Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 04:42 PM 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 X-ray Structure 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/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:

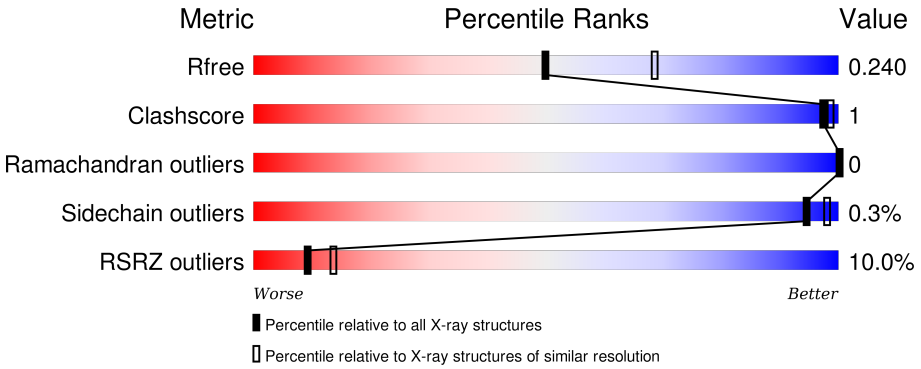
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

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}	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (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. 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 <=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	910	9% 88% · 9%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 13690 atoms, of which 6791 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 alpha catalytic subunit A.

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

- Molecule 2 is water.

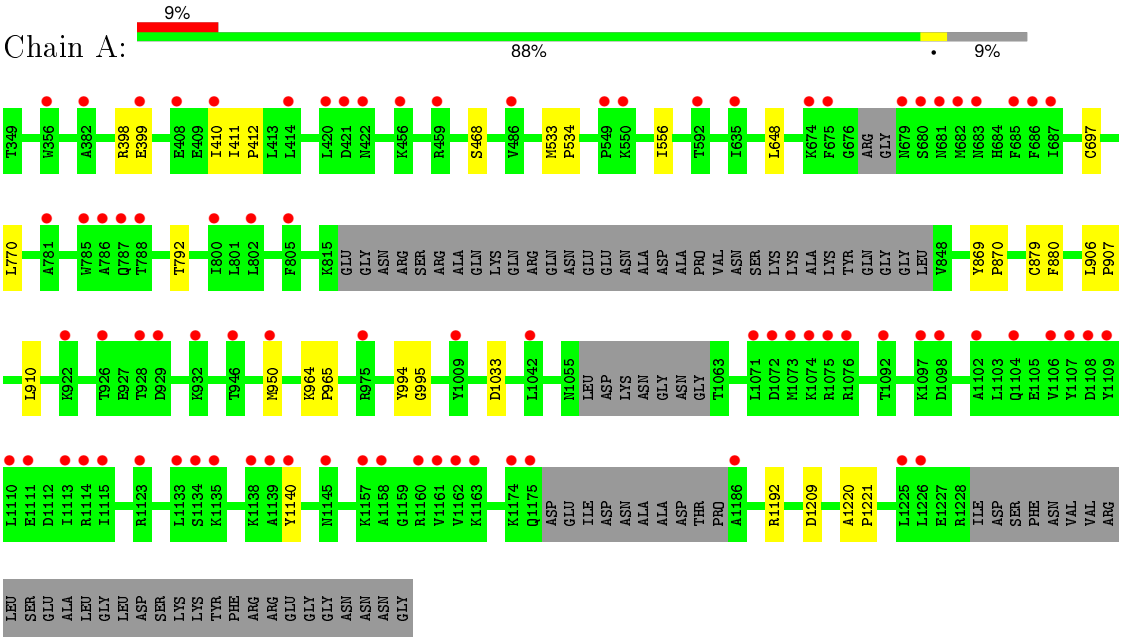
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	216	216	216	0	0

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 (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

Chain A:



4 Data and refinement statistics (i)

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
$< I/\sigma(I) >$ [1]	2.87 (at 2.29Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1078)	Depositor
R, R_{free}	0.205 , 0.236 0.217 , 0.240	Depositor DCC
R_{free} test set	2974 reflections (5.32%)	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 , 52.4	EDS
Estimated twinning fraction	0.022 for l,-k,h	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	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.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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

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 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	A	6683	6791	6743	16	0
2	A	216	0	0	1	0
All	All	6899	6791	6743	16	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 1.

