wwPDB X-ray Structure Validation Summary Report (i)

Aug 28, 2024 – 04:24 pm BST

PDB ID	:	8RT1
Title	:	BTV15 VP5 at pH 9.0
Authors	:	Sutton, G.C.; Stuart, D.I.
Deposited on	:	2024-01-25
Resolution	:	2.80 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 (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:

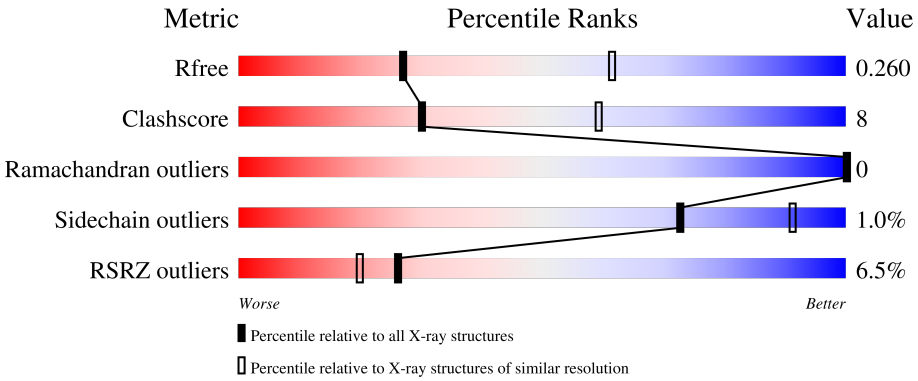
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.002 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.2

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 2.80 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}	164625	3657 (2.80-2.80)
Clashscore	180529	4123 (2.80-2.80)
Ramachandran outliers	177936	4071 (2.80-2.80)
Sidechain outliers	177891	4073 (2.80-2.80)
RSRZ outliers	164620	3659 (2.80-2.80)

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	490	5% 76% 16% • 8%
1	B	490	6% 74% 15% 10%
1	C	490	6% 72% 17% 10%
1	D	490	3% 75% 13% 12%
1	E	490	7% 74% 15% • 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	490	8% 74% 15% 10%
1	G	490	5% 73% 16% 11%
1	H	490	6% 72% 17% 10%
1	I	490	5% 74% 14% 12%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 31370 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 Outer capsid protein VP5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	451	Total	C	N	O	Se	0	0	0
			3585	2254	636	682	13			
1	B	440	Total	C	N	O	Se	0	0	0
			3485	2193	612	668	12			
1	C	439	Total	C	N	O	Se	0	0	0
			3479	2190	611	666	12			
1	D	432	Total	C	N	O	Se	0	0	0
			3431	2160	603	655	13			
1	E	442	Total	C	N	O	Se	0	0	0
			3516	2212	623	668	13			
1	F	439	Total	C	N	O	Se	0	0	0
			3479	2190	611	666	12			
1	G	438	Total	C	N	O	Se	0	0	0
			3471	2186	609	664	12			
1	H	440	Total	C	N	O	Se	0	0	0
			3484	2193	612	667	12			
1	I	433	Total	C	N	O	Se	0	0	0
			3440	2164	607	657	12			

