



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 15, 2017 – 12:54 am GMT

PDB ID : 4CYZ
Title : Structure of the A_mallard_Sweden_51_2002 H10 Avian Haemmagglutinin
in complex with avian receptor analog LSTA
Authors : Vachieri, S.G.; Xiong, X.; Collins, P.J.; Walker, P.A.; Martin, S.R.; Haire,
L.F.; McCauley, J.W.; Gamblin, S.J.; Skehel, J.J.
Deposited on : 2014-04-16
Resolution : 2.40 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

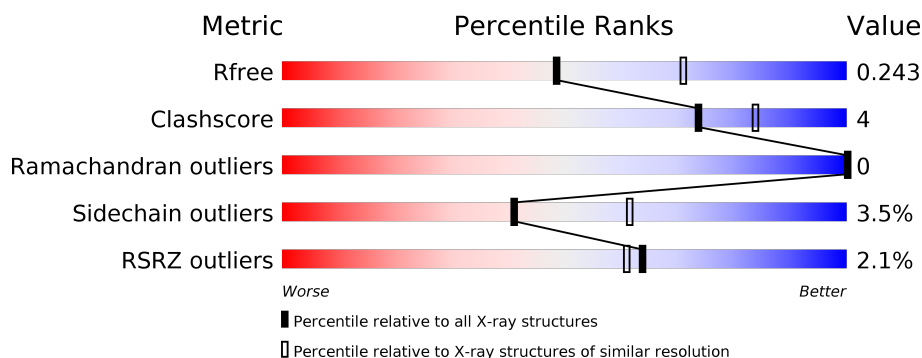
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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}	100719	3166 (2.40-2.40)
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (2.40-2.40)

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. 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	318	<div> <div>89%</div> <div>10% .</div> </div>
1	C	318	<div> <div>2%</div> <div>92%</div> <div>8% .</div> </div>
1	E	318	<div> <div>3%</div> <div>91%</div> <div>8% .</div> </div>
2	B	172	<div> <div>84%</div> <div>16% .</div> </div>
2	D	172	<div> <div>5%</div> <div>90%</div> <div>10% .</div> </div>
2	F	172	<div> <div>6%</div> <div>92%</div> <div>7% .</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
3	NAG	A	420	X	-	-	-
3	NAG	C	420	X	-	-	-
5	EDO	A	1326	-	-	-	X
5	EDO	A	1327	-	-	-	X
5	EDO	B	1173	-	-	-	X
5	EDO	C	1326	-	-	-	X
5	EDO	E	1326	-	-	-	X
5	EDO	E	1327	-	-	-	X
5	EDO	E	1328	-	-	-	X
5	EDO	F	1173	-	-	-	X
6	NAG	B	201	-	-	-	X
6	NAG	D	201	-	-	-	X
7	NAG	C	401	-	-	-	X

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 12227 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 HEMAGGLUTININ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	318	Total	C	N	O	S	0	0	0
			2423	1498	436	473	16			
1	C	318	Total	C	N	O	S	0	0	0
			2419	1496	435	472	16			
1	E	318	Total	C	N	O	S	0	0	0
			2419	1495	435	473	16			

- Molecule 2 is a protein called HEMAGGLUTININ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	172	Total	C	N	O	S	0	0	0
			1381	853	238	282	8			
2	D	172	Total	C	N	O	S	0	0	0
			1376	851	238	279	8			
2	F	172	Total	C	N	O	S	0	0	0
			1383	854	239	282	8			

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	E	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	3	Total	C	N	O	0	0
			46	25	2	19		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



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

- Molecule 6 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	5	Total	C	N	O	0	0
			61	34	2	25		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	D	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 7 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	C	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 8 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	E	2	Total	C	N	O	0	0
			31	17	1	13		

- Molecule 9 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	F	4	Total	C	N	O	0	0
			50	28	2	20		

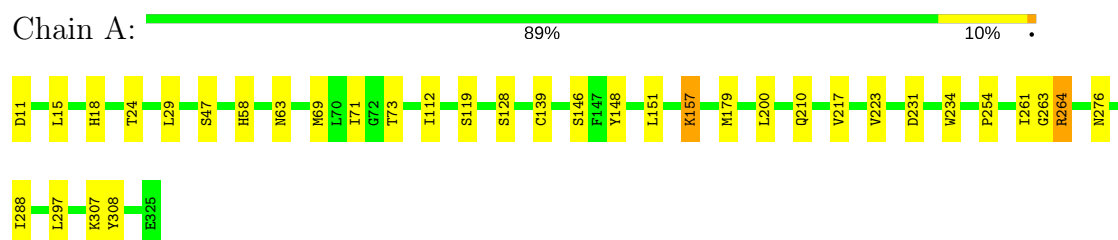
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	132	Total	O	0	0
			132	132		
10	B	61	Total	O	0	0
			61	61		
10	C	123	Total	O	0	0
			123	123		
10	D	26	Total	O	0	0
			26	26		
10	E	57	Total	O	0	0
			57	57		
10	F	22	Total	O	0	0
			22	22		

