-transcriptionally control the stability of a common set of mRNAs in yeast. *RNA Biol* **18**:1206–1220. doi:10.1080/15476286.2020.1839229

Gilbertson S, Federspiel JD, Hartenian E, Cristea IM, Glaunsinger B. 2018. Changes in mRNA abundance drive shuttling of RNA binding proteins, linking cytoplasmic RNA degradation to transcription. *Elife* **7**. doi:10.7554/eLife.37663

Goler-Baron V, Selitrennik M, Barkai O, Haimovich G, Lotan R, Choder M. 2008. Transcription in the nucleus and mRNA decay in the cytoplasm are coupled processes. *Genes Dev* **22**:2022–7. doi:10.1101/gad.473608

Gómez-Herreros F, De Miguel-Jiménez L, Millán-Zambrano G, Peñate X, Delgado-Ramos L, Muñoz-Centeno MC, Chávez S. 2012a. One step back before moving forward: regulation of transcription elongation by arrest and backtracking. *FEBS Lett* **586**:2820–2825. doi:10.1016/J.FEBSLET.2012.07.030

Gómez-Herreros F, De Miguel-Jiménez L, Morillo-Huesca M, Delgado-Ramos L, Muñoz-Centeno MC, Chávez S. 2012b. TFIIS is required for the balanced expression of the genes encoding ribosomal components under transcriptional stress. *Nucleic Acids Res* **40**:6508–6519. doi:10.1093/NAR/GKS340

Granneman S, Kudla G, Petfalski E, Tollervey D. 2009. Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc Natl Acad Sci U S A* **106**:9613–9618. doi:10.1073/pnas.0901997106

Haag S, Kretschmer J, Sloan KE, Bohnsack MT. 2017. Crosslinking methods to identify RNA methyltransferase targets in vivoMethods in Molecular Biology. Humana Press Inc. pp. 269–281. doi:10.1007/978-1-4939-6807-7_18

Haimovich G, Choder M, Singer RH, Trcek T. 2013a. The fate of the messenger is pre-determined: A new model for regulation of gene expression. *Biochim Biophys Acta Gene Regul Mech*. doi:10.1016/j.bbagrm.2013.01.004

Haimovich G, Medina DA, Causse SZ, Garber M, Millán-Zambrano G, Barkai O, Chávez S, Pérez-Ortín JE, Darzacq X, Choder M. 2013b. Gene Expression Is Circular: Factors for mRNA Degradation Also Foster mRNA Synthesis. *Cell* **153**:1000–1011. doi:10.1016/j.cell.2013.05.012

Harel-Sharvit L, Eldad N, Haimovich G, Barkai O, Duek L, Choder M. 2010. RNA Polymerase II Subunits Link Transcription and mRNA Decay to Translation. *Cell* **143**:552–563. doi:10.1016/J.CELL.2010.10.033

Hartenian E, Glaunsinger BA. 2019. Feedback to the central dogma : cytoplasmic mRNA decay and transcription are interdependent processes. *Crit Rev Biochem Mol Biol* **0**:1–14. doi:10.1080/10409238.2019.1679083

Jordán-Pla A, Gupta I, De Miguel-Jimnez L, Steinmetz LM, Chávez S, Pelechano V, Prez-Ortín JE. 2015. Chromatin-dependent regulation of RNA polymerases II and III activity throughout the transcription cycle. *Nucleic Acids Res* **43**:787–802. doi:10.1093/NAR/GKU1349

Jorgensen P, Rupeš I, Sharom JR, Schneper L, Broach JR, Tyers M. 2004. A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. *Genes Dev* **18**:2491–2505. doi:10.1101/gad.1228804

Kruk JA, Dutta A, Fu J, Gilmour DS, Reese JC. 2011a. The multifunctional Ccr4–Not complex directly promotes transcription elongation. *Genes & Development* **25**:581–593. doi:10.1101/gad.2020911

Kruk JA, Dutta A, Fu J, Gilmour DS, Reese JC. 2011b. The multifunctional Ccr4–Not complex directly promotes transcription elongation. *Genes & Development* **25**:581–593. doi:10.1101/gad.2020911

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, Fitzhugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, Levine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond Christina, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Hong ML, Dubois J, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, De La Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JGR, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin E V., Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran J V., Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AFA, Stupka E, Szustakowki J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson M V., Kaul R, Raymond Christopher, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Patrinos A, Morgan MJ. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**:860–921. doi:10.1038/35057062

Lascaris RF, Mager WH, Planta RJ. 1999. DNA-binding requirements of the yeast protein Rap1p as selected in silico from ribosomal protein gene promoter sequences. *Bioinformatics* **15**:267–277. doi:10.1093/BIOINFORMATICS/15.4.267

Lieb J, Liu X, Botstein D, Genetics PB-N, 2001 U. 2001. Promoter-specific binding of Rap1 revealed by genome-wide maps of protein–DNA association. *Nat Genet* **28**:327–334.

Lotan R, Bar-On VG, Harel-Sharvit L, Duek L, Melamed D, Choder M. 2005. The RNA polymerase II subunit Rpb4p mediates decay of a specific class of mRNAs. *Genes Dev* **19**:3004–16. doi:10.1101/gad.353205

Lotan R, Goler-Baron V, Duek L, Haimovich G, Choder M. 2007. The Rpb7p subunit of yeast RNA polymerase II plays roles in the two major cytoplasmic mRNA decay mechanisms. *J Cell Biol* **178**:1133–43. doi:10.1083/jcb.200701165

Luo Y, Na Z, Slavoff SA. 2018. P-Bodies: Composition, Properties, and Functions. *Biochemistry* **57**:2424–2431. doi:10.1021/ACS.BIOCHEM.7B01162

Ma Z, Zhu P, Shi H, Guo L, Zhang Q, Chen Y, Chen S, Zhang Z, Peng J, Chen J. 2019. PTC-bearing mRNA elicits a genetic compensation response via Upf3a and COMPASS components. *Nature 2019 568:7751* **568**:259–263. doi:10.1038/s41586-019-1057-y

Marion RM, Regev A, Segal E, Barash Y, Koller D, Friedman N, O'Shea EK. 2004. Sfp1 is a stress- and nutrient-sensitive regulator of ribosomal protein gene expression. *Proc Natl Acad Sci U S A* **101**:14315–14322. doi:10.1073/PNAS.0405353101/SUPPL_FILE/05353TABLE3.XLS

Nishimura K, Fukagawa T, Takisawa H, Kakimoto T, Kanemaki M. 2009. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat Methods* **6**:917–922. doi:10.1038/NMETH.1401

Parker R. 2012. RNA Degradation in Saccharomyces cerevisae. *Genetics* **191**:671–702. doi:10.1534/GENETICS.111.137265

Pelechano V, Jimeno-González S, Rodríguez-Gil A, García-Martínez J, Pérez-Ortín JE, Chávez S. 2009. Regulon-specific control of transcription elongation across the yeast genome. *PLoS Genet* **5**. doi:10.1371/JOURNAL.PGEN.1000614

Pérez-Ortín JE, Chávez S. 2022. Nucleo-cytoplasmic shuttling of RNA-binding factors: mRNA buffering and beyond. *Biochim Biophys Acta Gene Regul Mech* **1865**. doi:10.1016/J.BBAGRM.2022.194849

Reja R, Vinayachandran V, Ghosh S, Pugh BF. 2015. Molecular mechanisms of ribosomal protein gene coregulation. *Genes Dev* **29**:1942–1954. doi:10.1101/GAD.268896.115

Rosenheck S, Choder M. 1998. Rpb4, a subunit of RNA polymerase II, enables the enzyme to transcribe at temperature extremes in vitro. *J Bacteriol* **180**.

Sachs AB, Davis RW. 1989. The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. *Cell* **58**:857–867. doi:10.1016/0092-8674(89)90938-0

Schawalder S, Kabani M, Howald I, Choudhury U, Werner M, Shore D, Wade JT, Hall DB, Struhl K. 2004. Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh1. *Nature* **432**. doi:10.1038/nature03200ï

Shore D, Zencir S, Albert B. 2021. Transcriptional control of ribosome biogenesis in yeast: links to growth and stress signals. *Biochem Soc Trans* **49**:1589–1599. doi:10.1042/BST20201136

Sigurdsson S, Dirac-Svejstrup AB, Svejstrup JQ. 2010. Evidence that Transcript Cleavage Is Essential for RNA Polymerase II Transcription and Cell Viability. *Mol Cell* **38**:202–210. doi:10.1016/J.MOLCEL.2010.02.026

Sun M, Schwalb B, Schulz D, Pirkl N, Etzold S, Larivière L, Maier KC, Seizl M, Tresch A, Cramer P. 2012. Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome Res*. doi:10.1101/gr.130161.111

Timmers HTM, Tora L. 2018. Transcript Buffering: A Balancing Act between mRNA Synthesis and mRNA Degradation. *Mol Cell*. doi:10.1016/j.molcel.2018.08.023

Trcek T, Larson DR, Moldón A, Query CC, Singer RH. 2011. Single-molecule mRNA decay measurements reveal promoter-regulated mRNA stability in yeast. *Cell* **147**:1484–1497. doi:10.1016/j.cell.2011.11.051

van Nues R, Schweikert G, de Leau E, Selega A, Langford A, Franklin R, Iosub I, Wadsworth P, Sanguinetti G, Granneman S. 2017. Kinetic CRAC uncovers a role for Nab3 in determining gene expression profiles during stress. *Nat Commun* **8**:12. doi:10.1038/s41467-017-00025-5

Vera M, Pani B, Griffiths LA, Muchardt C, Abbott CM, Singer RH, Nudler E. 2014. The translation elongation factor eEF1A1 couples transcription to translation during heat shock response. *Elife* **3**:e03164. doi:10.7554/eLife.03164

Villanyi Z, Ribaud V, Kassem S, Panasenko OO, Pahi Z, Gupta I, Steinmetz L, Boros I, Collart MA. 2014. The Not5 subunit of the ccr4-not complex connects transcription and translation. *PLoS Genet* **10**. doi:10.1371/JOURNAL.PGEN.1004569

Wan R, Bai R, Zhan X, Shi Y. 2020. How Is Precursor Messenger RNA Spliced by the Spliceosome? *Annu Rev Biochem* **89**:333–358. doi:10.1146/ANNUREV-BIOCHEM-013118-111024

Warner JR. 1999. The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci* **24**:437–440. doi:10.1016/S0968-0004(99)01460-7

Xiao R, Chen JY, Liang Z, Luo D, Chen G, Lu ZJ, Chen Y, Zhou B, Li H, Du X, Yang Y, San M, Wei X, Liu W, Lécuyer E, Graveley BR, Yeo GW, Burge CB, Zhang MQ, Zhou Y, Fu XD. 2019. Pervasive Chromatin-RNA Binding Protein Interactions Enable RNA-Based Regulation of Transcription. *Cell* **178**:107-121.e18. doi:10.1016/J.CELL.2019.06.001

Zid BM, O'Shea EK. 2014. Promoter sequences direct cytoplasmic localization and translation of mRNAs during starvation in yeast. *Nature 2014 514:7520* **514**:117–121. doi:10.1038/nature13578

Figures and legends

Figure 1. **Sfp1 shuttles in a transcription-dependent manner and localizes to P- bodies.** **(A)** *Shuttling assay*. The assay used the temperature-sensitive (ts) *nup49-313* mutant and was performed as reported (Lee et al., 1996; Selitrennik et al., 2006). Wild type (WT, yMS119), *nup49-313*(ts) (yMS1) and *nup49-313*(ts) *rpb1-1*(ts) cells expressing GFP-Sfp1 were allowed to proliferate under optimal conditions at 24°C. Cycloheximide (CHX) (50 μg/ml) was then added, and the cultures were shifted to 37°C (to inhibit Nup49-313 and Rpb1-1). The proportion of cells expressing cytoplasmic GFP-Sfp1 was plotted as a function of time (N>200). Error bars represent the standard deviation of three replicates. **(B)** *GFP-Sfp1 is co-localized with P bodies markers***.** Cells expressing the indicated fluorescent proteins were allowed to proliferate till mid-logarithmic phase, followed by 24 h starvation in medium lacking glucose and amino acids. Live cells were inspected under fluorescent microscope. White arrows mark P-bodies. **(C)** *The number of GFP-Sfp1 containing foci per cell decrease in response to cycloheximide (CHX) treatment.* CHX (50 μg/ml) was added to exponentially proliferating cultures for the indicated time. Cells were then shifted to starvation medium as in B; Average of 2 replicates is shown. Student's t-test between time 0 and the indicated time points was performed; ** represents $p<0.001$ ($N\geq240$).

