



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 30, 2023 – 08:39 PM JST

PDB ID : 4YVS  
Title : crystal structure of the virus-like particle of a c4 strain EV71  
Authors : Chen, R.; Lyu, K.  
Deposited on : 2015-03-20  
Resolution : 3.65 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  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

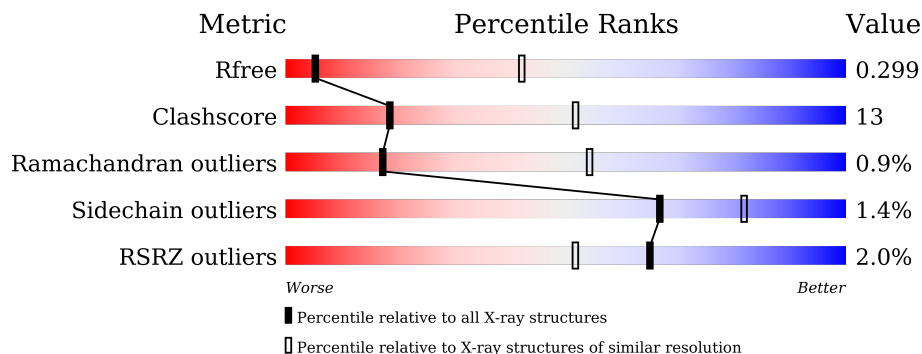
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.65 Å.

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











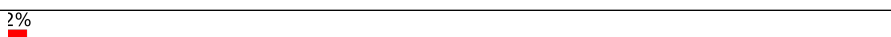
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1557 (3.82-3.50)
Clashscore	141614	1037 (3.80-3.52)
Ramachandran outliers	138981	1004 (3.80-3.52)
Sidechain outliers	138945	1002 (3.80-3.52)
RSRZ outliers	127900	1441 (3.82-3.50)

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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	297	
1	D	297	
1	G	297	
1	J	297	
1	M	297	
2	B	242	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	E	242	 64% 27% 7%
2	H	242	 65% 27% 7%
2	K	242	 63% 29% 7%
2	N	242	 66% 26% 7%
3	C	323	 2% 52% 19% 28%
3	F	323	 2% 48% 23% 28%
3	I	323	 2% 54% 17% 28%
3	L	323	 4% 46% 25% 28%
3	O	323	 2% 47% 23% 28%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 26430 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Capsid protein VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	226	1789	1145	301	332	11	0	0	0
1	D	226	1789	1145	301	332	11	0	0	0
1	G	226	1789	1145	301	332	11	0	0	0
1	J	226	1789	1145	301	332	11	0	0	0
1	M	226	1789	1145	301	332	11	0	0	0

- Molecule 2 is a protein called Capsid protein VP3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	224	1709	1099	282	317	11	0	0	0
2	E	224	1709	1099	282	317	11	0	0	0
2	H	224	1709	1099	282	317	11	0	0	0
2	K	224	1709	1099	282	317	11	0	0	0
2	N	224	1709	1099	282	317	11	0	0	0

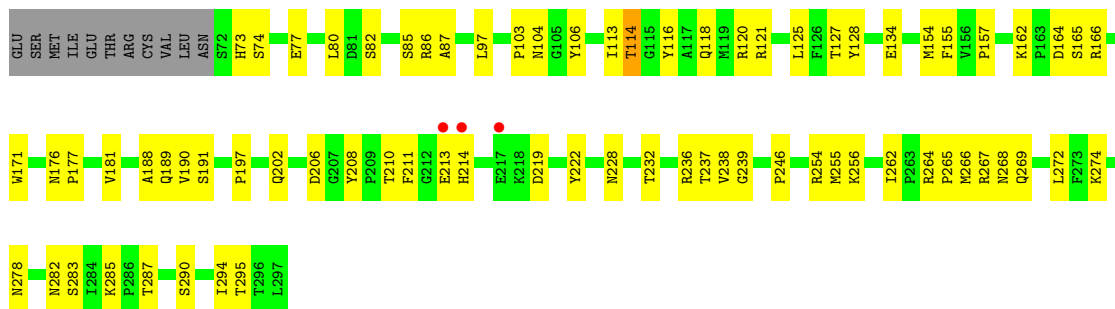
There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	227	GLN	LYS	engineered mutation	UNP F6KTB0
E	227	GLN	LYS	engineered mutation	UNP F6KTB0
H	227	GLN	LYS	engineered mutation	UNP F6KTB0
K	227	GLN	LYS	engineered mutation	UNP F6KTB0
N	227	GLN	LYS	engineered mutation	UNP F6KTB0

