



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 30, 2021 – 05:09 am GMT

PDB ID : 6Z3I  
Title : Structure of recombinant beta-glucocerebrosidase in complex with bifunctional cyclophellitol aziridine activity based probe  
Authors : Rowland, R.J.; Davies, G.J.  
Deposited on : 2020-05-20  
Resolution : 1.80 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4 (270009), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

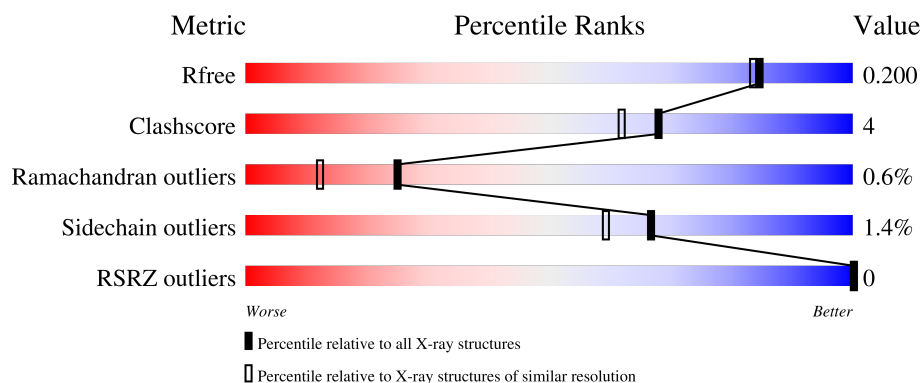
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*



The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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  $\leq 5\%$ . 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	BBB	497	 91% 9%
2	A	4	 75% 25%

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
5	EDO	BBB	524	-	-	X	-

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 8613 atoms, of which 4108 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 Lysosomal acid glucosylceramidase.

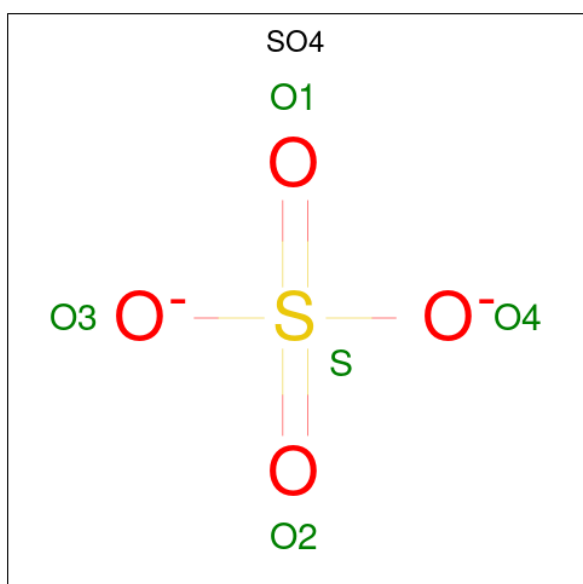
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	BBB	497	Total	C	H	N	O	S	199	6	0
			7869	2558	3897	678	720	16			

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-6)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



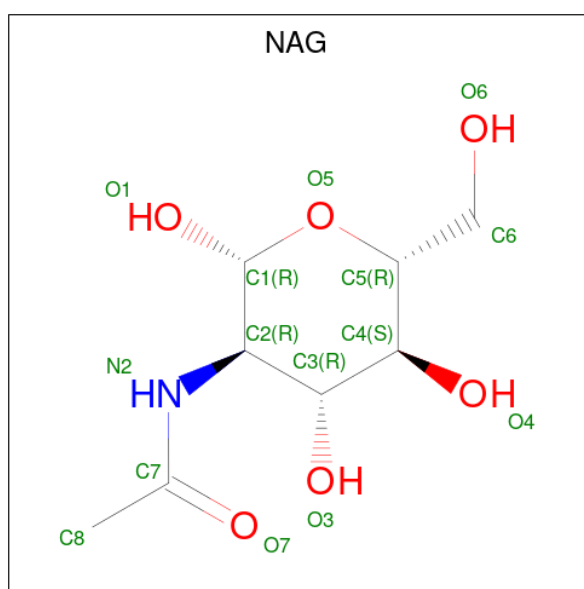
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	4	Total	C	H	N	O	11	0	0
			100	28	50	2	20			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	BBB	1	Total	O	S	0	0
			5	4	1		
3	BBB	1	Total	O	S	0	0
			5	4	1		
3	BBB	1	Total	O	S	0	0
			5	4	1		
3	BBB	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	BBB	1	Total	C	H	N	O	3	0
			28	8	14	1	5		
4	BBB	1	Total	C	H	N	O	3	0
			28	8	14	1	5		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



