



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 21, 2020 – 07:52 PM EST

PDB ID : 6XRR  
Title : Structure of SciW bound to the Rhs1 Transmembrane Domain from Salmonella typhimurium  
Authors : Sachar, K.; Ahmad, S.; Whitney, J.C.; Prehna, G.  
Deposited on : 2020-07-13  
Resolution : 1.90 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

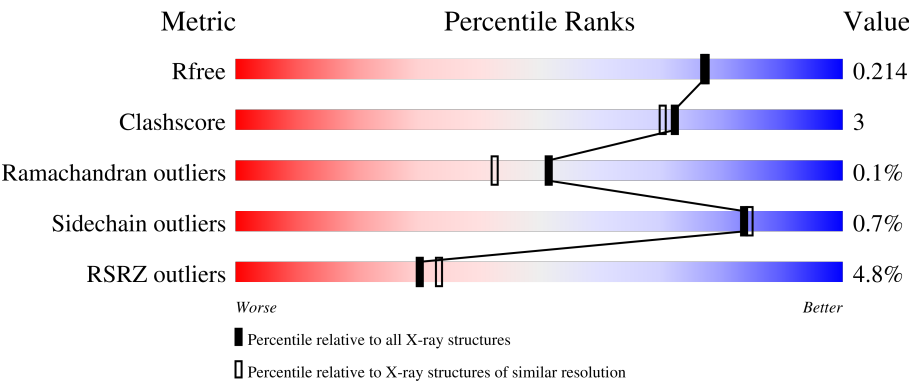
MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.16  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.16

# 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 1.90 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 <sub>free</sub>	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	148	<div><div>3%</div><div>95%</div><div>..</div></div>
1	B	148	<div><div>3%</div><div>91%</div><div>9% .</div></div>
1	D	148	<div><div>5%</div><div>94%</div><div>..</div></div>
1	E	148	<div><div>2%</div><div>93%</div><div>5% .</div></div>
1	F	148	<div><div>3%</div><div>91%</div><div>6% ..</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	148	
1	I	148	
1	K	148	
2	C	66	
2	G	66	
2	J	66	
2	L	66	

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
3	SO4	D	201	-	-	X	-
3	SO4	E	203	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 22160 atoms, of which 10488 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 Putative cytoplasmic protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	E	146	Total	C	H	N	O	S	0	0	0
			2297	726	1146	204	217	4			
1	I	144	Total	C	H	N	O	S	0	0	0
			2256	715	1122	201	215	3			
1	A	145	Total	C	H	N	O	S	0	0	0
			2278	721	1135	203	216	3			
1	B	147	Total	C	H	N	O	S	0	2	0
			2341	738	1168	210	220	5			
1	D	144	Total	C	H	N	O	S	22	2	0
			2290	724	1142	203	217	4			
1	F	145	Total	C	H	N	O	S	14	0	0
			2277	721	1134	203	216	3			
1	H	141	Total	C	H	N	O	S	6	0	0
			2222	704	1112	195	207	4			
1	K	146	Total	C	H	N	O	S	0	0	0
			2297	726	1146	204	217	4			

