



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 16, 2024 – 08:40 AM EDT

PDB ID : 5E7T  
Title : Structure of the tripod (BppUct-A-L) from the baseplate of bacteriophage Tuc2009  
Authors : Legrand, P.; Collins, B.; Blangy, S.; Murphy, J.; Spinelli, S.; Gutierrez, C.; Richet, N.; Kellenberger, C.; Desmyter, A.; Mahony, J.; van Sinderen, D.; Cambillau, C.  
Deposited on : 2015-10-13  
Resolution : 2.90 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

<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	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
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)	:	2.37.1

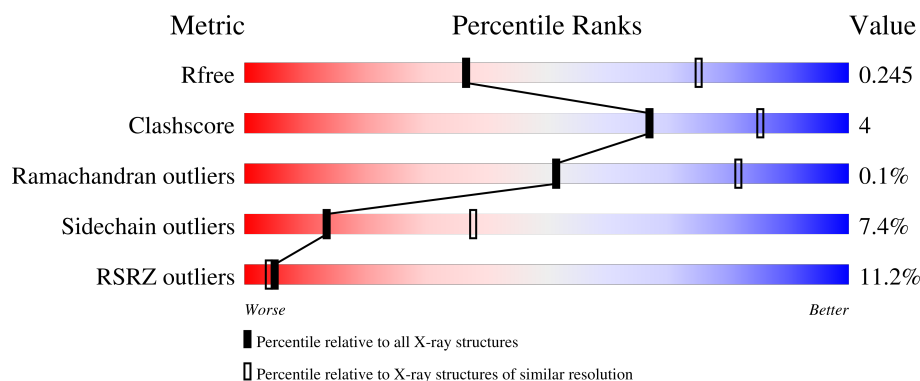
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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  $\geq 3$ , 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	130	<div> <div>90%</div> <div>8%</div> <div>..</div> </div>
2	B	286	<div> <div>82%</div> <div>17%</div> <div>.</div> </div>
3	G	173	<div> <div>23%</div> <div>88%</div> <div>11%</div> <div>.</div> </div>
4	H	174	<div> <div>17%</div> <div>86%</div> <div>13%</div> <div>..</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	I	174	
4	L	174	

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
5	SO4	L	202	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 8208 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 Minor structural protein 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	129	Total	C	N	O	S	0	0	0
			991	629	161	199	2			

- Molecule 2 is a protein called Minor structural protein 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	286	Total	C	N	O	S	0	0	0
			2238	1420	364	446	8			

- Molecule 3 is a protein called Major structural protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	172	Total	C	N	O	S	0	0	0
			1320	832	222	260	6			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	174	ALA	-	expression tag	UNP Q38610

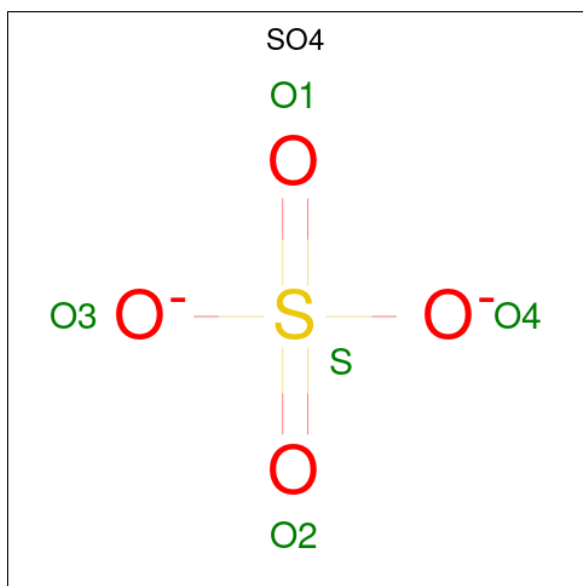
- Molecule 4 is a protein called Major structural protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	173	Total	C	N	O	S	0	0	0
			1325	835	223	261	6			
4	I	172	Total	C	N	O	S	0	0	0
			1320	832	222	260	6			
4	L	121	Total	C	N	O	S	0	0	0
			927	597	149	176	5			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	174	ALA	-	expression tag	UNP Q38610
I	174	ALA	-	expression tag	UNP Q38610
L	174	ALA	-	expression tag	UNP Q38610

