



wwPDB EM Model Validation Summary Report ⓘ

Mar 2, 2020 – 03:35 PM EST

PDB ID : 6VQW
EMDB ID : EMD-21359
Title : Type I-F CRISPR-Csy complex with its inhibitor AcrF8
Authors : Zhang, K.; Li, S.; Pintilie, G.; Zhu, Y.; Huang, Z.; Chiu, W.
Deposited on : 2020-02-06
Resolution : 3.42 Å(reported)

This is a wwPDB EM 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

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) : 2.8

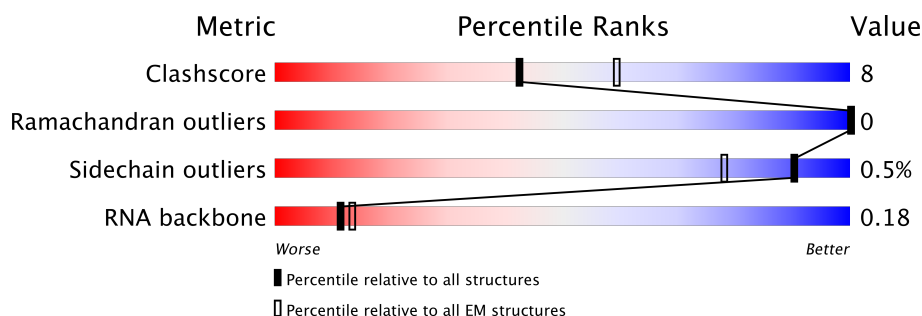
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.42 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 and their fit to the map. 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 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	C	327	74% 17% 9%
2	D	360	61% 20% 19%
2	E	360	75% 18% 8%
2	F	360	74% 18% 7%
2	G	360	70% 22% 8%
2	H	360	75% 16% 10%
2	I	360	72% 19% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	J	187	 45% 53%
4	A	92	 78% 9% 13%
5	B	434	 50% 6% 44%
6	K	60	 23% 25% 18% 33%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 20479 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 Type I-F CRISPR-associated protein Csy2.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	C	297	Total	C	N	O	S	0	0
			2309	1462	428	415	4		

- Molecule 2 is a protein called CRISPR-associated protein Csy3.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	D	293	Total	C	N	O	S	0	0
			2272	1430	409	431	2		
2	E	333	Total	C	N	O	S	0	0
			2575	1614	468	491	2		
2	F	334	Total	C	N	O	S	0	0
			2587	1623	470	492	2		
2	G	331	Total	C	N	O	S	0	0
			2559	1607	465	485	2		
2	H	325	Total	C	N	O	S	0	0
			2516	1581	457	476	2		
2	I	326	Total	C	N	O	S	0	0
			2523	1582	460	479	2		

There are 114 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1	MET	-	expression tag	UNP A0A444M080
D	2	LYS	-	expression tag	UNP A0A444M080
D	3	SER	-	expression tag	UNP A0A444M080
D	4	SER	-	expression tag	UNP A0A444M080
D	5	HIS	-	expression tag	UNP A0A444M080
D	6	HIS	-	expression tag	UNP A0A444M080
D	7	HIS	-	expression tag	UNP A0A444M080
D	8	HIS	-	expression tag	UNP A0A444M080
D	9	HIS	-	expression tag	UNP A0A444M080
D	10	HIS	-	expression tag	UNP A0A444M080
D	11	GLU	-	expression tag	UNP A0A444M080

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	12	ASN	-	expression tag	UNP A0A444M080
D	13	LEU	-	expression tag	UNP A0A444M080
D	14	TYR	-	expression tag	UNP A0A444M080
D	15	PHE	-	expression tag	UNP A0A444M080
D	16	GLN	-	expression tag	UNP A0A444M080
D	17	SER	-	expression tag	UNP A0A444M080
D	18	ASN	-	expression tag	UNP A0A444M080
D	19	ALA	-	expression tag	UNP A0A444M080
E	1	MET	-	expression tag	UNP A0A444M080
E	2	LYS	-	expression tag	UNP A0A444M080
E	3	SER	-	expression tag	UNP A0A444M080
E	4	SER	-	expression tag	UNP A0A444M080
E	5	HIS	-	expression tag	UNP A0A444M080
E	6	HIS	-	expression tag	UNP A0A444M080
E	7	HIS	-	expression tag	UNP A0A444M080
E	8	HIS	-	expression tag	UNP A0A444M080
E	9	HIS	-	expression tag	UNP A0A444M080
E	10	HIS	-	expression tag	UNP A0A444M080
E	11	GLU	-	expression tag	UNP A0A444M080
E	12	ASN	-	expression tag	UNP A0A444M080
E	13	LEU	-	expression tag	UNP A0A444M080
E	14	TYR	-	expression tag	UNP A0A444M080
E	15	PHE	-	expression tag	UNP A0A444M080
E	16	GLN	-	expression tag	UNP A0A444M080
E	17	SER	-	expression tag	UNP A0A444M080
E	18	ASN	-	expression tag	UNP A0A444M080
E	19	ALA	-	expression tag	UNP A0A444M080
F	1	MET	-	expression tag	UNP A0A444M080
F	2	LYS	-	expression tag	UNP A0A444M080
F	3	SER	-	expression tag	UNP A0A444M080
F	4	SER	-	expression tag	UNP A0A444M080
F	5	HIS	-	expression tag	UNP A0A444M080
F	6	HIS	-	expression tag	UNP A0A444M080
F	7	HIS	-	expression tag	UNP A0A444M080
F	8	HIS	-	expression tag	UNP A0A444M080
F	9	HIS	-	expression tag	UNP A0A444M080
F	10	HIS	-	expression tag	UNP A0A444M080
F	11	GLU	-	expression tag	UNP A0A444M080
F	12	ASN	-	expression tag	UNP A0A444M080
F	13	LEU	-	expression tag	UNP A0A444M080
F	14	TYR	-	expression tag	UNP A0A444M080
F	15	PHE	-	expression tag	UNP A0A444M080

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	16	GLN	-	expression tag	UNP A0A444M080
F	17	SER	-	expression tag	UNP A0A444M080
F	18	ASN	-	expression tag	UNP A0A444M080
F	19	ALA	-	expression tag	UNP A0A444M080
G	1	MET	-	expression tag	UNP A0A444M080
G	2	LYS	-	expression tag	UNP A0A444M080
G	3	SER	-	expression tag	UNP A0A444M080
G	4	SER	-	expression tag	UNP A0A444M080
G	5	HIS	-	expression tag	UNP A0A444M080
G	6	HIS	-	expression tag	UNP A0A444M080
G	7	HIS	-	expression tag	UNP A0A444M080
G	8	HIS	-	expression tag	UNP A0A444M080
G	9	HIS	-	expression tag	UNP A0A444M080
G	10	HIS	-	expression tag	UNP A0A444M080
G	11	GLU	-	expression tag	UNP A0A444M080
G	12	ASN	-	expression tag	UNP A0A444M080
G	13	LEU	-	expression tag	UNP A0A444M080
G	14	TYR	-	expression tag	UNP A0A444M080
G	15	PHE	-	expression tag	UNP A0A444M080
G	16	GLN	-	expression tag	UNP A0A444M080
G	17	SER	-	expression tag	UNP A0A444M080
G	18	ASN	-	expression tag	UNP A0A444M080
G	19	ALA	-	expression tag	UNP A0A444M080
H	1	MET	-	expression tag	UNP A0A444M080
H	2	LYS	-	expression tag	UNP A0A444M080
H	3	SER	-	expression tag	UNP A0A444M080
H	4	SER	-	expression tag	UNP A0A444M080
H	5	HIS	-	expression tag	UNP A0A444M080
H	6	HIS	-	expression tag	UNP A0A444M080
H	7	HIS	-	expression tag	UNP A0A444M080
H	8	HIS	-	expression tag	UNP A0A444M080
H	9	HIS	-	expression tag	UNP A0A444M080
H	10	HIS	-	expression tag	UNP A0A444M080
H	11	GLU	-	expression tag	UNP A0A444M080
H	12	ASN	-	expression tag	UNP A0A444M080
H	13	LEU	-	expression tag	UNP A0A444M080
H	14	TYR	-	expression tag	UNP A0A444M080
H	15	PHE	-	expression tag	UNP A0A444M080
H	16	GLN	-	expression tag	UNP A0A444M080
H	17	SER	-	expression tag	UNP A0A444M080
H	18	ASN	-	expression tag	UNP A0A444M080
H	19	ALA	-	expression tag	UNP A0A444M080

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
I	1	MET	-	expression tag	UNP A0A444M080
I	2	LYS	-	expression tag	UNP A0A444M080
I	3	SER	-	expression tag	UNP A0A444M080
I	4	SER	-	expression tag	UNP A0A444M080
I	5	HIS	-	expression tag	UNP A0A444M080
I	6	HIS	-	expression tag	UNP A0A444M080
I	7	HIS	-	expression tag	UNP A0A444M080
I	8	HIS	-	expression tag	UNP A0A444M080
I	9	HIS	-	expression tag	UNP A0A444M080
I	10	HIS	-	expression tag	UNP A0A444M080
I	11	GLU	-	expression tag	UNP A0A444M080
I	12	ASN	-	expression tag	UNP A0A444M080
I	13	LEU	-	expression tag	UNP A0A444M080
I	14	TYR	-	expression tag	UNP A0A444M080
I	15	PHE	-	expression tag	UNP A0A444M080
I	16	GLN	-	expression tag	UNP A0A444M080
I	17	SER	-	expression tag	UNP A0A444M080
I	18	ASN	-	expression tag	UNP A0A444M080
I	19	ALA	-	expression tag	UNP A0A444M080

- Molecule 3 is a protein called CRISPR-associated endonuclease Cas6/Csy4.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	J	87	Total	C	N	O	0	0
			350	175	87	88		

- Molecule 4 is a protein called AcrF8.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	A	80	Total	C	N	O	S	0	0
			552	343	92	116	1		

- Molecule 5 is a protein called CRISPR-associated protein Csy1.

Mol	Chain	Residues	Atoms				AltConf	Trace
5	B	242	Total	C	N	O	0	0
			1391	813	295	283		

- Molecule 6 is a RNA chain called CrRNA (40-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	K	40	Total	C	N	O	P	0	0
			845	379	148	279	39		

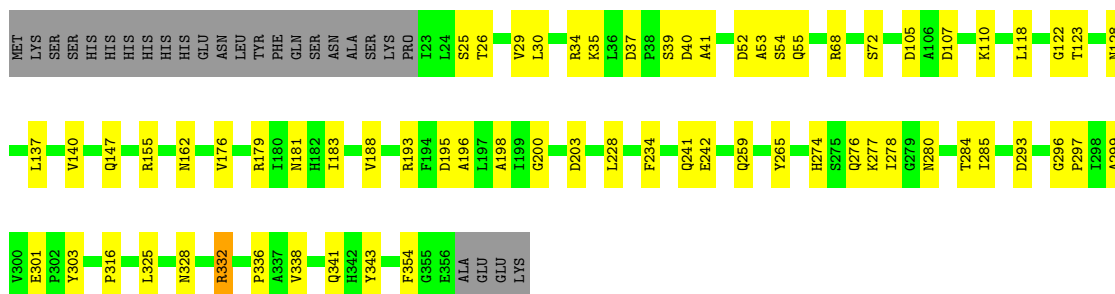
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	53	A	G	conflict	GB 313291946



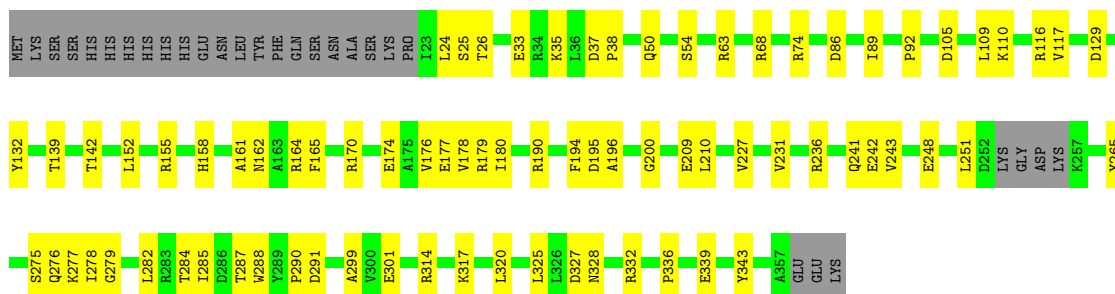
• Molecule 2: CRISPR-associated protein Csy3

Chain F: 74% 18% 7%



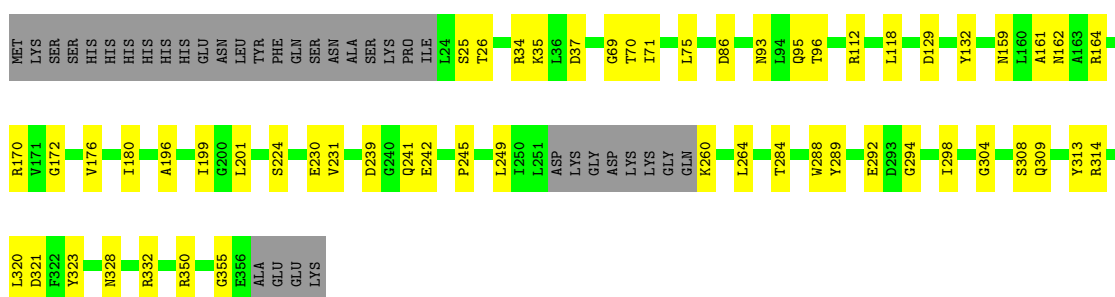
• Molecule 2: CRISPR-associated protein Csy3

Chain G: 70% 22% 8%



• Molecule 2: CRISPR-associated protein Csy3

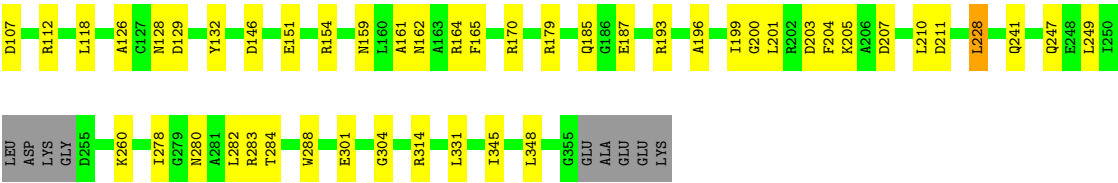
Chain H: 75% 16% 10%



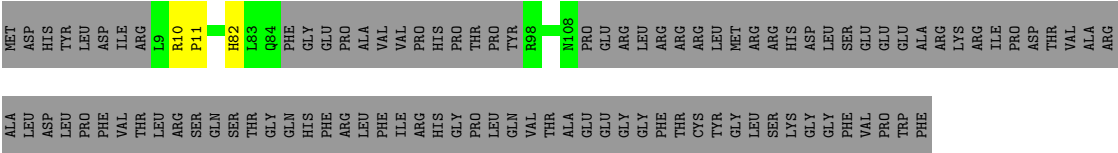
• Molecule 2: CRISPR-associated protein Csy3

Chain I: 72% 19% 9%

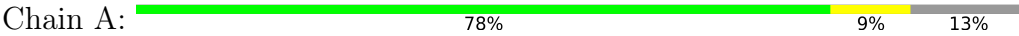




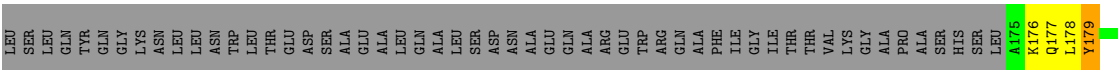
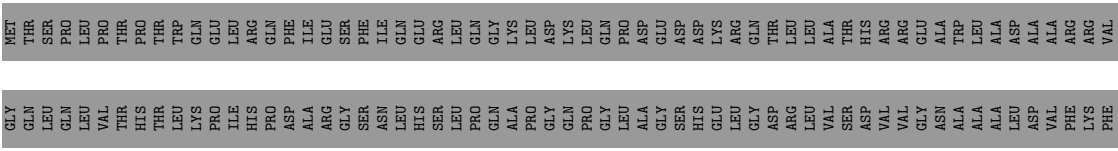
• Molecule 3: CRISPR-associated endonuclease Cas6/Csy4



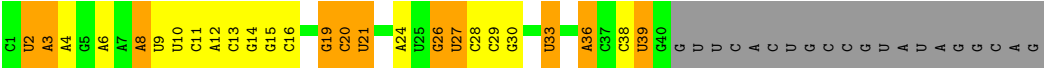
• Molecule 4: AcrF8



• Molecule 5: CRISPR-associated protein Csy1



• Molecule 6: CrRNA (40-MER)



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	91080	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 ($e^-/\text{\AA}^2$)	7.6	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 >5	RMSZ	# Z >5
1	C	0.41	0/2363	0.62	2/3216 (0.1%)
2	D	0.42	0/2315	0.53	0/3143
2	E	0.50	0/2622	0.57	0/3557
2	F	0.54	0/2634	0.57	0/3572
2	G	0.56	0/2605	0.62	2/3534 (0.1%)
2	H	0.55	0/2562	0.60	0/3477
2	I	0.47	0/2569	0.56	1/3485 (0.0%)
3	J	0.23	0/348	0.48	0/432
4	A	0.37	0/558	0.63	0/763
5	B	0.31	0/1416	0.51	0/1862
6	K	1.08	0/943	1.00	0/1468
All	All	0.52	0/20935	0.61	5/28509 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	I	228	LEU	CB-CG-CD1	-5.90	100.97	111.00
2	G	291	ASP	CB-CG-OD1	5.89	123.60	118.30
1	C	180	LEU	CA-CB-CG	5.59	128.16	115.30
2	G	251	LEU	CA-CB-CG	5.55	128.06	115.30
1	C	181	LEU	CA-CB-CG	5.26	127.39	115.30

There are no chirality outliers.

There are no planarity outliers.

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	C	2309	0	2289	36	0
2	D	2272	0	2232	48	0
2	E	2575	0	2554	42	0
2	F	2587	0	2576	46	0
2	G	2559	0	2545	52	0
2	H	2516	0	2501	37	0
2	I	2523	0	2500	47	0
3	J	350	0	95	2	0
4	A	552	0	489	4	0
5	B	1391	0	959	17	0
6	K	845	0	429	28	0
All	All	20479	0	19169	311	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 8.

The worst 5 of 311 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:C:271:ARG:HD2	6:K:4:A:H62	1.57	0.69
2:I:35:LYS:HG3	2:I:118:LEU:HB2	1.75	0.69
2:I:95:GLN:NE2	6:K:13:C:N3	2.43	0.65
2:D:286:ASP:HB3	2:D:298:ILE:HG22	1.79	0.64
5:B:255:ASN:HB3	5:B:258:ARG:HB2	1.79	0.64

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	C	293/327 (90%)	251 (86%)	42 (14%)	0	100	100
2	D	287/360 (80%)	257 (90%)	30 (10%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	E	331/360 (92%)	291 (88%)	40 (12%)	0	100	100
2	F	332/360 (92%)	293 (88%)	39 (12%)	0	100	100
2	G	327/360 (91%)	288 (88%)	39 (12%)	0	100	100
2	H	321/360 (89%)	285 (89%)	36 (11%)	0	100	100
2	I	322/360 (89%)	285 (88%)	37 (12%)	0	100	100
3	J	83/187 (44%)	66 (80%)	17 (20%)	0	100	100
4	A	78/92 (85%)	59 (76%)	19 (24%)	0	100	100
5	B	236/434 (54%)	193 (82%)	43 (18%)	0	100	100
All	All	2610/3200 (82%)	2268 (87%)	342 (13%)	0	100	100

There are no Ramachandran outliers to report.

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	C	239/270 (88%)	238 (100%)	1 (0%)	92	96
2	D	230/290 (79%)	230 (100%)	0	100	100
2	E	264/290 (91%)	262 (99%)	2 (1%)	83	92
2	F	266/290 (92%)	265 (100%)	1 (0%)	92	96
2	G	262/290 (90%)	262 (100%)	0	100	100
2	H	258/290 (89%)	258 (100%)	0	100	100
2	I	258/290 (89%)	257 (100%)	1 (0%)	92	96
3	J	1/160 (1%)	1 (100%)	0	100	100
4	A	52/77 (68%)	51 (98%)	1 (2%)	60	83
5	B	87/365 (24%)	83 (95%)	4 (5%)	29	65
All	All	1917/2612 (73%)	1907 (100%)	10 (0%)	90	95

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	I	96	THR
4	A	37	VAL
5	B	213	ASP
2	F	332	ARG
5	B	179	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	147	GLN
2	F	341	GLN
5	B	188	HIS
2	F	182	HIS
2	F	347	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
6	K	39/60 (65%)	20 (51%)	0

5 of 20 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
6	K	2	U
6	K	3	A
6	K	6	A
6	K	8	A
6	K	9	U

There are no RNA pucker outliers to report.

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.