



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 27, 2014 – 05:10 AM GMT

PDB ID : 4FVM
Title : Crystal structure of yeast DNA polymerase alpha
Authors : Perera, R.L.; Pellegrini, L.
Deposited on : 2012-06-29
Resolution : 2.30 Å(reported)

This is a full wwPDB validation 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/ValidationPDFNotes.html>

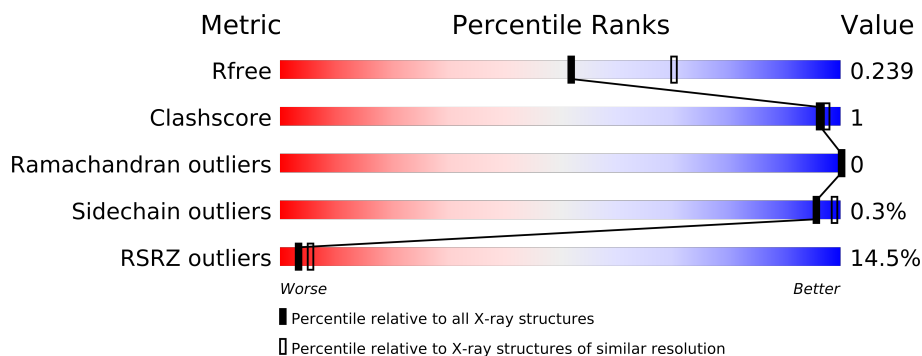
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

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}	66092	2929 (2.30-2.30)
Clashscore	79885	3679 (2.30-2.30)
Ramachandran outliers	78287	3642 (2.30-2.30)
Sidechain outliers	78261	3641 (2.30-2.30)
RSRZ outliers	66119	2930 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	910	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13690 atoms, of which 6791 are hydrogens and 0 are deuterium.

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 alpha catalytic subunit A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	829	Total	C	H	N	O	S	0	12	0
			13474	4229	6791	1158	1248	48			

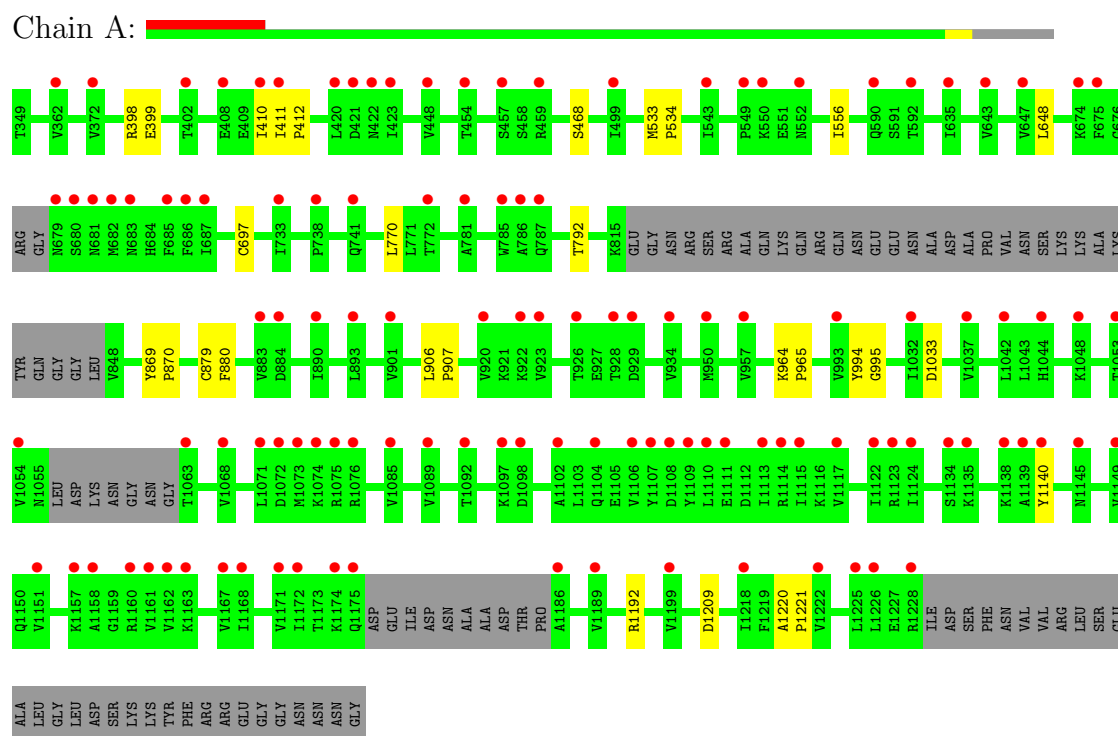
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	216	Total	O	0	0
			216	216		

