Supplementary Information

Fragment-based modelling of single stranded RNA bound to RNA recognition motif containing proteins

Isaure Chauvot de Beauchene, Sjoerd J. de Vries, Martin Zacharias*

Physics Department T38, Technical University of Munich, James-Franck-Str. 1, 85748 Garching, Germany

* To whom correspondence should be addressed, e-mail: martin.zacharias@ph.tum.de

Table S1. Sequence identity (%) between the RRM domains in the benchmark.

PDB		1B7F		1CVJ		2MGZ		2YH1		3NNH		4BS2		4N0T	
	RRM	1	2	1	2	1	2	1	2	1	2	1	2	2	3
1B7F	1	100	38	34	29	29	24	19	31	21	21	27	21	21	24
	2	38	100	26	24	28	28	25	29	32	32	25	23	19	24
1CVJ	1	34	26	100	31	27	22	17	19	25	25	27	27	16	16
	2	29	24	31	100	24	30	17	29	27	27	23	13	14	26
2MGZ	1	29	28	27	24	100	28	24	20	25	25	30	27	20	19
	2	24	28	22	30	28	100	13	26	36	36	32	30	8	28
2YH1	1	20	25	17	17	24	13	100	18	18	18	19	24	17	13
	2	31	29	19	29	20	26	18	100	23	23	21	18	13	20
3NNH	1	21	32	25	27	25	36	18	23	100	100	25	27	14	20
	2	21	32	25	27	25	36	18	23	100	100	25	27	14	20
4BS2	1	27	25	27	23	30	32	19	23	25	25	100	24	16	15
	2	21	23	27	13	27	30	24	18	27	27	24	100	11	17
4N0T	2	21	19	16	14	20	8	17	13	14	14	16	11	100	7
	3	24	24	16	26	19	28	13	20	20	20	15	17	7	100

Table S2. List of the 850 PDB structures used for building of the fragment library.

