



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 8, 2023 – 01:26 PM EDT

PDB ID : 4RET  
Title : Crystal structure of the Na,K-ATPase E2P-digoxin complex with bound magnesium  
Authors : Gregersen, J.L.; Laursen, M.; Yatime, L.; Nissen, P.; Fedosova, N.U.  
Deposited on : 2014-09-23  
Resolution : 4.00 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)  
Xtrriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

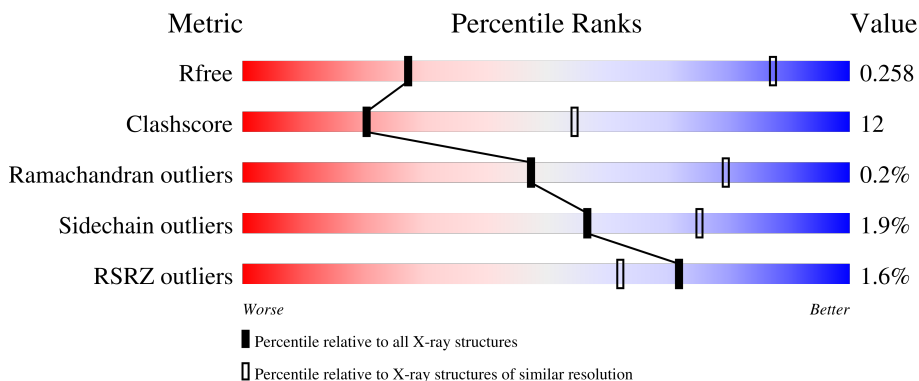
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.00 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.



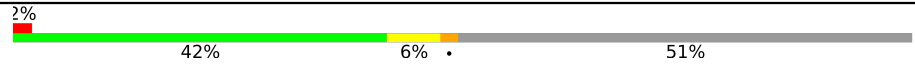


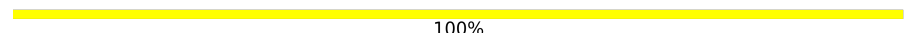
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1087 (4.30-3.70)
Clashscore	141614	1148 (4.30-3.70)
Ramachandran outliers	138981	1108 (4.30-3.70)
Sidechain outliers	138945	1099 (4.30-3.70)
RSRZ outliers	127900	1028 (4.34-3.66)

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	1021	
1	C	1021	
2	B	303	
2	D	303	
3	E	65	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	65	 2% 42% 6% 51%
4	F	2	 50% 50%
5	H	2	 50% 50%
5	I	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
5	GLC	H	1	-	-	X	-
9	17F	B	1004	-	-	-	X

## 2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 21123 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 Sodium/potassium-transporting ATPase subunit alpha-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	P				S
1	A	996	7730	4922	1301	1459	1	47	0	0	0
1	C	996	7730	4922	1301	1459	1	47	0	0	0

- Molecule 2 is a protein called Sodium/potassium-transporting ATPase subunit beta-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	290	2377	1540	388	436	13	0	0	0
2	D	288	2357	1525	385	434	13	0	0	0

- Molecule 3 is a protein called Na<sup>+</sup>/K<sup>+</sup> ATPase gamma subunit transcript variant a.

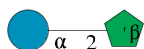
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	G	32	255	174	37	44	0	0	0
3	E	32	255	174	37	44	0	0	0

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	F	2	28	16	2	10	0	0	0

- Molecule 5 is an oligosaccharide called beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose.

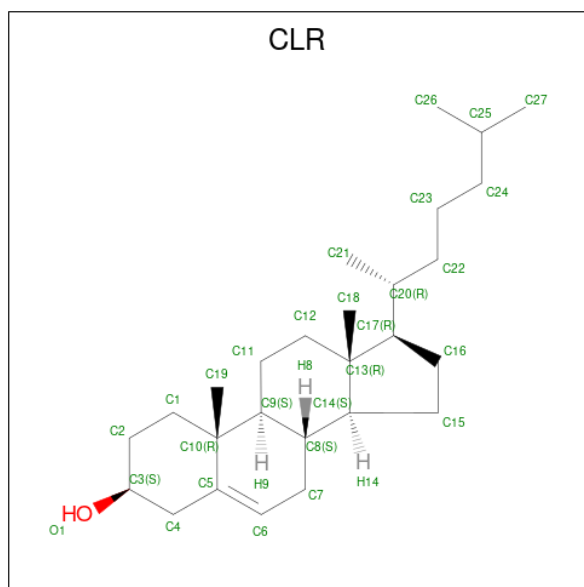


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
5	H	2	Total	C	O	0	0	0
			23	12	11			
5	I	2	Total	C	O	0	0	0
			23	12	11			

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	3	Total	Mg	0	0
			3	3		
6	C	3	Total	Mg	0	0
			3	3		

- Molecule 7 is CHOLESTEROL (three-letter code: CLR) (formula: C<sub>27</sub>H<sub>46</sub>O).



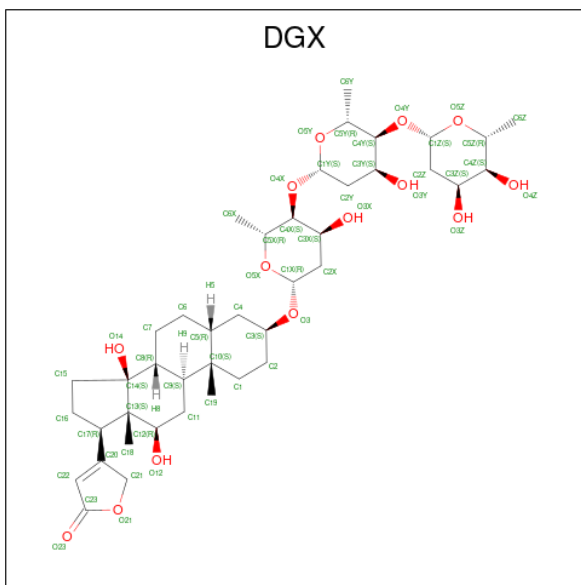
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			28	27	1		
7	G	1	Total	C	O	0	0
			28	27	1		

*Continued on next page...*

Continued from previous page...

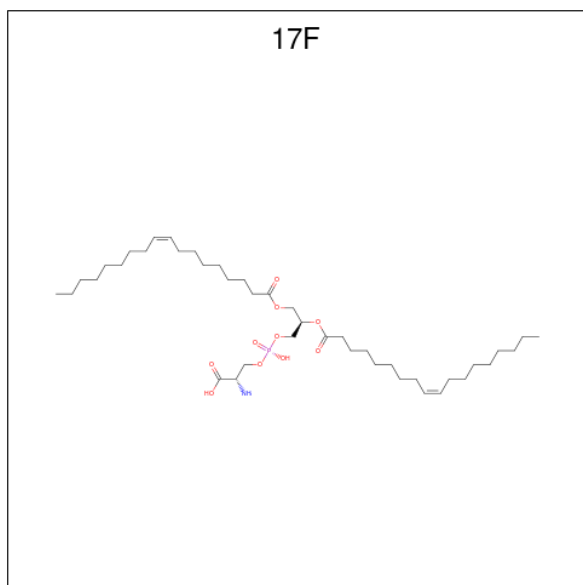
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	C	1	Total	C	O	0	0
			28	27	1		
7	C	1	Total	C	O	0	0
			28	27	1		

- Molecule 8 is DIGOXIN (three-letter code: DGX) (formula:  $C_{41}H_{64}O_{14}$ ).



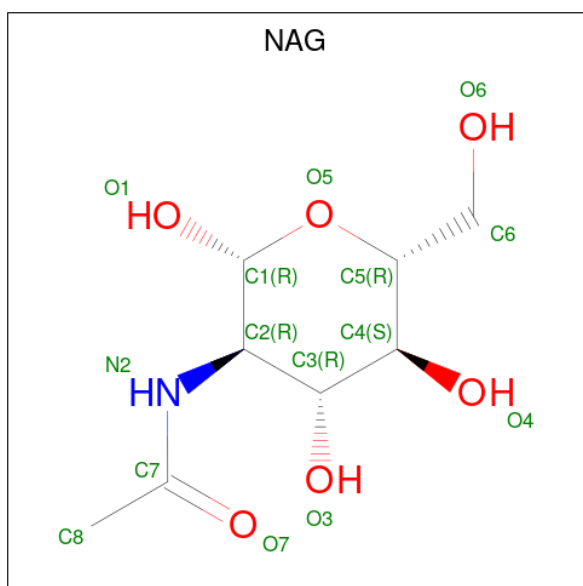
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			55	41	14		
8	C	1	Total	C	O	0	0
			55	41	14		

- Molecule 9 is O-[(S)-({(2R)-2,3-bis[(9Z)-octadec-9-enoyloxy]propyl}oxy)(hydroxy)phosphoryl]-L-serine (three-letter code: 17F) (formula:  $C_{42}H_{78}NO_{10}P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
9	A	1	16	6	1	8	1	0	0
9	A	1	15	6	1	7	1	0	0
9	B	1	24	13	1	9	1	0	0

- Molecule 10 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
10	B	1	Total	C	N	O	0	0
			14	8	1	5		
10	D	1	Total	C	N	O	0	0
			14	8	1	5		
10	D	1	Total	C	N	O	0	0
			14	8	1	5		
10	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 11 is water.

