



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2014 – 05:09 AM GMT

PDB ID : 3R5O  
Title : Crystal structure of the complex of bovine lactoperoxidase with 4-allyl-2-methoxyphenol at 2.6 Å resolution  
Authors : Pandey, N; Singh, A.K.; Singh, R.P.; Sinha, M.; Kaur, P.; Sharma, S.; Singh, T.P.  
Deposited on : 2011-03-19  
Resolution : 2.60 Å (reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

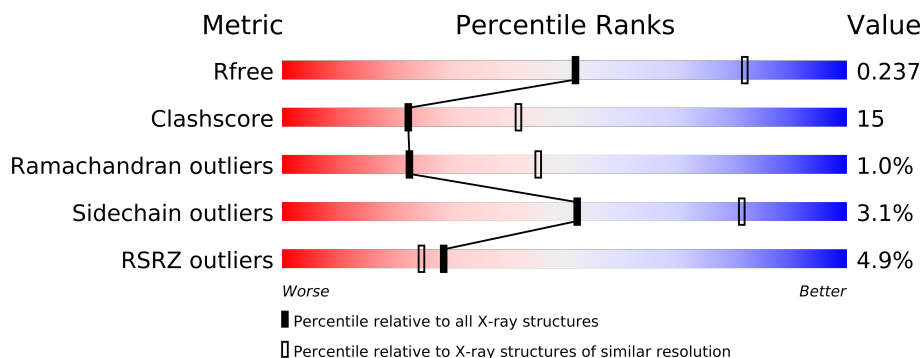
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	1718 (2.60-2.60)
Clashscore	79885	2154 (2.60-2.60)
Ramachandran outliers	78287	2113 (2.60-2.60)
Sidechain outliers	78261	2113 (2.60-2.60)
RSRZ outliers	66119	1718 (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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	595	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
11	GOL	A	610	-	X
4	NAG	A	604	-	X
7	EUG	A	597	-	X
9	SCN	A	700	-	X

## 2 Entry composition i

There are 12 unique types of molecules in this entry. The entry contains 5085 atoms, of which 0 are hydrogen and 0 are deuterium.

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.

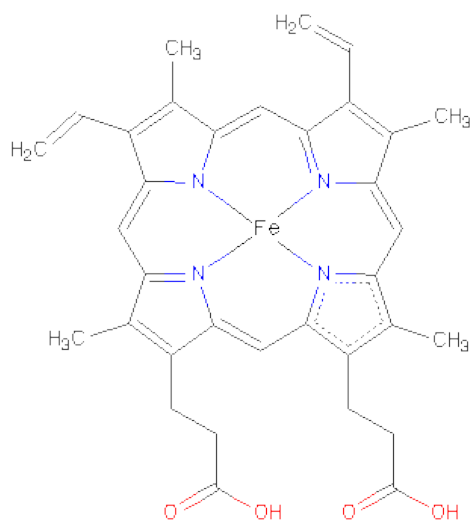
- Molecule 1 is a protein called Lactoperoxidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	595	Total	C	N	O	P	S	0	0	0
			4775	3037	847	864	1	26			

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

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

- 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
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		

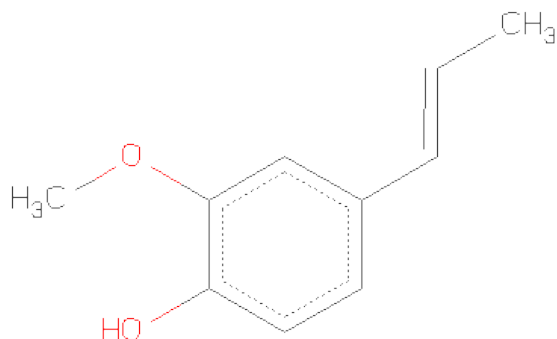
- Molecule 5 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 6 is IODIDE ION (three-letter code: IOD) (formula: I).

