Full wwPDB X-ray Structure Validation Report (i)

Jan 6, 2024 – 09:53 pm GMT

PDB ID	:	5MFA
Title	:	Crystal structure of human promyeloperoxidase (proMPO)
Authors	:	Grishkovskaya, I.; Furtmueller, P.G.; Obinger, C.; Djinovic-Carugo, K.
Deposited on	:	2016-11-17
Resolution	:	1.20 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

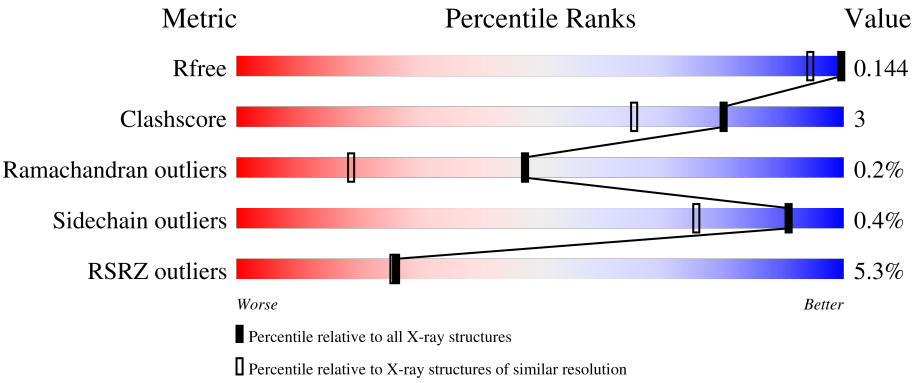
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.20 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	1223 (1.22-1.18)
Clashscore	141614	1286 (1.22-1.18)
Ramachandran outliers	138981	1240 (1.22-1.18)
Sidechain outliers	138945	1239 (1.22-1.18)
RSRZ outliers	127900	1200 (1.22-1.18)

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 >=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 <=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	697	4% 80% 5% 16%
2	B	2	50% 50%
2	C	2	100%
3	D	5	40% 60%

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
2	NAG	C	2	-	-	-	X
3	MAN	D	5	-	-	-	X
4	NAG	A	806	-	-	-	X

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 10867 atoms, of which 4895 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.

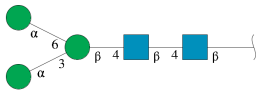
- Molecule 1 is a protein called Myeloperoxidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	588	9754	3138	4785	900	891	40	0	49	0

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

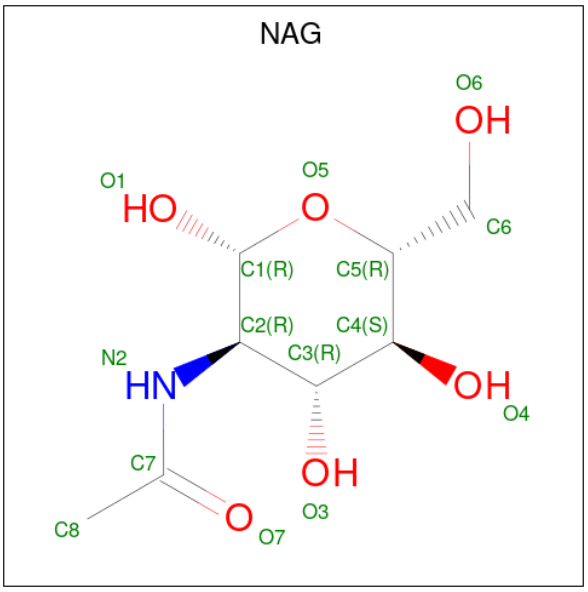
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
2	B	2	55	16	27	2	10	0	0	0
2	C	2	55	16	27	2	10	0	0	0

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



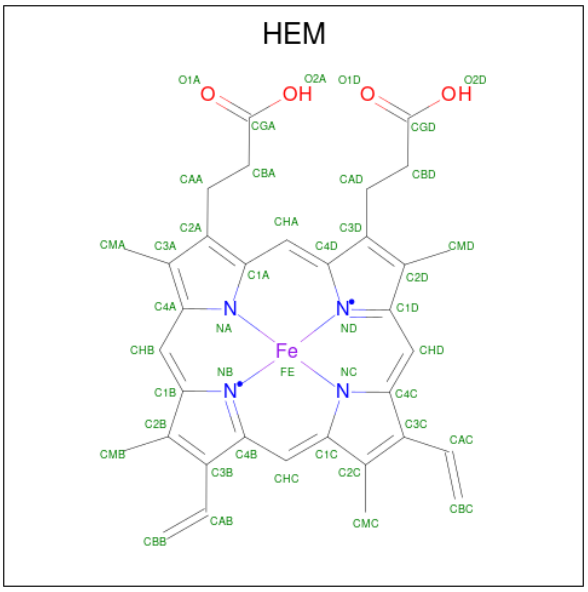
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	D	5	61	34	2	25	0	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 5 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	A	1	Total	C	Fe	H	N	O	0	0
			71	34	1	28	4	4		

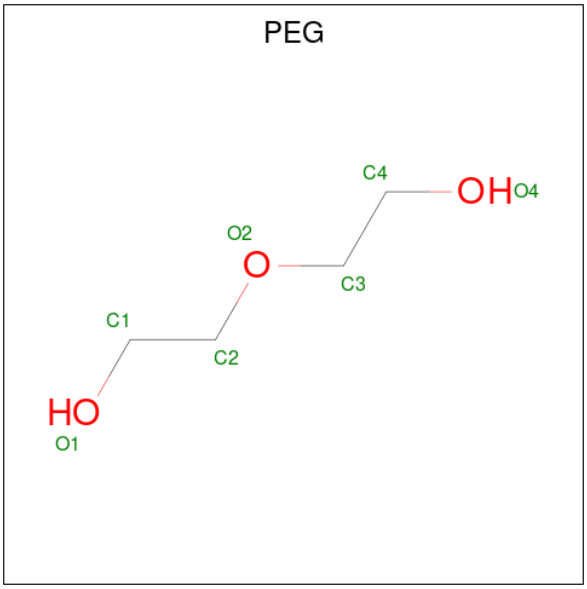
- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Ca 1 1	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

