



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 1, 2026 – 12:03 PM UTC

PDB ID : 6AC6 / pdb_00006ac6
Title : Ab initio crystal structure of Selenomethionine labelled Mycobacterium smegmatis Mfd
Authors : Putta, S.; Fox, G.C.; Walsh, M.A.; Rao, D.N.; Nagaraja, V.; Natesh, R.
Deposited on : 2018-07-25
Resolution : 2.99 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

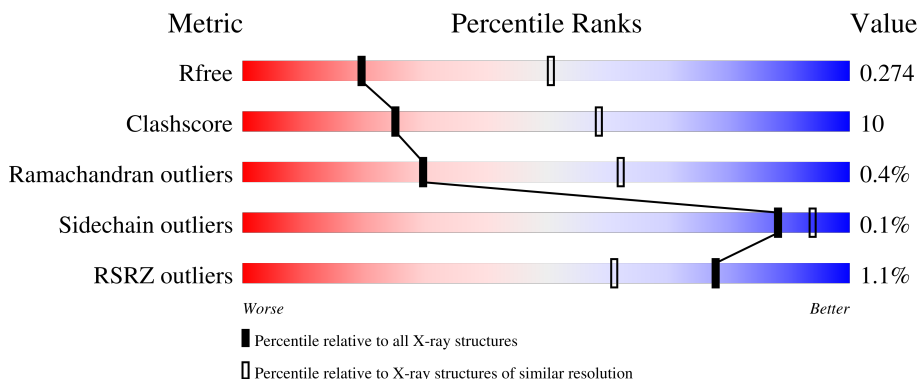
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.99 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}	180053	3580 (3.00-2.96)
Clashscore	190562	3904 (3.00-2.96)
Ramachandran outliers	187476	3761 (3.00-2.96)
Sidechain outliers	187428	3764 (3.00-2.96)
RSRZ outliers	180081	3579 (3.00-2.96)

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 $\leq 5\%$. 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	1235	76% 17% 7%
1	B	1235	71% 22% 7%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 16853 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 Mycobacterium smegmatis Mfd.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	1154	Total	C	N	O	S	Se	0	0	0
			8417	5293	1487	1606	6	25			
1	B	1153	Total	C	N	O	S	Se	0	0	0
			8337	5253	1473	1580	6	25			

There are 40 discrepancies between the modelled and reference sequences:

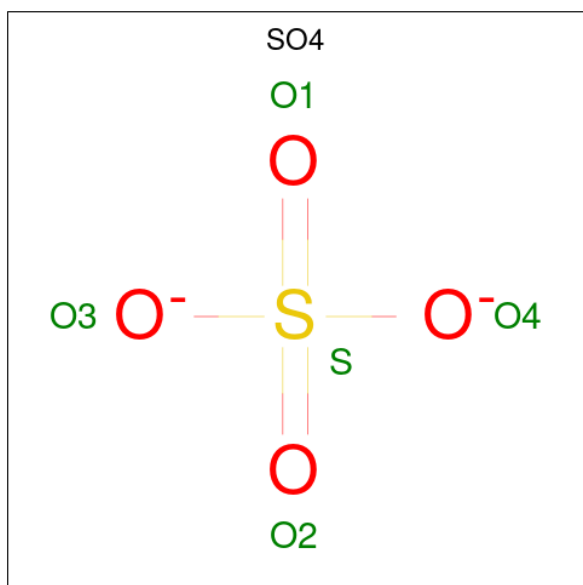
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MSE	-	initiating methionine	UNP I7G7M2
A	-18	GLY	-	expression tag	UNP I7G7M2
A	-17	SER	-	expression tag	UNP I7G7M2
A	-16	SER	-	expression tag	UNP I7G7M2
A	-15	HIS	-	expression tag	UNP I7G7M2
A	-14	HIS	-	expression tag	UNP I7G7M2
A	-13	HIS	-	expression tag	UNP I7G7M2
A	-12	HIS	-	expression tag	UNP I7G7M2
A	-11	HIS	-	expression tag	UNP I7G7M2
A	-10	HIS	-	expression tag	UNP I7G7M2
A	-9	SER	-	expression tag	UNP I7G7M2
A	-8	SER	-	expression tag	UNP I7G7M2
A	-7	GLY	-	expression tag	UNP I7G7M2
A	-6	LEU	-	expression tag	UNP I7G7M2
A	-5	VAL	-	expression tag	UNP I7G7M2
A	-4	PRO	-	expression tag	UNP I7G7M2
A	-3	ARG	-	expression tag	UNP I7G7M2
A	-2	GLY	-	expression tag	UNP I7G7M2
A	-1	SER	-	expression tag	UNP I7G7M2
A	0	HIS	-	expression tag	UNP I7G7M2
B	-19	MSE	-	initiating methionine	UNP I7G7M2
B	-18	GLY	-	expression tag	UNP I7G7M2
B	-17	SER	-	expression tag	UNP I7G7M2
B	-16	SER	-	expression tag	UNP I7G7M2
B	-15	HIS	-	expression tag	UNP I7G7M2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	HIS	-	expression tag	UNP I7G7M2
B	-13	HIS	-	expression tag	UNP I7G7M2
B	-12	HIS	-	expression tag	UNP I7G7M2
B	-11	HIS	-	expression tag	UNP I7G7M2
B	-10	HIS	-	expression tag	UNP I7G7M2
B	-9	SER	-	expression tag	UNP I7G7M2
B	-8	SER	-	expression tag	UNP I7G7M2
B	-7	GLY	-	expression tag	UNP I7G7M2
B	-6	LEU	-	expression tag	UNP I7G7M2
B	-5	VAL	-	expression tag	UNP I7G7M2
B	-4	PRO	-	expression tag	UNP I7G7M2
B	-3	ARG	-	expression tag	UNP I7G7M2
B	-2	GLY	-	expression tag	UNP I7G7M2
B	-1	SER	-	expression tag	UNP I7G7M2
B	0	HIS	-	expression tag	UNP I7G7M2

