



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 15, 2020 – 09:24 am BST

PDB ID : 4ZV9  
Title : 2.00 Angstrom resolution crystal structure of an uncharacterized protein from Escherichia coli O157:H7 str. Sakai  
Authors : Halavaty, A.S.; Wawrzak, Z.; Filippova, E.V.; Kiryukhina, O.; Endres, M.; Joachimiak, A.; Anderson, W.F.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2015-05-18  
Resolution : 2.00 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

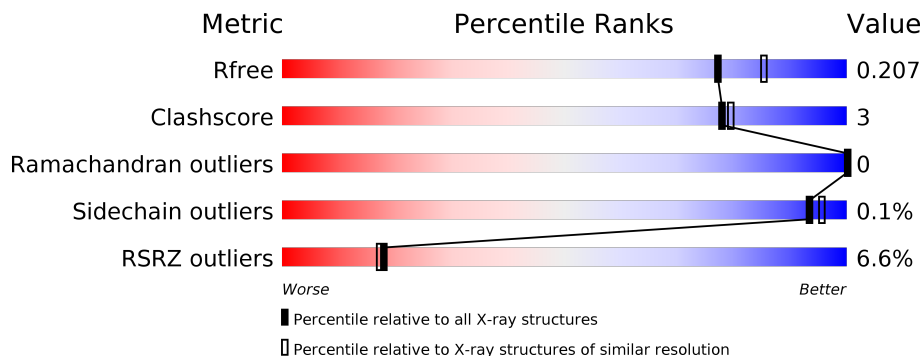
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	238	
1	B	238	
1	C	238	
1	D	238	
1	E	238	
1	F	238	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 12803 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 Uncharacterized protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	238	1910	1218	326	358	3	5	0	6	0
1	B	238	1926	1228	336	355	3	4	0	7	0
1	C	238	1856	1187	315	348	3	3	0	0	0
1	D	238	1898	1211	321	358	3	5	0	5	0
1	E	238	1900	1212	322	356	5	5	0	5	0
1	F	238	1892	1207	325	353	3	4	0	4	0

There are 18 discrepancies between the modelled and reference sequences:

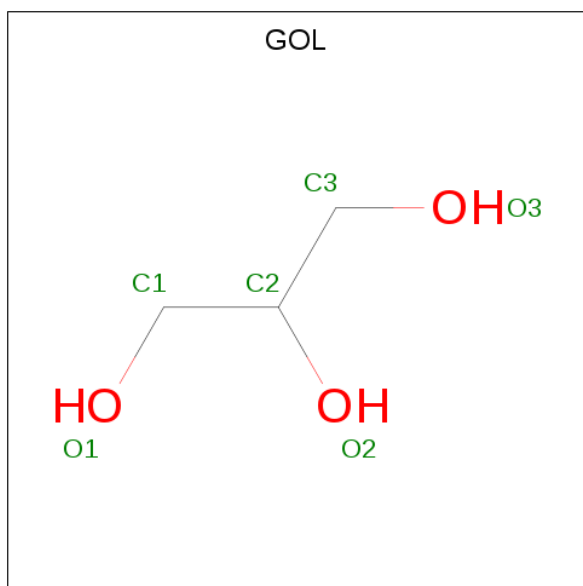
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q7AAT4
A	-1	ASN	-	expression tag	UNP Q7AAT4
A	0	ALA	-	expression tag	UNP Q7AAT4
B	57	SER	-	expression tag	UNP Q7AAT4
B	58	ASN	-	expression tag	UNP Q7AAT4
B	59	ALA	-	expression tag	UNP Q7AAT4
C	57	SER	-	expression tag	UNP Q7AAT4
C	58	ASN	-	expression tag	UNP Q7AAT4
C	59	ALA	-	expression tag	UNP Q7AAT4
D	57	SER	-	expression tag	UNP Q7AAT4
D	58	ASN	-	expression tag	UNP Q7AAT4
D	59	ALA	-	expression tag	UNP Q7AAT4
E	57	SER	-	expression tag	UNP Q7AAT4
E	58	ASN	-	expression tag	UNP Q7AAT4
E	59	ALA	-	expression tag	UNP Q7AAT4
F	57	SER	-	expression tag	UNP Q7AAT4
F	58	ASN	-	expression tag	UNP Q7AAT4

*Continued on next page...*

Continued from previous page...

