



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Oct 10, 2017 – 06:12 PM EDT

PDB ID : 6B47
EMDB ID: : EMD-7051
Title : Cryo-EM structure of Type I-F CRISPR crRNA-guided Csy surveillance complex with bound anti-CRISPR protein AcrF2
Authors : Guo, T.W.; Bartesaghi, A.; Yang, H.; Falconieri, V.; Rao, P.; Merk, A.; Fox, T.; Earl, L.; Patel, D.J.; Subramaniam, S.
Deposited on : unknown
Resolution : 3.20 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-20030345

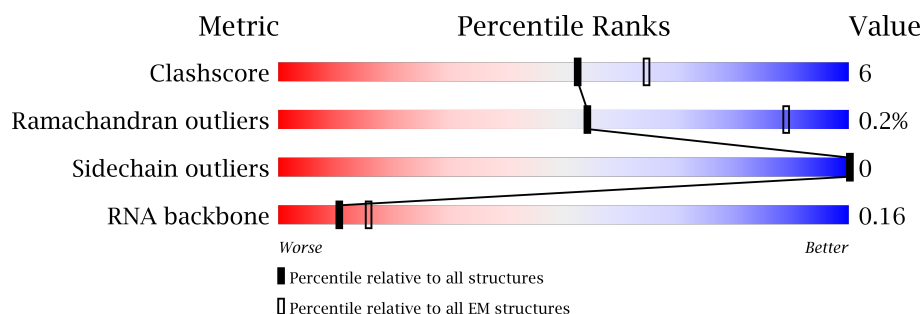
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.20 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	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	A	436	92% 5% .
2	B	329	72% 21% 7%
3	C	344	69% 16% 15%
3	D	344	81% 15% .
3	E	344	83% 14% .
3	F	344	83% 15% .
3	G	344	79% 17% .
3	H	344	85% 12% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	K	92	<div><div></div><div>68%</div><div>26%</div><div>5%</div></div>
5	L	189	<div><div></div><div>94%</div><div>6%</div></div>
6	M	60	<div><div></div><div>25%</div><div>52%</div><div>22%</div><div>.</div></div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 22213 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 Csy1.

Mol	Chain	Residues	Atoms				AltConf	Trace
1	A	424	Total	C	N	O	0	0
			2079	1145	473	461		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q02ML9
A	0	SER	-	expression tag	UNP Q02ML9

- Molecule 2 is a protein called CRISPR-associated protein Csy2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	305	Total	C	N	O	S	0	0
			2374	1504	440	425	5		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	MET	-	initiating methionine	UNP Q02MM0
B	0	ALA	-	expression tag	UNP Q02MM0

- Molecule 3 is a protein called CRISPR-associated protein Csy3.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	293	Total	C	N	O	S	0	0
			2272	1430	409	431	2		
3	D	333	Total	C	N	O	S	0	0
			2554	1603	466	483	2		
3	E	334	Total	C	N	O	S	0	0
			2561	1611	466	482	2		
3	F	334	Total	C	N	O	S	0	0
			2561	1611	466	482	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
3	G	333	Total	C	N	O	S	0	0
			2557	1608	466	481	2		
3	H	333	Total	C	N	O	S	0	0
			2557	1608	466	481	2		

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	MET	-	initiating methionine	UNP Q02MM1
C	0	ALA	-	expression tag	UNP Q02MM1
D	-1	MET	-	initiating methionine	UNP Q02MM1
D	0	ALA	-	expression tag	UNP Q02MM1
E	-1	MET	-	initiating methionine	UNP Q02MM1
E	0	ALA	-	expression tag	UNP Q02MM1
F	-1	MET	-	initiating methionine	UNP Q02MM1
F	0	ALA	-	expression tag	UNP Q02MM1
G	-1	MET	-	initiating methionine	UNP Q02MM1
G	0	ALA	-	expression tag	UNP Q02MM1
H	-1	MET	-	initiating methionine	UNP Q02MM1
H	0	ALA	-	expression tag	UNP Q02MM1

- Molecule 4 is a protein called Anti-CRISPR protein AcrF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	K	87	Total	C	N	O	S	0	0
			668	410	107	148	3		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	-1	GLY	-	expression tag	UNP Q6TM72
K	0	SER	-	expression tag	UNP Q6TM72

- Molecule 5 is a protein called CRISPR-associated endonuclease Cas6/Csy4.