3 Residue-property plots

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 ($RSRZ > 2$). 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 alpha catalytic subunit A



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	74.38Å 127.14Å 74.57Å 90.00° 104.78° 90.00°	Depositor
Resolution (Å)	36.51 – 2.30 36.51 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.9 (36.51-2.30) 99.0 (36.51-2.30)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.87 (at 2.29Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1078)	Depositor
R, R_{free}	0.205 , 0.236 0.206 , 0.239	Depositor DCC
R_{free} test set	2969 reflections (5.05%)	DCC
Wilson B-factor (Å ²)	42.1	Xtriage
Anisotropy	0.861	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 51.3	EDS
Estimated twinning fraction	0.022 for l,-k,h	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 58855 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	13690	wwPDB-VP
Average B, all atoms (Å ²)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.66% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

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	0.22	0/6843	0.38	0/9253

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6683	6791	0	15	0
2	A	216	0	0	1	0
All	All	6899	6791	0	15	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 1.

All (15) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:792:THR:O	2:A:1444:HOH:O	2.09	0.70
1:A:906:LEU:HB3	1:A:907:PRO:HD3	1.89	0.54
1:A:411:ILE:HB	1:A:412:PRO:HD3	1.90	0.53
1:A:410:ILE:HD12	1:A:411:ILE:N	2.26	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:398:ARG:NH2	1:A:468:SER:O	2.42	0.51
1:A:1220:ALA:HB3	1:A:1221:PRO:HD3	1.94	0.50
1:A:697:CYS:SG	1:A:770:LEU:HD23	2.55	0.46
1:A:964:LYS:N	1:A:965:PRO:CD	2.79	0.46
1:A:869:TYR:N	1:A:870:PRO:HD2	2.32	0.44
1:A:1192:ARG:NH2	1:A:1209:ASP:OD2	2.52	0.43
1:A:994:TYR:CG	1:A:995:GLY:N	2.87	0.42
1:A:879:CYS:SG	1:A:880:PHE:N	2.93	0.42
1:A:398:ARG:HG3	1:A:399:GLU:N	2.35	0.42
1:A:533:MET:HB2	1:A:534:PRO:HD2	2.02	0.42
1:A:556:ILE:HD11	1:A:648:LEU:HA	2.03	0.41

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	
1	A	831/910 (91%)	809 (97%)	22 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	
1	A	764/818 (93%)	762 (100%)	2 (0%)	96	99

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1033	ASP
1	A	1140	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	385	GLN
1	A	428	GLN
1	A	728	HIS
1	A	752	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no ligands in this entry.

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	829/910 (91%)	1.00	120 (14%) 3 5	32, 58, 99, 129	0

All (120) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1073	MET	9.6
1	A	1161	VAL	8.1
1	A	1139	ALA	6.8
1	A	1072	ASP	6.6
1	A	674	LYS	6.4
1	A	681	ASN	6.4
1	A	550	LYS	6.0
1	A	682	MET	5.7
1	A	1106	VAL	5.6
1	A	1092	THR	5.4
1	A	1186	ALA	5.3
1	A	1145	ASN	4.9
1	A	1111	GLU	4.9
1	A	1115	ILE	4.8
1	A	1162	VAL	4.7
1	A	680	SER	4.7
1	A	1113	ILE	4.6
1	A	1149	VAL	4.6
1	A	1160	ARG	4.6
1	A	1104	GLN	4.5
1	A	679	ASN	4.5
1	A	1068	VAL	4.5
1	A	928	THR	4.3
1	A	1117	VAL	4.2
1	A	883	VAL	4.0
1	A	1108	ASP	4.0
1	A	1074	LYS	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1071	LEU	4.0
1	A	1075	ARG	4.0
1	A	901	VAL	4.0
1	A	923	VAL	3.9
1	A	410	ILE	3.9
1	A	957	VAL	3.8
1	A	683	ASN	3.8
1	A	1122	ILE	3.7
1	A	1109	TYR	3.6
1	A	1175	GLN	3.6
1	A	1225	LEU	3.6
1	A	459	ARG	3.6
1	A	1163	LYS	3.5
1	A	1140	TYR	3.5
1	A	1135	LYS	3.4
1	A	1138	LYS	3.4
1	A	687	ILE	3.4
1	A	1042	LEU	3.4
1	A	686	PHE	3.4
1	A	738	PRO	3.4
1	A	993	VAL	3.4
1	A	1114	ARG	3.4
1	A	920	VAL	3.3
1	A	1054	VAL	3.3
1	A	922	LYS	3.3
1	A	1097	LYS	3.3
1	A	890	ILE	3.2
1	A	685	PHE	3.2
1	A	926	THR	3.2
1	A	1102	ALA	3.1
1	A	1076	ARG	3.1
1	A	592	THR	3.1
1	A	1167	VAL	3.1
1	A	647	VAL	3.1
1	A	421	ASP	3.0
1	A	420	LEU	3.0
1	A	1157	LYS	3.0
1	A	1174	LYS	3.0
1	A	1063	THR	2.9
1	A	1222	VAL	2.9
1	A	408	GLU	2.9
1	A	643	VAL	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1123	ARG	2.8
1	A	1044[A]	HIS	2.7
1	A	549	PRO	2.7
1	A	1151	VAL	2.7
1	A	423	ILE	2.7
1	A	741	GLN	2.7
1	A	1085	VAL	2.6
1	A	1089	VAL	2.6
1	A	1158	ALA	2.6
1	A	590	GLN	2.6
1	A	1218	ILE	2.6
1	A	448	VAL	2.6
1	A	1226	LEU	2.5
1	A	1189	VAL	2.5
1	A	422	ASN	2.5
1	A	552	ASN	2.5
1	A	1098	ASP	2.5
1	A	785	TRP	2.5
1	A	1134	SER	2.5
1	A	675	PHE	2.4
1	A	1053	THR	2.4
1	A	786	ALA	2.4
1	A	1168	ILE	2.4
1	A	1037	VAL	2.4
1	A	781	ALA	2.3
1	A	934	VAL	2.3
1	A	402	THR	2.3
1	A	1110	LEU	2.3
1	A	454	THR	2.3
1	A	1124	ILE	2.3
1	A	929	ASP	2.2
1	A	893	LEU	2.2
1	A	1171	VAL	2.2
1	A	411	ILE	2.2
1	A	543	ILE	2.2
1	A	635	ILE	2.2
1	A	1172	ILE	2.2
1	A	457	SER	2.2
1	A	787[A]	GLN	2.1
1	A	499	ILE	2.1
1	A	1228	ARG	2.1
1	A	1199	VAL	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	733	ILE	2.1
1	A	362	VAL	2.0
1	A	1048	LYS	2.0
1	A	1107	TYR	2.0
1	A	372	VAL	2.0
1	A	884	ASP	2.0
1	A	1032	ILE	2.0
1	A	950	MET	2.0
1	A	772	THR	2.0

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

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.