



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

Jan 1, 2017 – 05:35 AM EST

PDB ID : 5TFY  
EMDB ID: : EMD-8405  
Title : The archaeal flagellum of Methanospirillum hungatei strain JF1.  
Authors : Poweleit, N.; Peng, G.; Gunsalus, R.P.; Zhou, Z.H.  
Deposited on : 2016-09-27  
Resolution : 3.40 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  
Mogul : unknown  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20028442

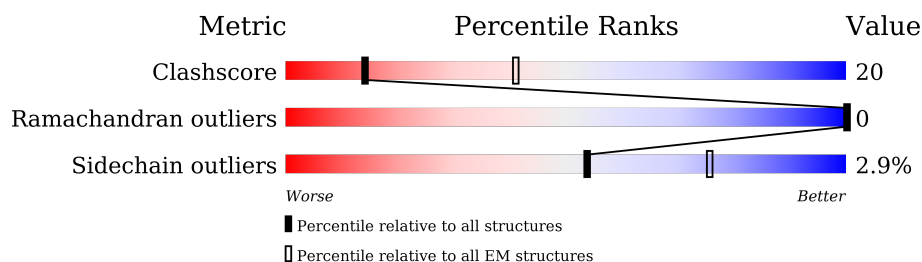
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.40 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.















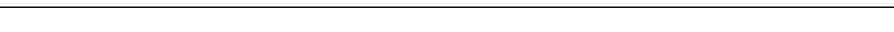



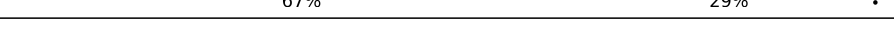
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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	164	67% 29% .
1	B	164	65% 30% .
1	C	164	66% 29% 5%
1	D	164	68% 27% 5%
1	E	164	65% 30% 5%
1	F	164	68% 27% 5%
1	G	164	66% 29% 5%
1	H	164	68% 27% 5%
1	I	164	70% 26% .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	J	164	 67%28%5%
1	K	164	 68%27%5%
1	L	164	 69%27%.
1	M	164	 70%26%5%
1	N	164	 68%28%.
1	O	164	 69%27%.
1	P	164	 70%25%5%
1	Q	164	 70%26%5%
1	R	164	 71%24%5%
1	S	164	 72%24%.
1	T	164	 72%24%.
1	U	164	 68%27%5%
1	V	164	 70%26%5%
1	W	164	 70%26%.
1	X	164	 70%26%.
1	Y	164	 67%29%.
1	Z	164	 67%28%5%

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 65182 atoms, of which 33046 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 Flagellin.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	I	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	A	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	B	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	C	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	D	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	E	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	F	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	G	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	H	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	J	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	K	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	L	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	M	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	N	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	O	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	P	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	Q	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		

*Continued on next page...*

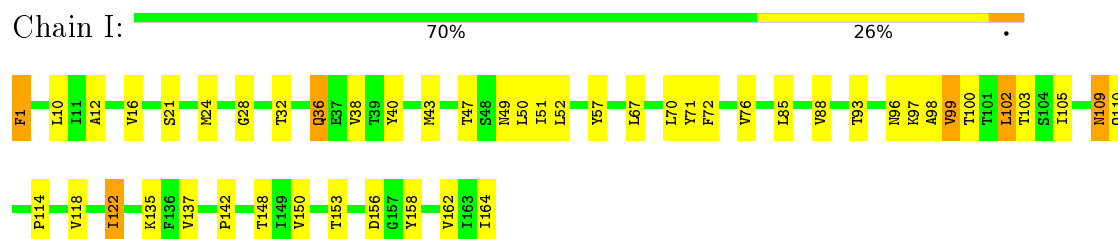
*Continued from previous page...*

Mol	Chain	Residues	Atoms						AltConf	Trace
1	R	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	S	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	T	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	U	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	V	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	W	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	X	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	Y	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		
1	Z	164	Total	C	H	N	O	S	0	0
			2507	802	1271	196	235	3		

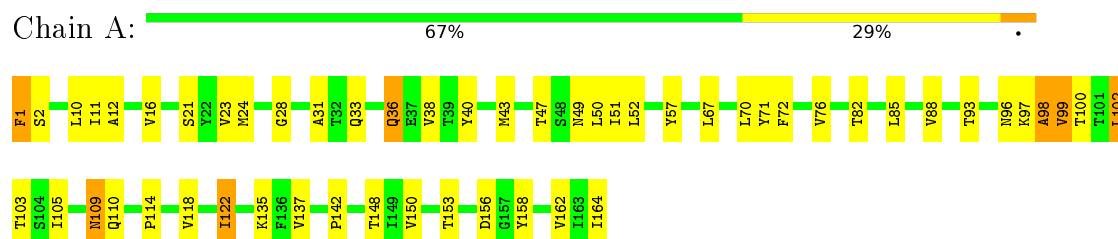
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

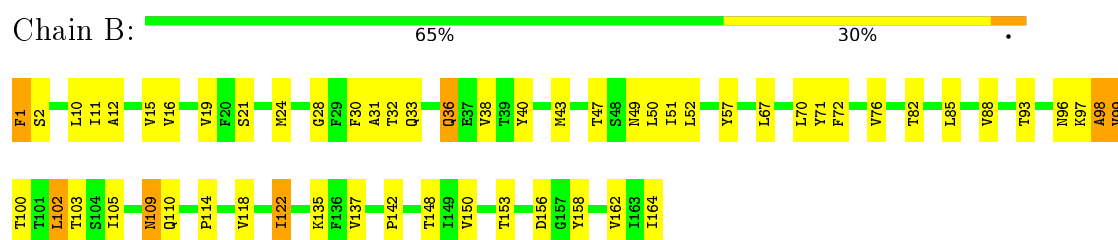
#### • Molecule 1: Flagellin



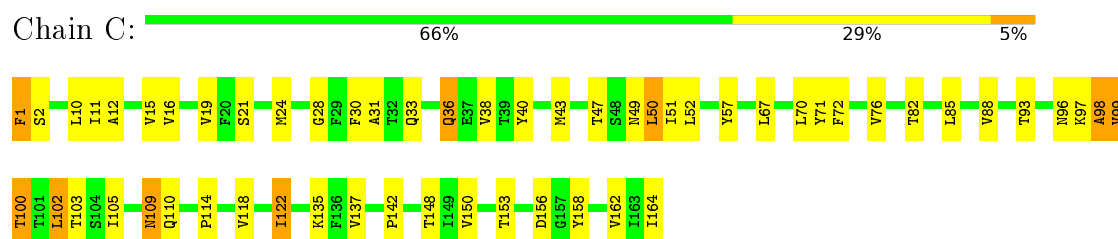
#### • Molecule 1: Flagellin



#### • Molecule 1: Flagellin

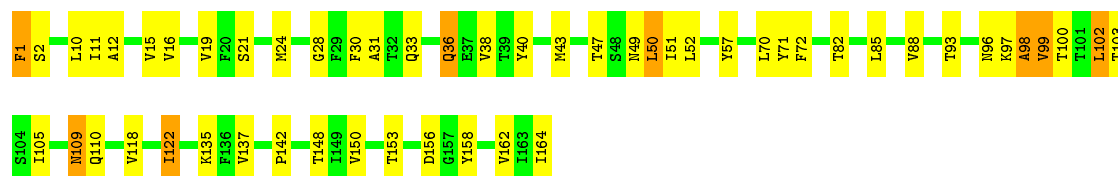


#### • Molecule 1: Flagellin



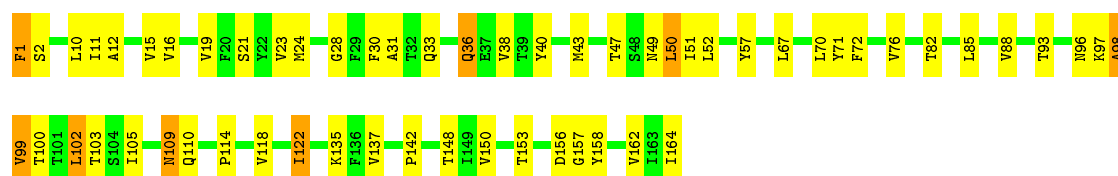
#### • Molecule 1: Flagellin

Chain D:  68% 27% 5%



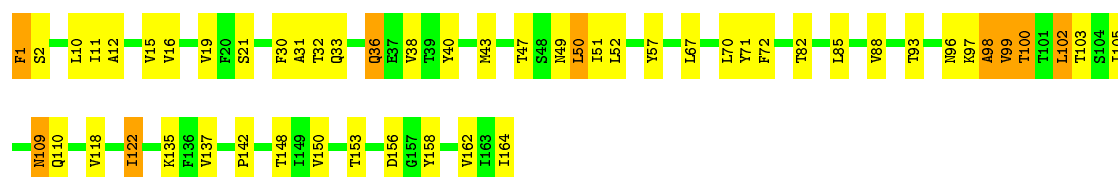
• Molecule 1: Flagellin

Chain E:  65% 30% 5%



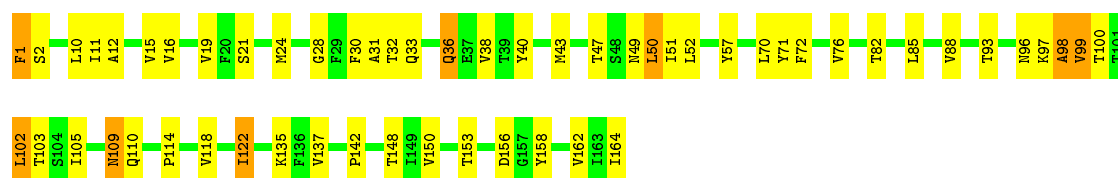
• Molecule 1: Flagellin

Chain F:  68% 27% 5%



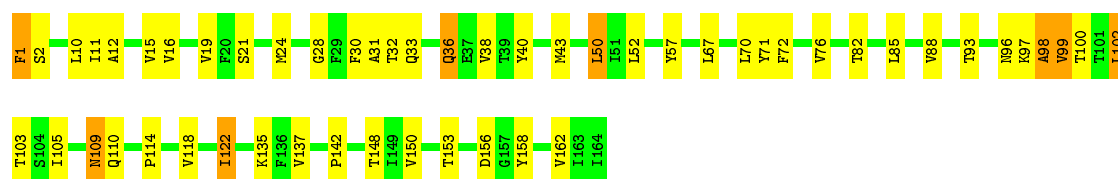
• Molecule 1: Flagellin

Chain G:  66% 29% 5%



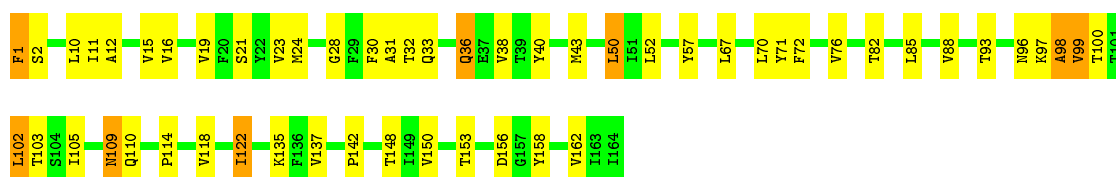
• Molecule 1: Flagellin

Chain H:  68% 27% 5%



• Molecule 1: Flagellin

Chain J:  67% 28% 5%



- Molecule 1: Flagellin



- Molecule 1: Flagellin



- Molecule 1: Flagellin

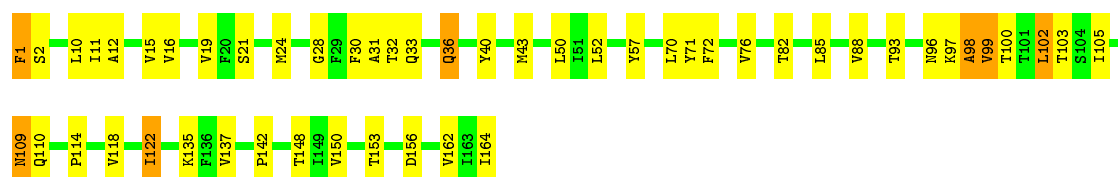


- Molecule 1: Flagellin



- Molecule 1: Flagellin





• Molecule 1: Flagellin



• Molecule 1: Flagellin



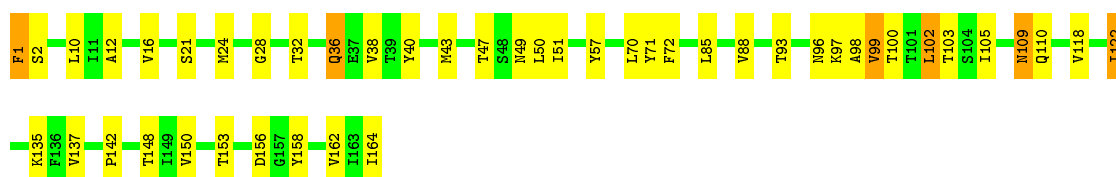
• Molecule 1: Flagellin



• Molecule 1: Flagellin

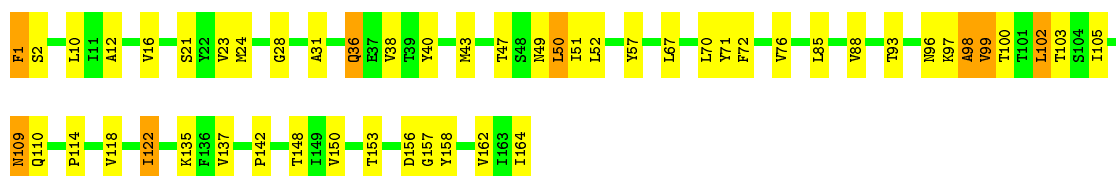


• Molecule 1: Flagellin



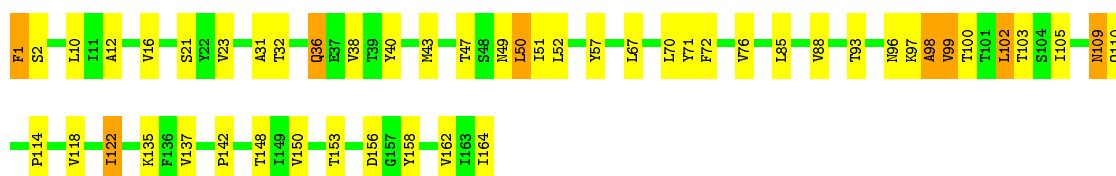
- Molecule 1: Flagellin

Chain U: 68% 27% 5%



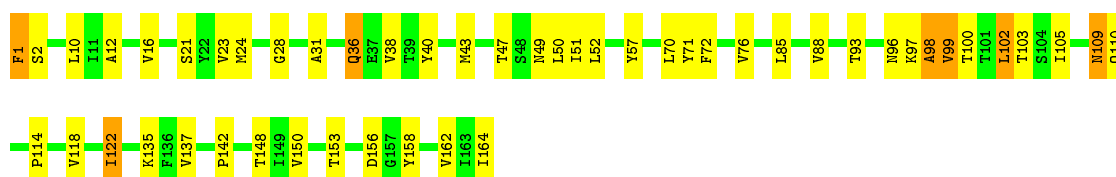
- Molecule 1: Flagellin

Chain V: 70% 26% 5%



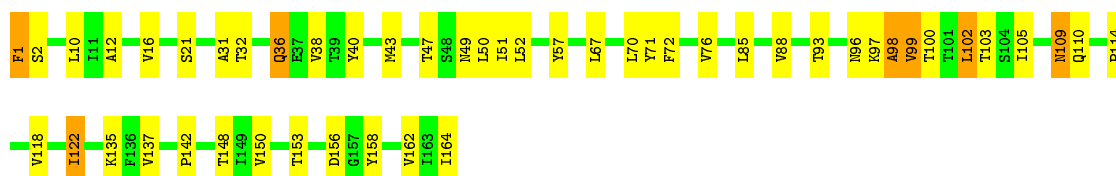
- Molecule 1: Flagellin

Chain W: 70% 26% 5%



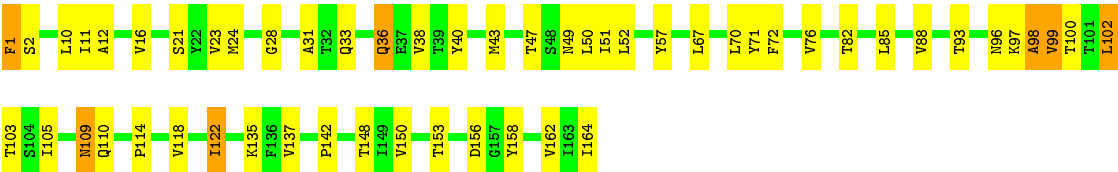
- Molecule 1: Flagellin

Chain X: 70% 26% 5%

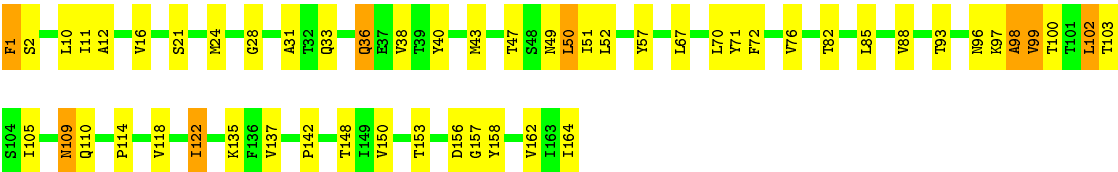


- Molecule 1: Flagellin

Chain Y: 67% 29% 5%



● Molecule 1: Flagellin



## 4 Experimental information

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of segments used	90118	Depositor
Resolution determination method	FSC 0.33 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	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	0.39	0/1260	0.86	8/1710 (0.5%)
1	B	0.39	0/1260	0.85	8/1710 (0.5%)
1	C	0.39	0/1260	0.85	9/1710 (0.5%)
1	D	0.39	0/1260	0.86	9/1710 (0.5%)
1	E	0.39	0/1260	0.86	9/1710 (0.5%)
1	F	0.40	0/1260	0.85	10/1710 (0.6%)
1	G	0.38	0/1260	0.86	8/1710 (0.5%)
1	H	0.38	0/1260	0.86	9/1710 (0.5%)
1	I	0.39	0/1260	0.86	8/1710 (0.5%)
1	J	0.38	0/1260	0.86	9/1710 (0.5%)
1	K	0.38	0/1260	0.86	9/1710 (0.5%)
1	L	0.38	0/1260	0.85	9/1710 (0.5%)
1	M	0.40	0/1260	0.85	9/1710 (0.5%)
1	N	0.39	0/1260	0.85	8/1710 (0.5%)
1	O	0.39	0/1260	0.86	8/1710 (0.5%)
1	P	0.39	0/1260	0.86	9/1710 (0.5%)
1	Q	0.39	0/1260	0.86	8/1710 (0.5%)
1	R	0.39	0/1260	0.85	9/1710 (0.5%)
1	S	0.38	0/1260	0.86	8/1710 (0.5%)
1	T	0.39	0/1260	0.86	9/1710 (0.5%)
1	U	0.39	0/1260	0.85	9/1710 (0.5%)
1	V	0.38	0/1260	0.86	8/1710 (0.5%)
1	W	0.39	0/1260	0.86	8/1710 (0.5%)
1	X	0.39	0/1260	0.86	8/1710 (0.5%)
1	Y	0.39	0/1260	0.86	9/1710 (0.5%)
1	Z	0.40	0/1260	0.85	8/1710 (0.5%)
All	All	0.39	0/32760	0.86	223/44460 (0.5%)

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
1	B	0	1
1	C	0	1
1	D	0	1
1	E	0	1
1	F	0	1
1	G	0	1
1	H	0	1
1	I	0	1
1	J	0	1
1	K	0	1
1	L	0	1
1	M	0	1
1	N	0	1
1	O	0	1
1	P	0	1
1	Q	0	1
1	R	0	1
1	S	0	1
1	T	0	1
1	U	0	1
1	V	0	1
1	W	0	1
1	X	0	1
1	Y	0	1
1	Z	0	1
All	All	0	26

There are no bond length outliers.

The worst 5 of 223 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	T	98	ALA	CB-CA-C	11.82	127.83	110.10
1	D	98	ALA	CB-CA-C	11.74	127.71	110.10
1	O	98	ALA	CB-CA-C	11.72	127.69	110.10
1	S	98	ALA	CB-CA-C	11.69	127.63	110.10
1	H	98	ALA	CB-CA-C	11.65	127.58	110.10

There are no chirality outliers.

5 of 26 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	156	ASP	Peptide
1	B	156	ASP	Peptide
1	C	156	ASP	Peptide
1	D	156	ASP	Peptide
1	I	156	ASP	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	1236	1271	1271	80	0
1	B	1236	1271	1270	92	0
1	C	1236	1271	1270	90	0
1	D	1236	1271	1270	91	0
1	E	1236	1271	1270	92	0
1	F	1236	1271	1270	87	0
1	G	1236	1271	1270	92	0
1	H	1236	1271	1270	80	0
1	I	1236	1271	1271	56	0
1	J	1236	1271	1270	80	0
1	K	1236	1271	1270	80	0
1	L	1236	1271	1271	65	0
1	M	1236	1271	1270	75	0
1	N	1236	1271	1270	76	0
1	O	1236	1271	1270	74	0
1	P	1236	1271	1270	77	0
1	Q	1236	1271	1270	72	0
1	R	1236	1271	1270	66	0
1	S	1236	1271	1270	55	0
1	T	1236	1271	1271	69	0
1	U	1236	1271	1271	78	0
1	V	1236	1271	1271	75	0
1	W	1236	1271	1271	78	0
1	X	1236	1271	1271	79	0
1	Y	1236	1271	1271	80	0
1	Z	1236	1271	1271	80	0
All	All	32136	33046	33030	1306	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 20.

The worst 5 of 1306 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:G:19:VAL:HG22	1:W:51:ILE:HD11	1.32	1.12
1:G:51:ILE:HD11	1:S:19:VAL:HG22	1.32	1.12
1:D:51:ILE:HD11	1:P:19:VAL:HG22	1.32	1.11
1:H:19:VAL:HG22	1:X:51:ILE:HD11	1.32	1.11
1:D:19:VAL:HG22	1:T:51:ILE:HD11	1.32	1.11

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	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	B	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	C	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	D	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	E	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	F	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	G	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	H	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	I	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	J	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	K	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	L	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	M	162/164 (99%)	154 (95%)	8 (5%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	N	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	O	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	P	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	Q	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	R	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	S	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	T	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	U	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	V	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	W	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	X	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	Y	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
1	Z	162/164 (99%)	154 (95%)	8 (5%)	0	100	100
All	All	4212/4264 (99%)	4004 (95%)	208 (5%)	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	A	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	B	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	C	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	D	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	E	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	F	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	G	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	H	132/132 (100%)	128 (97%)	4 (3%)	48	81

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	I	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	J	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	K	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	L	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	M	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	N	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	O	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	P	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	Q	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	R	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	S	132/132 (100%)	129 (98%)	3 (2%)	58	85
1	T	132/132 (100%)	129 (98%)	3 (2%)	58	85
1	U	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	V	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	W	132/132 (100%)	129 (98%)	3 (2%)	58	85
1	X	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	Y	132/132 (100%)	128 (97%)	4 (3%)	48	81
1	Z	132/132 (100%)	128 (97%)	4 (3%)	48	81
All	All	3432/3432 (100%)	3331 (97%)	101 (3%)	54	82

5 of 101 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	36	GLN
1	O	1	PHE
1	Y	1	PHE
1	L	148	THR
1	M	148	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
1	L	96	ASN
1	P	96	ASN
1	Y	49	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	L	111	GLN
1	N	96	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 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.