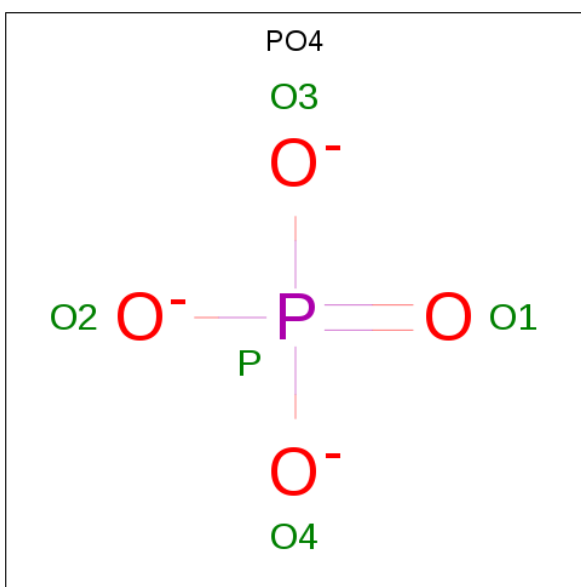
Chain	Residue	Modelled	Actual	Comment	Reference
F	59	ALA	-	expression tag	UNP Q7AAT4

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



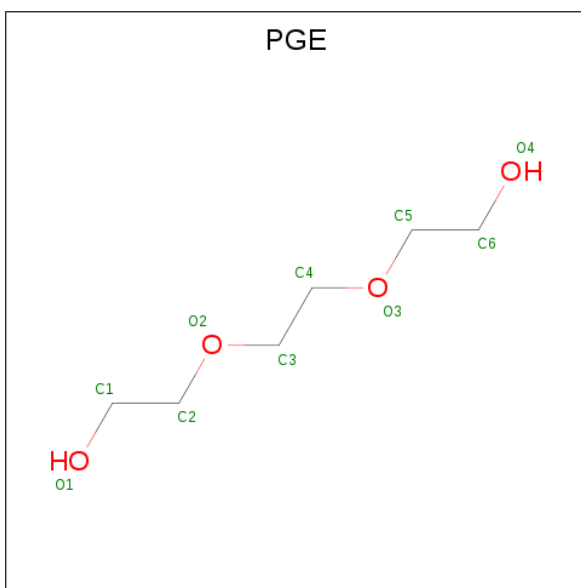
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	6	3	3	0	0
2	A	1	6	3	3	0	0
2	A	1	6	3	3	0	0
2	B	1	6	3	3	0	0
2	D	1	6	3	3	0	0
2	E	1	6	3	3	0	0
2	E	1	6	3	3	0	0
2	F	1	6	3	3	0	0
2	F	1	6	3	3	0	0
2	F	1	6	3	3	0	0

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



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

- Molecule 4 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	E	1	Total	C	O	0	0
			10	6	4		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	F	1	Total C O 7 4 3	0	0

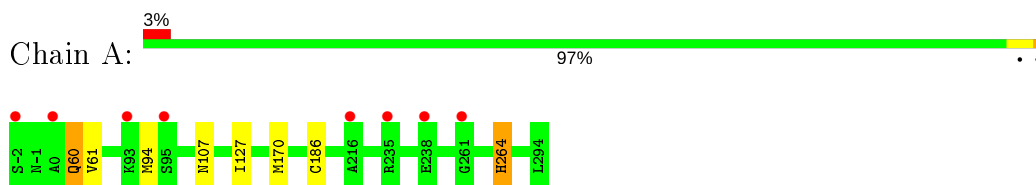
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	183	Total O 184 184	0	4
6	B	333	Total O 341 341	0	16
6	C	50	Total O 50 50	0	0
6	D	142	Total O 147 147	0	5
6	E	295	Total O 299 299	0	6
6	F	309	Total O 318 318	0	12

