W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

Full wwPDB NMR Structure Validation Report (i)

May 28, 2020 – 07:25 pm BST

PDB ID	:	1DWZ
Title	:	Bovine prion protein fragment 121-230
Authors	:	Lopez-Garcia, F.; Zahn, R.; Riek, R.; Billeter, M.; Wuthrich, K.
Deposited on	:	1999-12-15

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

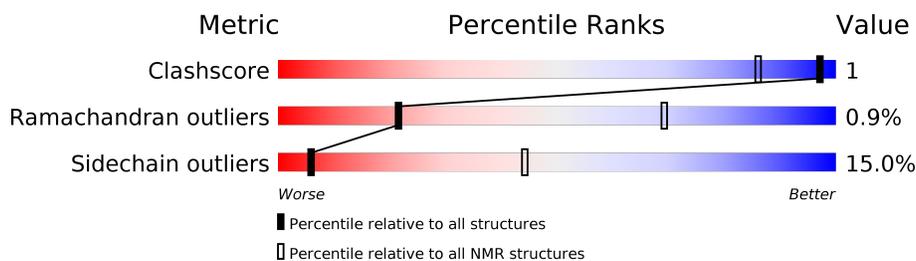|		
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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	Percentile Ranks	Value
Clashscore		1
Ramachandran outliers		0.9%
Sidechain outliers		15.0%

Worse   *Better*

■ Percentile relative to all structures

▯ Percentile relative to all NMR structures

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	112	79% 10% · 7%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 9 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:128-A:227 (100)	0.46	9

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 2 clusters and 1 single-model cluster was found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1675 atoms, of which 806 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called PRION PROTEIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	104	1675	540	806	151	170	8	0

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: 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

Chain A:

4.2.4 Score per residue for model 4

• Molecule 1: PRION PROTEIN

Chain A:

4.2.5 Score per residue for model 5

• Molecule 1: PRION PROTEIN

Chain A:

4.2.6 Score per residue for model 6

• Molecule 1: PRION PROTEIN

Chain A:

4.2.7 Score per residue for model 7

• Molecule 1: PRION PROTEIN

Chain A:

GLY SER VAL GLY G124 L125 G126 G127 Y128 Y149 Y150 R151 H155 R156 Y157 R164 Q172 F175 D178 C179 T190 T199 E200 T201 K204 I215 Q219 Q227 ARG GLY ALA

4.2.8   Score per residue for model 8

• Molecule 1: PRION PROTEIN

Chain A: 75%   13%   •   •   7%

GLY SER VAL GLY G124 L125 G126 G127 L130 Y149 Y150 Y157 R164 Q168 F175 V176 H177 K185 E186 T190 T193 K204 E211 R220 Y226 Q227 ARG GLY ALA

4.2.9   Score per residue for model 9 (medoid)

• Molecule 1: PRION PROTEIN

Chain A: 75%   12%   •   •   7%

GLY SER VAL GLY G124 L125 G126 G127 H140 S143 Y149 Y150 R151 Y163 V166 C179 T191 T199 K204 M205 V209 V210 E211 Y218 Y226 Q227 ARG GLY ALA

4.2.10   Score per residue for model 10

• Molecule 1: PRION PROTEIN

Chain A: 72%   14%   •   •   •   7%

GLY SER VAL GLY G124 L125 G126 G127 Y128 L138 D144 D147 R148 Y149 N153 M154 R164 Q168 Y169 S170 N171 F175 T191 T192 T199 M205 M206 Y218 Y226 Q227 ARG GLY ALA

4.2.11   Score per residue for model 11

• Molecule 1: PRION PROTEIN

Chain A: 73%   15%   •   •   7%

GLY SER VAL GLY G124 L125 G126 G127 L130 G131 S132 R136 P137 L138 R148 Y149 Y157 R164 Y169 F175 V176 H177 T190 T191 K204 C214 R220 Q223 Q227 ARG GLY ALA

4.2.12   Score per residue for model 12

• Molecule 1: PRION PROTEIN

Chain A: 69%   18%   •   •   •   7%

GLY SER VAL GLY G124 L125 G126 G127 Y128 M129 L130 R136 D144 Y145 Y149 Y150 R151 R156 Y157 Q160 R164 F175 C179 T190 T191 N197 K204 R208 V209 Y218 Q219 R220 Y225 Y226 Q227 ARG GLY ALA