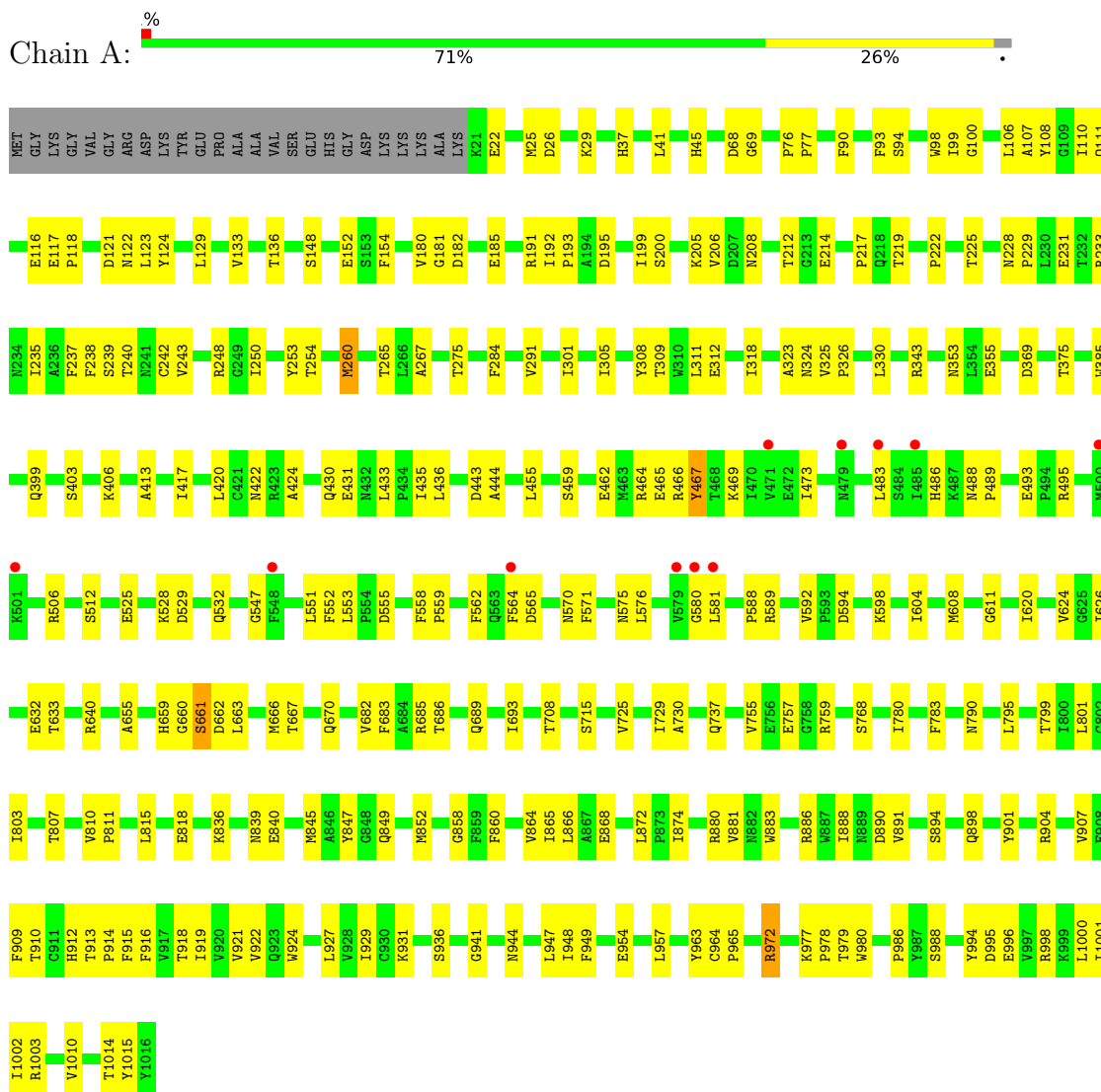
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	3	Total	O	0	0
			3	3		
11	C	3	Total	O	0	0
			3	3		



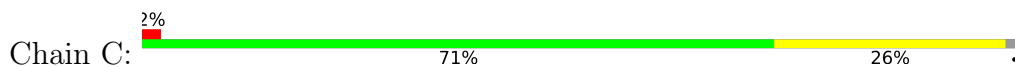
### 3 Residue-property plots

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: Sodium/potassium-transporting ATPase subunit alpha-1



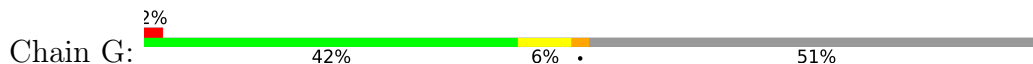
- Molecule 1: Sodium/potassium-transporting ATPase subunit alpha-1



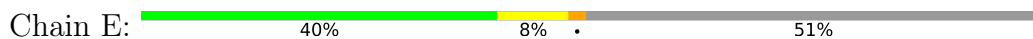




- Molecule 3: Na<sup>+</sup>/K<sup>+</sup> ATPase gamma subunit transcript variant a



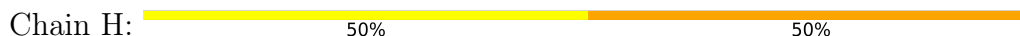
- Molecule 3: Na<sup>+</sup>/K<sup>+</sup> ATPase gamma subunit transcript variant a



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



- Molecule 5: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose



- Molecule 5: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	118.24Å 118.35Å 494.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.41 – 4.00 49.41 – 3.89	Depositor EDS
% Data completeness (in resolution range)	72.2 (49.41-4.00) 68.6 (49.41-3.89)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.45 (at 3.88Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: dev_1593)	Depositor
R, $R_{free}$	0.221 , 0.253 0.229 , 0.258	Depositor DCC
$R_{free}$ test set	2209 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	153.1	Xtrriage
Anisotropy	0.084	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 125.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	0.044 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	21123	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	198.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.94% 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: MG, FRU, PHD, NAG, GLC, CLR, 17F, DGX

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	0.29	2/7867 (0.0%)	0.46	0/10674
1	C	0.29	3/7867 (0.0%)	0.46	0/10674
2	B	0.29	0/2440	0.53	0/3290
2	D	0.29	0/2419	0.52	0/3263
3	E	0.26	0/261	0.45	0/354
3	G	0.27	0/261	0.45	0/354
All	All	0.29	5/21115 (0.0%)	0.48	0/28609

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

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	3
2	D	0	2
All	All	0	5

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	737	GLN	CD-NE2	-6.42	1.16	1.32
1	A	737	GLN	CD-NE2	-6.10	1.17	1.32
1	A	737	GLN	CD-OE1	-5.51	1.11	1.24
1	C	840	GLU	CB-CG	-5.21	1.42	1.52
1	C	737	GLN	CD-OE1	-5.13	1.12	1.24

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	15	PHE	Peptide
2	B	165	GLU	Peptide
2	B	222	GLU	Peptide
2	D	165	GLU	Peptide
2	D	222	GLU	Peptide

## 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	7730	0	7776	188	0
1	C	7730	0	7776	180	0
2	B	2377	0	2349	77	0
2	D	2357	0	2326	74	0
3	E	255	0	259	5	0
3	G	255	0	259	4	0
4	F	28	0	25	0	0
5	H	23	0	21	10	0
5	I	23	0	21	5	0
6	A	3	0	0	0	0
6	C	3	0	0	0	0
7	A	28	0	46	4	0
7	C	56	0	92	7	0
7	G	28	0	46	5	0
8	A	55	0	64	2	0
8	C	55	0	64	2	0
9	A	31	0	18	5	0
9	B	24	0	20	2	0
10	B	14	0	13	0	0
10	D	42	0	39	1	0
11	A	3	0	0	0	0
11	C	3	0	0	0	0
All	All	21123	0	21214	523	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 12.

The worst 5 of 523 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)
2:D:170:LYS:HB2	2:D:174:PRO:HA	1.57	0.86
2:B:170:LYS:HB2	2:B:174:PRO:HA	1.57	0.86
1:C:488:ASN:ND2	1:C:493:GLU:O	2.09	0.86
1:A:488:ASN:ND2	1:A:493:GLU:O	2.09	0.84
1:A:977:LYS:HB3	5:H:1:GLC:H62	1.60	0.82

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	993/1021 (97%)	941 (95%)	52 (5%)	0	100	100
1	C	993/1021 (97%)	942 (95%)	51 (5%)	0	100	100
2	B	288/303 (95%)	268 (93%)	17 (6%)	3 (1%)	15	53
2	D	286/303 (94%)	265 (93%)	18 (6%)	3 (1%)	15	53
3	E	30/65 (46%)	30 (100%)	0	0	100	100
3	G	30/65 (46%)	30 (100%)	0	0	100	100
All	All	2620/2778 (94%)	2476 (94%)	138 (5%)	6 (0%)	47	79

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	17	TRP
2	B	18	ASN
2	B	217	ARG
2	D	18	ASN
2	D	217	ARG

### 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	846/864 (98%)	832 (98%)	14 (2%)	60	78
1	C	846/864 (98%)	832 (98%)	14 (2%)	60	78
2	B	260/269 (97%)	253 (97%)	7 (3%)	44	66
2	D	258/269 (96%)	251 (97%)	7 (3%)	44	66
3	E	26/52 (50%)	25 (96%)	1 (4%)	33	59
3	G	26/52 (50%)	25 (96%)	1 (4%)	33	59
All	All	2262/2370 (95%)	2218 (98%)	44 (2%)	57	75

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	431	GLU
1	C	972	ARG
1	C	459	SER
1	C	564	PHE
2	D	36	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	241	ASN
1	A	486	HIS
1	A	737	GLN
1	C	241	ASN
1	C	486	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.



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

2 non-standard protein/DNA/RNA residues 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
1	PHD	C	369	6,1	9,11,12	1.36	2 (22%)	10,15,17	1.64	2 (20%)
1	PHD	A	369	6,1	9,11,12	1.32	2 (22%)	10,15,17	1.54	2 (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
1	PHD	C	369	6,1	-	0/8/11/13	-
1	PHD	A	369	6,1	-	0/8/11/13	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	369	PHD	CB-CA	-2.21	1.48	1.53
1	C	369	PHD	P-OD1	-2.20	1.55	1.59
1	A	369	PHD	P-OD1	-2.20	1.55	1.59
1	A	369	PHD	CB-CA	-2.04	1.49	1.53

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	369	PHD	OD1-CG-CB	3.18	119.85	111.11
1	A	369	PHD	OD1-CG-CB	3.01	119.40	111.11
1	C	369	PHD	OD2-CG-CB	-2.82	118.50	124.73
1	A	369	PHD	OD2-CG-CB	-2.54	119.12	124.73

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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
4	NAG	F	1	2,4	14,14,15	1.98	1 (7%)	17,19,21	0.74	0
4	NAG	F	2	4	14,14,15	0.25	0	17,19,21	0.36	0
5	GLC	H	1	5	11,11,12	0.57	0	15,15,17	1.32	3 (20%)
5	FRU	H	2	5	11,12,12	0.50	0	10,18,18	1.03	0
5	GLC	I	1	5	11,11,12	0.47	0	15,15,17	0.82	0
5	FRU	I	2	5	11,12,12	0.74	0	10,18,18	0.67	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	NAG	F	1	2,4	-	2/6/23/26	0/1/1/1
4	NAG	F	2	4	-	0/6/23/26	0/1/1/1
5	GLC	H	1	5	-	1/2/19/22	0/1/1/1
5	FRU	H	2	5	-	0/5/24/24	0/1/1/1
5	GLC	I	1	5	-	2/2/19/22	0/1/1/1
5	FRU	I	2	5	-	3/5/24/24	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	1	NAG	O5-C1	7.22	1.55	1.43

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	H	1	GLC	O5-C5-C6	2.91	111.76	107.20
5	H	1	GLC	C6-C5-C4	-2.10	108.08	113.00
5	H	1	GLC	O5-C1-C2	-2.04	107.61	110.77