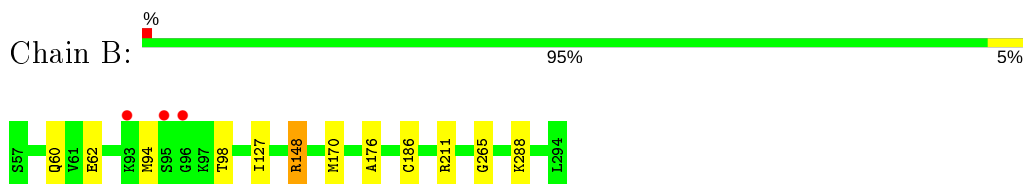
### 3 Residue-property plots [i](#)

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.

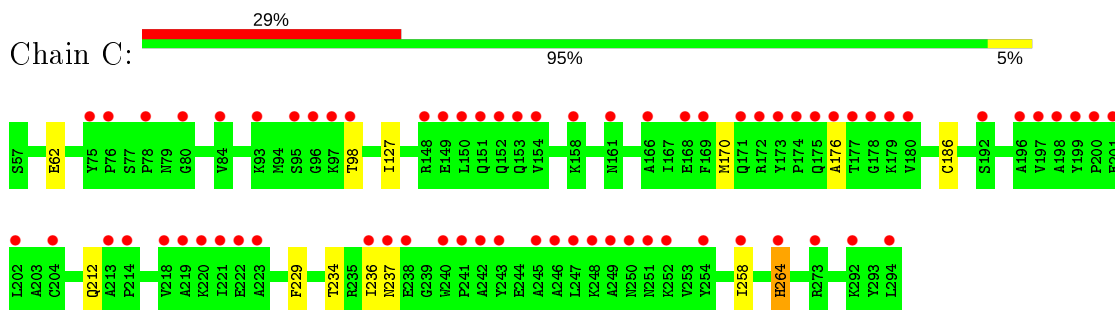
- Molecule 1: Uncharacterized protein



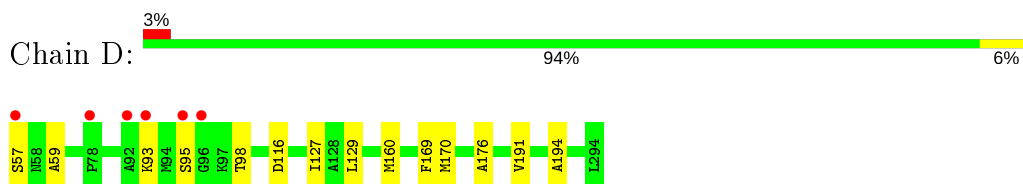
- Molecule 1: Uncharacterized protein



- Molecule 1: Uncharacterized protein



- Molecule 1: Uncharacterized protein



- Molecule 1: Uncharacterized protein





- Molecule 1: Uncharacterized protein





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.61Å 96.34Å 129.62Å 90.00° 98.22° 90.00°	Depositor
Resolution (Å)	38.79 – 2.00 38.79 – 2.00	Depositor EDS
% Data completeness (in resolution range)	92.0 (38.79-2.00) 92.0 (38.79-2.00)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.47 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.168 , 0.199 0.178 , 0.207	Depositor DCC
$R_{free}$ test set	5263 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.4	Xtrriage
Anisotropy	0.225	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 56.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	12803	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.35% 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: PEG, GOL, CME, PGE, PO4

