



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 09:38 am GMT

PDB ID : 1BL3  
Title : CATALYTIC DOMAIN OF HIV-1 INTEGRASE  
Authors : Maignan, S.; Guilloteau, J.P.; Zhou-Liu, Q.; Clement-Mella, C.; Mikol, V.  
Deposited on : 1998-07-23  
Resolution : 2.00 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  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
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) : recalc28949

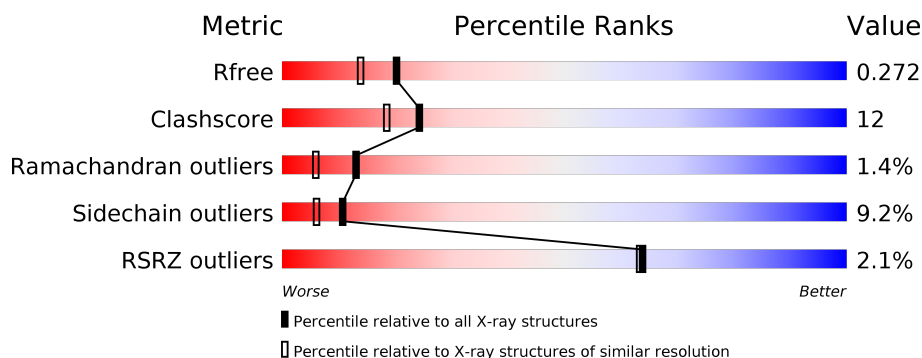
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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	160	<div> <div>%</div> <div> <div></div> <div>68%</div> <div>19%</div> <div>•</div> <div>11%</div> </div> </div>
1	B	160	<div> <div>4%</div> <div> <div></div> <div>63%</div> <div>24%</div> <div>•</div> <div>11%</div> </div> </div>
1	C	160	<div> <div>%</div> <div> <div></div> <div>73%</div> <div>25%</div> <div>•</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3593 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 INTEGRASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	142	Total	C	N	O	S	25	0	0
			1096	700	191	201	4			
1	B	143	Total	C	N	O	S	64	0	0
			1109	708	193	204	4			
1	C	160	Total	C	N	O	S	55	0	0
			1235	784	216	229	6			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	185	HIS	PHE	ENGINEERED	UNP P12497
B	185	HIS	PHE	ENGINEERED	UNP P12497
C	185	HIS	PHE	ENGINEERED	UNP P12497

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

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	58	Total	O	0	0
			58	58		
3	B	35	Total	O	0	0
			35	35		

*Continued on next page...*

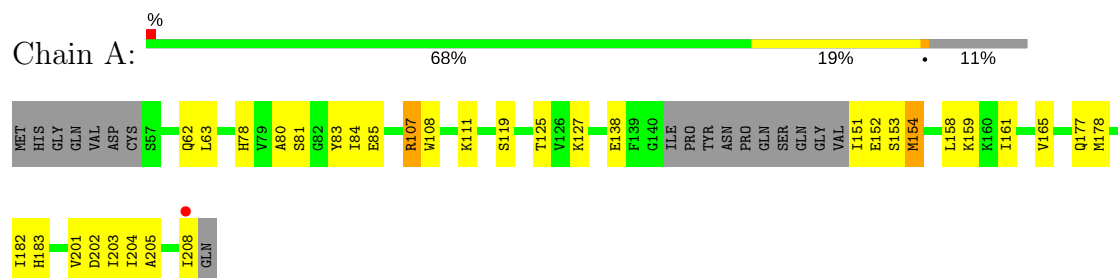
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	57	Total	O	0	0
			57	57		

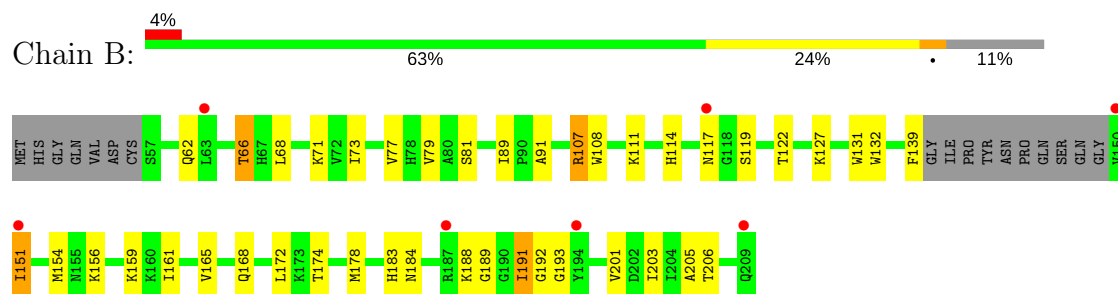
### 3 Residue-property plots [i](#)

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.

