



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:17 PM GMT

PDB ID : 4QYK  
Title : Crystal structure of a canine parvovirus variant  
Authors : Lukk, T.; Organtini, L.J.; Hafenstein, S.U.  
Deposited on : 2014-07-24  
Resolution : 3.50 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865



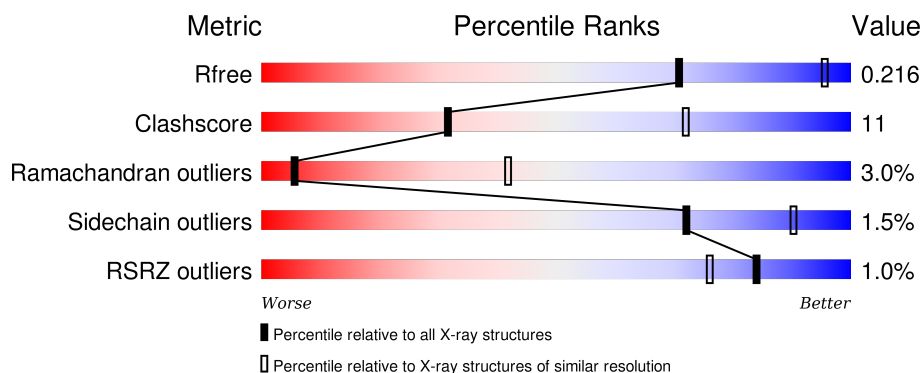
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	1051 (3.60-3.40)
Clashscore	102246	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)
RSRZ outliers	91569	1058 (3.60-3.40)





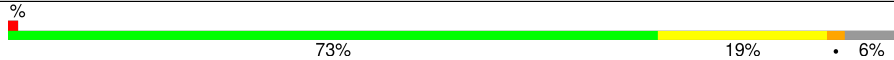
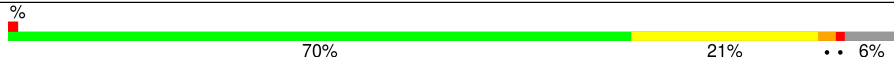
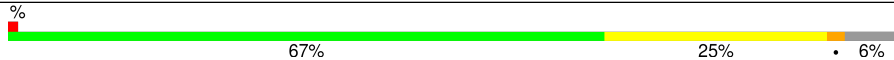
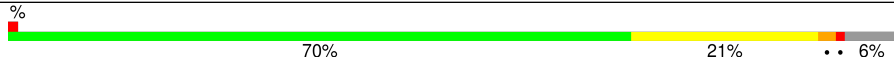
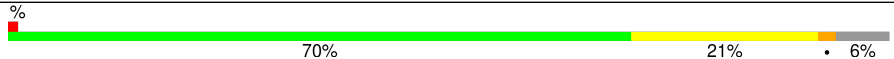
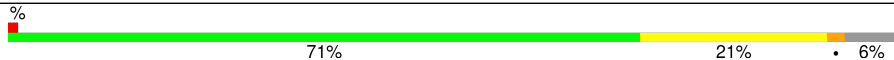
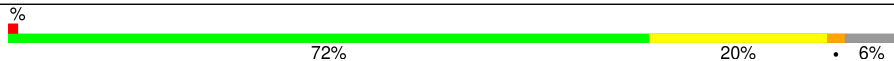
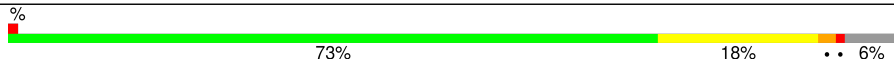


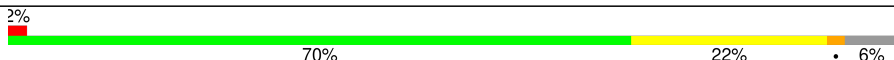
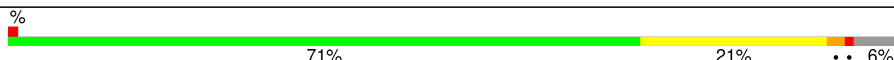
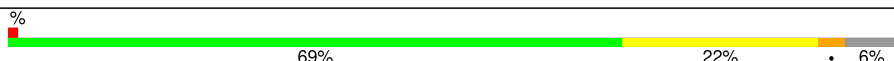
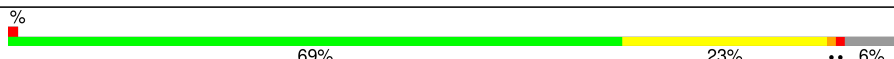
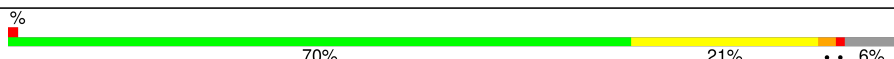

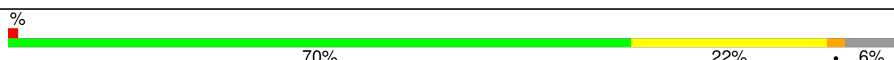
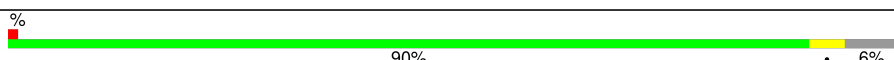
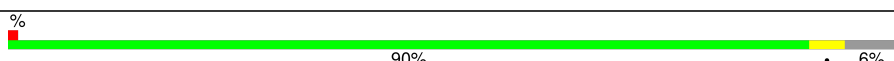
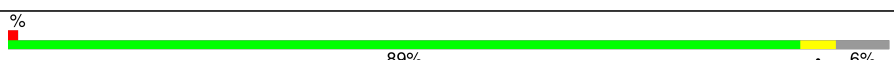
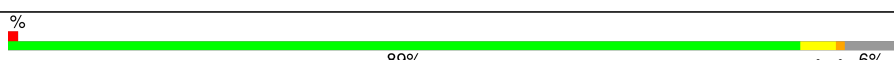
The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	584	<div> <div>%</div> <div> <div></div> <div>60%</div> <div>32%</div> <div>6%</div> </div> </div>
1	B	584	<div> <div></div> <div>62%</div> <div>30%</div> <div>6%</div> </div>
1	C	584	<div> <div>%</div> <div> <div></div> <div>60%</div> <div>31%</div> <div>6%</div> </div> </div>
1	D	584	<div> <div>%</div> <div> <div></div> <div>60%</div> <div>30%</div> <div>6%</div> </div> </div>
1	E	584	<div> <div>%</div> <div> <div></div> <div>68%</div> <div>23%</div> <div>6%</div> </div> </div>

*Continued on next page...*



Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	584	
1	G	584	
1	H	584	
1	I	584	
1	J	584	
1	K	584	
1	L	584	
1	M	584	
1	N	584	
1	O	584	
1	P	584	
1	Q	584	
1	R	584	
1	S	584	
1	T	584	
1	U	584	
1	V	584	
1	W	584	
1	X	584	
1	Y	584	
1	Z	584	
1	a	584	
1	b	584	
1	c	584	
1	d	584	



The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	MG	K	800	-	-	-	X
2	MG	Q	800	-	-	-	X
2	MG	T	800	-	-	-	X
2	MG	Y	601	-	-	-	X



## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 130590 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	A	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	E	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	J	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	N	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	Q	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	B	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	C	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	D	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	F	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	G	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	H	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	I	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	K	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	L	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	M	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	O	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	P	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	R	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	S	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	T	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	U	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	V	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	W	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	X	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	Y	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	Z	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	a	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	b	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	c	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			
1	d	548	Total	C	N	O	S	0	0	0
			4352	2767	740	830	15			

There are 150 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	60	TYR	GLU	CONFLICT	UNP P90456
A	104	GLU	GLN	CONFLICT	UNP P90456
A	426	ASP	ASN	CONFLICT	UNP P90456
A	509	GLN	GLU	CONFLICT	UNP P90456
A	555	VAL	ILE	CONFLICT	UNP P90456
E	60	TYR	GLU	CONFLICT	UNP P90456
E	104	GLU	GLN	CONFLICT	UNP P90456
E	426	ASP	ASN	CONFLICT	UNP P90456
E	509	GLN	GLU	CONFLICT	UNP P90456
E	555	VAL	ILE	CONFLICT	UNP P90456
J	60	TYR	GLU	CONFLICT	UNP P90456