There are 63 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	38	HIS	-	expression tag	UNP R4J9Y2
A	39	HIS	-	expression tag	UNP R4J9Y2
A	40	HIS	-	expression tag	UNP R4J9Y2
A	41	HIS	-	expression tag	UNP R4J9Y2
A	42	HIS	-	expression tag	UNP R4J9Y2
A	43	HIS	-	expression tag	UNP R4J9Y2
A	178	ALA	VAL	conflict	UNP R4J9Y2
B	38	HIS	-	expression tag	UNP R4J9Y2
B	39	HIS	-	expression tag	UNP R4J9Y2
B	40	HIS	-	expression tag	UNP R4J9Y2
B	41	HIS	-	expression tag	UNP R4J9Y2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	42	HIS	-	expression tag	UNP R4J9Y2
B	43	HIS	-	expression tag	UNP R4J9Y2
B	178	ALA	VAL	conflict	UNP R4J9Y2
C	38	HIS	-	expression tag	UNP R4J9Y2
C	39	HIS	-	expression tag	UNP R4J9Y2
C	40	HIS	-	expression tag	UNP R4J9Y2
C	41	HIS	-	expression tag	UNP R4J9Y2
C	42	HIS	-	expression tag	UNP R4J9Y2
C	43	HIS	-	expression tag	UNP R4J9Y2
C	178	ALA	VAL	conflict	UNP R4J9Y2
D	38	HIS	-	expression tag	UNP R4J9Y2
D	39	HIS	-	expression tag	UNP R4J9Y2
D	40	HIS	-	expression tag	UNP R4J9Y2
D	41	HIS	-	expression tag	UNP R4J9Y2
D	42	HIS	-	expression tag	UNP R4J9Y2
D	43	HIS	-	expression tag	UNP R4J9Y2
D	178	ALA	VAL	conflict	UNP R4J9Y2
E	38	HIS	-	expression tag	UNP R4J9Y2
E	39	HIS	-	expression tag	UNP R4J9Y2
E	40	HIS	-	expression tag	UNP R4J9Y2
E	41	HIS	-	expression tag	UNP R4J9Y2
E	42	HIS	-	expression tag	UNP R4J9Y2
E	43	HIS	-	expression tag	UNP R4J9Y2
E	178	ALA	VAL	conflict	UNP R4J9Y2
F	38	HIS	-	expression tag	UNP R4J9Y2
F	39	HIS	-	expression tag	UNP R4J9Y2
F	40	HIS	-	expression tag	UNP R4J9Y2
F	41	HIS	-	expression tag	UNP R4J9Y2
F	42	HIS	-	expression tag	UNP R4J9Y2
F	43	HIS	-	expression tag	UNP R4J9Y2
F	178	ALA	VAL	conflict	UNP R4J9Y2
G	38	HIS	-	expression tag	UNP R4J9Y2
G	39	HIS	-	expression tag	UNP R4J9Y2
G	40	HIS	-	expression tag	UNP R4J9Y2
G	41	HIS	-	expression tag	UNP R4J9Y2
G	42	HIS	-	expression tag	UNP R4J9Y2
G	43	HIS	-	expression tag	UNP R4J9Y2
G	178	ALA	VAL	conflict	UNP R4J9Y2
H	38	HIS	-	expression tag	UNP R4J9Y2
H	39	HIS	-	expression tag	UNP R4J9Y2
H	40	HIS	-	expression tag	UNP R4J9Y2
H	41	HIS	-	expression tag	UNP R4J9Y2

Continued on next page...

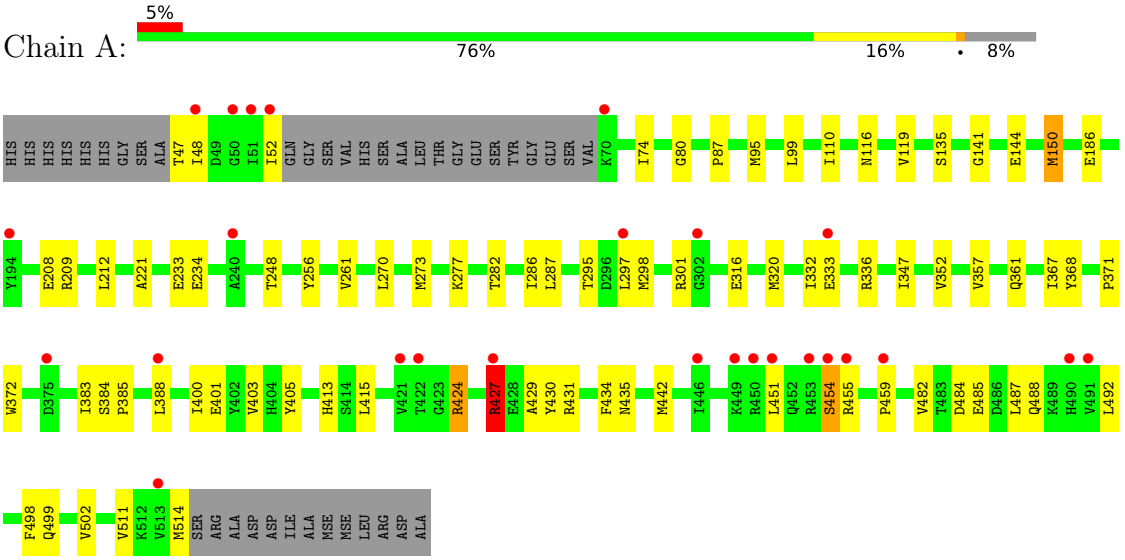
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	42	HIS	-	expression tag	UNP R4J9Y2
H	43	HIS	-	expression tag	UNP R4J9Y2
H	178	ALA	VAL	conflict	UNP R4J9Y2
I	38	HIS	-	expression tag	UNP R4J9Y2
I	39	HIS	-	expression tag	UNP R4J9Y2
I	40	HIS	-	expression tag	UNP R4J9Y2
I	41	HIS	-	expression tag	UNP R4J9Y2
I	42	HIS	-	expression tag	UNP R4J9Y2
I	43	HIS	-	expression tag	UNP R4J9Y2
I	178	ALA	VAL	conflict	UNP R4J9Y2

