



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 12:53 pm BST

PDB ID : 5Y9E  
Title : Crystal structure of HPV58 pentamer  
Authors : Li, S.W.; Li, Z.H.  
Deposited on : 2017-08-24  
Resolution : 2.04 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

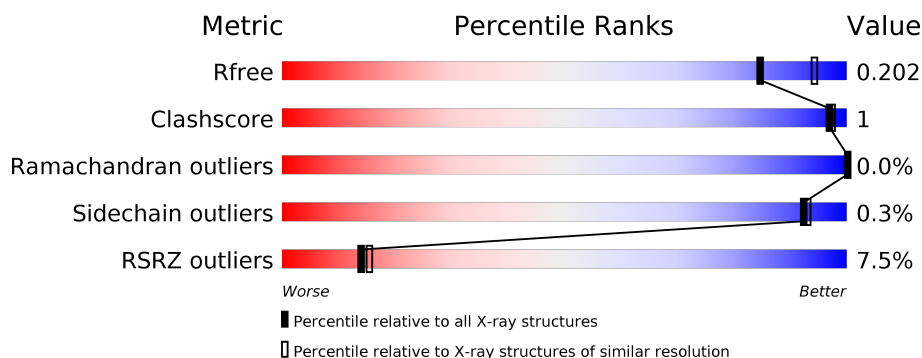
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.04 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	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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	490	<div> <div>5%</div> <div> <div></div> <div>82%</div> <div>•</div> <div>15%</div> </div> </div>
1	B	490	<div> <div>4%</div> <div> <div></div> <div>83%</div> <div>•</div> <div>16%</div> </div> </div>
1	C	490	<div> <div>7%</div> <div> <div></div> <div>83%</div> <div>•</div> <div>15%</div> </div> </div>
1	D	490	<div> <div>7%</div> <div> <div></div> <div>84%</div> <div>•</div> <div>15%</div> </div> </div>
1	E	490	<div> <div>9%</div> <div> <div></div> <div>83%</div> <div>•</div> <div>14%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GOL	B	501	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 17883 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 Major capsid protein L1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	416	Total	C	N	O	S	0	1	0
			3323	2120	551	632	20			
1	B	414	Total	C	N	O	S	0	1	0
			3312	2114	549	629	20			
1	C	415	Total	C	N	O	S	0	1	0
			3318	2117	550	631	20			
1	D	418	Total	C	N	O	S	0	1	0
			3332	2125	554	633	20			
1	E	420	Total	C	N	O	S	0	1	0
			3348	2133	558	637	20			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	MET	-	initiating methionine	UNP P26535
A	176	SER	CYS	engineered mutation	UNP P26535
B	9	MET	-	initiating methionine	UNP P26535
B	176	SER	CYS	engineered mutation	UNP P26535
C	9	MET	-	initiating methionine	UNP P26535
C	176	SER	CYS	engineered mutation	UNP P26535
D	9	MET	-	initiating methionine	UNP P26535
D	176	SER	CYS	engineered mutation	UNP P26535
E	9	MET	-	initiating methionine	UNP P26535
E	176	SER	CYS	engineered mutation	UNP P26535

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	D	1	Total	C	O	0	0
			6	3	3		
2	D	1	Total	C	O	0	0
			6	3	3		
2	E	1	Total	C	O	0	0
			6	3	3		
2	E	1	Total	C	O	0	0
			6	3	3		
2	E	1	Total	C	O	0	0
			6	3	3		
2	E	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	6	Total	Mg	0	0
			6	6		
3	A	7	Total	Mg	0	0
			7	7		
3	D	6	Total	Mg	0	0
			6	6		
3	C	6	Total	Mg	0	0
			6	6		
3	E	4	Total	Mg	0	0
			4	4		

