



wwPDB EM Validation Summary Report ⓘ

Sep 1, 2024 – 12:12 AM JST

PDB ID : 8YHA
EMDB ID : EMD-39286
Title : Type I-EHNNH Cascade-ssDNA complex
Authors : Li, Z.
Deposited on : 2024-02-27
Resolution : 3.40 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:

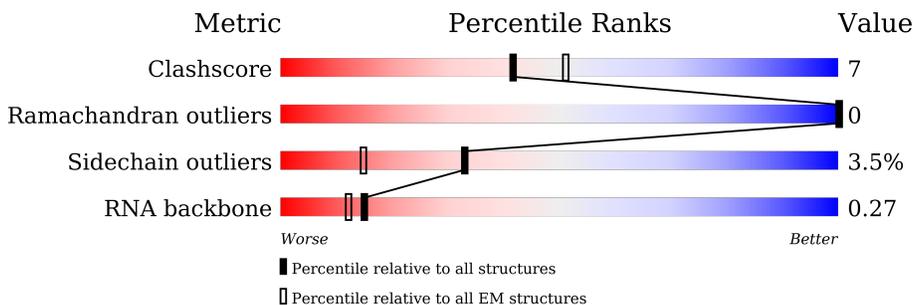
EMDB validation analysis : **FAILED**
MolProbity : 4.02b-467
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 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 3.40 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	388	72% (green), 17% (yellow), 10% (grey)
2	B	272	76% (green), 22% (yellow), 2% (grey)
3	C	61	44% (green), 41% (yellow), 15% (orange)
4	D	378	72% (green), 21% (yellow), 6% (grey)
4	E	378	84% (green), 14% (yellow), 2% (grey)
4	F	378	86% (green), 12% (yellow), 2% (grey)
4	G	378	80% (green), 18% (yellow), 2% (grey)
4	H	378	82% (green), 16% (yellow), 2% (grey)

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	I	378	
5	J	535	
6	K	174	
7	T	56	

2 Entry composition i

There are 8 unique types of molecules in this entry. The entry contains 28503 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 CRISPR system Cascade subunit CasD.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	349	Total	C	N	O	S	0	0
			2756	1736	512	489	19		

- Molecule 2 is a protein called CRISPR-associated endoribonuclease Cse3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	268	Total	C	N	O	S	0	0
			2174	1397	388	384	5		

- Molecule 3 is a RNA chain called 61-nt crRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	C	61	Total	C	N	O	P	0	0
			1303	584	239	420	60		

- Molecule 4 is a protein called CRISPR system Cascade subunit CasC.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	355	Total	C	N	O	S	0	0
			2792	1771	488	522	11		
4	E	375	Total	C	N	O	S	0	0
			2923	1848	510	553	12		
4	F	372	Total	C	N	O	S	0	0
			2897	1830	506	549	12		
4	G	370	Total	C	N	O	S	0	0
			2882	1821	503	546	12		
4	H	371	Total	C	N	O	S	0	0
			2891	1827	505	547	12		
4	I	266	Total	C	N	O	S	0	0
			2056	1308	356	385	7		

- Molecule 5 is a protein called CRISPR-associated protein Cse1 (CRISPR_cse1).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	J	442	3521	2260	602	640	19	0	0

- Molecule 6 is a protein called CRISPR-associated protein Cse2 (CRISPR_cse2).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	K	165	1356	878	239	233	6	0	0

- Molecule 7 is DNA/RNA hybrid called DNA/RNA (47-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
7	T	47	950	457	158	288	47	0	0

- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
8	A	1	Total	Zn	0
			1	1	
8	J	1	Total	Zn	0
			1	1	

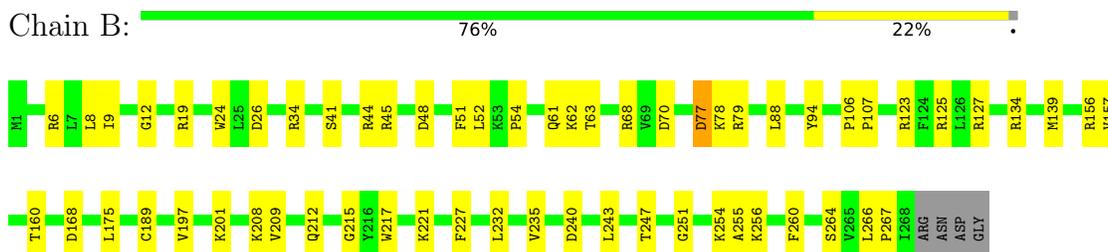
3 Residue-property plots

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.

