



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 7, 2023 – 11:32 AM EDT

PDB ID : 4F5X
Title : Location of the dsRNA-dependent polymerase, VP1, in rotavirus particles
Authors : Estrozi, L.F.; Settembre, E.C.; Goret, G.; McClain, B.; Zhang, X.; Chen, J.Z.; Grigorieff, N.; Harrison, S.C.
Deposited on : 2012-05-13
Resolution : 5.00 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 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:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : **FAILED**
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35.1

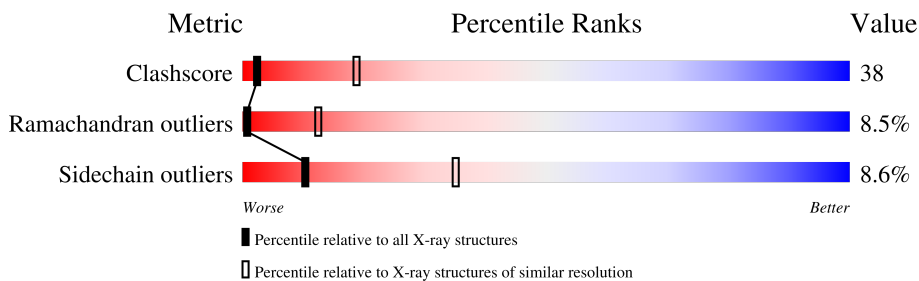
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 5.00 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))
Clashscore	141614	1000 (6.16-3.82)
Ramachandran outliers	138981	1146 (6.20-3.80)
Sidechain outliers	138945	1122 (6.20-3.80)


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 $\leq 5\%$.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	880	16% 49% 21% 11%
1	B	880	19% 49% 22% 8%
2	C	397	60% 31% 8%
2	D	397	57% 32% 9%
2	E	397	58% 33% 8%
2	F	397	59% 34% 7%
2	G	397	62% 30% 7%
2	H	397	60% 32% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	I	397	
2	J	397	
2	K	397	
2	L	397	
2	M	397	
2	N	397	
2	O	397	
3	W	1089	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 62014 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 VP2 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	781	Total 6374	C 4049	N 1099	O 1190	S 36	0	0	0
1	B	810	Total 6624	C 4211	N 1138	O 1239	S 36	0	0	0

- Molecule 2 is a protein called Intermediate capsid protein VP6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	397	Total 3162	C 2001	N 550	O 596	S 15	0	0	0
2	D	397	Total 3162	C 2001	N 550	O 596	S 15	0	0	0
2	E	397	Total 3162	C 2001	N 550	O 596	S 15	0	0	0
2	F	397	Total 3162	C 2001	N 550	O 596	S 15	0	0	0
2	G	397	Total 3162	C 2001	N 550	O 596	S 15	0	0	0
2	H	397	Total 3162	C 2001	N 550	O 596	S 15	0	0	0
2	I	397	Total 3162	C 2001	N 550	O 596	S 15	0	0	0
2	J	397	Total 3162	C 2001	N 550	O 596	S 15	0	0	0
2	K	397	Total 3162	C 2001	N 550	O 596	S 15	0	0	0
2	L	397	Total 3162	C 2001	N 550	O 596	S 15	0	0	0
2	M	397	Total 3162	C 2001	N 550	O 596	S 15	0	0	0
2	N	397	Total 3162	C 2001	N 550	O 596	S 15	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	O	397	3162	2001	550	596	15	0	0	0

- Molecule 3 is a protein called RNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	W	975	7905	5081	1308	1482	34	0	0	0

