



Full wwPDB NMR Structure Validation Report ⓘ

Mar 27, 2026 – 09:44 AM UTC

PDB ID : 7Q4L / pdb_00007q4l
BMRB ID : 34675
Title : The solution structure of hsDND1 RRM12 bound to CUUAUUUG RNA
Authors : Duszczuk, M.M.; Allain, F.H.T.
Deposited on : 2021-10-31

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
wwPDB-RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
wwPDB-ShiftChecker : v1.2
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

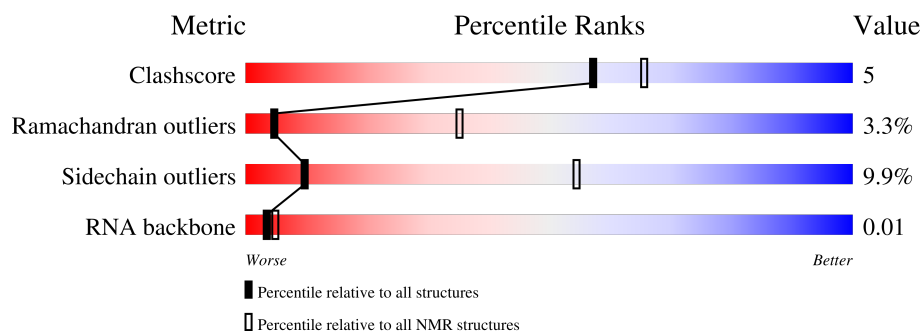
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 71%.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	229148	14424
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823
RNA backbone	8273	777

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	227	 71% 12% 15%
2	B	8	 12% 38% 50%

2 Ensemble composition and analysis

This entry contains 20 models. Model 13 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:15-A:43, A:58-A:92, A:97-A:224 (192)	0.94	13

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 4 clusters and 2 single-model clusters were found.

Cluster number	Models
1	1, 2, 5, 6, 8, 11, 13, 14
2	7, 9, 10, 12, 15
3	3, 16, 19
4	4, 18
Single-model clusters	17; 20

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3796 atoms, of which 1881 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Dead end protein homolog 1.

Mol	Chain	Residues	Atoms						Trace
1	A	227	Total	C	H	N	O	S	0
			3551	1114	1798	328	303	8	

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	GLY	-	expression tag	UNP Q8IYX4
A	10	ALA	-	expression tag	UNP Q8IYX4
A	11	MET	-	expression tag	UNP Q8IYX4

- Molecule 2 is a RNA chain called RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3').

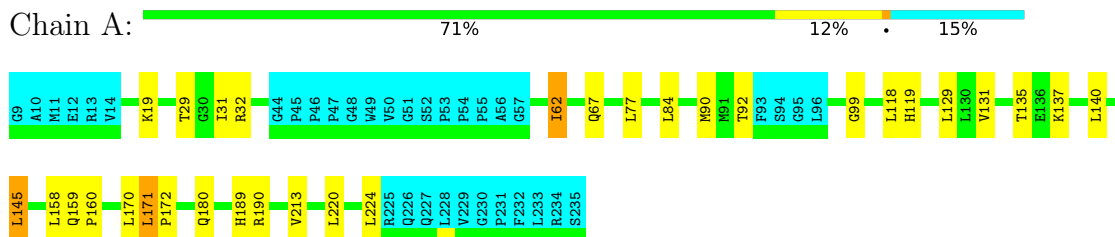
Mol	Chain	Residues	Atoms						Trace
2	B	8	Total	C	H	N	O	P	0
			245	74	83	23	58	7	

4 Residue-property plots

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: Dead end protein homolog 1



- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')

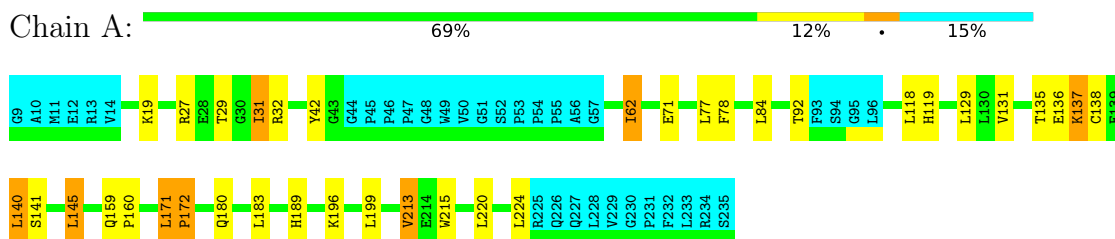


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

- Molecule 1: Dead end protein homolog 1



- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')

