



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 12:35 am BST

PDB ID : 4YVW
Title : crystal structure of an enterovirus 71/coxsackievirus A16 chimeric virus-like particle
Authors : Chen, R.; Lyu, K.
Deposited on : 2015-03-20
Resolution : 3.80 Å(reported)

This is a wwPDB X-ray Structure 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/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

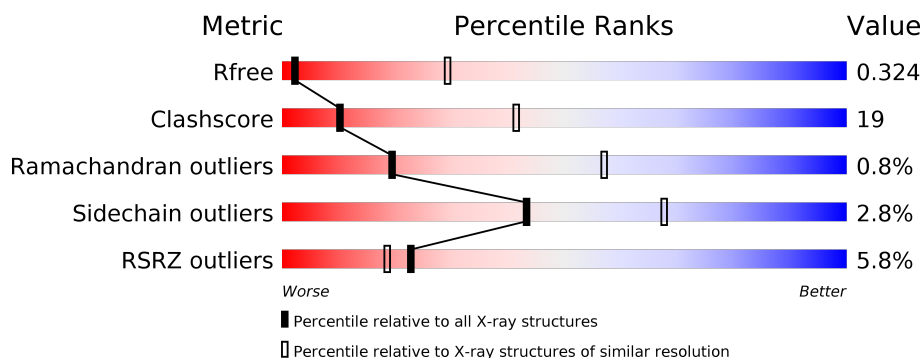
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.80 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1212 (4.00-3.60)
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)
RSRZ outliers	127900	1121 (4.00-3.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	297	<div> <div>4%</div> <div> <div></div> <div>50%</div> <div>25%</div> <div>24%</div> </div> </div>
1	D	297	<div> <div>4%</div> <div> <div></div> <div>47%</div> <div>28%</div> <div>•</div> <div>24%</div> </div> </div>
1	E	297	<div> <div>3%</div> <div> <div></div> <div>48%</div> <div>27%</div> <div>•</div> <div>24%</div> </div> </div>
1	J	297	<div> <div>4%</div> <div> <div></div> <div>48%</div> <div>27%</div> <div></div> <div>24%</div> </div> </div>
1	M	297	<div> <div>4%</div> <div> <div></div> <div>44%</div> <div>30%</div> <div>•</div> <div>24%</div> </div> </div>
2	B	242	<div> <div>%</div> <div> <div></div> <div>58%</div> <div>33%</div> <div>•</div> <div>7%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	242	<p>2% 49% 43% 7%</p>
2	H	242	<p>1% 56% 36% 7%</p>
2	K	242	<p>0% 59% 33% 7%</p>
2	N	242	<p>0% 58% 35% 7%</p>
3	C	323	<p>9% 47% 24% 27%</p>
3	G	323	<p>7% 45% 26% 27%</p>
3	I	323	<p>9% 42% 29% 27%</p>
3	L	323	<p>9% 44% 27% 27%</p>
3	O	323	<p>8% 42% 29% 27%</p>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 26525 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	225	Total	C	N	O	S	0	0	0
			1773	1134	299	329	11			
1	A	225	Total	C	N	O	S	0	0	0
			1773	1134	299	329	11			
1	D	225	Total	C	N	O	S	0	0	0
			1773	1134	299	329	11			
1	J	225	Total	C	N	O	S	0	0	0
			1773	1134	299	329	11			
1	M	225	Total	C	N	O	S	0	0	0
			1773	1134	299	329	11			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	215	LEU	LYS	engineered mutation	UNP F6KTB0
E	217	ALA	GLU	engineered mutation	UNP F6KTB0
E	218	ASN	LYS	engineered mutation	UNP F6KTB0
E	221	ASP	GLU	engineered mutation	UNP F6KTB0
A	215	LEU	LYS	engineered mutation	UNP F6KTB0
A	217	ALA	GLU	engineered mutation	UNP F6KTB0
A	218	ASN	LYS	engineered mutation	UNP F6KTB0
A	221	ASP	GLU	engineered mutation	UNP F6KTB0
D	215	LEU	LYS	engineered mutation	UNP F6KTB0
D	217	ALA	GLU	engineered mutation	UNP F6KTB0
D	218	ASN	LYS	engineered mutation	UNP F6KTB0
D	221	ASP	GLU	engineered mutation	UNP F6KTB0
J	215	LEU	LYS	engineered mutation	UNP F6KTB0
J	217	ALA	GLU	engineered mutation	UNP F6KTB0
J	218	ASN	LYS	engineered mutation	UNP F6KTB0
J	221	ASP	GLU	engineered mutation	UNP F6KTB0
M	215	LEU	LYS	engineered mutation	UNP F6KTB0
M	217	ALA	GLU	engineered mutation	UNP F6KTB0
M	218	ASN	LYS	engineered mutation	UNP F6KTB0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
M	221	ASP	GLU	engineered mutation	UNP F6KTB0