1A1T	1A34	1A4T	1A9N	1AQ3	1AQ4	1ASY	1AUD	1AV6	1B23	1B7F	1BIV	1BMV	1C0A	1CVJ	1CWP	1DDL	1DI2
1DRZ	1DZ5	1E7K	1E8O	1EC6	1EFW	1EIY	1EKZ	1EQQ	1ETF	1EUQ	1EXY	1F6U	1F7U	1F7V	1FJE	1FNX	1FXL
1G2E	1G59	1G70	1GAX	1GTF	1GTR	1H2C	1H3E	1H4Q	1HJI	1HVU	1I5L	1I9F	1IL2	1J1U	1J2B	1JBR	1JBS
1JID	1K1G	1K8W	1KNZ	1KOG	1KQ2	1L1C	1LAJ	1M5K	1M5O	1M8V	1M8X	1M8Y	1N1H	1N35	1N38	1N77	1NB7
1NYB	1O0B	1OB2	1OOA	1P6V	1PGL	1PVO	1Q2R	1Q2S	1QF6	1QFQ	1QRS	1QTQ	1QU2	1QZW	1R3E	1RGO	1RKJ
1SER	1SI3	1SJ3	1SZ1	1T2R	1TFY	1TRJ	1TTT	1U0B	1U1Y	1U6P	1UON	1URN	1UTD	1UTF	1UVI	1UVJ	1UVL
1UVM	1UVN	1VBX	1VVJ	1WMQ	1WNE	1WPU	1WSU	1WWD	1WWE	1WWF	1WWG	1WZ2	1XOK	1XPO	1XPR	1XPU	1YTU
1YTY	1YVP	1ZBH	1ZBN	1ZDH	1ZDI	1ZDJ	1ZDK	1ZE2	1ZJW	1ZL3	1ZZN	2A1R	2A8V	2A9X	2AB4	2AD9	2ADB
2ADC	2ANN	2ASB	2B3J	2B63	2B6G	2BBV	2BGG	2BH2	2BYT	2C06	2C0B	2C4R	2C50	2CJK	2CSX	2CZJ	2D6F
2DB3	2DER	2DET	2DLC	2DR2	2DRA	2DU4	2DXI	2E9R	2E9T	2E9Z	2ERR	2ESE	2F8K	2F8S	2FK6	2FMT	2FY1
2G4B	2GIC	2GJE	2GJW	2GTT	2GXB	2HT1	2HYI	2I2Y	2I91	2IHX	2IX1	2IY5	2J0Q	2J0S	2JEA	2JLU	2JLW
2JLX	2JLY	2JLZ	2JPP	2KDQ	2KFY	2KG0	2KG1	2KH9	2KM8	2KX5	2KXN	2L3C	2L41	2L5D	2LBS	2LEB	2LEC
2LI8	2LUP	2M5U	2M8D	2MB0	2MF0	2MF1	2MFC	2MFH	2MGZ	2MJH	2MKI	2MKK	2MKN	2MOB	2MQT	2MQV	2MS0
2MS1	2MTJ	2MTK	2MTV	2NQP	2NR0	2NZ4	2OIH	2OZB	2PJP	2PLY	2PO1	2PY9	2Q66	2QUX	2R7R	2R7S	2R7T
2R7U	2R7V	2R7W	2R7X	2R92	2R93	2RD2	2RE8	2RKJ	2RQC	2RRA	2RS2	2RSK	2RU3	2RU7	2UWM	2V0G	2V3C
2VNU	2VOD	2VON	2VOO	2VOP	2W2H	2WJ8	2XB2	2XBM	2XC7	2XD0	2XDB	2XDD	2XFM	2XGJ	2XNR	2XS2	2XS5
2XS7	2XXA	2XZL	2XZO	2Y8W	2Y8Y	2Y9A	2Y9H	2YH1	2YHM	2YJY	2YKG	2ZH1	2ZH7	2ZI0	2ZM5	2ZNI	2ZUF
2ZXU	2ZZM	2ZZN	3A2K	3A6P	3ADB	3AEV	3AGV	3AHU	3AKZ	3AM1	3AMT	3AMU	3AVT	3AVU	3AVV	3AVW	3AVX
3AVY	3B0U	3B0V	3BO2	3BOY	3BSB	3BSN	3BSX	3BT7	3BX2	3BX3	3CUL	3CZ3	3D2S	3DD2	3DKN	3EGZ	3EPH
3EQT	3EX7	3FHT	3FOZ	3FTE	3FTF	3G0H	3G9C	3G9Y	3H5X	3HHN	3HHZ	3HL2	3HSB	3HTX	3I5Y	3I61	3IAB
3ICE	3ICQ	3IE1	3IRW	3IWN	3IZZ	3J5Y	3J6V	3J7A	3J7Y	3J7Z	3J80	3J81	3J8G	3K0J	3K49	3K4E	3K5Q
3K5Y	3K5Z	3K61	3K62	3K64	3KFU	3KLV	3KMQ	3KMS	3KNA	3KOA	3KTW	3L25	3LQX	3M7N	3MDG	3MDI	3MJ0
3NDB	3NL0	3NMA	3NMR	3NNA	3NNC	3NNH	3O3I	3O8C	3O8R	3OIJ	3OIN	3OL6	3OUY	3OV7	3OVS	3P49	3P6Y
3PEW	3PEY	3PF4	3PF5	3PKM	3PTX	3PU0	3PU1	3PU4	3Q0L	3Q0M	3Q0N	3Q0O	3Q0P	3Q0Q	3Q0R	3Q0S	3Q1Q
