



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 28, 2021 – 02:09 PM JST

PDB ID : 7BRA  
Title : Bacillus subtilis IRG1  
Authors : Park, H.H.; Chun, H.L.  
Deposited on : 2020-03-27  
Resolution : 1.78 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	:	2.16
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)	:	2.16

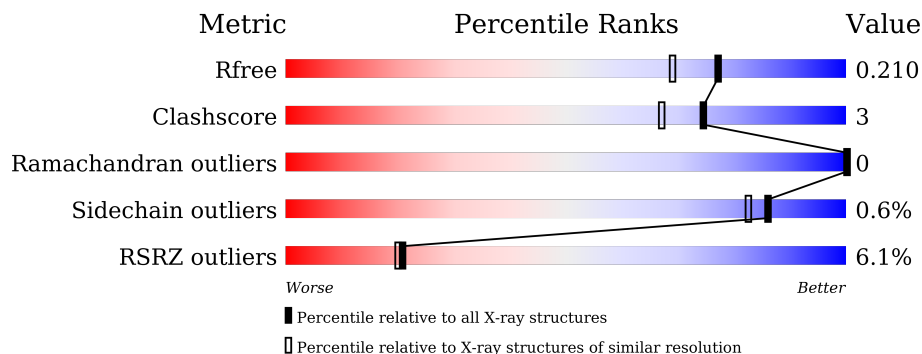
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.78 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	9185 (1.80-1.76)
Clashscore	141614	10184 (1.80-1.76)
Ramachandran outliers	138981	10051 (1.80-1.76)
Sidechain outliers	138945	10050 (1.80-1.76)
RSRZ outliers	127900	9032 (1.80-1.76)

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  $\geq 3$ , 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	445	<div> <div>7%</div> <div>94%</div> <div>6%</div> </div>
1	B	445	<div> <div>5%</div> <div>92%</div> <div>8%</div> </div>

## 2 Entry composition [i](#)

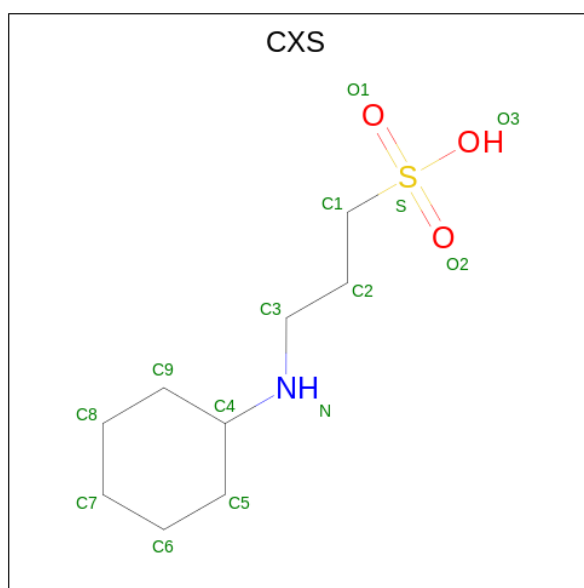
There are 4 unique types of molecules in this entry. The entry contains 7658 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Bacillus subtilis IRG1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	445	Total	C	N	O	S	0	0	0
			3367	2124	600	635	8			
1	B	445	Total	C	N	O	S	0	0	0
			3367	2124	600	635	8			

- Molecule 2 is 3-CYCLOHEXYL-1-PROPYLSULFONIC ACID (three-letter code: CXS) (formula:  $C_9H_{19}NO_3S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			14	9	1	3	1		
2	B	1	Total	C	N	O	S	0	0
			14	9	1	3	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

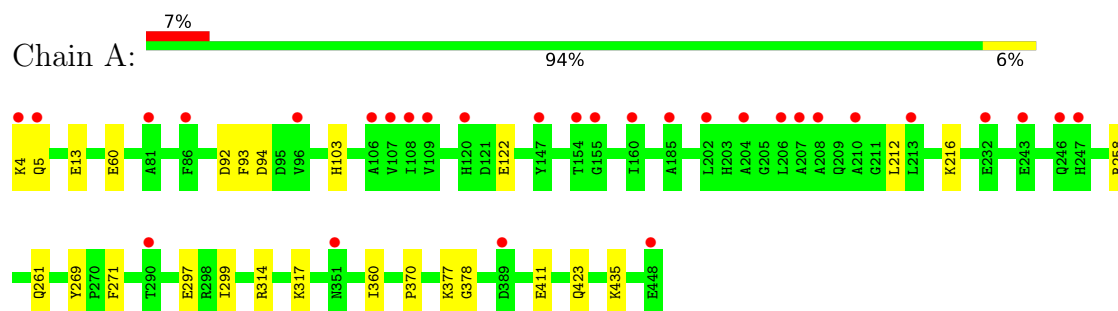
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	427	Total	O	0	0
			427	427		
4	B	459	Total	O	0	0
			459	459		

