



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 6, 2020 – 10:44 PM BST

PDB ID : 4XPT  
Title : X-ray structure of Drosophila dopamine transporter with subsiteB mutations D121G/S426M and EL2 deletion of 162-201 in complex with substrate analogue 3,4 dichlorophen ethylamine  
Authors : Aravind, P.; Wang, K.; Gouaux, E.  
Deposited on : 2015-01-17  
Resolution : 3.36 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

<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

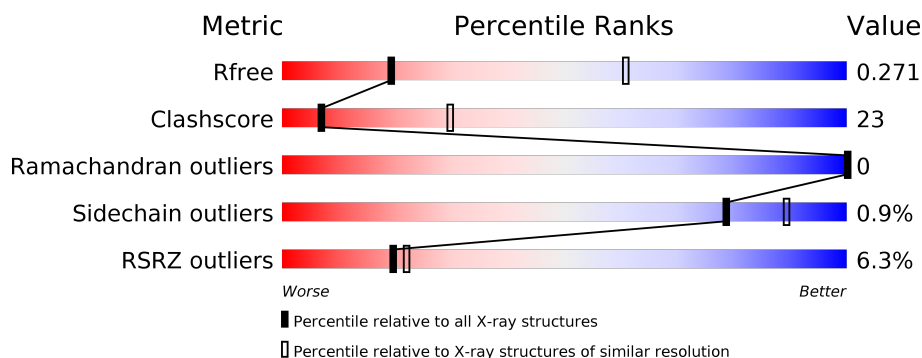
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.36 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	1558 (3.42-3.30)
Clashscore	141614	1627 (3.42-3.30)
Ramachandran outliers	138981	1599 (3.42-3.30)
Sidechain outliers	138945	1598 (3.42-3.30)
RSRZ outliers	127900	1507 (3.42-3.30)

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	535	<div> <div>9%</div> <div>53%</div> <div>46%</div> </div>
2	L	214	<div> <div>4%</div> <div>67%</div> <div>30%</div> <div>.</div> </div>
3	H	219	<div> <div>3%</div> <div>69%</div> <div>31%</div> </div>
4	B	2	<div> <div>50%</div> <div>50%</div> </div>
4	C	2	<div> <div>100%</div> </div>
4	D	2	<div> <div>100%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	E	2	 100%
4	F	2	 50%50%

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	BGC	F	1	-	-	-	X
5	Y01	A	601	-	-	-	X
7	NAG	A	603	-	-	-	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 7685 atoms, of which 14 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 Dopamine transporter.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	535	Total	C	N	O	S	0	2	0
			4216	2826	657	713	20			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	74	ALA	VAL	engineered mutation	UNP Q7K4Y6
A	121	GLY	ASP	engineered mutation	UNP Q7K4Y6
A	?	-	SER	deletion	UNP Q7K4Y6
A	?	-	GLN	deletion	UNP Q7K4Y6
A	?	-	ASN	deletion	UNP Q7K4Y6
A	?	-	ALA	deletion	UNP Q7K4Y6
A	?	-	SER	deletion	UNP Q7K4Y6
A	?	-	ARG	deletion	UNP Q7K4Y6
A	?	-	VAL	deletion	UNP Q7K4Y6
A	?	-	PRO	deletion	UNP Q7K4Y6
A	?	-	VAL	deletion	UNP Q7K4Y6
A	?	-	ILE	deletion	UNP Q7K4Y6
A	?	-	GLY	deletion	UNP Q7K4Y6
A	?	-	ASN	deletion	UNP Q7K4Y6
A	?	-	TYR	deletion	UNP Q7K4Y6
A	?	-	SER	deletion	UNP Q7K4Y6
A	?	-	ASP	deletion	UNP Q7K4Y6
A	?	-	LEU	deletion	UNP Q7K4Y6
A	?	-	TYR	deletion	UNP Q7K4Y6
A	?	-	ALA	deletion	UNP Q7K4Y6
A	?	-	MET	deletion	UNP Q7K4Y6
A	?	-	GLY	deletion	UNP Q7K4Y6
A	?	-	ASN	deletion	UNP Q7K4Y6
A	?	-	GLN	deletion	UNP Q7K4Y6
A	?	-	SER	deletion	UNP Q7K4Y6
A	?	-	LEU	deletion	UNP Q7K4Y6
A	?	-	LEU	deletion	UNP Q7K4Y6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	TYR	deletion	UNP Q7K4Y6
A	?	-	ASN	deletion	UNP Q7K4Y6
A	?	-	GLU	deletion	UNP Q7K4Y6
A	?	-	THR	deletion	UNP Q7K4Y6
A	?	-	TYR	deletion	UNP Q7K4Y6
A	?	-	MET	deletion	UNP Q7K4Y6
A	?	-	ASN	deletion	UNP Q7K4Y6
A	?	-	GLY	deletion	UNP Q7K4Y6
A	?	-	SER	deletion	UNP Q7K4Y6
A	?	-	SER	deletion	UNP Q7K4Y6
A	?	-	LEU	deletion	UNP Q7K4Y6
A	?	-	ASP	deletion	UNP Q7K4Y6
A	?	-	THR	deletion	UNP Q7K4Y6
A	?	-	SER	deletion	UNP Q7K4Y6
A	?	-	ALA	deletion	UNP Q7K4Y6
A	415	ALA	LEU	engineered mutation	UNP Q7K4Y6
A	426	MET	SER	engineered mutation	UNP Q7K4Y6

- Molecule 2 is a protein called antibody fragment light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	214	Total	C	N	O	S	0	0	0
			1626	1011	271	336	8			

- Molecule 3 is a protein called antibody fragment heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	219	Total	C	N	O	S	0	0	0
			1623	1022	275	318	8			

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



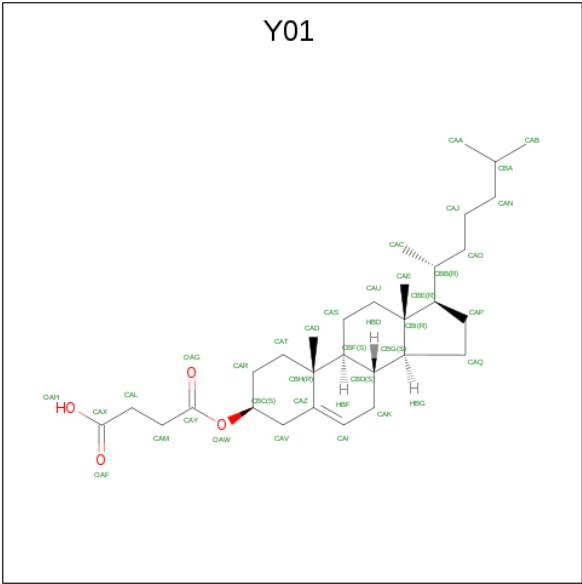
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
4	B	2	Total	C	O	0	0	0
			23	12	11			
4	C	2	Total	C	O	0	0	0
			23	12	11			

*Continued on next page...*

Continued from previous page...

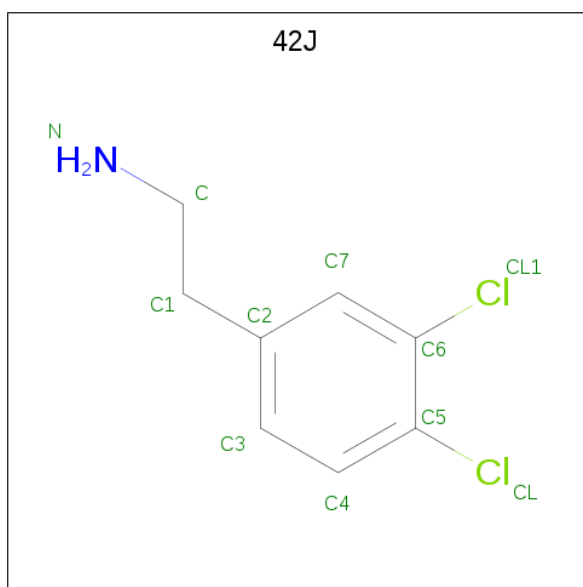
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
4	D	2	Total	C	O	0	0	0
			23	12	11			
4	E	2	Total	C	O	0	0	0
			23	12	11			
4	F	2	Total	C	O	0	0	0
			23	12	11			

- Molecule 5 is CHOLESTEROL HEMISUCCINATE (three-letter code: Y01) (formula: C<sub>31</sub>H<sub>50</sub>O<sub>4</sub>).



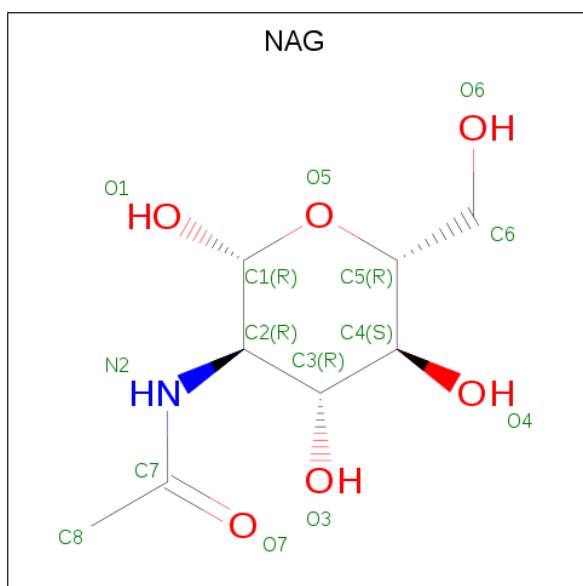
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			35	31	4		

- Molecule 6 is 2-(3,4-dichlorophenyl)ethanamine (three-letter code: 42J) (formula: C<sub>8</sub>H<sub>9</sub>Cl<sub>2</sub>N).



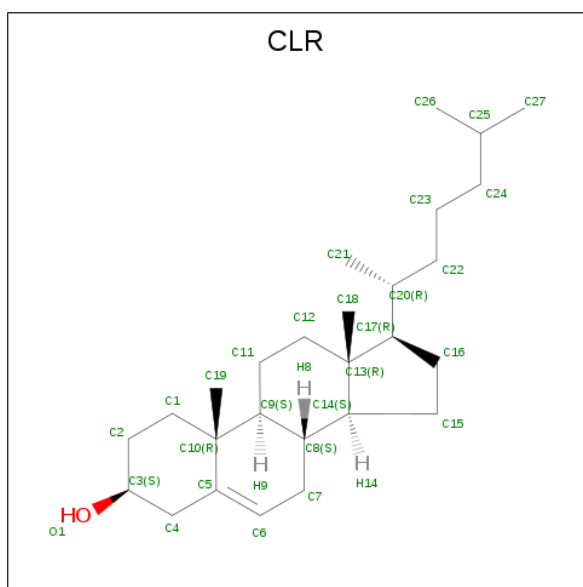
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	Cl	N	0	0
			11	8	2	1		

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	H	N	O	0	0
			28	8	14	1	5		

- Molecule 8 is CHOLESTEROL (three-letter code: CLR) (formula:  $C_{27}H_{46}O$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			28	27	1		

- Molecule 9 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	1	Total	Cl	0	0
			1	1		

- Molecule 10 is SODIUM ION (three-letter code: NA) (formula: Na).

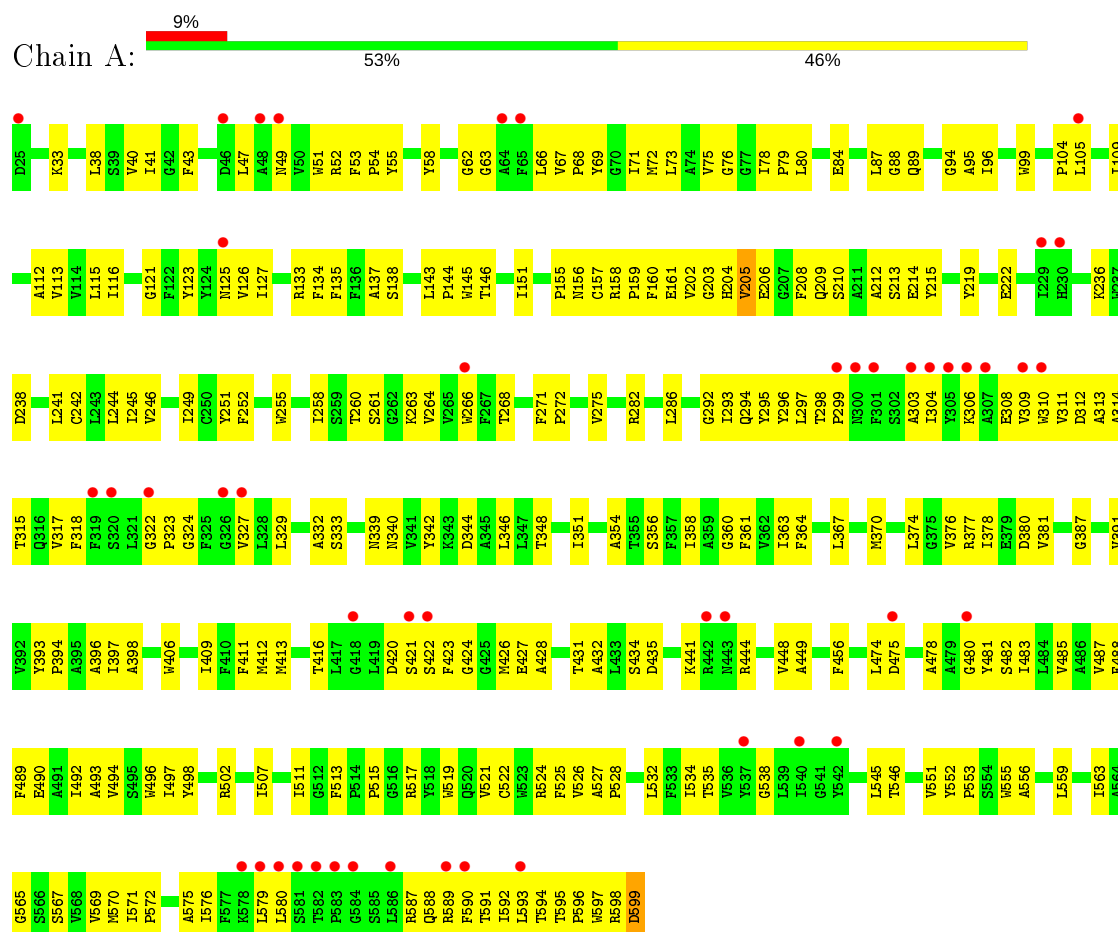
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	2	Total	Na	0	0
			2	2		



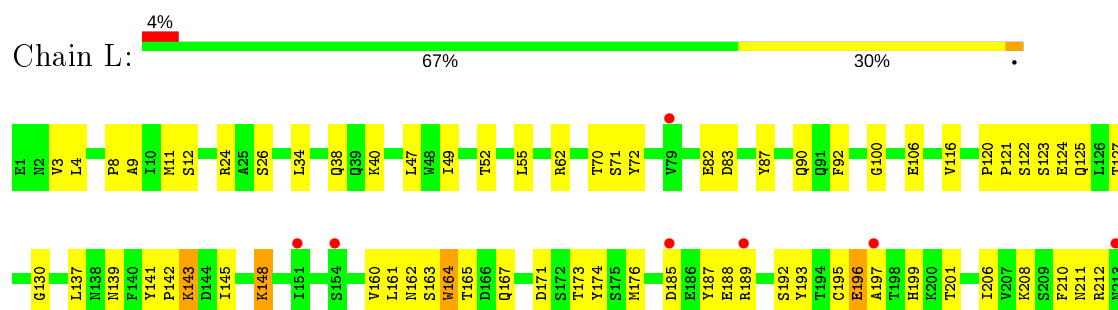
### 3 Residue-property plots

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: Dopamine transporter



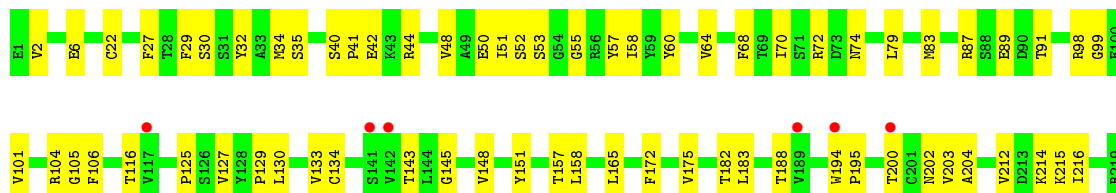
#### • Molecule 2: antibody fragment light chain



E214

- Molecule 3: antibody fragment heavy chain

Chain H:  3% 69% 31%



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

Chain B:  50% 50%

BGC1  
GLC2

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

Chain C:  100%

BGC1  
GLC2

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

Chain D:  100%

BGC1  
GLC2

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

Chain E:  100%

BGC1  
GLC2

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

Chain F:  50% 50%

BGC1  
GLC2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	96.61Å 139.15Å 166.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.60 – 3.36 38.60 – 3.35	Depositor EDS
% Data completeness (in resolution range)	(Not available) (38.60-3.36) 98.6 (38.60-3.35)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.27 (at 3.32Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.232 , 0.273 0.233 , 0.271	Depositor DCC
$R_{free}$ test set	1644 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	114.7	Xtriage
Anisotropy	0.164	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 76.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7685	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	105.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.48% 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: BGC, NAG, CL, NA, GLC, Y01, 42J, CLR

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	0.21	0/4360	0.38	0/5958
2	L	0.21	0/1664	0.39	0/2262
3	H	0.21	0/1662	0.38	0/2268
All	All	0.21	0/7686	0.38	0/10488

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4216	0	4128	214	0
2	L	1626	0	1536	68	0
3	H	1623	0	1557	58	0
4	B	23	0	21	5	0
4	C	23	0	21	5	0
4	D	23	0	21	4	0
4	E	23	0	21	3	0
4	F	23	0	21	2	0
5	A	35	0	49	5	0
6	A	11	0	9	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	14	14	13	0	0
8	A	28	0	46	6	0
9	A	1	0	0	0	0
10	A	2	0	0	0	0
All	All	7671	14	7443	347	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 23.

The worst 5 of 347 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)
4:C:1:BGC:H3	4:C:2:GLC:H3	1.34	1.10
4:B:1:BGC:H3	4:B:2:GLC:H3	1.25	1.10
1:A:204:HIS:HB2	1:A:205:VAL:HA	1.36	1.07
1:A:202:VAL:HG21	3:H:42:GLU:HG2	1.37	1.05
4:B:1:BGC:H3	4:B:2:GLC:C3	1.98	0.94

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	535/535 (100%)	511 (96%)	24 (4%)	0	100	100
2	L	212/214 (99%)	204 (96%)	8 (4%)	0	100	100
3	H	217/219 (99%)	213 (98%)	4 (2%)	0	100	100
All	All	964/968 (100%)	928 (96%)	36 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	429/439 (98%)	427 (100%)	2 (0%)	88	94
2	L	183/187 (98%)	178 (97%)	5 (3%)	44	72
3	H	176/187 (94%)	176 (100%)	0	100	100
All	All	788/813 (97%)	781 (99%)	7 (1%)	78	89

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	L	148	LYS
2	L	214	GLU
2	L	164	TRP
1	A	599	ASP
2	L	196	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	520	GLN

### 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](#)

10 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	$\# Z  > 2$
4	BGC	B	1	4	12,12,12	0.48	0	17,17,17	0.75	0
4	GLC	B	2	4	11,11,12	0.49	0	15,15,17	1.17	2 (13%)
4	BGC	C	1	4	12,12,12	0.51	0	17,17,17	0.76	1 (5%)
4	GLC	C	2	4	11,11,12	0.40	0	15,15,17	0.89	1 (6%)
4	BGC	D	1	4	12,12,12	0.52	0	17,17,17	1.00	2 (11%)
4	GLC	D	2	4	11,11,12	0.31	0	15,15,17	0.97	1 (6%)
4	BGC	E	1	4	12,12,12	0.57	0	17,17,17	1.06	1 (5%)
4	GLC	E	2	4	11,11,12	0.32	0	15,15,17	0.98	2 (13%)
4	BGC	F	1	4	12,12,12	0.55	0	17,17,17	1.15	1 (5%)
4	GLC	F	2	4	11,11,12	0.26	0	15,15,17	0.42	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
4	BGC	B	1	4	-	2/2/22/22	0/1/1/1
4	GLC	B	2	4	-	2/2/19/22	0/1/1/1
4	BGC	C	1	4	-	1/2/22/22	0/1/1/1
4	GLC	C	2	4	-	2/2/19/22	0/1/1/1
4	BGC	D	1	4	-	2/2/22/22	0/1/1/1
4	GLC	D	2	4	-	0/2/19/22	0/1/1/1
4	BGC	E	1	4	-	0/2/22/22	0/1/1/1
4	GLC	E	2	4	-	0/2/19/22	0/1/1/1
4	BGC	F	1	4	-	1/2/22/22	0/1/1/1
4	GLC	F	2	4	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	2	GLC	C1-C2-C3	2.71	113.00	109.67
4	E	2	GLC	O5-C1-C2	-2.45	107.00	110.77
4	C	2	GLC	O5-C5-C6	2.40	110.96	107.20
4	B	2	GLC	O5-C5-C6	2.39	110.94	107.20
4	D	1	BGC	C1-O5-C5	2.36	118.12	113.66

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

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

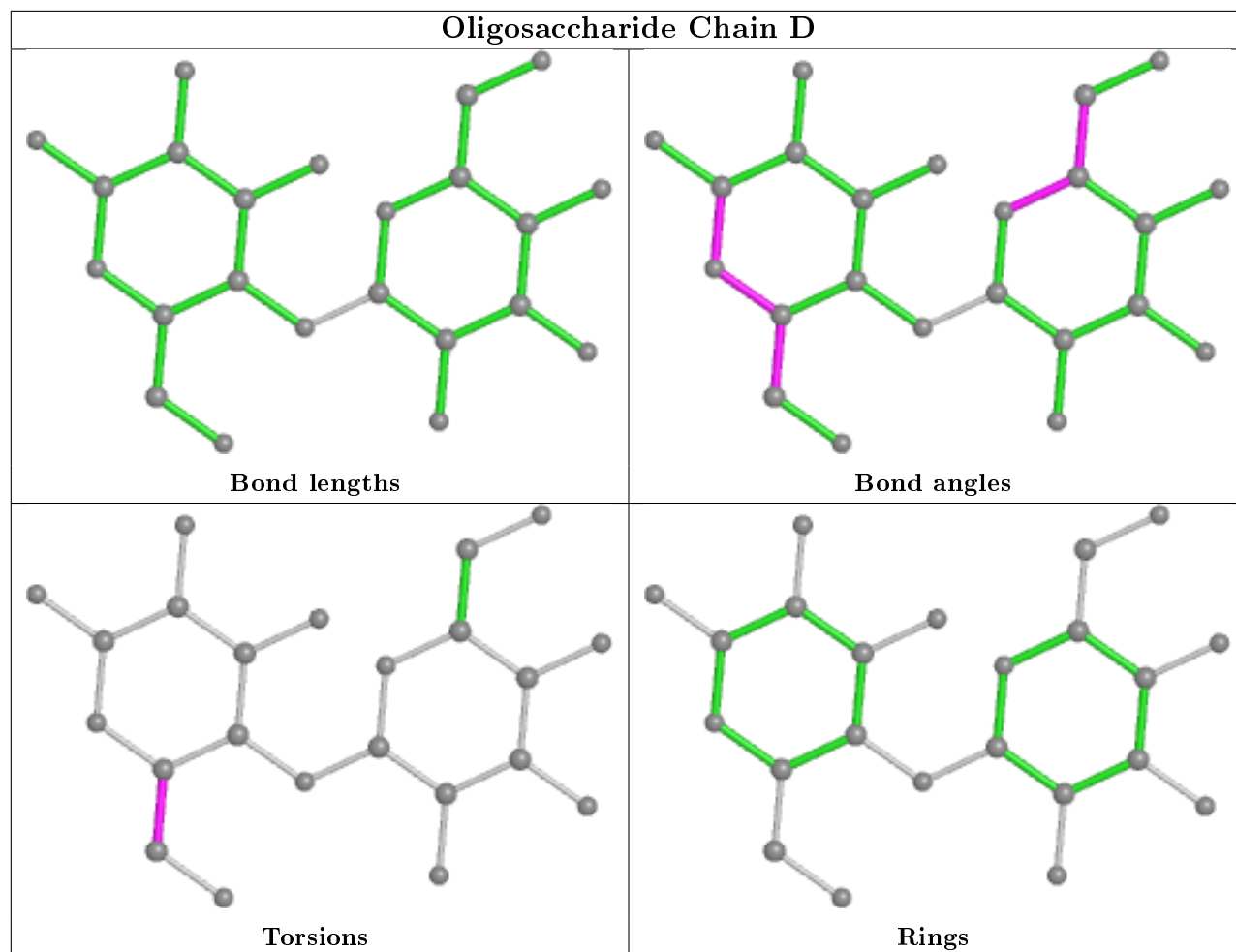
There are no ring outliers.

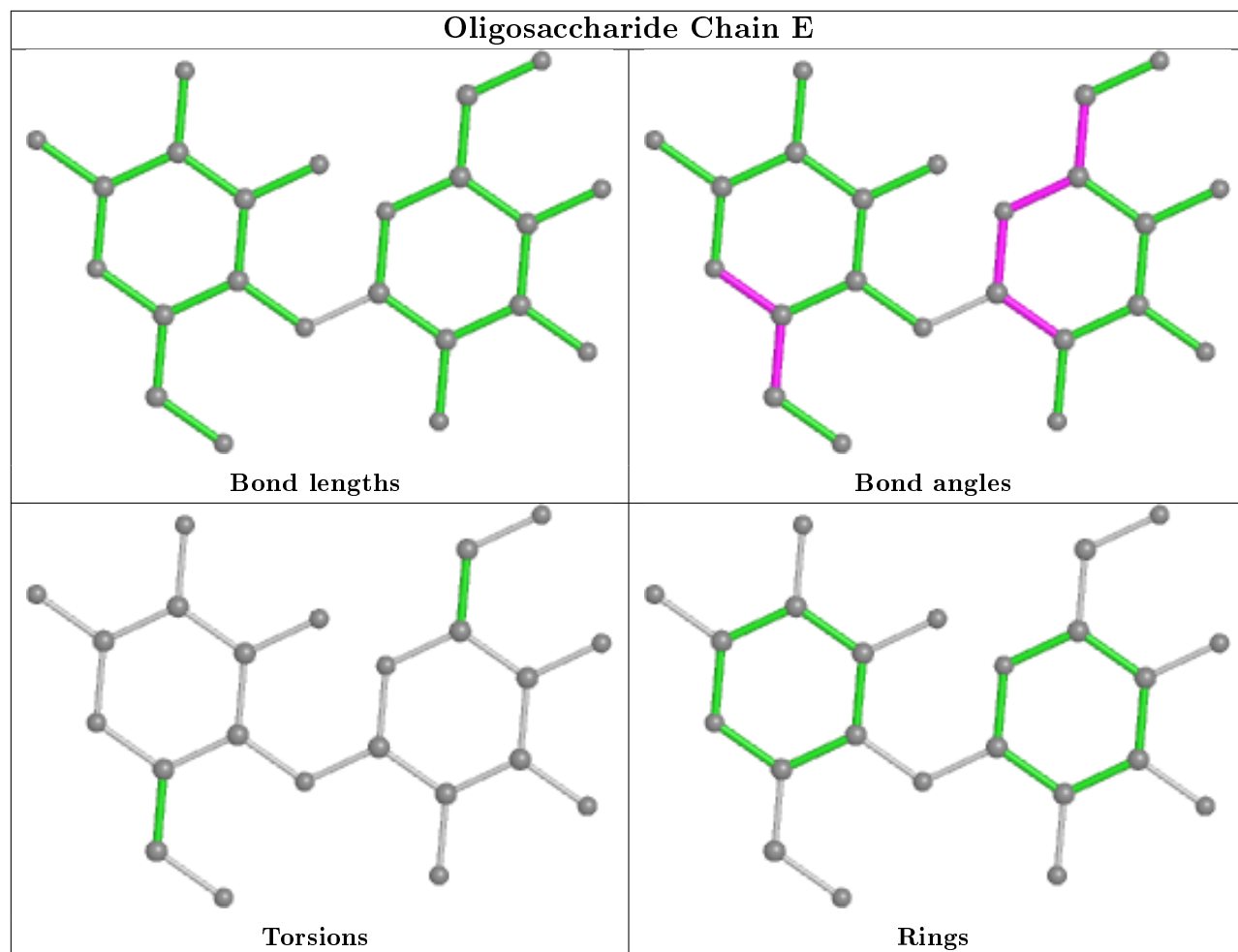
10 monomers are involved in 19 short contacts:

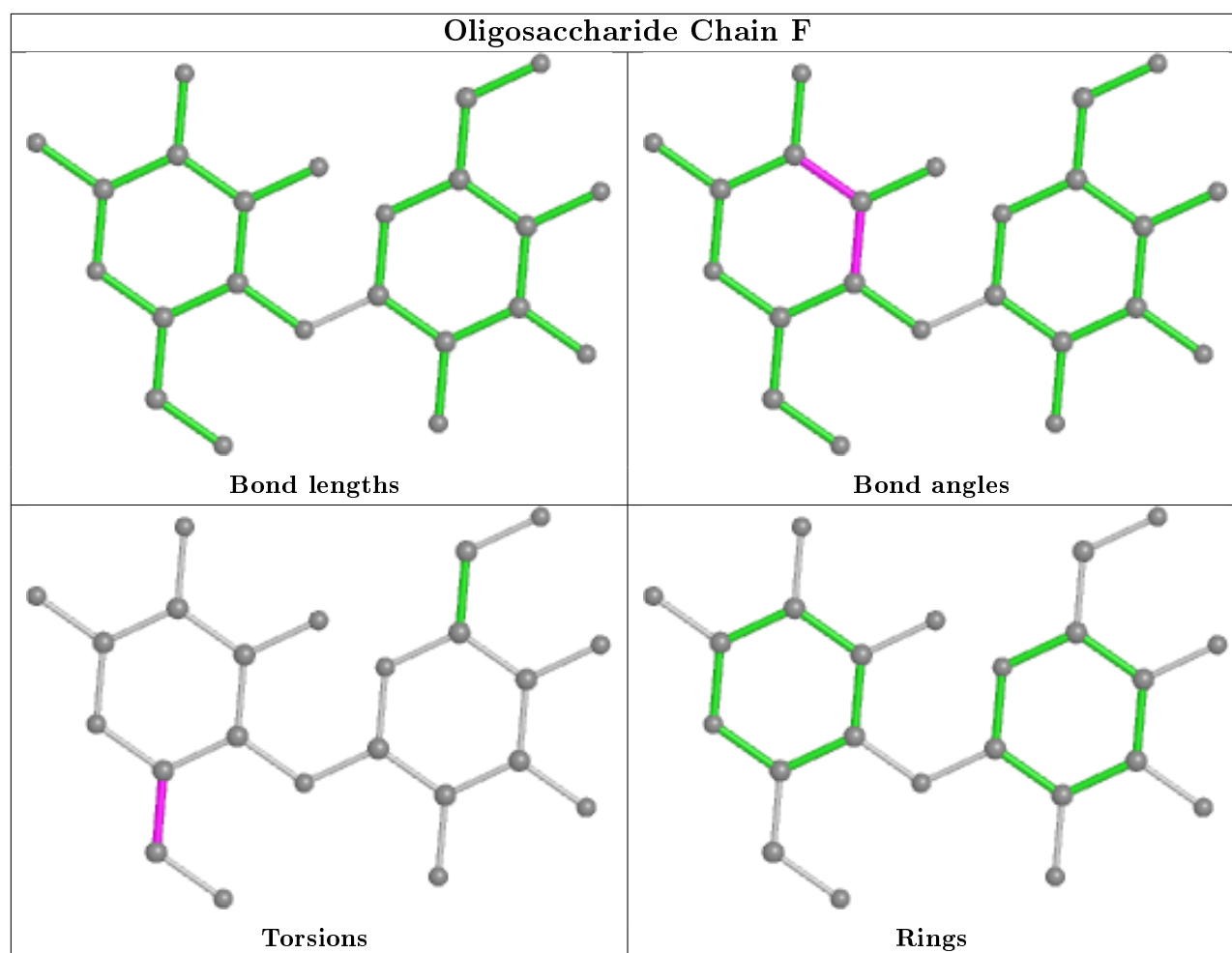
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	2	GLC	2	0
4	E	1	BGC	3	0
4	D	1	BGC	3	0
4	B	1	BGC	4	0
4	B	2	GLC	5	0
4	C	2	GLC	5	0
4	F	1	BGC	1	0
4	C	1	BGC	5	0
4	F	2	GLC	2	0
4	D	2	GLC	2	0

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









## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 3 are monoatomic - leaving 4 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	$\# Z  > 2$
8	CLR	A	609	-	31,31,31	0.74	0	48,48,48	1.16	5 (10%)
5	Y01	A	601	-	35,38,38	4.43	13 (37%)	54,57,57	1.99	15 (27%)
6	42J	A	602	-	10,11,11	1.21	2 (20%)	13,14,14	0.61	0
7	NAG	A	603	1	14,14,15	0.58	0	17,19,21	0.46	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
8	CLR	A	609	-	-	2/10/68/68	0/4/4/4
5	Y01	A	601	-	-	5/17/77/77	0/4/4/4
6	42J	A	602	-	-	2/3/3/3	0/1/1/1
7	NAG	A	603	1	-	0/6/23/26	0/1/1/1

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	601	Y01	CAI-CAZ	17.03	1.70	1.33
5	A	601	Y01	CBB-CBE	-10.25	1.36	1.54
5	A	601	Y01	CBH-CBF	7.67	1.68	1.56
5	A	601	Y01	CAU-CBI	-7.02	1.41	1.54
5	A	601	Y01	CAP-CBE	5.63	1.66	1.54

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	601	Y01	CBI-CBE-CBB	-5.65	110.64	119.49
5	A	601	Y01	CBH-CAZ-CAI	-4.07	116.67	122.90
5	A	601	Y01	CAU-CBI-CBG	4.03	113.53	107.27
5	A	601	Y01	OAW-CAY-CAM	3.93	119.97	111.50
5	A	601	Y01	CBG-CBI-CBE	3.56	104.30	100.07

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

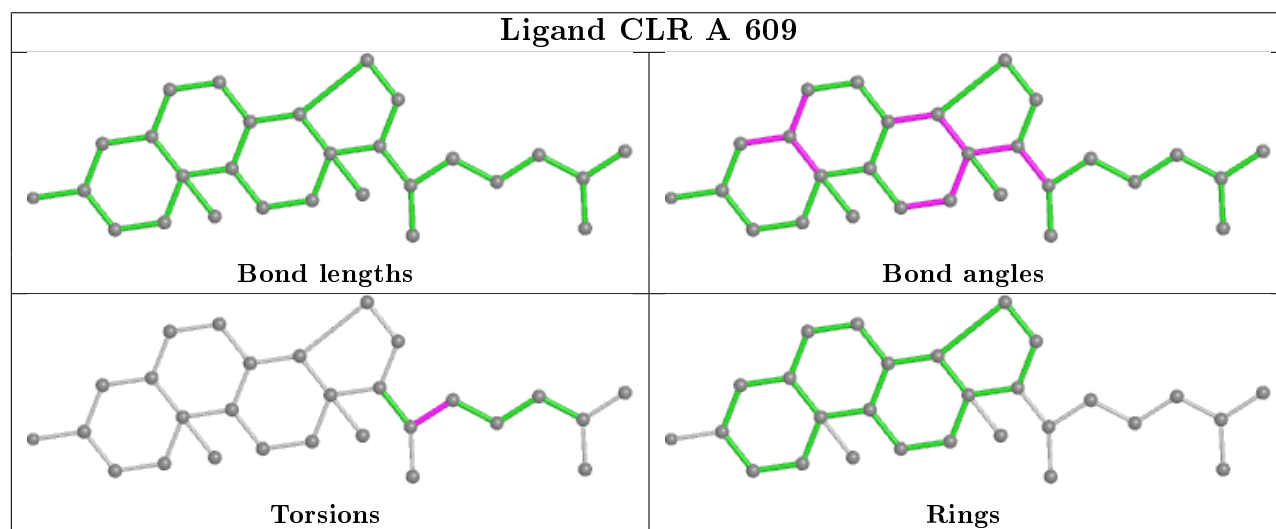
Mol	Chain	Res	Type	Atoms
5	A	601	Y01	CAX-CAL-CAM-CAY
8	A	609	CLR	C17-C20-C22-C23
8	A	609	CLR	C21-C20-C22-C23
5	A	601	Y01	CAM-CAY-OAW-CBC
5	A	601	Y01	OAG-CAY-OAW-CBC

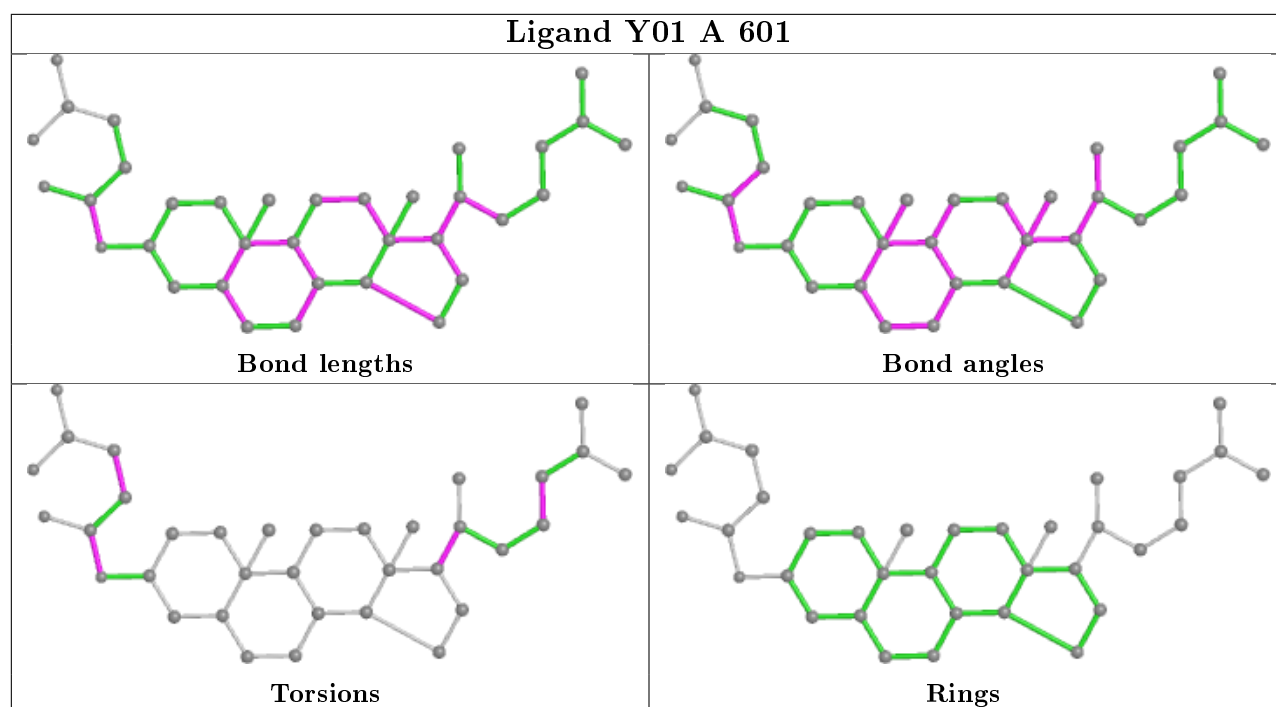
There are no ring outliers.

2 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	609	CLR	6	0
5	A	601	Y01	5	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 [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	A	535/535 (100%)	0.51	47 (8%) 10 11	76, 102, 143, 171	0
2	L	214/214 (100%)	0.24	8 (3%) 41 43	71, 96, 132, 176	0
3	H	219/219 (100%)	0.23	6 (2%) 54 57	75, 96, 143, 194	0
All	All	968/968 (100%)	0.39	61 (6%) 20 22	71, 99, 141, 194	0

The worst 5 of 61 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	304	ILE	7.6
1	A	583	PRO	5.4
1	A	307	ALA	5.2
1	A	582	THR	5.1
1	A	301	PHE	4.8

### 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.

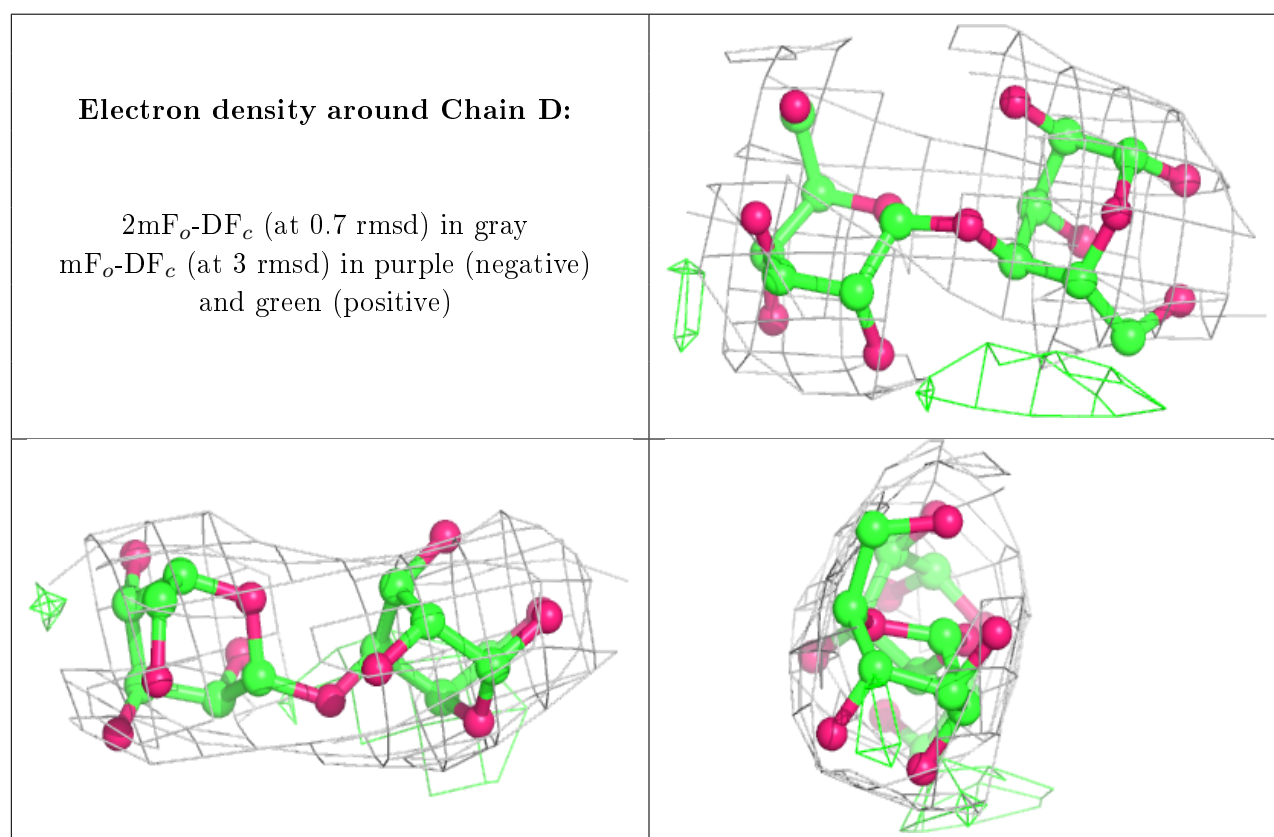
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	GLC	F	2	11/12	0.52	0.39	158,188,198,200	0
4	BGC	F	1	12/12	0.62	0.44	166,194,198,201	0
4	GLC	E	2	11/12	0.75	0.19	157,170,186,186	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	BGC	C	1	12/12	0.79	0.24	176,186,189,192	0
4	GLC	B	2	11/12	0.82	0.35	152,173,176,178	0
4	GLC	D	2	11/12	0.83	0.20	142,160,169,170	0
4	GLC	C	2	11/12	0.84	0.35	131,176,182,187	0
4	BGC	E	1	12/12	0.86	0.26	183,197,202,205	0
4	BGC	B	1	12/12	0.89	0.29	146,172,178,189	0
4	BGC	D	1	12/12	0.89	0.21	150,162,173,177	0

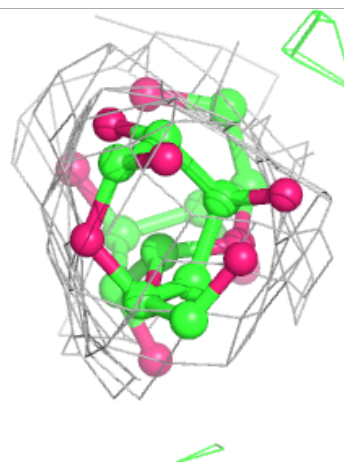
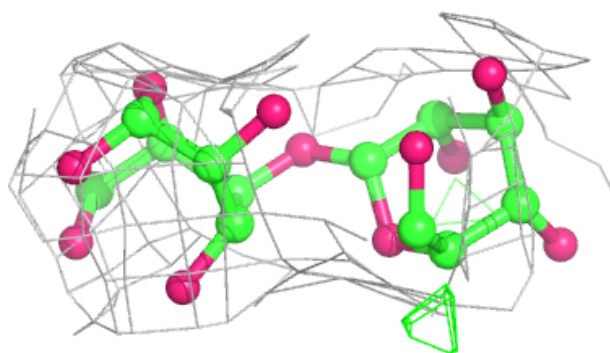
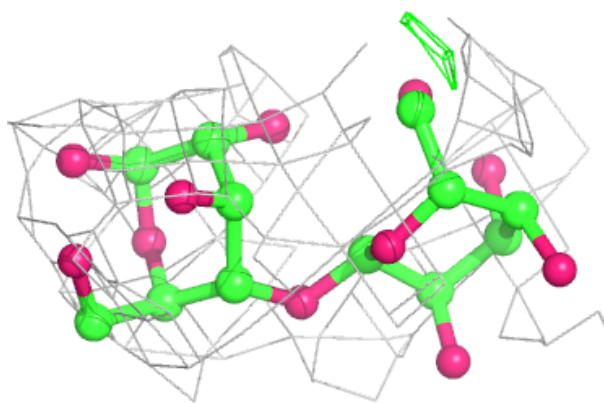
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

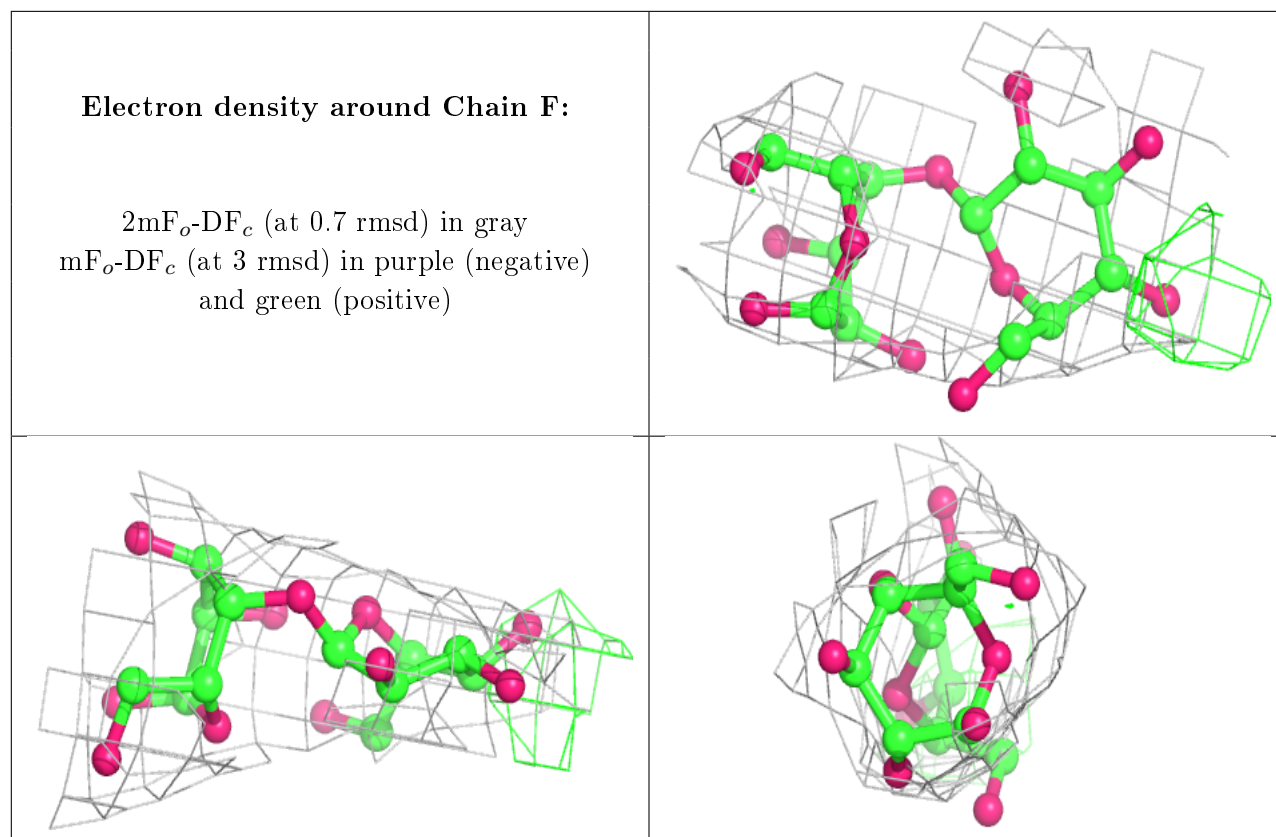




**Electron density around Chain E:**

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





## 6.4 Ligands ⓘ

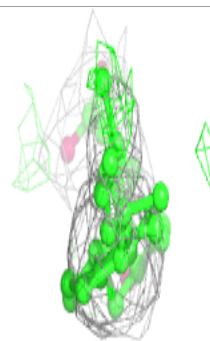
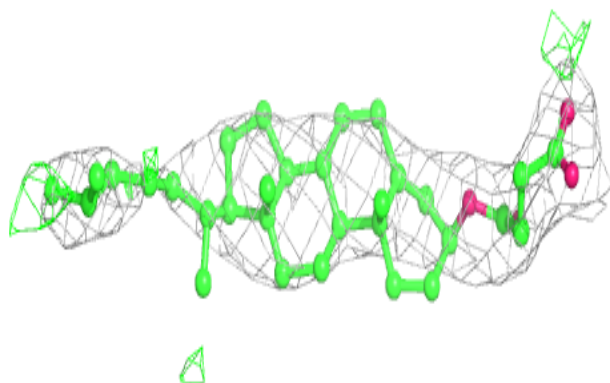
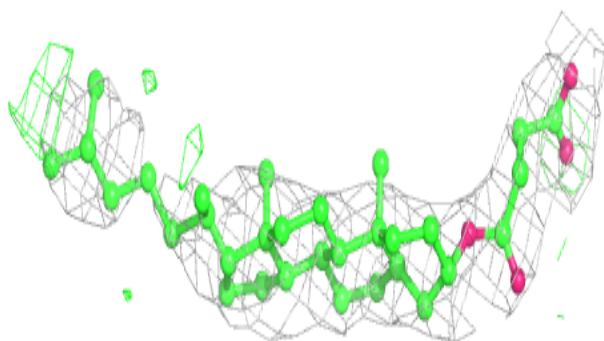
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
7	NAG	A	603	14/15	0.75	0.43	138,173,203,208	0
5	Y01	A	601	35/35	0.78	0.41	84,121,135,139	0
10	NA	A	612	1/1	0.89	0.53	91,91,91,91	0
10	NA	A	611	1/1	0.90	0.67	89,89,89,89	0
8	CLR	A	609	28/28	0.92	0.53	92,102,108,110	0
6	42J	A	602	11/11	0.94	0.40	85,90,105,122	0
9	CL	A	610	1/1	0.95	0.24	101,101,101,101	0

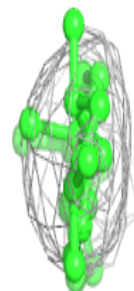
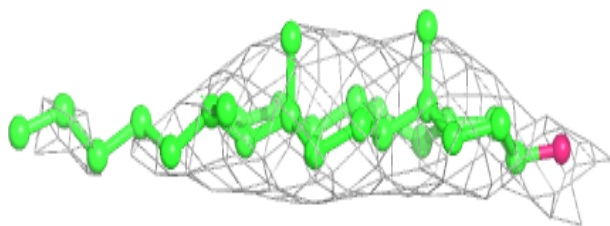
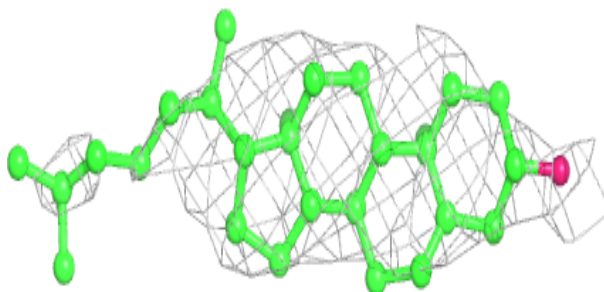
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around Y01 A 601:**

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

**Electron density around CLR A 609:**

$2mF_o-DF_c$  (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.