Extensive 5′-surveillance guards against non-canonical NAD-caps of nuclear mRNAs in yeast

Yaqing Zhang, David Kuster, Tobias Schmidt, Daniel Kirrmaier, Gabriele Nübel, David Ibberson, Vladimir Benes, Hans Hombauer, Michael Knop & Andres Jäschke

The ubiquitous redox coenzyme nicotinamide adenine dinucleotide (NAD) acts as a non-canonical cap structure on prokaryotic and eukaryotic ribonucleic acids. Here we find that in budding yeast, NAD-RNAs are abundant (>1400 species), short (<170 nt), and mostly correspond to mRNA 5′-ends. The modification percentage of transcripts is low (<5%). NAD incorporation occurs mainly during transcription initiation by RNA polymerase II, which uses distinct promoters with a YAAG core motif for this purpose. Most NAD-RNAs are 3′-truncated. At least three decapping enzymes, Rai1, Dxo1, and Npy1, guard against NAD-RNA at different cellular locations, targeting overlapping transcript populations. NAD-mRNAs are not translatable in vitro. Our work indicates that in budding yeast, most of the NAD incorporation into RNA seems to be disadvantageous to the cell, which has evolved a diverse surveillance machinery to prematurely terminate, decap and reject NAD-RNAs.

1 Institute of Pharmacy and Molecular Biotechnology (IPMB), Heidelberg University, 69120 Heidelberg, Germany. 2 DNA Repair Mechanisms and Cancer, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany. 3 Zentrum für Molekulare Biologie der Universität Heidelberg (ZMBH), DKFZ-ZMBH Alliance, Heidelberg University, 69120 Heidelberg, Germany. 4 Cell Morphogenesis and Signal Transduction, German Cancer Research Center (DKFZ), DKFZ-ZMBH Alliance, 69120 Heidelberg, Germany. 5 Deep Sequencing Core Facility, CellNetworks, Heidelberg University, 69120 Heidelberg, Germany. 6 Genomics Core Facility, European Molecular Biology Laboratory (EMBL), 69117 Heidelberg, Germany. ✉ Email: jaeschke@uni-hd.de
Results

Short NAD-RNAs are abundant in yeast. To comprehensively address NAD-RNAs in budding yeast, we isolated total RNA from yeast strain BY4742 and applied the original NAD captureSeq protocol\(^1\), in which the enzyme adenosine diphosphate-riboseylcyclase (ADPRC) tags NAD-RNAs at the NAD moiety, followed by click-chemistry biotinylation and selective isolation by streptavidin binding. After adapter ligation, reverse transcription (RT), and PCR amplification, amplicons with sizes between 150 and 300 bp were selected by gel electrophoresis; thus, this library represented mostly RNA species with sizes between 20 and 170 nt present in the original sample. Enrichment was determined by quantitative comparison with a minus ADPRC negative control. In this unfragmented library, 1460 RNAs were found to be enriched, with changes reaching up to 1200-fold (Fig. 1a). Sixty-nine percent of the genome-mapped reads corresponded to mRNA 5′-ends (Fig. 1b), whereas only little enrichment was observed for mRNA fragments starting further downstream (for details and validation, see Supplementary Methods and Supplementary Fig. 1a–d). Small nucleolar RNAs (snoRNAs) and ribosomal RNA (rRNA) fragments comprised 9.6% and 7.1% of the reads, but represented only 2 and 1 different RNA species, respectively. Thirteen percent corresponded to RNA fragments too small for unique genome mapping (12–17 nt, Fig. 1b), many of which showed homology to enriched members of the mRNA 5′-end group (Supplementary Fig. 1e). To probe the existence of full-length NAD-capped mRNAs, two additional datasets were generated from total RNA that was random-sheared prior to NAD captureSeq using different size selection windows (small fragmented: 20–170 nt; large fragmented: 170–350 nt). Consistently, these fragmented libraries revealed much lower numbers of enriched species and enrichment values (small fragmented: 145 RNAs, maximum fold change (FC) < 7 (Fig. 1c) and large fragmented: 200 RNAs, maximum FC < 9 (Supplementary Fig. 1d)), mainly due to increased background in the minus ADPRC controls. Extensive overlap was detected between the two fragmented libraries and the unfragmented one (83.8% and 76.0%, respectively; Fig. 1e). Five out of the 12 genes explicitly reported in the previous study\(^1\) (Supplementary Fig. 1h) overlap with the enriched species in our unfragmented library (COX2, LSM6, ERG2, UBC7, and YJR112W-A) and two with the fragmented library (LSM6 and UBC7; black dots in Fig. 1a, c).

To test whether highly expressed transcripts are generally more likely to be enriched in NAD captureSeq, we compared the enrichment levels observed in the unfragmented NAD captureSeq library with the transcript abundance determined by transcriptome sequencing. This analysis revealed no correlation (Supplementary Fig. 1i).

As NAD captureSeq enriches the 5′-ends of NAD-RNAs and sequences in 5′–3′ direction, the overall lengths of transcripts larger than the Illumina read length cannot be reliably inferred from the sequencing reads (Fig. 1d). We therefore carried out the first steps of the NAD captureSeq protocol (until the enriched RNAs were bound to streptavidin) and carried out RT quantitative PCR (RT-qPCR) on two RNA species that were enriched in the fragmented and unfragmented libraries (TDH3 and SED1) using four different primer pairs each. These data revealed roughly equal abundance from the 5′-end through ~300 nt, whereas their 3′-untranslated regions (UTRs) were reduced in abundance by several orders of magnitude (Fig. 2a, green bars). The preferential enrichment of smaller RNAs in NAD captureSeq was also confirmed by Bioanalyzer size analysis of the DNA amplicons (Supplementary Fig. 1j).

For two transcripts (TDH3 and POR1 mRNA), the 5′-NAD modification was directly identified and quantified by mass spectrometry (MS) after pulldown (Fig. 2b and Supplementary Fig. 1i), confirming the chemical identity of the NAD modification. Thus, NAD-RNAs are abundant, short, and mostly correspond to mRNA 5′-ends in budding yeast.

Nudix pyrophosphohydrolase Npy1 processes NAD-RNA. In *Escherichia coli*, the Nudix hydrolase NudC acts as an efficient decapping enzyme for NAD-RNA\(^4,12,13\). The yeast homolog Npy1 is known to hydrolyze the pyrophosphate bond in NAD to yield nicotinamide mononucleotide (NMN) and adenosine monophosphate (AMP)\(^14\), and was recently suggested as an NAD-RNA decapping enzyme\(^13\). The only support for this claim was, however, its in vitro processing of a synthetic NAD-RNA 12mer into a product that migrated on HPLC like a 12mer-5′-monophosphate RNA (p-RNA) and the inactivity of an active-site mutant to produce this product\(^13\). To characterize the in vitro activity of Npy1, we purified the protein from *E. coli* and analyzed its reaction kinetics with an in vitro-transcribed NAD-RNA (a 98 nt 5′-fragment of TDH3 RNA) on acryloylaminophenyl boronic acid (APB) gels, which separate NAD-RNA from p-RNA\(^15\). Purified Npy1 decapped NAD-RNA without inducing nucleolytic degradation and had no effect on m3G-RNA in vitro (Fig. 3a). Furthermore, efficient decapping of NAD-RNA required Mn2+ ions (Supplementary Fig. 2a–c). A Npy1 mutant...
in which a catalytic glutamate was replaced (E276Q) showed no decapping activity (Supplementary Fig. 2d). In addition to NAD-RNA, Npy1 also hydrolyzed NAD into NMN and AMP in a Mn$^{2+}$-dependent manner (Supplementary Fig. 2e), whereas the E276Q mutant was inactive (Supplementary Fig. 2f). Thus, recombinant Npy1 decaps NAD-RNA in vitro.