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
J	104	GLU	GLN	CONFLICT	UNP P90456
J	426	ASP	ASN	CONFLICT	UNP P90456
J	509	GLN	GLU	CONFLICT	UNP P90456
J	555	VAL	ILE	CONFLICT	UNP P90456
N	60	TYR	GLU	CONFLICT	UNP P90456
N	104	GLU	GLN	CONFLICT	UNP P90456
N	426	ASP	ASN	CONFLICT	UNP P90456
N	509	GLN	GLU	CONFLICT	UNP P90456
N	555	VAL	ILE	CONFLICT	UNP P90456
Q	60	TYR	GLU	CONFLICT	UNP P90456
Q	104	GLU	GLN	CONFLICT	UNP P90456
Q	426	ASP	ASN	CONFLICT	UNP P90456
Q	509	GLN	GLU	CONFLICT	UNP P90456
Q	555	VAL	ILE	CONFLICT	UNP P90456
B	60	TYR	GLU	CONFLICT	UNP P90456
B	104	GLU	GLN	CONFLICT	UNP P90456
B	426	ASP	ASN	CONFLICT	UNP P90456
B	509	GLN	GLU	CONFLICT	UNP P90456
B	555	VAL	ILE	CONFLICT	UNP P90456
C	60	TYR	GLU	CONFLICT	UNP P90456
C	104	GLU	GLN	CONFLICT	UNP P90456
C	426	ASP	ASN	CONFLICT	UNP P90456
C	509	GLN	GLU	CONFLICT	UNP P90456
C	555	VAL	ILE	CONFLICT	UNP P90456
D	60	TYR	GLU	CONFLICT	UNP P90456
D	104	GLU	GLN	CONFLICT	UNP P90456
D	426	ASP	ASN	CONFLICT	UNP P90456
D	509	GLN	GLU	CONFLICT	UNP P90456
D	555	VAL	ILE	CONFLICT	UNP P90456
F	60	TYR	GLU	CONFLICT	UNP P90456
F	104	GLU	GLN	CONFLICT	UNP P90456
F	426	ASP	ASN	CONFLICT	UNP P90456
F	509	GLN	GLU	CONFLICT	UNP P90456
F	555	VAL	ILE	CONFLICT	UNP P90456
G	60	TYR	GLU	CONFLICT	UNP P90456
G	104	GLU	GLN	CONFLICT	UNP P90456
G	426	ASP	ASN	CONFLICT	UNP P90456
G	509	GLN	GLU	CONFLICT	UNP P90456
G	555	VAL	ILE	CONFLICT	UNP P90456
H	60	TYR	GLU	CONFLICT	UNP P90456
H	104	GLU	GLN	CONFLICT	UNP P90456
H	426	ASP	ASN	CONFLICT	UNP P90456

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	509	GLN	GLU	CONFLICT	UNP P90456
H	555	VAL	ILE	CONFLICT	UNP P90456
I	60	TYR	GLU	CONFLICT	UNP P90456
I	104	GLU	GLN	CONFLICT	UNP P90456
I	426	ASP	ASN	CONFLICT	UNP P90456
I	509	GLN	GLU	CONFLICT	UNP P90456
I	555	VAL	ILE	CONFLICT	UNP P90456
K	60	TYR	GLU	CONFLICT	UNP P90456
K	104	GLU	GLN	CONFLICT	UNP P90456
K	426	ASP	ASN	CONFLICT	UNP P90456
K	509	GLN	GLU	CONFLICT	UNP P90456
K	555	VAL	ILE	CONFLICT	UNP P90456
L	60	TYR	GLU	CONFLICT	UNP P90456
L	104	GLU	GLN	CONFLICT	UNP P90456
L	426	ASP	ASN	CONFLICT	UNP P90456
L	509	GLN	GLU	CONFLICT	UNP P90456
L	555	VAL	ILE	CONFLICT	UNP P90456
M	60	TYR	GLU	CONFLICT	UNP P90456
M	104	GLU	GLN	CONFLICT	UNP P90456
M	426	ASP	ASN	CONFLICT	UNP P90456
M	509	GLN	GLU	CONFLICT	UNP P90456
M	555	VAL	ILE	CONFLICT	UNP P90456
O	60	TYR	GLU	CONFLICT	UNP P90456
O	104	GLU	GLN	CONFLICT	UNP P90456
O	426	ASP	ASN	CONFLICT	UNP P90456
O	509	GLN	GLU	CONFLICT	UNP P90456
O	555	VAL	ILE	CONFLICT	UNP P90456
P	60	TYR	GLU	CONFLICT	UNP P90456
P	104	GLU	GLN	CONFLICT	UNP P90456
P	426	ASP	ASN	CONFLICT	UNP P90456
P	509	GLN	GLU	CONFLICT	UNP P90456
P	555	VAL	ILE	CONFLICT	UNP P90456
R	60	TYR	GLU	CONFLICT	UNP P90456
R	104	GLU	GLN	CONFLICT	UNP P90456
R	426	ASP	ASN	CONFLICT	UNP P90456
R	509	GLN	GLU	CONFLICT	UNP P90456
R	555	VAL	ILE	CONFLICT	UNP P90456
S	60	TYR	GLU	CONFLICT	UNP P90456
S	104	GLU	GLN	CONFLICT	UNP P90456
S	426	ASP	ASN	CONFLICT	UNP P90456
S	509	GLN	GLU	CONFLICT	UNP P90456
S	555	VAL	ILE	CONFLICT	UNP P90456

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
T	60	TYR	GLU	CONFLICT	UNP P90456
T	104	GLU	GLN	CONFLICT	UNP P90456
T	426	ASP	ASN	CONFLICT	UNP P90456
T	509	GLN	GLU	CONFLICT	UNP P90456
T	555	VAL	ILE	CONFLICT	UNP P90456
U	60	TYR	GLU	CONFLICT	UNP P90456
U	104	GLU	GLN	CONFLICT	UNP P90456
U	426	ASP	ASN	CONFLICT	UNP P90456
U	509	GLN	GLU	CONFLICT	UNP P90456
U	555	VAL	ILE	CONFLICT	UNP P90456
V	60	TYR	GLU	CONFLICT	UNP P90456
V	104	GLU	GLN	CONFLICT	UNP P90456
V	426	ASP	ASN	CONFLICT	UNP P90456
V	509	GLN	GLU	CONFLICT	UNP P90456
V	555	VAL	ILE	CONFLICT	UNP P90456
W	60	TYR	GLU	CONFLICT	UNP P90456
W	104	GLU	GLN	CONFLICT	UNP P90456
W	426	ASP	ASN	CONFLICT	UNP P90456
W	509	GLN	GLU	CONFLICT	UNP P90456
W	555	VAL	ILE	CONFLICT	UNP P90456
X	60	TYR	GLU	CONFLICT	UNP P90456
X	104	GLU	GLN	CONFLICT	UNP P90456
X	426	ASP	ASN	CONFLICT	UNP P90456
X	509	GLN	GLU	CONFLICT	UNP P90456
X	555	VAL	ILE	CONFLICT	UNP P90456
Y	60	TYR	GLU	CONFLICT	UNP P90456
Y	104	GLU	GLN	CONFLICT	UNP P90456
Y	426	ASP	ASN	CONFLICT	UNP P90456
Y	509	GLN	GLU	CONFLICT	UNP P90456
Y	555	VAL	ILE	CONFLICT	UNP P90456
Z	60	TYR	GLU	CONFLICT	UNP P90456
Z	104	GLU	GLN	CONFLICT	UNP P90456
Z	426	ASP	ASN	CONFLICT	UNP P90456
Z	509	GLN	GLU	CONFLICT	UNP P90456
Z	555	VAL	ILE	CONFLICT	UNP P90456
a	60	TYR	GLU	CONFLICT	UNP P90456
a	104	GLU	GLN	CONFLICT	UNP P90456
a	426	ASP	ASN	CONFLICT	UNP P90456
a	509	GLN	GLU	CONFLICT	UNP P90456
a	555	VAL	ILE	CONFLICT	UNP P90456
b	60	TYR	GLU	CONFLICT	UNP P90456
b	104	GLU	GLN	CONFLICT	UNP P90456

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
b	426	ASP	ASN	CONFLICT	UNP P90456
b	509	GLN	GLU	CONFLICT	UNP P90456
b	555	VAL	ILE	CONFLICT	UNP P90456
c	60	TYR	GLU	CONFLICT	UNP P90456
c	104	GLU	GLN	CONFLICT	UNP P90456
c	426	ASP	ASN	CONFLICT	UNP P90456
c	509	GLN	GLU	CONFLICT	UNP P90456
c	555	VAL	ILE	CONFLICT	UNP P90456
d	60	TYR	GLU	CONFLICT	UNP P90456
d	104	GLU	GLN	CONFLICT	UNP P90456
d	426	ASP	ASN	CONFLICT	UNP P90456
d	509	GLN	GLU	CONFLICT	UNP P90456
d	555	VAL	ILE	CONFLICT	UNP P90456

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	P	1	Total Mg 1 1	0	0
2	K	1	Total Mg 1 1	0	0
2	B	1	Total Mg 1 1	0	0
2	c	1	Total Mg 1 1	0	0
2	W	1	Total Mg 1 1	0	0
2	N	1	Total Mg 1 1	0	0
2	X	1	Total Mg 1 1	0	0
2	S	1	Total Mg 1 1	0	0
2	J	2	Total Mg 2 2	0	0
2	E	2	Total Mg 2 2	0	0
2	A	1	Total Mg 1 1	0	0
2	R	1	Total Mg 1 1	0	0
2	M	1	Total Mg 1 1	0	0

