



wwPDB EM Validation Summary Report ⓘ

Dec 12, 2022 – 09:18 AM EST

PDB ID : 6W19
EMDB ID : EMD-21504
Title : Structures of Capsid and Capsid-Associated Tegument Complex inside the Epstein-Barr Virus
Authors : Liu, W.; Cui, Y.X.; Wang, C.Y.; Li, Z.H.; Gong, D.Y.; Dai, X.H.; Bi, G.Q.; Sun, R.; Zhou, Z.H.
Deposited on : 2020-03-03
Resolution : 5.50 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 : **FAILED**
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

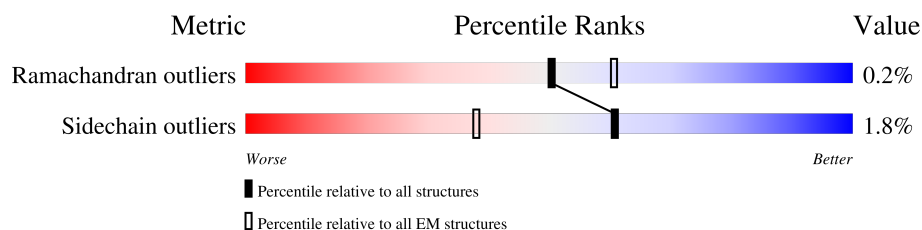
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 5.50 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.



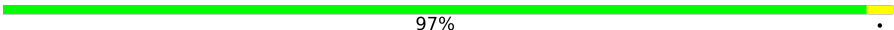
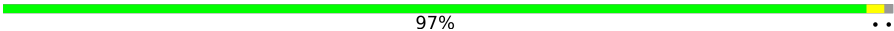
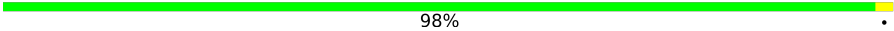
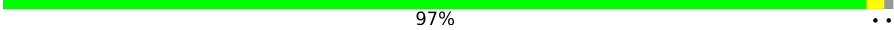
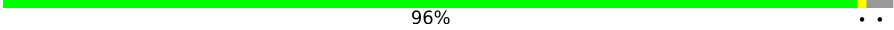






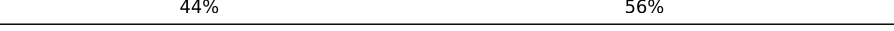







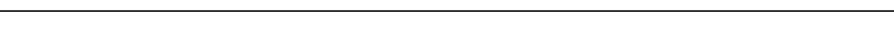

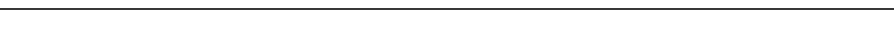
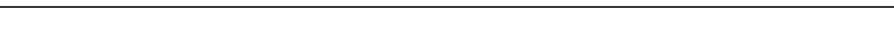


Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	1381	99% .
1	B	1381	98% ..
1	C	1381	97% ..
1	D	1381	99% .
1	E	1381	99% .
1	F	1381	97% ..
1	G	1381	98% .
1	H	1381	99% .
1	I	1381	97% ..
1	J	1381	96% ..



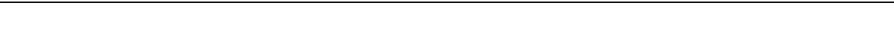

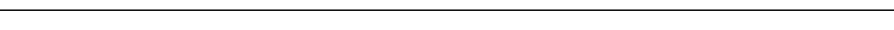
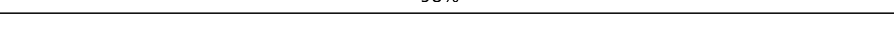
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain	
1	K	1381		97%
1	L	1381		97%
1	M	1381		98%
1	N	1381		97%
1	O	1381		96%
1	P	1381		91%
2	Q	176		44%
2	R	176		43%
2	S	176		43%
2	T	176		44%
2	U	176		43%
2	V	176		44%
2	W	176		43%
2	X	176		44%
2	Y	176		44%
2	Z	176		44%
2	a	176		42%
2	b	176		43%
2	c	176		42%
2	d	176		42%
2	e	176		43%
2	u	176		36%
3	1	364		92%
3	f	364		82%
3	g	364		90%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	h	364	 91% 8%
3	i	364	 91% 8%
3	j	364	 90% 8%
4	2	301	 98% ..
4	3	301	 99% .
4	k	301	 97% ..
4	l	301	 97% ..
4	m	301	 98% ..
4	n	301	 98% ..
4	o	301	 98% ..
4	p	301	 98% ..
4	q	301	 96% ..
4	r	301	 96% ...
4	s	301	 98% ..
4	t	301	 98% ..