To address whether Npy1 also functions on NAD-RNA in vivo, we investigated a yeast strain lacking Npy1. In agreement with the yeast SGA database, the absence of Npy1 caused no severe phenotypical changes under a variety of growth conditions (Supplementary Fig. 2g). Although gene expression analysis by transcriptome sequencing indicated changes in abundance for almost 50% of all detected transcripts (Supplementary Fig. 2h), mass spectrometric whole proteome analysis detected only very few proteins with significant (more than twofold) changes in expression, in comparison to the WT strain (Fig. 3b). Deletion of NPY1 slightly increased the total cellular concentration of NAD (by ~10%, Fig. 3c). When we applied NAD captureSeq to RNA purified from the npy1Δ strain, twice as many uniquely mapped RNAs (3028, unfragmented library) were NAD-capped (relative to WT), which were almost half of all detected RNA species (Fig. 3d). Consistent with the WT, NAD-RNAs from the npy1Δ strain were mostly short transcripts, as only 242 and 226 NAD-RNA species were enriched in the small and large fragmented RNA libraries, respectively (Supplementary Fig. 2i, j). Compared to the WT, a similar proportion of the reads allocated to mRNA 5′-ends in the unfragmented library (59.1%) but three times more on very small RNAs (40.3%), suggesting that Npy1 is involved in the decapping of small NAD-RNAs. rRNAs and snoRNAs disappeared almost completely (Fig. 3e).

Contrary to our expectations (but as also observed in Bacillus subtilis), removal of Npy1 reduced the total amount of NAD attached to RNA by ~60% (Fig. 4a and Supplementary Fig. 2k). We assessed the change in the apparent modification ratio (percentage of an RNA species that carries NAD) transcriptome-wide by integrating transcriptome and NAD captureSeq data, using the enrichment values in NAD captureSeq as proxy (Fig. 4b). This analysis indicated that upon NPY1 gene deletion, the NAD-modification ratio was reduced for 1013 species, whereas it increased for 164. There was no correlation between expression level (change) and modification ratio (change).
To compare the in decap NAD-RNA in vitro and in vivo by a mechanism different non-Nudix enzymes Rai1 and Dxo1 were previously reported to many weakly increased ones. The slope <1 (0.61) of this plot most cases between 1% and 3% were determined (Fig. 4c). Those background subtraction, NAD-modification values to the same amount of input RNA in WT and (N) for WT and mutant strain. After normalization of the cp preparation from cells containing this enzyme, we analyzed their hydrolyzed good substrates for Npy1, being severely depleted in RNA the latter two transcripts might indicate that they are particularly 16 genes with decreased NAD modification (LSM7 and RSM10) PCR-ampli fi cation (LSM7 and RSM10) PCR-amplified stronger (Fig. 4c). To test whether this increased modification ratio of the latter two transcripts might indicate that they are particularly good substrates for Npy1, being severely depleted in RNA preparation from cells containing this enzyme, we analyzed their decapping in vitro. Indeed, LSM7 and RSM10 were hydrolyzed much faster than TDH3 (Fig. 4d, e). Collectively, these data are consistent with a role of Npy1 in processing NAD caps in vivo.

Npy1, Rai1, and Dxo1 influence the NAD-RNA landscape. The non-Nudix enzymes Rai1 and Dxo1 were previously reported to decap NAD-RNA in vitro and in vivo by a mechanism different from Npy1, namely by removal of the entire NAD moiety en bloc24,47. To compare the influence of all three enzymes on the global NAD-modification landscape of RNAs in vivo, we created all possible combinations of rai1Δ, dxo1Δ, and npy1Δ deletion mutants. Phenotypically, the removal of Rai1 (from the WT and from mutant strains) had the strongest negative effect on growth in normal medium and in the presence of increasing concentrations of ethanol (Fig. 5a and Supplementary Fig. 3a). On the transcriptome level, we detected in all mutant strains ~1000 upregulated and ~1000 downregulated RNA species (at least fourfold, relative to WT), together corresponding to ~30% of all mRNAs (Supplementary Fig. 3b). There was over 60% overlap in regulated genes between the three different single-knockout mutants, whereas ~600 genes were selectively regulated by only one decapping enzyme (Supplementary Fig. 3c). A systematic analysis of the effect of the deletion of one particular enzyme in WT and mutant strains revealed high agreement within one group (e.g., all strains carrying a deletion of the NPY1 gene), and the strongest global effect on RNA expression was noticed for removal of RAII (Fig. 5b). Analysis of total RNA isolated from the knockout mutants by NAD captureSeq revealed enrichment of more than half of all detected RNA species (3765 in dxo1Δ; 3810 in rai1Δ), indicative of their modification with NAD. No significant further increase was observed in the double- and triple- deletion strains (Fig. 5c). In the triple-knockout dxo1Δ rai1Δ npy1Δ, only mRNA fragments (63%) and small RNAs (35%) were detected by NAD captureSeq (Fig. 5d). Unlike the WT, the top 250 enriched NAD-RNA species of all mutants functionally clustered (by Gene Ontology terms) as rRNA metabolic process and translation (Supplementary Fig. 3e). Thus, our analysis of the rai1Δ, dxo1Δ-, and npy1Δ-knockout strains may be consistent with a role of the affected gene products in processing NAD capping, but by no means demonstrative evidence.

NAD-RNAs have distinct TSSs. The above analysis suggested that the landscape of NAD-RNA transcripts is shaped by (at least) four enzymes as follows: RNAP II, Rai1, Dxo1, and Npy1. Using the deletion mutants, we first analyzed transcriptional preferences. Although the sequencing read profiles of some RNAs revealed homogenous 5′-ends (indicative of a defined TSS), others showed irregular patterns suggesting pervasive transcription or multiple TSSs (see Fig. 1d and Supplementary Fig. 4a for examples). From the NAD captureSeq data, we selected all significantly enriched RNAs starting with an A which had homogenous 5′-ends (“sharp A” selection). When we compared our experimentally determined 5′-ends of these NAD-RNAs with published next-generation sequencing (NGS)-derived and 5′-rapid amplification of cDNA end-validated TSSs for canonical (i.e., non-NAD-) RNAs18,20, for nearly half of all species the 5′-ends differed (Fig. 6a, b). For the WT strain, 98 RNAs were observed in which the 5′-transcript leader (TL) sequences were longer than in the database, whereas 63 species got shorter, in some cases by more than 100 nt (Fig. 6c). This TL length change was not only observed in the WT strain (in both unfragmented and fragmented libraries, Fig. 6c and Supplementary Fig. 4b, c), but also in all mutants, including the dxo1 Δrai1Δ npy1Δ triple mutant (Fig. 6b and Supplementary Fig. 4d), suggesting that RNAP II might select a different TSS for initiating transcription with NAD instead of ATP, compared to the canonical TSS19,20. The changed TL length upon NAD incorporation could be corroborated by qRT-PCR with primers targeting either our NAD captureSeq-observed TSS or the canonical ones from the database, comparing the RNAs enriched in NAD captureSeq with a non-enriched total RNA preparation (Fig. 6d). Thus, NAD-RNAs tend to have longer TL sequences than non-NAD-RNAs, indicative of their synthesis starting at a more distal TSS.

A YAAG promoter motif supports NAD incorporation by RNAP II. We supposed that analysis of the dxo1Δ rai1Δ npy1Δ
analyzed in the presence of Mn²⁺. An α³²P-body-labeled 5’-fragment of TDH3 RNA (98 nt) and a corresponding NAD-RNA control were treated with NudC in the presence of 2 mM Mg²⁺ and 1 mM Mn²⁺, and resolved by APB gel electrophoresis. Three independent experiments were performed, n = 3. b Whole proteome analysis of cell lysates (WT strain vs Δnpy1 strain). Performed in biologically independent replicates, p value = 0.011 (Student’s t-test, one-sided). d Enriched NAD-RNAs in the unfragmented NAD captureSeq library of the Npy1 depletion (npy1Δ) strain. Performed in biologically independent replicates, n = 3. e Reads of enriched NAD-RNA in the npy1Δ unfragmented NAD captureSeq library, mapped to genomically annotated RNAs. All parameters as in Fig. 1b. Source data are provided as a Source Data file.