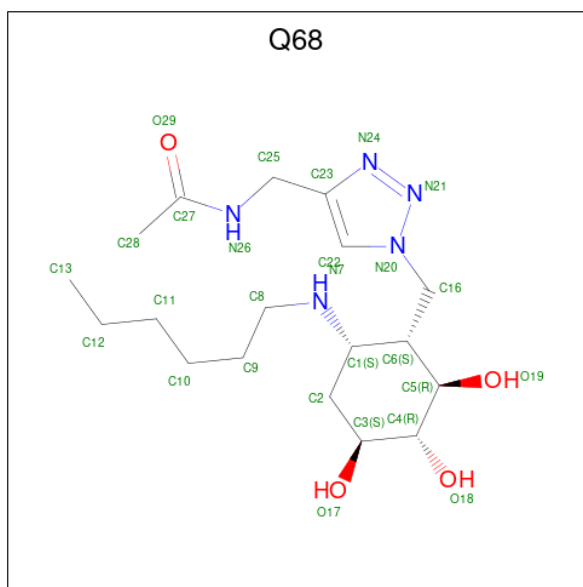
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
5	BBB	1	Total	C	H	O	2	1
			20	4	12	4		
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
5	BBB	1	Total	C	H	O	1	0
			10	2	6	2		

- Molecule 6 is {N}-[[1-[[1 {S},2 {R},3 {R},4 {S},6 {S}]-6-(hexylamino)-2,3,4-tris(oxidanyl)cyclohexyl)methyl]-1,2,3-triazol-4-yl)methyl]ethanamide (three-letter code: Q68) (formula:  $C_{18}H_{33}N_5O_4$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	BBB	1	Total	C	H	N	O	
			58	18	31	5	4	0

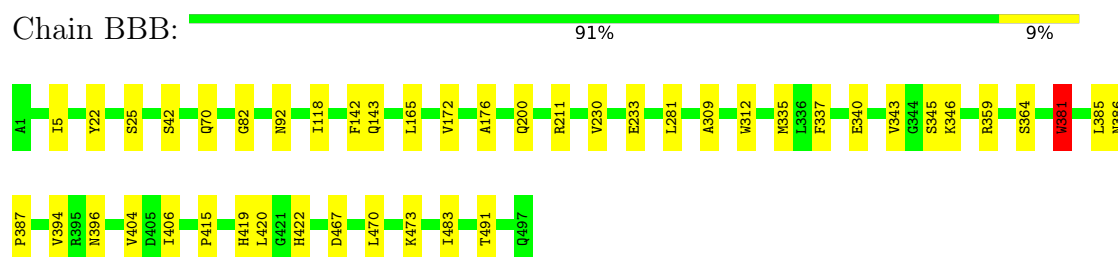
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	BBB	338	Total	O	0	2
			340	340		

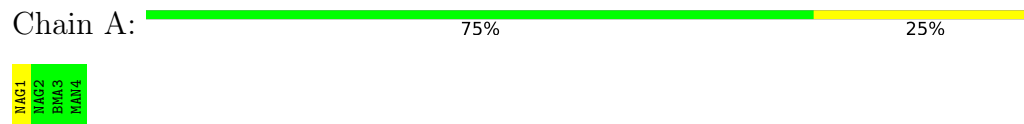
### 3 Residue-property plots

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: Lysosomal acid glucosylceramidase



- Molecule 2: alpha-D-mannopyranose-(1-6)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	53.14Å 76.67Å 67.99Å 90.00° 102.05° 90.00°	Depositor
Resolution (Å)	66.58 – 1.80 66.49 – 1.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (66.58-1.80) 100.0 (66.49-1.80)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.72 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.178 , 0.189 0.189 , 0.200	Depositor DCC
$R_{free}$ test set	2509 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.6	Xtriage
Anisotropy	0.168	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8613	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.18% 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: EDO, NAG, BMA, MAN, SO4, Q68

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	BBB	0.68	1/4091 (0.0%)	0.74	0/5578

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	BBB	340	GLU	CD-OE2	17.36	1.44	1.25

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	BBB	3972	3897	3866	30	1
2	A	50	50	43	0	1
3	BBB	20	0	0	0	0
4	BBB	28	28	26	0	0
5	BBB	68	102	102	7	0
6	BBB	27	31	0	5	0
7	BBB	340	0	0	4	0
All	All	4505	4108	4037	31	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including

hydrogen atoms). The all-atom clashscore for this structure is 4.

