wwPDB X-ray Structure Validation Summary Report (i)

Feb 4, 2024 – 11:25 PM EST

PDB ID	:	1U6M
Title	:	The crystal structure of acetyltransferase
Authors	:	Min, T.; Gorman, J.; Shapiro, L.; Burley, S.K.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2004-07-30
Resolution	:	2.40 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:

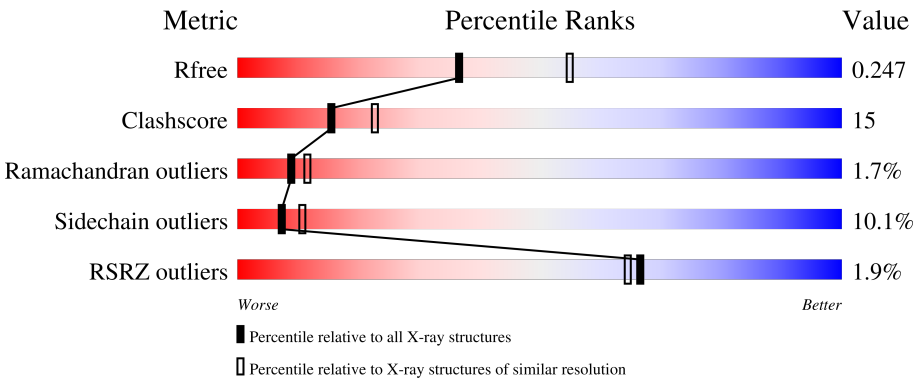
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.40 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	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 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	199	3% 68% 23% • • 5%
1	B	199	% 71% 17% 7% • 5%
1	C	199	2% 68% 21% 6% • 5%
1	D	199	3% 67% 21% 6% • 5%

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	SO4	A	1003	-	X	X	-
2	SO4	A	1005	-	X	X	-
2	SO4	D	1004	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6484 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 acetyltransferase, GNAT family.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	189	1474	939	246	284	5	0	0	0
1	B	189	1466	937	243	281	5	0	0	0
1	C	189	1447	926	246	270	5	0	0	0
1	D	189	1448	927	244	273	4	0	0	0

There are 52 discrepancies between the modelled and reference sequences:

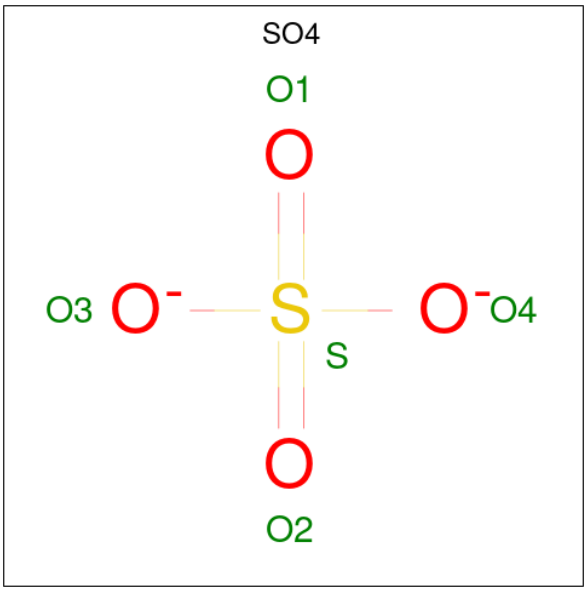
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	cloning artifact	UNP Q836Z8
A	2	SER	-	cloning artifact	UNP Q836Z8
A	3	LEU	-	cloning artifact	UNP Q836Z8
A	190	GLU	-	expression tag	UNP Q836Z8
A	191	GLY	-	expression tag	UNP Q836Z8
A	192	GLY	-	expression tag	UNP Q836Z8
A	193	SER	-	expression tag	UNP Q836Z8
A	194	HIS	-	expression tag	UNP Q836Z8
A	195	HIS	-	expression tag	UNP Q836Z8
A	196	HIS	-	expression tag	UNP Q836Z8
A	197	HIS	-	expression tag	UNP Q836Z8
A	198	HIS	-	expression tag	UNP Q836Z8
A	199	HIS	-	expression tag	UNP Q836Z8
B	1	MET	-	cloning artifact	UNP Q836Z8
B	2	SER	-	cloning artifact	UNP Q836Z8
B	3	LEU	-	cloning artifact	UNP Q836Z8
B	190	GLU	-	expression tag	UNP Q836Z8
B	191	GLY	-	expression tag	UNP Q836Z8
B	192	GLY	-	expression tag	UNP Q836Z8
B	193	SER	-	expression tag	UNP Q836Z8
B	194	HIS	-	expression tag	UNP Q836Z8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	195	HIS	-	expression tag	UNP Q836Z8
B	196	HIS	-	expression tag	UNP Q836Z8
B	197	HIS	-	expression tag	UNP Q836Z8
B	198	HIS	-	expression tag	UNP Q836Z8
B	199	HIS	-	expression tag	UNP Q836Z8
C	1	MET	-	cloning artifact	UNP Q836Z8
C	2	SER	-	cloning artifact	UNP Q836Z8
C	3	LEU	-	cloning artifact	UNP Q836Z8
C	190	GLU	-	expression tag	UNP Q836Z8
C	191	GLY	-	expression tag	UNP Q836Z8
C	192	GLY	-	expression tag	UNP Q836Z8
C	193	SER	-	expression tag	UNP Q836Z8
C	194	HIS	-	expression tag	UNP Q836Z8
C	195	HIS	-	expression tag	UNP Q836Z8
C	196	HIS	-	expression tag	UNP Q836Z8
C	197	HIS	-	expression tag	UNP Q836Z8
C	198	HIS	-	expression tag	UNP Q836Z8
C	199	HIS	-	expression tag	UNP Q836Z8
D	1	MET	-	cloning artifact	UNP Q836Z8
D	2	SER	-	cloning artifact	UNP Q836Z8
D	3	LEU	-	cloning artifact	UNP Q836Z8
D	190	GLU	-	expression tag	UNP Q836Z8
D	191	GLY	-	expression tag	UNP Q836Z8
D	192	GLY	-	expression tag	UNP Q836Z8
D	193	SER	-	expression tag	UNP Q836Z8
D	194	HIS	-	expression tag	UNP Q836Z8
D	195	HIS	-	expression tag	UNP Q836Z8
D	196	HIS	-	expression tag	UNP Q836Z8
D	197	HIS	-	expression tag	UNP Q836Z8
D	198	HIS	-	expression tag	UNP Q836Z8
D	199	HIS	-	expression tag	UNP Q836Z8

