Full wwPDB NMR Structure Validation Report (i)

Jun 16, 2024 – 12:57 AM EDT

PDB ID	:	2L1K
Title	:	Mouse prion protein (121-231) containing the substitutions Y169A, Y225A, and Y226A
Authors	:	Christen, B.; Damberger, F.F.; Perez, D.R.; Hornemann, S.; Wuthrich, K.
Deposited on	:	2010-07-28

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *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 (i) 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 (i)) were used in the production of this report:

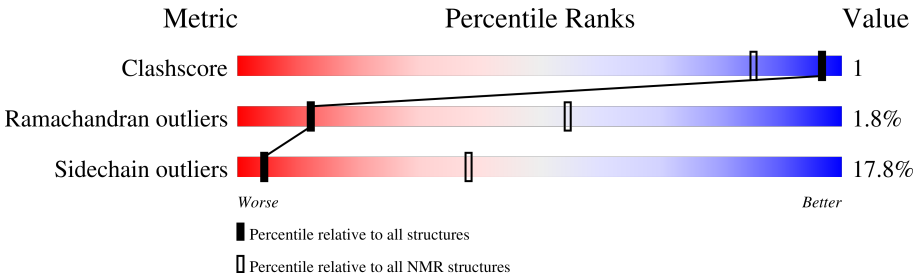
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

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	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	114	80% 12% 8%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 16 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:125-A:229 (105)	0.68	16

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 4 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1776 atoms, of which 860 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Major prion protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	114	1776	562	860	165	180	9	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	119	GLY	-	EXPRESSION TAG	UNP Q4FJQ7
A	120	SER	-	EXPRESSION TAG	UNP Q4FJQ7
A	169	ALA	TYR	ENGINEERED MUTATION	UNP Q4FJQ7
A	225	ALA	TYR	ENGINEERED MUTATION	UNP Q4FJQ7
A	226	ALA	TYR	ENGINEERED MUTATION	UNP Q4FJQ7

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: Major prion protein

Chain A: 80% 12% 8%

G119 S120 V121 V122 G123 G124 M129 S132 R136 P137 M138 V166 D167 S170 N174 I182 H187 T188 V189 N197 V209 Q223 R230 S231 S232

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

Chain A: 75% 15% • 8%

G119 S120 V121 V122 G123 G124 Y128 M129 S132 R136 P137 M138 D147 R148 Y155 V166 S170 N174 I182 V189 T193 E196 N197 V209 C214 Q223 R229 R230 S231 S232

4.2.2 Score per residue for model 2

• Molecule 1: Major prion protein

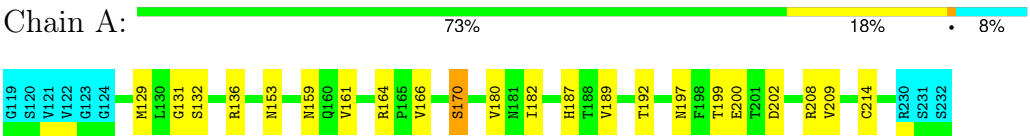
Chain A: 75% 17% 8%

G119 S120 V121 V122 G123 G124 M129 L130 M138 I139 H140 D144 N159 R164 P165 V166 N173 H187 T190 E196 N197 D202 V209 C214 Q217 Q223 R229 R230 S231 S232

4.2.3 Score per residue for model 3

- Molecule 1: Major prion protein

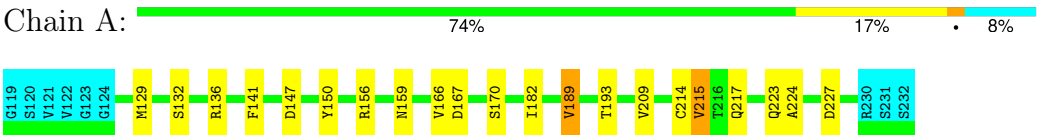
Chain A:



4.2.4 Score per residue for model 4

- Molecule 1: Major prion protein

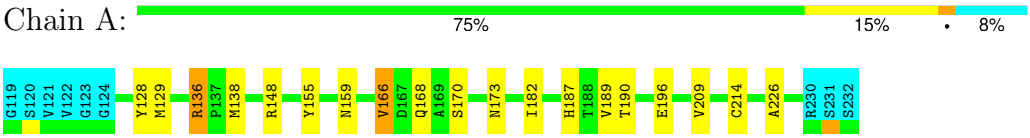
Chain A:



4.2.5 Score per residue for model 5

- Molecule 1: Major prion protein

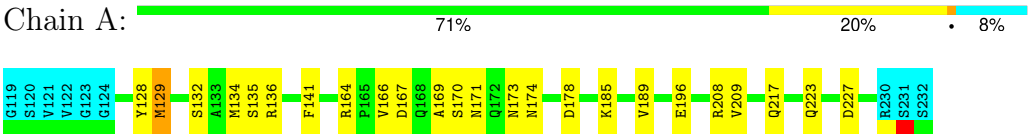
Chain A:



4.2.6 Score per residue for model 6

- Molecule 1: Major prion protein

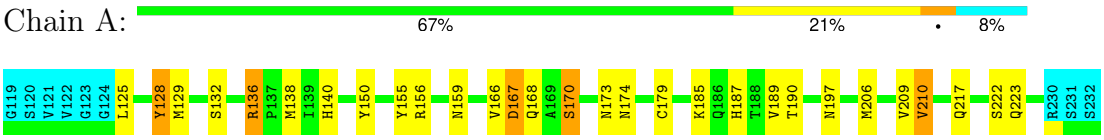
Chain A:



4.2.7 Score per residue for model 7

- Molecule 1: Major prion protein

Chain A:



4.2.8 Score per residue for model 8

- Molecule 1: Major prion protein

Chain A: 74% 16% • 8%

G119 S120 V121 V122 G123 G124 Y128 M129 R136 P137 M138 Y155 Y162 V166 D167 N174 Q186 V189 T190 T191 T192 T193 T199 D202 R208 V209 Q217 Q223 R229 R230 S231 S232

4.2.9 Score per residue for model 9

- Molecule 1: Major prion protein

Chain A: 68% 21% • 8%

G119 S120 V121 V122 G123 G124 L125 L130 G131 S132 R136 P137 M138 N143 Y150 Y155 R156 Y157 P158 N159 R164 P165 V166 D167 Q168 A169 S170 N174 C179 I182 H187 E196 T201 D202 M206 E207 R208 V209 V210 Q223 R229 R230 S231 S232

4.2.10 Score per residue for model 10

- Molecule 1: Major prion protein

Chain A: 72% 19% • 8%

G119 S120 V121 V122 G123 G124 Y128 S132 R136 P137 M138 N143 R148 R151 R156 V166 D167 S170 N174 H187 T190 T191 T192 E196 N197 V209 V215 T216 Q217 Y218 E221 S222 Q223 R230 S231 S232

4.2.11 Score per residue for model 11

- Molecule 1: Major prion protein

Chain A: 70% 20% • 8%

G119 S120 V121 V122 G123 G124 M129 S132 S135 R136 P137 M138 I139 H140 F141 G142 N143 Y149 Y150 Y157 P158 N159 V166 S170 N174 H187 T188 V189 E196 M206 V209 C214 Q217 Q223 R229 R230 S231 S232

4.2.12 Score per residue for model 12

- Molecule 1: Major prion protein

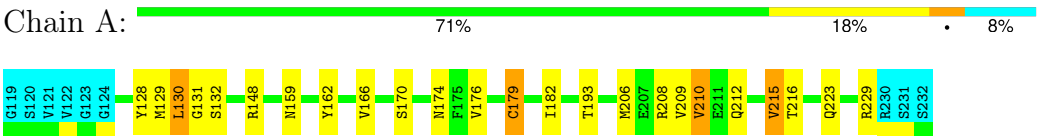
Chain A: 70% 19% • 8%

G119 S120 V121 V122 G123 G124 M129 S132 R136 P137 M138 F141 G142 N143 D144 W145 E146 V166 D167 S170 N173 I182 H187 T188 V189 T190 T191 T192 T193 N197 M206 E207 R208 V209 V210 E211 Q223 R230 S231 S232

