



# Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 04:34 pm BST

PDB ID : 4ZJM  
Title : Crystal Structure of Mycobacterium tuberculosis LpqH (Rv3763)  
Authors : Arbing, M.A.; Chan, S.; Kuo, E.; Harris, L.R.; Zhou, T.T.; Eisenberg, D.; TB Structural Genomics Consortium (TBSGC)  
Deposited on : 2015-04-29  
Resolution : 2.85 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 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	:	2.11
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

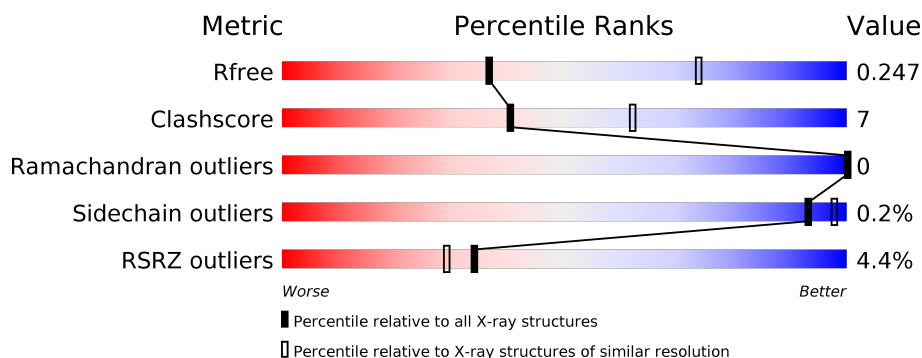
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.85 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	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

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	119	<div> <div>2%</div> <div> <div></div> <div>80%</div> <div>13%</div> <div>7%</div> </div> </div>
2	B	119	<div> <div>3%</div> <div> <div></div> <div>80%</div> <div>14%</div> <div>6%</div> </div> </div>
3	C	119	<div> <div>3%</div> <div> <div></div> <div>78%</div> <div>15%</div> <div>7%</div> </div> </div>
4	D	119	<div> <div>7%</div> <div> <div></div> <div>82%</div> <div>10%</div> <div>8%</div> </div> </div>
5	E	119	<div> <div>6%</div> <div> <div></div> <div>86%</div> <div>11%</div> <div>•</div> </div> </div>
6	F	119	<div> <div>9%</div> <div> <div></div> <div>85%</div> <div>10%</div> <div>5%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	G	119	<div><div></div><div>81%</div><div>13%</div><div>• 6%</div></div>

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 5497 atoms, of which 1 is hydrogen 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 Lipoprotein LpqH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	111	Total	C	N	O	S	0	0	0
			766	470	130	162	4			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	139	THR	-	expression tag	UNP P9WK61
A	140	HIS	-	expression tag	UNP P9WK61
A	141	HIS	-	expression tag	UNP P9WK61
A	142	HIS	-	expression tag	UNP P9WK61
A	143	HIS	-	expression tag	UNP P9WK61
A	144	HIS	-	expression tag	UNP P9WK61
A	145	HIS	-	expression tag	UNP P9WK61

- Molecule 2 is a protein called Lipoprotein LpqH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	112	Total	C	H	N	O	S	0	0
			766	467	1	131	163	4		

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	139	THR	-	expression tag	UNP P9WK61
B	140	HIS	-	expression tag	UNP P9WK61
B	141	HIS	-	expression tag	UNP P9WK61
B	142	HIS	-	expression tag	UNP P9WK61
B	143	HIS	-	expression tag	UNP P9WK61
B	144	HIS	-	expression tag	UNP P9WK61
B	145	HIS	-	expression tag	UNP P9WK61

- Molecule 3 is a protein called Lipoprotein LpqH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	111	Total	C	N	O	S	0	0	0
			757	462	130	161	4			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	139	THR	-	expression tag	UNP P9WK61
C	140	HIS	-	expression tag	UNP P9WK61
C	141	HIS	-	expression tag	UNP P9WK61
C	142	HIS	-	expression tag	UNP P9WK61
C	143	HIS	-	expression tag	UNP P9WK61
C	144	HIS	-	expression tag	UNP P9WK61
C	145	HIS	-	expression tag	UNP P9WK61

