



# Full wwPDB NMR Structure Validation Report ⓘ

Nov 16, 2017 – 02:42 PM EST

PDB ID : 2KFM  
Title : Mouse Prion Protein (121-231) with Mutations Y225A and Y226A  
Authors : Christen, B.; Hornemann, S.; Damberger, F.F.; Wuthrich, K.  
Deposited on : unknown

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 : rb-20030345  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

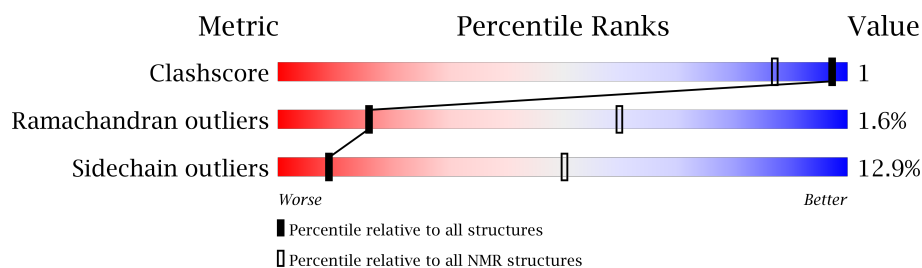
# 1 Overall quality at a glance

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	114	82% 7% 10% .

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 10 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:122-A:223 (102)	0.34	10

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 3 single-model clusters were found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1780 atoms, of which 861 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
			1780	566	861	164	180	9	

There are 4 discrepancies between the modelled and reference sequences:

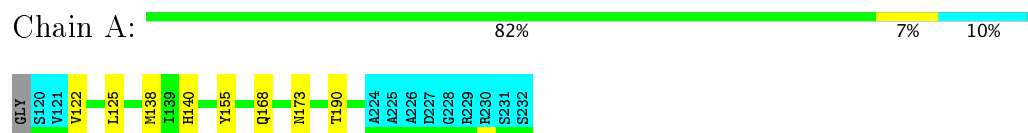
Chain	Residue	Modelled	Actual	Comment	Reference
A	119	GLY	-	EXPRESSION TAG	UNP P04925
A	120	SER	-	EXPRESSION TAG	UNP P04925
A	225	ALA	TYR	ENGINEERED	UNP P04925
A	226	ALA	TYR	ENGINEERED	UNP P04925

## 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

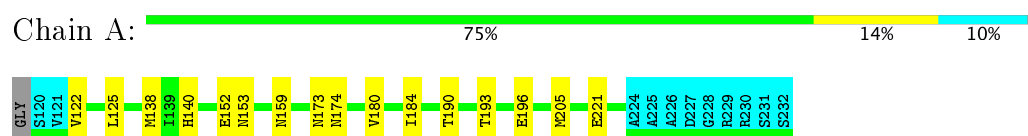


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section [4.1](#) above.

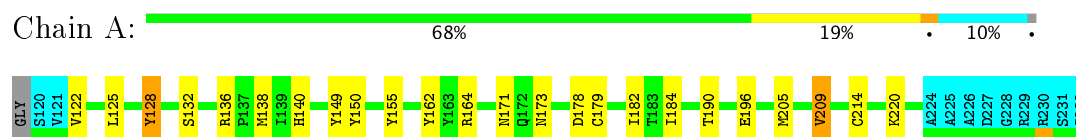
#### 4.2.1 Score per residue for model 1

- Molecule 1: Major prion protein



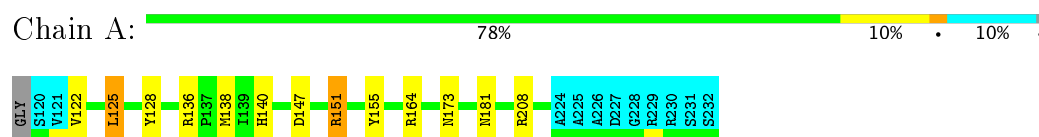
#### 4.2.2 Score per residue for model 2