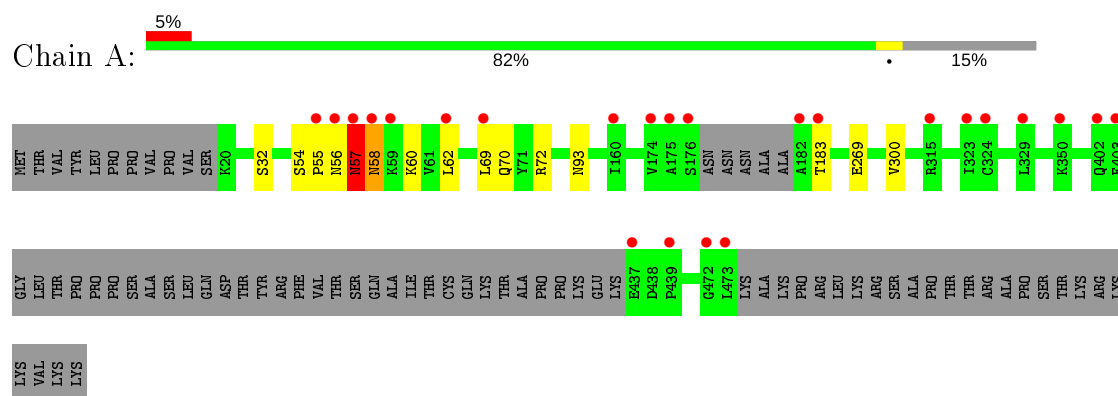
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	250	Total	O	0	0
			250	250		
4	B	238	Total	O	0	0
			238	238		
4	C	231	Total	O	0	0
			231	231		
4	D	205	Total	O	0	0
			205	205		
4	E	177	Total	O	0	0
			177	177		

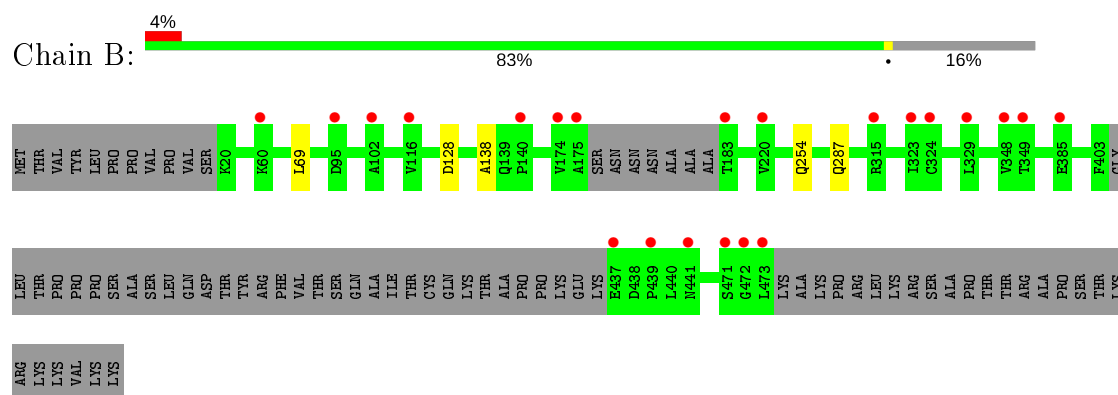
### 3 Residue-property plots [i](#)

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.

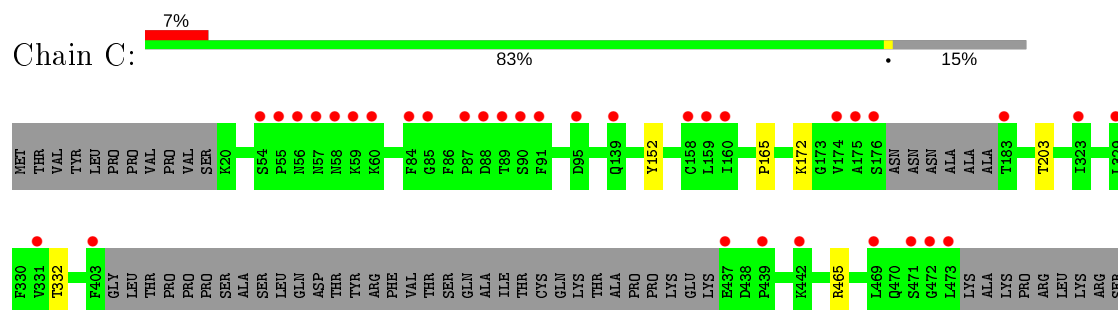
#### • Molecule 1: Major capsid protein L1