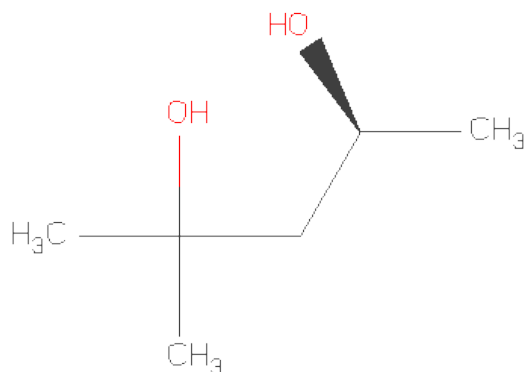
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	10	Total	I	0	0
			10	10		

- Molecule 7 is 2-METHOXY-4-[(1E)-PROP-1-EN-1-YL]PHENOL (three-letter code: EUG) (formula:  $C_{10}H_{12}O_2$ ).



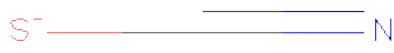
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			11	9	2		

- Molecule 8 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).



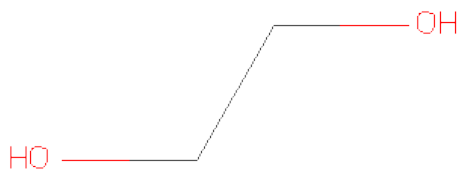
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			8	6	2		

- Molecule 9 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).



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

- Molecule 10 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

- Molecule 11 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

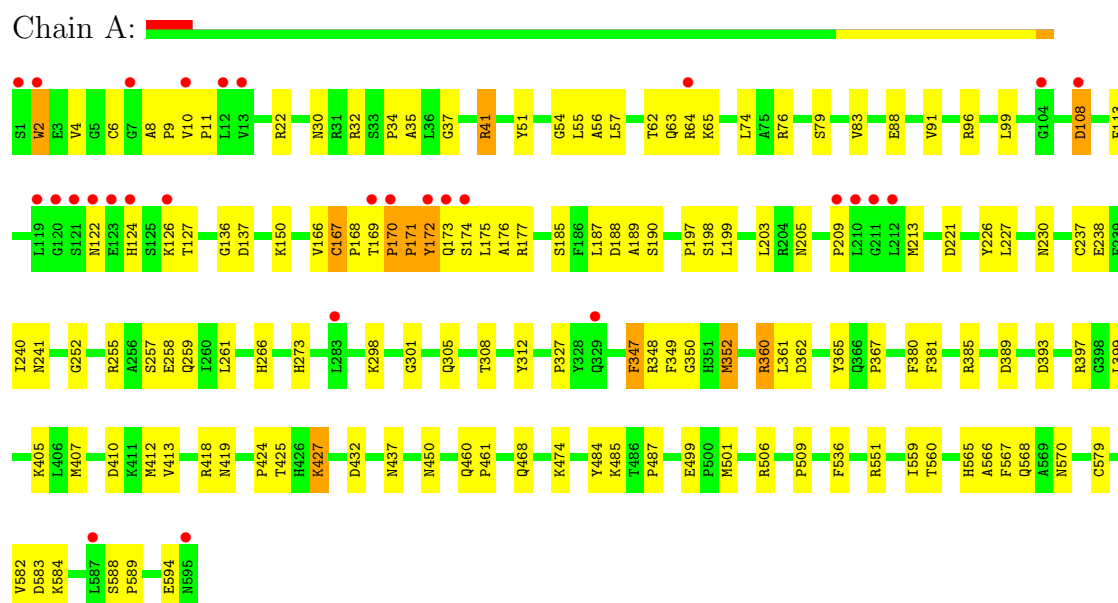
- Molecule 12 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	150	Total	O	0	0
			150	150		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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, $\alpha$ , $\beta$ , $\gamma$	54.01Å 80.05Å 75.68Å 90.00° 102.98° 90.00°	Depositor
Resolution (Å)	24.00 – 2.60 43.98 – 2.59	Depositor EDS
% Data completeness (in resolution range)	99.2 (24.00-2.60) 98.6 (43.98-2.59)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.09 (at 2.58Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.219 , 0.245 0.218 , 0.237	Depositor DCC
$R_{free}$ test set	944 reflections (5.09%)	DCC
Wilson B-factor (Å <sup>2</sup> )	49.1	Xtriage
Anisotropy	0.979	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 27.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 19486 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5085	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.71% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