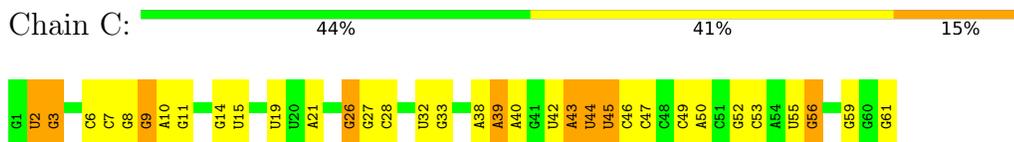
- Molecule 1: CRISPR system Cascade subunit CasD



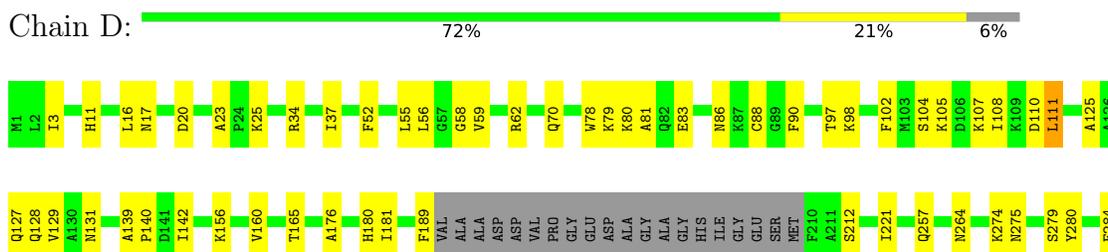
- Molecule 2: CRISPR-associated endoribonuclease Cse3



- Molecule 3: 61-nt crRNA

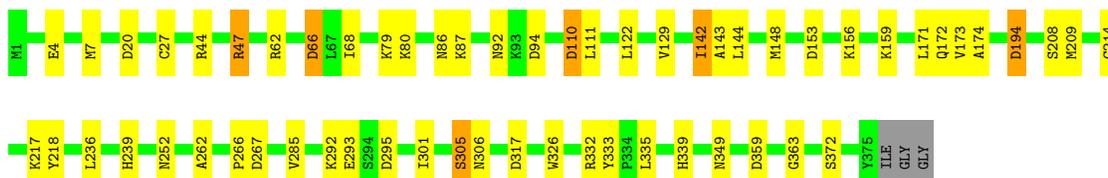
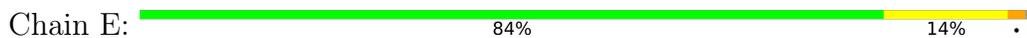


- Molecule 4: CRISPR system Cascade subunit CasC

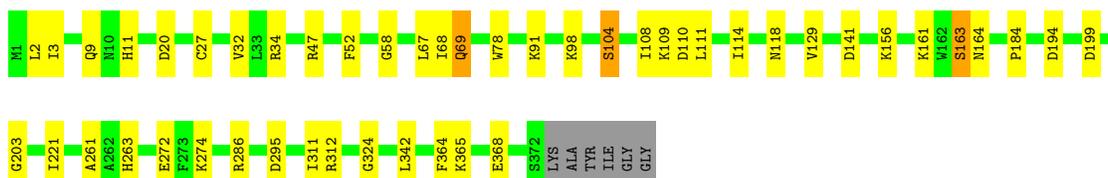
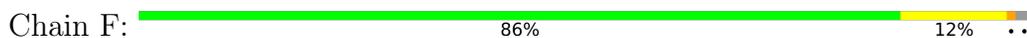




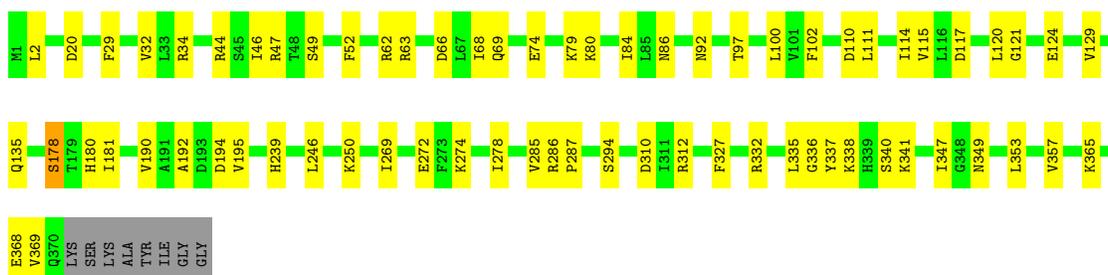
● Molecule 4: CRISPR system Cascade subunit CasC



