



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 14, 2023 – 08:07 PM JST

PDB ID : 6A4Y  
Title : Crystal structure of bovine lactoperoxidase with partial occupancies of iodide and SCN<sup>-</sup> ions at the substrate binding site on the distal heme side at 1.92 Å resolution  
Authors : Singh, P.K.; Sirohi, H.V.; kaur, P.; Sharma, S.; Singh, T.P.  
Deposited on : 2018-06-21  
Resolution : 1.92 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
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

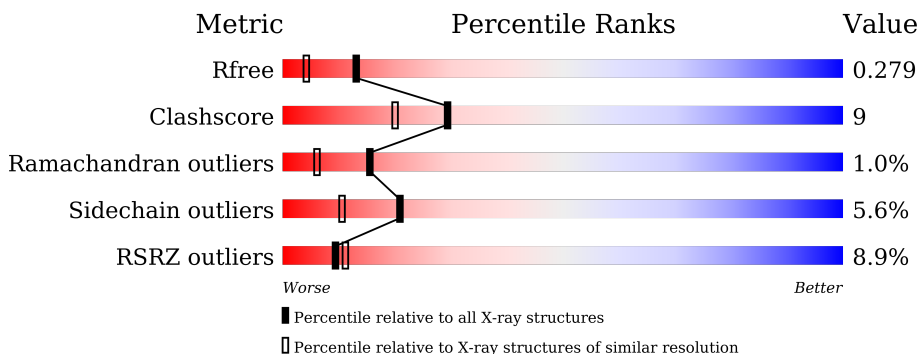
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.92 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	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	595	 9% 79% 18% .
2	B	2	 50% 50%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	IOD	A	607	-	-	X	-
7	SCN	A	614	-	-	X	-
7	SCN	A	619	-	-	X	-

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 5126 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.

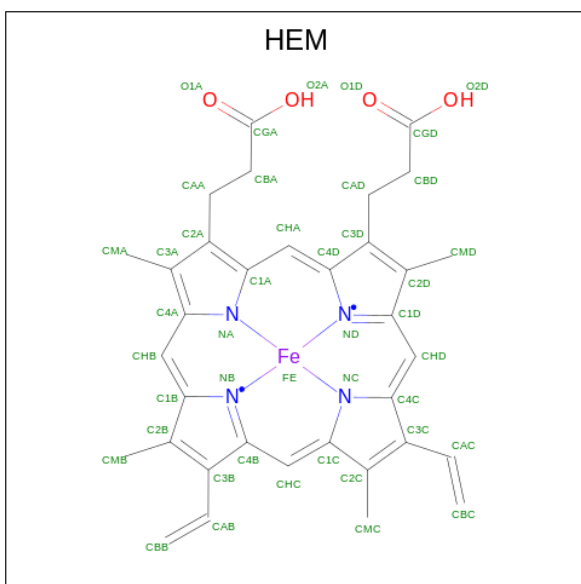
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	595	4774	3037	847	863	1	26	0	0	0

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



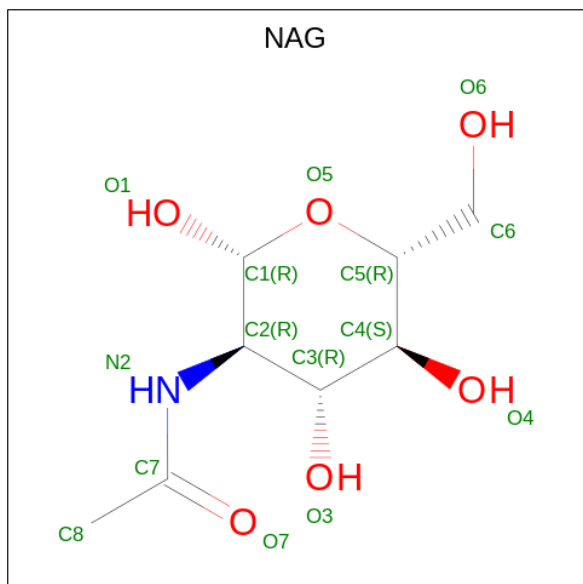
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	2	28	16	2	10	0	0	0

- Molecule 3 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
3	A	1	43	34	1	4	4	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
			Total	C	N	O		
4	A	1	14	8	1	5	0	0
4	A	1	14	8	1	5	0	0
4	A	1	14	8	1	5	0	0

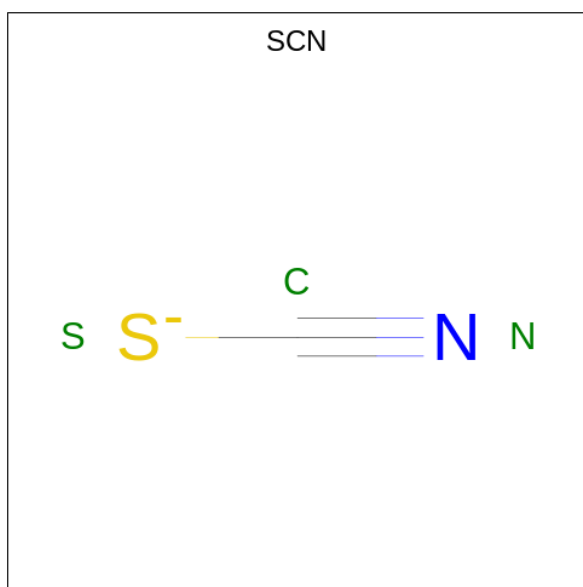
- Molecule 5 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	6	Total 6 I 6	0	0

