Full wwPDB NMR Structure Validation Report (i)

Feb 26, 2022 – 06:13 PM EST

PDB ID	:	2ADB
Title	:	Solution structure of Polypyrimidine Tract Binding protein RBD2 complexed with CUCUCU RNA
Authors	:	Oberstrass, F.C.; Auweter, S.D.; Erat, M.; Hargous, Y.; Henning, A.; Wenter, P.; Reymond, L.; Pitsch, S.; Black, D.L.; Allain, F.H.T.
Deposited on	:	2005-07-20

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 (i) symbol.

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

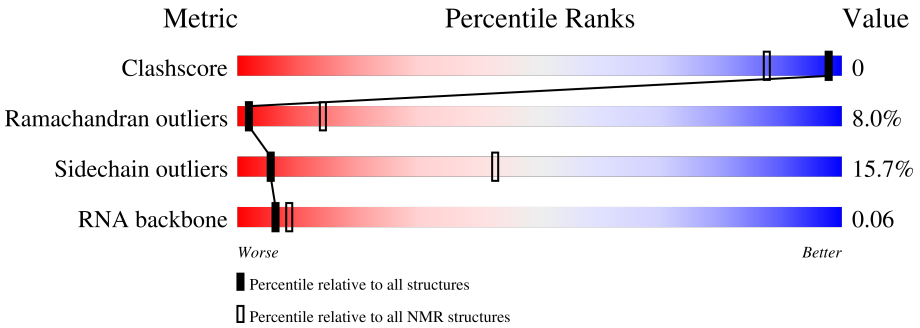
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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	B	6	17% 67% 17%
2	A	148	61% 15% • 9% 14%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 2 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:180-A:282 (103)	0.50	2
2	A:287-A:297 (11)	0.21	1

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 5 clusters and 2 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2154 atoms, of which 1048 are hydrogens and 0 are deuteriums.

- Molecule 1 is a RNA chain called 5'-R(*CP*UP*CP*UP*CP*U)-3'.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
1	B	6	182	54	65	15	43	5	0

- Molecule 2 is a protein called Polypyrimidine tract-binding protein 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	A	127	1972	627	983	167	190	5	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	151	MET	-	expression tag	UNP P26599
A	152	GLY	-	expression tag	UNP P26599
A	153	SER	-	expression tag	UNP P26599
A	154	SER	-	expression tag	UNP P26599
A	155	HIS	-	expression tag	UNP P26599
A	156	HIS	-	expression tag	UNP P26599
A	157	HIS	-	expression tag	UNP P26599
A	158	HIS	-	expression tag	UNP P26599
A	159	HIS	-	expression tag	UNP P26599
A	160	HIS	-	expression tag	UNP P26599
A	161	SER	-	expression tag	UNP P26599
A	162	SER	-	expression tag	UNP P26599
A	163	GLY	-	expression tag	UNP P26599
A	164	LEU	-	expression tag	UNP P26599
A	165	VAL	-	expression tag	UNP P26599
A	166	PRO	-	expression tag	UNP P26599
A	167	ARG	-	expression tag	UNP P26599
A	168	GLY	-	expression tag	UNP P26599
A	169	SER	-	expression tag	UNP P26599
A	170	HIS	-	expression tag	UNP P26599
A	171	MET	-	expression tag	UNP P26599

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

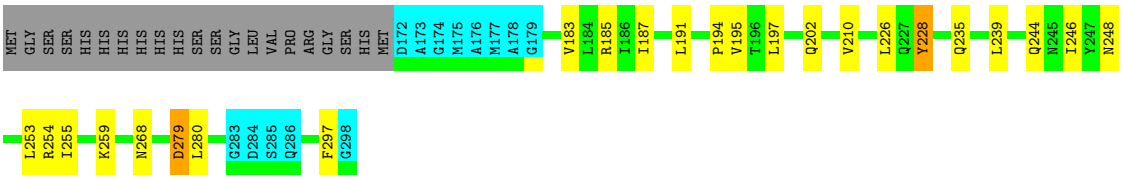
These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for 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: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 17% 67% 17%

• Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 61% 15% • 9% 14%



4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

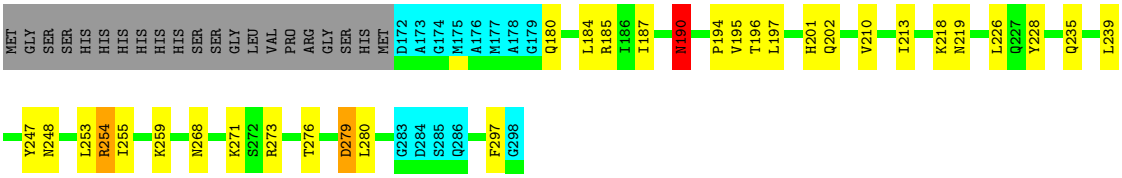
4.2.1 Score per residue for model 1

• Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 50% 17% 33%

• Molecule 2: Polypyrimidine tract-binding protein 1

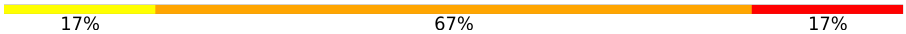
Chain A: 55% 20% •• 9% 14%



4.2.2 Score per residue for model 2 (medoid)

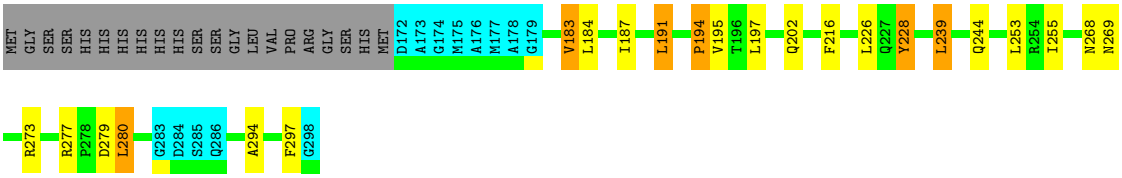
- Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B:



- Molecule 2: Polypyrimidine tract-binding protein 1

Chain A:



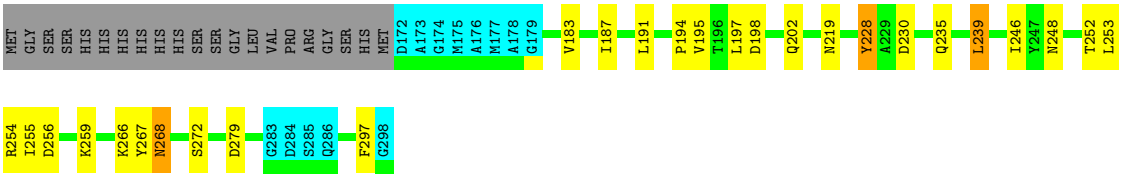
4.2.3 Score per residue for model 3

- Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B:

- Molecule 2: Polypyrimidine tract-binding protein 1

Chain A:



4.2.4 Score per residue for model 4

- Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 17% 50% 33%

C299 U300 C301 U302 G303 U304

• Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 55% 20% •• 9% 14%

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET D172 A173 G174 M175 A176 M177 A178 G179 V183 L184 R185 I186 I187 V188 E189 N190 Y193 P194 V195 Q202 N219 Y228 A229 D230 L239 Q244 N245 I246 Y247 N248 L253

R254 I255 D256 K259 V265 K266 Y267 N268 K271 S272 R273 D274 Y275 D279 L280 G283 D284 S285 Q286 P287 G298

4.2.5 Score per residue for model 5

• Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 17% 33% 17% 33%

C299 U300 C301 U302 G303 U304

• Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 55% 17% 5% 9% 14%

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET D172 A173 G174 M175 A176 M177 A178 G179 V183 L184 R185 I186 I187 V188 E189 N190 P194 V195 T196 L197 Q202 V210 N219 Q223 L226 Q227 Y228 Q235 K238 L239

