



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:42 PM GMT

PDB ID : 1Q57
Title : The Crystal Structure of the Bifunctional Primase-Helicase of Bacteriophage T7
Authors : Toth, E.A.; Li, Y.; Sawaya, M.R.; Cheng, Y.; Ellenberger, T.
Deposited on : 2003-08-06
Resolution : 3.45 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
<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

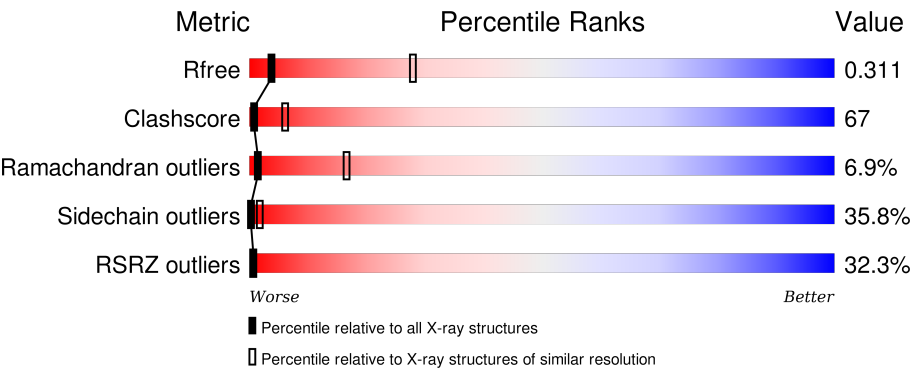
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.45 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	1000 (3.56-3.36)
Clashscore	102246	1090 (3.56-3.36)
Ramachandran outliers	100387	1057 (3.56-3.36)
Sidechain outliers	100360	1058 (3.56-3.36)
RSRZ outliers	91569	1005 (3.56-3.36)

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 $\leq 5\%$. 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	503	<div><div>20%</div><div><div>23%</div><div>47%</div><div>23%</div><div>.</div><div>.</div></div></div>
1	B	503	<div><div>33%</div><div><div>23%</div><div>49%</div><div>20%</div><div>.</div><div>.</div></div></div>
1	C	503	<div><div>32%</div><div><div>23%</div><div>47%</div><div>22%</div><div>.</div><div>.</div></div></div>
1	D	503	<div><div>33%</div><div><div>22%</div><div>47%</div><div>24%</div><div>.</div><div>.</div></div></div>
1	E	503	<div><div>39%</div><div><div>23%</div><div>46%</div><div>24%</div><div>.</div><div>.</div></div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	503	<div><div></div><div>25%</div><div></div><div>23%</div><div></div><div>46%</div><div></div><div>22%</div><div></div><div>5%</div><div></div><div></div></div>
1	G	503	<div><div></div><div>35%</div><div></div><div>23%</div><div></div><div>48%</div><div></div><div>22%</div><div></div><div></div><div></div><div></div></div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 26208 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 primase/helicase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	483	Total	C	N	O	S	0	0	0
			3744	2341	657	721	25			
1	B	483	Total	C	N	O	S	0	0	0
			3744	2341	657	721	25			
1	C	483	Total	C	N	O	S	0	0	0
			3744	2341	657	721	25			
1	D	483	Total	C	N	O	S	0	0	0
			3744	2341	657	721	25			
1	E	483	Total	C	N	O	S	0	0	0
			3744	2341	657	721	25			
1	F	483	Total	C	N	O	S	0	0	0
			3744	2341	657	721	25			
1	G	483	Total	C	N	O	S	0	0	0
			3744	2341	657	721	25			

