



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Oct 11, 2017 – 08:39 AM EDT

PDB ID : 6B48
EMDB ID: : EMD-7052
Title : Cryo-EM structure of Type I-F CRISPR crRNA-guided Csy surveillance complex with bound anti-CRISPR protein AcrF10
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 : 3.60 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

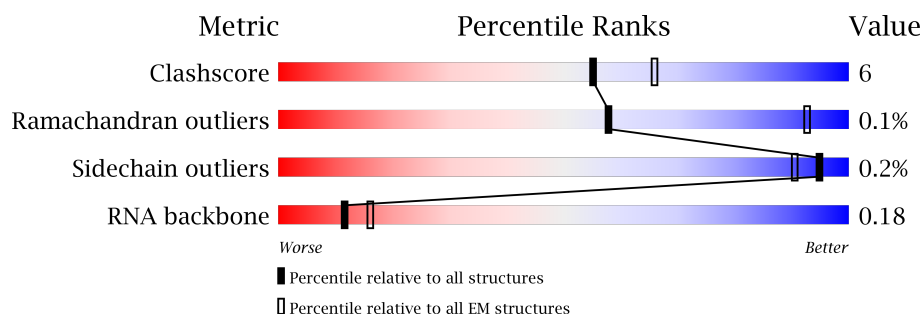
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.60 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	90% 7% .
2	B	329	72% 20% 7%
3	C	344	67% 18% 15%
3	D	344	76% 21% .
3	E	344	81% 16% .
3	F	344	84% 13% .
3	G	344	83% 14% .
3	H	344	81% 16% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	K	99	<div><div></div><div>86%</div><div>12%</div><div></div></div>
5	L	189	<div><div></div><div>98%</div><div></div></div>
6	M	60	<div><div></div><div>33%</div><div>48%</div><div>17%</div><div></div></div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 22333 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	335	Total	C	N	O	S	0	0
			2566	1614	467	483	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
			2554	1603	466	483	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 AcrF10.

Mol	Chain	Residues	Atoms						AltConf	Trace
4	K	97	Total	C	N	O	S	Se	0	0
			786	496	119	165	1	5		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	0	GLY	-	expression tag	UNP A0A073KP86
K	1	SER	-	expression tag	UNP A0A073KP86
K	2	MSE	-	expression tag	UNP A0A073KP86

- 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

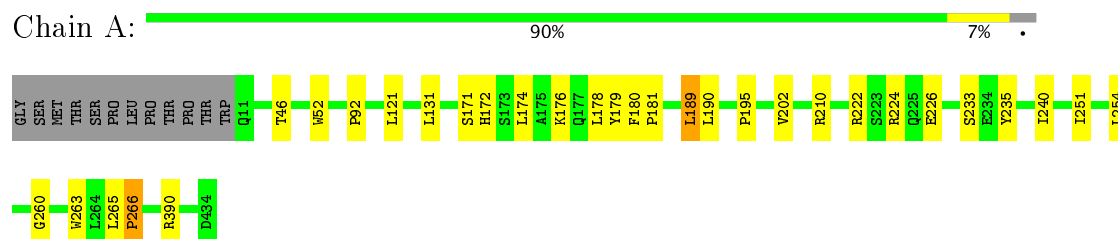
- Molecule 6 is a RNA chain called Pseudomonas aeruginosa strain SMC4485 CRISPR repeat sequence.

