



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 2, 2026 – 04:12 AM UTC

PDB ID : 1SKS / pdb_00001sk
Title : Binary 3' complex of T7 DNA polymerase with a DNA primer/template containing a cis-syn thymine dimer on the template
Authors : Li, Y.; Dutta, S.; Doublié, S.; Bdour, H.M.; Taylor, J.S.; Ellenberger, T.
Deposited on : 2004-03-05
Resolution : 2.30 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 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

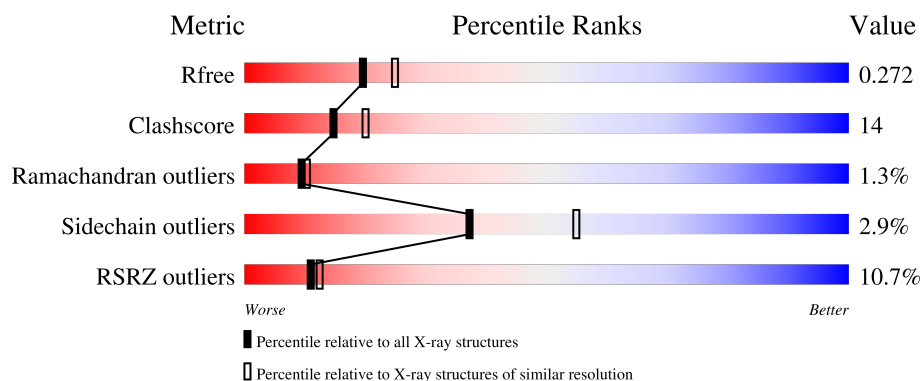
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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))
R_{free}	180053	6319 (2.30-2.30)
Clashscore	190562	6919 (2.30-2.30)
Ramachandran outliers	187476	6854 (2.30-2.30)
Sidechain outliers	187428	6854 (2.30-2.30)
RSRZ outliers	180081	6325 (2.30-2.30)

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	P	21	14% 19% 29% 5% 48%
2	T	25	4% 32% 12% 52%
3	A	698	10% 71% 25% ..
4	B	108	13% 74% 20% ...

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 6769 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 5'-D(*CP*GP*AP*AP*AP*AP*CP*GP*AP*C*GP*GP*CP*CP*AP*GP*TP*GP*CP*CP*(2DT))-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	P	11	Total	C	N	O	P	0	0	0
			199	96	31	62	10			

- Molecule 2 is a DNA chain called 5'-D(*CP*CP*CP*(TTD)P*AP*GP*GP*CP*AP*CP*T P*GP*GP*CP*CP*GP*TP*CP*GP*TP*TP*TP*TP*CP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	12	Total	C	N	O	P	0	0	0
			266	126	48	79	13			

- Molecule 3 is a protein called DNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	681	Total	C	N	O	S	0	0	0
			5308	3382	921	981	24			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	LYS	deletion	UNP P00581
A	?	-	ARG	deletion	UNP P00581
A	?	-	PHE	deletion	UNP P00581
A	?	-	GLY	deletion	UNP P00581
A	?	-	SER	deletion	UNP P00581
A	?	-	HIS	deletion	UNP P00581

- Molecule 4 is a protein called Thioredoxin 1.

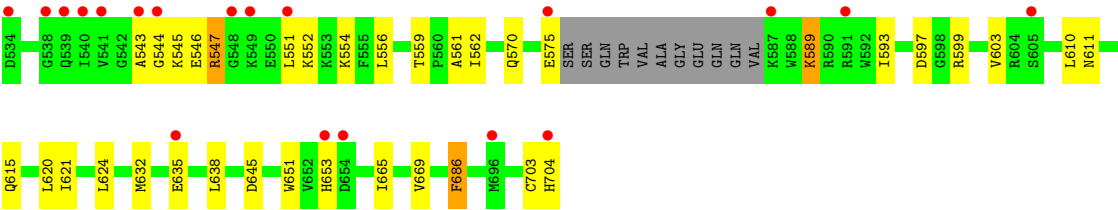
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	105	Total	C	N	O	S	0	0	0
			798	515	128	152	3			

- Molecule 5 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

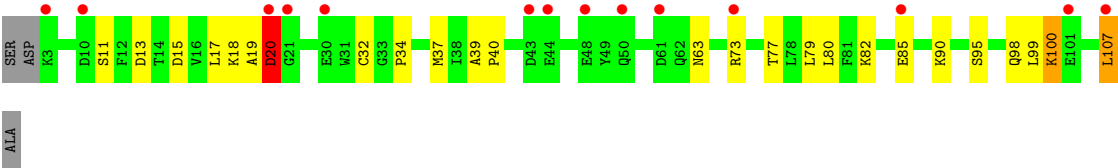
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total 1	Mg 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	P	7	Total 7	O 7	0	0
6	T	6	Total 6	O 6	0	0
6	A	172	Total 172	O 172	0	0
6	B	12	Total 12	O 12	0	0



● Molecule 4: Thioredoxin 1



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 ₁ 2 ₁ 2	Depositor
Cell constants a, b, c, α , β , γ	104.77Å 214.92Å 52.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.30 50.00 – 2.30	Depositor EDS
% Data completeness (in resolution range)	96.9 (50.00-2.30) 96.8 (50.00-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.91 (at 2.29Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.235 , 0.266 0.244 , 0.272	Depositor DCC
R_{free} test set	2546 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	35.7	Xtriage
Anisotropy	0.135	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 30.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	6769	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.79% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, 2DT, TTD