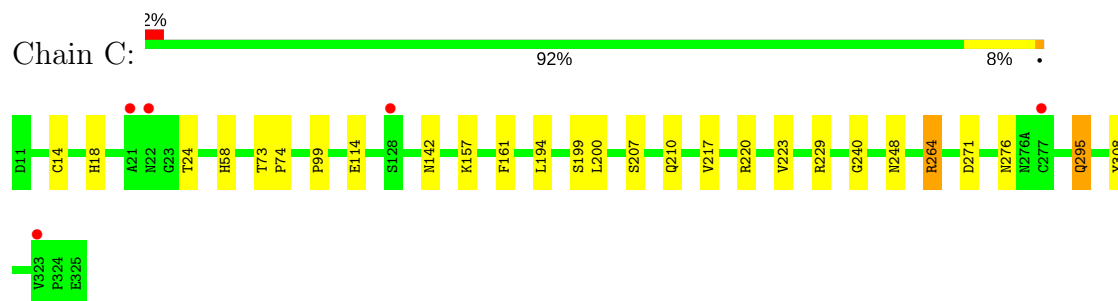
3 Residue-property plots [i](#)

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.

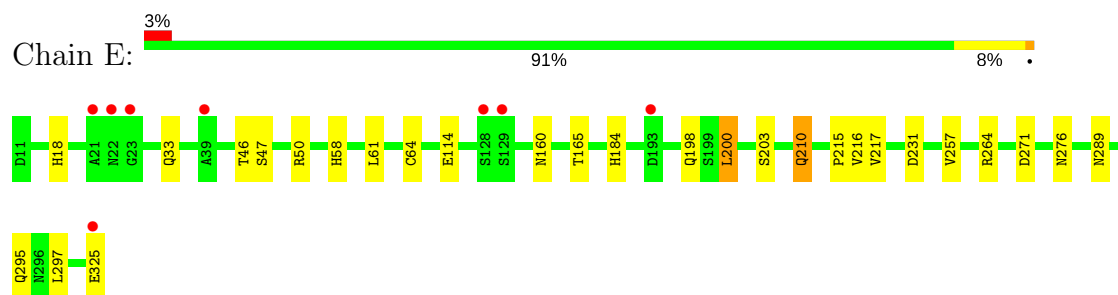
• Molecule 1: HEMAGGLUTININ



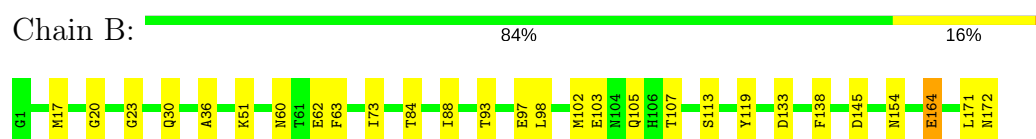
• Molecule 1: HEMAGGLUTININ



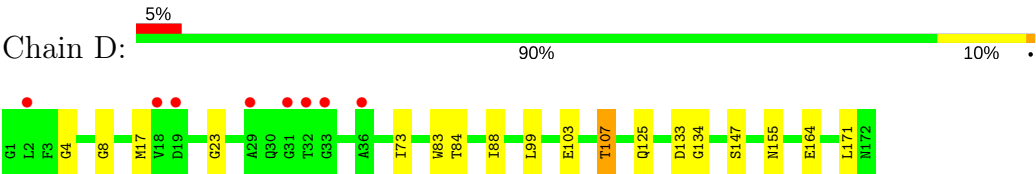
• Molecule 1: HEMAGGLUTININ



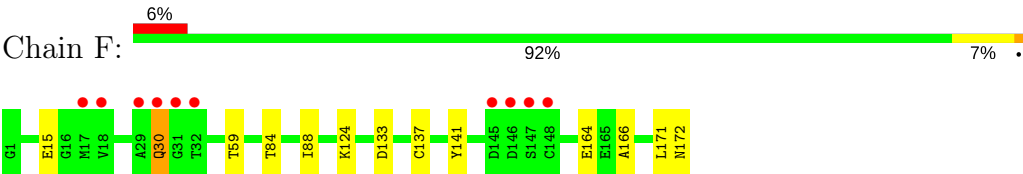
• Molecule 2: HEMAGGLUTININ



• Molecule 2: HEMAGGLUTININ



• Molecule 2: HEMAGGLUTININ



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	67.07Å 230.58Å 68.46Å 90.00° 110.41° 90.00°	Depositor
Resolution (Å)	115.29 – 2.40 48.66 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.0 (115.29-2.40) 99.0 (48.66-2.40)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.70 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.205 , 0.244 0.205 , 0.243	Depositor DCC
R_{free} test set	3771 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	37.4	Xtriage
Anisotropy	0.313	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 37.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.029 for l,-k,h	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	12227	wwPDB-VP
Average B, all atoms (Å ²)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% 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: BMA, NAG, EDO, SIA, GAL, MAN

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	0/2472	0.55	0/3350
1	C	0.39	1/2468 (0.0%)	0.56	0/3345
1	E	0.35	0/2468	0.53	0/3346
2	B	0.37	0/1406	0.52	0/1899
2	D	0.35	0/1401	0.50	0/1892
2	F	0.33	0/1408	0.50	0/1901
All	All	0.36	1/11623 (0.0%)	0.54	0/15733

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	114	GLU	CD-OE2	6.76	1.33	1.25

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2423	0	2356	25	0
1	C	2419	0	2350	16	0
1	E	2419	0	2346	16	0
2	B	1381	0	1268	21	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1376	0	1267	13	0
2	F	1383	0	1275	7	0
3	A	28	0	26	0	0
3	B	28	0	26	1	0
3	C	14	0	13	0	0
3	E	14	0	13	0	0
4	A	46	0	40	0	0
5	A	12	0	18	4	0
5	B	4	0	6	0	0
5	C	8	0	12	0	0
5	E	12	0	18	1	0
5	F	8	0	12	1	0
6	B	61	0	51	0	0
6	D	61	0	52	1	0
7	C	28	0	25	1	0
8	E	31	0	26	0	0
9	F	50	0	43	1	0
10	A	132	0	0	1	0
10	B	61	0	0	2	0
10	C	123	0	0	2	0
10	D	26	0	0	0	0
10	E	57	0	0	0	0
10	F	22	0	0	0	0
All	All	12227	0	11243	87	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 (87) 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:E:114:GLU:HG2	1:E:264:ARG:HH12	1.22	1.03
1:E:114:GLU:HG2	1:E:264:ARG:NH1	1.82	0.94
1:E:257:VAL:HA	5:E:1327:EDO:H11	1.56	0.88
1:C:264:ARG:HH11	1:C:264:ARG:HG3	1.43	0.83
1:A:264:ARG:HG3	1:A:264:ARG:HH11	1.45	0.80
2:F:171:LEU:O	2:F:172:ASN:HB2	1.84	0.77
2:D:103:GLU:O	2:D:107:THR:HG22	1.86	0.75
1:C:264:ARG:CG	1:C:264:ARG:HH11	2.03	0.70
1:A:308:TYR:H	5:A:1327:EDO:H12	1.56	0.69
2:B:103:GLU:O	2:B:107:THR:HG23	1.94	0.68