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

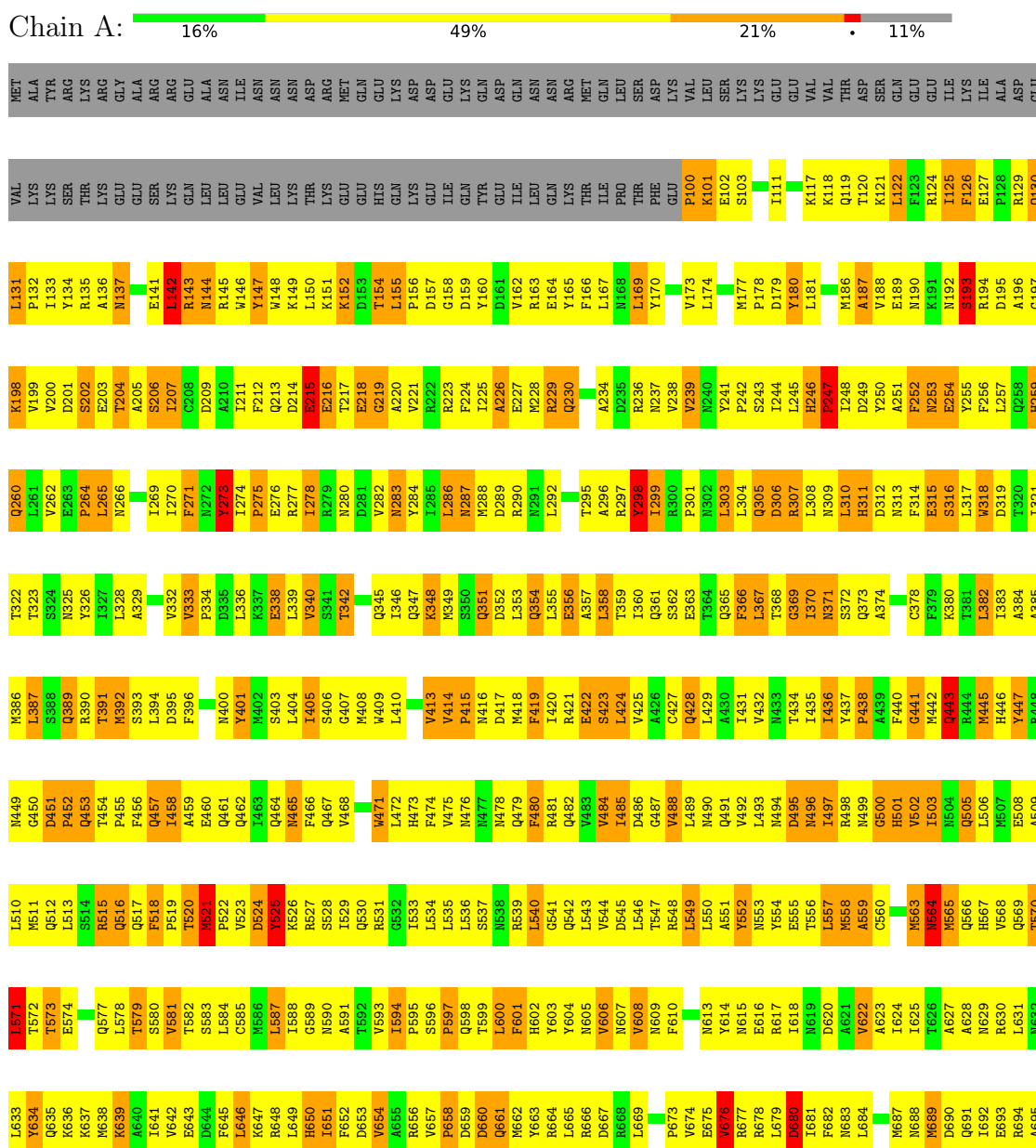
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
4	C	1	1	1	0	0
4	F	1	1	1	0	0
4	I	1	1	1	0	0
4	L	1	1	1	0	0
4	O	1	1	1	0	0

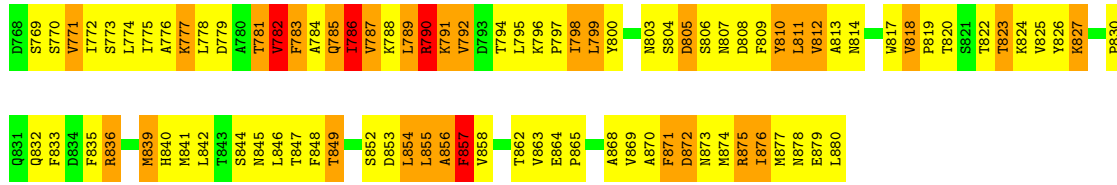
3 Residue-property plots

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.

