



# Full wwPDB NMR Structure Validation Report ⓘ

Oct 3, 2021 – 12:41 PM EDT

PDB ID : 2KU5  
Title : Mouse Prion Protein (121-231) with mutation D167S  
Authors : Perez, D.R.; Damberger, F.F.; Wuthrich, K.  
Deposited on : 2010-02-12

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

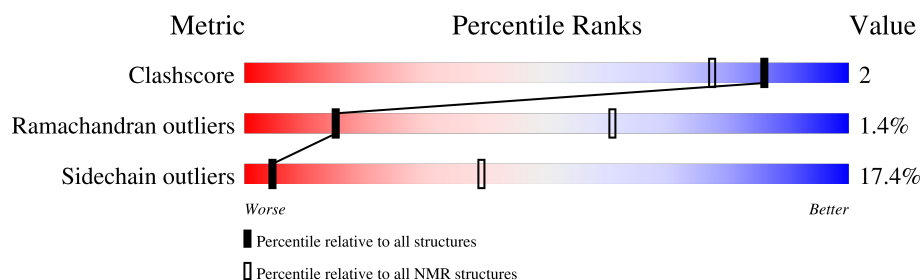
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 82%.

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	113	<div> <div></div> <div>75%</div> <div>15%</div> <div>10%</div> </div>

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 19 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

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:125-A:226 (102)	0.33	19

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 1 clusters and 1 single-model cluster was found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1801 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
			1801	577	870	164	181	9	

There are 2 discrepancies between the modelled and reference sequences:

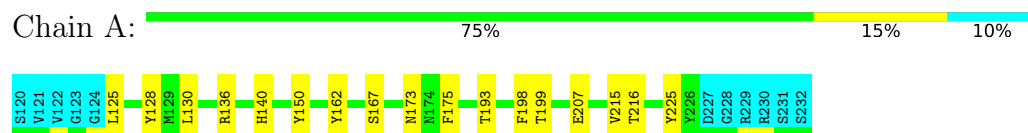
Chain	Residue	Modelled	Actual	Comment	Reference
A	120	SER	-	expression tag	UNP P04925
A	167	SER	ASP	engineered mutation	UNP P04925

## 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

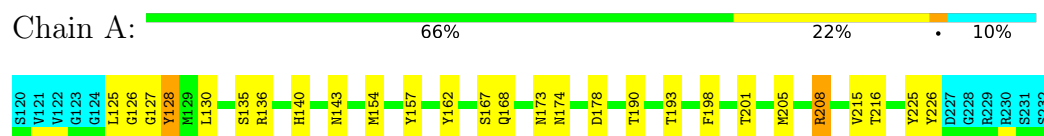


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

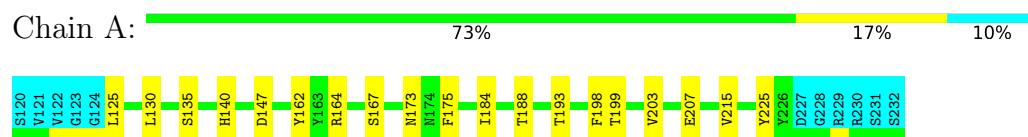
#### 4.2.1 Score per residue for model 1

- Molecule 1: Major prion protein



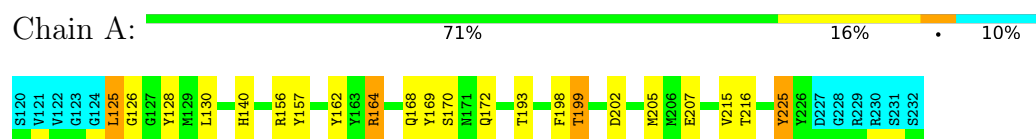
#### 4.2.2 Score per residue for model 2