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
10:C:2120:HOH:O	2:D:107:THR:HG21	1.95	0.67
1:C:58:HIS:HE1	1:C:276:ASN:HD21	1.43	0.66
1:E:50:ARG:HH11	1:E:50:ARG:HB2	1.62	0.65
2:B:51:LYS:HZ2	2:B:107:THR:HG22	1.61	0.64
1:A:47:SER:HB2	1:A:297:LEU:HD22	1.81	0.63
1:A:263:GLY:HA2	9:F:204:MAN:H62	1.79	0.62
1:A:264:ARG:HH11	1:A:264:ARG:CG	2.12	0.62
2:D:125:GLN:HE22	2:D:155:ASN:HA	1.65	0.61
1:E:200:LEU:HD23	1:E:215:PRO:HG2	1.82	0.61
1:C:240:GLY:O	7:C:401:NAG:H82	2.01	0.60
2:B:164:GLU:H	2:B:164:GLU:CD	2.04	0.60
2:B:154:ASN:HD21	3:B:211:NAG:H5	1.66	0.60
1:A:223:VAL:CG2	1:C:207:SER:HB2	2.31	0.60
1:A:15:LEU:HD13	2:B:119:TYR:HA	1.85	0.58
2:D:17:MET:HE3	2:D:23:GLY:HA3	1.86	0.57
2:F:30:GLN:HE21	2:F:30:GLN:H	1.53	0.57
2:D:17:MET:CE	2:D:23:GLY:HA3	2.34	0.57
2:B:30:GLN:HE22	2:B:145:ASP:HB2	1.68	0.57
1:A:11:ASP:HB3	10:B:2055:HOH:O	2.04	0.56
2:F:141:TYR:O	2:F:166:ALA:HA	2.06	0.56
2:D:164:GLU:CD	2:D:164:GLU:H	2.09	0.56
2:B:51:LYS:NZ	2:B:107:THR:HG22	2.21	0.55
1:A:307:LYS:HA	5:A:1327:EDO:H12	1.87	0.55
1:E:58:HIS:CE1	1:E:276:ASN:HD21	2.26	0.54
5:A:1328:EDO:H22	2:B:63:PHE:CG	2.43	0.54
2:D:84:THR:O	2:D:88:ILE:HG12	2.09	0.52
1:A:231:ASP:OD2	1:C:210:GLN:NE2	2.39	0.51
1:C:161:PHE:HB3	1:C:248:ASN:O	2.09	0.51
1:A:119:SER:OG	1:A:261:ILE:HD11	2.11	0.51
1:E:184:HIS:HB3	1:E:216:VAL:O	2.11	0.51
1:A:151:LEU:HD23	1:A:254:PRO:HA	1.93	0.51
2:D:125:GLN:NE2	2:D:155:ASN:HA	2.27	0.50
1:A:128:SER:O	1:A:157:LYS:HE2	2.12	0.50
1:E:50:ARG:NH1	1:E:50:ARG:HB2	2.26	0.49
1:A:69:MET:HE3	1:A:112:ILE:HG21	1.94	0.49
2:B:172:ASN:ND2	10:B:2059:HOH:O	2.45	0.49
1:A:69:MET:CE	1:A:112:ILE:HG21	2.43	0.49
1:C:14:CYS:HB3	10:C:2002:HOH:O	2.12	0.48
1:E:58:HIS:HE1	1:E:276:ASN:HD21	1.60	0.47
2:B:51:LYS:HZ2	2:B:107:THR:CG2	2.27	0.47
1:A:15:LEU:HD12	2:B:138:PHE:HE1	1.80	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:84:THR:O	2:B:88:ILE:HG12	2.14	0.46
1:C:73:THR:HG22	1:C:74:PRO:HD2	1.98	0.46
1:A:58:HIS:HE1	1:A:276:ASN:HD21	1.64	0.45
1:A:29:LEU:HD12	2:B:105:GLN:HG2	1.99	0.45
2:F:133:ASP:OD2	2:F:137:CYS:HB2	2.17	0.45
