



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 10:24 pm BST

PDB ID : 3HQ1  
Title : Crystal structure of Mycobacterium tuberculosis LeuA complexed with citrate and Mn<sup>2+</sup>  
Authors : Koon, N.; Squire, C.J.; Baker, E.N.  
Deposited on : 2009-06-05  
Resolution : 1.70 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

<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

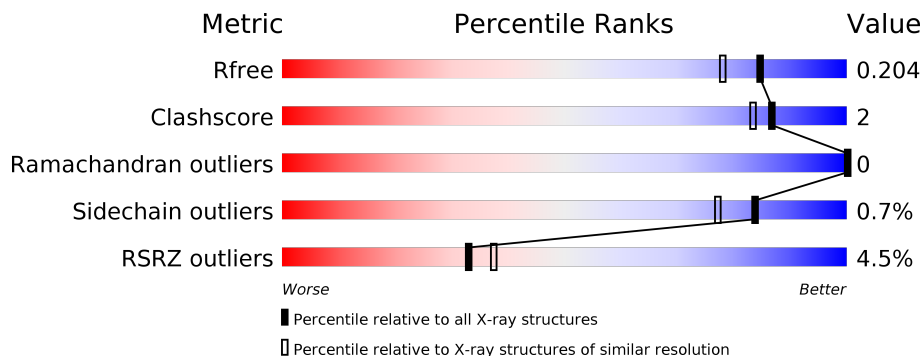
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 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	644	
1	B	644	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 9713 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 2-isopropylmalate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	573	4360	2749	748	843	20	0	5	0
1	B	576	4408	2779	764	843	22	0	9	0

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mn	0	0
			1	1		
2	A	1	Total	Mn	0	1
			2	2		

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

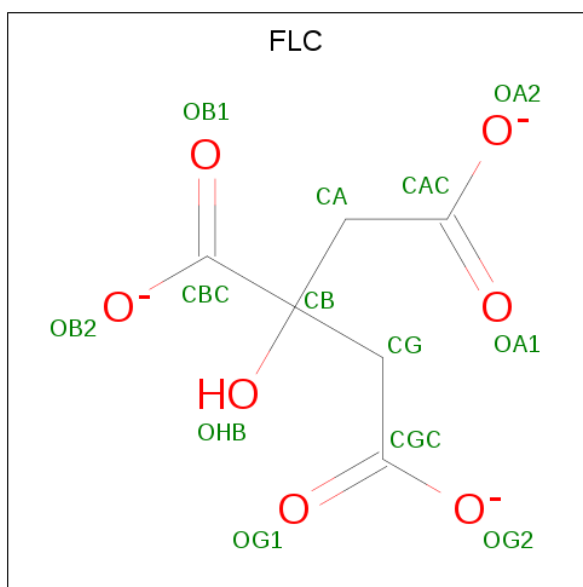
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Cl	0	0
			2	2		

- Molecule 4 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
4	A	1	Total	C O	0	0
			6	3 3		
4	B	1	Total	C O	0	0
			6	3 3		
4	B	1	Total	C O	0	0
			6	3 3		

- Molecule 5 is CITRATE ANION (three-letter code: FLC) (formula:  $C_6H_5O_7$ ).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	C O	0	1
			26	12 14		

