wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 6, 2023 – 03:48 am GMT

PDB ID	:	1O7O
Title	:	Roles of Individual Residues of Alpha-1,3 Galactosyltransferases in Substrate Binding and Catalysis
Authors	:	Zhang, Y.; Swaminathan, G.J.; Deshpande, A.; Natesh, R.; Xie, Z.; Acharya, K.R.; Brew, K.
Deposited on	:	2002-11-11
Resolution	:	1.97 Å(reported)

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

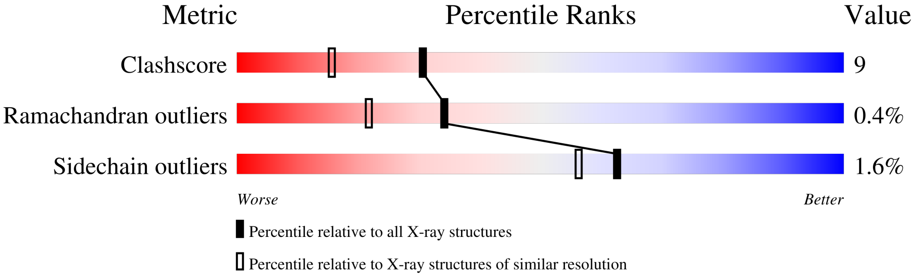
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.97 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	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	289	
1	B	289	
2	C	2	
2	D	2	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5478 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 2393 1564 396 421 12	0	0	0
1	B	287	Total C N O S 2393 1564 396 421 12	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	317	GLN	GLU	engineered mutation	UNP P14769
B	1317	GLN	GLU	engineered mutation	UNP P14769

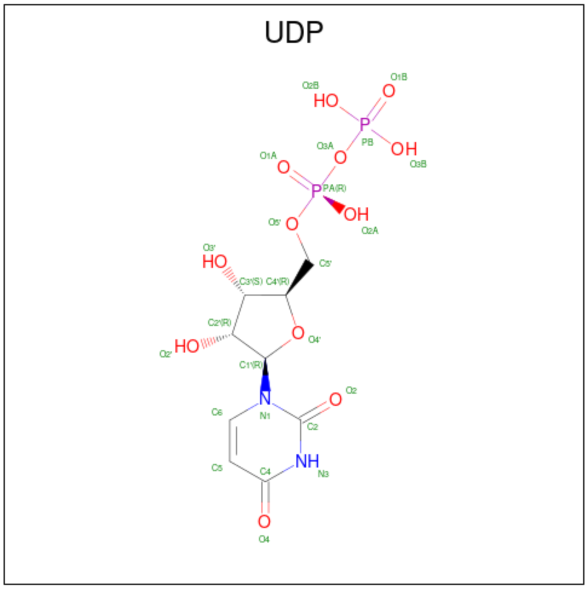
- Molecule 2 is an oligosaccharide called beta-D-galactopyranose-(1-4)-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	C	2	Total C O 23 12 11	0	0	0
2	D	2	Total C O 23 12 11	0	0	0

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

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

- Molecule 4 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: $C_9H_{14}N_2O_{12}P_2$).



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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	291	291	291	0	0
5	B	303	303	303	0	0

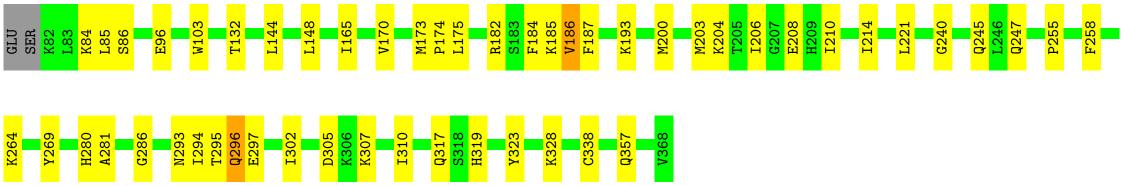
3 Residue-property plots

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. 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.

