



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:25 PM GMT

PDB ID : 4RR3
Title : Crystal structure of a recombinant EV71 virus particle
Authors : Chen, R.; Lyu, K.
Deposited on : 2014-11-06
Resolution : 3.10 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
<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

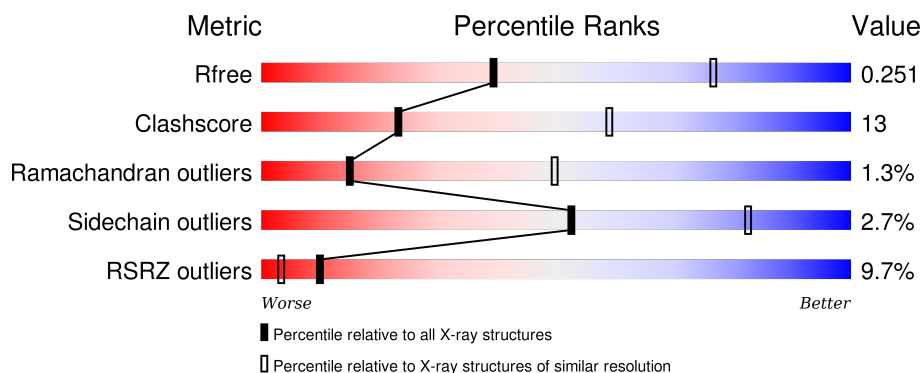
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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 ≥ 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	303	<div> <div>9%</div> <div>59%</div> <div>16%</div> <div>•</div> <div>23%</div> </div>
1	E	303	<div> <div>9%</div> <div>63%</div> <div>12%</div> <div>••</div> <div>23%</div> </div>
1	I	303	<div> <div>8%</div> <div>60%</div> <div>14%</div> <div>•</div> <div>23%</div> </div>
1	M	303	<div> <div>8%</div> <div>60%</div> <div>14%</div> <div>•</div> <div>23%</div> </div>
1	Q	303	<div> <div>9%</div> <div>59%</div> <div>15%</div> <div>•</div> <div>23%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	242	
2	F	242	
2	J	242	
2	N	242	
2	R	242	
3	C	323	
3	G	323	
3	K	323	
3	O	323	
3	S	323	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 27390 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	E	232	Total	C	N	O	S	0	0	0
			1844	1179	316	338	11			
1	Q	232	Total	C	N	O	S	0	0	0
			1844	1179	316	338	11			
1	I	232	Total	C	N	O	S	0	0	0
			1844	1179	316	338	11			
1	M	232	Total	C	N	O	S	0	0	0
			1844	1179	316	338	11			
1	A	232	Total	C	N	O	S	0	0	0
			1844	1179	316	338	11			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	101	GLU	-	EXPRESSION TAG	UNP F6KTB0
E	102	ARG	-	EXPRESSION TAG	UNP F6KTB0
E	103	LYS	-	EXPRESSION TAG	UNP F6KTB0
E	104	ARG	-	EXPRESSION TAG	UNP F6KTB0
E	105	ALA	-	EXPRESSION TAG	UNP F6KTB0
E	106	ARG	-	EXPRESSION TAG	UNP F6KTB0
E	107	LEU	-	EXPRESSION TAG	UNP F6KTB0
E	?	-	ASN	DELETION	UNP F6KTB0
Q	101	GLU	-	EXPRESSION TAG	UNP F6KTB0
Q	102	ARG	-	EXPRESSION TAG	UNP F6KTB0
Q	103	LYS	-	EXPRESSION TAG	UNP F6KTB0
Q	104	ARG	-	EXPRESSION TAG	UNP F6KTB0
Q	105	ALA	-	EXPRESSION TAG	UNP F6KTB0
Q	106	ARG	-	EXPRESSION TAG	UNP F6KTB0
Q	107	LEU	-	EXPRESSION TAG	UNP F6KTB0
Q	?	-	ASN	DELETION	UNP F6KTB0
I	101	GLU	-	EXPRESSION TAG	UNP F6KTB0
I	102	ARG	-	EXPRESSION TAG	UNP F6KTB0
I	103	LYS	-	EXPRESSION TAG	UNP F6KTB0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
I	104	ARG	-	EXPRESSION TAG	UNP F6KTB0
I	105	ALA	-	EXPRESSION TAG	UNP F6KTB0
I	106	ARG	-	EXPRESSION TAG	UNP F6KTB0
I	107	LEU	-	EXPRESSION TAG	UNP F6KTB0
I	?	-	ASN	DELETION	UNP F6KTB0
M	101	GLU	-	EXPRESSION TAG	UNP F6KTB0
M	102	ARG	-	EXPRESSION TAG	UNP F6KTB0
M	103	LYS	-	EXPRESSION TAG	UNP F6KTB0
M	104	ARG	-	EXPRESSION TAG	UNP F6KTB0
M	105	ALA	-	EXPRESSION TAG	UNP F6KTB0
M	106	ARG	-	EXPRESSION TAG	UNP F6KTB0
M	107	LEU	-	EXPRESSION TAG	UNP F6KTB0
M	?	-	ASN	DELETION	UNP F6KTB0
A	101	GLU	-	EXPRESSION TAG	UNP F6KTB0
A	102	ARG	-	EXPRESSION TAG	UNP F6KTB0
A	103	LYS	-	EXPRESSION TAG	UNP F6KTB0
A	104	ARG	-	EXPRESSION TAG	UNP F6KTB0
A	105	ALA	-	EXPRESSION TAG	UNP F6KTB0
A	106	ARG	-	EXPRESSION TAG	UNP F6KTB0
A	107	LEU	-	EXPRESSION TAG	UNP F6KTB0
A	?	-	ASN	DELETION	UNP F6KTB0

