Full wwPDB X-ray Structure Validation Report (i)

Aug 29, 2023 – 08:04 AM EDT

PDB ID	:	3MLJ
Title	:	Reduced (Cu+) peptidylglycine alpha-hydroxylating monooxygenase (PHM) with bound carbon monooxide (CO)
Authors	:	Prigge, S.T.; Chufan, E.E.; Eipper, B.A.; Mains, R.E.; Amzel, L.M.
Deposited on	:	2010-04-16
Resolution	:	2.15 Å(reported)

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

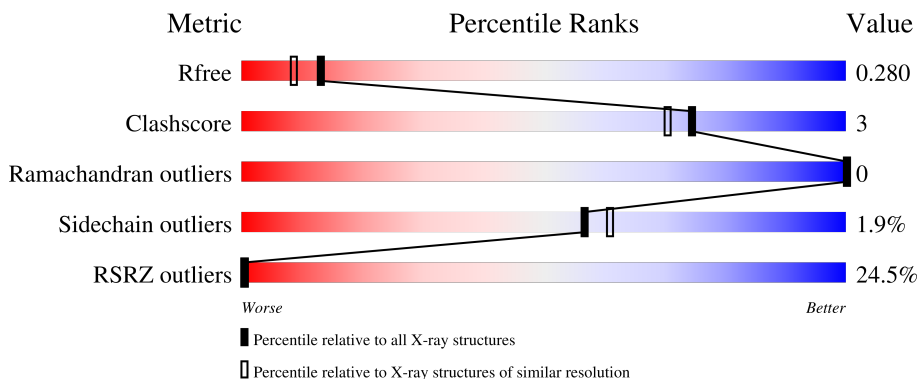|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.15 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	Percentile Ranks	Value
Rfree		0.280
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		1.9%
RSRZ outliers		24.5%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=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	A	314	24% 90% 8% ··

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 2667 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 Peptidyl-glycine alpha-amidating monooxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	310	Total	C	N	O	S	0	11	0
			2485	1590	419	448	28			

- Molecule 2 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Cu	0	0
			2	2		

- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

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

- Molecule 4 is CARBON MONOXIDE (three-letter code: CMO) (formula: CO).
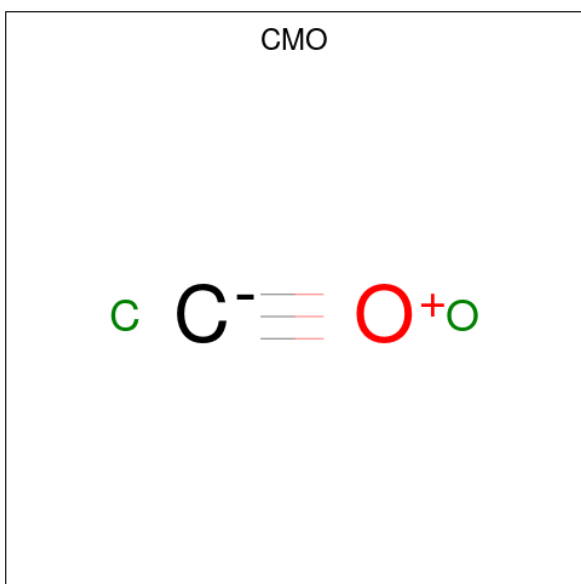
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total C O			0	0
			2	1	1		
4	A	1	Total C O			0	0
			2	1	1		

- Molecule 5 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).
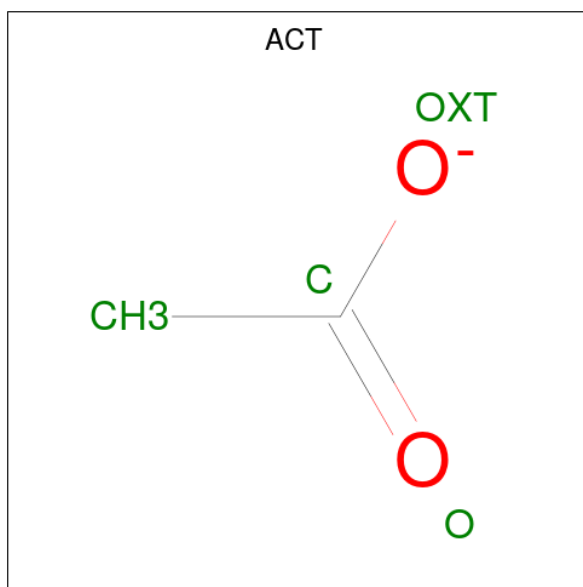
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total C O			0	0
			4	2	2		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
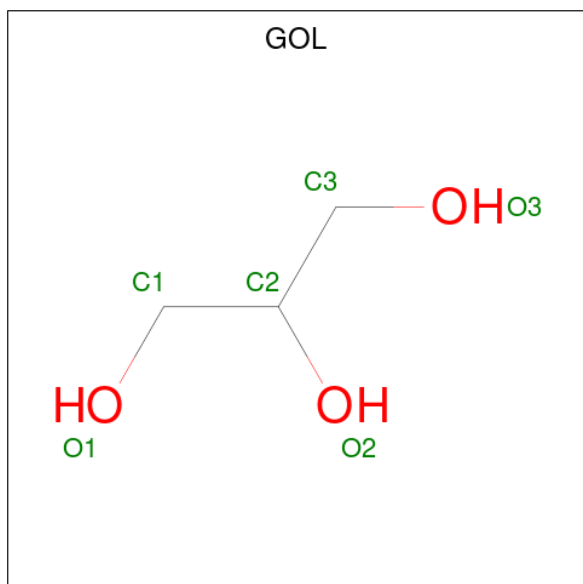
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total 6	C 3	O 3	0	0
6	A	1	Total 6	C 3	O 3	0	0
6	A	1	Total 6	C 3	O 3	0	0

• Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	153	Total 153	O 153	0	0

3 Residue-property plots ⓘ

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: Peptidyl-glycine alpha-amidating monooxygenase
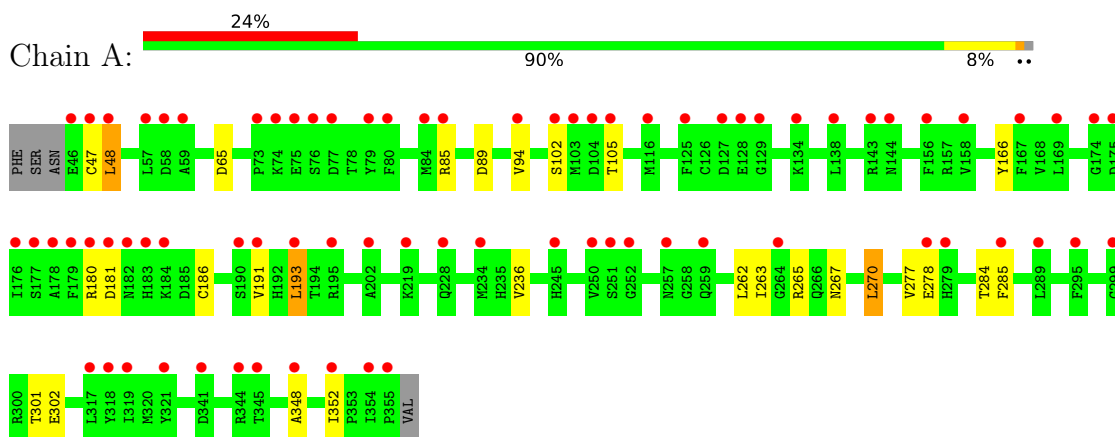
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	69.23Å 68.86Å 81.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	53.00 – 2.15 20.01 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.1 (53.00-2.15) 88.8 (20.01-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	1.14 (at 2.01Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.210 , 0.260 0.248 , 0.280	Depositor DCC
R_{free} test set	1311 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	45.4	Xtriage
Anisotropy	0.265	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 52.4	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.028 for k,h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	2667	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.24% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: GOL, NI, CU, ACT, CMO

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	0.48	2/2589 (0.1%)	0.59	2/3518 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	180	ARG	CZ-NH2	11.06	1.47	1.33
1	A	181	ASP	CG-OD2	5.04	1.36	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	180	ARG	NE-CZ-NH1	-11.20	114.70	120.30
1	A	193	LEU	CA-CB-CG	5.08	126.98	115.30

There are no chirality outliers.

There are no planarity outliers.

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	2485	0	2440	14	0
2	A	2	0	0	0	0
3	A	1	0	0	0	0
4	A	4	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	4	0	3	0	0
6	A	18	0	24	0	0
7	A	153	0	0	3	0
All	All	2667	0	2467	15	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 3.

All (15) 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:47:CYS:HG	1:A:186:CYS:HG	0.96	0.95
4:A:2:CMO:C	7:A:461:HOH:O	2.39	0.71
1:A:85[B]:ARG:HG3	1:A:166:TYR:CE2	2.38	0.59
1:A:236:VAL:HG12	1:A:277:VAL:HG21	1.88	0.55
1:A:48:LEU:HB3	7:A:407:HOH:O	2.11	0.51
1:A:348:ALA:O	1:A:352:ILE:HG12	2.15	0.47
1:A:191:VAL:HG23	1:A:193:LEU:CD2	2.47	0.45
1:A:102:SER:HB2	1:A:105:THR:HG22	1.99	0.45
1:A:94:VAL:HG22	1:A:193:LEU:HD13	2.00	0.44
1:A:236:VAL:HG11	1:A:263:ILE:HG23	2.00	0.44
1:A:301:THR:HG23	1:A:302:GLU:HG3	1.99	0.43
1:A:267:ASN:HB3	1:A:270:LEU:HD22	2.00	0.43
1:A:284:THR:HG22	1:A:285[B]:PHE:CD2	2.53	0.43
1:A:48:LEU:HD21	7:A:460:HOH:O	2.21	0.41
1:A:262:LEU:HD11	1:A:265:ARG:HB2	2.03	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	319/314 (102%)	309 (97%)	10 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	276/269 (103%)	270 (98%)	6 (2%)	52	55