*Continued on next page...*



*Continued from previous page...*

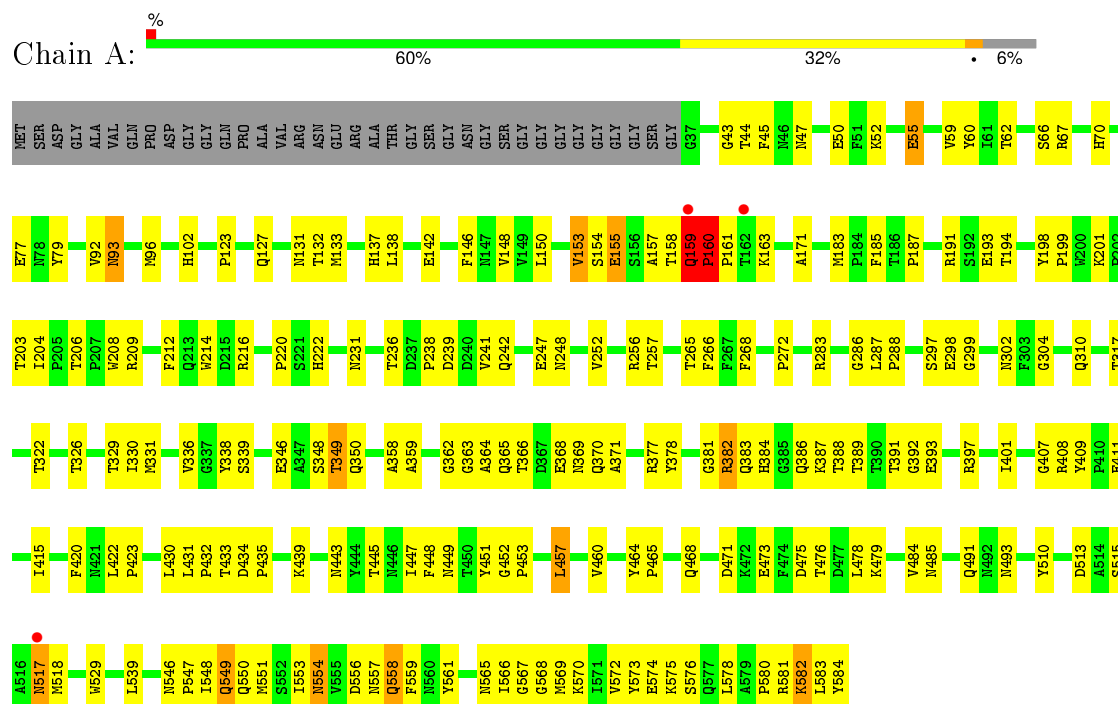
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	1	Total 1	Mg 1	0	0
2	I	1	Total 1	Mg 1	0	0
2	Z	1	Total 1	Mg 1	0	0
2	L	1	Total 1	Mg 1	0	0
2	G	1	Total 1	Mg 1	0	0
2	Q	1	Total 1	Mg 1	0	0
2	d	2	Total 2	Mg 2	0	0
2	H	2	Total 2	Mg 2	0	0
2	C	1	Total 1	Mg 1	0	0
2	T	1	Total 1	Mg 1	0	0
2	O	1	Total 1	Mg 1	0	0
2	Y	2	Total 2	Mg 2	0	0



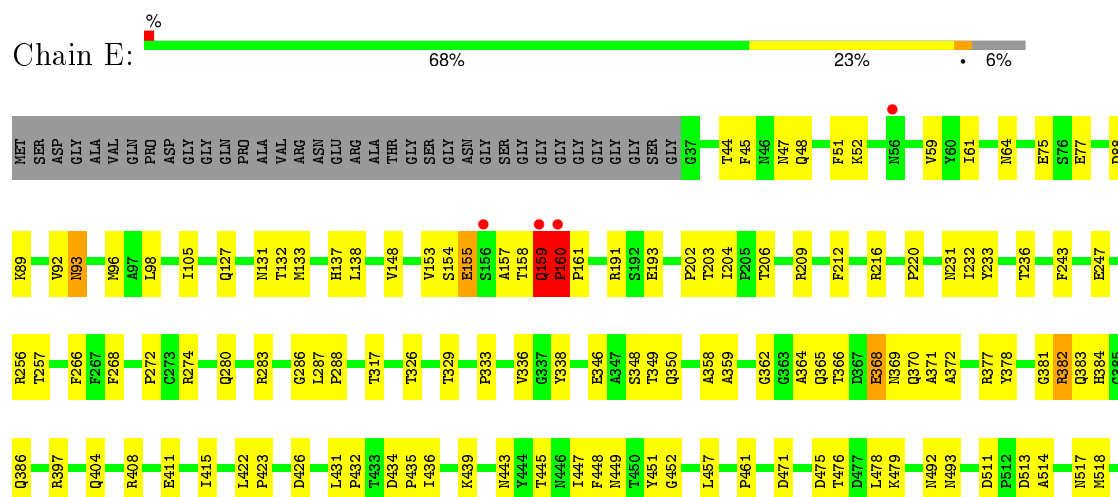
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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.

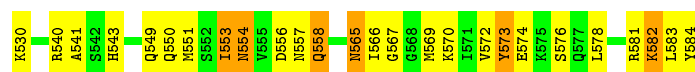
#### • Molecule 1: Capsid protein VP1



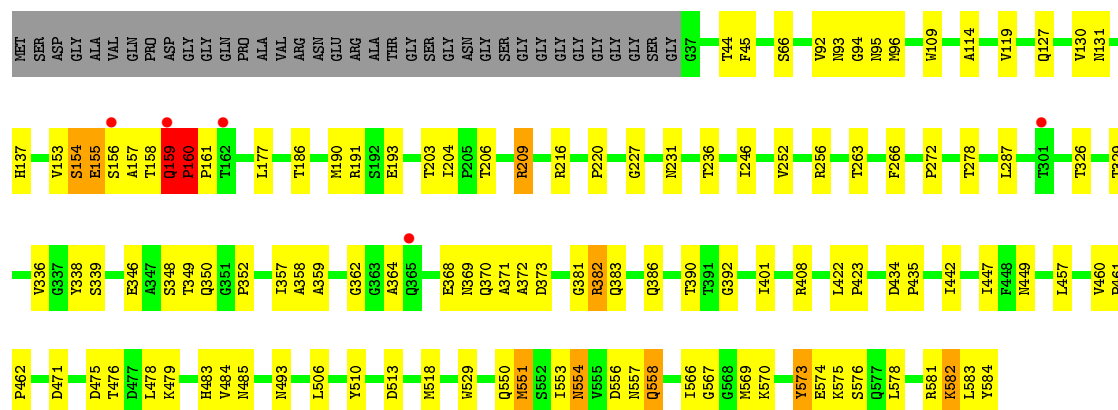
#### • Molecule 1: Capsid protein VP1



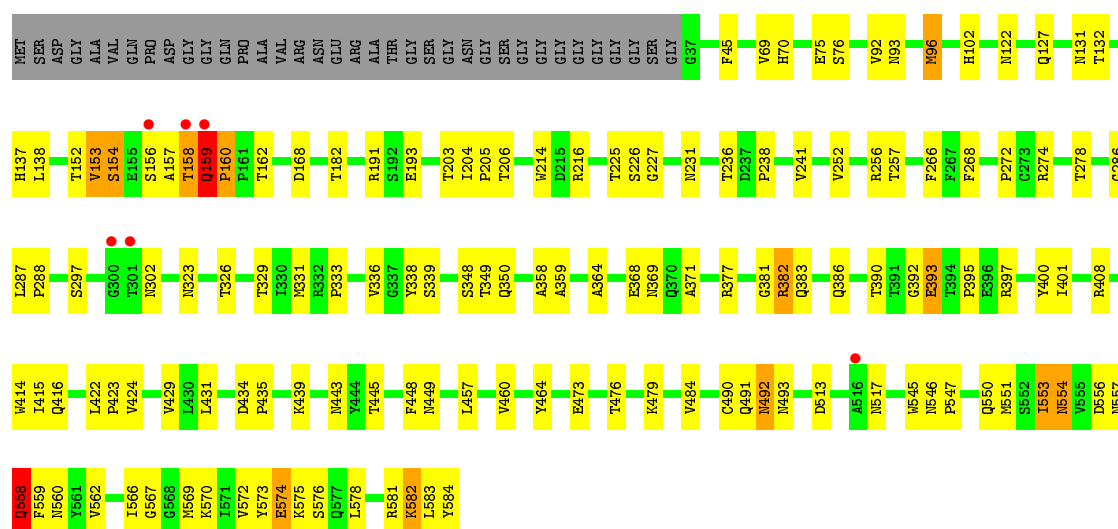




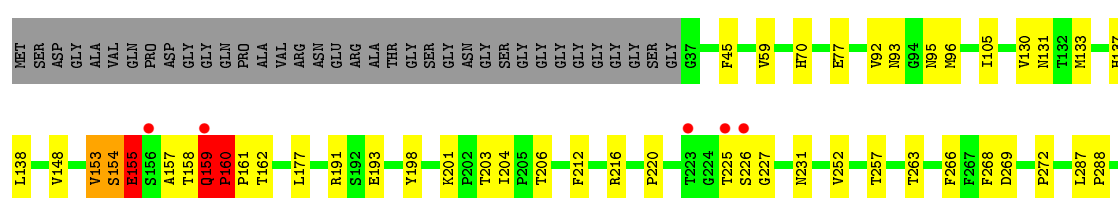
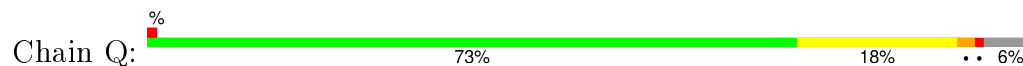
• Molecule 1: Capsid protein VP1



