



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 01:56 AM UTC

PDB ID : 2JGU / pdb_00002jgu
Title : crystal structure of DNA-directed DNA polymerase
Authors : Kim, D.U.; Cho, H.S.
Deposited on : 2007-02-15
Resolution : 2.60 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
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

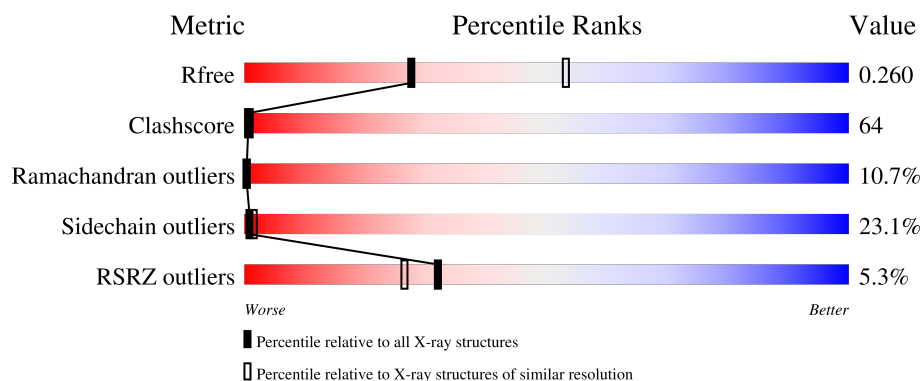
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	4008 (2.60-2.60)
Clashscore	190562	4347 (2.60-2.60)
Ramachandran outliers	187476	4277 (2.60-2.60)
Sidechain outliers	187428	4277 (2.60-2.60)
RSRZ outliers	180081	4008 (2.60-2.60)

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	A	775	5% 22% 35% 27% 8% 8%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5859 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	712	Total	C	N	O	S	0	0	1
			5826	3777	963	1072	14			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	96	LEU	ILE	conflict	UNP P61875
A	169	ARG	LYS	conflict	UNP P61875
A	183	SER	VAL	conflict	UNP P61875
A	186	THR	SER	conflict	UNP P61875
A	188	LYS	ARG	conflict	UNP P61875
A	265	ARG	THR	conflict	UNP P61875
A	266	THR	ARG	conflict	UNP P61875
A	406	TYR	PHE	conflict	UNP P61875
A	407	LYS	ARG	conflict	UNP P61875
A	408	SER	ALA	conflict	UNP P61875

- Molecule 2 is MANGANESE (II) ION (CCD ID: MN) (formula: Mn).

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

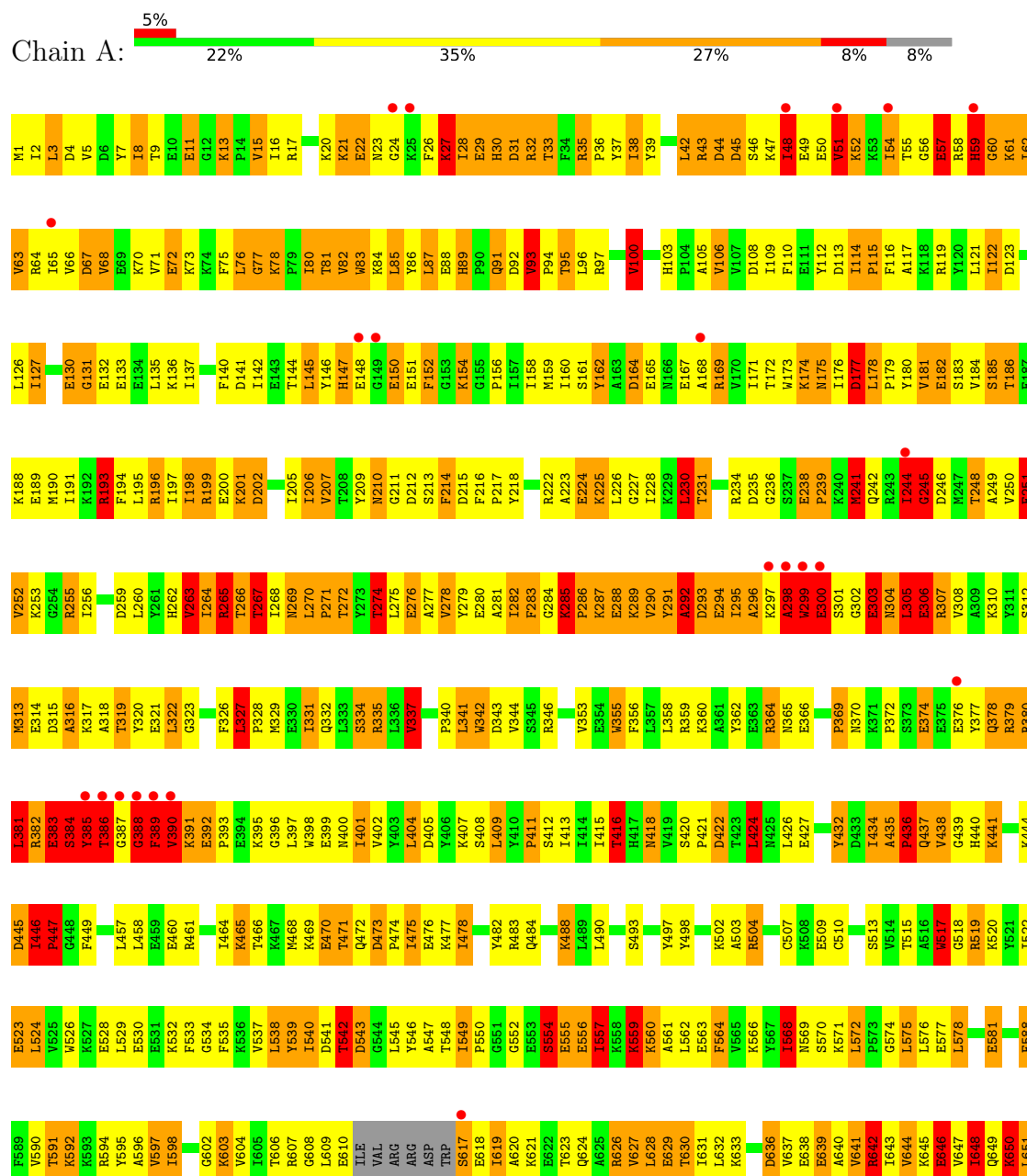
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	31	Total	O	0	0
			31	31		

3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 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.

• Molecule 1: DNA POLYMERASE



ILE	SER	A652
	ASN	M653
LYS	ARG	Y654
LYS	ALA	E655
SER	ILE	I656
	LEU	P657
LYS	ALA	P658
	GLU	E659
SER	GLU	K660
	TYR	L661
LYS	ASP	A662
	PRO	ILE
LYS	LYS	TYR
LYS	HIS	GLU
LYS	HIS	GLN
SER	LYS	ILE
	TYR	THR
LYS	D729	ARG
	A730	PRO
SER	E731	LEU
	Y732	HIS
LYS	Y733	GLU
	I734	TYR
SER	E735	LYS
	N736	ALA
LYS	Q737	ILE
	Y738	GLY
SER	L739	P679
	P740	H680
LYS	A741	V681
	Y742	A682
SER	L743	V683
	R744	A684
LYS	L745	K685
	L746	K686
SER	E747	L687
	G748	A688
LYS	F749	A689
	Y750	K690
SER	Y751	G691
	R752	V692
LYS	K753	K693
	E754	L694
SER	D755	K695
	L756	P696
LYS	R757	G697
	Y758	V698
SER	GLN	V699
	LYS	I700
LYS	THR	G701
	ARG	Y702
SER	GLN	I703
	VAL	V704
LYS	GLY	LEU
	LEU	ARG
SER	THR	GLY
	SER	ASP
LYS	TRP	GLY
	LEU	PRO
SER	ASN	ILE
	THR	THR

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	92.18Å 127.92Å 89.21Å 90.00° 109.11° 90.00°	Depositor
Resolution (Å)	15.00 – 2.60 15.00 – 2.60	Depositor EDS
% Data completeness (in resolution range)	94.2 (15.00-2.60) 93.6 (15.00-2.60)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.57 (at 2.42Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.275 , 0.271 (Not available) , 0.260	Depositor DCC
R_{free} test set	1413 reflections (4.16%)	wwPDB-VP
Wilson B-factor (Å ²)	45.6	Xtriage
Anisotropy	0.247	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 42.9	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	0.91	EDS
Total number of atoms	5859	wwPDB-VP
Average B, all atoms (Å ²)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.80% 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: 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	A	1.18	42/5953 (0.7%)	1.92	256/8024 (3.2%)

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	A	0	2

The worst 5 of 42 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	285	LYS	C-N	-11.54	1.06	1.33
1	A	299	TRP	CA-C	-8.99	1.40	1.52
1	A	386	THR	CA-C	8.43	1.64	1.52
1	A	385	TYR	C-N	8.30	1.45	1.33
1	A	384	SER	CA-C	-8.23	1.41	1.52

The worst 5 of 256 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	684	ALA	N-CA-C	-18.63	82.65	109.11
1	A	554	SER	N-CA-C	15.76	131.94	113.19
1	A	154	LYS	N-CA-C	-14.57	95.52	114.31
1	A	641	VAL	N-CA-C	-14.39	96.18	110.72
1	A	392	GLU	CA-C-N	14.12	133.88	119.76

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	162	TYR	Sidechain
1	A	389	PHE	Peptide

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	A	5826	0	5919	757	0
2	A	2	0	0	0	0
3	A	31	0	0	5	0
All	All	5859	0	5919	757	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 64.

The worst 5 of 757 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:389:PHE:CE1	1:A:519:ARG:HG2	1.30	1.65
1:A:389:PHE:HE1	1:A:519:ARG:CG	1.08	1.63
1:A:251:GLU:CG	1:A:251:GLU:CB	1.75	1.54
1:A:617:SER:OG	1:A:737:GLN:NE2	1.65	1.26
1:A:389:PHE:O	1:A:390:VAL:HG23	1.40	1.22

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	684/775 (88%)	503 (74%)	108 (16%)	73 (11%)	0 0

5 of 73 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	48	ILE
1	A	59	HIS
1	A	67	ASP
1	A	76	LEU
1	A	131	GLY

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
1	A	624/682 (92%)	480 (77%)	144 (23%)	1 1

5 of 144 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	629	GLU
1	A	752	ARG
1	A	646	GLU
1	A	694	ILE
1	A	252	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	437	GLN
1	A	462	GLN
1	A	737	GLN
1	A	569	ASN
1	A	736	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	285:LYS	C	286:PRO	N	1.06

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	A	712/775 (91%)	0.30	38 (5%) 32 26	21, 59, 91, 101	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	299	TRP	7.0
1	A	687	LEU	5.7
1	A	389	PHE	4.8
1	A	390	VAL	4.5
1	A	617	SER	4.0

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 oligosaccharides 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(Å ²)	Q<0.9
2	MN	A	1758	1/1	0.57	0.27	61,61,61,61	0
2	MN	A	1759	1/1	0.91	0.07	61,61,61,61	0

6.5 Other polymers [i](#)

There are no such residues in this entry.