



# Full wwPDB NMR Structure Validation Report ⓘ

Feb 12, 2017 – 06:13 pm GMT

PDB ID : 1H0L  
Title : HUMAN PRION PROTEIN 121-230 M166C/E221C  
Authors : Zahn, R.; Guntert, P.; Wuthrich, K.  
Deposited on : 2002-06-24

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 : trunk28760  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

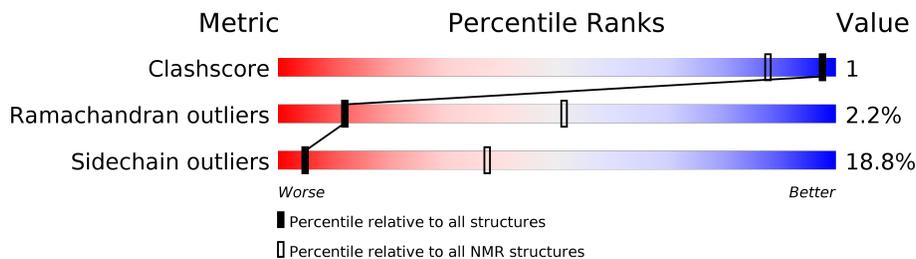
# 1 Overall quality at a glance

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	112	 67% 17% • 15%

## 2 Ensemble composition and analysis i

This entry contains 20 models. Model 6 is the overall representative, medoid model (most similar to other models). The authors have identified model 2 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:128-A:222 (95)	0.36	6

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 5 clusters and 2 single-model clusters were found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1761 atoms, of which 847 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	112	1761	564	847	161	179	10	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	166	CYS	MET	ENGINEERED MUTATION	UNP P04156
A	221	CYS	GLU	ENGINEERED MUTATION	UNP P04156

## 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

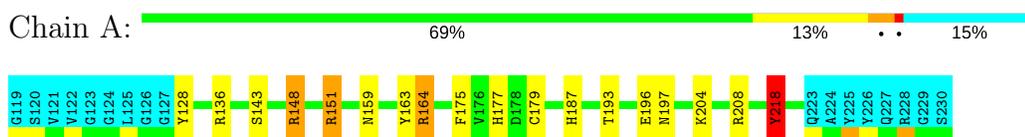


### 4.2 Scores per residue for each member of the ensemble

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



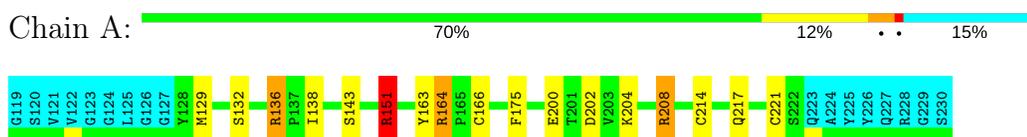
### 4.2.3 Score per residue for model 3

- Molecule 1: MAJOR PRION PROTEIN



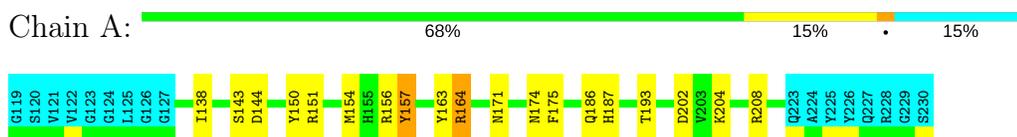
### 4.2.4 Score per residue for model 4

- Molecule 1: MAJOR PRION PROTEIN



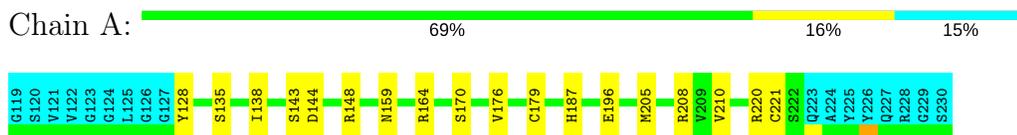
### 4.2.5 Score per residue for model 5

- Molecule 1: MAJOR PRION PROTEIN



### 4.2.6 Score per residue for model 6 (medoid)

- 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

Chain A: 65% 18% 15%



#### 4.2.9 Score per residue for model 9

- Molecule 1: MAJOR PRION PROTEIN

Chain A: 63% 21% 15%



#### 4.2.10 Score per residue for model 10

- Molecule 1: MAJOR PRION PROTEIN

Chain A: 67% 16% 15%



#### 4.2.11 Score per residue for model 11

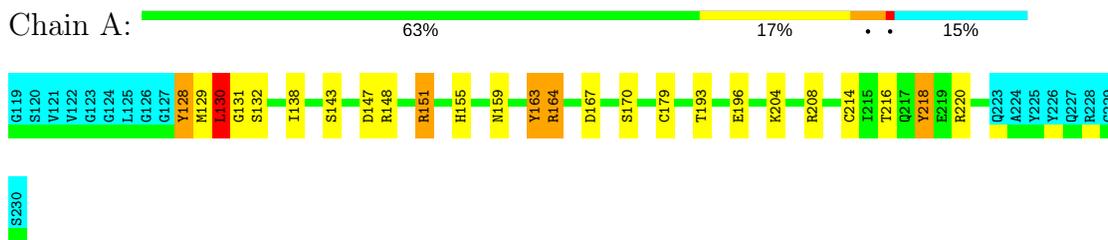
- Molecule 1: MAJOR PRION PROTEIN

Chain A: 64% 20% 15%



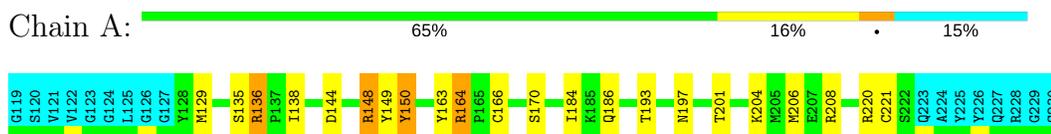
#### 4.2.12 Score per residue for model 12

- Molecule 1: MAJOR PRION PROTEIN



#### 4.2.13 Score per residue for model 13

- Molecule 1: MAJOR PRION PROTEIN



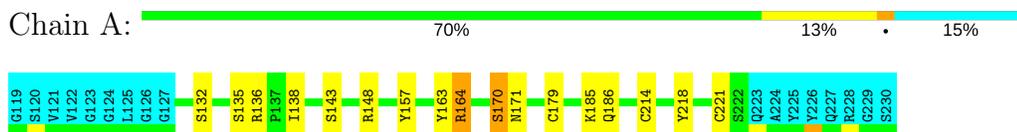
#### 4.2.14 Score per residue for model 14

- Molecule 1: MAJOR PRION PROTEIN



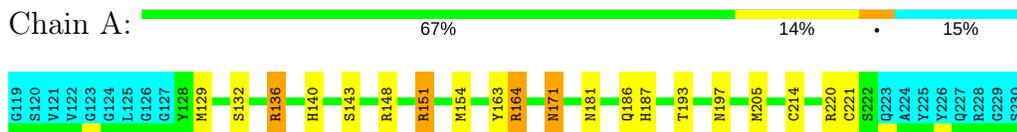
#### 4.2.15 Score per residue for model 15

- Molecule 1: MAJOR PRION PROTEIN



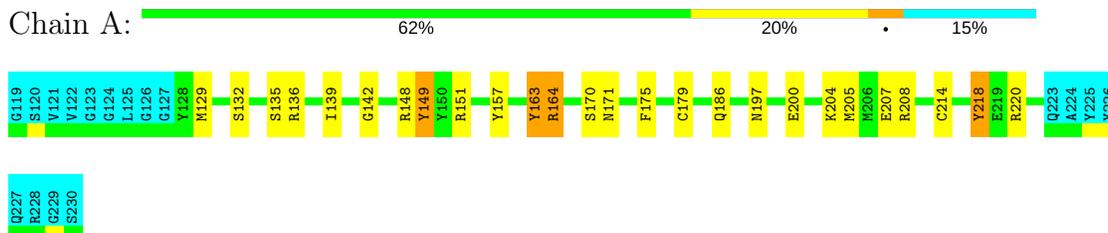
#### 4.2.16 Score per residue for model 16

- Molecule 1: MAJOR PRION PROTEIN



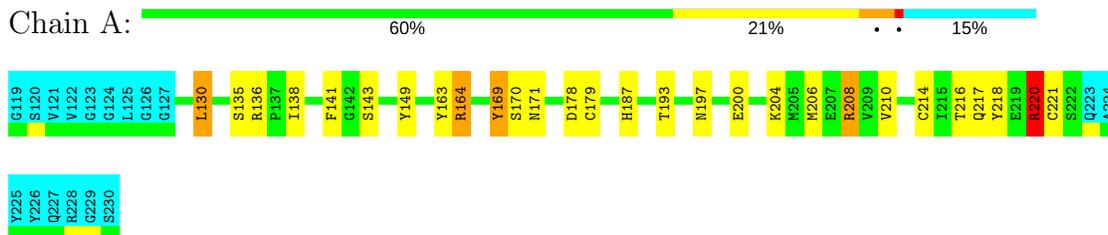
#### 4.2.17 Score per residue for model 17

- 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



5230

## 5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *LEAST RESTRAINT VIOLATION*.

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	0.66±0.01	0±0/817 (0.0±0.0%)	1.09±0.03	2±1/1105 (0.1±0.1%)
All	All	0.66	0/16340 (0.0%)	1.09	32/22100 (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	2.8±1.7
All	All	0	56

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	164	ARG	NE-CZ-NH1	6.70	123.65	120.30	3	3
1	A	169	TYR	CB-CG-CD2	-6.54	117.08	121.00	18	1
1	A	164	ARG	CD-NE-CZ	6.49	132.69	123.60	4	5
1	A	208	ARG	NE-CZ-NH1	6.45	123.52	120.30	7	1
1	A	148	ARG	NE-CZ-NH1	6.08	123.34	120.30	12	1
1	A	136	ARG	NE-CZ-NH2	-5.90	117.35	120.30	8	2
1	A	164	ARG	NE-CZ-NH2	-5.69	117.45	120.30	12	2
1	A	130	LEU	CB-CA-C	5.68	120.99	110.20	12	1
1	A	156	ARG	NE-CZ-NH2	-5.62	117.49	120.30	14	2
1	A	151	ARG	NE-CZ-NH2	-5.60	117.50	120.30	19	3
1	A	163	TYR	CB-CG-CD2	-5.59	117.65	121.00	12	1
1	A	156	ARG	NE-CZ-NH1	5.40	123.00	120.30	14	2
1	A	210	VAL	CG1-CB-CG2	-5.40	102.26	110.90	6	2
1	A	220	ARG	NE-CZ-NH2	-5.40	117.60	120.30	13	1
1	A	129	MET	C-N-CA	5.25	134.82	121.70	20	1

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	150	TYR	CB-CG-CD1	-5.20	117.88	121.00	13	1
1	A	166	CYS	CA-CB-SG	5.17	123.31	114.00	8	1
1	A	149	TYR	CB-CG-CD2	-5.15	117.91	121.00	17	1
1	A	136	ARG	NE-CZ-NH1	5.06	122.83	120.30	13	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	148	ARG	Sidechain	10
1	A	164	ARG	Sidechain	6
1	A	163	TYR	Sidechain	5
1	A	208	ARG	Sidechain	5
1	A	218	TYR	Sidechain	4
1	A	150	TYR	Sidechain	3
1	A	175	PHE	Sidechain	3
1	A	149	TYR	Sidechain	3
1	A	151	ARG	Sidechain	3
1	A	220	ARG	Sidechain	2
1	A	156	ARG	Sidechain	2
1	A	128	TYR	Sidechain	2
1	A	136	ARG	Sidechain	2
1	A	169	TYR	Sidechain	2
1	A	157	TYR	Sidechain	1
1	A	142	GLY	Peptide	1
1	A	141	PHE	Sidechain	1
1	A	162	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	798	738	738	1±1
All	All	15960	14760	14760	18

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:130:LEU:HD12	1:A:131:GLY:H	0.58	1.58	12	1
1:A:130:LEU:HD12	1:A:131:GLY:N	0.52	2.20	12	1
1:A:206:MET:O	1:A:210:VAL:HG23	0.48	2.08	18	4
1:A:216:THR:HG22	1:A:220:ARG:HD2	0.46	1.88	18	1
1:A:217:GLN:NE2	1:A:220:ARG:HH21	0.46	2.09	9	1
1:A:175:PHE:CE2	1:A:218:TYR:CD1	0.46	3.04	1	2
1:A:216:THR:CG2	1:A:220:ARG:HE	0.45	2.24	12	1
1:A:176:VAL:HA	1:A:179:CYS:SG	0.43	2.54	6	1
1:A:154:MET:HA	1:A:157:TYR:CD1	0.42	2.50	5	1
1:A:136:ARG:CD	1:A:136:ARG:H	0.41	2.27	4	1
1:A:139:ILE:CD1	1:A:208:ARG:HB3	0.41	2.45	17	1
1:A:164:ARG:HH22	1:A:170:SER:H	0.41	1.58	11	1
1:A:184:ILE:HG23	1:A:206:MET:SD	0.41	2.56	13	1
1:A:151:ARG:HA	1:A:154:MET:SD	0.40	2.56	16	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	95/112 (85%)	86±2 (91±2%)	7±2 (7±2%)	2±1 (2±1%)	12	51
All	All	1900/2240 (85%)	1724 (91%)	134 (7%)	42 (2%)	12	51

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	138	ILE	15
1	A	170	SER	6
1	A	169	TYR	5
1	A	128	TYR	4
1	A	171	ASN	4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	130	LEU	2
1	A	132	SER	2
1	A	166	CYS	2
1	A	221	CYS	1
1	A	198	PHE	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	91/101 (90%)	74±4 (81±4%)	17±4 (19±4%)	<b>5</b> 37
All	All	1820/2020 (90%)	1478 (81%)	342 (19%)	<b>5</b> 37

All 52 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	164	ARG	20
1	A	143	SER	17
1	A	204	LYS	15
1	A	193	THR	14
1	A	214	CYS	13
1	A	208	ARG	12
1	A	163	TYR	12
1	A	132	SER	12
1	A	187	HIS	11
1	A	151	ARG	11
1	A	221	CYS	10
1	A	179	CYS	10
1	A	186	GLN	10
1	A	171	ASN	10
1	A	136	ARG	10
1	A	135	SER	9
1	A	218	TYR	9
1	A	159	ASN	8
1	A	197	ASN	8
1	A	157	TYR	7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	167	ASP	7
1	A	217	GLN	7
1	A	200	GLU	7
1	A	202	ASP	7
1	A	148	ARG	6
1	A	205	MET	6
1	A	177	HIS	6
1	A	170	SER	6
1	A	129	MET	6
1	A	144	ASP	6
1	A	220	ARG	4
1	A	196	GLU	4
1	A	175	PHE	4
1	A	134	MET	4
1	A	178	ASP	3
1	A	207	GLU	3
1	A	174	ASN	3
1	A	172	GLN	3
1	A	201	THR	3
1	A	130	LEU	2
1	A	168	GLU	2
1	A	140	HIS	2
1	A	155	HIS	2
1	A	147	ASP	2
1	A	169	TYR	2
1	A	149	TYR	1
1	A	181	ASN	1
1	A	185	LYS	1
1	A	199	THR	1
1	A	222	SER	1
1	A	156	ARG	1
1	A	138	ILE	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