Figure 2. Sfp1 binds a group of mRNAs around GCTGCT motif. (A) *Metagene profile of CRAC analysis in two wild type replicates.* RPKMs plot around the average metagene region of all yeast genes. The 5'UTR and 3'UTR are shown in real scale in base pairs (bp) whereas the transcribed region is shown as percentage scale to normalize different gene lengths. **(B)** *Heatmap representation of CRAC reads around the polyadenylation site (pA) for the top 600 genes with the highest number of CRAC reads.* The genes with high CRAC signal density upstream of the pA site are considered as CRAC+ (n=264), indicated on the right. The chosen cut-off was somewhat arbitrary; additional analyses shown in the subsequent figures indicate that this choice was biologically significant. **(C)** *Average metagene analysis of two replicates of the CRAC+ signal in*

genes containing a GCTGCT motif in a region ±500 *bp around the motif*. CRAC reads were aligned by the center of the motif. **(D)** *Sfp1 pulls down CRAC+ mRNAs.* The extracts of isogenic cells, expressing the indicated tandem affinity purification tag (TAP), were subjected to tandem affinity purification (Puig et al., 2001), in the presence of RNase inhibitors. The RNA was extracted and was analyzed by Northern blot hybridization, using the probes indicated on the left. "+intron" denotes the position of intron-containing *RPL30* RNA.

Figure 3. The mRNA-binding specificity of Sfp1 depends on the Rap1 binding site (RapBS) within the promoter. **(A)** *Constructs used in this study*. The constructs were described previously (Bregman et al., 2011). To differentiate construct-encoded mRNAs from endogenous ones, we inserted an oligo$(G)_{18}$ in the 3' untranslated region. The constructs are identical except for the nature of their upstream activating sequence (UAS), located upstream of the *ACT1* core promoter that includes the TATA box (designated ''TATA''). The nucleotide boundaries of the *RPL30* sequences are depicted above the constructs, and those of the *ACT1* sequences are depicted below the constructs. The numbering referes to the translation start codon. These constructs encode identical mRNA (Bregman et al., 2011). **(B)** *Binding of Sfp1 to mRNA is dependent on Rap1-binding site (RapBS)*. Extracts of cells, expressing the indicated constructs, were subjected to RNA immunoprecipitation (RIP), using tandem affinity purification (TAP) of the indicated TAP-tagged proteins. RIP was followed by Northern blot hybridization using the probes indicated at the left. After the membrane was hybridized with oligo$(C)_{18}$-containing probes (to detect *RPL30pG* mRNA; see Bregman et al., 2011), the membrane was hybridized with probes to detect endogenous *RPL25* and *RPL29* mRNAs. +Intron represents the intron-containing *RPL30pG* RNA **(C)** *Most RiBi CRAC+ genes have Rap1 binding sites in their promoters.* Only 25 RiBi genes (including RiBi-like) are defined as CRAC+. Upper panel: Venn diagram showing overlap between these genes and genes carrying promoters with RapBS. Lower panel: Venn diagram showing overlap between all RiBi (+ RiBi-like) genes excluding CRAC+ ("non-CRAC+") and genes carrying promoters with RapBS. Data on Rap1 bound promoters were obtained from (Lieb et al., 2001). A hypergeometric test was applied to calculate the p-values, indicated underneath each diagram. Note that p-values were significant for both inclusion (upper diagram) and exclusion (lower diagram).

Figure 4. Sfp1 is required for efficient transcription of CRAC+ genes and stabilizes the deadenylation-dependent pathway of their mRNAs' decay. (A) *Sfp1 deletion affects SR (synthesis rate), RA (mRNA Abundance) and HL (mRNA half-life) differently for different subsets of genes.* CRAC+ (n=264), RiBi (including RiBi-like) CRAC+ (n=25), RiBi (+RiBi-like) excluding CRAC+ ("RiBi (non-CRAC+) (n=411), RPs (n=129). Statistical analyses were performed using the Wilcoxon test. Asterisks indicate significant results (* p-value < 0.05; ** p-value < 0.01; *** p-value < 0.001; **** p-value < 0.0001). Unless indicated otherwise, statistical comparisons were performed using the "Rest" group (all detectable genes exclusing CRAC+ ones) as a reference. In addition, RiBi CRAC+ and RiBi non-CRAC+ are compared against each other. **(B)** *Sfp1 depletion rapidly affects SR and RA of CRAC+ genes.* Scatterplot of changes in SR vs changes in RA at 20 min. (left) or 60 min. (right) after depleting Sfp1-degron by auxin. CRAC+ genes are highlighted in green. Spearman correlation values and the significance of the linear adjustment for the whole dataset are indicated inside the plot. Density curves are drawn on the margins of the plot to help evaluate the overlap between dots. **(C)** *RapBS confers Sfp1-dependent mRNA decay pathway*. Shown is the quantification of Northern blot hybridization results of mRNA decay assay (Methods), performed with WT or Δ*sfp1* cells that carried the indicated constructs (described in (Bregman et al., 2011) and shown at the top). The membrane was probed sequentially with an oligo$(C)_{18}$-containing probe, to detect the construct-encoded mRNA, and with probes to detect endogenous mRNAs. mRNA levels were normalized to the Pol III transcript *SCR1* mRNA (Methods). The band intensity at time 0, before transcription inhibition, was defined as 100% and the intensities at the other time points (min) were calculated relative to time 0. Error bars indicate the standard deviation the mean values of three independent replicates (for $(G)_{18}$-containing mRNAs), or of 12 replicates (for endogenous mRNAs). **(D)** The *deadenylation rate of CRAC+ mRNAs and the subsequent decay of the deadenylated RNAs is accelerated in sfp1Δ cells***.** Transcription was blocked as described in section C (Methods). RNA samples were analyzed using the polyacrylamide Northern technique (Sachs and Davis, 1989), using the probes indicated on the left, all are CRAC+. Half-lives were determined and the indicated ratios are depicted on the right. The asterisk (*) indicates the time point at which deadenylation is estimated to be complete.

Figure 5. Binding features of Sfp1 to chromatin. (A) *Sfp1 is present in the bodies of CRAC+ genes.* Average metagene of Sfp1 and Rap1 ChIP-exo signal, obtained from (Reja et al., 2015), for CRAC+ genes (n=264) and CRAC-CONTROL genes (a subset of 264 non-CRAC+ genes randomly selected from the entire genome, but excluding RP and RiBi genes). See also the plot with alternative scaling in Fig. S4A). **(B)** Left panel: *Positive correlation between Sfp1 binding to gene bodies and transcriptional activity.* Scatterplot comparing Sfp1 binding to gene bodies (measured by ChIP-exo, Reja et al., 2015) versus the density of actively elongating RNA pol II (measured by BioGRO-seq; Begley et al., 2021) in CRAC+ genes (n=264). Spearman correlation is indicated at the bottom of the plot. Right panel: *Correlation between Sfp1 binding to the bodies of all genes (n=4777) and transcription rate.* Spearman correlation for all genes (including CRAC+) or for all genes exclusing CRAC+ genes are indicated at the bottom of the plot. **(C)** The *Sfp1 ChIP-exo signal drops downstream of the GCTGCT motif.* Comparison of average metagene profiles of CRAC (blue) and ChIP-exo (orange) signals for genes with a GCTGCT motif (n=163).

Figure 6. Sfp1 induces Pol II backtracking, preferentially in CRAC+ genes. (A) *Sfp1 differently affects Pol II occupancy (left) and on Pol II activity (right).* Pol II levels were measured by Rpb3 ChIP and its activity - by genomic run-on (GRO), in WT (BY4741) and its isogenic *sfp1Δ* strain, growing exponentially in YPD. Anti Rpb3 (rabbit polyclonal) ChIP on chip experiments using Affymetrix® GeneChip *S. Cerevisiae* Tiling 1.0R custom arrays as described in Methods. For each gene, the average of signals corresponding to tiles covering 5' and 3' ends (250 bp) were calculated. Green dots represent the CRAC+ genes; grey dot – non-CRAC+. The tendency line, its equation, Pearson R and its *p*-value of the statistically significant deviation from the null hypothesis of no correlation are shown in grey, for the whole dataset, and in green, for the CRAC+ genes. (**B)** *Sfp1* promotes *Pol II backtracking of CRAC+ genes.* Backtracking index (BI), defined as the ratio of Rpb3-ChIP to GRO signals, is shown for different gene sets, indicated below, comparing WT and *sfp1*Δ strains. In order to compare data obtained from different types of experiments the values were normalized by the median and standard deviation (z-score). The bars represent standard errors. Statistical significance of the differences between the averages of the indicated samples was estimated using a two-tailed Student's t-test (* means $p < 0.01$). **(C)**

mRNA HL and BI of CRAC+ mRNAs/genes are affected by Sfp1. Box and whisker plots showing the effect of Sfp1 on mRNA HL and Pol II BI. A comparison between CRAC+ (green) and all genes excluding CRAC+ ones ("non-CRAC+" grey) genes is shown. HL was calculated from the mRNA abundance (RA) and synthesis rates (SR), using the data shown in Fig. 4A. The statistical significance of the differences between the averages of the CRAC+ and non-CRAC+was estimated using a two-tailed Student's t-test (*** means $p < 0.0001$). (**D)** *mRNA HL and BI are correlated via Sfp1: correlation between Sfp1-dependence of BI and mRNA HL ratios.* Data from C were represented in a scatter plot. Linear regression equations are shown for all (grey) and CRAC+ genes (green). Pearson correlation coefficient, *r*, and the *p*-value of the statistically significant deviation from the null hypothesis of no correlation ($r = 0$) are also indicated. All statistical correlations were determined using the ggpubr package in R.

