



Full wwPDB X-ray Structure Validation Report ⓘ

Jun 7, 2020 – 01:54 am BST

PDB ID : 6GJF
Title : Ancestral endocellulase Cel5A
Authors : Gavira, J.A.; Perez-Jimenez, R.; Barruetabena-Garate, N.
Deposited on : 2018-05-16
Resolution : 1.45 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

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

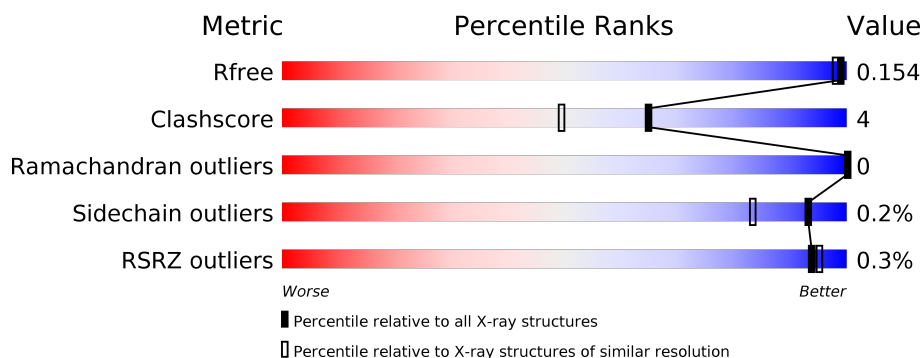
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.45 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	1156 (1.46-1.46)
Clashscore	141614	1202 (1.46-1.46)
Ramachandran outliers	138981	1178 (1.46-1.46)
Sidechain outliers	138945	1178 (1.46-1.46)
RSRZ outliers	127900	1139 (1.46-1.46)

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 ≥ 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	303	<div> <div>92%</div> <div>8%</div> </div>
1	B	303	<div> <div>92%</div> <div>8%</div> </div>
1	C	303	<div> <div>%</div> <div>95%</div> <div>5%</div> </div>
1	D	303	<div> <div>97%</div> <div>.</div> </div>
1	E	303	<div> <div>93%</div> <div>7%</div> </div>
1	F	303	<div> <div>90%</div> <div>10%</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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	EDO	A	300	-	-	X	-
2	EDO	A	306	-	-	X	-
2	EDO	E	302	-	-	X	-
2	EDO	F	301	-	-	-	X

2 Entry composition [i](#)

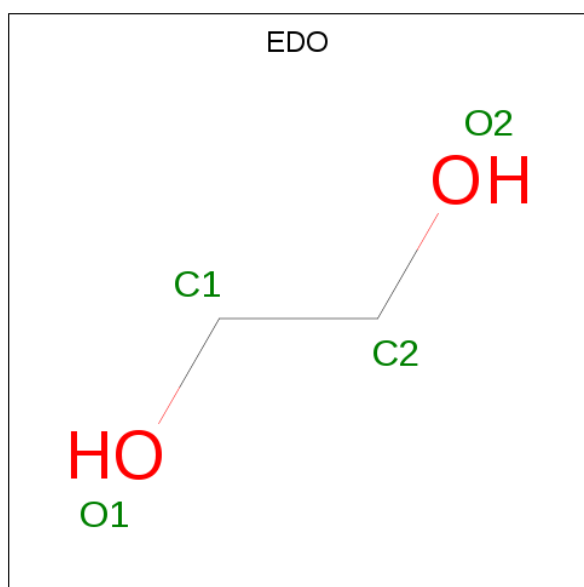
There are 6 unique types of molecules in this entry. The entry contains 32159 atoms, of which 14829 are hydrogens and 0 are deuteriums.

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 ENDOGLUCANASE.

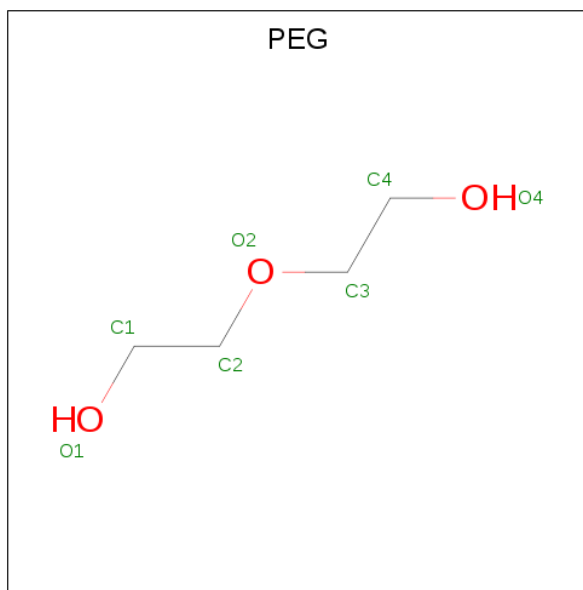
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	303	Total	C	H	N	O	S	0	24	0
			4908	1578	2410	428	481	11			
1	B	303	Total	C	H	N	O	S	0	27	0
			4936	1584	2434	431	477	10			
1	C	303	Total	C	H	N	O	S	0	25	0
			4895	1572	2406	430	477	10			
1	D	303	Total	C	H	N	O	S	0	26	0
			4869	1564	2388	430	478	9			
1	E	303	Total	C	H	N	O	S	1	34	0
			5018	1610	2474	438	486	10			
1	F	303	Total	C	H	N	O	S	0	42	0
			5100	1637	2517	441	493	12			

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C H O 10 2 6 2	0	0
2	A	1	Total C H O 10 2 6 2	0	0
2	A	1	Total C H O 10 2 6 2	0	0
2	A	1	Total C H O 10 2 6 2	0	0
2	B	1	Total C H O 10 2 6 2	0	0
2	B	1	Total C H O 10 2 6 2	0	0
2	C	1	Total C H O 10 2 6 2	0	0
2	D	1	Total C H O 10 2 6 2	0	0
2	E	1	Total C H O 10 2 6 2	0	0
2	E	1	Total C H O 10 2 6 2	0	0
2	E	1	Total C H O 10 2 6 2	0	0
2	F	1	Total C H O 10 2 6 2	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).

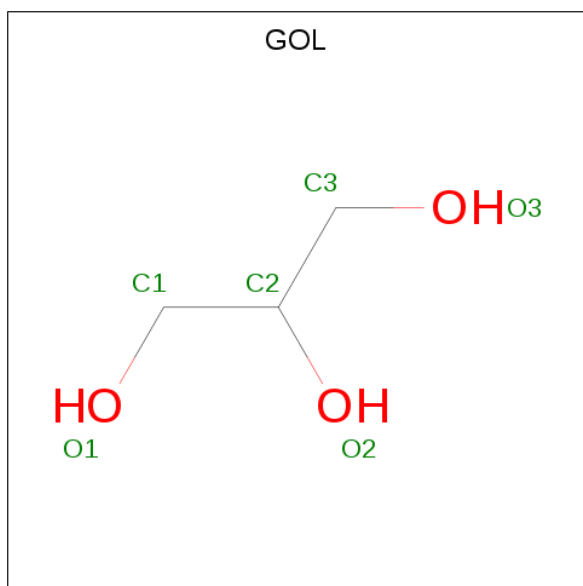


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	1
			34	8	20	6		

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

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

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	0	1
			28	6	16	6		
5	A	1	Total	C	H	O	0	0
			14	3	8	3		
5	A	1	Total	C	H	O	0	0
			14	3	8	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	H	O	0	1
			28	6	16	6		
5	B	1	Total	C	H	O	0	0
			12	3	6	3		
5	C	1	Total	C	H	O	0	0
			14	3	8	3		
5	C	1	Total	C	H	O	0	0
			13	3	7	3		
5	D	1	Total	C	H	O	0	0
			14	3	8	3		
5	E	1	Total	C	H	O	0	1
			28	6	16	6		
5	E	1	Total	C	H	O	0	0
			13	3	7	3		
5	F	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	367	Total	O	0	0
			367	367		
6	B	355	Total	O	0	0
			355	355		
6	C	317	Total	O	0	0
			317	317		
6	D	336	Total	O	0	0
			336	336		
6	E	369	Total	O	0	0
			369	369		
6	F	333	Total	O	0	0
			333	333		

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: ENDOGLUCANASE

Chain A: 



• Molecule 1: ENDOGLUCANASE

Chain B: 



• Molecule 1: ENDOGLUCANASE

Chain C: 



• Molecule 1: ENDOGLUCANASE

Chain D: 

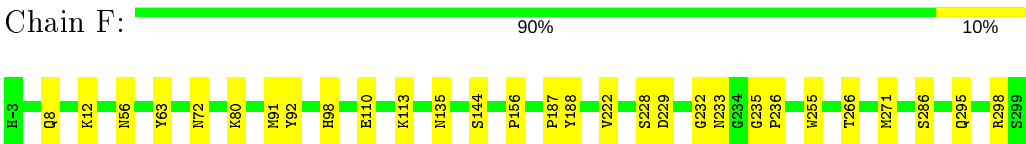


• Molecule 1: ENDOGLUCANASE

Chain E: 



• Molecule 1: ENDOGLUCANASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	51.44Å 235.64Å 74.56Å 90.00° 90.67° 90.00°	Depositor
Resolution (Å)	71.08 – 1.45 71.08 – 1.45	Depositor EDS
% Data completeness (in resolution range)	99.4 (71.08-1.45) 99.4 (71.08-1.45)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.12 (at 1.45Å)	Xtriage
Refinement program	PHENIX (1.13 _2998: ???)	Depositor
R, R_{free}	0.118 , 0.153 0.119 , 0.154	Depositor DCC
R_{free} test set	15359 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å ²)	14.5	Xtriage
Anisotropy	0.344	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 45.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.029 for h,-k,-l	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	32159	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.61% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: GOL, PEG, EDO, NA

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	0.45	0/2624	0.64	0/3564
1	B	0.43	0/2650	0.64	1/3597 (0.0%)
1	C	0.40	0/2637	0.61	0/3580
1	D	0.42	0/2629	0.62	0/3568
1	E	0.44	0/2724	0.64	1/3694 (0.0%)
1	F	0.44	0/2761	0.63	0/3744
All	All	0.43	0/16025	0.63	2/21747 (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
1	F	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	255	TRP	CA-CB-CG	5.21	123.61	113.70
1	B	255	TRP	CA-CB-CG	5.14	123.46	113.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	F	187	PRO	Mainchain

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	2498	2410	2379	26	0
1	B	2502	2434	2370	21	0
1	C	2489	2406	2338	11	0
1	D	2481	2388	2304	9	0
1	E	2544	2474	2386	20	0
1	F	2583	2517	2436	31	0
2	A	16	24	24	14	0
2	B	8	12	12	2	0
2	C	4	6	6	1	0
2	D	4	6	6	1	0
2	E	12	18	18	7	0
2	F	4	6	6	2	0
3	A	14	20	20	2	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	C	1	0	0	0	0
4	D	2	0	0	0	0
4	E	2	0	0	0	0
4	F	1	0	0	0	0
5	A	24	32	32	0	0
5	B	18	22	24	1	0
5	C	12	15	16	3	0
5	D	6	8	8	0	0
5	E	18	23	24	3	0
5	F	6	8	8	0	0
6	A	367	0	0	4	0
6	B	355	0	0	7	0
6	C	317	0	0	2	1
6	D	336	0	0	6	3
6	E	369	0	0	10	0
6	F	333	0	0	6	4
All	All	17330	14829	14417	122	4

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 (122) 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:C:22[B]:LYS:HA	1:C:22[B]:LYS:HE2	1.38	1.01
1:E:249[A]:ASN:OD1	6:E:401:HOH:O	1.83	0.95
1:E:152[B]:GLU:OE2	6:E:402:HOH:O	1.87	0.92
1:B:113[B]:LYS:NZ	1:B:153[B]:GLU:OE2	2.00	0.92
1:B:47[B]:LYS:NZ	6:B:402:HOH:O	2.10	0.84
1:B:144[B]:SER:OG	6:B:401:HOH:O	1.97	0.82
1:A:236:PRO:O	3:A:303[A]:PEG:H11	1.83	0.79
1:F:110[B]:GLU:OE2	1:F:113[B]:LYS:NZ	2.16	0.78
1:E:66[A]:GLU:OE1	6:E:403:HOH:O	2.01	0.77
1:E:152[A]:GLU:OE1	6:E:404:HOH:O	2.02	0.76
2:B:301:EDO:H11	6:B:460:HOH:O	1.85	0.75
1:C:22[B]:LYS:HE2	1:C:22[B]:LYS:CA	2.16	0.75
1:F:271[A]:MET:HE3	1:F:286:SER:HA	1.69	0.74
1:C:66:GLU:OE1	6:C:401:HOH:O	2.04	0.74
1:F:80:LYS:NZ	6:F:401:HOH:O	2.13	0.73
1:D:209[B]:LYS:HG3	6:D:402:HOH:O	1.92	0.70
1:F:232[B]:GLY:O	1:F:266[B]:THR:HG22	1.92	0.69
1:A:159[B]:ARG:HG2	1:A:163:PRO:HA	1.74	0.69
1:B:73:PRO:HB3	5:B:304:GOL:O1	1.92	0.69
1:A:236:PRO:O	3:A:303[B]:PEG:H12	1.93	0.68
1:B:12[A]:LYS:HG2	6:B:518:HOH:O	1.96	0.66
1:A:249[A]:ASN:ND2	6:A:401:HOH:O	2.15	0.65
1:F:271[A]:MET:CE	1:F:286:SER:HA	2.26	0.64
1:B:113[B]:LYS:HD2	1:B:153[B]:GLU:HG2	1.81	0.62
1:B:8:GLN:HG2	2:B:301:EDO:H12	1.82	0.61
1:F:228:SER:HB3	1:F:266[B]:THR:HG23	1.82	0.61
1:A:283:SER:CB	2:A:300:EDO:H12	2.31	0.59
2:E:301:EDO:H22	6:E:699:HOH:O	2.01	0.59
1:A:281:THR:HB	2:A:300:EDO:H22	1.85	0.58
1:E:239[B]:ASN:ND2	6:E:411:HOH:O	2.35	0.58
1:D:209[B]:LYS:CG	6:D:402:HOH:O	2.48	0.58
1:F:235[B]:GLY:C	6:F:405:HOH:O	2.41	0.58
1:A:8:GLN:O	2:A:306:EDO:C1	2.52	0.57
1:F:8:GLN:O	2:F:301:EDO:C1	2.53	0.57
1:B:252:ASN:OD1	6:B:403:HOH:O	2.18	0.56
1:B:173:THR:HG23	5:C:303:GOL:H32	1.86	0.56
1:E:15[A]:GLN:NE2	6:E:417:HOH:O	2.39	0.55
1:D:12[A]:LYS:HD3	6:D:506:HOH:O	2.06	0.55
1:A:10:SER:OG	2:A:306:EDO:H21	2.07	0.55
1:F:8:GLN:O	2:F:301:EDO:H12	2.07	0.54
1:F:271[A]:MET:HE1	6:F:733:HOH:O	2.07	0.54
1:D:19:GLU:HG3	2:D:303:EDO:H22	1.89	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:19:GLU:HA	2:A:306:EDO:H12	1.90	0.53
1:A:10:SER:CB	2:A:306:EDO:H21	2.38	0.53
1:C:176:GLN:O	1:C:209[A]:LYS:HE3	2.09	0.53
2:A:302:EDO:H11	6:A:626:HOH:O	2.09	0.52
1:F:233[B]:ASN:C	1:F:266[B]:THR:HG21	2.31	0.52
1:F:72[B]:ASN:O	1:F:72[B]:ASN:CG	2.48	0.51
1:A:50:ARG:HB2	1:A:91[B]:MET:CE	2.40	0.51
1:E:8:GLN:O	2:E:302:EDO:C1	2.59	0.51
1:B:176:GLN:O	1:B:209[A]:LYS:HE3	2.11	0.51
1:A:283:SER:OG	2:A:300:EDO:H12	2.11	0.51
1:F:188[B]:TYR:OH	6:F:402:HOH:O	2.17	0.50
1:E:8:GLN:O	2:E:302:EDO:H12	2.11	0.50
1:F:144[B]:SER:OG	6:F:403:HOH:O	2.18	0.50
1:A:159[B]:ARG:HG3	1:A:162:ASP:O	2.11	0.50
1:B:125[B]:ASN:OD1	5:E:300[B]:GOL:H11	2.12	0.50
6:E:540:HOH:O	1:F:188[B]:TYR:HE1	1.95	0.49
1:A:294[B]:GLU:HG2	6:A:409:HOH:O	2.13	0.49
1:E:18:ASP:O	2:E:302:EDO:H22	2.10	0.49
1:E:98:HIS:HA	1:E:135:ASN:HB2	1.94	0.49
1:A:283:SER:OG	2:A:300:EDO:C1	2.61	0.48
1:A:63:TYR:HA	1:A:98:HIS:HB3	1.95	0.48
1:E:10:SER:OG	2:E:302:EDO:H21	2.13	0.48
1:F:91[B]:MET:HG3	1:F:92:TYR:O	2.14	0.48
1:C:84:GLU:HG2	5:C:304:GOL:H2	1.95	0.47
1:D:209[B]:LYS:HD3	6:D:402:HOH:O	2.14	0.47
1:D:179:HIS:HB3	6:D:402:HOH:O	2.14	0.47
1:C:238:LEU:HB2	2:C:301:EDO:H12	1.96	0.47
1:C:263:LYS:HG2	1:C:265:GLU:HG3	1.95	0.47
1:F:229:ASP:OD1	1:F:233[B]:ASN:N	2.45	0.47
1:B:12[B]:LYS:HD3	6:B:438:HOH:O	2.15	0.47
1:E:113[B]:LYS:HB3	1:E:113[B]:LYS:NZ	2.30	0.47
1:B:98:HIS:HA	1:B:135:ASN:HB2	1.95	0.47
1:F:98:HIS:HA	1:F:135:ASN:HB2	1.96	0.47
1:F:228:SER:CB	1:F:266[B]:THR:HG23	2.43	0.47
1:C:179:HIS:ND1	1:C:209[A]:LYS:HD2	2.30	0.47
1:B:174:TRP:CZ3	5:C:303:GOL:H31	2.50	0.46
1:E:109[B]:LYS:NZ	1:E:109[B]:LYS:HB3	2.30	0.46
1:F:63:TYR:HA	1:F:98:HIS:HB3	1.97	0.46
1:E:63:TYR:HA	1:E:98:HIS:HB3	1.98	0.46
1:F:56:ASN:C	1:F:91[B]:MET:SD	2.94	0.46
1:A:283:SER:HB3	2:A:300:EDO:H12	1.96	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:E:300[B]:GOL:O1	6:E:405:HOH:O	2.18	0.46
1:E:10:SER:CB	2:E:302:EDO:H21	2.46	0.46
1:D:98:HIS:HA	1:D:135:ASN:HB2	1.97	0.45
1:A:98:HIS:HA	1:A:135:ASN:HB2	1.97	0.45
1:A:8:GLN:O	2:A:306:EDO:H12	2.16	0.45
1:B:63:TYR:HA	1:B:98:HIS:HB3	1.99	0.44
1:B:179:HIS:ND1	1:B:209[A]:LYS:HD2	2.33	0.44
1:E:295:GLN:HE22	1:E:298[A]:ARG:HH21	1.66	0.44
1:C:263:LYS:NZ	6:C:406:HOH:O	2.38	0.44
1:A:159[B]:ARG:NH1	1:A:189:SER:O	2.46	0.44
1:B:299:SER:HA	6:B:604:HOH:O	2.18	0.44
1:B:56:ASN:C	1:B:91[B]:MET:SD	2.97	0.43
1:A:298[A]:ARG:O	1:A:299[A]:SER:C	2.57	0.43
1:E:7:GLY:HA2	2:E:301:EDO:H21	1.99	0.43
1:D:258:TRP:HA	1:D:259:SER:HA	1.82	0.43
1:C:98:HIS:HA	1:C:135:ASN:HB2	2.00	0.43
1:F:235[B]:GLY:HA3	1:F:236[B]:PRO:C	2.37	0.43
1:A:282:ASP:HB3	2:A:300:EDO:O2	2.18	0.43
1:F:295:GLN:HE22	1:F:298[B]:ARG:CZ	2.32	0.43
1:E:249[A]:ASN:ND2	6:E:414:HOH:O	2.37	0.42
1:A:18:ASP:O	2:A:306:EDO:H22	2.20	0.42
1:C:63:TYR:HA	1:C:98:HIS:HB3	2.02	0.41
1:B:222:VAL:O	1:B:255:TRP:HA	2.20	0.41
1:F:271[A]:MET:CE	1:F:286:SER:CA	2.97	0.41
1:B:258:TRP:HA	1:B:259:SER:HA	1.84	0.41
1:F:222:VAL:O	1:F:255:TRP:HA	2.20	0.41
1:F:156:PRO:HG3	1:F:188[B]:TYR:OH	2.20	0.41
1:A:8:GLN:O	2:A:306:EDO:H11	2.20	0.41
1:D:209[B]:LYS:CD	6:D:402:HOH:O	2.67	0.41
1:F:232[B]:GLY:C	1:F:266[B]:THR:HG22	2.41	0.41
1:E:258:TRP:HA	1:E:259:SER:HA	1.84	0.40
1:A:15[B]:GLN:HG2	6:A:657:HOH:O	2.21	0.40
1:E:69:TYR:O	5:E:306:GOL:H32	2.21	0.40
1:F:12[B]:LYS:HD3	6:F:516:HOH:O	2.22	0.40
1:A:258:TRP:HA	1:A:259:SER:HA	1.77	0.40

All (4) 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)
6:D:669:HOH:O	6:F:725:HOH:O[2_554]	1.89	0.31

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:C:635:HOH:O	6:F:490:HOH:O[2_554]	1.94	0.26
6:D:588:HOH:O	6:F:662:HOH:O[2_554]	2.06	0.14
6:D:405:HOH:O	6:F:637:HOH:O[2_554]	2.16	0.04

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	323/303 (107%)	315 (98%)	8 (2%)	0	100	100
1	B	326/303 (108%)	320 (98%)	6 (2%)	0	100	100
1	C	325/303 (107%)	317 (98%)	8 (2%)	0	100	100
1	D	324/303 (107%)	317 (98%)	7 (2%)	0	100	100
1	E	335/303 (111%)	328 (98%)	7 (2%)	0	100	100
1	F	340/303 (112%)	333 (98%)	7 (2%)	0	100	100
All	All	1973/1818 (108%)	1930 (98%)	43 (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	274/252 (109%)	274 (100%)	0	100	100
1	B	275/252 (109%)	274 (100%)	1 (0%)	91	80

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	275/252 (109%)	271 (98%)	4 (2%)	65	35
1	D	274/252 (109%)	274 (100%)	0	100	100
1	E	282/252 (112%)	282 (100%)	0	100	100
1	F	288/252 (114%)	288 (100%)	0	100	100
All	All	1668/1512 (110%)	1663 (100%)	5 (0%)	93	82

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	4	GLU
1	C	195[A]	LEU
1	C	195[B]	LEU
1	C	233[A]	ASN
1	C	233[B]	ASN

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](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 38 ligands modelled in this entry, 10 are monoatomic - leaving 28 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	GOL	B	303[B]	-	5,5,5	0.67	0	5,5,5	0.78	0
5	GOL	B	303[C]	-	5,5,5	0.65	0	5,5,5	0.98	0
3	PEG	A	303[B]	-	6,6,6	0.50	0	5,5,5	0.28	0
3	PEG	A	303[A]	-	6,6,6	0.56	0	5,5,5	0.66	0
2	EDO	A	300	-	3,3,3	0.60	0	2,2,2	0.45	0
2	EDO	E	302	-	3,3,3	0.59	0	2,2,2	0.48	0
5	GOL	E	306	-	5,5,5	0.55	0	5,5,5	0.79	0
2	EDO	B	301	-	3,3,3	0.64	0	2,2,2	0.63	0
2	EDO	C	301	-	3,3,3	0.47	0	2,2,2	0.44	0
5	GOL	C	304	-	5,5,5	0.63	0	5,5,5	0.99	0
5	GOL	D	302	-	5,5,5	0.85	0	5,5,5	1.01	0
2	EDO	A	306	-	3,3,3	0.59	0	2,2,2	0.36	0
2	EDO	A	302	-	3,3,3	0.48	0	2,2,2	0.30	0
2	EDO	F	301	-	3,3,3	0.29	0	2,2,2	1.15	0
5	GOL	A	308	-	5,5,5	0.54	0	5,5,5	0.50	0
5	GOL	B	304	-	5,5,5	0.48	0	5,5,5	0.95	0
5	GOL	A	307[A]	-	5,5,5	0.63	0	5,5,5	0.69	0
2	EDO	D	303	-	3,3,3	0.49	0	2,2,2	0.40	0
5	GOL	A	307[B]	-	5,5,5	0.50	0	5,5,5	0.50	0
2	EDO	A	301	-	3,3,3	0.49	0	2,2,2	0.44	0
5	GOL	F	302	-	5,5,5	0.63	0	5,5,5	0.95	0
2	EDO	B	300	-	3,3,3	0.51	0	2,2,2	0.29	0
5	GOL	C	303	-	5,5,5	0.98	0	5,5,5	0.55	0
2	EDO	E	301	-	3,3,3	0.55	0	2,2,2	0.34	0
2	EDO	E	303	-	3,3,3	0.50	0	2,2,2	0.38	0
5	GOL	E	300[B]	-	5,5,5	0.79	0	5,5,5	1.00	0
5	GOL	A	309	-	5,5,5	0.62	0	5,5,5	0.85	0
5	GOL	E	300[A]	-	5,5,5	0.60	0	5,5,5	0.84	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
5	GOL	B	303[B]	-	-	4/4/4/4	-
5	GOL	B	303[C]	-	-	0/4/4/4	-
3	PEG	A	303[B]	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PEG	A	303[A]	-	-	2/4/4/4	-
2	EDO	A	300	-	-	1/1/1/1	-
2	EDO	E	302	-	-	1/1/1/1	-
5	GOL	E	306	-	-	4/4/4/4	-
2	EDO	B	301	-	-	0/1/1/1	-
2	EDO	C	301	-	-	1/1/1/1	-
5	GOL	C	304	-	-	0/4/4/4	-
5	GOL	D	302	-	-	0/4/4/4	-
2	EDO	A	306	-	-	1/1/1/1	-
2	EDO	A	302	-	-	1/1/1/1	-
2	EDO	F	301	-	-	1/1/1/1	-
5	GOL	A	308	-	-	2/4/4/4	-
5	GOL	B	304	-	-	2/4/4/4	-
5	GOL	A	307[A]	-	-	0/4/4/4	-
2	EDO	D	303	-	-	1/1/1/1	-
5	GOL	A	307[B]	-	-	4/4/4/4	-
2	EDO	A	301	-	-	0/1/1/1	-
5	GOL	F	302	-	-	2/4/4/4	-
2	EDO	B	300	-	-	1/1/1/1	-
5	GOL	C	303	-	-	4/4/4/4	-
2	EDO	E	301	-	-	1/1/1/1	-
2	EDO	E	303	-	-	0/1/1/1	-
5	GOL	E	300[B]	-	-	4/4/4/4	-
5	GOL	A	309	-	-	0/4/4/4	-
5	GOL	E	300[A]	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (41) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	304	GOL	C1-C2-C3-O3
5	A	307[B]	GOL	O1-C1-C2-C3
5	F	302	GOL	O1-C1-C2-C3
5	C	303	GOL	O1-C1-C2-C3
5	C	303	GOL	C1-C2-C3-O3
5	E	300[B]	GOL	C1-C2-C3-O3
3	A	303[A]	PEG	O1-C1-C2-O2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
5	E	306	GOL	O1-C1-C2-C3
5	E	306	GOL	C1-C2-C3-O3
5	B	303[B]	GOL	O1-C1-C2-C3
5	B	303[B]	GOL	C1-C2-C3-O3
5	A	307[B]	GOL	C1-C2-C3-O3
5	A	308	GOL	C1-C2-C3-O3
5	E	300[B]	GOL	O1-C1-C2-C3
5	B	304	GOL	O2-C2-C3-O3
5	A	307[B]	GOL	O1-C1-C2-O2
5	A	307[B]	GOL	O2-C2-C3-O3
5	E	300[B]	GOL	O2-C2-C3-O3
2	E	301	EDO	O1-C1-C2-O2
5	F	302	GOL	O1-C1-C2-O2
5	C	303	GOL	O1-C1-C2-O2
5	C	303	GOL	O2-C2-C3-O3
5	E	300[B]	GOL	O1-C1-C2-O2
5	B	303[B]	GOL	O1-C1-C2-O2
5	A	308	GOL	O2-C2-C3-O3
2	B	300	EDO	O1-C1-C2-O2
3	A	303[A]	PEG	C1-C2-O2-C3
5	E	300[A]	GOL	O2-C2-C3-O3
2	C	301	EDO	O1-C1-C2-O2
2	A	302	EDO	O1-C1-C2-O2
2	A	300	EDO	O1-C1-C2-O2
3	A	303[B]	PEG	O1-C1-C2-O2
2	E	302	EDO	O1-C1-C2-O2
5	B	303[B]	GOL	O2-C2-C3-O3
2	F	301	EDO	O1-C1-C2-O2
5	E	306	GOL	O2-C2-C3-O3
2	A	306	EDO	O1-C1-C2-O2
2	D	303	EDO	O1-C1-C2-O2
5	E	300[A]	GOL	C1-C2-C3-O3
3	A	303[B]	PEG	C1-C2-O2-C3
5	E	306	GOL	O1-C1-C2-O2

There are no ring outliers.

16 monomers are involved in 36 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	303[B]	PEG	1	0
3	A	303[A]	PEG	1	0
2	A	300	EDO	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	302	EDO	5	0
5	E	306	GOL	1	0
2	B	301	EDO	2	0
2	C	301	EDO	1	0
5	C	304	GOL	1	0
2	A	306	EDO	7	0
2	A	302	EDO	1	0
2	F	301	EDO	2	0
5	B	304	GOL	1	0
2	D	303	EDO	1	0
5	C	303	GOL	2	0
2	E	301	EDO	2	0
5	E	300[B]	GOL	2	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, 95th 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 ²)	Q<0.9
1	A	303/303 (100%)	-0.59	1 (0%) 94 95	10, 15, 26, 46	0
1	B	303/303 (100%)	-0.60	0 100 100	10, 15, 25, 51	0
1	C	303/303 (100%)	-0.53	2 (0%) 87 89	11, 18, 28, 44	1 (0%)
1	D	303/303 (100%)	-0.59	1 (0%) 94 95	10, 17, 27, 59	0
1	E	303/303 (100%)	-0.62	1 (0%) 94 95	9, 15, 24, 52	0
1	F	303/303 (100%)	-0.53	0 100 100	11, 17, 31, 53	0
All	All	1818/1818 (100%)	-0.58	5 (0%) 94 95	9, 16, 27, 59	1 (0%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	299[A]	SER	5.9
1	C	299[A]	SER	4.7
1	E	299	SER	2.9
1	C	298[A]	ARG	2.5
1	D	299[A]	SER	2.4

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

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

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, 95th 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 ²)	Q<0.9
2	EDO	F	301	4/4	0.52	0.59	120,151,165,181	0
5	GOL	F	302	6/6	0.60	0.28	54,68,80,84	0
2	EDO	B	301	4/4	0.64	0.23	54,65,67,69	0
3	PEG	A	303[A]	7/7	0.71	0.21	31,45,55,59	17
3	PEG	A	303[B]	7/7	0.71	0.21	26,35,43,44	17
5	GOL	A	307[B]	6/6	0.73	0.26	27,37,44,45	14
5	GOL	A	307[A]	6/6	0.73	0.26	45,55,61,61	14
5	GOL	B	304	6/6	0.76	0.14	35,57,84,97	0
2	EDO	A	302	4/4	0.80	0.12	57,68,72,73	0
5	GOL	A	309	6/6	0.81	0.20	48,59,67,70	0
2	EDO	B	300	4/4	0.82	0.15	52,62,67,71	0
5	GOL	E	306	6/6	0.82	0.19	52,59,70,71	0
2	EDO	D	303	4/4	0.83	0.21	36,45,53,54	0
5	GOL	B	303[B]	6/6	0.84	0.18	31,37,41,42	14
5	GOL	B	303[C]	6/6	0.84	0.18	41,49,50,52	14
5	GOL	A	308	6/6	0.84	0.17	42,57,66,69	0
2	EDO	A	300	4/4	0.84	0.37	47,57,61,65	0
5	GOL	E	300[B]	6/6	0.86	0.19	27,35,41,42	14
5	GOL	E	300[A]	6/6	0.86	0.19	52,63,64,64	14
2	EDO	E	301	4/4	0.88	0.26	39,47,49,49	0
5	GOL	C	304	6/6	0.89	0.13	42,65,84,88	0
2	EDO	A	306	4/4	0.90	0.19	35,42,44,48	0
5	GOL	C	303	6/6	0.91	0.14	28,68,104,119	0
2	EDO	A	301	4/4	0.91	0.23	43,51,54,56	0
2	EDO	E	302	4/4	0.92	0.26	34,41,48,48	0
2	EDO	E	303	4/4	0.92	0.14	35,43,46,47	0
2	EDO	C	301	4/4	0.94	0.16	52,63,69,72	0
5	GOL	D	302	6/6	0.97	0.11	30,36,40,40	0
4	NA	B	305	1/1	0.99	0.10	14,14,14,14	0
4	NA	B	302	1/1	0.99	0.06	17,17,17,17	0
4	NA	D	301	1/1	0.99	0.05	17,17,17,17	0
4	NA	F	300	1/1	0.99	0.08	14,14,14,14	0
4	NA	A	304	1/1	1.00	0.05	15,15,15,15	0
4	NA	E	305	1/1	1.00	0.05	16,16,16,16	0
4	NA	A	305	1/1	1.00	0.06	17,17,17,17	0
4	NA	E	304	1/1	1.00	0.05	16,16,16,16	0
4	NA	D	300	1/1	1.00	0.08	16,16,16,16	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	NA	C	302	1/1	1.00	0.05	17,17,17,17	0

6.5 Other polymers [i](#)

There are no such residues in this entry.