



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 10, 2023 – 10:38 AM EDT

PDB ID : 7KMV  
Title : Structure of Malaysian Banana Lectin F84T  
Authors : Meagher, J.L.; Stuckey, J.A.  
Deposited on : 2020-11-03  
Resolution : 1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

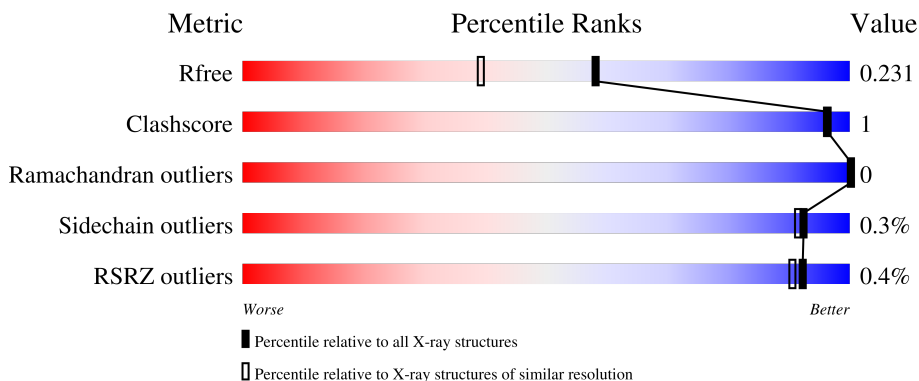
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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))
$R_{free}$	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	150	91% (0% poor fit, 7% not modelled)
1	B	150	91% (0% poor fit, 7% not modelled)
1	C	150	92% (0% poor fit, 7% not modelled)
1	D	150	91% (1% poor fit, 6% not modelled)
1	E	150	93% (1% poor fit, 6% not modelled)

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	150	<p>%</p> <p>87% 5% • 7%</p>
1	G	150	<p>93% • 7%</p>
1	H	150	<p>90% • 7%</p>
1	I	150	<p>91% • 6%</p>
1	J	150	<p>90% • 7%</p>
1	K	150	<p>90% • 6%</p>
1	L	150	<p>89% • 7%</p>
1	M	150	<p>3%</p> <p>93% • 6%</p>
1	N	150	<p>91% • 7%</p>
1	O	150	<p>94% 6%</p>
1	P	150	<p>89% • 7%</p>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 19033 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 Jacalin-type lectin domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	140	1052	674	172	202	4	0	0	0
1	B	140	1067	682	173	208	4	0	2	0
1	C	140	1052	674	172	202	4	0	0	0
1	D	141	1063	680	173	206	4	0	1	0
1	E	141	1056	676	173	203	4	0	0	0
1	F	140	1068	684	173	207	4	0	3	0
1	G	140	1059	679	172	204	4	0	1	0
1	H	140	1055	675	172	204	4	0	0	0
1	I	141	1063	680	173	206	4	0	1	0
1	J	140	1067	683	175	205	4	0	2	0
1	K	141	1061	681	173	203	4	0	1	0
1	L	140	1059	678	172	205	4	0	1	0
1	M	141	1050	673	170	203	4	0	0	0
1	N	140	1057	677	172	204	4	0	1	0
1	O	141	1064	682	173	205	4	0	1	0
1	P	139	1051	674	171	202	4	0	1	0

