wwPDB/EMDataBank EM Map/Model Validation Summary Report (i)

Sep 7, 2019 – 02:15 PM EDT

PDB ID	:	6PPD
EMDB ID:	:	EMD-20433
Title	:	Kaposi's sarcoma-associated herpesvirus (KSHV), C1 penton vertex register, CATC-absent structure
Authors	:	Gong, D.; Dai, X.; Jih, J.; Liu, Y.T.; Bi, G.Q.; Sun, R.; Zhou, Z.H.
Deposited on	:	2019-07-06
Resolution	:	3.70 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
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

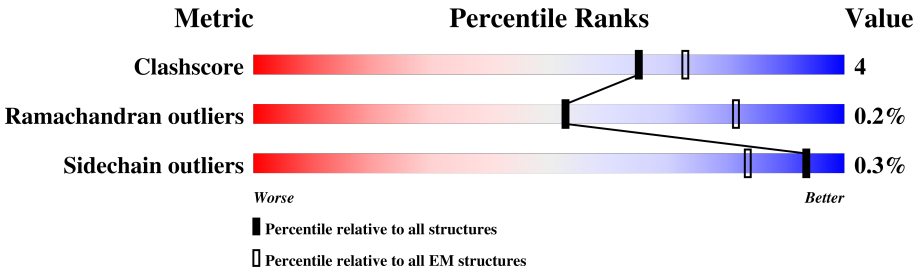
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.4

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 3.70 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	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

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 <=5%

Mol	Chain	Length	Quality of chain
1	0	170	42% · 54%
1	1	170	44% · 54%
1	2	170	42% · 54%
1	3	170	44% · 54%
1	A	170	21% 5% 74%
2	4	1376	73% 15% 11%
2	S	1376	82% 11% 7%
2	T	1376	86% 13% ·
2	W	1376	84% 14% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	X	1376	85% 13%
3	5	331	86% 9% 5%
3	b	331	96%
4	6	305	84% 13%
4	7	305	87% 8% 5%
4	c	305	96%
4	d	305	98%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 68769 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 Small capsomere-interacting protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	0	78	Total	C	N	O	S	0	0
			666	418	130	115	3		
1	2	78	Total	C	N	O	S	0	0
			666	418	130	115	3		
1	3	78	Total	C	N	O	S	0	0
			666	418	130	115	3		
1	A	44	Total	C	N	O	S	0	0
			380	244	72	62	2		
1	1	78	Total	C	N	O	S	0	0
			666	418	130	115	3		

- Molecule 2 is a protein called Major capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	W	1354	Total	C	N	O	S	0	0
			10622	6748	1842	1960	72		
2	S	1281	Total	C	N	O	S	0	0
			10061	6401	1739	1852	69		
2	T	1360	Total	C	N	O	S	0	0
			10667	6776	1851	1967	73		
2	4	1222	Total	C	N	O	S	0	0
			9634	6133	1673	1757	71		
2	X	1341	Total	C	N	O	S	0	0
			10519	6685	1824	1939	71		

- Molecule 3 is a protein called Triplex capsid protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	5	314	Total	C	N	O	S	0	0
			2425	1554	416	441	14		
3	b	321	Total	C	N	O	S	0	0
			2478	1586	424	453	15		

- Molecule 4 is a protein called Triplex capsid protein 2.

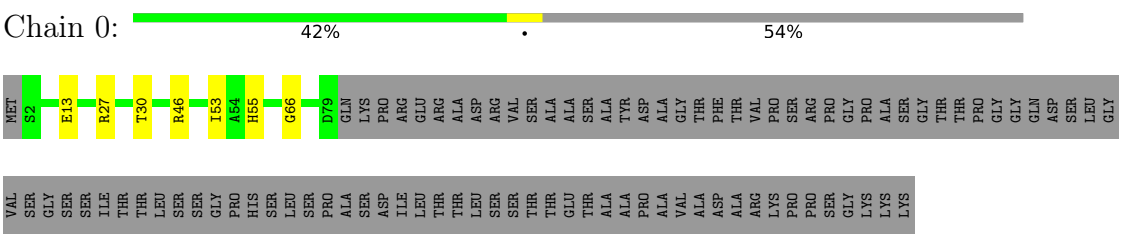
Mol	Chain	Residues	Atoms					AltConf	Trace
4	6	294	Total	C	N	O	S	0	0
			2330	1485	397	434	14		
4	7	291	Total	C	N	O	S	0	0
			2294	1465	388	426	15		
4	c	294	Total	C	N	O	S	0	0
			2330	1485	397	434	14		
4	d	300	Total	C	N	O	S	0	0
			2365	1505	401	444	15		

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 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.

