



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 30, 2020 – 01:39 PM BST

PDB ID : 2R24  
Title : Human Aldose Reductase structure  
Authors : Blakeley, M.P.; Ruiz, F.; Cachau, R.; Hazemann, I.; Meilleur, F.; Mitschler, A.; Ginell, S.; Afonine, P.; Ventura, O.N.; Cousido-Siah, A.; Haertlein, M.; Joachimiak, A.; Myles, D.; Podjarny, A.  
Deposited on : 2007-08-24  
Resolution : 1.75 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : **FAILED**  
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) : 2.13

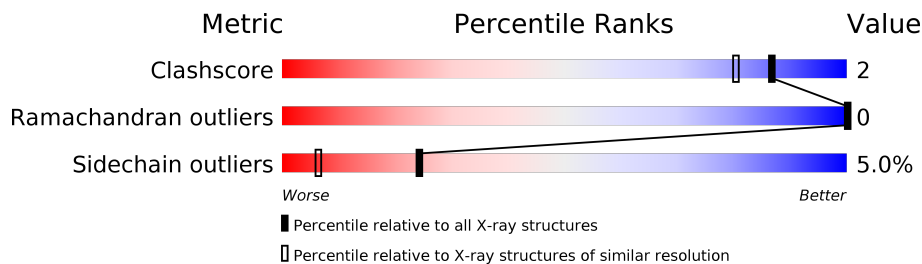
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION, NEUTRON DIFFRACTION*

The reported resolution of this entry is 1.75 Å.

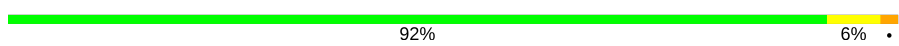
Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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 failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	316	 92% 6%

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	LDT	A	317	-	X	-	-

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 5434 atoms, of which 0 are hydrogens and 2552 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.

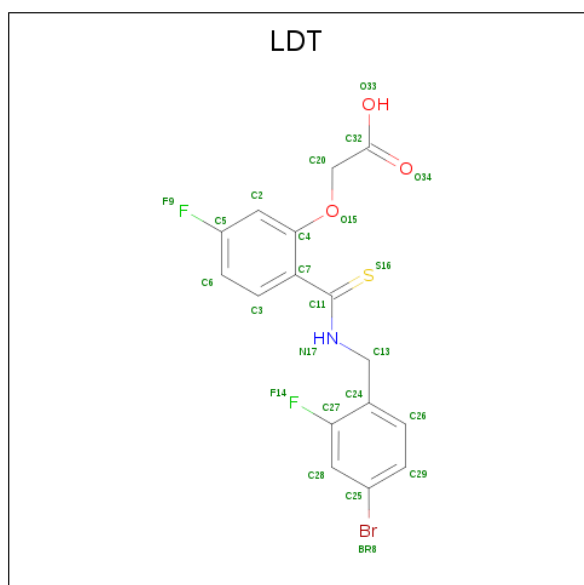
- Molecule 1 is a protein called Aldose reductase.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	D	N	O	S				
1	A	316	5067	1622	2542	426	464	13	0	0	0	

- Molecule 2 is a ligand with the chemical component id NAP but its atom names do not match the existing wwPDB Chemical Component Dictionary definition for NAP. ERROR THIS SHOULD NOT HAPPEN FOLLOWING ANNOTATION.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	D	N	O	P		
2	A	1	58	21	10	7	17	3	0	0

- Molecule 3 is IDD594 (three-letter code: LDT) (formula: C<sub>16</sub>H<sub>12</sub>BrF<sub>2</sub>NO<sub>3</sub>S).



Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
			Total	Br	C	F	N	O	S		
3	A	1	24	1	16	2	1	3	1	0	0

- Molecule 4 is water.

