Angewandte
Eine Zeitschrift der Gesellschaft Deutscher Chemiker
Chemie

Supporting Information

AlkAniline-Seq: Profiling of m^7G and m^3C RNA Modifications at Single Nucleotide Resolution

*Virginie Marchand, Lilia Ayadi, Felix G. M. Ernst, Jasmin Hertler, Valérie Bourguignon-Igel, Adeline Galvanin, Annika Kotter, Mark Helm, Denis L. J. Lafontaine, and Yuri Motorin**

anie_201810946_sm_miscellaneous_information.pdf

Author Contributions

V.M. and L.A. performed AlkAniline-Seq, F.E., A.G. and V.B.I. prepared strains/cell lines and extracted RNA, J.H. and A.K. performed LC-MS analysis, V.M., M.H., D.L.J.L. and Y.M. analysed the data and wrote the manuscript.

Supporting Information

Materials & Methods

Yeast and bacterial strains and human cells used in the study

Yeast strains used in this study were obtained from the EUROSCARF collection (Germany, see Supplementary Table S1). Cells were grown in standard Yeast Extract/Peptone/Dextrose (YPD) to 5.5–6.0 OD_{600} to stationary phase.

Bacterial strain E. coli DH5α used in this study is described in Supplementary Table S1. Cells were grown in standard Luria Bertani medium to 2.5–3.0 OD_{600} to stationary phase.

Supplementary Table S1

N°	Species	Name	Genotype
1	*S. cerevisiae*	WT BY4741	BY4741; MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0
3	*S. cerevisiae*	*bud23*Δ	BY4742; MATα; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; YCR047c::kanMX4
4	*S. cerevisiae*	*trm8Δ*	BY4742; MATα; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; YDL201w::kanMX4
5	*S. cerevisiae*	*trm82Δ*	BY4741; MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; YDR165w::kanMX4
6	*S. cerevisiae*	*trm140Δ*	BY4741; MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; YOR239w::kanMX4
7	*S. cerevisiae*	*dus1*Δ	BY4741; MATa; ura3Δ0; leu2Δ0; met15Δ0; YML080w::kanMX4
8	*S. cerevisiae*	*dus2*Δ	BY4742; MATα; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; YNR015w::kanMX4
9	*S. cerevisiae*	*dus3*Δ	BY4741; MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; YLR401c::kanMX4
10	*S. cerevisiae*	*dus4*Δ	BY4742; MATα; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; YLR405w::kanMX4
11	*E. coli*	WT DH5α	F– Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1

Human cells (HCT116 p53 +/+) were grown at 37 °C under 5% CO_2. Cell growth was monitored by direct cell counting with a Scepter (EMD Millipore, Billerica, MA, USA). The cell lines used in this work were obtained directly from ATCC and passaged in the laboratory for fewer than 6 months after receipt.

RNA extraction

Total RNA from yeast or bacterial cells grown to the stationary phase was isolated using hot acid phenol extraction protocol [1].

Total RNA from human cells (HCT116 p53 +/+) was extracted using Trizol according manufacturer's recommendations.

Alkaline hydrolysis fragmentation

RNA (300 ng) was subjected to alkaline hydrolysis in 50 mM bicarbonate buffer pH 9.2 for 5 min at 96°C. The reaction was stopped as previously described [2].

Extensive dephosphorylation

RNA fragments without any gel-purification step were directly 3'-end dephosphorylated using 5 U of Antarctic Phosphatase (NEB, UK) for 1h at 37°C. After inactivation of the phosphatase, RNA fragments were extracted by a phenol:chloroform:isoamyl alcohol mix (25:24:1) and ethanol precipitated. After centrifugation, the pellet was washed with 80% ethanol and air-dried for 5 min.

Aniline treatment

The RNA pellets were resuspended in 1M aniline pH4.5 and incubated for 15 min at 60°C in the dark. The reaction was stopped by ethanol precipitation, and treated as previously described [3].

Library preparation

RNA fragments were converted to library using NEBNext® Small RNA Library kit (NEB ref E7330S, UK, or equivalent from Illumina, USA) following the manufacturer's instructions. DNA library quality was assessed using a High Sensitivity DNA chip on a Bioanalyzer 2100. Library quantification was done using a fluorometer (Qubit 2.0 fluorometer, Invitrogen, USA).

Deep sequencing

Libraries were multiplexed and subjected for high-throughput sequencing using an Illumina HiSeq 1000 or MiSeq instrument with a 50 bp single-end read mode. Since clustering of short fragments was generally very efficient, libraries were loaded at 8 pM concentration per lane.

Bioinformatics pipeline

Raw reads were first trimmed using the Trimmomatic v32 software with default parameters [4]. Alignment to the reference sequence was performed using bowtie2 (v2.2.4) [5] in End-to-End mode with 'sensitive parameter' set. Counting of the mapped reads and positions of their 5'-extremities was performed using awk command [6]. Coverage for reference sequence was calculated using samtools mpileup command. 5'-end count was directly used for calculation of normalized cleavage. Stop-ratio for every position of the reference sequence was calculated using 5'-end count and coverage data. All other steps of analysis were performed in R-Studio 1.0.143 with R version 3.4.4.

Mg^{2+} fragmentation

Instead of alkaline hydrolysis, RNA was subjected to Mg^{2+} ion-based RNA cleavage in 100 mM Tris-HCl buffer, pH 8.0, containing 2 mM $MgCl_2$ for 3 min at 96°C. The reaction was stopped as for alkaline hydrolysis and all other steps in the procedure remained the same.

De-phosphorylation by T4 PNK

After alkaline hydrolysis, RNA was 3'-end dephosphorylated using T4 PNK. Briefly, RNA was mixed with T4 PNK (10U) in a 50 µl reaction containing 100 mM Tris-HCl pH6.5, 100 Mg-OAc, 5 mM β-mercaptoethanol and incubated for 6h at 37°C. The PNK was heat inactivated at 65°C for 20 min. The RNA was then extracted with a phenol:chloroform:isoamyl alcohol mix (25:24:1) following the same procedure as the one described after phosphatase treatment in the AlkAniline-Seq protocol.