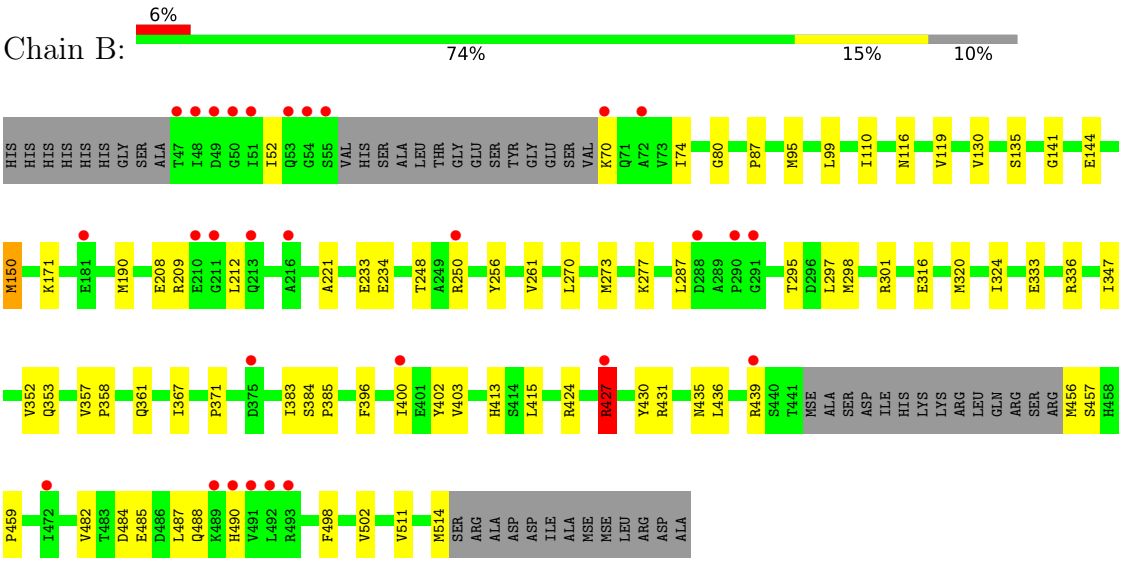
3 Residue-property plots (i)

