wwPDB/EMDataBank EM Map/Model Validation Summary Report (i)

Feb 10, 2019 – 09:33 AM EST

PDB ID	:	6MCC
EMDB ID:	:	EMD-9067
Title	:	CryoEM structure of AcrIIA2 homolog in complex with CRISPR-Cas9
Authors	:	Jiang, F.; Liu, J.J.; Doudna, J.A.
Deposited on	:	2018-08-31
Resolution	:	3.90 Å(reported)
Based on PDB ID	:	4ZT0

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
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

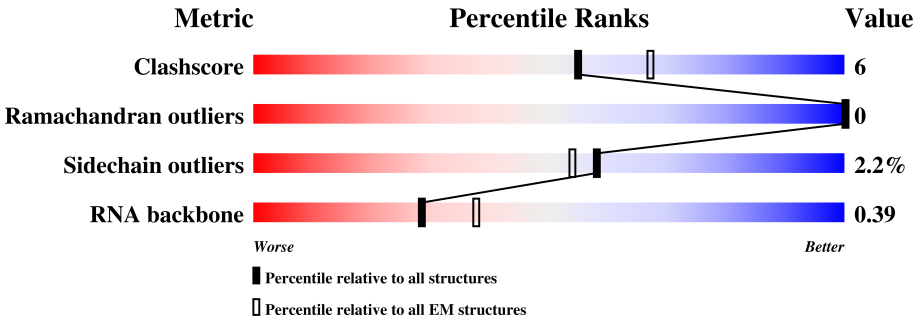
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20031633

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 3.90 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	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531
RNA backbone	3747	458

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 <=5%

Mol	Chain	Length	Quality of chain
1	B	116	43% 43% 12% .
2	A	1368	83% 16%
3	C	130	83% 16% .

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 14175 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
			Total	C	N	O	P		
1	B	116	2494	1112	455	809	118	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	A	1364	10715	6818	1848	2026	23	0	0

- Molecule 3 is a protein called Anti-CRISPR AcrIIA2 Homolog.

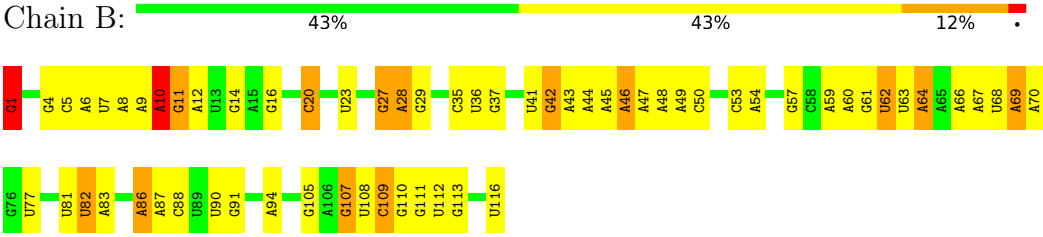
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	130	966	621	155	187	3	0	0

3 Residue-property plots (i)

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.

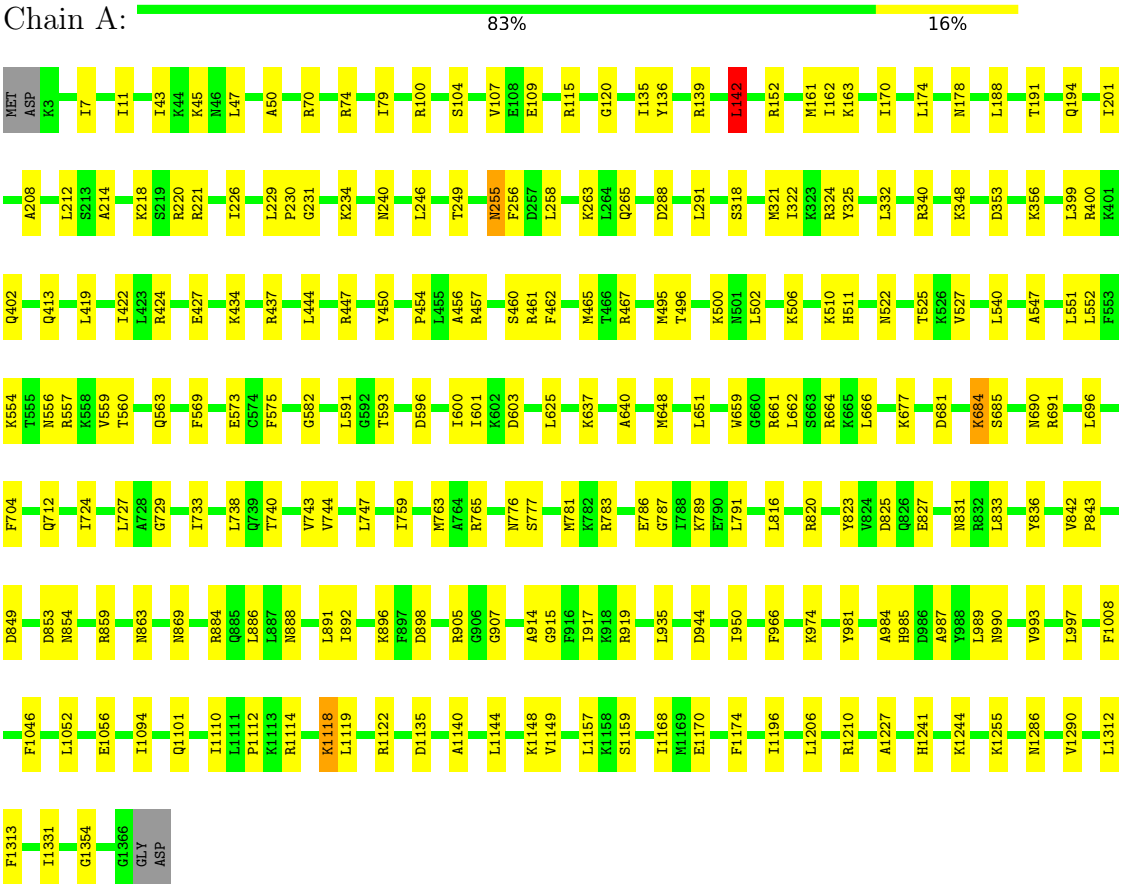
- Molecule 1: Single guide RNA (116-MER)

Chain B:



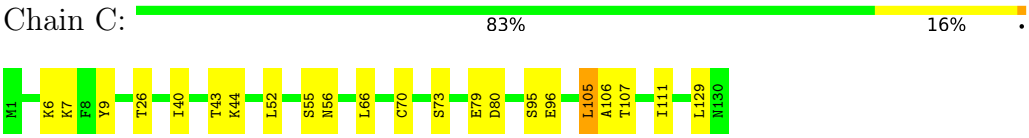
- Molecule 2: CRISPR-associated endonuclease Cas9

Chain A:



- Molecule 3: Anti-CRISPR AcrIIA2 Homolog

Chain C:


83%
16%
M1
K6
K7
F8
Y9
T26
I40
T43
K44
L52
S55
N56
L66
C70
S73
E79
D80
S95
E96
L105
A106
T107
I111
L129
N130

4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	167464	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 ($e^-/Å^2$)	46.2	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	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	0.58	0/2758	1.09	9/4297 (0.2%)
2	A	0.35	0/10907	0.61	2/14742 (0.0%)
3	C	0.38	0/983	0.70	1/1342 (0.1%)
All	All	0.40	0/14648	0.74	12/20381 (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	B	2	0
3	C	0	1
All	All	2	1

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	8	A	O4'-C1'-N9	7.39	114.11	108.20
1	B	20	C	N1-C2-O2	6.54	122.83	118.90
1	B	35	C	C2-N1-C1'	6.31	125.74	118.80
1	B	88	C	C2-N1-C1'	6.17	125.59	118.80
2	A	591	LEU	CA-CB-CG	5.95	128.99	115.30

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	1	GTP	C4',C2'

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	107	THR	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	B	2494	0	1248	31	0
2	A	10715	0	10475	131	0
3	C	966	0	870	11	0
All	All	14175	0	12593	155	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 6.

The worst 5 of 155 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:100:ARG:HH12	2:A:120:GLY:HA2	1.49	0.77
2:A:743:VAL:O	2:A:747:LEU:HB2	1.85	0.75
2:A:915:GLY:O	2:A:919:ARG:HB2	1.87	0.74
2:A:142:LEU:HG	2:A:422:ILE:HD11	1.76	0.68
2:A:139:ARG:NH1	2:A:161:MET:SD	2.69	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	
2	A	1362/1368 (100%)	1247 (92%)	115 (8%)	0	100	100
3	C	128/130 (98%)	116 (91%)	12 (9%)	0	100	100
All	All	1490/1498 (100%)	1363 (92%)	127 (8%)	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	
2	A	1120/1227 (91%)	1095 (98%)	25 (2%)	55	78
3	C	89/118 (75%)	88 (99%)	1 (1%)	76	88
All	All	1209/1345 (90%)	1183 (98%)	26 (2%)	58	78

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	522	ASN
2	A	648	MET
2	A	1313	PHE
2	A	554	LYS
2	A	600	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	831	ASN
2	A	899	ASN
2	A	990	ASN
2	A	674	GLN
2	A	985	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	B	115/116 (99%)	42 (36%)	4 (3%)

5 of 42 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	B	4	G
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	1	GTP
1	B	27	G
1	B	81	U
1	B	86	A

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.