There are 144 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	84	THR	PHE	engineered mutation	UNP M0TZ81
A	143	LEU	-	expression tag	UNP M0TZ81
A	144	GLU	-	expression tag	UNP M0TZ81
A	145	HIS	-	expression tag	UNP M0TZ81
A	146	HIS	-	expression tag	UNP M0TZ81
A	147	HIS	-	expression tag	UNP M0TZ81
A	148	HIS	-	expression tag	UNP M0TZ81
A	149	HIS	-	expression tag	UNP M0TZ81
A	150	HIS	-	expression tag	UNP M0TZ81
B	84	THR	PHE	engineered mutation	UNP M0TZ81
B	143	LEU	-	expression tag	UNP M0TZ81
B	144	GLU	-	expression tag	UNP M0TZ81
B	145	HIS	-	expression tag	UNP M0TZ81
B	146	HIS	-	expression tag	UNP M0TZ81
B	147	HIS	-	expression tag	UNP M0TZ81
B	148	HIS	-	expression tag	UNP M0TZ81
B	149	HIS	-	expression tag	UNP M0TZ81
B	150	HIS	-	expression tag	UNP M0TZ81
C	84	THR	PHE	engineered mutation	UNP M0TZ81
C	143	LEU	-	expression tag	UNP M0TZ81
C	144	GLU	-	expression tag	UNP M0TZ81
C	145	HIS	-	expression tag	UNP M0TZ81
C	146	HIS	-	expression tag	UNP M0TZ81
C	147	HIS	-	expression tag	UNP M0TZ81
C	148	HIS	-	expression tag	UNP M0TZ81
C	149	HIS	-	expression tag	UNP M0TZ81
C	150	HIS	-	expression tag	UNP M0TZ81
D	84	THR	PHE	engineered mutation	UNP M0TZ81
D	143	LEU	-	expression tag	UNP M0TZ81
D	144	GLU	-	expression tag	UNP M0TZ81
D	145	HIS	-	expression tag	UNP M0TZ81
D	146	HIS	-	expression tag	UNP M0TZ81
D	147	HIS	-	expression tag	UNP M0TZ81
D	148	HIS	-	expression tag	UNP M0TZ81
D	149	HIS	-	expression tag	UNP M0TZ81
D	150	HIS	-	expression tag	UNP M0TZ81
E	84	THR	PHE	engineered mutation	UNP M0TZ81
E	143	LEU	-	expression tag	UNP M0TZ81
E	144	GLU	-	expression tag	UNP M0TZ81
E	145	HIS	-	expression tag	UNP M0TZ81
E	146	HIS	-	expression tag	UNP M0TZ81
E	147	HIS	-	expression tag	UNP M0TZ81

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	148	HIS	-	expression tag	UNP M0TZ81
E	149	HIS	-	expression tag	UNP M0TZ81
E	150	HIS	-	expression tag	UNP M0TZ81
F	84	THR	PHE	engineered mutation	UNP M0TZ81
F	143	LEU	-	expression tag	UNP M0TZ81
F	144	GLU	-	expression tag	UNP M0TZ81
F	145	HIS	-	expression tag	UNP M0TZ81
F	146	HIS	-	expression tag	UNP M0TZ81
F	147	HIS	-	expression tag	UNP M0TZ81
F	148	HIS	-	expression tag	UNP M0TZ81
F	149	HIS	-	expression tag	UNP M0TZ81
F	150	HIS	-	expression tag	UNP M0TZ81
G	84	THR	PHE	engineered mutation	UNP M0TZ81
G	143	LEU	-	expression tag	UNP M0TZ81
G	144	GLU	-	expression tag	UNP M0TZ81
G	145	HIS	-	expression tag	UNP M0TZ81
G	146	HIS	-	expression tag	UNP M0TZ81
G	147	HIS	-	expression tag	UNP M0TZ81
G	148	HIS	-	expression tag	UNP M0TZ81
G	149	HIS	-	expression tag	UNP M0TZ81
G	150	HIS	-	expression tag	UNP M0TZ81
H	84	THR	PHE	engineered mutation	UNP M0TZ81
H	143	LEU	-	expression tag	UNP M0TZ81
H	144	GLU	-	expression tag	UNP M0TZ81
H	145	HIS	-	expression tag	UNP M0TZ81
H	146	HIS	-	expression tag	UNP M0TZ81
H	147	HIS	-	expression tag	UNP M0TZ81
H	148	HIS	-	expression tag	UNP M0TZ81
H	149	HIS	-	expression tag	UNP M0TZ81
H	150	HIS	-	expression tag	UNP M0TZ81
I	84	THR	PHE	engineered mutation	UNP M0TZ81
I	143	LEU	-	expression tag	UNP M0TZ81
I	144	GLU	-	expression tag	UNP M0TZ81
I	145	HIS	-	expression tag	UNP M0TZ81
I	146	HIS	-	expression tag	UNP M0TZ81
I	147	HIS	-	expression tag	UNP M0TZ81
I	148	HIS	-	expression tag	UNP M0TZ81
I	149	HIS	-	expression tag	UNP M0TZ81
I	150	HIS	-	expression tag	UNP M0TZ81
J	84	THR	PHE	engineered mutation	UNP M0TZ81
J	143	LEU	-	expression tag	UNP M0TZ81
J	144	GLU	-	expression tag	UNP M0TZ81

