



# Full wwPDB NMR Structure Validation Report ⓘ

Oct 28, 2024 – 02:00 pm GMT

PDB ID : 1E1G  
Title : Human prion protein variant M166V  
Authors : Calzolari, L.; Lysek, D.A.; Guntert, P.; Von Schroetter, C.; Zahn, R.; Riek, R.; Wuthrich, K.  
Deposited on : 2000-05-04

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

<https://www.wwpdb.org/validation/2017/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

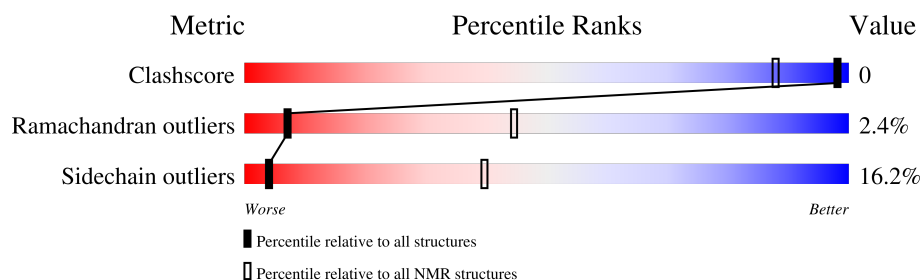
# 1 Overall quality at a glance

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	210492	14027
Ramachandran outliers	207382	12486
Sidechain outliers	206894	12463

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	<div> <div></div> <div>91%</div> <div>9%</div> </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 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)	1.12	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 4 clusters and 2 single-model clusters were found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1690 atoms, of which 814 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called PRION PROTEIN.

Mol	Chain	Residues	Atoms						Trace
1	A	104	Total	C	H	N	O	S	0
			1690	544	814	153	171	8	

There is a discrepancy between the modelled and reference sequences:

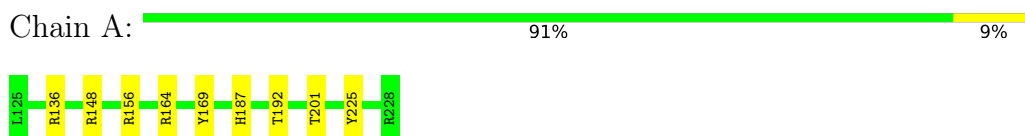
Chain	Residue	Modelled	Actual	Comment	Reference
A	166	VAL	MET	engineered mutation	UNP P78446

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

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

These plots are provided for all protein, RNA, DNA and oligosaccharide 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: PRION PROTEIN

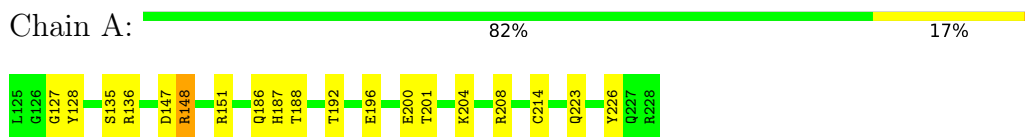


### 4.2 Scores per residue for each member of the ensemble

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

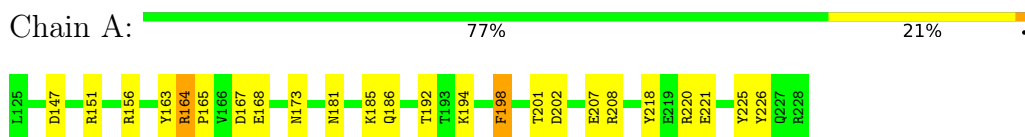
#### 4.2.1 Score per residue for model 1

- Molecule 1: PRION PROTEIN



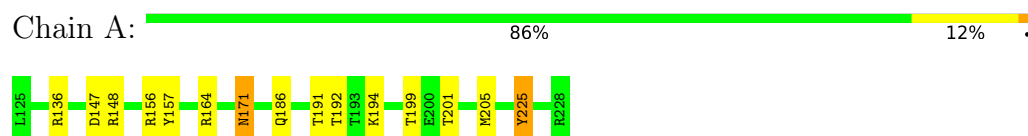
#### 4.2.2 Score per residue for model 2

