



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 09:23 pm BST

PDB ID : 4XB6  
Title : Structure of the E. coli C-P lyase core complex  
Authors : Brodersen, D.E.  
Deposited on : 2014-12-16  
Resolution : 1.70 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 : 2.11  
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

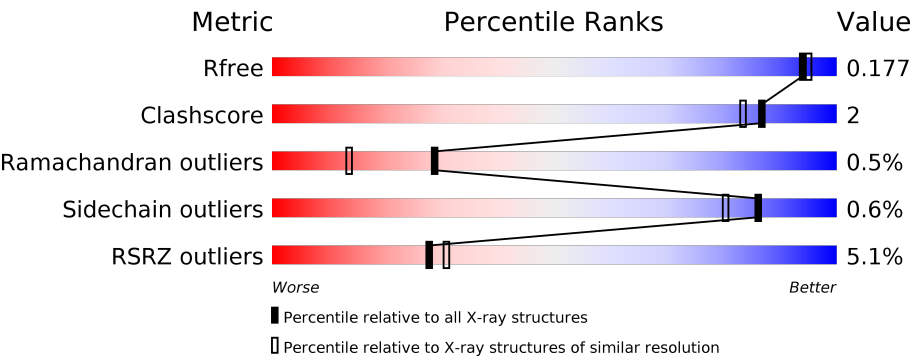
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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	150	<div><div>9%</div><div>85%9%</div></div>
1	E	150	<div><div>14%</div><div>93%5%</div></div>
2	B	194	<div><div>2%</div><div>92%7%</div></div>
2	F	194	<div><div>3%</div><div>94%6%</div></div>
3	C	354	<div><div>2%</div><div>96%</div></div>
3	G	354	<div><div>4%</div><div>95%5%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	281	<div><div></div><div>4%</div><div></div><div>93%</div><div></div><div>5%</div><div></div></div>
4	H	281	<div><div></div><div>8%</div><div></div><div>96%</div><div></div><div></div><div></div></div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 31865 atoms, of which 14894 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 Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnG.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	144	Total	C	H	N	O	S	0	1	0
			2233	692	1109	216	209	7			
1	E	149	Total	C	H	N	O	S	0	1	0
			2286	709	1130	223	218	6			

- Molecule 2 is a protein called Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnH.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	194	Total	C	H	N	O	S	0	3	0
			3027	947	1528	261	280	11			
2	F	193	Total	C	H	N	O	S	0	1	0
			2977	935	1496	258	278	10			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	152	ARG	GLN	engineered mutation	UNP P16686
F	152	ARG	GLN	engineered mutation	UNP P16686

- Molecule 3 is a protein called Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnI.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	353	Total	C	H	N	O	S	0	1	0
			5429	1717	2690	481	530	11			
3	G	353	Total	C	H	N	O	S	0	1	0
			5414	1714	2680	478	531	11			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	322	VAL	ALA	engineered mutation	UNP P16687
G	322	VAL	ALA	engineered mutation	UNP P16687

- Molecule 4 is a protein called Alpha-D-ribose 1-methylphosphonate 5-phosphate C-P lyase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	D	277	Total	C	H	N	O	S	0	2	0
			4336	1399	2124	377	421	15			
4	H	277	Total	C	H	N	O	S	0	2	0
			4347	1398	2137	376	421	15			

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



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

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	H	1	Total 1	Zn 1	0	0
6	G	1	Total 1	Zn 1	0	0
6	D	1	Total 1	Zn 1	0	0
6	C	1	Total 1	Zn 1	0	0


- Molecule 7 is water.

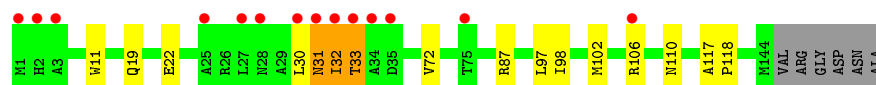
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	132	Total 132	O 132	0	0
7	B	162	Total 162	O 162	0	0
7	C	350	Total 350	O 350	0	0
7	D	259	Total 259	O 259	0	0
7	E	152	Total 152	O 152	0	0
7	F	135	Total 135	O 135	0	0
7	G	366	Total 366	O 366	0	0
7	H	236	Total 236	O 236	0	0

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

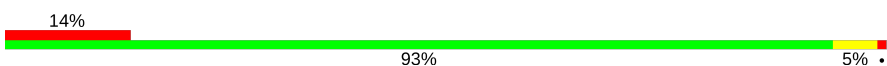
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