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

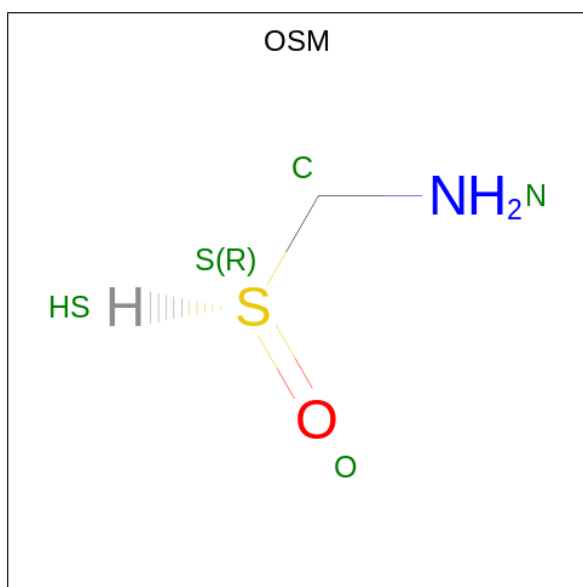
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total 1 Ca 1	0	0

- Molecule 7 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).



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

- Molecule 8 is 1-(OXIDOSULFANYL)METHANAMINE (three-letter code: OSM) (formula: CH<sub>5</sub>NOS).



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

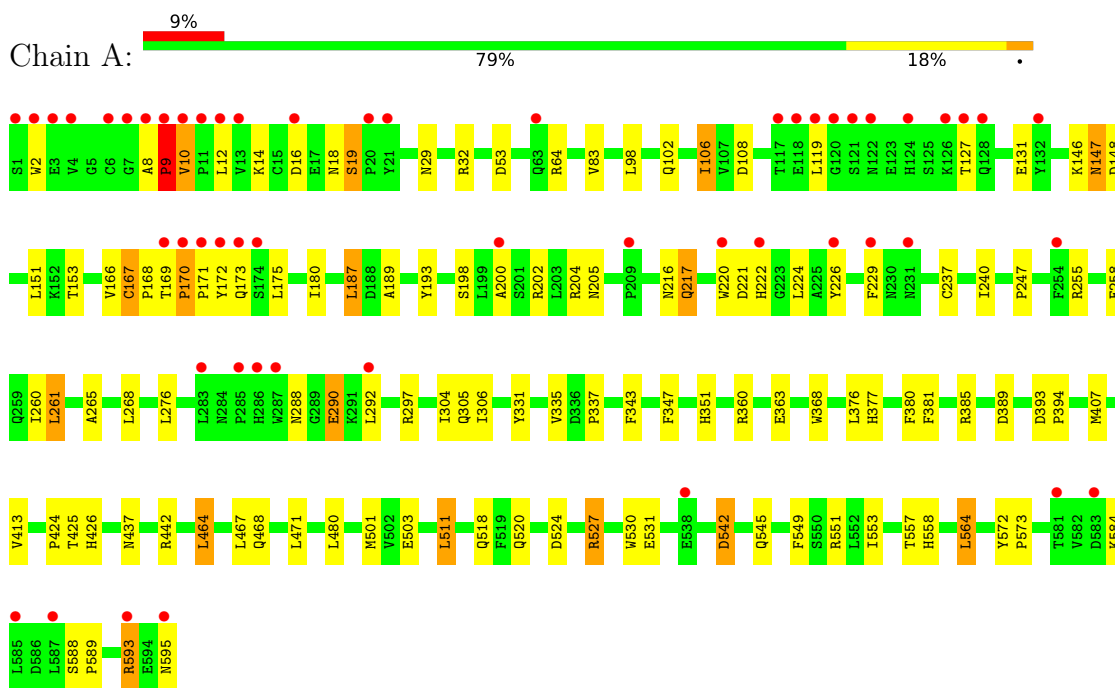
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
9	A	204	204	204	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



- 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, $\alpha$ , $\beta$ , $\gamma$	53.74Å 79.80Å 65.18Å 90.00° 91.43° 90.00°	Depositor
Resolution (Å)	50.52 – 1.92 50.47 – 1.92	Depositor EDS
% Data completeness (in resolution range)	95.2 (50.52-1.92) 95.3 (50.47-1.92)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.17 (at 1.91Å)	Xtrriage
Refinement program	REFMAC 5.8.0230	Depositor
R, $R_{free}$	0.222 , 0.279 0.227 , 0.279	Depositor DCC
$R_{free}$ test set	1899 reflections (4.70%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.8	Xtrriage
Anisotropy	0.577	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 55.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.045 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5126	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.50% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.56	0/4891	0.71	1/6634 (0.0%)

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	3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	442	ARG	NE-CZ-NH1	-5.01	117.80	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	10	VAL	Peptide
1	A	527	ARG	Sidechain
1	A	64	ARG	Sidechain

### 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	4774	0	4687	83	0
2	B	28	0	25	0	0
3	A	43	0	30	12	0
4	A	42	0	39	0	0
5	A	6	0	0	5	0
6	A	1	0	0	0	0
7	A	24	0	0	7	0
8	A	4	0	5	0	0
9	A	204	0	0	8	0
All	All	5126	0	4786	92	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 9.