Figure 7. Sfp1 alters Rpb4 stoichiometry/configuration within Pol II elongation complex and this alteration is linked to mRNA stabilization. (A) The *Rpb4 stoichiometry/configuration changes along the transcription units in Sfp1-mediated manner.* Top left panel - the values of Rpb3 and Rpb4 were obtained from ChIP on chip experiments, either against Rpb3 or against Rpb4-Myc in LMY3.1 cells proliferated exponentially in YPD. Rpb4-ChIP/ Rpb3-ChIP ratios were calculated and averages for the indicated genes sets were obtained for positions ranging from -100 to +250 (relative to TSS) and from -250 to +100 (relative to pA sites). Average ratios were normalised to the TSS -100 position, in order to represent profiles of Rpb4-ChIP changes after Pol II recruitment to promoters. Top right panel –profiles of Rpb4-ChIP/Rpb3-ChIP ratios were obtained as in top right panel, but from *sfp1*Δ strain (LMY7.1). Bottom left panel - Rpb4-ChIP/Rpb3-ChIP profiles of CRAC+ genes: comparing WT (blue) and *sfp1*Δ (red) strains. Bottom right panel - Rpb4-ChIP/Rpb3-ChIP profiles of rest of the genes (all detectable genes excluding CRAC+ ones) : comparing WT (blue) and *sfp1*Δ (red) strains. (**B)** *Correlation between the Rpb4-ChIP/Rpb3-ChIP ratios and Pol II BI in WT (left panel) or in sfp1Δ cells (right panel).* For each gene, average of Rpb4-ChIP/Rpb3-ChIP values corresponding to positions from TSS to +250 and from -250 to pA sites were calculated. BI values taken from Fig. 6C. Linear regression equations are shown for all (grey) and CRAC+ genes (green). Pearson correlation coefficients, *r*, and the *p*-values of the statistically significant deviation from the null hypothesis of no correlation ($r = 0$),

are also shown. All statistical correlations were determined using the *ggpubr* package in R. (**C)** *Correlation between the Rpb4-ChIP/Rpb3-ChIP ratios and mRNA HL in WT or sfp1Δ cells.* Values of Rpb4-ChIP/Rpb3-ChIP ratios ere determined as in B. mRNA HL was indirectly calculated from mRNA abundance and transcription rates taken from the data used in Fig. 4A and is shown in arbitrary units. CRAC+ genes are depicted in green. R was calculated as in B.

Figure 8. A model for Sfp1 functions in yeast. (A) ***Two modes of action of Sfp1 in transcription: CRAC+ genes recruit Sfp1 to their promoters, whereas non-CRAC+ genes recruit Sfp1 from the nuclear space directly to Pol II.*** Upper panel represents CRAC+ genes that recruit Sfp1 to their promoters. We discovered that Sfp1 appears to accompany Pol II of CRAC+ genes, in a manner proportional to the number of transcriptionally active Pol II (Fig. 5A-B). Its binding to all Pol II molecules, including backtracked Pol II, is even more apparent (Fig. 5B). In addition, we found that CRAC+ genes are enriched with Rap1-binding sites (Fig. 3). We therefore propose that, following binding to Rap1-containing promoters, Sfp1 binds Pol II. Specifically, it binds to Rpb4 (and possibly other Pol II subunits) and accompanies it until imprinting. This interaction influences Pol II configuration (Fig. 7A-D) and increases the likelihood of Pol II to undergo backtracking (Fig. 6B). Lower panel represents non-CRAC+ genes that also interact with Sfp1 (Fig. 5A CONTROL, Fig. 5B). We propose that promoters of non-CRAC+ genes recruit Sfp1 poorly (relative to CRAC+ promoters), except for small group of promoters, e.g., of RiBi genes lacking RapBS. The dashed arrow represents this minor group. For the majority of these genes, the nuclear Sfp1 interacts directly with their elongating Pol II, as its interaction correlates with the extent of chromatin-bound Pol II (Fig. 5B). This weak interaction also changes Pol II configuration (Fig. 7A, "non-CRAC+") and increases the propensity of Pol II to backtrack (Fig. 7B). However, it either does not result in imprinting or results in rare imprinting events that went undetected by our CRAC assay. **(B) Sfp1 mediates cross-talk between mRNA synthesis and decay via imprinting.** The backtracked configuration, induced by Sfp1 (A), is compatible with a movement of Sfp1 from Pol II to its transcripts (see text), which is enhanced in case the GCTGCT motif is localized near Sfp1 (Fig. 5C). Following co-transcriptional RNA binding, Sfp1 accompanies the mRNA to the cytoplasm and stabilizes the mRNAs. Following mRNA degradation, Sfp1 is imported back into the nucleus to initiate a new cycle. The model proposes that the specificity of

Sfp1-RNA interaction is determined, in part, by the promoter (Fig. 3A-B). Nevertheless, promoter binding is necessary, but not sufficient for RNA binding. See text for more details.

Supplemental Figures

Figure S1. Sfp1 binds Rpb4 and its efficient import is dependent on *RPB4*. (A) *RPB4 forms two hybrid interaction with SFP1, but not with FHL1, IFH1 or ABF1.* Two-hybrid interaction, using Rpb4 as the bait and the indicated proteins as the preys, was determine by growth on plates lacking leucine, tryptophan, adenine, and histidine supplemented with 5mM of 3-amino-1,2,4-triazole (Uetz et al., 2000). We verified that the growth on the indicator plates was dependent on both plasmids by evicting one plasmid at a time from each of the positive clones (results not shown). **(B)** *GFP-Sfp1 shuttles by a transcription-dependent mechanism.* Shuttling of GFP-Sfp1 was determined using *nup49-313*(ts) mutant cells that are defective in protein import at elevated temperatures (Lee et al., 1996). Wild type (WT, yMS119), *nup49-313*(ts) (yMS1) and *nup49-313 rpb1-1*(ts) (yMC4) cells (whose transcription is blocked following a temperature increase to 37ºC) expressing GFP-Sfp1 24°C. During mid-log phase, cycloheximide (CHX) (50 μg/ml) was added and the cultures were divided into two samples. One was incubated for at 24°C and the other - at 37°C. To monitor export kinetics, samples were examined microscopically at the indicated time points and photographs of random fields were taken. Cells were classified into those exhibiting nuclear or whole-cell (i.e., cytoplasmic) localization of GFP-Sfp1 (N>200). The proportion of cells exhibiting cytoplasmic localization was plotted as a function of time. Bars represent standard deviation of 3 replicates. **(C)** *Examples of the results at 0 or 3h post re-feeding*. Note that, in sated cells, no foci were observed (Fig. S1C).

Figure S2.

(A) *HTP-tagging does not affect yeast strain growth*. HTP-tagged strains were streak on YPD plate. Photo was taken after 2 days at 30ºC. Two individual cell lines, obtained during strain construction, are shown; both were used in this study as replicates. **(B)** *Gene ontology terms most significantly enriched in the 264 CRAC+ genes*. **(C)** *Overlaps between the genes activated by Sfp1 and the set of genes whose mRNA products are bound by Sfp1*. Overlap between CRAC+ genes and genes whose expression was reported to be sensitive (either positively – "Up" or negatively – "Down") to Sfp1 overexpression "Sfp1-OE" (Albert et al., 2019). Statistical figures, using

hypergeometric tests, are indicated. **(D)** The *Sfp1 ChIP-exo signal drops downstream of the GCTGCT motif.* Comparison of average metagene profiles of CRAC (pink) and ChIP-exo (green) signals for genes with a GCTGCT motif (n=163). **(E)** *Sfp1 ChIP-exo signal around the pA sites of control genes as compared with CRAC+ ones.* Average metagene profile of Sfp1 ChIP-exo signal around the pA sites of CRAC+ and CRAC control genes, as described in A. **(F)** Distribution of the distances of the GCTGCT motifs found in CRAC+ genes from to their corresponding pA sites. The motif logo is shown inserted inside the plot and the dashed vertical line indicates the length of the median 3' UTR. **(G)** *Genome browser screenshots of the CRAC signals in the indicated gene loci*. Bottom panels are zoom in version of the upper panels. Circles indicate the GCTGCT motif. Note that the Cs at positions 2 and 5 can be replaced with G (see inset in F).

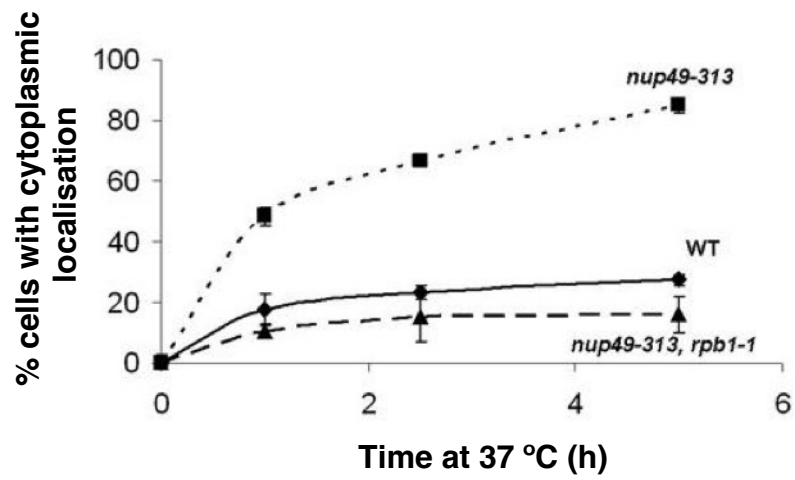
Figure S3. Sfp1 stabilizes CRAC+ mRNAs. (A) *Depletion of Sfp1 by "auxin induced degron" (AID) destabilized RPL30 mRNA having high CRAC index, but not MFA2, having the lowest CRAC index.* Cells expressing AID-*SFP1* or isogenic WT strain were allowed to proliferate under optimal environmental conditions. Auxin was added and cells were incubated for the duration indicated above the graphs. Transcription was then blocked by adding thiolutin and mRNA decay assay was performed (see Methods). (B) Sfp1 stabilizes CRAC+ mRNAs during heat shock (HS). The indicated cells were allowed to proliferate under optimal environmental conditions. At mid-log phase, the cultures were shifted to 42°C, to inactivate transcription (ref). mRNA decay assay was performed and quantified as described in Methods. Note the logarithmic scale of the graphs on the right.

Figure S4. Sfp1 binds CRAC+ gene bodies. (A) *Comparison of the average metagene profile of Sfp1 and Rap1 obtained by ChIP-exo (Reja et al., 2015), focusing on CRAC+ genes*. The original plots were artificially modified to equalize the heights of the profiles by their maximums to show that Sfp1 has higher gene body occupancy than Rap1**. (B)** *Comparison of the average metagene profile of Sfp1 occupancy obtained by two alternative methodologies*: ChIP-exo by Reja et al (2015) and ChEC-seq by Albert et al (2019).