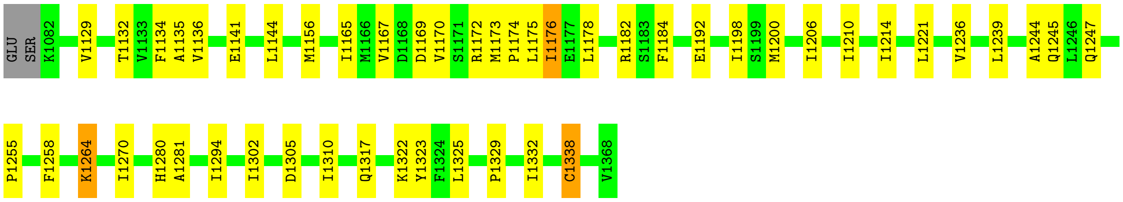
- Molecule 1: N-ACETYLLACTOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERAS E

Chain A:



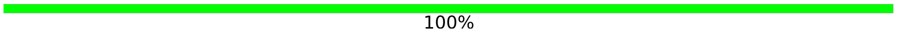
- Molecule 1: N-ACETYLLACTOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERAS E

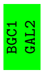
Chain B:



- Molecule 2: beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain C:





- Molecule 2: beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain D:

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	45.20Å 94.29Å 94.67Å 90.00° 99.10° 90.00°	Depositor
Resolution (Å)	50.00 – 1.97	Depositor
% Data completeness (in resolution range)	0.0 (50.00-1.97)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.181 , 0.206	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5478	wwPDB-VP
Average B, all atoms ($Å^2$)	16.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: UDP, BGC, GAL, 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	0.32	0/2466	0.60	1/3340 (0.0%)
1	B	0.32	0/2466	0.60	0/3340
All	All	0.32	0/4932	0.60	1/6680 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	296	GLN	N-CA-C	-5.43	96.34	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2393	0	2354	41	0
1	B	2393	0	2354	50	0
2	C	23	0	21	0	0
2	D	23	0	21	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	25	0	11	0	0
4	B	25	0	11	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	291	0	0	2	0
5	B	303	0	0	0	0
All	All	5478	0	4772	89	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 9.

The worst 5 of 89 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:1129:VAL:HG21	1:B:1156:MET:HE2	1.45	0.98
1:A:206:ILE:O	1:A:210:ILE:HD13	1.75	0.87
1:A:84:LYS:HD3	1:A:85:LEU:N	1.93	0.84
1:B:1281:ALA:H	1:B:1317:GLN:HE21	1.28	0.80
1:B:1329:PRO:CG	1:B:1332:ILE:HD11	2.19	0.72

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	285/289 (99%)	273 (96%)	11 (4%)	1 (0%)	34	22
1	B	285/289 (99%)	278 (98%)	6 (2%)	1 (0%)	34	22
All	All	570/578 (99%)	551 (97%)	17 (3%)	2 (0%)	34	22

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	186	VAL
1	B	1338	CYS

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	258/260 (99%)	254 (98%)	4 (2%)	62	56
1	B	258/260 (99%)	254 (98%)	4 (2%)	62	56
All	All	516/520 (99%)	508 (98%)	8 (2%)	62	56

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1264	LYS
1	B	1247	GLN
1	B	1176	ILE
1	A	357	GLN
1	B	1182	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1245	GLN
1	B	1296	GLN
1	B	1317	GLN
1	B	1313	GLN
1	A	327	ASN

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 ⓘ

4 monosaccharides 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	BGC	C	1	2	12,12,12	0.82	0	17,17,17	0.71	0
2	GAL	C	2	2	11,11,12	0.87	0	15,15,17	0.75	0
2	BGC	D	1	2	12,12,12	0.78	0	17,17,17	0.72	0
2	GAL	D	2	2	11,11,12	0.83	0	15,15,17	0.81	1 (6%)

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	BGC	C	1	2	-	2/2/22/22	0/1/1/1
2	GAL	C	2	2	-	0/2/19/22	0/1/1/1
2	BGC	D	1	2	-	0/2/22/22	0/1/1/1
2	GAL	D	2	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	2	GAL	C2-C3-C4	-2.01	107.42	110.89

There are no chirality outliers.

