wwPDB X-ray Structure Validation Summary Report (i)

Jun 15, 2024 – 08:43 PM EDT

PDB ID	:	4RQP
Title	:	Crystal structure of the natually occurring empty particle of a clinical C4 strain EV71
Authors	:	Chen, R.; Lyu, K.
Deposited on	:	2014-11-04
Resolution	:	3.15 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 (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 (1)) were used in the production of this report:

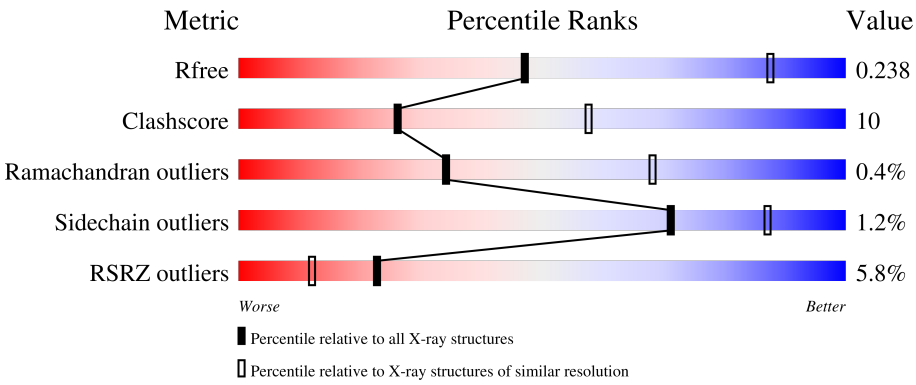
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.15 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	1665 (3.20-3.12)
Clashscore	141614	1804 (3.20-3.12)
Ramachandran outliers	138981	1770 (3.20-3.12)
Sidechain outliers	138945	1769 (3.20-3.12)
RSRZ outliers	127900	1616 (3.20-3.12)

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 of 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 <=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	3% 55% 19% • 25%
1	E	297	4% 57% 18% • 25%
1	I	297	4% 57% 18% • 25%
1	M	297	4% 56% 18% • 25%
1	Q	297	4% 57% 16% • 25%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	242	2% 74% 18% 7%
2	F	242	2% 76% 17% 7%
2	J	242	2% 74% 19% 7%
2	N	242	2% 75% 18% 7%
2	R	242	2% 75% 17% 7%
3	C	323	8% 57% 16% 26%
3	G	323	7% 59% 14% 26%
3	K	323	8% 54% 18% 26%
3	O	323	7% 55% 18% 26%
3	S	323	7% 56% 17% 26%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 26615 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
			Total	C	N	O	S			
1	E	223	1764	1130	296	327	11	0	0	0
1	Q	223	1764	1130	296	327	11	0	0	0
1	I	223	1764	1130	296	327	11	0	0	0
1	M	223	1764	1130	296	327	11	0	0	0
1	A	223	1764	1130	296	327	11	0	0	0

- Molecule 2 is a protein called Capsid protein VP3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	F	225	1718	1106	281	320	11	0	0	0
2	R	225	1718	1106	281	320	11	0	0	0
2	J	225	1718	1106	281	320	11	0	0	0
2	N	225	1718	1106	281	320	11	0	0	0
2	B	225	1718	1106	281	320	11	0	0	0

There are 5 discrepancies between the modelled and reference sequences:

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

- Molecule 3 is a protein called Capsid protein VP0.

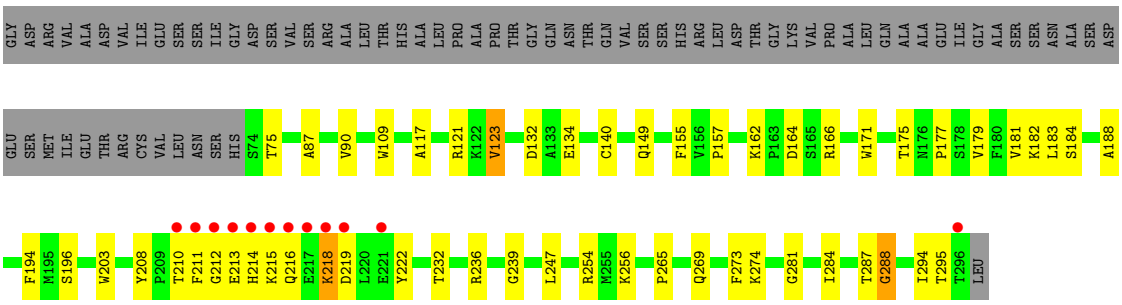
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	G	238	1841	1183	303	347	8	0	0	0
			Total	C	N	O	S			
3	S	238	1841	1183	303	347	8	0	0	0
			Total	C	N	O	S			
3	K	238	1841	1183	303	347	8	0	0	0
			Total	C	N	O	S			
3	O	238	1841	1183	303	347	8	0	0	0
			Total	C	N	O	S			
3	C	238	1841	1183	303	347	8	0	0	0