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

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	0.46	0/4892	0.68	2/6634 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	108	ASP	CB-CG-OD2	5.99	123.69	118.30
1	A	6	CYS	N-CA-C	-5.68	95.67	111.00

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4775	0	4687	141	0
2	A	1	0	0	0	0
3	A	43	0	30	5	0
4	A	42	0	39	0	0
5	A	28	0	25	1	0
6	A	10	0	0	4	0
7	A	11	0	8	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	A	8	0	13	1	0
9	A	3	0	0	0	0
10	A	8	0	12	0	0
11	A	6	0	8	0	0
12	A	150	0	0	10	0
All	All	5085	0	4822	145	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 15.

All (145) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:170:PRO:CG	1:A:171:PRO:HD3	1.82	1.10
1:A:150:LYS:HE2	1:A:419:ASN:HD22	1.17	1.10
1:A:170:PRO:CD	1:A:171:PRO:HD3	1.88	1.02
1:A:168:PRO:HG3	1:A:172:TYR:CG	1.97	1.00
1:A:168:PRO:HG3	1:A:172:TYR:HB2	1.44	0.98
1:A:170:PRO:HG2	1:A:171:PRO:HD3	1.45	0.97
1:A:168:PRO:HG3	1:A:172:TYR:CB	2.00	0.92
1:A:76:ARG:HH22	1:A:419:ASN:HD21	1.18	0.90
1:A:360:ARG:O	1:A:361:LEU:HD23	1.71	0.90
1:A:170:PRO:HD2	1:A:171:PRO:HD3	1.55	0.89
1:A:150:LYS:HE2	1:A:419:ASN:ND2	1.92	0.84
1:A:168:PRO:HG3	1:A:172:TYR:CD1	2.12	0.84
1:A:197:PRO:HD2	1:A:198:SEP:O3P	1.78	0.83
1:A:167:CYS:HB2	1:A:168:PRO:HD2	1.60	0.82
1:A:150:LYS:CE	1:A:419:ASN:HD22	1.94	0.80
1:A:360:ARG:C	1:A:361:LEU:HD23	2.02	0.79
1:A:167:CYS:CB	1:A:168:PRO:CD	2.62	0.77
1:A:170:PRO:HD2	1:A:171:PRO:CD	2.13	0.77
1:A:425:THR:O	1:A:425:THR:HG22	1.84	0.76
1:A:91:VAL:HG22	1:A:405:LYS:HG3	1.70	0.74
1:A:551:ARG:HD3	1:A:583:ASP:O	1.91	0.71
1:A:348:ARG:HH11	1:A:437:ASN:ND2	1.88	0.71
1:A:170:PRO:CD	1:A:171:PRO:CD	2.67	0.70
1:A:2:TRP:HB3	1:A:175:LEU:CD1	2.22	0.70
1:A:76:ARG:HH22	1:A:419:ASN:ND2	1.90	0.68
1:A:51:TYR:HD2	1:A:57:LEU:O	1.76	0.68
1:A:167:CYS:HB2	1:A:168:PRO:CD	2.24	0.67
1:A:167:CYS:CB	1:A:168:PRO:HD2	2.26	0.66
1:A:168:PRO:HB2	1:A:171:PRO:O	1.97	0.64
1:A:168:PRO:CB	1:A:171:PRO:O	2.45	0.64