- Molecule 1: PRION PROTEIN



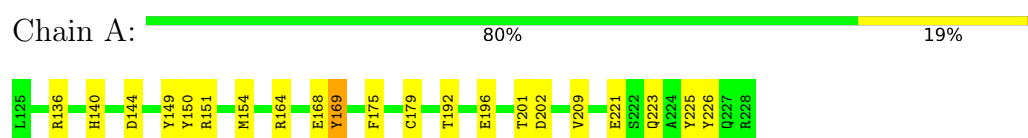
### 4.2.3 Score per residue for model 3

- Molecule 1: PRION PROTEIN



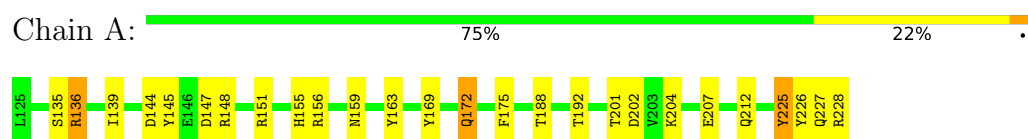
### 4.2.4 Score per residue for model 4

- Molecule 1: PRION PROTEIN



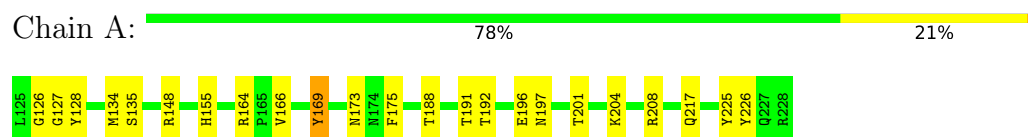
### 4.2.5 Score per residue for model 5

- Molecule 1: PRION PROTEIN



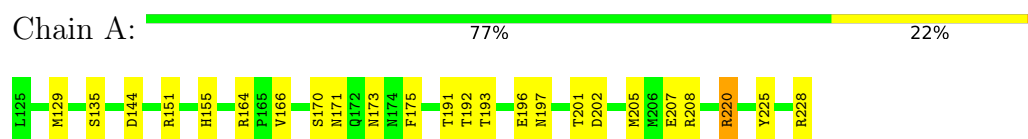
### 4.2.6 Score per residue for model 6

- Molecule 1: PRION PROTEIN



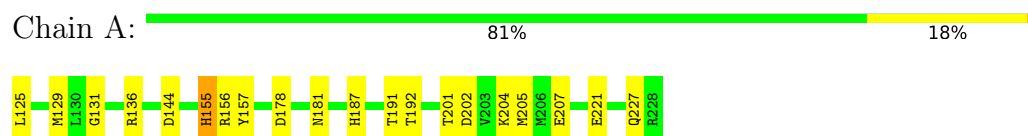
### 4.2.7 Score per residue for model 7

- Molecule 1: PRION PROTEIN



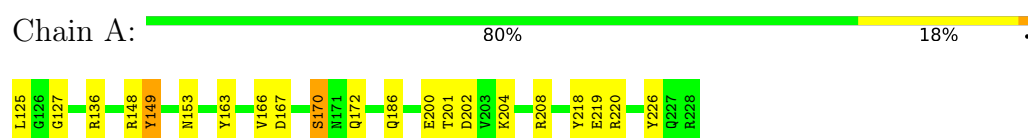
### 4.2.8 Score per residue for model 8

- Molecule 1: PRION PROTEIN



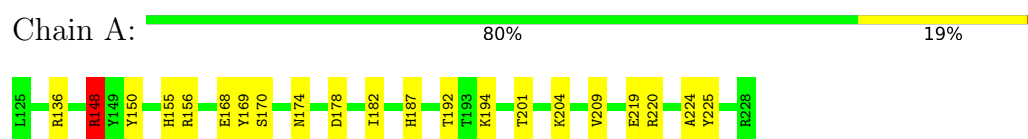
### 4.2.9 Score per residue for model 9

- Molecule 1: PRION PROTEIN



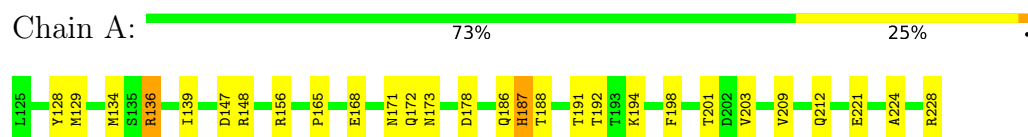
### 4.2.10 Score per residue for model 10