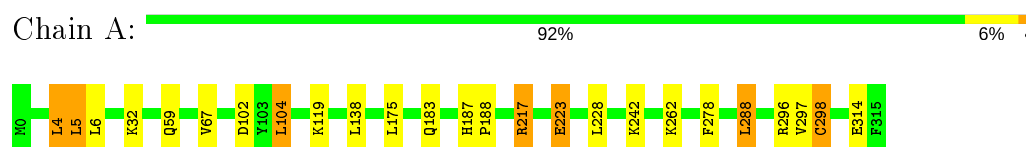
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	A	285	Total 285	O 285	0	0

### 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. 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 failed to run properly.

- Molecule 1: Aldose reductase



## 4 Data and refinement statistics i

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	50.07Å 67.13Å 47.86Å 90.00° 92.41° 90.00°	Depositor
Resolution (Å)	33.56 – 1.75	Depositor
% Data completeness (in resolution range)	98.8 (33.56-1.75)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.17 (at 1.75Å)	Xtrriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, $R_{free}$	0.129 , 0.166	Depositor
Wilson B-factor (Å <sup>2</sup> )	17.5	Xtrriage
Anisotropy	0.108	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.000 for -l,k,h 0.027 for -h,-k,l 0.015 for l,-k,h	Xtrriage
Total number of atoms	5434	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.20% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NAP, LDT

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	1.06	3/2587 (0.1%)	0.96	4/3512 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	298	CYS	CB-SG	-9.57	1.66	1.82
1	A	102	ASP	CB-CG	-5.83	1.39	1.51
1	A	32	LYS	CD-CE	-5.12	1.38	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	288	LEU	CB-CG-CD1	7.31	123.42	111.00
1	A	217	ARG	NE-CZ-NH1	6.85	123.72	120.30
1	A	104	LEU	CB-CG-CD1	6.59	122.20	111.00
1	A	32	LYS	CD-CE-NZ	-5.83	98.30	111.70

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	5067	0	2542	10	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	58	0	25	2	0
3	A	24	0	11	1	0
4	A	285	0	0	4	1
All	All	5434	0	2578	11	1

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 2.

The worst 5 of 11 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:A:59:GLN:HG2	4:A:366:HOH:O	1.93	0.62
1:A:4:LEU:HD11	4:A:504:HOH:O	1.96	0.59
1:A:223:GLU:OE1	1:A:223:GLU:HA	2.01	0.56
1:A:4:LEU:HD12	1:A:5:LEU:N	2.17	0.55
1:A:4:LEU:CD1	4:A:504:HOH:O	2.58	0.49

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:575:HOH:O	4:A:579:HOH:O 2_556	2.09	0.11

## 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	314/316 (99%)	310 (99%)	4 (1%)	0	<b>100</b> <b>100</b>

There are no Ramachandran outliers to report.



### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	281/281 (100%)	267 (95%)	14 (5%)	24 6

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	138	LEU
1	A	175	LEU
1	A	296	ARG
1	A	119	LYS
1	A	288	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 1 could not be matched to an existing wwPDB Chemical Component Dictionary definition at this stage - leaving 1 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
3	LDT	A	317	-	22,25,25	5.29	14 (63%)	31,34,34	4.17	16 (51%)

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
3	LDT	A	317	-	-	2/12/14/14	0/2/2/2

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	317	LDT	C11-N17	13.23	1.42	1.33
3	A	317	LDT	C2-C5	7.26	1.49	1.37
3	A	317	LDT	C24-C27	6.78	1.50	1.38
3	A	317	LDT	C29-C26	6.74	1.51	1.38
3	A	317	LDT	C3-C7	6.40	1.50	1.39

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	317	LDT	C13-N17-C11	-13.04	106.68	124.86
3	A	317	LDT	C7-C11-N17	7.44	121.83	115.39
3	A	317	LDT	C27-C28-C25	-6.91	111.92	117.40
3	A	317	LDT	C4-C7-C11	-6.69	111.29	122.27
3	A	317	LDT	C29-C25-C28	6.69	131.17	121.48

There are no chirality outliers.

