



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 6, 2017 – 05:46 PM EDT

PDB ID : 1CXP  
Title : CRYOGENIC CRYSTAL STRUCTURE OF HUMAN MYELOPEROXIDASE ISOFORM C  
Authors : Fiedler, T.J.; Fenna, R.E.  
Deposited on : unknown  
Resolution : 1.80 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	:	rb-20029824
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)	:	rb-20029824

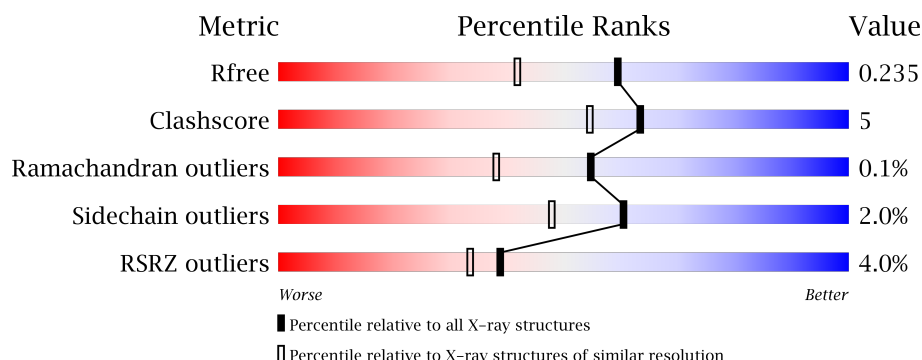
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	4827 (1.80-1.80)
Clashscore	112137	5742 (1.80-1.80)
Ramachandran outliers	110173	5676 (1.80-1.80)
Sidechain outliers	110143	5675 (1.80-1.80)
RSRZ outliers	101464	4906 (1.80-1.80)

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	104	<div> <div>5%</div> <div> <div></div> <div>91%</div> <div>9%</div> </div> </div>
1	B	104	<div> <div>3%</div> <div> <div></div> <div>97%</div> <div>.</div> </div> </div>
2	C	466	<div> <div>2%</div> <div> <div></div> <div>86%</div> <div>12%</div> <div>.</div> </div> </div>
2	D	466	<div> <div>6%</div> <div> <div></div> <div>86%</div> <div>14%</div> <div>.</div> </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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
11	ACT	C	1606	-	-	X	X
6	NAG	C	1620	-	-	-	X
6	NAG	C	1630	-	-	-	X

## 2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 10307 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 MYELOPEROXIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	104	Total	C	N	O	S	0	0	0
			838	529	148	156	5			
1	B	104	Total	C	N	O	S	0	0	0
			838	529	148	156	5			

- Molecule 2 is a protein called MYELOPEROXIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	466	Total	C	N	O	S	0	0	0
			3733	2351	687	668	27			
2	D	466	Total	C	N	O	S	0	0	0
			3733	2351	687	668	27			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	150	CSO	CYS	MODIFIED RESIDUE	UNP P05164
D	150	CSO	CYS	MODIFIED RESIDUE	UNP P05164

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

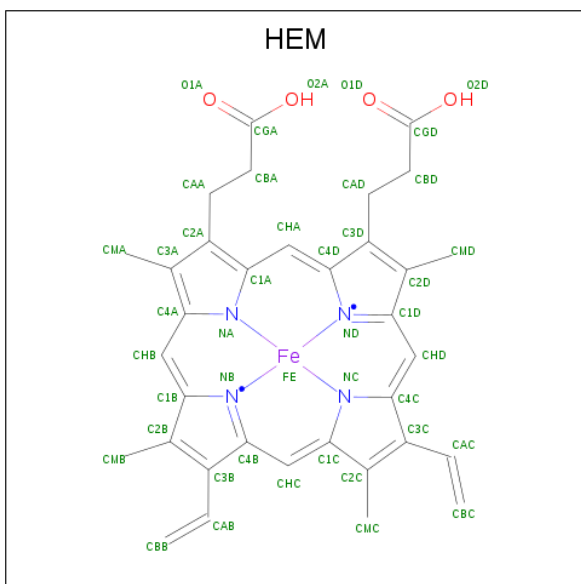
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Cl	0	0
			1	1		
3	A	1	Total	Cl	0	0
			1	1		

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



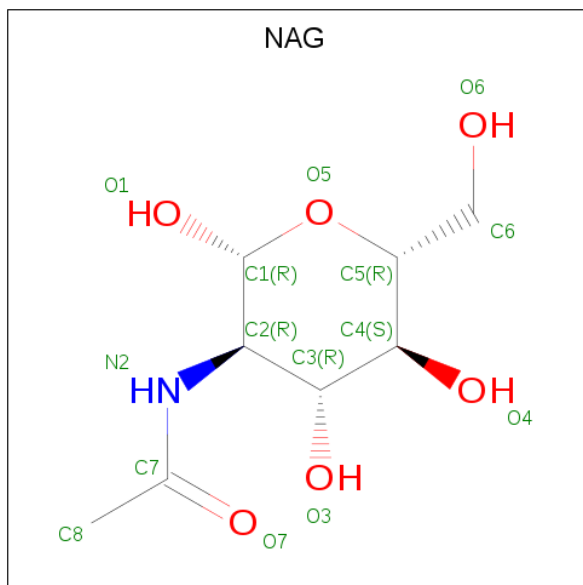
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 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
5	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
5	B	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