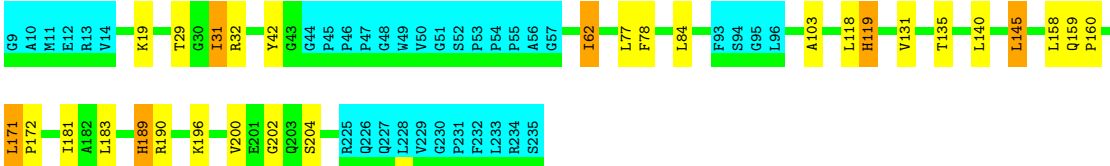
Chain B: 



4.2.2 Score per residue for model 2

- Molecule 1: Dead end protein homolog 1

Chain A: 



- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')

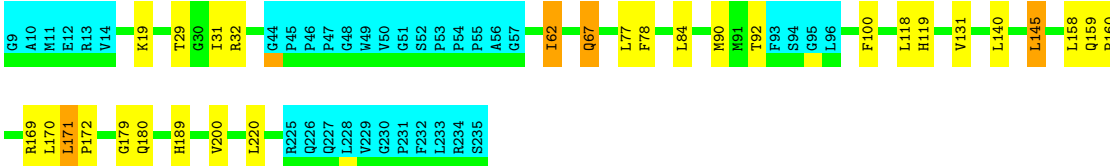
Chain B: 



4.2.3 Score per residue for model 3

- Molecule 1: Dead end protein homolog 1

Chain A: 



- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')

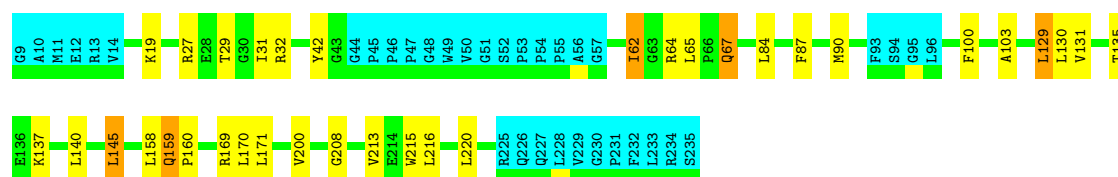
Chain B: 



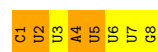
4.2.4 Score per residue for model 4

- Molecule 1: Dead end protein homolog 1

Chain A: 

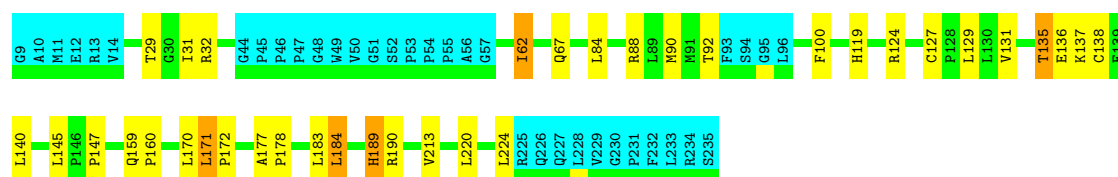


- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')

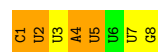


4.2.5 Score per residue for model 5

- Molecule 1: Dead end protein homolog 1

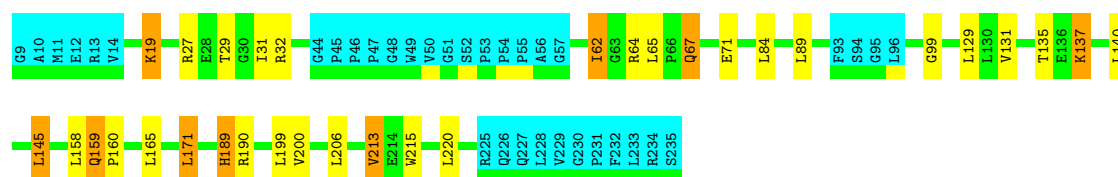


- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')



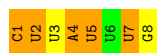
4.2.6 Score per residue for model 6

- Molecule 1: Dead end protein homolog 1



- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')

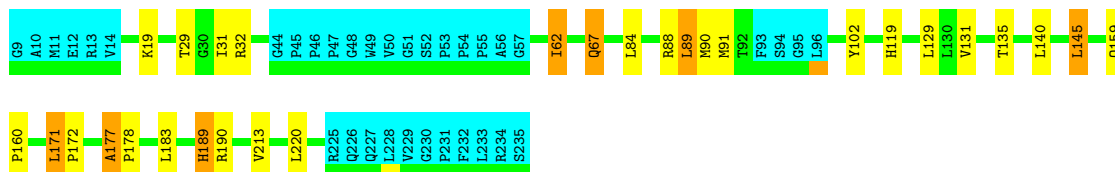




4.2.7 Score per residue for model 7

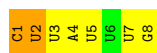
- Molecule 1: Dead end protein homolog 1