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 225047 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 Major capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1381	Total	C	N	O	S	0	0
			10832	6868	1884	2018	62		
1	B	1364	Total	C	N	O	S	0	0
			10701	6788	1859	1993	61		
1	C	1363	Total	C	N	O	S	0	0
			10690	6782	1855	1992	61		
1	D	1381	Total	C	N	O	S	0	0
			10832	6868	1884	2018	62		
1	E	1381	Total	C	N	O	S	0	0
			10832	6868	1884	2018	62		
1	F	1362	Total	C	N	O	S	0	0
			10683	6777	1854	1991	61		
1	G	1381	Total	C	N	O	S	0	0
			10831	6868	1884	2017	62		
1	H	1381	Total	C	N	O	S	0	0
			10832	6868	1884	2018	62		
1	I	1364	Total	C	N	O	S	0	0
			10702	6787	1860	1994	61		
1	J	1348	Total	C	N	O	S	0	0
			10601	6730	1844	1966	61		
1	K	1381	Total	C	N	O	S	0	0
			10832	6868	1884	2018	62		
1	L	1365	Total	C	N	O	S	0	0
			10705	6790	1860	1994	61		
1	M	1381	Total	C	N	O	S	0	0
			10832	6868	1884	2018	62		
1	N	1362	Total	C	N	O	S	0	0
			10683	6777	1854	1991	61		
1	O	1335	Total	C	N	O	S	0	0
			10473	6647	1819	1947	60		
1	P	1292	Total	C	N	O	S	0	0
			10173	6466	1764	1884	59		

- Molecule 2 is a protein called Small capsomere-interacting protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	Q	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	R	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	S	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	T	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	U	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	V	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	W	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	X	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	Y	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	Z	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	a	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	b	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	c	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	d	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	e	77	Total	C	N	O	S	0	0
			649	411	121	115	2		
2	u	63	Total	C	N	O	S	0	0
			528	339	90	98	1		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	f	315	Total	C	N	O	S	0	0
			2474	1586	436	444	8		
3	g	336	Total	C	N	O	S	0	0
			2604	1667	458	471	8		
3	h	336	Total	C	N	O	S	0	0
			2604	1667	458	471	8		
3	i	336	Total	C	N	O	S	0	0
			2604	1667	458	471	8		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
3	j	336	Total	C	N	O	S	0	0
			2604	1667	458	471	8		
3	1	336	Total	C	N	O	S	0	0
			2604	1667	458	471	8		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	k	299	Total	C	N	O	S	0	0
			2338	1500	386	434	18		
4	l	299	Total	C	N	O	S	0	0
			2338	1500	386	434	18		
4	m	299	Total	C	N	O	S	0	0
			2338	1500	386	434	18		
4	n	299	Total	C	N	O	S	0	0
			2338	1500	386	434	18		
4	o	299	Total	C	N	O	S	0	0
			2338	1500	386	434	18		
4	2	299	Total	C	N	O	S	0	0
			2338	1500	386	434	18		
4	p	299	Total	C	N	O	S	0	0
			2338	1500	386	434	18		
4	q	299	Total	C	N	O	S	0	0
			2338	1500	386	434	18		
4	r	299	Total	C	N	O	S	0	0
			2338	1500	386	434	18		
4	s	299	Total	C	N	O	S	0	0
			2338	1500	386	434	18		
4	t	299	Total	C	N	O	S	0	0
			2338	1500	386	434	18		
4	3	299	Total	C	N	O	S	0	0
			2338	1500	386	434	18		

