



## Full wwPDB EM Validation Report ⓘ

Dec 12, 2022 – 09:12 AM EST

PDB ID : 6X8M  
EMDB ID : EMD-22094  
Title : CryoEM structure of the holo-SrpI encapsulin complex from *Synechococcus elongatus* PCC 7942  
Authors : LaFrance, B.J.; Nichols, R.J.; Phillips, N.R.; Oltrogge, L.M.; Valentin-Alvarado, L.E.; Bischoff, A.J.; Savage, D.F.; Nogales, E.  
Deposited on : 2020-06-01  
Resolution : 2.20 Å(reported)

This is a Full wwPDB EM Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

EMDB validation analysis : **FAILED**  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.31.2

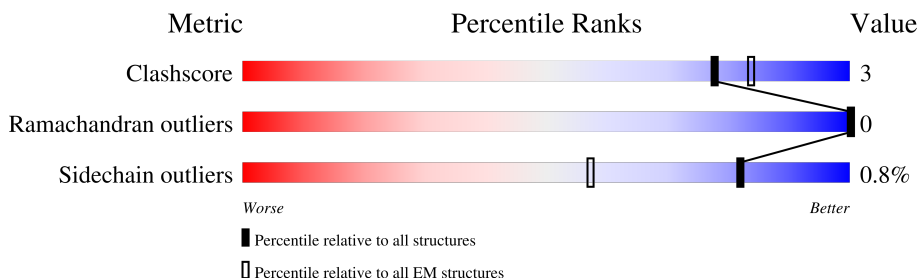
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*


The reported resolution of this entry is 2.20 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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 of the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	A	314	 82%      6% • 11%

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2224 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 Protein SrpL.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	280	Total	C	N	O	S	0	0
			2224	1416	398	405	5		


There are 8 discrepancies between the modelled and reference sequences:

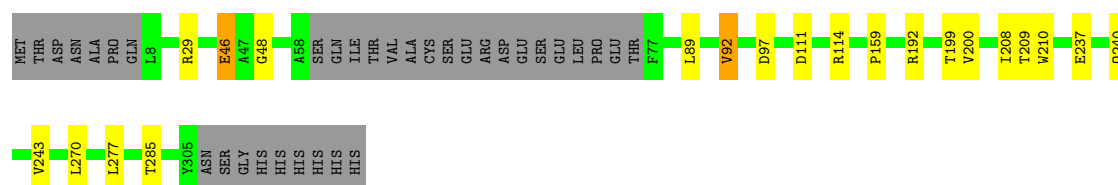
Chain	Residue	Modelled	Actual	Comment	Reference
A	307	SER	-	expression tag	UNP Q55032
A	308	GLY	-	expression tag	UNP Q55032
A	309	HIS	-	expression tag	UNP Q55032
A	310	HIS	-	expression tag	UNP Q55032
A	311	HIS	-	expression tag	UNP Q55032
A	312	HIS	-	expression tag	UNP Q55032
A	313	HIS	-	expression tag	UNP Q55032
A	314	HIS	-	expression tag	UNP Q55032

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

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Protein SrpI

Chain A:  82% 6% 11%



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	11980	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION; CTF refinement was performed in Relion/3.1	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	35	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2800	Depositor
Magnification	64000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	A	0.30	0/2271	0.46	0/3095

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	2224	0	2254	13	0
All	All	2224	0	2254	13	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 3.

All (13) 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:A:237:GLU:HG3	1:A:243:VAL:HG11	1.75	0.67
1:A:97:ASP:HB3	1:A:270:LEU:HD22	1.82	0.61
1:A:159:PRO:HB3	1:A:210:TRP:CZ2	2.37	0.60
1:A:92:VAL:HG23	1:A:277:LEU:HB3	1.85	0.57
1:A:48:GLY:HA2	1:A:89:LEU:HD12	1.87	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:199:THR:HG22	1:A:208:ILE:HD13	1.91	0.52
1:A:111:ASP:HB3	1:A:114:ARG:HB3	1.92	0.52
1:A:200:VAL:HG13	1:A:209:THR:HG21	1.94	0.49
1:A:46:GLU:H	1:A:46:GLU:HG2	1.50	0.44
1:A:192:ARG:HH11	1:A:192:ARG:HA	1.82	0.44
1:A:29:ARG:HE	1:A:29:ARG:HB2	1.71	0.43
1:A:208:ILE:HD13	1:A:208:ILE:HA	1.86	0.42
1:A:240:GLN:O	1:A:285:THR:OG1	2.35	0.41

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	276/314 (88%)	267 (97%)	9 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	243/274 (89%)	241 (99%)	2 (1%)	81	90

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	46	GLU
1	A	92	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	142	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 monosaccharides 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.