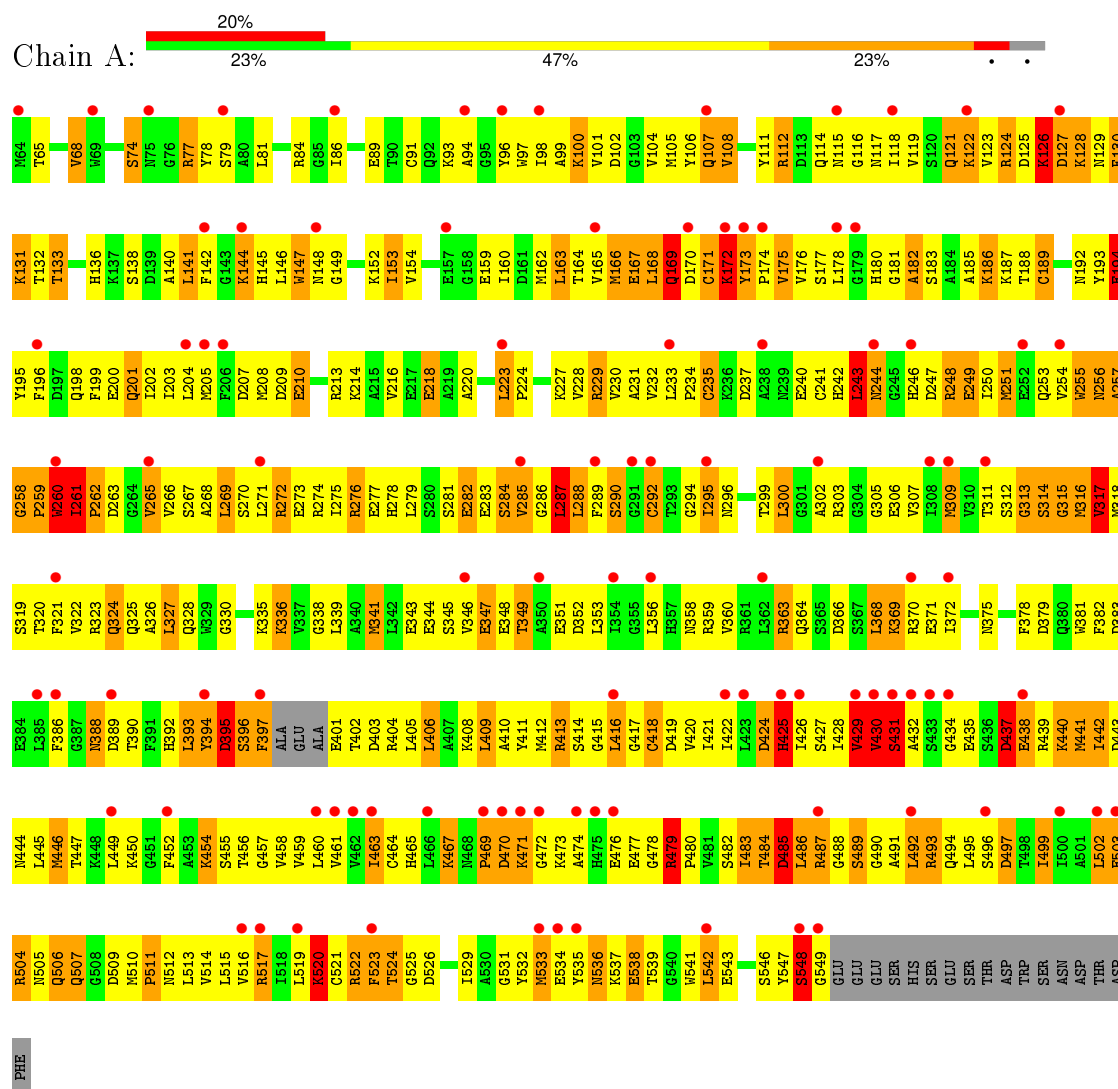
There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	317	VAL	GLY	ENGINEERED	UNP P03692
A	318	MET	LYS	ENGINEERED	UNP P03692
B	317	VAL	GLY	ENGINEERED	UNP P03692
B	318	MET	LYS	ENGINEERED	UNP P03692
C	317	VAL	GLY	ENGINEERED	UNP P03692
C	318	MET	LYS	ENGINEERED	UNP P03692
D	317	VAL	GLY	ENGINEERED	UNP P03692
D	318	MET	LYS	ENGINEERED	UNP P03692
E	317	VAL	GLY	ENGINEERED	UNP P03692
E	318	MET	LYS	ENGINEERED	UNP P03692
F	317	VAL	GLY	ENGINEERED	UNP P03692
F	318	MET	LYS	ENGINEERED	UNP P03692
G	317	VAL	GLY	ENGINEERED	UNP P03692
G	318	MET	LYS	ENGINEERED	UNP P03692

