



# Full wwPDB NMR Structure Validation Report ⓘ

Nov 1, 2021 – 03:56 PM EDT

PDB ID : 2K5O  
Title : Mouse Prion Protein (121-231) with Mutation S170N  
Authors : Perez, D.R.; Wuthrich, K.  
Deposited on : 2008-06-30

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

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 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 : 2.23.2  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

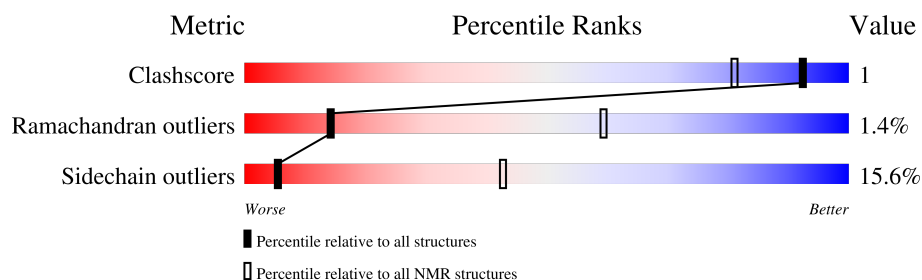
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 62%.

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

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  $\geq 3$ , 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	114	

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

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:127-A:226 (100)	0.47	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 4 clusters and 2 single-model clusters were found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1805 atoms, of which 870 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Major prion protein.

Mol	Chain	Residues	Atoms						Trace
1	A	113	Total	C	H	N	O	S	0
			1805	579	870	165	182	9	

There are 3 discrepancies between the modelled and reference sequences:

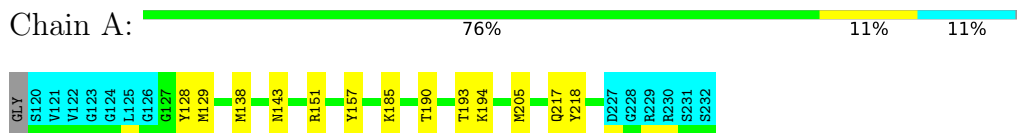
Chain	Residue	Modelled	Actual	Comment	Reference
A	119	GLY	-	expression tag	UNP P04925
A	120	SER	-	expression tag	UNP P04925
A	170	ASN	SER	engineered mutation	UNP P04925

## 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: Major prion protein

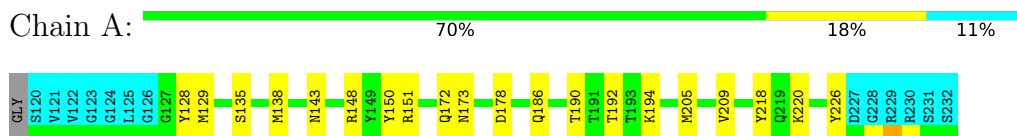


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

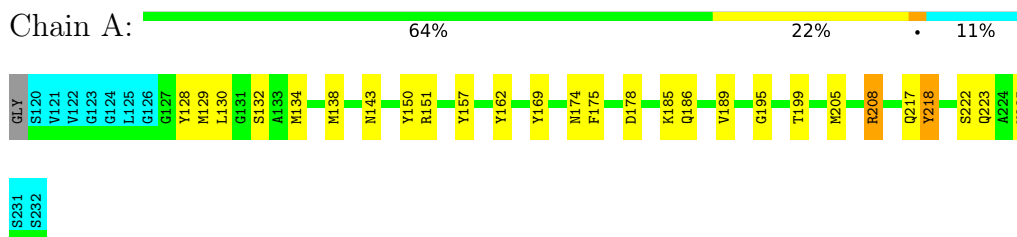
#### 4.2.1 Score per residue for model 1 (medoid)

