Full wwPDB NMR Structure Validation Report ⓘ

Feb 23, 2022 – 03:36 PM EST

PDB ID	:	1Y2S
Title	:	Ovine Prion Protein Variant R168
Authors	:	Christen, B.; Lysek, D.A.; Herrmann, T.; Wuthrich, K.
Deposited on	:	2004-11-23

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

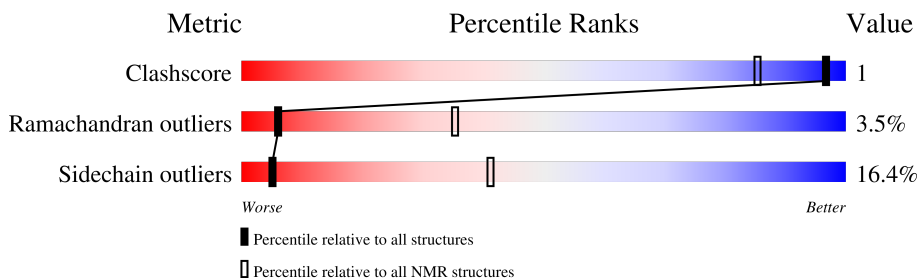|		
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

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 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	Percentile Ranks	Value
Clashscore		1
Ramachandran outliers		3.5%
Sidechain outliers		16.4%

Worse ... *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all NMR structures

Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

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	113	62% 13% 25%

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 12 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:128-A:164, A:173-A:220 (85)	0.51	12

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

3 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 1802 atoms, of which 873 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Major prion protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	113	1802	577	873	165	180	7	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	119	GLY	-	cloning artifact	UNP P23907
A	120	SER	-	cloning artifact	UNP P23907

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: Major prion protein
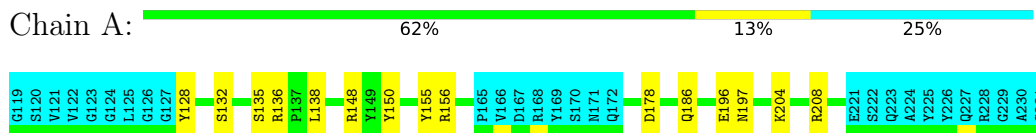
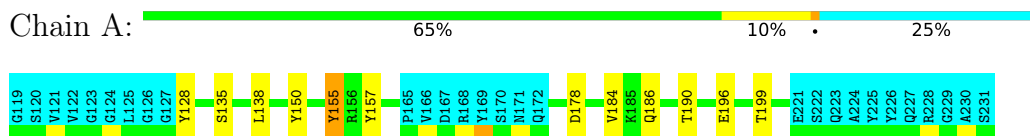## 4.2 Scores per residue for each member of the ensemble

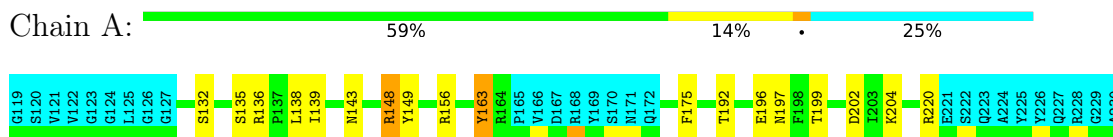Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Major prion protein

