nature chemical biology

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41589-018-0114-4

In the format provided by the authors and unedited.

ykkC riboswitches employ an add-on helix to adjust specificity for polyanionic ligands

Alla Peselis and Alexander Serganov*

Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, NY, USA.
*e-mail: alexander.serganov@nyumc.org

Supplementary Table 1 Data collection and refinement statistics for the *S. lipocalidus* PRPP riboswitch structures

	$[Ir(NH_3)_6]^{3+}$	Native	$MnCl_2$	Tl-acetate	apo
Data collection					
Space group	*C*2	*C*2	*C*2	*C*2	$P6_522$
Cell dimensions					
a, *b*, *c* (Å)	83.52, 92.68, 74.49	83.57, 90.76, 74.17	85.32, 90.51, 73.48	85.22, 91.98, 75.06	63.77, 63.77, 376.49
α, β, γ (°)	90.0, 121.0, 90.0	90.0, 122.0, 90.0	90.0, 122.5, 90.0	90.0, 122.8, 90.0	90.0, 90.0, 120.0
Resolution (Å)*	20.00-2.65 (2.74-2.65)	20.00-2.90 (3.00-2.90)	20.00-2.80 (2.90-2.80)	20.00-2.90 (3.00-2.90)	20.00-2.90 (2.95-2.90)
R_{merge}	0.119 (0.563)	0.156 (0.632)	0.082 (0.518)	0.098 (0.826)	0.154 (1.554)
R_{pim}	0.073 (0.349)	0.081 (0.298)	0.040 (0.269)	0.050 (0.454)	0.051 (0.388)
$I/\sigma I$	27.2 (2.1)	15.7 (2.7)	28.1 (1.9)	28.2 (1.6)	27.6 (1.1)
$CC_{1/2}$†	0.808	0.881	0.987	0.726	0.834
Completeness (%)	96.5 (90.0)	92.4 (96.9)	98.4 (91.4)	97.9 (93.1)	93.8 (94.5)
Redundancy	3.4 (3.0)	5.1 (5.4)	5.1 (4.0)	4.8 (3.8)	9.5 (11.5)
Refinement					
Resolution (Å)	20.00-2.65	20.00-2.90	20.00-2.80	20.00-2.90	20.00-2.89
No. reflections	13,280	9,366	11,318	10,696	10,251
R_{work}/ R_{free}	19.7/ 22.2	22.5/24.3	22.6/25.4	20.7/23.5	21.2/27.3
No. atoms					
RNA	2,301	2,310	2,310	2,310	2,310
Ligand	22	22	22	22	----
Ions	136	6	17	4	3
Water	6	3	8	1	----
B-factors					
RNA	63.04	64.69	86.18	94.62	130.09
Ligand	51.12	53.18	63.02	74.30	----
Ion	77.06	61.31	92.51	86.35	
Water	59.04	42.80	65.58	92.00	99.63
R.m.s deviations					----
Bond length (Å)	0.002	0.005	0.002	0.010	0.006
Bond angles (°)	0.551	1.237	0.580	1.400	1.241

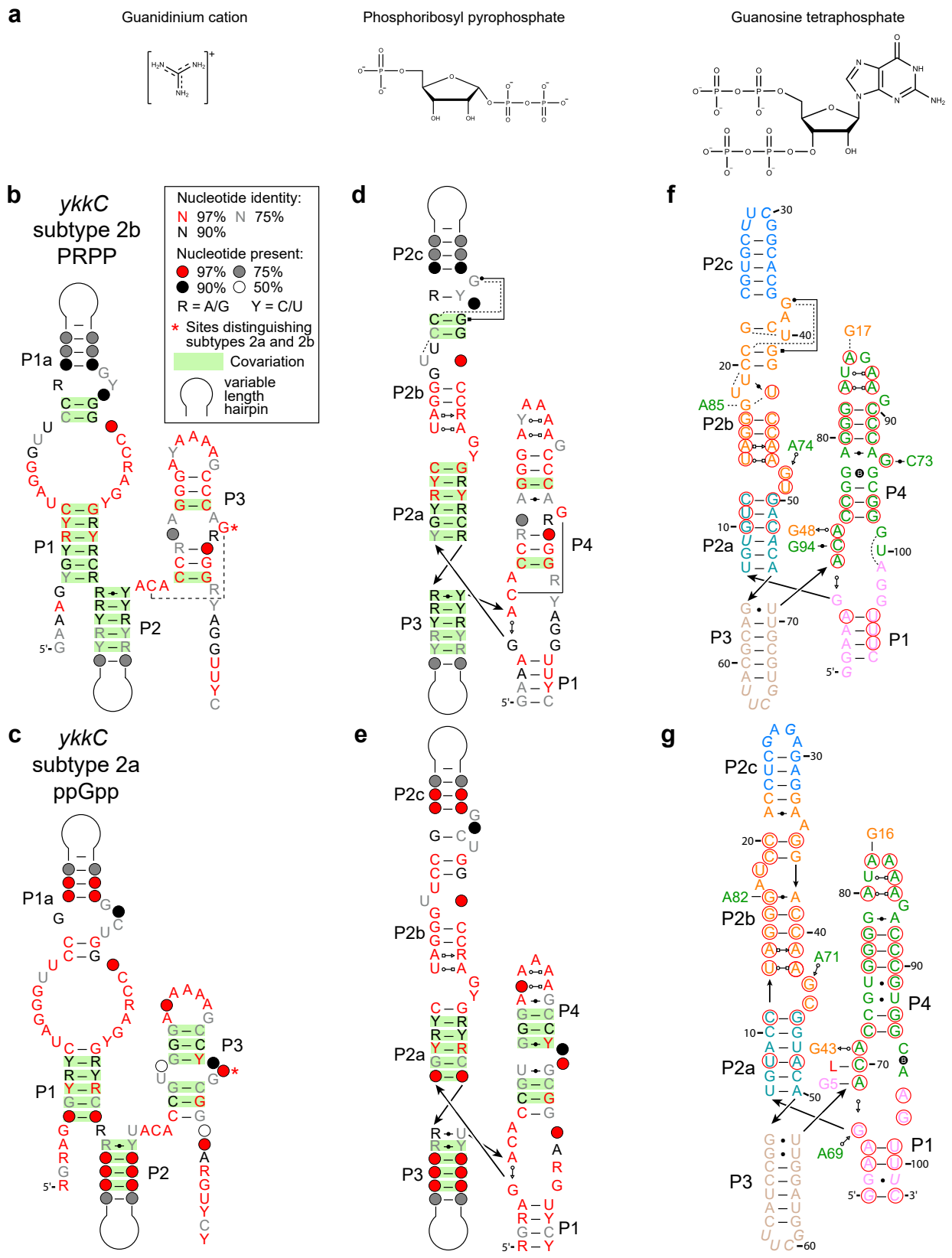
* Values in parentheses are for highest-resolution shell.

