Full wwPDB X-ray Structure Validation Report (i)

Aug 29, 2023 – 04:13 pm BST

PDB ID	:	8P0O
Title	:	Crystal structure of AaNGT complexed to UDP-Gal
Authors	:	Pinello, B.; Macias-Leon, J.; Rovira, C.; Hurtado-Guerrero, R.
Deposited on	:	2023-05-10
Resolution	:	1.76 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 (i) 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 (i)) were used in the production of this report:

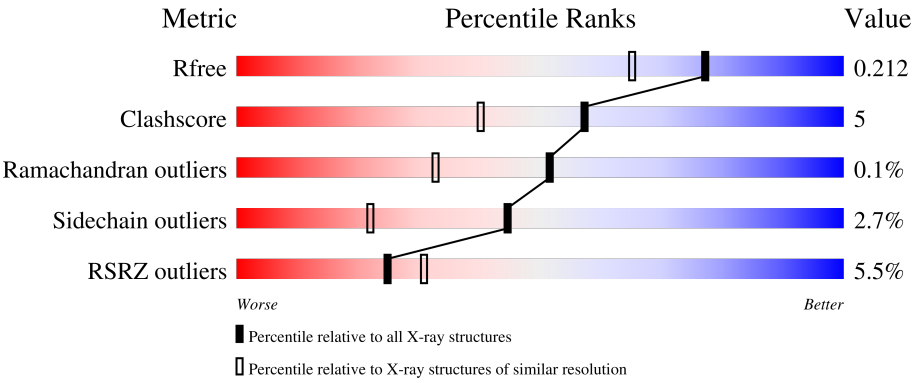
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.76 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	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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 >=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 <=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	622	4% 89% 9%
1	B	622	7% 91% 7%

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	801	-	-	X	-
2	EDO	B	801	-	-	X	-
2	EDO	B	809	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 10915 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.

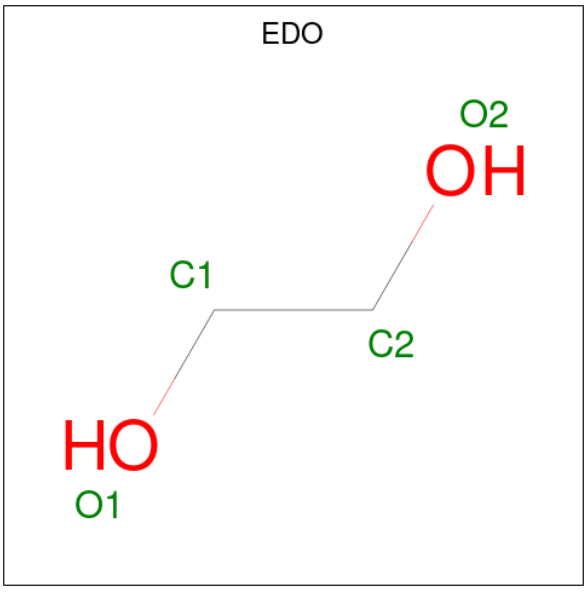
- Molecule 1 is a protein called Adhesin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	622	Total	C	N	O	S	0	0	0
			4983	3181	849	924	29			
1	B	621	Total	C	N	O	S	0	0	0
			4975	3175	848	923	29			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP A0A3M6PNT1
A	3	ARG	GLU	conflict	UNP A0A3M6PNT1
B	0	GLY	-	expression tag	UNP A0A3M6PNT1
B	3	ARG	GLU	conflict	UNP A0A3M6PNT1

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



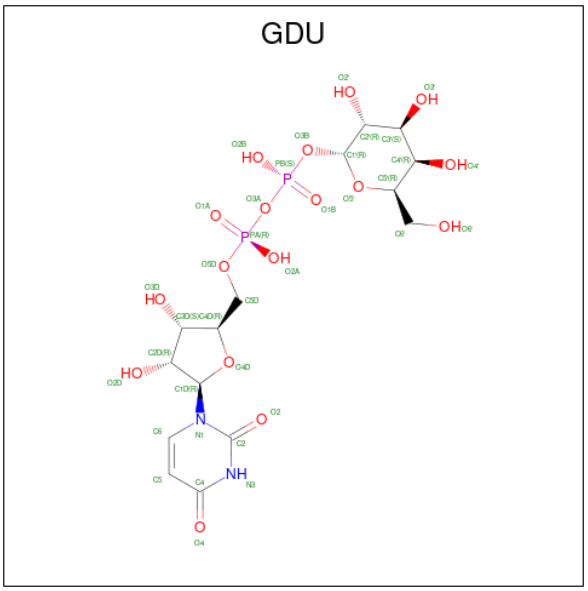
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0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

Continued on next page...

Continued from previous page...

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

- Molecule 3 is GALACTOSE-URIDINE-5'-DIPHOSPHATE (three-letter code: GDU) (formula: $C_{15}H_{24}N_2O_{17}P_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O P 36 15 2 17 2	0	0
3	B	1	Total C N O P 36 15 2 17 2	0	0

- Molecule 4 is water.

