



# Full wwPDB NMR Structure Validation Report ⓘ

Feb 12, 2017 – 11:34 pm GMT

PDB ID : 2L1K  
Title : Mouse prion protein (121-231) containing the substitutions Y169A, Y225A, and Y226A  
Authors : Christen, B.; Damberger, F.F.; Perez, D.R.; Hornemann, S.; Wuthrich, K.  
Deposited on : 2010-07-28

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

<http://wwpdb.org/validation/2016/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:

Cyrange : Kirchner and Güntert (2011)  
NmrClust : Kelley et al. (1996)  
MolProbity : 4.02b-467  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
RCI : v\_1n\_11\_5\_13\_A (Berjanski et al., 2005)  
PANAV : Wang et al. (2010)  
ShiftChecker : trunk28760  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

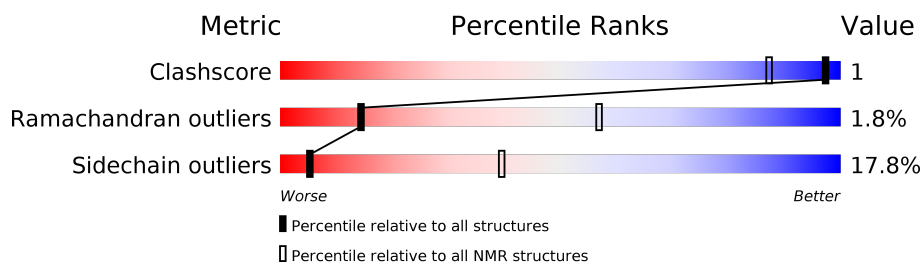
# 1 Overall quality at a glance

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	125131	11601
Ramachandran outliers	121729	10391
Sidechain outliers	121581	10367

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	 78% 14% 8%

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 16 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:229 (105)	0.39	16

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 3 clusters and 4 single-model clusters were found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1776 atoms, of which 860 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Major prion protein.

Mol	Chain	Residues	Atoms						Trace
1	A	114	Total	C	H	N	O	S	0
			1776	562	860	165	180	9	

There are 5 discrepancies between the modelled and reference sequences:

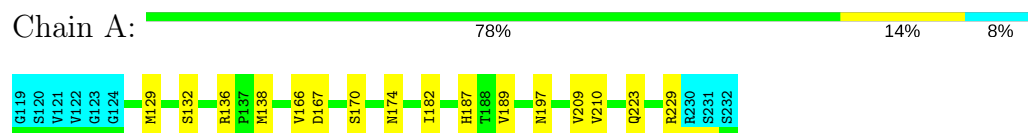
Chain	Residue	Modelled	Actual	Comment	Reference
A	119	GLY	-	EXPRESSION TAG	UNP Q4FJQ7
A	120	SER	-	EXPRESSION TAG	UNP Q4FJQ7
A	169	ALA	TYR	ENGINEERED MUTATION	UNP Q4FJQ7
A	225	ALA	TYR	ENGINEERED MUTATION	UNP Q4FJQ7
A	226	ALA	TYR	ENGINEERED MUTATION	UNP Q4FJQ7

## 4 Residue-property plots [i](#)

### 4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA 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

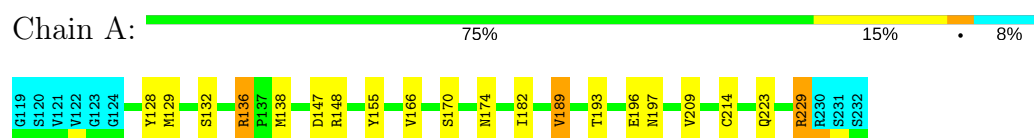


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

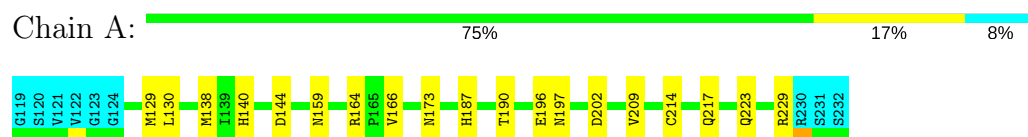
#### 4.2.1 Score per residue for model 1

