wwPDB EM Validation Summary Report (i)

May 4, 2024 – 12:37 pm BST

PDB ID	:	5LI2
EMDB ID	:	EMD-4051
Title	:	bacteriophage phi812K1-420 tail sheath and tail tube protein in native tail
Authors	:	Novacek, J.; Siborova, M.; Benesik, M.; Pantucek, R.; Doskar, J.; Plevka, P.
Deposited on	:	2016-07-14
Resolution	:	6.20 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 (i) 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 (i)) were used in the production of this report:

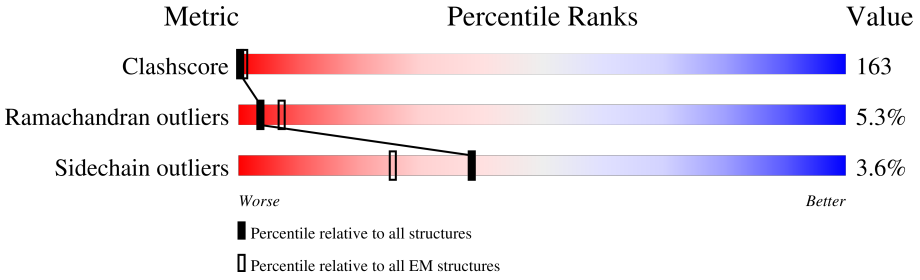
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 6.20 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

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 24732 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 tail sheath protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	427	3345	2110	567	661	7	0	0
1	B	427	3345	2110	567	661	7	0	0
1	C	427	3345	2110	567	661	7	0	0
1	D	427	3345	2110	567	661	7	0	0
1	E	427	3345	2110	567	661	7	0	0
1	F	427	3345	2110	567	661	7	0	0

- Molecule 2 is a protein called Phage-like element PBSX protein XkdM.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	G	100	777	489	124	161	3	0	0
2	H	100	777	489	124	161	3	0	0
2	I	100	777	489	124	161	3	0	0
2	J	100	777	489	124	161	3	0	0
2	K	100	777	489	124	161	3	0	0
2	L	100	777	489	124	161	3	0	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	148	LEU	-	expression tag	UNP P54332
G	149	GLU	-	expression tag	UNP P54332

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	150	HIS	-	expression tag	UNP P54332
G	151	HIS	-	expression tag	UNP P54332
G	152	HIS	-	expression tag	UNP P54332
G	153	HIS	-	expression tag	UNP P54332
G	154	HIS	-	expression tag	UNP P54332
G	155	HIS	-	expression tag	UNP P54332
H	148	LEU	-	expression tag	UNP P54332
H	149	GLU	-	expression tag	UNP P54332
H	150	HIS	-	expression tag	UNP P54332
H	151	HIS	-	expression tag	UNP P54332
H	152	HIS	-	expression tag	UNP P54332
H	153	HIS	-	expression tag	UNP P54332
H	154	HIS	-	expression tag	UNP P54332
H	155	HIS	-	expression tag	UNP P54332
I	148	LEU	-	expression tag	UNP P54332
I	149	GLU	-	expression tag	UNP P54332
I	150	HIS	-	expression tag	UNP P54332
I	151	HIS	-	expression tag	UNP P54332
I	152	HIS	-	expression tag	UNP P54332
I	153	HIS	-	expression tag	UNP P54332
I	154	HIS	-	expression tag	UNP P54332
I	155	HIS	-	expression tag	UNP P54332
J	148	LEU	-	expression tag	UNP P54332
J	149	GLU	-	expression tag	UNP P54332
J	150	HIS	-	expression tag	UNP P54332
J	151	HIS	-	expression tag	UNP P54332
J	152	HIS	-	expression tag	UNP P54332
J	153	HIS	-	expression tag	UNP P54332
J	154	HIS	-	expression tag	UNP P54332
J	155	HIS	-	expression tag	UNP P54332
K	148	LEU	-	expression tag	UNP P54332
K	149	GLU	-	expression tag	UNP P54332
K	150	HIS	-	expression tag	UNP P54332
K	151	HIS	-	expression tag	UNP P54332
K	152	HIS	-	expression tag	UNP P54332
K	153	HIS	-	expression tag	UNP P54332
K	154	HIS	-	expression tag	UNP P54332
K	155	HIS	-	expression tag	UNP P54332
L	148	LEU	-	expression tag	UNP P54332
L	149	GLU	-	expression tag	UNP P54332
L	150	HIS	-	expression tag	UNP P54332
L	151	HIS	-	expression tag	UNP P54332