- Molecule 6 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



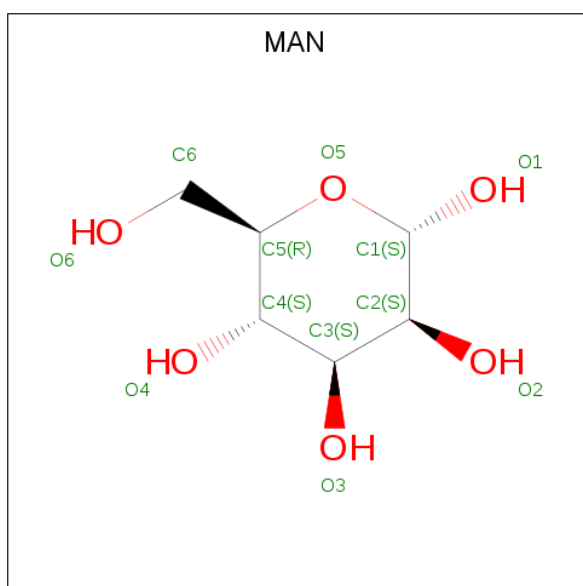
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	C	1	Total	C	N	O	0	0
			14	8	1	5		
6	C	1	Total	C	N	O	0	0
			14	8	1	5		
6	C	1	Total	C	N	O	0	0
			14	8	1	5		
6	C	1	Total	C	N	O	0	0
			14	8	1	5		
6	D	1	Total	C	N	O	0	0
			14	8	1	5		
6	D	1	Total	C	N	O	0	0
			14	8	1	5		
6	D	1	Total	C	N	O	0	0
			14	8	1	5		
6	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 7 is BETA-D-MANNOSE (three-letter code: BMA) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	C	1	Total	C	O	0	0
			11	6	5		
7	D	1	Total	C	O	0	0
			11	6	5		

- Molecule 8 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



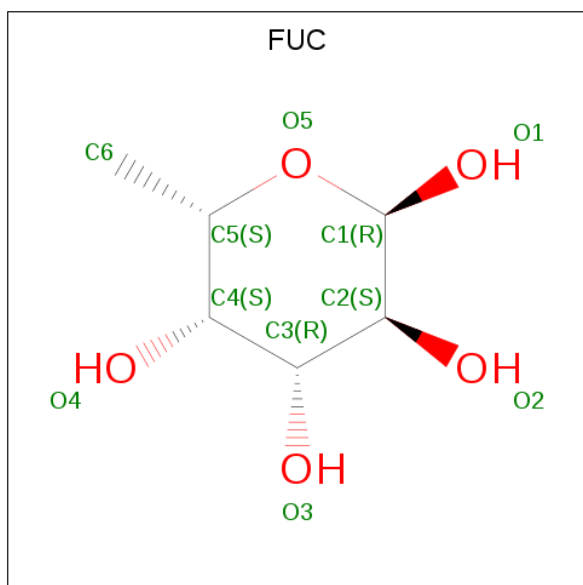
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	C	1	Total	C	O	0	0
			11	6	5		
8	C	1	Total	C	O	0	0
			11	6	5		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	D	1	Total	C	O	0	0
			11	6	5		
8	D	1	Total	C	O	0	0
			11	6	5		

- Molecule 9 is ALPHA-L-FUCOSE (three-letter code: FUC) (formula:  $C_6H_{12}O_5$ ).



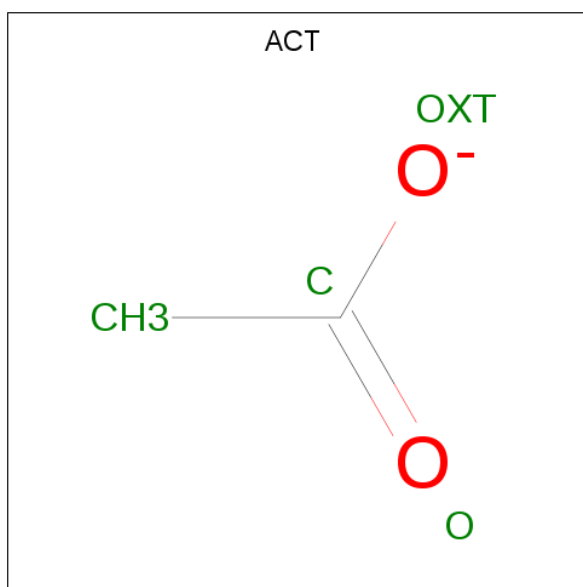
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	C	1	Total	C	O	0	0
			10	6	4		
9	D	1	Total	C	O	0	0
			10	6	4		