- Molecule 2 is a protein called RHS repeat protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	J	50	Total	C	H	N	O	S	0	0	0
			726	244	366	53	61	2			
2	C	50	Total	C	H	N	O	S	0	0	0
			725	244	365	53	61	2			
2	G	41	Total	C	H	N	O	S	0	0	0
			595	200	306	42	45	2			
2	L	46	Total	C	H	N	O	S	19	0	0
			678	228	346	49	53	2			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	0	MET	-	expression tag	UNP A0A5X2E0H7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
J	1	GLY	-	expression tag	UNP A0A5X2E0H7
J	60	HIS	-	expression tag	UNP A0A5X2E0H7
J	61	HIS	-	expression tag	UNP A0A5X2E0H7
J	62	HIS	-	expression tag	UNP A0A5X2E0H7
J	63	HIS	-	expression tag	UNP A0A5X2E0H7
J	64	HIS	-	expression tag	UNP A0A5X2E0H7
J	65	HIS	-	expression tag	UNP A0A5X2E0H7
C	0	MET	-	expression tag	UNP A0A5X2E0H7
C	1	GLY	-	expression tag	UNP A0A5X2E0H7
C	60	HIS	-	expression tag	UNP A0A5X2E0H7
C	61	HIS	-	expression tag	UNP A0A5X2E0H7
C	62	HIS	-	expression tag	UNP A0A5X2E0H7
C	63	HIS	-	expression tag	UNP A0A5X2E0H7
C	64	HIS	-	expression tag	UNP A0A5X2E0H7
C	65	HIS	-	expression tag	UNP A0A5X2E0H7
G	0	MET	-	expression tag	UNP A0A5X2E0H7
G	1	GLY	-	expression tag	UNP A0A5X2E0H7
G	60	HIS	-	expression tag	UNP A0A5X2E0H7
G	61	HIS	-	expression tag	UNP A0A5X2E0H7
G	62	HIS	-	expression tag	UNP A0A5X2E0H7
G	63	HIS	-	expression tag	UNP A0A5X2E0H7
G	64	HIS	-	expression tag	UNP A0A5X2E0H7
G	65	HIS	-	expression tag	UNP A0A5X2E0H7
L	0	MET	-	expression tag	UNP A0A5X2E0H7
L	1	GLY	-	expression tag	UNP A0A5X2E0H7
L	60	HIS	-	expression tag	UNP A0A5X2E0H7
L	61	HIS	-	expression tag	UNP A0A5X2E0H7
L	62	HIS	-	expression tag	UNP A0A5X2E0H7
L	63	HIS	-	expression tag	UNP A0A5X2E0H7
L	64	HIS	-	expression tag	UNP A0A5X2E0H7
L	65	HIS	-	expression tag	UNP A0A5X2E0H7

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	E	176	Total 176	O 176	0	0
4	I	152	Total 152	O 152	0	0
4	J	19	Total 19	O 19	0	0
4	A	168	Total 168	O 168	0	0
4	B	154	Total 154	O 154	0	0
4	C	36	Total 36	O 36	0	0
4	D	131	Total 131	O 131	0	0
4	F	104	Total 104	O 104	0	0
4	G	8	Total 8	O 8	0	0
4	H	124	Total 124	O 124	0	0
4	K	38	Total 38	O 38	0	0
4	L	8	Total 8	O 8	0	0

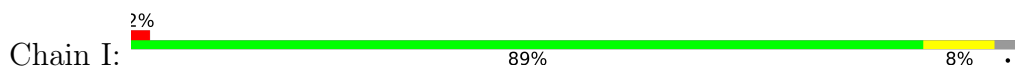
### 3 Residue-property plots [i](#)

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.

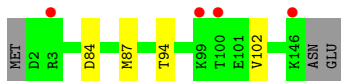
- Molecule 1: Putative cytoplasmic protein



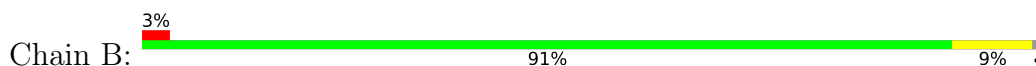
- Molecule 1: Putative cytoplasmic protein



- Molecule 1: Putative cytoplasmic protein



- Molecule 1: Putative cytoplasmic protein

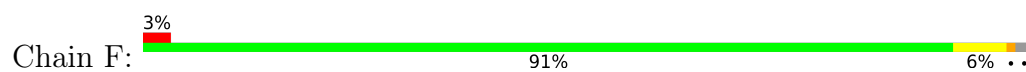


- Molecule 1: Putative cytoplasmic protein

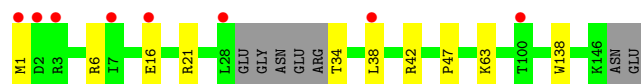
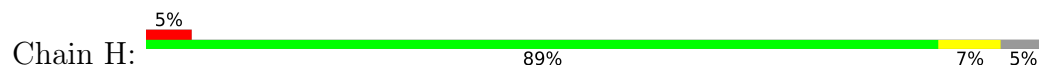


- Molecule 1: Putative cytoplasmic protein

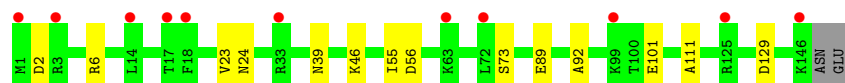
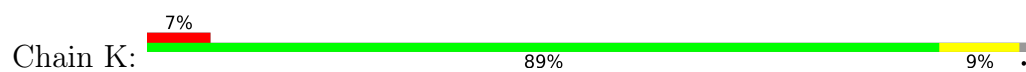




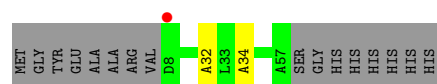
- Molecule 1: Putative cytoplasmic protein