- Molecule 1: Major prion protein



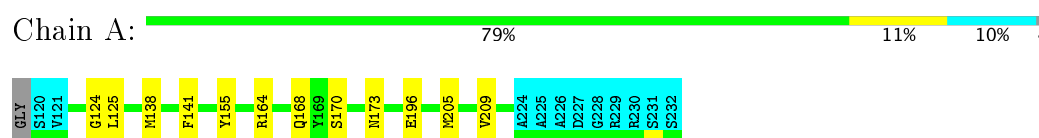
### 4.2.3 Score per residue for model 3

- Molecule 1: Major prion protein



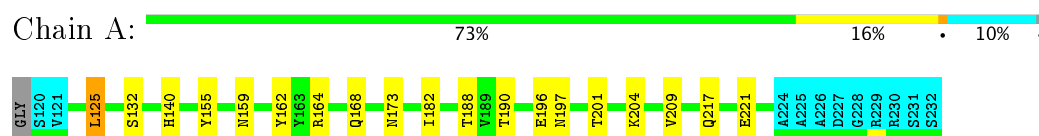
### 4.2.4 Score per residue for model 4

- Molecule 1: Major prion protein



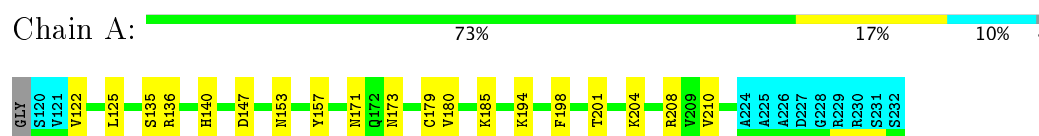
### 4.2.5 Score per residue for model 5

- Molecule 1: Major prion protein



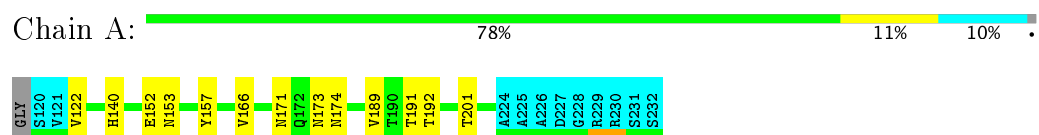
### 4.2.6 Score per residue for model 6

- Molecule 1: Major prion protein



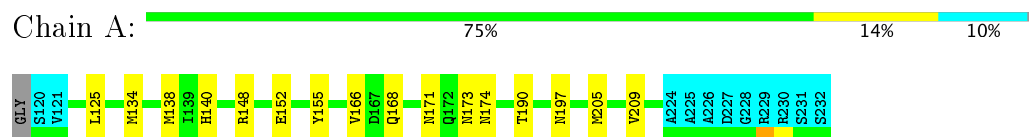
### 4.2.7 Score per residue for model 7

- Molecule 1: Major prion protein



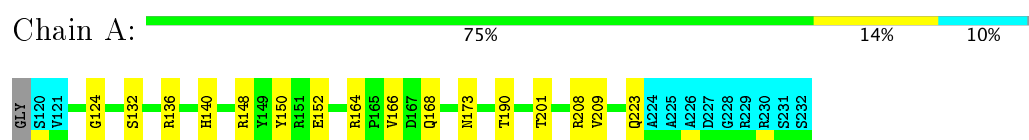
### 4.2.8 Score per residue for model 8

- Molecule 1: Major prion protein



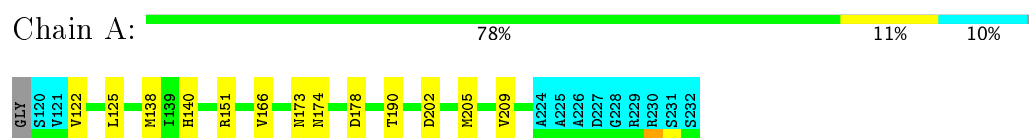
### 4.2.9 Score per residue for model 9

- Molecule 1: Major prion protein



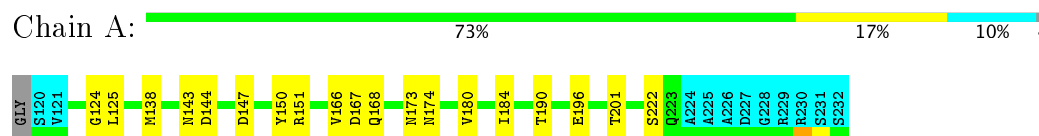
### 4.2.10 Score per residue for model 10 (medoid)

- Molecule 1: Major prion protein



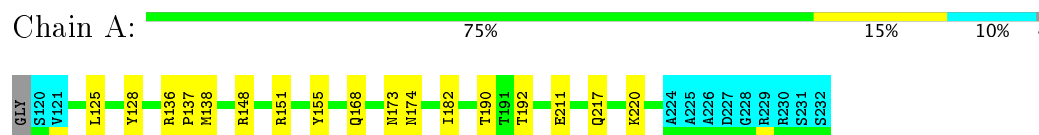
### 4.2.11 Score per residue for model 11

- Molecule 1: Major prion protein



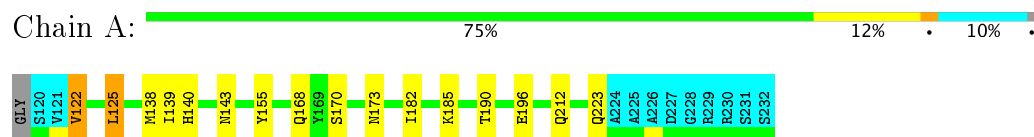
### 4.2.12 Score per residue for model 12