1:E:325:GLU:HG3	2:F:15:GLU:HG3	1.99	0.45
2:B:17:MET:CE	2:B:23:GLY:HA3	2.48	0.44
1:C:229:ARG:HG3	1:E:210:GLN:HG2	1.99	0.44
1:E:47:SER:HB2	1:E:297:LEU:HD22	1.99	0.44
1:C:264:ARG:NH1	1:C:264:ARG:CG	2.72	0.44
1:C:99:PRO:HB3	1:C:223:VAL:HB	1.99	0.44
1:A:179:MET:HG2	1:A:234:TRP:HB3	1.99	0.44
1:C:295:GLN:O	1:C:308:TYR:HA	2.17	0.44
2:D:134:GLY:HA2	2:F:124:LYS:HD3	2.00	0.44
2:D:99:LEU:O	2:D:103:GLU:HG2	2.18	0.43
2:D:83:TRP:HH2	5:F:1173:EDO:H11	1.83	0.43
10:A:2125:HOH:O	2:B:107:THR:HG21	2.17	0.43
1:C:58:HIS:CE1	1:C:276:ASN:HD21	2.29	0.43
1:A:210:GLN:NE2	1:E:231:ASP:OD2	2.52	0.43
1:C:264:ARG:NH1	1:C:264:ARG:HG3	2.22	0.43
1:E:160:ASN:HD21	1:E:198:GLN:H	1.65	0.43
2:B:171:LEU:HD22	2:D:171:LEU:HD11	1.99	0.43
6:D:201:NAG:O7	1:E:114:GLU:OE2	2.36	0.43
1:A:71:ILE:O	1:A:148:TYR:HB3	2.18	0.43
2:B:51:LYS:CE	2:B:107:THR:HG22	2.48	0.42
1:A:288:ILE:HG12	1:A:297:LEU:HD13	2.01	0.42
2:B:98:LEU:O	2:B:102:MET:HB2	2.21	0.41
1:E:61:LEU:O	1:E:64:CYS:HB3	2.20	0.41
2:F:84:THR:O	2:F:88:ILE:HG12	2.20	0.41
2:B:93:THR:O	2:B:97:GLU:HG3	2.19	0.41
1:A:223:VAL:HG21	1:C:207:SER:HB2	2.03	0.41
2:D:4:GLY:O	2:D:8:GLY:HA3	2.21	0.41
1:A:139:CYS:O	1:A:146:SER:HB3	2.21	0.40
2:B:60:ASN:OD1	2:B:60:ASN:N	2.55	0.40
1:A:308:TYR:N	5:A:1327:EDO:H12	2.28	0.40
2:B:20:GLY:HA3	2:B:36:ALA:HB1	2.04	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	316/318 (99%)	311 (98%)	5 (2%)	0	100	100
1	C	316/318 (99%)	310 (98%)	6 (2%)	0	100	100
1	E	316/318 (99%)	310 (98%)	6 (2%)	0	100	100
2	B	170/172 (99%)	166 (98%)	4 (2%)	0	100	100
2	D	170/172 (99%)	165 (97%)	5 (3%)	0	100	100
2	F	170/172 (99%)	165 (97%)	5 (3%)	0	100	100
All	All	1458/1470 (99%)	1427 (98%)	31 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	269/270 (100%)	261 (97%)	8 (3%)	46	67
1	C	268/270 (99%)	256 (96%)	12 (4%)	32	50
1	E	268/270 (99%)	257 (96%)	11 (4%)	35	54
2	B	144/146 (99%)	139 (96%)	5 (4%)	41	61
2	D	143/146 (98%)	139 (97%)	4 (3%)	49	70
2	F	145/146 (99%)	142 (98%)	3 (2%)	59	78
All	All	1237/1248 (99%)	1194 (96%)	43 (4%)	41	61

