



Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 01:48 pm BST

PDB ID : 4M8U
Title : The Structure of MalL mutant enzyme V200A from Bacillus subtilus
Authors : Hobbs, J.K.; Jiao, W.; Easter, A.D.; Parker, E.J.; Schipper, L.A.; Arcus, V.L.
Deposited on : 2013-08-13
Resolution : 1.45 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

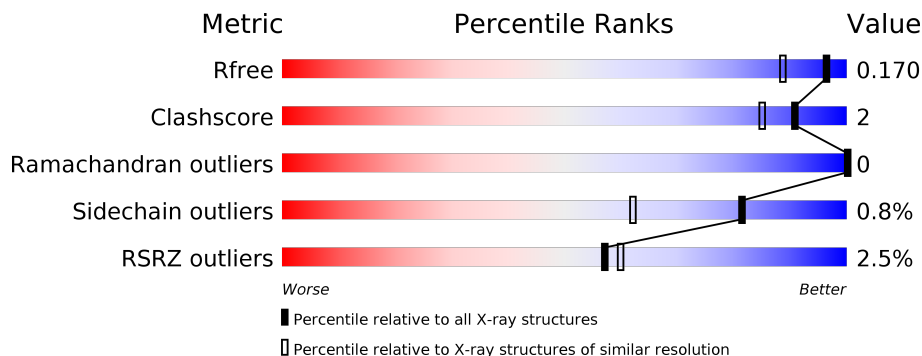
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.45 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}	130704	1156 (1.46-1.46)
Clashscore	141614	1202 (1.46-1.46)
Ramachandran outliers	138981	1178 (1.46-1.46)
Sidechain outliers	138945	1178 (1.46-1.46)
RSRZ outliers	127900	1139 (1.46-1.46)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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	560	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 5327 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.

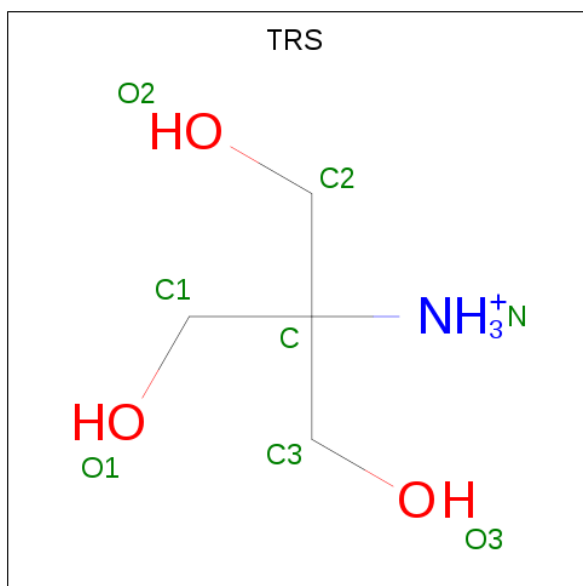
- Molecule 1 is a protein called Oligo-1,6-glucosidase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	559	4682	2990	779	888	25	8	11	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	200	ALA	VAL	ENGINEERED MUTATION	UNP O06994
A	?	-	GLN	DELETION	UNP O06994

- Molecule 2 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: $C_4H_{12}NO_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	8	4	1	3	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

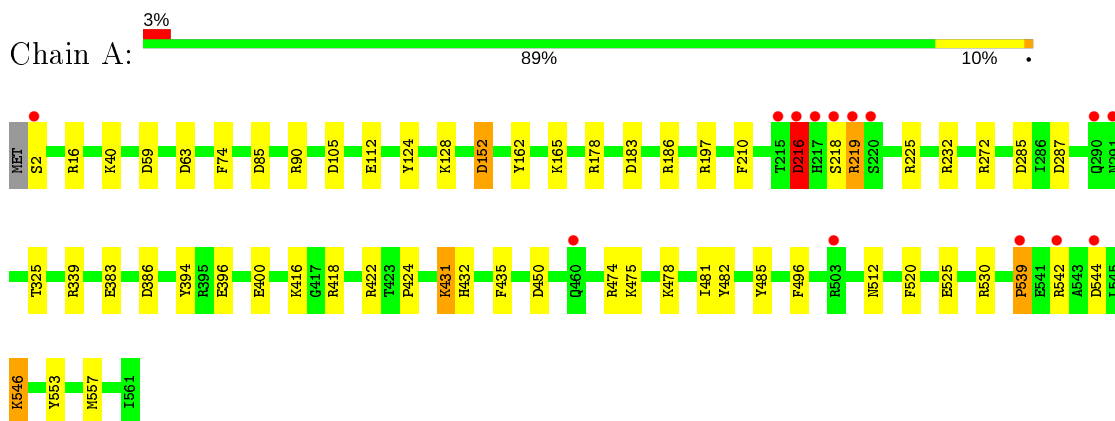
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	612	Total O 612 612	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: Oligo-1,6-glucosidase 1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	48.65Å 100.33Å 61.36Å 90.00° 112.59° 90.00°	Depositor
Resolution (Å)	27.28 – 1.45 27.26 – 1.45	Depositor EDS
% Data completeness (in resolution range)	99.5 (27.28-1.45) 99.5 (27.26-1.45)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.50 (at 1.45Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.138 , 0.164 0.146 , 0.170	Depositor DCC
R_{free} test set	4782 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	10.4	Xtrriage
Anisotropy	0.213	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 46.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.019 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	5327	wwPDB-VP
Average B, all atoms (Å ²)	12.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.05% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² 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: GOL, CA, TRS