All (2) torsion outliers are listed below:

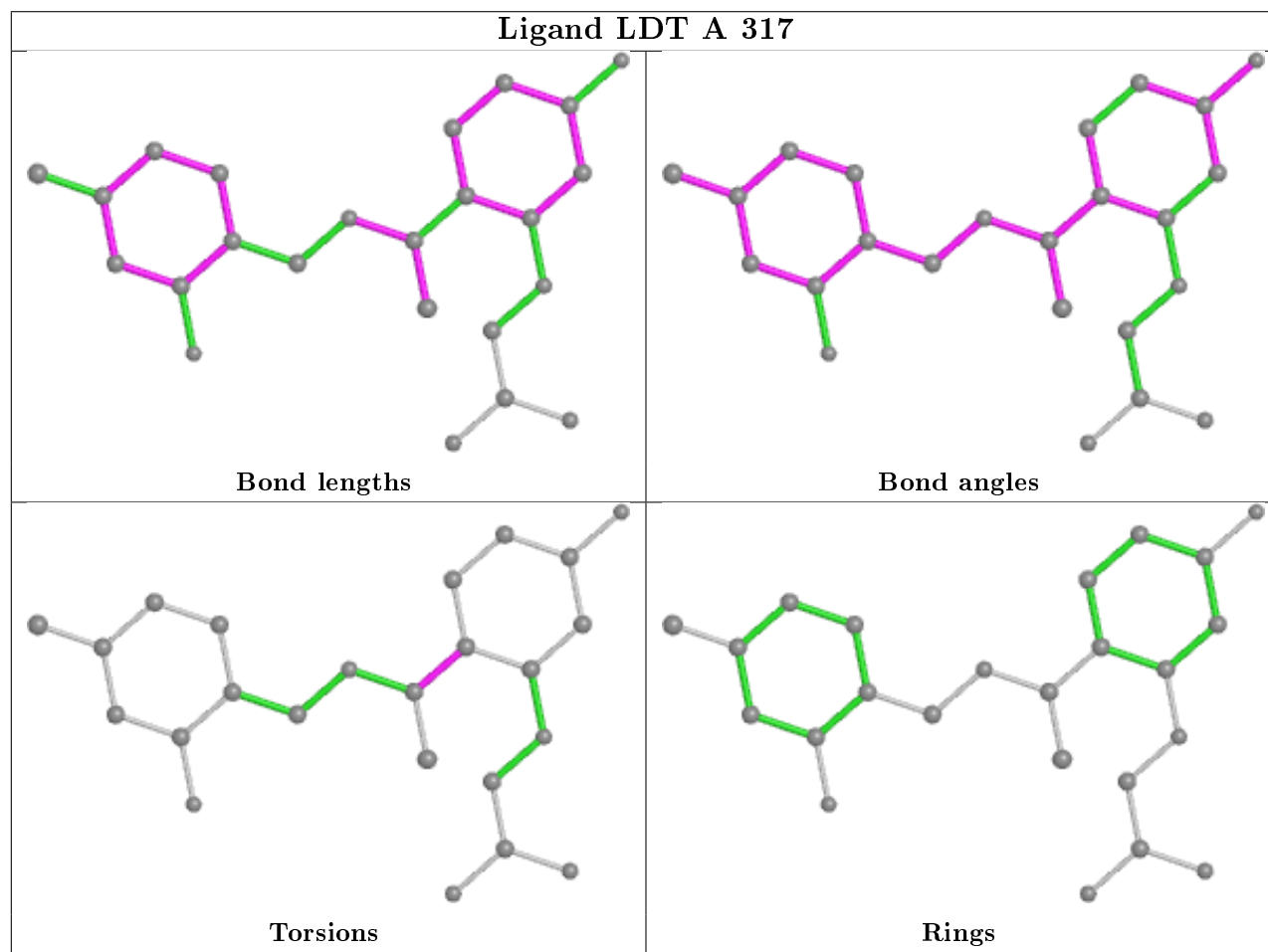
Mol	Chain	Res	Type	Atoms
3	A	317	LDT	N17-C11-C7-C3
3	A	317	LDT	N17-C11-C7-C4