- Molecule 4 is a protein called Lipoprotein LpqH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	110	Total	C	N	O	S	0	0	0
			751	459	129	159	4			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	139	THR	-	expression tag	UNP P9WK61
D	140	HIS	-	expression tag	UNP P9WK61
D	141	HIS	-	expression tag	UNP P9WK61
D	142	HIS	-	expression tag	UNP P9WK61
D	143	HIS	-	expression tag	UNP P9WK61
D	144	HIS	-	expression tag	UNP P9WK61
D	145	HIS	-	expression tag	UNP P9WK61

- Molecule 5 is a protein called Lipoprotein LpqH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	115	Total	C	N	O	S	0	0	0
			795	486	140	165	4			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	139	THR	-	expression tag	UNP P9WK61
E	140	HIS	-	expression tag	UNP P9WK61

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	141	HIS	-	expression tag	UNP P9WK61
E	142	HIS	-	expression tag	UNP P9WK61
E	143	HIS	-	expression tag	UNP P9WK61
E	144	HIS	-	expression tag	UNP P9WK61
E	145	HIS	-	expression tag	UNP P9WK61

- Molecule 6 is a protein called Lipoprotein LpqH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	113	Total	C	N	O	S	0	0	0
			772	471	132	165	4			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	139	THR	-	expression tag	UNP P9WK61
F	140	HIS	-	expression tag	UNP P9WK61
F	141	HIS	-	expression tag	UNP P9WK61
F	142	HIS	-	expression tag	UNP P9WK61
F	143	HIS	-	expression tag	UNP P9WK61
F	144	HIS	-	expression tag	UNP P9WK61
F	145	HIS	-	expression tag	UNP P9WK61

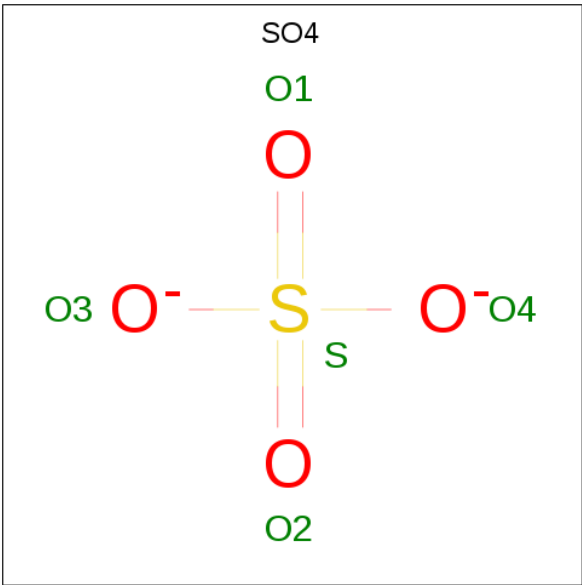
- Molecule 7 is a protein called Lipoprotein LpqH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	112	Total	C	N	O	S	0	0	0
			768	470	131	163	4			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	139	THR	-	expression tag	UNP P9WK61
G	140	HIS	-	expression tag	UNP P9WK61
G	141	HIS	-	expression tag	UNP P9WK61
G	142	HIS	-	expression tag	UNP P9WK61
G	143	HIS	-	expression tag	UNP P9WK61
G	144	HIS	-	expression tag	UNP P9WK61
G	145	HIS	-	expression tag	UNP P9WK61