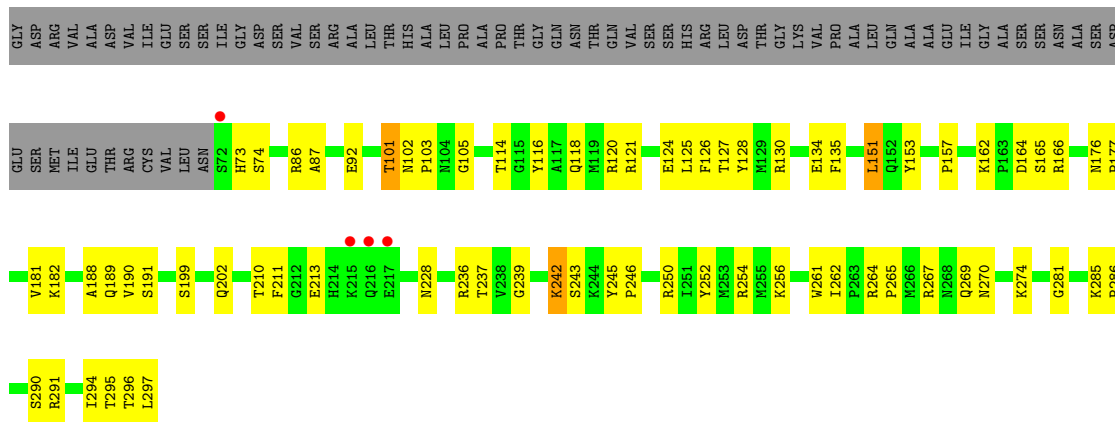
- Molecule 3 is a protein called Capsid protein VP0.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	231	Total 1788	C 1149	N 294	O 337	S 8	0	0	0
3	F	231	Total 1788	C 1149	N 294	O 337	S 8	0	0	0
3	I	231	Total 1788	C 1149	N 294	O 337	S 8	0	0	0
3	L	231	Total 1788	C 1149	N 294	O 337	S 8	0	0	0
3	O	231	Total 1788	C 1149	N 294	O 337	S 8	0	0	0

