Electronic Supplementary Material (ESI) for RSC Medicinal Chemistry.

A systematic analysis of the interactions driving small molecule-RNA recognition

Padroni, G.[a]; Patwardhan N.[a]; Schapira M. [b]; Hargrove A.E. [a]*

[a]Department of Chemistry, Duke University, Durham, North Carolina 27708, United States

[b]Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON M5S 1A8, Canada.

Supporting information

Table of Contents

1. Methods

Structures containing small molecules in complex with RNA solved by x-ray crystallography were extracted from the PDB database on the 9/1/2018 (97 structures). Structures with B-factor higher than 2.7 Å and/or containing the same ligand were filtered out to result in 37 structures that were used for the analysis. If complexes contained the same ligand, structures with better resolution were retained for analysis. Intermolecular contacts between small molecules and RNA were initially identified by using a script that assigned atom types of atoms within 4 Å in ICM (Molsoft). Tight water molecules were included in the analysis and considered as part of the receptor. Inorganic ions and buffer components were excluded by this analysis. The interactions were sorted according to the following classification:

Table S1 Parameters used for sorting intermolecular interactions in this study.

Type	***Atom***	***Distance (Å)***	***Angle***
Hydrophobic interaction	C-C (aliphatic and/or aromatic), C-S and Cl-C	≤ 4	-
Hydrogen bonding	N, O, S	≤3.9	≥90°
Weak hydrogen bonding	CH-O	≤3.6	≥130°
Stacking	C/N aromatic	≤ 4	Edge-to-face 60° ≤Φ≤ 120° Face-to-face Φ≤30°, Φ≥150°
	Amide/guanidine	≤ 4	Edge-to-face 60° ≤Φ≤ 120° Face-to-face Φ≤30°, Φ≥150°
Salt bridges	N(+) , O (-) phosphate	≤ 4	-
Cation-π	N(+) aromatic C-N	≤ 4	-

We also searched for halogen bonding, multipolar interactions, but none of these interactions were observed in our dataset. Coordination between inorganic ions and small molecule ligands

were manually determined for each structure by visually determined coordination of small molecules with K^+ or Mg^{2+} ions. To account for structural differences between RNA and protein interaction between aromatic carbons within 4 Å from each other were also considered as hydrophobic interactions if not already classified as stacking interactions. Electrostatic charges were determined by ICM and/or extracted from the PDB database. The total number of interactions identified was 2213. Figures were generated using UCSF Chimera molecular visualization software.[1]

2. Structures

Table S2 List of small molecule-RNA complexes used in this study.

	PDB_ID	RNA	Small molecules	Resolution (Å)	Dataset
1	1F27	RNA Pseudoknot	Biotin	1.3	**R**
2	1LC4	Eubacterial 16S rRNA A Site	Tobramycin	2.54	**A**
3	2BE0	16S rRNA A-Site	Paromomycin Derivative	2.63	**A**
4	2BEE	16S rRNA A-Site	Paromomycin Derivative	2.6	**A**
5	2F4U	A-Site RNA	Designer Antibiotic	2.6	**A**
6	2FCX	HIV-1-DIS Kissing Loop	Neamine	2	**A**
7	2FCY	HIV-1 DIS kissing loop	Neomycin	2.2	**A**
8	2FCZ	HIV-1 DIS kissing loop	Ribostamycin	2.01	**A**
9	2FD0	HIV-1 DIS kissing loop	Lividomycin A	1.8	**A**
10	2G9C	Purine Riboswitch	Modified Pyrimidines	1.7	**R**
11	2GDI	Thiamine pyrophosphate selective riboswitch	Thiamine pyrophosphate	2.05	**R**
12	2PWT	Bacterial Ribosomal Decoding Site	Aminoglycoside	1.8	**A**
13	2XNW	XPT-PBUX C74U RIBOSWITCH	Triazolo-triazole diamine	1.5	**R**
14	2XNZ	XPT-PBUX C74U RIBOSWITCH	Acetoguanamine	1.59	**R**
15	2XO0	XPT-PBUX C74U RIBOSWITCH	2,4-Diamino-1,3,5-triazine	1.7	**R**
16	2XO1	XPT-PBUX C74U RIBOSWITCH	N6-Methyladinine	1.6	**R**
17	3C44	HIV-1 DIS subtype F	Paromomycin	2	**A**
18	3D2V	Eukaryotic Thiamine pyrophosphate Riboswitch	Pyrithiamine	2	**R**
19	3D2X	Eukaryotic Thiamine pyrophosphate Riboswitch	oxythiamine pyrophosphate	2.5	**R**
20	3G4M	Guanine Riboswitch	2-Aminopurine	2.4	**R**
21	3GAO	Guanine Riboswitch	Xanthine	1.9	**R**
22	3GER	Guanine Riboswitch	6-Chloroguanine	1.7	**R**
23	3MIJ	Telomeric RNA-G Quadruplex	Acridine based inhibitor	2.6	**R**
24	3S4P	Bacterial Ribosomal Decoding Site	Paromomycin analogue	2.56	**A**
25	3SKI	2-Deoxyguanosine riboswitch	2-Deoxyguanosine	2.3	**R**
26	3SKZ	2-Deoxyguanosine riboswitch	Guanosine	2.61	**R**
27	3SUH	THF Riboswitch	5-Formyl THF	2.65	**R**
28	3WRU	Bacterial Ribosomal Decoding Site	Aminoglycoside	2.3	**A**
29	4F8U	Bacterial Ribosomal Decoding Site	Sisomicin	2	**A**
30	4GPX	Protozoal cytoplasmic ribosomal decoding site	6-hydroxysisomicin	2.6	**A**
31	4K31	leishmanial rRNA A-site	Apramycin	1.41	**A**
32	4P20	Bacterial Ribosomal Decoding Site	Amikacin	2.7	**A**
33	4QLN	ydao riboswitch	c-di-AMP	2.65	**R**
34	4YAZ	3',3'-cGAMP riboswitch	3',3'-cGAMP	2	**R**
35	4YB0	3',3'-cGAMP riboswitch	c-di-GMP	2.12	**R**
36	5NEF	Guanidine Riboswitch P1 StemLoop	Guanidine	1.91	**R**

