



wwPDB EM Map/Model Validation Report ⓘ

Oct 11, 2016 – 02:32 PM EDT

PDB ID : 5GQH
EMDB ID: : EMD-9535
Title : Cryo-EM structure of PaeCas3-AcrF3 complex
Authors : Zhang, X.; Ma, J.; Wang, Y.; Wang, J.
Deposited on : 2016-08-07
Resolution : 4.20 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>

MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027939

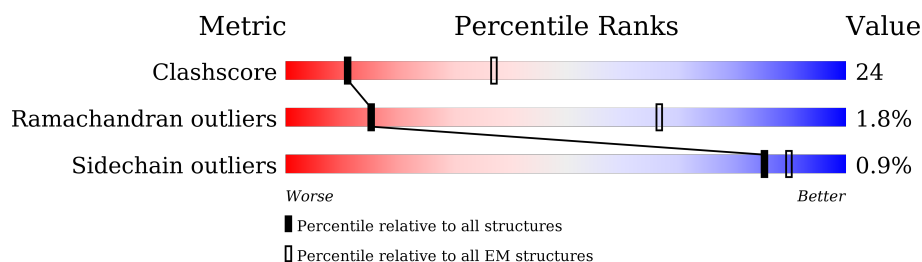
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.20 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	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	1090	
2	B	139	
2	C	139	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6641 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-associated nuclease/helicase Cas3 subtype I-F/YPEST.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	901	Total	C	N	O	S	0	0
			4473	2666	904	902	1		

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MET	-	expression tag	UNP Q02ML8
A	-12	GLY	-	expression tag	UNP Q02ML8
A	-11	SER	-	expression tag	UNP Q02ML8
A	-10	SER	-	expression tag	UNP Q02ML8
A	-9	HIS	-	expression tag	UNP Q02ML8
A	-8	HIS	-	expression tag	UNP Q02ML8
A	-7	HIS	-	expression tag	UNP Q02ML8
A	-6	HIS	-	expression tag	UNP Q02ML8
A	-5	HIS	-	expression tag	UNP Q02ML8
A	-4	HIS	-	expression tag	UNP Q02ML8
A	-3	SER	-	expression tag	UNP Q02ML8
A	-2	GLN	-	expression tag	UNP Q02ML8
A	-1	ASP	-	expression tag	UNP Q02ML8
A	0	PRO	-	expression tag	UNP Q02ML8

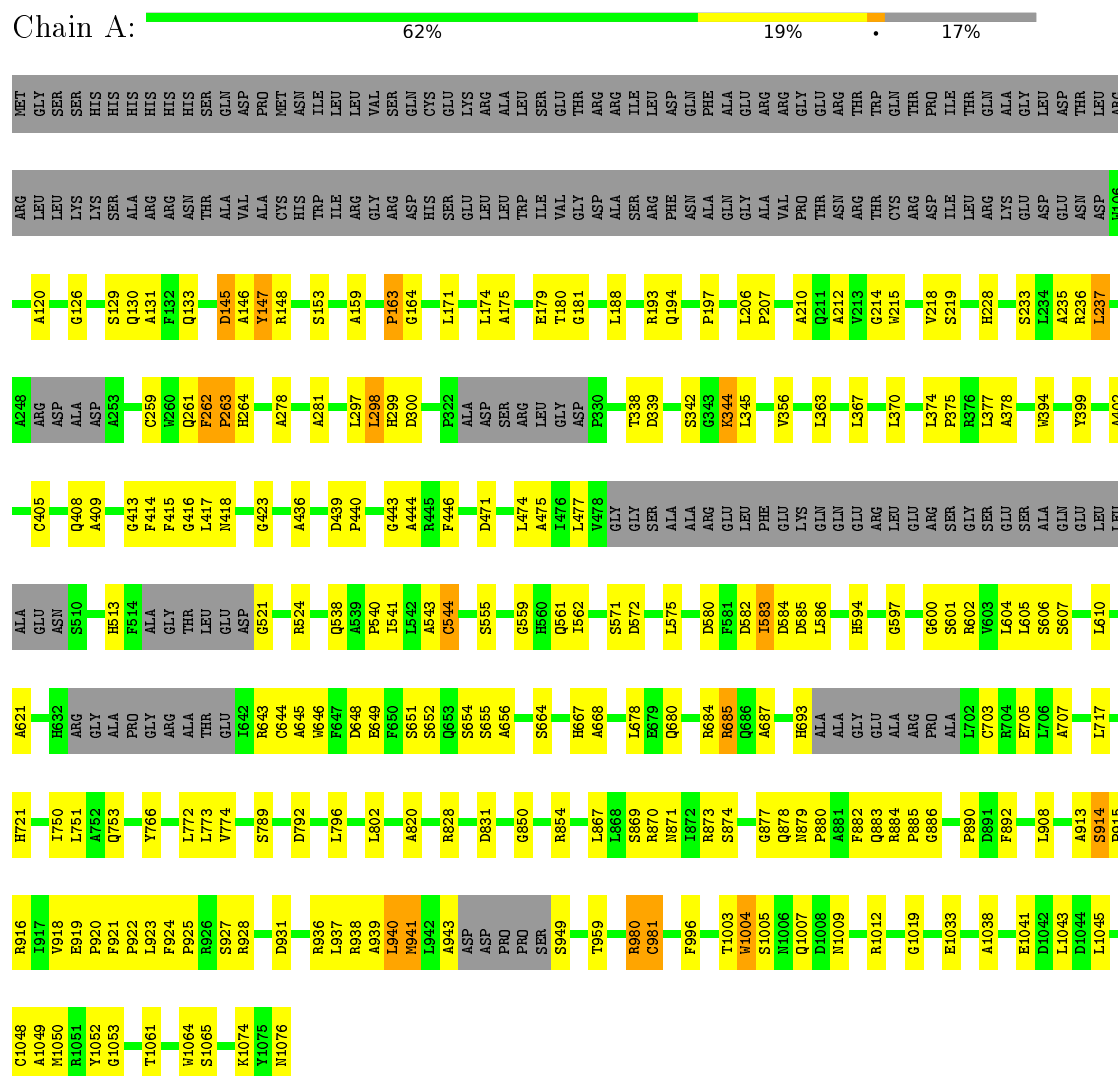
- Molecule 2 is a protein called anti-CRISPR protein 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	135	Total	C	N	O	S	0	0
			1084	682	195	205	2		
2	C	135	Total	C	N	O	S	0	0
			1084	682	195	205	2		

