



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Jan 25, 2018 – 07:14 PM EST

PDB ID : 5VZL
EMDB ID: : EMD-8749
Title : cryo-EM structure of the Cas9-sgRNA-AcrIIA4 anti-CRISPR complex
Authors : Jiang, F.; Liu, J.J.; Nogales, E.; Doudna, J.A.
Deposited on : 2017-05-29
Resolution : 3.90 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030736

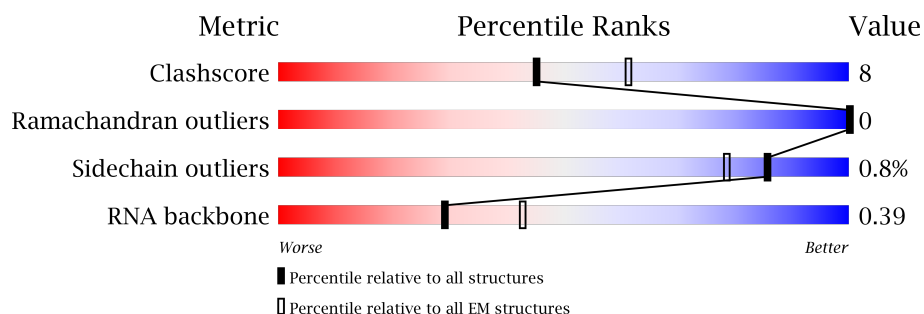
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.90 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026
RNA backbone	3398	335

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	B	118	<div> <div>44%</div> <div>40%</div> <div>13%</div> <div>..</div> </div>
2	A	1369	<div> <div>80%</div> <div>20%</div> </div>
3	C	87	<div> <div>71%</div> <div>28%</div> <div>.</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 13915 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 RNA chain called single guide RNA (116-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
1	B	116	Total	C	N	O	P	0	0
			2494	1112	455	809	118		

- Molecule 2 is a protein called CRISPR-associated endonuclease Cas9.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	A	1364	Total	C	N	O	S	0	0
			10711	6813	1851	2024	23		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP A0A0C6FZC2

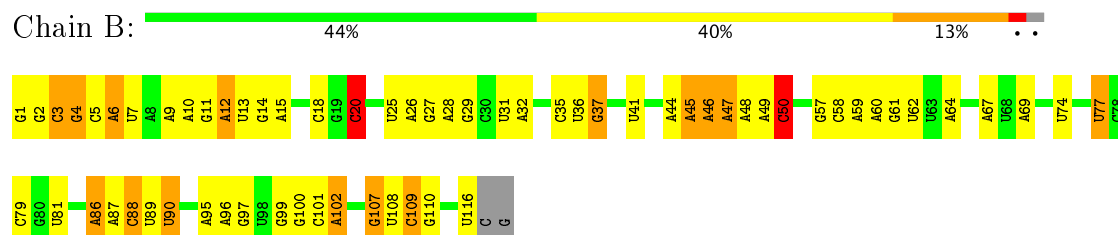
- Molecule 3 is a protein called phage anti-CRISPR AcrIIA4.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	87	Total	C	N	O	S	0	0
			710	439	114	155	2		

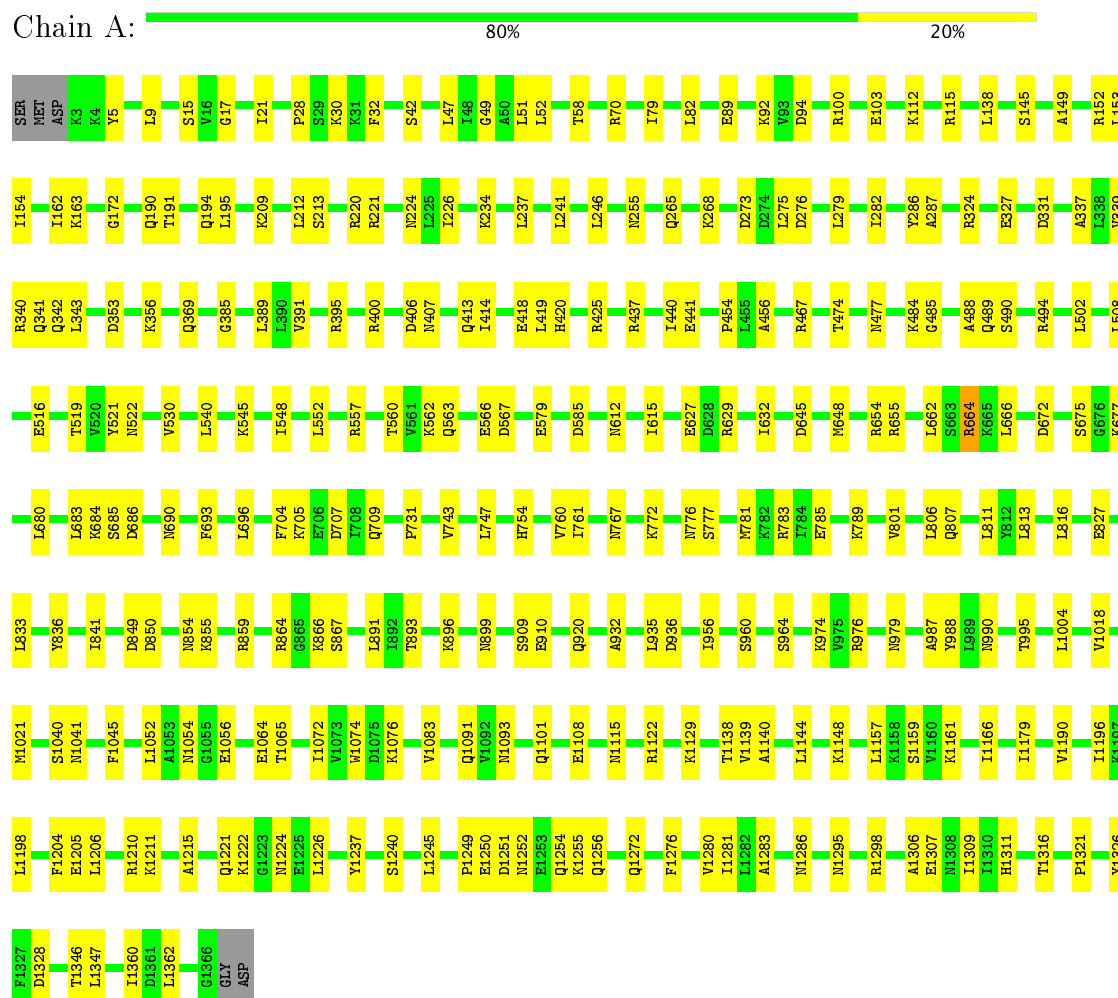
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