*Continued on next page...*

*Continued from previous page...*

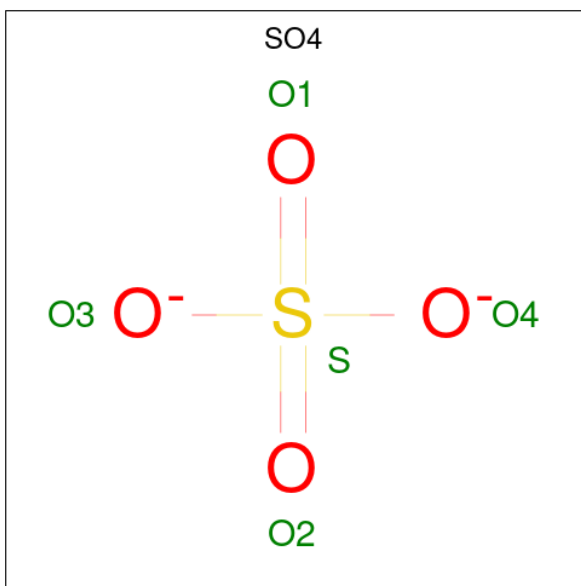
Chain	Residue	Modelled	Actual	Comment	Reference
J	145	HIS	-	expression tag	UNP M0TZ81
J	146	HIS	-	expression tag	UNP M0TZ81
J	147	HIS	-	expression tag	UNP M0TZ81
J	148	HIS	-	expression tag	UNP M0TZ81
J	149	HIS	-	expression tag	UNP M0TZ81
J	150	HIS	-	expression tag	UNP M0TZ81
K	84	THR	PHE	engineered mutation	UNP M0TZ81
K	143	LEU	-	expression tag	UNP M0TZ81
K	144	GLU	-	expression tag	UNP M0TZ81
K	145	HIS	-	expression tag	UNP M0TZ81
K	146	HIS	-	expression tag	UNP M0TZ81
K	147	HIS	-	expression tag	UNP M0TZ81
K	148	HIS	-	expression tag	UNP M0TZ81
K	149	HIS	-	expression tag	UNP M0TZ81
K	150	HIS	-	expression tag	UNP M0TZ81
L	84	THR	PHE	engineered mutation	UNP M0TZ81
L	143	LEU	-	expression tag	UNP M0TZ81
L	144	GLU	-	expression tag	UNP M0TZ81
L	145	HIS	-	expression tag	UNP M0TZ81
L	146	HIS	-	expression tag	UNP M0TZ81
L	147	HIS	-	expression tag	UNP M0TZ81
L	148	HIS	-	expression tag	UNP M0TZ81
L	149	HIS	-	expression tag	UNP M0TZ81
L	150	HIS	-	expression tag	UNP M0TZ81
M	84	THR	PHE	engineered mutation	UNP M0TZ81
M	143	LEU	-	expression tag	UNP M0TZ81
M	144	GLU	-	expression tag	UNP M0TZ81
M	145	HIS	-	expression tag	UNP M0TZ81
M	146	HIS	-	expression tag	UNP M0TZ81
M	147	HIS	-	expression tag	UNP M0TZ81
M	148	HIS	-	expression tag	UNP M0TZ81
M	149	HIS	-	expression tag	UNP M0TZ81
M	150	HIS	-	expression tag	UNP M0TZ81
N	84	THR	PHE	engineered mutation	UNP M0TZ81
N	143	LEU	-	expression tag	UNP M0TZ81
N	144	GLU	-	expression tag	UNP M0TZ81
N	145	HIS	-	expression tag	UNP M0TZ81
N	146	HIS	-	expression tag	UNP M0TZ81
N	147	HIS	-	expression tag	UNP M0TZ81
N	148	HIS	-	expression tag	UNP M0TZ81
N	149	HIS	-	expression tag	UNP M0TZ81
N	150	HIS	-	expression tag	UNP M0TZ81

