



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 02:27 PM BST

PDB ID : 5B72  
Title : Crystal structure of bovine lactoperoxidase with a broken covalent bond between Glu258 and heme moiety at 1.98 Å resolution.  
Authors : Singh, P.K.; Sirohi, H.V.; Kaur, P.; Sharma, S.; Singh, T.P.  
Deposited on : 2016-06-03  
Resolution : 1.98 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

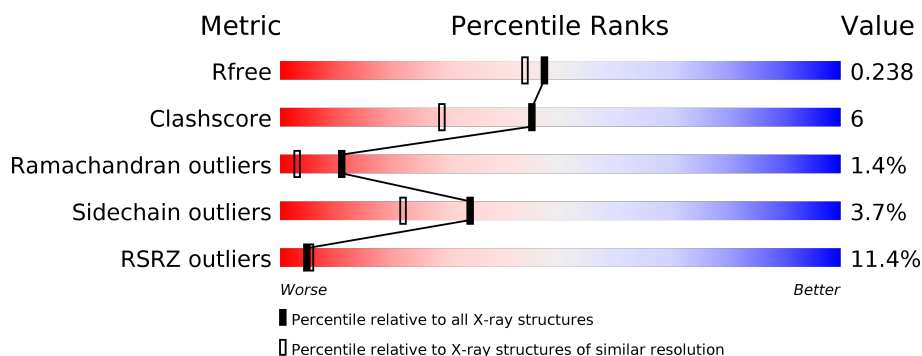
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.98 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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 on 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	595	
2	B	2	

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
1	SEP	A	198	-	X	-	-
2	NAG	B	2	-	-	-	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 5181 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	P	S			
1	A	595	4767	3027	847	865	1	27	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	220	SER	TRP	engineered mutation	UNP P80025
A	254	SER	PHE	engineered mutation	UNP P80025
A	410	LYS	ASP	engineered mutation	UNP P80025
A	547	MET	VAL	engineered mutation	UNP P80025

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



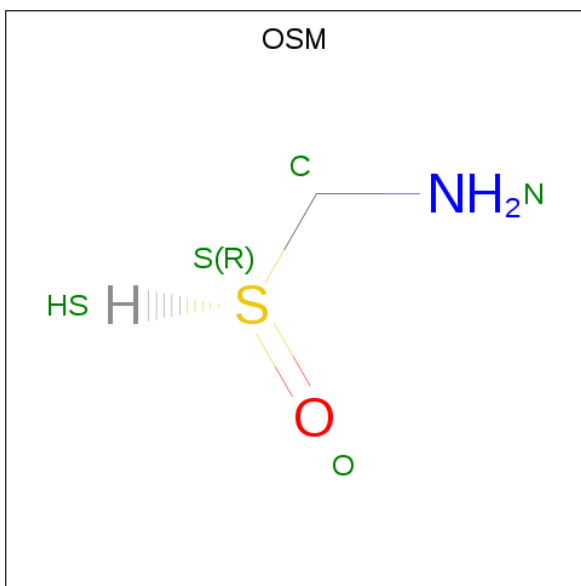
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	2	28	16	2	10	0	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 4 is 1-(OXIDOSULFANYL)METHANAMINE (three-letter code: OSM) (formula:  $\text{CH}_5\text{NOS}$ ).

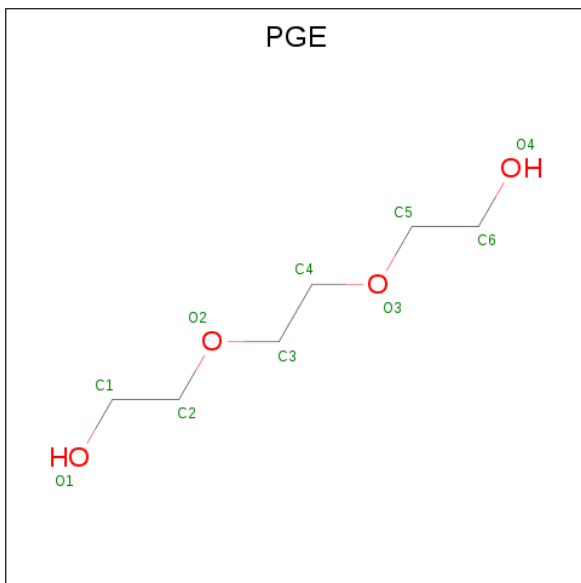


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0
			4	1	1	1	1	0

- Molecule 5 is IODIDE ION (three-letter code: IOD) (formula: I).

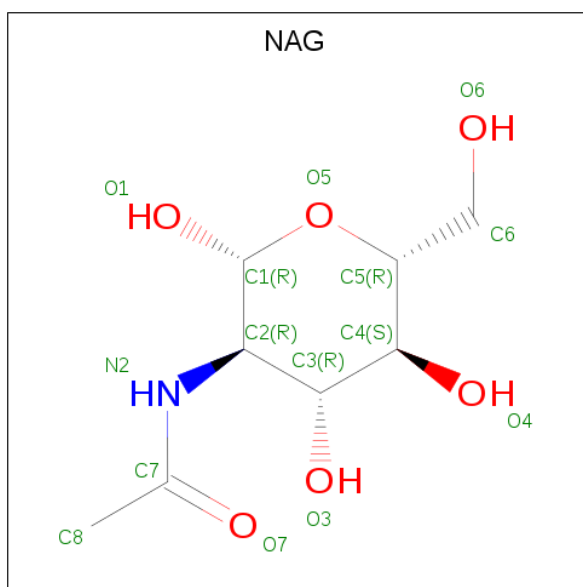
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	9	Total I 9 9	0	0

- Molecule 6 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula:  $C_6H_{14}O_4$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 10 6 4	0	0
6	A	1	Total C O 10 6 4	0	0

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

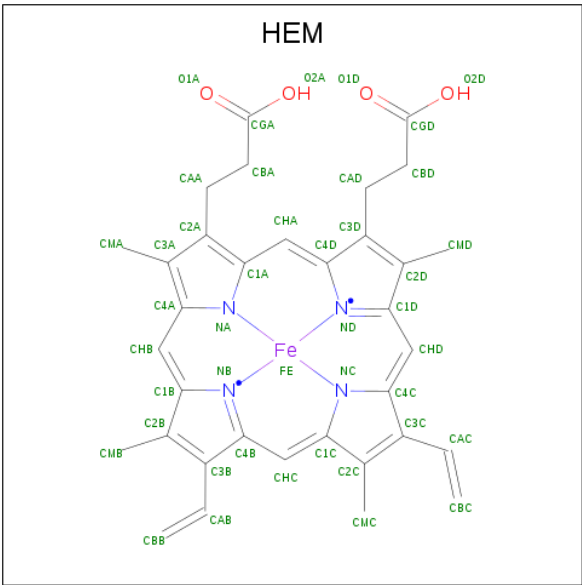


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total	C	N	O	0	0
			14	8	1	5		
7	A	1	Total	C	N	O	0	0
			14	8	1	5		
7	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 8 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- 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
9	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

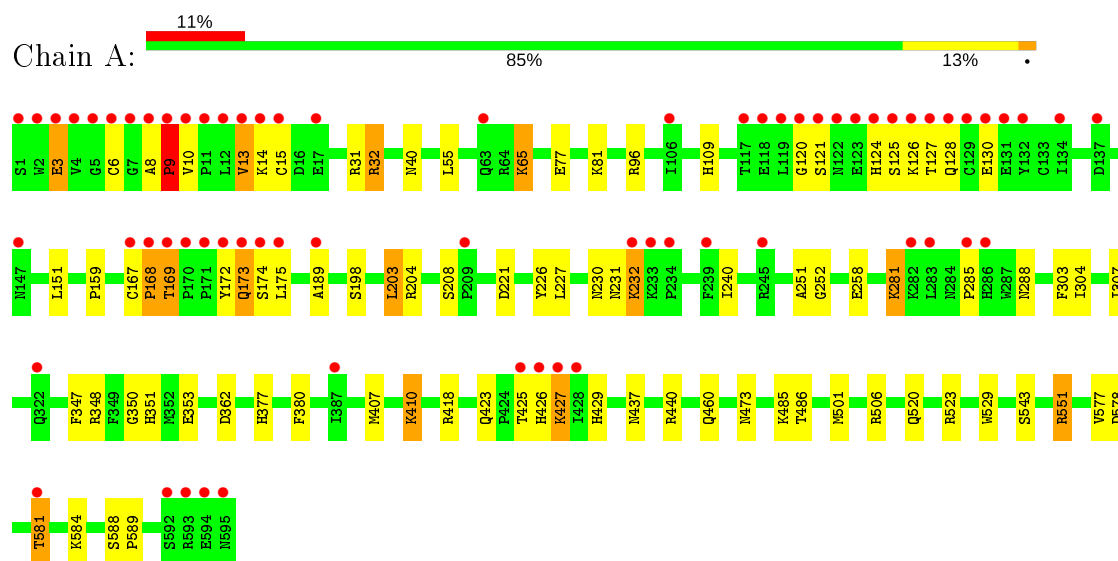
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	262	Total	O	0	0
			262	262		

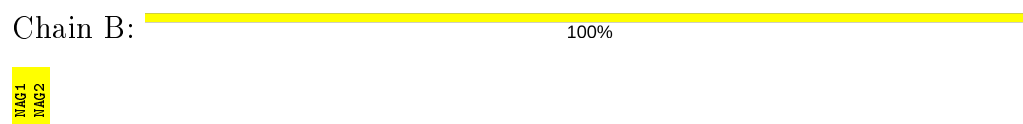
### 3 Residue-property plots [i](#)

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



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



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	53.85Å 80.19Å 75.09Å 90.00° 105.41° 90.00°	Depositor
Resolution (Å)	72.39 – 1.98 35.07 – 1.98	Depositor EDS
% Data completeness (in resolution range)	95.8 (72.39-1.98) 95.9 (35.07-1.98)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.15	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.55 (at 1.98Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.188 , 0.235 0.195 , 0.238	Depositor DCC
$R_{free}$ test set	1655 reflections (4.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.5	Xtriage
Anisotropy	0.427	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 70.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5181	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.04% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PGE, NAG, SEP, CA, OSM, SO4, HEM, IOD

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	0.83	1/4886 (0.0%)	0.99	18/6623 (0.3%)

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	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	353	GLU	CD-OE1	6.89	1.33	1.25

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	258[A]	GLU	CA-CB-CG	13.63	143.40	113.40
1	A	258[B]	GLU	CA-CB-CG	13.63	143.40	113.40
1	A	31	ARG	NE-CZ-NH1	7.39	124.00	120.30
1	A	362	ASP	CB-CG-OD1	7.30	124.87	118.30
1	A	440	ARG	NE-CZ-NH1	7.00	123.80	120.30
1	A	362	ASP	CB-CG-OD2	-6.59	112.37	118.30
1	A	232	LYS	N-CA-C	-6.35	93.84	111.00
1	A	31	ARG	NE-CZ-NH2	-6.07	117.27	120.30
1	A	506	ARG	NE-CZ-NH1	5.97	123.28	120.30
1	A	551	ARG	NE-CZ-NH2	-5.94	117.33	120.30
1	A	348	ARG	NE-CZ-NH2	-5.85	117.37	120.30
1	A	440	ARG	NE-CZ-NH2	-5.85	117.38	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	96	ARG	NE-CZ-NH1	5.71	123.15	120.30
1	A	551	ARG	NE-CZ-NH1	5.66	123.13	120.30
1	A	32	ARG	NE-CZ-NH1	-5.25	117.68	120.30
1	A	9	PRO	CA-N-CD	-5.17	104.26	111.50
1	A	418	ARG	NE-CZ-NH1	5.11	122.86	120.30
1	A	204	ARG	NE-CZ-NH2	-5.03	117.78	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	15	CYS	Peptide
1	A	167	CYS	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	4767	0	4689	58	0
2	B	28	0	25	0	0
3	A	5	0	0	0	0
4	A	4	0	5	0	0
5	A	9	0	0	1	0
6	A	20	0	28	1	0
7	A	42	0	39	0	0
8	A	1	0	0	0	0
9	A	43	0	30	5	0
10	A	262	0	0	9	1
All	All	5181	0	4816	63	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 6.

All (63) 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:159:PRO:HG3	1:A:426:HIS:CE1	2.07	0.89
1:A:126:LYS:O	1:A:130:GLU:OE1	2.00	0.80
1:A:159:PRO:HG3	1:A:426:HIS:HE1	1.47	0.77
1:A:189:ALA:HB2	1:A:304:ILE:HD12	1.66	0.76
1:A:172:TYR:HB2	1:A:173:GLN:HB3	1.69	0.75
1:A:303:PHE:CZ	1:A:307:ILE:HD13	2.27	0.70
1:A:230:ASN:O	10:A:902:HOH:O	2.10	0.69
1:A:159:PRO:CG	1:A:426:HIS:HE1	2.06	0.69
9:A:820:HEM:HMB1	9:A:820:HEM:HBB2	1.74	0.69
1:A:303:PHE:CE1	1:A:307:ILE:CD1	2.76	0.68
1:A:159:PRO:CG	1:A:426:HIS:CE1	2.76	0.68
1:A:351:HIS:HD1	1:A:437:ASN:HD21	1.41	0.66
9:A:820:HEM:HMC1	9:A:820:HEM:HBC2	1.79	0.65
1:A:127:THR:HA	1:A:130:GLU:CD	2.16	0.65
1:A:551:ARG:HD3	1:A:584:LYS:HA	1.80	0.62
1:A:303:PHE:CE1	1:A:307:ILE:HD11	2.34	0.62
1:A:303:PHE:CZ	1:A:307:ILE:CD1	2.84	0.61
1:A:3:GLU:HB2	1:A:175:LEU:HD21	1.81	0.61
1:A:127:THR:HA	1:A:130:GLU:OE2	2.00	0.60
1:A:8:ALA:N	1:A:9:PRO:HD2	2.17	0.60
1:A:231:ASN:O	1:A:232:LYS:HG3	2.04	0.58
1:A:65:LYS:NZ	10:A:906:HOH:O	2.35	0.57
1:A:281:LYS:HE2	1:A:285:PRO:O	2.08	0.53
1:A:460:GLN:H	6:A:807:PGE:H12	1.75	0.52
1:A:423:GLN:NE2	10:A:912:HOH:O	2.42	0.52
1:A:426:HIS:O	1:A:427:LYS:HB2	2.10	0.52
1:A:130:GLU:N	1:A:130:GLU:OE1	2.24	0.52
9:A:820:HEM:HBB2	9:A:820:HEM:CMB	2.39	0.51
1:A:377:HIS:HA	1:A:380:PHE:CE2	2.46	0.51
5:A:804:IOD:I	5:A:805:IOD:I	3.70	0.50
1:A:3:GLU:CB	1:A:175:LEU:HD21	2.40	0.50
1:A:125:SER:HA	1:A:128:GLN:HB3	1.95	0.49
1:A:523:ARG:HG3	1:A:529:TRP:CE2	2.48	0.49
1:A:127:THR:O	1:A:130:GLU:HG2	2.13	0.49
1:A:124:HIS:O	1:A:127:THR:N	2.47	0.48
1:A:426:HIS:CD2	1:A:426:HIS:N	2.82	0.48
1:A:588:SER:HB2	1:A:589:PRO:HD3	1.96	0.47
1:A:120:GLY:CA	1:A:126:LYS:HB2	2.45	0.47
1:A:423:GLN:O	10:A:903:HOH:O	2.20	0.47
1:A:231:ASN:O	1:A:232:LYS:CG	2.62	0.47
1:A:350:GLY:HA3	9:A:820:HEM:CBC	2.44	0.47
1:A:577:VAL:HG21	1:A:581:THR:CG2	2.44	0.47

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:14:LYS:HE2	1:A:14:LYS:HA	1.96	0.47
1:A:288:ASN:ND2	10:A:916:HOH:O	2.48	0.46
1:A:407:MET:HB3	1:A:501:MET:CE	2.46	0.46
1:A:410:LYS:HE2	1:A:473:ASN:HD21	1.80	0.46
1:A:10:VAL:HG11	1:A:40:ASN:HB2	1.98	0.45
1:A:203:LEU:HD21	1:A:252:GLY:HA2	1.99	0.45
1:A:9:PRO:HG3	1:A:168:PRO:CG	2.46	0.45
1:A:9:PRO:HG3	1:A:168:PRO:HG3	1.98	0.44
1:A:130:GLU:HB3	1:A:159:PRO:HG3	2.00	0.44
1:A:168:PRO:O	1:A:169:THR:CB	2.66	0.43
1:A:232:LYS:O	10:A:902:HOH:O	2.21	0.43
1:A:429:HIS:O	10:A:904:HOH:O	2.21	0.43
9:A:820:HEM:CMC	9:A:820:HEM:HBC2	2.47	0.43
1:A:120:GLY:HA2	1:A:126:LYS:HB2	2.01	0.43
1:A:77:GLU:OE2	1:A:81:LYS:NZ	2.42	0.43
1:A:240:ILE:HD12	10:A:1070:HOH:O	2.20	0.42
1:A:109:HIS:CD2	10:A:922:HOH:O	2.72	0.42
1:A:3:GLU:CD	1:A:175:LEU:HD11	2.40	0.42
1:A:221:ASP:HB2	1:A:226:TYR:CZ	2.56	0.41
1:A:227:LEU:CD2	1:A:251:ALA:HB2	2.51	0.40
1:A:172:TYR:CD2	1:A:173:GLN:HG3	2.56	0.40

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:964:HOH:O	10:A:1094:HOH:O[2_445]	1.91	0.29

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	593/595 (100%)	565 (95%)	20 (3%)	8 (1%)	12 4

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	9	PRO
1	A	13	VAL
1	A	169	THR
1	A	427	LYS
1	A	425	THR
1	A	174	SER
1	A	168	PRO
1	A	485	LYS

### 5.3.2 Protein sidechains ⓘ

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	518/517 (100%)	499 (96%)	19 (4%)	34 22

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	GLU
1	A	6	CYS
1	A	13	VAL
1	A	32	ARG
1	A	55	LEU
1	A	65	LYS
1	A	121	SER
1	A	151	LEU
1	A	173	GLN
1	A	203	LEU
1	A	208	SER
1	A	281	LYS
1	A	347	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	410	LYS
1	A	486	THR
1	A	520	GLN
1	A	543	SER
1	A	578	ASP
1	A	581	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	173	GLN
1	A	426	HIS
1	A	520	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	SEP	A	198	1	8,9,10	2.39	4 (50%)	8,12,14	8.18	6 (75%)

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	SEP	A	198	1	-	5/5/8/10	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	198	SEP	P-OG	3.81	1.72	1.60
1	A	198	SEP	OG-CB	3.49	1.58	1.44
1	A	198	SEP	CB-CA	3.28	1.61	1.52
1	A	198	SEP	P-O3P	-2.03	1.47	1.54

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	198	SEP	OG-CB-CA	18.64	126.29	108.14
1	A	198	SEP	P-OG-CB	12.42	152.50	118.30
1	A	198	SEP	O3P-P-O1P	-3.28	97.83	110.68
1	A	198	SEP	O3P-P-OG	2.84	114.29	106.73
1	A	198	SEP	O3P-P-O2P	2.53	117.30	107.64
1	A	198	SEP	O2P-P-O1P	2.32	119.75	110.68

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	198	SEP	N-CA-CB-OG
1	A	198	SEP	CA-CB-OG-P
1	A	198	SEP	CB-OG-P-O1P
1	A	198	SEP	CB-OG-P-O2P
1	A	198	SEP	CB-OG-P-O3P

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates ⓘ

2 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	0.69	0	17,19,21	1.28	3 (17%)
2	NAG	B	2	2	14,14,15	0.66	0	17,19,21	2.16	4 (23%)

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	-	0/6/23/26	0/1/1/1
2	NAG	B	2	2	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2	NAG	C1-O5-C5	7.05	121.74	112.19
2	B	2	NAG	C4-C3-C2	-3.16	106.39	111.02
2	B	2	NAG	C1-C2-N2	2.76	115.21	110.49
2	B	1	NAG	O4-C4-C5	2.76	116.15	109.30
2	B	1	NAG	C1-C2-N2	2.30	114.41	110.49
2	B	1	NAG	O4-C4-C3	-2.27	105.10	110.35
2	B	2	NAG	O5-C5-C6	2.10	110.49	107.20

There are no chirality outliers.

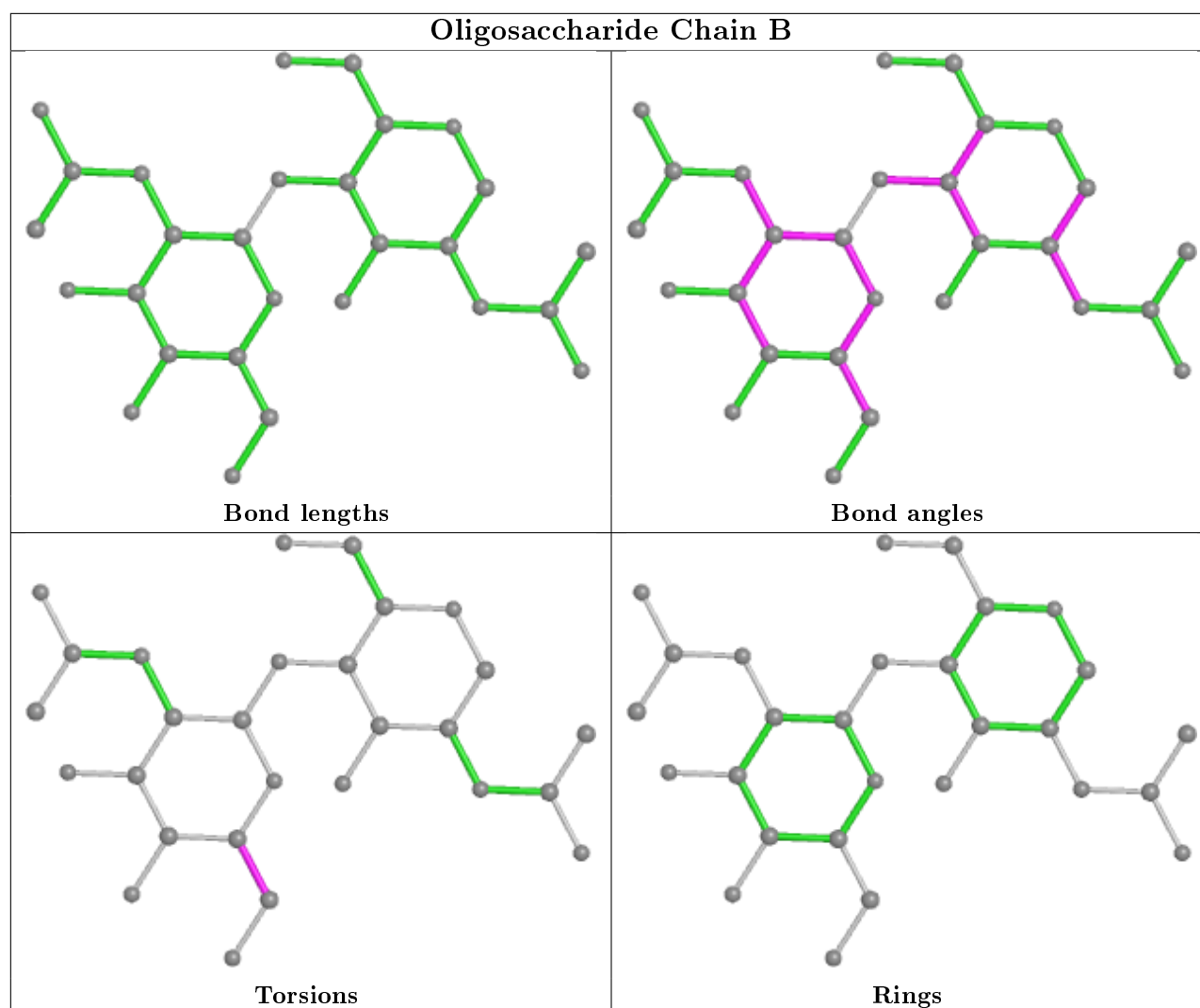
All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	2	NAG	O5-C5-C6-O6
2	B	2	NAG	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 ⓘ

Of 18 ligands modelled in this entry, 10 are monoatomic - leaving 8 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
9	HEM	A	820	1,10	27,50,50	0.83	1 (3%)	17,82,82	1.71	4 (23%)
3	SO4	A	801	-	4,4,4	0.37	0	6,6,6	0.68	0
6	PGE	A	806	-	9,9,9	0.50	0	8,8,8	0.82	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	NAG	A	812	1	14,14,15	0.70	0	17,19,21	1.60	2 (11%)
7	NAG	A	809	1	14,14,15	0.59	0	17,19,21	1.37	2 (11%)
4	OSM	A	802	-	1,3,3	0.58	0	0,2,2	0.00	-
6	PGE	A	807	-	9,9,9	0.59	0	8,8,8	1.01	0
7	NAG	A	808	1	14,14,15	0.58	0	17,19,21	1.61	4 (23%)

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	820	1,10	-	0/6/54/54	-
6	PGE	A	806	-	-	2/7/7/7	-
4	OSM	A	802	-	-	0/0/1/1	-
7	NAG	A	809	1	-	2/6/23/26	0/1/1/1
7	NAG	A	812	1	-	2/6/23/26	0/1/1/1
6	PGE	A	807	-	-	5/7/7/7	-
7	NAG	A	808	1	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	A	820	HEM	C3C-C2C	-2.02	1.37	1.40

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	812	NAG	C1-O5-C5	4.32	118.05	112.19
7	A	812	NAG	O5-C1-C2	-3.75	105.37	111.29
7	A	808	NAG	O6-C6-C5	-3.59	98.96	111.29
9	A	820	HEM	CBD-CAD-C3D	-3.30	106.40	112.48
9	A	820	HEM	CBA-CAA-C2A	-3.27	106.46	112.49
9	A	820	HEM	C1D-C2D-C3D	-3.04	104.88	107.00
9	A	820	HEM	C3B-C4B-NB	-2.65	105.78	109.21
7	A	808	NAG	C1-O5-C5	2.47	115.55	112.19
7	A	808	NAG	C3-C4-C5	-2.22	106.28	110.24
7	A	809	NAG	O7-C7-N2	2.18	125.96	121.95
7	A	808	NAG	O5-C1-C2	-2.10	107.97	111.29
7	A	809	NAG	O3-C3-C2	-2.10	105.12	109.47

There are no chirality outliers.

All (11) torsion outliers are listed below:

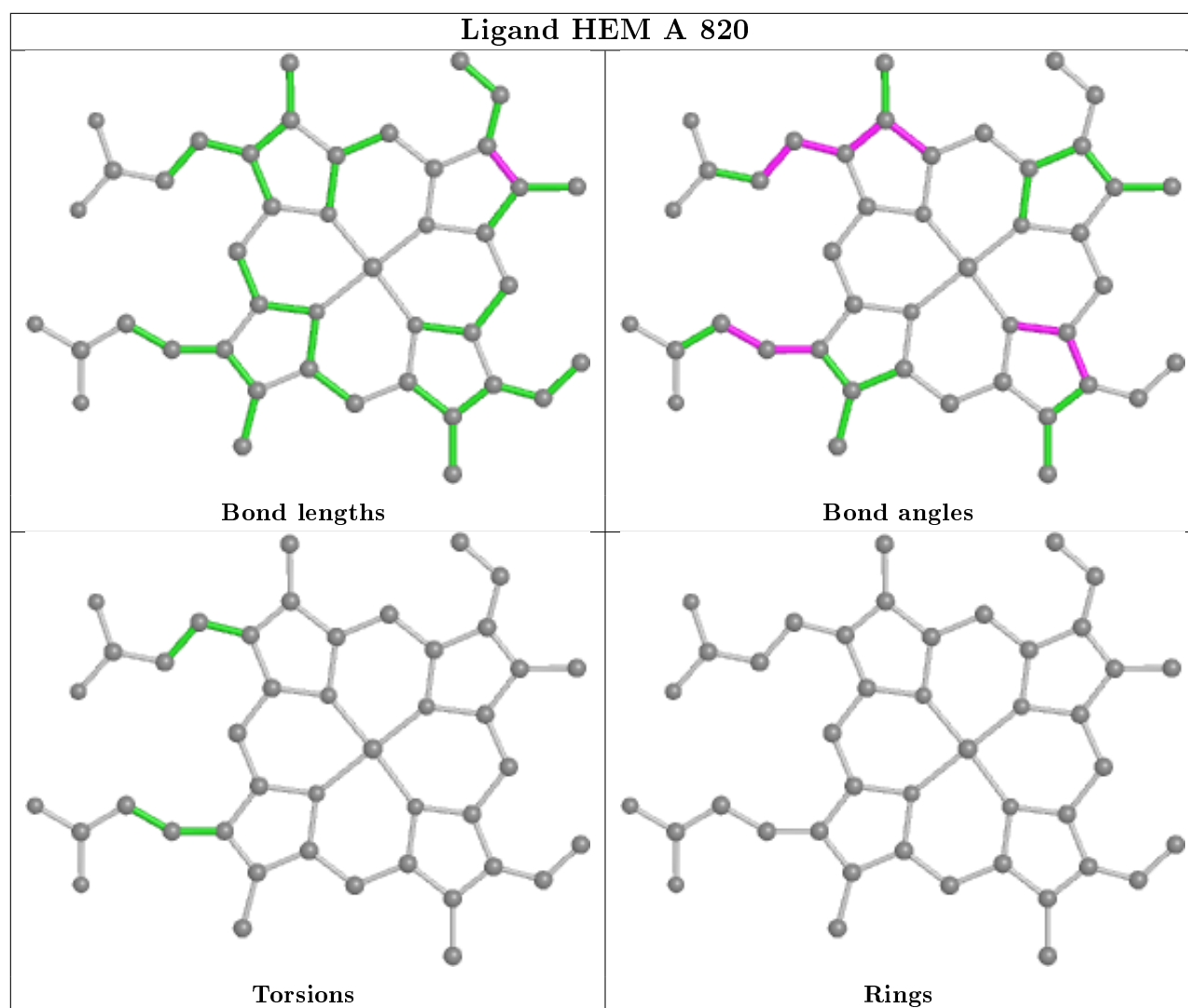
Mol	Chain	Res	Type	Atoms
6	A	806	PGE	O1-C1-C2-O2
6	A	807	PGE	O1-C1-C2-O2
6	A	807	PGE	O3-C5-C6-O4
6	A	807	PGE	O2-C3-C4-O3
7	A	809	NAG	O5-C5-C6-O6
7	A	809	NAG	C4-C5-C6-O6
6	A	807	PGE	C1-C2-O2-C3
6	A	807	PGE	C3-C4-O3-C5
6	A	806	PGE	C3-C4-O3-C5
7	A	812	NAG	O5-C5-C6-O6
7	A	812	NAG	C4-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	820	HEM	5	0
6	A	807	PGE	1	0

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



## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 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	594/595 (99%)	0.87	68 (11%) 5 5	19, 31, 106, 182	0

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	7	GLY	28.0
1	A	5	GLY	26.3
1	A	9	PRO	19.8
1	A	595	ASN	18.9
1	A	121	SER	18.5
1	A	2	TRP	17.7
1	A	169	THR	16.8
1	A	4	VAL	16.5
1	A	173	GLN	15.1
1	A	122	ASN	14.5
1	A	6	CYS	13.4
1	A	3	GLU	13.3
1	A	8	ALA	12.3
1	A	13	VAL	11.5
1	A	124	HIS	10.6
1	A	170	PRO	9.6
1	A	234	PRO	9.4
1	A	132	TYR	9.0
1	A	1	SER	9.0
1	A	172	TYR	8.6
1	A	127	THR	8.2
1	A	10	VAL	7.9
1	A	12	LEU	7.4
1	A	125	SER	7.3
1	A	119	LEU	7.1
1	A	118	GLU	7.0
1	A	120	GLY	6.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	123	GLU	6.5
1	A	594	GLU	6.2
1	A	11	PRO	6.1
1	A	117	THR	6.1
1	A	131	GLU	6.0
1	A	175	LEU	5.5
1	A	128	GLN	5.3
1	A	171	PRO	5.3
1	A	285	PRO	5.0
1	A	134	ILE	4.7
1	A	174	SER	4.4
1	A	593	ARG	4.4
1	A	168	PRO	4.4
1	A	428	ILE	4.4
1	A	167	CYS	4.0
1	A	17	GLU	3.9
1	A	126	LYS	3.8
1	A	129	CYS	3.6
1	A	63	GLN	3.5
1	A	322	GLN	3.3
1	A	15	CYS	3.2
1	A	426	HIS	3.2
1	A	427	LYS	3.1
1	A	245	ARG	3.1
1	A	14	LYS	3.1
1	A	147	ASN	3.1
1	A	239	PHE	3.0
1	A	581	THR	3.0
1	A	425	THR	2.9
1	A	286	HIS	2.9
1	A	232	LYS	2.8
1	A	592	SER	2.7
1	A	209	PRO	2.7
1	A	387	ILE	2.5
1	A	283	LEU	2.4
1	A	189	ALA	2.4
1	A	106	ILE	2.4
1	A	282	LYS	2.2
1	A	233	LYS	2.1
1	A	130	GLU	2.1
1	A	137	ASP	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.

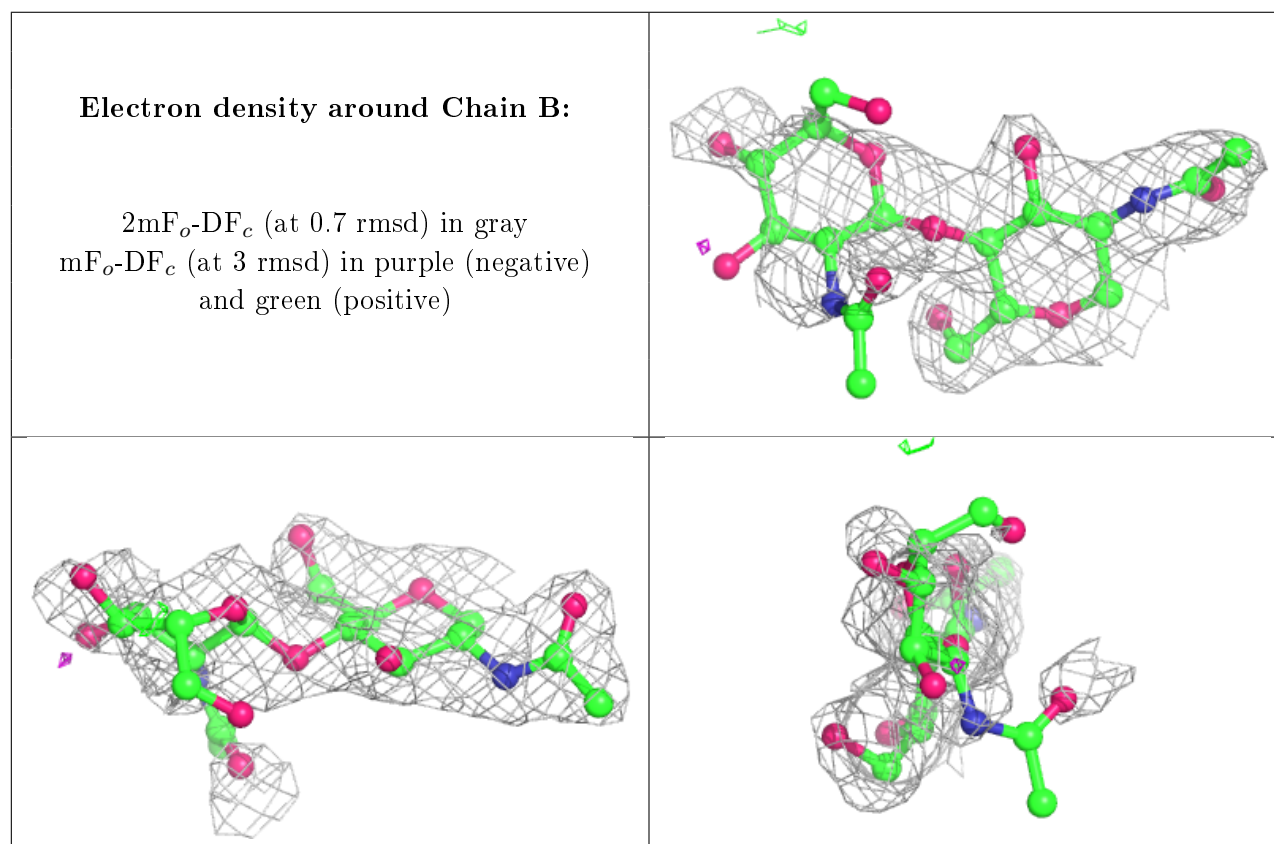
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
1	SEP	A	198	10/11	0.93	0.16	26,32,51,54	0

## 6.3 Carbohydrates [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	NAG	B	2	14/15	0.74	0.43	68,80,97,101	0
2	NAG	B	1	14/15	0.91	0.14	38,44,51,58	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.



## 6.4 Ligands ⓘ

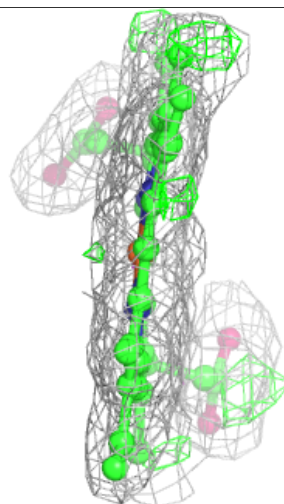
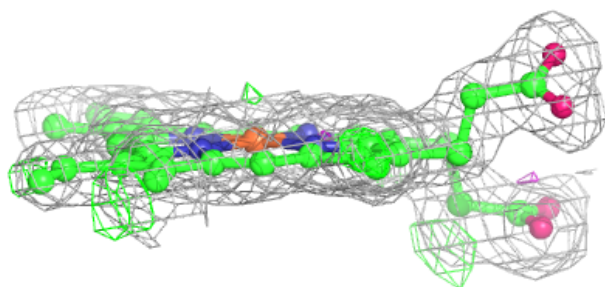
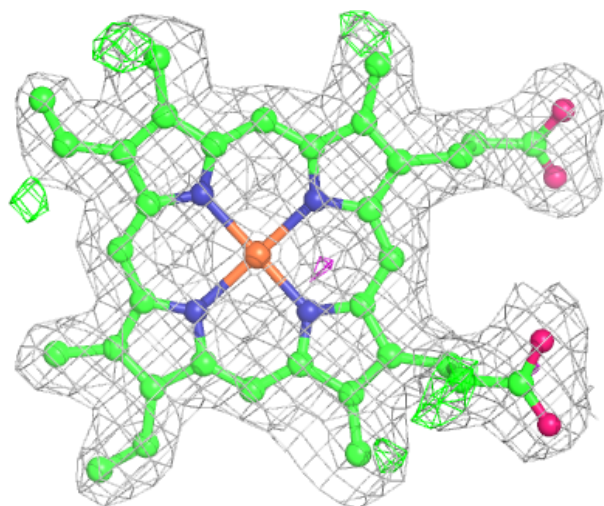
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
7	NAG	A	812	14/15	0.81	0.18	51,66,72,78	0
6	PGE	A	807	10/10	0.86	0.13	33,48,51,52	0
7	NAG	A	808	14/15	0.86	0.19	48,53,59,66	0
6	PGE	A	806	10/10	0.88	0.16	38,41,43,45	0
7	NAG	A	809	14/15	0.91	0.13	38,49,56,66	0
9	HEM	A	820	43/43	0.96	0.17	20,25,30,37	0
4	OSM	A	802	4/4	0.97	0.15	15,19,22,29	0
5	IOD	A	818	1/1	0.97	0.09	73,73,73,73	1
5	IOD	A	805	1/1	0.98	0.05	72,72,72,72	1
5	IOD	A	814	1/1	0.98	0.03	47,47,47,47	1
5	IOD	A	819	1/1	0.98	0.08	74,74,74,74	1
5	IOD	A	816	1/1	0.99	0.06	30,30,30,30	0
5	IOD	A	815	1/1	0.99	0.07	34,34,34,34	0
5	IOD	A	817	1/1	0.99	0.07	33,33,33,33	1
5	IOD	A	804	1/1	0.99	0.03	41,41,41,41	1
3	SO4	A	801	5/5	0.99	0.05	32,32,35,35	0
8	CA	A	813	1/1	1.00	0.11	24,24,24,24	0
5	IOD	A	803	1/1	1.00	0.07	23,23,23,23	1

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 820:**

$2mF_o-DF_c$  (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.