Mol	Chain	Residues	Atoms				AltConf	Trace
5	L	189	Total	C	N	O	0	0
			758	380	189	189		

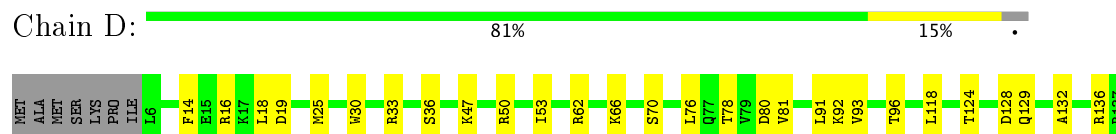
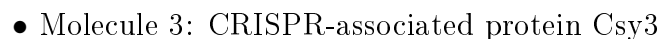
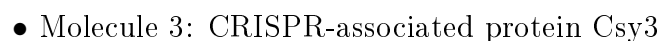
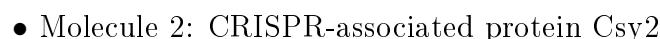
There are 2 discrepancies between the modelled and reference sequences:

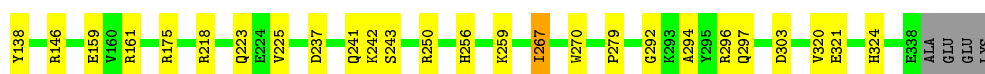
Chain	Residue	Modelled	Actual	Comment	Reference
L	-1	MET	-	initiating methionine	UNP Q02MM2
L	0	ALA	-	expression tag	UNP Q02MM2

- Molecule 6 is a RNA chain called Pseudomonas aeruginosa strain SMC4485 CRISPR repeat sequence.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	M	60	Total	C	N	O	P	0	0
			1272	569	223	421	59		

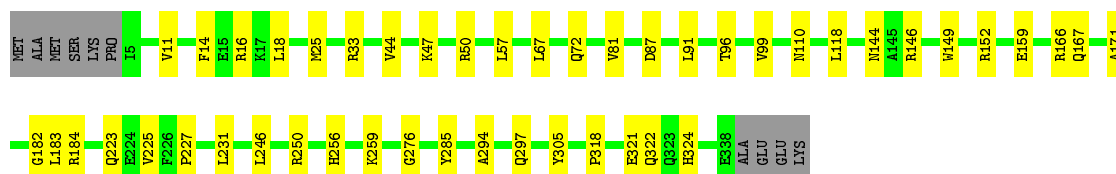
- Molecule 1: CRISPR-associated protein Csy1





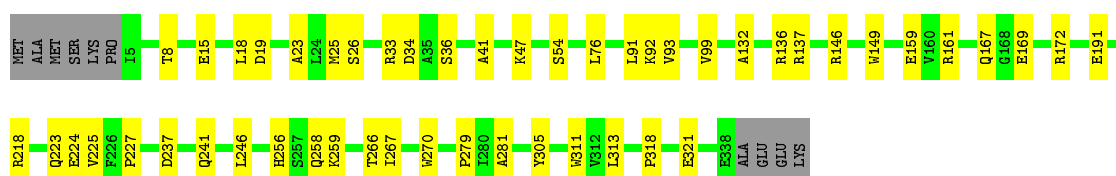
• Molecule 3: CRISPR-associated protein Csy3

Chain E: 83% 14% .



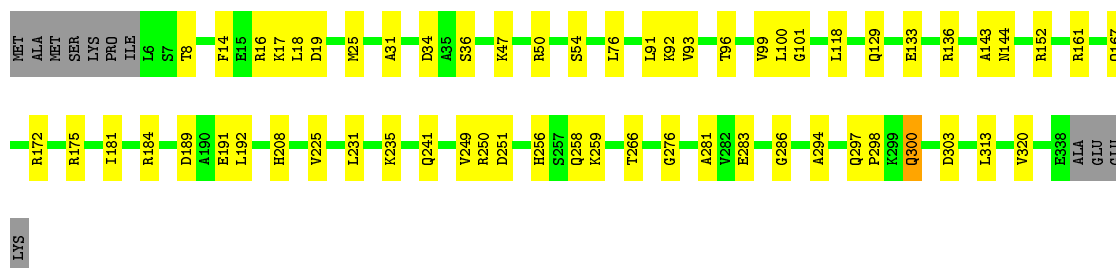
• Molecule 3: CRISPR-associated protein Csy3

Chain F: 83% 15% .



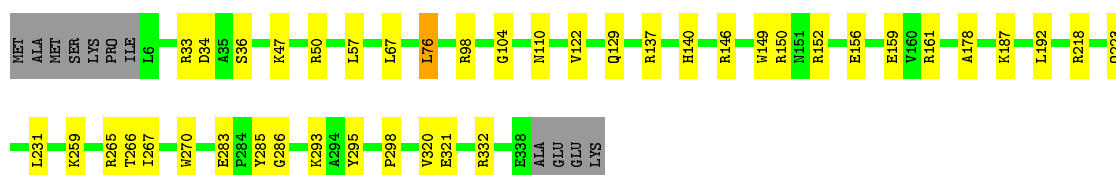
• Molecule 3: CRISPR-associated protein Csy3

Chain G: 79% 17% .



• Molecule 3: CRISPR-associated protein Csy3

Chain H: 85% 12% .