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

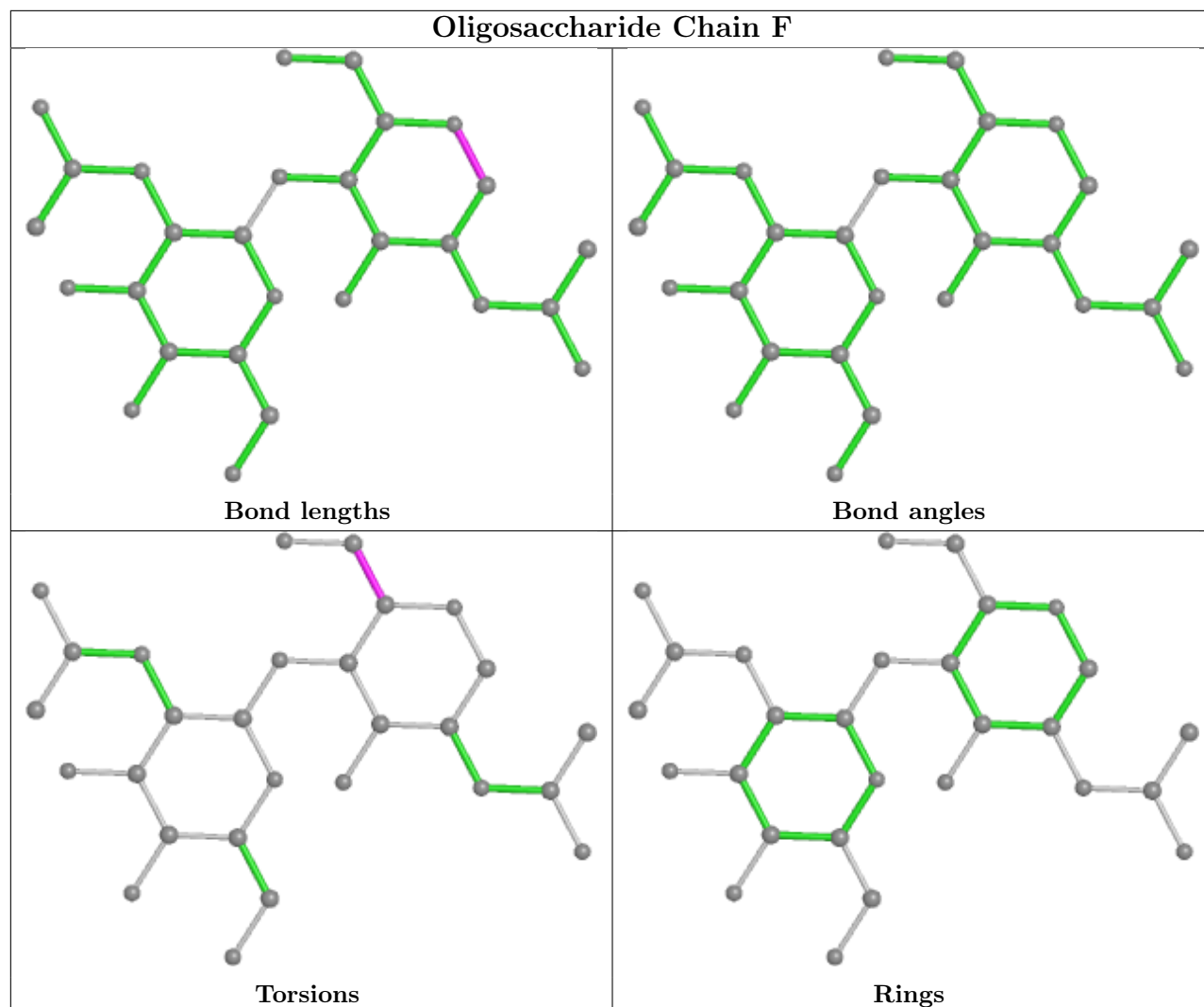
Mol	Chain	Res	Type	Atoms
5	I	2	FRU	O1-C1-C2-C3
5	I	2	FRU	O1-C1-C2-O2
5	I	2	FRU	O1-C1-C2-O5
4	F	1	NAG	C4-C5-C6-O6
4	F	1	NAG	O5-C5-C6-O6

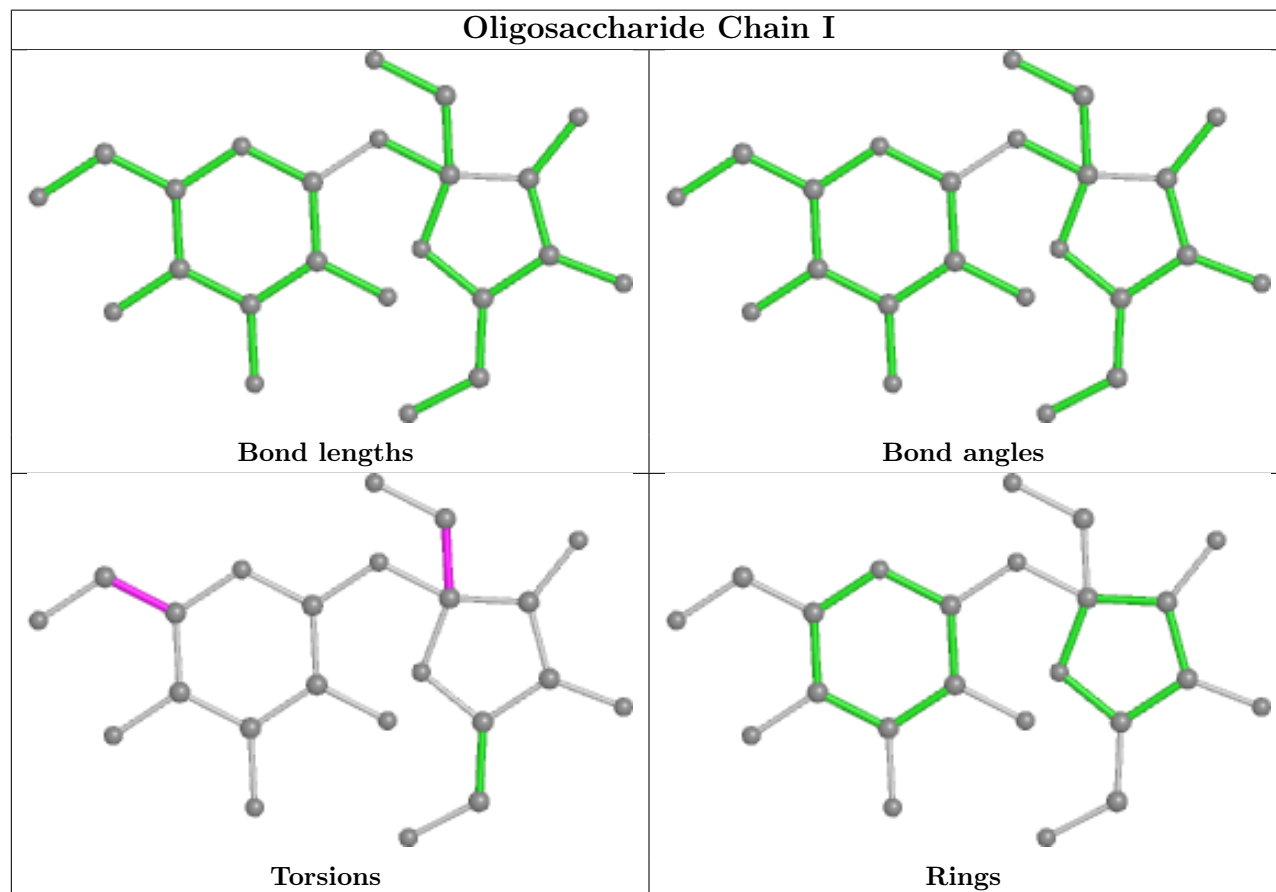
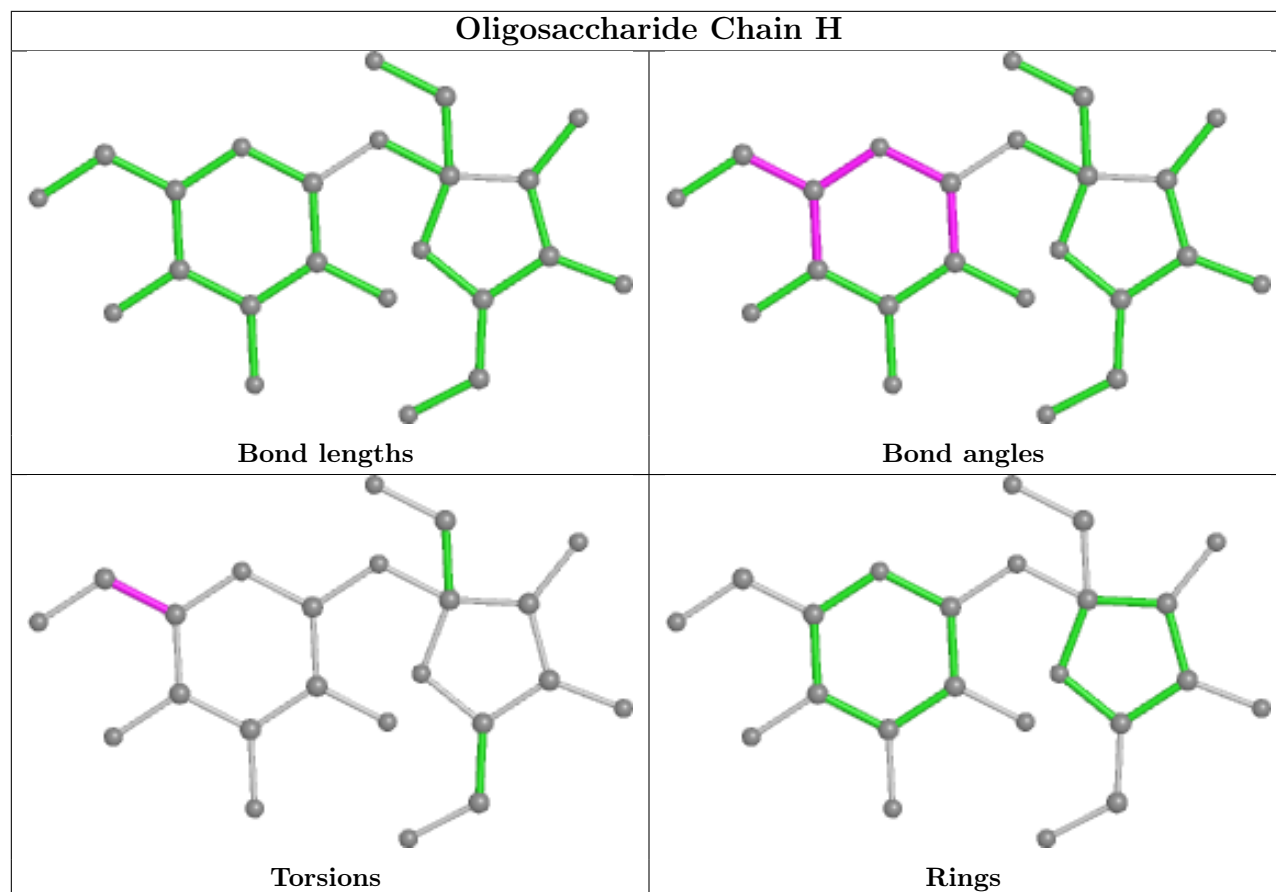
There are no ring outliers.

