Landes Bioscience

www.landesbioscience.com

Supplemental Material to:

Efrat Kligun and Yael Mandel-Gutfreund

Conformational readout of RNA by small ligands

2013; 10(6)
http://dx.doi.org/10.4161/rna.24682

www.landesbioscience.com/journals/rnabiology/article/24682/

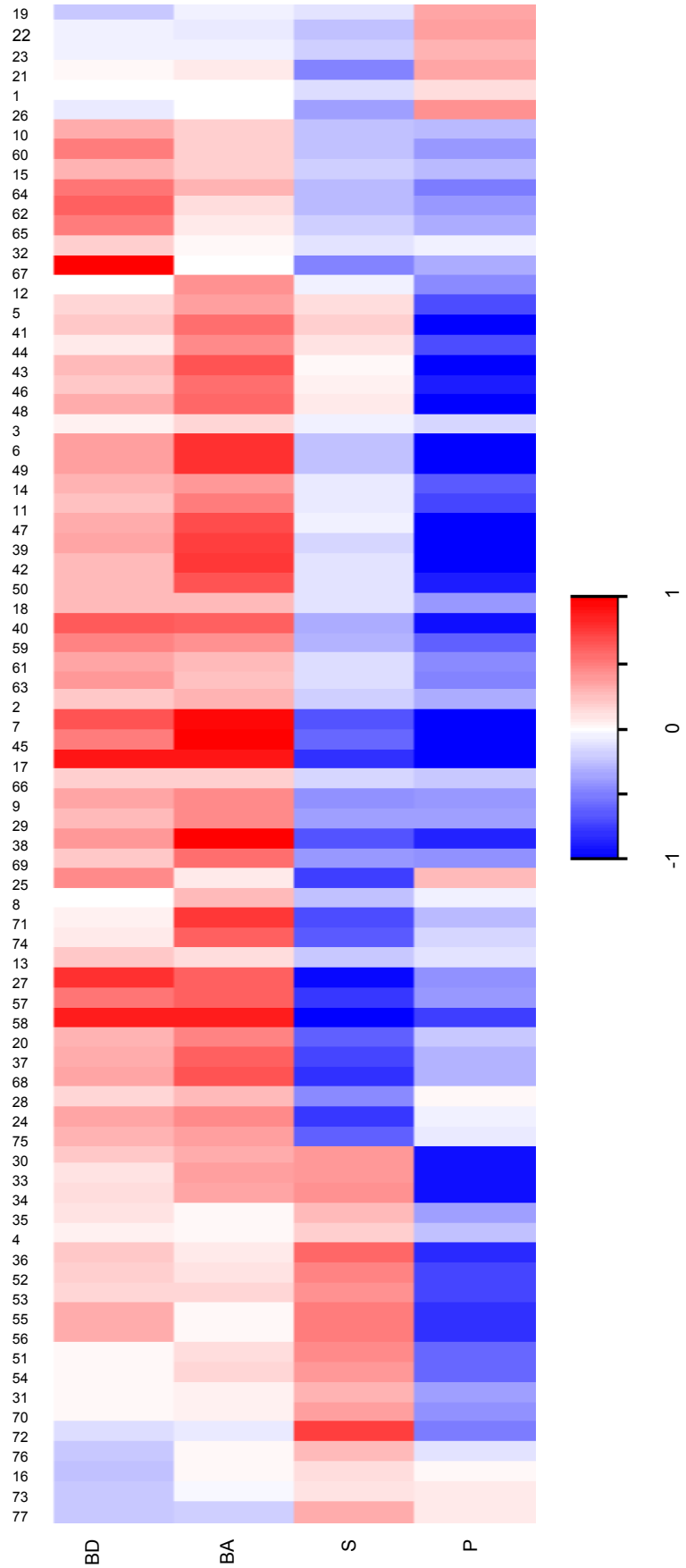

Figure S1: The physicochemical properties of ligand-binding pockets on RNA of 77 structures of RNA ligand complexes. BD denotes base donors, BA base acceptors, S sugar and P phosphate atoms. Numbers represent the index of the complex listed in Supplemental Table S1. The color scheme refers to the standardized score calculated against a background of 70,912 computed pockets (calculated by the *Solvent program). Scores were scaled to range from -1 to 1. Significant biases relative to the background average are colored red and blue for over- and under-representation, respectively*

