



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 17, 2023 – 02:26 PM EST

PDB ID : 2Q04  
Title : Crystal structure of acetoin utilization protein (ZP\_00540088.1) from *Exiguobacterium sibiricum* 255-15 at 2.33 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2007-05-18  
Resolution : 2.33 Å (reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.31.2  
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) : 2.31.2

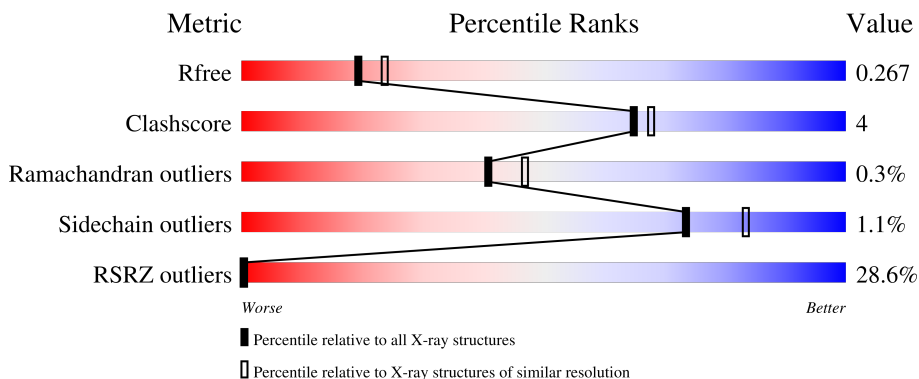
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.33 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	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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	211	
1	B	211	
1	C	211	
1	D	211	
1	E	211	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	211	

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
4	ACY	E	212	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10133 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 Acetoin utilization protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	211	1691	1092	284	307	1	7	0	2	0
1	B	207	1645	1058	282	298	1	6	0	2	0
1	C	208	1608	1030	272	297	1	8	0	2	0
1	D	210	1678	1083	280	306	1	8	0	1	0
1	E	210	1660	1072	277	302	1	8	0	0	0
1	F	204	1560	997	267	288	1	7	0	1	0

There are 54 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP Q41BL0
A	1	MSE	MET	modified residue	UNP Q41BL0
A	120	MSE	MET	modified residue	UNP Q41BL0
A	125	MSE	MET	modified residue	UNP Q41BL0
A	155	MSE	MET	modified residue	UNP Q41BL0
A	158	MSE	MET	modified residue	UNP Q41BL0
A	159	MSE	MET	modified residue	UNP Q41BL0
A	183	MSE	MET	modified residue	UNP Q41BL0
A	208	MSE	MET	modified residue	UNP Q41BL0
B	0	GLY	-	expression tag	UNP Q41BL0
B	1	MSE	MET	modified residue	UNP Q41BL0
B	120	MSE	MET	modified residue	UNP Q41BL0
B	125	MSE	MET	modified residue	UNP Q41BL0
B	155	MSE	MET	modified residue	UNP Q41BL0
B	158	MSE	MET	modified residue	UNP Q41BL0
B	159	MSE	MET	modified residue	UNP Q41BL0
B	183	MSE	MET	modified residue	UNP Q41BL0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	208	MSE	MET	modified residue	UNP Q41BL0
C	0	GLY	-	expression tag	UNP Q41BL0
C	1	MSE	MET	modified residue	UNP Q41BL0
C	120	MSE	MET	modified residue	UNP Q41BL0
C	125	MSE	MET	modified residue	UNP Q41BL0
C	155	MSE	MET	modified residue	UNP Q41BL0
C	158	MSE	MET	modified residue	UNP Q41BL0
C	159	MSE	MET	modified residue	UNP Q41BL0
C	183	MSE	MET	modified residue	UNP Q41BL0
C	208	MSE	MET	modified residue	UNP Q41BL0
D	0	GLY	-	expression tag	UNP Q41BL0
D	1	MSE	MET	modified residue	UNP Q41BL0
D	120	MSE	MET	modified residue	UNP Q41BL0
D	125	MSE	MET	modified residue	UNP Q41BL0
D	155	MSE	MET	modified residue	UNP Q41BL0
D	158	MSE	MET	modified residue	UNP Q41BL0
D	159	MSE	MET	modified residue	UNP Q41BL0
D	183	MSE	MET	modified residue	UNP Q41BL0
D	208	MSE	MET	modified residue	UNP Q41BL0
E	0	GLY	-	expression tag	UNP Q41BL0
E	1	MSE	MET	modified residue	UNP Q41BL0
E	120	MSE	MET	modified residue	UNP Q41BL0
E	125	MSE	MET	modified residue	UNP Q41BL0
E	155	MSE	MET	modified residue	UNP Q41BL0
E	158	MSE	MET	modified residue	UNP Q41BL0
E	159	MSE	MET	modified residue	UNP Q41BL0
E	183	MSE	MET	modified residue	UNP Q41BL0
E	208	MSE	MET	modified residue	UNP Q41BL0
F	0	GLY	-	expression tag	UNP Q41BL0
F	1	MSE	MET	modified residue	UNP Q41BL0
F	120	MSE	MET	modified residue	UNP Q41BL0
F	125	MSE	MET	modified residue	UNP Q41BL0
F	155	MSE	MET	modified residue	UNP Q41BL0
F	158	MSE	MET	modified residue	UNP Q41BL0
F	159	MSE	MET	modified residue	UNP Q41BL0
F	183	MSE	MET	modified residue	UNP Q41BL0
F	208	MSE	MET	modified residue	UNP Q41BL0