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: Outer capsid protein VP5

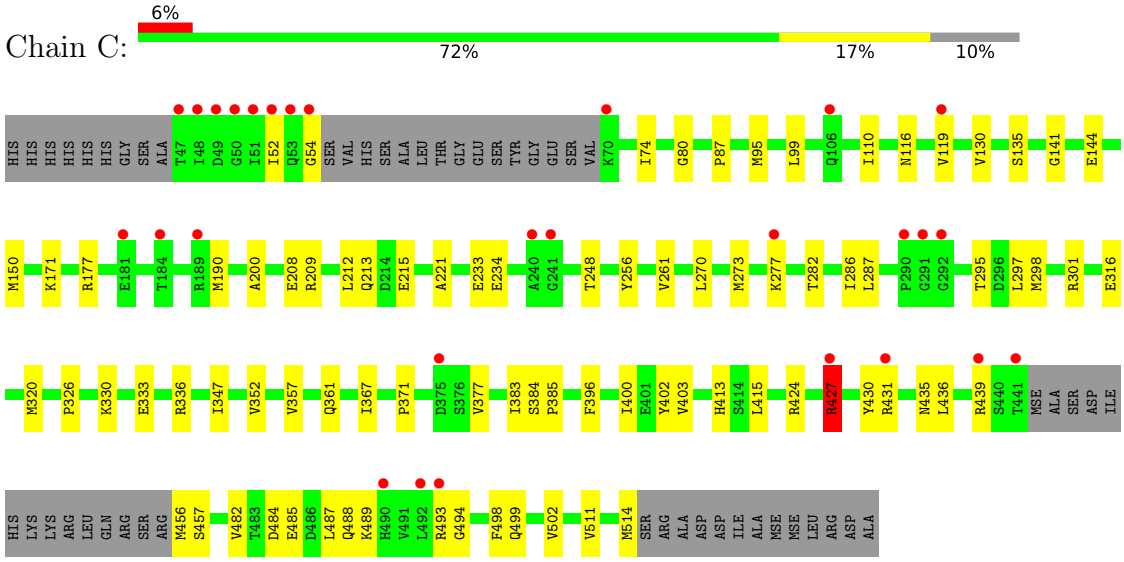


• Molecule 1: Outer capsid protein VP5



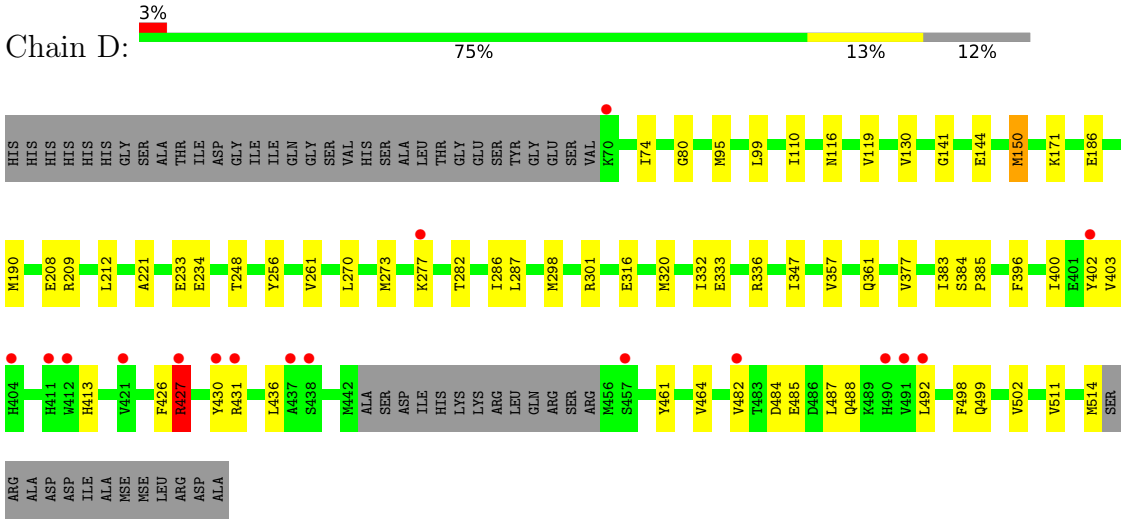
• Molecule 1: Outer capsid protein VP5

Chain C:



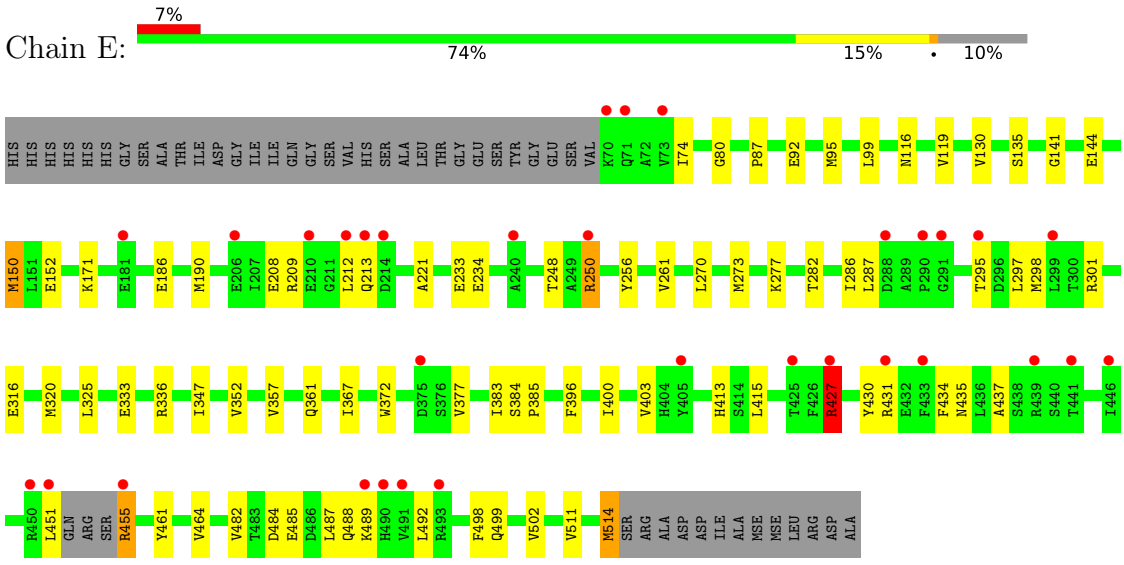
• Molecule 1: Outer capsid protein VP5

Chain D:



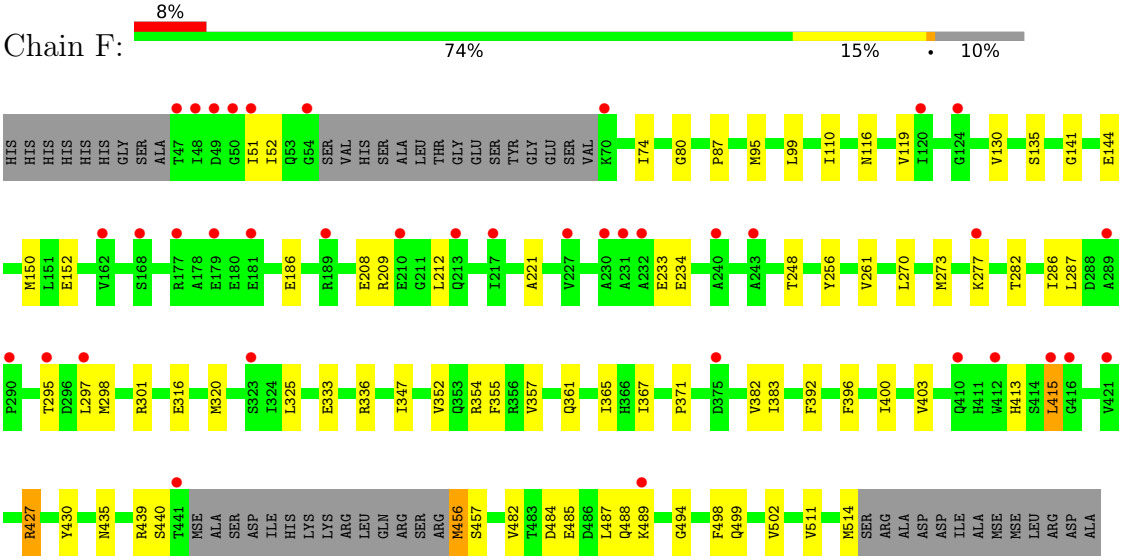
• Molecule 1: Outer capsid protein VP5

Chain E:



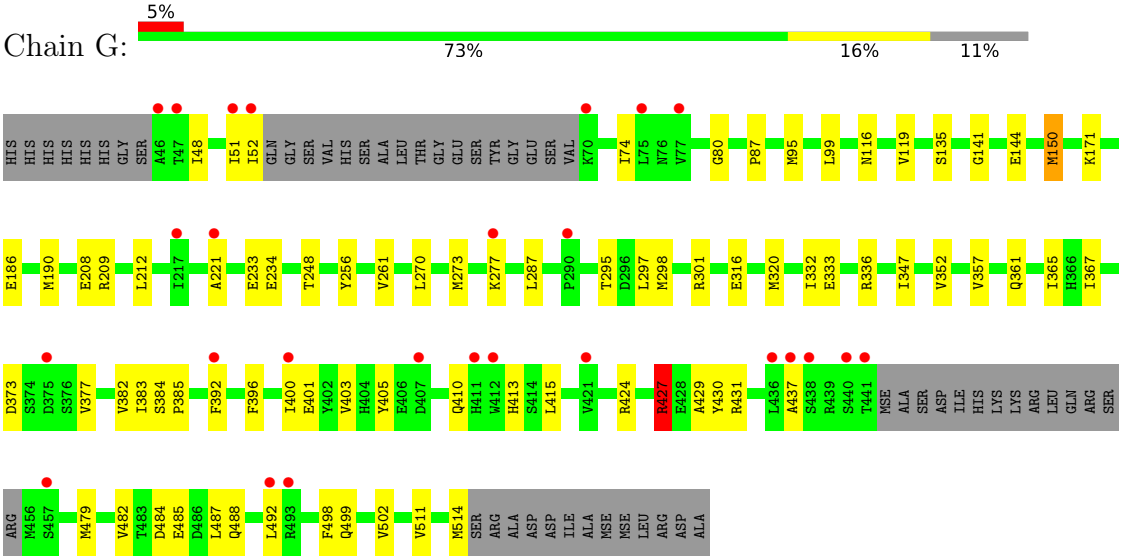
• Molecule 1: Outer capsid protein VP5

Chain F:



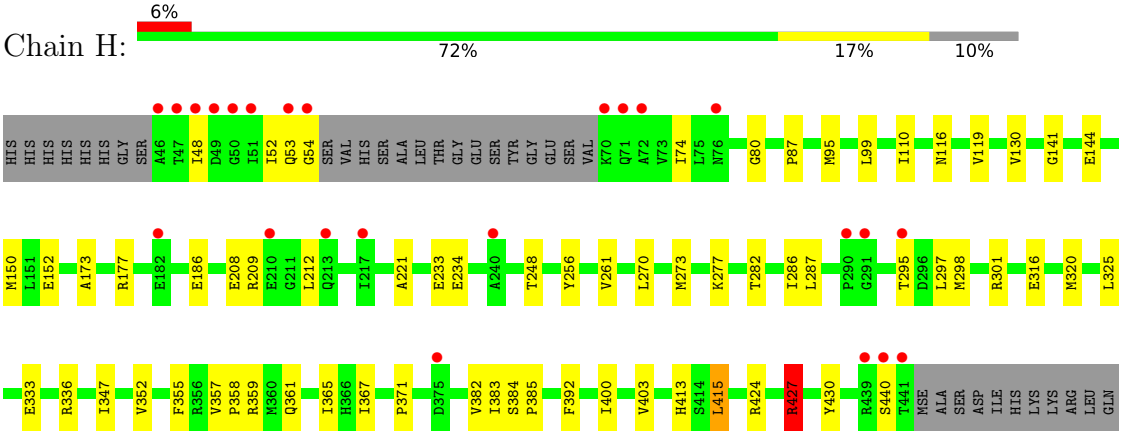
• Molecule 1: Outer capsid protein VP5

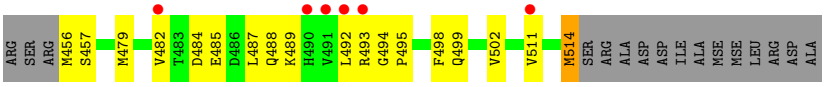
Chain G:



• Molecule 1: Outer capsid protein VP5

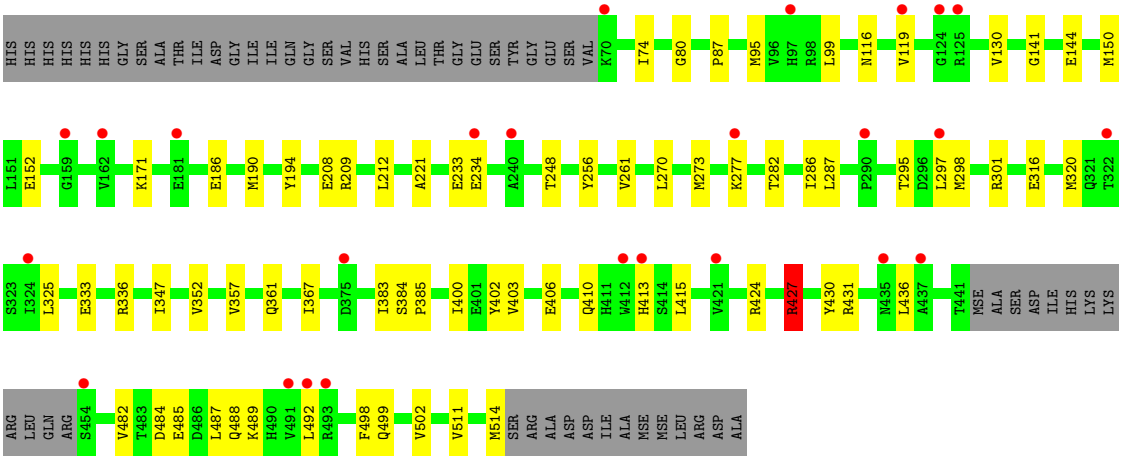
Chain H:





• Molecule 1: Outer capsid protein VP5

Chain I:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	91.40Å 142.21Å 142.15Å 119.54° 96.09° 95.80°	Depositor
Resolution (Å)	46.63 – 2.80 46.63 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.3 (46.63-2.80) 98.3 (46.63-2.80)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.24 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.240 , 0.261 0.239 , 0.260	Depositor DCC
R_{free} test set	7390 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	66.5	Xtriage
Anisotropy	0.029	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 31.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.009 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	31370	wwPDB-VP
Average B, all atoms (Å^2)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.40% 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

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	0.65	1/3634 (0.0%)	0.93	4/4878 (0.1%)
1	B	0.65	0/3534	0.94	5/4749 (0.1%)
1	C	0.65	0/3528	0.94	4/4741 (0.1%)
1	D	0.65	0/3480	0.92	2/4675 (0.0%)
1	E	0.65	0/3565	0.93	4/4785 (0.1%)
1	F	0.65	0/3528	0.93	4/4741 (0.1%)
1	G	0.65	0/3520	0.91	2/4731 (0.0%)
1	H	0.67	0/3533	0.95	2/4748 (0.0%)
1	I	0.64	0/3489	0.92	2/4687 (0.0%)
All	All	0.65	1/31811 (0.0%)	0.93	29/42735 (0.1%)

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	1
1	B	0	1
1	C	0	1
1	D	0	1
1	E	1	1
1	G	0	1
1	H	0	1
1	I	0	1
All	All	1	8

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	454	SER	CA-CB	-5.98	1.44	1.52

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	F	427	ARG	NE-CZ-NH2	9.92	125.26	120.30
1	C	177	ARG	NE-CZ-NH2	9.30	124.95	120.30
1	E	427	ARG	NE-CZ-NH1	-8.34	116.13	120.30
1	I	427	ARG	NE-CZ-NH1	-8.17	116.22	120.30
1	D	427	ARG	NE-CZ-NH1	-8.13	116.24	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	E	456	MSE	CA

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	427	ARG	Sidechain
1	B	427	ARG	Sidechain
1	C	427	ARG	Sidechain
1	D	427	ARG	Sidechain
1	E	427	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	3585	0	3589	66	0
1	B	3485	0	3475	67	0
1	C	3479	0	3470	62	0
1	D	3431	0	3422	49	0
1	E	3516	0	3517	60	0
1	F	3479	0	3470	68	0
1	G	3471	0	3464	62	0
1	H	3484	0	3475	64	0
1	I	3440	0	3431	52	0
All	All	31370	0	31313	520	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 8.