All (31) 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:BBB:396:ASN:ND2	6:BBB:520:Q68:N24	2.34	0.75
1:BBB:396:ASN:CG	6:BBB:520:Q68:N24	2.47	0.68
1:BBB:165:LEU:O	7:BBB:601[A]:HOH:O	2.11	0.68
1:BBB:396:ASN:OD1	6:BBB:520:Q68:N24	2.26	0.68
1:BBB:396:ASN:HD21	6:BBB:520:Q68:C23	2.09	0.66
1:BBB:70:GLN:C	5:BBB:524:EDO:H22	2.20	0.62
1:BBB:70:GLN:CA	5:BBB:524:EDO:H22	2.32	0.58
1:BBB:394:VAL:HG23	1:BBB:396:ASN:HB2	1.85	0.58
1:BBB:467:ASP:HB3	1:BBB:483:ILE:HD11	1.86	0.58
1:BBB:142:PHE:O	1:BBB:211:ARG:NH2	2.37	0.57
5:BBB:515:EDO:H12	7:BBB:716:HOH:O	2.04	0.56
1:BBB:25:SER:HA	5:BBB:509:EDO:H22	1.90	0.53
1:BBB:346:LYS:O	5:BBB:515:EDO:O1	2.27	0.52
1:BBB:404:VAL:HG12	1:BBB:406[A]:ILE:HG13	1.91	0.52
1:BBB:364:SER:OG	1:BBB:419:HIS:HD2	1.96	0.49
1:BBB:70:GLN:HA	5:BBB:524:EDO:C2	2.43	0.49
1:BBB:200:GLN:NE2	7:BBB:613:HOH:O	2.47	0.48
1:BBB:42:SER:OG	1:BBB:422:HIS:HE1	1.98	0.46
1:BBB:143:GLN:NE2	7:BBB:603:HOH:O	2.39	0.46
1:BBB:70:GLN:HA	5:BBB:524:EDO:H21	1.98	0.46
1:BBB:309:ALA:HA	1:BBB:337:PHE:O	2.17	0.45
1:BBB:176:ALA:HB3	1:BBB:230:VAL:HG12	1.99	0.44
1:BBB:422:HIS:HD2	1:BBB:491:THR:OG1	2.01	0.43
1:BBB:92:ASN:ND2	1:BBB:385:LEU:HA	2.34	0.43
1:BBB:82:GLY:HA3	1:BBB:118:ILE:O	2.19	0.43
1:BBB:381:TRP:HA	1:BBB:381:TRP:CE3	2.55	0.41
1:BBB:5:ILE:HG12	1:BBB:22:TYR:CE2	2.55	0.41
1:BBB:165:LEU:CD2	1:BBB:172:VAL:HB	2.50	0.41
1:BBB:345:SER:HB2	6:BBB:520:Q68:C22	2.50	0.41
1:BBB:343:VAL:HG11	1:BBB:359:ARG:HG2	2.01	0.40
1:BBB:386:ASN:HB2	1:BBB:387:PRO:CD	2.51	0.40

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:BBB:473:LYS:HZ3	2:A:1:NAG:HO3[1_455]	1.02	0.58

## 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	BBB	500/497 (101%)	482 (96%)	15 (3%)	3 (1%)	25 12

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	BBB	233	GLU
1	BBB	281	LEU
1	BBB	381	TRP

### 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	BBB	426/424 (100%)	420 (99%)	6 (1%)	67 59

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	312	TRP
1	BBB	335	MET
1	BBB	381	TRP
1	BBB	415	PRO
1	BBB	420	LEU
1	BBB	470	LEU

