wwPDB NMR Structure Validation Summary Report (i)

Mar 6, 2022 – 04:08 PM EST

PDB ID	:	2RQQ
Title	:	Structure of C-terminal region of Cdt1
Authors	:	Jee, J.G.; Mizuno, T.; Kamada, K.; Tochio, H.; Hiroaki, H.; Hanaoka, F.; Shirakawa, M.
Deposited on	:	2009-10-14

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:

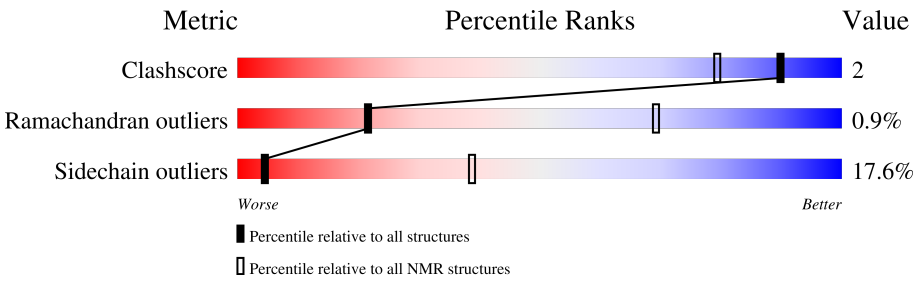
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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	113	73% 12% • 13%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 1 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:460-A:557 (98)	0.24	1

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. No single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1817 atoms, of which 927 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called DNA replication factor Cdt1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	113	1817	558	927	165	160	7	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	445	GLY	-	expression tag	UNP Q8R4E9
A	446	PRO	-	expression tag	UNP Q8R4E9
A	447	LEU	-	expression tag	UNP Q8R4E9
A	448	GLY	-	expression tag	UNP Q8R4E9
A	449	SER	-	expression tag	UNP Q8R4E9

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: DNA replication factor Cdt1

Chain A: 73% 12% • 13%

G445 P446 L447 G448 S449 M450 T451 R452 C453 P454 E455 Q456 E457 L458 R459 L460 L463 R471 V472 L473 R474 N475 V476 T486 M487 R493 M494 L513 I527 L535 L544 L548 L557

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

The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: DNA replication factor Cdt1

Chain A: 71% 13% • 13%

G445 P446 L447 G448 S449 M450 T451 R452 C453 P454 E455 Q456 E457 L458 R459 L460 Q461 R462 L463 E464 R471 R474 N475 V476 T486 M487 R493 L518 R526 I527 L535 L544 L548 H553 A554 E555 G556 L557

5 Refinement protocol and experimental data overview (i)

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

Of the 100 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
CYANA	structure solution	2.0
CYANA	refinement	2.0

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	0.63±0.00	0±0/788 (0.0± 0.0%)	1.03±0.02	1±1/1066 (0.1± 0.1%)
All	All	0.63	0/15760 (0.0%)	1.03	26/21320 (0.1%)

There are no bond-length outliers.

5 of 8 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	465	ARG	NE-CZ-NH1	6.99	123.80	120.30	2	7
1	A	474	ARG	NE-CZ-NH1	6.34	123.47	120.30	18	3
1	A	547	ARG	NE-CZ-NH1	6.14	123.37	120.30	14	5
1	A	493	ARG	NE-CZ-NH1	5.75	123.18	120.30	19	2
1	A	471	ARG	NE-CZ-NH1	5.56	123.08	120.30	14	1

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	776	811	811	3±1
All	All	15520	16220	16220	53

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 2.

5 of 14 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:460:LEU:HA	1:A:463:LEU:HD23	0.60	1.72	14	1
1:A:466:LEU:HD21	1:A:513:LEU:HD11	0.57	1.75	4	1
1:A:476:VAL:HG13	1:A:493:ARG:HB3	0.57	1.75	14	20
1:A:544:LEU:H	1:A:544:LEU:HD22	0.56	1.60	8	1
1:A:544:LEU:HD22	1:A:544:LEU:N	0.50	2.22	17	11

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	97/113 (86%)	91±1 (94±1%)	5±1 (5±1%)	1±0 (1±0%)	21	69
All	All	1940/2260 (86%)	1824 (94%)	98 (5%)	18 (1%)	21	69

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	527	ILE	18

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	85/98 (87%)	70±2 (82±2%)	15±2 (18±2%)	4	39
All	All	1700/1960 (87%)	1400 (82%)	300 (18%)	4	39

5 of 47 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	463	LEU	20
1	A	487	MET	19
1	A	548	LEU	19
1	A	460	LEU	17
1	A	471	ARG	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