- Molecule 2 is a protein called Capsid protein VP3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	224	Total	C	N	O	S	0	0	0
			1712	1103	281	317	11			
2	B	224	Total	C	N	O	S	0	0	0
			1712	1103	281	317	11			
2	H	224	Total	C	N	O	S	0	0	0
			1712	1103	281	317	11			
2	K	224	Total	C	N	O	S	0	0	0
			1712	1103	281	317	11			
2	N	224	Total	C	N	O	S	0	0	0
			1712	1103	281	317	11			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	227	GLN	LYS	engineered mutation	UNP F6KTB0
B	227	GLN	LYS	engineered mutation	UNP F6KTB0
H	227	GLN	LYS	engineered mutation	UNP F6KTB0
K	227	GLN	LYS	engineered mutation	UNP F6KTB0
N	227	GLN	LYS	engineered mutation	UNP F6KTB0

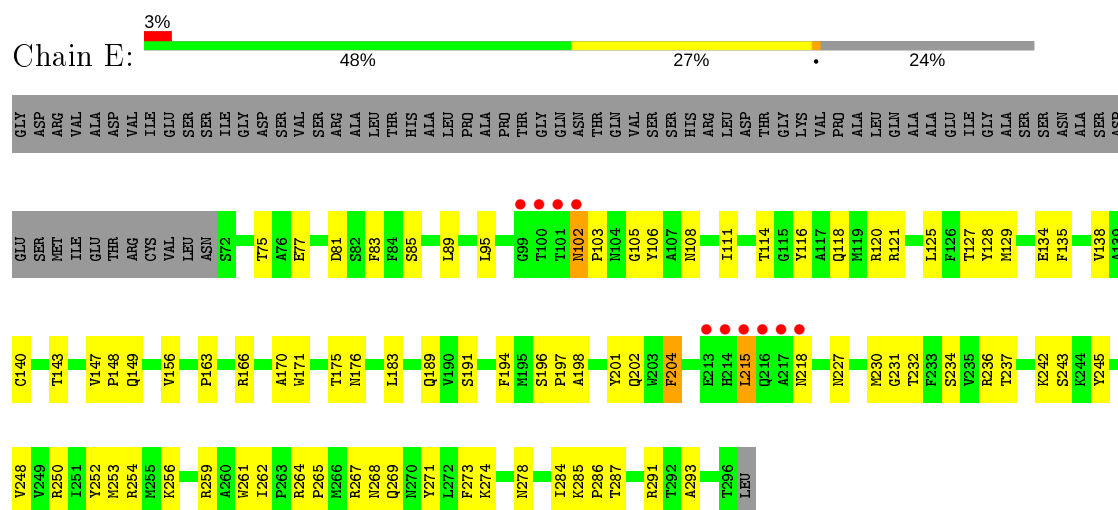
- Molecule 3 is a protein called Capsid protein VP0.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	235	Total	C	N	O	S	0	0	0
			1820	1170	299	343	8			
3	C	235	Total	C	N	O	S	0	0	0
			1820	1170	299	343	8			
3	I	235	Total	C	N	O	S	0	0	0
			1820	1170	299	343	8			
3	L	235	Total	C	N	O	S	0	0	0
			1820	1170	299	343	8			
3	O	235	Total	C	N	O	S	0	0	0
			1820	1170	299	343	8			