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	P	0.30	0/200	0.73	0/305
2	T	0.34	0/253	0.85	1/388 (0.3%)
3	A	0.45	0/5440	0.88	8/7382 (0.1%)
4	B	0.39	0/813	0.84	2/1104 (0.2%)
All	All	0.43	0/6706	0.87	11/9179 (0.1%)

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	P	0	1
2	T	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	589	LYS	N-CA-C	-7.14	104.70	113.41
3	A	127	ALA	N-CA-C	6.21	118.56	111.11
2	T	9	DC	C5'-C4'-C3'	-5.86	106.11	114.90
3	A	400	ALA	N-CA-C	5.79	121.22	114.62
3	A	9	ASN	N-CA-C	5.49	119.84	113.20

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	P	19	DC	Sidechain
2	T	8	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	P	199	0	115	10	0
2	T	266	0	147	4	0
3	A	5308	0	5090	151	0
4	B	798	0	805	18	0
5	A	1	0	0	0	0
6	A	172	0	0	2	0
6	B	12	0	0	0	0
6	P	7	0	0	0	0
6	T	6	0	0	1	0
All	All	6769	0	6157	177	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 177 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:P:15:DA:H2''	1:P:16:DG:H5'	1.58	0.83
2:T:5:TTD:H2'	2:T:5:TTD:H5R2	1.58	0.83
3:A:144:LYS:O	3:A:148:GLU:HG2	1.79	0.81
3:A:328:PRO:HB3	4:B:73:ARG:HH12	1.46	0.81
3:A:35:ARG:HB3	3:A:36:PRO:HD2	1.69	0.73

There are no symmetry-related clashes.

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	675/698 (97%)	639 (95%)	27 (4%)	9 (1%)	9	10
4	B	103/108 (95%)	97 (94%)	5 (5%)	1 (1%)	12	15
All	All	778/806 (96%)	736 (95%)	32 (4%)	10 (1%)	9	10

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	156	ASP
3	A	543	ALA
3	A	359	LYS
3	A	360	GLY
4	B	20	ASP

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	533/579 (92%)	520 (98%)	13 (2%)	43	62
4	B	84/87 (97%)	79 (94%)	5 (6%)	17	25
All	All	617/666 (93%)	599 (97%)	18 (3%)	37	55

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	B	80	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	B	107	LEU
4	B	100	LYS
3	A	501	LEU
4	B	20	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	611	ASN
3	A	667	GLN
4	B	50	GLN
3	A	704	HIS
3	A	506	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are 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	TTD	T	5	2	42,45,46	4.13	11 (26%)	61,74,77	3.47	31 (50%)
1	2DT	P	21	1,2	17,20,21	0.34	0	22,28,31	0.48	0

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	TTD	T	5	2	-	9/22/109/110	0/5/6/6
1	2DT	P	21	1,2	-	2/7/18/19	0/2/2/2

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	5	TTD	C5T-C6T	-21.16	1.32	1.55
2	T	5	TTD	C5-C6	-11.04	1.43	1.55
2	T	5	TTD	C1'-N1	5.93	1.53	1.45
2	T	5	TTD	C1R-N1T	4.76	1.51	1.45
2	T	5	TTD	C5A-C5	4.22	1.62	1.53

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	5	TTD	C5T-C5-C4	8.17	139.37	112.84
2	T	5	TTD	O4T-C4T-C5T	8.06	130.21	122.71
2	T	5	TTD	O4'-C1'-N1	-7.69	99.54	108.65
2	T	5	TTD	O5R-PB-O5P	-7.14	80.64	108.94
2	T	5	TTD	C5-C5T-C6T	6.94	96.48	88.36

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	P	21	2DT	O4'-C4'-C5'-O5'
2	T	5	TTD	C3'-C4'-C5R-O5R
2	T	5	TTD	O4R-C4'-C5R-O5R
2	T	5	TTD	C2'-C1'-N1-C6
2	T	5	TTD	C2'-C1'-N1-C2

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	T	5	TTD	2	0

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is 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	P	10/21 (47%)	1.75	3 (30%) 1 1	37, 49, 76, 79	0
2	T	11/25 (44%)	0.69	1 (9%) 15 16	33, 47, 69, 72	0
3	A	681/698 (97%)	0.66	68 (9%) 12 14	16, 36, 60, 69	0
4	B	105/108 (97%)	0.98	14 (13%) 7 8	33, 47, 59, 62	0
All	All	807/852 (94%)	0.72	86 (10%) 11 12	16, 38, 60, 79	0

The worst 5 of 86 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	P	12	DG	5.4
3	A	303	ASN	5.3
1	P	11	DG	4.8
3	A	704	HIS	4.4
4	B	21	GLY	4.1

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. 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(Å ²)	Q<0.9
2	TTD	T	5	40/41	0.45	0.24	72,79,86,87	0
1	2DT	P	21	19/20	0.94	0.11	39,41,43,44	0

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands

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($\text{\AA}^2$)	Q<0.9
5	MG	A	4003	1/1	0.75	0.33	68,68,68,68	0

6.5 Other polymers

There are no such residues in this entry.