- Molecule 1: Major prion protein



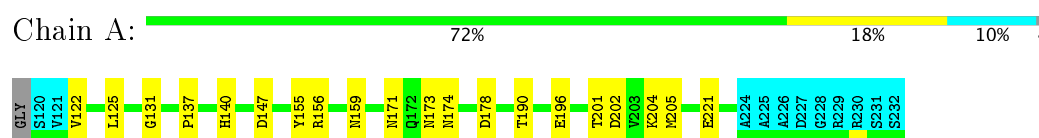
### 4.2.13 Score per residue for model 13

- Molecule 1: Major prion protein



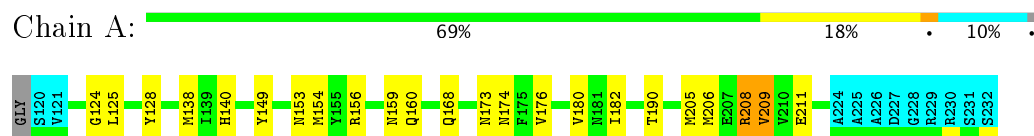
### 4.2.14 Score per residue for model 14

- Molecule 1: Major prion protein



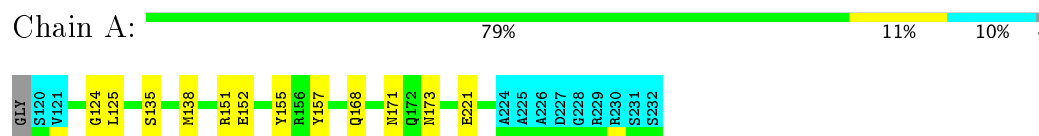
### 4.2.15 Score per residue for model 15

- Molecule 1: Major prion protein



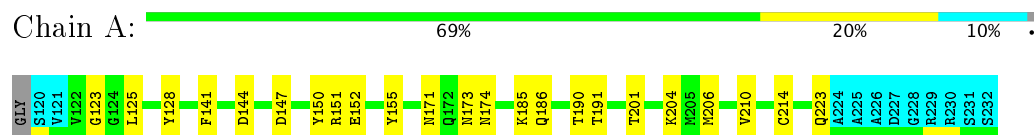
### 4.2.16 Score per residue for model 16

- Molecule 1: Major prion protein



### 4.2.17 Score per residue for model 17

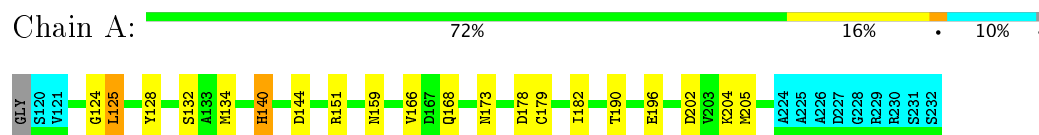
- Molecule 1: Major prion protein





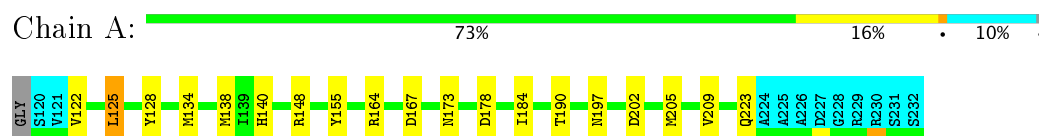
### 4.2.18 Score per residue for model 18

- Molecule 1: Major prion protein



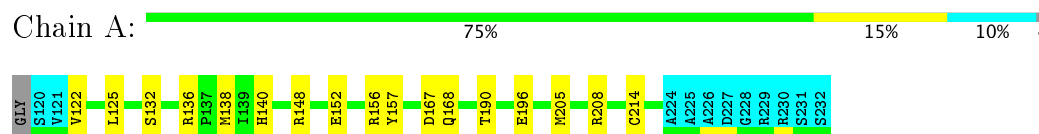
### 4.2.19 Score per residue for model 19

- 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: ?.

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	

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	0.65±0.01	0±0/865 (0.0±0.0%)	1.10±0.03	1±1/1170 (0.1±0.1%)
All	All	0.65	0/17300 (0.0%)	1.10	16/23400 (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	1.4±1.2
All	All	0	28