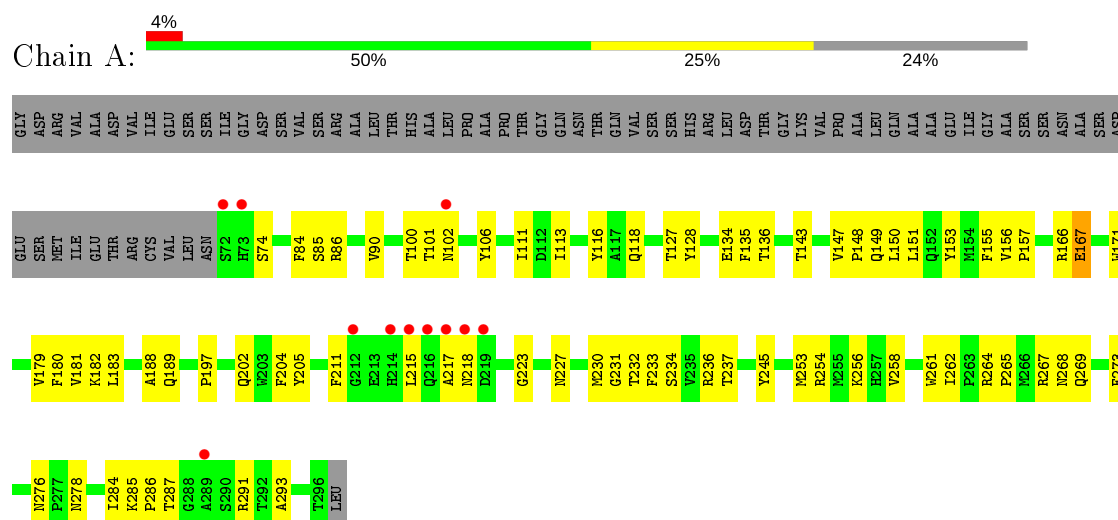
3 Residue-property plots

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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 VP1