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	L	1	Total	Ca	0	0
			1	1		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	14	Total 14	O 14	0	0
7	B	22	Total 22	O 22	0	0
7	G	3	Total 3	O 3	0	0
7	H	3	Total 3	O 3	0	0
7	I	6	Total 6	O 6	0	0
7	L	13	Total 13	O 13	0	0

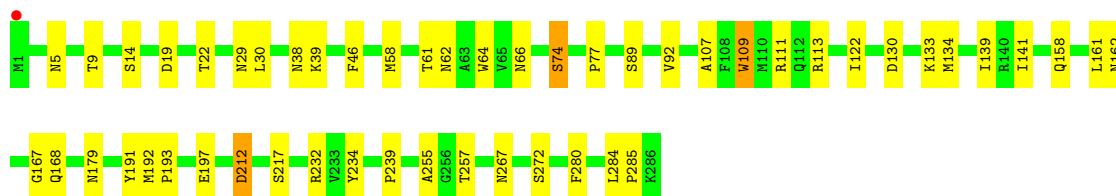
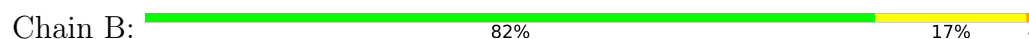
### 3 Residue-property plots [i](#)

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.

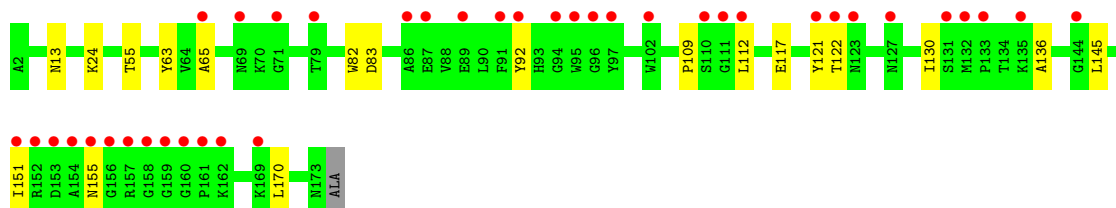
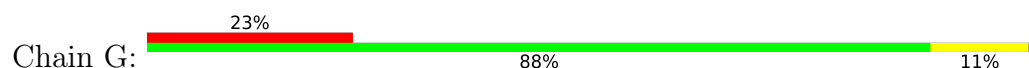
- Molecule 1: Minor structural protein 4



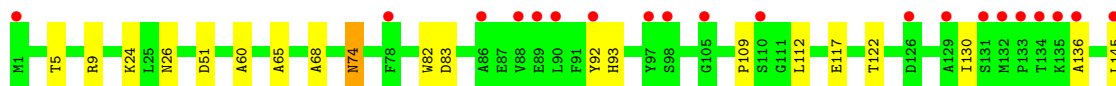
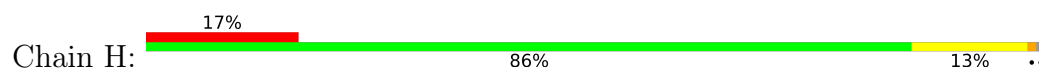
- Molecule 2: Minor structural protein 5

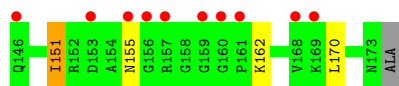


- Molecule 3: Major structural protein 1

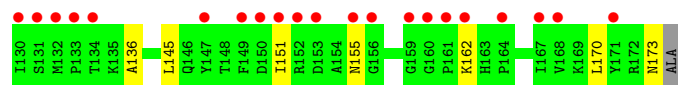
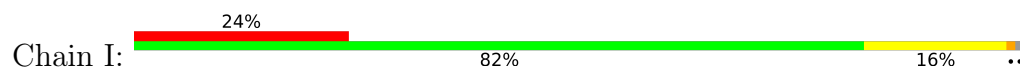


- Molecule 4: Major structural protein 1

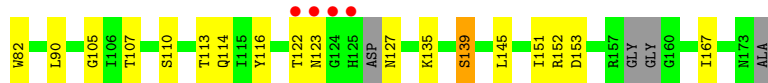
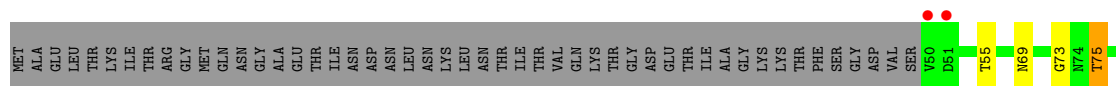