Chain A: 72% 10% 15%



- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')

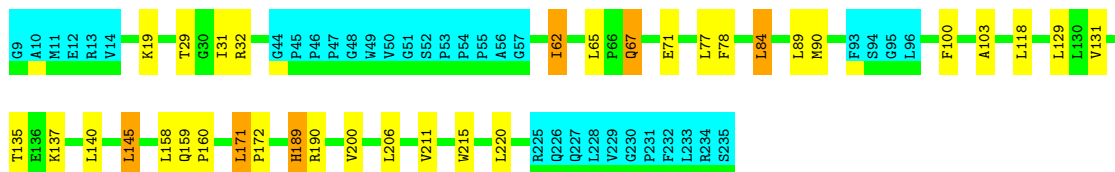
Chain B: 12% 62% 25%



4.2.8 Score per residue for model 8

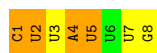
- Molecule 1: Dead end protein homolog 1

Chain A: 70% 12% 15%



- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')

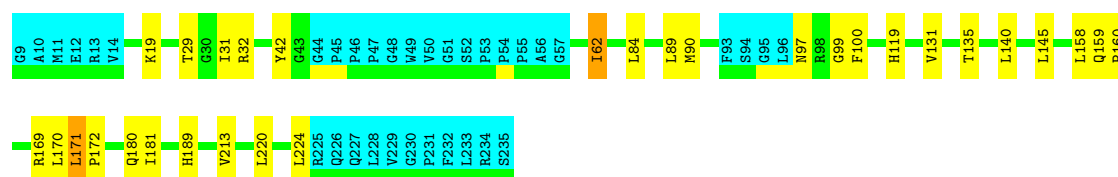
Chain B: 12% 38% 50%



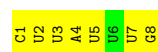
4.2.9 Score per residue for model 9

- Molecule 1: Dead end protein homolog 1

Chain A: 71% 12% 15%

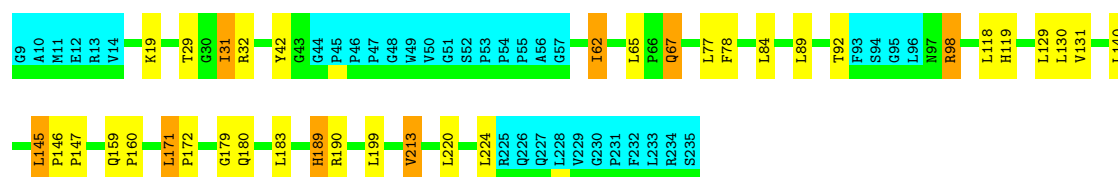


- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')

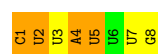


4.2.10 Score per residue for model 10

- Molecule 1: Dead end protein homolog 1

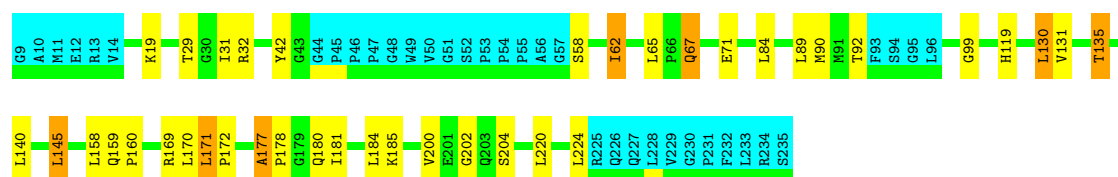


- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')



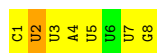
4.2.11 Score per residue for model 11

- Molecule 1: Dead end protein homolog 1



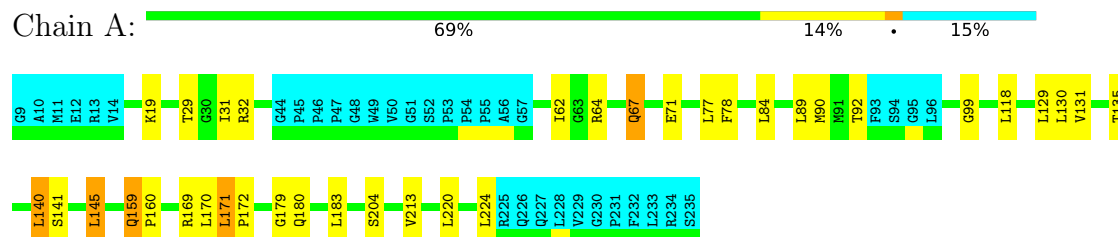
- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')