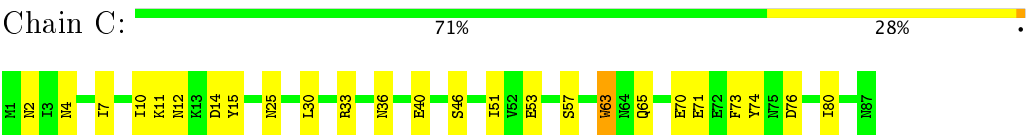
- Molecule 1: single guide RNA (116-MER)



- Molecule 2: CRISPR-associated endonuclease Cas9



● Molecule 3: phage anti-CRISPR AcrIIA4



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	185000	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 ($e^-/\text{\AA}^2$)	46	Depositor
Minimum defocus (nm)	900	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	B	0.65	0/2758	1.19	24/4297 (0.6%)
2	A	0.34	0/10903	0.54	0/14739
3	C	0.37	0/718	0.65	0/969
All	All	0.42	0/14379	0.74	24/20005 (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
2	A	0	2
3	C	0	2
All	All	0	4

There are no bond length outliers.

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	4	G	C4-N9-C1'	7.85	136.70	126.50
1	B	4	G	N3-C4-N9	7.76	130.66	126.00
1	B	4	G	N3-C4-C5	-7.58	124.81	128.60
1	B	88	C	C2-N1-C1'	7.06	126.56	118.80
1	B	20	C	N1-C2-O2	6.97	123.08	118.90

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	A	1316	THR	Peptide
2	A	1326	TYR	Peptide
3	C	46	SER	Peptide
3	C	63	TRP	Peptide

5.2 Too-close contacts

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	B	2494	0	1248	26	0
2	A	10711	0	10460	170	0
3	C	710	0	668	16	0
All	All	13915	0	12376	200	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 8.

The worst 5 of 200 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:A:32:PHE:O	2:A:42:SER:HA	1.77	0.83
2:A:760:VAL:HA	2:A:956:ILE:O	1.82	0.80
2:A:15:SER:HA	2:A:51:LEU:O	1.82	0.77
2:A:1251:ASP:O	2:A:1255:LYS:HB2	1.90	0.71
2:A:562:LYS:O	2:A:566:GLU:HB3	1.90	0.71

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	
2	A	1362/1369 (100%)	1256 (92%)	106 (8%)	0	100	100
3	C	85/87 (98%)	70 (82%)	15 (18%)	0	100	100
All	All	1447/1456 (99%)	1326 (92%)	121 (8%)	0	100	100

There are no Ramachandran outliers to report.

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	
2	A	1118/1228 (91%)	1112 (100%)	6 (0%)	91	96
3	C	82/83 (99%)	79 (96%)	3 (4%)	39	71
All	All	1200/1311 (92%)	1191 (99%)	9 (1%)	86	93

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	664	ARG
3	C	25	ASN
3	C	2	ASN
2	A	437	ARG
2	A	767	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	767	ASN
2	A	863	ASN
3	C	12	ASN
2	A	595	HIS
3	C	2	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	B	114/118 (96%)	43 (37%)	4 (3%)

5 of 43 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	B	3	C
1	B	5	C
1	B	6	A
1	B	7	U
1	B	9	A

All (4) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	B	27	G
1	B	35	C
1	B	47	A
1	B	86	A

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.