



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 2, 2017 – 11:48 PM EDT

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 : unknown  
Resolution : 2.40 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

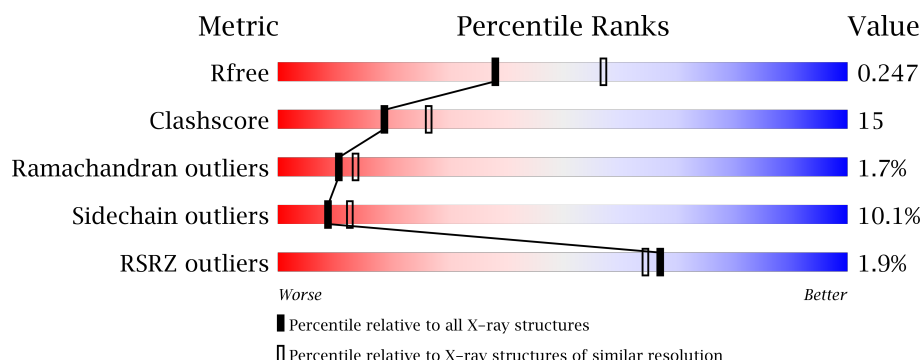
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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))
$R_{free}$	100719	3166 (2.40-2.40)
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	199	<div> <div>3%</div> <div> <div></div> <div>68%</div> <div>23%</div> <div>• • 5%</div> </div> </div>
1	B	199	<div> <div>%</div> <div> <div></div> <div>71%</div> <div>17%</div> <div>7% • 5%</div> </div> </div>
1	C	199	<div> <div>2%</div> <div> <div></div> <div>68%</div> <div>21%</div> <div>6% • 5%</div> </div> </div>
1	D	199	<div> <div>3%</div> <div> <div></div> <div>67%</div> <div>21%</div> <div>6% • 5%</div> </div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SO4	A	1003	-	X	X	-
2	SO4	A	1005	-	X	X	-
2	SO4	C	1007	-	-	-	X
2	SO4	D	1004	-	-	X	X

## 2 Entry composition

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

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<sub>4</sub>S).



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

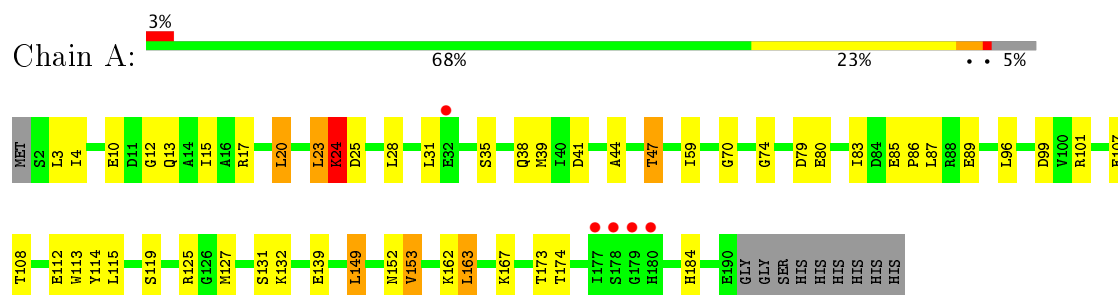
- Molecule 3 is water.

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

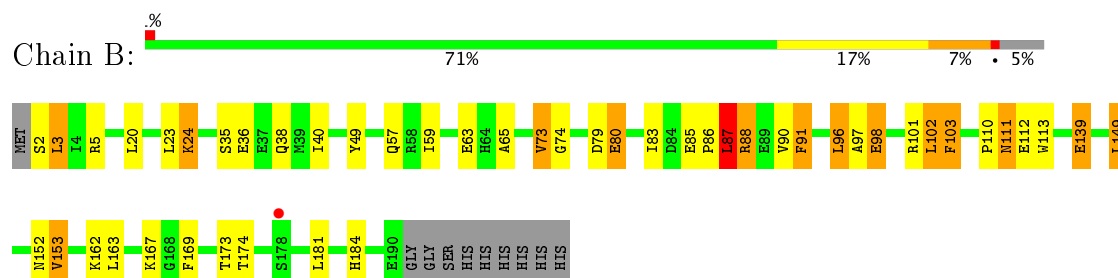
### 3 Residue-property plots [i](#)

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.

