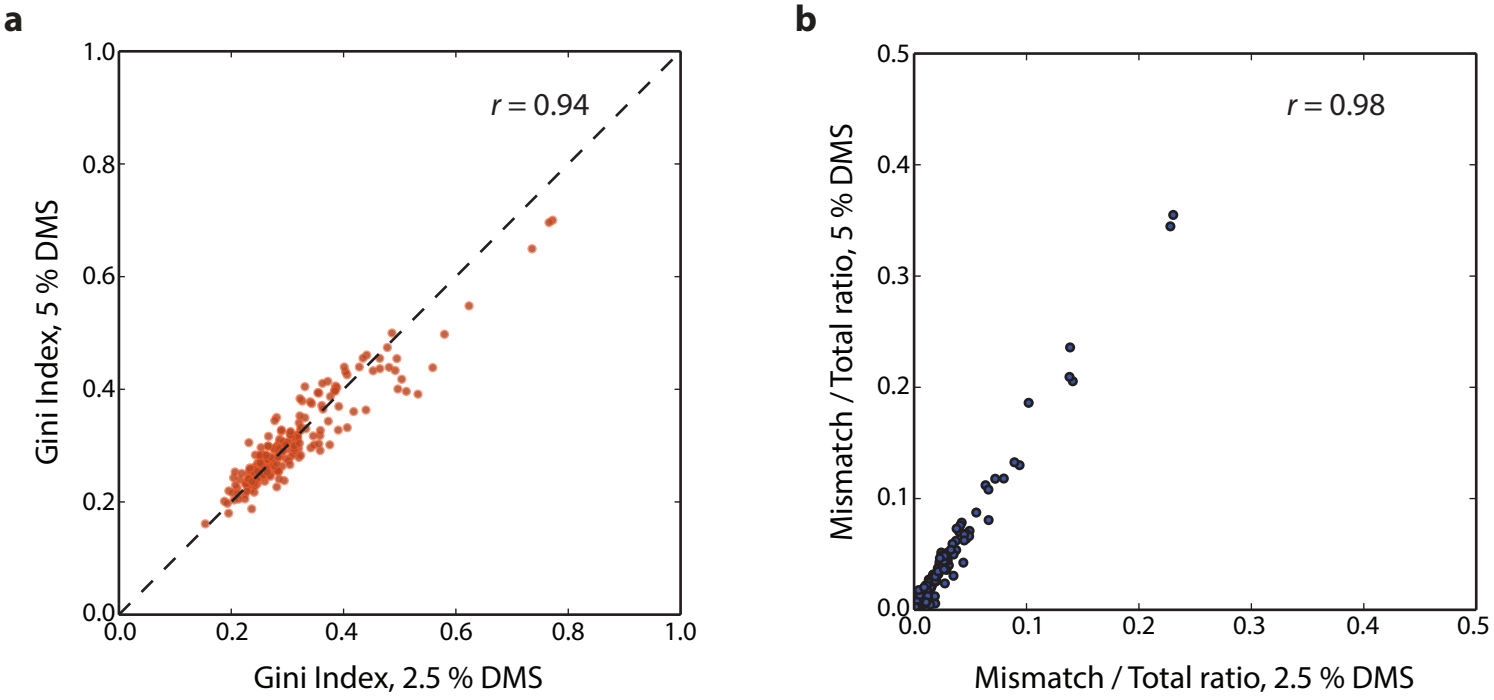
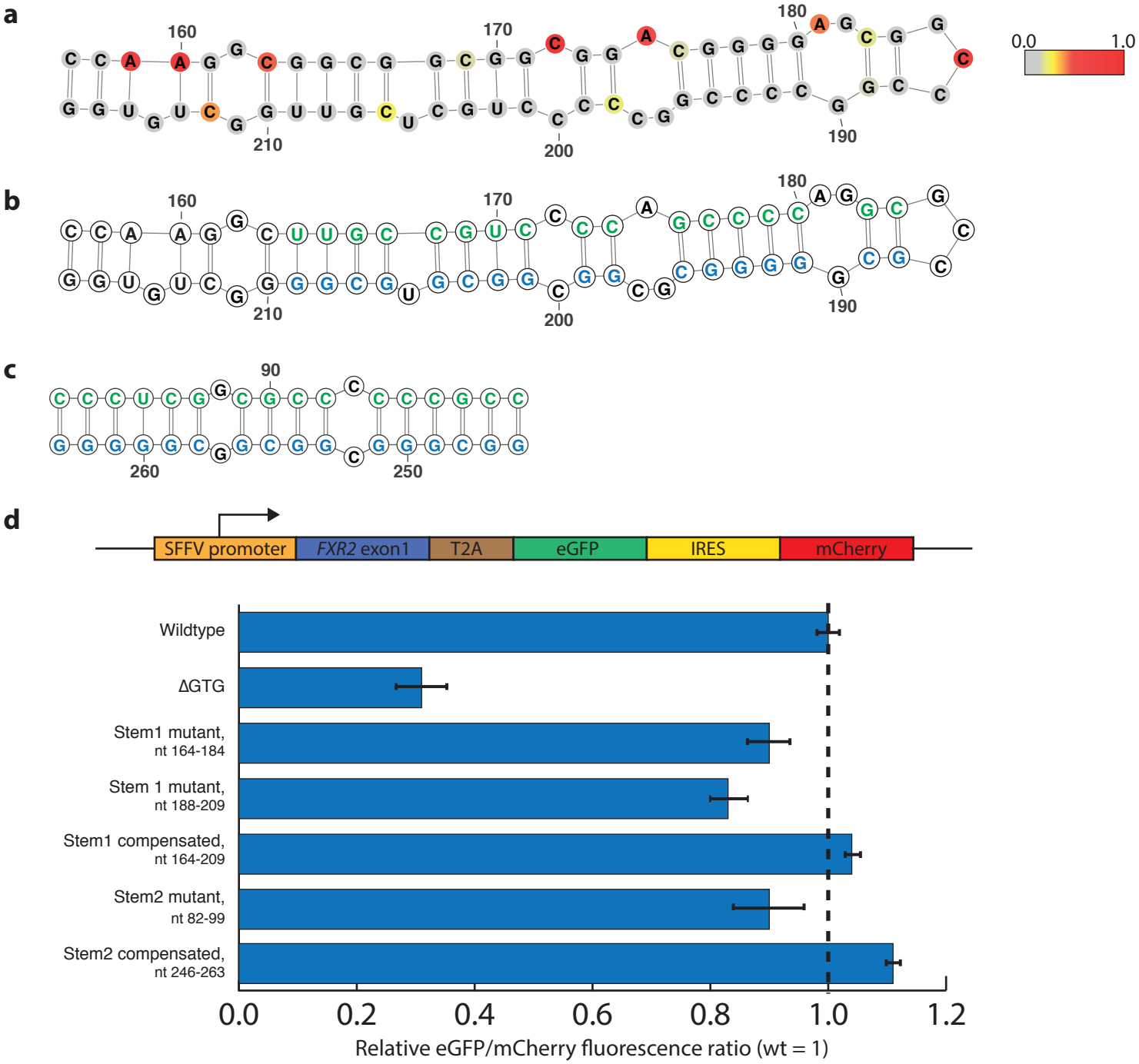
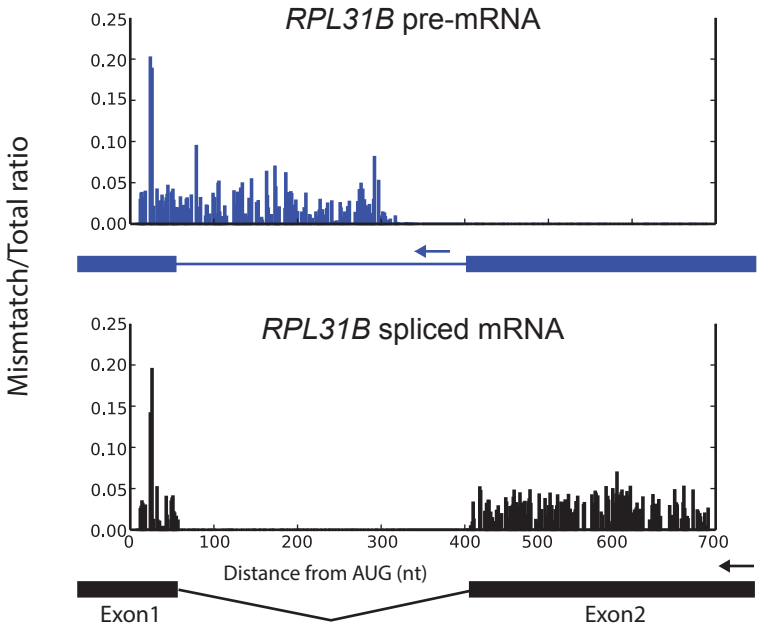