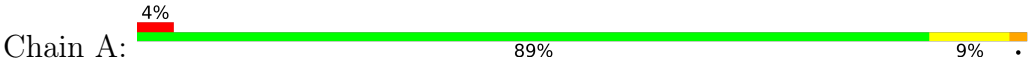
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	441	Total O 441 441	0	0
4	B	340	Total O 340 340	0	0

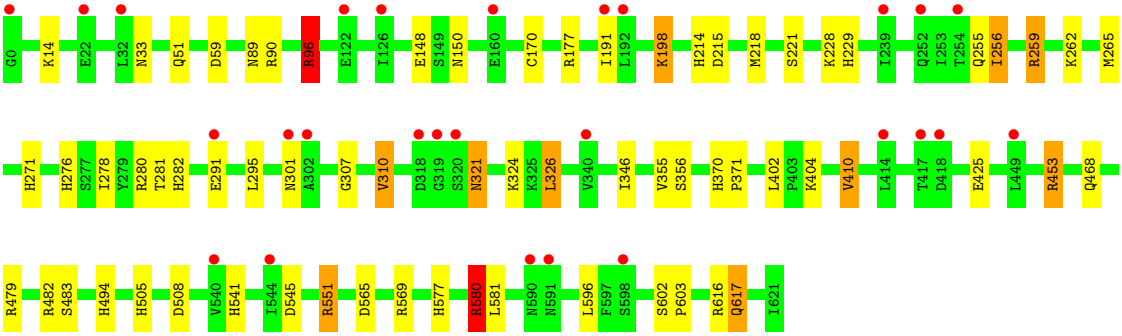
3 Residue-property plots (i)

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: Adhesin

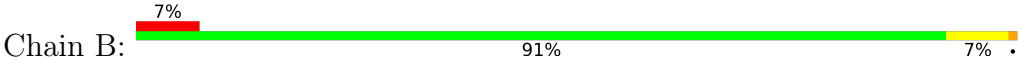
Chain A:

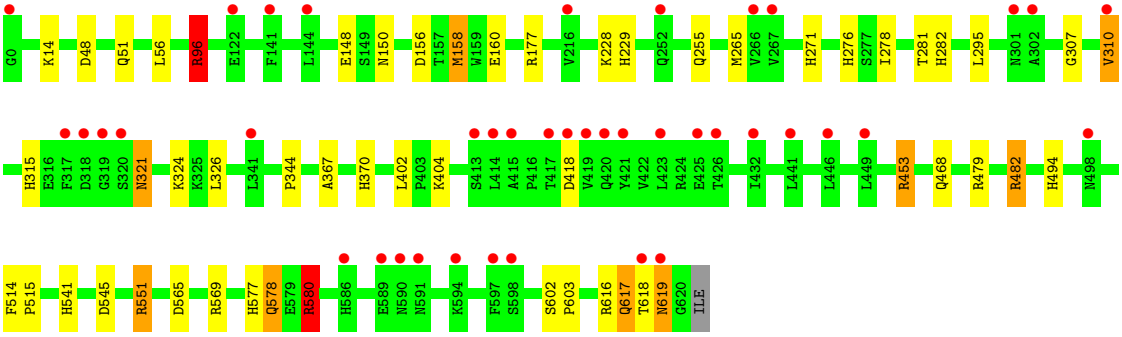




• Molecule 1: Adhesin

Chain B:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	47.34Å 113.33Å 260.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.76 20.00 – 1.76	Depositor EDS
% Data completeness (in resolution range)	99.9 (20.00-1.76) 100.0 (20.00-1.76)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.33 (at 1.76Å)	Xtriage
Refinement program	REFMAC 5.8.0411	Depositor
R, R_{free}	0.174 , 0.207 0.184 , 0.212	Depositor DCC
R_{free} test set	5584 reflections (3.99%)	wwPDB-VP
Wilson B-factor (Å^2)	35.1	Xtriage
Anisotropy	0.057	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 42.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	10915	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: EDO, GDU

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	0/5104	0.81	11/6927 (0.2%)
1	B	0.43	0/5096	0.76	9/6916 (0.1%)
All	All	0.44	0/10200	0.78	20/13843 (0.1%)

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	A	0	4
1	B	0	4
All	All	0	8

There are no bond length outliers.

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	479	ARG	NE-CZ-NH2	-14.00	113.30	120.30
1	A	453	ARG	NE-CZ-NH2	-12.50	114.05	120.30
1	A	580	ARG	NE-CZ-NH2	-11.63	114.48	120.30
1	B	479	ARG	NE-CZ-NH2	-11.24	114.68	120.30
1	B	580	ARG	NE-CZ-NH2	-10.26	115.17	120.30
1	A	551	ARG	NE-CZ-NH1	9.64	125.12	120.30
1	A	580	ARG	NE-CZ-NH1	9.06	124.83	120.30
1	B	96	ARG	NE-CZ-NH2	-8.87	115.86	120.30
1	B	158	MET	CG-SD-CE	-8.85	86.04	100.20
1	B	551	ARG	NE-CZ-NH1	8.84	124.72	120.30
1	A	96	ARG	NE-CZ-NH2	-8.77	115.91	120.30
1	B	551	ARG	NE-CZ-NH2	-8.46	116.07	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	453	ARG	NE-CZ-NH1	7.85	124.22	120.30
1	B	453	ARG	NE-CZ-NH2	-7.29	116.65	120.30
1	A	551	ARG	NE-CZ-NH2	-6.91	116.84	120.30
1	B	453	ARG	NE-CZ-NH1	5.96	123.28	120.30
1	A	617	GLN	CB-CA-C	5.80	121.99	110.40
1	B	580	ARG	NE-CZ-NH1	5.72	123.16	120.30
1	A	301	ASN	CB-CA-C	-5.56	99.29	110.40
1	A	479	ARG	NE-CZ-NH1	5.18	122.89	120.30

There are no chirality outliers.

All (8) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	453	ARG	Sidechain
1	A	551	ARG	Sidechain
1	A	580	ARG	Sidechain
1	A	96	ARG	Sidechain
1	B	482	ARG	Sidechain
1	B	551	ARG	Sidechain
1	B	580	ARG	Sidechain
1	B	96	ARG	Sidechain

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	4983	0	4903	52	0
1	B	4975	0	4892	44	0
2	A	52	0	78	8	0
2	B	52	0	78	6	0
3	A	36	0	22	0	0
3	B	36	0	22	0	0
4	A	441	0	0	16	0
4	B	340	0	0	11	0
All	All	10915	0	9995	98	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 5.

