



# wwPDB EM Validation Summary Report ⓘ

Jan 19, 2025 – 12:16 AM JST

PDB ID : 8YJZ  
EMDB ID : EMD-39354  
Title : Structure of the human endogenous PCNA-FEN1-RNase H2 complex - State D  
Authors : Tian, Y.; Gao, N.  
Deposited on : 2024-03-03  
Resolution : 5.15 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 : **FAILED**  
MolProbity : 4.02b-467  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

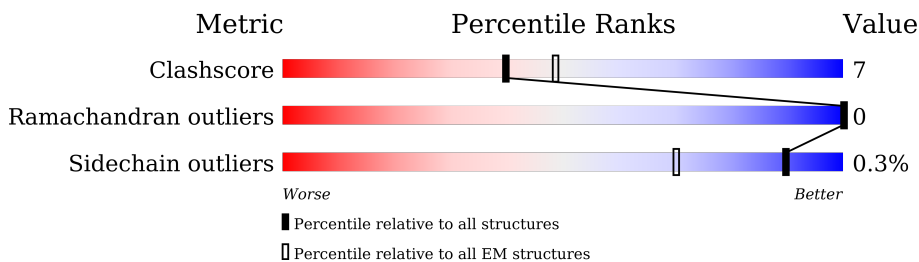
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 5.15 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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  $\geq 3$ , 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	A	261	75% (green), 23% (yellow), 2% (grey), 0% (orange), 0% (red)
1	B	261	75% (green), 23% (yellow), 2% (grey), 0% (orange), 0% (red)
1	C	261	75% (green), 24% (yellow), 1% (grey), 0% (orange), 0% (red)
2	D	380	76% (green), 17% (yellow), 7% (grey), 0% (orange), 0% (red)
3	H	299	86% (green), 13% (yellow), 1% (grey), 0% (orange), 0% (red)
4	G	312	54% (green), 9% (yellow), 37% (grey), 0% (orange), 0% (red)
5	I	164	62% (green), 13% (yellow), 24% (grey), 1% (orange), 0% (red)
6	J	20	80% (green), 20% (yellow), 0% (grey), 0% (orange), 0% (red)
7	E	31	74% (green), 26% (yellow), 0% (grey), 0% (orange), 0% (red)

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
8	F	14	 71% 29%

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 14948 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 Proliferating cell nuclear antigen.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	B	257	Total	C	N	O	S	0	0
			1980	1242	324	398	16		
1	C	257	Total	C	N	O	S	0	0
			1980	1242	324	398	16		
1	A	256	Total	C	N	O	S	0	0
			1972	1238	323	395	16		

- Molecule 2 is a protein called Flap endonuclease 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	D	354	Total	C	N	O	S	0	0
			2791	1756	489	530	16		

- Molecule 3 is a protein called Ribonuclease H2 subunit A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	H	299	Total	C	N	O	S	0	0
			2348	1483	401	453	11		

- Molecule 4 is a protein called Ribonuclease H2 subunit B.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	G	196	Total	C	N	O	S	0	0
			1598	1046	265	281	6		

- Molecule 5 is a protein called Ribonuclease H2 subunit C.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	I	125	Total	C	N	O	S	0	0
			951	609	172	167	3		

- Molecule 6 is a DNA chain called upstream DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
6	J	20	407	200	61	126	20	0	0

- Molecule 7 is a DNA chain called parent strand DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
7	E	31	637	310	122	175	30	0	0

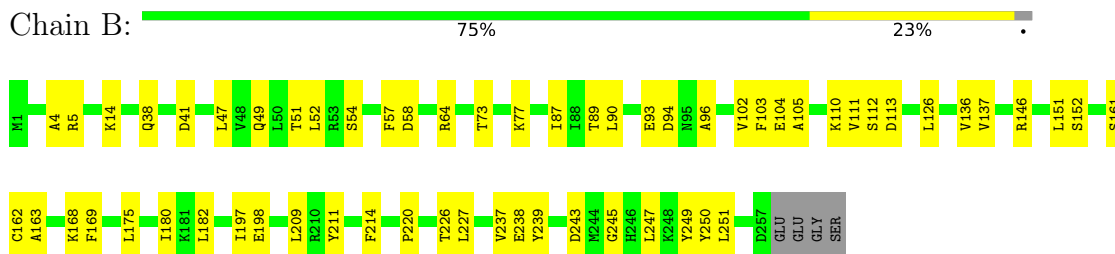
- Molecule 8 is a DNA chain called downstream DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
8	F	14	284	140	40	90	14	0	0

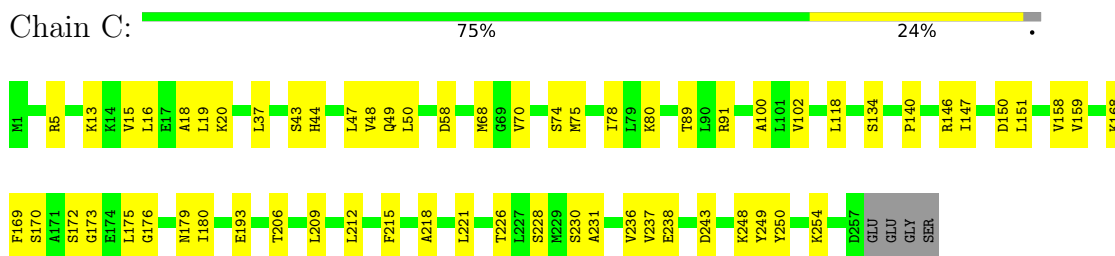
### 3 Residue-property plots [i](#)

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.

