



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 21, 2022 – 05:07 pm GMT

PDB ID : 7NI1  
Title : CRYSTAL STRUCTURE OF NATIVE HUMAN MYELOPEROXIDASE IN  
COMPLEX WITH CPD 9  
Authors : Sjogren, T.; Inghardt, T.  
Deposited on : 2021-02-11  
Resolution : 2.11 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.26  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.26

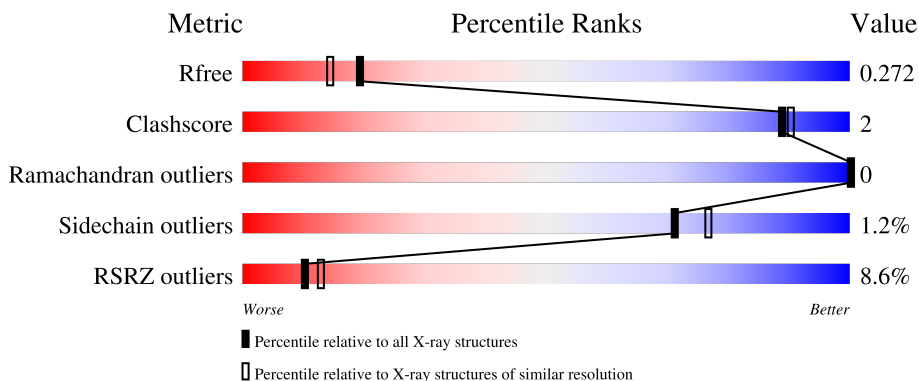
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.11 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	6241 (2.14-2.10)
Clashscore	141614	6778 (2.14-2.10)
Ramachandran outliers	138981	6705 (2.14-2.10)
Sidechain outliers	138945	6706 (2.14-2.10)
RSRZ outliers	127900	6112 (2.14-2.10)

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	105	 8% 93% 6%
1	B	105	 7% 90% 8%
2	C	466	 8% 94% 6%
2	D	466	 10% 94% 6%
3	E	5	 40% 60%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	5	

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
9	BMA	C	607	-	-	-	X
9	BMA	D	601	-	-	-	X

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 10097 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.

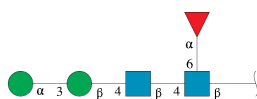
- Molecule 1 is a protein called Myeloperoxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	104	Total 837	C 529	N 148	O 155	S 5	0	0	0
1	B	103	Total 832	C 526	N 147	O 154	S 5	0	0	0

- Molecule 2 is a protein called Myeloperoxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	464	Total 3720	C 2343	N 685	O 665	S 27	0	0	0
2	D	464	Total 3720	C 2343	N 685	O 665	S 27	0	0	0

- Molecule 3 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	E	5	Total 60	C 34	N 2	O 24	0	0	0
3	F	5	Total 60	C 34	N 2	O 24	0	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Cl 2	0	0

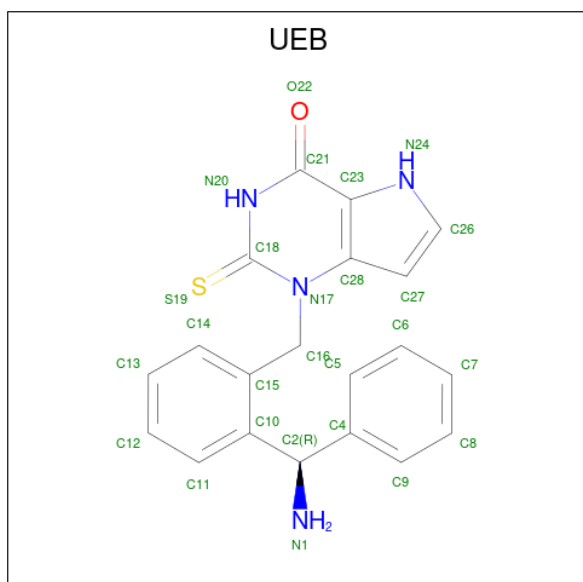
*Continued on next page...*





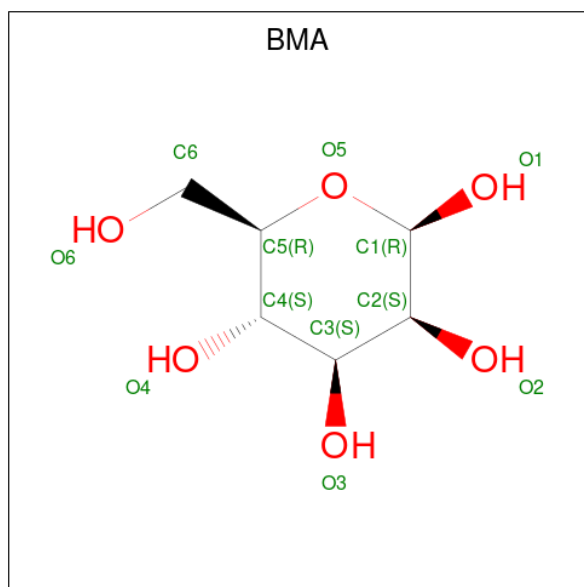
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
7	C	1	Total 14	C 8	N 1	O 5	0	0
7	C	1	Total 14	C 8	N 1	O 5	0	0
7	D	1	Total 14	C 8	N 1	O 5	0	0
7	D	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 8 is (S)-1-(2-(amino(phenyl)methyl)benzyl)-2-thioxo-1,2,3,5-tetrahydro-4H-pyrrolo[3,2-d]pyrimidin-4-one (three-letter code: UEB) (formula: C<sub>20</sub>H<sub>18</sub>N<sub>4</sub>OS) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
8	C	1	Total	C	N	O	S	0	0
			26	20	4	1	1		
8	D	1	Total	C	N	O	S	0	0
			26	20	4	1	1		

- Molecule 9 is beta-D-mannopyranose (three-letter code: BMA) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	C	1	Total	C	O	0	0
			11	6	5		
9	D	1	Total	C	O	0	0
			11	6	5		

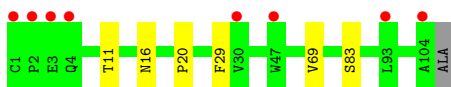
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	71	Total	O	0	0
			71	71		
10	B	74	Total	O	0	0
			74	74		
10	C	280	Total	O	0	0
			280	280		
10	D	218	Total	O	0	0
			218	218		

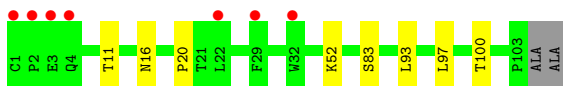
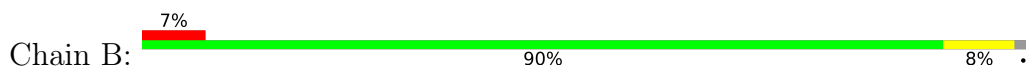
### 3 Residue-property plots

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.