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

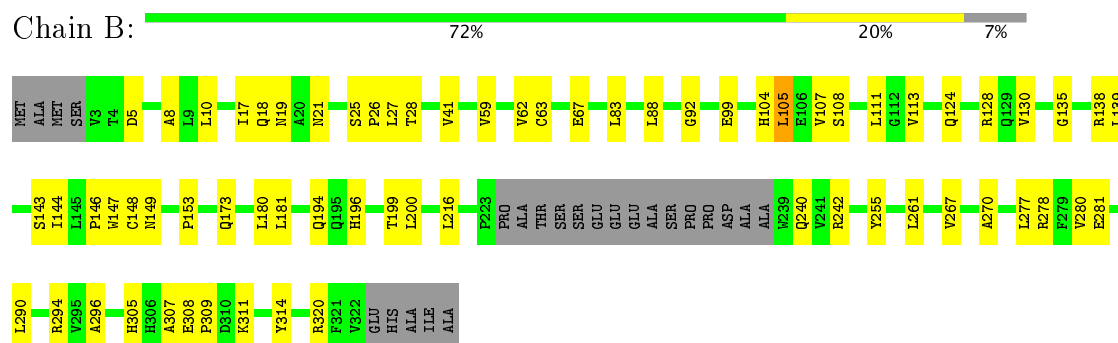
3 Residue-property plots [i](#)

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.

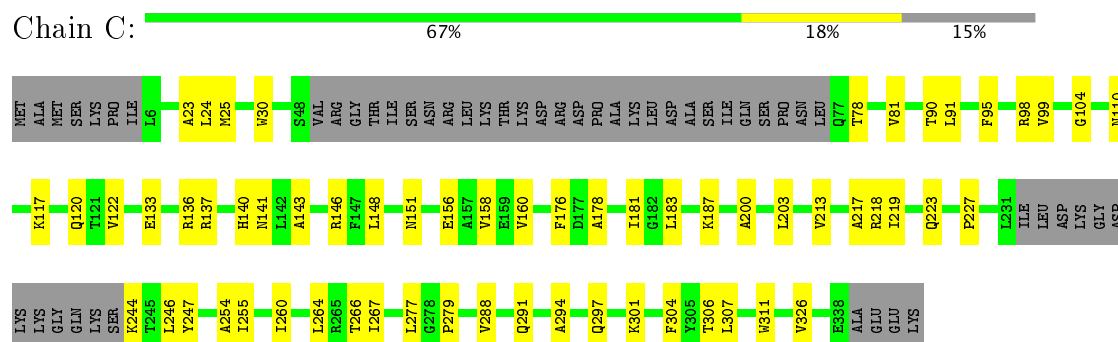
• Molecule 1: CRISPR-associated protein Csy1



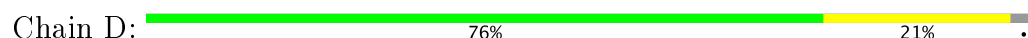
• Molecule 2: CRISPR-associated protein Csy2

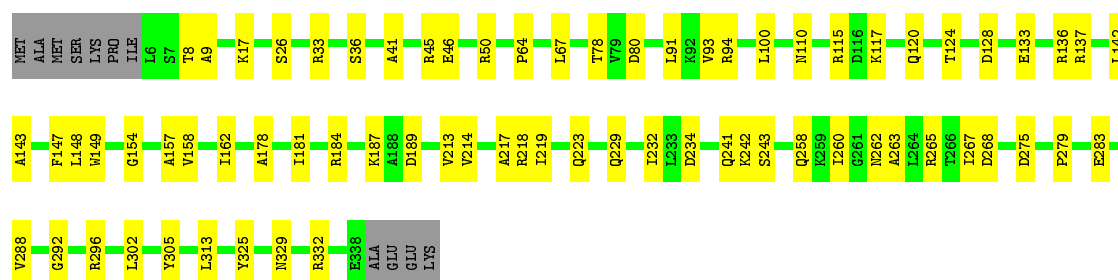


• Molecule 3: CRISPR-associated protein Csy3



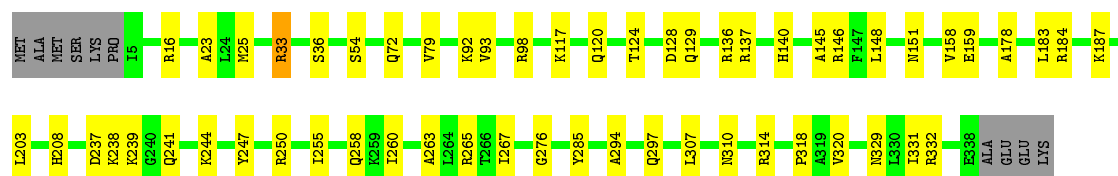
• Molecule 3: CRISPR-associated protein Csy3