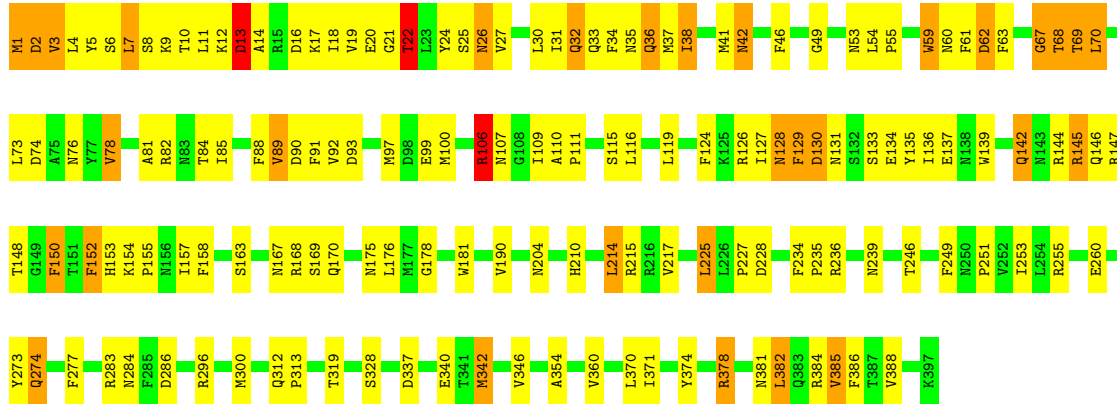
Note EDS failed to run properly.

- Molecule 1: VP2 protein

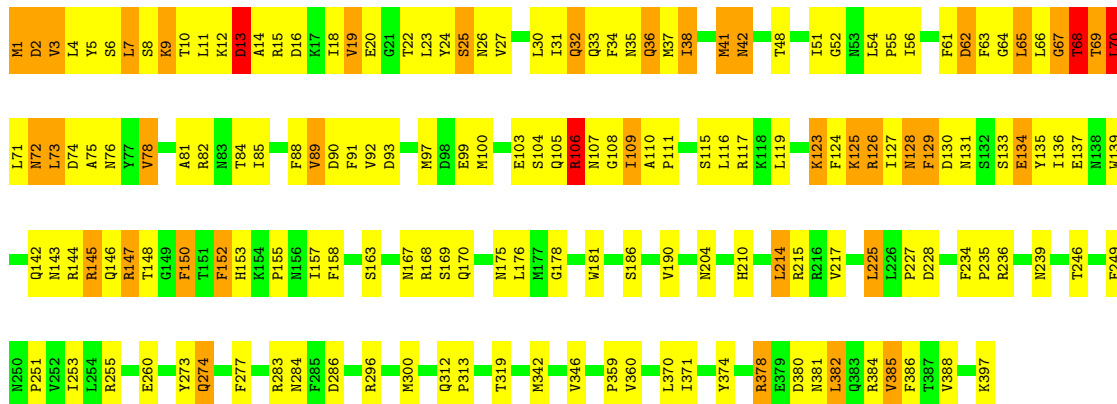




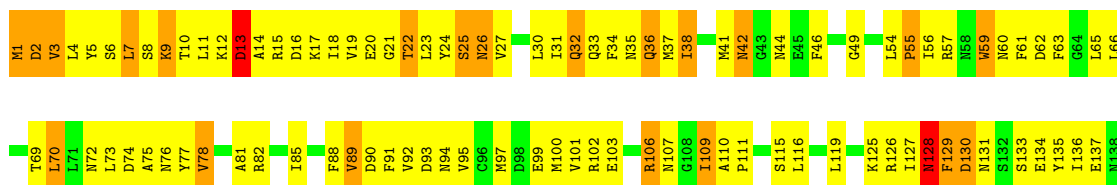
• Molecule 2: Intermediate capsid protein VP6

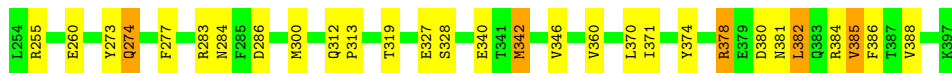


• Molecule 2: Intermediate capsid protein VP6



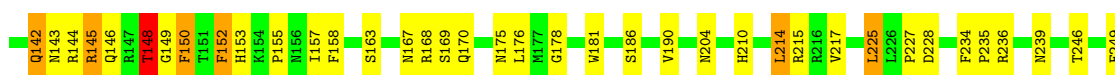
• Molecule 2: Intermediate capsid protein VP6





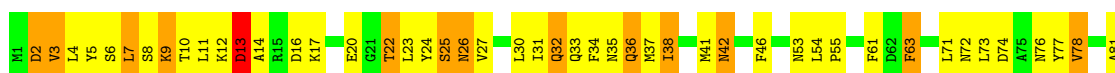
• Molecule 2: Intermediate capsid protein VP6