- Molecule 10 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	D	1	Total	Ca	0	0
			1	1		
10	C	1	Total	Ca	0	0
			1	1		

- Molecule 11 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	C	1	Total C O 4 2 2	0	0
11	C	1	Total C O 4 2 2	0	0
11	C	1	Total C O 4 2 2	0	0
11	D	1	Total C O 4 2 2	0	0
11	D	1	Total C O 4 2 2	0	0
11	D	1	Total C O 4 2 2	0	0

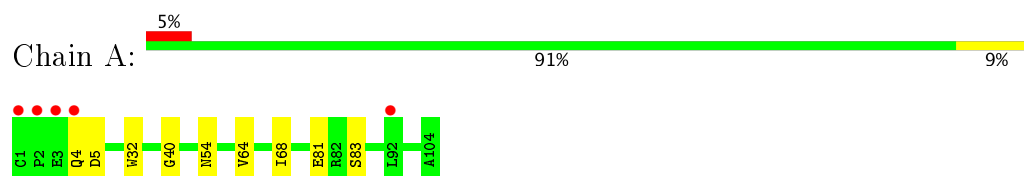
- Molecule 12 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	A	102	Total O 102 102	0	0
12	C	315	Total O 315 315	0	0
12	B	107	Total O 107 107	0	0
12	D	314	Total O 314 314	0	0

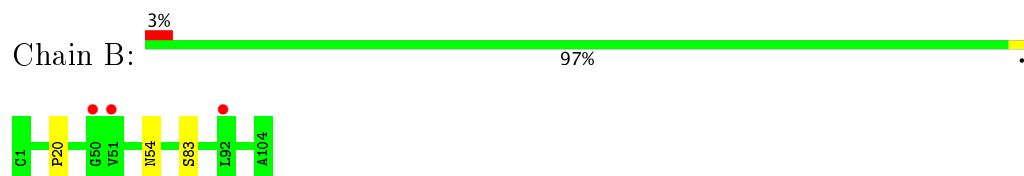
### 3 Residue-property plots [i](#)

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.