Supplementary Figure 1 | DMS-MaPseq data are highly reproducible at elevated DMS concentrations. **a,** Correlation of Gini index for 202 yeast mRNA regions with 15x coverage at 2.5 % or 5% v/v DMS concentrations reveals high reproducibility. **b,** Correlation of DMS-MaPseq signal for each of the 1800 nucleotides in the yeast 18S rRNA under varying DMS concentrations reveals high fidelity of RNA structure results at 5% v/v DMS. Pearson's *r* values are shown.



Supplementary Figure 10 | Fluorescent reporter constructs with RNA structure mutations confirm function of a highly stable structure in *FXR2* translation. **a, b,** Alternative stem 1 structure prediction from 0.06 constraint reactivity threshold from the human *FXR2* 5′ UTR and first exon, with nucleotides colored by DMS reactivity (a) or with the mutated sequence from (d) overlaid in green or blue text (b). DMS reactivity calculated as the ratiometric DMS signal normalized to the highest reactive base in the displayed region. **c,** Stem 2 structure model shown with mutated sequence used in (d) overlaid in green or blue text. **d,** *Top, FXR2* reporter construct design. The 5′ UTR and first exon of human *FXR2* ΔATG is fused to a T2A and in-frame eGFP lacking its initial AUG, such that mutations to the coding region of *FXR2* will not affect stability of the eGFP protein. To internally control for transfection and transcription efficiency, mCherry driven by an internal ribosome entry site was included downstream. *Bottom,* fluorescence measurements following transient transfection of *FXR2* reporter constructs into HEK 293T cells. The eGFP/mCherry ratio was calculated for transfection replications of each construct and scaled relative to the wildtype construct, which was set to 1.0. Error bars represent one standard deviation. This analysis reveals a drop in eGFP levels upon mutating the predicted *FXR2* structure and a full recovery of eGFP levels after compensatory mutation. Basal levels of protein expression in ΔGTG mutant likely reflect translation initiation at other NUG codons.