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	1.19	15/4840 (0.3%)	1.22	46/6546 (0.7%)

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

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	478	LYS	CG-CD	-7.43	1.27	1.52
1	A	383	GLU	CD-OE1	6.74	1.33	1.25
1	A	525	GLU	CB-CG	-6.03	1.40	1.52
1	A	394	TYR	CZ-OH	5.99	1.48	1.37
1	A	124	TYR	CE1-CZ	-5.98	1.30	1.38
1	A	394	TYR	CD2-CE2	5.88	1.48	1.39
1	A	485	TYR	CE1-CZ	-5.87	1.30	1.38
1	A	542	ARG	CZ-NH1	5.68	1.40	1.33
1	A	542	ARG	C-O	5.67	1.34	1.23
1	A	525	GLU	CD-OE1	5.65	1.31	1.25
1	A	112	GLU	CD-OE2	5.62	1.31	1.25
1	A	475	LYS	CD-CE	5.46	1.64	1.51
1	A	383	GLU	CG-CD	5.18	1.59	1.51
1	A	416	LYS	CD-CE	-5.07	1.38	1.51
1	A	525	GLU	CD-OE2	5.02	1.31	1.25

All (46) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	285	ASP	CB-CG-OD2	-15.31	104.52	118.30
1	A	386	ASP	CB-CG-OD2	10.77	127.99	118.30
1	A	386	ASP	CB-CG-OD1	-9.96	109.33	118.30
1	A	542	ARG	NE-CZ-NH2	-9.29	115.66	120.30
1	A	542	ARG	NE-CZ-NH1	8.49	124.54	120.30
1	A	431	LYS	CD-CE-NZ	7.98	130.05	111.70
1	A	539	PRO	O-C-N	-7.63	110.50	122.70
1	A	285	ASP	OD1-CG-OD2	7.56	137.66	123.30
1	A	478	LYS	CB-CG-CD	7.51	131.12	111.60
1	A	418	ARG	NE-CZ-NH1	-7.40	116.60	120.30
1	A	197	ARG	NE-CZ-NH2	-7.17	116.72	120.30
1	A	394	TYR	CB-CG-CD1	-7.15	116.71	121.00
1	A	74	PHE	CB-CG-CD2	6.91	125.64	120.80
1	A	474	ARG	NE-CZ-NH2	6.84	123.72	120.30
1	A	339	ARG	NE-CZ-NH2	6.73	123.66	120.30
1	A	162	TYR	CB-CG-CD1	6.70	125.02	121.00
1	A	186	ARG	NE-CZ-NH2	-6.58	117.01	120.30
1	A	530	ARG	NE-CZ-NH1	6.54	123.57	120.30
1	A	396	GLU	OE1-CD-OE2	6.47	131.06	123.30
1	A	183[A]	ASP	CB-CG-OD1	-6.44	112.50	118.30
1	A	183[B]	ASP	CB-CG-OD1	-6.44	112.50	118.30
1	A	216	ASP	CB-CG-OD1	6.26	123.93	118.30
1	A	16	ARG	NE-CZ-NH2	6.25	123.42	120.30
1	A	225	ARG	NE-CZ-NH1	6.02	123.31	120.30
1	A	124	TYR	CB-CG-CD2	-5.98	117.41	121.00
1	A	178	ARG	NE-CZ-NH1	5.91	123.25	120.30
1	A	557[A]	MET	CG-SD-CE	5.86	109.57	100.20
1	A	557[B]	MET	CG-SD-CE	5.86	109.57	100.20
1	A	272	ARG	NE-CZ-NH1	5.82	123.21	120.30
1	A	383	GLU	CG-CD-OE1	5.74	129.77	118.30
1	A	152	ASP	CB-CG-OD1	-5.69	113.18	118.30
1	A	162	TYR	CB-CG-CD2	-5.55	117.67	121.00
1	A	59	ASP	CB-CG-OD1	5.54	123.28	118.30
1	A	232	ARG	NE-CZ-NH1	5.51	123.06	120.30
1	A	63	ASP	CB-CG-OD2	-5.50	113.35	118.30
1	A	90	ARG	NE-CZ-NH1	5.49	123.04	120.30
1	A	422	ARG	NE-CZ-NH2	-5.43	117.58	120.30
1	A	394	TYR	CG-CD2-CE2	-5.42	116.97	121.30
1	A	287	ASP	CB-CG-OD1	5.34	123.11	118.30
1	A	272	ARG	NE-CZ-NH2	-5.32	117.64	120.30
1	A	525	GLU	OE1-CD-OE2	5.27	129.63	123.30
1	A	520	PHE	CB-CG-CD2	-5.26	117.12	120.80
1	A	105	ASP	CB-CG-OD2	-5.22	113.60	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	339	ARG	NE-CZ-NH1	-5.14	117.73	120.30
1	A	530	ARG	NE-CZ-NH2	-5.06	117.77	120.30
1	A	450	ASP	CB-CG-OD2	-5.05	113.75	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	216	ASP	Peptide
1	A	539	PRO	Peptide

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	4682	0	4462	16	0
2	A	8	0	12	0	0
3	A	24	0	32	1	0
4	A	1	0	0	0	0
5	A	612	0	0	6	0
All	All	5327	0	4506	16	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 2.

All (16) 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:218:SER:OG	1:A:219:ARG:HD3	1.70	0.90
1:A:218:SER:OG	1:A:219:ARG:CD	2.33	0.76
1:A:165:LYS:HE2	5:A:1152:HOH:O	1.88	0.74
1:A:2:SER:N	1:A:482:TYR:HH	1.87	0.71
1:A:85[A]:ASP:OD1	5:A:1010:HOH:O	2.10	0.69
1:A:400:GLU:OE1	5:A:1250:HOH:O	2.11	0.67
1:A:325:THR:O	3:A:606:GOL:H2	2.05	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:481:ILE:HA	5:A:1124:HOH:O	2.07	0.55
1:A:544:ASP:OD1	1:A:546:LYS:HE2	2.07	0.55
1:A:431:LYS:HE3	1:A:432:HIS:CD2	2.43	0.52
1:A:128:LYS:HE3	1:A:210:PHE:CE2	2.49	0.47
1:A:512:ASN:O	1:A:553:TYR:HA	2.14	0.47
1:A:40:LYS:HD2	5:A:965:HOH:O	2.14	0.47
1:A:218:SER:O	1:A:219:ARG:HG3	2.18	0.43
1:A:152:ASP:OD2	5:A:1172:HOH:O	2.22	0.42
1:A:424:PRO:HB3	1:A:435:PHE:CG	2.55	0.41

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	568/560 (101%)	555 (98%)	13 (2%)	0	100 100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	505/500 (101%)	501 (99%)	4 (1%)	81 62

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	216	ASP
1	A	219	ARG
1	A	496	PHE
1	A	546	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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	$\# Z > 2$
3	GOL	A	606	-	5,5,5	1.00	1 (20%)	5,5,5	1.44	1 (20%)
3	GOL	A	605	-	5,5,5	0.94	0	5,5,5	0.54	0
2	TRS	A	601	-	7,7,7	0.73	0	9,9,9	0.84	0
3	GOL	A	602	-	5,5,5	0.61	0	5,5,5	0.60	0
3	GOL	A	604	-	5,5,5	0.81	0	5,5,5	0.65	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	A	606	-	-	2/4/4/4	-
3	GOL	A	605	-	-	2/4/4/4	-
2	TRS	A	601	-	-	2/9/9/9	-
3	GOL	A	602	-	-	0/4/4/4	-
3	GOL	A	604	-	-	0/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	606	GOL	O2-C2	-2.10	1.37	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	606	GOL	O2-C2-C3	-2.52	98.00	109.12

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	605	GOL	O1-C1-C2-C3
2	A	601	TRS	N-C-C1-O1
3	A	606	GOL	C1-C2-C3-O3
3	A	605	GOL	O1-C1-C2-O2
2	A	601	TRS	C2-C-C1-O1
3	A	606	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	606	GOL	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	559/560 (99%)	-0.01	14 (2%) 57 60	5, 11, 21, 43	6 (1%)

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	216	ASP	6.4
1	A	218	SER	6.0
1	A	2	SER	5.9
1	A	542	ARG	4.5
1	A	539	PRO	4.2
1	A	217	HIS	3.5
1	A	219	ARG	3.4
1	A	503	ARG	3.2
1	A	291	ASN	3.2
1	A	220	SER	2.9
1	A	544	ASP	2.6
1	A	215	THR	2.5
1	A	460	GLN	2.4
1	A	290	GLN	2.4

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 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. 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
3	GOL	A	605	6/6	0.89	0.12	13,21,23,26	0
3	GOL	A	606	6/6	0.89	0.12	14,16,20,21	0
3	GOL	A	604	6/6	0.91	0.18	12,17,20,22	0
3	GOL	A	602	6/6	0.95	0.08	12,13,14,15	0
4	CA	A	603	1/1	0.99	0.06	10,10,10,10	0
2	TRS	A	601	8/8	0.99	0.12	6,6,7,9	0

6.5 Other polymers [i](#)

There are no such residues in this entry.