



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 08:47 am BST

PDB ID : 2WZP  
Title : Structures of Lactococcal Phage p2 Baseplate Shed Light on a Novel Mechanism of Host Attachment and Activation in Siphoviridae  
Authors : Sciara, G.; Bebeacua, C.; Bron, P.; Tremblay, D.; Ortiz-Lombardia, M.; Lichiere, J.; van Heel, M.; Campanacci, V.; Moineau, S.; Cambillau, C.  
Deposited on : 2009-12-01  
Resolution : 2.60 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 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	:	2.11
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)	:	2.11

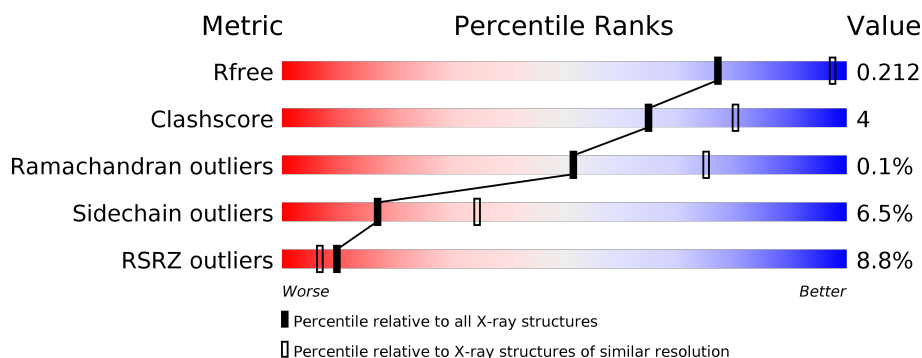
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 on 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	266	<div> <div>6%</div> <div> <div></div> <div>91%</div> <div>9%</div> </div> </div>
1	B	266	<div> <div>5%</div> <div> <div></div> <div>91%</div> <div>8%</div> </div> </div>
1	C	266	<div> <div>3%</div> <div> <div></div> <div>91%</div> <div>9%</div> </div> </div>
1	G	266	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>13%</div> </div> </div>
1	H	266	<div> <div>2%</div> <div> <div></div> <div>89%</div> <div>11%</div> </div> </div>
1	I	266	<div> <div>0%</div> <div> <div></div> <div>91%</div> <div>9%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	123	<div><div>5%</div><div><div></div><div>76%</div><div>23%</div><div></div></div><div>..</div></div>
2	E	123	<div><div>50%</div><div><div></div><div>71%</div><div>25%</div><div></div></div><div>..</div></div>
2	F	123	<div><div>55%</div><div><div></div><div>81%</div><div>16%</div><div></div></div><div>..</div></div>
2	J	123	<div><div>37%</div><div><div></div><div>77%</div><div>20%</div><div></div></div><div>..</div></div>
2	K	123	<div><div>35%</div><div><div></div><div>74%</div><div>23%</div><div></div></div><div>..</div></div>
2	L	123	<div><div>14%</div><div><div></div><div>78%</div><div>18%</div><div></div></div><div>..</div></div>
3	P	326	<div><div></div><div><div></div><div>81%</div><div>9%</div><div></div></div><div>. 9%</div></div>
3	Q	326	<div><div></div><div><div></div><div>78%</div><div>11%</div><div></div></div><div>. 9%</div></div>
4	R	375	<div><div></div><div><div></div><div>87%</div><div>11%</div><div></div></div><div>.</div></div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 27647 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 PUTATIVE RECEPTOR BINDING PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	266	Total	C	N	O	S	0	0	0
			2024	1269	349	400	6			
1	B	266	Total	C	N	O	S	0	0	0
			2024	1269	349	400	6			
1	C	266	Total	C	N	O	S	0	0	0
			2024	1269	349	400	6			
1	G	266	Total	C	N	O	S	0	0	0
			2024	1269	349	400	6			
1	H	266	Total	C	N	O	S	0	0	0
			2024	1269	349	400	6			
1	I	266	Total	C	N	O	S	0	0	0
			2024	1269	349	400	6			

