wwPDB NMR Structure Validation Summary Report

Feb 19, 2022 – 04:07 PM EST

PDB ID	:	1S40
Title	:	SOLUTION STRUCTURE OF THE CDC13 DNA-BINDING DOMAIN COMPLEXED WITH A SINGLE-STRANDED TELOMERIC DNA 11-MER
Authors	:	Mitton-Fry, R.M.; Anderson, E.M.; Theobald, D.L.; Glustrom, L.W.; Wuttke, D.S.
Deposited on	:	2004-01-14

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 ⓘ symbol.

The following versions of software and data (see references ⓘ) were used in the production of this report:

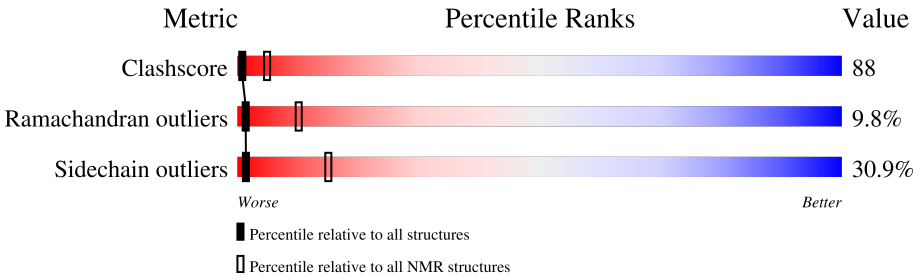
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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	B	11	100%
2	A	199	15% 54% 17% • 6% 6%

2 Ensemble composition and analysis

This entry contains 10 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

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:10-A:108, A:115-A:191 (176)	0.69	3

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, 5, 7, 8
2	6, 10
Single-model clusters	9

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3457 atoms, of which 1667 are hydrogens and 0 are deuteriums.

- Molecule 1 is a DNA chain called 5'-D(*GP*TP*GP*TP*GP*GP*GP*TP*GP*TP*G)-3'.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
1	B	11	358	110	127	43	68	10	0

- Molecule 2 is a protein called Cell division control protein 13.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	A	187	3099	1003	1540	262	283	11	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	cloning artifact	UNP P32797

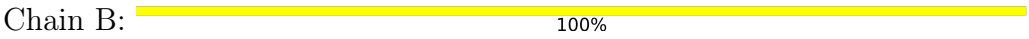
4 Residue-property plots

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.

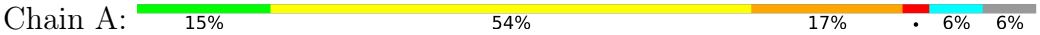
• Molecule 1: 5'-D(*GP*TP*GP*TP*GP*GP*GP*TP*GP*TP*G)-3'

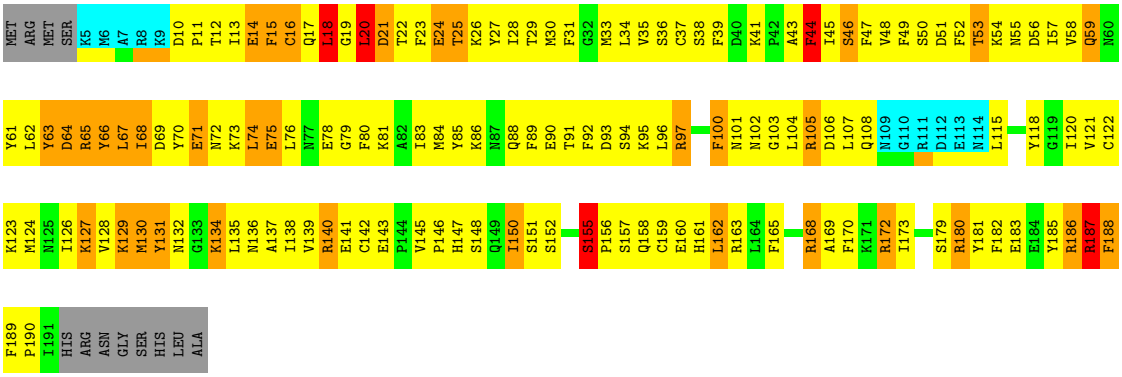
Chain B:



• Molecule 2: Cell division control protein 13

Chain A:



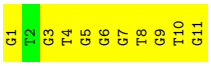


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 3. Colouring as in section 4.1 above.

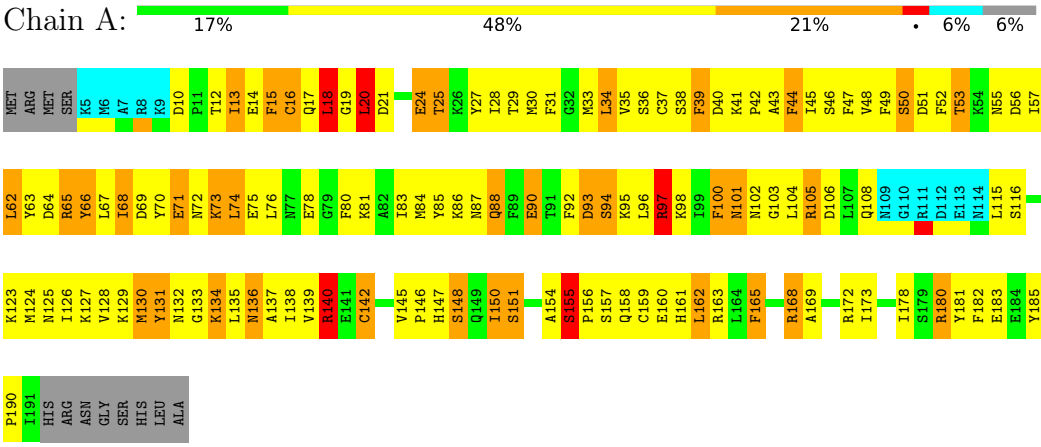
• Molecule 1: 5'-D(*GP*TP*GP*TP*GP*GP*GP*TP*GP*TP*G)-3'

Chain B:



• Molecule 2: Cell division control protein 13

Chain A:


17%
48%
21%
6%
6%
MET ARG MET SER K5 M6 A7 R8 K9 D10 P11 T12 I13 E14 F15 C16 Q17 L18 G19 L20 D21 E24 T25 K26 Y27 I28 T29 M30 F31 G32 M33 L34 V35 S36 C37 S38 F39 D40 K41 P42 A43 F44 I45 S46 F47 V48 F49 S50 D51 F52 T53 K54 N55 D56 I57 V58 Q59 N60 Y61
L62 Y63 D64 R65 Y66 L67 I68 D69 Y70 E71 N72 K73 L74 E75 L76 N77 E78 G79 F80 K81 A82 I83 M84 Y85 K86 N87 Q88 F89 E90 T91 F92 D93 S94 K95 L96 R97 K98 I99 F100 N101 N102 G103 L104 R105 D106 L107 Q108 N109 G110 R111 D112 E113 N114 L115 S116 G119 I120 V121 C122
K123 M124 N125 I126 K127 V128 K129 M130 Y131 N132 G133 K134 L135 N136 A137 I138 V139 R140 E141 C142 V145 P146 H147 S148 Q149 T150 S151 A154 S155 P156 S157 Q158 C159 E160 H161 L162 R163 L164 F165 R168 A169 R172 I173 I178 S179 R180 Y181 F182 E183 E184 Y185 R186 R187 F188 F189
P190 I191 HIS ARG ASN GLY SER HIS LEU ALA

5 Refinement protocol and experimental data overview

The models were refined using the following method: *distance geometry, simulated annealing, molecular dynamics, matrix relaxation, torsion angle dynamics.*

Of the 100 calculated structures, 10 were deposited, based on the following criterion: *structures with acceptable covalent geometry, structures with the lowest energy.*

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

Software name	Classification	Version
X-PLOR	structure solution	3.851
X-PLOR	refinement	3.851

No chemical shift data was provided.

6 Model quality

6.1 Standard geometry

There are no covalent bond-length or bond-angle outliers.

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
2	A	0.0±0.0	9.5±0.5
All	All	0	95

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

5 of 10 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
2	A	105	ARG	Sidechain	10
2	A	140	ARG	Sidechain	10
2	A	168	ARG	Sidechain	10
2	A	172	ARG	Sidechain	10
2	A	187	ARG	Sidechain	10

6.2 Too-close contacts

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	B	231	127	127	113±12
2	A	1469	1449	1449	236±15
All	All	17000	15760	15760	2873

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 88.

5 of 1492 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:B:7:DG:H2"	1:B:8:DT:O5'	1.13	1.38	4	5
2:A:83:ILE:HD11	2:A:138:ILE:HG22	1.11	1.13	1	6
2:A:13:ILE:HD11	2:A:28:ILE:HD13	1.10	1.17	10	4
1:B:10:DT:H1'	1:B:11:DG:N3	1.02	1.69	2	3
1:B:9:DG:H2"	1:B:10:DT:O5'	0.99	1.57	4	8

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	
2	A	175/199 (88%)	131±4 (75±2%)	27±4 (15±2%)	17±3 (10±2%)	1	10
All	All	1750/1990 (88%)	1308 (75%)	270 (15%)	172 (10%)	1	10

5 of 43 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	A	18	LEU	10
2	A	20	LEU	10
2	A	44	PHE	10
2	A	25	THR	9
2	A	59	GLN	9

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	
2	A	163/182 (90%)	113±5 (69±3%)	50±5 (31±3%)	1	15

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1630/1820 (90%)	1127 (69%)	503 (31%)	1 15

5 of 122 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)
2	A	18	LEU	10
2	A	20	LEU	10
2	A	150	ILE	10
2	A	185	TYR	10
2	A	105	ARG	9

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