4.2.13 Score per residue for model 13

• Molecule 1: PRION PROTEIN

Chain A: 73% 14% • • 7%

GLY SER VAL GLY G124 L125 G126 G127 S132 L138 R148 R151 R156 Y157 D167 Q168 N171 Q172 F175 H187 T190 T193 V209 Y218 Q219 R220 Y226 Q227 ARG GLY ALA

4.2.14 Score per residue for model 14

• Molecule 1: PRION PROTEIN

Chain A: 71% 17% • • 7%

GLY SER VAL GLY G124 L125 G126 G127 Y128 M129 L130 G131 S132 R136 P137 L138 E146 D147 R148 Y149 Y150 R151 R156 Y157 V166 Y169 S170 H177 V189 T199 K204 M205 C214 Q223 Q227 ARG GLY ALA

4.2.15 Score per residue for model 15

• Molecule 1: PRION PROTEIN

Chain A: 72% 14% • • 7%

GLY SER VAL GLY G124 L125 G126 G127 Y128 S132 L138 I139 H140 D144 Y145 E146 D147 R148 Y149 Y157 F175 I182 H187 T190 K204 M205 V209 C214 Q217 Y225 Y226 Q227 ARG GLY ALA
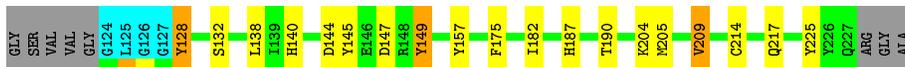
4.2.16 Score per residue for model 16

• Molecule 1: PRION PROTEIN

Chain A: 75% 13% • • 7%

GLY SER VAL GLY G124 L125 G126 G127 L130 R136 P137 L138 D144 Y149 M154 Y157 R164 S170 N173 N174 K185 T190 T199 S222 Q227 ARG GLY ALA

4.2.17 Score per residue for model 17

• Molecule 1: PRION PROTEIN

Chain A: 71% 16% • • • 7%

GLY SER VAL GLY G124 L125 G126 G127 L138 I139 Y150 R151 E152 N153 M154 Q160 Q168 N171 F175 D178 K185 T191 F198 T201 M205 V209 I215 Q219 Q227 ARG GLY ALA

4.2.18 Score per residue for model 18

• Molecule 1: PRION PROTEIN

Chain A: 72%   16% • • 7%

GLY SER VAL GLY G124 L125 G126 G127 Y128 M129 L130 L138 D144 Y149 Y150 R151 H155 R156 Y157 R164 Q168 F175 C179 T190 K204 M205 E211 Y218 Y225 Y226 Q227 ARG GLY ALA

4.2.19 Score per residue for model 19

• Molecule 1: PRION PROTEIN

Chain A: 73%   14% • • 7%

GLY SER VAL GLY G124 L125 G126 G127 Y128 S135 L138 R148 Y149 Y150 R151 Y157 R164 N171 Q172 F175 V176 H177 D178 C179 T191 K204 Y225 Y226 Q227 ARG GLY ALA

4.2.20 Score per residue for model 20

• Molecule 1: PRION PROTEIN

Chain A: 74%   15% • 7%

GLY SER VAL GLY G124 L125 G126 G127 L130 L138 D144 R148 R151 Y157 Q160 Y169 Q172 F175 H187 T188 V189 T199 K204 Y218 Y225 Y226 Q227 ARG GLY ALA

5   Refinement protocol and experimental data overview (i)

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

Of the 50 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.

COVALENT-GEOMETRY INFOmissingINFO

5.1   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	849	786	786	1 ± 1
All	All	16980	15720	15720	19

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:138:LEU:HD11	1:A:151:ARG:HE	0.72	1.43	14	1
1:A:150:TYR:OH	1:A:209:VAL:HG21	0.57	2.00	17	1
1:A:130:LEU:HD11	1:A:160:GLN:HB3	0.56	1.77	20	2
1:A:201:THR:HG22	1:A:205:MET:SD	0.54	2.43	17	1
1:A:166:VAL:HG12	1:A:218:TYR:CE1	0.53	2.39	9	1
1:A:175:PHE:CE2	1:A:218:TYR:HB2	0.52	2.38	18	2
1:A:138:LEU:HD11	1:A:151:ARG:NE	0.51	2.17	14	1
1:A:215:ILE:HG22	1:A:219:GLN:HE21	0.48	1.67	17	1
1:A:140:HIS:CD2	1:A:140:HIS:H	0.45	2.27	6	1
1:A:175:PHE:CZ	1:A:218:TYR:HB2	0.45	2.46	10	3

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:215:ILE:HG23	1:A:219:GLN:NE2	0.44	2.28	7	1
1:A:191:THR:HG21	1:A:198:PHE:CZ	0.43	2.48	17	1
1:A:130:LEU:HD23	1:A:162:TYR:CE1	0.43	2.48	3	1
1:A:175:PHE:CE1	1:A:218:TYR:HB2	0.42	2.49	13	1
1:A:198:PHE:CD2	1:A:202:ASP:HB3	0.40	2.52	6	1

5.2 Torsion angles (i)

5.2.1 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	99/112 (88%)	91±2 (92±2%)	7±2 (7±2%)	1±1 (1±1%)	21	69
All	All	1980/2240 (88%)	1816 (92%)	146 (7%)	18 (1%)	21	69

All 12 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	128	TYR	3
1	A	171	ASN	3
1	A	154	MET	2
1	A	167	ASP	2
1	A	166	VAL	1
1	A	169	TYR	1
1	A	172	GLN	1
1	A	131	GLY	1
1	A	220	ARG	1
1	A	170	SER	1
1	A	168	GLN	1
1	A	132	SER	1

5.2.2 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	95/100 (95%)	81±2 (85±3%)	14±2 (15±3%)	6	44
All	All	1900/2000 (95%)	1615 (85%)	285 (15%)	6	44

All 66 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	175	PHE	15
1	A	138	LEU	14
1	A	204	LYS	13
1	A	190	THR	13
1	A	149	TYR	13
1	A	164	ARG	11
1	A	227	GLN	10
1	A	144	ASP	10
1	A	130	LEU	8
1	A	199	THR	7
1	A	168	GLN	7
1	A	225	TYR	7
1	A	191	THR	7
1	A	205	MET	7
1	A	211	GLU	7
1	A	214	CYS	7
1	A	226	TYR	6
1	A	179	CYS	6
1	A	148	ARG	6
1	A	169	TYR	6
1	A	132	SER	5
1	A	178	ASP	5
1	A	172	GLN	5
1	A	185	LYS	4
1	A	186	GLU	4
1	A	177	HIS	4
1	A	140	HIS	4
1	A	136	ARG	4
1	A	209	VAL	4
1	A	147	ASP	3
1	A	187	HIS	3
1	A	146	GLU	3
1	A	171	ASN	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	193	THR	3
1	A	154	MET	3
1	A	222	SER	3
1	A	223	GLN	3
1	A	173	ASN	3
1	A	189	VAL	2
1	A	170	SER	2
1	A	192	THR	2
1	A	153	ASN	2
1	A	151	ARG	2
1	A	206	MET	2
1	A	155	HIS	2
1	A	217	GLN	2
1	A	145	TYR	2
1	A	197	ASN	2
1	A	150	TYR	2
1	A	201	THR	1
1	A	167	ASP	1
1	A	156	ARG	1
1	A	128	TYR	1
1	A	196	GLU	1
1	A	220	ARG	1
1	A	157	TYR	1
1	A	166	VAL	1
1	A	159	ASN	1
1	A	143	SER	1
1	A	207	GLU	1
1	A	139	ILE	1
1	A	135	SER	1
1	A	174	ASN	1
1	A	160	GLN	1
1	A	213	MET	1
1	A	129	MET	1

5.2.3 RNA (i)

There are no RNA molecules in this entry.

5.3 Non-standard residues in protein, DNA, RNA chains (i)

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

5.4 Carbohydrates (i)

There are no carbohydrates in this entry.

5.5 Ligand geometry (i)

There are no ligands in this entry.

5.6 Other polymers (i)

There are no such molecules in this entry.

5.7 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Chemical shift validation (i)

No chemical shift data were provided