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	152	HIS	-	expression tag	UNP P54332
L	153	HIS	-	expression tag	UNP P54332
L	154	HIS	-	expression tag	UNP P54332
L	155	HIS	-	expression tag	UNP P54332

SEQUENCE-PLOTS INFOmissingINFO

3 Experimental information (i)

Property	Value	Source
EM reconstruction method	HELICAL	Depositor
Imposed symmetry	HELICAL, twist=21.4°, rise=38.9 Å, axial sym=C6	Depositor
Number of segments used	3628	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	20	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON II (4k x 4k)	Depositor

4 Model quality (i)

4.1 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	3.49	230/3387 (6.8%)	3.05	104/4554 (2.3%)
1	B	3.49	231/3387 (6.8%)	3.05	104/4554 (2.3%)
1	C	3.48	235/3387 (6.9%)	3.05	104/4554 (2.3%)
1	D	3.49	231/3387 (6.8%)	3.05	104/4554 (2.3%)
1	E	3.49	232/3387 (6.8%)	3.05	104/4554 (2.3%)
1	F	3.49	233/3387 (6.9%)	3.05	104/4554 (2.3%)
2	G	0.64	2/783 (0.3%)	0.69	0/1046
2	H	0.64	2/783 (0.3%)	0.69	0/1046
2	I	0.64	2/783 (0.3%)	0.69	0/1046
2	J	0.64	2/783 (0.3%)	0.70	0/1046
2	K	0.64	2/783 (0.3%)	0.70	0/1046
2	L	0.64	2/783 (0.3%)	0.70	0/1046
All	All	3.15	1404/25020 (5.6%)	2.77	624/33600 (1.9%)

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	3
1	B	0	3
1	C	0	3
1	D	0	3
1	E	0	3
1	F	0	3
All	All	0	18

The worst 5 of 1404 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	546	PRO	N-CD	55.46	2.25	1.47
1	F	546	PRO	N-CD	55.45	2.25	1.47

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	546	PRO	N-CD	55.42	2.25	1.47
1	A	546	PRO	N-CD	55.41	2.25	1.47
1	D	546	PRO	N-CD	55.41	2.25	1.47

The worst 5 of 624 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	380	ARG	NE-CZ-NH1	-25.49	107.56	120.30
1	A	64	ARG	NE-CZ-NH2	-25.46	107.57	120.30
1	B	64	ARG	NE-CZ-NH2	-25.42	107.59	120.30
1	C	380	ARG	NE-CZ-NH1	-25.41	107.60	120.30
1	F	380	ARG	NE-CZ-NH1	-25.40	107.60	120.30

There are no chirality outliers.

5 of 18 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	392	VAL	Peptide
1	A	393	ALA	Peptide
1	A	394	ASN	Peptide
1	B	392	VAL	Peptide
1	B	393	ALA	Peptide

4.2 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	3345	0	3267	1020	0
1	B	3345	0	3267	1031	0
1	C	3345	0	3267	1038	0
1	D	3345	0	3267	1045	0
1	E	3345	0	3267	1033	0
1	F	3345	0	3267	1029	0
2	G	777	0	731	456	0
2	H	777	0	731	457	0
2	I	777	0	731	464	0
2	J	777	0	731	471	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	K	777	0	731	475	0
2	L	777	0	731	465	0
All	All	24732	0	23988	7933	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 163.

The worst 5 of 7933 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:B:320:TRP:CD1	1:B:344:VAL:HG23	1.31	1.65
1:A:252:PHE:CE2	1:A:257:LYS:HG2	1.11	1.63
1:B:252:PHE:CE2	1:B:257:LYS:HG2	1.11	1.63
1:C:320:TRP:CD1	1:C:344:VAL:HG23	1.31	1.63
1:A:320:TRP:CD1	1:A:344:VAL:HG23	1.31	1.63

There are no symmetry-related clashes.