Sometimes 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 ⓘ

4 monosaccharides are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

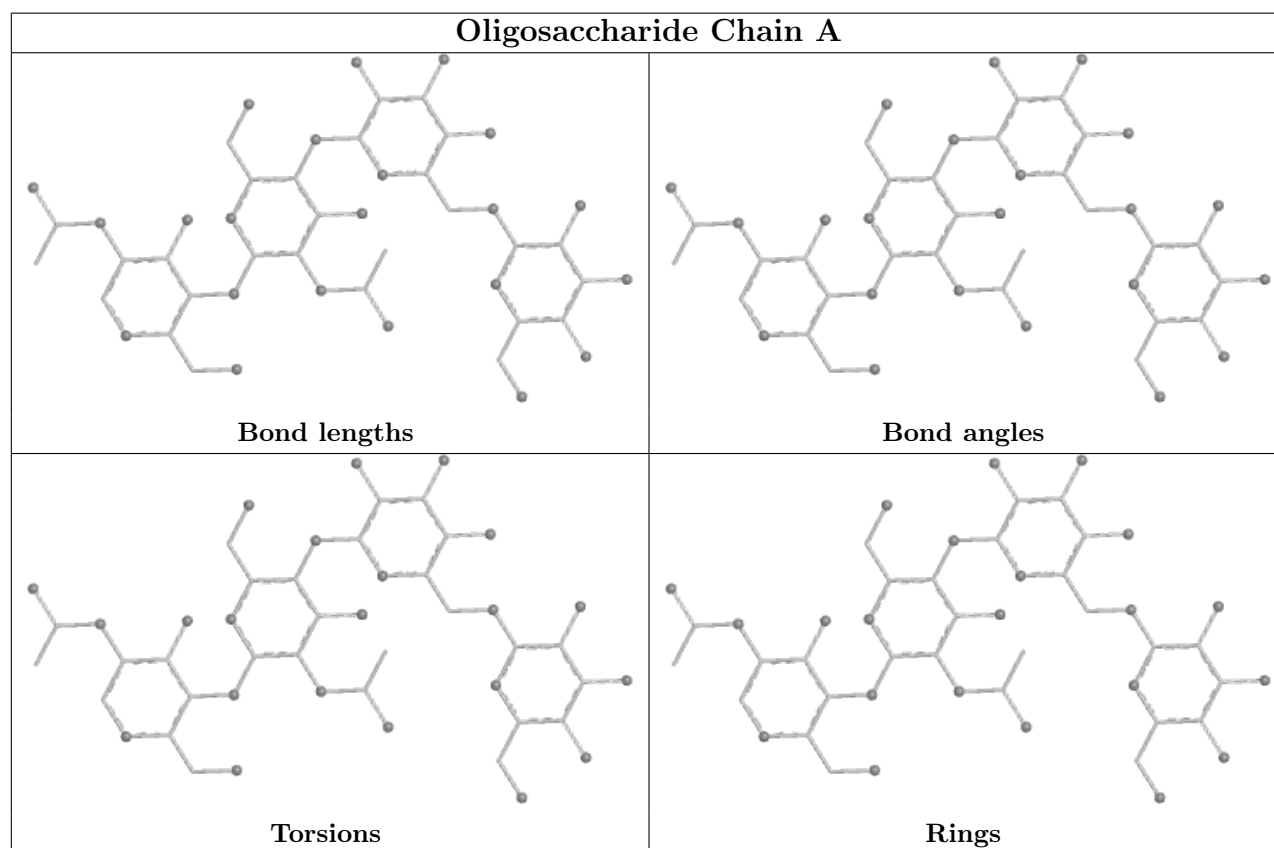
There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.



