



## wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Dec 13, 2016 – 07:56 PM EST

PDB ID : 5FKW  
EMDB ID: : EMD-3202  
Title : cryo-EM structure of the E. coli replicative DNA polymerase complex bound to DNA (DNA polymerase III alpha, beta, epsilon)  
Authors : Fernandez-Leiro, R.; Conrad, J.; Scheres, S.H.W.; Lamers, M.H.  
Deposited on : 2015-10-20  
Resolution : 7.30 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](mailto: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  
Mogul : unknown  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20028442



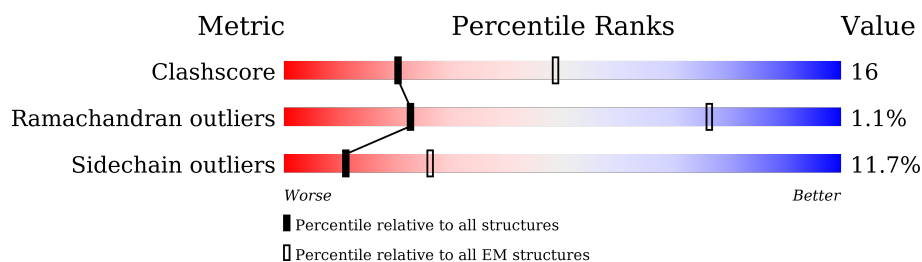
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 7.30 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)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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  $\geq 3$ , 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	P	25	
5	T	29	



## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 15726 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 ALPHA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	927	Total	C	N	O	S	0	0
			7274	4630	1240	1363	41		

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 DNA chain called PRIMER-TEMPLATE DUPLEX DNA.



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

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

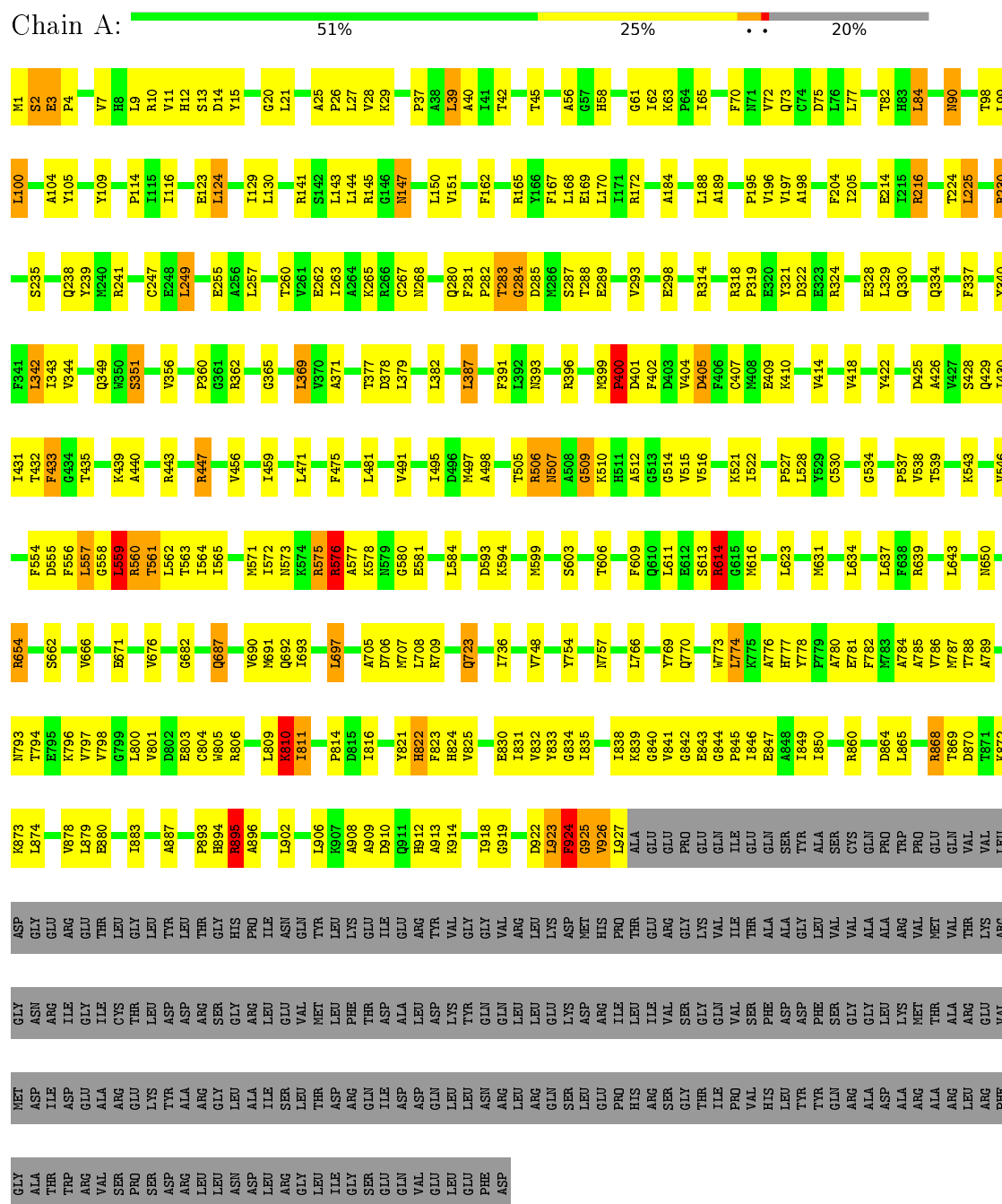
Mol	Chain	Residues	Atoms					AltConf	Trace
5	T	26	Total	C	N	O	P	0	0
			525	251	88	160	26		