- Molecule 2 is a protein called Capsid protein VP3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	230	Total	C	N	O	S	0	0	0
			1762	1133	291	327	11			
2	R	230	Total	C	N	O	S	0	0	0
			1762	1133	291	327	11			
2	J	230	Total	C	N	O	S	0	0	0
			1762	1133	291	327	11			
2	N	230	Total	C	N	O	S	0	0	0
			1762	1133	291	327	11			
2	B	230	Total	C	N	O	S	0	0	0
			1762	1133	291	327	11			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	227	GLN	LYS	engineered mutation	UNP F6KTB0
R	227	GLN	LYS	engineered mutation	UNP F6KTB0
J	227	GLN	LYS	engineered mutation	UNP F6KTB0

Continued on next page...

Continued from previous page...

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

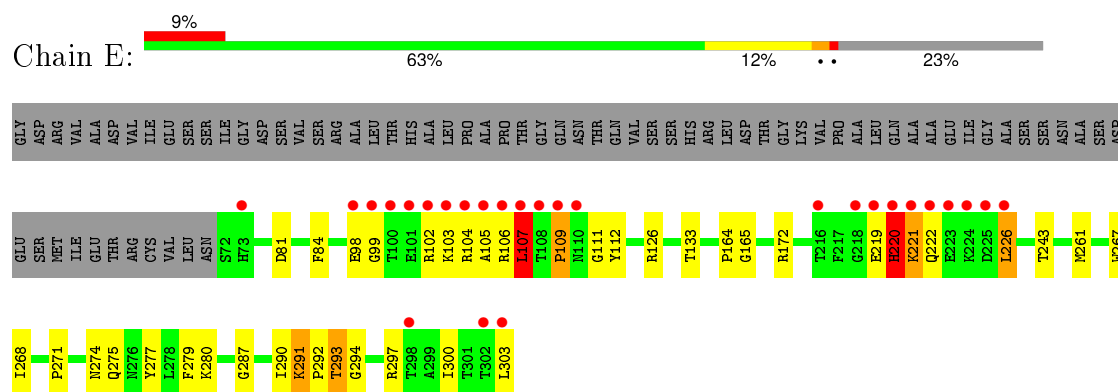
- Molecule 3 is a protein called Capsid protein VP0.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	242	Total 1872	C 1201	N 310	O 353	S 8	0	0	0
3	S	242	Total 1872	C 1201	N 310	O 353	S 8	0	0	0
3	K	242	Total 1872	C 1201	N 310	O 353	S 8	0	0	0
3	O	242	Total 1872	C 1201	N 310	O 353	S 8	0	0	0
3	C	242	Total 1872	C 1201	N 310	O 353	S 8	0	0	0

