Supplementary Data

A novel protocol for three-dimensional structure prediction of RNA-protein complexes

Yangyu Huang[1†], ShiyongLiu[1†], Dachuan Guo[1], Lin Li[2] and Yi Xiao[1*]

[1]Biomolecular Physics and Modeling Group, Department of Physics
Huazhong University of Science and Technology, Wuhan 430074, Hubei, China
[2]Computational Biophysics and Bioinformatics, Department of Physics and Astronomy
Kinard Laboratory Building, Clemson University, SC 29634, USA

[†] These authors made equal contributions.

[*] Corresponding author.

Supplementary Table S1 Protein-RNA complexes in our testing set I (66)

PDB id	Protein Chain	RNA Chain	PDB id	Protein Chain	RNA Chain	PDB id	Protein Chain	RNA Chain
1ASY	AB	R	1K8W	A	B	2BGG	A	PQ
1DFU	P	MN	1KOG	A	I	2BH2	A	C
1OB2	A	B	1KQ2	ABHIKM	R	2BTE	A	B
1R3E	A	C	1M5O	C	AB	2BU1	A	R
2FMT	A	C	1M8V	ABCDEFG	O	2CZJ	A	B
1B23	P	R	1M8W	A	C	2ERR	A	B
1MFQ	C	A	1MMS	A	C	2F8K	A	B
1QTQ	A	B	1N78	A	C	2GJE	AD	RS
1U0B	B	A	1Q2R	A	E	2GJW	AB	EFH
1B7F	A	P	1SER	AB	T	2HGH	A	B
1C9S	LMNOPQRSTUV	W	1T0K	AB	CD	2HW8	A	B
1DK1	A	B	1T4L	B	A	2I91	B	EF
1E7K	A	C	1U63	A	B	2IX1	A	B
1EC6	A	D	1WNE	A	BC	2PY9	A	E
1EKZ	A	B	1WPU	A	C	2QUX	AB	C
1F7U	A	B	1WSU	A	E	2R7R	A	X
1G1X	A	D	1YVP	A	CD	3BO2	A	BCDE
1H3E	A	B	2AD9	A	B	3BSB	B	C
1H4S	AB	T	2ADB	A	B	3BSO	A	PT
1HC8	A	C	2ADC	A	B	3BSX	A	C
1HVU	AB	C	2ASB	A	B	3BX2	A	C
1JBR	B	D	2AZ0	AB	CD	3CIY	A	CD

Supplementary Table S2 Protein-RNA complexes in our testing set II (72)

PDB id	Protein Chain	RNA Chain	PDB id	Protein Chain	RNA Chain	PDB id	Protein Chain	RNA Chain
1C0A	A	B	2AKE	A	B	3OVB	A	C
1DFU	P	MN	2ANR	A	B	1F7U	A	B
1E8O	CD	E	2AZ0	AB	CD	1IL2	A	C
1F7Y	A	B	2BH2	A	C	1R9F	A	BC
1FFY	A	T	2CSX	B	D	1RC7	A	BCDE
1G1X	FH	IJ	2CZJ	E	F	1SER	AB	T
1GAX	B	D	2DU3	A	D	1UN6	C	E
1H4S	AB	T	2FK6	A	R	2BTE	D	E
1HQ1	A	B	2GJW	AB	EFH	2FMT	A	C
1J1U	A	B	2QUX	DE	F	2NUG	AB	CDEF
1JBS	A	C	2RFK	A	DE	2UWM	A	C
1JID	A	B	2XDB	A	G	2VPL	A	B
1K8W	A	B	2ZM5	A	C	2ZKO	AB	CD
1KOG	CD	K	2ZNI	AB	C	2ZZM	A	B
1LNG	A	B	2ZUE	A	B	3ADD	A	C
1MMS	A	C	3CIY	A	CD	3FTF	A	CD
1N78	B	D	3DD2	H	B	3HL2	C	E
1Q2R	C	F	3EPH	A	E	1H3E	A	B
1QTQ	A	B	3FOZ	A	C	1OOA	B	D
1R3E	A	CDE	3HHZ	O	R	1U0B	B	A
1S03	H	A	3LRR	A	CD	2HW8	A	B
1SJ3	P	R	3LWR	ABC	DE	2IPY	A	C
1T0K	B	CD	3MOJ	B	A	2R8S	HL	R
1YVP	B	EF	3OL9	M	NO	2V3C	C	M

Supplementary Table S3 Success rates of different protocols over three categories of the testing set II. For different protocols, top 1000 structures are generated by corresponding docking procedures and evaluated by corresponding scoring functions. For N_P=10, the success rates are compared.

	easy	medium	hard
Number of complexes	49	16	7
RPDOCK & DECK-RP	49%	13%	0%
GRAMM & DARS-RNP	35%	13%	0%
FTDock & Li	16%	0%	0%

