



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 17, 2016 – 07:49 AM EDT

PDB ID : 5ELF  
Title : Cholera toxin El Tor B-pentamer in complex with A-pentasaccharide  
Authors : Heggelund, J.E.; Burschowsky, D.; Krengel, U.  
Deposited on : 2015-11-04  
Resolution : 1.55 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027257  
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) : rb-20027257

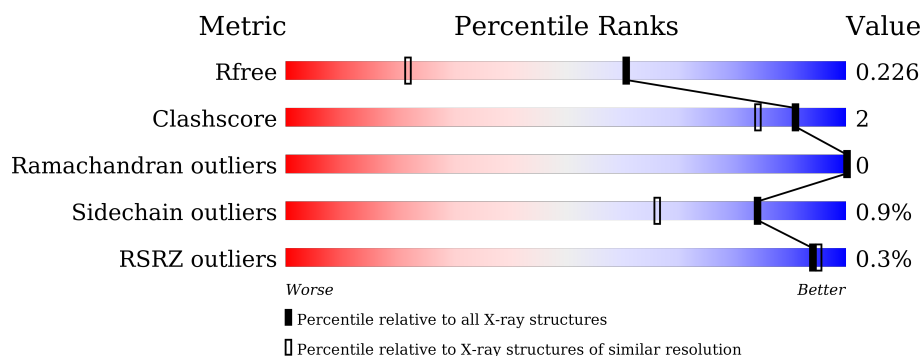
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.55 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.



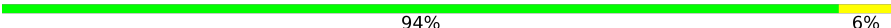


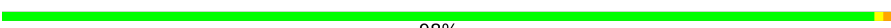
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	1665 (1.58-1.54)
Clashscore	102246	1014 (1.56-1.56)
Ramachandran outliers	100387	1704 (1.58-1.54)
Sidechain outliers	100360	1702 (1.58-1.54)
RSRZ outliers	91569	1668 (1.58-1.54)

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	A	103	<div> <div>%</div> <div>99%</div> <div>.</div> </div>
1	B	103	<div> <div>94%</div> <div>6%</div> </div>
1	C	103	<div> <div>%</div> <div>97%</div> <div>.</div> </div>
1	D	103	<div> <div>93%</div> <div>6%</div> <div>.</div> </div>
1	E	103	<div> <div>%</div> <div>94%</div> <div>6%</div> </div>
1	F	103	<div> <div>99%</div> <div>.</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	103	 94% 6%
1	H	103	 94% 6%
1	I	103	 92% 6% ..
1	J	103	 98% ..

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
4	GLC	G	206[A]	-	-	-	X
6	A2G	E	207	-	-	-	X
6	A2G	G	203[A]	-	-	-	X
6	A2G	G	203[B]	-	-	-	X
7	FUC	B	203	-	-	-	X
7	FUC	E	208	-	-	-	X
7	FUC	E	209	-	-	-	X
7	FUC	F	206	-	-	-	X
7	FUC	H	201	-	-	-	X
8	BGC	D	205[A]	-	-	-	X

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 10185 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 Cholera enterotoxin subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	103	Total	C	N	O	S	0	2	0
			829	525	142	157	5			
1	B	103	Total	C	N	O	S	0	2	0
			828	524	141	158	5			
1	C	103	Total	C	N	O	S	0	3	0
			834	531	142	156	5			
1	D	103	Total	C	N	O	S	0	2	0
			836	528	145	158	5			
1	E	103	Total	C	N	O	S	0	1	0
			823	521	141	156	5			
1	F	103	Total	C	N	O	S	0	1	0
			823	521	141	156	5			
1	G	103	Total	C	N	O	S	0	1	0
			826	522	142	157	5			
1	H	103	Total	C	N	O	S	0	0	0
			817	516	140	156	5			
1	I	103	Total	C	N	O	S	0	0	0
			817	516	140	156	5			
1	J	103	Total	C	N	O	S	0	2	0
			829	526	142	156	5			

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

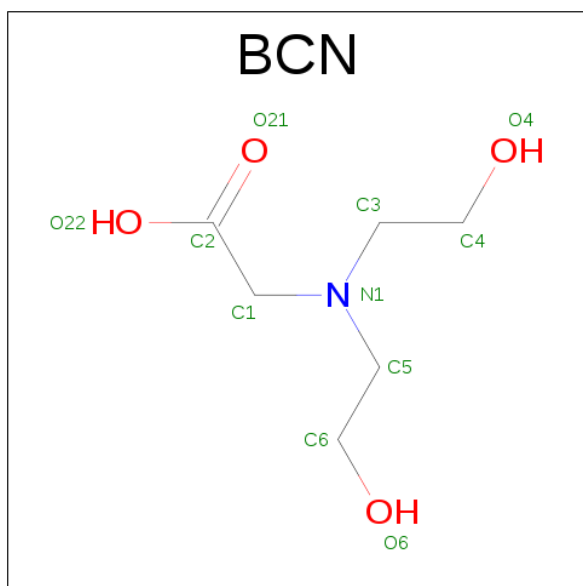
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total	Ca	0	0
			2	2		
2	A	2	Total	Ca	0	0
			2	2		
2	D	2	Total	Ca	0	0
			2	2		
2	C	2	Total	Ca	0	0
			2	2		

*Continued on next page...*

Continued from previous page...