- Molecule 2 is a protein called CAMELID VHH5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	122	Total	C	N	O	S	0	0	0
			938	581	165	186	6			
2	E	122	Total	C	N	O	S	0	0	0
			922	572	158	186	6			
2	F	122	Total	C	N	O	S	0	0	0
			931	575	165	185	6			
2	J	122	Total	C	N	O	S	0	0	0
			938	581	165	186	6			
2	K	122	Total	C	N	O	S	0	0	0
			938	581	165	186	6			
2	L	122	Total	C	N	O	S	0	0	0
			938	581	165	186	6			

- Molecule 3 is a protein called LACTOCOCCAL PHAGE P2 ORF15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	298	Total	C	N	O	S	0	0	0
			2432	1565	392	469	6			
3	Q	298	Total	C	N	O	S	0	0	0
			2432	1565	392	469	6			

- Molecule 4 is a protein called LACTOCOCCAL PHAGE P2 ORF16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	R	375	Total	C	N	O	S	0	0	0
			3025	1934	497	586	8			

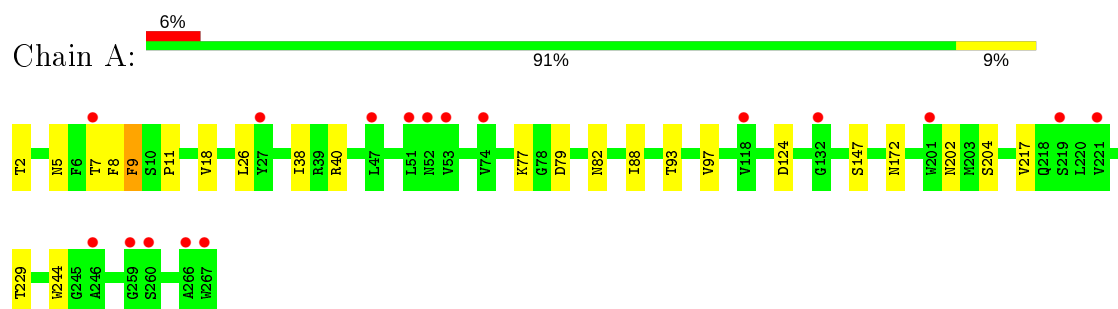
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	47	Total	O	0	0
			47	47		
5	B	85	Total	O	0	0
			85	85		
5	C	112	Total	O	0	0
			112	112		
5	D	11	Total	O	0	0
			11	11		
5	E	1	Total	O	0	0
			1	1		
5	F	1	Total	O	0	0
			1	1		
5	G	199	Total	O	0	0
			199	199		
5	H	156	Total	O	0	0
			156	156		
5	I	192	Total	O	0	0
			192	192		
5	K	3	Total	O	0	0
			3	3		
5	L	7	Total	O	0	0
			7	7		
5	P	363	Total	O	0	0
			363	363		
5	Q	265	Total	O	0	0
			265	265		
5	R	567	Total	O	0	0
			567	567		

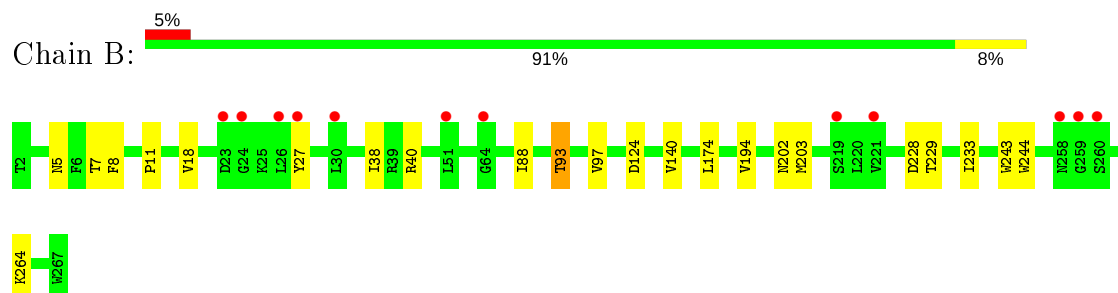
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