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: Major capsid protein

Chain A:  99%



- Molecule 1: Major capsid protein

Chain B:  98%



- Molecule 1: Major capsid protein

Chain C:  97%



- Molecule 1: Major capsid protein

Chain D:  99%



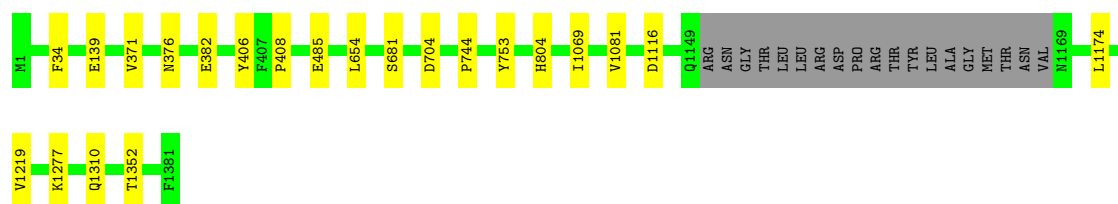
- Molecule 1: Major capsid protein

Chain E:  99%



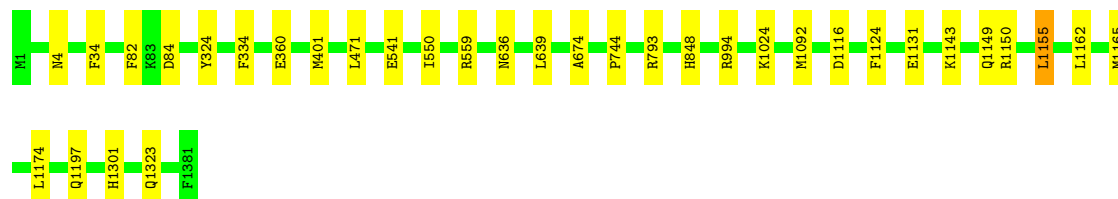
- Molecule 1: Major capsid protein

Chain F:  97%



- Molecule 1: Major capsid protein

Chain G: 98%



- Molecule 1: Major capsid protein

Chain H: 99%



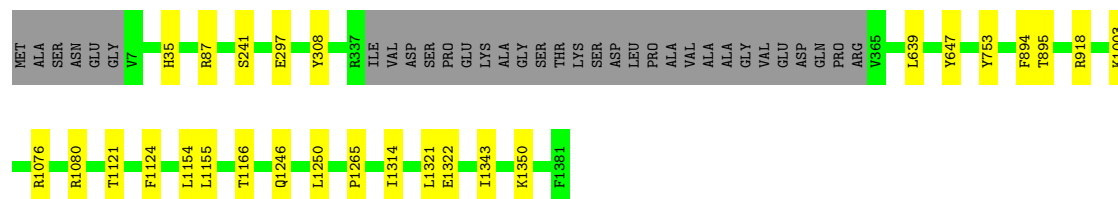
- Molecule 1: Major capsid protein

Chain I: 97%



- Molecule 1: Major capsid protein

Chain J: 96%



- Molecule 1: Major capsid protein

Chain K: 97%





- Molecule 1: Major capsid protein

Chain L: 97%



- Molecule 1: Major capsid protein

Chain M: 98%



- Molecule 1: Major capsid protein

Chain N: 97%



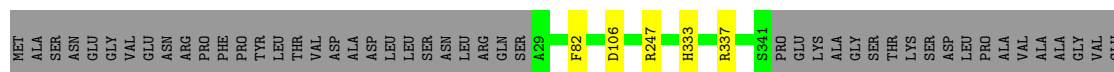
- Molecule 1: Major capsid protein

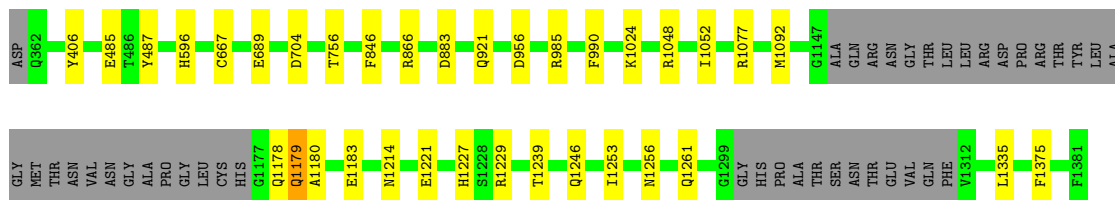
Chain O: 96%



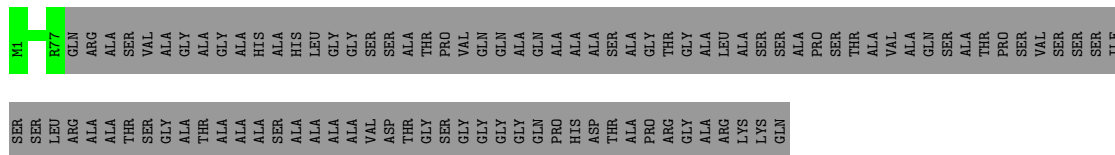
- Molecule 1: Major capsid protein

Chain P: 91% 6%

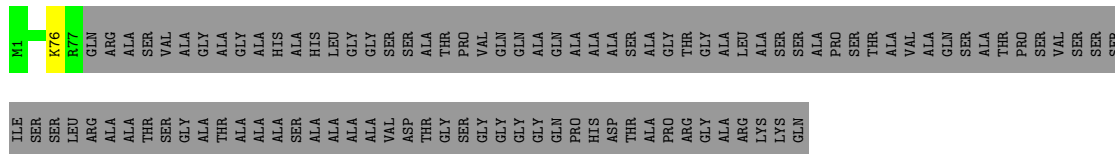




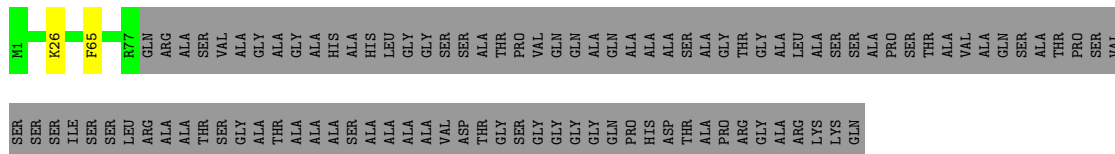
- Molecule 2: Small capsomere-interacting protein