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 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.

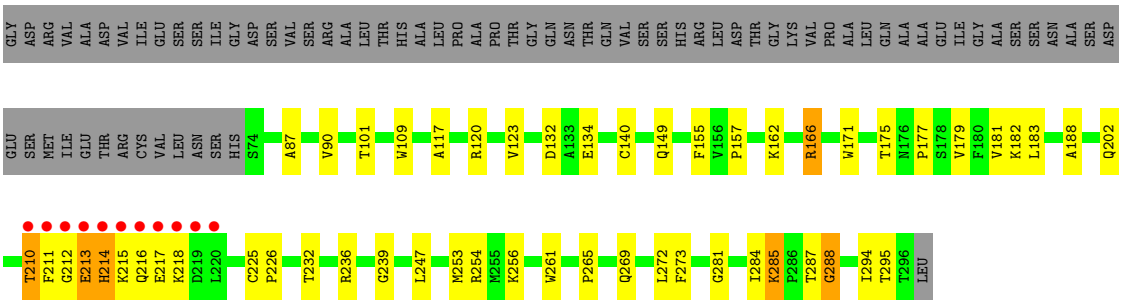
- Molecule 1: Capsid protein VP1

Chain E:



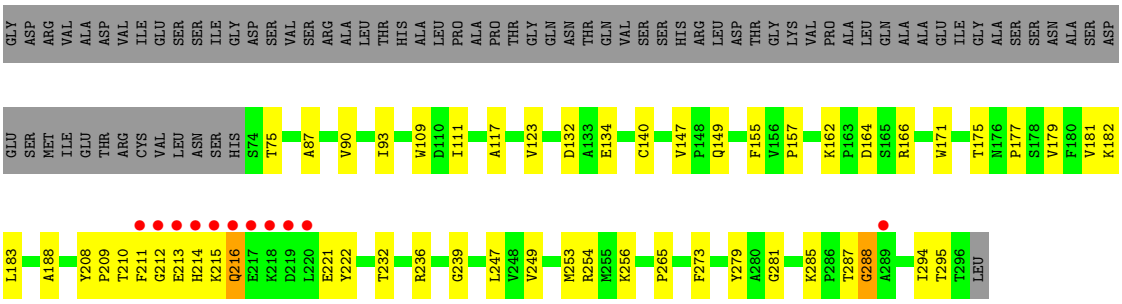
- Molecule 1: Capsid protein VP1

Chain Q:



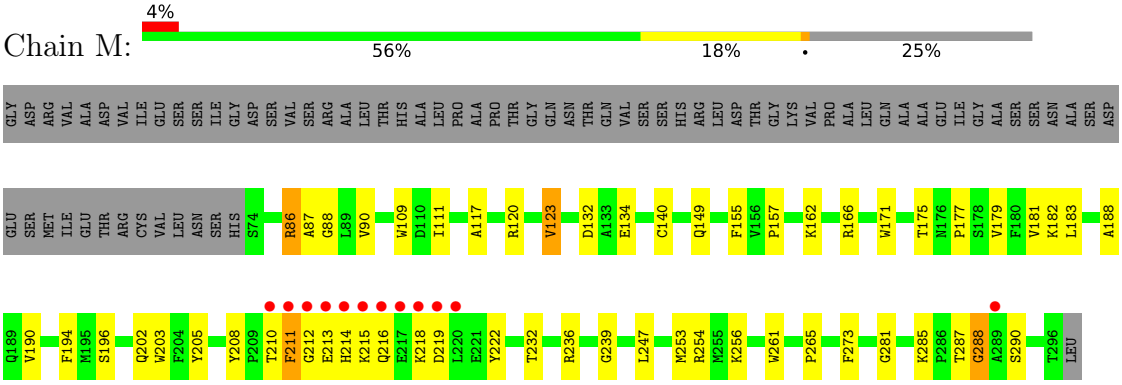
- Molecule 1: Capsid protein VP1

Chain I:



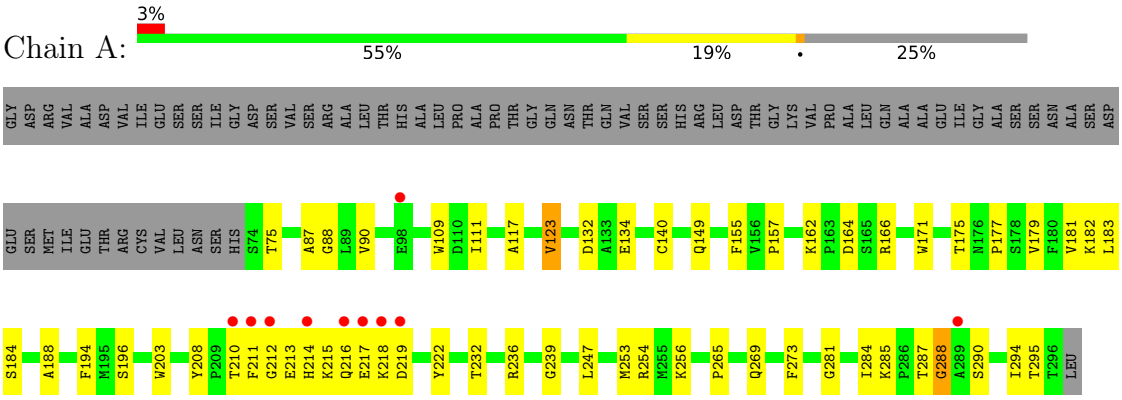
- Molecule 1: Capsid protein VP1

Chain M:



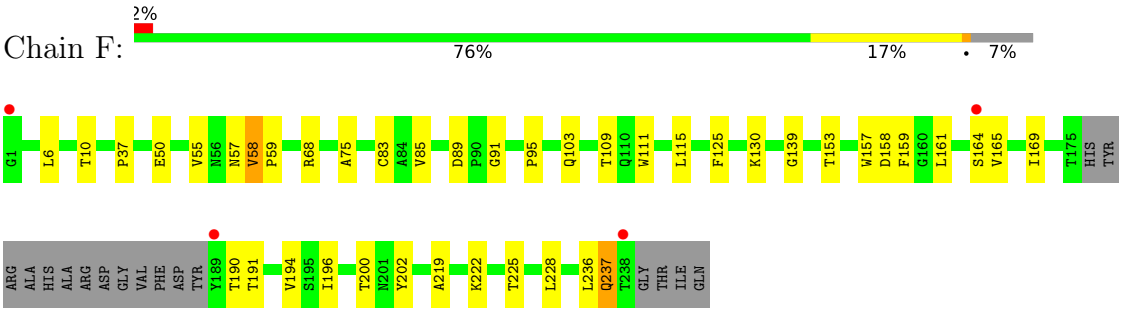
- Molecule 1: Capsid protein VP1

Chain A:



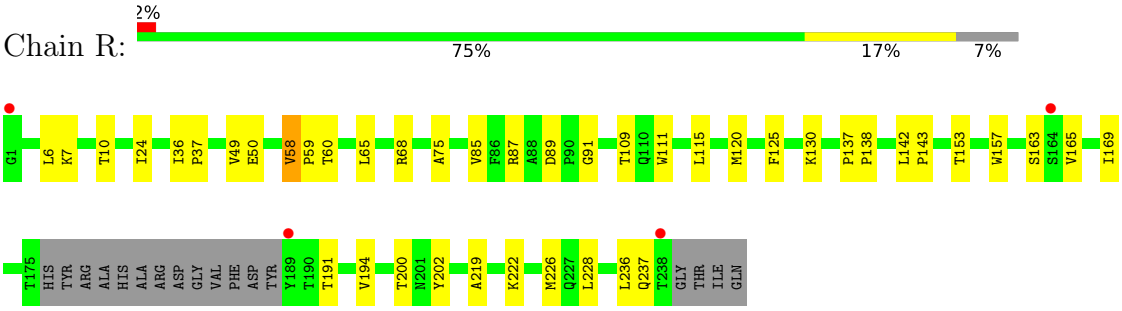
- Molecule 2: Capsid protein VP3

Chain F:



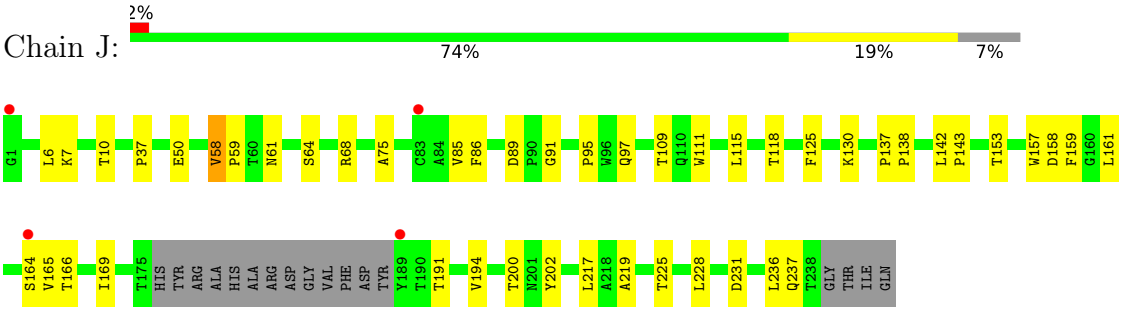
- Molecule 2: Capsid protein VP3

Chain R:



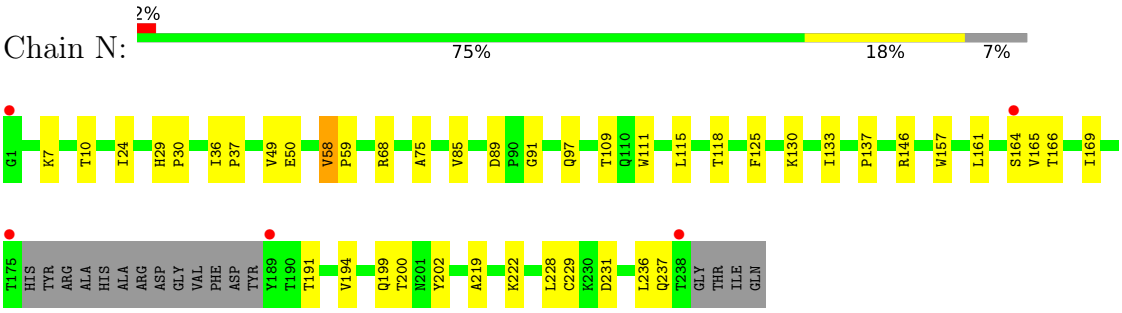
- Molecule 2: Capsid protein VP3

Chain J:



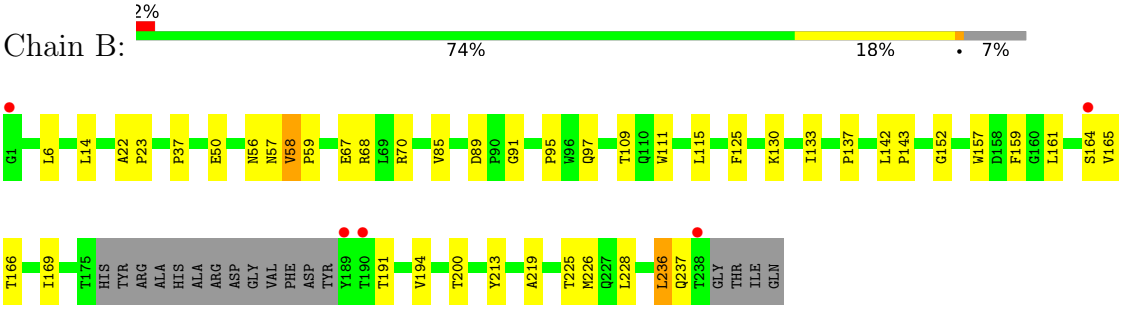
- Molecule 2: Capsid protein VP3

Chain N:



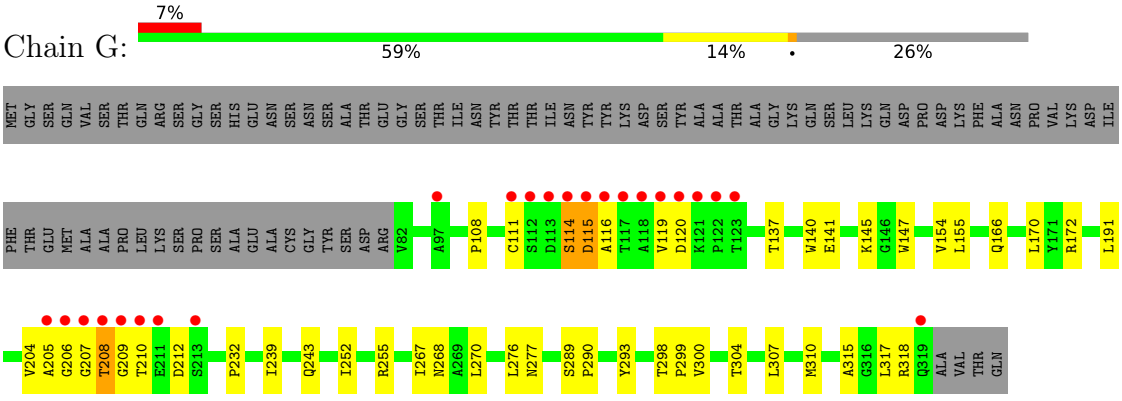
- Molecule 2: Capsid protein VP3

Chain B:



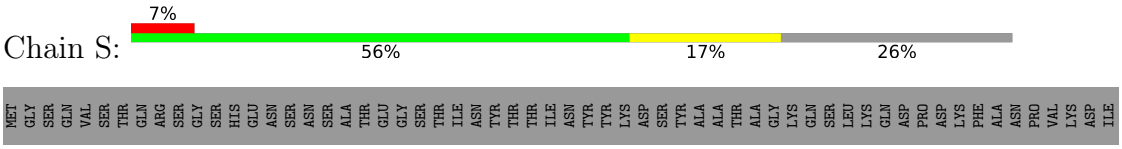
- Molecule 3: Capsid protein VP0

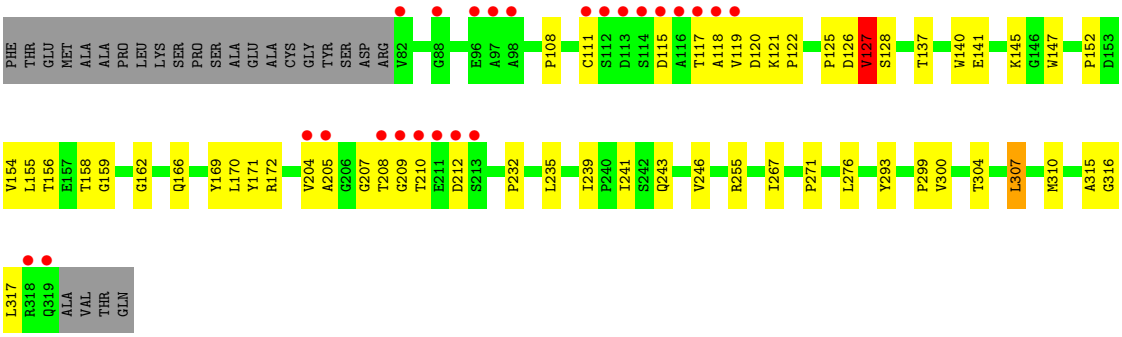
Chain G:



- Molecule 3: Capsid protein VP0

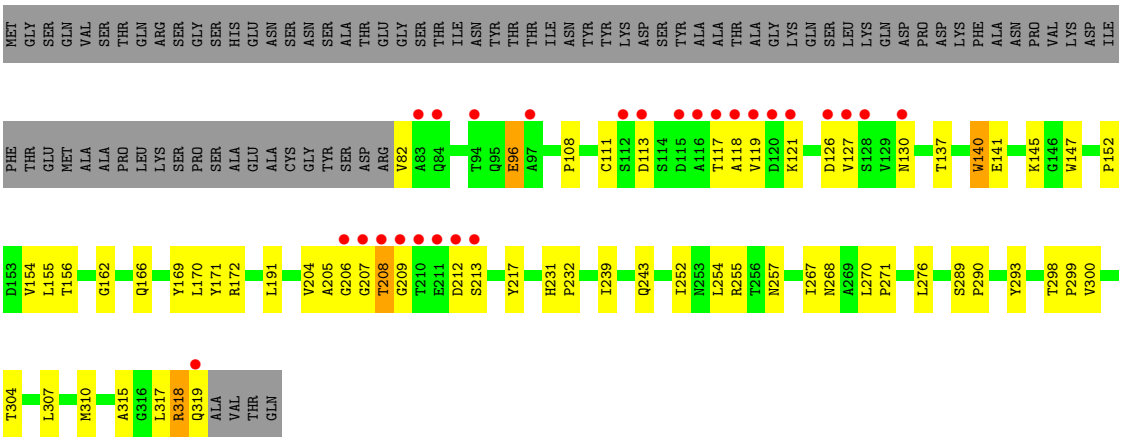
Chain S:





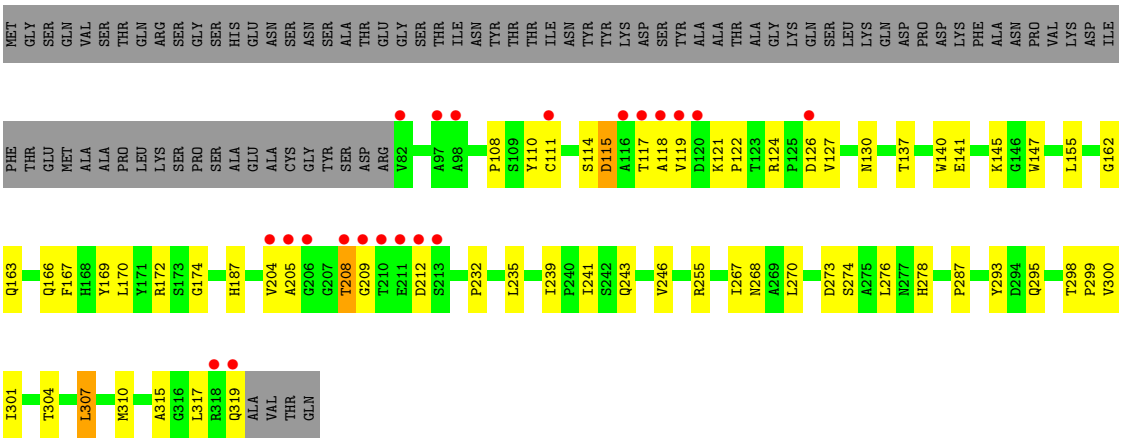
• Molecule 3: Capsid protein VP0

Chain K:



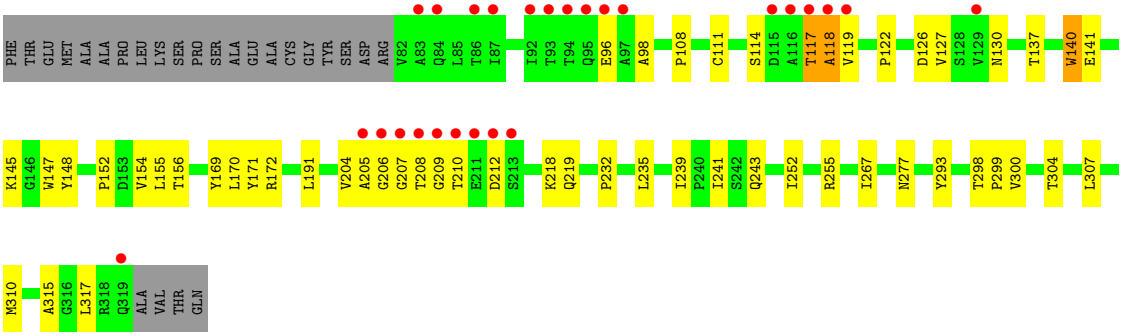
• Molecule 3: Capsid protein VP0

Chain O:



• Molecule 3: Capsid protein VP0

Chain C:


PHE
THR
GLU
MET
ALA
ALA
PRO
LEU
LYS
SER
PRO
SER
ALA
GLU
ALA
CYS
GLY
TYR
SER
ASP
ARG
V82
A83
Q84
L85
T86
I87
I92
T93
T94
Q95
E96
A97
A98
P108
C111
S114
D115
A116
T117
A118
V119
P122
D126
V127
S128
V129
N130
T137
W140
E141
K145
G146
W147
Y148
P152
D153
V154
L155
T156
Y169
L170
Y171
R172
L191
V204
A205
G206
G207
T208
G209
T210
E211
D212
S213
K218
Q219
P232
L235
I239
P240
I241
S242
Q243
I252
R255
I267
N277
Y293
T298
P299
V300
T304
L307
M310
A315
G316
L317
R318
Q319
ALA
VAL
THR
GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 42 3 2	Depositor
Cell constants a, b, c, α, β, γ	352.98Å 352.98Å 352.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.92 – 3.15 49.92 – 3.15	Depositor EDS
% Data completeness (in resolution range)	(Not available) (49.92-3.15) 78.4 (49.92-3.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.46 (at 3.12Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.216 , 0.234 0.216 , 0.238	Depositor DCC
R_{free} test set	2000 reflections (1.82%)	wwPDB-VP
Wilson B-factor (Å^2)	59.2	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 20.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	26615	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.28% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0/1818	0.51	0/2478
1	E	0.31	0/1818	0.50	0/2478
1	I	0.36	0/1818	0.53	0/2478
1	M	0.31	0/1818	0.50	0/2478
1	Q	0.35	0/1818	0.56	0/2478
2	B	0.31	0/1764	0.47	0/2415
2	F	0.27	0/1764	0.47	0/2415
2	J	0.30	0/1764	0.47	0/2415
2	N	0.29	0/1764	0.47	0/2415
2	R	0.28	0/1764	0.48	0/2415
3	C	0.31	0/1896	0.51	1/2602 (0.0%)
3	G	0.28	0/1896	0.50	1/2602 (0.0%)
3	K	0.30	0/1896	0.52	1/2602 (0.0%)
3	O	0.29	0/1896	0.51	1/2602 (0.0%)
3	S	0.32	0/1896	0.55	1/2602 (0.0%)
All	All	0.31	0/27390	0.50	5/37475 (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
1	I	0	1
3	G	0	1
All	All	0	2

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	O	115	ASP	CB-CG-OD1	5.60	123.34	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	115	ASP	N-CA-C	5.37	125.50	111.00
3	S	127	VAL	CG1-CB-CG2	-5.30	102.42	110.90
3	C	140	TRP	CA-CB-CG	5.02	123.23	113.70
3	K	140	TRP	CA-CB-CG	5.00	123.20	113.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	G	114	SER	Peptide
1	I	216	GLN	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	A	1764	0	1712	54	0
1	E	1764	0	1712	47	0
1	I	1764	0	1712	44	0
1	M	1764	0	1712	55	0
1	Q	1764	0	1712	55	0
2	B	1718	0	1707	33	0
2	F	1718	0	1707	38	0
2	J	1718	0	1707	37	0
2	N	1718	0	1707	39	0
2	R	1718	0	1707	40	0
3	C	1841	0	1780	42	0
3	G	1841	0	1780	41	0
3	K	1841	0	1780	51	0
3	O	1841	0	1780	46	0
3	S	1841	0	1780	42	0
All	All	26615	0	25995	549	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 10.

The worst 5 of 549 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:I:210:THR:HG22	1:I:211:PHE:HA	1.29	1.15
1:Q:213:GLU:HB2	1:Q:214:HIS:HA	1.34	1.05
1:A:215:LYS:H	1:A:216:GLN:HA	1.29	0.96
1:M:86:ARG:NH2	2:N:229:CYS:SG	2.42	0.92
1:Q:213:GLU:CB	1:Q:214:HIS:HA	2.03	0.88