All (98) 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)
2:A:801:EDO:H22	4:A:1173:HOH:O	1.35	1.26
1:B:468:GLN:H	1:B:494:HIS:HD2	1.13	0.95
1:A:468:GLN:H	1:A:494:HIS:HD2	1.14	0.94
1:B:617:GLN:C	1:B:617:GLN:HE21	1.75	0.88
2:B:801:EDO:C2	4:B:1081:HOH:O	2.32	0.77
1:B:14:LYS:HG3	4:B:1103:HOH:O	1.85	0.74
1:A:214:HIS:NE2	2:A:801:EDO:H21	2.02	0.74
1:B:271:HIS:O	1:B:278:ILE:HG21	1.89	0.72
1:A:468:GLN:H	1:A:494:HIS:CD2	2.05	0.71
1:A:577:HIS:HD2	1:A:580:ARG:HH12	1.37	0.70
1:B:577:HIS:HD2	1:B:580:ARG:HH12	1.39	0.69
1:A:271:HIS:O	1:A:278:ILE:HG21	1.92	0.69
2:A:802:EDO:H12	4:A:1218:HOH:O	1.93	0.68
2:B:801:EDO:H22	4:B:1081:HOH:O	1.93	0.66
2:A:801:EDO:H21	4:A:1190:HOH:O	1.95	0.66
1:A:89:ASN:ND2	4:A:901:HOH:O	2.21	0.65
1:B:156:ASP:O	1:B:160:GLU:HG3	1.97	0.65
2:A:802:EDO:C1	4:A:1218:HOH:O	2.45	0.64
1:B:578:GLN:H	1:B:578:GLN:CD	2.01	0.64
1:A:150:ASN:HD21	1:B:51:GLN:HE21	1.46	0.64
1:A:265:MET:HE2	1:A:295:LEU:HG	1.79	0.64
1:B:580:ARG:HD3	4:B:957:HOH:O	1.98	0.63
1:B:468:GLN:H	1:B:494:HIS:CD2	2.05	0.63
1:B:370:HIS:HD2	4:B:1210:HOH:O	1.81	0.62
1:A:89:ASN:HB2	4:A:1244:HOH:O	1.98	0.62
1:B:229:HIS:HD2	4:B:941:HOH:O	1.83	0.62
1:B:321:ASN:HD22	1:B:324:LYS:H	1.44	0.62
1:A:215:ASP:OD1	2:A:801:EDO:H11	2.00	0.61
1:A:259:ARG:HB2	1:A:259:ARG:HH11	1.65	0.61
1:A:214:HIS:CE1	2:A:801:EDO:H21	2.37	0.60
1:A:281:THR:OG1	1:A:282:HIS:HD2	1.84	0.60
1:B:281:THR:OG1	1:B:282:HIS:HD2	1.85	0.60
1:B:617:GLN:C	1:B:617:GLN:NE2	2.53	0.59
1:A:150:ASN:ND2	1:B:51:GLN:HE21	2.01	0.59
1:B:96:ARG:HH12	2:B:801:EDO:H12	1.67	0.59
1:A:51:GLN:HE21	1:B:150:ASN:HD21	1.52	0.58
2:A:801:EDO:C2	4:A:1190:HOH:O	2.50	0.57
1:B:617:GLN:HE21	1:B:618:THR:N	2.03	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:801:EDO:H21	4:B:1081:HOH:O	1.97	0.56
1:A:218:MET:HE2	1:A:370:HIS:ND1	2.21	0.55
1:A:602:SER:N	1:A:603:PRO:CD	2.69	0.55
1:B:177:ARG:HH21	1:B:468:GLN:NE2	2.04	0.55
1:A:177:ARG:HH21	1:A:468:GLN:NE2	2.05	0.54
1:B:602:SER:N	1:B:603:PRO:CD	2.70	0.54
1:A:262:LYS:HE2	1:A:291:GLU:O	2.07	0.54
1:A:280:ARG:HA	1:A:410:VAL:HG13	1.89	0.54
1:A:565:ASP:O	1:A:569:ARG:HG2	2.08	0.54
1:B:618:THR:O	1:B:619:ASN:HB2	2.09	0.53
1:A:90:ARG:HD2	4:A:1195:HOH:O	2.09	0.53
1:B:580:ARG:CD	4:B:957:HOH:O	2.55	0.53
1:B:276:HIS:ND1	1:B:278:ILE:HG22	2.25	0.52
1:A:218:MET:HE2	1:A:370:HIS:CE1	2.45	0.52
1:A:198:LYS:N	1:A:198:LYS:HD2	2.25	0.52
1:A:321:ASN:HD22	1:A:324:LYS:H	1.56	0.52
1:A:259:ARG:HB2	1:A:259:ARG:NH1	2.25	0.51
1:A:265:MET:HE1	1:A:295:LEU:HD21	1.91	0.51
1:A:326:LEU:HD22	1:A:355:VAL:HG22	1.92	0.51
1:B:96:ARG:HH12	2:B:801:EDO:C1	2.25	0.50
1:B:565:ASP:O	1:B:569:ARG:HG2	2.12	0.49
1:A:33:ASN:ND2	4:A:910:HOH:O	2.45	0.49
1:B:265:MET:HE2	1:B:295:LEU:HG	1.95	0.49
1:B:315:HIS:ND1	4:B:901:HOH:O	2.35	0.48
1:B:228:LYS:NZ	1:B:541:HIS:HE1	2.11	0.48
1:A:228:LYS:NZ	1:A:541:HIS:HE1	2.10	0.48
1:A:59:ASP:OD2	1:B:482:ARG:NH1	2.47	0.48
1:A:51:GLN:HE21	1:B:150:ASN:ND2	2.11	0.48
1:B:265:MET:HE1	1:B:295:LEU:HD21	1.95	0.47
1:A:14:LYS:HG3	4:A:1237:HOH:O	2.14	0.47
1:B:271:HIS:O	1:B:278:ILE:CG2	2.61	0.47
1:A:276:HIS:HB2	4:A:1305:HOH:O	2.14	0.47
1:A:580:ARG:CD	4:A:1042:HOH:O	2.61	0.47
1:A:170:CYS:SG	1:A:191:ILE:HD12	2.55	0.46
1:A:425:GLU:HA	1:A:581:LEU:HD21	1.97	0.46
1:B:617:GLN:HG2	1:B:617:GLN:O	2.15	0.46
1:A:580:ARG:HD3	4:A:1042:HOH:O	2.16	0.46
1:B:482:ARG:NH2	4:B:909:HOH:O	2.48	0.46
1:A:276:HIS:ND1	1:A:278:ILE:HG22	2.30	0.46
1:A:218:MET:CE	1:A:370:HIS:ND1	2.79	0.46
1:B:48:ASP:OD2	2:B:804:EDO:H21	2.16	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:307:GLY:O	1:B:310:VAL:HG13	2.16	0.45
1:A:229:HIS:HD2	4:A:956:HOH:O	2.00	0.45
1:A:482:ARG:HD3	4:A:1247:HOH:O	2.18	0.44
1:A:321:ASN:ND2	1:A:324:LYS:H	2.15	0.44
1:B:321:ASN:ND2	1:B:324:LYS:H	2.14	0.44
1:B:578:GLN:CD	1:B:578:GLN:N	2.70	0.44
1:B:418:ASP:OD1	1:B:418:ASP:N	2.51	0.43
1:A:508:ASP:O	1:A:580:ARG:HD3	2.19	0.43
1:B:276:HIS:HB2	4:B:1226:HOH:O	2.19	0.43
1:A:221:SER:OG	1:A:229:HIS:HE1	2.01	0.43
1:A:577:HIS:CD2	1:A:580:ARG:HH12	2.27	0.43
1:A:255:GLN:C	1:A:256:ILE:HD12	2.39	0.42
1:B:514:PHE:HB2	1:B:515:PRO:HA	2.01	0.42
1:A:483:SER:HA	1:B:56:LEU:HD21	2.00	0.42
1:B:344:PRO:HA	1:B:367:ALA:HB3	2.00	0.42
1:A:307:GLY:O	1:A:310:VAL:HG13	2.18	0.42
1:A:577:HIS:HE1	4:A:1280:HOH:O	2.02	0.41
1:A:371:PRO:O	1:A:541:HIS:HD2	2.03	0.41
1:A:346:ILE:CD1	1:A:356:SER:HB3	2.52	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	620/622 (100%)	609 (98%)	11 (2%)	0	100	100
1	B	619/622 (100%)	606 (98%)	12 (2%)	1 (0%)	47	29
All	All	1239/1244 (100%)	1215 (98%)	23 (2%)	1 (0%)	51	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	619	ASN