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:168:PRO:CG	1:A:172:TYR:HB2	2.21	0.64
1:A:4:VAL:O	1:A:4:VAL:HG23	1.97	0.63
1:A:348:ARG:HH11	1:A:437:ASN:HD22	1.47	0.62
1:A:205:ASN:HA	12:A:719:HOH:O	1.99	0.62
1:A:227:LEU:HD11	1:A:266:HIS:HB3	1.83	0.61
1:A:168:PRO:CG	1:A:172:TYR:CD1	2.81	0.61
1:A:551:ARG:HD3	1:A:584:LYS:HA	1.84	0.60
1:A:83:VAL:HG12	1:A:413:VAL:HB	1.83	0.60
1:A:172:TYR:O	1:A:173:GLN:HG3	2.03	0.59
1:A:450:ASN:ND2	1:A:487:PRO:HB2	2.17	0.59
1:A:99:LEU:HD23	1:A:566:ALA:HB1	1.85	0.58
1:A:51:TYR:CD2	1:A:57:LEU:O	2.55	0.58
1:A:2:TRP:CD1	1:A:175:LEU:HD11	2.38	0.58
1:A:169:THR:HB	1:A:170:PRO:HD3	1.86	0.57
1:A:30:ASN:O	1:A:34:PRO:HA	2.03	0.57
1:A:237:CYS:HA	1:A:381:PHE:O	2.04	0.57
1:A:385:ARG:O	1:A:389:ASP:HB3	2.04	0.57
1:A:259:GLN:OE1	1:A:261:LEU:HB2	2.04	0.57
1:A:425:THR:CG2	1:A:425:THR:O	2.53	0.56
1:A:63:GLN:OE1	1:A:63:GLN:N	2.30	0.56
3:A:605:HEM:HMC2	3:A:605:HEM:HBC2	1.88	0.56
1:A:407:MET:HB3	1:A:501:MET:CE	2.36	0.55
1:A:450:ASN:HD21	1:A:487:PRO:HB2	1.72	0.55
1:A:301:GLY:O	1:A:305:GLN:HG3	2.07	0.55
1:A:427:LYS:H	1:A:427:LYS:HD3	1.72	0.54
1:A:32:ARG:HD3	12:A:742:HOH:O	2.08	0.54
1:A:230:ASN:OD1	1:A:230:ASN:C	2.46	0.54
1:A:461:PRO:O	1:A:487:PRO:HG2	2.07	0.54
1:A:255:ARG:HB3	7:A:597:EUG:C5	2.38	0.53
1:A:32:ARG:HD2	12:A:675:HOH:O	2.07	0.53
1:A:170:PRO:HG2	1:A:171:PRO:CD	2.31	0.53
1:A:468:GLN:HG2	1:A:474:LYS:HA	1.89	0.53
1:A:168:PRO:CG	1:A:172:TYR:CG	2.85	0.53
3:A:605:HEM:HBB2	3:A:605:HEM:HMB2	1.91	0.52
1:A:2:TRP:HD1	1:A:175:LEU:HD11	1.75	0.52
1:A:484:TYR:O	1:A:485:LYS:HB2	2.09	0.52
1:A:188:ASP:O	1:A:189:ALA:HB3	2.11	0.51
1:A:2:TRP:HB3	1:A:175:LEU:HD12	1.93	0.51
1:A:582:VAL:HA	12:A:768:HOH:O	2.09	0.51
1:A:361:LEU:O	1:A:397:ARG:HD2	2.11	0.51
1:A:177:ARG:HA	12:A:634:HOH:O	2.10	0.51
1:A:352:MET:HB3	1:A:407:MET:HG2	1.93	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:91:VAL:O	1:A:91:VAL:HG22	2.11	0.50
1:A:124:HIS:CD2	12:A:749:HOH:O	2.64	0.50
1:A:74:LEU:HB3	12:A:722:HOH:O	2.12	0.50
1:A:221:ASP:HB2	1:A:226:TYR:CZ	2.47	0.49
1:A:568:GLN:HE21	1:A:570:ASN:HD21	1.61	0.49
1:A:551:ARG:CD	1:A:583:ASP:O	2.60	0.49
1:A:240:ILE:HG13	1:A:241:ASN:H	1.77	0.49
1:A:559:ILE:HA	6:A:612:IOD:I	2.83	0.49
1:A:175:LEU:HD23	1:A:176:ALA:O	2.12	0.49
1:A:352:MET:CB	1:A:407:MET:HG2	2.43	0.49
1:A:393:ASP:HB2	12:A:686:HOH:O	2.12	0.49
3:A:605:HEM:HBB2	3:A:605:HEM:CMB	2.42	0.49
1:A:62:THR:HG21	1:A:65:LYS:HG3	1.95	0.48
1:A:175:LEU:HD23	1:A:176:ALA:N	2.28	0.48
1:A:258:GLU:HB2	7:A:597:EUG:H6	1.97	0.47
1:A:362:ASP:O	1:A:365:TYR:N	2.40	0.47