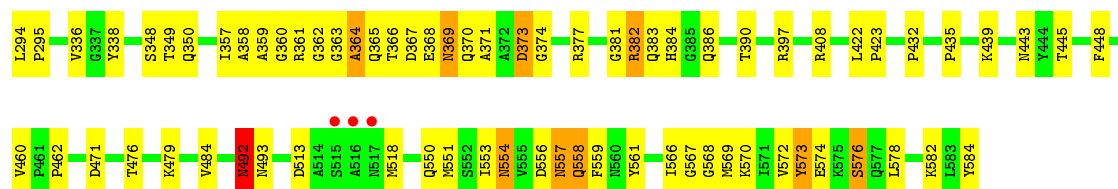
• Molecule 1: Capsid protein VP1



• Molecule 1: Capsid protein VP1

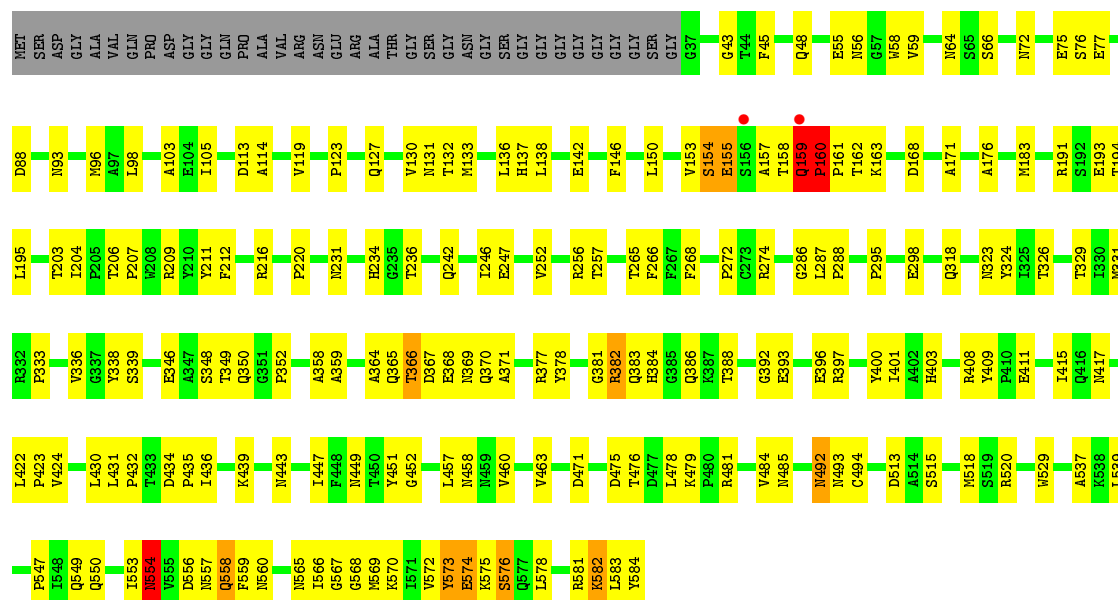






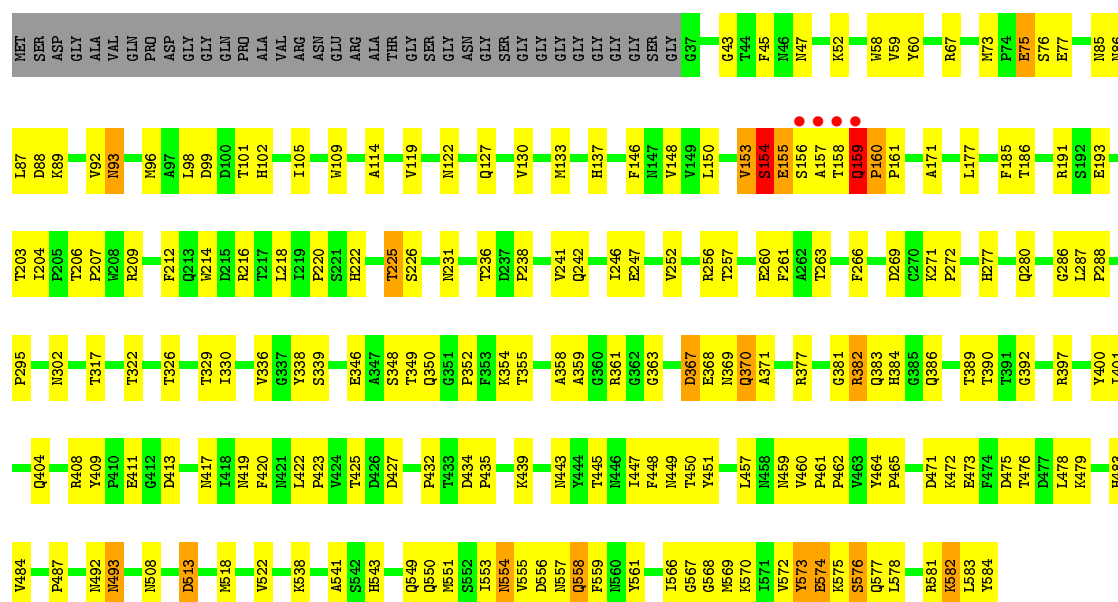
• Molecule 1: Capsid protein VP1

Chain B: 62% 30% 6%



• Molecule 1: Capsid protein VP1

Chain C: 60% 31% 6%

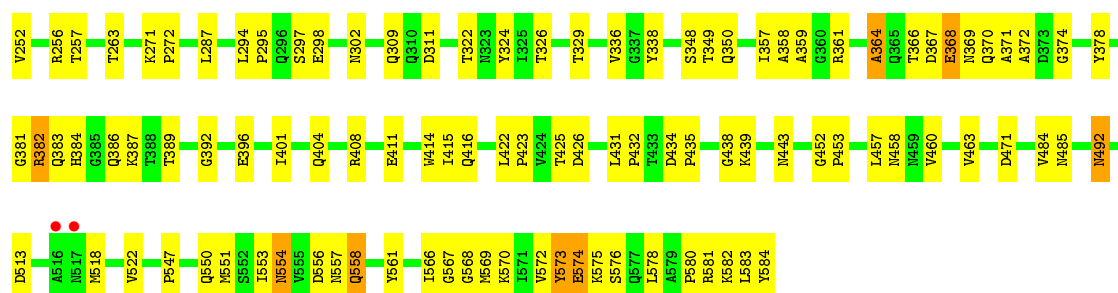




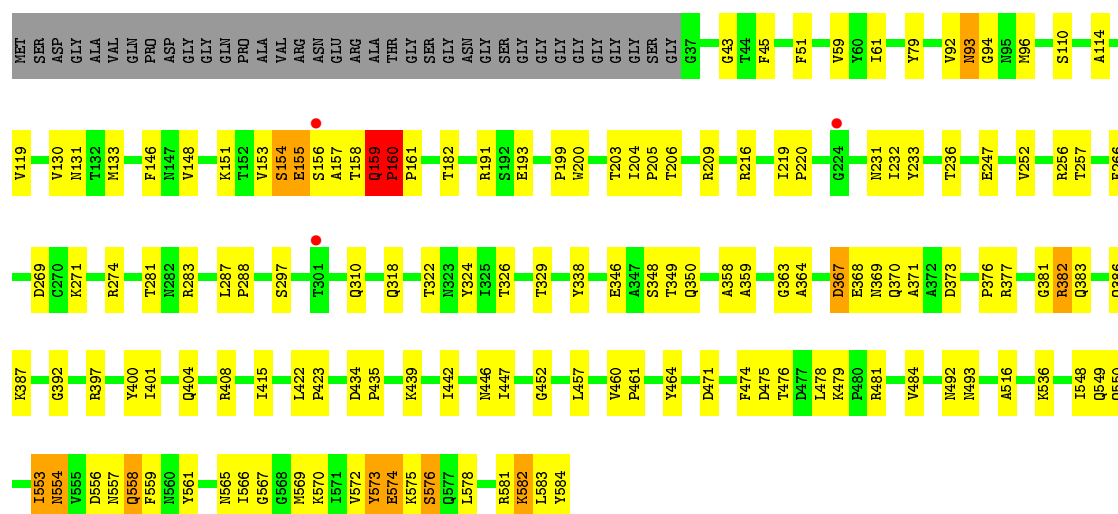
Chain D:  60% 30% 6%



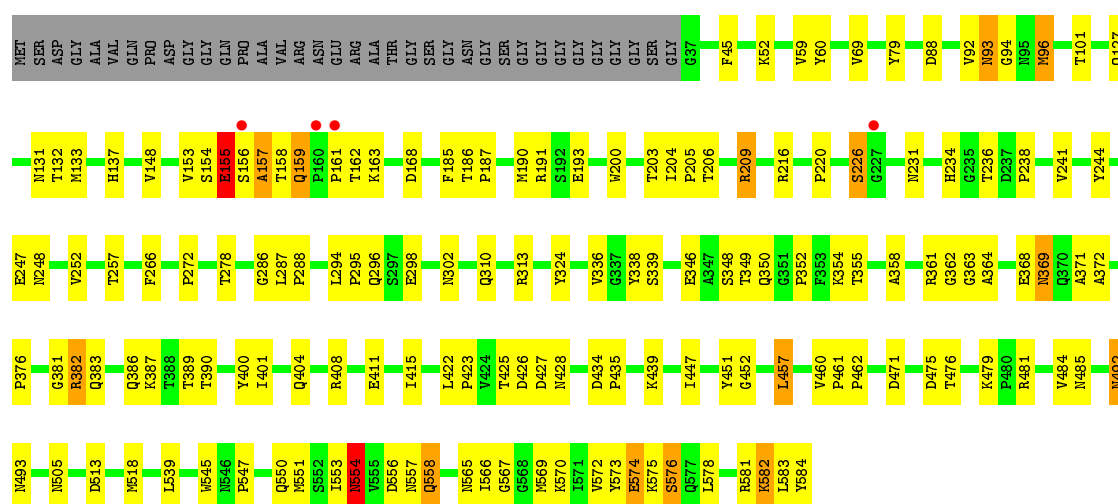




• Molecule 1: Capsid protein VP1



• Molecule 1: Capsid protein VP1



• Molecule 1: Capsid protein VP1



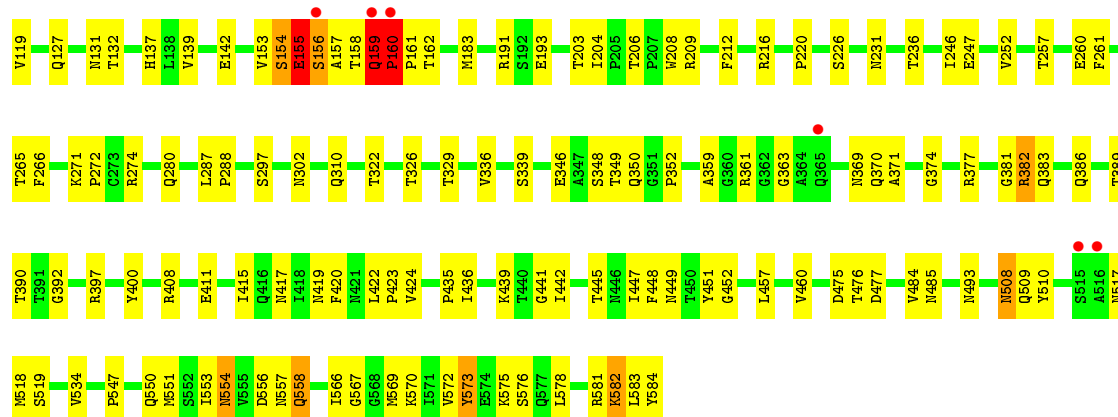






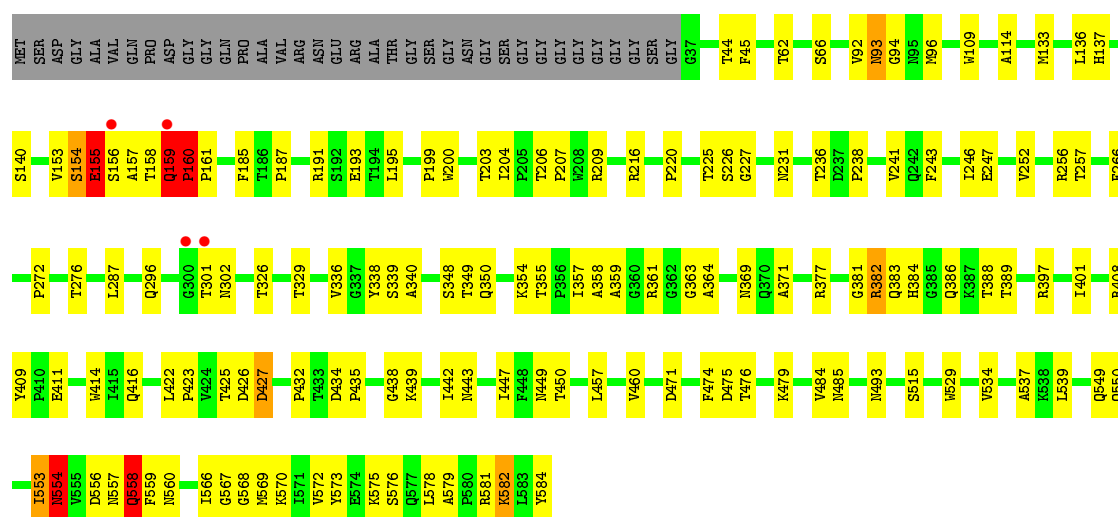






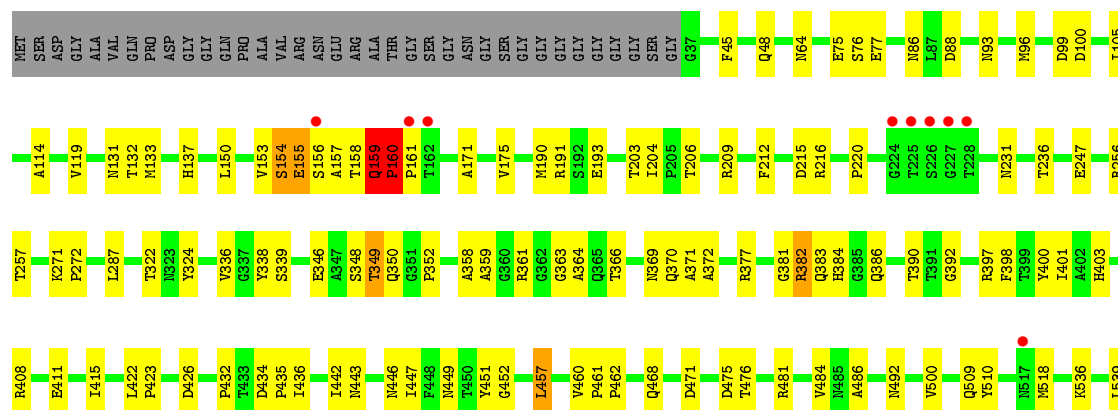
• Molecule 1: Capsid protein VP1

Chain S: 69% 23% 6%



• Molecule 1: Capsid protein VP1

