wwPDB EM Validation Summary Report

Dec 12, 2022 – 11:25 AM EST

PDB ID	:	6WKK
EMDB ID	:	EMD-21695
Title	:	Phage G gp27 major capsid proteins and gp26 decoration proteins
Authors	:	Monroe, L.; Gonzalez, B.; Jiang, W.; Kihara, D.
Deposited on	:	2020-04-16
Resolution	:	6.10 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:

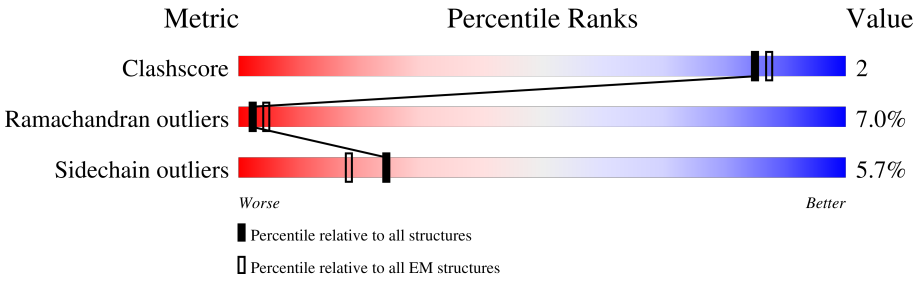
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 6.10 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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	280	74% 20% 6% •
1	B	280	73% 21% 5%
1	C	280	74% 22% • •
1	D	280	70% 24% 5% •
1	E	280	68% 26% 5% •
1	F	280	74% 20% 6%
2	G	150	67% 29% • •
2	H	150	69% 23% 7% •
2	I	150	69% 25% 5% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	150	70% 23% 6%
2	K	150	63% 33%
2	L	150	60% 30% 9%
2	M	150	64% 26% 9%
2	N	150	75% 21%
2	O	150	60% 30% 9%
2	P	150	69% 25% 5%
2	Q	150	72% 23% 5%
2	R	150	69% 25%
2	S	150	71% 23% 5%
2	T	150	67% 28%
2	U	150	68% 29%
2	V	150	71% 21% 6%
2	W	150	75% 20% 5%
2	X	150	69% 21% 9%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 33000 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 Gp27 major capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	280	Total	C	N	O	S	0	0
			2182	1390	374	413	5		
1	B	280	Total	C	N	O	S	0	0
			2182	1390	374	413	5		
1	C	280	Total	C	N	O	S	0	0
			2182	1390	374	413	5		
1	D	280	Total	C	N	O	S	0	0
			2182	1390	374	413	5		
1	E	280	Total	C	N	O	S	0	0
			2182	1390	374	413	5		
1	F	280	Total	C	N	O	S	0	0
			2182	1390	374	413	5		

- Molecule 2 is a protein called Gp26 capsid decoration protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	G	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	H	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	I	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	J	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	K	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	L	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	M	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	N	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	O	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		

Continued on next page...

Continued from previous page...

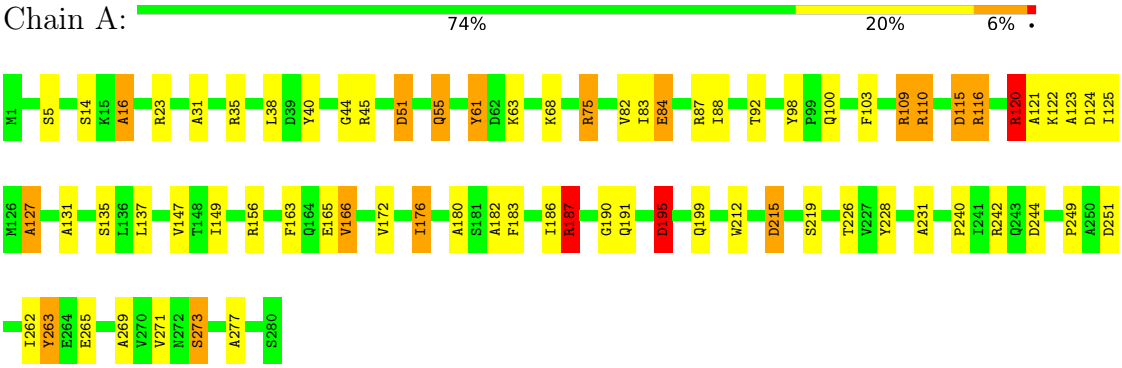
Mol	Chain	Residues	Atoms					AltConf	Trace
2	P	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	Q	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	R	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	S	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	T	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	U	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	V	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	W	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		
2	X	150	Total	C	N	O	S	0	0
			1106	696	186	219	5		

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.

