



# Full wwPDB NMR Structure Validation Report ⓘ

Apr 26, 2016 – 03:02 PM BST

PDB ID : 1HJN  
Title : HUMAN PRION PROTEIN AT PH 7.0  
Authors : Calzolari, L.; Zahn, R.  
Deposited on : 2003-02-27

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

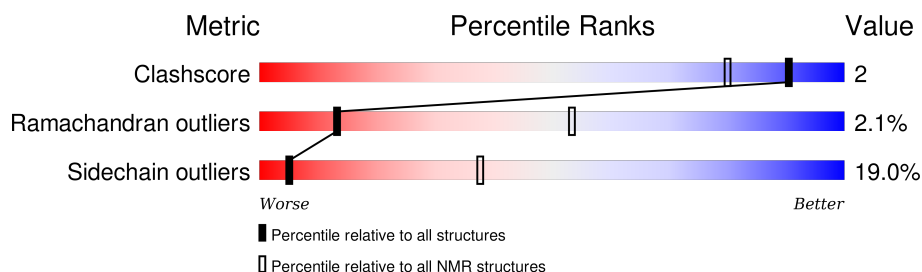
# 1 Overall quality at a glance

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	104	

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 20 is the overall representative, medoid model (most similar to other models).

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:228 (104)	0.52	20

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

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

There is only 1 type of molecule in this entry. The entry contains 1691 atoms, of which 814 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called MAJOR PRION PROTEIN PRECURSOR.

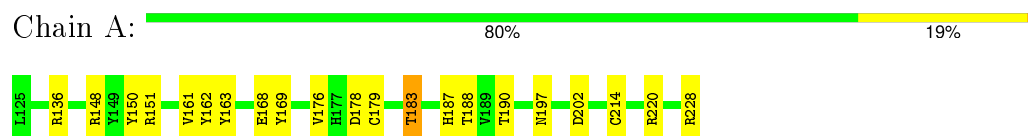
Mol	Chain	Residues	Atoms						Trace
1	A	104	Total	C	H	N	O	S	0
			1691	544	814	153	171	9	

## 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 PRECURSOR

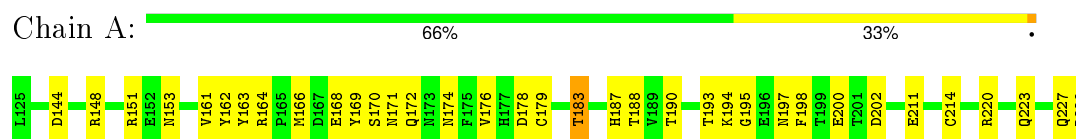


### 4.2 Scores per residue for each member of the ensemble

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

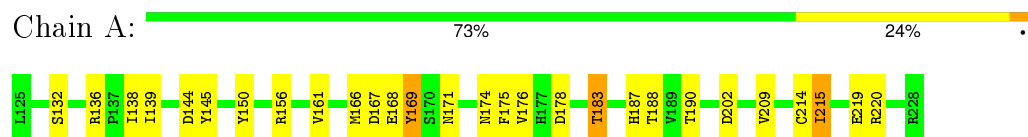
#### 4.2.1 Score per residue for model 1

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



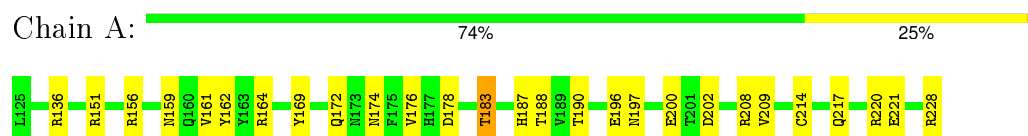
#### 4.2.2 Score per residue for model 2