LC-MS measurements of m^7G content in rRNA and tRNA

Samples for LC-MS analysis were prepared from WT and mutant yeast strains using gel-purification for 18S rRNA [1] or the NucleoBond RNA/DNA 80 kit (Macherey Nagel) according to manufacturer's recommendations. 300 ng of tRNA or 18S RNA were digested into nucleotides using 0.3 U nuclease P1 from *P. citrinum* (Sigma-Aldrich), 0.1 U snake venom phosphodiesterase from *C. adamanteus* (Worthington), 200 ng Pentostatin (Sigma-Aldrich) and 500 ng Tetrahydrouridine (Merck-Millipore) in 5 mM ammonium acetate (pH 5.3; Sigma-Aldrich) for two hours at 37°C. The remaining phosphates were removed by 1 U FastAP (Thermo Scientific) in 10 mM ammonium acetate (pH 8) for one hour at 37°C. The nucleosides were then spiked with internal standard (^{13}C stable isotope-labeled nucleosides from *S. cerevisiae*, SIL-IS) and subjected to analysis. Technical triplicates with 50 ng digested RNA and 30 ng internal standard were analyzed via LC-MS (Agilent 1260 series and Agilent 6460 Triple Quadrupole mass spectrometer equipped with an electrospray ion source (ESI)). The solvents consisted of 5 mM ammonium acetate buffer (pH 5.3; solvent A) and LC-MS grade acetonitrile (solvent B; Honeywell). The elution started with 100% solvent A with a flow rate of 0.35 ml/min, followed by a linear gradient to 8% solvent B at 10 min and 40% solvent B after 20 min. Initial conditions were regenerated with 100% solvent A for 10 min. The column used was a Synergi Fusion (4 µM particle size, 80 Å pore size, 250 × 2.0 mm; Phenomenex). The UV signal at 254 nm was recorded via a diode array detector (DAD) to monitor the main nucleosides. ESI parameters were as follows: gas temperature 350°C, gas flow 8 l/min, nebulizer pressure 50 psi, sheath gas temperature 350°C, sheath gas flow 12 l/min, capillary voltage 3000 V. The MS was operated in the positive ion mode using Agilent MassHunter software in the dynamic MRM (multiple reaction monitoring) mode. For quantification, a combination of external and internal calibration was applied as described previously [7].

Kinetic studies of m^3C and m^7G instability in bicarbonate buffer system

150 pmol (3 µl of 50 µM solution) or 450 pmol (9 µl of 50 µM solution) of the respective nucleoside sample (m^3C or m^7G) was diluted in RNAse-free water to a total volume of 10 µl. To this mixture 10 µl of 100 mM bicarbonate buffer (pH 9.2, 50 mM final concentration) were added and incubated either at room temperature or at 96°C. After indicated time

points, the 20 µl sample was neutralized with 3.4 µl of a 1 % acetic acid solution (acetic acid in H_2O) and 22 µl of the final mixture were submitted to HPLC separation and monitored by UV absorption at 254 nm.

The HPLC measurements were performed on an Agilent 1100 HPLC series coupled with diode array detector and the MS measurements were performed on an Agilent 1100 LC/MSD Ion-Trap equipped electrospray ion source (parameters are displayed in Table ESI settings) (electron spray ionization and measurement in positive mode using a Synergi Fusion-RP column with 4 µm particle size, 80 Å pore size, 250 mm length and 2 mm inner diameter from Phenomenex (Aschaffenburg, Germany)). The elution was performed with a column temperature of 35°C and a flow rate of 0.35 ml/min. The solvents applied were a 5 mM ammonium acetate buffer adjusted to pH 5.3 (solvent A) and LC-MS grade acetonitrile (solvent B). A linear gradient from 0 % to 8 % solvent B at 10 min, 40 % solvent B at 20 min and 0 % solvent B at 23 min was used.

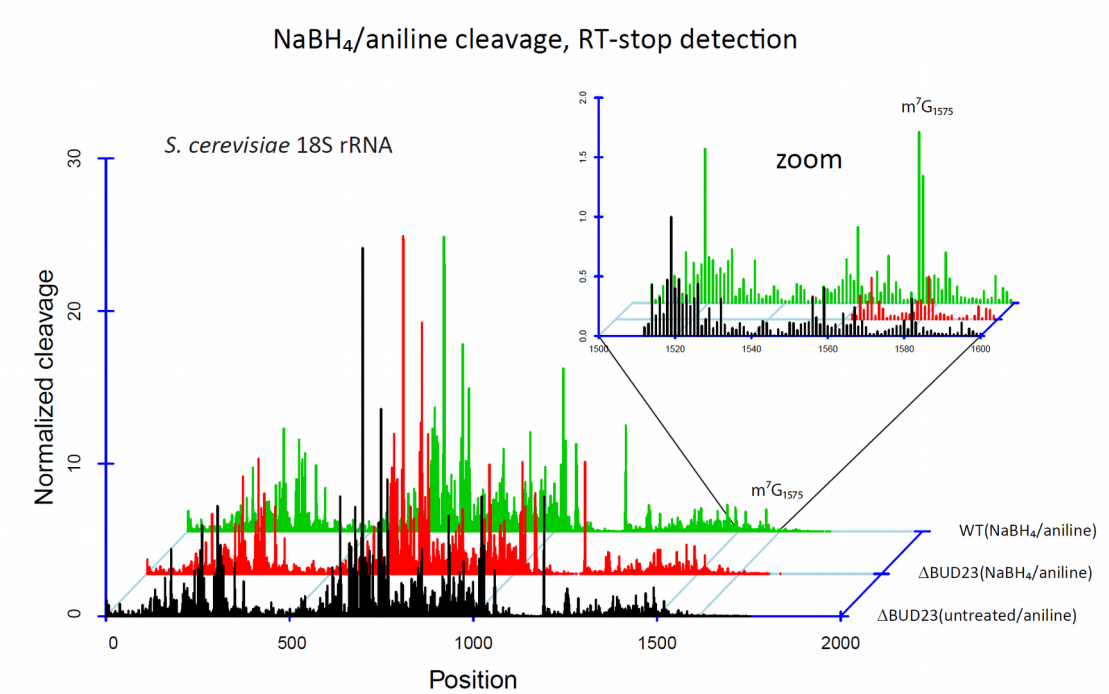
Table ESI settings.

