



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:38 PM GMT

PDB ID : 4TVP  
Title : Crystal Structure of the HIV-1 BG505 SOSIP.664 Env Trimer Ectodomain, Comprising Atomic-Level Definition of Pre-Fusion gp120 and gp41, in Complex with Human Antibodies PGT122 and 35O22  
Authors : Pancera, M.; Zhou, T.; Kwong, P.D.  
Deposited on : 2014-06-27  
Resolution : 3.10 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

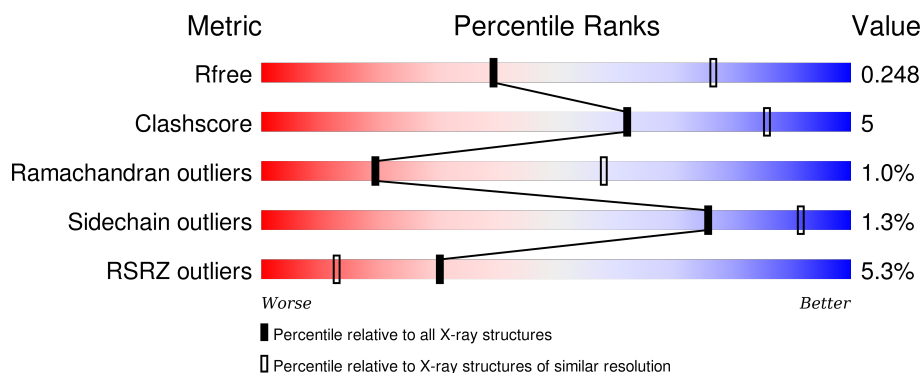
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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))
$R_{free}$	91344	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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. 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	G	481	<div> <div>15%</div> <div>78%</div> <div>15%</div> <div>6%</div> </div>
2	B	153	<div> <div>66%</div> <div>14%</div> <div>•</div> <div>18%</div> </div>
3	L	213	<div> <div>81%</div> <div>17%</div> <div>••</div> </div>
4	H	235	<div> <div>3%</div> <div>81%</div> <div>16%</div> <div>•</div> </div>
5	D	243	<div> <div>15%</div> <div>88%</div> <div>12%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	E	216	<div><div></div><div>12%</div><div>86%</div><div>13%</div><div></div></div>

## 2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 12188 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 Envelope glycoprotein gp160.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	453	Total	C	N	O	S	0	0	0
			3565	2236	630	671	28			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	332	ASN	THR	engineered mutation	UNP Q2N0S5
G	501	CYS	ALA	engineered mutation	UNP Q2N0S5
G	509	ARG	-	expression tag	UNP Q2N0S5
G	510	ARG	-	expression tag	UNP Q2N0S5
G	511	ARG	-	expression tag	UNP Q2N0S5
G	512	ARG	-	expression tag	UNP Q2N0S5
G	513	ARG	-	expression tag	UNP Q2N0S5

- Molecule 2 is a protein called Envelope glycoprotein gp160.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	126	Total	C	N	O	S	0	0	0
			1001	633	172	190	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	559	PRO	ILE	engineered mutation	UNP Q2N0S5
B	605	CYS	THR	engineered mutation	UNP Q2N0S5

- Molecule 3 is a protein called PGT122 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	210	Total	C	N	O	S	0	0	0
			1589	998	267	320	4			

- Molecule 4 is a protein called PGT122 Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	228	Total	C	N	O	S	0	0	0
			1742	1109	295	333	5			

- Molecule 5 is a protein called 35O22 Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	242	Total	C	N	O	S	0	0	0
			1832	1165	306	353	8			

- Molecule 6 is a protein called 35O22 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	E	213	Total	C	N	O	S	0	0	0
			1615	1012	267	328	8			

- Molecule 7 is a polymer of unknown type called SUGAR (7-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	G	7	Total	C	N	O	0	0
			83	46	2	35		

- Molecule 8 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	G	4	Total	C	N	O	0	0
			50	28	2	20		

- Molecule 9 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	G	2	Total	C	N	O	0	0
			28	16	2	10		
9	G	2	Total	C	N	O	0	0
			28	16	2	10		
9	G	2	Total	C	N	O	0	0
			28	16	2	10		
9	G	2	Total	C	N	O	0	0
			28	16	2	10		
9	G	2	Total	C	N	O	0	0
			28	16	2	10		

