Use of bacterial Rho helicase to gain new insights into the targeting mechanism of nuclear RNAs by the exosome-associated exoribonuclease Rrp6 and its cofactors in yeast

Abstract

In eukaryotic cells, aberrant mRNPs with processing and packaging defects are targeted co-transcriptionally by a surveillance system that triggers their nuclear retention and ultimately the degradation of their mRNA component by the 3'-5' activity of the exosome-associated exoribonuclease Rrp6 along with its associated factors Rrp47 and Mpp6. This mRNP quality control process is stimulated by the NNS complex (Nrd1-Nab3-Sen1), which otherwise mediates termination, processing and decay of ncRNAs and the process involves additionally the exosome co-activator TRAMP complex (Tf4-Air2-Mtr4). Here, we describe a genome-wide approach to visualize the dynamic recruitment of these quality control components over the yeast chromosomes upon global perturbation of mRNP biogenesis by the RNA-dependent helicase/translocase activity of the bacterial Rho factor.

The method provides valuable information on how the surveillance system coordinates with the transcription machinery in order to detect faulty events during perturbation of mRNP biogenesis. In addition, our overview shows that the assembly of the quality control components for the affected mRNA genes takes place at the expense of their commitment to be recruited to ncRNAs genomic features, which in turn leads to termination and processing defects of ncRNAs.

INTRODUCTION

In eukaryotic cells messenger RNA production is a multistep process in which the genetic message is transcribed from DNA into a pre-mRNA molecule that undergoes numerous processing and packaging reactions co-transcriptionally (1). This process is coordinated by the unique features of the C-terminal domain (CTD) of the largest subunit of the RNA polymerase II (RNAP II) which serves as a platform for sequential recruitment of acting factors (2). The RNAP II CTD is a highly conserved and unstructured domain that forms a long tail of 26 (in yeast) to 52 (in mammals) repeats of the heptapeptide (Y1S2P3T4S5P6S7) where a dynamic process of phosphorylation/dephosphorylation of serines S2, S5, and S7 during transcription constitutes a code that adjusts the sequential recruitment of the different factors (3, 4). Emerging pre-mRNA molecule undergoes numerous modifications such as 5'end capping, splicing, 3'end cleavage and polyadenylation while being assembled into a ribonucleoprotein particle (mRNP) that will be exported to the
cytoplasm for translation (5). The first event in co-transcriptional processing is addition of a protective 7-methylguanosine (m7G) cap moiety to the 5′-end of the nascent transcript (6). During the transcription elongation, the emerging transcript is sequentially loaded with various factors that accomplish further processing reactions including the removal of introns and the cleavage of the RNA chain at the 3′-end of the message followed by the addition of a polyA tail (7). The final step is transport of mRNP to the nuclear pore for export into the cytoplasm which includes both a recruitment of a large set of proteins important for mRNP assembly and proteins important for transport (1, 5). Each of these steps is subject to errors, thus cells have developed mechanisms to detect aberrant mRNPs resulting from faulty events, which leads to their targeting and degradation inside the nucleus before transport to the cytoplasm. The mRNA component of the transcripts recognized as faulty is degraded by ribonuclease activities associated by RNA exosome (8, 9). RNA exosome is a eukaryotic ribonuclease complex conserved from yeast to human, which is essential for regulation of gene expression and RNA quality control leading to regulation of diverse cellular processes in eukaryotes. The overall molecular architecture of the eukaryotic exosome is made of six different subunits to form a ring-shaped structure. The six-member ring (Rrp41, Rrp42, Mtr4, Rrp43, Rrp45 and Rrp46) does not assemble into a stable structure in the absence of Rrp4, Rrp40 and Csd4 indicating that the eukaryotic exosome core is composed of nine subunits. The nine subunits exosome core is catalytically inert in most tested organisms and it associates with the Rrp6 and Dis3, which are hydrolytic exoribonucleases that display activity on unstructured and poly(A)-extended RNAs as well as structured substrates (10). Known targets of the exosome/Rrp6p include short RNAP II transcripts such as small noncoding nuclear or nucleolar RNAs (sn/snoRNAs), cryptic unstable transcripts (CUTs) and improperly terminated RNAs. Exosome/Rrp6p is also required for the regulation of polyA tail length, proper 3′-end trimming of small nuclear/nucleolar RNAs (sn/snoRNAs) and termination of specific short transcripts. While for discriminating degradation of different RNAs by Rrp6, an important role is played by various co-factors, it is still not well understood how Rrp6 accomplishes this task (11). However, it is largely recognized that the exosome/Rrp6p interacting partners (co-factors) play key roles in these distinctive processes. These associating partners include Rrp47, Mpp6, some components of the TRAMP complex (Trf4-Air2-Mtr4) as well as components of the NNS complex Nrd1 and Nab3 (12, 13). In Drosophila melanogaster, Rrp6-dependent processing activity is required for cell proliferation and mitosis without core exosome (14) but the mechanism by which Rrp6 acts independently from the core exosome in nuclear quality control (QC) is still unknown.

