



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 22, 2020 – 09:34 AM BST

PDB ID : 6O3G  
Title : Crystal structure of the Fab fragment of the human HIV-1 neutralizing antibody PGZL1 in complex with its MPER peptide epitope (region 671-683 of HIV-1 gp41).  
Authors : Irimia, A.; Wilson, I.A.  
Deposited on : 2019-02-26  
Resolution : 3.65 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
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)	:	2.13.1

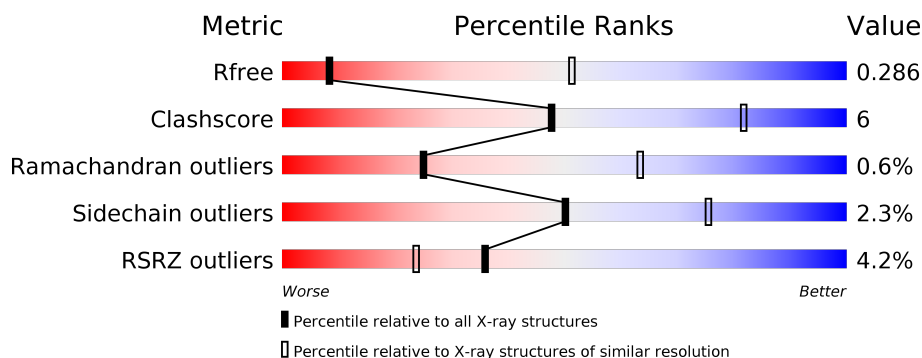
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.65 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	1341 (3.78-3.50)
Clashscore	141614	1439 (3.78-3.50)
Ramachandran outliers	138981	1391 (3.78-3.50)
Sidechain outliers	138945	1391 (3.78-3.50)
RSRZ outliers	127900	1242 (3.78-3.50)

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	B	215	<div> <div></div> <div>85%14%</div> </div>
1	D	215	<div> <div>4%</div> <div>84%15%</div> </div>
1	F	215	<div> <div>18%</div> <div>82%15%</div> </div>
1	L	215	<div> <div>%</div> <div>85%13%</div> </div>
2	A	227	<div> <div>2%</div> <div>86%13%</div> </div>
2	C	227	<div> <div>3%</div> <div>78%21%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	E	227	 5% 87% 12%
2	H	227	 2% 82% 15%
3	G	16	 69% 25% 6%
3	I	16	 63% 31% 6%
3	Q	16	 63% 31% 6%
3	S	16	 75% 25%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 13916 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 PGZL1 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	214	Total	C	N	O	S	0	2	0
			1639	1017	281	335	6			
1	B	214	Total	C	N	O	S	0	1	0
			1633	1014	280	333	6			
1	D	214	Total	C	N	O	S	0	1	0
			1633	1014	280	333	6			
1	F	210	Total	C	N	O	S	0	2	0
			1612	1003	275	328	6			

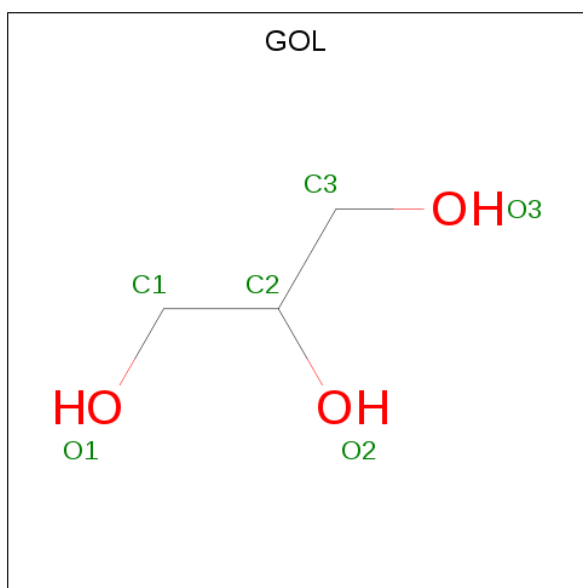
- Molecule 2 is a protein called PGZL1 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	222	Total	C	N	O	S	0	5	0
			1672	1065	277	322	8			
2	A	226	Total	C	N	O	S	0	3	0
			1686	1071	281	326	8			
2	C	227	Total	C	N	O	S	0	4	0
			1695	1076	282	328	9			
2	E	227	Total	C	N	O	S	0	3	0
			1692	1074	282	327	9			

- Molecule 3 is a protein called MEPER peptide, region 671-683 of HIV-1 gp41.