†$CC_{1/2}$ values shown for high resolution shell

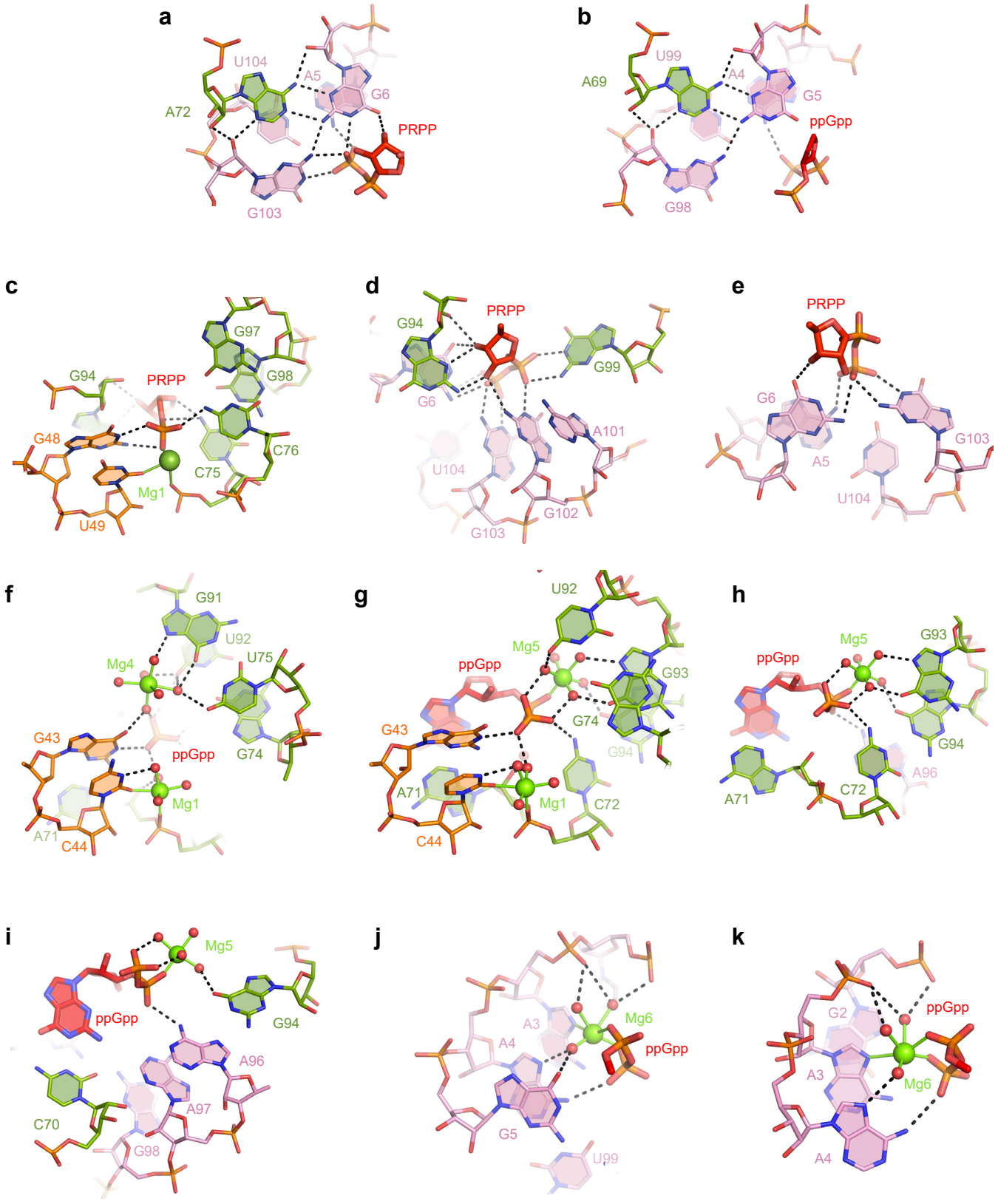
Supplementary Table 2 Data collection and refinement statistics for the *S. acidophilus* ppGpp riboswitch structures