Chain F: 59% 34% 7%



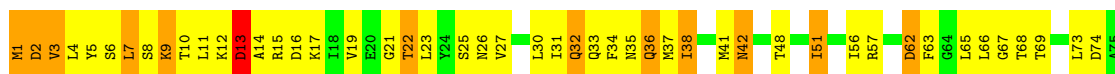
• Molecule 2: Intermediate capsid protein VP6

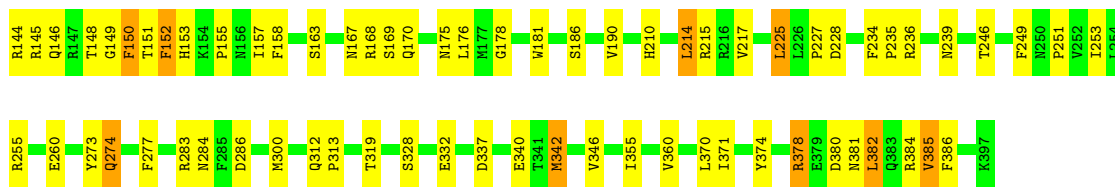
Chain G: 62% 30% 7%



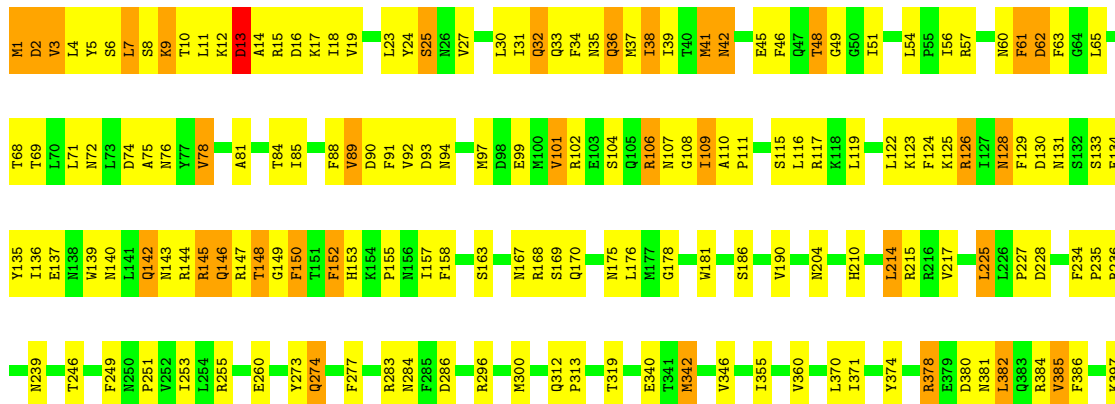
• Molecule 2: Intermediate capsid protein VP6

Chain H: 60% 32% 7%

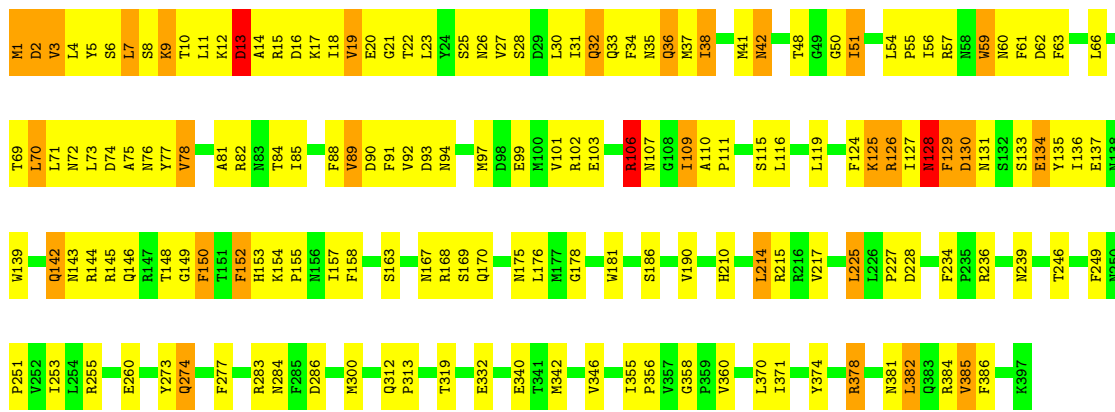




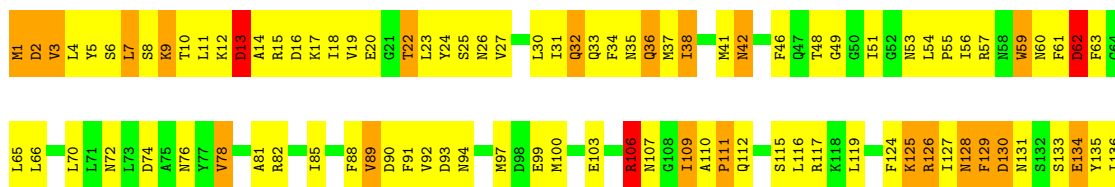
• Molecule 2: Intermediate capsid protein VP6