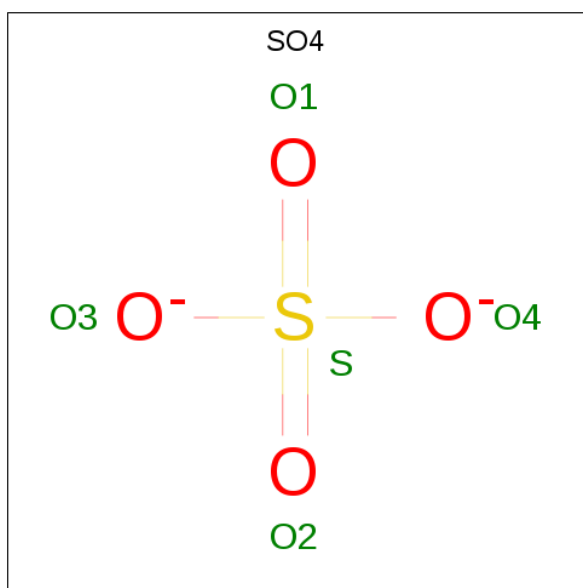
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	Q	16	Total	C	N	O	0	0	0
			157	109	25	23			
3	S	16	Total	C	N	O	0	0	0
			157	109	25	23			
3	G	16	Total	C	N	O	0	0	0
			157	109	25	23			
3	I	16	Total	C	N	O	0	0	0
			157	109	25	23			

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	L	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



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

*Continued on next page...*

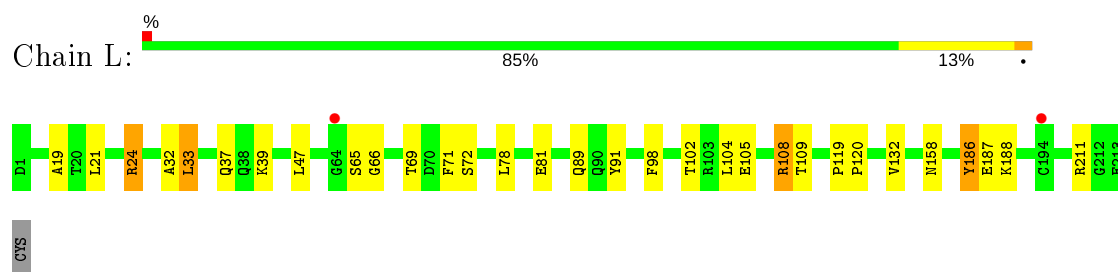
*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	O	S	0	0
			5	4	1		
5	E	1	Total	O	S	0	0
			5	4	1		

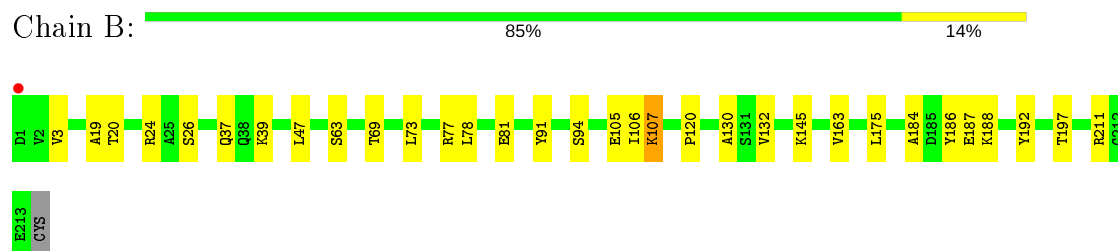
### 3 Residue-property plots [i](#)

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.