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



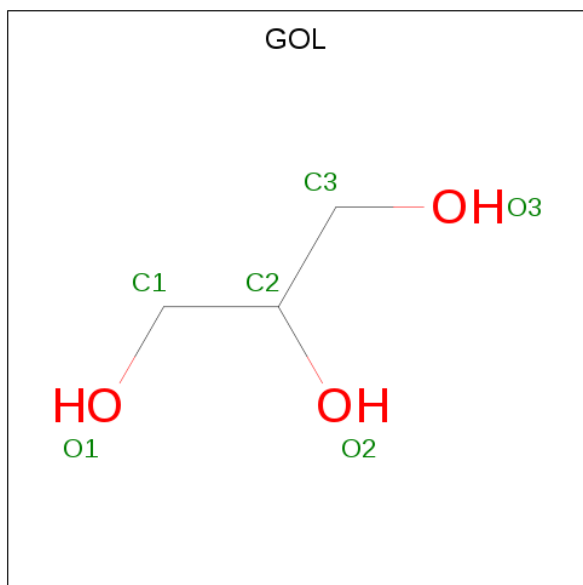
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	O	S	0	0
			5	4	1		
8	A	1	Total	O	S	0	0
			5	4	1		
8	B	1	Total	O	S	0	0
			5	4	1		
8	B	1	Total	O	S	0	0
			5	4	1		
8	B	1	Total	O	S	0	0
			5	4	1		
8	B	1	Total	O	S	0	0
			5	4	1		
8	C	1	Total	O	S	0	0
			5	4	1		
8	C	1	Total	O	S	0	0
			5	4	1		
8	D	1	Total	O	S	0	0
			5	4	1		
8	E	1	Total	O	S	0	0
			5	4	1		
8	E	1	Total	O	S	0	0
			5	4	1		
8	F	1	Total	O	S	0	0
			5	4	1		
8	F	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	G	1	Total	O	S	0	0
			5	4	1		
8	G	1	Total	O	S	0	0
			5	4	1		
8	G	1	Total	O	S	0	0
			5	4	1		

- Molecule 9 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			6	3	3		
9	A	1	Total	C	O	0	0
			6	3	3		
9	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 10 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	D	1	Total	Cl	0	0
			1	1		

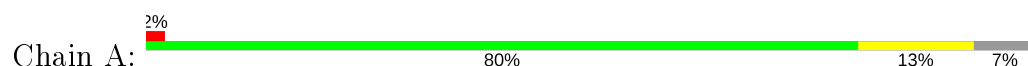
- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	2	Total 2	O 2	0	0
11	B	4	Total 4	O 4	0	0
11	C	2	Total 2	O 2	0	0
11	D	1	Total 1	O 1	0	0
11	E	3	Total 3	O 3	0	0
11	F	3	Total 3	O 3	0	0
11	G	3	Total 3	O 3	0	0

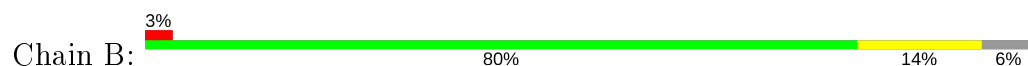
### 3 Residue-property plots [i](#)

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



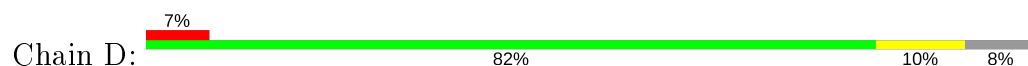
#### • Molecule 2: Lipoprotein LpqH



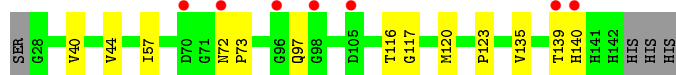
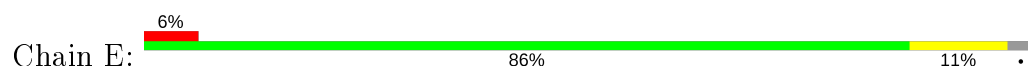
#### • Molecule 3: Lipoprotein LpqH



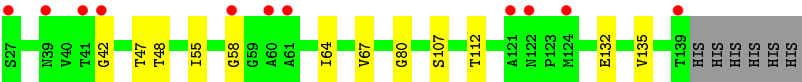
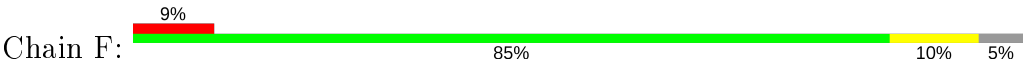
#### • Molecule 4: Lipoprotein LpqH



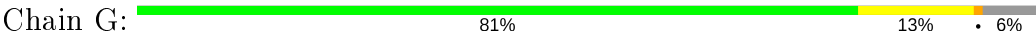
#### • Molecule 5: Lipoprotein LpqH



#### • Molecule 6: Lipoprotein LpqH



• Molecule 7: Lipoprotein LpqH



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.37Å 100.37Å 239.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.83 – 2.85 19.83 – 2.85	Depositor EDS
% Data completeness (in resolution range)	99.7 (19.83-2.85) 99.9 (19.83-2.85)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.54 (at 2.83Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.202 , 0.246 0.204 , 0.247	Depositor DCC
$R_{free}$ test set	2924 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	69.2	Xtriage
Anisotropy	0.675	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 55.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5497	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.44% 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 ⓘ