4.2.12 Score per residue for model 12

- Molecule 1: Dead end protein homolog 1

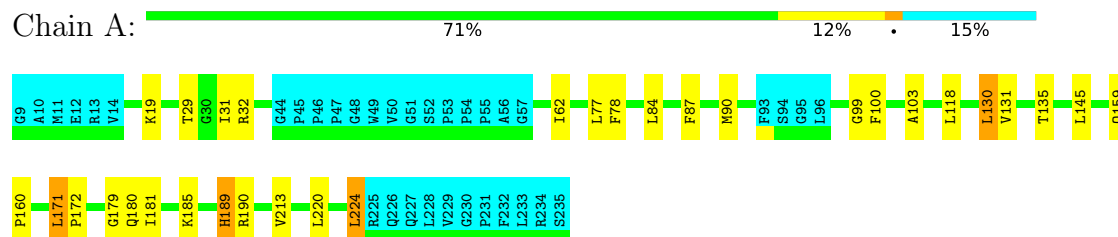


- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')



4.2.13 Score per residue for model 13 (medoid)

- Molecule 1: Dead end protein homolog 1



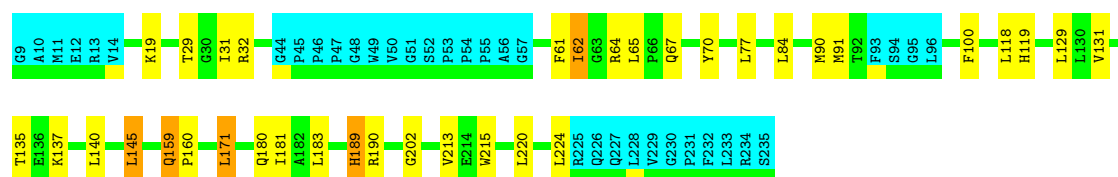
- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')



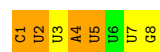
4.2.14 Score per residue for model 14

- Molecule 1: Dead end protein homolog 1



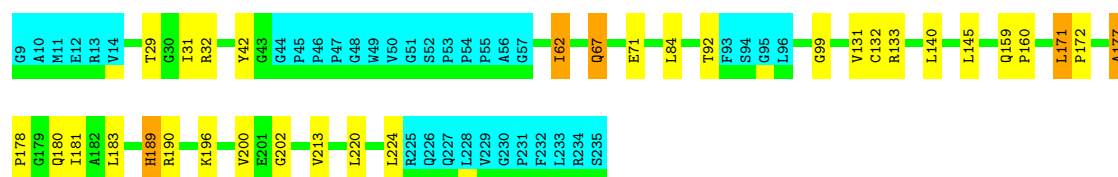


- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')

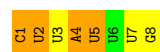


4.2.15 Score per residue for model 15

- Molecule 1: Dead end protein homolog 1

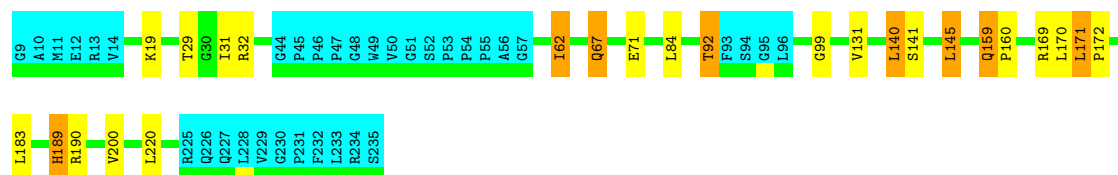


- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')



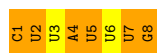
4.2.16 Score per residue for model 16

- Molecule 1: Dead end protein homolog 1



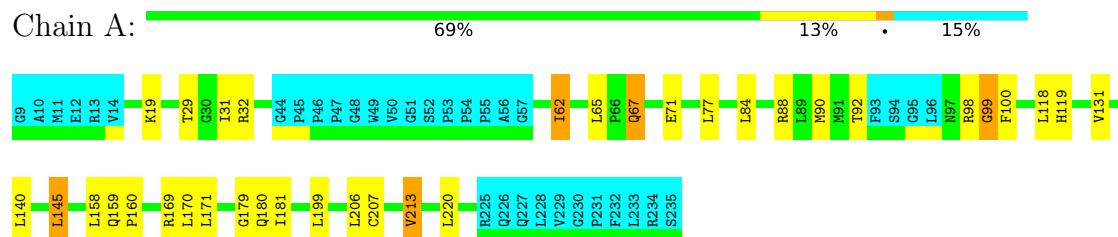
- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')