3 Residue-property plots

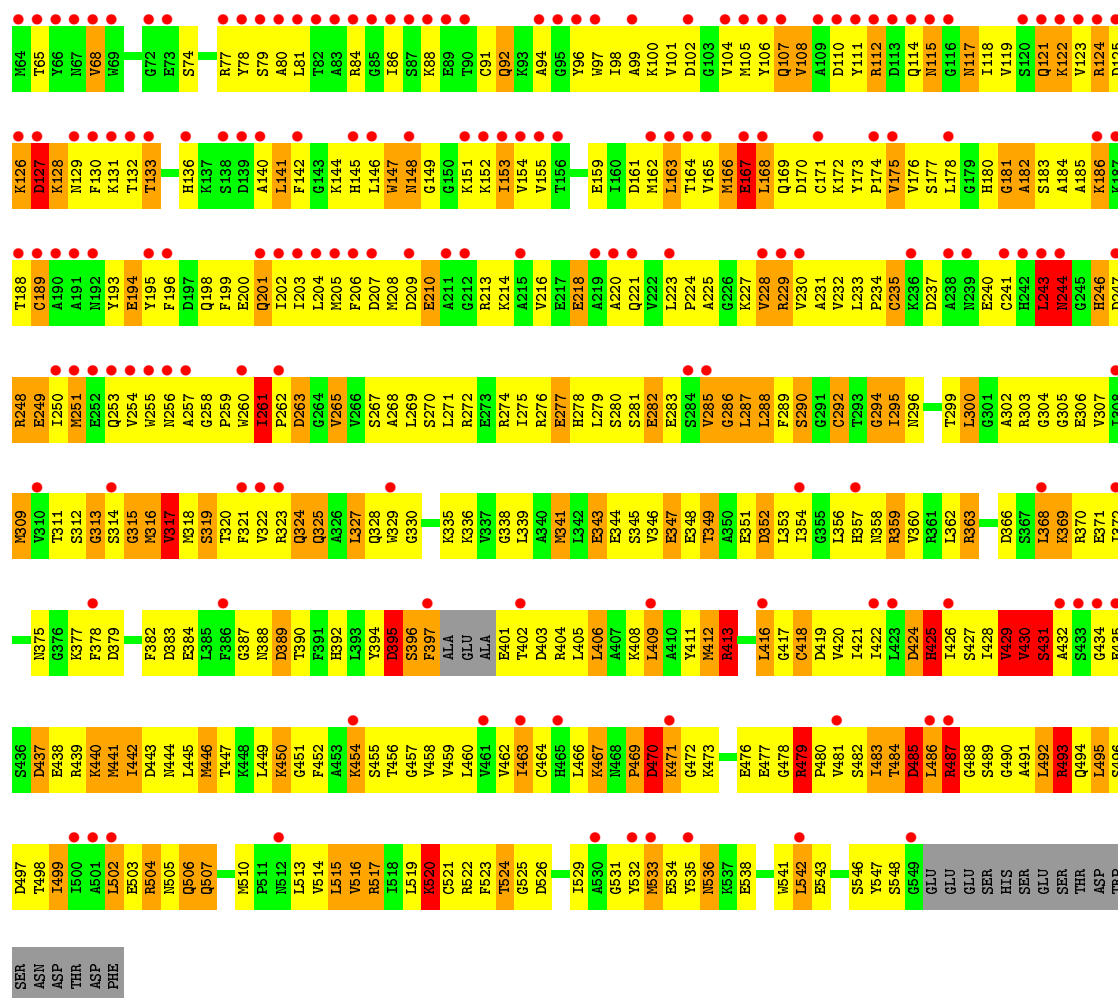
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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 ($\text{RSRZ} > 2$). 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 primase/helicase