4 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	I	2	FRU	3	0
5	I	1	GLC	2	0
5	H	2	FRU	3	0
5	H	1	GLC	7	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

Of 19 ligands modelled in this entry, 6 are monoatomic - leaving 13 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
9	17F	A	1109	-	12,14,53	1.78	1 (8%)	14,19,60	1.63	2 (14%)
7	CLR	G	1001	-	31,31,31	2.40	9 (29%)	48,48,48	3.02	17 (35%)
9	17F	A	1108	-	14,15,53	1.52	2 (14%)	16,20,60	1.87	2 (12%)
9	17F	B	1004	-	22,23,53	1.61	2 (9%)	25,29,60	1.33	3 (12%)
10	NAG	D	2002	2	14,14,15	0.28	0	17,19,21	0.42	0
8	DGX	A	1107	-	62,62,62	1.23	5 (8%)	89,98,98	1.75	20 (22%)
10	NAG	B	1003	2	14,14,15	0.37	0	17,19,21	0.43	0
10	NAG	D	2003	2	14,14,15	0.60	0	17,19,21	0.85	2 (11%)
7	CLR	C	2004	-	31,31,31	2.40	9 (29%)	48,48,48	2.96	17 (35%)
10	NAG	D	2001	2	14,14,15	0.39	0	17,19,21	0.35	0
8	DGX	C	2005	-	62,62,62	1.19	4 (6%)	89,98,98	1.73	18 (20%)
7	CLR	A	1104	-	31,31,31	2.39	9 (29%)	48,48,48	3.09	19 (39%)
7	CLR	C	2006	-	31,31,31	2.39	7 (22%)	48,48,48	2.80	16 (33%)

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
9	17F	A	1109	-	-	11/16/16/59	-
7	CLR	G	1001	-	-	3/10/68/68	0/4/4/4
9	17F	A	1108	-	-	11/18/18/59	-
9	17F	B	1004	-	-	11/28/28/59	-
10	NAG	D	2002	2	-	4/6/23/26	0/1/1/1
8	DGX	A	1107	-	-	3/16/141/141	0/8/8/8
10	NAG	B	1003	2	-	4/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
10	NAG	D	2003	2	-	4/6/23/26	0/1/1/1
7	CLR	C	2004	-	-	4/10/68/68	0/4/4/4
10	NAG	D	2001	2	-	2/6/23/26	0/1/1/1
8	DGX	C	2005	-	-	3/16/141/141	0/8/8/8
7	CLR	A	1104	-	-	3/10/68/68	0/4/4/4
7	CLR	C	2006	-	-	4/10/68/68	0/4/4/4

The worst 5 of 48 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	C	2004	CLR	C6-C5	9.11	1.53	1.33
7	G	1001	CLR	C6-C5	9.09	1.53	1.33
7	C	2006	CLR	C6-C5	9.04	1.52	1.33
7	A	1104	CLR	C6-C5	8.97	1.52	1.33
9	B	1004	17F	P1-O6	5.41	1.81	1.59

The worst 5 of 116 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	G	1001	CLR	C19-C10-C9	-8.13	101.99	111.68
7	C	2006	CLR	C4-C5-C6	-8.12	108.91	120.61
7	C	2004	CLR	C4-C5-C6	-7.86	109.28	120.61
7	A	1104	CLR	C4-C5-C6	-7.83	109.32	120.61
7	A	1104	CLR	C7-C6-C5	-7.62	111.01	125.06

There are no chirality outliers.

5 of 67 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	A	1107	DGX	C2Z-C1Z-O4Y-C4Y
8	A	1107	DGX	O5Z-C1Z-O4Y-C4Y
8	C	2005	DGX	C2Z-C1Z-O4Y-C4Y
8	C	2005	DGX	O5Z-C1Z-O4Y-C4Y
9	A	1108	17F	O3-C1-C2-N1

There are no ring outliers.

10 monomers are involved in 28 short contacts:

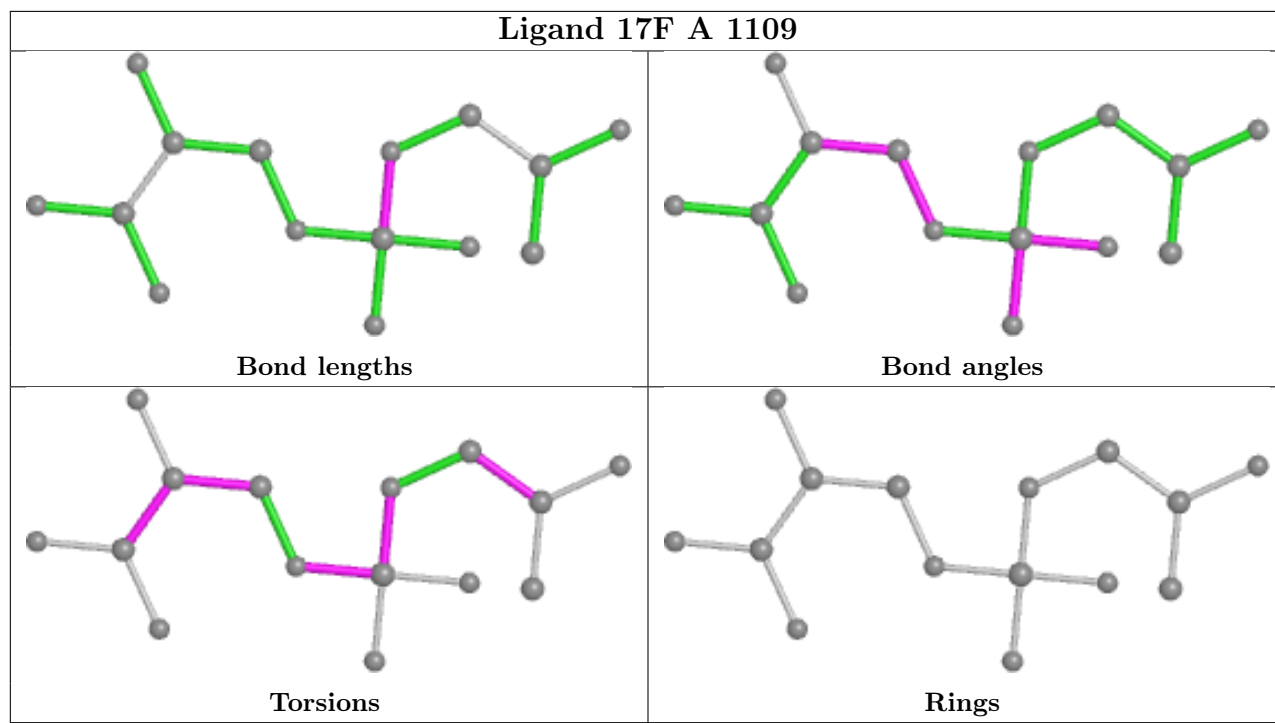
Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	1109	17F	2	0

