Published in final edited form as:

Biochim Biophys Acta. 2016 October ; 1859(10): 1314–1321. doi:10.1016/j.bbagr.2016.06.009.

The fission yeast CENP-B protein Abp1 prevents pervasive transcription of repetitive DNA elements

Anne Daulny^{1,2}, Eva Mejía-Ramírez^{1,2,4}, Oscar Reina², Jesus Rosado-Lugo³, Lorena Aguilar-Arnal^{1,2,4}, Herbert Auer^{2,4}, Mikel Zaratiegui^{3,*}, and Fernando Azorin^{1,2,*}

¹Institute of Molecular Biology of Barcelona, CSIC. Baldiri Reixac, 4. 08028 Barcelona. Spain

²Institute for Research in Biomedicine, IRB Barcelona. The Barcelona Institute of Science and Technology. Baldiri Reixac, 10. 08028 Barcelona. Spain

³Rutgers, the State University of New Jersey, Piscataway, U.S.A

Abstract

It is well established that eukaryotic genomes are pervasively transcribed producing cryptic unstable transcripts (CUTs). However, the mechanisms regulating pervasive transcription are not well understood. Here, we report that the fission yeast CENP-B homologue Abp1 plays an important role in preventing pervasive transcription. We show that loss of *abp1* results in the accumulation of CUTs, which are targeted for degradation by the exosome pathway. These CUTs originate from different types of genomic features, but the highest increase corresponds to *Tf2* retrotransposons and rDNA repeats, where they map along the entire elements. In the absence of *abp1*, increased RNAPII-Ser5P occupancy is observed throughout the *Tf2* coding region and, unexpectedly, RNAPII-Ser5P is enriched at rDNA repeats. Loss of *abp1* also results in *Tf2* derepression and increased nucleolus size. Altogether these results suggest that Abp1 prevents pervasive RNAPII transcription of repetitive DNA elements (ie, *Tf2* and rDNA repeats) from internal cryptic sites.

Keywords

Abp1; CENP-B; Tf2; rDNA; cryptic transcription; CUTs; chromatin; epigenetics; *Schizosaccharomyces pombe*

1. INTRODUCTION

High throughput transcriptome analyses have revealed that pervasive transcription is more frequent than previously anticipated (reviewed in [1–3]). Pervasive transcription initiates at

*To whom correspondence should be addressed: Dr. F. Azorin, Institute of Molecular Biology of Barcelona, CSIC, Institute for Research in Biomedicine of Barcelona, Baldiri Reixac, 10. 08028 Barcelona. Spain, Phone#: 3493 4034958; Fax#: 3493 4034979, fambmc@ibmb.csic.es. Dr. M. Zaratiegui, Department of Molecular Biology and Biochemistry, Rutgers, the State University of New Jersey, Nelson Research Building, A139, 604 Allison Rd, Piscataway NJ 08854, Phone#: (848) 445-1497; Fax#: (732) 445-2447, Zaratiegui@dls.rutgers.edu.

⁴present addresses: Eva Mejía-Ramírez, Center of Regenerative Medicine in Barcelona. Dr. Aiguader, 88 08003 Barcelona. Spain Lorena Aguilar-Arnal, Institute of Biomedical Research, UNAM, 04510 Mexico D.F., Mexico. Herbert Auer, Functional GenOmics Consulting, Bellavista 53, 08753 Palleja. Spain

cryptic sites and results in short non-coding RNAs collectively known as Cryptic Unstable Transcripts (CUTs). CUTs are rapidly degraded by the exosome, but the factors and mechanisms that regulate their expression are not fully understood. Here, we report that, in the fission yeast *Schizosaccharomyces pombe*, the CENP-B protein Abp1 prevents pervasive transcription of repetitive DNA elements. CENP-B is a sequence-specific DNA binding factor that derives from the domestication of *pogo*-like transposases, an event that occurred independently in mammals and *S. pombe* [4–6]. In contrast to mammals that generally have a single CENP-B, the *S. pombe* genome contains three CENP-B homologues (Abp1, Cbh1 and Cbh2) that play largely redundant functions, although Abp1 usually shows stronger effects [5, 7–10]. Originally identified as a centromere-associated protein [11–13], CENP-B is important for *de novo* centromere assembly in mammalian artificial chromosomes [14–18], and for heterochromatin formation and silencing in *S. pombe* [19]. Also in *S. pombe*, CENP-B contributes to silencing of *Tf2* retrotransposons [20], regulates directionality of mating-type switching [21–23] and stabilizes stalled replication forks [24, 25]. Our results show that loss of *abp1* results in the accumulation of CUTs. These CUTs are mostly produced from *Tf2* retrotransposons and rDNA repeats. We also show that, in the absence of Abp1, RNAPII gains access to multiple internal cryptic sites in *Tf2* retrotransposons and, most remarkably, in rDNA repeats. Altogether, these results suggest that, in the absence of Abp1, strong cryptic transcription of *Tf2* retrotransposons and rDNA repeats takes place, overloading the exosome degradation pathway and resulting in CUTs accumulation.

2. MATERIALS AND METHODS

2.1. Strains and Media

S. pombe cells were grown in complete medium (YES) or Edinburgh minimal medium (EMM), according to standard procedures. Strains used in these experiments are summarized in Supplementary Table S1. *abp1* ::*KAN*, *abp1* ::*NAT*, *cbh1* ::*NAT* and *cbh2* ::*NAT* were obtained by replacing the entire ORF of the corresponding gene with PCR products carrying KanMX6 or NatMX6 markers obtained from pFA6a plasmids [26, 27]. Abp1 was tagged with 3HA using plasmids pFA6a–3HA [26, 28]. The *abp1 dis3-54* strain was obtained using conventional genetic crosses.