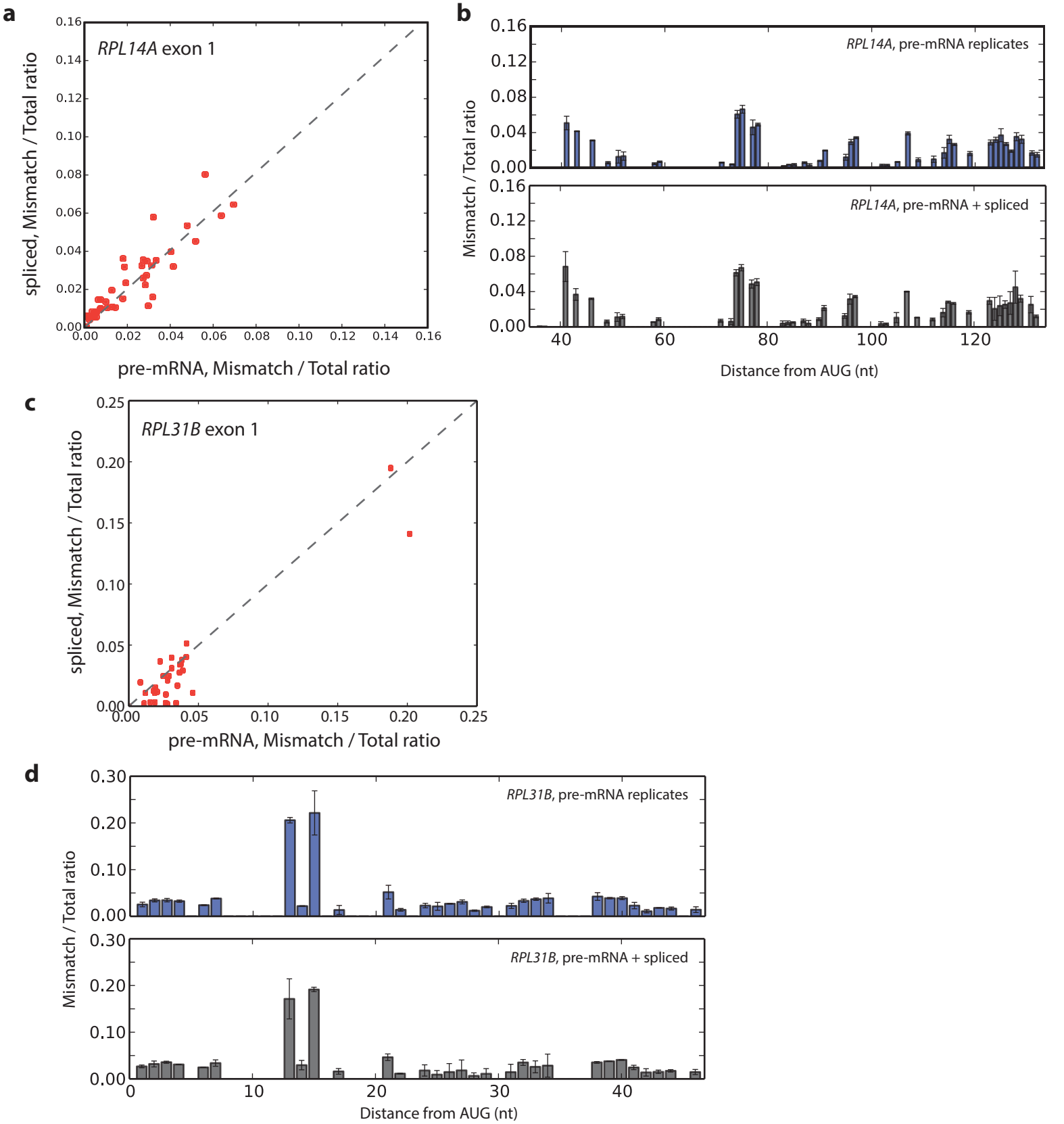
a



b

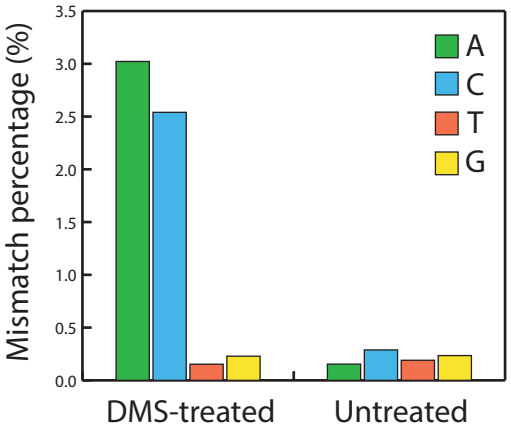
	r value	
Gene	**pre-mRNA replicates**	**pre- vs spliced**
RPL14A	0.96	0.88
RPL31B	0.98	0.95

Supplementary Figure 11 | RNA structure does not vary between the pre-mRNA and spliced mRNA isoforms of yeast ribosomal protein genes. **a,** Targeted DMS-MaPseq data specific for the yeast *RPL31B* pre-mRNA and spliced mRNA isoforms reveal minimal structure difference in the common exon1 sequence. Ratiometric DMS-MaPseq data is plotted with isoform-specific RT primer locations noted with arrows. **b,** Exon1 DMS-MaPseq structure signal correlation (Pearson's *r* value) across pre-mRNA and spliced mRNA isoforms and between isoform-specific replicates.

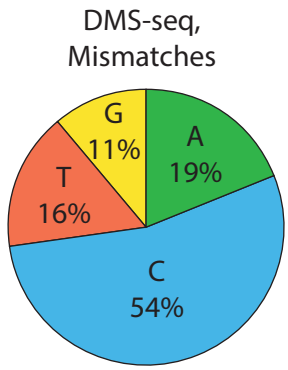


Supplementary Figure 12 | Investigation of RNA structure data in ribosomal protein exon 1 regions reveals nucleotide resolution consistency in structure signal between pre-mRNA and spliced isoforms. **a,** Ratiometric DMS-MaPseq data for the *RPL14A* exon 1 from its pre-mRNA versus spliced isoform reveals highly correlated signal for each A/C nucleotide. **b,** *Top*, Average ratiometric DMS-MaPseq signal from two *RPL14A* pre-mRNA technical replicates. *Bottom*, Average signal across *RPL14A* pre-mRNA and spliced isoforms reveals tight error distribution. **c,** Ratiometric DMS-MaPseq data for the *RPL31B* exon 1 from its pre-mRNA versus spliced isoform reveals highly correlated signal for each A/C nucleotide. **d,** *Top*, Average ratiometric DMS-MaPseq signal from two *RPL31B* pre-mRNA technical replicates. *Bottom*, Average signal across *RPL31B* pre-mRNA and spliced isoforms reveals tight error distribution. Error bars represent one standard deviation.

