



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 11, 2017 – 05:36 PM EDT

PDB ID : 5ULW
Title : Structure of human DNA polymerase iota bound to template 1-methyl-deoxy adenosine
Authors : Jain, R.; Aggarwal, A.K.
Deposited on : unknown
Resolution : 2.62 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

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
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20029824
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029824

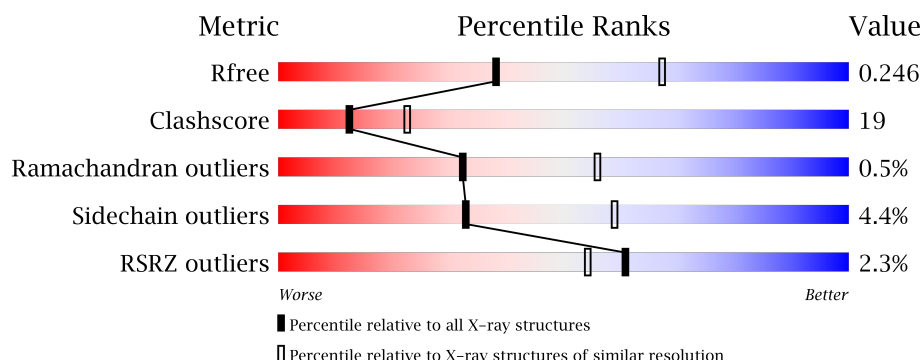
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.62 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}	100719	2983 (2.64-2.60)
Clashscore	112137	3351 (2.64-2.60)
Ramachandran outliers	110173	3298 (2.64-2.60)
Sidechain outliers	110143	3298 (2.64-2.60)
RSRZ outliers	101464	2992 (2.64-2.60)

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. 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	P	7	
2	T	11	
3	A	420	

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 3263 atoms, of which 39 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(*AP*GP*GP*AP*CP*CP*(DOC))-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	P	7	Total	C	H	N	O	P	0	0	0
			151	67	12	29	37	6			

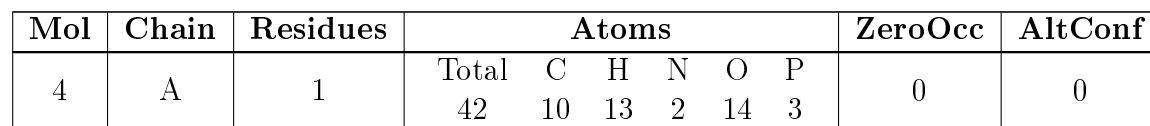
- Molecule 2 is a DNA chain called DNA (5'-D(*TP*(MA7)P*GP*GP*GP*TP*CP*CP*T)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	T	9	Total	C	H	N	O	P	0	0	0
			197	89	14	32	54	8			

- Molecule 3 is a protein called DNA polymerase iota.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	373	Total	C	N	O	S	0	2	0
			2784	1763	482	518	21			

- Molecule 4 is THYMIDINE-5'-TRIPHOSPHATE (three-letter code: TTP) (formula: C₁₀H₁₇N₂O₁₄P₃).



- | Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|-----|-------|----------|-----------------|---------|---------|
| 5 | A | 1 | Total Mg
1 1 | 0 | 0 |

- | Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|-----|-------|----------|-----------------|---------|---------|
| 6 | A | 1 | Total Cl
1 1 | 0 | 0 |

- | Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|-----|-------|----------|------------------|---------|---------|
| 7 | P | 2 | Total O
2 2 | 0 | 0 |
| 7 | T | 5 | Total O
5 5 | 0 | 0 |
| 7 | A | 80 | Total O
80 80 | 0 | 0 |

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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 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: DNA (5'-D(*AP*GP*GP*AP*CP*CP*(DOC))-3')

Chain P: 



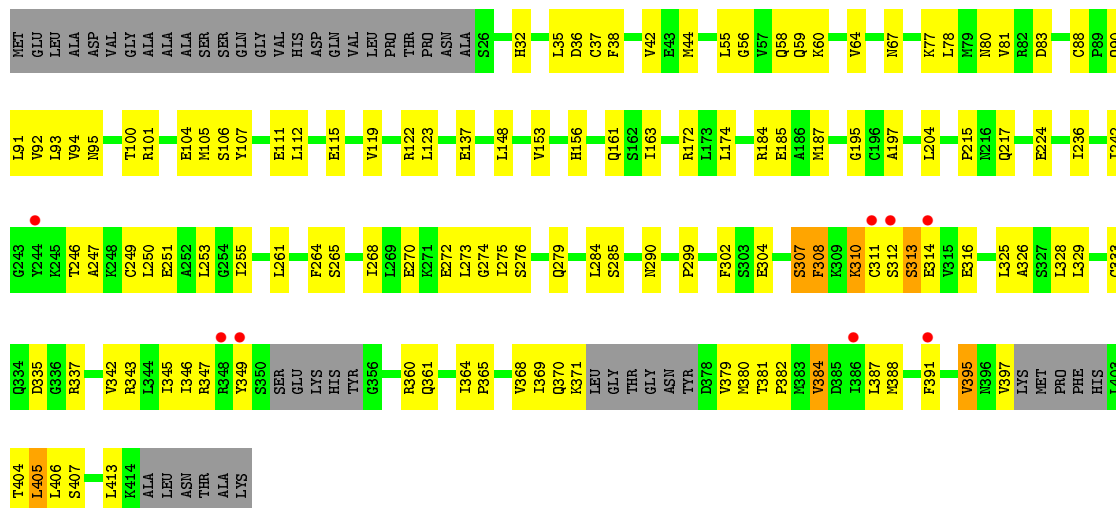
- Molecule 2: DNA (5'-D(*TP*(MA7)P*GP*GP*GP*TP*CP*CP*T)-3')

Chain T: 



- Molecule 3: DNA polymerase iota

Chain A: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α , β , γ	97.92Å 97.92Å 202.47Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.96 – 2.62 48.96 – 2.62	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.96-2.62) 99.2 (48.96-2.62)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.54 (at 2.61Å)	Xtriage
Refinement program	PHENIX (1.10.1 _2155: ???)	Depositor
R, R_{free}	0.204 , 0.248 0.200 , 0.246	Depositor DCC
R_{free} test set	1413 reflections (7.89%)	DCC
Wilson B-factor (Å ²)	53.4	Xtriage
Anisotropy	0.269	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 56.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	3263	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.45% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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: DOC, MG, TTP, MA7, CL

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	P	1.04	0/136	0.88	1/208 (0.5%)
2	T	1.07	0/178	1.06	0/271
3	A	0.41	0/2827	0.57	0/3836
All	All	0.51	0/3141	0.63	1/4315 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	P	871	DC	O4'-C4'-C3'	-5.20	102.42	104.50

There are no chirality outliers.

There are no planarity outliers.

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	P	139	12	79	5	1
2	T	183	14	107	7	1
3	A	2784	0	2730	105	0
4	A	29	13	13	0	0
5	A	1	0	0	0	0
6	A	1	0	0	1	0
7	A	80	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	P	2	0	0	0	0
7	T	5	0	0	0	0
All	All	3224	39	2929	114	1

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 19.

The worst 5 of 114 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)
1:P:873:DOC:O4'	1:P:873:DOC:C1'	1.65	1.20
2:T:840:MA7:O4'	2:T:840:MA7:C4'	1.69	1.18
3:A:44:MET:HE1	3:A:67:ASN:HD22	1.16	1.10
3:A:35:LEU:HD21	3:A:187:MET:HE1	1.52	0.92
3:A:185:GLU:HG2	7:A:656:HOH:O	1.71	0.88

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:P:867:DA:O5'	2:T:847:DT:O3'[10_665]	1.73	0.47

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	367/420 (87%)	355 (97%)	10 (3%)	2 (0%)	32 56

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	276	SER
3	A	370	GLN

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	296/376 (79%)	283 (96%)	13 (4%)	33 59

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	312	SER
3	A	313	SER
3	A	384	VAL
3	A	310	LYS
3	A	360	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	128	ASN
3	A	156	HIS

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	DOC	P	873	1,2	13,19,20	5.25	10 (76%)	12,26,29	1.02	1 (8%)
2	MA7	T	840	2	17,24,25	3.84	9 (52%)	15,35,38	2.49	4 (26%)

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
1	DOC	P	873	1,2	-	0/3/18/19	0/2/2/2
2	MA7	T	840	2	-	0/3/21/22	0/3/3/3

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	P	873	DOC	C2'-C1'	-7.36	1.33	1.51
2	T	840	MA7	C3'-C4'	-6.71	1.34	1.53
1	P	873	DOC	O4'-C4'	-6.58	1.30	1.44
2	T	840	MA7	O5'-C5'	-4.55	1.38	1.44
2	T	840	MA7	O4'-C1'	-4.09	1.33	1.42

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	840	MA7	C4'-O4'-C1'	-2.04	104.43	109.42
2	T	840	MA7	O4'-C1'-C2'	-2.01	102.40	106.25
1	P	873	DOC	C3'-C2'-C1'	2.05	104.96	102.69
2	T	840	MA7	C2'-C3'-C4'	2.53	108.11	102.73
2	T	840	MA7	N6-C6-N1	8.19	125.88	120.73

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	P	873	DOC	2	0
2	T	840	MA7	3	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 for Mogul analysis.

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	TTP	A	501	5	21,30,30	0.77	1 (4%)	24,47,47	1.27	4 (16%)

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	TTP	A	501	5	-	0/18/34/34	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	501	TTP	C4-N3	-2.66	1.33	1.36

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	501	TTP	C5M-C5-C4	-2.83	118.09	121.08
4	A	501	TTP	C5-C6-N1	-2.68	119.25	122.15
4	A	501	TTP	C2'-C1'-N1	-2.56	108.18	114.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
4	A	501	TTP	C5M-C5-C6	2.55	123.76	118.67

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](#)

There are no chain breaks in this entry.

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

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

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	P	6/7 (85%)	-0.59	0 100 100	47, 58, 68, 69	0
2	T	8/11 (72%)	-0.03	1 (12%) 4 2	46, 49, 54, 110	0
3	A	373/420 (88%)	-0.05	8 (2%) 64 58	31, 59, 94, 109	0
All	All	387/438 (88%)	-0.05	9 (2%) 61 54	31, 59, 94, 110	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	311	CYS	2.9
3	A	391	PHE	2.8
3	A	244	TYR	2.3
3	A	386	ILE	2.3
3	A	314	GLU	2.2

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

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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
1	DOC	P	873	18/19	0.98	0.15	-	34,48,59,59	0
2	MA7	T	840	22/23	0.94	0.16	-	43,63,101,112	0

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
4	TTP	A	501	29/29	0.98	0.14	-0.63	37,45,53,58	0
5	MG	A	502	1/1	0.97	0.11	-	41,41,41,41	0
6	CL	A	503	1/1	0.89	0.11	-	96,96,96,96	0

6.5 Other polymers [i](#)

There are no such residues in this entry.