	Native	$MnCl_2$	Tl-acetate
Data collection			
Space group	$P2_1$	$P2_1$	$C2$
Cell dimensions			
a, *b*, *c* (Å)	78.16, 48.92, 78.72	78.74, 49.59, 79.36	100.18, 51.82, 77.86
α, β, γ (°)	90.0, 97.3, 90.0	90.0, 97.3, 90.0	90.0, 128.1, 90.0
Resolution (Å)	20.00-2.20 (2.24-2.20)	30.00-2.65 (2.74-2.65)	30.00-2.70 (2.80-2.70)
R_{merge}	0.128 (0.886)	0.107 (0.455)	0.095 (0.237)
R_{pim}	0.069 (0.633)	0.049 (0.243)	0.049 (0.129)
$I/\sigma I$	15.8 (1.0)	22.5 (1.7)	21.4 (5.0)
$CC_{1/2}$†	0.540	0.876	0.961
Completeness (%)	96.1 (91.1)	97.4 (91.0)	98.4 (94.0)
Redundancy	4.0 (2.5)	5.3 (3.7)	4.6 (3.9)
Refinement			
Resolution (Å)	20.00-2.20	30.00-2.65	30.00-2.70
No. reflections	27,688	16,626	8,587
R_{work}/ R_{free}	22.8/25.9	21.7/27.2	19.9/24.3
No. atoms			
RNA	4,414	4,414	2,203
Ligand	72	72	36
Ions	32	27	14
Water	168	43	8
B-factors			
RNA	46.98	52.59	22.49
Ligand	38.14	37.27	26.53
Ions	47.46	48.41	25.03
Water	33.37	29.28	18.85
R.m.s deviations			
Bond lengths (Å)	0.004	0.005	0.003
Bond angles (°)	1.103	1.182	0.782

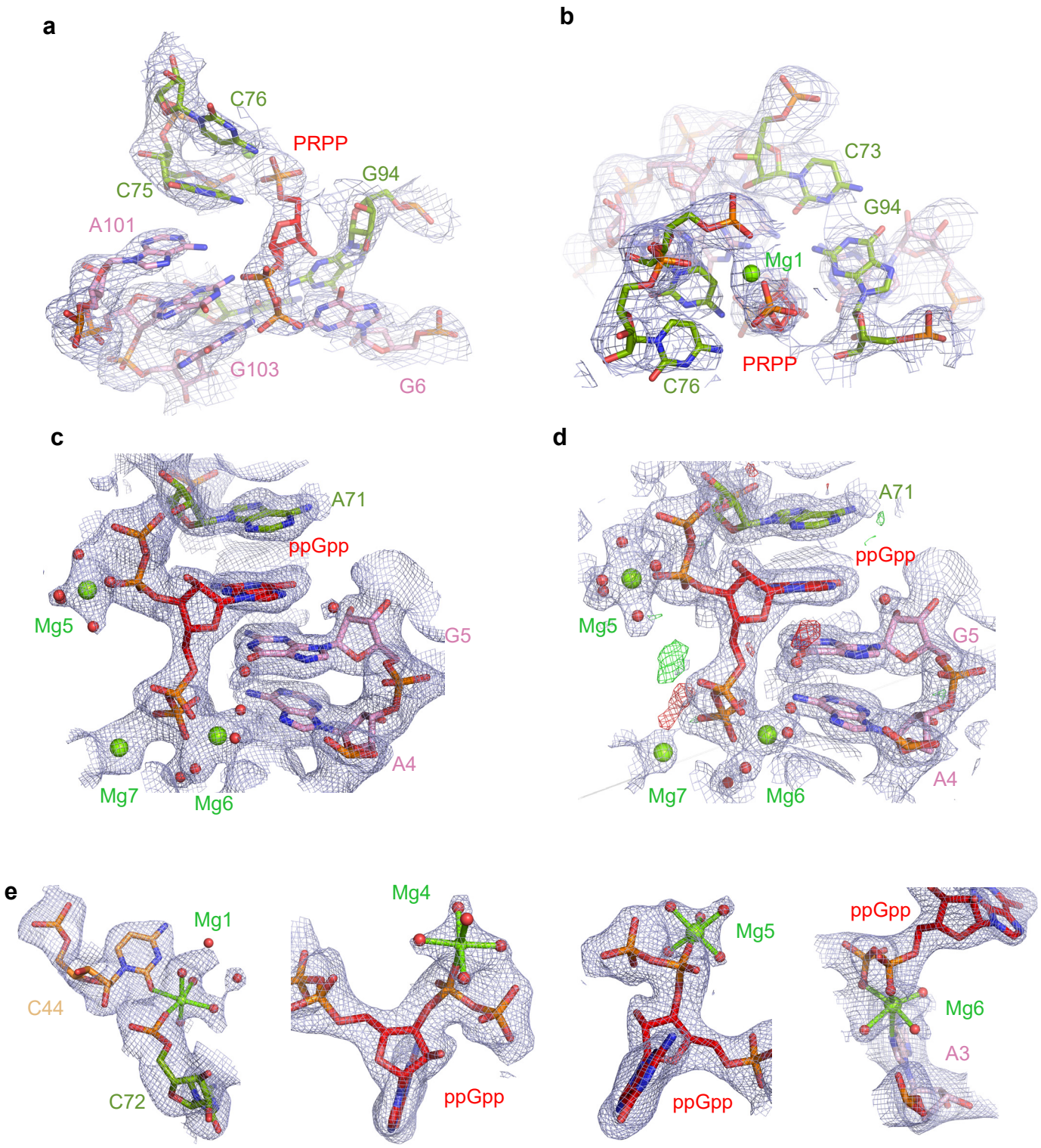
*Values in parentheses are for highest-resolution shell.
†$CC_{1/2}$ values shown for high resolution shell