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	551/551 (100%)	535 (97%)	16 (3%)	42 19
1	B	550/551 (100%)	536 (98%)	14 (2%)	47 25
All	All	1101/1102 (100%)	1071 (97%)	30 (3%)	44 22

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	96	ARG
1	A	148	GLU
1	A	198	LYS
1	A	256	ILE
1	A	259	ARG
1	A	310	VAL
1	A	321	ASN
1	A	326	LEU
1	A	402	LEU
1	A	404	LYS
1	A	410	VAL
1	A	505	HIS
1	A	545	ASP
1	A	596	LEU
1	A	616	ARG
1	A	617	GLN
1	B	96	ARG
1	B	148	GLU
1	B	158	MET
1	B	255	GLN
1	B	310	VAL
1	B	321	ASN
1	B	326	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	402	LEU
1	B	404	LYS
1	B	453	ARG
1	B	545	ASP
1	B	578	GLN
1	B	616	ARG
1	B	617	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (32) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	33	ASN
1	A	134	ASN
1	A	150	ASN
1	A	219	HIS
1	A	229	HIS
1	A	282	HIS
1	A	292	GLN
1	A	321	ASN
1	A	357	ASN
1	A	468	GLN
1	A	494	HIS
1	A	506	ASN
1	A	512	ASN
1	A	541	HIS
1	A	576	ASN
1	A	577	HIS
1	B	33	ASN
1	B	134	ASN
1	B	150	ASN
1	B	229	HIS
1	B	282	HIS
1	B	321	ASN
1	B	357	ASN
1	B	370	HIS
1	B	431	ASN
1	B	468	GLN
1	B	494	HIS
1	B	512	ASN
1	B	541	HIS
1	B	576	ASN
1	B	577	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	617	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

There are no monosaccharides in this entry.