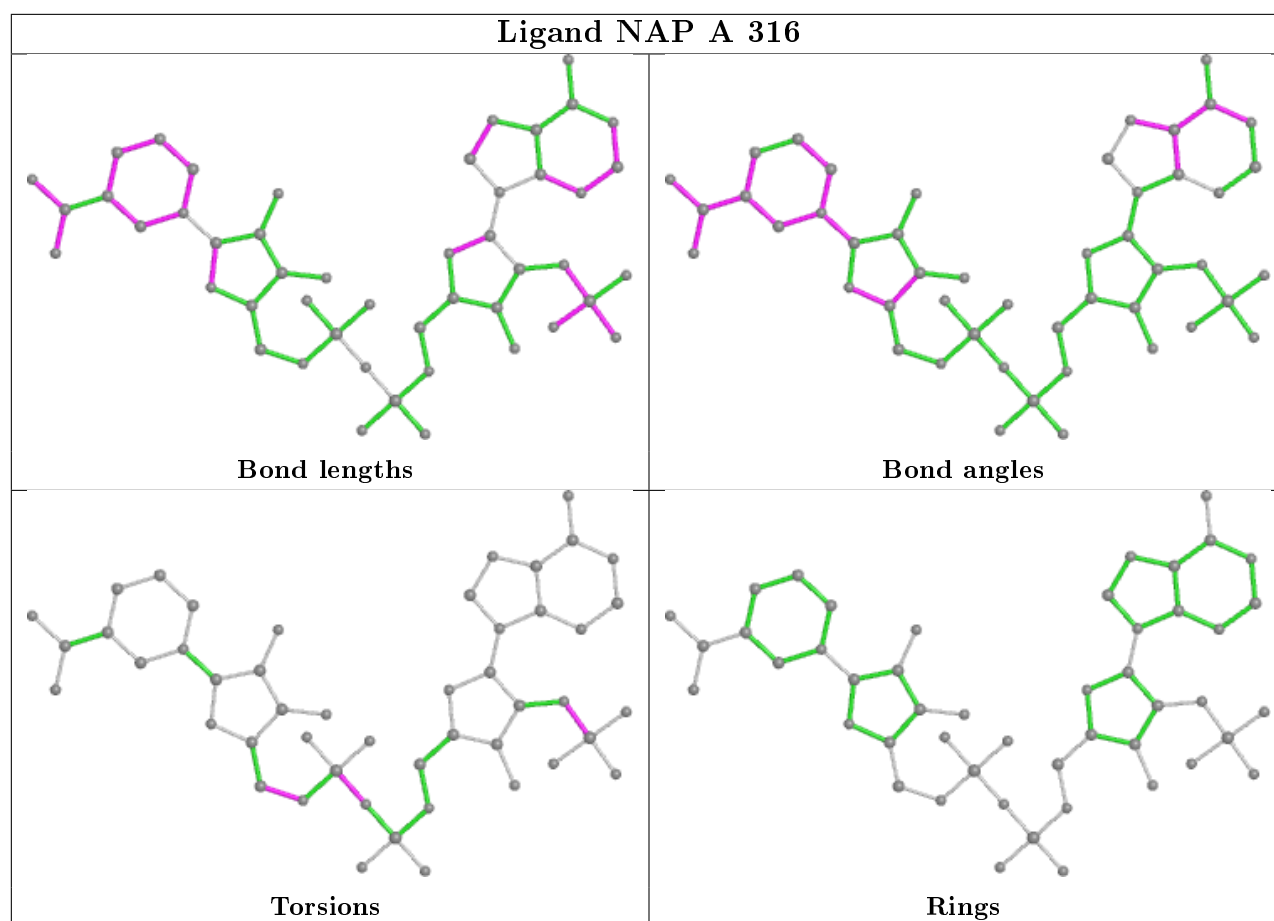
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	317	LDT	1	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 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 failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

### 6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

### 6.4 Ligands

EDS failed to run properly - this section is therefore empty.

### 6.5 Other polymers

EDS failed to run properly - this section is therefore empty.