- Molecule 4: Major structural protein 1



- Molecule 4: Major structural protein 1





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	211.96Å 211.96Å 211.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.85 – 2.90 34.85 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (34.85-2.90) 99.9 (34.85-2.90)	Depositor EDS
$R_{merge}$	0.19	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.34 (at 2.90Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.214 , 0.237 0.225 , 0.245	Depositor DCC
$R_{free}$ test set	3508 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	80.9	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 58.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.023 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	8208	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.14% of the height of the origin peak. No significant pseudotranslation is detected.*

<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, CA

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	0.49	0/1010	0.76	0/1373
2	B	0.48	0/2295	0.74	0/3125
3	G	0.47	0/1349	0.66	0/1824
4	H	0.47	0/1354	0.68	0/1831
4	I	0.48	0/1349	0.67	0/1824
4	L	0.55	0/952	0.79	0/1288
All	All	0.49	0/8309	0.72	0/11265

There are no bond length outliers.

There are no bond angle outliers.

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	991	0	976	8	0
2	B	2238	0	2149	26	0
3	G	1320	0	1272	7	0
4	H	1325	0	1277	12	0
4	I	1320	0	1272	17	0
4	L	927	0	864	9	0
5	B	15	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	L	10	0	0	0	6
6	L	1	0	0	0	0
7	A	14	0	0	0	0
7	B	22	0	0	0	0
7	G	3	0	0	0	0
7	H	3	0	0	0	0
7	I	6	0	0	0	0
7	L	13	0	0	1	0
All	All	8208	0	7810	62	6

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 4.