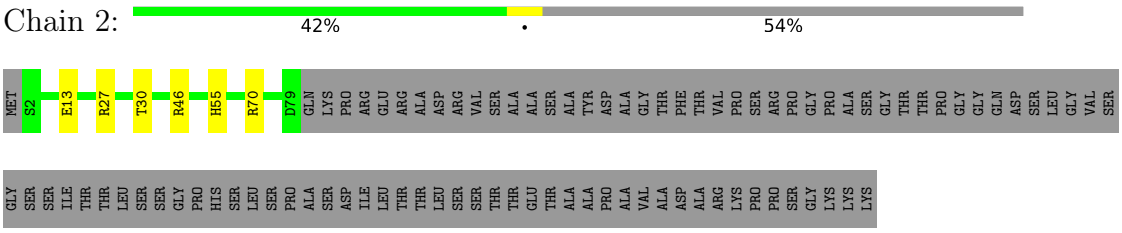
- Molecule 1: Small capsomere-interacting protein

Chain 0:



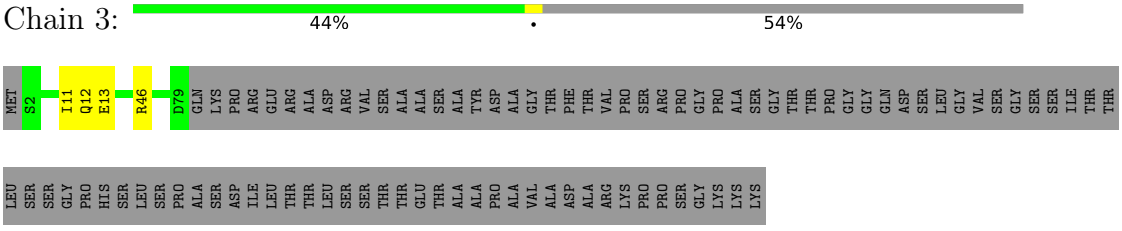
- Molecule 1: Small capsomere-interacting protein

Chain 2:



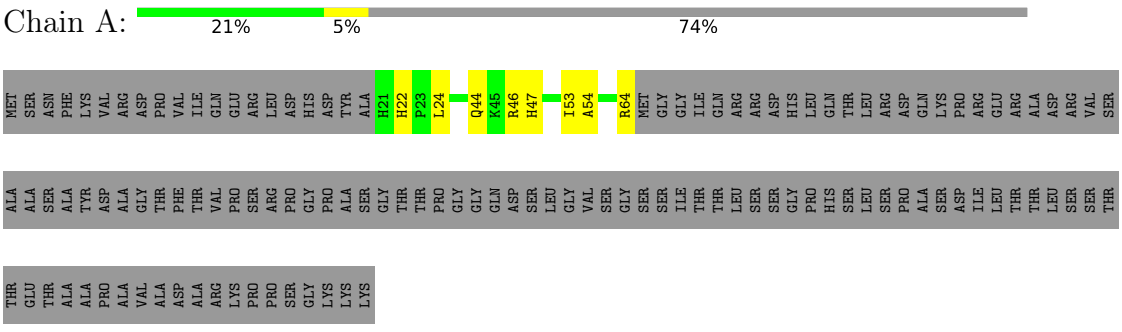
- Molecule 1: Small capsomere-interacting protein

Chain 3:



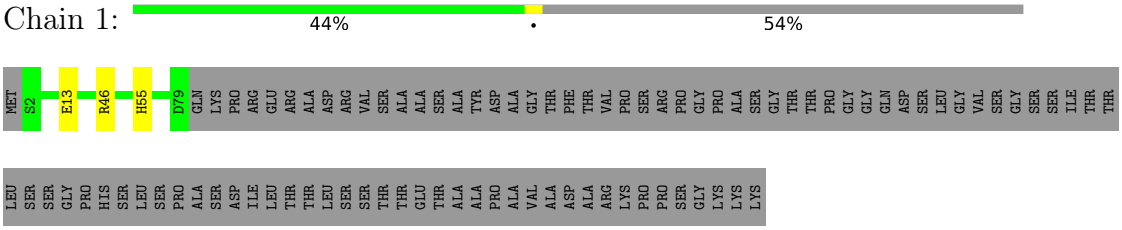
- Molecule 1: Small capsomere-interacting protein