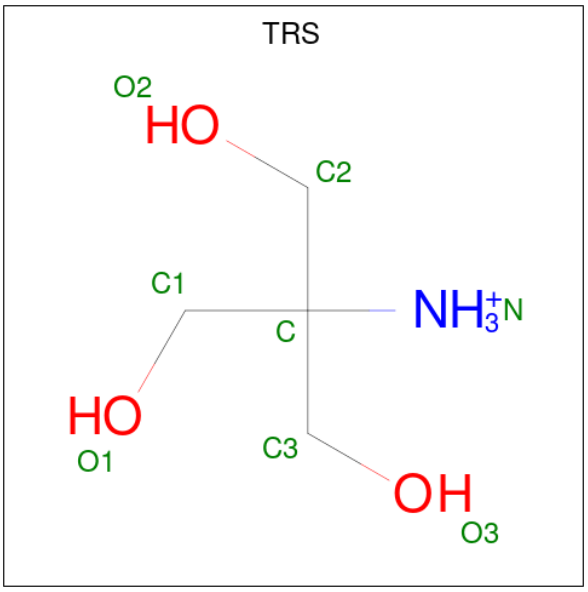
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	5	Total Cl 5 5	0	0

- Molecule 8 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total C O 7 4 3	0	0
8	A	1	Total C O 7 4 3	0	0
8	A	1	Total C O 7 4 3	0	0
8	A	1	Total C O 7 4 3	0	0
8	A	1	Total C O 7 4 3	0	0

- Molecule 9 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: $C_4H_{12}NO_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	A	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 10 is water.

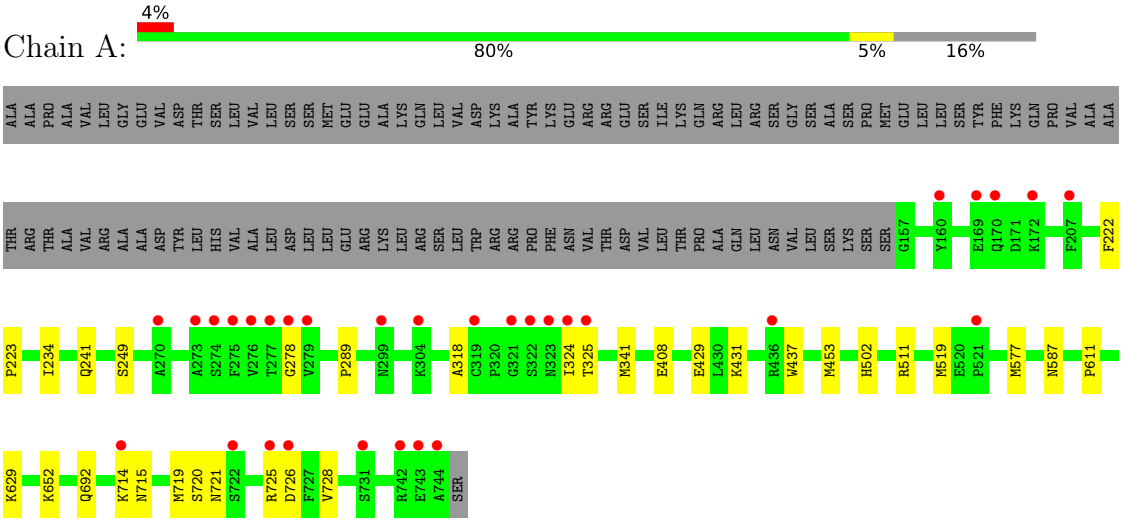
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	766	Total	O	0	0
			766	766		

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: Myeloperoxidase

Chain A:

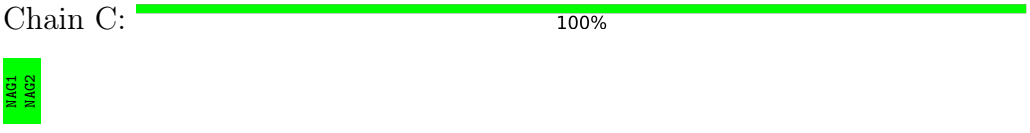


• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

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

Chain C:



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

Chain D:


NAG1
NAG2
BMA3
MAN4
MAN5

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	106.15Å 109.15Å 83.96Å 90.00° 122.66° 90.00°	Depositor
Resolution (Å)	44.68 – 1.20 44.68 – 1.20	Depositor EDS
% Data completeness (in resolution range)	90.2 (44.68-1.20) 90.2 (44.68-1.20)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.39 (at 1.20Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, R_{free}	0.123 , 0.143 0.124 , 0.144	Depositor DCC
R_{free} test set	11279 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	12.7	Xtriage
Anisotropy	0.245	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 52.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	10867	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.39% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: MAN, HEM, CL, CSO, NAG, TRS, PEG, CA, 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	0.38	0/5214	0.65	0/7066

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
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	715[A]	ASN	Peptide

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4969	4785	4888	25	0
2	B	28	27	25	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	28	27	25	0	0
3	D	61	0	52	1	0
4	A	28	28	26	0	0
5	A	43	28	30	1	0
6	A	1	0	0	0	0
7	A	5	0	0	0	0
8	A	35	0	46	3	0
9	A	8	0	12	0	0
10	A	766	0	0	15	1
All	All	5972	4895	5104	28	1

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 3.

All (28) 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:A:714[B]:LYS:HE2	1:A:726[B]:ASP:HA	1.41	1.00
1:A:692[C]:GLN:HG3	10:A:989:HOH:O	1.68	0.94
1:A:721[A]:ASN:ND2	10:A:903:HOH:O	2.13	0.81
1:A:408[B]:GLU:OE1	10:A:901:HOH:O	2.00	0.79
1:A:714[B]:LYS:CE	1:A:726[B]:ASP:HA	2.15	0.74
1:A:577:MET:O	8:A:823:PEG:H22	1.87	0.74
1:A:502:HIS:HD1	1:A:587:ASN:HD21	1.34	0.73
1:A:721[A]:ASN:O	10:A:902:HOH:O	2.10	0.70
1:A:429[B]:GLU:OE2	10:A:904:HOH:O	2.13	0.65
10:A:905:HOH:O	3:D:2:NAG:O7	2.16	0.61
1:A:714[B]:LYS:HB2	1:A:728:VAL:HG13	1.82	0.61
1:A:241[A]:GLN:NE2	10:A:912:HOH:O	2.33	0.60
1:A:453[A]:MET:HE1	10:A:1432:HOH:O	2.03	0.59
1:A:725[B]:ARG:NE	1:A:726[B]:ASP:OD1	2.37	0.58
1:A:249:SER:HB3	1:A:720[A]:SER:O	2.07	0.54
1:A:721[A]:ASN:HB2	10:A:902:HOH:O	2.08	0.53
8:A:819:PEG:H21	10:A:1143:HOH:O	2.09	0.52
1:A:234[B]:ILE:CD1	1:A:629[B]:LYS:HB3	2.40	0.52
1:A:222[B]:PHE:CG	1:A:223:PRO:HD2	2.46	0.51
5:A:812:HEM:HMC2	5:A:812:HEM:HBC2	1.95	0.48
1:A:289:PRO:HA	10:A:1333:HOH:O	2.13	0.48
1:A:431:LYS:HD2	10:A:906:HOH:O	2.14	0.47
1:A:652[B]:LYS:CE	10:A:1439:HOH:O	2.64	0.46
1:A:324:ILE:HD12	1:A:325:THR:N	2.32	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:511:ARG:HB3	1:A:519[A]:MET:HE2	2.00	0.44
1:A:437:TRP:O	10:A:906:HOH:O	2.21	0.44
1:A:611:PRO:HA	8:A:820:PEG:H11	2.01	0.42
1:A:318:ALA:O	10:A:907:HOH:O	2.22	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
10:A:1423:HOH:O	10:A:1566:HOH:O[2_455]	1.86	0.34

