



# Full wwPDB NMR Structure Validation Report ⓘ

Feb 12, 2017 – 08:44 pm GMT

PDB ID : 1XYX  
Title : mouse prion protein fragment 121-231  
Authors : Gossert, A.D.; Bonjour, S.; Lysek, D.A.; Fiorito, F.; Wuthrich, K.  
Deposited on : 2004-11-11

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

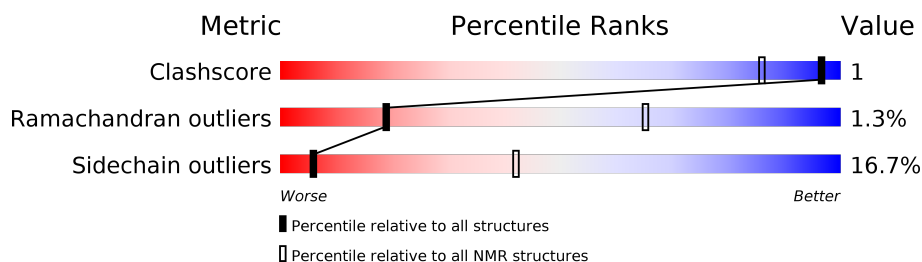
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment 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	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	112	<div> <div style="width: 69%; background-color: green;"></div> <div style="width: 12%; background-color: yellow;"></div> <div style="width: 20%; background-color: cyan;"></div> </div> <div>69% 12% 20%</div>

## 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, 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:128-A:166, A:172-A:222 (90)	0.31	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 3 clusters and 2 single-model clusters were found.

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

### 3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 1791 atoms, of which 864 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Major prion protein.

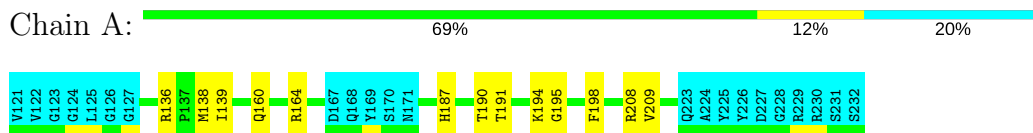
Mol	Chain	Residues	Atoms						Trace
1	A	112	Total	C	H	N	O	S	0
			1791	575	864	163	180	9	

## 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

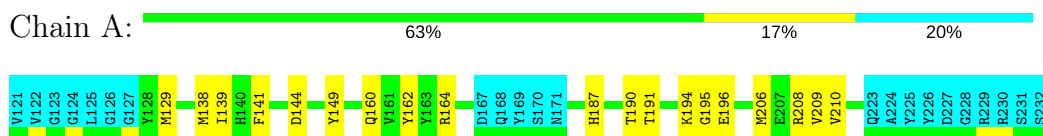


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

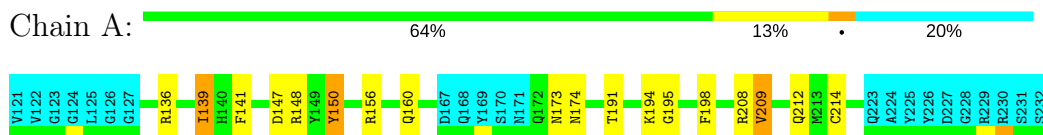
#### 4.2.1 Score per residue for model 1

- Molecule 1: Major prion protein



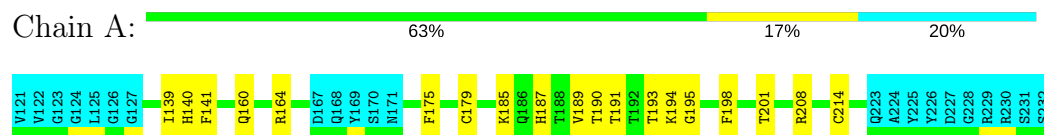
#### 4.2.2 Score per residue for model 2