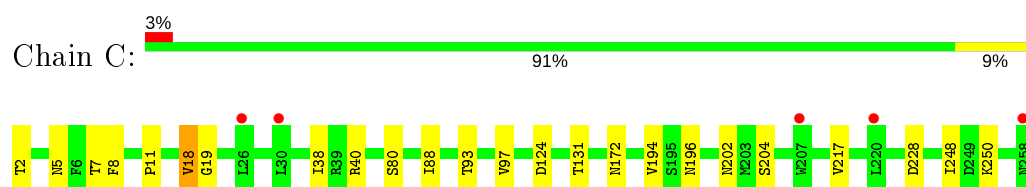
#### • Molecule 1: PUTATIVE RECEPTOR BINDING PROTEIN



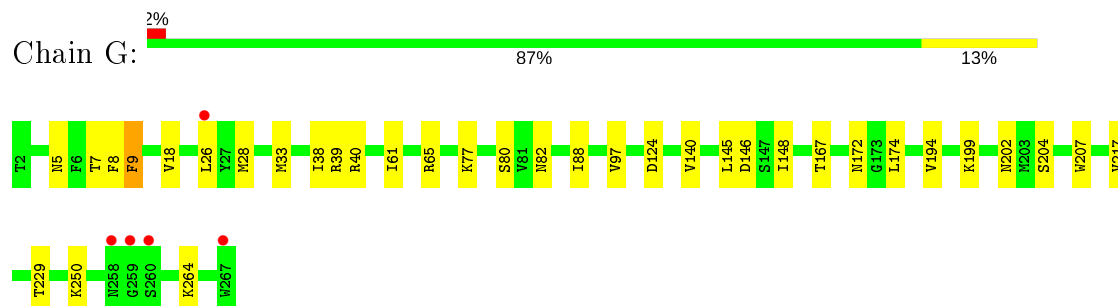
#### • Molecule 1: PUTATIVE RECEPTOR BINDING PROTEIN



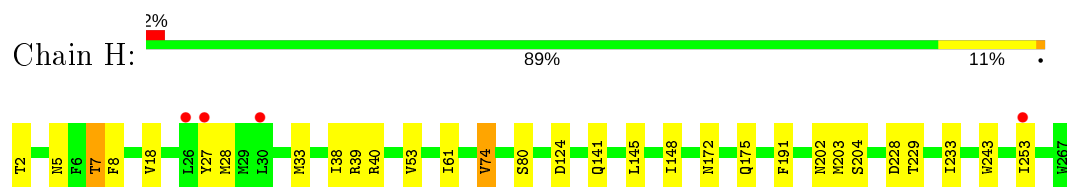
#### • Molecule 1: PUTATIVE RECEPTOR BINDING PROTEIN



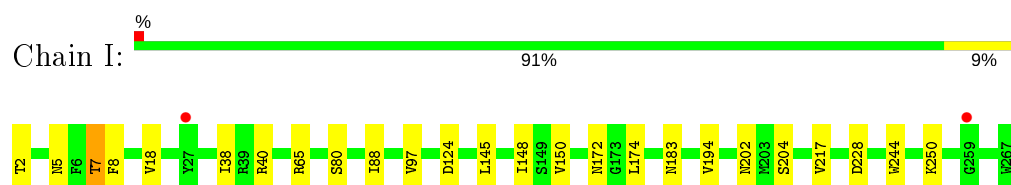
#### • Molecule 1: PUTATIVE RECEPTOR BINDING PROTEIN