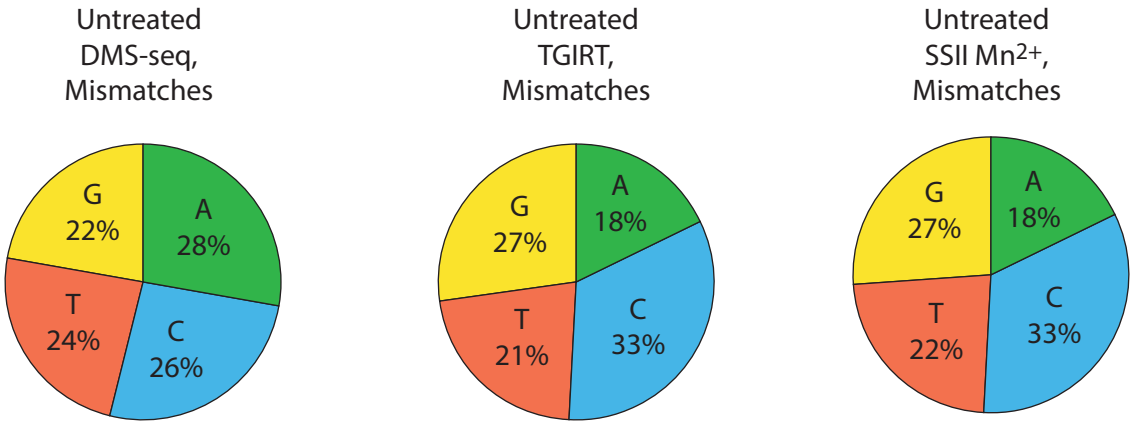
a



b



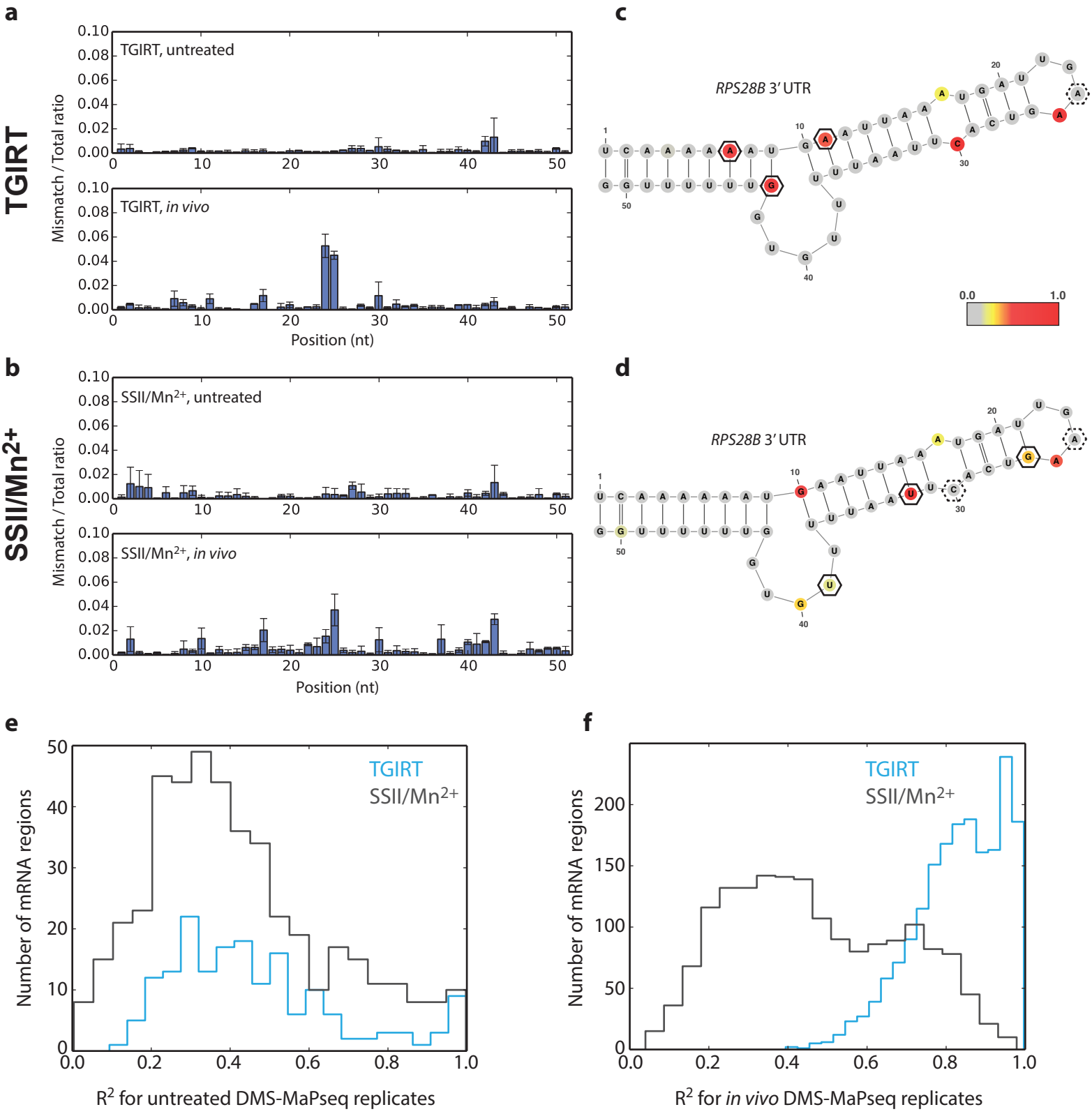
c



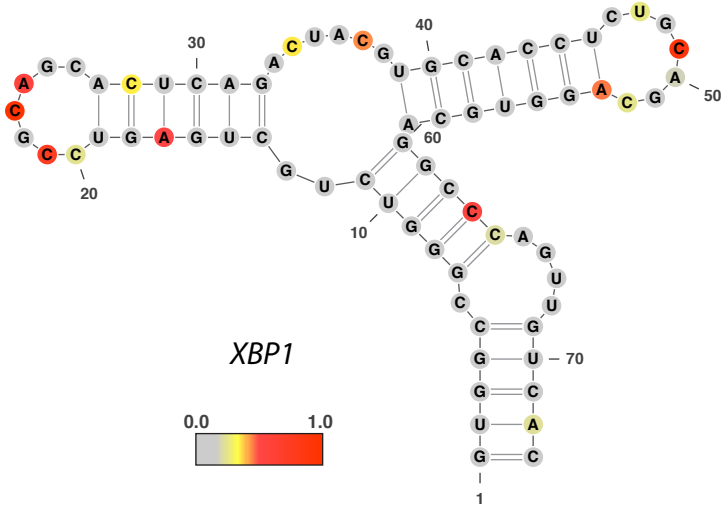
d

	Mismatch Frequency (%)		Indel Frequency (%)	
Library Type	**Untreated**	**DMS-treated**	**Untreated**	**DMS-treated**
TGIRT	0.21%	1.50%	0.02%	0.10%
SSII, Mn^{2+}	0.26%	1.09%	0.08%	0.44%
DMS-seq	0.33%	0.51%	0.03%	0.03%

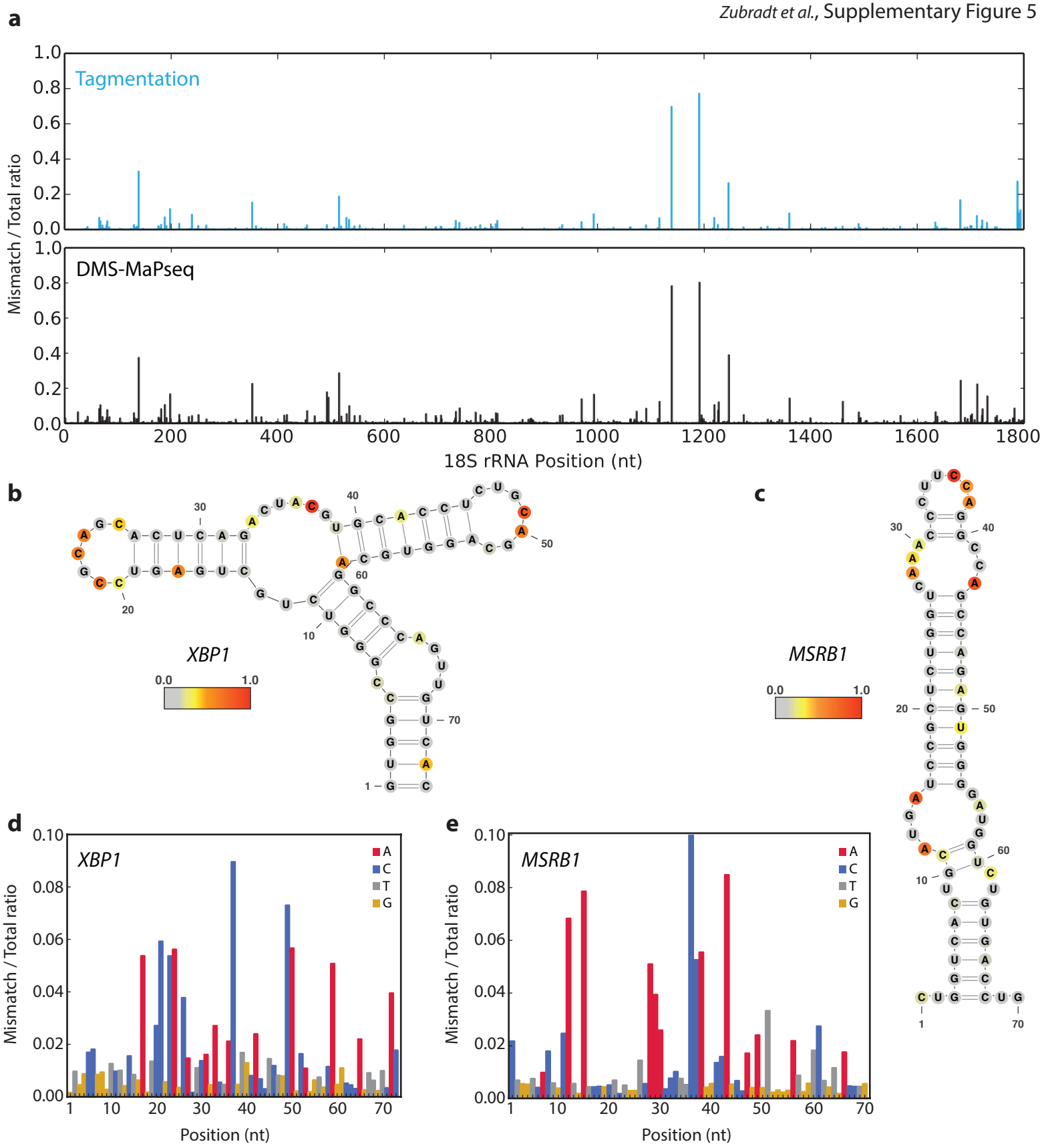
Supplementary Figure 2 | Mutations produced by reverse transcription on *in vivo* DMS-treated and untreated templates. **a,** Total mismatch percentage on each nucleotide from *in vivo* DMS-MaPseq with TGIRT on yeast mRNA. **b,** Nucleotide composition of mismatches in DMS-seq from Rouskin et al. for *in vivo* DMS-treated yeast mRNA, revealing a preference to generate mismatches on cytosines. **c,** Nucleotide composition of mismatches as detected by existing RNA structure probing approaches for untreated yeast mRNA, revealing no strong mismatch biases independent of DMS modification. **d,** Mutation frequency from DMS-treated and untreated yeast mRNA templates, derived from the same RNA source for TGIRT and SSII/Mn^{2+} data. Mutation frequency was calculated as the number of mismatches or indels detected via sequencing divided by the total number of bases sequenced.