- Molecule 1: Major prion protein



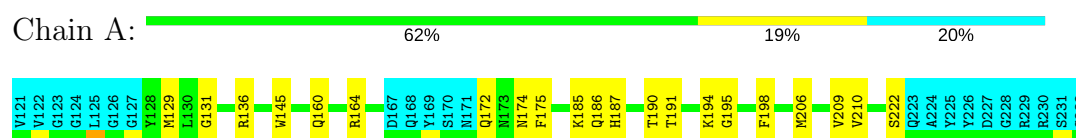
### 4.2.3 Score per residue for model 3

- Molecule 1: Major prion protein



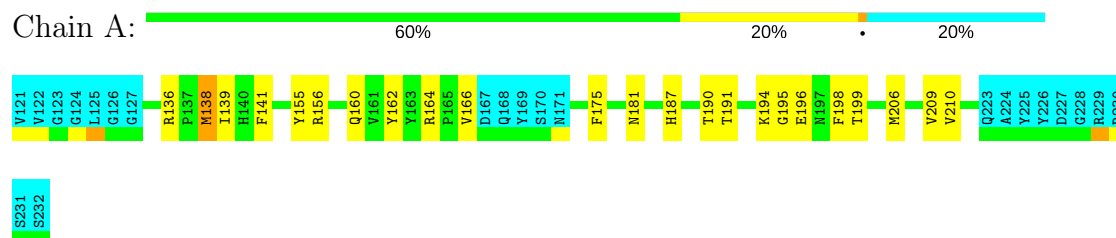
### 4.2.4 Score per residue for model 4

- Molecule 1: Major prion protein



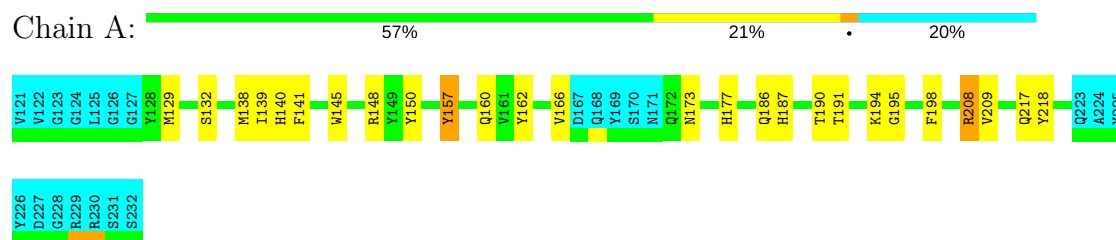
### 4.2.5 Score per residue for model 5

- Molecule 1: Major prion protein



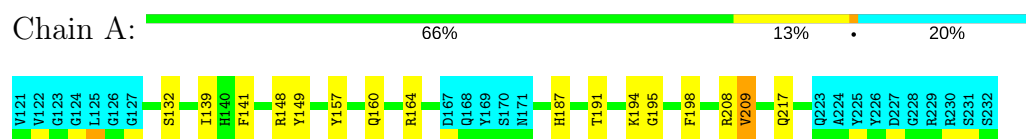
### 4.2.6 Score per residue for model 6

- Molecule 1: Major prion protein



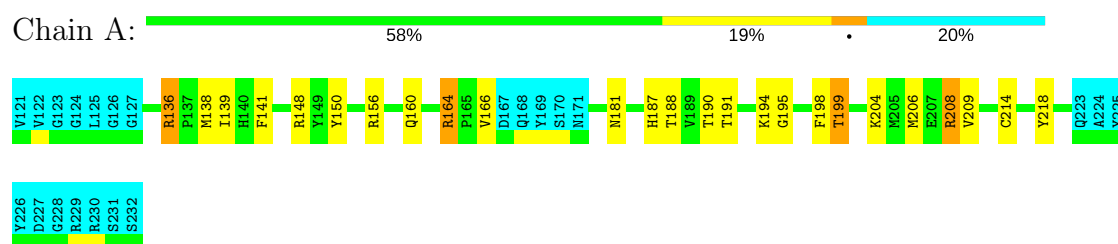
### 4.2.7 Score per residue for model 7

- Molecule 1: Major prion protein



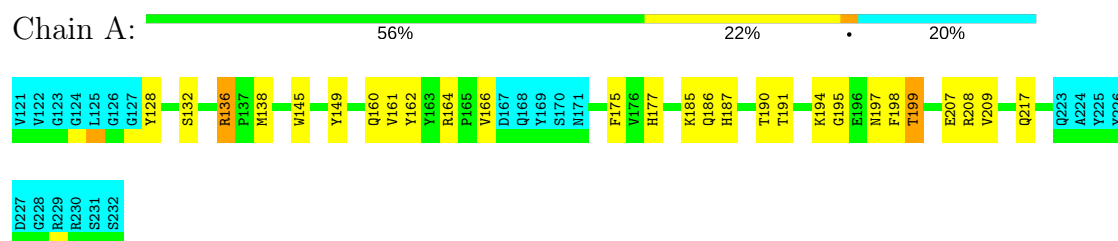