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
O	84	THR	PHE	engineered mutation	UNP M0TZ81
O	143	LEU	-	expression tag	UNP M0TZ81
O	144	GLU	-	expression tag	UNP M0TZ81
O	145	HIS	-	expression tag	UNP M0TZ81
O	146	HIS	-	expression tag	UNP M0TZ81
O	147	HIS	-	expression tag	UNP M0TZ81
O	148	HIS	-	expression tag	UNP M0TZ81
O	149	HIS	-	expression tag	UNP M0TZ81
O	150	HIS	-	expression tag	UNP M0TZ81
P	84	THR	PHE	engineered mutation	UNP M0TZ81
P	143	LEU	-	expression tag	UNP M0TZ81
P	144	GLU	-	expression tag	UNP M0TZ81
P	145	HIS	-	expression tag	UNP M0TZ81
P	146	HIS	-	expression tag	UNP M0TZ81
P	147	HIS	-	expression tag	UNP M0TZ81
P	148	HIS	-	expression tag	UNP M0TZ81
P	149	HIS	-	expression tag	UNP M0TZ81
P	150	HIS	-	expression tag	UNP M0TZ81

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

Continued on next page...



*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		
2	G	1	Total	O	S	0	0
			5	4	1		
2	H	1	Total	O	S	0	0
			5	4	1		
2	I	1	Total	O	S	0	0
			5	4	1		
2	J	1	Total	O	S	0	0
			5	4	1		
2	K	1	Total	O	S	0	0
			5	4	1		
2	K	1	Total	O	S	0	0
			5	4	1		
2	L	1	Total	O	S	0	0
			5	4	1		
2	M	1	Total	O	S	0	0
			5	4	1		
2	N	1	Total	O	S	0	0
			5	4	1		
2	O	1	Total	O	S	0	0
			5	4	1		
2	P	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	132	Total	O	0	0
			132	132		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	143	Total 143	O 143	0	0
3	C	132	Total 132	O 132	0	0
3	D	134	Total 134	O 134	0	0
3	E	98	Total 98	O 98	0	0
3	F	147	Total 147	O 147	0	0
3	G	122	Total 122	O 122	0	0
3	H	109	Total 109	O 109	0	0
3	I	97	Total 97	O 97	0	0
3	J	141	Total 141	O 141	0	0
3	K	130	Total 130	O 130	0	0
3	L	146	Total 146	O 146	0	0
3	M	90	Total 90	O 90	0	0
3	N	101	Total 101	O 101	0	0
3	O	149	Total 149	O 149	0	0
3	P	118	Total 118	O 118	0	0

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


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 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.

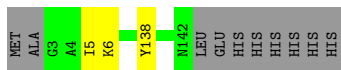
- Molecule 1: Jacalin-type lectin domain-containing protein

Chain A:  91% 7%



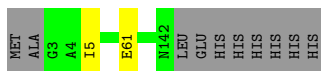
- Molecule 1: Jacalin-type lectin domain-containing protein

Chain B:  91% 7%

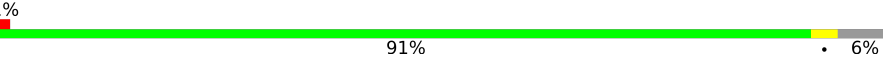


- Molecule 1: Jacalin-type lectin domain-containing protein

Chain C:  92% 7%

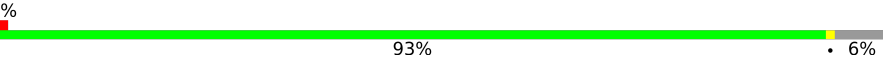


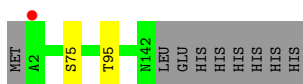
- Molecule 1: Jacalin-type lectin domain-containing protein