All (2) torsion outliers are listed below:

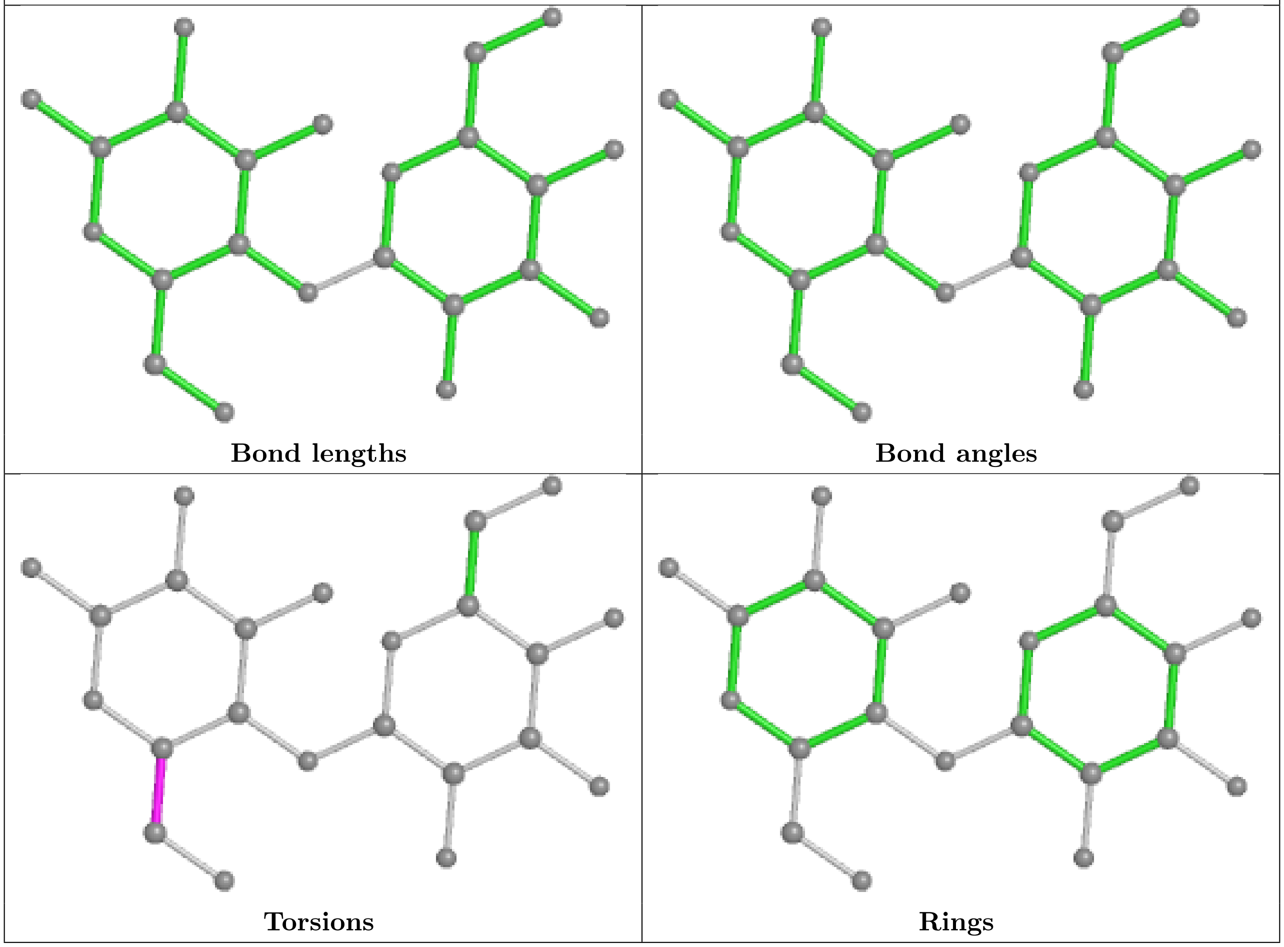
Mol	Chain	Res	Type	Atoms
2	C	1	BGC	C4-C5-C6-O6
2	C	1	BGC	O5-C5-C6-O6

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 oligosaccharide.

