



Full wwPDB NMR Structure Validation Report ⓘ

Mar 30, 2026 – 12:56 AM UTC

PDB ID : 2YH1 / pdb_00002yh1
BMRB ID : 17623
Title : Model of human U2AF65 tandem RRM1 and RRM2 domains with eight-site uridine binding
Authors : Mackereth, C.D.; Madl, T.; Simon, B.; Zanier, K.; Gasch, A.; Sattler, M.
Deposited on : 2011-04-26

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

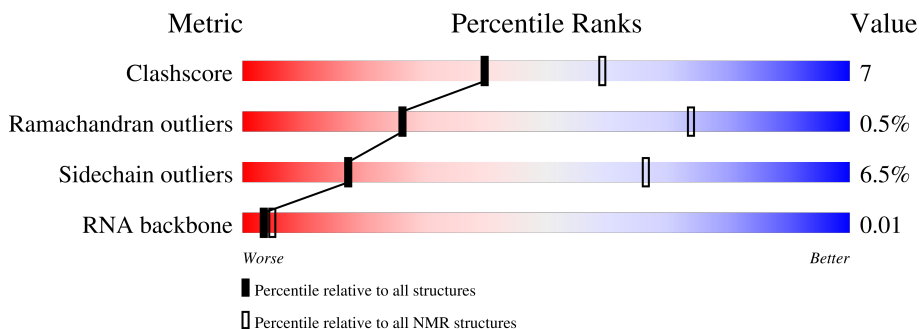
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 58%.

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	198	
2	B	9	

2 Ensemble composition and analysis

This entry contains 10 models. Model 5 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative.

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:149-A:230 (82)	0.08	5
2	A:259-A:337 (79)	0.06	2

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 2 clusters. No single-model clusters were found.

Cluster number	Models
1	2, 3, 4, 5, 7, 10
2	1, 6, 8, 9

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3255 atoms, of which 1584 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called SPLICING FACTOR U2AF 65 KDA SUBUNIT.

Mol	Chain	Residues	Atoms						Trace
1	A	195	Total	C	H	N	O	S	0
			2986	952	1492	254	281	7	

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	145	GLY	-	expression tag	UNP P26368
A	146	ALA	-	expression tag	UNP P26368
A	147	MET	-	expression tag	UNP P26368

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

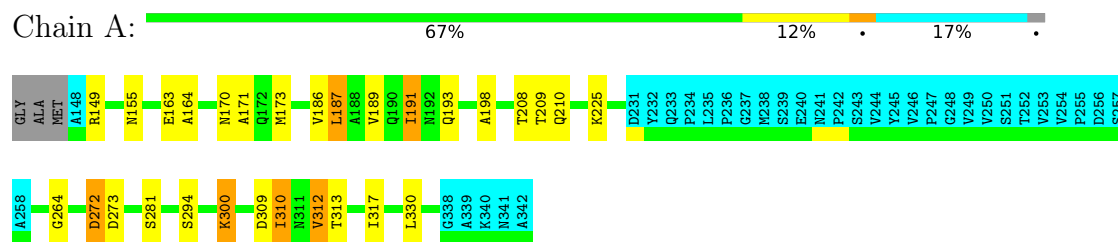
Mol	Chain	Residues	Atoms						Trace
2	B	9	Total	C	H	N	O	P	0
			269	81	92	18	70	8	

4 Residue-property plots [i](#)

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: SPLICING FACTOR U2AF 65 KDA SUBUNIT



- Molecule 2: 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP)-3'

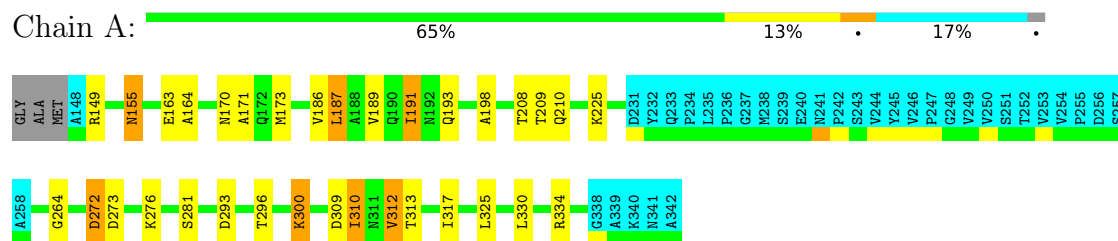


4.2 Scores per residue for each member of the ensemble

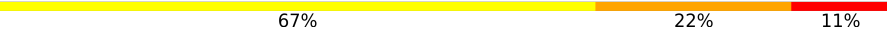
Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