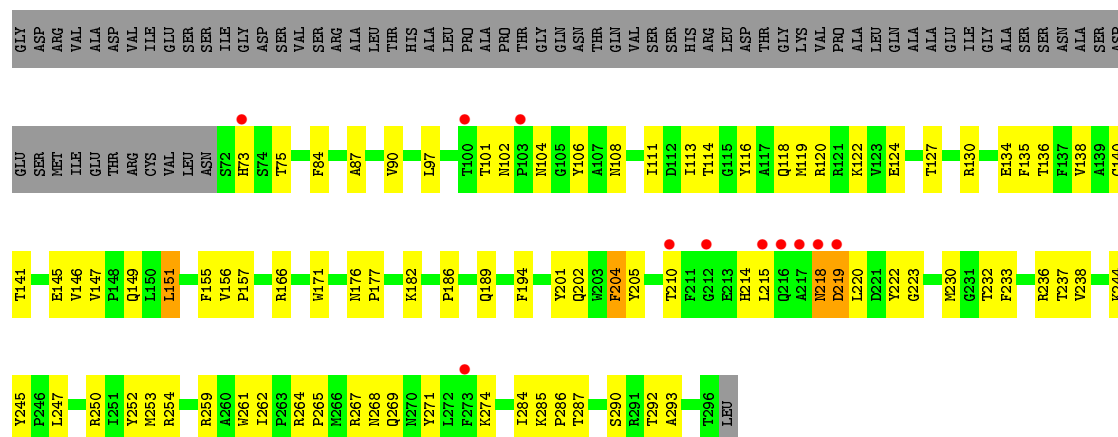


• Molecule 1: Capsid protein VP1

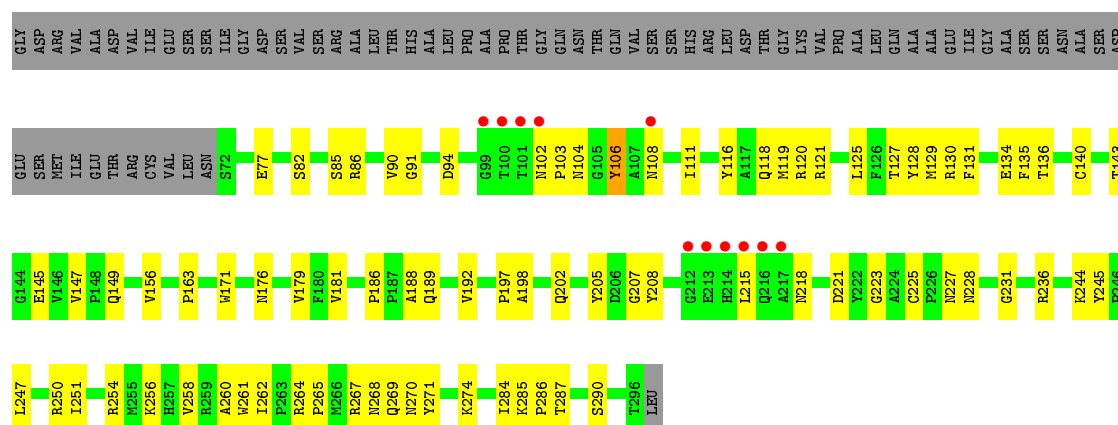


• Molecule 1: Capsid protein VP1

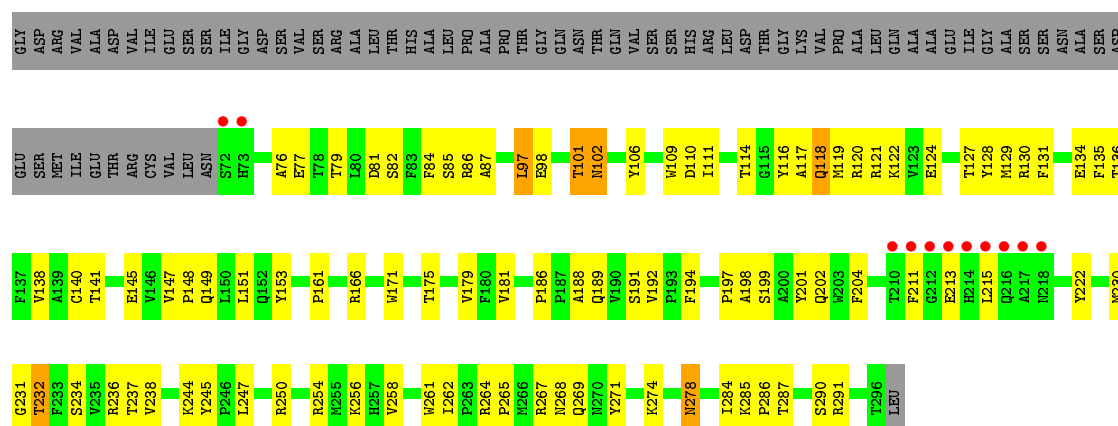
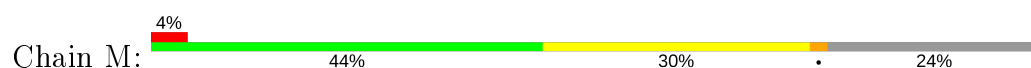