The worst 5 of 520 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:B:431:ARG:CG	1:B:456:MSE:HE1	1.37	1.51
1:F:427:ARG:NH1	1:F:456:MSE:HE1	1.18	1.45
1:C:456:MSE:SE	1:C:456:MSE:CE	2.17	1.43
1:B:431:ARG:CD	1:B:456:MSE:HE1	1.54	1.37
1:D:402:TYR:CZ	1:D:436:LEU:HD23	1.68	1.28

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	447/490 (91%)	432 (97%)	15 (3%)	0	100	100
1	B	434/490 (89%)	421 (97%)	13 (3%)	0	100	100
1	C	433/490 (88%)	423 (98%)	10 (2%)	0	100	100
1	D	428/490 (87%)	413 (96%)	15 (4%)	0	100	100
1	E	438/490 (89%)	426 (97%)	12 (3%)	0	100	100
1	F	433/490 (88%)	419 (97%)	14 (3%)	0	100	100
1	G	432/490 (88%)	415 (96%)	17 (4%)	0	100	100
1	H	434/490 (89%)	423 (98%)	11 (2%)	0	100	100
1	I	429/490 (88%)	414 (96%)	15 (4%)	0	100	100
All	All	3908/4410 (89%)	3786 (97%)	122 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	383/398 (96%)	380 (99%)	3 (1%)	79	93
1	B	372/398 (94%)	368 (99%)	4 (1%)	70	90
1	C	371/398 (93%)	369 (100%)	2 (0%)	86	95
1	D	366/398 (92%)	363 (99%)	3 (1%)	79	93
1	E	375/398 (94%)	371 (99%)	4 (1%)	70	90
1	F	371/398 (93%)	367 (99%)	4 (1%)	70	90
1	G	370/398 (93%)	366 (99%)	4 (1%)	70	90
1	H	371/398 (93%)	363 (98%)	8 (2%)	47	79
1	I	367/398 (92%)	365 (100%)	2 (0%)	86	95
All	All	3346/3582 (93%)	3312 (99%)	34 (1%)	73	91

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	424	ARG
1	H	492	LEU
1	I	186	GLU
1	E	150	MSE
1	D	234	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	404	HIS
1	G	404	HIS
1	E	435	ASN
1	E	404	HIS
1	F	404	HIS

5.3.3 RNA

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 oligosaccharides 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, 95^{th} 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	438/490 (89%)	0.40	26 (5%)	29	22	39, 67, 141, 228	0
1	B	428/490 (87%)	0.28	29 (6%)	25	19	36, 60, 105, 157	0
1	C	427/490 (87%)	0.51	28 (6%)	26	19	43, 67, 108, 148	0
1	D	419/490 (85%)	0.35	17 (4%)	42	34	39, 67, 126, 164	0
1	E	429/490 (87%)	0.40	32 (7%)	22	16	35, 68, 131, 181	0
1	F	427/490 (87%)	0.57	38 (8%)	17	13	44, 70, 113, 166	0
1	G	426/490 (86%)	0.46	26 (6%)	28	21	40, 79, 133, 193	0
1	H	428/490 (87%)	0.34	30 (7%)	24	18	37, 65, 112, 152	0
1	I	421/490 (85%)	0.52	25 (5%)	29	22	45, 79, 133, 176	0
All	All	3843/4410 (87%)	0.42	251 (6%)	26	20	35, 69, 125, 228	0

The worst 5 of 251 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	230	ALA	6.4
1	F	48	ILE	6.0
1	A	454	SER	5.5
1	E	213	GLN	5.5
1	B	55	SER	5.4

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.