1:A:560:THR:HA	1:A:579:CYS:SG	2.54	0.47
1:A:168:PRO:CG	1:A:171:PRO:O	2.63	0.47
1:A:484:TYR:O	1:A:485:LYS:CB	2.62	0.47
1:A:2:TRP:HD1	1:A:175:LEU:CD1	2.27	0.47
1:A:427:LYS:H	1:A:427:LYS:CD	2.26	0.47
1:A:565:HIS:HB3	6:A:614:IOD:I	2.84	0.47
1:A:199:LEU:HB2	8:A:598:MPD:HM1	1.97	0.47
1:A:367:PRO:HG3	6:A:621:IOD:I	2.85	0.47
1:A:113:PHE:HE1	7:A:597:EUG:H93	1.79	0.46
1:A:10:VAL:HG11	1:A:41:ARG:HE	1.81	0.46
1:A:169:THR:HG23	12:A:758:HOH:O	2.15	0.45
1:A:54:GLY:C	1:A:55:LEU:HD12	2.37	0.45
1:A:79:SER:HA	1:A:83:VAL:HG23	1.98	0.45
1:A:167:CYS:HB3	1:A:168:PRO:HD3	1.98	0.45
1:A:238:GLU:O	1:A:241:ASN:O	2.35	0.45
1:A:167:CYS:CB	1:A:168:PRO:HD3	2.47	0.44
1:A:99:LEU:HG	1:A:567:PHE:CE1	2.52	0.44
1:A:136:GLY:O	1:A:137:ASP:C	2.53	0.44
1:A:308:THR:HA	1:A:312:TYR:HB3	2.00	0.44
1:A:240:ILE:HG13	1:A:241:ASN:N	2.33	0.44
1:A:88:GLU:O	1:A:91:VAL:HG12	2.18	0.44
1:A:348:ARG:NH2	3:A:605:HEM:HAD1	2.32	0.43
1:A:258:GLU:O	1:A:380:PHE:HA	2.18	0.43
1:A:559:ILE:HG23	6:A:612:IOD:I	2.88	0.43
1:A:173:GLN:HB3	1:A:174:SER:HA	2.01	0.43
1:A:203:LEU:HD11	1:A:252:GLY:HA2	2.00	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:113:PHE:CE1	7:A:597:EUG:H93	2.52	0.43
1:A:418:ARG:O	1:A:432:ASP:HA	2.18	0.43
1:A:213:MET:HG2	1:A:273:HIS:CD2	2.54	0.43
1:A:2:TRP:HB3	1:A:175:LEU:HD11	2.00	0.42
1:A:79:SER:HA	1:A:83:VAL:CG2	2.49	0.42
1:A:257:SER:O	1:A:381:PHE:HA	2.19	0.42
1:A:350:GLY:HA3	3:A:605:HEM:CBC	2.49	0.42
1:A:424:PRO:O	1:A:425:THR:HB	2.19	0.42
1:A:261:LEU:HD22	1:A:399:LEU:HG	2.02	0.42
1:A:568:GLN:HE21	1:A:570:ASN:ND2	2.18	0.42
1:A:37:GLY:HA2	1:A:185:SER:HB2	2.02	0.42
1:A:8:ALA:N	1:A:9:PRO:CD	2.82	0.42
1:A:588:SER:N	1:A:589:PRO:CD	2.83	0.41
1:A:187:LEU:HD23	1:A:305:GLN:HA	2.02	0.41
1:A:96:ARG:HG3	1:A:506:ARG:HE	1.85	0.41
1:A:499:GLU:OE2	1:A:509:PRO:HD2	2.19	0.41
5:A:601:NAG:H61	5:A:602:NAG:N2	2.35	0.41
1:A:56:ALA:HB1	1:A:177:ARG:NE	2.35	0.41
1:A:298:LYS:HG2	1:A:536:PHE:CZ	2.55	0.41
1:A:62:THR:CG2	1:A:65:LYS:HG3	2.49	0.41
1:A:347:PHE:C	1:A:349:PHE:H	2.24	0.41
1:A:63:GLN:HA	12:A:743:HOH:O	2.20	0.41
1:A:166:VAL:O	1:A:167:CYS:HB2	2.21	0.41
1:A:2:TRP:CD1	1:A:175:LEU:CD1	3.02	0.41
1:A:460:GLN:HA	1:A:461:PRO:HD2	1.89	0.41
1:A:35:ALA:HB1	1:A:41:ARG:CD	2.51	0.41
1:A:126:LYS:HE3	1:A:126:LYS:HB2	1.94	0.41
1:A:76:ARG:NH2	1:A:419:ASN:ND2	2.63	0.41
1:A:167:CYS:HB3	1:A:168:PRO:CD	2.44	0.40
1:A:57:LEU:HD23	1:A:57:LEU:HA	1.92	0.40
1:A:348:ARG:NH1	1:A:437:ASN:HD22	2.16	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%)	565 (95%)	21 (4%)	6 (1%)	22	45

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	11	PRO
1	A	167	CYS
1	A	170	PRO
1	A	171	PRO
1	A	327	PRO
1	A	209	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	517/517 (100%)	501 (97%)	16 (3%)	52	81

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	TRP
1	A	22	ARG
1	A	41	ARG
1	A	64	ARG
1	A	108	ASP
1	A	122	ASN
1	A	127	THR
1	A	172	TYR
1	A	190	SER
1	A	347	PHE
1	A	352	MET
1	A	360	ARG
1	A	410	ASP
1	A	412	MET
1	A	427	LYS
1	A	594	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	18	ASN
1	A	122	ASN
1	A	128	GLN
1	A	138	ASN
1	A	147	ASN
1	A	364	ASN
1	A	419	ASN
1	A	437	ASN
1	A	497	ASN
1	A	570	ASN
1	A	595	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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	6.08	1 (11%)	10,12,14	2.16	3 (30%)

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	-	0/6/8/10	0/0/0/0



All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	198	SEP	O-C	18.09	1.23	1.11

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	198	SEP	C-CA-N	-4.77	109.06	113.83
1	A	198	SEP	OG-CB-CA	3.00	112.94	108.69
1	A	198	SEP	O3P-P-OG	-2.06	100.95	106.65

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates ⓘ

2 carbohydrates 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	$\# Z  > 2$
5	NAG	A	601	1,5	12,14,15	1.95	4 (33%)	15,19,21	2.17	7 (46%)
5	NAG	A	602	5	12,14,15	0.89	1 (8%)	15,19,21	1.35	2 (13%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	A	601	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	602	5	-	0/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	601	NAG	O5-C5	-4.07	1.37	1.45
5	A	601	NAG	O4-C4	-3.52	1.34	1.43
5	A	601	NAG	C4-C5	-3.14	1.46	1.53
5	A	602	NAG	C3-C2	-2.45	1.47	1.52
5	A	601	NAG	C3-C2	-2.01	1.48	1.52

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	601	NAG	O4-C4-C5	-4.27	98.04	109.28
5	A	601	NAG	O5-C5-C4	-4.07	105.48	110.65
5	A	602	NAG	C3-C4-C5	3.21	115.94	110.20
5	A	601	NAG	O5-C5-C6	3.13	110.27	106.98
5	A	601	NAG	C3-C4-C5	2.66	114.94	110.20
5	A	602	NAG	O5-C5-C6	2.35	109.45	106.98
5	A	601	NAG	O4-C4-C3	-2.12	105.59	110.35
5	A	601	NAG	C6-C5-C4	2.03	117.91	113.00
5	A	601	NAG	C3-C2-N2	-2.02	108.69	111.76

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry ⓘ

Of 21 ligands modelled in this entry, 11 are monoatomic - leaving 10 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
4	NAG	A	596	1	12,14,15	0.97	1 (8%)	15,19,21	0.90	0
7	EUG	A	597	-	10,11,12	3.15	4 (40%)	13,14,15	2.22	5 (38%)
8	MPD	A	598	-	7,7,7	2.07	1 (14%)	10,10,10	1.12	1 (10%)
4	NAG	A	599	1	12,14,15	1.08	1 (8%)	15,19,21	0.83	0
10	EDO	A	600	-	3,3,3	0.33	0	2,2,2	0.18	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
10	EDO	A	603	-	3,3,3	0.53	0	2,2,2	0.38	0
4	NAG	A	604	1	12,14,15	0.64	0	15,19,21	0.81	0
3	HEM	A	605	1,12	49,50,50	3.27	16 (32%)	46,82,82	1.76	8 (17%)
11	GOL	A	610	-	5,5,5	0.47	0	5,5,5	0.54	0
9	SCN	A	700	-	2,2,2	2.37	1 (50%)	1,1,1	0.11	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
4	NAG	A	596	1	-	0/6/23/26	0/1/1/1
7	EUG	A	597	-	-	0/4/4/5	0/1/1/1
8	MPD	A	598	-	-	0/5/5/5	0/0/0/0
4	NAG	A	599	1	-	0/6/23/26	0/1/1/1
10	EDO	A	600	-	-	0/1/1/1	0/0/0/0
10	EDO	A	603	-	-	0/1/1/1	0/0/0/0
4	NAG	A	604	1	-	0/6/23/26	0/1/1/1
3	HEM	A	605	1,12	-	0/14/114/114	0/0/8/8
11	GOL	A	610	-	-	0/4/4/4	0/0/0/0
9	SCN	A	700	-	-	0/0/0/0	0/0/0/0

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	605	HEM	C3D-C4D	12.98	1.47	1.44
3	A	605	HEM	C2B-C1B	7.92	1.46	1.44
7	A	597	EUG	C4-C3	7.10	1.52	1.40
3	A	605	HEM	FE-NB	6.54	2.21	1.97
3	A	605	HEM	C3B-C2B	-6.08	1.33	1.43
3	A	605	HEM	C4A-C3A	5.44	1.47	1.40
3	A	605	HEM	C3B-CAB	5.41	1.57	1.40
3	A	605	HEM	C3C-CAC	5.11	1.56	1.40
8	A	598	MPD	O2-C2	-4.98	1.30	1.44
7	A	597	EUG	C5-C4	-4.80	1.30	1.39
3	A	605	HEM	CMC-C2C	4.29	1.60	1.47
3	A	605	HEM	C3C-C2C	-3.70	1.37	1.43
7	A	597	EUG	O3-C3	3.54	1.42	1.37
3	A	605	HEM	C4D-ND	-3.53	1.32	1.39
9	A	700	SCN	C-S	3.27	1.84	1.63
3	A	605	HEM	FE-NA	3.25	2.06	1.92

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	599	NAG	O5-C5	-3.10	1.39	1.45
3	A	605	HEM	CHD-C4C	2.87	1.41	1.36
7	A	597	EUG	O4-C4	2.85	1.42	1.36
4	A	596	NAG	O5-C5	-2.52	1.40	1.45
3	A	605	HEM	CMB-C2B	2.41	1.54	1.47
3	A	605	HEM	O1A-CGA	2.17	1.30	1.22
3	A	605	HEM	CAA-C2A	2.02	1.55	1.52
3	A	605	HEM	CHC-C4B	-2.01	1.34	1.39

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	605	HEM	C3B-C4B-NB	-5.60	109.99	114.00
3	A	605	HEM	CAD-C3D-C4D	5.57	134.54	124.53
7	A	597	EUG	C5-C6-C1	5.03	127.63	121.30
3	A	605	HEM	C4D-ND-C1D	4.17	109.43	105.16
7	A	597	EUG	C6-C5-C4	-3.01	117.22	120.50
3	A	605	HEM	CAD-C3D-C2D	-2.87	120.85	127.25
7	A	597	EUG	O3-C3-C4	2.72	118.17	114.57
3	A	605	HEM	CBD-CAD-C3D	-2.70	108.47	114.37
7	A	597	EUG	C9-O3-C3	2.67	121.53	117.59
8	A	598	MPD	O2-C2-C3	2.57	118.19	109.16
3	A	605	HEM	CMD-C2D-C3D	-2.55	119.83	125.60
3	A	605	HEM	C2D-C1D-ND	-2.11	110.44	112.93
7	A	597	EUG	C2-C1-C7	2.09	124.98	120.95
3	A	605	HEM	C1A-CHA-C4D	2.07	130.19	127.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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	595/595 (100%)	0.14	29 (4%) 28 25	33, 56, 97, 100	0

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	173	GLN	11.1
1	A	174	SER	10.5
1	A	121	SER	9.7
1	A	122	ASN	6.8
1	A	120	GLY	6.7
1	A	13	VAL	5.2
1	A	7	GLY	4.9
1	A	119	LEU	4.9
1	A	172	TYR	4.6
1	A	2	TRP	4.5
1	A	595	ASN	4.3
1	A	10	VAL	4.0
1	A	1	SER	4.0
1	A	211	GLY	3.8
1	A	212	LEU	3.5
1	A	210	LEU	3.3
1	A	587	LEU	3.3
1	A	64	ARG	3.2
1	A	12	LEU	3.2
1	A	124	HIS	3.0
1	A	123	GLU	2.8
1	A	169	THR	2.6
1	A	283	LEU	2.4
1	A	126	LYS	2.3
1	A	329	GLN	2.3
1	A	108	ASP	2.2
1	A	104	GLY	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	209	PRO	2.1
1	A	170	PRO	2.1

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

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	SEP	A	198	10/11	0.12	-0.94	51,54,58,58	0

## 6.3 Carbohydrates ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	NAG	A	602	14/15	0.23	1.42	73,79,81,84	0
5	NAG	A	601	14/15	0.18	0.79	65,69,71,71	0

## 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. 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
9	SCN	A	700	3/3	0.36	6.27	61,61,62,64	3
7	EUG	A	597	11/12	0.50	5.42	30,42,45,45	11
4	NAG	A	604	14/15	0.20	5.37	89,92,94,94	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
11	GOL	A	610	6/6	0.27	2.93	74,76,77,78	6
4	NAG	A	596	14/15	0.28	1.36	78,79,80,81	0
8	MPD	A	598	8/8	0.20	1.11	24,32,41,46	0
3	HEM	A	605	43/43	0.24	0.41	30,37,43,48	0
6	IOD	A	614	1/1	0.18	0.41	79,79,79,79	1
2	CA	A	606	1/1	0.20	0.29	54,54,54,54	0
10	EDO	A	600	4/4	0.17	-0.27	74,74,76,76	0
10	EDO	A	603	4/4	0.24	-0.60	75,75,75,77	4
4	NAG	A	599	14/15	0.14	-0.70	63,75,77,77	0
6	IOD	A	609	1/1	0.12	-0.99	84,84,84,84	1
6	IOD	A	607	1/1	0.08	-1.47	56,56,56,56	0
6	IOD	A	624	1/1	0.06	-1.56	95,95,95,95	1
6	IOD	A	621	1/1	0.11	-1.58	91,91,91,91	1
6	IOD	A	623	1/1	0.06	-1.64	88,88,88,88	1
6	IOD	A	612	1/1	0.07	-2.90	97,97,97,97	0
6	IOD	A	608	1/1	0.03	-3.46	84,84,84,84	0
6	IOD	A	611	1/1	0.05	-3.91	74,74,74,74	0
6	IOD	A	622	1/1	0.04	-4.02	75,75,75,75	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.