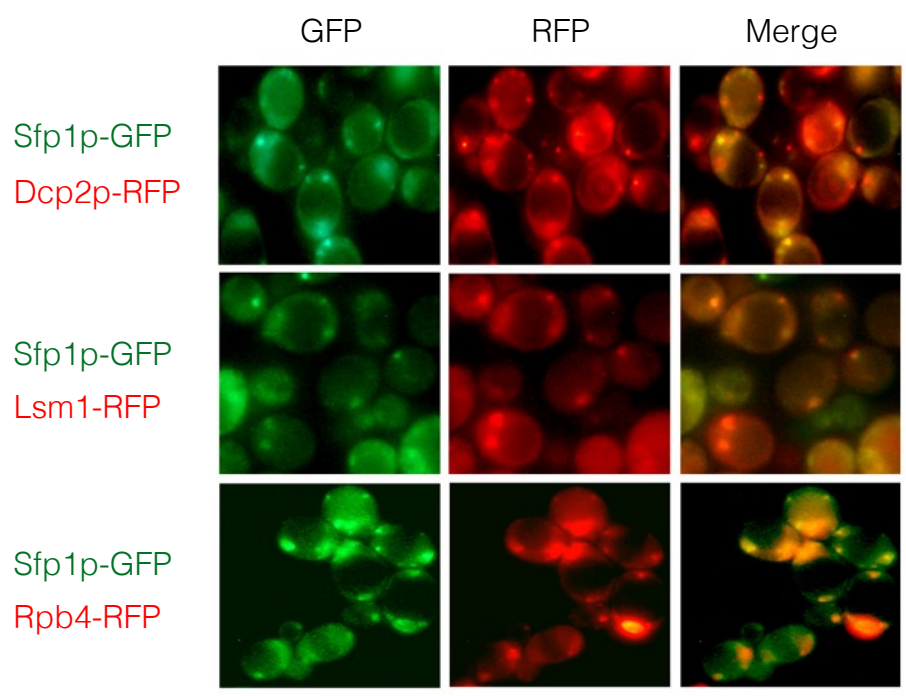
Figure S5. Correlation between the effects of Sfp1 on Rpb4-ChIP/Rpb3-ChIP ratios, Pol II BI and mRNA HL. (**A)** WT/*sfp1*Δ ratios for Rpb4-ChIP/Rpb3-ChIP and Pol II BI were calculated from values represented in Figure 7 B and C, and represented in a scatter plot. CRAC+ genes are depicted in green. Linear regression equations are shown for all (grey) and CRAC+ genes (green). Spearman correlation coefficient and p-value calculated by the stat_cor function of the ggpubr package in R. Correlation coefficients and *p* values of statistical significance are also shown. Note that correlation between Sfp1 effects in BI and Rpb4-ChIP/Rpb3-ChIP ratio was triple in CRAC+ than in CRAC- genes. **(B)** WT/*sfp1*Δ ratios for Rpb4-ChIP/Rpb3-ChIP and mRNA HL were calculated from values represented in Figure 7 B and C, and represented in a scatter plot. CRAC+ genes are depicted in green. The effect of Sfp1 on Rpb4ChIP/Rpb3-ChIP ratio correlated with its effect of mRNA HL in CRAC+ genes, whereas in CRAC- genes no correlation was detected. Linear regression equations are shown for all (grey) and CRAC+ genes (green). Correlation coefficients and *p* values of statistical significance were done as in B.

Figure 1

A Sfp1 shuttles in a transcription-dependent manner



B Sfp1 co-localizes with Rpb4-containing P-bodies



C Number of P-bodies decreases after CHX treatment

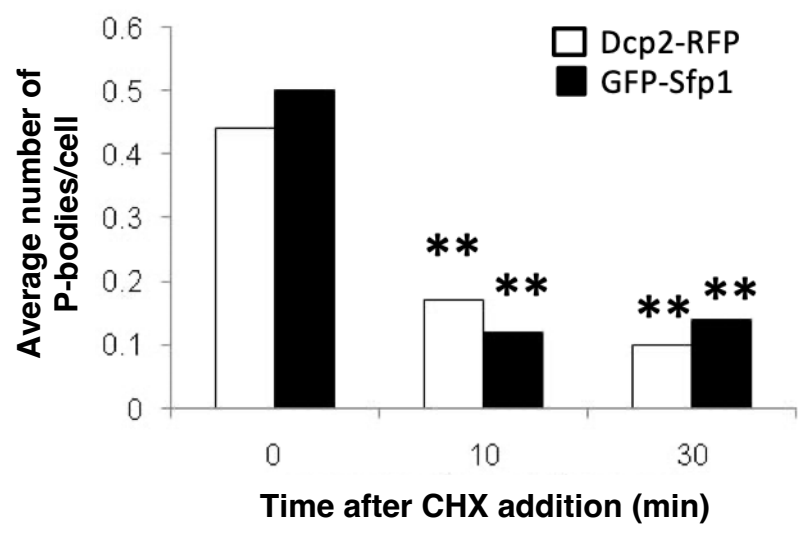
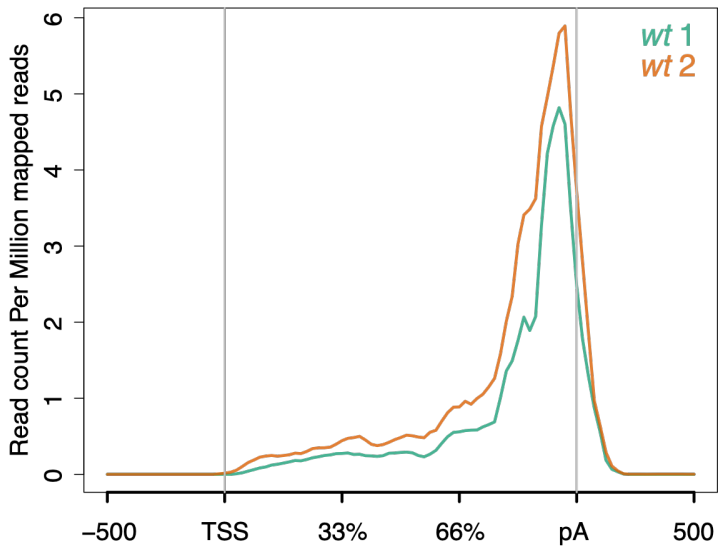


Figure 2

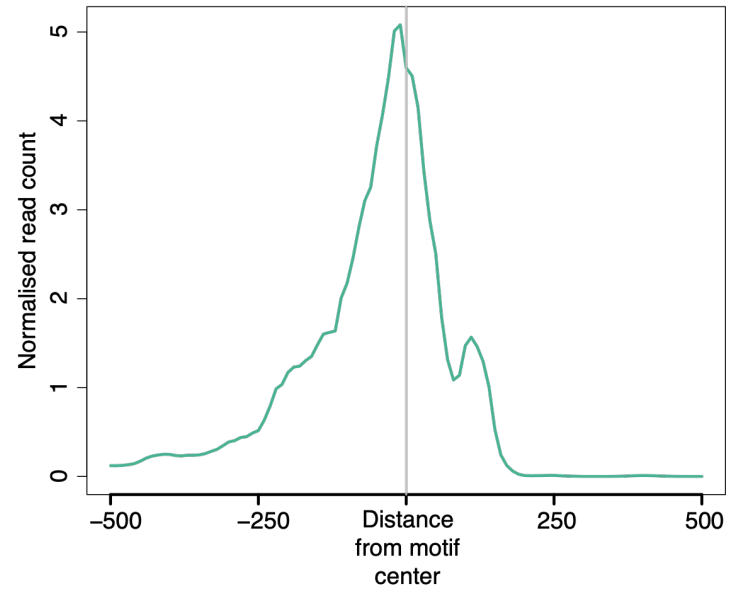
A

CRAC-seq replicate average metagene



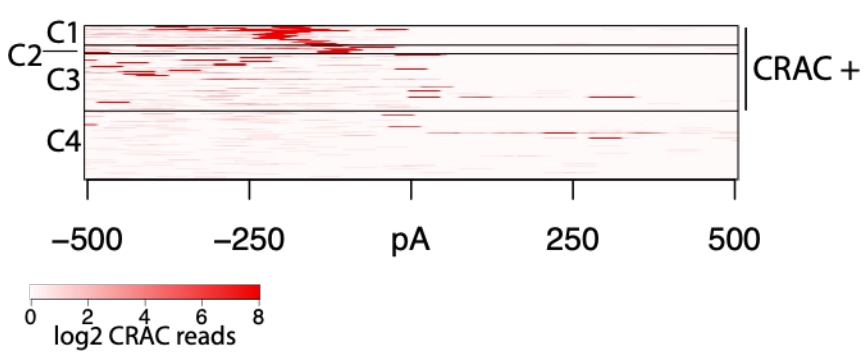
C

Average CRAC signal around GCTGCT motif



B

Heatmap to identify CRAC+ genes



D

Sfp1 binds CRAC+ mRNAs

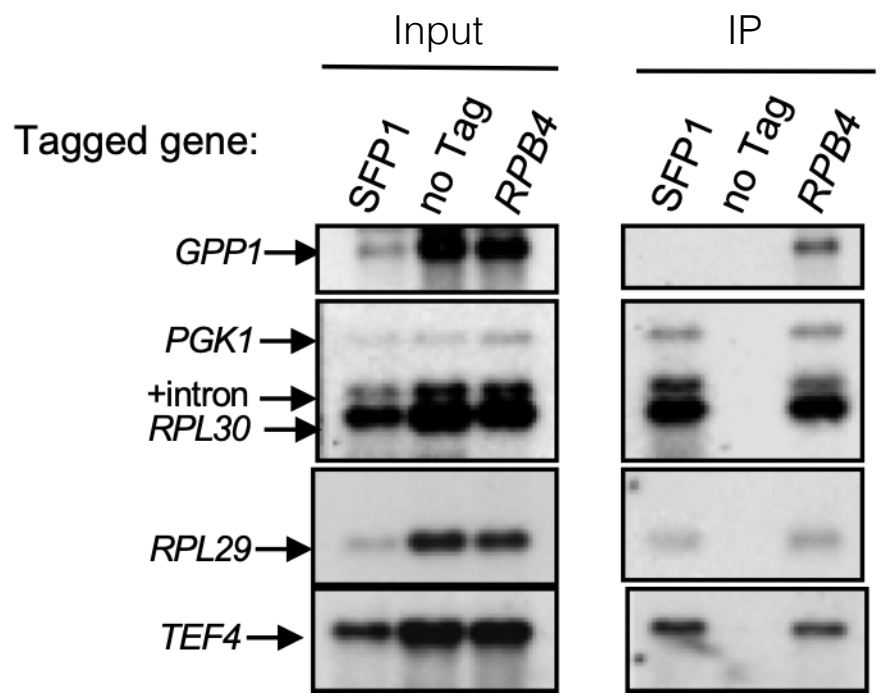
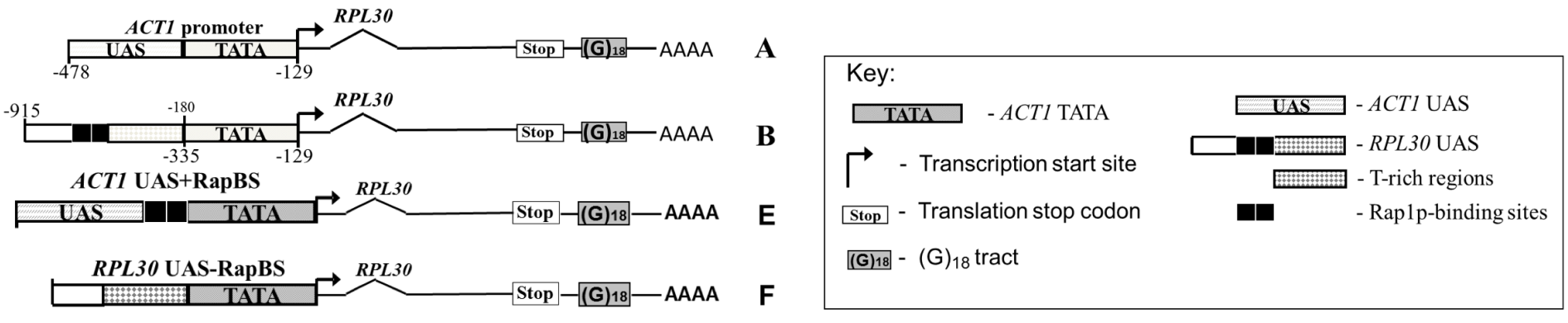


