Supplemental Information

Structural and biochemical characterization of linear dinucleotide analogs bound to the c-di-GMP-I aptamer[†,‡]

Kathryn D. Smith[1], Sarah V. Lipchock[1,§], Scott A. Strobel[1,2,*]

[1]Department of Chemistry, Yale University, New Haven, Connecticut, USA.

[2]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut, USA.

[†]Supported by National Institutes of Health Grant GM022778.

[‡]Atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.rcsb.org, under accession codes 3UCZ, 3UCU, 3UD4, and 3UD3.

*To whom correspondence should be addressed. Phone: (203) 432-9772. Fax: (203) 432-5767. Email: scott.strobel@yale.edu.

[§]Current address: Monell Chemical Senses Center, Philadelphia, Pennsylvania, USA.

Supplemental Results

Structural refinement of GpG

The structure of GpG bound to the c-di-GMP-I riboswitch was solved using molecular replacement in Phaser as described in the main text. The search model did not contain a ligand molecule. The map from Phaser showed two peaks of equal height corresponding to the positions of P_{G_α} and P_{G_β}. After initial restrained refinement with no ligand modeled, the heights of these peaks show that approximately 60% of the molecules are bound in the P_{G_α} orientation and the remaining 40% are bound in the P_{G_β} orientation. When each orientation was included individually, a large negative peak was observed over the modeled phosphate while a positive peak appeared in the position of the second phosphate. Neither the sizes of these peaks nor the B-factors revealed a strong preference for one orientation over the other. Modeling the two orientations with equal occupancy fit the electron density well and this is what is included in the final model.

During refinement of the structure of GpG, the difference between R_{work} and R_{free} was consistently >6%. A similar trend was observed in the refinement of the GpA and pGpA structures. One explanation for this large difference is that this crystal form has a very low solvent content (~47%) and a small unit cell volume. At the resolution of these data sets (2.7-3.1 Å) the number of unique reflections is therefore lower than might be expected for a different (high solvent content) crystal form. When the ratio of the number of observations to the number of atoms being refined was calculated, it was approximately 3 in these

cases, suggesting that the system is slightly underdetermined. We believe that this contributes to the larger than average difference between R_{work} and R_{free}.

Binding of ApG

Binding of ApG to the c-di-GMP-I riboswitch was first assessed by the gel-shift assay as described in the main text. No binding was observed upon addition of up to 100 μM RNA. In order to confirm this result, we also tested binding of ApG using a competition assay previously reported (*1*). In this experiment, radiolabeled c-di-GMP and increasing concentrations of the unlabeled competitor (ApG in this case) were added to a solution of RNA in folding buffer and allowed to reach equilibrium as previously described. A small amount of displacement of c-di-GMP was observed at very high concentrations of ApG (> 300 μM) that suggests that the affinity of ApG is > 100 μM. When this same experiment was performed using GpA, a binding constant of 2 μM (compared to 8 μM from the gel-shift assay) was observed.

Supplemental References

1. Shanahan, C. A., Gaffney, B. L., Jones, R. A., and Strobel, S. A. (2011) Differential analog binding by two classes of c-di-GMP riboswitches, *J. Am. Chem. Soc. 133*, 15578-15592.

Table S1. Crystallographic data collection and refinement statistics.

	GpG	pGpG	GpA	pGpA
Data collection				
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimensions				
a, *b*, *c* (Å)	49.8, 45.4, 77.3	49.1, 45.4, 77.2	49.6, 45.1, 76.1	50.0, 45.4, 78.7
α, β, γ (°)	90.0, 95.8, 90.0	90.0, 96.2, 90.0	90.0, 96.1, 90.0	90.0, 95.0, 90.0
Resolution (Å)	50-2.8 (2.85-2.80)[a]	50–2.8 (2.85–2.80)	50-2.7 (2.75-2.70)	50-3.1 (3.15-3.10)
R_{merge}	0.086 (0.85)	0.12 (1.0)	0.083 (0.63)	0.14 (0.73)
$I/\sigma I$	21.8 (2.3)	16.9 (1.7)	15.7 (1.9)	15.0 (2.2)
Completeness (%)	99.7 (99.8)	99.7 (99.0)	99.6 (99.8)	99.7 (100)
Redundancy	7.0 (6.6)	7.1 (6.4)	3.9 (3.8)	6.9 (5.3)
Refinement				
Resolution (Å)	77-2.8 (2.87-2.80)	77–2.8 (2.86-2.79)	76-2.7 (2.767-2.697)	79-3.1 (3.16-3.08)
No. reflections	8233 (586)	8150 (530)	8878 (615)	6320 (444)
R_{work}/R_{free}	20.2/26.9 (41.7/47.8)	21.5/26.2 (42.0/44.4)	22.2/29.0 (41.2/36.0)	18.3/25.8 (24.9/33.1)
No. atoms				
RNA	1984	1984	1984	1984
Protein	742	712	712	712
Ligand	43	47	42	46
Ions[b]	18	6	6	7
Water	17	9	6	5
B-factors				
RNA	62.9	60.0	54.8	80.0
Protein	66.8	69.3	48.4	96.3
Ligand	45.9	52.4	46.6	59.5
Ions	54.3	67.4	35.6	78.1
Water	47.4	42.3	41.2	58.1
R.m.s. deviations				
Bond lengths (Å)	0.0054	0.0057	0.0054	0.0059
Bond angles (°)	1.072	1.038	1.123	1.039

[a]Data in parentheses is for the highest resolution bin. [b]The number of atoms counted as ions include inner sphere coordinated water molecules.