- Molecule 1: Major prion protein



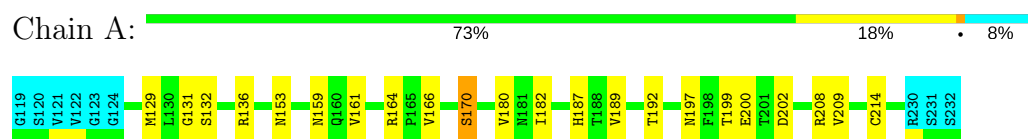
#### 4.2.2 Score per residue for model 2

- Molecule 1: Major prion protein



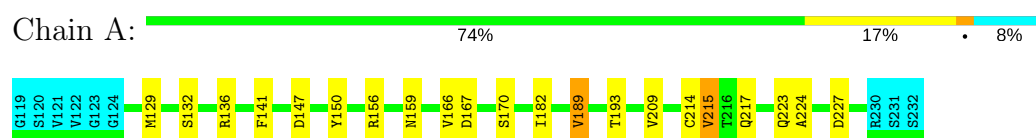
### 4.2.3 Score per residue for model 3

- Molecule 1: Major prion protein



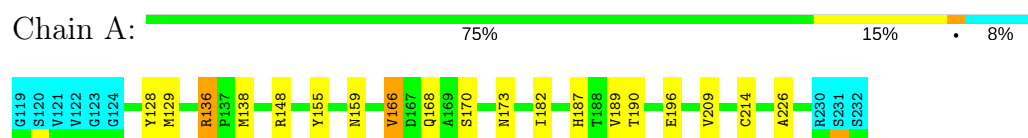
### 4.2.4 Score per residue for model 4

- Molecule 1: Major prion protein



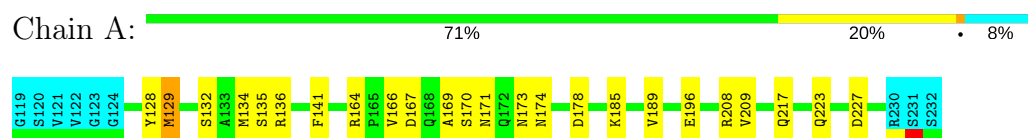
### 4.2.5 Score per residue for model 5

- Molecule 1: Major prion protein



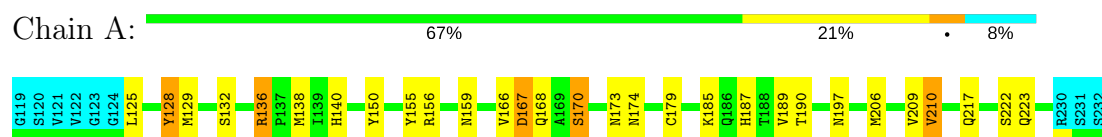
### 4.2.6 Score per residue for model 6

- Molecule 1: Major prion protein



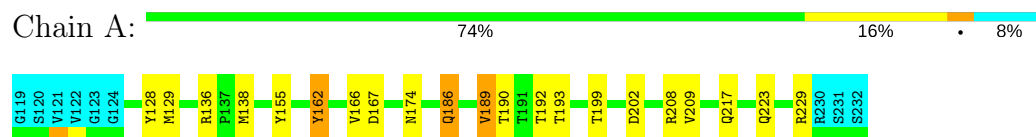
### 4.2.7 Score per residue for model 7

- Molecule 1: Major prion protein



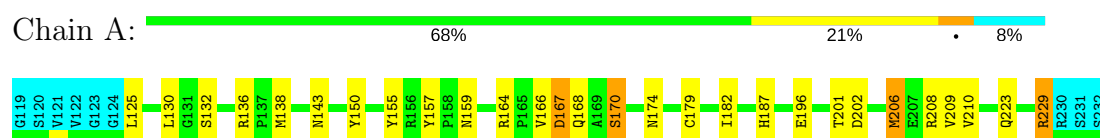