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	122	VAL	CA-CB-CG1	6.78	121.08	110.90	13	2
1	A	211	GLU	CA-CB-CG	6.07	126.75	113.40	12	1
1	A	198	PHE	CB-CG-CD2	-5.98	116.62	120.80	6	1
1	A	179	CYS	CA-CB-SG	5.88	124.59	114.00	18	1
1	A	136	ARG	NE-CZ-NH2	-5.78	117.41	120.30	12	1
1	A	140	HIS	CA-CB-CG	5.73	123.34	113.60	18	1
1	A	209	VAL	CA-CB-CG1	5.70	119.45	110.90	2	1
1	A	209	VAL	CA-CB-CG2	5.59	119.29	110.90	5	2
1	A	136	ARG	NE-CZ-NH1	5.58	123.09	120.30	2	1
1	A	150	TYR	CB-CG-CD1	-5.53	117.68	121.00	2	1
1	A	151	ARG	NE-CZ-NH2	-5.50	117.55	120.30	3	2
1	A	198	PHE	CB-CG-CD1	5.42	124.60	120.80	6	1
1	A	122	VAL	CB-CA-C	5.05	120.99	111.40	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	157	TYR	Sidechain	4
1	A	148	ARG	Sidechain	3
1	A	208	ARG	Sidechain	3
1	A	156	ARG	Sidechain	3
1	A	150	TYR	Sidechain	3
1	A	136	ARG	Sidechain	3
1	A	151	ARG	Sidechain	2
1	A	162	TYR	Sidechain	2
1	A	149	TYR	Sidechain	2
1	A	190	THR	Peptide	1
1	A	170	SER	Peptide	1
1	A	128	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	845	789	789	1±1
All	All	16900	15780	15780	22

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:180:VAL:HG13	1:A:210:VAL:CG2	0.54	2.32	6	1
1:A:125:LEU:HD11	1:A:182:ILE:HD12	0.53	1.81	5	3
1:A:208:ARG:HA	1:A:211:GLU:HG2	0.52	1.81	15	1
1:A:186:GLN:O	1:A:190:THR:HG22	0.49	2.07	17	1
1:A:128:TYR:CE2	1:A:182:ILE:HG13	0.47	2.44	2	4
1:A:125:LEU:HD23	1:A:128:TYR:CG	0.47	2.44	3	3
1:A:206:MET:HA	1:A:209:VAL:HG12	0.46	1.87	15	1
1:A:180:VAL:HG12	1:A:184:ILE:HD12	0.46	1.88	11	2
1:A:184:ILE:O	1:A:184:ILE:HG22	0.46	2.10	19	1
1:A:139:ILE:HD11	1:A:212:GLN:NE2	0.43	2.29	13	1

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:125:LEU:HD21	1:A:182:ILE:CG2	0.41	2.44	12	1
1:A:206:MET:O	1:A:210:VAL:HG23	0.41	2.15	17	1
1:A:176:VAL:O	1:A:180:VAL:HG23	0.41	2.16	15	1
1:A:125:LEU:HD12	1:A:128:TYR:CD1	0.41	2.50	18	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	102/114 (89%)	88±2 (86±2%)	13±2 (12±2%)	2±1 (2±1%)	16	60
All	All	2040/2280 (89%)	1755 (86%)	252 (12%)	33 (2%)	16	60

All 13 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	122	VAL	8
1	A	124	GLY	6
1	A	168	GLN	4
1	A	167	ASP	3
1	A	141	PHE	2
1	A	137	PRO	2
1	A	132	SER	2
1	A	170	SER	1
1	A	184	ILE	1
1	A	189	VAL	1
1	A	131	GLY	1
1	A	123	GLY	1
1	A	192	THR	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	94/101 (93%)	82±2 (87±3%)	12±2 (13±3%)	9	50	
All	All	1880/2020 (93%)	1638 (87%)	242 (13%)	9	50	

All 48 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	173	ASN	19
1	A	140	HIS	15
1	A	125	LEU	15
1	A	190	THR	13
1	A	138	MET	13
1	A	155	TYR	11
1	A	205	MET	10
1	A	196	GLU	9
1	A	174	ASN	9
1	A	168	GLN	7
1	A	152	GLU	7
1	A	171	ASN	7
1	A	201	THR	7
1	A	164	ARG	6
1	A	166	VAL	6
1	A	209	VAL	6
1	A	159	ASN	5
1	A	178	ASP	5
1	A	147	ASP	5
1	A	204	LYS	5
1	A	221	GLU	4
1	A	202	ASP	4
1	A	153	ASN	4
1	A	151	ARG	4
1	A	185	LYS	3
1	A	134	MET	3
1	A	144	ASP	3
1	A	197	ASN	3
1	A	214	CYS	3
1	A	132	SER	3
1	A	223	GLN	3
1	A	135	SER	2
1	A	220	LYS	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	148	ARG	2
1	A	191	THR	2
1	A	179	CYS	2
1	A	217	GLN	2
1	A	143	ASN	2
1	A	208	ARG	2
1	A	136	ARG	1
1	A	193	THR	1
1	A	194	LYS	1
1	A	222	SER	1
1	A	181	ASN	1
1	A	154	MET	1
1	A	188	THR	1
1	A	192	THR	1
1	A	160	GLN	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