Chain D:  91% 6%

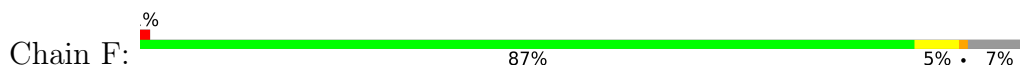


- Molecule 1: Jacalin-type lectin domain-containing protein

Chain E:  93% 6%



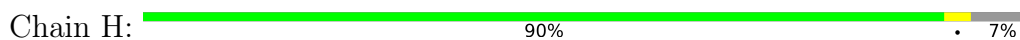
- Molecule 1: Jacalin-type lectin domain-containing protein



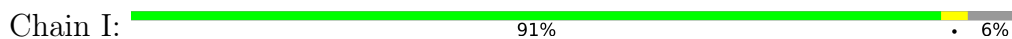
- Molecule 1: Jacalin-type lectin domain-containing protein



- Molecule 1: Jacalin-type lectin domain-containing protein



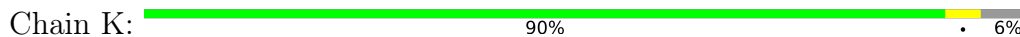
- Molecule 1: Jacalin-type lectin domain-containing protein



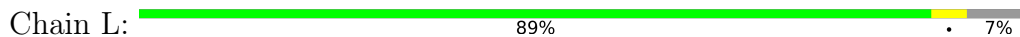
- Molecule 1: Jacalin-type lectin domain-containing protein



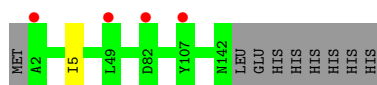
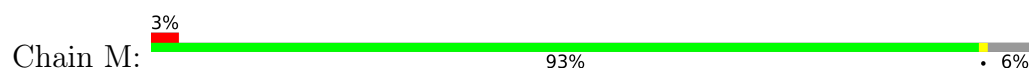
- Molecule 1: Jacalin-type lectin domain-containing protein



- Molecule 1: Jacalin-type lectin domain-containing protein



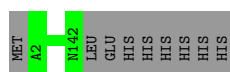
- Molecule 1: Jacalin-type lectin domain-containing protein



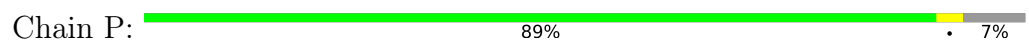
- Molecule 1: Jacalin-type lectin domain-containing protein



- Molecule 1: Jacalin-type lectin domain-containing protein



- Molecule 1: Jacalin-type lectin domain-containing protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.25Å 157.87Å 204.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.06 – 1.80 49.06 – 1.80	Depositor EDS
% Data completeness (in resolution range)	97.7 (49.06-1.80) 97.6 (49.06-1.80)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.56 (at 1.81Å)	Xtrriage
Refinement program	BUSTER 2.10.3	Depositor
R, $R_{free}$	0.198 , 0.230 0.196 , 0.231	Depositor DCC
$R_{free}$ test set	11266 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.3	Xtrriage
Anisotropy	0.018	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 47.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	19033	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 38.94 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.4186e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.45	0/1079	0.66	0/1462
1	B	0.47	0/1097	0.65	0/1487
1	C	0.46	0/1079	0.66	0/1462
1	D	0.47	0/1093	0.67	0/1481
1	E	0.44	0/1083	0.66	0/1468
1	F	0.48	0/1104	0.68	0/1496
1	G	0.46	0/1089	0.67	0/1476
1	H	0.44	0/1082	0.67	0/1466
1	I	0.46	0/1093	0.66	0/1481
1	J	0.46	0/1100	0.68	0/1490
1	K	0.47	0/1091	0.65	0/1479
1	L	0.45	0/1089	0.67	0/1476
1	M	0.46	0/1077	0.69	0/1461
1	N	0.42	0/1087	0.66	0/1473
1	O	0.46	0/1094	0.67	0/1483
1	P	0.46	0/1081	0.68	0/1465
All	All	0.46	0/17418	0.67	0/23606