5.6 Ligand geometry

28 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	EDO	B	812	-	3,3,3	0.20	0	2,2,2	0.34	0
2	EDO	A	802	-	3,3,3	0.28	0	2,2,2	0.42	0
2	EDO	A	806	-	3,3,3	0.08	0	2,2,2	0.11	0
2	EDO	A	809	-	3,3,3	0.07	0	2,2,2	0.18	0
3	GDU	B	814	-	35,38,38	0.41	0	53,58,58	0.74	0
3	GDU	A	814	-	35,38,38	0.45	0	53,58,58	0.73	0
2	EDO	B	801	-	3,3,3	0.25	0	2,2,2	0.30	0
2	EDO	A	813	-	3,3,3	0.10	0	2,2,2	0.13	0
2	EDO	B	803	-	3,3,3	0.08	0	2,2,2	0.31	0
2	EDO	B	804	-	3,3,3	0.25	0	2,2,2	0.17	0
2	EDO	B	806	-	3,3,3	0.24	0	2,2,2	0.18	0
2	EDO	A	805	-	3,3,3	0.21	0	2,2,2	0.35	0
2	EDO	B	802	-	3,3,3	0.11	0	2,2,2	0.28	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	EDO	A	803	-	3,3,3	0.14	0	2,2,2	0.04	0
2	EDO	B	809	-	3,3,3	0.18	0	2,2,2	0.18	0
2	EDO	B	810	-	3,3,3	0.34	0	2,2,2	0.46	0
2	EDO	B	808	-	3,3,3	0.18	0	2,2,2	0.25	0
2	EDO	A	801	-	3,3,3	0.74	0	2,2,2	0.78	0
2	EDO	A	807	-	3,3,3	0.16	0	2,2,2	0.21	0
2	EDO	B	805	-	3,3,3	0.11	0	2,2,2	0.38	0
2	EDO	A	811	-	3,3,3	0.08	0	2,2,2	0.13	0
2	EDO	B	811	-	3,3,3	0.15	0	2,2,2	0.35	0
2	EDO	A	810	-	3,3,3	0.18	0	2,2,2	0.24	0
2	EDO	B	813	-	3,3,3	0.14	0	2,2,2	0.23	0
2	EDO	A	808	-	3,3,3	0.40	0	2,2,2	0.60	0
2	EDO	B	807	-	3,3,3	0.31	0	2,2,2	0.61	0
2	EDO	A	812	-	3,3,3	0.12	0	2,2,2	0.21	0
2	EDO	A	804	-	3,3,3	0.21	0	2,2,2	0.42	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EDO	B	812	-	-	1/1/1/1	-
2	EDO	A	802	-	-	1/1/1/1	-
2	EDO	A	806	-	-	0/1/1/1	-
2	EDO	A	809	-	-	1/1/1/1	-
3	GDU	B	814	-	-	3/23/59/59	0/3/3/3
3	GDU	A	814	-	-	5/23/59/59	0/3/3/3
2	EDO	B	801	-	-	0/1/1/1	-
2	EDO	A	813	-	-	1/1/1/1	-
2	EDO	B	803	-	-	1/1/1/1	-
2	EDO	B	804	-	-	1/1/1/1	-
2	EDO	B	806	-	-	1/1/1/1	-
2	EDO	A	805	-	-	0/1/1/1	-
2	EDO	B	802	-	-	1/1/1/1	-
2	EDO	A	803	-	-	0/1/1/1	-
2	EDO	B	809	-	-	1/1/1/1	-
2	EDO	B	810	-	-	1/1/1/1	-
2	EDO	B	808	-	-	0/1/1/1	-
2	EDO	A	801	-	-	1/1/1/1	-
2	EDO	A	807	-	-	1/1/1/1	-
2	EDO	B	805	-	-	1/1/1/1	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EDO	A	811	-	-	0/1/1/1	-
2	EDO	B	811	-	-	1/1/1/1	-
2	EDO	A	810	-	-	0/1/1/1	-
2	EDO	B	813	-	-	0/1/1/1	-
2	EDO	A	808	-	-	1/1/1/1	-
2	EDO	B	807	-	-	1/1/1/1	-
2	EDO	A	812	-	-	1/1/1/1	-
2	EDO	A	804	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (25) torsion outliers are listed below:

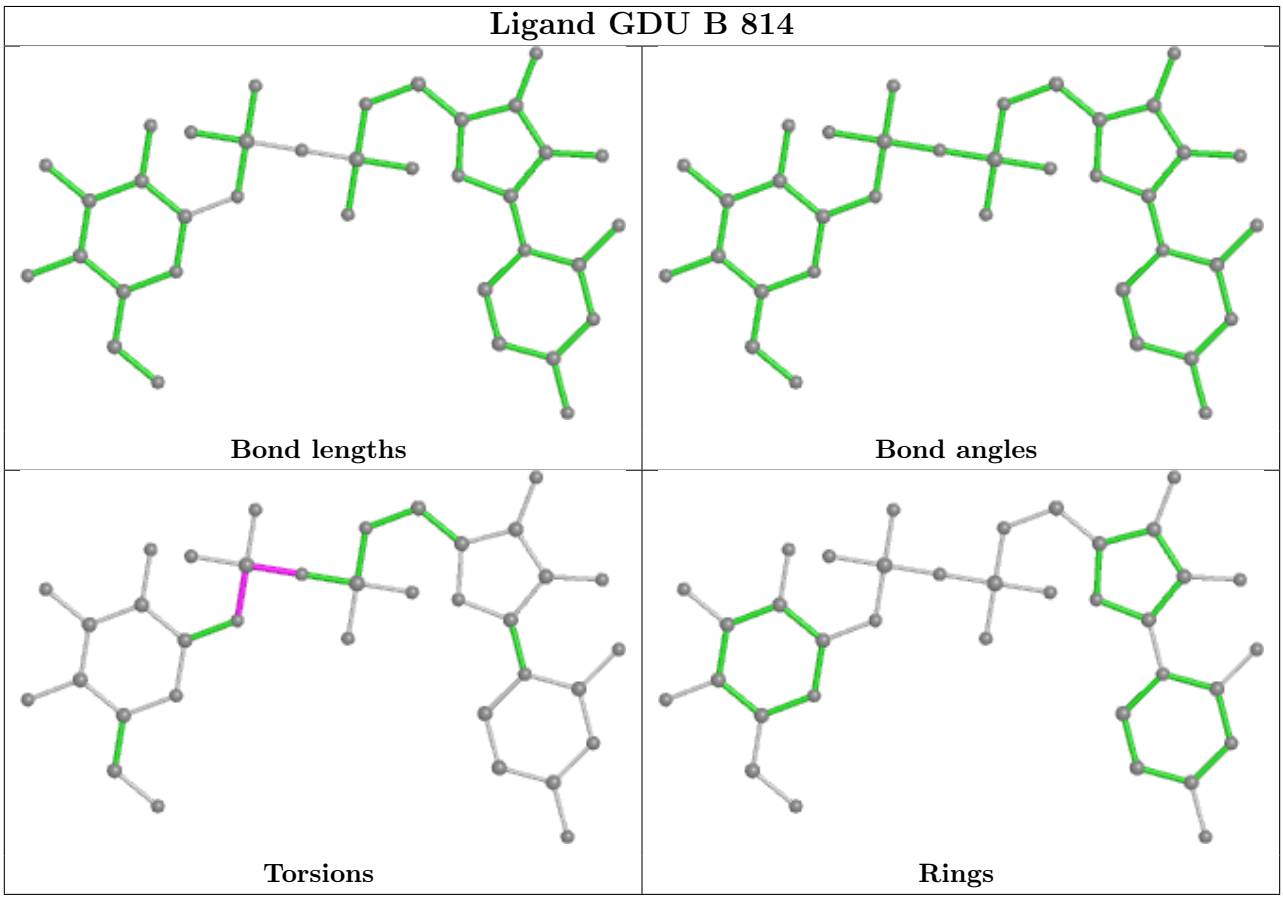
Mol	Chain	Res	Type	Atoms
2	B	805	EDO	O1-C1-C2-O2
3	A	814	GDU	PA-O3A-PB-O3B
3	B	814	GDU	PA-O3A-PB-O3B
3	B	814	GDU	C1'-O3B-PB-O3A
3	A	814	GDU	C1'-O3B-PB-O3A
2	A	801	EDO	O1-C1-C2-O2
2	A	809	EDO	O1-C1-C2-O2
2	B	803	EDO	O1-C1-C2-O2
2	B	810	EDO	O1-C1-C2-O2
2	B	812	EDO	O1-C1-C2-O2
2	B	802	EDO	O1-C1-C2-O2
3	B	814	GDU	C1'-O3B-PB-O2B
2	B	804	EDO	O1-C1-C2-O2
2	B	806	EDO	O1-C1-C2-O2
3	A	814	GDU	PB-O3A-PA-O5D
2	B	809	EDO	O1-C1-C2-O2
2	A	802	EDO	O1-C1-C2-O2
2	A	807	EDO	O1-C1-C2-O2
2	B	807	EDO	O1-C1-C2-O2
3	A	814	GDU	C1'-O3B-PB-O2B
2	A	812	EDO	O1-C1-C2-O2
2	A	808	EDO	O1-C1-C2-O2
2	A	813	EDO	O1-C1-C2-O2
2	B	811	EDO	O1-C1-C2-O2
3	A	814	GDU	C1'-O3B-PB-O1B

There are no ring outliers.

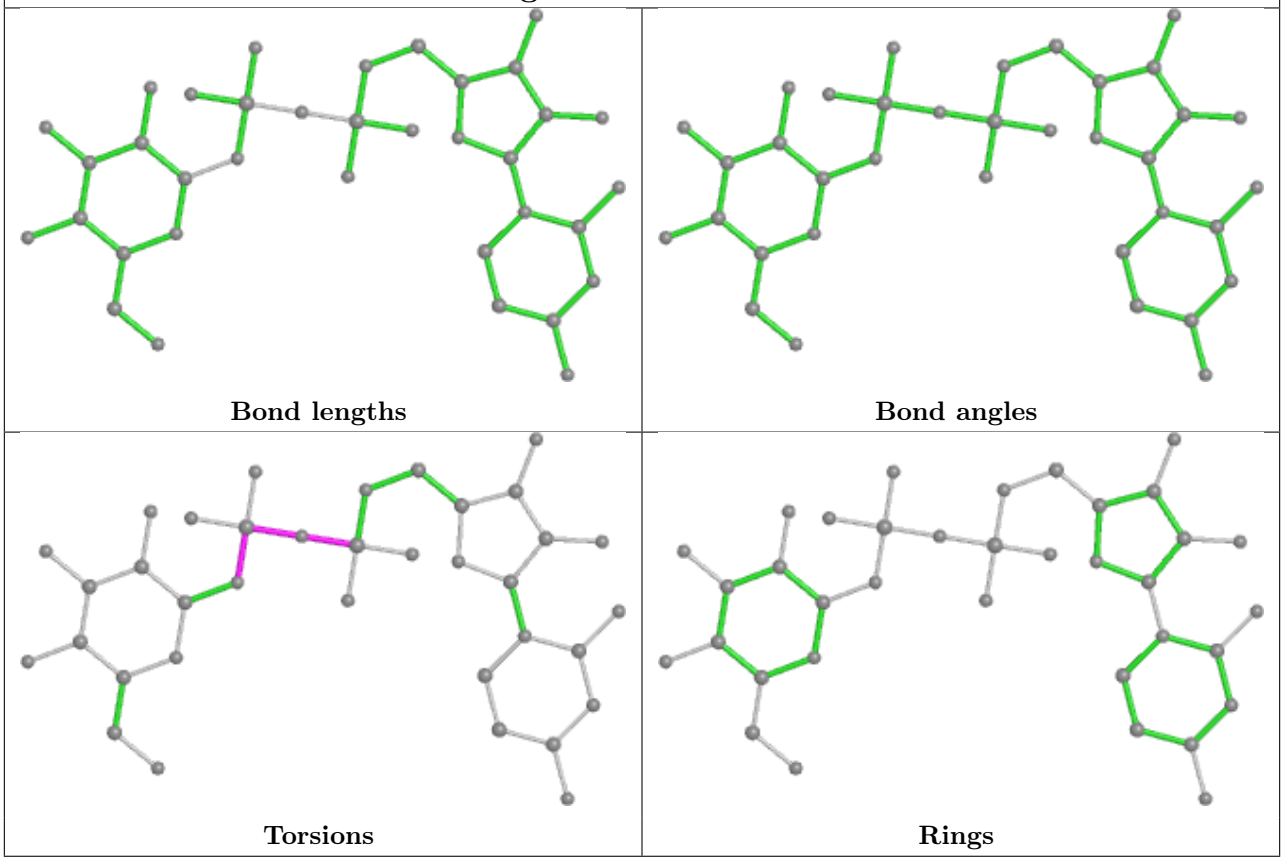
4 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	802	EDO	2	0
2	B	801	EDO	5	0
2	B	804	EDO	1	0
2	A	801	EDO	6	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



Ligand GDU A 814



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} 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	622/622 (100%)	0.20	27 (4%)	35	41	23, 35, 56, 86	0
1	B	621/622 (99%)	0.37	41 (6%)	18	24	23, 41, 66, 108	0
All	All	1243/1244 (99%)	0.28	68 (5%)	25	31	23, 38, 62, 108	0

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	318	ASP	8.6
1	B	619	ASN	7.1
1	B	414	LEU	6.6
1	B	418	ASP	5.4
1	B	319	GLY	4.9
1	B	417	THR	4.8
1	B	594	LYS	4.5
1	A	417	THR	4.0
1	B	618	THR	4.0
1	B	413	SER	3.9
1	A	318	ASP	3.9
1	B	302	ALA	3.8
1	B	426	THR	3.7
1	B	317	PHE	3.7
1	B	301	ASN	3.5
1	B	586	HIS	3.4
1	B	591	ASN	3.4
1	B	598	SER	3.4
1	B	266	VAL	3.4
1	B	590	ASN	3.3
1	A	122	GLU	3.3
1	B	425	GLU	3.0
1	A	590	ASN	3.0
1	B	320	SER	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	589	GLU	2.9
1	B	597	PHE	2.8
1	B	415	ALA	2.8
1	B	420	GLN	2.7
1	A	191	ILE	2.6
1	B	310	VAL	2.6
1	A	319	GLY	2.6
1	B	341	LEU	2.6
1	B	419	VAL	2.5
1	B	421	TYR	2.5
1	A	449	LEU	2.5
1	B	498	ASN	2.4
1	B	0	GLY	2.4
1	A	414	LEU	2.4
1	B	449	LEU	2.4
1	A	252	GLN	2.4
1	A	32	LEU	2.3
1	A	598	SER	2.3
1	B	216	VAL	2.3
1	A	160	GLU	2.3
1	A	302	ALA	2.3
1	B	267	VAL	2.3
1	A	239	ILE	2.3
1	A	418	ASP	2.2
1	A	544	ILE	2.2
1	B	122	GLU	2.2
1	B	252	GLN	2.1
1	A	301	ASN	2.1
1	A	22	GLU	2.1
1	B	423	LEU	2.1
1	A	0	GLY	2.1
1	A	126	ILE	2.1
1	B	144	LEU	2.1
1	B	441	LEU	2.1
1	B	446	LEU	2.1
1	A	320	SER	2.1
1	A	291	GLU	2.0
1	A	591	ASN	2.0
1	A	192	LEU	2.0
1	B	432	ILE	2.0
1	A	340	VAL	2.0
1	A	540	VAL	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	141	PHE	2.0
1	A	254	THR	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^{th} 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	B	809	4/4	0.43	0.42	79,89,91,91	0
2	EDO	A	808	4/4	0.65	0.19	66,70,72,73	0
2	EDO	B	803	4/4	0.71	0.17	69,74,74,75	0
2	EDO	B	807	4/4	0.72	0.23	57,65,67,67	0
2	EDO	B	805	4/4	0.72	0.29	61,61,62,63	0
2	EDO	B	801	4/4	0.75	0.28	51,61,62,68	0
2	EDO	A	802	4/4	0.77	0.29	57,72,73,77	0
2	EDO	A	806	4/4	0.77	0.16	47,49,51,53	0
2	EDO	A	810	4/4	0.79	0.24	53,57,58,67	0
2	EDO	B	808	4/4	0.80	0.36	56,62,63,67	0
2	EDO	B	811	4/4	0.81	0.25	60,72,74,75	0
2	EDO	A	813	4/4	0.83	0.30	72,73,73,75	0
2	EDO	B	812	4/4	0.83	0.44	58,64,69,71	0
2	EDO	B	806	4/4	0.84	0.24	59,60,61,67	0
2	EDO	A	807	4/4	0.84	0.16	54,55,59,62	0
2	EDO	A	801	4/4	0.86	0.27	44,47,48,63	0
2	EDO	B	802	4/4	0.87	0.26	57,64,68,76	0
2	EDO	A	812	4/4	0.88	0.23	59,68,70,72	0
2	EDO	A	809	4/4	0.88	0.26	81,82,82,86	0
2	EDO	B	804	4/4	0.89	0.16	49,56,58,61	0
2	EDO	A	811	4/4	0.89	0.10	61,64,68,71	0

Continued on next page...

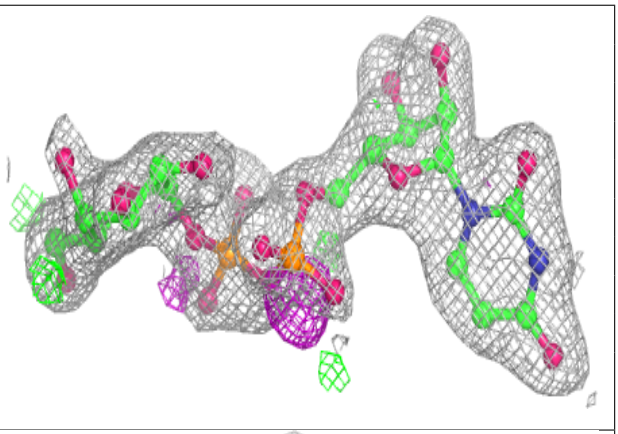
Continued from previous page...

