wwPDB NMR Structure Validation Summary Report (i)

Feb 23, 2022 – 09:56 AM EST

PDB ID	:	1WYO
Title	:	Solution structure of the CH domain of human microtubule-associated protein RP/EB family member 3
Authors	:	Tomizawa, T.; Kigawa, T.; Koshiba, S.; Inoue, M.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2005-02-15

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:

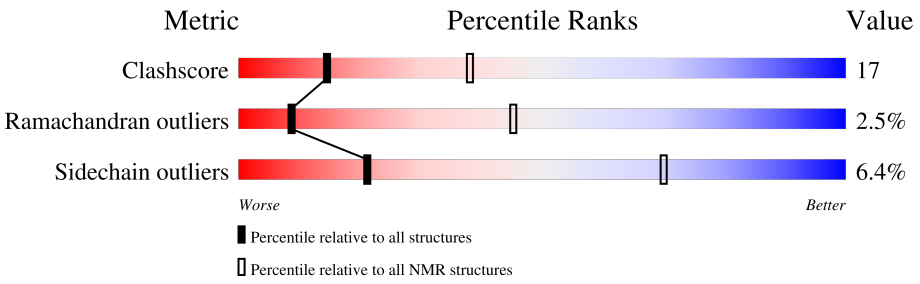
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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	159	44% 26% • 30%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 2 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:9-A:10, A:23-A:132 (112)	0.22	2

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 4 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2490 atoms, of which 1230 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Microtubule-associated protein RP/EB family member 3.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	159	2490	809	1230	212	231	8	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	cloning artifact	UNP Q9UPY8
A	2	SER	-	cloning artifact	UNP Q9UPY8
A	3	SER	-	cloning artifact	UNP Q9UPY8
A	4	GLY	-	cloning artifact	UNP Q9UPY8
A	5	SER	-	cloning artifact	UNP Q9UPY8
A	6	SER	-	cloning artifact	UNP Q9UPY8
A	7	GLY	-	cloning artifact	UNP Q9UPY8
A	154	SER	-	cloning artifact	UNP Q9UPY8
A	155	GLY	-	cloning artifact	UNP Q9UPY8
A	156	PRO	-	cloning artifact	UNP Q9UPY8
A	157	SER	-	cloning artifact	UNP Q9UPY8
A	158	SER	-	cloning artifact	UNP Q9UPY8
A	159	GLY	-	cloning artifact	UNP Q9UPY8

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: Microtubule-associated protein RP/EB family member 3

Chain A: 44% 26% • 30%

G1 S2 S3 G4 S5 S6 G7 M8 A9 V10 N11 V12 Y13 S14 T15 S16 V17 T18 S19 E20 N21 L22 S23 R24 H25 D26 M27 W30 L35 I42 L45 C46 S47 G48 A49 R66 E75 H76 E77 Y78 I79 H80 N81 F82 K90 K91 M92 G93 V94 D95 K96 V100

E101 K102 L103 V104 F108 Q109 F112 I115 Q116 W117 F118 K119 K120 F121 F122 Y126 D127 G128 K129 P133 L134 L135 A136 R137 Q138 G139 Q140 D141 V142 A143 P144 P145 P146 N147 P148 G149 D150 Q151 I152 F153 S154 G155 P156 S157 S158 G159

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

The representative model is number 2. Colouring as in section 4.1 above.

- Molecule 1: Microtubule-associated protein RP/EB family member 3

Chain A: 44% 25% • 30%

G1 S2 S3 G4 S5 S6 G7 M8 A9 V10 N11 V12 Y13 S14 T15 S16 V17 T18 S19 E20 N21 L22 S23 R24 M27 I42 L45 C46 S47 G48 A49 L58 R66 K67 E75 H76 E77 Y78 I79 H80 N81 F82 K90 K91 M92 K96 V100 E101 K102 L103 V104

K107 F108 Q109 D110 N111 F112 I115 Q116 W117 F118 K119 K120 F121 F122 D127 G128 P133 L134 L135 A136 R137 Q138 G139 Q140 D141 V142 A143 P144 P145 P146 N147 P148 G149 D150 Q151 I152 F153 S154 G155 P156 S157 S158 G159

5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function,structures with the lowest energy,structures with the least restraint violations.*

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

Software name	Classification	Version
CYANA	structure solution	2.0.17
CYANA	refinement	2.0.17

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	935	926	921	32±6
All	All	18700	18520	18420	636

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 17.

5 of 181 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:9:ALA:HB2	1:A:92:MET:SD	0.81	2.15	12	12
1:A:122:PHE:CZ	1:A:126:TYR:CD2	0.73	2.76	18	2
1:A:82:PHE:CD2	1:A:100:VAL:HG13	0.71	2.20	20	12
1:A:58:LEU:HD21	1:A:122:PHE:CD1	0.70	2.21	1	3
1:A:108:PHE:CE2	1:A:112:PHE:CZ	0.69	2.81	12	18

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	112/159 (70%)	88±3 (78±3%)	22±3 (19±3%)	3±1 (2±1%)	9 45
All	All	2240/3180 (70%)	1755 (78%)	430 (19%)	55 (2%)	9 45

5 of 13 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	101	GLU	13
1	A	128	GLY	13
1	A	45	LEU	8
1	A	108	PHE	5
1	A	91	LYS	3

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	99/137 (72%)	93±1 (94±1%)	6±1 (6±1%)	21 70
All	All	1980/2740 (72%)	1853 (94%)	127 (6%)	21 70

5 of 34 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	122	PHE	14
1	A	81	ASN	11
1	A	129	LYS	10
1	A	35	LEU	9
1	A	120	LYS	7

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