3Q2T	3QG9	3QGB	3QGC	3QJJ	3QJL	3QJP	3QRP	3QRR	3QSU	3QSY	3R1H	3R2C	3R9W	3RC8	3RER	3RW6	3SIU
3SIV	3SN2	3SNP	3SQW	3SQX	3T3N	3T3O	3T5N	3T5Q	3TRZ	3TS0	3TS2	3TUP	3U2E	3UCU	3UZS	3UZT	3V11
3V6Y	3V71	3V74	3VNU	3VNV	3VYX	3W1K	3W3S	3WC2	3WFQ	3WQY	3WQZ	3WRU	3ZC0	3ZD6	3ZGZ	3ZJT	3ZLA
3ZN8	484D	4A2X	4A76	4AL5	4ALP	4AM3	4AQ7	4AS1	4ATO	4AY2	4B3G	4B8T	4BHH	4BKK	4BPB	4BS2	4C7O
4C8Y	4C8Z	4C9D	4CIO	4CQN	4CS1	4CSF	4CSU	4CXG	4CXH	4D25	4D26	4D5L	4D5N	4DZS	4E78	4ED5	4EI1
4EI3	4ENN	4ERD	4EYA	4F02	4F1N	4F3T	4FWT	4GCW	4GHA	4GL2	4GV3	4GV9	4H5O	4H5P	4H6F	4HOR	4HOS
4HOT	4HT8	4HT9	4I67	4I9Q	4IFD	4II9	4IJS	4ILL	4ILM	4IQX	4J1G	4J7L	4J7M	4JGN	4JK0	4JNG	4JVH
4JVY	4JXX	4JYZ	4K4S	4K4U	4K4W	4K4X	4K50	4KJI	4KNQ	4KQ0	4KR2	4KR3	4KR6	4KR7	4KR9	4KRE	4KRF
4KTG	4KVB	4KXT	4KZD	4KZE	4L47	4L71	4L81	4L8H	4L8R	4LEL	4LF4	4LF5	4LF6	4LF7	4LF8	4LF9	4LFA
4LFB	4LFC	4LFZ	4LG2	4LJ0	4LMZ	4LNT	4LSK	4LT8	4LX5	4LX6	4M30	4M4O	4M59	4M6D	4M7A	4MCE	4MCF
4MDX	4N0T	4N2Q	4N2S	4N48	4NFO	4NFP	4NFQ	4NGB	4NIA	4NL3	4NLF	4NMG	4NXH	4NXM	4NXN	4NYA	4NYB
4NYC	4NYD	4NYG	4O26	4O41	4O8J	4OAV	4OE1	4OHZ	4OJI	4OLA	4OO1	4OO8	4OOG	4OQ8	4OQ9	4OQU	4OX9
4P20	4P3E	4P5J	4P8Z	4P95	4P9R	4PCJ	4PDB	4PDQ	4PEH	4PEI	4PJO	4PKD	4PLX	4PMI	4PMW	4PQU	4PQV
4PR6	4PRF	4PUO	4PWD	4PY5	4Q0B	4Q5S	4Q5V	4Q9Q	4Q9R	4QI2	4QIK	4QIL	4QJD	4QJH	4QK8	4QK9	4QKA
4QLM	4QLN	4QM6	4QOZ	4QQB	4QU6	4QU7	4QYZ	4R0D	4R3I	4RCJ	4RGE	4RGF	4RJ1	4RKV	4RNE	4S20	4TS0
4TS2	4TU0	4TUW	4TUX	4TV0	4TVX	4TYW	4TYY	4TZ0	4TZ6	4TZP	4TZV	4TZW	4TZX	4TZY	4TZZ	4U1U	4U1V
4U20	4U24	4U25	4U26	4U27	4U34	4U35	4U37	4U38	4U7U	4U8T	4UDV	4UN3	4UN4	4UN5	4UY8	4UYJ	4UYK
4V19	4V1M	4V1N	4V1O	4V49	4V4A	4V4B	4V4H	4V4P	4V4Q	4V4R	4V4S	4V4T	4V4V	4V4W	4V4X	4V4Y	4V4Z
4V50	4V52	4V53	4V54	4V55	4V56	4V57	4V5B	4V5F	4V5G	4V5L	4V5O	4V5P	4V5Q	4V5R	4V5S	4V5Y	4V61
4V64	4V65	4V66	4V6C	4V6D	4V6E	4V6K	4V6T	4V6Y	4V70	4V71	4V72	4V73	4V74	4V75	4V76	4V77	4V78
4V79	4V7A	4V7B	4V7C	4V7D	4V7H	4V7I	4V7P	4V7S	4V7T	4V7U	4V7V	4V85	4V89	4V8A	4V8N	4V8O	4V8Q
4V8U	4V90	4V9C	4V9D	4V9H	4V9I	4V9J	4V9K	4V9L	4V9M	4W29	4W5N	4W5O	4W5Q	4W5R	4W5T	4W90	4W92
4WAL	4WAN	4WF1	4WFL	4WFM	4WJ3	4WJ4	4WKJ	4WRT	4WSA	4WSB	4WT8	4WUS	4WWE	4WWT	4WWW	4WZJ	4X67
4X6A	5MSF	6MSF	7MSF														

