



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 23, 2023 – 08:45 AM EDT

PDB ID : 3EE5  
Title : Crystal structure of human M340H-Beta1,4-Galactosyltransferase-I (M340H-B4GAL-T1) in complex with GLCNAC-Beta1,3-Gal-Beta-Naphthalenemethanol  
Authors : Ramakrishnan, B.; Qasba, P.K.  
Deposited on : 2008-09-04  
Resolution : 2.20 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 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)  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

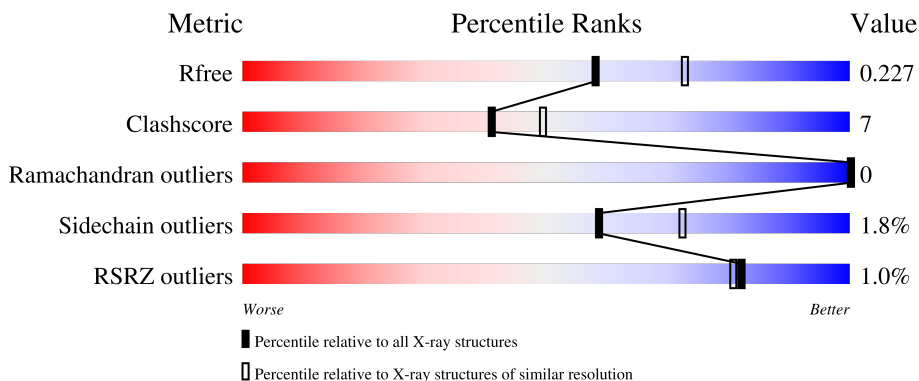
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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.



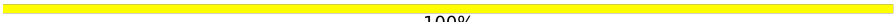
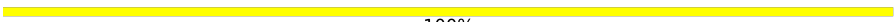
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	A	287	
1	B	287	
1	C	287	
2	D	2	

Continued on next page...

*Continued from previous page...*

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

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
7	GOL	B	415	-	-	-	X
7	GOL	B	416	-	-	X	-

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 7469 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 Beta-1,4-galactosyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	273	2218	1420	386	401	11	0	0	0
1	B	273	2218	1420	386	401	11	0	0	0
1	C	273	2218	1420	386	401	11	0	0	0

There are 51 discrepancies between the modelled and reference sequences:

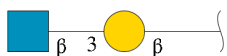
Chain	Residue	Modelled	Actual	Comment	Reference
A	112	ALA	-	expression tag	UNP P15291
A	113	SER	-	expression tag	UNP P15291
A	114	MET	-	expression tag	UNP P15291
A	115	THR	-	expression tag	UNP P15291
A	116	GLY	-	expression tag	UNP P15291
A	117	GLY	-	expression tag	UNP P15291
A	118	GLN	-	expression tag	UNP P15291
A	119	GLN	-	expression tag	UNP P15291
A	120	MET	-	expression tag	UNP P15291
A	121	GLY	-	expression tag	UNP P15291
A	122	ARG	-	expression tag	UNP P15291
A	123	GLY	-	expression tag	UNP P15291
A	124	SER	-	expression tag	UNP P15291
A	125	ALA	-	expression tag	UNP P15291
A	337	THR	ARG	engineered mutation	UNP P15291
A	338	THR	CYS	engineered mutation	UNP P15291
A	340	HIS	MET	engineered mutation	UNP P15291
B	112	ALA	-	expression tag	UNP P15291
B	113	SER	-	expression tag	UNP P15291
B	114	MET	-	expression tag	UNP P15291
B	115	THR	-	expression tag	UNP P15291
B	116	GLY	-	expression tag	UNP P15291
B	117	GLY	-	expression tag	UNP P15291

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	118	GLN	-	expression tag	UNP P15291
B	119	GLN	-	expression tag	UNP P15291
B	120	MET	-	expression tag	UNP P15291
B	121	GLY	-	expression tag	UNP P15291
B	122	ARG	-	expression tag	UNP P15291
B	123	GLY	-	expression tag	UNP P15291
B	124	SER	-	expression tag	UNP P15291
B	125	ALA	-	expression tag	UNP P15291
B	337	THR	ARG	engineered mutation	UNP P15291
B	338	THR	CYS	engineered mutation	UNP P15291
B	340	HIS	MET	engineered mutation	UNP P15291
C	112	ALA	-	expression tag	UNP P15291
C	113	SER	-	expression tag	UNP P15291
C	114	MET	-	expression tag	UNP P15291
C	115	THR	-	expression tag	UNP P15291
C	116	GLY	-	expression tag	UNP P15291
C	117	GLY	-	expression tag	UNP P15291
C	118	GLN	-	expression tag	UNP P15291
C	119	GLN	-	expression tag	UNP P15291
C	120	MET	-	expression tag	UNP P15291
C	121	GLY	-	expression tag	UNP P15291
C	122	ARG	-	expression tag	UNP P15291
C	123	GLY	-	expression tag	UNP P15291
C	124	SER	-	expression tag	UNP P15291
C	125	ALA	-	expression tag	UNP P15291
C	337	THR	ARG	engineered mutation	UNP P15291
C	338	THR	CYS	engineered mutation	UNP P15291
C	340	HIS	MET	engineered mutation	UNP P15291

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose.



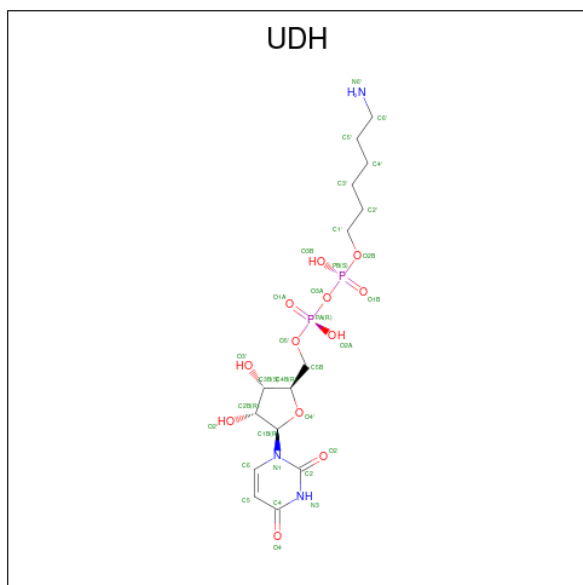
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	2	Total	C	N	O	0	0	0
			25	14	1	10			
2	E	2	Total	C	N	O	0	0	0
			25	14	1	10			

Continued on next page...

Continued from previous page...

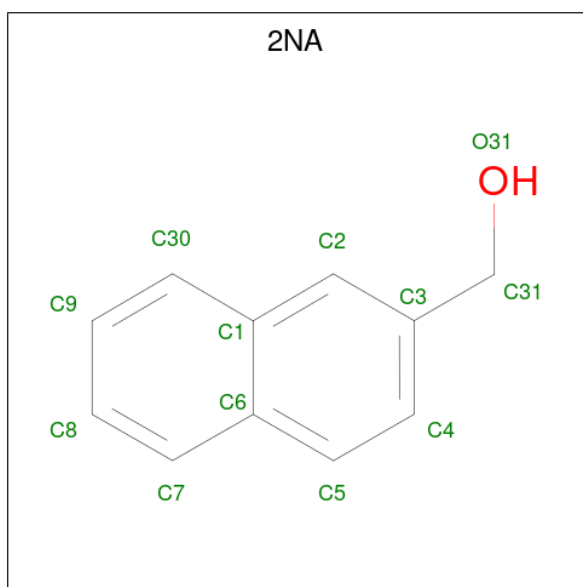
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	F	2	25	14	1	10	0	0	0

- Molecule 3 is 6-AMINOHEXYL-URIDINE-C1,5'-DIPHOSPHATE (three-letter code: UDH) (formula:  $C_{15}H_{27}N_3O_{12}P_2$ ).