All (43) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	HIS
1	A	24	THR
1	A	63	ASN
1	A	73	THR
1	A	157	LYS
1	A	200	LEU
1	A	217	VAL
1	A	264	ARG
2	B	62	GLU
2	B	73	ILE
2	B	113	SER
2	B	133	ASP
2	B	164	GLU
1	C	18	HIS
1	C	24	THR
1	C	142	ASN
1	C	157	LYS
1	C	194	LEU
1	C	199	SER
1	C	200	LEU
1	C	217	VAL
1	C	220	ARG
1	C	264	ARG
1	C	271	ASP
1	C	295	GLN
2	D	73	ILE
2	D	107	THR
2	D	133	ASP
2	D	147	SER
1	E	18	HIS
1	E	33	GLN
1	E	46	THR
1	E	165	THR
1	E	200	LEU
1	E	203	SER
1	E	210	GLN
1	E	217	VAL
1	E	271	ASP
1	E	289	ASN
1	E	295	GLN
2	F	30	GLN
2	F	59	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	164	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	17	HIS
1	A	276	ASN
2	B	26	HIS
2	B	30	GLN
2	B	154	ASN
2	B	172	ASN
1	C	58	HIS
2	D	30	GLN
2	D	125	GLN
1	E	58	HIS
1	E	160	ASN
1	E	212	ASN
2	F	30	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

21 carbohydrates 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
4	SIA	A	409	4	17,20,21	0.39	0	19,28,31	0.69	0
4	GAL	A	410	4	11,11,12	0.63	0	13,15,17	0.90	0
4	NAG	A	411	4	15,15,15	0.44	0	21,21,21	0.83	0
6	NAG	B	201	2,6	14,14,15	0.64	0	15,19,21	2.23	5 (33%)
6	NAG	B	202	6	14,14,15	0.66	0	15,19,21	0.99	1 (6%)
6	BMA	B	203	6	11,11,12	0.45	0	13,15,17	1.06	2 (15%)
6	MAN	B	204	6	11,11,12	0.78	0	13,15,17	1.87	2 (15%)
6	MAN	B	207	6	11,11,12	0.56	0	13,15,17	1.10	1 (7%)
7	NAG	C	401	1,7	14,14,15	0.49	0	15,19,21	2.39	3 (20%)
7	NAG	C	402	7	14,14,15	0.63	0	15,19,21	1.04	1 (6%)
6	NAG	D	201	2,6	14,14,15	0.57	0	15,19,21	1.21	1 (6%)
6	NAG	D	202	6	14,14,15	0.62	0	15,19,21	0.82	0
6	BMA	D	203	6	11,11,12	0.41	0	13,15,17	1.54	3 (23%)
6	MAN	D	204	6	11,11,12	0.61	0	13,15,17	0.84	0
6	MAN	D	207	6	11,11,12	0.59	0	13,15,17	1.96	4 (30%)
8	SIA	E	409	8	17,20,21	0.30	0	19,28,31	0.63	0
8	GAL	E	410	8	11,11,12	0.66	0	13,15,17	0.89	1 (7%)
9	NAG	F	201	9,2	14,14,15	0.61	0	15,19,21	0.80	0
9	NAG	F	202	9	14,14,15	0.59	0	15,19,21	1.27	1 (6%)
9	BMA	F	203	9	11,11,12	0.70	0	13,15,17	2.44	3 (23%)
9	MAN	F	204	9	11,11,12	0.57	0	13,15,17	1.15	1 (7%)

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
4	SIA	A	409	4	-	0/14/34/38	0/1/1/1
4	GAL	A	410	4	-	0/2/19/22	0/1/1/1
4	NAG	A	411	4	-	0/6/26/26	0/1/1/1
6	NAG	B	201	2,6	-	0/6/23/26	0/1/1/1
6	NAG	B	202	6	-	0/6/23/26	0/1/1/1
6	BMA	B	203	6	-	0/2/19/22	0/1/1/1
6	MAN	B	204	6	-	0/2/19/22	0/1/1/1
6	MAN	B	207	6	-	0/2/19/22	0/1/1/1
7	NAG	C	401	1,7	-	0/6/23/26	0/1/1/1
7	NAG	C	402	7	-	0/6/23/26	0/1/1/1
6	NAG	D	201	2,6	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	D	202	6	-	0/6/23/26	0/1/1/1
6	BMA	D	203	6	-	0/2/19/22	0/1/1/1
6	MAN	D	204	6	-	0/2/19/22	0/1/1/1
6	MAN	D	207	6	-	0/2/19/22	0/1/1/1
8	SIA	E	409	8	-	0/14/34/38	0/1/1/1
8	GAL	E	410	8	-	0/2/19/22	0/1/1/1
9	NAG	F	201	9,2	-	0/6/23/26	0/1/1/1
9	NAG	F	202	9	-	0/6/23/26	0/1/1/1
9	BMA	F	203	9	-	0/2/19/22	0/1/1/1
9	MAN	F	204	9	-	0/2/19/22	1/1/1/1

There are no bond length outliers.

All (29) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	201	NAG	O7-C7-N2	-3.65	114.90	121.92
7	C	401	NAG	O5-C1-C2	-3.22	106.99	111.47
6	B	201	NAG	O5-C1-C2	-2.72	107.69	111.47
6	B	203	BMA	C1-O5-C5	2.04	114.98	112.17
6	B	202	NAG	C4-C3-C2	2.18	114.22	111.02
6	B	203	BMA	C1-C2-C3	2.20	112.43	109.65
6	D	203	BMA	C1-O5-C5	2.34	115.39	112.17
6	D	207	MAN	C2-C3-C4	2.57	115.35	110.88
6	B	204	MAN	C2-C3-C4	2.62	115.44	110.88
6	D	203	BMA	C3-C4-C5	2.63	114.84	110.22
6	D	201	NAG	C1-O5-C5	2.68	115.86	112.17
6	D	207	MAN	C3-C4-C5	2.79	115.13	110.22
8	E	410	GAL	C1-C2-C3	2.87	113.29	109.65
7	C	402	NAG	C4-C3-C2	2.97	115.36	111.02
6	B	201	NAG	C1-O5-C5	3.00	116.30	112.17
9	F	204	MAN	C1-O5-C5	3.06	116.38	112.17
6	D	203	BMA	C1-C2-C3	3.08	113.56	109.65
7	C	401	NAG	C8-C7-N2	3.09	121.69	116.11
9	F	203	BMA	O5-C1-C2	3.38	116.08	110.79
6	B	207	MAN	C1-O5-C5	3.42	116.88	112.17
9	F	203	BMA	C1-C2-C3	3.43	113.99	109.65
9	F	202	NAG	C4-C3-C2	3.78	116.56	111.02
6	B	201	NAG	C2-N2-C7	3.80	128.48	122.94
6	D	207	MAN	C1-O5-C5	4.02	117.71	112.17
6	D	207	MAN	C1-C2-C3	4.04	114.77	109.65
6	B	201	NAG	C8-C7-N2	4.91	124.97	116.11
6	B	204	MAN	C1-C2-C3	5.54	116.68	109.65

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	F	203	BMA	C1-O5-C5	6.65	121.33	112.17
7	C	401	NAG	C2-N2-C7	7.17	133.41	122.94

There are no chirality outliers.