Table S3. Index of the anchored nucleotides in the RNA-RRM complexes structures.

	RRM 1		RRM2	
	RNP-2	RNP-1	RNP-2	RNP-1
1B7F	2	3	No aromatic	9
1CVJ	2	3	5	7
2MGZ	5	6	10	11
2YH1	3	4	6	7
3NNH	4	Exotic rotamer	9	10
4BS2	No aromatic	5	8	9
4N0T	No aromatic	4	10	11

The numbering of a RRM refers to the 5' (RRM1) or 3' (RRM2) part of the RNA that it binds, and not to their order in the protein sequence. The indexes refer to the nucleotide sequence that was docked, I. e. after removal of the non-bounded nucleotide from the PDB structures.

Table S4. RMSD between each bound fragment of the benchmark and the closest conformer of the corresponding sub-library, after withdrawal of the bound form from the sub-library if present.

	1cvj	1b7f	2mgz	2yh1	3nnh	4bs2	4n0t
frag. 1	1.73	0.92	0.37	0.27	0.73	0.21	1.45
frag. 2	0.43	0.76	0.44	0.43	0.42	0.17	1.84
frag. 3	0.25	0.57	0.67	0.18	1.38	0.45	1.08
frag. 4	0.19	0.51	0.53	0.71	0.61	0.32	1.20
frag. 5	1.20	0.85	0.68	0.28	0.73	0.55	1.65
frag. 6	0.95	1.03	0.22	0.44	0.53	0.42	0.80
frag. 7		0.45	0.29		0.35	0.53	0.67
frag. 8			0.66		0.59	0.58	0.89
frag. 9			0.58				1.53
frag. 10			0.74				1.93
frag. 11							2.17

Table S5. Closest pose and number of hits (RMSD ≤ 2 Å) in the top-scored docking poses for each anchored fragments, compared to the closest conformer in library after perfect fitting to the bound form, and to the conformer corresponding to best docking pose after perfect fitting to the bound form.

