



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 04:05 am BST

PDB ID : 4UYP  
Title : High resolution structure of the third cohesin ScaC in complex with the ScaB dockerin with a mutation in the N-terminal helix (IN to SI) from *Acetivibrio cellulolyticus* displaying a type I interaction.  
Authors : Cameron, K.; Alves, V.D.; Bule, P.; Ferreira, L.M.A.; Fontes, C.M.G.A.; Najmudin, S.  
Deposited on : 2014-09-02  
Resolution : 1.49 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](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	:	2.11
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)	:	2.11

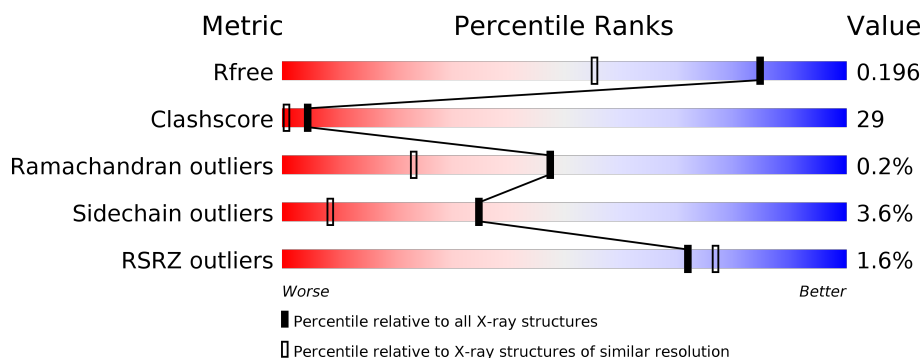
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.49 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	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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 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	151	<div> <div>%</div> <div> <div></div> <div>71%</div> <div>23%</div> <div>• 5%</div> </div> </div>
1	C	151	<div> <div>%</div> <div> <div></div> <div>71%</div> <div>20%</div> <div>5% • •</div> </div> </div>
2	B	75	<div> <div>3%</div> <div> <div></div> <div>68%</div> <div>24%</div> <div>• • •</div> </div> </div>
2	D	75	<div> <div>4%</div> <div> <div></div> <div>79%</div> <div>11%</div> <div>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
3	SO4	C	1151	-	-	X	X
4	EPE	A	1148	-	-	X	X
4	EPE	C	1152[B]	-	X	-	-
5	MPD	A	1149	-	-	X	-
5	MPD	C	1153	-	-	X	-
5	MPD	C	1154	-	-	X	-
5	MPD	D	1077	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 4392 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 Cellulosomal scaffoldin anchoring protein C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	144	Total	C	N	O	S	0	18	0
			1180	754	180	243	3			
1	C	146	Total	C	N	O	S	0	18	0
			1189	763	177	245	4			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP Q7WYN2
A	144	LEU	-	expression tag	UNP Q7WYN2
A	145	GLU	-	expression tag	UNP Q7WYN2
A	146	HIS	-	expression tag	UNP Q7WYN2
A	147	HIS	-	expression tag	UNP Q7WYN2
A	148	HIS	-	expression tag	UNP Q7WYN2
A	149	HIS	-	expression tag	UNP Q7WYN2
A	150	HIS	-	expression tag	UNP Q7WYN2
A	151	HIS	-	expression tag	UNP Q7WYN2
C	1	MET	-	initiating methionine	UNP Q7WYN2
C	144	LEU	-	expression tag	UNP Q7WYN2
C	145	GLU	-	expression tag	UNP Q7WYN2
C	146	HIS	-	expression tag	UNP Q7WYN2
C	147	HIS	-	expression tag	UNP Q7WYN2
C	148	HIS	-	expression tag	UNP Q7WYN2
C	149	HIS	-	expression tag	UNP Q7WYN2
C	150	HIS	-	expression tag	UNP Q7WYN2
C	151	HIS	-	expression tag	UNP Q7WYN2

- Molecule 2 is a protein called Cellulosomal scaffoldin adaptor protein B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	73	Total	C	N	O	S	0	8	0
			604	400	91	111	2			