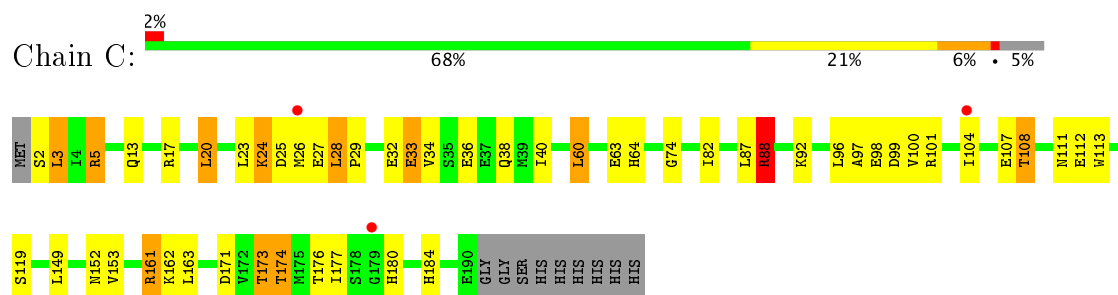
- Molecule 1: acetyltransferase, GNAT family



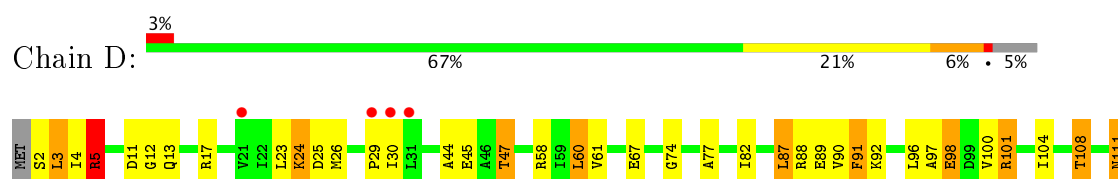
- Molecule 1: acetyltransferase, GNAT family

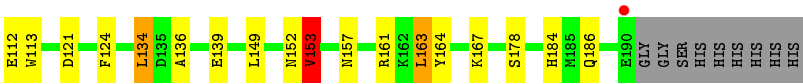


- Molecule 1: acetyltransferase, GNAT family



- Molecule 1: acetyltransferase, GNAT family







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	115.26 Å   115.26 Å   225.26 Å 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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	7.32 (at 2.41 Å)	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 (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	32.5	Xtriage
Anisotropy	0.213	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 45.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	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 <sup>2</sup> )	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.*

<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: 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%)

All (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
1	C	88	ARG	CG-CD	5.35	1.65	1.51
1	A	89	GLU	CD-OE1	5.29	1.31	1.25
1	B	73	VAL	CB-CG1	5.03	1.63	1.52
1	B	80	GLU	CD-OE2	5.00	1.31	1.25

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
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
1	B	101	ARG	NE-CZ-NH2	-8.84	115.88	120.30
1	C	5	ARG	NE-CZ-NH1	-8.80	115.90	120.30
1	D	88	ARG	NE-CZ-NH1	-8.59	116.00	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	88	ARG	NE-CZ-NH1	-7.61	116.50	120.30
1	B	101	ARG	NE-CZ-NH1	7.42	124.01	120.30
1	D	153	VAL	CB-CA-C	-7.16	97.80	111.40
1	C	153	VAL	CB-CA-C	-6.92	98.25	111.40
1	D	60	LEU	CB-CG-CD1	6.91	122.75	111.00
1	A	96	LEU	CA-CB-CG	6.85	131.06	115.30
1	C	5	ARG	NE-CZ-NH2	6.29	123.45	120.30
1	D	5	ARG	CG-CD-NE	-6.20	98.78	111.80
1	D	5	ARG	CD-NE-CZ	6.19	132.26	123.60
1	B	153	VAL	CB-CA-C	-6.08	99.85	111.40
1	B	5	ARG	NE-CZ-NH2	-5.98	117.31	120.30
1	A	149	LEU	CA-CB-CG	5.96	129.02	115.30
1	B	103	PHE	CB-CA-C	-5.86	98.68	110.40
1	A	153	VAL	CB-CA-C	-5.85	100.29	111.40
1	B	102	LEU	CA-CB-CG	5.69	128.38	115.30
1	C	60	LEU	CB-CG-CD1	5.67	120.64	111.00
1	D	134	LEU	CB-CG-CD2	-5.65	101.39	111.00
1	A	23	LEU	C-N-CA	-5.52	107.91	121.70
1	D	58	ARG	NE-CZ-NH1	5.23	122.91	120.30
1	A	99	ASP	CB-CG-OD1	5.22	123.00	118.30
1	B	96	LEU	CA-CB-CG	5.20	127.26	115.30
1	B	87	LEU	CB-CG-CD1	5.08	119.64	111.00
1	C	3	LEU	CA-CB-CG	5.04	126.89	115.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
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.