4.2.17 Score per residue for model 17

- Molecule 1: Dead end protein homolog 1

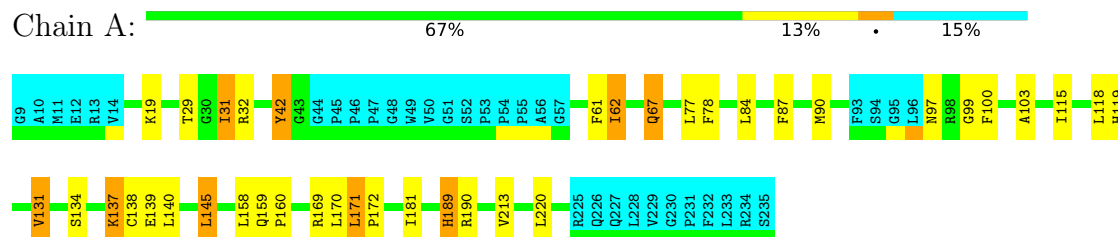


- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')

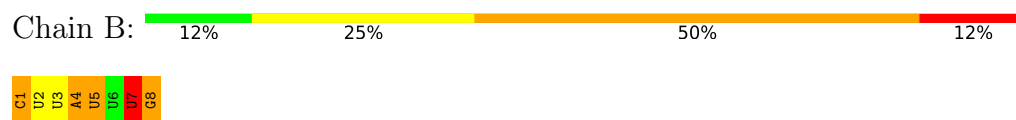


4.2.18 Score per residue for model 18

- Molecule 1: Dead end protein homolog 1



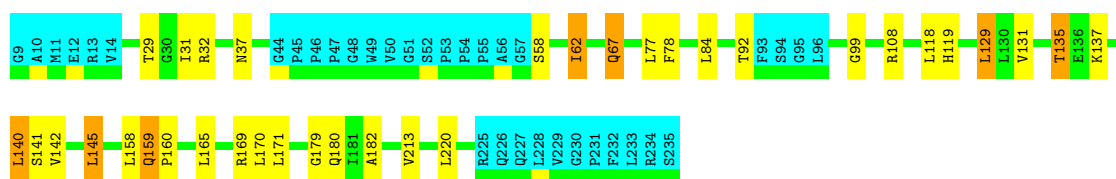
- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')



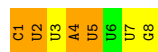
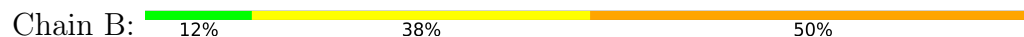
4.2.19 Score per residue for model 19

- Molecule 1: Dead end protein homolog 1



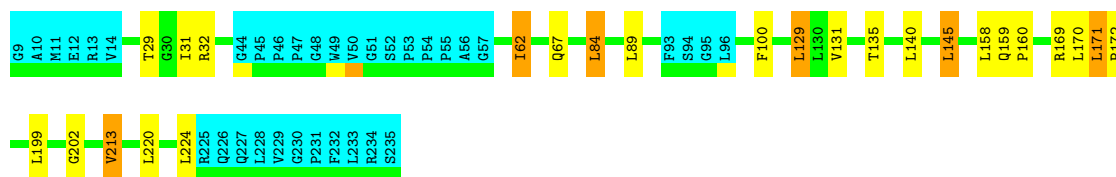
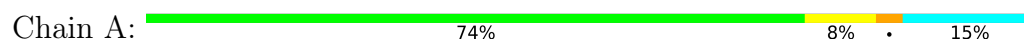


- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')

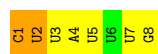
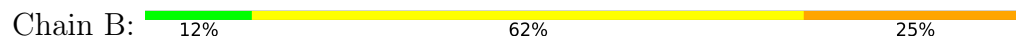


4.2.20 Score per residue for model 20

- Molecule 1: Dead end protein homolog 1



- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*UP*UP*UP*G)-3')



5 Refinement protocol and experimental data overview

The models were refined using the following method: *torsion angle dynamics*.

Of the 50 calculated structures, 20 were deposited, based on the following criterion: *structures with the least restraint violations and lowest energy*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CYANA	structure calculation	3.98
Amber	structure calculation	12
Amber	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	2300
Number of shifts mapped to atoms	2300
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	71%

6 Model quality i

6.1 Standard geometry i

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.57±0.00	0±0/1533 (0.0± 0.0%)	0.98±0.01	1±1/2078 (0.0± 0.0%)
2	B	0.75±0.01	0±0/179 (0.0± 0.0%)	1.21±0.03	0±0/276 (0.1± 0.2%)
All	All	0.59	0/34240 (0.0%)	1.01	18/47080 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
2	B	0.0±0.0	0.2±0.4
All	All	0	4

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
2	B	7	U	C5'-C4'-C3'	-5.81	106.49	115.20	16	4
1	A	135	THR	N-CA-C	-5.53	105.69	113.20	11	1
1	A	119	HIS	CB-CG-CD2	-5.47	124.09	131.20	14	12
2	B	7	U	O4'-C1'-N1	5.01	115.72	108.20	18	1

