



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2024 – 01:08 PM EST

PDB ID : 3KZ4  
Title : Crystal Structure of the Rotavirus Double Layered Particle  
Authors : McClain, B.; Settembre, E.C.; Bellamy, A.R.; Harrison, S.C.  
Deposited on : 2009-12-07  
Resolution : 3.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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  
Xtrriage (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) : 2.36

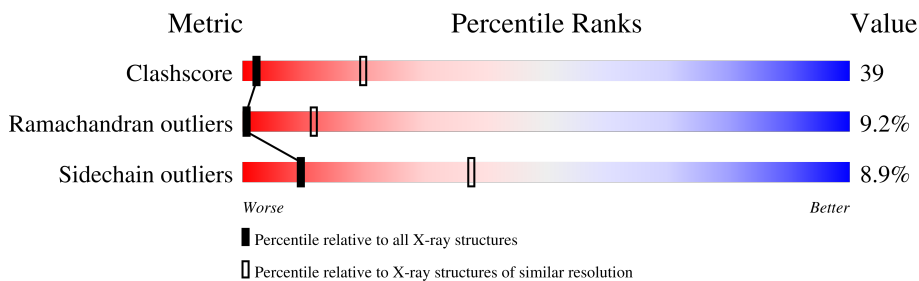
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)



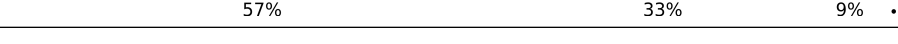
The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria 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% 50% 21% 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	 58% 33% 9%
2	J	397	 58% 34% 8%
2	K	397	 59% 33% 8%
2	L	397	 57% 35% 7%
2	M	397	 58% 35% 6%
2	N	397	 56% 34% 9%
2	O	397	 57% 33% 9%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 54109 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 Inner capsid protein VP2.

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
2	O	397	Total	C	N	O	S	0	0	0
			3162	2001	550	596	15			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

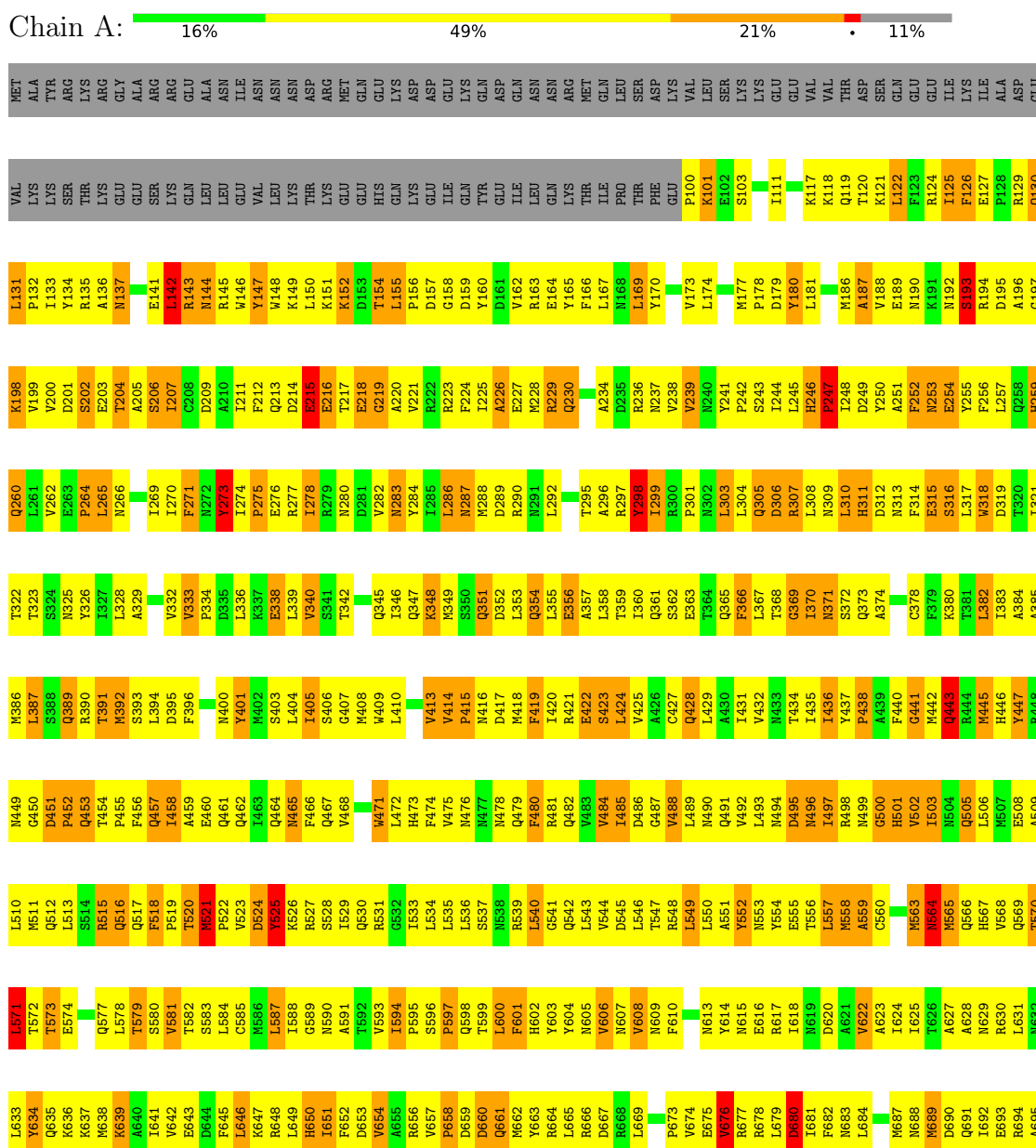
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	1	Total	Zn	0	0
			1	1		
3	F	1	Total	Zn	0	0
			1	1		
3	I	1	Total	Zn	0	0
			1	1		
3	L	1	Total	Zn	0	0
			1	1		
3	O	1	Total	Zn	0	0
			1	1		