• Molecule 1: Capsid protein VP1



• Molecule 1: Capsid protein VP1

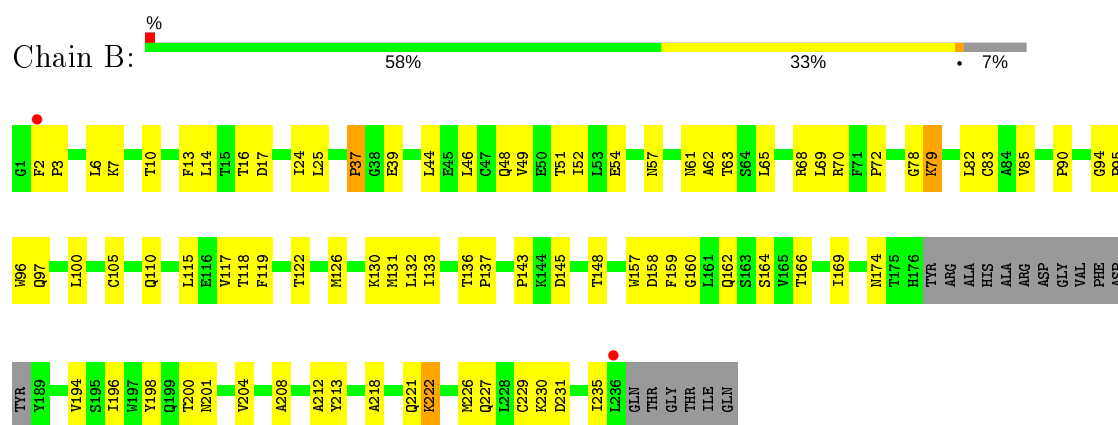


• Molecule 2: Capsid protein VP3

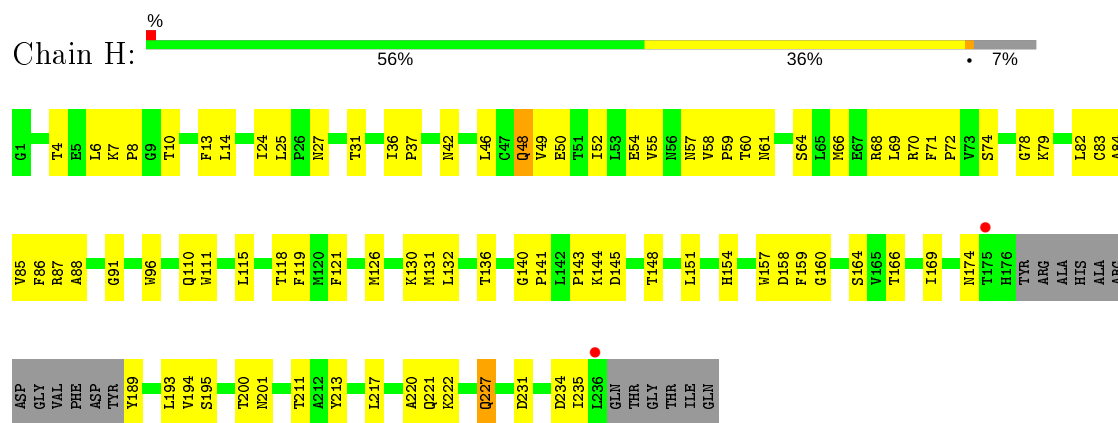




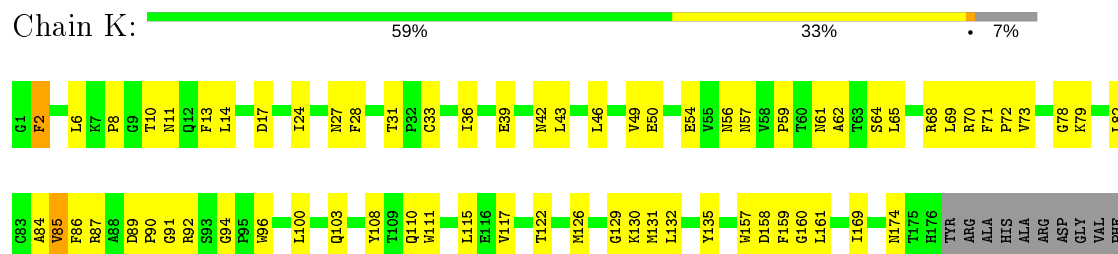
• Molecule 2: Capsid protein VP3



• Molecule 2: Capsid protein VP3



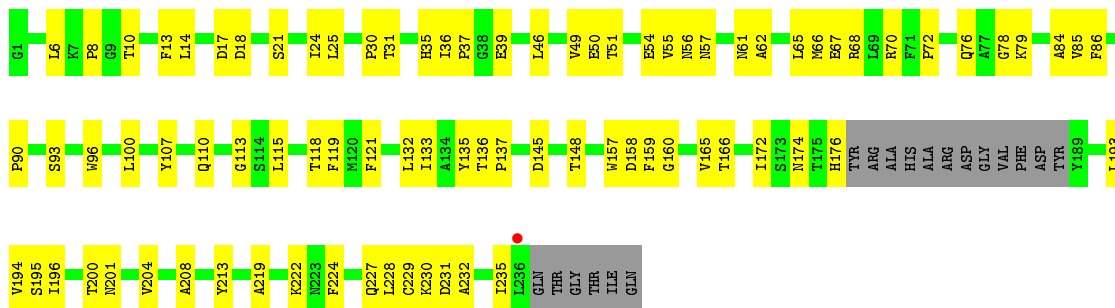
• Molecule 2: Capsid protein VP3





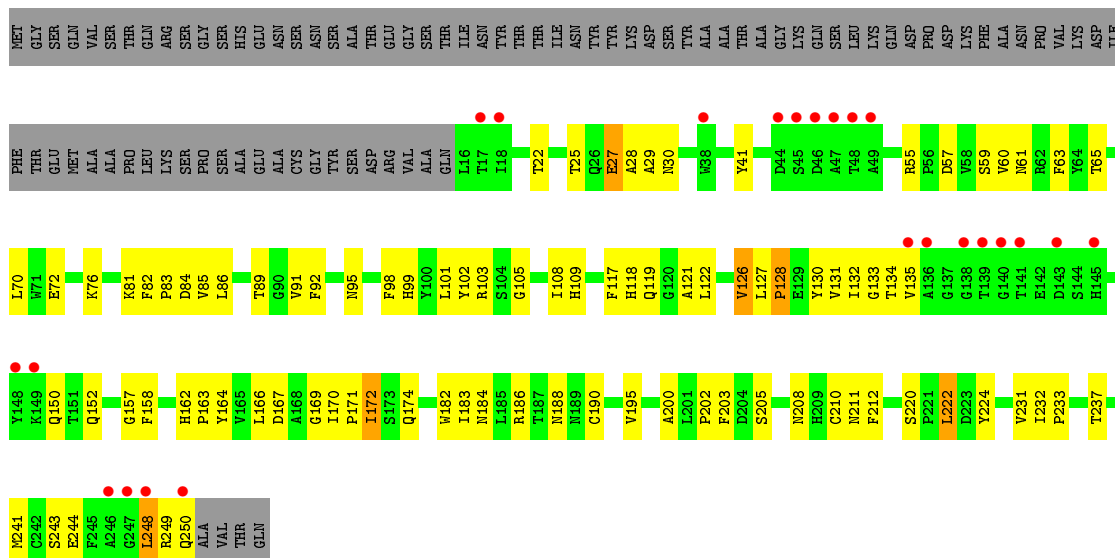
• Molecule 2: Capsid protein VP3

Chain N: 58% 35% 7%