### 4.2.8 Score per residue for model 8

- Molecule 1: Major prion protein



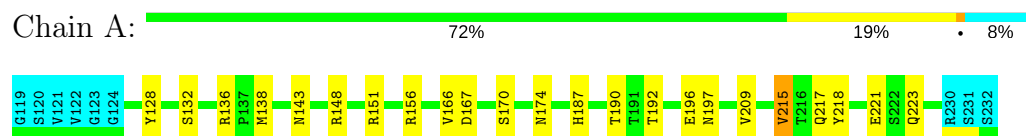
### 4.2.9 Score per residue for model 9

- Molecule 1: Major prion protein



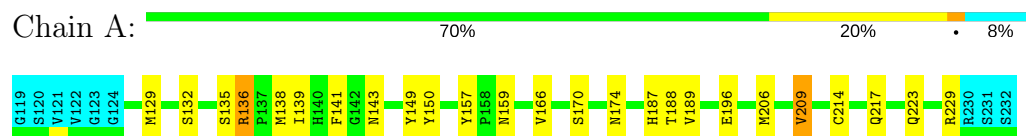
### 4.2.10 Score per residue for model 10

- Molecule 1: Major prion protein



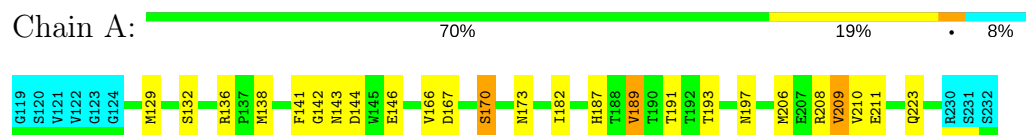
### 4.2.11 Score per residue for model 11

- Molecule 1: Major prion protein



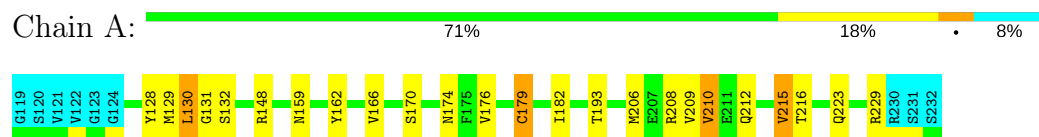
### 4.2.12 Score per residue for model 12

- Molecule 1: Major prion protein



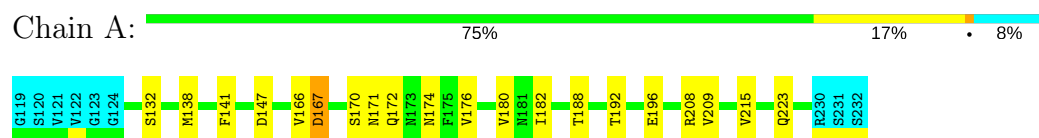
### 4.2.13 Score per residue for model 13

- Molecule 1: Major prion protein



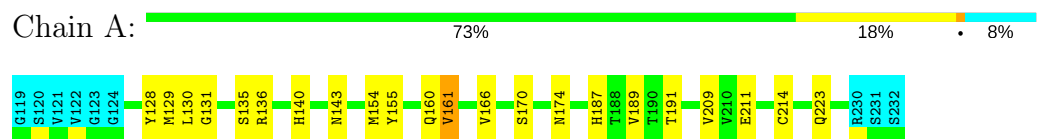
### 4.2.14 Score per residue for model 14

