



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 06:23 pm GMT

PDB ID : 4MSF  
Title : Crystal structure of the complex of goat lactoperoxidase with 3-hydroxymethyl phenol at 1.98 Angstrom resolution  
Authors : Singh, A.; Singh, R.P.; Sinha, M.; Singh, A.K.; Bhushan, A.; Kaur, P.; Sharma, S.; Singh, T.P.  
Deposited on : 2013-09-18  
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

<http://wwpdb.org/validation/2016/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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

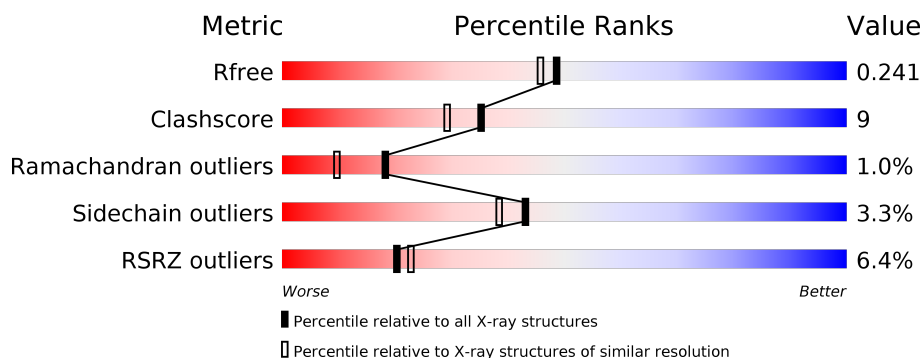
# 1 Overall quality at a glance

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))
$R_{free}$	100719	9293 (2.00-1.96)
Clashscore	112137	10621 (2.00-1.96)
Ramachandran outliers	110173	10502 (2.00-1.96)
Sidechain outliers	110143	10501 (2.00-1.96)
RSRZ outliers	101464	9395 (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. 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	<div> <div>6%</div> <div>86%</div> <div>13%</div> </div>

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
5	IOD	A	612	-	-	X	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	IOD	A	615	-	-	X	-
6	SCN	A	622	-	-	-	X
7	PEG	A	623	-	-	X	X
7	PEG	A	624	-	-	-	X
8	EDO	A	626	-	-	-	X
9	3HY	A	627	-	-	X	X

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 5414 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
1	A	595	Total	C	N	O	P	S	0	0	0
			4757	3021	844	865	1	26			

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



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

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

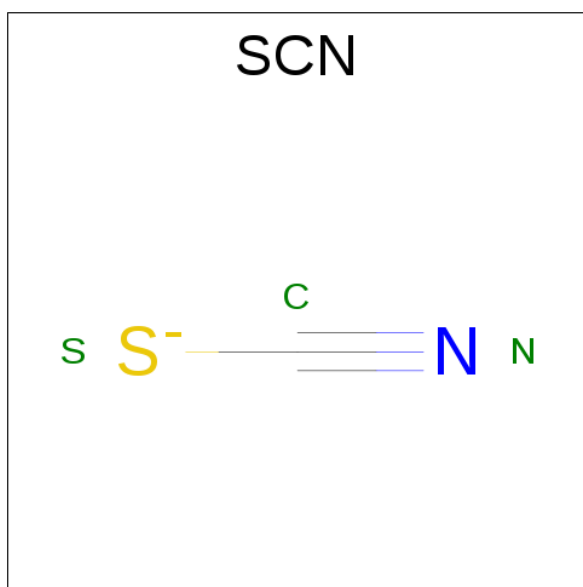


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			15	8	1	6		
4	A	1	Total	C	N	O	0	0
			15	8	1	6		
4	A	1	Total	C	N	O	0	0
			15	8	1	6		
4	A	1	Total	C	N	O	0	0
			15	8	1	6		

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	15	Total	I	0	0
			15	15		

- Molecule 6 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).