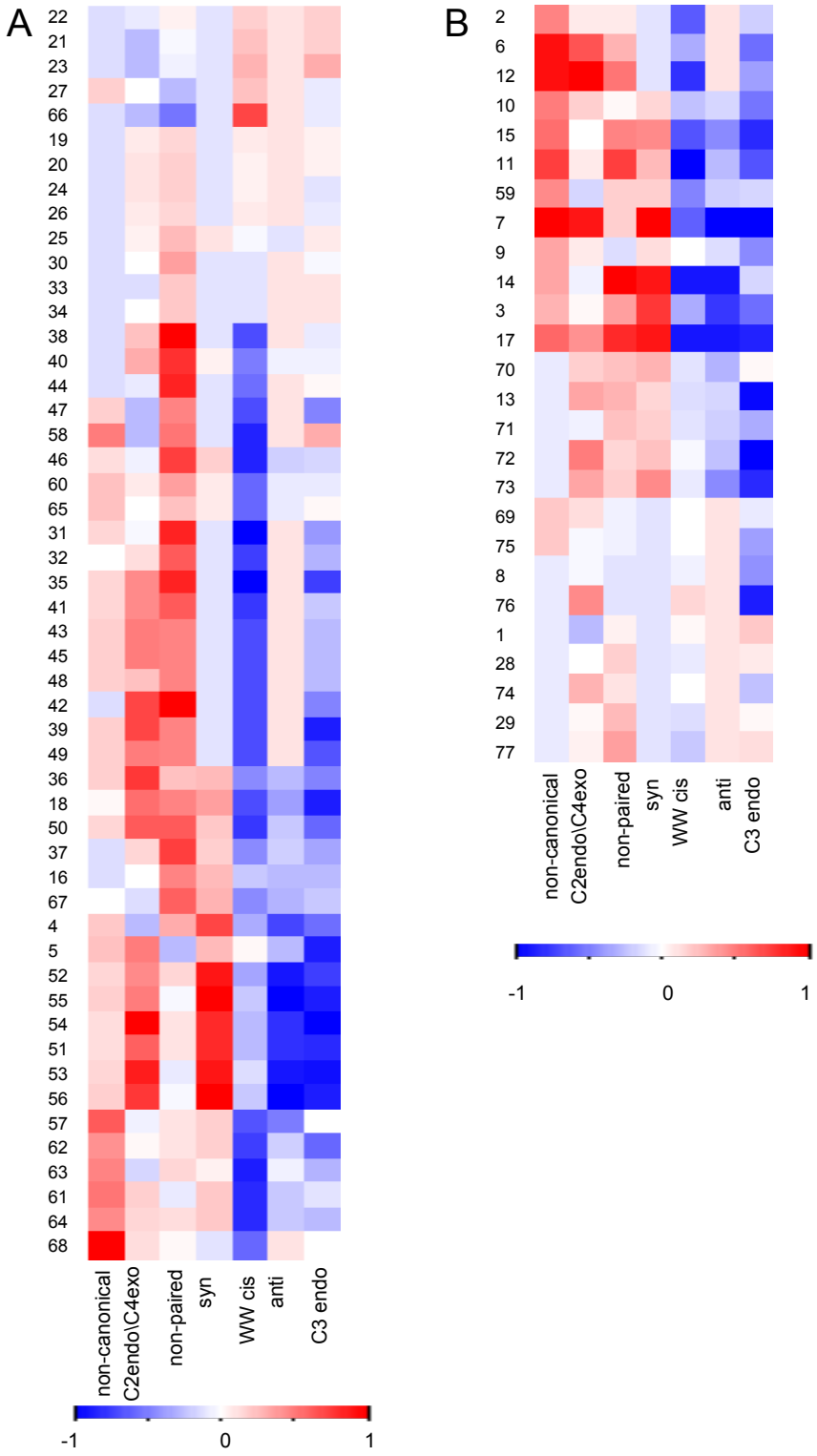

Figure S2: (A) The structural properties of ligand-binding pockets on RNA of X-ray structures of RNA ligand complexes (indices are given in Supplemental Table S1). The color scheme refers to the standardized score calculated against a background of computed pockets derived from the same technique (X-ray) as the structure of the RNA-ligand complex (calculated by the *Solvent program). Scores were scaled to range from -1 to 1. Significant preferences of properties relative to the background of all RNA pockets are colored red (1) while blue denotes under-representation (-1).*
(B) The structural properties of ligand-binding pockets on RNA of NMR structures of RNA ligand complexes