### 4.2.8 Score per residue for model 8

- Molecule 1: Major prion protein



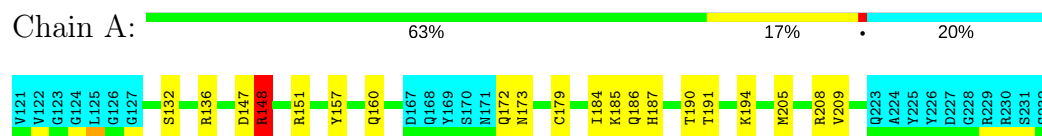
### 4.2.9 Score per residue for model 9

- Molecule 1: Major prion protein



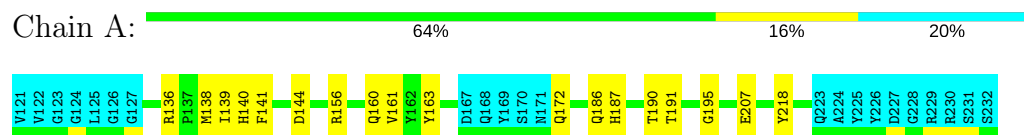
### 4.2.10 Score per residue for model 10

- Molecule 1: Major prion protein



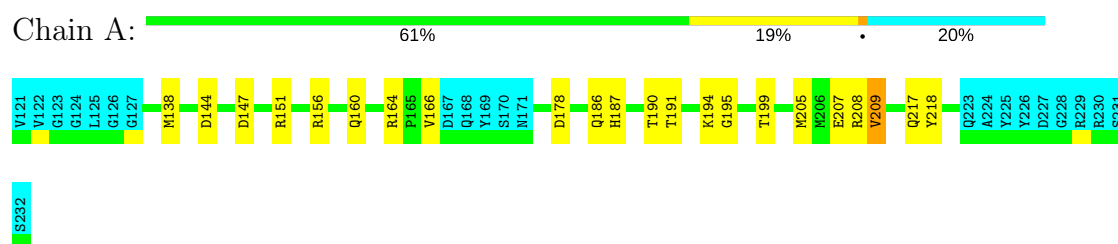
### 4.2.11 Score per residue for model 11

- Molecule 1: Major prion protein



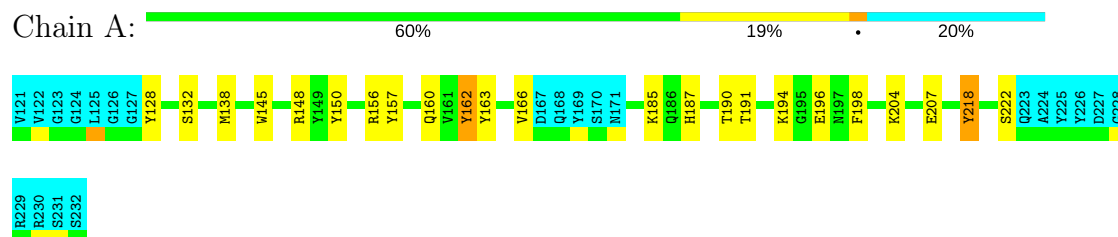
### 4.2.12 Score per residue for model 12 (medoid)

- Molecule 1: Major prion protein



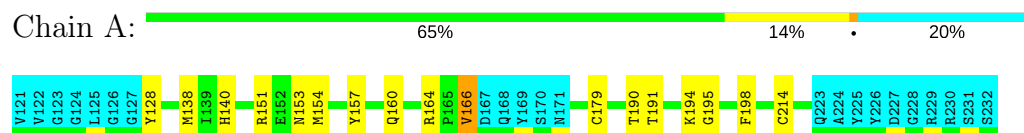
### 4.2.13 Score per residue for model 13