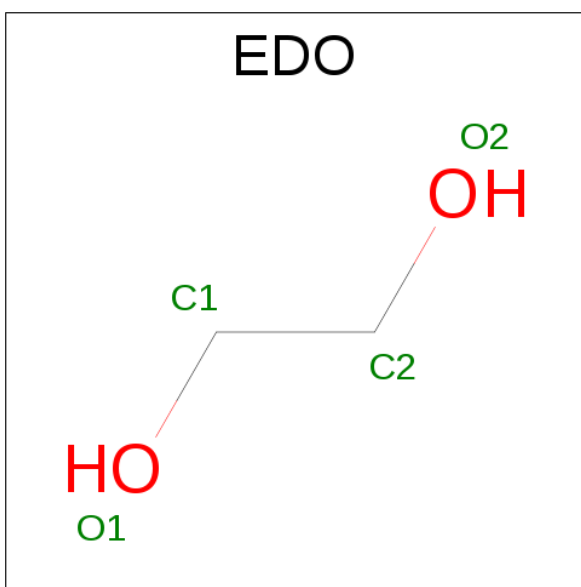
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	S	0	0
			3	1	1	1		

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



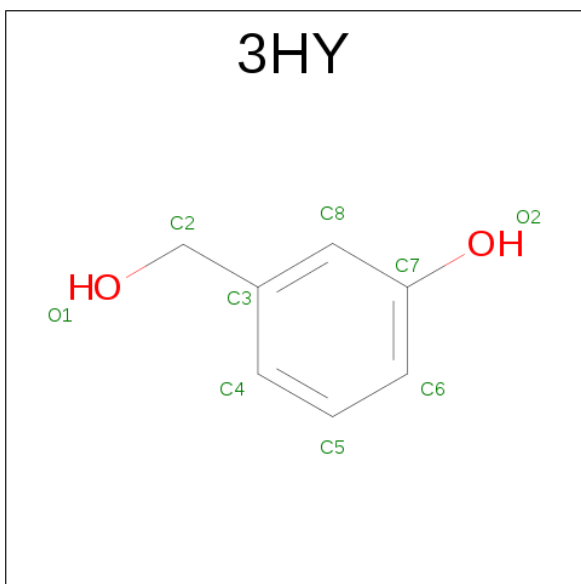
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			7	4	3		
7	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			4	2	2		
8	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 9 is 3-(HYDROXYMETHYL)PHENOL (three-letter code: 3HY) (formula:  $C_7H_8O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			9	7	2		

- Molecule 10 is water.

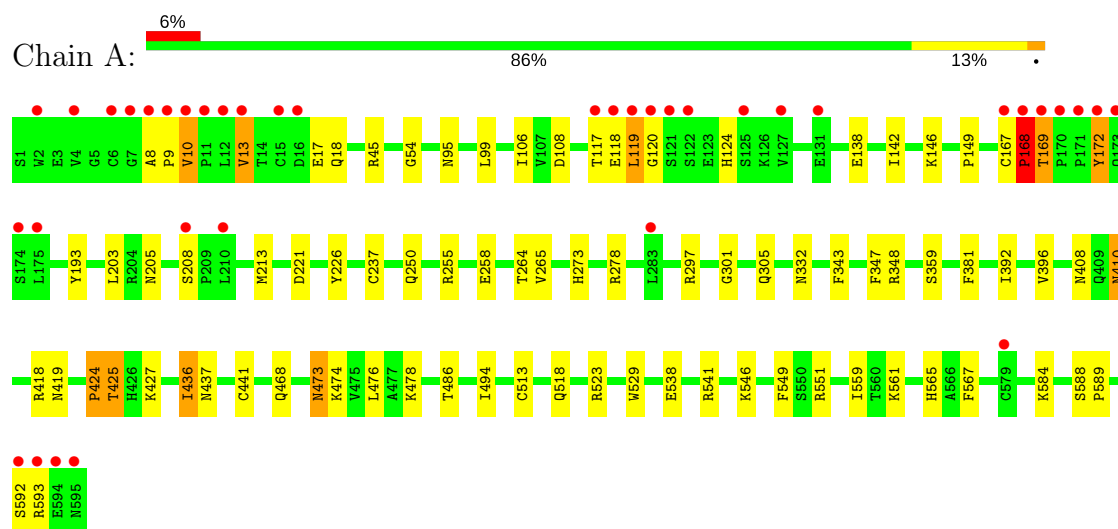
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	504	Total	O	0	0
			504	504		