Parameter	
Dry Temp	350°C
Dry Gas	12 l/min
Nebulizer pressure	30 psi
Capillary Exit	95.5 Volt

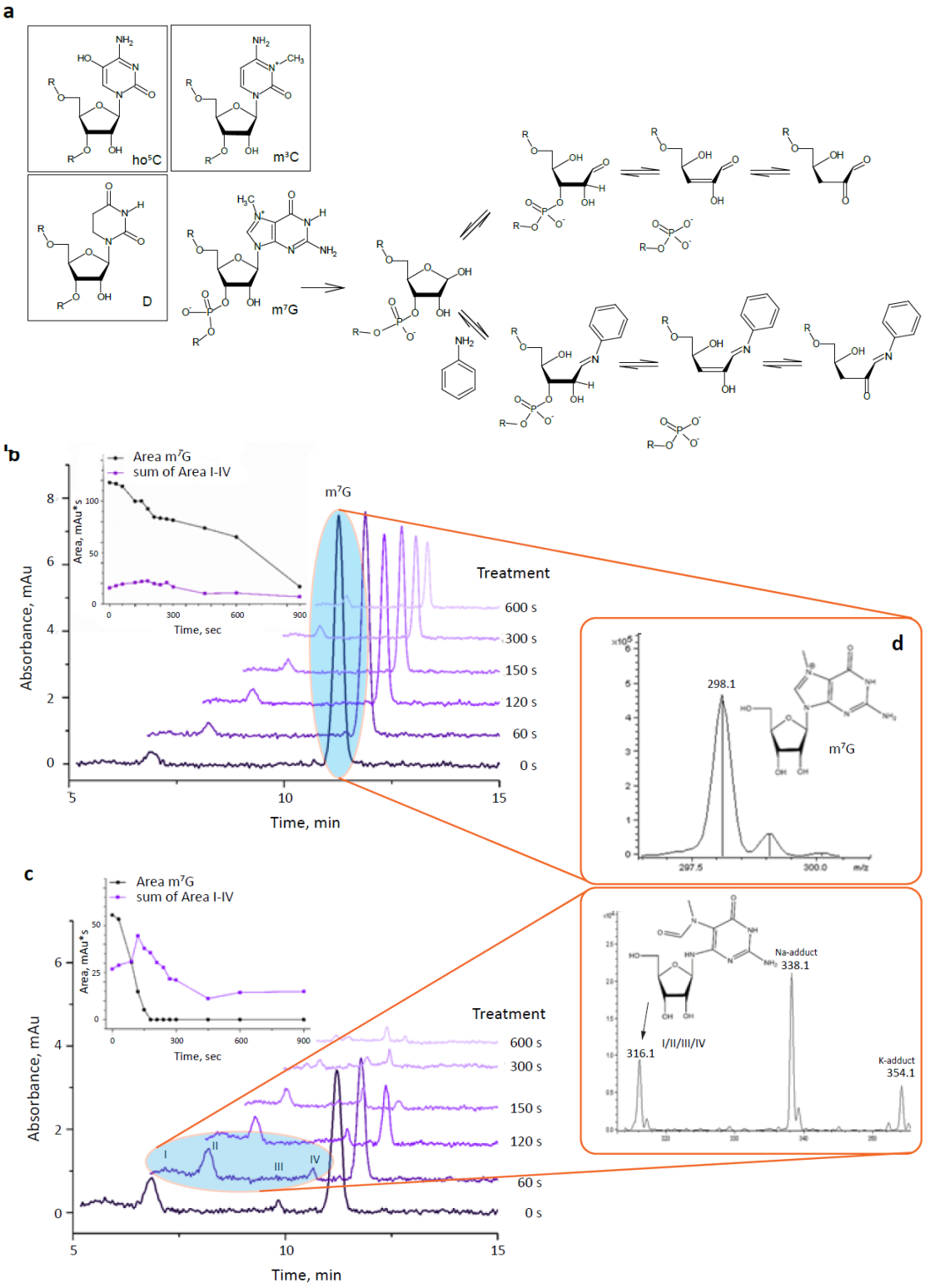
1. Collart, M. A. & Oliviero, S. Preparation of yeast RNA. *Curr. Protoc. Mol. Biol.* **Chapter 13,** Unit13.12 (2001).
2. Ayadi, L., Motorin, Y. & Marchand, V. Quantification of 2'-O-Me Residues in RNA Using Next-Generation Sequencing (Illumina RiboMethSeq Protocol). *Methods Mol. Biol. Clifton NJ* **1649,** 29–48 (2018).
3. Zueva, V. S., Mankin, A. S., Bogdanov, A. A. & Baratova, L. A. Specific fragmentation of tRNA and rRNA at a 7-methylguanine residue in the presence of methylated carrier RNA. *Eur. J. Biochem.* **146,** 679–687 (1985).
4. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinforma. Oxf. Engl.* **30,** 2114–2120 (2014).
5. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).
6. Marchand, V. *et al.* Next-Generation Sequencing-Based RiboMethSeq Protocol for Analysis of tRNA 2'-O-Methylation. *Biomolecules* **7,** (2017).

7. Kellner, S. *et al.* Absolute and relative quantification of RNA modifications via biosynthetic isotopomers. *Nucleic Acids Res.* **42,** e142 (2014).
8. Zorbas, C. *et al.* The human 18S rRNA base methyltransferases DIMT1L and WBSCR22-TRMT112 but not rRNA modification are required for ribosome biogenesis. *Mol. Biol. Cell* **26,** 2080–2095 (2015).
9. Figaro, S. *et al.* Trm112 is required for Bud23-mediated methylation of the 18S rRNA at position G1575. *Mol. Cell. Biol.* **32,** 2254–2267 (2012).
10. Létoquart, J. *et al.* Structural and functional studies of Bud23-Trm112 reveal 18S rRNA N7-G1575 methylation occurs on late 40S precursor ribosomes. *Proc. Natl. Acad. Sci. U. S. A.* **111,** E5518-5526 (2014).
11. Havelund, J. F. *et al.* Identification of 5-hydroxycytidine at position 2501 concludes characterization of modified nucleotides in E. coli 23S rRNA. *J. Mol. Biol.* **411,** 529–536 (2011).

