



Full wwPDB NMR Structure Validation Report ⓘ

Mar 7, 2022 – 02:26 AM EST

PDB ID : 2KKG
Title : NMR structure of the octarepeat region of prion protein bound to pentosan polysulfate
Authors : Taubner, L.M.; Caughey, B.; Copie, V.
Deposited on : 2009-06-19

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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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)
RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
ShiftChecker : 2.27
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.27

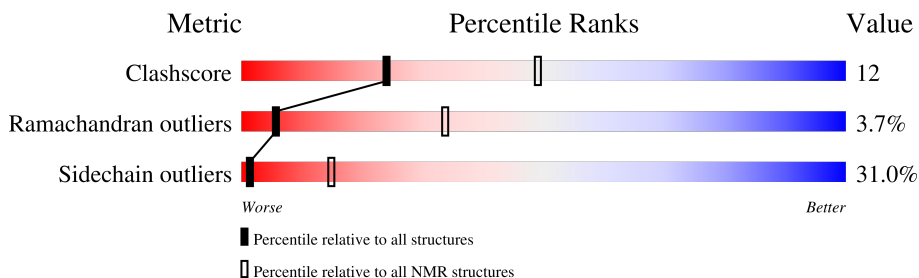
1 Overall quality at a glance

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	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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	85	

2 Ensemble composition and analysis

This entry contains 20 models. Model 19 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 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:72-A:91 (20)	0.49	19

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 11 single-model clusters were found.

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

3 Entry composition

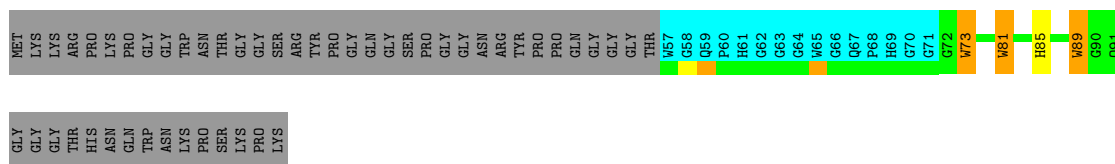
There is only 1 type of molecule in this entry. The entry contains 452 atoms, of which 201 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Major prion protein.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
1	A	35	452	158	201	53	40	0

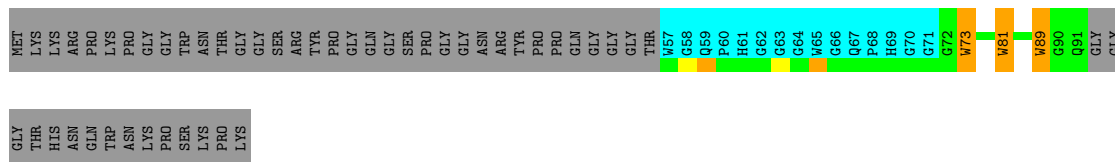
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	MET	-	initiating methionine	UNP P04273



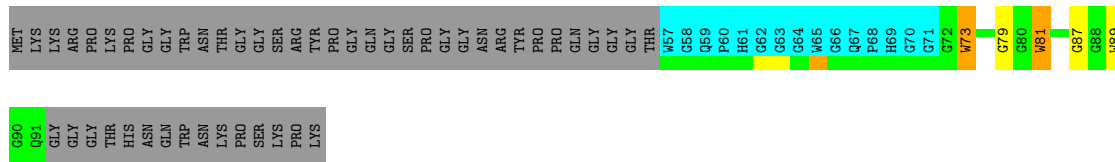
4.2.3 Score per residue for model 3

- Molecule 1: Major prion protein



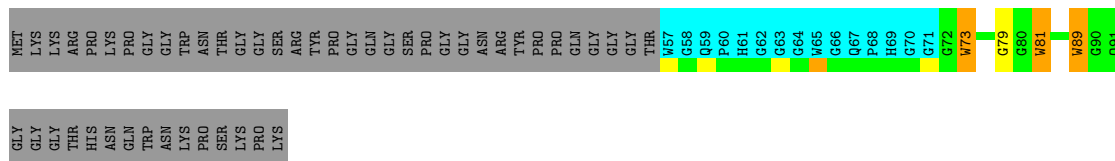
4.2.4 Score per residue for model 4

- 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

- Molecule 1: Major prion protein

Chain A: 19% 18% 59%

MET LYS LYS ARG ARG PRO LYS PRO GLY GLY TRP SER ASN THR THR GLY SER ARG ARG PRO TYR PRO GLY GLN GLY SER PRO PRO PRO PRO GLN GLY GLY THR W57 W58 Q59 P60 H61 G62 G63 G64 G65 G66 Q67 P68 H69 G70 G71 G72 W73 W81 W89 G90 Q91 GLY GLY

GLY THR HIS ASN GLN TRP ASH LYS PRO PRO SER SER LYS PRO LYS

4.2.7 Score per residue for model 7

- Molecule 1: Major prion protein

Chain A: 15% 5% 18% 59%

MET LYS LYS ARG ARG PRO LYS PRO GLY GLY TRP SER ASN THR THR GLY SER ARG ARG PRO TYR PRO GLY GLN GLY SER PRO PRO PRO PRO GLN GLY GLY THR W57 W58 Q59 P60 H61 G62 G63 G64 G65 G66 Q67 P68 H69 G70 G71 G72 W73 H77 G78 G79 G80 W81 G87

G88 W89 G90 Q91 GLY GLY THR HIS ASN GLN TRP ASH LYS PRO SER ARG ARG PRO TYR PRO GLY GLN GLY SER PRO PRO PRO LYS

4.2.8 Score per residue for model 8

- Molecule 1: Major prion protein

Chain A: 20% 18% 59%

MET LYS LYS ARG ARG PRO LYS PRO GLY GLY TRP SER ASN THR THR GLY SER ARG ARG PRO TYR PRO GLY GLN GLY SER PRO PRO PRO PRO GLN GLY GLY THR W57 W58 Q59 P60 H61 G62 G63 G64 G65 G66 Q67 P68 H69 G70 G71 G72 W73 W81 W89 G90 Q91 GLY GLY

GLY THR HIS ASN GLN TRP ASH LYS PRO PRO SER SER LYS PRO LYS

4.2.9 Score per residue for model 9

- Molecule 1: Major prion protein

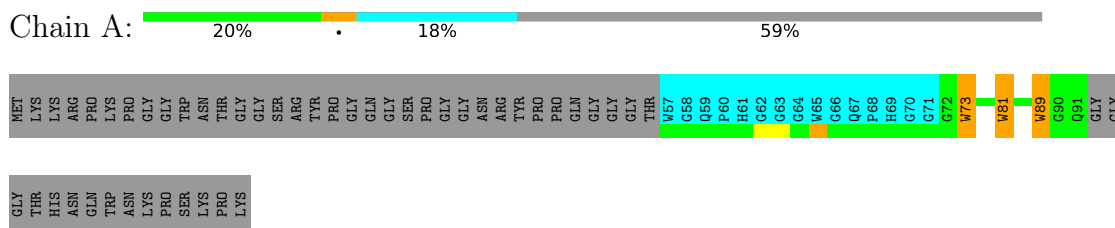
Chain A: 18% 18% 59%

MET LYS LYS ARG ARG PRO LYS PRO GLY GLY TRP SER ASN THR THR GLY SER ARG ARG PRO TYR PRO GLY GLN GLY SER PRO PRO PRO PRO GLN GLY GLY THR W57 W58 Q59 P60 H61 G62 G63 G64 G65 G66 Q67 P68 H69 G70 G71 G72 W73 W81 W89 G85 G86 G87 G88 W89

G90 Q91 GLY GLY THR HIS ASN GLN TRP ASH LYS PRO PRO SER SER LYS PRO LYS

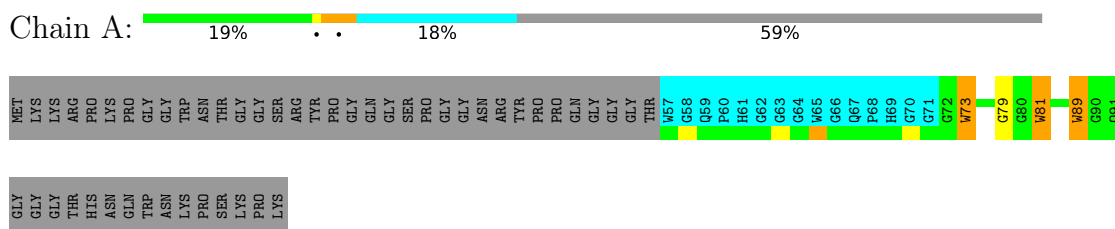
4.2.14 Score per residue for model 14

- Molecule 1: Major prion protein



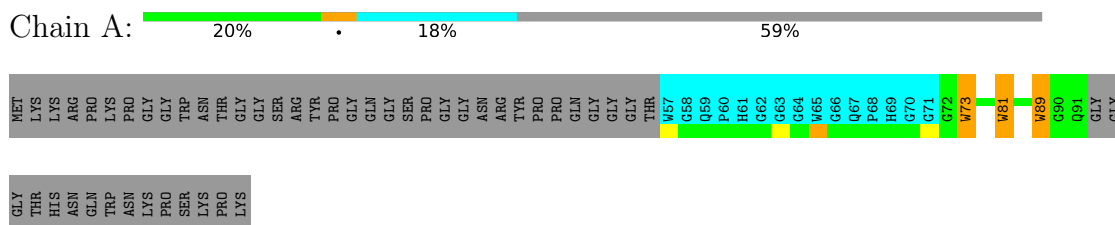
4.2.15 Score per residue for model 15

- Molecule 1: Major prion protein



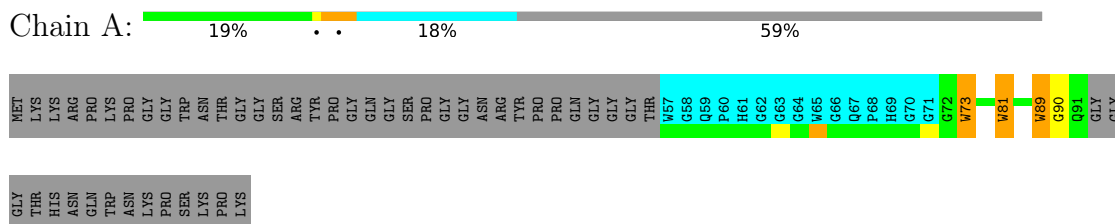
4.2.16 Score per residue for model 16

- Molecule 1: Major prion protein



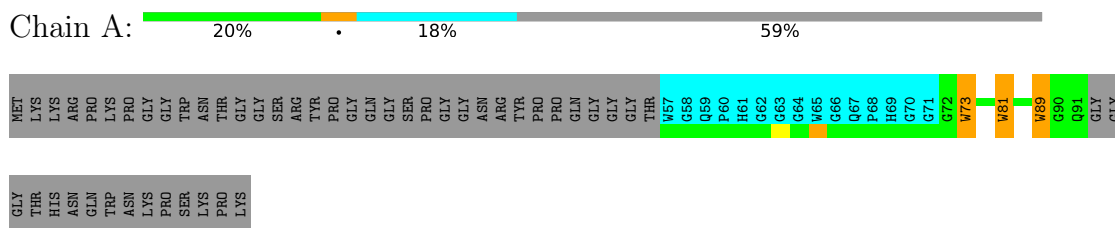
4.2.17 Score per residue for model 17

- Molecule 1: Major prion protein



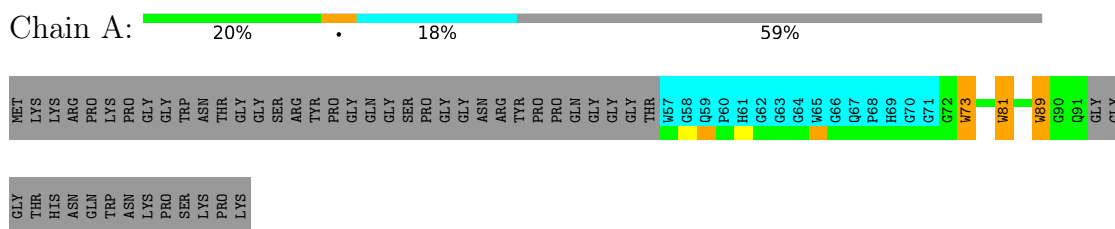
4.2.18 Score per residue for model 18

- Molecule 1: Major prion protein



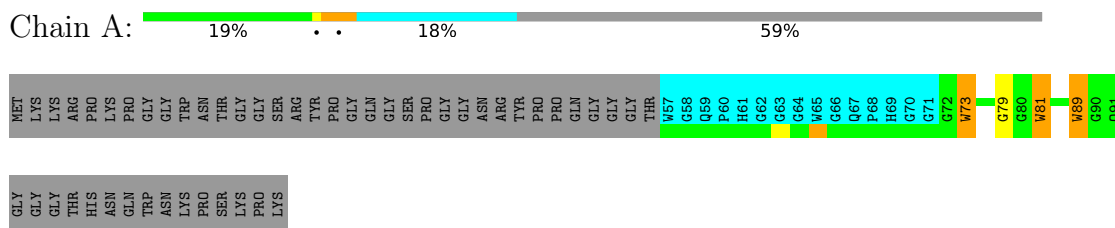
4.2.19 Score per residue for model 19 (medoid)

- Molecule 1: Major prion protein



4.2.20 Score per residue for model 20

- Molecule 1: Major prion protein



5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing, molecular dynamics*.

Of the 40 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
ProcheckNMR	geometry optimization	
CNS	structure solution	
CNS	refinement	

No chemical shift data was provided.

6 Model quality [i](#)

6.1 Standard geometry [i](#)

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	143	114	112	3±0
All	All	2860	2280	2240	61

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 12.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:73:TRP:N	1:A:73:TRP:CD1	0.60	2.69	14	14
1:A:81:TRP:CD1	1:A:81:TRP:N	0.55	2.75	5	20
1:A:89:TRP:CD1	1:A:89:TRP:N	0.52	2.77	17	19
1:A:73:TRP:CD1	1:A:73:TRP:N	0.50	2.80	3	6
1:A:85:HIS:CD2	1:A:85:HIS:N	0.40	2.89	9	1
1:A:77:HIS:CD2	1:A:77:HIS:O	0.40	2.75	7	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	19/85 (22%)	13±1 (69±5%)	5±1 (27±5%)	1±1 (4±4%)	6	34
All	All	380/1700 (22%)	262 (69%)	104 (27%)	14 (4%)	6	34

All 4 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	79	GLY	8
1	A	87	GLY	4
1	A	78	GLY	1
1	A	90	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	10/54 (19%)	7±0 (69±3%)	3±0 (31±3%)	1	15
All	All	200/1080 (19%)	138 (69%)	62 (31%)	1	15

All 5 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	73	TRP	20
1	A	81	TRP	20
1	A	89	TRP	20
1	A	85	HIS	1
1	A	91	GLN	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

No chemical shift data were provided