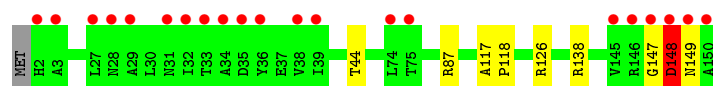
- Molecule 1: Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnG

Chain A: 



- Molecule 1: Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnG

Chain E: 



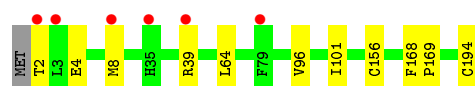
- Molecule 2: Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnH

Chain B: 



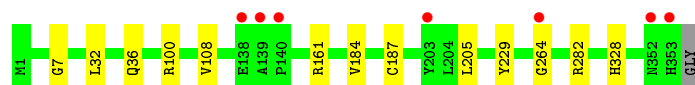
- Molecule 2: Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnH

Chain F: 



- Molecule 3: Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnI

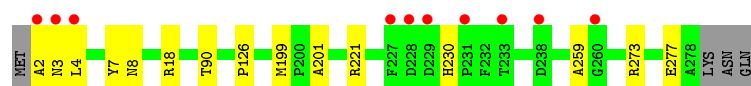
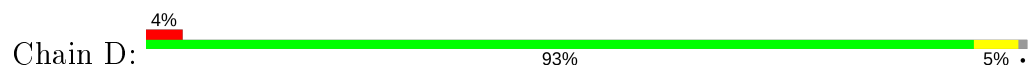
Chain C: 



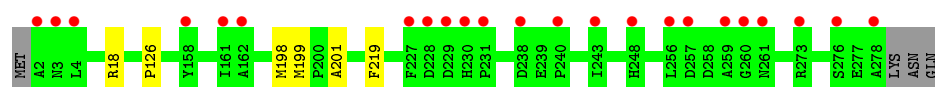
- Molecule 3: Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnI



- Molecule 4: Alpha-D-ribose 1-methylphosphonate 5-phosphate C-P lyase



- Molecule 4: Alpha-D-ribose 1-methylphosphonate 5-phosphate C-P lyase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	95.51Å 133.71Å 176.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	58.36 – 1.70 58.36 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.6 (58.36-1.70) 94.9 (58.36-1.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.15 (at 1.70Å)	Xtriage
Refinement program	PHENIX dev_1593	Depositor
R, $R_{free}$	0.148 , 0.176 0.151 , 0.177	Depositor DCC
$R_{free}$ test set	12423 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.6	Xtriage
Anisotropy	0.136	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 48.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	31865	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, CME, 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	0.37	0/1144	0.57	1/1546 (0.1%)
1	E	0.37	0/1176	0.56	0/1591
2	B	0.36	0/1515	0.61	0/2067
2	F	0.34	0/1490	0.53	0/2034
3	C	0.41	0/2796	0.60	0/3791
3	G	0.43	0/2791	0.62	0/3785
4	D	0.41	0/2271	0.61	0/3087
4	H	0.38	0/2269	0.58	0/3084
All	All	0.39	0/15452	0.59	1/20985 (0.0%)

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
2	B	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	32	ILE	C-N-CA	7.57	140.61	121.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	34	LEU	Peptide

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
2	B	35	HIS	Peptide

## 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	1124	1109	1111	9	0
1	E	1156	1130	1132	7	0
2	B	1499	1528	1539	10	0
2	F	1481	1496	1515	5	0
3	C	2739	2690	2690	13	0
3	G	2734	2680	2682	16	0
4	D	2212	2124	2153	8	0
4	H	2210	2137	2152	3	0
5	C	5	0	0	0	0
5	D	5	0	0	0	0
5	G	5	0	0	0	0
5	H	5	0	0	0	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
6	G	1	0	0	0	0
6	H	1	0	0	0	0
7	A	132	0	0	3	0
7	B	162	0	0	0	1
7	C	350	0	0	3	0
7	D	259	0	0	3	1
7	E	152	0	0	2	0
7	F	135	0	0	1	0
7	G	366	0	0	4	0
7	H	236	0	0	1	0
All	All	16971	14894	14974	63	1

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 2.