### 3 Residue-property plots

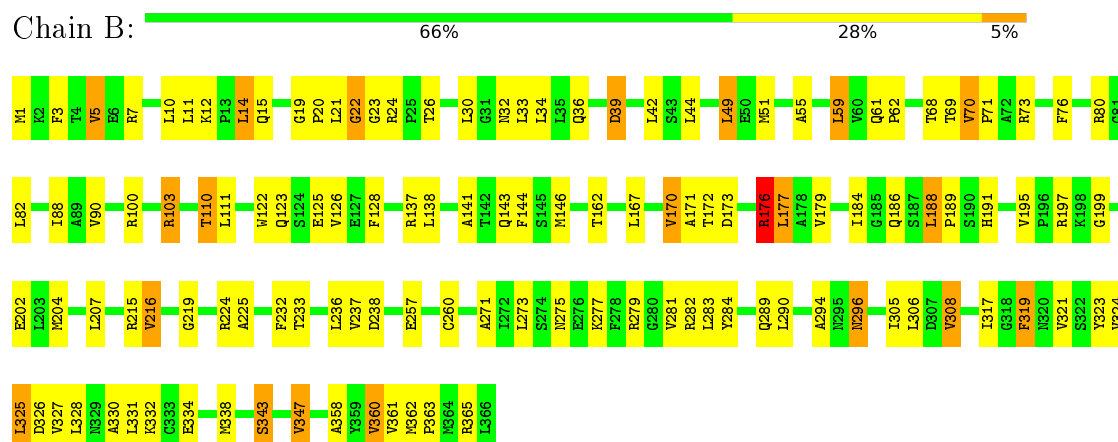
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric 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 ALPHA

