



## wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Jan 4, 2018 – 08:22 AM EST

PDB ID : 5XLO  
EMDB ID: : EMD-6729  
Title : Anti-CRISPR proteins AcrF1/2 bound to Csy surveillance complex with a  
32nt spacer crRNA backbone region  
Authors : Peng, R.; Shi, Y.; Gao, G.F.  
Deposited on : 2017-05-11  
Resolution : 3.80 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](mailto: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

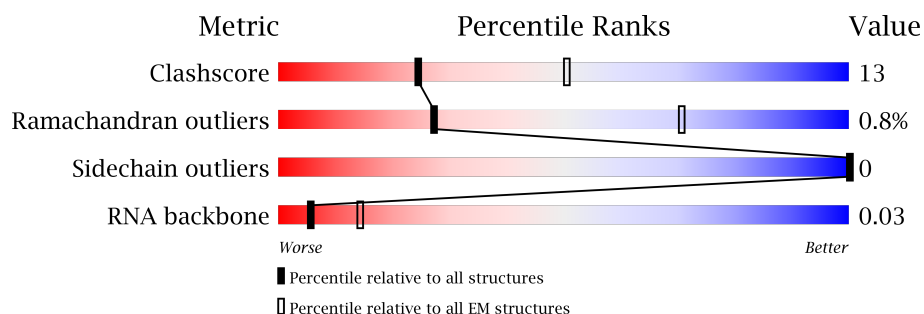
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.80 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  $\geq 3$ , 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	342	62% 21% • 16%
1	B	342	70% 24% • 5%
1	C	342	69% 25% • 5%
1	D	342	68% 26% • 5%
1	E	342	67% 27% • 5%
1	F	342	67% 27% • 5%
2	K	60	7% 25% 18% 13% 37%
3	M	78	68% 26% 6%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	N	78	 A horizontal bar chart showing the quality of the chain. The bar is divided into three segments: green (69%), yellow (26%), and orange (5%).

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 15413 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 protein Csy3.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	287	Total	C	N	O	S	0	0
			2004	1251	349	402	2		
1	B	326	Total	C	N	O	S	0	0
			2273	1414	396	461	2		
1	C	326	Total	C	N	O	S	0	0
			2273	1414	396	461	2		
1	D	326	Total	C	N	O	S	0	0
			2273	1414	396	461	2		
1	E	326	Total	C	N	O	S	0	0
			2273	1414	396	461	2		
1	F	326	Total	C	N	O	S	0	0
			2273	1414	396	461	2		

- Molecule 2 is a RNA chain called crRNA with 32nt spacer sequence.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	K	38	Total	C	N	O	P	0	0
			809	361	143	267	38		

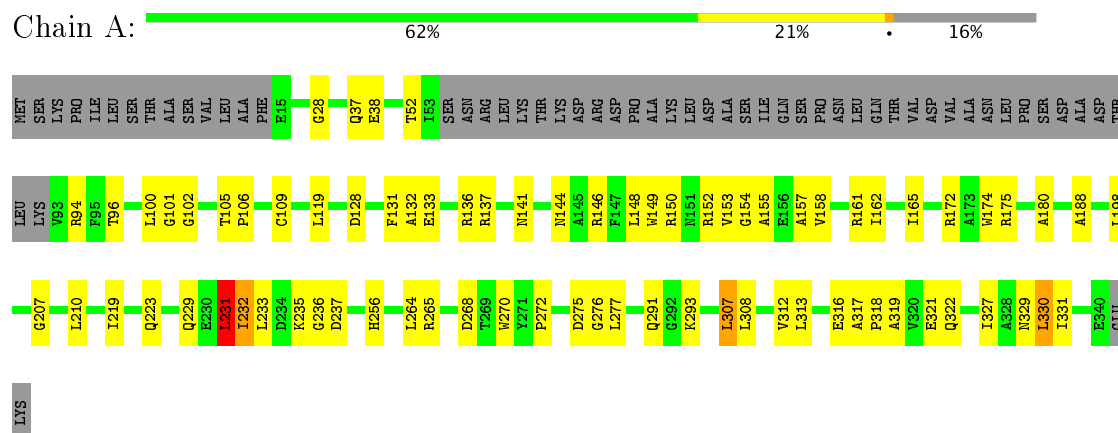
- Molecule 3 is a protein called Uncharacterized protein AcrF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	N	78	Total	C	N	O	S	0	0
			615	385	108	119	3		
3	M	78	Total	C	N	O	S	0	0
			620	388	110	119	3		