All (63) 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)
4:D:221:ARG:NH1	7:D:401:HOH:O	2.08	0.86
1:E:87:ARG:NH2	3:G:58:GLU:OE1	2.09	0.86
3:G:6:LYS:NZ	7:G:782:HOH:O	2.18	0.77
3:C:187:CYS:SG	7:C:683:HOH:O	2.51	0.67
2:F:194:CYS:SG	7:F:314:HOH:O	2.53	0.66
4:D:8[A]:ASN:OD1	7:D:617:HOH:O	2.15	0.65
4:H:18:ARG:NH2	7:H:519:HOH:O	2.32	0.62
3:G:351:GLN:O	3:G:353:HIS:N	2.33	0.62
4:D:18:ARG:NH2	7:D:611:HOH:O	2.34	0.61
2:B:1:MET:O	2:B:3:LEU:N	2.34	0.61
4:D:2:ALA:HB3	4:D:90:THR:O	2.02	0.60
1:A:87[A]:ARG:NH1	7:A:282:HOH:O	2.29	0.57
4:D:273:ARG:NH1	4:D:277:GLU:OE2	2.37	0.56
4:D:3:ASN:ND2	4:D:7:TYR:O	2.40	0.54
2:B:35:HIS:HB3	2:B:46:ILE:HG21	1.89	0.53
3:G:265:GLU:O	7:G:526:HOH:O	2.19	0.53
4:D:4:LEU:HD12	4:D:4:LEU:O	2.10	0.52
3:C:32:LEU:HD22	3:G:32:LEU:HD22	1.91	0.52
2:F:4:GLU:OE1	2:F:39:ARG:NH1	2.44	0.51
2:F:101:ILE:O	2:F:156:CME:HE3	2.11	0.51
4:H:199:MET:HG2	4:H:201:ALA:H	1.76	0.51
3:C:264:GLY:O	7:C:834:HOH:O	2.20	0.51
3:C:32:LEU:HD22	3:G:32:LEU:CD2	2.41	0.50
1:E:44:THR:HG21	1:E:87:ARG:HG2	1.93	0.50
3:C:32:LEU:CD2	3:G:32:LEU:HD22	2.41	0.50
1:A:19:GLN:O	1:A:22:GLU:HG2	2.12	0.49
3:G:262:PRO:O	7:G:526:HOH:O	2.18	0.49
1:A:106:ARG:NH2	7:A:310:HOH:O	2.45	0.48
1:A:72:VAL:HG21	1:A:97:LEU:HA	1.94	0.48
1:E:117:ALA:HB3	1:E:118:PRO:HD3	1.96	0.48
2:B:35:HIS:ND1	2:B:180:GLU:OE1	2.47	0.47
3:C:100[A]:ARG:HG2	3:C:229:TYR:HB2	1.95	0.47
2:B:135:MET:HG3	2:B:148[A]:MET:HE2	1.96	0.47
1:E:126:ARG:NH1	7:E:227:HOH:O	2.41	0.47
1:A:117:ALA:HB3	1:A:118:PRO:HD3	1.97	0.46
1:A:98:ILE:O	1:A:102:MET:HG2	2.15	0.46
1:A:30:LEU:O	1:A:31:ASN:ND2	2.49	0.46
2:B:135:MET:HG3	2:B:148[A]:MET:CE	2.46	0.46
3:C:161:ARG:NH2	7:C:738:HOH:O	2.39	0.45
2:B:168:PHE:CG	2:B:169:PRO:HA	2.51	0.45
2:B:47:ALA:HB3	2:B:177:THR:HB	1.98	0.45
3:G:349:GLN:NE2	7:G:866:HOH:O	2.49	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:110:ASN:ND2	7:A:287:HOH:O	2.43	0.44
3:C:205:LEU:HD11	3:C:282:ARG:CD	2.47	0.44
1:E:138:ARG:NH1	3:G:114:GLY:O	2.44	0.44
2:F:64:LEU:HD23	2:F:96:VAL:HB	2.00	0.44
2:F:168:PHE:CG	2:F:169:PRO:HA	2.54	0.43
2:B:64:LEU:HD23	2:B:96:VAL:HB	2.00	0.43
4:D:199:MET:HG2	4:D:201:ALA:H	1.83	0.43
2:B:168:PHE:CD1	2:B:169:PRO:HA	2.54	0.42
3:C:108:VAL:HB	3:G:337:PHE:CD2	2.55	0.42
1:A:32:ILE:HG22	1:A:33:THR:N	2.35	0.41
3:C:36:GLN:HB3	3:G:32:LEU:HD21	2.02	0.41
1:E:147:GLY:O	1:E:148:ASP:HB2	2.21	0.41
3:C:184:VAL:HG23	3:C:184:VAL:O	2.21	0.41
3:G:146:ASP:N	3:G:146:ASP:OD1	2.54	0.41
4:H:198:MET:HA	4:H:219:PHE:HA	2.03	0.41
3:C:32:LEU:HD21	3:G:36:GLN:HB3	2.03	0.40
1:E:149:ASN:HA	7:E:201:HOH:O	2.19	0.40
3:G:319:ASN:HA	3:G:322:VAL:HB	2.03	0.40
3:G:351:GLN:O	3:G:352:ASN:C	2.60	0.40
3:C:100[A]:ARG:CG	3:C:229:TYR:HB2	2.52	0.40

