



wwPDB EM Validation Summary Report ⓘ

Sep 1, 2024 – 12:09 AM JST

PDB ID : 8YEO
EMDB ID : EMD-39200
Title : Type I-FHNN Cascade-dsDNA R-loop complex
Authors : Li, Z.
Deposited on : 2024-02-22
Resolution : 3.44 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

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:

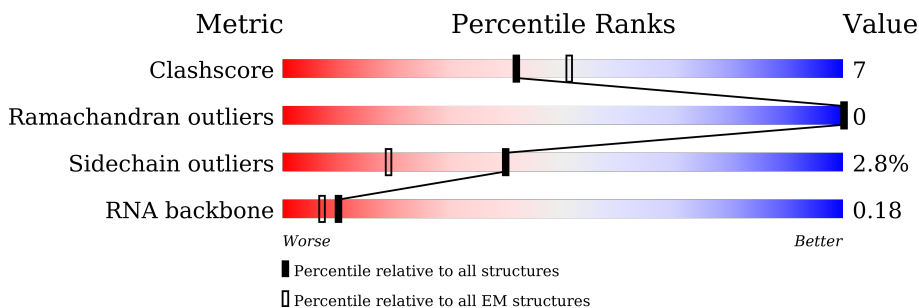
EMDB validation analysis : **FAILED**
MolProbity : 4.02b-467
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.38.2

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.44 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.






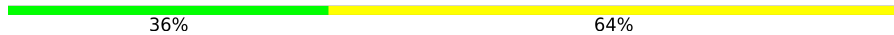
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

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	A	255	76% (Green), 23% (Yellow), . (Grey)
2	J	344	68% (Green), 30% (Yellow), .. (Grey)
3	B	181	68% (Green), 27% (Yellow), .. (Grey)
4	D	335	72% (Green), 16% (Yellow), 12% (Grey)
4	E	335	81% (Green), 14% (Yellow), . (Grey)
4	F	335	84% (Green), 12% (Yellow), . (Grey)
4	G	335	79% (Green), 17% (Yellow), .. (Grey)
4	H	335	83% (Green), 14% (Yellow), . (Grey)

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	I	335	 84% 12%
5	C	60	 35% 53% 12%
6	T	48	 73% 27%
7	N	11	 36% 64%

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 24197 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 Cas5f.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	255	2010	1281	338	377	14	0	0

- Molecule 2 is a protein called Cas8f fusion with HNH.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	J	340	2718	1719	468	518	13	0	0

- Molecule 3 is a protein called Cas6f.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	B	175	1415	913	239	258	5	0	0

- Molecule 4 is a protein called Cas7f.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	I	324	2639	1688	436	500	15	0	0
4	D	295	2412	1546	397	456	13	0	0
4	E	322	2621	1677	433	496	15	0	0
4	F	323	2630	1682	434	499	15	0	0
4	G	323	2630	1682	434	499	15	0	0
4	H	323	2630	1682	434	499	15	0	0

- Molecule 5 is a RNA chain called 60-nt crRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
5	C	60	1287	577	240	411	59	0	0

- Molecule 6 is DNA/RNA hybrid called TS.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
6	T	48	975	468	168	291	48	0	0

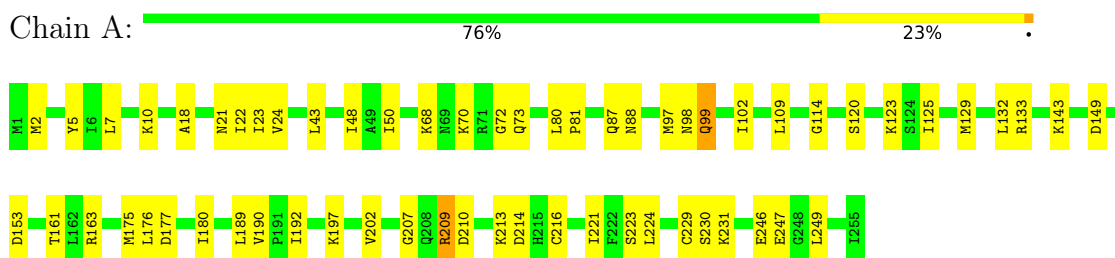
- Molecule 7 is DNA/RNA hybrid called NTS.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
7	N	11	230	109	44	66	11	0	0