Supplementary figures

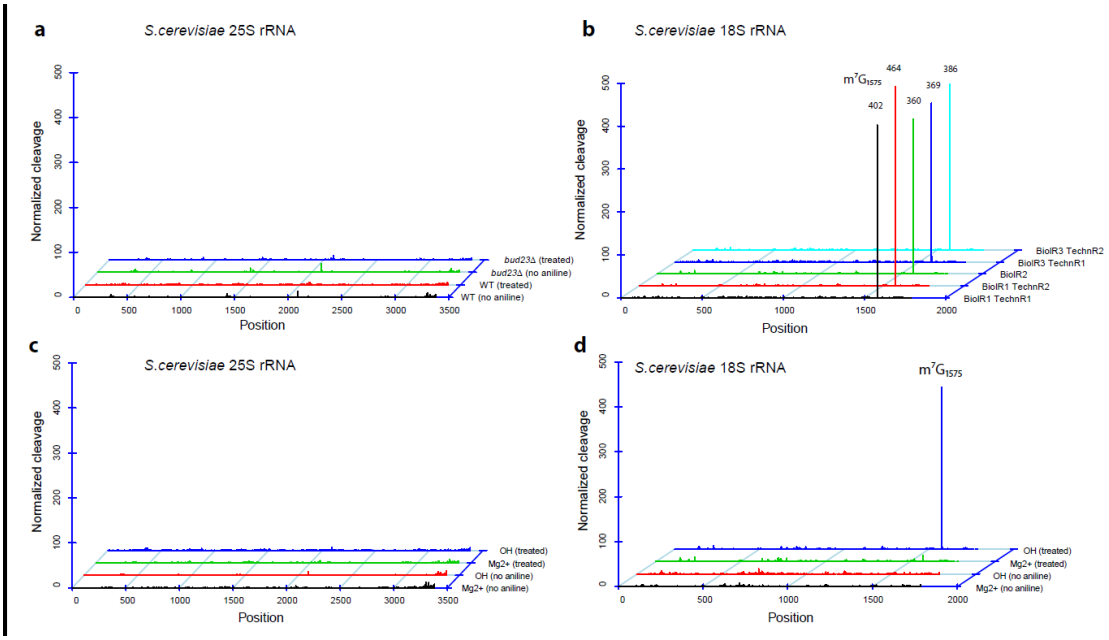


Supplementary Figure 1: Detection of m^7G1575 residue in *S. cerevisiae* 18S rRNA using a 'traditional' m^7G detection protocol which involves RNA strand scission by $NaBH_4$ reduction followed by aniline cleavage; a method which was formely used to detect the 18S rRNA m7G residue by reverse-transcription [8–10]. Here, the purified 18S rRNA was treated as described in the Materials & Methods section, followed by library preparation and sequencing. Analysis was performed with the same pipeline used for AlkAniline-Seq. Normalized cleavage profiles obtained for treated WT sample, and two controls: $NaBH_4$-reduced and un-reduced unmodified ΔBUD23 18S rRNA, are shown. Specific signal for m^7G1575 was detected, but the signal-to-noise ratio is barely higher than background (zoom).

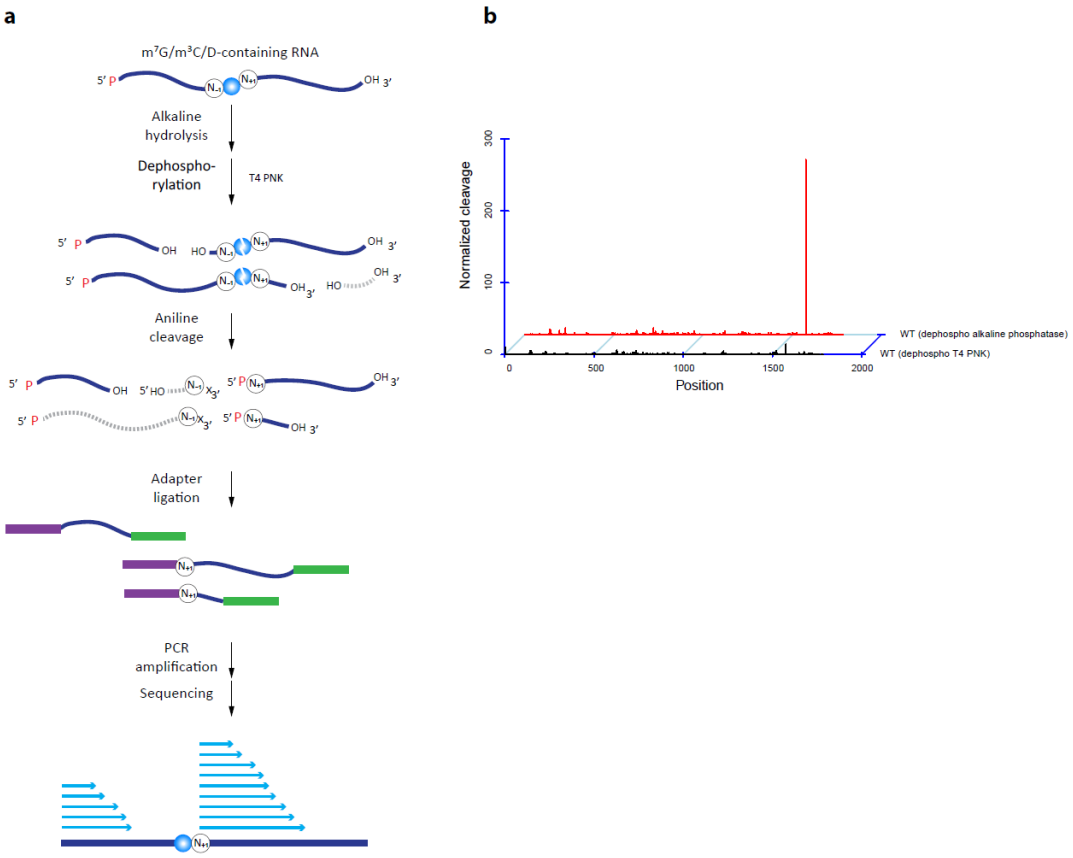


Supplementary Figure 2: **(a)** Chemistry of $m^7G/m^3C/ho^5C$ and D cleavage in RNA under alkaline conditions to create an abasic site and its subsequent decomposition in the presence of aniline. **(b,c)** Kinetics of nucleoside m^7G degradation in bicarbonate buffer (pH 9.2) at room temperature (b) and at 96°C (c). Aliquots drawn at the indicated time points were submitted to HPLC separation on an RP-18 column and monitored by UV absorption at 254 nm. LC-MS analysis in separate runs identified identical m/z values and fragmentation patterns for signals I-IV in (c). The area under the UV-peaks were used to plot the kinetics