All (62) 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)
2:B:130:ASP:O	4:L:139:SER:HB3	1.76	0.86
2:B:113:ARG:HE	2:B:158:GLN:HE21	1.29	0.80
4:L:105:GLY:HA3	4:L:152:ARG:NH1	2.00	0.76
4:L:122:THR:HA	4:L:127:ASN:HA	1.67	0.76
2:B:30:LEU:O	2:B:74:SER:HB2	1.90	0.71
1:A:202:HIS:HD2	1:A:204:SER:H	1.45	0.64
2:B:62:ASN:ND2	2:B:167:GLY:H	1.97	0.63
2:B:193:PRO:HB2	2:B:197:GLU:HG3	1.81	0.63
2:B:5:ASN:HB3	2:B:212:ASP:OD1	2.04	0.57
2:B:14:SER:HB3	2:B:19:ASP:HB3	1.86	0.56
2:B:29:ASN:HB2	2:B:192:MET:O	2.06	0.55
4:H:93:HIS:CE1	4:I:121:TYR:H	2.24	0.55
2:B:234:TYR:CE2	2:B:239:PRO:HG3	2.41	0.55
2:B:179:ASN:HB3	2:B:191:TYR:CD1	2.41	0.55
4:I:74:ASN:HD21	4:I:92:TYR:HE1	1.54	0.55
2:B:133:LYS:HG3	4:L:139:SER:OG	2.06	0.55
1:A:306:VAL:HG21	2:B:280:PHE:HB3	1.89	0.55
4:H:60:ALA:HB3	4:I:173:ASN:HB2	1.90	0.54
2:B:62:ASN:HD21	2:B:167:GLY:H	1.56	0.54
4:H:51:ASP:HA	4:I:57:LYS:HD3	1.89	0.54
2:B:109:TRP:CE3	2:B:111:ARG:HB3	2.44	0.53
2:B:58:MET:HE1	2:B:161:LEU:HG	1.91	0.51
4:H:65:ALA:HB3	4:H:170:LEU:HB2	1.93	0.51
2:B:232:ARG:HH11	2:B:232:ARG:HG3	1.76	0.51
3:G:121:TYR:H	4:I:93:HIS:CE1	2.29	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:H:74:ASN:HD22	4:H:151:ILE:HG22	1.74	0.50
4:H:26:ASN:OD1	4:I:2:ALA:HB3	2.13	0.49
1:A:201:GLU:HG2	1:A:250:LYS:HG2	1.94	0.49
1:A:216:LYS:HG2	1:A:278:THR:HG23	1.95	0.48
1:A:197:ASN:HD22	1:A:299:LYS:NZ	2.12	0.48
2:B:64:TRP:CZ3	2:B:107:ALA:HB3	2.49	0.48
2:B:66:ASN:HB3	2:B:162:ASN:HD22	1.79	0.47
4:I:65:ALA:HB3	4:I:170:LEU:HB2	1.95	0.47
3:G:130:ILE:HD12	4:H:130:ILE:HG21	1.95	0.47
4:H:109:PRO:HD2	4:H:112:LEU:HD12	1.97	0.46
1:A:202:HIS:CD2	1:A:204:SER:H	2.29	0.46
3:G:13:ASN:HA	4:H:9:ARG:HH22	1.80	0.46
4:H:68:ALA:HB3	4:I:116:TYR:CD1	2.51	0.46
4:I:109:PRO:HD2	4:I:112:LEU:HD12	1.97	0.45
2:B:255:ALA:HB1	2:B:285:PRO:HD3	1.99	0.45
4:I:24:LYS:O	4:I:28:ILE:HG12	2.17	0.44
2:B:113:ARG:HE	2:B:158:GLN:NE2	2.08	0.44
3:G:65:ALA:HB3	3:G:170:LEU:HB2	1.99	0.44
4:I:54:PHE:HE2	4:I:56:MET:HG3	1.82	0.44
2:B:139:ILE:HG12	4:L:116:TYR:CE2	2.53	0.44
1:A:298:VAL:HG12	2:B:272:SER:HB3	2.00	0.43
2:B:77:PRO:HB2	2:B:122:ILE:HD13	2.01	0.43
2:B:141:ILE:HG12	4:L:114:GLN:HG2	2.00	0.43
3:G:109:PRO:HD2	3:G:112:LEU:HD12	2.00	0.43
4:L:73:GLY:HA3	7:L:308:HOH:O	2.19	0.42
3:G:117:GLU:HG3	3:G:136:ALA:HB2	2.01	0.42
4:I:2:ALA:HB2	4:I:28:ILE:CD1	2.49	0.42
4:I:117:GLU:HG3	4:I:136:ALA:HB2	2.02	0.42
4:I:74:ASN:HD22	4:I:74:ASN:HA	1.63	0.42
4:L:69:ASN:HB2	4:L:75:THR:O	2.19	0.41
2:B:39:LYS:HE2	2:B:46:PHE:O	2.20	0.41
2:B:134:MET:SD	4:L:139:SER:HB2	2.59	0.41
4:H:68:ALA:HB3	4:I:116:TYR:CG	2.55	0.41
4:H:117:GLU:HG3	4:H:136:ALA:HB2	2.03	0.41
3:G:63:TYR:OH	4:I:56:MET:HB2	2.21	0.40
4:I:2:ALA:HB2	4:I:28:ILE:HD13	2.03	0.40
1:A:202:HIS:CD2	1:A:204:SER:HB3	2.56	0.40

All (6) 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)
5:L:202:SO4:O1	5:L:202:SO4:O2[5_555]	0.00	2.20
5:L:202:SO4:O1	5:L:202:SO4:O3[9_555]	0.00	2.20
5:L:202:SO4:O2	5:L:202:SO4:O3[5_555]	0.00	2.20
5:L:202:SO4:S	5:L:202:SO4:O1[5_555]	1.46	0.74
5:L:202:SO4:S	5:L:202:SO4:O2[5_555]	1.46	0.74
5:L:202:SO4:S	5:L:202:SO4:O3[5_555]	1.46	0.74

## 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	127/130 (98%)	120 (94%)	7 (6%)	0	100	100
2	B	284/286 (99%)	268 (94%)	15 (5%)	1 (0%)	34	66
3	G	170/173 (98%)	161 (95%)	9 (5%)	0	100	100
4	H	171/174 (98%)	163 (95%)	8 (5%)	0	100	100
4	I	170/174 (98%)	161 (95%)	9 (5%)	0	100	100
4	L	115/174 (66%)	114 (99%)	1 (1%)	0	100	100
All	All	1037/1111 (93%)	987 (95%)	49 (5%)	1 (0%)	51	82