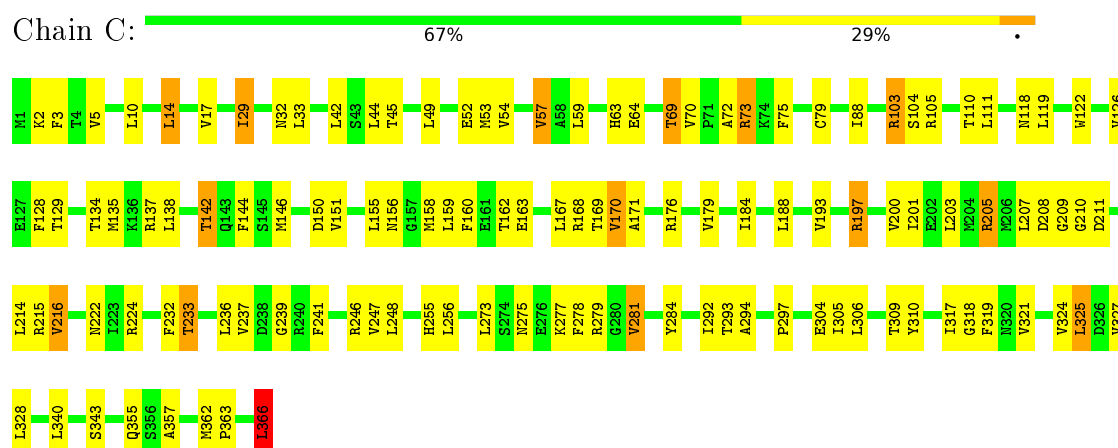




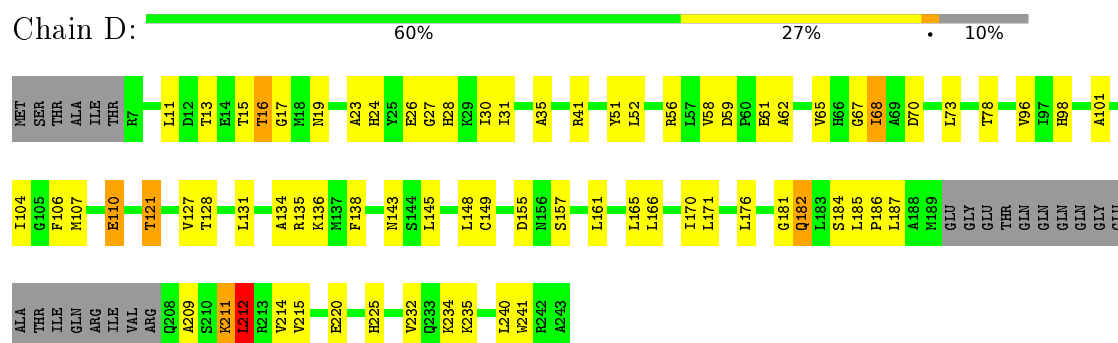
- Molecule 2: DNA POLYMERASE III BETA



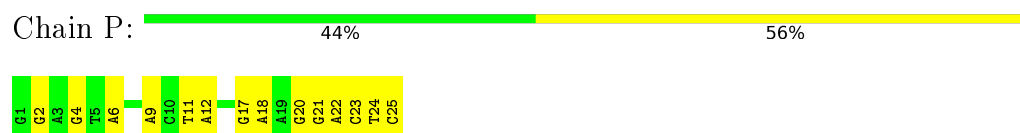
- Molecule 2: DNA POLYMERASE III BETA



- Molecule 3: DNA POLYMERASE III EPSILON



- Molecule 4: PRIMER-TEMPLATE DUPLEX DNA





- Molecule 5: PRIMER-TEMPLATE DUPLEX DNA

Chain T:  31% 59% 10%





## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	40582	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 ( $e^-/\text{\AA}^2$ )	40	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	28409	Depositor
Image detector	GATAN K2 QUANTUM (4K X 4K)	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	0.57	2/7427 (0.0%)	0.85	11/10040 (0.1%)
2	B	0.54	0/2893	0.94	6/3915 (0.2%)
2	C	0.55	0/2893	0.94	3/3915 (0.1%)
3	D	0.56	1/1747 (0.1%)	0.80	3/2358 (0.1%)
4	P	0.42	0/588	0.89	0/907
5	T	0.46	0/585	0.98	0/899
All	All	0.55	3/16133 (0.0%)	0.89	23/22034 (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	11
2	B	0	1
3	D	0	4
All	All	0	16

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	687	GLN	CD-NE2	9.65	1.56	1.32
3	D	110	GLU	CD-OE2	6.77	1.33	1.25
1	A	687	GLN	CD-OE1	5.48	1.36	1.24

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	614	ARG	NE-CZ-NH1	7.76	124.18	120.30
1	A	924	PHE	N-CA-C	7.16	130.33	111.00
3	D	212	LEU	N-CA-C	6.92	129.69	111.00
3	D	212	LEU	N-CA-CB	-6.74	96.93	110.40

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	924	PHE	N-CA-CB	-6.17	99.50	110.60

There are no chirality outliers.

5 of 16 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	2	SER	Peptide
1	A	400	PRO	Peptide
1	A	433	PHE	Peptide
1	A	509	GLY	Peptide
1	A	559	LEU	Mainchain

## 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	7274	0	7219	282	0
2	B	2844	0	2861	66	0
2	C	2844	0	2861	61	0
3	D	1717	0	1715	54	0
4	P	522	0	280	40	0
5	T	525	0	295	31	0
All	All	15726	0	15231	493	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 16.

The worst 5 of 493 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:810:LYS:HA	1:A:811:ILE:CG1	1.45	1.46
1:A:810:LYS:CA	1:A:811:ILE:HG12	1.69	1.22
1:A:810:LYS:CA	1:A:811:ILE:CG1	2.24	1.15
1:A:810:LYS:CB	1:A:811:ILE:HG12	1.78	1.12
1:A:399:MET:CE	1:A:401:ASP:HB2	1.83	1.08



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	925/1160 (80%)	856 (92%)	57 (6%)	12 (1%)	15	60
2	B	364/366 (100%)	341 (94%)	20 (6%)	3 (1%)	24	69
2	C	364/366 (100%)	344 (94%)	18 (5%)	2 (0%)	34	77
3	D	215/243 (88%)	198 (92%)	13 (6%)	4 (2%)	10	52
All	All	1868/2135 (88%)	1739 (93%)	108 (6%)	21 (1%)	23	63

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	284	GLY
1	A	559	LEU
1	A	561	THR
1	A	895	ARG
1	A	924	PHE

### 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	765/965 (79%)	694 (91%)	71 (9%)	11	42
2	B	313/313 (100%)	263 (84%)	50 (16%)	3	21
2	C	313/313 (100%)	265 (85%)	48 (15%)	3	22

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	D	180/200 (90%)	165 (92%)	15 (8%)	14	48
All	All	1571/1791 (88%)	1387 (88%)	184 (12%)	11	32

5 of 184 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	100	ARG
2	B	216	VAL
3	D	16	THR
2	B	110	THR
2	B	170	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	824	HIS
2	B	16	GLN
2	C	348	GLN
1	A	894	HIS
2	B	36	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](#)

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](#)

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.