All (92) 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:108:ASP:OD2	3:A:601:HEM:CMD	1.70	1.40
1:A:258:GLU:OE2	3:A:601:HEM:HMB1	1.31	1.28
5:A:607:IOD:I	7:A:619:SCN:N	2.43	1.19
1:A:108:ASP:OD2	3:A:601:HEM:HMD1	0.82	0.98
1:A:258:GLU:OE2	3:A:601:HEM:CMB	2.16	0.93
7:A:614:SCN:N	9:A:702:HOH:O	2.01	0.92
5:A:607:IOD:I	7:A:619:SCN:C	1.31	0.90
1:A:351:HIS:HD1	1:A:437:ASN:HD21	1.20	0.88
5:A:607:IOD:I	7:A:619:SCN:S	0.62	0.88
1:A:29:ASN:OD1	9:A:701:HOH:O	1.97	0.81
1:A:385:ARG:NE	9:A:704:HOH:O	2.15	0.78
1:A:8:ALA:O	1:A:10:VAL:N	2.18	0.77
1:A:424:PRO:O	5:A:610:IOD:I	2.74	0.76
1:A:167:CYS:HB2	1:A:168:PRO:HD2	1.69	0.74
1:A:204:ARG:NH1	1:A:290:GLU:OE2	2.19	0.74
1:A:108:ASP:CG	3:A:601:HEM:HMD1	2.01	0.73
1:A:306:ILE:HA	7:A:616:SCN:S	2.29	0.73
1:A:32:ARG:NH1	9:A:705:HOH:O	2.26	0.69
1:A:53:ASP:OD2	9:A:703:HOH:O	2.13	0.66
3:A:601:HEM:HBB2	3:A:601:HEM:HMB2	1.79	0.65
1:A:10:VAL:O	1:A:10:VAL:HG12	1.97	0.65
1:A:10:VAL:O	1:A:10:VAL:CG1	2.47	0.63
1:A:167:CYS:HB2	1:A:168:PRO:CD	2.30	0.61

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:335:VAL:O	1:A:337:PRO:HD3	2.01	0.61
1:A:407:MET:HB3	1:A:501:MET:CE	2.31	0.60
1:A:108:ASP:OD2	3:A:601:HEM:C2D	2.51	0.60
1:A:8:ALA:N	1:A:9:PRO:CD	2.65	0.60
1:A:167:CYS:CB	1:A:168:PRO:CD	2.81	0.59
1:A:258:GLU:CD	3:A:601:HEM:HMB1	2.21	0.58
1:A:146:LYS:O	1:A:147:ASN:HB2	2.03	0.57
1:A:168:PRO:HB2	1:A:171:PRO:HD2	1.86	0.57
1:A:503:GLU:HG2	9:A:882:HOH:O	2.05	0.55
1:A:108:ASP:CG	3:A:601:HEM:CMD	2.70	0.54
1:A:343:PHE:CD2	1:A:518:GLN:HG2	2.43	0.54
1:A:542:ASP:O	1:A:545:GLN:HG2	2.07	0.54
3:A:601:HEM:HMC2	3:A:601:HEM:HBC2	1.89	0.53
1:A:148:ASP:O	1:A:151:LEU:HB2	2.09	0.53
1:A:530:TRP:CE2	1:A:531:GLU:HG3	2.45	0.51
1:A:168:PRO:HG2	1:A:172:TYR:HA	1.92	0.51
1:A:224:LEU:HD12	1:A:558:HIS:CE1	2.46	0.51
1:A:229:PHE:CD1	1:A:247:PRO:HG2	2.46	0.51
1:A:551:ARG:HD3	1:A:584:LYS:HA	1.92	0.50
1:A:168:PRO:HD3	1:A:173:GLN:HG2	1.93	0.50
1:A:168:PRO:CD	1:A:173:GLN:HG2	2.42	0.50
1:A:588:SER:N	1:A:589:PRO:CD	2.74	0.50
1:A:520:GLN:NE2	1:A:524:ASP:OD2	2.36	0.50
1:A:255:ARG:HD2	5:A:607:IOD:I	2.82	0.49
7:A:614:SCN:C	9:A:702:HOH:O	2.51	0.49
1:A:146:LYS:O	1:A:147:ASN:CB	2.60	0.48
1:A:189:ALA:HB2	1:A:304:ILE:HD12	1.94	0.48
1:A:200:ALA:O	1:A:204:ARG:HG3	2.12	0.48
1:A:377:HIS:HA	1:A:380:PHE:CE2	2.49	0.48
1:A:260:ILE:HG23	1:A:261:LEU:HD13	1.95	0.48
1:A:222:HIS:HB2	1:A:558:HIS:CE1	2.49	0.48
1:A:221:ASP:HB2	1:A:226:TYR:CZ	2.49	0.48
1:A:102:GLN:HE21	1:A:106:ILE:HD12	1.79	0.47
1:A:237:CYS:HA	1:A:381:PHE:O	2.14	0.47
1:A:368:TRP:CH2	1:A:389:ASP:O	2.68	0.47
1:A:102:GLN:HE21	1:A:106:ILE:CD1	2.29	0.46
1:A:83:VAL:HG12	1:A:413:VAL:HB	1.97	0.46
1:A:8:ALA:N	1:A:9:PRO:HD2	2.30	0.46
1:A:216:ASN:CG	1:A:217:GLN:H	2.18	0.46
1:A:393:ASP:N	1:A:394:PRO:CD	2.79	0.45
1:A:288:ASN:OD1	1:A:288:ASN:N	2.50	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:166:VAL:HG12	1:A:180:ILE:HG12	1.98	0.44
1:A:221:ASP:O	1:A:222:HIS:C	2.56	0.44
1:A:217:GLN:HE21	1:A:217:GLN:HB2	1.64	0.44
1:A:221:ASP:HB2	1:A:226:TYR:CE1	2.52	0.44
1:A:572:TYR:HA	1:A:573:PRO:HA	1.80	0.44
1:A:467:LEU:HG	1:A:471:LEU:HD22	1.99	0.44
1:A:106:ILE:HD11	1:A:265:ALA:HB3	2.00	0.43
1:A:407:MET:CB	1:A:501:MET:CE	2.97	0.43
1:A:187:LEU:HD13	1:A:305:GLN:HA	1.99	0.43
1:A:511:LEU:HD12	1:A:511:LEU:HA	1.91	0.43
1:A:18:ASN:O	1:A:19:SER:O	2.36	0.43
1:A:425:THR:HB	1:A:426:HIS:CD2	2.54	0.43
3:A:601:HEM:CMB	3:A:601:HEM:HBB2	2.46	0.42
1:A:193:TYR:OH	1:A:297:ARG:HA	2.19	0.42
1:A:216:ASN:C	1:A:217:GLN:HG3	2.40	0.42
1:A:464:LEU:O	1:A:468:GLN:HG3	2.20	0.42
1:A:204:ARG:NH1	9:A:736:HOH:O	2.53	0.42
1:A:425:THR:HB	1:A:426:HIS:NE2	2.35	0.42
1:A:127:THR:O	1:A:131:GLU:HB2	2.19	0.41
1:A:169:THR:N	1:A:170:PRO:CD	2.83	0.41
1:A:424:PRO:O	7:A:618:SCN:C	2.58	0.41
1:A:549:PHE:HE1	1:A:553:ILE:HD11	1.86	0.41
1:A:593:ARG:C	1:A:595:ASN:HD22	2.24	0.41
1:A:393:ASP:OD1	1:A:557:THR:HB	2.21	0.41
1:A:331:TYR:CD1	1:A:527:ARG:HA	2.57	0.40
1:A:564:LEU:HD12	1:A:564:LEU:HA	1.89	0.40
1:A:205:ASN:OD1	1:A:205:ASN:C	2.60	0.40
3:A:601:HEM:HBC2	3:A:601:HEM:CMC	2.52	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	592/595 (100%)	542 (92%)	44 (7%)	6 (1%)	15 6