All (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
1:B:97:ALA:O	1:B:98:GLU:HB3	1.53	1.08
1:D:90:VAL:O	1:D:91:PHE:HB2	1.49	1.06
1:A:23:LEU:O	1:A:24:LYS:CB	2.04	1.02
1:C:88:ARG:HG3	1:C:88:ARG:NH1	1.73	1.00
1:C:2:SER:OG	1:C:64:HIS:HD2	1.47	0.97
1:B:90:VAL:C	3:B:1032:HOH:O	1.87	0.96
1:B:90:VAL:CA	3:B:1032:HOH:O	2.11	0.96
1:A:173:THR:HG22	1:A:174:THR:H	1.29	0.95
1:B:90:VAL:O	1:B:91:PHE:HB2	1.64	0.95
1:A:163:LEU:HA	1:C:82:ILE:HD11	1.47	0.94
1:A:23:LEU:O	1:A:24:LYS:HB3	1.65	0.94
1:C:23:LEU:O	1:C:24:LYS:HB2	1.65	0.94
2:A:1003:SO4:O3	2:A:1003:SO4:O2	1.83	0.93
1:B:139:GLU:CD	3:B:1072:HOH:O	2.03	0.89
1:C:88:ARG:HH11	1:C:88:ARG:CG	1.76	0.89
1:B:23:LEU:O	1:B:24:LYS:CB	2.13	0.87
2:D:1004:SO4:O4	3:D:1021:HOH:O	1.93	0.84
1:D:90:VAL:O	1:D:91:PHE:CB	2.26	0.83
1:D:12:GLY:HA2	1:D:47:THR:HG21	1.61	0.81
1:C:97:ALA:HB3	1:C:100:VAL:HG22	1.62	0.81
1:A:163:LEU:HA	1:C:82:ILE:CD1	2.11	0.80
2:C:1007:SO4:O4	3:C:1064:HOH:O	2.00	0.79

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:5:ARG:NE	3:C:1052:HOH:O	2.14	0.79
2:A:1005:SO4:O4	3:A:1099:HOH:O	1.98	0.79
1:D:112:GLU:OE2	1:D:184:HIS:HE1	1.66	0.78
1:C:88:ARG:NH1	1:C:88:ARG:CG	2.37	0.78
1:A:12:GLY:HA2	1:A:47:THR:HG21	1.65	0.78
1:C:5:ARG:CZ	3:C:1052:HOH:O	2.34	0.75
1:B:110:PRO:O	1:B:111:ASN:HB3	1.87	0.75
1:D:111:ASN:H	1:D:111:ASN:HD22	1.31	0.75
1:D:92:LYS:HB2	1:D:98:GLU:HG2	1.68	0.74
1:B:90:VAL:HA	3:B:1032:HOH:O	1.77	0.74
1:A:23:LEU:O	1:A:24:LYS:CG	2.36	0.74
1:B:97:ALA:O	1:B:98:GLU:CB	2.28	0.73
1:A:112:GLU:OE2	1:A:184:HIS:HE1	1.70	0.72
1:D:92:LYS:NZ	3:D:1061:HOH:O	2.19	0.72
1:C:5:ARG:NH1	1:C:63:GLU:HB2	2.05	0.71
1:A:173:THR:HG22	1:A:174:THR:N	2.03	0.70
1:C:108:THR:HG23	3:C:1061:HOH:O	1.91	0.70
1:B:174:THR:CG2	1:B:181:LEU:HD22	2.22	0.69
1:B:110:PRO:O	1:B:111:ASN:CB	2.33	0.69
1:B:174:THR:HG22	1:B:181:LEU:HD22	1.74	0.69
1:C:112:GLU:OE2	1:C:184:HIS:HE1	1.77	0.68
1:B:112:GLU:OE2	1:B:184:HIS:HE1	1.77	0.68
1:A:47:THR:HG22	3:A:1061:HOH:O	1.95	0.67
1:A:23:LEU:O	1:A:24:LYS:HG2	1.95	0.65
1:A:173:THR:CG2	1:A:174:THR:H	2.08	0.65
2:D:1004:SO4:O2	3:D:1060:HOH:O	2.14	0.64
1:D:101:ARG:HB2	1:D:101:ARG:HH11	1.62	0.64
1:B:90:VAL:O	1:B:91:PHE:CB	2.36	0.64
1:D:13:GLN:HB3	3:D:1070:HOH:O	1.99	0.63
1:B:59:ILE:HG12	1:B:73:VAL:HG22	1.81	0.63
1:C:161:ARG:CG	1:C:161:ARG:HH11	2.12	0.63
1:C:5:ARG:NH2	3:C:1052:HOH:O	2.32	0.63
1:D:97:ALA:O	1:D:98:GLU:HB2	2.00	0.62
1:D:112:GLU:OE2	1:D:184:HIS:CE1	2.52	0.60
1:D:47:THR:HG22	3:D:1014:HOH:O	2.01	0.60
1:A:4:ILE:HG13	1:B:3:LEU:HD21	1.83	0.60
1:C:2:SER:OG	1:C:64:HIS:CD2	2.40	0.60
1:C:101:ARG:NE	3:C:1066:HOH:O	2.02	0.60
1:B:102:LEU:O	1:B:103:PHE:HB2	2.01	0.60
1:C:34:VAL:HG22	1:C:38:GLN:HG3	1.84	0.59
1:D:5:ARG:NH2	1:D:11:ASP:OD2	2.32	0.59