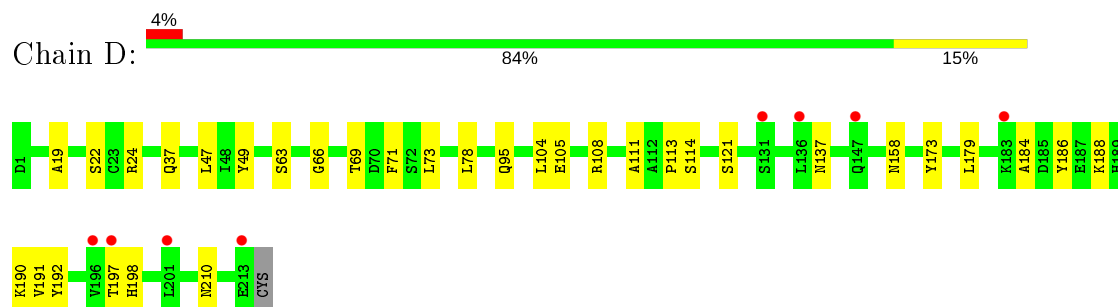
- Molecule 1: PGZL1 light chain



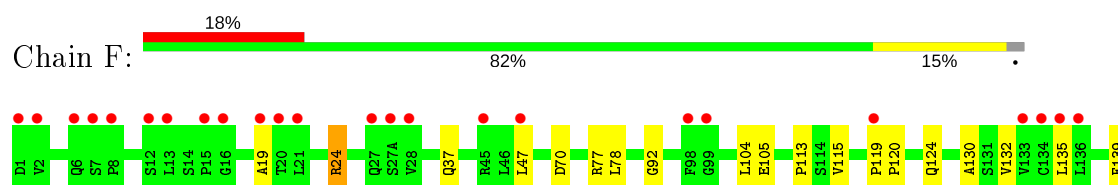
- Molecule 1: PGZL1 light chain

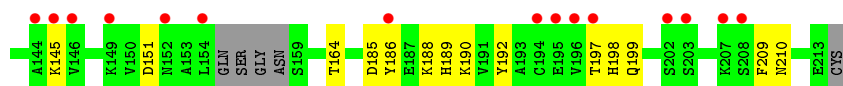


- Molecule 1: PGZL1 light chain

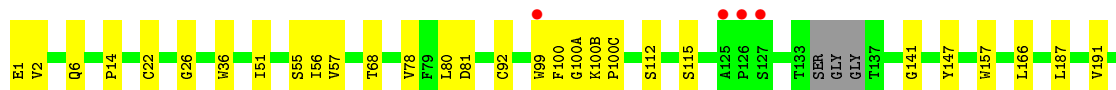
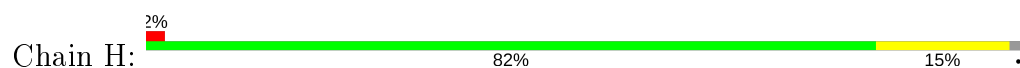


- Molecule 1: PGZL1 light chain

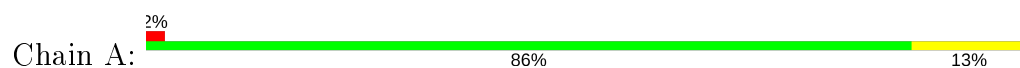




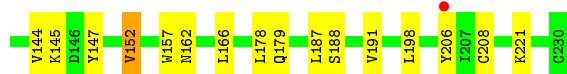
• Molecule 2: PGZL1 heavy chain



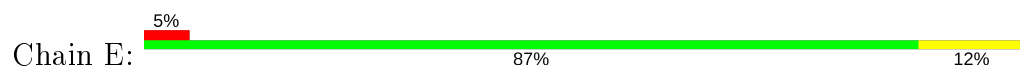
• Molecule 2: PGZL1 heavy chain



• Molecule 2: PGZL1 heavy chain



• Molecule 2: PGZL1 heavy chain



• Molecule 3: MEPER peptide, region 671-683 of HIV-1 gp41



Chain Q:  63% 31% 6%



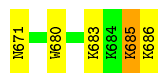
- Molecule 3: MEPER peptide, region 671-683 of HIV-1 gp41