- Molecule 1: Major prion protein



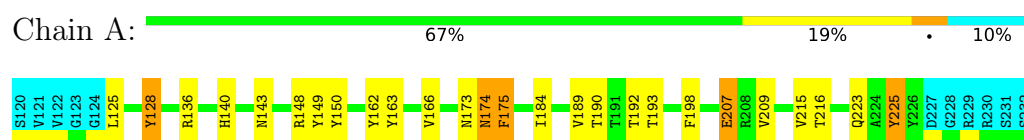
### 4.2.3 Score per residue for model 3

- Molecule 1: Major prion protein



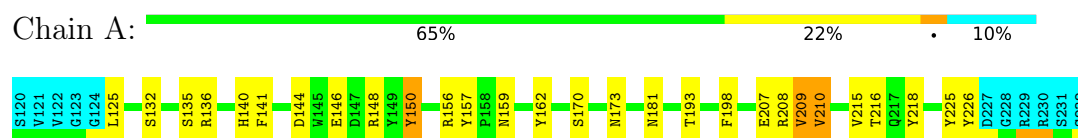
### 4.2.4 Score per residue for model 4

- Molecule 1: Major prion protein



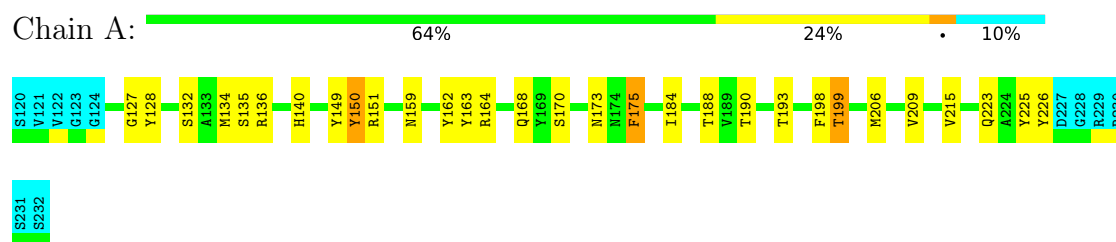
### 4.2.5 Score per residue for model 5

- Molecule 1: Major prion protein



### 4.2.6 Score per residue for model 6

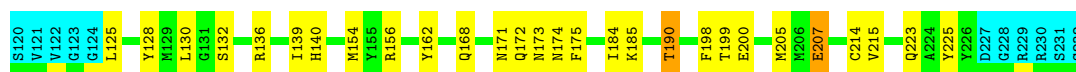
- Molecule 1: Major prion protein



### 4.2.7 Score per residue for model 7

- Molecule 1: Major prion protein





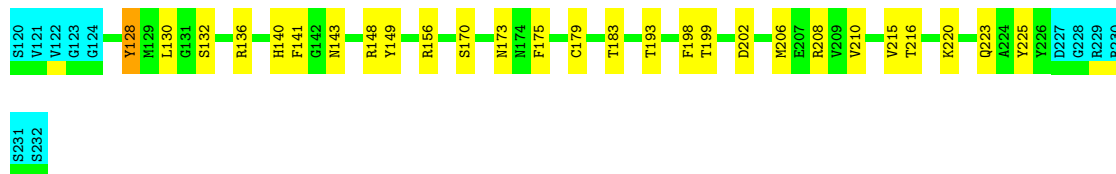
#### 4.2.8 Score per residue for model 8

- Molecule 1: Major prion protein



#### 4.2.9 Score per residue for model 9

- Molecule 1: Major prion protein



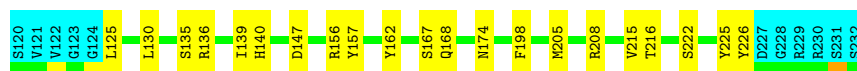
#### 4.2.10 Score per residue for model 10

- Molecule 1: Major prion protein



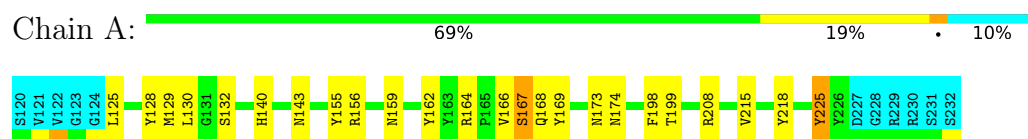
#### 4.2.11 Score per residue for model 11

- Molecule 1: Major prion protein



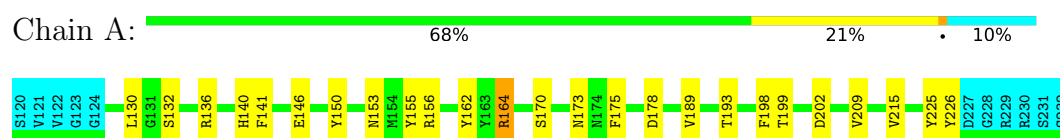
### 4.2.12 Score per residue for model 12

