



## wwPDB EM Validation Summary Report ⓘ

Aug 6, 2023 – 06:08 PM JST

PDB ID : 7XKJ  
Title : Kras-G12D-GDP-MRTX1133 by FIB-MicroED  
Authors : Li, X.M.  
Deposited on : 2022-04-19  
Resolution : 3.00 Å(reported)  
Based on initial model : 7RPZ

This is a wwPDB EM Validation Summary Report for a publicly released PDB/EMDB 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/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
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

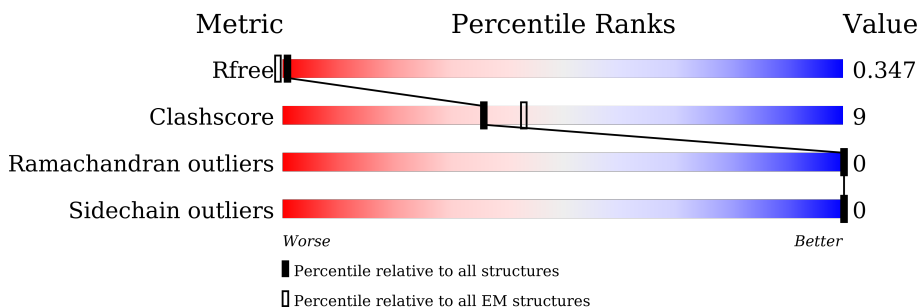
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON CRYSTALLOGRAPHY*


The reported resolution of this entry is 3.00 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)	EM structures (#Entries)
$R_{free}$	130704	0
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	169	 79% 21% .

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 1409 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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.

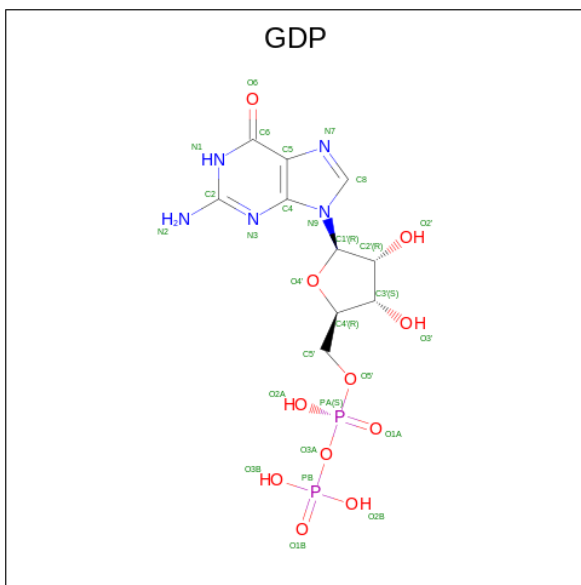
- Molecule 1 is a protein called KRAS proto-oncogene, GTPase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	168	1330	830	228	267	5	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	12	ASP	GLY	engineered mutation	UNP A0A7J7Z4L6
A	118	SER	CYS	engineered mutation	UNP A0A7J7Z4L6

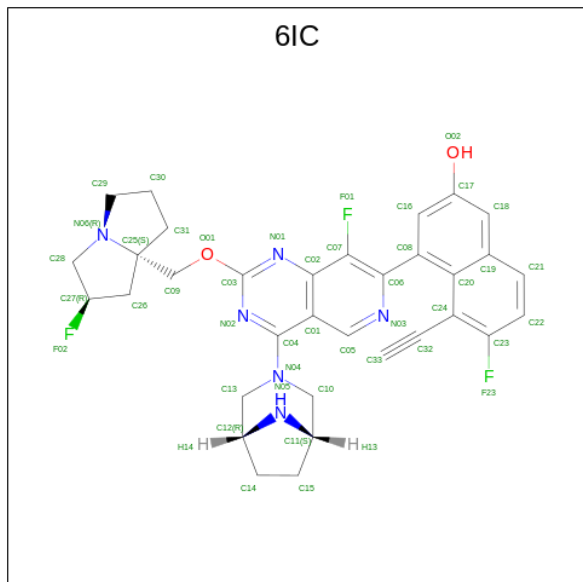
- Molecule 2 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: C<sub>10</sub>H<sub>15</sub>N<sub>5</sub>O<sub>11</sub>P<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
2	A	1	28	10	5	11	2	0