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	9	PRO
1	A	167	CYS
1	A	16	ASP
1	A	147	ASN
1	A	19	SER
1	A	170	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	517/517 (100%)	488 (94%)	29 (6%)	21 10

All (29) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	TRP
1	A	9	PRO
1	A	12	LEU
1	A	14	LYS
1	A	98	LEU
1	A	106	ILE
1	A	119	LEU
1	A	153	THR
1	A	175	LEU
1	A	187	LEU
1	A	202	ARG
1	A	217	GLN
1	A	220	TRP
1	A	240	ILE
1	A	261	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	268	LEU
1	A	276	LEU
1	A	290	GLU
1	A	292	LEU
1	A	347	PHE
1	A	360	ARG
1	A	363	GLU
1	A	376	LEU
1	A	464	LEU
1	A	480	LEU
1	A	511	LEU
1	A	542	ASP
1	A	564	LEU
1	A	593	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	217	GLN
1	A	322	GLN
1	A	468	GLN
1	A	595	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](#)

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	0.95	0	8,12,14	1.45	2 (25%)

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	-	1/5/8/10	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	198	SEP	P-OG-CB	3.20	127.10	118.30
1	A	198	SEP	O2P-P-O1P	2.09	118.85	110.68

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	198	SEP	N-CA-CB-OG

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

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	2,1	14,14,15	0.34	0	17,19,21	1.67	4 (23%)
2	NAG	B	2	2	14,14,15	0.48	0	17,19,21	1.08	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
2	NAG	B	1	2,1	-	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 (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	1	NAG	C3-C4-C5	3.65	116.75	110.24
2	B	1	NAG	O5-C1-C2	-2.88	106.73	111.29
2	B	1	NAG	O4-C4-C3	-2.78	103.92	110.35
2	B	1	NAG	C4-C3-C2	2.11	114.12	111.02

There are no chirality outliers.

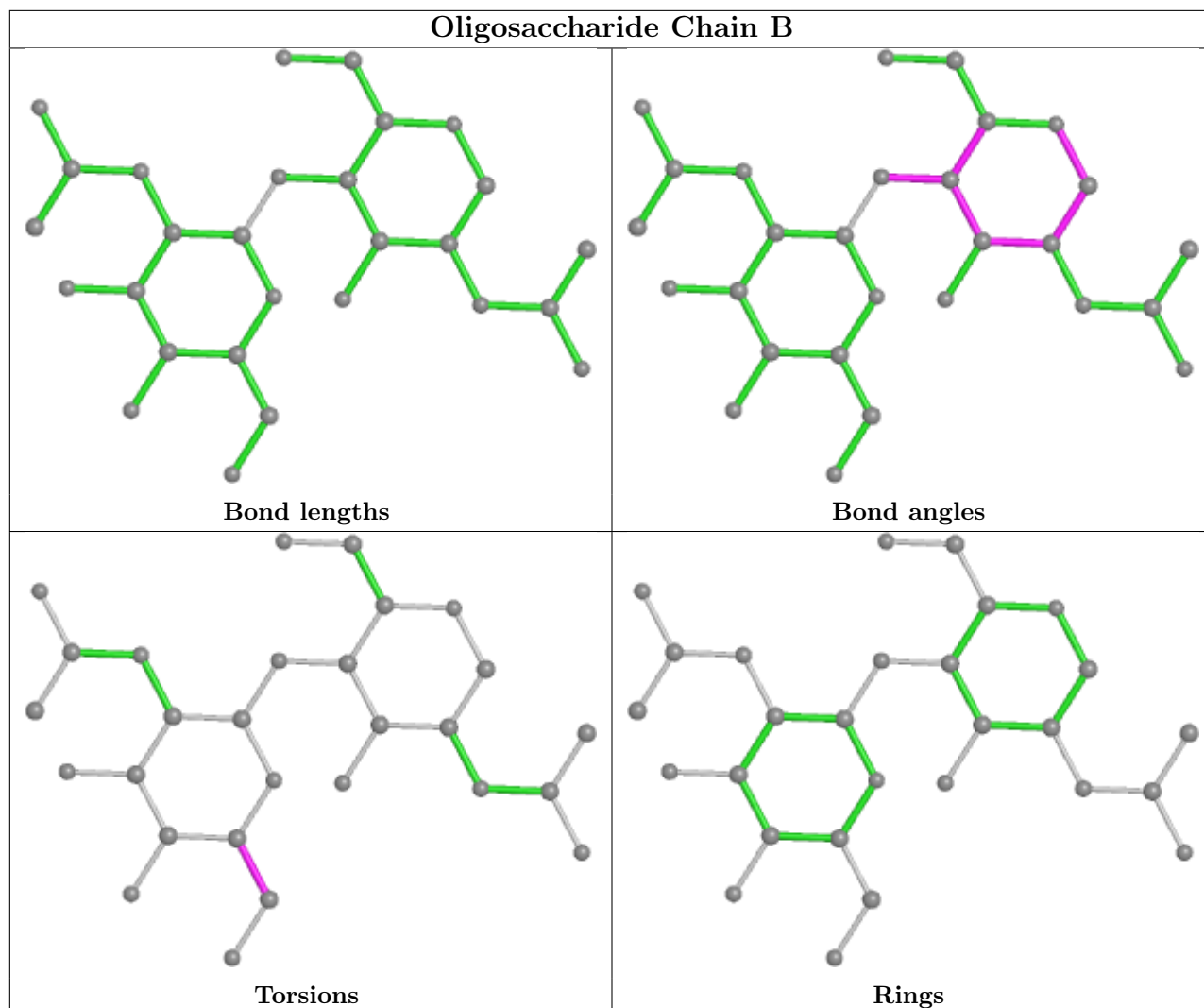
All (2) torsion outliers are listed below:

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 [i](#)

Of 20 ligands modelled in this entry, 7 are monoatomic - leaving 13 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
7	SCN	A	613	-	1,2,2	0.45	0	0,1,1	-	-
8	OSM	A	622	-	1,3,3	0.12	0	0,2,2	-	-
7	SCN	A	612	-	1,2,2	1.61	0	0,1,1	-	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	HEM	A	601	1,9	41,50,50	2.36	18 (43%)	45,82,82	2.81	21 (46%)
4	NAG	A	602	1	14,14,15	0.62	0	17,19,21	1.51	2 (11%)
4	NAG	A	606	1	14,14,15	0.79	0	17,19,21	1.34	1 (5%)
7	SCN	A	616	-	1,2,2	1.00	0	0,1,1	-	-
7	SCN	A	618	5	1,2,2	1.55	0	0,1,1	-	-
7	SCN	A	617	-	1,2,2	3.94	1 (100%)	0,1,1	-	-
7	SCN	A	615	-	1,2,2	0.38	0	0,1,1	-	-
7	SCN	A	614	-	1,2,2	0.43	0	0,1,1	-	-
7	SCN	A	619	-	1,2,2	0.71	0	0,1,1	-	-
4	NAG	A	605	1	14,14,15	0.51	0	17,19,21	0.71	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	OSM	A	622	-	-	0/0/1/1	-
3	HEM	A	601	1,9	-	4/12/54/54	-
4	NAG	A	602	1	-	1/6/23/26	0/1/1/1
4	NAG	A	606	1	-	2/6/23/26	0/1/1/1
4	NAG	A	605	1	-	0/6/23/26	0/1/1/1

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	601	HEM	C3D-C2D	5.02	1.47	1.36
3	A	601	HEM	CHB-C1B	4.89	1.47	1.35
3	A	601	HEM	CHA-C4D	4.34	1.46	1.35
3	A	601	HEM	C4D-ND	-4.15	1.33	1.40
3	A	601	HEM	C3C-C2C	4.01	1.45	1.40
7	A	617	SCN	C-N	3.94	1.28	1.15
3	A	601	HEM	C3B-C2B	3.88	1.45	1.37
3	A	601	HEM	CBB-CAB	3.71	1.48	1.30
3	A	601	HEM	C1B-NB	-3.59	1.34	1.40
3	A	601	HEM	C4B-NB	-2.98	1.32	1.38
3	A	601	HEM	CBC-CAC	2.81	1.47	1.29
3	A	601	HEM	C2A-C3A	2.80	1.46	1.37
3	A	601	HEM	FE-NB	2.73	2.10	1.96
3	A	601	HEM	C1A-NA	-2.26	1.31	1.36
3	A	601	HEM	FE-ND	2.15	2.07	1.96

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	601	HEM	C1D-ND	-2.13	1.34	1.38
3	A	601	HEM	CHD-C1D	2.10	1.47	1.41
3	A	601	HEM	CHC-C4B	2.03	1.46	1.41
3	A	601	HEM	O2A-CGA	-2.02	1.24	1.30

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	601	HEM	C3B-C2B-C1B	-8.14	100.45	106.49
3	A	601	HEM	C4A-C3A-C2A	-7.33	101.89	107.00
3	A	601	HEM	C2B-C1B-NB	5.17	115.97	109.84
3	A	601	HEM	CBA-CAA-C2A	-4.54	104.87	112.62
3	A	601	HEM	CMC-C2C-C3C	4.50	133.10	124.68
3	A	601	HEM	C1D-C2D-C3D	-4.03	102.72	106.96
3	A	601	HEM	C2D-C1D-ND	3.77	114.40	109.88
3	A	601	HEM	C4B-CHC-C1C	3.75	127.51	122.56
3	A	601	HEM	C3D-C4D-ND	3.46	114.02	110.17
4	A	606	NAG	O5-C5-C6	3.38	112.51	107.20
4	A	602	NAG	C1-C2-N2	-3.36	104.75	110.49
3	A	601	HEM	CMD-C2D-C1D	3.32	130.09	125.04
4	A	602	NAG	O5-C5-C6	3.29	112.36	107.20
3	A	601	HEM	C2C-C3C-C4C	-3.11	104.72	106.90
3	A	601	HEM	CBD-CAD-C3D	-2.98	104.34	112.63
3	A	601	HEM	CMA-C3A-C4A	2.85	132.84	128.46
3	A	601	HEM	CAD-C3D-C4D	2.69	129.35	124.66
3	A	601	HEM	CHD-C1D-C2D	-2.67	120.81	124.98
3	A	601	HEM	C4D-C3D-C2D	-2.65	103.03	106.90
3	A	601	HEM	O1D-CGD-CBD	-2.60	114.74	123.08
3	A	601	HEM	O2D-CGD-CBD	2.52	122.12	114.03
3	A	601	HEM	CHB-C1B-C2B	-2.38	120.13	126.72
3	A	601	HEM	CHA-C4D-C3D	-2.23	121.14	125.33
3	A	601	HEM	CMB-C2B-C1B	2.04	128.15	125.04

