



Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 12:54 pm BST

PDB ID : 1K4V
Title : 1.53 Å Crystal Structure of the Beta-Galactoside-alpha-1,3-galactosyltransferase in Complex with UDP
Authors : Boix, E.; Swaminathan, G.J.; Zhang, Y.; Natesh, R.; Brew, K.; Acharya, K.R.
Deposited on : 2001-10-09
Resolution : 1.53 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) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

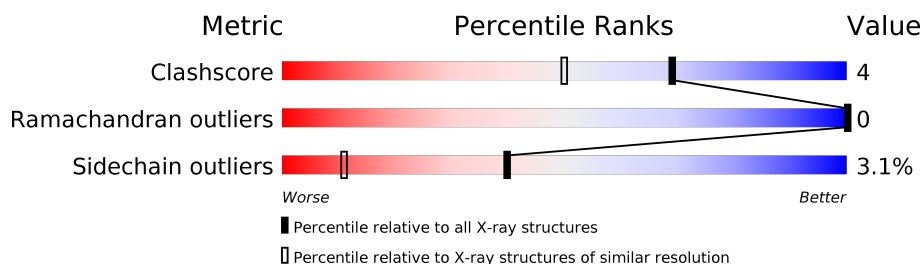
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.53 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))
Clashscore	141614	2634 (1.56-1.52)
Ramachandran outliers	138981	2580 (1.56-1.52)
Sidechain outliers	138945	2577 (1.56-1.52)

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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	289	
1	B	289	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	GOL	A	404	-	X	-	-
4	GOL	B	1404	-	X	-	-

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 5665 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.

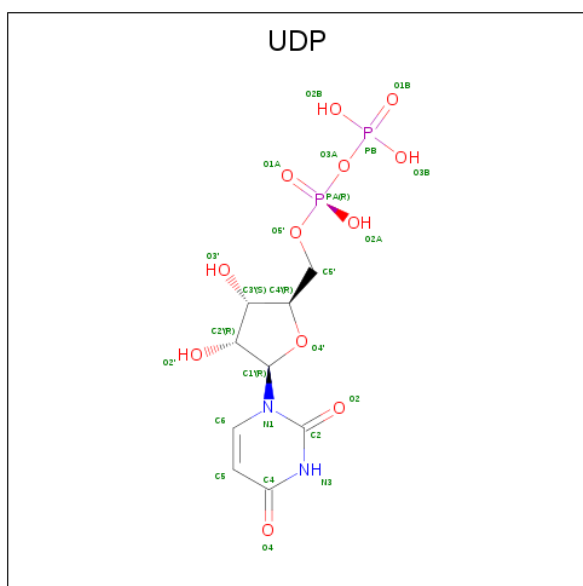
- Molecule 1 is a protein called N-ACETYLLACTOSAMINIDE ALPHA-1,3-GALACTOSYL TRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	287	Total	C	N	O	S	0	5	0
			2426	1592	396	426	12			
1	B	287	Total	C	N	O	S	0	4	0
			2418	1583	395	428	12			

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mn	0	0
			1	1		
2	A	1	Total	Mn	0	0
			1	1		

- Molecule 3 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: C₉H₁₄N₂O₁₂P₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
3	B	1	Total	C	N	O	P	0	0
			25	9	2	12	2		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is water.

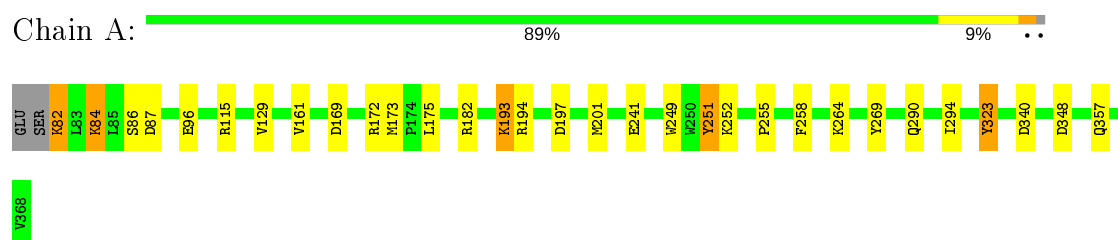
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	382	Total	O	0	3
			385	385		
5	B	365	Total	O	0	7
			372	372		