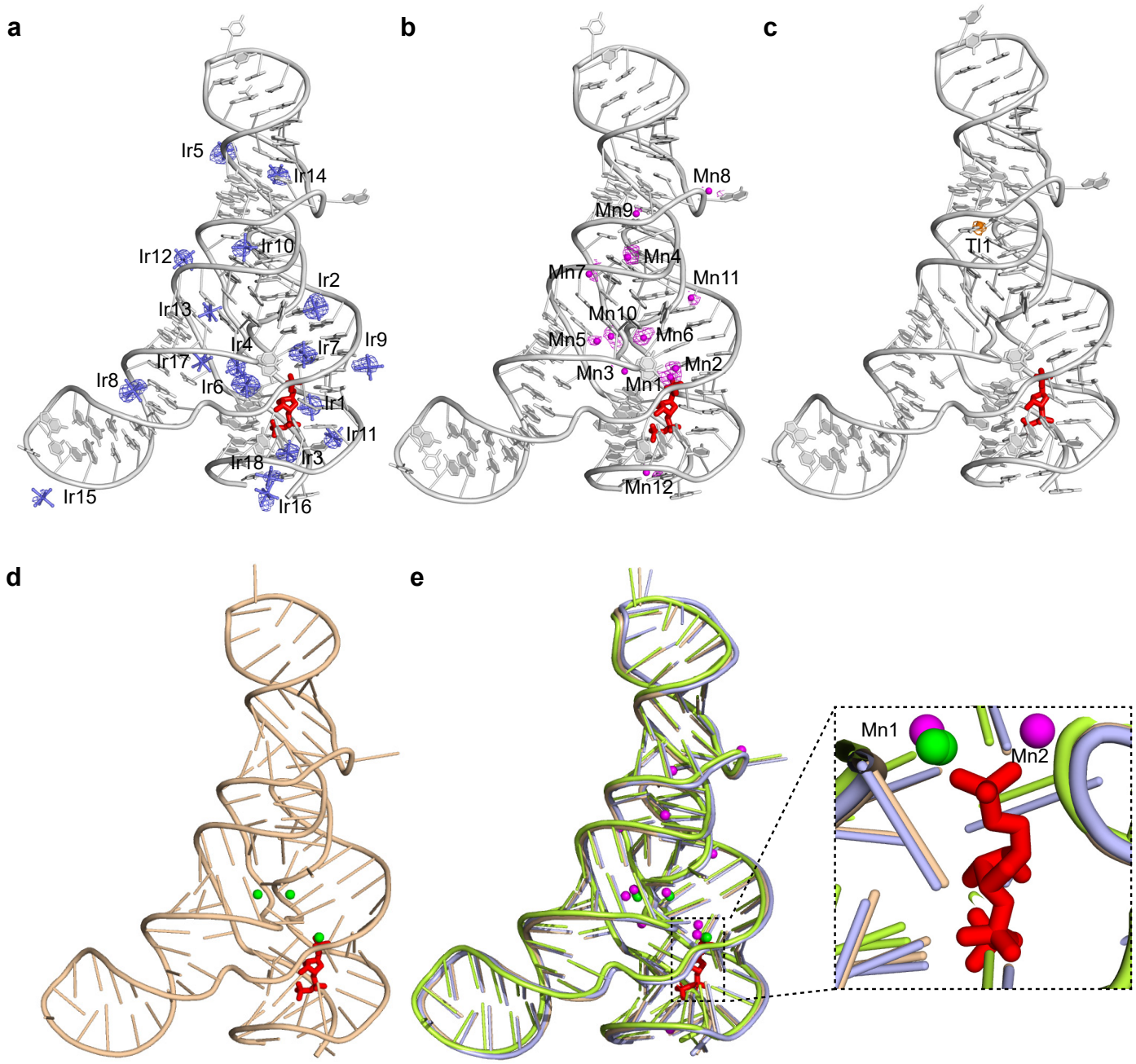
Supplementary Figure 1. Consensus secondary structure of the *ykkC* subclass 2a and 2b riboswitches. (**a**) The chemical structures of the guanidinium cation, phosphoribosyl pyrophosphate (PRPP), and guanosine tetraphosphate (ppGpp). (**b**) Phylogeny-based consensus secondary structure of the *ykkC* subclass 2b riboswitch responsive to PRPP[6]. (**c**) The consensus motif of the *ykkC* subclass 2a ppGpp riboswitch[7]. (**d-e**) Consensus motifs updated on the basis of the crystal structures of the PRPP (d) and ppGpp (e) riboswitches. Base-pairing notation is according to Leontis and Westhof (*RNA* 7, 499-512 (2001)). **(f-g)** The secondary structures of the PRPP (f) and ppGpp (g) riboswitches from Figure 1 annotated with the Leontis and Westhof pairing nomenclature. Other symbols as in Figure 1.