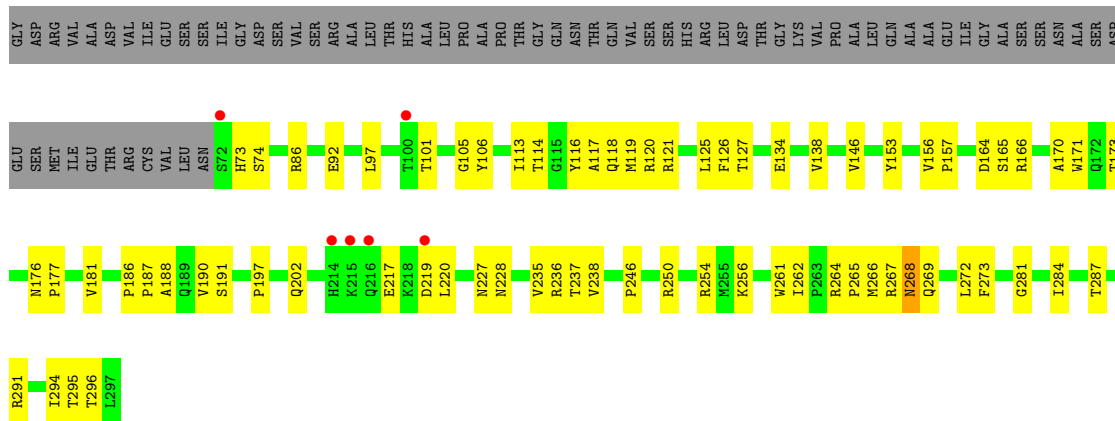




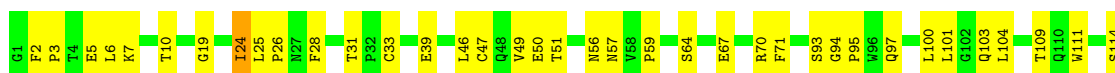
• Molecule 1: Capsid protein VP1

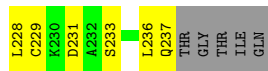
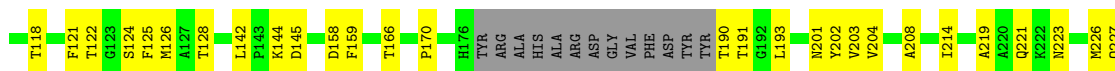


• Molecule 1: Capsid protein VP1

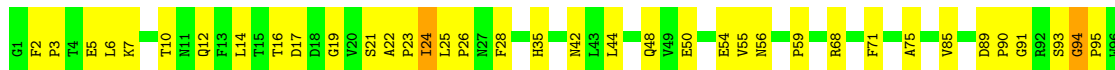


• Molecule 2: Capsid protein VP3

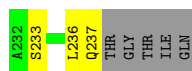
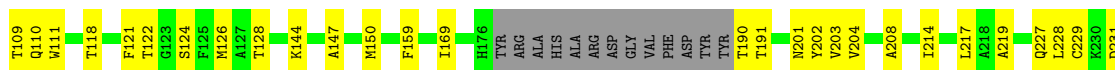
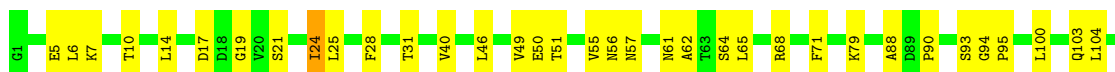




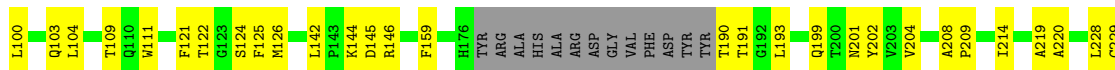
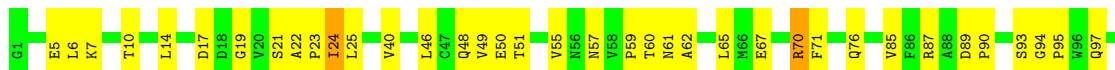
• Molecule 2: Capsid protein VP3