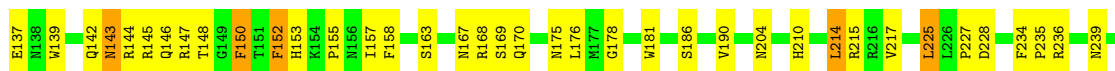


• Molecule 2: Intermediate capsid protein VP6



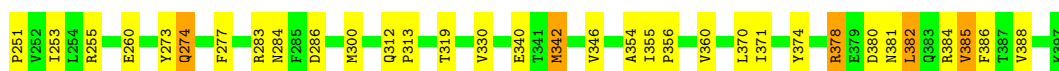
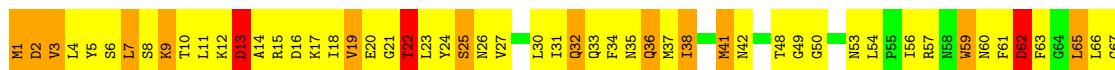
• Molecule 2: Intermediate capsid protein VP6





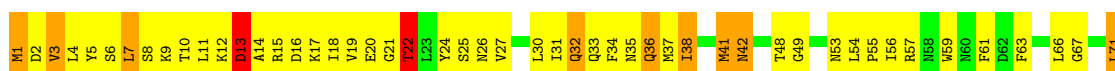
- Molecule 2: Intermediate capsid protein VP6

Chain L: 57% 35% 7%



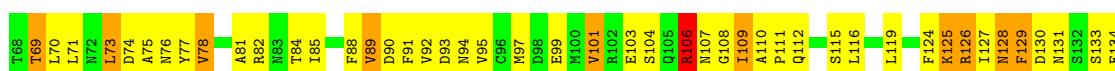
- Molecule 2: Intermediate capsid protein VP6

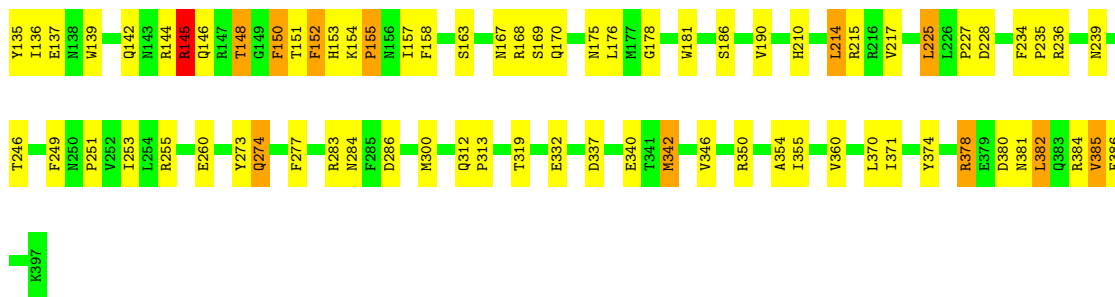
Chain M: 58% 35% 6%



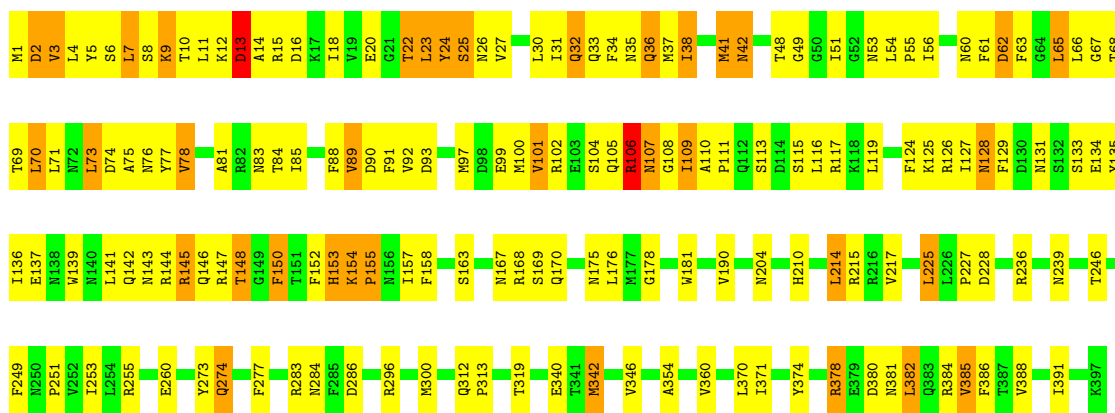
- Molecule 2: Intermediate capsid protein VP6