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	606	NAG	O5-C5-C6-O6
4	A	606	NAG	C4-C5-C6-O6
4	A	602	NAG	O5-C5-C6-O6
3	A	601	HEM	CAA-CBA-CGA-O1A
3	A	601	HEM	CAA-CBA-CGA-O2A

Continued on next page...

*Continued from previous page...*

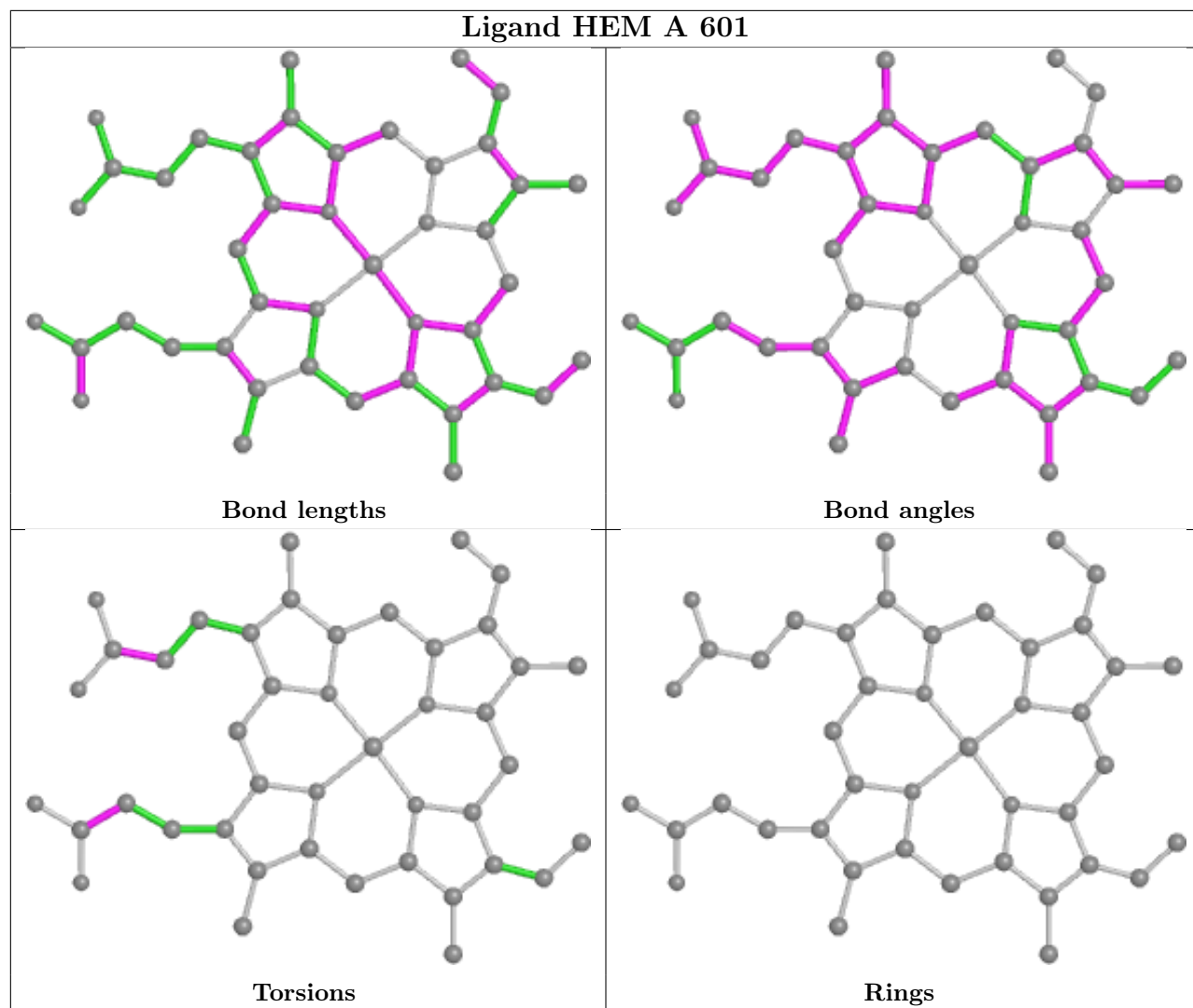
Mol	Chain	Res	Type	Atoms
3	A	601	HEM	CAD-CBD-CGD-O1D
3	A	601	HEM	CAD-CBD-CGD-O2D

There are no ring outliers.