All (16) 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:792:THR:O	2:A:1444:HOH:O	2.09	0.70
1:A:906:LEU:HB3	1:A:907:PRO:HD3	1.89	0.53
1:A:411:ILE:HB	1:A:412:PRO:HD3	1.90	0.53
1:A:398:ARG:NH2	1:A:468:SER:O	2.42	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:410:ILE:HD12	1:A:411:ILE:N	2.26	0.51
1:A:1220:ALA:HB3	1:A:1221:PRO:HD3	1.95	0.49
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:556:ILE:HD11	1:A:648:LEU:HA	2.01	0.42
1:A:879:CYS:SG	1:A:880:PHE:N	2.93	0.42
1:A:910:LEU:HD21	1:A:950:MET:HG3	2.03	0.41
1:A:398:ARG:HG3	1:A:399:GLU:N	2.36	0.41
1:A:533:MET:HB2	1:A:534:PRO:HD2	2.03	0.41

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	831/910 (91%)	809 (97%)	22 (3%)	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	764/818 (93%)	762 (100%)	2 (0%)	94 98

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 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

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, 95^{th} 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%)	0.77	83 (10%)	9	14	32, 57, 98, 129	0

All (83) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1073	MET	9.1
1	A	1161	VAL	6.5
1	A	681	ASN	6.3
1	A	1072	ASP	5.5
1	A	550	LYS	5.4
1	A	674	LYS	5.2
1	A	1104	GLN	5.1
1	A	1111	GLU	4.9
1	A	682	MET	4.8
1	A	680	SER	4.7
1	A	1145	ASN	4.6
1	A	1160	ARG	4.3
1	A	1139	ALA	4.2
1	A	1108	ASP	4.2
1	A	1106	VAL	4.1
1	A	1186	ALA	4.1
1	A	679	ASN	4.0
1	A	928	THR	3.8
1	A	1092	THR	3.7
1	A	683	ASN	3.6
1	A	459	ARG	3.6
1	A	1074	LYS	3.5
1	A	1109	TYR	3.4
1	A	1163	LYS	3.3
1	A	1115	ILE	3.3
1	A	592	THR	3.3
1	A	1225	LEU	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	420	LEU	3.1
1	A	1114	ARG	3.1
1	A	1075	ARG	3.0
1	A	1175	GLN	3.0
1	A	1042	LEU	3.0
1	A	1071	LEU	2.9
1	A	922	LYS	2.9
1	A	421	ASP	2.9
1	A	1113	ILE	2.9
1	A	1158	ALA	2.8
1	A	785	TRP	2.8
1	A	926	THR	2.8
1	A	1157	LYS	2.8
1	A	408	GLU	2.8
1	A	410	ILE	2.8
1	A	1135	LYS	2.8
1	A	1102	ALA	2.7
1	A	686	PHE	2.6
1	A	1123	ARG	2.6
1	A	414	LEU	2.6
1	A	1140	TYR	2.6
1	A	781	ALA	2.6
1	A	1138	LYS	2.6
1	A	1097	LYS	2.5
1	A	675	PHE	2.5
1	A	929	ASP	2.5
1	A	787[A]	GLN	2.4
1	A	1098	ASP	2.4
1	A	422	ASN	2.4
1	A	549	PRO	2.4
1	A	1076	ARG	2.4
1	A	1174	LYS	2.3
1	A	382	ALA	2.3
1	A	1226	LEU	2.3
1	A	635	ILE	2.3
1	A	1134	SER	2.3
1	A	805	PHE	2.2
1	A	1110	LEU	2.2
1	A	356	TRP	2.2
1	A	786	ALA	2.2
1	A	1107	TYR	2.2
1	A	685	PHE	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	975	ARG	2.2
1	A	800	ILE	2.1
1	A	1162	VAL	2.1
1	A	950	MET	2.1
1	A	399	GLU	2.1
1	A	486	VAL	2.1
1	A	1009	TYR	2.1
1	A	802	LEU	2.0
1	A	1133	LEU	2.0
1	A	456	LYS	2.0
1	A	932	LYS	2.0
1	A	687	ILE	2.0
1	A	788[A]	THR	2.0
1	A	946	THR	2.0

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 carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.