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	0.43	0/1949	0.63	1/2644 (0.0%)
1	B	0.50	0/1965	0.66	2/2662 (0.1%)
1	C	0.38	0/1895	0.65	1/2574 (0.0%)
1	D	0.39	0/1937	0.61	0/2629
1	E	0.49	0/1929	0.64	0/2618
1	F	0.47	0/1931	0.63	0/2620
All	All	0.45	0/11606	0.64	4/15747 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	264	HIS	N-CA-CB	-5.48	100.73	110.60
1	A	264	HIS	N-CA-CB	-5.30	101.07	110.60
1	B	148[A]	ARG	CA-CB-CG	5.03	124.45	113.40
1	B	148[B]	ARG	CA-CB-CG	5.03	124.45	113.40

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	1910	0	1834	7	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1926	0	1863	13	0
1	C	1856	0	1783	8	0
1	D	1898	0	1814	14	0
1	E	1900	0	1820	14	0
1	F	1892	0	1819	13	0
2	A	18	0	24	0	0
2	B	6	0	8	0	0
2	D	6	0	8	0	0
2	E	12	0	16	0	0
2	F	18	0	24	0	0
3	B	5	0	0	0	0
4	E	10	0	14	3	0
5	F	7	0	10	0	0
6	A	184	0	0	2	0
6	B	341	0	0	8	0
6	C	50	0	0	0	0
6	D	147	0	0	2	0
6	E	299	0	0	3	0
6	F	318	0	0	7	0
All	All	12803	0	11037	65	0

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 (65) 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:D:129:LEU:HD22	1:D:170[B]:MSE:HE3	1.42	1.02
1:D:169:PHE:CD2	1:D:170[B]:MSE:HE2	2.01	0.94
1:E:94[A]:MSE:HE2	6:E:508:HOH:O	1.76	0.82
1:D:169:PHE:HD2	1:D:170[B]:MSE:HE2	1.47	0.79
1:D:95:SER:HB3	6:D:496:HOH:O	1.87	0.73
1:D:169:PHE:CE2	1:D:170[B]:MSE:HE2	2.23	0.73
1:C:212:GLN:NE2	1:C:236:ILE:O	2.21	0.73
1:A:94:MSE:HE2	6:A:457:HOH:O	1.88	0.71
1:A:186:CME:SG	1:A:264:HIS:NE2	2.59	0.71
1:F:212:GLN:HG2	6:F:568:HOH:O	1.90	0.70
4:E:303:PGE:H2	1:F:174:PRO:HD3	1.75	0.69
1:B:62:GLU:OE2	1:E:57:SER:HA	1.93	0.68
4:E:303:PGE:C2	1:F:174:PRO:HD3	2.22	0.68
1:D:129:LEU:HD22	1:D:170[B]:MSE:CE	2.22	0.65