	Docking run	Docking			RMSD(Å)after fitting		RMSD (Å, coarse-grain) of the predicted anchor(s)
		Min RMSD(Å)	Nb hits	Nb poses	Closest conformer in library	Conformer of the best docking pose	
1B7F	Frag 1 AMB	1.5	3	1000	0.9	0.9	1.0 – 1.3
	Frag 2 AMF	1.0	20	1000	0.8	0.8	1.0 – 1.3
	Frag 3 AF	1.3	22	10000	0.6	0.6	1.3
	Frag 7 AB	0.8	73	10000	0.4	0.4	1.4
1CVJ	Frag 1 AMB	2.2	0	1000	1.7	2.0	1.0 – 3.2
	Frag 2 AMF	1.9	2	1000	0.4	1.5	1.0 – 3.2
	Frag 3 AFB	1.8	1	100	0.3	0.9	3.2 – 0.8
	Frag 4 AM	1.3	1	1000	0.2	0.2	0.8
	Frag 5 AFB	1.7	4	100	1.2	1.5	0.8 – 1.4
	Frag 6 AM	1.3	1	1000	0.9	0.9	1.4
2MGZ	Frag 3 AB	1.7	12	10000	0.7	1.5	1.6
	Frag 4 AMB	1.1	6	1000	0.5	0.6	1.6 – 1.4
	Frag 5 AMF	1.0	7	1000	0.7	0.7	1.6 – 1.4
	Frag 6 AF	0.9	11	10000	0.2	0.2	1.4
	Frag 8 AB	2.4	0	10000	0.7	1.0	1.2
	Frag 9 AMB	1.0	16	1000	0.6	0.6	1.2 – 1.4
	Frag 10 AMF	1.2	27	1000	0.7	0.7	1.2 – 1.4
2yh1	Frag 1 AB	1.2	10	10000	0.3	0.3	1.3
	Frag 2 AMB	0.7	13	1000	0.4	0.4	1.3 – 1.7
	Frag 3 AMF	0.8	7	1000	0.2	0.2	1.3 – 1.7
	Frag 4 AFB	0.9	1	100	0.7	0.7	1.7 – 1.8
	Frag 5 AMB	0.8	23	1000	0.3	0 3	1.8 – 1.3
	Frag 6 AMF	1.4	28	1000	0.4	1.1	1.8 – 1.3
3NNH	Frag 2 AB	1.6	5	10000	0.4	1.0	1.3
	Frag 3 AM	1.6	9	1000	1.4	1.4	1.3
	Frag 4 AF	1.3	15	10000	0.6	0.6	1.3
	Frag 7 AB	1.6	10	10000	0.4	1.2	1.3
	Frag 8 AMB	1.1	9	1000	0.6	0.6	1.3 – 1.9
4BS2	Frag 3 AB	2.8	0	1000	0.5	2.3	3.2
	Frag 4 AM	2.5	0	1000	0.3	2.0	3.2
	Frag 5 AF	2.5	0	10000	0.5	2.3	3.2
	Frag 6 AB	1.0	16	10000	0.4	0.4	1.4
	Frag 7 AMB	1.0	6	1000	0.5	0.5	1.4 – 1.4
	Frag 8 AMF	1.1	12	1000	0.6	0.6	1.4 – 1.4
4N0T	Frag 2 AB	2.5	0	10000	1.8	2.0	1.4
	Frag 3 AM	1.6	6	1000	1.1	1.1	1.4
	Frag 4 AF	2.2	0	10000	1.2	1.3	1.4
	Frag 8 AB	1.3	69	10000	0.9	1.1	0.8
	Frag 9 AMB	2.3	0	1000	1.5	2.2	0.8 – 1.8
	Frag 10 AMF	2.5	0	1000	1.9	2.1	0 8 – 1.8
	Frag 11 AF	3.1	0	10000	2.2	2.6	1.8
Unbound docking							

1B7F	Frag 1 AMB	1.3	3	1000	0.9	0.9	1.3 – 0.9
	Frag 2 AMF	1.0	14	1000	0.8	0.8	1.3 – 0.9
	Frag 3 AF	1.0	15	10000	0.6	0.6	1.3
	Frag 7 AB	2.1	0	10000	0.4	1.7	1.3
2YH1	Frag 1 AB	1.6	14	10000	0.3	0.3	2.0
	Frag 2 AMB	0.8	13	1000	0.4	0.7	2.0 – 2.9
	Frag 3 AMF	1.3	6	1000	0.2	0.8	2.0 – 2.9
	Frag 4 AFB	1.0	1	1000	0.7	0.7	2.9 – 1.7
	Frag 5 AMB	1.3	15	1000	0.3	0.3	1.7 – 2.1
	Frag 6 AMF	1.1	23	1000	0.4	0.4	1.7 – 2.1

* The best docked conformer is also the closest conformer in library after perfect fitting.

Table S6. Closest pose and number of good solutions in the top-scored docking poses for each sub- anchored fragments, compared to the RMSDs of the closest sub-anchor (2 nucleotides), and of the closest conformer in library after perfect fitting.