Chain S:  75% 25%



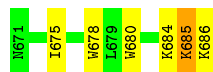
- Molecule 3: MEPER peptide, region 671-683 of HIV-1 gp41

Chain G:  69% 25% 6%



- Molecule 3: MEPER peptide, region 671-683 of HIV-1 gp41

Chain I:  63% 31% 6%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	113.53Å 113.53Å 300.54Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.13 – 3.65 46.72 – 3.65	Depositor EDS
% Data completeness (in resolution range)	93.2 (44.13-3.65) 93.2 (46.72-3.65)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.19	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.07 (at 3.66Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, $R_{free}$	0.260 , 0.289 0.260 , 0.286	Depositor DCC
$R_{free}$ test set	1218 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	95.4	Xtriage
Anisotropy	0.016	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 59.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.043 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	13916	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.28% 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, 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	B	0.27	0/1670	0.57	0/2266
1	D	0.28	0/1670	0.56	0/2266
1	F	0.27	0/1648	0.56	0/2235
1	L	0.28	0/1676	0.58	0/2274
2	A	0.28	0/1737	0.53	0/2366
2	C	0.29	0/1749	0.58	0/2382
2	E	0.28	0/1743	0.57	1/2374 (0.0%)
2	H	0.28	0/1728	0.58	0/2353
3	G	0.23	0/164	0.30	0/220
3	I	0.25	0/164	0.34	0/220
3	Q	0.23	0/164	0.32	0/220
3	S	0.23	0/164	0.42	0/220
All	All	0.28	0/14277	0.56	1/19396 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	187	LEU	CA-CB-CG	5.03	126.87	115.30

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	B	1633	0	1582	18	0
1	D	1633	0	1582	16	0
1	F	1612	0	1563	23	0
1	L	1639	0	1586	21	0
2	A	1686	0	1671	16	0
2	C	1695	0	1681	29	0
2	E	1692	0	1676	21	0
2	H	1672	0	1664	21	0
3	G	157	0	155	3	0
3	I	157	0	155	6	0
3	Q	157	0	155	4	0
3	S	157	0	155	3	0
4	L	6	0	8	0	0
5	A	5	0	0	0	0
5	C	5	0	0	0	0
5	E	5	0	0	0	0
5	H	5	0	0	0	0
All	All	13916	0	13633	167	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 6.

The worst 5 of 167 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)
2:H:212:HIS:CD2	2:H:214:PRO:HD2	2.19	0.78
2:H:212:HIS:HD2	2:H:214:PRO:HD2	1.51	0.76
2:E:99:TRP:O	2:E:100(A):GLY:N	2.23	0.71
2:E:48:MET:HG2	2:E:63:PHE:HE2	1.58	0.69
1:L:24:ARG:HA	1:L:69:THR:O	1.92	0.68

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	B	213/215 (99%)	208 (98%)	5 (2%)	0	100	100
1	D	213/215 (99%)	208 (98%)	5 (2%)	0	100	100
1	F	208/215 (97%)	204 (98%)	4 (2%)	0	100	100
1	L	214/215 (100%)	202 (94%)	10 (5%)	2 (1%)	17	56
2	A	227/227 (100%)	220 (97%)	6 (3%)	1 (0%)	34	70
2	C	229/227 (101%)	215 (94%)	12 (5%)	2 (1%)	17	56
2	E	228/227 (100%)	218 (96%)	9 (4%)	1 (0%)	34	70
2	H	223/227 (98%)	211 (95%)	10 (4%)	2 (1%)	17	56
3	G	14/16 (88%)	13 (93%)	0	1 (7%)	1	14
3	I	14/16 (88%)	12 (86%)	2 (14%)	0	100	100
3	Q	14/16 (88%)	12 (86%)	1 (7%)	1 (7%)	1	14
3	S	14/16 (88%)	14 (100%)	0	0	100	100
All	All	1811/1832 (99%)	1737 (96%)	64 (4%)	10 (1%)	25	63

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	100(B)	LYS
2	A	182	SER
2	C	30	SER
2	E	100	PHE
3	G	685	LYS