## 5.6 Ligand geometry

24 ligands are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

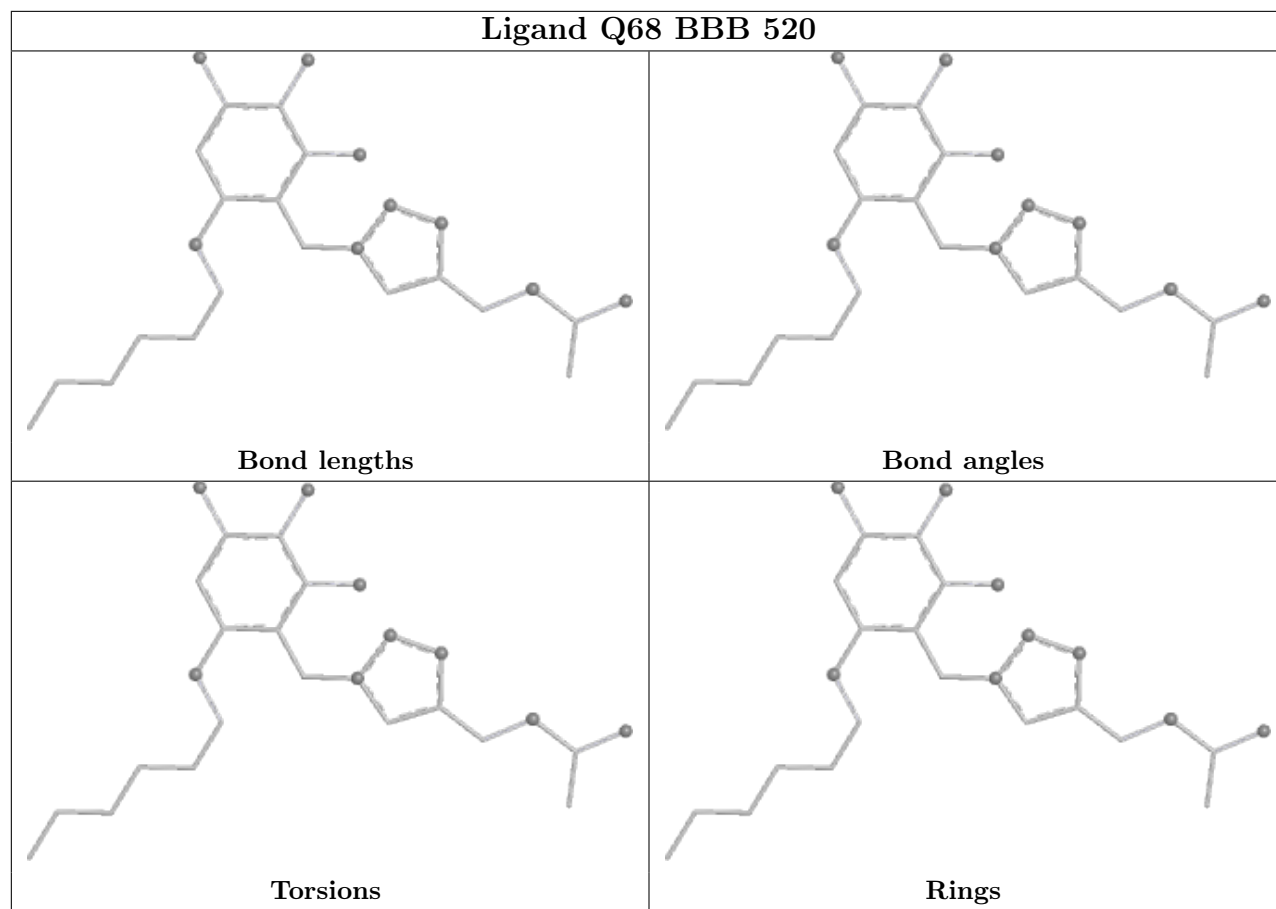
There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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	BBB	497/497 (100%)	-0.52	0 100 100	19, 25, 35, 45	3 (0%)

There are no RSRZ outliers to report.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