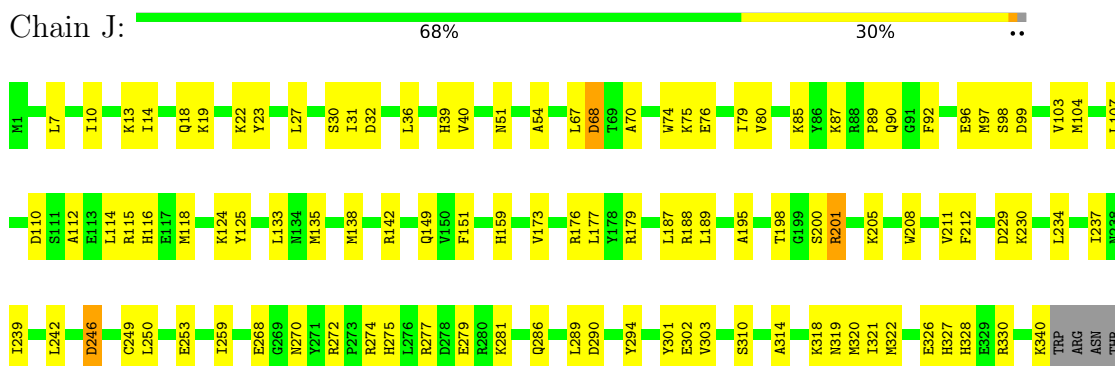
3 Residue-property plots [i](#)

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.