4.2.13 Score per residue for model 13

- Molecule 1: Major prion protein

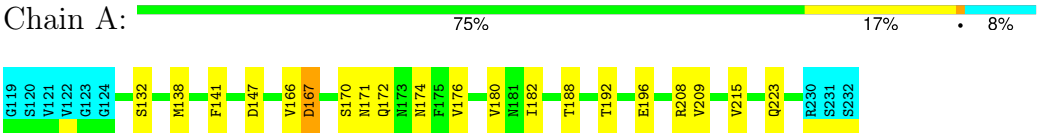
Chain A:



4.2.14 Score per residue for model 14

- Molecule 1: Major prion protein

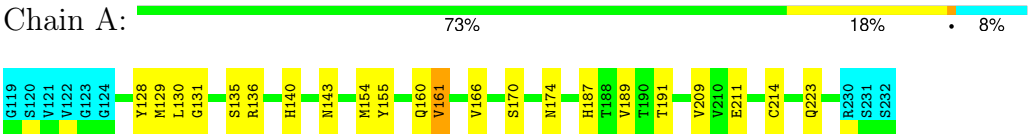
Chain A:



4.2.15 Score per residue for model 15

- Molecule 1: Major prion protein

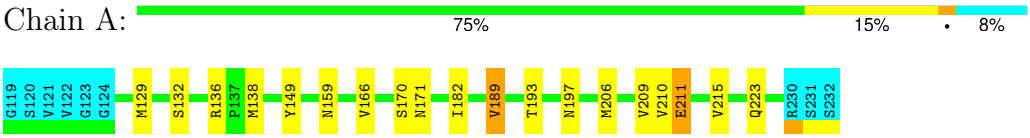
Chain A:



4.2.16 Score per residue for model 16 (medoid)

- Molecule 1: Major prion protein

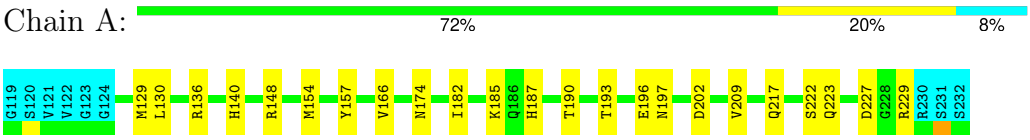
Chain A:



4.2.17 Score per residue for model 17

- Molecule 1: Major prion protein

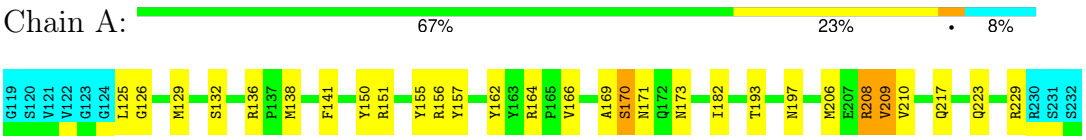
Chain A:



4.2.18 Score per residue for model 18

- Molecule 1: Major prion protein

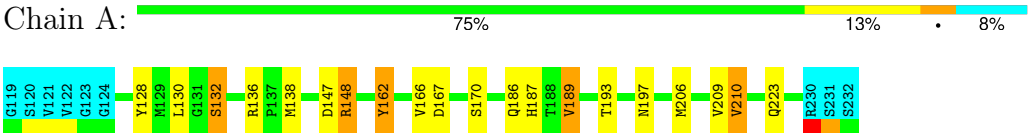
Chain A:



4.2.19 Score per residue for model 19

- Molecule 1: Major prion protein

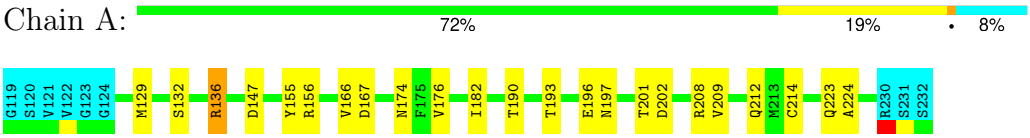
Chain A:



4.2.20 Score per residue for model 20

- Molecule 1: Major prion protein

Chain A:



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *torsion angle dynamics*.