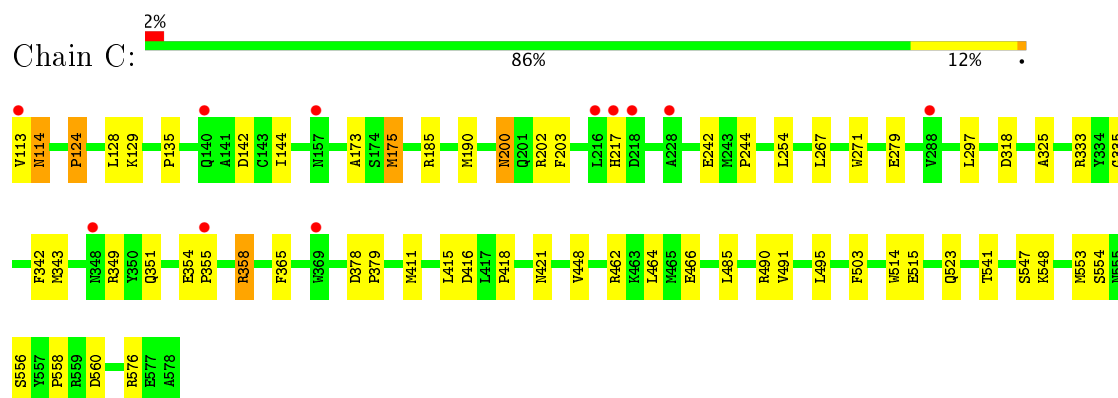
#### • Molecule 1: MYELOPEROXIDASE



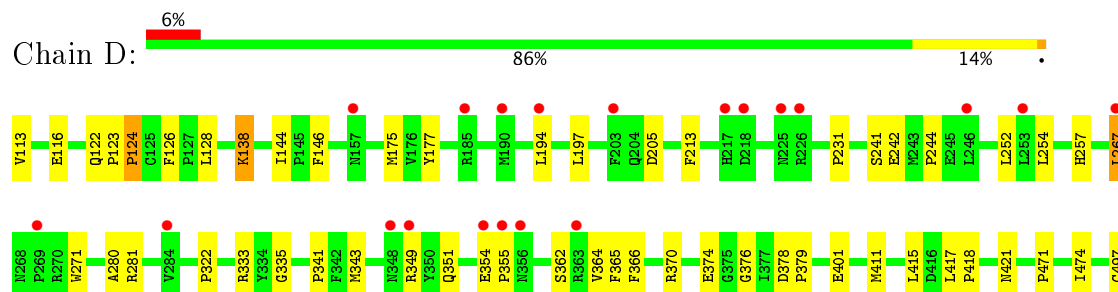
#### • Molecule 1: MYELOPEROXIDASE



#### • Molecule 2: MYELOPEROXIDASE



#### • Molecule 2: MYELOPEROXIDASE





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.03 Å 63.38 Å 92.17 Å 90.00° 97.36° 90.00°	Depositor
Resolution (Å)	30.00 – 1.80 34.08 – 1.75	Depositor EDS
% Data completeness (in resolution range)	96.1 (30.00-1.80) 99.6 (34.08-1.75)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.55 (at 1.75 Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.197 , 0.239 0.195 , 0.235	Depositor DCC
$R_{free}$ test set	5940 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	12.8	Xtriage
Anisotropy	0.406	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 53.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	10307	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% 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: CSO, BMA, NAG, CL, CA, SO4, ACT, HEM, MAN, FUC

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	0.67	0/863	0.71	0/1174
1	B	0.68	0/863	0.72	0/1174
2	C	0.62	0/3811	0.64	0/5168
2	D	0.61	0/3811	0.61	0/5168
All	All	0.63	0/9348	0.64	0/12684

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	A	838	0	798	8	0
1	B	838	0	798	2	0
2	C	3733	0	3725	42	0
2	D	3733	0	3725	43	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	5	0	0	0	0
4	B	5	0	0	0	0
4	C	5	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	43	0	30	4	0
5	B	43	0	30	3	0
6	C	56	0	49	0	0
6	D	56	0	49	0	0
7	C	11	0	8	0	0
7	D	11	0	8	0	0
8	C	22	0	20	0	0
8	D	22	0	20	0	0
9	C	10	0	10	0	0
9	D	10	0	10	0	0
10	C	1	0	0	0	0
10	D	1	0	0	0	0
11	C	12	0	9	2	0
11	D	12	0	9	0	0
12	A	102	0	0	1	0
12	B	107	0	0	0	0
12	C	315	0	0	6	0
12	D	314	0	0	1	0
All	All	10307	0	9298	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 5.