*Continued on next page...*

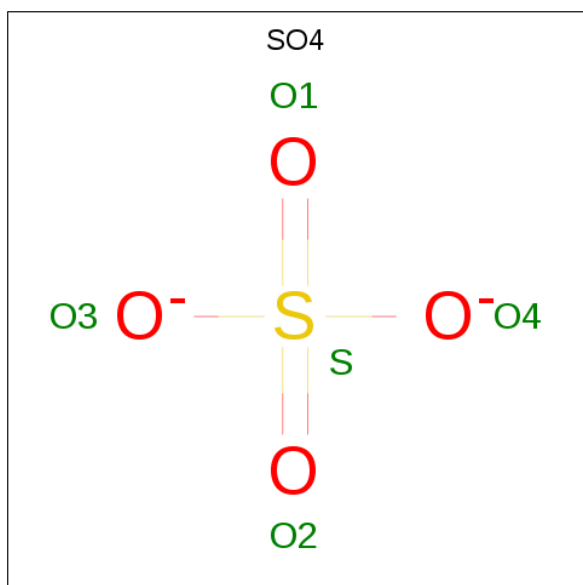
*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	73	Total	C	N	O	S	0	7	0
			602	396	92	112	2			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	15	SER	ILE	engineered mutation	UNP Q7WYN3
B	16	ILE	ASN	engineered mutation	UNP Q7WYN3
D	15	SER	ILE	engineered mutation	UNP Q7WYN3
D	16	ILE	ASN	engineered mutation	UNP Q7WYN3

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



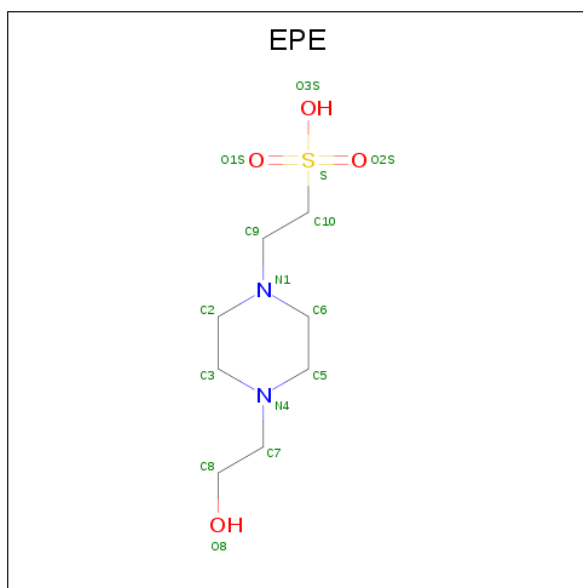
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		

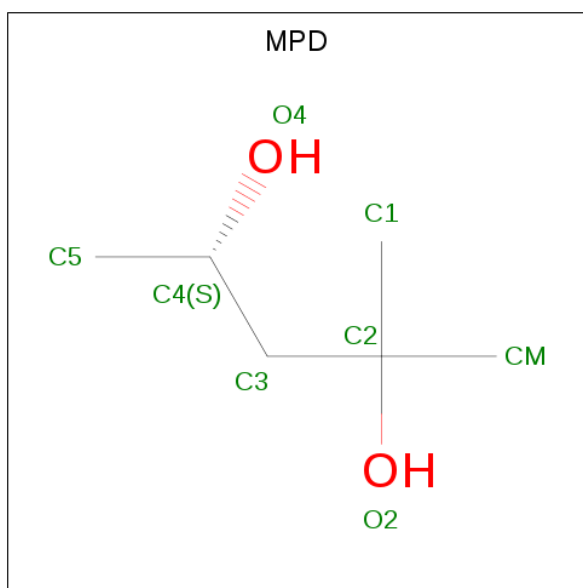
- Molecule 4 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula:  $C_8H_{18}N_2O_4S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
4	B	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
4	C	1	Total	C	N	O	S	0	1
			30	16	4	8	2		

- Molecule 5 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula:

C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			8	6	2		
5	C	1	Total	C	O	0	0
			8	6	2		
5	C	1	Total	C	O	0	0
			8	6	2		
5	C	1	Total	C	O	0	0
			8	6	2		
5	D	1	Total	C	O	0	0
			8	6	2		
5	D	1	Total	C	O	0	0
			8	6	2		

