Full wwPDB X-ray Structure Validation Report (i)

Oct 18, 2023 – 04:06 AM EDT

PDB ID	:	2GJM
Title	:	Crystal structure of Buffalo lactoperoxidase at 2.75A resolution
Authors	:	Sheikh, I.A.; Ethayathulla, A.S.; Singh, A.K.; Singh, N.; Sharma, S.; Singh, T.P.
Deposited on	:	2006-03-31
Resolution	:	2.75 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:

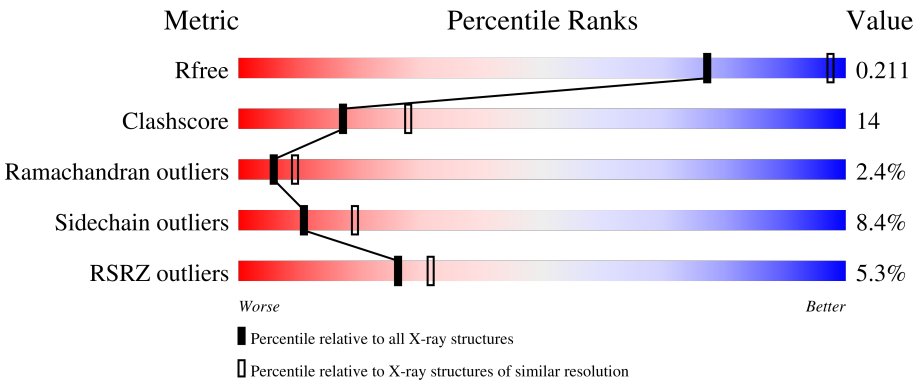
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.75 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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	583	5%, 65%, 30%, • •
2	B	2	100%
2	C	2	50%, 50%
3	D	3	33%, 33%, 33%
4	E	4	75%, 25%

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
4	NAG	E	2	-	-	-	X
4	BMA	E	3	-	-	-	X
4	MAN	E	4	-	-	-	X
8	IOD	A	2006	-	-	X	-

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 5088 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.

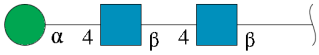
- Molecule 1 is a protein called Lactoperoxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	583	4701	2994	835	847	25	0	1	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	N	O			
2	B	2	28	16	2	10	0	0	0
2	C	2	28	16	2	10	0	0	0

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

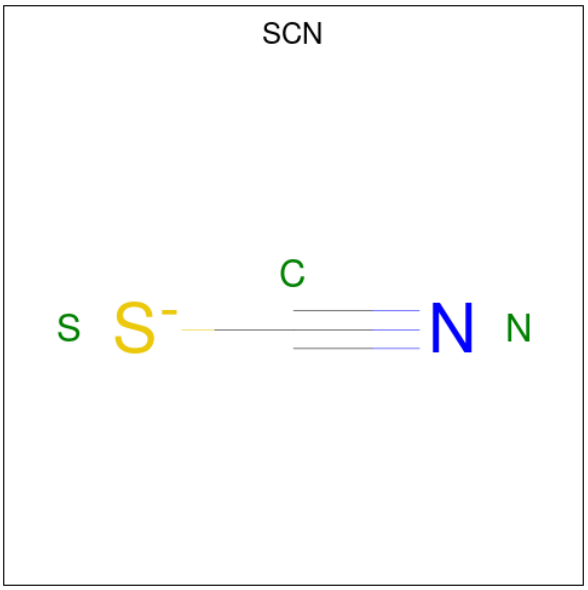


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	D	3	39	22	2	15	0	0	0

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

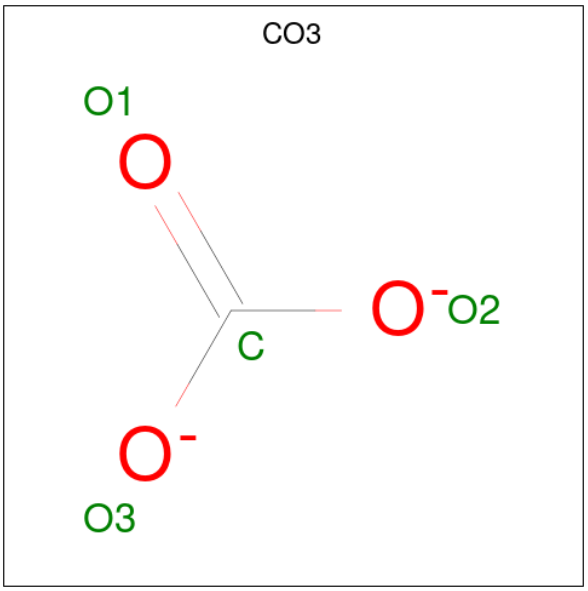
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	E	4	50	28	2	20	0	0	0

- Molecule 5 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	S		
5	A	1	3	1	1	1	0	0

- Molecule 6 is CARBONATE ION (three-letter code: CO3) (formula: CO_3).



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

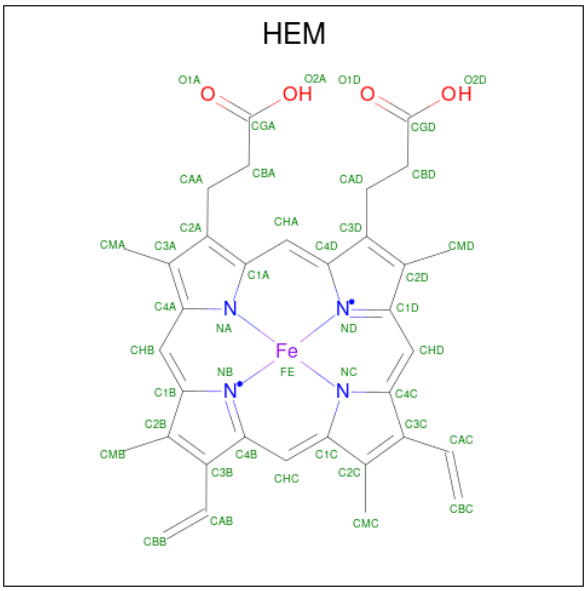
- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 8 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	8	Total I 8 8	0	0

- Molecule 9 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Fe	N	O		
9	A	1	43	34	1	4	4	0	0

- Molecule 10 is water.

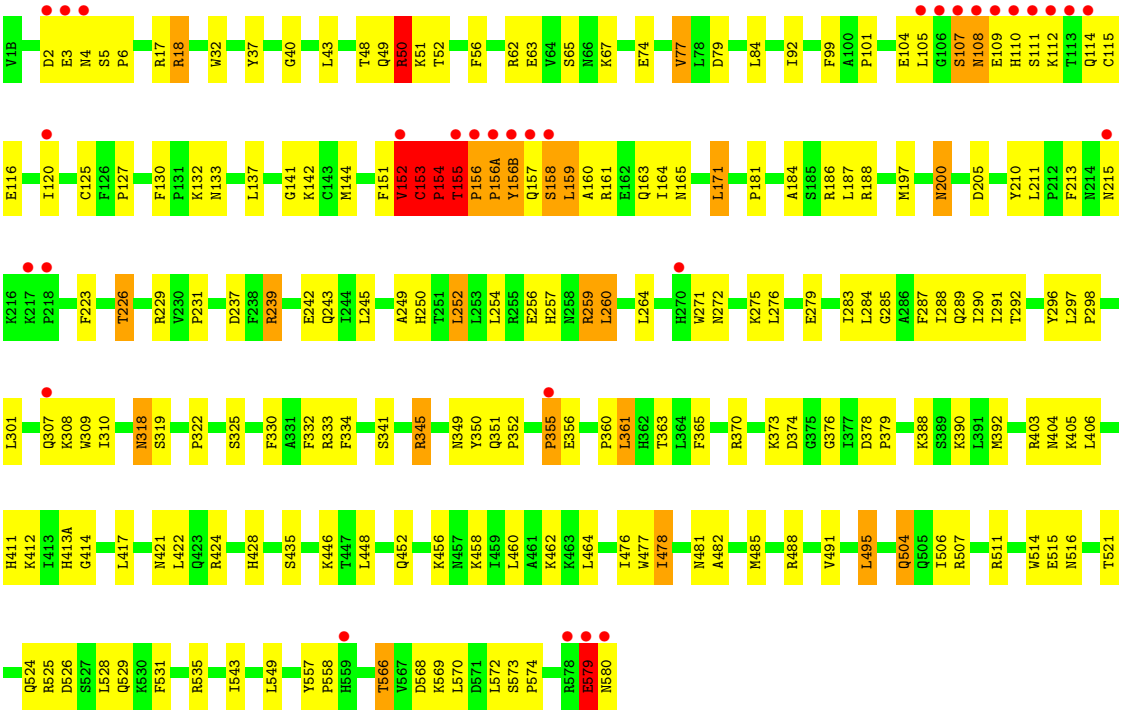
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
10	A	183	183	183	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