Supplementary Tables

Table S1: List of RNA-ligand complexes*

Index	PDB	Ligand	RNA group	NMR/ X-ray**
1	**2KX8**	**Arginine**	**7SK snRNA**	**NMR**
2	**1FMN**	**Flavin mononucleotide**	**FMN aptamer**	**NMR**
3	**1Q8N**	**Malachite green**	**Malachite green aptamer**	**NMR**
4	1F1T	Tetramethylrosamine	Malachite green aptamer	2.80Å
5	**1ET4**	**Vitamin B12**	**Vitamin B12 aptamer**	**2.30Å**
6	1O15	Theophylline	Theophylline aptamer	NMR
7	**1KOC**	**Arginine**	**Arginine aptamer**	**NMR**
8	**1NEM**	**Neomycin B**	**Neomycin aptamer**	**NMR**
9	**1NBK**	**Argininamide**	**TAT protein aptamer**	**NMR**
10	**2TOB**	**Tobramycin**	**Tobramycin aptamer**	**NMR**
11	**1AM0**	**Adenosine monophosphate**	**AMP aptamer**	**NMR**
12	**1EHT**	**Theophylline**	**Theophylline aptamer**	**NMR**
13	1TOB	Tobramycin	Tobramycin aptamer	NMR
14	**1RAW**	**Adenosine monophosphate**	**ATP aptamer**	**NMR**
15	**2AU4**	**Guanosine triphosphate**	**GTP aptamer**	**NMR**

16	**1F27**	**Biotin**	**Biotin aptamer**	**1.30Å**
17	**1KOD**	**Citrulline**	**Citrulline aptamer**	**NMR**
18	**1NTA**	**Streptomycin**	**Streptomycin aptamer**	**2.90Å**
19	2FD0	Lividomycin	DIS-HIV1	1.80Å
20	2FCX	Neomycin A	DIS-HIV1	2.00Å
21	**3C7R**	**Neomycin**	**DIS-HIV1**	**1.70Å**
22	3C5D	Lividomycin	DIS-HIV1	1.80Å
23	3C44	Paromomycin	DIS-HIV1	2.00Å
24	2FCZ	Ribostamycin	DIS-HIV1	2.01Å
25	1XPF	Spermine	DIS-HIV1	2.30Å
26	2FCY	Neomycin	DIS-HIV1	2.20Å
27	**1FUF**	**Spermine**	**Duplex**	**1.70Å**
28	**2JUK**	**Guandinoneomycin B**	**HIV1 helix**	**NMR**
29	**2KTZ**	**ISH**	**HCV IRES Domain IIa RNA**	**NMR**
30	3GX3	S-Adenosyl-L-homocysteine	SAM-I riboswitch	2.70Å
31	3E5F	SE-adenosyl-serenomethionine	SAM-III riboswitch	2.70Å
32	3E5E	S-Adenosyl-L-homocysteine	SAM-III riboswitch	2.90Å
33	**3GX5**	**S-Adenosylmethionine**	**SAM-I riboswitch**	**2.40Å**
34	3GX2	Sinefungin	SAM-I riboswitch	2.90Å
35	**3E5C**	**S-Adenosylmethionine**	**SAM-III riboswitch**	**2.25Å**
36	3NPQ	S-Adenosyl-L-homocysteine	SAM riboswitch	2.18Å
68	**2QWY**	**S-Adenosylmethionine**	**SAM-II riboswitch**	**2.80Å**
37	**3OWW**	**Glycine**	**Glycine riboswitch**	**2.80Å**
38	3LA5	5-Azacytosine	Purine riboswitch	1.70Å