- Molecule 1: PRION PROTEIN



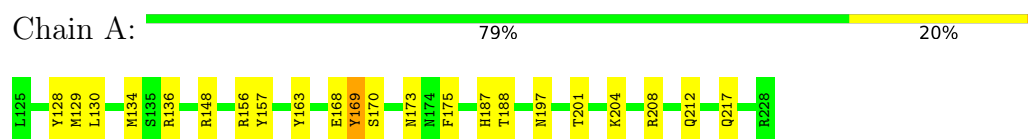
### 4.2.11 Score per residue for model 11

- Molecule 1: PRION PROTEIN



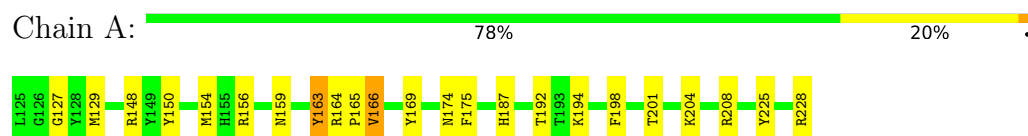
### 4.2.12 Score per residue for model 12 (medoid)

- Molecule 1: PRION PROTEIN



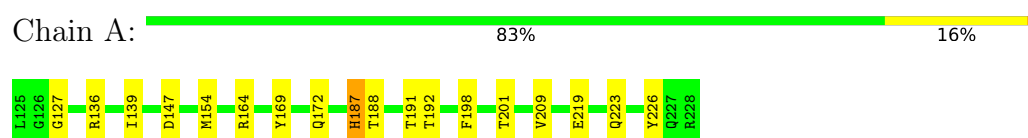
#### 4.2.13 Score per residue for model 13

- Molecule 1: PRION PROTEIN



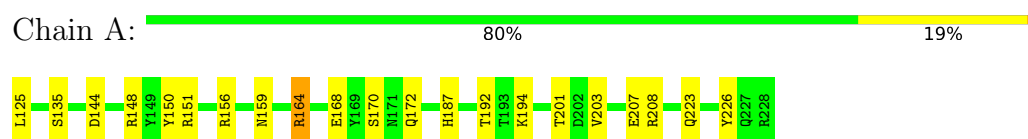
#### 4.2.14 Score per residue for model 14

- Molecule 1: PRION PROTEIN



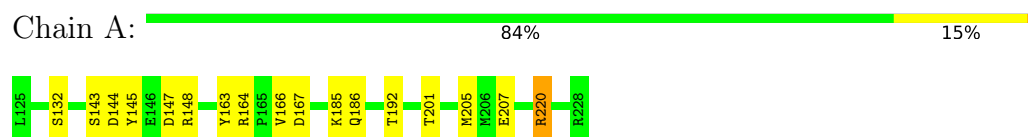
#### 4.2.15 Score per residue for model 15

- Molecule 1: PRION PROTEIN



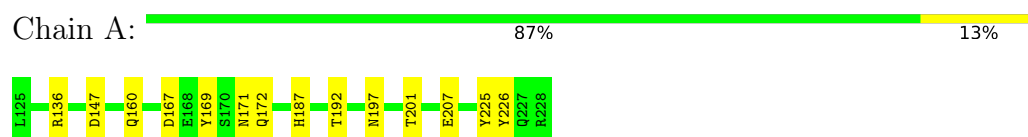
#### 4.2.16 Score per residue for model 16