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	2	Total	Ca	0	0
			2	2		
6	D	2	Total	Ca	0	0
			2	2		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	243	Total	O	0	0
			243	243		

*Continued on next page...*

*Continued from previous page...*

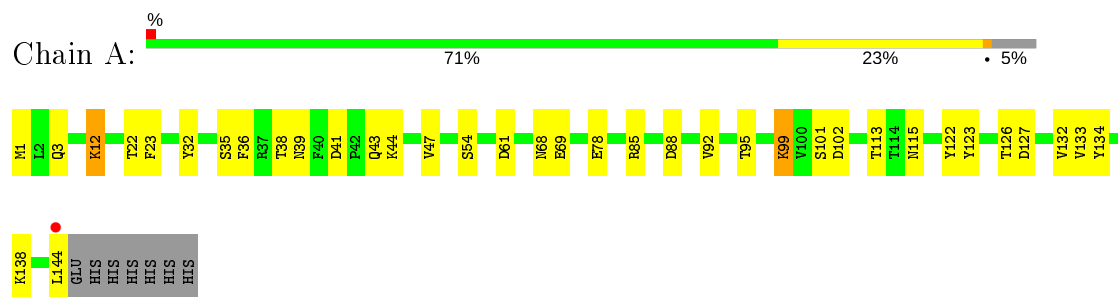
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	109	Total 109	O 109	0	0
7	C	211	Total 211	O 211	0	0
7	D	77	Total 77	O 77	0	0



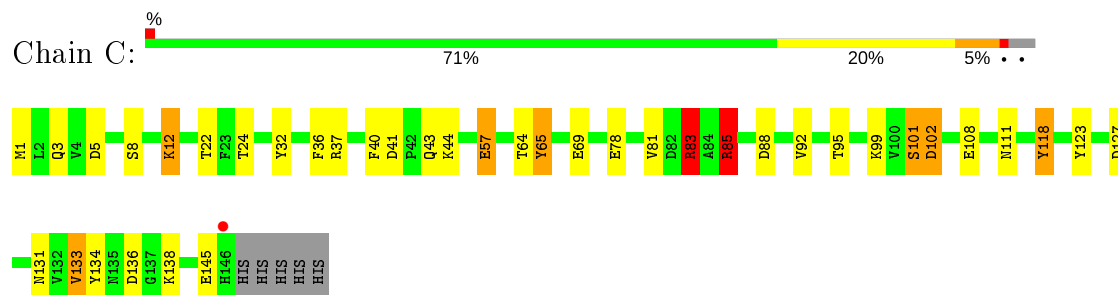
### 3 Residue-property plots [i](#)

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.

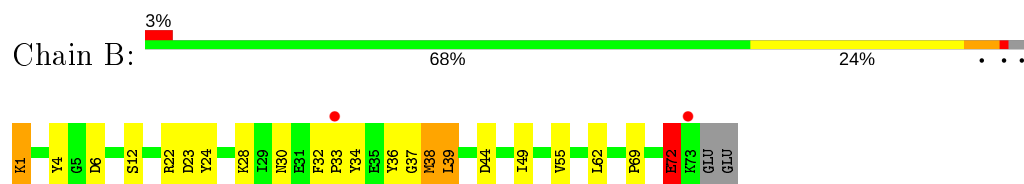
- Molecule 1: Cellulosomal scaffoldin anchoring protein C



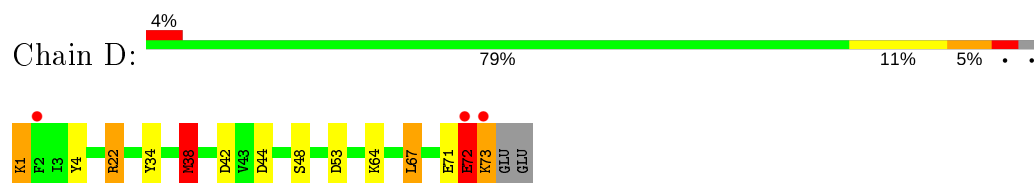
- Molecule 1: Cellulosomal scaffoldin anchoring protein C