The difficulty to decipher the actual mechanism of action underlying the QC process resides in the scarcity of mRNP biogenesis defects occurring under physiological context. Indeed, the putative prevalence of mRNP aberrancies occurring under physiological conditions does not allow a direct study of the QC mechanism within the cellular context. To circumvent this limitation, we have implemented a new experimental approach based on global perturbation of mRNP biogenesis by the RNA-dependent helicase/translocase activity of the bacterial Rho factor. The conditionally heterologous expression of Rho in yeast induces a dose-dependent growth defect that stems from its interference in the co-transcriptional recruitment of mRNA processing and packaging factors yielding mRNPs that are recognized as aberrant and eliminated by the Rrp6-dependent nuclear QC system (15). In this approach, sufficient amounts of defective mRNPs can be produced by Rho action in the cell nucleus to serve as substrates for the QC, which facilitates deep investigations of the surveillance process (reviewed in (16)). Using this methodology, we have shown that the QC process is coordinated by Nrd1, a component of the NNS (Nrd1-Nab3-Sen1) complex known to mediate termination and processing of the non-coding RNAs (ncRNAs) and which is recruited by the CTD of RNAP II through its CTD-interacting domain (CID) (17). The Rho-mediated production of aberrant mRNPs stimulates the recruitment of Nrd1 with its partner Nab3 to the transcription complex, which promotes the attachment of Rrp6 along with its co-activating components Rrp47 and Trf4 (12, 13, 18).

In this report, we describe our Rho-based experimental approach and outline its use to investigate the mRNP QC activities over the whole yeast genome. High-throughput RNA sequencing (RNA-seq) was used to survey a broad collection of mRNPs whose biogenesis is affected by Rho action. This genome-wide perspective of affected mRNPs was extended further by generating high-resolution maps (ChIP-seq) of the distribution of the QC components along the yeast chromosomes before and after perturbation of mRNP biogenesis by Rho. The results show that the perturbation of mRNP biogenesis redistributes the binding of QC components over the genome with a large shift of Nrd1 and Nab3 from genomic loci producing ncRNAs prone to processing and decay such as CUTs (Cryptic Unstable Transcripts), SUTs (Stable Unannotated Transcripts) and snoRNAs (small nucleolar RNAs) to protein-coding genes. Nrd1 and Nab3 are apparently titrated out from ncRNAs genomic features by a large recruitment to Rho-affected mRNA gene loci, leading to transcriptional termination defects of ncRNAs (19).

**Rho-based perturbation of global mRNP biogenesis in yeast**

Bacterial Rho factor is a hexameric helicase/translocase that binds to naked, C-rich and preferentially unstructured portions of RNA. This binding activates the ATPase activity of the protein leading to its translocation along the RNA chain in a 5′ to 3′ direction. Numerous studies in our laboratory have shown that Rho has a loose sequence requirement for binding to a transcript. Any RNA se-
The targeting mechanism of nuclear RNAs
Igor Stuparević et al.

The targeting mechanism of nuclear RNAs
Igor Stuparević et al.


quence with relatively well spaced Cs can serve for Rho loading (20). Also, certain RNA structures such as hairpins can serve as efficient loading sites (21). Rho is able to compete with other proteins for binding, a process widely used in prokaryotic gene regulation (22). Furthermore, Rho functions as a powerful helicase/translocase that can track on an RNA chain and melt nucleic acid base pairs or disrupt a protein-RNA complex present on its path (23–25). Given these functional features, we surmised that once in the nucleus of a eukaryotic cell, Rho could bind to nascent transcripts by competing with endogenous mRNA processing and packaging factors and even displace them during translocation along the RNA chain. Such action would potentially lead to the production of mRNPs aberrantly processed or depleted of some crucial proteins.