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



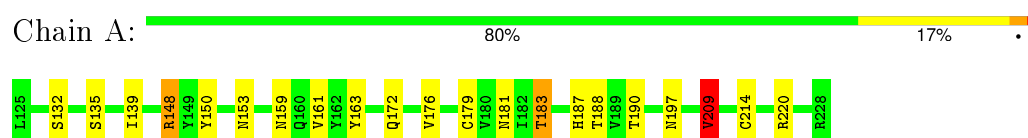
### 4.2.3 Score per residue for model 3

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



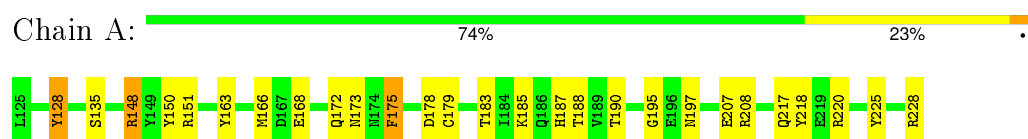
### 4.2.4 Score per residue for model 4

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



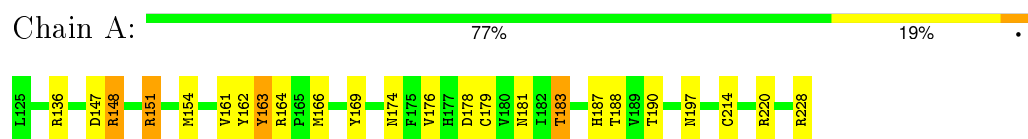
### 4.2.5 Score per residue for model 5

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



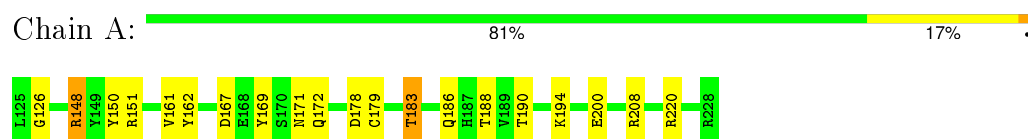
### 4.2.6 Score per residue for model 6

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



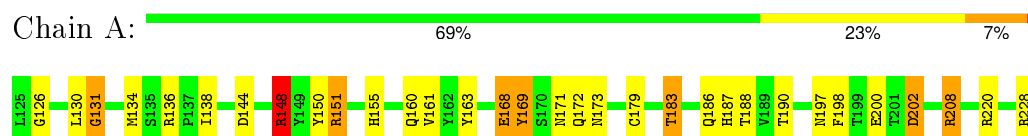
### 4.2.7 Score per residue for model 7

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



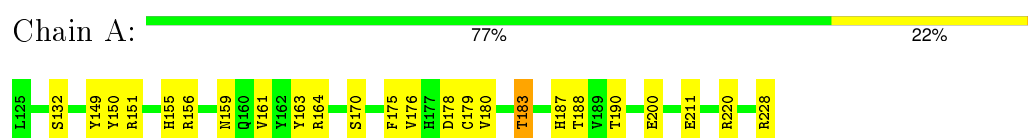
### 4.2.8 Score per residue for model 8

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



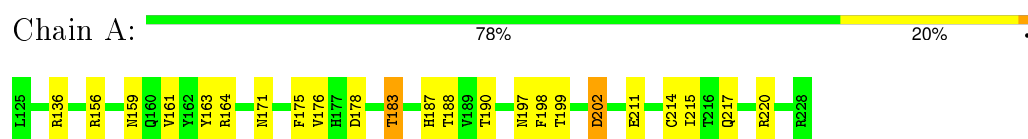
### 4.2.9 Score per residue for model 9

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



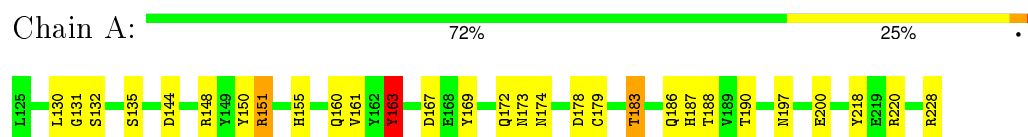
### 4.2.10 Score per residue for model 10

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



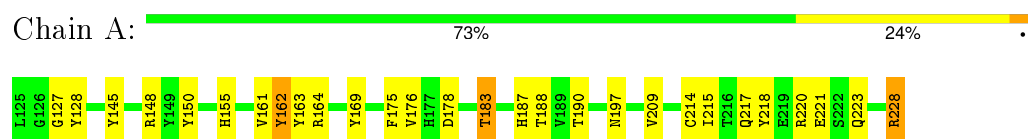
### 4.2.11 Score per residue for model 11