- Molecule 2: Cellulosomal scaffoldin adaptor protein B



- Molecule 2: Cellulosomal scaffoldin adaptor protein B



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.75Å 107.75Å 100.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	76.19 – 1.49 76.19 – 1.49	Depositor EDS
% Data completeness (in resolution range)	99.8 (76.19-1.49) 99.8 (76.19-1.49)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.45 (at 1.49Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.155 , 0.185 0.164 , 0.196	Depositor DCC
$R_{free}$ test set	4843 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.2	Xtriage
Anisotropy	0.196	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 46.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	4392	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.34% 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: MPD, EPE, CA, 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	1.32	5/1245 (0.4%)	1.41	13/1690 (0.8%)
1	C	1.43	11/1263 (0.9%)	1.75	20/1717 (1.2%)
2	B	1.24	2/638 (0.3%)	1.40	8/862 (0.9%)
2	D	1.22	0/633	1.62	9/853 (1.1%)
All	All	1.33	18/3779 (0.5%)	1.57	50/5122 (1.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	134	TYR	CE1-CZ	8.77	1.50	1.38
1	C	108	GLU	CD-OE1	-7.27	1.17	1.25
1	A	85	ARG	CZ-NH1	6.91	1.42	1.33
1	C	83	ARG	CD-NE	-6.67	1.35	1.46
1	C	85	ARG	CZ-NH1	6.58	1.41	1.33

The worst 5 of 50 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	83	ARG	NE-CZ-NH2	28.59	134.60	120.30
1	C	83	ARG	NE-CZ-NH1	-27.90	106.35	120.30
2	D	22	ARG	NE-CZ-NH2	-18.90	110.85	120.30
1	C	83	ARG	CD-NE-CZ	12.53	141.15	123.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	42	ASP	CB-CG-OD1	11.28	128.45	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	83	ARG	Sidechain

## 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	1180	0	1188	70	0
1	C	1189	0	1203	50	0
2	B	604	0	634	44	0
2	D	602	0	630	29	0
3	A	20	0	0	1	0
3	B	10	0	0	1	0
3	C	25	0	0	4	0
3	D	10	0	0	0	0
4	A	15	0	18	21	0
4	B	15	0	18	1	0
4	C	30	0	35	7	0
5	A	8	0	14	8	0
5	C	24	0	41	20	0
5	D	16	0	28	12	0
6	B	2	0	0	0	0
6	D	2	0	0	0	0
7	A	243	0	0	42	4
7	B	109	0	0	11	1
7	C	211	0	0	30	5
7	D	77	0	0	5	1
All	All	4392	0	3809	216	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 29.

The worst 5 of 216 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:D:22:ARG:HH22	5:D:1077:MPD:C5	1.21	1.48
2:D:22:ARG:NH2	5:D:1077:MPD:C5	1.96	1.29
7:C:2140:HOH:O	2:D:38[B]:MET:HE1	1.26	1.26
2:D:22:ARG:NH2	5:D:1077:MPD:H52	1.53	1.23
2:B:38[A]:MET:HE2	2:B:49[A]:ILE:CD1	1.68	1.22

The worst 5 of 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)
7:A:2109:HOH:O	7:C:2207:HOH:O[4_544]	1.37	0.83
7:A:2138:HOH:O	7:C:2206:HOH:O[4_544]	1.93	0.27
7:C:2099:HOH:O	7:D:2041:HOH:O[3_445]	2.01	0.19
7:C:2008:HOH:O	7:C:2202:HOH:O[8_555]	2.05	0.15
7:A:2143:HOH:O	7:C:2207:HOH:O[4_544]	2.07	0.13

## 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	159/151 (105%)	159 (100%)	0	0	100	100
1	C	161/151 (107%)	161 (100%)	0	0	100	100
2	B	79/75 (105%)	79 (100%)	0	0	100	100
2	D	78/75 (104%)	77 (99%)	0	1 (1%)	12	2
All	All	477/452 (106%)	476 (100%)	0	1 (0%)	47	23