Indeed, in this experimental approach, large amounts of defective mRNPs can be generated by the activity of bacterial Rho factor in yeast cells to serve as substrates to study the QC mechanism (reviewed in (16)). As illustrated in Figure 1A, Rho expressed from a centromeric plasmid under the control of the Doxycycline (Doxy) repressible TetO7 (Tet-off) promoter is directed to the nucleus by a nuclear localization signal fused to the C-terminus of the protein. The growth of S. cerevisiae BMA41 strain transformed with the Rho expression construct was monitored by serial dilutions on agar plates in the presence or absence of Doxy (Figure 1B). The growth of yeast cells expressing Rho-NLS protein is severely reduced under inducing conditions (−Doxy). However, the growth defect is readily alleviated in yeast strains having a deletion of the RRP6 gene (rrp6∆) or harboring a point mutation in the catalytic site of Rrp6 protein (Rrp6-Y361A). Our previous investigations with a model gene (PMA1) showed that Rho action along the nascent transcript interferes with the normal recruitment of mRNA processing and packaging factors yielding mRNPs that are recognized as defective and eliminated by the QC apparatus in a process involving the 3′→5′ exonuclease activity of Rrp6 (Figure 1C).
monitoring the growth defect alleviation and the rescue of the PMA1 mRNA level in different deletion strains, we showed that the targeting and degradation of Rho-induced aberrant mRNPs by Rrp6 activity is stimulated by the NNS components Nrd1 and Nab3, the exosome cofactors Rrp47 and Mpp6 as well as the TRAMP complex (12, 13, 19). Deletion or alteration of any of these components alleviates the Rho-induced growth defect phenotype of the yeast cells and rescue the model transcript PMA1. In this connection, the rrp47Δ strain was used regularly as a QC-defective representative strain for the ease of manipulation (absence of genetic instability and absence of temperature sensitivity for growth).

**High throughput determination of Rho-induced aberrant mRNPs targeted by the Rrp6-dependent QC**

To survey the effect of Rho on a large collection of mRNAs, we performed stranded RNA-seq using rRNA-depleted RNAs isolated from wild-type and rrp47Δ yeast cells grown under Rho-repressing or Rho-inducing conditions (19). For each condition, two biologically independent samples were used for libraries preparation and sequencing. In general, the results of the duplicates were highly correlated with Spearman correlation coefficients above 0.98. The RNA-seq single reads were aligned on the *S. cerevisiae* genome and the mapped reads matching annotated transcripts were distributed into different RNA biotypes. About 60% of the mapped transcripts were protein-coding mRNAs and the remaining 40% represented ncRNAs including CUTs, SUTs, XUTs, snRNAs and snoRNAs. Upon Rho expression in the wild-type cells, a substantial fraction of the transcripts within all RNA biotypes showed variations, either up or down, in their steady-state levels. We focused our analyses on the Rho-induced variations in the abundance of mRNAs to have insights into the fate of defective mRNPs. We calculated the statistical significance of the variations (mRNA ratios between +Rho and −Rho conditions) and plotted the log2 fold change of the level of each mRNA as a function of its log2 mean expression. Among the

---

**Figure 2** Fate of subpopulations of mRNAs down regulated by Rho action in yeast wild-type strain (Wt) and their rescue in a strain defective in QC (rrp47Δ). Meta-transcript profiles variations of mRNAs obtained from Wt and the Rrp6 depleted rrp47Δ strain in the absence or presence of Rho. The top graph represents the down-regulated mRNAs subpopulation in the presence of Rho in Wt cells. The bottom graph displays the meta-transcript profiles of the pool of mRNAs down regulated by Rho action in Wt strain and which were rescued upon Rrp47 depletion (rrp47Δ strain). Meta-profiles plots were made using bedtools 24 V2.25.0 with computeMatrix (scale-regions mode, with a bin size of 1 base) on BigWig files and plotProfile for graphical output (19).
5,132 mRNAs mapped with the RNA-seq results, 1,015 mRNAs showed decreased levels upon Rho expression in wild-type cells with a log2 fold change ≤ −0.5. When the statistical significance calculations (ratios between +Rho and −Rho mRNA levels) and plotting were applied to the RNA-seq datasets of the mRNAs produced in the rrp47Δ strain, about half (491 transcripts) of the mRNAs that were diminished by Rho action in wild-type cells were

Figure 3. Rho activity mediates the stimulation of massive recruitment of the QC components (Nrd1, Nab3, Rrp6 and Trf4) to rrp6-sensitive mRNAs genomic loci at the expense of their locations within ncRNAs genomic features. (A) IGV (Integrative Genomics Viewer) snapshots of ChIP-seq peaks mapped over two rescued mRNAs genomic loci. The ChIP signals for each protein are shown both for −Rho and +Rho conditions. (B) Circos plots summarizing the observed dynamic landscape of the four QC components over the 16 yeast chromosomes analyzed under −Rho (left plots) and +Rho (right plots) conditions. The two top plots show the ChIP signals detected for the ncRNAs genomic loci (CUTs, SUTs and snoRNAs). The two plots at the bottom show the ChIP signals detected for the genomic features of the Rho-affected mRNAs that were rescued by Rrp6 depletion. The red arrow symbolizes the hijacking of Nrd1 and Nab3 from ncRNAs genomic loci to the Rho-affected mRNA genes upon perturbation of mRNP biogenesis by Rho (19).
partially reverted close to their normal levels by the deletion of the RRP47 gene, i.e. defective QC.