39	**1Y26**	**Adenine**	**Purine riboswitch**	**2.10Å**
41	3GAO	Xanthine	Purine riboswitch	1.90Å
42	3G4M	9H-purin-2-amine	Purine riboswitch	2.40Å
43	3GOT	2-Fluoroadenine	Purine riboswitch	1.95Å
44	3DS7	2'-Deoxyguanosine	Purine riboswitch	1.85Å
45	2G9C	Pyrimidine-2,4,6-truamine	Purine riboswitch	1.70Å
46	3GER	6-Chloroguanine	Purine riboswitch	1.70Å
47	3FO6	6-O-methylguanine	Purine riboswitch	1.90Å
48	2B57	9H-purine-2,6-diamine	Purine riboswitch	2.15Å
49	2EES	Hypoxanthine	Purine riboswitch	1.75Å
50	**1Y27**	**Guanine**	**Purine riboswitch**	**2.40Å**
40	**3F2Q**	**Flavin mononucleotide**	**FMN riboswitch**	**2.95Å**
51	3DIR	Lysine derivatives	Lysine riboswitch	2.90Å
52	**3D0U**	**Lysine**	**Lysine riboswitch**	**2.80Å**
53	3DIG	Lysine derivatives	Lysine riboswitch	2.80Å
54	3DIQ	Lysine derivatives	LLysine riboswitc	2.70Å
55	3DJ0	Lysine derivatives	Lysine riboswitch	2.50Å
56	3DIL	Lysine	Lysine riboswitch	1.90Å
57	3K1V	7-Deaza-7-aminomethyl-guanine	PreQ1 riboswitch	2.20Å
58	3GCA	PreQ0	PreQ1 riboswitch	2.75Å
59	**2L1V**	**7-Deaza-7-aminomethyl-guanine**	**PreQ1 riboswitch**	**NMR**
60	2HOM	Thiamine phosphate	TPP riboswitch	2.89Å
61	3D2X	Oxythiamine	TPP riboswitch	2.50Å

		pylophosphate		
62	**2CKY**	**Thiamine diphosphate**	**TPP riboswitch**	**2.90Å**
63	3D2G	Thiamine diphosphate	TPP riboswitch	2.25Å
64	3D2V	Pylithiamine pylophosphate	TPP riboswitch	2.00Å
65	2GDI	Thiamine diphosphate	TPP riboswitch	2.05Å
66	**1NYI**	**Guanosine-5'-Monophosphate**	**Hammerhead Ribozyme**	**2.85Å**
67	**2HO7**	**Glucose-6-phosphate**	**Glms ribozyme**	**2.90Å**
69	1UUD	Designer ligands	TAR-HIV1	NMR
70	**1LVJ**	**Acetylpromazine**	**TAR-HIV1**	**NMR**
71	1ARJ	Arginine	TAR-HIV1	NMR
72	1QD3	Neomycin B	TAR-HIV1	NMR
73	1UTS	Designer ligands	TAR-HIV1	NMR
74	1AKX	Arginine	TAR-HIV2	NMR
75	1UUI	Designer ligands	TAR-HIV2	NMR
76	2KGP	Mitoxantrone	Splicing regulatory element	NMR
77	**1EI2**	**Neomycin**	**Splicing regulatory element**	**NMR**

*** Representative pockets are highlighted in gray**

**For structures solved by x-ray crystallography the resolution of the structures is given

Table S2: Comparison between the predicted and known pockets using the program 'Solvent'