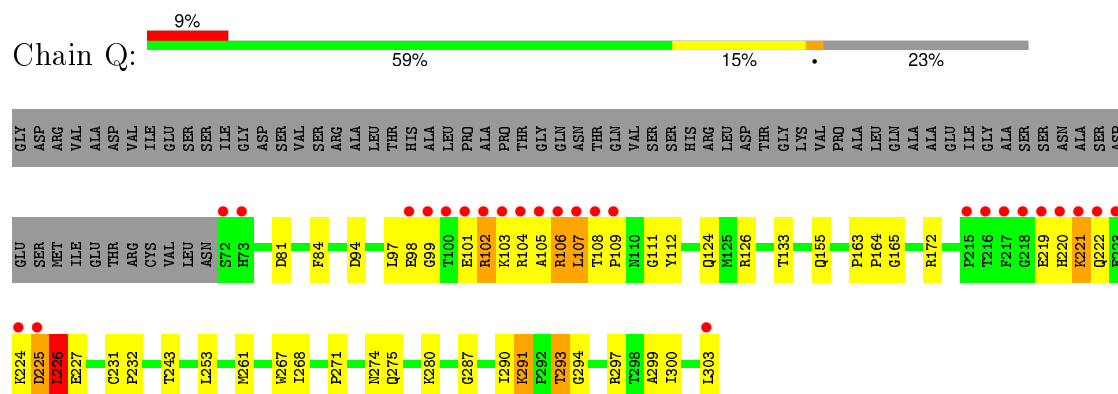
3 Residue-property plots [i](#)

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 ($RSRZ > 2$). 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.