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:132:LYS:NZ	1:C:98:GLU:OE2	2.36	0.58
1:B:88:ARG:NH1	3:B:1040:HOH:O	2.37	0.58
1:C:119:SER:HB3	3:C:1087:HOH:O	2.03	0.58
1:B:139:GLU:HA	1:B:139:GLU:OE1	2.03	0.58
1:C:24:LYS:C	1:C:26:MET:H	2.06	0.58
1:C:97:ALA:CB	1:C:100:VAL:HG22	2.33	0.57
1:C:5:ARG:HH11	1:C:63:GLU:HB2	1.70	0.57
1:A:101:ARG:HD2	3:A:1141:HOH:O	2.04	0.56
1:B:102:LEU:O	1:B:103:PHE:CB	2.52	0.56
1:A:127:MET:CE	1:C:99:ASP:HB3	2.35	0.56
1:A:112:GLU:OE2	1:A:184:HIS:CE1	2.55	0.56
1:D:134:LEU:HD23	1:D:167:LYS:HE3	1.88	0.56
2:D:1004:SO4:O3	3:D:1060:HOH:O	2.17	0.56
1:B:152:ASN:OD1	1:B:184:HIS:HD2	1.89	0.56
1:C:36:GLU:O	1:C:40:ILE:HG12	2.06	0.55
1:C:111:ASN:ND2	1:C:111:ASN:H	2.06	0.53
1:A:85:GLU:HB2	1:A:86:PRO:HD3	1.90	0.53
1:D:163:LEU:HD22	1:D:167:LYS:HE2	1.91	0.53
1:D:74:GLY:HA3	1:D:113:TRP:CH2	2.44	0.53
1:C:5:ARG:NH1	1:C:63:GLU:OE1	2.41	0.53
1:C:111:ASN:HD22	1:C:111:ASN:H	1.58	0.52
1:A:125:ARG:HB3	1:C:104:ILE:HD13	1.92	0.52
1:D:12:GLY:CA	1:D:47:THR:HG21	2.38	0.52
1:C:74:GLY:HA3	1:C:113:TRP:CH2	2.45	0.51
1:D:121:ASP:HB3	1:D:124:PHE:HD2	1.76	0.51
1:A:12:GLY:CA	1:A:47:THR:HG21	2.39	0.51
1:C:107:GLU:HB3	1:C:177:ILE:HD13	1.93	0.51
1:A:35:SER:OG	1:A:38:GLN:HG3	2.11	0.51
1:A:127:MET:HE3	1:C:99:ASP:HB3	1.92	0.51
1:B:65:ALA:HA	1:C:92:LYS:O	2.12	0.50
1:D:45:GLU:HG2	1:D:90:VAL:HG22	1.91	0.50
1:D:108:THR:CG2	3:D:1099:HOH:O	2.60	0.50
1:C:173:THR:HG22	1:C:174:THR:H	1.76	0.50
1:D:111:ASN:H	1:D:111:ASN:ND2	2.06	0.50
1:D:89:GLU:O	1:D:92:LYS:HB3	2.12	0.50
1:D:153:VAL:HG13	1:D:164:TYR:HE1	1.78	0.49
1:B:2:SER:N	1:C:98:GLU:HG3	2.27	0.49
1:D:111:ASN:HD21	1:D:186:GLN:HE22	1.60	0.49
1:C:2:SER:N	3:C:1114:HOH:O	2.46	0.49
1:A:24:LYS:HG3	1:A:25:ASP:N	2.24	0.49
1:C:28:LEU:HA	1:C:29:PRO:HD2	1.59	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:112:GLU:OE2	1:B:184:HIS:CE1	2.62	0.48
1:C:32:GLU:O	1:C:33:GLU:CB	2.61	0.48
1:A:31:LEU:HD23	1:A:39:MET:HE3	1.96	0.47
1:B:35:SER:H	1:B:38:GLN:NE2	2.12	0.47
1:C:34:VAL:CG2	1:C:38:GLN:HG3	2.44	0.47
1:D:24:LYS:C	1:D:26:MET:H	2.17	0.47
1:D:87:LEU:O	1:D:90:VAL:O	2.33	0.47
1:D:92:LYS:HB2	1:D:98:GLU:CG	2.42	0.47
1:A:80:GLU:OE1	1:A:107:GLU:HG2	2.14	0.46
1:A:44:ALA:O	1:A:47:THR:HB	2.15	0.46