PDB	Number of pockets	Number of Common nucleotides	PPV	Sensitivity
2KGP	9	2/9	0.17	0.22
3DIR	546	9/12	0.75	0.75
3GX3	249	11/17	0.92	0.65
1FMN	32	10/10	0.83	1
2KX8	63	11/11	0.92	1
3D0U	489	8/11	0.67	0.73
1Q8N	61	12/15	1	0.8
1UUD	23	12/15	1	0.8
2FD0	54	12/15	1	0.8
1LVJ	6	10/11	0.83	0.91
2JUK	13	12/16	1	0.75
1F1T	39	9/9	0.75	1
3E5F	97	11/11	0.92	1
3E5E	94	12/13	1	0.92
1ET4	432	10/10	0.83	1
3K1V	28	10/10	0.83	0.83
1ARJ	31	12/15	1	0.8
1O15	51	10/10	0.83	1
3LA5	165	9/10	0.75	0.9
1NYI	44	1/5	**NA**	**NA**
3F2Q	212	11/21	0.92	0.52
1KOC	14	11/12	0.92	0.92
1NEM	29	12/17	1	0.71
2FCX	68	12/14	1	0.86

3GAO	134	8/11	0.67	0.73
3G4M	143	9/10	0.75	0.9
2HO7	381	11/18	0.92	0.61
1NBK	40	12/20	1	0.6
3GOT	130	7/10	0.58	0.7
3DIG	542	7/11	0.58	0.64
3C7R	50	12/15	1	0.8
1QD3	34	10/13	0.83	0.77
3C5D	51	12/26	1	0.46
2QWY	255	12/18	1	0.67
1FUF	24	12/20	1	0.6
3C44	76	12/16	1	0.75
2HOM	155	11/20	0.92	0.55
3DS7	378	11/13	0.92	0.85
3GCA	39	10/12	0.83	0.83
2TOB	12	12/17	1	0.71
2FCZ	158	12/14	1	0.86
1UTS	8	5/8	0.42	0.63
2G9C	138	7/10	0.58	0.7
1AM0	5	12/12	1	1
1Y26	162	7/10	0.58	0.7
1EHT	29	10/10	0.83	1
3GX5	266	12/17	1	0.71
1AKX	19	12/14	1	0.86
1TOB	29	12/17	1	0.71
3GER	129	9/12	0.75	0.75
1UUI	13	12/14	1	0.86

3OWW	437	8/12	0.67	0.67
3FO6	142	8/10	0.67	0.8
3D2X	499	12/23	1	0.52
2CKY	346	12/24	1	0.5
1RAW	55	12/13	1	0.92
1XPF	56	12/16	1	0.75
2B57	131	7/10	0.58	0.7
2AU4	53	12/16	1	0.75
1F27	35	10/10	0.83	1
3D2G	496	12/22	1	0.55
3GX2	270	12/17	1	0.71
1KOD	23	12/13	1	0.92
2FCY	60	12/15	1	0.8
2L1V	48	1216	1	0.75
3DIQ	521	9/12	0.75	0.75
2EES	138	7/10	0.58	0.7
1Y27	149	8/11	0.67	0.73
3E5C	90	11/11	0.92	1
3D2V	507	12/23	1	0.52
3DJ0	530	7/10	0.58	0.7
1EI2	25	12/18	1	0.67
2KTZ	55	12/14	1	0.86
2GDI	372	12/20	1	0.6
3DIL	537	7/10	0.58	0.7
3NPQ	261	9/10	0.75	0.9
1NTA	39	12/15	1	0.8

Sensitivity and PPV were calculated for known binding sites that consist more than 5 nucleotides. Number of pockets represents the number of putative pockets that were predicted per each structure. The number of overlapping nucleotides were calculated for the pocket with the maximum number of overlapping nucleotides. Size of the predicted pocket was predefined as 12 nucleotides.

Table S3: Comparison between the predicted and the known pockets using the program 'Fpocket'

PDB	Number of pockets	Size of putative pocket	Size of the known binding site	Number of overlapping nucleotides	PPV	Sensitivity
2KGP	1	5	9	5	NA	NA
3DIR	11	7	12	4	0.57	0.33
3GX3	9	10	17	8	0.8	0.47
1FMN	1	6	10	6	1	0.6
2KX8	3	9	11	9	1	0.82
3D0U	17	12	11	9	0.75	0.82
1Q8N	3	14	15	13	0.93	0.87
1UUD	3	5	15	5	NA	NA
2FD0	6	12	15	9	0.75	0.6
1LVJ	2	6	11	6	1	0.55
2JUK	0	0	16	0	NA	NA
1F1T	3	4	9	4	NA	NA
3E5F	3	7	11	6	0.86	0.55
3E5E	5	4	13	4	NA	NA
1ET4	16	5	10	3	NA	NA
3K1V	2	6	12	6	1	0.5
1ARJ	2	13	15	11	0.85	0.73
1O15	2	11	10	9	0.82	0.9