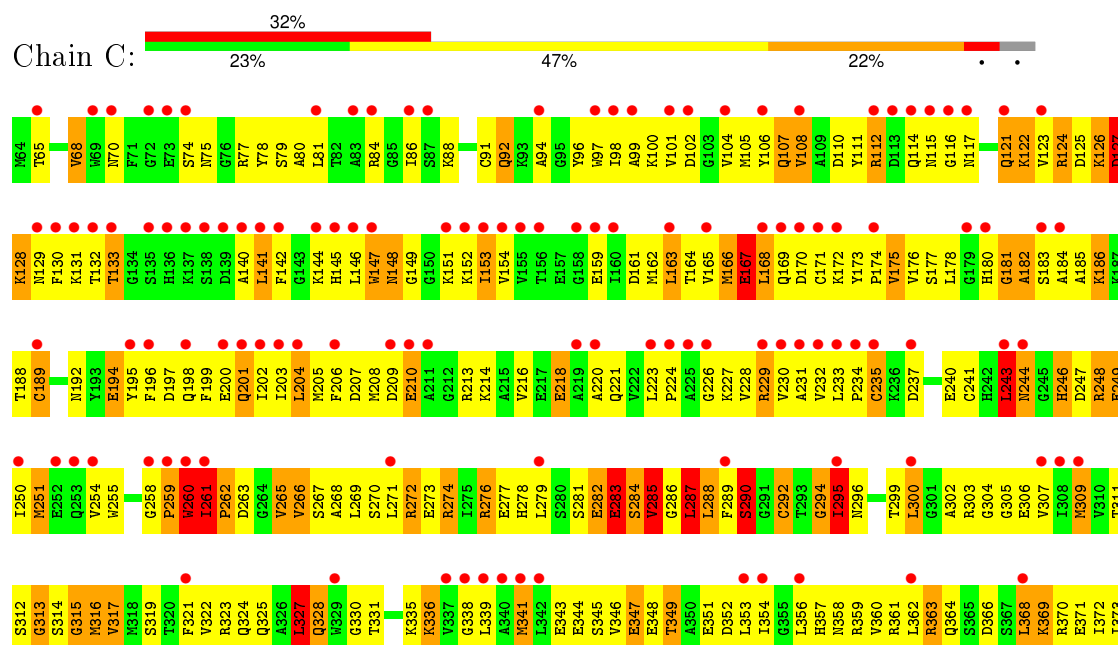


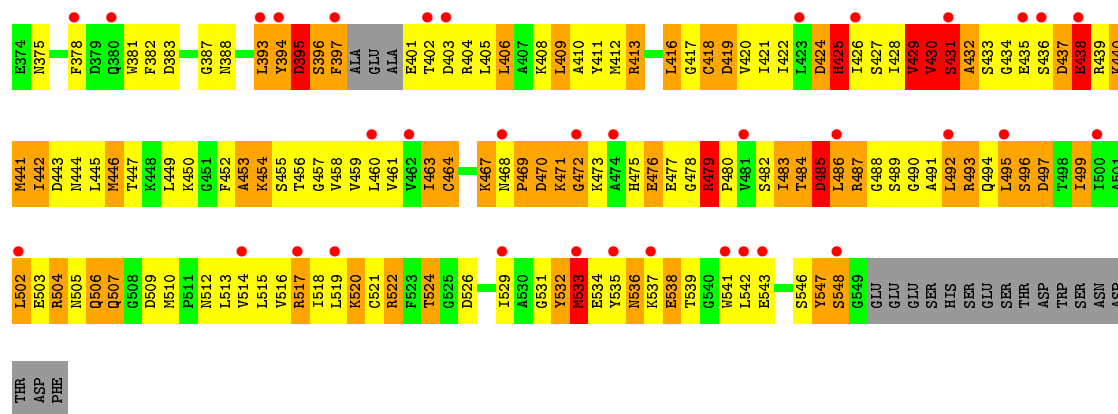
- Molecule 1: DNA primase/helicase



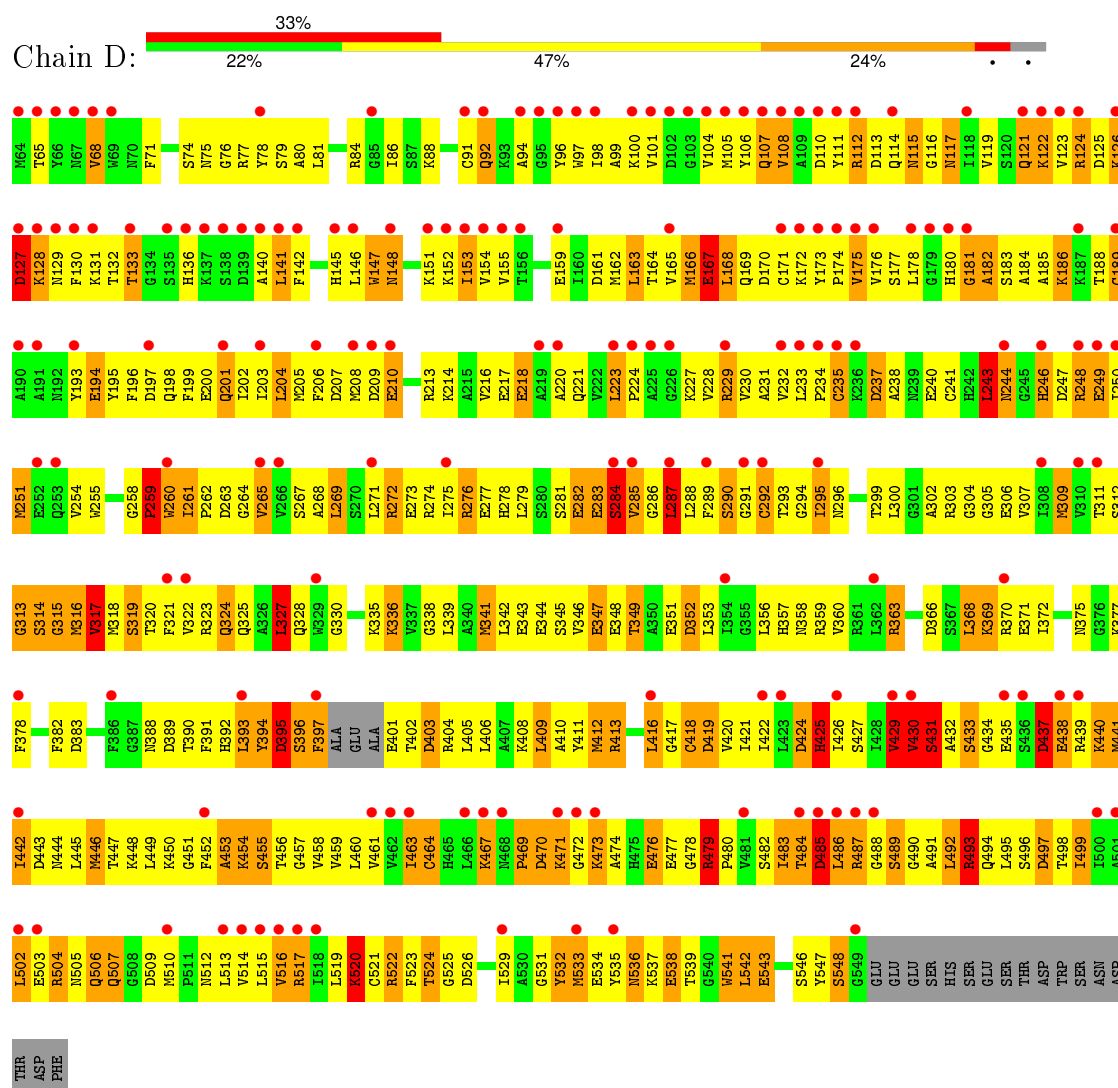


• Molecule 1: DNA primase/helicase

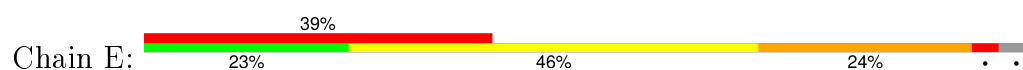


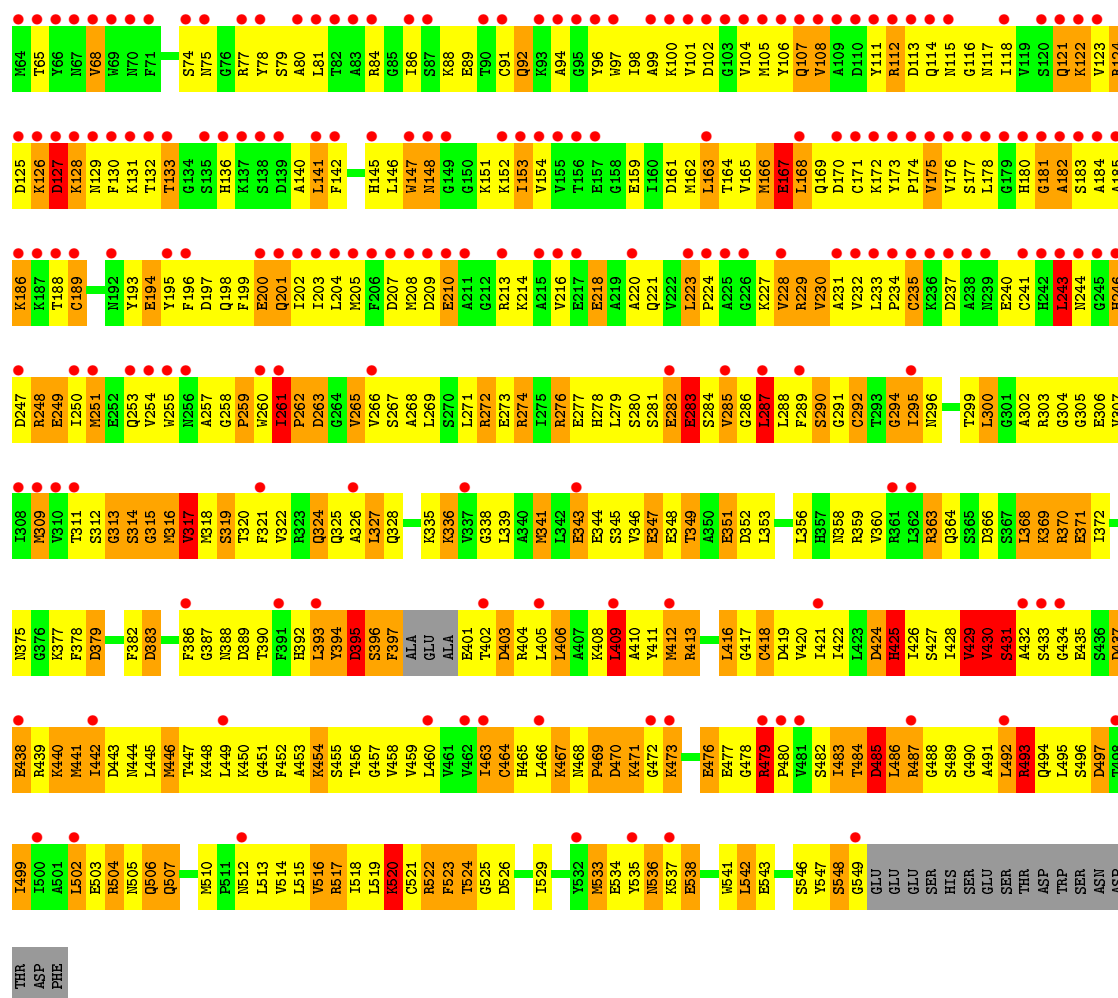


• Molecule 1: DNA primase/helicase



• Molecule 1: DNA primase/helicase







4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	117.18Å 171.57Å 118.58Å 90.00° 99.86° 90.00°	Depositor
Resolution (Å)	20.00 – 3.45 95.78 – 3.45	Depositor EDS
% Data completeness (in resolution range)	99.7 (20.00-3.45) 99.6 (95.78-3.45)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.04	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.91 (at 3.49Å)	Xtriage
Refinement program	REFMAC 5.1.80	Depositor