- Molecule 1: Major prion protein



### 4.2.14 Score per residue for model 14

- Molecule 1: Major prion protein





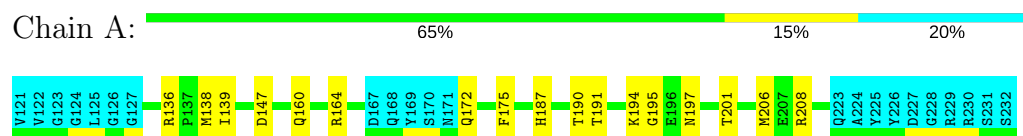
### 4.2.15 Score per residue for model 15

- Molecule 1: Major prion protein



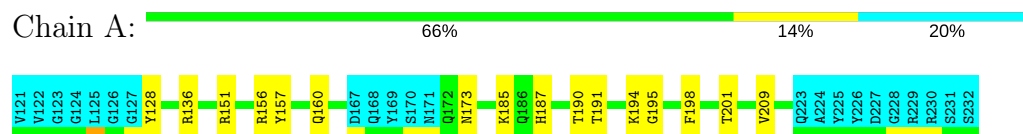
### 4.2.16 Score per residue for model 16

- Molecule 1: Major prion protein



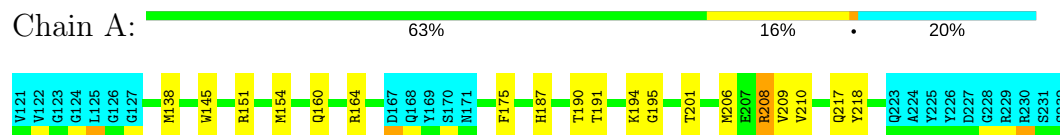
### 4.2.17 Score per residue for model 17

- Molecule 1: Major prion protein



### 4.2.18 Score per residue for model 18

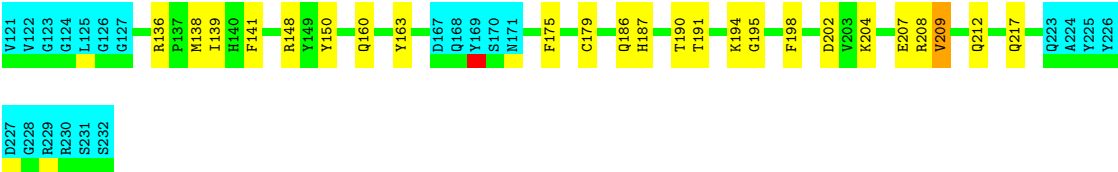
- Molecule 1: Major prion protein



### 4.2.19 Score per residue for model 19

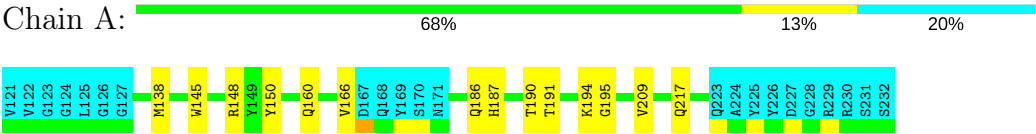
- Molecule 1: Major prion protein





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 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
CYANA	structure solution	1.0
CANDID	refinement	1.0

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

## 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	0.67±0.01	0±0/781 (0.0±0.0%)	1.09±0.04	2±1/1057 (0.1±0.1%)
All	All	0.67	0/15620 (0.0%)	1.09	30/21140 (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	1.8±1.4
All	All	0	36

