



wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Jan 19, 2017 – 06:52 PM EST

PDB ID : 5TC1
EMDB ID: : EMD-8397
Title : In situ structures of the genome and genome-delivery apparatus in ssRNA bacteriophage MS2
Authors : Dai, X.H.; Li, Z.H.; Lai, M.; Shu, S.; Du, Y.S.; Zhou, Z.H.; Sun, R.
Deposited on : 2016-09-13
Resolution : 3.60 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
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

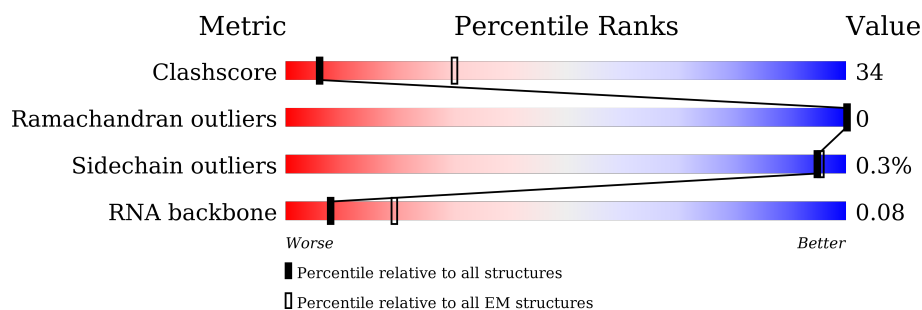
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.60 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
RNA backbone	3027	244

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 ≥ 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	130	89% 10% .
1	B	130	80% 19% .
1	C	130	90% 9% .
1	D	130	75% 25% .
1	E	130	75% 25% .
1	F	130	74% 25% ..
1	G	130	86% 5% 8%
1	H	130	85% 14% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	M	393	<div><div></div><div>70%</div><div>14%</div><div>16%</div></div>
3	R	3569	<div><div></div><div>93%</div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 15393 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 Capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	129	Total	C	N	O	S	0	0
			965	606	165	190	4		
1	B	129	Total	C	N	O	S	0	0
			965	606	165	190	4		
1	C	129	Total	C	N	O	S	0	0
			965	606	165	190	4		
1	D	129	Total	C	N	O	S	0	0
			965	606	165	190	4		
1	E	129	Total	C	N	O	S	0	0
			965	606	165	190	4		
1	F	129	Total	C	N	O	S	0	0
			965	606	165	190	4		
1	G	119	Total	C	N	O	S	0	0
			898	565	154	175	4		
1	H	129	Total	C	N	O	S	0	0
			965	606	165	190	4		

- Molecule 2 is a protein called Maturation protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	M	330	Total	C	N	O	S	0	0
			2608	1664	465	468	11		

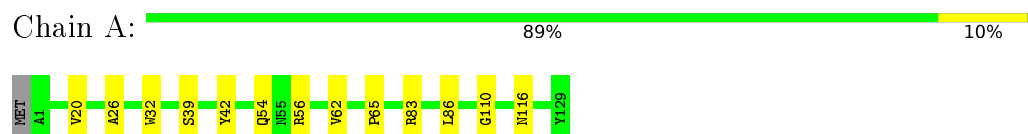
- Molecule 3 is a RNA chain called phage MS2 genome.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	R	242	Total	C	N	O	P	0	0
			5132	2297	897	1696	242		

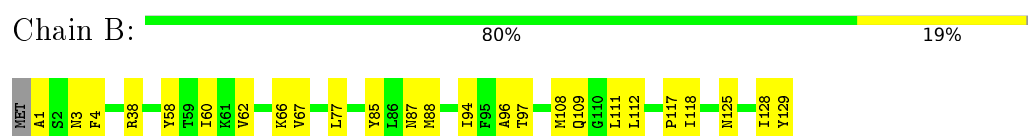
3 Residue-property plots [i](#)

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.