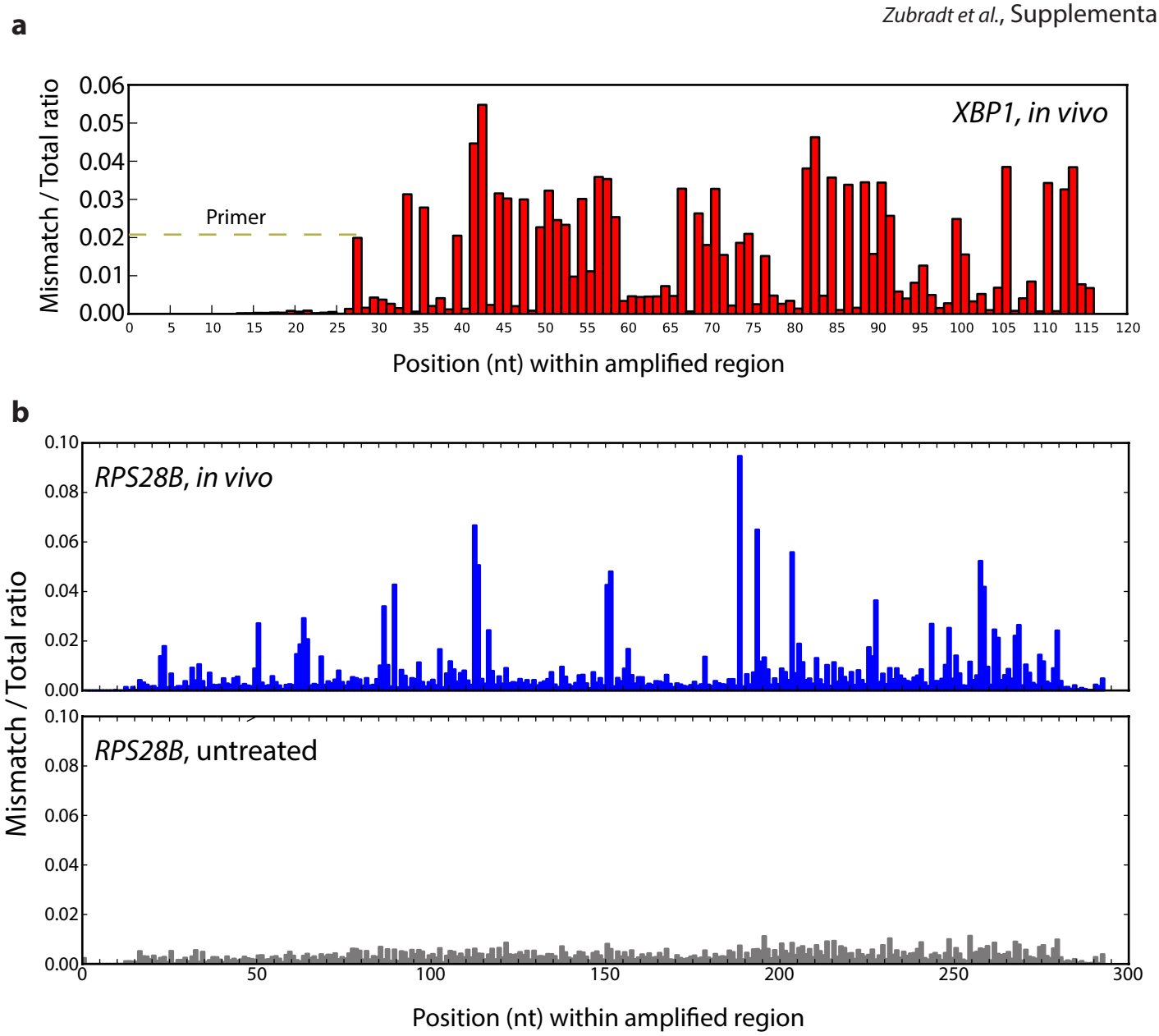
Supplementary Figure 3 | The variable background signal in DMS-MaPseq data can produce additional noise when used for correction. **a, b,** *RPS28B* 3′ UTR positive control region shown with average ratiometric DMS-MaPseq signal from two untreated and *in vivo* DMS-treated biological replicates prepared with TGIRT (a) or SSII/Mn^{2+}(c). Error bars represent one standard deviation. **c, d,** Background correction of TGIRT (c) or SSII/Mn^{2+} (d) RPS28B structure signal as ($S_{in\ vivo}$ - $S_{untreated}$) / $S_{denatured}$ produces false positives (solid lines) and false negatives (dashed lines) at certain positions. DMS reactivity calculated as the ratiometric DMS signal per position normalized to the highest number of reads in displayed region, which is set to 1.0. **e, f,** Histogram of R^2 values for DMS-MaPseq data from yeast mRNA regions for untreated (e) and *in vivo* treated (f) genome-wide biological replicates prepared with TGIRT or SSII/Mn^{2+} reveals high variability in background signal.