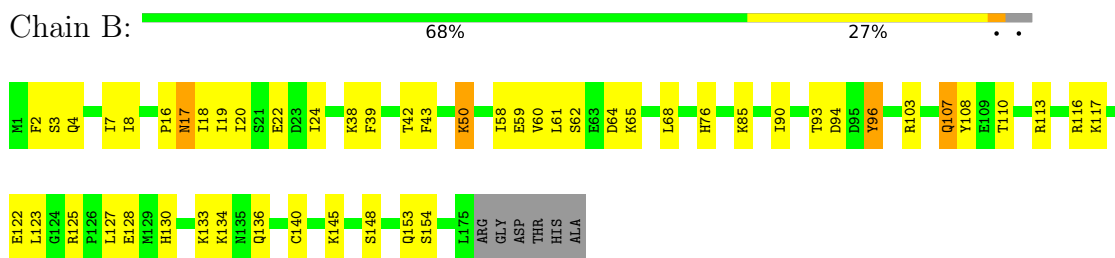
- Molecule 1: Cas5f



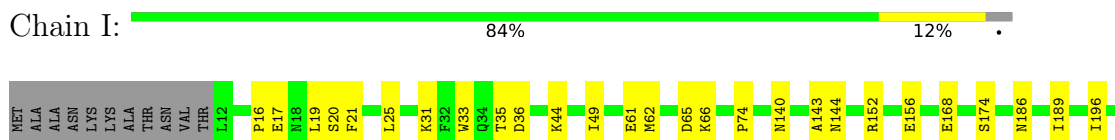
- Molecule 2: Cas8f fusion with HNH

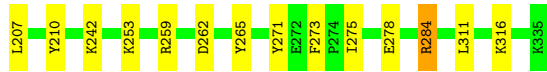


- Molecule 3: Cas6f

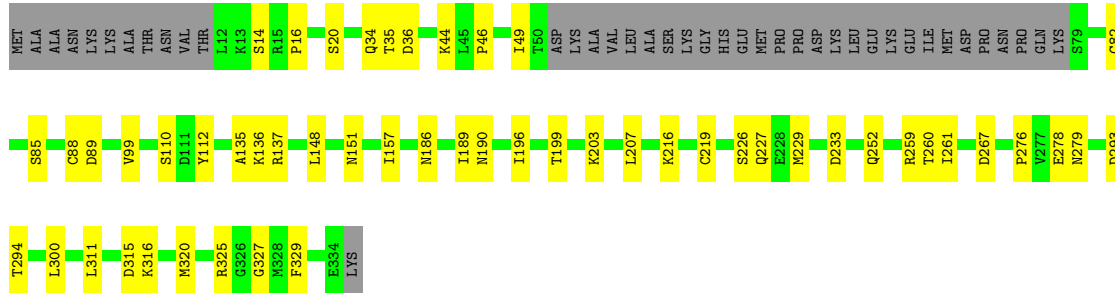


- Molecule 4: Cas7f

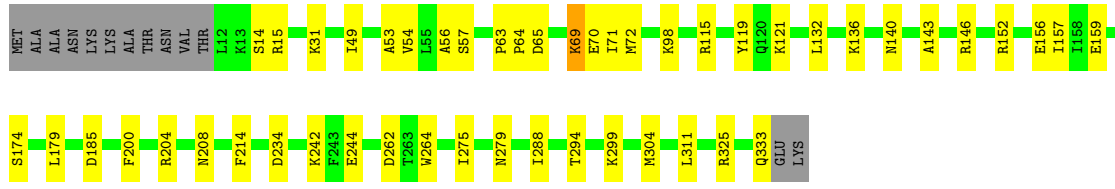
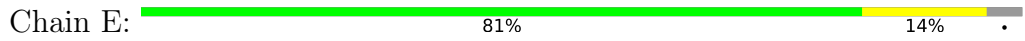




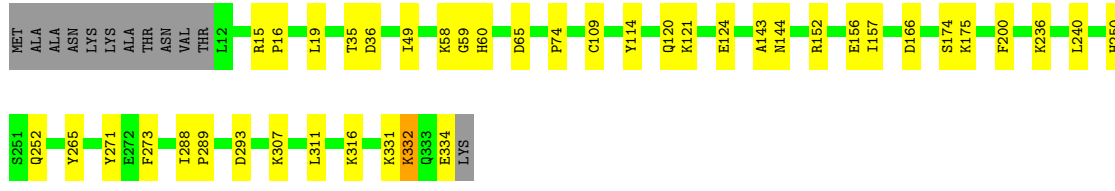
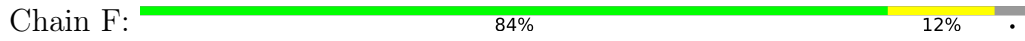
• Molecule 4: Cas7f



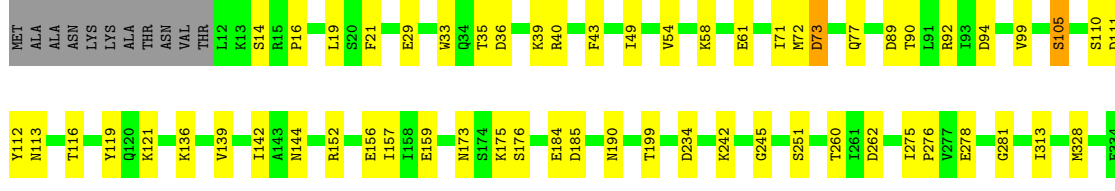
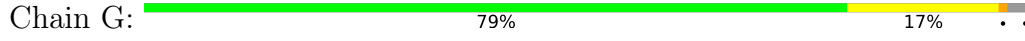
• Molecule 4: Cas7f



• Molecule 4: Cas7f



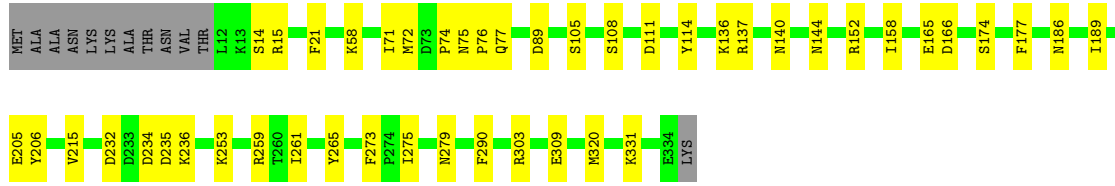
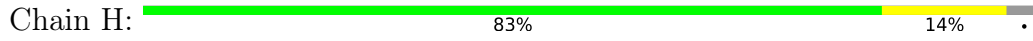
• Molecule 4: Cas7f



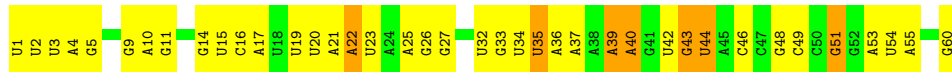
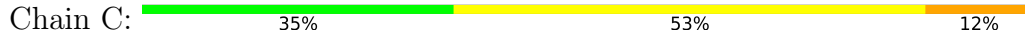
• Molecule 4: Cas7f