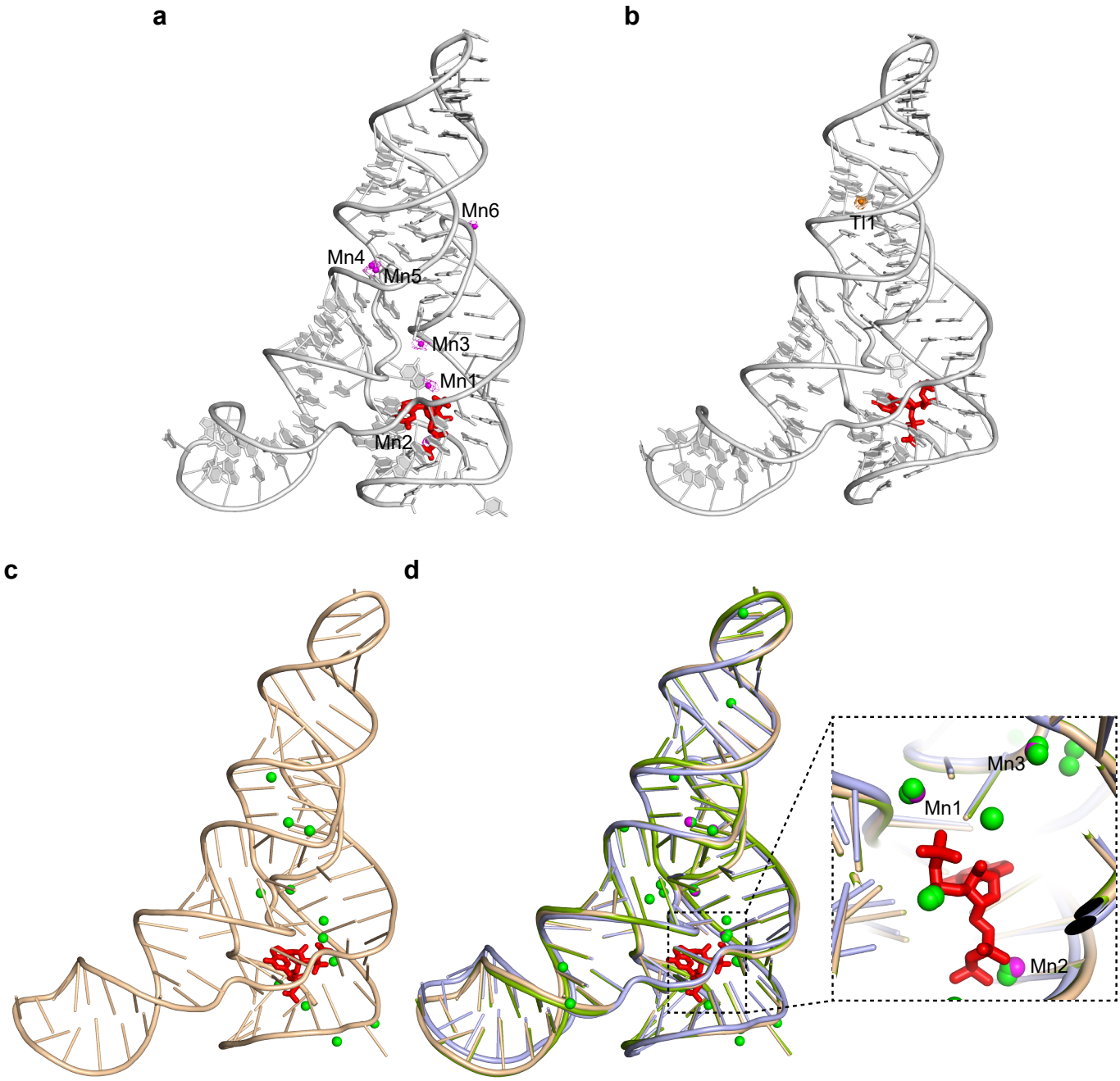
Supplementary Figure 2. Molecular details of ligand recognition by the PRPP and ppGpp riboswitch RNAs. Molecules are depicted in atomic colors and stick representation. Mg^{2+} cations and water molecules are shown by green and red spheres, respectively. Putative hydrogen bonds are shown in black dashed lines. Mg^{2+} coordination bonds are shown in thin green sticks. (**a**,**b**) Junction-stabilizing triples in the PRPP (a) and ppGpp (b) riboswitches. (**c**-**e**) Zoomed-in views of the PRPP ligand binding site from top to bottom of the ligand. (**f**-**k**) Zoomed-in views of the ppGpp ligand binding site from top to bottom of the ligand.



Supplementary Figure 3. Electron density maps for the PRPP and ppGpp riboswitch structures. (**a-b**) Validation of the PRPP riboswitch structure. $2F_o$-F_c simulated annealing omit map, calculated by PHENIX[3], contoured at 1.0 σ level is shown with the refined structural model of the riboswitch. Views are centered on the ligand-binding sites. (**c**) $2F_o$-F_c simulated annealing omit map contoured at 1.0 σ level shown with the refined structural model of the ppGpp riboswitch structure. (**d**) Refined structural model of the ppGpp riboswitch shown with $2F_o$-F_c map, after refinement in Refmac[4], contoured at 1.0 σ level (blue) and F_o-F_c maps contoured at 3.0 σ (green) and -3.0 σ (red) levels. (**e**) $2F_o$-F_c simulated annealing omit map contoured at 1.0 σ level shown with the refined structural model of the ppGpp riboswitch structure. Views highlight octahedral geometry of the ppGpp-bound Mg^{2+} cations. Coordination bonds are depicted by green sticks.