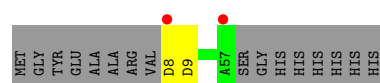
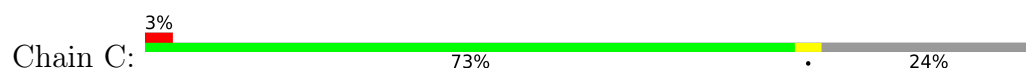
- Molecule 1: Putative cytoplasmic protein



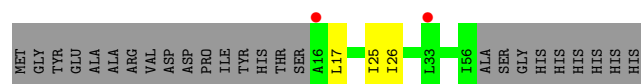
- Molecule 2: RHS repeat protein



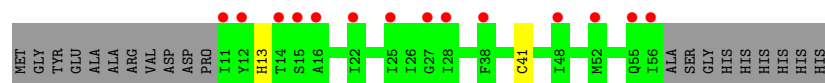
- Molecule 2: RHS repeat protein



- Molecule 2: RHS repeat protein



- Molecule 2: RHS repeat protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.31Å 105.31Å 248.36Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	91.20 – 1.90 91.20 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (91.20-1.90) 99.9 (91.20-1.90)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.19 (at 1.90Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.186 , 0.213 0.186 , 0.214	Depositor DCC
$R_{free}$ test set	6333 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.9	Xtriage
Anisotropy	0.053	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 47.5	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 -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	22160	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.72% 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

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	0.32	0/1168	0.53	0/1584
1	B	0.32	0/1201	0.52	0/1627
1	D	0.31	0/1172	0.53	0/1587
1	E	0.34	0/1176	0.54	0/1594
1	F	0.29	0/1168	0.52	0/1584
1	H	0.32	0/1134	0.54	0/1537
1	I	0.30	0/1159	0.53	0/1573
1	K	0.29	0/1176	0.50	0/1594
2	C	0.30	0/368	0.49	0/499
2	G	0.29	0/294	0.43	0/396
2	J	0.29	0/368	0.46	0/499
2	L	0.28	0/339	0.42	0/458
All	All	0.31	0/10723	0.52	0/14532

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	1143	1135	1134	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1173	1168	1169	11	0
1	D	1148	1142	1140	2	1
1	E	1151	1146	1146	11	0
1	F	1143	1134	1134	11	0
1	H	1110	1112	1111	13	0
1	I	1134	1122	1121	9	0
1	K	1151	1146	1146	8	1
2	C	360	365	365	2	0
2	G	289	306	306	3	0
2	J	360	366	365	2	0
2	L	332	346	345	2	0
3	A	10	0	0	0	0
3	B	20	0	0	1	0
3	D	5	0	0	2	0
3	E	15	0	0	0	0
3	H	5	0	0	0	0
3	K	5	0	0	0	0
4	A	168	0	0	3	2
4	B	154	0	0	9	0
4	C	36	0	0	2	0
4	D	131	0	0	3	0
4	E	176	0	0	4	0
4	F	104	0	0	4	0
4	G	8	0	0	0	0
4	H	124	0	0	2	1
4	I	152	0	0	3	0
4	J	19	0	0	1	0
4	K	38	0	0	3	0
4	L	8	0	0	0	0
All	All	11672	10488	10482	72	3

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 3.

All (72) 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:I:96:LYS:NZ	4:I:201:HOH:O	1.91	1.02
1:B:11[A]:CYS:SG	4:B:422:HOH:O	2.16	1.01
1:K:111:ALA:O	4:K:301:HOH:O	1.81	0.97
1:E:1:MET:HE1	4:E:473:HOH:O	1.65	0.96

