



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Apr 12, 2017 – 09:46 PM EDT

PDB ID : 5FKV
EMDB ID: : EMD-3198
Title : cryo-EM structure of the E. coli replicative DNA polymerase complex bound to DNA (DNA polymerase III alpha, beta, epsilon, tau complex)
Authors : Fernandez-Leiro, R.; Conrad, J.; Scheres, S.H.W.; Lamers, M.H.
Deposited on : 2015-10-20
Resolution : 8.00 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029077

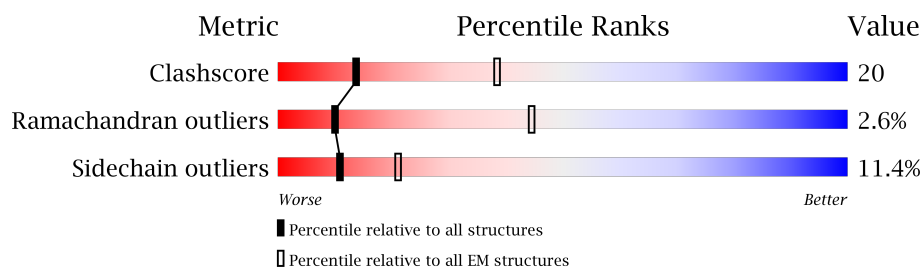
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 8.00 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)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	1160	
2	B	366	
2	C	366	
3	D	243	
4	E	144	
5	P	25	
6	T	29	

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 18386 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 III SUBUNIT ALPHA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1145	Total	C	N	O	S	0	0
			9020	5721	1562	1690	47		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	921	LEU	ALA	ENGINEERED MUTATION	UNP P10443
A	923	LEU	MET	ENGINEERED MUTATION	UNP P10443

- Molecule 2 is a protein called DNA POLYMERASE III BETA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	366	Total	C	N	O	S	0	0
			2844	1786	498	541	19		
2	C	366	Total	C	N	O	S	0	0
			2844	1786	498	541	19		

- Molecule 3 is a protein called DNA POLYMERASE III EPSILON.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	219	Total	C	N	O	S	0	0
			1717	1090	299	319	9		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	183	LEU	THR	ENGINEERED MUTATION	UNP P03007
D	185	LEU	MET	ENGINEERED MUTATION	UNP P03007
D	186	PRO	ALA	ENGINEERED MUTATION	UNP P03007
D	187	LEU	PHE	ENGINEERED MUTATION	UNP P03007

- Molecule 4 is a protein called DNA POLYMERASE III TAU.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	117	Total	C	N	O	S	0	0
			914	567	165	180	2		

- Molecule 5 is a DNA chain called PRIMER-TEMPLATE DUPLEX DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	P	25	Total	C	N	O	P	0	0
			522	246	105	146	25		