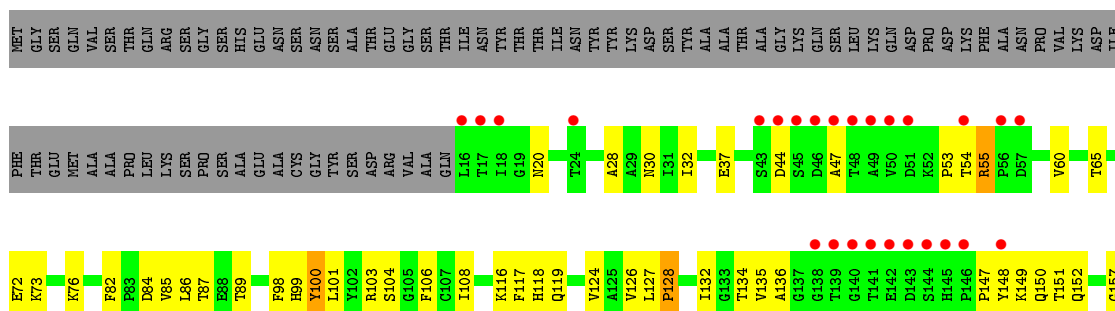
• Molecule 3: Capsid protein VP0

Chain G: 45% 26% 27%



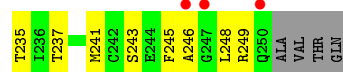
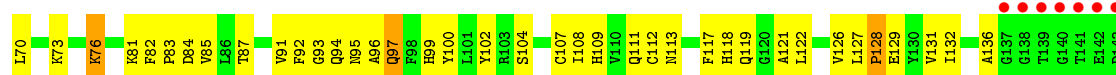
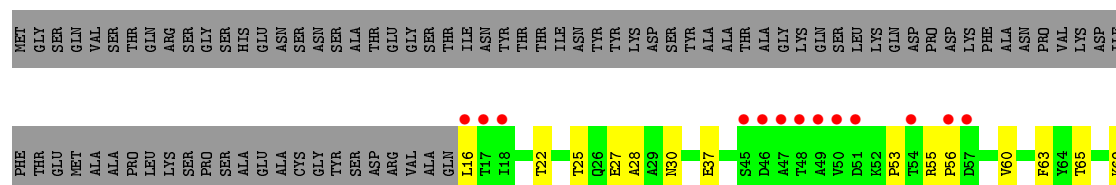
• Molecule 3: Capsid protein VP0

Chain C: 47% 24% 27%

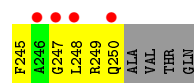
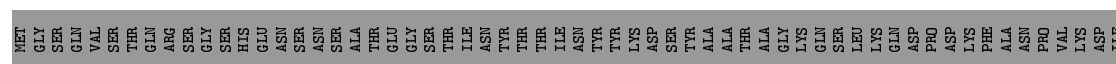




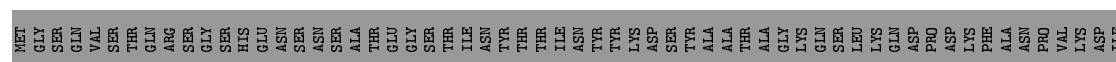
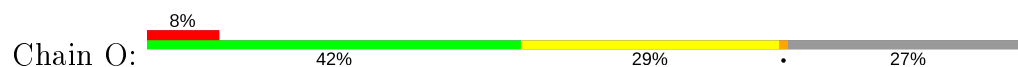
• Molecule 3: Capsid protein VP0

