



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 23, 2024 – 06:10 AM EDT

PDB ID : 1NOY
Title : DNA POLYMERASE (E.C.2.7.7.7)/DNA COMPLEX
Authors : Wang, J.; Yu, P.; Lin, T.C.; Konigsberg, W.H.; Steitz, T.A.
Deposited on : 1996-02-16
Resolution : 2.20 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
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

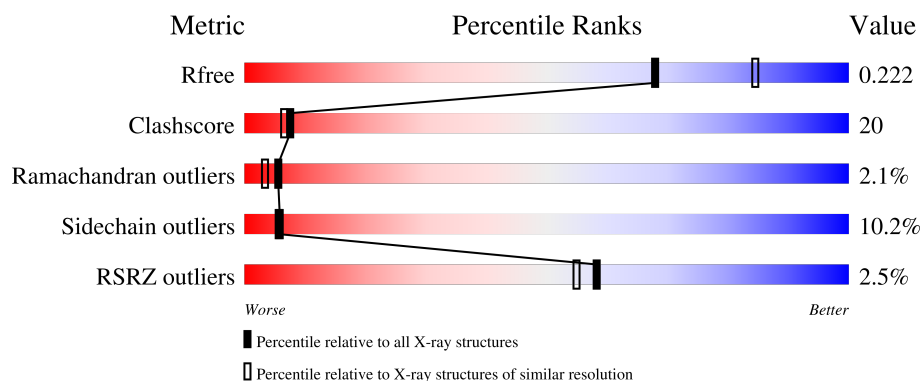
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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}	164625	5791 (2.20-2.20)
Clashscore	180529	6634 (2.20-2.20)
Ramachandran outliers	177936	6560 (2.20-2.20)
Sidechain outliers	177891	6561 (2.20-2.20)
RSRZ outliers	164620	5791 (2.20-2.20)

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 $\leq 5\%$. 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	S	3	 67% 33%
2	A	388	 2% 60% 32% . .
2	B	388	 3% 48% 35% . . 11%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	ZN	B	2000	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6130 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(*TP*TP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	S	3	Total	C	N	O	P	0	0	0
			57	30	6	19	2			

- Molecule 2 is a protein called PROTEIN (DNA POLYMERASE (E.C.2.7.7.7)).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	372	Total	C	N	O	S	0	0	0
			3052	1956	497	578	21			
2	B	346	Total	C	N	O	S	0	0	0
			2840	1822	466	533	19			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	ASP	LYS	conflict	UNP P04415
A	250	LEU	ILE	conflict	UNP P04415
B	2	ASP	LYS	conflict	UNP P04415
B	250	LEU	ILE	conflict	UNP P04415

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Zn	0	0
			1	1		

- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Mn	0	0
			1	1		

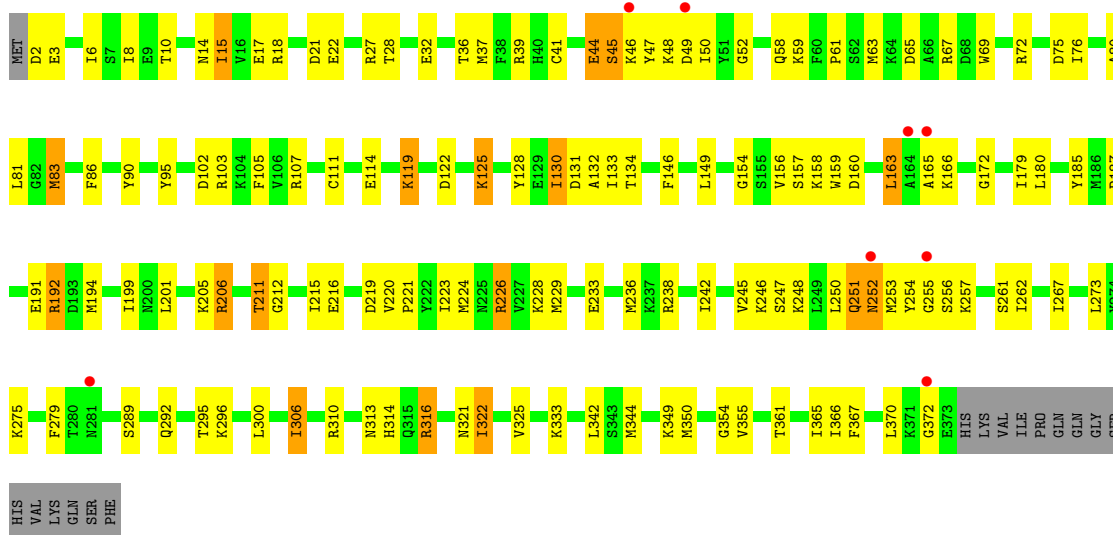
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	S	4	Total 4	O 4	0	0
5	A	112	Total 112	O 112	0	0
5	B	63	Total 63	O 63	0	0

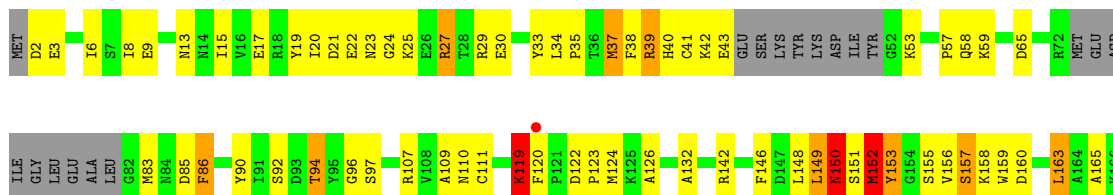
- Molecule 1: DNA (5'-D(*TP*TP*T)-3')

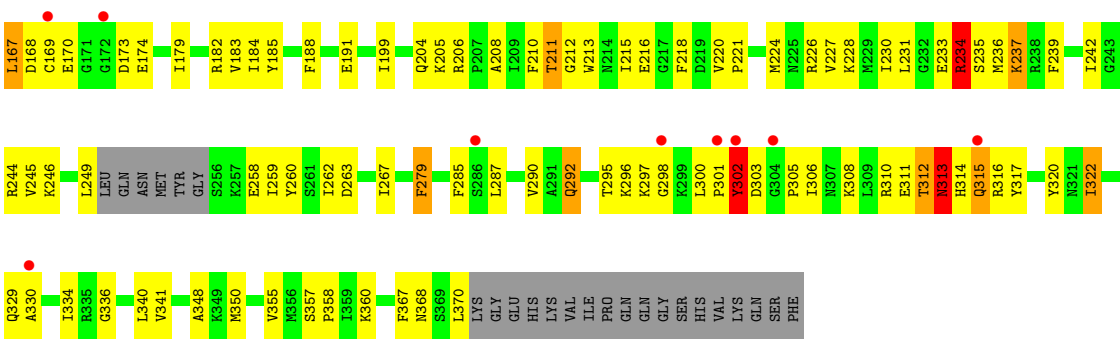
T3	T4	T5
----	----	----

Chain A: 



Chain B: 





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	114.70Å 109.66Å 70.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.00 – 2.20 6.00 – 2.49	Depositor EDS
% Data completeness (in resolution range)	(Not available) (6.00-2.20) 72.6 (6.00-2.49)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.75 (at 2.48Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.222 , (Not available) (Not available) , 0.222	Depositor DCC
R_{free} test set	1414 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å ²)	45.1	Xtriage
Anisotropy	0.440	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 127.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.015 for k,h,-l	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	6130	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.91% 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: ZN, MN

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	S	3.81	8/62 (12.9%)	4.48	20/94 (21.3%)
2	A	0.54	0/3122	0.73	0/4206
2	B	0.49	0/2904	0.71	0/3911
All	All	0.64	8/6088 (0.1%)	0.86	20/8211 (0.2%)

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	S	0	1
2	A	1	0
2	B	1	0
All	All	2	1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	S	3	DT	C5-C7	13.34	1.58	1.50
1	S	5	DT	C5-C7	12.16	1.57	1.50
1	S	4	DT	C5'-C4'	9.10	1.61	1.51
1	S	4	DT	C5-C7	8.99	1.55	1.50
1	S	3	DT	C5-C6	7.29	1.39	1.34

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	S	3	DT	OP2-P-O3'	14.82	137.80	105.20
1	S	4	DT	O5'-P-OP1	-14.27	92.86	105.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	S	4	DT	O3'-P-O5'	-13.68	78.01	104.00
1	S	4	DT	OP2-P-O3'	12.91	133.60	105.20
1	S	4	DT	O5'-P-OP2	11.74	124.79	110.70

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	211	THR	CB
2	B	211	THR	CB

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	S	5	DT	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	S	57	0	38	12	0
2	A	3052	0	2993	102	0
2	B	2840	0	2783	135	0
3	B	1	0	0	0	0
4	B	1	0	0	1	0
5	A	112	0	0	3	0
5	B	63	0	0	6	0
5	S	4	0	0	3	0
All	All	6130	0	5814	234	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 20.

The worst 5 of 234 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:S:4:DT:H4'	5:S:108:HOH:O	1.42	1.18

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:S:3:DT:O4'	5:S:116:HOH:O	1.67	1.13
1:S:5:DT:O2	2:B:119:LYS:NZ	1.84	1.10
2:B:119:LYS:HA	2:B:119:LYS:HZ3	1.15	1.08
2:A:37:MET:SD	2:A:372:GLY:HA2	2.08	0.93

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	
2	A	370/388 (95%)	344 (93%)	24 (6%)	2 (0%)	25	28
2	B	338/388 (87%)	294 (87%)	31 (9%)	13 (4%)	2	1
All	All	708/776 (91%)	638 (90%)	55 (8%)	15 (2%)	5	3

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	45	SER
2	B	22	GLU
2	B	152	MET
2	B	313	ASN
2	B	150	ASN

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	
2	A	335/350 (96%)	303 (90%)	32 (10%)	7	6
2	B	312/350 (89%)	278 (89%)	34 (11%)	5	5
All	All	647/700 (92%)	581 (90%)	66 (10%)	6	6

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	279	PHE
2	B	292	GLN
2	B	322	ILE
2	A	289	SER
2	A	254	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	292	GLN
2	B	313	ASN
2	B	321	ASN
2	A	313	ASN
2	A	321	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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

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](#)

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	S	3/3 (100%)	1.05	0 100 100	20, 20, 20, 20	0
2	A	372/388 (95%)	-0.03	8 (2%) 62 58	14, 38, 73, 80	0
2	B	346/388 (89%)	0.19	10 (2%) 54 51	23, 56, 80, 80	0
All	All	721/779 (92%)	0.08	18 (2%) 58 55	14, 46, 80, 80	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	298	GLY	4.1
2	B	330	ALA	4.0
2	B	169	CYS	3.5
2	A	255	GLY	3.3
2	A	46	LYS	3.1

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

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

6.3 Carbohydrates [i](#)

There are no monosaccharides 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. 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	B-factors(\AA^2)	Q<0.9
3	ZN	B	2000	1/1	-0.17	0.51	20,20,20,20	0
4	MN	B	3000	1/1	0.92	0.31	20,20,20,20	0

6.5 Other polymers [i](#)

There are no such residues in this entry.