



wwPDB EM Validation Summary Report i

Jul 15, 2025 – 10:04 PM JST

PDB ID : 8K26 / pdb_00008k26
EMDB ID : EMD-36829
Title : Structure of Cas1-Cas2 complex
Authors : Zhang, L.X.; Feng, Y.
Deposited on : 2023-07-12
Resolution : 3.60 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 i 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](#) i) were used in the production of this report:

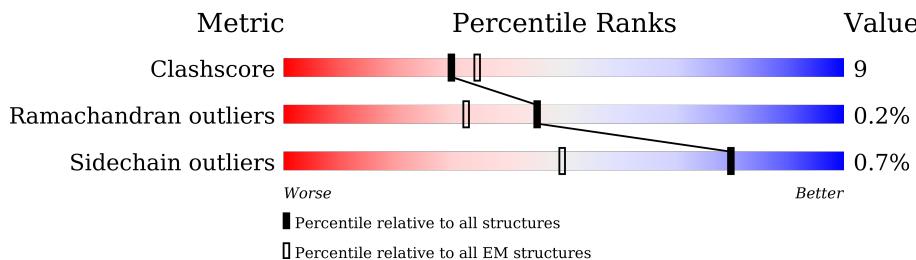
EMDB validation analysis : **FAILED**
MolProbity : 4-5-2 with Phenix2.0rc1
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.44

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 3.60 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 ≥ 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 $\leq 5\%$.



2 Entry composition [\(i\)](#)

There are 2 unique types of molecules in this entry. The entry contains 10549 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 Cas1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	D	289	Total	C	N	O	S		
			2292	1458	395	427	12	2	0
1	E	295	Total	C	N	O	S		
			2343	1494	403	434	12	1	0
1	d	289	Total	C	N	O	S		
			2296	1466	392	425	13	2	0
1	e	295	Total	C	N	O	S		
			2326	1480	401	432	13	1	0

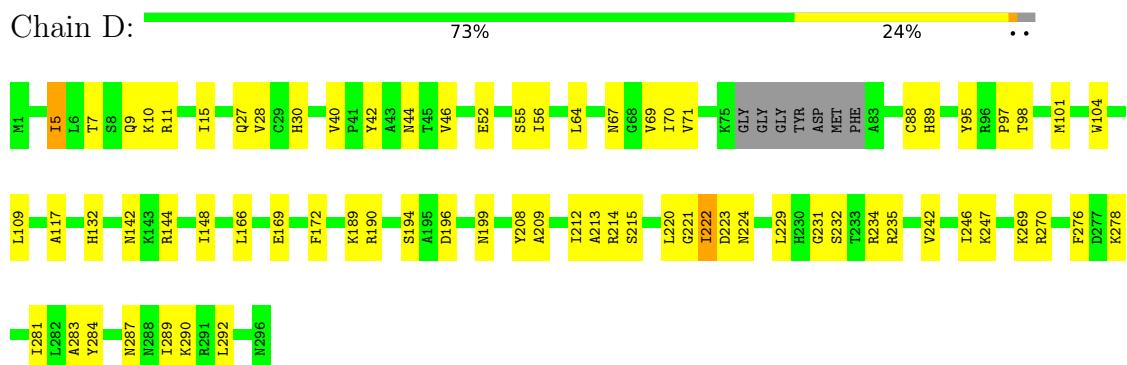
- Molecule 2 is a protein called HD Cas3-type domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	A	84	Total	C	N	O	S		
			649	414	112	119	4	0	0
2	a	84	Total	C	N	O	S		
			643	410	112	117	4	0	0