**Fig. 3 Npy1 affects NAD-RNA in vitro and in vivo.** a S. cerevisiae Npy1 in vitro enzyme kinetics of decapping of 5’-NAD- and 5’-m⁷G modified RNAs was analyzed in the presence of Mn²⁺. A α³²P-body-labeled 5’-fragment of TDH3 RNA (98 nt) and a corresponding NAD-RNA control were treated with NudC in the presence of 2 mM Mg²⁺ and 1 mM Mn²⁺, and resolved by APB gel electrophoresis. b Whole proteome analysis of cell lysates (WT strain vs Δnpy1Δ). Red dots indicate enriched hits (fold change >1.5 and FDR < 0.05), whereas green dots are hit candidates (fold change >1.4, 0.05 ≤ FDR < 0.2) and gray dots are proteins without significant change. c Quantification of cellular NAD content in WT (blue bar) and npy1Δ (orange bar) yeast strains, determined using an enzyme cycling assay. The amount of NAD was normalized to the amount of total RNA determined in cell lysates, obtained from each sample. Error bars represent the mean ± SD. Performed in biologically independent replicates, n = 4. p-values are denoted by asterisks: *p < 0.05 (Student’s t-test, one-sided).

A triple-knockout mutant strain would reveal the least biased information about the factors that govern transcriptional NAD incorporation by RNAP II. We mapped nucleotides –10 to +10, relative to the RNA 5’-end inferred from the NAD captureSeq reads, for the 25 most enriched “sharp A” NAD-RNAs (log₂FC > 8, false discovery rate (FDR) < 0.00002) and for appropriate control groups (i.e., RNA species not enriched in NAD captureSeq). In the enriched fraction, we observed a highly conserved motif YAA (with the first A being the 5’-terminal nucleotide of the transcript, i.e., the site where the NAD is incorporated), followed by an A-rich stretch of lower significance, whereas in the non-enriched fraction no preferences were found (Fig. 7a). The motif was not observed when for the same top 25 candidate RNAs the published canonical TSSs¹⁸ were mapped (Fig. 7b). When for those 25 genes all TSSs listed in the yeast TSS database²¹ (top 5 abundant TSSs per gene) were analyzed, only a YA motif²² was identified (Supplementary Fig. 5a). However, when only the TSS (from this database) closest to our observed one was utilized, the YAAG motif appeared prominently (Supplementary Fig. 5b). This analysis implies that our identification procedure revealed real TSSs and further supports the assumption that transcriptional NAD incorporation is the predominant biosynthetic pathway to furnish NAD-RNAs. The YAAG motif was also observed (although less prominently) in the top 100 and 200 enriched RNAs, and its prominence decreased with decreasing NAD captureSeq enrichment values (Supplementary Fig. 5c, d). It could also be detected in WT and all mutant strains, whereby generally the significance decreased with increasing number of decapping enzymes present (Supplementary Fig. 5e–k). The motif was not observed when the NAD captureSeq-enriched snoRNAs or transfer RNAs (tRNAs) were mapped (Supplementary Fig. 5l, m), suggesting that these candidates may have a different biogenesis. To exclude the possibility that the motif reflects a bias introduced by the enzymes applied in NAD captureSeq (ADPRC, reverse transcriptase, terminal deoxynucleotidyl transferase, two ligases), we mapped the top 25 enriched sequences from our previously published E. coli, B. subtilis, and Staphylococcus aureus NAD captureSeq datasets by the same procedure, finding neither YAAG nor an A-rich tail (Supplementary Fig. 5n–p). Further analysis revealed that this motif constitutes a fraction of known “good” RNAP II core promoter sequences, having all conserved features²³–²⁵, namely: (1) being A/T-rich between positions –30 and +10, (2) a switch from T-rich to A-rich in the coding strand around position –8, (3) a pyrimidine at position –1, and (4) an A at position +1. In addition, two specific features distinguish good NAD-incorporating promoters, namely a slightly increased probability for an A at position +2 and a strongly conserved G at +3.

To test whether this motif actually modulates NAD incorporation by RNAP II in vivo, we deleted gene TDH3, a highly enriched...
NAD-RNA observed in every strain, and added a low-copy plasmid in which we inserted a DNA fragment containing the 600 bp upstream of the TDH3 gene, containing the entire promoter region, plus the first 54 bp after the experimentally observed TSS of the TDH3 RNA (39 nt 5′-UTR, 15 nt coding sequence), followed by the ORF of superfold-GFP to monitor gene expression (Supplementary Fig. 5g). Mutants were prepared in which the Y at position −1, the A at position +2 and the G at position +3 were individually varied. An additional mutant was generated in which all A’s in the tail region (+4, +5 and +9) were replaced. Cells were transfected with these plasmids, and harvested around OD₆₀₀ = 0.8. Total RNA was isolated, treated with ADPRC, followed by click biotinylation, streptavidin purification, and RT. qPCR with gene-specific primers was used to assess the percentage of NAD-modified TDH3 RNA in each strain, using pure synthetic spike-in NAD-RNA and ppp-RNA, to ensure equal reactivity of each sample. This analysis revealed indeed strong (~2-fold) reduction of relative NAD incorporation upon mutating positions −1 and +3, whereas for positions +2 and the A-rich tail the observed effects were not statistically significant (Fig. 7c and Supplementary Fig. 5r). Quantification of green fluorescent protein (GFP) expression levels revealed that mutating position −1 significantly decreases both NAD-RNA and non-NAD-RNA, whereas mutating position +3 modulates exclusively NAD-RNA (Fig. 7d and Supplementary Fig. 5s). Thus, a specific promoter sequence and particularly a G at position +3 are responsible for efficient NAD incorporation in vivo.

Most NAD-RNAs are 3′-truncated. The observation that most yeast RNAs enriched in NAD captureSeq are much shorter than full-length mRNAs and their preferential mapping to mRNA 5′-ends lead us to ask whether there is an influence of NAD incorporation on the transcript length. Globally, we determined the percentage of mRNA-mapped full-length reads for WT and
Fig. 5 Enzyme deletions influence the NAD-RNA landscape. a Growth phenotype of the WT and mutant yeast strains in the presence of different ethanol concentrations. The rank of approximate growth density is given on the right side. Each condition was tested in triplicates. Strain names are denoted on the left.

b Effect of the deletion of one particular enzyme in different strains on the upregulation (red arrows) or downregulation (blue arrows) of RNA species, assessed by transcriptome sequencing. Group 1 summarizes the deletion of the NPY1 gene from the single mutants (dxo1Δ, rai1Δ) and from the double mutant (dxo1Δ rai1Δ). Groups 2 and 3 show deletion of the DXO1 gene and of the RAI1 gene, respectively. The number of RNA species with log2 FC > 2 is shown on the left panel, whereas with log2 FC > 0.5 is shown on the right panel. e Enriched NAD-RNAs in the unfragmented NAD captureSeq library of the dxo1Δ rai1Δ npy1Δ strain. All parameters are as in Fig. 1a. Performed in biologically independent replicates, n = 3. d Radar plot of the distribution of different classes of enriched NAD-RNAs (in %) in different deletion strains. The colors indicate the type of RNAs, enumerated on the left. Source data are provided as a Source Data file.

all mutants and compared this value for sample (+ADPRC, S) and negative control (−ADPRC, N). After normalization to non-enriched species, in all eight libraries the sample group contained less full-length reads than the negative control and more abortive fragments (Supplementary Fig. 6a). At the individual transcript level, we determined for the highly expressed (and enriched) TDH3 RNA the transcript start and end nucleotide analyzing each read individually26. According to this analysis, both S and N distribution of truncated 3′-ends was observed between S group and N group. The proportion of full-length Illumina reads with identical TSS differed by a factor of 2.7 (33.4% in S and 88.8% in N, Fig. 8a–b, histograms to the right of the two-dimensional plot). Although the reasons for this increased proportion of 3′-truncated NAD-RNAs remain unclear, these findings may suggest that unidentified quality control (QC) mechanisms detect NAD incorporation into RNA as an error quite early and interfere with efficient transcript elongation.