*Continued on next page...*

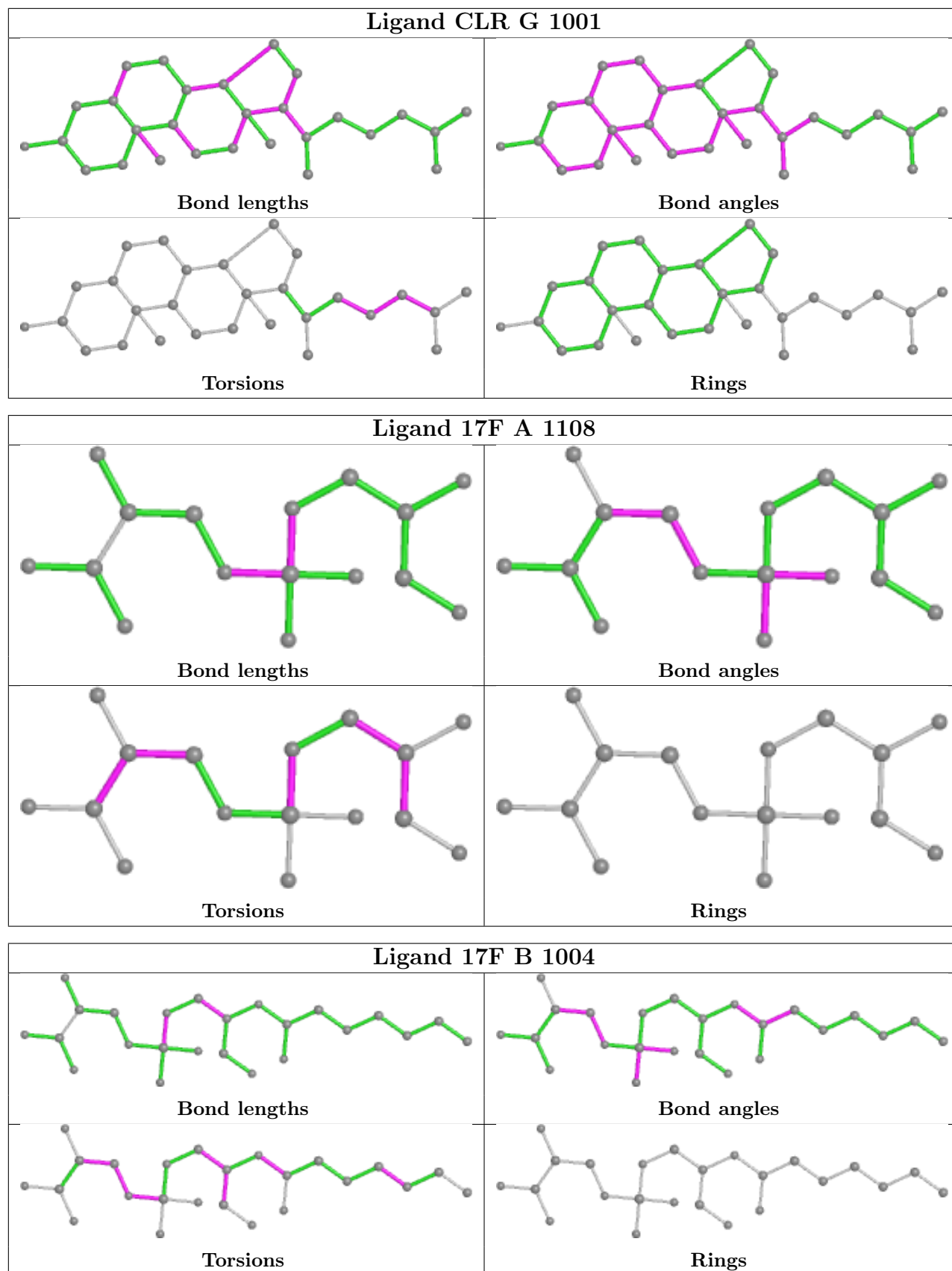
*Continued from previous page...*

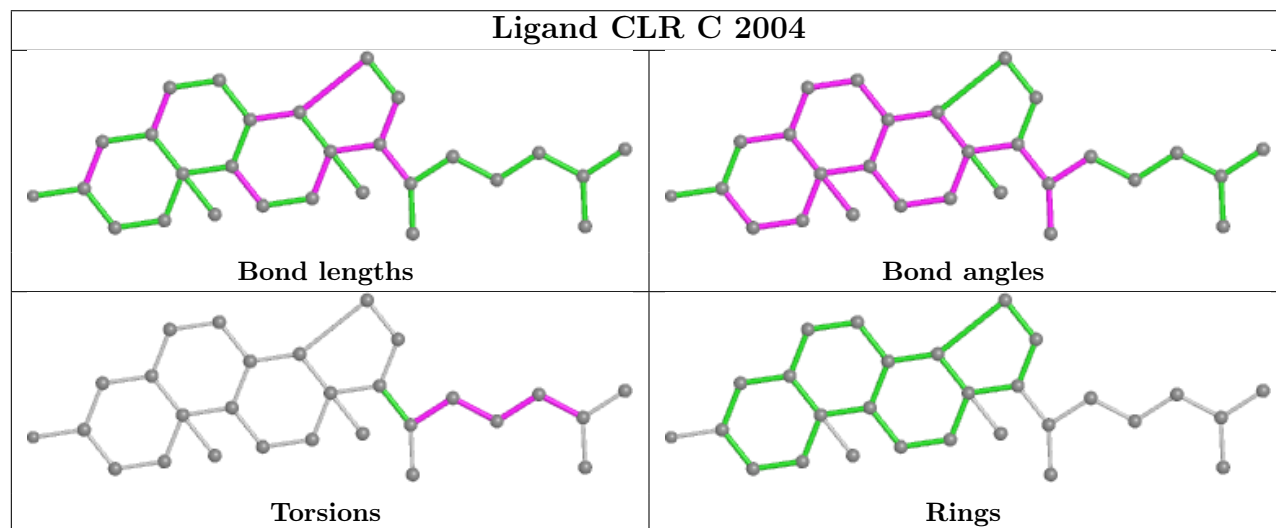
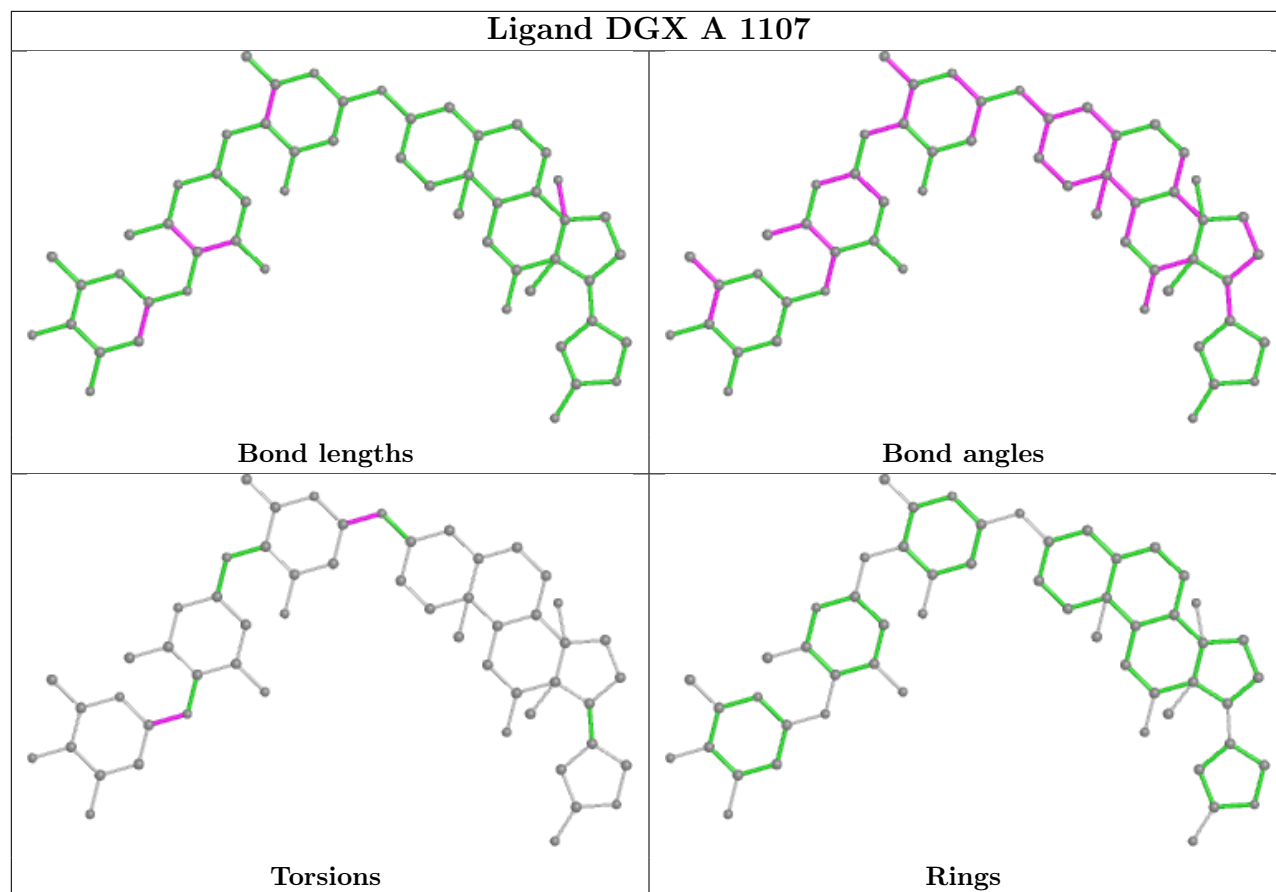
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	G	1001	CLR	5	0
9	A	1108	17F	5	0
9	B	1004	17F	2	0
8	A	1107	DGX	2	0
10	D	2003	NAG	1	0
7	C	2004	CLR	5	0
8	C	2005	DGX	2	0
7	A	1104	CLR	4	0
7	C	2006	CLR	2	0

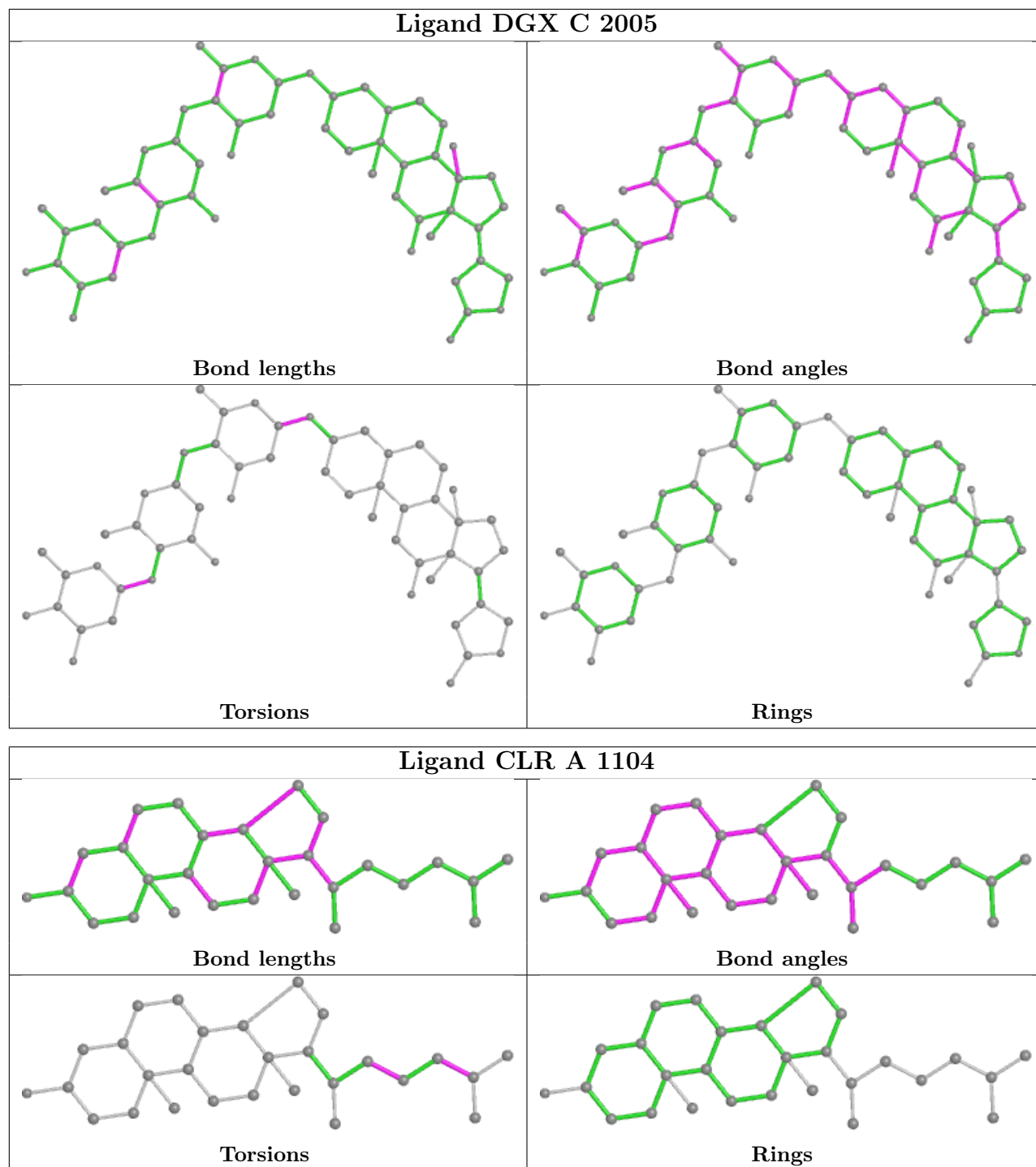
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