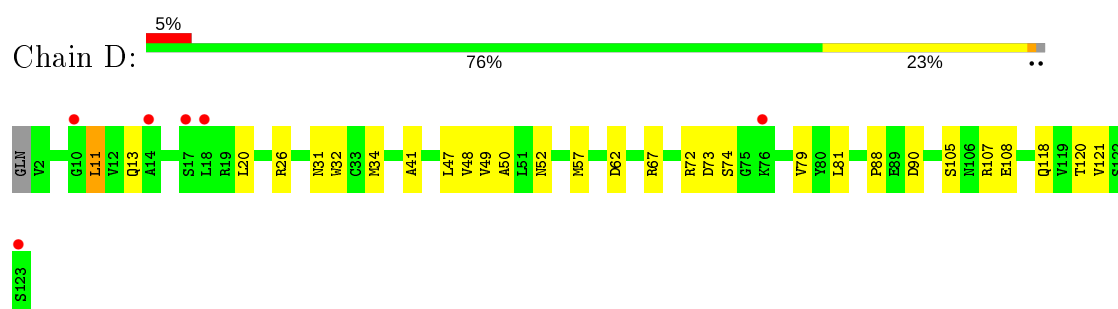
- Molecule 1: PUTATIVE RECEPTOR BINDING PROTEIN



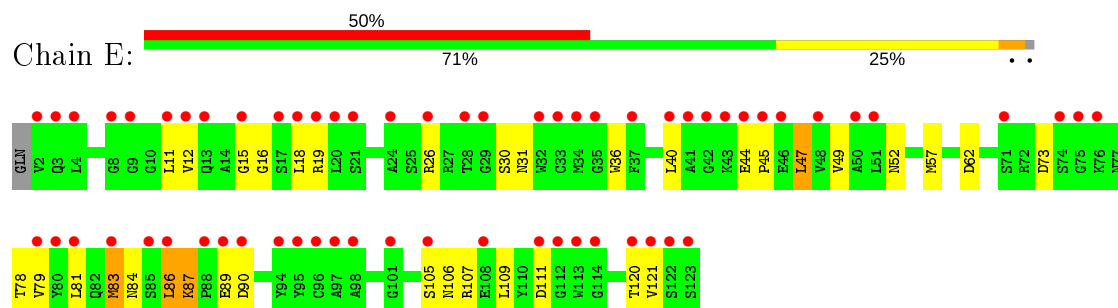
- Molecule 1: PUTATIVE RECEPTOR BINDING PROTEIN