- Molecule 1: Major prion protein



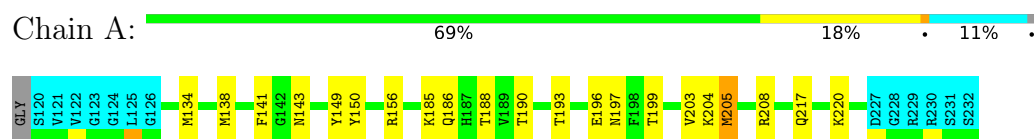
#### 4.2.2 Score per residue for model 2

- Molecule 1: Major prion protein



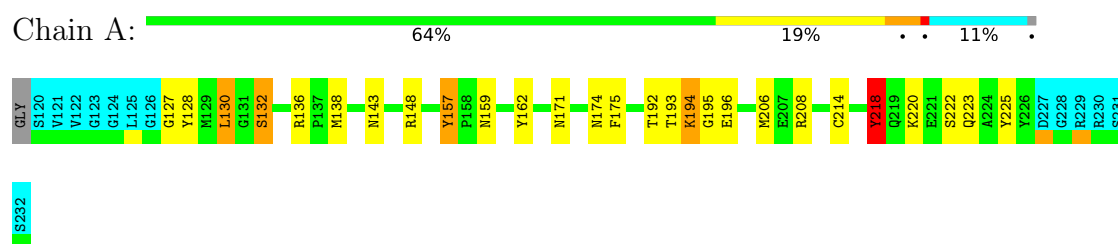
### 4.2.3 Score per residue for model 3

- Molecule 1: Major prion protein



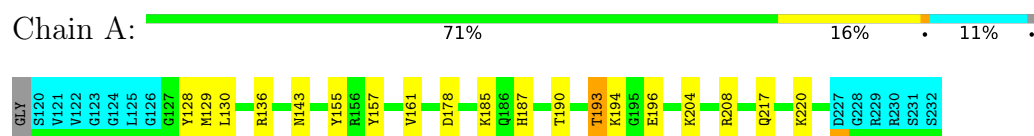
### 4.2.4 Score per residue for model 4

- Molecule 1: Major prion protein



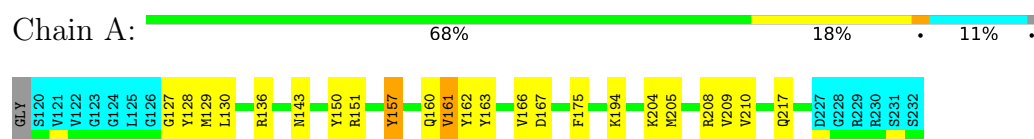
### 4.2.5 Score per residue for model 5

- Molecule 1: Major prion protein



### 4.2.6 Score per residue for model 6

- Molecule 1: Major prion protein



### 4.2.7 Score per residue for model 7

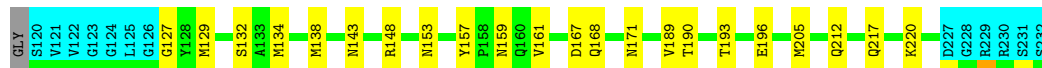
- Molecule 1: Major prion protein





#### 4.2.8 Score per residue for model 8

- Molecule 1: Major prion protein



#### 4.2.9 Score per residue for model 9

- Molecule 1: Major prion protein



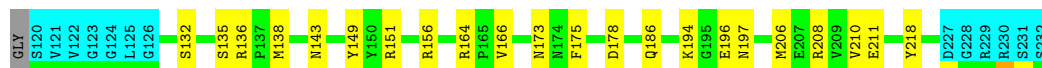
#### 4.2.10 Score per residue for model 10