• Molecule 2: Capsid protein VP3



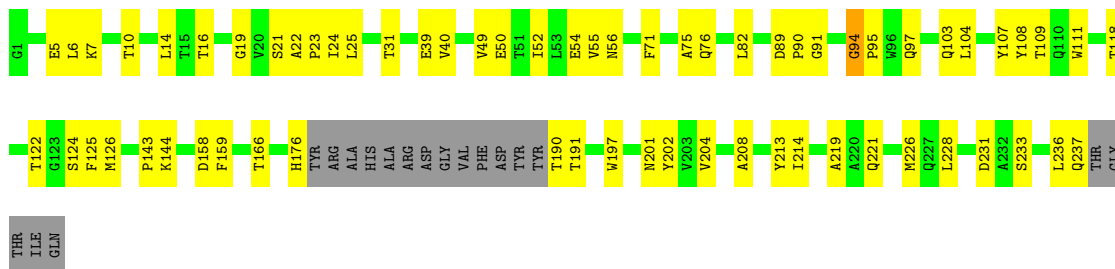
• Molecule 2: Capsid protein VP3



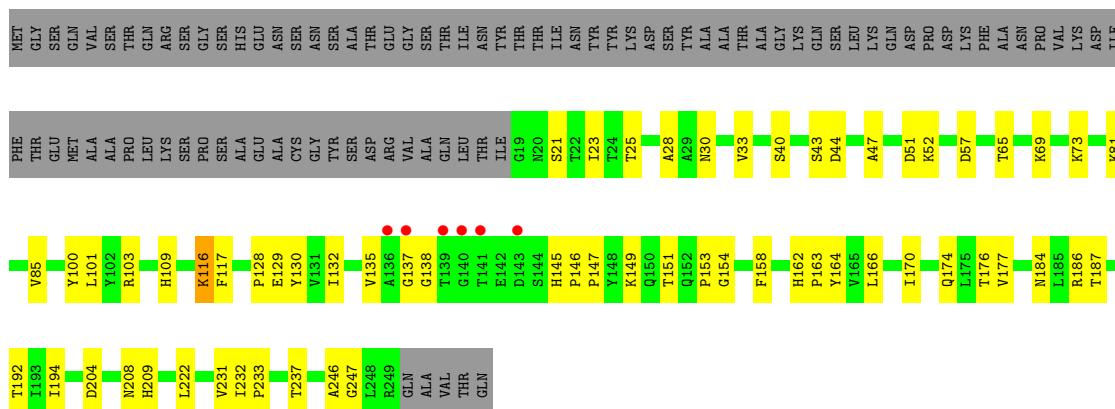
• Molecule 2: Capsid protein VP3



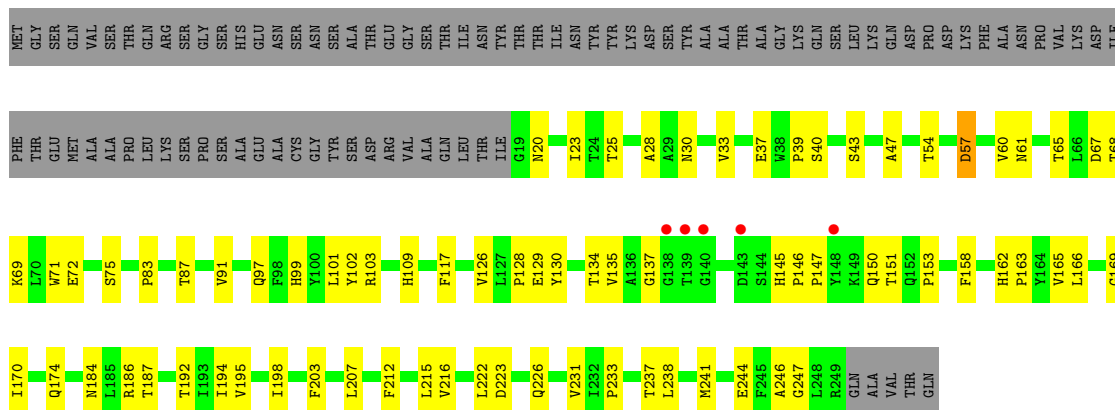




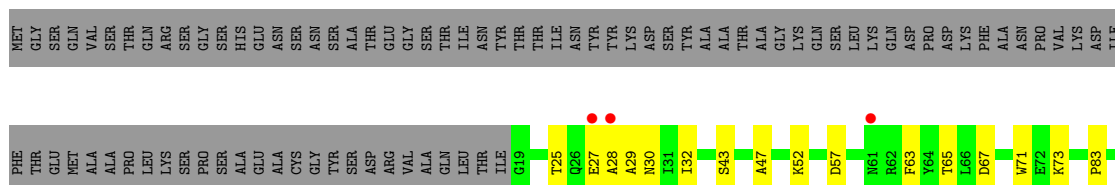
- Molecule 3: Capsid protein VP0

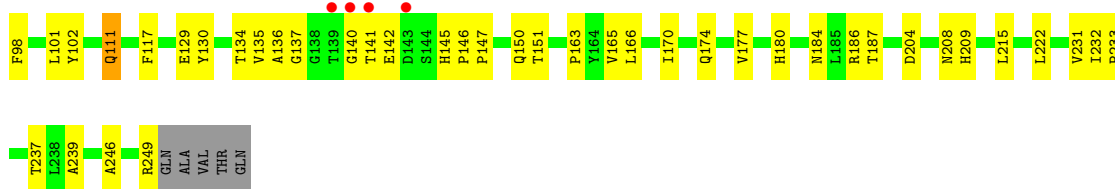


- Molecule 3: Capsid protein VP0

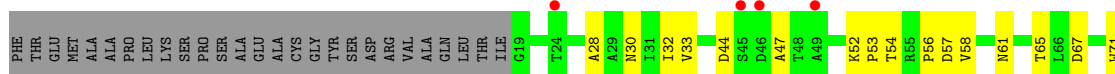
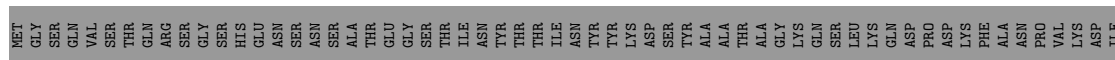