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	212	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	111/111 (100%)	107 (96%)	4 (4%)	35	69
2	B	249/251 (99%)	236 (95%)	13 (5%)	23	55
3	G	138/138 (100%)	129 (94%)	9 (6%)	17	45
4	H	138/139 (99%)	127 (92%)	11 (8%)	12	33
4	I	138/139 (99%)	125 (91%)	13 (9%)	8	26
4	L	92/139 (66%)	78 (85%)	14 (15%)	3	8
All	All	866/917 (94%)	802 (93%)	64 (7%)	13	38

All (64) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	278	THR
1	A	291	THR
1	A	306	VAL
1	A	309	VAL
2	B	9	THR
2	B	22	THR
2	B	38	ASN
2	B	61	THR
2	B	74	SER
2	B	89	SER
2	B	92	VAL
2	B	109	TRP
2	B	168	GLN
2	B	217	SER
2	B	257	THR
2	B	267	ASN
2	B	284	LEU
3	G	24	LYS
3	G	55	THR
3	G	82	TRP
3	G	83	ASP
3	G	92	TYR
3	G	122	THR
3	G	145	LEU
3	G	151	ILE
3	G	155	ASN
4	H	5	THR
4	H	24	LYS
4	H	74	ASN
4	H	82	TRP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	H	83	ASP
4	H	92	TYR
4	H	122	THR
4	H	145	LEU
4	H	151	ILE
4	H	155	ASN
4	H	162	LYS
4	I	5	THR
4	I	20	ASP
4	I	33	THR
4	I	35	ASP
4	I	74	ASN
4	I	82	TRP
4	I	83	ASP
4	I	92	TYR
4	I	122	THR
4	I	145	LEU
4	I	151	ILE
4	I	155	ASN
4	I	162	LYS
4	L	55	THR
4	L	75	THR
4	L	82	TRP
4	L	90	LEU
4	L	107	THR
4	L	110	SER
4	L	113	THR
4	L	123	ASN
4	L	135	LYS
4	L	139	SER
4	L	145	LEU
4	L	151	ILE
4	L	153	ASP
4	L	167	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (19) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	197	ASN
1	A	202	HIS
2	B	51	ASN
2	B	62	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	158	GLN
2	B	162	ASN
2	B	182	GLN
3	G	93	HIS
3	G	125	HIS
4	H	31	GLN
4	H	74	ASN
4	H	93	HIS
4	H	125	HIS
4	H	163	HIS
4	I	74	ASN
4	I	93	HIS
4	I	125	HIS
4	I	163	HIS
4	L	123	ASN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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
5	SO4	L	202	-	4,4,4	0.15	0	6,6,6	0.04	0
5	SO4	B	301	-	4,4,4	0.19	0	6,6,6	0.21	0
5	SO4	L	203	-	4,4,4	0.16	0	6,6,6	0.17	0
5	SO4	B	303	-	4,4,4	0.25	0	6,6,6	0.20	0
5	SO4	B	302	-	4,4,4	0.18	0	6,6,6	0.18	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 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	L	202	SO4	0	6

## 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

### 6.1 Protein, DNA and RNA chains

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	129/130 (99%)	-0.56	1 (0%) 86 86	48, 64, 90, 129	0
2	B	286/286 (100%)	-0.44	1 (0%) 94 94	51, 68, 108, 173	0
3	G	172/173 (99%)	0.94	39 (22%) 0 0	54, 153, 177, 193	0
4	H	173/174 (99%)	0.80	30 (17%) 1 1	58, 150, 188, 202	0
4	I	172/174 (98%)	0.94	41 (23%) 0 0	53, 152, 190, 203	0
4	L	121/174 (69%)	0.01	6 (4%) 28 25	49, 62, 137, 181	0
All	All	1053/1111 (94%)	0.25	118 (11%) 5 4	48, 86, 181, 203	0