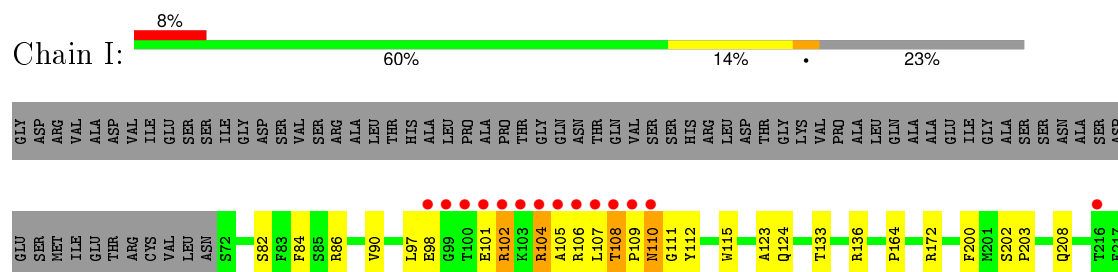
- Molecule 1: Capsid protein VP1



- Molecule 1: Capsid protein VP1

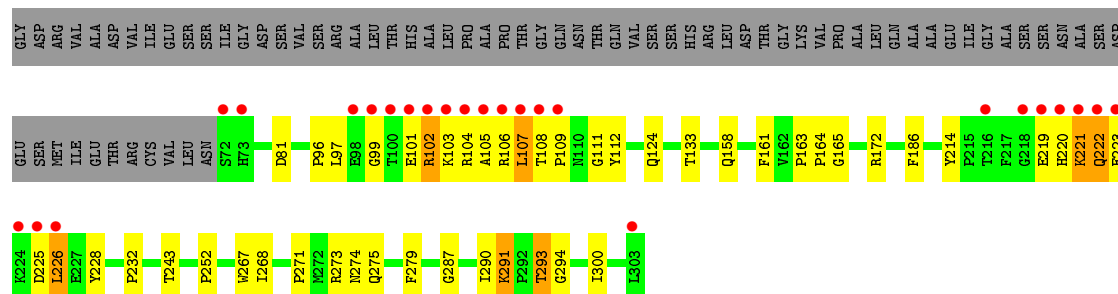


- Molecule 1: Capsid protein VP1

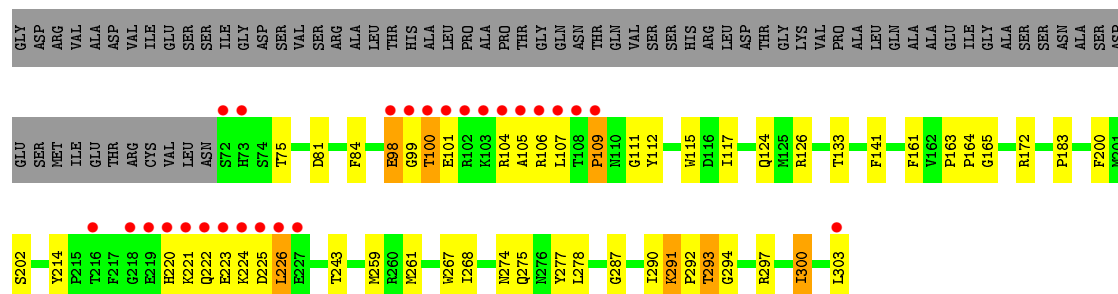




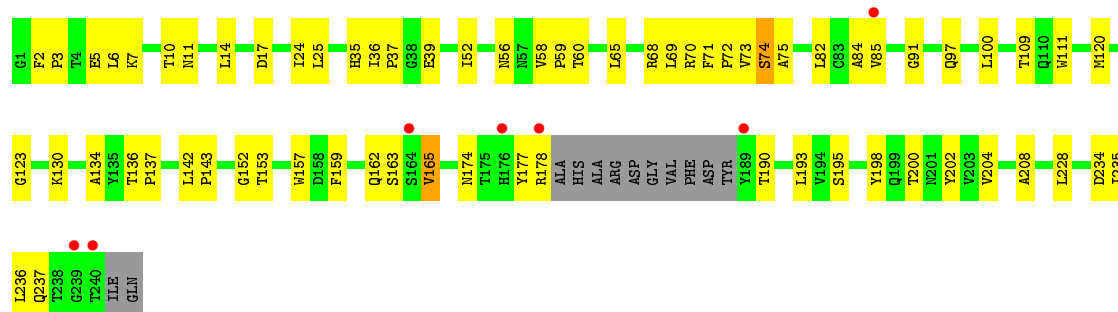
• Molecule 1: Capsid protein VP1



• Molecule 1: Capsid protein VP1



• Molecule 2: Capsid protein VP3

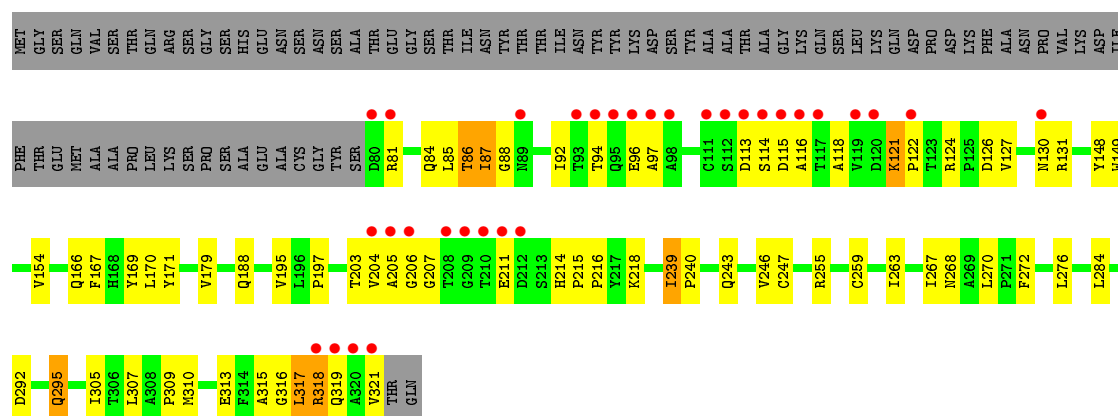


• Molecule 2: Capsid protein VP3

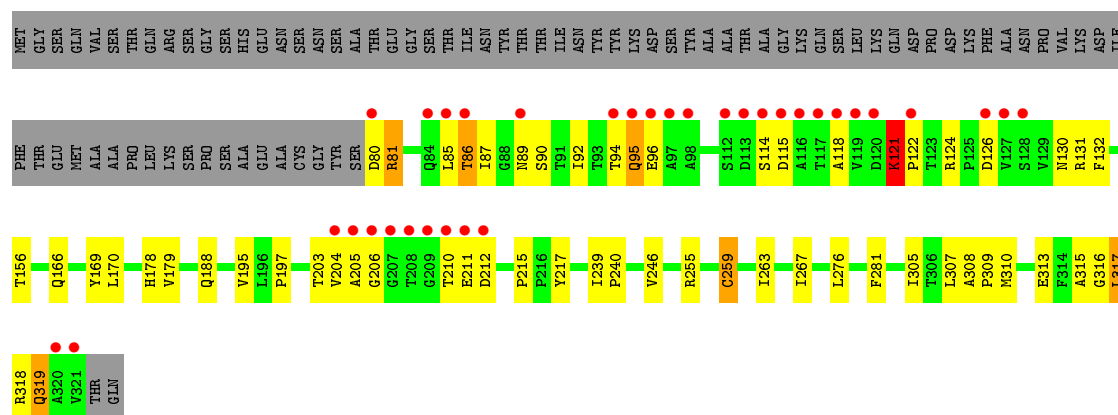




- Molecule 3: Capsid protein VP0

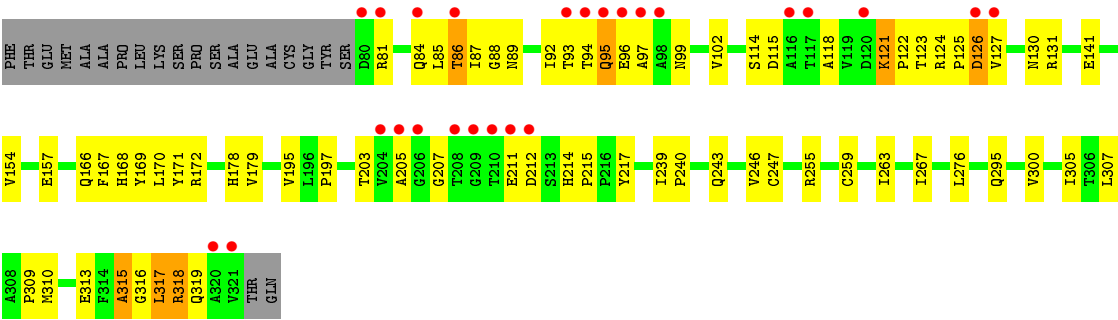


- Molecule 3: Capsid protein VP0

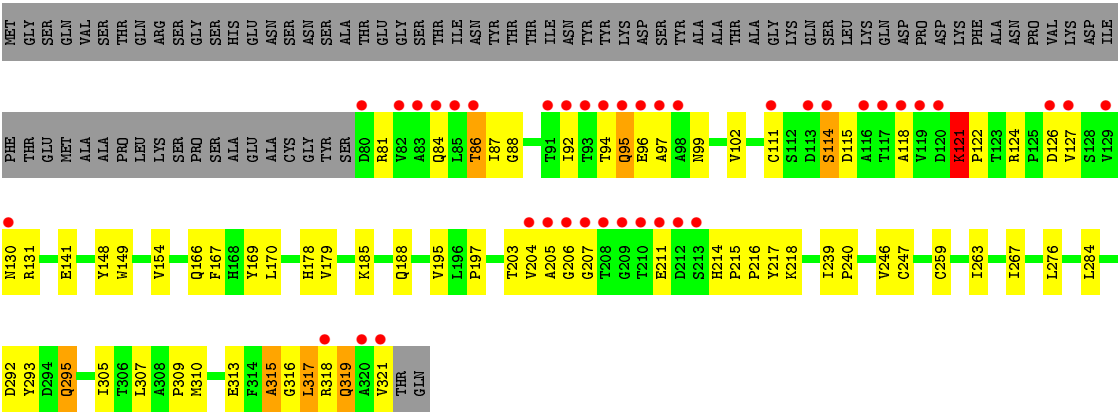


- Molecule 3: Capsid protein VP0





● Molecule 3: Capsid protein VP0



4 Data and refinement statistics

Property	Value	Source
Space group	P 42 3 2	Depositor
Cell constants a, b, c, α , β , γ	350.60Å 350.60Å 350.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.44 – 3.10 47.71 – 3.10	Depositor EDS
% Data completeness (in resolution range)	(Not available) (46.44-3.10) 90.6 (47.71-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.28 (at 3.12Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.235 , 0.265 0.219 , 0.251	Depositor DCC
R_{free} test set	1802 reflections (1.51%)	DCC
Wilson B-factor (Å ²)	61.9	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 30.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.36$	Xtriage
