

# wwPDB X-ray Structure Validation Summary Report (i)

#### Aug 20, 2020 – 02:05 PM BST

PDB ID	:	4GM7
Title	:	Structure of cinnamic acid bound bovine lactoperoxidase at 2.6A resolution.
Authors	:	Pandey, N.; Singh, A.K.; Sinha, M.; Kaur, P.; Sharma, S.; Singh, T.P.
Deposited on		
Resolution	:	2.60  Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 following versions of software and data (see references (1)) were used in the production of this report:

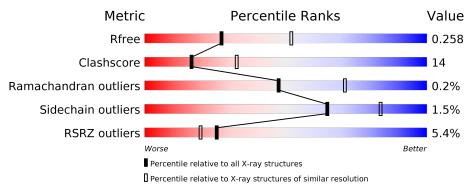
Xtriage (Phenix) : EDS : buster-report : Percentile statistics :	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : CCP4 : Ideal geometry (proteins) : Ideal geometry (DNA, RNA) : Validation Pipeline (wwPDB-VP) :	Engh & Huber (2001)

# 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 2.60 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries},{ m resolution\ range}({ m \AA}))$
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455(2.60-2.60)
Sidechain outliers	138945	3455(2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 >=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	А	595	5%	27% •				
2	В	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
11	HCI	А	622	-	-	Х	-



# 2 Entry composition (i)

There are 12 unique types of molecules in this entry. The entry contains 5078 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.

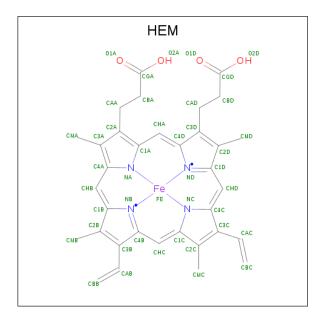
Mol	Chain	Residues		ŀ	Atom	5			ZeroOcc	AltConf	Trace
1	A	595	Total 4775	$\mathrm{C}$ 3037	N 847	0 864	Р 1	S 26	0	0	0

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



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

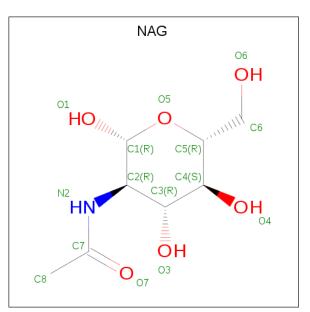
• Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: C<sub>34</sub>H<sub>32</sub>FeN<sub>4</sub>O<sub>4</sub>).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
3	А	1	Total 43	С 34	Fe 1	N 4	0 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
4	А	1	Total C N O 14 8 1 5	0	0
4	А	1	Total         C         N         O           14         8         1         5	0	0
4	А	1	Total C N O 14 8 1 5	0	0

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

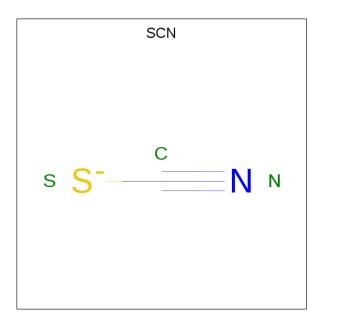
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	8	Total I 8 8	0	0

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

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

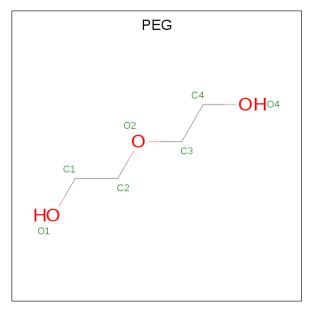
• Molecule 7 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	А	1	Total 3	С 1	N 1	${ m S}$ 1	0	0

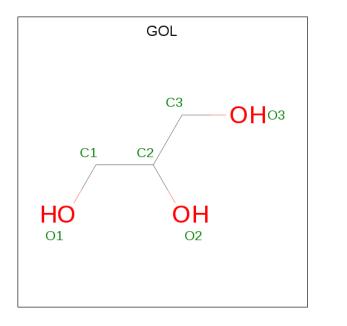
• Molecule 8 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	
8	А	1	Total 7	С 4	O 3	0	0

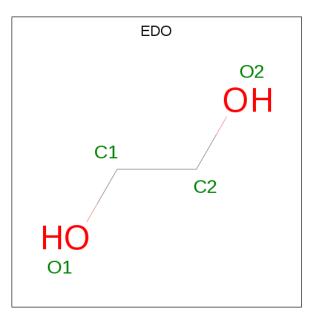
• Molecule 9 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).





Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	А	1	Total C 6 3	O 3	0	0

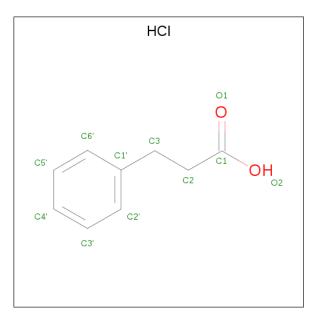
• Molecule 10 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	А	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 4 & 2 & 2 \end{array}$	0	0
10	А	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 4 & 2 & 2 \end{array}$	0	0
10	А	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 4 & 2 & 2 \end{array}$	0	0



• Molecule 11 is HYDROCINNAMIC ACID (three-letter code: HCI) (formula:  $C_9H_{10}O_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	А	1	Total         C         O           11         9         2	0	0

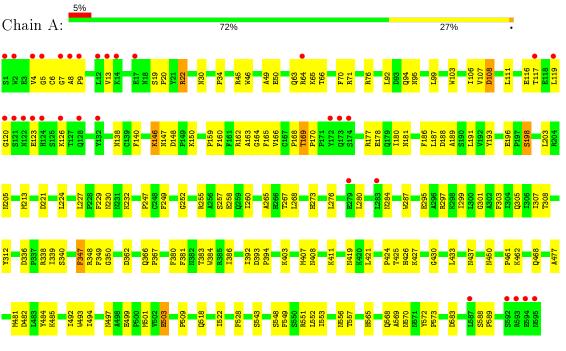
• Molecule 12 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	А	142	Total O 142 142	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-gluc opyranose

Chain B:

50%

50%



# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	54.05Å 80.23Å 76.04Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $103.30^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	43.99 - 2.60	Depositor
Resolution (A)	43.99 - 2.60	EDS
% Data completeness	97.8 (43.99-2.60)	Depositor
(in resolution range)	97.9(43.99-2.60)	EDS
R <sub>merge</sub>	(Not available)	Depositor
R <sub>sym</sub>	0.10	Depositor
$< I/\sigma(I) > 1$	$2.73$ (at $2.61\text{\AA}$ )	Xtriage
Refinement program	CNS 1.1	Depositor
B B.	0.221 , $0.235$	Depositor
$R, R_{free}$	0.218 , $0.258$	DCC
$R_{free}$ test set	983 reflections $(5.12\%)$	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.9	Xtriage
Anisotropy	0.901	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.27, 35.4	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	0.94	EDS
Total number of atoms	5078	wwPDB-VP
Average B, all atoms $(Å^2)$	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 5.98% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: GOL, SCN, NAG, SEP, CA, EDO, PEG, HEM, HCI, 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
	Cham	RMSZ	# Z  > 5	RMSZ	# Z  > 5
1	A	0.42	0/4892	0.67	0/6634

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	А	4775	0	4685	132	0
2	В	28	0	25	1	0
3	А	43	0	30	6	0
4	А	42	0	39	1	0
5	А	8	0	0	0	0
6	А	1	0	0	0	0
7	А	3	0	0	1	0
8	А	7	0	9	0	0
9	А	6	0	8	3	0
10	А	12	0	15	0	0
11	А	11	0	9	8	0
12	А	142	0	0	5	0
All	All	5078	0	4820	138	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.

The worst 5 of 138 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:168:PRO:HB2	1:A:169:THR:OG1	1.71	0.91
1:A:482:ASP:O	1:A:485:LYS:HE2	1.69	0.91
1:A:350:GLY:HA3	3:A:601:HEM:CBC	2.02	0.89
1:A:381:PHE:HZ	11:A:622:HCI:H21	1.38	0.87
1:A:518:GLN:O	1:A:522:ILE:HG12	1.77	0.84

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	А	592/595~(100%)	550~(93%)	41 (7%)	1 (0%)	47 71

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	169	THR

#### 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	А	517/517~(100%)	509~(98%)	8 (2%)	65 83

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	146	LYS
1	А	503	GLU
1	А	347	PHE
1	А	108	ASP
1	А	232	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	А	437	ASN
1	А	468	GLN
1	А	521	GLN
1	А	426	HIS
1	А	497	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).