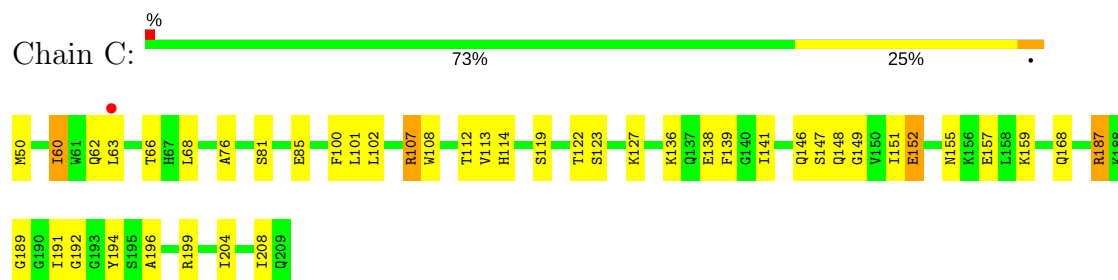
#### • Molecule 1: INTEGRASE



#### • Molecule 1: INTEGRASE



#### • Molecule 1: INTEGRASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.10Å 72.50Å 79.40Å 90.00° 108.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.00 19.80 – 2.00	Depositor EDS
% Data completeness (in resolution range)	95.7 (20.00-2.00) 93.7 (19.80-2.00)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.76 (at 2.01Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.216 , 0.279 0.211 , 0.272	Depositor DCC
$R_{free}$ test set	2383 reflections (7.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	30.1	Xtriage
Anisotropy	0.399	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 80.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3593	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.87% 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: 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	0.39	0/1118	0.61	0/1512
1	B	0.36	0/1131	0.58	0/1529
1	C	0.39	0/1262	0.62	0/1709
All	All	0.38	0/3511	0.60	0/4750

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	1096	0	1101	22	0
1	B	1109	0	1115	36	0
1	C	1235	0	1228	30	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	58	0	0	0	0
3	B	35	0	0	2	0
3	C	57	0	0	2	0
All	All	3593	0	3444	81	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 12.

All (81) 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:77:VAL:HG23	1:B:151:ILE:HD12	1.34	1.03
1:A:84:ILE:HD11	1:A:154:MET:HG3	1.58	0.85
1:A:205:ALA:HB2	1:B:201:VAL:HG12	1.59	0.82
1:B:62:GLN:NE2	1:B:151:ILE:HG22	2.01	0.76
1:A:107:ARG:HD3	1:B:107:ARG:HD3	1.69	0.74
1:C:50:MET:HE1	1:C:157:GLU:HG2	1.68	0.74
1:A:83:TYR:OH	1:B:107:ARG:HB2	1.89	0.73
1:B:165:VAL:HB	1:B:178:MET:HE2	1.76	0.67
1:C:68:LEU:HD13	1:C:159:LYS:HD3	1.77	0.67
1:A:165:VAL:HG11	1:A:182:ILE:CD1	2.26	0.65
1:B:165:VAL:HB	1:B:178:MET:CE	2.28	0.64
1:A:62:GLN:NE2	1:A:151:ILE:HB	2.15	0.61
1:B:165:VAL:O	1:B:168:GLN:HB2	2.00	0.61
1:B:107:ARG:HG3	1:B:108:TRP:CD1	2.36	0.60
1:B:77:VAL:CG2	1:B:151:ILE:HD12	2.22	0.60
1:C:81:SER:O	1:C:199:ARG:NH1	2.35	0.59
1:C:107:ARG:HG2	1:C:108:TRP:CD1	2.37	0.59
1:A:127:LYS:HE3	1:C:141:ILE:O	2.03	0.58
1:B:62:GLN:HE22	1:B:151:ILE:H	1.50	0.57
1:A:201:VAL:HG12	1:B:205:ALA:HB2	1.85	0.57
1:B:62:GLN:NE2	1:B:151:ILE:H	2.02	0.57
1:B:117:ASN:OD1	1:B:139:PHE:HB3	2.06	0.56
1:A:165:VAL:HG11	1:A:182:ILE:HD12	1.87	0.56
1:C:66:THR:OG1	1:C:159:LYS:HE2	2.06	0.56
1:A:204:ILE:O	1:A:208:ILE:HG23	2.07	0.55
1:B:62:GLN:HE21	1:B:151:ILE:HG22	1.71	0.55
1:B:62:GLN:HE21	1:B:151:ILE:CG2	2.20	0.55
1:C:155:ASN:O	1:C:159:LYS:HG3	2.07	0.55
1:A:78:HIS:CE1	1:A:80:ALA:HB3	2.42	0.54
1:C:60:ILE:HD13	1:C:60:ILE:N	2.23	0.54
1:B:89:ILE:HD12	1:B:91:ALA:O	2.07	0.54
1:C:187:ARG:HG3	1:C:187:ARG:HH11	1.73	0.54
1:A:62:GLN:HE21	1:A:151:ILE:HB	1.73	0.53
1:B:66:THR:HG22	1:B:73:ILE:HB	1.91	0.53
1:B:114:HIS:HB3	3:B:525:HOH:O	2.09	0.53
1:B:168:GLN:HG3	3:B:533:HOH:O	2.09	0.53
1:C:122:THR:HG22	1:C:139:PHE:CE2	2.45	0.52

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:149:GLY:O	1:C:152:GLU:HB3	2.09	0.52
1:A:107:ARG:HG2	1:A:108:TRP:CD1	2.44	0.52
1:A:178:MET:HE3	1:B:132:TRP:CE2	2.45	0.51
1:C:60:ILE:HD12	1:C:112:THR:HB	1.93	0.51
1:A:161:ILE:HD12	1:A:183:HIS:HB2	1.93	0.51
1:B:122:THR:HG22	1:B:139:PHE:CE2	2.46	0.51
1:A:201:VAL:CG1	1:B:205:ALA:HB2	2.42	0.49
1:C:85:GLU:OE1	1:C:107:ARG:NH1	2.45	0.49
1:A:84:ILE:HD12	1:A:183:HIS:CD2	2.48	0.48
1:C:76:ALA:HB2	1:C:100:PHE:HE1	1.79	0.48
1:B:81:SER:HB3	1:B:203:ILE:HD12	1.95	0.48
1:C:196:ALA:HA	1:C:199:ARG:NH1	2.28	0.48
1:B:68:LEU:HD23	1:B:172:LEU:HD11	1.96	0.47
1:B:154:MET:HE2	1:B:154:MET:HB3	1.77	0.47
1:A:84:ILE:CG2	1:A:85:GLU:N	2.77	0.46
1:C:187:ARG:CG	1:C:187:ARG:HH11	2.28	0.46
1:B:68:LEU:HD13	1:B:159:LYS:HE2	1.98	0.45
1:A:177:GLN:OE1	1:A:177:GLN:HA	2.17	0.45
1:B:127:LYS:HE3	1:B:131:TRP:CZ2	2.52	0.44
1:C:85:GLU:OE2	1:C:107:ARG:HD2	2.17	0.44
1:C:62:GLN:HG2	1:C:114:HIS:HB2	1.97	0.44
1:C:204:ILE:O	1:C:208:ILE:HG13	2.17	0.44
1:C:187:ARG:NH2	3:C:557:HOH:O	2.50	0.43
1:A:81:SER:HB3	1:A:203:ILE:HD12	1.99	0.43
1:B:151:ILE:O	1:B:151:ILE:HG13	2.19	0.43
1:A:84:ILE:CD1	1:A:154:MET:HG3	2.38	0.42
1:C:50:MET:N	3:C:514:HOH:O	2.52	0.42
1:C:62:GLN:NE2	1:C:151:ILE:H	2.17	0.42
1:C:68:LEU:HA	1:C:68:LEU:HD12	1.83	0.42
1:B:183:HIS:HD2	1:B:184:ASN:OD1	2.03	0.41
1:C:147:SER:O	1:C:148:GLN:CB	2.68	0.41
1:C:63:LEU:HD13	1:C:113:VAL:HG13	2.03	0.41
1:A:84:ILE:HD11	1:A:154:MET:CG	2.40	0.41
1:B:62:GLN:NE2	1:B:151:ILE:CG2	2.76	0.41
1:C:50:MET:CE	1:C:157:GLU:HG2	2.46	0.41
1:B:161:ILE:O	1:B:165:VAL:HG22	2.21	0.41
1:C:123:SER:O	1:C:127:LYS:HG3	2.20	0.41
1:B:107:ARG:CG	1:B:108:TRP:CD1	3.03	0.40
1:B:62:GLN:HE21	1:B:151:ILE:CB	2.34	0.40
1:C:147:SER:C	1:C:149:GLY:H	2.25	0.40
1:C:147:SER:O	1:C:148:GLN:HB2	2.21	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:136:LYS:HD3	1:C:138:GLU:OE2	2.21	0.40
1:B:111:LYS:HD3	1:B:111:LYS:HA	1.88	0.40
1:B:62:GLN:HG3	1:B:79:VAL:CG2	2.52	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	138/160 (86%)	137 (99%)	1 (1%)	0	100	100
1	B	139/160 (87%)	129 (93%)	6 (4%)	4 (3%)	5	1
1	C	158/160 (99%)	150 (95%)	6 (4%)	2 (1%)	14	7
All	All	435/480 (91%)	416 (96%)	13 (3%)	6 (1%)	13	6

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	189	GLY
1	B	189	GLY
1	B	192	GLY
1	C	192	GLY
1	B	193	GLY
1	B	191	ILE

### 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	114/130 (88%)	102 (90%)	12 (10%)	8	4
1	B	116/130 (89%)	106 (91%)	10 (9%)	12	7
1	C	130/130 (100%)	119 (92%)	11 (8%)	12	7
All	All	360/390 (92%)	327 (91%)	33 (9%)	11	6

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	63	LEU
1	A	107	ARG
1	A	111	LYS
1	A	119	SER
1	A	125	THR
1	A	138	GLU
1	A	152	GLU
1	A	153	SER
1	A	154	MET
1	A	158	LEU
1	A	159	LYS
1	A	202	ASP
1	B	66	THR
1	B	71	LYS
1	B	107	ARG
1	B	119	SER
1	B	151	ILE
1	B	156	LYS
1	B	174	THR
1	B	188	LYS
1	B	191	ILE
1	B	206	THR
1	C	60	ILE
1	C	101	LEU
1	C	102	LEU
1	C	107	ARG
1	C	119	SER
1	C	146	GLN
1	C	152	GLU
1	C	168	GLN
1	C	187	ARG
1	C	191	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	194	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	62	GLN
1	A	67	HIS
1	A	164	GLN
1	B	62	GLN
1	B	67	HIS
1	B	183	HIS
1	C	62	GLN
1	C	67	HIS
1	C	137	GLN
1	C	164	GLN
1	C	185	HIS

### 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 3 ligands modelled in this entry, 3 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 [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 [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	142/160 (88%)	-0.29	1 (0%) 87 87	17, 32, 51, 67	10 (7%)
1	B	138/160 (86%)	0.03	7 (5%) 29 29	21, 36, 66, 76	13 (9%)
1	C	155/160 (96%)	-0.26	1 (0%) 89 88	17, 28, 46, 67	10 (6%)
All	All	435/480 (90%)	-0.18	9 (2%) 64 63	17, 32, 55, 76	33 (7%)

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	194	TYR	4.0
1	B	187	ARG	3.4
1	B	150	VAL	3.4
1	A	208	ILE	2.6
1	B	151	ILE	2.6
1	B	117	ASN	2.3
1	C	63	LEU	2.3
1	B	209	GLN	2.1
1	B	63	LEU	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. 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	MG	C	502	1/1	0.89	0.06	-	35,35,35,35	0
2	MG	A	501	1/1	0.81	0.12	-	45,45,45,45	0
2	MG	B	503	1/1	0.95	0.07	-	52,52,52,52	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.