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:34:THR:OG1	4:B:301:HOH:O	1.85	0.93
1:E:11:CYS:SG	4:E:417:HOH:O	2.28	0.89
1:E:1:MET:SD	4:F:286:HOH:O	2.34	0.86
1:B:47:PRO:O	4:B:302:HOH:O	1.94	0.84
1:B:29:GLU:OE2	4:B:303:HOH:O	1.99	0.80
3:B:201:SO4:O1	4:B:304:HOH:O	1.99	0.79
1:B:145:ASP:O	4:B:305:HOH:O	2.00	0.79
3:D:201:SO4:O3	4:D:301:HOH:O	1.99	0.79
1:B:64:LYS:NZ	4:B:306:HOH:O	2.11	0.79
1:A:87:MET:HE2	4:A:417:HOH:O	1.85	0.75
1:A:87:MET:CE	4:A:417:HOH:O	2.37	0.72
1:K:101:GLU:O	4:K:302:HOH:O	2.06	0.72
1:I:11:CYS:SG	4:I:308:HOH:O	2.48	0.71
1:F:23:VAL:HG11	2:G:26:ILE:HD12	1.74	0.69
1:F:23:VAL:HG11	2:G:26:ILE:CD1	2.23	0.69
1:B:134:LEU:HD12	4:B:405:HOH:O	1.92	0.69
1:E:1:MET:SD	1:F:70:ARG:NH2	2.67	0.68
1:H:6:ARG:HE	1:K:6:ARG:HH21	1.41	0.67
1:D:84[B]:ASP:OD1	4:D:302:HOH:O	2.15	0.65
1:H:38:LEU:HD13	1:H:138:TRP:CD2	2.32	0.64
1:B:147:ASN:N	4:B:305:HOH:O	2.30	0.63
2:J:34:ALA:O	4:J:101:HOH:O	2.15	0.63
1:H:47:PRO:HA	4:H:302:HOH:O	2.01	0.60
1:H:1:MET:H2	1:H:16:GLU:HG3	1.66	0.59
1:F:28:LEU:HD12	1:F:35:SER:CB	2.34	0.58
1:F:28:LEU:HD12	1:F:35:SER:HB3	1.89	0.55
1:E:1:MET:CG	1:F:70:ARG:HH21	2.21	0.54
1:F:133:LEU:O	1:F:137:THR:HG23	2.09	0.53
1:I:100:THR:HG21	2:J:32:ALA:HB2	1.92	0.51
1:E:2:ASP:HB3	1:E:13:VAL:HB	1.92	0.51
1:E:3:ARG:HB2	4:E:351:HOH:O	2.11	0.50
1:H:34:THR:HG23	2:L:13:HIS:NE2	2.27	0.49
1:H:38:LEU:HD22	1:H:138:TRP:CZ3	2.47	0.49
1:K:73:SER:HB3	1:K:92:ALA:HB3	1.95	0.49
1:H:1:MET:H2	1:H:16:GLU:CG	2.26	0.48
1:K:129:ASP:OD2	4:K:303:HOH:O	2.20	0.48
1:H:63:LYS:HE3	4:H:409:HOH:O	2.13	0.48
1:I:81:THR:HG23	1:I:140:ALA:O	2.14	0.47
1:E:1:MET:HG3	1:F:70:ARG:HH21	1.80	0.47
1:K:23:VAL:HG23	1:K:23:VAL:O	2.14	0.47
1:H:1:MET:N	1:H:16:GLU:HG3	2.29	0.47

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:1:MET:CE	4:F:286:HOH:O	2.63	0.47
1:I:97:SER:O	1:I:100:THR:HG22	2.15	0.46
1:B:1:MET:O	1:B:3:ARG:NH2	2.49	0.45
2:G:17:LEU:HD13	2:G:25:ILE:HD12	1.98	0.45
1:I:105:ARG:HG3	1:I:127:PHE:CE2	2.51	0.45
1:F:63:LYS:NZ	4:F:210:HOH:O	2.50	0.44
1:K:55:ILE:HD12	1:K:89:GLU:HB2	1.99	0.44
1:F:29:GLU:N	4:F:205:HOH:O	2.42	0.44
1:H:38:LEU:HD13	1:H:138:TRP:CE2	2.52	0.44
1:B:121:LEU:HD11	1:B:134:LEU:HD23	1.98	0.44
2:C:8:ASP:HA	4:C:127:HOH:O	2.17	0.43
1:H:21:ARG:HD2	2:L:41:CYS:HB3	1.99	0.43
1:E:70:ARG:NH2	4:E:308:HOH:O	2.49	0.43
1:I:3:ARG:HD2	1:I:18:PHE:HE1	1.82	0.42
1:I:87:MET:HG2	4:I:278:HOH:O	2.20	0.42
1:H:38:LEU:HD13	1:H:138:TRP:CE3	2.55	0.42
1:I:3:ARG:HD2	1:I:18:PHE:CE1	2.53	0.42
2:C:8:ASP:N	4:C:104:HOH:O	2.53	0.42
1:E:24:ASN:O	1:E:39:ASN:HA	2.19	0.42
1:A:94:THR:HA	1:A:102:VAL:O	2.20	0.42
3:D:201:SO4:O2	4:D:303:HOH:O	2.18	0.41
1:F:68:GLN:HA	1:F:70:ARG:HH12	1.84	0.41
1:K:24:ASN:O	1:K:39:ASN:HA	2.20	0.41
1:D:121:LEU:C	1:D:121:LEU:HD23	2.41	0.41
1:A:84:ASP:HB2	4:A:304:HOH:O	2.21	0.40
1:B:51:LEU:HB3	1:B:52:PRO:HD3	2.04	0.40
1:H:1:MET:N	1:H:16:GLU:CG	2.85	0.40