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	48	LEU
1	A	65	ASP
1	A	89	ASP
1	A	270	LEU
1	A	278[A]	GLU
1	A	278[B]	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	222	ASN
1	A	316	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 9 ligands modelled in this entry, 3 are monoatomic - leaving 6 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	#$|Z| > 2$
6	GOL	A	362	-	5,5,5	0.44	0	5,5,5	0.48	0
6	GOL	A	3	-	5,5,5	0.36	0	5,5,5	0.23	0
5	ACT	A	360	-	3,3,3	0.78	0	3,3,3	1.36	0
4	CMO	A	2	3	0,1,1	-	-	-		
4	CMO	A	1	2	0,1,1	-	-	-		
6	GOL	A	361	-	5,5,5	0.40	0	5,5,5	0.19	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
6	GOL	A	361	-	-	0/4/4/4	-
6	GOL	A	362	-	-	0/4/4/4	-
6	GOL	A	3	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	3	GOL	O1-C1-C2-O2
6	A	3	GOL	O1-C1-C2-C3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	2	CMO	1	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, 95th 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^2)	Q<0.9
1	A	310/314 (98%)	1.24	76 (24%)	0	0	42, 48, 54, 67	11 (3%)

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	285[A]	PHE	7.7
1	A	48	LEU	7.0
1	A	179	PHE	6.8
1	A	180	ARG	6.1
1	A	85[A]	ARG	5.3
1	A	116[A]	MET	5.3
1	A	259[A]	GLN	5.1
1	A	355	PRO	4.9
1	A	279[A]	HIS	4.9
1	A	57	LEU	4.7
1	A	76	SER	4.6
1	A	182	ASN	4.5
1	A	175	ASP	4.5
1	A	59	ALA	4.2
1	A	102	SER	4.2
1	A	75	GLU	4.2
1	A	344	ARG	4.1
1	A	58	ASP	3.9
1	A	143[A]	ARG	3.9
1	A	158	VAL	3.9
1	A	190[A]	SER	3.8
1	A	47	CYS	3.7
1	A	181	ASP	3.6
1	A	104	ASP	3.6
1	A	127	ASP	3.6
1	A	144	ASN	3.5
1	A	178	ALA	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	138	LEU	3.4
1	A	177	SER	3.4
1	A	183	HIS	3.3
1	A	319	ILE	3.3
1	A	317	LEU	3.3
1	A	46	GLU	3.3
1	A	94	VAL	3.3
1	A	202	ALA	3.2
1	A	341	ASP	3.2
1	A	129	GLY	3.2
1	A	134[A]	LYS	3.2
1	A	354	ILE	3.0
1	A	176	ILE	3.0
1	A	299	GLY	3.0
1	A	103	MET	2.9
1	A	219	LYS	2.9
1	A	125	PHE	2.9
1	A	193	LEU	2.9
1	A	228[A]	GLN	2.8
1	A	252	GLY	2.8
1	A	174	GLY	2.8
1	A	264	GLY	2.7
1	A	191	VAL	2.7
1	A	234	MET	2.7
1	A	80	PHE	2.7
1	A	184	LYS	2.7
1	A	167	PHE	2.7
1	A	105	THR	2.7
1	A	245	HIS	2.6
1	A	156	PHE	2.6
1	A	278[A]	GLU	2.6
1	A	73	PRO	2.6
1	A	352	ILE	2.5
1	A	79	TYR	2.5
1	A	74	LYS	2.5
1	A	345	THR	2.4
1	A	128	GLU	2.4
1	A	289	LEU	2.4
1	A	250	VAL	2.4
1	A	195	ARG	2.4
1	A	257	ASN	2.3
1	A	348	ALA	2.3

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	84[A]	MET	2.2
1	A	169	LEU	2.2
1	A	295	PHE	2.2
1	A	321	TYR	2.1
1	A	318	TYR	2.1
1	A	77	ASP	2.0
1	A	251	SER	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides 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. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	GOL	A	3	6/6	0.88	0.21	46,48,50,52	0
6	GOL	A	361	6/6	0.93	0.16	38,41,41,42	0
6	GOL	A	362	6/6	0.94	0.21	39,42,43,43	0
5	ACT	A	360	4/4	0.95	0.20	52,52,52,52	0
4	CMO	A	1	2/2	0.97	0.11	53,53,53,54	0
4	CMO	A	2	2/2	0.97	0.07	49,49,49,50	0
2	CU	A	357	1/1	0.98	0.07	55,55,55,55	1
3	NI	A	359	1/1	0.98	0.06	49,49,49,49	0
2	CU	A	358	1/1	0.99	0.03	51,51,51,51	0

6.5 Other polymers ⓘ

There are no such residues in this entry.