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 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	221/297 (74%)	207 (94%)	13 (6%)	1 (0%)	29	65
1	E	221/297 (74%)	206 (93%)	14 (6%)	1 (0%)	29	65
1	I	221/297 (74%)	208 (94%)	12 (5%)	1 (0%)	29	65
1	M	221/297 (74%)	207 (94%)	13 (6%)	1 (0%)	29	65
1	Q	221/297 (74%)	207 (94%)	12 (5%)	2 (1%)	17	53
2	B	221/242 (91%)	202 (91%)	18 (8%)	1 (0%)	29	65
2	F	221/242 (91%)	201 (91%)	20 (9%)	0	100	100
2	J	221/242 (91%)	201 (91%)	20 (9%)	0	100	100
2	N	221/242 (91%)	200 (90%)	21 (10%)	0	100	100
2	R	221/242 (91%)	201 (91%)	20 (9%)	0	100	100
3	C	236/323 (73%)	219 (93%)	16 (7%)	1 (0%)	34	68
3	G	236/323 (73%)	219 (93%)	17 (7%)	0	100	100
3	K	236/323 (73%)	220 (93%)	15 (6%)	1 (0%)	34	68
3	O	236/323 (73%)	219 (93%)	17 (7%)	0	100	100
3	S	236/323 (73%)	218 (92%)	15 (6%)	3 (1%)	12	44
All	All	3390/4310 (79%)	3135 (92%)	243 (7%)	12 (0%)	34	68

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	S	127	VAL
1	Q	288	GLY
1	M	288	GLY
1	A	288	GLY
3	K	113	ASP

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 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	191/251 (76%)	190 (100%)	1 (0%)	88	95
1	E	191/251 (76%)	189 (99%)	2 (1%)	76	89
1	I	191/251 (76%)	189 (99%)	2 (1%)	76	89
1	M	191/251 (76%)	188 (98%)	3 (2%)	62	83
1	Q	191/251 (76%)	184 (96%)	7 (4%)	34	66
2	B	189/202 (94%)	188 (100%)	1 (0%)	88	95
2	F	189/202 (94%)	187 (99%)	2 (1%)	73	88
2	J	189/202 (94%)	188 (100%)	1 (0%)	88	95
2	N	189/202 (94%)	188 (100%)	1 (0%)	88	95
2	R	189/202 (94%)	188 (100%)	1 (0%)	88	95
3	C	202/272 (74%)	199 (98%)	3 (2%)	65	84
3	G	202/272 (74%)	201 (100%)	1 (0%)	88	95
3	K	202/272 (74%)	198 (98%)	4 (2%)	55	79
3	O	202/272 (74%)	197 (98%)	5 (2%)	47	75
3	S	202/272 (74%)	201 (100%)	1 (0%)	88	95
All	All	2910/3625 (80%)	2875 (99%)	35 (1%)	71	87

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	O	208	THR
3	O	295	GLN
3	C	117	THR
1	I	123	VAL
1	Q	285	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	G	277	ASN
1	I	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 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 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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	223/297 (75%)	-0.32	10 (4%)	33	19	28, 37, 132, 220	0
1	E	223/297 (75%)	-0.25	12 (5%)	25	13	27, 37, 133, 222	0
1	I	223/297 (75%)	-0.30	11 (4%)	29	16	29, 38, 132, 219	0
1	M	223/297 (75%)	-0.31	12 (5%)	25	13	28, 38, 132, 219	0
1	Q	223/297 (75%)	-0.26	11 (4%)	29	16	28, 38, 133, 219	0
2	B	225/242 (92%)	-0.38	5 (2%)	62	47	30, 42, 79, 153	0
2	F	225/242 (92%)	-0.42	4 (1%)	68	55	30, 41, 77, 151	0
2	J	225/242 (92%)	-0.42	4 (1%)	68	55	30, 42, 79, 153	0
2	N	225/242 (92%)	-0.48	5 (2%)	62	47	30, 42, 81, 155	0
2	R	225/242 (92%)	-0.45	4 (1%)	68	55	30, 42, 79, 152	0
3	C	238/323 (73%)	0.18	26 (10%)	5	3	30, 50, 166, 220	0
3	G	238/323 (73%)	0.01	23 (9%)	7	4	30, 49, 161, 218	0
3	K	238/323 (73%)	0.24	26 (10%)	5	3	30, 50, 163, 218	0
3	O	238/323 (73%)	0.03	21 (8%)	10	5	30, 50, 163, 217	0
3	S	238/323 (73%)	0.22	24 (10%)	7	3	31, 52, 164, 221	0
All	All	3430/4310 (79%)	-0.19	198 (5%)	23	12	27, 42, 137, 222	0

The worst 5 of 198 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	210	THR	14.1
3	O	210	THR	10.7
3	S	210	THR	10.0
3	K	115	ASP	7.9
3	C	116	ALA	7.3

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 monosaccharides 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.