### 5.1 Standard geometry ⓘ

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

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	0/730	0.63	0/1000
2	B	0.43	0/751	0.62	0/1026
3	C	0.41	0/754	0.63	0/1029
4	D	0.36	0/748	0.56	0/1021
5	E	0.36	0/784	0.55	0/1072
6	F	0.36	0/758	0.60	0/1036
7	G	0.44	0/743	0.64	0/1017
All	All	0.40	0/5268	0.60	0/7201

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	766	0	758	20	0
2	B	765	1	754	14	0
3	C	757	0	747	19	0
4	D	751	0	742	10	0
5	E	795	0	774	11	0
6	F	772	0	762	13	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	G	768	0	759	13	0
8	A	10	0	0	0	0
8	B	25	0	0	1	0
8	C	10	0	0	0	0
8	D	5	0	0	0	0
8	E	10	0	0	0	0
8	F	10	0	0	1	0
8	G	15	0	0	1	0
9	A	12	0	16	5	0
9	B	6	0	8	0	0
10	D	1	0	0	0	0
11	A	2	0	0	0	0
11	B	4	0	0	0	0
11	C	2	0	0	0	0
11	D	1	0	0	0	0
11	E	3	0	0	0	0
11	F	3	0	0	0	0
11	G	3	0	0	0	0
All	All	5496	1	5320	78	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 7.

All (78) 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)
3:C:96:GLY:HA2	6:F:67:VAL:HG11	1.51	0.92
2:B:103:THR:HG22	6:F:47:THR:HG22	1.52	0.90
1:A:48:THR:HG21	7:G:75:GLU:HG2	1.52	0.90
5:E:97:GLN:HE21	5:E:116:THR:HG21	1.40	0.85
2:B:75:GLU:HG2	6:F:48:THR:HG21	1.61	0.82
4:D:129:MLY:HH22	4:D:129:MLY:HG2	1.61	0.81
7:G:42:GLY:HA3	7:G:58:GLY:HA2	1.63	0.80
6:F:42:GLY:HA3	6:F:58:GLY:HA2	1.64	0.80
3:C:112:THR:HG22	3:C:132:GLU:HG3	1.68	0.76
1:A:47:THR:HG23	3:C:126:PRO:HG3	1.69	0.73
1:A:47:THR:CG2	3:C:126:PRO:HG3	2.22	0.69
1:A:107:SER:HB2	9:A:203:GOL:H2	1.79	0.65
2:B:103:THR:CG2	6:F:47:THR:HG22	2.27	0.65
6:F:55:ILE:HG21	6:F:135:VAL:HG21	1.79	0.65
2:B:73:PRO:HG2	6:F:48:THR:OG1	1.97	0.64

