



wwPDB NMR Structure Validation Summary Report ⓘ

Apr 26, 2016 – 09:16 PM BST

PDB ID : 2H3C
Title : Structural basis for nucleic acid and toxin recognition of the bacterial antitoxin CcdA
Authors : Madl, T.; Van Melderren, L.; Respondek, M.; Oberer, M.; Keller, W.; Zangger, K.
Deposited on : 2006-05-22

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

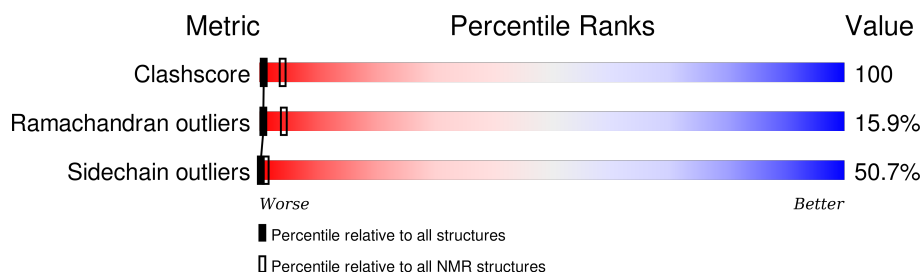
1 Overall quality at a glance

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	C	13	
2	D	13	
3	A	72	
3	B	72	

2 Ensemble composition and analysis

This entry contains 20 models. Model 16 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:2-A:38, B:102-B:138 (74)	0.41	16

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 4 clusters and 5 single-model clusters were found.

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

3 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3159 atoms, of which 1448 are hydrogens and 0 are deuteriums.

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

Mol	Chain	Residues	Atoms						Trace
1	C	13	Total	C	H	N	O	P	0
			415	127	149	47	79	13	

- Molecule 2 is a DNA chain called 5'-D(P*TP*CP*GP*GP*GP*TP*AP*TP*AP*CP*AP*TP*A)-3'.

Mol	Chain	Residues	Atoms						Trace
2	D	13	Total	C	H	N	O	P	0
			418	128	149	49	79	13	

- Molecule 3 is a protein called CcdA.

Mol	Chain	Residues	Atoms						Trace
3	A	72	Total	C	H	N	O	S	0
			1163	362	575	105	117	4	
3	B	72	Total	C	H	N	O	S	0
			1163	362	575	105	117	4	

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	70	LYS	ARG	ENGINEERED	UNP Q9S0Z5
B	170	LYS	ARG	ENGINEERED	UNP Q9S0Z5

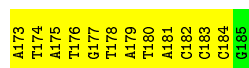
4 Residue-property plots

4.1 Average score per residue in the NMR ensemble

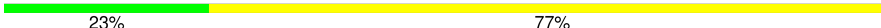
These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain 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.

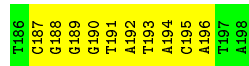
- Molecule 1: 5'-D(P*AP*TP*AP*TP*GP*TP*AP*TP*AP*CP*CP*CP*G)-3'

Chain C: 


A173 T174 T175 T176 T177 T178 T179 T180 A181 C182 C183 C184 G185

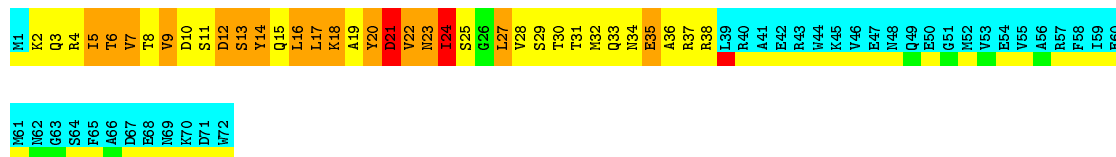
- Molecule 2: 5'-D(P*TP*CP*GP*GP*GP*TP*AP*TP*AP*CP*AP*TP*A)-3'

Chain D: 


T186 C187 G188 G189 G190 T191 A192 T193 A194 C195 A196 T197 A198


- Molecule 3: CcdA

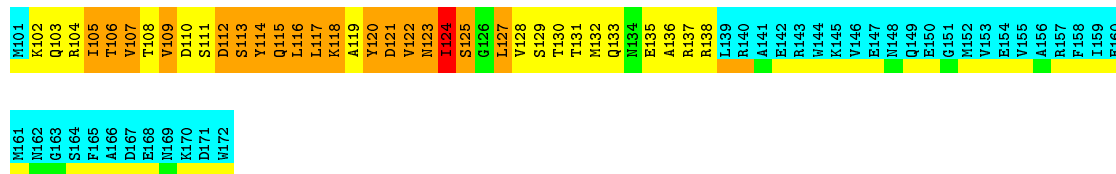
Chain A: 


M1 K2 Q3 R4 I5 T6 T7 V7 T8 V9 D10 S11 D12 S13 Y14 Q15 L16 L17 L18 K18 A19 Y20 Y21 D22 V22 N23 N24 I24 S25 G26 L27 V28 S29 S30 T30 T31 M32 M33 Q33 N34 E35 A36 R37 R38 L39 R40 A41 E42 R43 R44 K45 V46 E47 N48 Q49 E50 G51 M52 V53 E54 V55 A56 R57 F58 I59 E60


M61 M62 G63 S64 F65 A66 D67 E68 M69 K70 D71 W72

- Molecule 3: CcdA

Chain B: 


M101 K102 Q103 R104 I105 T106 T107 V108 T109 D110 S111 D112 S113 Y114 Q115 L116 L117 L118 K118 A119 Y120 Y121 D122 V122 N123 N124 I124 S125 G126 L127 V128 S129 S130 T130 T131 M132 M133 Q133 N134 E135 A136 R137 R138 L139 R140 A141 E142 R143 R144 K145 V146 E147 N148 Q149 E150 G151 M152 V153 E154 V155 A156 R157 F158 I159 E160

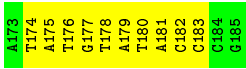

M161 M162 G163 S164 F165 A166 D167 E168 M169 K170 D171 W172

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

The representative model is number 16. Colouring as in section 4.1 above.

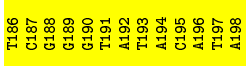
- Molecule 1: 5'-D(P*AP*TP*AP*TP*GP*TP*AP*TP*AP*CP*CP*CP*G)-3'

Chain C:  23% 77%

 A173
T174
A175
T176
A177
G177
T178
A179
T180
A181
C182
C183
G184
G185

- Molecule 2: 5'-D(P*TP*CP*GP*GP*GP*TP*AP*TP*AP*CP*AP*TP*A)-3'

Chain D:  100%

 T186
C187
G188
G189
G190
T191
A192
T193
A194
C195
A196
T197
A198


- Molecule 3: CcdA

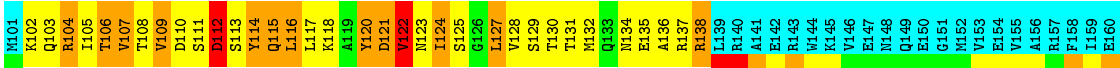
Chain A:  21% 26% 49%

 M1
R2
Q3
R4
I5
T6
V7
T8
V9
D10
S11
D12
S13
Y14
Q15
L16
L17
K18
A19
Y20
D21
V22
N23
I24
S25
G26
L27
V28
S29
T30
T31
M32
Q33
N34
E35
A36
R37
E38
L39
R40
A41
E42
R43
W44
K45
V46
E47
M48
Q49
E50
G51
V52
V53
E54
V55
A56
R57
F58
I59
E60

 M61
M62
G63
S64
F65
A66
D67
E68
M69
K70
D71
W72

- Molecule 3: CcdA

Chain B:  28% 17% 49%

 M101
K102
Q103
R104
I105
T106
V107
T108
V109
D110
S111
D112
S113
Y114
Q115
L116
L117
K118
A119
Y120
D121
V122
N123
I124
S125
G126
L127
V128
S129
T130
T131
M132
Q133
N134
E135
A136
R137
R138
L139
R140
A141
E142
R143
W144
K145
V146
E147
N148
Q149
E150
G151
M152
V153
E154
V155
A156
R157
F158
I159
E160

 M161
M162
G163
S164
F165
A166
D167
E168
M169
K170
D171
W172

5 Refinement protocol and experimental data overview ⓘ

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
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

6.1 Standard geometry

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

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	C	266	149	148	16±5
2	D	269	149	148	13±4
3	A	293	298	298	132±7
3	B	293	298	298	132±9
All	All	22420	17880	17840	4044

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 100.

5 of 1222 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
3:B:109:VAL:HG21	3:B:114:TYR:CD1	1.22	1.69	13	1
3:B:128:VAL:HG13	3:B:132:MET:SD	1.08	1.88	1	3
3:A:9:VAL:HG21	3:A:14:TYR:CD1	1.07	1.82	11	2
3:A:5:ILE:HD11	3:B:107:VAL:HG13	1.06	1.23	15	6
3:A:32:MET:HA	3:B:117:LEU:HD12	1.06	1.15	4	6

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	
3	A	37/72 (51%)	25±1 (66±2%)	6±1 (17±3%)	6±1 (16±3%)	0	4
3	B	37/72 (51%)	25±1 (68±3%)	6±2 (17±4%)	6±1 (15±3%)	1	4
All	All	1480/2880 (51%)	994 (67%)	251 (17%)	235 (16%)	1	4

5 of 21 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
3	A	9	VAL	20
3	A	21	ASP	20
3	B	121	ASP	20
3	B	109	VAL	18
3	B	113	SER	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	
3	A	34/64 (53%)	17±3 (49±7%)	17±3 (51±7%)	0	1
3	B	34/64 (53%)	17±2 (49±7%)	17±2 (51±7%)	0	1
All	All	1360/2560 (53%)	670 (49%)	690 (51%)	0	1

5 of 67 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)
3	A	6	THR	20
3	B	106	THR	20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
3	A	24	ILE	20
3	A	14	TYR	19
3	B	116	LEU	19

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 [i](#)

There are no chain breaks in this entry.

7 Chemical shift validation

No chemical shift data were provided