Q244 N245 N248 A249 C250 C251 R254 I255 L263 N264 V265 N268 D279 L280 P281 S282 G283 D284 S285 Q286 A294 F297 G298

4.2.6 Score per residue for model 6

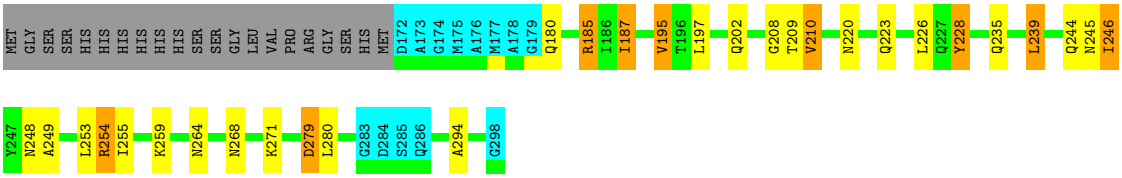
• Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 33% 33% 33%

C299 U300 C301 U302 C303 U304

• Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 57% 14% 6% 9% 14%



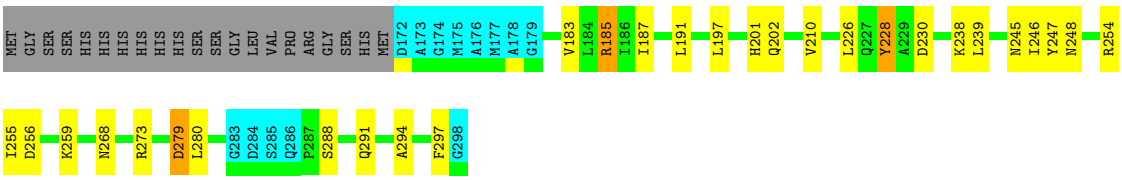
4.2.7 Score per residue for model 7

- Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 33% 50% 17%

- Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 57% 18% • 9% 14%



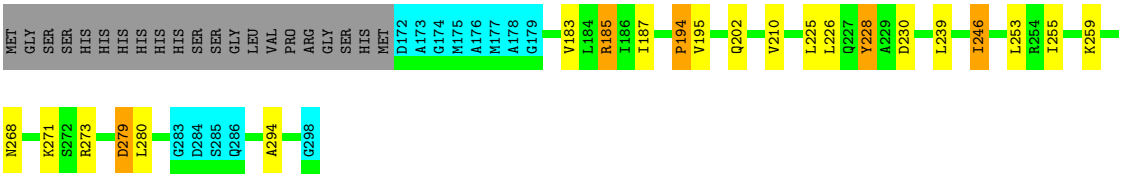
4.2.8 Score per residue for model 8

- Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 17% 33% 50%

- Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 62% 11% • 9% 14%



4.2.9 Score per residue for model 9

- Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 50% 17% 33%

C299 U300 G301 U302 G303 U304

• Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 59% 17% • 9% 14%

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET D172 A173 G174 M175 A176 M177 A178 G179 V183 L184 I187 V195 T196 L197 D198 Q202 V210 I213 L226 K238 L239 Q244 N245 I246 L253 R254 I255 K259

Y267 N268 R273 D279 L280 G283 D284 S285 Q286 A294 F297 G298

4.2.10 Score per residue for model 10

• Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 50% 33% 17%

C299 U300 G301 U302 C303 U304

• Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 55% 18% • 9% 14%

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET D172 A173 G174 M175 A176 M177 A178 G179 Q180 V183 L184 R185 I186 I187 N190 V195 T196 L197 H201 Q202 I213 F216 N219 Y228 Q235 L239 Q244 N245

I246 Y247 N248 L253 R254 I255 D256 K259 N268 N269 D270 D279 L280 P281 S282 G283 D284 S285 Q286 Q291 T292 W293 F297 G298

4.2.11 Score per residue for model 11

• Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 33% 50% 17%

C299 U300 G301 U302 C303 U304

• Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 57% 17% • 9% 14%

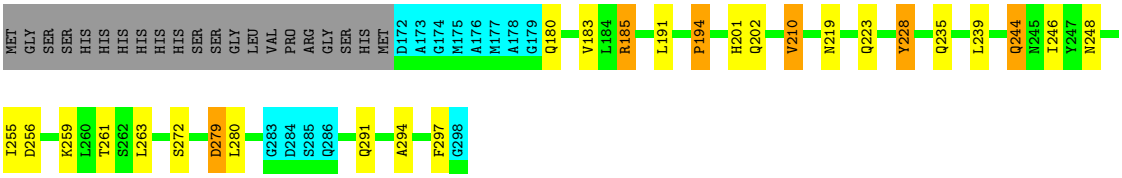
4.2.12 Score per residue for model 12

- Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B:

- Molecule 2: Polypyrimidine tract-binding protein 1

Chain A:



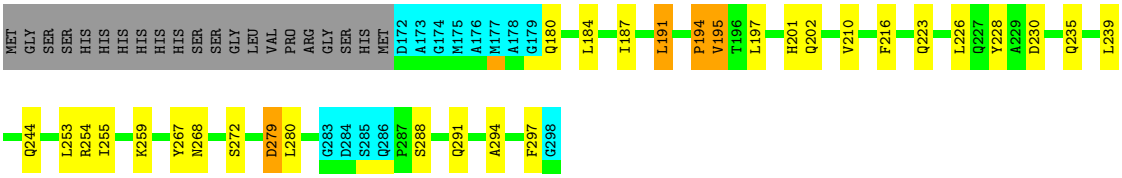
4.2.13 Score per residue for model 13

- Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B:

- Molecule 2: Polypyrimidine tract-binding protein 1

Chain A:



4.2.14 Score per residue for model 14

- Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 33% 33% 33%

C299 U300 C301 U302 C303 U304

• Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 56% 17% • • 9% 14%

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET D172 A173 G174 M175 A176 M177 A178 G179 Q180 R185 I186 I187 L191 F192 Y193 P194 V195 T196 L197 D198 V199 L200 H201 Q202 V210 I213 L225 Y228 Q235 K238 L239

L253 R254 I255 D256 K259 L263 S272 R273 D279 L280 G283 D284 S285 Q286 Q291 F297 G298

4.2.15 Score per residue for model 15

• Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 33% 50% 17%

C299 U300 C301 U302 C303 U304

• Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 59% 14% 5% 9% 14%

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET D172 A173 G174 M175 A176 M177 A178 G179 Q180 L184 I187 P194 L197 G208 T209 V210 Q223 L226 Q227 Y228 Q235 K238 L239 Q244 N245 I246 L253 R254

I255 D256 N268 K271 D274 D279 L280 G283 D284 S285 Q286 Q291 F297 G298

4.2.16 Score per residue for model 16

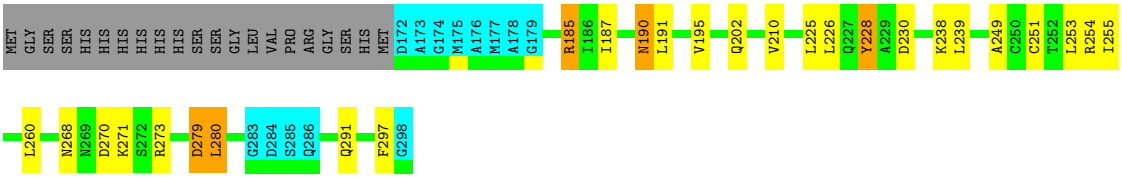
• Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 17% 17% 67%