### 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.

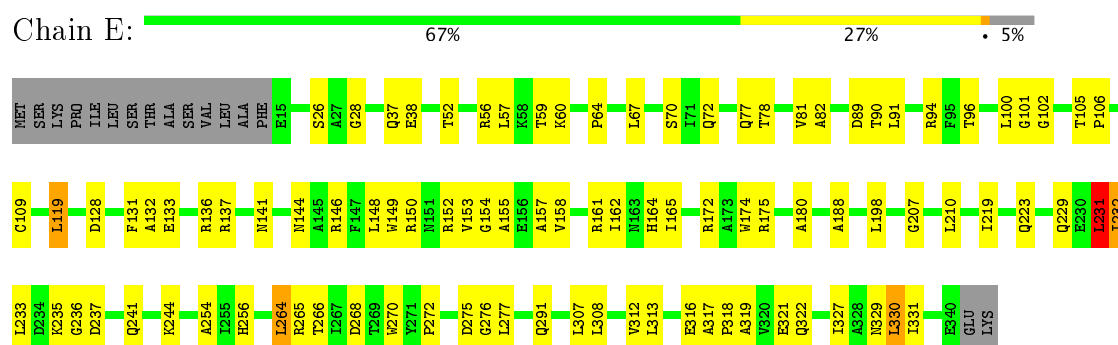
#### • Molecule 1: CRISPR-associated protein Csy3



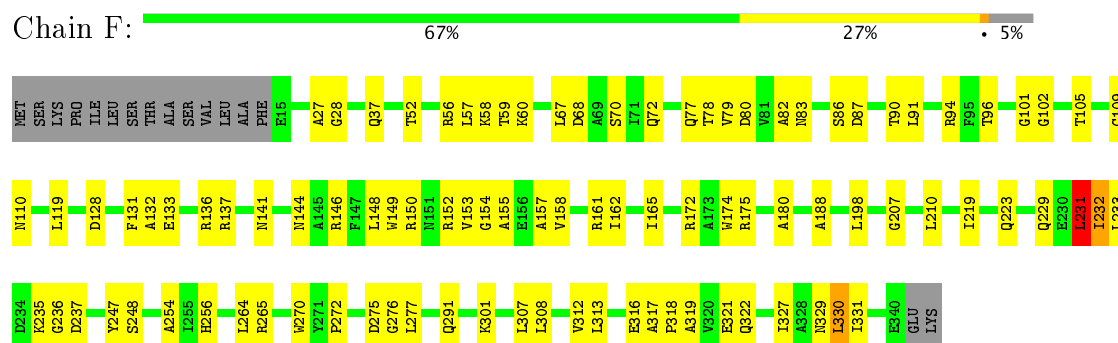
- Molecule 1: CRISPR-associated protein Csy3



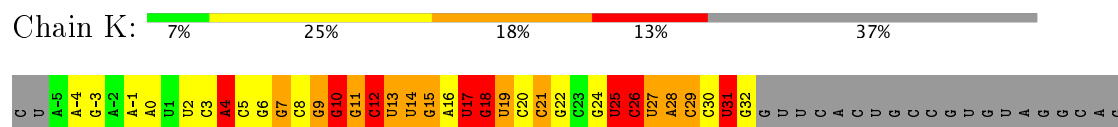
- Molecule 1: CRISPR-associated protein Csy3



- Molecule 1: CRISPR-associated protein Csy3

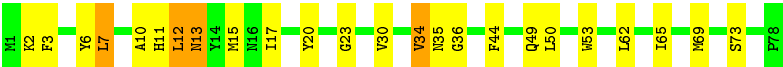


- Molecule 2: crRNA with 32nt spacer sequence

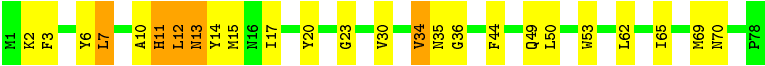


- Molecule 3: Uncharacterized protein AcrF1





● Molecule 3: Uncharacterized protein AcrF1



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	155099	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 ( $e^-/\text{\AA}^2$ )	1.0	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