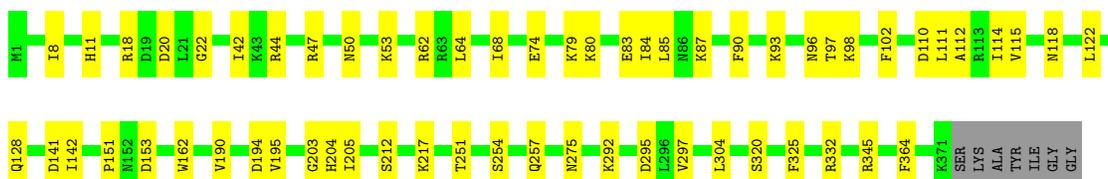
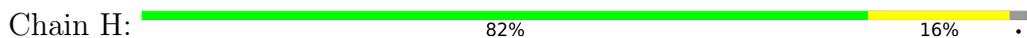
● Molecule 4: CRISPR system Cascade subunit CasC



● Molecule 4: CRISPR system Cascade subunit CasC



● Molecule 4: CRISPR system Cascade subunit CasC



● Molecule 4: CRISPR system Cascade subunit CasC



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	63162	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](#)

Bond lengths and bond angles in the following residue types are not validated in this section:
ZN

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	0/2824	0.57	0/3831
2	B	0.25	0/2232	0.54	0/3022
3	C	0.26	0/1459	0.86	0/2273
4	D	0.26	0/2847	0.51	0/3847
4	E	0.25	0/2981	0.48	0/4031
4	F	0.26	0/2954	0.49	0/3995
4	G	0.25	0/2939	0.48	0/3976
4	H	0.26	0/2948	0.48	0/3987
4	I	0.26	0/2102	0.49	0/2849
5	J	0.25	0/3621	0.52	0/4927
6	K	0.25	0/1386	0.52	0/1866
7	T	0.54	0/1060	1.02	0/1631
All	All	0.27	0/29353	0.56	0/40235

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	2756	0	2752	65	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	2174	0	2189	35	0
3	C	1303	0	659	23	0
4	D	2792	0	2791	47	0
4	E	2923	0	2905	38	0
4	F	2897	0	2878	30	0
4	G	2882	0	2860	42	0
4	H	2891	0	2873	38	0
4	I	2056	0	2016	25	0
5	J	3521	0	3453	50	0
6	K	1356	0	1398	17	0
7	T	950	0	535	44	0
8	A	1	0	0	0	0
8	J	1	0	0	0	0
All	All	28503	0	27309	366	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 366 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:349:ARG:HH22	7:T:6:DG:H1'	1.13	1.11
1:A:287:ARG:NH2	7:T:10:DA:OP1	1.87	1.06
6:K:49:ARG:NH2	7:T:7:DC:H5	1.57	1.02
6:K:49:ARG:HH22	7:T:7:DC:H5	1.01	0.96
1:A:313:TYR:CD2	7:T:6:DG:C4	2.57	0.93

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	345/388 (89%)	321 (93%)	24 (7%)	0	100	100
2	B	266/272 (98%)	240 (90%)	26 (10%)	0	100	100
4	D	351/378 (93%)	322 (92%)	29 (8%)	0	100	100
4	E	373/378 (99%)	357 (96%)	16 (4%)	0	100	100
4	F	370/378 (98%)	352 (95%)	18 (5%)	0	100	100
4	G	368/378 (97%)	359 (98%)	9 (2%)	0	100	100
4	H	369/378 (98%)	355 (96%)	14 (4%)	0	100	100
4	I	260/378 (69%)	249 (96%)	11 (4%)	0	100	100
5	J	440/535 (82%)	396 (90%)	44 (10%)	0	100	100
6	K	163/174 (94%)	152 (93%)	11 (7%)	0	100	100
All	All	3305/3637 (91%)	3103 (94%)	202 (6%)	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	289/322 (90%)	280 (97%)	9 (3%)	35	60
2	B	235/238 (99%)	225 (96%)	10 (4%)	25	50
4	D	300/313 (96%)	288 (96%)	12 (4%)	27	52
4	E	312/313 (100%)	299 (96%)	13 (4%)	25	51
4	F	310/313 (99%)	301 (97%)	9 (3%)	37	61
4	G	308/313 (98%)	298 (97%)	10 (3%)	34	59
4	H	309/313 (99%)	302 (98%)	7 (2%)	45	67
4	I	218/313 (70%)	212 (97%)	6 (3%)	38	62
5	J	377/459 (82%)	361 (96%)	16 (4%)	25	51
6	K	144/153 (94%)	139 (96%)	5 (4%)	31	56
All	All	2802/3050 (92%)	2705 (96%)	97 (4%)	33	56

5 of 97 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	G	110	ASP
4	I	212	SER
4	G	135	GLN
4	H	102	PHE
4	I	365	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
4	D	127	GLN
4	E	86	ASN
4	F	9	GLN
5	J	178	GLN
5	J	200	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	C	60/61 (98%)	25 (41%)	1 (1%)
7	T	0/56	-	-
All	All	60/117 (51%)	25 (41%)	1 (1%)

5 of 25 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	C	2	U
3	C	3	G
3	C	9	G
3	C	10	A
3	C	14	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	C	45	U

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](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.