- Molecule 2 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



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

Continued on next page...

Continued from previous page...

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

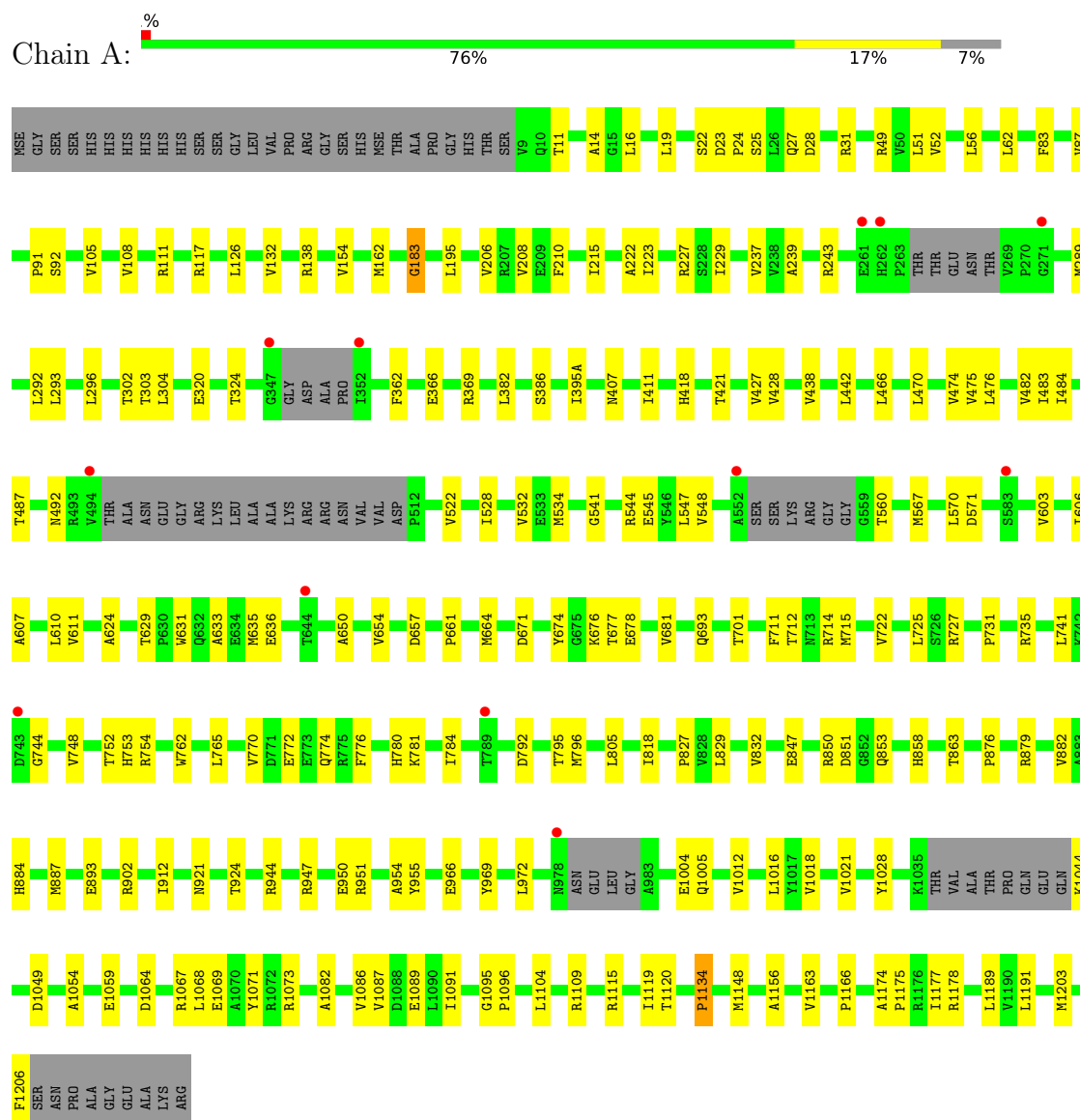
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	37	Total	O	0	0
			37	37		
3	B	37	Total	O	0	0
			37	37		

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: Mycobacterium smegmatis Mfd



• Molecule 1: Mycobacterium smegmatis Mfd



V1086	R951	V840	R714	V611	GLY	P353	F210	L79	MSE
E1089	Y958	R846	G717	Y614	GLY	L354	S216	R80	GLY
Y1094	P959	Q853	T721	R617	LYS	R369	F221	F83	SER
A1101	L964	I857	K723	Q618	LEU	V379	R227	S86	HIS
V1105	R971	H858	S726	A624	ALA	V380	S228	V87	HIS
C1114	L972	R860	R727	T629	LYS	T381	A239	F90	HIS
I1119	A973	V861	F728	P830	ARG	S386	V240	P91	HIS
T1129	I975	R862	D743	K631	D511	V391	P241	W93	SER
M985	H979	R872	V746	M635	A514	E392	T248	T95	GLY
S1133	M985	R879	G751	E636	L515	A401	T264	L96	LEU
P1134	A988	V880	L755	D637	D519	ARG	THR	E99	VAL
R1144	E993	V882	L756	A638	V622	GLY	GLU	R100	PRO
L1145	N999	M887	T758	PHE	I629	SER	ASN	P103	ARG
R1155	H1008	E889	I769	THR	F531	GLN	THR	T107	GLY
T1158	V1012	K894	V770	ILE	E533	H406	P270	R111	MSE
P1164	G1013	F899	V654	D646	H534	M414	VAL	T125	ALA
V1172	D1015	D905	E773	V654	G540	L415	M276	R120	GLY
A1174	L1016	I906	Q774	M658	L547	L418	L277	D123	ALA
I1177	Y1017	I907	F776	P661	A552	H418	A278	E124	PRO
R1178	R1019	C909	H780	P662	SER	A445	E282	T135	GLY
D1179	L1020	T910	K781	P663	SER	D446	V286	T136	THR
L1182	V1021	T911	R788	P667	LYS	A448	M289	T137	SER
L1193	G1022	E914	V793	C669	ARG	L452	T305	R138	ALA
S1207	Y1028	T915	L794	V672	GLY	V463	L308	V154	THR
PRO	T1036	I926	M796	Q675	T560	L466	A312	F164	GLY
ALA	VAL	R929	S797	E678	V565	P469	D318	L174	ALA
GLY	THR	A930	P802	I679	D571	V474	P319	R178	VAL
ALA	PRO	D931	S808	A690	Q672	V475	V322	D180	GLY
LYS	GLU	L935	R813	V681	S574	L476	A326	G183	S47
ARG	GLN	S936	E814	R682	R575	N480	I330	R184	R49
	K1044	Q937	Q693	Q693	L587	L481	K331	G186	L56
	V1052	L938	T320	T700	G588	V482	F336	E187	P61
	S1063	H939	P821	T701	N594	I483	L337	V190	L62
	E1069	Q940	P822	L702	T595	I484	E338	R191	L63
	R1073	L941	E823	L703	K596	H492	A339	I194	V64
	D1079	R944	V832	H707	A599	ARG	S340	A202	A66
		S948	D836	T710	R600	VAL	S341	E209	A67
		R949	Q839			THR	S342		E78
		E950				ASN	I352		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	82.71Å 158.87Å 207.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.19 – 2.99 47.19 – 2.99	Depositor EDS
% Data completeness (in resolution range)	99.0 (47.19-2.99) 96.2 (47.19-2.99)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.69 (at 3.01Å)	Xtriage
Refinement program	PHENIX (1.11.1 _2575: ???)	Depositor
R, R_{free}	0.218 , 0.270 0.223 , 0.274	Depositor DCC
R_{free} test set	2788 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	78.7	Xtriage
Anisotropy	0.389	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 56.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	16853	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.68% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: 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	0.15	0/8542	0.34	0/11635
1	B	0.18	1/8461 (0.0%)	0.37	2/11529 (0.0%)
All	All	0.17	1/17003 (0.0%)	0.35	2/23164 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	289	MSE	SE-CE	-5.33	1.79	1.95

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	531	PHE	CA-C-N	-5.76	116.72	122.77
1	B	531	PHE	C-N-CA	-5.76	116.72	122.77

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	8417	0	8131	145	0
1	B	8337	0	8022	176	0
2	A	15	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	10	0	0	0	0
3	A	37	0	0	2	0
3	B	37	0	0	2	0
All	All	16853	0	16153	318	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 10.

The worst 5 of 318 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:452:LEU:HD13	1:B:466:LEU:HG	1.43	0.95
1:B:330:ILE:HD12	1:B:331:LYS:N	1.91	0.86
1:A:674:TYR:HB3	1:A:818:ILE:HD11	1.62	0.81
1:B:636:GLU:HG2	1:B:682:ARG:HH11	1.48	0.79
1:B:532:VAL:HG23	1:B:533:GLU:HB3	1.64	0.77

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	1140/1235 (92%)	1081 (95%)	56 (5%)	3 (0%)	36	67
1	B	1137/1235 (92%)	1070 (94%)	61 (5%)	6 (0%)	24	58
All	All	2277/2470 (92%)	2151 (94%)	117 (5%)	9 (0%)	30	62

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1134	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	588	GLY
1	A	560	THR
1	A	1134	PRO
1	B	248	THR

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	827/961 (86%)	826 (100%)	1 (0%)	88	94
1	B	807/961 (84%)	806 (100%)	1 (0%)	88	94
All	All	1634/1922 (85%)	1632 (100%)	2 (0%)	88	94

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1203	MSE
1	B	240	VAL

Sometimes 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	226	GLN
1	B	839	GLN
1	A	780	HIS
1	A	783	HIS
1	A	901	ASN

5.3.3 RNA ⓘ

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

5 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	$\# Z > 2$
2	SO4	A	1302	-	4,4,4	0.24	0	6,6,6	0.07	0
2	SO4	B	1301	-	4,4,4	0.24	0	6,6,6	0.12	0
2	SO4	A	1303	-	4,4,4	0.24	0	6,6,6	0.09	0
2	SO4	B	1302	-	4,4,4	0.23	0	6,6,6	0.08	0
2	SO4	A	1301	-	4,4,4	0.24	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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1302	SO4	1	0
2	A	1301	SO4	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 [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	1129/1235 (91%)	-0.02	12 (1%) 78 61	43, 76, 115, 138	0
1	B	1128/1235 (91%)	0.07	12 (1%) 78 61	46, 82, 116, 153	0
All	All	2257/2470 (91%)	0.02	24 (1%) 78 61	43, 78, 115, 153	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	822	PRO	3.5
1	B	682	ARG	3.1
1	A	347	GLY	2.8
1	B	401	ALA	2.7
1	A	552	ALA	2.6

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
2	SO4	B	1302	5/5	0.81	0.12	121,128,130,135	0
2	SO4	B	1301	5/5	0.95	0.07	61,66,69,76	0
2	SO4	A	1303	5/5	0.96	0.06	78,81,90,90	0
2	SO4	A	1301	5/5	0.96	0.05	63,69,78,79	0
2	SO4	A	1302	5/5	0.96	0.06	63,63,64,64	0

6.5 Other polymers [i](#)

There are no such residues in this entry.