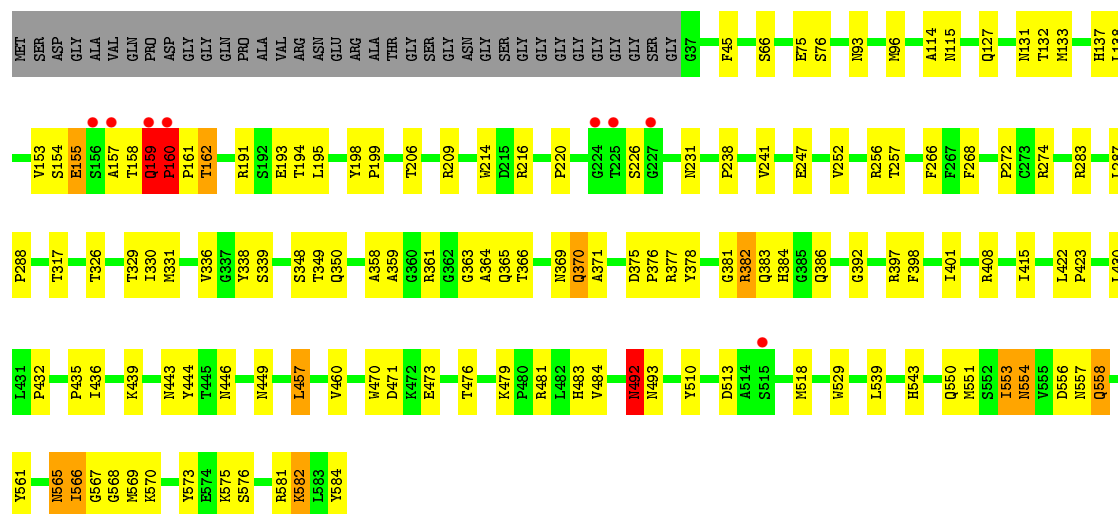
Chain T: 70% 22% 6%



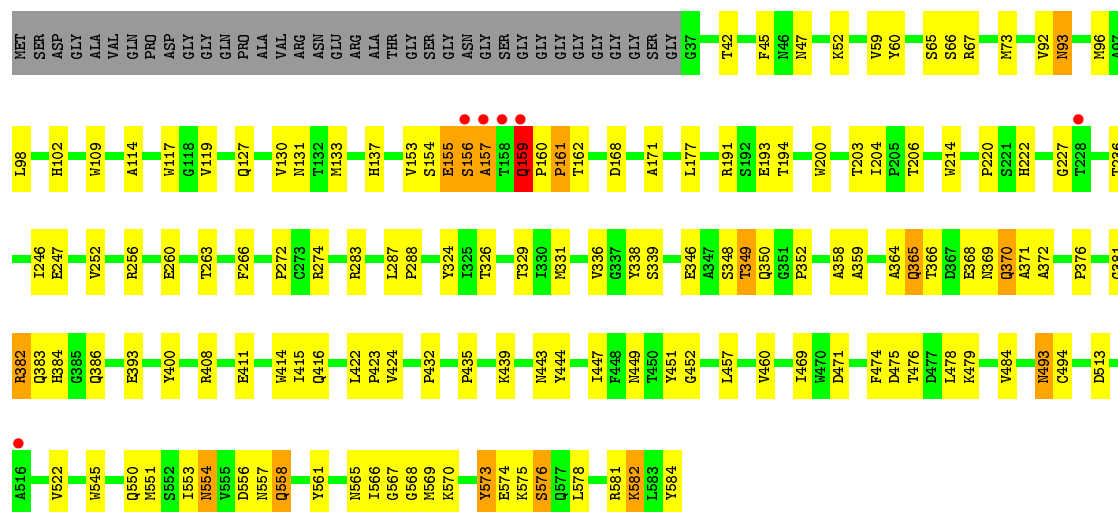




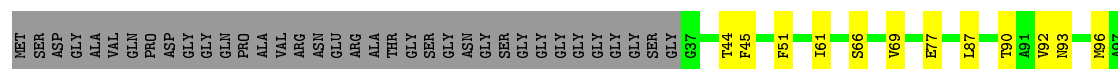
• Molecule 1: Capsid protein VP1



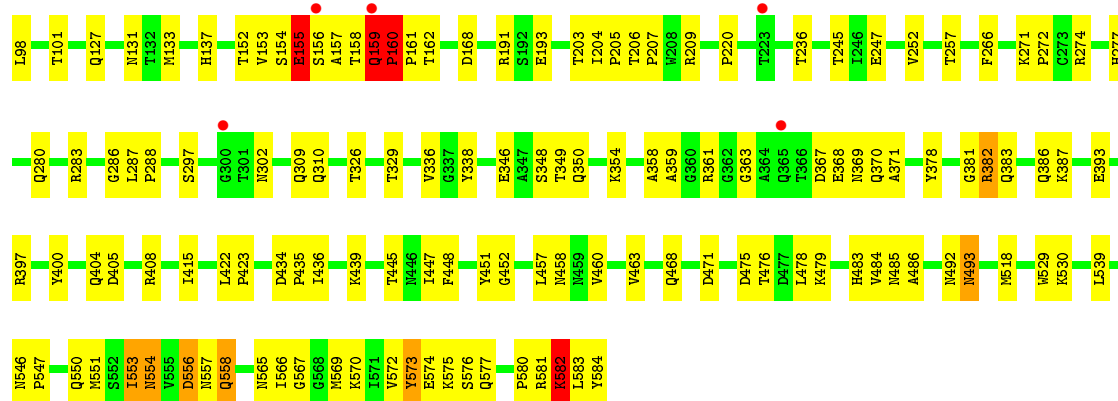
• Molecule 1: Capsid protein VP1



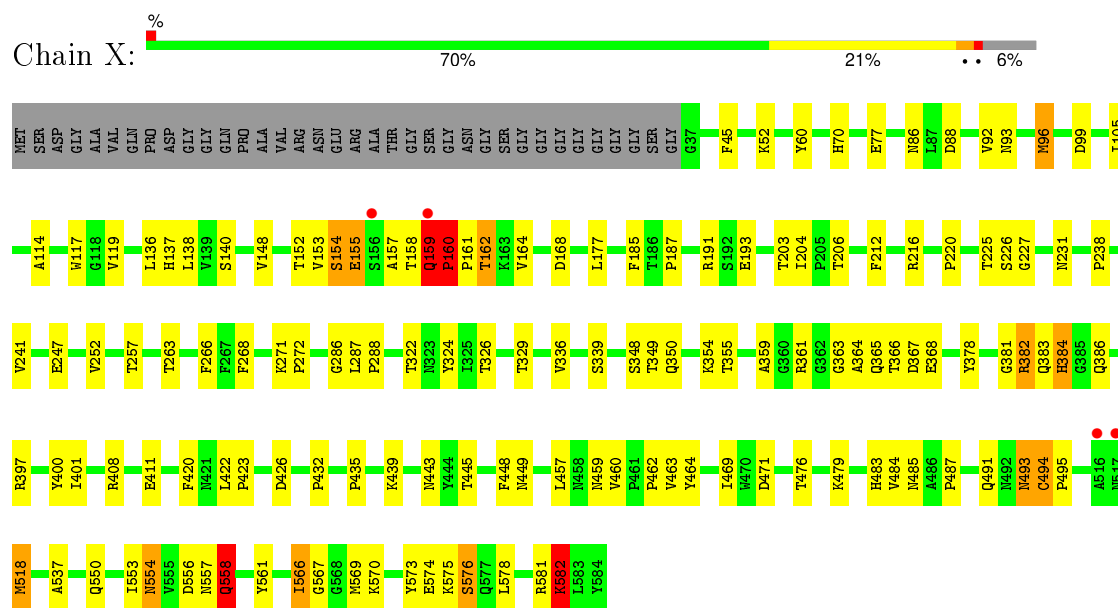
• Molecule 1: Capsid protein VP1



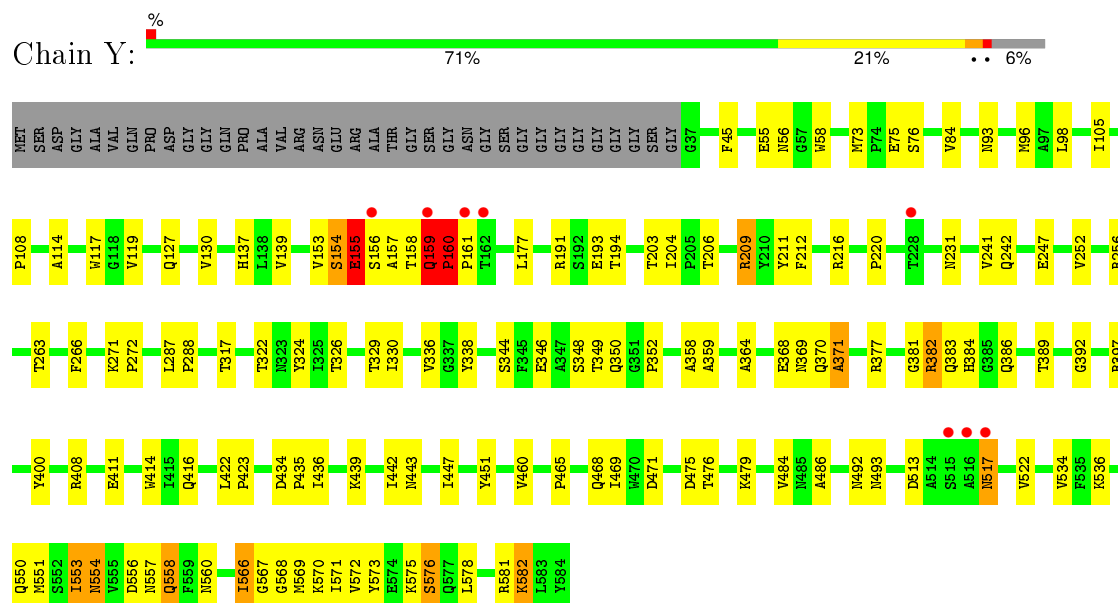




• Molecule 1: Capsid protein VP1



• Molecule 1: Capsid protein VP1

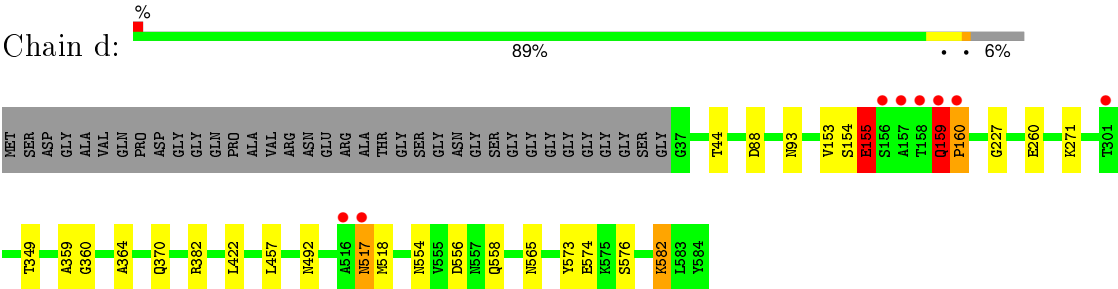








● Molecule 1: Capsid protein VP1





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	453.10Å 453.10Å 319.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.99 – 3.50 49.99 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.99-3.50) 96.4 (49.99-3.50)	Depositor EDS
$R_{merge}$	0.22	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.65 (at 3.48Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, $R_{free}$	0.180 , 0.221 0.175 , 0.216	Depositor DCC
$R_{free}$ test set	19783 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	42.7	Xtriage
Anisotropy	0.298	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 43.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	2 of 410636 reflections (0.000%)	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	130590	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.48% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

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



## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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