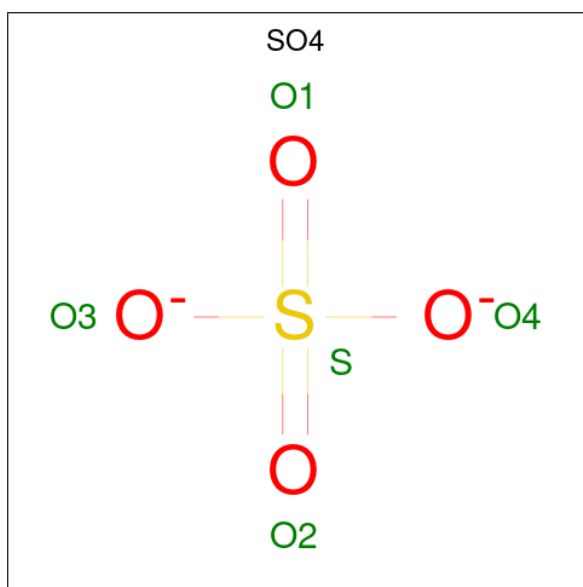
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	32	15	3	12	2	0	0
3	B	1	32	15	3	12	2	0	0
3	C	1	32	15	3	12	2	0	0

- Molecule 4 is naphthalen-2-ylmethanol (three-letter code: 2NA) (formula:  $C_{11}H_{10}O$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			12	11	1		
4	B	1	Total	C	O	0	0
			12	11	1		
4	C	1	Total	C	O	0	0
			12	11	1		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Mn	0	0
			1	1		
6	B	1	Total	Mn	0	0
			1	1		
6	C	1	Total	Mn	0	0
			1	1		

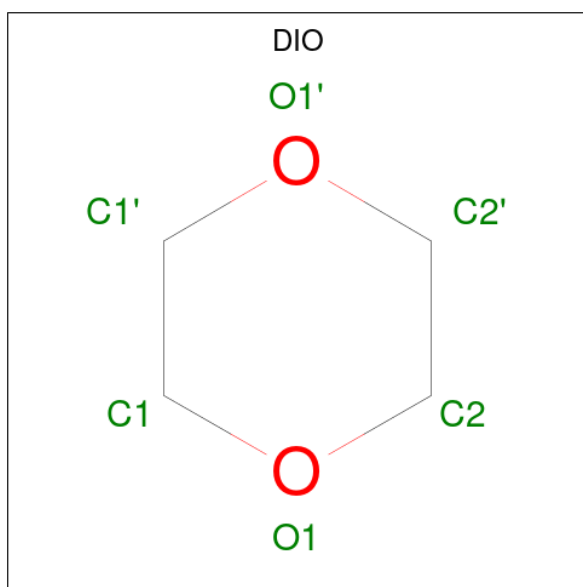
- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).





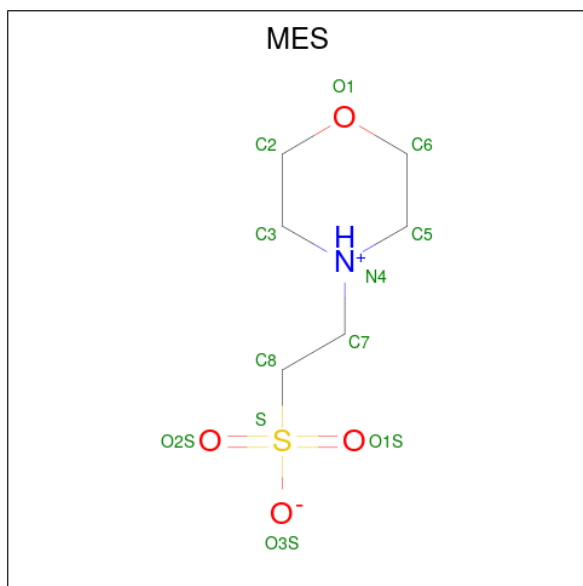
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 6 3 3	0	0
7	A	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	C	1	Total C O 6 3 3	0	0

- Molecule 8 is 1,4-DIETHYLENE DIOXIDE (three-letter code: DIO) (formula: C<sub>4</sub>H<sub>8</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	C	1	Total	C	O	0	0
			6	4	2		

- Molecule 9 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula:  $C_6H_{13}NO_4S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
9	C	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

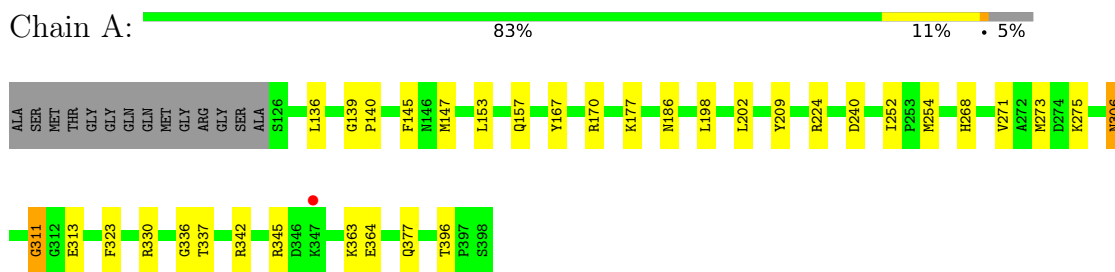
- Molecule 10 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
10	A	152	Total 152	O 152	0	0
10	B	192	Total 192	O 192	0	0
10	C	131	Total 131	O 131	0	0

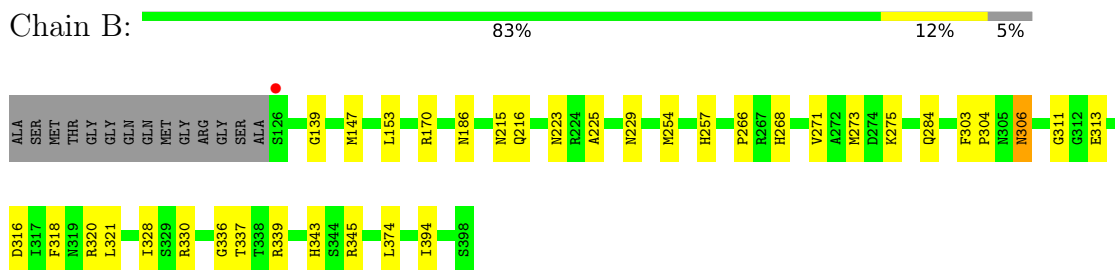
### 3 Residue-property plots [i](#)

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 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.

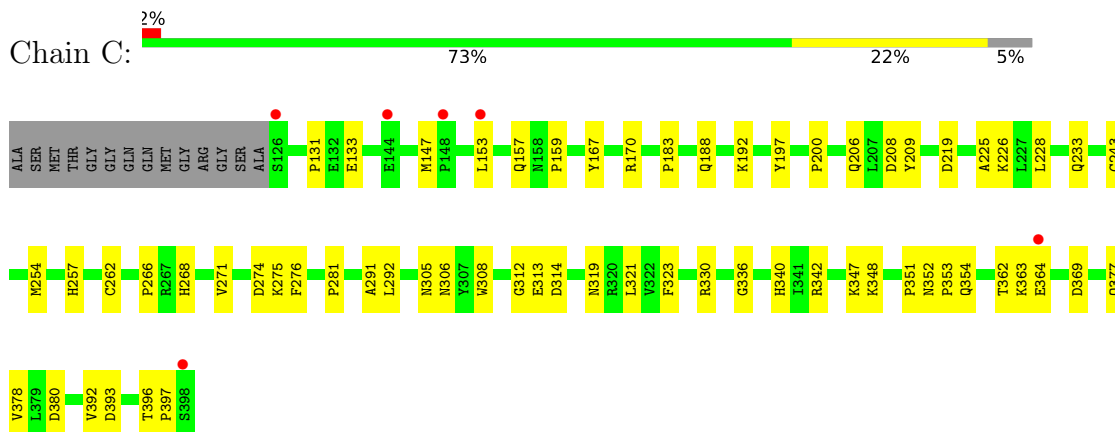
- Molecule 1: Beta-1,4-galactosyltransferase 1



- Molecule 1: Beta-1,4-galactosyltransferase 1



- Molecule 1: Beta-1,4-galactosyltransferase 1



- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose



GAL1  
MAG2

- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose

Chain E:

100%

GAL1  
MAG2

- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose

Chain F:

100%

GAL1  
MAG2

## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.13Å 195.74Å 143.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.33 – 2.20 46.33 – 2.20	Depositor EDS
% Data completeness (in resolution range)	98.4 (46.33-2.20) 98.4 (46.33-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.99 (at 2.20Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.206 , 0.240 0.191 , 0.227	Depositor DCC
$R_{free}$ test set	7621 reflections (10.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.0	Xtrriage
Anisotropy	0.659	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 43.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.000 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.009 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7469	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.96% 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: SO4, GAL, GOL, 2NA, MN, UDH, MES, NAG, DIO

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	0.61	0/2280	0.78	1/3097 (0.0%)
1	B	0.66	0/2280	0.77	0/3097
1	C	0.55	0/2280	0.73	2/3097 (0.1%)
All	All	0.61	0/6840	0.76	3/9291 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	208	ASP	N-CA-C	-5.25	96.82	111.00
1	A	311	GLY	N-CA-C	5.13	125.92	113.10
1	C	312	GLY	N-CA-C	5.12	125.89	113.10

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	2218	0	2162	31	0
1	B	2218	0	2162	34	0
1	C	2218	0	2162	39	0
2	D	25	0	22	0	0
2	E	25	0	22	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	25	0	22	0	0
3	A	32	0	25	0	0
3	B	32	0	25	1	0
3	C	32	0	25	0	0
4	A	12	0	9	0	0
4	B	12	0	9	0	0
4	C	12	0	9	0	0
5	A	25	0	0	0	0
5	B	35	0	0	1	0
5	C	10	0	0	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
7	A	12	0	16	4	0
7	B	24	0	31	10	0
7	C	6	0	8	1	0
8	C	6	0	8	1	0
9	C	12	0	13	2	0
10	A	152	0	0	0	0
10	B	192	0	0	1	0
10	C	131	0	0	0	0
All	All	7469	0	6730	97	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 7.

The worst 5 of 97 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:186:ASN:HD21	7:B:416:GOL:H11	1.29	0.96
1:B:313:GLU:OE1	7:B:414:GOL:H31	1.71	0.90
1:A:170:ARG:HD2	1:C:342:ARG:HE	1.42	0.84
1:B:186:ASN:ND2	7:B:416:GOL:H11	1.92	0.84
1:B:313:GLU:OE2	7:B:414:GOL:H11	1.81	0.81

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	271/287 (94%)	263 (97%)	8 (3%)	0	100	100
1	B	271/287 (94%)	267 (98%)	4 (2%)	0	100	100
1	C	271/287 (94%)	255 (94%)	16 (6%)	0	100	100
All	All	813/861 (94%)	785 (97%)	28 (3%)	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	244/252 (97%)	240 (98%)	4 (2%)	62	76
1	B	244/252 (97%)	241 (99%)	3 (1%)	71	83
1	C	244/252 (97%)	238 (98%)	6 (2%)	47	60
All	All	732/756 (97%)	719 (98%)	13 (2%)	59	72

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	133	GLU
1	C	147	MET
1	C	380	ASP
1	C	257	HIS
1	C	364	GLU

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

Mol	Chain	Res	Type
1	C	206	GLN
1	C	257	HIS
1	C	377	GLN
1	A	377	GLN
1	B	186	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](#)

6 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	GAL	D	1	4,2	11,11,12	1.51	2 (18%)	15,15,17	1.23	2 (13%)
2	NAG	D	2	2	14,14,15	1.00	1 (7%)	17,19,21	0.93	0
2	GAL	E	1	4,2	11,11,12	1.24	1 (9%)	15,15,17	1.02	1 (6%)
2	NAG	E	2	2	14,14,15	1.06	1 (7%)	17,19,21	0.90	1 (5%)
2	GAL	F	1	4,2	11,11,12	1.09	2 (18%)	15,15,17	0.82	1 (6%)
2	NAG	F	2	2	14,14,15	1.10	1 (7%)	17,19,21	1.03	2 (11%)

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	GAL	D	1	4,2	-	0/2/19/22	0/1/1/1
2	NAG	D	2	2	-	0/6/23/26	0/1/1/1
2	GAL	E	1	4,2	-	0/2/19/22	0/1/1/1
2	NAG	E	2	2	-	0/6/23/26	0/1/1/1
2	GAL	F	1	4,2	-	0/2/19/22	0/1/1/1
2	NAG	F	2	2	-	0/6/23/26	0/1/1/1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1	GAL	O2-C2	3.33	1.50	1.43
2	F	2	NAG	C4-C5	2.57	1.58	1.53
2	E	2	NAG	C4-C5	2.47	1.58	1.53
2	D	1	GAL	C2-C3	2.33	1.55	1.52
2	E	1	GAL	C4-C5	2.24	1.57	1.53

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	1	GAL	C1-O5-C5	3.24	116.58	112.19
2	D	1	GAL	C1-C2-C3	-2.56	106.52	109.67
2	E	1	GAL	C1-O5-C5	2.38	115.42	112.19
2	F	2	NAG	C8-C7-N2	-2.36	112.10	116.10
2	F	1	GAL	C1-O5-C5	2.31	115.32	112.19

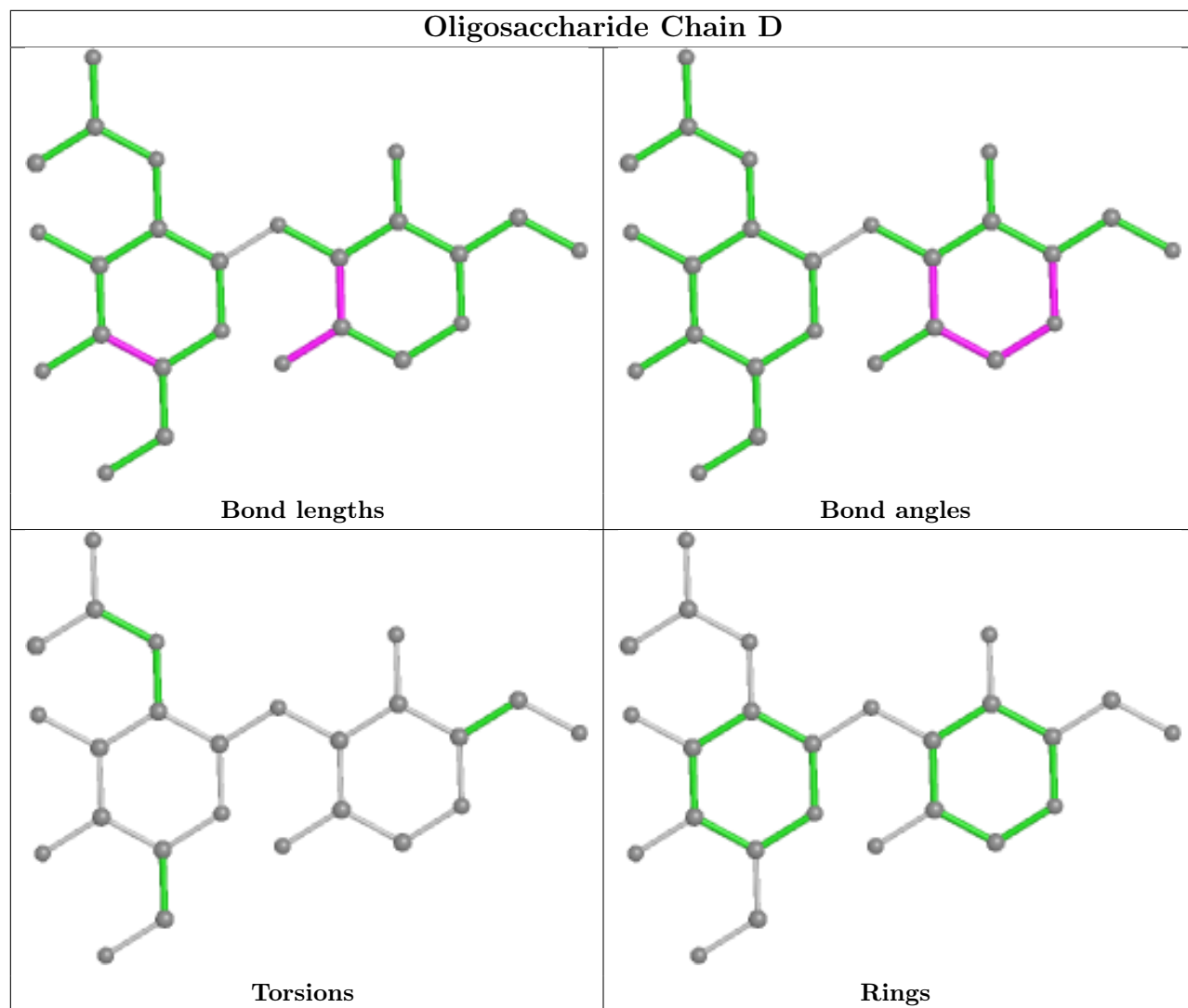
There are no chirality outliers.