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



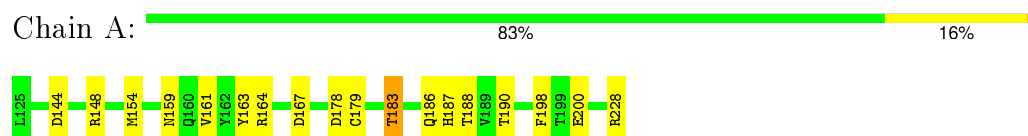
### 4.2.12 Score per residue for model 12

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



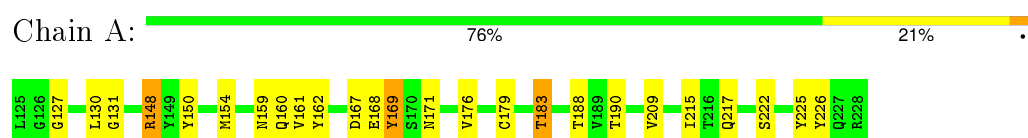
#### 4.2.13 Score per residue for model 13

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



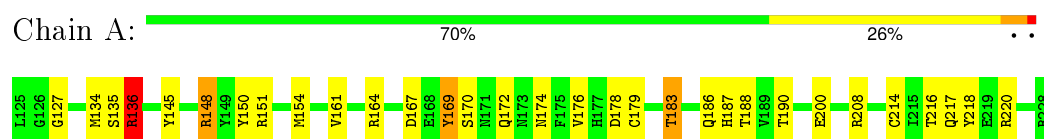
#### 4.2.14 Score per residue for model 14

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



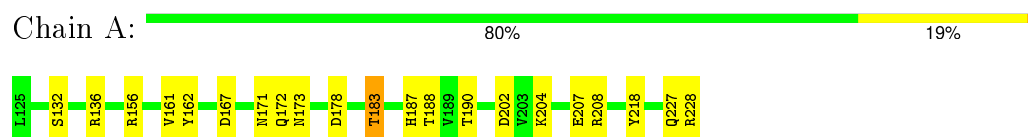
#### 4.2.15 Score per residue for model 15

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



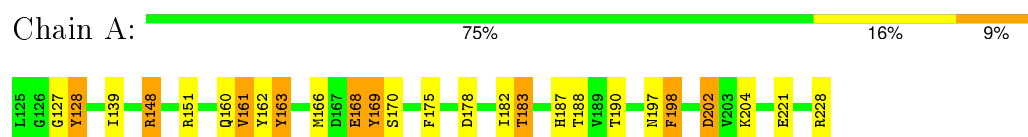
#### 4.2.16 Score per residue for model 16

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



#### 4.2.17 Score per residue for model 17

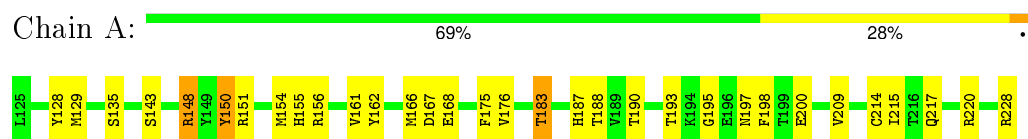
- Molecule 1: MAJOR PRION PROTEIN PRECURSOR





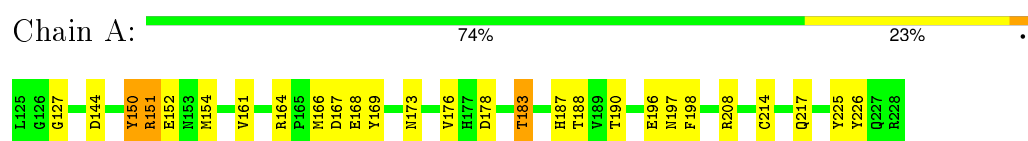
### 4.2.18 Score per residue for model 18

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



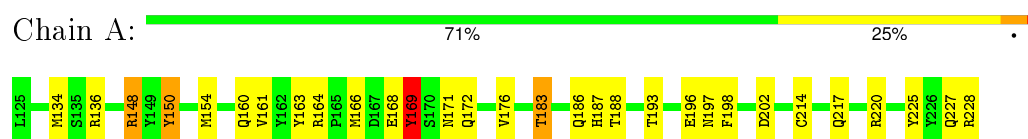
### 4.2.19 Score per residue for model 19

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



### 4.2.20 Score per residue for model 20 (medoid)

- Molecule 1: MAJOR PRION PROTEIN PRECURSOR



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *TORSION ANGLE DYNAMICS*.