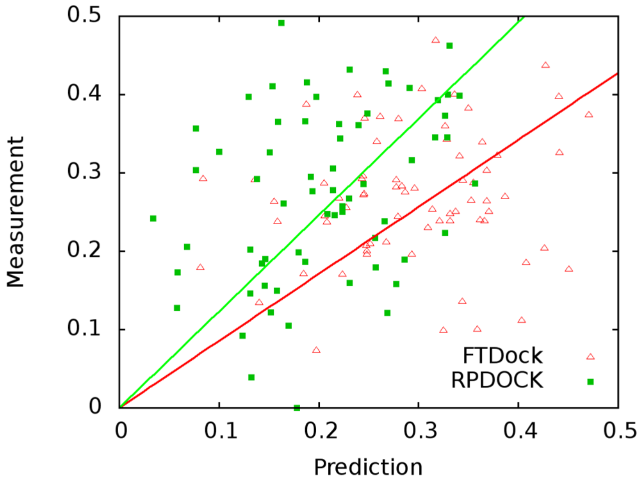
Supplementary Table S4 Protein-RNA complexes in our training set (118)

PDB id	Protein Chain	RNA chain	PDB id	Protein Chain	RNA chain	PDB id	Protein Chain	RNA chain
1A9N	AB	Q	2JLU	A	C	3MOJ	B	A
1AV6	A	B	2PJP	A	B	3MQK	A	DE
1C0A	A	B	2R8S	HL	R	3NMR	A	B
1CVJ	A	M	2V3C	C	M	3NMU	AG	DI
1CX0	A	B	2VNU	D	B	3O3I	X	A
1DI2	AB	CD	2XD0	Y	V	3OIJ	AB	C
1DUL	A	B	2XLI	A	B	3OL6	A	BC
1FEU	A	BC	2XS2	A	B	3Q0N	A	C
1FFY	A	T	2XZO	A	D	3QJJ	B	R
1GAX	A	C	2Y8W	A	B	3R2C	A	R
1H2C	A	R	2YKG	A	CD	3RC8	A	E
1J1U	A	B	2ZI0	A	CD	3RW6	A	H
1JID	A	B	2ZKO	AB	CD	3SIU	AB	C
1L9A	A	B	3FOZ	AB	C	3SN2	A	B
1LNG	A	B	2ZZM	A	B	3SNP	A	C
1N35	A	BC	2ZZN	A	C	3T5N	A	C
1OOA	A	C	3A6P	A	DE	3TRZ	A	U
1R9F	A	BC	3ADB	A	C	3V7E	A	C
1YYK	AB	CDEF	3AEV	AB	C	4E78	A	PT
1RLG	A	C	3HSB	ABCDEF	X	4ERD	AB	CD
1S03	H	A	3AMT	A	B	1A1T	A	B
1SI3	A	B	3AVU	A	TG	1AUD	A	B
1TFW	A	GJ	3AVY	A	GT	1FJE	B	A
1VFG	A	C	3DD2	H	B	1K1G	A	B
1YTY	AB	CD	3DH3	B	F	1L1C	AB	C

1YYW	A	EFGH	3EPH	A	E	1RGO	A	D
1ZBH	AD	F	3EQT	AB	CD	1RKJ	A	B
2ANN	A	B	3FTE	A	CD	2CJK	A	B
2AZX	A	C	3G0H	A	E	2FY1	A	B
2B3J	AB	E	3GIB	ABC	H	2JPP	AB	C
2CSX	A	C	3HL2	AB	E	2KFY	A	B
2DB3	A	E	3I5X	A	B	2KH9	A	B
2DR5	A	B	3IAB	AB	R	2L3C	A	B
2DU3	A	D	3ICE	ABCDEF	G	2LEB	A	B
2EZ6	AB	CD	3IEV	A	D	2LI8	A	B
2F8S	A	CD	3K5Q	A	B	2RQC	A	B
2FK6	A	R	3KS8	AB	EF	2XFM	A	B
2GXB	A	EF	3L25	ABDE	CF	2YH1	A	B
2HVY	ABCD	E	3LRR	A	CD			
2I82	A	E	3MDG	AB	C			

Supplementary Table S5 Classification of amino acids based on the dipoles and volumes of the side chains (based on Shen, J. et al. *Proc Natl Acad Sci U S A* **104**, 4337-41 (2007))

Type	Residues
1	ALA，GLY，VAL
2	ILE, LEU, PHE, PRO
3	TYR, MET, THR, SER
4	HIS, ASN, GLN, TRP
5	ARG, LYS
6	ASP, GLU
7	CYS



Supplementary Figure S1 The comparison of the fractions of the fully buried atoms in the interface atoms for the top-one complex structures predicted by FTDock and RPDOCK with those for the measured ones in the bound training set.