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	185/185 (100%)	182 (98%)	3 (2%)	62	82
1	D	185/185 (100%)	181 (98%)	4 (2%)	52	76

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	183/185 (99%)	180 (98%)	3 (2%)	62	82
1	L	186/185 (100%)	181 (97%)	5 (3%)	44	72
2	A	190/188 (101%)	187 (98%)	3 (2%)	62	82
2	C	192/188 (102%)	184 (96%)	8 (4%)	30	62
2	E	191/188 (102%)	186 (97%)	5 (3%)	46	73
2	H	190/188 (101%)	187 (98%)	3 (2%)	62	82
3	G	16/16 (100%)	16 (100%)	0	100	100
3	I	16/16 (100%)	15 (94%)	1 (6%)	18	52
3	Q	16/16 (100%)	15 (94%)	1 (6%)	18	52
3	S	16/16 (100%)	16 (100%)	0	100	100
All	All	1566/1556 (101%)	1530 (98%)	36 (2%)	50	75

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	121	SER
2	C	100	PHE
2	E	205	THR
2	C	1	GLU
2	C	152	VAL

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

Mol	Chain	Res	Type
1	D	155	GLN
1	D	158	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

5 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	$\# Z  > 2$
5	SO4	H	301	-	4,4,4	0.14	0	6,6,6	0.06	0
5	SO4	E	301	-	4,4,4	0.14	0	6,6,6	0.05	0
4	GOL	L	301	-	5,5,5	0.91	0	5,5,5	0.96	0
5	SO4	C	301	-	4,4,4	0.13	0	6,6,6	0.06	0
5	SO4	A	301	-	4,4,4	0.14	0	6,6,6	0.05	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
4	GOL	L	301	-	-	4/4/4/4	-

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
4	L	301	GOL	O1-C1-C2-C3
4	L	301	GOL	O1-C1-C2-O2
4	L	301	GOL	O2-C2-C3-O3
4	L	301	GOL	C1-C2-C3-O3

There are no ring outliers.

No monomer is involved in short contacts.

## 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	B	214/215 (99%)	0.01	1 (0%) 91 84	54, 88, 117, 153	0
1	D	214/215 (99%)	0.43	8 (3%) 41 27	52, 104, 151, 172	0
1	F	210/215 (97%)	1.08	39 (18%) 1 0	87, 142, 168, 184	0
1	L	214/215 (99%)	0.12	2 (0%) 84 73	45, 86, 131, 148	0
2	A	226/227 (99%)	0.09	4 (1%) 68 53	47, 71, 131, 166	1 (0%)
2	C	227/227 (100%)	0.20	6 (2%) 56 40	42, 78, 159, 198	1 (0%)
2	E	227/227 (100%)	0.41	12 (5%) 26 17	64, 104, 186, 205	1 (0%)
2	H	222/227 (97%)	0.18	4 (1%) 68 53	47, 83, 136, 163	0
3	G	16/16 (100%)	-0.03	0 100 100	51, 79, 102, 105	0
3	I	16/16 (100%)	-0.12	0 100 100	58, 72, 99, 106	0
3	Q	16/16 (100%)	-0.09	0 100 100	52, 76, 83, 99	0
3	S	16/16 (100%)	0.01	0 100 100	57, 79, 97, 100	0
All	All	1818/1832 (99%)	0.30	76 (4%) 36 24	42, 91, 160, 205	3 (0%)

The worst 5 of 76 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	191	VAL	7.5
1	F	196	VAL	5.9
1	F	152	ASN	5.5
2	E	138	ALA	5.4
2	A	134	SER	5.3

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

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( $\text{\AA}^2$ )	Q<0.9
5	SO4	E	301	5/5	0.82	0.19	94,94,129,159	0
5	SO4	C	301	5/5	0.83	0.24	91,92,109,141	0
4	GOL	L	301	6/6	0.88	0.18	72,74,75,98	0
5	SO4	H	301	5/5	0.90	0.28	80,81,86,123	0
5	SO4	A	301	5/5	0.95	0.16	91,93,97,131	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.