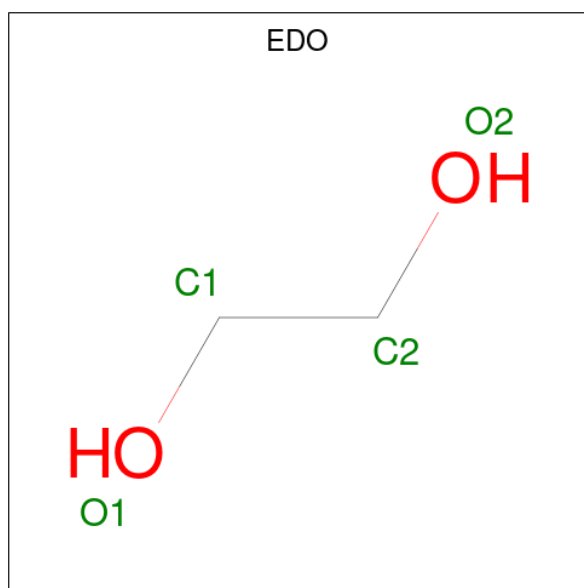
- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

*Continued on next page...*

Continued from previous page...

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

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



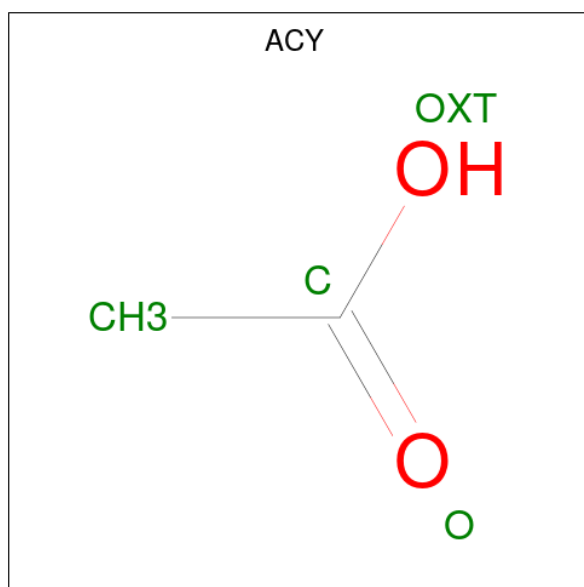
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	C	1	Total C O 4 2 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	D	1	Total C O 4 2 2	0	0
3	D	1	Total C O 4 2 2	0	0
3	F	1	Total C O 4 2 2	0	0

- Molecule 4 is ACETIC ACID (three-letter code: ACY) (formula: C<sub>2</sub>H<sub>4</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	E	1	Total C O 4 2 2	0	0
4	E	1	Total C O 4 2 2	0	0
4	F	1	Total C O 4 2 2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	57	Total O 57 57	0	0
5	B	45	Total O 45 45	0	0

Continued on next page...

*Continued from previous page...*

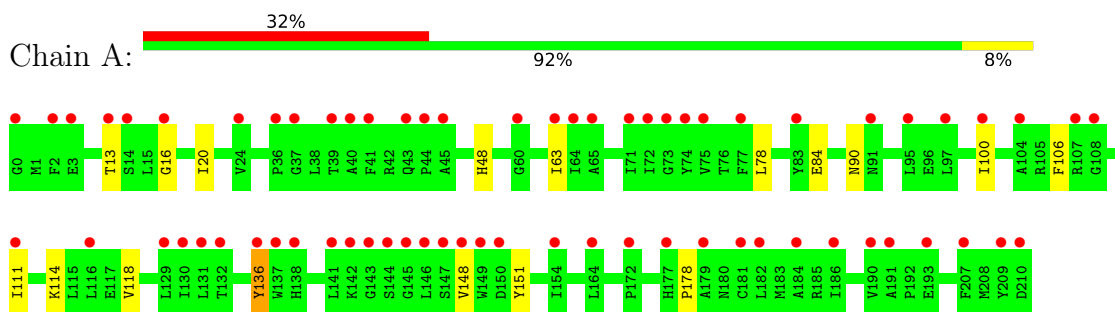
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	C	23	Total 23	O 23	0	0
5	D	57	Total 57	O 57	0	0
5	E	40	Total 40	O 40	0	0
5	F	18	Total 18	O 18	0	0



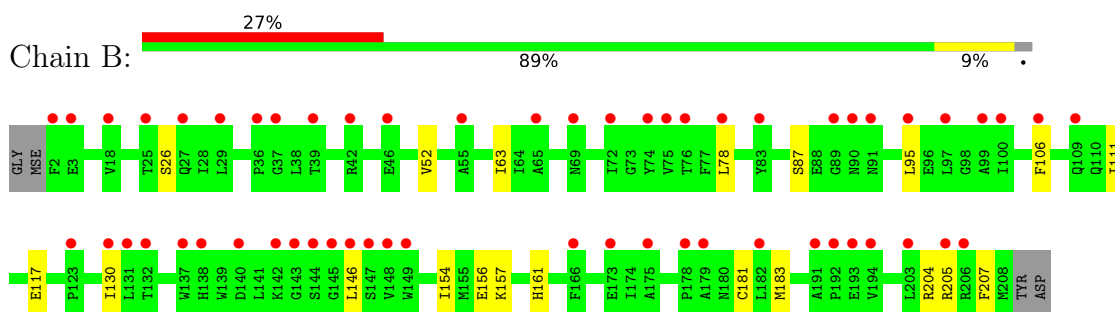
### 3 Residue-property plots [i](#)

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.