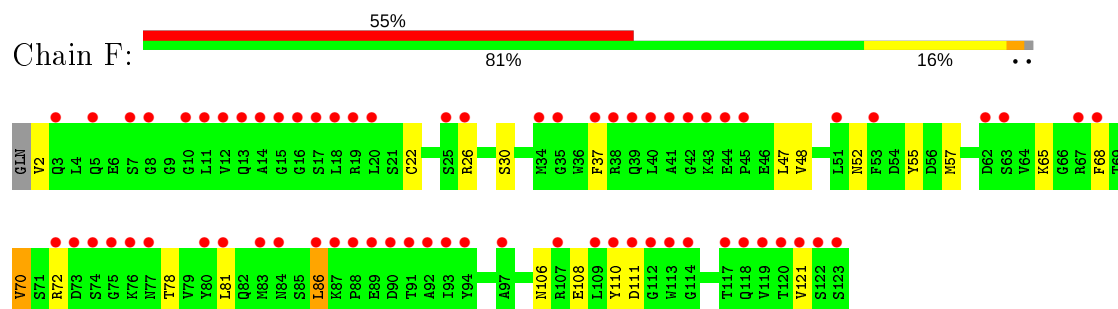
- Molecule 2: CAMELID VHH5



- Molecule 2: CAMELID VHH5

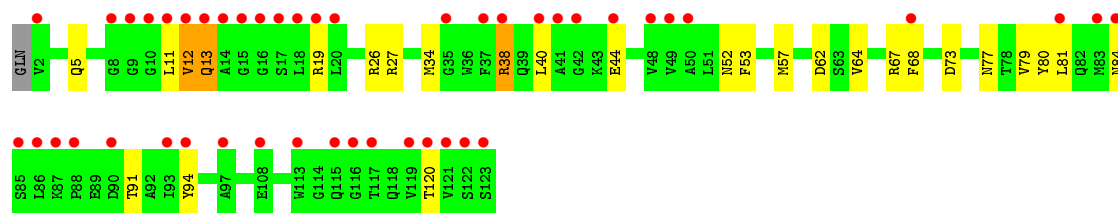


- Molecule 2: CAMELID VHH5

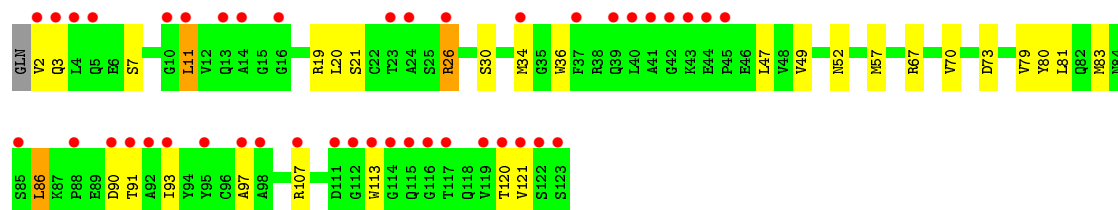


- Molecule 2: CAMELID VHH5

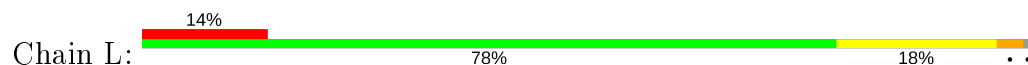




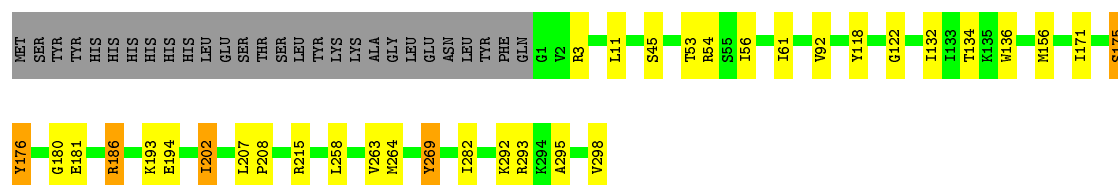
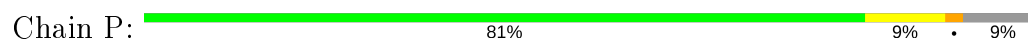
• Molecule 2: CAMELID VHH5



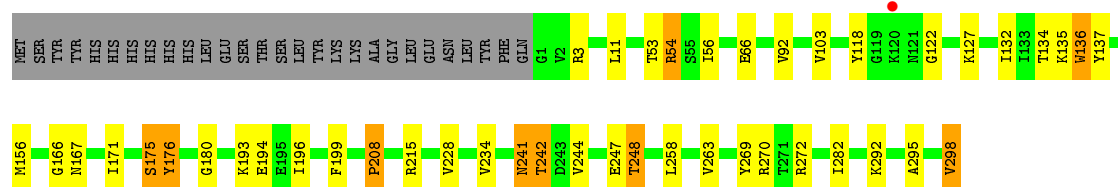
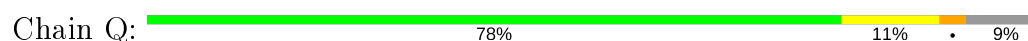
• Molecule 2: CAMELID VHH5



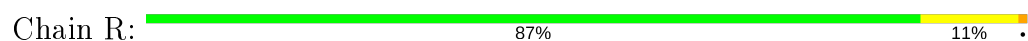
• Molecule 3: LACTOCOCCAL PHAGE P2 ORF15



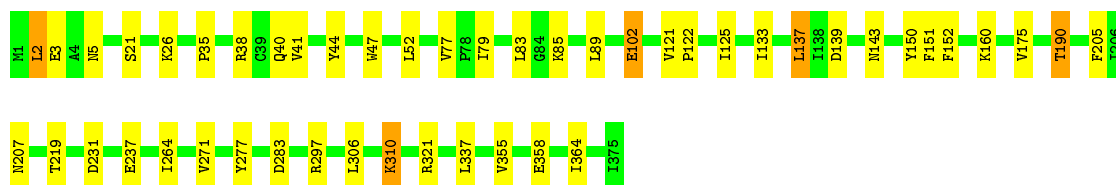
• Molecule 3: LACTOCOCCAL PHAGE P2 ORF15