5 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	601	HEM	12	0
7	A	616	SCN	1	0
7	A	618	SCN	1	0
7	A	614	SCN	2	0
7	A	619	SCN	3	0

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



## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	594/595 (99%)	0.64	53 (8%) <b>9</b> <b>11</b>	30, 52, 110, 178	0

All (53) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	7	GLY	22.1
1	A	121	SER	21.4
1	A	2	TRP	15.3
1	A	1	SER	14.6
1	A	11	PRO	10.6
1	A	119	LEU	9.9
1	A	120	GLY	9.6
1	A	172	TYR	9.3
1	A	174	SER	8.9
1	A	122	ASN	8.9
1	A	117	THR	8.6
1	A	171	PRO	7.9
1	A	118	GLU	7.0
1	A	170	PRO	6.2
1	A	169	THR	5.8
1	A	173	GLN	5.4
1	A	10	VAL	5.1
1	A	13	VAL	4.6
1	A	20	PRO	4.2
1	A	124	HIS	4.1
1	A	595	ASN	3.9
1	A	587	LEU	3.7
1	A	21	TYR	3.5
1	A	583	ASP	3.5
1	A	3	GLU	3.4
1	A	222	HIS	3.2
1	A	8	ALA	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	593	ARG	3.0
1	A	132	TYR	3.0
1	A	200	ALA	3.0
1	A	287	TRP	3.0
1	A	283	LEU	2.9
1	A	285	PRO	2.9
1	A	16	ASP	2.8
1	A	231	ASN	2.8
1	A	6	CYS	2.7
1	A	126	LYS	2.7
1	A	128	GLN	2.7
1	A	209	PRO	2.6
1	A	229	PHE	2.5
1	A	292	LEU	2.5
1	A	4	VAL	2.4
1	A	226	TYR	2.4
1	A	9	PRO	2.2
1	A	581	THR	2.2
1	A	538	GLU	2.2
1	A	220	TRP	2.2
1	A	585	LEU	2.2
1	A	254	PHE	2.1
1	A	286	HIS	2.1
1	A	127	THR	2.1
1	A	63	GLN	2.1
1	A	12	LEU	2.1

## 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	0.94	0.13	48,61,69,76	0

## 6.3 Carbohydrates [i](#)

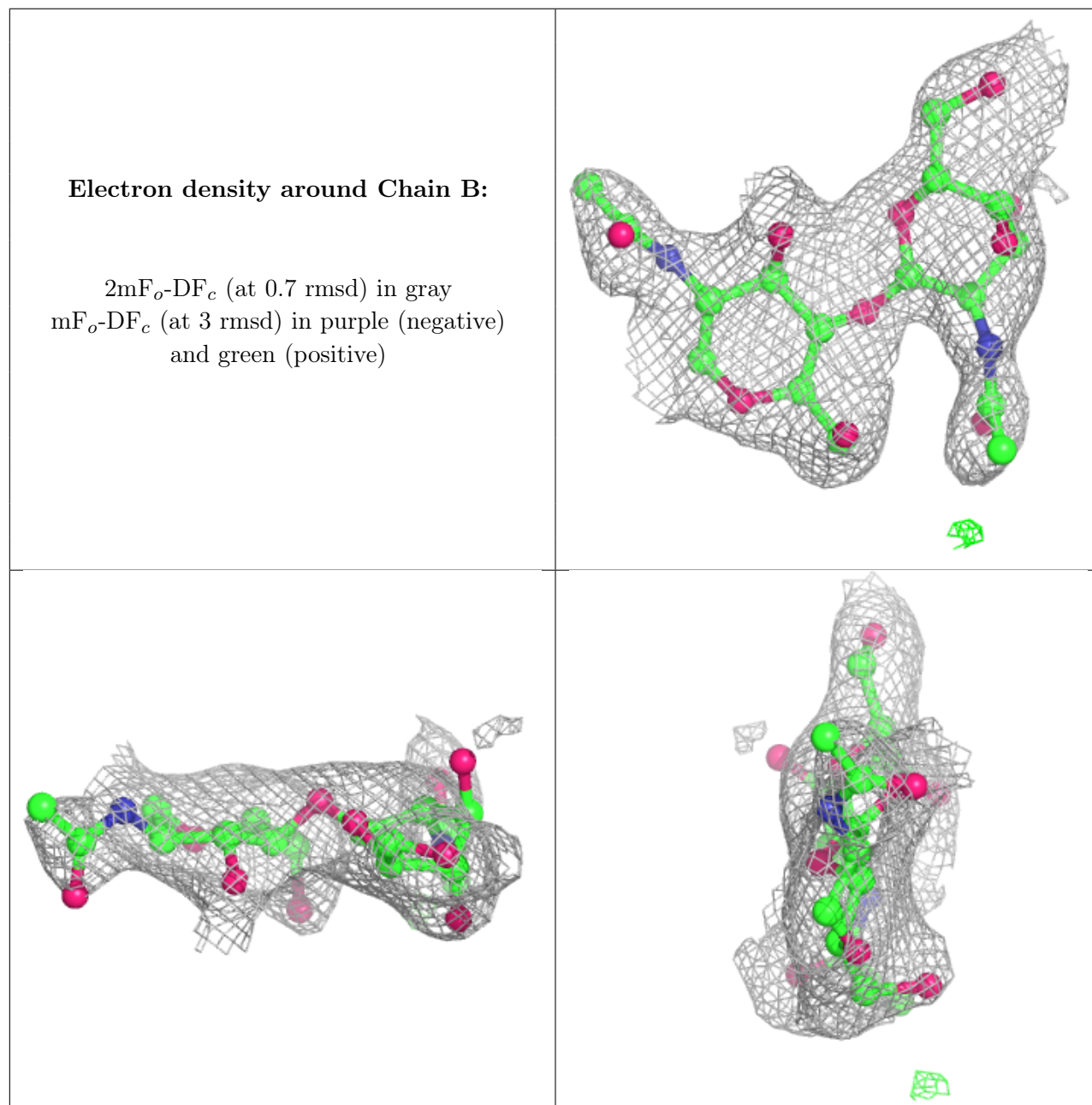
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,



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	0.20	73,83,88,93	0
2	NAG	B	1	14/15	0.86	0.16	64,75,81,81	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.