R, R_{free}	0.299 , 0.326 0.284 , 0.311	Depositor DCC
R_{free} test set	3055 reflections (5.34%)	DCC
Wilson B-factor (Å ²)	127.0	Xtriage
Anisotropy	0.322	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 197.2	EDS
Estimated twinning fraction	0.001 for l,-k,h	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 60583 reflections	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	26208	wwPDB-VP
Average B, all atoms (Å ²)	138.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.65% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.96	5/3806 (0.1%)	1.12	33/5124 (0.6%)
1	B	0.87	2/3806 (0.1%)	1.05	35/5124 (0.7%)
1	C	0.85	5/3806 (0.1%)	1.04	33/5124 (0.6%)
1	D	0.78	2/3806 (0.1%)	1.04	33/5124 (0.6%)
1	E	0.75	1/3806 (0.0%)	1.00	33/5124 (0.6%)
1	F	0.85	1/3806 (0.0%)	1.04	32/5124 (0.6%)
1	G	0.73	1/3806 (0.0%)	1.02	36/5124 (0.7%)
All	All	0.83	17/26642 (0.1%)	1.05	235/35868 (0.7%)

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	4
1	B	0	3
1	C	0	5
1	D	0	4
1	E	0	4
1	F	0	5
1	G	0	4
All	All	0	29

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	130	PHE	CE1-CZ	6.54	1.49	1.37
1	A	317	VAL	CB-CG2	6.50	1.66	1.52
1	F	260	TRP	CB-CG	6.18	1.61	1.50
1	B	277	GLU	CD-OE1	6.07	1.32	1.25
1	A	130	PHE	CE2-CZ	5.88	1.48	1.37

The worst 5 of 235 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	259	PRO	CA-N-CD	-13.27	92.93	111.50
1	A	262	PRO	N-CD-CG	-11.76	85.56	103.20
1	A	261	ILE	C-N-CD	-10.84	96.75	120.60
1	A	173	TYR	N-CA-CB	-10.40	91.88	110.60
1	C	167	GLU	CA-C-N	-10.27	94.60	117.20

There are no chirality outliers.