All (3) 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)
4:A:303:HOH:O	4:A:303:HOH:O[6_554]	1.86	0.34
1:D:125:ARG:HE	1:K:56:ASP:OD2[5_664]	1.55	0.05
4:A:333:HOH:O	4:H:347:HOH:O[6_554]	2.18	0.02

## 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	143/148 (97%)	142 (99%)	1 (1%)	0	100	100
1	B	147/148 (99%)	146 (99%)	1 (1%)	0	100	100
1	D	142/148 (96%)	139 (98%)	3 (2%)	0	100	100
1	E	144/148 (97%)	143 (99%)	1 (1%)	0	100	100
1	F	143/148 (97%)	143 (100%)	0	0	100	100
1	H	137/148 (93%)	135 (98%)	2 (2%)	0	100	100
1	I	142/148 (96%)	142 (100%)	0	0	100	100
1	K	144/148 (97%)	143 (99%)	1 (1%)	0	100	100
2	C	48/66 (73%)	47 (98%)	0	1 (2%)	7	1
2	G	39/66 (59%)	39 (100%)	0	0	100	100
2	J	48/66 (73%)	48 (100%)	0	0	100	100
2	L	44/66 (67%)	44 (100%)	0	0	100	100
All	All	1321/1448 (91%)	1311 (99%)	9 (1%)	1 (0%)	51	43

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	9	ASP

### 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	123/126 (98%)	123 (100%)	0	100	100
1	B	127/126 (101%)	127 (100%)	0	100	100
1	D	124/126 (98%)	122 (98%)	2 (2%)	62	60
1	E	124/126 (98%)	124 (100%)	0	100	100
1	F	123/126 (98%)	121 (98%)	2 (2%)	62	60
1	H	120/126 (95%)	119 (99%)	1 (1%)	81	82
1	I	122/126 (97%)	121 (99%)	1 (1%)	81	82
1	K	124/126 (98%)	122 (98%)	2 (2%)	62	60
2	C	33/45 (73%)	33 (100%)	0	100	100
2	G	25/45 (56%)	25 (100%)	0	100	100
2	J	33/45 (73%)	33 (100%)	0	100	100
2	L	30/45 (67%)	30 (100%)	0	100	100
All	All	1108/1188 (93%)	1100 (99%)	8 (1%)	84	84

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	42	ARG
1	D	35	SER
1	D	42	ARG
1	F	32	GLU
1	F	63	LYS
1	H	42	ARG
1	K	2	ASP
1	K	46	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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

12 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	$\# Z  > 2$
3	SO4	B	204	-	4,4,4	0.15	0	6,6,6	0.05	0
3	SO4	A	202	-	4,4,4	0.15	0	6,6,6	0.07	0
3	SO4	E	201	-	4,4,4	0.14	0	6,6,6	0.08	0
3	SO4	H	201	-	4,4,4	0.14	0	6,6,6	0.05	0
3	SO4	K	201	-	4,4,4	0.14	0	6,6,6	0.05	0
3	SO4	E	203	-	4,4,4	0.14	0	6,6,6	0.05	0
3	SO4	B	203	-	4,4,4	0.13	0	6,6,6	0.09	0
3	SO4	B	202	-	4,4,4	0.15	0	6,6,6	0.06	0
3	SO4	A	201	-	4,4,4	0.14	0	6,6,6	0.08	0
3	SO4	D	201	-	4,4,4	0.12	0	6,6,6	0.06	0
3	SO4	B	201	-	4,4,4	0.14	0	6,6,6	0.09	0
3	SO4	E	202	-	4,4,4	0.13	0	6,6,6	0.05	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 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	201	SO4	2	0
3	B	201	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 ⓘ