• Molecule 4: LACTOCOCCAL PHAGE P2 ORF16







## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	202.88Å 202.88Å 760.52Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	38.07 – 2.60 37.91 – 2.60	Depositor EDS
% Data completeness (in resolution range)	(Not available) (38.07-2.60) 99.3 (37.91-2.60)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.56 (at 2.61Å)	Xtriage
Refinement program	BUSTER 2.9.2	Depositor
R, $R_{free}$	0.179 , 0.208 0.184 , 0.212	Depositor DCC
$R_{free}$ test set	3689 reflections (2.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.4	Xtriage
Anisotropy	0.479	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 63.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	27647	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.81% 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	0.38	0/2064	0.64	0/2813
1	B	0.38	0/2064	0.64	0/2813
1	C	0.39	0/2064	0.63	0/2813
1	G	0.42	0/2064	0.67	0/2813
1	H	0.41	0/2064	0.67	0/2813
1	I	0.42	0/2064	0.66	0/2813
2	D	0.39	0/956	0.63	0/1289
2	E	0.43	0/940	0.73	0/1271
2	F	0.43	0/948	0.67	0/1278
2	J	0.43	0/956	0.67	0/1289
2	K	0.41	0/956	0.66	0/1289
2	L	0.42	0/956	0.66	0/1289
3	P	0.52	0/2485	0.79	0/3356
3	Q	0.50	0/2485	0.79	1/3356 (0.0%)
4	R	0.57	0/3094	0.81	0/4208
All	All	0.45	0/26160	0.71	1/35503 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	Q	208	PRO	N-CA-C	5.52	126.44	112.10

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	2024	0	1983	17	0
1	B	2024	0	1983	18	0
1	C	2024	0	1983	16	0
1	G	2024	0	1983	19	0
1	H	2024	0	1983	16	0
1	I	2024	0	1983	11	0
2	D	938	0	891	12	0
2	E	922	0	858	17	0
2	F	931	0	884	8	0
2	J	938	0	891	10	0
2	K	938	0	891	16	0
2	L	938	0	891	12	0
3	P	2432	0	2394	19	0
3	Q	2432	0	2394	27	0
4	R	3025	0	2984	23	0
5	A	47	0	0	0	0
5	B	85	0	0	1	0
5	C	112	0	0	0	0
5	D	11	0	0	0	0
5	E	1	0	0	0	0
5	F	1	0	0	0	0
5	G	199	0	0	2	0
5	H	156	0	0	0	0
5	I	192	0	0	0	0
5	K	3	0	0	0	0
5	L	7	0	0	0	0
5	P	363	0	0	0	0
5	Q	265	0	0	1	0
5	R	567	0	0	3	0
All	All	27647	0	24976	209	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 4.

The worst 5 of 209 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:A:93:THR:HG23	1:C:11:PRO:HG2	1.28	1.08
1:B:203:MET:HE1	1:B:243:TRP:HE3	1.40	0.86
1:A:93:THR:CG2	1:C:11:PRO:HG2	2.06	0.86
1:A:93:THR:HG23	1:C:11:PRO:CG	2.09	0.83
1:A:77:LYS:H	1:A:82:ASN:HD21	1.31	0.77

There are no symmetry-related clashes.

## 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	264/266 (99%)	255 (97%)	9 (3%)	0	100	100
1	B	264/266 (99%)	255 (97%)	9 (3%)	0	100	100
1	C	264/266 (99%)	255 (97%)	9 (3%)	0	100	100
1	G	264/266 (99%)	254 (96%)	10 (4%)	0	100	100
1	H	264/266 (99%)	252 (96%)	12 (4%)	0	100	100
1	I	264/266 (99%)	253 (96%)	11 (4%)	0	100	100
2	D	120/123 (98%)	117 (98%)	3 (2%)	0	100	100
2	E	120/123 (98%)	116 (97%)	4 (3%)	0	100	100
2	F	120/123 (98%)	118 (98%)	2 (2%)	0	100	100
2	J	120/123 (98%)	114 (95%)	6 (5%)	0	100	100
2	K	120/123 (98%)	118 (98%)	2 (2%)	0	100	100
2	L	120/123 (98%)	117 (98%)	3 (2%)	0	100	100
3	P	296/326 (91%)	285 (96%)	10 (3%)	1 (0%)	41	64
3	Q	296/326 (91%)	283 (96%)	12 (4%)	1 (0%)	41	64
4	R	373/375 (100%)	355 (95%)	18 (5%)	0	100	100
All	All	3269/3361 (97%)	3147 (96%)	120 (4%)	2 (0%)	51	75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	P	175	SER
3	Q	175	SER