- Molecule 1: Major prion protein



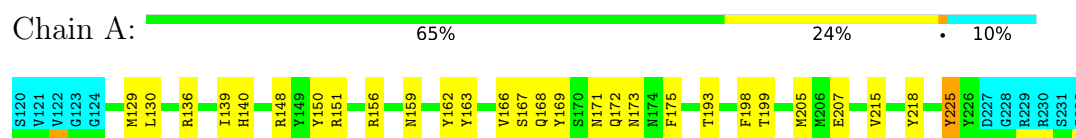
### 4.2.13 Score per residue for model 13

- Molecule 1: Major prion protein



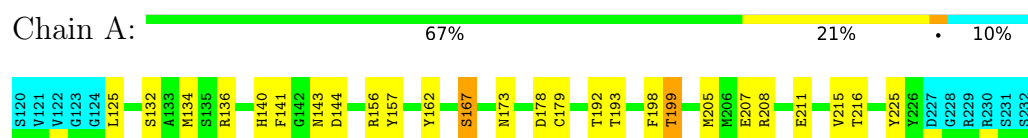
### 4.2.14 Score per residue for model 14

- Molecule 1: Major prion protein



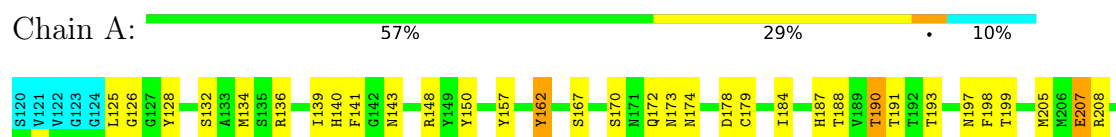
### 4.2.15 Score per residue for model 15

- Molecule 1: Major prion protein



### 4.2.16 Score per residue for model 16

- Molecule 1: Major prion protein



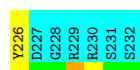
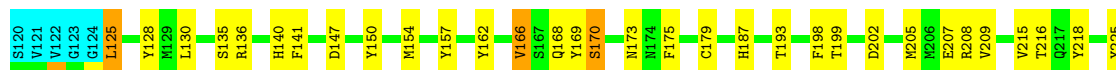




#### 4.2.17 Score per residue for model 17

- Molecule 1: Major prion protein

Chain A: 61% 27% 10%



#### 4.2.18 Score per residue for model 18

- Molecule 1: Major prion protein

Chain A: 66% 21% 10%



#### 4.2.19 Score per residue for model 19 (medoid)

- Molecule 1: Major prion protein

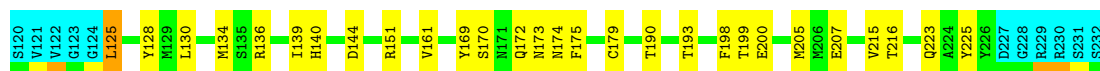
Chain A: 68% 21% 10%



#### 4.2.20 Score per residue for model 20

- Molecule 1: Major prion protein

Chain A: 65% 24% 10%



## 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
OPALp	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	1291
Number of shifts mapped to atoms	1291
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	82%

## 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	0.69±0.01	0±0/879 ( 0.0± 0.0%)	1.12±0.03	1±1/1190 ( 0.1± 0.1%)
All	All	0.69	0/17580 ( 0.0%)	1.12	29/23800 ( 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	3.2±1.5
All	All	0	64