#### • Molecule 1: Major capsid protein L1

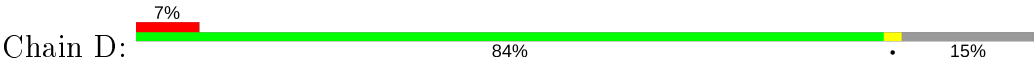


#### • Molecule 1: Major capsid protein L1



ALA  
PRO  
THR  
THR  
THR  
ARG  
ALA  
PRO  
SER  
THR  
LYS  
ARG  
LYS  
LYS  
VAL  
LYS

● Molecule 1: Major capsid protein L1

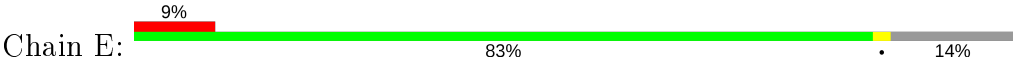


MET  
THR  
VAL  
TYR  
LEU  
PRO  
PRO  
VAL  
SER  
GLN  
SER  
K20  
S51  
S54  
P55  
N56  
N57  
N58  
K59  
K60  
Q70  
Y71  
R72  
P87  
D88  
A102  
G111  
Q112  
V174  
A175  
S176  
ASN  
ASN  
N179  
A180  
A181  
A182  
T183  
R315  
I323  
C324  
L329  
T340  
E385  
I386  
N387  
F403

GLY  
LEU  
THR  
PRO  
PRO  
PRO  
SER  
ALA  
SER  
LEU  
GLN  
ASP  
THR  
TYR  
ARG  
PHE  
VAL  
THR  
SER  
GLN  
ALA  
ILE  
THR  
CYS  
GLN  
LYS  
THR  
ALA  
PRO  
PRO  
GLY  
LYS  
D438  
P439  
L440  
N441  
K442  
S471  
Q472  
L473  
LYS  
ALA  
LYS  
PRO  
ARG  
LEU  
LYS  
SER  
SER  
ALA  
PRO  
THR  
THR  
ARG  
ALA  
PRO  
SER

THR  
LYS  
ARG  
LYS  
VAL  
LYS  
LYS

● Molecule 1: Major capsid protein L1



MET  
THR  
VAL  
TYR  
LEU  
PRO  
PRO  
VAL  
SER  
VAL  
SER  
K20  
Y21  
S54  
P55  
N56  
N57  
N58  
K59  
K60  
V61  
L62  
Q70  
Y71  
R72  
G85  
F86  
P87  
D88  
T89  
S90  
T96  
G103  
Q139  
G161  
Y174  
A175  
S176  
A180  
A181  
A182  
T183  
D184  
E269  
K309  
R315  
I323

C324  
L329  
L382  
T383  
A384  
E385  
F403  
GLY  
LEU  
THR  
PRO  
PRO  
PRO  
SER  
ALA  
SER  
LEU  
GLN  
ASP  
TYR  
ARG  
PHE  
VAL  
THR  
THR  
GLN  
ALA  
ILE  
THR  
CYS  
GLN  
LYS  
THR  
ALA  
PRO  
PRO  
LYS  
LYS  
GLU  
GLU  
D438  
P439  
L440  
N441  
K442  
Y443  
E452  
S471  
Q472  
L473  
LYS  
ALA  
LYS  
PRO

ARG  
LEU  
LYS  
ARG  
SER  
ALA  
PRO  
THR  
THR  
ARG  
ALA  
PRO  
SER  
THR  
LYS  
ARG  
LYS  
LYS  
VAL  
LYS  
LYS