3 Residue-property plots [i](#)

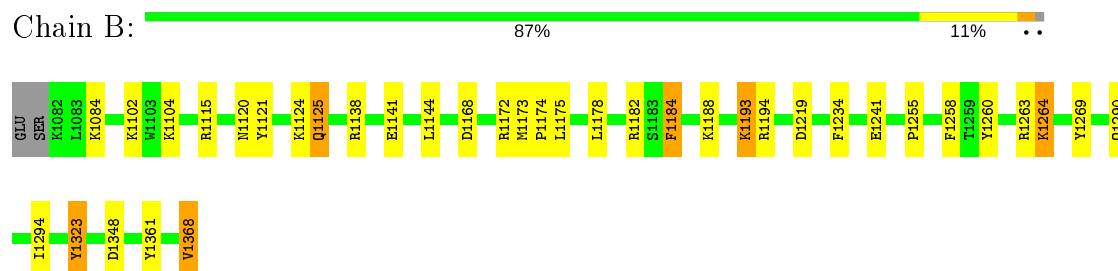
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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. 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.

Note EDS was not executed.

- Molecule 1: N-ACETYLLACTOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE



- Molecule 1: N-ACETYLLACTOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	45.02Å 94.14Å 94.38Å 90.00° 98.92° 90.00°	Depositor
Resolution (Å)	40.00 – 1.53	Depositor
% Data completeness (in resolution range)	97.2 (40.00-1.53)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.141 , 0.191	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5665	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, UDP, MN

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	0.51	0/2519	1.14	5/3413 (0.1%)
1	B	0.97	1/2505 (0.0%)	1.20	12/3392 (0.4%)
All	All	0.78	1/5024 (0.0%)	1.17	17/6805 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1368	VAL	C-OXT	41.82	2.02	1.23

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1263	ARG	NE-CZ-NH1	-8.20	116.20	120.30
1	B	1194	ARG	NE-CZ-NH1	8.05	124.33	120.30
1	B	1194	ARG	NE-CZ-NH2	-7.44	116.58	120.30
1	B	1115	ARG	NE-CZ-NH2	6.92	123.76	120.30
1	B	1361	TYR	CB-CG-CD2	6.27	124.76	121.00
1	B	1219	ASP	CB-CG-OD2	6.14	123.83	118.30
1	B	1323	TYR	CB-CG-CD1	5.89	124.53	121.00
1	A	115	ARG	CD-NE-CZ	5.85	131.79	123.60
1	A	194	ARG	NE-CZ-NH2	5.77	123.19	120.30
1	B	1184[A]	PHE	CB-CG-CD2	-5.44	116.99	120.80
1	B	1184[B]	PHE	CB-CG-CD2	-5.44	116.99	120.80
1	B	1260	TYR	CB-CG-CD2	5.35	124.21	121.00
1	B	1361	TYR	CB-CG-CD1	-5.25	117.85	121.00
1	A	323	TYR	CB-CG-CD2	-5.22	117.86	121.00
1	B	1348	ASP	CB-CG-OD2	-5.15	113.66	118.30
1	A	251	TYR	CB-CG-CD2	5.14	124.08	121.00
1	A	348	ASP	CB-CG-OD2	-5.03	113.77	118.30

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	2426	0	2384	17	0
1	B	2418	0	2375	20	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	25	0	11	0	0
3	B	25	0	11	0	0
4	A	6	0	5	0	0
4	B	6	0	5	0	0
5	A	385	0	0	3	0
5	B	372	0	0	4	0
All	All	5665	0	4791	36	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 (36) 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:B:1368:VAL:C	1:B:1368:VAL:OXT	2.02	0.96
1:A:193:LYS:HD2	5:A:2412:HOH:O	1.85	0.78
1:A:82:LYS:O	1:A:82:LYS:HD3	1.89	0.73
1:B:1141:GLU:OE2	1:B:1172:ARG:HD2	1.96	0.66
1:B:1193:LYS:HE2	5:B:2190:HOH:O	1.98	0.63
1:A:255:PRO:HA	1:A:258:PHE:CD1	2.40	0.56
1:B:1368:VAL:OXT	1:B:1368:VAL:O	2.25	0.53
1:A:241[B]:GLU:HG2	1:A:290:GLN:HG3	1.92	0.52
1:A:169:ASP:OD2	1:A:172:ARG:HD3	2.10	0.51
1:B:1368:VAL:OXT	1:B:1368:VAL:CA	2.60	0.50
1:B:1241[B]:GLU:HG2	1:B:1290:GLN:HG3	1.93	0.49
1:A:255:PRO:HA	1:A:258:PHE:CE1	2.48	0.48
1:A:175:LEU:HD21	1:B:1175:LEU:HD21	1.95	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1255:PRO:HA	1:B:1258:PHE:CD1	2.50	0.47
1:A:258:PHE:HA	5:A:2628:HOH:O	2.15	0.47
1:B:1121:TYR:O	1:B:1125:GLN:HG3	2.15	0.47
1:B:1178:LEU:HD11	1:B:1184[B]:PHE:CD1	2.50	0.46
1:A:241[B]:GLU:CG	1:A:290:GLN:HG3	2.45	0.46
1:A:251:TYR:CE1	1:A:252:LYS:HG3	2.51	0.46
1:B:1264:LYS:HA	1:B:1269:TYR:CG	2.51	0.46
1:A:84:LYS:HE2	5:A:2574:HOH:O	2.14	0.46
1:B:1120:ASN:O	1:B:1124:LYS:HG3	2.16	0.46
1:B:1138:ARG:NH2	5:B:2676:HOH:O	2.50	0.45
1:A:264:LYS:HA	1:A:269:TYR:CG	2.51	0.45
1:A:197:ASP:O	1:A:201:MET:HG2	2.16	0.45
1:B:1175:LEU:HD22	1:B:1184[A]:PHE:HZ	1.81	0.44
1:B:1102:LYS:NZ	5:B:2746:HOH:O	2.50	0.44
1:A:129:VAL:O	1:A:161:VAL:HA	2.19	0.42
1:A:84:LYS:O	1:A:87:ASP:HB2	2.20	0.42
1:A:249[A]:TRP:NE1	1:A:340:ASP:OD2	2.50	0.42
1:B:1144:LEU:HD23	1:B:1174:PRO:HD2	2.02	0.42
1:A:294:ILE:HG12	1:A:323:TYR:CE1	2.55	0.41
1:B:1294:ILE:HG12	1:B:1323:TYR:CE1	2.56	0.41
1:B:1241[A]:GLU:HG2	5:B:2536:HOH:O	2.20	0.41
1:B:1255:PRO:HA	1:B:1258:PHE:CE1	2.55	0.41
1:B:1168:ASP:O	1:B:1188:LYS:NZ	2.55	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	290/289 (100%)	286 (99%)	4 (1%)	0	100	100
1	B	289/289 (100%)	286 (99%)	3 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	579/578 (100%)	572 (99%)	7 (1%)	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	263/260 (101%)	255 (97%)	8 (3%)	41	12
1	B	262/260 (101%)	254 (97%)	8 (3%)	40	11
All	All	525/520 (101%)	509 (97%)	16 (3%)	40	12

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	82	LYS
1	A	84	LYS
1	A	86	SER
1	A	96	GLU
1	A	173	MET
1	A	182	ARG
1	A	193	LYS
1	A	357	GLN
1	B	1084	LYS
1	B	1104	LYS
1	B	1125	GLN
1	B	1173	MET
1	B	1182	ARG
1	B	1193	LYS
1	B	1234	PHE
1	B	1264	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	125	GLN
1	A	296	GLN
1	A	357	GLN
1	B	1231	GLN
1	B	1296	GLN

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 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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$
4	GOL	A	404	-	5,5,5	4.35	4 (80%)	5,5,5	3.42	3 (60%)
3	UDP	B	1401	2	20,26,26	2.05	4 (20%)	25,40,40	1.16	3 (12%)
4	GOL	B	1404	-	5,5,5	4.27	4 (80%)	5,5,5	3.41	3 (60%)
3	UDP	A	401	2	20,26,26	2.08	4 (20%)	25,40,40	1.32	3 (12%)

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	GOL	A	404	-	-	1/4/4/4	-
3	UDP	B	1401	2	-	1/14/32/32	0/2/2/2
4	GOL	B	1404	-	-	0/4/4/4	-
3	UDP	A	401	2	-	1/14/32/32	0/2/2/2

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1404	GOL	C3-C2	-6.88	1.23	1.51
4	A	404	GOL	C3-C2	-6.85	1.23	1.51
3	A	401	UDP	C4-N3	5.86	1.43	1.33
3	B	1401	UDP	C4-N3	5.36	1.42	1.33
3	A	401	UDP	C6-N1	5.31	1.42	1.35
3	B	1401	UDP	C6-N1	5.04	1.42	1.35
4	A	404	GOL	O1-C1	4.86	1.62	1.42
4	B	1404	GOL	O1-C1	4.69	1.62	1.42
4	A	404	GOL	O3-C3	3.69	1.58	1.42
4	B	1404	GOL	O3-C3	3.61	1.57	1.42
3	B	1401	UDP	O4'-C1'	2.90	1.45	1.41
3	B	1401	UDP	PB-O1B	2.89	1.59	1.50
4	A	404	GOL	C1-C2	-2.63	1.40	1.51
3	A	401	UDP	PB-O1B	2.59	1.58	1.50
3	A	401	UDP	O4'-C1'	2.57	1.44	1.41
4	B	1404	GOL	C1-C2	-2.44	1.41	1.51

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	404	GOL	O3-C3-C2	5.72	137.65	110.20
4	B	1404	GOL	O3-C3-C2	5.41	136.15	110.20
4	B	1404	GOL	O2-C2-C3	4.19	127.59	109.12
3	A	401	UDP	C5-C4-N3	-3.90	114.73	123.31
4	A	404	GOL	O2-C2-C3	3.76	125.67	109.12
3	B	1401	UDP	PA-O3A-PB	-3.46	120.95	132.83
3	A	401	UDP	PA-O3A-PB	-3.05	122.37	132.83
4	A	404	GOL	O1-C1-C2	2.97	124.44	110.20
3	A	401	UDP	O3B-PB-O2B	2.83	118.44	107.64
4	B	1404	GOL	C3-C2-C1	-2.60	101.61	111.70
3	B	1401	UDP	O3B-PB-O2B	2.33	116.54	107.64
3	B	1401	UDP	C5-C4-N3	-2.16	118.56	123.31

There are no chirality outliers.

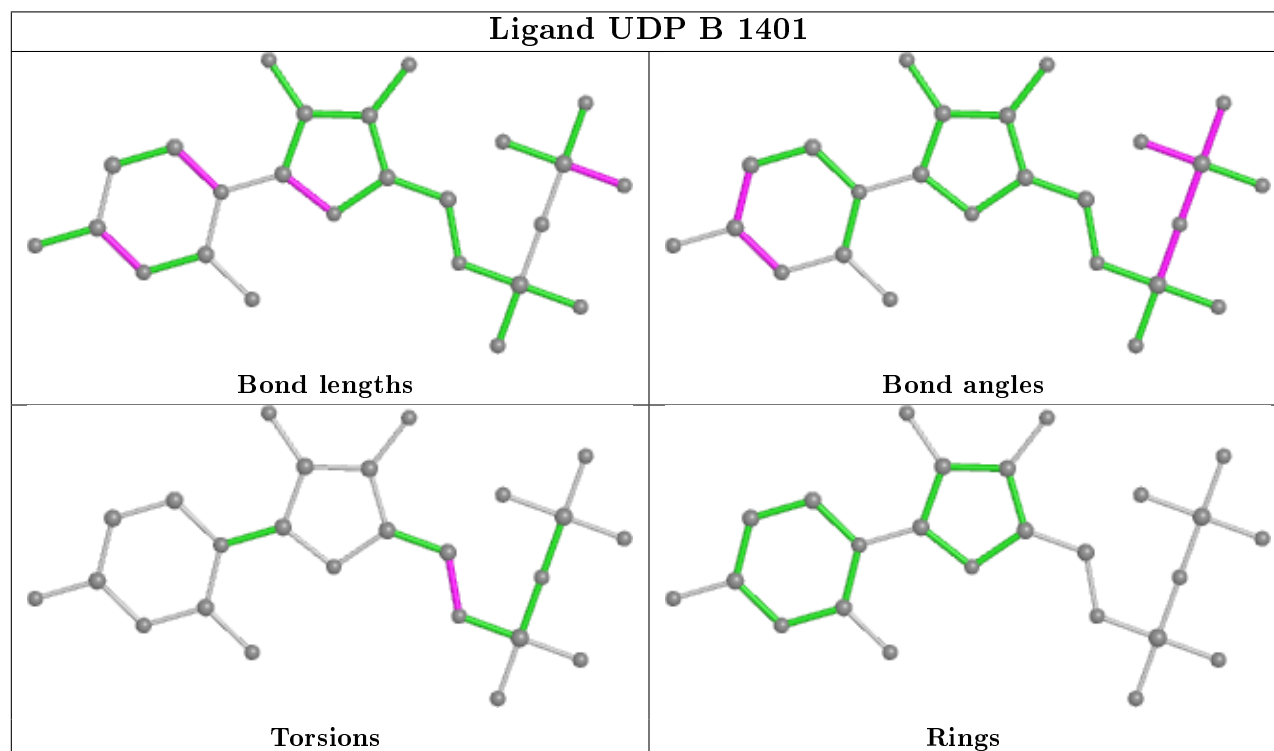
All (3) torsion outliers are listed below:

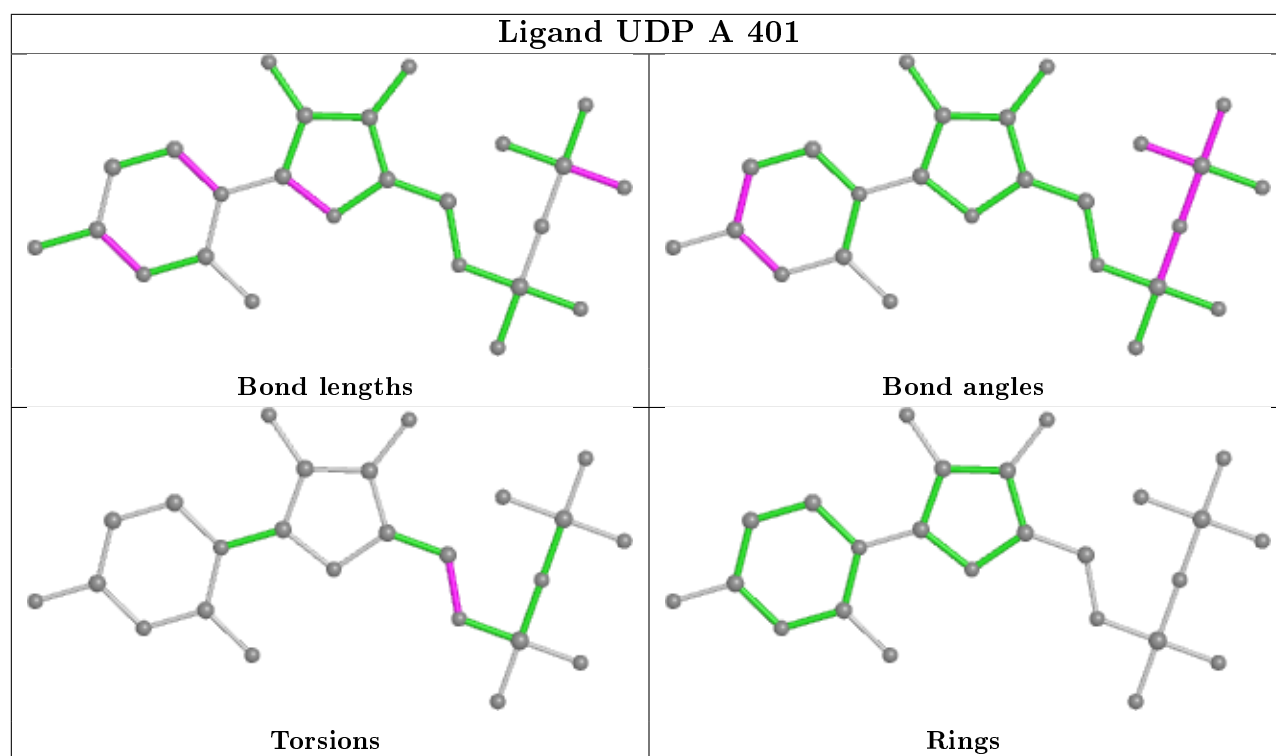
Mol	Chain	Res	Type	Atoms
4	A	404	GOL	C1-C2-C3-O3
3	A	401	UDP	C4'-C5'-O5'-PA
3	B	1401	UDP	C4'-C5'-O5'-PA

There are no ring outliers.

No monomer is involved in short contacts.

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 than 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.





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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.