



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 11:11 PM GMT

PDB ID : 1J59
Title : CATABOLITE GENE ACTIVATOR PROTEIN (CAP)/DNA COMPLEX +
ADENOSINE-3',5'-CYCLIC-MONOPHOSPHATE
Authors : Parkinson, G.; Wilson, C.; Gunasekera, A.; Ebright, Y.W.; Ebright, R.H.;
Berman, H.M.
Deposited on : 2002-03-01
Resolution : 2.50 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>

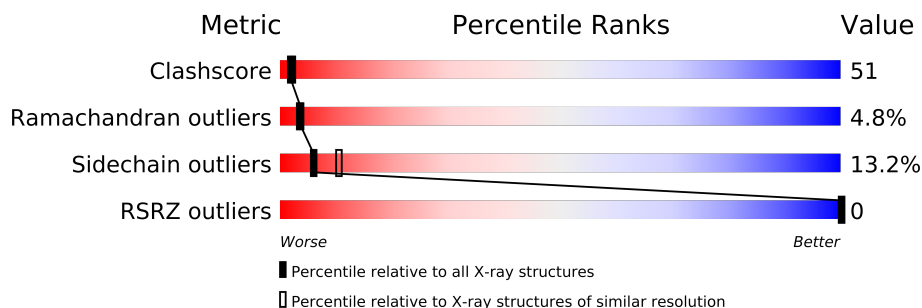
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.50 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))
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (2.50-2.50)

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	C	14	
1	E	14	
2	D	17	
2	F	17	
3	A	209	
3	B	209	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4718 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(*GP*CP*GP*AP*AP*AP*AP*GP*TP*GP*TP*GP*AP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	14	Total	C	N	O	P	0	0	0
			290	138	60	79	13			
1	E	14	Total	C	N	O	P	0	0	0
			290	138	60	79	13			

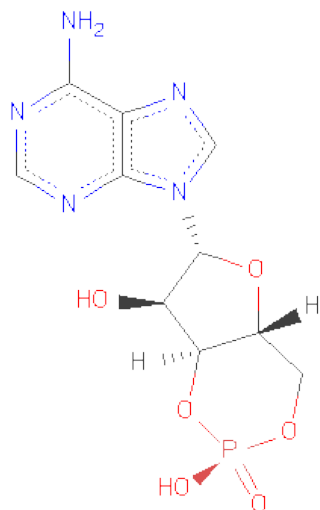
- Molecule 2 is a DNA chain called 5'-D(*AP*TP*AP*TP*GP*TP*CP*AP*CP*AP*CP*TP*TP*TP*TP*CP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	17	Total	C	N	O	P	0	0	0
			341	166	56	103	16			
2	F	17	Total	C	N	O	P	0	0	0
			341	166	56	103	16			

- Molecule 3 is a protein called CATABOLITE GENE ACTIVATOR PROTEIN (CAP).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	199	Total	C	N	O	S	0	0	0
			1572	997	275	291	9			
3	B	197	Total	C	N	O	S	0	0	0
			1556	986	273	288	9			

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



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total	C	N	O	P	0	0
			21	10	5	5	1		
4	A	1	Total	C	N	O	P	0	0
			21	10	5	5	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	21	Total	O	0	0
			21	21		
5	B	22	Total	O	0	0
			22	22		
5	C	50	Total	O	0	0
			50	50		
5	D	71	Total	O	0	0
			71	71		
5	E	47	Total	O	0	0
			47	47		
5	F	75	Total	O	0	0
			75	75		



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	136.99Å 152.80Å 76.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 2.50 26.97 – 2.50	Depositor EDS
% Data completeness (in resolution range)	86.0 (10.00-2.50) 85.2 (26.97-2.50)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.19 (at 2.50Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.199 , 0.279 0.200 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	33.2	Xtriage
Anisotropy	0.303	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 428.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtriage
Outliers	0 of 23876 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	4718	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.11% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality i

5.1 Standard geometry i

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	C	1.21	2/327 (0.6%)	1.27	4/504 (0.8%)
1	E	1.63	3/327 (0.9%)	1.56	6/504 (1.2%)
2	D	1.62	3/380 (0.8%)	1.54	5/584 (0.9%)
2	F	1.42	2/380 (0.5%)	1.59	8/584 (1.4%)
3	A	0.82	0/1597	1.12	2/2150 (0.1%)
3	B	0.91	1/1580 (0.1%)	1.15	2/2127 (0.1%)
All	All	1.09	11/4591 (0.2%)	1.27	27/6453 (0.4%)

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	1	DT	C4-C5	6.23	1.50	1.45
1	C	3	DG	C6-O6	6.18	1.29	1.24
1	E	25	DG	N9-C4	5.74	1.42	1.38
2	D	26	DG	C5-C6	5.68	1.48	1.42
1	E	17	DT	C4-O4	5.54	1.28	1.23

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	25	DG	N1-C6-O6	-11.25	113.15	119.90
2	D	23	DT	O5'-P-OP1	-8.29	98.24	105.70
1	E	25	DG	C5-C6-O6	6.90	132.74	128.60
1	C	4	DT	OP2-P-O3'	6.84	120.26	105.20
2	D	22	DT	C5'-C4'-C3'	6.62	126.02	114.10

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	C	290	0	158	28	0
1	E	290	0	158	53	0
2	D	341	0	196	48	0
2	F	341	0	194	30	0
3	A	1572	0	1612	142	0
3	B	1556	0	1600	145	0
4	A	21	0	9	6	0
4	B	21	0	9	4	0
5	A	21	0	0	9	0
5	B	22	0	0	7	0
5	C	50	0	0	4	0
5	D	71	0	0	2	0
5	E	47	0	0	8	0
5	F	75	0	0	0	0
All	All	4718	0	3936	419	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 51.

The worst 5 of 419 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
4:B:761:CMP:H2	4:B:761:CMP:C2	0.97	1.48
4:A:762:CMP:H2	4:A:762:CMP:C2	0.97	1.46
3:B:200:GLY:HA2	5:B:532:HOH:O	1.21	1.31
3:A:166:LYS:HD3	5:A:719:HOH:O	1.26	1.29
3:B:18:CYS:SG	3:B:97:ILE:HG12	1.92	1.08

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	197/209 (94%)	165 (84%)	27 (14%)	5 (2%)	9	12
3	B	195/209 (93%)	168 (86%)	13 (7%)	14 (7%)	2	1
All	All	392/418 (94%)	333 (85%)	40 (10%)	19 (5%)	4	3

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	199	HIS
3	A	206	TYR
3	A	75	LEU
3	B	155	ASP
3	B	158	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	171/180 (95%)	146 (85%)	25 (15%)	5	8
3	B	170/180 (94%)	150 (88%)	20 (12%)	8	14
All	All	341/360 (95%)	296 (87%)	45 (13%)	6	10

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	163	MET
3	B	13	TRP
3	B	182	THR
3	A	171	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	B	15	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	174	GLN
3	B	32	GLN
3	B	65	ASN
3	B	193	GLN

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	762	-	23,24,25	0.99	1 (4%)	33,37,39	1.62	6 (18%)
4	CMP	B	761	-	23,24,25	1.40	4 (17%)	33,37,39	2.17	8 (24%)

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	762	-	-	0/4/27/31	0/0/4/4
4	CMP	B	761	-	-	0/4/27/31	0/0/4/4

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	761	CMP	O5'-C5'	-3.62	1.40	1.46
4	B	761	CMP	P-O3'	2.42	1.62	1.58
4	B	761	CMP	O3'-C3'	-2.33	1.43	1.46
4	B	761	CMP	C5'-C4'	-2.31	1.47	1.51
4	A	762	CMP	O3'-C3'	-2.30	1.43	1.46

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	761	CMP	C2'-C1'-N9	6.77	120.59	113.39
4	B	761	CMP	O2P-P-O3'	5.38	118.43	107.39
4	A	762	CMP	O2P-P-O3'	4.19	116.00	107.39
4	B	761	CMP	O3'-C3'-C4'	-3.97	106.80	110.94
4	B	761	CMP	O3'-P-O1P	-3.62	102.26	109.99

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	C	14/14 (100%)	-0.74	0 100 100	23, 45, 50, 53	0
1	E	14/14 (100%)	-0.37	0 100 100	36, 47, 53, 54	0
2	D	17/17 (100%)	-0.55	0 100 100	27, 45, 54, 54	0
2	F	17/17 (100%)	-0.82	0 100 100	29, 39, 50, 52	0
3	A	199/209 (95%)	-0.74	0 100 100	6, 28, 48, 52	0
3	B	197/209 (94%)	-0.65	0 100 100	7, 28, 50, 54	0
All	All	458/480 (95%)	-0.68	0 100 100	6, 30, 50, 54	0

There are no RSRZ outliers to report.

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(\AA^2)	Q<0.9
4	CMP	A	762	21/22	0.10	-0.30	9,16,24,26	0
4	CMP	B	761	21/22	0.09	-0.70	7,18,22,29	0

6.5 Other polymers ⓘ

There are no such residues in this entry.