Chain A:



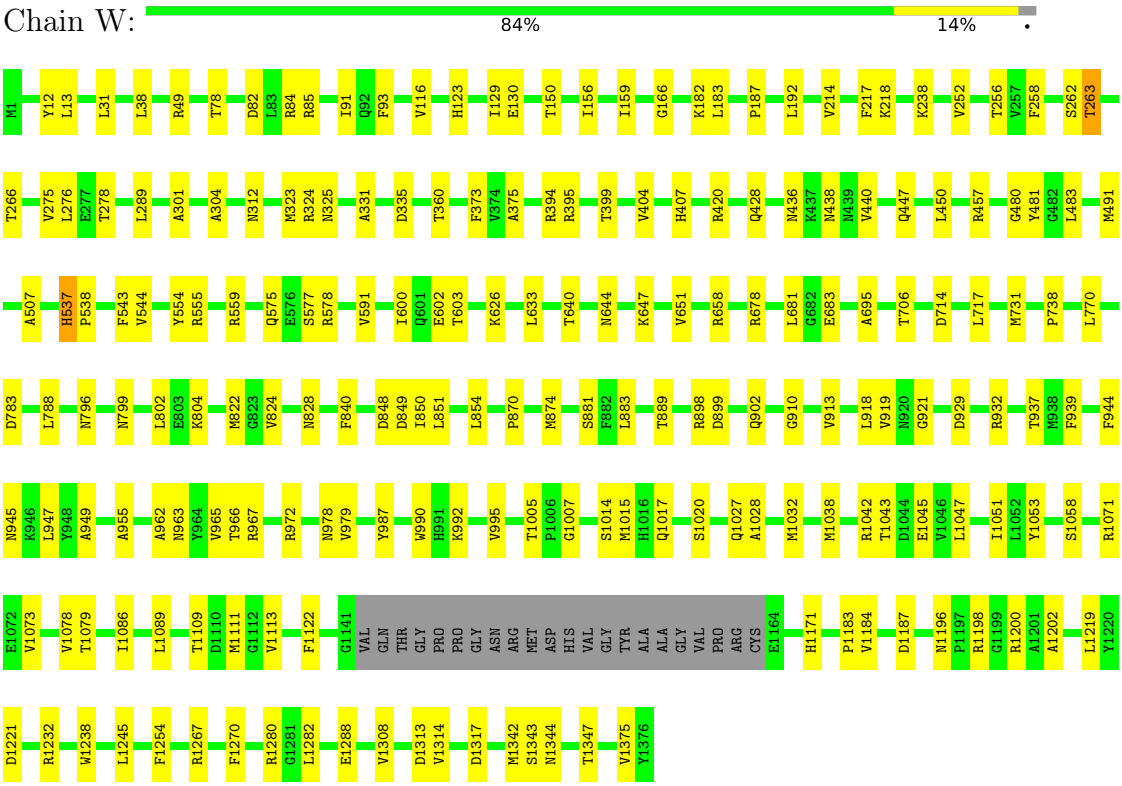
• Molecule 1: Small capsomere-interacting protein

Chain 1:



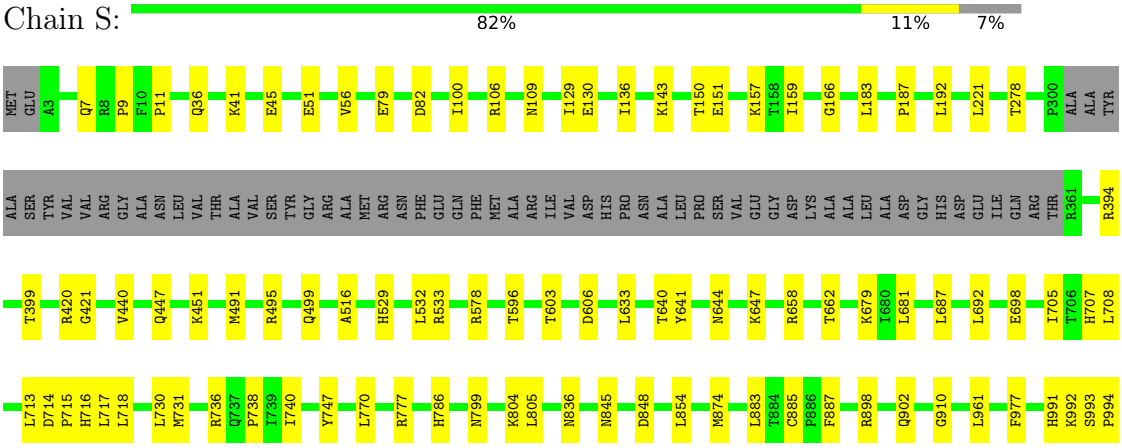
• Molecule 2: Major capsid protein

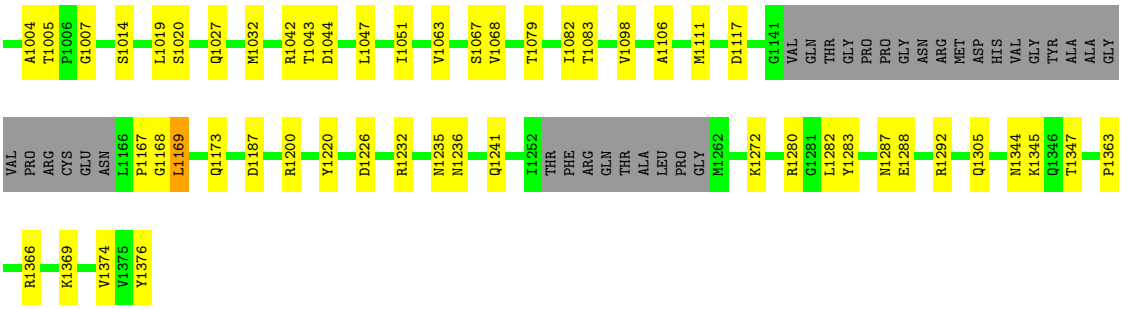
Chain W:



• Molecule 2: Major capsid protein

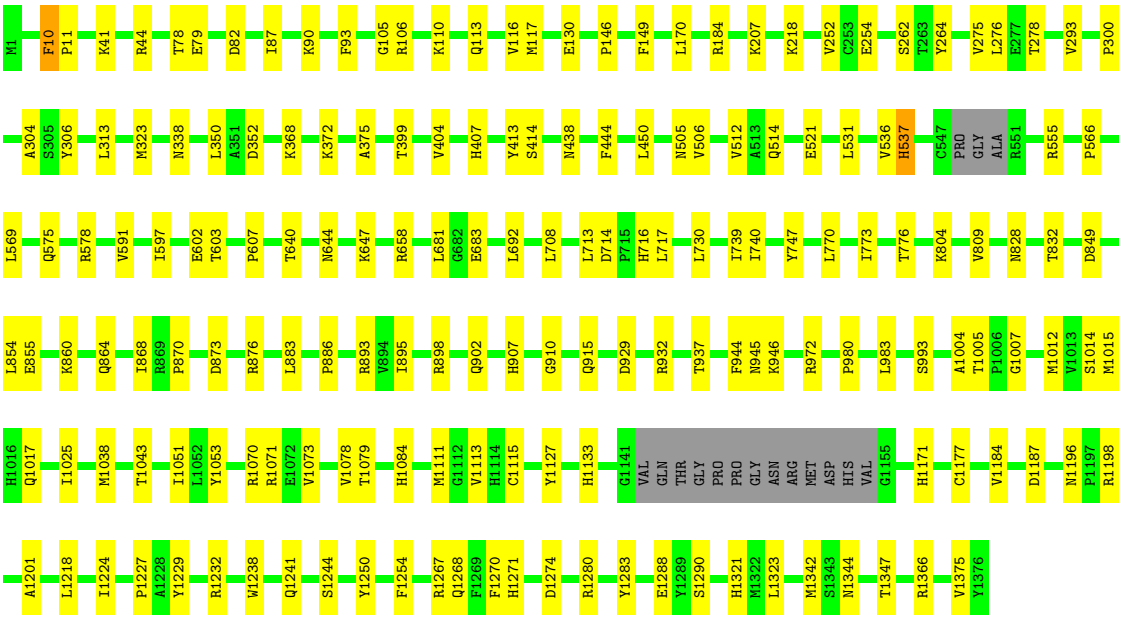
Chain S:





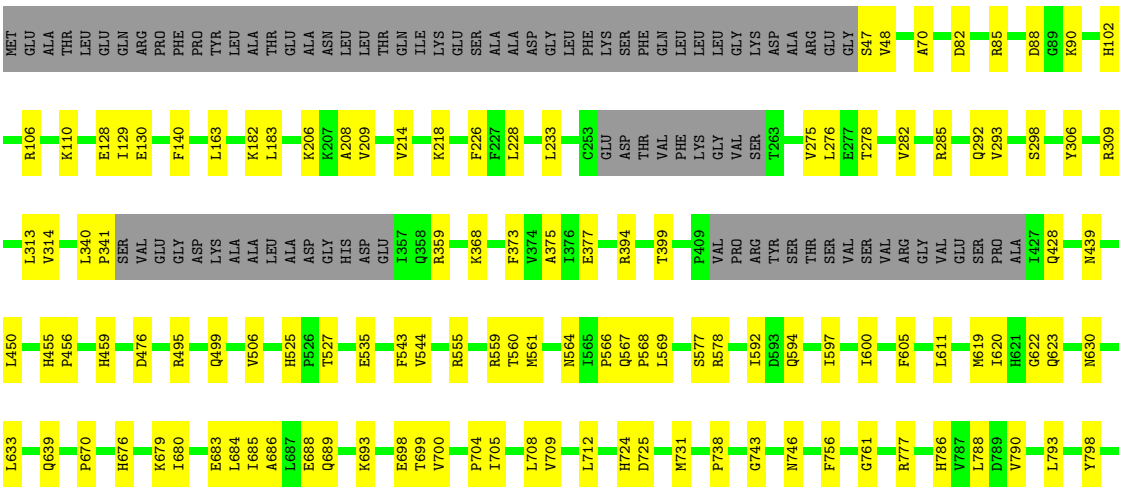
• Molecule 2: Major capsid protein

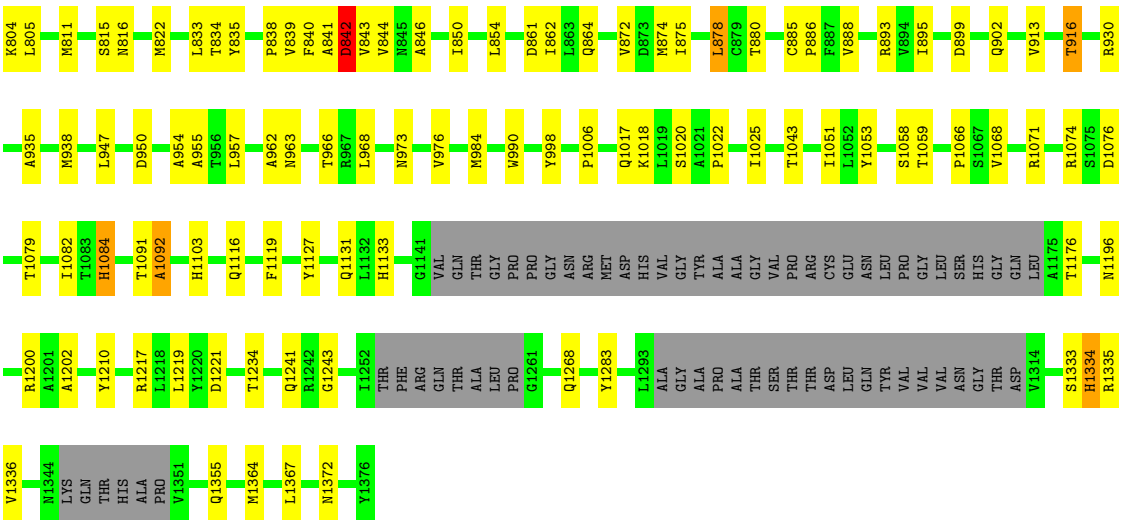
Chain T:



• Molecule 2: Major capsid protein

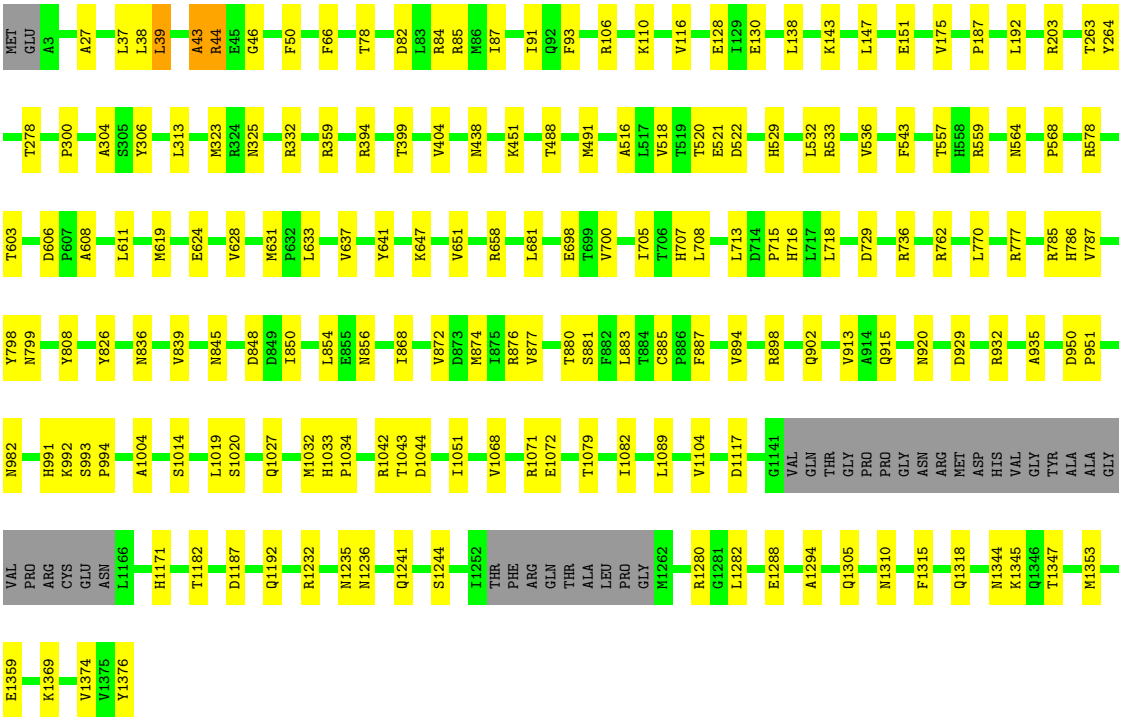
Chain 4:





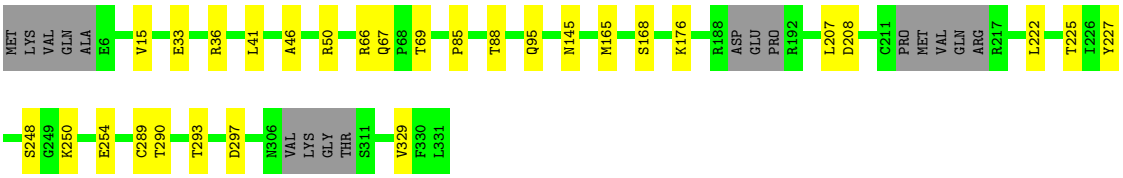
• Molecule 2: Major capsid protein

Chain X:



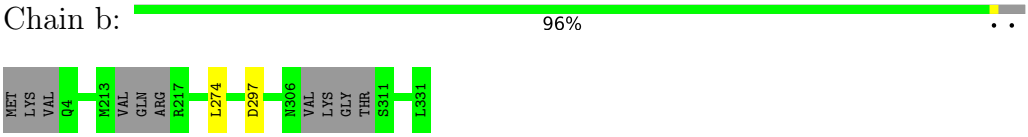
• Molecule 3: Triplex capsid protein 1

Chain 5:



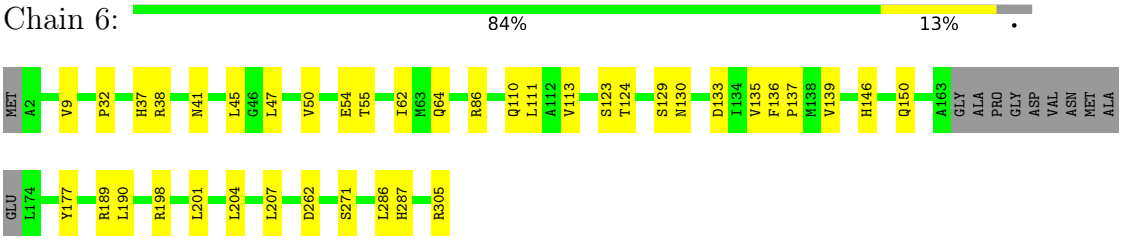
• Molecule 3: Triplex capsid protein 1

Chain b:



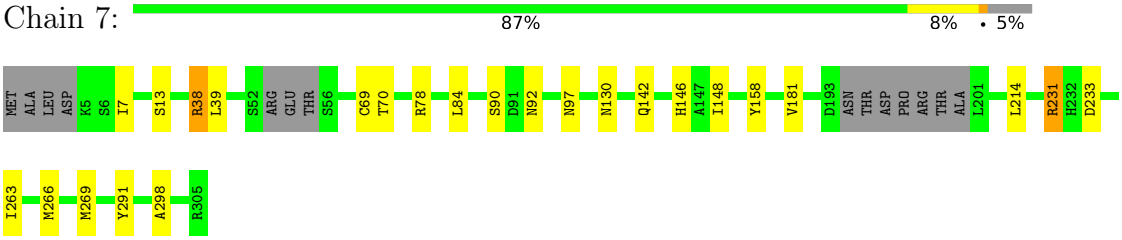
• Molecule 4: Triplex capsid protein 2

Chain 6:



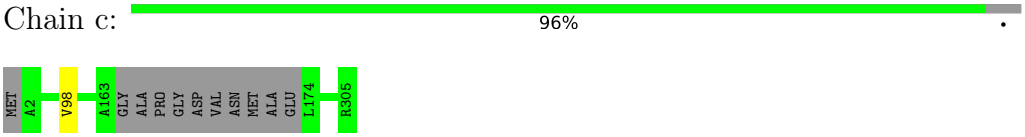
• Molecule 4: Triplex capsid protein 2

Chain 7:



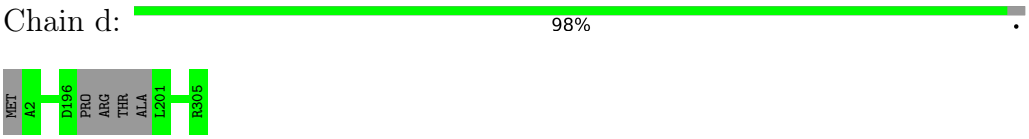
• Molecule 4: Triplex capsid protein 2

Chain c:



• Molecule 4: Triplex capsid protein 2

Chain d:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	1521505	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 ($e^-/Å^2$)	25	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	24271	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

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	0	0.29	0/682	0.49	0/919
1	1	0.29	0/682	0.49	0/919
1	2	0.29	0/682	0.49	0/919
1	3	0.29	0/682	0.49	0/919
1	A	0.24	0/391	0.46	0/528
2	4	0.34	0/9861	0.61	4/13391 (0.0%)
2	S	0.39	0/10300	0.62	5/13998 (0.0%)
2	T	0.42	0/10919	0.64	2/14839 (0.0%)
2	W	0.40	0/10874	0.62	2/14780 (0.0%)
2	X	0.39	0/10768	0.61	6/14635 (0.0%)
3	5	0.34	0/2484	0.58	1/3373 (0.0%)
3	b	0.37	0/2540	0.63	2/3452 (0.1%)
4	6	0.35	0/2376	0.63	1/3234 (0.0%)
4	7	0.31	0/2339	0.58	1/3181 (0.0%)
4	c	0.33	0/2376	0.61	0/3234
4	d	0.35	0/2411	0.60	0/3281
All	All	0.38	0/70367	0.61	24/95602 (0.0%)

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
2	4	0	12
2	S	0	1
2	T	0	7
2	W	0	5
2	X	0	2
All	All	0	27

There are no bond length outliers.

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	S	1004	ALA	C-N-CA	9.52	145.49	121.70
2	X	1004	ALA	C-N-CA	8.30	142.46	121.70
3	b	297	ASP	CB-CG-OD1	7.52	125.07	118.30
2	S	708	LEU	CA-CB-CG	7.13	131.69	115.30
2	X	708	LEU	CA-CB-CG	7.11	131.64	115.30

There are no chirality outliers.

