



# Full wwPDB NMR Structure Validation Report ⓘ

Apr 26, 2016 – 10:59 PM BST

PDB ID : 2K56  
Title : Bank Vole Prion Protein (121-231)  
Authors : Christen, B.; Hornemann, S.; Wuthrich, K.  
Deposited on : 2008-06-26

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  
Mogul : unknown  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
RCI : v\_1n\_11\_5\_13\_A (Berjanski et al., 2005)  
PANAV : Wang et al. (2010)  
ShiftChecker : rb-20027457  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027457

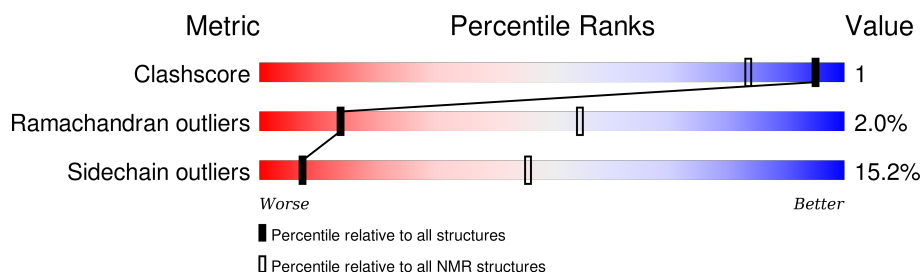
# 1 Overall quality at a glance

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	 76% 14% 10%

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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)	0.52	3

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 1 single-model cluster was found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1784 atoms, of which 859 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
			1784	571	859	163	182	9	

There are 2 discrepancies between the modelled and reference sequences:

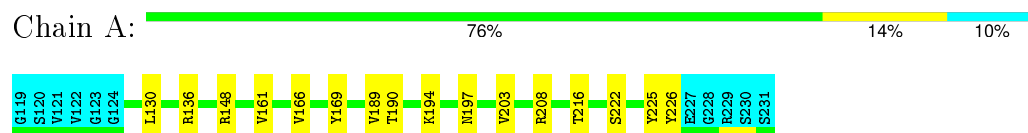
Chain	Residue	Modelled	Actual	Comment	Reference
A	119	GLY	-	EXPRESSION TAG	UNP Q8VHV5
A	120	SER	-	EXPRESSION TAG	UNP Q8VHV5

## 4 Residue-property plots

### 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

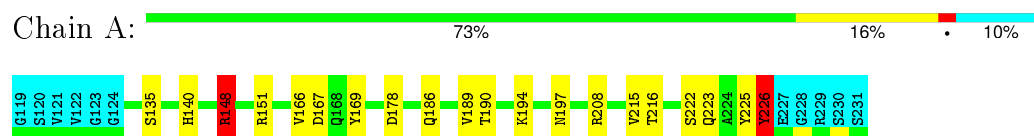


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

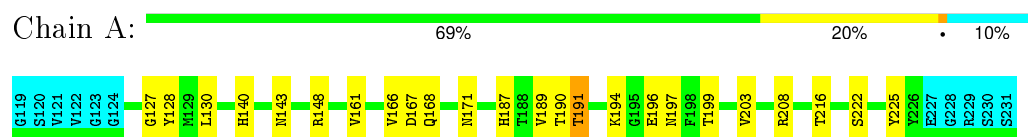
#### 4.2.1 Score per residue for model 1

- Molecule 1: Major prion protein



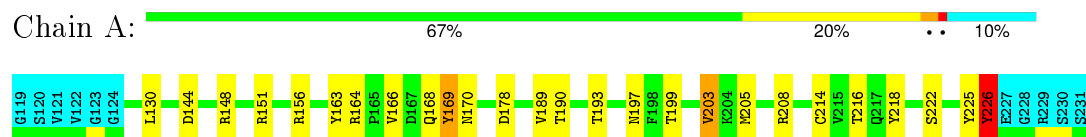
#### 4.2.2 Score per residue for model 2