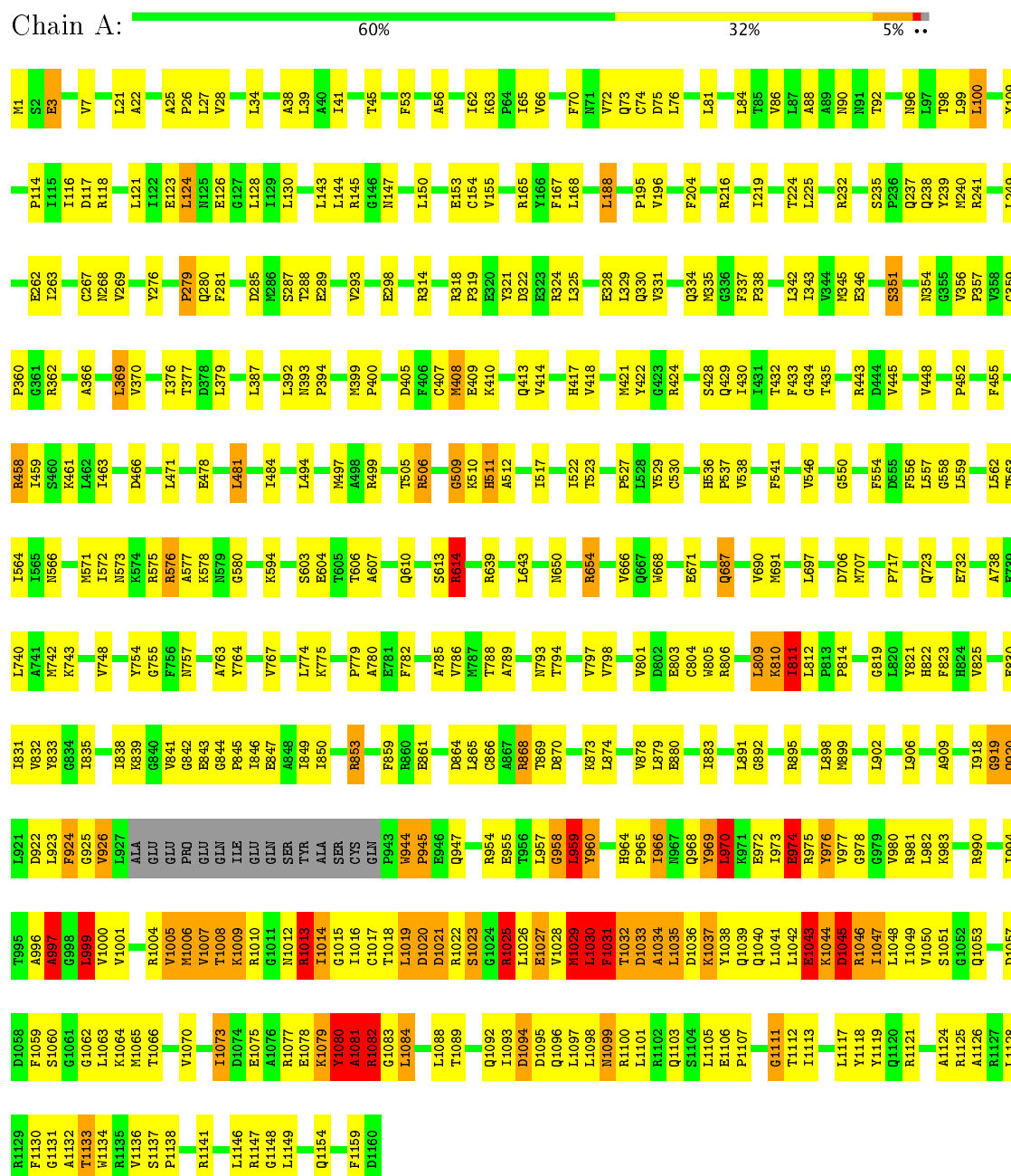
- Molecule 6 is a DNA chain called PRIMER-TEMPLATE DUPLEX DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	T	26	Total	C	N	O	P	0	0
			525	251	88	160	26		

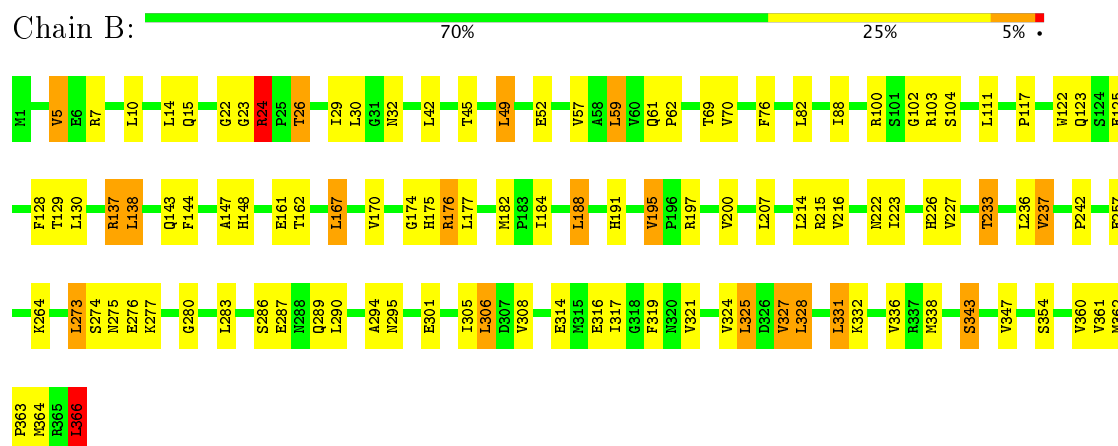
3 Residue-property plots

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. 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. 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.