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:52:ASN:OD1	3:C:95:THR:HB	1.97	0.63
5:E:139:THR:HG23	5:E:140:HIS:ND1	2.14	0.63
2:B:75:GLU:HG2	6:F:48:THR:CG2	2.29	0.62
1:A:48:THR:HG21	7:G:75:GLU:CG	2.29	0.61
3:C:96:GLY:CA	6:F:67:VAL:HG11	2.27	0.60
2:B:112:THR:HG22	2:B:132:GLU:HG3	1.84	0.60
1:A:107:SER:H	9:A:203:GOL:C3	2.14	0.60
2:B:68:LEU:HD22	2:B:76:VAL:HG22	1.83	0.60
3:C:124:MET:SD	7:G:103:THR:HG23	2.42	0.59
1:A:60:ALA:HB3	4:D:60:ALA:HB2	1.86	0.58
1:A:116:THR:HG22	1:A:117:GLY:N	2.19	0.58
1:A:107:SER:H	9:A:203:GOL:H32	1.70	0.57
7:G:50:ALA:N	8:G:202:SO4:O2	2.37	0.56
5:E:97:GLN:NE2	5:E:116:THR:HG21	2.16	0.56
6:F:64:ILE:HA	6:F:80:GLY:O	2.06	0.56
2:B:47:THR:HG22	2:B:54:ASN:HB2	1.87	0.55
4:D:47:THR:HG22	5:E:123:PRO:HB2	1.89	0.55
1:A:63:GLY:HA3	9:A:204:GOL:H11	1.90	0.53
1:A:45:VAL:HG13	3:C:123:PRO:HB3	1.89	0.53
3:C:47:THR:HG22	3:C:54:ASN:HB2	1.90	0.53
3:C:96:GLY:HA2	6:F:67:VAL:CG1	2.34	0.53
5:E:116:THR:HG22	5:E:117:GLY:N	2.26	0.51
4:D:51:GLY:HA2	4:D:70:ASP:OD2	2.10	0.51
3:C:32:VAL:O	3:C:131:PHE:HA	2.12	0.49
6:F:112:THR:HG22	6:F:132:GLU:HG3	1.95	0.49
1:A:60:ALA:O	9:A:204:GOL:H12	2.12	0.49
7:G:62:THR:CG2	7:G:64:ILE:HD12	2.44	0.48
1:A:102:ALA:HA	1:A:110:LYS:O	2.14	0.48
3:C:124:MET:HA	7:G:103:THR:CG2	2.45	0.47
6:F:107:SER:N	8:F:201:SO4:O4	2.40	0.47
5:E:44:VAL:HG11	5:E:135:VAL:HB	1.97	0.46
1:A:60:ALA:CB	4:D:60:ALA:HB2	2.45	0.46
2:B:116:THR:HA	2:B:127:VAL:O	2.16	0.46
4:D:112:THR:HG22	4:D:132:GLU:HG3	1.97	0.46
4:D:62:THR:CG2	4:D:64:ILE:HD12	2.46	0.46
2:B:72:ASN:HA	2:B:73:PRO:HA	1.77	0.45
2:B:61:ALA:HB3	8:B:201:SO4:O1	2.15	0.45
7:G:102:ALA:HA	7:G:110:LYS:O	2.16	0.45
1:A:112:THR:HG22	1:A:132:GLU:HG3	1.99	0.45
3:C:57:ILE:HD12	3:C:64:ILE:HG21	1.99	0.45
1:A:77:MLY:HH12	1:A:77:MLY:HD3	1.42	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:G:107:SER:HB2	7:G:137:CYS:O	2.17	0.44
1:A:45:VAL:CG1	3:C:123:PRO:HB3	2.47	0.44
7:G:42:GLY:HA3	7:G:58:GLY:CA	2.41	0.44
5:E:72:ASN:HA	5:E:73:PRO:HA	1.82	0.43
5:E:97:GLN:HE21	5:E:116:THR:CG2	2.20	0.43
4:D:72:ASN:HA	4:D:73:PRO:HA	1.83	0.43
7:G:68:LEU:HD12	7:G:109:TYR:CE2	2.54	0.43
4:D:129:MLY:HG2	4:D:129:MLY:CH2	2.40	0.43
3:C:58:GLY:HA3	3:C:62:THR:HB	2.01	0.42
2:B:31:VAL:HG11	2:B:64:ILE:HD13	2.02	0.42
4:D:128:ASN:O	4:D:129:MLY:HD2	2.20	0.42
3:C:57:ILE:HD12	3:C:64:ILE:CG2	2.49	0.41
7:G:53:VAL:HG23	7:G:70:ASP:HA	2.02	0.41
1:A:77:MLY:HH21	2:B:99:ASN:OD1	2.20	0.41
3:C:47:THR:CG2	3:C:54:ASN:HB2	2.50	0.41
3:C:85:ASN:HB3	5:E:120:MET:HB3	2.02	0.41
5:E:116:THR:HG22	5:E:117:GLY:H	1.85	0.41
5:E:40:VAL:HG11	5:E:57:ILE:CG2	2.50	0.41
7:G:74:PRO:HG2	7:G:102:ALA:O	2.21	0.41
1:A:32:VAL:O	1:A:131:PHE:HA	2.21	0.40
2:B:57:ILE:HD12	2:B:64:ILE:HG21	2.03	0.40
3:C:91:TYR:CE2	3:C:100:ALA:HB3	2.57	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	105/119 (88%)	103 (98%)	2 (2%)	0	100	100
2	B	108/119 (91%)	106 (98%)	2 (2%)	0	100	100
3	C	108/119 (91%)	105 (97%)	3 (3%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	D	107/119 (90%)	105 (98%)	2 (2%)	0	100	100
5	E	111/119 (93%)	111 (100%)	0	0	100	100
6	F	109/119 (92%)	105 (96%)	4 (4%)	0	100	100
7	G	107/119 (90%)	106 (99%)	1 (1%)	0	100	100
All	All	755/833 (91%)	741 (98%)	14 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	79/86 (92%)	79 (100%)	0	100	100
2	B	81/88 (92%)	81 (100%)	0	100	100
3	C	81/89 (91%)	81 (100%)	0	100	100
4	D	80/89 (90%)	80 (100%)	0	100	100
5	E	83/88 (94%)	83 (100%)	0	100	100
6	F	82/88 (93%)	82 (100%)	0	100	100
7	G	80/87 (92%)	79 (99%)	1 (1%)	69	88
All	All	566/615 (92%)	565 (100%)	1 (0%)	93	98