Supplementary Figure 4. Metal cation binding sites in the PRPP riboswitch. (**a-c**) Refined PRPP riboswitch structure superposed with anomalous maps of $[Ir(NH_3)_6]^{3+}$ (a), Mn^{2+} (b), and Tl^{+} (c) shown in blue, magenta, and orange colors and contoured at 3.0, 2.5, and 3.0 σ level, respectively. Cations are depicted as spheres. (**d**) Native PRPP riboswitch structure with Mg^{2+} cations in green. **(e)** Superposition of native (beige), Mn^{2+} (green) and Tl^{+}-soaked (purple) structures. Mg^{2+} cations found in the native and Tl^{+}-soaked structures are indicated in green while Mn^{2+} cations found in the Mn^{2+}-soaked structure are colored in magenta.



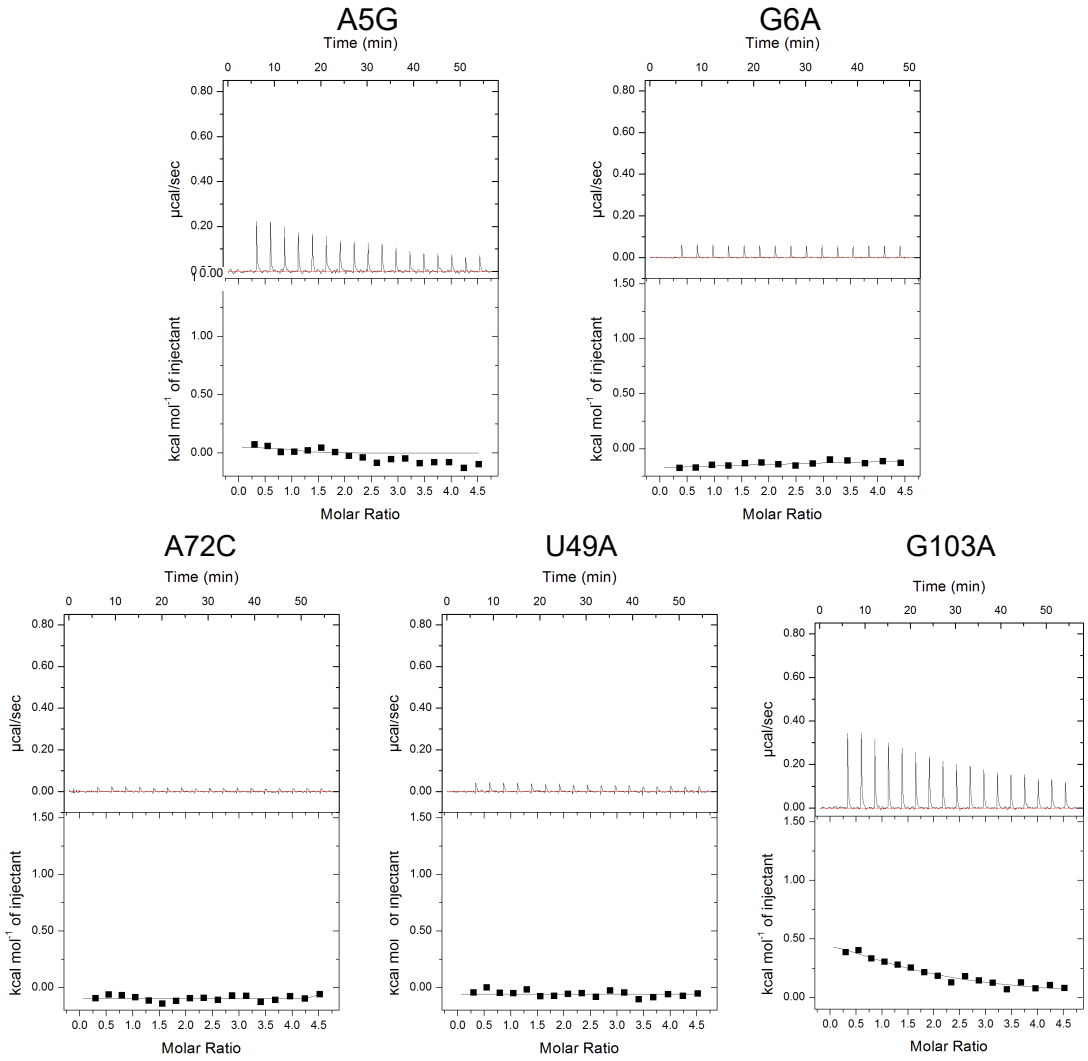
Supplementary Figure 5. Metal cation binding sites in the ppGpp riboswitch. (**a**-**b**) Refined ppGpp riboswitch structure superposed with anomalous maps of Mn^{2+} (a) and Tl^{+} (b) shown in magenta and orange colors and contoured at 2.0 and 2.5 σ level, respectively. Cations are depicted as spheres. Molecules "A" are shown for the structures in the $P2_1$ space group. (**c**) Native ppGpp riboswitch structure with Mg^{2+} cations in green. **(d)** Superposition of native (beige), Mn^{2+}-(green) and Tl^{+}-soaked (purple) structures. Mg^{2+} cations found in the native and Tl^{+}-soaked structures are indicated in green while Mn^{2+} cations found in the Mn^{2+}-soaked structure are colored in magenta.