All (118) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	I	123	ASN	8.0
4	I	156	GLY	6.7
4	I	155	ASN	5.5
4	I	122	THR	5.4
4	H	97	TYR	5.2
3	G	160	GLY	5.1
4	I	134	THR	5.0
4	H	157	ARG	4.8
4	I	133	PRO	4.8
3	G	92	TYR	4.7
3	G	135	LYS	4.5
4	H	134	THR	4.5
4	I	132	MET	4.4
3	G	91	PHE	4.4
4	I	160	GLY	4.3
4	H	98	SER	4.3
3	G	154	ALA	4.3
3	G	131	SER	4.3
4	I	131	SER	4.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	G	111	GLY	4.2
4	I	152	ARG	4.1
3	G	121	TYR	4.0
4	H	131	SER	4.0
3	G	155	ASN	3.9
4	H	155	ASN	3.8
3	G	95	TRP	3.8
3	G	132	MET	3.7
3	G	89	GLU	3.7
4	H	105	GLY	3.7
4	H	135	LYS	3.7
4	I	162	LYS	3.7
2	B	1	MET	3.7
3	G	65	ALA	3.7
4	H	89	GLU	3.6
4	I	116	TYR	3.6
3	G	169	LYS	3.6
4	I	129	ALA	3.6
4	L	123	ASN	3.5
4	I	105	GLY	3.5
4	I	151	ILE	3.5
4	I	159	GLY	3.4
3	G	97	TYR	3.4
4	H	169	LYS	3.3
1	A	193	PRO	3.3
3	G	151	ILE	3.3
4	L	125	HIS	3.3
4	H	90	LEU	3.3
4	I	167	ILE	3.3
3	G	152	ARG	3.3
3	G	102	TRP	3.2
4	H	132	MET	3.2
4	H	126	ASP	3.2
4	I	112	LEU	3.2
4	H	86	ALA	3.2
3	G	144	GLY	3.1
3	G	153	ASP	3.1
4	I	164	PRO	3.1
3	G	112	LEU	3.1
4	H	156	GLY	3.1
4	I	92	TYR	3.0
4	I	115	ILE	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	I	90	LEU	3.0
4	H	78	PHE	3.0
4	H	110	SER	2.9
4	H	133	PRO	2.8
3	G	157	ARG	2.8
4	I	127	ASN	2.8
4	H	161	PRO	2.7
4	H	129	ALA	2.7
4	H	136	ALA	2.7
3	G	156	GLY	2.7
3	G	94	GLY	2.7
3	G	69	ASN	2.7
4	I	147	TYR	2.7
3	G	161	PRO	2.6
4	I	107	THR	2.6
4	I	130	ILE	2.6
4	I	168	VAL	2.6
4	L	124	GLY	2.6
4	I	91	PHE	2.6
4	I	98	SER	2.6
4	I	71	GLY	2.6
4	I	93	HIS	2.5
4	H	88	VAL	2.5
4	I	161	PRO	2.5
4	L	50	VAL	2.5
3	G	86	ALA	2.5
3	G	71	GLY	2.5
3	G	133	PRO	2.5
4	I	153	ASP	2.4
3	G	158	GLY	2.4
4	H	1	MET	2.4
4	H	168	VAL	2.4
4	L	122	THR	2.4
3	G	127	ASN	2.4
3	G	87	GLU	2.3
4	H	146	GLN	2.3
3	G	162	LYS	2.3
3	G	79	THR	2.2
3	G	123	ASN	2.2
4	I	149	PHE	2.2
4	H	145	LEU	2.2
4	I	111	GLY	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	G	96	GLY	2.2
4	I	150	ASP	2.2
4	I	121	TYR	2.2
3	G	159	GLY	2.2
4	H	160	GLY	2.2
4	H	92	TYR	2.2
4	I	97	TYR	2.2
4	I	120	GLY	2.1
4	H	153	ASP	2.1
4	L	51	ASP	2.1
4	I	171	TYR	2.1
4	H	159	GLY	2.1
3	G	122	THR	2.1
4	I	86	ALA	2.0
3	G	110	SER	2.0

## 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, 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	B-factors( $\text{\AA}^2$ )	Q<0.9
5	SO4	B	303	5/5	0.92	0.22	146,146,147,147	0
5	SO4	L	203	5/5	0.93	0.21	142,143,144,144	0
5	SO4	B	302	5/5	0.94	0.11	131,132,133,134	0
6	CA	L	201	1/1	0.95	0.41	75,75,75,75	0
5	SO4	B	301	5/5	0.98	0.18	132,133,133,134	0
5	SO4	L	202	5/5	0.99	0.27	156,156,157,157	2

## 6.5 Other polymers [i](#)

There are no such residues in this entry.