- Molecule 1: SPLICING FACTOR U2AF 65 KDA SUBUNIT



- Molecule 2: 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP)-3'

Chain B: 

U601
U602
U603
U604
U605
U606
U607
U608
U609

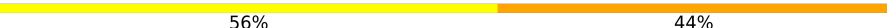
4.2.2 Score per residue for model 2

- Molecule 1: SPLICING FACTOR U2AF 65 KDA SUBUNIT

Chain A: 

GLY
ALA
MET
A148
R149
R150
N155
E163
A164
N170
A171
Q172
M173
V186
L187
A188
V189
Q190
I191
N192
Q193
A198
T208
T209
Q210
K225
D231
V232
Q233
P234
L235
P236
G237
M238
S239
E240
N241
P242
S243
V244
Y245
V246
P247
G248
V249
V250
S251
T252
V253
V254
P255
D256

- Molecule 2: 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP)-3'

Chain B: 

U601
U602
U603
U604
U605
U606
U607
U608
U609

4.2.3 Score per residue for model 3

- Molecule 1: SPLICING FACTOR U2AF 65 KDA SUBUNIT

Chain A: 

GLY
ALA
MET
A148
R149
R150
L151
N155
E163
A164
N170
A171
M173
V186
L187
A188
V189
Q190
I191
N192
Q193
A198
T208
T209
Q210
K225
D231
V232
Q233
P234
L235
P236
G237
M238
S239
E240
N241
P242
S243
V244
Y245
V246
P247
G248
V249
V250
S251
T252
V253
V254
P255

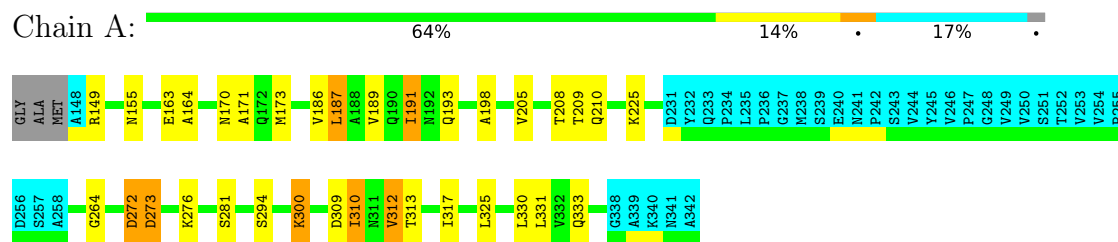
- Molecule 2: 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP)-3'

Chain B: 

U601
U602
U603
U604
U605
U606
U607
U608
U609

4.2.4 Score per residue for model 4

- Molecule 1: SPLICING FACTOR U2AF 65 KDA SUBUNIT

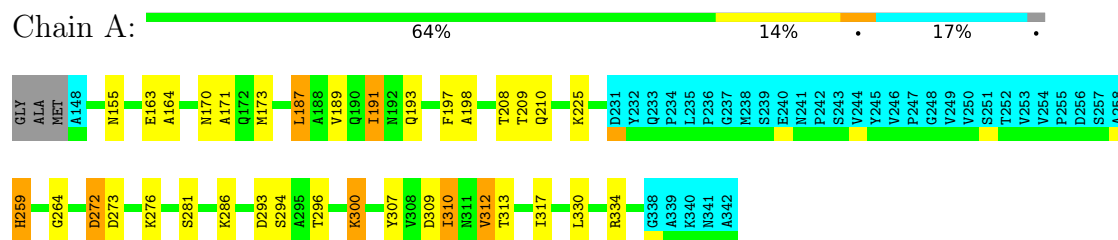


- Molecule 2: 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP)-3'