Chain N: 56% 34% 9%

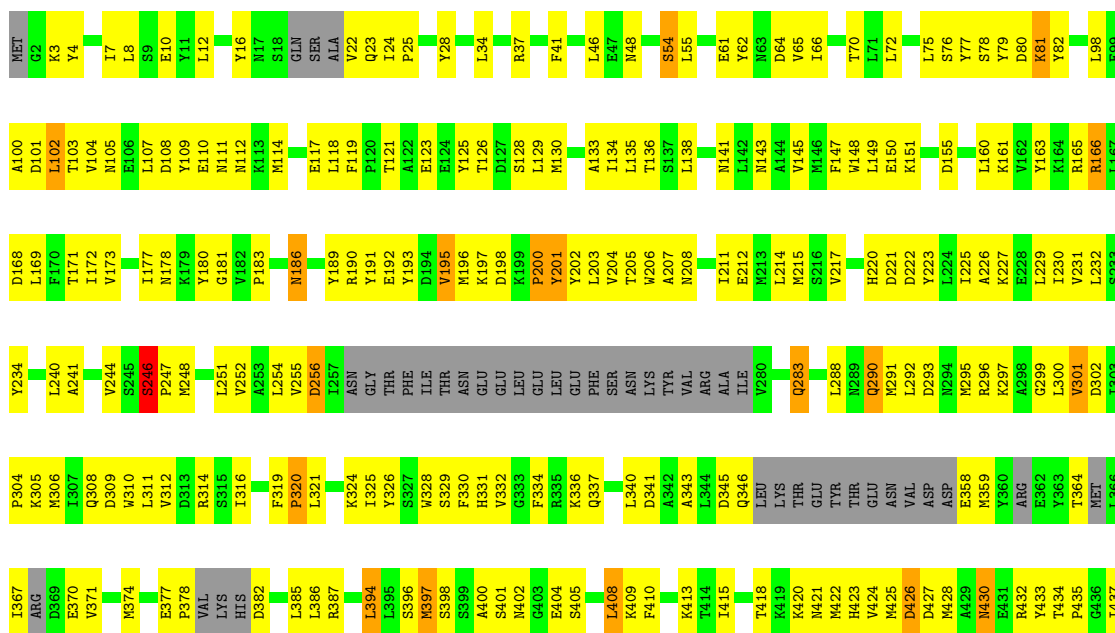




• Molecule 2: Intermediate capsid protein VP6



• Molecule 3: RNA-directed RNA polymerase



LYS	GLN	F884	T813	I736	I648	V580	ASN	I438
ILE	LEU	F885	R814	M737	Q649	I581	THR	P439
PRO	Y961	T886	L815	T740	D650	K582	MET	P440
ALA	L962	V887	S816	S741	V651	K583		V441
VAL	ILE		L819	I744	D654	I584		D444
THR	S964	A890	L820	T745	V655	Y586		K445
F1031	L965	H891	L820	G746	Y659	G587		P446
	G966	L892	K823	T748	Y659	A588		I447
	I967	I894	N824	S747	V589	A588		D520
	PRO	I894	N825	L748	V589	A588		P448
	LYS	Q897	I826	R749	N663	A590		L449
	ILE	K898	V827	L750	V666	S591		G450
	ASP	F899	S828	F751	K667	E593		R451
	ALA	M900	R829	E754		K594		R452
	ASP	P901	G830		S671	K594		
	THR	Q902	R831		THR	N530		R458
	THR	N906	I832		T758	Q595		T459
	VAL	V907	A832		T759	T531		R460
	VAL	Q908	E835		M760	K597		P466
	GLY	Q908	K836		S761	S601		L465
	SER	Q912	A837		T762	G537		L465
	LYS	Q912	K838		R680	I602		I461
	ILE	Q912	L839		K686	L605		I462
	Y981	R917	N840		I687	A686		F463
	S982	T918			F688	L607		K536
	ARG	Y919			P689			L466
	D984	Q920	A843		E789			P467
	K985	Q920	P844		D770			E468
	Y986	I921	P844		F771			Y469
	R987	E922	I845		I772			F470
	I988	E922	A691		T777			I471
	R988	D923	S846		T778			A472
	I988	D924	L847		V779			Q473
	S991	G925	E848		E781			A475
	Y992	S926			D780			A475
	Y992	R927	Q853		E781			V476
	L996	S928	I854		V782			V477
	L996	A929	L857		Y783			ASN
	I999	I930	L858		I784			ASN
	I1000	S931	L858		Q702			ASP
	Y1001	R932	L861		Q703			ALA
	G1002	L933	Q862		S704			ALA
	C1003	I934	K863		T705			K553
	Y1004	S935	K863		S790			V554
	Q1005	LYS	P864		L791			L555
	L1006	TYR	V865		Q709			L558
	F1009	S936	T866		I712			L559
	N1010	V939	R867		Y717			L560
	N1010	Y940	K868		I718			Y561
	D1013	I944	S869		Y718			K562
	L1014	E945	K871		R721			Q563
	GLU	Y948	L872		L722			T564
	K1016		T873		R723			Q565
	L1017		L874		Y719			I566
	I1018		N875		R729			N567
	R1019		R879		F808			L568
	PRO		D880		T807			S571
	ILE		F881		R808			THR
	PRO		K882		K889			ARC
	PHE		P883		M810			Y572
	LYS				L733			F509
	GLY				T734			LEU
					K735			LEU
								SER
								ASN

4 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	740.75Å 1198.07Å 1345.41Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 5.00	Depositor