5 of 27 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	W	262	SER	Peptide
2	W	263	THR	Mainchain
2	W	537	HIS	Peptide
2	W	848	ASP	Peptide
2	W	849	ASP	Peptide

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	0	666	0	647	5	0
1	1	666	0	647	2	0
1	2	666	0	647	5	0
1	3	666	0	647	3	0
1	A	380	0	369	6	0
2	4	9634	0	9527	124	0
2	S	10061	0	9955	83	0
2	T	10667	0	10543	101	0
2	W	10622	0	10500	118	0
2	X	10519	0	10398	103	0
3	5	2425	0	2419	20	0
3	b	2478	0	2466	0	0
4	6	2330	0	2354	24	0
4	7	2294	0	2315	15	0
4	c	2330	0	2354	0	0
4	d	2365	0	2379	0	0
All	All	68769	0	68167	546	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 4.

The worst 5 of 546 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:T:536:VAL:HG13	2:T:1244:SER:HA	1.72	0.72
1:2:70:ARG:NH2	1:3:12:GLN:O	2.26	0.68
2:S:770:LEU:HD21	2:S:883:LEU:HD22	1.75	0.67
1:1:55:HIS:HD2	2:X:770:LEU:HD22	1.59	0.67
2:W:799:ASN:HD21	2:W:802:LEU:HD12	1.61	0.66

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	0	76/170 (45%)	74 (97%)	2 (3%)	0	100	100
1	1	76/170 (45%)	74 (97%)	2 (3%)	0	100	100
1	2	76/170 (45%)	74 (97%)	2 (3%)	0	100	100
1	3	76/170 (45%)	74 (97%)	2 (3%)	0	100	100
1	A	42/170 (25%)	42 (100%)	0	0	100	100
2	4	1206/1376 (88%)	1120 (93%)	80 (7%)	6 (0%)	31	71
2	S	1273/1376 (92%)	1188 (93%)	83 (6%)	2 (0%)	49	83
2	T	1354/1376 (98%)	1276 (94%)	74 (6%)	4 (0%)	43	78
2	W	1350/1376 (98%)	1282 (95%)	66 (5%)	2 (0%)	53	87
2	X	1335/1376 (97%)	1246 (93%)	86 (6%)	3 (0%)	49	83
3	5	306/331 (92%)	293 (96%)	13 (4%)	0	100	100
3	b	315/331 (95%)	307 (98%)	8 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	6	290/305 (95%)	273 (94%)	17 (6%)	0	100	100
4	7	285/305 (93%)	268 (94%)	17 (6%)	0	100	100
4	c	290/305 (95%)	277 (96%)	13 (4%)	0	100	100
4	d	296/305 (97%)	292 (99%)	4 (1%)	0	100	100
All	All	8646/9612 (90%)	8160 (94%)	469 (5%)	17 (0%)	53	83

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	W	263	THR
2	4	843	VAL
2	S	992	LYS
2	T	10	PHE
2	X	992	LYS

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	0	70/141 (50%)	70 (100%)	0	100	100
1	1	70/141 (50%)	70 (100%)	0	100	100
1	2	70/141 (50%)	70 (100%)	0	100	100
1	3	70/141 (50%)	70 (100%)	0	100	100
1	A	39/141 (28%)	38 (97%)	1 (3%)	49	77
2	4	1044/1166 (90%)	1041 (100%)	3 (0%)	93	97
2	S	1094/1166 (94%)	1088 (100%)	6 (0%)	90	96
2	T	1154/1166 (99%)	1154 (100%)	0	100	100
2	W	1150/1166 (99%)	1149 (100%)	1 (0%)	94	98
2	X	1139/1166 (98%)	1131 (99%)	8 (1%)	85	93
3	5	266/281 (95%)	266 (100%)	0	100	100
3	b	272/281 (97%)	272 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	6	267/274 (97%)	265 (99%)	2 (1%)	85 93
4	7	262/274 (96%)	258 (98%)	4 (2%)	67 86
4	c	267/274 (97%)	266 (100%)	1 (0%)	92 97
4	d	270/274 (98%)	270 (100%)	0	100 100
All	All	7504/8193 (92%)	7478 (100%)	26 (0%)	93 97

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	6	130	ASN
4	7	78	ARG
2	X	1104	VAL
4	6	198	ARG
4	7	38	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 98 such sidechains are listed below:

Mol	Chain	Res	Type
2	4	498	HIS
2	4	1196	ASN
2	X	716	HIS
2	4	564	ASN
2	4	652	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 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.