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	633/697 (91%)	622 (98%)	10 (2%)	1 (0%)	47 19

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	278	GLY

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	560/606 (92%)	557 (100%)	3 (0%)	88 67

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	341	MET
1	A	719[A]	MET
1	A	719[B]	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

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	CSO	A	316[B]	-	3,6,7	0.67	0	0,6,8	-	-
1	CSO	A	316[A]	-	3,6,7	0.52	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
1	CSO	A	316[B]	-	-	0/1/5/7	-
1	CSO	A	316[A]	-	-	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)

9 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	NAG	B	1	2,1	14,14,15	0.62	1 (7%)	17,19,21	0.54	0
2	NAG	B	2	2	14,14,15	0.21	0	17,19,21	0.52	0
2	NAG	C	1	2,1	14,14,15	0.25	0	17,19,21	0.48	0
2	NAG	C	2	2	14,14,15	0.29	0	17,19,21	0.36	0
3	NAG	D	1	3,1	14,14,15	0.51	0	17,19,21	0.49	0
3	NAG	D	2	3	14,14,15	0.58	0	17,19,21	0.52	0
3	BMA	D	3	3	11,11,12	0.70	0	15,15,17	0.76	0
3	MAN	D	4	3	11,11,12	0.71	0	15,15,17	1.11	2 (13%)
3	MAN	D	5	3	11,11,12	0.75	0	15,15,17	0.98	2 (13%)

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	NAG	B	1	2,1	-	0/6/23/26	0/1/1/1
2	NAG	B	2	2	-	0/6/23/26	0/1/1/1
2	NAG	C	1	2,1	-	0/6/23/26	0/1/1/1
2	NAG	C	2	2	-	0/6/23/26	0/1/1/1
3	NAG	D	1	3,1	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	D	2	3	-	0/6/23/26	0/1/1/1
3	BMA	D	3	3	-	0/2/19/22	0/1/1/1
3	MAN	D	4	3	-	0/2/19/22	0/1/1/1
3	MAN	D	5	3	-	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	NAG	O5-C1	-2.02	1.40	1.43

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	4	MAN	C1-O5-C5	2.71	115.86	112.19
3	D	5	MAN	C1-O5-C5	2.33	115.34	112.19
3	D	4	MAN	O2-C2-C3	-2.13	105.86	110.14
3	D	5	MAN	O2-C2-C3	-2.02	106.09	110.14

There are no chirality outliers.

There are no torsion outliers.

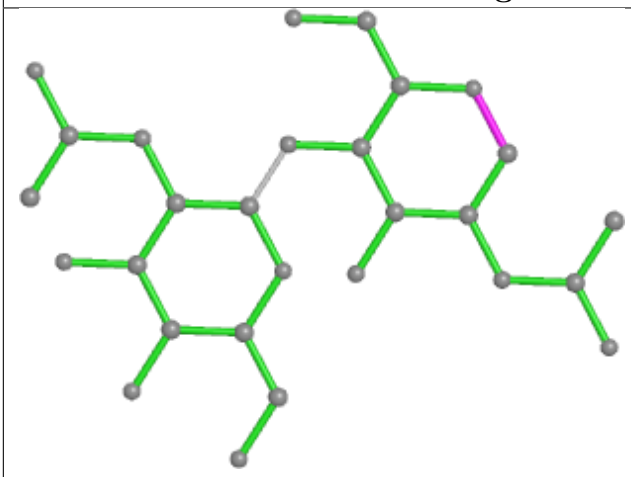
There are no ring outliers.

1 monomer is involved in 1 short contact:

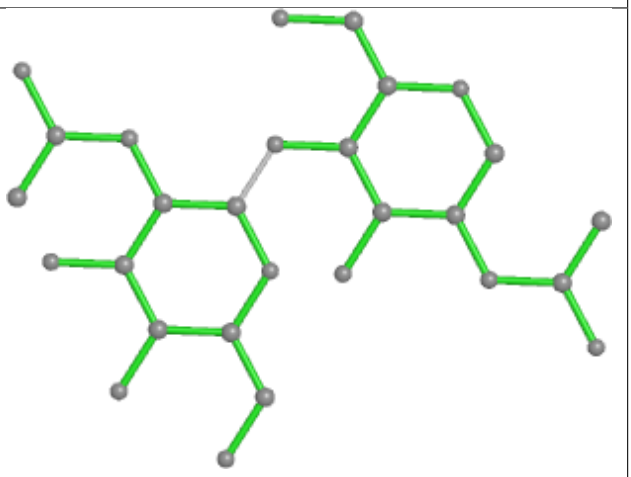
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	2	NAG	1	0