- Molecule 3: Capsid protein VP0





• Molecule 3: Capsid protein VP0



• Molecule 3: Capsid protein VP0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 42 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	349.75Å 349.75Å 349.75Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.96 – 3.65 49.96 – 3.65	Depositor EDS
% Data completeness (in resolution range)	88.5 (49.96-3.65) 70.2 (49.96-3.65)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.04 (at 3.67Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: dev_1839)	Depositor
R, $R_{free}$	0.271 , 0.300 0.271 , 0.299	Depositor DCC
$R_{free}$ test set	2000 reflections (2.76%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	86.9	Xtrriage
Anisotropy	0.000	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 18.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.85	EDS
Total number of atoms	26430	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	88.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.95% 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](#)

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	0.27	0/1844	0.49	0/2512
1	D	0.27	0/1844	0.49	0/2512
1	G	0.27	0/1844	0.49	0/2512
1	J	0.27	0/1844	0.48	0/2512
1	M	0.29	0/1844	0.51	0/2512
2	B	0.27	0/1755	0.54	0/2402
2	E	0.28	0/1755	0.52	0/2402
2	H	0.28	0/1755	0.53	0/2402
2	K	0.29	0/1755	0.55	0/2402
2	N	0.28	0/1755	0.54	0/2402
3	C	0.28	0/1843	0.50	0/2529
3	F	0.29	0/1843	0.51	0/2529
3	I	0.28	0/1843	0.52	0/2529
3	L	0.29	0/1843	0.51	0/2529
3	O	0.27	0/1843	0.51	0/2529
All	All	0.28	0/27210	0.51	0/37215

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	1789	0	1735	69	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1789	0	1735	63	0
1	G	1789	0	1735	65	0
1	J	1789	0	1735	58	0
1	M	1789	0	1735	62	0
2	B	1709	0	1698	61	0
2	E	1709	0	1698	60	1
2	H	1709	0	1698	62	0
2	K	1709	0	1698	63	0
2	N	1709	0	1698	58	0
3	C	1788	0	1721	46	0
3	F	1788	0	1721	56	0
3	I	1788	0	1721	43	1
3	L	1788	0	1721	59	0
3	O	1788	0	1721	52	0
All	All	26430	0	25770	689	1

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 13.

The worst 5 of 689 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:M:181:VAL:HG21	1:M:188:ALA:HB2	1.58	0.86
1:J:210:THR:O	3:L:208:ASN:ND2	2.10	0.85
1:M:268:ASN:OD1	1:M:269:GLN:HG2	1.77	0.84
1:G:134:GLU:OE2	1:G:189:GLN:NE2	2.10	0.84
3:F:40:SER:O	3:F:103:ARG:NH2	2.12	0.81

All (1) 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)
2:E:130:LYS:NZ	3:I:111:GLN:OE1[4_555]	2.10	0.10

## 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	224/297 (75%)	200 (89%)	21 (9%)	3 (1%)	12	47
1	D	224/297 (75%)	198 (88%)	23 (10%)	3 (1%)	12	47
1	G	224/297 (75%)	198 (88%)	24 (11%)	2 (1%)	17	54
1	J	224/297 (75%)	199 (89%)	23 (10%)	2 (1%)	17	54
1	M	224/297 (75%)	199 (89%)	25 (11%)	0	100	100
2	B	220/242 (91%)	203 (92%)	17 (8%)	0	100	100
2	E	220/242 (91%)	201 (91%)	18 (8%)	1 (0%)	29	66
2	H	220/242 (91%)	201 (91%)	19 (9%)	0	100	100
2	K	220/242 (91%)	202 (92%)	18 (8%)	0	100	100
2	N	220/242 (91%)	203 (92%)	16 (7%)	1 (0%)	29	66
3	C	229/323 (71%)	204 (89%)	23 (10%)	2 (1%)	17	54
3	F	229/323 (71%)	204 (89%)	21 (9%)	4 (2%)	9	42
3	I	229/323 (71%)	206 (90%)	19 (8%)	4 (2%)	9	42
3	L	229/323 (71%)	206 (90%)	19 (8%)	4 (2%)	9	42
3	O	229/323 (71%)	206 (90%)	18 (8%)	5 (2%)	6	37
All	All	3365/4310 (78%)	3030 (90%)	304 (9%)	31 (1%)	17	54