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	72	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	139/129 (108%)	137 (99%)	2 (1%)	67	42
1	C	141/129 (109%)	137 (97%)	4 (3%)	43	14
2	B	68/63 (108%)	63 (93%)	5 (7%)	13	1
2	D	68/63 (108%)	62 (91%)	6 (9%)	10	0
All	All	416/384 (108%)	399 (96%)	17 (4%)	35	6

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	85	ARG
1	C	133[A]	VAL
2	D	38[B]	MET
2	B	72	GLU
2	D	67	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	39	ASN
1	A	68	ASN
1	A	70	ASN
1	A	135	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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 27 ligands modelled in this entry, 4 are monoatomic - leaving 23 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	MPD	C	1153	-	7,7,7	1.94	2 (28%)	9,10,10	2.37	4 (44%)
5	MPD	A	1149	-	7,7,7	0.44	0	9,10,10	0.63	0
4	EPE	C	1152[A]	-	15,15,15	1.53	1 (6%)	18,20,20	2.84	10 (55%)
5	MPD	D	1076	-	7,7,7	0.69	0	9,10,10	0.83	0
5	MPD	C	1154	-	7,7,7	0.99	0	9,10,10	1.43	2 (22%)
3	SO4	A	1146	-	4,4,4	0.79	0	6,6,6	0.90	0
4	EPE	A	1148	-	15,15,15	2.10	2 (13%)	18,20,20	4.65	6 (33%)
3	SO4	A	1145	-	4,4,4	0.31	0	6,6,6	0.54	0
3	SO4	B	1075	-	4,4,4	1.06	0	6,6,6	0.99	0
3	SO4	C	1148	-	4,4,4	0.85	0	6,6,6	1.12	0
3	SO4	A	1147	-	4,4,4	0.88	0	6,6,6	1.34	1 (16%)
3	SO4	C	1150	-	4,4,4	1.10	0	6,6,6	0.90	0
3	SO4	B	1074	-	4,4,4	0.60	0	6,6,6	0.39	0
3	SO4	A	1150	-	4,4,4	0.94	0	6,6,6	1.52	1 (16%)
3	SO4	C	1149	-	4,4,4	1.63	1 (25%)	6,6,6	1.33	1 (16%)
3	SO4	C	1147	-	4,4,4	1.13	0	6,6,6	0.98	0
5	MPD	C	1155	-	7,7,7	0.96	0	9,10,10	0.82	0
5	MPD	D	1077	-	7,7,7	0.85	0	9,10,10	2.04	3 (33%)
3	SO4	D	1074	-	4,4,4	0.61	0	6,6,6	0.41	0
3	SO4	C	1151	-	4,4,4	1.09	0	6,6,6	0.68	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	EPE	B	1076	-	15,15,15	0.87	0	18,20,20	2.90	9 (50%)
3	SO4	D	1075	-	4,4,4	1.01	0	6,6,6	0.87	0
4	EPE	C	1152[B]	-	15,15,15	1.39	2 (13%)	18,20,20	3.45	15 (83%)

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
5	MPD	C	1153	-	-	2/5/5/5	-
5	MPD	A	1149	-	-	0/5/5/5	-
5	MPD	C	1155	-	-	0/5/5/5	-
4	EPE	A	1148	-	-	3/9/19/19	0/1/1/1
5	MPD	D	1077	-	-	1/5/5/5	-
5	MPD	D	1076	-	-	0/5/5/5	-
4	EPE	C	1152[A]	-	-	3/9/19/19	0/1/1/1
5	MPD	C	1154	-	-	0/5/5/5	-
4	EPE	B	1076	-	-	2/9/19/19	0/1/1/1
4	EPE	C	1152[B]	-	-	5/9/19/19	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1148	EPE	C10-S	-7.01	1.67	1.77
4	C	1152[A]	EPE	C10-S	-5.08	1.70	1.77
4	C	1152[B]	EPE	C10-S	-4.54	1.71	1.77
5	C	1153	MPD	O2-C2	-3.96	1.34	1.44
5	C	1153	MPD	CM-C2	-2.77	1.43	1.52

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1148	EPE	O3S-S-O1S	-9.22	88.73	111.27
4	A	1148	EPE	O1S-S-C10	9.22	118.01	106.92
4	A	1148	EPE	O3S-S-C10	-9.13	91.01	105.77
4	A	1148	EPE	O3S-S-O2S	-8.27	91.07	111.27
4	A	1148	EPE	O2S-S-C10	6.90	115.22	106.92

There are no chirality outliers.