Of the 80 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
DYANA	structure solution	1.0.3
OPALP	refinement	

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	0.64±0.01	0±0/880 (0.0± 0.0%)	1.10±0.02	2±2/1190 (0.2± 0.1%)
All	All	0.64	0/17600 (0.0%)	1.10	43/23800 (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	2.4±1.5
All	All	0	47

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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	209	VAL	CA-CB-CG2	9.70	125.45	110.90	11	1
1	A	229	ARG	NE-CZ-NH2	-7.15	116.73	120.30	2	4
1	A	210	VAL	CA-CB-CG1	6.85	121.17	110.90	7	2
1	A	162	TYR	CB-CG-CD1	-6.70	116.98	121.00	18	1
1	A	189	VAL	CA-CB-CG1	6.51	120.66	110.90	11	7
1	A	148	ARG	NE-CZ-NH2	-6.34	117.13	120.30	19	3
1	A	136	ARG	NE-CZ-NH1	6.32	123.46	120.30	1	1
1	A	176	VAL	CA-CB-CG1	6.20	120.20	110.90	20	2
1	A	208	ARG	NE-CZ-NH2	-5.77	117.42	120.30	18	2
1	A	136	ARG	NE-CZ-NH2	-5.70	117.45	120.30	1	1
1	A	156	ARG	NE-CZ-NH2	-5.70	117.45	120.30	18	2
1	A	170	SER	CB-CA-C	5.60	120.75	110.10	14	1
1	A	210	VAL	CA-CB-CG2	5.59	119.28	110.90	7	1
1	A	215	VAL	CA-CB-CG1	5.56	119.24	110.90	13	2
1	A	156	ARG	NE-CZ-NH1	5.51	123.06	120.30	18	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	209	VAL	CA-CB-CG1	5.47	119.10	110.90	18	2
1	A	192	THR	CA-CB-CG2	5.35	119.89	112.40	14	1
1	A	157	TYR	CB-CG-CD2	-5.33	117.81	121.00	9	1
1	A	191	THR	CA-CB-CG2	5.24	119.74	112.40	12	2
1	A	215	VAL	CG1-CB-CG2	-5.18	102.61	110.90	10	1
1	A	164	ARG	NE-CZ-NH2	-5.13	117.73	120.30	3	1
1	A	130	LEU	CB-CG-CD2	5.07	119.61	111.00	13	1
1	A	190	THR	CA-CB-CG2	5.06	119.48	112.40	5	1
1	A	208	ARG	NE-CZ-NH1	5.04	122.82	120.30	18	1
1	A	164	ARG	NE-CZ-NH1	5.01	122.80	120.30	18	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	9
1	A	136	ARG	Sidechain	5
1	A	150	TYR	Sidechain	5
1	A	208	ARG	Sidechain	4
1	A	229	ARG	Sidechain	3
1	A	164	ARG	Sidechain	3
1	A	162	TYR	Sidechain	3
1	A	148	ARG	Sidechain	3
1	A	169	ALA	Peptide	2
1	A	196	GLU	Peptide	2
1	A	149	TYR	Sidechain	2
1	A	157	TYR	Sidechain	2
1	A	226	ALA	Peptide	1
1	A	156	ARG	Sidechain	1
1	A	218	TYR	Sidechain	1
1	A	151	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	861	805	805	1±1
All	All	17220	16100	16100	23