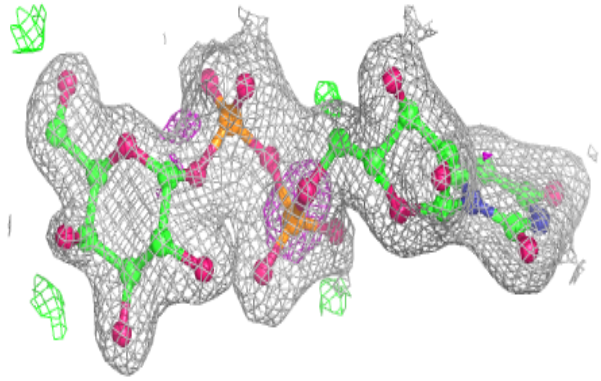
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	EDO	B	813	4/4	0.89	0.17	55,59,59,66	0
3	GDU	A	814	36/36	0.89	0.13	33,46,59,62	0
2	EDO	A	803	4/4	0.90	0.17	48,48,54,57	0
2	EDO	A	804	4/4	0.91	0.15	52,55,56,56	0
2	EDO	B	810	4/4	0.91	0.12	52,57,60,62	0
2	EDO	A	805	4/4	0.92	0.18	36,47,48,51	0
3	GDU	B	814	36/36	0.92	0.14	47,54,60,61	0

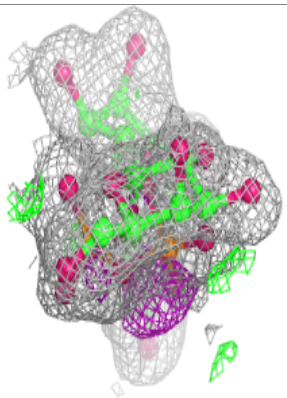
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around GDU A 814:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

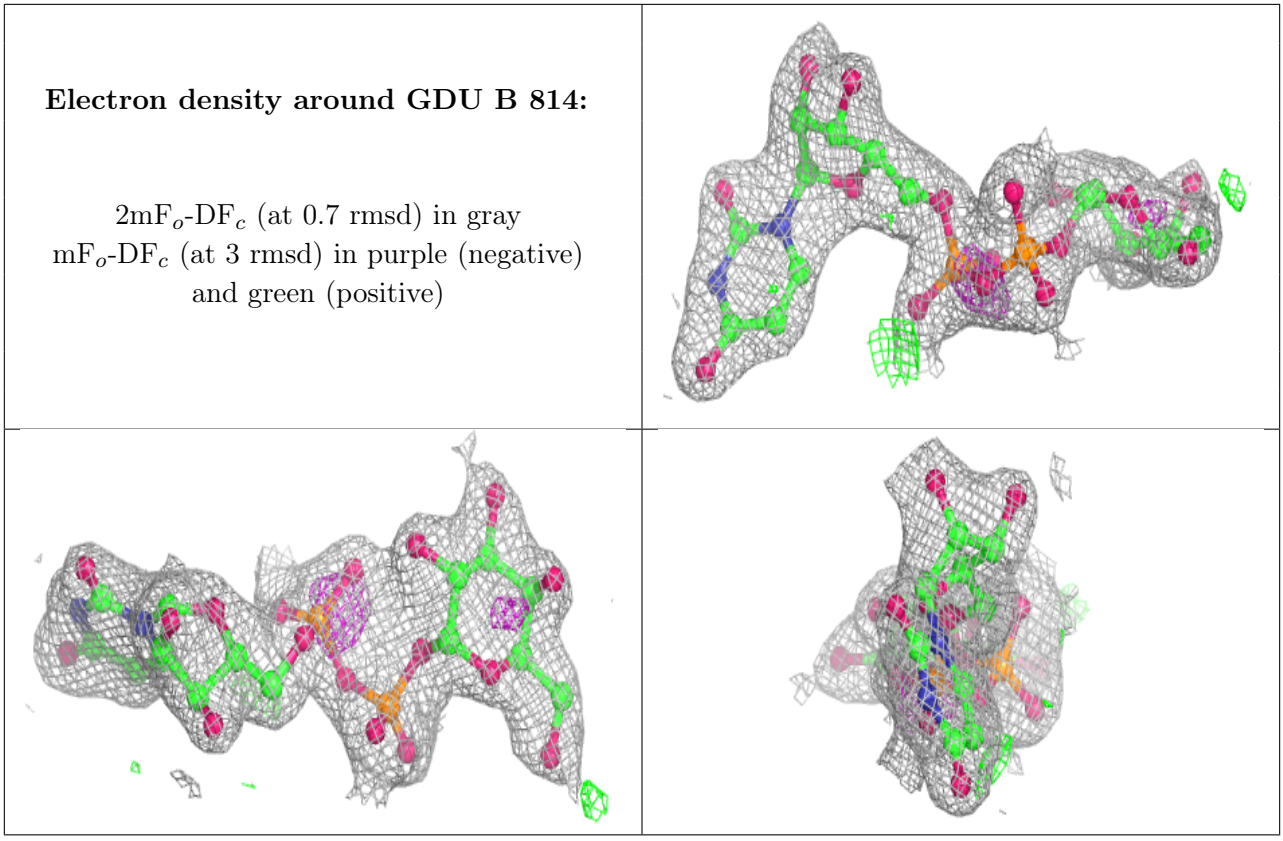






Electron density around GDU B 814:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.