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)
12:A:1243(A):HOH:O	2:C:129:LYS:HD3	1.83	0.78
2:D:354:GLU:HB3	2:D:355:PRO:HA	1.77	0.67
2:D:349:ARG:HG3	2:D:351:GLN:HG2	1.80	0.64
2:C:333:ARG:HH11	2:C:421:ASN:HD22	1.47	0.63
2:C:200:ASN:ND2	2:C:202:ARG:H	1.97	0.62
2:C:354:GLU:HB3	2:C:355:PRO:HA	1.82	0.62
2:C:200:ASN:HD22	2:C:202:ARG:H	1.48	0.60
2:C:142:ASP:HB3	12:C:1239(A):HOH:O	2.02	0.59
2:C:135:PRO:HG2	12:C:1204(A):HOH:O	2.03	0.58
2:C:485:LEU:HD13	2:C:490:ARG:HA	1.84	0.58
1:A:4:GLN:HG2	1:A:5:ASP:N	2.18	0.58
2:D:333:ARG:HH11	2:D:421:ASN:ND2	2.01	0.57
2:D:333:ARG:HH11	2:D:421:ASN:HD22	1.50	0.57
5:A:605:HEM:HMC2	5:A:605:HEM:HBC2	1.87	0.56
2:D:411:MET:HE2	2:D:415:LEU:HD21	1.87	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:349:ARG:HE	2:D:351:GLN:HG3	1.71	0.56
2:D:411:MET:CE	2:D:415:LEU:HD21	2.35	0.56
1:A:4:GLN:HG2	1:A:5:ASP:H	1.71	0.55
2:C:349:ARG:HG3	2:C:351:GLN:HG2	1.87	0.55
2:D:113:VAL:HG21	2:D:122:GLN:HB2	1.88	0.54
2:D:378:ASP:HB2	2:D:379:PRO:HD3	1.90	0.54
2:C:333:ARG:HH11	2:C:421:ASN:ND2	2.06	0.53
2:C:128:LEU:HB2	2:C:144:ILE:HB	1.91	0.53
1:A:68:ILE:HD13	2:C:464:LEU:HD23	1.91	0.53
2:D:138:LYS:NZ	2:D:138:LYS:HA	2.24	0.52
2:D:417:LEU:HB3	2:D:418:PRO:HD3	1.92	0.52
1:B:83:SER:HB3	2:D:554:SER:O	2.09	0.52
2:D:197:LEU:HG	2:D:257:HIS:CD2	2.45	0.52
2:D:194:LEU:HD13	12:D:948(B):HOH:O	2.09	0.52
2:C:548:LYS:HD2	2:C:560:ASP:HA	1.90	0.52
2:D:122:GLN:HA	2:D:122:GLN:HE21	1.76	0.51
2:D:244:PRO:HD3	2:D:364:VAL:O	2.10	0.51
5:B:605:HEM:CBC	2:D:335:GLY:HA3	2.41	0.51
2:C:113:VAL:HG12	2:C:114:ASN:N	2.26	0.51
2:C:411:MET:HE2	2:C:415:LEU:HD21	1.93	0.51
2:D:241:SER:O	2:D:366:PHE:HA	2.11	0.51
5:A:605:HEM:HBB2	2:C:242:GLU:OE1	2.10	0.50
2:D:514:TRP:CE2	2:D:515:GLU:HG3	2.47	0.50
2:C:556:SER:HA	12:C:862(A):HOH:O	2.12	0.50
2:C:491:VAL:HB	2:C:495:LEU:HB2	1.94	0.49
2:C:271:TRP:CZ3	2:C:279:GLU:HG3	2.48	0.48
2:C:514:TRP:CE2	2:C:515:GLU:HG3	2.49	0.48
2:D:244:PRO:HB2	2:D:343:MET:SD	2.54	0.47
1:A:83:SER:HB3	2:C:554:SER:O	2.14	0.47
5:B:605:HEM:HBB2	2:D:242:GLU:OE1	2.14	0.47
2:C:378:ASP:HB2	2:C:379:PRO:HD3	1.96	0.47
2:C:462:ARG:O	2:C:466:GLU:HG2	2.16	0.47
2:C:378:ASP:OD1	2:C:541:THR:HB	2.14	0.46
2:D:257:HIS:CE1	2:D:280:ALA:HB3	2.50	0.46
2:D:242:GLU:O	2:D:365:PHE:HA	2.16	0.46
5:B:605:HEM:HBC2	2:D:335:GLY:HA3	1.98	0.46
5:A:605:HEM:HBC2	5:A:605:HEM:CMC	2.46	0.45
2:D:177:TYR:OH	2:D:281:ARG:HA	2.17	0.45
11:C:1606:ACT:H2	12:C:1213(A):HOH:O	2.16	0.45
2:C:342:PHE:CD1	2:C:358:ARG:NH1	2.85	0.45
2:C:135:PRO:CG	12:C:1204(A):HOH:O	2.64	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:267:LEU:HD22	2:D:576:ARG:CZ	2.48	0.44
1:A:40:GLY:HA2	1:B:20:PRO:HD2	1.99	0.44
2:D:126:PHE:HB3	2:D:146:PHE:CD2	2.52	0.44
2:C:448:VAL:HG21	2:C:462:ARG:HG2	1.99	0.44
1:A:81:GLU:O	2:C:553:MET:HG3	2.17	0.43
2:C:124:PRO:HA	11:C:1606:ACT:CH3	2.49	0.43
2:D:471:PRO:HA	2:D:474:ILE:HG13	2.01	0.43
2:D:123:PRO:HA	2:D:124:PRO:HA	1.85	0.43
2:D:370:ARG:O	2:D:374:GLU:HB2	2.17	0.43
2:C:267:LEU:HD12	2:C:576:ARG:HB2	2.01	0.43
1:A:64:VAL:HG13	1:A:68:ILE:HD12	2.00	0.43
2:C:342:PHE:CE1	2:C:358:ARG:NH1	2.87	0.42
2:D:213:PHE:CD2	2:D:231:PRO:HG2	2.55	0.42
2:D:128:LEU:HB2	2:D:144:ILE:HB	2.02	0.42
2:C:242:GLU:O	2:C:365:PHE:HA	2.20	0.42
2:D:205:ASP:OD2	2:D:376:GLY:HA3	2.18	0.42
2:D:122:GLN:HA	2:D:122:GLN:NE2	2.34	0.42
2:C:200:ASN:HD22	2:C:203:PHE:H	1.66	0.41
2:C:244:PRO:HB2	2:C:343:MET:SD	2.60	0.41
1:A:32:TRP:CE2	2:C:325:ALA:HB2	2.55	0.41
2:D:362:SER:HA	2:D:365:PHE:CE1	2.55	0.41
2:D:341:PRO:O	2:D:401:GLU:HG3	2.20	0.41
2:C:114:ASN:HA	12:C:973(A):HOH:O	2.21	0.41
2:D:116:GLU:OE1	2:D:411:MET:HE3	2.19	0.41
2:D:534:PRO:HG2	2:D:549:ASN:ND2	2.36	0.41
2:D:271:TRP:HE1	2:D:577:GLU:CD	2.24	0.41
5:A:605:HEM:CBC	2:C:335:GLY:HA3	2.51	0.41
2:C:333:ARG:HD3	2:C:421:ASN:ND2	2.35	0.41
2:C:173:ALA:HA	2:C:175:MET:SD	2.60	0.41
2:C:297:LEU:HD11	2:C:503:PHE:CD1	2.55	0.40
2:D:507:ARG:HG3	2:D:513:TRP:CE2	2.56	0.40
2:C:185:ARG:HG3	2:C:190:MET:CE	2.51	0.40
2:D:533:LEU:HB3	2:D:534:PRO:HD3	2.02	0.40
2:D:252:LEU:HD11	2:D:537:ILE:HA	2.03	0.40
2:D:378:ASP:OD1	2:D:541:THR:HB	2.22	0.40
2:C:416:ASP:OD2	2:C:418:PRO:HD2	2.21	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	102/104 (98%)	100 (98%)	2 (2%)	0	100	100
1	B	102/104 (98%)	100 (98%)	2 (2%)	0	100	100
2	C	463/466 (99%)	449 (97%)	13 (3%)	1 (0%)	51	35
2	D	463/466 (99%)	450 (97%)	13 (3%)	0	100	100
All	All	1130/1140 (99%)	1099 (97%)	30 (3%)	1 (0%)	55	38

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	114	ASN

### 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	90/90 (100%)	89 (99%)	1 (1%)	78	72
1	B	90/90 (100%)	89 (99%)	1 (1%)	78	72
2	C	410/410 (100%)	400 (98%)	10 (2%)	54	40
2	D	410/410 (100%)	402 (98%)	8 (2%)	60	48
All	All	1000/1000 (100%)	980 (98%)	20 (2%)	60	48

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	54	ASN
2	C	124	PRO
2	C	175	MET
2	C	200	ASN
2	C	217	HIS
2	C	254	LEU
2	C	318	ASP
2	C	358	ARG
2	C	523	GLN
2	C	547	SER
2	C	558	PRO
1	B	54	ASN
2	D	124	PRO
2	D	138	LYS
2	D	175	MET
2	D	254	LEU
2	D	267	LEU
2	D	322	PRO
2	D	497	CYS
2	D	523	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	200	ASN
2	C	421	ASN
2	C	530	GLN
1	B	54	ASN
2	D	122	GLN
2	D	421	ASN
2	D	549	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	CSO	C	150	2	4,6,7	1.35	1 (25%)	1,6,8	1.76	0
2	CSO	D	150	2	4,6,7	1.68	1 (25%)	1,6,8	1.40	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	CSO	C	150	2	-	0/1/5/7	0/0/0/0
2	CSO	D	150	2	-	0/1/5/7	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	150	CSO	CA-C	2.27	1.53	1.50
2	D	150	CSO	CA-C	2.91	1.54	1.50

There are no bond angle outliers.

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 31 ligands modelled in this entry, 4 are monoatomic - leaving 27 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	SO4	A	1602	-	4,4,4	0.30	0	6,6,6	0.11	0
5	HEM	A	605	1,12,2	28,50,50	1.65	7 (25%)	17,82,82	1.06	1 (5%)
4	SO4	B	1003	-	4,4,4	0.37	0	6,6,6	0.05	0
5	HEM	B	605	1,12,2	28,50,50	1.61	6 (21%)	17,82,82	1.13	1 (5%)
4	SO4	C	1603	-	4,4,4	0.33	0	6,6,6	0.12	0
11	ACT	C	1604	-	1,3,3	0.13	0	0,3,3	0.00	-
11	ACT	C	1606	-	1,3,3	0.53	0	0,3,3	0.00	-
11	ACT	C	1607	-	1,3,3	0.46	0	0,3,3	0.00	-
6	NAG	C	1620	2	14,14,15	0.49	0	15,19,21	0.92	1 (6%)
6	NAG	C	1630	2	14,14,15	0.47	0	15,19,21	0.66	0
6	NAG	C	1640	9,2,6	14,14,15	0.81	0	15,19,21	0.91	0
6	NAG	C	1641	7,6	14,14,15	0.66	0	15,19,21	0.91	0
7	BMA	C	1642	8,6	11,11,12	0.68	0	13,15,17	0.49	0
8	MAN	C	1643	7	11,11,12	0.75	0	13,15,17	0.77	1 (7%)
8	MAN	C	1644	7	11,11,12	0.57	0	13,15,17	0.71	1 (7%)
9	FUC	C	1645	6	9,10,11	0.86	1 (11%)	13,14,16	0.48	0
11	ACT	D	1004	-	1,3,3	0.67	0	0,3,3	0.00	-
11	ACT	D	1005	-	1,3,3	0.04	0	0,3,3	0.00	-
11	ACT	D	1006	-	1,3,3	0.13	0	0,3,3	0.00	-
6	NAG	D	2620	2	14,14,15	0.76	1 (7%)	15,19,21	0.76	1 (6%)
6	NAG	D	2630	2	14,14,15	0.47	0	15,19,21	0.78	0
6	NAG	D	2640	9,2,6	14,14,15	0.62	0	15,19,21	0.96	0
6	NAG	D	2641	7,6	14,14,15	0.73	0	15,19,21	1.01	2 (13%)
7	BMA	D	2642	8,6	11,11,12	0.51	0	13,15,17	0.59	0
8	MAN	D	2643	7	11,11,12	0.61	0	13,15,17	0.63	0
8	MAN	D	2644	7	11,11,12	0.92	0	13,15,17	0.55	0
9	FUC	D	2645	6	9,10,11	0.81	0	13,14,16	0.59	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	SO4	A	1602	-	-	0/0/0/0	0/0/0/0
5	HEM	A	605	1,12,2	-	0/6/54/54	0/0/8/8
4	SO4	B	1003	-	-	0/0/0/0	0/0/0/0
5	HEM	B	605	1,12,2	-	0/6/54/54	0/0/8/8
4	SO4	C	1603	-	-	0/0/0/0	0/0/0/0
11	ACT	C	1604	-	-	0/0/0/0	0/0/0/0
11	ACT	C	1606	-	-	0/0/0/0	0/0/0/0
11	ACT	C	1607	-	-	0/0/0/0	0/0/0/0
6	NAG	C	1620	2	-	0/6/23/26	0/1/1/1
6	NAG	C	1630	2	-	0/6/23/26	0/1/1/1
6	NAG	C	1640	9,2,6	-	0/6/23/26	0/1/1/1
6	NAG	C	1641	7,6	-	0/6/23/26	0/1/1/1
7	BMA	C	1642	8,6	-	0/2/19/22	0/1/1/1
8	MAN	C	1643	7	-	0/2/19/22	0/1/1/1
8	MAN	C	1644	7	-	0/2/19/22	0/1/1/1
9	FUC	C	1645	6	-	0/0/17/20	0/1/1/1
11	ACT	D	1004	-	-	0/0/0/0	0/0/0/0
11	ACT	D	1005	-	-	0/0/0/0	0/0/0/0
11	ACT	D	1006	-	-	0/0/0/0	0/0/0/0
6	NAG	D	2620	2	-	0/6/23/26	0/1/1/1
6	NAG	D	2630	2	-	0/6/23/26	0/1/1/1
6	NAG	D	2640	9,2,6	-	0/6/23/26	0/1/1/1
6	NAG	D	2641	7,6	-	0/6/23/26	0/1/1/1
7	BMA	D	2642	8,6	-	0/2/19/22	0/1/1/1
8	MAN	D	2643	7	-	0/2/19/22	0/1/1/1
8	MAN	D	2644	7	-	0/2/19/22	0/1/1/1
9	FUC	D	2645	6	-	0/0/17/20	0/1/1/1

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	605	HEM	C3B-C2B	-3.80	1.35	1.40
5	A	605	HEM	C3B-CAB	-3.19	1.41	1.47
5	A	605	HEM	C3C-CAC	-3.11	1.41	1.47
5	B	605	HEM	C3B-CAB	-3.09	1.41	1.47
5	B	605	HEM	C3C-CAC	-2.64	1.42	1.47
5	A	605	HEM	C3B-C2B	-2.46	1.37	1.40
5	A	605	HEM	C3C-C2C	-2.03	1.37	1.40
6	D	2620	NAG	C1-C2	2.09	1.55	1.52
9	C	1645	FUC	C2-C3	2.22	1.55	1.52
5	A	605	HEM	CBC-CAC	2.32	1.45	1.28
5	B	605	HEM	CBC-CAC	2.40	1.45	1.28
5	B	605	HEM	CBB-CAB	2.48	1.46	1.28

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	605	HEM	CBB-CAB	2.51	1.46	1.28
5	B	605	HEM	C1C-NC	2.53	1.39	1.36
5	A	605	HEM	C1B-NB	3.26	1.40	1.36

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	D	2641	NAG	C4-C3-C2	-2.29	107.67	111.02
6	D	2641	NAG	C2-N2-C7	-2.27	119.63	122.94
6	C	1620	NAG	C1-O5-C5	2.05	114.99	112.17
8	C	1644	MAN	C1-O5-C5	2.11	115.07	112.17
6	D	2620	NAG	C1-O5-C5	2.12	115.09	112.17
5	A	605	HEM	C3B-C4B-NB	2.29	112.18	109.21
5	B	605	HEM	C3B-C4B-NB	2.54	112.50	109.21
8	C	1643	MAN	C1-O5-C5	2.55	115.68	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	605	HEM	4	0
5	B	605	HEM	3	0
11	C	1606	ACT	2	0

## 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	104/104 (100%)	0.37	5 (4%) 31 26	5, 8, 22, 44	0
1	B	104/104 (100%)	0.33	3 (2%) 52 47	5, 9, 18, 23	0
2	C	465/466 (99%)	0.34	11 (2%) 59 55	4, 9, 22, 38	0
2	D	465/466 (99%)	0.57	26 (5%) 25 21	4, 12, 27, 42	0
All	All	1138/1140 (99%)	0.44	45 (3%) 39 33	4, 10, 24, 44	0

All (45) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	113	VAL	6.8
2	D	355	PRO	6.6
2	D	217	HIS	6.4
2	D	578	ALA	6.3
2	D	190	MET	4.9
2	C	355	PRO	4.4
1	A	4	GLN	3.9
2	C	217	HIS	3.9
2	D	568	PRO	3.7
1	A	2	PRO	3.5
1	A	3	GLU	3.4
2	C	140	GLN	3.4
1	B	50	GLY	3.4
2	D	576	ARG	3.3
2	D	269	PRO	3.1
1	A	1	CYS	3.1
2	D	563	ASN	3.1
2	D	157	ASN	3.0
2	C	218	ASP	2.8
2	D	562	VAL	2.8
2	D	218	ASP	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	349	ARG	2.8
2	D	226	ARG	2.7
2	C	348	ASN	2.7
2	D	194	LEU	2.6
2	D	267	LEU	2.6
2	D	253	LEU	2.5
2	D	354	GLU	2.5
2	D	348	ASN	2.4
2	D	203	PHE	2.4
2	D	246	LEU	2.3
1	A	92	LEU	2.3
2	D	225	ASN	2.3
2	D	522	MET	2.3
2	C	228	ALA	2.3
2	C	157	ASN	2.3
2	D	356	ASN	2.2
1	B	92	LEU	2.2
2	C	369	TRP	2.2
2	D	185	ARG	2.1
2	C	288	VAL	2.1
2	C	216	LEU	2.1
1	B	51	VAL	2.1
2	D	284	VAL	2.1
2	D	363	ARG	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
2	CSO	C	150	7/8	0.92	0.11	-	6,8,10,13	0
2	CSO	D	150	7/8	0.92	0.11	-	5,7,8,12	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.



## 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	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
11	ACT	C	1606	4/4	0.78	0.15	2.90	23,25,25,26	0
6	NAG	C	1630	14/15	0.78	0.24	2.45	20,27,35,37	0
6	NAG	C	1620	14/15	0.84	0.15	2.35	13,17,21,22	0
6	NAG	D	2630	14/15	0.82	0.24	1.62	30,33,36,37	0
7	BMA	C	1642	11/12	0.93	0.14	1.56	9,12,14,19	0
7	BMA	D	2642	11/12	0.92	0.11	0.79	10,11,14,17	0
6	NAG	D	2641	14/15	0.93	0.11	0.65	6,9,11,11	0
6	NAG	D	2620	14/15	0.84	0.20	0.61	14,23,28,31	0
6	NAG	C	1641	14/15	0.91	0.12	0.06	7,9,10,10	0
5	HEM	B	605	43/43	0.94	0.14	0.03	2,7,9,9	0
5	HEM	A	605	43/43	0.95	0.12	-0.15	4,6,8,9	0
10	CA	D	1001	1/1	1.00	0.11	-1.21	7,7,7,7	0
3	CL	B	1002	1/1	0.99	0.12	-1.62	7,7,7,7	0
10	CA	C	1600	1/1	1.00	0.10	-1.94	8,8,8,8	0
3	CL	A	1601	1/1	1.00	0.10	-2.19	6,6,6,6	0
6	NAG	C	1640	14/15	0.90	0.13	-	8,9,12,15	0
11	ACT	D	1005	4/4	0.73	0.32	-	35,35,35,36	0
8	MAN	D	2643	11/12	0.81	0.21	-	20,22,26,28	0
9	FUC	C	1645	10/11	0.85	0.16	-	14,16,17,17	0
4	SO4	B	1003	5/5	0.65	0.27	-	46,46,47,47	5
4	SO4	A	1602	5/5	0.95	0.11	-	24,24,26,29	0
9	FUC	D	2645	10/11	0.93	0.15	-	11,13,16,17	0
8	MAN	C	1643	11/12	0.75	0.26	-	22,26,27,27	0
11	ACT	C	1607	4/4	0.69	0.21	-	28,28,29,30	0
6	NAG	D	2640	14/15	0.93	0.13	-	8,10,17,18	0
8	MAN	C	1644	11/12	0.94	0.12	-	13,15,16,16	0
4	SO4	C	1603	5/5	0.94	0.14	-	24,24,28,29	0
11	ACT	D	1004	4/4	0.90	0.12	-	20,20,22,23	0
11	ACT	D	1006	4/4	0.88	0.20	-	28,29,29,30	0
8	MAN	D	2644	11/12	0.90	0.16	-	11,13,16,16	0
11	ACT	C	1604	4/4	0.89	0.15	-	17,19,19,22	0

## 6.5 Other polymers

There are no such residues in this entry.