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:60:GLN:OE1	1:E:58:ASN:HB3	1.96	0.65
1:E:160[B]:MSE:HE3	6:E:632:HOH:O	1.97	0.65
1:A:60[A]:GLN:HG3	1:A:61:VAL:N	2.11	0.64
1:C:229:PHE:HB2	1:C:258:ILE:HD13	1.81	0.62
1:C:212:GLN:N	1:C:212:GLN:OE1	2.31	0.62
1:F:171:GLN:NE2	6:F:402:HOH:O	2.33	0.60
1:B:211[A]:ARG:NH2	6:B:404:HOH:O	2.34	0.60
1:F:94:MSE:HE2	6:F:466:HOH:O	2.01	0.60
1:B:94:MSE:HE2	6:B:508:HOH:O	2.02	0.59
1:C:186:CME:SG	1:C:264:HIS:NE2	2.61	0.59
1:B:186:CME:HB3	6:B:622[B]:HOH:O	2.02	0.58
1:F:160[B]:MSE:HE3	6:F:675:HOH:O	2.04	0.56
1:E:160[B]:MSE:HE2	1:E:191:VAL:HG22	1.86	0.56
1:F:108:ARG:NH1	6:F:404:HOH:O	2.39	0.55
6:B:643:HOH:O	1:E:57:SER:HB2	2.07	0.55
1:B:288[A]:LYS:HE2	6:B:545:HOH:O	2.07	0.54
4:E:303:PGE:H22	1:F:174:PRO:HD3	1.90	0.54
1:B:60:GLN:OE1	1:E:58:ASN:CB	2.55	0.54
1:D:160[B]:MSE:HE3	1:D:191:VAL:HG22	1.91	0.53
1:A:107:ASN:HB3	1:A:186:CME:HZ2	1.90	0.53
1:C:234:THR:HA	1:C:237:ASN:OD1	2.08	0.53
1:F:97:LYS:NZ	6:F:405:HOH:O	2.41	0.52
1:B:288[A]:LYS:HD2	6:B:660:HOH:O	2.09	0.51
1:D:160[B]:MSE:HE2	1:D:194:ALA:HB3	1.95	0.49
1:E:91:PRO:HG2	1:E:94[B]:MSE:HE2	1.94	0.49
1:A:60[A]:GLN:NE2	6:A:401[A]:HOH:O	2.32	0.48
1:D:93:LYS:NZ	6:D:403:HOH:O	2.38	0.47
1:B:148[B]:ARG:HD2	6:B:650:HOH:O	2.16	0.45
1:E:186[B]:CME:HE3	1:E:186[B]:CME:HB2	1.47	0.45
1:B:186:CME:HZ2	1:B:265:GLY:HA2	1.98	0.45
1:A:127:ILE:HG21	1:A:170[B]:MSE:HE1	1.99	0.45
1:E:127:ILE:HG21	1:E:170:MSE:HE1	1.99	0.45
1:A:170[A]:MSE:HE2	1:A:170[A]:MSE:HA	1.99	0.44
1:D:127:ILE:HG21	1:D:170[A]:MSE:HE1	1.99	0.44
1:B:186:CME:HZ2	1:B:265:GLY:CA	2.48	0.43
1:B:127:ILE:HG21	1:B:170:MSE:HE1	2.00	0.43
1:E:91:PRO:HB2	1:E:94[B]:MSE:HG2	2.01	0.43
6:B:561:HOH:O	1:E:58:ASN:HB2	2.18	0.43
1:F:74:THR:HG23	6:F:631:HOH:O	2.19	0.43
1:E:186[A]:CME:HB2	6:E:549[A]:HOH:O	2.18	0.42
1:F:127:ILE:HG21	1:F:170:MSE:HE1	2.02	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:98:THR:O	1:B:176:ALA:HA	2.19	0.42
1:C:127:ILE:HG21	1:C:170:MSE:HE1	2.00	0.42
1:F:98:THR:O	1:F:176:ALA:HA	2.19	0.42
1:D:169:PHE:CE2	1:D:170[B]:MSE:CE	3.00	0.41
1:C:62:GLU:OE1	1:D:57:SER:HA	2.21	0.41
1:D:98:THR:O	1:D:176:ALA:HA	2.21	0.41
1:C:98:THR:O	1:C:176:ALA:HA	2.20	0.41
1:F:186:CME:HZ2	1:F:273[B]:ARG:HE	1.85	0.41
1:D:59:ALA:HB1	1:D:116:ASP:OD1	2.21	0.41
1:E:98:THR:O	1:E:176:ALA:HA	2.21	0.41

There are no symmetry-related clashes.

## 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	241/238 (101%)	236 (98%)	5 (2%)	0	100	100
1	B	242/238 (102%)	237 (98%)	5 (2%)	0	100	100
1	C	235/238 (99%)	230 (98%)	5 (2%)	0	100	100
1	D	240/238 (101%)	234 (98%)	6 (2%)	0	100	100
1	E	239/238 (100%)	234 (98%)	5 (2%)	0	100	100
1	F	239/238 (100%)	232 (97%)	7 (3%)	0	100	100
All	All	1436/1428 (101%)	1403 (98%)	33 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	191/182 (105%)	189 (99%)	2 (1%)	76	81
1	B	192/182 (106%)	192 (100%)	0	100	100
1	C	185/182 (102%)	185 (100%)	0	100	100
1	D	190/182 (104%)	190 (100%)	0	100	100
1	E	189/182 (104%)	189 (100%)	0	100	100
1	F	189/182 (104%)	189 (100%)	0	100	100
All	All	1136/1092 (104%)	1134 (100%)	2 (0%)	93	95