Oligosaccharide Chain C



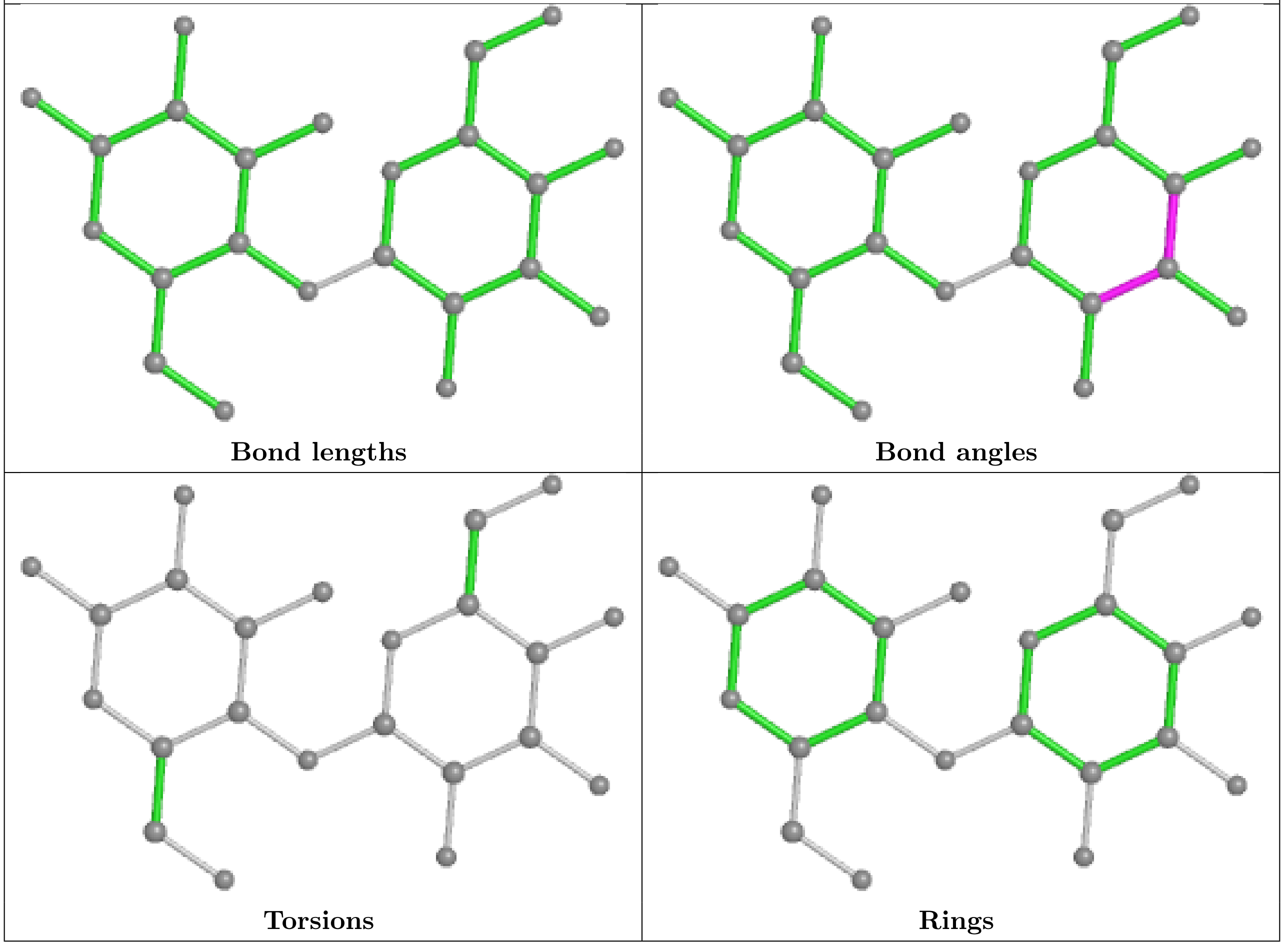
Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain D



Bond lengths

Bond angles

Torsions

Rings

5.6 Ligand geometry

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	UDP	B	2371	3	24,26,26	1.23	1 (4%)	37,40,40	1.06	2 (5%)
4	UDP	A	1371	3	24,26,26	1.22	2 (8%)	37,40,40	1.05	2 (5%)

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	UDP	B	2371	3	-	3/16/32/32	0/2/2/2
4	UDP	A	1371	3	-	3/16/32/32	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	2371	UDP	PB-O1B	3.41	1.61	1.50
4	A	1371	UDP	PB-O1B	3.33	1.61	1.50
4	A	1371	UDP	C4-N3	2.08	1.42	1.38

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	2371	UDP	O3B-PB-O2B	3.35	120.46	107.64
4	A	1371	UDP	O3B-PB-O2B	3.32	120.34	107.64
4	A	1371	UDP	O3B-PB-O3A	2.28	112.27	104.64
4	B	2371	UDP	O3B-PB-O3A	2.21	112.05	104.64

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

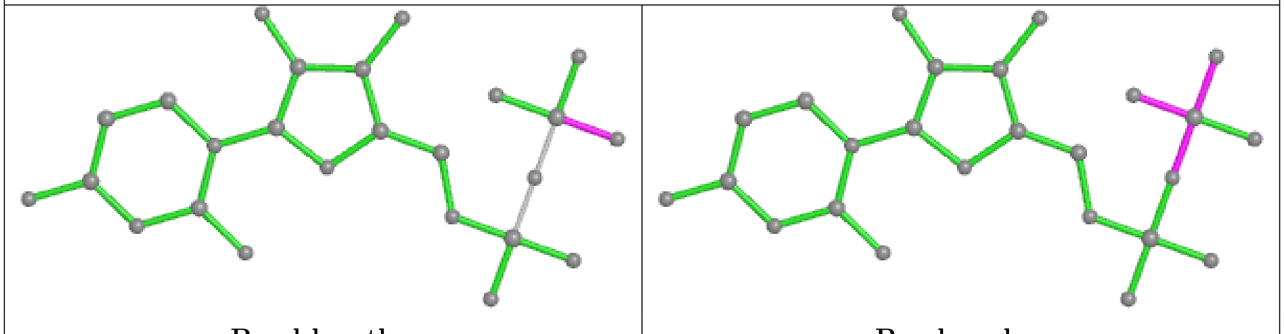
Mol	Chain	Res	Type	Atoms
4	A	1371	UDP	C3'-C4'-C5'-O5'
4	A	1371	UDP	C4'-C5'-O5'-PA
4	A	1371	UDP	O4'-C4'-C5'-O5'
4	B	2371	UDP	C3'-C4'-C5'-O5'
4	B	2371	UDP	C4'-C5'-O5'-PA

There are no ring outliers.