1:D:97:ALA:HB3	1:D:100:VAL:CG1	2.46	0.46
1:D:17:ARG:NE	3:D:1070:HOH:O	2.48	0.46
1:C:20:LEU:O	1:C:23:LEU:O	2.33	0.46
1:C:13:GLN:HE21	1:C:17:ARG:NH2	2.14	0.46
1:B:87:LEU:O	1:B:90:VAL:O	2.34	0.45
1:C:34:VAL:HG22	1:C:38:GLN:CG	2.45	0.45
1:A:3:LEU:HD23	1:A:3:LEU:C	2.36	0.45
1:A:74:GLY:HA3	1:A:113:TRP:CH2	2.51	0.45
1:C:13:GLN:HE21	1:C:17:ARG:HH21	1.65	0.45
1:D:101:ARG:NH1	1:D:101:ARG:HB2	2.30	0.45
1:A:20:LEU:HD22	1:A:24:LYS:HE3	1.98	0.45
1:A:79:ASP:O	1:A:83:ILE:HG23	2.17	0.45
1:C:161:ARG:HH11	1:C:161:ARG:HG2	1.81	0.45
1:A:115:LEU:HG	2:A:1005:SO4:O3	2.17	0.45
1:B:74:GLY:HA3	1:B:113:TRP:CH2	2.52	0.45
1:B:36:GLU:O	1:B:40:ILE:HG12	2.17	0.45
1:C:152:ASN:OD1	1:C:184:HIS:HD2	2.00	0.44
1:C:2:SER:HA	1:C:63:GLU:O	2.17	0.44
1:A:24:LYS:CG	1:A:25:ASP:N	2.80	0.44
1:C:23:LEU:O	1:C:24:LYS:CB	2.44	0.44
1:D:152:ASN:OD1	1:D:184:HIS:HD2	2.00	0.44
1:D:5:ARG:HH22	1:D:11:ASP:CG	2.20	0.44
1:B:20:LEU:HD22	1:B:40:ILE:HD11	2.00	0.44
1:B:149:LEU:HD13	1:B:169:PHE:HZ	1.83	0.43
1:D:44:ALA:O	1:D:47:THR:HB	2.18	0.43
1:D:77:ALA:HB1	1:D:108:THR:HG23	1.99	0.43
1:A:108:THR:HG22	1:A:114:TYR:CG	2.52	0.43
1:A:107:GLU:HB2	1:A:152:ASN:ND2	2.33	0.43
1:A:24:LYS:CB	1:A:31:LEU:CD1	2.96	0.43
1:B:80:GLU:HG2	1:B:80:GLU:O	2.18	0.43
1:B:152:ASN:OD1	1:B:184:HIS:CD2	2.69	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:79:ASP:O	1:B:83:ILE:HG23	2.18	0.43
1:D:17:ARG:NH2	3:D:1070:HOH:O	2.52	0.43
1:B:35:SER:H	1:B:38:GLN:HE21	1.67	0.43
1:D:77:ALA:HB1	1:D:108:THR:CG2	2.49	0.43
1:D:111:ASN:N	1:D:111:ASN:HD22	2.06	0.42
1:A:20:LEU:HD22	1:A:24:LYS:CE	2.49	0.42
1:A:47:THR:CG2	3:A:1061:HOH:O	2.62	0.42
1:B:49:TYR:CG	1:B:86:PRO:HB3	2.55	0.42
1:A:132:LYS:HD2	3:A:1048:HOH:O	2.20	0.42
1:A:15:ILE:HD13	1:A:59:ILE:HD12	2.01	0.42
1:A:70:GLY:HA2	1:A:119:SER:O	2.20	0.41
1:D:97:ALA:HB3	1:D:100:VAL:HG13	2.02	0.41
1:D:4:ILE:HA	1:D:61:VAL:O	2.21	0.41
1:A:13:GLN:OE1	1:A:17:ARG:NH2	2.53	0.41
1:A:24:LYS:HB2	1:A:31:LEU:HD13	2.01	0.41
1:C:112:GLU:OE2	1:C:184:HIS:CE1	2.66	0.41
1:C:152:ASN:OD1	1:C:184:HIS:CD2	2.74	0.41
1:C:171:ASP:HB3	3:C:1097:HOH:O	2.19	0.41
1:C:176:THR:HA	1:C:180:HIS:O	2.21	0.41
1:C:92:LYS:HE2	3:C:1106:HOH:O	2.22	0.40
1:D:136:ALA:O	1:D:139:GLU:HG2	2.22	0.40
1:D:24:LYS:O	1:D:26:MET:N	2.52	0.40