- Molecule 1: PRION PROTEIN



#### 4.2.17 Score per residue for model 17

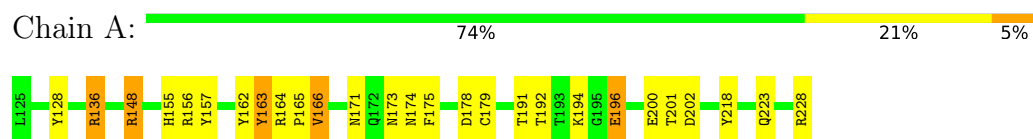
- Molecule 1: PRION PROTEIN





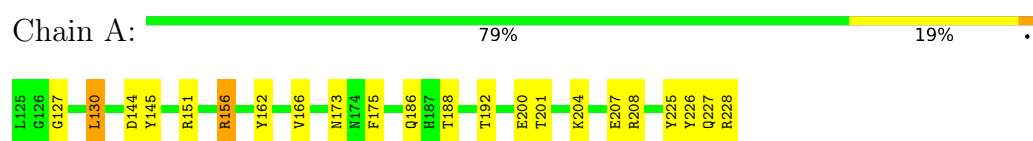
#### 4.2.18 Score per residue for model 18

- Molecule 1: PRION PROTEIN



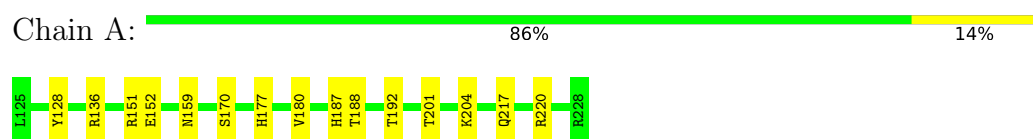
#### 4.2.19 Score per residue for model 19

- Molecule 1: PRION PROTEIN



#### 4.2.20 Score per residue for model 20

- Molecule 1: PRION PROTEIN



## 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: ?.

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.

## 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	0.68±0.01	0±0/896 ( 0.0± 0.0%)	1.06±0.04	1±1/1210 ( 0.1± 0.1%)
All	All	0.68	0/17920 ( 0.0%)	1.06	26/24200 ( 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	2.2±1.9
All	All	0	45

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	148	ARG	NE-CZ-NH2	-8.22	116.19	120.30	6	3
1	A	169	TYR	CB-CG-CD2	-6.82	116.91	121.00	13	1
1	A	164	ARG	NE-CZ-NH2	-6.68	116.96	120.30	7	3
1	A	149	TYR	CB-CG-CD2	-6.51	117.09	121.00	9	1
1	A	157	TYR	CB-CG-CD2	-6.50	117.10	121.00	3	3
1	A	156	ARG	NE-CZ-NH1	6.41	123.50	120.30	19	2
1	A	148	ARG	NE-CZ-NH1	6.03	123.31	120.30	10	1
1	A	220	ARG	NE-CZ-NH2	-5.68	117.46	120.30	16	2
1	A	180	VAL	CA-CB-CG1	5.64	119.36	110.90	20	1
1	A	136	ARG	CD-NE-CZ	5.62	131.47	123.60	11	1
1	A	136	ARG	NE-CZ-NH1	5.59	123.09	120.30	5	1
1	A	225	TYR	CB-CG-CD2	-5.36	117.78	121.00	5	1
1	A	150	TYR	CB-CG-CD1	-5.35	117.79	121.00	15	1
1	A	202	ASP	CB-CG-OD1	5.29	123.06	118.30	8	1
1	A	151	ARG	NE-CZ-NH2	-5.26	117.67	120.30	20	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	136	ARG	NE-CZ-NH2	-5.07	117.77	120.30	17	1
1	A	228	ARG	NE-CZ-NH2	-5.06	117.77	120.30	19	1
1	A	157	TYR	CB-CG-CD1	5.00	124.00	121.00	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	128	TYR	Sidechain	5
1	A	148	ARG	Sidechain	4
1	A	164	ARG	Sidechain	4
1	A	220	ARG	Sidechain	4
1	A	225	TYR	Sidechain	4
1	A	151	ARG	Sidechain	3
1	A	218	TYR	Sidechain	3
1	A	150	TYR	Sidechain	3
1	A	156	ARG	Sidechain	2
1	A	169	TYR	Sidechain	2
1	A	163	TYR	Sidechain	2
1	A	198	PHE	Sidechain	1
1	A	149	TYR	Sidechain	1
1	A	208	ARG	Sidechain	1
1	A	167	ASP	Peptide	1
1	A	228	ARG	Sidechain	1
1	A	166	VAL	Peptide	1
1	A	136	ARG	Sidechain	1
1	A	162	TYR	Sidechain	1
1	A	226	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	876	814	813	0±1
All	All	17520	16280	16260	10