Figure 3

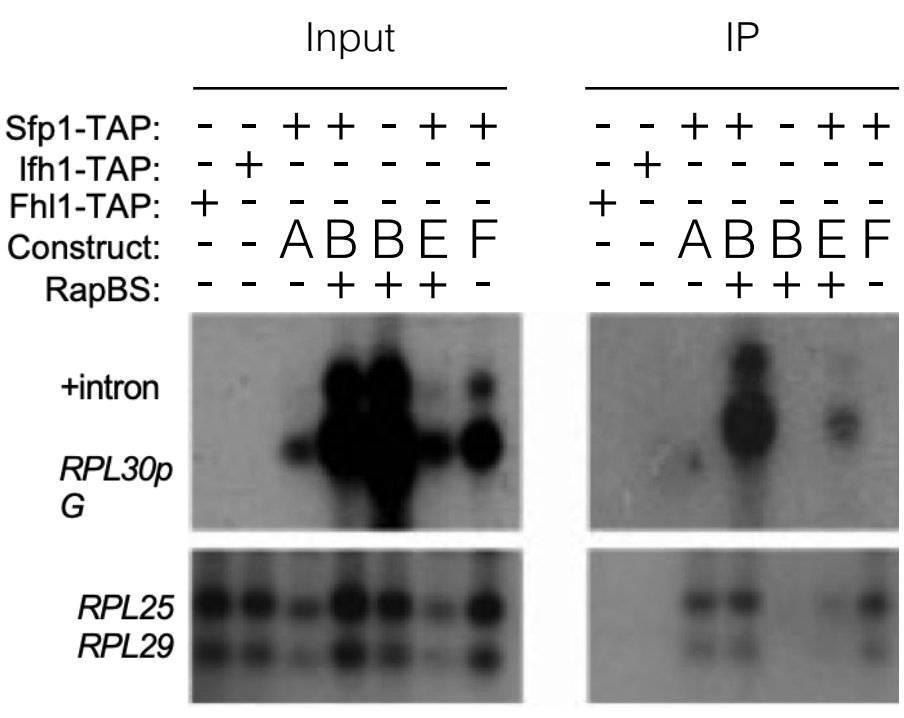
A

Schematic illustration of the constructs used in B



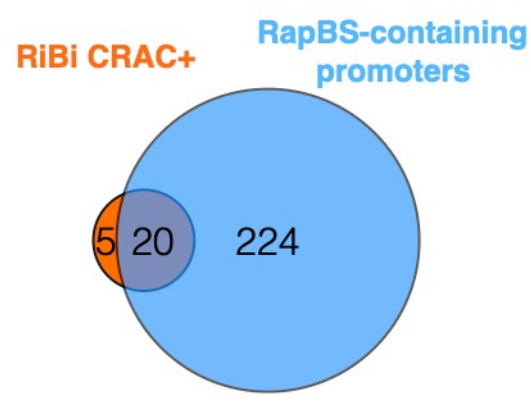
B

Binding of Sfp1 to mRNAs is dependent on Rap1-binding sites (RapBS)

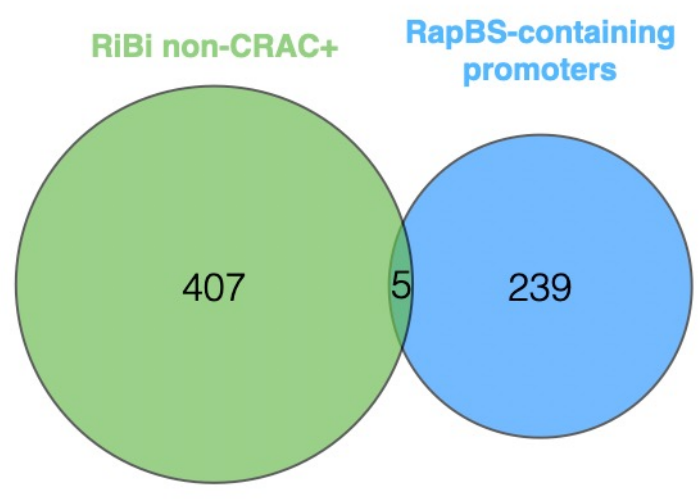


C

Most RiBi CRAC+ genes, but not non-CRAC+, have Rap1 binding sites in their promoters