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	216	GLN
3	C	57	ASP
3	F	57	ASP
3	I	57	ASP
3	L	57	ASP

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	194/251 (77%)	191 (98%)	3 (2%)	65	81
1	D	194/251 (77%)	192 (99%)	2 (1%)	76	86
1	G	194/251 (77%)	193 (100%)	1 (0%)	88	94
1	J	194/251 (77%)	190 (98%)	4 (2%)	53	73
1	M	194/251 (77%)	191 (98%)	3 (2%)	65	81
2	B	188/202 (93%)	187 (100%)	1 (0%)	88	94
2	E	188/202 (93%)	186 (99%)	2 (1%)	73	85
2	H	188/202 (93%)	186 (99%)	2 (1%)	73	85
2	K	188/202 (93%)	185 (98%)	3 (2%)	62	79
2	N	188/202 (93%)	184 (98%)	4 (2%)	53	73
3	C	196/272 (72%)	194 (99%)	2 (1%)	76	86
3	F	196/272 (72%)	191 (97%)	5 (3%)	46	68
3	I	196/272 (72%)	194 (99%)	2 (1%)	76	86
3	L	196/272 (72%)	193 (98%)	3 (2%)	65	81
3	O	196/272 (72%)	192 (98%)	4 (2%)	55	74
All	All	2890/3625 (80%)	2849 (99%)	41 (1%)	67	82

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	54	THR
3	L	116	LYS
3	F	216	VAL
3	I	111	GLN
3	O	22	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	61	ASN
3	F	162	HIS
3	O	61	ASN
3	F	208	ASN
1	M	176	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](#)

There are no ligands in this entry.

### 5.7 Other polymers [i](#)

There are no such residues in this entry.

### 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.



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

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

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	226/297 (76%)	-0.24	4 (1%) 68 55	60, 81, 134, 274	0
1	D	226/297 (76%)	-0.11	12 (5%) 26 18	61, 82, 146, 274	0
1	G	226/297 (76%)	-0.17	3 (1%) 77 65	63, 85, 146, 274	0
1	J	226/297 (76%)	-0.31	4 (1%) 68 55	55, 78, 147, 274	0
1	M	226/297 (76%)	-0.23	6 (2%) 54 40	64, 83, 144, 274	0
2	B	224/242 (92%)	-0.33	0 100 100	60, 81, 113, 216	0
2	E	224/242 (92%)	-0.24	0 100 100	67, 84, 114, 217	0
2	H	224/242 (92%)	-0.36	0 100 100	65, 83, 112, 219	0
2	K	224/242 (92%)	-0.34	0 100 100	56, 76, 113, 216	0
2	N	224/242 (92%)	-0.40	0 100 100	59, 82, 112, 217	0
3	C	231/323 (71%)	-0.05	6 (2%) 56 42	63, 83, 151, 185	0
3	F	231/323 (71%)	-0.11	5 (2%) 62 48	60, 86, 152, 183	0
3	I	231/323 (71%)	-0.02	7 (3%) 50 36	66, 88, 154, 180	0
3	L	231/323 (71%)	-0.07	12 (5%) 27 19	58, 78, 154, 185	0
3	O	231/323 (71%)	-0.17	8 (3%) 44 31	63, 88, 154, 179	0
All	All	3405/4310 (79%)	-0.21	67 (1%) 65 52	55, 83, 146, 274	0

The worst 5 of 67 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	139	THR	6.5
1	D	216	GLN	5.1
3	O	139	THR	5.0
3	C	139	THR	4.9
3	L	136	ALA	4.4

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

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

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.