- Molecule 3 is 4-(4-[(1R,5S)-3,8-diazabicyclo[3.2.1]octan-3-yl]-8-fluoro-2-([(2R,4R,7aS)-2-fluorotetrahydro-1H-pyrrolizin-7a(5H)-yl]methoxy)pyrido[4,3-d]pyrimidin-7-yl)-5-ethynyl-6-fluo

ronaphthalen-2-ol (three-letter code: 6IC) (formula: C<sub>33</sub>H<sub>31</sub>F<sub>3</sub>N<sub>6</sub>O<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	F	N	O	
3	A	1	44	33	3	6	2	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
4	A	1	1	1	0


- Molecule 5 is water.

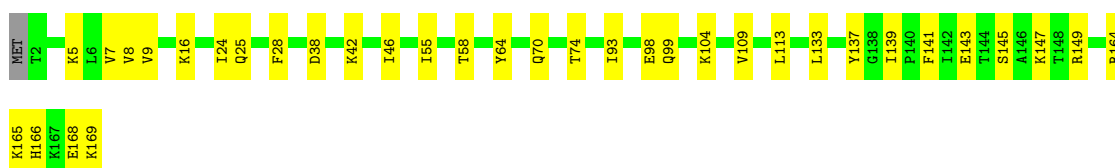
Mol	Chain	Residues	Atoms		AltConf
			Total	O	
5	A	6	6	6	0

### 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.

- Molecule 1: KRAS proto-oncogene, GTPase

Chain A:  79% 21%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	38.88Å 49.28Å 87.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.51 – 3.00 42.90 – 2.57	Depositor EDS
% Data completeness (in resolution range)	99.2 (35.51-3.00) 79.3 (42.90-2.57)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.09 (at 2.58Å)	Xtrriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, $R_{free}$	0.310 , 0.337 0.327 , 0.347	Depositor DCC
$R_{free}$ test set	508 reflections (10.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	30.7	Xtrriage
Anisotropy	0.227	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 21.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.72	EDS
Total number of atoms	1409	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.09% 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: MG, GDP, 6IC

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	0.52	0/1354	0.70	0/1826

There are no bond length outliers.

There are no bond angle outliers.

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	1330	0	1279	23	7
2	A	28	0	12	0	4
3	A	44	0	0	0	0
4	A	1	0	0	0	0
5	A	6	0	0	1	0
All	All	1409	0	1291	23	7

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 23 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:166:HIS:HA	1:A:169:LYS:HE3	1.79	0.64
1:A:137:TYR:HB2	1:A:139:ILE:HG12	1.82	0.61
1:A:145:SER:O	1:A:149:ARG:N	2.33	0.61
1:A:104:LYS:HG3	1:A:109:VAL:HG11	1.83	0.60
1:A:165:LYS:O	1:A:169:LYS:HE2	2.03	0.59

The worst 5 of 7 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)
1:A:25:GLN:OE1	1:A:64:TYR:O[3_544]	1.59	0.61
1:A:168:GLU:OE2	2:A:201:GDP:N2[1_455]	1.60	0.60
1:A:38:ASP:OD2	1:A:99:GLN:NE2[3_544]	1.66	0.54
1:A:168:GLU:OE1	2:A:201:GDP:C2[1_455]	2.02	0.18
1:A:25:GLN:CD	1:A:64:TYR:O[3_544]	2.03	0.17

### 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 PDB entries followed by that with respect to all EM entries.

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	167/169 (99%)	159 (95%)	8 (5%)	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 PDB entries followed by that with respect to all EM entries.