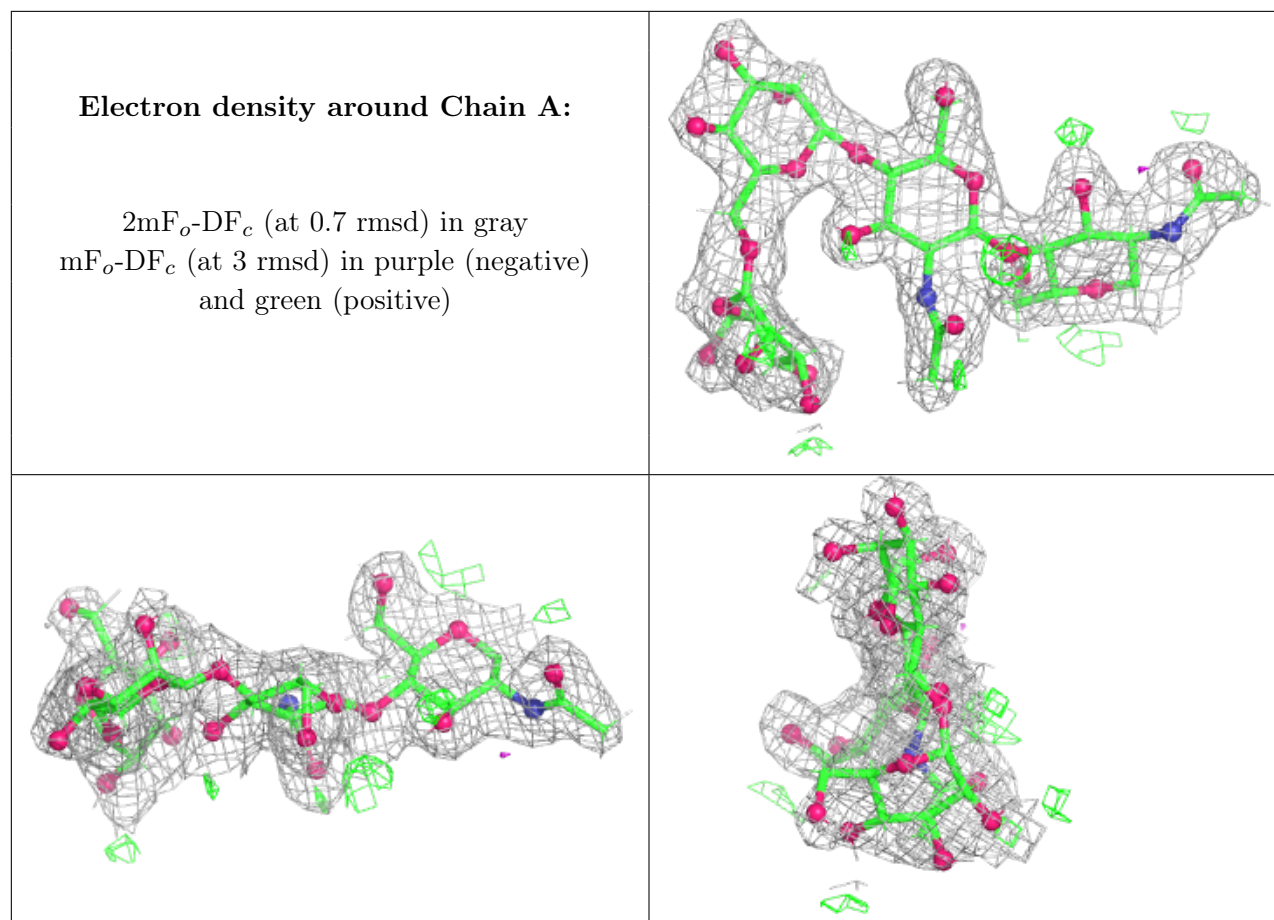
### 6.3 Carbohydrates [i](#)

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. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	MAN	A	4	11/12	0.83	0.15	75,77,77,77	4
2	BMA	A	3	11/12	0.85	0.10	61,64,67,69	3
2	NAG	A	1	14/15	0.90	0.11	34,38,43,44	2
2	NAG	A	2	14/15	0.90	0.09	40,47,52,55	2

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.