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

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	1052	0	1003	2	0
1	B	1067	0	1015	2	0
1	C	1052	0	1003	3	0
1	D	1063	0	1015	4	0
1	E	1056	0	1006	1	0
1	F	1068	0	1025	5	0
1	G	1059	0	1014	1	0
1	H	1055	0	1005	4	0
1	I	1063	0	1015	2	0
1	J	1067	0	1025	4	0
1	K	1061	0	1017	3	0
1	L	1059	0	1012	3	0
1	M	1050	0	995	1	0
1	N	1057	0	1007	2	0
1	O	1064	0	1019	0	0
1	P	1051	0	1005	3	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	D	10	0	0	0	0
2	E	5	0	0	0	0
2	F	15	0	0	0	0
2	G	5	0	0	0	0
2	H	5	0	0	0	0
2	I	5	0	0	0	0
2	J	5	0	0	0	0
2	K	10	0	0	0	0
2	L	5	0	0	0	0
2	M	5	0	0	0	0
2	N	5	0	0	0	0
2	O	5	0	0	0	0
2	P	5	0	0	0	0
3	A	132	0	0	0	0
3	B	143	0	0	0	0
3	C	132	0	0	1	0
3	D	134	0	0	0	0
3	E	98	0	0	0	0
3	F	147	0	0	0	0
3	G	122	0	0	0	0
3	H	109	0	0	0	0
3	I	97	0	0	0	0
3	J	141	0	0	0	0
3	K	130	0	0	0	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	L	146	0	0	0	0
3	M	90	0	0	0	0
3	N	101	0	0	0	0
3	O	149	0	0	0	0
3	P	118	0	0	0	0
All	All	19033	0	16181	33	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 1.

All (33) 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:P:27:ASP:OD2	1:P:45:THR:HG23	1.84	0.77
1:L:27:ASP:OD2	1:L:45[A]:THR:HG23	1.89	0.72
1:J:27:ASP:OD2	1:J:45[A]:THR:HG23	1.90	0.71
1:F:45[A]:THR:HG22	1:F:50:THR:OG1	1.91	0.71
1:L:45[A]:THR:HG22	1:L:50:THR:OG1	1.94	0.68
1:J:45[A]:THR:HG22	1:J:50:THR:OG1	1.94	0.68
1:G:5:ILE:HD12	1:H:5:ILE:HD12	1.77	0.66
1:F:27:ASP:OD2	1:F:45[A]:THR:HG23	1.97	0.64
1:M:5:ILE:HD12	1:N:5:ILE:HD12	1.80	0.63
1:P:6:LYS:HD3	1:P:138:TYR:CZ	2.37	0.59
1:P:45:THR:HG22	1:P:50:THR:OG1	2.04	0.58
1:L:6:LYS:HD3	1:L:138:TYR:CZ	2.39	0.58
1:J:6:LYS:HD3	1:J:138:TYR:CZ	2.40	0.56
1:A:5:ILE:HD12	1:B:5:ILE:HD12	1.89	0.55
1:H:27:ASP:OD2	1:H:45:THR:HG23	2.06	0.55
1:D:6:LYS:HD3	1:D:138:TYR:CZ	2.43	0.54
1:F:6:LYS:HD3	1:F:138:TYR:CZ	2.44	0.52
1:I:75:SER:HB3	1:I:95:THR:HB	1.92	0.51
1:H:45:THR:HG22	1:H:50:THR:OG1	2.11	0.50
1:B:6:LYS:HD3	1:B:138:TYR:CZ	2.47	0.50
1:C:61:GLU:HG2	3:C:419:HOH:O	2.12	0.50
1:E:75:SER:HB3	1:E:95:THR:HB	1.93	0.49
1:K:61:GLU:HG2	1:K:62:PRO:HD2	1.96	0.47
1:F:81:VAL:HG11	1:F:107:TYR:HB3	1.95	0.47
1:D:140:MET:HG2	1:H:47:TYR:HB3	1.95	0.46
1:F:120:LYS:HE3	1:F:140:MET:HE2	1.99	0.45
1:I:91:LYS:HA	1:I:104:PHE:O	2.19	0.43
1:C:5:ILE:CD1	1:D:5:ILE:HD12	2.49	0.42