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	G	137	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
4	D	54	ASN
5	E	97	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

15 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
1	MLY	A	30	1	9,10,11	0.71	0	6,11,13	0.59	0
7	MLY	G	77	7	9,10,11	0.56	0	6,11,13	0.95	0
2	MLY	B	36	2	9,10,11	1.02	0	6,11,13	1.06	0
3	MLY	C	77	3	9,10,11	0.68	0	6,11,13	0.67	0
1	MLY	A	129	1	9,10,11	0.48	0	6,11,13	1.25	0
1	MLY	A	77	1	9,10,11	0.69	0	6,11,13	1.14	0
5	MLY	E	30	5	9,10,11	0.65	0	6,11,13	1.00	0
1	MLY	A	104	1	9,10,11	0.62	0	6,11,13	1.01	0
7	MLY	G	129	7	9,10,11	0.60	0	6,11,13	1.13	0
7	MLY	G	36	7	9,10,11	0.74	0	6,11,13	0.50	0
6	MLY	F	110	6	9,10,11	0.61	0	6,11,13	0.56	0
4	MLY	D	129	4	9,10,11	0.59	0	6,11,13	1.12	0
2	MLY	B	129	2	9,10,11	0.61	0	6,11,13	0.91	0
5	MLY	E	36	5	9,10,11	0.84	0	6,11,13	0.49	0
6	MLY	F	77	6	9,10,11	0.68	0	6,11,13	0.96	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	MLY	A	30	1	-	1/8/9/11	-
7	MLY	G	77	7	-	3/8/9/11	-
2	MLY	B	36	2	-	0/8/9/11	-
3	MLY	C	77	3	-	3/8/9/11	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	MLY	A	129	1	-	3/8/9/11	-
1	MLY	A	77	1	-	4/8/9/11	-
5	MLY	E	30	5	-	0/8/9/11	-
1	MLY	A	104	1	-	3/8/9/11	-
7	MLY	G	129	7	-	0/8/9/11	-
7	MLY	G	36	7	-	0/8/9/11	-
6	MLY	F	110	6	-	1/8/9/11	-
4	MLY	D	129	4	-	1/8/9/11	-
2	MLY	B	129	2	-	2/8/9/11	-
5	MLY	E	36	5	-	2/8/9/11	-
6	MLY	F	77	6	-	1/8/9/11	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (24) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	G	77	MLY	O-C-CA-CB
3	C	77	MLY	O-C-CA-CB
7	G	77	MLY	CD-CE-NZ-CH1
7	G	77	MLY	CD-CE-NZ-CH2
2	B	129	MLY	CD-CE-NZ-CH1
2	B	129	MLY	CD-CE-NZ-CH2
5	E	36	MLY	CD-CE-NZ-CH1
5	E	36	MLY	CD-CE-NZ-CH2
1	A	30	MLY	CD-CE-NZ-CH1
3	C	77	MLY	CD-CE-NZ-CH1
1	A	77	MLY	CD-CE-NZ-CH1
1	A	104	MLY	CD-CE-NZ-CH2
6	F	110	MLY	CD-CE-NZ-CH1
1	A	77	MLY	CG-CD-CE-NZ
3	C	77	MLY	CD-CE-NZ-CH2
1	A	129	MLY	CA-CB-CG-CD
1	A	129	MLY	CD-CE-NZ-CH2
1	A	77	MLY	CA-CB-CG-CD
6	F	77	MLY	CD-CE-NZ-CH1
4	D	129	MLY	CA-CB-CG-CD
1	A	77	MLY	CD-CE-NZ-CH2
1	A	104	MLY	CD-CE-NZ-CH1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
1	A	104	MLY	CE-CD-CG-CB
1	A	129	MLY	CE-CD-CG-CB