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
2	B	6	U	Sidechain	2
2	B	7	U	Sidechain	1
2	B	2	U	Sidechain	1

6.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1499	1543	1543	14±3
2	B	162	83	85	3±2
All	All	33220	32520	32560	318

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:137:LYS:HE3	1:A:215:TRP:CZ2	0.64	2.28	8	2
1:A:137:LYS:HZ3	1:A:215:TRP:CG	0.60	2.14	1	1
1:A:177:ALA:N	1:A:178:PRO:CD	0.59	2.66	11	4
1:A:177:ALA:H	1:A:178:PRO:CD	0.58	2.10	15	2
1:A:90:MET:HE3	1:A:100:PHE:CE2	0.57	2.35	13	8
1:A:177:ALA:N	1:A:178:PRO:HD3	0.57	2.15	7	1
2:B:4:A:C2	2:B:5:U:C4	0.56	2.93	18	13
1:A:90:MET:HG3	1:A:100:PHE:CZ	0.56	2.36	14	1
1:A:189:HIS:CG	1:A:190:ARG:N	0.55	2.75	13	11
1:A:19:LYS:H	1:A:19:LYS:HD2	0.53	1.63	6	1
2:B:6:U:H3'	2:B:7:U:H5''	0.53	1.80	2	2
1:A:137:LYS:HE2	1:A:215:TRP:CE3	0.52	2.39	4	2
1:A:220:LEU:HD13	1:A:220:LEU:C	0.52	2.29	4	16
1:A:159:GLN:N	1:A:160:PRO:HD2	0.52	2.20	7	20
2:B:1:C:H2'	2:B:2:U:C5	0.52	2.39	17	16
1:A:224:LEU:C	1:A:224:LEU:HD13	0.51	2.30	13	1
1:A:137:LYS:HE3	1:A:215:TRP:CE2	0.51	2.41	14	1
1:A:92:THR:HG22	2:B:8:G:C5	0.50	2.41	16	1
1:A:159:GLN:N	1:A:160:PRO:CD	0.50	2.75	7	20
1:A:220:LEU:C	1:A:220:LEU:HD13	0.50	2.30	18	3
2:B:1:C:C5	2:B:2:U:C4	0.50	2.99	13	1
1:A:77:LEU:HD11	1:A:118:LEU:CD2	0.49	2.37	19	11
1:A:67:GLN:H	1:A:67:GLN:CD	0.49	2.15	7	13
1:A:62:ILE:H	1:A:62:ILE:HD13	0.49	1.68	20	18
1:A:78:PHE:CE1	1:A:118:LEU:CD1	0.49	2.96	18	9
1:A:137:LYS:HE2	1:A:215:TRP:CD2	0.49	2.42	6	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:140:LEU:HD23	1:A:141:SER:H	0.48	1.68	1	4
1:A:84:LEU:HD21	1:A:103:ALA:HB1	0.48	1.85	2	2
2:B:4:A:C2	2:B:5:U:N3	0.48	2.81	10	4
2:B:7:U:O2'	2:B:8:G:C8	0.48	2.66	18	1
1:A:78:PHE:HB3	1:A:84:LEU:HD13	0.47	1.86	19	1
1:A:171:LEU:CB	1:A:172:PRO:CD	0.47	2.92	15	9
1:A:177:ALA:H	1:A:178:PRO:HD3	0.47	1.70	11	2
1:A:169:ARG:C	1:A:170:LEU:HD12	0.47	2.34	9	10
1:A:171:LEU:HB3	1:A:172:PRO:CD	0.47	2.40	15	10
1:A:137:LYS:HZ3	1:A:215:TRP:CD1	0.46	2.29	1	1
2:B:1:C:H5''	2:B:1:C:C6	0.46	2.46	15	4
1:A:137:LYS:HE3	1:A:215:TRP:CH2	0.46	2.45	8	1
1:A:42:TYR:CD2	1:A:42:TYR:C	0.46	2.94	18	3
1:A:199:LEU:HD12	1:A:213:VAL:CG1	0.45	2.42	1	1
1:A:158:LEU:HD12	1:A:158:LEU:C	0.45	2.36	6	9