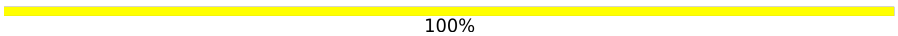
• Molecule 1: Lactoperoxidase

Chain A:



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

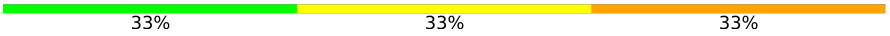
Chain B:



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

Chain D: 

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

Chain E:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	54.24Å 80.47Å 77.35Å 90.00° 102.71° 90.00°	Depositor
Resolution (Å)	20.00 – 2.75 25.27 – 2.73	Depositor EDS
% Data completeness (in resolution range)	96.0 (20.00-2.75) 95.4 (25.27-2.73)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	1.51 (at 2.72Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.182 , 0.216 0.178 , 0.211	Depositor DCC
R_{free} test set	798 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å^2)	28.8	Xtriage
Anisotropy	0.711	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 63.6	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	0.93	EDS
Total number of atoms	5088	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.77% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: HEM, CO3, IOD, NAG, SCN, MAN, 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	0.40	0/4827	0.72	5/6544 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	153	CYS	C-N-CD	-8.60	101.68	120.60
1	A	355	PRO	CA-N-CD	-7.25	101.35	111.50
1	A	154	PRO	CA-N-CD	-7.25	101.35	111.50
1	A	154	PRO	CA-C-N	-6.28	103.39	117.20
1	A	158	SER	N-CA-C	5.87	126.85	111.00

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	4701	0	4618	131	0
2	B	28	0	25	0	0
2	C	28	0	25	0	0
3	D	39	0	34	1	0
4	E	50	0	43	1	0
5	A	3	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	4	0	0	0	0
7	A	1	0	0	0	0
8	A	8	0	0	5	0
9	A	43	0	30	2	0
10	A	183	0	0	10	0
All	All	5088	0	4775	135	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 14.

All (135) 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:211:LEU:HD23	1:A:254:LEU:HD22	1.44	1.00
1:A:109:GLU:HB3	1:A:112:LYS:HE3	1.48	0.94
1:A:153:CYS:HB2	1:A:154:PRO:HD2	1.47	0.93
1:A:153:CYS:CB	1:A:154:PRO:HD2	1.99	0.93
8:A:2005:IOD:I	10:A:2191:HOH:O	2.61	0.88
1:A:50:ARG:HG2	1:A:51:LYS:N	1.89	0.86
1:A:114:GLN:HG3	10:A:2161:HOH:O	1.77	0.83
1:A:17:ARG:HD2	10:A:2170:HOH:O	1.77	0.81
1:A:120:ILE:HB	10:A:2161:HOH:O	1.82	0.80
1:A:153:CYS:CB	1:A:154:PRO:CD	2.62	0.77
1:A:333:ARG:HH11	1:A:421:ASN:ND2	1.85	0.74
1:A:49:GLN:O	1:A:50:ARG:HB3	1.89	0.71
1:A:109:GLU:HG3	1:A:111:SER:H	1.55	0.71
1:A:573:SER:OG	1:A:574:PRO:HD3	1.90	0.71
1:A:452:GLN:NE2	1:A:458:LYS:HG2	2.06	0.70
1:A:154:PRO:HG2	1:A:156(B):TYR:CD1	2.26	0.69
1:A:197:MET:HG2	1:A:257:HIS:CD2	2.28	0.69
1:A:79:ASP:O	1:A:388:LYS:HD3	1.95	0.67
1:A:99:PHE:O	1:A:101:PRO:HD3	1.93	0.67
1:A:521:THR:OG1	1:A:524:GLN:HG3	1.95	0.66
1:A:171:LEU:HD21	1:A:288:ILE:HG22	1.77	0.66
1:A:155:THR:O	1:A:156:PRO:O	2.13	0.66
1:A:2:ASP:HB3	10:A:2162:HOH:O	1.95	0.65
1:A:77:VAL:HG13	1:A:390:LYS:HG3	1.79	0.65
1:A:514:TRP:CE2	1:A:515:GLU:HG3	2.32	0.65
1:A:355:PRO:HB2	1:A:356:GLU:HG3	1.78	0.64
1:A:237:ASP:OD2	1:A:239:ARG:HD3	1.98	0.64
1:A:159:LEU:HD22	1:A:160:ALA:H	1.63	0.63