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	187.21Å 101.80Å 136.19Å 90.00° 95.68° 90.00°	Depositor
Resolution (Å)	29.78 – 2.04 29.78 – 2.04	Depositor EDS
% Data completeness (in resolution range)	98.9 (29.78-2.04) 98.9 (29.78-2.04)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.43 (at 2.04Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575)	Depositor
R, $R_{free}$	0.168 , 0.198 0.173 , 0.202	Depositor DCC
$R_{free}$ test set	12229 reflections (7.67%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.1	Xtriage
Anisotropy	0.906	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 46.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	17883	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.99% 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](#)

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

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	0.26	0/3412	0.46	0/4627
1	B	0.26	0/3401	0.46	0/4612
1	C	0.25	0/3407	0.45	0/4620
1	D	0.25	0/3421	0.45	0/4640
1	E	0.25	0/3438	0.45	0/4665
All	All	0.25	0/17079	0.45	0/23164

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3323	0	3204	9	1
1	B	3312	0	3194	3	0
1	C	3318	0	3199	5	0
1	D	3332	0	3214	4	1
1	E	3348	0	3227	7	0
2	A	24	0	32	0	0
2	B	30	0	40	0	0
2	C	18	0	24	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	24	0	32	1	0
2	E	24	0	32	0	0
3	A	7	0	0	0	0
3	B	6	0	0	0	0
3	C	6	0	0	0	0
3	D	6	0	0	0	0
3	E	4	0	0	0	0
4	A	250	0	0	1	0
4	B	238	0	0	1	0
4	C	231	0	0	2	0
4	D	205	0	0	2	0
4	E	177	0	0	3	0
All	All	17883	0	16198	28	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All (28) 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:E:20:LYS:N	4:E:601:HOH:O	2.15	0.79
1:A:70:GLN:OE1	1:A:72:ARG:NH2	2.17	0.77
1:E:70:GLN:OE1	1:E:72:ARG:NH2	2.17	0.77
1:D:70:GLN:OE1	1:D:72:ARG:NH2	2.26	0.68
1:A:55:PRO:HG2	1:A:62:LEU:HD22	1.80	0.63
1:E:309:LYS:NZ	4:E:605:HOH:O	2.33	0.62
1:D:51:SER:O	4:D:601:HOH:O	2.16	0.61
1:A:32:SER:O	4:A:601:HOH:O	2.17	0.58
2:D:504:GOL:O2	4:D:602:HOH:O	2.18	0.56
1:E:184:ASP:OD2	4:E:602:HOH:O	2.18	0.55
1:A:56:ASN:O	1:A:57:ASN:HB2	2.08	0.53
1:C:172:LYS:NZ	4:C:611:HOH:O	2.45	0.50
1:A:54:SER:HB2	1:A:58:ASN:O	2.16	0.45
1:A:58:ASN:O	1:A:60:LYS:N	2.48	0.44
1:A:183:THR:HG21	1:A:269:GLU:OE1	2.19	0.43
1:D:54:SER:HB3	1:D:57:ASN:HB3	2.01	0.43
1:E:452:GLU:OE1	1:E:452:GLU:N	2.51	0.43
1:E:56:ASN:O	1:E:58:ASN:N	2.52	0.43
1:C:465:ARG:NH2	4:C:609:HOH:O	2.42	0.43
1:B:138:ALA:HB2	1:B:287:GLN:HG2	2.00	0.42
1:B:128:ASP:OD1	4:B:601:HOH:O	2.21	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:152:TYR:CD1	1:C:203:THR:HB	2.54	0.42
1:A:300:VAL:O	1:B:254:GLN:HA	2.20	0.41
1:C:152:TYR:CG	1:C:203:THR:HB	2.56	0.41
1:A:56:ASN:O	1:A:57:ASN:CB	2.69	0.40
1:C:165:PRO:HG3	1:C:332:THR:OG1	2.21	0.40
1:D:111:GLY:O	1:D:112:GLN:HB2	2.22	0.40
1:E:183:THR:HG21	1:E:269:GLU:OE1	2.22	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:93:ASN:ND2	1:D:55:PRO:O[3_445]	2.01	0.19