- Molecule 1: Major prion protein



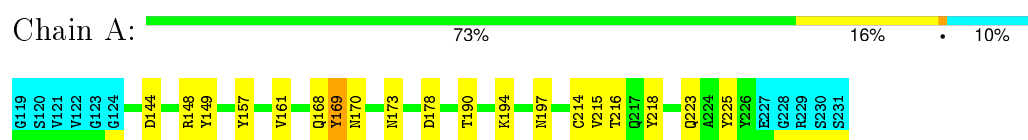
### 4.2.3 Score per residue for model 3 (medoid)

- Molecule 1: Major prion protein



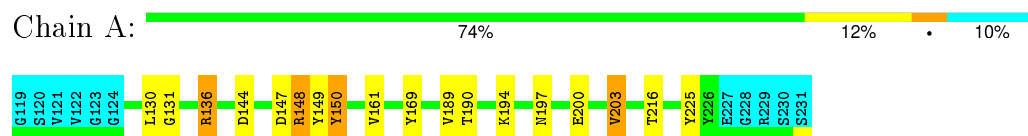
### 4.2.4 Score per residue for model 4

- Molecule 1: Major prion protein



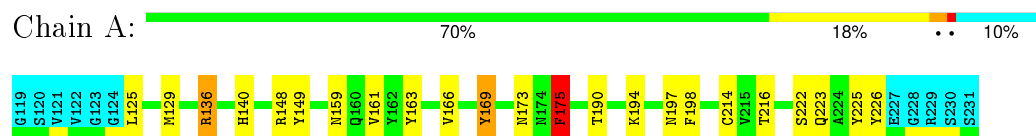
### 4.2.5 Score per residue for model 5

- Molecule 1: Major prion protein



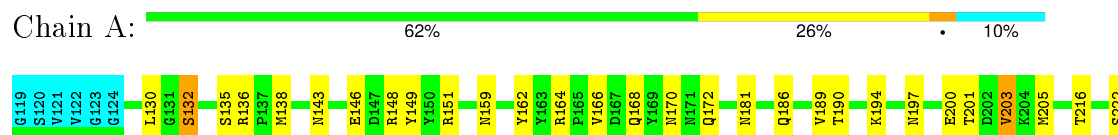
### 4.2.6 Score per residue for model 6

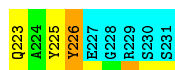
- Molecule 1: Major prion protein



### 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

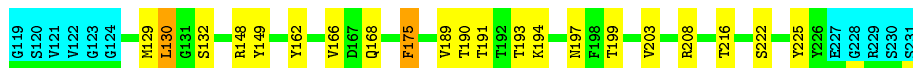
Chain A: 72% 19% 10%



#### 4.2.9 Score per residue for model 9

- Molecule 1: Major prion protein

Chain A: 72% 17% 10%



#### 4.2.10 Score per residue for model 10

- Molecule 1: Major prion protein

Chain A: 75% 14% 10%



#### 4.2.11 Score per residue for model 11

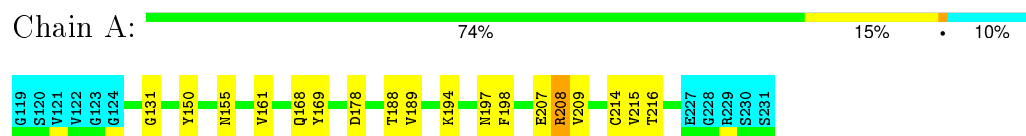
- Molecule 1: Major prion protein

Chain A: 75% 13% 10%



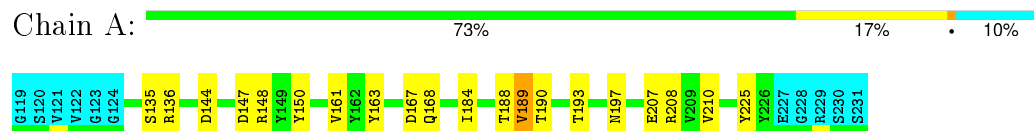
#### 4.2.12 Score per residue for model 12

- Molecule 1: Major prion protein



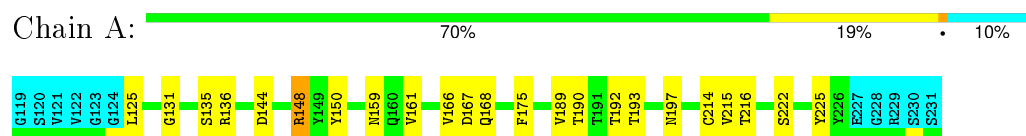
#### 4.2.13 Score per residue for model 13

