wwPDB X-ray Structure Validation Summary Report (i)

Feb 6, 2024 – 11:31 AM EST

PDB ID	:	2FLP
Title	:	Binary complex of the catalytic core of human DNA polymerase iota with DNA (template G)
Authors	:	Nair, D.T.; Johnson, R.E.; Prakash, L.; Prakash, S.; Aggarwal, A.K.
Deposited on	:	2006-01-06
Resolution	:	2.40 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 (i) 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 (i)) were used in the production of this report:

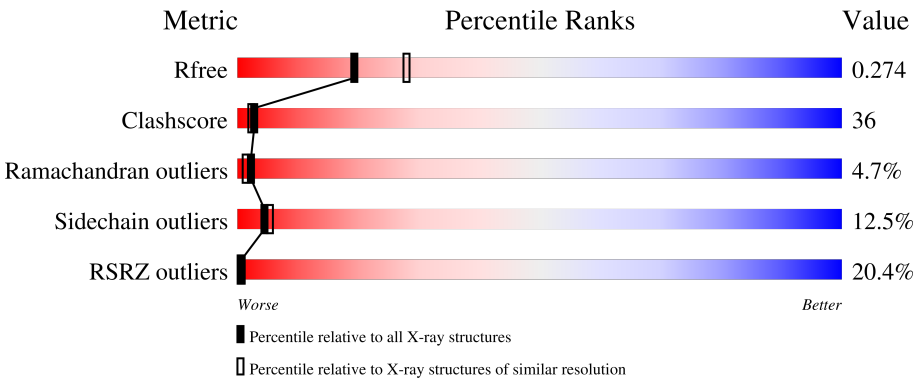
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 <=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	T	11	9%; 36%, 36%, 9%, 18%
2	P	7	57%, 43%
3	A	420	18%; 43%, 28%, 13%, •, 13%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3250 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 DNA template strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	T	9	183	88	32	55	8	0	0	0

- Molecule 2 is a DNA chain called DNA primer strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	P	7	139	67	29	37	6	0	0	0

- Molecule 3 is a protein called DNA polymerase iota.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	367	2813	1768	495	529	21	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	T	6	6	6	0	0
4	P	2	2	2	0	0
4	A	107	107	107	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 template strand

Chain T: 9% 36% 36% 9% 18%

DT DC T839 G840 G841 G842 G843 T844 C845 C846 T847

• Molecule 2: DNA primer strand

Chain P: 57% 43%

A867 G868 G869 A870 C871 C872 C873

• Molecule 3: DNA polymerase iota

Chain A: 18% 43% 28% 13% • 13%

MET GLU LEU ALA ASP VAL GLY ALA ALA ALA SER SER GLN GLY VAL HIS ASP GLN VAL LEU PRO THR PRO ASN ALA S26 S27 R28 V31 H32 V33 D36 C37 F38 Y39 E43 M44 L50 K51 D52 K53 P54 L55 Q58 Q59 K60 Y61 L62 V63 C66 N67 Y68 E69 A70

R71 K72 L73 G74 V75 K76 K77 L78 M79 N80 V81 R82 D83 A84 K85 E86 K87 Q90 L91 V92 L93 V94 N95 D98 R101 Y102 R103 E104 M105 S106 Y107 K108 E111 L112 S117 P118 V119 V120 E121 R122 L123 G124 F125 D126 E127 N128 F129 V130 E134 M135 V136 E137 Q141

Q142 L143 Q144 S145 D146 E147 L148 S149 V153 Y158 N159 N164 H170 I171 R172 S177 E182 M183 R184 E185 Y188 N189 C196 N202 K203 L204 L205 A206 K207 F213 K214 P215 N216 Q217 E224 L229 I230 H231 S232 H235 I236 Y244 K245 T246 A247 K248

C249 L250 E251 A252 L253 G254 I255 S265 P266 K267 I268 E272 Q279 K283 F286 G287 E288 D289 N290 S291 P292 L295 P298 P299 F302 S303 E304 E305 D306 S307 F308 K309 K310 C311 S312 S313 E314 V315 E316 A317 K318 N319 K320 I321 E322 E323 L324 L325 A326 S327 L328 L329

N330 R331 V332 C333 Q334 D335 G336 R337 K338 P339 V342 R343 L344 I345 I346 R347 R348 Y349 SER SER GLU LYS HIS Y355 G356 R357 E358 S359 R360 Q361 C362 P365 S366 H367 V368 I369 Q370 LYS LEU GLY THR GLY ASN TYR ASP V379 N380 T381 P382 M383 V384 D385 I386 L387 M388 K389 L390 F391

R392 N393 N394 VAL ASN VAL LYS MET PRO PHE HIS LEU T404 L405 L406 S407 F410 C411 N412 L413 K414 ALA LEU ASN THR ALA LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	98.37Å 98.37Å 202.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	36.51 – 2.40 32.59 – 2.40	Depositor EDS
% Data completeness (in resolution range)	96.2 (36.51-2.40) 96.2 (32.59-2.40)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.65 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.228 , 0.276 0.229 , 0.274	Depositor DCC
R_{free} test set	2280 reflections (9.88%)	wwPDB-VP
Wilson B-factor (Å^2)	43.8	Xtriage
Anisotropy	0.148	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 71.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	3250	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.87% 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, 0.333 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: DOC

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	T	3.87	32/204 (15.7%)	3.78	52/314 (16.6%)
2	P	3.30	13/136 (9.6%)	3.58	24/208 (11.5%)
3	A	3.23	114/2850 (4.0%)	1.52	46/3851 (1.2%)
All	All	3.28	159/3190 (5.0%)	1.91	122/4373 (2.8%)

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	T	0	1
3	A	0	3
All	All	0	4

The worst 5 of 159 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	313	SER	CB-OG	99.73	2.71	1.42
3	A	318	LYS	CE-NZ	42.89	2.56	1.49
3	A	319	ASN	CG-ND2	28.92	2.05	1.32
3	A	314	GLU	CB-CG	28.28	2.05	1.52
3	A	319	ASN	CG-OD1	26.82	1.82	1.24

The worst 5 of 122 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	T	844	DT	C4-C5-C7	18.14	129.88	119.00
2	P	870	DA	O4'-C1'-N9	16.06	119.24	108.00
2	P	867	DA	O4'-C1'-N9	15.30	118.71	108.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	122	ARG	NE-CZ-NH2	-13.75	113.43	120.30
3	A	306	ASP	CB-CG-OD1	12.61	129.65	118.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	213	PHE	Peptide
3	A	305	GLU	Peptide
3	A	388	MET	Peptide
1	T	843	DG	Sidechain

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	T	183	0	104	9	2
2	P	139	0	79	2	2
3	A	2813	0	2813	212	0
4	A	107	0	0	8	0
4	P	2	0	0	0	0
4	T	6	0	0	0	0
All	All	3250	0	2996	219	2

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 36.

The worst 5 of 219 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)
3:A:392:ARG:CG	3:A:392:ARG:CB	1.79	1.58
3:A:318:LYS:CB	3:A:318:LYS:CG	1.79	1.57
3:A:230:ILE:CD1	3:A:230:ILE:CG1	1.77	1.55
3:A:105:MET:SD	3:A:105:MET:CE	2.01	1.48
3:A:382:PRO:C	3:A:382:PRO:CA	1.84	1.46

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:T:847:DT:O3'	2:P:867:DA:O5'[10_665]	1.55	0.65
1:T:847:DT:O3'	2:P:867:DA:C5'[10_665]	2.14	0.06

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
3	A	359/420 (86%)	315 (88%)	27 (8%)	17 (5%)	2 1

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	145	SER
3	A	146	ASP
3	A	333	CYS
3	A	334	GLN
3	A	366	SER

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
3	A	312/376 (83%)	273 (88%)	39 (12%)	4 5

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	347	ARG
3	A	404	THR
3	A	357	ARG
3	A	362	CYS
3	A	411	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	367	HIS
3	A	340	HIS
3	A	262	GLN
3	A	217	GLN
3	A	279	GLN

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.

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	DOC	P	873	1,2	16,19,20	1.66	3 (18%)	20,26,29	1.93	3 (15%)

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	DOC	P	873	1,2	-	2/7/18/19	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	873	DOC	C2'-C1'	4.23	1.61	1.51
2	P	873	DOC	O4'-C4'	3.03	1.50	1.44
2	P	873	DOC	O4'-C1'	-2.21	1.37	1.42

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	P	873	DOC	C2'-C1'-N1	-4.79	103.33	112.40
2	P	873	DOC	O4'-C4'-C3'	4.32	111.98	104.80
2	P	873	DOC	C4'-O4'-C1'	-3.99	106.04	109.81

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	P	873	DOC	O4'-C4'-C5'-O5'
2	P	873	DOC	C3'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	P	873	DOC	2	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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
3	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	314:GLU	C	315:VAL	N	1.64

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	T	9/11 (81%)	0.71	1 (11%) 5 4	32, 34, 64, 101	0
2	P	6/7 (85%)	0.20	0 100 100	32, 41, 44, 53	0
3	A	367/420 (87%)	1.06	77 (20%) 1 0	9, 39, 89, 101	0
All	All	382/438 (87%)	1.04	78 (20%) 1 0	9, 39, 89, 101	0

The worst 5 of 78 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	355	TYR	12.0
3	A	313	SER	10.2
3	A	349	TYR	9.1
3	A	319	ASN	8.6
3	A	84	ALA	6.8

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	DOC	P	873	18/19	0.93	0.18	27,34,39,42	0

6.3 Carbohydrates (i)

There are no monosaccharides 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.