Enrichment p-value 1.9×10^{-22}

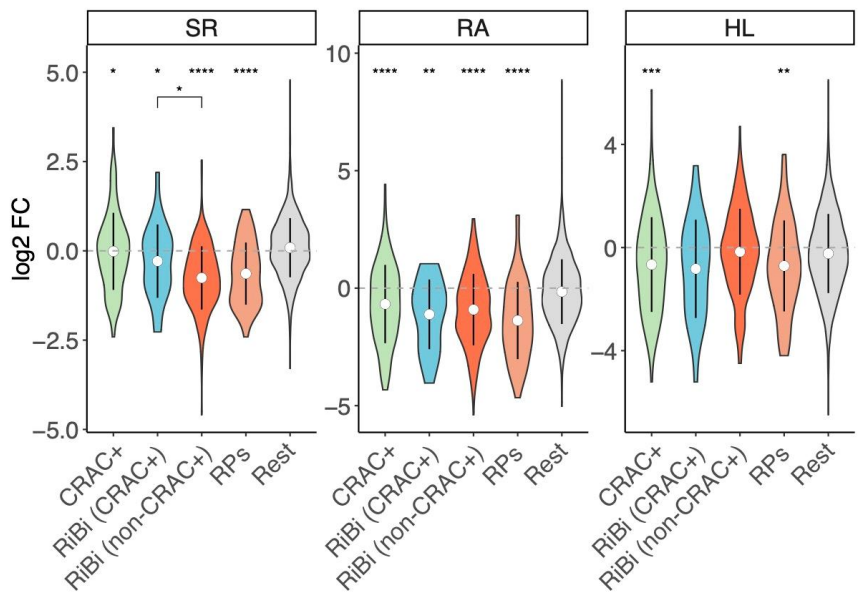


Exclusion p-value 1.5×10^{-5}

Figure 4

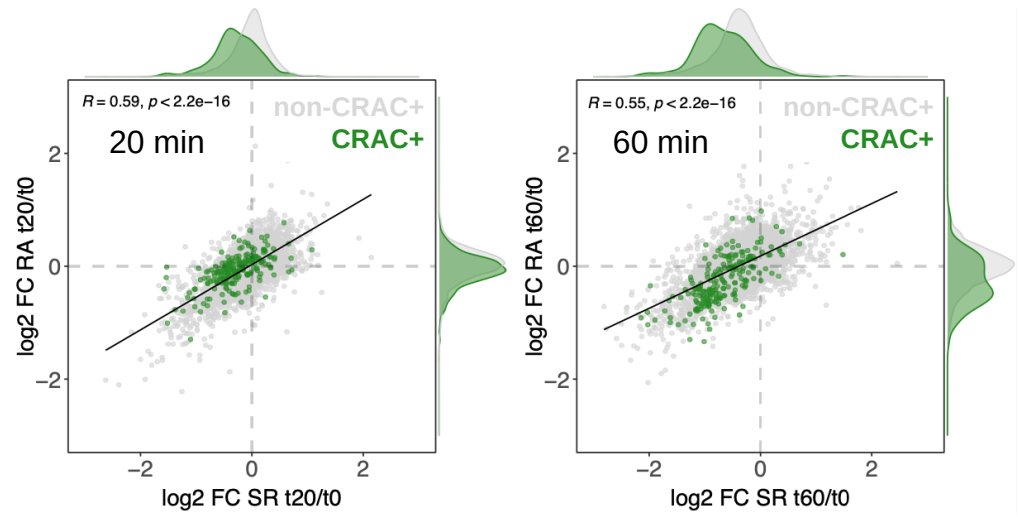
A

Effect of deleting *SFP1* on SR, RA and HL



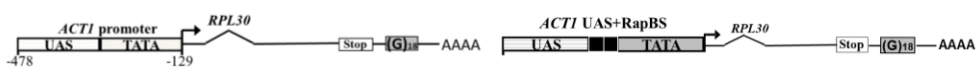
B

Effect of depleting *SFP1* on SR and RA

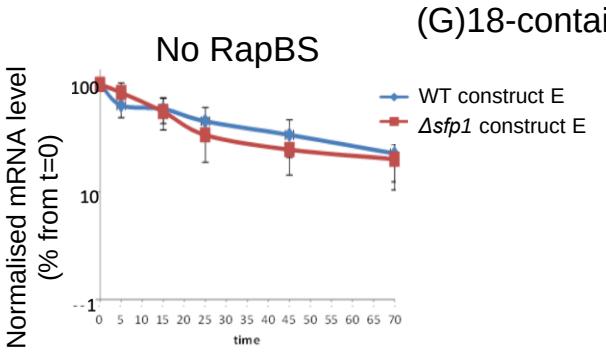


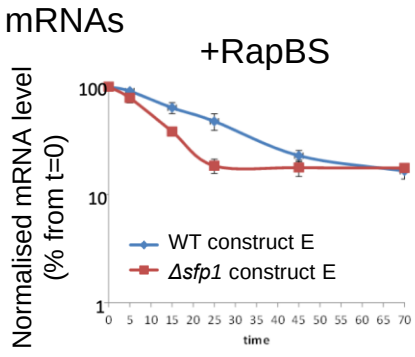
C

RapBS confer Sfp1-dependent mRNA decay pathway

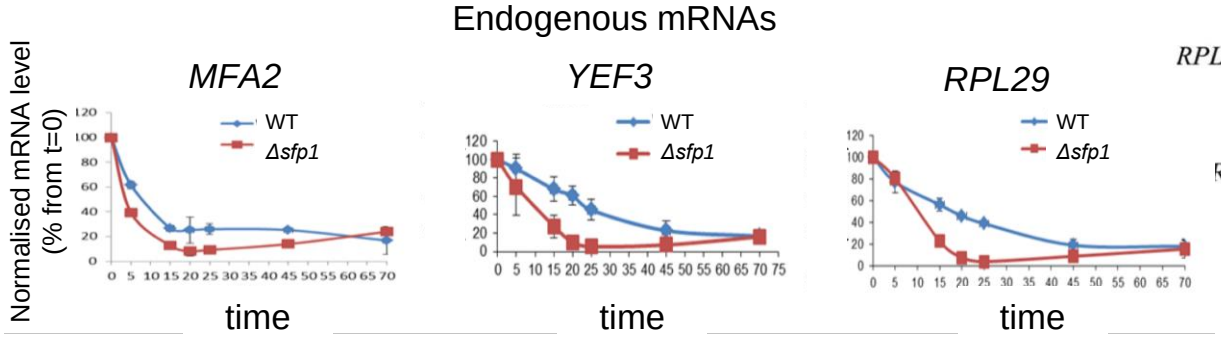


(G)18-containing mRNAs





Endogenous mRNAs



D

Deadenylation and decay of CRAC+ mRNAs are accelerated in *Δsfp1* cells

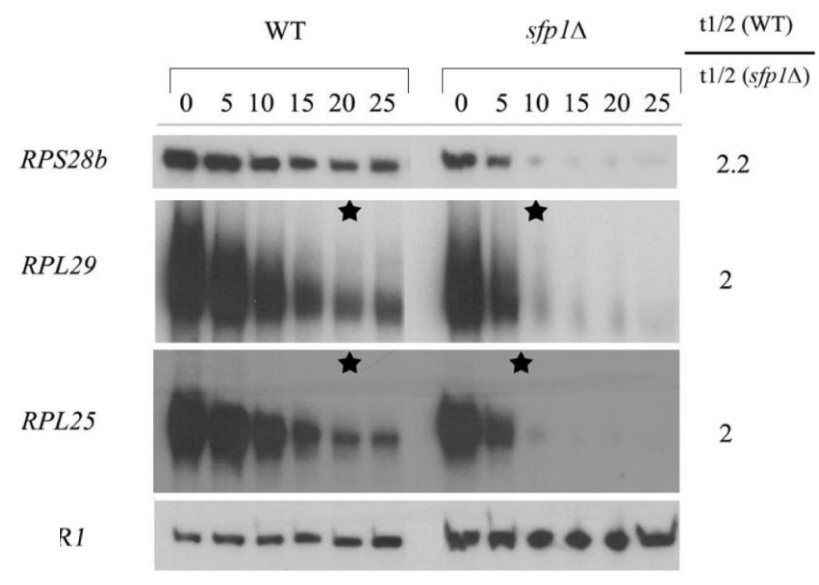
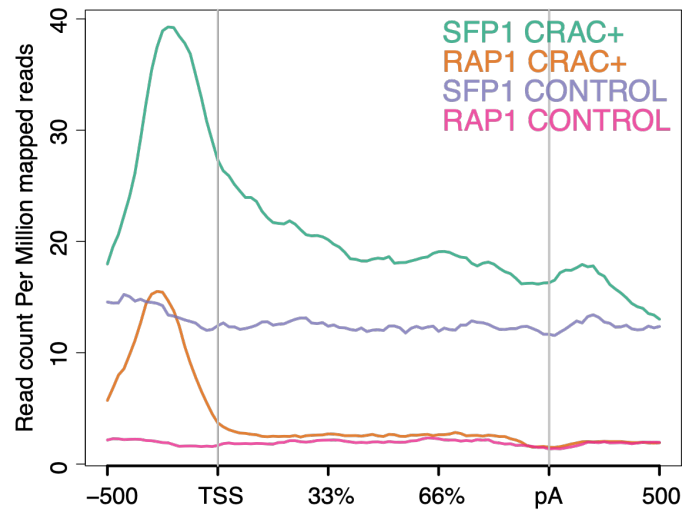


Figure 5

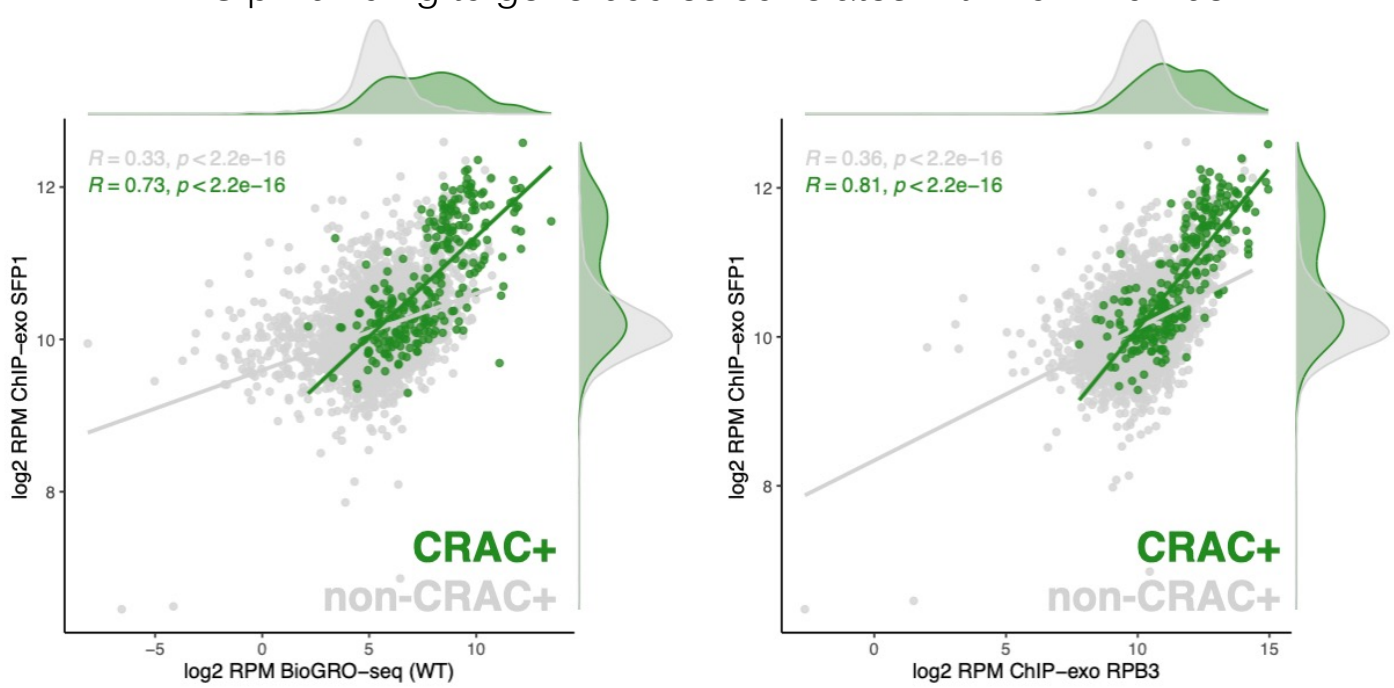
A

Sfp1 is present in bodies
of CRAC+ genes



B

Sfp1 binding to gene bodies correlates with Pol II number



C

Motif locally antagonizes Sfp1
binding to chromatin in CRAC+ genes

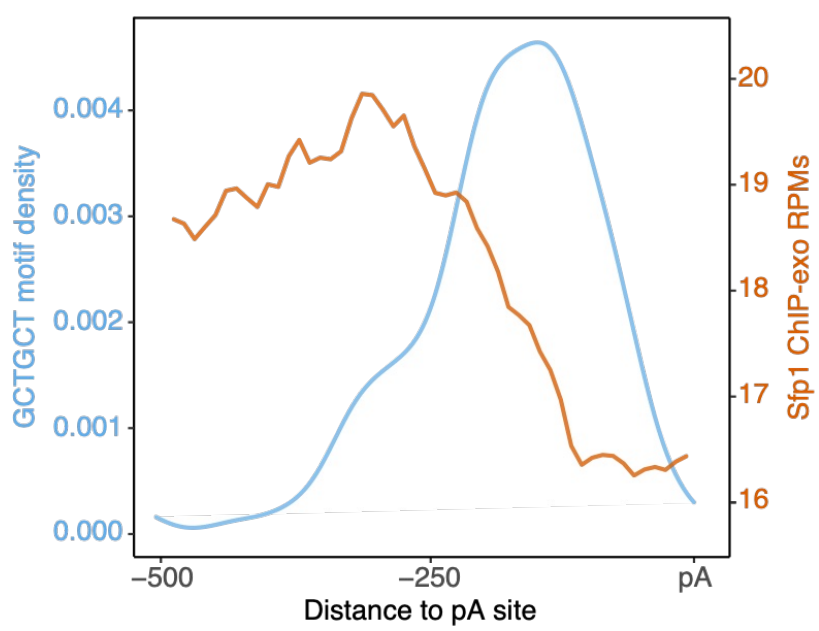
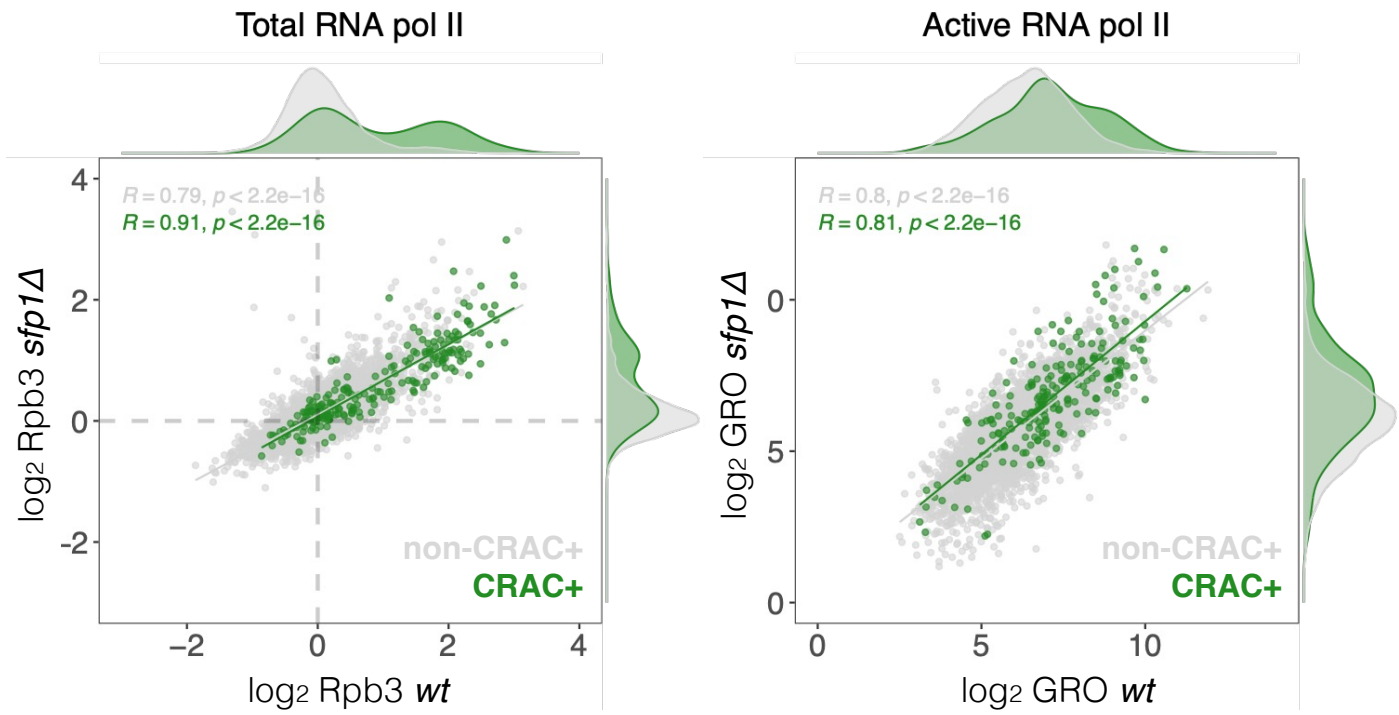