- Molecule 1: Major prion protein



#### 4.2.11 Score per residue for model 11

- Molecule 1: Major prion protein



#### 4.2.12 Score per residue for model 12

- Molecule 1: Major prion protein

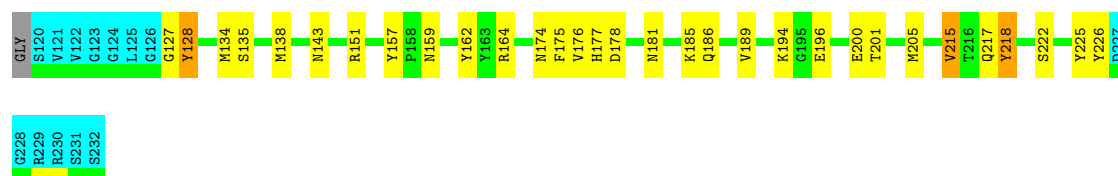
Chain A:  70% 18% 11%



#### 4.2.13 Score per residue for model 13

- Molecule 1: Major prion protein

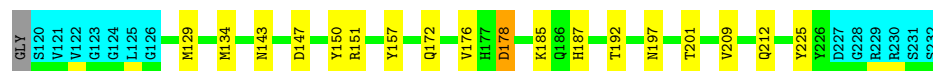
Chain A:  61% 25% 11%



#### 4.2.14 Score per residue for model 14

- Molecule 1: Major prion protein

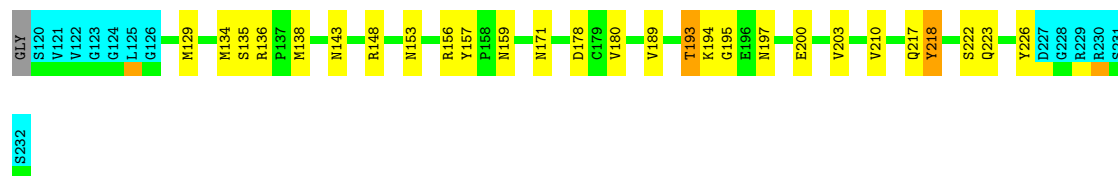
Chain A:  72% 15% 11%



#### 4.2.15 Score per residue for model 15

- Molecule 1: Major prion protein

Chain A:  64% 22% 11%



#### 4.2.16 Score per residue for model 16

- Molecule 1: Major prion protein

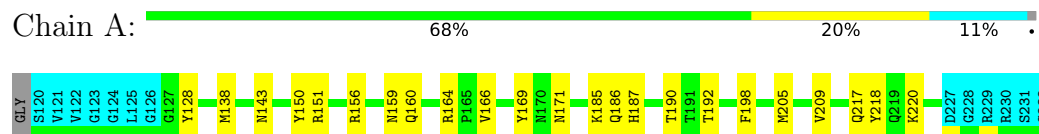
Chain A:  72% 12% 11%





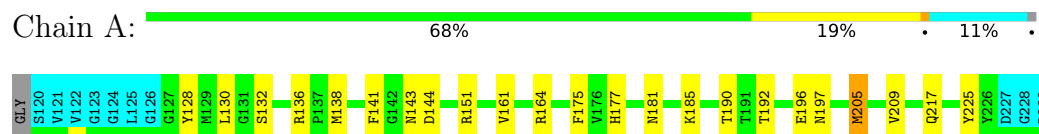
#### 4.2.17 Score per residue for model 17

- Molecule 1: Major prion protein



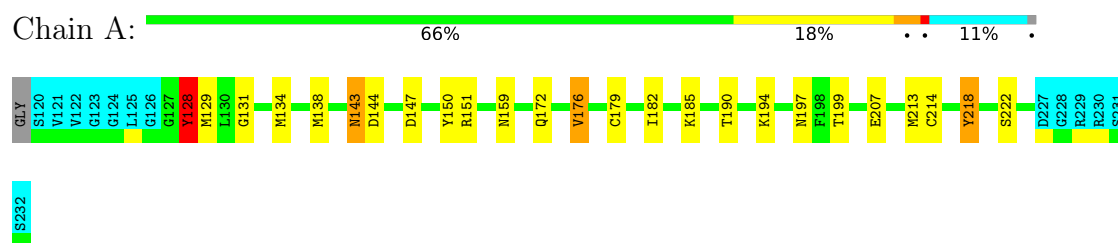
#### 4.2.18 Score per residue for model 18