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

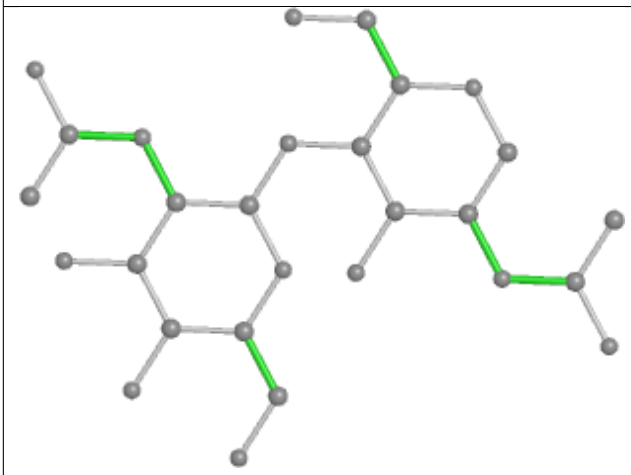
Oligosaccharide Chain B



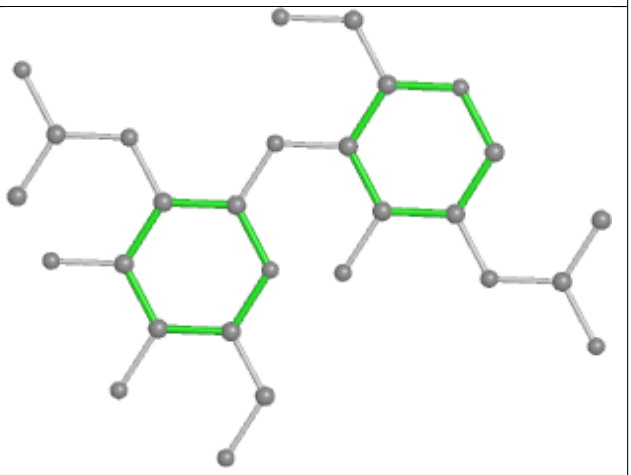
Bond lengths



Bond angles

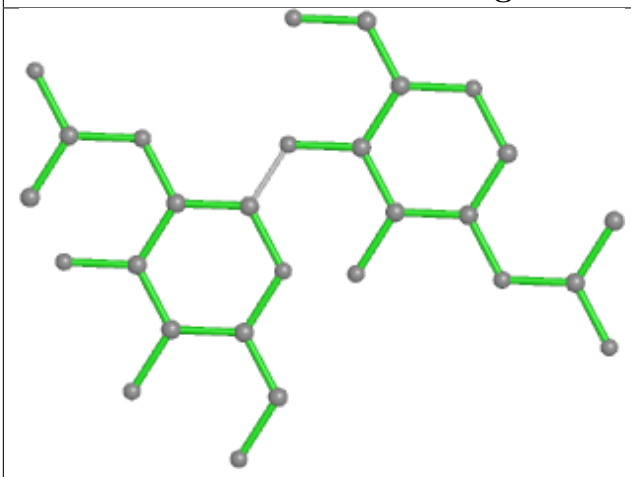


Torsions

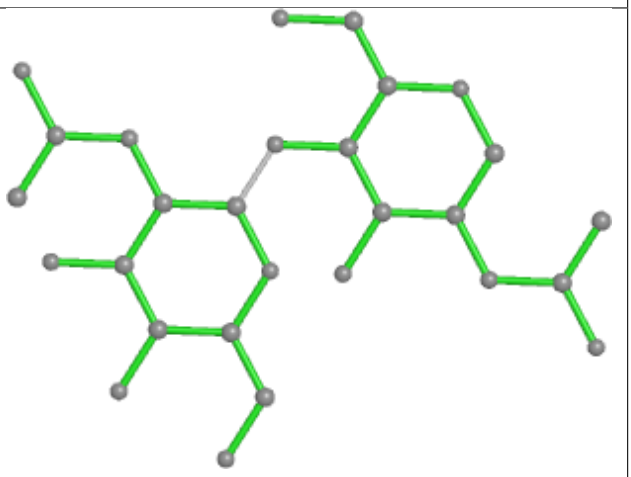


Rings

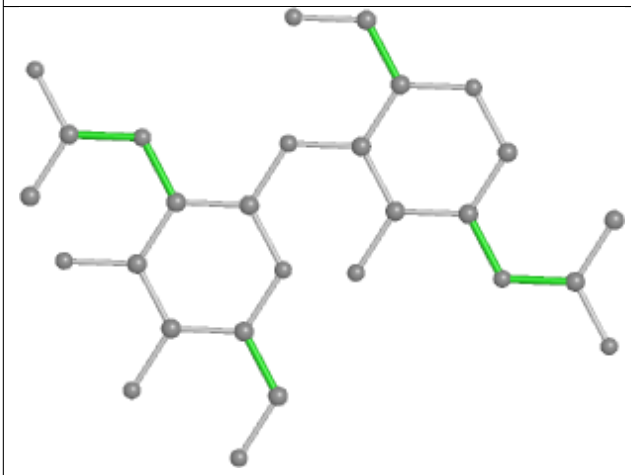
Oligosaccharide Chain C



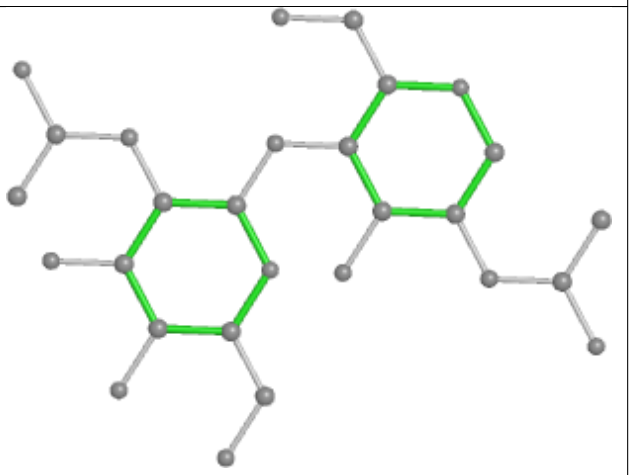
Bond lengths



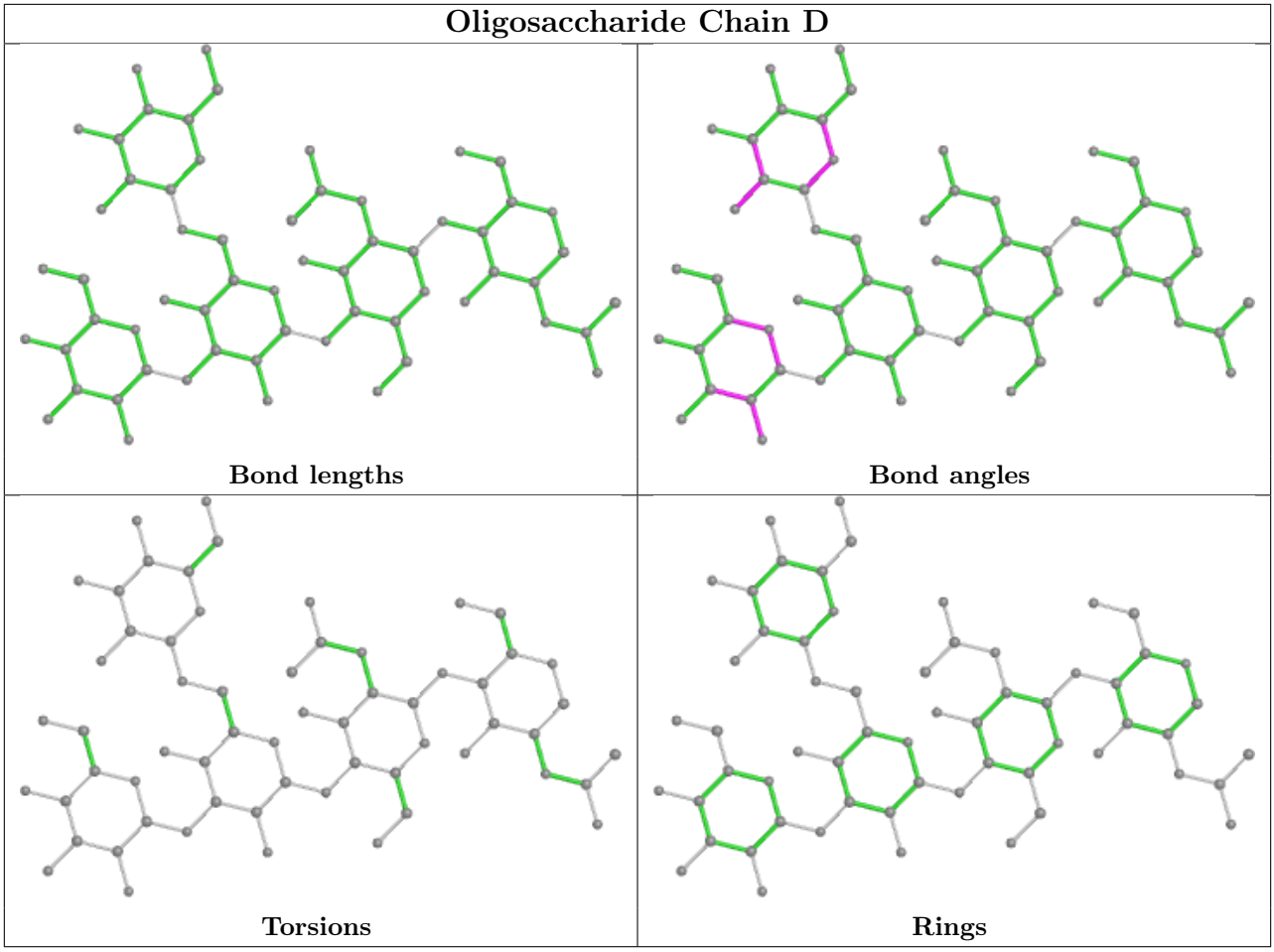
Bond angles



Torsions



Rings



5.6 Ligand geometry (i)

Of 15 ligands modelled in this entry, 6 are monoatomic - leaving 9 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
8	PEG	A	820	-	6,6,6	0.46	0	5,5,5	0.46	0
8	PEG	A	822	-	6,6,6	0.44	0	5,5,5	0.51	0
4	NAG	A	801	1	14,14,15	0.27	0	17,19,21	0.32	0
9	TRS	A	824	-	7,7,7	0.30	0	9,9,9	0.33	0
8	PEG	A	819	-	6,6,6	0.48	0	5,5,5	0.53	0
4	NAG	A	806	1	14,14,15	0.43	0	17,19,21	0.41	0
8	PEG	A	823	-	6,6,6	0.47	0	5,5,5	0.38	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	HEM	A	812	1	41,50,50	1.40	5 (12%)	45,82,82	1.49	7 (15%)
8	PEG	A	821	-	6,6,6	0.48	0	5,5,5	0.49	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	PEG	A	820	-	-	1/4/4/4	-
8	PEG	A	822	-	-	0/4/4/4	-
4	NAG	A	801	1	-	2/6/23/26	0/1/1/1
9	TRS	A	824	-	-	6/9/9/9	-
8	PEG	A	819	-	-	1/4/4/4	-
4	NAG	A	806	1	-	0/6/23/26	0/1/1/1
8	PEG	A	823	-	-	3/4/4/4	-
5	HEM	A	812	1	-	4/12/54/54	-
8	PEG	A	821	-	-	0/4/4/4	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	812	HEM	FE-NB	3.26	2.13	1.96
5	A	812	HEM	FE-ND	2.86	2.11	1.96
5	A	812	HEM	C3C-C2C	-2.73	1.36	1.40
5	A	812	HEM	C3C-CAC	2.59	1.53	1.47
5	A	812	HEM	CAB-C3B	2.01	1.52	1.47

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	812	HEM	C4B-CHC-C1C	4.26	128.18	122.56
5	A	812	HEM	C4C-CHD-C1D	3.05	126.59	122.56
5	A	812	HEM	C1B-NB-C4B	2.67	107.83	105.07
5	A	812	HEM	CMC-C2C-C3C	2.52	129.40	124.68
5	A	812	HEM	CMB-C2B-C1B	-2.27	121.58	125.04
5	A	812	HEM	C4A-C3A-C2A	2.19	108.52	107.00
5	A	812	HEM	CMA-C3A-C4A	-2.19	125.09	128.46

There are no chirality outliers.

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	A	824	TRS	C2-C-C1-O1
9	A	824	TRS	C3-C-C1-O1
9	A	824	TRS	N-C-C1-O1
8	A	823	PEG	C4-C3-O2-C2
8	A	819	PEG	C4-C3-O2-C2
8	A	820	PEG	O1-C1-C2-O2
8	A	823	PEG	O1-C1-C2-O2
9	A	824	TRS	C1-C-C3-O3
8	A	823	PEG	O2-C3-C4-O4
4	A	801	NAG	C4-C5-C6-O6
9	A	824	TRS	N-C-C3-O3
4	A	801	NAG	O5-C5-C6-O6
9	A	824	TRS	C2-C-C3-O3
5	A	812	HEM	CAD-CBD-CGD-O1D
5	A	812	HEM	CAA-CBA-CGA-O2A
5	A	812	HEM	CAD-CBD-CGD-O2D
5	A	812	HEM	CAA-CBA-CGA-O1A

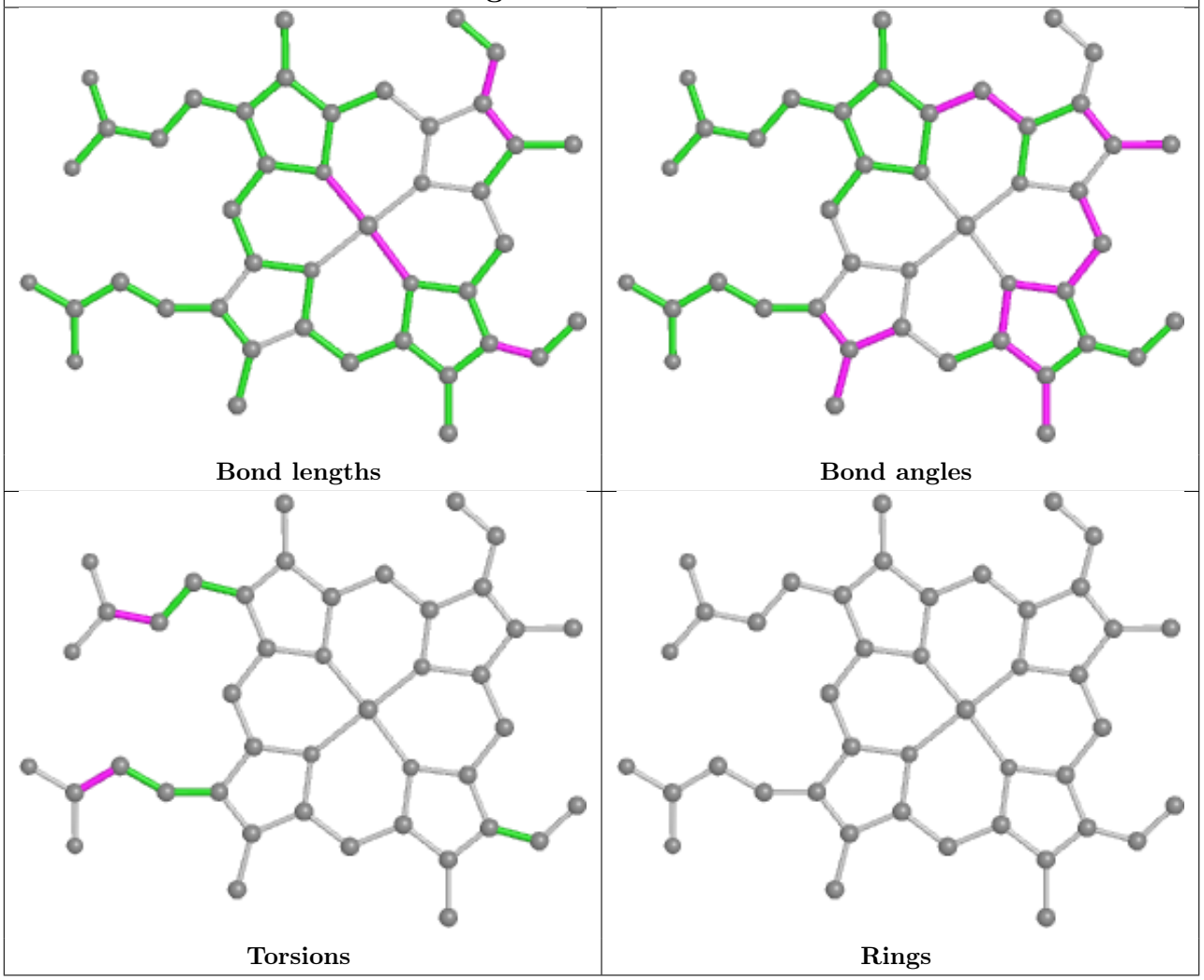
There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	820	PEG	1	0
8	A	819	PEG	1	0
8	A	823	PEG	1	0
5	A	812	HEM	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 then 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.

Ligand HEM A 812



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^{th} 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^2)	Q<0.9
1	A	587/697 (84%)	0.19	31 (5%) 26 26	10, 15, 31, 71	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	324	ILE	13.2
1	A	277	THR	11.6
1	A	276	VAL	11.1
1	A	744	ALA	10.1
1	A	278	GLY	7.8
1	A	275	PHE	6.7
1	A	321	GLY	5.8
1	A	273	ALA	5.6
1	A	279	VAL	4.7
1	A	725[A]	ARG	4.5
1	A	322	SER	4.3
1	A	169	GLU	3.9
1	A	274	SER	3.7
1	A	207[A]	PHE	3.7
1	A	160	TYR	3.7
1	A	436	ARG	3.5
1	A	325	THR	3.5
1	A	521	PRO	3.4
1	A	323	ASN	3.3
1	A	742	ARG	3.3
1	A	170	GLN	3.1
1	A	270	ALA	2.5
1	A	722[A]	SER	2.4
1	A	319	CYS	2.3
1	A	726[A]	ASP	2.3
1	A	731[A]	SER	2.3
1	A	743	GLU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	304	LYS	2.2
1	A	714[A]	LYS	2.2
1	A	299	ASN	2.1
1	A	172	LYS	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^{th} 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^2)	Q<0.9
1	CSO	A	316[A]	7/8	0.97	0.09	12,14,16,17	9
1	CSO	A	316[B]	7/8	0.97	0.09	12,17,20,22	9

6.3 Carbohydrates (i)

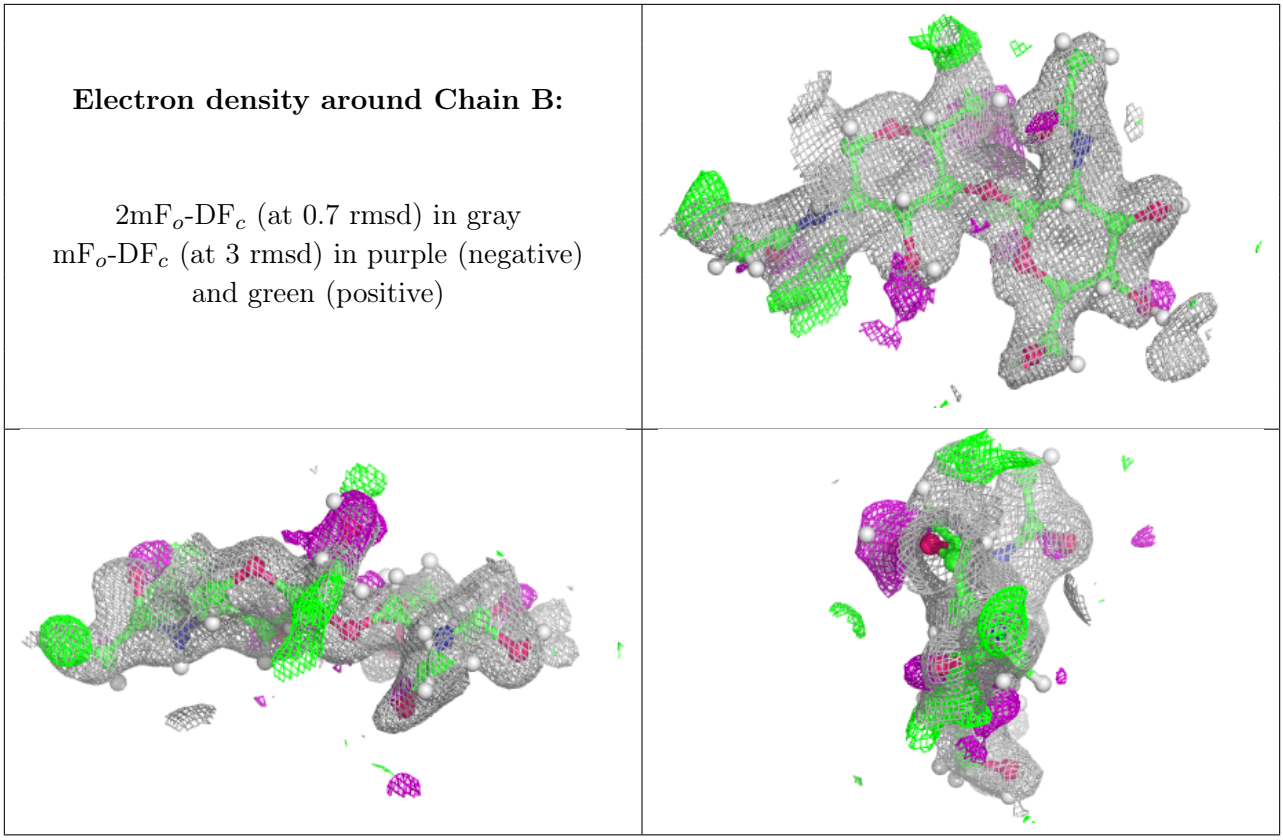
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	BMA	D	3	11/12	0.62	0.25	54,65,73,76	0
2	NAG	C	2	14/15	0.67	0.42	51,56,67,68	28
3	MAN	D	5	11/12	0.67	0.41	52,74,92,92	0
3	MAN	D	4	11/12	0.70	0.22	40,62,79,80	0
2	NAG	C	1	14/15	0.84	0.24	38,46,55,57	0
3	NAG	D	2	14/15	0.85	0.17	29,52,77,81	0
2	NAG	B	2	14/15	0.86	0.32	44,53,62,63	0
2	NAG	B	1	14/15	0.90	0.16	27,37,45,45	0
3	NAG	D	1	14/15	0.93	0.08	19,24,40,48	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.

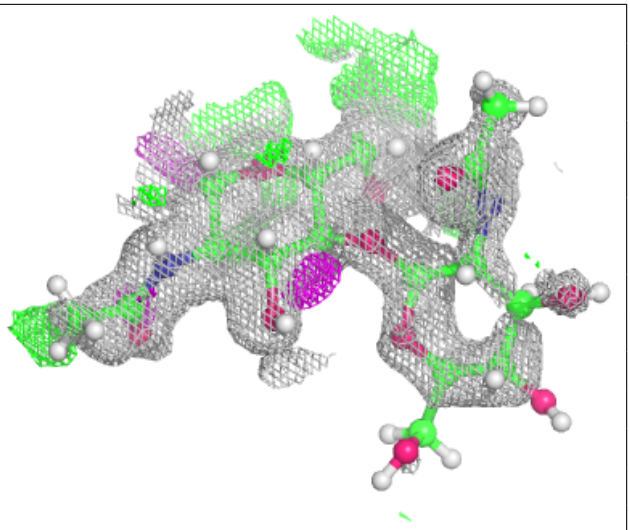
Electron density around Chain B:

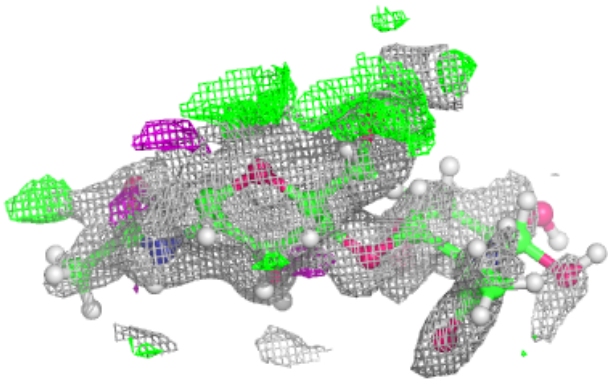
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

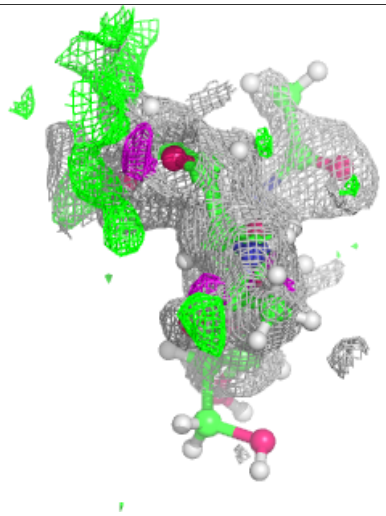


Electron density around Chain C:

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

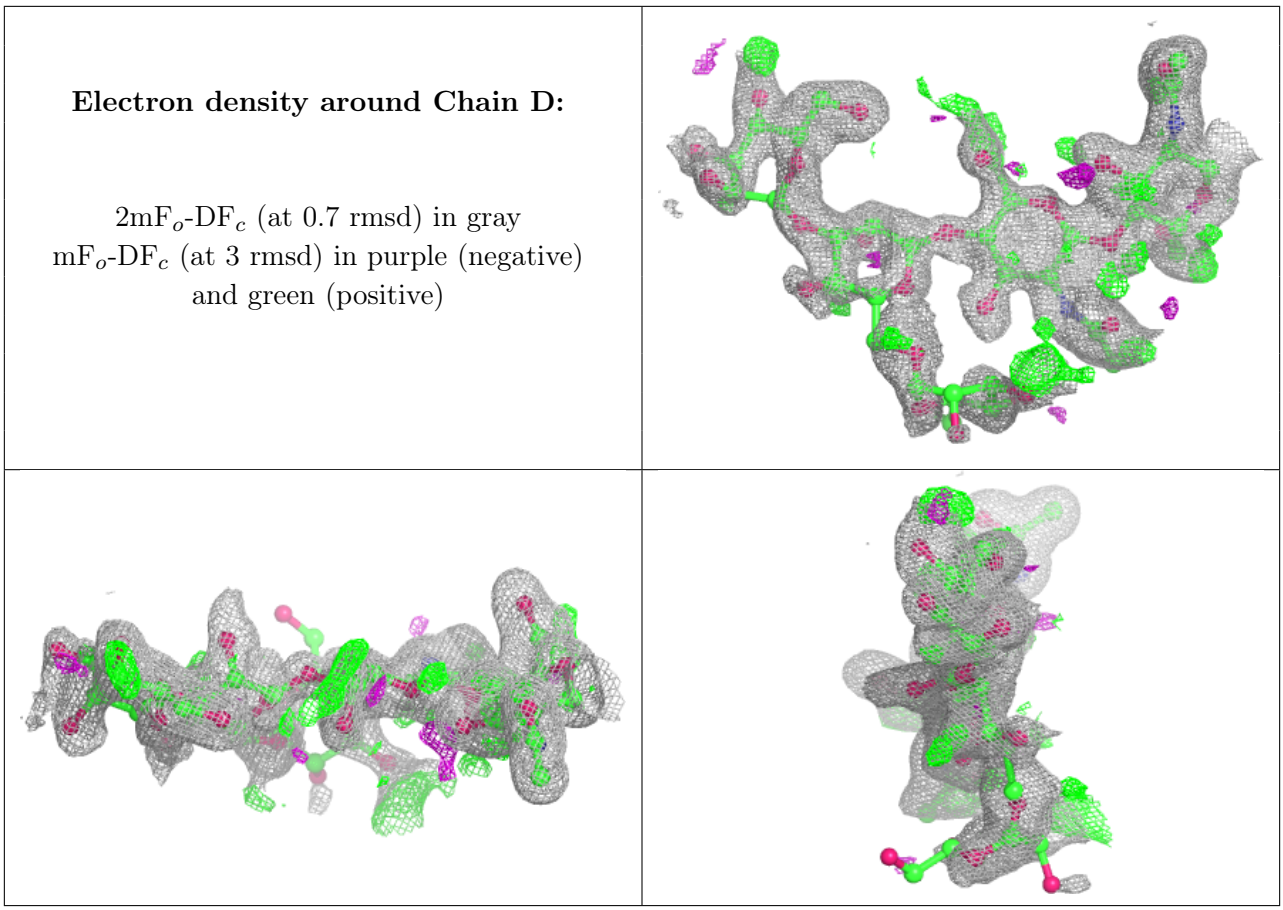






Electron density around Chain D:

$2mF_o$-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^{th} 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^2)	Q<0.9
4	NAG	A	806	14/15	0.56	0.41	47,55,65,66	0
4	NAG	A	801	14/15	0.71	0.32	44,49,58,59	0
9	TRS	A	824	8/8	0.76	0.20	39,53,63,63	0
8	PEG	A	821	7/7	0.79	0.20	37,49,60,61	0
8	PEG	A	823	7/7	0.81	0.17	48,49,49,49	0
8	PEG	A	820	7/7	0.81	0.16	38,39,51,56	0
8	PEG	A	819	7/7	0.87	0.15	38,39,44,45	0
8	PEG	A	822	7/7	0.91	0.10	34,35,39,41	0
7	CL	A	815	1/1	0.98	0.07	38,38,38,38	0
5	HEM	A	812	43/43	0.99	0.10	10,12,17,20	0
7	CL	A	816	1/1	0.99	0.05	25,25,25,25	1
7	CL	A	817	1/1	1.00	0.05	22,22,22,22	1

Continued on next page...

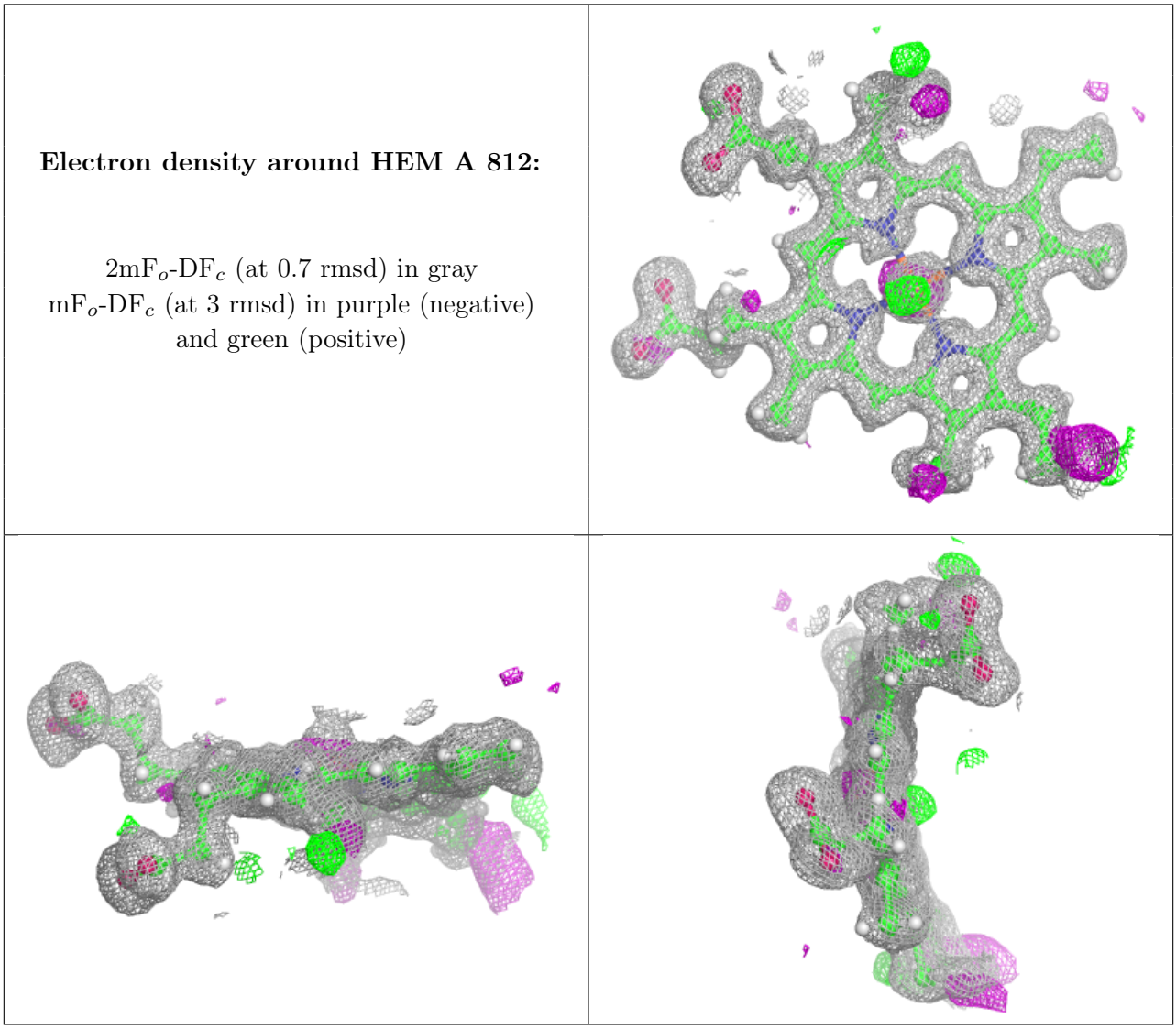
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	CL	A	818	1/1	1.00	0.04	20,20,20,20	0
6	CA	A	813	1/1	1.00	0.09	9,9,9,9	0
7	CL	A	814	1/1	1.00	0.09	11,11,11,11	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 812:

$2mF_o$-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.