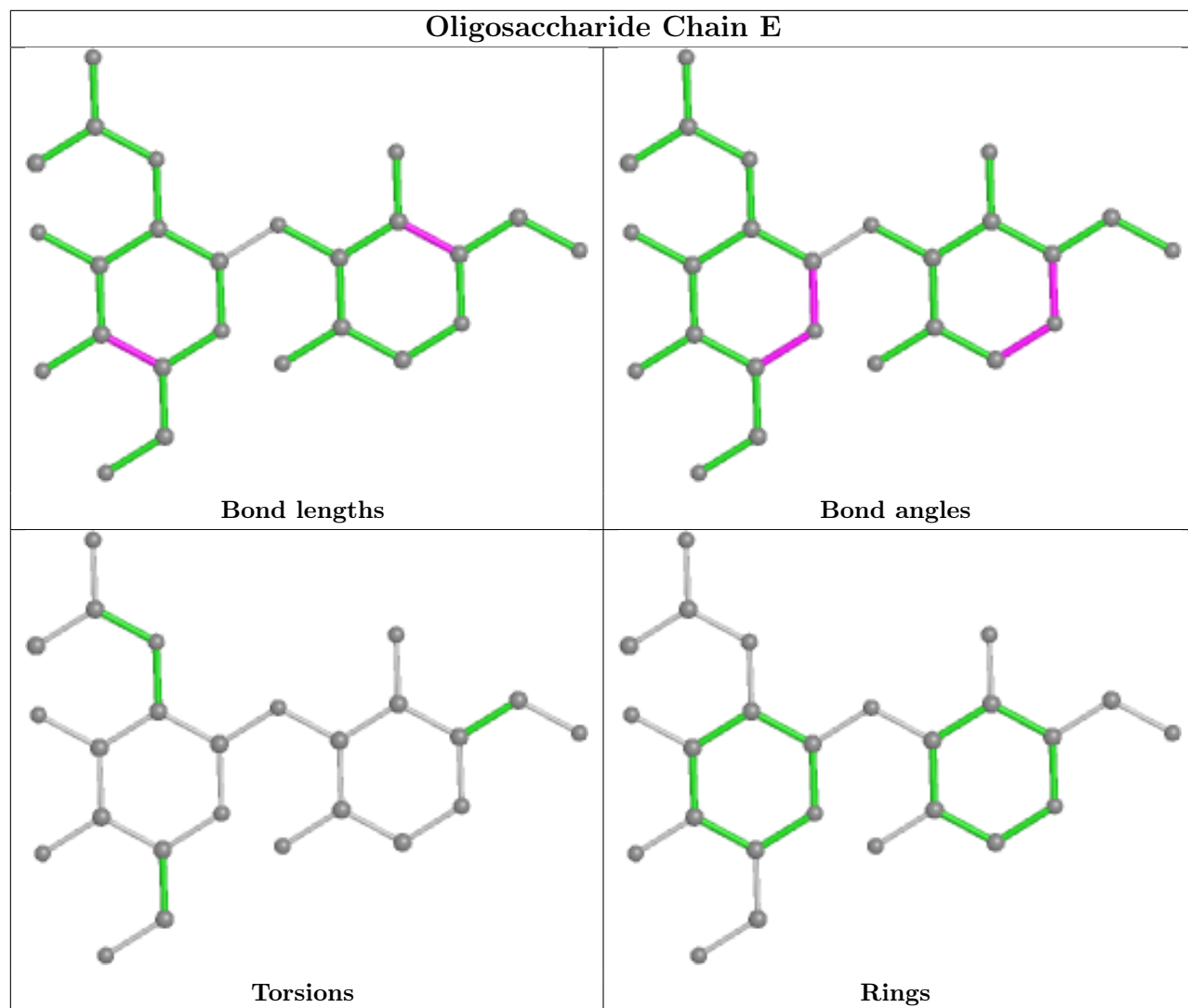
There are no torsion outliers.

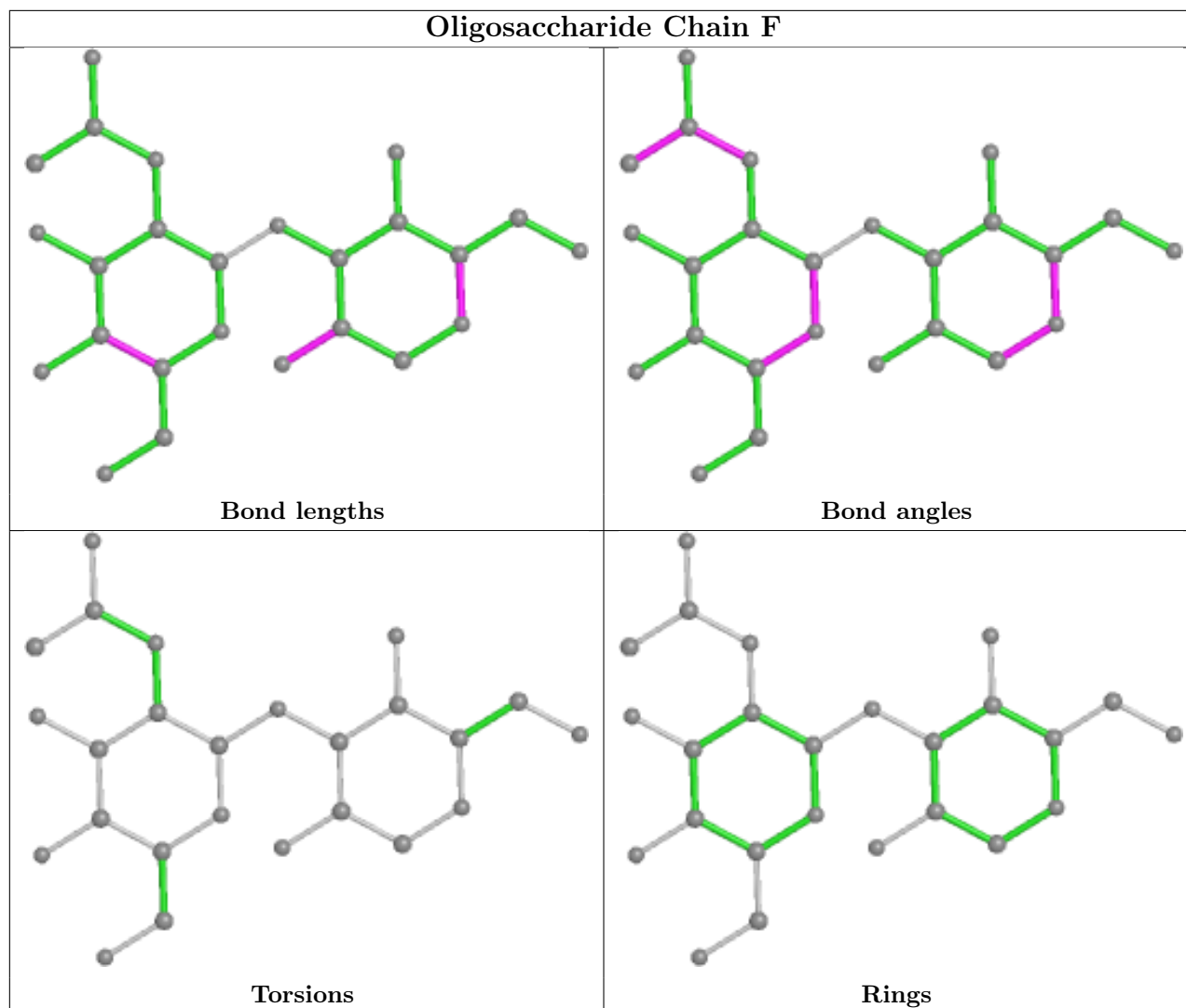
There are no ring outliers.