	ſol	Type	Chain	Res Link		B	ond leng	gths	В	ond ang	gles
IV	101	rybe	Ullalli	nes	LIIIK	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
	1	SEP	А	198	1	$8,\!9,\!10$	0.81	0	8,12,14	1.50	1 (12%)

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	А	198	1	-	0/5/8/10	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	А	198	SEP	OG-CB-CA	3.10	111.16	108.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	$\mathbf{Res}$	Type	Clashes	Symm-Clashes
1	А	198	SEP	1	0

### 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	Tune	Chain	$\mathbf{Res}$	Link	Bo	ond leng	$\mathbf{ths}$	B	ond ang	les
INIOI	Type	Cham	nes		Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NAG	В	1	1,2	14, 14, 15	2.25	5 (35%)	17,19,21	1.98	6(35%)
2	NAG	В	2	2	14,14,15	1.05	1 (7%)	17,19,21	2.17	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	В	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	В	2	2	-	1/6/23/26	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	$\mathbf{Res}$	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
2	В	1	NAG	O4-C4	-5.96	1.29	1.43
2	В	1	NAG	O5-C5	-4.08	1.35	1.43
2	В	1	NAG	C4-C5	-2.69	1.47	1.53
2	В	1	NAG	C3-C2	-2.38	1.47	1.52
2	В	2	NAG	C3-C2	-2.27	1.47	1.52

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	В	2	NAG	C4-C3-C2	-6.17	101.97	111.02
2	В	1	NAG	O4-C4-C5	-4.03	99.29	109.30
2	В	2	NAG	C1-O5-C5	4.02	117.64	112.19
2	В	1	NAG	C3-C4-C5	3.58	116.62	110.24
2	В	1	NAG	C1-O5-C5	-3.12	107.97	112.19

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	В	2	NAG	C4-C5-C6-O6
2	В	1	NAG	C4-C5-C6-O6
2	В	1	NAG	O5-C5-C6-O6

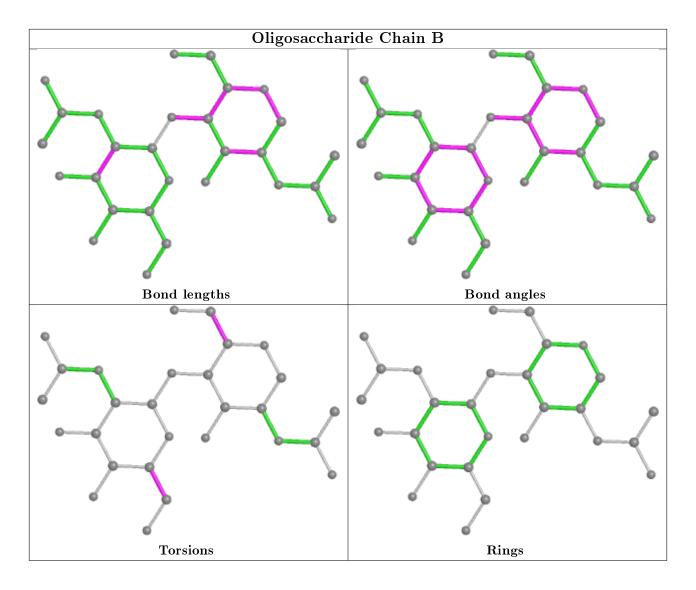
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	В	1	NAG	1	0

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





### 5.6 Ligand geometry (i)

Of 20 ligands modelled in this entry, 9 are monoatomic - leaving 11 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).

Mal	Mol Type	Chain	Res	Link	B	ond leng	gths	Bond angles		
	Type	Chain	nes		Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
11	HCI	А	622	-	8,11,11	0.14	0	$10,\!13,\!13$	0.77	0
4	NAG	А	602	1	14, 14, 15	1.17	2 (14%)	17,19,21	1.12	1(5%)
10	EDO	А	621	-	$^{3,3,3}$	1.67	1 (33%)	2,2,2	0.95	0



Mol	Turne	Chain	Res	Link	B	ond leng	gths	B	ond ang	les
MOI	Type	Unam	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z  > 2
3	HEM	А	601	1,12	27,50,50	2.36	12 (44%)	17,82,82	1.80	<mark>6 (35%)</mark>
4	NAG	А	606	1	14,14,15	1.10	2 (14%)	17,19,21	0.99	1(5%)
8	PEG	А	617	-	$^{6,6,6}$	1.29	1 (16%)	5, 5, 5	1.11	1 (20%)
10	EDO	А	620	-	$^{3,3,3}$	1.57	1 (33%)	2,2,2	0.88	0
9	GOL	А	618	-	$5,\!5,\!5$	0.75	0	5, 5, 5	0.43	0
7	SCN	А	616	-	$1,\!2,\!2$	3.54	1 (100%)	$0,\!1,\!1$	0.00	-
4	NAG	А	603	1	14,14,15	1.15	2 (14%)	17,19,21	1.04	1(5%)
10	EDO	А	619	-	$^{3,3,3}$	1.45	1 (33%)	2,2,2	0.65	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	$\mathbf{Res}$	$\mathbf{Link}$	Chirals	Torsions	Rings
11	HCI	А	622	-	-	3/3/5/5	0/1/1/1
4	NAG	А	602	1	-	0/6/23/26	0/1/1/1
10	EDO	A	621	-	-	1/1/1/1	-
3	HEM	А	601	1,12	-	0/6/54/54	-
4	NAG	А	606	1	-	0/6/23/26	0/1/1/1
8	PEG	А	617	-	-	2/4/4/4	-
10	EDO	А	620	-	-	1/1/1/1	-
9	GOL	А	618	-	-	1/4/4/4	-
4	NAG	А	603	1	-	0/6/23/26	0/1/1/1
10	EDO	А	619	_	-	1/1/1/1	-

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
3	А	601	HEM	C3B-CAB	5.20	1.58	1.47
3	А	601	HEM	C3C-CAC	5.17	1.58	1.47
3	А	601	HEM	CAD-C3D	3.61	1.58	1.52
7	А	616	SCN	C-N	3.54	1.27	1.15
3	А	601	HEM	C1A-CHA	-3.52	1.31	1.41

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
3	А	601	HEM	CMD-C2D-C1D	4.25	135.00	128.46
4	А	602	NAG	C4-C3-C2	-2.97	106.67	111.02

Continued on next page...



Mol	Chain	$\mathbf{Res}$	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
3	А	601	HEM	CMD-C2D-C3D	-2.91	119.46	124.94
4	А	603	NAG	O5-C1-C2	2.28	114.89	111.29
8	А	617	PEG	O2-C3-C4	2.26	120.02	110.07

Continued from previous page...

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
11	А	622	HCI	C1-C2-C3-C1'
8	А	617	PEG	O2-C3-C4-O4
10	А	621	EDO	O1-C1-C2-O2
10	А	620	EDO	O1-C1-C2-O2
8	А	617	PEG	O1-C1-C2-O2

There are no ring outliers.

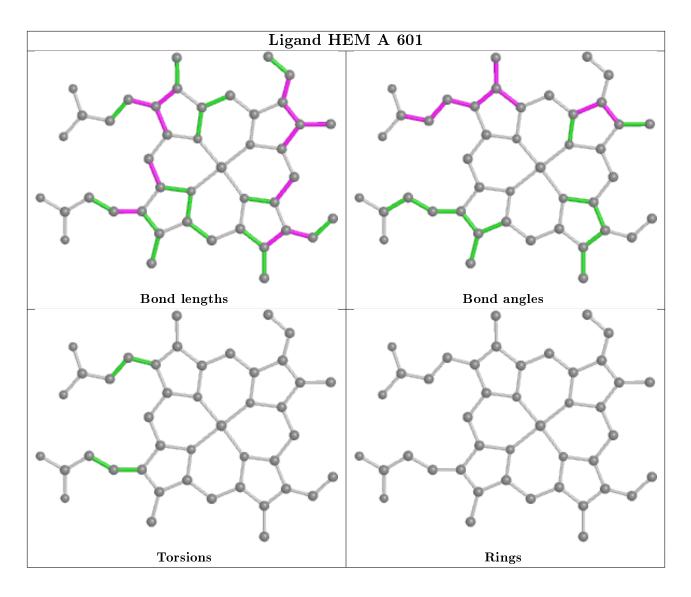
5 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	А	622	HCI	8	0
4	А	602	NAG	1	0
3	А	601	HEM	6	0
9	А	618	GOL	3	0
7	А	616	SCN	1	0

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







## 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,  $95^{th}$  percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ $>$	#RSRZ>	$\cdot 2$	$\mathbf{OWAB}(\mathbf{\AA}^2)$	$\mathbf{Q}{<}0.9$
1	А	594/595~(99%)	0.05	32 (5%) 25	20	28,48,93,163	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	173	GLN	11.2
1	А	2	TRP	10.4
1	А	7	GLY	9.2
1	А	1	SER	8.4
1	А	121	SER	7.7

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

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-factors}(\mathbf{\AA}^2)$	$Q{<}0.9$
1	SEP	А	198	10/11	0.81	0.27	$49,\!54,\!58,\!59$	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,  $95^{th}$  percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	$\mathbf{Res}$	Atoms	RSCC	RSR	$\mathbf{B} ext{-factors}(\mathbf{\AA}^2)$	Q < 0.9
2	NAG	В	2	14/15	0.80	0.31	$59,\!66,\!73,\!73$	0

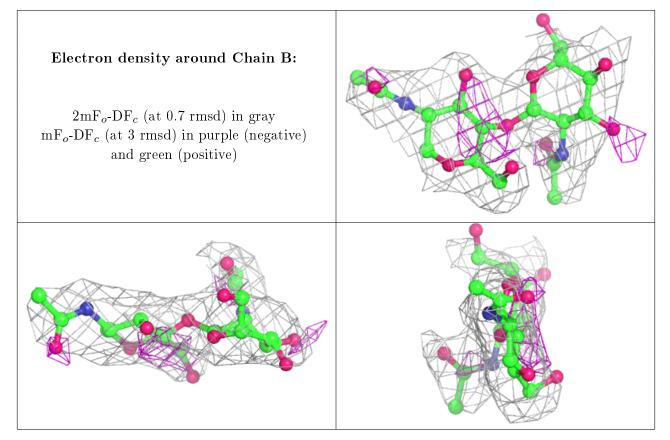
Continued on next page...



Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q<0.9
2	NAG	В	1	14/15	0.81	0.22	$43,\!47,\!54,\!55$	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 (i)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-factors}(\mathbf{\AA}^2)$	$Q{<}0.9$
4	NAG	А	602	14/15	0.76	0.34	$53,\!67,\!77,\!86$	0
4	NAG	А	606	14/15	0.79	0.27	$49,\!59,\!63,\!64$	0
10	EDO	А	619	4/4	0.79	0.29	$70,\!73,\!75,\!75$	0
8	PEG	А	617	7/7	0.82	0.31	$67,\!71,\!73,\!73$	0
11	HCI	А	622	11/11	0.87	0.29	$31,\!37,\!45,\!46$	0
4	NAG	А	603	14/15	0.88	0.17	$41,\!51,\!53,\!54$	0
9	GOL	А	618	6/6	0.90	0.15	$47,\!49,\!51,\!51$	0

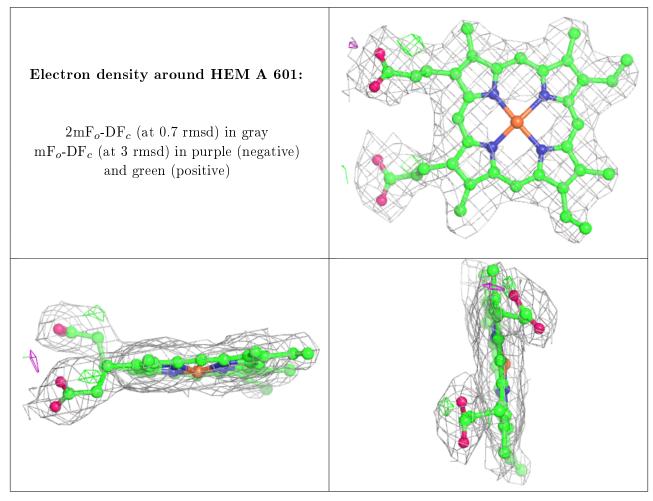
Continued on next page...



Mol	Type	Chain	Res	Atoms	RSCC	RSR	$B$ -factors( $Å^2$ )	Q<0.9
10	EDO	А	621	4/4	0.90	0.24	$61,\!61,\!62,\!62$	0
10	EDO	А	620	4/4	0.91	0.17	$48,\!50,\!51,\!52$	0
5	IOD	А	611	1/1	0.95	0.08	72,72,72,72	1
3	HEM	А	601	43/43	0.96	0.20	$25,\!30,\!34,\!37$	0
5	IOD	А	610	1/1	0.96	0.08	72,72,72,72	1
5	IOD	А	614	1/1	0.96	0.05	86,86,86,86	0
7	SCN	А	616	3/3	0.97	0.18	$57,\!57,\!57,\!59$	0
5	IOD	А	613	1/1	0.97	0.08	67,67,67,67	0
5	IOD	А	607	1/1	0.98	0.05	$90,\!90,\!90,\!90$	0
6	CA	А	615	1/1	0.98	0.17	40,40,40,40	0
5	IOD	А	612	1/1	0.99	0.11	49,49,49,49	0
5	IOD	А	609	1/1	0.99	0.06	73,73,73,73	0
5	IOD	А	608	1/1	1.00	0.05	79,79,79,79	0

Continued from previous page...

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.