C299 U300 C301 U302 C303 U304

• Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 59% 15% • 9% 14%



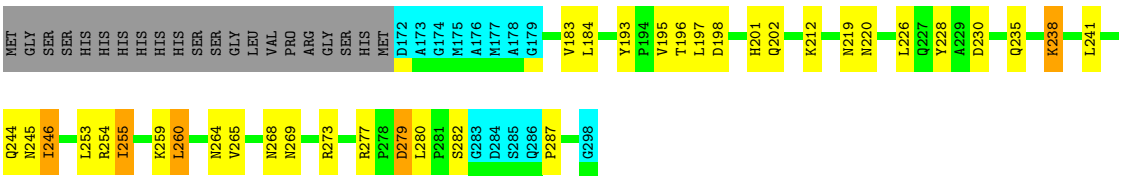
4.2.17 Score per residue for model 17

- Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 17% 33% 50%

- Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 53% 21% • 9% 14%



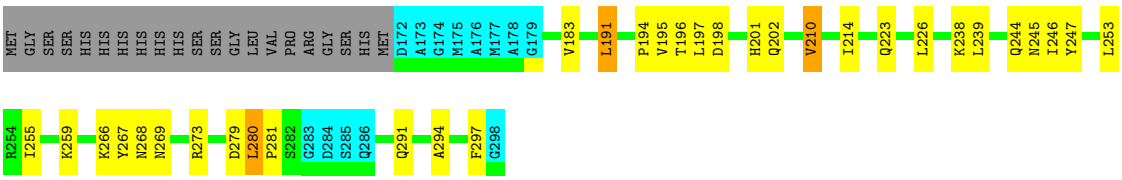
4.2.18 Score per residue for model 18

- Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 50% 17% 33%

- Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 55% 20% • 9% 14%



4.2.19 Score per residue for model 19

- Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 17% 50% 33%

C299 U300 C301 U302 G303 U304

• Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 58% 14% • • 9% 14%

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET D172 A173 G174 M175 A176 M177 A178 G179 V183 L184 R185 L191 P194 V195 Q202 V210 Q223 L226 Q227 Y228 A229 D230 L239 Q244 N245 I246 T252 L253 R254

I255 K259 Y267 N268 N269 D270 K271 S272 D279 L280 G283 D284 S285 Q286 Q291 F297 G298

4.2.20 Score per residue for model 20

• Molecule 1: 5'-R(*CP*UP*CP*UP*CP*U)-3'

Chain B: 17% 67% 17%

C299 U300 C301 U302 C303 U304

• Molecule 2: Polypyrimidine tract-binding protein 1

Chain A: 55% 20% • 9% 14%

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET D172 A173 G174 M175 A176 M177 A178 G179 Q180 V183 L184 R185 I186 I187 P194 V195 T196 L197 D198 H201 Q202 I213 L226 Q227 Y228 A229 D230 L239 Q244 N245 I246 Y247

N248 A249 C250 L253 R254 I255 D256 K259 L263 Y267 D279 L280 P281 S282 G283 D284 S285 Q286 L289 F297 G298

5 Refinement protocol and experimental data overview (i)

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

Of the 40 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
DYANA	structure solution	3.02
Amber	refinement	7

No chemical shift data was provided.

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	B	1.34±0.04	0±0/128 (0.0± 0.0%)	2.20±0.11	7±1/196 (3.5± 0.6%)
2	A	0.71±0.00	0±0/930 (0.0± 0.0%)	1.22±0.03	2±2/1265 (0.2± 0.1%)
All	All	0.81	0/21160 (0.0%)	1.40	188/29220 (0.6%)

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
1	B	0.0±0.0	2.1±0.9
2	A	0.0±0.0	5.2±2.0
All	All	0	145