3LA5	7	8	10	7	0.88	0.7
1NYI	1	0	5	0	NA	NA
3F2Q	5	35	21	15	0.43	0.71
1KOC	3	9	12	8	0.89	0.67
1NEM	1	13	17	13	1	0.76
2FCX	4	4	14	4	NA	NA
3GAO	8	10	11	10	1	0.91
3G4M	6	14	10	4	0.29	0.4
2HO7	10	32	18	10	0.31	0.56
1NBK	3	8	20	5	0.63	0.25
3GOT	6	6	10	6	1	0.6
3DIG	17	7	11	6	0.86	0.55
3C7R	5	11	15	10	0.91	0.67
1QD3	5	12	13	7	0.58	0.54
3C5D	4	6	26	6	1	0.23
2QWY	17	16	18	16	1	0.89
1FUF	3	14	20	14	1	0.7
3C44	3	16	16	8	0.5	0.5
2HOM	8	25	20	14	0.56	0.7
3DS7	13	13	13	6	0.46	0.46
3GCA	1	16	12	9	0.56	0.75
2TOB	1	12	17	11	0.92	0.65
2FCZ	11	15	14	6	0.4	0.43
1UTS	5	7	8	7	1	0.88
2G9C	4	12	10	9	0.75	0.9
1AM0	2	8	12	6	0.75	0.5
1Y26	7	20	10	4	0.2	0.4
1EHT	2	5	10	5	NA	NA

3GX5	9	8	17	6	0.75	0.35
1AKX	3	7	14	7	1	0.5
1TOB	2	5	17	5	NA	NA
3GER	6	8	12	4	0.5	0.33
1UUI	3	8	14	5	0.63	0.36
3OWW	21	5	12	3	NA	NA
3FO6	7	5	10	5	NA	NA
3D2X	15	27	23	15	0.56	0.65
2CKY	14	23	24	18	0.78	0.75
1RAW	4	8	13	5	0.63	0.38
1XPF	4	7	16	7	1	0.44
2B57	8	5	10	5	NA	NA
2AU4	2	14	16	12	0.86	0.75
1F27	3	10	10	5	0.5	0.5
3D2G	16	30	22	18	0.6	0.82
3GX2	8	14	17	6	0.43	0.35
1KOD	6	5	13	4	NA	NA
2FCY	7	8	15	7	0.88	0.47
2L1V	3	8	16	8	1	0.5
3DIQ	16	29	12	4	0.14	0.33
2EES	9	7	10	7	1	0.7
1Y27	6	10	11	5	0.5	0.45
3E5C	2	10	11	10	1	0.91
3D2V	10	15	23	11	0.73	0.48
3DJ0	16	7	10	4	0.57	0.4
1EI2	2	13	18	13	1	0.72
2KTZ	2	8	14	8	1	0.57
2GDI	17	11	20	11	1	0.55

3DIL	15	12	10	9	0.75	0.9
3NPQ	7	10	10	9	0.9	0.9
1NTA	4	10	15	10	1	0.67

Sensitivity and PPV were calculated between the known binding sites and the computed pockets that consist more than 5 nucleotides. Number of pockets represents the number of putative pockets that were calculated for each structure. The number of overlapping nucleotides were calculated for the pocket with the highest overlap.