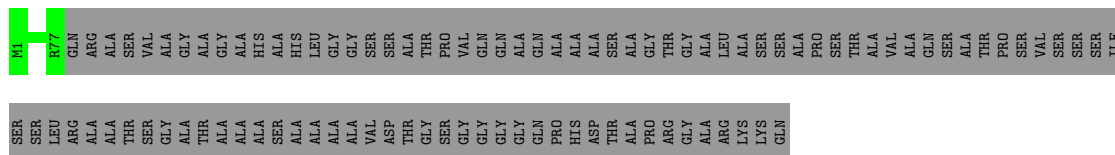
- Molecule 2: Small capsomere-interacting protein



- Molecule 2: Small capsomere-interacting protein



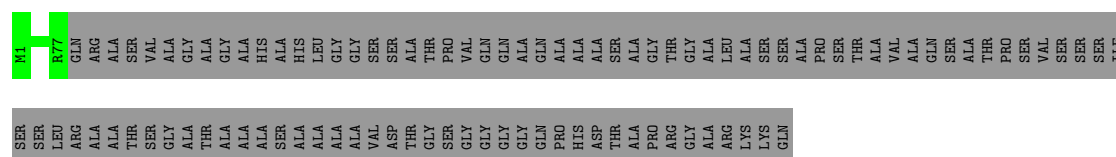
- Molecule 2: Small capsomere-interacting protein



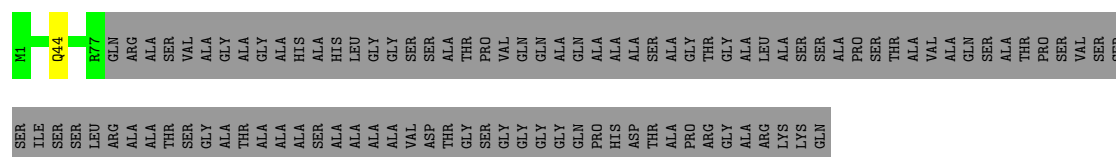
- Molecule 2: Small capsomere-interacting protein



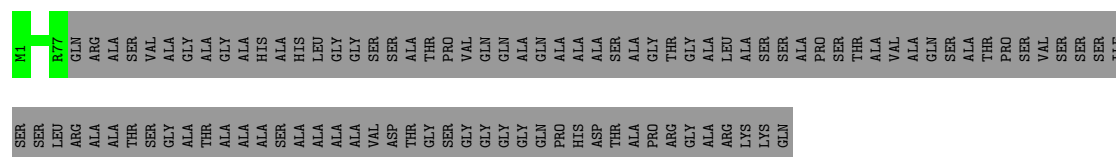
- Molecule 2: Small capsomere-interacting protein



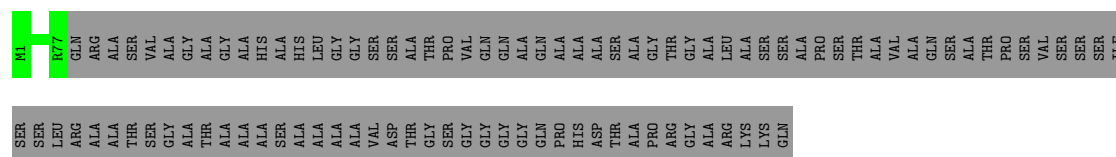
- Molecule 2: Small capsomere-interacting protein



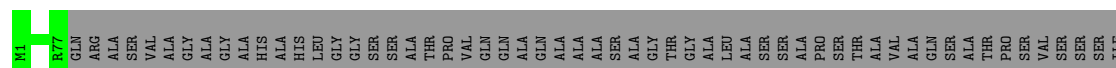
- Molecule 2: Small capsomere-interacting protein



- Molecule 2: Small capsomere-interacting protein



- Molecule 2: Small capsomere-interacting protein



SER SER LEU ARG ALA THR SER GLY THR ALA ALA SER ALA ALA ALA VAL ASP THR GLY SER GLY GLY GLY GLN PRO PRO HIS ASP THR ALA PRO PRO ARG GLY ALA ARG LYS LYS GLN

- Molecule 2: Small capsomere-interacting protein

Chain a:  42% 56%

W1 R13 L14 R69 R77 GLN ARG ALA THR VAL ALA GLY ALA ALA HIS VAL HIS LEU GLY SER SER ALA THR PRO GLN GLN ALA ALA PRO ALA ALA ARG GLY THR GLY ALA LEU ALA SER SER SER SER THR VAL ALA VAL ALA THR SER SER PRO SER

