



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 27, 2024 – 11:03 PM EST

PDB ID : 1DIZ
Title : CRYSTAL STRUCTURE OF E. COLI 3-METHYLADENINE DNA GLYCOSYLASE (ALKA) COMPLEXED WITH DNA
Authors : Hollis, T.; Ichikawa, Y.; Ellenberger, T.E.
Deposited on : 1999-11-30
Resolution : 2.50 Å(reported)

This is a wwPDB X-ray 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/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtrriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

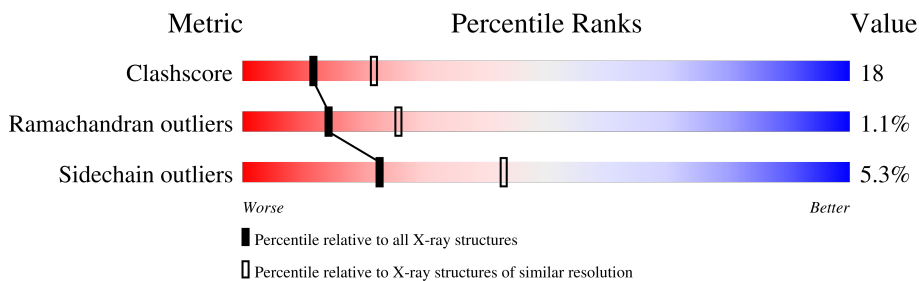
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

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	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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 $\leq 5\%$

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	13	
1	E	13	
2	D	13	
2	F	13	
3	A	282	
3	B	282	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 5643 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 DNA (5'-D(*GP*AP*CP*AP*TP*GP*AP*(NRI)P*TP*GP*CP*CP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	13	Total 254	C 122	N 46	O 74	P 12	0	0	0
1	E	13	Total 254	C 122	N 46	O 74	P 12	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	13	Total 264	C 127	N 50	O 75	P 12	0	0	0
2	F	13	Total 264	C 127	N 50	O 75	P 12	0	0	0

- Molecule 3 is a protein called 3-METHYLADENINE DNA GLYCOSYLASE II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	282	Total 2199	C 1418	N 380	O 389	S 12	0	0	0
3	B	282	Total 2199	C 1418	N 380	O 389	S 12	0	0	0

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	Na 1	0	0
4	B	1	Total 1	Na 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	15	Total O 15 15	0	0
5	D	11	Total O 11 11	0	0
5	E	14	Total O 14 14	0	0
5	F	10	Total O 10 10	0	0
5	A	73	Total O 73 73	0	0
5	B	84	Total O 84 84	0	0

3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

Note EDS was not executed.

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

Chain C: 



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

Chain E: 



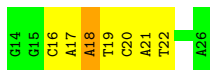
- Molecule 2: DNA (5'-D(*GP*GP*CP*AP*AP*TP*CP*AP*TP*GP*TP*CP*A)-3')

Chain D: 



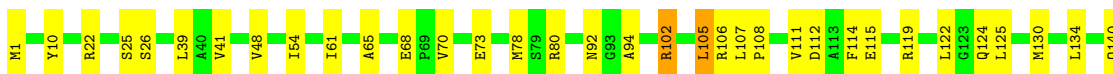
- Molecule 2: DNA (5'-D(*GP*GP*CP*AP*AP*TP*CP*AP*TP*GP*TP*CP*A)-3')

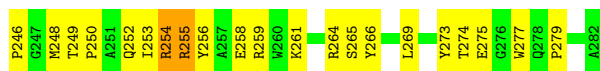
Chain F: 



- Molecule 3: 3-METHYLADENINE DNA GLYCOSYLASE II

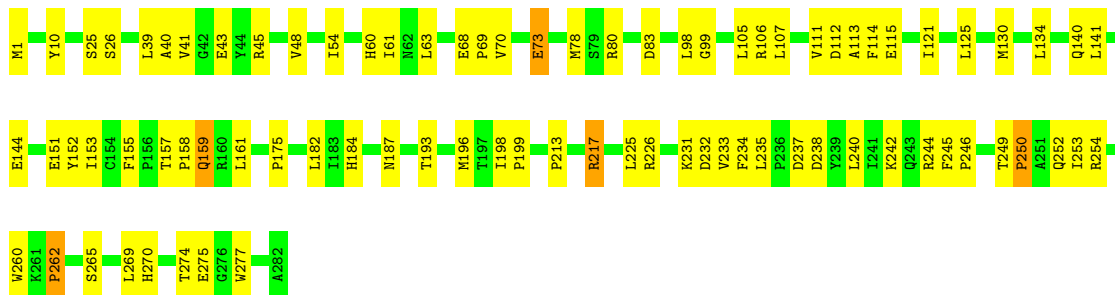
Chain A: 





● Molecule 3: 3-METHYLADENINE DNA GLYCOSYLASE II

Chain B: 71% 28%



4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	82.40Å 82.40Å 199.70Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	10.00 – 2.50	Depositor
% Data completeness (in resolution range)	97.8 (10.00-2.50)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.250 , 0.290	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	5643	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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

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	0.43	0/271	0.76	1/414 (0.2%)
1	E	0.39	0/271	0.65	0/414
2	D	0.32	0/296	0.69	0/455
2	F	0.36	0/296	0.74	0/455
3	A	0.41	0/2261	0.61	0/3086
3	B	0.41	0/2261	0.63	0/3086
All	All	0.40	0/5656	0.64	1/7910 (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 outliers	#Planarity outliers
2	D	0	1
2	F	0	1
All	All	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	9	DT	P-O5'-C5'	5.95	130.42	120.90

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	18	DA	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
2	F	18	DA	Sidechain

5.2 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 the 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 within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	254	0	146	9	0
1	E	254	0	146	8	0
2	D	264	0	148	13	0
2	F	264	0	148	9	0
3	A	2199	0	2176	79	0
3	B	2199	0	2176	70	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	73	0	0	3	0
5	B	84	0	0	4	0
5	C	15	0	0	0	0
5	D	11	0	0	0	0
5	E	14	0	0	0	0
5	F	10	0	0	0	0
All	All	5643	0	4940	182	0

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 18.

The worst 5 of 182 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:232:ASP:HB2	5:B:850:HOH:O	1.32	1.28
1:E:2:DA:H2''	1:E:3:DC:H5''	1.21	1.19
1:E:2:DA:H2''	1:E:3:DC:C5'	1.88	1.04
3:A:255:ARG:HH21	3:A:256:TYR:HA	1.23	1.00
2:F:17:DA:H2''	2:F:18:DA:H5'	1.48	0.96

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 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 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	280/282 (99%)	258 (92%)	19 (7%)	3 (1%)	14	26
3	B	280/282 (99%)	261 (93%)	16 (6%)	3 (1%)	14	26
All	All	560/564 (99%)	519 (93%)	35 (6%)	6 (1%)	14	26

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	125	LEU
3	A	171	ALA
3	B	125	LEU
3	A	275	GLU
3	B	275	GLU

5.3.2 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 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	219/222 (99%)	205 (94%)	14 (6%)	17	33
3	B	219/222 (99%)	210 (96%)	9 (4%)	30	55
All	All	438/444 (99%)	415 (95%)	23 (5%)	22	43

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	73	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	B	112	ASP
3	B	111	VAL
3	B	140	GLN
3	A	140	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	184	HIS
3	B	243	GLN
3	B	221	ASN
3	A	243	GLN
3	B	159	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

2 non-standard protein/DNA/RNA residues 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
1	NRI	C	8	1	6,11,12	5.85	1 (16%)	4,14,17	5.59	4 (100%)
1	NRI	E	8	1	6,11,12	5.23	2 (33%)	4,14,17	4.21	3 (75%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	NRI	C	8	1	-	2/3/15/16	0/1/1/1
1	NRI	E	8	1	-	2/3/15/16	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	8	NRI	O5'-C5'	-14.08	1.10	1.44
1	E	8	NRI	O5'-C5'	-12.48	1.14	1.44
1	E	8	NRI	C6'-N1'	-2.37	1.39	1.46

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	8	NRI	O3'-C3'-C2'	-8.64	90.57	110.90
1	E	8	NRI	O5'-C5'-C4'	6.09	148.72	110.62
1	C	8	NRI	O5'-C5'-C4'	5.70	146.30	110.62
1	E	8	NRI	O3'-C3'-C2'	-4.34	100.69	110.90
1	E	8	NRI	C6'-N1'-C2'	3.63	114.00	105.42

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	E	8	NRI	C6'-C4'-C5'-O5'
1	C	8	NRI	C3'-C4'-C5'-O5'
1	C	8	NRI	C6'-C4'-C5'-O5'
1	E	8	NRI	C3'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	8	NRI	2	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	E	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	E	8:NRI	O3'	9:DT	P	1.78

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.