Outliers	1 of 131764 reflections (0.001%)	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	27390	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.19% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.36	0/1899	0.63	0/2584
1	E	0.35	0/1899	0.57	1/2584 (0.0%)
1	I	0.36	0/1899	0.64	2/2584 (0.1%)
1	M	0.33	0/1899	0.60	2/2584 (0.1%)
1	Q	0.37	0/1899	0.66	2/2584 (0.1%)
2	B	0.32	0/1810	0.57	0/2477
2	F	0.29	0/1810	0.53	0/2477
2	J	0.32	0/1810	0.57	0/2477
2	N	0.29	0/1810	0.53	0/2477
2	R	0.31	0/1810	0.55	0/2477
3	C	0.39	0/1927	0.60	0/2644
3	G	0.29	0/1927	0.54	1/2644 (0.0%)
3	K	0.30	0/1927	0.55	0/2644
3	O	0.29	0/1927	0.56	0/2644
3	S	0.30	0/1927	0.54	0/2644
All	All	0.33	0/28180	0.58	8/38525 (0.0%)

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	I	0	1
1	Q	0	2
3	C	0	2
3	G	0	2
3	K	0	2
3	O	0	1
All	All	0	10

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	104	ARG	NE-CZ-NH1	-9.68	115.46	120.30
1	Q	106	ARG	NE-CZ-NH1	-9.04	115.78	120.30
1	Q	106	ARG	NE-CZ-NH2	8.34	124.47	120.30
1	M	226	LEU	CA-CB-CG	-6.01	101.48	115.30
1	E	107	LEU	CA-CB-CG	5.83	128.72	115.30

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	G	121	LYS	Peptide
3	G	95	GLN	Peptide
1	I	219	GLU	Peptide
1	Q	225	ASP	Peptide
1	Q	226	LEU	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	1844	0	1803	60	0