	Docking run	Docking					Fitting
		Min RMSD (Å)	Hits	Nb poses	closest sub-anchor (Å)	Nb sub-anchors	Min RMSD (Å)
1B7F	Frag 4 SF	0.9	321	2.10^6	1.4	5183	0.5
	Frag 6 SB	1.2	413	2.10^6	0.8	5109	1.0
2MGZ	Frag 2 SB	1.1	71	2.10^6	1.5	6680	0.4
	Frag 7 SF	1.2	10	2.10^6	2.6	5735	0.3
3NNH	Frag 1 SB	3.4	0	2.10^6	1.9	5370	0.7
	Frag 5 SF	1.6	9	2.10^6	1.3	5127	0.7
	Frag 6 SB	1.2	6	2.10^6	1.9	5441	0.5
4BS2	Frag 2 SB	2.6	0	2.10^6	2.9	7027	0.2
4N0T	Frag 1 SB	2.3	0	2.10^6	2.9	5695	1.4
	Frag 5 SF	2.5	0	2.10^6	2.6	5688	1.7
	Frag 7 SB	0.9	7273	2.10^6	1.0	5688	0.8

Table S7. Closest pose and number of good solutions in the top-scored docking poses for each (sub-) sub-sub-anchored fragments, compared to the RMSDs of the closest sub-anchor and of the closest conformer in library after perfect fitting.

		Docking						Fitting
	Docking run	Min RMSD(Å)	Hits	Near-hits	Nb poses	Closest sub-anchor(Å)	Nb sub-anchors	Min RMSD (Å)
1B7F	Frag 5 SSF	1.1	122	665	2.10^6	1.1	10^4	0.9
	Frag 5 SSB	1.2	129	507	2.10^6	0.8	10^4	0.9
2MGZ	Frag 1 SSB	2.1	0	19	2.10^6	1.3	10^4	0.3
4BS2	Frag 1 SSB	4.2	0	0	2.10^6	4.3	10^4	0.2
4N0T	Frag 6 SSF	3.1	0	0	1.10^6	4.2	10^4	0.8
	Frag 6 SSB	1.6	267	13750	1.10^6	1.0	10^4	0.8

Table S8. Results of all-atoms chains building, before and after refinement of the 1000 top-ranked chains.