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	0.45	0/2031	0.72	3/2774 (0.1%)
1	B	0.43	0/2304	0.71	3/3152 (0.1%)
1	C	0.44	0/2304	0.72	4/3152 (0.1%)
1	D	0.44	0/2304	0.73	4/3152 (0.1%)
1	E	0.44	0/2304	0.71	4/3152 (0.1%)
1	F	0.44	0/2304	0.72	3/3152 (0.1%)
2	K	0.95	0/903	1.80	35/1405 (2.5%)
3	M	0.46	0/631	0.71	1/846 (0.1%)
3	N	0.46	0/625	0.71	1/838 (0.1%)
All	All	0.49	0/15710	0.83	58/21623 (0.3%)

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	3
1	B	0	3
1	C	0	3
1	D	0	3
1	E	0	3
1	F	0	3
3	M	0	5
3	N	0	5
All	All	0	28

There are no bond length outliers.

The worst 5 of 58 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	K	29	C	N1-C2-O2	13.41	126.95	118.90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	K	29	C	C2-N1-C1'	13.07	133.18	118.80
2	K	31	U	N1-C2-O2	11.75	131.03	122.80
2	K	31	U	C2-N1-C1'	11.41	131.39	117.70
2	K	29	C	N3-C2-O2	-11.17	114.08	121.90

There are no chirality outliers.

5 of 28 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	154	GLY	Peptide
1	A	231	LEU	Peptide
1	A	264	LEU	Peptide
1	B	154	GLY	Peptide
1	B	231	LEU	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	A	2004	0	1837	54	0
1	B	2273	0	2086	62	0
1	C	2273	0	2086	64	0
1	D	2273	0	2086	69	0
1	E	2273	0	2086	72	0
1	F	2273	0	2086	73	0
2	K	809	0	410	16	0
3	M	620	0	609	14	0
3	N	615	0	604	11	0
All	All	15413	0	13890	392	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 13.

The worst 5 of 392 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:B:94:ARG:HH12	1:B:229:GLN:HB2	1.29	0.98
1:C:94:ARG:HH12	1:C:229:GLN:HB2	1.29	0.98
1:D:94:ARG:HH12	1:D:229:GLN:HB2	1.29	0.98
1:A:94:ARG:HH12	1:A:229:GLN:HB2	1.29	0.97
1:E:77:GLN:O	1:E:81:VAL:HB	1.64	0.96

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	283/342 (83%)	234 (83%)	47 (17%)	2 (1%)	25	68
1	B	324/342 (95%)	267 (82%)	55 (17%)	2 (1%)	28	70
1	C	324/342 (95%)	266 (82%)	56 (17%)	2 (1%)	28	70
1	D	324/342 (95%)	268 (83%)	54 (17%)	2 (1%)	28	70
1	E	324/342 (95%)	268 (83%)	54 (17%)	2 (1%)	28	70
1	F	324/342 (95%)	267 (82%)	55 (17%)	2 (1%)	28	70
3	M	76/78 (97%)	57 (75%)	17 (22%)	2 (3%)	6	43
3	N	76/78 (97%)	57 (75%)	17 (22%)	2 (3%)	6	43
All	All	2055/2208 (93%)	1684 (82%)	355 (17%)	16 (1%)	27	65

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	232	ILE
1	B	232	ILE
1	C	232	ILE
1	D	232	ILE
1	E	232	ILE

### 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	181/273 (66%)	181 (100%)	0	100	100
1	B	209/273 (77%)	209 (100%)	0	100	100
1	C	209/273 (77%)	209 (100%)	0	100	100
1	D	209/273 (77%)	209 (100%)	0	100	100
1	E	209/273 (77%)	209 (100%)	0	100	100
1	F	209/273 (77%)	209 (100%)	0	100	100
3	M	66/66 (100%)	66 (100%)	0	100	100
3	N	65/66 (98%)	65 (100%)	0	100	100
All	All	1357/1770 (77%)	1357 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	141	ASN
1	D	300	GLN
1	F	83	ASN
1	D	83	ASN
1	F	141	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	K	37/60 (61%)	32 (86%)	6 (16%)

5 of 32 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	K	-4	A
2	K	-3	G

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	K	-1	A
2	K	0	A
2	K	2	U

5 of 6 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	K	10	G
2	K	28	A
2	K	21	C
2	K	9	G
2	K	25	U

#### 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.