1	E	1844	0	1803	46	0
1	I	1844	0	1803	53	0
1	M	1844	0	1803	67	0
1	Q	1844	0	1803	65	0
2	B	1762	0	1746	67	0
2	F	1762	0	1746	68	0
2	J	1762	0	1746	62	0
2	N	1762	0	1746	57	0
2	R	1762	0	1746	65	0
3	C	1872	0	1811	58	0
3	G	1872	0	1811	54	0
3	K	1872	0	1811	45	0
3	O	1872	0	1811	64	0
3	S	1872	0	1811	64	0
All	All	27390	0	26800	724	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 13.

The worst 5 of 724 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:I:98:GLU:HA	1:I:104:ARG:HH12	1.19	1.04
2:R:85:VAL:HG21	2:R:195:SER:HA	1.40	1.02
3:C:295:GLN:HA	3:C:295:GLN:HE21	1.19	1.00
1:M:105:ALA:HA	1:M:108:THR:HA	1.44	0.99
1:I:105:ALA:HA	1:I:108:THR:HA	1.43	0.98

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	230/303 (76%)	207 (90%)	19 (8%)	4 (2%)	11	43
1	E	230/303 (76%)	207 (90%)	18 (8%)	5 (2%)	8	36
1	I	230/303 (76%)	208 (90%)	17 (7%)	5 (2%)	8	36
1	M	230/303 (76%)	208 (90%)	19 (8%)	3 (1%)	15	50
1	Q	230/303 (76%)	208 (90%)	18 (8%)	4 (2%)	11	43
2	B	226/242 (93%)	206 (91%)	18 (8%)	2 (1%)	21	61
2	F	226/242 (93%)	204 (90%)	20 (9%)	2 (1%)	21	61
2	J	226/242 (93%)	204 (90%)	20 (9%)	2 (1%)	21	61
2	N	226/242 (93%)	204 (90%)	20 (9%)	2 (1%)	21	61
2	R	226/242 (93%)	204 (90%)	20 (9%)	2 (1%)	21	61
3	C	240/323 (74%)	218 (91%)	19 (8%)	3 (1%)	15	50
3	G	240/323 (74%)	221 (92%)	16 (7%)	3 (1%)	15	50
3	K	240/323 (74%)	221 (92%)	16 (7%)	3 (1%)	15	50