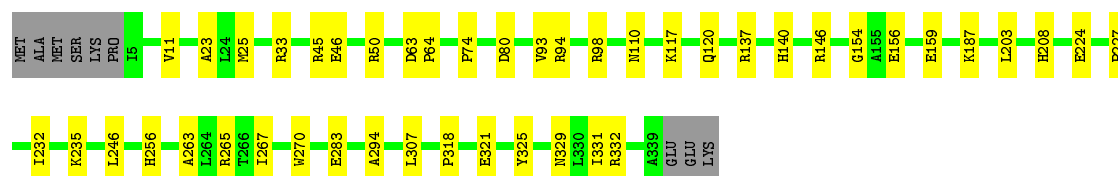
- Molecule 3: CRISPR-associated protein Csy3

Chain E: 81% 16% .



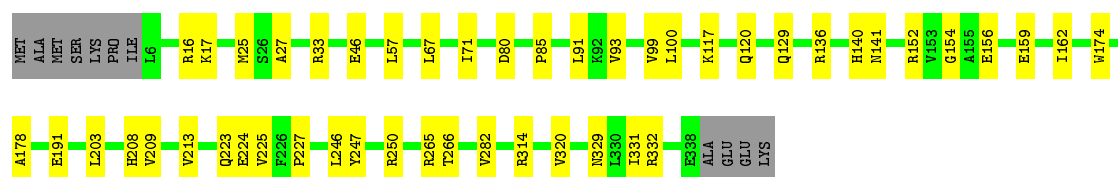
- Molecule 3: CRISPR-associated protein Csy3

Chain F: 84% 13% .



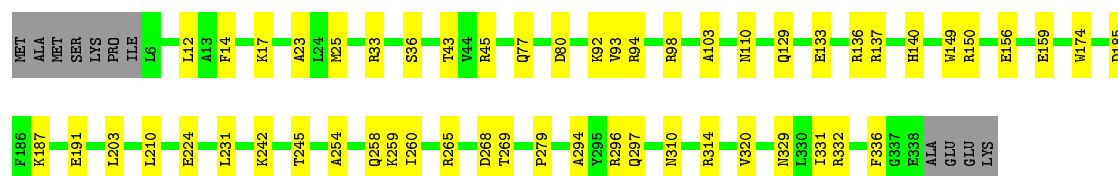
- Molecule 3: CRISPR-associated protein Csy3

Chain G: 83% 14% .




- Molecule 3: CRISPR-associated protein Csy3

Chain H: 81% 16% .



- Molecule 4: Anti-CRISPR protein AcrF10

Chain K:  86% 12% .



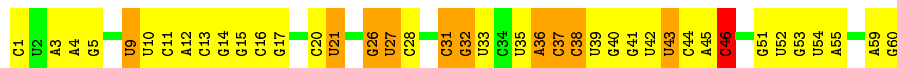
- Molecule 5: CRISPR-associated endonuclease Cas6/Csy4

Chain L:  98% .



- Molecule 6: Pseudomonas aeruginosa strain SMC4485 CRISPR repeat sequence

Chain M:  33% 48% 17% .