% Data completeness (in resolution range)	(Not available) (30.00-5.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.49 (at 3.78Å)	Xtrriage
Refinement program	CNS	Depositor
R, R_{free}	0.293 , 0.296	Depositor
Wilson B-factor (Å ²)	167.1	Xtrriage
Anisotropy	0.043	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	62014	wwPDB-VP
Average B, all atoms (Å ²)	146.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.22% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section:
ZN

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	0.48	0/6491	0.84	9/8806 (0.1%)
1	B	0.50	0/6745	0.82	10/9149 (0.1%)
2	C	0.51	0/3232	0.78	5/4397 (0.1%)
2	D	0.51	0/3232	0.78	5/4397 (0.1%)
2	E	0.51	0/3232	0.77	5/4397 (0.1%)
2	F	0.50	0/3232	0.76	5/4397 (0.1%)
2	G	0.50	0/3232	0.77	5/4397 (0.1%)
2	H	0.50	0/3232	0.76	5/4397 (0.1%)
2	I	0.50	0/3232	0.77	5/4397 (0.1%)
2	J	0.51	0/3232	0.78	5/4397 (0.1%)
2	K	0.51	0/3232	0.77	5/4397 (0.1%)
2	L	0.51	0/3232	0.77	5/4397 (0.1%)
2	M	0.51	0/3232	0.77	5/4397 (0.1%)
2	N	0.51	0/3232	0.77	5/4397 (0.1%)
2	O	0.52	0/3232	0.79	7/4397 (0.2%)
3	W	0.41	0/8045	0.62	3/10847 (0.0%)
All	All	0.49	0/63297	0.77	89/85963 (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

There are no bond length outliers.

The worst 5 of 89 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	273	TYR	CB-CG-CD2	10.70	127.42	121.00
1	A	273	TYR	CB-CG-CD1	-9.86	115.08	121.00
1	A	273	TYR	CA-CB-CG	9.85	132.12	113.40
1	B	273	TYR	CB-CG-CD1	9.85	126.91	121.00
1	B	273	TYR	CB-CG-CD2	-9.21	115.48	121.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	273	TYR	Sidechain

5.2 Too-close contacts

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	6374	0	6394	1006	0
1	B	6624	0	6652	1133	0
2	C	3162	0	3101	170	0
2	D	3162	0	3101	179	0
2	E	3162	0	3101	164	0
2	F	3162	0	3101	160	0
2	G	3162	0	3101	149	0
2	H	3162	0	3101	175	0
2	I	3162	0	3101	215	0
2	J	3162	0	3101	187	0
2	K	3162	0	3101	166	0
2	L	3162	0	3101	178	0
2	M	3162	0	3101	173	0
2	N	3162	0	3101	191	0
2	O	3162	0	3101	163	0
3	W	7905	0	7966	543	0
4	C	1	0	0	0	0
4	F	1	0	0	0	0
4	I	1	0	0	0	0
4	L	1	0	0	0	0
4	O	1	0	0	0	0
All	All	62014	0	61325	4691	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 38.