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	O	240/323 (74%)	225 (94%)	11 (5%)	4 (2%)	11	43
3	S	240/323 (74%)	223 (93%)	15 (6%)	2 (1%)	24	63
All	All	3480/4340 (80%)	3168 (91%)	266 (8%)	46 (1%)	15	50

5 of 46 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	G	121	LYS
3	G	317	LEU
1	Q	226	LEU
1	I	226	LEU
1	M	221	LYS

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	199/256 (78%)	193 (97%)	6 (3%)	48	81
1	E	199/256 (78%)	194 (98%)	5 (2%)	55	84
1	I	199/256 (78%)	194 (98%)	5 (2%)	55	84
1	M	199/256 (78%)	195 (98%)	4 (2%)	63	86
1	Q	199/256 (78%)	195 (98%)	4 (2%)	63	86
2	B	193/202 (96%)	189 (98%)	4 (2%)	61	86
2	F	193/202 (96%)	191 (99%)	2 (1%)	82	93
2	J	193/202 (96%)	189 (98%)	4 (2%)	61	86
2	N	193/202 (96%)	190 (98%)	3 (2%)	70	89
2	R	193/202 (96%)	190 (98%)	3 (2%)	70	89
3	C	205/272 (75%)	196 (96%)	9 (4%)	35	71
3	G	205/272 (75%)	198 (97%)	7 (3%)	44	79
3	K	205/272 (75%)	198 (97%)	7 (3%)	44	79
3	O	205/272 (75%)	197 (96%)	8 (4%)	39	75

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	S	205/272 (75%)	196 (96%)	9 (4%)	35 71
All	All	2985/3650 (82%)	2905 (97%)	80 (3%)	52 82

5 of 80 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	J	70	ARG
2	B	76	GLN
3	C	141	GLU
2	J	85	VAL
2	N	70	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
3	S	188	GLN
3	C	319	GLN
3	S	295	GLN
1	M	222	GLN
3	C	295	GLN

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](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	232/303 (76%)	0.05	26 (11%) 7 2	25, 40, 140, 156	0
1	E	232/303 (76%)	0.04	27 (11%) 6 2	27, 41, 140, 154	0
1	I	232/303 (76%)	0.08	25 (10%) 8 2	30, 41, 142, 157	0
1	M	232/303 (76%)	0.10	25 (10%) 8 2	28, 41, 142, 156	0
1	Q	232/303 (76%)	0.06	26 (11%) 7 2	29, 42, 140, 156	0
2	B	230/242 (95%)	-0.16	9 (3%) 43 21	27, 41, 73, 147	0
2	F	230/242 (95%)	-0.26	7 (3%) 54 29	32, 44, 80, 142	0
2	J	230/242 (95%)	-0.16	10 (4%) 39 18	31, 44, 79, 145	0
2	N	230/242 (95%)	-0.22	12 (5%) 31 13	28, 44, 76, 149	0
2	R	230/242 (95%)	-0.24	10 (4%) 39 18	30, 45, 81, 145	0
3	C	242/323 (74%)	0.33	39 (16%) 3 1	25, 47, 130, 150	0
3	G	242/323 (74%)	0.23	33 (13%) 4 2	31, 50, 131, 149	0
3	K	242/323 (74%)	0.30	34 (14%) 4 2	29, 50, 130, 148	0
3	O	242/323 (74%)	0.16	25 (10%) 9 3	31, 52, 131, 148	0
3	S	242/323 (74%)	0.30	32 (13%) 4 2	31, 52, 133, 148	0
All	All	3520/4340 (81%)	0.04	340 (9%) 10 3	25, 44, 129, 157	0

The worst 5 of 340 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	S	80	ASP	10.5
2	J	240	THR	9.5
3	S	97	ALA	9.2
2	B	240	THR	8.6
1	E	108	THR	8.3

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 carbohydrates 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.