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	143/150 (95%)	143 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	25	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 1 is 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
3	6IC	A	202	-	47,51,51	5.14	21 (44%)	53,78,78	3.81	28 (52%)
2	GDP	A	201	1,4	24,30,30	0.98	1 (4%)	30,47,47	1.24	3 (10%)

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	6IC	A	202	-	-	1/14/55/55	0/8/8/8
2	GDP	A	201	1,4	-	3/12/32/32	0/3/3/3

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	202	6IC	C31-C25	-20.53	1.33	1.54
3	A	202	6IC	C25-N06	13.81	1.66	1.49
3	A	202	6IC	C10-N04	10.24	1.67	1.46
3	A	202	6IC	C13-N04	9.47	1.65	1.46
3	A	202	6IC	C28-N06	-9.24	1.34	1.47

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	202	6IC	O01-C03-N02	11.78	153.83	116.13
3	A	202	6IC	O01-C03-N01	-9.55	85.55	116.13
3	A	202	6IC	C23-C24-C32	-8.66	114.29	121.93
3	A	202	6IC	F01-C07-C02	6.96	132.84	118.12
3	A	202	6IC	C16-C08-C06	-6.87	105.51	118.78

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	201	GDP	C5'-O5'-PA-O1A
3	A	202	6IC	N02-C04-N04-C10
2	A	201	GDP	C5'-O5'-PA-O3A
2	A	201	GDP	C5'-O5'-PA-O2A

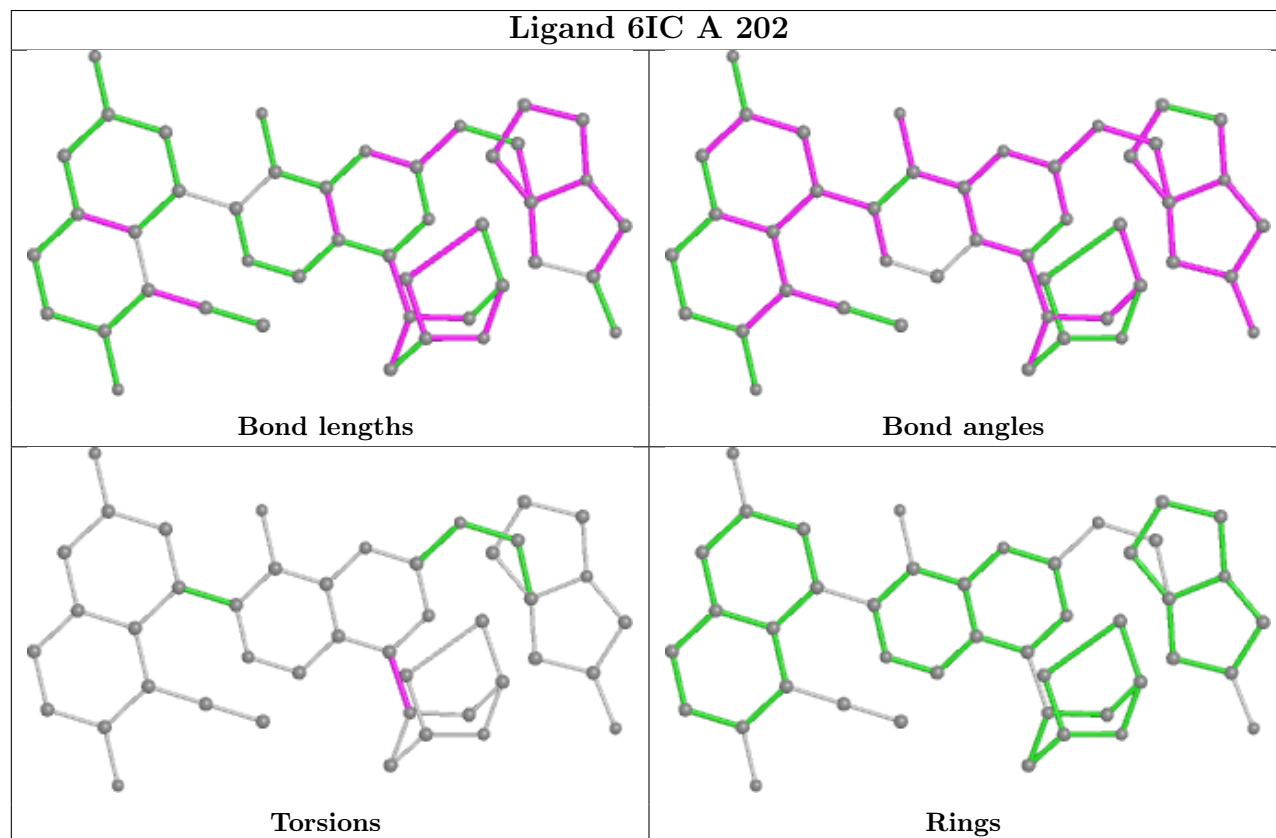
There are no ring outliers.

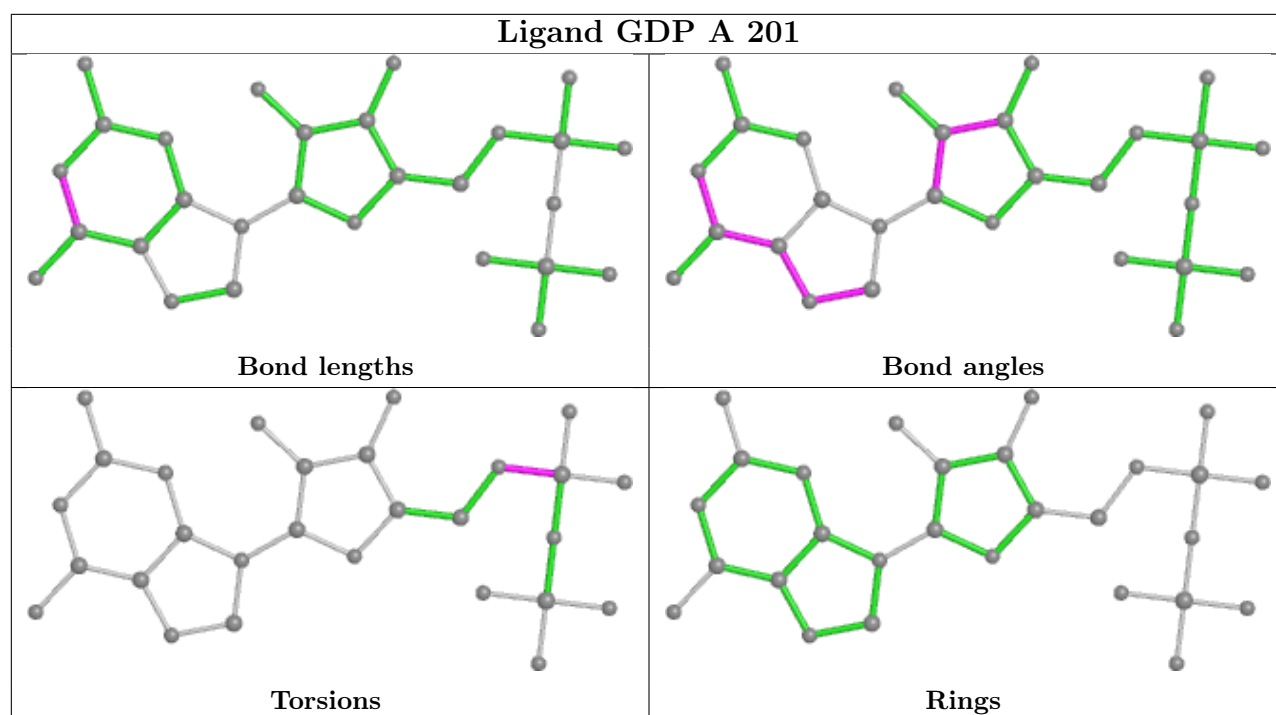
1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	201	GDP	0	4

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.