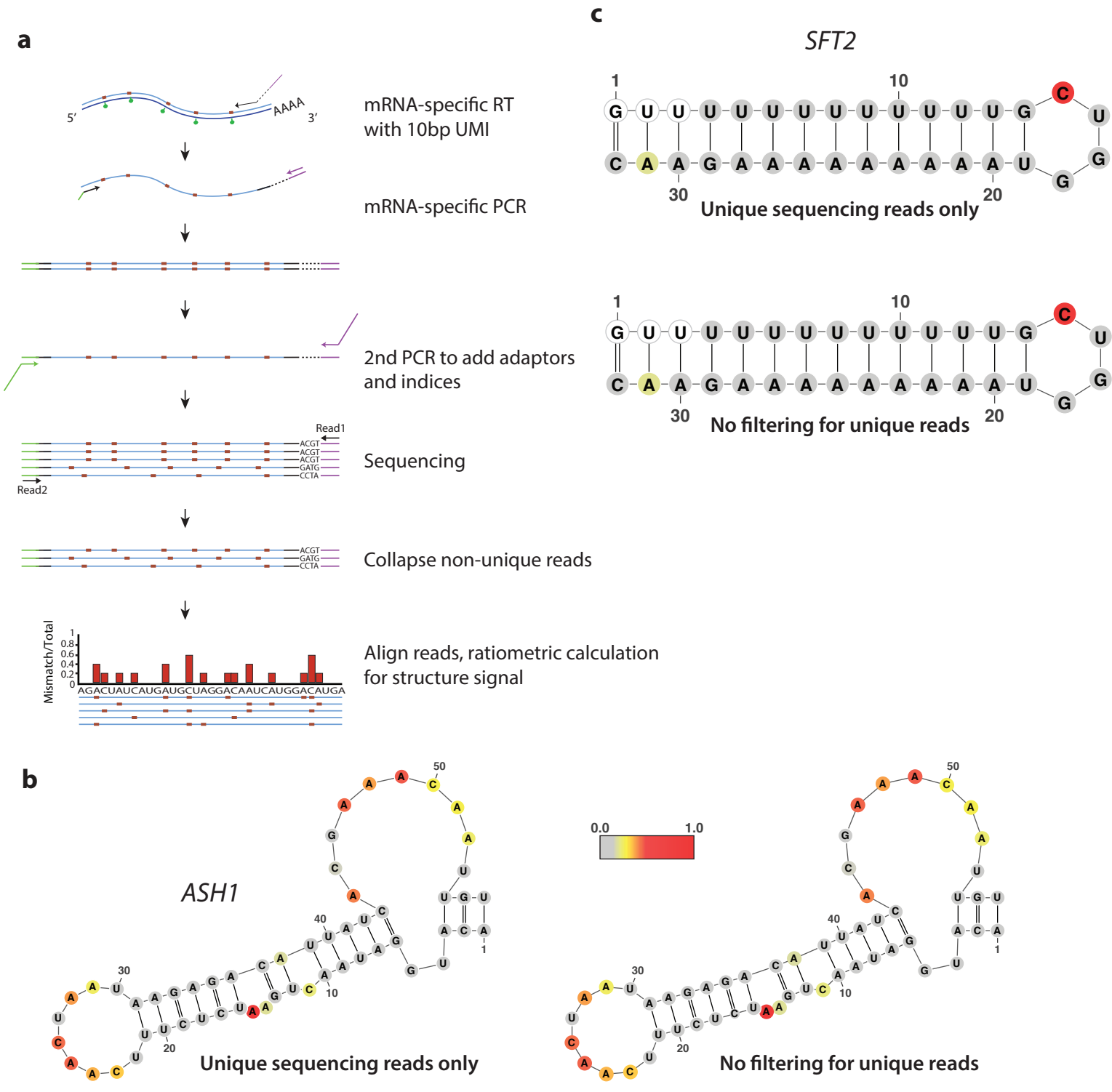
Supplementary Figure 4 | *XBP1* mRNA positive control structure with nucleotides colored by DMS reactivity from *in vivo* genome-wide DMS-MaPseq in HEK 293Ts. Note that the sequencing coverage for this region (~10x) is lower than suggested for high data reproducibility (~20x). DMS reactivity calculated as the ratiometric DMS signal per position normalized to the highest number of reads in displayed region, which is set to 1.0.



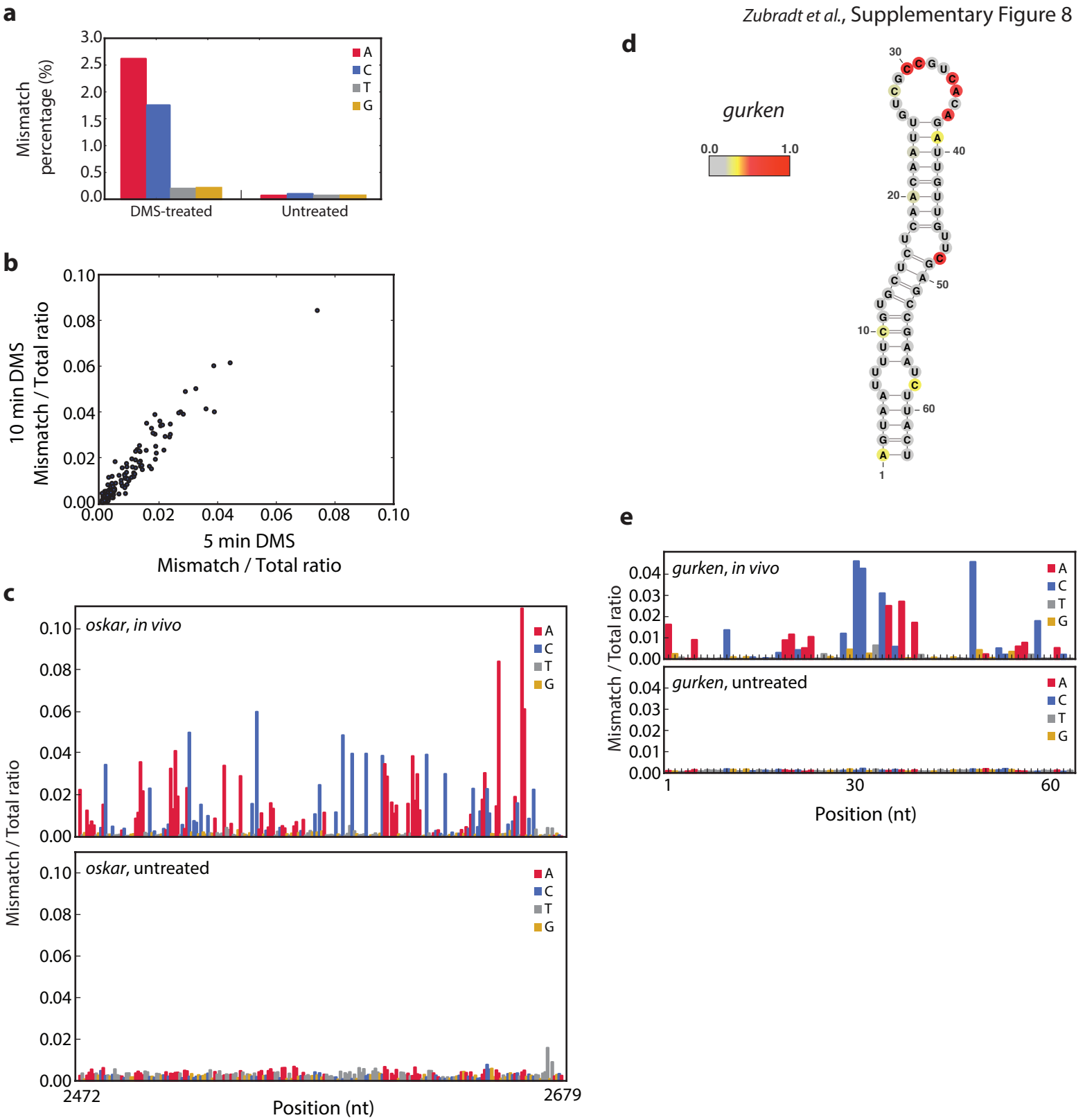
Supplementary Figure 5 | Target-specific DMS-MaPseq data for positive control RNA structures *in vivo*. **a,** Raw ratiometric data from yeast 18S rRNA prepared with the target-specific tagmentation or genome-wide DMS-MaPseq approach. Data is the same as used for the ROC curve in Figure 4d. **b, c,** *XBP1* and *MSRB1* mRNA positive control structures with nucleotides colored by DMS reactivity from target-specific DMS-MaPseq in HEK 293Ts. DMS reactivity calculated as the ratiometric DMS signal per position normalized to the highest number of reads in displayed region, which is set to 1.0. **d, e,** Raw ratiometric data for *XBP1* and *MSRB1* positive control regions in (b) and (c), respectively. *XBP1* region spans nt 520-592 from the NCBI NM_005080.3 transcript, and the *MSRB1* region spans nt 966-1036 from the NCBI NM_016332.2 transcript.