## 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.

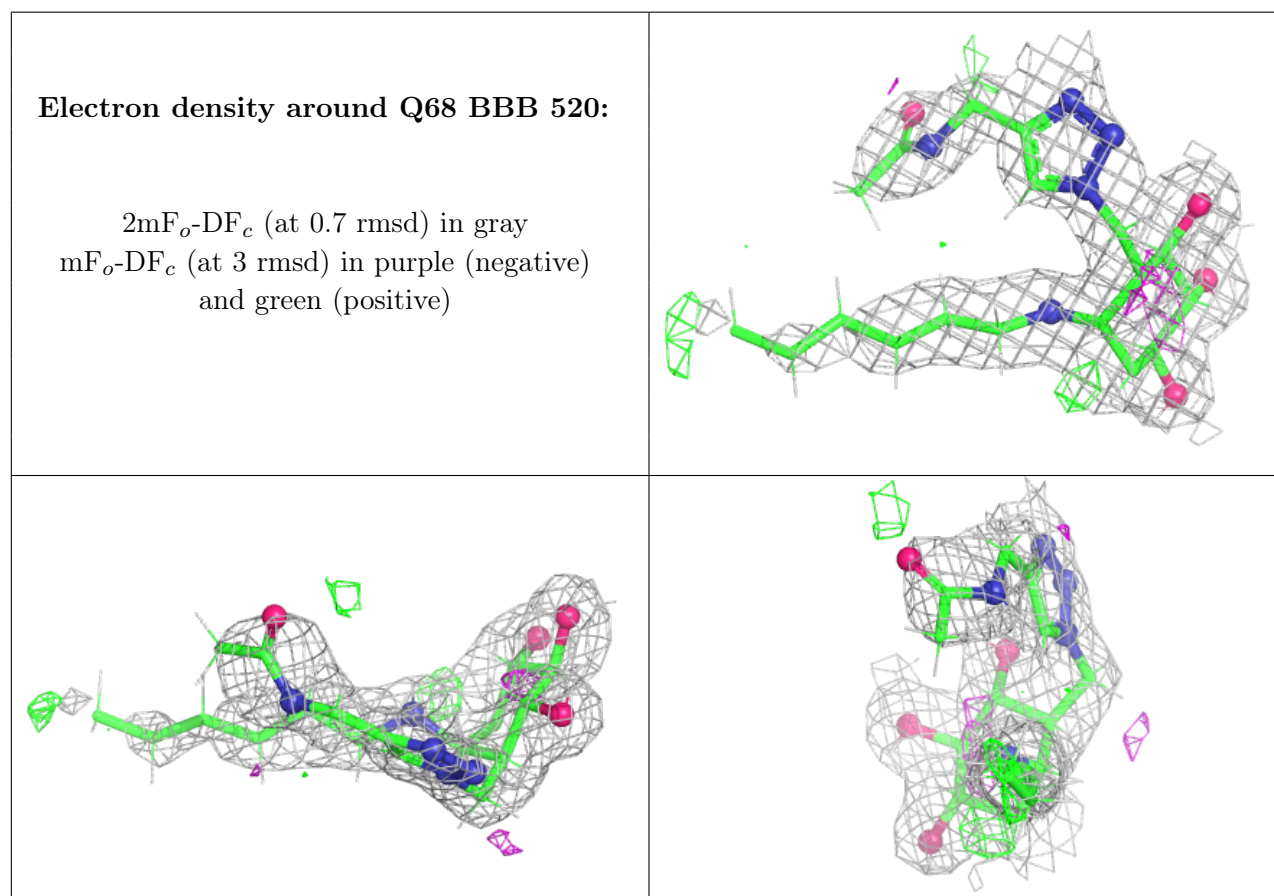
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	EDO	BBB	526	4/4	0.67	0.18	46,49,49,50	1
4	NAG	BBB	519	14/15	0.75	0.16	69,73,79,83	3
4	NAG	BBB	502	14/15	0.77	0.15	42,45,49,49	28
5	EDO	BBB	527	4/4	0.81	0.13	50,54,55,56	1
5	EDO	BBB	516	4/4	0.83	0.15	51,52,52,52	1
5	EDO	BBB	524	4/4	0.83	0.21	35,38,41,43	1
5	EDO	BBB	518	4/4	0.85	0.23	42,47,49,49	1
5	EDO	BBB	513	4/4	0.85	0.14	45,45,48,49	1
5	EDO	BBB	523	4/4	0.88	0.13	40,46,52,53	1
5	EDO	BBB	522	4/4	0.88	0.13	42,44,46,48	1
5	EDO	BBB	512	4/4	0.89	0.09	47,49,51,51	1
5	EDO	BBB	525	4/4	0.91	0.16	34,39,45,47	1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	Q68	BBB	520	27/27	0.91	0.18	23,38,57,61	10
5	EDO	BBB	521[B]	4/4	0.92	0.12	33,36,36,36	10
5	EDO	BBB	517	4/4	0.92	0.17	37,37,38,39	1
5	EDO	BBB	515	4/4	0.92	0.13	32,33,33,34	1
5	EDO	BBB	521[A]	4/4	0.92	0.12	42,44,44,44	10
5	EDO	BBB	509	4/4	0.93	0.15	33,34,37,37	1
3	SO4	BBB	511	5/5	0.94	0.11	67,67,69,70	0
3	SO4	BBB	508	5/5	0.96	0.18	35,35,37,38	5
5	EDO	BBB	510	4/4	0.97	0.07	30,31,31,32	1
5	EDO	BBB	514	4/4	0.98	0.08	25,26,26,26	1
3	SO4	BBB	501	5/5	0.99	0.10	33,33,37,38	0
3	SO4	BBB	503	5/5	0.99	0.09	39,41,41,42	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.