4.2.5 Score per residue for model 5 (medoid)

- Molecule 1: SPLICING FACTOR U2AF 65 KDA SUBUNIT

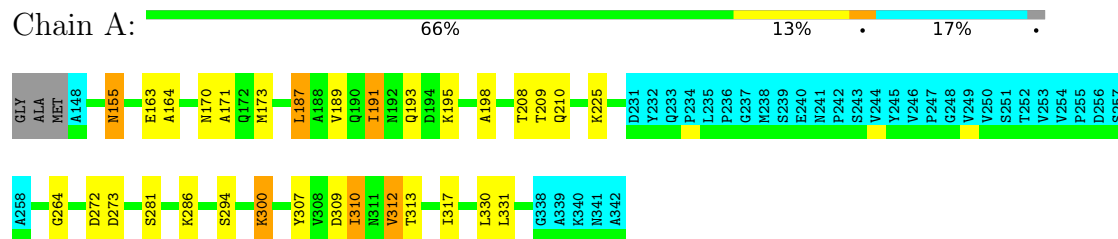


- Molecule 2: 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP)-3'

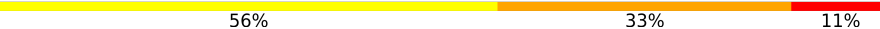


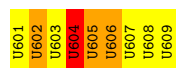
4.2.6 Score per residue for model 6

- Molecule 1: SPLICING FACTOR U2AF 65 KDA SUBUNIT



- Molecule 2: 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP)-3'

Chain B: 



4.2.7 Score per residue for model 7

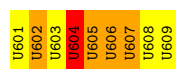
- Molecule 1: SPLICING FACTOR U2AF 65 KDA SUBUNIT

Chain A: 



- Molecule 2: 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP)-3'

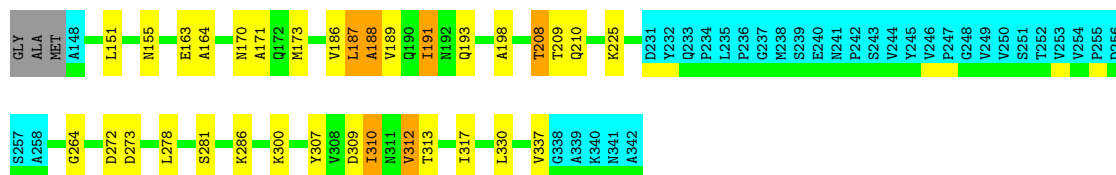
Chain B: 



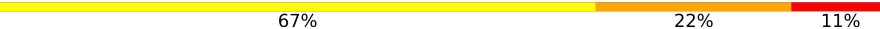
4.2.8 Score per residue for model 8

- Molecule 1: SPLICING FACTOR U2AF 65 KDA SUBUNIT

Chain A: 



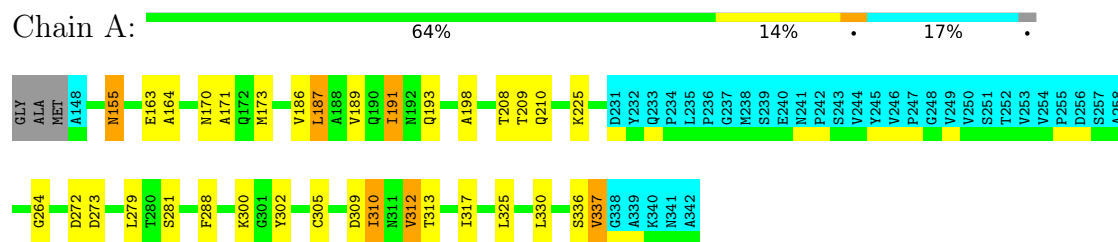
- Molecule 2: 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP)-3'

Chain B: 



4.2.9 Score per residue for model 9

- Molecule 1: SPLICING FACTOR U2AF 65 KDA SUBUNIT

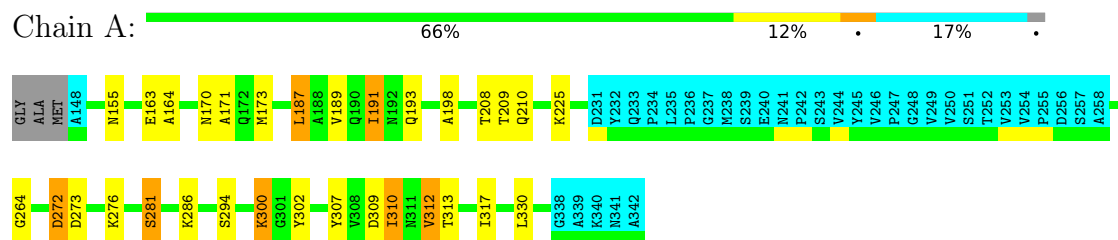


- Molecule 2: 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP)-3'



4.2.10 Score per residue for model 10

- Molecule 1: SPLICING FACTOR U2AF 65 KDA SUBUNIT



- Molecule 2: 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP)-3'



5 Refinement protocol and experimental data overview

The models were refined using the following method: *MODIFIED ARIA*.

