



# Full wwPDB NMR Structure Validation Report ⓘ

Apr 26, 2016 – 02:20 PM BST

PDB ID : 1E1P  
Title : HUMAN PRION PROTEIN VARIANT S170N  
Authors : Calzolari, L.; Lysek, D.A.; Guntert, P.; Von Schroetter, C.; Zahn, R.; Riek, R.; Wuthrich, K.  
Deposited on : 2000-05-09

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

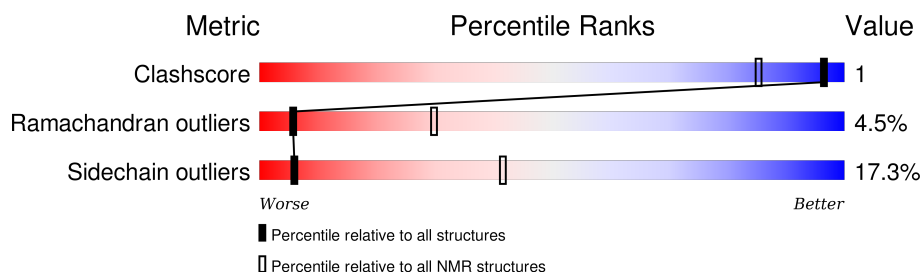
# 1 Overall quality at a glance

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	<div> <div>71%</div> <div>21%</div> <div>8%</div> </div>

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 5 is the overall representative, medoid model (most similar to other models). The authors have identified model 9 as representative.

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:127-A:222 (96)	0.49	5

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1694 atoms, of which 815 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
			1694	545	815	154	171	9	

There is a discrepancy between the modelled and reference sequences:

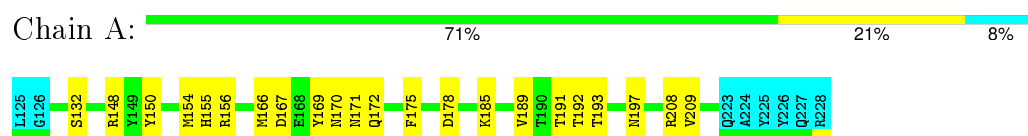
Chain	Residue	Modelled	Actual	Comment	Reference
A	170	ASN	SER	ENGINEERED	UNP P78446

## 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: PRION PROTEIN

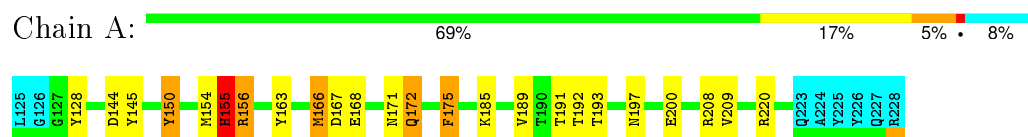


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

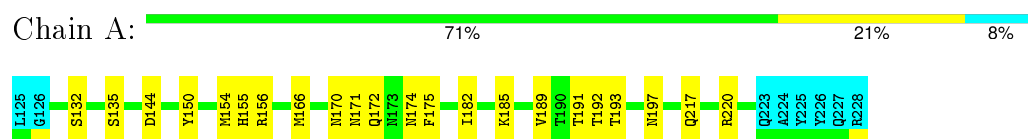
#### 4.2.1 Score per residue for model 1

- Molecule 1: PRION PROTEIN



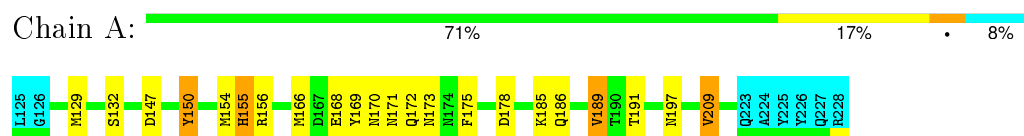
#### 4.2.2 Score per residue for model 2

- Molecule 1: PRION PROTEIN



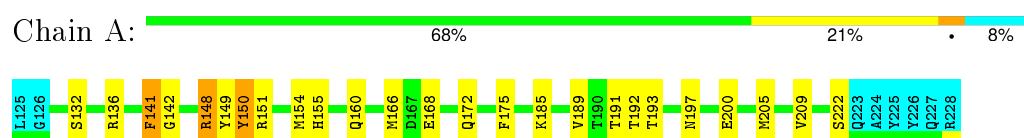
### 4.2.3 Score per residue for model 3

- Molecule 1: PRION PROTEIN



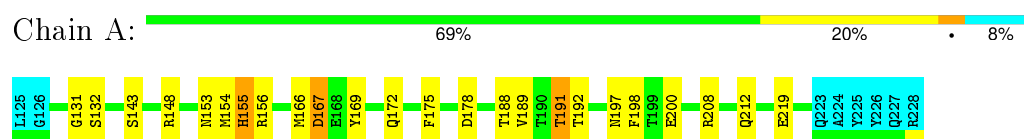
### 4.2.4 Score per residue for model 4

- Molecule 1: PRION PROTEIN



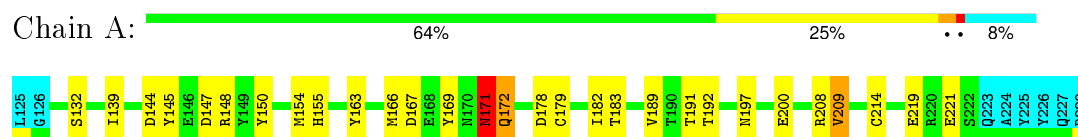
### 4.2.5 Score per residue for model 5 (medoid)

- Molecule 1: PRION PROTEIN



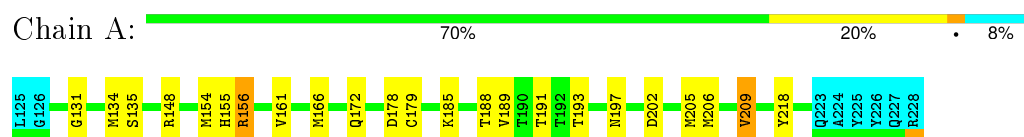
### 4.2.6 Score per residue for model 6

- Molecule 1: PRION PROTEIN



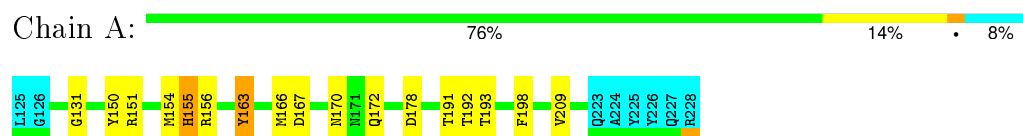
### 4.2.7 Score per residue for model 7

- Molecule 1: PRION PROTEIN



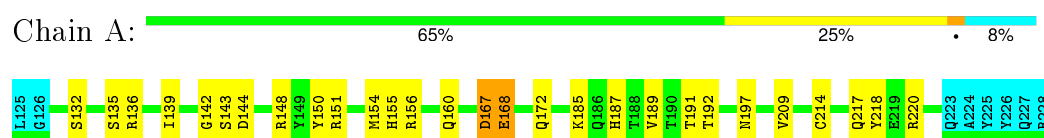
#### 4.2.8 Score per residue for model 8

- Molecule 1: PRION PROTEIN



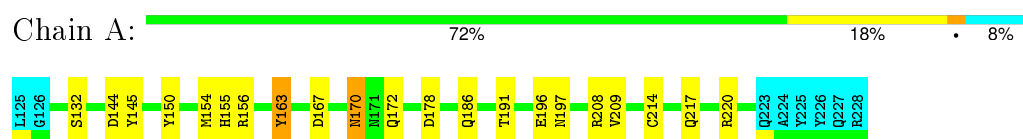
#### 4.2.9 Score per residue for model 9

- Molecule 1: PRION PROTEIN



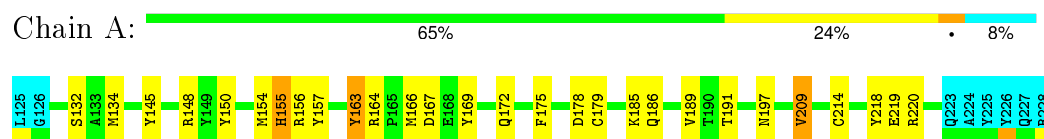
#### 4.2.10 Score per residue for model 10

- Molecule 1: PRION PROTEIN



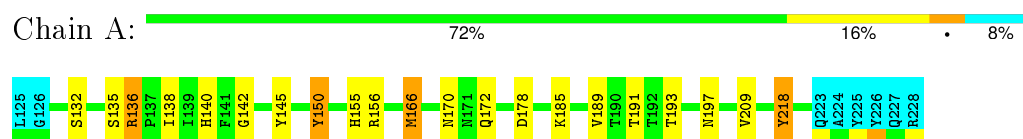
#### 4.2.11 Score per residue for model 11

- Molecule 1: PRION PROTEIN



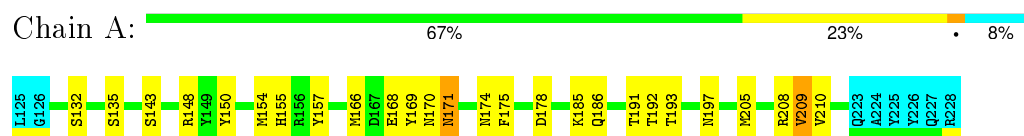
#### 4.2.12 Score per residue for model 12

- Molecule 1: PRION PROTEIN



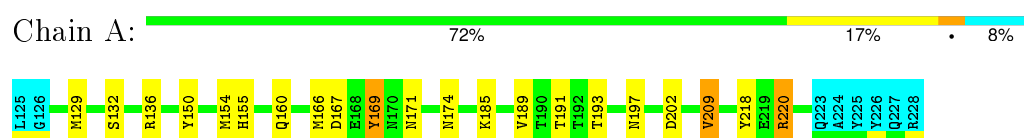
### 4.2.13 Score per residue for model 13

- Molecule 1: PRION PROTEIN



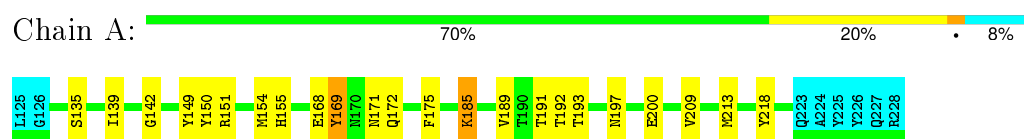
### 4.2.14 Score per residue for model 14

- Molecule 1: PRION PROTEIN



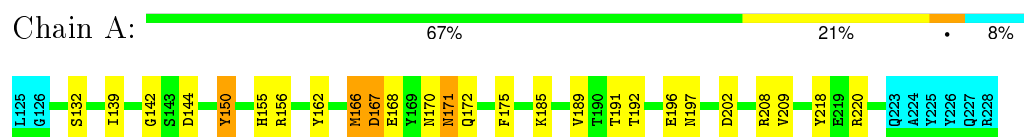
### 4.2.15 Score per residue for model 15

- Molecule 1: PRION PROTEIN



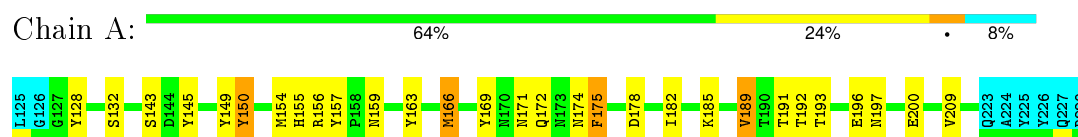
### 4.2.16 Score per residue for model 16

- Molecule 1: PRION PROTEIN



### 4.2.17 Score per residue for model 17

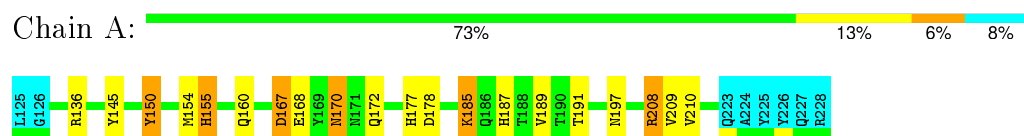
- Molecule 1: PRION PROTEIN





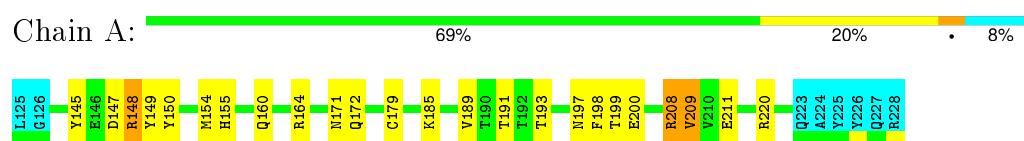
#### 4.2.18 Score per residue for model 18

- Molecule 1: PRION PROTEIN



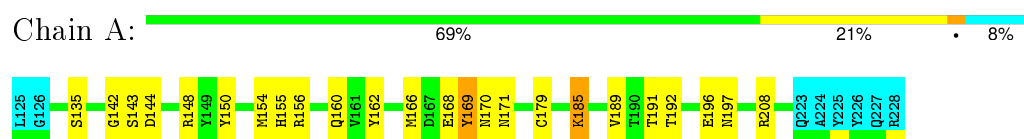
#### 4.2.19 Score per residue for model 19

- 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. 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±0/828 (0.0±0.0%)	1.07±0.03	1±1/1119 (0.1±0.1%)
All	All	0.67	0/16560 (0.0%)	1.07	23/22380 (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±1.2
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	157	TYR	CB-CG-CD2	-7.04	116.78	121.00	17	2
1	A	220	ARG	NE-CZ-NH2	-6.89	116.85	120.30	19	2
1	A	218	TYR	CB-CG-CD2	-6.12	117.33	121.00	11	1
1	A	209	VAL	CA-CB-CG1	5.99	119.88	110.90	19	5
1	A	210	VAL	CG1-CB-CG2	-5.85	101.54	110.90	18	2
1	A	148	ARG	NE-CZ-NH2	-5.63	117.48	120.30	5	1
1	A	163	TYR	CB-CG-CD2	-5.56	117.66	121.00	8	3
1	A	150	TYR	CB-CG-CD1	-5.38	117.77	121.00	4	1
1	A	151	ARG	NE-CZ-NH2	-5.36	117.62	120.30	15	1
1	A	136	ARG	NE-CZ-NH2	-5.30	117.65	120.30	18	2
1	A	208	ARG	NE-CZ-NH1	5.20	122.90	120.30	19	1
1	A	214	CYS	CA-CB-SG	-5.12	104.78	114.00	11	1
1	A	150	TYR	CB-CG-CD2	-5.12	117.93	121.00	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	145	TYR	Sidechain	7
1	A	148	ARG	Sidechain	6
1	A	208	ARG	Sidechain	6
1	A	169	TYR	Sidechain	5
1	A	175	PHE	Sidechain	4
1	A	220	ARG	Sidechain	3
1	A	149	TYR	Sidechain	3
1	A	136	ARG	Sidechain	3
1	A	150	TYR	Sidechain	3
1	A	162	TYR	Sidechain	2
1	A	164	ARG	Sidechain,Peptide	2
1	A	151	ARG	Sidechain	2
1	A	153	ASN	Peptide	1
1	A	166	MET	Peptide	1
1	A	141	PHE	Sidechain	1
1	A	218	TYR	Sidechain	1
1	A	157	TYR	Sidechain	1
1	A	131	GLY	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	809	749	749	2±1
All	All	16180	14980	14980	34

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:150:TYR:CE1	1:A:209:VAL:HG21	0.61	2.31	13	5
1:A:185:LYS:O	1:A:189:VAL:HG23	0.57	1.98	16	14
1:A:163:TYR:CD2	1:A:175:PHE:CE2	0.52	2.97	1	2

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:171:ASN:ND2	1:A:172:GLN:H	0.50	2.05	6	1
1:A:171:ASN:HD22	1:A:172:GLN:H	0.49	1.50	6	1
1:A:172:GLN:HA	1:A:172:GLN:HE21	0.48	1.68	1	1
1:A:189:VAL:O	1:A:192:THR:HG22	0.46	2.10	5	1
1:A:171:ASN:ND2	1:A:174:ASN:H	0.44	2.11	13	1
1:A:150:TYR:CZ	1:A:209:VAL:HG21	0.44	2.47	1	1
1:A:191:THR:HG21	1:A:198:PHE:CE2	0.44	2.48	5	1
1:A:156:ARG:HD3	1:A:156:ARG:H	0.43	1.73	7	1
1:A:141:PHE:CE1	1:A:205:MET:HG2	0.42	2.49	4	1
1:A:138:ILE:HD12	1:A:138:ILE:N	0.42	2.30	12	1
1:A:155:HIS:CG	1:A:156:ARG:N	0.41	2.88	11	1
1:A:155:HIS:CD2	1:A:156:ARG:HD3	0.41	2.50	1	1
1:A:171:ASN:HD22	1:A:174:ASN:H	0.40	1.60	17	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	96/104 (92%)	83±3 (86±4%)	9±3 (10±3%)	4±1 (4±1%)	6	30
All	All	1920/2080 (92%)	1651 (86%)	183 (10%)	86 (4%)	6	30

All 14 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	155	HIS	20
1	A	154	MET	18
1	A	167	ASP	7
1	A	168	GLU	7
1	A	142	GLY	6
1	A	135	SER	5
1	A	170	ASN	5
1	A	166	MET	5
1	A	171	ASN	5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	169	TYR	2
1	A	198	PHE	2
1	A	131	GLY	2
1	A	196	GLU	1
1	A	128	TYR	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	91/97 (94%)	75±3 (83±3%)	16±3 (17±3%)	6	41
All	All	1820/1940 (94%)	1506 (83%)	314 (17%)	6	41

All 63 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	191	THR	20
1	A	197	ASN	19
1	A	172	GLN	17
1	A	150	TYR	14
1	A	166	MET	13
1	A	132	SER	13
1	A	209	VAL	13
1	A	156	ARG	12
1	A	192	THR	11
1	A	193	THR	11
1	A	178	ASP	11
1	A	200	GLU	7
1	A	171	ASN	7
1	A	144	ASP	7
1	A	167	ASP	7
1	A	175	PHE	6
1	A	160	GLN	6
1	A	218	TYR	6
1	A	170	ASN	6
1	A	185	LYS	5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	169	TYR	5
1	A	143	SER	5
1	A	155	HIS	5
1	A	179	CYS	5
1	A	208	ARG	4
1	A	186	GLN	4
1	A	163	TYR	4
1	A	189	VAL	4
1	A	139	ILE	4
1	A	220	ARG	4
1	A	148	ARG	4
1	A	182	ILE	3
1	A	168	GLU	3
1	A	196	GLU	3
1	A	214	CYS	3
1	A	219	GLU	3
1	A	217	GLN	3
1	A	147	ASP	3
1	A	202	ASP	3
1	A	205	MET	2
1	A	174	ASN	2
1	A	187	HIS	2
1	A	188	THR	2
1	A	135	SER	2
1	A	134	MET	2
1	A	129	MET	2
1	A	222	SER	1
1	A	173	ASN	1
1	A	221	GLU	1
1	A	183	THR	1
1	A	145	TYR	1
1	A	206	MET	1
1	A	177	HIS	1
1	A	211	GLU	1
1	A	161	VAL	1
1	A	212	GLN	1
1	A	159	ASN	1
1	A	149	TYR	1
1	A	199	THR	1
1	A	151	ARG	1
1	A	140	HIS	1
1	A	136	ARG	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	213	MET	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