### 3 Residue-property plots

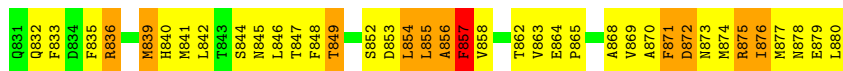
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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: Inner capsid protein VP2



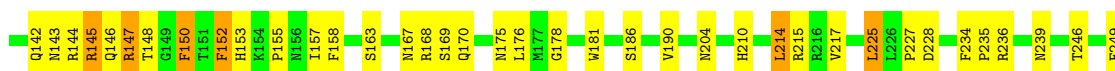
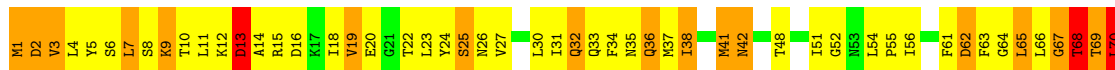




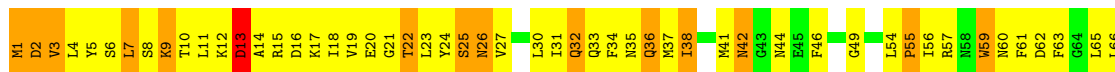
• Molecule 2: Intermediate capsid protein VP6



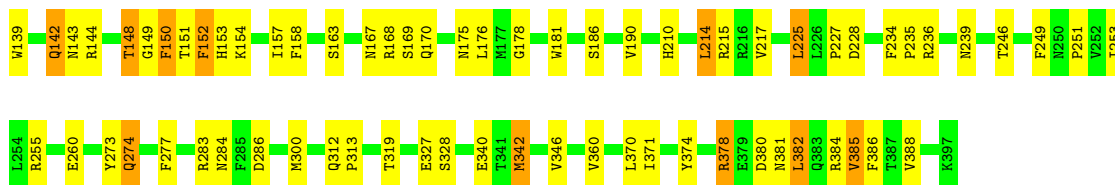
• Molecule 2: Intermediate capsid protein VP6