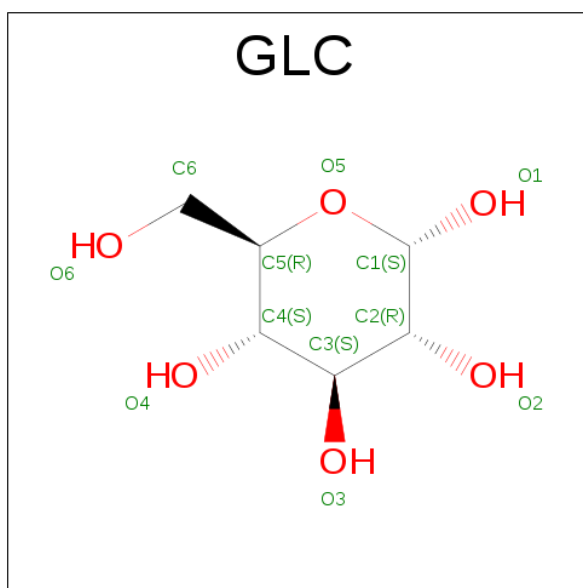
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	E	2	Total	Ca	0	0
			2	2		

- Molecule 3 is BICINE (three-letter code: BCN) (formula:  $C_6H_{13}NO_4$ ).



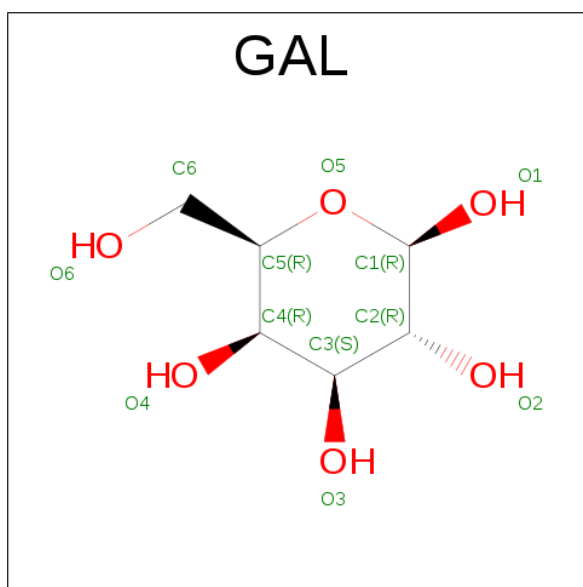
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			11	6	1	4		
3	A	1	Total	C	N	O	0	0
			11	6	1	4		
3	B	1	Total	C	N	O	0	0
			11	6	1	4		
3	B	1	Total	C	N	O	0	0
			11	6	1	4		
3	C	1	Total	C	N	O	0	0
			11	6	1	4		
3	C	1	Total	C	N	O	0	0
			11	6	1	4		
3	D	1	Total	C	N	O	0	0
			11	6	1	4		
3	D	1	Total	C	N	O	0	0
			11	6	1	4		
3	E	1	Total	C	N	O	0	0
			11	6	1	4		
3	E	1	Total	C	N	O	0	0
			11	6	1	4		

- Molecule 4 is ALPHA-D-GLUCOSE (three-letter code: GLC) (formula:  $C_6H_{12}O_6$ ).



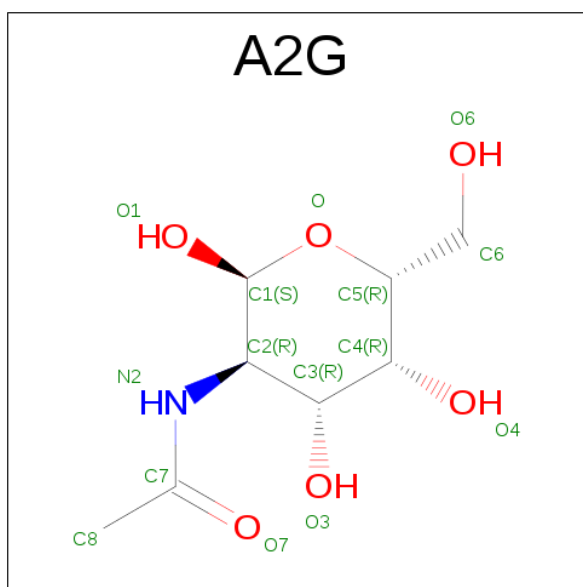
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			12	6	6		
4	D	1	Total	C	O	0	1
			12	6	6		
4	E	1	Total	C	O	0	0
			12	6	6		
4	G	1	Total	C	O	0	1
			12	6	6		
4	I	1	Total	C	O	0	0
			12	6	6		
4	J	1	Total	C	O	0	1
			24	12	12		

- Molecule 5 is BETA-D-GALACTOSE (three-letter code: GAL) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			11	6	5		
5	D	1	Total	C	O	0	1
			22	12	10		
5	E	1	Total	C	O	0	0
			11	6	5		
5	F	1	Total	C	O	0	0
			11	6	5		
5	G	1	Total	C	O	0	1
			22	12	10		
5	I	1	Total	C	O	0	0
			11	6	5		
5	J	1	Total	C	O	0	1
			22	12	10		

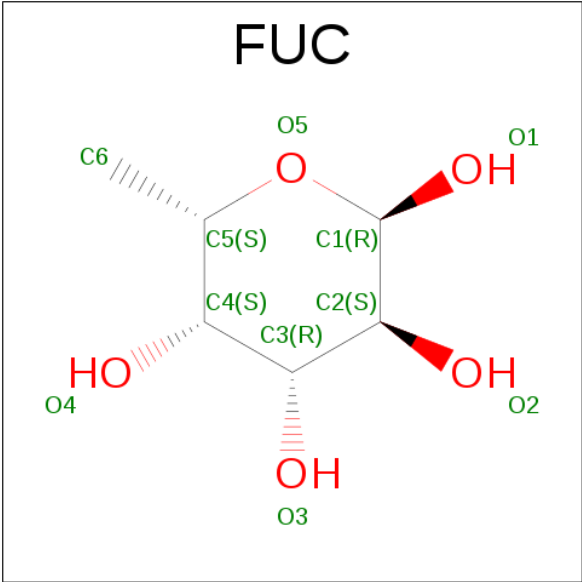
- Molecule 6 is N-ACETYL-2-DEOXY-2-AMINO-GALACTOSE (three-letter code: A2G) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	D	1	Total	C	N	O	0	1
			28	16	2	10		
6	E	1	Total	C	N	O	0	0
			14	8	1	5		
6	F	1	Total	C	N	O	0	0
			14	8	1	5		
6	G	1	Total	C	N	O	0	1
			28	16	2	10		
6	I	1	Total	C	N	O	0	0
			14	8	1	5		
6	J	1	Total	C	N	O	0	1
			28	16	2	10		

- Molecule 7 is ALPHA-L-FUCOSE (three-letter code: FUC) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>5</sub>).





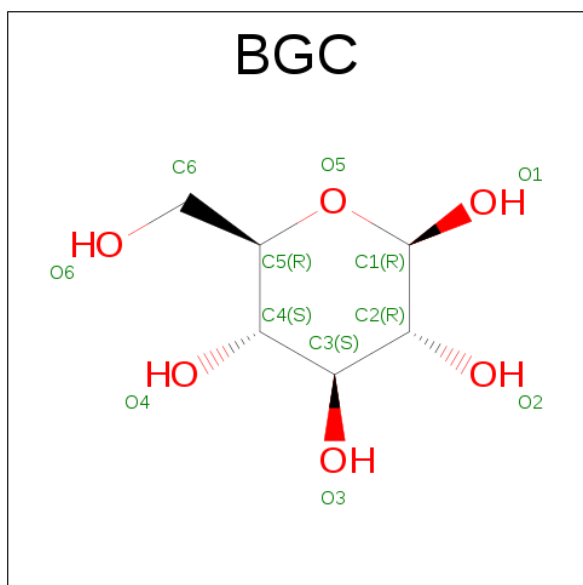
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			10	6	4		
7	A	1	Total	C	O	0	0
			10	6	4		
7	B	1	Total	C	O	0	0
			11	6	5		
7	D	1	Total	C	O	0	1
			20	12	8		
7	D	1	Total	C	O	0	1
			20	12	8		
7	E	1	Total	C	O	0	0
			10	6	4		
7	E	1	Total	C	O	0	0
			10	6	4		
7	F	1	Total	C	O	0	0
			10	6	4		
7	F	1	Total	C	O	0	0
			10	6	4		
7	G	1	Total	C	O	0	1
			20	12	8		
7	G	1	Total	C	O	0	1
			20	12	8		
7	H	1	Total	C	O	0	0
			11	6	5		
7	I	1	Total	C	O	0	0
			10	6	4		
7	I	1	Total	C	O	0	0
			10	6	4		

Continued on next page...

Continued from previous page...

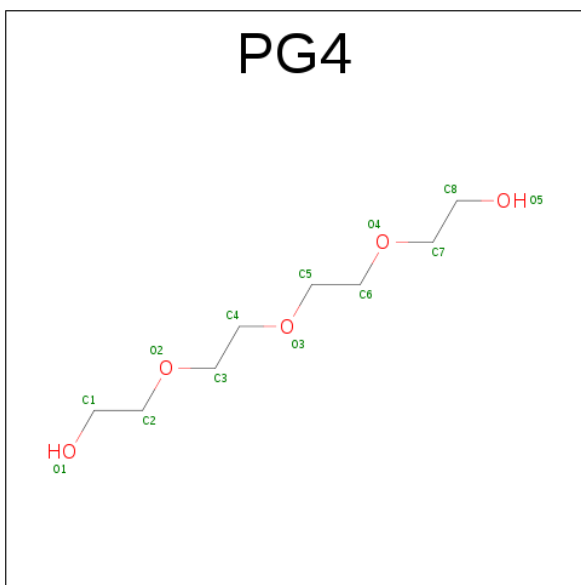
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	J	1	Total	C	O	0	1
			20	12	8		
7	J	1	Total	C	O	0	1
			20	12	8		

- Molecule 8 is BETA-D-GLUCOSE (three-letter code: BGC) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	D	1	Total	C	O	0	1
			12	6	6		
8	F	1	Total	C	O	0	0
			12	6	6		
8	G	1	Total	C	O	0	1
			12	6	6		

- Molecule 9 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula:  $C_8H_{18}O_5$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	F	1	Total	C	O	0	0
			13	8	5		

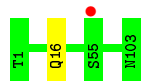
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	106	Total	O	0	0
			106	106		
10	B	102	Total	O	0	1
			103	103		
10	C	111	Total	O	0	0
			111	111		
10	D	128	Total	O	0	0
			128	128		
10	E	126	Total	O	0	0
			126	126		
10	F	123	Total	O	0	2
			125	125		
10	G	131	Total	O	0	1
			132	132		
10	H	115	Total	O	0	0
			115	115		
10	I	127	Total	O	0	0
			127	127		
10	J	125	Total	O	0	0
			125	125		

### 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: Cholera enterotoxin subunit B



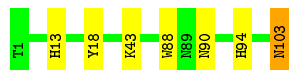
- Molecule 1: Cholera enterotoxin subunit B



- Molecule 1: Cholera enterotoxin subunit B



- Molecule 1: Cholera enterotoxin subunit B

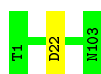


- Molecule 1: Cholera enterotoxin subunit B



- Molecule 1: Cholera enterotoxin subunit B





- Molecule 1: Cholera enterotoxin subunit B

Chain G:  94% 6%




- Molecule 1: Cholera enterotoxin subunit B

Chain H:  94% 6%



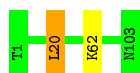
- Molecule 1: Cholera enterotoxin subunit B

Chain I:  92% 6% ..



- Molecule 1: Cholera enterotoxin subunit B

Chain J:  98% ..



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.63Å 84.72Å 195.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.92 – 1.55 45.55 – 1.55	Depositor EDS
% Data completeness (in resolution range)	99.6 (48.92-1.55) 99.6 (45.55-1.55)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.07 (at 1.55Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.179 , 0.220 0.189 , 0.226	Depositor DCC
$R_{free}$ test set	7664 reflections (5.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	15.4	Xtriage
Anisotropy	0.641	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 52.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10185	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.34% 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: BCN, BGC, CA, GLC, PG4, GAL, FUC, A2G

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	0.75	0/849	0.80	0/1145
1	B	0.71	0/848	0.81	3/1144 (0.3%)
1	C	0.75	0/857	0.77	0/1155
1	D	0.75	0/851	0.85	0/1148
1	E	0.79	0/840	0.89	3/1133 (0.3%)
1	F	0.82	0/840	0.85	1/1133 (0.1%)
1	G	0.80	0/840	0.86	1/1133 (0.1%)
1	H	0.83	0/831	0.84	0/1122
1	I	0.80	0/831	0.88	3/1122 (0.3%)
1	J	0.83	0/849	0.86	1/1144 (0.1%)
All	All	0.78	0/8436	0.84	12/11379 (0.1%)

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	35	ARG	NE-CZ-NH1	5.69	123.15	120.30
1	J	20	LEU	CA-CB-CG	-5.58	102.46	115.30
1	E	35	ARG	NE-CZ-NH2	-5.57	117.52	120.30
1	I	35	ARG	NE-CZ-NH1	5.56	123.08	120.30
1	E	35	ARG	NE-CZ-NH1	5.51	123.06	120.30

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	A	829	0	844	1	0
1	B	828	0	840	7	0
1	C	834	0	860	3	0
1	D	836	0	841	5	0
1	E	823	0	836	2	0
1	F	823	0	836	0	0
1	G	826	0	835	4	0
1	H	817	0	823	4	0
1	I	817	0	823	7	0
1	J	829	0	849	1	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
2	E	2	0	0	0	0
3	A	22	0	20	0	0
3	B	22	0	20	0	0
3	C	22	0	20	0	0
3	D	22	0	20	0	0
3	E	22	0	20	0	0
4	A	12	0	10	0	0
4	D	12	0	10	3	0
4	E	12	0	10	0	0
4	G	12	0	10	0	0
4	I	12	0	10	0	0
4	J	24	0	20	2	0
5	A	11	0	8	0	0
5	D	22	0	16	1	0
5	E	11	0	8	1	0
5	F	11	0	8	0	0
5	G	22	0	16	0	0
5	I	11	0	8	0	0
5	J	22	0	16	1	0
6	A	14	0	13	0	0
6	D	28	0	26	2	0
6	E	14	0	13	2	0
6	F	14	0	13	0	0
6	G	28	0	26	1	0
6	I	14	0	13	0	0
6	J	28	0	26	1	0
7	A	20	0	20	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	11	0	12	1	0
7	D	40	0	40	3	0
7	E	20	0	20	0	0
7	F	20	0	20	0	0
7	G	40	0	40	1	0
7	H	11	0	12	0	0
7	I	20	0	20	0	0
7	J	40	0	40	2	0
8	D	12	0	10	0	0
8	F	12	0	10	0	0
8	G	12	0	10	1	0
9	F	13	0	18	0	0
10	A	106	0	0	1	0
10	B	103	0	0	1	0
10	C	111	0	0	1	0
10	D	128	0	0	3	0
10	E	126	0	0	1	0
10	F	125	0	0	0	0
10	G	132	0	0	1	0
10	H	115	0	0	2	0
10	I	127	0	0	2	0
10	J	125	0	0	0	0
All	All	10185	0	9039	39	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 2.

The worst 5 of 39 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:A:16[B]:GLN:OE1	10:A:301:HOH:O	1.56	1.22
6:D:207[A]:A2G:O6	10:D:301:HOH:O	1.53	1.21
1:B:103:ASN:HD21	1:I:23:LYS:CE	2.12	0.61
1:H:35:ARG:NH1	10:H:301:HOH:O	2.33	0.61
1:D:103:ASN:O	1:G:43:LYS:NZ	2.30	0.59

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	A	103/103 (100%)	102 (99%)	1 (1%)	0	100	100
1	B	103/103 (100%)	101 (98%)	2 (2%)	0	100	100
1	C	104/103 (101%)	103 (99%)	1 (1%)	0	100	100
1	D	103/103 (100%)	102 (99%)	1 (1%)	0	100	100
1	E	102/103 (99%)	101 (99%)	1 (1%)	0	100	100
1	F	102/103 (99%)	101 (99%)	1 (1%)	0	100	100
1	G	102/103 (99%)	101 (99%)	1 (1%)	0	100	100
1	H	101/103 (98%)	99 (98%)	2 (2%)	0	100	100
1	I	101/103 (98%)	99 (98%)	2 (2%)	0	100	100
1	J	103/103 (100%)	101 (98%)	2 (2%)	0	100	100
All	All	1024/1030 (99%)	1010 (99%)	14 (1%)	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	91/89 (102%)	91 (100%)	0	100	100
1	B	91/89 (102%)	91 (100%)	0	100	100
1	C	92/89 (103%)	92 (100%)	0	100	100
1	D	91/89 (102%)	89 (98%)	2 (2%)	60	27

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	90/89 (101%)	88 (98%)	2 (2%)	60	27
1	F	90/89 (101%)	90 (100%)	0	100	100
1	G	90/89 (101%)	90 (100%)	0	100	100
1	H	89/89 (100%)	88 (99%)	1 (1%)	80	58
1	I	89/89 (100%)	87 (98%)	2 (2%)	60	27
1	J	91/89 (102%)	90 (99%)	1 (1%)	80	58
All	All	904/890 (102%)	896 (99%)	8 (1%)	84	66

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	29	GLU
1	J	20	LEU
1	I	18	TYR
1	E	4	ASN
1	H	89	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	14	ASN
1	G	103	ASN
1	H	16	GLN
1	F	94	HIS
1	H	4	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 carbohydrates in this entry.

## 5.6 Ligand geometry

Of 73 ligands modelled in this entry, 10 are monoatomic - leaving 63 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$
3	BCN	A	202	2	7,10,10	0.60	0	8,11,11	0.92	0
3	BCN	A	204	2	7,10,10	0.37	0	8,11,11	4.26	4 (50%)
4	GLC	A	205	5,7	12,12,12	1.03	1 (8%)	17,17,17	1.38	3 (17%)
5	GAL	A	206	4,7,6	11,11,12	1.02	2 (18%)	15,15,17	1.49	1 (6%)
6	A2G	A	207	5	14,14,15	1.45	3 (21%)	15,19,21	2.01	4 (26%)
7	FUC	A	208	5	10,10,11	1.16	1 (10%)	13,14,16	0.96	1 (7%)
7	FUC	A	209	4	10,10,11	1.53	2 (20%)	13,14,16	1.37	2 (15%)
3	BCN	B	201	2	7,10,10	1.32	1 (14%)	8,11,11	4.60	4 (50%)
7	FUC	B	203	-	11,11,11	1.09	1 (9%)	16,16,16	1.41	3 (18%)
3	BCN	B	204	2	7,10,10	0.37	0	8,11,11	0.73	0
3	BCN	C	202	2	7,10,10	0.59	0	8,11,11	0.56	0
3	BCN	C	203	2	7,10,10	0.50	0	8,11,11	0.81	0
3	BCN	D	202	2	7,10,10	0.37	0	8,11,11	1.00	1 (12%)
3	BCN	D	203	2	7,10,10	0.50	0	8,11,11	1.15	1 (12%)
8	BGC	D	205[A]	5,7	12,12,12	0.86	1 (8%)	17,17,17	1.40	3 (17%)
5	GAL	D	206[A]	8,7,6	11,11,12	1.70	3 (27%)	15,15,17	1.80	5 (33%)
5	GAL	D	206[B]	4,7,6	11,11,12	1.09	1 (9%)	15,15,17	1.45	2 (13%)
6	A2G	D	207[A]	5	14,14,15	0.84	0	15,19,21	1.39	2 (13%)
6	A2G	D	207[B]	5	14,14,15	1.20	2 (14%)	15,19,21	1.40	3 (20%)
7	FUC	D	208[A]	5	10,10,11	0.87	0	13,14,16	0.91	1 (7%)
7	FUC	D	208[B]	5	10,10,11	0.86	0	13,14,16	1.03	0
7	FUC	D	209[A]	8	10,10,11	1.23	1 (10%)	13,14,16	0.65	0
7	FUC	D	209[B]	4	10,10,11	0.82	0	13,14,16	1.03	0
4	GLC	D	210[B]	5,7	12,12,12	0.83	0	17,17,17	1.25	3 (17%)
3	BCN	E	201	2	7,10,10	0.42	0	8,11,11	3.78	6 (75%)
3	BCN	E	203	2	7,10,10	0.67	0	8,11,11	2.89	5 (62%)
4	GLC	E	205	5,7	12,12,12	0.63	0	17,17,17	1.26	2 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	GAL	E	206	4,7,6	11,11,12	1.06	1 (9%)	15,15,17	1.08	1 (6%)
6	A2G	E	207	5	14,14,15	1.71	3 (21%)	15,19,21	3.24	4 (26%)
7	FUC	E	208	5	10,10,11	1.01	0	13,14,16	1.18	2 (15%)
7	FUC	E	209	4	10,10,11	1.44	1 (10%)	13,14,16	1.39	4 (30%)
9	PG4	F	201	-	12,12,12	0.33	0	11,11,11	0.43	0
8	BGC	F	202	5,7	12,12,12	1.52	3 (25%)	17,17,17	1.33	2 (11%)
5	GAL	F	203	8,7,6	11,11,12	1.35	2 (18%)	15,15,17	1.39	2 (13%)
6	A2G	F	204	5	14,14,15	1.38	1 (7%)	15,19,21	1.48	2 (13%)
7	FUC	F	205	5	10,10,11	0.74	0	13,14,16	0.97	1 (7%)
7	FUC	F	206	8	10,10,11	2.29	5 (50%)	13,14,16	0.61	0
8	BGC	G	201[B]	5,7	12,12,12	0.88	0	17,17,17	1.62	1 (5%)
5	GAL	G	202[A]	4,7,6	11,11,12	1.36	1 (9%)	15,15,17	1.22	0
5	GAL	G	202[B]	8,7,6	11,11,12	1.21	1 (9%)	15,15,17	1.08	1 (6%)
6	A2G	G	203[A]	5	14,14,15	1.45	4 (28%)	15,19,21	1.26	1 (6%)
6	A2G	G	203[B]	5	14,14,15	1.90	1 (7%)	15,19,21	2.58	4 (26%)
7	FUC	G	204[A]	5	10,10,11	1.02	0	13,14,16	1.09	0
7	FUC	G	204[B]	5	10,10,11	0.74	0	13,14,16	1.35	2 (15%)
7	FUC	G	205[A]	4	10,10,11	1.38	2 (20%)	13,14,16	1.19	2 (15%)
7	FUC	G	205[B]	8	10,10,11	1.68	3 (30%)	13,14,16	1.43	2 (15%)
4	GLC	G	206[A]	5,7	12,12,12	1.64	2 (16%)	17,17,17	1.07	1 (5%)
7	FUC	H	201	-	11,11,11	0.87	0	16,16,16	1.35	3 (18%)
4	GLC	I	201	5,7	12,12,12	1.10	2 (16%)	17,17,17	1.33	3 (17%)
5	GAL	I	202	4,7,6	11,11,12	0.95	1 (9%)	15,15,17	1.48	3 (20%)
6	A2G	I	203	5	14,14,15	1.38	2 (14%)	15,19,21	1.43	1 (6%)
7	FUC	I	204	5	10,10,11	0.86	0	13,14,16	1.15	1 (7%)
7	FUC	I	205	4	10,10,11	1.39	2 (20%)	13,14,16	1.12	1 (7%)
4	GLC	J	201[A]	5,7	12,12,12	0.95	1 (8%)	17,17,17	1.14	1 (5%)
4	GLC	J	201[B]	5,7	12,12,12	0.88	0	17,17,17	1.80	3 (17%)
5	GAL	J	202[A]	4,7,6	11,11,12	1.13	2 (18%)	15,15,17	1.55	4 (26%)
5	GAL	J	202[B]	4,7,6	11,11,12	1.12	1 (9%)	15,15,17	0.98	1 (6%)
6	A2G	J	203[A]	5	14,14,15	1.62	3 (21%)	15,19,21	1.47	2 (13%)
6	A2G	J	203[B]	5	14,14,15	1.06	1 (7%)	15,19,21	1.46	2 (13%)
7	FUC	J	204[A]	5	10,10,11	1.21	1 (10%)	13,14,16	1.65	3 (23%)
7	FUC	J	204[B]	5	10,10,11	1.64	1 (10%)	13,14,16	1.50	3 (23%)
7	FUC	J	205[A]	4	10,10,11	1.19	2 (20%)	13,14,16	0.84	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	FUC	J	205[B]	4	10,10,11	0.66	0	13,14,16	0.92	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
3	BCN	A	202	2	-	0/8/10/10	0/0/0/0
3	BCN	A	204	2	-	0/8/10/10	0/0/0/0
4	GLC	A	205	5,7	-	0/2/22/22	0/1/1/1
5	GAL	A	206	4,7,6	-	0/2/19/22	0/1/1/1
6	A2G	A	207	5	-	0/6/23/26	0/1/1/1
7	FUC	A	208	5	-	0/0/17/20	0/1/1/1
7	FUC	A	209	4	-	0/0/17/20	0/1/1/1
3	BCN	B	201	2	-	0/8/10/10	0/0/0/0
7	FUC	B	203	-	-	0/0/20/20	0/1/1/1
3	BCN	B	204	2	-	0/8/10/10	0/0/0/0
3	BCN	C	202	2	-	0/8/10/10	0/0/0/0
3	BCN	C	203	2	-	0/8/10/10	0/0/0/0
3	BCN	D	202	2	-	0/8/10/10	0/0/0/0
3	BCN	D	203	2	-	0/8/10/10	0/0/0/0
8	BGC	D	205[A]	5,7	-	0/2/22/22	0/1/1/1
5	GAL	D	206[A]	8,7,6	-	0/2/19/22	0/1/1/1
5	GAL	D	206[B]	4,7,6	-	0/2/19/22	0/1/1/1
6	A2G	D	207[A]	5	-	0/6/23/26	0/1/1/1
6	A2G	D	207[B]	5	-	0/6/23/26	0/1/1/1
7	FUC	D	208[A]	5	-	0/0/17/20	0/1/1/1
7	FUC	D	208[B]	5	-	0/0/17/20	0/1/1/1
7	FUC	D	209[A]	8	-	0/0/17/20	0/1/1/1
7	FUC	D	209[B]	4	-	0/0/17/20	0/1/1/1
4	GLC	D	210[B]	5,7	-	0/2/22/22	0/1/1/1
3	BCN	E	201	2	-	0/8/10/10	0/0/0/0
3	BCN	E	203	2	-	0/8/10/10	0/0/0/0
4	GLC	E	205	5,7	-	0/2/22/22	0/1/1/1
5	GAL	E	206	4,7,6	-	0/2/19/22	0/1/1/1
6	A2G	E	207	5	-	0/6/23/26	0/1/1/1
7	FUC	E	208	5	-	0/0/17/20	0/1/1/1
7	FUC	E	209	4	-	0/0/17/20	0/1/1/1
9	PG4	F	201	-	-	0/10/10/10	0/0/0/0
8	BGC	F	202	5,7	-	0/2/22/22	0/1/1/1
5	GAL	F	203	8,7,6	-	0/2/19/22	0/1/1/1
6	A2G	F	204	5	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	FUC	F	205	5	-	0/0/17/20	0/1/1/1
7	FUC	F	206	8	-	0/0/17/20	0/1/1/1
8	BGC	G	201[B]	5,7	-	0/2/22/22	0/1/1/1
5	GAL	G	202[A]	4,7,6	-	0/2/19/22	0/1/1/1
5	GAL	G	202[B]	8,7,6	-	0/2/19/22	0/1/1/1
6	A2G	G	203[A]	5	-	0/6/23/26	0/1/1/1
6	A2G	G	203[B]	5	-	0/6/23/26	0/1/1/1
7	FUC	G	204[A]	5	-	0/0/17/20	0/1/1/1
7	FUC	G	204[B]	5	-	0/0/17/20	0/1/1/1
7	FUC	G	205[A]	4	-	0/0/17/20	0/1/1/1
7	FUC	G	205[B]	8	-	0/0/17/20	0/1/1/1
4	GLC	G	206[A]	5,7	-	0/2/22/22	0/1/1/1
7	FUC	H	201	-	-	0/0/20/20	0/1/1/1
4	GLC	I	201	5,7	-	0/2/22/22	0/1/1/1
5	GAL	I	202	4,7,6	-	0/2/19/22	0/1/1/1
6	A2G	I	203	5	-	0/6/23/26	0/1/1/1
7	FUC	I	204	5	-	0/0/17/20	0/1/1/1
7	FUC	I	205	4	-	0/0/17/20	0/1/1/1
4	GLC	J	201[A]	5,7	-	0/2/22/22	0/1/1/1
4	GLC	J	201[B]	5,7	-	0/2/22/22	0/1/1/1
5	GAL	J	202[A]	4,7,6	-	0/2/19/22	0/1/1/1
5	GAL	J	202[B]	4,7,6	-	0/2/19/22	0/1/1/1
6	A2G	J	203[A]	5	-	0/6/23/26	0/1/1/1
6	A2G	J	203[B]	5	-	0/6/23/26	0/1/1/1
7	FUC	J	204[A]	5	-	0/0/17/20	0/1/1/1
7	FUC	J	204[B]	5	-	0/0/17/20	0/1/1/1
7	FUC	J	205[A]	4	-	0/0/17/20	0/1/1/1
7	FUC	J	205[B]	4	-	0/0/17/20	0/1/1/1

The worst 5 of 68 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	J	204[B]	FUC	O5-C1	-4.68	1.36	1.43
7	A	209	FUC	C2-C3	-3.07	1.48	1.52
7	G	205[B]	FUC	C2-C3	-3.00	1.48	1.52
7	G	205[B]	FUC	O5-C1	-2.79	1.39	1.43
5	J	202[B]	GAL	O5-C1	-2.50	1.39	1.43

The worst 5 of 119 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	G	203[B]	A2G	C4-C3-C2	-6.01	102.01	111.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	G	203[B]	A2G	C2-N2-C7	-5.91	115.42	123.11
8	G	201[B]	BGC	C1-O5-C5	-5.79	102.47	113.54
4	J	201[B]	GLC	O5-C5-C6	-4.76	94.06	106.38
5	F	203	GAL	O5-C5-C6	-3.94	98.91	107.34

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

16 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	203	FUC	1	0
5	D	206[B]	GAL	1	0
6	D	207[A]	A2G	1	0
6	D	207[B]	A2G	1	0
7	D	208[B]	FUC	2	0
7	D	209[A]	FUC	1	0
4	D	210[B]	GLC	3	0
5	E	206	GAL	1	0
6	E	207	A2G	2	0
8	G	201[B]	BGC	1	0
6	G	203[B]	A2G	1	0
7	G	205[B]	FUC	1	0
4	J	201[A]	GLC	2	0
5	J	202[A]	GAL	1	0
6	J	203[A]	A2G	1	0
7	J	204[A]	FUC	2	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 ⓘ

### 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	103/103 (100%)	-0.28	1 (0%) 84 86	13, 20, 33, 38	0
1	B	103/103 (100%)	-0.28	0 100 100	13, 23, 35, 51	0
1	C	103/103 (100%)	-0.36	1 (0%) 84 86	12, 18, 30, 44	0
1	D	103/103 (100%)	-0.35	0 100 100	12, 17, 30, 43	0
1	E	103/103 (100%)	-0.32	1 (0%) 84 86	12, 18, 31, 44	0
1	F	103/103 (100%)	-0.55	0 100 100	12, 15, 24, 39	0
1	G	103/103 (100%)	-0.47	0 100 100	12, 16, 29, 45	0
1	H	103/103 (100%)	-0.31	0 100 100	12, 18, 31, 44	0
1	I	103/103 (100%)	-0.41	0 100 100	12, 16, 29, 34	0
1	J	103/103 (100%)	-0.47	0 100 100	11, 15, 25, 43	0
All	All	1030/1030 (100%)	-0.38	3 (0%) 94 95	11, 18, 32, 51	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	55	SER	3.7
1	C	55	SER	2.2
1	A	55	SER	2.0

### 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 ⓘ

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. 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(Å <sup>2</sup> )	Q<0.9
7	FUC	E	208	10/11	0.67	0.21	15.73	37,40,43,43	10
6	A2G	E	207	14/15	0.73	0.23	7.33	34,39,42,43	14
7	FUC	H	201	11/11	0.74	0.18	5.10	31,39,42,44	0
7	FUC	E	209	10/11	0.89	0.11	3.91	22,26,27,28	10
7	FUC	B	203	11/11	0.72	0.17	3.53	39,46,50,52	0
4	GLC	G	206[A]	12/12	0.95	0.09	2.67	14,15,15,17	12
7	FUC	F	206	10/11	0.92	0.09	2.60	18,22,24,25	0
6	A2G	G	203[A]	14/15	0.83	0.18	2.51	26,29,33,38	14
6	A2G	G	203[B]	14/15	0.83	0.18	2.32	19,21,24,27	14
8	BGC	D	205[A]	12/12	0.96	0.09	2.29	14,16,17,18	12
4	GLC	D	210[B]	12/12	0.96	0.08	1.55	15,17,18,19	12
9	PG4	F	201	13/13	0.92	0.10	1.43	29,35,41,45	0
6	A2G	D	207[A]	14/15	0.94	0.09	1.42	18,19,21,21	14
7	FUC	J	205[A]	10/11	0.95	0.08	1.18	14,15,17,17	10
7	FUC	J	205[B]	10/11	0.95	0.08	1.18	13,14,15,16	10
7	FUC	D	208[B]	10/11	0.97	0.07	1.03	14,15,16,16	10
7	FUC	D	208[A]	10/11	0.97	0.07	1.03	16,17,17,18	10
4	GLC	J	201[B]	12/12	0.97	0.09	0.95	15,16,17,17	12
3	BCN	C	202	11/11	0.94	0.12	0.94	17,22,25,26	0
4	GLC	J	201[A]	12/12	0.97	0.09	0.90	19,23,26,31	12
6	A2G	I	203	14/15	0.93	0.09	0.80	24,27,33,35	0
3	BCN	B	201	11/11	0.90	0.10	0.62	19,24,30,32	0
7	FUC	A	209	10/11	0.92	0.09	0.48	23,27,29,30	10
3	BCN	D	202	11/11	0.95	0.11	0.47	16,23,25,27	0
3	BCN	E	203	11/11	0.94	0.10	0.41	17,23,26,27	0
3	BCN	C	203	11/11	0.96	0.10	0.39	16,22,25,26	0
3	BCN	E	201	11/11	0.95	0.11	0.39	17,19,23,24	0
3	BCN	B	204	11/11	0.92	0.09	0.36	20,24,28,29	0
7	FUC	D	209[B]	10/11	0.96	0.07	0.25	13,14,15,15	10
7	FUC	D	209[A]	10/11	0.96	0.07	0.25	11,13,13,14	10
3	BCN	A	204	11/11	0.93	0.11	0.20	18,23,26,28	0
6	A2G	D	207[B]	14/15	0.94	0.09	0.13	17,17,19,19	14
7	FUC	I	205	10/11	0.97	0.06	0.00	14,15,16,17	0
3	BCN	D	203	11/11	0.97	0.10	-0.20	17,22,27,27	0
7	FUC	G	205[B]	10/11	0.98	0.06	-0.27	12,13,13,14	10

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	FUC	G	205[A]	10/11	0.98	0.06	-0.27	11,12,12,13	10
3	BCN	A	202	11/11	0.94	0.08	-0.31	18,21,25,26	0
7	FUC	G	204[B]	10/11	0.98	0.06	-0.74	11,11,12,12	10
7	FUC	G	204[A]	10/11	0.98	0.06	-0.74	14,14,15,15	10
4	GLC	A	205	12/12	0.85	0.15	-	28,36,40,40	12
5	GAL	I	202	11/12	0.96	0.05	-	18,20,22,24	0
2	CA	B	205	1/1	1.00	0.03	-	15,15,15,15	0
7	FUC	A	208	10/11	0.85	0.13	-	35,39,41,45	10
6	A2G	A	207	14/15	0.70	0.23	-	43,47,57,60	14
7	FUC	I	204	10/11	0.93	0.08	-	26,30,35,36	0
4	GLC	I	201	12/12	0.94	0.11	-	16,26,35,36	0
5	GAL	D	206[B]	11/12	0.97	0.07	-	14,14,15,15	11
5	GAL	A	206	11/12	0.86	0.17	-	37,39,40,41	11
2	CA	A	203	1/1	0.99	0.04	-	15,15,15,15	0
5	GAL	D	206[A]	11/12	0.97	0.07	-	16,17,17,19	11
2	CA	B	202	1/1	1.00	0.04	-	15,15,15,15	0
6	A2G	J	203[B]	14/15	0.90	0.14	-	20,21,22,22	14
6	A2G	F	204	14/15	0.83	0.20	-	33,38,42,44	0
5	GAL	F	203	11/12	0.94	0.13	-	23,26,30,30	0
5	GAL	J	202[A]	11/12	0.95	0.08	-	20,21,21,23	11
6	A2G	J	203[A]	14/15	0.90	0.14	-	24,27,29,30	14
7	FUC	F	205	10/11	0.88	0.13	-	26,30,32,33	0
8	BGC	F	202	12/12	0.91	0.14	-	22,27,33,34	0
2	CA	E	202	1/1	1.00	0.05	-	14,14,14,14	0
2	CA	C	201	1/1	1.00	0.05	-	15,15,15,15	0
7	FUC	J	204[A]	10/11	0.89	0.14	-	20,23,23,24	10
8	BGC	G	201[B]	12/12	0.95	0.10	-	13,14,14,17	12
5	GAL	E	206	11/12	0.82	0.28	-	37,40,43,47	11
5	GAL	J	202[B]	11/12	0.95	0.08	-	18,19,20,20	11
2	CA	D	201	1/1	0.99	0.04	-	14,14,14,14	0
5	GAL	G	202[A]	11/12	0.98	0.07	-	16,18,19,21	11
2	CA	A	201	1/1	1.00	0.04	-	13,13,13,13	0
7	FUC	J	204[B]	10/11	0.89	0.14	-	19,21,22,22	10
5	GAL	G	202[B]	11/12	0.98	0.07	-	11,13,13,15	11
2	CA	C	204	1/1	1.00	0.04	-	16,16,16,16	0
2	CA	D	204	1/1	0.99	0.06	-	15,15,15,15	0
4	GLC	E	205	12/12	0.84	0.28	-	31,38,43,44	12
2	CA	E	204	1/1	1.00	0.03	-	14,14,14,14	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.