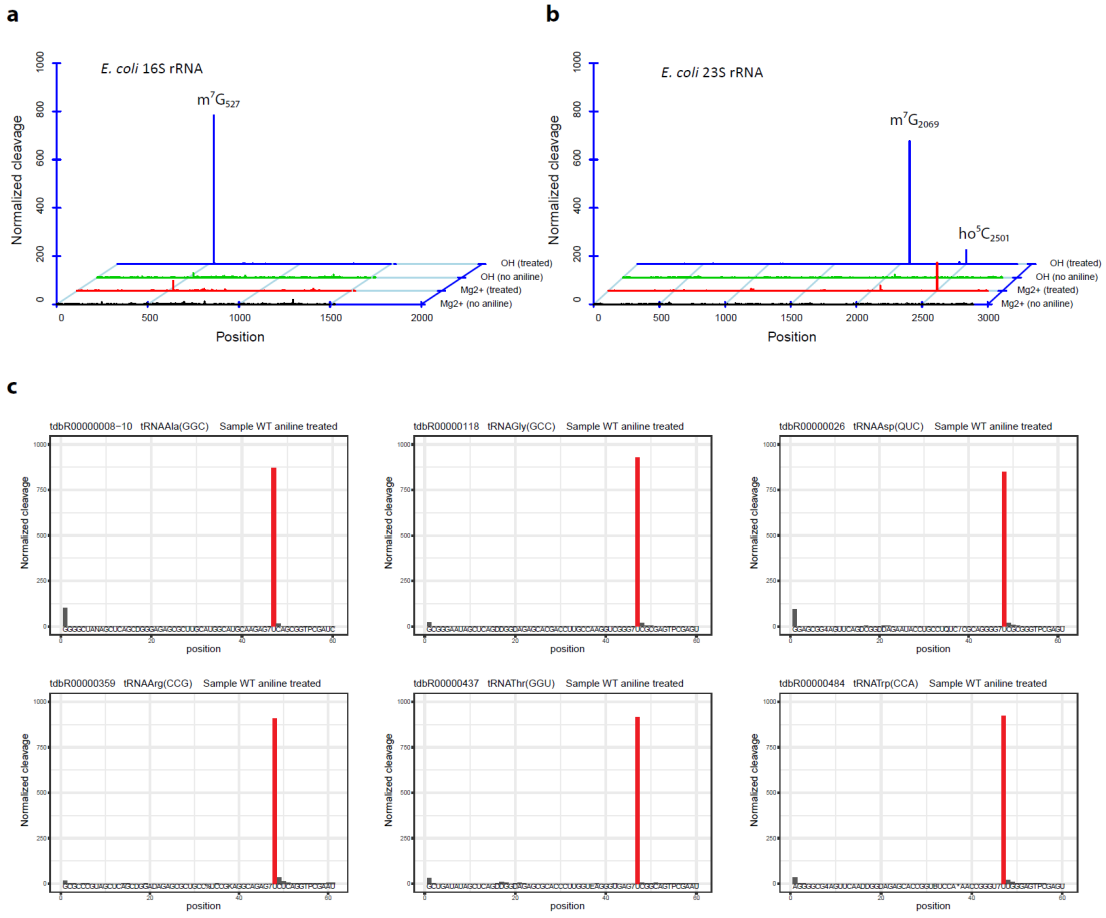
displayed in the inset, where the areas from signals I, II, III and IV were summed up. **(d)** The hypothetical structure was attributed to signals I-IV based on literature, but structural isomers are equally possible. Collection of material from peaks I-IV and subsequent reinjection produced an identical elution signal pattern, suggesting isomerization among I-IV.



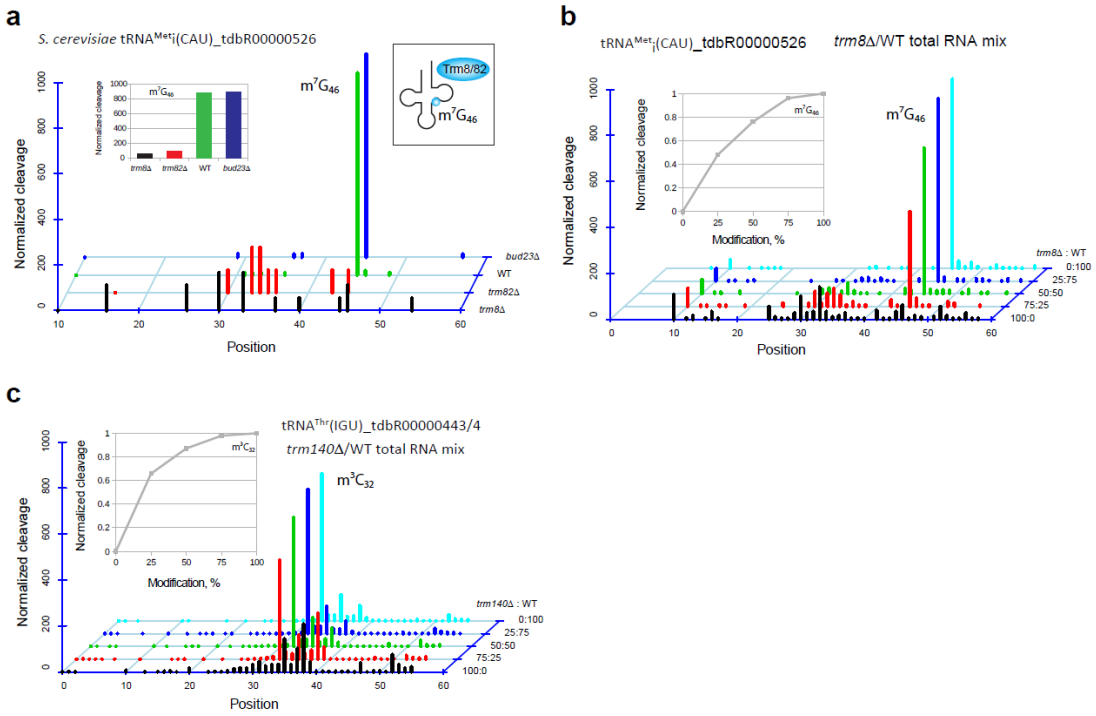
Supplementary Figure 3: **AlkAniline-Seq signals detected in *S. cerevisiae* 18S and 25S rRNA under different treatment conditions. (a)** Total RNA from WT and *bud23Δ* strains was subjected to AlkAniline-Seq in the absence or in the presence of aniline treatment. **(b)** AlkAniline-Seq signals in 18S rRNA (m^7G1575) obtained for two technical (TechR1/TechR2) and three biological replicates (BiolR1-R3). **(c,d)** Influence of RNA cleavage conditions on the AlkAniline-Seq signals, standard OH^- cleavage compared to Mg^{2+}-induced cleavage at neutral pH.