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	208	ARG	NE-CZ-NH2	-7.65	116.47	120.30	18	2
1	A	148	ARG	NE-CZ-NH2	-7.43	116.59	120.30	6	5
1	A	157	TYR	CB-CG-CD2	-7.02	116.79	121.00	17	7
1	A	156	ARG	NE-CZ-NH2	-6.72	116.94	120.30	5	2
1	A	148	ARG	NE-CZ-NH1	6.56	123.58	120.30	20	1
1	A	162	TYR	CB-CG-CD1	-5.88	117.47	121.00	1	3
1	A	218	TYR	CB-CG-CD2	-5.47	117.72	121.00	6	1
1	A	136	ARG	NE-CZ-NH2	-5.47	117.57	120.30	5	1
1	A	199	THR	CA-CB-CG2	-5.42	104.82	112.40	8	2
1	A	149	TYR	CB-CG-CD2	-5.27	117.84	121.00	7	1
1	A	188	THR	CA-CB-CG2	5.21	119.69	112.40	8	1
1	A	151	ARG	NE-CZ-NH2	-5.14	117.73	120.30	12	1
1	A	150	TYR	CB-CG-CD1	-5.08	117.95	121.00	2	1
1	A	157	TYR	CB-CG-CD1	5.05	124.03	121.00	17	1
1	A	209	VAL	CA-CB-CG1	5.04	118.46	110.90	12	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	151	ARG	Sidechain	5
1	A	150	TYR	Sidechain	5
1	A	156	ARG	Sidechain	4
1	A	136	ARG	Sidechain	4
1	A	218	TYR	Sidechain	3
1	A	164	ARG	Sidechain	2
1	A	208	ARG	Sidechain	2
1	A	163	TYR	Sidechain	2
1	A	148	ARG	Sidechain	2
1	A	149	TYR	Sidechain	2
1	A	155	TYR	Sidechain	1
1	A	139	ILE	Peptide	1
1	A	162	TYR	Sidechain	1
1	A	175	PHE	Sidechain	1
1	A	157	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	762	717	719	1±1
All	All	15240	14340	14380	27

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:O	1:A:210:VAL:HG23	0.56	2.00	4	4
1:A:139:ILE:HD13	1:A:212:GLN:HG3	0.52	1.80	2	2
1:A:211:GLU:O	1:A:215:VAL:HG23	0.51	2.06	15	1
1:A:166:VAL:HG12	1:A:218:TYR:CE1	0.51	2.41	12	2
1:A:139:ILE:HG21	1:A:141:PHE:CZ	0.50	2.41	7	8

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:139:ILE:HD11	1:A:209:VAL:HA	0.48	1.84	7	3
1:A:186:GLN:O	1:A:190:THR:HG22	0.45	2.11	11	1
1:A:139:ILE:HG21	1:A:141:PHE:CE1	0.44	2.47	11	2
1:A:179:CYS:SG	1:A:214:CYS:HB3	0.43	2.53	14	1
1:A:138:MET:CE	1:A:138:MET:H	0.42	2.28	5	1
1:A:189:VAL:O	1:A:193:THR:HG22	0.41	2.15	3	1
1:A:141:PHE:CE2	1:A:208:ARG:HG3	0.41	2.51	6	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	90/112 (80%)	82±2 (91±2%)	7±2 (8±2%)	1±1 (1±1%)	19	65
All	All	1800/2240 (80%)	1641 (91%)	136 (8%)	23 (1%)	19	65

All 5 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	195	GLY	18
1	A	129	MET	2
1	A	166	VAL	1
1	A	131	GLY	1
1	A	196	GLU	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	86/102 (84%)	72±3 (83±4%)	14±3 (17±4%)	6	42
All	All	1720/2040 (84%)	1432 (83%)	288 (17%)	6	42

All 50 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	160	GLN	20
1	A	191	THR	20
1	A	194	LYS	19
1	A	187	HIS	18
1	A	190	THR	17
1	A	209	VAL	15
1	A	138	MET	14
1	A	198	PHE	13
1	A	208	ARG	11
1	A	164	ARG	11
1	A	136	ARG	8
1	A	217	GLN	7
1	A	175	PHE	7
1	A	186	GLN	7
1	A	145	TRP	6
1	A	185	LYS	6
1	A	166	VAL	6
1	A	132	SER	5
1	A	173	ASN	5
1	A	147	ASP	5
1	A	207	GLU	5
1	A	199	THR	4
1	A	140	HIS	4
1	A	172	GLN	4
1	A	201	THR	4
1	A	174	ASN	3
1	A	179	CYS	3
1	A	214	CYS	3
1	A	144	ASP	3
1	A	204	LYS	3
1	A	196	GLU	3
1	A	206	MET	2
1	A	154	MET	2
1	A	148	ARG	2
1	A	161	VAL	2
1	A	205	MET	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	139	ILE	2
1	A	197	ASN	2
1	A	222	SER	2
1	A	181	ASN	2
1	A	177	HIS	2
1	A	156	ARG	1
1	A	153	ASN	1
1	A	218	TYR	1
1	A	184	ILE	1
1	A	129	MET	1
1	A	178	ASP	1
1	A	202	ASP	1
1	A	163	TYR	1
1	A	150	TYR	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

No chemical shift data were provided