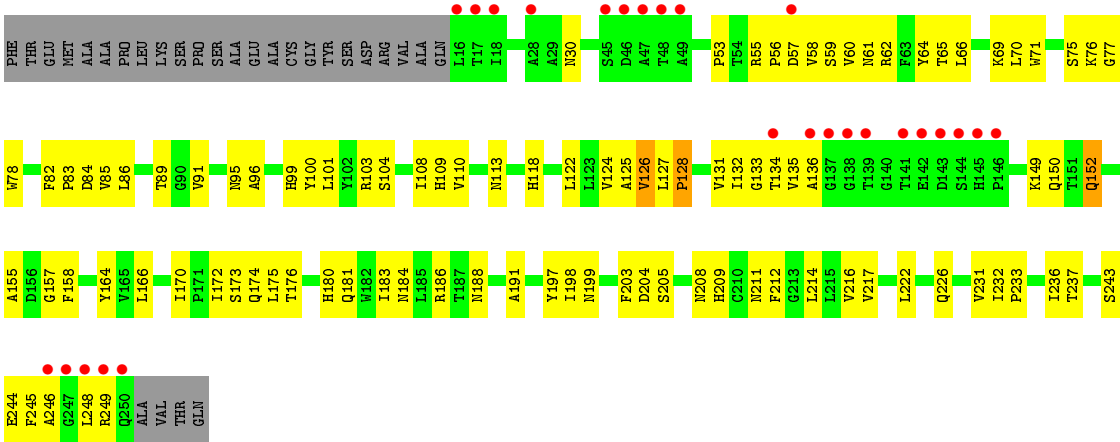


• Molecule 3: Capsid protein VP0



• Molecule 3: Capsid protein VP0





4 Data and refinement statistics

Property	Value	Source
Space group	P 42 3 2	Depositor
Cell constants a, b, c, α , β , γ	349.15Å 349.15Å 349.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.96 – 3.80 47.96 – 3.80	Depositor EDS
% Data completeness (in resolution range)	78.7 (47.96-3.80) 61.3 (47.96-3.80)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.52 (at 3.77Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1839)	Depositor
R, R_{free}	0.286 , 0.324 0.286 , 0.324	Depositor DCC
R_{free} test set	2000 reflections (3.52%)	wwPDB-VP
Wilson B-factor (Å ²)	94.0	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 37.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.84	EDS
Total number of atoms	26525	wwPDB-VP
Average B, all atoms (Å ²)	112.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.63% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	0.32	0/1828	0.57	0/2495
1	D	0.30	0/1828	0.53	0/2495
1	E	0.28	0/1828	0.53	0/2495
1	J	0.28	0/1828	0.52	0/2495
1	M	0.29	0/1828	0.53	0/2495
2	B	0.28	0/1759	0.55	0/2408
2	F	0.28	0/1759	0.57	0/2408
2	H	0.28	0/1759	0.57	0/2408
2	K	0.26	0/1759	0.56	0/2408
2	N	0.28	0/1759	0.54	0/2408
3	C	0.27	0/1875	0.56	0/2573
3	G	0.29	0/1875	0.56	1/2573 (0.0%)
3	I	0.30	0/1875	0.59	0/2573
3	L	0.29	0/1875	0.58	0/2573
3	O	0.30	0/1875	0.56	0/2573
All	All	0.29	0/27310	0.56	1/37380 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	248	LEU	CA-CB-CG	5.30	127.50	115.30

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	1773	0	1712	69	0
1	D	1773	0	1712	81	0
1	E	1773	0	1712	83	0
1	J	1773	0	1712	84	0
1	M	1773	0	1712	89	0
2	B	1712	0	1699	68	0
2	F	1712	0	1699	90	0
2	H	1712	0	1699	76	2
2	K	1712	0	1699	66	0
2	N	1712	0	1699	72	0
3	C	1820	0	1758	66	0
3	G	1820	0	1758	81	0
3	I	1820	0	1758	93	1
3	L	1820	0	1758	83	0
3	O	1820	0	1758	91	0
All	All	26525	0	25845	984	2

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 19.

The worst 5 of 984 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:M:189:GLN:HE21	2:N:21:SER:HB3	1.28	0.95
3:L:26:GLN:HA	3:L:26:GLN:HE21	1.33	0.93
1:J:262:ILE:HG21	3:L:128:PRO:HG2	1.52	0.90
3:G:132:ILE:HG12	3:G:150:GLN:HE21	1.40	0.87
1:E:262:ILE:HG21	3:G:128:PRO:HG2	1.57	0.87

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:60:THR:OG1	2:H:60:THR:OG1[14_777]	2.02	0.18
2:H:158:ASP:OD2	3:I:113:ASN:ND2[5_636]	2.19	0.01