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	77	MLY	2	0
4	D	129	MLY	3	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 21 ligands modelled in this entry, 1 is monoatomic - leaving 20 for Mogul analysis.

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	$\# Z  > 2$
8	SO4	C	201	-	4,4,4	0.15	0	6,6,6	0.17	0
8	SO4	E	201	-	4,4,4	0.13	0	6,6,6	0.25	0
8	SO4	A	202	-	4,4,4	0.14	0	6,6,6	0.11	0
8	SO4	G	201	-	4,4,4	0.15	0	6,6,6	0.25	0
8	SO4	D	201	-	4,4,4	0.15	0	6,6,6	0.10	0
8	SO4	B	202	-	4,4,4	0.15	0	6,6,6	0.11	0
8	SO4	F	202	-	4,4,4	0.15	0	6,6,6	0.09	0
8	SO4	B	201	-	4,4,4	0.12	0	6,6,6	0.31	0
8	SO4	C	202	-	4,4,4	0.15	0	6,6,6	0.19	0
8	SO4	B	203	-	4,4,4	0.14	0	6,6,6	0.15	0
8	SO4	G	203	-	4,4,4	0.15	0	6,6,6	0.16	0
9	GOL	A	203	-	5,5,5	0.36	0	5,5,5	0.40	0
9	GOL	B	206	-	5,5,5	0.37	0	5,5,5	0.28	0
8	SO4	E	202	-	4,4,4	0.14	0	6,6,6	0.09	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
9	GOL	A	204	-	5,5,5	0.45	0	5,5,5	0.59	0
8	SO4	F	201	-	4,4,4	0.15	0	6,6,6	0.12	0
8	SO4	B	204	-	4,4,4	0.13	0	6,6,6	0.13	0
8	SO4	G	202	-	4,4,4	0.14	0	6,6,6	0.15	0
8	SO4	A	201	-	4,4,4	0.15	0	6,6,6	0.20	0
8	SO4	B	205	-	4,4,4	0.15	0	6,6,6	0.18	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	GOL	A	204	-	-	2/4/4/4	-
9	GOL	A	203	-	-	2/4/4/4	-
9	GOL	B	206	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	A	204	GOL	C1-C2-C3-O3
9	A	203	GOL	O1-C1-C2-C3
9	A	204	GOL	O2-C2-C3-O3
9	A	203	GOL	O1-C1-C2-O2

There are no ring outliers.