- Molecule 1: Major prion protein



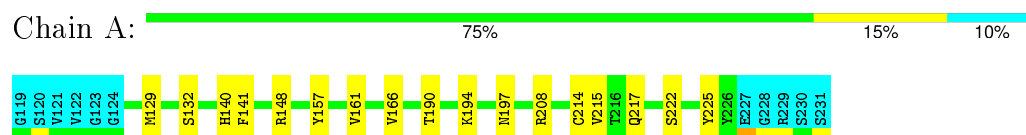
#### 4.2.14 Score per residue for model 14

- Molecule 1: Major prion protein



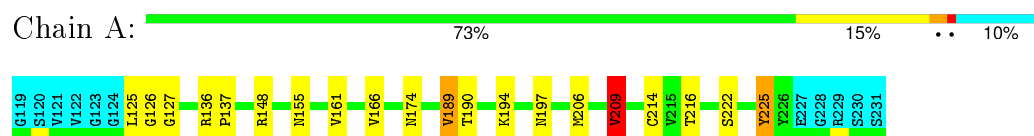
#### 4.2.15 Score per residue for model 15

- Molecule 1: Major prion protein



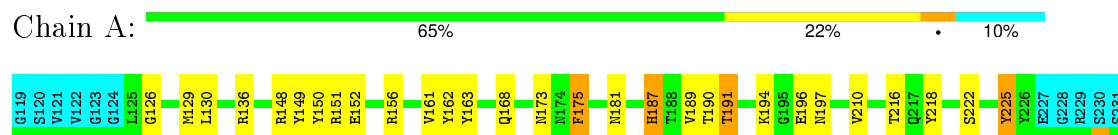
#### 4.2.16 Score per residue for model 16

- Molecule 1: Major prion protein



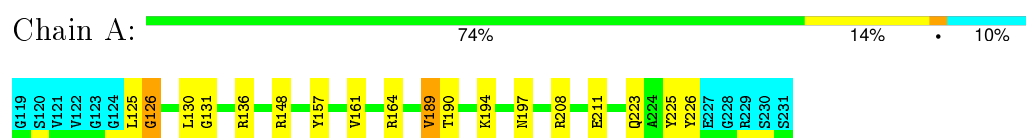
### 4.2.17 Score per residue for model 17

- Molecule 1: Major prion protein



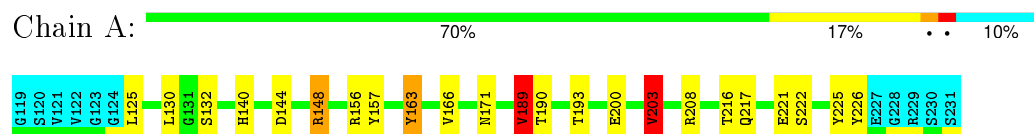
### 4.2.18 Score per residue for model 18

- Molecule 1: Major prion protein



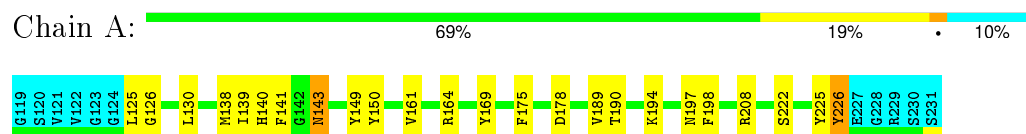
### 4.2.19 Score per residue for model 19

- 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), 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
OPAL	refinement	1.2
DYANA	structure solution	1.0.3

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 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	0.67±0.01	0±0/878 (0.0±0.0%)	1.09±0.03	2±1/1189 (0.2±0.1%)
All	All	0.67	0/17560 (0.0%)	1.09	44/23780 (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	2.0±1.3
All	All	0	41