VAL SER SER ILE SER LEU ARG ALA THR SER GLY THR ALA ALA ALA ALA SER ALA THR VAL ASP THR THR GLY GLY GLN GLN GLN GLN HIS THR ALA ARG LYS GLN

- Molecule 2: Small capsomere-interacting protein

Chain b:  43% 56%

W1 K76 GLN ARG ALA SER VAL GLY ALA GLY HIS ALA HIS LEU GLY THR PRO VAL GLN GLN GLN ALA ALA ALA SER GLY THR GLY ALA PRO ALA SER SER SER SER THR VAL ALA GLN ALA THR PRO SER SER

ILE SER SER LEU ARG ALA THR SER GLY THR ALA ALA ALA SER ALA ALA ALA ALA VAL THR ASP THR GLY GLY GLN GLN GLN GLN HIS THR ALA ARG GLY ALA ARG LYS LYS GLN

- Molecule 2: Small capsomere-interacting protein

Chain c:  42% 56%

W1 Q32 N33 N37 R77 GLN ARG ALA THR VAL ALA GLY ALA HIS ALA HIS LEU GLY THR PRO VAL GLN GLN GLN ALA ALA ALA SER GLY THR GLY ALA PRO ALA SER SER SER SER THR VAL ALA GLN ALA THR PRO SER SER

VAL SER SER ILE SER SER LEU ARG ALA THR SER GLY THR ALA ALA ALA ALA SER ALA THR VAL ASP THR THR GLY GLY GLN GLN GLN GLN HIS THR ALA ARG GLY ALA ARG LYS GLN

- Molecule 2: Small capsomere-interacting protein

Chain d:  42% 56%

W1 A2 P19 L24 R77 GLN ARG ALA THR VAL ALA GLY ALA HIS ALA HIS LEU GLY THR PRO VAL GLN GLN GLN ALA ALA ALA SER GLY THR GLY ALA PRO ALA SER SER SER SER THR VAL ALA GLN ALA THR PRO SER SER

VAL SER SER ILE SER SER LEU ARG ALA THR SER GLY THR ALA ALA ALA ALA SER ALA THR VAL ASP THR THR GLY GLY GLN GLN GLN GLN HIS THR ALA ARG GLY ALA ARG LYS GLN

- Molecule 2: Small capsomere-interacting protein

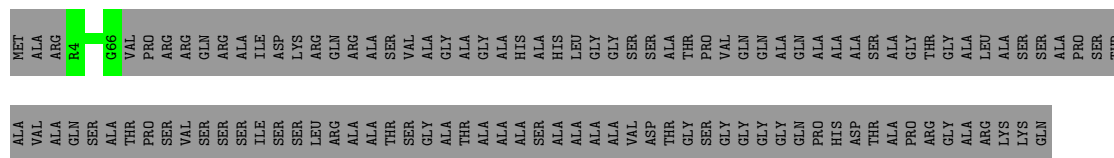
Chain e:  43% 56%

W1 N31 R77 GLN ARG ALA THR VAL ALA GLY ALA HIS ALA HIS LEU GLY THR PRO VAL GLN GLN GLN ALA ALA ALA SER GLY THR GLY ALA PRO ALA SER SER SER SER THR VAL ALA GLN ALA THR PRO SER SER


SER ILE SER SER LEU ARG ALA THR SER GLY THR ALA ALA ALA SER ALA THR VAL ASP THR GLY GLY GLN GLN GLN GLN HIS THR ALA ARG GLY ALA ARG LYS LYS GLN

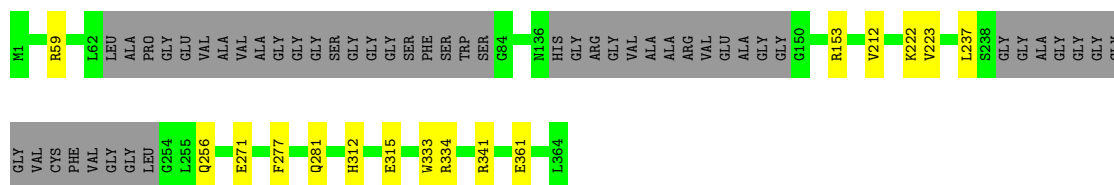
- Molecule 2: Small capsomere-interacting protein

Chain u:  36% 64%

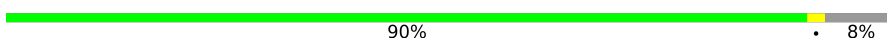


- Molecule 3: Triplex capsid protein 1

Chain f:  82% 13%



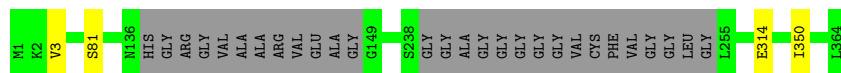
- Molecule 3: Triplex capsid protein 1

Chain g:  90% 8%

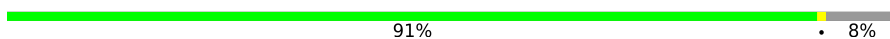


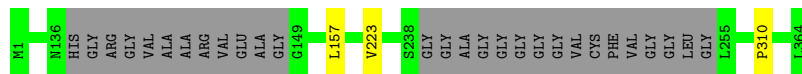
- Molecule 3: Triplex capsid protein 1

Chain h:  91% 8%




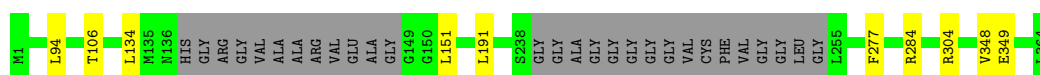
- Molecule 3: Triplex capsid protein 1

Chain i:  91% 8%




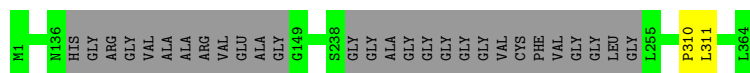
- Molecule 3: Triplex capsid protein 1

Chain j:  90% 8%



- Molecule 3: Triplex capsid protein 1

Chain 1:  92% 8%



- Molecule 4: Triplex capsid protein 2

Chain k:  97% ..



- Molecule 4: Triplex capsid protein 2

Chain l:  97% ..



- Molecule 4: Triplex capsid protein 2

Chain m:  98% ..



- Molecule 4: Triplex capsid protein 2

Chain n:  98% ..



- Molecule 4: Triplex capsid protein 2

Chain o:  98% ..



- Molecule 4: Triplex capsid protein 2

Chain 2:  98% ..



- Molecule 4: Triplex capsid protein 2

Chain p:  98% ..



- Molecule 4: Triplex capsid protein 2

Chain q:  96% ..



- Molecule 4: Triplex capsid protein 2

Chain r:  96% ..



- Molecule 4: Triplex capsid protein 2

Chain s:  98% ..



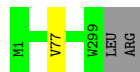
- Molecule 4: Triplex capsid protein 2

Chain t:  98% ..



- Molecule 4: Triplex capsid protein 2

Chain 3:  99% .



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	2048	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 ($e^-/\text{\AA}^2$)	28	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	105000	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	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.41	0/11085	0.49	1/15066 (0.0%)
1	B	0.42	0/10951	0.50	0/14882
1	C	0.42	0/10940	0.50	0/14868
1	D	0.42	0/11085	0.51	0/15066
1	E	0.41	0/11085	0.49	0/15066
1	F	0.40	0/10933	0.50	0/14858
1	G	0.38	0/11084	0.50	1/15065 (0.0%)
1	H	0.37	0/11085	0.46	0/15066
1	I	0.38	0/10952	0.48	0/14883
1	J	0.54	0/10850	0.55	0/14745
1	K	0.53	0/11085	0.53	0/15066
1	L	0.35	0/10955	0.46	0/14887
1	M	0.35	0/11085	0.46	0/15066
1	N	0.37	0/10933	0.49	0/14858
1	O	0.40	0/10719	0.50	0/14565
1	P	0.46	0/10410	0.54	0/14140
2	Q	0.32	0/664	0.44	0/896
2	R	0.35	0/664	0.50	0/896
2	S	0.41	0/664	0.51	0/896
2	T	0.33	0/664	0.43	0/896
2	U	0.36	0/664	0.44	0/896
2	V	0.33	0/664	0.41	0/896
2	W	0.30	0/664	0.43	0/896
2	X	0.32	0/664	0.41	0/896
2	Y	0.31	0/664	0.43	0/896
2	Z	0.39	0/664	0.44	0/896
2	a	0.38	0/664	0.50	0/896
2	b	0.28	0/664	0.41	0/896
2	c	0.30	0/664	0.44	0/896
2	d	0.28	0/664	0.45	0/896
2	e	0.32	0/664	0.47	0/896
2	u	0.32	0/542	0.45	0/735
3	l	0.33	0/2672	0.48	0/3635
3	f	0.43	0/2537	0.50	0/3450

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
3	g	0.35	0/2672	0.46	0/3635
3	h	0.37	0/2672	0.48	0/3635
3	i	0.37	0/2672	0.48	0/3635
3	j	0.35	0/2672	0.48	0/3635
4	2	0.37	0/2388	0.53	1/3254 (0.0%)
4	3	0.34	0/2388	0.49	0/3254
4	k	0.34	0/2388	0.50	0/3254
4	l	0.36	0/2388	0.48	0/3254
4	m	0.33	0/2388	0.51	0/3254
4	n	0.37	0/2388	0.54	0/3254
4	o	0.34	0/2388	0.47	0/3254
4	p	0.32	0/2388	0.51	0/3254
4	q	0.35	0/2388	0.48	0/3254
4	r	0.36	0/2388	0.50	0/3254
4	s	0.37	0/2388	0.51	0/3254
4	t	0.35	0/2388	0.47	0/3254
All	All	0.40	0/230292	0.50	3/312995 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	2	182	LEU	CA-CB-CG	5.74	128.49	115.30
1	G	1155	LEU	CA-CB-CG	5.62	128.22	115.30
1	A	1264	VAL	C-N-CD	5.22	139.36	128.40

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	1379/1381 (100%)	1297 (94%)	78 (6%)	4 (0%)	41	76
1	B	1360/1381 (98%)	1287 (95%)	72 (5%)	1 (0%)	51	85
1	C	1359/1381 (98%)	1277 (94%)	80 (6%)	2 (0%)	51	85
1	D	1379/1381 (100%)	1301 (94%)	78 (6%)	0	100	100
1	E	1379/1381 (100%)	1286 (93%)	91 (7%)	2 (0%)	51	85
1	F	1358/1381 (98%)	1272 (94%)	84 (6%)	2 (0%)	51	85
1	G	1379/1381 (100%)	1296 (94%)	80 (6%)	3 (0%)	47	81
1	H	1379/1381 (100%)	1301 (94%)	75 (5%)	3 (0%)	47	81
1	I	1360/1381 (98%)	1288 (95%)	70 (5%)	2 (0%)	51	85
1	J	1344/1381 (97%)	1238 (92%)	104 (8%)	2 (0%)	51	85
1	K	1379/1381 (100%)	1268 (92%)	110 (8%)	1 (0%)	51	85
1	L	1361/1381 (99%)	1299 (95%)	60 (4%)	2 (0%)	51	85
1	M	1379/1381 (100%)	1299 (94%)	79 (6%)	1 (0%)	51	85
1	N	1358/1381 (98%)	1312 (97%)	42 (3%)	4 (0%)	41	76
1	O	1331/1381 (96%)	1274 (96%)	52 (4%)	5 (0%)	34	72
1	P	1284/1381 (93%)	1183 (92%)	97 (8%)	4 (0%)	41	76
2	Q	75/176 (43%)	68 (91%)	7 (9%)	0	100	100
2	R	75/176 (43%)	71 (95%)	4 (5%)	0	100	100
2	S	75/176 (43%)	68 (91%)	6 (8%)	1 (1%)	12	47
2	T	75/176 (43%)	71 (95%)	4 (5%)	0	100	100
2	U	75/176 (43%)	69 (92%)	6 (8%)	0	100	100
2	V	75/176 (43%)	71 (95%)	4 (5%)	0	100	100
2	W	75/176 (43%)	71 (95%)	4 (5%)	0	100	100
2	X	75/176 (43%)	73 (97%)	2 (3%)	0	100	100
2	Y	75/176 (43%)	72 (96%)	3 (4%)	0	100	100
2	Z	75/176 (43%)	69 (92%)	6 (8%)	0	100	100
2	a	75/176 (43%)	70 (93%)	5 (7%)	0	100	100
2	b	75/176 (43%)	69 (92%)	6 (8%)	0	100	100
2	c	75/176 (43%)	71 (95%)	4 (5%)	0	100	100
2	d	75/176 (43%)	71 (95%)	2 (3%)	2 (3%)	5	31

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	e	75/176 (43%)	75 (100%)	0	0	100	100
2	u	61/176 (35%)	56 (92%)	5 (8%)	0	100	100
3	1	330/364 (91%)	303 (92%)	25 (8%)	2 (1%)	25	65
3	f	307/364 (84%)	280 (91%)	27 (9%)	0	100	100
3	g	330/364 (91%)	308 (93%)	22 (7%)	0	100	100
3	h	330/364 (91%)	318 (96%)	10 (3%)	2 (1%)	25	65
3	i	330/364 (91%)	306 (93%)	23 (7%)	1 (0%)	41	76
3	j	330/364 (91%)	299 (91%)	31 (9%)	0	100	100
4	2	297/301 (99%)	253 (85%)	42 (14%)	2 (1%)	22	62
4	3	297/301 (99%)	279 (94%)	17 (6%)	1 (0%)	41	76
4	k	297/301 (99%)	284 (96%)	12 (4%)	1 (0%)	41	76
4	l	297/301 (99%)	286 (96%)	11 (4%)	0	100	100
4	m	297/301 (99%)	287 (97%)	10 (3%)	0	100	100
4	n	297/301 (99%)	277 (93%)	20 (7%)	0	100	100
4	o	297/301 (99%)	275 (93%)	22 (7%)	0	100	100
4	p	297/301 (99%)	293 (99%)	3 (1%)	1 (0%)	41	76
4	q	297/301 (99%)	283 (95%)	14 (5%)	0	100	100
4	r	297/301 (99%)	279 (94%)	15 (5%)	3 (1%)	15	53
4	s	297/301 (99%)	283 (95%)	12 (4%)	2 (1%)	22	62
4	t	297/301 (99%)	284 (96%)	12 (4%)	1 (0%)	41	76
All	All	28475/30708 (93%)	26770 (94%)	1648 (6%)	57 (0%)	50	81