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 [i](#)

Of 32 ligands modelled in this entry, 3 are monoatomic - leaving 29 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	$\# Z  > 2$
5	SO4	A	609	-	4,4,4	1.96	2 (50%)	6,6,6	0.80	0
5	SO4	B	613	-	4,4,4	2.00	1 (25%)	6,6,6	0.88	0
7	GOL	C	418	-	5,5,5	0.49	0	5,5,5	0.54	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	SO4	B	606	-	4,4,4	1.93	1 (25%)	6,6,6	0.84	0
7	GOL	B	415	-	5,5,5	0.26	0	5,5,5	0.58	0
3	UDH	A	399	6	31,33,33	1.28	3 (9%)	41,47,47	1.73	13 (31%)
5	SO4	B	603	-	4,4,4	1.82	2 (50%)	6,6,6	0.86	0
5	SO4	B	605	-	4,4,4	1.90	1 (25%)	6,6,6	0.90	0
5	SO4	A	610	-	4,4,4	1.90	1 (25%)	6,6,6	0.86	0
7	GOL	A	420	-	5,5,5	0.42	0	5,5,5	0.25	0
5	SO4	A	602	-	4,4,4	1.83	2 (50%)	6,6,6	0.92	0
4	2NA	A	100	2	13,13,13	3.65	4 (30%)	17,17,17	0.68	0
5	SO4	B	614	-	4,4,4	1.98	1 (25%)	6,6,6	0.90	0
7	GOL	B	416	-	5,5,5	0.72	0	5,5,5	0.56	0
5	SO4	C	608	-	4,4,4	1.95	2 (50%)	6,6,6	0.85	0
4	2NA	C	102	2	13,13,13	3.54	4 (30%)	17,17,17	0.59	0
5	SO4	C	612	-	4,4,4	1.89	1 (25%)	6,6,6	0.92	0
5	SO4	A	607	-	4,4,4	1.98	2 (50%)	6,6,6	0.78	0
4	2NA	B	101	2	13,13,13	3.82	5 (38%)	17,17,17	0.57	0
3	UDH	B	1399	6	31,33,33	1.36	5 (16%)	41,47,47	1.74	13 (31%)
7	GOL	B	414	-	5,5,5	1.03	0	5,5,5	0.72	0
3	UDH	C	2399	6	31,33,33	1.18	3 (9%)	41,47,47	1.64	11 (26%)
9	MES	C	436	-	12,12,12	1.32	2 (16%)	14,16,16	0.79	0
7	GOL	A	417	-	5,5,5	0.58	0	5,5,5	0.68	0
8	DIO	C	435	-	6,6,6	0.94	0	6,6,6	1.66	2 (33%)
5	SO4	A	604	-	4,4,4	1.95	1 (25%)	6,6,6	0.92	0
5	SO4	B	601	-	4,4,4	1.86	1 (25%)	6,6,6	0.94	0
5	SO4	B	611	-	4,4,4	1.97	1 (25%)	6,6,6	0.90	0
7	GOL	B	419	-	5,5,5	0.54	0	5,5,5	0.68	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	DIO	C	435	-	-	-	0/1/1/1
4	2NA	C	102	2	-	0/2/2/2	0/2/2/2
7	GOL	B	415	-	-	0/4/4/4	-
4	2NA	B	101	2	-	1/2/2/2	0/2/2/2
3	UDH	B	1399	6	-	5/25/41/41	0/2/2/2

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	GOL	A	420	-	-	0/4/4/4	-
7	GOL	B	414	-	-	0/4/4/4	-
3	UDH	C	2399	6	-	8/25/41/41	0/2/2/2
7	GOL	B	419	-	-	0/4/4/4	-
7	GOL	C	418	-	-	0/4/4/4	-
4	2NA	A	100	2	-	0/2/2/2	0/2/2/2
9	MES	C	436	-	-	2/6/14/14	0/1/1/1
7	GOL	B	416	-	-	0/4/4/4	-
7	GOL	A	417	-	-	0/4/4/4	-
3	UDH	A	399	6	-	4/25/41/41	0/2/2/2

The worst 5 of 45 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	101	2NA	C2-C3	8.40	1.57	1.37
4	C	102	2NA	C2-C3	8.05	1.56	1.37
4	A	100	2NA	C2-C3	7.87	1.55	1.37
4	A	100	2NA	C8-C7	7.76	1.54	1.36
4	B	101	2NA	C8-C7	7.70	1.54	1.36

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1399	UDH	O4'-C4B-C3B	-4.38	96.44	105.11
3	A	399	UDH	O4'-C4B-C3B	-4.33	96.54	105.11
3	B	1399	UDH	PA-O5'-C5B	3.72	143.51	121.68
3	C	2399	UDH	O4'-C4B-C3B	-3.46	98.27	105.11
3	A	399	UDH	PA-O5'-C5B	3.26	140.77	121.68

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	399	UDH	C5B-O5'-PA-O1A
3	A	399	UDH	C5B-O5'-PA-O2A
3	A	399	UDH	C5B-O5'-PA-O3A
3	B	1399	UDH	C5B-O5'-PA-O1A
3	B	1399	UDH	C5B-O5'-PA-O2A

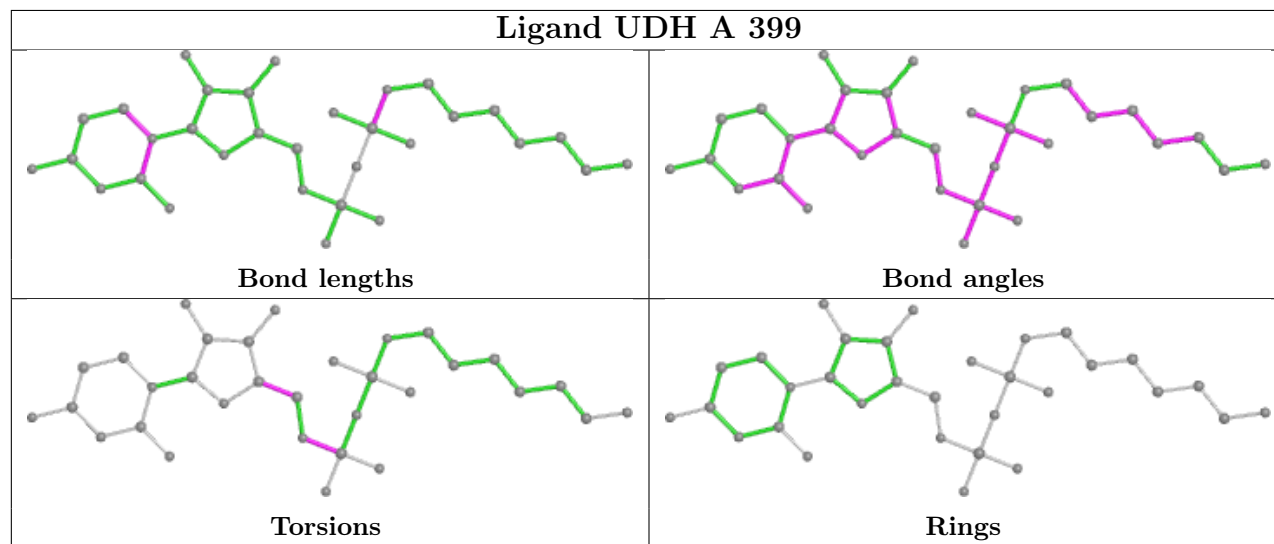
There are no ring outliers.