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:HA	1:A:209:VAL:HG22	0.72	1.62	11	1
1:A:206:MET:O	1:A:210:VAL:HG12	0.63	1.94	9	6
1:A:130:LEU:HD21	1:A:160:GLN:HB3	0.58	1.76	15	1
1:A:206:MET:O	1:A:210:VAL:HG13	0.52	2.04	7	1
1:A:125:LEU:HB3	1:A:128:TYR:CD1	0.48	2.44	7	1
1:A:224:ALA:HA	1:A:227:ASP:OD1	0.47	2.10	4	1
1:A:129:MET:CE	1:A:129:MET:HA	0.46	2.41	7	2
1:A:211:GLU:O	1:A:215:VAL:HG23	0.45	2.12	16	1
1:A:134:MET:CE	1:A:217:GLN:HG2	0.43	2.43	6	1
1:A:206:MET:HG3	1:A:210:VAL:HG12	0.43	1.89	7	1
1:A:172:GLN:HE21	1:A:215:VAL:HG13	0.42	1.74	14	1
1:A:176:VAL:HA	1:A:179:CYS:SG	0.41	2.54	13	1
1:A:130:LEU:HD22	1:A:161:VAL:O	0.41	2.16	15	1
1:A:130:LEU:HD21	1:A:160:GLN:CB	0.41	2.45	15	1
1:A:154:MET:HA	1:A:157:TYR:CE2	0.41	2.51	17	1
1:A:130:LEU:HD13	1:A:162:TYR:CZ	0.41	2.51	19	1
1:A:192:THR:HG23	1:A:197:ASN:HA	0.40	1.93	10	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	105/114 (92%)	89±3 (84±3%)	14±3 (14±3%)	2±1 (2±1%)	12	54
All	All	2100/2280 (92%)	1774 (84%)	289 (14%)	37 (2%)	12	54

All 10 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	170	SER	15
1	A	141	PHE	6
1	A	167	ASP	5
1	A	131	GLY	3
1	A	132	SER	3
1	A	166	VAL	1
1	A	185	LYS	1
1	A	142	GLY	1
1	A	171	ASN	1
1	A	126	GLY	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	94/100 (94%)	77±3 (82±3%)	17±3 (18±3%)	4 38
All	All	1880/2000 (94%)	1545 (82%)	335 (18%)	4 38

All 61 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	166	VAL	20
1	A	209	VAL	19
1	A	223	GLN	18
1	A	136	ARG	16
1	A	129	MET	15
1	A	132	SER	13
1	A	138	MET	13
1	A	174	ASN	12
1	A	182	ILE	12
1	A	189	VAL	11
1	A	187	HIS	11
1	A	193	THR	10
1	A	197	ASN	10
1	A	159	ASN	9
1	A	155	TYR	8
1	A	196	GLU	8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	214	CYS	8
1	A	217	GLN	8
1	A	167	ASP	8
1	A	173	ASN	6
1	A	190	THR	6
1	A	202	ASP	6
1	A	147	ASP	5
1	A	170	SER	5
1	A	143	ASN	5
1	A	130	LEU	4
1	A	140	HIS	4
1	A	208	ARG	4
1	A	229	ARG	3
1	A	215	VAL	3
1	A	168	GLN	3
1	A	135	SER	3
1	A	171	ASN	3
1	A	179	CYS	3
1	A	211	GLU	3
1	A	144	ASP	2
1	A	161	VAL	2
1	A	180	VAL	2
1	A	192	THR	2
1	A	199	THR	2
1	A	156	ARG	2
1	A	227	ASP	2
1	A	185	LYS	2
1	A	222	SER	2
1	A	186	GLN	2
1	A	125	LEU	2
1	A	201	THR	2
1	A	188	THR	2
1	A	212	GLN	2
1	A	153	ASN	1
1	A	200	GLU	1
1	A	148	ARG	1
1	A	178	ASP	1
1	A	206	MET	1
1	A	151	ARG	1
1	A	221	GLU	1
1	A	139	ILE	1
1	A	146	GLU	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	216	THR	1
1	A	154	MET	1
1	A	210	VAL	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 monosaccharides 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 (i)

No chemical shift data were provided