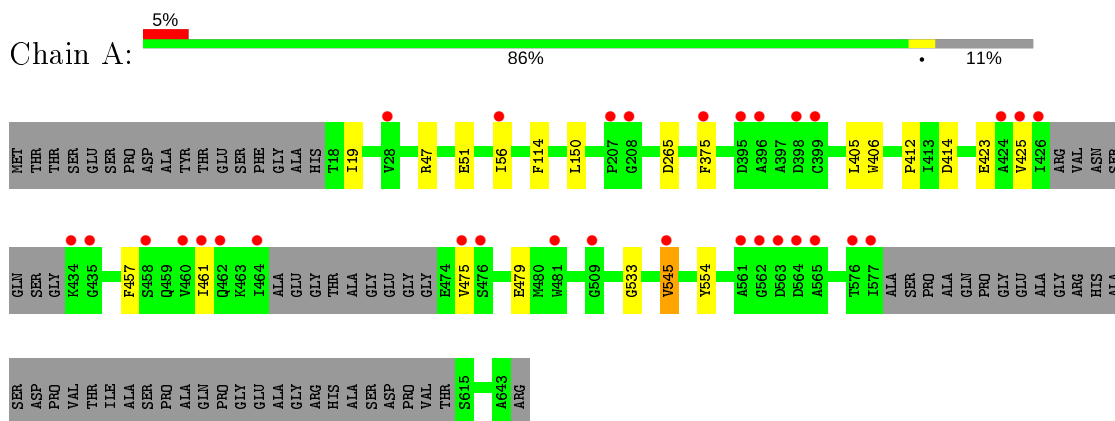
- Molecule 6 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	A	470	Total 470	O 470	0	0
6	B	426	Total 426	O 426	0	0

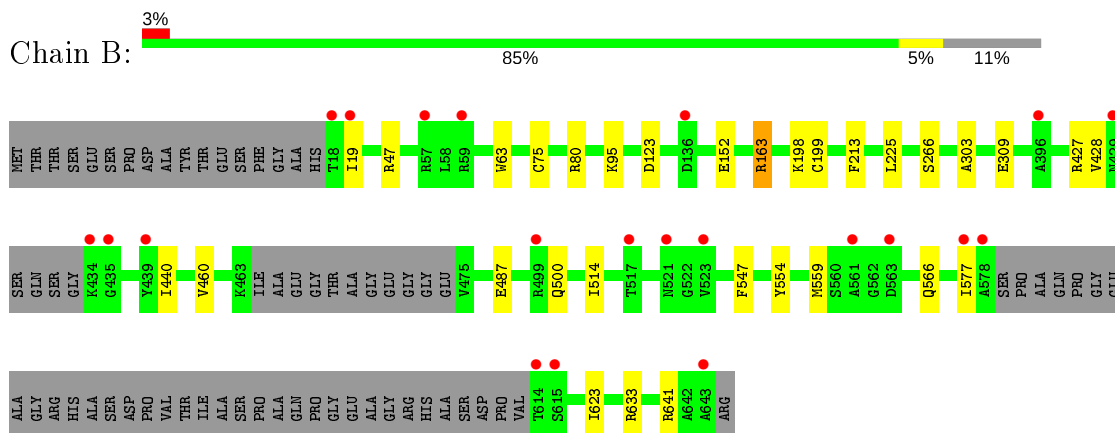
### 3 Residue-property plots [i](#)

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: 2-isopropylmalate synthase



- Molecule 1: 2-isopropylmalate synthase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.36Å 154.80Å 69.11Å 90.00° 98.04° 90.00°	Depositor
Resolution (Å)	38.69 – 1.70 38.45 – 1.70	Depositor EDS
% Data completeness (in resolution range)	94.7 (38.69-1.70) 94.7 (38.45-1.70)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.43 (at 1.70Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.170 , 0.204 0.170 , 0.204	Depositor DCC
$R_{free}$ test set	5974 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.3	Xtrriage
Anisotropy	0.065	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 49.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9713	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.45% 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: GOL, FLC, MN, CL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.57	0/4470	0.63	0/6102
1	B	0.57	0/4530	0.64	2/6180 (0.0%)
All	All	0.57	0/9000	0.64	2/12282 (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	B	163	ARG	NE-CZ-NH2	-5.56	117.52	120.30
1	B	163	ARG	CG-CD-NE	-5.15	100.98	111.80

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	4360	0	4163	15	0
1	B	4408	0	4234	22	0
2	A	2	0	0	0	0
2	B	1	0	0	0	0
3	A	2	0	0	1	0
4	A	6	0	8	0	0
4	B	12	0	16	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	26	0	8	1	0
6	A	470	0	0	5	0
6	B	426	0	0	4	0
All	All	9713	0	8429	36	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 (36) 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:19:ILE:HG12	6:B:776:HOH:O	1.91	0.69
6:A:786:HOH:O	1:B:19:ILE:HG12	1.93	0.68
1:B:559:MET:HE3	1:B:566:GLN:HB3	1.76	0.67
1:B:559:MET:CE	1:B:623:ILE:HG23	2.25	0.66
1:A:423:GLU:HG3	6:A:804:HOH:O	1.96	0.64
1:A:412:PRO:HG2	6:A:988:HOH:O	1.99	0.60
1:A:405:LEU:HD11	1:B:47:ARG:HG3	1.84	0.59
1:A:457:PHE:CE2	1:A:461:ILE:HD11	2.38	0.59
1:B:559:MET:CE	1:B:566:GLN:HB3	2.34	0.57
1:B:460:VAL:HG11	1:B:487:GLU:HG3	1.86	0.56
1:B:559:MET:HE1	1:B:623:ILE:HG23	1.88	0.56
1:B:500:GLN:HG3	1:B:514:ILE:HD11	1.89	0.55
1:A:47:ARG:HD2	1:A:51:GLU:HG2	1.89	0.54
1:B:487:GLU:OE2	1:B:641:ARG:NH1	2.42	0.52
1:B:152:GLU:HG3	1:B:198:LYS:HE3	1.95	0.49
1:A:406:TRP:CE3	1:A:414:ASP:HB2	2.49	0.48
1:B:95:LYS:HE2	6:B:962:HOH:O	2.14	0.48
1:A:375:PHE:CE2	1:A:425:VAL:HB	2.49	0.47
1:B:559:MET:HE3	1:B:623:ILE:HG23	1.97	0.46
1:B:63:TRP:HB3	1:B:303:ALA:HA	1.97	0.46
1:B:75:CYS:HB3	1:B:309:GLU:HG2	1.98	0.46
1:A:475:VAL:HG13	1:A:479:GLU:HB3	1.98	0.45
1:B:123:ASP:OD2	6:B:718:HOH:O	2.20	0.45
1:A:545:VAL:O	1:A:545:VAL:HG13	2.17	0.45
1:B:225:LEU:HD21	1:B:266:SER:HB2	2.00	0.44
1:A:423:GLU:CG	6:A:804:HOH:O	2.60	0.43
1:B:428:VAL:HG11	1:B:440:ILE:HD12	2.00	0.43
1:B:559:MET:HE3	1:B:566:GLN:CB	2.45	0.43
1:A:56:ILE:HD13	1:A:265:ASP:OD2	2.18	0.43
1:B:199:CYS:HB3	1:B:213:PHE:CE2	2.54	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:547:PHE:HD2	1:B:577:ILE:HG12	1.84	0.42
1:B:80:ARG:NH1	5:B:705[B]:FLC:OA2	2.50	0.42
1:A:545:VAL:O	1:A:545:VAL:CG1	2.68	0.41
1:A:150:LEU:HD13	6:A:1046:HOH:O	2.20	0.40
1:B:633:ARG:NE	6:B:751:HOH:O	2.54	0.40
1:A:533:GLY:HA3	3:A:703:CL:CL	2.58	0.40

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	570/644 (88%)	559 (98%)	11 (2%)	0	100	100
1	B	577/644 (90%)	565 (98%)	12 (2%)	0	100	100
All	All	1147/1288 (89%)	1124 (98%)	23 (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	447/520 (86%)	444 (99%)	3 (1%)	84	77
1	B	453/520 (87%)	450 (99%)	3 (1%)	84	77

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	900/1040 (86%)	894 (99%)	6 (1%)	84	77

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

Mol	Chain	Res	Type
1	A	114	PHE
1	A	545	VAL
1	A	554	TYR
1	B	163	ARG
1	B	427	ARG
1	B	554	TYR

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 10 ligands modelled in this entry, 5 are 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
4	GOL	B	708	-	5,5,5	0.32	0	5,5,5	0.82	0
4	GOL	A	706	-	5,5,5	0.29	0	5,5,5	0.26	0
5	FLC	B	705[A]	2	3,12,12	0.87	0	3,17,17	1.31	0
4	GOL	B	707	-	5,5,5	0.28	0	5,5,5	0.36	0
5	FLC	B	705[B]	2	3,12,12	1.03	0	3,17,17	1.59	1 (33%)

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	GOL	B	708	-	-	2/4/4/4	-
4	GOL	A	706	-	-	0/4/4/4	-
5	FLC	B	705[A]	2	-	0/6/16/16	-
4	GOL	B	707	-	-	0/4/4/4	-
5	FLC	B	705[B]	2	-	2/6/16/16	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	705[B]	FLC	CB-CG-CGC	-2.39	111.15	114.98

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	708	GOL	O1-C1-C2-C3
4	B	708	GOL	O1-C1-C2-O2
5	B	705[B]	FLC	CBC-CB-CG-CGC
5	B	705[B]	FLC	OHB-CB-CG-CGC

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	705[B]	FLC	1	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	573/644 (88%)	0.07	31 (5%) 25 28	7, 17, 35, 56	0
1	B	576/644 (89%)	-0.02	21 (3%) 42 47	7, 17, 33, 42	0
All	All	1149/1288 (89%)	0.03	52 (4%) 33 37	7, 17, 34, 56	0

All (52) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	461	ILE	6.0
1	A	426	ILE	5.2
1	A	435	GLY	4.9
1	B	614	THR	4.7
1	A	464	ILE	4.4
1	B	18	THR	4.4
1	B	434	LYS	4.3
1	B	429	ASN	4.1
1	A	399	CYS	4.1
1	A	398	ASP	4.1
1	A	565	ALA	3.9
1	B	57	ARG	3.8
1	A	434	LYS	3.7
1	A	425	VAL	3.5
1	B	396	ALA	3.4
1	A	458	SER	3.4
1	A	28	VAL	3.3
1	A	562	GLY	3.1
1	B	578	ALA	3.1
1	B	643	ALA	3.0
1	A	576	THR	3.0
1	A	475	VAL	2.9
1	B	563	ASP	2.8
1	A	56	ILE	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	545	VAL	2.8
1	A	208	GLY	2.8
1	A	395	ASP	2.8
1	B	435	GLY	2.7
1	A	396	ALA	2.7
1	B	521	ASN	2.7
1	B	561	ALA	2.7
1	A	564	ASP	2.6
1	A	577	ILE	2.5
1	A	462	GLN	2.5
1	B	577	ILE	2.5
1	B	615	SER	2.3
1	A	460	VAL	2.3
1	B	517	THR	2.2
1	A	509	GLY	2.2
1	B	19	ILE	2.2
1	A	207	PRO	2.2
1	A	563	ASP	2.2
1	A	476	SER	2.1
1	A	424	ALA	2.1
1	B	136	ASP	2.1
1	B	59	ARG	2.1
1	B	523	VAL	2.1
1	B	439	TYR	2.1
1	A	375	PHE	2.1
1	A	481	TRP	2.1
1	A	561	ALA	2.0
1	B	499	ARG	2.0

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

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

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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	B-factors(Å <sup>2</sup> )	Q<0.9
4	GOL	B	708	6/6	0.84	0.13	35,36,37,38	0
3	CL	A	704	1/1	0.95	0.07	45,45,45,45	0
4	GOL	A	706	6/6	0.95	0.09	14,15,17,18	0
5	FLC	B	705[A]	13/13	0.96	0.10	10,15,16,16	13
5	FLC	B	705[B]	13/13	0.96	0.10	11,15,16,16	13
3	CL	A	703	1/1	0.97	0.06	27,27,27,27	0
4	GOL	B	707	6/6	0.97	0.07	12,15,16,16	0
2	MN	A	701[A]	1/1	0.98	0.12	10,10,10,10	1
2	MN	A	701[B]	1/1	0.98	0.12	12,12,12,12	1
2	MN	B	702	1/1	1.00	0.05	12,12,12,12	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.