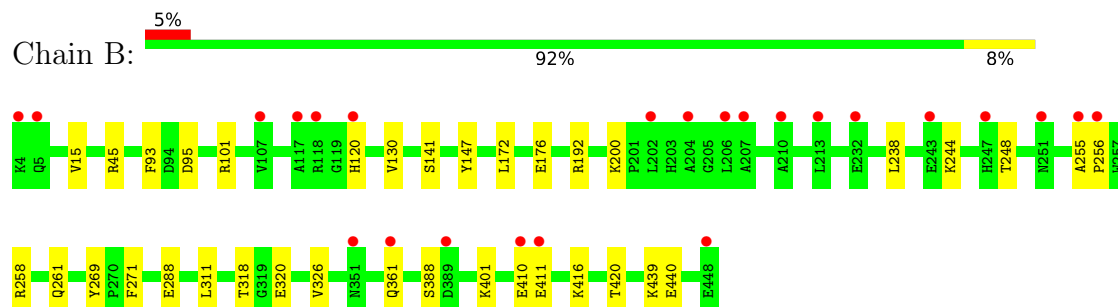
### 3 Residue-property plots [i](#)

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: *Bacillus subtilis* IRG1



- Molecule 1: *Bacillus subtilis* IRG1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.21Å 110.87Å 167.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.39 – 1.78 39.36 – 1.78	Depositor EDS
% Data completeness (in resolution range)	99.4 (39.39-1.78) 99.4 (39.36-1.78)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.31 (at 1.78Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, $R_{free}$	0.175 , 0.201 0.182 , 0.210	Depositor DCC
$R_{free}$ test set	5250 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.5	Xtriage
Anisotropy	0.023	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 45.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7658	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.07% 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: CXS, SO4

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	0.38	0/3437	0.57	1/4658 (0.0%)
1	B	0.39	0/3437	0.57	1/4658 (0.0%)
All	All	0.38	0/6874	0.57	2/9316 (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	A	297	GLU	CA-CB-CG	6.88	128.54	113.40
1	B	95	ASP	CB-CG-OD1	5.27	123.04	118.30

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	3367	0	3340	16	0
1	B	3367	0	3340	26	0
2	A	14	0	19	1	0
2	B	14	0	19	1	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
4	A	427	0	0	9	4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	459	0	0	9	5
All	All	7658	0	6718	43	5

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 (43) 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:388:SER:O	4:B:601:HOH:O	1.89	0.91
1:A:92:ASP:OD1	4:A:601:HOH:O	2.00	0.78
1:A:423:GLN:NE2	4:A:607:HOH:O	2.27	0.68
1:B:288:GLU:OE2	4:B:602:HOH:O	2.11	0.68
1:B:318:THR:HG22	1:B:320:GLU:H	1.60	0.67
1:B:440:GLU:OE1	4:B:603:HOH:O	2.13	0.67
1:B:258:ARG:NH1	4:B:606:HOH:O	2.30	0.64
1:B:244:LYS:O	1:B:248:THR:HG22	2.02	0.60
1:B:318:THR:HG22	1:B:320:GLU:N	2.16	0.60
1:B:361:GLN:HA	1:B:361:GLN:OE1	2.04	0.58
1:A:317:LYS:HE2	4:A:850:HOH:O	2.05	0.56
1:A:435:LYS:HG3	4:A:631:HOH:O	2.06	0.55
1:B:141:SER:OG	1:B:248:THR:HG23	2.06	0.55
1:B:416:LYS:HB2	2:B:501:CXS:H72	1.89	0.54
1:B:101:ARG:NH1	4:B:613:HOH:O	2.36	0.53
1:B:439:LYS:NZ	4:B:616:HOH:O	2.42	0.53
1:B:401:LYS:NZ	4:B:604:HOH:O	2.19	0.51
1:A:60:GLU:OE2	4:A:602:HOH:O	2.19	0.50
1:B:200:LYS:NZ	4:B:610:HOH:O	2.34	0.49
1:A:94:ASP:OD2	1:A:103:HIS:HD2	1.95	0.49
1:B:45:ARG:HH11	1:B:420:THR:HG23	1.78	0.48
1:A:122:GLU:OE2	4:A:603:HOH:O	2.20	0.47
2:A:501:CXS:H91	2:A:501:CXS:H32	1.56	0.47
1:B:120:HIS:HE1	4:B:828:HOH:O	1.97	0.46
1:B:45:ARG:HD2	1:B:420:THR:HG21	1.97	0.45
1:A:299:ILE:HB	1:A:360:ILE:HG12	1.99	0.45
1:A:212:LEU:HG	1:A:216:LYS:HE2	2.00	0.44
1:A:269:TYR:HB3	1:A:271:PHE:CE1	2.53	0.44
1:A:314:ARG:NH1	4:A:619:HOH:O	2.51	0.44
1:A:377:LYS:HG2	4:A:736:HOH:O	2.18	0.43
1:B:258:ARG:HA	1:B:261:GLN:O	2.19	0.43
1:B:147:TYR:CZ	1:B:192:ARG:HD3	2.54	0.42

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:258:ARG:HA	1:A:261:GLN:O	2.19	0.42
1:A:370:PRO:HG2	1:A:378:GLY:HA2	2.01	0.42
1:A:411:GLU:OE1	4:A:604:HOH:O	2.22	0.42
1:B:255:ALA:HA	1:B:256:PRO:HA	1.78	0.42
1:B:311:LEU:HD13	1:B:326:VAL:HB	2.01	0.42
1:B:15:VAL:HG13	1:B:130:VAL:HG11	2.02	0.41
1:A:4:LYS:N	1:A:13:GLU:OE1	2.53	0.41
1:B:269:TYR:HB3	1:B:271:PHE:CE1	2.56	0.41
1:B:172:LEU:HG	1:B:176:GLU:HB3	2.03	0.40
1:B:410:GLU:HB2	1:B:411:GLU:OE2	2.21	0.40
1:B:147:TYR:CE2	1:B:192:ARG:HD3	2.56	0.40

All (5) 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)
4:A:908:HOH:O	4:B:929:HOH:O[2_665]	1.97	0.23
4:A:1023:HOH:O	4:B:960:HOH:O[2_665]	1.99	0.21
4:A:908:HOH:O	4:B:885:HOH:O[2_665]	2.03	0.17
4:B:999:HOH:O	4:B:1028:HOH:O[4_566]	2.12	0.08
4:A:993:HOH:O	4:B:960:HOH:O[2_665]	2.14	0.06

## 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	443/445 (100%)	440 (99%)	3 (1%)	0	100	100
1	B	443/445 (100%)	436 (98%)	7 (2%)	0	100	100
All	All	886/890 (100%)	876 (99%)	10 (1%)	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	340/341 (100%)	338 (99%)	2 (1%)	86	82
1	B	340/341 (100%)	338 (99%)	2 (1%)	86	82
All	All	680/682 (100%)	676 (99%)	4 (1%)	86	82

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

Mol	Chain	Res	Type
1	A	5	GLN
1	A	93	PHE
1	B	93	PHE
1	B	238	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	5	GLN
1	A	369	GLN
1	A	423	GLN
1	B	347	GLN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry

4 ligands are modelled in this entry.

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	SO4	B	502	-	4,4,4	0.17	0	6,6,6	0.19	0
3	SO4	A	502	-	4,4,4	0.14	0	6,6,6	0.42	0
2	CXS	A	501	-	14,14,14	1.45	2 (14%)	18,18,18	3.29	7 (38%)
2	CXS	B	501	-	14,14,14	1.42	3 (21%)	18,18,18	2.63	8 (44%)

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
2	CXS	A	501	-	-	3/8/16/16	0/1/1/1
2	CXS	B	501	-	-	2/8/16/16	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	CXS	C1-S	3.68	1.82	1.77
2	B	501	CXS	O2-S	3.45	1.55	1.45
2	A	501	CXS	O1-S	2.79	1.53	1.45
2	B	501	CXS	C1-S	2.28	1.80	1.77
2	B	501	CXS	O1-S	2.11	1.51	1.45

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	CXS	O3-S-C1	8.69	119.83	105.77
2	B	501	CXS	O1-S-C1	6.55	114.81	106.92
2	A	501	CXS	O2-S-C1	6.07	114.22	106.92
2	A	501	CXS	C7-C8-C9	5.32	122.27	111.42

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	501	CXS	O3-S-O1	-5.17	98.64	111.27
2	A	501	CXS	O3-S-O2	-4.12	101.22	111.27
2	B	501	CXS	O3-S-C1	3.64	111.66	105.77
2	A	501	CXS	C2-C3-N	-3.20	100.11	112.02
2	B	501	CXS	C3-C2-C1	-3.07	102.68	112.79
2	A	501	CXS	C2-C1-S	-2.37	109.61	113.25
2	A	501	CXS	O3-S-O1	-2.37	105.48	111.27
2	B	501	CXS	C7-C8-C9	2.34	116.18	111.42
2	B	501	CXS	C9-C4-C5	2.22	114.67	110.82
2	B	501	CXS	C7-C6-C5	2.07	115.63	111.42
2	B	501	CXS	O2-S-C1	2.03	109.36	106.92

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	CXS	C9-C4-N-C3
2	A	501	CXS	C1-C2-C3-N
2	A	501	CXS	C5-C4-N-C3
2	B	501	CXS	C1-C2-C3-N
2	B	501	CXS	C2-C3-N-C4

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	CXS	1	0
2	B	501	CXS	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	445/445 (100%)	0.38	30 (6%)	17 17	16, 24, 38, 54	0
1	B	445/445 (100%)	0.28	24 (5%)	25 24	15, 23, 36, 56	0
All	All	890/890 (100%)	0.33	54 (6%)	21 20	15, 24, 38, 56	0

All (54) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	5	GLN	4.5
1	A	448	GLU	4.5
1	B	4	LYS	4.5
1	A	107	VAL	4.2
1	A	247	HIS	4.1
1	A	4	LYS	4.1
1	B	118	ARG	3.8
1	A	120	HIS	3.7
1	B	247	HIS	3.5
1	B	206	LEU	3.5
1	B	5	GLN	3.4
1	A	202	LEU	3.4
1	A	206	LEU	3.4
1	A	106	ALA	3.4
1	B	410	GLU	3.2
1	A	208	ALA	3.2
1	A	108	ILE	3.1
1	A	290	THR	3.1
1	A	389	ASP	3.0
1	A	210	ALA	3.0
1	B	107	VAL	3.0
1	B	448	GLU	2.9
1	B	411	GLU	2.8
1	A	207	ALA	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	207	ALA	2.8
1	A	213	LEU	2.8
1	A	243	GLU	2.8
1	B	389	ASP	2.7
1	A	81	ALA	2.7
1	B	213	LEU	2.7
1	B	361	GLN	2.6
1	B	255	ALA	2.6
1	B	256	PRO	2.6
1	B	202	LEU	2.6
1	B	351	ASN	2.6
1	A	246	GLN	2.6
1	B	120	HIS	2.5
1	A	109	VAL	2.5
1	A	155	GLY	2.4
1	B	210	ALA	2.4
1	A	204	ALA	2.3
1	A	351	ASN	2.3
1	A	154	THR	2.3
1	A	96	VAL	2.2
1	A	160	ILE	2.2
1	A	147	TYR	2.2
1	B	251	ASN	2.2
1	B	204	ALA	2.2
1	B	243	GLU	2.2
1	A	185	ALA	2.2
1	B	117	ALA	2.2
1	B	232	GLU	2.1
1	A	232	GLU	2.0
1	A	86	PHE	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 monosaccharides 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	CXS	B	501	14/14	0.94	0.12	31,35,42,46	0
2	CXS	A	501	14/14	0.95	0.12	29,33,40,40	0
3	SO4	A	502	5/5	0.98	0.15	34,36,40,43	0
3	SO4	B	502	5/5	0.99	0.11	36,38,41,41	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.