



wwPDB NMR Structure Validation Summary Report ⓘ

Apr 26, 2016 – 05:03 PM BST

PDB ID : 1T3O
Title : Solution structure of CsrA, a bacterial carbon storage regulatory protein
Authors : Koharudin, L.M.I.; Georgiou, T.; Kleanthous, C.; Geoffrey, R.; Kaptein, R.; Boelens, R.
Deposited on : 2004-04-27

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
<http://wwpdb.org/validation/2016/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:

Cyrange : Kirchner and Güntert (2011)
NmrClust : Kelley et al. (1996)
MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
ShiftChecker : rb-20027457
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027457

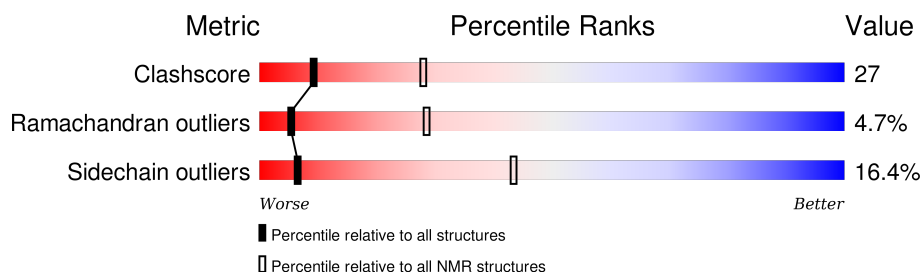
1 Overall quality at a glance

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	95	

2 Ensemble composition and analysis

This entry contains 14 models. Model 13 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:-1-A:56 (58)	0.35	13

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 9 single-model clusters were found.

Cluster number	Models
1	3, 5, 6
2	9, 13
Single-model clusters	1; 2; 4; 7; 8; 10; 11; 12; 14

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1323 atoms, of which 674 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Carbon storage regulator.

Mol	Chain	Residues	Atoms						Trace
1	A	84	Total	C	H	N	O	S	0
			1323	402	674	120	126	1	

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	INITIATING METHIONINE	UNP P33911
A	-19	GLY	-	EXPRESSION TAG	UNP P33911
A	-18	HIS	-	EXPRESSION TAG	UNP P33911
A	-17	HIS	-	EXPRESSION TAG	UNP P33911
A	-16	HIS	-	EXPRESSION TAG	UNP P33911
A	-15	HIS	-	EXPRESSION TAG	UNP P33911
A	-14	HIS	-	EXPRESSION TAG	UNP P33911
A	-13	HIS	-	EXPRESSION TAG	UNP P33911
A	-12	HIS	-	EXPRESSION TAG	UNP P33911
A	-11	HIS	-	EXPRESSION TAG	UNP P33911
A	-10	HIS	-	EXPRESSION TAG	UNP P33911
A	-9	HIS	-	EXPRESSION TAG	UNP P33911
A	-8	SER	-	EXPRESSION TAG	UNP P33911
A	-7	SER	-	EXPRESSION TAG	UNP P33911
A	-6	GLY	-	EXPRESSION TAG	UNP P33911
A	-5	HIS	-	EXPRESSION TAG	UNP P33911
A	-4	ILE	-	EXPRESSION TAG	UNP P33911
A	-3	GLU	-	EXPRESSION TAG	UNP P33911
A	-2	GLY	-	EXPRESSION TAG	UNP P33911
A	-1	ARG	-	EXPRESSION TAG	UNP P33911
A	0	HIS	-	EXPRESSION TAG	UNP P33911

5 Refinement protocol and experimental data overview

The models were refined using the following method: *Automatic peak assignment by ARIA; simulated annealing in CNS*;

Of the 30 calculated structures, 14 were deposited, based on the following criterion: *structures with the least restraint violations, structures with the lowest energy*.

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

Software name	Classification	Version
ARIA	structure solution	1.2
CNS	structure solution	1.1
CNS	refinement	1.1

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

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	0.48±0.15	1±1/469 (0.2±0.2%)	0.55±0.04	0±0/632 (0.0±0.0%)
All	All	0.50	10/6566 (0.2%)	0.55	0/8848 (0.0%)

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	0.4±0.5
All	All	0	5

All unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	48	TYR	CE1-CZ	10.79	1.52	1.38	1	5
1	A	48	TYR	CE2-CZ	-10.27	1.25	1.38	1	5

There are no bond-angle outliers.

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	48	TYR	Sidechain	5

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	464	487	484	25±5
All	All	6496	6818	6776	353

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 27.

5 of 155 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:3:VAL:HB	1:A:31:LYS:HE2	0.89	1.44	13	9
1:A:21:LYS:HE2	1:A:32:LEU:HD22	0.88	1.43	12	1
1:A:2:LEU:HB3	1:A:34:ILE:HD11	0.88	1.45	2	2
1:A:14:ILE:HG13	1:A:49:LEU:HG	0.87	1.44	9	3
1:A:41:ASP:HB3	1:A:52:GLN:HE21	0.85	1.29	8	2

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	58/95 (61%)	44±2 (76±3%)	11±1 (19±2%)	3±1 (5±2%)	5	28
All	All	812/1330 (61%)	620 (76%)	154 (19%)	38 (5%)	5	28

5 of 11 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	20	VAL	11
1	A	16	ALA	8
1	A	47	ILE	4
1	A	39	HIS	3
1	A	45	LYS	2

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	51/81 (63%)	43±2 (84±3%)	8±2 (16±3%)	6	43
All	All	714/1134 (63%)	597 (84%)	117 (16%)	6	43

5 of 28 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	48	TYR	14
1	A	14	ILE	14
1	A	40	ILE	14
1	A	39	HIS	11
1	A	4	LEU	10

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 carbohydrates 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 ⓘ

There are no chain breaks in this entry.

7 Chemical shift validation

No chemical shift data were provided