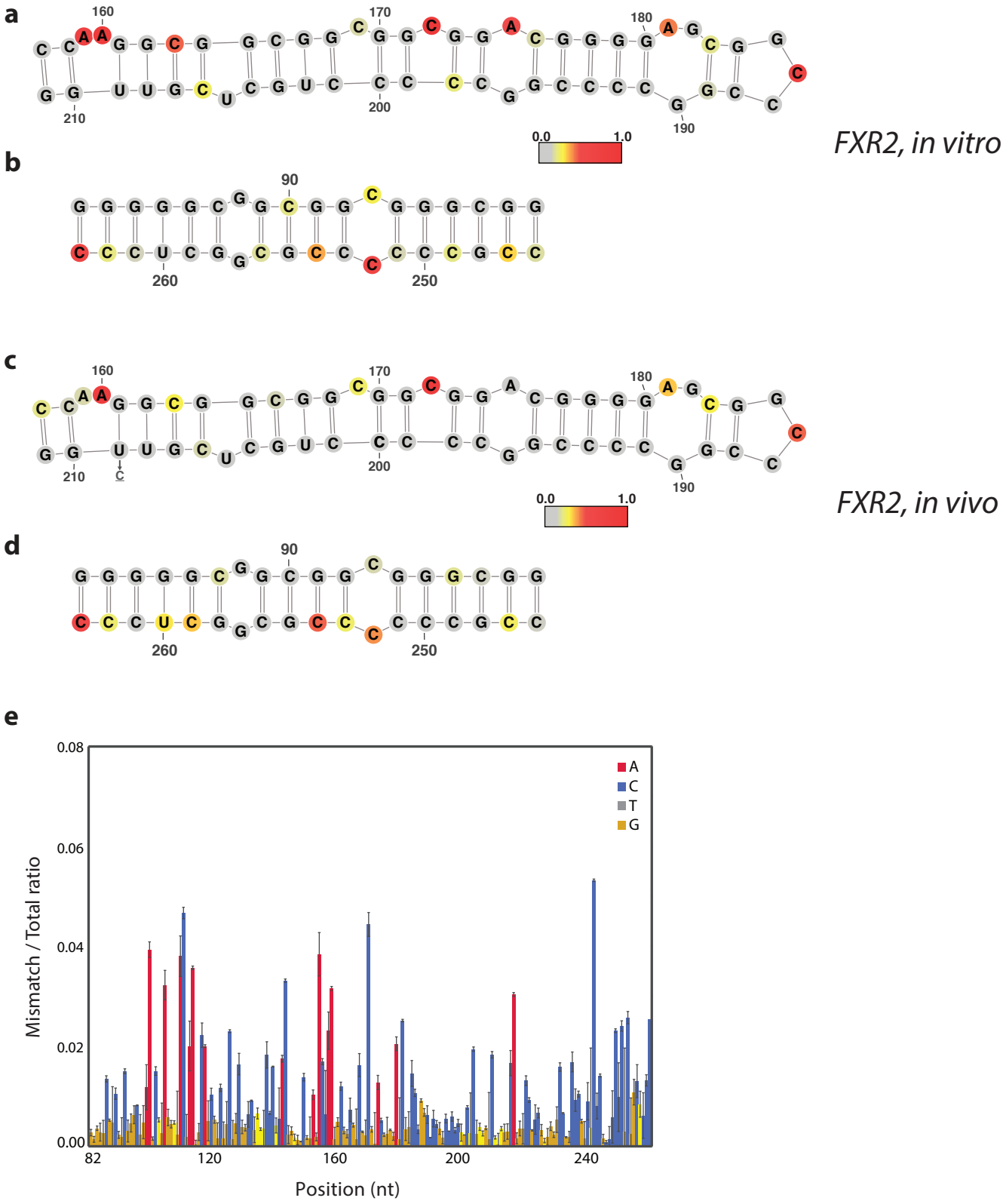
Supplementary Figure 6 | Target-specific DMS-MaPseq data show no signal drop-off and low background signal. **a,** Raw ratiometric data from HEK 293Ts for an *in vivo* DMS-treated *XBP1* mRNA region reveals no loss of signal across amplicon until the sequence bound by the PCR primer. (Position 1 corresponds nt 340 in the NCBI NM_005080.3 transcript annotation.) **b,** Ratiometric data plotted for *RPS28B* yeast mRNA positive control region for *in vivo* DMS-treated or untreated RNA reveals low background relative to *in vivo* signal. *In vivo* data is the same used for the structure model overlay in Figure 4f. Position 1 corresponds to chr XII, position 673546.