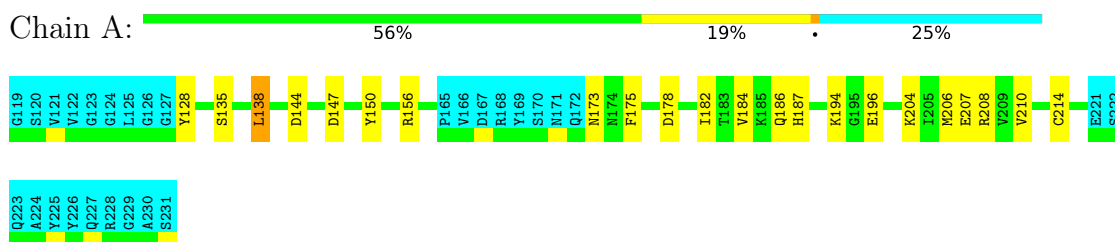### 4.2.2 Score per residue for model 2

• Molecule 1: Major prion protein

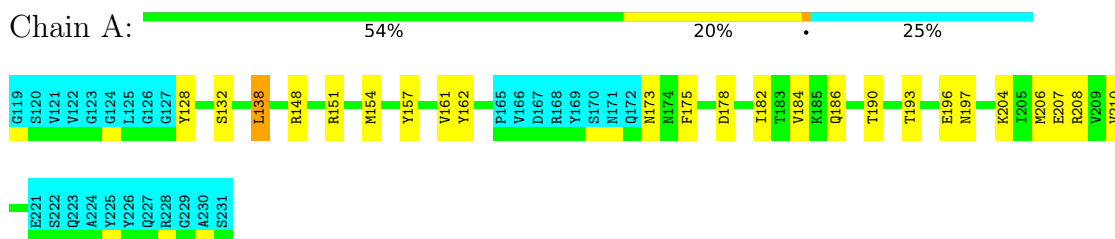### 4.2.3 Score per residue for model 3

• Molecule 1: Major prion protein

Chain A: 59% 13% • 25%

G119 S120 V121 V122 G123 G124 L125 G126 G127 S132 A133 L138 M143 D147 R148 Y149 Y150 M154 Y155 R156 Y157 P165 V166 D167 R168 Y169 S170 N171 Q172 D178 Q186 G195 E196 N197 F198 R208 E221 S222 Q223 A224 Y225 Y226 Q227 R228 G229 A230 S231

4.2.4 Score per residue for model 4

• Molecule 1: Major prion protein

Chain A: 56% 19% • 25%

G119 S120 V121 V122 G123 G124 L125 G126 G127 Y128 S135 L138 D144 D147 Y150 R156 P165 V166 D167 R168 Y169 S170 N171 Q172 N173 N174 F175 D178 I182 T183 V184 K185 Q186 H187 K194 G195 E196 K204 I205 M206 E207 R208 V209 V210 C214 E221 S222

Q223 A224 Y225 Y226 Q227 R228 G229 A230 S231

4.2.5 Score per residue for model 5

• Molecule 1: Major prion protein

Chain A: 54% 20% • 25%

G119 S120 V121 V122 G123 G124 L125 G126 G127 Y128 S132 L138 R148 R151 M154 Y157 V161 Y162 P165 V166 D167 R168 Y169 S170 N171 Q172 N173 N174 F175 D178 I182 T183 V184 K185 Q186 T190 T193 E196 N197 K204 I205 M206 E207 R208 V209 V210

E221 S222 Q223 A224 Y225 Y226 Q227 R228 G229 A230 S231

4.2.6 Score per residue for model 6

• Molecule 1: Major prion protein

Chain A: 56% 19% 25%

G119 S120 V121 V122 G123 G124 L125 G126 G127 Y128 S132 R136 P137 L138 Y150 Y155 N159 Y162 P165 V166 D167 R168 Y169 S170 N171 Q172 N173 N174 F175 D178 N181 Q186 T192 E196 N197 D202 I203 K204 E207 R208 C214 Q217 E221

S222 Q223 A224 Y225 Y226 Q227 R228 G229 A230 S231

4.2.7 Score per residue for model 7

• Molecule 1: Major prion protein

Chain A:

4.2.8 Score per residue for model 8

• Molecule 1: Major prion protein

Chain A:

4.2.9 Score per residue for model 9

• Molecule 1: Major prion protein

Chain A:

4.2.10 Score per residue for model 10

• Molecule 1: Major prion protein

Chain A:

4.2.11 Score per residue for model 11

• Molecule 1: Major prion protein

Chain A:

4.2.12 Score per residue for model 12 (medoid)

• Molecule 1: Major prion protein

4.2.13 Score per residue for model 13

• Molecule 1: Major prion protein

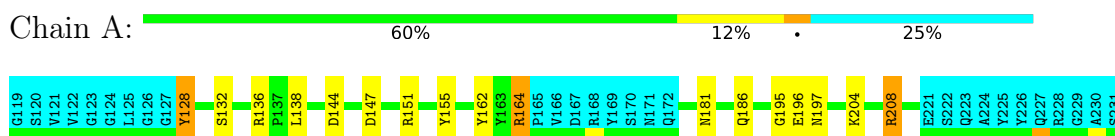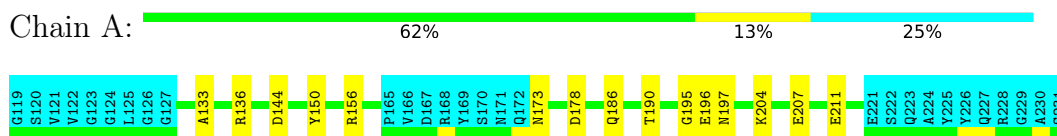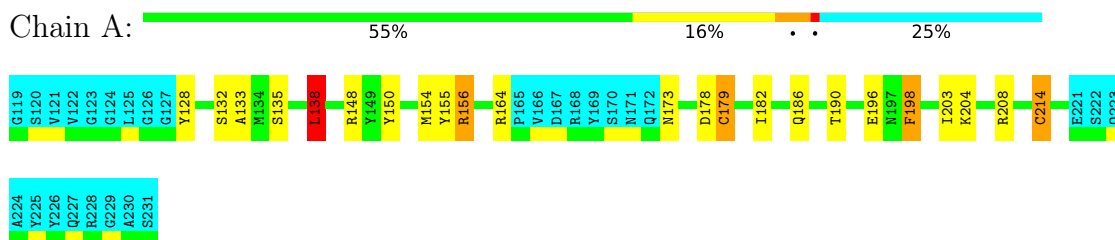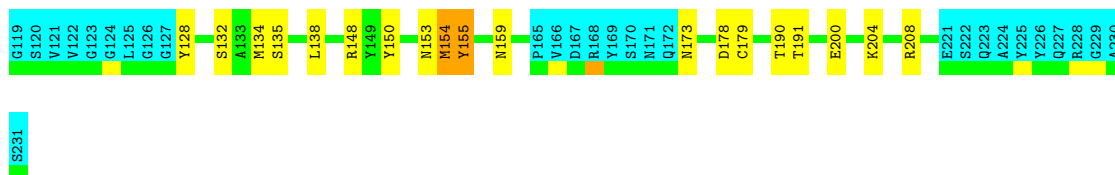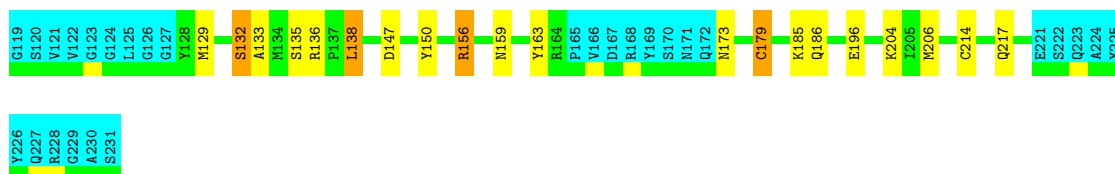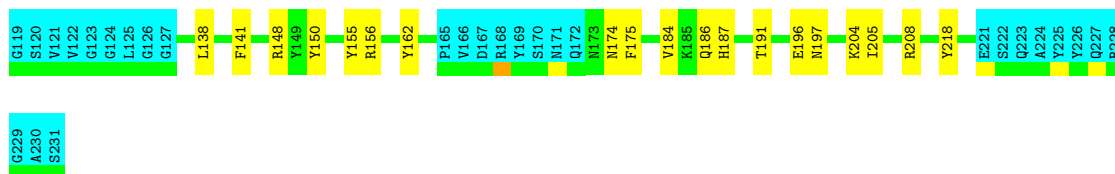## 4.2.14 Score per residue for model 14