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	60[A]	GLN
1	A	60[B]	GLN

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

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

7 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
1	CME	E	186[B]	1	8,9,10	0.84	0	5,9,11	1.32	1 (20%)
1	CME	F	186	1	8,9,10	0.82	0	5,9,11	1.07	0
1	CME	C	186	1	8,9,10	0.71	0	5,9,11	0.93	0
1	CME	A	186	1	8,9,10	0.85	0	5,9,11	1.14	0
1	CME	D	186	1	8,9,10	0.76	0	5,9,11	0.76	0
1	CME	B	186	1	8,9,10	1.64	1 (12%)	5,9,11	2.57	2 (40%)
1	CME	E	186[A]	1	8,9,10	0.92	0	5,9,11	0.90	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
1	CME	E	186[B]	1	-	1/5/8/10	-
1	CME	F	186	1	-	2/5/8/10	-
1	CME	C	186	1	-	3/5/8/10	-
1	CME	A	186	1	-	3/5/8/10	-
1	CME	D	186	1	-	2/5/8/10	-
1	CME	B	186	1	-	4/5/8/10	-
1	CME	E	186[A]	1	-	2/5/8/10	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	186	CME	CB-SG	-4.23	1.68	1.81

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	186	CME	CB-SG-SD	-4.36	92.53	103.82
1	B	186	CME	CA-CB-SG	-3.10	101.31	114.55
1	E	186[B]	CME	CB-SG-SD	-2.81	96.55	103.82

There are no chirality outliers.

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	F	186	CME	CZ-CE-SD-SG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
1	B	186	CME	N-CA-CB-SG
1	B	186	CME	SD-CE-CZ-OH
1	B	186	CME	CE-SD-SG-CB
1	F	186	CME	CE-SD-SG-CB
1	E	186[A]	CME	CE-SD-SG-CB
1	E	186[B]	CME	CE-SD-SG-CB
1	C	186	CME	SD-CE-CZ-OH
1	A	186	CME	SD-CE-CZ-OH
1	C	186	CME	N-CA-CB-SG
1	A	186	CME	N-CA-CB-SG
1	C	186	CME	CZ-CE-SD-SG
1	B	186	CME	CZ-CE-SD-SG
1	E	186[A]	CME	CZ-CE-SD-SG
1	D	186	CME	SD-CE-CZ-OH
1	A	186	CME	CZ-CE-SD-SG
1	D	186	CME	CA-CB-SG-SD

There are no ring outliers.

6 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	E	186[B]	CME	1	0
1	F	186	CME	1	0
1	C	186	CME	1	0
1	A	186	CME	2	0
1	B	186	CME	3	0
1	E	186[A]	CME	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

