



wwPDB EM Validation Summary Report ⓘ

Mar 20, 2026 – 06:45 AM UTC

PDB ID : 7OE2 / pdb_00007oe2
EMDB ID : EMD-12859
Title : Model of closed pentamer of the Haliangium ochraceum encapsulin from symmetry expansion of icosahedral single particle reconstruction
Authors : Marles-Wright, J.; Basle, A.; Clarke, D.J.; Ross, J.
Deposited on : 2021-05-01
Resolution : 2.40 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB 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 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 : 0.0.1.dev132
MolProbity : 4-5-2 with Phenix2.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics : **NOT EXECUTED**
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

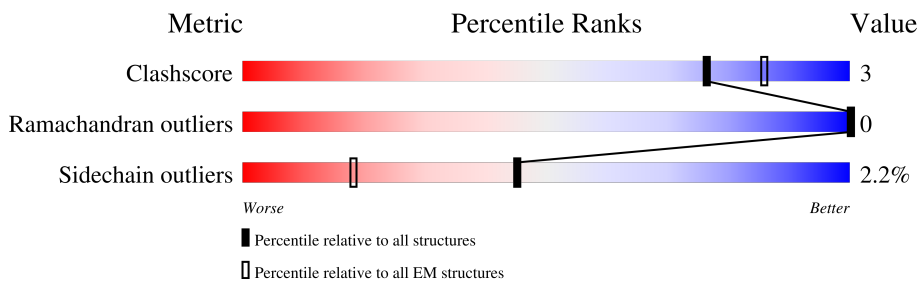
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 2.40 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	229148	23984
Ramachandran outliers	224038	23583
Sidechain outliers	223484	23102

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 ≥ 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	266	92% 8%
1	B	266	90% 10%
1	C	266	91% 9%
1	D	266	92% 8%
1	E	266	90% 10%
2	1	131	7% 93%
2	2	131	7% 93%
2	3	131	7% 93%
2	4	131	7% 93%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	5	131	 7% 93%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 10589 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 Linocin_M18 bacteriocin protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	266	Total	C	N	O	S	0	0
			2034	1290	352	387	5		
1	B	266	Total	C	N	O	S	0	0
			2034	1290	352	387	5		
1	C	266	Total	C	N	O	S	0	0
			2034	1290	352	387	5		
1	D	266	Total	C	N	O	S	0	0
			2034	1290	352	387	5		
1	E	266	Total	C	N	O	S	0	0
			2034	1290	352	387	5		

- Molecule 2 is a protein called Haliangium ochraceum Encapsulated ferritin localisation sequence.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
2	1	9	Total	C	N	O	0	0
			53	33	9	11		
2	2	9	Total	C	N	O	0	0
			53	33	9	11		
2	3	9	Total	C	N	O	0	0
			53	33	9	11		
2	4	9	Total	C	N	O	0	0
			53	33	9	11		
2	5	9	Total	C	N	O	0	0
			53	33	9	11		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		AltConf
			Total	O	
3	A	26	Total	O	0
			26	26	
3	B	27	Total	O	0
			27	27	

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	AltConf
3	C	32	Total O 32 32	0
3	D	32	Total O 32 32	0
3	E	27	Total O 27 27	0
3	1	2	Total O 2 2	0
3	2	2	Total O 2 2	0
3	3	1	Total O 1 1	0
3	4	3	Total O 3 3	0
3	5	2	Total O 2 2	0

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: Linocin_M18 bacteriocin protein

Chain A:  92% 8%




- Molecule 1: Linocin_M18 bacteriocin protein

Chain B:  90% 10%



- Molecule 1: Linocin_M18 bacteriocin protein

Chain C:  91% 9%




- Molecule 1: Linocin_M18 bacteriocin protein

Chain D:  92% 8%



- Molecule 1: Linocin_M18 bacteriocin protein

Chain E:  90% 10%



- Molecule 2: Haliangium ochraceum Encapsulated ferritin localisation sequence

Chain 1:  7% 93%

VAL
GLU
HIS
ALA
MET
MET
THR
LEU
GLU
TRP
ILE
ARG
ARG
SER
PRO
VAL
PHE
ASP
ALA
HIS
MET
ARG
THR
TYR
LEU
LEU
PHE
THR
GLU
ARG
PRO
ILE
LEU
GLU
LEU
GLU
GLU
GLU
ASP
THR
GLY
SER
SER
SER
SER
VAL
ALA
ALA
SER
PRO
THR
SER
ALA
PRO
SER
HIS
G917
R925
GLN

GLU
GLY
LYS
GLU
ASP

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	4734551	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION; CTF correction after motion correction followed by CFT refinement after 3D Refinement.	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	40.509	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 (6k 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	0/2078	0.41	0/2828
1	B	0.30	0/2078	0.41	0/2828
1	C	0.30	0/2078	0.41	0/2828
1	D	0.30	0/2078	0.41	0/2828
1	E	0.30	0/2078	0.41	0/2828
2	1	0.17	0/52	0.32	0/68
2	2	0.16	0/52	0.32	0/68
2	3	0.17	0/52	0.31	0/68
2	4	0.17	0/52	0.31	0/68
2	5	0.17	0/52	0.32	0/68
All	All	0.30	0/10650	0.41	0/14480

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	2034	0	2006	10	0
1	B	2034	0	2006	14	0
1	C	2034	0	2006	13	0
1	D	2034	0	2006	12	0
1	E	2034	0	2006	14	0
2	1	53	0	53	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	2	53	0	53	0	0
2	3	53	0	53	0	0
2	4	53	0	53	0	0
2	5	53	0	53	0	0
3	1	2	0	0	0	0
3	2	2	0	0	0	0
3	3	1	0	0	0	0
3	4	3	0	0	0	0
3	5	2	0	0	0	0
3	A	26	0	0	0	0
3	B	27	0	0	0	0
3	C	32	0	0	0	0
3	D	32	0	0	1	0
3	E	27	0	0	1	0
All	All	10589	0	10295	63	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.

The worst 5 of 63 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:D:264:ARG:HH11	1:D:264:ARG:HG3	1.63	0.64
1:A:264:ARG:HH11	1:A:264:ARG:HG3	1.63	0.63
1:B:264:ARG:HH11	1:B:264:ARG:HG3	1.63	0.63
1:C:264:ARG:HG3	1:C:264:ARG:HH11	1.63	0.63
1:E:264:ARG:HH11	1:E:264:ARG:HG3	1.63	0.63

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	264/266 (99%)	258 (98%)	6 (2%)	0	100	100
1	B	264/266 (99%)	258 (98%)	6 (2%)	0	100	100
1	C	264/266 (99%)	258 (98%)	6 (2%)	0	100	100
1	D	264/266 (99%)	258 (98%)	6 (2%)	0	100	100
1	E	264/266 (99%)	258 (98%)	6 (2%)	0	100	100
2	1	7/131 (5%)	7 (100%)	0	0	100	100
2	2	7/131 (5%)	7 (100%)	0	0	100	100
2	3	7/131 (5%)	7 (100%)	0	0	100	100
2	4	7/131 (5%)	7 (100%)	0	0	100	100
2	5	7/131 (5%)	7 (100%)	0	0	100	100
All	All	1355/1985 (68%)	1325 (98%)	30 (2%)	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	208/208 (100%)	202 (97%)	6 (3%)	37	60
1	B	208/208 (100%)	204 (98%)	4 (2%)	50	71
1	C	208/208 (100%)	204 (98%)	4 (2%)	50	71
1	D	208/208 (100%)	204 (98%)	4 (2%)	50	71
1	E	208/208 (100%)	203 (98%)	5 (2%)	43	65
2	1	5/115 (4%)	5 (100%)	0	100	100
2	2	5/115 (4%)	5 (100%)	0	100	100
2	3	5/115 (4%)	5 (100%)	0	100	100
2	4	5/115 (4%)	5 (100%)	0	100	100
2	5	5/115 (4%)	5 (100%)	0	100	100
All	All	1065/1615 (66%)	1042 (98%)	23 (2%)	45	67

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	2	ASP
1	D	92	VAL
1	D	61	THR
1	E	2	ASP
1	B	2	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	73	GLN
1	B	73	GLN
1	C	73	GLN
1	D	73	GLN
1	E	73	GLN

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 oligosaccharides 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

There are no chain breaks in this entry.

6 Map visualisation

This section contains visualisations of the EMDB entry EMD-12859. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal standard-deviation projections (False-color)

This section was not generated.

6.5 Orthogonal surface views

This section was not generated.

6.6 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution

This section was not generated.

7.2 Volume estimate versus contour level

This section was not generated.

7.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

This section was not generated.