*Continued on next page...*

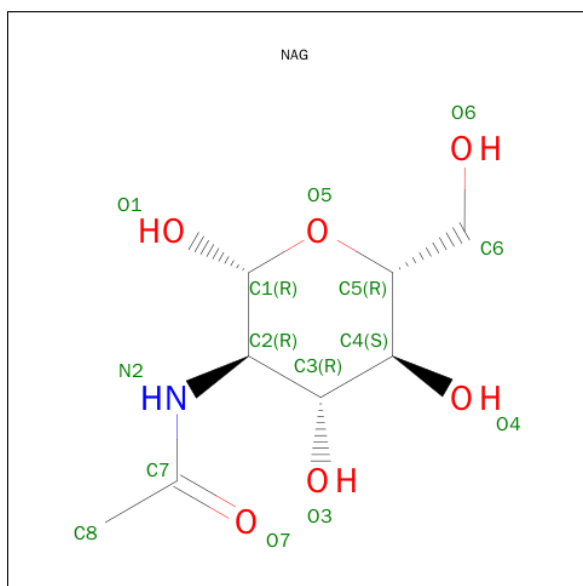
*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	G	2	Total	C	N	O	0	0
			28	16	2	10		
9	G	2	Total	C	N	O	0	0
			28	16	2	10		
9	G	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 10 is a polymer of unknown type called SUGAR (6-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
10	G	6	Total	C	N	O	0	0
			72	40	2	30		

- Molecule 11 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
11	G	1	Total	C	N	O	0	0
			14	8	1	5		
11	G	1	Total	C	N	O	0	0
			14	8	1	5		
11	G	1	Total	C	N	O	0	0
			14	8	1	5		
11	G	1	Total	C	N	O	0	0
			14	8	1	5		
11	B	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
11	B	1	Total	C	N	O	0	0
			14	8	1	5		
11	B	1	Total	C	N	O	0	0
			14	8	1	5		
11	H	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 12 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
12	G	3	Total	C	N	O	0	0
			39	22	2	15		

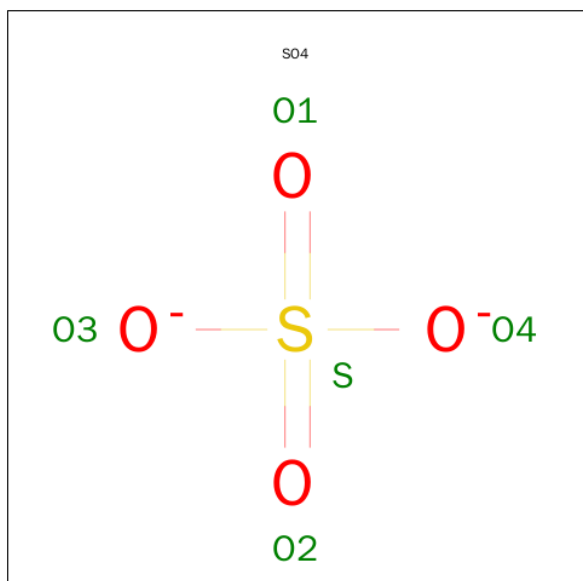
- Molecule 13 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
13	G	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 14 is a polymer of unknown type called SUGAR (10-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
14	G	10	Total	C	N	O	0	0
			116	64	2	50		

- Molecule 15 is SULFATE ION (three-letter code: SO<sub>4</sub>) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	G	1	Total O S 5 4 1	0	0
15	G	1	Total O S 5 4 1	0	0
15	G	1	Total O S 5 4 1	0	0
15	G	1	Total O S 5 4 1	0	0
15	G	1	Total O S 5 4 1	0	0
15	G	1	Total O S 5 4 1	0	0
15	G	1	Total O S 5 4 1	0	0
15	B	1	Total O S 5 4 1	0	0
15	B	1	Total O S 5 4 1	0	0
15	L	1	Total O S 5 4 1	0	0

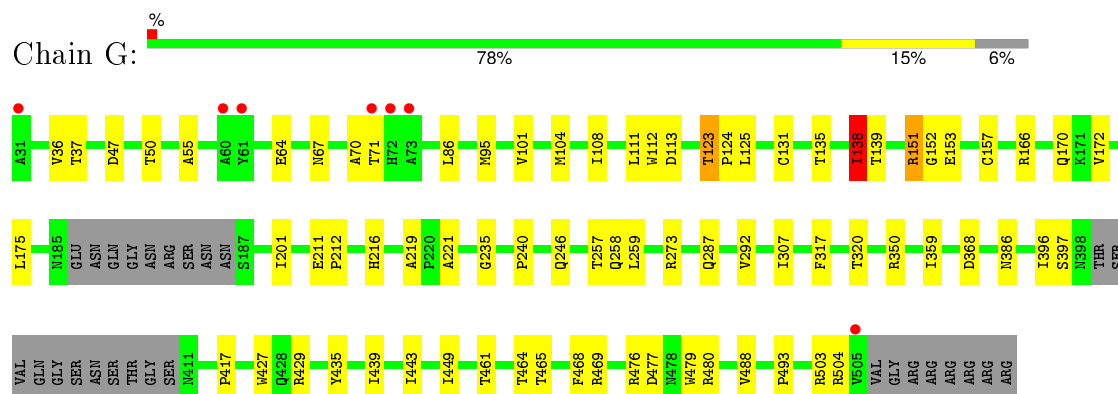
- Molecule 16 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	G	17	Total O 17 17	0	0
16	B	2	Total O 2 2	0	0
16	L	3	Total O 3 3	0	0
16	H	2	Total O 2 2	0	0
16	D	11	Total O 11 11	0	0
16	E	2	Total O 2 2	0	0