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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	B	304	U	O4'-C1'-N1	10.53	116.62	108.20	19	8
2	A	254	ARG	NE-CZ-NH1	8.14	124.37	120.30	10	13
2	A	185	ARG	NE-CZ-NH1	8.04	124.32	120.30	14	9
1	B	302	U	O4'-C1'-N1	7.58	114.26	108.20	11	16
1	B	304	U	N3-C2-O2	-7.50	116.95	122.20	12	10
2	A	273	ARG	NE-CZ-NH1	7.44	124.02	120.30	14	9
1	B	299	C	N3-C2-O2	-7.30	116.79	121.90	9	20
1	B	301	C	N3-C2-O2	-7.29	116.80	121.90	16	20
1	B	301	C	C5'-C4'-O4'	7.19	117.73	109.10	20	1
1	B	300	U	O4'-C1'-N1	6.64	113.51	108.20	20	3
1	B	303	C	N3-C2-O2	-6.53	117.33	121.90	3	19
1	B	301	C	O4'-C1'-N1	5.98	112.99	108.20	12	2
2	A	277	ARG	NE-CZ-NH1	5.97	123.28	120.30	2	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	B	301	C	N1-C2-O2	5.93	122.46	118.90	5	5
1	B	302	U	C5'-C4'-O4'	5.85	116.12	109.10	9	3
2	A	193	TYR	CB-CG-CD2	-5.85	117.49	121.00	4	1
2	A	252	THR	C-N-CA	5.84	136.31	121.70	19	1
1	B	303	C	C5'-C4'-O4'	5.81	116.07	109.10	9	1
2	A	216	PHE	CB-CG-CD1	-5.66	116.83	120.80	2	2
2	A	194	PRO	C-N-CA	5.64	135.81	121.70	19	7
1	B	301	C	C5-C6-N1	-5.63	118.18	121.00	16	1
1	B	299	C	N1-C2-O2	5.58	122.25	118.90	1	14
2	A	225	LEU	CB-CA-C	5.41	120.48	110.20	8	1
1	B	301	C	N3-C4-N4	-5.38	114.23	118.00	5	1
2	A	280	LEU	CB-CA-C	5.34	120.35	110.20	4	1
1	B	300	U	C5-C6-N1	-5.21	120.09	122.70	17	2
1	B	300	U	P-O3'-C3'	5.20	125.94	119.70	2	6
1	B	302	U	C5'-C4'-C3'	-5.20	107.69	116.00	3	1
1	B	303	C	N1-C2-O2	5.14	121.98	118.90	9	2
2	A	267	TYR	CB-CG-CD2	-5.11	117.93	121.00	19	1
2	A	272	SER	C-N-CA	5.10	134.45	121.70	14	1
1	B	302	U	N1-C2-N3	5.09	117.95	114.90	8	1
2	A	275	TYR	CB-CG-CD2	-5.06	117.96	121.00	4	1
1	B	300	U	N3-C2-O2	-5.04	118.67	122.20	12	1
1	B	300	U	N1-C1'-C2'	-5.03	106.47	112.00	3	1
2	A	288	SER	C-N-CA	5.01	134.23	121.70	7	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	A	228	TYR	Sidechain	18
1	B	300	U	Sidechain	15
1	B	301	C	Sidechain	13
2	A	279	ASP	Peptide	13
2	A	253	LEU	Peptide	12
2	A	195	VAL	Peptide	10
1	B	303	C	Sidechain	9
2	A	238	LYS	Peptide	6
2	A	184	LEU	Peptide	5
2	A	272	SER	Peptide	4
2	A	223	GLN	Peptide	4
2	A	180	GLN	Peptide	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
2	A	191	LEU	Peptide	3
2	A	194	PRO	Peptide	3
2	A	190	ASN	Peptide	3
2	A	185	ARG	Sidechain	3
2	A	270	ASP	Peptide	3
2	A	266	LYS	Peptide	2
1	B	304	U	Sidechain	2
2	A	271	LYS	Peptide	2
1	B	302	U	Sidechain	2
2	A	276	THR	Peptide	1
1	B	299	C	Sidechain	1
2	A	183	VAL	Peptide	1
2	A	188	VAL	Peptide	1
2	A	265	VAL	Peptide	1
2	A	209	THR	Peptide	1
2	A	267	TYR	Peptide	1
2	A	241	LEU	Peptide	1
2	A	255	ILE	Peptide	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	B	117	65	65	0±0
2	A	910	917	917	1±1
All	All	20540	19640	19640	18