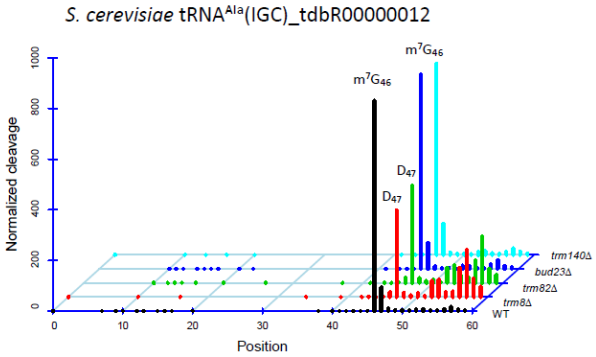
147
148 **Supplementary Figure 4**: **Procedure for AlkAniline-Seq with omitted 5'-dephosphorylation**
149 **step. (a)** To avoid 5'-phosphate removal, alkaline phosphatase was replaced by T4 PNK 3'-
150 phosphatase activity in the absence of ATP. Other steps remain as previously described in
151 Figure 1a. **(b)** Comparative cleavage profile for *S. cerevisiae* 18S rRNA with 3'-
152 dephosphorylation by T4 PNK and with alkaline phosphatase dephosphorylation.
153
154



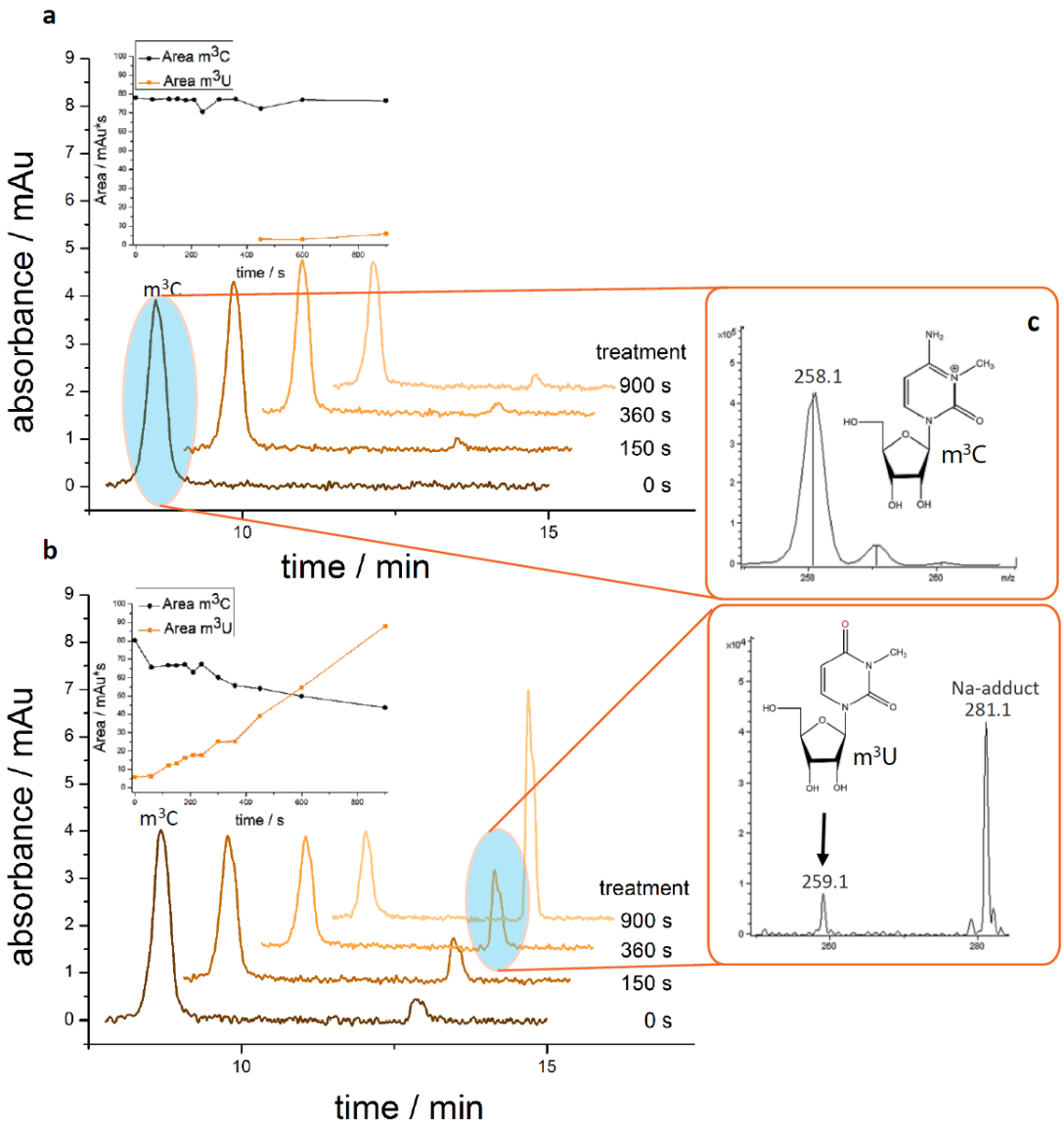
Supplementary Figure 5: **Specificity of AlkAniline-Seq protocol. (a,b)** AlkAniline-Seq signals obtained for m^7G527 and m^7G2069 in 16S (a) and 23S rRNA (b) from *E. coli*, respectively. Graphs show dependence on the cleavage conditions (OH^- vs Mg^{2+}) and aniline treatment. Only one additional minor signal at position 2501 in 23S rRNA was detected, which, interestingly, corresponds to ho^5C, which is a rare RNA modification only found so far in bacterial rRNA [11]. **(c)** Representative signals for m^7G46 residues in *E. coli* tRNAs. Only normalized cleavage profiles of OH^-/aniline treated samples are shown.