Of the 125 calculated structures, 10 were deposited, based on the following criterion: *LOWEST ENERGY*.

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

Software name	Classification	Version
CNS	refinement	1.2
Sparky	structure solution	

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	1473
Number of shifts mapped to atoms	1473
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	58%

6 Model quality

6.1 Standard geometry

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	0.95±0.01	14±1/1279 (1.1± 0.1%)	0.84±0.01	1±0/1726 (0.1± 0.0%)
2	B	0.65±0.06	1±1/194 (0.5± 0.4%)	0.83±0.04	0±0/298 (0.0± 0.0%)
All	All	0.92	151/14730 (1.0%)	0.84	10/20240 (0.0%)

All unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	155	ASN	C-N	8.95	1.43	1.33	8	10
1	A	208	THR	C-N	8.46	1.45	1.34	4	10
1	A	281	SER	C-N	7.48	1.45	1.33	5	10
1	A	273	ASP	C-N	6.97	1.43	1.33	10	10
1	A	164	ALA	C-N	6.85	1.43	1.34	3	10
1	A	171	ALA	C-N	6.65	1.43	1.33	9	10
1	A	317	ILE	C-N	6.64	1.42	1.33	9	10
1	A	209	THR	CA-CB	6.63	1.64	1.53	4	10
1	A	209	THR	C-N	6.40	1.43	1.33	4	10
2	B	604	U	O3'-P	6.31	1.70	1.61	6	7
1	A	187	LEU	N-CA	6.13	1.54	1.46	1	8
2	B	604	U	C3'-O3'	5.90	1.52	1.43	4	3
1	A	163	GLU	C-N	5.74	1.41	1.34	2	10
1	A	186	VAL	C-N	5.62	1.39	1.33	2	7
1	A	209	THR	N-CA	5.43	1.53	1.46	4	1
1	A	187	LEU	CA-C	5.38	1.58	1.53	2	1
1	A	272	ASP	C-N	5.35	1.41	1.34	9	10
1	A	170	ASN	C-N	5.24	1.41	1.33	1	10
1	A	188	ALA	N-CA	5.19	1.52	1.46	2	2
1	A	187	LEU	C-N	5.07	1.40	1.33	2	2

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)	Models	
								Worst	Total
1	A	191	ILE	N-CA-C	6.30	116.98	108.17	3	10

There are no chirality outliers.

There are no planarity outliers.

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	1255	1260	1258	18 $\pm$ 5
2	B	177	92	92	12 $\pm$ 5
All	All	14320	13520	13500	197

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 7.