Npy1, Dxo1, and Rai target different NAD-RNA populations. The observation that the promoter motif got increasingly “blurry” with increasing number of decapping enzymes present (Fig. 7a and Supplementary Fig. 5e–k) supported our assumption that the NAD captureSeq data actually reflect a superposition of RNAP II and decapping enzyme preferences. The comparison of the datasets of the three single mutants revealed extensive overlap, and 1544 species (>60%) were enriched in all three mutants, compared to the WT (Supplementary Fig. 6b). Similar findings were observed comparing the three double mutants. Computational sequence and secondary structure analysis of RNAs of uniquely or commonly enriched RNAs did not reveal specific features indicative of substrate preferences of these enzymes. However, for Rai1 we observed a slightly decreased minimum free energy of folding27 for preferred RNA substrates, compared to poor ones (Supplementary Fig. 6c). This finding may suggest that Rai1 tends to have a preference for less structured 5′-ends.

We noticed that the removal of decapping enzymes not only influenced the number of RNA species enriched in NAD captureSeq and their enrichment values, but also the (apparent) length of their 5′-ends (TL). This phenomenon was observed for ~20% of all RNA species and occurred in both directions, namely (apparent) TL lengthening and shortening upon knockout. For example, among the 1100 enriched sequences in common between the WT and npy1Δ strain, 152 apparently got shorter and 75 got longer TLs (Fig. 8c). For all other mutants, similar observations were made. For several candidate RNAs, these length differences could be confirmed by qRT-PCR with the cDNA from the NAD captureSeq samples (Fig. 8d). This
phenomenon was observed almost exclusively for RNAs with read patterns indicative of pervasive transcription or multiple TSSs, and not for those with homogenous TLs. We assumed that the most likely explanation for these results may be that the decapping enzyme, when presented with a transcript mixture with different TLs, decaps some more rapidly than the others, due to sequence or structural preferences, thereby causing changes in the NAD captureSeq read profiles that look like shifted TSSs. A direct modulation of transcription (e.g., as transcription factors) is difficult to reconcile with the currently assumed roles and locations of these proteins, at least for Dxo1 and Npy1.

Rai1 has been reported as a nuclear protein and was detected as a component of the RNAP II elongation complex, whereas for Dxo1 both nuclear and cytosolic locations were claimed. Npy1 was described as a peroxysomal protein. Localization microscopy using three different C-SWAT fluorescent protein fusions for each candidate gene revealed strong localized fluorescence in the nucleus for Rai1, whereas Dxo1 showed only a very weak and ubiquitous fluorescence (Fig. 9a), consistent with the reported localizations of these enzymes. For Npy1, however, a rather homogenous cellular distribution without enrichment at specific sites was observed, consistent with cytosolic localization (Fig. 9a). This localization may imply a temporal order, in which Rai1 processes its NAD-RNA substrates during or shortly after transcription, whereas Npy1 can only act once the transcripts arrive in the cytosol. For Dxo1, both options are conceivable. Therefore, we tried to find evidence in our NAD captureSeq data for a temporal order of the pathways—whether the cap is translated in the nucleus for Rai1, whereas Dxo1 showed only a very weak and ubiquitous fluorescence (Fig. 9a). Consistently, an example of a temporal order of the pathways was observed almost exclusively for RNAs with read patterns indicative of pervasive transcription or multiple TSSs, and not for those with homogenous TLs. We assumed that the most likely explanation for these results may be that the decapping enzyme, when presented with a transcript mixture with different TLs, decaps some more rapidly than the others, due to sequence or structural preferences, thereby causing changes in the NAD captureSeq read profiles that look like shifted TSSs. A direct modulation of transcription (e.g., as transcription factors) is difficult to reconcile with the currently assumed roles and locations of these proteins, at least for Dxo1 and Npy1.

Indeed, from the 84 species with TL length changes between triple knockout and dxo1Δ npy1Δ double knockout, 61 disappeared in the npy1Δ single knockout and 38 in the dxo1Δ single knockout. Importantly, hardly any examples were found for the pathways via the other double mutants (one example for npy1Δ rai1Δ and 0 for dxo1Δ rai1Δ) (Fig. 9b). These findings are consistent with our assumption that Rai1 is the first factor in NAD-RNA decapping.

**NAD-RNAs are not translatable in vitro.** Finally, we tested whether the NAD cap in combination with different TL lengths and sequences may modulate translation. Reports on NAD-RNA translatability are conflicting: Jiao et al. had reported that NAD-RNA is not translated in human (HEK293T) cell extracts, based on a single mRNA luciferase construct with a single fixed TL sequence, while a recent study in the model plant *Arabidopsis thaliana* demonstrated that NAD-capped mRNAs are enriched in the polysomal fraction, associate with translating ribosomes, and can probably be translated. No data for yeast have been reported yet. For seven different mRNAs, we prepared luciferase fusion constructs with long and short TLs by in vitro transcription, followed by the removal of the accompanying ppp-RNA by treatment with polyphosphatase and exonuclease Xrn-1. Although the control constructs harboring an m7G-capped 5'-end were efficiently translated in a yeast in vitro extract and showed significant differences in luminescence depending on the TL length, NAD-capped RNA was not translated to any significant extent, even less than ppp-RNA and p-RNA of the same sequence (Fig. 10a). These results suggest that NAD-capped RNAs (at least the nuclear transcripts investigated here) are not translated in budding yeast.

**Discussion**

Taken together, our results indicate that in budding yeast, NADylation of RNAs is a very common phenomenon. A previous study reported only 37 species enriched in NAD captureSeq in budding yeast grown in the same medium. This study, however, focused on full-length mRNAs and used a library preparation

---

**Fig. 6 NAD-RNAs have longer 5'-UTRs than non-NAD-RNAs.** a Scheme of TSS shifting in proximal or distal direction. b Global TSS shifting between NAD-RNA (according to NAD captureSeq, unfragmented libraries) in all strains and canonical RNA (according to the dataset of Nagalakshmi et al.). The boxplot shows from bottom to top minimum (Q1-1.5 interquartile range (IQR, 25–75%)), first quartile (Q1, 25%), median (solid line, 50%), third quartile (Q3, 75%), maximum (Q3 + 1.5 IQR), and outliers (black dots). The numbers above each box group indicates the number of RNA species analyzed in the corresponding strain. c Detailed TSS shifting between NAD-RNA and All-cap RNA in WT strain, 5'-UTR length difference denotes NAD-RNA 5'-UTR length minus canonical (all-cap) RNA 5'-UTR length. Red dots represent NAD-RNA species with distal TSS (> 10 nt difference) and blue dots represent NAD-RNA species with proximal TSS (> 10 nt difference). FDR<sub>holm</sub> < 0.1. The sample size (n) is 347. The black line represents the linear correlation between 5'-UTR length and NAD-RNA enrichment level. d qRT-PCR with different primer pairs confirming different transcript leader (TL) lengths between NAD-capped RNAs (enriched fraction from NAD captureSeq library, blue boxes) and all-capped RNAs (from non-enriched input RNA, gray boxes). The boxplot displays the log<sub>2</sub> of the ratio of the number of transcripts derived from the distal TSS to that from the proximal TSS. The layout of boxplot is as defined in b. Dots represent individual biological triplicate measurements. Source data are provided as a Source Data file.
Fig. 7 NAD incorporation by RNAP II is promoter dependent. a Motif analysis of the −10 to +10 region around the TSS based on NAD captureSeq data. Top25 and background25 represent the 25 NAD-mRNA species with the highest enrichment values and the 25 most abundant, but not significantly enriched mRNA species (0.707 ≤ S/N ≤ 1.414) with a sharp A feature at position +1 (TSS) in the dxo1Δ rai1Δ npy1Δ strain. S/N is the enrichment of NAD-RNA in the sample group (+ADPRC) compared to the negative control group (−ADPRC). Sharp A means an “A” at position +1 with a more than three-fold higher signal than that of the nucleotide in the −1 position. The letters, representing the four nucleobases, were colored dark orange for A/T and blue for G/C. b Motif analysis of the same Top25 enriched RNA species, using the canonical TSS from Nagalakshmi dataset18. c Quantification of the NAD-modification ratio using the TDH3 gene promoter and the relevant mutant promoter sequences in vivo. The height of the green bar indicates the TDH3 RNA NAD-modification ratio in the sample group (+ADPRC), whereas the gray bar denotes the same in the N group (−ADPRC). The Pori represents the original TDH3 promoter and promoter mutants are enumerated below the bar chart. Dots represent biological triplicates and error bars represent SDs. p-values are denoted by asterisks: *p < 0.05 (Student’s t-test, one sided). d Expression of sfGFP under a tdh3 promoter dependence, for many RNAs the NADylated species

