



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:34 AM GMT

PDB ID : 2HPM  
Title : Eubacterial and Eukaryotic Replicative DNA Polymerases are not Homologous: X-ray Structure of DNA Polymerase III  
Authors : Bailey, S.; Wing, R.A.; Steitz, T.A.  
Deposited on : 2006-07-17  
Resolution : 3.70 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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

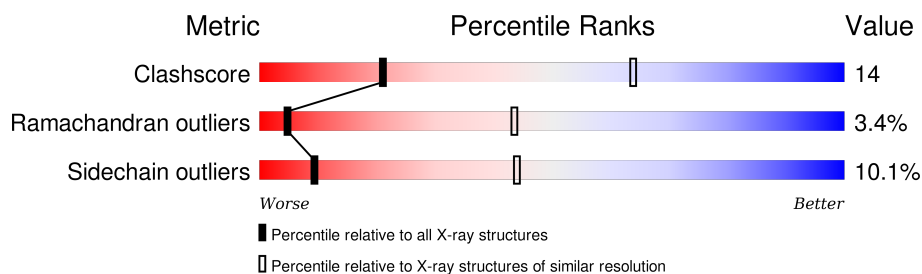
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.70 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	102246	1224 (3.90-3.50)
Ramachandran outliers	100387	1172 (3.90-3.50)
Sidechain outliers	100360	1170 (3.90-3.50)

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  $\geq 3$ , 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	1220	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9146 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 protein called DNA Polymerase III alpha subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1143	Total	C	N	O	S	0	0	0
			9127	5820	1598	1681	28			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Zn	0	0
			2	2		

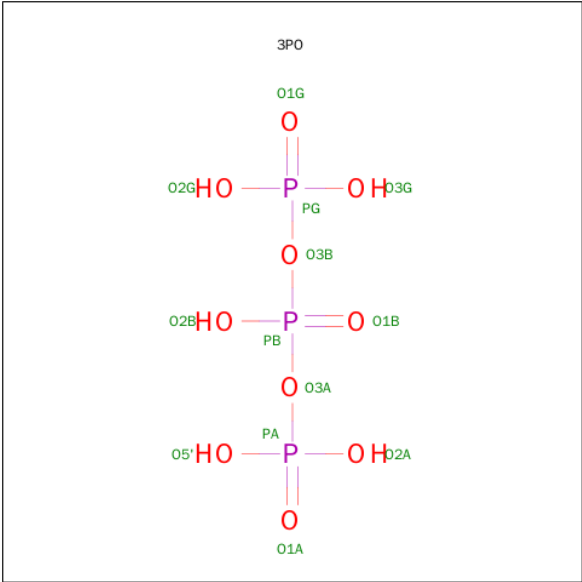
- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Mg	0	0
			2	2		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Cl	0	0
			2	2		

- Molecule 5 is TRIPHOSPHATE (three-letter code: 3PO) (formula: H<sub>5</sub>O<sub>10</sub>P<sub>3</sub>).



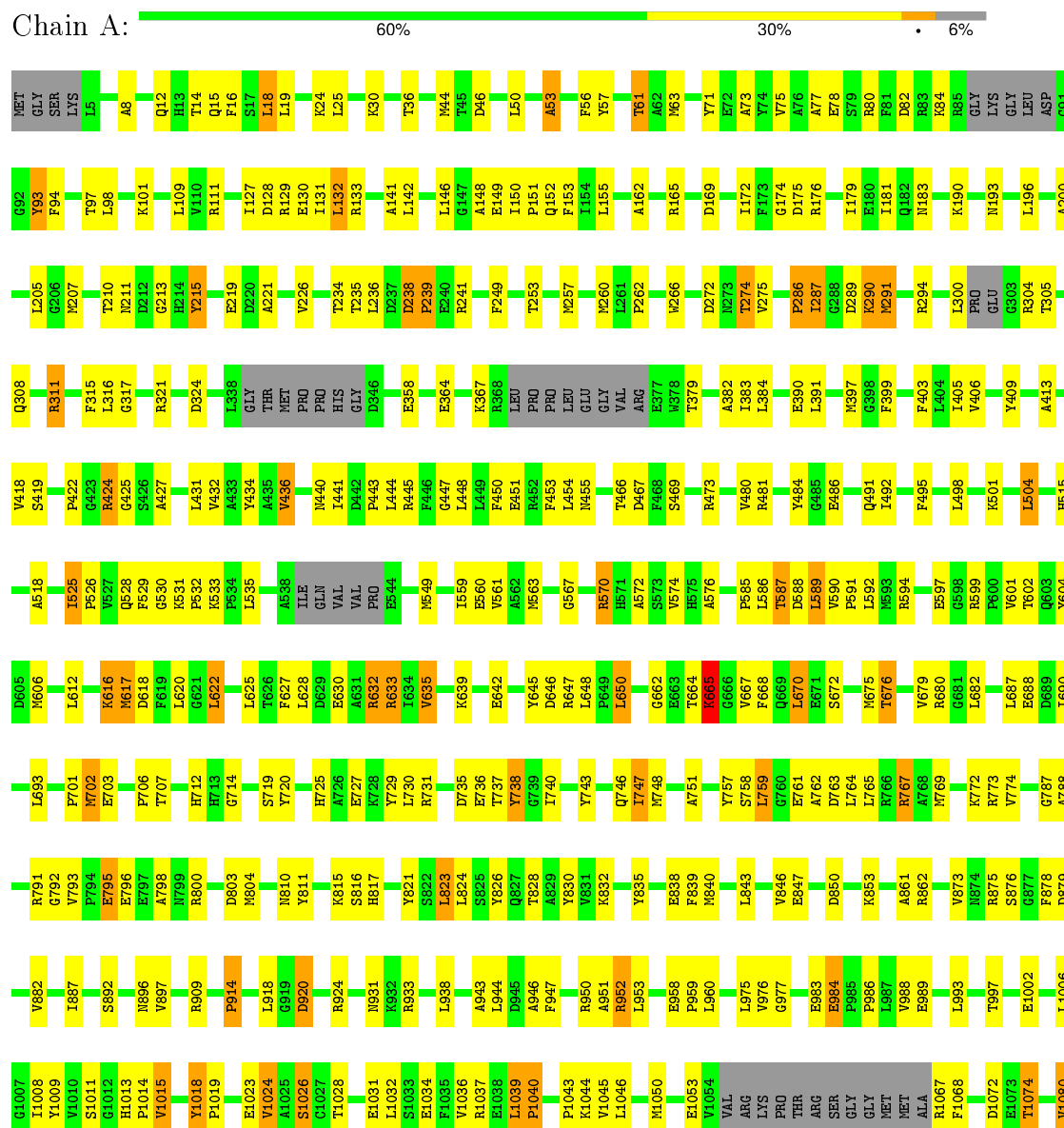
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	P	0	0
			13	10	3		

### 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.

Note EDS was not executed.

- Molecule 1: DNA Polymerase III alpha subunit



VAL	PHE	GLY	ARG	ALA	TYR	GLU	GLY	VAL	SER	PRO	LYS	LEU	K1094	E1095	D1096	I1097	P1098	L1099	L1100	V1101	L1102	A1103	E1104	V1105	E1106	LYS	GLY	GLU	GLU	LEU	R1112	A1115	V1118	W1119	T1120	K1129	A1130	L1131	V1135	D1136	K1143	G1144	VAL	A1146	R1147	G1157	S1158	L1159	F1169	L1173	F1174																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
A1175	L1176	R1177	E1178	V1179	R1180	V1181	G1182	E1183	E1184	A1185	L1189	E1192	G1193	Y1194	P1200	D1201	R1202	E1203	V1204	F1205	LEU	GLN	GLY	ASN	GLY	GLY	GLY	GLY	PRO	LYS	GLU	GLU	VAL	VAL	PRO	PRO	PHE																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						

## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	174.44Å 185.72Å 125.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.70	Depositor
% Data completeness (in resolution range)	99.9 (20.00-3.70)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.240 , 0.311	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9146	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	131.0	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: ZN, CL, MG, 3PO

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	A	0.55	5/9307 (0.1%)	0.69	3/12562 (0.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	767	ARG	CZ-NH2	16.71	1.54	1.33
1	A	767	ARG	CZ-NH1	-15.07	1.13	1.33
1	A	1034	GLU	CD-OE1	6.77	1.33	1.25
1	A	219	GLU	CD-OE2	5.80	1.32	1.25
1	A	219	GLU	CD-OE1	5.77	1.31	1.25

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	767	ARG	NE-CZ-NH1	13.13	126.86	120.30
1	A	767	ARG	NE-CZ-NH2	-10.99	114.80	120.30
1	A	1159	LEU	CA-CB-CG	5.12	127.08	115.30

There are no chirality outliers.

There are no planarity outliers.

### 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	A	9127	0	9158	261	0
2	A	2	0	0	0	0
3	A	2	0	0	0	0
4	A	2	0	0	0	0
5	A	13	0	0	0	0
All	All	9146	0	9158	261	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 14.

The worst 5 of 261 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:A:1143:LYS:HA	1:A:1147:ARG:HB2	1.43	0.96
1:A:1039:LEU:HB3	1:A:1040:PRO:CD	2.01	0.90
1:A:772:LYS:HE3	1:A:810:ASN:HD21	1.36	0.90
1:A:672:SER:HB2	1:A:675:MET:HB3	1.56	0.85
1:A:1053:GLU:HB3	1:A:1068:PHE:HA	1.58	0.84

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
1	A	1123/1220 (92%)	939 (84%)	146 (13%)	38 (3%)	<b>5</b> 43

5 of 38 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	53	ALA
1	A	175	ASP
1	A	221	ALA

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	702	MET
1	A	774	VAL

### 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
1	A	949/1009 (94%)	853 (90%)	96 (10%)	9 44

5 of 96 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	616	LYS
1	A	650	LEU
1	A	1120	THR
1	A	617	MET
1	A	633	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	746	GLN
1	A	753	GLN
1	A	896	ASN
1	A	725	HIS
1	A	814	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules 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

Of 7 ligands modelled in this entry, 6 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	$\# Z  > 2$
5	3PO	A	1227	-	8,12,12	0.87	0	15,20,20	1.50	2 (13%)

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
5	3PO	A	1227	-	-	0/12/12/12	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1227	3PO	PB-O3A-PA	-2.97	122.72	132.67
5	A	1227	3PO	PB-O3B-PG	-2.85	123.10	132.67

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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.