WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 16, 2020 – 06:26 am BST

PDB ID	:	6S8Q
Title	:	Human Brr2 Helicase Region in complex with C-tail deleted Jab1
Authors	:	Santos, K.F.; Vester, K.; Wahl, M.C.
Deposited on	:	2019-07-10
Resolution	:	2.39 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

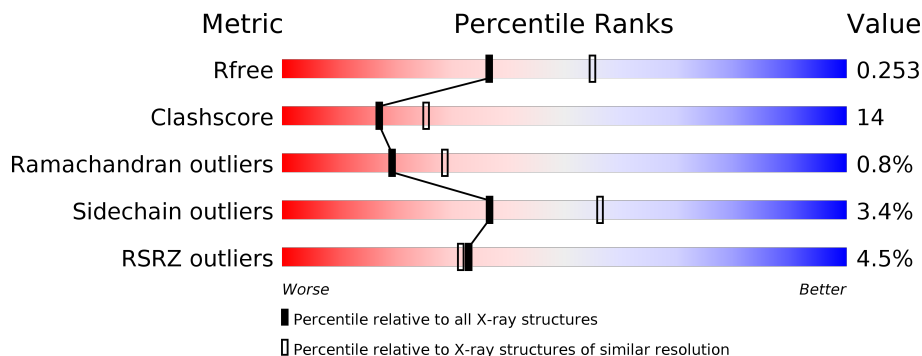|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.39 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	Percentile Ranks	Value
Rfree		0.253
Clashscore		14
Ramachandran outliers		0.8%
Sidechain outliers		3.4%
RSRZ outliers		4.5%

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 <=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	B	1747	4% 70% 27% ··
2	J	263	6% 78% 19% ·

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 16866 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 U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	1725	Total	C	N	O	S	0	22	0
			14071	8979	2416	2602	74			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	390	GLY	-	expression tag	UNP O75643
B	391	ALA	-	expression tag	UNP O75643
B	392	GLU	-	expression tag	UNP O75643
B	393	PHE	-	expression tag	UNP O75643

- Molecule 2 is a protein called Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	J	262	Total	C	N	O	S	0	1	0
			2129	1362	368	387	12			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	2058	GLY	-	expression tag	UNP Q6P2Q9
J	2059	PRO	-	expression tag	UNP Q6P2Q9
J	2060	LEU	-	expression tag	UNP Q6P2Q9
J	2061	GLY	-	expression tag	UNP Q6P2Q9
J	2062	SER	-	expression tag	UNP Q6P2Q9
J	2063	MET	-	expression tag	UNP Q6P2Q9

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).
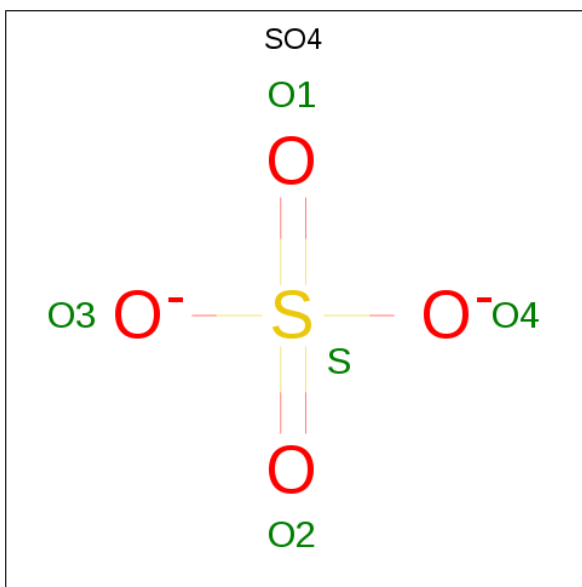
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total 5	O 4	S 1	0	0
3	B	1	Total 5	O 4	S 1	0	0

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	572	Total 574	O 574	0	2
4	J	82	Total 82	O 82	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: U5 small nuclear ribonucleoprotein 200 kDa helicase

Chain B:
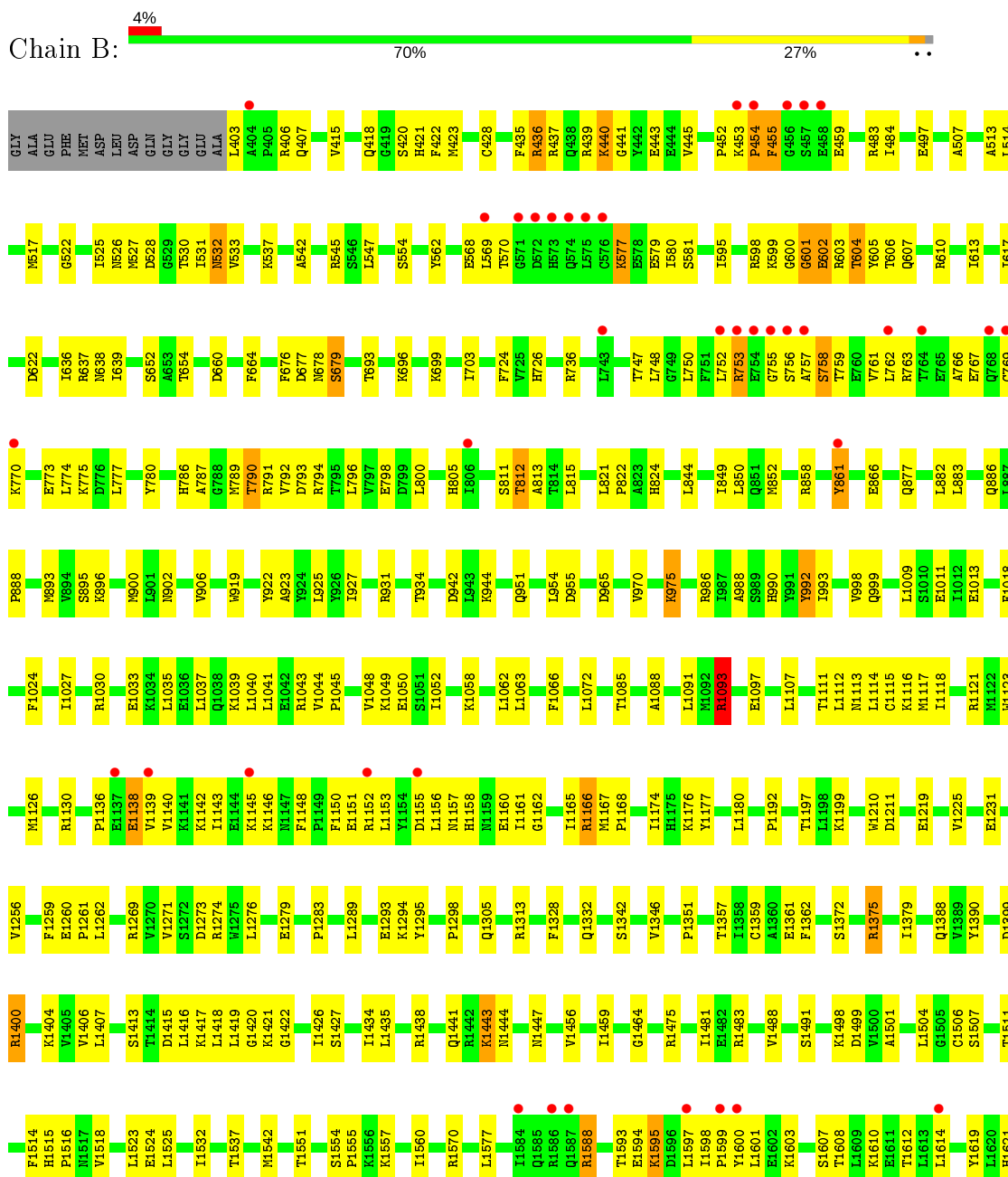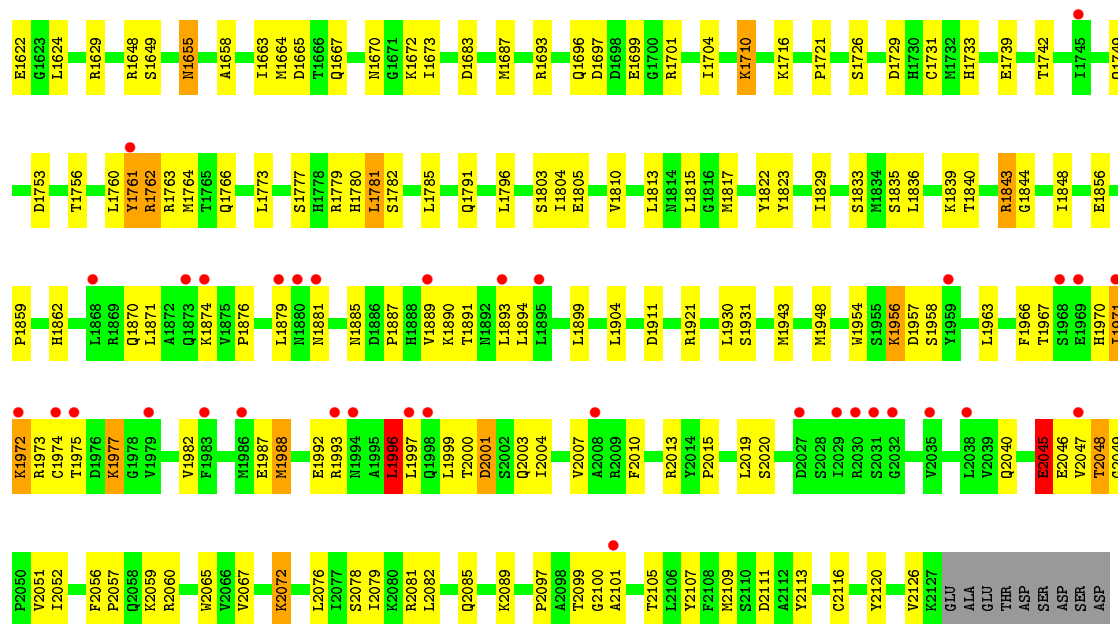
• Molecule 2: Pre-mRNA-processing-splicing factor 8

Chain J:

wwPDB X-ray Structure Validation Summary Report
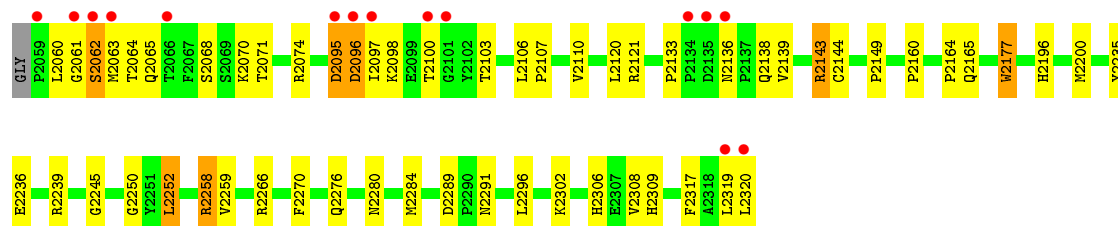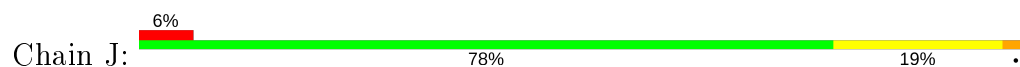
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	99.48Å 118.63Å 187.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.08 – 2.39 48.36 – 2.39	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.08-2.39) 99.7 (48.36-2.39)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.78 (at 2.39Å)	Xtriage
Refinement program	PHENIX 1.15.2_3472	Depositor
R, R_{free}	0.179 , 0.253 0.179 , 0.253	Depositor DCC
R_{free} test set	2100 reflections (2.38%)	wwPDB-VP
Wilson B-factor (Å^2)	41.1	Xtriage
Anisotropy	0.218	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 56.0	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	16866	wwPDB-VP
Average B, all atoms (Å^2)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.97% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.
[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: 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	B	0.48	3/14366 (0.0%)	0.65	11/19457 (0.1%)
2	J	0.44	0/2196	0.60	0/2989
All	All	0.48	3/16562 (0.0%)	0.64	11/22446 (0.0%)

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	B	0	5

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1731	CYS	CB-SG	-5.65	1.72	1.81
1	B	1761	TYR	CB-CG	-5.59	1.43	1.51
1	B	861	TYR	CD1-CE1	-5.33	1.31	1.39

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1595	LYS	CD-CE-NZ	-12.50	82.94	111.70
1	B	944	LYS	CD-CE-NZ	-11.57	85.10	111.70
1	B	1977	LYS	CD-CE-NZ	-8.57	91.98	111.70
1	B	1166	ARG	CG-CD-NE	-8.53	93.90	111.80
1	B	1166	ARG	NE-CZ-NH1	-6.87	116.87	120.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	1971	ILE	Peptide
1	B	1987	GLU	Peptide
1	B	1996	LEU	Peptide
1	B	601	GLY	Peptide
1	B	602	GLU	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	B	14071	0	14173	409	2
2	J	2129	0	2073	46	1
3	B	10	0	0	1	0
4	B	574	0	0	44	0
4	J	82	0	0	10	0
All	All	16866	0	16246	450	2

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 14.

The worst 5 of 450 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:753:ARG:HB3	1:B:756:SER:HB3	1.45	0.99
1:B:406:ARG:NH1	1:B:955:ASP:OD1	1.98	0.97
1:B:406:ARG:NH2	1:B:951:GLN:OE1	1.98	0.96
1:B:2065:TRP:HE1	1:B:2081[B]:ARG:HE	1.08	0.94
1:B:602:GLU:HB2	1:B:603:ARG:HA	1.47	0.93

All (2) 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:B:1166:ARG:NH2	1:B:1988:MET:O[2_555]	2.09	0.11
1:B:1782:SER:OG	2:J:2065:GLN:NE2[4_555]	2.16	0.04

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	B	1743/1747 (100%)	1659 (95%)	70 (4%)	14 (1%)	19	29
2	J	261/263 (99%)	245 (94%)	13 (5%)	3 (1%)	14	20
All	All	2004/2010 (100%)	1904 (95%)	83 (4%)	17 (1%)	19	29

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	440	LYS
1	B	790	THR
1	B	1988	MET
1	B	2048	THR
2	J	2061	GLY

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	B	1566/1560 (100%)	1511 (96%)	55 (4%)	36	55
2	J	237/236 (100%)	228 (96%)	9 (4%)	33	51
All	All	1803/1796 (100%)	1739 (96%)	64 (4%)	37	55

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1413	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	1699	GLU
2	J	2121	ARG
1	B	1443	LYS
1	B	1607	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	2085	GLN
2	J	2247	ASN
2	J	2154	HIS
1	B	877	GLN
2	J	2196	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)

2 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	5702	-	4,4,4	0.24	0	6,6,6	0.50	0
3	SO4	B	5701	-	4,4,4	0.15	0	6,6,6	0.11	0

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	5702	SO4	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	2046:GLU	C	2047:VAL	N	5.64

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, 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	B	1725/1747 (98%)	-0.03	74 (4%)	35	33	19, 51, 112, 196	0
2	J	262/263 (99%)	-0.07	15 (5%)	23	22	29, 52, 113, 215	0
All	All	1987/2010 (98%)	-0.04	89 (4%)	33	31	19, 52, 113, 215	0

The worst 5 of 89 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	J	2062	SER	8.2
2	J	2097	ILE	6.8
2	J	2061	GLY	6.8
1	B	769	CYS	6.4
1	B	1971	ILE	5.9

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, 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	SO4	B	5702	5/5	0.89	0.22	93,95,96,99	0
3	SO4	B	5701	5/5	0.89	0.26	138,145,147,150	0

6.5 Other polymers (i)

There are no such residues in this entry.