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

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	223/297 (75%)	199 (89%)	22 (10%)	2 (1%)	17	54
1	D	223/297 (75%)	202 (91%)	21 (9%)	0	100	100
1	E	223/297 (75%)	198 (89%)	23 (10%)	2 (1%)	17	54
1	J	223/297 (75%)	199 (89%)	23 (10%)	1 (0%)	34	70
1	M	223/297 (75%)	199 (89%)	22 (10%)	2 (1%)	17	54
2	B	220/242 (91%)	198 (90%)	20 (9%)	2 (1%)	17	54
2	F	220/242 (91%)	201 (91%)	17 (8%)	2 (1%)	17	54
2	H	220/242 (91%)	200 (91%)	20 (9%)	0	100	100
2	K	220/242 (91%)	201 (91%)	19 (9%)	0	100	100
2	N	220/242 (91%)	199 (90%)	20 (9%)	1 (0%)	29	66
3	C	233/323 (72%)	196 (84%)	34 (15%)	3 (1%)	12	48
3	G	233/323 (72%)	197 (84%)	32 (14%)	4 (2%)	9	43
3	I	233/323 (72%)	198 (85%)	32 (14%)	3 (1%)	12	48
3	L	233/323 (72%)	195 (84%)	36 (16%)	2 (1%)	17	54
3	O	233/323 (72%)	196 (84%)	34 (15%)	3 (1%)	12	48
All	All	3380/4310 (78%)	2978 (88%)	375 (11%)	27 (1%)	19	57

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	128	PRO
3	I	145	HIS
3	G	27	GLU
3	G	128	PRO
3	I	128	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 X-ray entries followed by that with respect to entries of similar resolution.

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	192/250 (77%)	185 (96%)	7 (4%)	35	63
1	D	192/250 (77%)	184 (96%)	8 (4%)	30	58
1	E	192/250 (77%)	186 (97%)	6 (3%)	40	65
1	J	192/250 (77%)	189 (98%)	3 (2%)	62	79
1	M	192/250 (77%)	183 (95%)	9 (5%)	26	56
2	B	188/202 (93%)	186 (99%)	2 (1%)	73	85
2	F	188/202 (93%)	183 (97%)	5 (3%)	44	69
2	H	188/202 (93%)	184 (98%)	4 (2%)	53	74
2	K	188/202 (93%)	183 (97%)	5 (3%)	44	69
2	N	188/202 (93%)	181 (96%)	7 (4%)	34	62
3	C	200/272 (74%)	193 (96%)	7 (4%)	36	64
3	G	200/272 (74%)	195 (98%)	5 (2%)	47	70
3	I	200/272 (74%)	197 (98%)	3 (2%)	65	81
3	L	200/272 (74%)	194 (97%)	6 (3%)	41	66
3	O	200/272 (74%)	195 (98%)	5 (2%)	47	70
All	All	2900/3620 (80%)	2818 (97%)	82 (3%)	43	68

5 of 82 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	106	TYR
2	H	48	GLN
3	L	94	GLN
1	M	118	GLN
1	M	232	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	J	218	ASN
1	M	214	HIS
3	O	181	GLN
1	J	269	GLN
1	M	189	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 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.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	225/297 (75%)	-0.04	11 (4%) 29 25	75, 100, 167, 219	0
1	D	225/297 (75%)	-0.11	11 (4%) 29 25	74, 96, 172, 219	0
1	E	225/297 (75%)	-0.18	10 (4%) 34 29	77, 98, 163, 222	0
1	J	225/297 (75%)	-0.09	11 (4%) 29 25	76, 99, 170, 212	0
1	M	225/297 (75%)	-0.07	11 (4%) 29 25	75, 100, 172, 220	0
2	B	224/242 (92%)	-0.29	2 (0%) 84 79	77, 101, 128, 163	0
2	F	224/242 (92%)	-0.27	4 (1%) 68 61	78, 96, 124, 153	0
2	H	224/242 (92%)	-0.33	2 (0%) 84 79	76, 96, 122, 156	0
2	K	224/242 (92%)	-0.35	1 (0%) 92 89	78, 101, 127, 168	0
2	N	224/242 (92%)	-0.38	1 (0%) 92 89	79, 103, 127, 161	0
3	C	235/323 (72%)	0.36	29 (12%) 4 4	76, 121, 195, 213	0
3	G	235/323 (72%)	0.23	23 (9%) 7 7	75, 116, 191, 212	0
3	I	235/323 (72%)	0.36	28 (11%) 4 5	74, 117, 200, 223	0
3	L	235/323 (72%)	0.31	29 (12%) 4 4	78, 122, 201, 226	0
3	O	235/323 (72%)	0.25	26 (11%) 5 5	79, 124, 193, 222	0
All	All	3420/4310 (79%)	-0.03	199 (5%) 23 18	74, 102, 183, 226	0

The worst 5 of 199 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	145	HIS	8.3
3	I	138	GLY	7.8
3	I	139	THR	7.4
3	C	246	ALA	7.0
3	O	247	GLY	6.8

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.