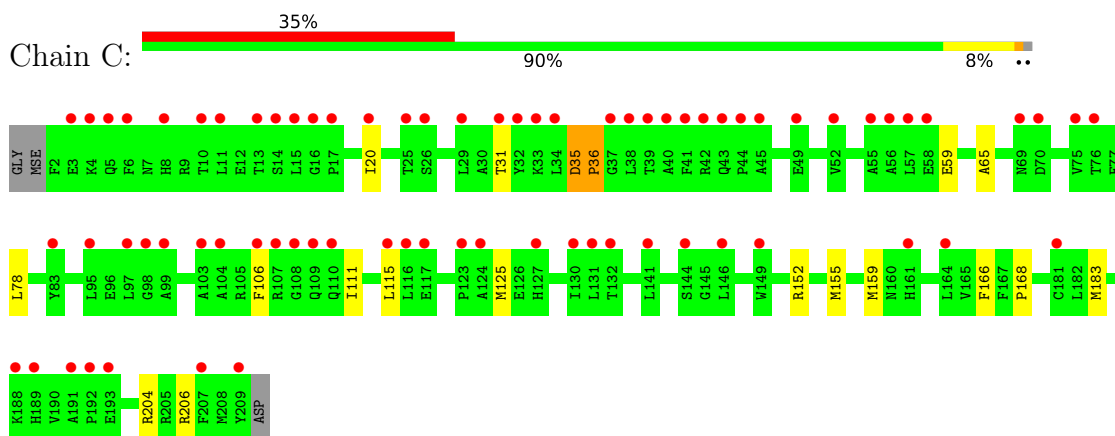
- Molecule 1: Acetoin utilization protein



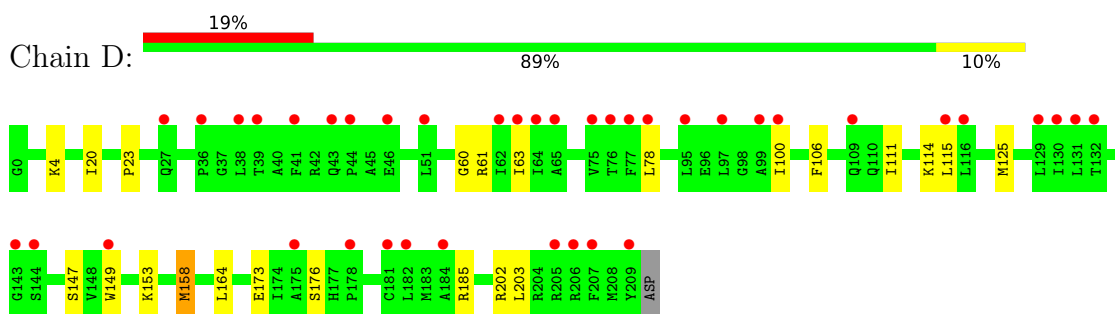
- Molecule 1: Acetoin utilization protein



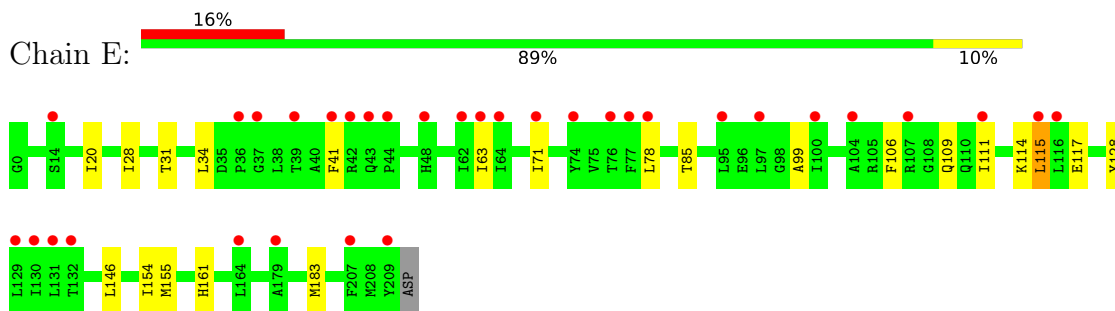
- Molecule 1: Acetoin utilization protein



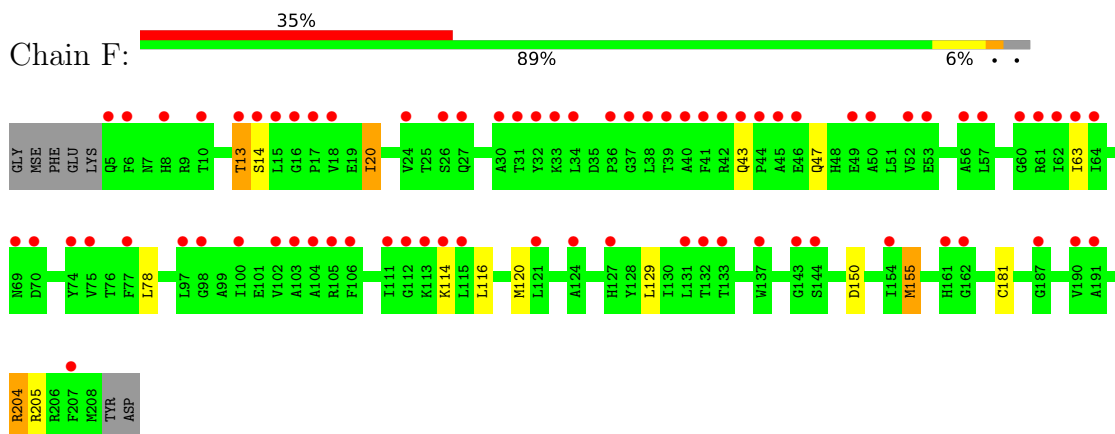
- Molecule 1: Acetoin utilization protein