Of the 20 calculated structures, 20 were deposited, based on the following criterion: *LOWER TARGET FUNCTION*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
OPALP	refinement	
DYANA	structure solution	

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	0.69±0.01	0±0/897 (0.0±0.0%)	1.08±0.03	1±1/1210 (0.1±0.1%)
All	All	0.69	0/17940 (0.0%)	1.08	23/24200 (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	2.6±1.5
All	All	0	53

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	163	TYR	CB-CG-CD2	-6.83	116.90	121.00	11	3
1	A	136	ARG	NE-CZ-NH2	-6.71	116.94	120.30	16	2
1	A	148	ARG	NE-CZ-NH1	6.47	123.53	120.30	5	2
1	A	164	ARG	NE-CZ-NH2	-6.34	117.13	120.30	19	3
1	A	208	ARG	NE-CZ-NH2	-5.66	117.47	120.30	8	1
1	A	164	ARG	CD-NE-CZ	5.64	131.49	123.60	9	1
1	A	209	VAL	CA-CB-CG2	5.56	119.24	110.90	4	1
1	A	169	TYR	CB-CG-CD2	-5.48	117.71	121.00	8	1
1	A	150	TYR	CB-CG-CD1	-5.34	117.80	121.00	20	1
1	A	164	ARG	NE-CZ-NH1	5.21	122.91	120.30	19	2
1	A	220	ARG	NE-CZ-NH1	5.21	122.90	120.30	8	1
1	A	148	ARG	NE-CZ-NH2	-5.08	117.76	120.30	8	1
1	A	139	ILE	CA-CB-CG1	5.07	120.62	111.00	17	1
1	A	156	ARG	NE-CZ-NH1	5.04	122.82	120.30	18	1
1	A	134	MET	CG-SD-CE	-5.03	92.16	100.20	8	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	156	ARG	NE-CZ-NH2	-5.00	117.80	120.30	3	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	150	TYR	Sidechain	12
1	A	148	ARG	Sidechain	9
1	A	228	ARG	Sidechain	5
1	A	151	ARG	Sidechain	5
1	A	169	TYR	Sidechain	5
1	A	162	TYR	Sidechain	4
1	A	136	ARG	Sidechain	4
1	A	208	ARG	Sidechain	2
1	A	225	TYR	Sidechain	1
1	A	226	TYR	Sidechain	1
1	A	198	PHE	Sidechain	1
1	A	149	TYR	Sidechain	1
1	A	128	TYR	Sidechain	1
1	A	218	TYR	Sidechain	1
1	A	220	ARG	Sidechain	1

## 6.2 Too-close contacts [i](#)

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	877	814	813	4±2
All	All	17540	16280	16260	71