13 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the



expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	GOL	A	303	-	5,5,5	0.28	0	5,5,5	0.37	0
2	GOL	F	302	-	5,5,5	0.41	0	5,5,5	0.96	0
2	GOL	B	302	-	5,5,5	0.10	0	5,5,5	0.55	0
2	GOL	D	301	-	5,5,5	0.29	0	5,5,5	0.48	0
2	GOL	F	301	-	5,5,5	0.22	0	5,5,5	0.34	0
2	GOL	E	301	-	5,5,5	0.30	0	5,5,5	0.37	0
3	PO4	B	301	-	4,4,4	0.97	0	6,6,6	0.38	0
2	GOL	A	301	-	5,5,5	0.36	0	5,5,5	0.35	0
2	GOL	E	302	-	5,5,5	0.29	0	5,5,5	0.85	0
5	PEG	F	304	-	6,6,6	0.50	0	5,5,5	0.26	0
2	GOL	A	302	-	5,5,5	0.31	0	5,5,5	0.25	0
2	GOL	F	303	-	5,5,5	0.18	0	5,5,5	0.35	0
4	PGE	E	303	-	9,9,9	0.62	0	8,8,8	0.72	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	GOL	A	303	-	-	2/4/4/4	-
2	GOL	F	302	-	-	0/4/4/4	-
2	GOL	B	302	-	-	4/4/4/4	-
2	GOL	D	301	-	-	0/4/4/4	-
2	GOL	F	301	-	-	3/4/4/4	-
2	GOL	E	301	-	-	2/4/4/4	-
2	GOL	A	301	-	-	2/4/4/4	-
2	GOL	E	302	-	-	4/4/4/4	-
5	PEG	F	304	-	-	2/4/4/4	-
2	GOL	A	302	-	-	0/4/4/4	-
2	GOL	F	303	-	-	2/4/4/4	-
4	PGE	E	303	-	-	3/7/7/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (24) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	303	GOL	O1-C1-C2-C3
2	F	301	GOL	C1-C2-C3-O3
2	F	301	GOL	O2-C2-C3-O3
2	A	301	GOL	C1-C2-C3-O3
2	E	302	GOL	O1-C1-C2-O2
2	B	302	GOL	O2-C2-C3-O3
2	E	302	GOL	O2-C2-C3-O3
4	E	303	PGE	O1-C1-C2-O2
2	B	302	GOL	C1-C2-C3-O3
2	E	302	GOL	O1-C1-C2-C3
2	E	302	GOL	C1-C2-C3-O3
2	F	303	GOL	C1-C2-C3-O3
5	F	304	PEG	O2-C3-C4-O4
2	A	301	GOL	O2-C2-C3-O3
5	F	304	PEG	O1-C1-C2-O2
2	F	303	GOL	O2-C2-C3-O3
2	A	303	GOL	O1-C1-C2-O2
2	E	301	GOL	O2-C2-C3-O3
2	B	302	GOL	O1-C1-C2-C3
4	E	303	PGE	O3-C5-C6-O4
2	F	301	GOL	O1-C1-C2-O2
2	B	302	GOL	O1-C1-C2-O2
4	E	303	PGE	O2-C3-C4-O3
2	E	301	GOL	C1-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	303	PGE	3	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	234/238 (98%)	-0.07	8 (3%) 45 44	20, 33, 58, 94	0
1	B	234/238 (98%)	-0.23	3 (1%) 77 76	14, 22, 35, 80	0
1	C	234/238 (98%)	1.41	70 (29%) 0 0	33, 76, 130, 144	0
1	D	234/238 (98%)	-0.09	6 (2%) 56 54	27, 42, 63, 129	0
1	E	234/238 (98%)	-0.36	3 (1%) 77 76	14, 23, 38, 79	0
1	F	234/238 (98%)	-0.48	2 (0%) 84 83	15, 22, 38, 84	0
All	All	1404/1428 (98%)	0.03	92 (6%) 18 17	14, 30, 92, 144	0

All (92) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	251	ASN	7.7
1	C	247	LEU	7.4
1	C	249	ALA	7.4
1	D	95	SER	6.8
1	C	172	ARG	6.2
1	C	238	GLU	6.1
1	C	80	GLY	5.9
1	C	243	TYR	5.6
1	C	221	ILE	5.4
1	C	175	GLN	5.1
1	C	219	ALA	4.7
1	C	93	LYS	4.5
1	C	237	ASN	4.4
1	D	57	SER	4.4
1	B	95	SER	4.3
1	C	153	GLN	4.3
1	C	241	PRO	4.3
1	C	152	GLN	4.2
1	C	218	VAL	4.1