## 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	411/490 (84%)	397 (97%)	13 (3%)	1 (0%)	47	39
1	B	409/490 (84%)	397 (97%)	12 (3%)	0	100	100
1	C	410/490 (84%)	399 (97%)	11 (3%)	0	100	100
1	D	413/490 (84%)	400 (97%)	13 (3%)	0	100	100
1	E	417/490 (85%)	401 (96%)	16 (4%)	0	100	100
All	All	2060/2450 (84%)	1994 (97%)	65 (3%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	57	ASN

### 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	371/435 (85%)	368 (99%)	3 (1%)	81	82
1	B	370/435 (85%)	369 (100%)	1 (0%)	92	93
1	C	371/435 (85%)	371 (100%)	0	100	100
1	D	371/435 (85%)	371 (100%)	0	100	100
1	E	373/435 (86%)	372 (100%)	1 (0%)	92	93
All	All	1856/2175 (85%)	1851 (100%)	5 (0%)	92	93

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

Mol	Chain	Res	Type
1	A	57	ASN
1	A	58	ASN
1	A	69	LEU
1	B	69	LEU
1	E	59	LYS

Some 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 49 ligands modelled in this entry, 29 are 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
2	GOL	D	503	-	5,5,5	0.34	0	5,5,5	0.22	0
2	GOL	A	503	-	5,5,5	0.35	0	5,5,5	0.26	0
2	GOL	C	503	-	5,5,5	0.39	0	5,5,5	0.09	0
2	GOL	B	504	-	5,5,5	0.38	0	5,5,5	0.21	0
2	GOL	B	503	-	5,5,5	0.37	0	5,5,5	0.26	0
2	GOL	E	502	-	5,5,5	0.36	0	5,5,5	0.21	0
2	GOL	D	504	-	5,5,5	0.35	0	5,5,5	0.16	0
2	GOL	A	502	-	5,5,5	0.38	0	5,5,5	0.17	0
2	GOL	C	502	-	5,5,5	0.38	0	5,5,5	0.18	0
2	GOL	E	501	-	5,5,5	0.38	0	5,5,5	0.20	0
2	GOL	B	505	-	5,5,5	0.37	0	5,5,5	0.24	0
2	GOL	C	501	-	5,5,5	0.37	0	5,5,5	0.18	0
2	GOL	D	501	3	5,5,5	0.39	0	5,5,5	0.16	0
2	GOL	A	501	-	5,5,5	0.39	0	5,5,5	0.14	0
2	GOL	B	501	-	5,5,5	0.37	0	5,5,5	0.24	0
2	GOL	D	502	-	5,5,5	0.39	0	5,5,5	0.23	0
2	GOL	A	504	-	5,5,5	0.38	0	5,5,5	0.20	0
2	GOL	B	502	-	5,5,5	0.37	0	5,5,5	0.20	0
2	GOL	E	503	-	5,5,5	0.38	0	5,5,5	0.16	0
2	GOL	E	504	-	5,5,5	0.38	0	5,5,5	0.13	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
2	GOL	D	503	-	-	2/4/4/4	-
2	GOL	A	503	-	-	0/4/4/4	-
2	GOL	C	503	-	-	4/4/4/4	-
2	GOL	B	504	-	-	2/4/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	B	503	-	-	2/4/4/4	-
2	GOL	E	502	-	-	2/4/4/4	-
2	GOL	D	504	-	-	2/4/4/4	-
2	GOL	A	502	-	-	2/4/4/4	-
2	GOL	C	502	-	-	2/4/4/4	-
2	GOL	E	501	-	-	2/4/4/4	-
2	GOL	B	505	-	-	2/4/4/4	-
2	GOL	C	501	-	-	1/4/4/4	-
2	GOL	D	501	3	-	4/4/4/4	-
2	GOL	A	501	-	-	2/4/4/4	-
2	GOL	B	501	-	-	2/4/4/4	-
2	GOL	D	502	-	-	2/4/4/4	-
2	GOL	A	504	-	-	3/4/4/4	-
2	GOL	B	502	-	-	2/4/4/4	-
2	GOL	E	503	-	-	2/4/4/4	-
2	GOL	E	504	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (42) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	502	GOL	O1-C1-C2-C3
2	B	502	GOL	O1-C1-C2-C3
2	E	503	GOL	O1-C1-C2-C3
2	E	504	GOL	O1-C1-C2-C3
2	D	501	GOL	O1-C1-C2-C3
2	B	501	GOL	O1-C1-C2-C3
2	B	504	GOL	O1-C1-C2-C3
2	E	502	GOL	O1-C1-C2-O2
2	E	502	GOL	O1-C1-C2-C3
2	C	503	GOL	O1-C1-C2-C3
2	B	503	GOL	O1-C1-C2-O2
2	B	503	GOL	O1-C1-C2-C3
2	A	504	GOL	O1-C1-C2-C3
2	B	505	GOL	O1-C1-C2-O2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	B	505	GOL	O1-C1-C2-C3
2	A	502	GOL	O1-C1-C2-C3
2	E	501	GOL	O1-C1-C2-C3
2	D	501	GOL	C1-C2-C3-O3
2	A	501	GOL	C1-C2-C3-O3
2	C	503	GOL	C1-C2-C3-O3
2	E	503	GOL	O1-C1-C2-O2
2	E	501	GOL	O1-C1-C2-O2
2	B	501	GOL	O1-C1-C2-O2
2	B	504	GOL	O1-C1-C2-O2
2	C	502	GOL	O1-C1-C2-O2
2	B	502	GOL	O1-C1-C2-O2
2	E	504	GOL	O1-C1-C2-O2
2	D	501	GOL	O1-C1-C2-O2
2	C	503	GOL	O1-C1-C2-O2
2	C	503	GOL	O2-C2-C3-O3
2	C	501	GOL	O1-C1-C2-C3
2	A	502	GOL	O1-C1-C2-O2
2	D	504	GOL	O1-C1-C2-O2
2	A	504	GOL	O1-C1-C2-O2
2	D	503	GOL	O1-C1-C2-O2
2	D	503	GOL	O1-C1-C2-C3
2	A	501	GOL	O2-C2-C3-O3
2	D	502	GOL	O1-C1-C2-C3
2	D	501	GOL	O2-C2-C3-O3
2	D	502	GOL	O1-C1-C2-O2
2	D	504	GOL	O1-C1-C2-C3
2	A	504	GOL	C1-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	504	GOL	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	416/490 (84%)	0.13	24 (5%)	23	25	25, 39, 76, 125	0
1	B	414/490 (84%)	0.06	22 (5%)	26	28	27, 41, 69, 101	0
1	C	415/490 (84%)	0.29	34 (8%)	11	12	30, 43, 84, 118	0
1	D	418/490 (85%)	0.29	33 (7%)	12	13	30, 46, 81, 125	0
1	E	420/490 (85%)	0.33	44 (10%)	6	6	28, 46, 86, 120	0
All	All	2083/2450 (85%)	0.22	157 (7%)	14	15	25, 43, 81, 125	0

All (157) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	181	ALA	7.7
1	C	175	ALA	7.5
1	C	473	LEU	7.3
1	D	175	ALA	7.3
1	D	473	LEU	6.9
1	A	175	ALA	6.9
1	A	176	SER	6.2
1	D	182	ALA	6.2
1	D	472	GLY	6.1
1	E	180	ALA	6.0
1	D	180	ALA	5.9
1	E	56	ASN	5.8
1	E	87	PRO	5.8
1	D	174	VAL	5.6
1	A	56	ASN	5.3
1	C	176	SER	5.3
1	A	182	ALA	5.2
1	E	403	PHE	5.1
1	A	403	PHE	5.1
1	C	59	LYS	5.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	473	LEU	4.9
1	E	88	ASP	4.9
1	A	473	LEU	4.8
1	E	58	ASN	4.6
1	D	179	ASN	4.6
1	C	174	VAL	4.6
1	A	59	LYS	4.5
1	D	59	LYS	4.3
1	C	55	PRO	4.3
1	C	472	GLY	4.3
1	B	473	LEU	4.2
1	E	439	PRO	4.1
1	E	59	LYS	4.1
1	A	183	THR	4.1
1	E	86	PHE	4.0
1	A	57	ASN	4.0
1	C	90	SER	4.0
1	E	384	ALA	3.9
1	C	87	PRO	3.9
1	E	57	ASN	3.9
1	A	62	LEU	3.9
1	A	58	ASN	3.9
1	D	55	PRO	3.8
1	C	471	SER	3.8
1	A	437	GLU	3.8
1	E	55	PRO	3.8
1	D	441	ASN	3.7
1	A	472	GLY	3.7
1	E	85	GLY	3.7
1	B	437	GLU	3.7
1	A	174	VAL	3.6
1	B	472	GLY	3.6
1	D	471	SER	3.6
1	C	403	PHE	3.6
1	A	55	PRO	3.5
1	E	441	ASN	3.5
1	C	139	GLN	3.5
1	C	95	ASP	3.4
1	D	56	ASN	3.3
1	E	442	LYS	3.2
1	E	323	ILE	3.2
1	E	324	CYS	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	89	THR	3.1
1	E	139	GLN	3.0
1	E	329	LEU	3.0
1	C	437	GLU	3.0
1	E	54	SER	3.0
1	C	88	ASP	3.0
1	D	54	SER	3.0
1	E	176	SER	3.0
1	E	315	ARG	3.0
1	E	382	LEU	3.0
1	B	441	ASN	3.0
1	D	58	ASN	2.9
1	E	443	TYR	2.9
1	B	174	VAL	2.9
1	C	56	ASN	2.9
1	B	95	ASP	2.9
1	D	88	ASP	2.9
1	D	403	PHE	2.8
1	E	89	THR	2.8
1	C	57	ASN	2.8
1	D	438	ASP	2.8
1	E	472	GLY	2.7
1	B	385	GLU	2.7
1	B	175	ALA	2.7
1	B	324	CYS	2.7
1	C	329	LEU	2.7
1	B	349	THR	2.7
1	D	440	LEU	2.7
1	E	174	VAL	2.7
1	C	439	PRO	2.7
1	E	438	ASP	2.7
1	A	329	LEU	2.7
1	C	58	ASN	2.7
1	D	176	SER	2.6
1	D	439	PRO	2.6
1	C	91	PHE	2.6
1	C	442	LYS	2.6
1	B	348	VAL	2.6
1	E	96	THR	2.6
1	D	87	PRO	2.6
1	A	315	ARG	2.6
1	C	160	ILE	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	183	THR	2.6
1	C	60	LYS	2.6
1	B	439	PRO	2.6
1	D	315	ARG	2.5
1	A	439	PRO	2.5
1	C	84	PHE	2.5
1	E	60	LYS	2.5
1	B	183	THR	2.4
1	B	323	ILE	2.4
1	D	329	LEU	2.4
1	D	442	LYS	2.4
1	E	90	SER	2.4
1	D	385	GLU	2.4
1	D	387	MET	2.4
1	D	324	CYS	2.4
1	B	471	SER	2.4
1	D	340	THR	2.4
1	E	62	LEU	2.3
1	A	350	LYS	2.3
1	C	331	VAL	2.3
1	E	471	SER	2.3
1	E	161	GLY	2.3
1	E	182	ALA	2.3
1	B	140	PRO	2.3
1	B	116	VAL	2.3
1	D	60	LYS	2.3
1	C	323	ILE	2.3
1	D	102	ALA	2.3
1	E	181	ALA	2.3
1	C	183	THR	2.3
1	B	102	ALA	2.2
1	E	175	ALA	2.2
1	A	69	LEU	2.2
1	A	323	ILE	2.2
1	E	452	GLU	2.2
1	B	315	ARG	2.2
1	B	329	LEU	2.1
1	E	21	VAL	2.1
1	B	220	VAL	2.1
1	C	158[A]	CYS	2.1
1	E	20	LYS	2.1
1	E	61	VAL	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	54	SER	2.1
1	D	323	ILE	2.1
1	C	159	LEU	2.1
1	C	469	LEU	2.1
1	A	324	CYS	2.1
1	E	385	GLU	2.0
1	A	160	ILE	2.0
1	A	402	GLN	2.0
1	E	103	CYS	2.0
1	B	60	LYS	2.0
1	C	85	GLY	2.0

