



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 11:47 am BST

PDB ID : 7ICU
Title : DNA POLYMERASE BETA (POL B) (E.C.2.7.7.7) COMPLEXED WITH SIX BASE PAIRS OF DNA; SOAKED IN THE PRESENCE OF CDCL₂ (0.1 MILLIMOLAR)
Authors : Pelletier, H.; Sawaya, M.R.
Deposited on : 1996-04-19
Resolution : 3.30 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

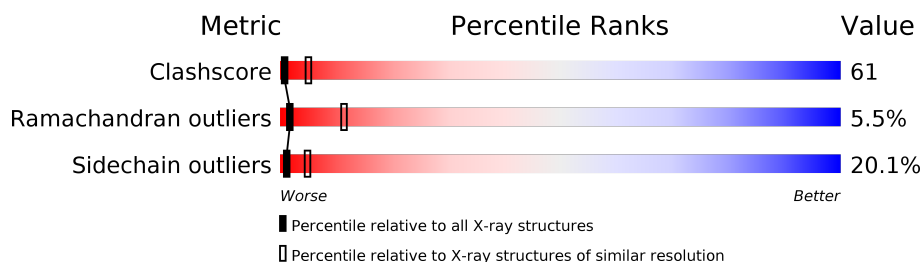
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)

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 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	T	7	
2	P	6	
3	A	335	

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 3025 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(*CP*AP*TP*CP*TP*GP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	T	7	Total	C	N	O	P	0	0	0
			122	58	20	38	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	6	Total	C	N	O	P	0	0	0
			126	59	25	36	6			

- Molecule 3 is a protein called PROTEIN (DNA POLYMERASE BETA (E.C.2.7.7.7)).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	327	Total	C	N	O	S	26	0	0
			2623	1657	458	499	9			

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

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

- Molecule 5 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Cd	0	0
			1	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	T	11	Total	O	0	0
			11	11		

Continued on next page...

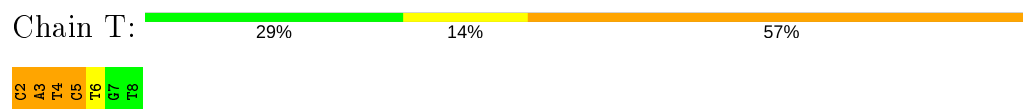
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	P	23	Total 23	O 23	0	0
6	A	117	Total 117	O 117	0	0

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 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.

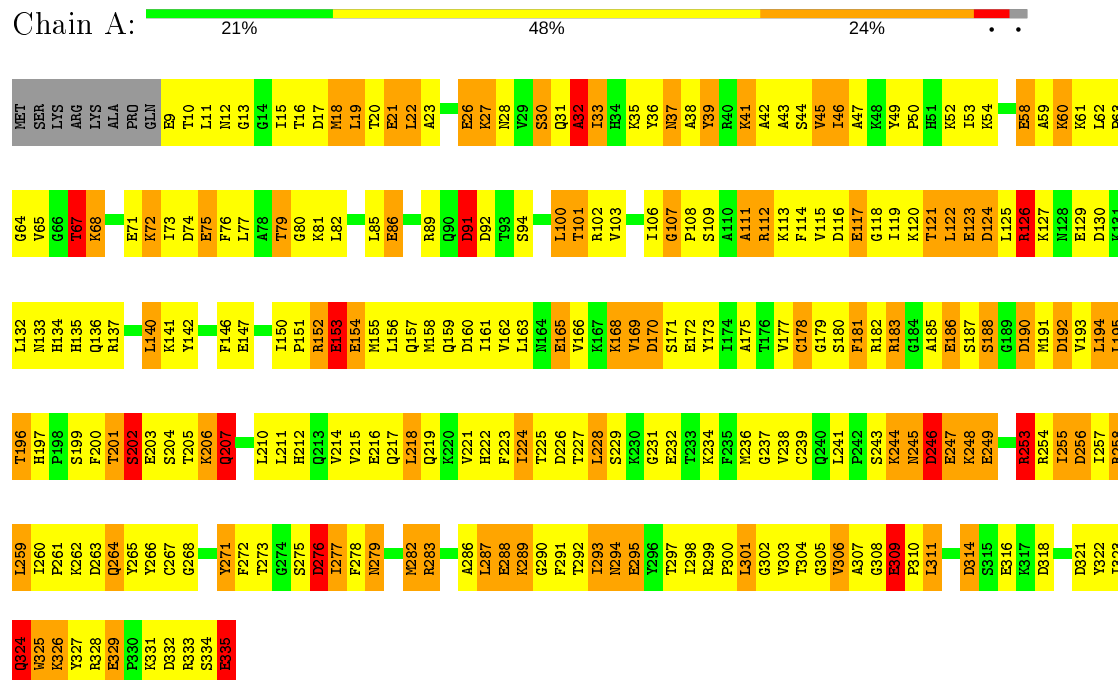
- Molecule 1: DNA (5'-D(*CP*AP*TP*CP*TP*GP*T)-3')



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



- Molecule 3: PROTEIN (DNA POLYMERASE BETA (E.C.2.7.7.7))



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	177.49 Å 57.68 Å 48.51 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.30 15.75 – 2.80	Depositor EDS
% Data completeness (in resolution range)	89.0 (20.00-3.30) 82.4 (15.75-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.32 (at 2.82 Å)	Xtriage
Refinement program	TNT 5-D	Depositor
R, R_{free}	0.165 , (Not available) 0.156 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	24.8	Xtriage
Anisotropy	0.460	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 303.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.21$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	3025	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.78% 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 ⓘ

5.1 Standard geometry ⓘ

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

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	1.96	4/135 (3.0%)	3.87	16/207 (7.7%)
2	P	2.23	4/141 (2.8%)	3.93	23/214 (10.7%)
3	A	1.22	29/2672 (1.1%)	1.65	53/3590 (1.5%)
All	All	1.33	37/2948 (1.3%)	2.01	92/4011 (2.3%)

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
3	A	2	0

The worst 5 of 37 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	5	DT	C1'-N1	11.17	1.63	1.49
2	P	5	DT	N1-C2	9.71	1.45	1.38
1	T	6	DT	C1'-N1	-8.86	1.34	1.47
3	A	75	GLU	CD-OE1	7.67	1.34	1.25
3	A	129	GLU	CD-OE1	7.24	1.33	1.25

The worst 5 of 92 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	P	5	DT	C6-N1-C1'	-25.82	81.66	120.40
2	P	5	DT	C2-N1-C1'	25.34	158.74	118.20
1	T	4	DT	C6-N1-C1'	-23.75	84.78	120.40
1	T	4	DT	C2-N1-C1'	21.98	153.37	118.20
1	T	5	DC	C2-N1-C1'	17.47	138.02	118.80

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	32	ALA	CA
3	A	86	GLU	CA

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	T	122	0	69	6	0
2	P	126	0	68	12	0
3	A	2623	0	2641	325	0
4	A	2	0	0	0	0
5	A	1	0	0	0	0
6	A	117	0	0	19	0
6	P	23	0	0	3	0
6	T	11	0	0	3	0
All	All	3025	0	2778	339	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 61.

The worst 5 of 339 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:155:MET:HA	3:A:158:MET:HE3	1.26	1.13
3:A:12:ASN:HB3	3:A:46:ILE:HD12	1.30	1.13
2:P:1:DC:H2"	2:P:2:DA:H5'	1.28	1.09
3:A:15:ILE:HD11	3:A:77:LEU:HD11	1.34	1.08
3:A:277:ILE:HG13	3:A:335:GLU:HB3	1.42	1.01

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	325/335 (97%)	273 (84%)	34 (10%)	18 (6%)	2 11

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	32	ALA
3	A	206	LYS
3	A	207	GLN
3	A	244	LYS
3	A	309	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	288/295 (98%)	230 (80%)	58 (20%)	1 5

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	194	LEU
3	A	228	LEU
3	A	324	GLN
3	A	201	THR
3	A	207	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such

sidechains are listed below:

Mol	Chain	Res	Type
3	A	134	HIS
3	A	136	GLN
3	A	245	ASN
3	A	133	ASN
3	A	264	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](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 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](#)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.