• Molecule 1: Major prion protein

4.2.15 Score per residue for model 15

• Molecule 1: Major prion protein

G119 S120 V121 V122 G123 G124 L125 G126 G127 A133 M134 S135 R136 D144 D147 R148 Y149 Y150 Y155 Q160 P165 V166 D167 R168 Y169 S170 N171 Q172 H177 D178 C179 V180 N181 Q186 T190 G195 E196 N197 F198 T199 K204 E207 E221 S222 Q223 A224 Y225

Y226 Q227 R228 G229 A230 S231
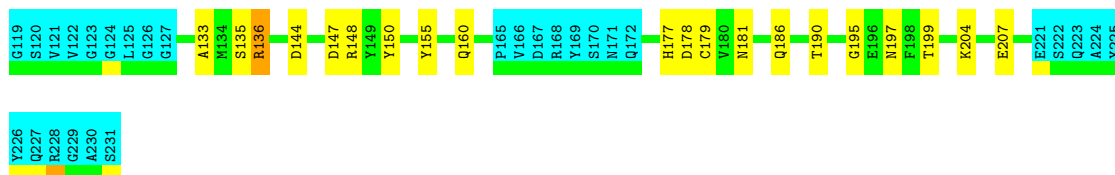
4.2.16 Score per residue for model 16

• Molecule 1: Major prion protein

Chain A: 54% 19% • 25%

G119 S120 V121 V122 G123 G124 L125 G126 G127 S132 A133 M134 S135 R136 P137 L138 N153 R156 Y157 Y162 P165 V166 D167 R168 Y169 S170 N171 Q172 N173 N174 D178 Q186 H187 T188 T192 T193 K194 G195 E196 N197 F198 T199 E200 T201 D202 I203 K204 I205 M206 E207 R208

Y218 E221 S222 Q223 A224 Y225 Y226 Q227 R228 G229 A230 S231
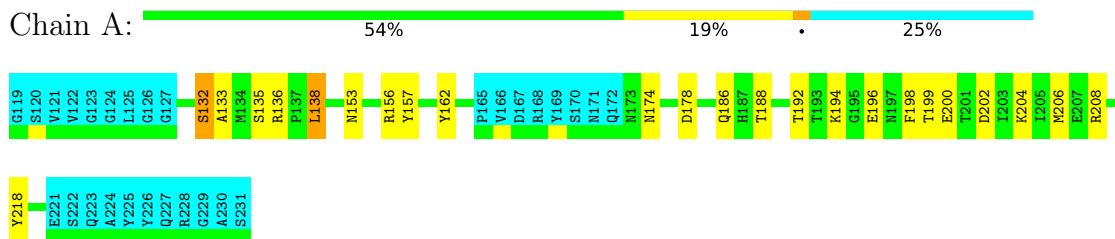
4.2.17 Score per residue for model 17

• Molecule 1: Major prion protein

Chain A: 56% 18% • 25%

G119 S120 V121 V122 G123 G124 L125 G126 G127 Y128 S132 A133 R136 P137 L138 D147 Y150 R151 N159 Y162 Y163 R164 P165 V166 D167 R168 Y169 S170 N171 Q172 N173 N174 F175 V176 H177 N181 V184 K185 Q186 T190 N197 F198 D202 I203 K204 E221 S222 Q223