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	210	VAL	CA-CB-CG2	12.43	129.55	110.90	5	1
1	A	198	PHE	CB-CG-CD2	-8.51	114.84	120.80	18	1
1	A	203	VAL	CA-CB-CG2	7.46	122.09	110.90	2	1
1	A	226	TYR	CB-CG-CD2	-6.90	116.86	121.00	17	2
1	A	178	ASP	CB-CG-OD1	6.42	124.08	118.30	13	1
1	A	148	ARG	NE-CZ-NH2	-6.38	117.11	120.30	9	2
1	A	199	THR	CA-CB-CG2	-6.38	103.47	112.40	3	6
1	A	218	TYR	CB-CG-CD2	-6.14	117.32	121.00	12	1
1	A	208	ARG	NE-CZ-NH1	5.98	123.29	120.30	15	2
1	A	175	PHE	CB-CG-CD1	5.81	124.86	120.80	6	1
1	A	147	ASP	CB-CG-OD1	5.73	123.45	118.30	2	1
1	A	156	ARG	NE-CZ-NH1	5.67	123.14	120.30	3	1
1	A	169	TYR	CB-CG-CD2	-5.66	117.60	121.00	12	1
1	A	164	ARG	NE-CZ-NH1	5.66	123.13	120.30	3	1
1	A	198	PHE	CB-CG-CD1	5.41	124.59	120.80	18	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	198	PHE	C-N-CA	5.36	135.09	121.70	18	1
1	A	156	ARG	NE-CZ-NH2	-5.29	117.66	120.30	9	1
1	A	208	ARG	NE-CZ-NH2	-5.24	117.68	120.30	9	1
1	A	225	TYR	CB-CG-CD2	-5.21	117.88	121.00	4	1
1	A	209	VAL	CA-CB-CG2	5.08	118.52	110.90	5	1
1	A	128	TYR	CB-CG-CD2	-5.02	117.99	121.00	20	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	9
1	A	157	TYR	Sidechain	9
1	A	156	ARG	Sidechain	8
1	A	150	TYR	Sidechain	6
1	A	226	TYR	Sidechain	4
1	A	162	TYR	Sidechain	4
1	A	208	ARG	Sidechain	3
1	A	164	ARG	Sidechain	3
1	A	148	ARG	Sidechain	3
1	A	151	ARG	Sidechain	3
1	A	225	TYR	Sidechain	2
1	A	149	TYR	Sidechain	2
1	A	163	TYR	Sidechain	2
1	A	169	TYR	Sidechain	2
1	A	218	TYR	Sidechain	1
1	A	179	CYS	Peptide	1
1	A	170	SER	Peptide	1
1	A	187	HIS	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	857	798	800	3±2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes
All	All	17140	15960	16000	69

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 2.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:130:LEU:HD13	1:A:162:TYR:CE1	0.71	2.21	17	2
1:A:125:LEU:HD11	1:A:162:TYR:CD1	0.71	2.21	1	6
1:A:125:LEU:HD21	1:A:162:TYR:CD1	0.70	2.21	17	1
1:A:125:LEU:HD21	1:A:162:TYR:CE1	0.59	2.32	19	4
1:A:166:VAL:HG11	1:A:218:TYR:CD1	0.57	2.34	17	1
1:A:206:MET:O	1:A:210:VAL:HG12	0.57	2.00	9	1
1:A:184:ILE:O	1:A:188:THR:HG23	0.55	2.00	2	3
1:A:184:ILE:HD11	1:A:207:GLU:HA	0.54	1.79	16	5
1:A:125:LEU:HD12	1:A:128:TYR:HB2	0.54	1.79	1	2
1:A:134:MET:CE	1:A:216:THR:HG21	0.52	2.34	16	3
1:A:130:LEU:HD13	1:A:162:TYR:CE2	0.52	2.38	13	1
1:A:125:LEU:HD11	1:A:162:TYR:CG	0.52	2.39	7	4
1:A:130:LEU:HD13	1:A:162:TYR:CD1	0.51	2.40	17	1
1:A:190:THR:O	1:A:193:THR:HG22	0.49	2.08	4	1
1:A:141:PHE:CD1	1:A:146:GLU:HB3	0.48	2.43	5	4
1:A:174:ASN:ND2	1:A:175:PHE:H	0.48	2.07	4	1
1:A:187:HIS:CE1	1:A:191:THR:HG21	0.47	2.44	16	1
1:A:150:TYR:CE1	1:A:209:VAL:HG21	0.47	2.45	17	1
1:A:172:GLN:HG3	1:A:218:TYR:CD2	0.46	2.45	14	1
1:A:150:TYR:CE1	1:A:209:VAL:HG11	0.46	2.44	18	5
1:A:167:SER:HA	1:A:225:TYR:CE1	0.46	2.46	18	3
1:A:211:GLU:O	1:A:215:VAL:HG23	0.45	2.11	15	2
1:A:134:MET:CE	1:A:216:THR:CG2	0.45	2.95	16	2
1:A:166:VAL:HG13	1:A:175:PHE:CD2	0.45	2.47	17	1
1:A:129:MET:SD	1:A:163:TYR:CE2	0.44	3.10	14	1
1:A:176:VAL:HA	1:A:179:CYS:SG	0.43	2.53	18	1
1:A:206:MET:O	1:A:209:VAL:HG22	0.43	2.13	6	1
1:A:216:THR:HG23	1:A:217:GLN:N	0.43	2.29	8	1
1:A:179:CYS:SG	1:A:214:CYS:SG	0.42	3.18	16	1
1:A:168:GLN:HG2	1:A:169:TYR:CD2	0.42	2.49	17	1
1:A:125:LEU:HD11	1:A:162:TYR:CD2	0.41	2.50	7	1
1:A:133:ALA:HB1	1:A:159:ASN:ND2	0.41	2.30	10	1
1:A:168:GLN:HG2	1:A:169:TYR:CD1	0.41	2.50	3	2
1:A:145:TRP:CE2	1:A:146:GLU:HG3	0.41	2.50	8	1