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	0.66	0/4483	0.70	2/6134 (0.0%)
1	B	0.63	0/4483	0.73	5/6134 (0.1%)
1	C	0.59	0/4483	0.73	4/6134 (0.1%)
1	D	0.58	0/4483	0.73	5/6134 (0.1%)
1	E	0.62	0/4483	0.73	4/6134 (0.1%)
1	F	0.63	1/4483 (0.0%)	0.84	9/6134 (0.1%)
1	G	0.59	0/4483	0.74	6/6134 (0.1%)
1	H	0.57	0/4483	0.72	3/6134 (0.0%)
1	I	0.64	0/4483	0.71	1/6134 (0.0%)
1	J	0.61	0/4483	0.71	5/6134 (0.1%)
1	K	0.67	0/4483	0.74	2/6134 (0.0%)
1	L	0.61	0/4483	0.73	3/6134 (0.0%)
1	M	0.66	0/4483	0.74	5/6134 (0.1%)
1	N	0.66	0/4483	0.72	6/6134 (0.1%)
1	O	0.58	0/4483	0.71	3/6134 (0.0%)
1	P	0.61	0/4483	0.73	4/6134 (0.1%)
1	Q	0.61	0/4483	0.72	5/6134 (0.1%)
1	R	0.63	0/4483	0.73	3/6134 (0.0%)
1	S	0.63	1/4483 (0.0%)	0.76	9/6134 (0.1%)
1	T	0.65	0/4483	0.72	4/6134 (0.1%)
1	U	0.60	0/4483	0.73	4/6134 (0.1%)
1	V	0.60	0/4483	0.70	1/6134 (0.0%)
1	W	0.55	0/4483	0.72	3/6134 (0.0%)
1	X	0.61	1/4483 (0.0%)	0.76	7/6134 (0.1%)
1	Y	0.56	0/4483	0.73	5/6134 (0.1%)
1	Z	0.60	2/4483 (0.0%)	0.74	4/6134 (0.1%)
1	a	0.55	0/4483	0.75	2/6134 (0.0%)
1	b	0.57	0/4483	0.73	3/6134 (0.0%)
1	c	0.58	0/4483	0.73	4/6134 (0.1%)
1	d	0.60	0/4483	0.76	7/6134 (0.1%)
All	All	0.61	5/134490 (0.0%)	0.73	128/184020 (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	C	0	2
1	D	0	2
1	E	0	1
1	F	0	1
1	G	0	1
1	H	0	1
1	I	0	1
1	L	0	2
1	M	0	3
1	N	0	1
1	Q	0	1
1	R	0	2
1	S	0	2
1	U	0	1
1	W	0	1
1	X	0	1
1	Y	0	2
1	Z	0	1
1	a	0	1
1	b	0	1
1	c	0	2
1	d	0	3
All	All	0	34

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	Z	222	HIS	CG-CD2	-8.08	1.22	1.35
1	F	426	ASP	CB-CG	-7.62	1.35	1.51
1	S	438	GLY	C-O	-5.81	1.14	1.23
1	X	426	ASP	CB-CG	-5.05	1.41	1.51
1	Z	273	CYS	CB-SG	-5.00	1.73	1.81

The worst 5 of 128 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	426	ASP	CB-CG-OD2	-27.25	93.78	118.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	a	159	GLN	C-N-CD	-18.61	79.65	120.60
1	X	159	GLN	C-N-CD	-16.06	85.27	120.60
1	F	426	ASP	CB-CG-OD1	15.33	132.10	118.30
1	C	159	GLN	C-N-CD	-13.92	89.97	120.60

There are no chirality outliers.

5 of 34 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	159	GLN	Peptide
1	C	159	GLN	Peptide
1	E	368	GLU	Peptide
1	N	226	SER	Peptide
1	Q	159	GLN	Peptide

## 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	4352	0	4143	207	0
1	B	4352	0	4143	216	0
1	C	4352	0	4143	176	2
1	D	4352	0	4143	190	4
1	E	4352	0	4143	128	0
1	F	4352	0	4143	131	0
1	G	4352	0	4143	122	3
1	H	4352	0	4143	125	0
1	I	4352	0	4143	131	0
1	J	4352	0	4143	83	2
1	K	4352	0	4143	122	0
1	L	4352	0	4143	130	0
1	M	4352	0	4143	118	0
1	N	4352	0	4143	102	2
1	O	4352	0	4143	98	5
1	P	4352	0	4143	104	0
1	Q	4352	0	4143	81	5
1	R	4352	0	4143	120	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	S	4352	0	4143	92	2
1	T	4352	0	4143	110	0
1	U	4352	0	4143	106	3
1	V	4352	0	4143	128	0
1	W	4352	0	4143	124	0
1	X	4352	0	4145	93	4
1	Y	4352	0	4143	111	2
1	Z	4352	0	4143	87	2
1	a	4352	0	4143	0	3
1	b	4352	0	4143	0	0
1	c	4352	0	4143	0	1
1	d	4352	0	4143	0	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	2	0	0	0	0
2	G	1	0	0	0	0
2	H	2	0	0	0	0
2	I	1	0	0	0	0
2	J	2	0	0	0	0
2	K	1	0	0	0	0
2	L	1	0	0	0	0
2	M	1	0	0	0	0
2	N	1	0	0	0	0
2	O	1	0	0	0	0
2	P	1	0	0	0	0
2	Q	1	0	0	0	0
2	R	1	0	0	0	0
2	S	1	0	0	0	0
2	T	1	0	0	0	0
2	W	1	0	0	0	0
2	X	1	0	0	0	0
2	Y	2	0	0	0	0
2	Z	1	0	0	0	0
2	c	1	0	0	0	0
2	d	2	0	0	0	0
All	All	130590	0	124292	2724	20

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 11.

The worst 5 of 2724 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:L:553:ILE:HA	1:L:554:ASN:HB2	1.35	1.08
1:B:553:ILE:HA	1:B:554:ASN:HB2	1.42	1.07
1:V:154:SER:HA	1:V:155:GLU:HB2	1.36	1.06
1:R:154:SER:HA	1:R:155:GLU:HG3	1.33	1.06
1:Z:553:ILE:HA	1:Z:554:ASN:HB2	1.37	1.05

The worst 5 of 20 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)
1:O:209:ARG:NH1	1:Y:288:PRO:O[8_555]	2.04	0.16
1:N:182:THR:O	1:X:582:LYS:NZ[8_555]	2.06	0.14
1:S:584:TYR:OH	1:X:471:ASP:OD1[8_555]	2.07	0.13
1:O:191:ARG:NH1	1:Y:287:LEU:O[8_555]	2.10	0.10
1:C:287:LEU:O	1:Z:191:ARG:NH1[8_555]	2.10	0.10

## 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	546/584 (94%)	500 (92%)	31 (6%)	15 (3%)	6	44
1	B	546/584 (94%)	496 (91%)	32 (6%)	18 (3%)	5	39
1	C	546/584 (94%)	493 (90%)	39 (7%)	14 (3%)	7	45
1	D	546/584 (94%)	497 (91%)	32 (6%)	17 (3%)	5	41
1	E	546/584 (94%)	497 (91%)	33 (6%)	16 (3%)	6	42
1	F	546/584 (94%)	496 (91%)	33 (6%)	17 (3%)	5	41
1	G	546/584 (94%)	497 (91%)	32 (6%)	17 (3%)	5	41
1	H	546/584 (94%)	491 (90%)	38 (7%)	17 (3%)	5	41
1	I	546/584 (94%)	499 (91%)	34 (6%)	13 (2%)	7	47
1	J	546/584 (94%)	498 (91%)	33 (6%)	15 (3%)	6	44

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	K	546/584 (94%)	499 (91%)	31 (6%)	16 (3%)	6	42
1	L	546/584 (94%)	494 (90%)	35 (6%)	17 (3%)	5	41
1	M	546/584 (94%)	491 (90%)	40 (7%)	15 (3%)	6	44
1	N	546/584 (94%)	497 (91%)	34 (6%)	15 (3%)	6	44
1	O	546/584 (94%)	492 (90%)	35 (6%)	19 (4%)	4	38
1	P	546/584 (94%)	497 (91%)	32 (6%)	17 (3%)	5	41
1	Q	546/584 (94%)	496 (91%)	34 (6%)	16 (3%)	6	42
1	R	546/584 (94%)	490 (90%)	39 (7%)	17 (3%)	5	41
1	S	546/584 (94%)	500 (92%)	32 (6%)	14 (3%)	7	45
1	T	546/584 (94%)	501 (92%)	30 (6%)	15 (3%)	6	44
1	U	546/584 (94%)	496 (91%)	32 (6%)	18 (3%)	5	39
1	V	546/584 (94%)	492 (90%)	37 (7%)	17 (3%)	5	41
1	W	546/584 (94%)	491 (90%)	40 (7%)	15 (3%)	6	44
1	X	546/584 (94%)	495 (91%)	34 (6%)	17 (3%)	5	41
1	Y	546/584 (94%)	498 (91%)	30 (6%)	18 (3%)	5	39
1	Z	546/584 (94%)	497 (91%)	31 (6%)	18 (3%)	5	39
1	a	546/584 (94%)	497 (91%)	33 (6%)	16 (3%)	6	42
1	b	546/584 (94%)	489 (90%)	41 (8%)	16 (3%)	6	42
1	c	546/584 (94%)	499 (91%)	31 (6%)	16 (3%)	6	42
1	d	546/584 (94%)	498 (91%)	32 (6%)	16 (3%)	6	42
All	All	16380/17520 (94%)	14873 (91%)	1020 (6%)	487 (3%)	5	42