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	C	171	GLN	4.0
1	C	258	ILE	4.0
1	E	95	SER	4.0
1	C	95	SER	3.9
1	C	240	TRP	3.8
1	C	294	LEU	3.8
1	C	78	PRO	3.8
1	C	169	PHE	3.7
1	C	97	LYS	3.7
1	C	178	GLY	3.6
1	C	292	LYS	3.6
1	E	57	SER	3.5
1	C	220	LYS	3.5
1	C	252	LYS	3.5
1	C	174	PRO	3.4
1	C	76	PRO	3.4
1	C	98	THR	3.4
1	C	84	VAL	3.4
1	A	238	GLU	3.3
1	C	197	VAL	3.3
1	C	176	ALA	3.3
1	C	168	GLU	3.2
1	C	246	ALA	3.2
1	C	242	ALA	3.2
1	C	177	THR	3.2
1	C	222	GLU	3.2
1	C	254	TYR	3.2
1	C	192	SER	3.1
1	C	149	GLU	3.1
1	D	96	GLY	3.0
1	C	245	ALA	3.0
1	C	273	ARG	2.9
1	C	96	GLY	2.9
1	C	202	LEU	2.9
1	D	92	ALA	2.9
1	B	96	GLY	2.8
1	C	154	VAL	2.8
1	B	93	LYS	2.8
1	C	250	ASN	2.8
1	C	179	LYS	2.8
1	A	0	ALA	2.7
1	F	95	SER	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	158	LYS	2.7
1	A	-2	SER	2.7
1	F	93	LYS	2.6
1	A	216	ALA	2.6
1	A	93	LYS	2.6
1	C	148	ARG	2.5
1	C	214	PRO	2.5
1	C	199	TYR	2.5
1	C	150	LEU	2.4
1	C	200	PRO	2.4
1	C	173	TYR	2.4
1	C	223	ALA	2.4
1	A	95	SER	2.4
1	A	235	ARG	2.4
1	C	198	ALA	2.3
1	C	201	GLU	2.3
1	E	93	LYS	2.3
1	C	75	TYR	2.3
1	D	93	LYS	2.3
1	C	151	GLN	2.3
1	C	180	VAL	2.2
1	D	78	PRO	2.2
1	C	248	LYS	2.2
1	C	166	ALA	2.1
1	C	204	CYS	2.1
1	C	213	ALA	2.1
1	C	161	ASN	2.1
1	C	264	HIS	2.1
1	C	236	ILE	2.1
1	C	196	ALA	2.0
1	A	261	GLY	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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
1	CME	C	186	10/11	0.74	0.21	71,78,91,92	0
1	CME	E	186[B]	10/11	0.91	0.20	18,27,34,34	10

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
1	CME	E	186[A]	10/11	0.91	0.20	19,32,46,52	10
1	CME	B	186	10/11	0.92	0.16	16,33,46,57	0
1	CME	A	186	10/11	0.92	0.17	29,42,71,79	0
1	CME	D	186	10/11	0.96	0.10	27,38,66,77	0
1	CME	F	186	10/11	0.97	0.10	16,25,65,75	0

### 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(Å <sup>2</sup> )	Q<0.9
4	PGE	E	303	10/10	0.67	0.27	48,60,63,65	0
2	GOL	E	301	6/6	0.72	0.31	62,73,77,86	0
2	GOL	A	301	6/6	0.73	0.22	73,75,78,79	0
2	GOL	F	301	6/6	0.74	0.24	58,63,66,67	0
5	PEG	F	304	7/7	0.79	0.21	52,55,62,63	0
2	GOL	A	303	6/6	0.81	0.26	51,52,56,57	0
3	PO4	B	301	5/5	0.83	0.26	48,49,54,63	5
2	GOL	B	302	6/6	0.85	0.20	39,49,59,64	0
2	GOL	E	302	6/6	0.87	0.18	37,43,48,50	0
2	GOL	D	301	6/6	0.88	0.17	45,51,57,63	0
2	GOL	F	303	6/6	0.92	0.20	40,47,66,72	0
2	GOL	A	302	6/6	0.92	0.17	45,49,50,53	0
2	GOL	F	302	6/6	0.94	0.17	28,35,37,40	0

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

There are no such residues in this entry.