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:318:ASN:HD22	1:A:318:ASN:C	2.02	0.63
1:A:460:LEU:HD21	1:A:482:ALA:HB1	1.83	0.61
1:A:406:LEU:HD22	1:A:417:LEU:HB2	1.82	0.60
1:A:318:ASN:HD22	1:A:319:SER:N	2.00	0.60
1:A:285:GLY:O	1:A:289:GLN:HG3	2.02	0.59
1:A:334:PHE:CB	1:A:481:ASN:HD21	2.17	0.57
1:A:256:GLU:OE2	1:A:259:ARG:NH1	2.38	0.57
3:D:2:NAG:H61	3:D:3:MAN:H2	1.86	0.57
1:A:107:SER:O	1:A:108:ASN:CB	2.54	0.55
1:A:525:ARG:O	1:A:529:GLN:HG3	2.06	0.55
1:A:171:LEU:HD21	1:A:288:ILE:CG2	2.35	0.55
1:A:272:ASN:O	1:A:276:LEU:HD23	2.06	0.55
1:A:152:VAL:O	1:A:153:CYS:C	2.45	0.55
1:A:290:ILE:HD12	1:A:531:PHE:HD1	1.72	0.55
1:A:213:PHE:CD1	1:A:231:PRO:HG2	2.42	0.54
1:A:242:GLU:OE1	1:A:243:GLN:HG2	2.07	0.54
1:A:116:GLU:CD	1:A:411:HIS:HD1	2.10	0.54
1:A:349:ASN:O	1:A:350:TYR:HB2	2.06	0.54
1:A:205:ASP:HB2	1:A:210:TYR:CZ	2.43	0.54
1:A:242:GLU:OE1	1:A:243:GLN:CG	2.56	0.54
1:A:132:LYS:O	1:A:133:ASN:HB2	2.08	0.53
1:A:92:ILE:HD11	1:A:249:ALA:HB1	1.91	0.53
1:A:153:CYS:HB3	1:A:154:PRO:HD2	1.89	0.53
1:A:531:PHE:CE2	1:A:570:LEU:HD22	2.44	0.53
1:A:153:CYS:HB3	1:A:154:PRO:CD	2.38	0.52
1:A:287:PHE:O	1:A:291:ILE:HG12	2.10	0.52
1:A:18:ARG:CB	1:A:18:ARG:HH11	2.22	0.52
1:A:200:ASN:ND2	8:A:2004:IOD:I	3.13	0.51
1:A:2:ASP:O	1:A:4:ASN:N	2.37	0.51
1:A:226:THR:O	1:A:229:ARG:HG3	2.11	0.51
1:A:164:ILE:HG22	1:A:165:ASN:N	2.25	0.51
1:A:40:GLY:HA2	10:A:2132:HOH:O	2.10	0.51
1:A:163:GLN:HG2	1:A:428:HIS:CE1	2.46	0.51
1:A:535:ARG:HD3	1:A:568:ASP:O	2.11	0.50
1:A:184:ALA:O	1:A:188:ARG:HG3	2.11	0.49
1:A:271:TRP:CG	1:A:275:LYS:HG2	2.48	0.49
1:A:378:ASP:N	1:A:379:PRO:CD	2.76	0.49
1:A:528:LEU:HD22	1:A:531:PHE:CZ	2.47	0.48
1:A:322:PRO:O	1:A:506:ILE:HG22	2.13	0.48
9:A:605:HEM:HMB1	9:A:605:HEM:HBB2	1.96	0.48
1:A:109:GLU:CB	1:A:112:LYS:HE3	2.32	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:360:PRO:HB2	1:A:363:THR:HG23	1.95	0.47
1:A:63:GLU:OE2	1:A:67:LYS:NZ	2.44	0.47
1:A:531:PHE:HE2	1:A:570:LEU:HD22	1.79	0.47
1:A:181:PRO:HD2	8:A:2002:IOD:I	2.84	0.47
1:A:404:ASN:O	1:A:414:GLY:HA2	2.15	0.47
1:A:290:ILE:HD12	1:A:531:PHE:CD1	2.50	0.46
1:A:74:GLU:HG2	1:A:341:SER:OG	2.15	0.46
1:A:260:LEU:O	1:A:264:LEU:HG	2.16	0.46
1:A:48:THR:CG2	1:A:50:ARG:HD3	2.45	0.46
1:A:528:LEU:HD22	1:A:531:PHE:CE1	2.51	0.46
1:A:535:ARG:HD3	1:A:569:LYS:HA	1.97	0.46
1:A:566:THR:O	1:A:566:THR:CG2	2.64	0.46
1:A:107:SER:O	1:A:108:ASN:HB3	2.15	0.46
1:A:155:THR:C	1:A:156:PRO:O	2.54	0.46
1:A:370:ARG:O	1:A:374:ASP:HB3	2.16	0.46
1:A:405:LYS:HA	1:A:413(A):HIS:O	2.15	0.45
1:A:491:VAL:HB	1:A:495:LEU:HB3	1.98	0.45
1:A:215:ASN:N	10:A:2140:HOH:O	2.43	0.45
1:A:308:LYS:HE3	1:A:309:TRP:CZ2	2.50	0.45
1:A:210:TYR:OH	1:A:376:GLY:HA2	2.17	0.45
1:A:310:ILE:HG12	1:A:504:GLN:HB2	1.99	0.45
1:A:543:ILE:HG23	8:A:2006:IOD:I	2.87	0.45
1:A:334:PHE:HA	1:A:481:ASN:HD21	1.82	0.45
1:A:345:ARG:NH2	1:A:374:ASP:OD2	2.49	0.44
1:A:112:LYS:HB2	1:A:112:LYS:NZ	2.32	0.44
1:A:524:GLN:O	1:A:528:LEU:HG	2.17	0.44
1:A:422:LEU:HD21	1:A:478:ILE:HB	1.98	0.44
1:A:32:TRP:CE2	1:A:325:SER:HB3	2.52	0.44
1:A:215:ASN:HB3	10:A:2140:HOH:O	2.17	0.44
1:A:223:PHE:CZ	1:A:412:LYS:HB3	2.53	0.44
1:A:125:CYS:SG	1:A:127:PRO:HG3	2.58	0.43
1:A:279:GLU:O	1:A:283:ILE:HG13	2.18	0.43
1:A:105:LEU:N	1:A:105:LEU:HD12	2.33	0.43
1:A:5:SER:HA	1:A:6:PRO:HD3	1.87	0.43
1:A:37:TYR:CZ	1:A:161:ARG:HB3	2.53	0.43
1:A:297:LEU:HB2	1:A:298:PRO:HD3	2.00	0.43
1:A:334:PHE:HB2	1:A:481:ASN:HD21	1.82	0.43
1:A:462:LYS:HB2	1:A:462:LYS:HE2	1.82	0.42
1:A:62:ARG:O	1:A:65:SER:HB3	2.19	0.42
1:A:92:ILE:HD11	1:A:249:ALA:CB	2.49	0.42
1:A:130:PHE:HE2	1:A:144:MET:CE	2.33	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:579:GLU:H	1:A:579:GLU:HG2	1.46	0.42
1:A:260:LEU:HD12	1:A:572:LEU:HD11	2.01	0.42
1:A:292:THR:HA	1:A:296:TYR:HB3	2.00	0.42
4:E:2:NAG:H62	4:E:3:BMA:C1	2.50	0.42
1:A:333:ARG:NH1	1:A:421:ASN:ND2	2.61	0.42
1:A:120:ILE:CB	10:A:2161:HOH:O	2.57	0.42
1:A:392:MET:HB3	1:A:485:MET:CE	2.49	0.42
1:A:226:THR:O	1:A:229:ARG:NH2	2.52	0.41
1:A:330:PHE:HE2	1:A:424:ARG:HB3	1.85	0.41
1:A:52:THR:HB	1:A:56:PHE:N	2.36	0.41
1:A:351:GLN:O	1:A:352:PRO:C	2.58	0.41
1:A:543:ILE:HA	8:A:2006:IOD:I	2.91	0.41
1:A:137:LEU:HA	1:A:141:GLY:O	2.19	0.41
1:A:211:LEU:HD11	1:A:250:HIS:HB3	2.03	0.41
1:A:151:PHE:CD1	1:A:151:PHE:N	2.88	0.41
1:A:476:ILE:HG23	1:A:477:TRP:N	2.36	0.41
1:A:579:GLU:HB2	1:A:580:ASN:H	1.45	0.41
1:A:115:CYS:HB2	10:A:2015:HOH:O	2.21	0.41
1:A:452:GLN:O	1:A:456:LYS:N	2.54	0.41
1:A:284:LEU:O	1:A:287:PHE:HB3	2.21	0.40
1:A:333:ARG:NH1	9:A:605:HEM:HBA2	2.36	0.40
1:A:361:LEU:HD13	1:A:365:PHE:CZ	2.56	0.40
1:A:511:ARG:O	1:A:516:ASN:ND2	2.48	0.40
1:A:252:LEU:HD12	1:A:252:LEU:HA	1.94	0.40
1:A:557:TYR:CD1	1:A:558:PRO:HA	2.57	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	582/583 (100%)	528 (91%)	40 (7%)	14 (2%)	6 9

All (14) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	GLU
1	A	107	SER
1	A	108	ASN
1	A	152	VAL
1	A	153	CYS
1	A	155	THR
1	A	156	PRO
1	A	158	SER
1	A	579	GLU
1	A	50	ARG
1	A	154	PRO
1	A	156(B)	TYR
1	A	157	GLN
1	A	156(A)	PRO

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	510/509 (100%)	467 (92%)	43 (8%)	11 19

All (43) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	ARG
1	A	43	LEU
1	A	50	ARG
1	A	77	VAL
1	A	84	LEU
1	A	104	GLU
1	A	110	HIS
1	A	142	LYS
1	A	152	VAL
1	A	155	THR
1	A	156(A)	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	159	LEU
1	A	171	LEU
1	A	186	ARG
1	A	187	LEU
1	A	200	ASN
1	A	226	THR
1	A	239	ARG
1	A	245	LEU
1	A	252	LEU
1	A	259	ARG
1	A	260	LEU
1	A	301	LEU
1	A	307	GLN
1	A	318	ASN
1	A	332	PHE
1	A	345	ARG
1	A	361	LEU
1	A	373	LYS
1	A	403	ARG
1	A	435	SER
1	A	446	LYS
1	A	448	LEU
1	A	464	LEU
1	A	478	ILE
1	A	488	ARG
1	A	495	LEU
1	A	504	GLN
1	A	507	ARG
1	A	526	ASP
1	A	549	LEU
1	A	566	THR
1	A	579	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	133	ASN
1	A	200	ASN
1	A	307	GLN
1	A	314	GLN
1	A	318	ASN
1	A	349	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	421	ASN
1	A	452	GLN
1	A	481	ASN
1	A	504	GLN
1	A	529	GLN
1	A	553	GLN
1	A	555	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

11 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	1,2	14,14,15	0.50	0	17,19,21	1.53	1 (5%)
2	NAG	B	2	2	14,14,15	0.57	0	17,19,21	1.22	1 (5%)
2	NAG	C	1	1,2	14,14,15	0.50	0	17,19,21	0.93	0
2	NAG	C	2	2	14,14,15	0.64	0	17,19,21	1.01	1 (5%)
3	NAG	D	1	3,1	14,14,15	0.47	0	17,19,21	0.77	0
3	NAG	D	2	3	14,14,15	0.65	0	17,19,21	0.87	1 (5%)
3	MAN	D	3	3	11,11,12	0.56	0	15,15,17	0.38	0
4	NAG	E	1	4,1	14,14,15	0.75	0	17,19,21	1.38	1 (5%)
4	NAG	E	2	4	14,14,15	0.71	0	17,19,21	0.82	0
4	BMA	E	3	4	11,11,12	1.00	1 (9%)	15,15,17	2.05	4 (26%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	MAN	E	4	4	11,11,12	1.04	0	15,15,17	1.10	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	1,2	-	2/6/23/26	0/1/1/1
2	NAG	B	2	2	-	4/6/23/26	0/1/1/1
2	NAG	C	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	C	2	2	-	2/6/23/26	0/1/1/1
3	NAG	D	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	D	2	3	-	4/6/23/26	0/1/1/1
3	MAN	D	3	3	-	0/2/19/22	0/1/1/1
4	NAG	E	1	4,1	-	2/6/23/26	0/1/1/1
4	NAG	E	2	4	-	2/6/23/26	0/1/1/1
4	BMA	E	3	4	-	0/2/19/22	0/1/1/1
4	MAN	E	4	4	-	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)
4	E	3	BMA	O5-C5	2.25	1.48	1.43

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	E	3	BMA	C1-O5-C5	5.93	120.22	112.19
2	B	1	NAG	C2-N2-C7	-5.03	115.74	122.90
4	E	1	NAG	C4-C3-C2	4.42	117.49	111.02
2	B	2	NAG	C2-N2-C7	-4.27	116.83	122.90
4	E	3	BMA	O5-C1-C2	3.45	116.10	110.77
2	C	2	NAG	C4-C3-C2	2.75	115.05	111.02
4	E	4	MAN	C1-C2-C3	-2.69	106.36	109.67
4	E	3	BMA	C1-C2-C3	2.26	112.45	109.67
4	E	4	MAN	C1-O5-C5	-2.24	109.16	112.19
4	E	3	BMA	O5-C5-C4	2.10	115.93	110.83
3	D	2	NAG	C2-N2-C7	-2.03	120.01	122.90

There are no chirality outliers.

All (18) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	2	NAG	O7-C7-N2-C2
4	E	2	NAG	C8-C7-N2-C2
4	E	2	NAG	O7-C7-N2-C2
3	D	2	NAG	C8-C7-N2-C2
2	B	2	NAG	O5-C5-C6-O6
2	B	2	NAG	C8-C7-N2-C2
3	D	2	NAG	O5-C5-C6-O6
2	B	2	NAG	C4-C5-C6-O6
2	C	1	NAG	C4-C5-C6-O6
2	C	1	NAG	O5-C5-C6-O6
3	D	2	NAG	C4-C5-C6-O6
2	B	2	NAG	O7-C7-N2-C2
4	E	1	NAG	C8-C7-N2-C2
4	E	1	NAG	O7-C7-N2-C2
2	B	1	NAG	O5-C5-C6-O6
2	B	1	NAG	C4-C5-C6-O6
2	C	2	NAG	C4-C5-C6-O6
2	C	2	NAG	O5-C5-C6-O6

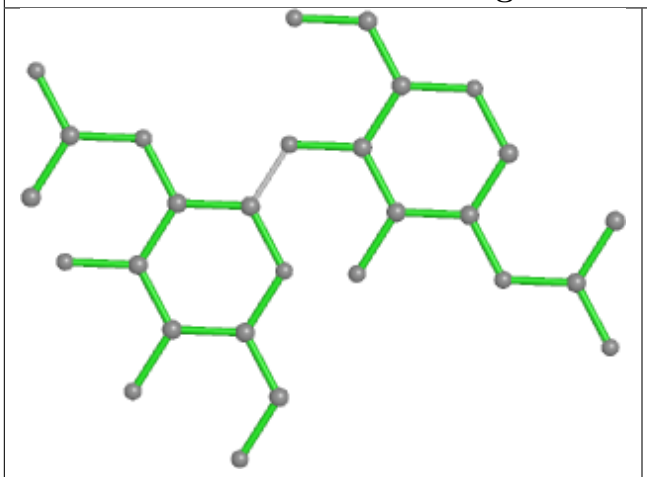
There are no ring outliers.

4 monomers are involved in 2 short contacts:

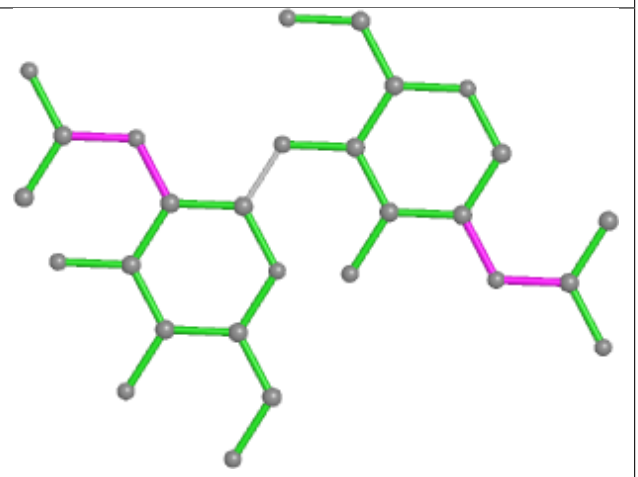
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	2	NAG	1	0
4	E	3	BMA	1	0
3	D	2	NAG	1	0
3	D	3	MAN	1	0

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

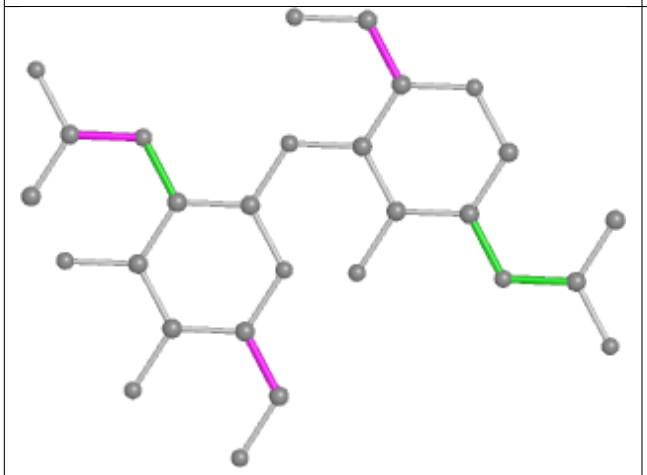
Oligosaccharide Chain B



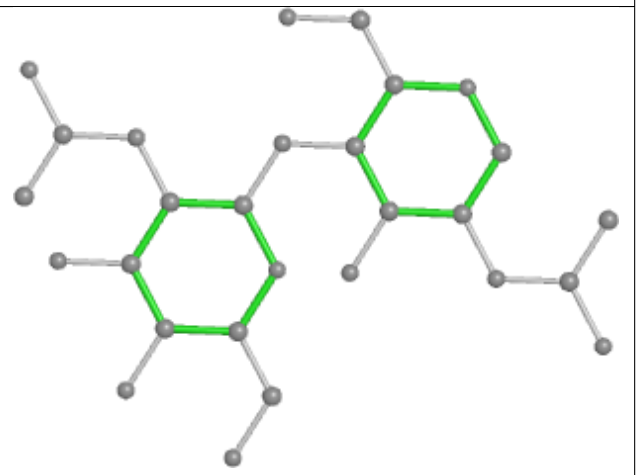
Bond lengths



Bond angles

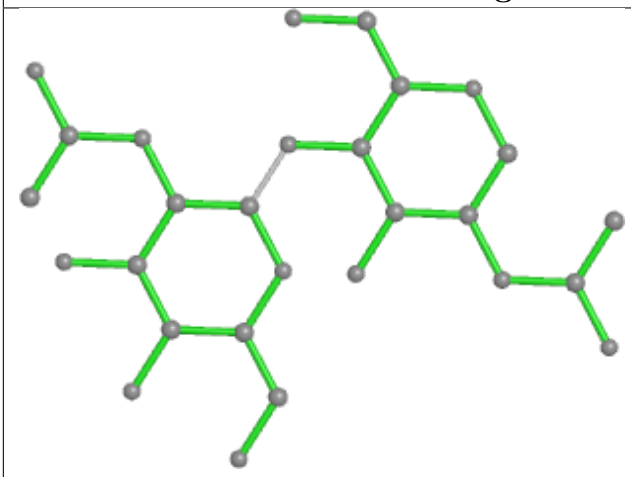


Torsions

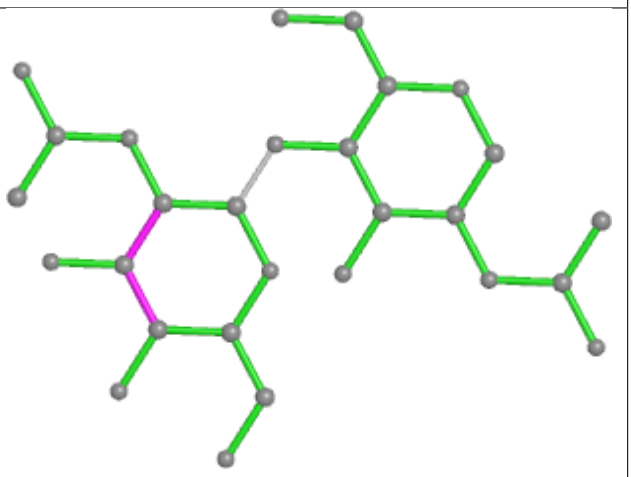


Rings

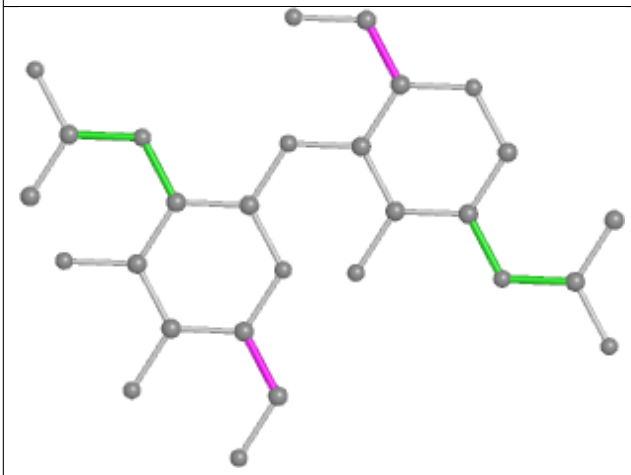
Oligosaccharide Chain C



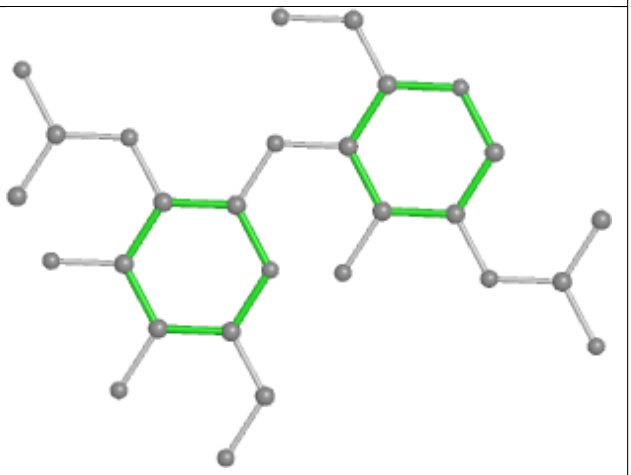
Bond lengths



Bond angles

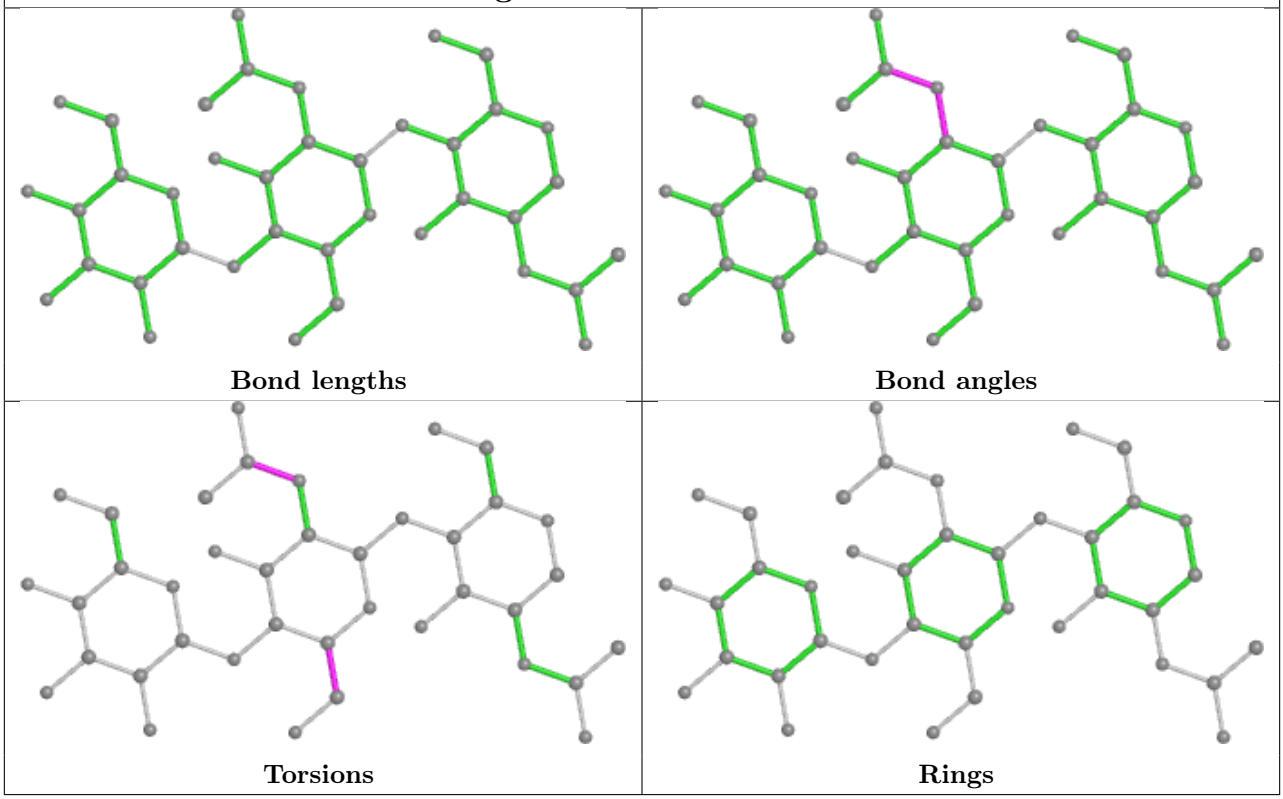


Torsions

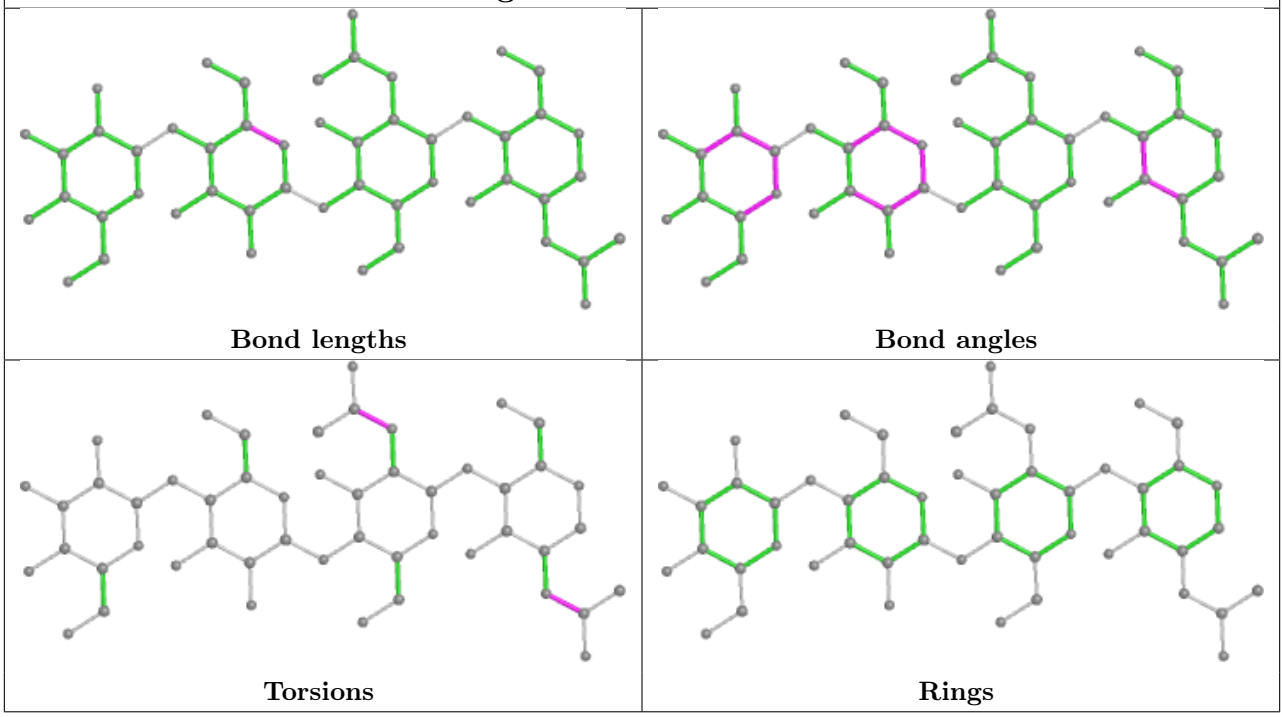


Rings

Oligosaccharide Chain D



Oligosaccharide Chain E



5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 9 are monoatomic - leaving 3 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	SCN	A	1502	-	1,2,2	1.60	0	0,1,1	-	-
6	CO3	A	1001	-	2,3,3	0.71	0	2,3,3	0.37	0
9	HEM	A	605	1	41,50,50	2.06	10 (24%)	45,82,82	2.90	13 (28%)

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	HEM	A	605	1	-	3/12/54/54	-

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	A	605	HEM	C3D-C2D	6.61	1.50	1.36
9	A	605	HEM	FE-ND	5.02	2.21	1.96
9	A	605	HEM	CAB-C3B	4.38	1.59	1.47
9	A	605	HEM	C3C-CAC	3.99	1.56	1.47
9	A	605	HEM	CAA-C2A	3.03	1.56	1.52
9	A	605	HEM	C1D-C2D	2.43	1.49	1.44
9	A	605	HEM	CHB-C1B	2.35	1.41	1.35
9	A	605	HEM	FE-NB	2.16	2.07	1.96
9	A	605	HEM	CMC-C2C	2.15	1.56	1.51
9	A	605	HEM	CHA-C4D	2.05	1.40	1.35

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	A	605	HEM	CAD-CBD-CGD	11.31	137.94	113.60
9	A	605	HEM	CBD-CAD-C3D	6.27	130.05	112.63
9	A	605	HEM	C4D-ND-C1D	6.07	111.34	105.07
9	A	605	HEM	C4B-CHC-C1C	4.61	128.65	122.56
9	A	605	HEM	CMD-C2D-C1D	4.55	131.96	125.04

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	A	605	HEM	C1B-NB-C4B	4.47	109.69	105.07
9	A	605	HEM	C4C-CHD-C1D	4.25	128.17	122.56
9	A	605	HEM	CMC-C2C-C3C	3.83	131.85	124.68
9	A	605	HEM	CBA-CAA-C2A	-3.46	106.72	112.62
9	A	605	HEM	C3B-C2B-C1B	3.44	109.03	106.49
9	A	605	HEM	CAD-C3D-C4D	2.78	129.51	124.66
9	A	605	HEM	C2D-C1D-ND	-2.66	106.69	109.88
9	A	605	HEM	C2B-C1B-NB	-2.57	106.79	109.84

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	A	605	HEM	CAD-CBD-CGD-O2D
9	A	605	HEM	CAD-CBD-CGD-O1D
9	A	605	HEM	CAA-CBA-CGA-O2A

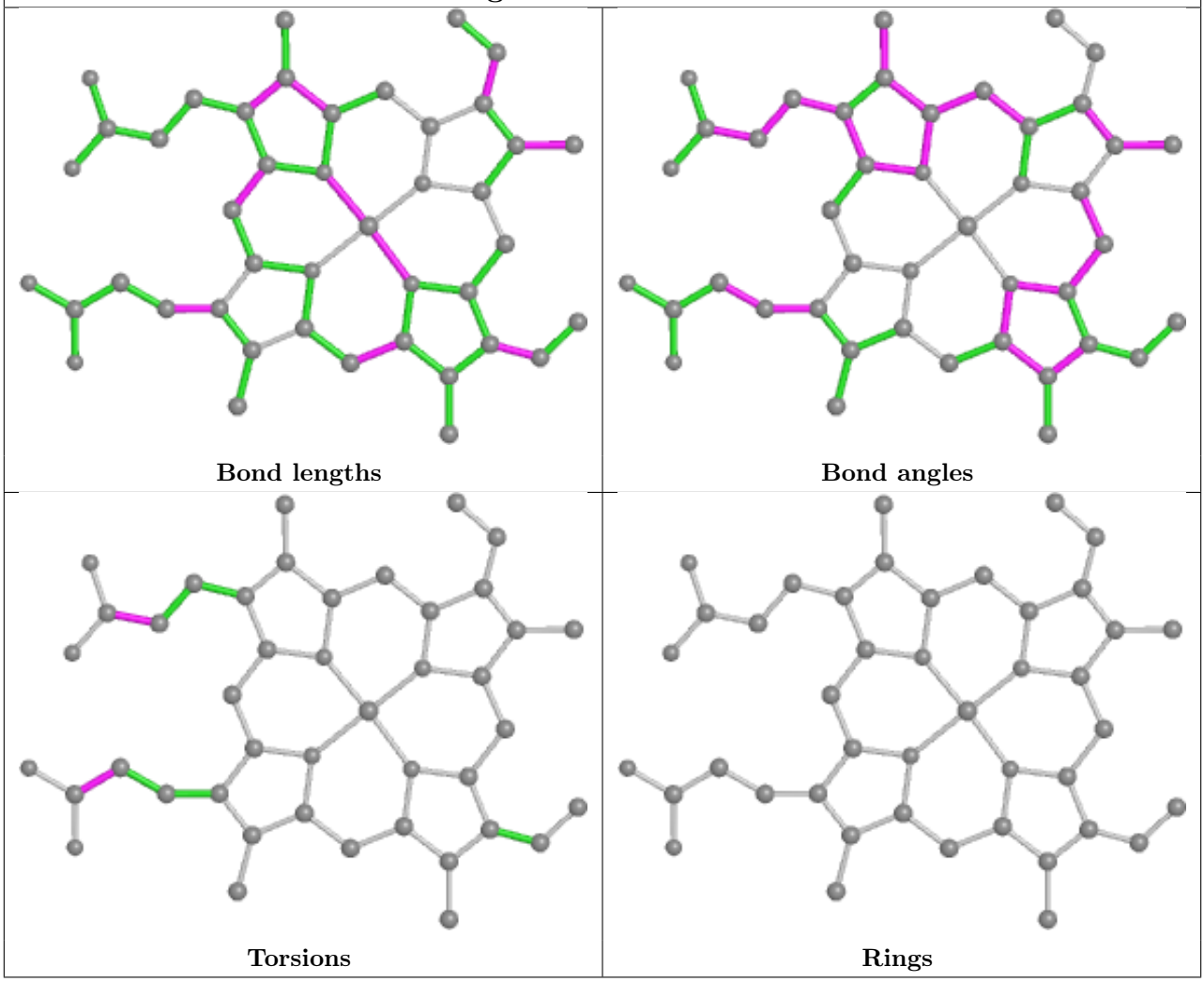
There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	605	HEM	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 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 605



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, 95th 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	583/583 (100%)	-0.15	31 (5%) 26 31	8, 26, 52, 85	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	578	ARG	7.9
1	A	108	ASN	7.8
1	A	109	GLU	6.0
1	A	158	SER	5.9
1	A	157	GLN	5.7
1	A	579	GLU	5.7
1	A	111	SER	5.5
1	A	107	SER	5.3
1	A	156(B)	TYR	5.1
1	A	156	PRO	4.9
1	A	105	LEU	4.8
1	A	106	GLY	4.7
1	A	218	PRO	4.0
1	A	580	ASN	4.0
1	A	110	HIS	3.9
1	A	112	LYS	3.8
1	A	113	THR	3.2
1	A	355	PRO	3.1
1	A	156(A)	PRO	2.8
1	A	2	ASP	2.7
1	A	3	GLU	2.4
1	A	114	GLN	2.3
1	A	559	HIS	2.3
1	A	155	THR	2.2
1	A	270	HIS	2.2
1	A	307	GLN	2.1
1	A	4	ASN	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	120	ILE	2.1
1	A	215	ASN	2.1
1	A	152	VAL	2.1
1	A	217	LYS	2.1

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

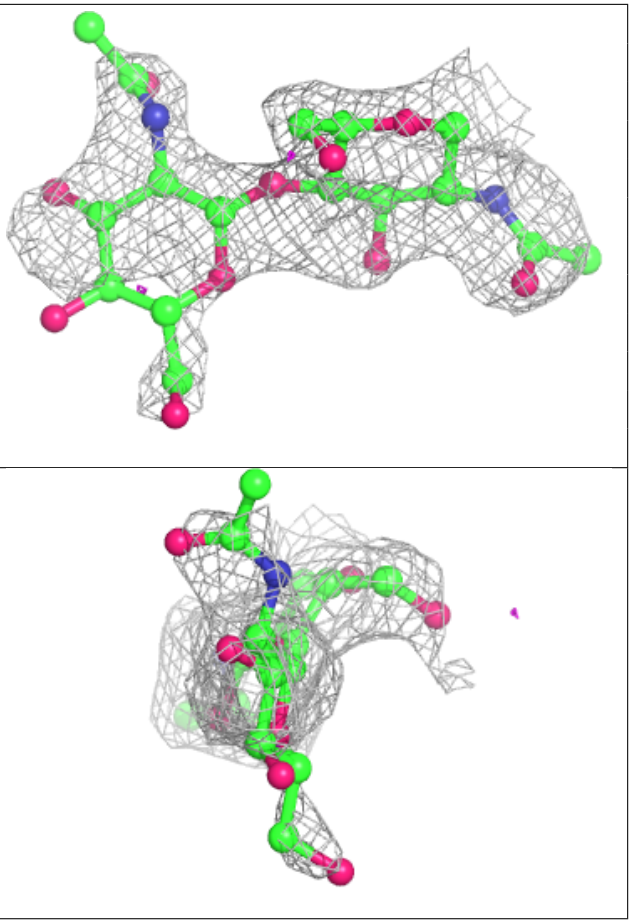
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

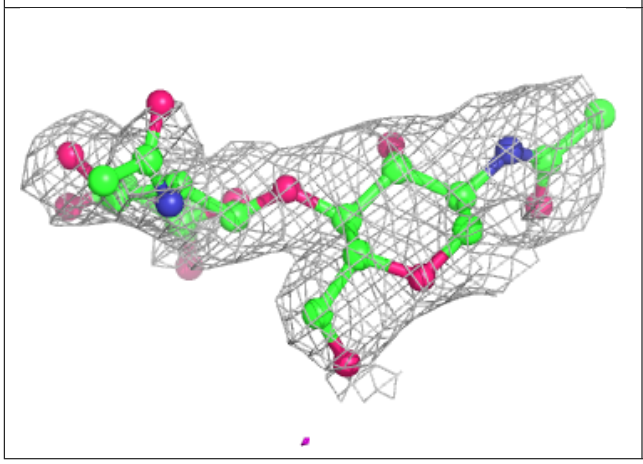
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	MAN	E	4	11/12	0.59	0.55	48,49,49,49	11
3	MAN	D	3	11/12	0.60	0.34	46,47,48,48	11
4	BMA	E	3	11/12	0.66	0.48	49,49,49,49	11
4	NAG	E	2	14/15	0.67	0.43	48,49,50,50	14
2	NAG	C	2	14/15	0.80	0.48	56,57,58,58	0
3	NAG	D	2	14/15	0.82	0.35	47,50,51,51	0
4	NAG	E	1	14/15	0.84	0.27	45,49,51,51	0
2	NAG	B	2	14/15	0.86	0.43	63,66,67,67	0
2	NAG	B	1	14/15	0.87	0.17	49,51,55,60	0
2	NAG	C	1	14/15	0.91	0.20	45,49,51,53	0
3	NAG	D	1	14/15	0.94	0.17	39,40,44,45	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:

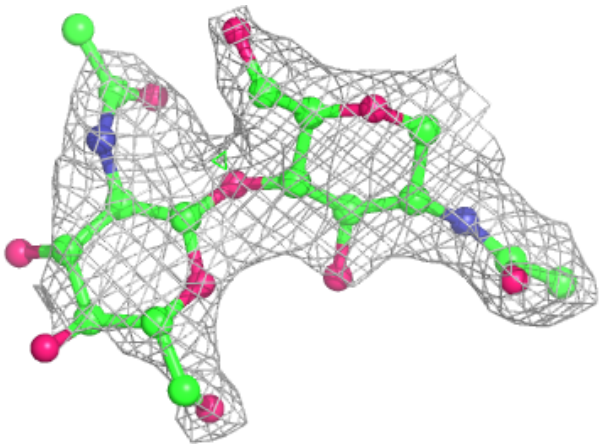
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

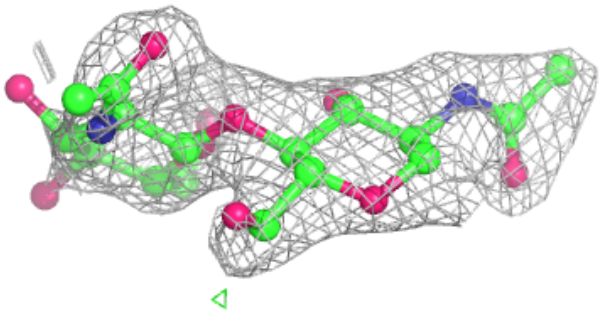




Electron density around Chain C:

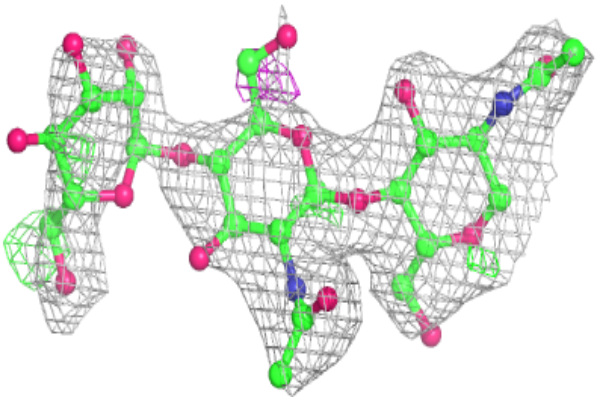
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

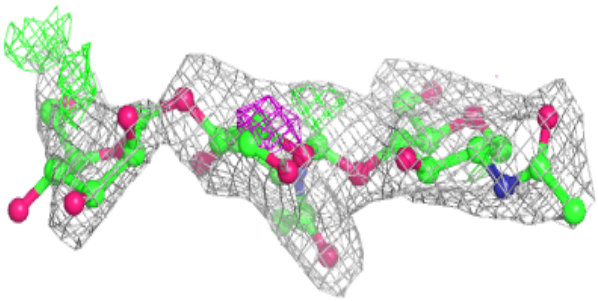


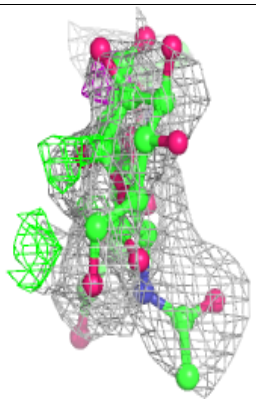


Electron density around Chain D:

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

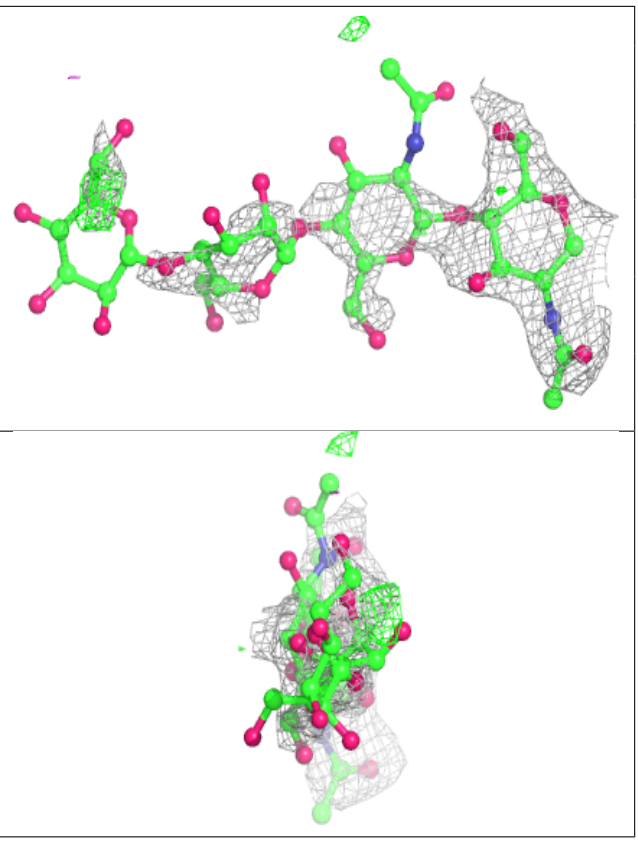






Electron density around Chain E:

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



6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th 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.

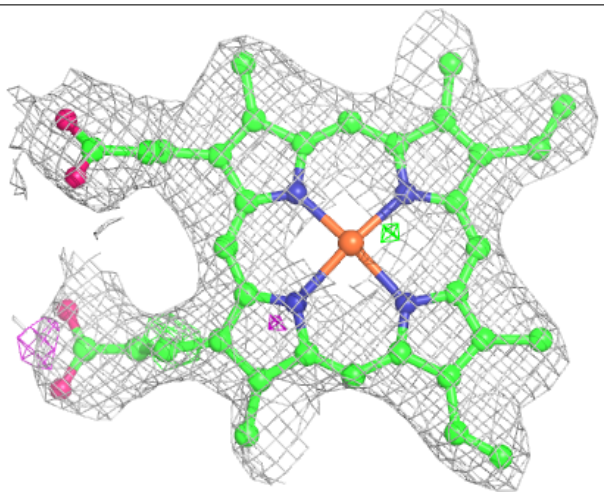
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
8	IOD	A	2005	1/1	0.92	0.10	37,37,37,37	1
8	IOD	A	2008	1/1	0.92	0.12	24,24,24,24	1
8	IOD	A	2006	1/1	0.93	0.12	37,37,37,37	1
6	CO3	A	1001	4/4	0.95	0.19	31,32,33,33	0
5	SCN	A	1502	3/3	0.95	0.12	14,14,14,15	0
9	HEM	A	605	43/43	0.95	0.14	13,19,23,25	0
8	IOD	A	2003	1/1	0.97	0.07	47,47,47,47	1
8	IOD	A	2007	1/1	0.97	0.13	10,10,10,10	1
8	IOD	A	2004	1/1	0.97	0.12	33,33,33,33	1
8	IOD	A	2001	1/1	0.97	0.06	47,47,47,47	1
8	IOD	A	2002	1/1	0.99	0.05	42,42,42,42	1
7	CA	A	1503	1/1	0.99	0.10	14,14,14,14	0

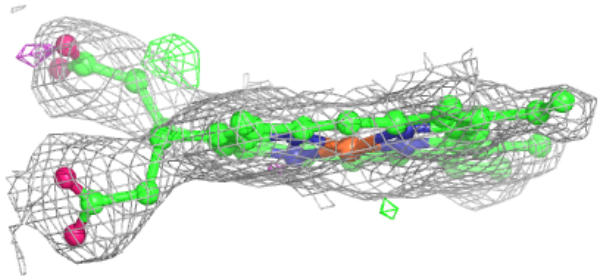
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

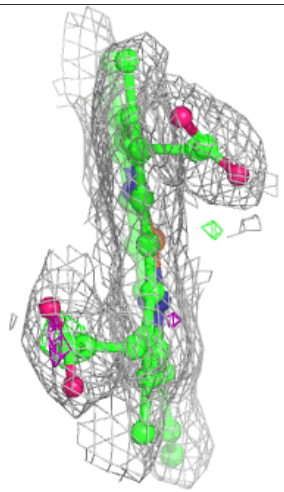
as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around HEM A 605:

$2mF_o$-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.