2.2. Short RNAs purification and analyses

Total RNA from *S. pombe* cells was obtained by acid-phenol extraction as described in [29] and quantified using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific). Capillary gel electrophoretic analyses of total RNA were performed at the Functional Genomics facility of IRB (Barcelona) using the RNA 6000 Nano kit for the Bioanalyzer 2100 capillary electrophoresis system (Agilent). For purification of short RNAs, 3 µg of total RNA were loaded on a 1% TAE-low melting point agarose gel. After electrophoretic separation, RNAs of sizes lower and higher than 500nt were selected, eluted from the gel by digestion with gelase (Epicentre) and purified according to manufacturer's instructions. The fraction of purified RNAs of sizes < 500 nt was used to estimate the levels of *Tf2* and 28S rRNA CUTs, while the fraction of purified RNAs > 500 nt was used to estimate the RNAs levels of sizes bigger than CUTs, including the full-length transcripts. After digestion by the gelase, total cDNA was prepared using 5µL of the purified RNAs using hexamer primers and

the Transcriptor First Strand cDNA Synthesis Kit (Roche). The resulting cDNAs were used as templates for quantitative real-time PCR (qPCR). qPCR reactions were performed by using the LightCycler 480 Real-Time PCR System (Roche) and the LightCycler 480 SYBR Green I master mix (Roche). Primers used for RT-qPCR amplification are listed in Supplementary Table S2.

For the sequencing of the short RNAs (RNAseq) a single sequencing run was performed, using two biological replicates per condition. Library construction, cluster generation and sequencing analysis using the Illumina GAII Genome Analyzer followed the manufacturer's protocols (www.illumina.com). In brief, total RNA was quantified by Qubit fluorometer using the Quant-iT RNA Assay Kit (Life Technologies). Using Illumina's small RNA Sample Prep Kit v1.5, 3' and 5' adapters were ligated to 5 µg of total RNA, followed by reverse transcription and 12 cycles of amplification. Amplification products between 125 to 250 nt size were resolved in a 4% agarose gel and purified. Libraries were quantified by Qubit fluorometer using the Quant-iT DNA HS Assay Kit (Life Technologies). Using Illumina's Cluster Generation Kit, libraries were subjected to cluster generation at 8 pM concentration as one sample per lane. Sequencing-by-synthesis was performed for 38 cycles.

2.3. Chromatin-immunoprecipitation (ChIP) experiments

ChIP assays of Abp1 and PolII-Ser5P were essentially performed as previously described [30, 31], using an Abp1-HA tagged strain (SpA129) for Abp1 ChIPseq, and *wt* (spA64) and *abp1* (spA160) strains for PolII-Ser5P ChIPseq. In brief, formaldehyde was added to the cells at a final concentration of 1% for 15 min. After quenching with 125mM Glycine, cells were harvested and lysed with a FastPrep cell disrupter (FP120, Bio101 Thermo Savant). DNA was sheared to a range of 250–400bp with a Bioruptor Standard (Diagenode). Monoclonal anti-HA tag antibody [12CA5] (Abcam16918) was used to retrieve HA-tagged Abp1 combined with Dynabeads Protein A (Life Technologies). For PolII-Ser5P ChIPseq we used αPolIIser5 monoclonal antibody (Abcam/ab5131). After reversing the crosslinks DNA was purified using phenol-chloroform and precipitated with LiCl and isopropanol. One sequencing ChIPseq run was performed for HA-tagged Abp1 and PolII-Ser5P respectively, using one immunoprecipitated sample per condition, together with their respective input samples. For ChIP-qPCR, triplicates for at least 3 biological replicates were analyzed following the Ct method (see Supplementary Table S2 for primers used in each experiment). ChIPseq analyses were performed with 10 ng of ChIP/Input DNA as previously described [32].

ChIP assays were performed on the following histone modifications: H3K9Ac (antibody AB4441, Abcam); H3K14Ac (antibody AB52946, Abcam) and H3K36Me3 (antibody AB9050, Abcam) using the strains ZB1396 (*wt*), ZB1397 (*abp1*), and ZB1655 (*clr3*) grown at 32°C and ZB295 (*clr6-I*) grown at 26°C. A volume of 200 mL cells grown to OD600 = 0.8 were crosslinked with 1% formaldehyde for 20 min. The reaction was stopped by addition of glycine to 125 mM, for 5 min. The cell pellets were washed with 5 ml of TBS buffer containing 1% Triton X-100 and resuspended in 700 µl of FA-lysis buffer (50 mM Tris-HCl [pH 7.5], 200 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, and 1 mM PMSF). Cells were lysed in a mini-bead beater

(MiniBeadBeater-16, Model: 607). Cell lysates were centrifuged for 10 min at 4°C and 13,000rpm. The crude chromatin pellet was washed with 1 mL of FA-lysis buffer and resuspended in 2 mL of FA-lysis buffer. Chromatin was sonicated for 10 cycles for 30s with 30s intervals in FA lysis buffer in a Bioruptor (Diagenode). Samples were spun at 13,000 rpm at 4°C for 10 minutes. FA lysis buffer was added to each supernatant to a final volume of 9.5 mL and 1 mL sample was used for immunoprecipitation. Chromatin was pre-cleared then incubated overnight at 4°C with pre-bound antibody and beads. The antigen-antibody complex was washed successively with 1 ml each of FA-lysis buffer, FA-lysis buffer containing 500 mM NaCl, ChIP wash buffer (10 mM Tris-HCl [pH 7.5], 250 mM LiCl, 0.5% Nonidet-P40, 0.5% sodium deoxycholate, and 1 mM EDTA), and TE buffer. The beads were incubated with 10 µg of DNase-free RNase (Thermo Scientific) for 30 min at 42°C followed by 20 µg Proteinase K (Thermo Scientific) for 1hr at 55°C. The crosslinks were reversed by overnight incubation at 65°C in the same buffer. Samples were extracted with phenol-chloroform and precipitated with ethanol in the presence of linear acrylamide. DNA pellets were resuspended in 50 µl H₂O and used as template for qPCR amplification. For ChIP-qPCR, 3 biological replicates were analyzed following the Ct method (see Supplementary Table S2 for primers used in each experiment).