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:81:VAL:HG11	1:A:107:TYR:HB3	2.00	0.42
1:K:75:SER:HB3	1:K:95:THR:HB	2.02	0.42
1:N:81:VAL:HG11	1:N:107:TYR:HB3	2.02	0.41
1:J:50:THR:O	1:K:47:TYR:HE1	2.04	0.41
1:C:5:ILE:HD12	1:D:5:ILE:HD12	2.03	0.40

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	138/150 (92%)	133 (96%)	5 (4%)	0	100	100
1	B	140/150 (93%)	135 (96%)	5 (4%)	0	100	100
1	C	138/150 (92%)	134 (97%)	4 (3%)	0	100	100
1	D	140/150 (93%)	136 (97%)	4 (3%)	0	100	100
1	E	139/150 (93%)	135 (97%)	4 (3%)	0	100	100
1	F	141/150 (94%)	135 (96%)	6 (4%)	0	100	100
1	G	139/150 (93%)	135 (97%)	4 (3%)	0	100	100
1	H	138/150 (92%)	133 (96%)	5 (4%)	0	100	100
1	I	140/150 (93%)	135 (96%)	5 (4%)	0	100	100
1	J	140/150 (93%)	137 (98%)	3 (2%)	0	100	100
1	K	140/150 (93%)	135 (96%)	5 (4%)	0	100	100
1	L	139/150 (93%)	135 (97%)	4 (3%)	0	100	100
1	M	139/150 (93%)	134 (96%)	5 (4%)	0	100	100
1	N	139/150 (93%)	134 (96%)	5 (4%)	0	100	100
1	O	140/150 (93%)	135 (96%)	5 (4%)	0	100	100
1	P	138/150 (92%)	133 (96%)	5 (4%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	2228/2400 (93%)	2154 (97%)	74 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	104/114 (91%)	103 (99%)	1 (1%)	76	71
1	B	107/114 (94%)	107 (100%)	0	100	100
1	C	104/114 (91%)	104 (100%)	0	100	100
1	D	106/114 (93%)	106 (100%)	0	100	100
1	E	104/114 (91%)	104 (100%)	0	100	100
1	F	108/114 (95%)	107 (99%)	1 (1%)	78	75
1	G	106/114 (93%)	106 (100%)	0	100	100
1	H	105/114 (92%)	105 (100%)	0	100	100
1	I	106/114 (93%)	105 (99%)	1 (1%)	78	75
1	J	107/114 (94%)	107 (100%)	0	100	100
1	K	105/114 (92%)	104 (99%)	1 (1%)	76	71
1	L	106/114 (93%)	105 (99%)	1 (1%)	78	75
1	M	103/114 (90%)	103 (100%)	0	100	100
1	N	105/114 (92%)	105 (100%)	0	100	100
1	O	106/114 (93%)	106 (100%)	0	100	100
1	P	105/114 (92%)	105 (100%)	0	100	100
All	All	1687/1824 (92%)	1682 (100%)	5 (0%)	92	91

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	142	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	107	TYR
1	I	69	ASP
1	K	114	ILE
1	L	75	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

20 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	SO4	G	201	-	4,4,4	0.16	0	6,6,6	0.05	0
2	SO4	L	201	-	4,4,4	0.16	0	6,6,6	0.06	0
2	SO4	F	201	-	4,4,4	0.16	0	6,6,6	0.05	0
2	SO4	J	201	-	4,4,4	0.13	0	6,6,6	0.06	0
2	SO4	C	201	-	4,4,4	0.17	0	6,6,6	0.06	0
2	SO4	D	202	-	4,4,4	0.18	0	6,6,6	0.08	0
2	SO4	D	201	-	4,4,4	0.17	0	6,6,6	0.12	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SO4	M	201	-	4,4,4	0.16	0	6,6,6	0.07	0
2	SO4	F	203	-	4,4,4	0.20	0	6,6,6	0.11	0
2	SO4	P	201	-	4,4,4	0.11	0	6,6,6	0.07	0
2	SO4	K	202	-	4,4,4	0.11	0	6,6,6	0.14	0
2	SO4	E	201	-	4,4,4	0.13	0	6,6,6	0.07	0
2	SO4	B	201	-	4,4,4	0.24	0	6,6,6	0.07	0
2	SO4	H	201	-	4,4,4	0.15	0	6,6,6	0.05	0
2	SO4	N	201	-	4,4,4	0.15	0	6,6,6	0.09	0
2	SO4	O	201	-	4,4,4	0.15	0	6,6,6	0.11	0
2	SO4	K	201	-	4,4,4	0.25	0	6,6,6	0.10	0
2	SO4	F	202	-	4,4,4	0.38	0	6,6,6	0.11	0
2	SO4	A	201	-	4,4,4	0.15	0	6,6,6	0.06	0
2	SO4	I	201	-	4,4,4	0.16	0	6,6,6	0.10	0