### 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	145/148 (97%)	-0.02	4 (2%)	53	56	20, 29, 56, 106	0
1	B	147/148 (99%)	0.02	4 (2%)	54	57	21, 33, 63, 87	0
1	D	143/148 (96%)	0.15	7 (4%)	29	33	23, 34, 70, 106	0
1	E	146/148 (98%)	0.01	3 (2%)	63	66	19, 28, 59, 86	0
1	F	144/148 (97%)	0.20	5 (3%)	44	47	32, 44, 76, 103	0
1	H	141/148 (95%)	0.20	8 (5%)	23	26	23, 34, 73, 124	1 (0%)
1	I	144/148 (97%)	-0.01	3 (2%)	63	66	21, 32, 58, 78	0
1	K	146/148 (98%)	0.54	11 (7%)	14	15	28, 59, 86, 98	0
2	C	50/66 (75%)	0.20	2 (4%)	38	41	22, 36, 67, 78	0
2	G	41/66 (62%)	0.58	2 (4%)	29	33	30, 54, 82, 92	0
2	J	50/66 (75%)	0.13	1 (2%)	65	68	25, 36, 70, 80	0
2	L	44/66 (66%)	1.41	14 (31%)	0	0	32, 61, 99, 112	0
All	All	1341/1448 (92%)	0.19	64 (4%)	30	33	19, 37, 78, 124	1 (0%)

All (64) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	K	33	ARG	7.4
1	H	3	ARG	6.7
2	L	56	ILE	6.5
1	K	1	MET	6.2
2	L	12	TYR	5.3
1	H	2	ASP	5.0
1	K	14	LEU	4.9
1	D	33	ARG	4.6
2	L	14	THR	4.5
1	B	146	LYS	4.4
1	E	1	MET	4.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	147	ASN	4.2
1	B	1	MET	4.2
1	E	146	LYS	4.0
1	H	1	MET	3.9
1	D	29	GLU	3.8
2	J	8	ASP	3.7
2	C	8	ASP	3.6
1	F	3	ARG	3.6
1	K	99	LYS	3.6
1	K	18	PHE	3.6
1	K	146	LYS	3.6
1	D	1	MET	3.6
1	F	146	LYS	3.6
2	L	11	ILE	3.4
1	H	100	THR	3.3
1	F	8	GLN	3.1
1	H	7	ILE	3.1
2	L	48	ILE	3.1
2	C	57	ALA	3.0
1	I	3	ARG	2.9
2	L	16	ALA	2.9
2	L	25	ILE	2.9
1	D	34	THR	2.8
1	A	99	LYS	2.8
1	A	3	ARG	2.7
1	F	16	GLU	2.7
1	D	28	LEU	2.7
1	D	100	THR	2.7
1	F	14	LEU	2.5
2	L	52	MET	2.5
1	A	146	LYS	2.4
1	K	17	THR	2.4
2	L	28	ILE	2.4
1	H	16	GLU	2.4
1	D	3	ARG	2.4
1	H	28	LEU	2.4
2	L	55	GLN	2.3
1	E	99	LYS	2.3
2	G	33	LEU	2.3
1	I	2	ASP	2.3
2	L	38	PHE	2.3
1	K	72	LEU	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	33	ARG	2.2
1	K	3	ARG	2.2
2	L	15	SER	2.2
2	L	27	GLY	2.2
1	H	38	LEU	2.2
2	G	16	ALA	2.2
2	L	22	ILE	2.1
1	I	15	PRO	2.1
1	A	100	THR	2.1
1	K	63	LYS	2.0
1	K	125	ARG	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(Å <sup>2</sup> )	Q<0.9
3	SO4	E	203	5/5	0.62	0.55	278,279,279,280	0
3	SO4	H	201	5/5	0.70	0.30	140,155,162,162	0
3	SO4	E	202	5/5	0.78	0.23	133,133,133,134	0
3	SO4	K	201	5/5	0.81	0.20	137,138,139,139	0
3	SO4	B	202	5/5	0.86	0.23	117,119,120,122	0
3	SO4	B	204	5/5	0.86	0.17	124,125,125,127	0
3	SO4	B	203	5/5	0.87	0.15	90,91,91,92	0
3	SO4	A	201	5/5	0.89	0.17	93,94,97,98	0
3	SO4	D	201	5/5	0.91	0.15	66,71,77,81	0
3	SO4	B	201	5/5	0.92	0.14	77,85,87,88	0
3	SO4	A	202	5/5	0.93	0.19	104,106,107,108	0
3	SO4	E	201	5/5	0.98	0.10	70,71,72,73	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.