5 of 29 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	260	TRP	Mainchain
1	A	261	ILE	Peptide
1	A	282	GLU	Peptide
1	A	396	SER	Peptide
1	B	167	GLU	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	3744	0	3697	553	0
1	B	3744	0	3697	482	0
1	C	3744	0	3697	545	0
1	D	3744	0	3697	527	0
1	E	3744	0	3697	524	0
1	F	3744	0	3697	527	0
1	G	3744	0	3697	505	0
All	All	26208	0	25879	3503	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 67.

The worst 5 of 3503 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:G:276:ARG:CB	1:G:276:ARG:HH11	1.16	1.55
1:B:248:ARG:HH11	1:B:248:ARG:CB	1.19	1.51
1:F:260:TRP:HB2	1:F:262:PRO:CD	1.49	1.40
1:G:276:ARG:HB3	1:G:276:ARG:NH1	1.31	1.40
1:F:260:TRP:CB	1:F:262:PRO:HD2	1.56	1.35

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	479/503 (95%)	377 (79%)	68 (14%)	34 (7%)	1	16
1	B	479/503 (95%)	384 (80%)	65 (14%)	30 (6%)	2	19
1	C	479/503 (95%)	385 (80%)	57 (12%)	37 (8%)	1	14
1	D	479/503 (95%)	377 (79%)	66 (14%)	36 (8%)	1	14
1	E	479/503 (95%)	381 (80%)	65 (14%)	33 (7%)	1	16
1	F	479/503 (95%)	380 (79%)	63 (13%)	36 (8%)	1	14
1	G	479/503 (95%)	386 (81%)	67 (14%)	26 (5%)	2	23
All	All	3353/3521 (95%)	2670 (80%)	451 (14%)	232 (7%)	1	16

5 of 232 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	128	LYS
1	A	182	ALA
1	A	244	ASN
1	A	258	GLY
1	A	261	ILE

5.3.2 Protein sidechains ⓘ

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	401/421 (95%)	254 (63%)	147 (37%)	0	2
1	B	401/421 (95%)	264 (66%)	137 (34%)	0	2
1	C	401/421 (95%)	257 (64%)	144 (36%)	0	2
1	D	401/421 (95%)	254 (63%)	147 (37%)	0	2
1	E	401/421 (95%)	257 (64%)	144 (36%)	0	2
1	F	401/421 (95%)	254 (63%)	147 (37%)	0	2
1	G	401/421 (95%)	261 (65%)	140 (35%)	0	2
All	All	2807/2947 (95%)	1801 (64%)	1006 (36%)	0	2

5 of 1006 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	167	GLU
1	D	517	ARG
1	G	287	LEU
1	D	218	GLU
1	D	370	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 146 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	117	ASN
1	D	512	ASN
1	G	201	GLN
1	D	145	HIS
1	D	325	GLN

5.3.3 RNA ⓘ

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 [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	483/503 (96%)	1.17	102 (21%) 1 1	59, 136, 136, 136	0
1	B	483/503 (96%)	1.92	166 (34%) 0 0	94, 94, 178, 178	0
1	C	483/503 (96%)	1.65	159 (32%) 0 0	120, 120, 178, 178	0
1	D	483/503 (96%)	1.81	168 (34%) 0 0	132, 132, 178, 178	0
1	E	483/503 (96%)	2.79	194 (40%) 0 0	115, 115, 178, 178	0
1	F	483/503 (96%)	1.47	127 (26%) 1 1	103, 164, 164, 164	0
1	G	483/503 (96%)	2.22	176 (36%) 0 0	129, 129, 178, 178	0
All	All	3381/3521 (96%)	1.86	1092 (32%) 1 0	59, 132, 178, 178	0

The worst 5 of 1092 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	136	HIS	32.6
1	G	109	ALA	30.3
1	E	65	THR	26.0
1	G	157	GLU	23.1
1	E	181	GLY	21.7

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.