• Molecule 2: Intermediate capsid protein VP6

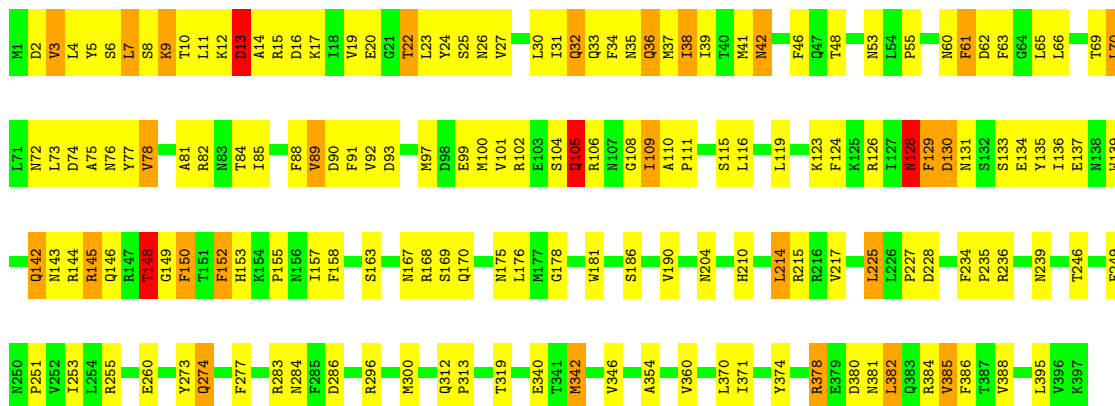






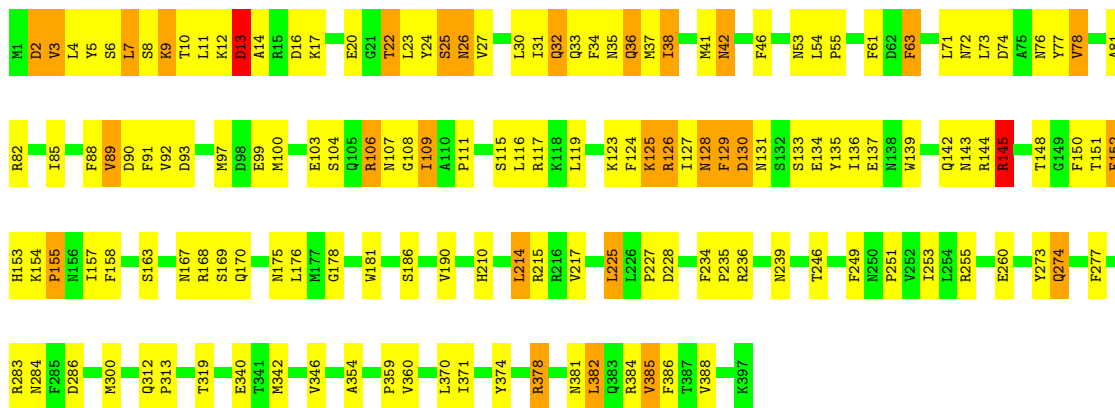
- Molecule 2: Intermediate capsid protein VP6

Chain F: 59% 34% 7%



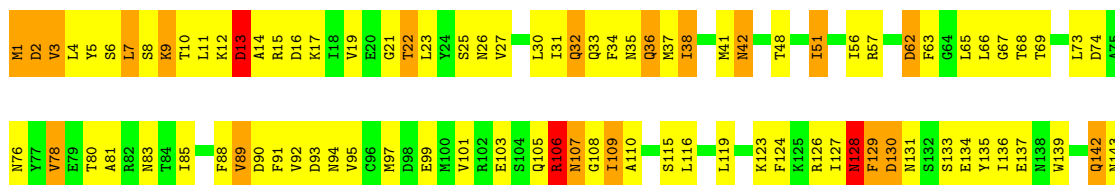
- Molecule 2: Intermediate capsid protein VP6

Chain G: 62% 30% 7%



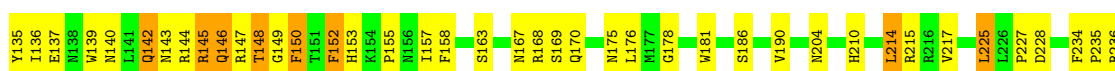
- Molecule 2: Intermediate capsid protein VP6

Chain H: 60% 32% 7%

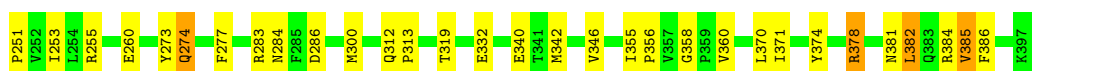
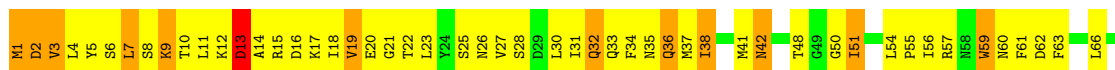




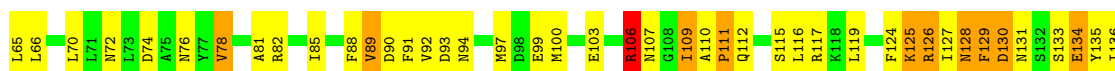
- Molecule 2: Intermediate capsid protein VP6