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-CG1	8.88	124.22	110.90	16	2
1	A	157	TYR	CB-CG-CD2	-7.86	116.28	121.00	10	3
1	A	203	VAL	CA-CB-CG2	7.31	121.87	110.90	3	3
1	A	175	PHE	CB-CG-CD1	-7.02	115.89	120.80	17	3
1	A	226	TYR	CB-CG-CD1	-6.60	117.04	121.00	1	2
1	A	148	ARG	NE-CZ-NH2	-6.59	117.01	120.30	1	3
1	A	151	ARG	NE-CZ-NH2	-6.46	117.07	120.30	7	2
1	A	150	TYR	CB-CG-CD2	-6.22	117.27	121.00	13	3
1	A	164	ARG	NE-CZ-NH2	-6.12	117.24	120.30	7	1
1	A	208	ARG	NE-CZ-NH2	-6.09	117.26	120.30	12	2
1	A	210	VAL	CA-CB-CG1	5.81	119.62	110.90	17	1
1	A	225	TYR	CB-CG-CD1	-5.79	117.53	121.00	17	3
1	A	150	TYR	CB-CG-CD1	-5.79	117.53	121.00	5	1
1	A	156	ARG	NE-CZ-NH1	5.70	123.15	120.30	17	1
1	A	163	TYR	CB-CG-CD2	-5.68	117.59	121.00	19	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	169	TYR	CB-CG-CD1	-5.54	117.67	121.00	3	1
1	A	198	PHE	CB-CG-CD2	-5.54	116.92	120.80	8	2
1	A	191	THR	CA-CB-CG2	5.50	120.10	112.40	17	1
1	A	164	ARG	NE-CZ-NH1	5.38	122.99	120.30	7	1
1	A	226	TYR	CB-CG-CD2	-5.24	117.85	121.00	11	1
1	A	189	VAL	CB-CA-C	5.22	121.32	111.40	19	2
1	A	148	ARG	CD-NE-CZ	5.18	130.86	123.60	14	1
1	A	136	ARG	NE-CZ-NH2	-5.13	117.73	120.30	8	1
1	A	148	ARG	NE-CZ-NH1	5.11	122.86	120.30	14	1
1	A	157	TYR	CB-CG-CD1	5.06	124.04	121.00	10	1
1	A	163	TYR	CB-CG-CD1	-5.01	117.99	121.00	13	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	169	TYR	Sidechain	5
1	A	148	ARG	Sidechain	4
1	A	136	ARG	Sidechain	4
1	A	150	TYR	Sidechain	4
1	A	226	TYR	Sidechain	4
1	A	208	ARG	Sidechain	3
1	A	128	TYR	Sidechain	2
1	A	218	TYR	Sidechain	2
1	A	149	TYR	Sidechain	2
1	A	156	ARG	Sidechain	2
1	A	225	TYR	Sidechain	2
1	A	157	TYR	Sidechain	2
1	A	164	ARG	Sidechain	2
1	A	163	TYR	Sidechain	2
1	A	126	GLY	Peptide	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	795	795	2±2
All	All	17140	15900	15900	47