protocol that discarded the small RNA fraction (≤ 200 nt). Our work confirms that there are hardly any full-length NAD-mRNAs, but additionally reveals a rich landscape of thousands of short NAD-mRNA fragments whose purpose is apparently not to

targeting sequence and are absent in yeast mitochondrial proteomic data33,34. NAD is initially incorporated into several thousands of transcripts by transcription initiation by RNAP II in a largely statistical manner reflecting the competition of NAD with the canonical initiator ATP. As in prokaryotes5,35,36, the promoter sequence determines the efficiency of NAD incorporation, which is for most yeast nuclear transcripts between 1% and 5%. We observe that a YAAG motif supports efficient NAD incorporation by RNAP II in vivo, with the G at position 3 being particularly important. This finding does not rule out the existence of unknown alternative post-transcriptional pathways for NAD incorporation, e.g., for enriched snoRNAs or rRNAs, or for enriched mRNAs without the YAAG motif. In contrast to the promoter requirements determined here, a preference for HRRASWW was reported for E. coli RNAP36, WARR for B. subtilis RNAP3, and RA for yeast mtRNAP32, with the underlined A always indicating the TSS. It should, however, be noted that the E. coli RNAP and yeast mtRNAP consensus motifs were established using an entirely different methodology, making a direct comparison difficult. We find that, as a consequence of this promoter dependence, for many RNAs the NADylated species originate from different TSSs and have therefore different (shorter or longer) 5‘-UTRs than the canonical ones. This phenomenon may modulate the secondary structure of these RNAs

In addition, the identity of a dedicated surveillance machinery within the mitochondrial matrix is of considerable interest. The three decapping enzymes investigated here, however, are unlikely to encounter mitochondrial transcripts, as they lack the required
and hence their stability, molecular interactions, and biological fate. Of note, the discovery of alternative TSS selection and the YAAG core promoter motif have been made possible by the combination of 5′-end selection by ADPRC treatment and ligation-based attachment of the 5′-adapter, which allowed the determination of NAD-RNA 5′-ends with single-nucleotide precision, in contrast to random-primed library preparation methods that create heterogeneous ends. An increased affinity of RNAP II for the YAAG motif, while the polymerase transiently harbors NAD in its catalytic site, could potentially also explain the enrichment of non-canonical TSSs upon execution of NAD captureSeq. An in-depth biochemical investigation should explore....

Fig. 8 The 5′- and 3′-end heterogeneity of NAD-RNA. a, b 2D NAD-capping single transcript plot for TDH3 RNA in the dxo1Δ rai1Δ npy1Δ strain for the sample group (+ADPRC, S, (a)) and the N group (−ADPRC, N, (b)). O denotes the RefSeq TSS (−39 nt of the TTS site). Each bin represents a unique 5′ (initiation site, TSS, x axis) and 3′ (transcript end site, TES, y axis) pairing colored by the number of reads mapped to that bin. Expanded view below. For the 5′-end initiation site histogram (above the 2D plot), the red bar indicates the “A” position with YAAG feature, whereas orange bars indicate “A’s” without this feature and gray bars indicate U/C/G. For the 3′-end histogram, no color differentiation was performed. Arrows represent preferred TSSs in the 5′-dimension and TESs in the 3′-dimension. c Global effect of the deletion of the NPY1 gene on the transcript leader length by comparing NAD captureSeq read starts in the npy1Δ mutant with the WT (unfragmented libraries). A Transcript Leader length Index (TLI) > 0 indicates that upon npy1Δ mutation the TSS towards more distal positions, while a TLI < 0 indicates a proximal shift; RNAs with significantly (FDR < 0.05) shifted TSS are shown as colored dots. Red: distal shift (TLI > 0.1) Blue: proximal shift (TLI < −0.1). The y axis indicates the relative NAD-RNA enrichment difference (ratio) from NAD captureSeq libraries between the two strains. The bar charts represent the total number of RNA species with distal (red) or proximal (blue) TSS shift. d qRT-PCR with different primer pairs confirming different NAD-RNA transcript leader (TL) lengths in the enriched fractions from NAD captureSeq libraries between WT (blue boxes) and npy1Δ strain (orange boxes). All parameters and the layout of boxplot are as in Fig. 6d. In the table, symbol “−” and “+” indicate a decrease or increase of the NAD-modification ratio, respectively. “NS” indicates no significant difference. Source data are provided as a Source Data file.
the possibility of NCIN-mediated guidance of RNAPII, and other RNAPs, to distinct TSSs. NAD-RNAs are—on average—shorter than non-NAD-RNAs and only rarely reach the size of a typical primary mRNA transcript. The most likely explanation is that some unidentified QC mechanism detects 5′-NADylation of RNA as an error early during transcription and prevents efficient elongation, as it does with uncapped or incompletely capped transcripts. Alternatively, NAD-RNAs might be subject to accelerated degradation after transcription is complete, but it is unclear how 5′-NAD can accelerate degradation at the 3′-end.

The discovery that budding yeast maintains at least three different, partly redundant, pathways for NAD cap removal, using enzymes with different chemistry and cellular localization, implies that decapping unwanted NAD-RNAs is important for the cell. Our data are in agreement with the hypothesis that Rai1 acts earlier than the other two enzymes. As the nuclear protein Rai1 is known to associate with RNAPII during elongation and to act in RNA surveillance by assisting the 5′- to 3′- exonuclease Rat1 in the co-transcriptional degradation of uncapped transcripts, such an order appears plausible. The Rat1-Rai1 complex is hereby believed to play an important role by mediating 5′-end cap QC (5′-QC) in the yeast RNAPII transcription cycle, following the transcription checkpoint pause stage, whereby RNAPII enters transcription elongation upon phosphorylation of distinct serines within the C-terminal repeat domain of the polymerase. Surveillance and hydrolysis of the accidentally incorporated 5′-NAD cap could be enacted in a mechanistically similar manner, mirroring the clearance of unmethylated, aberrantly capped mRNAs by the Rat1-Rai1 heterodimer.

An earlier report, providing evidence that dinucleotide hydrolysis mediated by *Saccharomyces cerevisiae* Npy1 is not entirely restricted to NAD, but also includes the redox cofactor flavin adenine dinucleotide among others, warrants a thorough biochemical characterization to define the set of RNA 5′-metabolite caps, targeted by this enzyme. A corresponding study should hereby follow the example set by the systematic and meticulous elucidation of RNA 5′-cap specificities of mammalian DXO and *Schizosaccharomyces pombe* Rai1.

The observed combination of the low efficiency of RNAPII transcription initiation by NAD, the reduced length of NAD-RNAs, and the abundance of NAD-RNA decapping enzymes warrants that hardly any NAD-RNAs occur in the cell that could give rise to translation into proteins. Our data indicate, however, that yeast ribosomes, such as mammalian ones, hardly translate synthetic NAD-mRNAs, suggesting that the ribosomal machinery contains additional safeguards against NAD-mRNAs. Thus, budding yeast protects itself at different stages of gene expression against NAD-RNA.