1:A:115:ILE:CB	1:A:131:VAL:HG11	0.45	2.41	18	1
1:A:135:THR:HG22	2:B:5:U:O4	0.45	2.11	4	2
1:A:199:LEU:HD12	1:A:213:VAL:HG11	0.45	1.88	17	3
2:B:1:C:C6	2:B:1:C:H5''	0.45	2.46	18	1
1:A:87:PHE:CD2	1:A:103:ALA:HB2	0.44	2.48	18	3
1:A:78:PHE:CB	1:A:84:LEU:HD22	0.44	2.42	19	1
1:A:135:THR:CG2	1:A:137:LYS:HE3	0.44	2.42	4	1
1:A:171:LEU:HD13	1:A:171:LEU:C	0.44	2.38	8	2
1:A:158:LEU:C	1:A:158:LEU:HD12	0.44	2.37	19	2
1:A:61:PHE:CG	2:B:4:A:C6	0.44	3.05	18	1
1:A:42:TYR:C	1:A:42:TYR:CD2	0.43	2.96	15	3
1:A:137:LYS:HD3	2:B:5:U:C4	0.43	2.47	18	1
1:A:89:LEU:HD22	1:A:91:MET:HE3	0.43	1.90	7	1
1:A:196:LYS:HZ1	1:A:215:TRP:CD1	0.43	2.31	1	1
1:A:98:ARG:O	1:A:99:GLY:C	0.43	2.62	17	1
1:A:171:LEU:HB3	1:A:172:PRO:HD3	0.43	1.91	2	6
1:A:206:LEU:C	1:A:206:LEU:HD13	0.43	2.38	17	2
1:A:137:LYS:C	1:A:139:GLU:H	0.43	2.20	18	1
2:B:4:A:H1'	2:B:5:U:C6	0.43	2.49	10	3
1:A:31:ILE:HG13	1:A:42:TYR:CD1	0.42	2.48	2	4
1:A:136:GLU:C	1:A:138:CYS:H	0.42	2.21	1	2
1:A:134:SER:HA	2:B:4:A:H61	0.42	1.74	18	1
1:A:171:LEU:C	1:A:171:LEU:HD13	0.42	2.40	6	3
1:A:90:MET:CG	1:A:100:PHE:CZ	0.42	3.02	14	1
1:A:77:LEU:HD11	1:A:118:LEU:HD22	0.42	1.91	10	6
1:A:130:LEU:CD1	2:B:2:U:H2'	0.42	2.45	11	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:62:ILE:HD12	1:A:78:PHE:CE1	0.42	2.50	18	1
1:A:184:LEU:HD12	1:A:184:LEU:H	0.42	1.74	5	1
1:A:61:PHE:CD1	2:B:4:A:C2	0.42	3.08	18	1
1:A:90:MET:HE1	1:A:102:TYR:CE1	0.42	2.50	7	1
1:A:199:LEU:HB2	1:A:213:VAL:HG21	0.42	1.90	20	1
1:A:61:PHE:CB	2:B:4:A:C6	0.42	3.03	14	1
1:A:132:CYS:SG	1:A:133:ARG:N	0.42	2.93	15	1
2:B:4:A:H1'	2:B:5:U:C5	0.41	2.50	12	1
1:A:142:VAL:HB	1:A:182:ALA:HB3	0.41	1.91	19	1
1:A:62:ILE:HD12	1:A:78:PHE:CZ	0.41	2.50	3	1
1:A:206:LEU:HD23	1:A:206:LEU:C	0.41	2.41	6	1
1:A:129:LEU:C	1:A:129:LEU:HD23	0.41	2.40	14	1
1:A:146:PRO:CB	1:A:147:PRO:HD2	0.41	2.45	10	1
1:A:115:ILE:HB	1:A:131:VAL:HG11	0.41	1.92	18	1
1:A:171:LEU:HD13	1:A:172:PRO:N	0.41	2.30	8	1
1:A:100:PHE:CD2	1:A:100:PHE:C	0.41	2.98	20	1
1:A:129:LEU:HD23	1:A:129:LEU:C	0.41	2.40	8	1
1:A:137:LYS:HE2	1:A:189:HIS:ND1	0.41	2.31	8	1
1:A:130:LEU:HD11	2:B:2:U:C4'	0.40	2.46	13	1
2:B:1:C:H2'	2:B:2:U:C6	0.40	2.51	12	1
1:A:92:THR:HG23	1:A:98:ARG:HB2	0.40	1.94	10	1
1:A:224:LEU:C	1:A:224:LEU:HD23	0.40	2.41	20	1