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:207:GLU:HA	1:A:210:VAL:HG12	0.64	1.69	13	1
1:A:166:VAL:HG11	1:A:222:SER:HA	0.58	1.75	3	9
1:A:199:THR:O	1:A:203:VAL:HG13	0.57	2.00	9	2
1:A:137:PRO:CD	1:A:209:VAL:HG23	0.54	2.33	11	2
1:A:175:PHE:CZ	1:A:218:TYR:CE1	0.54	2.96	17	1
1:A:130:LEU:HD13	1:A:162:TYR:CE2	0.53	2.39	7	1
1:A:146:GLU:HG2	1:A:205:MET:CE	0.52	2.35	7	1
1:A:200:GLU:HA	1:A:203:VAL:HG22	0.50	1.82	5	2
1:A:200:GLU:HA	1:A:203:VAL:HG13	0.49	1.84	19	1
1:A:125:LEU:HD22	1:A:125:LEU:H	0.49	1.68	14	2
1:A:187:HIS:O	1:A:191:THR:HG22	0.49	2.06	17	1
1:A:206:MET:HA	1:A:209:VAL:CG1	0.49	2.38	11	2
1:A:200:GLU:HA	1:A:203:VAL:CG1	0.48	2.38	19	1
1:A:189:VAL:HA	1:A:192:THR:CG2	0.48	2.39	14	1
1:A:200:GLU:CA	1:A:203:VAL:HG13	0.46	2.41	19	1
1:A:189:VAL:HA	1:A:192:THR:HG22	0.46	1.87	14	1
1:A:184:ILE:HD11	1:A:207:GLU:CB	0.44	2.42	13	1
1:A:166:VAL:HG21	1:A:222:SER:HA	0.44	1.88	1	2
1:A:199:THR:O	1:A:203:VAL:HG23	0.43	2.13	2	1
1:A:125:LEU:H	1:A:125:LEU:HD22	0.43	1.74	18	1
1:A:169:TYR:CE1	1:A:175:PHE:HA	0.43	2.48	6	1
1:A:146:GLU:HG2	1:A:205:MET:HE2	0.42	1.91	7	1
1:A:125:LEU:CD2	1:A:125:LEU:H	0.42	2.28	14	1
1:A:184:ILE:HD11	1:A:207:GLU:HB2	0.41	1.92	13	1
1:A:222:SER:HB3	1:A:226:TYR:CE1	0.41	2.49	20	1
1:A:184:ILE:CD1	1:A:207:GLU:HB2	0.41	2.46	13	1
1:A:130:LEU:HD23	1:A:162:TYR:CE2	0.41	2.51	9	1
1:A:169:TYR:CZ	1:A:175:PHE:CD2	0.41	3.09	20	1
1:A:206:MET:HA	1:A:209:VAL:HG12	0.41	1.93	16	2
1:A:191:THR:HG23	1:A:196:GLU:HG3	0.40	1.93	2	1
1:A:175:PHE:CE1	1:A:218:TYR:CZ	0.40	3.10	17	1
1:A:161:VAL:HG23	1:A:163:TYR:CE1	0.40	2.52	8	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%)	89±3 (87±3%)	11±3 (11±2%)	2±1 (2±1%)	14	55
All	All	2040/2260 (90%)	1784 (87%)	216 (11%)	40 (2%)	14	55

All 17 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	189	VAL	8
1	A	126	GLY	4
1	A	131	GLY	4
1	A	226	TYR	4
1	A	127	GLY	3
1	A	140	HIS	2
1	A	188	THR	2
1	A	171	ASN	2
1	A	125	LEU	2
1	A	141	PHE	2
1	A	132	SER	1
1	A	170	ASN	1
1	A	128	TYR	1
1	A	181	ASN	1
1	A	143	ASN	1
1	A	138	MET	1
1	A	203	VAL	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	95/102 (93%)	81±2 (85±3%)	14±2 (15±3%)	7	46
All	All	1900/2040 (93%)	1611 (85%)	289 (15%)	7	46

All 58 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	190	THR	19
1	A	197	ASN	18
1	A	216	THR	16
1	A	148	ARG	16
1	A	225	TYR	16
1	A	194	LYS	15
1	A	161	VAL	13
1	A	189	VAL	11
1	A	214	CYS	10
1	A	168	GLN	10
1	A	208	ARG	8
1	A	130	LEU	8
1	A	178	ASP	7
1	A	136	ARG	7
1	A	149	TYR	7
1	A	193	THR	6
1	A	144	ASP	6
1	A	215	VAL	5
1	A	140	HIS	5
1	A	223	GLN	5
1	A	135	SER	5
1	A	167	ASP	4
1	A	143	ASN	4
1	A	132	SER	4
1	A	129	MET	4
1	A	209	VAL	3
1	A	147	ASP	3
1	A	175	PHE	3
1	A	217	GLN	3
1	A	226	TYR	3
1	A	159	ASN	3
1	A	173	ASN	3
1	A	170	ASN	2
1	A	181	ASN	2
1	A	187	HIS	2
1	A	191	THR	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	163	TYR	2
1	A	155	ASN	2
1	A	169	TYR	2
1	A	151	ARG	2
1	A	186	GLN	2
1	A	164	ARG	2
1	A	221	GLU	2
1	A	174	ASN	2
1	A	198	PHE	2
1	A	211	GLU	1
1	A	205	MET	1
1	A	203	VAL	1
1	A	196	GLU	1
1	A	138	MET	1
1	A	172	GLN	1
1	A	152	GLU	1
1	A	162	TYR	1
1	A	139	ILE	1
1	A	207	GLU	1
1	A	222	SER	1
1	A	125	LEU	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 [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