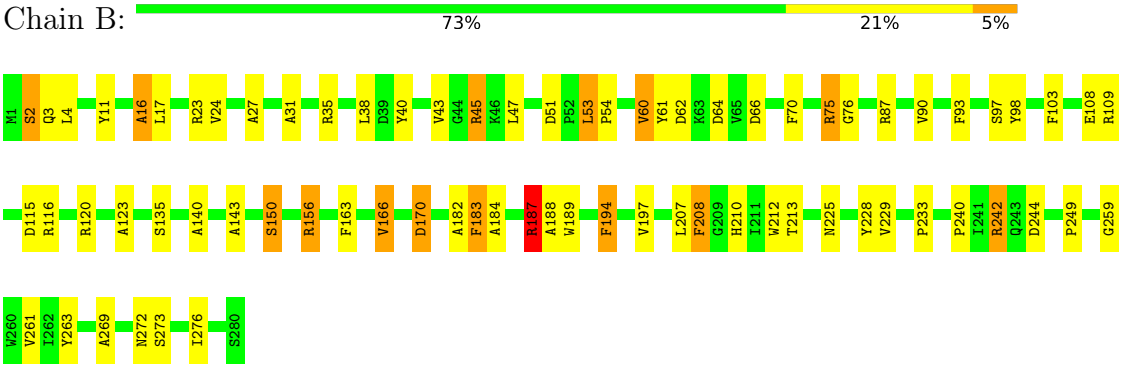
• Molecule 1: Gp27 major capsid protein

Chain A:



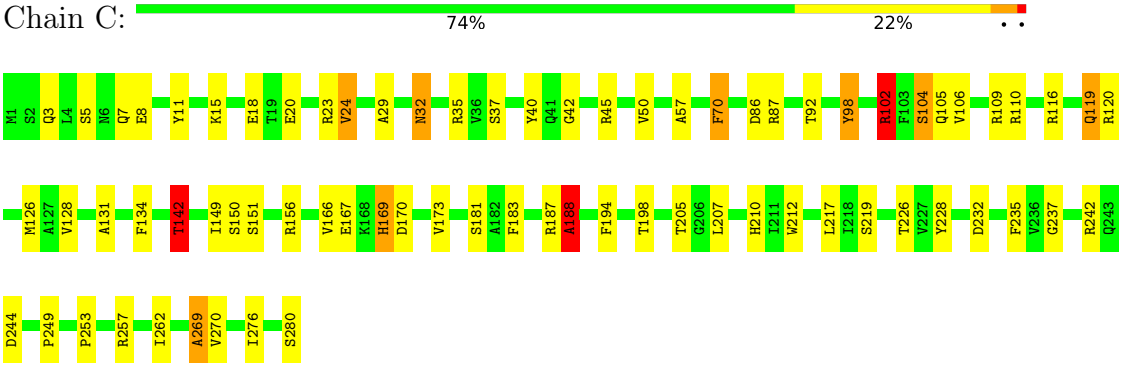
• Molecule 1: Gp27 major capsid protein

Chain B:



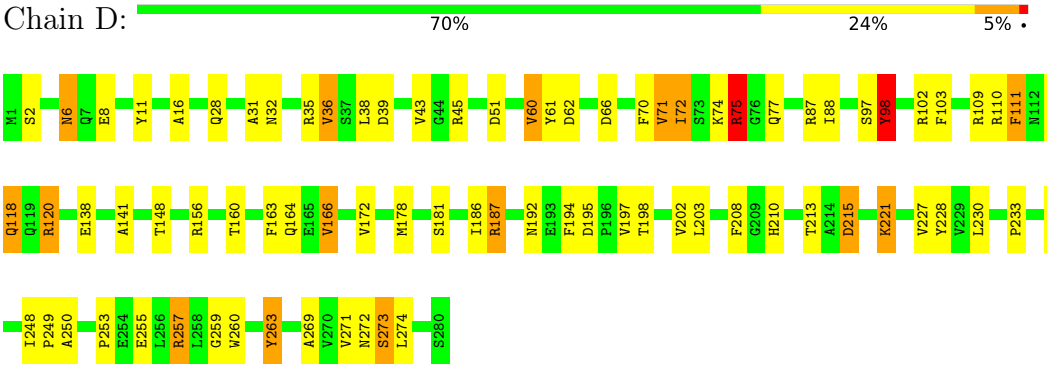
• Molecule 1: Gp27 major capsid protein

Chain C:



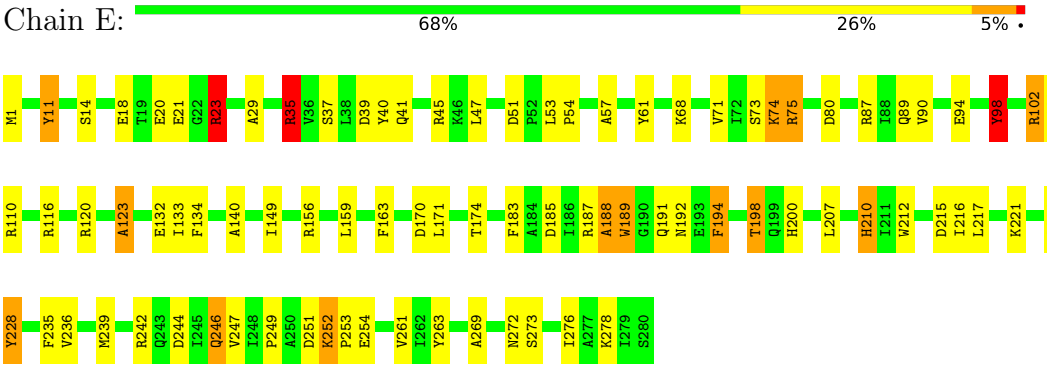
• Molecule 1: Gp27 major capsid protein

Chain D:



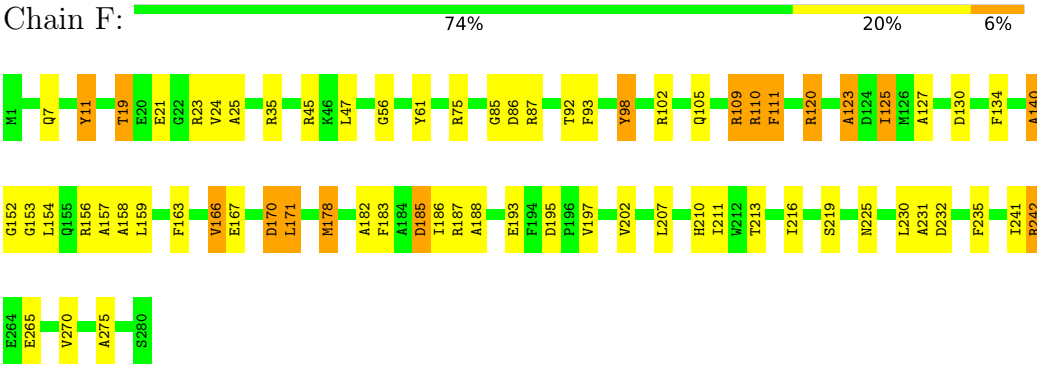
• Molecule 1: Gp27 major capsid protein

Chain E:



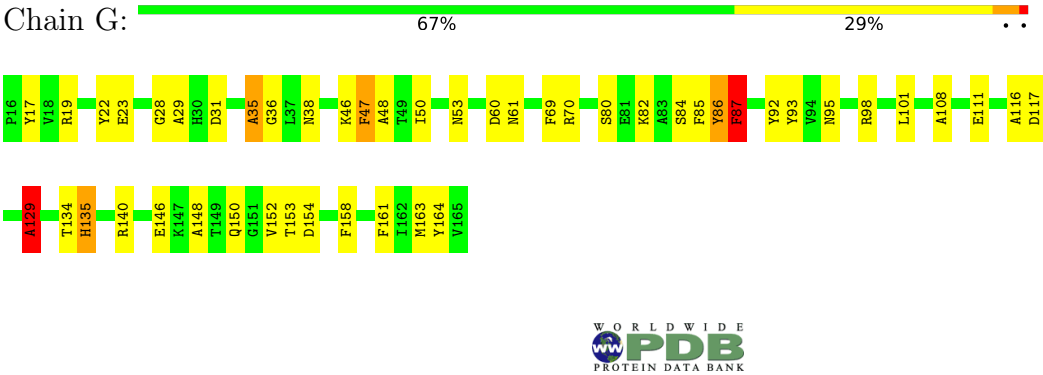
• Molecule 1: Gp27 major capsid protein

Chain F:



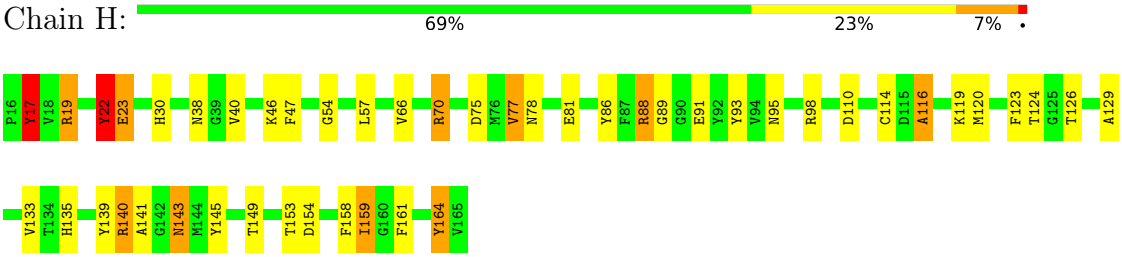
• Molecule 2: Gp26 capsid decoration protein

Chain G:



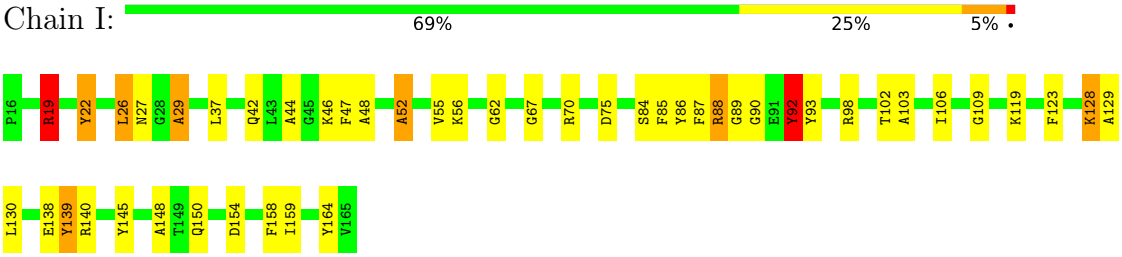
• Molecule 2: Gp26 capsid decoration protein

Chain H:



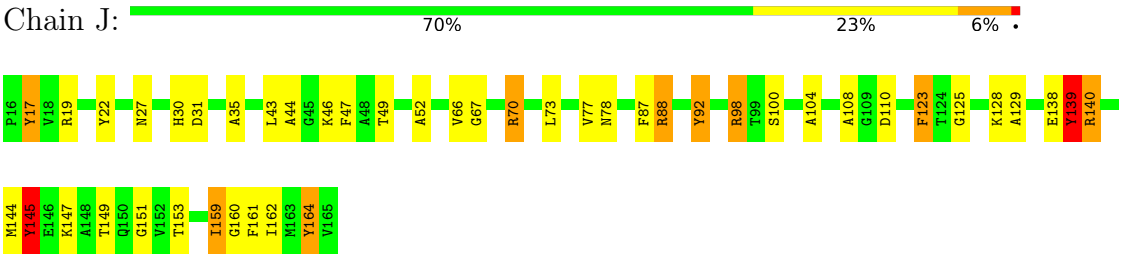
• Molecule 2: Gp26 capsid decoration protein

Chain I:



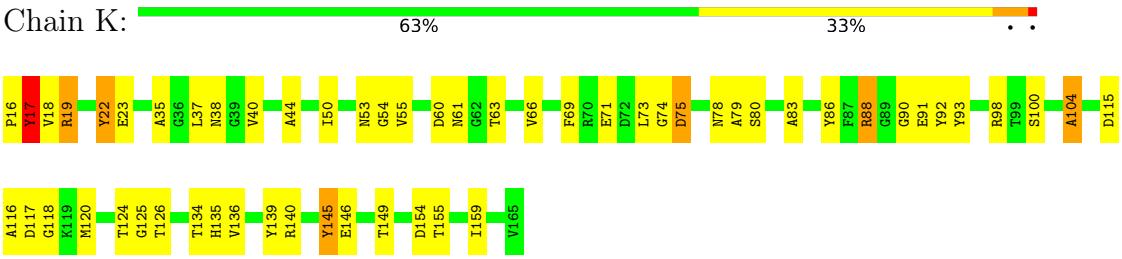
• Molecule 2: Gp26 capsid decoration protein

Chain J:



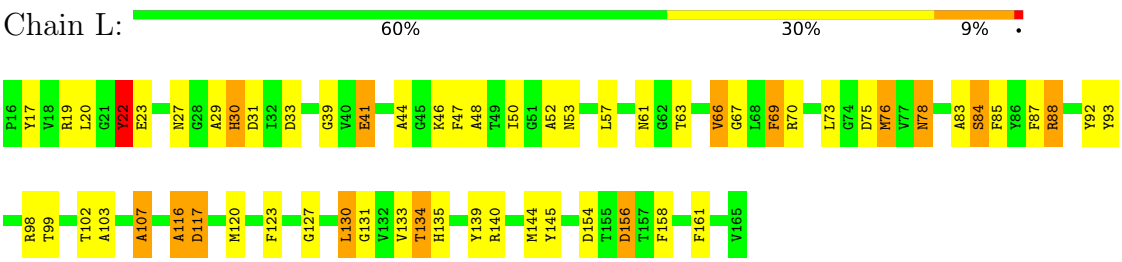
• Molecule 2: Gp26 capsid decoration protein

Chain K:



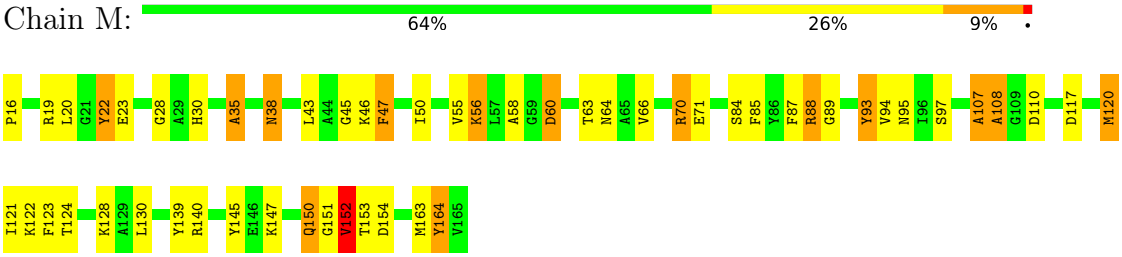
• Molecule 2: Gp26 capsid decoration protein

Chain L:



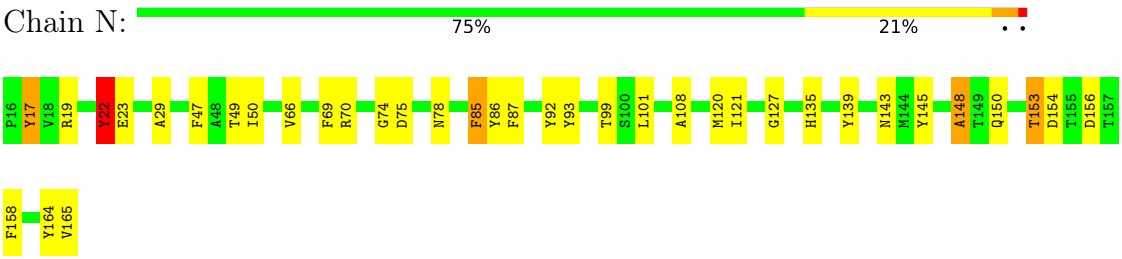
• Molecule 2: Gp26 capsid decoration protein

Chain M:



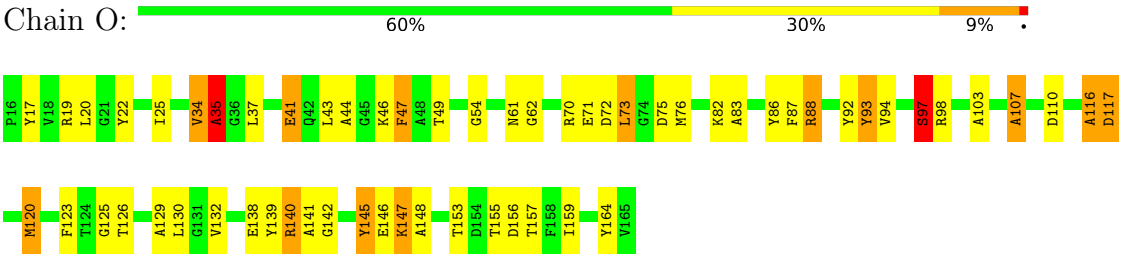
• Molecule 2: Gp26 capsid decoration protein

Chain N:



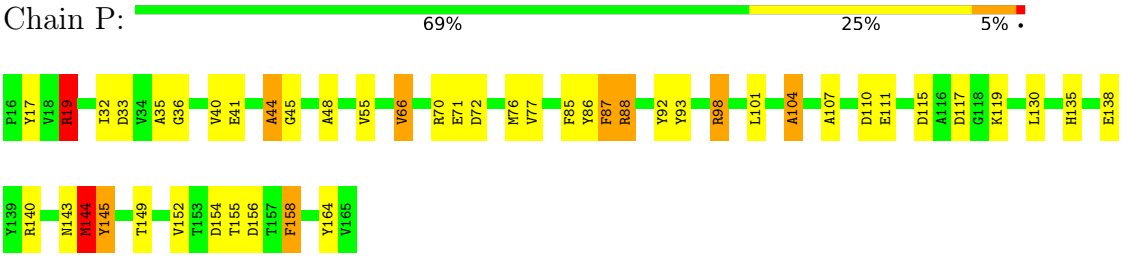
• Molecule 2: Gp26 capsid decoration protein

Chain O:



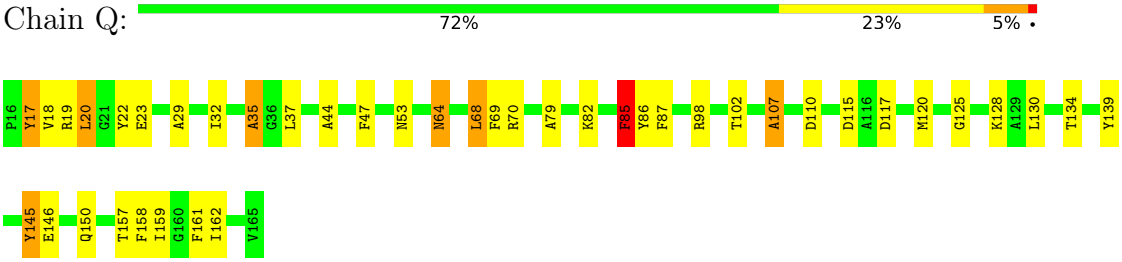
• Molecule 2: Gp26 capsid decoration protein

Chain P:



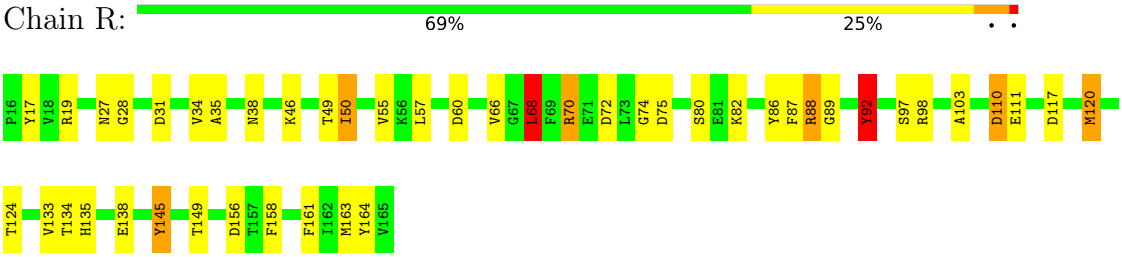
• Molecule 2: Gp26 capsid decoration protein

Chain Q:



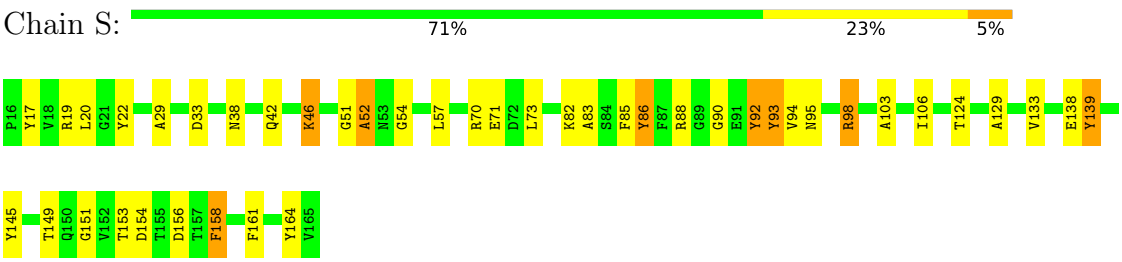
• Molecule 2: Gp26 capsid decoration protein