A224 Y225 Y226 Q227 R228 G229 A230 S231
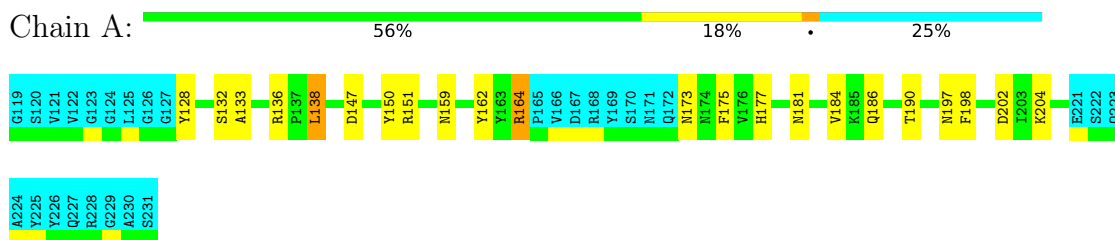
4.2.18 Score per residue for model 18

• Molecule 1: Major prion protein

Chain A: 58% 14% • 25%

G119 S120 V121 V122 G123 G124 L125 G126 G127 L130 S135 R136 P137 L138 D147 R148 Y149 Y150 R151 E152 N153 M154 Y155 Y163 R164 P165 V166 D167 R168 Y169 S170 N171 Q172 N173 Q186 H187 T190 G195 E196 N197 R208 E221 S222 Q223 A224 Y225 Y226 Q227 R228 G229 A230

S231
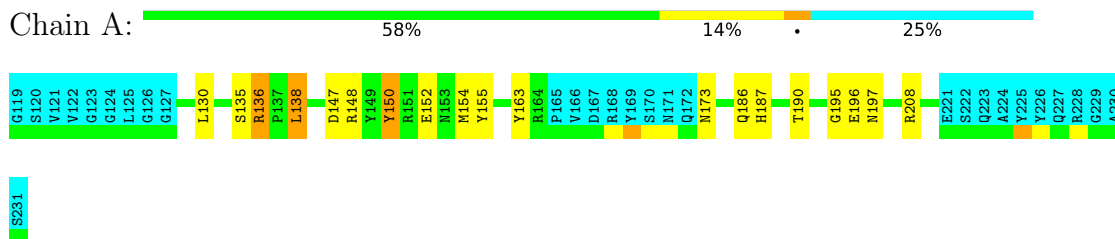
4.2.19 Score per residue for model 19

• Molecule 1: Major prion protein

Chain A:

52%   19%   . .   25%

G119 S120 V121 V122 G123 G124 L125 G126 G127 Y128 M129 L130 G131 S132 A133 M134 L138 D147 R148 Y155 R156 Y157 Q160 Y163 R164 P165 V166 D167 R168 Y169 S170 N171 Q172 F175 H187 T191 E196 N197 F198 K204 E207 R208 Q217 Y218 Q219 R220 E221 S222

Q223 A224 Y225 Y226 Q227 R228 G229 A230 S231
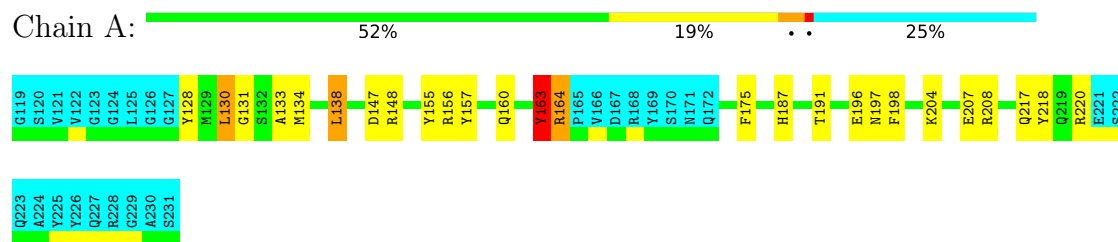
4.2.20   Score per residue for model 20

• Molecule 1: Major prion protein

Chain A:

62%   12%   •   25%

G119 S120 V121 V122 G123 G124 L125 G126 G127 Y128 S132 A133 M134 S135 R136 P137 L138 Y150 Y162 Y163 R164 P165 V166 D167 R168 Y169 S170 N171 Q172 D178 Q186 F198 T201 K204 E211 E221 S222 Q223 A224 Y225 Y226 Q227 R228 G229 A230 S231
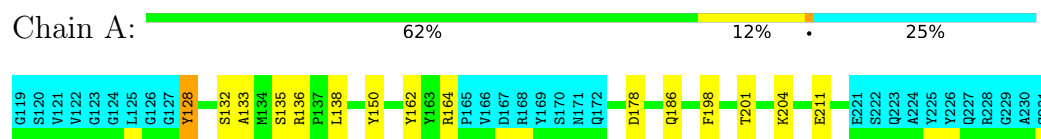
5 Refinement protocol and experimental data overview ⓘ

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function*.

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

Software name	Classification	Version
DYANA	structure solution	
OPALp	refinement	

No chemical shift data was provided.

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	0.68±0.01	0±0/742 (0.0± 0.0%)	1.07±0.03	1±1/1005 (0.1± 0.1%)
All	All	0.68	1/14840 (0.0%)	1.07	22/20100 (0.1%)

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	A	0.0±0.0	2.4±1.6
All	All	0	47

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	179	CYS	CB-SG	-6.31	1.71	1.82	10	1

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
1	A	156	ARG	NE-CZ-NH2	-6.79	116.90	120.30	9	3
1	A	184	VAL	CA-CB-CG1	6.71	120.96	110.90	1	5
1	A	151	ARG	CD-NE-CZ	6.14	132.19	123.60	5	1
1	A	162	TYR	CB-CG-CD2	-5.97	117.42	121.00	5	1
1	A	148	ARG	NE-CZ-NH2	-5.74	117.43	120.30	2	1
1	A	150	TYR	CB-CG-CD2	-5.59	117.65	121.00	20	2
1	A	179	CYS	CA-CB-SG	5.55	123.99	114.00	12	1
1	A	156	ARG	NE-CZ-NH1	5.42	123.01	120.30	9	2
1	A	163	TYR	CB-CG-CD2	-5.33	117.80	121.00	2	1
1	A	132	SER	C-N-CA	5.24	134.80	121.70	12	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models Worst	Models Total
1	A	151	ARG	NE-CZ-NH2	-5.17	117.72	120.30	17	1
1	A	130	LEU	CB-CG-CD2	5.06	119.60	111.00	19	1
1	A	128	TYR	CB-CG-CD1	-5.04	117.98	121.00	8	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)
1	A	155	TYR	Sidechain	7
1	A	162	TYR	Sidechain	6
1	A	128	TYR	Sidechain	5
1	A	157	TYR	Sidechain	5
1	A	208	ARG	Sidechain	5
1	A	156	ARG	Sidechain	4
1	A	220	ARG	Sidechain	2
1	A	164	ARG	Sidechain	2
1	A	163	TYR	Sidechain,Peptide	2
1	A	148	ARG	Sidechain	2
1	A	218	TYR	Sidechain	2
1	A	149	TYR	Sidechain	1
1	A	132	SER	Peptide	1
1	A	151	ARG	Sidechain	1
1	A	195	GLY	Peptide	1
1	A	136	ARG	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	725	687	687	1±1
All	All	14500	13740	13740	21