- Molecule 1: Major prion protein



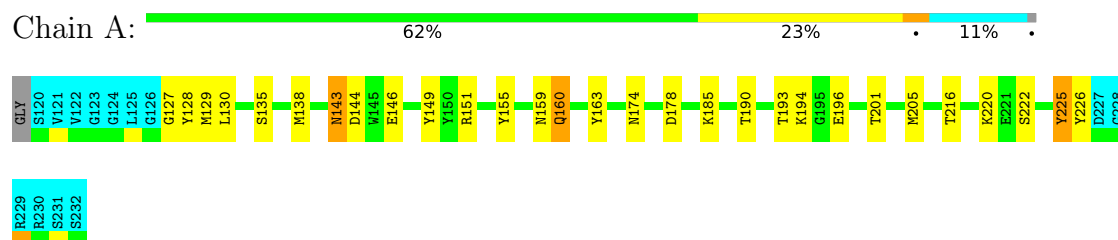
#### 4.2.19 Score per residue for model 19

- Molecule 1: Major prion protein



#### 4.2.20 Score per residue for model 20

- Molecule 1: Major prion protein



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *OPAL (water shell)*.

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	1.0.3
OPAL	refinement	1.2

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	898
Number of shifts mapped to atoms	898
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	62%

## 6 Model quality i

### 6.1 Standard geometry i

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.68±0.01	0±0/871 ( 0.0± 0.0%)	1.10±0.04	2±1/1180 ( 0.2± 0.1%)
All	All	0.68	0/17420 ( 0.0%)	1.10	36/23600 ( 0.2%)

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	2.5±1.6
All	All	0	49

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
1	A	156	ARG	NE-CZ-NH1	8.50	124.55	120.30	16	3
1	A	128	TYR	CB-CG-CD2	-7.71	116.37	121.00	19	1
1	A	209	VAL	CA-CB-CG2	7.32	121.87	110.90	18	1
1	A	157	TYR	CB-CG-CD2	-7.05	116.77	121.00	10	7
1	A	156	ARG	NE-CZ-NH2	-7.03	116.78	120.30	16	2
1	A	150	TYR	CB-CG-CD2	-6.93	116.84	121.00	7	1
1	A	161	VAL	CA-CB-CG1	6.55	120.72	110.90	18	2
1	A	215	VAL	CG1-CB-CG2	-6.30	100.81	110.90	13	1
1	A	166	VAL	CA-CB-CG2	6.22	120.24	110.90	17	1
1	A	136	ARG	NE-CZ-NH2	-6.05	117.27	120.30	9	1
1	A	161	VAL	CA-CB-CG2	5.65	119.37	110.90	10	3
1	A	157	TYR	CB-CG-CD1	5.58	124.35	121.00	10	2
1	A	178	ASP	CB-CG-OD1	5.44	123.20	118.30	14	1
1	A	176	VAL	CG1-CB-CG2	-5.43	102.21	110.90	13	1
1	A	208	ARG	NE-CZ-NH2	-5.34	117.63	120.30	5	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	218	TYR	CB-CG-CD1	-5.27	117.84	121.00	16	2
1	A	193	THR	CA-CB-CG2	5.25	119.75	112.40	15	1
1	A	136	ARG	NE-CZ-NH1	5.20	122.90	120.30	11	1
1	A	208	ARG	NE-CZ-NH1	5.17	122.89	120.30	11	1
1	A	226	TYR	C-N-CA	5.06	134.34	121.70	12	2
1	A	176	VAL	CA-CB-CG1	5.02	118.43	110.90	19	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	128	TYR	Sidechain	10
1	A	157	TYR	Sidechain	6
1	A	169	TYR	Sidechain	3
1	A	151	ARG	Sidechain	3
1	A	164	ARG	Sidechain	3
1	A	148	ARG	Sidechain	2
1	A	226	TYR	Sidechain,Peptide	2
1	A	162	TYR	Sidechain	2
1	A	208	ARG	Sidechain	2
1	A	150	TYR	Sidechain	2
1	A	163	TYR	Sidechain	2
1	A	187	HIS	Sidechain	2
1	A	156	ARG	Sidechain	1
1	A	218	TYR	Sidechain	1
1	A	155	TYR	Sidechain	1
1	A	127	GLY	Peptide	1
1	A	143	ASN	Peptide	1
1	A	192	THR	Peptide	1
1	A	198	PHE	Sidechain	1
1	A	131	GLY	Peptide	1
1	A	225	TYR	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	849	784	784	2±1
All	All	16980	15680	15680	45