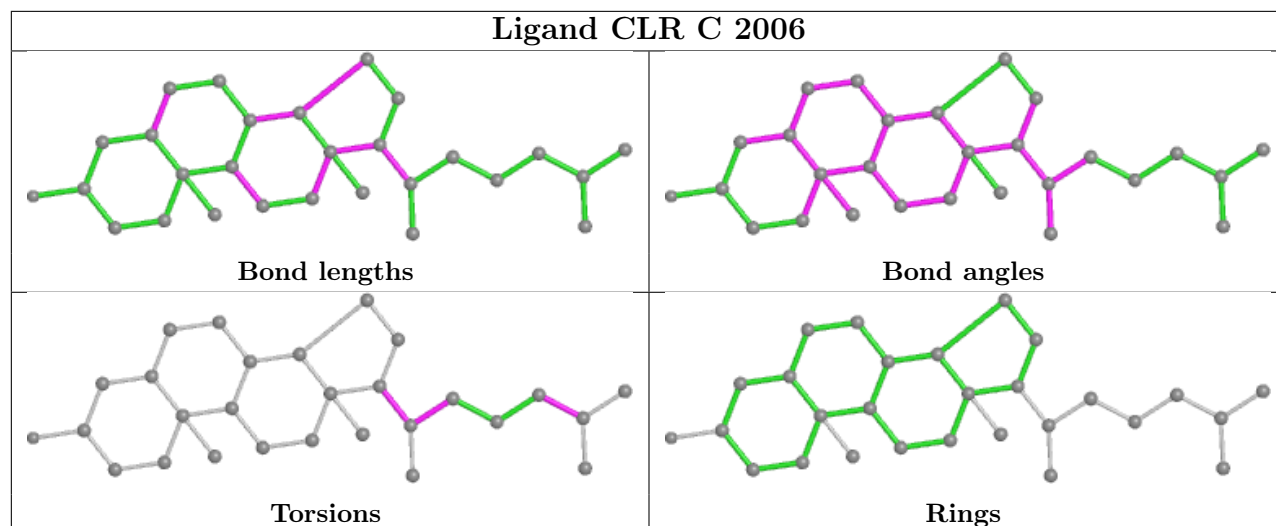












## 5.7 Other polymers [i](#)

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 [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	995/1021 (97%)	-0.33	11 (1%) 80 72	109, 155, 307, 408	0
1	C	995/1021 (97%)	-0.27	21 (2%) 63 54	116, 190, 348, 473	0
2	B	290/303 (95%)	-0.33	3 (1%) 82 74	125, 222, 307, 421	0
2	D	288/303 (95%)	-0.18	5 (1%) 70 60	120, 198, 300, 381	0
3	E	32/65 (49%)	-0.71	0 100 100	97, 148, 217, 297	0
3	G	32/65 (49%)	-0.42	1 (3%) 49 38	128, 149, 228, 268	0
All	All	2632/2778 (94%)	-0.29	41 (1%) 72 62	97, 181, 324, 473	0

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	218	ASP	11.4
2	D	217	ARG	7.0
2	B	218	ASP	4.6
1	A	471	VAL	4.4
2	D	219	GLU	4.3

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

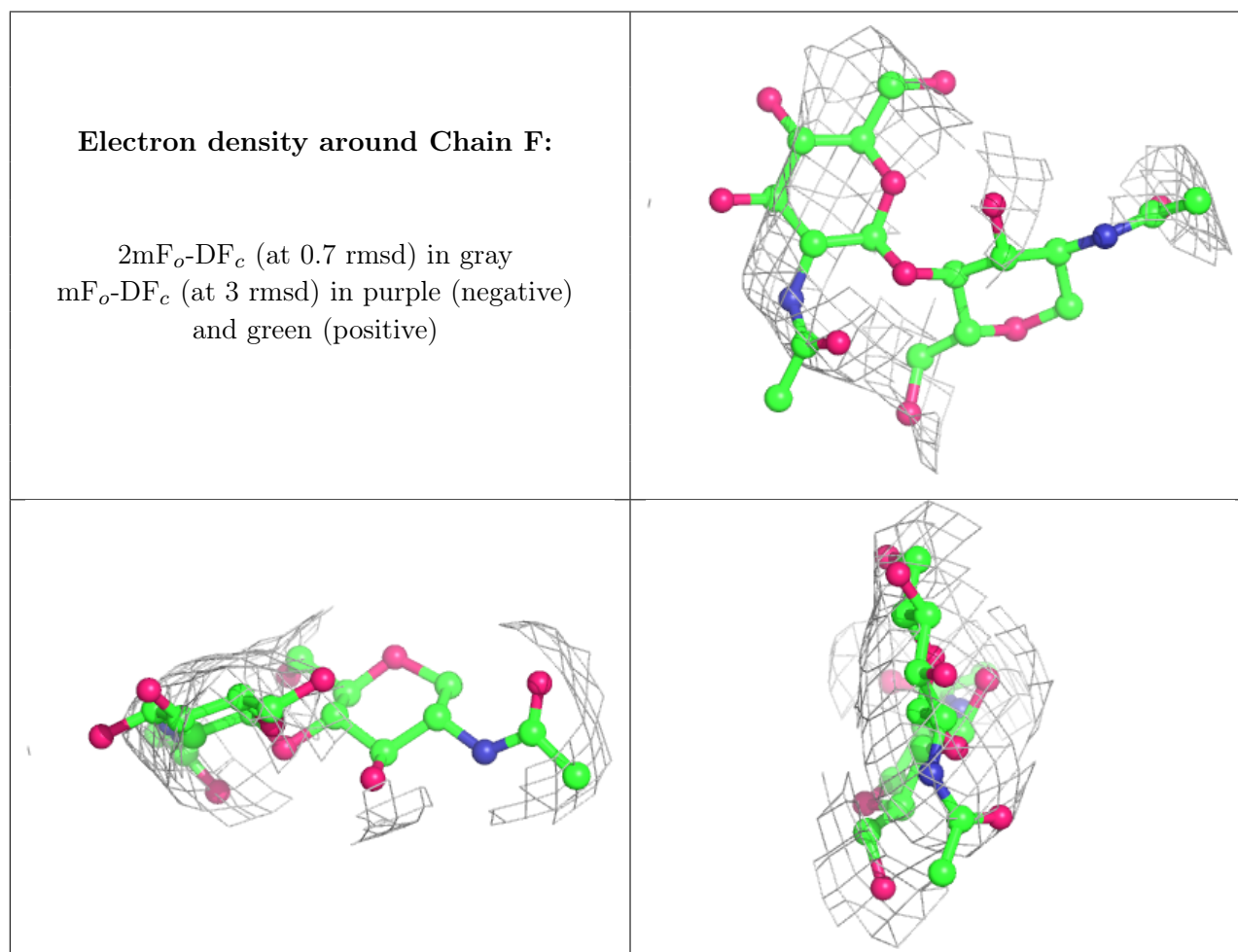
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
1	PHD	C	369	12/13	0.94	0.18	164,202,204,205	0
1	PHD	A	369	12/13	0.97	0.18	120,125,129,130	0

### 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.

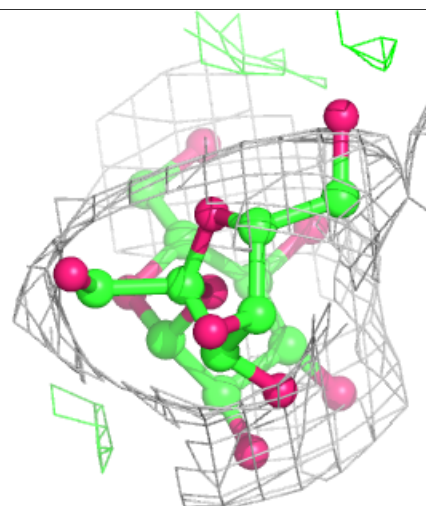
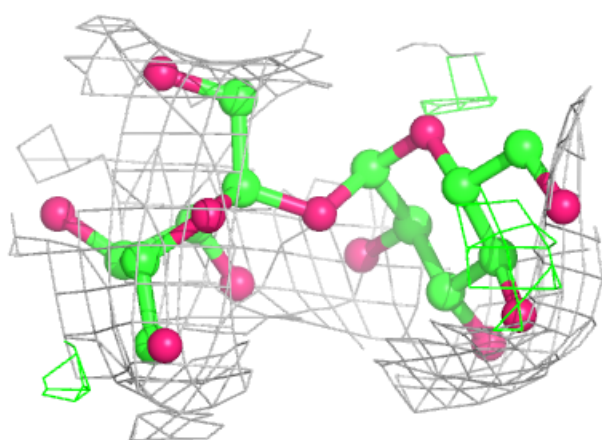
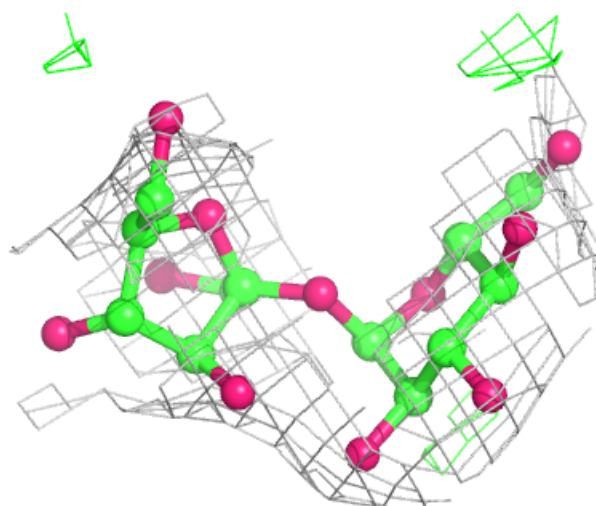
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	GLC	H	1	11/12	0.83	0.31	200,210,213,214	0
5	FRU	H	2	12/12	0.84	0.25	187,204,211,215	0
4	NAG	F	2	14/15	0.85	0.44	245,257,261,266	0
4	NAG	F	1	14/15	0.87	0.18	226,235,245,249	0
5	FRU	I	2	12/12	0.87	0.21	171,183,188,189	0
5	GLC	I	1	11/12	0.91	0.18	162,163,171,176	0

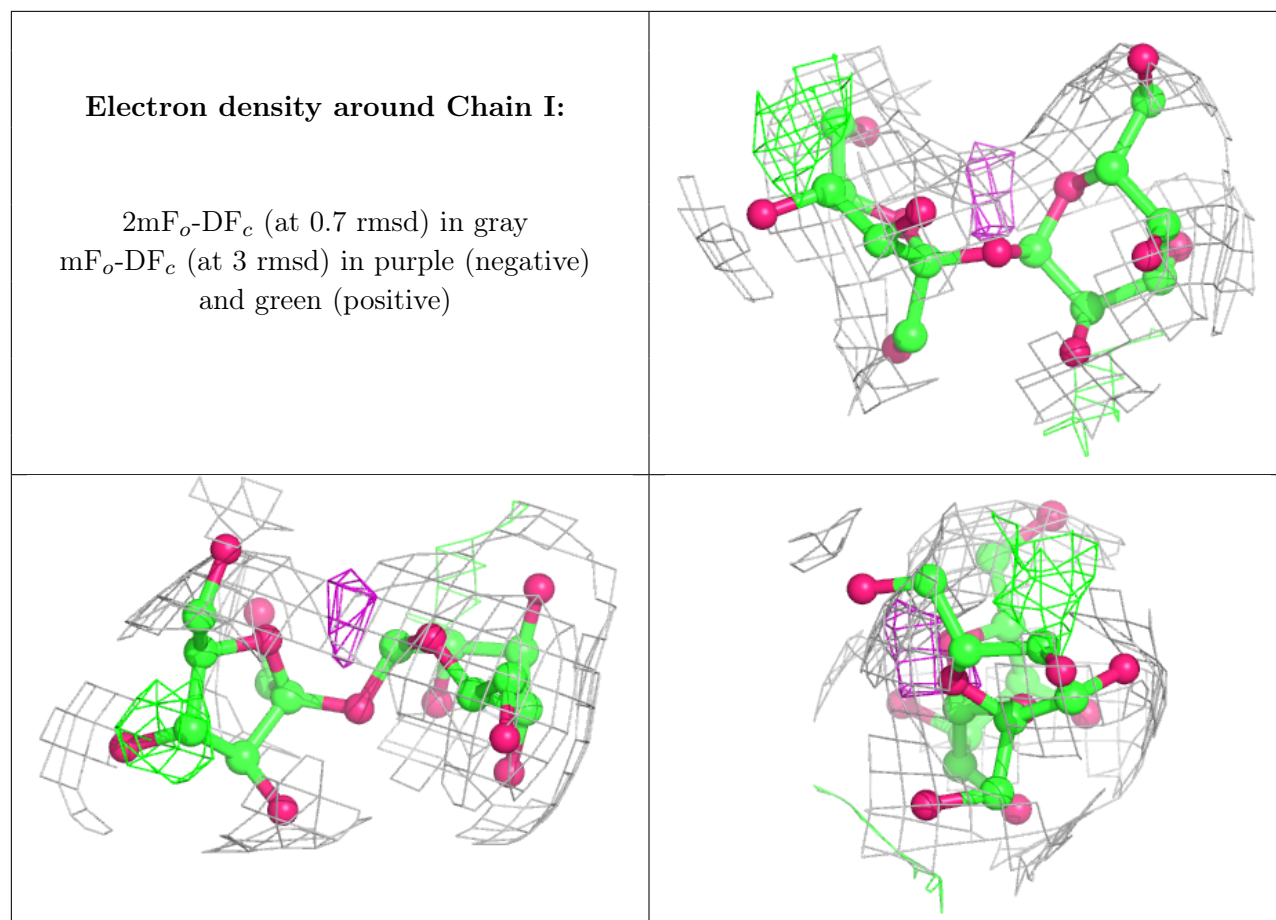
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



**Electron density around Chain H:**

$2mF_o-DF_c$  (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	17F	A	1108	16/54	0.65	0.36	196,222,271,271	0
9	17F	B	1004	24/54	0.67	0.40	128,215,240,263	0
9	17F	A	1109	15/54	0.71	0.36	143,217,306,306	0
10	NAG	D	2002	14/15	0.80	0.15	179,191,200,204	0
7	CLR	C	2004	28/28	0.82	0.45	146,162,169,218	0
8	DGX	A	1107	55/55	0.82	0.35	128,187,256,259	0
10	NAG	D	2001	14/15	0.83	0.29	201,213,221,222	0
7	CLR	A	1104	28/28	0.84	0.31	128,137,188,190	0
10	NAG	D	2003	14/15	0.84	0.30	228,238,245,250	0
10	NAG	B	1003	14/15	0.86	0.45	209,222,229,230	0
8	DGX	C	2005	55/55	0.87	0.26	153,182,232,234	0
7	CLR	C	2006	28/28	0.90	0.25	143,143,143,143	0

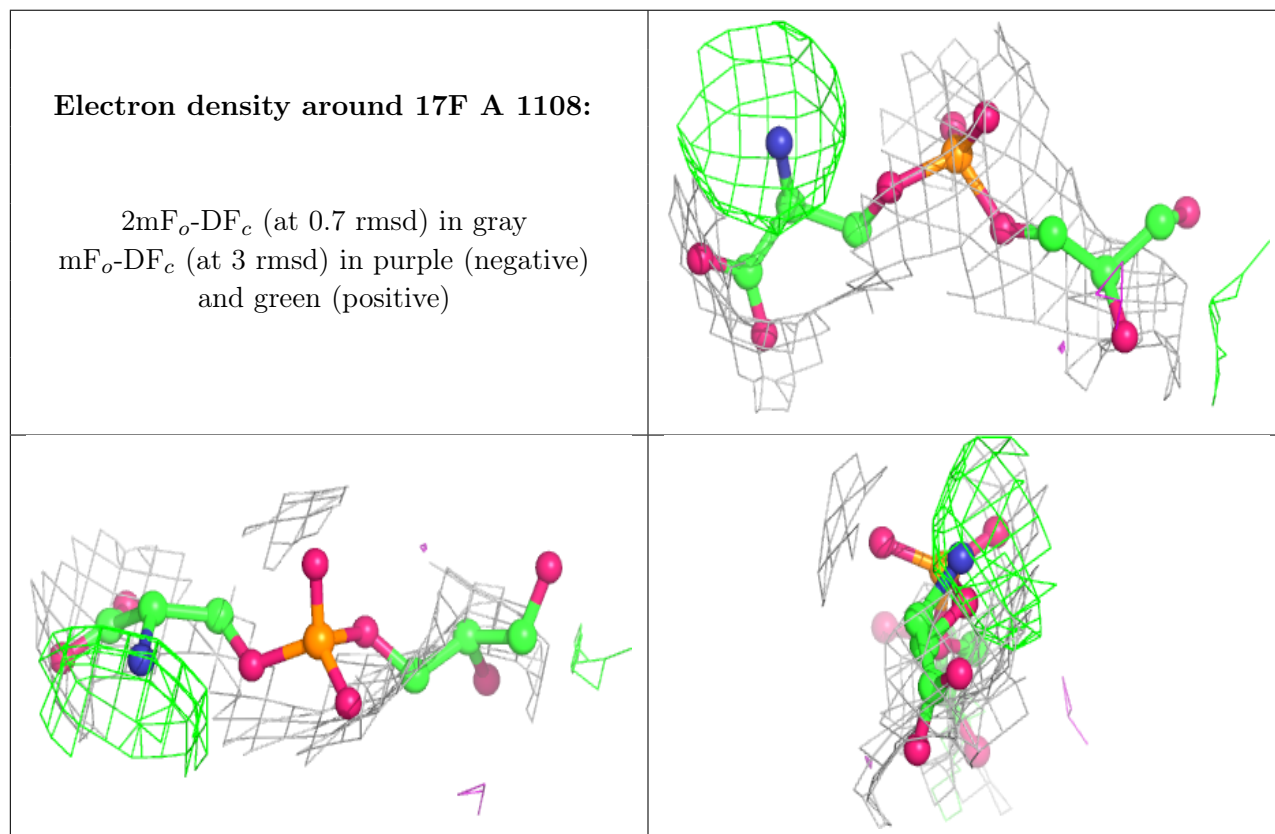
*Continued on next page...*



Continued from previous page...

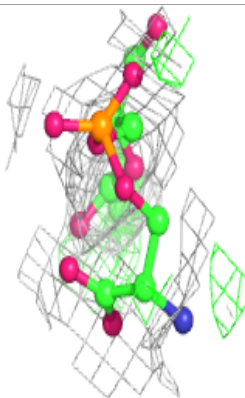
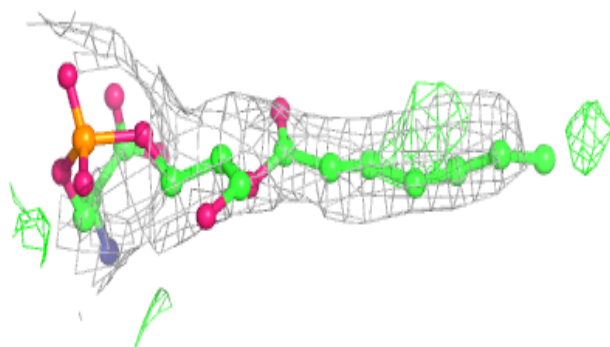
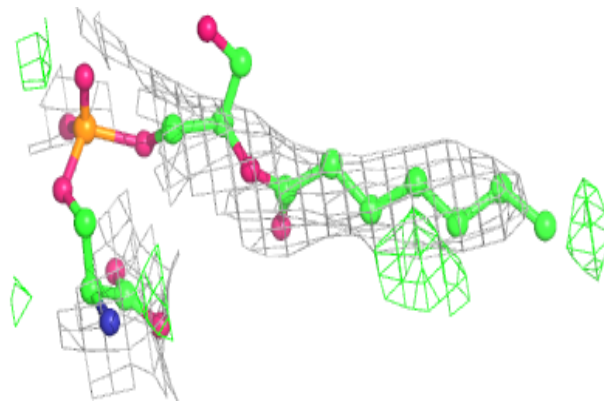
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	MG	A	1102	1/1	0.90	0.63	115,115,115,115	0
6	MG	C	2001	1/1	0.91	0.17	147,147,147,147	0
7	CLR	G	1001	28/28	0.93	0.25	143,143,143,143	0
6	MG	C	2002	1/1	0.94	0.63	148,148,148,148	0
6	MG	C	2003	1/1	0.95	0.31	130,130,130,130	0
6	MG	A	1103	1/1	0.95	0.28	112,112,112,112	0
6	MG	A	1101	1/1	0.99	0.14	123,123,123,123	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.

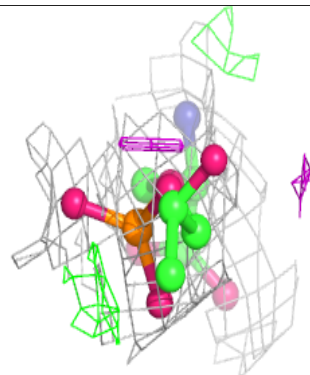
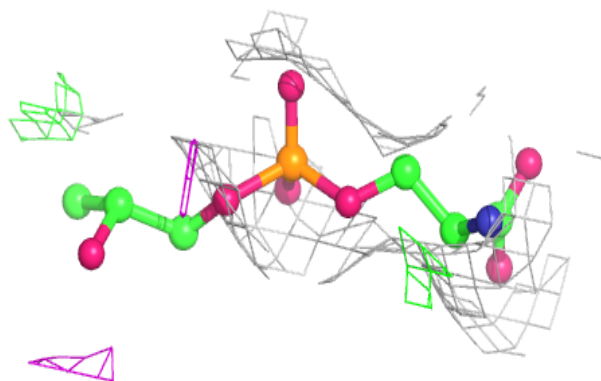
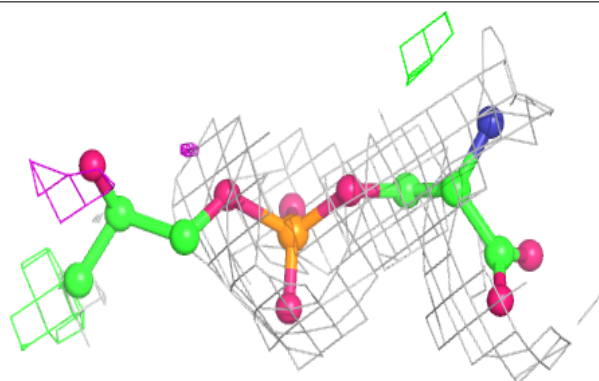


**Electron density around 17F B 1004:**

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

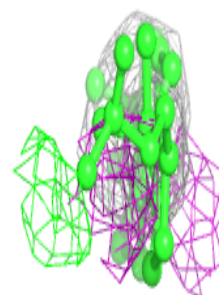
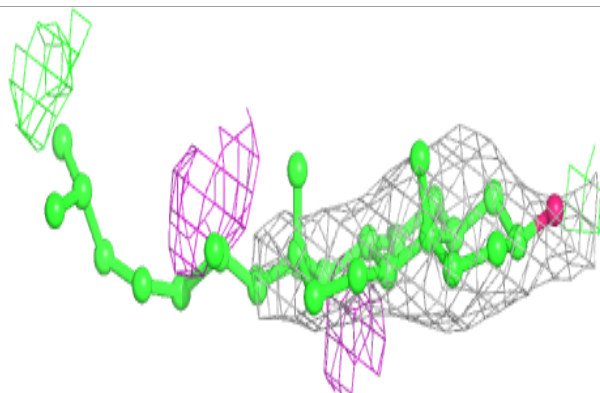
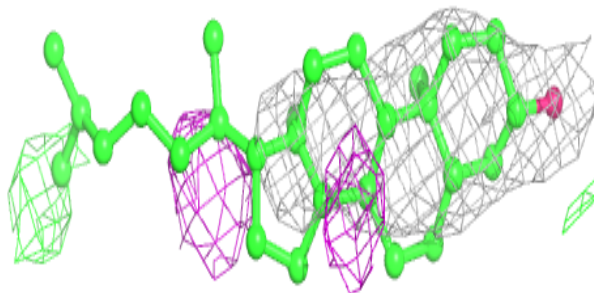
**Electron density around 17F A 1109:**

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

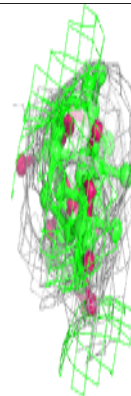
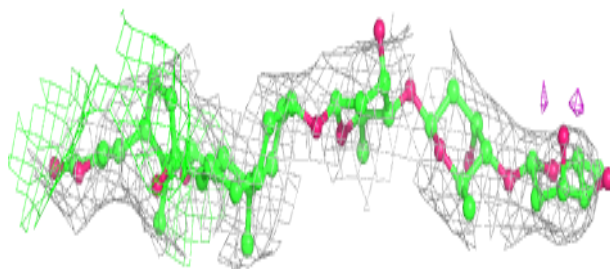
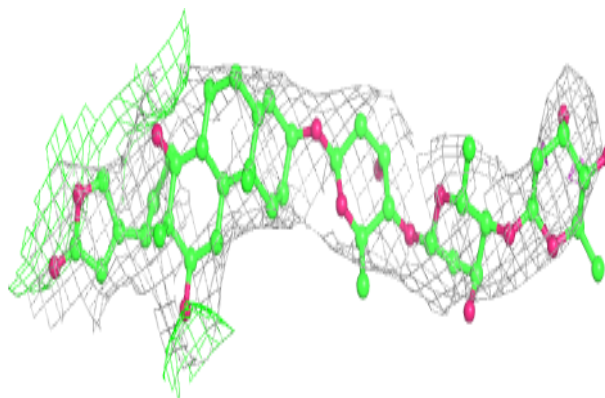


**Electron density around CLR C 2004:**

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

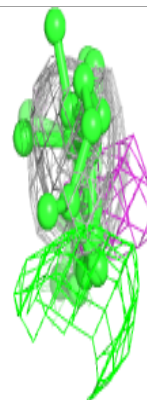
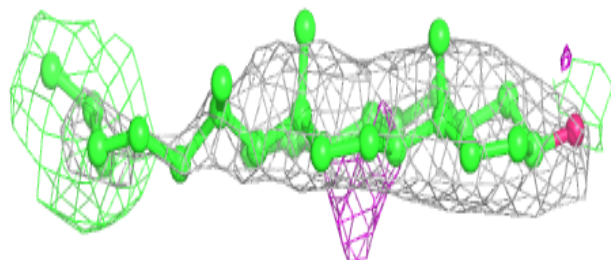
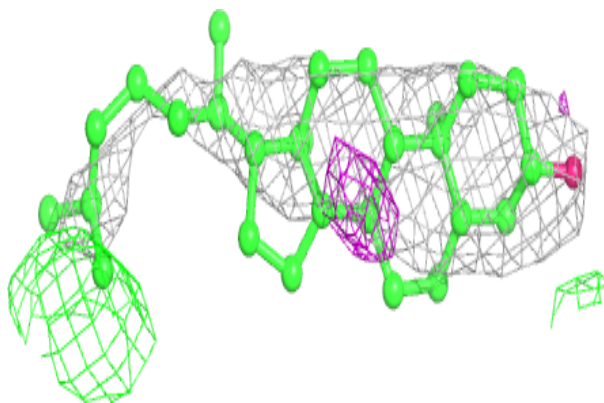
**Electron density around DGX A 1107:**

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

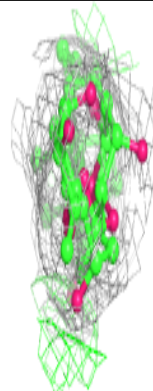
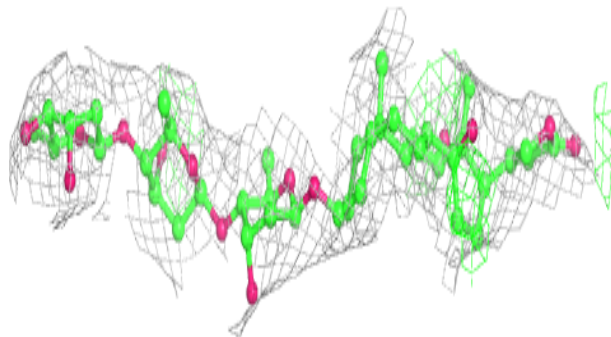
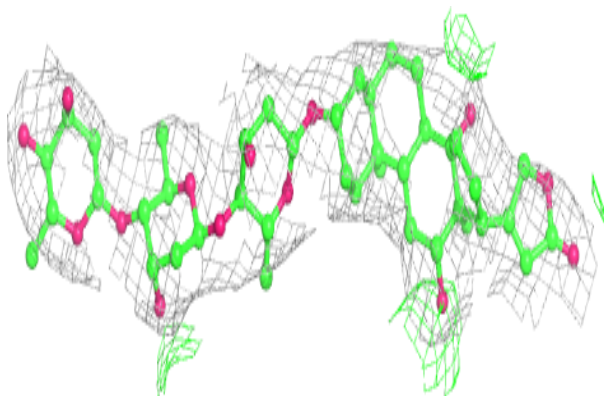


**Electron density around CLR A 1104:**

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

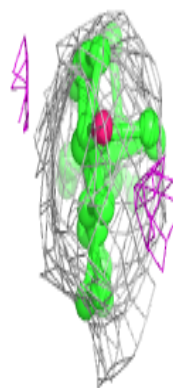
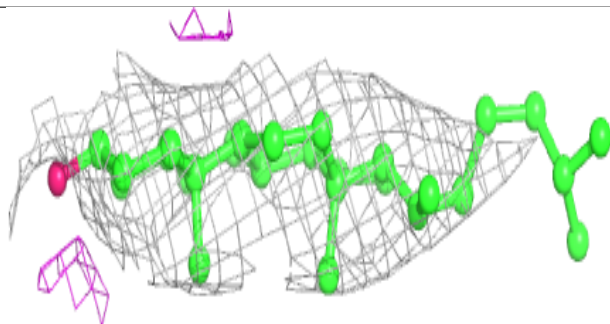
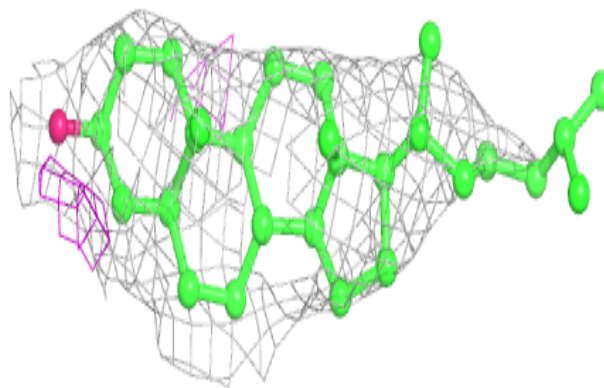
**Electron density around DGX C 2005:**

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

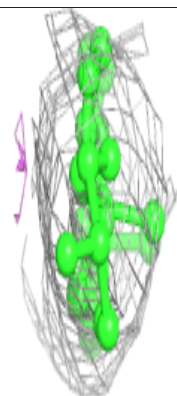
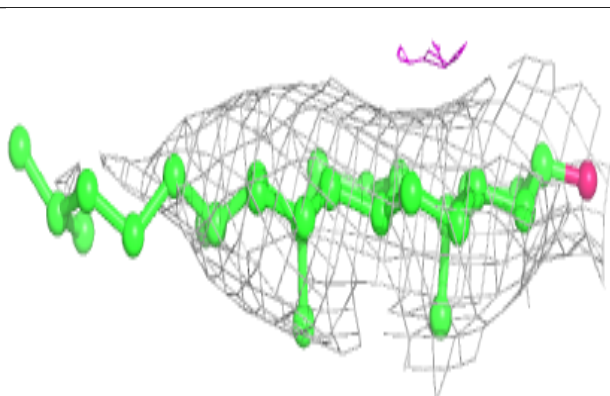
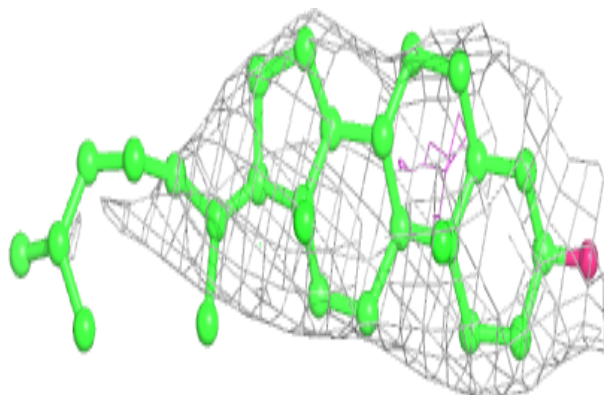


**Electron density around CLR C 2006:**

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

**Electron density around CLR G 1001:**

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



## 6.5 Other polymers [i](#)

There are no such residues in this entry.