Table S4: Detailed information on the distribution of nucleotides possessing the unique conformations in the representative ligand binding pockets

Name	Total number in pocket	syn	C2_endo	c4_exo	syn+ c2_endo	syn+ c4_exo
2KX8	11	0	0	0	0	0
1FMN	10	0	0	2	0	0
1Q8N	15	3	1	1	0	1
1F1T	9	2	0	0	0	0
1ET4	10	1	3	0	1	0
1O15	10	0	3	1	0	0
1KOC	12	3	3	4	2	1
1NEM	17	0	1	1	0	0
1NBK	20	1	0	4	0	0
2TOB	17	1	1	3	1	0
1AM0	12	1	1	1	0	0
1EHT	10	0	5	0	0	0
1TOB	17	1	4	0	0	0
1RAW	13	3	1	0	1	0
2AU4	16	2	1	1	0	0
1F27	10	1	1	0	0	0
1KOD	13	3	1	4	1	2
1NTA	15	2	3	1	0	1
2FD0	15	0	2	0	0	0
2FCX	14	0	2	0	0	0
3C7R	15	0	0	0	0	0
3C5D	26	0	2	0	0	0
3C44	16	0	0	0	0	0
2FCZ	14	0	2	0	0	0
1XPF	16	1	2	0	1	0
2FCY	15	0	2	0	0	0
1FUF	20	0	2	0	0	0
2JUK	16	0	1	1	0	0

2KTZ	14	0	1	1	0	0
3GX3	17	0	0	1	0	0
3E5F	11	0	1	0	0	0
3E5E	13	0	2	0	0	0
3GX5	17	0	1	0	0	0
3GX2	17	0	0	1	0	0
3E5C	11	0	3	0	0	0
3NPQ	10	1	4	0	1	0
3OWW	12	1	2	0	0	0
3LA5	10	0	2	0	0	0
1Y26	10	0	2	1	0	0
3F2Q	21	1	3	1	1	0
3GAO	11	0	3	0	0	0
3G4M	10	0	2	1	0	0
3GOT	10	0	3	0	0	0
3DS7	13	0	1	0	0	0
2G9C	10	0	3	0	0	0
3GER	12	1	1	0	0	0
3FO6	10	0	0	0	0	0
2B57	10	0	2	0	0	0
2EES	10	0	3	0	0	0
1Y27	11	1	2	1	0	0
3DIR	12	3	4	0	2	0
3D0U	11	3	3	0	1	0
3DIG	11	3	3	1	1	0
3DIQ	12	3	4	1	1	0
3DJ0	10	3	3	0	1	0
3DIL	10	3	4	0	2	0
3K1V	12	1	1	0	1	0
3GCA	12	0	0	0	0	0
2L1V	16	1	0	1	0	0
2HOM	20	1	1	1	0	0
3D2X	23	2	4	0	1	0

2CKY	24	2	1	1	0	0
3D2G	22	1	1	0	0	0
3D2V	23	2	2	1	0	0
2GDI	20	1	2	0	0	0
1NYI	5	0	0	0	0	0
2HO7	18	2	1	0	0	0
2QWY	18	0	1	1	0	0
1UUD	15	0	1	2	0	0
1LVJ	11	1	2	0	0	0
1ARJ	15	1	0	2	0	0
1QD3	13	1	4	0	0	0
1UTS	8	1	0	3	0	0
1AKX	14	0	3	0	0	0
1UUI	14	0	0	2	0	0
2KGP	9	0	0	4	0	0
1EI2	18	0	1	2	0	0

Table S5: RNA ligands interactions

		C2endo+ C4exo	syn	C2endo+C 4exo+syn	non-paired	non-canonical
Hydrogen Bonds (306)*	Total Pockets	186	66	25	333	150
	Overlap	55	28	13	107	50
	P value	0.361	0.0076	0.0096	0.035	0.083
	Percent of bonds	17.97	9.15	4.25	34.97	16.34
Van der Waals (466)*	Total Pockets	186	66	25	333	150
	Overlap	94	43	18	136	80
	P value	0.014	0.0002	0.0029	0.850	0.004
	Percent of bonds	20.17	9.23	3.86	29.18	17.17
Hydrophobic (261)*	Total Pockets	186	66	25	333	150
	Overlap	63	30	15	60	43
	P value	0.0006	7.13e-05	0.0001	0.999	0.097
	Percent of bonds	24.14	11.49	5.75	22.99	16.48
Aromatic (118)*	Total Pockets	186	66	25	333	150
	Overlap	24	9	1	23	15
	P value	0.1996	0.285	0.946	0.999	0.692
	Percent of bonds	20.34	7.63	0.85	19.49	12.71

* Total number of bonds of this category