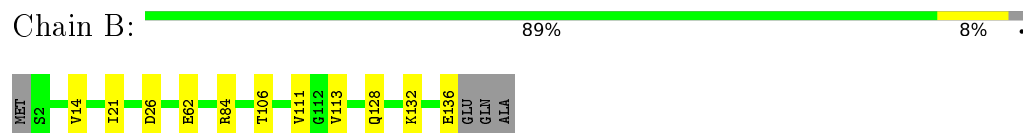
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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-associated nuclease/helicase Cas3 subtype I-F/YPEST



- Molecule 2: anti-CRISPR protein 3



- Molecule 2: anti-CRISPR protein 3

Chain C:  89% 8% .



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	78904	Depositor
Resolution determination method	FSC 0.5 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.62	7/4466 (0.2%)	0.73	8/6202 (0.1%)
2	B	0.49	0/1109	0.57	1/1505 (0.1%)
2	C	0.49	0/1109	0.57	1/1505 (0.1%)
All	All	0.58	7/6684 (0.1%)	0.68	10/9212 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	5

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1004	TRP	N-CA	17.58	1.81	1.46
1	A	1009	ASN	N-CA	-12.69	1.21	1.46
1	A	1050	MET	C-O	-12.54	0.99	1.23
1	A	1004	TRP	CA-C	11.16	1.81	1.52
1	A	1009	ASN	C-O	-9.86	1.04	1.23

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1004	TRP	N-CA-C	7.00	129.91	111.00
1	A	1004	TRP	N-CA-CB	-6.77	98.42	110.60
1	A	1050	MET	N-CA-C	-6.43	93.63	111.00
1	A	571	SER	N-CA-C	-6.42	93.67	111.00
1	A	940	LEU	CA-C-O	6.40	133.53	120.10

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	408	GLN	Peptide
1	A	575	LEU	Mainchain,Peptide
1	A	867	LEU	Mainchain
1	A	939	ALA	Peptide

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	4473	0	2152	251	0
2	B	1084	0	1030	13	0
2	C	1084	0	1030	7	0
All	All	6641	0	4212	258	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 24.

The worst 5 of 258 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:126:GLY:HA3	1:A:153:SER:CB	1.39	1.50
1:A:1004:TRP:C	1:A:1004:TRP:CA	1.82	1.49
1:A:1004:TRP:N	1:A:1004:TRP:CA	1.81	1.43
1:A:996:PHE:O	1:A:1007:GLN:CB	1.66	1.42
1:A:126:GLY:CA	1:A:153:SER:CA	2.01	1.36

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	885/1090 (81%)	837 (95%)	27 (3%)	21 (2%)	7	49
2	B	133/139 (96%)	129 (97%)	4 (3%)	0	100	100
2	C	133/139 (96%)	129 (97%)	4 (3%)	0	100	100
All	All	1151/1368 (84%)	1095 (95%)	35 (3%)	21 (2%)	15	55

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	145	ASP
1	A	147	TYR
1	A	197	PRO
1	A	262	PHE
1	A	263	PRO

5.3.2 Protein sidechains ⓘ

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	7/898 (1%)	5 (71%)	2 (29%)	0	4
2	B	109/112 (97%)	109 (100%)	0	100	100
2	C	109/112 (97%)	109 (100%)	0	100	100
All	All	225/1122 (20%)	223 (99%)	2 (1%)	85	93

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	772	LEU
1	A	980	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 carbohydrates 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.