The worst 5 of 4691 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:75:VAL:HG22	3:W:673:VAL:CG1	1.63	1.28
1:B:771:VAL:HB	1:B:809:PHE:HB3	1.23	1.18
1:B:75:VAL:CG2	3:W:673:VAL:HG12	1.72	1.18
1:A:333:VAL:HG11	1:A:380:LYS:HA	1.27	1.15
1:A:428:GLN:OE1	1:A:456:PHE:HB2	1.46	1.13

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	779/880 (88%)	431 (55%)	199 (26%)	149 (19%)	0	2
1	B	808/880 (92%)	458 (57%)	204 (25%)	146 (18%)	0	2
2	C	395/397 (100%)	318 (80%)	54 (14%)	23 (6%)	1	20
2	D	395/397 (100%)	315 (80%)	50 (13%)	30 (8%)	1	15
2	E	395/397 (100%)	323 (82%)	49 (12%)	23 (6%)	1	20
2	F	395/397 (100%)	324 (82%)	47 (12%)	24 (6%)	1	18
2	G	395/397 (100%)	323 (82%)	49 (12%)	23 (6%)	1	20
2	H	395/397 (100%)	323 (82%)	50 (13%)	22 (6%)	2	20
2	I	395/397 (100%)	320 (81%)	51 (13%)	24 (6%)	1	18
2	J	395/397 (100%)	320 (81%)	53 (13%)	22 (6%)	2	20
2	K	395/397 (100%)	318 (80%)	52 (13%)	25 (6%)	1	18
2	L	395/397 (100%)	322 (82%)	47 (12%)	26 (7%)	1	17

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	M	395/397 (100%)	322 (82%)	52 (13%)	21 (5%)	2	21
2	N	395/397 (100%)	317 (80%)	50 (13%)	28 (7%)	1	16
2	O	395/397 (100%)	317 (80%)	47 (12%)	31 (8%)	1	14
3	W	933/1089 (86%)	807 (86%)	93 (10%)	33 (4%)	3	28
All	All	7655/8010 (96%)	5858 (76%)	1147 (15%)	650 (8%)	1	12

5 of 650 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	101	LYS
1	A	130	GLN
1	A	193	SER
1	A	198	LYS
1	A	220	ALA

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	715/809 (88%)	603 (84%)	112 (16%)	2	15
1	B	744/809 (92%)	635 (85%)	109 (15%)	3	16
2	C	350/350 (100%)	325 (93%)	25 (7%)	14	41
2	D	350/350 (100%)	322 (92%)	28 (8%)	12	37
2	E	350/350 (100%)	325 (93%)	25 (7%)	14	41
2	F	350/350 (100%)	327 (93%)	23 (7%)	16	43
2	G	350/350 (100%)	327 (93%)	23 (7%)	16	43
2	H	350/350 (100%)	329 (94%)	21 (6%)	19	46
2	I	350/350 (100%)	327 (93%)	23 (7%)	16	43
2	J	350/350 (100%)	327 (93%)	23 (7%)	16	43
2	K	350/350 (100%)	327 (93%)	23 (7%)	16	43
2	L	350/350 (100%)	321 (92%)	29 (8%)	11	36

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	M	350/350 (100%)	325 (93%)	25 (7%)	14	41
2	N	350/350 (100%)	324 (93%)	26 (7%)	13	40
2	O	350/350 (100%)	329 (94%)	21 (6%)	19	46
3	W	885/990 (89%)	828 (94%)	57 (6%)	17	44
All	All	6894/7158 (96%)	6301 (91%)	593 (9%)	10	35

5 of 593 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	L	385	VAL
3	W	717	TYR
2	M	150	PHE
2	L	382	LEU
2	O	107	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 253 such sidechains are listed below:

Mol	Chain	Res	Type
2	G	128	ASN
3	W	36	ASN
2	I	146	GLN
2	O	345	ASN
3	W	604	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 5 ligands modelled in this entry, 5 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates [i](#)

EDS failed to run properly - this section is therefore empty.

6.4 Ligands [i](#)

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers [i](#)

EDS failed to run properly - this section is therefore empty.