Supplementary Figure 7 | Targeted amplification of low-abundance RNA targets using a unique molecular index. **a,** Schematic for targeted RNA structure probing via gene-specific RT-PCR using a unique molecular index (UMI) on the RT primer. Using a gene-specific RT primer with a 5′ overhang comprised of an N_{10} random index and defined PCR primer binding site, each cDNA is labeled with a UMI. After gene-specific PCR amplification and limited-cycle second PCR to add sequencing adaptors and indexes, the PCR amplicon is sequenced on a MiSeq for a read length specified by the size of the region of interest. **b, c,** Yeast *ASH1* (b) and *SFT2* (c) mRNA positive control structures from target-specific UMI approach with nucleotides colored by DMS reactivity *in vivo*. Data are presented after collapsing to unique reads based on UMI and internal DMS-induced mismatches and are presented without collapsing, with nearly identical results. DMS reactivity calculated as the ratiometric DMS signal per position normalized to the highest number of reads in displayed region, which is set to 1.0. Uncolored nucleotides had no data collected.



Supplementary Figure 8 | Targeted DMS-MaPseq in *D. melanogaster* ovaries. **a,** Total mismatch percentage observed on each nucleotide from *in vivo* DMS-treated and untreated *D. melanogaster* RNA. **b,** Correlation of DMS-MaPseq signal for each nucleotide in the *oskar* PCR amplicon under varied DMS treatment times reveals high fidelty of RNA structure results under both conditions. **c,** Raw ratiometric data for the *oskar* positive control structure in Figure 5a reveals low background signal in an untreated versus *in vivo* treated sample at 50% v/v DMS for 10min. Nucleotide positions are noted relative to transcription start site of FlyBase transcript annotation FBtr0081956. **d,** *gurken* mRNA positive control structure with nucleotides colored by *in vivo* DMS reactivity, after 25% v/v DMS treatment for 5 min. DMS reactivity calculated as the ratiometric DMS signal per position normalized to the highest number of reads in displayed region, which is set to 1.0. **e,** Raw ratiometric data for the gurken positive control region shown in (d) reveals low background signal in an untreated versus *in vivo* treated sample.



Supplementary Figure 9 | *In vitro* and *in vivo* DMS-MaPseq reactivity of *FXR2* stem1 and stem2 regions within the human *FXR2* 5′ UTR and first exon. **a, b,** Stem 1 (a) and stem 2 (b) structure predictions from 0.04 constraint reactivity threshold, with nucleotides colored by DMS reactivity with nucleotides colored by *in vitro* DMS reactivity. **c, d,** Stem 1 (c) and stem 2 (d) structure predictions with nucleotides colored by *in vivo* DMS reactivity. DMS reactivity calculated as the ratiometric DMS signal normalized to the highest reactive base in the 82-263nt region. **e,** Raw ratiometric signal from the *FXR2* 5′ UTR and first exon *in vivo*, displaying the mean between two technical replcates and error bars representing one standard deviation. Position 1 corresponds to chrXVII: 7614897.

Table S1.

Primers used in this study.

name	purpose	sequence (5' to 3')
Linker 2	3' Cloning adaptor for RNA footprints	5rApp/CACTCGGGCACCAAGGA/3ddC
oCJ200-link2	Primer for reverse transcription of sequencing libraries	5'/5phos/GATCGTCGGACTGTAGAACTCTGAACCTGTCG/iSp18/CAAGCAGAAGACGGCATACGAGATTCCTTGGTGCCCGAGTG
oNTI231	Amplification of sequencing libraries, paired with indexing primer	caagcagaagacggcatacga
Indexing primer with 6bp TruSeq index		aatgatacggcgaccaccgagatctacacgatcggaagagcacacgtctgaactccagtcacNNNNNNcgacaggttcagagttc
oNTI202	Read1 sequencing primer	CGACAGGTTCAGAGTTCTACAGTCCGACGATC
oMZ282	Reverse transcription and reverse primer (1st PCR) for targeted amplification with a unique molecular index	GCAGCGACAGGTTCAGAGTTCTACAGTCCGACGATC – $(N)_{10}$ – gene-specific primer
oMZ283	Forward primer (1st PCR) for targeted amplification with a unique molecular index	CTGAACCGCTCTTCCGATCT–gene-specific primer
oMZ284	Forward primer (2nd PCR) for targeted amplification with a unique molecular index	CAAGCAGAAGACGGCATACGAGACGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
oMZ285	*ASH1* RT primer	oMZ282 + TTTGTAGTTTATTTAGCACAGACAAGGAGAGAAATGT
oMZ286	*ASH1* Fwd PCR primer	oMZ283 + TGGAATAGACAAAGAATTCGTCCCCGAACGCAACT
oMZ287	*SFT2* RT primer	oMZ282 + CGAGTTCTGCTGTCTTGTTTGATTCCATCG
oMZ288	*SFT2* Fwd PCR primer	oMZ283 + GAGTAGGACTTAGATTACCTGTATTGTCTGCAGTTGCGTT
oMZ289	*RPL14A* RT/Rev PCR primer, intron	ACGATGGAAGGCATGGTTTAATATTTGAGGAAACATGG
oMZ290	*RPL14A* RT/Rev PCR primer, exon	AGCCTTAGCCAAAGCCTTCTTGACAGTGTA
oMZ291	*RPL14A* Fwd PCR primer	TGTCCACCGATTCTATTGTCAAGGCTTCTAACTGG
oMZ292	*RPL31B* RT/Rev PCR primer, intron	AAGGTGGAACTAAAGCATCACGCCAAAAACATCG
oMZ293	*RPL31B* RT/Rev PCR primer, exon	CGTACCCCGAAAGCAGCTCTGTTTGTGTAAT
oMZ294	*RPL31B* Fwd PCR primer	TGCACGAGCAGATAATCCAAAGTACTTGAAAATGGCC

Table S2.

Plasmids used in this study.

name	description
pJW1643	pLeGO-ic2 pSFFV-FXR2exon1(wt)-3xFLAG-TEV-HA-T2A-eGFP-IRES-mCherry
pJW1644	pLeGO-ic2 pSFFV-FXR2exon1(ΔGTG)-3xFLAG-TEV-HA-T2A-eGFP-IRES-mCherry
pJW1645	pLeGO-ic2 pSFFV-FXR2exon1(mut, 164-184)-3xFLAG-TEV-HA-T2A-eGFP-IRES-mCherry
pJW1646	pLeGO-ic2 pSFFV-FXR2exon1(mut, 188-217)-3xFLAG-TEV-HA-T2A-eGFP-IRES-mCherry
pJW1647	pLeGO-ic2 pSFFV-FXR2exon1(comp, 164-217)-3xFLAG-TEV-HA-T2A-eGFP-IRES-mCherry
pJW1648	pLeGO-ic2 pSFFV-FXR2exon1(mut, 82-99)-3xFLAG-TEV-HA-T2A-eGFP-IRES-mCherry
pJW1649	pLeGO-ic2 pSFFV-FXR2exon1(comp, 246-263)-3xFLAG-TEV-HA-T2A-eGFP-IRES-mCherry