wwPDB NMR Structure Validation Summary Report (i)

Feb 26, 2022 – 08:27 AM EST

PDB ID	:	2CT2
Title	:	Solution Structure of the RING domain of the Tripartite motif protein 32
Authors	:	Miyamoto, K.; Tochio, N.; Sato, M.; Koshiba, S.; Inoue, M.; Kigawa, T.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2005-05-23

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:

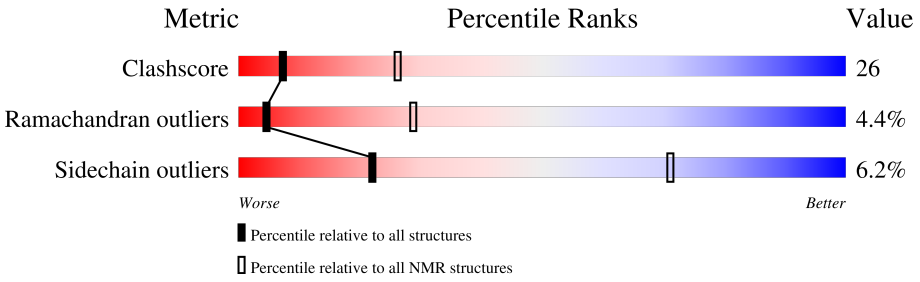
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

1 Overall quality at a glance (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	88	38% 32% • 28%

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 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:15-A:51, A:56-A:81 (63)	0.18	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 2 clusters and 1 single-model cluster was found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1323 atoms, of which 665 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Tripartite motif protein 32.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	88	1321	402	665	117	129	8	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	cloning artifact	UNP Q13049
A	2	SER	-	cloning artifact	UNP Q13049
A	3	SER	-	cloning artifact	UNP Q13049
A	4	GLY	-	cloning artifact	UNP Q13049
A	5	SER	-	cloning artifact	UNP Q13049
A	6	SER	-	cloning artifact	UNP Q13049
A	7	GLY	-	cloning artifact	UNP Q13049
A	83	SER	-	cloning artifact	UNP Q13049
A	84	GLY	-	cloning artifact	UNP Q13049
A	85	PRO	-	cloning artifact	UNP Q13049
A	86	SER	-	cloning artifact	UNP Q13049
A	87	SER	-	cloning artifact	UNP Q13049
A	88	GLY	-	cloning artifact	UNP Q13049

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
2	A	2	2	2

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: Tripartite motif protein 32

Chain A: 38% 32% • 28%

G1 S2 S3 G4 S5 S6 G7 N8 L9 D10 A11 L12 R13 E14 V15 L16 P19 S24 F25 T26 E27 L30 R31 P32 K33 L34 L35 H36 C37 G38 H39 T40 R43 L46 E47 K48 S52 S53 I54 N55 C59 P60 F61 C62 S63 K64 I65 T66 R67 I68 L71

L74 T75 D76 N77 L78 T79 V80 L81 K82 S83 G84 P85 S86 S87 G88

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

The representative model is number 6. Colouring as in section 4.1 above.

- Molecule 1: Tripartite motif protein 32

Chain A: 35% 32% 5% 28%

G1 S2 S3 G4 S5 S6 G7 N8 L9 D10 A11 L12 R13 E14 V15 L16 E17 C18 P19 E23 S24 F25 T26 E27 L30 R31 P32 K33 L34 L35 H36 C37 G38 H39 T40 R43 Q44 C45 L46 E47 K48 L49 L50 A51 S52 S53 I54 N55 C59 P60 F61 C62 S63 K64 I65 T66

R67 I68 T69 S70 L71 T72 Q73 L74 L78 T79 V80 L81 K82 S83 G84 P85 S86 S87 G88

5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function, 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)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	493	512	512	27±4
All	All	9900	10240	10240	532

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 26.

5 of 132 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:68:ILE:CD1	1:A:74:LEU:HD23	0.93	1.94	18	12
1:A:47:GLU:OE2	1:A:71:LEU:HD11	0.85	1.71	9	5
1:A:68:ILE:HD13	1:A:74:LEU:HD23	0.81	1.53	18	9
1:A:68:ILE:HD13	1:A:74:LEU:HB3	0.79	1.54	10	9
1:A:34:LEU:HD11	1:A:38:GLY:O	0.79	1.78	3	20

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	63/88 (72%)	48±2 (77±3%)	12±2 (19±3%)	3±1 (4±2%)	4 29
All	All	1260/1760 (72%)	968 (77%)	237 (19%)	55 (4%)	4 29

5 of 9 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	74	LEU	20
1	A	16	LEU	8
1	A	67	ARG	8
1	A	65	ILE	8
1	A	24	SER	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	60/79 (76%)	56±2 (94±3%)	4±2 (6±3%)	22 71
All	All	1200/1580 (76%)	1126 (94%)	74 (6%)	22 71

5 of 18 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	78	LEU	20
1	A	43	ARG	11
1	A	48	LYS	11
1	A	74	LEU	6
1	A	67	ARG	4

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)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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