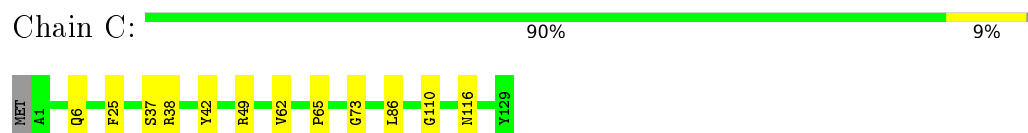
- Molecule 1: Capsid protein



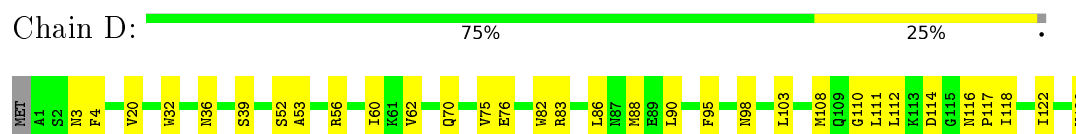
- Molecule 1: Capsid protein



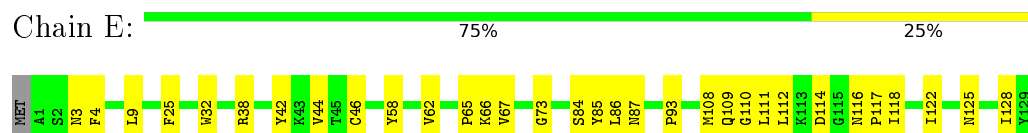
- Molecule 1: Capsid protein



- Molecule 1: Capsid protein

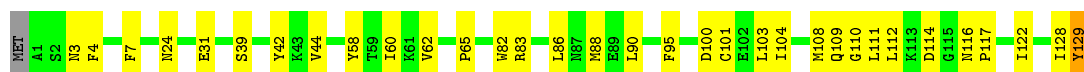


- Molecule 1: Capsid protein



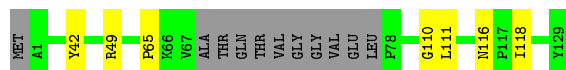
- Molecule 1: Capsid protein





- Molecule 1: Capsid protein

Chain G: 86% 5% 8%



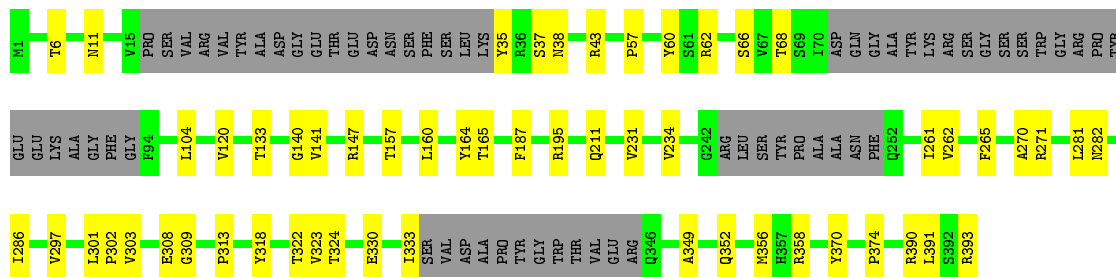
- Molecule 1: Capsid protein

Chain H: 85% 14% .



- Molecule 2: Maturation protein

Chain M: 70% 14% 16%



- Molecule 3: phage MS2 genome

Chain R: 93%



[illegible]




[illegible]

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	339718	Depositor
Resolution determination method	FSC 0.143 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	47170	Depositor
Image detector	Not provided	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 >2	RMSZ	# Z >2
1	A	0.30	0/982	0.50	0/1337
1	B	0.30	0/982	0.51	0/1337
1	C	0.31	0/982	0.51	0/1337
1	D	0.35	0/982	0.52	0/1337
1	E	0.31	0/982	0.52	0/1337
1	F	0.31	0/982	0.51	0/1337
1	G	0.30	0/914	0.49	0/1241
1	H	0.29	0/982	0.50	0/1337
2	M	0.32	0/2665	0.53	1/3617 (0.0%)
3	R	0.17	0/5719	0.61	0/8867
All	All	0.27	0/16172	0.55	1/23084 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	M	104	LEU	C-N-CA	7.43	140.27	121.70

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	965	0	964	14	0
1	B	965	0	964	49	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	965	0	964	10	0
1	D	965	0	963	49	0
1	E	965	0	962	71	0
1	F	965	0	964	72	0
1	G	898	0	896	7	0
1	H	965	0	964	31	0
2	M	2608	0	2619	75	0
3	R	5132	0	2611	759	0
All	All	15393	0	12871	945	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 34.

The worst 5 of 945 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:56:ARG:CZ	1:E:38:ARG:NH1	1.91	1.32
1:E:58:TYR:CE2	1:F:122:ILE:HG12	1.63	1.31
2:M:356:MET:SD	3:R:3554:G:H4'	1.71	1.29
3:R:106:U:C5	3:R:108:C:C5	2.20	1.29
1:E:114:ASP:OD1	1:F:109:GLN:NE2	1.65	1.26

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	127/130 (98%)	121 (95%)	6 (5%)	0	100	100
1	B	127/130 (98%)	124 (98%)	3 (2%)	0	100	100
1	C	127/130 (98%)	121 (95%)	6 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	127/130 (98%)	122 (96%)	5 (4%)	0	100	100
1	E	127/130 (98%)	125 (98%)	2 (2%)	0	100	100
1	F	127/130 (98%)	122 (96%)	5 (4%)	0	100	100
1	G	115/130 (88%)	108 (94%)	7 (6%)	0	100	100
1	H	127/130 (98%)	122 (96%)	5 (4%)	0	100	100
2	M	320/393 (81%)	315 (98%)	5 (2%)	0	100	100
All	All	1324/1433 (92%)	1280 (97%)	44 (3%)	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	106/107 (99%)	106 (100%)	0	100	100
1	B	106/107 (99%)	106 (100%)	0	100	100
1	C	106/107 (99%)	105 (99%)	1 (1%)	84	94
1	D	106/107 (99%)	105 (99%)	1 (1%)	84	94
1	E	106/107 (99%)	106 (100%)	0	100	100
1	F	106/107 (99%)	105 (99%)	1 (1%)	84	94
1	G	99/107 (92%)	99 (100%)	0	100	100
1	H	106/107 (99%)	106 (100%)	0	100	100
2	M	280/330 (85%)	280 (100%)	0	100	100
All	All	1121/1186 (94%)	1118 (100%)	3 (0%)	95	99

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	49	ARG
1	D	36	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	129	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	116	ASN
1	F	109	GLN
2	M	211	GLN
1	D	36	ASN
2	M	11	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	228/3569 (6%)	164 (71%)	16 (7%)

5 of 164 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	R	104	A
3	R	106	U
3	R	107	U
3	R	109	C
3	R	110	A

5 of 16 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	R	1755	U
3	R	1782	U
3	R	2849	A
3	R	1729	A
3	R	2851	C

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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.