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 ⓘ

### 5.3.1 Protein backbone ⓘ

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%)	32	46
1	B	187/199 (94%)	172 (92%)	12 (6%)	3 (2%)	11	15
1	C	187/199 (94%)	172 (92%)	11 (6%)	4 (2%)	8	9
1	D	187/199 (94%)	169 (90%)	13 (7%)	5 (3%)	6	6
All	All	748/796 (94%)	690 (92%)	45 (6%)	13 (2%)	11	13

All (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
1	D	30	ILE
1	B	91	PHE
1	B	98	GLU
1	D	24	LYS
1	A	24	LYS
1	C	25	ASP
1	C	27	GLU
1	C	33	GLU

### 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	153/166 (92%)	140 (92%)	13 (8%)	12	19
1	B	151/166 (91%)	139 (92%)	12 (8%)	14	22
1	C	146/166 (88%)	132 (90%)	14 (10%)	10	14
1	D	146/166 (88%)	125 (86%)	21 (14%)	4	4
All	All	596/664 (90%)	536 (90%)	60 (10%)	9	12

All (60) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	GLU
1	A	20	LEU
1	A	24	LYS
1	A	28	LEU
1	A	41	ASP
1	A	47	THR
1	A	87	LEU
1	A	131	SER
1	A	149	LEU
1	A	153	VAL
1	A	162	LYS
1	A	163	LEU
1	A	167	LYS
1	B	3	LEU
1	B	57	GLN
1	B	87	LEU
1	B	96	LEU
1	B	111	ASN
1	B	139	GLU
1	B	149	LEU
1	B	153	VAL
1	B	162	LYS
1	B	163	LEU
1	B	167	LYS
1	B	173	THR
1	C	3	LEU
1	C	20	LEU
1	C	28	LEU
1	C	60	LEU
1	C	87	LEU
1	C	88	ARG
1	C	96	LEU
1	C	108	THR
1	C	149	LEU
1	C	161	ARG
1	C	162	LYS
1	C	163	LEU
1	C	173	THR
1	C	174	THR
1	D	2	SER
1	D	3	LEU
1	D	5	ARG
1	D	23	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	47	THR
1	D	60	LEU
1	D	67	GLU
1	D	82	ILE
1	D	87	LEU
1	D	96	LEU
1	D	98	GLU
1	D	101	ARG
1	D	104	ILE
1	D	108	THR
1	D	111	ASN
1	D	149	LEU
1	D	153	VAL
1	D	157	ASN
1	D	161	ARG
1	D	163	LEU
1	D	178	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	64	HIS
1	A	152	ASN
1	A	184	HIS
1	A	186	GLN
1	B	38	GLN
1	B	184	HIS
1	B	186	GLN
1	C	13	GLN
1	C	57	GLN
1	C	64	HIS
1	C	111	ASN
1	C	184	HIS
1	C	186	GLN
1	D	64	HIS
1	D	111	ASN
1	D	157	ASN
1	D	184	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

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	A	1003	-	4,4,4	4.94	4 (100%)	6,6,6	1.83	2 (33%)
2	SO4	A	1005	-	4,4,4	3.52	4 (100%)	6,6,6	2.31	2 (33%)
2	SO4	B	1001	-	4,4,4	0.22	0	6,6,6	0.52	0
2	SO4	B	1006	-	4,4,4	0.36	0	6,6,6	0.63	0
2	SO4	C	1002	-	4,4,4	0.10	0	6,6,6	1.05	0
2	SO4	C	1007	-	4,4,4	0.79	0	6,6,6	0.76	0
2	SO4	D	1004	-	4,4,4	0.60	0	6,6,6	1.16	1 (16%)
2	SO4	D	1008	-	4,4,4	0.33	0	6,6,6	0.78	0

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	SO4	A	1003	-	-	0/0/0/0	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SO4	A	1005	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1001	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1006	-	-	0/0/0/0	0/0/0/0
2	SO4	C	1002	-	-	0/0/0/0	0/0/0/0
2	SO4	C	1007	-	-	0/0/0/0	0/0/0/0
2	SO4	D	1004	-	-	0/0/0/0	0/0/0/0
2	SO4	D	1008	-	-	0/0/0/0	0/0/0/0

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1003	SO4	O4-S	-5.86	0.97	1.47
2	A	1003	SO4	O2-S	-5.30	1.16	1.45
2	A	1003	SO4	O3-S	-5.02	1.04	1.47
2	A	1005	SO4	O2-S	-4.22	1.22	1.45
2	A	1005	SO4	O3-S	-3.99	1.13	1.47
2	A	1005	SO4	O4-S	-2.43	1.26	1.47
2	A	1003	SO4	O1-S	3.14	1.62	1.45
2	A	1005	SO4	O1-S	3.16	1.62	1.45

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1003	SO4	O4-S-O2	-3.82	88.21	109.26
2	A	1005	SO4	O3-S-O1	-3.55	89.70	109.26
2	D	1004	SO4	O4-S-O3	-2.11	99.45	108.96
2	A	1003	SO4	O2-S-O1	2.03	123.83	109.64
2	A	1005	SO4	O3-S-O2	3.39	127.96	109.26

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	A	1003	SO4	2	0
2	A	1005	SO4	2	0
2	C	1007	SO4	1	0
2	D	1004	SO4	3	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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, 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	189/199 (94%)	-0.44	5 (2%) 56 54	15, 29, 61, 91	0
1	B	189/199 (94%)	-0.50	1 (0%) 90 89	13, 26, 55, 78	0
1	C	189/199 (94%)	-0.44	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%) 67 64	13, 29, 74, 122	0

All (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
1	A	178	SER	3.0
1	C	26	MET	2.7
1	D	31	LEU	2.7
1	A	32	GLU	2.7
1	D	21	VAL	2.4
1	D	29	PRO	2.2
1	D	190	GLU	2.2
1	A	177	ILE	2.2
1	C	179	GLY	2.1

### 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 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SO4	D	1004	5/5	0.90	0.28	5.31	62,64,65,66	0
2	SO4	C	1007	5/5	0.93	0.23	2.41	56,57,60,61	0
2	SO4	B	1006	5/5	0.95	0.21	1.89	61,63,63,65	0
2	SO4	A	1003	5/5	0.99	0.19	0.76	19,27,44,44	2
2	SO4	A	1005	5/5	0.98	0.11	0.06	27,38,44,59	2
2	SO4	D	1008	5/5	0.99	0.10	-0.41	45,45,47,48	0
2	SO4	C	1002	5/5	0.98	0.10	-0.50	42,43,43,44	0
2	SO4	B	1001	5/5	0.99	0.08	-0.92	45,45,48,48	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.