4.3 Torsion angles (i)

4.3.1 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	401/587 (68%)	316 (79%)	63 (16%)	22 (6%)	2 19
1	B	401/587 (68%)	316 (79%)	63 (16%)	22 (6%)	2 19
1	C	401/587 (68%)	316 (79%)	63 (16%)	22 (6%)	2 19
1	D	401/587 (68%)	316 (79%)	63 (16%)	22 (6%)	2 19
1	E	401/587 (68%)	316 (79%)	63 (16%)	22 (6%)	2 19
1	F	401/587 (68%)	316 (79%)	63 (16%)	22 (6%)	2 19
2	G	88/155 (57%)	72 (82%)	12 (14%)	4 (4%)	2 21
2	H	88/155 (57%)	72 (82%)	12 (14%)	4 (4%)	2 21

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	I	88/155 (57%)	72 (82%)	12 (14%)	4 (4%)	2	21
2	J	88/155 (57%)	72 (82%)	12 (14%)	4 (4%)	2	21
2	K	88/155 (57%)	72 (82%)	12 (14%)	4 (4%)	2	21
2	L	88/155 (57%)	72 (82%)	12 (14%)	4 (4%)	2	21
All	All	2934/4452 (66%)	2328 (79%)	450 (15%)	156 (5%)	3	19

5 of 156 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	79	ASN
1	A	337	PRO
1	A	361	PRO
1	A	409	PRO
1	A	429	SER

4.3.2 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	363/495 (73%)	358 (99%)	5 (1%)	67	80
1	B	363/495 (73%)	358 (99%)	5 (1%)	67	80
1	C	363/495 (73%)	358 (99%)	5 (1%)	67	80
1	D	363/495 (73%)	358 (99%)	5 (1%)	67	80
1	E	363/495 (73%)	358 (99%)	5 (1%)	67	80
1	F	363/495 (73%)	358 (99%)	5 (1%)	67	80
2	G	84/134 (63%)	73 (87%)	11 (13%)	4	18
2	H	84/134 (63%)	73 (87%)	11 (13%)	4	18
2	I	84/134 (63%)	73 (87%)	11 (13%)	4	18
2	J	84/134 (63%)	73 (87%)	11 (13%)	4	18
2	K	84/134 (63%)	73 (87%)	11 (13%)	4	18
2	L	84/134 (63%)	73 (87%)	11 (13%)	4	18

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2682/3774 (71%)	2586 (96%)	96 (4%)	38 59

5 of 96 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	I	76	ASP
2	J	135	GLU
2	I	103	ARG
2	J	17	PHE
2	K	10	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	407	HIS
2	J	7	ASN
1	E	81	ASN
2	L	7	ASN
2	G	7	ASN

4.3.3 RNA (i)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

There are no ligands in this entry.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	9
1	B	9
1	C	9
1	D	9
1	E	9
1	F	9

The worst 5 of 54 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	427:GLY	C	428:GLU	N	15.07
1	B	427:GLY	C	428:GLU	N	15.07
1	C	427:GLY	C	428:GLU	N	15.07
1	D	427:GLY	C	428:GLU	N	15.07
1	E	427:GLY	C	428:GLU	N	15.07

5 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-4051. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

5.1 Orthogonal projections (i)

This section was not generated.

5.2 Central slices (i)

This section was not generated.

5.3 Largest variance slices (i)

This section was not generated.

5.4 Orthogonal standard-deviation projections (False-color) (i)

This section was not generated.

5.5 Orthogonal surface views (i)

This section was not generated.

5.6 Mask visualisation (i)

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

6 Map analysis (i)

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

6.1 Map-value distribution (i)

This section was not generated.

6.2 Volume estimate versus contour level (i)

This section was not generated.

6.3 Rotationally averaged power spectrum (i)

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

7 Fourier-Shell correlation (i)

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

8 Map-model fit (i)

This section was not generated.