	Nucl.	Nb chains	Before refinement				After refinement				% improved models
			Min RMSD (Å)	% with RMSD			Min RMSD (Å)	% with RMSD			
				≤2 Å	≤3 Å	≤4 Å		≤3 Å	≤4 Å	≤5 Å	
1b7f	2-9	1000	1.6	3	16	51	1.7	3	18	56	66
1cvj	2-7	1000	1.8	11	99	99	1.8	14	99	99	94
2mgz	5-11	1000	1.4	9	13	89	1.5	5	13	93	50
2yh1	3-7	1000	0.8	11	14	24	1.0	12	15	42	67
3nnh	4-10	1000	2.8	0	0	27	2.7	0	0	30	82
4bs2	5-9	909	1.8	3	11	13	1.7	2	12	13	52
Unbound docking											
1b7f	2-9	1000	1.6	4	16	52	1.7	4	19	58	68
2yh1	2-7	1000	1.0	13	36	67	1.1	12	38	76	77

```
                 7574566       67  8 559   7            7  6 6          7576698 7
4n0t  1  --TECTLWMTNFPPSY----TQRNIRDLLQDINV----------VALSIRLPS-LRF--NTSRRFAYIDV   51
1cvj  1  ----ASLYVGDLHPDV----TEAMLYEKFSPAG-----------PILSIRVCRDMIT--RRSLGYAYVNF   49
4bs2  1  -----DLIVLGLPWKT----TEQDLKEYFSTFG-----------EVLMVQVKKDLKT--GHSKGFGFVRF   48
2mgz  1  ----RRLHVSNIPFKY----REPDLTAMFEKVG-----------PVVDVEIIFNE----RGSKGFGFVTM   47
2yh1  1  ---ARRLYVGNIPFGI----TEEAMMDFFNAQMRLGGLTQAPGNPVLAVQIN--------QDKNFAFLEF   55
1b7f  1  --SNTNLIVNYLPQDM----TDRELYALFRAIG-----------PINTCRIMRDYKT--GYSYGYAFVDF   51
3nnh  1  --DAIKMFVGQVPRTW----SEKDLRELFEQYG-----------AVYEINVLRDRSQNPPQSKGCCFVTF   53

                     6         6   75 5  5 6                              5556
4n0t  52  TSKEDARYCVEKLNG-LKIE--GYTLVTKVSN------PLEKSKR-------TDSATLEGREIMIRNLST  105
1cvj  50  QQPADAERALDTMNF-DVIK--GKPVRIMWSQRD------------------PSLRKSGVGNIFIKNLDK   98
4bs2  49  TEYETQVKVMSQR---HMID--GRWCDCKLPNSK--------------QSQDE---PLRSRKVFVGRCTE   96
2mgz  48  QNPDDADRARAEFNG-TTIE--GRRVEVNLATQ-------------------------RVTKIFVGGLPY   89
2yh1  56  RSVDETTQAM-AFDG-IIFQ--GQSLKIRRPHDYQPLPGMSENPSVYVPGVVSTVVPDSAHKLFIGGLPN  121
1b7f  52  TSEMDSQRAIKVLNG-ITVR--NKRLKVSYARPG-------------GES-------IKDTNLYVTNLPR   98
3nnh  54  YTRKAALEAQNALHNMKVLPGMHHPIQMKPADS-------------------------DAIKMFVGQVPR   98

               5 5 6  57 5    5           55  75 5     5  5
4n0t 106  ELLDENLLRESFEGFGSIEKINIPAGQK-EHSFNNCCAFMVFENKDSAERALQ-MNR-SLL--GNREISV  170
1cvj  99  SI-DNKALYDTFSAFGNILSCKVVCD----ENGSKGYGFVHFETQEAAERAIEKMNG-MLL--NDRKVFV  160
4bs2  97  DM-TEDELREFFSQYGDVMDVFIPK-------PFRAFAFVTFADDQIAQSLCGED---LII--KGISVHI  153
2mgz  90  HT-SDKTLHEYFEQFGDIEEAVVITDR--NTQKSRGYGFVTMKDRASAERACK-DPN-PII--DGRKANV  152
2yh1 122  YL-NDDQVKELLTSFGPLKAFNLVKDS--ATGLSKGYAFCEYVDINVTDQAIAGLNG-MQL--GDKKLLV  185
1b7f  99  TI-TDDQLDTIFGKYGSIVQKNILRDK--LTGRPRGVAFVRYNKREEAQEAISALNN-VIPEGGSQPLSV  164
3nnh  99  TW-SEKDLRELFEQYGAVYEINVLRDRSQNPPQSKGCCFVTFYTRKAALEAQNALHNMKVLPGMHHPIQM  167

               7
4n0t 171  SLADK----  175
1cvj 161  GRFKSRKER  169
4bs2 154  SNAEPK---  159
2mgz 153  NLAYLGAK-  160
2yh1 186  QRASVG---  191
1b7f 165  RLA------  167
3nnh 168  KPAD-----  171
```

Figure S1. Sequence comparison of the binding sites of the proteins in the benchmark. The sequence structurally aligned with PROMALS3D (1) are coloured according to the number of proteins were it binds RNA (contact cutoff 4 A). The conservation score *scorescons* (1) is reported on the 1st line.

(1) Pei, J. and Grishin, N. V. (2014) PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information. *Methods Mol. Biol. Clifton NJ*, **1079**, 263–271.