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:150:TYR:CZ	1:A:209:VAL:HG21	0.62	2.29	14	4
1:A:143:ASN:HD22	1:A:143:ASN:C	0.53	2.07	20	2
1:A:161:VAL:HG21	1:A:210:VAL:HG13	0.53	1.80	6	1
1:A:200:GLU:HA	1:A:203:VAL:CG2	0.52	2.34	15	1
1:A:200:GLU:HA	1:A:203:VAL:HG22	0.51	1.82	15	1
1:A:130:LEU:HD22	1:A:162:TYR:CE1	0.51	2.41	4	1
1:A:172:GLN:HE22	1:A:176:VAL:HG21	0.51	1.66	14	1
1:A:130:LEU:HD11	1:A:160:GLN:HB3	0.49	1.85	20	1
1:A:150:TYR:CE2	1:A:209:VAL:HG21	0.48	2.43	17	3
1:A:143:ASN:HD21	1:A:146:GLU:HB2	0.48	1.69	20	1
1:A:199:THR:O	1:A:203:VAL:HG23	0.48	2.08	3	1
1:A:206:MET:O	1:A:210:VAL:HG12	0.48	2.08	11	1
1:A:161:VAL:CG2	1:A:210:VAL:HG13	0.47	2.40	6	1
1:A:149:TYR:CE1	1:A:153:ASN:ND2	0.47	2.83	9	1
1:A:176:VAL:HG23	1:A:214:CYS:SG	0.46	2.50	19	1
1:A:192:THR:HG21	1:A:196:GLU:HB2	0.46	1.87	18	1
1:A:166:VAL:HG22	1:A:175:PHE:CE2	0.46	2.45	6	2
1:A:130:LEU:HD13	1:A:162:TYR:CE1	0.46	2.45	6	1
1:A:187:HIS:HA	1:A:190:THR:HG22	0.46	1.87	5	2
1:A:141:PHE:CZ	1:A:205:MET:HG3	0.45	2.46	18	2
1:A:128:TYR:CZ	1:A:182:ILE:HG13	0.45	2.46	19	2
1:A:218:TYR:CZ	1:A:222:SER:HB3	0.45	2.47	19	2
1:A:177:HIS:CD2	1:A:181:ASN:ND2	0.44	2.85	18	2
1:A:211:GLU:O	1:A:215:VAL:HG23	0.43	2.13	16	1
1:A:218:TYR:CE2	1:A:222:SER:HB2	0.43	2.49	15	1
1:A:128:TYR:CZ	1:A:164:ARG:HG3	0.42	2.49	13	1
1:A:218:TYR:CE1	1:A:222:SER:HB2	0.42	2.50	13	1
1:A:161:VAL:HG11	1:A:214:CYS:SG	0.42	2.55	12	1
1:A:157:TYR:CD2	1:A:206:MET:HG2	0.41	2.49	4	1
1:A:128:TYR:CE1	1:A:164:ARG:HG3	0.41	2.50	18	1
1:A:218:TYR:CZ	1:A:222:SER:HB2	0.41	2.49	4	1
1:A:161:VAL:HG21	1:A:214:CYS:SG	0.41	2.56	10	1
1:A:180:VAL:HG22	1:A:210:VAL:HG12	0.40	1.92	15	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	100/114 (88%)	90±2 (90±2%)	8±2 (8±2%)	1±1 (1±1%)	14	59
All	All	2000/2280 (88%)	1807 (90%)	164 (8%)	29 (1%)	14	59