**Methods**

**Yeast strains.** Unless otherwise stated, *S. cerevisiae* strains were grown in yeast extract/peptone/dextrose media (YPD). All strains used in this study except of YDK587-1 and its derivatives, YDK53-7, and C-SWAT mNeonGreen (mNG-I) strains were derivatives of the S288C strain BY4741 (MATα his3Δ1 leu2Δ0 0 lys2Δ0 0 ura3Δ0) and are listed in the key resource table. YDK53-7 and C-SWAT mNeonGreen (mNG-I) strains were derivatives of the S288C strain BY4741
**Fig. 10 In vitro translation of NAD-RNA with shifted TLs.** In vitro translation of NAD- RNA, p-RNA, ppp-RNA, and m7G-RNA with short or long TL sequences. Probed mRNAs, bearing a 5′-NAD-, 5′-p-, or 5′-ppp-terminus, contain the alternative TL sequences identified in NAD captureSeq, followed by 22 nt of the CDS of the corresponding gene, followed by the firefly luciferase CDS (1653 nt) and a poly(A)30 tail. Firefly luciferase activity was normalized to Renilla luciferase activity, then normalized to the C0-value of the full-length mRNA determined by qRT-PCR. NAD-capped mRNAs, p-mRNAs, and ppp-mRNAs contain the same sequence as the corresponding m7G-capped RNAs. Logarithmic representation of normalized luciferase activity. Blue, NAD-RNA; purple, p-RNA; orange, ppp-RNA; red, m7G-RNA. The percent the full-length mRNA determined by qRT-PCR. NAD-capped mRNAs, p-mRNAs, and ppp-mRNAs contain the same sequence as the corresponding m7G-capped RNAs. Logarithmic representation of normalized luciferase activity. Blue, NAD-RNA; purple, p-RNA; orange, ppp-RNA; red, m7G-RNA. The percent values below the bars indicate relative luciferase activity normalized to the m7G-capped mRNA of that species with the higher luciferase expression. L.S. denotes that the luciferase signal was not significantly above the background. Error bars represent mean ± SD. Three independent experiments were performed, n = 3. p-values are denoted by asterisks: *p < 0.05; **p < 0.01; ***p < 0.005 (Student’s t-test, NAD-RNA vs. m7G-RNA, one sided). Source data are provided as a Source Data file.
distribution, the library pools were either supplemented with 20% v/v Phix Control (Illumina) or custom Illumina sequencing primers, which bear three Gs at their 5′-end, prepared by standard procedures to mitigate library imbalances before NextSeq 500 75 bp single-end (SE) sequencing.

Fragmented NAD captureSeq: Biological triplicates of total RNA (gDNA free) were used as library starting material. Total RNA (100 μg), supplemented with 5 ng NAD-RNA (as optional spike-in control), were randomly sheared in a 65 μL reaction volume that contained 25 μL of 0.5x fragmentation solution (2 μM EDTA, 10 mM Na2CO3, 90 mM NaHCO3, pH 9.3) at 94 °C (5 min for WT strain and 20 min for npy1A strain to approach similar fragment size). Sheared RNA fragments were visualized on a 1.2% formaldehyde-denaturing agarose gel. Next, the sheared RNA was rapidly chilled, either the total amount of 200 μL or by the addition of 600 μL ethanol, 20 μL 3 M NaOAc pH 5.5, and 1 μL glycogen at ~20 °C overnight. The precipitated RNA was washed with 75% ethanol and subsequently dissolved in 50 μL dH2O. An equivalent of 100 μg sheared RNA was then treated with 100 U T4 PNK, along with 0.1 mM ATP, 100 mM imidazole-HCl pH 6.0, 10 mM MgCl2, 10 mM β-mercaptoethanol, and 20 μg/ml RNase-free bovine serum albumin (RNA) in a total volume of 200 μL at 37 °C for 5.5 h. RNA extraction was performed twice, employing the P/C/I approach, and followed by triple ether extraction and ethanol precipitation. Precipitated RNA was the washed again, using 75% ethanol, and ultimately dissolved in 20 μL dH2O, yielding the library input for the standard NAD captureSeq protocol. PCR products were size-selected within a range of 150 to 300 bp (referred to as “small fragmented NAD captureSeq library”) and of 300–500 bp (referred to as “large fragmented NAD captureSeq library”), enabled by 10% native PAGE. Bioanalyzer QC library multiplexing, and the overall sequencing strategy were executed in a similar manner, as done for the unfragmented NAD captureSeq. Again, the NextSeq 500 75 bp single-end protocol was chosen for sequence analysis.

Transcriptome libraries: Biological triplicates of total RNA (gDNA free) served as library input material. Total RNA (1 μg) was subjected to rRNA depletion by Ribo-Zero RNA Removal Kit (yeast). rRNA-depleted RNA was randomly sheared in 10 μL dH2O, at 94 °C for 10 min. Fragmented RNA was processed using the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB). The sequence of raw reads was converted into its reverse complementation using NEBNext Multiplex Oligos for Illumina. Further cDNA size selection in the range of 300–500 bp was performed, employing the Agencourt RNAClean XP kit. Primer-depleted cDNA was examined by Bioanalyzer and the concentration was measured by Qubit. Multiplexed libraries were sequenced by NextSeq 500 75 bp SE.

NGS analysis: Unfragmented NAD captureSeq analysis: Original reads were demultiplexed, based on the PCR barcode, not allowing for mismatches and subjected to standard read trimming procedures. Further 5′-end adapters (C NNNNNN AGATCG) were trimmed (minimum length 12 nt) by in-house scripts. Reads that mapped to the spike-in internal standard (IS) RNA sequence (bowtie -v 2, version 1.2.1) were then counted first. IS unmapped reads were subsequently classified as “small RNA reads” (12–17 nt) or “normal RNA reads” (≥18 nt).

Normal RNA reads were mapped to the reference genome S. cerevisiae BY4742 strain (BY4742_Toronto_2012, SGD) (bowtie -v 2). Normal RNA reads that mapped to rRNA genes or tRNA genes were separated and counted as rRNA and tRNA reads, respectively. Normal RNA reads were further analyzed for a consensus motif. The sequence sharpA feature was screened, based on the premise that the 5′-UTR region accumulating at the position, within the reference genome, was further analyzed for a consensus motif. Fragmented NAD captureSeq was analyzed in the same way, as done for the unfragmented NAD captureSeq protocol.

RNA pulldown and UPLC-MS analysis: Streptavidin Sepharose High Performance Beads (250 μL) were loaded on Mobicil Classic columns. The column was washed three times with 1× phosphate-buffered saline (PBS) buffer, then five times 25 μL of 25 μM biotin-DNA probe (Biomers, Supplementary Table 1) were added sequentially. The mixture was incubated at 25 °C, for 10 min. Next, the column was washed with three times 30 μL PBS, followed by equilibration in 300 μL pull-down buffer (10 mM Tris-HCl pH 7.8, 0.9 M tetramethylammoniumchloride, 0.1 M EDTA pH 8.0). Total RNA (gDNA free; 200–500 μg) was added into the column and incubated at 65 °C for 10 min and then rotated (Tube Rotator, VMR) at 20 °C for 25 min. Next, the column was washed six times with 200 μL dH2O, to remove uncleaved RNA and rRNA, by adding four times 200 μL pull-down buffer (75 °C, pre-heated) under 10 s per min shaking (350 r.p.m.) at 75 °C for 10 min. The eluate was precipitated with 0.5 M ammonium acetate pH 5.5 and 50% isopropanol. Precipitated RNA was dissolved in dH2O for UPLC/MS analysis.

To determine the amount of NAD that is covalently linked to RNA, the RNA samples were washed three times with 400 μL of 8.3 M urea, one time with dH2O, two times with 4.15 M urea, and again four times with dH2O in Amicon Ultra-0.5 mL Centrifugal Filter Units 10 kDa, to remove non-covalently bound cellular NAD. The recovered RNA was subsequently concentrated. The pulldown RNA samples or 10 μg urea-washed total RNA samples were treated with 10 μM NudC in the presence of 10 μM MgCl2, 0.6 ng/mL of d1-4 ribose, isolated using the UPLC-MS/MS analysis protocol using a 1500 V capillary voltage, 11 V cone voltage, 150 °C ionization was performed with a 1500 V capillary voltage, 11 V cone voltage, 150 °C
source temperature, 200 °C desolvation temperature, 150 L h⁻¹ desolvation gas flow, and 800 L h⁻¹ desolvation gas flow (N₂). The Xevo TQ-S was automatically tuned to acquire the MassLynx V4.1 with the IntelliStart standard procedures. Multiple reaction monitoring measurements were conducted, using collision argon (argon, 0.15 mL min⁻¹) for collision-induced decomposition and MS/MS transitions were monitored in the positive ion mode (N-ribosylnicotinamide: m/z 254.94–122.81, d4-N-ribosylnicotinamide m/z 258.94–126.81, 20 V, 50 mV dwell time for each mass transition).

