



wwPDB EM Validation Summary Report ⓘ

May 14, 2024 – 10:11 am BST

PDB ID : 8Q3R
EMDB ID : EMD-18134
Title : Cryo-EM structure of the DNA polymerase holoenzyme E9-A20-D4 of vaccinia virus
Authors : Burmeister, W.P.; Ballandras-Colas, A.; Boettcher, B.; Grimm, C.
Deposited on : 2023-08-04
Resolution : 3.80 Å (reported)
Based on initial models : 4od8, 8hg1, .

This is a wwPDB EM 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/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

EMDB validation analysis : 0.0.1.dev92
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36.2

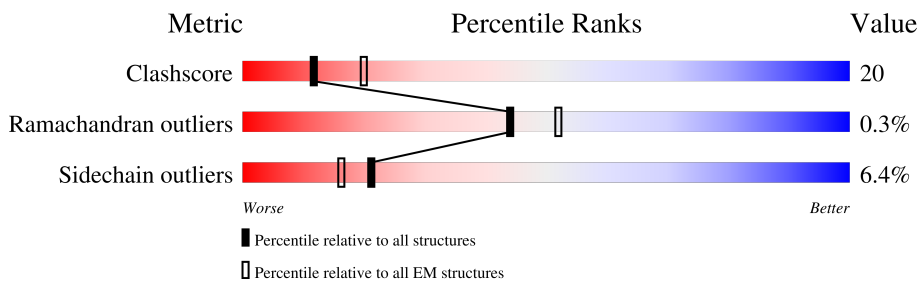
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.80 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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

Mol	Chain	Length	Quality of chain
1	D	242	55% (green), 34% (yellow), 10% (grey)
2	A	426	56% (green), 41% (yellow), 2% (orange), 1% (red)
3	E	1033	42% (green), 35% (yellow), 1% (orange), 20% (grey)

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 11921 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 Uracil-DNA glycosylase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	D	218	1771	1147	293	325	6	0	0

There are 25 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-23	MET	-	initiating methionine	UNP P20536
D	-22	ALA	-	expression tag	UNP P20536
D	-21	SER	-	expression tag	UNP P20536
D	-20	TRP	-	expression tag	UNP P20536
D	-19	SER	-	expression tag	UNP P20536
D	-18	HIS	-	expression tag	UNP P20536
D	-17	PRO	-	expression tag	UNP P20536
D	-16	GLN	-	expression tag	UNP P20536
D	-15	PHE	-	expression tag	UNP P20536
D	-14	GLU	-	expression tag	UNP P20536
D	-13	LYS	-	expression tag	UNP P20536
D	-12	SER	-	expression tag	UNP P20536
D	-11	GLY	-	expression tag	UNP P20536
D	-10	GLY	-	expression tag	UNP P20536
D	-9	GLY	-	expression tag	UNP P20536
D	-8	GLY	-	expression tag	UNP P20536
D	-7	GLY	-	expression tag	UNP P20536
D	-6	LEU	-	expression tag	UNP P20536
D	-5	VAL	-	expression tag	UNP P20536
D	-4	PRO	-	expression tag	UNP P20536
D	-3	ARG	-	expression tag	UNP P20536
D	-2	GLY	-	expression tag	UNP P20536
D	-1	SER	-	expression tag	UNP P20536
D	0	ALA	-	expression tag	UNP P20536
D	208	ALA	VAL	conflict	UNP P20536

- Molecule 2 is a protein called DNA polymerase processivity factor component OPG148.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	A	420	3430	2209	560	651	10	0	0

- Molecule 3 is a protein called DNA polymerase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	E	827	6720	4298	1119	1261	42	0	0

There are 28 discrepancies between the modelled and reference sequences:

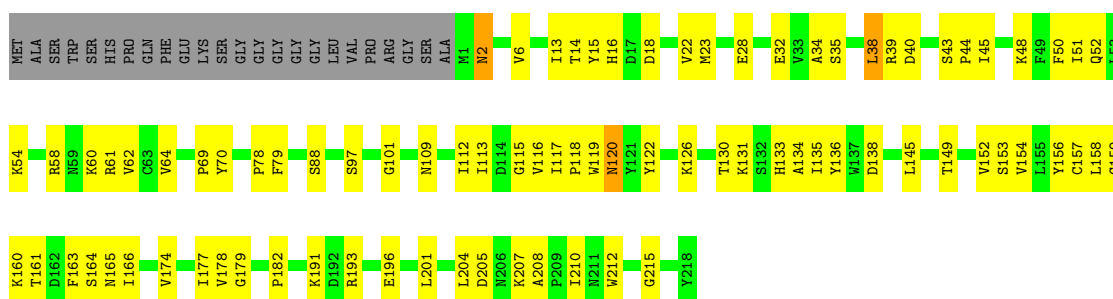
Chain	Residue	Modelled	Actual	Comment	Reference
E	-26	MET	-	initiating methionine	UNP P20509
E	-25	SER	-	expression tag	UNP P20509
E	-24	TYR	-	expression tag	UNP P20509
E	-23	TYR	-	expression tag	UNP P20509
E	-22	HIS	-	expression tag	UNP P20509
E	-21	HIS	-	expression tag	UNP P20509
E	-20	HIS	-	expression tag	UNP P20509
E	-19	HIS	-	expression tag	UNP P20509
E	-18	HIS	-	expression tag	UNP P20509
E	-17	HIS	-	expression tag	UNP P20509
E	-16	ASP	-	expression tag	UNP P20509
E	-15	TYR	-	expression tag	UNP P20509
E	-14	ASP	-	expression tag	UNP P20509
E	-13	ILE	-	expression tag	UNP P20509
E	-12	PRO	-	expression tag	UNP P20509
E	-11	THR	-	expression tag	UNP P20509
E	-10	THR	-	expression tag	UNP P20509
E	-9	GLU	-	expression tag	UNP P20509
E	-8	ASN	-	expression tag	UNP P20509
E	-7	LEU	-	expression tag	UNP P20509
E	-6	TYR	-	expression tag	UNP P20509
E	-5	PHE	-	expression tag	UNP P20509
E	-4	GLN	-	expression tag	UNP P20509
E	-3	GLY	-	expression tag	UNP P20509
E	-2	ALA	-	expression tag	UNP P20509
E	-1	MET	-	expression tag	UNP P20509
E	0	ASP	-	expression tag	UNP P20509
E	1	PRO	-	expression tag	UNP P20509

3 Residue-property plots

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. 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.