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:183:THR:HG21	1:A:210:VAL:HG21	0.41	1.93	9	1
1:A:134:MET:HE3	1:A:216:THR:CG2	0.41	2.46	16	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	102/113 (90%)	88±2 (86±2%)	13±2 (12±2%)	1±1 (1±1%)	15	61
All	All	2040/2260 (90%)	1756 (86%)	255 (12%)	29 (1%)	15	61

All 9 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	167	SER	7
1	A	170	SER	6
1	A	141	PHE	5
1	A	190	THR	4
1	A	127	GLY	2
1	A	132	SER	2
1	A	126	GLY	1
1	A	166	VAL	1
1	A	187	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	95/103 (92%)	78±4 (83±4%)	17±4 (17±4%)	4	39

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1900/2060 (92%)	1569 (83%)	331 (17%)	<b>4</b> 39

All 60 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	140	HIS	20
1	A	225	TYR	20
1	A	198	PHE	19
1	A	173	ASN	18
1	A	215	VAL	18
1	A	136	ARG	16
1	A	193	THR	15
1	A	207	GLU	12
1	A	216	THR	11
1	A	130	LEU	11
1	A	199	THR	11
1	A	205	MET	10
1	A	175	PHE	9
1	A	135	SER	8
1	A	132	SER	8
1	A	143	ASN	7
1	A	174	ASN	7
1	A	223	GLN	6
1	A	139	ILE	6
1	A	154	MET	5
1	A	168	GLN	5
1	A	178	ASP	5
1	A	202	ASP	5
1	A	159	ASN	5
1	A	208	ARG	4
1	A	125	LEU	4
1	A	172	GLN	4
1	A	192	THR	4
1	A	164	ARG	4
1	A	166	VAL	3
1	A	144	ASP	3
1	A	149	TYR	3
1	A	190	THR	3
1	A	200	GLU	3
1	A	179	CYS	3
1	A	201	THR	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	128	TYR	2
1	A	189	VAL	2
1	A	134	MET	2
1	A	170	SER	2
1	A	171	ASN	2
1	A	185	LYS	2
1	A	214	CYS	2
1	A	147	ASP	2
1	A	155	TYR	2
1	A	167	SER	2
1	A	181	ASN	1
1	A	210	VAL	1
1	A	226	TYR	1
1	A	141	PHE	1
1	A	221	GLU	1
1	A	220	LYS	1
1	A	206	MET	1
1	A	222	SER	1
1	A	129	MET	1
1	A	153	ASN	1
1	A	197	ASN	1
1	A	196	GLU	1
1	A	151	ARG	1
1	A	161	VAL	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 [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 82% for the well-defined parts and 81% 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	1291
Number of shifts mapped to atoms	1291
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	$0.27 \pm 0.17$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	105	$0.57 \pm 0.09$	Should be applied
$^{13}\text{C}'$	0	—	None (insufficient data)
$^{15}\text{N}$	110	$0.23 \pm 0.20$	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 82%, i.e. 1075 atoms were assigned a chemical shift out of a possible 1310. 11 out of 11 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	397/504 (79%)	196/201 (98%)	102/204 (50%)	99/99 (100%)
Sidechain	568/658 (86%)	357/389 (92%)	189/231 (82%)	22/38 (58%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	110/148 (74%)	71/77 (92%)	38/67 (57%)	1/4 (25%)
Overall	1075/1310 (82%)	624/667 (94%)	329/502 (66%)	122/141 (87%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 81%, i.e. 1162 atoms were assigned a chemical shift out of a possible 1426. 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	441/559 (79%)	218/223 (98%)	113/226 (50%)	110/110 (100%)
Sidechain	611/719 (85%)	384/425 (90%)	204/250 (82%)	23/44 (52%)
Aromatic	110/148 (74%)	71/77 (92%)	38/67 (57%)	1/4 (25%)
Overall	1162/1426 (81%)	673/725 (93%)	355/543 (65%)	134/158 (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:

