wwPDB NMR Structure Validation Summary Report (i)

Feb 7, 2022 – 06:52 PM EST

PDB ID	:	1B10
Title	:	SOLUTION NMR STRUCTURE OF RECOMBINANT SYRIAN HAMSTER PRION PROTEIN RPRP(90-231) , 25 STRUCTURES
Authors	:	James, T.L.; Liu, H.; Ulyanov, N.B.; Farr-Jones, S.
Deposited on	:	1998-11-25

This is a wwPDB NMR Structure Validation Summary 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 following versions of software and data (see references (i)) were used in the production of this report:

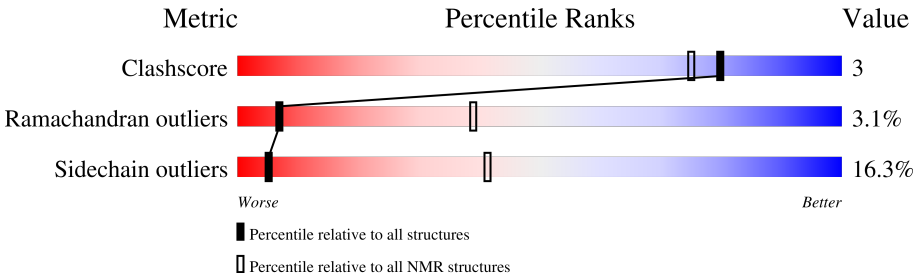
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

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	142	52% 18% • • 27%

2 Ensemble composition and analysis (i)

This entry contains 25 models. Model 25 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:127-A:226 (100)	0.46	25

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, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25
2	2, 17
Single-model clusters	8

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1675 atoms, of which 805 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called PROTEIN (PRION PROTEIN).

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	104	1675	539	805	151	171	9	0

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: PROTEIN (PRION PROTEIN)

Chain A: 52% 18% • • 27%

GLY GLN GLY GLY GLY THR HIS ASN GLN TRP ASN LYS PRO SER LYS PRO LYS THR ASN MET LYS HIS MET ALA GLY ALA ALA ALA ALA GLY ALA VAL VAL GLY GLY L125 G126 G127 Y128 G131 S132 R136 N143 D147 R148 Y149 Y150 R151 R156 Y157 P158 N159 Q160 V161

R164 P165 V166 D167 Q168 Y169 V180 N181 I182 T201 M206 E207 R208 V209 V210 T215 D227 G228 ARG ARG SER

4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 25. Colouring as in section 4.1 above.

• Molecule 1: PROTEIN (PRION PROTEIN)

Chain A: 47% 22% •• • 27%

GLY GLN GLY GLY GLY THR HIS ASN GLN TRP ASN LYS PRO SER LYS PRO LYS THR ASN MET LYS HIS MET ALA GLY ALA ALA ALA ALA GLY ALA VAL VAL GLY GLY L125 G126 G127 Y128 M129 M134 S135 R136 P137 H140 N143 D147 R148 Y149 Y150 R151 M154 N155 R156 Y157

Q160 V161 R164 P165 V166 D167 Q168 Y169 N170 D178 C179 V180 N181 I182 T183 T191 T201 M206 E207 R208 V209 V210 T215 Y226 D227 G228 ARG ARG SER

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *NMR-RESTRAINED SIMULATED ANNEALING, MOLECULAR DYNAMICS, AND ENERGY MINIMIZATION.*.

Of the 100 calculated structures, 25 were deposited, based on the following criterion: *TOTAL ENERGY*.

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

Software name	Classification	Version
ARIA	refinement	
XPLOR	refinement	
AMBER4.1	refinement	
ARIA	structure solution	
XPLOR	structure solution	
Amber	structure solution	

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	0.70±0.01	0±0/867 (0.0± 0.0%)	1.28±0.04	8±2/1174 (0.7± 0.2%)
All	All	0.70	0/21675 (0.0%)	1.28	207/29350 (0.7%)

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	2.2±1.4
All	All	0	55

There are no bond-length outliers.

5 of 42 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	164	ARG	NE-CZ-NH1	10.31	125.45	120.30	4	16
1	A	151	ARG	NE-CZ-NH1	10.15	125.38	120.30	23	14
1	A	156	ARG	NE-CZ-NH1	9.73	125.17	120.30	16	14
1	A	136	ARG	NE-CZ-NH2	-9.45	115.58	120.30	13	6
1	A	151	ARG	NE-CZ-NH2	-9.38	115.61	120.30	1	4

There are no chirality outliers.

5 of 16 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	218	TYR	Sidechain	7
1	A	169	TYR	Sidechain	7
1	A	226	TYR	Sidechain	5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	225	TYR	Sidechain	5
1	A	208	ARG	Sidechain	4

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	846	784	781	4±2
All	All	21150	19600	19525	108

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 3.

5 of 44 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:162:TYR:CD2	1:A:182:ILE:HG22	0.79	2.13	9	2
1:A:180:VAL:HG23	1:A:210:VAL:HG13	0.72	1.59	23	1
1:A:128:TYR:CD2	1:A:182:ILE:HG21	0.67	2.25	2	13
1:A:206:MET:O	1:A:210:VAL:HG12	0.63	1.94	1	23
1:A:166:VAL:HG22	1:A:169:TYR:CD2	0.58	2.33	17	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	100/142 (70%)	84±3 (84±3%)	13±3 (13±3%)	3±1 (3±1%)	7	39
All	All	2500/3550 (70%)	2097 (84%)	326 (13%)	77 (3%)	7	39

5 of 15 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	127	GLY	19
1	A	131	GLY	12
1	A	136	ARG	11
1	A	207	GLU	5
1	A	140	HIS	5

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	94/120 (78%)	79±2 (84±2%)	15±2 (16±2%)	5	41
All	All	2350/3000 (78%)	1967 (84%)	383 (16%)	5	41

5 of 57 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	136	ARG	20
1	A	148	ARG	19
1	A	149	TYR	19
1	A	201	THR	17
1	A	215	THR	17

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