### 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	228/229 (100%)	219 (96%)	9 (4%)	32	58
1	B	228/229 (100%)	221 (97%)	7 (3%)	40	66
1	C	228/229 (100%)	220 (96%)	8 (4%)	36	62
1	G	228/229 (100%)	217 (95%)	11 (5%)	25	49
1	H	228/229 (100%)	218 (96%)	10 (4%)	28	53
1	I	228/229 (100%)	218 (96%)	10 (4%)	28	53
2	D	99/100 (99%)	89 (90%)	10 (10%)	7	14
2	E	96/100 (96%)	83 (86%)	13 (14%)	4	6
2	F	98/100 (98%)	86 (88%)	12 (12%)	5	9
2	J	99/100 (99%)	85 (86%)	14 (14%)	3	6
2	K	99/100 (99%)	88 (89%)	11 (11%)	6	11
2	L	99/100 (99%)	88 (89%)	11 (11%)	6	11
3	P	264/290 (91%)	250 (95%)	14 (5%)	22	45
3	Q	264/290 (91%)	243 (92%)	21 (8%)	12	24
4	R	340/340 (100%)	318 (94%)	22 (6%)	17	34
All	All	2826/2894 (98%)	2643 (94%)	183 (6%)	17	34

5 of 183 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	18	VAL
2	J	79	VAL
4	R	83	LEU
1	I	80	SER
2	J	12	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 45 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	31	ASN
1	G	196	ASN
3	Q	278	GLN
1	G	71	ASN
1	H	48	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](#)

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, 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(Å²)	Q<0.9
1	A	266/266 (100%)	0.27	17 (6%)	19	14	53, 82, 115, 135	0
1	B	266/266 (100%)	-0.05	12 (4%)	33	26	41, 63, 86, 118	0
1	C	266/266 (100%)	-0.12	7 (2%)	56	50	36, 57, 110, 146	0
1	G	266/266 (100%)	-0.33	5 (1%)	66	62	28, 42, 79, 111	0
1	H	266/266 (100%)	-0.18	4 (1%)	73	70	27, 43, 90, 113	0
1	I	266/266 (100%)	-0.40	2 (0%)	86	84	25, 40, 83, 103	0
2	D	122/123 (99%)	0.21	6 (4%)	29	23	61, 80, 97, 108	0
2	E	122/123 (99%)	2.19	62 (50%)	0	0	102, 138, 158, 170	0
2	F	122/123 (99%)	2.67	68 (55%)	0	0	112, 156, 177, 197	0
2	J	122/123 (99%)	1.73	46 (37%)	0	0	93, 131, 152, 168	1 (0%)
2	K	122/123 (99%)	1.55	43 (35%)	0	0	80, 119, 142, 159	0
2	L	122/123 (99%)	0.56	17 (13%)	2	1	72, 101, 129, 145	0
3	P	298/326 (91%)	-0.67	0	100	100	20, 34, 57, 89	0
3	Q	298/326 (91%)	-0.73	1 (0%)	94	93	23, 40, 63, 93	0
4	R	375/375 (100%)	-0.59	0	100	100	17, 26, 45, 77	0
All	All	3299/3361 (98%)	0.07	290 (8%)	10	7	17, 57, 143, 197	1 (0%)

The worst 5 of 290 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	109	LEU	9.6
2	F	11	LEU	9.4
2	E	18	LEU	8.6
2	K	112	GLY	8.4
2	K	114	GLY	8.1



## 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.