Figure 6

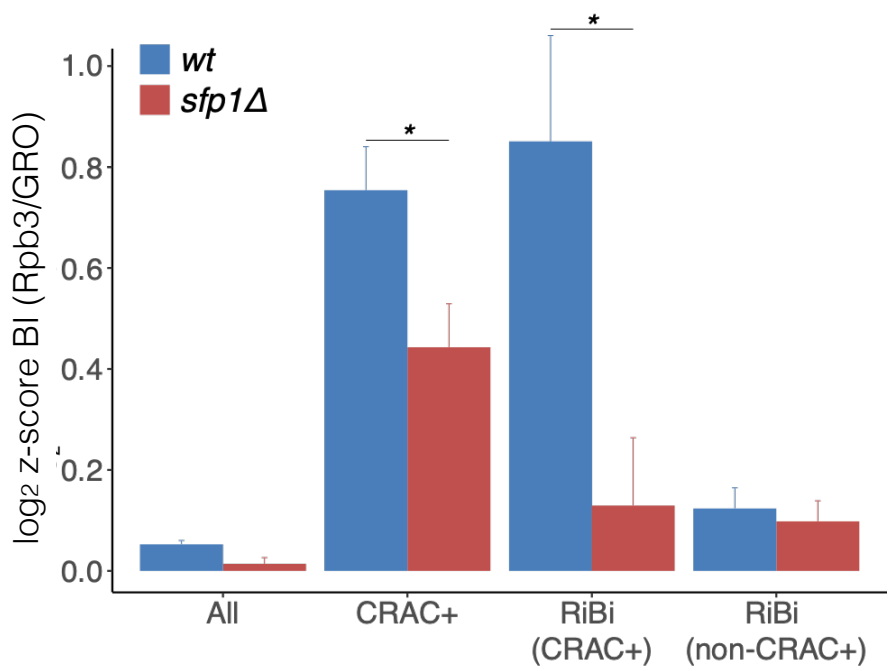
A

Impact of Sfp1 on Pol II occupancy (left) and on Pol II activity (right)



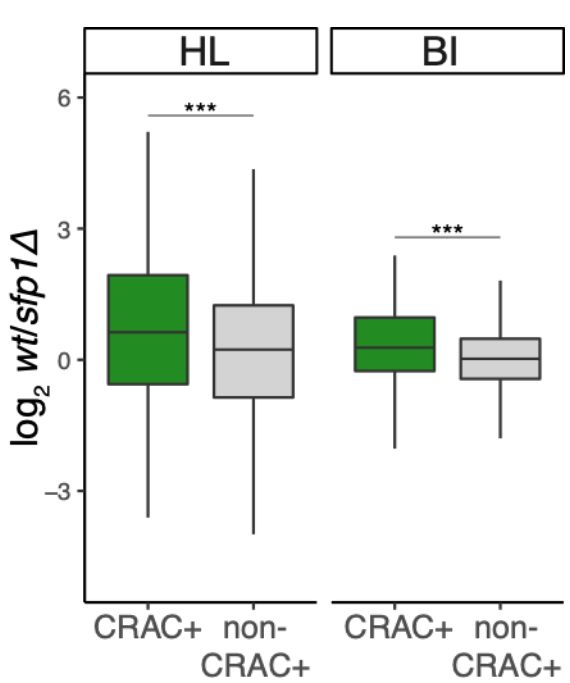
B

Backtracking index (BI) is affected by Sfp1



C

mRNA HL and BI of CRAC+ genes are affected by Sfp1



D

Correlated influence of Sfp1 on BI and mRNA HL

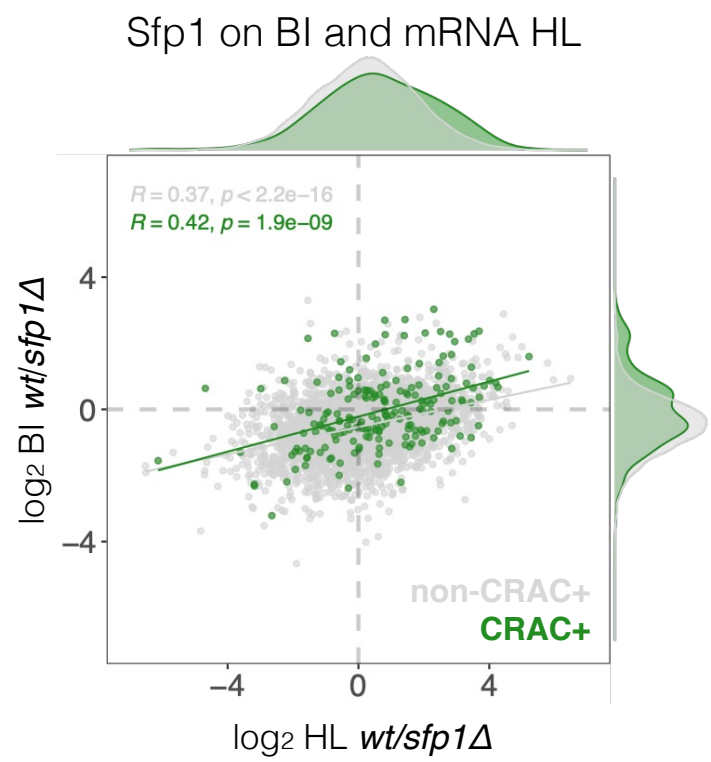
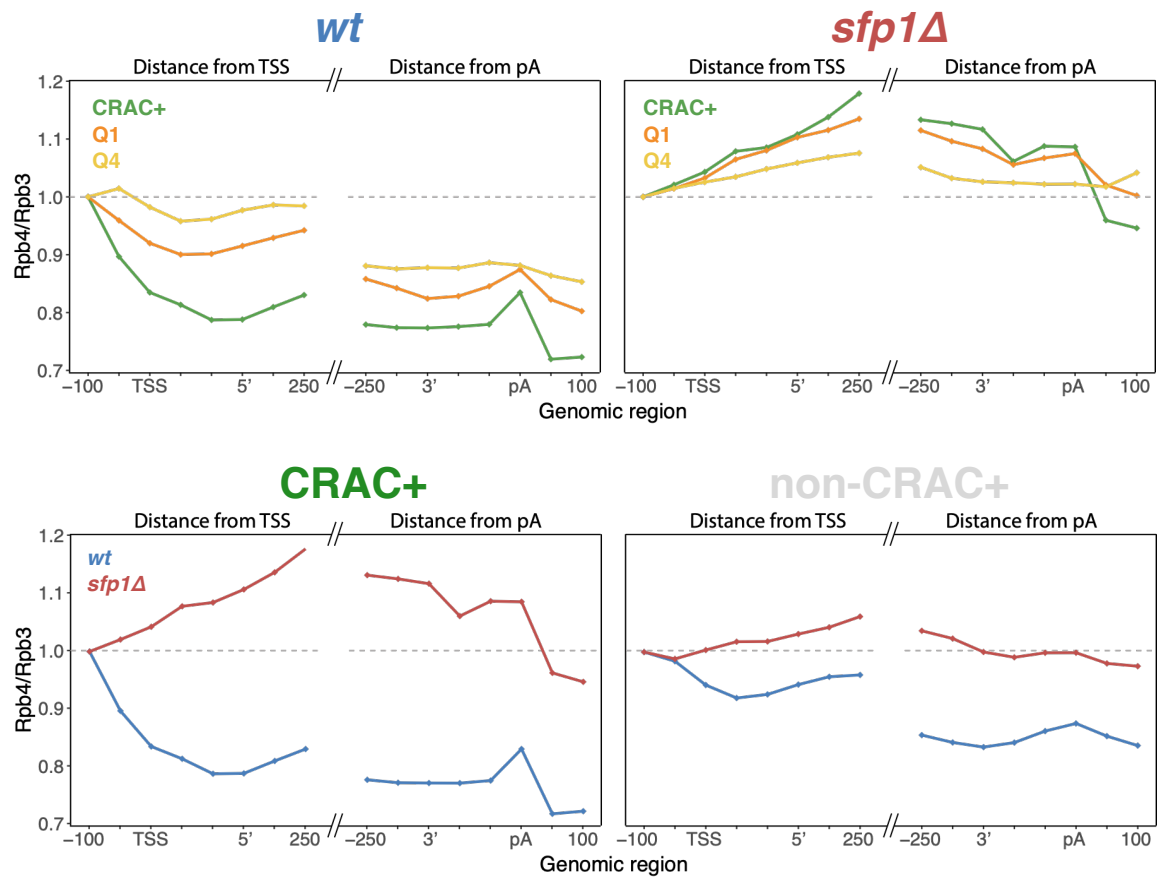
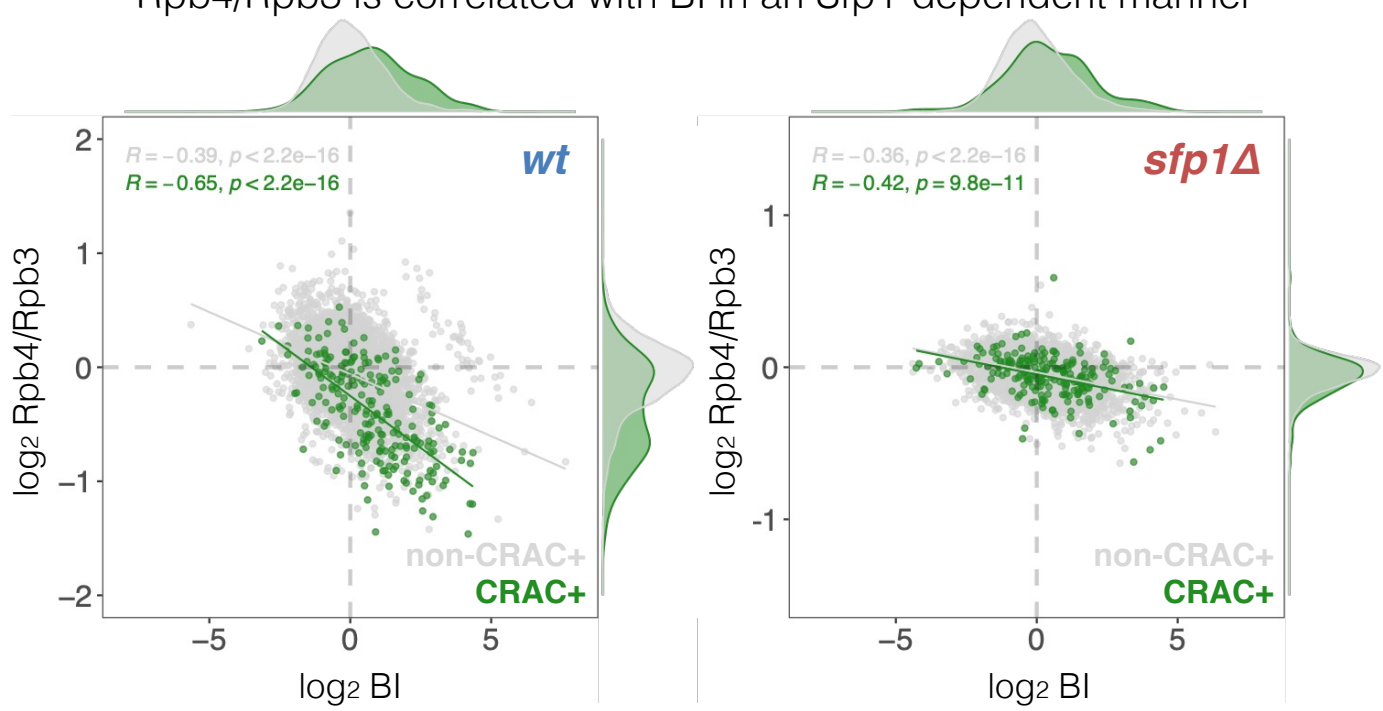