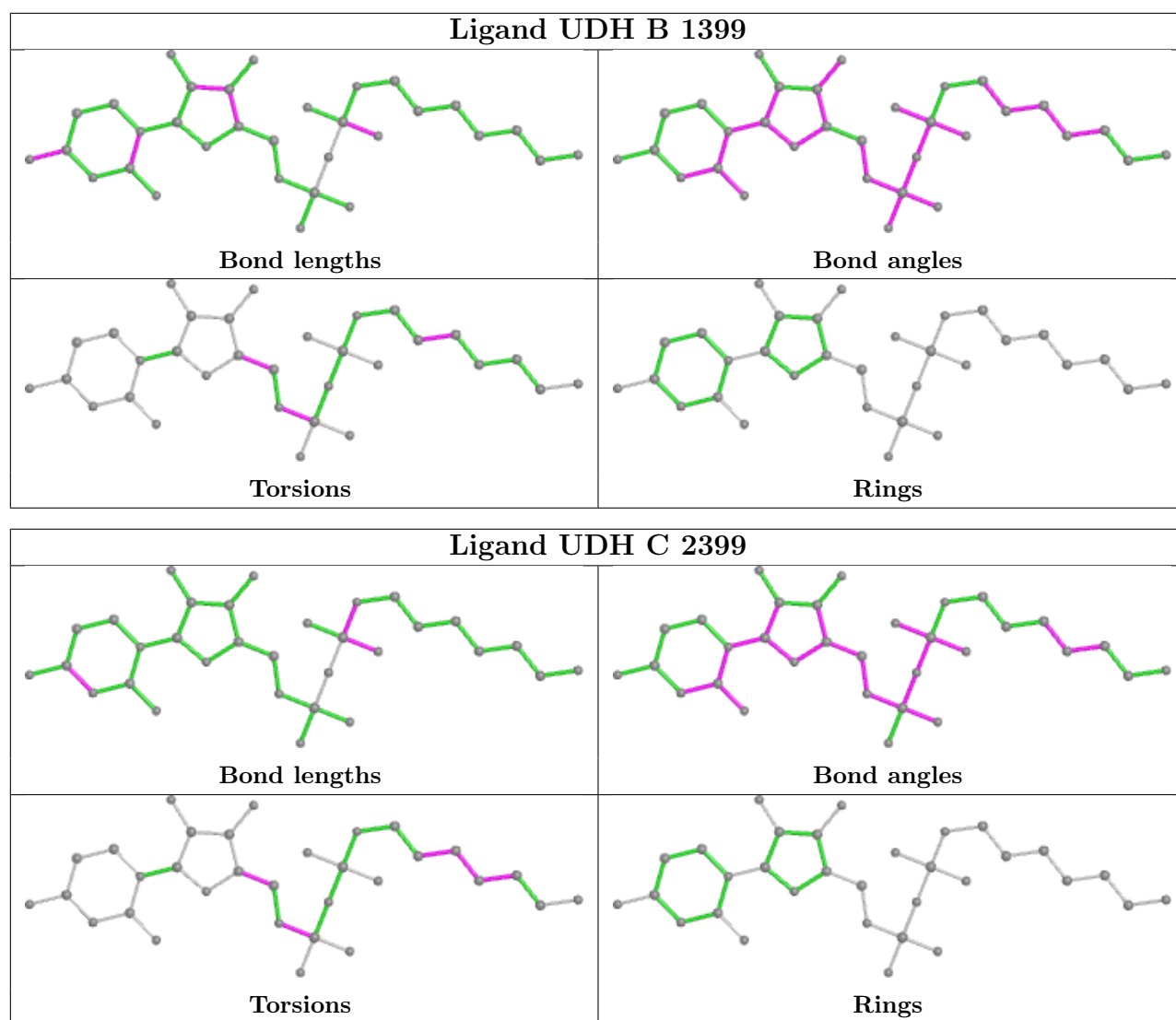
10 monomers are involved in 20 short contacts:



Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	C	418	GOL	1	0
7	B	415	GOL	1	0
7	A	420	GOL	1	0
7	B	416	GOL	6	0
3	B	1399	UDH	1	0
7	B	414	GOL	3	0
9	C	436	MES	2	0
7	A	417	GOL	3	0
8	C	435	DIO	1	0
5	B	611	SO4	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 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	273/287 (95%)	-0.47	1 (0%) 92 91	23, 32, 51, 63	0
1	B	273/287 (95%)	-0.45	1 (0%) 92 91	22, 30, 44, 56	0
1	C	273/287 (95%)	-0.09	6 (2%) 62 59	29, 41, 61, 70	0
All	All	819/861 (95%)	-0.34	8 (0%) 82 81	22, 34, 55, 70	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	126	SER	3.2
1	C	126	SER	3.1
1	C	148	PRO	2.9
1	C	364	GLU	2.6
1	C	398	SER	2.6

### 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.

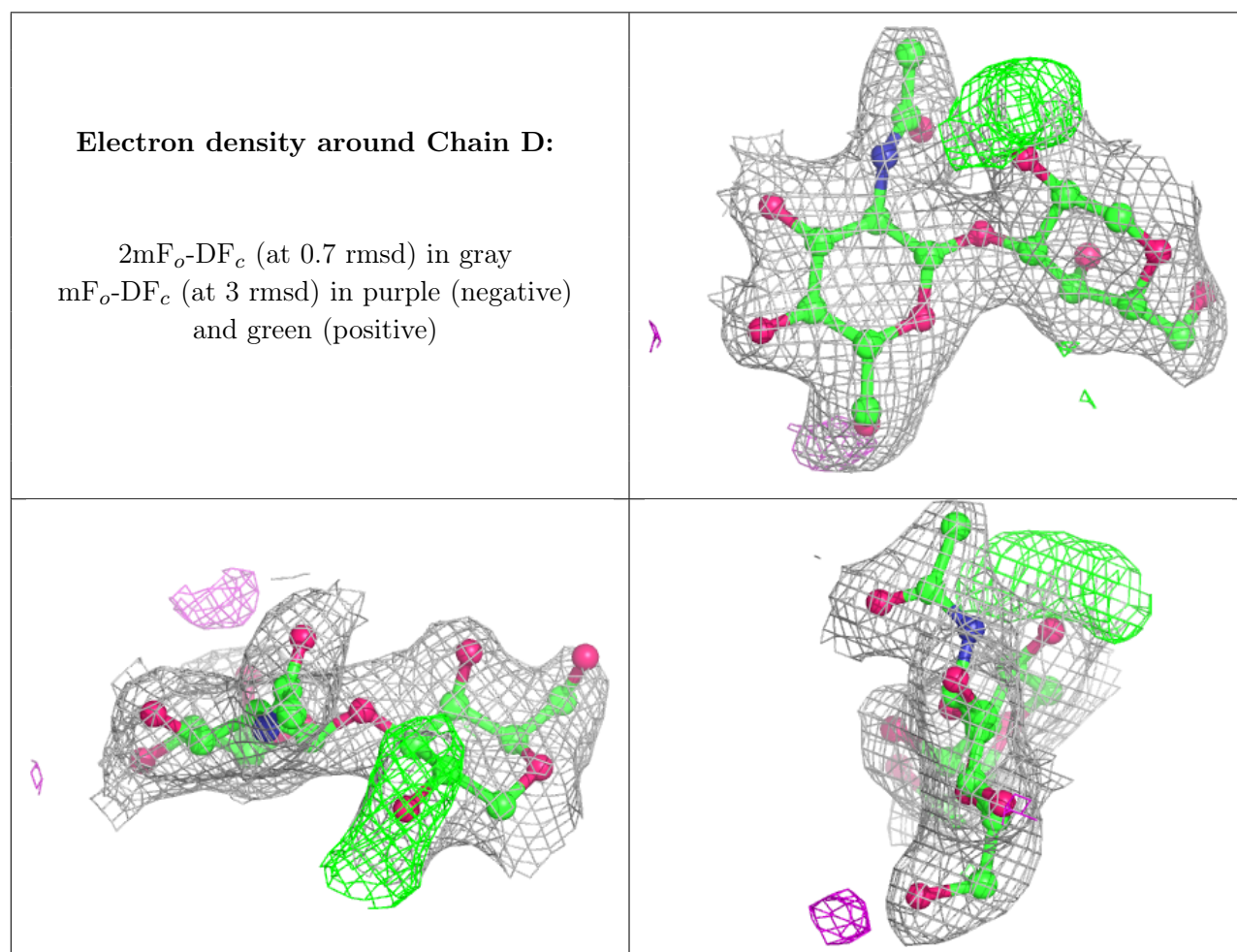
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	GAL	D	1	11/12	0.84	0.14	45,52,57,61	0
2	GAL	E	1	11/12	0.84	0.23	42,52,58,62	0
2	GAL	F	1	11/12	0.85	0.15	59,64,68,71	0

*Continued on next page...*

Continued from previous page...