- Molecule 1: Major prion protein



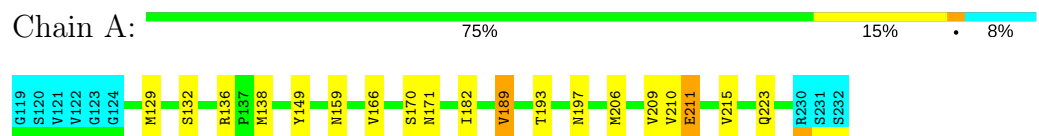
### 4.2.15 Score per residue for model 15

- Molecule 1: Major prion protein



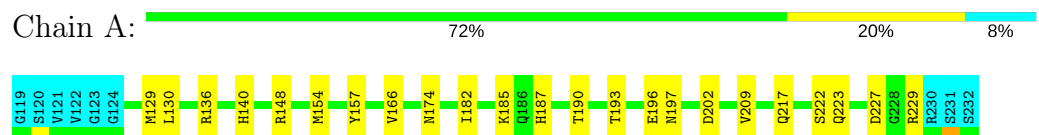
### 4.2.16 Score per residue for model 16 (medoid)

- Molecule 1: Major prion protein



### 4.2.17 Score per residue for model 17

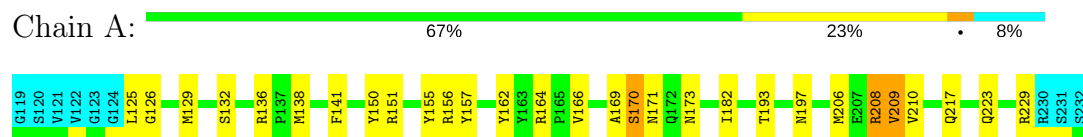
- Molecule 1: Major prion protein





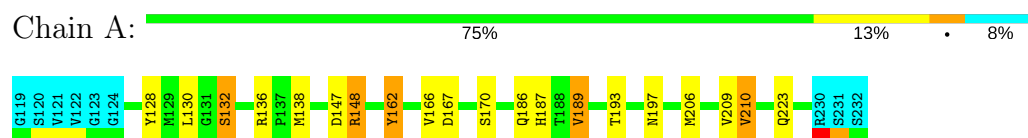
### 4.2.18 Score per residue for model 18

- Molecule 1: Major prion protein



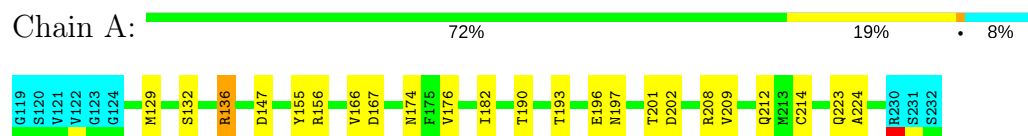
### 4.2.19 Score per residue for model 19

- 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: *TORSION ANGLE DYNAMICS*.

Of the 80 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
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)	BMRB entry 17087
Number of chemical shift lists	1
Total number of shifts	1282
Number of shifts mapped to atoms	1282
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%

No validations of the models with respect to experimental NMR restraints is performed at this time.

## 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	0.64±0.01	0±0/880 (0.0±0.0%)	1.10±0.02	2±2/1190 (0.2±0.1%)
All	All	0.64	0/17600 (0.0%)	1.10	43/23800 (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	2.4±1.5
All	All	0	47

