



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 08:24 PM BST

PDB ID : 3PBL  
Title : Structure of the human dopamine D3 receptor in complex with eticlopride  
Authors : Chien, E.Y.T.; Liu, W.; Han, G.W.; Katritch, V.; Zhao, Q.; Cherezov, V.; Stevens, R.C.; Accelerated Technologies Center for Gene to 3D Structure (ATCG3D); GPCR Network (GPCR)  
Deposited on : 2010-10-20  
Resolution : 2.89 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

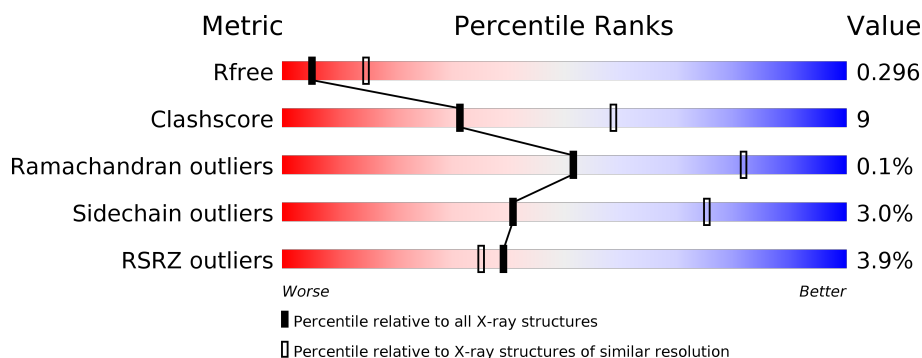
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.89 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	481	<div> <div>3%</div> <div> <div></div> <div>71%</div> <div>17%</div> <div>•</div> <div>10%</div> </div> </div>
1	B	481	<div> <div>4%</div> <div> <div></div> <div>70%</div> <div>16%</div> <div>•</div> <div>12%</div> </div> </div>
2	C	2	<div> <div></div> <div>100%</div> </div>
2	D	2	<div> <div></div> <div> <div>50%</div> <div>50%</div> </div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GLC	C	1	X	-	-	-
2	GLC	D	1	X	-	-	-

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 6787 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(3) dopamine receptor, Lysozyme chimera.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	432	Total	C	N	O	S	0	0	0
			3389	2197	575	593	24			
1	B	423	Total	C	N	O	S	0	0	0
			3306	2149	552	581	24			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	ASP	-	expression tag	UNP P35462
A	-7	TYR	-	expression tag	UNP P35462
A	-6	LYS	-	expression tag	UNP P35462
A	-5	ASP	-	expression tag	UNP P35462
A	-4	ASP	-	expression tag	UNP P35462
A	-3	ASP	-	expression tag	UNP P35462
A	-2	ASP	-	expression tag	UNP P35462
A	-1	GLY	-	expression tag	UNP P35462
A	0	ALA	-	expression tag	UNP P35462
A	1	PRO	-	expression tag	UNP P35462
A	119	TRP	LEU	engineered mutation	UNP P35462
A	1054	THR	CYS	engineered mutation	UNP P00720
A	1097	ALA	CYS	engineered mutation	UNP P00720
A	401	GLY	-	expression tag	UNP P35462
A	402	ARG	-	expression tag	UNP P35462
A	403	PRO	-	expression tag	UNP P35462
A	404	LEU	-	expression tag	UNP P35462
A	405	GLU	-	expression tag	UNP P35462
A	406	VAL	-	expression tag	UNP P35462
A	407	LEU	-	expression tag	UNP P35462
A	408	PHE	-	expression tag	UNP P35462
A	409	GLN	-	expression tag	UNP P35462
B	-8	ASP	-	expression tag	UNP P35462
B	-7	TYR	-	expression tag	UNP P35462
B	-6	LYS	-	expression tag	UNP P35462

*Continued on next page...*

Continued from previous page...

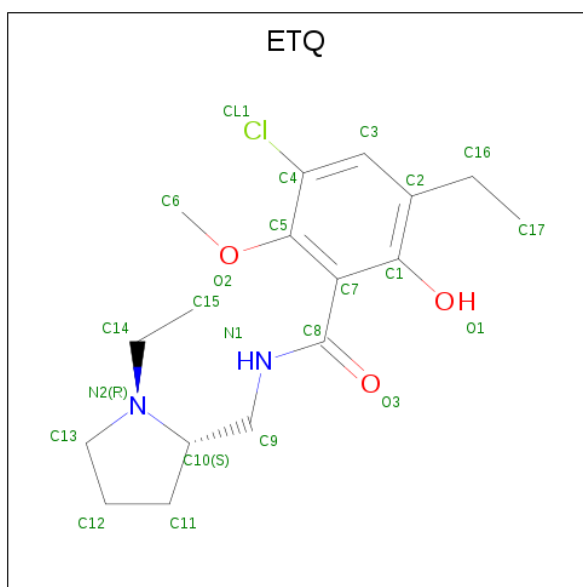
Chain	Residue	Modelled	Actual	Comment	Reference
B	-5	ASP	-	expression tag	UNP P35462
B	-4	ASP	-	expression tag	UNP P35462
B	-3	ASP	-	expression tag	UNP P35462
B	-2	ASP	-	expression tag	UNP P35462
B	-1	GLY	-	expression tag	UNP P35462
B	0	ALA	-	expression tag	UNP P35462
B	1	PRO	-	expression tag	UNP P35462
B	119	TRP	LEU	engineered mutation	UNP P35462
B	1054	THR	CYS	engineered mutation	UNP P00720
B	1097	ALA	CYS	engineered mutation	UNP P00720
B	401	GLY	-	expression tag	UNP P35462
B	402	ARG	-	expression tag	UNP P35462
B	403	PRO	-	expression tag	UNP P35462
B	404	LEU	-	expression tag	UNP P35462
B	405	GLU	-	expression tag	UNP P35462
B	406	VAL	-	expression tag	UNP P35462
B	407	LEU	-	expression tag	UNP P35462
B	408	PHE	-	expression tag	UNP P35462
B	409	GLN	-	expression tag	UNP P35462

- Molecule 2 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	C	2	Total	C	O	0	0	0
			23	12	11			
2	D	2	Total	C	O	0	0	0
			23	12	11			

- Molecule 3 is 3-chloro-5-ethyl-N-{[(2S)-1-ethylpyrrolidin-2-yl]methyl}-6-hydroxy-2-methoxy benzamide (three-letter code: ETQ) (formula: C<sub>17</sub>H<sub>25</sub>ClN<sub>2</sub>O<sub>3</sub>).

