



wwPDB EM Validation Summary Report ⓘ

Nov 7, 2022 – 01:03 AM EST

PDB ID : 6O0X
EMDB ID : EMD-0583
Title : Conformational states of Cas9-sgRNA-DNA ternary complex in the presence of magnesium
Authors : Zhu, X.; Clarke, R.; Puppala, A.K.; Chittori, S.; Merk, A.; Merrill, B.J.; Simonovic, M.; Subramaniam, S.
Deposited on : 2019-02-17
Resolution : 3.28 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 : 0.0.1.dev43
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.31.2

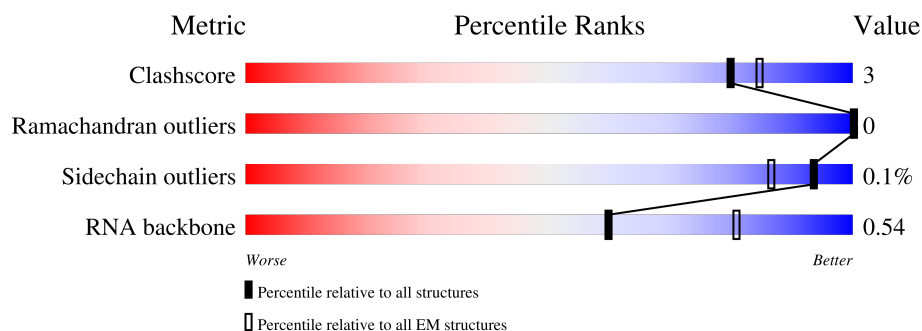
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.28 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	1372	 76% 5% 19%
2	B	102	 60% 30% 5% . .
3	C	13	 69% 31%
4	D	40	 32% 8% 60%
5	c	27	 81% . 15%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11257 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 endonuclease Cas9/Csn1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1112	Total	C	N	O	S	1	0
			8098	5226	1433	1426	13		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q99ZW2
A	-2	ALA	-	expression tag	UNP Q99ZW2
A	-1	ALA	-	expression tag	UNP Q99ZW2
A	0	SER	-	expression tag	UNP Q99ZW2

- Molecule 2 is a RNA chain called single guide RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	98	Total	C	N	O	P	0	0
			2104	941	389	676	98		

- Molecule 3 is a DNA chain called 5' product of target strand DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	13	Total	C	N	O	P	0	0
			261	125	49	75	12		

- Molecule 4 is a DNA chain called non-target strand DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	16	Total	C	N	O	P	0	0
			331	157	59	99	16		

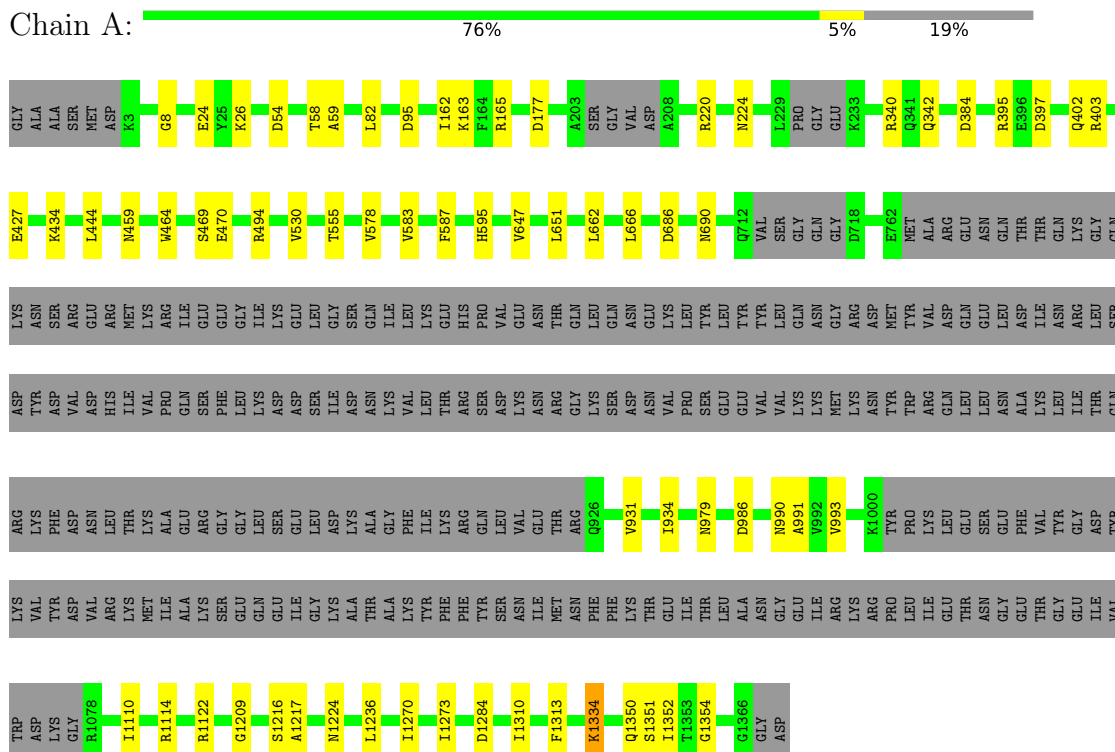
- Molecule 5 is a DNA chain called 3' product of target strand DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	c	23	Total 463	C 222	N 75	O 143	P 23	0	0