**Gel electrophoresis.** Native PAGE (10%) was utilized to size select cDNA for NGS. Briefly, 10% acrylamide/Bis solution (19:1), 0.1% Ammonium persulfate (APS) (w/v), and 0.1% N,N,N′-Tetramethylethylenediamine (TEMED) (w/v) along with 1× Tris-borate-EDTA (TBE) in 50 mL volume were mixed and poured between glass plates (19 cm × 27 cm). Gel mixtures were polymerized at room temperature for 45 min. The electrophoresis was conducted at a stable 27 mA current for 2.5 h. The gel was then stained with SYBR Gold (Thermo Scientific) in 1× TAE buffer, 5 min. The signal intensities were read out by scanning the gel at 400 V, 50 or 100 µm resolution using a Typhoon FLA 9500. The printout picture of the gel (in its original size) was used for excision of desired size ranges within the corresponding gel lanes. APB gel electrophoresis was utilized to separate NAD-RNA, mG-G-RNA, and ppp/RNA from each other. Briefly, 0.5% (w/v) APB (Lab stock), 10% acrylamide/Bis solution (19:1), 0.1% (w/v) APS, and 0.1% (v/v) TEMED with 2× Tris-acetate-EDTA (TAE) buffer in 50 mL volume were mixed and poured between glass plates (Bio-Rad). Gels were run at a stable current (15–25 mA per gel) in 1× TAE buffer. Gels were then stained with SYBR Gold or gels, containing 32P-labeled nucleic acids, were exposed to storage phosphor screens (GE Healthcare) and visualized using a Typhoon FLA 9500. Signal quantification was performed using the ImageQuant software (GE Healthcare).

**In vitro transcription variants.** Radio-labeled RNA: 400 nM double-stranded DNA (dsDNA) served as a template in a transcription buffer containing 40 mM Tris-HCl pH 7.9, 7 mM spermidine, 22 mM MgCl₂, 0.01% Triton X-100, 10 mM dithiothreitol (DTT), 5 mM dimethyl sulfoxide, along with 60 µCi α−32P-ATP, 10 nCi α−32P-CTP, 4 mM each CTP/GTP/UTP, 2 mM ATP, 6 mM NAD (NAD-RNA) or without NAD (ppp-RNA), 0.7 mM T7 polymerase (self-prepared) within a total overall volume of 10 µL for 40 min at 37 °C. Remaining NAD-capped RNA was added subsequently to the reaction and incubated for an additional 30 min at 37 °C. The modiﬁed NAD amounts were normalized by the amount of measured lysate RNA. mRNAs was adjusted accordingly.

**Cellular NAD quantification.** The general experimental procedure was based on the manufacturer’s protocol (NAD/NADH Quantitation Kit) with minor modiﬁcation. Yeast cells were cultured in 50 mL YPD medium in biological quadruplicates (n = 4). One milliliter of these cultures, at an OD₆₀₀ of ~0.8, were pelleted by the centrifugation at 4000 × g, 1 min, 4 °C. The cell pellets were washed twice with 1 mL iced cold breaking Buffer (Epicentre) and transferred into 1 mL Ice-cold Breaking Buffer (containing Sephadex G-25 Fine resin) were pre-equilibrated with 100 mL Breaking Buffer (Epicentre) with 0.5 mL PMSE. The injected sample was resolved by the indicated column matrix running on a FPLC system (flow rate 1 mg min⁻¹, 0.5 mL collected fractions) employing the same equilibration buffer. Absorption values of each fraction were determined at 260 nm (A₂₆₀) and appropriate fractions, exceeding 75% of the highest A₂₆₀ value, were pooled together. Next, 1 mL CaCl₂ and 50 µL 1 M micrococcal nucleases were added to the pooled fractions, followed by incubation at 26 °C for 15 min. The reaction was stopped by adding 1 M EDTA to a ﬁnal concentration. Aliquots of 100 µL each were then frozen in liquid nitrogen and stored at −80 °C. In vitro translation: Master Translation Solution was freshly prepared by mixing 25 mM HEPES-KOH pH 7.6, 1.25 mM ATP, 0.125 mM GTP, 0.15 µM 32P-c creativity phosphokinase, 2.5 mM DTT, 125 mM KOAc, 5 mM MgOAc pH 7.0, 25 mM Amino Acid Mixtures, 1 U mL⁻¹ RNas in Ribonuclease Inhibitors, 80 mM mG-capped Renilla mRNA, and dH₂O in a final volume 80 µL. The solution was gently mixed and 4 µL of the Master Translation Solution aliquoted to individual PCR tubes. Then, 1 µL mRNA (200 ng or accordingly) and 5 µL of the cell-free extract were added to thus prepared reaction volumes. The “ready-for-transcription” reactions were then mixed gently and incubated again at 26 °C for 30 min. Next, 90 µL dH₂O were added to each in vitro translation reaction. Seventy ﬁve microliters of this dilution were then used to conduct a ﬁrefly luciferase activity assay (Bright-Glo Luciferase Assay System). The remaining 25 µL, with additional 50 µL of dH₂O, were subjected to a Renilla luciferase activity assay (Renilla-Glo Luciferase Assay System), executed according to the manufacturer instructions. Emitted luminescence was read out using a TECAN plate reader.

**RNA in vitro decapping and NAD hydrolysis kinetic assays.** In general, 32P-body-labeled NAD/mG-capped RNAs or S-32P-labeled NAD were incubated with 0.5 mL NADK-1. The amount of purified rabbit reticulocyte lysate was added to bring the total volume of the reaction to 25 µL Tris-HCl pH 7.5, 50 mM NaCl, 50 mM KCl, 2 mM MgCl₂, 1 mM DTT, 1 mM MnCl₂ (was absent in reactions refered to as without Mn²⁺) kinetics)
incubated at 37 °C for 120 min. Samples that were treated with the E. coli Nudix
hydrolase NudC are referred to as “positive control.” Aliquots of 5 µL were
then transferred to 96-well plates (at indicated time points) and treated with
mixed with the same volume of 2× APB Gel Loading Buffer (50 mM Tris-HCl pH
7.5, 8 M Urea, 20 mM EDTA, 20% Glycerol, 0.01% Xylenol, 0.01% Bro-
mophenol Blue) and placed on ice for further APB gel electrophoresis. Reaction
mixtures, containing NAD, were stored on ice before performing thin-layer
chromatography. Hydrophilic Magnets (Eppendorf ALDMAG) were added to the
suspension (1 µL GIL UV/25 cm x 20 cm). Resolution of nucleotides via TLC, at room temperature for 5.5 h, was
achieved employing a flow phase of 1 M NH₄OAc and ethanol (4:6).

**Determination of RNA NAD-capping ratios.** For each assay, 100 µg total RNA
gDNA was subjected to ADPRC treatment as fully treated group. An equal
amount of total RNA was applied to the same treatment without ADPRC as
background group. The subsequent copper-click reaction, capture by streptavidin
beads (streptavidin-unbound flow-through RNAs (non-NAD-RNA) were collected
and precipitated with ethanol), as in the standard NAD captureSeq protocol11.
For RT on beads, to each sample of the fully treated group, as well as the back-
ground group, were added 2.5 µm random hexamers, 0.5 mM dNTP mix, and
and precipitated with ethanol), as in the standard NAD captureSeq protocol11.

**RT-PCR and standard PCR procedures.** Real-time PCR was performed using 250
µM Fw/Rev primer, 5 µl diluted cDNA, and 1x SsoAdvanced Universal SYBR
Green Supermix in 20 µl reaction. PCR conditions were the following, denaturing
step at 95°C (2 min), 40 cycles of consecutive annealing/extension steps at 95°C
for 7 s and 60°C for 15 s, respectively. Melting curves were generated by heating
from 65°C to 95°C with an incremental increase of 0.5 °C s⁻¹. Fluorescence was
measured throughout using a LifeCycle 480 Instrument.