a

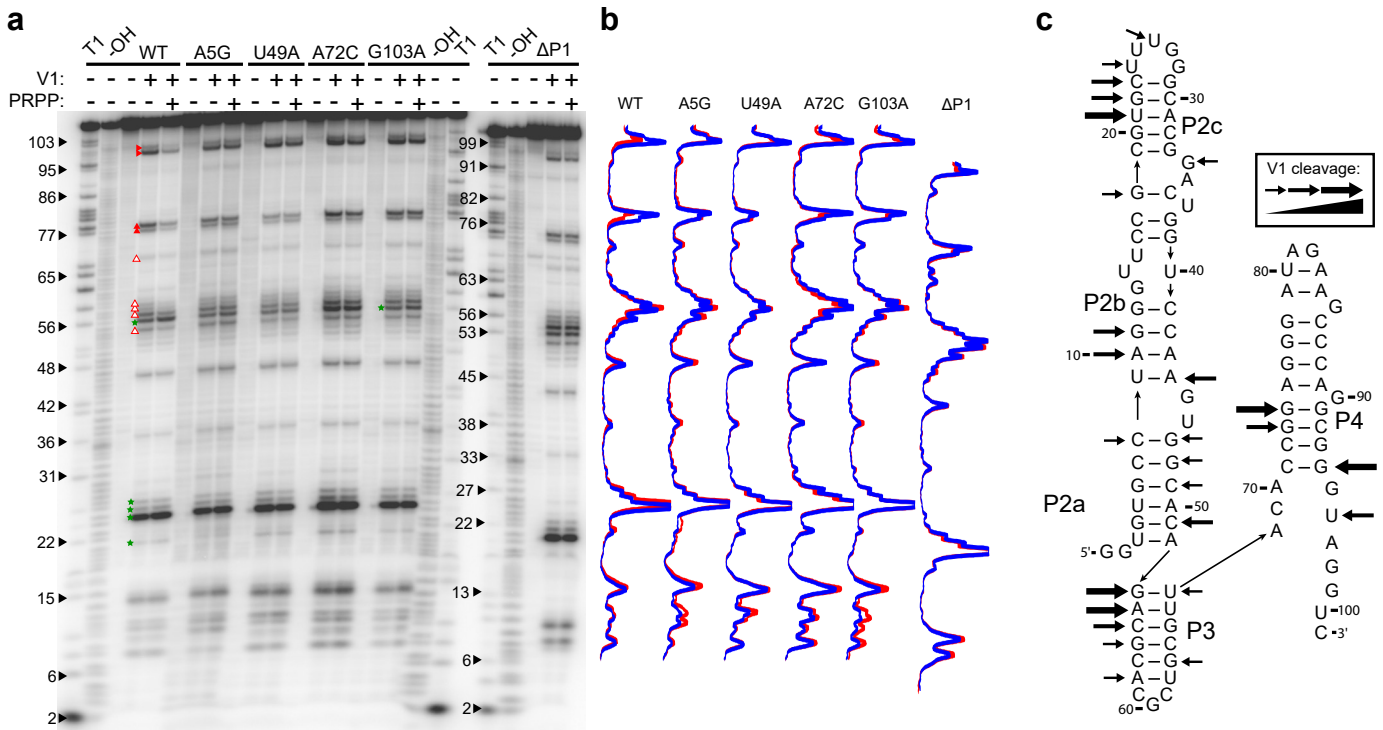
Sl RNA	Ligand	K_D ± SE, μM
WT	PRPP	25.2 ± 6.3
WT	ppGpp	NB
WT	guanidine	NB
A5G	PRPP	weak
G6A	PRPP	NB
U49A	PRPP	NB
A72C	PRPP	NB
ΔG94	PRPP	NB
ΔG94	ppGpp	80 ± 9.5
G94U	PRPP	NB
G94U	ppGpp	115 ± 34.5
G102A	PRPP	41 ± 2.0
G103A	PRPP	weak
ΔP1	PRPP	NB

Sa RNA	Ligand	K_D ± SE, μM
WT/CC	ppGpp	0.5 ± 0.02
WT/CC	PRPP	NB
WT/CC	guanidine	NB