## 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](#)

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	GOL	B	501	6/6	0.42	0.43	100,103,106,107	0
2	GOL	E	503	6/6	0.43	0.32	79,80,82,86	0
2	GOL	B	505	6/6	0.45	0.33	131,134,134,135	0
2	GOL	E	504	6/6	0.54	0.35	81,84,86,86	0
2	GOL	C	503	6/6	0.55	0.28	85,87,88,89	0
2	GOL	C	502	6/6	0.58	0.23	78,83,85,86	0
2	GOL	B	503	6/6	0.62	0.30	85,91,93,95	0
2	GOL	C	501	6/6	0.64	0.25	86,88,90,90	0
2	GOL	D	504	6/6	0.66	0.24	65,78,80,81	0
2	GOL	D	502	6/6	0.71	0.32	78,82,86,86	0
2	GOL	D	503	6/6	0.74	0.38	60,67,68,72	0
3	MG	A	510	1/1	0.75	0.27	76,76,76,76	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MG	C	505	1/1	0.78	0.28	58,58,58,58	0
2	GOL	A	502	6/6	0.78	0.18	81,83,86,86	0
3	MG	E	507	1/1	0.78	0.20	64,64,64,64	0
2	GOL	B	504	6/6	0.79	0.25	71,74,76,77	0
2	GOL	A	503	6/6	0.79	0.38	58,60,62,64	0
2	GOL	A	504	6/6	0.80	0.19	54,60,62,62	0
3	MG	A	508	1/1	0.82	0.15	65,65,65,65	0
2	GOL	E	502	6/6	0.82	0.25	58,67,69,71	0
3	MG	D	508	1/1	0.82	0.46	72,72,72,72	0
2	GOL	B	502	6/6	0.84	0.18	92,92,94,94	0
3	MG	C	508	1/1	0.84	0.12	91,91,91,91	0
2	GOL	A	501	6/6	0.85	0.28	38,62,69,71	0
2	GOL	E	501	6/6	0.85	0.27	72,73,74,75	0
3	MG	D	510	1/1	0.87	0.07	63,63,63,63	0
3	MG	B	509	1/1	0.87	0.54	62,62,62,62	0
3	MG	A	507	1/1	0.88	0.13	69,69,69,69	0
3	MG	D	506	1/1	0.89	0.15	50,50,50,50	0
3	MG	B	511	1/1	0.89	0.34	67,67,67,67	0
3	MG	C	509	1/1	0.89	0.13	60,60,60,60	0
2	GOL	D	501	6/6	0.90	0.16	33,59,66,71	0
3	MG	B	506	1/1	0.91	0.15	52,52,52,52	0
3	MG	C	507	1/1	0.92	0.25	55,55,55,55	0
3	MG	C	506	1/1	0.93	0.14	56,56,56,56	0
3	MG	A	509	1/1	0.93	0.15	45,45,45,45	0
3	MG	A	505	1/1	0.94	0.15	42,42,42,42	0
3	MG	A	511	1/1	0.94	0.09	45,45,45,45	0
3	MG	E	505	1/1	0.94	0.07	61,61,61,61	0
3	MG	B	508	1/1	0.94	0.20	50,50,50,50	0
3	MG	A	506	1/1	0.95	0.09	42,42,42,42	0
3	MG	B	510	1/1	0.95	0.10	46,46,46,46	0
3	MG	D	507	1/1	0.95	0.49	66,66,66,66	0
3	MG	E	506	1/1	0.97	0.28	59,59,59,59	0
3	MG	B	507	1/1	0.97	0.12	52,52,52,52	0
3	MG	D	509	1/1	0.97	0.15	60,60,60,60	0
3	MG	D	505	1/1	0.97	0.09	46,46,46,46	0
3	MG	E	508	1/1	0.97	0.09	58,58,58,58	0
3	MG	C	504	1/1	0.98	0.06	46,46,46,46	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.