5 of 57 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	978	ARG
1	K	1011	SER
1	O	1284	ASN
1	P	1179	GLN
2	d	19	PRO

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	1171/1171 (100%)	1160 (99%)	11 (1%)	78	88
1	B	1157/1171 (99%)	1142 (99%)	15 (1%)	69	82
1	C	1156/1171 (99%)	1140 (99%)	16 (1%)	67	80
1	D	1171/1171 (100%)	1155 (99%)	16 (1%)	67	80
1	E	1171/1171 (100%)	1156 (99%)	15 (1%)	69	82
1	F	1155/1171 (99%)	1135 (98%)	20 (2%)	60	78
1	G	1170/1171 (100%)	1139 (97%)	31 (3%)	46	67
1	H	1171/1171 (100%)	1154 (98%)	17 (2%)	65	80
1	I	1157/1171 (99%)	1141 (99%)	16 (1%)	67	80
1	J	1146/1171 (98%)	1121 (98%)	25 (2%)	52	71
1	K	1171/1171 (100%)	1126 (96%)	45 (4%)	33	57
1	L	1157/1171 (99%)	1138 (98%)	19 (2%)	62	79
1	M	1171/1171 (100%)	1148 (98%)	23 (2%)	55	73
1	N	1155/1171 (99%)	1136 (98%)	19 (2%)	62	79
1	O	1131/1171 (97%)	1120 (99%)	11 (1%)	76	86
1	P	1100/1171 (94%)	1063 (97%)	37 (3%)	37	60
2	Q	71/128 (56%)	71 (100%)	0	100	100
2	R	71/128 (56%)	70 (99%)	1 (1%)	67	80
2	S	71/128 (56%)	70 (99%)	1 (1%)	67	80
2	T	71/128 (56%)	71 (100%)	0	100	100
2	U	71/128 (56%)	70 (99%)	1 (1%)	67	80
2	V	71/128 (56%)	71 (100%)	0	100	100
2	W	71/128 (56%)	70 (99%)	1 (1%)	67	80
2	X	71/128 (56%)	71 (100%)	0	100	100
2	Y	71/128 (56%)	71 (100%)	0	100	100
2	Z	71/128 (56%)	71 (100%)	0	100	100
2	a	71/128 (56%)	68 (96%)	3 (4%)	30	54
2	b	71/128 (56%)	70 (99%)	1 (1%)	67	80
2	c	71/128 (56%)	68 (96%)	3 (4%)	30	54

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	d	71/128 (56%)	70 (99%)	1 (1%)	67	80
2	e	71/128 (56%)	70 (99%)	1 (1%)	67	80
2	u	59/128 (46%)	59 (100%)	0	100	100
3	1	278/289 (96%)	278 (100%)	0	100	100
3	f	267/289 (92%)	251 (94%)	16 (6%)	19	46
3	g	278/289 (96%)	270 (97%)	8 (3%)	42	64
3	h	278/289 (96%)	276 (99%)	2 (1%)	84	90
3	i	278/289 (96%)	276 (99%)	2 (1%)	84	90
3	j	278/289 (96%)	268 (96%)	10 (4%)	35	59
4	2	265/267 (99%)	263 (99%)	2 (1%)	81	89
4	3	265/267 (99%)	265 (100%)	0	100	100
4	k	265/267 (99%)	258 (97%)	7 (3%)	46	67
4	l	265/267 (99%)	257 (97%)	8 (3%)	41	63
4	m	265/267 (99%)	260 (98%)	5 (2%)	57	75
4	n	265/267 (99%)	261 (98%)	4 (2%)	65	80
4	o	265/267 (99%)	260 (98%)	5 (2%)	57	75
4	p	265/267 (99%)	261 (98%)	4 (2%)	65	80
4	q	265/267 (99%)	255 (96%)	10 (4%)	33	57
4	r	265/267 (99%)	257 (97%)	8 (3%)	41	63
4	s	265/267 (99%)	262 (99%)	3 (1%)	73	84
4	t	265/267 (99%)	262 (99%)	3 (1%)	73	84
All	All	24471/25722 (95%)	24025 (98%)	446 (2%)	61	77

5 of 446 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	113	GLN
4	t	49	TYR
1	N	1295	SER
4	r	204	ASN
4	l	171	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 608 such sidechains are listed below:

Mol	Chain	Res	Type
1	O	1056	ASN
4	2	38	GLN
1	P	454	ASN
1	O	997	HIS
2	b	44	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 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.