All (1) 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:B:226:HOH:O	7:D:419:HOH:O[4_555]	2.12	0.08

## 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	143/150 (95%)	138 (96%)	3 (2%)	2 (1%)	<b>11</b> <b>2</b>

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	148/150 (99%)	144 (97%)	3 (2%)	1 (1%)	22	8
2	B	193/194 (100%)	185 (96%)	7 (4%)	1 (0%)	29	13
2	F	190/194 (98%)	186 (98%)	4 (2%)	0	100	100
3	C	352/354 (99%)	349 (99%)	2 (1%)	1 (0%)	41	24
3	G	352/354 (99%)	348 (99%)	2 (1%)	2 (1%)	25	11
4	D	277/281 (99%)	267 (96%)	8 (3%)	2 (1%)	22	8
4	H	277/281 (99%)	268 (97%)	8 (3%)	1 (0%)	34	18
All	All	1932/1958 (99%)	1885 (98%)	37 (2%)	10 (0%)	29	13

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	G	352	ASN
2	B	2	THR
1	A	31	ASN
4	D	126	PRO
4	D	259	ALA
3	C	7	GLY
1	E	148	ASP
3	G	7	GLY
4	H	126	PRO
1	A	33	THR

### 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	112/115 (97%)	111 (99%)	1 (1%)	78	70
1	E	115/115 (100%)	114 (99%)	1 (1%)	78	70
2	B	165/162 (102%)	164 (99%)	1 (1%)	86	80
2	F	162/162 (100%)	159 (98%)	3 (2%)	57	41
3	C	287/286 (100%)	286 (100%)	1 (0%)	92	89

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	G	287/286 (100%)	285 (99%)	2 (1%)	84	77
4	D	242/244 (99%)	241 (100%)	1 (0%)	91	87
4	H	242/244 (99%)	242 (100%)	0	100	100
All	All	1612/1614 (100%)	1602 (99%)	10 (1%)	86	80

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11	TRP
2	B	35	HIS
3	C	328	HIS
4	D	230	HIS
1	E	148	ASP
2	F	2	THR
2	F	8[A]	MET
2	F	8[B]	MET
3	G	113	PRO
3	G	328	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	149	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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	CME	F	156	2	8,9,10	1.00	0	5,9,11	1.18	1 (20%)
2	CME	B	156	2	8,9,10	0.99	0	5,9,11	0.92	0
2	CME	B	178	2	8,9,10	1.03	0	5,9,11	0.96	0
2	CME	F	178	2	8,9,10	0.98	0	5,9,11	0.65	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	CME	F	156	2	-	1/5/8/10	-
2	CME	B	156	2	-	1/5/8/10	-
2	CME	B	178	2	-	0/5/8/10	-
2	CME	F	178	2	-	3/5/8/10	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	156	CME	CB-SG-SD	2.11	109.28	103.82

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	178	CME	SD-CE-CZ-OH
2	F	156	CME	CZ-CE-SD-SG
2	B	156	CME	CZ-CE-SD-SG
2	F	178	CME	CZ-CE-SD-SG
2	F	178	CME	CE-SD-SG-CB

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	156	CME	1	0



## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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	D	301	-	4,4,4	0.46	0	6,6,6	0.34	0
5	SO4	C	401	-	4,4,4	0.17	0	6,6,6	0.07	0
5	SO4	G	401	-	4,4,4	0.20	0	6,6,6	0.24	0
5	SO4	H	301	-	4,4,4	0.20	0	6,6,6	0.30	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.

No monomer is involved in short contacts.

## 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	144/150 (96%)	0.14	14 (9%) 7 8	20, 29, 54, 88	0
1	E	149/150 (99%)	0.57	21 (14%) 2 3	16, 31, 66, 98	0
2	B	192/194 (98%)	-0.21	4 (2%) 63 67	19, 29, 50, 78	0
2	F	191/194 (98%)	-0.19	6 (3%) 49 53	17, 36, 56, 71	0
3	C	353/354 (99%)	-0.01	7 (1%) 65 69	14, 22, 52, 71	0
3	G	353/354 (99%)	-0.02	13 (3%) 41 46	13, 21, 49, 86	0
4	D	277/281 (98%)	-0.24	10 (3%) 42 47	15, 22, 59, 77	0
4	H	277/281 (98%)	0.03	23 (8%) 11 13	16, 26, 60, 86	0
All	All	1936/1958 (98%)	-0.02	98 (5%) 28 31	13, 26, 55, 98	0