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	42498	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 ($e^-/\text{\AA}^2$)	40	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	0.28	0/2104	0.56	1/2715 (0.0%)
2	B	0.37	0/2431	0.68	1/3310 (0.0%)
3	C	0.26	0/2315	0.51	0/3143
3	D	0.31	0/2601	0.58	0/3532
3	E	0.37	0/2608	0.59	0/3540
3	F	0.41	0/2613	0.59	0/3547
3	G	0.41	0/2604	0.62	0/3533
3	H	0.40	0/2601	0.62	1/3532 (0.0%)
4	K	0.38	0/793	0.71	2/1066 (0.2%)
5	L	0.24	0/757	0.50	0/946
6	M	0.52	0/1420	1.04	7/2212 (0.3%)
All	All	0.37	0/22847	0.64	12/31076 (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
1	A	0	1
2	B	0	2
3	E	0	1
All	All	0	4

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	M	1	C	O4'-C1'-N1	9.54	115.83	108.20
4	K	17	ASP	CB-CG-OD1	9.49	126.84	118.30
2	B	181	LEU	CA-CB-CG	7.79	133.21	115.30
3	H	231	LEU	CA-CB-CG	7.12	131.66	115.30
6	M	46	C	N1-C2-O2	6.55	122.83	118.90

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	222	ARG	Peptide
2	B	105	LEU	Peptide
2	B	280	VAL	Peptide
3	E	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	21	0
2	B	2374	0	2345	43	0
3	C	2272	0	2232	39	0
3	D	2554	0	2522	45	0
3	E	2561	0	2542	36	0
3	F	2566	0	2547	28	0
3	G	2557	0	2542	32	0
3	H	2554	0	2522	35	0
4	K	786	0	749	6	0
5	L	758	0	209	2	0
6	M	1272	0	644	18	0
All	All	22333	0	19948	259	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 259 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)
3:G:25:MET:HG2	3:G:93:VAL:HG22	1.78	0.65
2:B:148:CYS:SG	2:B:149:ASN:N	2.70	0.64
3:D:262:ASN:ND2	6:M:32:G:OP2	2.30	0.64
3:F:232:ILE:HG13	3:F:235:LYS:HB2	1.80	0.63
2:B:25:SER:HB3	2:B:28:THR:HG22	1.82	0.60

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	422/436 (97%)	348 (82%)	72 (17%)	2 (0%)	32	73
2	B	301/329 (92%)	259 (86%)	42 (14%)	0	100	100
3	C	287/344 (83%)	260 (91%)	27 (9%)	0	100	100
3	D	331/344 (96%)	292 (88%)	39 (12%)	0	100	100
3	E	332/344 (96%)	304 (92%)	27 (8%)	1 (0%)	44	80
3	F	333/344 (97%)	297 (89%)	36 (11%)	0	100	100
3	G	331/344 (96%)	300 (91%)	31 (9%)	0	100	100
3	H	331/344 (96%)	298 (90%)	33 (10%)	0	100	100
4	K	95/99 (96%)	84 (88%)	11 (12%)	0	100	100
5	L	187/189 (99%)	149 (80%)	38 (20%)	0	100	100
All	All	2950/3117 (95%)	2591 (88%)	356 (12%)	3 (0%)	58	88

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	265	LEU
3	E	250	ARG
1	A	266	PRO

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	78/366 (21%)	78 (100%)	0	100	100
2	B	245/271 (90%)	244 (100%)	1 (0%)	93	98
3	C	230/274 (84%)	230 (100%)	0	100	100
3	D	258/274 (94%)	257 (100%)	1 (0%)	93	98
3	E	259/274 (94%)	258 (100%)	1 (0%)	93	98
3	F	259/274 (94%)	259 (100%)	0	100	100
3	G	259/274 (94%)	258 (100%)	1 (0%)	93	98
3	H	258/274 (94%)	258 (100%)	0	100	100
4	K	91/86 (106%)	91 (100%)	0	100	100
All	All	1937/2367 (82%)	1933 (100%)	4 (0%)	95	99

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	138	ARG
3	D	115	ARG
3	E	33	ARG
3	G	314	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	120	GLN
3	E	120	GLN
3	G	120	GLN
3	C	329	ASN
3	G	129	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
6	M	59/60 (98%)	34 (57%)	0

5 of 34 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
6	M	3	A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
6	M	9	U
6	M	10	U
6	M	14	G
6	M	15	G

There are no RNA pucker outliers to report.

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.