Chain R:



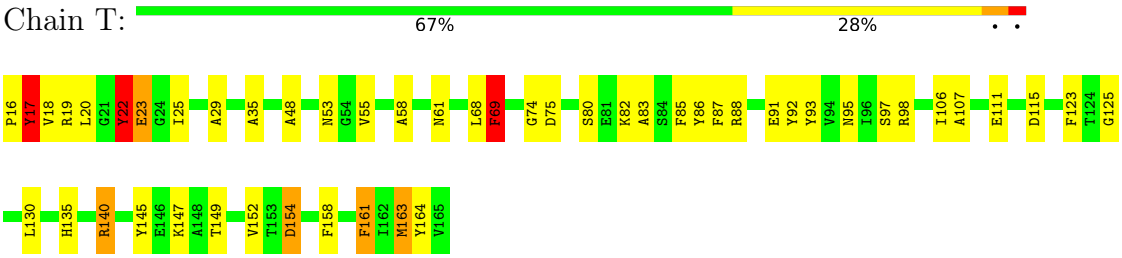
• Molecule 2: Gp26 capsid decoration protein

Chain S:



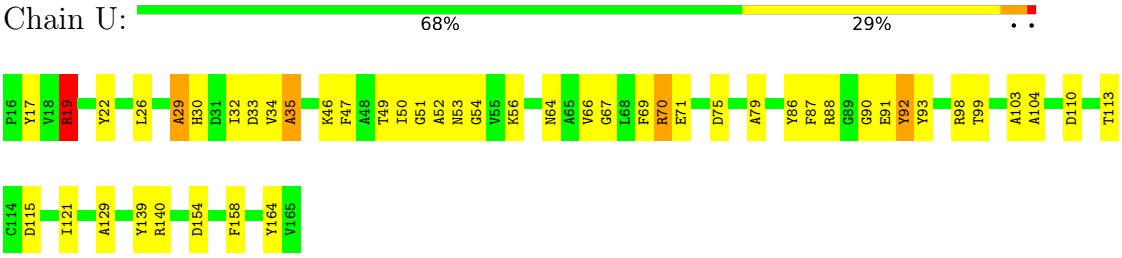
• Molecule 2: Gp26 capsid decoration protein

Chain T:



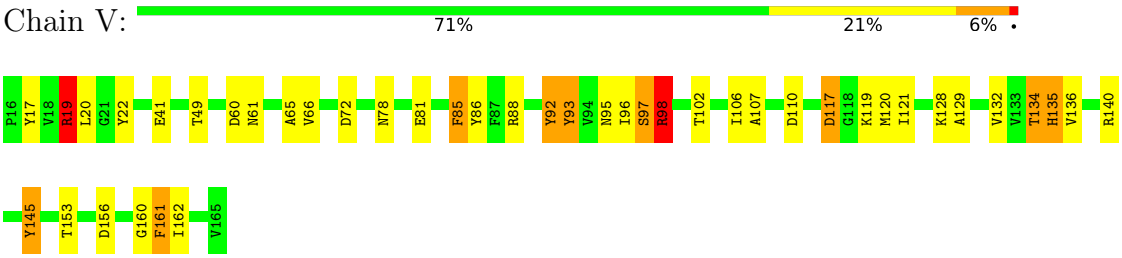
• Molecule 2: Gp26 capsid decoration protein

Chain U:



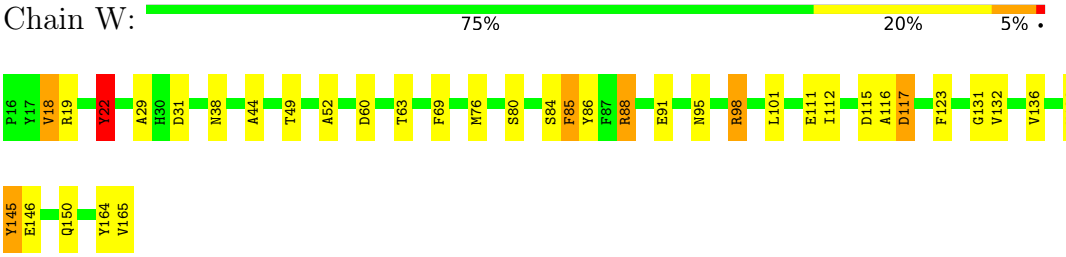
• Molecule 2: Gp26 capsid decoration protein

Chain V:



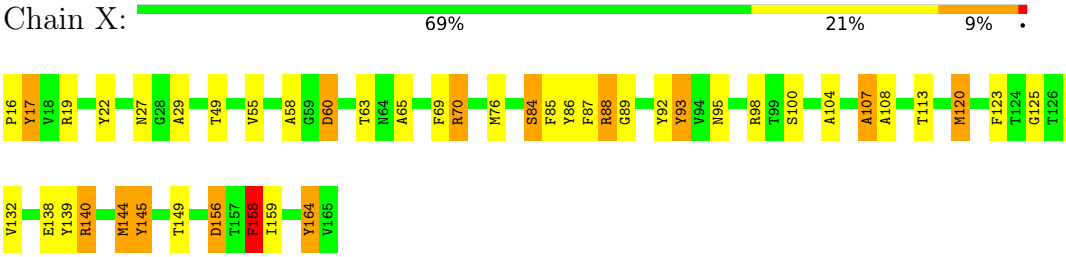
• Molecule 2: Gp26 capsid decoration protein