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

These plots are drawn for all protein, RNA and DNA 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



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	53.77Å 80.36Å 75.76Å 90.00° 102.71° 90.00°	Depositor
Resolution (Å)	54.39 – 1.98 80.36 – 1.98	Depositor EDS
% Data completeness (in resolution range)	97.4 (54.39-1.98) 97.4 (80.36-1.98)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.91 (at 1.98Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.192 , 0.238 0.201 , 0.241	Depositor DCC
$R_{free}$ test set	2148 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	26.9	Xtriage
Anisotropy	0.117	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 51.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5414	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 5.54% 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: SCN, NAG, SEP, 3HY, CA, EDO, PEG, 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	0.57	0/4875	0.72	4/6621 (0.1%)

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

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	168	PRO	CB-CA-C	-7.53	93.17	112.00
1	A	424	PRO	CA-C-O	-5.72	106.46	120.20
1	A	486	THR	N-CA-CB	5.46	120.67	110.30
1	A	424	PRO	O-C-N	5.01	130.72	122.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	167	CYS	Peptide
1	A	168	PRO	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	4757	0	4648	78	0
2	A	43	0	30	11	0
3	A	1	0	0	0	0
4	A	60	0	59	14	0
5	A	15	0	0	8	0
6	A	3	0	0	0	0
7	A	14	0	20	7	0
8	A	8	0	12	1	0
9	A	9	0	7	7	0
10	A	504	0	0	11	0
All	All	5414	0	4776	90	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 (90) 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	2:A:601:HEM:CMD	1.74	1.35
1:A:258:GLU:OE2	2:A:601:HEM:HMB1	1.09	1.23
1:A:95:ASN:ND2	4:A:603:NAG:O1	1.72	1.20
1:A:258:GLU:OE2	2:A:601:HEM:CMB	1.90	1.19
1:A:95:ASN:ND2	4:A:603:NAG:C1	2.18	1.06
1:A:108:ASP:OD2	2:A:601:HEM:HMD1	0.88	1.05
1:A:205:ASN:CG	4:A:604:NAG:O1	1.97	1.03
5:A:614:IOD:I	10:A:723:HOH:O	2.53	0.97
5:A:615:IOD:I	10:A:807:HOH:O	2.54	0.94
1:A:332:ASN:ND2	4:A:606:NAG:O1	2.01	0.93
1:A:95:ASN:HD21	4:A:603:NAG:C1	1.86	0.86
1:A:205:ASN:ND2	4:A:604:NAG:O1	2.12	0.82
2:A:601:HEM:CMA	9:A:627:3HY:H1	2.10	0.80
1:A:332:ASN:HD21	4:A:606:NAG:C1	1.93	0.80
1:A:95:ASN:CG	4:A:603:NAG:O1	2.23	0.77
1:A:108:ASP:CG	2:A:601:HEM:HMD1	2.03	0.71
1:A:106:ILE:CD1	1:A:265:VAL:HG11	2.20	0.71