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

Atom-1	Atom-2	Clash($\text{\AA}$)	Distance($\text{\AA}$)	Models	
				Worst	Total
1:A:331:LEU:CD1	2:B:602:U:H3	1.34	1.34	4	3
1:A:331:LEU:HD13	2:B:602:U:N3	1.09	1.61	4	2
1:A:331:LEU:CD1	2:B:602:U:N3	1.08	2.16	4	3
1:A:331:LEU:HD13	2:B:602:U:H3	1.03	1.11	6	2
1:A:300:LYS:NZ	2:B:601:U:O2'	0.99	1.96	1	7
1:A:225:LYS:NZ	2:B:604:U:H2'	0.97	1.74	8	4
1:A:331:LEU:HD11	2:B:602:U:H3	0.93	1.23	6	2
1:A:300:LYS:NZ	2:B:601:U:C2	0.93	2.37	10	3
1:A:331:LEU:HD12	2:B:602:U:C2	0.91	1.99	3	1
1:A:331:LEU:HD13	2:B:602:U:C2	0.86	2.04	4	2
1:A:331:LEU:HD12	2:B:602:U:O2	0.83	1.73	3	1
1:A:331:LEU:HD12	2:B:602:U:N3	0.80	1.92	3	1
1:A:225:LYS:NZ	2:B:604:U:C2'	0.79	2.45	8	3
1:A:300:LYS:HE2	2:B:601:U:O2	0.78	1.77	1	6
1:A:225:LYS:NZ	2:B:604:U:H1'	0.75	1.95	10	6
1:A:300:LYS:NZ	2:B:601:U:O2	0.74	2.19	10	7
1:A:225:LYS:HZ2	2:B:604:U:H2'	0.73	1.39	1	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:225:LYS:HZ2	2:B:604:U:C2'	0.73	1.95	8	1
1:A:225:LYS:HZ3	2:B:604:U:C2'	0.71	1.98	9	1
1:A:300:LYS:CE	2:B:601:U:O2	0.71	2.38	1	6
1:A:225:LYS:NZ	2:B:604:U:O2'	0.69	2.22	4	6
1:A:225:LYS:HZ3	2:B:604:U:H2'	0.67	1.49	8	4
1:A:331:LEU:HD13	2:B:602:U:O2	0.66	1.91	6	2
1:A:309:ASP:O	1:A:312:VAL:HG22	0.65	1.92	2	10
1:A:300:LYS:CE	2:B:601:U:C2	0.65	2.80	10	2
1:A:300:LYS:HE2	2:B:601:U:C2	0.64	2.26	1	2
2:B:605:U:O2'	2:B:606:U:P	0.64	2.55	6	2
1:A:294:SER:HB3	2:B:605:U:C2	0.64	2.27	5	6
1:A:149:ARG:HB2	1:A:208:THR:OG1	0.61	1.93	3	3
2:B:605:U:HO2'	2:B:606:U:P	0.61	2.17	6	1
1:A:225:LYS:HZ2	2:B:604:U:H1'	0.59	1.55	10	4
1:A:331:LEU:CD1	2:B:602:U:C2	0.59	2.83	3	1
1:A:302:TYR:OH	2:B:603:U:H4'	0.59	1.98	9	2
1:A:331:LEU:HD11	2:B:602:U:N3	0.58	1.98	6	3
1:A:260:LYS:HD3	1:A:306:GLU:HB2	0.57	1.74	2	1
1:A:149:ARG:HB3	1:A:208:THR:HG21	0.57	1.75	3	3
1:A:331:LEU:HD12	2:B:602:U:H3	0.56	1.50	3	1
1:A:302:TYR:CE1	2:B:603:U:OP1	0.55	2.59	3	1
1:A:225:LYS:NZ	2:B:604:U:C1'	0.54	2.68	10	4
1:A:325:LEU:HD23	1:A:330:LEU:HB2	0.50	1.83	4	4
1:A:195:LYS:HD3	2:B:605:U:O2'	0.50	2.07	6	2
1:A:286:LYS:HE3	1:A:307:TYR:O	0.49	2.08	5	5
1:A:310:ILE:O	1:A:313:THR:HB	0.48	2.09	2	10
1:A:149:ARG:CB	1:A:208:THR:OG1	0.48	2.62	3	1
1:A:149:ARG:HA	1:A:149:ARG:HE	0.48	1.69	3	1
1:A:302:TYR:CZ	2:B:603:U:H4'	0.47	2.44	9	2
2:B:602:U:H4'	2:B:603:U:OP1	0.47	2.09	3	1
1:A:310:ILE:HD13	1:A:310:ILE:H	0.46	1.70	5	1
1:A:272:ASP:O	1:A:276:LYS:HG3	0.46	2.10	10	5
1:A:336:SER:O	1:A:337:VAL:HB	0.45	2.11	9	1
1:A:325:LEU:HD13	1:A:330:LEU:HB2	0.44	1.88	7	1
1:A:264:GLY:O	1:A:330:LEU:HD12	0.44	2.13	4	10
1:A:294:SER:HB3	2:B:605:U:N3	0.44	2.27	6	1
1:A:191:ILE:HG12	1:A:198:ALA:HB2	0.43	1.91	7	10
1:A:293:ASP:HB3	1:A:296:THR:HB	0.43	1.91	1	4
2:B:601:U:O2'	2:B:602:U:OP1	0.42	2.37	3	10
1:A:197:PHE:CD2	2:B:606:U:OP2	0.42	2.72	5	1
1:A:149:ARG:O	1:A:150:ARG:CG	0.41	2.68	2	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:288:PHE:HD1	1:A:305:CYS:HB3	0.41	1.75	9	2
1:A:151:LEU:HG	1:A:208:THR:HG23	0.41	1.92	8	2
1:A:279:LEU:HD13	1:A:305:CYS:SG	0.40	2.56	9	1
1:A:149:ARG:O	1:A:150:ARG:HG2	0.40	2.16	2	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	161/198 (81%)	150±1 (93±1%)	10±1 (6±0%)	1±1 (0±0%)	26	74
All	All	1610/1980 (81%)	1499 (93%)	103 (6%)	8 (0%)	26	74