Two rounds of PCR were carried out to generate linear DNA templates for
in vitro translation reactions as follows: the first round aimed to bridge the 5′-UTR/
CDS to the firefly Luc sequence. The reactions contained 6.4 µM 5′-UTR/CDS
template, 6.4 µM bridge region DNA (Supplementary Table 2), and 6.4 µM firefly
Luc2 DNA template. 200 nM dNTPs, 1x Q5 Reaction Buffer (NEB), 0.02 U µL⁻¹
hot-start Q5 high-fidelity DNA polymerase in a total volume 50 µL. The PCR was
initiated by heating to 98 °C for 40 s and followed by 5 cycles (98 °C 10 s, 65 °C 20 s,
72 °C 2 min). The second round of the PCR aimed to specifically amplify the
bridged 5′-UTR/CDS, bearing the Luc2 template extension, by additionally adding
corresponding 500 nM Fw_5UTR primer, 500 nM Rev_Luc2_polyA primer, and
1x Q5 Reaction Buffer (NEB) to the final volume of 56.2 µL.

All other PCRs were performed using the Q5 high-fidelity DNA polymerase
for amplifying NAD captureSeq cDNA with barcodes, as described11 and related
dNA templates for in vitro translation assay, as linear templates from plasmids.
Otherwise, Taq polymerase/Q5 high-fidelity DNA polymerase were employed to
amplify 5′-UTR/CDS sequences from gDNA and from plasmids, obtained by using
standard PCR procedures.

**Quantification.** PAGE gel, APB gel, and TLC intensities were quantified using
the ImageQuant software (GE Healthcare).

**Statistical analysis.** Samples for NAD captureSeq, transcriptomics, UPLC-MS,
RT-qPCR, and proteomics were prepared as biological triplicates. The outlined
in vitro experiments, including NAD kinetics, RNA decapping kinetics, and in vitro
translations, were performed as technical triplicates. Mean values (± 2 SEM)
were set using Student’s t-test (single-tail and double tail) only (FCM data), unequal variance
was calculated by R/python. For NGS TSS switching relevant math, covariance with
two variables, linear trend test, and corresponding p-values, as well as FDR value,
were calculated using python scripts. For NGS NAD captureSeq, NAD-RNA
enrichment and transcriptome differential expression analysis, DESeq2 was
utilized, as described above.

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

**Data availability**
NRG raw data and analyzed files are available in the GEO repository under the GEO
Accession: GSE146368. Proteomics raw data and analyzed files are deposited in the
ProteomeXchange Consortium via the PRIDE repository: PXD017893. The data
supporting the findings of this study are available from the corresponding authors upon
reasonable request. Source data is provided with this paper.

Received: 4 June 2020; Accepted: 24 September 2020;
Published online: 02 November 2020

**References**
1. Shatkin, A. J. Capping of eucaryotic mRNAs. Cell 9, 645–653 (1976).
2. Topisirovic, I., Sviliv, Y. V., Sonenberg, N. & Shatkin, A. J. Cap and capping
proteins in the control of gene expression. Wiley Interdiscip. Rev. RNA
2, 277–298 (2011).
3. Belasco, J. G. All things must pass: contrasts and commonalities in eukaryotic
and bacterial mRNA decay. Nat. Rev. Mol. Cell Biol. 11, 467–478 (2010).
4. Cahová, H., Winz, M. L., Höfer, K., Nübel, G. & Jäschke, A. NAD captureSeq
indicates NAD as a bacterial cap for a subset of regulatory RNAs. Nature
519, 374–377 (2015).
5. Frindt, J. et al. Identification, biosynthesis, and decapping of NAD-capped
RNAs in E. coli. Cell Rep. 24, 1890–1901 (2018).
10. Zhang, H. et al. NAD tagSeq reveals that NAD with NADK. J. Bacteriol. 202, e00591–00519 (2020).

11. Zhang, D. et al. Structural basis of prokaryotic NAD-RNA decapping by a subset of the transcription initiation sites at the CYC-1 promoter of Saccharomyces cerevisiae using 5′-3′ exoribonuclease Rat1 and its activating partner Rai1. Nature 458, 784–789 (2009).

12. Schiestl, R. H. & Gietz, R. D. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. Curr. Genet. 16, 339–346 (1989).

13. Kiledjian, M., Zhou, M. & Jiao, X. Normal and aberrantly capped mRNA decapping and 5′-3′ exoribonuclease activity of NudC. Nat. Struct. Mol. Biol. 18, 1224–1230 (2011).

14. Gilbert, W. V., Zhou, K., Butler, T. K. & Doudna, J. A. Cap-independent mRNA decay as a default fate. Nature 501, 12102 (2019).

15. Jiao, X. et al. Identiﬁcation of a quality-control mechanism for mRNA 5′-3′ capping with NADK. Mol. Cell 70, 553–564 (2018).

16. Meurer, M. et al. Deﬁnition of a high-conﬁdence mitochondrial proteome at quantitative scale. Cell Rep. 19, 2836–2852 (2017).

17. Bird, J. G. et al. The mechanism of RNA 5′ capping with NADK. NADH and dephospho-CoA. Nature 535, 444–447 (2016).

18. Vvedenskaya, I. O. et al. CapZyme-Seq comprehensively deﬁnes promoter-sequence determinants for RNA 5′ capping with NADK. Mol. Cell 70, 553–564 (2018).

19. Bresson, S. & Tolleryev, D. Surveillance-ready transcription: nuclear RNA decay as a default fate. Open Biol. 8, 17027 (2018).

20. Xiao, X. et al. Identiﬁcation of a quality-control mechanism for mRNA 5′-end capping. Nature 467, 608–611 (2010).

21. Kim, M. et al. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature 432, 517–522 (2004).

22. Jiao, X. et al. Identiﬁcation of a quality-control mechanism for mRNA 5′-end capping. Nature 467, 608–611 (2010).

23. Hahn, S., Hoar, E. T. & Guarente, L. Each of three TATA elements upstream of the yeast CYC-1 promoter requires a speciﬁc region-speciﬁc microRNA-mediated gene regulation. J. Biol. Chem. 289, 179 (2010).

24. Gilbert, W. V., Zhou, K., Butler, T. K. & Doudna, J. A. Cap-independent mRNA decay as a default fate. Nature 501, 12102 (2019).

25. Xu, W., Dunn, C. A. & Bessman, M. J. Cloning and characterization of the Saccharomyces cerevisiae NudC. Nucleic Acids Res. 33, 8566 (2005).

26. Tome, J. M., Tippens, N. D. & Lis, J. T. Single-molecule nascent RNA cleavage by G. F. RNA cleavage by the 10-23 DNA enzyme. Methods Enzymol. 341, 503–517 (2001).

Acknowledgements

We thank Jäschke Lab members, V. Winkler (Heidelberg University), and A. Hotz-Wagenblatt (DFKZ, Heidelberg) for discussions; H.C. Lee and D. Grimme for strains and plasmids; M. Brunner for access to LightCycler, ZMBH Flow Cytometry & FACS Core Facility for the FCM measurement; M. Rettel and F. Stein (EMBL Proteomics Core Facility, Heidelberg) for proteomics analysis; and bwHPC (bwForCluster MLS&WISO and bwUniCluster) for cluster computation resources. This work was supported by the German Research Foundation (DFG, grant number Ja794/10, SPP 1784, to A.J.).

Author contributions

Conceptualization: Y.Z. and A.J. Methodology: Y.Z., D. Kuster, T.S., D. Kirrmaier, G.N., D.I., and H.H. Formal analysis: Y.Z., M.K., H.H., and A.J. Supervision: H.H., M.K., and A.J. Administrative support: D.I., and V.B. Investigation: Y.Z., D. Kuster, T.S., D. Kirrmaier, G.N., D.I., and H.H. Conceptualization: Y.Z. and A.J. Methodology: Y.Z., D. Kuster, T.S., D. Kirrmaier, G.N., D.I., and V.B. Formal analysis: Y.Z., M.K., H.H., and A.J. Supervision: H.H., M.K., and A.J. Administrative support: D.I., and V.B. Funding: Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-19326-3.

Correspondence and requests for materials should be addressed to A.J.

Peer review information Nature Communications thanks Tatsutomu Suzuki, Toshiyuki Inada and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

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