• Molecule 1: Acetoin utilization protein



• Molecule 1: Acetoin utilization protein



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.87Å 128.58Å 108.43Å 90.00° 108.43° 90.00°	Depositor
Resolution (Å)	29.34 – 2.33 29.33 – 2.33	Depositor EDS
% Data completeness (in resolution range)	93.7 (29.34-2.33) 93.7 (29.33-2.33)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.06 (at 2.34Å)	Xtrriage
Refinement program	REFMAC 5.2.0005, PHENIX	Depositor
R, $R_{free}$	0.229 , 0.260 0.235 , 0.267	Depositor DCC
$R_{free}$ test set	3696 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.1	Xtrriage
Anisotropy	0.091	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 51.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.017 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	10133	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 46.81 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.0849e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: EDO, CA, ACY

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	0.60	0/1734	0.42	0/2344
1	B	0.61	0/1689	0.42	0/2286
1	C	0.53	0/1650	0.42	0/2235
1	D	0.74	1/1718 (0.1%)	0.44	0/2319
1	E	0.65	0/1698	0.44	0/2298
1	F	0.54	0/1598	0.43	0/2168
All	All	0.62	1/10087 (0.0%)	0.43	0/13650

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	158	MSE	SE-CE	5.33	2.26	1.95

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	1691	0	1622	15	0
1	B	1645	0	1586	12	0
1	C	1608	0	1486	14	0
1	D	1678	0	1633	18	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	1660	0	1597	15	0
1	F	1560	0	1440	11	0
2	A	2	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	12	0	18	1	0
3	C	4	0	6	1	0
3	D	8	0	12	0	0
3	F	4	0	6	0	0
4	A	4	0	3	1	0
4	E	8	0	6	2	0
4	F	4	0	3	0	0
5	A	57	0	0	0	0
5	B	45	0	0	0	0
5	C	23	0	0	0	0
5	D	57	0	0	1	0
5	E	40	0	0	1	0
5	F	18	0	0	0	0
All	All	10133	0	9418	79	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 4.

All (79) 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:125:MSE:SE	1:D:125:MSE:CE	2.15	1.44
1:D:158:MSE:CE	1:D:158:MSE:SE	2.26	1.33
1:E:155:MSE:SE	1:E:183:MSE:HE1	2.20	0.92
1:B:157:LYS:NZ	1:F:150:ASP:OD2	2.12	0.82
1:A:178:PRO:HG2	1:D:149:TRP:CH2	2.22	0.75
1:C:65:ALA:HB2	1:C:115:LEU:HD11	1.70	0.72
1:E:146:LEU:HD21	1:E:154:ILE:HD12	1.74	0.69
1:B:181:CYS:SG	1:B:183:MSE:HE3	2.32	0.68
1:E:117:GLU:OE2	1:E:161:HIS:NE2	2.27	0.67
1:E:128:TYR:OH	4:E:212:ACY:H3	1.97	0.65
1:C:159:MSE:HG3	1:C:183:MSE:HE1	1.78	0.64
1:C:125:MSE:HA	1:C:125:MSE:HE2	1.85	0.58

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:106:PHE:HB3	1:B:111:ILE:HD12	1.86	0.57
1:F:155:MSE:HE1	1:F:181:CYS:SG	2.45	0.56
1:C:20:ILE:HG23	1:C:115:LEU:HD12	1.88	0.55
1:F:20:ILE:CD1	1:F:114:LYS:HB3	2.37	0.55
1:B:117:GLU:OE2	1:B:161[B]:HIS:NE2	2.42	0.53
1:A:13:THR:OG1	1:A:16:GLY:N	2.41	0.53
1:D:173:GLU:O	1:D:176:SER:OG	2.26	0.52
1:E:85:THR:HG21	5:E:233:HOH:O	2.09	0.52
1:C:166:PHE:HA	1:C:183:MSE:HG2	1.93	0.51
1:C:35:ASP:CG	1:C:36:PRO:CA	2.81	0.49
1:D:4:LYS:NZ	1:D:60:GLY:O	2.30	0.49
1:E:20:ILE:HD13	1:E:114:LYS:HB3	1.94	0.49
1:C:168:PRO:HB3	1:D:147:SER:HA	1.94	0.49
1:D:106:PHE:HB3	1:D:111:ILE:HD12	1.94	0.48
1:A:84:GLU:O	1:B:204:ARG:NH2	2.47	0.48
1:C:35:ASP:CB	1:C:36:PRO:CA	2.92	0.47
1:D:202:ARG:HG3	1:D:203:LEU:HG	1.97	0.47
1:A:100:ILE:O	1:A:100:ILE:HG23	2.15	0.46
1:B:156:GLU:HA	1:B:183:MSE:HE1	1.97	0.46
1:C:78:LEU:HD23	1:C:78:LEU:N	2.30	0.46
1:F:155:MSE:CE	1:F:181:CYS:SG	3.04	0.46
1:A:136:TYR:CD1	1:A:136:TYR:C	2.90	0.46
1:A:48:HIS:CG	4:A:216:ACY:H3	2.51	0.45
1:D:4:LYS:HB3	1:D:23:PRO:CG	2.46	0.45
1:D:61:ARG:CZ	1:D:125:MSE:HE3	2.47	0.45
1:F:78:LEU:N	1:F:78:LEU:HD23	2.31	0.45
1:F:63:ILE:HD12	1:F:63:ILE:N	2.32	0.45
1:A:20:ILE:HG22	1:A:118:VAL:HG21	1.99	0.45
1:C:106:PHE:HB3	1:C:111:ILE:HD12	1.98	0.45
1:C:155:MSE:SE	1:C:183:MSE:HE1	2.67	0.44
1:D:63:ILE:HD12	1:D:63:ILE:N	2.32	0.44
1:C:206:ARG:HB3	3:C:212:EDO:H12	1.99	0.44
1:C:20:ILE:CG2	1:C:115:LEU:HD12	2.47	0.44
1:D:78:LEU:HD23	1:D:78:LEU:N	2.33	0.44
1:A:106:PHE:HB3	1:A:111:ILE:HD12	1.98	0.44
1:B:146:LEU:HD21	1:B:154:ILE:HD12	1.99	0.44
1:F:43:GLN:O	1:F:47:GLN:N	2.46	0.44
1:B:63:ILE:HD12	1:B:63:ILE:N	2.34	0.43
1:A:148:VAL:HG23	5:D:235:HOH:O	2.18	0.43
1:A:78:LEU:N	1:A:78:LEU:HD23	2.33	0.43
1:D:100:ILE:O	1:D:100:ILE:HG23	2.19	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:78:LEU:HD23	1:E:78:LEU:N	2.34	0.43
1:D:164:LEU:HD12	1:D:185:ARG:HB2	2.00	0.43
1:B:26:SER:OG	1:B:52:VAL:HG13	2.19	0.42
1:F:204:ARG:O	1:F:205:ARG:C	2.57	0.42
1:E:63:ILE:HD12	1:E:63:ILE:N	2.34	0.42
1:E:106:PHE:HB3	1:E:111:ILE:HD12	2.02	0.42
1:D:20:ILE:HD13	1:D:114:LYS:HB3	2.01	0.42
1:B:95:LEU:O	1:B:130:ILE:HA	2.19	0.42
1:E:128:TYR:CZ	4:E:212:ACY:H3	2.54	0.42
1:A:151:TYR:OH	3:A:214:EDO:C1	2.68	0.42
1:A:20:ILE:HD13	1:A:114:LYS:HB3	2.02	0.42
1:D:149:TRP:O	1:D:153:LYS:HG3	2.20	0.41
1:E:155:MSE:SE	1:E:183:MSE:CE	3.08	0.41
1:B:78:LEU:N	1:B:78:LEU:HD23	2.35	0.41
1:D:4:LYS:HB3	1:D:23:PRO:HG2	2.02	0.41
1:A:178:PRO:HG2	1:D:149:TRP:CZ2	2.52	0.41
1:F:116:LEU:O	1:F:120:MSE:HG2	2.21	0.41
1:E:34:LEU:HD12	1:E:71:ILE:HB	2.01	0.41
1:A:63:ILE:HD12	1:A:63:ILE:N	2.34	0.41
1:E:115:LEU:HD12	1:E:115:LEU:HA	1.94	0.41
1:F:13:THR:OG1	1:F:14:SER:N	2.53	0.41
1:C:59:GLU:OE1	1:C:59:GLU:N	2.50	0.41
1:E:41:PHE:CZ	1:E:99:ALA:HB1	2.55	0.41
1:A:90:ASN:HB2	1:B:87:SER:O	2.21	0.40
1:E:28:ILE:O	1:E:31:THR:OG1	2.36	0.40
1:F:129:LEU:C	1:F:129:LEU:HD23	2.42	0.40

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	211/211 (100%)	206 (98%)	5 (2%)	0	100	100
1	B	207/211 (98%)	201 (97%)	4 (2%)	2 (1%)	15	14
1	C	208/211 (99%)	196 (94%)	10 (5%)	2 (1%)	15	14
1	D	209/211 (99%)	202 (97%)	7 (3%)	0	100	100
1	E	208/211 (99%)	205 (99%)	3 (1%)	0	100	100
1	F	203/211 (96%)	198 (98%)	5 (2%)	0	100	100
All	All	1246/1266 (98%)	1208 (97%)	34 (3%)	4 (0%)	41	47

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	207	PHE
1	C	35	ASP
1	C	36	PRO
1	B	205	ARG

### 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	172/170 (101%)	171 (99%)	1 (1%)	86	92
1	B	167/170 (98%)	167 (100%)	0	100	100
1	C	154/170 (91%)	151 (98%)	3 (2%)	57	68
1	D	173/170 (102%)	172 (99%)	1 (1%)	86	92
1	E	169/170 (99%)	167 (99%)	2 (1%)	71	82
1	F	150/170 (88%)	146 (97%)	4 (3%)	44	55
All	All	985/1020 (97%)	974 (99%)	11 (1%)	73	83

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	136	TYR

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	C	31	THR
1	C	152	ARG
1	C	204	ARG
1	D	115	LEU
1	E	109	GLN
1	E	115	LEU
1	F	13	THR
1	F	20	ILE
1	F	155	MSE
1	F	204	ARG

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

Mol	Chain	Res	Type
1	B	110	GLN
1	C	160	ASN
1	D	160	ASN
1	E	160	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 7 are monoatomic - leaving 11 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
3	EDO	D	212	-	3,3,3	0.42	0	2,2,2	0.25	0
4	ACY	E	212	-	3,3,3	0.72	0	3,3,3	1.07	0
3	EDO	A	214	-	3,3,3	0.47	0	2,2,2	0.26	0
4	ACY	A	216	-	3,3,3	0.80	0	3,3,3	0.78	0
3	EDO	A	213	-	3,3,3	0.52	0	2,2,2	0.27	0
3	EDO	C	212	-	3,3,3	0.53	0	2,2,2	0.19	0
3	EDO	A	215	-	3,3,3	0.64	0	2,2,2	0.22	0
4	ACY	F	213	-	3,3,3	0.71	0	3,3,3	0.86	0
3	EDO	D	213	-	3,3,3	0.52	0	2,2,2	0.36	0
4	ACY	E	213	2	3,3,3	0.96	0	3,3,3	0.66	0
3	EDO	F	212	-	3,3,3	0.48	0	2,2,2	0.31	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
3	EDO	D	212	-	-	0/1/1/1	-
3	EDO	A	214	-	-	1/1/1/1	-
3	EDO	A	213	-	-	0/1/1/1	-
3	EDO	C	212	-	-	0/1/1/1	-
3	EDO	A	215	-	-	1/1/1/1	-
3	EDO	D	213	-	-	1/1/1/1	-
3	EDO	F	212	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	213	EDO	O1-C1-C2-O2
3	F	212	EDO	O1-C1-C2-O2
3	A	214	EDO	O1-C1-C2-O2
3	A	215	EDO	O1-C1-C2-O2

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	212	ACY	2	0
3	A	214	EDO	1	0
4	A	216	ACY	1	0
3	C	212	EDO	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 [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	203/211 (96%)	1.58	67 (33%) 0 0	52, 57, 67, 82	0
1	B	200/211 (94%)	1.58	57 (28%) 0 0	51, 57, 69, 76	0
1	C	201/211 (95%)	1.70	74 (36%) 0 0	50, 57, 66, 83	0
1	D	202/211 (95%)	1.20	40 (19%) 1 2	47, 57, 71, 90	0
1	E	202/211 (95%)	1.03	33 (16%) 1 3	48, 57, 67, 85	0
1	F	197/211 (93%)	1.87	74 (37%) 0 0	51, 57, 64, 85	0
All	All	1205/1266 (95%)	1.49	345 (28%) 0 0	47, 57, 69, 90	0

All (345) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	146	LEU	8.7
1	C	40	ALA	7.7
1	F	69	ASN	7.2
1	F	27	GLN	7.2
1	F	56	ALA	7.1
1	E	44	PRO	6.9
1	D	209	TYR	6.8
1	A	210	ASP	6.6
1	D	207	PHE	6.6
1	F	36	PRO	6.5
1	F	26	SER	6.4
1	F	70	ASP	6.3
1	F	15	LEU	6.3
1	B	2	PHE	6.2
1	B	83	TYR	6.2
1	C	209	TYR	6.2
1	F	17	PRO	6.2
1	F	45	ALA	6.2
1	F	5	GLN	6.1

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	D	36	PRO	6.1
1	F	143	GLY	5.9
1	A	71	ILE	5.8
1	C	39	THR	5.7
1	F	63	ILE	5.6
1	F	75	VAL	5.6
1	F	6	PHE	5.5
1	B	89	GLY	5.5
1	F	97	LEU	5.5
1	D	97	LEU	5.5
1	A	39	THR	5.4
1	A	83	TYR	5.3
1	D	100	ILE	5.3
1	C	131	LEU	5.2
1	F	41	PHE	5.1
1	C	45	ALA	5.1
1	B	143	GLY	5.1
1	A	3	GLU	5.1
1	C	123	PRO	5.0
1	A	43	GLN	5.0
1	E	97	LEU	4.9
1	C	44	PRO	4.9
1	F	43	GLN	4.9
1	C	41	PHE	4.8
1	C	189	HIS	4.8
1	D	63	ILE	4.8
1	D	75	VAL	4.8
1	F	31	THR	4.7
1	A	150	ASP	4.7
1	A	2	PHE	4.7
1	E	37	GLY	4.7
1	D	143	GLY	4.7
1	B	147	SER	4.6
1	C	107	ARG	4.6
1	C	149	TRP	4.6
1	A	190	VAL	4.6
1	F	46	GLU	4.6
1	C	69	ASN	4.5
1	F	42	ARG	4.5
1	A	64	ILE	4.4
1	E	39	THR	4.4
1	E	130	ILE	4.4