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 0.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:139:ILE:HD11	1:A:209:VAL:HG22	0.53	1.80	14	1
1:A:130:LEU:HD23	1:A:162:TYR:CE1	0.51	2.39	19	1
1:A:191:THR:HG23	1:A:196:GLU:HG2	0.51	1.83	18	1
1:A:187:HIS:CD2	1:A:198:PHE:CD2	0.47	3.03	11	2
1:A:169:TYR:CD2	1:A:175:PHE:CE1	0.46	3.04	6	1
1:A:139:ILE:HD11	1:A:212:GLN:HG2	0.43	1.88	5	1
1:A:139:ILE:HD11	1:A:212:GLN:CG	0.43	2.43	11	2
1:A:196:GLU:HG2	1:A:197:ASN:H	0.40	1.76	7	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%)	89±3 (87±3%)	11±3 (11±3%)	2±1 (2±1%)	7	44
All	All	2040/2080 (98%)	1771 (87%)	220 (11%)	49 (2%)	7	44

All 19 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	127	GLY	6
1	A	172	GLN	6
1	A	170	SER	6
1	A	166	VAL	5
1	A	165	PRO	4
1	A	171	ASN	3
1	A	169	TYR	3
1	A	198	PHE	2
1	A	135	SER	2
1	A	155	HIS	2
1	A	224	ALA	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	167	ASP	1
1	A	126	GLY	1
1	A	128	TYR	1
1	A	131	GLY	1
1	A	157	TYR	1
1	A	219	GLU	1
1	A	156	ARG	1
1	A	200	GLU	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%)	81±3 (84±3%)	16±3 (16±3%)	4	40
All	All	1940/1940 (100%)	1625 (84%)	315 (16%)	4	40

All 70 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	201	THR	20
1	A	192	THR	18
1	A	136	ARG	12
1	A	187	HIS	10
1	A	204	LYS	10
1	A	226	TYR	9
1	A	147	ASP	8
1	A	188	THR	8
1	A	208	ARG	8
1	A	207	GLU	8
1	A	148	ARG	8
1	A	186	GLN	7
1	A	163	TYR	7
1	A	173	ASN	7
1	A	194	LYS	7
1	A	225	TYR	7
1	A	156	ARG	7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	144	ASP	7
1	A	175	PHE	7
1	A	168	GLU	6
1	A	202	ASP	6
1	A	191	THR	6
1	A	223	GLN	5
1	A	164	ARG	5
1	A	155	HIS	5
1	A	129	MET	5
1	A	196	GLU	4
1	A	151	ARG	4
1	A	221	GLU	4
1	A	205	MET	4
1	A	169	TYR	4
1	A	159	ASN	4
1	A	228	ARG	4
1	A	178	ASP	4
1	A	135	SER	3
1	A	200	GLU	3
1	A	171	ASN	3
1	A	154	MET	3
1	A	209	VAL	3
1	A	145	TYR	3
1	A	227	GLN	3
1	A	134	MET	3
1	A	197	ASN	3
1	A	217	GLN	3
1	A	125	LEU	3
1	A	174	ASN	3
1	A	166	VAL	3
1	A	181	ASN	2
1	A	185	LYS	2
1	A	179	CYS	2
1	A	220	ARG	2
1	A	219	GLU	2
1	A	203	VAL	2
1	A	130	LEU	2
1	A	167	ASP	2
1	A	214	CYS	1
1	A	199	THR	1
1	A	140	HIS	1
1	A	172	GLN	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	193	THR	1
1	A	149	TYR	1
1	A	153	ASN	1
1	A	170	SER	1
1	A	182	ILE	1
1	A	212	GLN	1
1	A	132	SER	1
1	A	143	SER	1
1	A	160	GLN	1
1	A	152	GLU	1
1	A	177	HIS	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 oligosaccharides 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