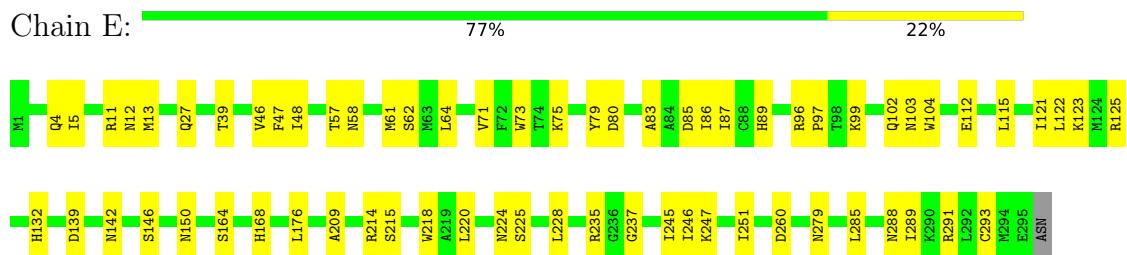
3 Residue-property plots [i](#)

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: Cas1



- Molecule 1: Cas1



- Molecule 1: Cas1



- Molecule 1: Cas1





- Molecule 2: HD Cas3-type domain-containing protein

Chain A: 76% 24%



- Molecule 2: HD Cas3-type domain-containing protein

Chain a: 88% 11%



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	101197	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 ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	1500	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	D	0.35	0/2342	0.85	5/3163 (0.2%)
1	E	0.28	0/2395	0.71	6/3235 (0.2%)
1	d	0.30	0/2346	0.74	1/3170 (0.0%)
1	e	0.30	0/2378	0.74	2/3214 (0.1%)
2	A	0.30	0/660	0.70	0/888
2	a	0.27	0/654	0.74	0/880
All	All	0.31	0/10775	0.76	14/14550 (0.1%)

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	D	222	ILE	N-CA-C	13.47	127.62	108.36
1	D	221	GLY	N-CA-C	6.33	128.19	113.18
1	e	154	LYS	CB-CG-CD	6.16	125.47	111.30
1	d	237	GLY	N-CA-C	5.82	117.26	111.21
1	e	279	ASN	CA-CB-CG	5.68	118.28	112.60

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	2292	0	2273	49	0
1	E	2343	0	2316	43	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	d	2296	0	2271	53	0
1	e	2326	0	2272	35	0
2	A	649	0	668	17	0
2	a	643	0	657	6	0
All	All	10549	0	10457	186	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 9.

The worst 5 of 186 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:d:98:THR:O	1:d:102:GLN:HB2	1.72	0.90
2:a:18:SER:O	2:a:22:SER:HB3	1.91	0.70
1:D:213:ALA:HB1	1:D:242:VAL:HG23	1.73	0.70
1:e:48:ILE:O	1:e:71:VAL:HA	1.93	0.69
1:D:220:LEU:HD22	1:D:289:ILE:HG23	1.76	0.68

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	287/296 (97%)	261 (91%)	25 (9%)	1 (0%)	37 67
1	E	294/296 (99%)	285 (97%)	9 (3%)	0	100 100
1	d	286/296 (97%)	263 (92%)	22 (8%)	1 (0%)	37 67
1	e	294/296 (99%)	284 (97%)	10 (3%)	0	100 100
2	A	82/84 (98%)	78 (95%)	4 (5%)	0	100 100
2	a	82/84 (98%)	76 (93%)	5 (6%)	1 (1%)	11 43

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1325/1352 (98%)	1247 (94%)	75 (6%)	3 (0%)	45 73

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	a	82	PRO
1	D	5	ILE
1	d	5	ILE

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	247/255 (97%)	243 (98%)	4 (2%)	58 76
1	E	250/255 (98%)	250 (100%)	0	100 100
1	d	247/255 (97%)	247 (100%)	0	100 100
1	e	245/255 (96%)	244 (100%)	1 (0%)	89 95
2	A	72/72 (100%)	72 (100%)	0	100 100
2	a	70/72 (97%)	67 (96%)	3 (4%)	25 54
All	All	1131/1164 (97%)	1123 (99%)	8 (1%)	80 90

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	a	68	PHE
2	a	44	ASN
1	e	279	ASN
1	D	222	ILE
2	a	40	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	e	197	ASN
1	e	257	HIS
2	A	51	ASN
1	e	279	ASN
2	A	40	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.