2.4. Analysis of nucleolar size

Nucleolar size was determined by fluorescence microscopy, using the non-ribosomal protein Gar2-GFP as a marker [33, 34]. Gar2-GFP was expressed from the endogenous locus. Cells were grown in YES medium at 30°C and, without fixation, visualized with an epifluorescence microscope Nikon E1000 using the 100X objective with MetaMorph image analysis software. DNA was stained with Hoechst 33342. The relative nucleolar size was determined as the ratio of the Gar2-GFP area to the total nuclear area determined as the sum of the Hoechst and Gar2-GFP areas together.

2.5. Bioinformatics and Biostatistics analyses

Illumina ChIPseq data for Abp1 and RNAPII-Ser5, together with their respective Input samples, and strand-specific RNAseq data for short RNAs from *wt* and *abp1* cells were aligned against the reference *S. pombe* genome (Sanger release 09052011) using Bowtie 0.12.5, allowing for 2 (ChIPseq) and 0 (RNAseq) mismatches in the read seed. In order to compare read proportions to estimate enrichment and depletion in feature types with multiple copies along the genome (i.e. transposons, rRNA units), all possible alignments for multiple hits were considered for all samples and controls. For ChIPseq data, strand shift bias and duplicated reads were estimated and removed with the htSeqTools package [35] using the filterDuplReads functions with default options, and significantly enriched regions were detected with the enrichedRegions function, considering only regions within the IP sample with a coverage above 2.5 times the first quartile, and using a Benjamini Hochberg p-value threshold of 0.05. Afterwards the enrichedPeaks function was used to detect significantly enriched peaks (putative binding sites), setting the minimum coverage difference between IP and Input at 200. For RNAseq data, *abp1* and *wt* replicates were pooled, and significantly enriched regions in short RNAseq in *abp1* and *wt* samples independently (without control) were detected using the enrichedRegions function from the htSeqTools package, using the one tailed option and a Benjamini Hochbergh pvalue

threshold of 0.05, and leaving the other settings as default. Afterwards, significantly enriched and depleted regions in small RNAseq in *abp1* mutant against *wt* (thus, up and down-regulated short RNAs in *abp1*) were detected using the `enrichedRegions` function from the `htSeqTools` package, using the two tailed option and a Benjamini Hochberg pvalue threshold of 0.05, and leaving the other settings as default.

Annotation of all resulting enriched/depleted regions (RNAseq) and putative binding sites (ChIPseq) against the *S. pombe* annotation 294 v2.20 [36] was performed using the `AnnotatePeakInBatch` function from the `ChIPpeakAnno` package [37], and setting the `maxgap` parameter at 100 bp. ChIPseq (IP/Input) and short RNAs (*abp1* / *wt*) log2 ratios were computed in R, and peak location, mean coverage and coverage ratio plots were produced with the `PeakLocation`, `plotMeanCoverage` and `RegionsCoverage` functions of the `htSeqTools` package respectively, normalizing with respect to the mean value across genomic feature length when necessary. Unless otherwise stated, all downstream statistical analysis of ChIPseq and RNAseq data was performed in R version 2.15.3. Microarray data for *abp1* de-regulated probesets was retrieved from NCBI GEO accession GSE39404 and probeset chromosomal location was annotated against the *S. pombe* annotation 294 v2.20 genome.

3. RESULTS and DISCUSSION

3.1. *abp1* cells accumulate CUTs from *Tf2* retrotransposons and rDNA repeats

Capillary gel electrophoresis analyses of total RNAs obtained from *abp1* cells showed an accumulation of short RNA molecules of sizes ranging between 50 nt and 200 nt (Figure 1). No such increase was observed upon deletion of the other two *S. pombe* CENP-B homologs *cbh1* or *cbh2* (Figure 1). Furthermore, no synergistic effects were detected in *abp1* / *cbh1* or *abp1* / *cbh2* double mutants (Figure 1), suggesting that Cbh1/2 have only minor effects on this accumulation.

RNAseq analyses showed that, as expected from their length, a significant proportion of short RNAs detected in *abp1* corresponded to short non-coding RNAs species, such as 5S and 5.8S rRNAs, and tRNAs (Figure 2A). In addition, we also detected short RNAs corresponding to genomic features of larger size, such as protein-coding genes and, in particular, *Tf2* retrotransposons and 28S and 18S rDNAs, their full-length forms being respectively of 4000 nt, 3485 nt and 1842 nt (Figure 2A), suggesting that a large proportion of these short RNAs corresponded to CUTs species. Note that, although the highest proportion of short RNAs mapped to protein-coding genes (~55%), we detected them in only <5% of the *S. pombe* protein-coding genes (not shown). From the RNAseq data we defined significantly enriched regions in short RNAs in *abp1* and *wt* samples (see material and method). The short-RNAs enriched regions found in *abp1* were smaller in size than the one found in *wt* (Supplementary Figure S3A). We observed that ~65% of the genomic regions enriched in short RNAs were detected in both *abp1* and *wt* cells and 35% (169) were present in *abp1* only (Figure 2B). Compared with *wt* and *abp1*, the 169 short-RNAs enriched regions specific to *abp1* contained a higher proportions of regions mapping to tRNA but were depleted in regions corresponding to the rDNA repeats (Supplementary Figure S3B). Overall, the short RNAs were generally enriched in *abp1* cells in comparison to *wt* cells (Figure 2C), with the highest accumulation corresponding to *Tf2* retrotransposons

and 28S rDNAs (Figure 2C). We confirmed this accumulation by RT-qPCR analyses on gel purified-RNAs of sizes <500 nt (Figure 2D). *Tf2* and rDNA constitute two major types of repetitive DNA elements in the *S. pombe* genome, which contains thirteen *Tf2* copies scattered along the genome [38] and approximately 100–150 rDNA repeats located at both ends of chromosome III [39]. CUTs are usually degraded through the exosome pathway [1–3]. Indeed, we observed that *dis3-54*, a mutation that impairs ribonuclease activity of the essential Dis3 exosome subunit [40], and deletion of *cid14*, a poly(A)polymerase targeting nuclear RNAs to the exosome [41], showed synergistic effects with *abp1* regarding the accumulation of CUTs from both *Tf2* and rDNA elements (Figure 2E). Altogether, these results suggest that Abp1 prevents the accumulation of CUTs from repetitive DNAs, mainly *Tf2* and rDNA elements. Abp1 has recently been shown to collaborate with Set1 to silence sense and antisense *Tf2* transcription [42]. We tested whether Set1 had an effect on the accumulation of short RNAs seen in the *abp1* strain but we saw no increase in CUTs proportion in *set1* compared to *wt* or in the double mutant *set1 /abp1* compared to *abp1* (Supplementary Figure S2A and B), indicating that the Abp1-mediated silencing of cryptic transcription is independent of Set1.