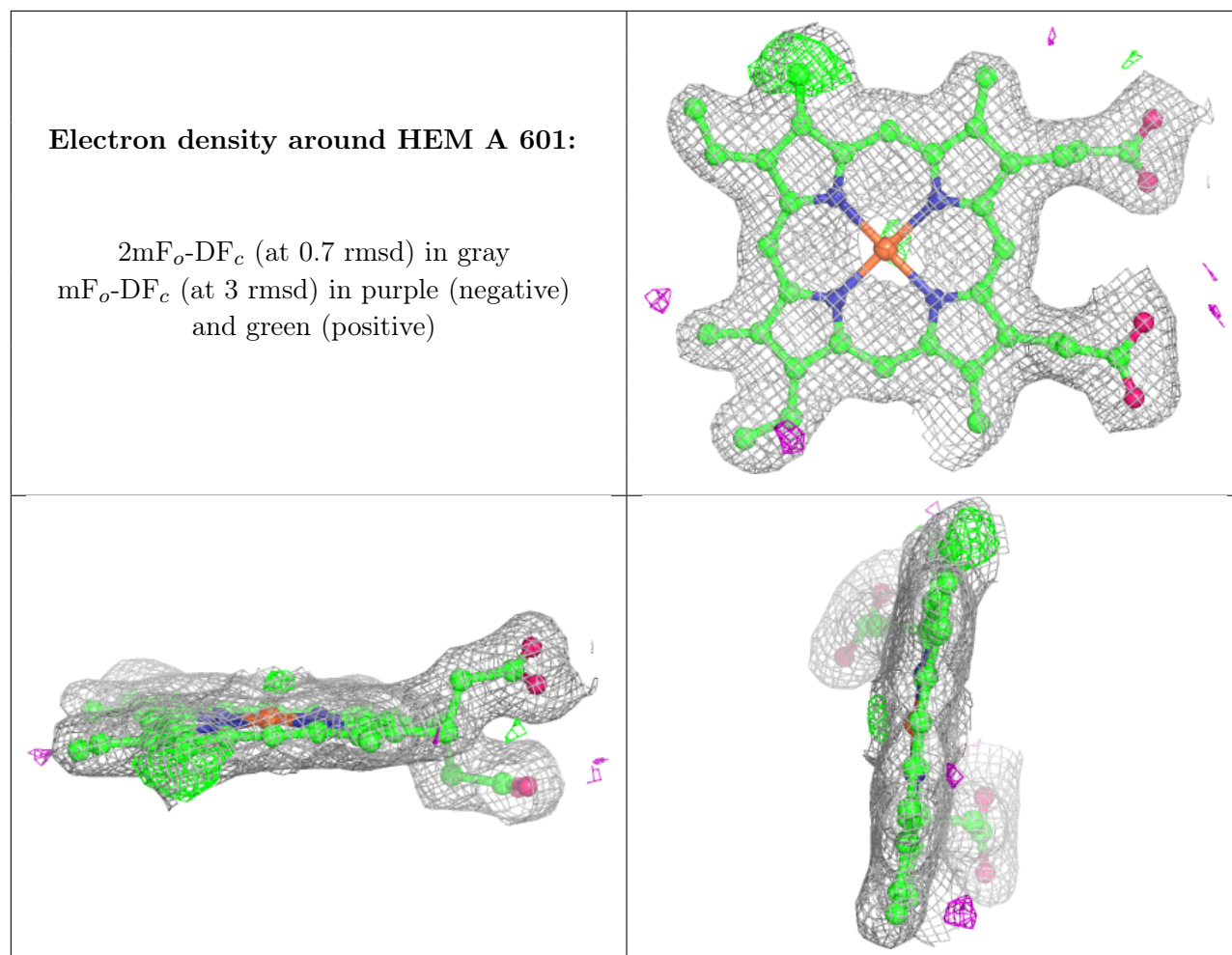


## 6.4 Ligands

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
4	NAG	A	606	14/15	0.71	0.20	66,80,100,112	0
4	NAG	A	602	14/15	0.73	0.16	73,80,82,83	0
7	SCN	A	615	3/3	0.75	0.22	40,40,40,69	3
8	OSM	A	622	4/4	0.88	0.10	36,43,47,54	0
4	NAG	A	605	14/15	0.90	0.12	51,63,68,75	0
6	CA	A	611	1/1	0.96	0.07	42,42,42,42	0
7	SCN	A	612	3/3	0.96	0.15	60,60,61,68	0
7	SCN	A	613	3/3	0.97	0.17	52,52,53,61	0
7	SCN	A	614	3/3	0.97	0.12	51,51,52,58	0
3	HEM	A	601	43/43	0.97	0.11	25,30,41,43	0
5	IOD	A	621	1/1	0.97	0.04	62,62,62,62	1
5	IOD	A	610	1/1	0.98	0.09	64,64,64,64	1
5	IOD	A	609	1/1	0.99	0.09	55,55,55,55	1
5	IOD	A	607	1/1	0.99	0.08	70,70,70,70	1
7	SCN	A	616	3/3	0.99	0.09	21,21,23,30	3
7	SCN	A	617	3/3	0.99	0.15	23,23,30,33	0
7	SCN	A	618	3/3	0.99	0.10	42,42,45,50	3
7	SCN	A	619	3/3	0.99	0.08	39,39,45,52	3
5	IOD	A	608	1/1	0.99	0.10	53,53,53,53	1
5	IOD	A	620	1/1	1.00	0.05	52,52,52,52	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.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.