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 1.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:138:LEU:HD21	1:A:154:MET:CE	0.65	2.21	11	1
1:A:206:MET:O	1:A:210:VAL:HG23	0.49	2.06	5	2
1:A:138:LEU:HD21	1:A:154:MET:HE1	0.47	1.87	11	1
1:A:138:LEU:HA	1:A:150:TYR:CE2	0.45	2.47	18	1
1:A:198:PHE:CE2	1:A:203:ILE:HD11	0.45	2.47	10	1
1:A:188:THR:O	1:A:192:THR:HG23	0.45	2.11	16	1
1:A:179:CYS:SG	1:A:214:CYS:CB	0.45	3.05	10	2
1:A:138:LEU:HD23	1:A:154:MET:CE	0.45	2.42	5	2
1:A:128:TYR:CE2	1:A:162:TYR:CD2	0.44	3.05	20	1
1:A:141:PHE:CZ	1:A:205:ILE:HG23	0.43	2.48	14	1
1:A:138:LEU:HD21	1:A:154:MET:HE3	0.43	1.90	11	1
1:A:187:HIS:CE1	1:A:198:PHE:CZ	0.42	3.07	13	1
1:A:128:TYR:CE2	1:A:162:TYR:CE2	0.42	3.08	7	1
1:A:187:HIS:O	1:A:191:THR:HG23	0.42	2.14	19	1
1:A:130:LEU:HD22	1:A:160:GLN:CG	0.42	2.45	19	1
1:A:179:CYS:CB	1:A:214:CYS:SG	0.41	3.06	12	1
1:A:131:GLY:HA2	1:A:163:TYR:CE1	0.41	2.51	19	1
1:A:128:TYR:CE2	1:A:182:ILE:HG21	0.40	2.51	5	1

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	85/113 (75%)	75±2 (88±2%)	7±2 (8±2%)	3±1 (4±2%)	6	35
All	All	1700/2260 (75%)	1504 (88%)	136 (8%)	60 (4%)	6	35

All 12 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	196	GLU	16
1	A	138	LEU	10
1	A	133	ALA	9
1	A	136	ARG	7
1	A	195	GLY	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	132	SER	4
1	A	164	ARG	4
1	A	128	TYR	2
1	A	198	PHE	1
1	A	192	THR	1
1	A	135	SER	1
1	A	152	GLU	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	81/101 (80%)	68±3 (84±4%)	13±3 (16±4%)	5	41
All	All	1620/2020 (80%)	1355 (84%)	265 (16%)	5	41

All 56 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	186	GLN	16
1	A	204	LYS	15
1	A	138	LEU	13
1	A	150	TYR	12
1	A	178	ASP	12
1	A	197	ASN	12
1	A	135	SER	11
1	A	173	ASN	10
1	A	148	ARG	9
1	A	190	THR	8
1	A	132	SER	8
1	A	147	ASP	8
1	A	208	ARG	8
1	A	175	PHE	7
1	A	207	GLU	7
1	A	198	PHE	6
1	A	159	ASN	6
1	A	155	TYR	5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	199	THR	5
1	A	136	ARG	5
1	A	163	TYR	4
1	A	202	ASP	4
1	A	144	ASP	4
1	A	128	TYR	4
1	A	181	ASN	4
1	A	143	ASN	3
1	A	156	ARG	3
1	A	154	MET	3
1	A	187	HIS	3
1	A	214	CYS	3
1	A	217	GLN	3
1	A	179	CYS	3
1	A	164	ARG	3
1	A	177	HIS	3
1	A	192	THR	2
1	A	182	ILE	2
1	A	194	LYS	2
1	A	161	VAL	2
1	A	193	THR	2
1	A	160	GLN	2
1	A	211	GLU	2
1	A	134	MET	2
1	A	153	ASN	2
1	A	191	THR	2
1	A	200	GLU	2
1	A	206	MET	2
1	A	174	ASN	2
1	A	139	ILE	1
1	A	220	ARG	1
1	A	196	GLU	1
1	A	145	TYR	1
1	A	129	MET	1
1	A	185	LYS	1
1	A	130	LEU	1
1	A	218	TYR	1
1	A	201	THR	1

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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 ⓘ

There are no ligands in this entry.

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 ⓘ

No chemical shift data were provided