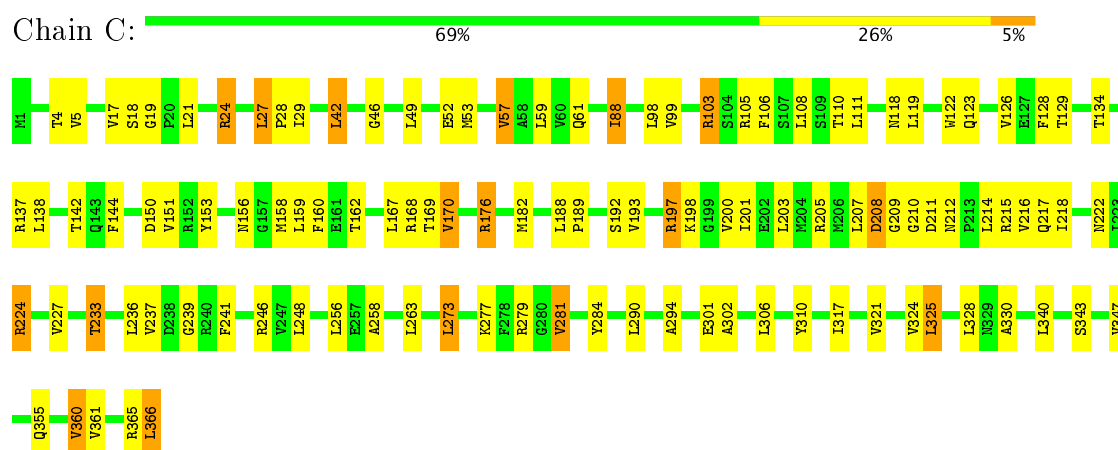
• Molecule 1: DNA POLYMERASE III SUBUNIT ALPHA



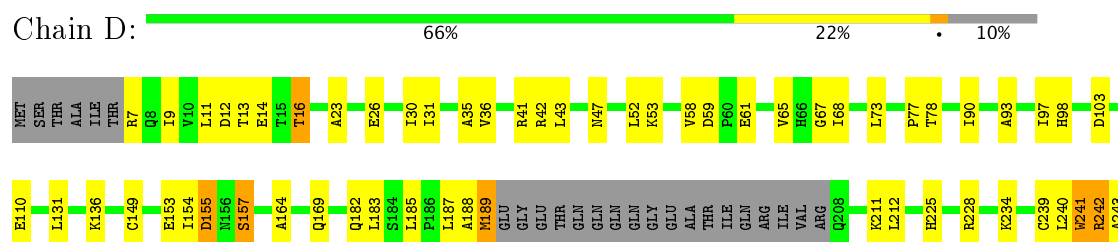
- Molecule 2: DNA POLYMERASE III BETA



- Molecule 2: DNA POLYMERASE III BETA

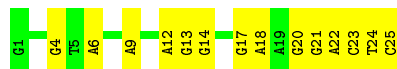


- Molecule 3: DNA POLYMERASE III EPSILON




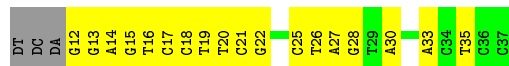
- Molecule 5: PRIMER-TEMPLATE DUPLEX DNA

Chain P:  44% 56%



- Molecule 6: PRIMER-TEMPLATE DUPLEX DNA

Chain T:  28% 62% 10%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	5663	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	each particle	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	4	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	28409	Depositor
Image detector	GATAN K2 SUMMIT	Depositor

