



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:09 PM GMT

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
<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

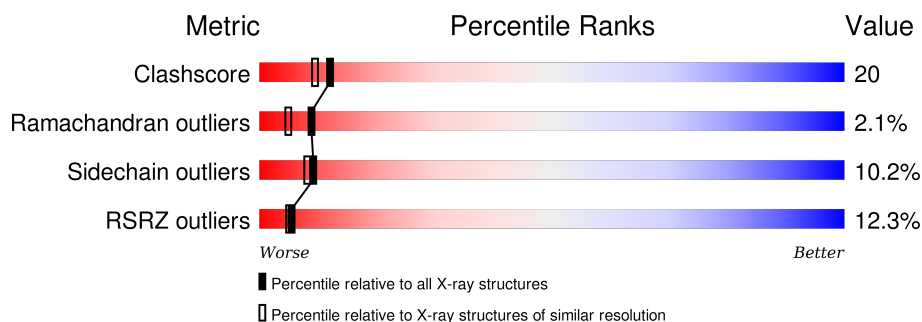
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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))
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (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 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	S	3	<div> <div>100%</div> <div>67% 33%</div> </div>
2	A	388	<div> <div>8%</div> <div>60% 32% . .</div> </div>
2	B	388	<div> <div>14%</div> <div>48% 35% . . 11%</div> </div>

2 Entry composition [i](#)

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		

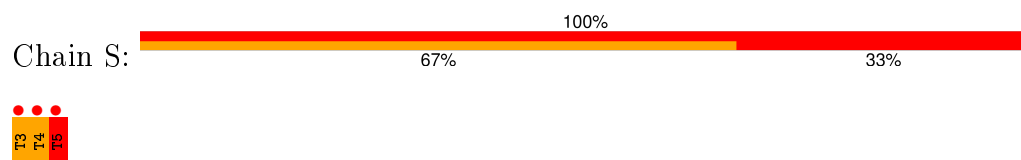
- Molecule 5 is water.

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

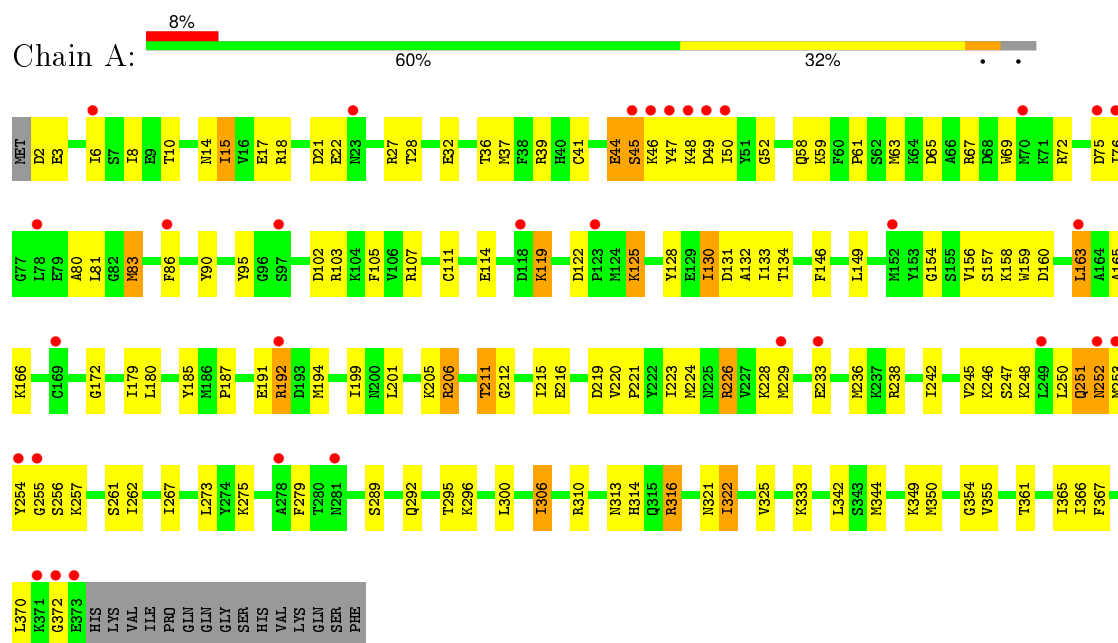
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 ($\text{RSRZ} > 2$). 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.

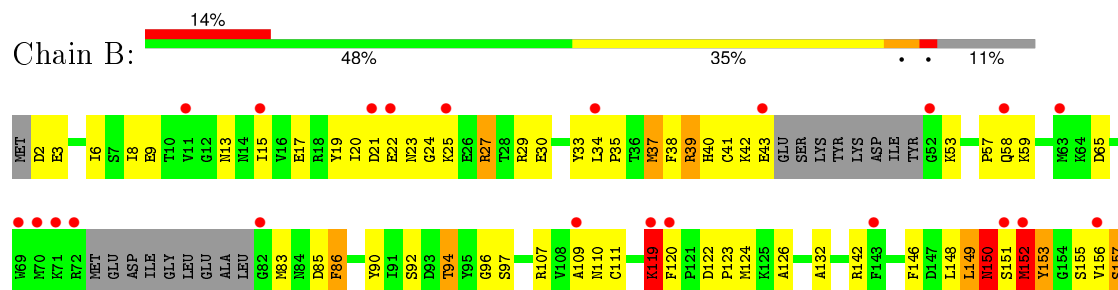
- Molecule 1: DNA (5'-D(*TP*TP*T)-3')



- Molecule 2: PROTEIN (DNA POLYMERASE (E.C.2.7.7.7))



- Molecule 2: PROTEIN (DNA POLYMERASE (E.C.2.7.7.7))



K158	W159	D160	A161	K162	L163	A164	A165	K166	L167	D168	K169	E170	G171	G172	D173	E174	I179	R182	V183	I184	Y185	F188	E191	R192	M196	I199	Q204	K205	R206	P207	A208	I209	F210	T211	G212	H213	H214	I215	E216	G217	F218	D219	V220	F221	M224	N225	R226	V227	K228	M229	I230		
L231	G232	E233	R234	S235	M236	K237	R238	F239	I242	G243	R244	V245	K246	S247	K248	L249	LEU	GLN	ASN	MET	TYR	GLY	S256	K257	E258	I259	Y260	S261	I262	D263	I267	F279	T280	N281	L282	F285	S286	L287	V290	A291	Q292	T295	K296	K297	G298	K299	L300	P301	Y302	D303	G304	P305	I306
R307	K308	L309	R310	E311	T312	R313	H314	Q315	R316	Y317	Y320	R321	I322	Q329	A330	I334	R335	G336	L340	V341	A348	K349	M350	V355	K356	S357	P358	I359	K360	F367	N368	S369	L370	LVS	GLY	GLU	HIS	LVS	VAL	ILE	PRO	GLN	GLN	GLY	SER	HIS	VAL	LVS	GLN	SER	PHE		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	114.70 Å 109.66 Å 70.27 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.00 – 2.20 59.16 – 2.49	Depositor EDS
% Data completeness (in resolution range)	(Not available) (6.00-2.20) 79.8 (59.16-2.49)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.49 (at 2.48 Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.222 , (Not available) (Not available) , (Not available)	Depositor DCC
R_{free} test set	NotAvailable	DCC
Wilson B-factor (Å ²)	44.1	Xtriage
Anisotropy	0.427	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 87.3	EDS
Estimated twinning fraction	0.013 for k,h,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 27185 reflections	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 6.93% 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, 0.375 respectively for untwinned datasets, and 0.333, 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%)	34	35
2	B	338/388 (87%)	294 (87%)	31 (9%)	13 (4%)	4	1
All	All	708/776 (91%)	638 (90%)	55 (8%)	15 (2%)	9	5

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%)	10	9
2	B	312/350 (89%)	278 (89%)	34 (11%)	8	7
All	All	647/700 (92%)	581 (90%)	66 (10%)	9	8

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

Mol	Chain	Res	Type
2	A	322	ILE
2	B	43	GLU
2	B	302	TYR
2	A	333	LYS
2	B	27	ARG

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

Mol	Chain	Res	Type
2	A	321	ASN
2	B	14	ASN
2	B	292	GLN
2	A	313	ASN
2	B	110	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 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

There are no such residues in this entry.

5.8 Polymer linkage issues

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%)	10.78	3 (100%) 0 0	20, 20, 20, 20	0
2	A	372/388 (95%)	0.88	32 (8%) 13 12	14, 38, 73, 80	0
2	B	346/388 (89%)	1.17	54 (15%) 3 2	23, 56, 80, 80	0
All	All	721/779 (92%)	1.06	89 (12%) 5 5	14, 46, 80, 80	0

The worst 5 of 89 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	372	GLY	18.3
1	S	3	DT	12.3
1	S	5	DT	10.2
1	S	4	DT	9.9
2	B	249	LEU	8.6

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 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(\AA^2)	Q<0.9
3	ZN	B	2000	1/1	0.47	0.27	-0.41	20,20,20,20	0
4	MN	B	3000	1/1	0.94	0.10	-0.98	20,20,20,20	0

6.5 Other polymers [i](#)

There are no such residues in this entry.