All 5 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	188	ALA	2
1	A	281	SER	2
1	A	337	VAL	2
1	A	149	ARG	1
1	A	259	HIS	1

6.3.2 Protein sidechains [i](#)

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	133/161 (83%)	124±1 (93±1%)	9±1 (7±1%)	17	66
All	All	1330/1610 (83%)	1243 (93%)	87 (7%)	17	66

All 12 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	173	MET	10
1	A	187	LEU	10
1	A	189	VAL	10
1	A	193	GLN	10
1	A	210	GLN	10
1	A	310	ILE	10
1	A	300	LYS	9
1	A	312	VAL	9
1	A	149	ARG	3
1	A	155	ASN	3
1	A	278	LEU	2
1	A	273	ASP	1

6.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers	Suiteness
2	B	8/9 (89%)	8±0 (100±0%)	2±1 (20±8%)	0.00±0.00
All	All	80/90 (89%)	80 (100%)	16 (20%)	0.00

The overall RNA backbone suiteness is 0.01.

All unique RNA backbone outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
2	B	602	U	10
2	B	603	U	10
2	B	604	U	10
2	B	605	U	10
2	B	606	U	10
2	B	607	U	10
2	B	608	U	10
2	B	609	U	10

All unique RNA pucker outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
2	B	606	U	8
2	B	605	U	7
2	B	603	U	1

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 [i](#)

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 58% for the well-defined parts and 53% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chemical_shifts_1*

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	1473
Number of shifts mapped to atoms	1473
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	4

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	129	-0.05 ± 0.14	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	118	0.02 ± 0.09	None needed (< 0.5 ppm)
$^{13}\text{C}'$	0	—	None (insufficient data)
^{15}N	171	0.35 ± 0.43	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 58%, i.e. 1365 atoms were assigned a chemical shift out of a possible 2362. 0 out of 28 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	546/808 (68%)	278/331 (84%)	121/322 (38%)	147/155 (95%)
Sidechain	780/1229 (63%)	533/800 (67%)	236/378 (62%)	11/51 (22%)

Continued on next page...

Continued from previous page...

	Total	¹ H	¹³ C	¹⁵ N
Aromatic	39/172 (23%)	39/84 (46%)	0/84 (0%)	0/4 (0%)
Sugar	0/99 (0%)	0/54 (0%)	0/45 (0%)	0/0 (—%)
Base	0/54 (0%)	0/27 (0%)	0/18 (0%)	0/9 (0%)
Overall	1365/2362 (58%)	850/1296 (66%)	357/847 (42%)	158/219 (72%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 53%, i.e. 1472 atoms were assigned a chemical shift out of a possible 2763. 0 out of 35 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	610/971 (63%)	310/397 (78%)	129/390 (33%)	171/184 (93%)
Sidechain	819/1449 (57%)	560/946 (59%)	246/448 (55%)	13/55 (24%)
Aromatic	43/190 (23%)	43/92 (47%)	0/94 (0%)	0/4 (0%)
Sugar	0/99 (0%)	0/54 (0%)	0/45 (0%)	0/0 (—%)
Base	0/54 (0%)	0/27 (0%)	0/18 (0%)	0/9 (0%)
Overall	1472/2763 (53%)	913/1516 (60%)	375/995 (38%)	184/252 (73%)

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	276	LYS	HD3	0.15	0.54 – 2.65	-6.9
1	A	172	GLN	HG3	0.60	0.91 – 3.68	-6.1
1	A	276	LYS	HD2	0.43	0.58 – 2.64	-5.7
1	A	208	THR	HA	1.95	2.11 – 6.79	-5.3

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:

