



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 6, 2024 – 02:25 PM EST

PDB ID : 2GBP  
Title : SUGAR AND SIGNAL-TRANSDUCER BINDING SITES OF THE ES-  
CHERICHIA COLI GALACTOSE CHEMORECEPTOR PROTEIN  
Authors : Vyas, N.K.; Vyas, M.N.; Quioco, F.A.  
Deposited on : 1989-02-23  
Resolution : 1.90 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

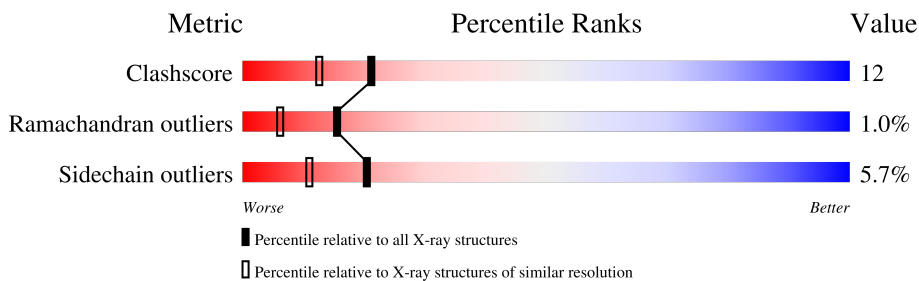
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)

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 of 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	309	

## 2 Entry composition [i](#)

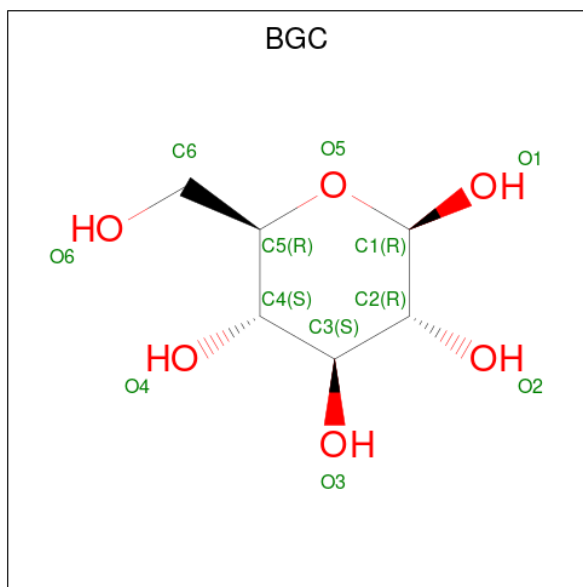
There are 4 unique types of molecules in this entry. The entry contains 2575 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 D-GALACTOSE/D-GLUCOSE BINDING PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	309	2348	1473	406	463	6	0	0	0

- Molecule 2 is beta-D-glucopyranose (three-letter code: BGC) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	12	6	6	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		

- Molecule 4 is water.

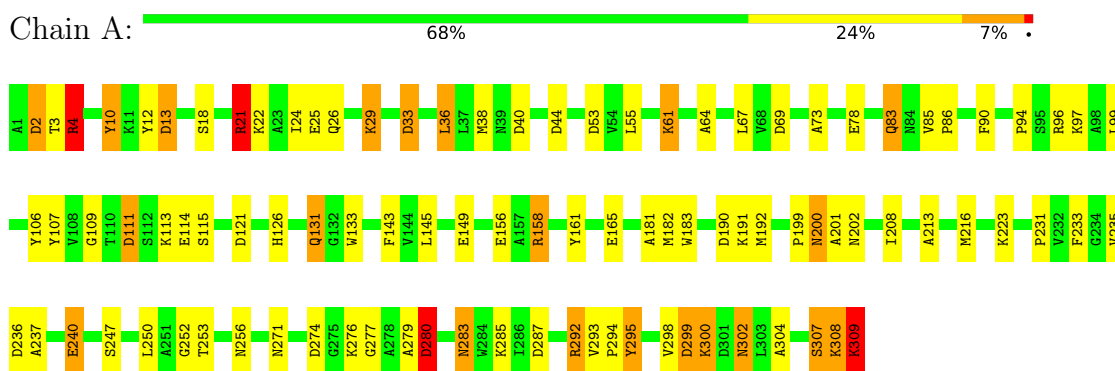
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	A	214	Total 214	O 214	0	0

### 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. 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. 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.

Note EDS was not executed.

- Molecule 1: D-GALACTOSE/D-GLUCOSE BINDING PROTEIN



## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.00Å 37.05Å 61.57Å 90.00° 106.80° 90.00°	Depositor
Resolution (Å)	10.00 – 1.90	Depositor
% Data completeness (in resolution range)	(Not available) (10.00-1.90)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, $R_{free}$	0.146 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	2575	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

## 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: CA, BGC

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	1.25	3/2384 (0.1%)	1.95	76/3229 (2.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	277	GLY	N-CA	8.53	1.58	1.46
1	A	143	PHE	CE2-CZ	5.41	1.47	1.37
1	A	156	GLU	CD-OE1	-5.06	1.20	1.25

All (76) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	158	ARG	NE-CZ-NH2	-13.42	113.59	120.30
1	A	69	ASP	CB-CG-OD1	12.03	129.13	118.30
1	A	280	ASP	CB-CG-OD2	-10.54	108.81	118.30
1	A	13	ASP	CB-CG-OD2	-10.18	109.14	118.30
1	A	165	GLU	OE1-CD-OE2	-9.02	112.47	123.30
1	A	40	ASP	CB-CG-OD1	-8.24	110.89	118.30
1	A	21	ARG	CD-NE-CZ	-8.20	112.12	123.60
1	A	299	ASP	CB-CG-OD2	-8.14	110.97	118.30
1	A	69	ASP	CB-CG-OD2	-8.07	111.04	118.30
1	A	161	TYR	CB-CG-CD1	-8.01	116.19	121.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	202	ASN	CA-CB-CG	-7.80	96.24	113.40
1	A	4	ARG	NE-CZ-NH2	-7.45	116.57	120.30
1	A	78	GLU	OE1-CD-OE2	-7.35	114.48	123.30
1	A	53	ASP	CB-CG-OD1	7.25	124.83	118.30
1	A	158	ARG	NE-CZ-NH1	7.17	123.89	120.30
1	A	292	ARG	CD-NE-CZ	-7.17	113.56	123.60
1	A	131	GLN	CA-CB-CG	7.14	129.10	113.40
1	A	190	ASP	CB-CG-OD1	6.97	124.58	118.30
1	A	44	ASP	CB-CG-OD2	6.91	124.52	118.30
1	A	307	SER	CA-C-N	-6.79	102.27	117.20
1	A	181	ALA	N-CA-CB	6.78	119.59	110.10
1	A	111	ASP	CB-CG-OD1	-6.77	112.21	118.30
1	A	121	ASP	CB-CG-OD1	6.77	124.39	118.30
1	A	276	LYS	CA-C-N	6.74	129.68	116.20
1	A	90	PHE	CB-CG-CD2	6.71	125.49	120.80
1	A	233	PHE	CB-CG-CD1	-6.66	116.14	120.80
1	A	29	LYS	N-CA-CB	6.62	122.53	110.60
1	A	309	LYS	N-CA-CB	6.49	122.27	110.60
1	A	252	GLY	O-C-N	6.44	133.01	122.70
1	A	309	LYS	N-CA-C	-6.39	93.76	111.00
1	A	256	ASN	N-CA-C	-6.37	93.81	111.00
1	A	240	GLU	OE1-CD-OE2	-6.34	115.69	123.30
1	A	29	LYS	CB-CA-C	-6.27	97.86	110.40
1	A	237	ALA	N-CA-CB	6.17	118.74	110.10
1	A	2	ASP	CA-C-O	6.14	132.99	120.10
1	A	33	ASP	CB-CG-OD2	6.13	123.81	118.30
1	A	280	ASP	OD1-CG-OD2	6.12	134.93	123.30
1	A	247	SER	CB-CA-C	-6.12	98.47	110.10
1	A	256	ASN	CB-CA-C	6.12	122.63	110.40
1	A	106	TYR	CB-CG-CD1	-6.11	117.33	121.00
1	A	109	GLY	O-C-N	6.09	132.45	122.70
1	A	307	SER	N-CA-CB	5.97	119.46	110.50
1	A	90	PHE	CB-CG-CD1	-5.97	116.62	120.80
1	A	149	GLU	OE1-CD-OE2	5.92	130.41	123.30
1	A	277	GLY	N-CA-C	-5.86	98.46	113.10
1	A	276	LYS	C-N-CA	-5.80	110.12	122.30
1	A	295	TYR	CB-CG-CD2	-5.80	117.52	121.00
1	A	10	TYR	CG-CD2-CE2	-5.77	116.69	121.30
1	A	4	ARG	CD-NE-CZ	-5.76	115.54	123.60
1	A	21	ARG	NE-CZ-NH1	-5.71	117.44	120.30
1	A	192	MET	CG-SD-CE	-5.65	91.16	100.20
1	A	307	SER	O-C-N	5.60	131.66	122.70

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	53	ASP	CB-CG-OD2	-5.54	113.31	118.30
1	A	240	GLU	CG-CD-OE1	5.47	129.24	118.30
1	A	309	LYS	CA-C-O	-5.44	108.67	120.10
1	A	235	VAL	CA-CB-CG2	5.40	119.00	110.90
1	A	253	THR	CA-CB-CG2	5.35	119.89	112.40
1	A	67	LEU	CB-CG-CD2	-5.30	101.99	111.00
1	A	114	GLU	CG-CD-OE2	-5.30	107.70	118.30
1	A	287	ASP	CB-CG-OD2	-5.30	113.53	118.30
1	A	199	PRO	O-C-N	-5.30	114.23	122.70
1	A	299	ASP	OD1-CG-OD2	5.24	133.26	123.30
1	A	231	PRO	O-C-N	5.24	131.08	122.70
1	A	287	ASP	CB-CA-C	5.23	120.87	110.40
1	A	25	GLU	CG-CD-OE1	5.23	128.76	118.30
1	A	2	ASP	CA-C-N	-5.21	105.73	117.20
1	A	200	ASN	CB-CA-C	-5.21	99.98	110.40
1	A	78	GLU	CG-CD-OE1	5.21	128.71	118.30
1	A	309	LYS	CB-CA-C	5.21	120.81	110.40
1	A	96	ARG	NE-CZ-NH2	5.20	122.90	120.30
1	A	106	TYR	CG-CD2-CE2	-5.11	117.21	121.30
1	A	64	ALA	N-CA-C	-5.06	97.33	111.00
1	A	107	TYR	CB-CG-CD1	-5.06	117.97	121.00
1	A	280	ASP	CA-CB-CG	-5.05	102.29	113.40
1	A	114	GLU	CG-CD-OE1	5.04	128.37	118.30
1	A	12	TYR	CG-CD2-CE2	-5.03	117.28	121.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	158	ARG	Sidechain
1	A	21	ARG	Sidechain
1	A	4	ARG	Sidechain

## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2348	0	2348	57	0
2	A	12	0	12	0	0
3	A	1	0	0	0	0
4	A	214	0	0	15	0
All	All	2575	0	2360	57	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 12.

All (57) 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:85:VAL:O	4:A:507:HOH:O	1.55	1.23
1:A:61:LYS:HG2	4:A:348:HOH:O	1.39	1.18
1:A:85:VAL:C	4:A:507:HOH:O	1.85	1.15
1:A:26:GLN:HE22	1:A:29:LYS:NZ	1.62	0.97
1:A:308:LYS:HA	4:A:526:HOH:O	1.71	0.89
1:A:4:ARG:HB2	1:A:61:LYS:HZ2	1.45	0.82
1:A:21:ARG:HG3	1:A:38:MET:HE1	1.64	0.77
1:A:26:GLN:HE22	1:A:29:LYS:HZ1	1.35	0.73
1:A:26:GLN:HE22	1:A:29:LYS:HZ3	1.42	0.67
1:A:13:ASP:OD2	1:A:13:ASP:N	2.29	0.64
1:A:309:LYS:HB2	4:A:525:HOH:O	1.98	0.62
1:A:4:ARG:HB2	1:A:61:LYS:NZ	2.15	0.60
1:A:21:ARG:HG2	1:A:21:ARG:HH11	1.68	0.59
1:A:283:ASN:ND2	1:A:283:ASN:O	2.37	0.58
1:A:111:ASP:OD1	4:A:509:HOH:O	2.16	0.58
1:A:94:PRO:HG2	1:A:99:LEU:HD13	1.86	0.57
1:A:26:GLN:NE2	1:A:29:LYS:NZ	2.45	0.56
1:A:21:ARG:HG3	1:A:38:MET:CE	2.34	0.55
1:A:4:ARG:H	1:A:61:LYS:HD2	1.72	0.55
1:A:4:ARG:H	1:A:61:LYS:NZ	2.05	0.54
1:A:182:MET:O	1:A:183:TRP:HB2	2.08	0.54
1:A:300:LYS:HE3	1:A:300:LYS:CA	2.39	0.53
1:A:126:HIS:HB3	1:A:133:TRP:CH2	2.45	0.51
1:A:274:ASP:OD2	4:A:488:HOH:O	2.18	0.51
1:A:86:PRO:CA	4:A:507:HOH:O	2.59	0.51
1:A:299:ASP:H	1:A:302:ASN:HD21	1.59	0.51
1:A:2:ASP:HA	1:A:33:ASP:O	2.10	0.51
1:A:21:ARG:NE	4:A:494:HOH:O	2.28	0.51
1:A:86:PRO:N	4:A:507:HOH:O	2.26	0.51
1:A:3:THR:HA	1:A:61:LYS:HD3	1.94	0.50

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:4:ARG:H	1:A:61:LYS:CD	2.25	0.50
1:A:145:LEU:HA	1:A:208:ILE:O	2.12	0.49
1:A:298:VAL:HA	1:A:302:ASN:HD21	1.79	0.48
1:A:300:LYS:HE3	1:A:300:LYS:HA	1.95	0.48
1:A:83:GLN:NE2	4:A:328:HOH:O	2.47	0.48
1:A:285:LYS:O	1:A:285:LYS:HG3	2.13	0.47
1:A:86:PRO:HA	4:A:507:HOH:O	2.14	0.46
1:A:299:ASP:H	1:A:302:ASN:ND2	2.11	0.46
1:A:191:LYS:HE3	4:A:515:HOH:O	2.15	0.46
1:A:308:LYS:H	1:A:308:LYS:HG3	1.31	0.45
1:A:131:GLN:HB3	4:A:510:HOH:O	2.17	0.44
1:A:26:GLN:NE2	1:A:29:LYS:HZ3	2.09	0.44
1:A:21:ARG:HH11	1:A:21:ARG:CG	2.27	0.43
1:A:73:ALA:HB2	1:A:94:PRO:HB3	2.00	0.43
1:A:293:VAL:HA	1:A:294:PRO:HD3	1.88	0.43
1:A:200:ASN:O	1:A:201:ALA:C	2.56	0.43
1:A:223:LYS:O	4:A:408:HOH:O	2.20	0.43
1:A:279:ALA:O	1:A:280:ASP:C	2.55	0.42
1:A:18:SER:O	1:A:22:LYS:HG3	2.19	0.42
1:A:24:ILE:HG22	1:A:36:LEU:HD21	2.02	0.41
1:A:302:ASN:H	1:A:302:ASN:HD22	1.68	0.41
1:A:213:ALA:HA	1:A:216:MET:HE2	2.03	0.41
1:A:292:ARG:HH11	1:A:292:ARG:HD3	1.56	0.41
1:A:115:SER:HB3	1:A:295:TYR:CZ	2.56	0.41
1:A:250:LEU:HD12	1:A:250:LEU:HA	1.86	0.41
1:A:271:ASN:HD22	1:A:271:ASN:HA	1.69	0.41
1:A:4:ARG:CB	1:A:61:LYS:HZ2	2.24	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	307/309 (99%)	297 (97%)	7 (2%)	3 (1%)	15 6

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	236	ASP
1	A	308	LYS
1	A	304	ALA

### 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	245/245 (100%)	231 (94%)	14 (6%)	20 11

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	TYR
1	A	36	LEU
1	A	55	LEU
1	A	61	LYS
1	A	83	GLN
1	A	97	LYS
1	A	113	LYS
1	A	240	GLU
1	A	280	ASP
1	A	283	ASN
1	A	300	LYS
1	A	302	ASN
1	A	307	SER
1	A	309	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	26	GLN
1	A	43	ASN
1	A	83	GLN
1	A	131	GLN
1	A	200	ASN
1	A	260	ASN
1	A	271	ASN
1	A	283	ASN
1	A	302	ASN

### 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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	BGC	A	310	-	12,12,12	1.27	2 (16%)	17,17,17	2.01	5 (29%)

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	BGC	A	310	-	-	0/2/22/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	310	BGC	O2-C2	2.41	1.48	1.43
2	A	310	BGC	O5-C1	2.19	1.48	1.42

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	310	BGC	O5-C1-C2	-6.23	99.17	110.28
2	A	310	BGC	O2-C2-C1	-2.53	103.29	109.16
2	A	310	BGC	O1-C1-O5	-2.35	103.32	110.38
2	A	310	BGC	C1-O5-C5	2.18	117.78	113.66
2	A	310	BGC	C4-C3-C2	-2.11	107.14	110.82

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.