ChIPseq analyses showed a significant co-occurrence of CUTs with Abp1 binding sites (Fisher's test $p < 10^{-42}$) (Figure 3A), suggesting a direct contribution of Abp1 to their regulation. Actually, ~50% of CUTs-enriched regions localized at Abp1 binding sites (Figure 3B). The other ~50% of the CUTs don't appear to be directly associated with an Abp1 binding sites, either because we could not detect the binding sites or because Abp1 was acting on the CUTs production from a bigger distance. In particular, Abp1 binding was detected at the LTRs of *Tf2* elements (Figure 3C, *left*) [20, 25] and all along the rDNA repeats (Figure 3C, *right*). ChIP-qPCR experiments confirmed binding of Abp1 at LTRs of *Tf2* elements and at the 28S rDNA gene (Figure 3D). Abp1 binding was also detected at tRNA and 5S rDNA genes, and at ~10% of the *S. pombe* protein-coding genes (Supplementary Figure S1A). Except in the rDNA repeats, Abp1 binding was generally restricted to the 5' region of the corresponding genomic feature (Supplementary Figure S1B).

3.2. In the absence of Abp1, RNAPII transcribes *Tf2* and rDNA elements from multiple internal cryptic sites

At *Tf2* elements, CUTs were enriched along the entire element including the LTRs (Figure 4A). Similarly, CUTs enrichment was also detected all along the rDNA repeats (Figure 5A), with the highest enrichment observed at the 28S rDNA gene. Notice that at *Tf2* elements CUTs were mostly sense, with antisense CUTs restricted to the LTRs (Figure 4A). On the other hand, in rDNA repeats, we detected both sense and antisense CUTs, the later being slightly more abundant (Figure 5A). Furthermore, in ChIPseq experiments, we detected the active RNAPolII form phosphorylated in Serine 5 (RNAPII-Ser5P), which is normally restricted to promoters (reviewed in [43]), all along *Tf2* elements (Figure 4B). Most remarkably, we also detected RNAPII-Ser5P across rDNA repeats, which are normally transcribed by RNAPI [44] (Figure 5B). ChIP-qPCR analyses confirmed these results (Figures 4C and 5C). Altogether these results suggest that, in the absence of Abp1, RNAPII

initiates transcription from multiple internal cryptic sites distributed along *Tf2* elements and rDNA repeats.

We observed that most differentially expressed Abp1 target genes were upregulated in *abp1* cells (Supplementary Figure S1C). A third of the differentially affected genes were not localized near an Abp1-binding site (Supplementary Figure S1C), suggesting that the absence of Abp1 could also have an indirect role on gene expression, for example through its ability to regulate nuclear organization ([20, 42]). Transcription from internal cryptic sites generally occurs when genes are actively transcribed [45, 46]. In this regard, full-length *Tf2* mRNA levels increased in *abp1* cells (Figure 2D, *left*) [20, 47]. Abp1 has been shown to repress transcription by promoting histone deacetylation through the recruitment of the HDACs Crl3 and Clr6 [20, 42, 47, 48]. Thus, the function of Abp1 in preventing cryptic transcription of *Tf2* elements might reflect its role in establishing a repressive chromatin state. Consistent with this hypothesis, CUTs also accumulated in *crl3* */clr6* mutants as well as in the *hip1* and *hip3* mutants, which are histones chaperones belonging to the HIRA complex involved in nucleosome assembly and repressing *Tf2* and cryptic transcriptions [45, 49] (Figures 4D and 4E). The extent of CUTs accumulation generally correlated with the extent of *Tf2* activation (Figures 4D and 4E). We explored the state of H3K9 and H3K14 acetylation in *Tf2* (Figures 4F and 4G), but did not observe an increase in acetylation in the *abp1* mutant. H3K9Ac was nevertheless increased within the coding sequence of *Tf2* in the *clr6-1* mutant, indicating that this chromatin mark could mediate the increased transcription of CUTs in this strain (Figure 4E). Increased RNAPII transcription in *abp1* might recruit Set2, resulting in H3K36me deposition that in turn could recruit the Clr6 complex II to repress CUT expression [50, 51]. This model could potentially explain the unexpected lack of increased H3 acetylation in *abp1*. We measured the presence of H3K36me on *Tf2* elements, but observed no changes in *abp1*, *clr3* or *clr6* mutants. This suggests that the short abortive transcription of CUTs does not stably recruit Set2, possibly because RNAPII does not reach a ser2-phosphorylated elongating form.

In contrast to *Tf2*, expression of full-length 28S RNAs was unaltered in *abp1* cells (Figure 2D), suggesting that transcription by RNAPI was not affected. However, as shown above (Figure 5B), CUTs accumulation at rDNA repeats is associated with cryptic transcription by RNAPII, which normally is not detected at rDNAs repeats. H3K14ac is significantly increased over the rDNA gene in *abp1*, supporting a model of Abp1-mediated recruitment of Clr3 to repress CUTs (Figure 5F) [52]. However, neither *clr3* nor *clr6* mutants phenocopied *abp1* in this respect, and unlike *clr3* [52], *abp1* mutants do not lose silencing of a *ura4+* reporter gene embedded in the rDNA (See Supplementary Figure S3). Altogether this indicates that other HDACs might be implicated in Abp1-mediated CUT repression in rDNA. Also, in contrast to the case of *Tf2*, rDNA genes accumulated high levels of H3K36me (Figure 5G). Together, these results suggest that *abp1* represses CUT expression through different mechanisms in *Tf2* and rDNA.