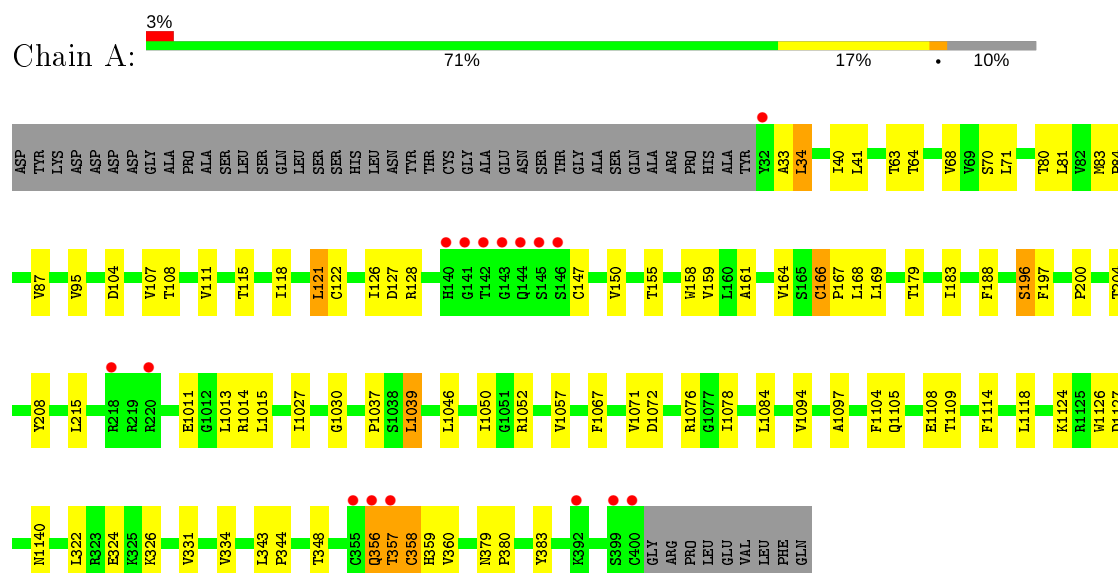


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Cl	N	O		
3	A	1	23	17	1	2	3	0	0
3	B	1	23	17	1	2	3	0	0

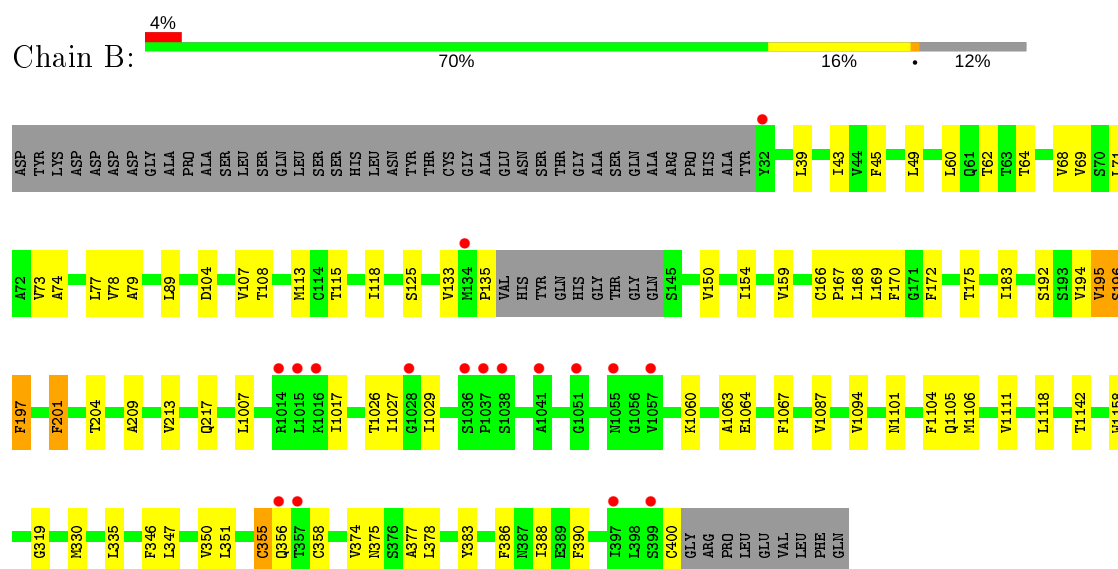
### 3 Residue-property plots

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: D(3) dopamine receptor, Lysozyme chimera



- Molecule 1: D(3) dopamine receptor, Lysozyme chimera



- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain C:  100%

GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain D:  50%  50%

GLC1  
GLC2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.83Å 92.49Å 176.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.45 – 2.89 36.45 – 2.89	Depositor EDS
% Data completeness (in resolution range)	(Not available) (36.45-2.89) 77.8 (36.45-2.89)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.17 (at 2.90Å)	Xtriage
Refinement program	BUSTER 2.8.0	Depositor
R, $R_{free}$	0.243 , 0.272 0.257 , 0.296	Depositor DCC
$R_{free}$ test set	1318 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	59.2	Xtriage
Anisotropy	0.145	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 67.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.028 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	6787	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.55% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GLC, ETQ

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	0.45	0/3464	0.66	6/4718 (0.1%)
1	B	0.41	0/3377	0.60	3/4600 (0.1%)
All	All	0.43	0/6841	0.63	9/9318 (0.1%)

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	358	CYS	N-CA-CB	-13.93	85.52	110.60
1	B	197	PHE	N-CA-CB	-11.83	89.31	110.60
1	A	197	PHE	N-CA-CB	-9.45	93.58	110.60
1	B	196	SER	CB-CA-C	9.41	127.99	110.10
1	A	196	SER	CB-CA-C	8.06	125.41	110.10
1	B	197	PHE	N-CA-C	7.05	130.03	111.00
1	A	357	THR	N-CA-CB	-6.23	98.46	110.30
1	A	197	PHE	N-CA-C	5.46	125.74	111.00
1	A	359	HIS	N-CA-C	-5.21	96.94	111.00

There are no chirality outliers.

There are no planarity outliers.

### 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	3389	0	3455	62	0
1	B	3306	0	3374	62	0
2	C	23	0	21	4	0
2	D	23	0	21	1	0
3	A	23	0	24	2	0
3	B	23	0	24	2	0
All	All	6787	0	6919	123	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 9.