Figure 7

A Rpb4/Rpb3 changes along transcription units in an Sfp1-dependent manner



B Rpb4/Rpb3 is correlated with BI in an Sfp1-dependent manner



C Rpb4/Rpb3 is correlated with mRNA HL in an Sfp1-dependent manner

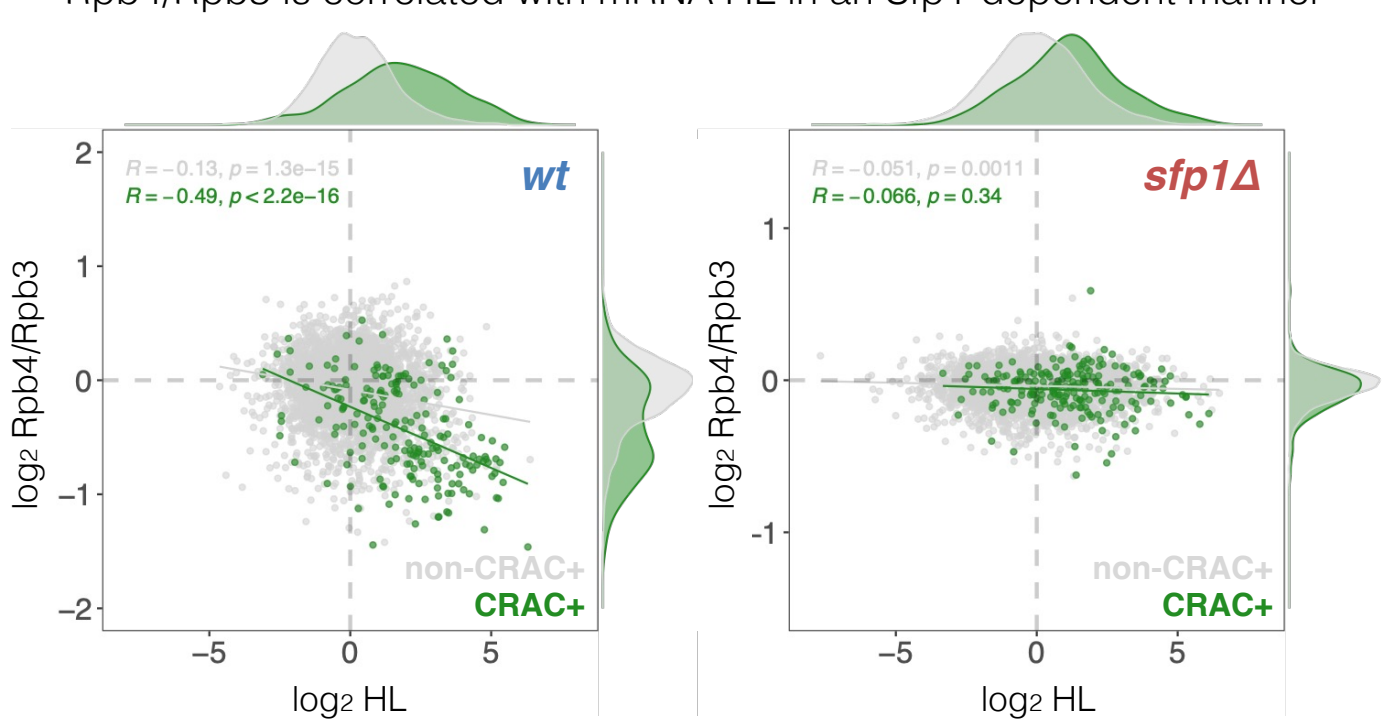
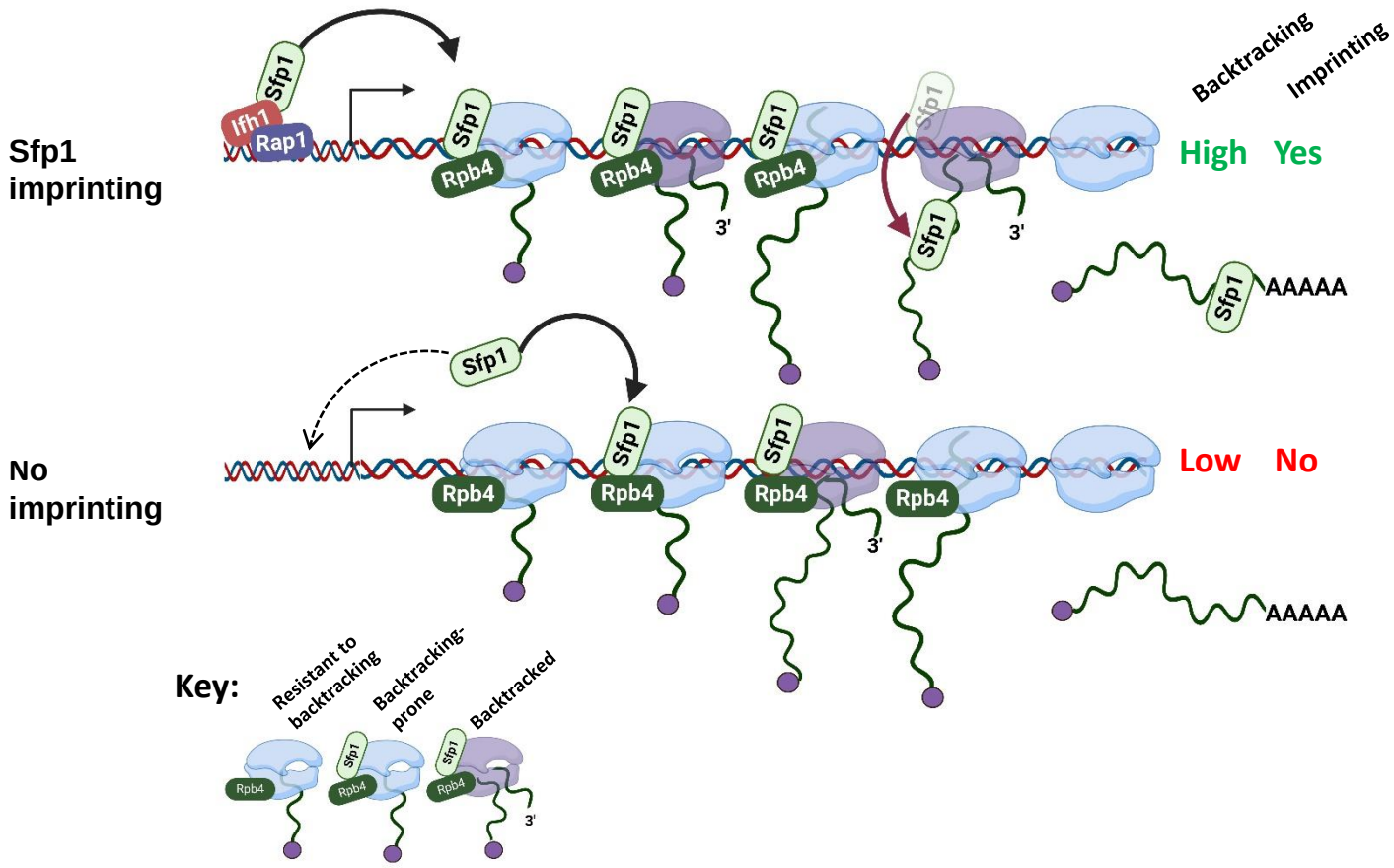
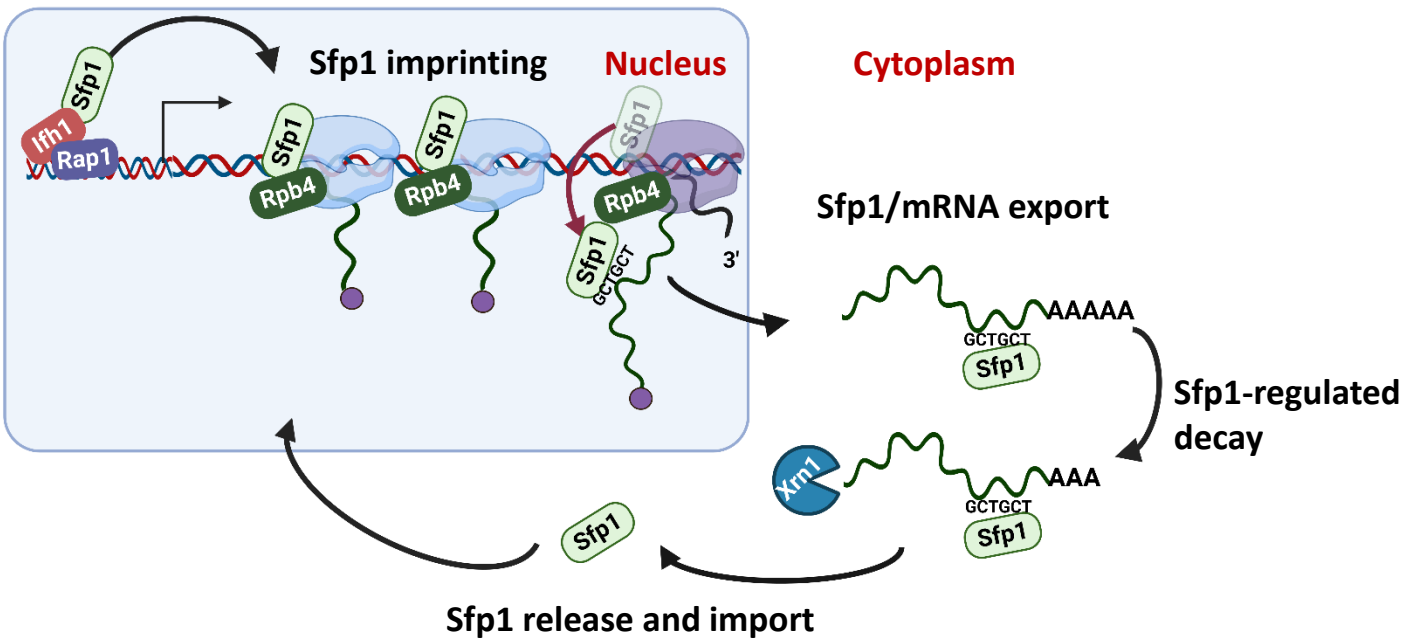


Figure 8

A Two modes of action od Sfp1 in transcription



B Cross-talk between mRNA synthesis & decay by Sfp1 imprinting