Moreover, we observed that nucleolus size is enlarged by ~20% in *abp1* cells (Figure 5D). The augmentation in nucleolus size in *abp1* is not due to an increase in rDNA copy number, because the length of the rDNA repeats on chromosome III remains unchanged in *abp1* ([25] and our unpublished results). Altogether these results suggest that nucleolus

organization is disturbed in the absence of Abp1 and that this alteration does not affect RNAPII transcription but results in deregulated RNAPII binding and transcription. Interestingly, although normally transcribed by RNAPII, cryptic RNAPII transcription of rDNAs genes has been reported in both budding and fission yeasts, and is believed to be important for stability and expression of the rDNA array [53, 54]. The nucleolus constitutes a highly structured nuclear compartment [55]. However, the factors and mechanisms that determine nucleolus organization are still poorly understood. In this regard, our results suggest a contribution of Abp1 to the structural organization of nucleolar chromatin. Interestingly, Abp1 is known to mediate clustering of *Tf2* elements [20] and it locates at 5S rRNA and tRNA genes that, being transcribed by RNAPIII, also form clusters in *S. pombe* [56–58]. Altogether these observations suggest that Abp1 might play a general structural role in the organization of the *S. pombe* genome.

3.3. Conclusions

Here, we have shown that deletion of *abp1* induces the accumulation of CUTs from *Tf2* and rDNA repeats. These CUTs are product of RNAPII transcription, initiate at multiple internal cryptic sites and are degraded by the exosome pathway. Altogether, these results suggest that Abp1 prevents pervasive RNAPII transcription at *Tf2* elements and rDNA repeats. Abp1 is also detected at promoters of ~10% of protein-coding genes, and at most 5S rRNA and tRNA genes. Whether Abp1 is also preventing pervasive transcription at these sites remains to be determined.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Drs F. Antequera, G. Thon, D. Moazed, S. Whitehall and S. Wang for yeast strains. We are thankful to Evarist Planet and Alicia Vera for technical assistance. We also thank Dr. Tomás Morán for very useful discussions. This work was supported by grants from MICINN (BFU2012-30724 and BFU2015-65082-P), the Generalitat de Catalunya (SGR2009-1023 and SGR2014-204) and the European Community FEDER program. This work was carried out within the framework of the “Centre de Referència en Biotecnologia” of the “Generalitat de Catalunya”. AD was funded by a Marie Curie Reintegration Grant within the 7th European Community Framework Program PIRG7-GA-2010-268400.

References

1. Jacquier A. The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. *Nat Rev Genet.* 2009; 10:833–844. [PubMed: 19920851]
2. Jensen TH, Jacquier A, Libri D. Dealing with pervasive transcription. *Mol Cell.* 2013; 52:473–484. [PubMed: 24267449]
3. Rougemaille M, Libri D. Control of cryptic transcription in eukaryotes. *Adv Exp Med Biol.* 2010; 702:122–131. [PubMed: 21618879]
4. Casola C, Hucks D, Feschotte C. Convergent domestication of pogo-like transposases into centromere-binding proteins in fission yeast and mammals. *Mol Biol Evol.* 2008; 25:29–41. [PubMed: 17940212]
5. Ireland JT, Gutkin GI, Clarke L. Functional redundancies, distinct localizations and interactions among three fission yeast homologs of centromere protein-B. *Genetics.* 2001; 157:1191–1203. [PubMed: 11238404]