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:A:280:LEU:HD22	2:A:281:PRO:HD2	0.59	1.74	18	4
2:A:238:LYS:HB3	2:A:255:ILE:HD13	0.53	1.80	17	1
2:A:210:VAL:HG21	2:A:213:ILE:HD11	0.49	1.83	14	1
2:A:266:LYS:HE3	2:A:267:TYR:CE2	0.48	2.44	3	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:A:273:ARG:HA	2:A:280:LEU:HD12	0.47	1.87	16	1
2:A:190:ASN:ND2	2:A:190:ASN:H	0.45	2.07	1	1
2:A:185:ARG:HD2	2:A:187:ILE:HD11	0.44	1.90	6	1
1:B:301:C:H1'	2:A:225:LEU:HD11	0.44	1.90	14	1
2:A:183:VAL:C	2:A:184:LEU:HD22	0.43	2.34	17	2
2:A:260:LEU:HD13	2:A:260:LEU:H	0.42	1.75	17	1
1:B:302:U:H5'	2:A:216:PHE:CZ	0.41	2.50	13	1
2:A:263:LEU:CD1	2:A:263:LEU:H	0.41	2.29	20	1
1:B:303:C:H5"	1:B:303:C:C6	0.40	2.50	8	1
2:A:267:TYR:HA	2:A:280:LEU:HD13	0.40	1.92	19	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	
2	A	114/148 (77%)	83±3 (73±3%)	22±3 (19±3%)	9±2 (8±2%)	2	14
All	All	2280/2960 (77%)	1657 (73%)	441 (19%)	182 (8%)	2	14

All 30 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	A	279	ASP	20
2	A	239	LEU	19
2	A	280	LEU	19
2	A	268	ASN	15
2	A	297	PHE	15
2	A	245	ASN	10
2	A	194	PRO	9
2	A	195	VAL	9
2	A	294	ALA	9
2	A	219	ASN	7
2	A	246	ILE	6
2	A	196	THR	5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
2	A	267	TYR	5
2	A	210	VAL	5
2	A	247	TYR	4
2	A	248	ASN	3
2	A	265	VAL	2
2	A	287	PRO	2
2	A	250	CYS	2
2	A	251	CYS	2
2	A	208	GLY	2
2	A	220	ASN	2
2	A	249	ALA	2
2	A	244	GLN	2
2	A	190	ASN	1
2	A	272	SER	1
2	A	192	PHE	1
2	A	269	ASN	1
2	A	252	THR	1
2	A	289	LEU	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
2	A	104/128 (81%)	88±2 (84±2%)	16±2 (16±2%)	5 42
All	All	2080/2560 (81%)	1753 (84%)	327 (16%)	5 42

All 51 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)
2	A	202	GLN	19
2	A	255	ILE	19
2	A	187	ILE	16
2	A	259	LYS	16
2	A	197	LEU	15
2	A	226	LEU	15
2	A	228	TYR	15

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
2	A	244	GLN	14
2	A	210	VAL	13
2	A	183	VAL	13
2	A	246	ILE	13
2	A	235	GLN	11
2	A	201	HIS	10
2	A	191	LEU	10
2	A	230	ASP	10
2	A	291	GLN	9
2	A	248	ASN	8
2	A	256	ASP	8
2	A	254	ARG	7
2	A	239	LEU	7
2	A	198	ASP	6
2	A	185	ARG	6
2	A	180	GLN	5
2	A	213	ILE	5
2	A	271	LYS	5
2	A	253	LEU	5
2	A	190	ASN	4
2	A	268	ASN	4
2	A	263	LEU	4
2	A	223	GLN	4
2	A	193	TYR	3
2	A	238	LYS	3
2	A	269	ASN	2
2	A	280	LEU	2
2	A	264	ASN	2
2	A	212	LYS	2
2	A	261	THR	2
2	A	260	LEU	2
2	A	218	LYS	1
2	A	247	TYR	1
2	A	252	THR	1
2	A	189	GLU	1
2	A	297	PHE	1
2	A	293	MET	1
2	A	288	SER	1
2	A	200	LEU	1
2	A	274	ASP	1
2	A	225	LEU	1
2	A	282	SER	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
2	A	214	ILE	1
2	A	267	TYR	1

6.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers	Suiteness
1	B	6/6 (100%)	4±1 (68±13%)	2±1 (27±15%)	0.07±0.10
All	All	104/120 (87%)	82 (79%)	32 (31%)	0.06

The overall RNA backbone suiteness is 0.06.

All unique RNA backbone outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
1	B	301	C	20
1	B	300	U	18
1	B	304	U	16
1	B	303	C	16
1	B	302	U	12

All unique RNA pucker outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
1	B	300	U	15
1	B	301	C	7
1	B	303	C	6
1	B	299	C	4

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 monosaccharides 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 (i)

No chemical shift data were provided