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	F	104	ALA	4.4
1	C	127	HIS	4.3
1	C	14	SER	4.3
1	A	41	PHE	4.3
1	E	132	THR	4.3
1	B	138[A]	HIS	4.2
1	F	100	ILE	4.2
1	A	149	TRP	4.2
1	B	149	TRP	4.2
1	A	154	ILE	4.2
1	A	146	LEU	4.2
1	A	16	GLY	4.2
1	B	142	LYS	4.1
1	A	75	VAL	4.1
1	A	97	LEU	4.1
1	B	137	TRP	4.1
1	F	105	ARG	4.0
1	B	90	ASN	4.0
1	E	115	LEU	4.0
1	B	144	SER	4.0
1	A	44	PRO	4.0
1	A	145	GLY	4.0
1	E	63	ILE	4.0
1	B	123	PRO	4.0
1	F	39	THR	3.9
1	F	77	PHE	3.9
1	C	34	LEU	3.9
1	F	37	GLY	3.9
1	F	16	GLY	3.9
1	A	181	CYS	3.8
1	A	100	ILE	3.8
1	C	38	LEU	3.8
1	F	40	ALA	3.8
1	D	95	LEU	3.8
1	B	205	ARG	3.8
1	D	77	PHE	3.8
1	C	57	LEU	3.7
1	A	36	PRO	3.7
1	A	143	GLY	3.7
1	A	14	SER	3.7
1	E	131	LEU	3.7
1	C	144	SER	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	34	LEU	3.6
1	C	132	THR	3.6
1	C	130	ILE	3.6
1	E	64	ILE	3.6
1	B	148	VAL	3.6
1	D	131	LEU	3.6
1	C	16	GLY	3.6
1	F	57	LEU	3.6
1	F	106	PHE	3.6
1	F	50	ALA	3.5
1	D	206	ARG	3.5
1	C	192	PRO	3.5
1	F	103	ALA	3.5
1	D	130	ILE	3.4
1	C	52	VAL	3.4
1	C	104	ALA	3.4
1	B	192	PRO	3.4
1	F	24	VAL	3.4
1	C	42	ARG	3.4
1	E	76	THR	3.4
1	E	62	ILE	3.4
1	D	44	PRO	3.4
1	F	112	GLY	3.4
1	B	100	ILE	3.4
1	F	111	ILE	3.4
1	B	140	ASP	3.3
1	F	124	ALA	3.3
1	F	190	VAL	3.3
1	A	142	LYS	3.3
1	B	27	GLN	3.3
1	D	46	GLU	3.3
1	C	116	LEU	3.3
1	E	77	PHE	3.3
1	C	4	LYS	3.2
1	F	121	LEU	3.2
1	A	37	GLY	3.2
1	C	13	THR	3.2
1	D	132	THR	3.2
1	F	8	HIS	3.2
1	A	144	SER	3.2
1	F	161	HIS	3.2
1	D	76	THR	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	108	GLY	3.2
1	C	31	THR	3.1
1	F	144	SER	3.1
1	B	193	GLU	3.1
1	A	104	ALA	3.1
1	D	65	ALA	3.1
1	E	95	LEU	3.1
1	D	41	PHE	3.1
1	A	138[A]	HIS	3.1
1	B	179	ALA	3.1
1	F	191	ALA	3.1
1	C	75	VAL	3.1
1	E	74	TYR	3.1
1	E	209	TYR	3.0
1	F	13	THR	3.0
1	B	130	ILE	3.0
1	E	104	ALA	3.0
1	A	209	TYR	3.0
1	A	13	THR	3.0
1	C	10	THR	3.0
1	C	161	HIS	3.0
1	A	147	SER	3.0
1	A	91	ASN	3.0
1	E	43	GLN	3.0
1	F	53	GLU	3.0
1	A	24	VAL	3.0
1	C	76	THR	3.0
1	F	30	ALA	3.0
1	A	182	LEU	2.9
1	B	97	LEU	2.9
1	B	65	ALA	2.9
1	F	207	PHE	2.9
1	B	25	THR	2.9
1	D	182	LEU	2.9
1	A	193	GLU	2.9
1	B	132	THR	2.9
1	C	97	LEU	2.9
1	C	95	LEU	2.9
1	C	99	ALA	2.9
1	F	38	LEU	2.8
1	B	36	PRO	2.8
1	F	132	THR	2.8

*Continued on next page...*



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	D	27	GLN	2.8
1	E	41	PHE	2.8
1	A	45	ALA	2.8
1	D	144	SER	2.7
1	C	29	LEU	2.7
1	E	164	LEU	2.7
1	D	62	ILE	2.7
1	E	100	ILE	2.7
1	B	78	LEU	2.7
1	A	184	ALA	2.7
1	F	61	ARG	2.7
1	C	25	THR	2.7
1	F	131	LEU	2.7
1	A	0	GLY	2.7
1	E	14	SER	2.7
1	D	51	LEU	2.7
1	C	5	GLN	2.6
1	B	166	PHE	2.6
1	D	109	GLN	2.6
1	A	63	ILE	2.6
1	B	191	ALA	2.6
1	A	207	PHE	2.6
1	F	33	LYS	2.6
1	C	55	ALA	2.6
1	F	162	GLY	2.6
1	B	206	ARG	2.6
1	B	95	LEU	2.6
1	B	75	VAL	2.6
1	C	17	PRO	2.6
1	E	179	ALA	2.6
1	D	178	PRO	2.6
1	A	130	ILE	2.6
1	F	52	VAL	2.5
1	C	70	ASP	2.5
1	B	46	GLU	2.5
1	C	109	GLN	2.5
1	C	43	GLN	2.5
1	A	131	LEU	2.5
1	C	3	GLU	2.5
1	C	26	SER	2.5
1	C	117	GLU	2.5
1	D	78	LEU	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	116	LEU	2.5
1	D	149	TRP	2.5
1	A	132	THR	2.5
1	F	154	ILE	2.5
1	A	141	LEU	2.5
1	F	115	LEU	2.5
1	C	6	PHE	2.5
1	A	74	TYR	2.5
1	F	74	TYR	2.5
1	B	3	GLU	2.5
1	B	72	ILE	2.5
1	B	39	THR	2.4
1	A	179	ALA	2.4
1	A	73	GLY	2.4
1	C	49	GLU	2.4
1	D	64	ILE	2.4
1	D	129	LEU	2.4
1	B	91	ASN	2.4
1	C	56	ALA	2.4
1	C	37	GLY	2.4
1	C	32	TYR	2.4
1	A	116	LEU	2.4
1	B	69	ASN	2.4
1	E	116	LEU	2.4
1	C	58	GLU	2.4
1	F	98	GLY	2.4
1	A	136	TYR	2.4
1	C	115	LEU	2.4
1	B	42	ARG	2.4
1	D	181	CYS	2.4
1	F	18	VAL	2.4
1	A	60	GLY	2.4
1	B	145	GLY	2.4
1	C	164	LEU	2.4
1	A	40	ALA	2.4
1	D	39	THR	2.4
1	A	107	ARG	2.3
1	C	146	LEU	2.3
1	B	173	GLU	2.3
1	A	191	ALA	2.3
1	B	203	LEU	2.3
1	B	37	GLY	2.3

*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	181	CYS	2.3
1	A	65	ALA	2.3
1	F	49	GLU	2.3
1	B	178	PRO	2.3
1	D	99	ALA	2.3
1	F	14	SER	2.3
1	C	106	PHE	2.3
1	E	207	PHE	2.3
1	E	42	ARG	2.3
1	A	111	ILE	2.3
1	A	177	HIS	2.3
1	C	193	GLU	2.3
1	A	137	TRP	2.3
1	B	106	PHE	2.3
1	F	137	TRP	2.3
1	F	44	PRO	2.3
1	E	111	ILE	2.3
1	B	131	LEU	2.2
1	F	187	GLY	2.2
1	C	188	LYS	2.2
1	F	102	VAL	2.2
1	B	109	GLN	2.2
1	B	29	LEU	2.2
1	C	98	GLY	2.2
1	D	205	ARG	2.2
1	C	124	ALA	2.2
1	D	175	ALA	2.2
1	F	64	ILE	2.2
1	A	129	LEU	2.2
1	F	127	HIS	2.2
1	B	55	ALA	2.2
1	C	103	ALA	2.2
1	D	115	LEU	2.2
1	E	36	PRO	2.2
1	A	148	VAL	2.2
1	C	207	PHE	2.2
1	B	99	ALA	2.2
1	D	43	GLN	2.2
1	F	32	TYR	2.2
1	B	18	VAL	2.2
1	C	110	GLN	2.2
1	A	108	GLY	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	20	ILE	2.1
1	A	95	LEU	2.1
1	F	113	LYS	2.1
1	B	175	ALA	2.1
1	F	62	ILE	2.1
1	B	74	TYR	2.1
1	B	194	VAL	2.1
1	E	71	ILE	2.1
1	D	38	LEU	2.1
1	C	191	ALA	2.1
1	C	33	LYS	2.1
1	E	129	LEU	2.1
1	A	164	LEU	2.1
1	B	76	THR	2.1
1	D	184	ALA	2.1
1	A	172	PRO	2.1
1	C	83	TYR	2.1
1	E	48	HIS	2.0
1	B	182	LEU	2.0
1	E	78	LEU	2.0
1	F	60	GLY	2.0
1	A	77	PHE	2.0
1	A	186	ILE	2.0
1	C	15	LEU	2.0
1	F	10	THR	2.0
1	F	133	THR	2.0
1	E	107	ARG	2.0
1	F	114	LYS	2.0
1	C	8	HIS	2.0
1	A	72	ILE	2.0
1	C	11	LEU	2.0
1	C	141	LEU	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	EDO	A	214	4/4	0.78	0.34	76,79,80,81	0
3	EDO	A	213	4/4	0.83	0.29	82,83,83,83	0
4	ACY	F	213	4/4	0.84	0.34	88,88,88,89	0
2	CA	C	211	1/1	0.87	0.10	73,73,73,73	0
4	ACY	E	213	4/4	0.89	0.21	56,56,57,58	0
3	EDO	C	212	4/4	0.89	0.22	62,65,67,68	0
3	EDO	D	213	4/4	0.91	0.19	44,45,49,53	0
4	ACY	A	216	4/4	0.91	0.22	79,79,80,80	0
2	CA	E	211	1/1	0.94	0.10	51,51,51,51	0
3	EDO	F	212	4/4	0.95	0.07	69,71,73,75	0
4	ACY	E	212	4/4	0.95	0.14	48,49,49,50	0
2	CA	F	211	1/1	0.96	0.17	68,68,68,68	0
3	EDO	D	212	4/4	0.96	0.10	37,40,43,48	0
2	CA	A	212	1/1	0.97	0.04	68,68,68,68	0
3	EDO	A	215	4/4	0.97	0.10	38,48,48,55	0
2	CA	D	211	1/1	0.98	0.04	39,39,39,39	0
2	CA	A	211	1/1	0.98	0.04	48,48,48,48	0
2	CA	B	211	1/1	0.99	0.03	48,48,48,48	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.