37	5ZEI	Mutated bacterial rRNA decoding site		Geneticin	2.1	**A**

3. Interactions counts

Table S3 Stacking interactions count (distance ≤ 4Å. π stacking angle 60° ≤Φ≤ 120°. Edge-to-face angle Φ≤30°, Φ≥150°).

Interaction	RNA residue	# interactions	# interactions
π-stacking	A	482	764
	C	39	
	G	163	
	U	80	
Edge-to-face	A	1	3
	C	-	
	G	2	
	U	-	
Guanidine/Amide stacking	A	2	3
	C	-	
	G	1	
	U	-	
Total 770			

Table S4 Strong hydrogen bonding interaction count (distance between heavy atoms ≤ 3.9Å. angle Φ≥90°).

Interaction	Ligand atom	RNA Atom	RNA residue	# interactions	# interactions
2'OH	O and N	OH sugar	A	5	21
			C	3	
			G	9	

			U	4	
O-Phosphate	OH and NH	O-(phosphate)	A	47	93
			C	12	
			G	26	
			U	8	
O-sugar	OH and NH	O-sugar O5'	A	12	21
			C	1	
			G	7	
			U	1	
Imino	O and N	NH	A	/	54
			C	/	
			G	11	
			U	43	
Carbonyl Base	NH and OH	O	A	/	106
			C	18	
			G	26	
			U	62	
N aromatic	NH and OH	N7/N1/N3	A	17	87
			C	15	
			G	55	
			U	/	
Exocyclic amine	O or N	NH2	A	44	139
			C	67	

			G	28	
			U	/	
N-wat	O	N	-	-	104
O-wat	O	O	-	-	136
S-wat	O	S			1
Total 762					

Table S5 Hydrophobic interactions count.

Interaction	Ligand atom	RNA Atom	RNA residue	# interactions	# interactions
Aromatic- aliphatic	C aromatic	C sugar	A	11	36
			C	4	
			G	14	
			U	7	
sp2C-Aromatic	Sp2 C	C base	A	38	180
			C	15	
			G	83	
			U	44	
Aliphatic-Aliphatic	C aliphatic	C sugar	A/m6A	3	14
			C	1	
			G	5	
			U	5	
Aliphatic-Aromatic	C aliphatic	C base	A	16	151
			C	6	

			G	128	
			U	1	
Sulphur-sp2/sp3 carbon		C sugar, C base	A	6	12
			C	-	
			G	6	
			U	-	
Halogen-carbon			A	1	1
			C	-	
			G	-	
			U	-	

Total 395

Table S6 Weak hydrogen bonding (CH-O) interaction count (distance between heavy atoms ≤ 3.6Å. angle Φ≥130°)

Interaction	Ligand atom	RNA Atom	RNA residue	# interactions	# interactions
Aliphatic	O	C-H sugar	A	3	7
			C	-	
			G	4	
			U	-	
Aromatic	O	C-H base	A	4	14
			C	3	
			G	2	
			U	5	
Aliphatic	C-H		A	13	61
			C	4	

		O sugar phosphate, carbonyl	G	34	
			U	10	
Aromatic	CH	O sugar phosphate, carbonyl	A	4	6
			C	-	
			G	2	
			U	-	
Water	CH	O water			46

Total 134

Table S7 Distribution of cation-π interactions.

Interaction	Ligand atom	RNA residue	# interactions	# interactions
Cation-π	N	A	6	66
		C	6	
		G	39	
		U	15	

Total 66

Table S8 Distribution of salt-bridge interactions.

Interaction	Ligand atom	RNA residue	# interactions
Salt Bridge	N (positively charged)	O phosphate	63

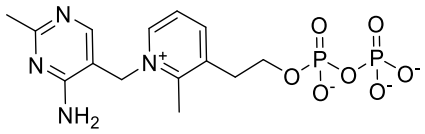
Table S9 Distribution of interactions with inorganic ions.