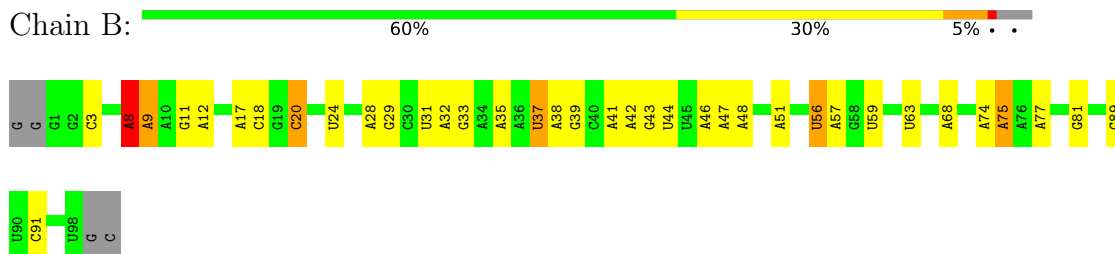
3 Residue-property plots

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-associated endonuclease Cas9/Csn1

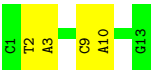


• Molecule 2: single guide RNA

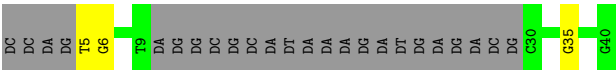


• Molecule 3: 5' product of target strand DNA

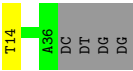
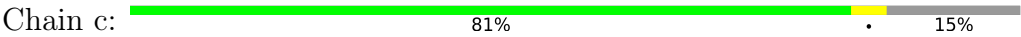




- Molecule 4: non-target strand DNA



- Molecule 5: 3' product of target strand DNA



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	30513	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 ($e^-/\text{\AA}^2$)	73	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k 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	A	0.32	0/8242	0.48	1/11200 (0.0%)
2	B	0.61	0/2358	1.03	8/3675 (0.2%)
3	C	0.79	0/292	0.85	0/448
4	D	0.73	0/369	0.97	0/566
5	c	0.92	1/515 (0.2%)	1.02	0/789
All	All	0.46	1/11776 (0.0%)	0.70	9/16678 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	c	14	DT	OP3-P	-10.87	1.48	1.61

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1334	LYS	CA-CB-CG	6.34	127.36	113.40
2	B	3	C	N1-C2-O2	5.92	122.45	118.90
2	B	18	C	C2-N1-C1'	5.57	124.92	118.80
2	B	8	A	P-O3'-C3'	5.32	126.08	119.70
2	B	18	C	N1-C2-O2	5.20	122.02	118.90

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	8098	0	7438	40	0
2	B	2104	0	1056	11	0
3	C	261	0	147	2	0
4	D	331	0	183	2	0
5	c	463	0	262	0	0
All	All	11257	0	9086	49	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 3.

The worst 5 of 49 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:530:VAL:O	1:A:578:VAL:HA	1.92	0.70
1:A:464:TRP:HB3	1:A:494:ARG:HH11	1.63	0.64
1:A:8:GLY:HA3	1:A:991:ALA:HB2	1.84	0.59
1:A:1351:SER:OG	1:A:1352:ILE:N	2.36	0.58
1:A:54:ASP:N	1:A:54:ASP:OD1	2.38	0.56

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	1100/1372 (80%)	1032 (94%)	68 (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	698/1228 (57%)	697 (100%)	1 (0%)	93 97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1334	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1262	HIS
1	A	990	ASN
1	A	295	ASN
1	A	224	ASN
1	A	402	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	97/102 (95%)	22 (22%)	3 (3%)

5 of 22 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	9	A
2	B	17	A
2	B	24	U
2	B	28	A
2	B	29	G

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	B	8	A
2	B	38	A
2	B	42	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 monosaccharides 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.

6 Map visualisation

This section contains visualisations of the EMDB entry EMD-0583. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal surface views

This section was not generated.

6.5 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis ⓘ

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution ⓘ

This section was not generated.

7.2 Volume estimate versus contour level ⓘ

This section was not generated.

7.3 Rotationally averaged power spectrum ⓘ

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

This section was not generated.