6. Smit AF, Riggs AD. Tiggers and DNA transposon fossils in the human genome. *Proc Natl Acad Sci USA*. 1996; 93:1443–1448. [PubMed: 8643651]
7. Baum M, Clarke L. Fission yeast homologs of human CENP-B have redundant functions affecting cell growth and chromosome segregation. *Mol Cell Biol*. 2000; 20:2852–2864. [PubMed: 10733588]
8. Halverson D, Baum M, Stryker J, Carbon J, Clarke L. A centromere DNA-binding protein from fission yeast affects chromosome segregation and has homology to human CENP-B. *J Cell Biol*. 1997; 136:487–500. [PubMed: 9024682]
9. Lee JK, Huberman JA, Hurwitz J. Purification and characterization of a CENP-B homologue protein that binds the centromeric K-type repeat DNA of *Schizosaccharomyces pombe*. *Proc Natl Acad Sci USA*. 1997; 94:8427–8432. [PubMed: 9237993]
10. Murakami Y, Huberman JA, Hurwitz J. Identification, purification, and molecular cloning of autonomously replicating sequence-binding protein 1 from fission yeast *Schizosaccharomyces pombe*. *Proc Natl Acad Sci USA*. 1996; 93:502–507. [PubMed: 8552670]
11. Earnshaw WC, Sullivan KF, Machlin PS, Cooke CA, Kaiser DA, Pollard TD, Rothfield NF, Cleveland DW. Molecular cloning of cDNA for CENP-B, the major human centromere autoantigen. *J Cell Biol*. 1987; 104:817–829. [PubMed: 2435739]
12. Masumoto H, Masukata H, Muro Y, Nozaki N, Okazaki T. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J Cell Biol*. 1989; 109:1963–1973. [PubMed: 2808515]
13. Muro Y, Masumoto H, Yoda K, Nozaki N, Ohashi M, Okazaki T. Centromere protein B assembles human centromeric alpha-satellite DNA at the 17-bp sequence, CENP-B box. *J Cell Biol*. 1992; 192:585–596.
14. Ebersole TA, Ross A, Clark E, McGill N, Schindelbauer D, Cooke H, Grimes B. Mammalian artificial chromosome formation from circular alphoid input DNA does not require telomere repeats. *Hum Mol Genet*. 2000; 9:1623–1631. [PubMed: 10861289]
15. Grimes BR, Rhoades AA, Willard HF. Alpha-satellite DNA and vector composition influence rates of human artificial chromosome formation. *Mol Ther*. 2002; 5:798–805. [PubMed: 12027565]
16. Harrington JJ, Van Bokkelen G, Mays RW, Gustashaw K, Willard HF. Formation of *de novo* centromeres and construction of first-generation human artificial microchromosomes. *Nat Genet*. 1997; 15:345–355. [PubMed: 9090378]
17. Ikeno M, Grimes B, Okazaki T, Nakano M, Saitoh K, Hoshino H, McGill NI, Cooke H, Masumoto H. Construction of YAC-based mammalian artificial chromosomes. *Nat Biotechnol*. 1998; 16:431–439. [PubMed: 9592390]
18. Okada T, Ohzeki J, Nakano M, Yoda K, Brinkley WR, Larionov V, Masumoto H. CENP-B controls centromere formation depending on the chromatin context. *Cell*. 2007; 131:1287–1300. [PubMed: 18160038]
19. Nakagawa H, Lee J-K, Hurwitz J, Allshire RC, Nakayama J-i, Grewal SIS, Tanaka K, Murakami Y. Fission yeast CENP-B homologs nucleate centromeric heterochromatin by promoting heterochromatin-specific histone tail modifications. *Genes Dev*. 2002; 16:1766–1778. [PubMed: 12130537]
20. Cam HP, Noma K-i, Ebina H, Levin HL, Grewal SIS. Host genome surveillance for retrotransposons by transposon-derived proteins. *Nature*. 2007; 451:431–436. [PubMed: 18094683]
21. Aguilar-Arnal L, Marsellach FX, Azorín F. The fission yeast homologue of CENP-B, Abp1, regulates directionality of mating-type switching. *EMBO J*. 2008; 27:1029–1038. [PubMed: 18354497]
22. Matsuda E, Sugioka-Sugiyama R, Mizuguchi T, Mehta S, Cui B, Grewal SI. A homolog of male sex-determining factor SRY cooperates with a transposon-derived CENP-B protein to control sex-specific directed recombination. *Proc Natl Acad Sci USA*. 2011; 108:18754–18759. [PubMed: 22042869]
23. Yu C, Bonaduce MJ, Klar AJ. Going in the right direction: mating-type switching of *Schizosaccharomyces pombe* is controlled by judicious expression of two different *swi2* transcripts. *Genetics*. 2012; 190:977–987. [PubMed: 22209903]

24. Locovei AM, Spiga M-G, Tanaka K, Murakami Y, D'Urso G. The CENP-B homolog, Abp1, interacts with the initiation protein Cdc23(MCM10) and is required for efficient DNA replication in fission yeast. *Cell Division*. 2006; 1:27. [PubMed: 17112379]
25. Zaratiegui M, Vaughn MW, Irvine DV, Goto D, Watt S, Bähler J, Arcangioli B, Martienssen RA. CENP-B preserves genome integrity at replication forks paused by retrotransposon LTR. *Nature*. 2011; 469:112–115. [PubMed: 21151105]
26. Bahler J, Wu JQ, Longtine MS, Shah NG, McKenzie A 3rd, Steever AB, Wach A, Philippsen P, Pringle JR. Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast*. 1998; 14:943–951. [PubMed: 9717240]
27. Hentges P, Van Driessche B, Tafforeau L, Vandenhoute J, Carr AM. Three novel antibiotic marker cassettes for gene disruption and marker switching in *Schizosaccharomyces pombe*. *Yeast*. 2005; 22:1013–1019. [PubMed: 16200533]
28. Van Driessche B, Tafforeau L, Hentges P, Carr AM, Vandenhoute J. Additional vectors for PCR-based gene tagging in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* using nourseothricin resistance. *Yeast*. 2005; 22:1061–1068. [PubMed: 16200506]
29. Schmitt ME, Brown TA, Trumpower BL. A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucleic acids research*. 1990; 18:3091–3092. [PubMed: 2190191]
30. Du LL, Nakamura TM, Russell P. Histone modification-dependent and -independent pathways for recruitment of checkpoint protein Crb2 to double-strand breaks. *Genes Dev*. 2006; 20:1583–1596. [PubMed: 16778077]
31. Langerak P, Mejia-Ramirez E, Limbo O, Russell P. Release of Ku and MRN from DNA ends by Mre11 nuclease activity and Ctp1 is required for homologous recombination repair of double-strand breaks. *PLoS Genet*. 2011; 7:e1002271. [PubMed: 21931565]
32. Lloret-Llinares M, Perez-Lluch S, Rossell D, Moran T, Ponsa-Cobas J, Auer H, Corominas M, Azorin F. dKDM5/LID regulates H3K4me3 dynamics at the transcription-start site (TSS) of actively transcribed developmental genes. *Nucleic acids research*. 2012; 40:9493–9505. [PubMed: 22904080]
33. Gulli MP, Girard JP, Zabetakis D, Lapeyre B, Melese T, Caizergues-Ferrer M. gar2 is a nucleolar protein from *Schizosaccharomyces pombe* required for 18S rRNA and 40S ribosomal subunit accumulation. *Nucleic acids research*. 1995; 23:1912–1918. [PubMed: 7596817]
34. Shimada T, Yamashita A, Yamamoto M. The fission yeast meiotic regulator Mei2p forms a dot structure in the horse-tail nucleus in association with the sme2 locus on chromosome II. *Molecular biology of the cell*. 2003; 14:2461–2469. [PubMed: 12808043]
35. Planet E, Attolini CS, Reina O, Flores O, Rossell D. htSeqTools: high-throughput sequencing quality control, processing and visualization in R. *Bioinformatics*. 2012; 28:589–590. [PubMed: 22199381]
36. Wood V, Harris MA, McDowall MD, Rutherford K, Vaughan BW, Staines DM, Aslett M, Lock A, Bahler J, Kersey PJ, Oliver SG. PomBase: a comprehensive online resource for fission yeast. *Nucleic acids research*. 2012; 40:D695–D699. [PubMed: 22039153]
37. Zhu LJ, Gazin C, Lawson ND, Pages H, Lin SM, Lapointe DS, Green MR. ChIPpeakAnno: a Bioconductor package to annotate ChIPseq and ChIP-chip data. *BMC Bioinformatics*. 2010; 11:237. [PubMed: 20459804]
38. Bowen NJ, Jordan IK, Epstein JA, Wood V, Levin HL. Retrotransposons and their recognition of pol II promoters: a comprehensive survey of the transposable elements from the complete genome sequence of *Schizosaccharomyces pombe*. *Genome Res*. 2003; 13:1984–1997. [PubMed: 12952871]
39. Maleszka R, Clark-Walker GD. Yeasts have a four-fold variation in ribosomal DNA copy number. *Yeast*. 1993; 9:53–58. [PubMed: 8442387]
40. Murakami H, Goto DB, Toda T, Chen ES, Grewal SI, Martienssen RA, Yanagida M. Ribonuclease activity of Dis3 is required for mitotic progression and provides a possible link between heterochromatin and kinetochore function. *PLoS One*. 2007; 2:e317. [PubMed: 17380189]
41. Wang SW, Stevenson AL, Kearsey SE, Watt S, Bähler J. Global role for polyadenylation-assisted nuclear RNA degradation in posttranscriptional gene silencing. *Mol Cell Biol*. 2008; 28:656–665. [PubMed: 18025105]