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:255:ARG:HG2	9:A:627:3HY:H4	1.74	0.70
5:A:616:IOD:I	10:A:1131:HOH:O	2.80	0.69
1:A:424:PRO:O	1:A:425:THR:HB	1.95	0.67
5:A:612:IOD:I	10:A:823:HOH:O	2.84	0.65
5:A:615:IOD:I	10:A:1003:HOH:O	2.83	0.64
1:A:538:GLU:HG3	1:A:541:ARG:NH2	2.13	0.64
1:A:106:ILE:HD13	1:A:265:VAL:HG11	1.79	0.63
1:A:551:ARG:HD3	1:A:584:LYS:HA	1.81	0.61
1:A:264:THR:HG23	1:A:392:ILE:HG23	1.83	0.60
1:A:332:ASN:CG	4:A:606:NAG:O1	2.40	0.59
1:A:142:ILE:HD12	1:A:436:ILE:HD12	1.85	0.59
2:A:601:HEM:HMA2	9:A:627:3HY:H1	1.85	0.58
1:A:106:ILE:HD11	1:A:265:VAL:HG11	1.86	0.57
1:A:95:ASN:HD21	4:A:603:NAG:H1	1.69	0.57
2:A:601:HEM:HMA1	9:A:627:3HY:H1	1.85	0.56
1:A:565:HIS:HB3	5:A:612:IOD:I	2.75	0.56
1:A:108:ASP:CG	2:A:601:HEM:CMD	2.68	0.56
1:A:142:ILE:CD1	1:A:436:ILE:HD12	2.35	0.56
1:A:213:MET:HG2	1:A:273:HIS:CD2	2.42	0.55
1:A:205:ASN:ND2	4:A:604:NAG:C1	2.70	0.55
1:A:203:LEU:HD23	1:A:250:GLN:NE2	2.22	0.54
1:A:95:ASN:ND2	4:A:603:NAG:O5	2.41	0.54
1:A:237:CYS:HA	1:A:381:PHE:O	2.08	0.53
1:A:9:PRO:O	1:A:10:VAL:C	2.47	0.53
1:A:119:LEU:HD12	1:A:138:GLU:HG3	1.91	0.53
2:A:601:HEM:HBC2	2:A:601:HEM:HMC1	1.91	0.52
1:A:523:ARG:HG3	1:A:529:TRP:CE2	2.45	0.51
1:A:348:ARG:HH11	1:A:437:ASN:ND2	2.08	0.51
8:A:626:EDO:H11	10:A:861:HOH:O	2.10	0.51
1:A:255:ARG:HG2	9:A:627:3HY:C8	2.40	0.51
1:A:348:ARG:HH11	1:A:437:ASN:HD22	1.58	0.51
1:A:567:PHE:HB2	5:A:612:IOD:I	2.81	0.50
1:A:419:ASN:HD21	7:A:623:PEG:H11	1.76	0.50
1:A:255:ARG:CG	9:A:627:3HY:H4	2.41	0.50
1:A:221:ASP:HB2	1:A:226:TYR:CZ	2.47	0.49
1:A:149:PRO:HG2	7:A:623:PEG:H22	1.96	0.48
1:A:54:GLY:HA2	10:A:943:HOH:O	2.14	0.47
7:A:624:PEG:H11	10:A:777:HOH:O	2.13	0.47
1:A:119:LEU:HG	1:A:169:THR:HG23	1.98	0.46
1:A:106:ILE:HD11	1:A:265:VAL:CG1	2.45	0.45
1:A:255:ARG:CB	9:A:627:3HY:H4	2.45	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:106:ILE:HD11	1:A:265:VAL:CB	2.47	0.45
1:A:8:ALA:HB3	1:A:9:PRO:HD3	1.98	0.45
1:A:119:LEU:HG	1:A:169:THR:CG2	2.46	0.45
1:A:468:GLN:HG2	1:A:474:LYS:HA	1.98	0.45
1:A:278:ARG:NH2	10:A:1005:HOH:O	2.50	0.45
2:A:601:HEM:HMB2	2:A:601:HEM:HBB2	1.99	0.44
1:A:203:LEU:CD2	1:A:250:GLN:NE2	2.80	0.44
1:A:45:ARG:HB3	7:A:624:PEG:H21	2.00	0.44
1:A:99:LEU:HD21	1:A:549:PHE:CD1	2.52	0.44
1:A:418:ARG:HH22	7:A:623:PEG:H12	1.82	0.44
1:A:473:ASN:ND2	1:A:476:LEU:H	2.15	0.44
1:A:343:PHE:CD1	1:A:518:GLN:HG2	2.53	0.43
1:A:408:ASN:OD1	1:A:410:ASN:ND2	2.51	0.43
1:A:588:SER:N	1:A:589:PRO:CD	2.82	0.43
1:A:332:ASN:CG	4:A:606:NAG:HO1	2.21	0.43
1:A:396:VAL:HB	1:A:559:ILE:HD11	2.01	0.42
1:A:538:GLU:HB2	10:A:1081:HOH:O	2.19	0.42
1:A:117:THR:HG21	1:A:138:GLU:CB	2.49	0.42
1:A:172:TYR:HD1	1:A:172:TYR:HA	1.71	0.42
1:A:494:ILE:HA	1:A:494:ILE:HD12	1.89	0.42
1:A:418:ARG:HH22	7:A:623:PEG:C1	2.32	0.42
1:A:301:GLY:O	1:A:305:GLN:HG3	2.20	0.42
1:A:118:GLU:O	1:A:120:GLY:N	2.53	0.42
1:A:193:TYR:CE2	1:A:297:ARG:HG3	2.55	0.42
1:A:8:ALA:N	1:A:9:PRO:CD	2.83	0.41
1:A:117:THR:HG21	1:A:138:GLU:HB3	2.00	0.41
1:A:18:GLN:HB2	10:A:831:HOH:O	2.20	0.41
1:A:45:ARG:CB	7:A:624:PEG:H21	2.50	0.41
1:A:142:ILE:HD12	1:A:436:ILE:CD1	2.51	0.41
1:A:359:SER:OG	5:A:615:IOD:I	2.99	0.40
1:A:205:ASN:OD1	4:A:604:NAG:O1	2.36	0.40
1:A:205:ASN:ND2	1:A:208:SER:OG	2.54	0.40