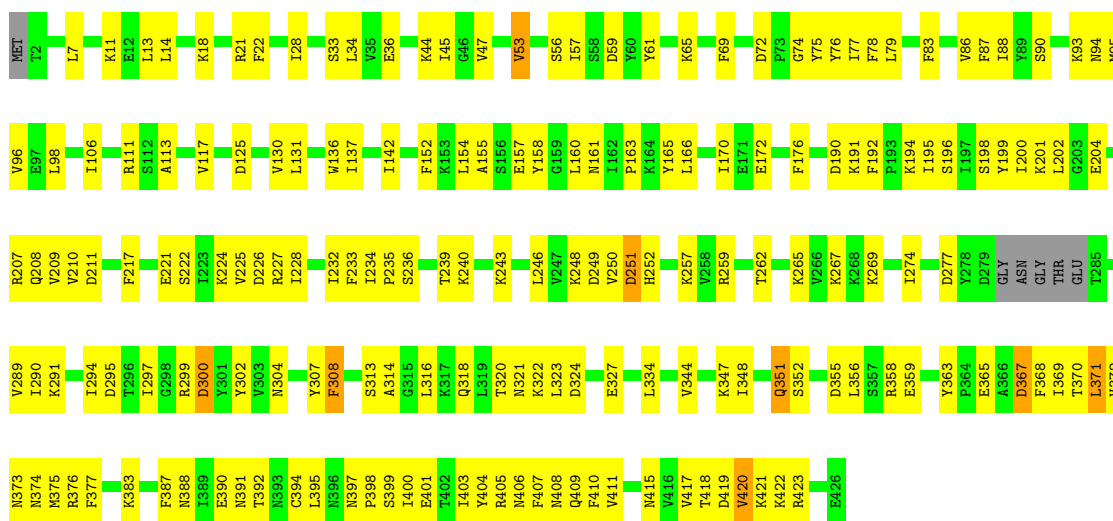
- Molecule 1: Uracil-DNA glycosylase

Chain D: 



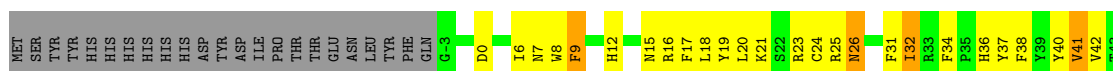
- Molecule 2: DNA polymerase processivity factor component OPG148

Chain A: 



- Molecule 3: DNA polymerase

Chain E: 



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	104239	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	68	Depositor
Minimum defocus (nm)	1400	Depositor
Maximum defocus (nm)	2400	Depositor
Magnification	42000	Depositor
Image detector	FEI FALCON III (4k x 4k)	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	D	0.29	0/1820	0.47	0/2474
2	A	0.27	0/3496	0.47	0/4713
3	E	0.37	0/6864	0.52	0/9278
All	All	0.33	0/12180	0.50	0/16465

There are no bond length outliers.

There are no bond angle outliers.

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	D	1771	0	1766	57	0
2	A	3430	0	3438	132	0
3	E	6720	0	6687	292	0
All	All	11921	0	11891	469	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 469 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)
2:A:368:PHE:HA	2:A:371:LEU:HD23	1.60	0.82

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:E:8:TRP:HB2	3:E:489:PRO:HB3	1.62	0.82
2:A:21:ARG:HB3	2:A:56:SER:H	1.45	0.82
2:A:78:PHE:HB2	2:A:166:LEU:HB2	1.60	0.81
2:A:371:LEU:HB2	2:A:400:ILE:HG23	1.64	0.79

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	D	216/242 (89%)	205 (95%)	11 (5%)	0	100	100
2	A	416/426 (98%)	388 (93%)	26 (6%)	2 (0%)	29	66
3	E	823/1033 (80%)	765 (93%)	56 (7%)	2 (0%)	47	79
All	All	1455/1701 (86%)	1358 (93%)	93 (6%)	4 (0%)	44	74

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	E	365	VAL
2	A	72	ASP
3	E	709	ASN
2	A	53	VAL

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	D	199/215 (93%)	190 (96%)	9 (4%)	27	57
2	A	390/394 (99%)	376 (96%)	14 (4%)	35	63
3	E	755/951 (79%)	692 (92%)	63 (8%)	11	40
All	All	1344/1560 (86%)	1258 (94%)	86 (6%)	21	48

5 of 86 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	E	365	VAL
3	E	654	ASP
3	E	459	CYS
3	E	564	SER
3	E	665	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
3	E	323	ASN
3	E	583	ASN
3	E	538	GLN
3	E	589	GLN
2	A	115	ASN

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 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](#)

There are no chain breaks in this entry.

6 Map visualisation

This section contains visualisations of the EMDB entry EMD-18134. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal standard-deviation projections (False-color)

This section was not generated.

6.5 Orthogonal surface views

This section was not generated.

6.6 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution

This section was not generated.

7.2 Volume estimate versus contour level

This section was not generated.

7.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

This section was not generated.