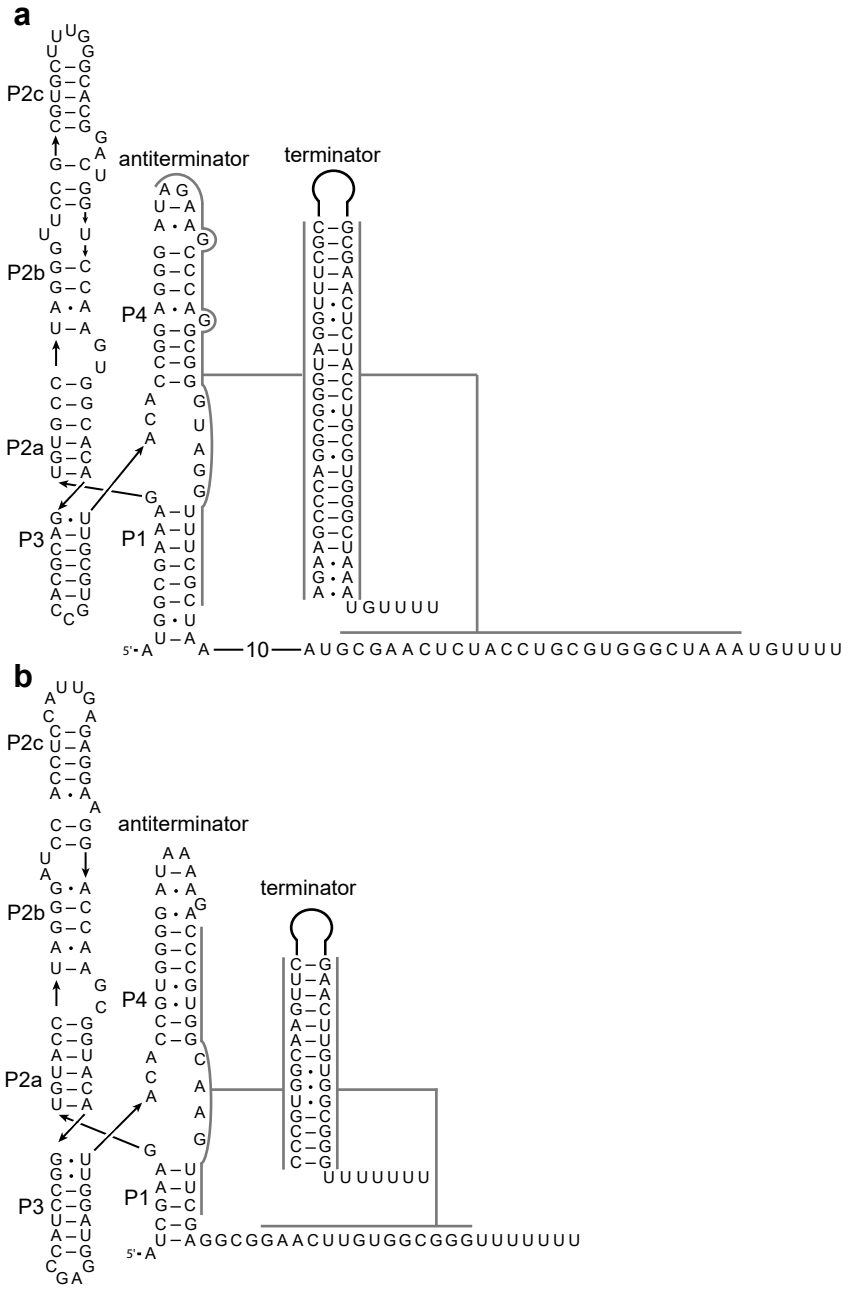
b



Supplementary Figure 6. Binding affinity of the mutant PRPP riboswitch RNAs. (a) Summary of affinity measurements for the PRPP (left table) and ppGpp (right table) riboswitch RNAs. The tables report the average of two experiments ± standard error. WT/CC is a wild-type RNA used in crystallization. NB, no binding. Weak, weak binding. **(b)** Representative ITC titrations and integrated heat plots for binding of the mutant *S. lipocalidus* (Sl) PRPP riboswitch RNAs to PRPP. Due to small heat changes, binding affinities for A5G and G103A RNAs could not be determined accurately. We estimate that K_Ds for these RNAs exceed 0.2 mM. Numbering of the mutated residues corresponds to the crystallized RNA. Note that U49 is U50 in the context of the wild-type RNA used for mutagenesis. Numerical positions of all other mutated nucleotides in RNAs used for ITC measurements are identical to those of the crystallized RNA.



Supplementary Figure 7. Ribonuclease V1 probing of the PRPP riboswitch RNA structure. (**a**) Gel showing results of the V1 nuclease cleavage of the 5′ ^{32}P-labelled 107-nt wild type (WT) and mutant PRPP riboswitch RNAs (0.5 μM) in the absence and presence of 2 mM PRPP. T1 and -OH designate RNase T1 and alkaline ladders, respectively. Weak and strong PRPP-induced cleavage reductions are shown in full and empty red triangles, respectively; cleavage enhancements are shown in green stars. Note that poorly reproducible PRPP-induced effects on short (<15 nt) RNAs likely represent artifacts of the different ethanol precipitation efficiency of the RNA in the absence and presence of the ligand. (**b**) Cleavage intensity profiles of the V1-digested RNAs in the absence (blue line) and presence (red line) of PRPP, determined for each lane of the gel in panel (a). (**c**) Summary of the V1 cleavage results on the ΔP1 mutant. Symbols are described in the inset.



Supplementary Figure 8. Predicted alternative structures in the *S. lipocalidus* PRPP and *S. acidophilus* ppGpp riboswitches. (**a**) The PRPP riboswitch forms an antiterminator hairpin in the presence of PRPP. In the absence of the ligand, the 3′ region of the ligand-sensing domain participates in alternative base-pairing with the dowstream sequence and forms a transcriptional terminator that aborts transcription elongation. The PRPP riboswitch is located upstream of the gene (*purE* in *E. coli*) that encodes for 5-(carboxyamino)imidazole ribonucleotide mutase, an enzyme involved in purine metabolism. (**b**) Same model as (a) for the ppGpp riboswitch. The ppGpp riboswitch is located upsteam of the short putative protein followed by a putative amino-acid binding protein and ketol-acid reductoisomerase (NADP(+)). The last protein is encoded by the *ilvC* gene in *E. coli* and is involved in the biosynthesis of branched-chain amino acids.