6.3 Torsion angles [i](#)

6.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	192/227 (85%)	168±3 (88±1%)	17±3 (9±1%)	6±2 (3±1%)	5	35
All	All	3840/4540 (85%)	3370 (88%)	342 (9%)	128 (3%)	5	35

All 24 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	32	ARG	20
1	A	145	LEU	20
1	A	171	LEU	20
1	A	180	GLN	11
1	A	99	GLY	10
1	A	71	GLU	8
1	A	129	LEU	7
1	A	179	GLY	6
1	A	202	GLY	5
1	A	204	SER	3
1	A	177	ALA	3
1	A	137	LYS	2
1	A	58	SER	2
1	A	172	PRO	1
1	A	119	HIS	1
1	A	208	GLY	1
1	A	124	ARG	1
1	A	127	CYS	1
1	A	135	THR	1
1	A	147	PRO	1
1	A	207	CYS	1
1	A	138	CYS	1
1	A	37	ASN	1
1	A	84	LEU	1

6.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	158/184 (86%)	142±3 (90±2%)	16±3 (10±2%)	10	54
All	All	3160/3680 (86%)	2848 (90%)	312 (10%)	10	54

All 39 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	29	THR	20
1	A	31	ILE	20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	62	ILE	20
1	A	131	VAL	20
1	A	140	LEU	19
1	A	84	LEU	18
1	A	19	LYS	16
1	A	145	LEU	16
1	A	67	GLN	16
1	A	213	VAL	15
1	A	189	HIS	14
1	A	135	THR	13
1	A	92	THR	9
1	A	183	LEU	9
1	A	224	LEU	9
1	A	181	ILE	8
1	A	200	VAL	8
1	A	89	LEU	8
1	A	65	LEU	7
1	A	159	GLN	6
1	A	129	LEU	5
1	A	130	LEU	5
1	A	64	ARG	4
1	A	27	ARG	3
1	A	88	ARG	3
1	A	137	LYS	3
1	A	196	LYS	2
1	A	184	LEU	2
1	A	165	LEU	2
1	A	90	MET	2
1	A	185	LYS	2
1	A	216	LEU	1
1	A	170	LEU	1
1	A	211	VAL	1
1	A	97	ASN	1
1	A	98	ARG	1
1	A	70	TYR	1
1	A	91	MET	1
1	A	42	TYR	1

6.3.3 RNA ⓘ

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers	Suiteness
2	B	8/8 (100%)	6±0 (75±0%)	3±0 (38±3%)	0.02±0.00
All	All	160/160 (100%)	120 (75%)	61 (38%)	0.02

The overall RNA backbone suiteness is 0.01.

All unique RNA backbone outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
2	B	2	U	20
2	B	3	U	20
2	B	4	A	20
2	B	5	U	20
2	B	7	U	20
2	B	8	G	20

All unique RNA pucker outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
2	B	1	C	20
2	B	2	U	20
2	B	7	U	19
2	B	6	U	2

6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.6 Ligand geometry [i](#)

There are no ligands in this entry.

6.7 Other polymers [i](#)

There are no such molecules in this entry.

6.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 71% for the well-defined parts and 70% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *Combined_shifts_all.bmrB*

7.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	2300
Number of shifts mapped to atoms	2300
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	191	0.12 ± 0.08	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	143	0.06 ± 0.11	None needed (< 0.5 ppm)
$^{13}\text{C}'$	0	—	None (insufficient data)
^{15}N	196	0.16 ± 0.18	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 71%, i.e. 2010 atoms were assigned a chemical shift out of a possible 2832. 0 out of 40 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	705/944 (75%)	367/383 (96%)	166/384 (43%)	172/177 (97%)
Sidechain	1107/1591 (70%)	788/1039 (76%)	316/475 (67%)	3/77 (4%)

Continued on next page...

Continued from previous page...

	Total	¹H	¹³C	¹⁵N
Aromatic	103/154 (67%)	62/77 (81%)	39/70 (56%)	2/7 (29%)
Sugar	85/88 (97%)	46/48 (96%)	39/40 (98%)	0/0 (—%)
Base	10/55 (18%)	10/31 (32%)	0/15 (0%)	0/9 (0%)
Overall	2010/2832 (71%)	1273/1578 (81%)	560/984 (57%)	177/270 (66%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 70%, i.e. 2295 atoms were assigned a chemical shift out of a possible 3281. 0 out of 46 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹H	¹³C	¹⁵N
Backbone	810/1112 (73%)	423/453 (93%)	191/454 (42%)	196/205 (96%)
Sidechain	1260/1840 (68%)	899/1203 (75%)	358/549 (65%)	3/88 (3%)
Aromatic	130/186 (70%)	78/93 (84%)	49/85 (58%)	3/8 (38%)
Sugar	85/88 (97%)	46/48 (96%)	39/40 (98%)	0/0 (—%)
Base	10/55 (18%)	10/31 (32%)	0/15 (0%)	0/9 (0%)
Overall	2295/3281 (70%)	1456/1828 (80%)	637/1143 (56%)	202/310 (65%)

7.1.4 Statistically unusual chemical shifts ⓘ

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	71	GLU	H	12.45	5.45 – 11.20	7.2
1	A	196	LYS	HD3	0.20	0.54 – 2.65	-6.6
1	A	73	GLN	HG3	0.56	0.91 – 3.68	-6.3

7.1.5 Random Coil Index (RCI) plots ⓘ

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition. If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:

