wwPDB X-ray Structure Validation Summary Report (i)

May 29, 2020 – 03:11 am BST

PDB ID	:	3MGH
Title	:	Binary complex of a DNA polymerase lambda loop mutant
Authors	:	Garcia-Diaz, M.; Bebenek, K.; Zhou, R.Z.; Povirk, L.F.; Kunkel, T.
Deposited on	:	2010-04-06
Resolution	:	2.40 Å(reported)

The following versions of software and data (see references (1)) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

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 2.40 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.
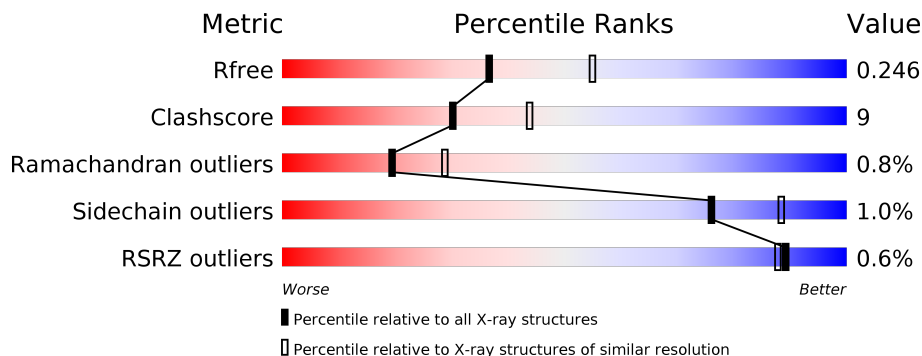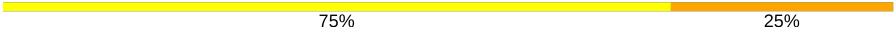

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 on 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	A	329	80%　　16%　· ·
1	C	329	%　81%　　16%　··
2	E	11	45%　45%　9%
2	T	11	45%　55%
3	F	6	17%　83%
3	P	6	67%　33%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	4	75% · 25%
4	G	4	50% · 25% · 25%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6194 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 lambda.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	319	Total 2484	C 1565	N 449	O 458	S 12	0	0	0
1	C	321	Total 2493	C 1572	N 451	O 458	S 12	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	463	LYS	-	SEE REMARK 999	UNP Q9UGP5
A	464	GLY	-	SEE REMARK 999	UNP Q9UGP5
A	470	GLU	-	SEE REMARK 999	UNP Q9UGP5
A	471	THR	-	SEE REMARK 999	UNP Q9UGP5
C	463	LYS	-	SEE REMARK 999	UNP Q9UGP5
C	464	GLY	-	SEE REMARK 999	UNP Q9UGP5
C	470	GLU	-	SEE REMARK 999	UNP Q9UGP5
C	471	THR	-	SEE REMARK 999	UNP Q9UGP5

- Molecule 2 is a DNA chain called DNA (5'-D(*CP*GP*GP*CP*AP*GP*TP*AP*CP*TP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	11	Total 224	C 107	N 43	O 64	P 10	0	0	0
2	E	11	Total 224	C 107	N 43	O 64	P 10	0	0	0

- Molecule 3 is a DNA chain called DNA (5'-D(*CP*AP*GP*TP*AP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	6	Total 119	C 58	N 23	O 33	P 5	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	6	Total 119	C 58	N 23	O 33	P 5	0	0	0

- Molecule 4 is a DNA chain called DNA (5'-D(P*GP*CP*CP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	4	Total 83	C 38	N 16	O 25	P 4	0	0	0
4	G	4	Total 83	C 38	N 16	O 25	P 4	0	0	0

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total 1	Na 1	0	0
5	C	3	Total 3	Na 3	0	0

- Molecule 6 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	171	Total 171	O 171	0	0
6	T	19	Total 19	O 19	0	0
6	P	11	Total 11	O 11	0	0
6	D	7	Total 7	O 7	0	0
6	C	126	Total 126	O 126	0	0
6	E	14	Total 14	O 14	0	0
6	F	7	Total 7	O 7	0	0
6	G	6	Total 6	O 6	0	0

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 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.
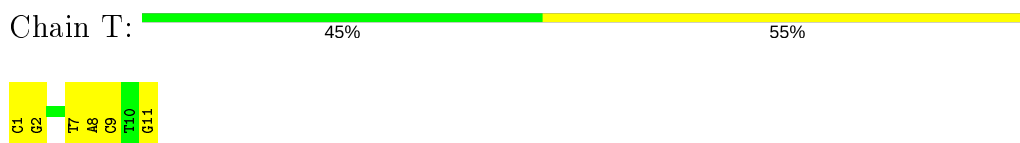
- Molecule 1: DNA polymerase lambda
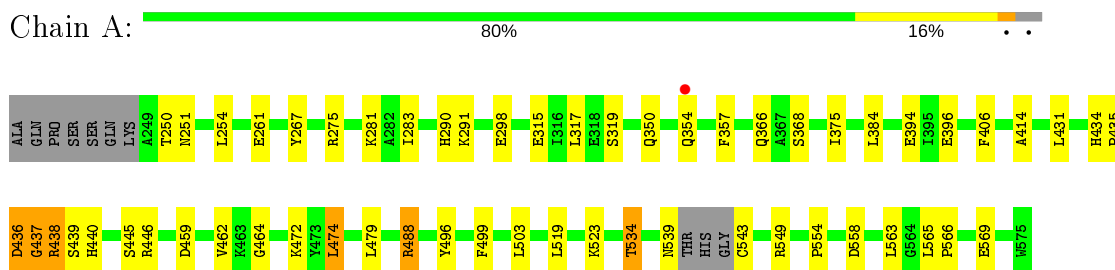
Chain A:

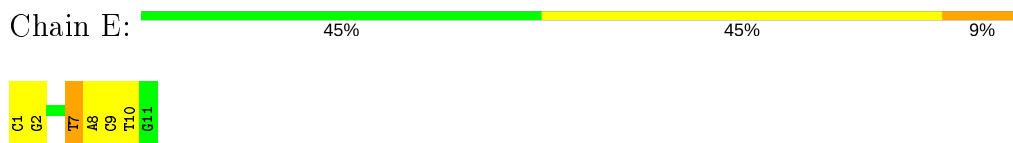- Molecule 1: DNA polymerase lambda
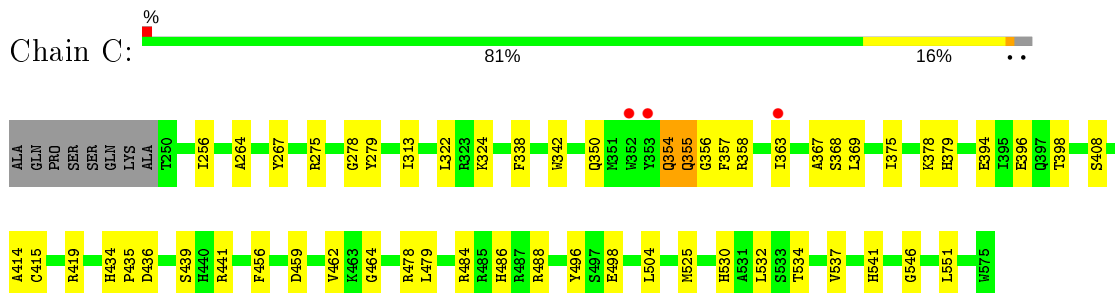
Chain C:

- Molecule 2: DNA (5'-D(*CP*GP*GP*CP*AP*GP*TP*AP*CP*TP*G)-3')

Chain T:

- Molecule 2: DNA (5'-D(*CP*GP*GP*CP*AP*GP*TP*AP*CP*TP*G)-3')

Chain E:

- Molecule 3: DNA (5'-D(*CP*AP*GP*TP*AP*C)-3')

Chain P:

C1 A2 C6

● Molecule 3: DNA (5'-D(*CP*AP*GP*TP*AP*C)-3')

Chain F: 17%   83%

C1 A2 G3 T4 A5 C6

● Molecule 4: DNA (5'-D(P*GP*CP*CP*G)-3')

Chain D: 75%   25%

G1 C2 C3 G4

● Molecule 4: DNA (5'-D(P*GP*CP*CP*G)-3')

Chain G: 50%   25%   25%

G1 C2 C3 G4
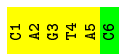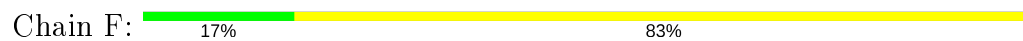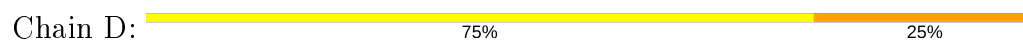
4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	96.00Å 190.89Å 58.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.46 – 2.40	Depositor
	48.46 – 2.40	EDS
% Data completeness (in resolution range)	95.6 (48.46-2.40)	Depositor
	95.6 (48.46-2.40)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.47 (at 2.39Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.209 , 0.261	Depositor
	0.195 , 0.246	DCC
R_{free} test set	2051 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	44.8	Xtriage
Anisotropy	0.247	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 39.5	EDS
L-test for twinning[2]	$<|L|> = 0.46$, $<L^2> = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6194	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.16% 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: NA

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	0.39	0/2535	0.54	1/3424 (0.0%)
1	C	0.36	0/2545	0.50	0/3438
2	E	0.76	0/251	1.35	1/386 (0.3%)
2	T	0.75	0/251	1.58	5/386 (1.3%)
3	F	0.63	0/133	1.21	0/203
3	P	0.71	0/133	1.33	1/203 (0.5%)
4	D	1.24	1/92 (1.1%)	0.99	0/138
4	G	1.32	1/92 (1.1%)	1.12	0/138
All	All	0.48	2/6032 (0.0%)	0.74	8/8316 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	G	1	DG	OP3-P	-10.56	1.48	1.61
4	D	1	DG	OP3-P	-10.30	1.48	1.61

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	T	7	DT	O4'-C1'-N1	6.91	112.84	108.00
2	T	7	DT	P-O5'-C5'	-6.57	110.38	120.90
2	E	7	DT	O4'-C1'-C2'	-6.48	100.71	105.90
2	T	11	DG	O4'-C1'-N9	6.39	112.47	108.00
2	T	7	DT	C1'-O4'-C4'	-5.76	104.34	110.10

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	2484	0	2465	47	0
1	C	2493	0	2472	43	0
2	E	224	0	125	7	0
2	T	224	0	125	4	0
3	F	119	0	69	4	0
3	P	119	0	69	1	0
4	D	83	0	45	3	0
4	G	83	0	45	3	0
5	A	1	0	0	0	0
5	C	3	0	0	0	0
6	A	171	0	0	6	0
6	C	126	0	0	1	0
6	D	7	0	0	0	0
6	E	14	0	0	0	0
6	F	7	0	0	1	0
6	G	6	0	0	0	0
6	P	11	0	0	0	0
6	T	19	0	0	0	0
All	All	6194	0	5415	104	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 9.

The worst 5 of 104 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:534:THR:HG22	1:A:549:ARG:H	1.42	0.85
1:A:565:LEU:HD12	1:A:566:PRO:HD2	1.57	0.84
1:C:396:GLU:HG3	1:C:414:ALA:HB2	1.68	0.73
1:A:291:LYS:HE3	1:A:298:GLU:OE2	1.91	0.69
1:A:488:ARG:NH1	1:A:488:ARG:HG3	2.08	0.69

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	315/329 (96%)	306 (97%)	7 (2%)	2 (1%)	25	36
1	C	319/329 (97%)	301 (94%)	15 (5%)	3 (1%)	17	25
All	All	634/658 (96%)	607 (96%)	22 (4%)	5 (1%)	19	29

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	355	GLN
1	C	541	HIS
1	C	354	GLN
1	A	438	ARG
1	A	437	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	261/275 (95%)	257 (98%)	4 (2%)	65	80
1	C	260/275 (94%)	259 (100%)	1 (0%)	91	96
All	All	521/550 (95%)	516 (99%)	5 (1%)	76	88

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	250	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	436	ASP
1	A	474	LEU
1	A	534	THR
1	C	408	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	290	HIS
1	A	539	ASN
1	C	379	HIS

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 4 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)

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	319/329 (96%)	-0.30	1 (0%)	94	93	28, 40, 61, 80	0
1	C	321/329 (97%)	-0.16	3 (0%)	84	82	26, 46, 69, 82	0
2	E	11/11 (100%)	-0.43	0	100	100	37, 47, 59, 65	0
2	T	11/11 (100%)	-0.28	0	100	100	37, 41, 51, 52	0
3	F	6/6 (100%)	-0.38	0	100	100	36, 42, 60, 67	0
3	P	6/6 (100%)	-0.46	0	100	100	31, 34, 40, 48	0
4	D	4/4 (100%)	-0.48	0	100	100	35, 36, 44, 50	0
4	G	4/4 (100%)	-0.32	0	100	100	39, 43, 44, 49	0
All	All	682/700 (97%)	-0.24	4 (0%)	89	88	26, 42, 67, 82	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	354	GLN	2.7
1	C	363	ILE	2.3
1	C	353	TYR	2.2
1	C	352	TRP	2.1

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)

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
5	NA	C	803	1/1	0.73	0.17	60,60,60,60	0
5	NA	C	1	1/1	0.80	0.22	43,43,43,43	0
5	NA	A	800	1/1	0.80	0.12	39,39,39,39	0
5	NA	C	802	1/1	0.90	0.13	42,42,42,42	0

6.5 Other polymers (i)

There are no such residues in this entry.