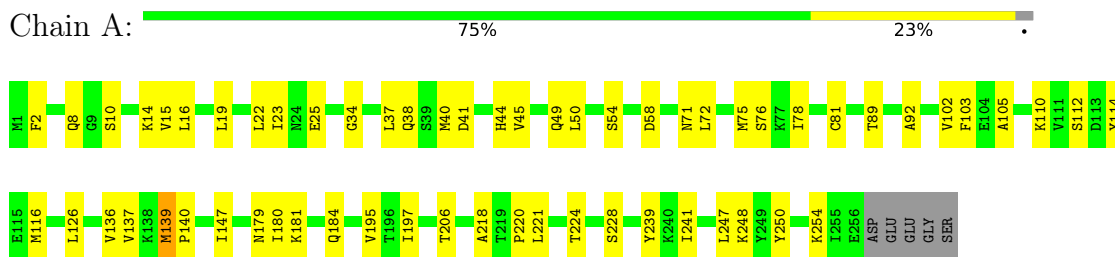
- Molecule 1: Proliferating cell nuclear antigen



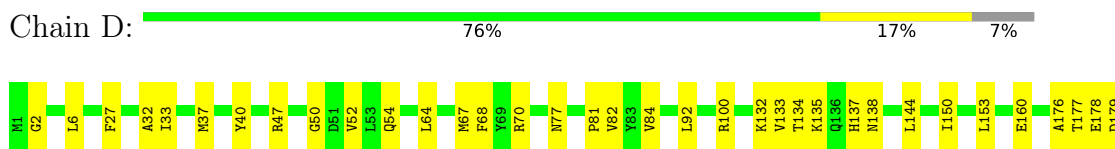
- Molecule 1: Proliferating cell nuclear antigen



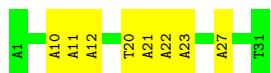
- Molecule 1: Proliferating cell nuclear antigen



- Molecule 2: Flap endonuclease 1







- Molecule 8: downstream DNA





## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	13906	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	60	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k 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	A	0.29	0/1998	0.67	1/2699 (0.0%)
1	B	0.29	0/2006	0.63	2/2710 (0.1%)
1	C	0.30	0/2006	0.72	3/2710 (0.1%)
2	D	0.29	0/2838	0.65	2/3813 (0.1%)
3	H	0.30	0/2397	0.65	2/3253 (0.1%)
4	G	0.28	0/1634	0.60	0/2201
5	I	0.34	0/977	0.72	2/1334 (0.1%)
6	J	0.56	0/453	1.13	0/697
7	E	0.55	0/718	1.05	0/1106
8	F	0.55	0/315	1.21	0/484
All	All	0.33	0/15342	0.72	12/21007 (0.1%)

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	I	135	GLU	CA-CB-CG	7.39	129.66	113.40
1	A	139	MET	CA-CB-CG	6.43	124.23	113.30
1	C	175	LEU	CA-CB-CG	6.35	129.90	115.30
2	D	67	MET	CA-CB-CG	5.97	123.44	113.30
5	I	134	LEU	CA-CB-CG	5.95	128.98	115.30

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	1972	0	1979	37	0
1	B	1980	0	1983	32	0
1	C	1980	0	1983	36	0
2	D	2791	0	2822	35	0
3	H	2348	0	2323	27	0
4	G	1598	0	1641	21	0
5	I	951	0	947	18	0
6	J	407	0	234	4	0
7	E	637	0	354	6	0
8	F	284	0	165	3	0
All	All	14948	0	14431	197	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 7.

The worst 5 of 197 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:D:177:THR:HG22	2:D:179:ASP:H	1.59	0.64
2:D:192:ARG:HB3	2:D:205:GLN:HB2	1.80	0.64
1:A:103:PHE:HB2	1:A:112:SER:HB2	1.81	0.62
4:G:36:MET:HG3	5:I:15:HIS:HA	1.84	0.60
1:C:89:THR:HB	1:C:102:VAL:HB	1.85	0.59

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	254/261 (97%)	245 (96%)	9 (4%)	0	100	100
1	B	255/261 (98%)	250 (98%)	5 (2%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	255/261 (98%)	243 (95%)	12 (5%)	0	100	100
2	D	352/380 (93%)	338 (96%)	14 (4%)	0	100	100
3	H	297/299 (99%)	277 (93%)	20 (7%)	0	100	100
4	G	190/312 (61%)	183 (96%)	7 (4%)	0	100	100
5	I	121/164 (74%)	118 (98%)	3 (2%)	0	100	100
All	All	1724/1938 (89%)	1654 (96%)	70 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	224/228 (98%)	224 (100%)	0	100	100
1	B	225/228 (99%)	224 (100%)	1 (0%)	89	91
1	C	225/228 (99%)	224 (100%)	1 (0%)	89	91
2	D	303/322 (94%)	301 (99%)	2 (1%)	81	87
3	H	257/257 (100%)	256 (100%)	1 (0%)	89	91
4	G	179/277 (65%)	179 (100%)	0	100	100
5	I	98/132 (74%)	98 (100%)	0	100	100
All	All	1511/1672 (90%)	1506 (100%)	5 (0%)	90	92

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

Mol	Chain	Res	Type
1	B	146	ARG
2	D	77	ASN
2	D	345	LYS
3	H	280	ARG
1	C	20	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	49	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 oligosaccharides 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.