All (98) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	150	ALA	8.7
3	G	1	MET	7.6
1	E	149	ASN	7.4
4	D	4	LEU	7.0
1	E	148	ASP	6.9
3	G	353	HIS	6.4
1	A	32	ILE	6.3
4	H	2	ALA	6.2
3	G	139	ALA	6.0
4	H	4	LEU	5.7
4	H	259	ALA	5.6
3	G	352	ASN	5.4
4	H	158	TYR	5.2
4	H	227	PHE	5.1
1	A	33	THR	5.1
4	D	227	PHE	4.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	33	THR	4.8
1	A	30	LEU	4.6
1	E	34	ALA	4.6
4	D	228	ASP	4.5
2	B	1	MET	4.3
1	E	146	ARG	4.2
3	G	140	PRO	4.1
1	E	147	GLY	4.1
3	C	139	ALA	4.1
4	H	230	HIS	4.0
1	E	32	ILE	4.0
1	E	75	THR	4.0
1	E	35	ASP	3.9
1	E	39	ILE	3.7
1	E	2	HIS	3.7
3	C	353	HIS	3.7
1	E	28	ASN	3.7
3	C	352	ASN	3.6
3	C	264	GLY	3.6
1	E	31	ASN	3.5
2	F	8[A]	MET	3.5
1	A	31	ASN	3.4
1	A	34	ALA	3.4
4	H	228	ASP	3.3
1	E	36	TYR	3.3
2	B	2	THR	3.3
1	A	28	ASN	3.3
3	C	140	PRO	3.3
4	D	238	ASP	3.3
2	B	194	CYS	3.2
1	E	145	VAL	3.2
4	H	229	ASP	3.0
4	H	278	ALA	3.0
4	D	2	ALA	3.0
4	D	229	ASP	3.0
1	E	38	VAL	2.9
4	H	273	ARG	2.9
4	H	276	SER	2.9
4	H	248	HIS	2.9
2	F	35	HIS	2.8
4	H	260	GLY	2.8
2	B	35	HIS	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	D	3	ASN	2.8
2	F	2	THR	2.8
1	A	25	ALA	2.8
4	D	233	THR	2.6
4	H	162	ALA	2.6
4	H	256	LEU	2.6
4	H	231	PRO	2.6
3	G	141	THR	2.6
4	H	261	ASN	2.5
4	D	231	PRO	2.5
1	E	27	LEU	2.5
2	F	79	PHE	2.5
1	A	106	ARG	2.5
4	H	238	ASP	2.5
1	E	74	LEU	2.5
1	A	27	LEU	2.4
3	G	138	GLU	2.4
4	D	260	GLY	2.4
1	A	2	HIS	2.4
1	A	3	ALA	2.3
3	G	147	SER	2.3
2	F	39	ARG	2.3
4	H	243	ILE	2.3
3	G	349	GLN	2.3
2	F	3	LEU	2.2
3	C	138	GLU	2.2
3	G	2	TYR	2.2
4	H	161	ILE	2.2
3	C	203	TYR	2.2
1	E	29	ALA	2.2
3	G	204	LEU	2.1
3	G	263	PRO	2.1
4	H	240	PRO	2.1
3	G	203	TYR	2.1
4	H	3	ASN	2.1
1	E	3	ALA	2.1
1	A	35	ASP	2.1
1	A	75	THR	2.1
1	A	1	MET	2.0
4	H	257	ASP	2.0

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

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
2	CME	F	156	10/11	0.89	0.14	31,51,113,113	0
2	CME	B	156	10/11	0.92	0.11	32,48,115,115	0
2	CME	F	178	10/11	0.94	0.10	37,42,77,104	0
2	CME	B	178	10/11	0.96	0.09	24,36,84,106	0

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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
5	SO4	C	401	5/5	0.85	0.21	72,79,83,84	1
5	SO4	G	401	5/5	0.94	0.10	37,45,49,62	0
6	ZN	G	402	1/1	0.96	0.10	21,21,21,21	1
6	ZN	H	302	1/1	0.98	0.05	32,32,32,32	0
5	SO4	D	301	5/5	0.99	0.10	14,15,16,18	0
6	ZN	C	402	1/1	0.99	0.09	20,20,20,20	1
6	ZN	D	302	1/1	0.99	0.07	28,28,28,28	0
5	SO4	H	301	5/5	0.99	0.12	14,15,17,21	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.