5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	B	201	SO4	1	0
9	A	203	GOL	3	0
9	A	204	GOL	2	0
8	F	201	SO4	1	0
8	G	202	SO4	1	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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(Å <sup>2</sup> )	Q<0.9
1	A	107/119 (89%)	-0.44	2 (1%) 66 64	54, 65, 90, 110	0
2	B	110/119 (92%)	-0.35	3 (2%) 54 50	58, 67, 98, 128	0
3	C	110/119 (92%)	-0.29	3 (2%) 54 50	61, 78, 113, 136	0
4	D	109/119 (91%)	0.24	8 (7%) 15 11	72, 109, 144, 154	0
5	E	113/119 (94%)	0.13	7 (6%) 20 16	69, 95, 138, 157	0
6	F	111/119 (93%)	0.18	11 (9%) 7 5	64, 88, 143, 188	0
7	G	109/119 (91%)	-0.32	0 100 100	58, 71, 100, 116	0
All	All	769/833 (92%)	-0.12	34 (4%) 34 29	54, 80, 131, 188	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	106	GLY	6.2
5	E	139	THR	5.6
3	C	138	SER	4.8
2	B	27	SER	4.4
6	F	122	ASN	4.3
4	D	107	SER	4.2
4	D	60	ALA	3.9
6	F	41	THR	3.7
6	F	139	THR	3.4
4	D	61	ALA	3.2
5	E	140	HIS	3.2
1	A	124	MET	3.1
6	F	39	ASN	3.0
2	B	124	MET	3.0
6	F	124	MET	3.0
6	F	121	ALA	2.9
5	E	72	ASN	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	D	96	GLY	2.8
3	C	60	ALA	2.8
3	C	105	ASP	2.7
5	E	105	ASP	2.7
4	D	66	ALA	2.5
2	B	125	SER	2.5
4	D	50	ALA	2.4
4	D	128	ASN	2.4
6	F	42	GLY	2.4
6	F	27	SER	2.4
6	F	60	ALA	2.4
1	A	139	THR	2.2
6	F	58	GLY	2.1
6	F	61	ALA	2.1
5	E	70	ASP	2.0
5	E	96	GLY	2.0
5	E	98	GLY	2.0

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

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
7	MLY	G	36	11/12	0.84	0.25	78,101,115,115	0
5	MLY	E	36	11/12	0.84	0.28	76,94,123,149	0
6	MLY	F	110	11/12	0.85	0.32	77,83,102,102	0
2	MLY	B	36	11/12	0.89	0.20	62,69,93,95	0
7	MLY	G	77	11/12	0.92	0.27	58,63,73,76	0
6	MLY	F	77	11/12	0.93	0.21	57,61,76,78	0
5	MLY	E	30	11/12	0.94	0.25	85,94,112,113	0
1	MLY	A	30	11/12	0.95	0.14	55,61,91,93	0
1	MLY	A	104	11/12	0.95	0.14	70,73,88,100	0
4	MLY	D	129	11/12	0.97	0.17	84,90,99,101	0
2	MLY	B	129	11/12	0.97	0.16	58,66,67,67	0
1	MLY	A	77	11/12	0.97	0.17	53,64,75,77	0
7	MLY	G	129	11/12	0.97	0.18	62,64,74,84	0
3	MLY	C	77	11/12	0.98	0.13	59,62,79,82	0
1	MLY	A	129	11/12	0.98	0.15	55,62,67,71	0

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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
10	CL	D	202	1/1	0.83	0.19	119,119,119,119	0
8	SO4	B	204	5/5	0.83	0.36	146,148,150,150	0
8	SO4	E	202	5/5	0.84	0.56	169,171,173,174	0
8	SO4	F	201	5/5	0.89	0.30	150,154,156,158	0
8	SO4	F	202	5/5	0.89	0.32	90,96,97,98	5
9	GOL	B	206	6/6	0.90	0.24	70,72,77,81	0
9	GOL	A	203	6/6	0.91	0.23	77,89,91,93	0
9	GOL	A	204	6/6	0.91	0.37	74,83,85,86	0
8	SO4	D	201	5/5	0.91	0.13	146,149,152,157	0
8	SO4	B	202	5/5	0.92	0.14	115,115,120,126	0
8	SO4	B	203	5/5	0.92	0.15	121,122,124,126	0
8	SO4	C	201	5/5	0.93	0.10	99,99,104,111	0
8	SO4	G	203	5/5	0.93	0.20	115,116,118,119	0
8	SO4	B	205	5/5	0.93	0.15	124,124,127,129	0
8	SO4	G	202	5/5	0.94	0.14	124,125,128,132	0
8	SO4	G	201	5/5	0.94	0.13	119,119,122,123	0
8	SO4	A	201	5/5	0.95	0.12	66,74,77,84	0
8	SO4	C	202	5/5	0.95	0.18	104,110,112,122	0
8	SO4	E	201	5/5	0.97	0.09	79,81,85,90	0
8	SO4	B	201	5/5	0.98	0.11	56,66,75,77	0
8	SO4	A	202	5/5	0.98	0.12	71,73,76,94	0

### 6.5 Other polymers ⓘ

There are no such residues in this entry.