There are no symmetry-related clashes.

## 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	592/595 (100%)	563 (95%)	23 (4%)	6 (1%)	18 10

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	13	VAL
1	A	119	LEU
1	A	169	THR
1	A	425	THR
1	A	10	VAL
1	A	168	PRO

### 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	516/516 (100%)	499 (97%)	17 (3%)	43 38

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	13	VAL
1	A	17	GLU
1	A	124	HIS
1	A	146	LYS
1	A	172	TYR
1	A	347	PHE
1	A	410	ASN
1	A	427	LYS
1	A	436	ILE
1	A	441	CYS
1	A	473	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	478	LYS
1	A	513	CYS
1	A	546	LYS
1	A	561	LYS
1	A	592	SER
1	A	593	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	95	ASN
1	A	205	ASN
1	A	410	ASN
1	A	419	ASN
1	A	437	ASN
1	A	468	GLN
1	A	473	ASN
1	A	497	ASN
1	A	521	GLN
1	A	570	ASN

### 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	9,9,10	1.37	2 (22%)	9,12,14	6.37	4 (44%)



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	-	0/5/8/10	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	198	SEP	P-O1P	-2.69	1.41	1.50
1	A	198	SEP	P-O2P	-2.53	1.44	1.54

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	198	SEP	OG-P-O1P	-4.49	93.87	106.47
1	A	198	SEP	O2P-P-O1P	2.40	119.87	110.50
1	A	198	SEP	P-OG-CB	7.30	138.42	118.30
1	A	198	SEP	OG-CB-CA	16.54	124.48	108.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 27 ligands modelled in this entry, 16 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).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	HEM	A	601	1,10	28,50,50	1.26	3 (10%)	17,82,82	1.39	3 (17%)
4	NAG	A	603	-	15,15,15	0.45	0	21,21,21	2.57	2 (9%)
4	NAG	A	604	-	15,15,15	0.81	1 (6%)	21,21,21	1.59	3 (14%)
4	NAG	A	605	-	15,15,15	0.69	1 (6%)	21,21,21	1.95	4 (19%)
4	NAG	A	606	-	15,15,15	0.59	0	21,21,21	1.91	4 (19%)
6	SCN	A	622	-	1,2,2	0.38	0	0,1,1	0.00	-
7	PEG	A	623	-	6,6,6	0.33	0	5,5,5	0.40	0
7	PEG	A	624	-	6,6,6	0.51	0	5,5,5	0.38	0
8	EDO	A	625	-	3,3,3	0.41	0	2,2,2	0.23	0
8	EDO	A	626	-	3,3,3	0.37	0	2,2,2	0.34	0
9	3HY	A	627	-	9,9,9	1.83	2 (22%)	11,11,11	2.00	5 (45%)

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	HEM	A	601	1,10	-	0/6/54/54	0/0/8/8
4	NAG	A	603	-	-	0/6/26/26	0/1/1/1
4	NAG	A	604	-	-	0/6/26/26	0/1/1/1
4	NAG	A	605	-	-	0/6/26/26	0/1/1/1
4	NAG	A	606	-	-	0/6/26/26	0/1/1/1
6	SCN	A	622	-	-	0/0/0/0	0/0/0/0
7	PEG	A	623	-	-	0/4/4/4	0/0/0/0
7	PEG	A	624	-	-	0/4/4/4	0/0/0/0
8	EDO	A	625	-	-	0/1/1/1	0/0/0/0
8	EDO	A	626	-	-	0/1/1/1	0/0/0/0
9	3HY	A	627	-	-	0/2/2/2	0/1/1/1

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	A	627	3HY	C4-C3	-3.57	1.31	1.38
2	A	601	HEM	C1B-NB	-3.57	1.32	1.36
2	A	601	HEM	C4D-ND	-3.06	1.33	1.36
9	A	627	3HY	C5-C4	-2.98	1.33	1.38
4	A	604	NAG	C1-C2	-2.44	1.50	1.52
4	A	605	NAG	C1-C2	-2.35	1.50	1.52
2	A	601	HEM	C1C-NC	-2.26	1.34	1.36

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	603	NAG	C1-C2-N2	-9.49	99.74	110.73
4	A	605	NAG	C1-C2-N2	-6.68	102.98	110.73
4	A	604	NAG	C1-C2-N2	-5.08	104.84	110.73
4	A	606	NAG	O5-C1-C2	-4.23	105.27	109.52
4	A	606	NAG	C1-C2-C3	-3.49	105.79	110.54
4	A	605	NAG	C1-O5-C5	-3.39	107.28	113.39
4	A	604	NAG	O1-C1-C2	-2.86	103.27	109.22
9	A	627	3HY	C6-C7-C8	-2.85	117.04	120.20
2	A	601	HEM	CBA-CAA-C2A	-2.82	107.10	112.48
2	A	601	HEM	CBD-CAD-C3D	-2.78	107.16	112.47
4	A	605	NAG	O1-C1-C2	-2.49	104.05	109.22
4	A	604	NAG	C1-O5-C5	-2.35	109.16	113.39
4	A	605	NAG	O5-C5-C6	2.01	111.22	106.41
4	A	606	NAG	O5-C5-C4	2.27	113.84	109.66
9	A	627	3HY	C2-C3-C8	2.30	125.80	120.20
9	A	627	3HY	O1-C2-C3	2.31	121.14	112.08
9	A	627	3HY	C5-C4-C3	2.57	124.55	120.64
2	A	601	HEM	CMC-C2C-C3C	2.65	129.81	124.89
9	A	627	3HY	C7-C8-C3	3.51	122.55	120.41
4	A	606	NAG	C3-C4-C5	5.38	119.69	110.22
4	A	603	NAG	O5-C1-C2	6.14	115.69	109.52

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 37 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	HEM	11	0
4	A	603	NAG	6	0
4	A	604	NAG	4	0
4	A	606	NAG	4	0
7	A	623	PEG	4	0
7	A	624	PEG	3	0
8	A	626	EDO	1	0
9	A	627	3HY	7	0

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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%)	0.44	38 (6%) 20 23	13, 34, 82, 145	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	172	TYR	10.6
1	A	171	PRO	10.2
1	A	122	SER	8.5
1	A	121	SER	8.3
1	A	173	GLN	8.1
1	A	174	SER	7.1
1	A	11	PRO	6.9
1	A	170	PRO	6.6
1	A	7	GLY	6.4
1	A	12	LEU	6.3
1	A	120	GLY	6.3
1	A	10	VAL	6.0
1	A	119	LEU	5.9
1	A	2	TRP	4.9
1	A	167	CYS	4.8
1	A	169	THR	4.8
1	A	13	VAL	4.7
1	A	175	LEU	4.7
1	A	595	ASN	4.3
1	A	168	PRO	3.9
1	A	593	ARG	3.6
1	A	6	CYS	3.6
1	A	592	SER	3.4
1	A	118	GLU	3.3
1	A	16	ASP	3.3
1	A	283	LEU	3.1
1	A	9	PRO	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	208	SER	3.1
1	A	125	SER	3.0
1	A	8	ALA	2.9
1	A	127	VAL	2.9
1	A	594	GLU	2.8
1	A	210	LEU	2.5
1	A	117	THR	2.4
1	A	131	GLU	2.4
1	A	4	VAL	2.4
1	A	579	CYS	2.4
1	A	15	CYS	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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	SEP	A	198	10/11	0.86	0.19	-	43,46,48,49	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
9	3HY	A	627	9/9	0.79	0.24	8.31	19,33,40,44	0
7	PEG	A	623	7/7	0.93	0.20	5.54	24,29,30,34	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
8	EDO	A	626	4/4	0.81	0.20	5.37	38,43,46,48	0
6	SCN	A	622	3/3	0.81	0.19	4.78	56,56,61,62	0
7	PEG	A	624	7/7	0.79	0.22	2.86	28,32,35,37	0
5	IOD	A	607	1/1	1.00	0.12	2.24	18,18,18,18	0
4	NAG	A	605	15/15	0.86	0.14	0.09	31,37,41,45	0
2	HEM	A	601	43/43	0.97	0.11	0.05	13,13,16,20	0
4	NAG	A	604	15/15	0.83	0.15	-0.37	43,44,46,46	0
5	IOD	A	621	1/1	0.96	0.08	-1.38	46,46,46,46	1
5	IOD	A	618	1/1	0.97	0.11	-1.43	45,45,45,45	1
5	IOD	A	619	1/1	0.95	0.06	-1.57	56,56,56,56	1
5	IOD	A	611	1/1	0.97	0.05	-2.17	45,45,45,45	0
5	IOD	A	620	1/1	0.98	0.05	-2.20	51,51,51,51	1
5	IOD	A	610	1/1	0.97	0.07	-3.20	40,40,40,40	0
5	IOD	A	616	1/1	0.97	0.06	-3.62	45,45,45,45	0
3	CA	A	602	1/1	0.99	0.03	-5.48	12,12,12,12	0
5	IOD	A	615	1/1	0.97	0.06	-5.54	47,47,47,47	1
5	IOD	A	608	1/1	0.94	0.07	-	42,42,42,42	0
5	IOD	A	609	1/1	0.87	0.06	-	55,55,55,55	1
5	IOD	A	612	1/1	0.91	0.07	-	57,57,57,57	1
5	IOD	A	617	1/1	0.95	0.06	-	41,41,41,41	1
4	NAG	A	606	15/15	0.75	0.26	-	49,55,59,61	14
4	NAG	A	603	15/15	0.84	0.22	-	51,54,59,60	0
5	IOD	A	613	1/1	0.96	0.06	-	65,65,65,65	1
5	IOD	A	614	1/1	0.94	0.14	-	66,66,66,66	1
8	EDO	A	625	4/4	0.89	0.20	-	32,34,38,42	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.