



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 11, 2018 – 11:41 PM EST

PDB ID : 3N4F  
Title : CRYSTAL STRUCTURE OF Mandelate racemase/muconate lactonizing protein from *Geobacillus* sp. Y412MC10  
Authors : Malashkevich, V.N.; Toro, R.; Sauder, J.M.; Burley, S.K.; Almo, S.C.; New York SGX Research Center for Structural Genomics (NYSGXRC)  
Deposited on : 2010-05-21  
Resolution : 1.88 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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
Xtriage (Phenix)	:	1.13
EDS	:	rb-20031633
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20031633

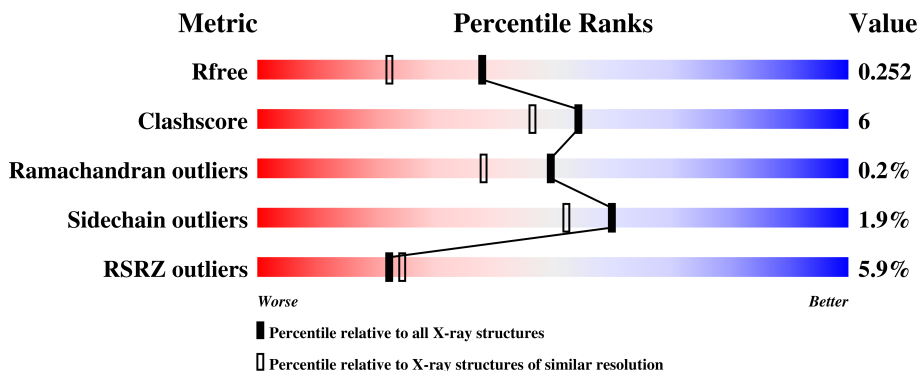
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.88 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))
$R_{free}$	111664	8255 (1.90-1.86)
Clashscore	122126	9028 (1.90-1.86)
Ramachandran outliers	120053	8930 (1.90-1.86)
Sidechain outliers	120020	8930 (1.90-1.86)
RSRZ outliers	108989	8087 (1.90-1.86)

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	392	<div> <div>5%</div> <div>87%</div> <div>11% ..</div> </div>
1	B	392	<div> <div>6%</div> <div>86%</div> <div>11% ..</div> </div>
1	C	392	<div> <div>3%</div> <div>88%</div> <div>9% ..</div> </div>
1	D	392	<div> <div>9%</div> <div>84%</div> <div>11% ..</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 13659 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 Mandelate racemase/muconate lactonizing protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	390	Total	C	N	O	S	Se	0	5	0
			3111	1970	550	581	3	7			
1	B	386	Total	C	N	O	S	Se	0	4	0
			3069	1941	540	577	3	8			
1	C	386	Total	C	N	O	S	Se	0	3	0
			3061	1940	537	574	3	7			
1	D	376	Total	C	N	O	S	Se	0	4	0
			3001	1905	525	561	3	7			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	-	expression tag	UNP D3EID5
A	2	SER	-	expression tag	UNP D3EID5
A	3	LEU	-	expression tag	UNP D3EID5
A	385	GLU	-	expression tag	UNP D3EID5
A	386	GLY	-	expression tag	UNP D3EID5
A	387	HIS	-	expression tag	UNP D3EID5
A	388	HIS	-	expression tag	UNP D3EID5
A	389	HIS	-	expression tag	UNP D3EID5
A	390	HIS	-	expression tag	UNP D3EID5
A	391	HIS	-	expression tag	UNP D3EID5
A	392	HIS	-	expression tag	UNP D3EID5
B	1	MSE	-	expression tag	UNP D3EID5
B	2	SER	-	expression tag	UNP D3EID5
B	3	LEU	-	expression tag	UNP D3EID5
B	385	GLU	-	expression tag	UNP D3EID5
B	386	GLY	-	expression tag	UNP D3EID5
B	387	HIS	-	expression tag	UNP D3EID5
B	388	HIS	-	expression tag	UNP D3EID5
B	389	HIS	-	expression tag	UNP D3EID5
B	390	HIS	-	expression tag	UNP D3EID5
B	391	HIS	-	expression tag	UNP D3EID5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	392	HIS	-	expression tag	UNP D3EID5
C	1	MSE	-	expression tag	UNP D3EID5
C	2	SER	-	expression tag	UNP D3EID5
C	3	LEU	-	expression tag	UNP D3EID5
C	385	GLU	-	expression tag	UNP D3EID5
C	386	GLY	-	expression tag	UNP D3EID5
C	387	HIS	-	expression tag	UNP D3EID5
C	388	HIS	-	expression tag	UNP D3EID5
C	389	HIS	-	expression tag	UNP D3EID5
C	390	HIS	-	expression tag	UNP D3EID5
C	391	HIS	-	expression tag	UNP D3EID5
C	392	HIS	-	expression tag	UNP D3EID5
D	1	MSE	-	expression tag	UNP D3EID5
D	2	SER	-	expression tag	UNP D3EID5
D	3	LEU	-	expression tag	UNP D3EID5
D	385	GLU	-	expression tag	UNP D3EID5
D	386	GLY	-	expression tag	UNP D3EID5
D	387	HIS	-	expression tag	UNP D3EID5
D	388	HIS	-	expression tag	UNP D3EID5
D	389	HIS	-	expression tag	UNP D3EID5
D	390	HIS	-	expression tag	UNP D3EID5
D	391	HIS	-	expression tag	UNP D3EID5
D	392	HIS	-	expression tag	UNP D3EID5

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total Mg 2 2	0	0
2	A	2	Total Mg 2 2	0	0
2	D	2	Total Mg 2 2	0	0
2	C	2	Total Mg 2 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	394	Total O 394 394	0	0

*Continued on next page...*

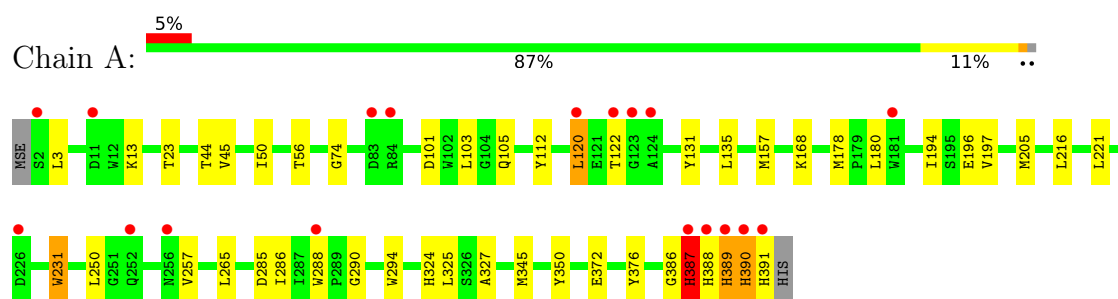
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	328	Total 328	O 328	0	0
3	C	450	Total 450	O 450	0	0
3	D	237	Total 237	O 237	0	0

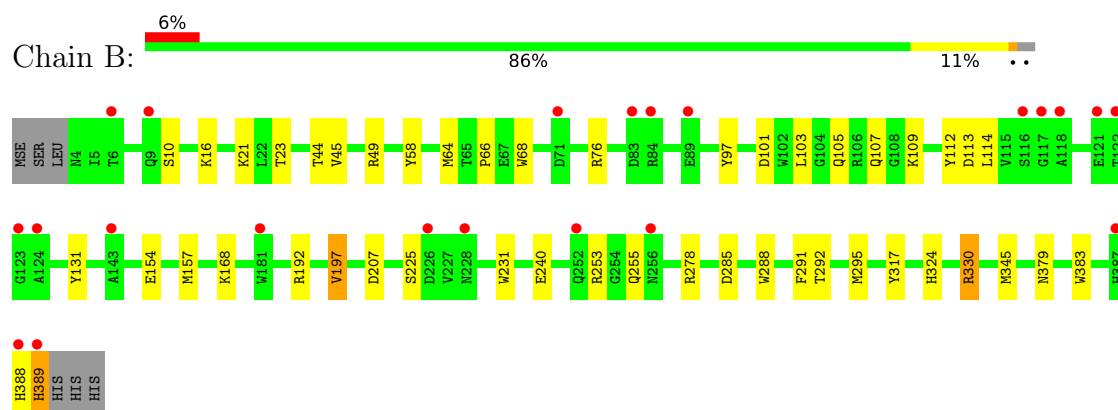
### 3 Residue-property plots [i](#)

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.

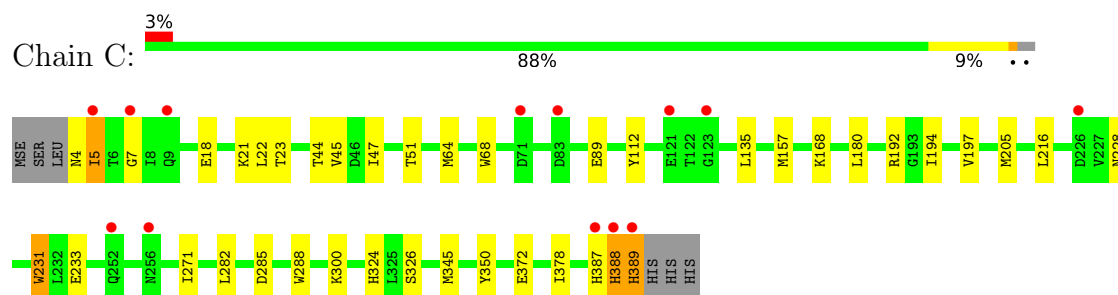
- Molecule 1: Mandelate racemase/muconate lactonizing protein



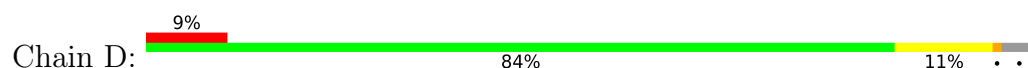
- Molecule 1: Mandelate racemase/muconate lactonizing protein

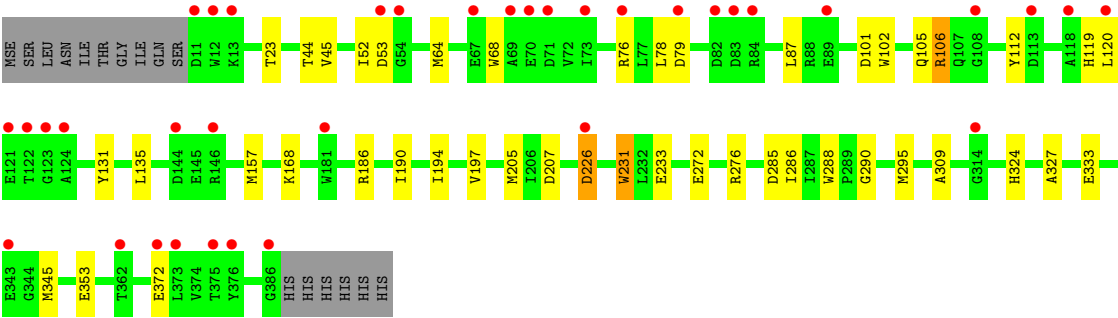


- Molecule 1: Mandelate racemase/muconate lactonizing protein



- Molecule 1: Mandelate racemase/muconate lactonizing protein





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	155.39Å 66.10Å 151.54Å 90.00° 94.33° 90.00°	Depositor
Resolution (Å)	8.00 – 1.88 8.00 – 1.88	Depositor EDS
% Data completeness (in resolution range)	93.1 (8.00-1.88) 93.1 (8.00-1.88)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	19.12 (at 1.88Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.185 , 0.242 0.194 , 0.252	Depositor DCC
$R_{free}$ test set	5793 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	15.1	Xtriage
Anisotropy	0.253	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.50 , 57.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13659	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 66.71 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.7857e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: MG

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	0.58	0/3194	0.64	1/4317 (0.0%)
1	B	0.54	1/3147 (0.0%)	0.64	2/4253 (0.0%)
1	C	0.59	0/3136	0.65	0/4240
1	D	0.44	0/3079	0.57	0/4162
All	All	0.54	1/12556 (0.0%)	0.63	3/16972 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	330	ARG	CZ-NH2	-5.51	1.25	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	330	ARG	NE-CZ-NH1	9.88	125.24	120.30
1	B	330	ARG	NE-CZ-NH2	-9.71	115.44	120.30
1	A	389	HIS	N-CA-C	5.46	125.73	111.00

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	3111	0	3012	40	1
1	B	3069	0	2962	33	0
1	C	3061	0	2963	35	0
1	D	3001	0	2901	31	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
3	A	394	0	0	8	0
3	B	328	0	0	7	0
3	C	450	0	0	7	0
3	D	237	0	0	7	0
All	All	13659	0	11838	136	1

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 6.

The worst 5 of 136 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:B:157[B]:MSE:HE3	1:B:197:VAL:HG13	1.36	1.06
1:C:345:MSE:SE	3:C:1613:HOH:O	2.23	1.05
1:B:345:MSE:SE	3:B:1560:HOH:O	2.26	1.02
1:A:390:HIS:ND1	1:A:391:HIS:CD2	2.30	0.99
1:A:345:MSE:SE	3:A:494:HOH:O	2.35	0.93

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:376:TYR:OH	1:A:376:TYR:OH[2_555]	1.80	0.40

## 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	393/392 (100%)	382 (97%)	9 (2%)	2 (0%)	31	18
1	B	388/392 (99%)	379 (98%)	9 (2%)	0	100	100
1	C	387/392 (99%)	380 (98%)	6 (2%)	1 (0%)	43	31
1	D	378/392 (96%)	372 (98%)	6 (2%)	0	100	100
All	All	1546/1568 (99%)	1513 (98%)	30 (2%)	3 (0%)	49	38

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	387	HIS
1	C	388	HIS
1	A	390	HIS

### 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	319/308 (104%)	313 (98%)	6 (2%)	60	52
1	B	314/308 (102%)	307 (98%)	7 (2%)	55	46
1	C	313/308 (102%)	309 (99%)	4 (1%)	71	67
1	D	305/308 (99%)	298 (98%)	7 (2%)	53	44
All	All	1251/1232 (102%)	1227 (98%)	24 (2%)	60	52

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	225	SER
1	C	4	ASN
1	D	231	TRP
1	B	231	TRP
1	B	389	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	324	HIS
1	D	324	HIS
1	C	9	GLN
1	A	391	HIS
1	C	4	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.

### 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

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.

### 5.7 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	383/392 (97%)	0.17	18 (4%) 31 33	6, 14, 36, 58	0
1	B	379/392 (96%)	0.32	22 (5%) 23 25	8, 17, 37, 51	0
1	C	379/392 (96%)	0.05	13 (3%) 45 47	6, 12, 29, 49	0
1	D	369/392 (94%)	0.54	36 (9%) 7 8	11, 25, 50, 61	0
All	All	1510/1568 (96%)	0.27	89 (5%) 22 24	6, 17, 41, 61	0

The worst 5 of 89 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	388	HIS	6.5
1	B	181	TRP	6.1
1	A	123	GLY	5.0
1	C	7	GLY	5.0
1	B	123	GLY	4.9

### 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( $\text{\AA}^2$ )	Q<0.9
2	MG	D	501	1/1	0.99	0.07	13,13,13,13	0
2	MG	A	502	1/1	0.99	0.05	10,10,10,10	0
2	MG	B	502	1/1	0.99	0.06	15,15,15,15	0
2	MG	C	501	1/1	0.99	0.10	9,9,9,9	0
2	MG	A	501	1/1	0.99	0.05	8,8,8,8	0
2	MG	C	502	1/1	1.00	0.04	8,8,8,8	0
2	MG	B	501	1/1	1.00	0.08	13,13,13,13	0
2	MG	D	502	1/1	1.00	0.03	14,14,14,14	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.