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	F	204	MAN	C1-C2-C3-C4-C5-O5

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	C	401	NAG	1	0
6	D	201	NAG	1	0
9	F	204	MAN	1	0

5.6 Ligand geometry

17 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	$\# Z > 2$
5	EDO	A	1326	-	3,3,3	0.46	0	2,2,2	0.37	0
5	EDO	A	1327	-	3,3,3	0.47	0	2,2,2	0.40	0
5	EDO	A	1328	-	3,3,3	0.41	0	2,2,2	0.35	0
3	NAG	A	401	1	14,14,15	0.47	0	15,19,21	1.12	1 (6%)
3	NAG	A	420	1	14,14,15	0.52	0	15,19,21	1.80	2 (13%)
5	EDO	B	1173	-	3,3,3	0.40	0	2,2,2	0.62	0
3	NAG	B	208	-	14,14,15	0.57	0	15,19,21	1.58	1 (6%)
3	NAG	B	211	-	14,14,15	0.51	0	15,19,21	0.81	0
5	EDO	C	1326	-	3,3,3	0.45	0	2,2,2	0.41	0
5	EDO	C	1327	-	3,3,3	0.47	0	2,2,2	0.23	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	NAG	C	420	1	14,14,15	0.58	0	15,19,21	1.00	2 (13%)
5	EDO	E	1326	-	3,3,3	0.48	0	2,2,2	0.18	0
5	EDO	E	1327	-	3,3,3	0.43	0	2,2,2	0.42	0
5	EDO	E	1328	-	3,3,3	0.55	0	2,2,2	0.29	0
3	NAG	E	401	1	14,14,15	0.58	0	15,19,21	0.99	1 (6%)
5	EDO	F	1173	-	3,3,3	0.47	0	2,2,2	0.32	0
5	EDO	F	1174	-	3,3,3	0.47	0	2,2,2	0.34	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	EDO	A	1326	-	-	0/1/1/1	0/0/0/0
5	EDO	A	1327	-	-	0/1/1/1	0/0/0/0
5	EDO	A	1328	-	-	0/1/1/1	0/0/0/0
3	NAG	A	401	1	-	0/6/23/26	0/1/1/1
3	NAG	A	420	1	1/1/5/7	0/6/23/26	0/1/1/1
5	EDO	B	1173	-	-	0/1/1/1	0/0/0/0
3	NAG	B	208	-	-	0/6/23/26	0/1/1/1
3	NAG	B	211	-	-	0/6/23/26	0/1/1/1
5	EDO	C	1326	-	-	0/1/1/1	0/0/0/0
5	EDO	C	1327	-	-	0/1/1/1	0/0/0/0
3	NAG	C	420	1	1/1/5/7	0/6/23/26	0/1/1/1
5	EDO	E	1326	-	-	0/1/1/1	0/0/0/0
5	EDO	E	1327	-	-	0/1/1/1	0/0/0/0
5	EDO	E	1328	-	-	0/1/1/1	0/0/0/0
3	NAG	E	401	1	-	0/6/23/26	0/1/1/1
5	EDO	F	1173	-	-	0/1/1/1	0/0/0/0
5	EDO	F	1174	-	-	0/1/1/1	0/0/0/0

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	420	NAG	O5-C1-C2	-5.95	103.20	111.47
3	C	420	NAG	O5-C1-C2	-2.21	108.40	111.47
3	A	420	NAG	C1-C2-N2	2.00	113.91	110.49
3	C	420	NAG	C4-C3-C2	2.05	114.02	111.02
3	E	401	NAG	C4-C3-C2	2.19	114.23	111.02

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	401	NAG	C1-O5-C5	2.39	115.46	112.17
3	B	208	NAG	C2-N2-C7	4.59	129.64	122.94

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	C	420	NAG	C1
3	A	420	NAG	C1

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1327	EDO	3	0
5	A	1328	EDO	1	0
3	B	211	NAG	1	0
5	E	1327	EDO	1	0
5	F	1173	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