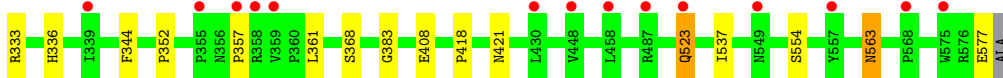
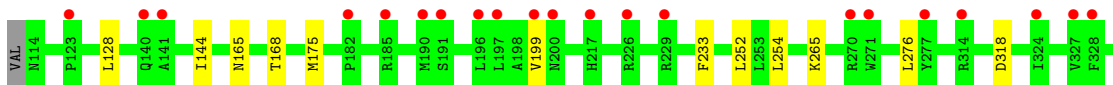
- Molecule 1: Myeloperoxidase



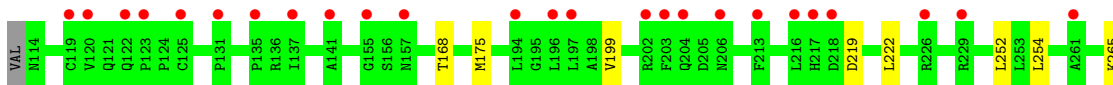
- Molecule 1: Myeloperoxidase



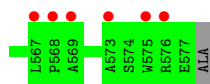
- Molecule 2: Myeloperoxidase



- Molecule 2: Myeloperoxidase







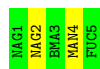
- Molecule 3: alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E: 40% 60%



- Molecule 3: alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F: 60% 40%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.22Å 63.82Å 111.34Å 90.00° 97.24° 90.00°	Depositor
Resolution (Å)	46.24 – 2.11 46.24 – 2.11	Depositor EDS
% Data completeness (in resolution range)	97.9 (46.24-2.11) 97.9 (46.24-2.11)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.32 (at 2.10Å)	Xtriage
Refinement program	BUSTER 2.11.7	Depositor
R, $R_{free}$	0.217 , 0.256 0.228 , 0.272	Depositor DCC
$R_{free}$ test set	3581 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.1	Xtriage
Anisotropy	0.707	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	10097	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.20% 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: NAG, MAN, UEB, CSO, CL, CA, HEM, FUC, BMA

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	0.51	0/862	0.68	0/1174
1	B	0.50	0/857	0.68	0/1167
2	C	0.52	0/3798	0.64	0/5151
2	D	0.51	0/3798	0.64	0/5151
All	All	0.51	0/9315	0.65	0/12643

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	837	0	798	5	0
1	B	832	0	793	5	0
2	C	3720	0	3711	16	0
2	D	3720	0	3711	16	0
3	E	60	0	52	0	0
3	F	60	0	52	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	C	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	1	0	0	0	0
5	A	43	0	30	2	0
5	B	43	0	30	3	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
7	C	28	0	26	0	0
7	D	28	0	26	0	0
8	C	26	0	0	0	0
8	D	26	0	0	0	0
9	C	11	0	10	0	0
9	D	11	0	10	0	0
10	A	71	0	0	0	0
10	B	74	0	0	0	0
10	C	280	0	0	1	0
10	D	218	0	0	0	0
All	All	10097	0	9249	35	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 2.

All (35) 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:C:336:HIS:HD1	2:C:421:ASN:HD21	1.40	0.67
2:D:336:HIS:HD1	2:D:421:ASN:HD21	1.40	0.67
2:D:393:ARG:HB2	2:D:396:GLN:HB2	1.83	0.60
2:D:219:ASP:HB3	2:D:222:LEU:HD12	1.90	0.53
2:C:563:ASN:HB2	10:C:898:HOH:O	2.09	0.52
1:A:83:SER:HB3	2:C:554:SER:O	2.09	0.51
2:D:265:LYS:HD3	2:D:276:LEU:HD11	1.93	0.50
1:A:29:PHE:CE1	2:C:165:ASN:HB2	2.47	0.50
2:C:265:LYS:HD3	2:C:276:LEU:HD11	1.93	0.50
1:B:100:THR:HG21	2:D:428:HIS:CE1	2.49	0.47
1:A:69:VAL:HG11	2:C:418:PRO:HG2	1.96	0.47
2:C:344:PHE:O	2:C:383:GLY:HA3	2.15	0.46
2:C:128:LEU:HB2	2:C:144:ILE:HB	1.98	0.46
2:D:344:PHE:O	2:D:383:GLY:HA3	2.16	0.46
5:B:203:HEM:HAD1	2:D:333:ARG:NH2	2.30	0.45
5:B:203:HEM:HAD1	2:D:333:ARG:HH21	1.81	0.45
2:D:491:VAL:HB	2:D:495:LEU:HB2	2.01	0.43
1:B:93:LEU:O	1:B:97:LEU:HG	2.19	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:11:THR:O	2:D:168:THR:HG22	2.19	0.42
1:A:11:THR:O	2:C:168:THR:HG22	2.19	0.42
1:B:16:ASN:O	1:B:20:PRO:HA	2.20	0.42
1:B:83:SER:HB3	2:D:554:SER:O	2.20	0.41
5:B:203:HEM:CBC	2:D:335:GLY:HA3	2.50	0.41
2:C:233:PHE:CE2	2:C:368:SER:HB2	2.55	0.41
2:C:252:LEU:HD11	2:C:537:ILE:HA	2.01	0.41
2:D:199:VAL:HG12	2:D:254:LEU:HD21	2.02	0.41
1:A:16:ASN:O	1:A:20:PRO:HA	2.21	0.41
5:A:203:HEM:HAD1	2:C:333:ARG:NH2	2.36	0.41
2:C:523:GLN:NE2	2:C:523:GLN:H	2.19	0.41
2:D:523:GLN:H	2:D:523:GLN:NE2	2.20	0.41
2:C:199:VAL:HG12	2:C:254:LEU:HD21	2.03	0.40
2:D:252:LEU:HD11	2:D:537:ILE:HA	2.03	0.40
5:A:203:HEM:HAD1	2:C:333:ARG:HH21	1.87	0.40
2:C:352:PRO:HB3	2:C:357:PRO:HB3	2.02	0.40
2:D:417:LEU:HB3	2:D:418:PRO:HD3	2.04	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	102/105 (97%)	99 (97%)	3 (3%)	0	100	100
1	B	101/105 (96%)	98 (97%)	3 (3%)	0	100	100
2	C	461/466 (99%)	453 (98%)	8 (2%)	0	100	100
2	D	461/466 (99%)	452 (98%)	9 (2%)	0	100	100
All	All	1125/1142 (98%)	1102 (98%)	23 (2%)	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	90/90 (100%)	90 (100%)	0	100	100
1	B	90/90 (100%)	89 (99%)	1 (1%)	73	79
2	C	409/410 (100%)	402 (98%)	7 (2%)	60	66
2	D	409/410 (100%)	405 (99%)	4 (1%)	76	81
All	All	998/1000 (100%)	986 (99%)	12 (1%)	71	77

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	52	LYS
2	C	175	MET
2	C	318	ASP
2	C	361	LEU
2	C	408	GLU
2	C	523	GLN
2	C	563	ASN
2	C	577	GLU
2	D	175	MET
2	D	318	ASP
2	D	523	GLN
2	D	563	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	523	GLN
2	D	523	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](#)

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
2	CSO	C	150	2	3,6,7	1.00	0	0,6,8	-	-
2	CSO	D	150	2	3,6,7	0.81	0	0,6,8	-	-

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	CSO	C	150	2	-	0/1/5/7	-
2	CSO	D	150	2	-	0/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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$
3	NAG	E	1	3,2	14,14,15	0.26	0	17,19,21	0.81	1 (5%)
3	NAG	E	2	3	14,14,15	0.25	0	17,19,21	0.96	1 (5%)
3	BMA	E	3	3	11,11,12	0.36	0	15,15,17	0.86	0
3	MAN	E	4	3	11,11,12	0.39	0	15,15,17	1.16	1 (6%)
3	FUC	E	5	3	10,10,11	0.47	0	14,14,16	0.61	0
3	NAG	F	1	3,2	14,14,15	0.29	0	17,19,21	0.75	0
3	NAG	F	2	3	14,14,15	0.28	0	17,19,21	0.80	1 (5%)
3	BMA	F	3	3	11,11,12	0.38	0	15,15,17	0.81	0
3	MAN	F	4	3	11,11,12	0.41	0	15,15,17	1.21	1 (6%)
3	FUC	F	5	3	10,10,11	0.50	0	14,14,16	0.55	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
3	NAG	E	1	3,2	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	0/6/23/26	0/1/1/1
3	BMA	E	3	3	-	2/2/19/22	0/1/1/1
3	MAN	E	4	3	-	0/2/19/22	0/1/1/1
3	FUC	E	5	3	-	-	0/1/1/1
3	NAG	F	1	3,2	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	0/6/23/26	0/1/1/1
3	BMA	F	3	3	-	2/2/19/22	0/1/1/1
3	MAN	F	4	3	-	0/2/19/22	0/1/1/1
3	FUC	F	5	3	-	-	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	4	MAN	C1-O5-C5	4.32	118.04	112.19
3	E	4	MAN	C1-O5-C5	4.20	117.88	112.19
3	E	2	NAG	C1-O5-C5	2.92	116.15	112.19
3	E	1	NAG	C1-O5-C5	2.41	115.46	112.19
3	F	2	NAG	C1-O5-C5	2.14	115.09	112.19



There are no chirality outliers.

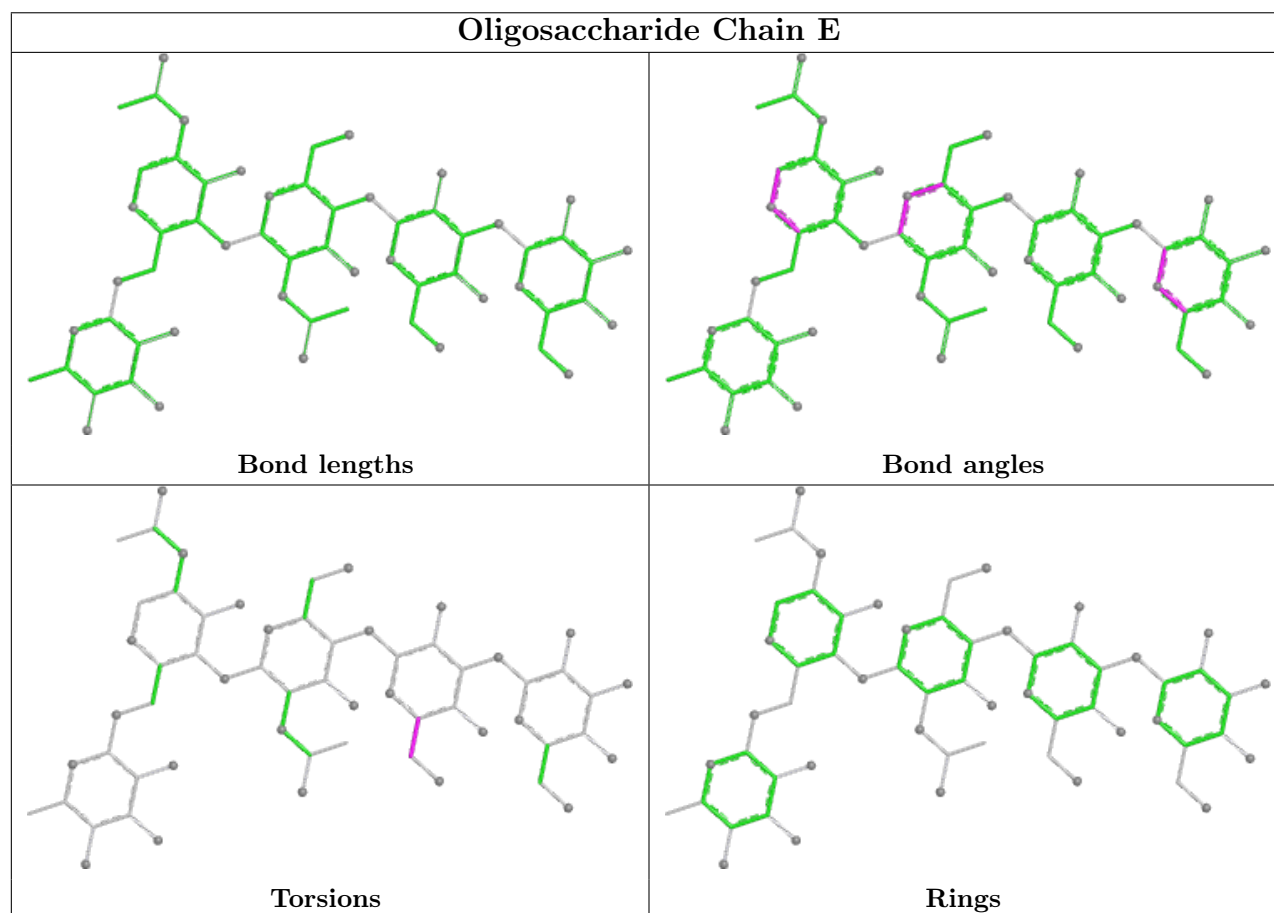
All (4) torsion outliers are listed below:

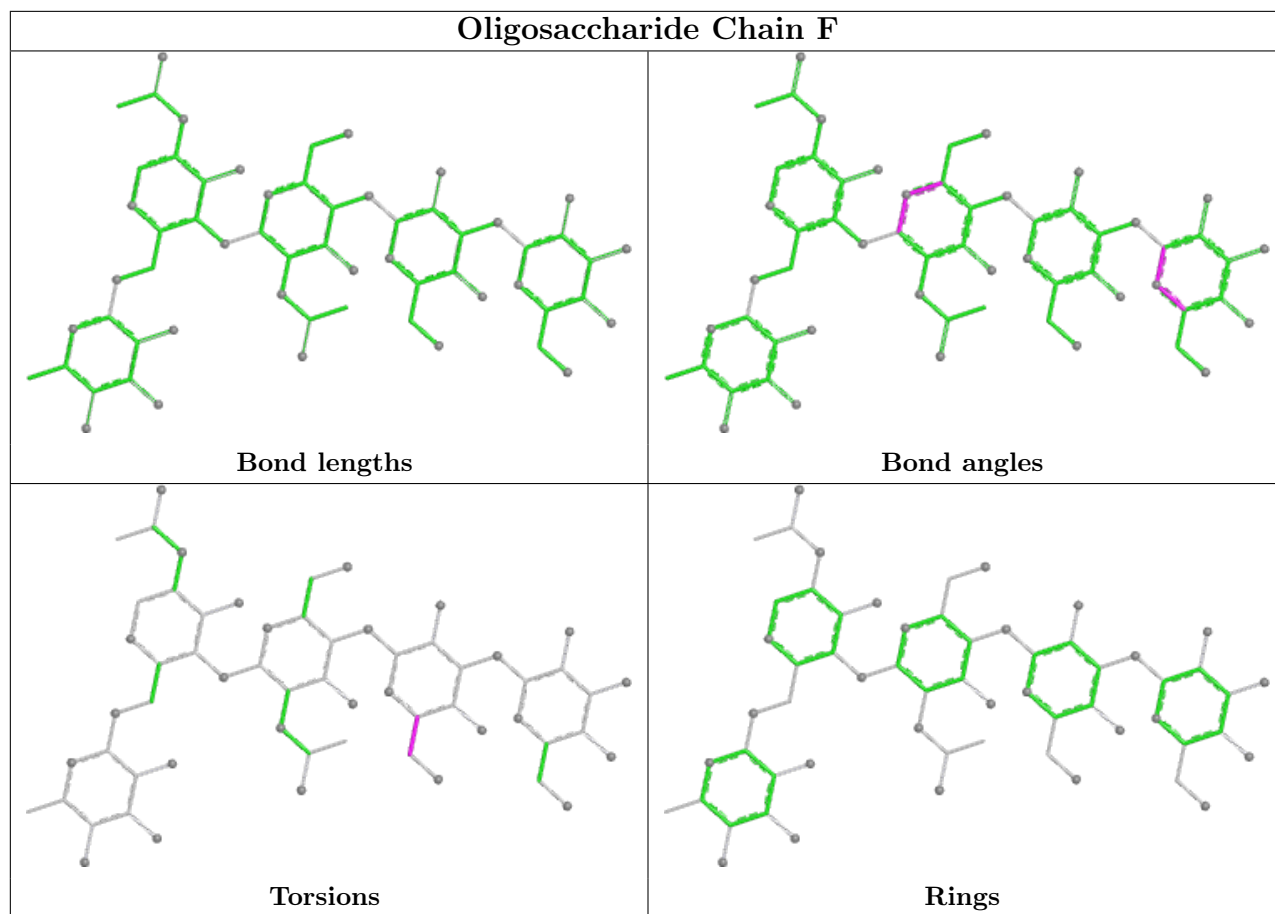
Mol	Chain	Res	Type	Atoms
3	E	3	BMA	O5-C5-C6-O6
3	F	3	BMA	O5-C5-C6-O6
3	E	3	BMA	C4-C5-C6-O6
3	F	3	BMA	C4-C5-C6-O6

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 19 ligands modelled in this entry, 9 are monoatomic - leaving 10 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$
7	NAG	C	605	2	14,14,15	0.34	0	17,19,21	0.76	1 (5%)
8	UEB	C	606	-	24,29,29	1.35	2 (8%)	25,41,41	3.47	3 (12%)
7	NAG	D	604	2	14,14,15	0.33	0	17,19,21	0.63	0
7	NAG	C	604	2	14,14,15	0.33	0	17,19,21	0.70	1 (5%)
7	NAG	D	603	2	14,14,15	0.32	0	17,19,21	0.73	1 (5%)
5	HEM	A	203	1,2	27,50,50	1.07	4 (14%)	17,82,82	2.21	8 (47%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	UEB	D	606	-	24,29,29	1.30	3 (12%)	25,41,41	3.67	4 (16%)
9	BMA	D	601	-	11,11,12	0.34	0	15,15,17	0.67	0
5	HEM	B	203	1,2	27,50,50	1.14	1 (3%)	17,82,82	1.86	6 (35%)
9	BMA	C	607	-	11,11,12	0.36	0	15,15,17	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
7	NAG	C	605	2	-	0/6/23/26	0/1/1/1
8	UEB	C	606	-	-	2/12/12/12	0/4/4/4
7	NAG	D	604	2	-	0/6/23/26	0/1/1/1
7	NAG	C	604	2	-	0/6/23/26	0/1/1/1
7	NAG	D	603	2	-	0/6/23/26	0/1/1/1
5	HEM	A	203	1,2	-	0/6/54/54	-
8	UEB	D	606	-	-	4/12/12/12	0/4/4/4
9	BMA	D	601	-	-	0/2/19/22	0/1/1/1
5	HEM	B	203	1,2	-	0/6/54/54	-
9	BMA	C	607	-	-	0/2/19/22	0/1/1/1

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	C	606	UEB	C21-N20	4.16	1.40	1.33
8	D	606	UEB	C21-N20	3.49	1.39	1.33
5	B	203	HEM	C3B-C2B	-3.38	1.35	1.40
8	D	606	UEB	C4-C2	-2.64	1.49	1.52
5	A	203	HEM	C3B-C2B	-2.62	1.36	1.40
8	C	606	UEB	C4-C2	-2.41	1.50	1.52
5	A	203	HEM	C3D-C2D	-2.31	1.30	1.37
5	A	203	HEM	C3C-CAC	2.15	1.52	1.47
5	A	203	HEM	C3D-C4D	2.11	1.47	1.42
8	D	606	UEB	C16-N17	-2.08	1.44	1.48

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	D	606	UEB	C21-N20-C18	14.21	130.46	115.93
8	C	606	UEB	C21-N20-C18	13.65	129.89	115.93

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	D	606	UEB	C23-C21-N20	-10.11	109.61	123.43
8	C	606	UEB	C23-C21-N20	-9.41	110.57	123.43
5	A	203	HEM	CBD-CAD-C3D	-4.17	104.80	112.48
8	D	606	UEB	C28-C23-C21	3.81	124.43	120.80
8	C	606	UEB	C28-C23-C21	3.75	124.38	120.80
5	A	203	HEM	CBA-CAA-C2A	-3.40	106.21	112.49
5	B	203	HEM	CBD-CAD-C3D	-3.40	106.22	112.48
5	A	203	HEM	C4A-C3A-C2A	-3.12	104.82	107.00
5	A	203	HEM	C1D-C2D-C3D	-2.91	104.97	107.00
5	B	203	HEM	C3B-C4B-NB	2.88	112.93	109.21
5	A	203	HEM	C3B-C4B-NB	2.72	112.73	109.21
5	B	203	HEM	CMB-C2B-C3B	2.65	129.64	124.68
5	A	203	HEM	CMC-C2C-C3C	2.63	129.59	124.68
5	A	203	HEM	CMB-C2B-C3B	2.62	129.58	124.68
7	C	605	NAG	C1-O5-C5	2.59	115.71	112.19
5	A	203	HEM	C2C-C3C-C4C	-2.58	105.10	106.90
7	D	603	NAG	C1-O5-C5	2.57	115.67	112.19
5	B	203	HEM	CMC-C2C-C3C	2.56	129.47	124.68
7	C	604	NAG	C1-O5-C5	2.52	115.61	112.19
8	D	606	UEB	C28-N17-C18	-2.24	117.08	120.83
5	B	203	HEM	CBA-CAA-C2A	-2.22	108.38	112.49
5	B	203	HEM	CAD-CBD-CGD	2.08	116.15	112.67

There are no chirality outliers.

All (6) torsion outliers are listed below:

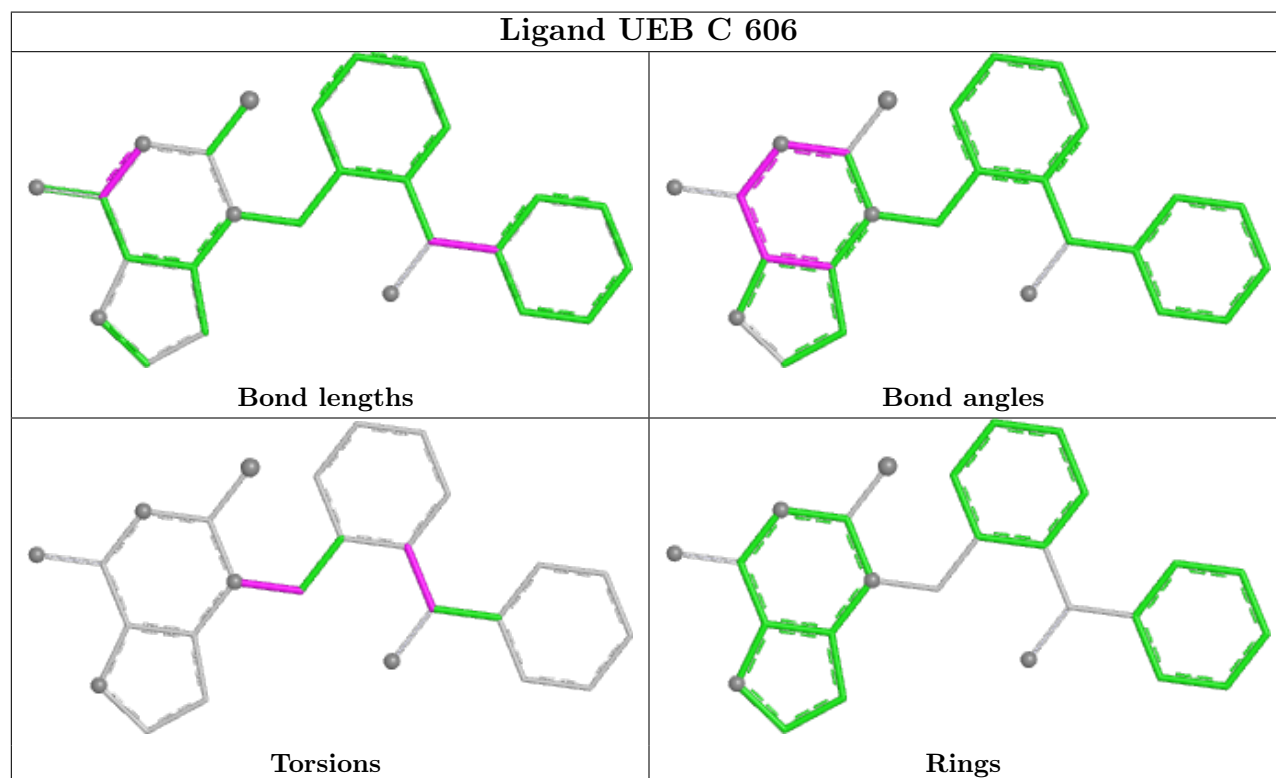
Mol	Chain	Res	Type	Atoms
8	C	606	UEB	C15-C10-C2-N1
8	C	606	UEB	C15-C16-N17-C28
8	D	606	UEB	C15-C16-N17-C28
8	D	606	UEB	N1-C2-C4-C9
8	D	606	UEB	C15-C10-C2-N1
8	D	606	UEB	N1-C2-C4-C5

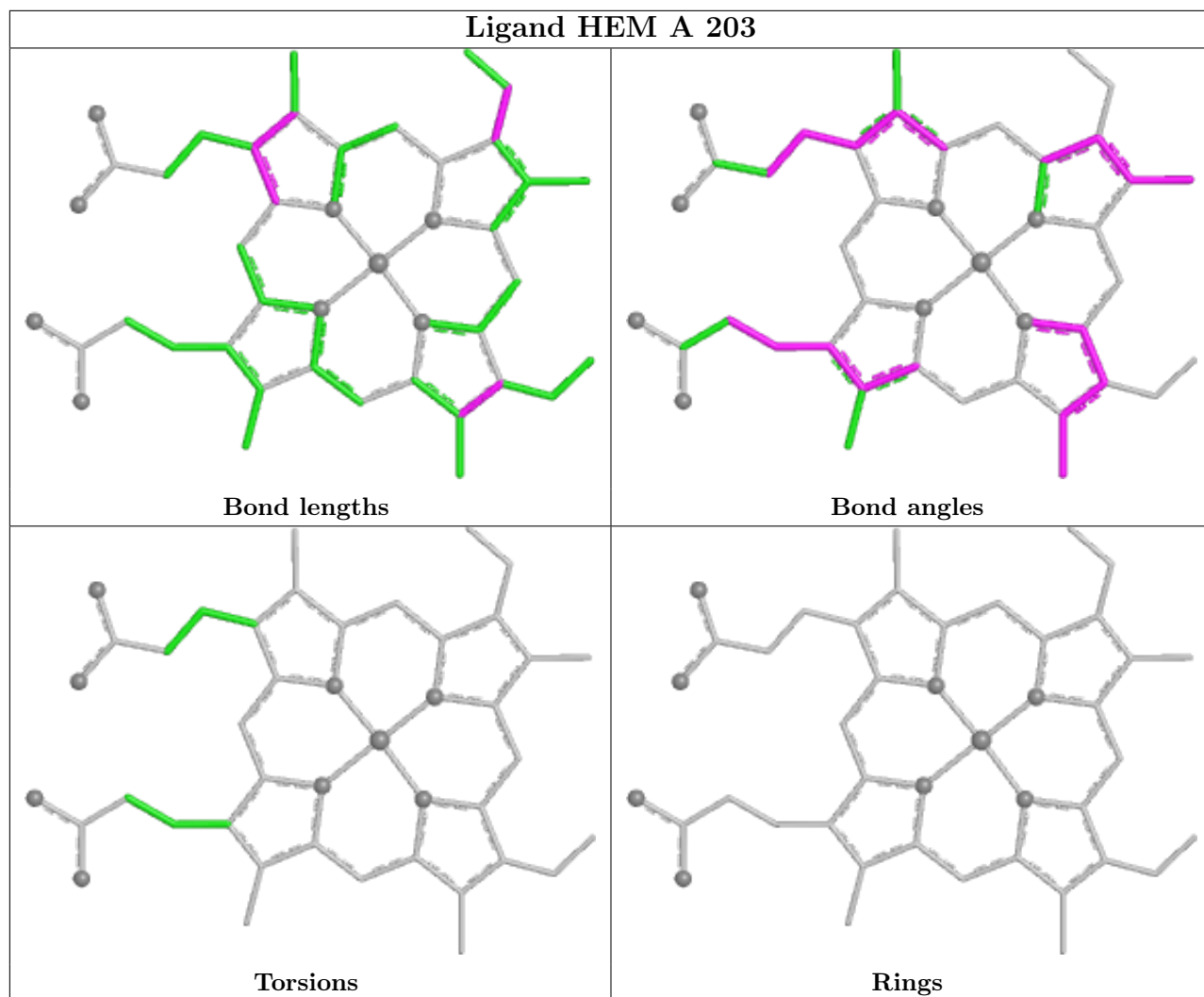
There are no ring outliers.

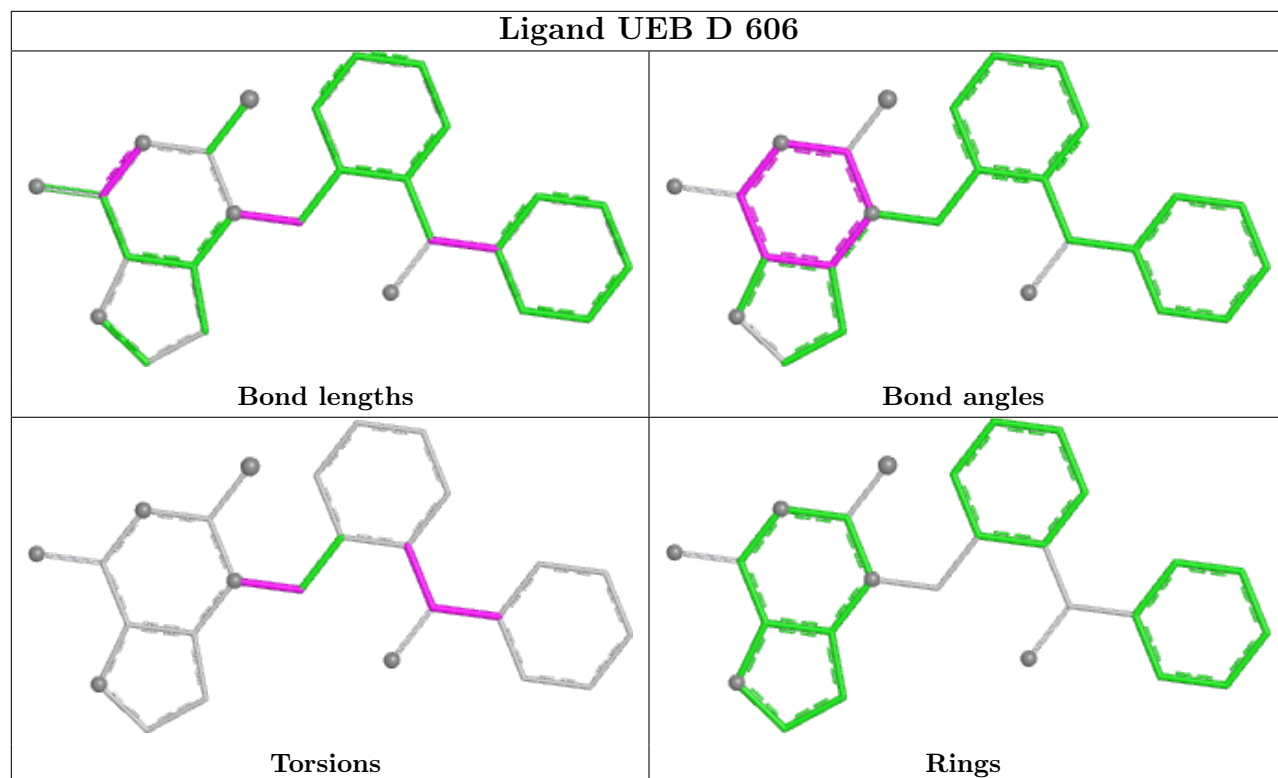
2 monomers are involved in 5 short contacts:

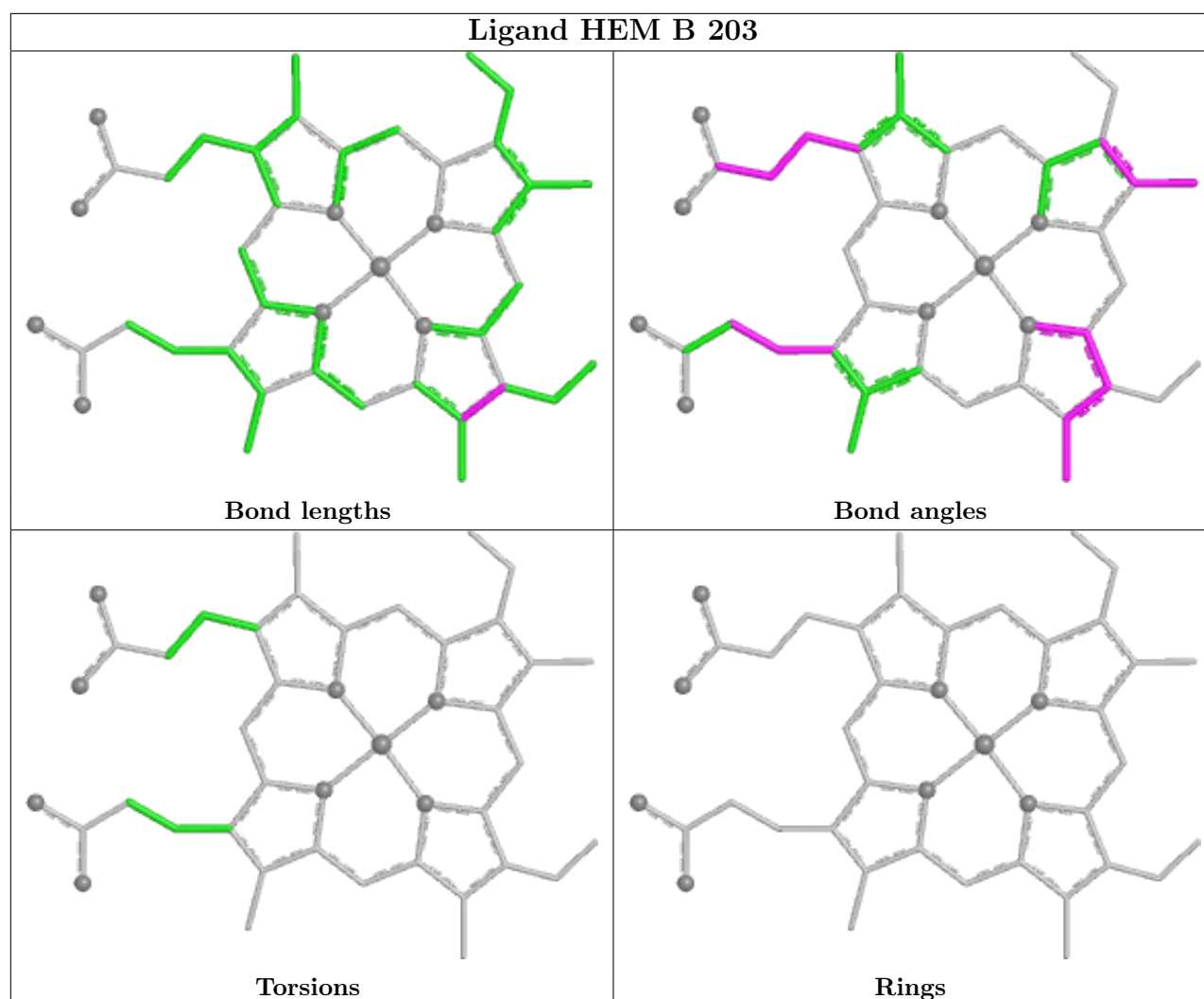
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	203	HEM	2	0
5	B	203	HEM	3	0

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









## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.



## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	104/105 (99%)	0.80	8 (7%) 13 17	10, 17, 43, 65	0
1	B	103/105 (98%)	0.81	7 (6%) 17 21	12, 20, 40, 67	0
2	C	463/466 (99%)	0.78	35 (7%) 13 17	8, 19, 39, 61	0
2	D	463/466 (99%)	0.96	47 (10%) 6 8	11, 24, 45, 60	0
All	All	1133/1142 (99%)	0.86	97 (8%) 10 13	8, 21, 43, 67	0

All (97) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	355	PRO	11.0
2	D	217	HIS	7.8
1	A	1	CYS	6.1
2	C	226	ARG	5.8
2	D	355	PRO	5.5
1	B	3	GLU	5.4
1	B	4	GLN	5.2
2	D	218	ASP	4.8
2	C	141	ALA	4.7
2	C	270	ARG	4.5
1	A	104	ALA	4.4
2	D	269	PRO	4.1
2	D	216	LEU	4.0
2	D	568	PRO	4.0
2	D	576	ARG	3.8
2	D	194	LEU	3.8
2	D	573	ALA	3.5
2	D	267	LEU	3.5
2	C	271	TRP	3.5
2	D	569	ALA	3.4
2	D	226	ARG	3.3

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
2	C	140	GLN	3.3
2	D	157	ASN	3.3
2	D	549	ASN	3.2
1	B	1	CYS	3.2
2	D	197	LEU	3.2
1	B	32	TRP	3.1
2	D	268	ASN	3.0
1	A	4	GLN	3.0
2	C	191	SER	2.9
2	C	217	HIS	2.9
2	C	575	TRP	2.9
2	D	122	GLN	2.8
2	C	359	VAL	2.8
2	D	141	ALA	2.8
2	C	197	LEU	2.8
2	D	324	ILE	2.8
2	D	356	ASN	2.7
1	B	2	PRO	2.7
2	C	568	PRO	2.7
2	C	327	VAL	2.7
2	D	562	VAL	2.7
2	D	206	ASN	2.6
2	C	557	TYR	2.6
2	D	283	ILE	2.6
2	D	522	MET	2.6
1	A	93	LEU	2.6
2	D	271	TRP	2.6
2	D	135	PRO	2.6
2	C	314	ARG	2.5
2	C	523	GLN	2.5
2	C	229	ARG	2.5
2	D	202	ARG	2.5
1	B	22	LEU	2.5
2	C	357	PRO	2.5
2	D	229	ARG	2.5
2	D	125	CYS	2.5
2	D	261	ALA	2.5
2	D	528	LEU	2.4
2	D	575	TRP	2.4
2	D	203	PHE	2.4
2	C	185	ARG	2.4
1	A	2	PRO	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	C	549	ASN	2.4
2	C	358	ARG	2.3
2	C	487	ARG	2.3
2	D	120	VAL	2.3
1	A	30	VAL	2.3
2	D	155	GLY	2.3
2	D	530	GLN	2.3
2	D	213	PHE	2.3
2	D	137	ILE	2.3
1	B	29	PHE	2.3
2	C	448	VAL	2.2
2	C	324	ILE	2.2
2	D	204	GLN	2.2
2	C	328	PHE	2.2
2	C	200	ASN	2.2
2	D	123	PRO	2.2
1	A	47	TRP	2.1
2	C	430	LEU	2.1
2	C	458	LEU	2.1
2	C	123	PRO	2.1
2	C	196	LEU	2.1
2	D	513	TRP	2.1
2	C	182	PRO	2.1
1	A	3	GLU	2.1
2	D	131	PRO	2.1
2	D	354	GLU	2.1
2	C	199	VAL	2.1
2	C	277	TYR	2.1
2	D	567	LEU	2.1
2	C	190	MET	2.1
2	D	119	CYS	2.1
2	D	563	ASN	2.0
2	D	196	LEU	2.0
2	C	339	ILE	2.0

## 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( $\text{\AA}^2$ )	Q<0.9
2	CSO	D	150	7/8	0.94	0.12	14,18,25,27	0
2	CSO	C	150	7/8	0.95	0.11	14,16,20,22	0

### 6.3 Carbohydrates [i](#)

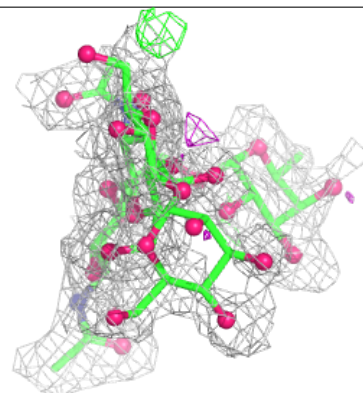
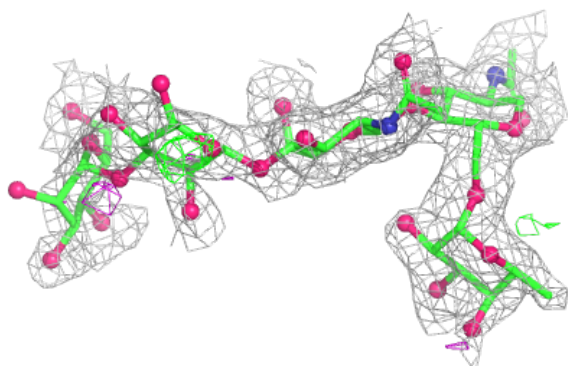
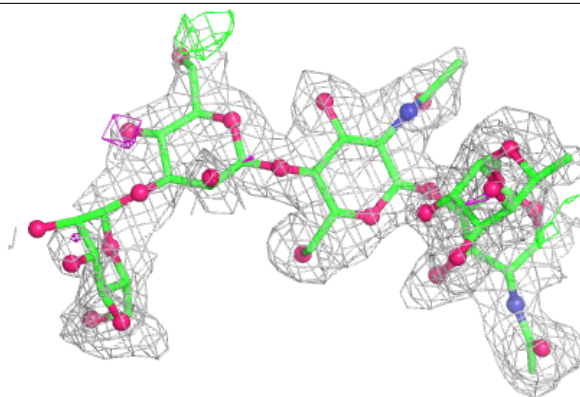
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MAN	E	4	11/12	0.15	0.37	61,63,64,64	0
3	MAN	F	4	11/12	0.40	0.35	58,60,63,63	0
3	FUC	F	5	10/11	0.61	0.34	49,53,55,55	0
3	BMA	E	3	11/12	0.65	0.31	40,46,50,56	0
3	BMA	F	3	11/12	0.73	0.22	43,49,55,58	0
3	NAG	F	2	14/15	0.83	0.17	13,23,29,36	0
3	FUC	E	5	10/11	0.87	0.17	28,31,35,38	0
3	NAG	F	1	14/15	0.87	0.15	25,28,34,40	0
3	NAG	E	1	14/15	0.90	0.13	18,21,26,26	0
3	NAG	E	2	14/15	0.90	0.16	17,22,29,35	0

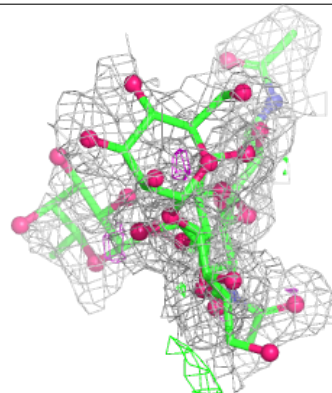
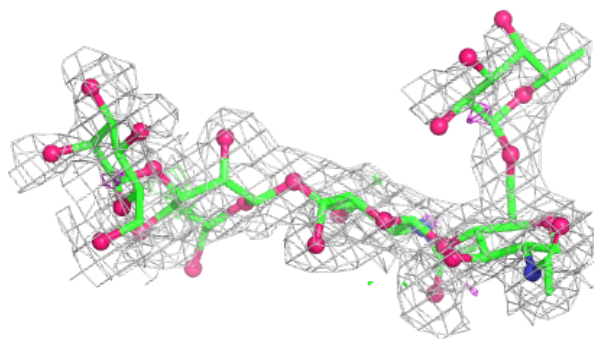
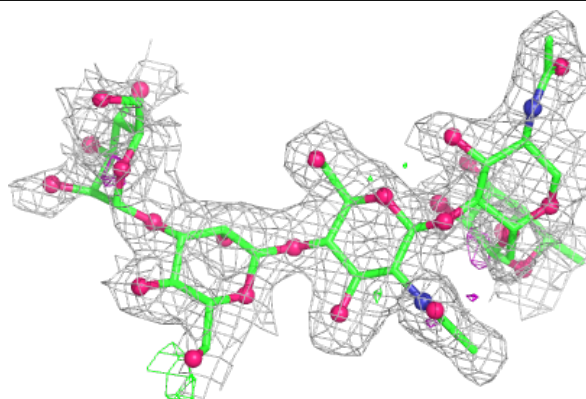
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**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)

**Electron density around Chain F:**

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

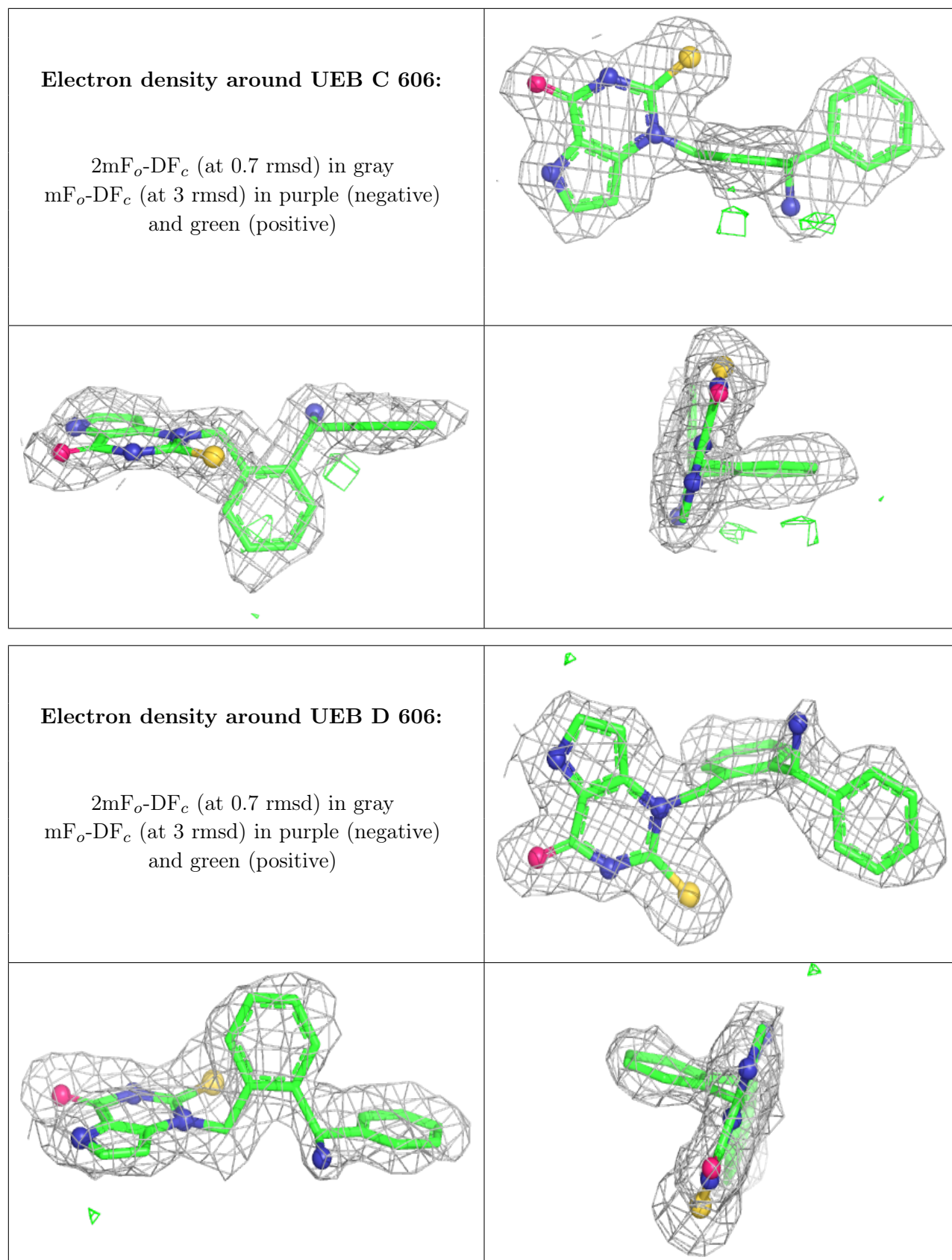


## 6.4 Ligands

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
9	BMA	C	607	11/12	0.23	0.69	116,118,119,119	0
9	BMA	D	601	11/12	0.49	0.47	86,87,89,90	0
7	NAG	D	604	14/15	0.81	0.23	42,44,46,47	0
7	NAG	C	604	14/15	0.83	0.23	29,32,33,36	0
4	CL	D	602	1/1	0.87	0.40	81,81,81,81	0
7	NAG	D	603	14/15	0.88	0.21	31,35,37,38	0
7	NAG	C	605	14/15	0.88	0.18	26,33,34,34	0
8	UEB	C	606	26/26	0.89	0.20	18,34,37,38	0
8	UEB	D	606	26/26	0.90	0.21	18,26,32,33	0
5	HEM	A	203	43/43	0.92	0.23	19,20,22,33	0
5	HEM	B	203	43/43	0.92	0.20	22,23,25,31	0
6	CA	D	605	1/1	0.97	0.16	10,10,10,10	1
4	CL	C	603	1/1	0.98	0.18	33,33,33,33	0
4	CL	C	602	1/1	0.98	0.12	23,23,23,23	0
6	CA	C	601	1/1	0.98	0.07	12,12,12,12	0
4	CL	B	202	1/1	0.99	0.09	21,21,21,21	0
4	CL	A	201	1/1	0.99	0.16	11,11,11,11	0
4	CL	A	202	1/1	0.99	0.13	21,21,21,21	0
4	CL	B	201	1/1	0.99	0.15	12,12,12,12	0

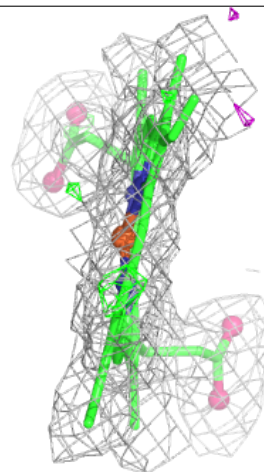
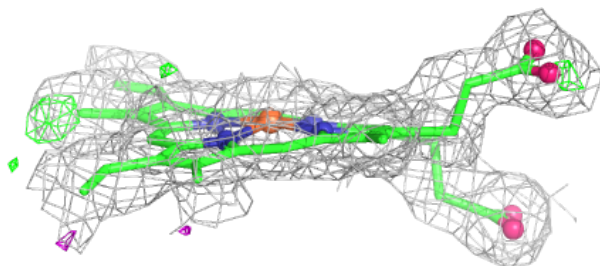
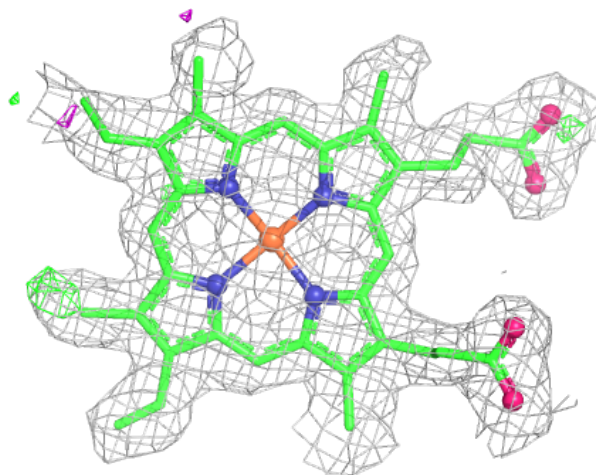
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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



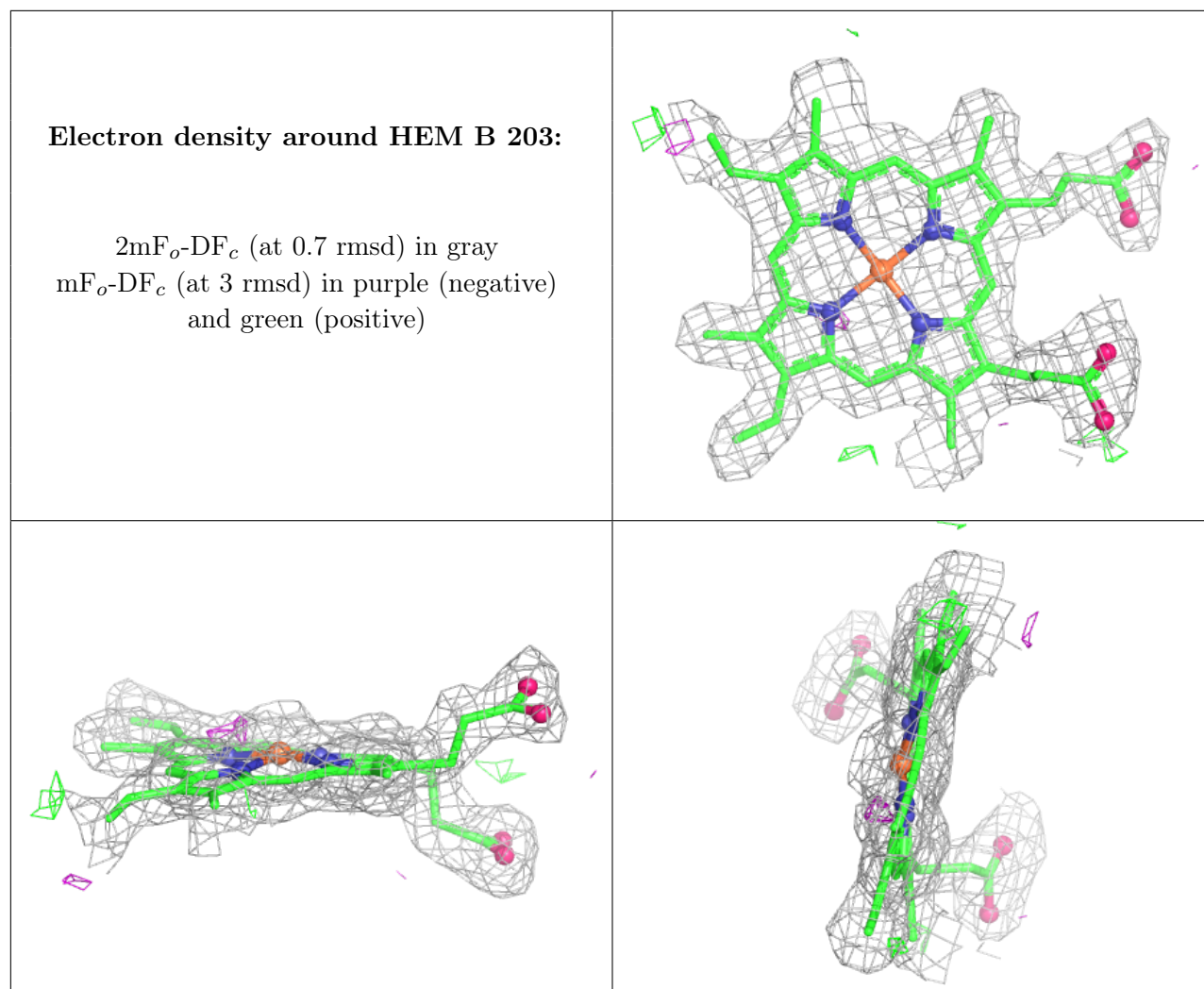


**Electron density around HEM A 203:**

$2mF_o-DF_c$  (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.