5 of 487 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	153	VAL
1	A	155	GLU
1	A	159	GLN
1	A	160	PRO
1	A	349	THR

### 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	476/495 (96%)	468 (98%)	8 (2%)	68	89
1	B	476/495 (96%)	472 (99%)	4 (1%)	86	95
1	C	476/495 (96%)	467 (98%)	9 (2%)	65	87
1	D	476/495 (96%)	470 (99%)	6 (1%)	76	91
1	E	476/495 (96%)	471 (99%)	5 (1%)	80	92
1	F	476/495 (96%)	470 (99%)	6 (1%)	76	91
1	G	476/495 (96%)	468 (98%)	8 (2%)	68	89
1	H	476/495 (96%)	468 (98%)	8 (2%)	68	89
1	I	476/495 (96%)	463 (97%)	13 (3%)	52	83
1	J	476/495 (96%)	471 (99%)	5 (1%)	80	92
1	K	476/495 (96%)	467 (98%)	9 (2%)	65	87
1	L	476/495 (96%)	467 (98%)	9 (2%)	65	87
1	M	476/495 (96%)	470 (99%)	6 (1%)	76	91
1	N	476/495 (96%)	470 (99%)	6 (1%)	76	91
1	O	476/495 (96%)	473 (99%)	3 (1%)	90	97
1	P	476/495 (96%)	472 (99%)	4 (1%)	86	95
1	Q	476/495 (96%)	470 (99%)	6 (1%)	76	91
1	R	476/495 (96%)	468 (98%)	8 (2%)	68	89
1	S	476/495 (96%)	465 (98%)	11 (2%)	58	85
1	T	476/495 (96%)	468 (98%)	8 (2%)	68	89
1	U	476/495 (96%)	468 (98%)	8 (2%)	68	89
1	V	476/495 (96%)	471 (99%)	5 (1%)	80	92
1	W	476/495 (96%)	468 (98%)	8 (2%)	68	89
1	X	476/495 (96%)	468 (98%)	8 (2%)	68	89
1	Y	476/495 (96%)	474 (100%)	2 (0%)	93	98
1	Z	476/495 (96%)	471 (99%)	5 (1%)	80	92
1	a	476/495 (96%)	471 (99%)	5 (1%)	80	92
1	b	476/495 (96%)	468 (98%)	8 (2%)	68	89
1	c	476/495 (96%)	463 (97%)	13 (3%)	52	83

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	d	476/495 (96%)	465 (98%)	11 (2%)	58	85
All	All	14280/14850 (96%)	14065 (98%)	215 (2%)	72	90

5 of 215 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	490	CYS
1	R	509	GLN
1	c	384	HIS
1	L	573	TYR
1	O	457	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 102 such sidechains are listed below:

Mol	Chain	Res	Type
1	L	384	HIS
1	P	558	GLN
1	a	554	ASN
1	L	558	GLN
1	O	557	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 30 ligands modelled in this entry, 30 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	548/584 (93%)	-0.64	3 (0%) 91 88	3, 17, 60, 119	0
1	B	548/584 (93%)	-0.63	2 (0%) 93 90	4, 19, 64, 124	0
1	C	548/584 (93%)	-0.61	4 (0%) 89 82	3, 20, 64, 121	0
1	D	548/584 (93%)	-0.58	6 (1%) 82 73	5, 20, 63, 131	0
1	E	548/584 (93%)	-0.65	4 (0%) 89 82	4, 18, 62, 120	0
1	F	548/584 (93%)	-0.66	3 (0%) 91 88	3, 18, 61, 118	0
1	G	548/584 (93%)	-0.58	8 (1%) 76 67	4, 20, 62, 126	0
1	H	548/584 (93%)	-0.57	3 (0%) 91 88	4, 19, 62, 118	0
1	I	548/584 (93%)	-0.66	4 (0%) 89 82	5, 19, 62, 127	0
1	J	548/584 (93%)	-0.64	5 (0%) 85 78	3, 19, 61, 122	0
1	K	548/584 (93%)	-0.62	8 (1%) 76 67	3, 17, 61, 122	0
1	L	548/584 (93%)	-0.64	3 (0%) 91 88	4, 19, 60, 125	0
1	M	548/584 (93%)	-0.61	6 (1%) 82 73	3, 18, 62, 121	0
1	N	548/584 (93%)	-0.64	6 (1%) 82 73	5, 19, 62, 126	0
1	O	548/584 (93%)	-0.53	5 (0%) 85 78	3, 20, 64, 123	0
1	P	548/584 (93%)	-0.62	4 (0%) 89 82	3, 18, 61, 119	0
1	Q	548/584 (93%)	-0.59	8 (1%) 76 67	4, 19, 61, 130	0
1	R	548/584 (93%)	-0.63	6 (1%) 82 73	4, 18, 62, 121	0
1	S	548/584 (93%)	-0.64	4 (0%) 89 82	5, 19, 61, 131	0
1	T	548/584 (93%)	-0.61	9 (1%) 74 65	4, 18, 64, 120	0
1	U	548/584 (93%)	-0.58	8 (1%) 76 67	3, 18, 65, 118	0
1	V	548/584 (93%)	-0.59	6 (1%) 82 73	4, 19, 62, 126	0
1	W	548/584 (93%)	-0.56	5 (0%) 85 78	4, 21, 66, 121	0
1	X	548/584 (93%)	-0.64	4 (0%) 89 82	4, 19, 60, 131	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	Y	548/584 (93%)	-0.55	8 (1%) 76 67	4, 19, 64, 123	0
1	Z	548/584 (93%)	-0.59	3 (0%) 91 88	3, 20, 64, 116	0
1	a	548/584 (93%)	-0.55	8 (1%) 76 67	4, 21, 67, 123	0
1	b	548/584 (93%)	-0.56	7 (1%) 79 70	3, 20, 66, 124	0
1	c	548/584 (93%)	-0.54	7 (1%) 79 70	4, 22, 61, 130	0
1	d	548/584 (93%)	-0.62	8 (1%) 76 67	4, 19, 64, 125	0
All	All	16440/17520 (93%)	-0.61	165 (1%) 84 76	3, 19, 64, 131	0

The worst 5 of 165 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	O	156	SER	4.3
1	T	156	SER	4.0
1	Q	515	SER	3.8
1	X	156	SER	3.7
1	K	156	SER	3.7

## 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 carbohydrates in this entry.

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	MG	K	800	1/1	0.93	0.36	4.05	39,39,39,39	0
2	MG	T	800	1/1	0.90	0.31	3.54	43,43,43,43	0
2	MG	Q	800	1/1	0.93	0.33	3.40	41,41,41,41	0

Continued on next page...



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	MG	Y	601	1/1	0.94	0.29	2.60	47,47,47,47	0
2	MG	O	800	1/1	0.90	0.17	1.84	41,41,41,41	0
2	MG	M	800	1/1	0.89	0.27	1.75	41,41,41,41	0
2	MG	Y	602	1/1	0.93	0.25	1.40	52,52,52,52	0
2	MG	I	800	1/1	0.97	0.24	1.28	43,43,43,43	0
2	MG	B	800	1/1	0.96	0.24	1.09	41,41,41,41	0
2	MG	D	800	1/1	0.94	0.26	1.04	43,43,43,43	0
2	MG	H	601	1/1	0.92	0.20	0.87	36,36,36,36	0
2	MG	P	800	1/1	0.95	0.19	0.85	39,39,39,39	0
2	MG	J	601	1/1	0.96	0.19	0.62	40,40,40,40	0
2	MG	A	800	1/1	0.95	0.19	0.45	32,32,32,32	0
2	MG	G	800	1/1	0.93	0.18	0.42	44,44,44,44	0
2	MG	d	602	1/1	0.90	0.18	0.26	56,56,56,56	0
2	MG	W	800	1/1	0.92	0.18	0.13	47,47,47,47	0
2	MG	H	602	1/1	0.96	0.18	0.08	54,54,54,54	0
2	MG	E	601	1/1	0.90	0.16	0.07	40,40,40,40	0
2	MG	C	800	1/1	0.94	0.16	0.04	41,41,41,41	0
2	MG	J	602	1/1	0.95	0.16	-0.23	38,38,38,38	0
2	MG	L	800	1/1	0.93	0.14	-0.34	41,41,41,41	0
2	MG	N	800	1/1	0.88	0.16	-0.65	39,39,39,39	0
2	MG	Z	800	1/1	0.97	0.12	-0.74	47,47,47,47	0
2	MG	R	800	1/1	0.91	0.14	-0.77	37,37,37,37	0
2	MG	E	602	1/1	0.97	0.10	-1.42	38,38,38,38	0
2	MG	S	800	1/1	0.96	0.09	-1.71	39,39,39,39	0
2	MG	d	601	1/1	0.96	0.09	-1.71	37,37,37,37	0
2	MG	X	800	1/1	0.97	0.10	-2.16	43,43,43,43	0
2	MG	c	800	1/1	0.96	0.16	-	53,53,53,53	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.