Interaction	Ligand atom	Ion	# interactions	# interactions

Inorganic	O (negatively charged)	K_+	8	23
		Mg_{2+}	15	

4. Supplementary figures

A **B**



S1

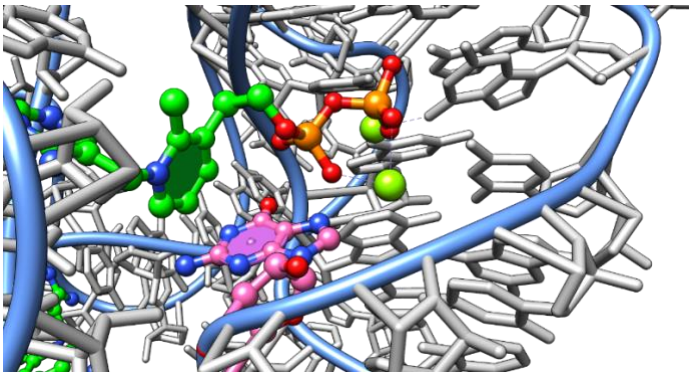
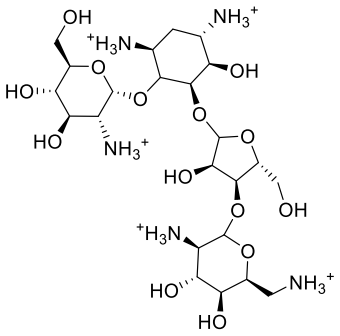
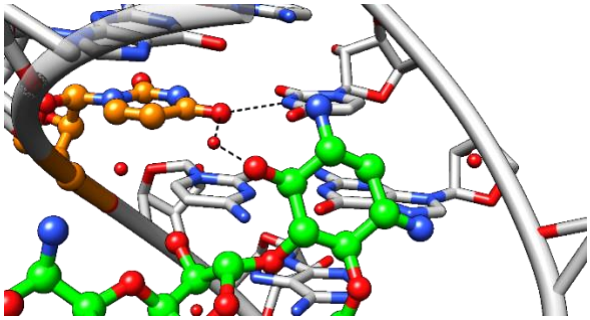


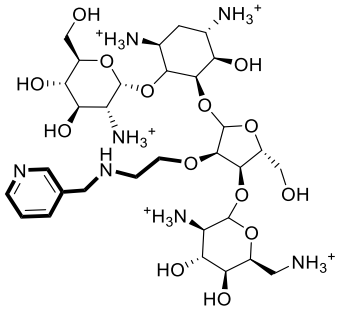
Figure S1 Depiction of the T-shaped (edge to face) interactions observed between the pyridine ring of the antibacterial compound thiamine pyrophosphate with a G residue of the thiamine riboswitch (PDB ID 3D2V).

A **B**



S2





S3

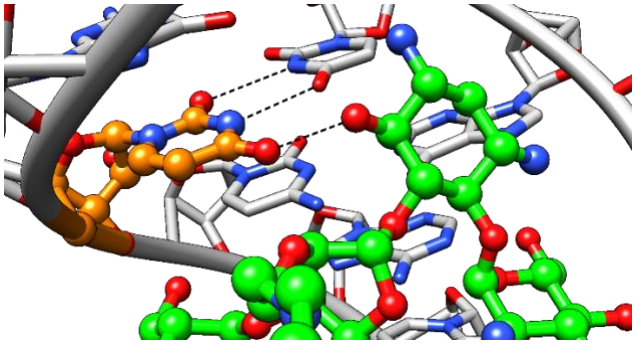


*Figure S2 Structures and depiction of hydrogen bonding pattern after directionality optimization of paramomycin (**S2**, PDBID 1J7T) and its analogue (S3, PDB ID 2BE0, chemical derivatization is bolded).*

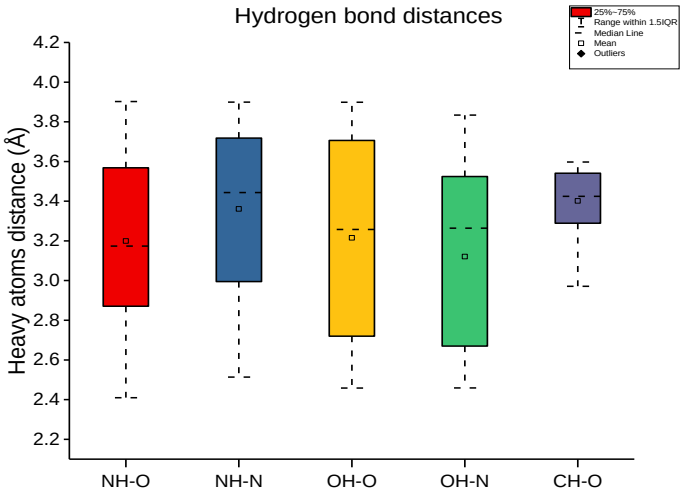


Figure S3 Whisker plots of the distribution of weak and strong hydrogen bonds distance between heavy atoms of small molecule-RNA recognition.

5. References

(1) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612. https://doi.org/10.1002/jcc.20084.