wwPDB X-ray Structure Validation Summary Report (i)

Nov 1, 2023 – 12:29 PM JST

PDB ID	:	5HAA
Title	:	Crystal structure of mpy-RNase J, an archaeal RNase J from Methanolobus psychrophilus R15
Authors	:	Feng, N.; Li, D.F.
Deposited on	:	2015-12-30
Resolution	:	2.90 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

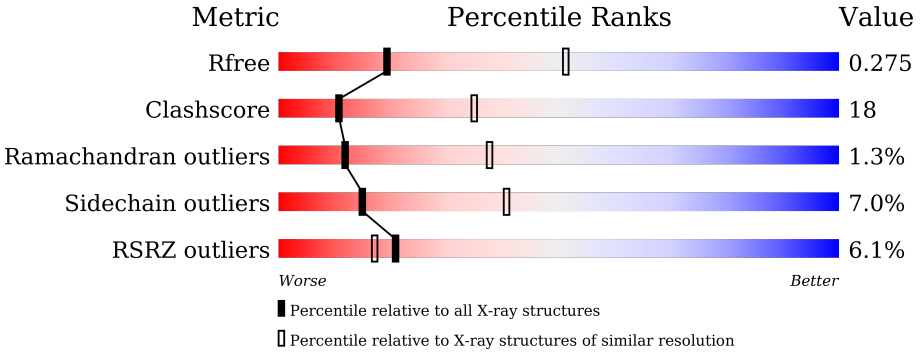
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 >=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	A	470	5% 57% 34% 5% 5%
1	B	470	7% 60% 35% • •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	ZN	A	501	-	-	-	X
3	SO4	A	506	-	-	X	-
3	SO4	B	507	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7113 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 Ribonuclease J.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	447	3482	2192	608	662	20	0	0	0
1	B	460	3580	2249	630	681	20	0	0	0

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-21	MET	-	initiating methionine	UNP K4MAF9
A	-20	GLY	-	expression tag	UNP K4MAF9
A	-19	SER	-	expression tag	UNP K4MAF9
A	-18	SER	-	expression tag	UNP K4MAF9
A	-17	HIS	-	expression tag	UNP K4MAF9
A	-16	HIS	-	expression tag	UNP K4MAF9
A	-15	HIS	-	expression tag	UNP K4MAF9
A	-14	HIS	-	expression tag	UNP K4MAF9
A	-13	HIS	-	expression tag	UNP K4MAF9
A	-12	HIS	-	expression tag	UNP K4MAF9
A	-11	SER	-	expression tag	UNP K4MAF9
A	-10	SER	-	expression tag	UNP K4MAF9
A	-9	GLY	-	expression tag	UNP K4MAF9
A	-8	LEU	-	expression tag	UNP K4MAF9
A	-7	VAL	-	expression tag	UNP K4MAF9
A	-6	PRO	-	expression tag	UNP K4MAF9
A	-5	ARG	-	expression tag	UNP K4MAF9
A	-4	GLY	-	expression tag	UNP K4MAF9
A	-3	SER	-	expression tag	UNP K4MAF9
A	-2	HIS	-	expression tag	UNP K4MAF9
A	-1	MET	-	expression tag	UNP K4MAF9
A	0	ALA	-	expression tag	UNP K4MAF9
A	1	SER	-	expression tag	UNP K4MAF9
B	-21	MET	-	initiating methionine	UNP K4MAF9
B	-20	GLY	-	expression tag	UNP K4MAF9

Continued on next page...