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	140/150 (93%)	-0.31	0 100 100	13, 19, 35, 42	0
1	B	140/150 (93%)	-0.50	0 100 100	10, 16, 28, 47	0
1	C	140/150 (93%)	-0.42	0 100 100	11, 17, 32, 41	0
1	D	141/150 (94%)	-0.48	2 (1%) 75 72	10, 17, 28, 42	0
1	E	141/150 (94%)	-0.25	1 (0%) 87 86	11, 21, 37, 52	0
1	F	140/150 (93%)	-0.44	1 (0%) 87 86	12, 17, 31, 49	0
1	G	140/150 (93%)	-0.43	0 100 100	9, 18, 29, 44	0
1	H	140/150 (93%)	-0.22	0 100 100	15, 22, 37, 47	0
1	I	141/150 (94%)	-0.36	0 100 100	11, 19, 31, 43	0
1	J	140/150 (93%)	-0.41	0 100 100	13, 19, 32, 42	0
1	K	141/150 (94%)	-0.41	0 100 100	10, 17, 29, 41	0
1	L	140/150 (93%)	-0.45	0 100 100	12, 17, 32, 41	0
1	M	141/150 (94%)	0.03	4 (2%) 53 47	13, 25, 41, 51	0
1	N	140/150 (93%)	-0.32	0 100 100	14, 23, 37, 43	0
1	O	141/150 (94%)	-0.45	0 100 100	10, 17, 29, 43	0
1	P	139/150 (92%)	-0.34	0 100 100	13, 20, 35, 43	0
All	All	2245/2400 (93%)	-0.36	8 (0%) 92 90	9, 19, 34, 52	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	2	ALA	3.3
1	M	49	LEU	3.0
1	D	2	ALA	2.8
1	F	24	TYR	2.4
1	M	107	TYR	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	24	TYR	2.3
1	M	82	ASP	2.2
1	M	2	ALA	2.1

## 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 monosaccharides in this entry.

## 6.4 Ligands [i](#)

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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
2	SO4	H	201	5/5	0.48	0.30	146,146,146,146	0
2	SO4	F	202	5/5	0.61	0.30	51,56,57,58	0
2	SO4	C	201	5/5	0.65	0.18	95,95,95,96	0
2	SO4	A	201	5/5	0.67	0.17	100,100,101,101	0
2	SO4	D	202	5/5	0.72	0.36	81,83,83,84	0
2	SO4	L	201	5/5	0.73	0.18	108,108,109,109	0
2	SO4	F	203	5/5	0.76	0.25	70,73,74,74	0
2	SO4	F	201	5/5	0.80	0.14	105,105,105,105	0
2	SO4	P	201	5/5	0.80	0.18	112,112,112,112	0
2	SO4	I	201	5/5	0.82	0.13	81,81,82,82	0
2	SO4	N	201	5/5	0.83	0.16	81,81,82,83	0
2	SO4	K	201	5/5	0.85	0.17	58,58,59,61	0
2	SO4	J	201	5/5	0.85	0.15	106,106,106,107	0
2	SO4	E	201	5/5	0.87	0.17	90,90,90,90	0
2	SO4	G	201	5/5	0.89	0.14	71,72,72,73	0
2	SO4	B	201	5/5	0.89	0.14	56,57,58,59	0
2	SO4	O	201	5/5	0.90	0.17	57,58,60,61	0
2	SO4	D	201	5/5	0.90	0.14	67,67,68,69	0
2	SO4	M	201	5/5	0.93	0.10	87,88,88,88	0
2	SO4	K	202	5/5	0.96	0.21	57,57,58,58	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.