Chain W:



• Molecule 2: Gp26 capsid decoration protein

Chain X:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	2564	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	14.5	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
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	1.87	13/2219 (0.6%)	2.00	56/3012 (1.9%)
1	B	1.69	15/2219 (0.7%)	2.06	63/3012 (2.1%)
1	C	1.69	12/2219 (0.5%)	2.04	54/3012 (1.8%)
1	D	1.67	11/2219 (0.5%)	1.96	57/3012 (1.9%)
1	E	1.67	17/2219 (0.8%)	1.99	60/3012 (2.0%)
1	F	2.45	19/2219 (0.9%)	1.94	55/3012 (1.8%)
2	G	1.80	13/1123 (1.2%)	1.96	27/1515 (1.8%)
2	H	1.70	7/1123 (0.6%)	2.08	37/1515 (2.4%)
2	I	1.73	10/1123 (0.9%)	2.03	26/1515 (1.7%)
2	J	1.77	9/1123 (0.8%)	2.10	31/1515 (2.0%)
2	K	1.73	4/1123 (0.4%)	2.10	31/1515 (2.0%)
2	L	1.69	8/1123 (0.7%)	2.25	58/1515 (3.8%)
2	M	2.43	14/1123 (1.2%)	2.09	43/1515 (2.8%)
2	N	1.67	4/1123 (0.4%)	2.04	30/1515 (2.0%)
2	O	2.86	11/1123 (1.0%)	2.18	46/1515 (3.0%)
2	P	1.70	9/1123 (0.8%)	2.12	39/1515 (2.6%)
2	Q	1.73	5/1123 (0.4%)	1.99	25/1515 (1.7%)
2	R	1.70	11/1123 (1.0%)	2.02	34/1515 (2.2%)
2	S	1.76	8/1123 (0.7%)	2.10	35/1515 (2.3%)
2	T	1.77	7/1123 (0.6%)	1.98	31/1515 (2.0%)
2	U	1.78	7/1123 (0.6%)	2.05	31/1515 (2.0%)
2	V	1.72	5/1123 (0.4%)	1.92	20/1515 (1.3%)
2	W	1.73	6/1123 (0.5%)	1.96	22/1515 (1.5%)
2	X	1.76	14/1123 (1.2%)	2.10	39/1515 (2.6%)
All	All	1.86	239/33528 (0.7%)	2.04	950/45342 (2.1%)

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	5

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	9
1	C	0	2
1	D	0	11
1	E	0	6
1	F	0	6
2	G	1	6
2	H	3	5
2	I	0	8
2	J	1	7
2	K	3	3
2	L	0	4
2	M	1	3
2	N	3	3
2	O	0	4
2	P	1	4
2	Q	3	3
2	R	0	3
2	S	1	4
2	T	2	7
2	U	0	7
2	V	2	3
2	W	3	3
2	X	0	7
All	All	24	123

The worst 5 of 239 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	O	35	ALA	N-CA	74.41	2.95	1.46
1	F	111	PHE	CG-CD1	43.98	2.04	1.38
1	F	111	PHE	CG-CD2	40.20	1.99	1.38
1	A	55	GLN	CD-NE2	39.08	2.30	1.32
1	F	111	PHE	CE1-CZ	32.88	1.99	1.37

The worst 5 of 950 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	156	ARG	NE-CZ-NH1	18.74	129.67	120.30
1	C	242	ARG	NE-CZ-NH1	17.05	128.82	120.30
2	P	98	ARG	NE-CZ-NH1	16.72	128.66	120.30
1	A	55	GLN	CG-CD-OE1	-16.37	88.87	121.60
1	D	242	ARG	NE-CZ-NH2	16.24	128.42	120.30

5 of 24 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	G	19	ARG	CA
2	H	17	TYR	CA
2	H	22	TYR	CA
2	H	23	GLU	CA
2	J	19	ARG	CA

5 of 123 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	109	ARG	Sidechain
1	A	110	ARG	Sidechain
1	A	116	ARG	Sidechain
1	A	120	ARG	Sidechain
1	A	187	ARG	Sidechain

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	2182	0	2222	27	0
1	B	2182	0	2222	5	0
1	C	2182	0	2222	12	0
1	D	2182	0	2222	8	0
1	E	2182	0	2222	8	0
1	F	2182	0	2222	32	0
2	G	1106	0	1072	2	0
2	H	1106	0	1072	4	0
2	I	1106	0	1072	4	0
2	J	1106	0	1072	5	0
2	K	1106	0	1072	5	0
2	L	1106	0	1072	3	0
2	M	1106	0	1072	33	0
2	N	1106	0	1072	2	0
2	O	1106	0	1072	32	0
2	P	1106	0	1072	1	0
2	Q	1106	0	1072	3	0
2	R	1106	0	1072	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	S	1106	0	1072	4	0
2	T	1106	0	1072	3	0
2	U	1106	0	1072	1	0
2	V	1106	0	1072	4	0
2	W	1106	0	1072	2	0
2	X	1106	0	1072	2	0
All	All	33000	0	32628	144	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 2.

The worst 5 of 144 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)
2:M:47:PHE:CG	2:M:47:PHE:CD1	1.79	1.70
2:M:47:PHE:CE1	2:M:47:PHE:CZ	1.79	1.65
2:M:47:PHE:CZ	2:M:47:PHE:CE2	1.84	1.63
2:M:47:PHE:CG	2:M:47:PHE:CD2	1.85	1.63
1:F:111:PHE:CE2	1:F:111:PHE:CZ	1.91	1.57

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	278/280 (99%)	240 (86%)	20 (7%)	18 (6%)	1 16
1	B	278/280 (99%)	242 (87%)	25 (9%)	11 (4%)	3 23
1	C	278/280 (99%)	244 (88%)	23 (8%)	11 (4%)	3 23
1	D	278/280 (99%)	245 (88%)	19 (7%)	14 (5%)	2 20
1	E	278/280 (99%)	233 (84%)	25 (9%)	20 (7%)	1 14

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	F	278/280 (99%)	248 (89%)	20 (7%)	10 (4%)	3 25
2	G	148/150 (99%)	115 (78%)	20 (14%)	13 (9%)	1 11
2	H	148/150 (99%)	114 (77%)	24 (16%)	10 (7%)	1 15
2	I	148/150 (99%)	118 (80%)	18 (12%)	12 (8%)	1 12
2	J	148/150 (99%)	118 (80%)	21 (14%)	9 (6%)	1 16
2	K	148/150 (99%)	115 (78%)	15 (10%)	18 (12%)	0 5
2	L	148/150 (99%)	119 (80%)	14 (10%)	15 (10%)	0 9
2	M	148/150 (99%)	113 (76%)	22 (15%)	13 (9%)	1 11
2	N	148/150 (99%)	119 (80%)	19 (13%)	10 (7%)	1 15
2	O	148/150 (99%)	114 (77%)	20 (14%)	14 (10%)	0 10
2	P	148/150 (99%)	115 (78%)	23 (16%)	10 (7%)	1 15
2	Q	148/150 (99%)	118 (80%)	15 (10%)	15 (10%)	0 9
2	R	148/150 (99%)	123 (83%)	16 (11%)	9 (6%)	1 16
2	S	148/150 (99%)	120 (81%)	20 (14%)	8 (5%)	2 19
2	T	148/150 (99%)	121 (82%)	14 (10%)	13 (9%)	1 11
2	U	148/150 (99%)	116 (78%)	19 (13%)	13 (9%)	1 11
2	V	148/150 (99%)	117 (79%)	16 (11%)	15 (10%)	0 9
2	W	148/150 (99%)	116 (78%)	19 (13%)	13 (9%)	1 11
2	X	148/150 (99%)	114 (77%)	25 (17%)	9 (6%)	1 16
All	All	4332/4380 (99%)	3557 (82%)	472 (11%)	303 (7%)	2 14

5 of 303 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	SER
1	A	16	ALA
1	A	98	TYR
1	A	122	LYS
1	A	123	ALA

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	235/235 (100%)	222 (94%)	13 (6%)	21 47
1	B	235/235 (100%)	226 (96%)	9 (4%)	33 57
1	C	235/235 (100%)	226 (96%)	9 (4%)	33 57
1	D	235/235 (100%)	221 (94%)	14 (6%)	19 44
1	E	235/235 (100%)	218 (93%)	17 (7%)	14 39
1	F	235/235 (100%)	225 (96%)	10 (4%)	29 53
2	G	109/109 (100%)	105 (96%)	4 (4%)	34 58
2	H	109/109 (100%)	101 (93%)	8 (7%)	14 39
2	I	109/109 (100%)	102 (94%)	7 (6%)	17 42
2	J	109/109 (100%)	101 (93%)	8 (7%)	14 39
2	K	109/109 (100%)	101 (93%)	8 (7%)	14 39
2	L	109/109 (100%)	101 (93%)	8 (7%)	14 39
2	M	109/109 (100%)	97 (89%)	12 (11%)	6 23
2	N	109/109 (100%)	107 (98%)	2 (2%)	59 77
2	O	109/109 (100%)	98 (90%)	11 (10%)	7 25
2	P	109/109 (100%)	103 (94%)	6 (6%)	21 47
2	Q	109/109 (100%)	101 (93%)	8 (7%)	14 39
2	R	109/109 (100%)	103 (94%)	6 (6%)	21 47
2	S	109/109 (100%)	107 (98%)	2 (2%)	59 77
2	T	109/109 (100%)	107 (98%)	2 (2%)	59 77
2	U	109/109 (100%)	105 (96%)	4 (4%)	34 58
2	V	109/109 (100%)	98 (90%)	11 (10%)	7 25
2	W	109/109 (100%)	103 (94%)	6 (6%)	21 47
2	X	109/109 (100%)	102 (94%)	7 (6%)	17 42
All	All	3372/3372 (100%)	3180 (94%)	192 (6%)	24 46

5 of 192 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	M	16	PRO
2	P	85	PHE
2	M	88	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	O	46	LYS
2	Q	130	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
2	K	53	ASN
2	O	38	ASN
2	K	61	ASN
2	L	78	ASN
2	Q	53	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.

6 Map visualisation

This section contains visualisations of the EMDB entry EMD-21695. 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 surface views

This section was not generated.

6.5 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.