• Molecule 4: Anti-CRISPR protein AcrF2

Chain K: 68% 26% 5%

- Molecule 5: CRISPR-associated endonuclease Cas6/Csy4



- Molecule 6: *Pseudomonas aeruginosa* strain SMC4485 CRISPR repeat sequence



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	219718	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 ($e^-/\text{\AA}^2$)	84	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	0.29	0/2104	0.53	0/2715
2	B	0.38	1/2431 (0.0%)	0.64	0/3310
3	C	0.38	0/2315	0.60	0/3143
3	D	0.45	0/2601	0.61	2/3532 (0.1%)
3	E	0.50	0/2608	0.62	0/3540
3	F	0.50	0/2608	0.62	0/3540
3	G	0.50	0/2604	0.63	0/3533
3	H	0.47	0/2604	0.62	1/3533 (0.0%)
4	K	0.29	0/677	0.57	0/918
5	L	0.24	0/757	0.51	0/946
6	M	0.87	0/1420	1.15	8/2212 (0.4%)
All	All	0.47	1/22729 (0.0%)	0.66	11/30922 (0.0%)

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
3	D	0	1
3	E	0	1
3	G	0	2
3	H	0	1
All	All	0	5

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	152	PHE	C-N	5.80	1.45	1.34

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	M	1	C	O4'-C1'-N1	10.52	116.62	108.20
6	M	21	U	C2-N1-C1'	7.06	126.18	117.70
3	H	76	LEU	CB-CG-CD2	-6.85	99.35	111.00
6	M	1	C	N1-C2-O2	5.86	122.41	118.90
6	M	43	U	C2-N1-C1'	5.72	124.56	117.70

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	267	ILE	Peptide
3	E	285	TYR	Peptide
3	G	300	GLN	Peptide
3	G	31	ALA	Peptide
3	H	285	TYR	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	2079	0	1094	16	0
2	B	2374	0	2345	47	0
3	C	2272	0	2232	34	0
3	D	2554	0	2522	32	0
3	E	2561	0	2542	31	0
3	F	2561	0	2542	35	0
3	G	2557	0	2542	36	0
3	H	2557	0	2542	32	0
4	K	668	0	615	21	0
5	L	758	0	209	16	0
6	M	1272	0	644	20	0
All	All	22213	0	19829	273	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 273 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)
4:K:1:MET:CE	4:K:4:GLN:NE2	1.82	1.42
5:L:0:ALA:CB	5:L:170:GLU:H	1.31	1.39
4:K:1:MET:HE3	4:K:4:GLN:NE2	0.99	1.30
4:K:1:MET:HE3	4:K:4:GLN:CD	1.55	1.26
4:K:1:MET:CE	4:K:4:GLN:HE22	1.44	1.24

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	
1	A	422/436 (97%)	357 (85%)	64 (15%)	1 (0%)	51	86
2	B	301/329 (92%)	273 (91%)	28 (9%)	0	100	100
3	C	287/344 (83%)	255 (89%)	31 (11%)	1 (0%)	44	81
3	D	331/344 (96%)	299 (90%)	31 (9%)	1 (0%)	44	81
3	E	332/344 (96%)	308 (93%)	23 (7%)	1 (0%)	44	81
3	F	332/344 (96%)	299 (90%)	33 (10%)	0	100	100
3	G	331/344 (96%)	306 (92%)	24 (7%)	1 (0%)	44	81
3	H	331/344 (96%)	299 (90%)	32 (10%)	0	100	100
4	K	85/92 (92%)	72 (85%)	13 (15%)	0	100	100
5	L	187/189 (99%)	153 (82%)	34 (18%)	0	100	100
All	All	2939/3110 (94%)	2621 (89%)	313 (11%)	5 (0%)	54	86

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	250	ARG
3	G	250	ARG
3	C	250	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	E	250	ARG
1	A	34	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	78/366 (21%)	78 (100%)	0	100	100
2	B	245/271 (90%)	245 (100%)	0	100	100
3	C	230/274 (84%)	230 (100%)	0	100	100
3	D	258/274 (94%)	258 (100%)	0	100	100
3	E	259/274 (94%)	259 (100%)	0	100	100
3	F	259/274 (94%)	259 (100%)	0	100	100
3	G	259/274 (94%)	259 (100%)	0	100	100
3	H	259/274 (94%)	259 (100%)	0	100	100
4	K	68/72 (94%)	68 (100%)	0	100	100
All	All	1915/2353 (81%)	1915 (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 25 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	164	HIS
3	E	164	HIS
4	K	25	GLN
3	D	223	GLN
3	E	329	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
6	M	59/60 (98%)	36 (61%)	1 (1%)

5 of 36 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
6	M	2	U
6	M	3	A
6	M	5	G
6	M	7	A
6	M	9	U

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
6	M	16	C

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.