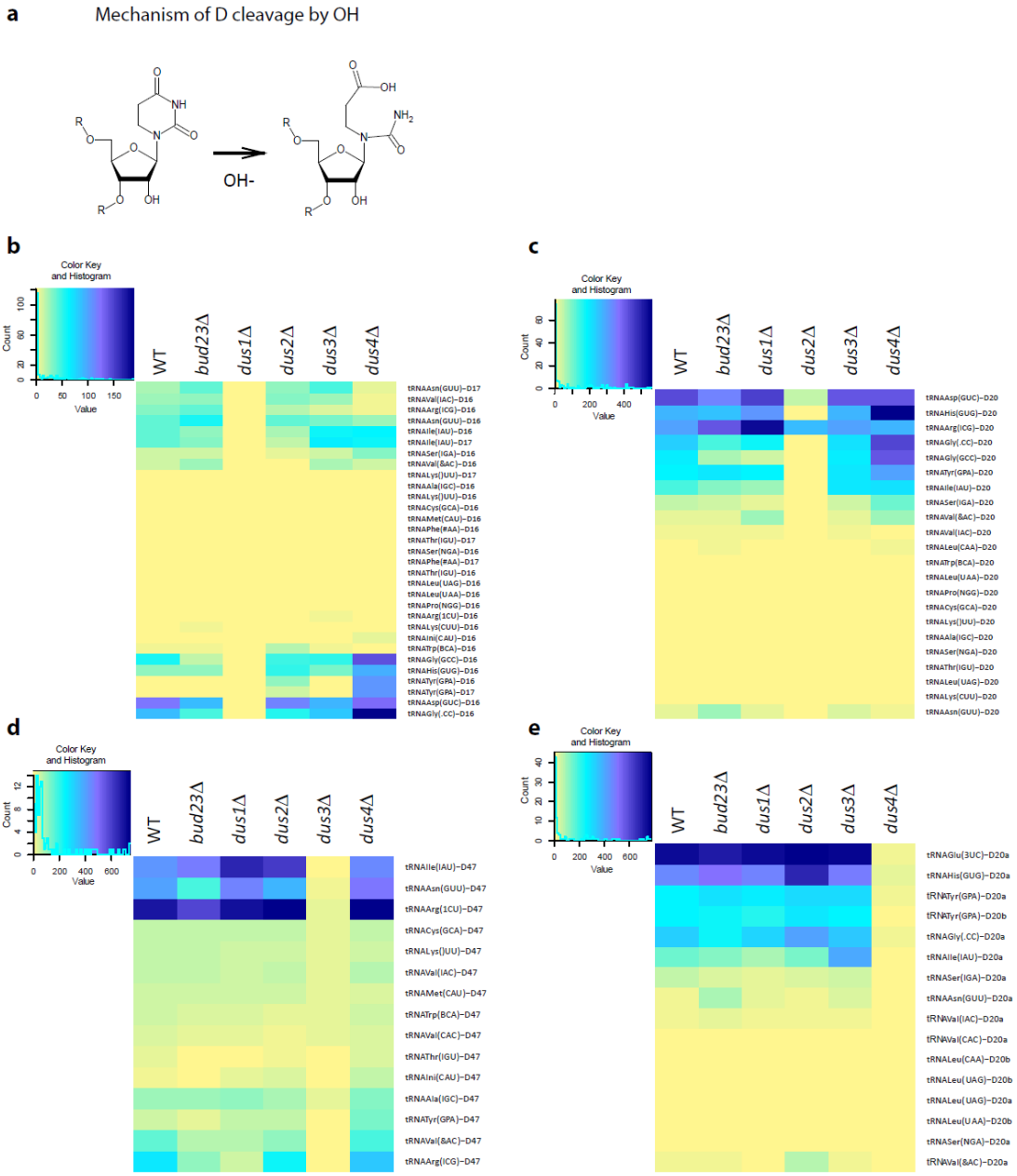
178 **Supplementary Figure 6**: **Detection and quantification of m^7G46 and m^3C32 residues in**
179 ***S. cerevisiae* tRNAs**. **(a)** Normalized cleavage signals for *S. cerevisiae* $tRNA^{Met}_i$(CAU) in *trm8Δ*,
180 *trm82Δ*, WT and *bud23Δ* yeast strains. Inset on the right shows the cloverleaf structure of
181 tRNA with the modified position and the corresponding enzymatic activity (heterocomplex
182 Trm8/Trm82). Inset on the left provides quantification of normalized cleavage in the four
183 strains used.
184 **(b)** Sensitivity of m^7G detection in $tRNA^{Met}_i$(CAU). Mixes of total RNA from *trm8Δ* and WT
185 strains at the proportions indicated at the right were subjected to AlkAniline-Seq. The inset
186 shows the relative normalized cleavage in function of molar ratio of modified WT tRNA.
187 **(c)** Sensitivity of m^3C detection in $tRNA^{Thr}$(IGU). Mixes of total RNA from *trm140Δ* and WT
188 strains at the proportions indicated at the right were subjected to AlkAniline-Seq. The inset
189 shows the relative normalized cleavage in function of molar ratio of modified RNA. The same
190 type of 'non-linear' calibration curve as observed both for m^7G and m^3C in RNA.
191
192
193
194
195
196



198
199 **Supplementary Figure 7**: **Detection and validation of the m^{7}G46 residue in yeast**
200 **tRNA^{Ala}(IGC).** Traces for WT and 4 mutant yeast strains are shown (labeled to the right of the
201 graph). The signal for m^{7}G46 disappears in Δ*trm8* and Δ*trm82* strains, lacking tRNA:m^{7}G46-
202 methyltransferase activity. In the absence of m^{7}G46, the predominant signal originates from
203 D47.
204
205
206
207