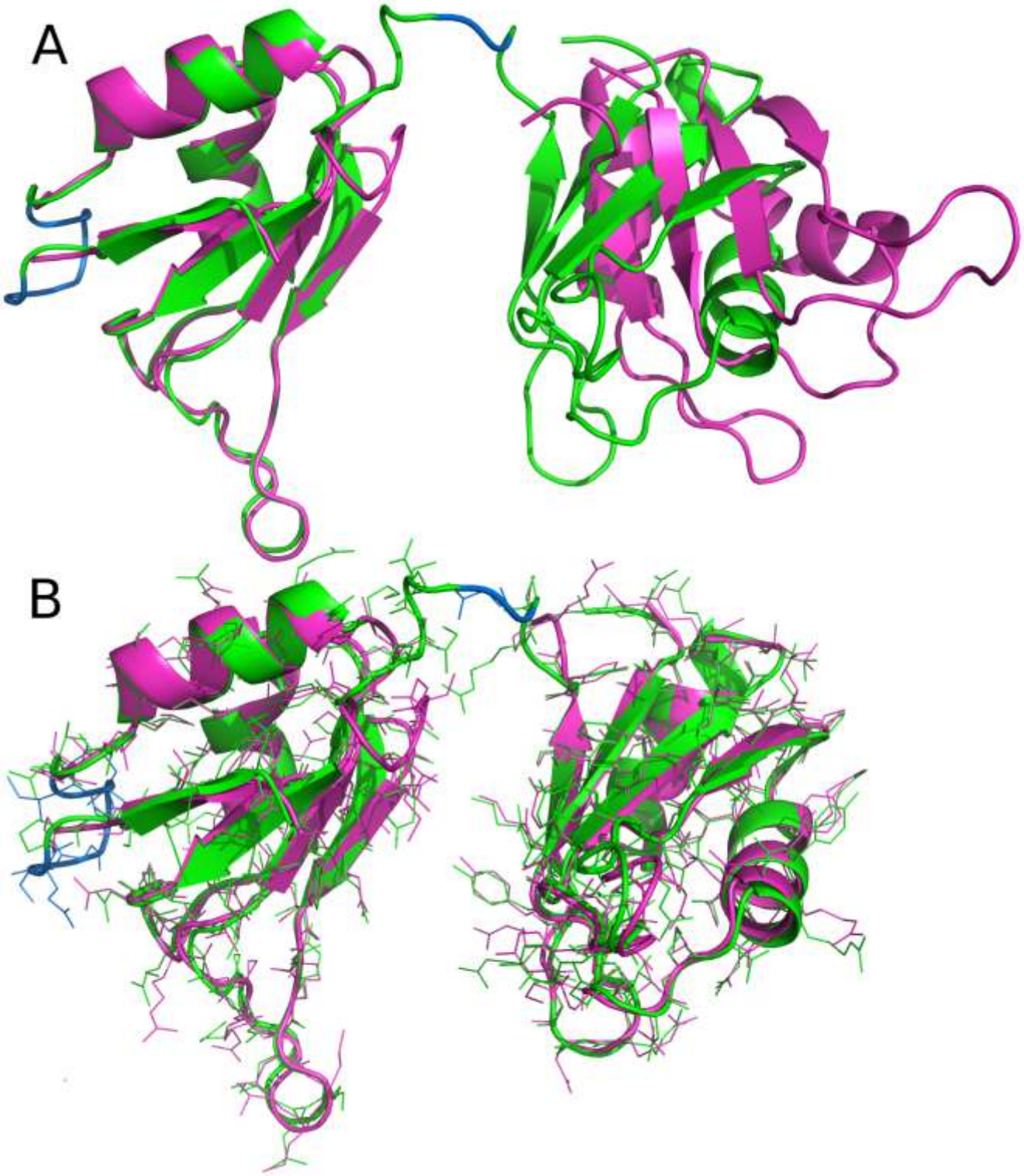

Figure S2. Differences between RNA-bound and unbound structures 1B7F and 3SXL. The bound and unbound structures are represented in green and magenta respectively. The parts resolved only in the bound structure are distinguished in blue. (A) The structures are aligned on their RMM1 domain. (B) Teach RRM domain of the unbound structure are aligned on the corresponding RRM domain of the unbound structure.