



wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 02:49 AM GMT

PDB ID : 1ZRD
Title : 4 crystal structures of CAP-DNA with all base-pair substitutions at position 6, CAP-[6A;17T]ICAP38 DNA
Authors : Berman, H.M.; Napoli, A.A.
Deposited on : 2005-05-19
Resolution : 2.80 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

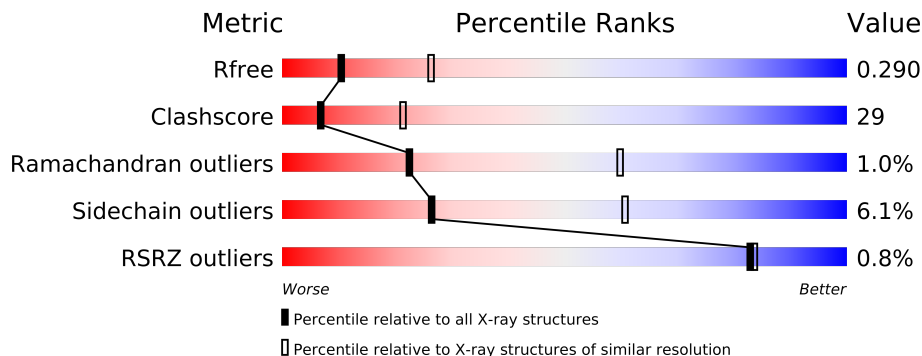
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.80 Å.

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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1799 (2.80-2.80)
Clashscore	79885	2295 (2.80-2.80)
Ramachandran outliers	78287	2252 (2.80-2.80)
Sidechain outliers	78261	2254 (2.80-2.80)
RSRZ outliers	66119	1802 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	W	17	
1	Y	17	
2	X	21	
2	Z	21	
3	A	209	
3	B	209	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4781 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a DNA chain called 5'-D(*AP*TP*TP*TP*CP*GP*AP*AP*AP*AP*T
P*GP*AP*GP*AP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	W	17	Total	C	N	O	P	0	0	0
			350	169	68	97	16			
1	Y	17	Total	C	N	O	P	0	0	0
			350	169	68	97	16			

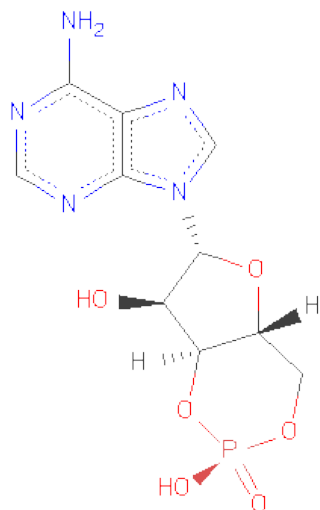
- Molecule 2 is a DNA chain called 5'-D(*CP*TP*AP*GP*AP*TP*CP*TP*CP*AP*TP*TP
*TP*TP*TP*CP*GP*AP*AP*AP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	X	21	Total	C	N	O	P	0	0	0
			423	206	70	127	20			
2	Z	21	Total	C	N	O	P	0	0	0
			423	206	70	127	20			

- Molecule 3 is a protein called Catabolite gene activator.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	200	Total	C	N	O	S	0	0	0
			1580	1001	276	294	9			
3	B	200	Total	C	N	O	S	0	0	0
			1580	1001	276	294	9			

- Molecule 4 is ADENOSINE-3',5'-CYCLIC-MONOPHOSPHATE (three-letter code: CMP) (formula: C₁₀H₁₂N₅O₆P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
4	B	1	Total	C	N	O	P	0	0
			22	10	5	6	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	16	Total	O	0	0
			16	16		
5	B	9	Total	O	0	0
			9	9		
5	W	2	Total	O	0	0
			2	2		
5	X	3	Total	O	0	0
			3	3		
5	Z	1	Total	O	0	0
			1	1		

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: 5'-D(*AP*TP*TP*TP*CP*GP*AP*AP*AP*AP*TP*GP*AP*GP*AP*T)-3',

Chain W: 

A-8
T-7
T-6
T-5
C-4
G-3
A-2
A-1
A1
A2
A3
T4
G5
A6
G7
A8
T9

- Molecule 1: 5'-D(*AP*TP*TP*TP*CP*GP*AP*AP*AP*AP*TP*GP*AP*GP*AP*T)-3',

Chain Y: 

A-8
T-7
T-6
T-5
C-4
G-3
A-2
A-1
A1
A2
A3
T4
G5
A6
G7
A8
T9

- Molecule 2: 5'-D(*CP*TP*AP*GP*AP*TP*CP*TP*CP*AP*TP*TP*TP*TP*CP*GP*AP*AP*AP*T)-3',

Chain X: 

C13
T12
A11
G10
A9
T8
C7
T6
C5
A4
T3
T2
T1
T-1
T-2
A-6
A-7
T-8

- Molecule 2: 5'-D(*CP*TP*AP*GP*AP*TP*CP*TP*CP*AP*TP*TP*TP*TP*CP*GP*AP*AP*AP*T)-3',

Chain Z: 

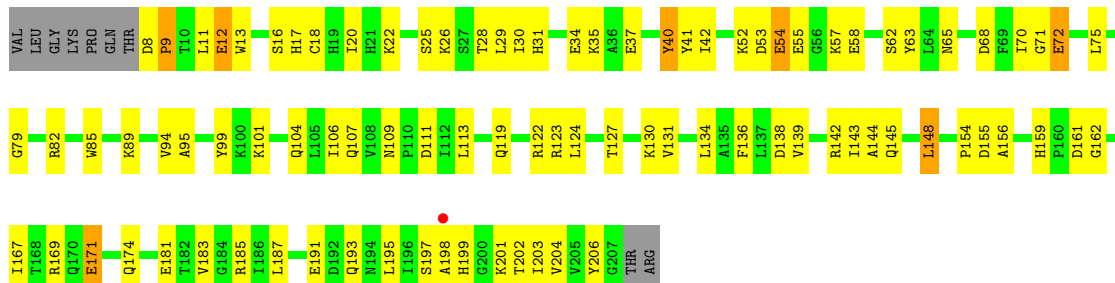
C13
T12
A11
G10
A9
T8
C7
T6
C5
A4
T3
T2
T1
T-1
T-2
C-3
G-4
A-5
A-6
A-7
T-8

- Molecule 3: Catabolite gene activator

Chain A: 

VAL
LEU
GLY
LYS
PRO
GLN
THR
D8
P9
T10
L11
F102
R103
W13
F14
L105
L106
Q107
H108
N109
C18
H19
I20
Y23
P24
S25
K26
L29
A36
E37
L137
L39
Y40
Y41
I43
A144
S46
V47
K52
D53
E54
K57
E58
M59
Y63
L64
M65
F69
G177
I70
G71
E72
S179
R180
E181
T182
L75
Q80
S83
A84
W85
I97
S98
Y99
K100
F102
R103
Q104
L105
L106
Q107
H108
N109
P110
D111
L112
L113
L116
M120
K130
F136
L137
D138
V139
I143
A144
L148
N149
L150
A151
K152
H159
P160
D161
T168
R169
Q170
E171
G177
C178
S179
R180
E181
T182
V183

Chain B:



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	61.29Å 75.34Å 178.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.63 – 2.80 46.63 – 2.77	Depositor EDS
% Data completeness (in resolution range)	77.9 (46.63-2.80) 76.9 (46.63-2.77)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.75 (at 2.77Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.224 , 0.297 0.223 , 0.290	Depositor DCC
R_{free} test set	1599 reflections (10.82%)	DCC
Wilson B-factor (Å ²)	67.3	Xtriage
Anisotropy	0.346	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 52.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 17903 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	4781	wwPDB-VP
Average B, all atoms (Å ²)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.22% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

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

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 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	W	0.33	0/394	0.78	0/607
1	Y	0.34	0/394	0.76	0/607
2	X	0.38	0/472	0.71	0/726
2	Z	0.37	0/472	0.74	0/726
3	A	0.46	0/1605	0.73	1/2162 (0.0%)
3	B	0.38	0/1605	0.66	0/2162
All	All	0.40	0/4942	0.71	1/6990 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	116	LEU	CA-CB-CG	5.99	129.08	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	W	350	0	194	27	0
1	Y	350	0	194	20	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	X	423	0	242	28	0
2	Z	423	0	242	39	0
3	A	1580	0	1615	72	0
3	B	1580	0	1615	84	0
4	A	22	0	11	3	0
4	B	22	0	11	3	0
5	A	16	0	0	1	0
5	B	9	0	0	0	0
5	W	2	0	0	0	0
5	X	3	0	0	0	0
5	Z	1	0	0	0	0
All	All	4781	0	4124	257	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 29.

The worst 5 of 257 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
4:A:210:CMP:C2	4:A:210:CMP:H2	0.97	1.49
4:B:210:CMP:C2	4:B:210:CMP:H2	0.97	1.48
2:X:11:DA:H2''	2:X:10:DG:H5''	1.08	1.08
2:X:11:DA:C2'	2:X:10:DG:H5''	1.94	0.96
3:B:42:ILE:HD11	3:B:70:ILE:HD11	1.50	0.92

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. 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	198/209 (95%)	183 (92%)	12 (6%)	3 (2%)	15	46
3	B	198/209 (95%)	178 (90%)	19 (10%)	1 (0%)	38	76
All	All	396/418 (95%)	361 (91%)	31 (8%)	4 (1%)	22	60

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	109	ASN
3	B	9	PRO
3	A	177	GLY
3	A	10	THR

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. 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	172/180 (96%)	162 (94%)	10 (6%)	28	63
3	B	172/180 (96%)	161 (94%)	11 (6%)	25	58
All	All	344/360 (96%)	323 (94%)	21 (6%)	26	61

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	182	THR
3	B	28	THR
3	B	148	LEU
3	A	170	GLN
3	B	155	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	107	GLN
3	B	109	ASN
3	B	174	GLN
3	B	65	ASN
3	B	159	HIS

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
4	CMP	A	210	-	25,25,25	0.95	1 (4%)	39,39,39	1.50	3 (7%)
4	CMP	B	210	-	25,25,25	0.89	1 (4%)	39,39,39	1.66	5 (12%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	CMP	A	210	-	-	0/4/31/31	0/0/4/4
4	CMP	B	210	-	-	0/4/31/31	0/0/4/4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	210	CMP	O3'-C3'	-2.39	1.41	1.44
4	A	210	CMP	P-O3'	2.12	1.62	1.58

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	210	CMP	O3'-C3'-C4'	-7.16	104.98	110.73
4	A	210	CMP	O3'-C3'-C4'	-6.68	105.37	110.73
4	B	210	CMP	O3'-C3'-C2'	4.03	120.25	115.62
4	A	210	CMP	O2P-P-O1P	3.24	119.51	108.75
4	B	210	CMP	O2P-P-O1P	3.16	119.26	108.75

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	W	17/17 (100%)	-0.03	0 100 100	62, 107, 140, 142	0
1	Y	17/17 (100%)	0.12	1 (5%) 22 21	71, 107, 145, 147	0
2	X	21/21 (100%)	-0.09	0 100 100	59, 88, 150, 151	0
2	Z	21/21 (100%)	0.10	2 (9%) 8 7	61, 95, 140, 145	0
3	A	200/209 (95%)	-0.06	0 100 100	36, 51, 73, 92	0
3	B	200/209 (95%)	0.10	1 (0%) 88 90	38, 65, 104, 130	0
All	All	476/494 (96%)	0.02	4 (0%) 83 83	36, 60, 122, 151	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	198	ALA	3.5
1	Y	-8	DA	2.4
2	Z	-8	DT	2.2
2	Z	-5	DA	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron

density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å ²)	Q<0.9
4	CMP	A	210	22/22	0.20	-0.01	41,44,48,51	0
4	CMP	B	210	22/22	0.21	-0.27	36,42,43,45	0

6.5 Other polymers

There are no such residues in this entry.