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	318/318 (100%)	-0.24	0 100 100	27, 37, 52, 69	0
1	C	318/318 (100%)	-0.24	5 (1%) 72 70	24, 37, 57, 94	0
1	E	318/318 (100%)	0.08	8 (2%) 58 55	34, 51, 73, 95	0
2	B	172/172 (100%)	0.02	0 100 100	21, 45, 61, 70	0
2	D	172/172 (100%)	0.32	8 (4%) 32 30	23, 52, 76, 88	0
2	F	172/172 (100%)	0.46	10 (5%) 24 22	23, 58, 86, 94	0
All	All	1470/1470 (100%)	0.01	31 (2%) 64 61	21, 45, 74, 95	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	21	ALA	3.9
1	C	21	ALA	3.8
2	D	33	GLY	3.6
2	F	146	ASP	3.5
2	D	32	THR	3.2
2	D	19	ASP	3.0
2	F	29	ALA	2.9
1	E	193	ASP	2.9
2	D	29	ALA	2.9
2	F	30	GLN	2.8
1	E	128	SER	2.8
1	E	325	GLU	2.8
2	F	148	CYS	2.8
1	C	22	ASN	2.8
2	F	31	GLY	2.7
1	E	23	GLY	2.6
2	D	31	GLY	2.5
2	F	147	SER	2.5
1	C	128	SER	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	145	ASP	2.4
2	D	18	VAL	2.3
1	C	277	CYS	2.3
2	F	18	VAL	2.2
1	C	323	VAL	2.2
1	E	129	SER	2.2
2	D	2	LEU	2.1
2	F	32	THR	2.1
1	E	39	ALA	2.1
1	E	22	ASN	2.0
2	F	17	MET	2.0
2	D	36	ALA	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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
6	NAG	B	201	14/15	0.85	0.26	5.33	58,59,61,62	0
6	NAG	D	201	14/15	0.89	0.26	4.28	62,64,65,67	0
7	NAG	C	401	14/15	0.80	0.25	3.31	93,97,100,102	0
8	SIA	E	409	20/21	0.81	0.24	1.68	102,111,120,120	0
9	NAG	F	201	14/15	0.93	0.22	1.40	42,43,44,45	0
4	SIA	A	409	20/21	0.94	0.12	-0.38	52,56,59,60	0
8	GAL	E	410	11/12	0.72	0.26	-	122,132,137,139	0
9	BMA	F	203	11/12	0.76	0.26	-	54,57,59,61	0
6	NAG	B	202	14/15	0.90	0.19	-	62,64,66,68	0
4	GAL	A	410	11/12	0.94	0.14	-	61,67,70,71	0
6	MAN	D	207	11/12	0.80	0.23	-	81,84,90,91	0
7	NAG	C	402	14/15	0.77	0.36	-	103,107,113,113	0
6	BMA	B	203	11/12	0.85	0.19	-	69,73,76,77	0
6	MAN	B	204	11/12	0.79	0.29	-	74,79,86,88	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	NAG	A	411	15/15	0.84	0.24	-	73,79,87,88	0
6	MAN	D	204	11/12	0.83	0.32	-	84,87,89,91	0
6	NAG	D	202	14/15	0.83	0.20	-	68,71,75,75	0
9	NAG	F	202	14/15	0.91	0.17	-	46,48,50,52	0
6	BMA	D	203	11/12	0.76	0.17	-	78,82,85,86	0
9	MAN	F	204	11/12	0.82	0.32	-	63,65,66,67	0
6	MAN	B	207	11/12	0.87	0.32	-	72,74,78,79	0

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
5	EDO	E	1328	4/4	0.73	0.56	14.82	66,75,75,76	0
5	EDO	E	1327	4/4	0.96	0.51	12.66	47,47,48,48	0
5	EDO	A	1327	4/4	0.90	0.38	11.51	45,46,46,47	0
5	EDO	B	1173	4/4	0.94	0.31	7.60	43,43,44,44	0
5	EDO	A	1326	4/4	0.95	0.21	3.18	36,37,37,37	0
5	EDO	C	1326	4/4	0.94	0.28	3.13	38,40,41,45	0
5	EDO	F	1173	4/4	0.94	0.23	3.07	53,55,55,58	0
5	EDO	E	1326	4/4	0.94	0.19	2.06	50,51,52,52	0
5	EDO	A	1328	4/4	0.95	0.23	1.94	53,54,55,58	0
3	NAG	B	211	14/15	0.79	0.27	1.70	80,90,103,104	0
3	NAG	A	401	14/15	0.80	0.20	1.33	87,92,98,100	0
5	EDO	C	1327	4/4	0.97	0.12	-0.90	40,41,42,42	0
3	NAG	C	420	14/15	0.82	0.21	-	148,164,180,184	0
5	EDO	F	1174	4/4	0.91	0.13	-	53,54,54,55	0
3	NAG	A	420	14/15	0.79	0.25	-	82,91,97,102	0
3	NAG	B	208	14/15	0.73	0.42	-	81,92,102,102	0
3	NAG	E	401	14/15	0.73	0.34	-	156,165,171,179	0

6.5 Other polymers [i](#)

There are no such residues in this entry.