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 3 is water.

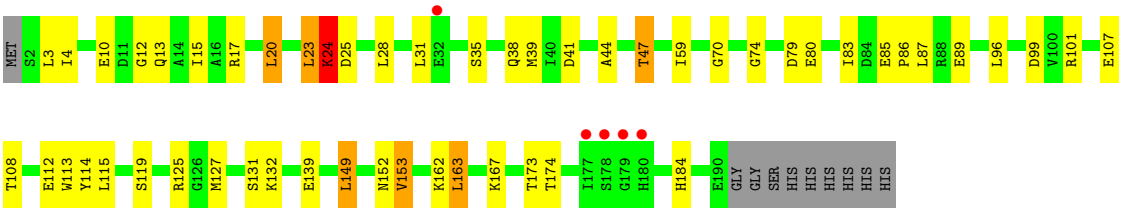
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	179	Total O 179 179	0	0
3	B	194	Total O 194 194	0	0
3	C	120	Total O 120 120	0	0
3	D	116	Total O 116 116	0	0

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.

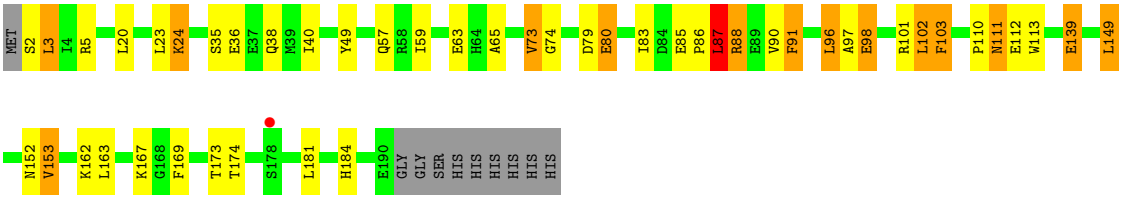
• Molecule 1: acetyltransferase, GNAT family

Chain A:



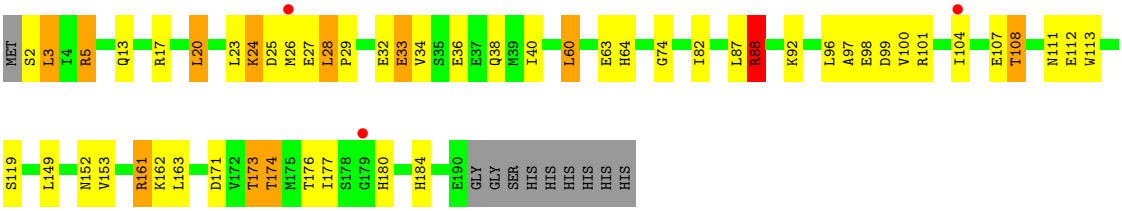
• Molecule 1: acetyltransferase, GNAT family

Chain B:



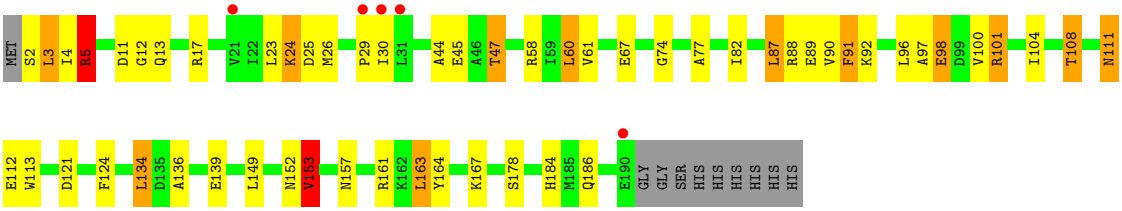
• Molecule 1: acetyltransferase, GNAT family

Chain C:



• Molecule 1: acetyltransferase, GNAT family

Chain D:


MET
S2
L3
I4
R5
D11
G12
Q13
R17
V21
I22
L23
K24
D25
M26
P29
I30
L31
A44
E45
A46
T47
R58
I59
L60
V61
E67
G74
A77
I82
L87
R88
E89
V90
F91
K92
L96
A97
E98
D99
V100
R101
I104
T108
N111
E112
W113
D121
F124
L134
D135
A136
E139
L149
N152
V153
N157
R161
K162
L163
Y164
K167
S178
H184
M185
Q186
E190
GLY
GLY
SER
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	115.26Å 115.26Å 225.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	102.60 – 2.40 20.01 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.7 (102.60-2.40) 99.8 (20.01-2.40)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	7.32 (at 2.41Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.203 , 0.247 0.203 , 0.247	Depositor DCC
R_{free} test set	3015 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	32.5	Xtriage
Anisotropy	0.213	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 45.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6484	wwPDB-VP
Average B, all atoms (Å^2)	34.0	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 38.00 % 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 3.9373e-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

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	1.03	3/1500 (0.2%)	1.01	5/2028 (0.2%)
1	B	1.11	5/1491 (0.3%)	1.18	10/2016 (0.5%)
1	C	0.95	1/1472 (0.1%)	1.02	6/1992 (0.3%)
1	D	1.00	0/1474	1.21	10/1997 (0.5%)
All	All	1.02	9/5937 (0.2%)	1.11	31/8033 (0.4%)

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	103	PHE	N-CA	6.48	1.59	1.46
1	A	139	GLU	CG-CD	6.25	1.61	1.51
1	B	85	GLU	CG-CD	5.74	1.60	1.51
1	A	139	GLU	CB-CG	5.65	1.62	1.52
1	B	63	GLU	CB-CG	-5.50	1.41	1.52

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	5	ARG	NE-CZ-NH2	-19.75	110.43	120.30
1	B	88	ARG	NE-CZ-NH1	-15.85	112.37	120.30
1	D	5	ARG	NE-CZ-NH1	15.50	128.05	120.30
1	D	3	LEU	CA-CB-CG	9.49	137.12	115.30
1	B	88	ARG	NE-CZ-NH2	9.41	125.01	120.30

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	1474	0	1460	42	0
1	B	1466	0	1458	38	0
1	C	1447	0	1442	53	0
1	D	1448	0	1429	41	1
2	A	10	0	0	4	0
2	B	10	0	0	0	0
2	C	10	0	0	1	0
2	D	10	0	0	3	0
3	A	179	0	0	5	0
3	B	194	0	0	7	1
3	C	120	0	0	10	1
3	D	116	0	0	9	1
All	All	6484	0	5789	172	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 15.

The worst 5 of 172 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:91:PHE:N	3:B:1032:HOH:O	1.62	1.26
1:B:139:GLU:OE1	3:B:1072:HOH:O	1.65	1.14
1:C:88:ARG:HH11	1:C:88:ARG:HG3	1.00	1.12
2:A:1003:SO4:O3	2:A:1003:SO4:O4	1.68	1.09
1:B:23:LEU:O	1:B:24:LYS:HB2	1.30	1.08

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)
3:C:1026:HOH:O	3:D:1092:HOH:O[7_555]	1.91	0.29
1:D:101:ARG:NE	3:B:1193:HOH:O[5_455]	2.11	0.09

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	187/199 (94%)	177 (95%)	9 (5%)	1 (0%)	29 41
1	B	187/199 (94%)	172 (92%)	12 (6%)	3 (2%)	9 13
1	C	187/199 (94%)	172 (92%)	11 (6%)	4 (2%)	7 8
1	D	187/199 (94%)	169 (90%)	13 (7%)	5 (3%)	5 5
All	All	748/796 (94%)	690 (92%)	45 (6%)	13 (2%)	9 11

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	24	LYS
1	C	24	LYS
1	D	91	PHE
1	D	25	ASP
1	D	29	PRO

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	153/166 (92%)	140 (92%)	13 (8%)	10 16
1	B	151/166 (91%)	139 (92%)	12 (8%)	12 19
1	C	146/166 (88%)	132 (90%)	14 (10%)	8 12
1	D	146/166 (88%)	125 (86%)	21 (14%)	3 3
All	All	596/664 (90%)	536 (90%)	60 (10%)	7 11

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	60	LEU
1	D	153	VAL
1	C	162	LYS
1	D	149	LEU
1	D	178	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	111	ASN
1	D	184	HIS
1	C	13	GLN
1	C	57	GLN
1	C	64	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

8 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	C	1007	-	4,4,4	0.71	0	6,6,6	0.78	0
2	SO4	D	1004	-	4,4,4	0.53	0	6,6,6	1.22	1 (16%)
2	SO4	B	1006	-	4,4,4	0.28	0	6,6,6	0.66	0
2	SO4	D	1008	-	4,4,4	0.43	0	6,6,6	0.80	0
2	SO4	A	1005	-	4,4,4	3.63	4 (100%)	6,6,6	2.44	3 (50%)
2	SO4	A	1003	-	4,4,4	5.10	4 (100%)	6,6,6	1.90	1 (16%)
2	SO4	C	1002	-	4,4,4	0.19	0	6,6,6	1.10	0
2	SO4	B	1001	-	4,4,4	0.23	0	6,6,6	0.51	0

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1003	SO4	O4-S	-6.10	0.97	1.47
2	A	1003	SO4	O2-S	-5.45	1.16	1.46
2	A	1003	SO4	O3-S	-5.24	1.04	1.47
2	A	1005	SO4	O2-S	-4.36	1.22	1.46
2	A	1005	SO4	O3-S	-4.19	1.13	1.47

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1003	SO4	O4-S-O2	-4.04	88.21	109.31
2	A	1005	SO4	O3-S-O1	-3.76	89.70	109.31
2	A	1005	SO4	O3-S-O2	3.57	127.96	109.31
2	D	1004	SO4	O4-S-O3	-2.25	99.45	109.06
2	A	1005	SO4	O4-S-O2	2.05	119.98	109.31

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1007	SO4	1	0
2	D	1004	SO4	3	0
2	A	1005	SO4	2	0
2	A	1003	SO4	2	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	189/199 (94%)	-0.44	5 (2%)	56	54	14, 29, 61, 90	0
1	B	189/199 (94%)	-0.50	1 (0%)	91	89	13, 26, 55, 78	0
1	C	189/199 (94%)	-0.43	3 (1%)	72	70	17, 31, 77, 96	0
1	D	189/199 (94%)	-0.40	5 (2%)	56	54	17, 29, 91, 122	0
All	All	756/796 (94%)	-0.44	14 (1%)	66	64	13, 29, 74, 122	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	30	ILE	3.7
1	A	180	HIS	3.7
1	B	178	SER	3.4
1	A	179	GLY	3.3
1	C	104	ILE	3.1

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
2	SO4	D	1004	5/5	0.90	0.28	62,64,65,66	0
2	SO4	C	1007	5/5	0.93	0.23	56,57,60,61	0
2	SO4	B	1006	5/5	0.95	0.21	61,63,63,65	0
2	SO4	A	1005	5/5	0.98	0.11	27,38,44,59	2
2	SO4	C	1002	5/5	0.98	0.10	42,43,43,44	0
2	SO4	A	1003	5/5	0.99	0.19	19,27,44,44	2
2	SO4	B	1001	5/5	0.99	0.08	45,45,48,48	0
2	SO4	D	1008	5/5	0.99	0.10	45,45,47,48	0

6.5 Other polymers (i)

There are no such residues in this entry.