42. Lorenz DR, Mikheyeva IV, Johansen P, Meyer L, Berg A, Grewal SI, Cam HP. CENP-B cooperates with Set1 in bidirectional transcriptional silencing and genome organization of retrotransposons. *Mol Cell Biol.* 2012; 32:4215–4225. [PubMed: 22907751]
43. Heidemann M, Hintermair C, Voss K, Eick D. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim Biophys Acta.* 2013; 1829:55–62. [PubMed: 22982363]
44. Boukhgalter B, Liu M, Guo A, Tripp M, Tran K, Huynh C, Pape L. Characterization of a fission yeast subunit of an RNA polymerase I essential transcription initiation factor, SpRrn7h/TAF(I)68, that bridges yeast and mammals: association with SpRrn11h and the core ribosomal RNA gene promoter. *Gene.* 2002; 291:187–201. [PubMed: 12095692]
45. Anderson HE, Wardle J, Korkut SV, Murton HE, Lopez-Maury L, Bahler J, Whitehall SK. The fission yeast HIRA histone chaperone is required for promoter silencing and the suppression of cryptic antisense transcripts. *Mol Cell Biol.* 2009; 29:5158–5167. [PubMed: 19620282]
46. Nicolas E, Yamada T, Cam HP, Fitzgerald PC, Kobayashi R, Grewal SI. Distinct roles of HDAC complexes in promoter silencing, antisense suppression and DNA damage protection. *Nat Struct Mol Biol.* 2007; 14:372–380. [PubMed: 17450151]
47. Hansen KR, Burns G, Mata J, Volpe TA, Martienssen RA, Bahler J, Thon G. Global effects on gene expression in fission yeast by silencing and RNA interference machineries. *Mol Cell Biol.* 2005; 25:590–601. [PubMed: 15632061]
48. Sugiyama T, Cam HP, Sugiyama R, Noma K, Zofall M, Kobayashi R, Grewal SI. SHREC, an effector complex for heterochromatic transcriptional silencing. *Cell.* 2007; 128:491–504. [PubMed: 17289569]
49. Greenall A, Williams ES, Martin KA, Palmer JM, Gray J, Liu C, Whitehall SK. Hip3 interacts with the HIRA proteins Hip1 and Slm9 and is required for transcriptional silencing and accurate chromosome segregation. *J Biol Chem.* 2006; 281:8732–8739. [PubMed: 16428807]
50. Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP, Workman JL. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell.* 2005; 123:581–592. [PubMed: 16286007]
51. Keogh MC, Kurdastani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ, Boone C, Emili A, Weissman JS, Hughes TR, Strahl BD, Grunstein M, Greenblatt JF, Buratowski S, Krogan NJ. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell.* 2005; 123:593–605. [PubMed: 16286008]
52. Bjerling P, Silverstein RA, Thon G, Caudy A, Grewal S, Ekwall K. Functional divergence between histone deacetylases in fission yeast by distinct cellular localization and in vivo specificity. *Mol Cell Biol.* 2002; 22:2170–2181. [PubMed: 11884604]
53. Castel SE, Ren J, Bhattacharjee S, Chang AY, Sanchez M, Valbuena A, Antequera F, Martienssen RA. Dicer promotes transcription termination at sites of replication stress to maintain genome stability. *Cell.* 2014; 159:572–583. [PubMed: 25417108]
54. Vasiljeva L, Kim M, Terzi N, Soares LM, Buratowski S. Transcription termination and RNA degradation contribute to silencing of RNA polymerase II transcription within heterochromatin. *Mol Cell.* 2008; 29:313–323. [PubMed: 18280237]
55. Nemeth A, Langst G. Genome organization in and around the nucleolus. *Trends Genet.* 2011; 27:149–156. [PubMed: 21295884]
56. Iwasaki O, Tanaka A, Tanizawa H, Grewal SI, Noma K. Centromeric localization of dispersed Pol III genes in fission yeast. *Molecular biology of the cell.* 2010; 21:254–265. [PubMed: 19910488]
57. Kirkland JG, Raab JR, Kamakaka RT. TFIIC bound DNA elements in nuclear organization and insulation. *Biochim Biophys Acta.* 2013; 1829:418–424. [PubMed: 23000638]
58. Thompson M, Haeusler RA, Good PD, Engelke DR. Nucleolar clustering of dispersed tRNA genes. *Science.* 2003; 302:1399–1401. [PubMed: 14631041]

Figure 1.

abp1 cells show increased levels of short RNAs. A) Capillary gel-electrophoresis analyses of total RNA obtained from *wt* (lane 2), *abp1* (lane 3), *cbh1* (lane 4), *cbh2* (lane 5), *abp1 cbh1* (lane 6), *abp1 cbh2* (lane 7) and *cbh1 cbh2* (lane 8) cells. Lane 1 corresponds to molecular weight markers (bp). The positions corresponding to the 28S and 18S rRNAs are indicated. The region corresponding to short RNAs is also indicated. B) Quantitative analysis of the results shown in A. The percentage of short RNAs respect to total RNA is presented for the indicated strains. The data presented are the averages of at least 4 biological replicates. Error bars indicate SEM. P values report significance by one-way ANOVA followed by Dunnett's multiple comparison test (** $P < 0.01$, *** $P < 0.001$).

