



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 6, 2021 – 05:07 am GMT

PDB ID : 7OM3
Title : Crystal structure of KOD DNA Polymerase in a binary complex with Hypoxanthine containing template
Authors : Betz, K.; Kropp, H.M.; Diederichs, K.; Marx, A.
Deposited on : 2021-05-21
Resolution : 1.92 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
Mogul : 1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.23.2
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.23.2

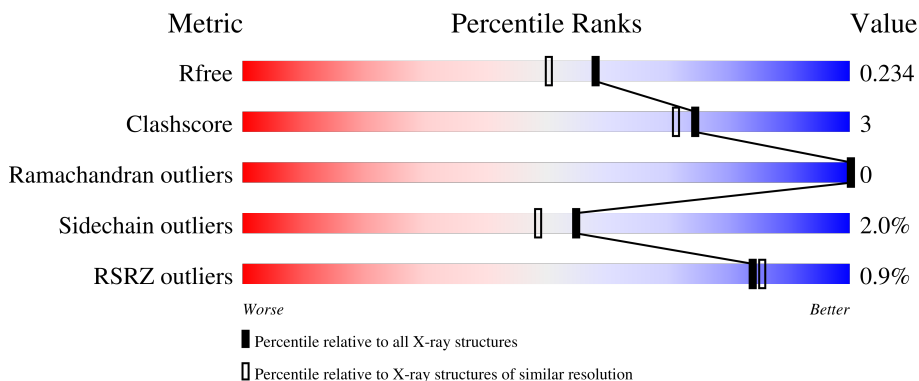
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.92 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}	130704	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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	774	
2	T	21	
3	P	12	

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 14024 atoms, of which 6700 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,DNA polymerase,DNA polymerase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	759	12540	4024	6292	1056	1152	16	0	2	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	141	ALA	ASP	engineered mutation	UNP P77933
A	143	ALA	GLU	engineered mutation	UNP P77933

- Molecule 2 is a DNA chain called 21nt Template.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	P			
2	T	21	666	205	238	76	127	20	0	0	0

- Molecule 3 is a DNA chain called Primer.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	P			
3	P	12	373	114	134	48	66	11	0	0	0

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
4	A	1	17	4	10	3	0	0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

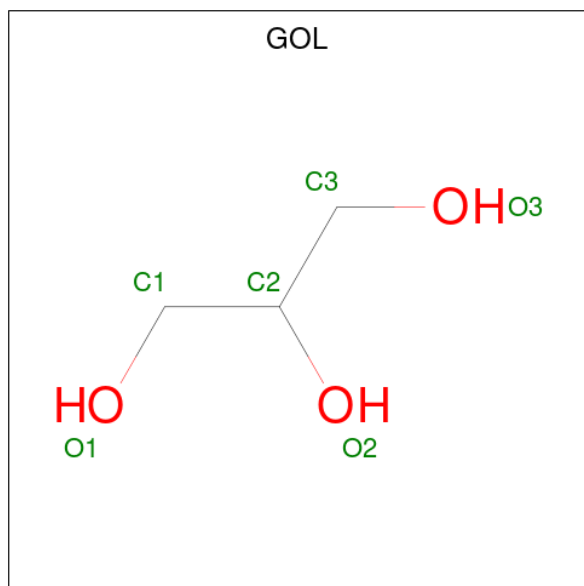


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
5	A	1	10	2	6	2	0	0
5	A	1	10	2	6	2	0	0
5	A	1	10	2	6	2	0	0

- Molecule 6 is BROMIDE ION (three-letter code: BR) (formula: Br).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	2	Total Br 2 2	0	0

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C H O 14 3 8 3	0	0

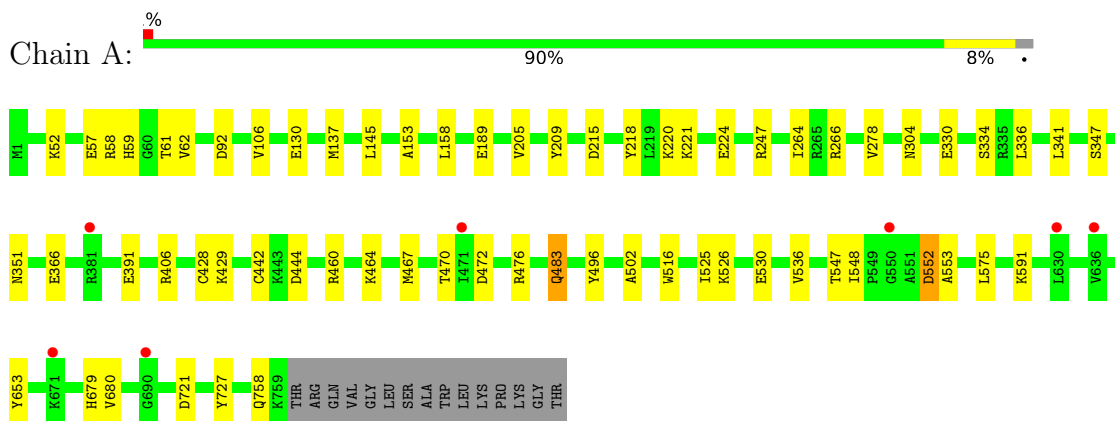
- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	337	Total O 337 337	0	0
8	T	34	Total O 34 34	0	0
8	P	11	Total O 11 11	0	0

3 Residue-property plots [i](#)

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.

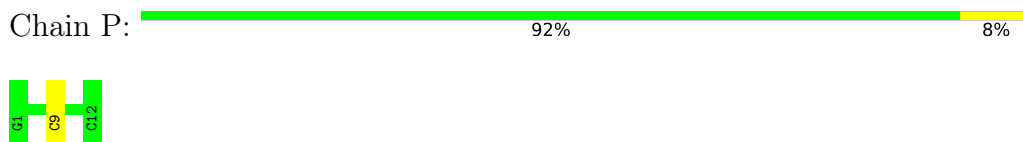
- Molecule 1: DNA polymerase,DNA polymerase,DNA polymerase



- Molecule 2: 21nt Template



- Molecule 3: Primer



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	113.36Å 142.94Å 65.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.45 – 1.92 48.32 – 1.91	Depositor EDS
% Data completeness (in resolution range)	92.4 (44.45-1.92) 79.4 (48.32-1.91)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.24 (at 1.91Å)	Xtrriage
Refinement program	PHENIX 1.19_4085, PHENIX 1.19_4085	Depositor
R, R_{free}	0.183 , 0.234 0.183 , 0.234	Depositor DCC
R_{free} test set	1979 reflections (2.57%)	wwPDB-VP
Wilson B-factor (Å ²)	26.9	Xtrriage
Anisotropy	0.352	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	14024	wwPDB-VP
Average B, all atoms (Å ²)	50.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.38% 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: BR, GOL, PEG, DOC, EDO

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	0.59	1/6397 (0.0%)	0.71	0/8628
2	T	1.30	1/454 (0.2%)	1.46	8/697 (1.1%)
3	P	1.12	1/248 (0.4%)	1.09	1/380 (0.3%)
All	All	0.68	3/7099 (0.0%)	0.81	9/9705 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	P	9	DC	C3'-O3'	-5.69	1.36	1.44
2	T	8	DC	C1'-N1	5.48	1.56	1.49
1	A	189	GLU	CG-CD	5.24	1.59	1.51

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	5	DC	O5'-P-OP1	-10.42	96.32	105.70
2	T	14	DC	OP2-P-O3'	6.70	119.94	105.20
2	T	15	DC	O5'-P-OP2	-6.02	100.29	105.70
2	T	19	DG	O5'-P-OP2	-5.87	100.42	105.70
2	T	2	DA	OP1-P-O3'	5.86	118.10	105.20

There are no chirality outliers.

There are no planarity outliers.

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	6248	6292	6278	37	0
2	T	428	238	238	4	0
3	P	239	134	134	0	0
4	A	7	10	10	2	0
5	A	12	18	18	0	0
6	A	2	0	0	1	0
7	A	6	8	8	1	0
8	A	337	0	0	12	0
8	P	11	0	0	0	0
8	T	34	0	0	1	0
All	All	7324	6700	6686	42	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 3.

The worst 5 of 42 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)
2:T:21:DC:OP2	8:T:101:HOH:O	1.91	0.88
1:A:266:ARG:NH2	1:A:330:GLU:OE1	2.16	0.78
1:A:406:ARG:NH1	8:A:1005:HOH:O	2.23	0.72
1:A:428:CYS:HG	1:A:442:CYS:HG	0.72	0.67
1:A:721:ASP:OD2	8:A:1001:HOH:O	2.14	0.65

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	759/774 (98%)	743 (98%)	16 (2%)	0	100 100

There are no Ramachandran outliers to report.

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	662/672 (98%)	649 (98%)	13 (2%)	55 49

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	347	SER
1	A	351	ASN
1	A	575	LEU
1	A	516	TRP
1	A	552	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 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	A	759/774 (98%)	-0.02	7 (0%) 84 85	27, 41, 71, 122	0
2	T	20/21 (95%)	-0.26	0 100 100	35, 49, 106, 107	0
3	P	11/12 (91%)	-0.12	0 100 100	36, 48, 76, 80	0
All	All	790/807 (97%)	-0.03	7 (0%) 84 85	27, 42, 72, 122	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	550	GLY	4.3
1	A	471	ILE	2.4
1	A	381	ARG	2.3
1	A	690	GLY	2.3
1	A	671	LYS	2.2

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
3	DOC	P	12	18/19	0.97	0.12	29,36,45,47	0

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(Å ²)	Q<0.9
4	PEG	A	901	7/7	0.67	0.18	48,57,71,71	0
7	GOL	A	907	6/6	0.80	0.15	33,52,59,67	0
5	EDO	A	903	4/4	0.86	0.16	50,60,71,71	0
5	EDO	A	904	4/4	0.91	0.19	44,62,77,77	0
5	EDO	A	902	4/4	0.95	0.11	37,48,58,58	0
6	BR	A	906	1/1	0.96	0.09	93,93,93,93	0
6	BR	A	905	1/1	0.98	0.15	97,97,97,97	0

6.5 Other polymers [i](#)

There are no such residues in this entry.