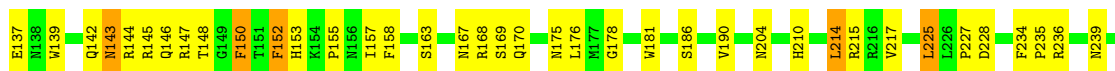


- Molecule 2: Intermediate capsid protein VP6

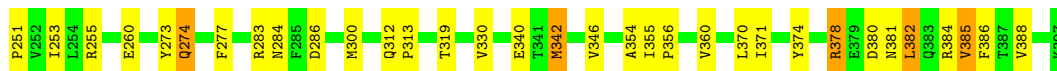
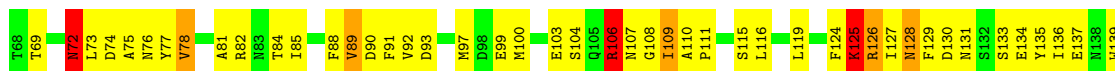
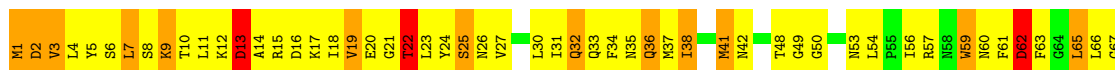


- Molecule 2: Intermediate capsid protein VP6

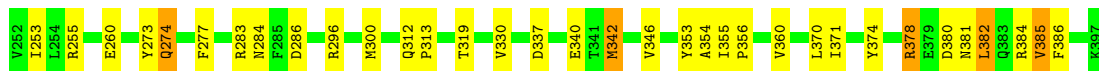
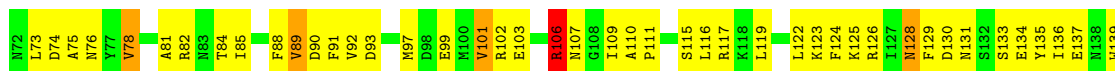
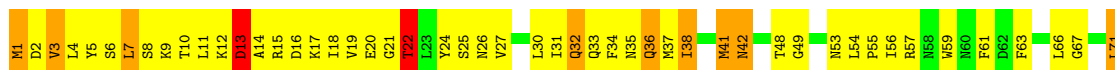




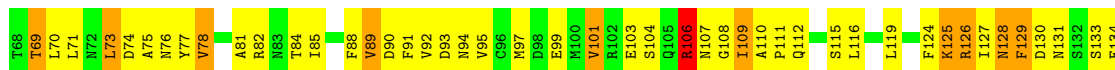
• Molecule 2: Intermediate capsid protein VP6