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	209	VAL	CA-CB-CG2	9.70	125.45	110.90	11	1
1	A	229	ARG	NE-CZ-NH2	-7.15	116.73	120.30	2	4
1	A	210	VAL	CA-CB-CG1	6.85	121.17	110.90	7	2
1	A	162	TYR	CB-CG-CD1	-6.70	116.98	121.00	18	1
1	A	189	VAL	CA-CB-CG1	6.51	120.66	110.90	11	7
1	A	148	ARG	NE-CZ-NH2	-6.34	117.13	120.30	19	3
1	A	136	ARG	NE-CZ-NH1	6.32	123.46	120.30	1	1
1	A	176	VAL	CA-CB-CG1	6.20	120.20	110.90	20	2
1	A	208	ARG	NE-CZ-NH2	-5.77	117.42	120.30	18	2
1	A	136	ARG	NE-CZ-NH2	-5.70	117.45	120.30	1	1
1	A	156	ARG	NE-CZ-NH2	-5.70	117.45	120.30	18	2
1	A	170	SER	CB-CA-C	5.60	120.75	110.10	14	1
1	A	210	VAL	CA-CB-CG2	5.59	119.28	110.90	7	1
1	A	215	VAL	CA-CB-CG1	5.56	119.24	110.90	13	2
1	A	156	ARG	NE-CZ-NH1	5.51	123.06	120.30	18	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	209	VAL	CA-CB-CG1	5.47	119.10	110.90	18	2
1	A	192	THR	CA-CB-CG2	5.35	119.89	112.40	14	1
1	A	157	TYR	CB-CG-CD2	-5.33	117.81	121.00	9	1
1	A	191	THR	CA-CB-CG2	5.24	119.74	112.40	12	2
1	A	215	VAL	CG1-CB-CG2	-5.18	102.61	110.90	10	1
1	A	164	ARG	NE-CZ-NH2	-5.13	117.73	120.30	3	1
1	A	130	LEU	CB-CG-CD2	5.07	119.61	111.00	13	1
1	A	190	THR	CA-CB-CG2	5.06	119.48	112.40	5	1
1	A	208	ARG	NE-CZ-NH1	5.04	122.82	120.30	18	1
1	A	164	ARG	NE-CZ-NH1	5.01	122.80	120.30	18	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	150	TYR	Sidechain	5
1	A	136	ARG	Sidechain	5
1	A	208	ARG	Sidechain	4
1	A	229	ARG	Sidechain	3
1	A	164	ARG	Sidechain	3
1	A	162	TYR	Sidechain	3
1	A	148	ARG	Sidechain	3
1	A	196	GLU	Peptide	2
1	A	169	ALA	Peptide	2
1	A	149	TYR	Sidechain	2
1	A	157	TYR	Sidechain	2
1	A	151	ARG	Sidechain	1
1	A	226	ALA	Peptide	1
1	A	156	ARG	Sidechain	1
1	A	218	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	861	805	805	1±1
All	All	17220	16100	16100	23

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:206:MET:HA	1:A:209:VAL:HG22	0.72	1.62	11	1
1:A:206:MET:O	1:A:210:VAL:HG12	0.63	1.94	9	6
1:A:130:LEU:HD21	1:A:160:GLN:HB3	0.58	1.76	15	1
1:A:206:MET:O	1:A:210:VAL:HG13	0.52	2.04	7	1
1:A:125:LEU:HB3	1:A:128:TYR:CD1	0.48	2.44	7	1
1:A:224:ALA:HA	1:A:227:ASP:OD1	0.47	2.10	4	1
1:A:129:MET:CE	1:A:129:MET:HA	0.46	2.41	7	1
1:A:211:GLU:O	1:A:215:VAL:HG23	0.45	2.12	16	1
1:A:134:MET:CE	1:A:217:GLN:HG2	0.43	2.43	6	1
1:A:206:MET:HG3	1:A:210:VAL:HG12	0.43	1.89	7	1
1:A:172:GLN:HE21	1:A:215:VAL:HG13	0.42	1.74	14	1
1:A:129:MET:HA	1:A:129:MET:CE	0.42	2.45	6	1
1:A:176:VAL:HA	1:A:179:CYS:SG	0.41	2.54	13	1
1:A:130:LEU:HD22	1:A:161:VAL:O	0.41	2.16	15	1
1:A:130:LEU:HD13	1:A:162:TYR:CZ	0.41	2.51	19	1
1:A:130:LEU:HD21	1:A:160:GLN:CB	0.41	2.45	15	1
1:A:154:MET:HA	1:A:157:TYR:CE2	0.41	2.51	17	1
1:A:192:THR:HG23	1:A:197:ASN:HA	0.40	1.93	10	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	105/114 (92%)	89±3 (84±3%)	14±3 (14±3%)	2±1 (2±1%)	14	57
All	All	2100/2280 (92%)	1774 (84%)	289 (14%)	37 (2%)	14	57