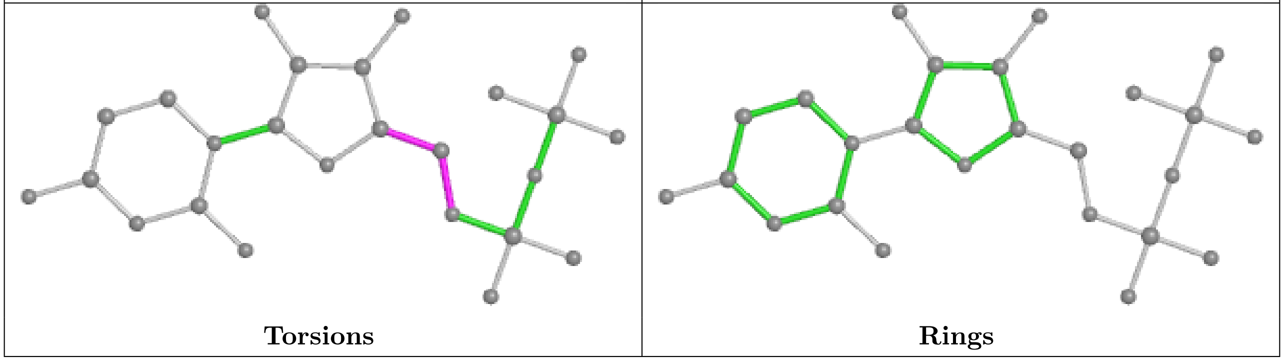
No monomer is involved in short contacts.

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 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 UDP B 2371

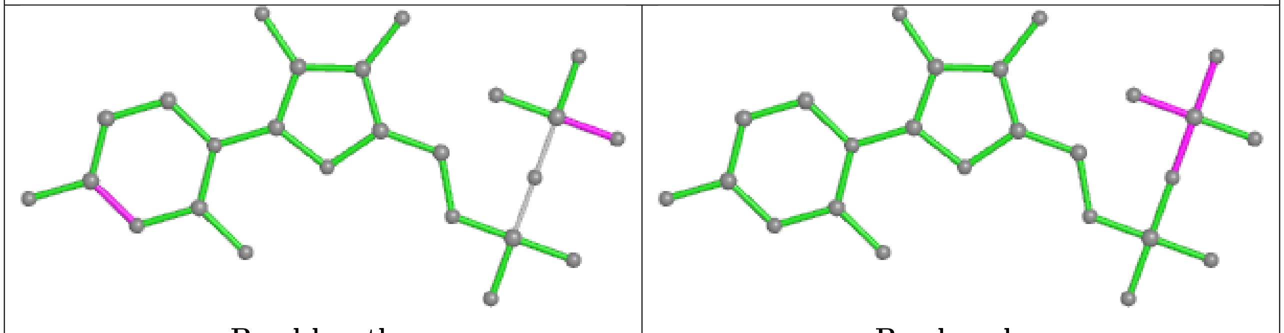


Bond lengths **Bond angles**

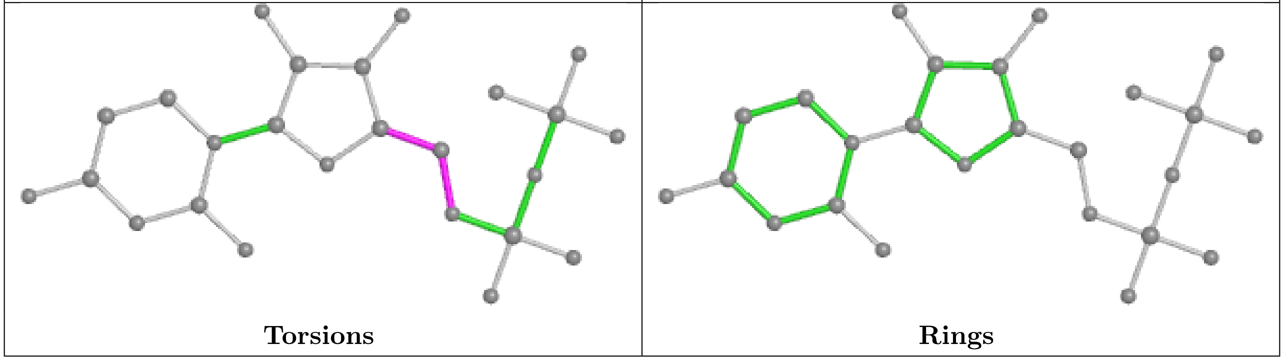


Torsions **Rings**

Ligand UDP A 1371



Bond lengths **Bond angles**



Torsions **Rings**

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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.