• Molecule 4: Cas7f



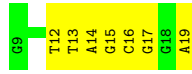
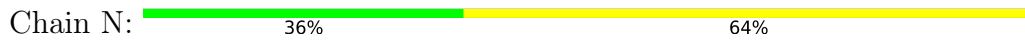
• Molecule 5: 60-nt crRNA



• Molecule 6: TS



• Molecule 7: NTS



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	48756	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	54	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (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	0.26	0/2045	0.53	0/2756
2	J	0.25	0/2773	0.53	0/3745
3	B	0.26	0/1446	0.53	0/1945
4	D	0.26	0/2465	0.51	0/3327
4	E	0.25	0/2680	0.48	0/3618
4	F	0.26	0/2689	0.48	0/3630
4	G	0.25	0/2689	0.48	0/3630
4	H	0.25	0/2689	0.47	0/3630
4	I	0.25	0/2698	0.46	0/3641
5	C	0.25	0/1443	0.89	0/2250
6	T	0.50	0/1090	1.01	0/1678
7	N	0.40	0/258	0.91	0/397
All	All	0.27	0/24965	0.57	0/34247

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	A	2010	0	2030	40	0
2	J	2718	0	2708	69	0
3	B	1415	0	1437	33	0
4	D	2412	0	2379	28	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	2621	0	2596	26	0
4	F	2630	0	2602	27	0
4	G	2630	0	2602	38	0
4	H	2630	0	2602	27	0
4	I	2639	0	2615	23	0
5	C	1287	0	645	24	0
6	T	975	0	545	15	0
7	N	230	0	125	4	0
All	All	24197	0	22886	311	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 311 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)
6:T:8:DC:H2''	6:T:9:DA:H5'	1.64	0.80
2:J:103:VAL:O	2:J:107:LEU:HB2	1.93	0.69
4:I:74:PRO:HG2	6:T:38:A:H1'	1.76	0.68
1:A:50:ILE:HG22	1:A:114:GLY:HA3	1.77	0.66
1:A:80:LEU:HD12	1:A:81:PRO:HD2	1.78	0.66

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	253/255 (99%)	236 (93%)	17 (7%)	0	100 100
2	J	338/344 (98%)	316 (94%)	22 (6%)	0	100 100
3	B	173/181 (96%)	153 (88%)	20 (12%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	D	291/335 (87%)	279 (96%)	12 (4%)	0	100	100
4	E	320/335 (96%)	305 (95%)	15 (5%)	0	100	100
4	F	321/335 (96%)	304 (95%)	17 (5%)	0	100	100
4	G	321/335 (96%)	308 (96%)	13 (4%)	0	100	100
4	H	321/335 (96%)	311 (97%)	10 (3%)	0	100	100
4	I	322/335 (96%)	314 (98%)	8 (2%)	0	100	100
All	All	2660/2790 (95%)	2526 (95%)	134 (5%)	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	219/219 (100%)	212 (97%)	7 (3%)	34	62
2	J	299/304 (98%)	283 (95%)	16 (5%)	18	46
3	B	156/160 (98%)	150 (96%)	6 (4%)	28	56
4	D	265/299 (89%)	258 (97%)	7 (3%)	41	67
4	E	289/299 (97%)	281 (97%)	8 (3%)	38	65
4	F	290/299 (97%)	286 (99%)	4 (1%)	62	79
4	G	290/299 (97%)	283 (98%)	7 (2%)	44	68
4	H	290/299 (97%)	284 (98%)	6 (2%)	48	71
4	I	291/299 (97%)	285 (98%)	6 (2%)	48	71
All	All	2389/2477 (96%)	2322 (97%)	67 (3%)	40	65

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	G	105	SER
4	G	185	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	H	279	ASN
3	B	50	LYS
3	B	17	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	12	ASN
4	G	77	GLN
4	H	279	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
5	C	59/60 (98%)	24 (40%)	0
6	T	0/48	-	-
7	N	0/11	-	-
All	All	59/119 (49%)	24 (40%)	0

5 of 24 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
5	C	3	U
5	C	9	G
5	C	15	U
5	C	16	C
5	C	19	U

There are no RNA pucker outliers to report.

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.