To have a global view on these sorting results, we used the RNA-seq datasets obtained under the different conditions to generate averaged meta-transcript profiles (Figure 2). For the whole group of down-regulated mRNAs (1,015 transcripts), the action of Rho in the wild-type cells leads to a general decrease in the intensity of the averaged meta-transcript, indicating that, on average, a fraction of corresponding mRNPs has been recognized as aberrant and eliminated. This conclusion is supported by the absence of any marked effect of Rho on the level of the averaged profile for mRNAs produced in the *rpr47*Δ strain. Sorting the group of down-regulated mRNAs into two categories (rescued and non-rescued by the QC alteration) showed two different trends regarding the Rho-induced variations as well as the shape of the averaged meta-transcript profiles. The pool of rescued mRNAs appears to contain relatively less expressed transcripts with a higher average reads coverage at the 5'end of the metagene which could be ascribed to the accumulation of short transcripts resulting from RNAP II pausing at the early part of the genes. Such paused transcription complexes are likely terminated by the action of Rho, which could explain the absence of their full recovery in the *rpr47*Δ strain.

**High resolution mapping of Rrp6-dependent QC components association with chromatin**

Our previous work on a model mRNA transcript (PMA1) revealed that the co-transcriptional targeting of Rho-induced aberrant mRNPs is mediated by a stimulation of Nrd1 recruitment to the gene locus with a concomitant enrichment by components of the Rtp6-dependent QC. To explore the general nature of the process, we analyzed the genome-wide association with chromatin in −Rho and +Rho conditions of Nrd1, Nab3, Rtp6 and Trf4. Yeast strains harboring C-terminally tagged proteins expressed from their chromosomal locations under their own promoters were used for chromatin immunoprecipitation followed by single-end deep sequencing (ChIP-seq). These ChIP-seq results indicate that the Rtp6-dependent QC components are involved physically in the recognition and targeting of aberrant mRNPs at the co-transcriptional level. Indeed, close visual examinations of individual genomic loci, as shown by representative examples in Figure 3A, highlight the Rho-mediated stimulation of QC recruitment to chromatin. The landscape of Nrd1 association with chromatin reveals broad occupancy profiles that cover the whole body of the affected genes. This underlines the known coupling of Nrd1 with transcription RNAP II following its recruitment by the CTD. Alternatively, the occupancy profiles for Nab3, Rtp6 and Trf4 display rather sharp peaks with significant cross-overlapping, suggesting a possible co-localization of the QC components over portions of the Rho-affected genes. Similar close visual examinations of ncRNAs genomic features (CUTs, SUTs and snRNAs) revealed Nrd1 and Nab3 enrichment signals in the absence of Rho supporting the known close association of the two proteins at the ncRNAs termination sites. However, the presence of Rho leads to a drastic decrease of both the number of Nrd1 and Nab3 recruitment peaks as well as their intensity over the ncRNAs genomic loci.

To monitor the interplay between the loss of QC proteins from ncRNAs loci and their enrichment within mRNA genes as a result of Rho action, we sought to visualize the protein distributions over the whole genome. This global view was obtained using the Omic Cursos R package that generates a circular layout of the 16 yeast chromosomes wreathed with the ChIP-seq peaks detected for the four proteins (Figure 3B). Comparisons between the plots obtained for the ncRNAs and Rho-affected mRNAs under −Rho and +Rho conditions show clearly a wide shift reminiscent of a migration of Nrd1 and Nab3 from ncRNAs genomic positions to mRNA genes in the presence of Rho. This suggests that Nrd1 and Nab3 are probably titrated-out from ncRNAs genomic features by their attachment to Rho-affected mRNA genes, generating extended ncRNAs. In contrast, it appears that the Rho-induced stimulation of Rrp6 recruitment to mRNA genes does not depopulate significantly its presence within ncRNAs loci. The plots show also the large recruitment of Trf4 to both ncRNAs and mRNA genomic features in the presence of Rho, substantiating the role of the TRAMP complex in the targeting of aberrant mRNPs as well as extended ncRNAs (26).

**Acknowledgements:** The authors worked supported by the Croatian Science Foundation (Grant UIP-2017-05-4411) to Igor Stuparević and by the CNRS France recurrent funding to A. Rachid Rahmouni. Partial support was also provided by Région Centre-Val de Loire through research grants (Anti-Rho-Biotics, and ARD2020 Biomédicaments) to A. Rachid Rahmouni. For the genome-wide study, Kévin Moreau was a recipient of a PhD fellowship from the Région Centre-Val de Loire. The genome-wide sequencing data were deposited at Array Express Database under the following accession numbers: RNA-seq: E-MTAB-7538, ChIP-seq: E-MTAB-7539.

**REFERENCES**

The targeting mechanism of nuclear RNAs