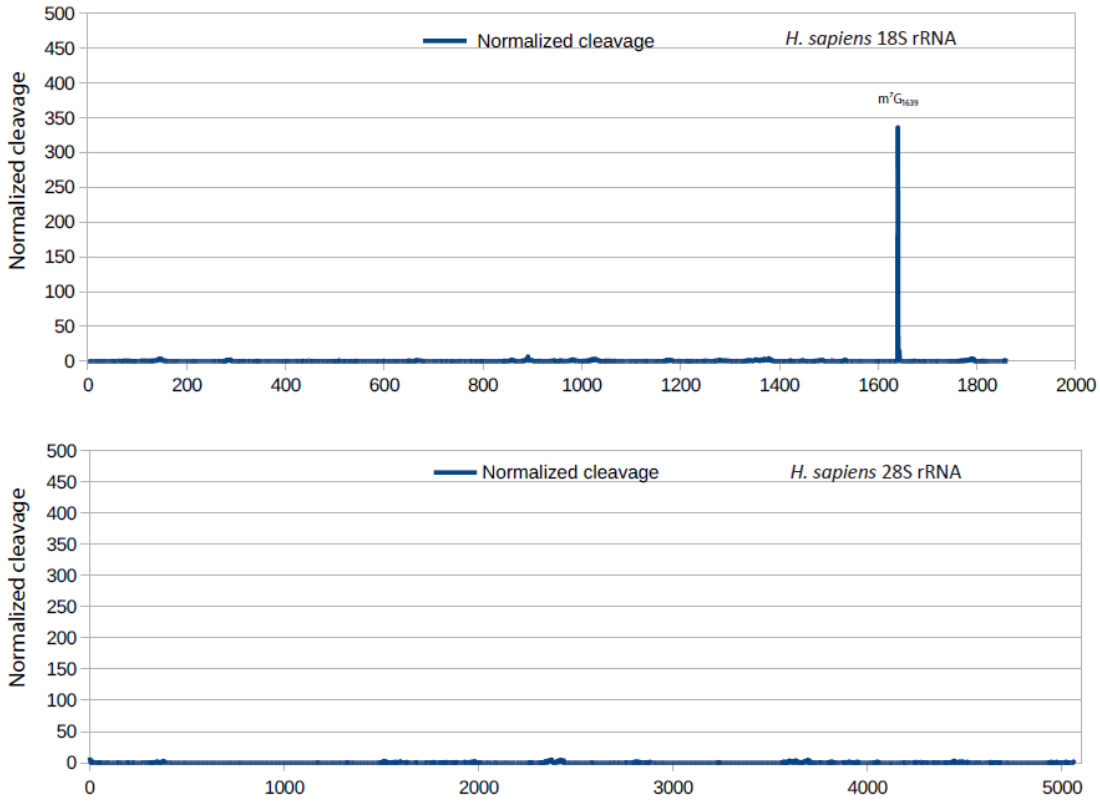


Supplementary Figure 8: **Kinetics of nucleoside m^{3}C deamination. (a,b)** The reaction was performed in bicarbonate buffer (pH 9.2) at room temperature (a) and at 96°C (b). Aliquots drawn at the indicated time points were submitted to HPLC separation on an RP-18 column and monitored by UV absorption at 254 nm. The area under the UV-peaks were used to plot the kinetics displayed in the inset. **(c)** LC-MS analysis in separate runs confirmed the structures displayed, and the identity of m^{3}U as major product was confirmed with an authentic sample (data not shown).

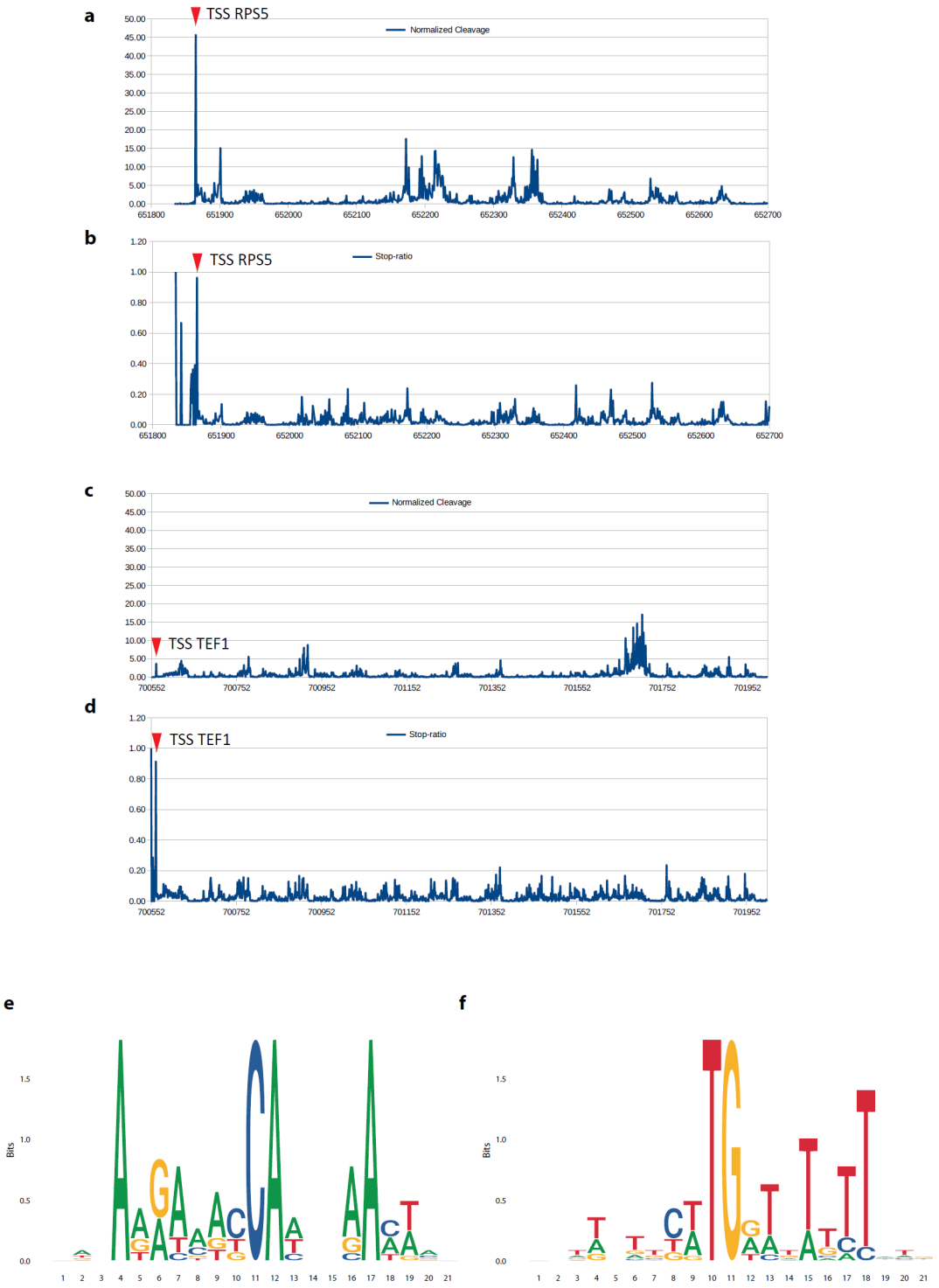


225 **Supplementary Figure 9**: **Detection of dihydrouridine residues by AlkAniline-Seq.** (a)
226 Chemical mechanism of D ring opening under alkaline conditions. **(b,c,d,e)** Heat maps of
227 AlkAniline-Seq signals for D residues in yeast *S. cerevisiae* tRNAs, grouped by their respective
228 enzymatic activity (*dus1-4*Δ). As shown in the figure, the majority of detectable signals are
229 decreased in the respective deleted strain.

230
231
232
233
234
235



Supplementary Figure 10: **Detection of a unique m^{7}G1639 residue in H. *sapiens* 18S rRNA by AlkAnilineSeq.** Graphs show Normalized cleavage for 18S (a) and 28S (b) human rRNA.



Supplementary Figure 11: **Transcription Start Signals (TSS) hits detected by AlkAniline-Seq in yeast *S. cerevisiae* mRNAs.** Graphs show Normalized cleavage (**a**,**c**) and Stop-ratio (**b**,**d**) for two representative hits: RPS5 (**a**,**b**) and TEF1 (**c**,**d**) mRNAs. The TSS is indicated by a red arrow. Consensus sequences obtained for selected sites in the positive (**e**) and negative (**f**) strands. Consensus were obtained using MIME software package.