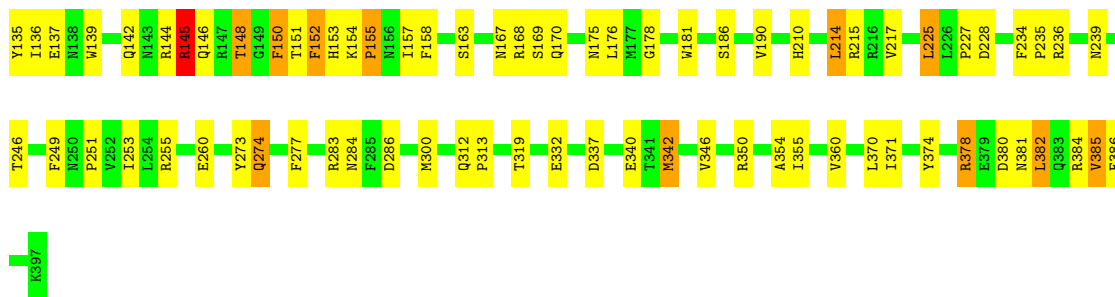


• Molecule 2: Intermediate capsid protein VP6

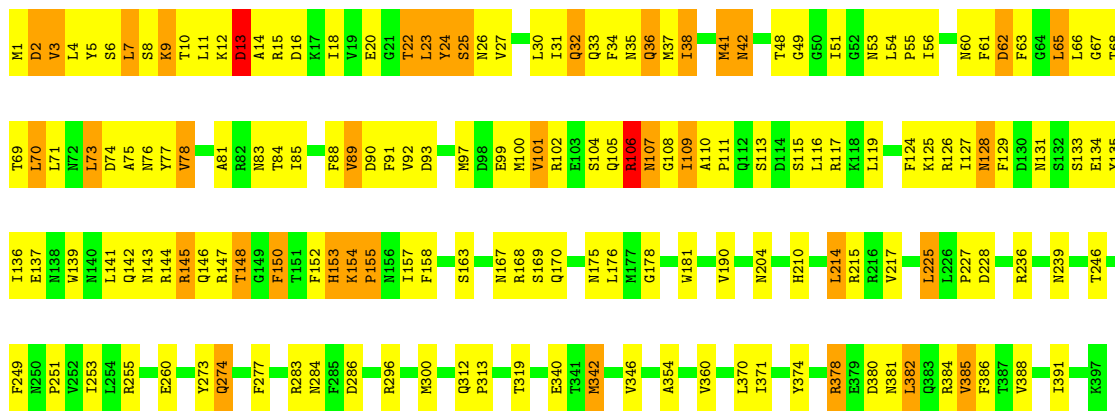


• Molecule 2: Intermediate capsid protein VP6





• Molecule 2: Intermediate capsid protein VP6



## 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, $\alpha$ , $\beta$ , $\gamma$	740.75Å 1198.07Å 1345.41Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 3.80	Depositor
% Data completeness (in resolution range)	15.9 (30.00-3.80)	Depositor
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.49 (at 3.78Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.328 , 0.339	Depositor
Wilson B-factor (Å <sup>2</sup> )	167.1	Xtrriage
Anisotropy	0.043	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	54109	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	154.0	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.*

<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, 0.333 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%)
All	All	0.50	0/55252	0.79	86/75116 (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 86 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

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
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	992	0
1	B	6624	0	6652	1122	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	C	1	0	0	0	0
3	F	1	0	0	0	0
3	I	1	0	0	0	0
3	L	1	0	0	0	0
3	O	1	0	0	0	0
All	All	54109	0	53359	4148	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 39.

The worst 5 of 4148 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:771:VAL:HB	1:B:809:PHE:HB3	1.23	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
1:A:563:MET:HE3	1:A:563:MET:HA	1.22	1.12
2:G:145:ARG:HB3	2:G:145:ARG:HH11	1.10	1.12

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	21
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	21
2	F	395/397 (100%)	324 (82%)	47 (12%)	24 (6%)	1	20
2	G	395/397 (100%)	323 (82%)	49 (12%)	23 (6%)	1	21
2	H	395/397 (100%)	323 (82%)	50 (13%)	22 (6%)	2	21
2	I	395/397 (100%)	320 (81%)	51 (13%)	24 (6%)	1	20
2	J	395/397 (100%)	320 (81%)	53 (13%)	22 (6%)	2	21
2	K	395/397 (100%)	318 (80%)	52 (13%)	25 (6%)	1	20
2	L	395/397 (100%)	322 (82%)	47 (12%)	26 (7%)	1	19

*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	23
2	N	395/397 (100%)	317 (80%)	50 (13%)	28 (7%)	1	17
2	O	395/397 (100%)	317 (80%)	47 (12%)	31 (8%)	1	15
All	All	6722/6921 (97%)	5051 (75%)	1054 (16%)	617 (9%)	1	12

5 of 617 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	17
1	B	744/809 (92%)	635 (85%)	109 (15%)	3	19
2	C	350/350 (100%)	325 (93%)	25 (7%)	14	45
2	D	350/350 (100%)	322 (92%)	28 (8%)	12	42
2	E	350/350 (100%)	325 (93%)	25 (7%)	14	45
2	F	350/350 (100%)	327 (93%)	23 (7%)	16	48
2	G	350/350 (100%)	327 (93%)	23 (7%)	16	48
2	H	350/350 (100%)	329 (94%)	21 (6%)	19	50
2	I	350/350 (100%)	327 (93%)	23 (7%)	16	48
2	J	350/350 (100%)	327 (93%)	23 (7%)	16	48
2	K	350/350 (100%)	327 (93%)	23 (7%)	16	48
2	L	350/350 (100%)	321 (92%)	29 (8%)	11	40
2	M	350/350 (100%)	325 (93%)	25 (7%)	14	45

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	N	350/350 (100%)	324 (93%)	26 (7%)	13	44
2	O	350/350 (100%)	329 (94%)	21 (6%)	19	50
All	All	6009/6168 (97%)	5473 (91%)	536 (9%)	9	38

5 of 536 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	L	385	VAL
2	M	150	PHE
2	L	382	LEU
2	O	107	ASN
1	B	744	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 220 such sidechains are listed below:

Mol	Chain	Res	Type
2	G	128	ASN
2	I	146	GLN
2	O	345	ASN
2	N	131	ASN
2	G	239	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 [i](#)

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 [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

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

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

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

### 6.3 Carbohydrates

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

### 6.4 Ligands

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

### 6.5 Other polymers

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