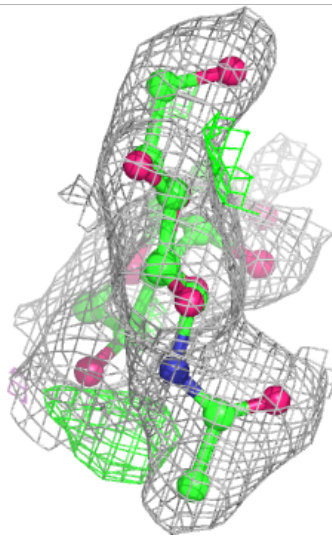
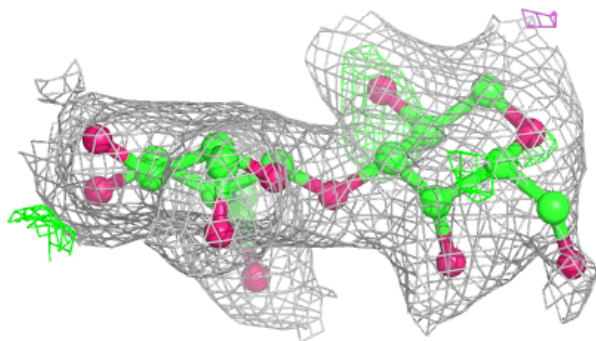
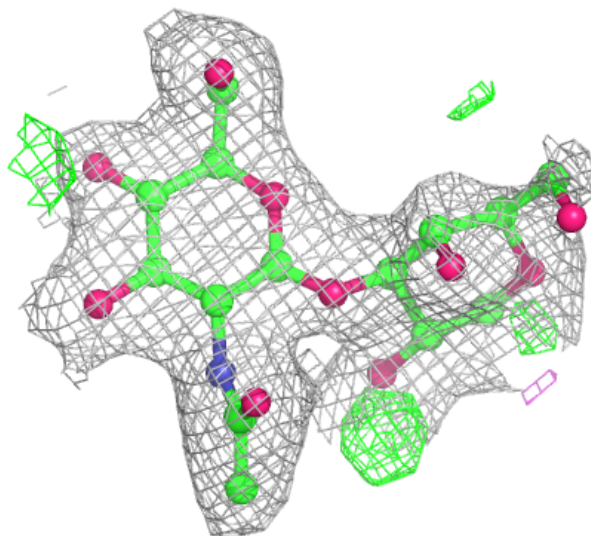
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	F	2	14/15	0.92	0.10	49,51,54,54	0
2	NAG	D	2	14/15	0.95	0.10	37,38,41,43	0
2	NAG	E	2	14/15	0.96	0.09	26,32,34,35	0

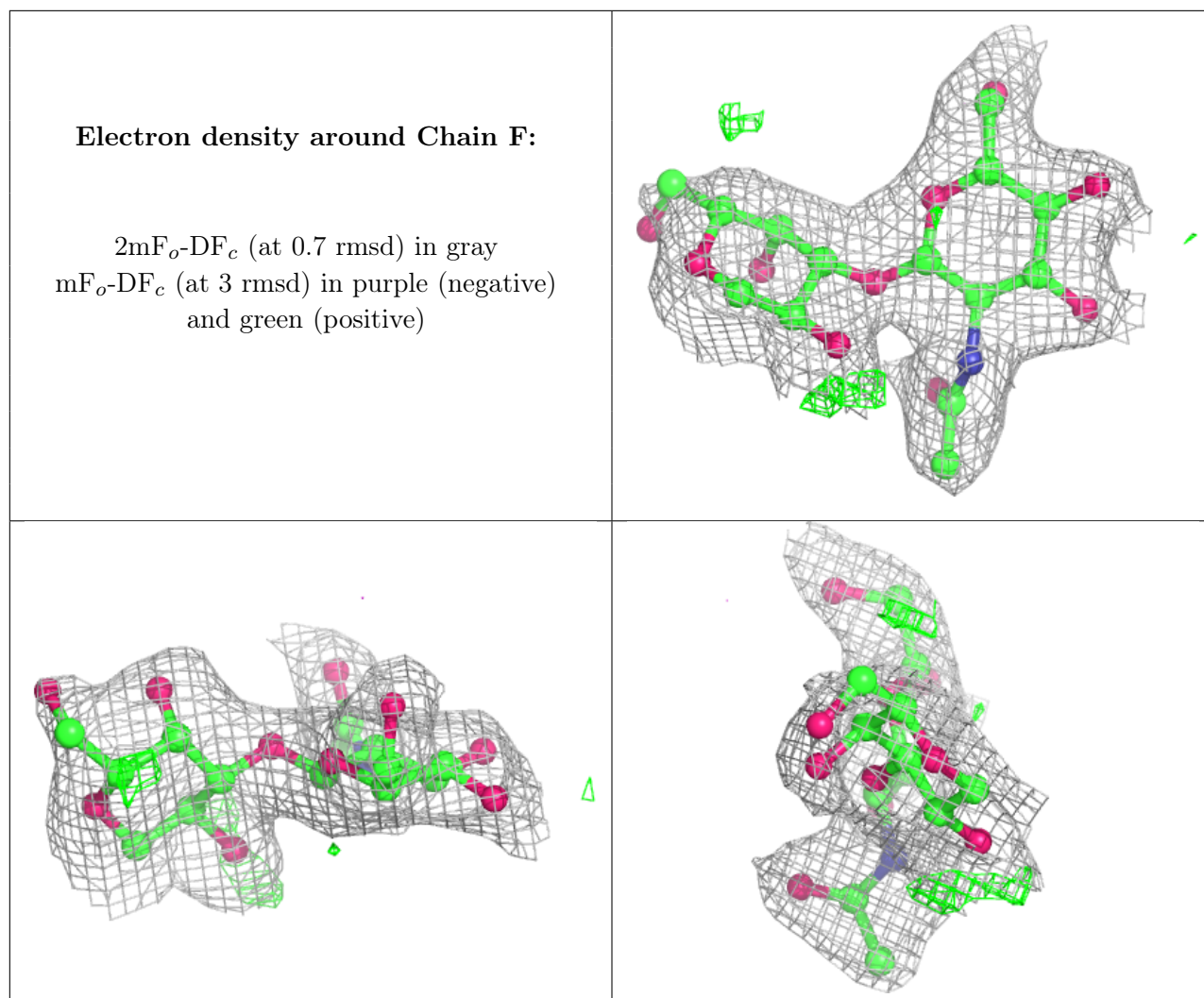
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



**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 [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
9	MES	C	436	12/12	0.55	0.31	90,92,99,100	0
7	GOL	A	420	6/6	0.69	0.33	68,71,71,72	0
7	GOL	B	415	6/6	0.73	0.43	70,71,74,74	0
7	GOL	B	419	6/6	0.74	0.24	56,61,62,64	0
5	SO4	B	614	5/5	0.76	0.40	87,88,89,90	0
7	GOL	A	417	6/6	0.76	0.27	59,61,62,63	0
7	GOL	C	418	6/6	0.78	0.29	57,57,58,59	0
4	2NA	B	101	12/12	0.79	0.28	57,59,59,60	0
5	SO4	A	607	5/5	0.79	0.22	80,83,83,84	0

*Continued on next page...*

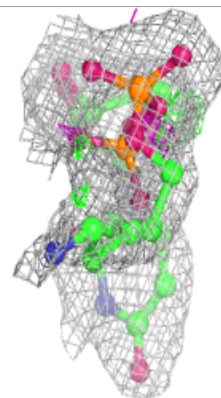
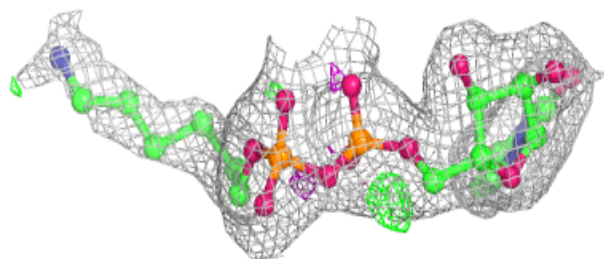
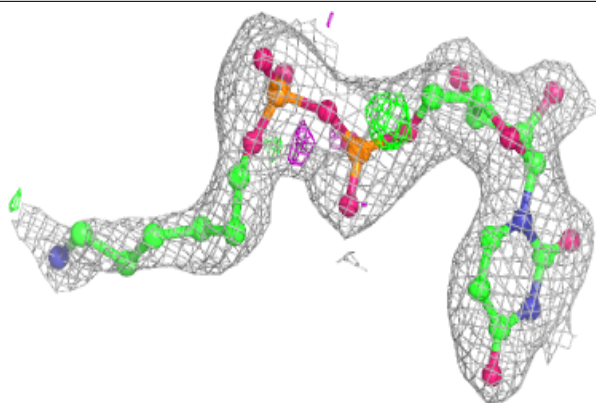
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	GOL	B	416	6/6	0.79	0.30	36,38,44,50	0
7	GOL	B	414	6/6	0.82	0.25	43,49,50,53	0
8	DIO	C	435	6/6	0.84	0.16	67,69,69,69	0
5	SO4	A	609	5/5	0.88	0.25	80,81,82,83	0
4	2NA	C	102	12/12	0.89	0.28	64,65,66,66	0
5	SO4	B	611	5/5	0.89	0.42	86,86,86,87	0
4	2NA	A	100	12/12	0.90	0.18	48,51,52,54	0
5	SO4	B	613	5/5	0.92	0.15	78,78,80,80	0
5	SO4	B	606	5/5	0.92	0.23	78,78,80,80	0
5	SO4	C	608	5/5	0.92	0.23	83,84,84,84	0
5	SO4	A	604	5/5	0.92	0.10	75,76,76,76	0
5	SO4	B	603	5/5	0.93	0.26	74,75,76,76	0
5	SO4	A	602	5/5	0.94	0.19	75,75,76,77	0
3	UDH	C	2399	32/32	0.94	0.14	37,46,59,63	0
5	SO4	A	610	5/5	0.94	0.27	85,86,86,86	0
5	SO4	C	612	5/5	0.95	0.21	85,86,86,86	0
6	MN	C	400	1/1	0.97	0.07	48,48,48,48	0
3	UDH	B	1399	32/32	0.97	0.12	23,28,56,59	0
5	SO4	B	605	5/5	0.97	0.13	83,83,83,83	0
3	UDH	A	399	32/32	0.97	0.11	28,34,57,60	0
5	SO4	B	601	5/5	0.97	0.13	74,75,75,75	0
6	MN	B	400	1/1	0.99	0.09	30,30,30,30	0
6	MN	A	611	1/1	1.00	0.07	38,38,38,38	0

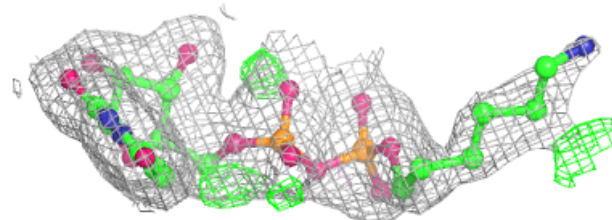
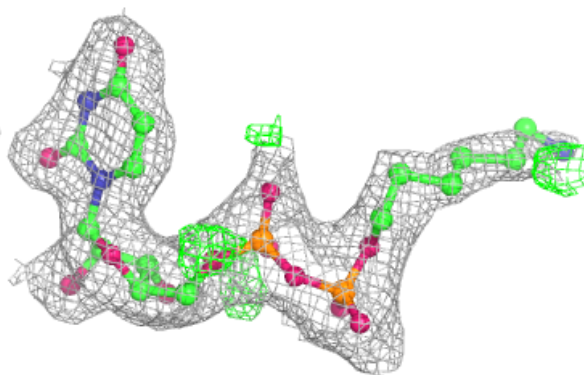
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around UDH C 2399:**

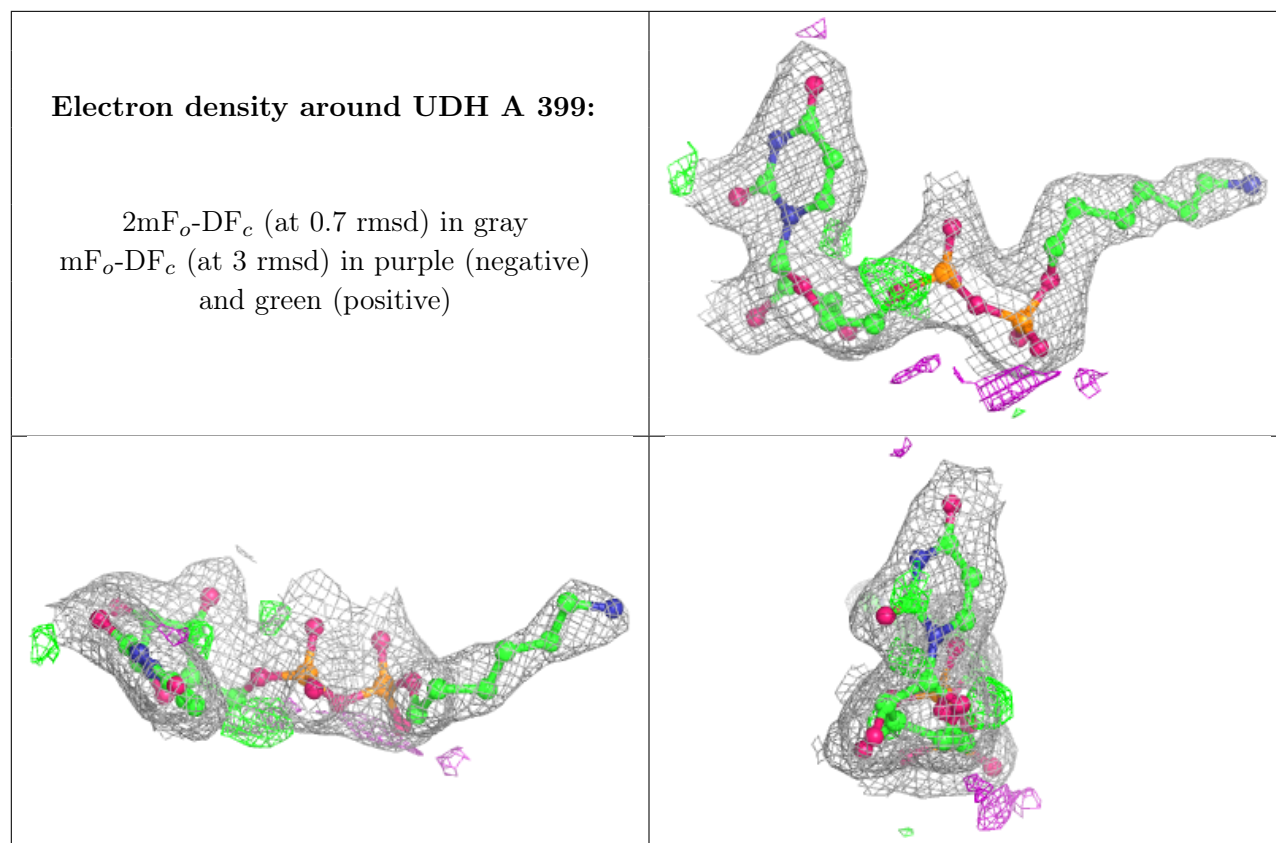
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around UDH B 1399:**

$2mF_o-DF_c$  (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.