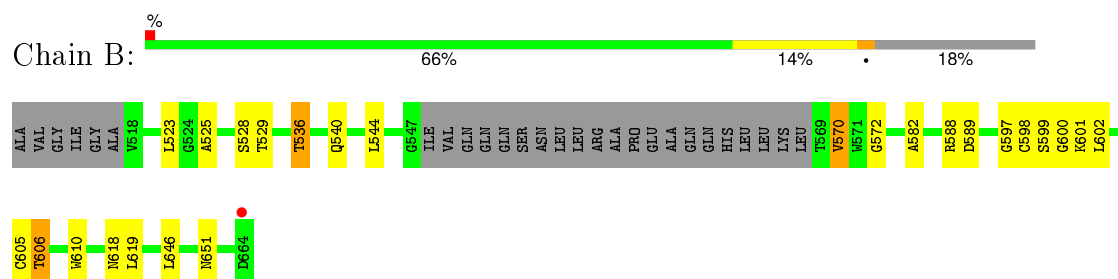
### 3 Residue-property plots [i](#)

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

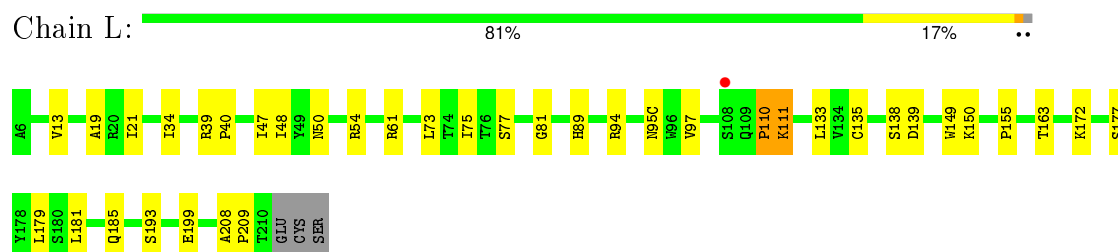
- Molecule 1: Envelope glycoprotein gp160



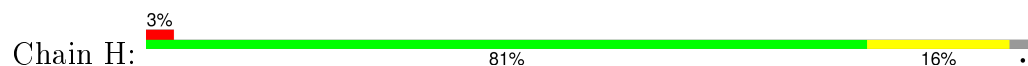
- Molecule 2: Envelope glycoprotein gp160

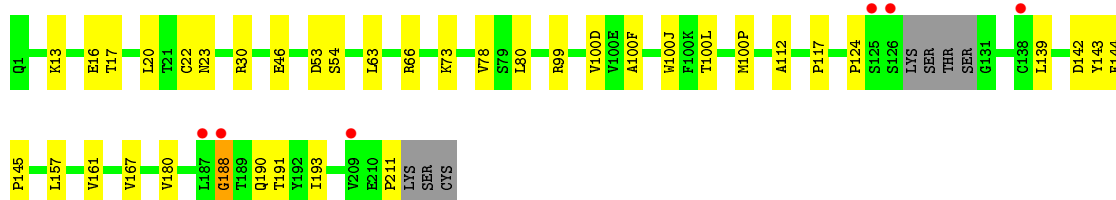


- Molecule 3: PGT122 Light chain

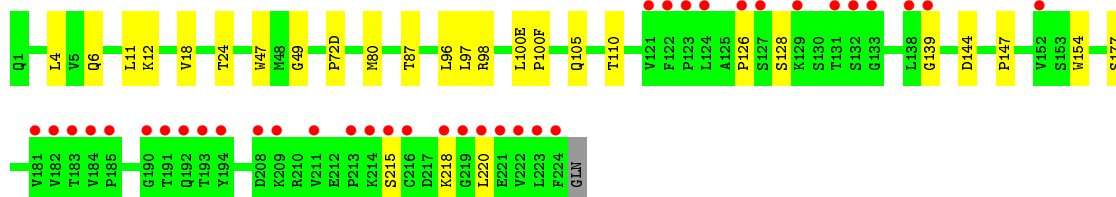
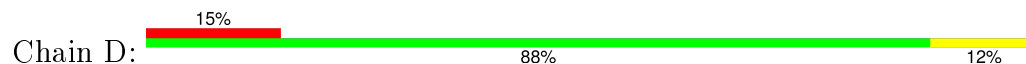


- Molecule 4: PGT122 Heavy chain

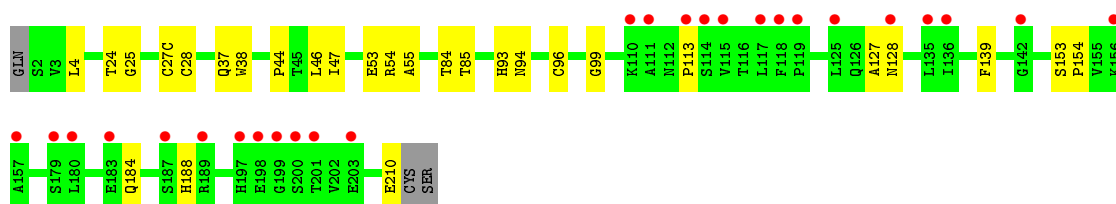
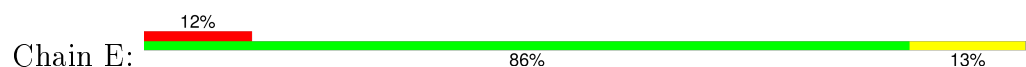