All (123) 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:B:108:THR:HG22	1:B:168:LEU:HB2	1.40	1.03
1:A:108:THR:HG22	1:A:168:LEU:HB2	1.50	0.91
1:A:34:LEU:CD2	1:A:34:LEU:N	2.34	0.89
1:B:104:ASP:O	1:B:108:THR:HG23	1.77	0.85
1:A:34:LEU:HD23	1:A:34:LEU:N	1.94	0.81
1:B:108:THR:HG22	1:B:168:LEU:CB	2.10	0.80
1:A:83:MET:O	1:A:87:VAL:HG23	1.86	0.76
1:B:43:ILE:HG21	1:B:79:ALA:HB2	1.67	0.76
1:A:104:ASP:O	1:A:108:THR:HG23	1.89	0.73
1:A:200:PRO:O	1:A:204:THR:HG23	1.89	0.73
1:A:208:TYR:OH	1:A:331:VAL:HG23	1.90	0.70
1:A:155:THR:O	1:A:159:VAL:HG23	1.92	0.69
1:A:64:THR:OG1	1:A:128:ARG:NH1	2.28	0.67
1:A:34:LEU:HD22	1:A:34:LEU:N	2.11	0.65
1:A:40:ILE:HG22	1:A:41:LEU:HD23	1.81	0.63
1:B:346:PHE:O	1:B:350:VAL:HG23	1.99	0.63
1:B:209:ALA:O	1:B:213:VAL:HG23	1.98	0.63
1:B:118:ILE:HG21	1:B:196:SER:O	1.99	0.62
1:A:356:GLN:O	1:A:357:THR:CB	2.48	0.61
1:A:1114:PHE:O	1:A:1118:LEU:HG	2.00	0.61
1:A:118:ILE:HG21	1:A:196:SER:O	2.02	0.60
1:B:204:THR:CG2	1:B:335:LEU:HD22	2.31	0.60
1:B:355:CYS:SG	1:B:356:GLN:N	2.75	0.60
1:B:388:ILE:H	1:B:388:ILE:HD12	1.67	0.59
1:A:322:LEU:HD11	1:A:326:LYS:NZ	2.17	0.58
1:A:108:THR:HG21	1:A:169:LEU:HD13	1.84	0.58
1:A:1011:GLU:OE1	2:C:1:GLC:H62	2.03	0.58
1:B:330:MET:HE3	1:B:386:PHE:HB2	1.84	0.58

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:115:THR:CG2	1:A:161:ALA:HA	2.34	0.58
1:B:1007:LEU:HD21	1:B:1101:ASN:HA	1.86	0.57
1:B:45:PHE:O	1:B:49:LEU:HD13	2.05	0.56
1:B:1017:ILE:HG12	1:B:1027:ILE:HD12	1.88	0.56
1:A:356:GLN:NE2	1:A:357:THR:N	2.54	0.55
1:A:1105:GLN:HA	2:C:1:GLC:O6	2.06	0.55
1:A:33:ALA:C	1:A:34:LEU:HD22	2.27	0.55
1:A:1104:PHE:O	2:C:1:GLC:H61	2.05	0.55
1:A:115:THR:HG21	1:A:161:ALA:HA	1.89	0.54
1:A:1084:LEU:HD11	1:A:1108:GLU:O	2.08	0.54
1:B:125:SER:HB3	1:B:204:THR:HG23	1.90	0.53
1:B:1106:MET:HE3	1:B:1111:VAL:HG22	1.90	0.53
1:A:1015:LEU:O	1:A:1027:ILE:HD11	2.10	0.52
1:A:80:THR:O	1:A:81:LEU:HD23	2.10	0.52
1:B:71:LEU:O	1:B:74:ALA:HB3	2.10	0.52
1:A:360:VAL:HG13	1:A:360:VAL:O	2.10	0.52
1:A:379:ASN:HB2	1:A:380:PRO:HD3	1.92	0.52
1:B:375:ASN:HA	1:B:378:LEU:HD12	1.91	0.51
1:B:115:THR:O	1:B:118:ILE:HG22	2.11	0.51
1:A:33:ALA:C	1:A:34:LEU:CD2	2.79	0.51
1:B:77:LEU:HB2	1:B:113:MET:CE	2.40	0.51
1:B:204:THR:HG22	1:B:335:LEU:HD22	1.92	0.51
1:B:346:PHE:O	1:B:350:VAL:CG2	2.58	0.51
1:A:158:TRP:O	1:A:161:ALA:HB3	2.11	0.50
1:B:192:SER:HB3	3:B:1200:ETQ:H16	1.92	0.50
1:B:69:VAL:O	1:B:73:VAL:HG23	2.11	0.50
1:B:108:THR:CG2	1:B:168:LEU:HB2	2.27	0.50
1:B:107:VAL:HG12	1:B:168:LEU:HD13	1.92	0.50
1:A:1039:LEU:HD12	1:A:1039:LEU:O	2.11	0.50
1:B:1029:ILE:HD12	1:B:1063:ALA:HB1	1.93	0.50
1:B:355:CYS:SG	1:B:358:CYS:N	2.81	0.50
1:B:1106:MET:CE	1:B:1111:VAL:HG22	2.42	0.50
1:A:127:ASP:OD2	1:A:128:ARG:NH1	2.41	0.49
1:A:95:VAL:HG13	1:A:179:THR:HB	1.95	0.49
1:B:204:THR:HG21	1:B:335:LEU:HD22	1.94	0.49
1:B:150:VAL:HG12	1:B:154:ILE:HD12	1.96	0.48
1:A:63:THR:HG23	1:A:150:VAL:HG22	1.96	0.48
1:B:194:VAL:HG12	1:B:195:VAL:N	2.28	0.48
1:A:83:MET:N	1:A:84:PRO:CD	2.77	0.48
1:B:1026:THR:HG22	1:B:1027:ILE:N	2.28	0.48
1:B:213:VAL:HG12	1:B:217:GLN:OE1	2.13	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:122:CYS:HA	1:A:204:THR:HG22	1.95	0.48
1:A:1104:PHE:O	2:C:2:GLC:H5	2.14	0.47
1:B:77:LEU:HB2	1:B:113:MET:HE2	1.96	0.47
1:B:197:PHE:O	1:B:201:PHE:HB3	2.14	0.47
1:B:64:THR:O	1:B:68:VAL:HG23	2.14	0.47
1:A:1067:PHE:O	1:A:1071:VAL:HG23	2.15	0.47
1:A:111:VAL:HA	3:A:1200:ETQ:CL1	2.52	0.47
1:A:1124:LYS:HD3	1:A:1126:TRP:CZ2	2.50	0.47
1:A:1094:VAL:O	1:A:1097:ALA:HB3	2.15	0.47
1:A:164:VAL:HG22	1:B:159:VAL:HG11	1.96	0.46
1:A:111:VAL:HG21	1:A:168:LEU:HD11	1.98	0.46
1:A:166:CYS:N	1:A:167:PRO:CD	2.79	0.46
1:B:1105:GLN:NE2	1:B:1142:THR:HG21	2.31	0.46
1:A:1027:ILE:HG21	1:A:1046:LEU:HD13	1.98	0.46
1:A:108:THR:HA	1:A:168:LEU:HD12	1.98	0.46
1:B:62:THR:HG22	1:B:64:THR:H	1.81	0.45
1:A:1072:ASP:HB3	1:A:1076:ARG:NH1	2.31	0.45
1:A:107:VAL:HG11	1:A:183:ILE:HG13	1.99	0.45
1:B:108:THR:HG21	1:B:169:LEU:HG	1.99	0.44
1:B:1104:PHE:O	2:D:1:GLC:H61	2.16	0.44
1:B:1029:ILE:HD13	1:B:1067:PHE:HB2	1.98	0.44
1:B:39:LEU:HD21	1:B:374:VAL:HA	1.99	0.44
1:B:68:VAL:HG13	1:B:383:TYR:OH	2.18	0.44
1:A:126:ILE:HD12	1:B:170:PHE:HD2	1.83	0.44
1:B:1087:VAL:HG21	1:B:1118:LEU:HD22	2.00	0.44
1:B:74:ALA:O	1:B:78:VAL:HG23	2.17	0.44
1:A:356:GLN:C	1:A:356:GLN:NE2	2.70	0.44
1:B:1094:VAL:HG22	1:B:1158:TRP:NE1	2.33	0.44
1:B:1026:THR:CG2	1:B:1027:ILE:N	2.81	0.43
1:A:71:LEU:HD22	1:A:121:LEU:HD13	2.00	0.43
1:A:121:LEU:HA	1:A:121:LEU:HD12	1.71	0.43
1:B:60:LEU:HD11	1:B:390:PHE:HA	2.01	0.43
1:B:133:VAL:C	1:B:135:PRO:HD3	2.39	0.43
1:A:168:LEU:HD21	1:A:188:PHE:CE2	2.54	0.42
3:A:1200:ETQ:C6	3:A:1200:ETQ:H15B	2.49	0.42
1:B:78:VAL:HG22	1:B:113:MET:CB	2.50	0.42
1:A:1013:LEU:HD23	1:A:1014:ARG:N	2.34	0.42
1:A:334:VAL:HG13	1:A:379:ASN:OD1	2.19	0.42
1:B:43:ILE:HG13	1:B:377:ALA:HB2	2.02	0.42
1:A:1011:GLU:OE1	1:A:1030:GLY:HA3	2.19	0.42
1:A:1078:ILE:HG23	1:A:1084:LEU:HB3	2.02	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:68:VAL:HG13	1:A:383:TYR:OH	2.20	0.41
1:B:347:LEU:O	1:B:351:LEU:HB2	2.20	0.41
1:B:330:MET:HE1	1:B:383:TYR:HA	2.03	0.41
1:A:1078:ILE:HG23	1:A:1084:LEU:CB	2.51	0.41
1:A:128:ARG:NH2	1:A:324:GLU:OE1	2.53	0.41
1:B:1029:ILE:HD12	1:B:1029:ILE:N	2.36	0.41
1:A:1050:ILE:HG22	1:A:1052:ARG:HG2	2.03	0.41
1:B:166:CYS:SG	1:B:167:PRO:HD3	2.61	0.41
1:B:183:ILE:HD12	3:B:1200:ETQ:C1	2.51	0.41
1:A:343:LEU:N	1:A:344:PRO:CD	2.84	0.40
1:B:43:ILE:CG2	1:B:79:ALA:HB2	2.44	0.40
1:B:330:MET:CE	1:B:386:PHE:HB2	2.49	0.40
1:B:1060:LYS:O	1:B:1064:GLU:HB2	2.22	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	430/481 (89%)	404 (94%)	26 (6%)	0	100	100
1	B	419/481 (87%)	402 (96%)	16 (4%)	1 (0%)	47	78
All	All	849/962 (88%)	806 (95%)	42 (5%)	1 (0%)	51	82

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	319	GLY

### 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	367/410 (90%)	352 (96%)	15 (4%)	30	64
1	B	358/410 (87%)	351 (98%)	7 (2%)	55	82
All	All	725/820 (88%)	703 (97%)	22 (3%)	41	75

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	34	LEU
1	A	70	SER
1	A	121	LEU
1	A	147	CYS
1	A	166	CYS
1	A	215	LEU
1	A	1037	PRO
1	A	1039	LEU
1	A	1057	VAL
1	A	1109	THR
1	A	1127	ASP
1	A	1140	ASN
1	A	348	THR
1	A	356	GLN
1	A	358	CYS
1	B	89	LEU
1	B	172	PHE
1	B	175	THR
1	B	195	VAL
1	B	201	PHE
1	B	355	CYS
1	B	400	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	97	ASN
1	A	1140	ASN
1	A	354	HIS
1	A	356	GLN
1	B	375	ASN
1	B	387	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 ⓘ

4 monosaccharides 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	$\# Z  > 2$
2	GLC	C	1	2	12,12,12	0.86	0	17,17,17	2.41	5 (29%)
2	GLC	C	2	2	11,11,12	0.54	0	15,15,17	1.06	1 (6%)
2	GLC	D	1	2	12,12,12	0.73	0	17,17,17	1.31	2 (11%)
2	GLC	D	2	2	11,11,12	0.83	0	15,15,17	1.60	3 (20%)

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	GLC	C	1	2	1/1/5/5	2/2/22/22	0/1/1/1

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GLC	C	2	2	-	0/2/19/22	0/1/1/1
2	GLC	D	1	2	1/1/5/5	1/2/22/22	0/1/1/1
2	GLC	D	2	2	-	2/2/19/22	0/1/1/1

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	1	GLC	O5-C5-C4	5.92	120.44	109.69
2	C	1	GLC	C6-C5-C4	-5.43	100.28	113.00
2	D	2	GLC	O5-C5-C6	3.85	113.24	107.20
2	D	1	GLC	C1-O5-C5	-3.47	107.11	113.66
2	D	2	GLC	C1-C2-C3	3.05	113.41	109.67
2	C	1	GLC	O4-C4-C3	2.98	117.24	110.35
2	C	1	GLC	O5-C1-C2	2.91	115.47	110.28
2	C	2	GLC	O5-C1-C2	-2.71	106.58	110.77
2	D	2	GLC	O5-C1-C2	2.60	114.79	110.77
2	C	1	GLC	C3-C4-C5	2.44	114.59	110.24
2	D	1	GLC	O5-C5-C6	2.08	111.61	106.44

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	1	GLC	C1
2	D	1	GLC	C1

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1	GLC	O5-C5-C6-O6
2	C	1	GLC	C4-C5-C6-O6
2	D	2	GLC	O5-C5-C6-O6
2	D	2	GLC	C4-C5-C6-O6
2	D	1	GLC	O5-C5-C6-O6

There are no ring outliers.

3 monomers are involved in 5 short contacts:

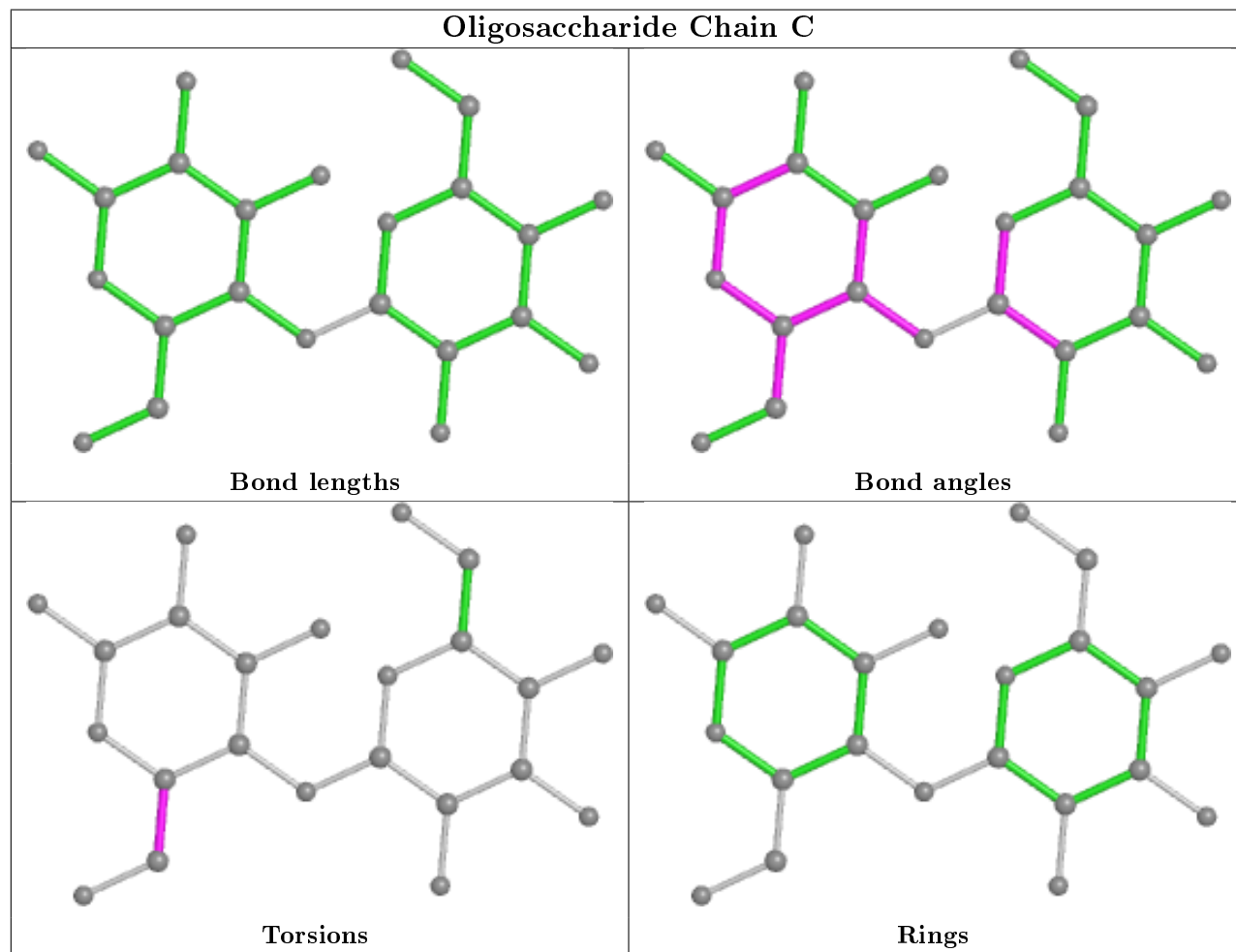
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	2	GLC	1	0

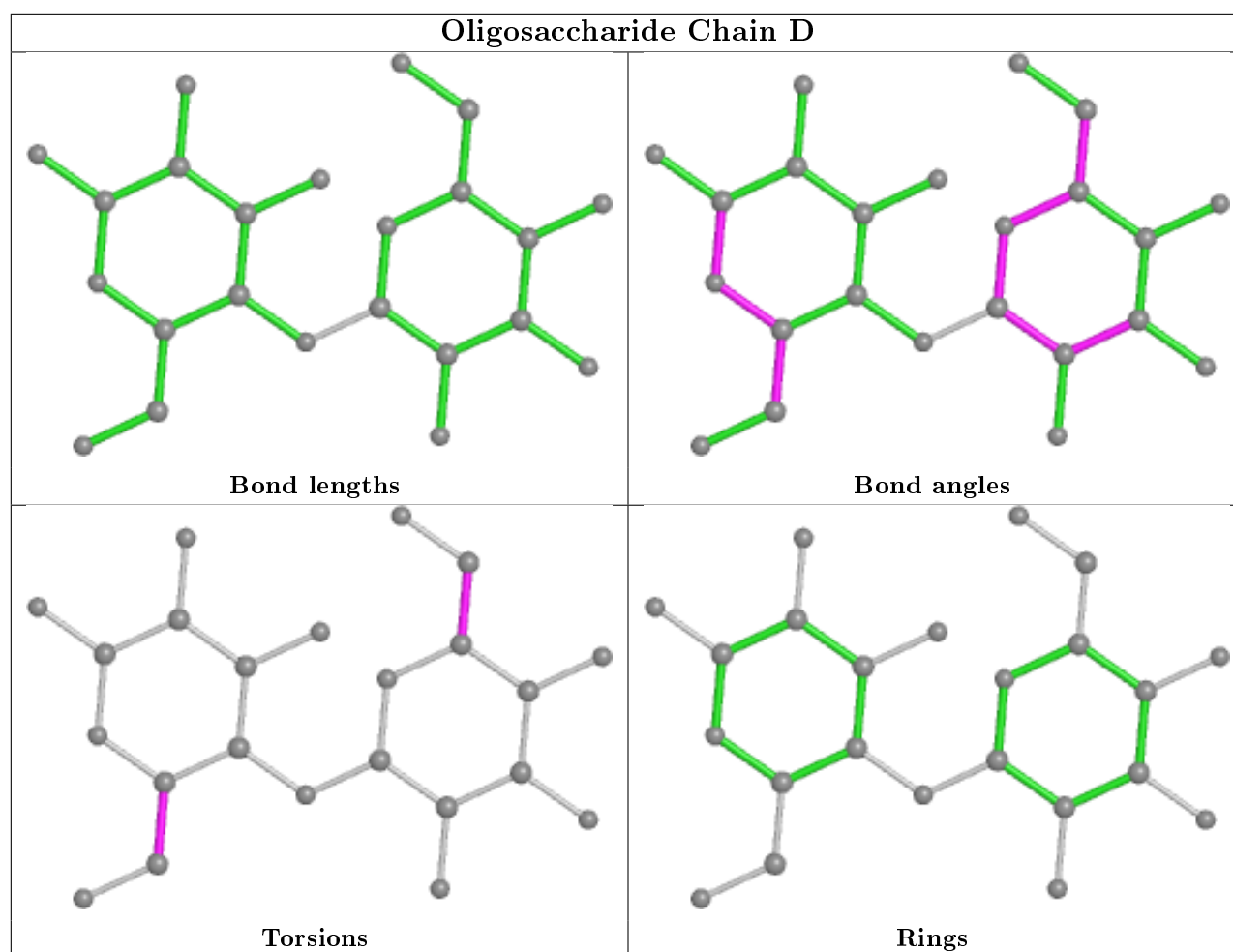
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1	GLC	3	0
2	D	1	GLC	1	0

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 [i](#)

2 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	$\# Z  > 2$
3	ETQ	B	1200	-	24,24,24	0.72	1 (4%)	32,33,33	1.03	2 (6%)
3	ETQ	A	1200	-	24,24,24	0.72	1 (4%)	32,33,33	1.06	2 (6%)

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	ETQ	B	1200	-	-	5/15/25/25	0/2/2/2
3	ETQ	A	1200	-	-	7/15/25/25	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1200	ETQ	C7-C8	2.81	1.55	1.51
3	A	1200	ETQ	C7-C8	2.80	1.55	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1200	ETQ	C16-C2-C1	3.46	124.53	120.14
3	B	1200	ETQ	C16-C2-C1	3.06	124.03	120.14
3	B	1200	ETQ	C6-O2-C5	-2.22	108.70	114.78
3	A	1200	ETQ	C6-O2-C5	-2.20	108.75	114.78

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	1200	ETQ	C10-C9-N1-C8
3	A	1200	ETQ	C10-C9-N1-C8
3	A	1200	ETQ	C5-C7-C8-O3
3	A	1200	ETQ	C17-C16-C2-C3
3	A	1200	ETQ	C17-C16-C2-C1
3	B	1200	ETQ	C5-C7-C8-O3
3	B	1200	ETQ	C5-C7-C8-N1
3	A	1200	ETQ	C5-C7-C8-N1
3	B	1200	ETQ	N2-C10-C9-N1
3	A	1200	ETQ	C15-C14-N2-C13
3	B	1200	ETQ	C1-C7-C8-N1
3	A	1200	ETQ	C1-C7-C8-N1

There are no ring outliers.

2 monomers are involved in 4 short contacts:

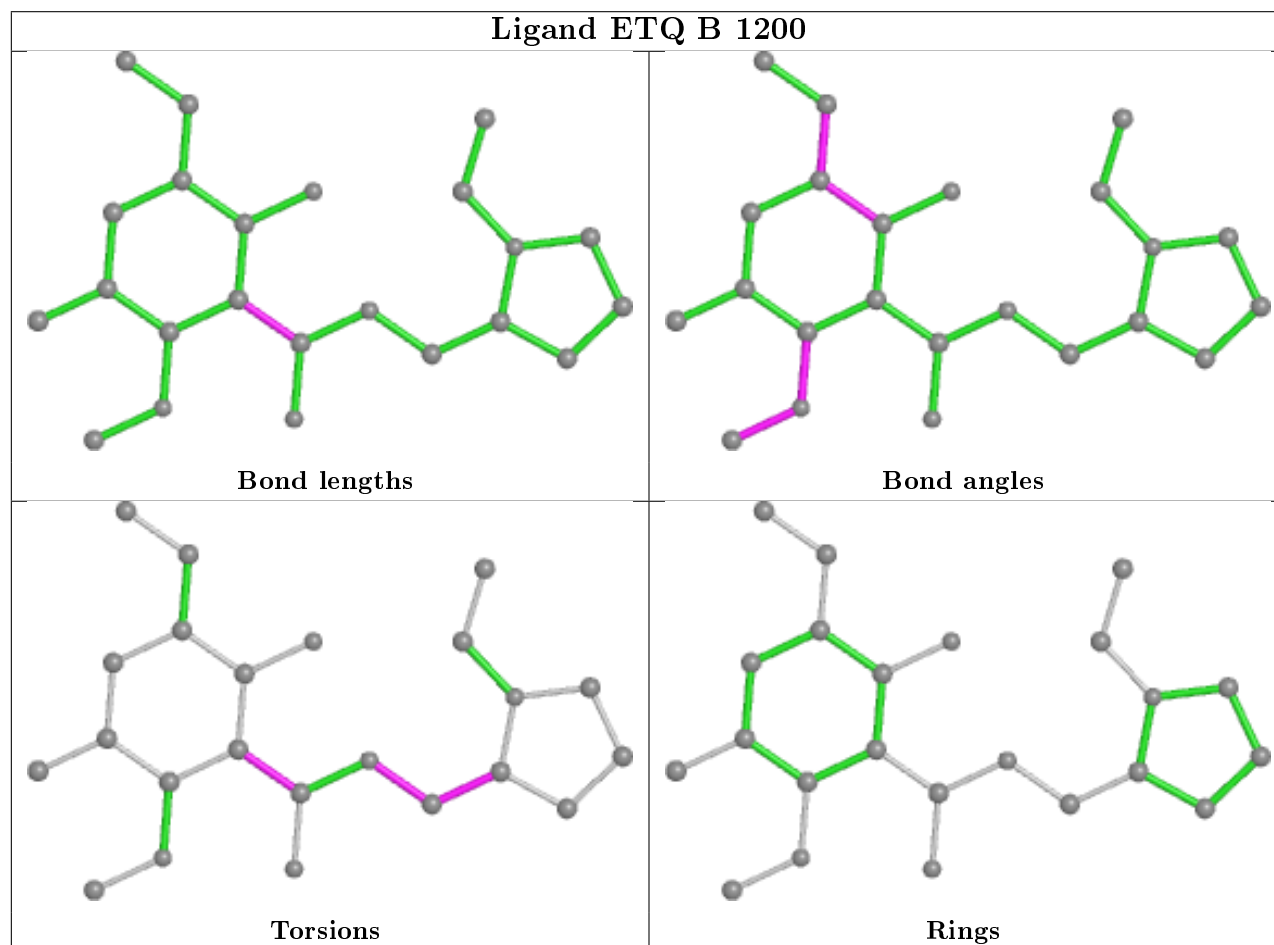
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1200	ETQ	2	0

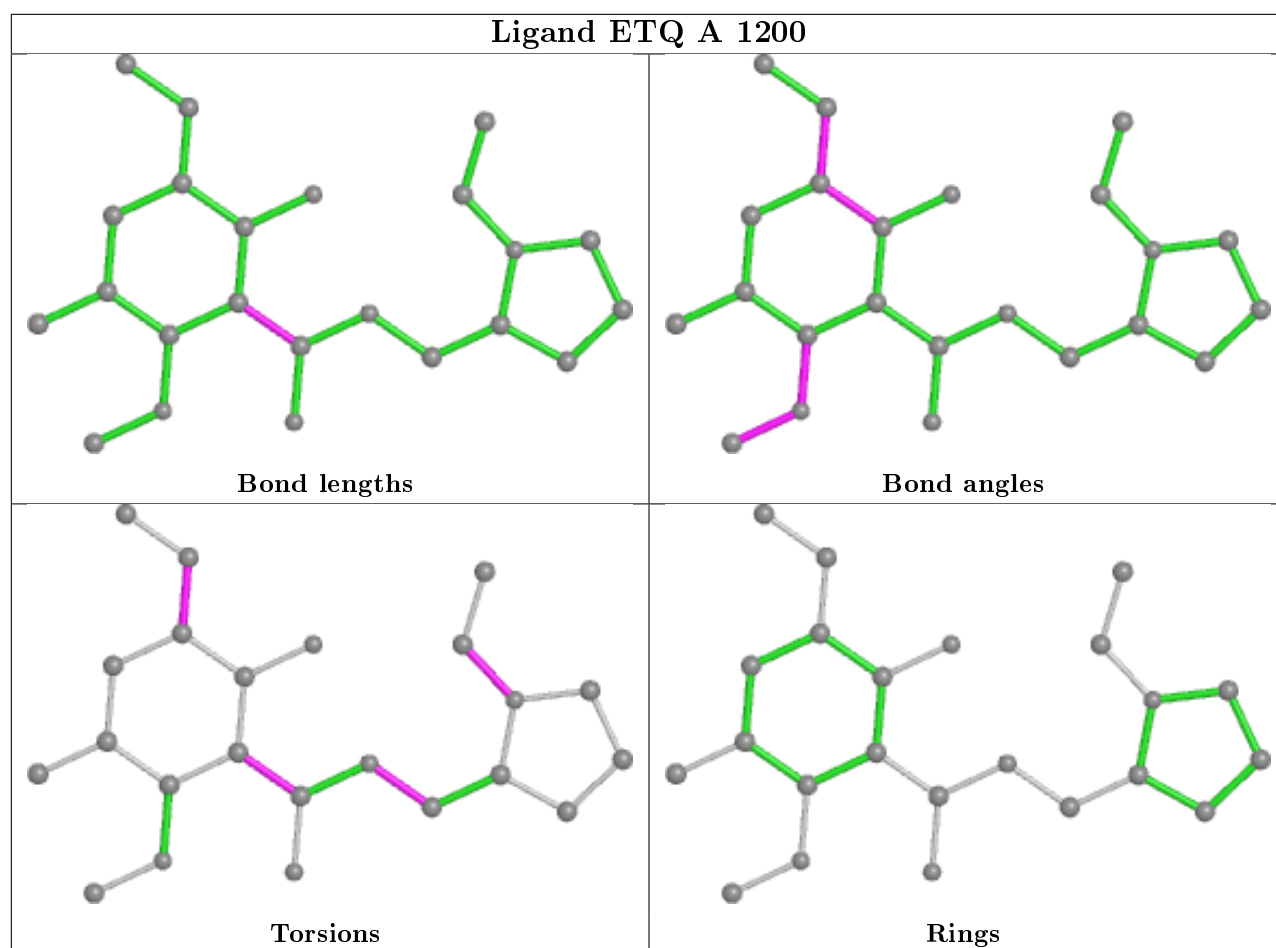
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1200	ETQ	2	0

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 ⓘ

### 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	432/481 (89%)	0.01	16 (3%) 41 37	35, 65, 117, 179	0
1	B	423/481 (87%)	0.18	17 (4%) 38 33	35, 75, 151, 182	0
All	All	855/962 (88%)	0.09	33 (3%) 39 35	35, 69, 133, 182	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1037	PRO	6.2
1	A	400	CYS	5.0
1	B	1036	SER	4.6
1	B	1038	SER	3.9
1	A	355	CYS	3.8
1	A	356	GLN	3.5
1	A	144	GLN	3.4
1	B	1041	ALA	3.4
1	A	357	THR	3.1
1	B	397	ILE	3.0
1	B	32	TYR	2.9
1	A	220	ARG	2.9
1	B	356	GLN	2.9
1	B	1055	ASN	2.9
1	B	1051	GLY	2.8
1	A	32	TYR	2.7
1	B	357	THR	2.7
1	A	142	THR	2.6
1	A	143	GLY	2.6
1	A	399	SER	2.6
1	A	392	LYS	2.5
1	B	1014	ARG	2.4
1	A	145	SER	2.4
1	B	1015	LEU	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	146	SER	2.2
1	B	1016	LYS	2.2
1	A	141	GLY	2.1
1	A	218	ARG	2.1
1	B	1057	VAL	2.1
1	B	134	MET	2.0
1	A	140	HIS	2.0
1	B	1028	GLY	2.0
1	B	399	SER	2.0

## 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.

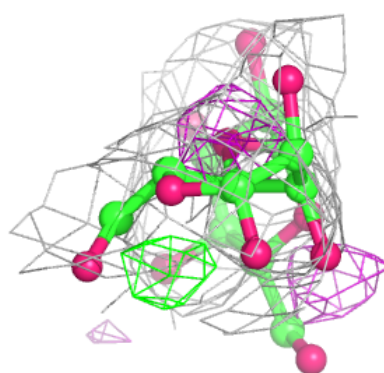
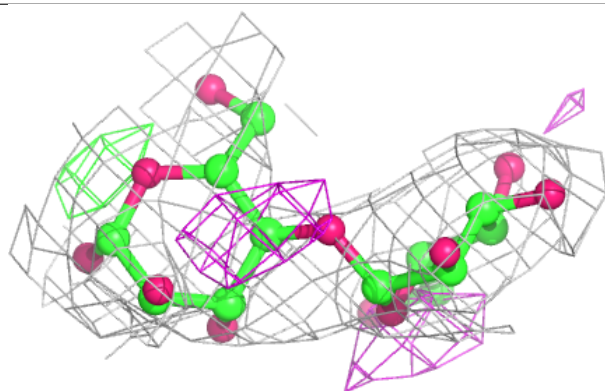
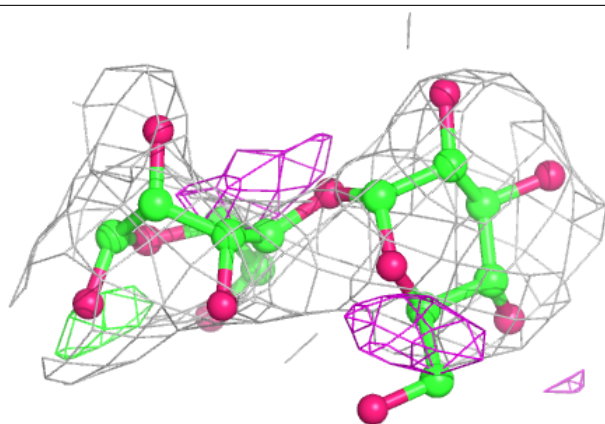
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	GLC	C	1	12/12	0.89	0.23	23,96,144,165	0
2	GLC	C	2	11/12	0.90	0.29	38,87,140,152	0
2	GLC	D	2	11/12	0.91	0.30	35,126,163,193	0
2	GLC	D	1	12/12	0.93	0.16	60,92,135,145	0

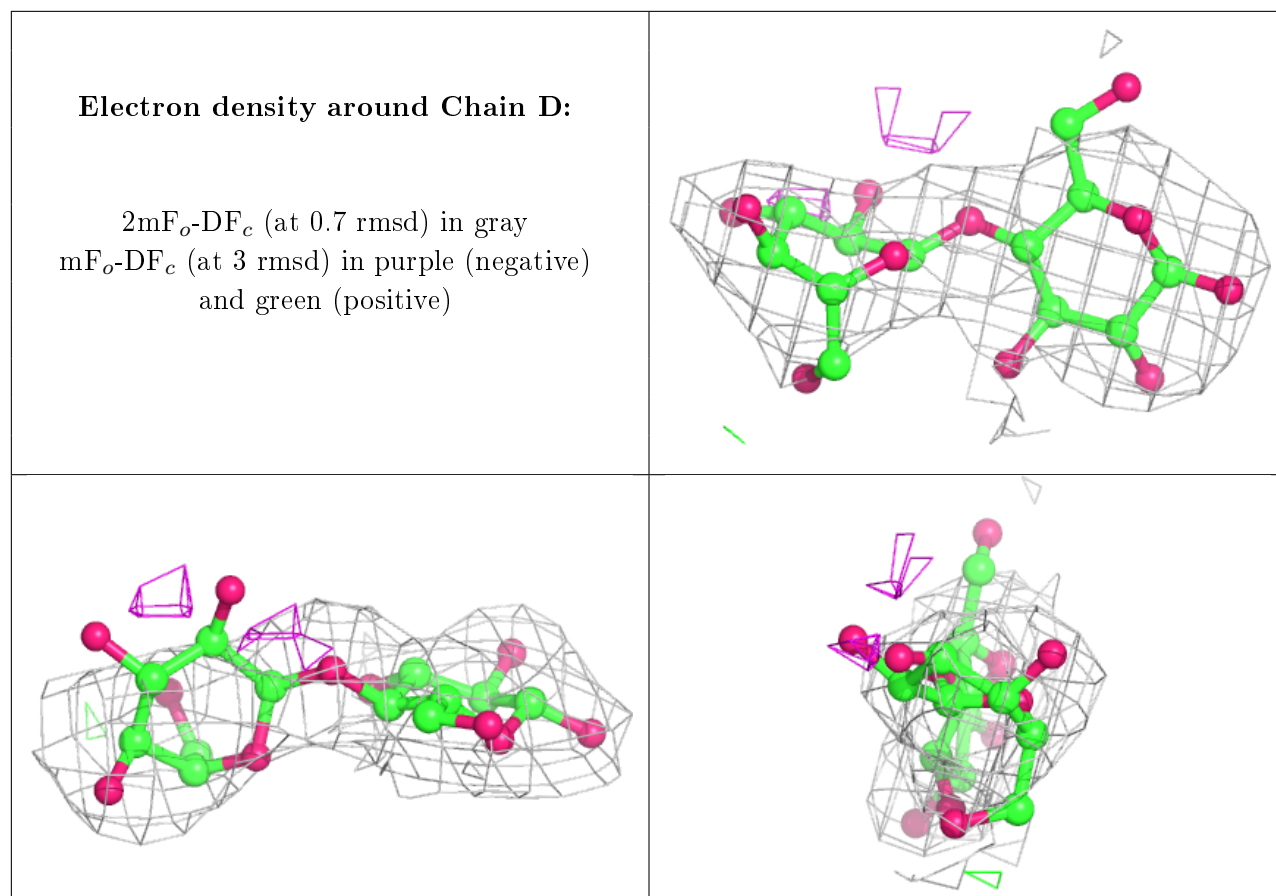
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



**Electron density around Chain C:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





## 6.4 Ligands ⓘ

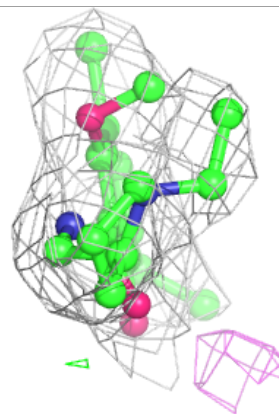
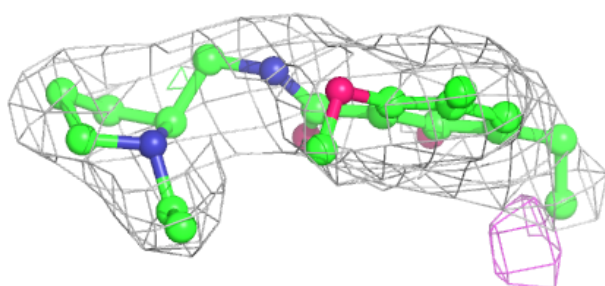
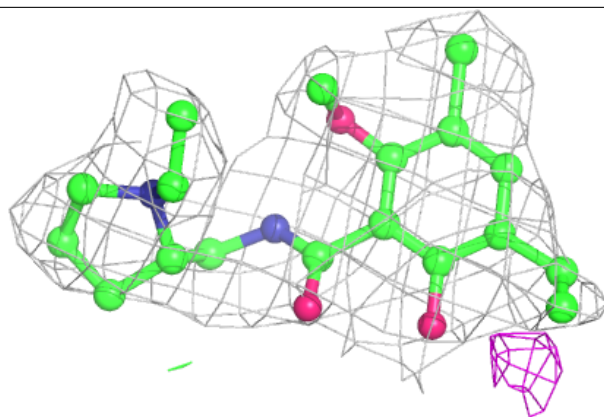
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
3	ETQ	A	1200	23/23	0.93	0.20	7,57,100,122	0
3	ETQ	B	1200	23/23	0.96	0.19	11,47,91,134	0

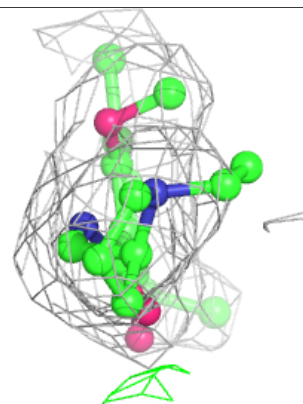
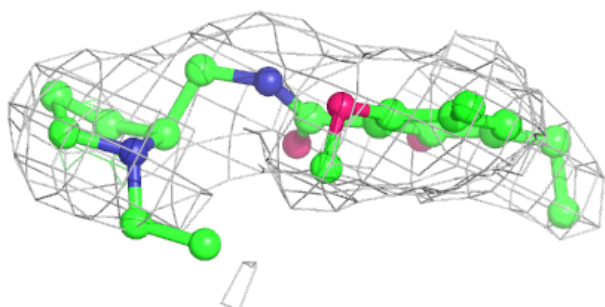
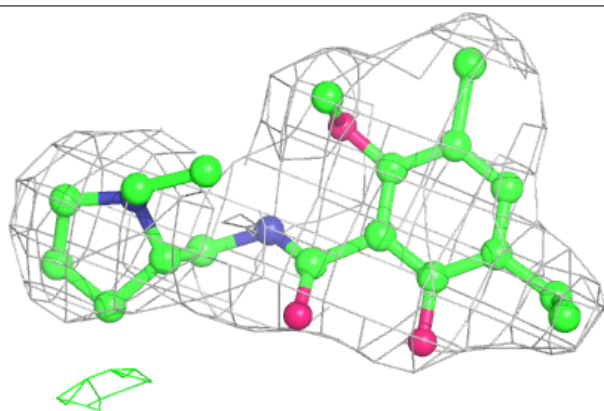
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around ETQ A 1200:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around ETQ B 1200:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers [i](#)

There are no such residues in this entry.