5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	1153	MPD	C2-C3-C4-O4
4	C	1152[A]	EPE	C10-C9-N1-C2
4	B	1076	EPE	C10-C9-N1-C6
4	C	1152[B]	EPE	N4-C7-C8-O8
4	B	1076	EPE	N4-C7-C8-O8

There are no ring outliers.

13 monomers are involved in 75 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	1153	MPD	10	0
5	A	1149	MPD	8	0
4	C	1152[A]	EPE	2	0
5	C	1154	MPD	7	0
4	A	1148	EPE	21	0
3	B	1075	SO4	1	0
3	A	1150	SO4	1	0
3	C	1149	SO4	1	0
5	C	1155	MPD	3	0
5	D	1077	MPD	12	0
3	C	1151	SO4	3	0
4	B	1076	EPE	1	0
4	C	1152[B]	EPE	5	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, 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	144/151 (95%)	-0.24	1 (0%) 87 90	9, 14, 26, 39	0
1	C	146/151 (96%)	-0.28	1 (0%) 87 90	9, 12, 22, 58	0
2	B	73/75 (97%)	-0.14	2 (2%) 54 59	10, 16, 36, 71	0
2	D	73/75 (97%)	-0.20	3 (4%) 37 41	10, 17, 33, 63	0
All	All	436/452 (96%)	-0.23	7 (1%) 72 77	9, 14, 28, 71	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	144	LEU	11.6
2	B	73	LYS	5.4
2	B	33	PRO	4.0
2	D	73	LYS	3.6
2	D	2	PHE	2.9

### 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. 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
3	SO4	C	1151	5/5	0.69	0.52	30,30,32,33	0
4	EPE	A	1148	15/15	0.78	0.40	21,41,118,119	0
5	MPD	D	1077	8/8	0.79	0.28	17,34,42,46	0
3	SO4	B	1075	5/5	0.81	0.30	30,36,64,66	0
5	MPD	C	1154	8/8	0.81	0.20	13,20,26,31	0
5	MPD	A	1149	8/8	0.83	0.17	25,43,50,54	0
3	SO4	C	1149	5/5	0.86	0.24	16,33,37,42	0
3	SO4	C	1148	5/5	0.89	0.23	33,36,46,54	0
5	MPD	D	1076	8/8	0.92	0.10	15,18,22,30	0
3	SO4	A	1150	5/5	0.92	0.23	17,31,43,48	0
4	EPE	C	1152[A]	15/15	0.92	0.17	5,15,27,29	15
4	EPE	C	1152[B]	15/15	0.92	0.17	8,26,38,39	15
3	SO4	A	1147	5/5	0.93	0.20	29,36,38,41	0
3	SO4	C	1150	5/5	0.93	0.22	31,38,41,51	0
5	MPD	C	1153	8/8	0.93	0.15	10,18,28,35	0
4	EPE	B	1076	15/15	0.94	0.17	23,44,62,66	0
3	SO4	C	1147	5/5	0.94	0.25	24,25,37,43	0
3	SO4	D	1075	5/5	0.95	0.14	26,30,36,49	0
5	MPD	C	1155	8/8	0.96	0.07	10,13,15,25	0
3	SO4	A	1146	5/5	0.96	0.19	32,37,40,44	0
3	SO4	B	1074	5/5	0.98	0.16	29,37,41,46	0
6	CA	D	200	1/1	0.99	0.09	14,14,14,14	0
3	SO4	A	1145	5/5	0.99	0.06	17,17,19,28	0
6	CA	D	201	1/1	0.99	0.11	14,14,14,14	0
3	SO4	D	1074	5/5	0.99	0.05	13,13,16,20	0
6	CA	B	201	1/1	1.00	0.10	14,14,14,14	0
6	CA	B	200	1/1	1.00	0.09	13,13,13,13	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.