All 10 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	170	SER	15
1	A	141	PHE	6
1	A	167	ASP	5
1	A	131	GLY	3
1	A	132	SER	3
1	A	126	GLY	1
1	A	142	GLY	1
1	A	166	VAL	1
1	A	185	LYS	1
1	A	171	ASN	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	94/100 (94%)	77±3 (82±3%)	17±3 (18±3%)	5	39
All	All	1880/2000 (94%)	1545 (82%)	335 (18%)	5	39

All 61 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	166	VAL	20
1	A	209	VAL	19
1	A	223	GLN	18
1	A	136	ARG	16
1	A	129	MET	15
1	A	132	SER	13
1	A	138	MET	13
1	A	174	ASN	12
1	A	182	ILE	12
1	A	189	VAL	11
1	A	187	HIS	11
1	A	197	ASN	10
1	A	193	THR	10
1	A	159	ASN	9
1	A	217	GLN	8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	196	GLU	8
1	A	214	CYS	8
1	A	167	ASP	8
1	A	155	TYR	8
1	A	173	ASN	6
1	A	202	ASP	6
1	A	190	THR	6
1	A	143	ASN	5
1	A	170	SER	5
1	A	147	ASP	5
1	A	208	ARG	4
1	A	130	LEU	4
1	A	140	HIS	4
1	A	135	SER	3
1	A	168	GLN	3
1	A	179	CYS	3
1	A	211	GLU	3
1	A	229	ARG	3
1	A	171	ASN	3
1	A	215	VAL	3
1	A	212	GLN	2
1	A	125	LEU	2
1	A	144	ASP	2
1	A	227	ASP	2
1	A	180	VAL	2
1	A	199	THR	2
1	A	185	LYS	2
1	A	201	THR	2
1	A	222	SER	2
1	A	192	THR	2
1	A	161	VAL	2
1	A	188	THR	2
1	A	156	ARG	2
1	A	186	GLN	2
1	A	154	MET	1
1	A	216	THR	1
1	A	139	ILE	1
1	A	221	GLU	1
1	A	206	MET	1
1	A	146	GLU	1
1	A	153	ASN	1
1	A	148	ARG	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	200	GLU	1
1	A	210	VAL	1
1	A	151	ARG	1
1	A	178	ASP	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 carbohydrates 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 82% for the entire structure.

### 7.1 Chemical shift list 1

File name: BMRB entry 17087

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	1282
Number of shifts mapped to atoms	1282
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
$^{13}\text{C}_\alpha$	114	$-0.33 \pm 0.21$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	105	$0.32 \pm 0.16$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	0	—	None (insufficient data)
$^{15}\text{N}$	110	$-0.11 \pm 0.15$	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. 1080 atoms were assigned a chemical shift out of a possible 1321. 11 out of 11 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	409/519 (79%)	202/207 (98%)	105/210 (50%)	102/102 (100%)
Sidechain	579/678 (85%)	363/399 (91%)	193/238 (81%)	23/41 (56%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	92/124 (74%)	58/65 (89%)	33/55 (60%)	1/4 (25%)
Overall	1080/1321 (82%)	623/671 (93%)	331/503 (66%)	126/147 (86%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 82%, i.e. 1146 atoms were assigned a chemical shift out of a possible 1405. 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	443/564 (79%)	219/225 (97%)	114/228 (50%)	110/111 (99%)
Sidechain	611/717 (85%)	382/422 (91%)	205/251 (82%)	24/44 (55%)
Aromatic	92/124 (74%)	58/65 (89%)	33/55 (60%)	1/4 (25%)
Overall	1146/1405 (82%)	659/712 (93%)	352/534 (66%)	135/159 (85%)

#### 7.1.4 Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

#### 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.

Random coil index (RCI) for chain A:

