Full wwPDB NMR Structure Validation Report ⓘ

Jun 3, 2023 – 06:41 AM EDT

PDB ID	:	5U9B
BMRB ID	:	30210
Title	:	Solution structure of the zinc fingers 1 and 2 of MBNL1 in complex with human cardiac troponin T pre-mRNA
Authors	:	Phukan, P.D.; Park, S.; Martinez-Yamout, M.M.; Zeeb, M.; Dyson, H.J.; Wright, P.E.
Deposited on	:	2016-12-15

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

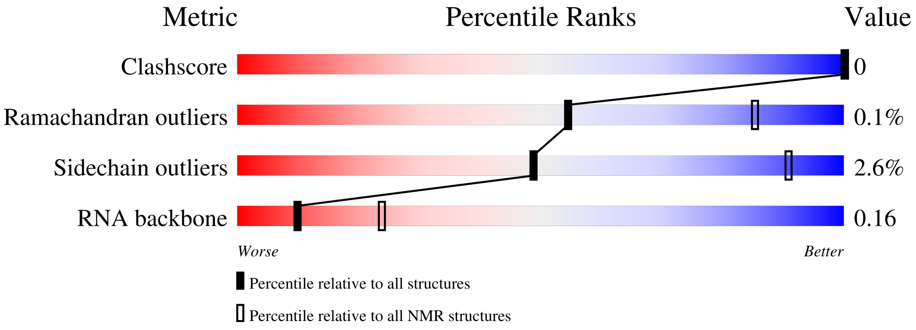
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

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 59%.

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	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428
RNA backbone	4643	676

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	92	
2	B	15	

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: *lowest energy*.

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:10-A:62, A:66-A:83 (71)	0.51	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 3 clusters and 1 single-model cluster was found.

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

3 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 1934 atoms, of which 891 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Muscleblind-like protein 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	92	1467	452	732	141	134	8	0

- Molecule 2 is a RNA chain called RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP *UP*CP*CP*CP*C)-3').

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	B	15	465	137	159	43	111	15	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
3	A	2	2	2

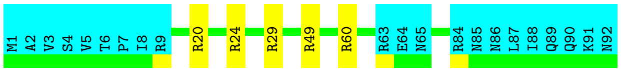
4 Residue-property plots ⓘ

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.

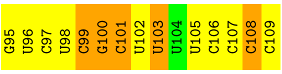
- Molecule 1: Muscleblind-like protein 1

Chain A:



- Molecule 2: RNA (5’-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3’)

Chain B:



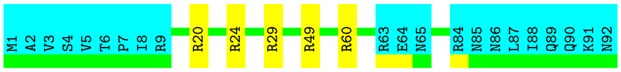
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

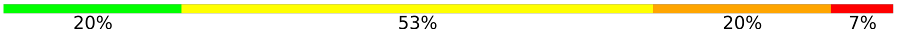
- Molecule 1: Muscleblind-like protein 1

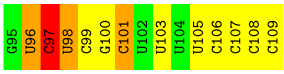
Chain A:



- Molecule 2: RNA (5’-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3’)

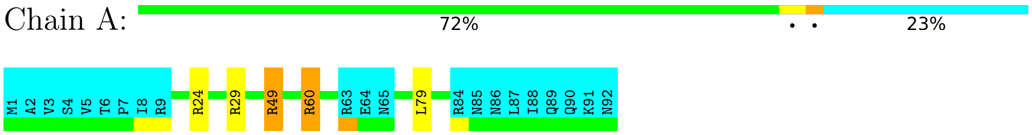
Chain B:



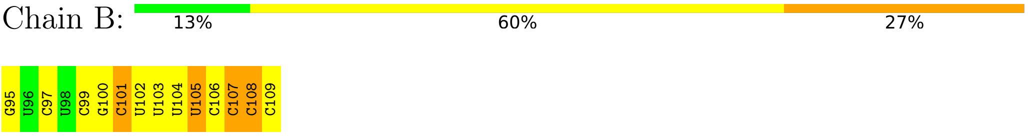


4.2.2 Score per residue for model 2

- Molecule 1: Muscleblind-like protein 1

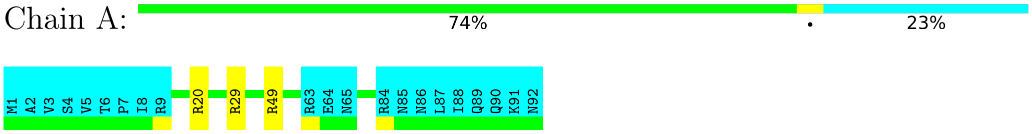


- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

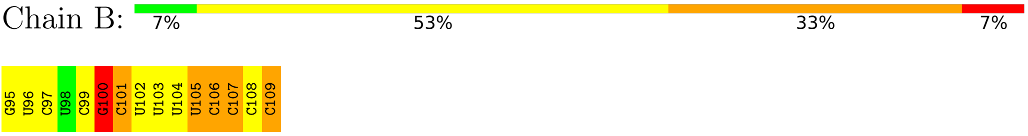


4.2.3 Score per residue for model 3

- Molecule 1: Muscleblind-like protein 1

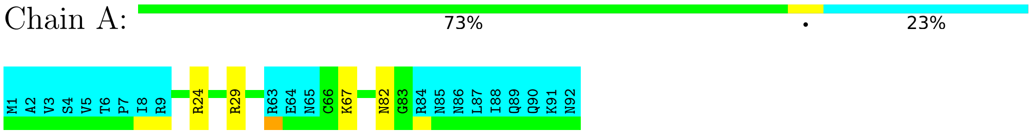


- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')



4.2.4 Score per residue for model 4

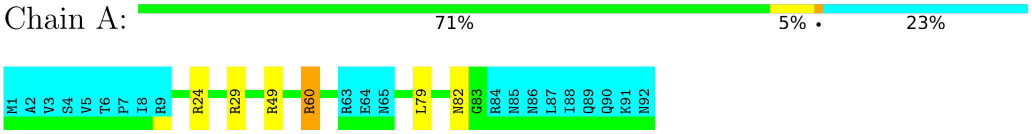
- Molecule 1: Muscleblind-like protein 1



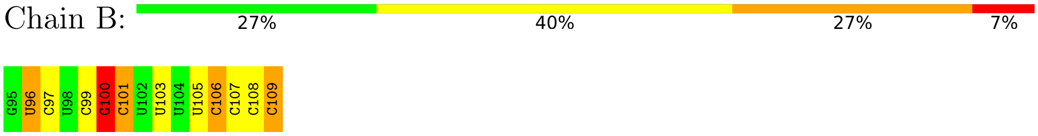
- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

4.2.5 Score per residue for model 5

- Molecule 1: Muscleblind-like protein 1

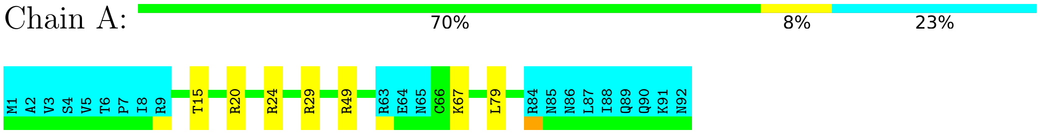


- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

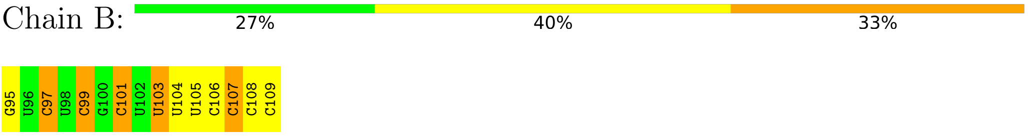


4.2.6 Score per residue for model 6

- Molecule 1: Muscleblind-like protein 1

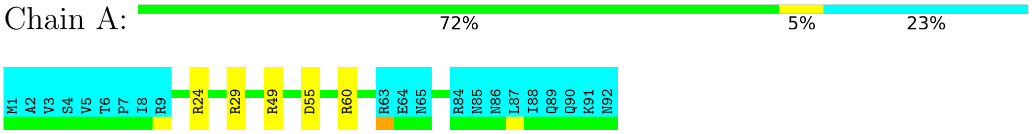


- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

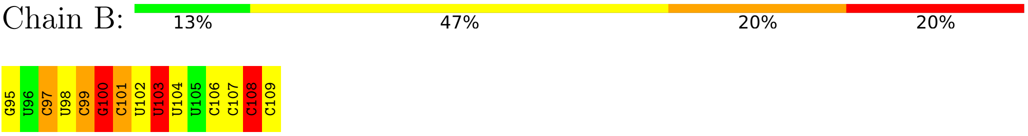


4.2.7 Score per residue for model 7

- Molecule 1: Muscleblind-like protein 1

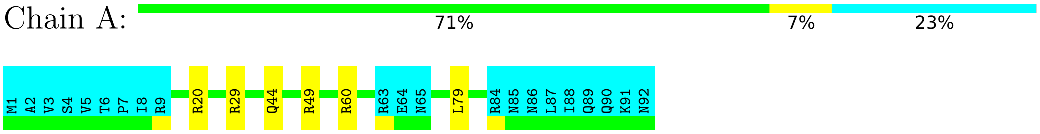


- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

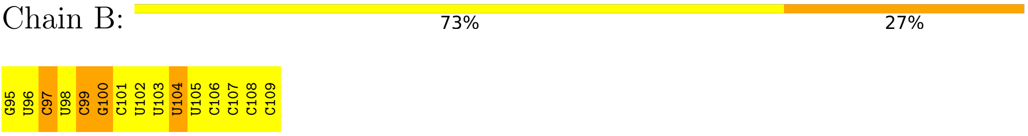


4.2.8 Score per residue for model 8

- Molecule 1: Muscleblind-like protein 1

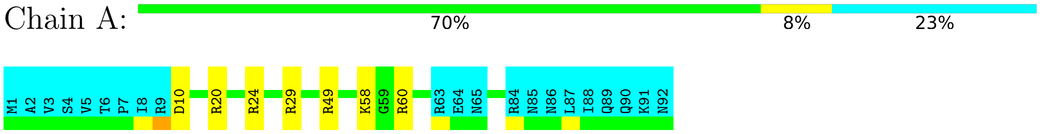


- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

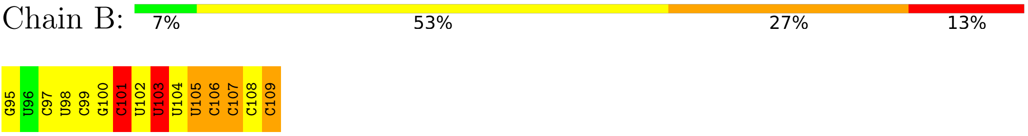


4.2.9 Score per residue for model 9

- Molecule 1: Muscleblind-like protein 1



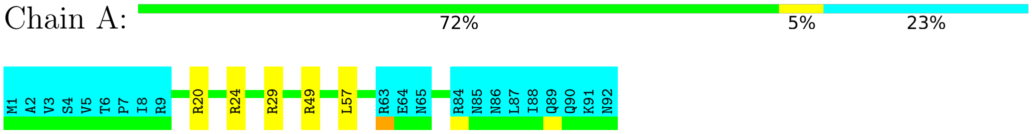
- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')



4.2.10 Score per residue for model 10

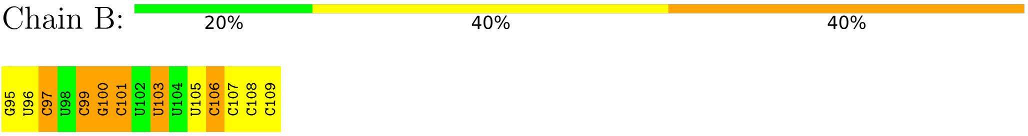
- Molecule 1: Muscleblind-like protein 1

Chain A:



- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

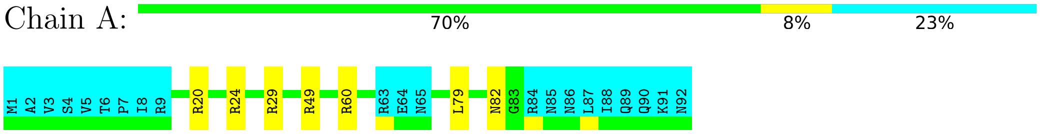
Chain B:



4.2.11 Score per residue for model 11

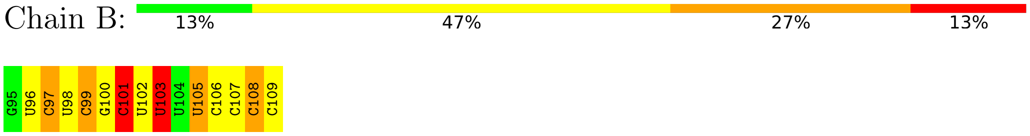
- Molecule 1: Muscleblind-like protein 1

Chain A:



- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

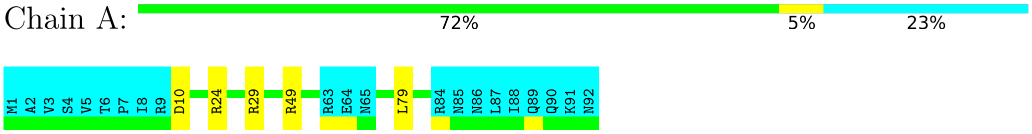
Chain B:



4.2.12 Score per residue for model 12

- Molecule 1: Muscleblind-like protein 1

Chain A:



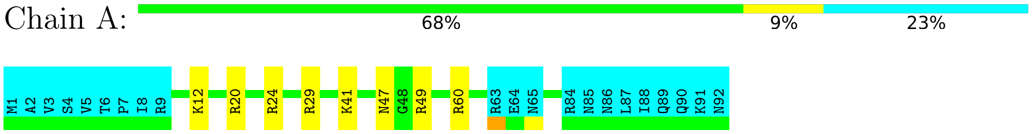
- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

Chain B:

4.2.13 Score per residue for model 13 (medoid)

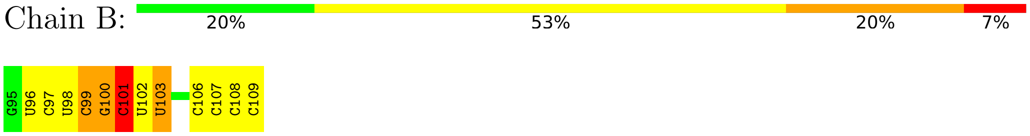
- Molecule 1: Muscleblind-like protein 1

Chain A:



- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

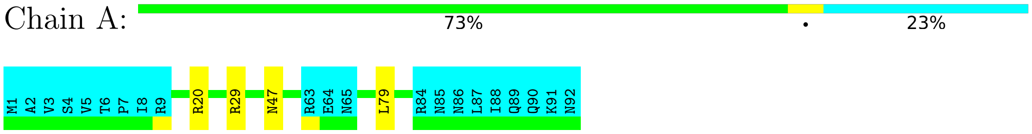
Chain B:



4.2.14 Score per residue for model 14

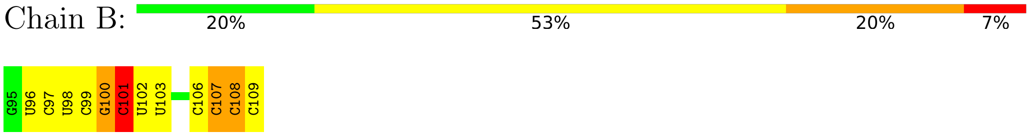
- Molecule 1: Muscleblind-like protein 1

Chain A:



- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

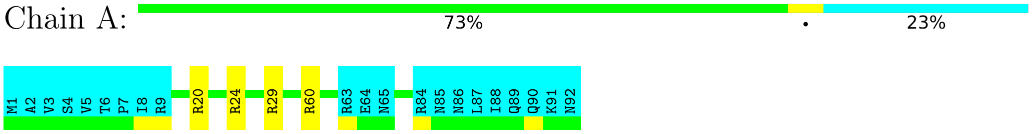
Chain B:



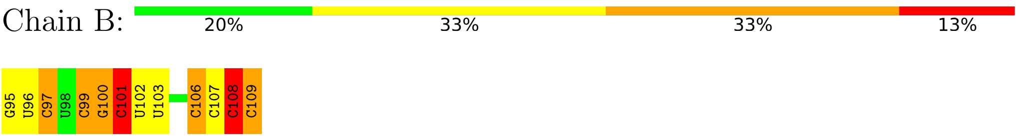
4.2.15 Score per residue for model 15

- Molecule 1: Muscleblind-like protein 1

Chain A:

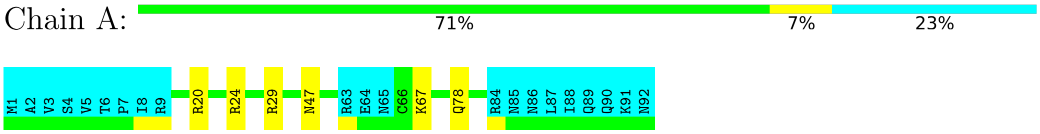


- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

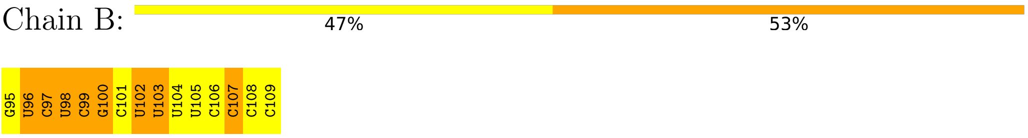


4.2.16 Score per residue for model 16

- Molecule 1: Muscleblind-like protein 1

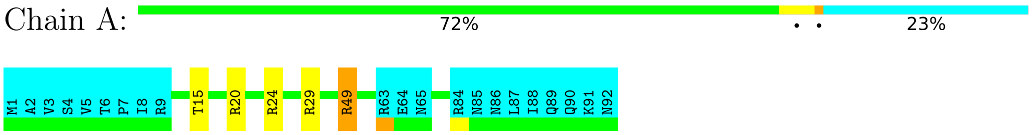


- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

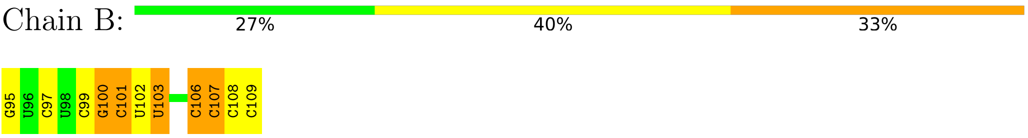


4.2.17 Score per residue for model 17

- Molecule 1: Muscleblind-like protein 1



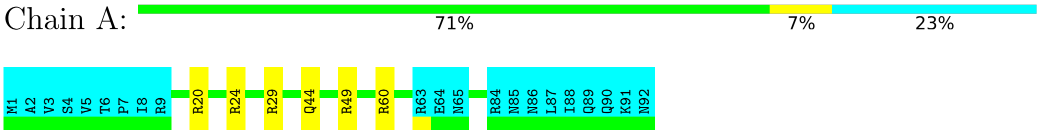
- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')



4.2.18 Score per residue for model 18

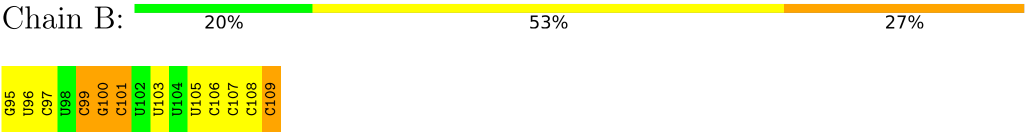
- Molecule 1: Muscleblind-like protein 1

Chain A:



- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

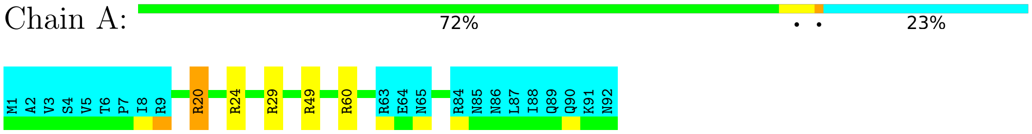
Chain B:



4.2.19 Score per residue for model 19

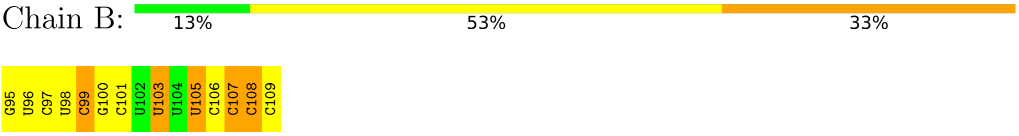
- Molecule 1: Muscleblind-like protein 1

Chain A:



- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

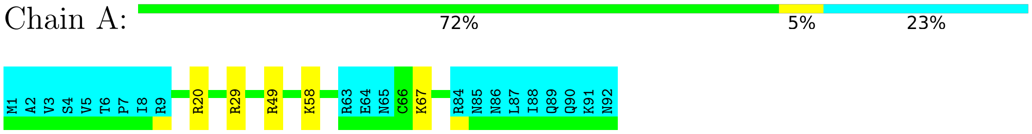
Chain B:



4.2.20 Score per residue for model 20

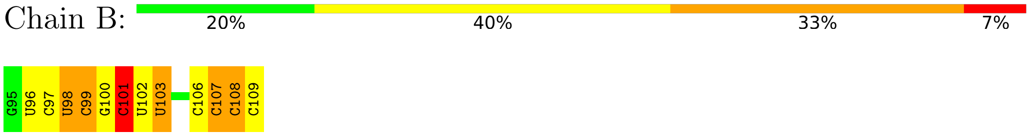
- Molecule 1: Muscleblind-like protein 1

Chain A:



- Molecule 2: RNA (5'-R(P*GP*UP*CP*UP*CP*GP*CP*UP*UP*UP*UP*CP*CP*CP*C)-3')

Chain B:



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing*.

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	**Classification**	**Version**
NMRDraw	structure calculation	
CYANA	structure calculation	
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	939
Number of shifts mapped to atoms	939
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	59%

6 Model quality ⓘ

6.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	0.68±0.00	0±0/577 (0.0± 0.0%)	1.02±0.02	4±1/779 (0.5± 0.1%)
2	B	1.40±0.02	0±0/337 (0.0± 0.0%)	2.24±0.08	19±3/520 (3.7± 0.5%)
All	All	1.00	0/18280 (0.0%)	1.63	459/25980 (1.8%)

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	2.1±0.8
All	All	0	42

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	105	U	O4'-C1'-N1	11.99	117.80	108.20	9	4
2	B	96	U	O4'-C1'-N1	10.91	116.93	108.20	16	5
2	B	98	U	O4'-C1'-N1	10.88	116.91	108.20	12	6
2	B	97	C	O4'-C1'-N1	9.84	116.07	108.20	1	6
2	B	100	G	O4'-C1'-N9	9.60	115.88	108.20	3	9
2	B	103	U	O4'-C1'-N1	9.51	115.80	108.20	15	12
2	B	108	C	O4'-C1'-N1	9.45	115.76	108.20	7	6
2	B	99	C	O4'-C1'-N1	9.44	115.75	108.20	12	10
2	B	107	C	O4'-C1'-N1	9.20	115.56	108.20	1	5
2	B	95	G	O4'-C1'-N9	8.71	115.17	108.20	10	8
2	B	107	C	N3-C2-O2	-8.60	115.88	121.90	4	20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
2	B	100	G	P-O3'-C3'	8.00	129.30	119.70	4	2
2	B	97	C	N3-C2-O2	-7.96	116.33	121.90	1	20
2	B	107	C	N1-C2-O2	7.93	123.66	118.90	4	15
2	B	97	C	N1-C2-O2	7.92	123.65	118.90	1	11
1	A	49	ARG	NE-CZ-NH2	7.83	124.22	120.30	7	11
2	B	99	C	N3-C2-O2	-7.80	116.44	121.90	1	20
2	B	108	C	N3-C2-O2	-7.65	116.55	121.90	3	20
2	B	106	C	N3-C2-O2	-7.61	116.57	121.90	9	20
2	B	99	C	N1-C2-O2	7.60	123.46	118.90	1	12
1	A	20	ARG	NE-CZ-NH2	7.58	124.09	120.30	15	14
2	B	103	U	N3-C2-O2	-7.56	116.91	122.20	13	13
2	B	105	U	N3-C2-O2	-7.47	116.97	122.20	2	6
2	B	109	C	N3-C2-O2	-7.47	116.67	121.90	1	20
2	B	105	U	C1'-O4'-C4'	-7.33	104.04	109.90	9	1
1	A	60	ARG	NE-CZ-NH1	7.20	123.90	120.30	2	6
2	B	108	C	N1-C2-O2	7.13	123.18	118.90	3	11
2	B	96	U	N3-C2-O2	-7.11	117.22	122.20	16	6
2	B	106	C	N1-C2-O2	7.03	123.12	118.90	9	11
2	B	109	C	O4'-C1'-N1	6.86	113.69	108.20	19	6
1	A	49	ARG	NE-CZ-NH1	6.82	123.71	120.30	2	5
1	A	60	ARG	NE-CZ-NH2	6.80	123.70	120.30	19	4
2	B	104	U	O4'-C1'-N1	6.74	113.59	108.20	9	5
2	B	109	C	N1-C2-O2	6.69	122.92	118.90	1	15
2	B	98	U	N3-C2-O2	-6.65	117.54	122.20	12	7
2	B	100	G	C1'-O4'-C4'	-6.51	104.69	109.90	3	1
1	A	29	ARG	NE-CZ-NH2	6.45	123.53	120.30	12	11
2	B	106	C	C1'-O4'-C4'	-6.33	104.84	109.90	5	1
2	B	108	C	C1'-O4'-C4'	-6.32	104.85	109.90	7	1
2	B	106	C	O4'-C1'-N1	6.30	113.24	108.20	5	7
1	A	29	ARG	NE-CZ-NH1	6.28	123.44	120.30	1	9
2	B	95	G	N3-C4-C5	-6.27	125.47	128.60	7	4
1	A	24	ARG	NE-CZ-NH1	6.26	123.43	120.30	7	5
2	B	102	U	O4'-C1'-N1	6.21	113.17	108.20	13	8
2	B	99	C	C1'-O4'-C4'	-6.20	104.94	109.90	19	1
1	A	24	ARG	NE-CZ-NH2	6.07	123.33	120.30	13	11
2	B	101	C	N3-C2-O2	-6.04	117.67	121.90	20	17
2	B	102	U	C3'-C2'-C1'	5.97	106.28	101.50	9	6
1	A	20	ARG	NE-CZ-NH1	5.78	123.19	120.30	1	1
2	B	98	U	C3'-C2'-C1'	5.72	106.08	101.50	1	1
2	B	100	G	N1-C6-O6	-5.72	116.47	119.90	13	9
2	B	96	U	C3'-C2'-C1'	5.64	106.01	101.50	16	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
2	B	107	C	C1'-O4'-C4'	-5.55	105.46	109.90	14	1
2	B	104	U	C5'-C4'-O4'	5.49	115.69	109.10	12	1
2	B	100	G	N3-C4-C5	-5.45	125.87	128.60	1	3
2	B	95	G	N1-C6-O6	-5.42	116.64	119.90	6	1
2	B	101	C	N1-C2-O2	5.41	122.14	118.90	10	4
2	B	101	C	N3-C4-C5	5.27	124.01	121.90	18	3

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	101	C	Sidechain	20
2	B	100	G	Sidechain	5
2	B	105	U	Sidechain	4
2	B	102	U	Sidechain	3
2	B	98	U	Sidechain	2
2	B	104	U	Sidechain	2
2	B	108	C	Sidechain	2
2	B	97	C	Sidechain	1
2	B	107	C	Sidechain	1
2	B	103	U	Sidechain	1
2	B	99	C	Sidechain	1

6.2 Too-close contacts (i)

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	563	550	549	0±0
2	B	306	159	160	0±0
All	All	17420	14180	14180	-

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 -.

There are no clashes.

6.3 Torsion angles ⓘ

6.3.1 Protein backbone ⓘ

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	71/92 (77%)	67±1 (94±1%)	4±1 (6±1%)	0±0 (0±0%)	54	85
All	All	1420/1840 (77%)	1338 (94%)	80 (6%)	2 (0%)	54	85

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	10	ASP	2

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	65/85 (76%)	63±1 (97±2%)	2±1 (3±2%)	49	91
All	All	1300/1700 (76%)	1266 (97%)	34 (3%)	49	91

All 15 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	79	LEU	7
1	A	67	LYS	4
1	A	60	ARG	3
1	A	82	ASN	3
1	A	47	ASN	3
1	A	49	ARG	2
1	A	15	THR	2
1	A	44	GLN	2
1	A	58	LYS	2
1	A	55	ASP	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	57	LEU	1
1	A	12	LYS	1
1	A	41	LYS	1
1	A	78	GLN	1
1	A	20	ARG	1

6.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers	Suiteness
2	B	14/15 (93%)	6±1 (41±9%)	2±1 (12±8%)	0.16±0.07
All	All	285/300 (95%)	116 (41%)	34 (12%)	0.16

The overall RNA backbone suiteness is 0.16.

All unique RNA backbone outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
2	B	103	U	18
2	B	100	G	15
2	B	96	U	12
2	B	105	U	11
2	B	99	C	11
2	B	108	C	9
2	B	101	C	8
2	B	97	C	7
2	B	107	C	7
2	B	106	C	7
2	B	109	C	5
2	B	98	U	4
2	B	104	U	2

All unique RNA pucker outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
2	B	100	G	8
2	B	99	C	6
2	B	95	G	5
2	B	107	C	4
2	B	97	C	3
2	B	96	U	2
2	B	106	C	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
2	B	105	U	2
2	B	104	U	2

6.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

6.7 Other polymers ⓘ

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 (i)

The completeness of assignment taking into account all chemical shift lists is 59% for the well-defined parts and 61% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *chemical_shifts_zf12RNA_str.txt*

7.1.1 Bookkeeping (i)

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	939
Number of shifts mapped to atoms	939
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	**# values**	**Correction ± precision, *ppm***	**Suggested action**
$^{13}C_{\alpha}$	90	0.09 ± 0.15	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	84	0.22 ± 0.14	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	78	-0.94 ± 0.55	None needed (imprecise)

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 59%, i.e. 728 atoms were assigned a chemical shift out of a possible 1233. 0 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	**^{1}H**	**^{13}C**	**^{15}N**
Backbone	267/349 (77%)	136/141 (96%)	70/142 (49%)	61/66 (92%)
Sidechain	417/536 (78%)	270/344 (78%)	143/166 (86%)	4/26 (15%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	44/73 (60%)	22/37 (59%)	21/31 (68%)	1/5 (20%)
Sugar	0/165 (0%)	0/90 (0%)	0/75 (0%)	0/0 (—%)
Base	0/110 (0%)	0/65 (0%)	0/28 (0%)	0/17 (0%)
Overall	728/1233 (59%)	428/677 (63%)	234/442 (53%)	66/114 (58%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 61%, i.e. 938 atoms were assigned a chemical shift out of a possible 1549. 0 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	341/452 (75%)	173/182 (95%)	90/184 (49%)	78/86 (91%)
Sidechain	553/749 (74%)	359/480 (75%)	188/227 (83%)	6/42 (14%)
Aromatic	44/73 (60%)	22/37 (59%)	21/31 (68%)	1/5 (20%)
Sugar	0/165 (0%)	0/90 (0%)	0/75 (0%)	0/0 (—%)
Base	0/110 (0%)	0/65 (0%)	0/28 (0%)	0/17 (0%)
Overall	938/1549 (61%)	554/854 (65%)	299/545 (55%)	85/150 (57%)

7.1.4 Statistically unusual chemical shifts ⓘ

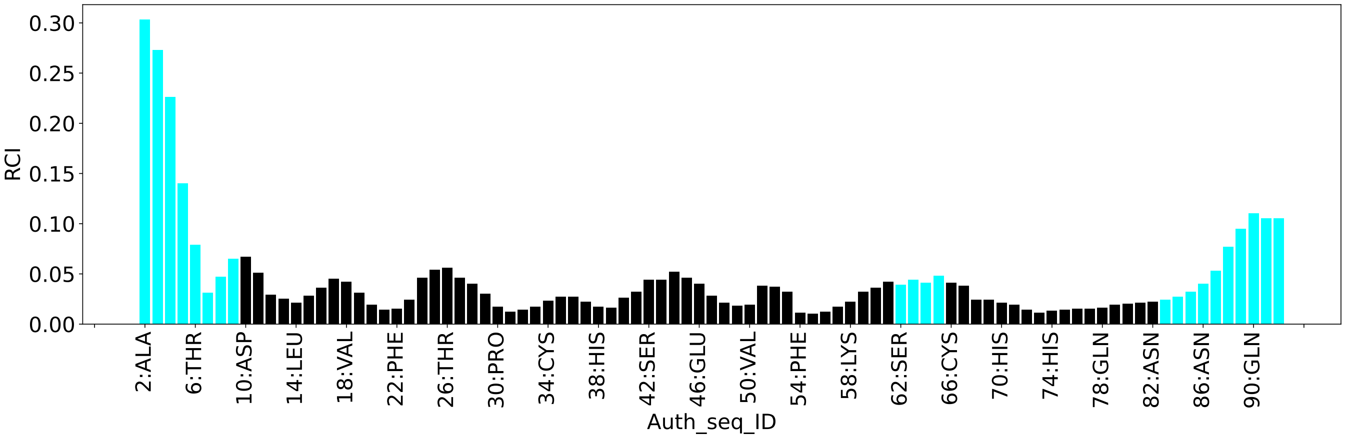
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	39	PRO	CB	62.60	26.06 – 37.61	26.6
1	A	30	PRO	CB	62.56	26.06 – 37.61	26.6
1	A	73	PRO	CB	62.56	26.06 – 37.61	26.6
1	A	7	PRO	CB	62.26	26.06 – 37.61	26.3
1	A	14	LEU	HG	-0.94	-0.13 – 3.16	-7.4

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:


RCI
0.30
0.25
0.20
0.15
0.10
0.05
0.00
2:ALA
6:THR
10:ASP
14:LEU
18:VAL
22:PHE
26:THR
30:PRO
34:CYS
38:HIS
42:SER
46:GLU
50:VAL
54:PHE
58:LYS
62:SER
66:CYS
70:HIS
74:HIS
78:GLN
82:ASN
86:ASN
90:GLN
Auth_seq_ID

8 NMR restraints analysis ⓘ

8.1 Conformationally restricting restraints ⓘ

The following table provides the summary of experimentally observed NMR restraints in different categories. Restraints are classified into different categories based on the sequence separation of the atoms involved.

Description	Value
Total distance restraints	16
Intra-residue (|i-j|=0)	3
Sequential (|i-j|=1)	13
Medium range (|i-j|>1 and |i-j|<5)	0
Long range (|i-j|≥5)	0
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	0
Number of restraints per residue	0.1
Number of long range restraints per residue[1]	0.0

[1] Long range hydrogen bonds and disulfide bonds are counted as long range restraints while calculating the number of long range restraints per residue

8.2 Residual restraint violations ⓘ

This section provides the overview of the restraint violations analysis. The violations are binned as small, medium and large violations based on its absolute value. Average number of violations per model is calculated by dividing the total number of violations in each bin by the size of the ensemble.

8.2.1 Average number of distance violations per model ⓘ

Distance violations less than 0.1 Å are not included in the calculation.There are no distance violations

8.2.2 Average number of dihedral-angle violations per model ⓘ

Dihedral-angle violations less than 1° are not included in the calculation.There are no dihedral-angle violations

9 Distance violation analysis

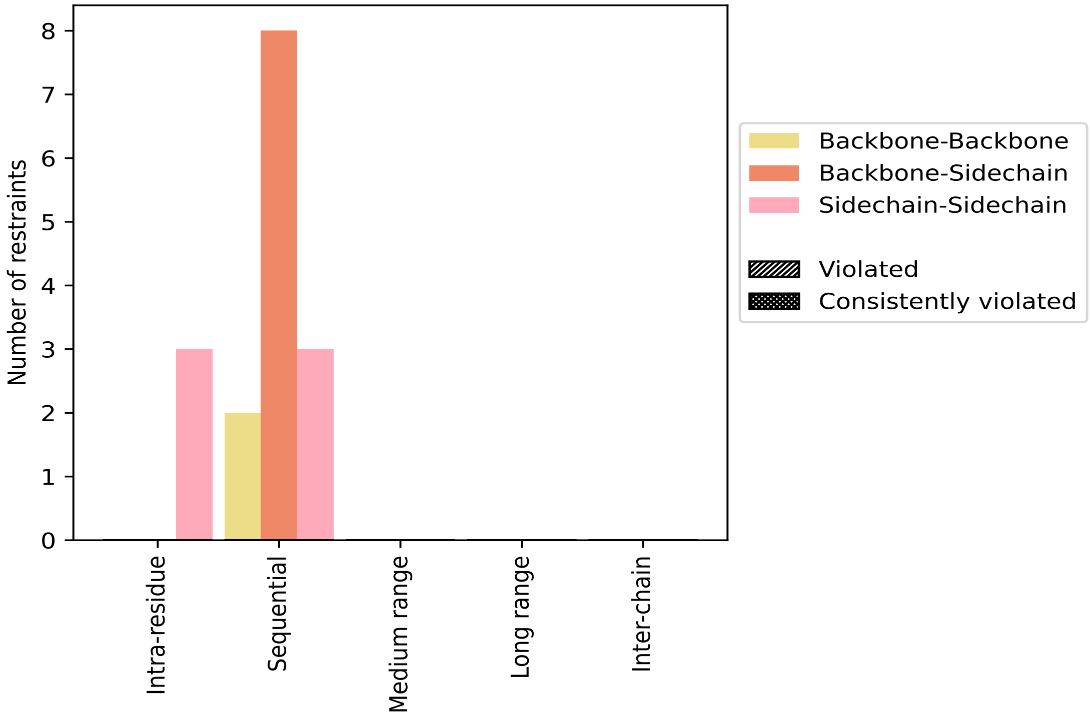
9.1 Summary of distance violations

The following table shows the summary of distance violations in different restraint categories based on the sequence separation of the atoms involved. Each category is further sub-divided into three sub-categories based on the atoms involved. Violations less than 0.1 Å are not included in the statistics.

Restraints type	**Count**	%[1]	**Violated**[3]			**Consistently Violated**[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	3	18.8	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	3	18.8	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	13	81.2	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	2	12.5	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	8	50.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	3	18.8	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Long range (|i-j|≥5)	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Inter-chain	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Hydrogen bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	16	100.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	2	12.5	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	8	50.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	6	37.5	0	0.0	0.0	0	0.0	0.0

[1] percentage calculated with respect to the total number of distance restraints, [2] percentage calculated with respect to the number of restraints in a particular restraint category, [3] violated in at least one model, [4] violated in all the models

9.1.1 Bar chart : Distribution of distance restraints and violations ⓘ



Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and disulfied bonds are counted in their appropriate category on the x-axis

9.2 Distance violation statistics for each model ⓘ

No violations found

9.3 Distance violation statistics for the ensemble ⓘ

No violations found

9.4 Most violated distance restraints in the ensemble ⓘ

No violations found

9.5 All violated distance restraints ⓘ

No violations found

10 Dihedral-angle violation analysis ⓘ

Dihedral angle analysis failed due to data error in the dihedral angle restraints, possibly missing target value