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:161:VAL:HA	1:A:183:THR:HG21	0.66	1.67	14	19
1:A:130:LEU:HD13	1:A:131:GLY:N	0.59	2.12	11	3
1:A:176:VAL:HG22	1:A:214:CYS:HB3	0.58	1.75	10	9
1:A:199:THR:H	1:A:202:ASP:CG	0.57	2.03	10	1
1:A:130:LEU:HD11	1:A:160:GLN:NE2	0.56	2.15	14	3
1:A:151:ARG:CA	1:A:151:ARG:HE	0.54	2.14	8	1
1:A:138:ILE:H	1:A:138:ILE:HD12	0.53	1.64	2	1
1:A:138:ILE:HG23	1:A:151:ARG:NH1	0.53	2.18	8	1
1:A:131:GLY:HA2	1:A:163:TYR:CE1	0.50	2.41	11	1
1:A:176:VAL:HG21	1:A:215:ILE:HG12	0.50	1.82	14	3
1:A:198:PHE:CD2	1:A:202:ASP:HB3	0.49	2.43	17	2
1:A:151:ARG:N	1:A:151:ARG:HE	0.48	2.05	8	1
1:A:151:ARG:HE	1:A:151:ARG:N	0.48	2.07	19	1
1:A:134:MET:CE	1:A:216:THR:HG21	0.47	2.40	15	1
1:A:176:VAL:HG22	1:A:214:CYS:CB	0.46	2.41	3	6
1:A:216:THR:HG23	1:A:217:GLN:N	0.45	2.26	15	1
1:A:138:ILE:HG23	1:A:151:ARG:HH12	0.45	1.71	8	1
1:A:180:VAL:HG21	1:A:211:GLU:HG3	0.45	1.89	9	1
1:A:180:VAL:HG21	1:A:211:GLU:HA	0.44	1.90	9	1
1:A:150:TYR:CD2	1:A:151:ARG:NH2	0.43	2.86	19	1
1:A:139:ILE:CG1	1:A:209:VAL:HG13	0.43	2.44	4	1
1:A:175:PHE:CE2	1:A:218:TYR:CE1	0.43	3.07	5	1
1:A:161:VAL:HA	1:A:183:THR:CG2	0.42	2.43	19	2
1:A:169:TYR:CE2	1:A:175:PHE:HA	0.42	2.49	2	1
1:A:150:TYR:CE1	1:A:209:VAL:HG21	0.42	2.50	12	1
1:A:128:TYR:CD2	1:A:182:ILE:HG21	0.42	2.50	17	1
1:A:139:ILE:CG1	1:A:209:VAL:HG23	0.41	2.45	2	1
1:A:172:GLN:HA	1:A:218:TYR:CE2	0.41	2.51	11	1
1:A:151:ARG:CZ	1:A:154:MET:CE	0.41	2.98	19	1
1:A:169:TYR:CD1	1:A:175:PHE:CD1	0.41	3.09	17	1
1:A:175:PHE:CD2	1:A:218:TYR:CE1	0.41	3.08	5	1
1:A:150:TYR:OH	1:A:209:VAL:HG21	0.40	2.17	18	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	102/104 (98%)	90±3 (88±3%)	10±3 (10±3%)	2±1 (2±1%)	13	53
All	All	2040/2080 (98%)	1799 (88%)	198 (10%)	43 (2%)	13	53

All 15 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	166	MET	7
1	A	168	GLU	5
1	A	127	GLY	5
1	A	135	SER	4
1	A	128	TYR	4
1	A	172	GLN	4
1	A	195	GLY	3
1	A	132	SER	2
1	A	170	SER	2
1	A	126	GLY	2
1	A	131	GLY	1
1	A	169	TYR	1
1	A	167	ASP	1
1	A	198	PHE	1
1	A	134	MET	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	97/97 (100%)	79±3 (81±3%)	18±3 (19±3%)	5	38
All	All	1940/1940 (100%)	1572 (81%)	368 (19%)	5	38

All 65 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	183	THR	20
1	A	188	THR	20
1	A	190	THR	19

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	187	HIS	18
1	A	178	ASP	15
1	A	220	ARG	13
1	A	197	ASN	13
1	A	148	ARG	13
1	A	163	TYR	12
1	A	169	TYR	11
1	A	179	CYS	11
1	A	151	ARG	9
1	A	200	GLU	9
1	A	228	ARG	9
1	A	202	ASP	8
1	A	217	GLN	8
1	A	171	ASN	8
1	A	167	ASP	8
1	A	168	GLU	6
1	A	186	GLN	6
1	A	154	MET	6
1	A	174	ASN	6
1	A	144	ASP	6
1	A	159	ASN	6
1	A	162	TYR	6
1	A	175	PHE	5
1	A	155	HIS	5
1	A	173	ASN	5
1	A	208	ARG	5
1	A	164	ARG	5
1	A	172	GLN	5
1	A	198	PHE	5
1	A	156	ARG	4
1	A	136	ARG	4
1	A	193	THR	3
1	A	145	TYR	3
1	A	215	ILE	3
1	A	221	GLU	3
1	A	132	SER	3
1	A	196	GLU	3
1	A	209	VAL	3
1	A	225	TYR	3
1	A	227	GLN	3
1	A	160	GLN	2
1	A	153	ASN	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	223	GLN	2
1	A	211	GLU	2
1	A	207	GLU	2
1	A	218	TYR	2
1	A	204	LYS	2
1	A	194	LYS	2
1	A	181	ASN	2
1	A	170	SER	2
1	A	147	ASP	1
1	A	143	SER	1
1	A	166	MET	1
1	A	152	GLU	1
1	A	226	TYR	1
1	A	135	SER	1
1	A	129	MET	1
1	A	176	VAL	1
1	A	222	SER	1
1	A	219	GLU	1
1	A	185	LYS	1
1	A	161	VAL	1

### 6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

### 6.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.6 Ligand geometry ⓘ

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