• Molecule 5: 35O22 Heavy chain



• Molecule 6: 35O22 Light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	128.89Å 128.89Å 313.42Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.74 – 3.10 40.74 – 3.10	Depositor EDS
% Data completeness (in resolution range)	55.0 (40.74-3.10) 55.1 (40.74-3.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.51 (at 3.12Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1702)	Depositor
R, $R_{free}$	0.213 , 0.248 0.214 , 0.248	Depositor DCC
$R_{free}$ test set	1449 reflections (4.94%)	DCC
Wilson B-factor (Å <sup>2</sup> )	76.8	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , 56.7	EDS
Estimated twinning fraction	0.086 for h,-h-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	1 of 29370 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	12188	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	109.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.14% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: SO4, BMA, NAG, MAN

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	G	0.21	0/3639	0.40	0/4941
2	B	0.20	0/1019	0.37	0/1382
3	L	0.20	0/1632	0.39	0/2236
4	H	0.20	0/1789	0.38	0/2443
5	D	0.20	0/1880	0.37	0/2560
6	E	0.20	0/1659	0.37	0/2269
All	All	0.20	0/11618	0.39	0/15831

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	G	3565	0	3495	43	0
2	B	1001	0	975	15	0
3	L	1589	0	1530	22	0
4	H	1742	0	1715	21	0
5	D	1832	0	1806	16	0
6	E	1615	0	1542	16	0
7	G	83	0	70	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	G	50	0	43	2	0
9	G	224	0	200	2	0
10	G	72	0	61	0	0
11	B	42	0	39	2	0
11	G	56	0	52	1	0
11	H	14	0	13	0	0
12	G	39	0	34	0	0
13	G	61	0	52	0	0
14	G	116	0	97	2	0
15	B	10	0	0	1	0
15	G	35	0	0	2	0
15	L	5	0	0	0	0
16	B	2	0	0	0	0
16	D	11	0	0	0	0
16	E	2	0	0	0	0
16	G	17	0	0	0	0
16	H	2	0	0	0	0
16	L	3	0	0	0	0
All	All	12188	0	11724	128	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 5.

The worst 5 of 128 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:G:70:ALA:HB2	1:G:111:LEU:HD11	1.69	0.75
1:G:55:ALA:HB3	1:G:216:HIS:HB2	1.71	0.71
6:E:37:GLN:HB2	6:E:47:ILE:HD11	1.74	0.68
1:G:350:ARG:NH2	1:G:396:ILE:O	2.27	0.66
5:D:6:GLN:H	5:D:105:GLN:HE22	1.43	0.66

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	G	447/481 (93%)	405 (91%)	35 (8%)	7 (2%)	12	44
2	B	122/153 (80%)	108 (88%)	11 (9%)	3 (2%)	7	32
3	L	208/213 (98%)	196 (94%)	10 (5%)	2 (1%)	19	58
4	H	224/235 (95%)	211 (94%)	11 (5%)	2 (1%)	21	61
5	D	240/243 (99%)	228 (95%)	12 (5%)	0	100	100
6	E	211/216 (98%)	197 (93%)	14 (7%)	0	100	100
All	All	1452/1541 (94%)	1345 (93%)	93 (6%)	14 (1%)	19	58

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	138	ILE
3	L	110	PRO
1	G	71	THR
1	G	152	GLY
2	B	602	LEU

### 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	G	404/428 (94%)	398 (98%)	6 (2%)	72	90
2	B	108/129 (84%)	105 (97%)	3 (3%)	51	82
3	L	178/181 (98%)	176 (99%)	2 (1%)	80	93
4	H	198/205 (97%)	195 (98%)	3 (2%)	72	90
5	D	205/206 (100%)	204 (100%)	1 (0%)	92	96
6	E	186/189 (98%)	185 (100%)	1 (0%)	92	96
All	All	1279/1338 (96%)	1263 (99%)	16 (1%)	76	91

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	570	VAL
2	B	606	THR
4	H	100(J)	TRP
2	B	536	THR
4	H	139	LEU

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

### 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 ⓘ

51 carbohydrates 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	$\# Z  > 2$
7	NAG	G	1088	1,7	14,14,15	0.32	0	15,19,21	0.31	0
7	NAG	G	1089	7	14,14,15	0.42	0	15,19,21	1.32	1 (6%)
7	BMA	G	1090	7	11,11,12	0.66	0	14,15,17	0.73	0
7	MAN	G	1091	7	11,11,12	0.65	0	14,15,17	1.25	2 (14%)
7	MAN	G	1092	7	11,11,12	0.82	1 (9%)	14,15,17	1.37	3 (21%)
7	MAN	G	1093	7	11,11,12	0.61	0	14,15,17	1.13	2 (14%)
7	MAN	G	1094	7	11,11,12	0.69	0	14,15,17	1.04	2 (14%)
8	NAG	G	1137	1,8	14,14,15	0.42	0	15,19,21	0.40	0
8	NAG	G	1138	8	14,14,15	0.39	0	15,19,21	0.25	0
8	BMA	G	1139	8	11,11,12	0.51	0	14,15,17	0.81	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	MAN	G	1140	8	11,11,12	0.68	0	14,15,17	1.08	2 (14%)
13	NAG	G	1156	1,13	14,14,15	0.25	0	15,19,21	0.26	0
13	NAG	G	1157	13	14,14,15	0.22	0	15,19,21	0.26	0
13	BMA	G	1158	13	11,11,12	0.59	0	14,15,17	0.75	0
13	MAN	G	1159	13	11,11,12	0.70	0	14,15,17	1.09	2 (14%)
9	NAG	G	1160	1,9	14,14,15	0.28	0	15,19,21	0.37	0
9	NAG	G	1161	9	14,14,15	0.44	0	15,19,21	1.28	1 (6%)
13	MAN	G	1169	13	11,11,12	0.69	0	14,15,17	1.15	2 (14%)
9	NAG	G	1197	1,9	14,14,15	0.20	0	15,19,21	0.31	0
9	NAG	G	1198	9	14,14,15	0.21	0	15,19,21	0.33	0
9	NAG	G	1234	1,9	14,14,15	0.17	0	15,19,21	0.33	0
9	NAG	G	1235	9	14,14,15	0.21	0	15,19,21	0.29	0
10	NAG	G	1262	1,10	14,14,15	0.24	0	15,19,21	0.41	0
10	NAG	G	1263	10	14,14,15	0.33	0	15,19,21	0.32	0
10	BMA	G	1264	10	11,11,12	0.65	0	14,15,17	0.71	0
10	MAN	G	1265	10	11,11,12	0.63	0	14,15,17	1.14	2 (14%)
10	MAN	G	1268	10	11,11,12	0.67	0	14,15,17	1.09	2 (14%)
10	MAN	G	1269	10	11,11,12	0.75	0	14,15,17	1.28	2 (14%)
9	NAG	G	1295	1,9	14,14,15	0.25	0	15,19,21	0.40	0
9	NAG	G	1296	9	14,14,15	0.18	0	15,19,21	0.31	0
9	NAG	G	1301	1,9	14,14,15	0.19	0	15,19,21	0.27	0
9	NAG	G	1302	9	14,14,15	0.22	0	15,19,21	0.34	0
14	NAG	G	1331	1,14	14,14,15	0.25	0	15,19,21	0.27	0
14	NAG	G	1332	14	14,14,15	0.29	0	15,19,21	0.43	0
14	BMA	G	1333	14	11,11,12	0.76	0	14,15,17	1.31	2 (14%)
14	MAN	G	1334	14	11,11,12	0.79	0	14,15,17	1.03	2 (14%)
14	MAN	G	1335	14	11,11,12	0.70	0	14,15,17	1.05	2 (14%)
14	MAN	G	1336	14	11,11,12	0.69	0	14,15,17	1.05	2 (14%)
14	MAN	G	1337	14	11,11,12	0.62	0	14,15,17	1.18	1 (7%)
14	MAN	G	1338	14	11,11,12	0.57	0	14,15,17	1.09	1 (7%)
14	MAN	G	1339	14	11,11,12	0.58	0	14,15,17	1.49	3 (21%)
14	MAN	G	1340	14	11,11,12	0.78	0	14,15,17	1.26	2 (14%)
9	NAG	G	1363	1,9	14,14,15	0.21	0	15,19,21	0.25	0
9	NAG	G	1364	9	14,14,15	0.20	0	15,19,21	0.35	0
9	NAG	G	1386	1,9	14,14,15	0.22	0	15,19,21	0.57	0
9	NAG	G	1387	9	14,14,15	0.22	0	15,19,21	0.31	0
9	NAG	G	1392	1,9	14,14,15	0.24	0	15,19,21	0.33	0
9	NAG	G	1393	9	14,14,15	0.29	0	15,19,21	0.38	0
12	NAG	G	1448	1,12	14,14,15	0.19	0	15,19,21	0.45	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
12	NAG	G	1449	12	14,14,15	0.25	0	15,19,21	0.22	0
12	BMA	G	1450	12	11,11,12	0.55	0	14,15,17	0.84	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
7	NAG	G	1088	1,7	-	0/6/23/26	0/1/1/1
7	NAG	G	1089	7	-	0/6/23/26	0/1/1/1
7	BMA	G	1090	7	-	0/2/19/22	0/1/1/1
7	MAN	G	1091	7	-	0/2/19/22	0/1/1/1
7	MAN	G	1092	7	-	0/2/19/22	0/1/1/1
7	MAN	G	1093	7	-	0/2/19/22	0/1/1/1
7	MAN	G	1094	7	-	0/2/19/22	0/1/1/1
8	NAG	G	1137	1,8	-	0/6/23/26	0/1/1/1
8	NAG	G	1138	8	-	0/6/23/26	0/1/1/1
8	BMA	G	1139	8	-	0/2/19/22	0/1/1/1
8	MAN	G	1140	8	-	0/2/19/22	0/1/1/1
13	NAG	G	1156	1,13	-	0/6/23/26	0/1/1/1
13	NAG	G	1157	13	-	0/6/23/26	0/1/1/1
13	BMA	G	1158	13	-	0/2/19/22	0/1/1/1
13	MAN	G	1159	13	-	0/2/19/22	0/1/1/1
9	NAG	G	1160	1,9	-	0/6/23/26	0/1/1/1
9	NAG	G	1161	9	-	0/6/23/26	0/1/1/1
13	MAN	G	1169	13	-	0/2/19/22	0/1/1/1
9	NAG	G	1197	1,9	-	0/6/23/26	0/1/1/1
9	NAG	G	1198	9	-	0/6/23/26	0/1/1/1
9	NAG	G	1234	1,9	-	0/6/23/26	0/1/1/1
9	NAG	G	1235	9	-	0/6/23/26	0/1/1/1
10	NAG	G	1262	1,10	-	0/6/23/26	0/1/1/1
10	NAG	G	1263	10	-	0/6/23/26	0/1/1/1
10	BMA	G	1264	10	-	0/2/19/22	0/1/1/1
10	MAN	G	1265	10	-	0/2/19/22	0/1/1/1
10	MAN	G	1268	10	-	0/2/19/22	0/1/1/1
10	MAN	G	1269	10	-	0/2/19/22	1/1/1/1
9	NAG	G	1295	1,9	-	0/6/23/26	0/1/1/1
9	NAG	G	1296	9	-	0/6/23/26	0/1/1/1
9	NAG	G	1301	1,9	-	0/6/23/26	0/1/1/1
9	NAG	G	1302	9	-	0/6/23/26	0/1/1/1
14	NAG	G	1331	1,14	-	0/6/23/26	0/1/1/1
14	NAG	G	1332	14	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
14	BMA	G	1333	14	-	0/2/19/22	0/1/1/1
14	MAN	G	1334	14	-	0/2/19/22	0/1/1/1
14	MAN	G	1335	14	-	0/2/19/22	0/1/1/1
14	MAN	G	1336	14	-	0/2/19/22	0/1/1/1
14	MAN	G	1337	14	-	0/2/19/22	0/1/1/1
14	MAN	G	1338	14	-	0/2/19/22	0/1/1/1
14	MAN	G	1339	14	-	0/2/19/22	0/1/1/1
14	MAN	G	1340	14	-	0/2/19/22	1/1/1/1
9	NAG	G	1363	1,9	-	0/6/23/26	0/1/1/1
9	NAG	G	1364	9	-	0/6/23/26	0/1/1/1
9	NAG	G	1386	1,9	-	0/6/23/26	0/1/1/1
9	NAG	G	1387	9	-	0/6/23/26	0/1/1/1
9	NAG	G	1392	1,9	-	0/6/23/26	0/1/1/1
9	NAG	G	1393	9	-	0/6/23/26	0/1/1/1
12	NAG	G	1448	1,12	-	0/6/23/26	0/1/1/1
12	NAG	G	1449	12	-	0/6/23/26	0/1/1/1
12	BMA	G	1450	12	-	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	G	1092	MAN	C1-C2	2.23	1.57	1.52

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	G	1268	MAN	O2-C2-C3	-2.69	104.71	110.12
7	G	1091	MAN	O2-C2-C3	-2.29	105.50	110.12
10	G	1265	MAN	O2-C2-C3	-2.27	105.55	110.12
7	G	1093	MAN	O2-C2-C3	-2.27	105.55	110.12
13	G	1169	MAN	O2-C2-C3	-2.25	105.59	110.12

There are no chirality outliers.

There are no torsion outliers.

All (2) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
14	G	1340	MAN	C1-C2-C3-C4-C5-O5
10	G	1269	MAN	C1-C2-C3-C4-C5-O5

12 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	G	1089	NAG	1	0
7	G	1090	BMA	1	0
7	G	1091	MAN	1	0
7	G	1092	MAN	1	0
7	G	1093	MAN	1	0
8	G	1137	NAG	1	0
8	G	1140	MAN	1	0
9	G	1160	NAG	1	0
9	G	1161	NAG	2	0
14	G	1332	NAG	1	0
14	G	1336	MAN	1	0
14	G	1340	MAN	1	0

## 5.6 Ligand geometry [i](#)

18 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	$\# Z  > 2$
11	NAG	B	1611	2	14,14,15	0.22	0	15,19,21	0.33	0
11	NAG	B	1618	2	14,14,15	0.20	0	15,19,21	0.22	0
11	NAG	B	1637	2	14,14,15	0.24	0	15,19,21	0.34	0
15	SO4	B	701	-	4,4,4	0.22	0	6,6,6	0.07	0
15	SO4	B	702	-	4,4,4	0.23	0	6,6,6	0.08	0
11	NAG	G	1133	1	14,14,15	0.23	0	15,19,21	0.30	0
11	NAG	G	1276	1	14,14,15	0.43	0	15,19,21	1.30	1 (6%)
11	NAG	G	1355	1	14,14,15	0.22	0	15,19,21	0.26	0
11	NAG	G	1839	1	14,14,15	0.25	0	15,19,21	0.31	0
15	SO4	G	601	-	4,4,4	0.23	0	6,6,6	0.08	0
15	SO4	G	602	-	4,4,4	0.23	0	6,6,6	0.08	0
15	SO4	G	603	-	4,4,4	0.22	0	6,6,6	0.08	0
15	SO4	G	604	-	4,4,4	0.23	0	6,6,6	0.09	0
15	SO4	G	605	-	4,4,4	0.23	0	6,6,6	0.08	0
15	SO4	G	606	-	4,4,4	0.26	0	6,6,6	0.19	0
15	SO4	G	607	-	4,4,4	0.24	0	6,6,6	0.08	0
11	NAG	H	523	4	14,14,15	0.23	0	15,19,21	0.29	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
15	SO4	L	301	-	4,4,4	0.24	0	6,6,6	0.08	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
11	NAG	B	1611	2	-	0/6/23/26	0/1/1/1
11	NAG	B	1618	2	-	0/6/23/26	0/1/1/1
11	NAG	B	1637	2	-	0/6/23/26	0/1/1/1
15	SO4	B	701	-	-	0/0/0/0	0/0/0/0
15	SO4	B	702	-	-	0/0/0/0	0/0/0/0
11	NAG	G	1133	1	-	0/6/23/26	0/1/1/1
11	NAG	G	1276	1	-	0/6/23/26	0/1/1/1
11	NAG	G	1355	1	-	0/6/23/26	0/1/1/1
11	NAG	G	1839	1	-	0/6/23/26	0/1/1/1
15	SO4	G	601	-	-	0/0/0/0	0/0/0/0
15	SO4	G	602	-	-	0/0/0/0	0/0/0/0
15	SO4	G	603	-	-	0/0/0/0	0/0/0/0
15	SO4	G	604	-	-	0/0/0/0	0/0/0/0
15	SO4	G	605	-	-	0/0/0/0	0/0/0/0
15	SO4	G	606	-	-	0/0/0/0	0/0/0/0
15	SO4	G	607	-	-	0/0/0/0	0/0/0/0
11	NAG	H	523	4	-	0/6/23/26	0/1/1/1
15	SO4	L	301	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	G	1276	NAG	C2-N2-C7	4.71	129.09	123.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	B	1618	NAG	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	B	701	SO4	1	0
11	G	1276	NAG	1	0
15	G	602	SO4	1	0
15	G	606	SO4	1	0

## 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	G	453/481 (94%)	-0.35	7 (1%) 76 58	33, 72, 145, 193	0
2	B	126/153 (82%)	-0.26	1 (0%) 87 75	38, 77, 145, 176	0
3	L	210/213 (98%)	-0.46	1 (0%) 91 83	60, 96, 142, 172	0
4	H	228/235 (97%)	-0.20	6 (2%) 59 35	65, 111, 160, 199	0
5	D	242/243 (99%)	0.45	37 (15%) 3 1	51, 123, 284, 386	0
6	E	213/216 (98%)	0.39	26 (12%) 5 2	82, 146, 238, 257	0
All	All	1472/1541 (95%)	-0.09	78 (5%) 30 13	33, 97, 220, 386	0

The worst 5 of 78 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	D	218	LYS	7.3
5	D	214	LYS	7.2
6	E	157	ALA	7.2
5	D	185	PRO	6.8
6	E	197	HIS	6.5

### 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
14	MAN	G	1338	11/12	0.93	0.18	0.67	76,80,93,100	0
7	MAN	G	1093	11/12	0.92	0.22	0.23	97,104,117,118	0
9	NAG	G	1234	14/15	0.91	0.17	0.09	85,99,116,124	0
9	NAG	G	1197	14/15	0.91	0.17	-0.03	75,106,125,140	0
9	NAG	G	1363	14/15	0.92	0.17	-0.45	81,90,114,132	0
9	NAG	G	1160	14/15	0.94	0.16	-0.51	66,90,98,111	0
14	MAN	G	1337	11/12	0.97	0.14	-0.53	49,55,72,92	0
7	NAG	G	1088	14/15	0.98	0.15	-0.67	35,43,56,57	0
14	NAG	G	1332	14/15	0.91	0.16	-0.79	71,88,108,114	0
8	NAG	G	1137	14/15	0.95	0.15	-0.80	72,98,116,118	0
13	NAG	G	1156	14/15	0.96	0.12	-0.86	27,73,91,94	0
10	NAG	G	1262	14/15	0.98	0.14	-0.89	34,54,70,80	0
9	NAG	G	1295	14/15	0.96	0.11	-1.01	59,75,100,110	0
7	NAG	G	1089	14/15	0.97	0.13	-1.29	44,56,72,76	0
14	NAG	G	1331	14/15	0.97	0.12	-1.31	51,62,79,84	0
14	MAN	G	1335	11/12	0.90	0.16	-	94,109,115,117	0
8	BMA	G	1139	11/12	0.84	0.30	-	123,133,142,149	0
9	NAG	G	1301	14/15	0.96	0.13	-	62,77,90,104	0
14	MAN	G	1336	11/12	0.91	0.24	-	135,138,146,152	0
13	NAG	G	1157	14/15	0.90	0.20	-	91,105,116,127	0
14	MAN	G	1340	11/12	0.79	0.25	-	142,149,153,154	0
10	BMA	G	1264	11/12	0.80	0.14	-	123,129,144,150	0
10	MAN	G	1269	11/12	0.86	0.29	-	131,135,139,140	0
13	MAN	G	1169	11/12	0.79	0.32	-	153,155,158,160	0
14	MAN	G	1334	11/12	0.88	0.11	-	100,117,123,132	0
9	NAG	G	1392	14/15	0.90	0.37	-	110,120,137,140	0
8	MAN	G	1140	11/12	0.74	0.59	-	139,156,166,167	0
12	NAG	G	1449	14/15	0.89	0.28	-	102,124,150,177	0
13	MAN	G	1159	11/12	0.88	0.18	-	110,122,126,127	0
7	MAN	G	1092	11/12	0.86	0.16	-	107,114,122,123	0
7	BMA	G	1090	11/12	0.94	0.14	-	66,77,101,104	0
12	NAG	G	1448	14/15	0.91	0.16	-	73,91,103,106	0
13	BMA	G	1158	11/12	0.84	0.21	-	126,135,145,155	0
10	MAN	G	1265	11/12	0.74	0.38	-	122,141,149,150	0
7	MAN	G	1091	11/12	0.95	0.15	-	68,74,91,105	0
9	NAG	G	1387	14/15	0.73	0.41	-	131,149,156,158	0
14	MAN	G	1339	11/12	0.85	0.24	-	115,120,125,126	0
9	NAG	G	1161	14/15	0.88	0.24	-	93,107,118,125	0
9	NAG	G	1235	14/15	0.86	0.49	-	135,137,140,141	0
7	MAN	G	1094	11/12	0.88	0.19	-	116,133,144,153	0
9	NAG	G	1386	14/15	0.85	0.22	-	74,111,130,137	0
9	NAG	G	1302	14/15	0.83	0.39	-	117,131,141,146	0
12	BMA	G	1450	11/12	0.77	0.27	-	161,197,227,281	0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
10	NAG	G	1263	14/15	0.96	0.12	-	77,92,102,112	0
9	NAG	G	1393	14/15	0.82	0.34	-	121,142,144,145	0
8	NAG	G	1138	14/15	0.90	0.21	-	109,118,126,127	0
14	BMA	G	1333	11/12	0.95	0.13	-	65,82,97,99	0
10	MAN	G	1268	11/12	0.91	0.14	-	117,128,138,144	0
9	NAG	G	1296	14/15	0.87	0.40	-	117,131,142,150	0
9	NAG	G	1198	14/15	0.81	0.33	-	143,152,162,166	0
9	NAG	G	1364	14/15	0.83	0.35	-	142,148,156,161	0

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
15	SO4	G	607	5/5	0.84	0.24	1.97	155,157,158,160	0
11	NAG	B	1618	14/15	0.82	0.38	1.60	95,108,114,122	0
11	NAG	G	1133	14/15	0.93	0.28	0.60	104,119,131,133	0
11	NAG	B	1611	14/15	0.80	0.23	-0.09	110,127,140,144	0
15	SO4	G	603	5/5	0.91	0.19	-0.42	147,147,152,153	0
15	SO4	G	604	5/5	0.92	0.13	-0.90	135,139,141,142	0
15	SO4	G	601	5/5	0.95	0.15	-1.07	123,126,128,132	0
15	SO4	L	301	5/5	0.98	0.11	-	90,90,92,93	5
11	NAG	G	1355	14/15	0.84	0.29	-	121,137,144,146	0
11	NAG	G	1839	14/15	0.87	0.30	-	107,125,132,137	0
15	SO4	G	605	5/5	0.94	0.14	-	147,147,151,157	0
11	NAG	B	1637	14/15	0.91	0.26	-	99,117,130,131	0
15	SO4	G	602	5/5	0.96	0.18	-	107,114,116,121	0
11	NAG	H	523	14/15	0.90	0.24	-	113,124,132,136	0
15	SO4	G	606	5/5	0.45	0.41	-	229,229,230,246	5
15	SO4	B	702	5/5	0.90	0.17	-	154,155,156,156	0
15	SO4	B	701	5/5	0.97	0.15	-	81,82,86,101	0
11	NAG	G	1276	14/15	0.90	0.33	-	113,122,128,128	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.