Figure 2.

Loss of Abp1 results in the accumulation of CUTs produced from *Tf2* retrotransposons and rDNA repeats. A) The percentage of short RNAs-enriched regions mapping at different genomic features in *wt* and *abp1* cells is shown (results correspond to RNAseq analyses). B) Venn diagrams showing the overlap of short RNAs-enriched regions in *wt* and *abp1* (results correspond to RNAseq analyses). C) The fold-change expression of differentially expressed short RNAs in *abp1* vs *wt* is presented for different genomic features (results correspond to RNAseq analyses). Error bars indicate SEM. D) RT-qPCR analyses from purified RNAs of sizes <500 nt (CUTs) and >500 nt of *Tf2* (left) and 28S rRNA (right) expression levels in *abp1* compared to *wt*. The data presented are the averages of at least 16 biological replicates. Error bars indicate SEM. P values report significance by two-tailed one-sample t-test (***P<0.001). E) RT-qPCR analyses from purified RNAs of sizes <500 nt (CUTs) of *Tf2* (left) and 28S rRNA (right) expression levels in different mutant strains. The data presented are the averages of at least 3 biological replicates. Error bars indicate SEM. P values report significance by two-tailed one-sample t-test on the log fold changes (*P<0.05, ***P<0.001). D) and E) Primers amplify a 3121–3193 bp region of the *Tf2* transcript and a 673–756 bp region of the 28S rRNA transcript.

Figure 3.

Abp1 binding correlates with the pattern of CUTs accumulation. A) Location of Abp1 binding sites determined by ChIPseq relative to the start of CUTs determined by RNAseq. B) Venn diagrams showing the overlap between Abp1-binding sites and CUTs-enriched regions. C) Normalized Abp1 enrichment ratio (IP/Input) along *Tf2* retrotransposons (left) and rDNA repeats (right). D) ChIP-qPCR analysis of Abp1 binding at *Tf2* LTRs and 28S rDNA. Actin was used as a negative control. The data presented are the averages of at least 3 biological replicates. Error bars indicate SEM. P values report significance by two-tailed one-sample t-test (* $P < 0.05$, ** $P < 0.01$).

Figure 4.

In *abp1* cells, CUTs originate across the entire *Tf2* element. A) Normalized CUTs coverage ratio (*abp1* /wt) along *Tf2*. Both sense and antisense CUTs are shown. B) Normalized RNAII-Ser5P coverage ratio (IP/Input) along *Tf2*. C) ChIP-qPCR analysis of RNAII-Ser5P binding along *Tf2*. The location of the positions analyzed is indicated. The data presented are the averages of at least 3 biological replicates. Error bars indicate SEM. D) RT-qPCR analyses from purified RNAs of sizes <500 nt (CUTs) and >500 nt of *Tf2* expression levels in *hip1* and *hip3* compared to *wt*. The data presented are the averages of at least 4 biological replicates. Error bars indicate SEM. E) RT-qPCR analyses of *Tf2* expression levels from purified RNAs of sizes <500 nt (CUTs) and >500 nt in *crl3*, *crl6-1* and *crl3 crl6-1* compared to *wt*. The data presented are the averages of at least 2 biological replicates. Error bars indicate SD. D) and E) Primers amplify a 3121–3193 bp region of the *Tf2* transcript C) D) and E) P values report significance by two-tailed one-sample t-test (* $P < 0.05$, *** $P < 0.001$). F) G) H) ChIP-qPCR analysis of H3K9Ac (F) H3K14Ac (G) H3K36Me (H) along *Tf2* in *wt*, *abp1*, *crl3* and *crl6-1*. The location of the positions analyzed is indicated. The data presented are the averages of at least 3 biological replicates. Error bars indicate SEM. P values report significance of differences with *wt* by Tukey's HSD test.

Figure 5.

In *abp1* cells, rDNA repeats are transcribed by RNAPII. A) Normalized CUTs coverage ratio (*abp1*Δ/wt) along rDNA repeats. Both sense and antisense CUTs are shown. The positions of the 5.8S, 18S and 28S rDNA genes are indicated. B) Normalized RNAPII-Ser5P coverage ratio (IP/input *abp1*Δ vs wt) along rDNA repeats. The positions of the 5.8S, 18S and 28S rDNA genes are indicated. C) ChIP-qPCR analysis of RNAPII-Ser5P binding along rDNA repeats. The location of the positions analyzed is indicated. Error bars indicate SEM. P values report significance by two-tailed one-sample t-test (*P<0.05). D) At the top, fluorescence micrographs showing GFP-Gar2 and DNA (Hoestch 33342) localization in *wt* and *abp1*Δ. At the bottom, the relative nucleolar size in *wt* and *abp1*Δ is shown as the ratio of the GFP signal area to the total DNA area. Measurements were performed using Image J software. The data presented are the averages of least 300 cells (scored in 2 independent experiments). Error bars indicate SEM. P values report significance by two-tailed unpaired t-test (***P<0.001). E) F) G) ChIP-qPCR analysis of H3K9Ac (E) H3K14Ac (F) H3K36me (G) along the rDNA repeats in *wt*, *abp1*Δ, *clr3*Δ and *clr6-1*. The location of the positions analyzed is indicated. The data presented are the averages of at least 3 biological replicates.

Error bars indicate SEM. P values report significance of differences with *wt* by Tukey's HSD test.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript