



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:05 AM GMT

PDB ID : 2FLP  
Title : Binary complex of the catalytic core of human DNA polymerase iota with DNA (template G)  
Authors : Nair, D.T.; Johnson, R.E.; Prakash, L.; Prakash, S.; Aggarwal, A.K.  
Deposited on : 2006-01-06  
Resolution : 2.40 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](mailto: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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

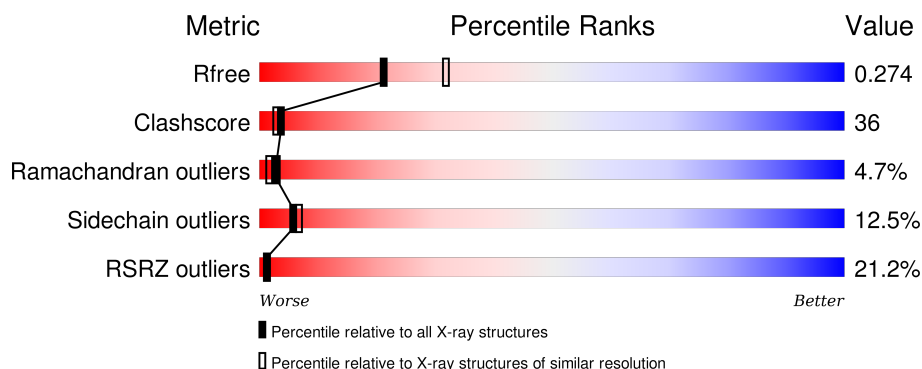
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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))
$R_{free}$	91344	2919 (2.40-2.40)
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (2.40-2.40)

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  $\geq 3$ , 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	T	11	<div> <div>9%</div> <div>36%</div> <div>36%</div> <div>9%</div> <div>18%</div> </div>
2	P	7	<div> <div>71%</div> <div>29%</div> </div>
3	A	420	<div> <div>19%</div> <div>43%</div> <div>28%</div> <div>13%</div> <div>•</div> <div>13%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3250 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 template strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	T	9	Total	C	N	O	P	0	0	0
			183	88	32	55	8			

- Molecule 2 is a DNA chain called DNA primer strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	7	Total	C	N	O	P	0	0	0
			139	67	29	37	6			

- Molecule 3 is a protein called DNA polymerase iota.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	367	Total	C	N	O	S	0	0	0
			2813	1768	495	529	21			

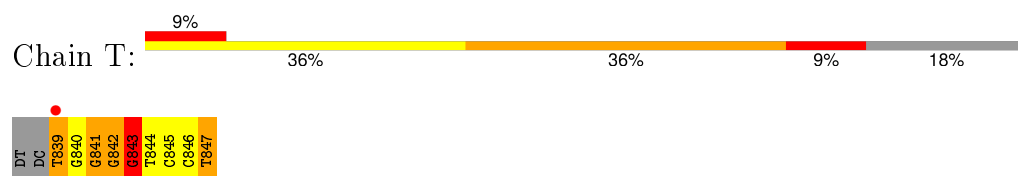
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	107	Total	O	0	0
			107	107		
4	P	2	Total	O	0	0
			2	2		
4	T	6	Total	O	0	0
			6	6		

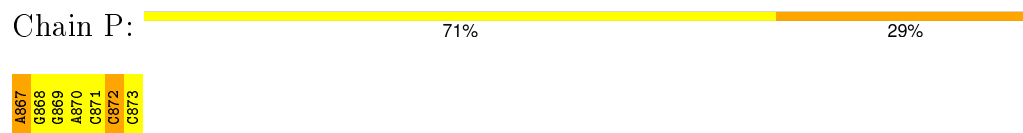
### 3 Residue-property plots

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.

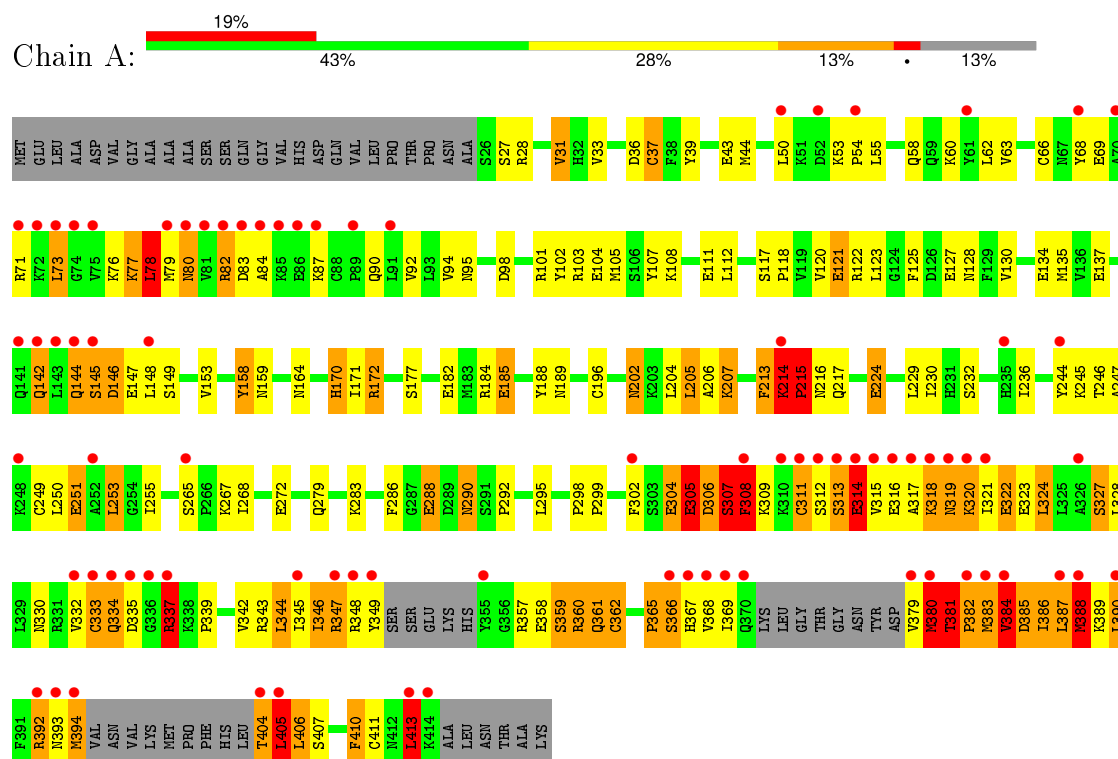
- Molecule 1: DNA template strand



- Molecule 2: DNA primer strand



- Molecule 3: DNA polymerase iota



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	98.37Å 98.37Å 202.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	36.51 – 2.40 32.59 – 2.40	Depositor EDS
% Data completeness (in resolution range)	96.2 (36.51-2.40) 96.2 (32.59-2.40)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.65 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.228 , 0.276 0.229 , 0.274	Depositor DCC
$R_{free}$ test set	2280 reflections (11.22%)	DCC
Wilson B-factor (Å <sup>2</sup> )	43.8	Xtriage
Anisotropy	0.148	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 71.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 23087 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3250	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.87% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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

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	T	3.87	32/204 (15.7%)	3.78	52/314 (16.6%)
2	P	3.30	13/136 (9.6%)	3.58	24/208 (11.5%)
3	A	3.23	114/2850 (4.0%)	1.52	46/3851 (1.2%)
All	All	3.28	159/3190 (5.0%)	1.91	122/4373 (2.8%)

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
1	T	0	1
3	A	0	3
All	All	0	4

The worst 5 of 159 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	313	SER	CB-OG	99.73	2.71	1.42
3	A	318	LYS	CE-NZ	42.89	2.56	1.49
3	A	319	ASN	CG-ND2	28.92	2.05	1.32
3	A	314	GLU	CB-CG	28.28	2.05	1.52
3	A	319	ASN	CG-OD1	26.82	1.82	1.24

The worst 5 of 122 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	T	844	DT	C4-C5-C7	18.14	129.88	119.00
2	P	870	DA	O4'-C1'-N9	16.06	119.24	108.00
2	P	867	DA	O4'-C1'-N9	15.30	118.71	108.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	122	ARG	NE-CZ-NH2	-13.75	113.43	120.30
3	A	306	ASP	CB-CG-OD1	12.61	129.65	118.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	213	PHE	Peptide
3	A	305	GLU	Peptide
3	A	388	MET	Peptide
1	T	843	DG	Sidechain

## 5.2 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 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	T	183	0	104	9	2
2	P	139	0	79	2	2
3	A	2813	0	2813	212	0
4	A	107	0	0	8	0
4	P	2	0	0	0	0
4	T	6	0	0	0	0
All	All	3250	0	2996	219	2

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 36.

The worst 5 of 219 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:A:392:ARG:CB	3:A:392:ARG:CG	1.79	1.58
3:A:318:LYS:CG	3:A:318:LYS:CB	1.79	1.57
3:A:230:ILE:CG1	3:A:230:ILE:CD1	1.77	1.55
3:A:105:MET:CE	3:A:105:MET:SD	2.01	1.48
3:A:382:PRO:C	3:A:382:PRO:CA	1.84	1.46

All (2) 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:T:847:DT:O3'	2:P:867:DA:O5'[10_665]	1.55	0.65
1:T:847:DT:O3'	2:P:867:DA:C5'[10_665]	2.14	0.06

## 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	359/420 (86%)	315 (88%)	27 (8%)	17 (5%)	<b>3</b> <b>2</b>

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	145	SER
3	A	146	ASP
3	A	333	CYS
3	A	334	GLN
3	A	366	SER

### 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	312/376 (83%)	273 (88%)	39 (12%)	<b>6</b> <b>7</b>

5 of 39 residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
3	A	229	LEU
3	A	309	LYS
3	A	406	LEU
3	A	295	LEU
3	A	306	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	217	GLN
3	A	367	HIS
3	A	279	GLN
3	A	202	ASN
3	A	262	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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
2	DOC	P	873	1,2	11,19,20	2.22	5 (45%)	14,26,29	2.18	3 (21%)

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
2	DOC	P	873	1,2	-	0/3/18/19	0/2/2/2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	873	DOC	O4'-C1'	-2.15	1.37	1.42
2	P	873	DOC	C5-C4	2.74	1.47	1.40
2	P	873	DOC	O4'-C4'	2.77	1.50	1.44
2	P	873	DOC	C2'-C1'	3.75	1.61	1.51
2	P	873	DOC	C6-N1	4.18	1.41	1.35

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	P	873	DOC	C2'-C1'-N1	-4.52	103.33	112.49
2	P	873	DOC	C2-N3-C4	3.45	120.48	115.61
2	P	873	DOC	O4'-C4'-C3'	4.35	111.98	104.69

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	P	873	DOC	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	T	9/11 (81%)	0.77	1 (11%) <b>7</b> <b>7</b>	32, 34, 64, 101	0
2	P	6/7 (85%)	0.27	0 <b>100</b> <b>100</b>	32, 41, 44, 53	0
3	A	367/420 (87%)	1.15	80 (21%) <b>1</b> <b>1</b>	9, 39, 89, 101	0
All	All	382/438 (87%)	1.13	81 (21%) <b>1</b> <b>1</b>	9, 39, 89, 101	0

The worst 5 of 81 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	355	TYR	12.4
3	A	313	SER	10.5
3	A	349	TYR	9.3
3	A	319	ASN	8.9
3	A	84	ALA	7.1

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	DOC	P	873	18/19	0.93	0.18	-	27,34,39,42	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.