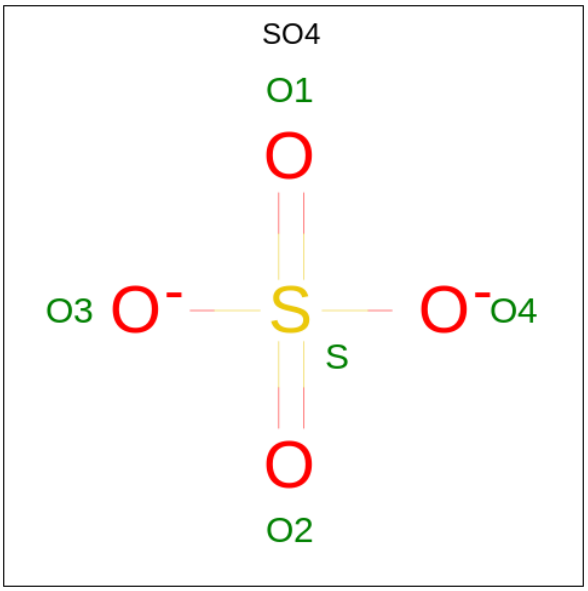
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-19	SER	-	expression tag	UNP K4MAF9
B	-18	SER	-	expression tag	UNP K4MAF9
B	-17	HIS	-	expression tag	UNP K4MAF9
B	-16	HIS	-	expression tag	UNP K4MAF9
B	-15	HIS	-	expression tag	UNP K4MAF9
B	-14	HIS	-	expression tag	UNP K4MAF9
B	-13	HIS	-	expression tag	UNP K4MAF9
B	-12	HIS	-	expression tag	UNP K4MAF9
B	-11	SER	-	expression tag	UNP K4MAF9
B	-10	SER	-	expression tag	UNP K4MAF9
B	-9	GLY	-	expression tag	UNP K4MAF9
B	-8	LEU	-	expression tag	UNP K4MAF9
B	-7	VAL	-	expression tag	UNP K4MAF9
B	-6	PRO	-	expression tag	UNP K4MAF9
B	-5	ARG	-	expression tag	UNP K4MAF9
B	-4	GLY	-	expression tag	UNP K4MAF9
B	-3	SER	-	expression tag	UNP K4MAF9
B	-2	HIS	-	expression tag	UNP K4MAF9
B	-1	MET	-	expression tag	UNP K4MAF9
B	0	ALA	-	expression tag	UNP K4MAF9
B	1	SER	-	expression tag	UNP K4MAF9

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
2	A	2	2	2	0	0
			Total	Zn		
2	B	2	2	2	0	0

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



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

- Molecule 4 is water.

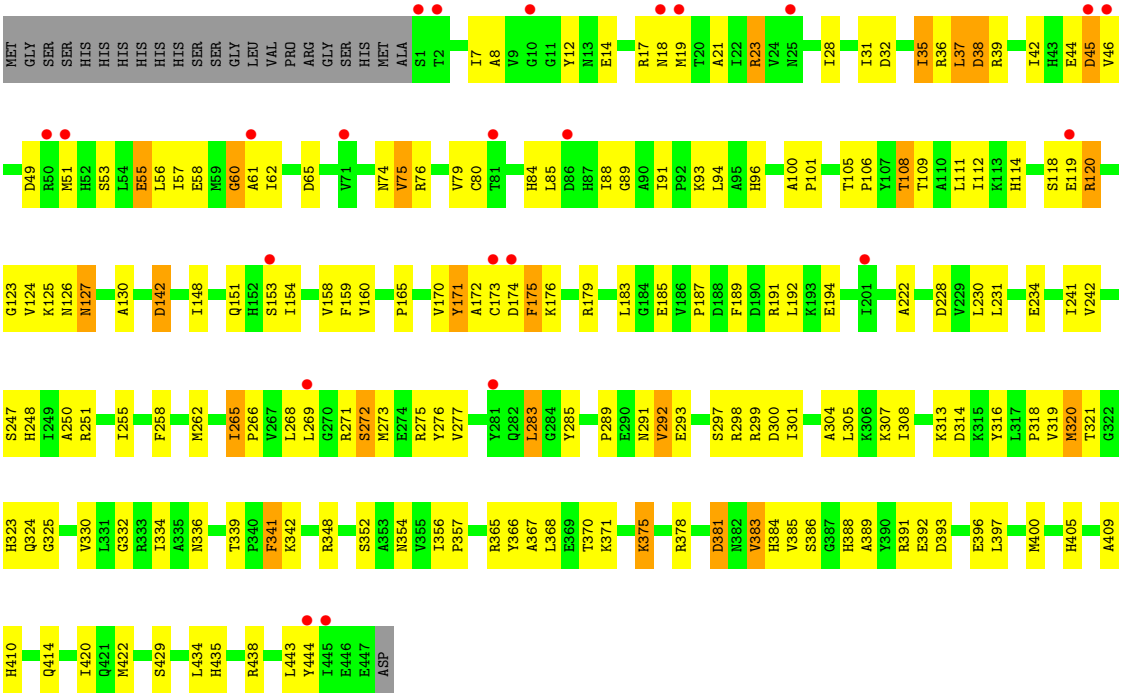
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O 1 1	0	0
4	B	1	Total O 1 1	0	0

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.

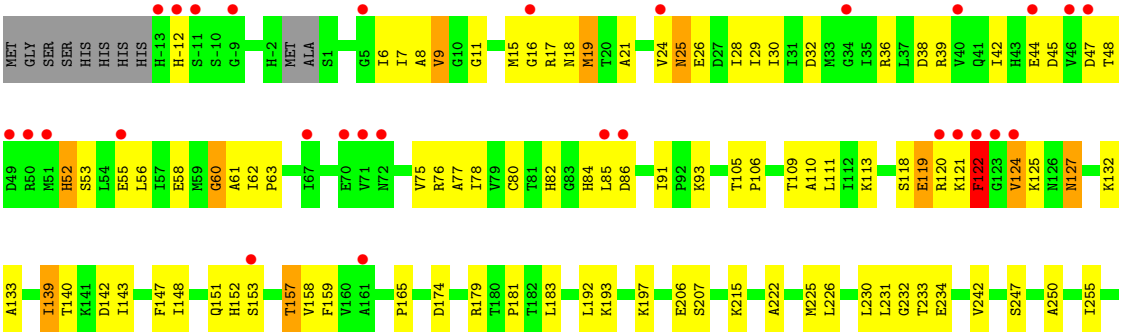
• Molecule 1: Ribonuclease J

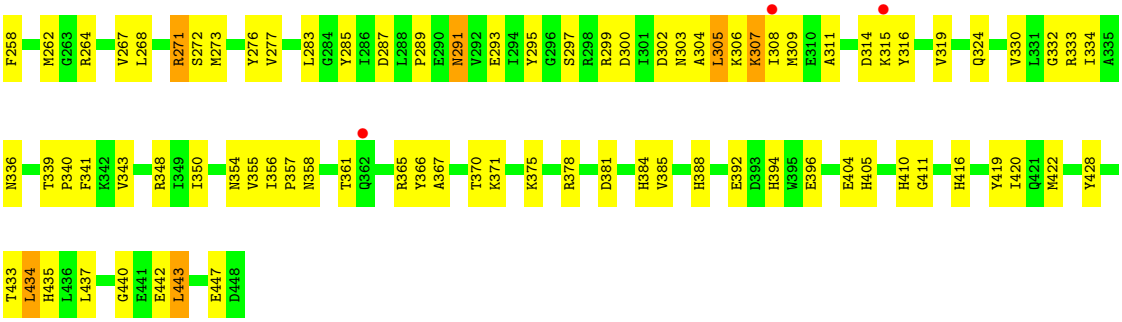
Chain A:



• Molecule 1: Ribonuclease J

Chain B:




F258
M262
G263
R264
V267
L268
R271
S272
M273
Y276
V277
L283
G284
Y285
I286
D287
L288
P289
E290
N291
V292
E293
I294
Y295
G296
S297
R298
R299
D300
I301
D302
N303
A304
L305
K306
K307
I308
M309
E310
A311
D314
K315
Y316
V319
Q324
V330
L331
G332
R333
I334
A335
N336
T339
P340
F341
K342
V343
R348
I349
I350
N354
V355
I356
P357
N358
T361
Q362
R365
Y366
A367
T370
K371
K375
R378
D381
H384
V385
H388
E392
D393
H394
W395
E396
E404
H405
H410
G411
H416
Y419
I420
Q421
M422
Y428
T433
L434
H435
L436
L437
G440
E441
E442
L443
E447
D448

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 62 2 2	Depositor
Cell constants a, b, c, α, β, γ	162.51Å 162.51Å 180.66Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.03 – 2.90 39.03 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.1 (39.03-2.90) 98.2 (39.03-2.90)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	1.95 (at 2.90Å)	Xtriage
Refinement program	PHENIX dev_1980	Depositor
R, R_{free}	0.237 , 0.274 0.238 , 0.275	Depositor DCC
R_{free} test set	1998 reflections (6.43%)	wwPDB-VP
Wilson B-factor (Å^2)	92.8	Xtriage
Anisotropy	0.143	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 67.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7113	wwPDB-VP
Average B, all atoms (Å^2)	110.0	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 45.34 % 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 1.3398e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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, ZN

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	0.31	0/3544	0.62	1/4798 (0.0%)
1	B	0.37	0/3645	0.66	5/4932 (0.1%)
All	All	0.34	0/7189	0.64	6/9730 (0.1%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	124	VAL	CG1-CB-CG2	6.88	121.91	110.90
1	B	305	LEU	CB-CG-CD2	-5.82	101.11	111.00
1	B	122	PHE	CB-CG-CD1	-5.55	116.92	120.80
1	A	443	LEU	CA-CB-CG	5.33	127.56	115.30
1	B	124	VAL	N-CA-C	-5.22	96.90	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	3482	0	3488	145	0
1	B	3580	0	3577	122	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	20	0	0	4	0
3	B	25	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
All	All	7113	0	7065	259	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 18.

The worst 5 of 259 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:39:ARG:HA	1:A:42:ILE:HD13	1.56	0.85
1:B:305:LEU:HA	1:B:308:ILE:HG22	1.60	0.84
1:B:305:LEU:HD21	1:B:341:PHE:H	1.46	0.81
1:A:179:ARG:NH1	1:B:232:GLY:O	2.15	0.80
1:B:76:ARG:NH1	1:B:139:ILE:O	2.16	0.78

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	445/470 (95%)	411 (92%)	29 (6%)	5 (1%)	14 42
1	B	456/470 (97%)	417 (91%)	32 (7%)	7 (2%)	10 34
All	All	901/940 (96%)	828 (92%)	61 (7%)	12 (1%)	12 37

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	120	ARG
1	B	45	ASP
1	A	118	SER
1	A	45	ASP
1	B	32	ASP

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	375/394 (95%)	349 (93%)	26 (7%)	15 41
1	B	386/394 (98%)	359 (93%)	27 (7%)	15 41
All	All	761/788 (97%)	708 (93%)	53 (7%)	15 41

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	19	MET
1	B	122	PHE
1	B	422	MET
1	B	42	ILE
1	B	52	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	209	ASN
1	B	324	GLN
1	B	282	GLN
1	B	336	ASN
1	A	257	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 (i)

Of 13 ligands modelled in this entry, 4 are monoatomic - leaving 9 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	$\#\|Z\| > 2$
3	SO4	B	506	-	4,4,4	0.14	0	6,6,6	0.06	0
3	SO4	B	505	-	4,4,4	0.15	0	6,6,6	0.05	0
3	SO4	B	503	-	4,4,4	0.14	0	6,6,6	0.09	0
3	SO4	B	504	-	4,4,4	0.14	0	6,6,6	0.04	0
3	SO4	A	505	-	4,4,4	0.17	0	6,6,6	0.07	0
3	SO4	B	507	-	4,4,4	0.14	0	6,6,6	0.05	0
3	SO4	A	503	-	4,4,4	0.16	0	6,6,6	0.13	0
3	SO4	A	504	1	4,4,4	0.11	0	6,6,6	0.07	0
3	SO4	A	506	1	4,4,4	0.14	0	6,6,6	0.06	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 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	506	SO4	4	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 (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	A	447/470 (95%)	0.39	23 (5%)	28	24	63, 101, 148, 208	0
1	B	460/470 (97%)	0.43	32 (6%)	16	12	63, 109, 168, 246	0
All	All	907/940 (96%)	0.41	55 (6%)	21	17	63, 105, 158, 246	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	44	GLU	7.7
1	B	124	VAL	6.7
1	B	123	GLY	5.4
1	A	46	VAL	5.3
1	B	122	PHE	4.8

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, 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	507	5/5	0.58	1.02	428,428,430,431	0
3	SO4	A	504	5/5	0.63	0.39	137,159,165,165	0
2	ZN	A	501	1/1	0.72	0.40	152,152,152,152	0
3	SO4	B	506	5/5	0.76	0.25	187,187,187,187	0
3	SO4	B	504	5/5	0.80	0.43	167,167,168,168	0
3	SO4	A	506	5/5	0.81	0.28	130,163,169,170	0
3	SO4	A	505	5/5	0.86	0.14	108,123,132,132	0
3	SO4	B	505	5/5	0.86	0.17	128,148,152,153	0
3	SO4	A	503	5/5	0.94	0.21	67,78,83,88	0
2	ZN	B	501	1/1	0.94	0.35	139,139,139,139	0
2	ZN	B	502	1/1	0.96	0.27	114,114,114,114	0
3	SO4	B	503	5/5	0.96	0.18	76,87,97,104	0
2	ZN	A	502	1/1	0.98	0.20	93,93,93,93	0

6.5 Other polymers (i)

There are no such residues in this entry.