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 >2	RMSZ	# Z >2
1	A	0.59	8/9199 (0.1%)	0.88	20/12427 (0.2%)
2	B	0.51	0/2893	0.92	3/3915 (0.1%)
2	C	0.52	0/2893	0.92	4/3915 (0.1%)
3	D	0.57	1/1747 (0.1%)	0.75	4/2358 (0.2%)
4	E	0.65	0/927	0.82	0/1259
5	P	0.39	0/588	0.81	0/907
6	T	0.44	0/585	0.94	0/899
All	All	0.56	9/18832 (0.0%)	0.88	31/25680 (0.1%)

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	A	0	8
2	B	0	1
All	All	0	9

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	242	ARG	N-CA	-12.04	1.22	1.46
1	A	687	GLN	CG-CD	10.32	1.74	1.51
1	A	1045	ASP	C-N	-8.38	1.14	1.34
1	A	687	GLN	CD-OE1	6.76	1.38	1.24
1	A	1080	TYR	CB-CG	-6.71	1.41	1.51

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1045	ASP	O-C-N	-12.69	102.40	122.70
1	A	1081	ALA	N-CA-C	-10.74	82.01	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	687	GLN	CB-CG-CD	8.18	132.87	111.60
1	A	1082	ARG	CA-CB-CG	7.99	130.99	113.40
1	A	614	ARG	NE-CZ-NH1	7.55	124.08	120.30

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1027	GLU	Peptide
1	A	1033	ASP	Peptide
1	A	509	GLY	Peptide
1	A	511	HIS	Peptide
1	A	996	ALA	Peptide

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	9020	0	8985	497	0
2	B	2844	0	2861	67	0
2	C	2844	0	2861	46	0
3	D	1717	0	1715	58	0
4	E	914	0	915	30	0
5	P	522	0	280	39	0
6	T	525	0	295	31	0
All	All	18386	0	17912	725	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 20.

The worst 5 of 725 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:687:GLN:CD	1:A:687:GLN:CG	1.74	1.55
1:A:1040:GLN:O	1:A:1043:GLU:HG3	1.29	1.32

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1079:LYS:NZ	1:A:1083:GLY:O	1.66	1.29
1:A:1040:GLN:O	1:A:1043:GLU:CG	1.81	1.26
1:A:977:VAL:HB	1:A:981:ARG:HB2	1.20	1.18

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 PDB entries followed by that with respect to all EM entries.

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	1141/1160 (98%)	1028 (90%)	70 (6%)	43 (4%)	4	32
2	B	364/366 (100%)	338 (93%)	21 (6%)	5 (1%)	13	54
2	C	364/366 (100%)	339 (93%)	19 (5%)	6 (2%)	11	51
3	D	215/243 (88%)	199 (93%)	14 (6%)	2 (1%)	20	63
4	E	115/144 (80%)	108 (94%)	5 (4%)	2 (2%)	11	50
All	All	2199/2279 (96%)	2012 (92%)	129 (6%)	58 (3%)	10	40

5 of 58 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	GLU
1	A	279	PRO
1	A	280	GLN
1	A	466	ASP
1	A	576	ARG

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 PDB entries followed by that with respect to all EM entries.

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	952/965 (99%)	860 (90%)	92 (10%)	9	35
2	B	313/313 (100%)	268 (86%)	45 (14%)	4	22
2	C	313/313 (100%)	261 (83%)	52 (17%)	2	16
3	D	180/200 (90%)	170 (94%)	10 (6%)	25	57
4	E	97/122 (80%)	84 (87%)	13 (13%)	4	24
All	All	1855/1913 (97%)	1643 (89%)	212 (11%)	11	28

5 of 212 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	42	LEU
2	B	216	VAL
3	D	157	SER
2	B	70	VAL
2	B	138	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	920	GLN
2	B	143	GLN
3	D	98	HIS
1	A	1039	GLN
2	B	149	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 [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
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1045:ASP	C	1046:ARG	N	1.14