All 6 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	193	THR	8
1	A	194	LYS	8
1	A	127	GLY	5
1	A	195	GLY	4
1	A	192	THR	2
1	A	132	SER	2

### 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	94/103 (91%)	79±4 (84±4%)	15±4 (16±4%)	5	43
All	All	1880/2060 (91%)	1587 (84%)	293 (16%)	5	43

All 62 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	143	ASN	20
1	A	138	MET	14
1	A	129	MET	12
1	A	217	GLN	12

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	185	LYS	11
1	A	205	MET	10
1	A	178	ASP	9
1	A	190	THR	9
1	A	218	TYR	9
1	A	220	LYS	9
1	A	134	MET	9
1	A	159	ASN	9
1	A	151	ARG	8
1	A	186	GLN	8
1	A	189	VAL	8
1	A	225	TYR	8
1	A	196	GLU	8
1	A	197	ASN	8
1	A	135	SER	6
1	A	208	ARG	6
1	A	130	LEU	5
1	A	174	ASN	5
1	A	175	PHE	5
1	A	132	SER	5
1	A	136	ARG	5
1	A	144	ASP	5
1	A	194	LYS	4
1	A	193	THR	4
1	A	171	ASN	4
1	A	160	GLN	4
1	A	173	ASN	3
1	A	223	GLN	3
1	A	149	TYR	3
1	A	204	LYS	3
1	A	167	ASP	3
1	A	201	THR	3
1	A	172	GLN	2
1	A	150	TYR	2
1	A	199	THR	2
1	A	148	ARG	2
1	A	153	ASN	2
1	A	212	GLN	2
1	A	211	GLU	2
1	A	226	TYR	2
1	A	147	ASP	2
1	A	156	ARG	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	188	THR	1
1	A	192	THR	1
1	A	214	CYS	1
1	A	168	GLN	1
1	A	198	PHE	1
1	A	200	GLU	1
1	A	215	VAL	1
1	A	141	PHE	1
1	A	170	ASN	1
1	A	181	ASN	1
1	A	179	CYS	1
1	A	207	GLU	1
1	A	213	MET	1
1	A	155	TYR	1
1	A	216	THR	1
1	A	222	SER	1

### 6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 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 ⓘ

There are no chain breaks in this entry.

## 7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 62% 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: *assigned\_chem\_shift\_list\_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	898
Number of shifts mapped to atoms	898
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

#### 7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	113	$0.23 \pm 0.09$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	105	$0.75 \pm 0.09$	Should be applied
$^{13}\text{C}'$	0	—	None (insufficient data)
$^{15}\text{N}$	110	$0.18 \pm 0.23$	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 62%, i.e. 809 atoms were assigned a chemical shift out of a possible 1296. 10 out of 10 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	293/494 (59%)	96/197 (49%)	100/200 (50%)	97/97 (100%)
Sidechain	462/654 (71%)	256/386 (66%)	184/229 (80%)	22/39 (56%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	54/148 (36%)	14/77 (18%)	39/67 (58%)	1/4 (25%)
Overall	809/1296 (62%)	366/660 (55%)	323/496 (65%)	120/140 (86%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 61%, i.e. 879 atoms were assigned a chemical shift out of a possible 1431. 13 out of 13 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	328/559 (59%)	105/223 (47%)	113/226 (50%)	110/110 (100%)
Sidechain	497/724 (69%)	269/427 (63%)	204/252 (81%)	24/45 (53%)
Aromatic	54/148 (36%)	14/77 (18%)	39/67 (58%)	1/4 (25%)
Overall	879/1431 (61%)	388/727 (53%)	356/545 (65%)	135/159 (85%)

#### 7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

#### 7.1.5 Random Coil Index (RCI) plots [i](#)

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.

Random coil index (RCI) for chain A:

