



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 4, 2023 – 11:34 PM EDT

PDB ID : 6UM8
Title : HIV Integrase in complex with Compound-14
Authors : Khan, J.A.; Kish, K.
Deposited on : 2019-10-09
Resolution : 2.33 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 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 : **FAILED**
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtrriage (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) : 2.35.1

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 2.33 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 2552 atoms, of which 60 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.

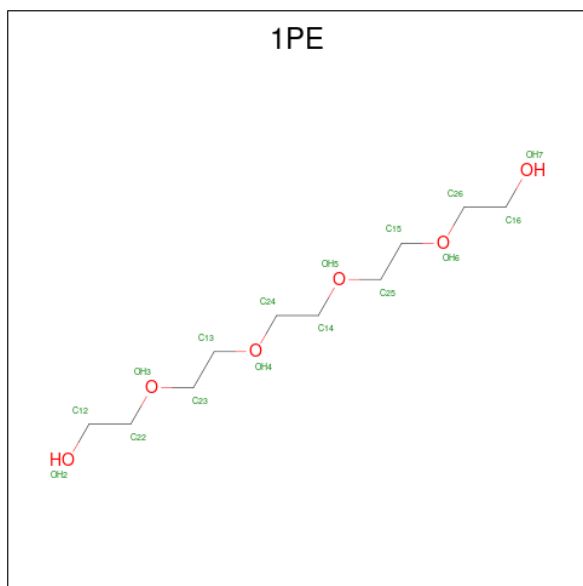
- Molecule 1 is a protein called Integrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	139	1090	694	187	205	4	0	3	0
1	B	141	1094	699	188	203	4	0	3	0

There are 24 discrepancies between the modelled and reference sequences:

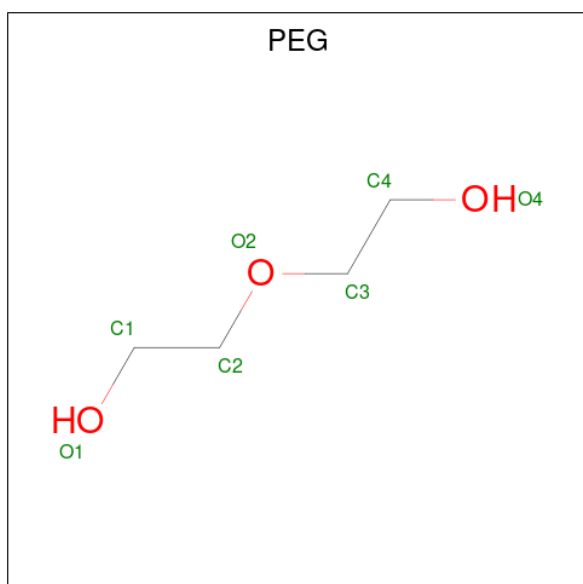
Chain	Residue	Modelled	Actual	Comment	Reference
A	47	GLY	-	expression tag	UNP Q76353
A	48	SER	-	expression tag	UNP Q76353
A	49	HIS	-	expression tag	UNP Q76353
A	50	MET	-	expression tag	UNP Q76353
A	51	HIS	-	expression tag	UNP Q76353
A	52	GLY	-	expression tag	UNP Q76353
A	53	GLN	-	expression tag	UNP Q76353
A	54	VAL	-	expression tag	UNP Q76353
A	55	ASP	-	expression tag	UNP Q76353
A	56	SER	-	expression tag	UNP Q76353
A	131	GLU	TRP	conflict	UNP Q76353
A	185	LYS	PHE	conflict	UNP Q76353
B	47	GLY	-	expression tag	UNP Q76353
B	48	SER	-	expression tag	UNP Q76353
B	49	HIS	-	expression tag	UNP Q76353
B	50	MET	-	expression tag	UNP Q76353
B	51	HIS	-	expression tag	UNP Q76353
B	52	GLY	-	expression tag	UNP Q76353
B	53	GLN	-	expression tag	UNP Q76353
B	54	VAL	-	expression tag	UNP Q76353
B	55	ASP	-	expression tag	UNP Q76353
B	56	SER	-	expression tag	UNP Q76353
B	131	GLU	TRP	conflict	UNP Q76353
B	185	LYS	PHE	conflict	UNP Q76353

- Molecule 2 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$).



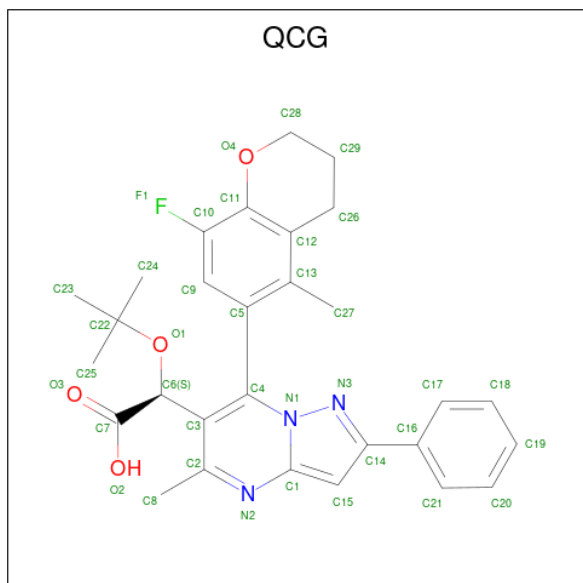
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	C O	0	0
			10	6 4		
2	A	1	Total	C O	0	0
			16	10 6		
2	B	1	Total	C O	0	0
			16	10 6		
2	B	1	Total	C O	0	0
			13	8 5		

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			7	4	3		
3	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 4 is (2S)-tert-butoxy[7-(8-fluoro-5-methyl-3,4-dihydro-2H-1-benzopyran-6-yl)-5-methyl-2-phenylpyrazolo[1,5-a]pyrimidin-6-yl]acetic acid (three-letter code: QCG) (formula: C₂₉H₃₀FN₃O₄) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
4	A	1	Total	C	F	H	N	O	30	0
			67	29	1	30	3	4		
4	B	1	Total	C	F	H	N	O	30	0
			67	29	1	30	3	4		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	54	Total	O	0	0
			54	54		
6	B	81	Total	O	0	0
			81	81		

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

3 Data and refinement statistics

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

Property	Value	Source
Space group	P 43 3 2	Depositor
Cell constants a, b, c, α , β , γ	164.63Å 164.63Å 164.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	116.41 – 2.33	Depositor
% Data completeness (in resolution range)	99.9 (116.41-2.33)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.36 (at 2.32Å)	Xtrriage
Refinement program	BUSTER 2.11.7	Depositor
R, R_{free}	0.196 , 0.220	Depositor
Wilson B-factor (Å ²)	53.7	Xtrriage
Anisotropy	0.000	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	2552	wwPDB-VP
Average B, all atoms (Å ²)	62.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.56% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

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

4.2 Too-close contacts [i](#)

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

4.3 Torsion angles [i](#)

4.3.1 Protein backbone [i](#)

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

4.3.2 Protein sidechains [i](#)

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

4.3.3 RNA [i](#)

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

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

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

4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

4.6 Ligand geometry [i](#)

14 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
5	SO4	B	304	-	4,4,4	0.12	0	6,6,6	0.21	0
2	1PE	B	302	-	12,12,15	0.26	0	11,11,14	0.35	0
2	1PE	A	302	-	15,15,15	0.26	0	14,14,14	0.30	0
4	QCG	B	303	-	37,41,41	0.79	2 (5%)	42,62,62	0.63	0
5	SO4	A	308	-	4,4,4	0.08	0	6,6,6	0.21	0
4	QCG	A	305	-	37,41,41	0.81	2 (5%)	42,62,62	0.61	0
5	SO4	B	306	-	4,4,4	0.12	0	6,6,6	0.71	0
3	PEG	A	304	-	6,6,6	0.21	0	5,5,5	0.10	0
5	SO4	A	307	-	4,4,4	0.24	0	6,6,6	0.16	0
5	SO4	B	305	-	4,4,4	0.05	0	6,6,6	0.20	0
2	1PE	B	301	-	15,15,15	0.21	0	14,14,14	0.23	0
5	SO4	A	306	-	4,4,4	0.16	0	6,6,6	0.25	0
2	1PE	A	301	-	9,9,15	0.23	0	8,8,14	0.16	0
3	PEG	A	303	-	6,6,6	0.10	0	5,5,5	0.17	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	1PE	B	302	-	-	5/10/10/13	-
2	1PE	A	302	-	-	8/13/13/13	-
4	QCG	B	303	-	-	5/20/28/28	0/5/5/5
4	QCG	A	305	-	-	5/20/28/28	0/5/5/5
3	PEG	A	304	-	-	2/4/4/4	-
2	1PE	B	301	-	-	2/13/13/13	-
2	1PE	A	301	-	-	4/7/7/13	-
3	PEG	A	303	-	-	1/4/4/4	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	305	QCG	C4-N1	-2.30	1.37	1.41
4	B	303	QCG	C4-N1	-2.18	1.37	1.41
4	B	303	QCG	C4-C3	-2.08	1.35	1.39
4	A	305	QCG	C4-C3	-2.04	1.35	1.39

There are no bond angle outliers.

There are no chirality outliers.

All (32) torsion outliers are listed below:

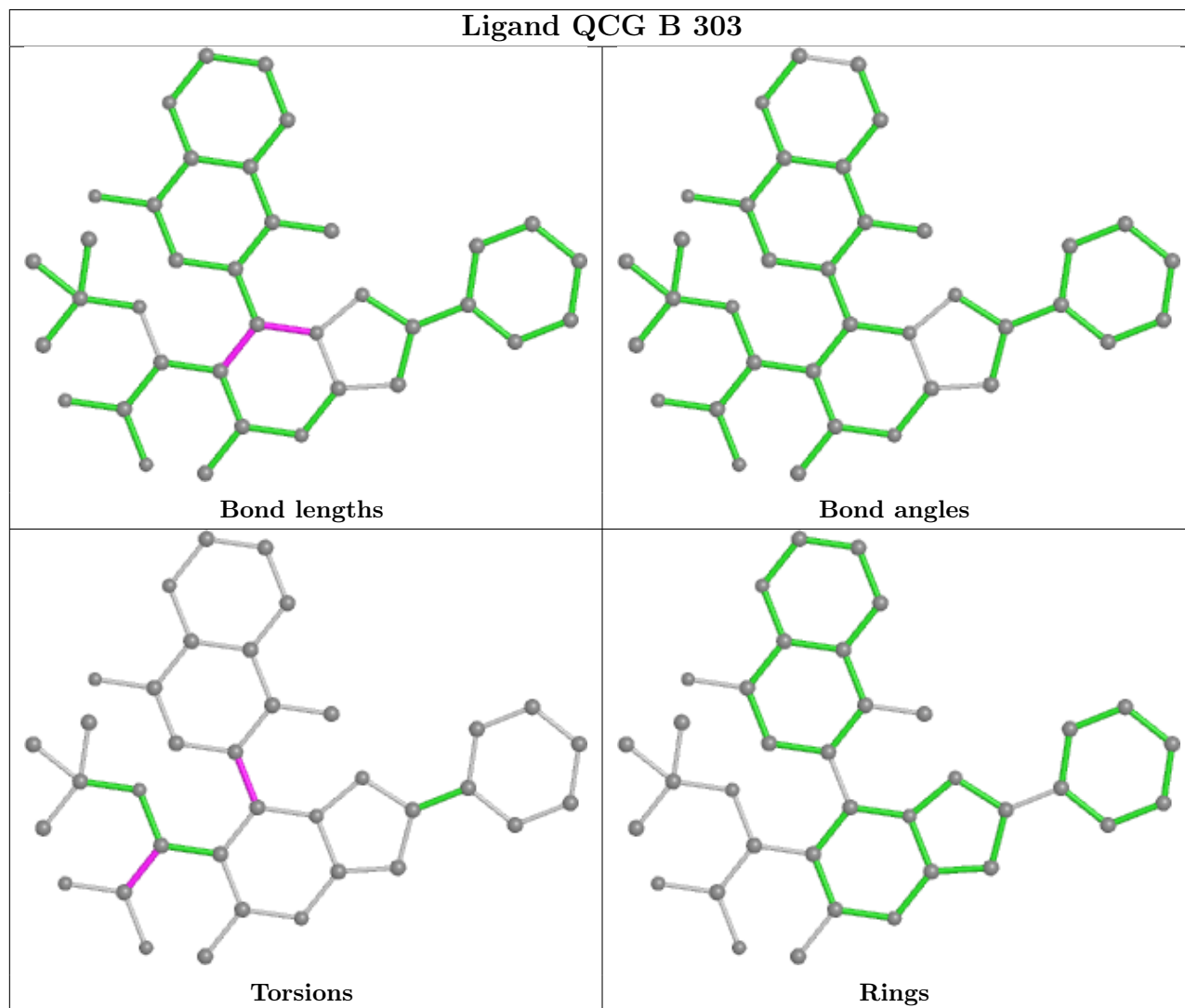
Mol	Chain	Res	Type	Atoms
4	A	305	QCG	O1-C6-C7-O2
4	B	303	QCG	C3-C6-C7-O2
4	B	303	QCG	C3-C6-C7-O3
2	A	302	1PE	OH5-C14-C24-OH4
2	B	302	1PE	OH5-C14-C24-OH4
4	A	305	QCG	C3-C6-C7-O2
4	A	305	QCG	C3-C6-C7-O3
2	A	301	1PE	OH2-C12-C22-OH3
2	A	302	1PE	C14-C24-OH4-C13
2	A	301	1PE	C23-C13-OH4-C24
2	B	302	1PE	C15-C25-OH5-C14
2	B	302	1PE	C14-C24-OH4-C13
2	A	302	1PE	C16-C26-OH6-C15
3	A	303	PEG	C1-C2-O2-C3
2	A	302	1PE	OH2-C12-C22-OH3
2	A	302	1PE	OH6-C15-C25-OH5
2	A	301	1PE	OH4-C13-C23-OH3
2	A	302	1PE	C24-C14-OH5-C25
2	B	302	1PE	C16-C26-OH6-C15
4	B	303	QCG	C3-C4-C5-C13
3	A	304	PEG	C1-C2-O2-C3
2	B	301	1PE	C14-C24-OH4-C13
2	B	301	1PE	C23-C13-OH4-C24
3	A	304	PEG	O2-C3-C4-O4
2	B	302	1PE	OH6-C15-C25-OH5
2	A	302	1PE	C23-C13-OH4-C24
4	A	305	QCG	O1-C6-C7-O3
2	A	302	1PE	OH4-C13-C23-OH3
4	A	305	QCG	C3-C4-C5-C9
4	B	303	QCG	C3-C4-C5-C9
4	B	303	QCG	O1-C6-C7-O2
2	A	301	1PE	C13-C23-OH3-C22

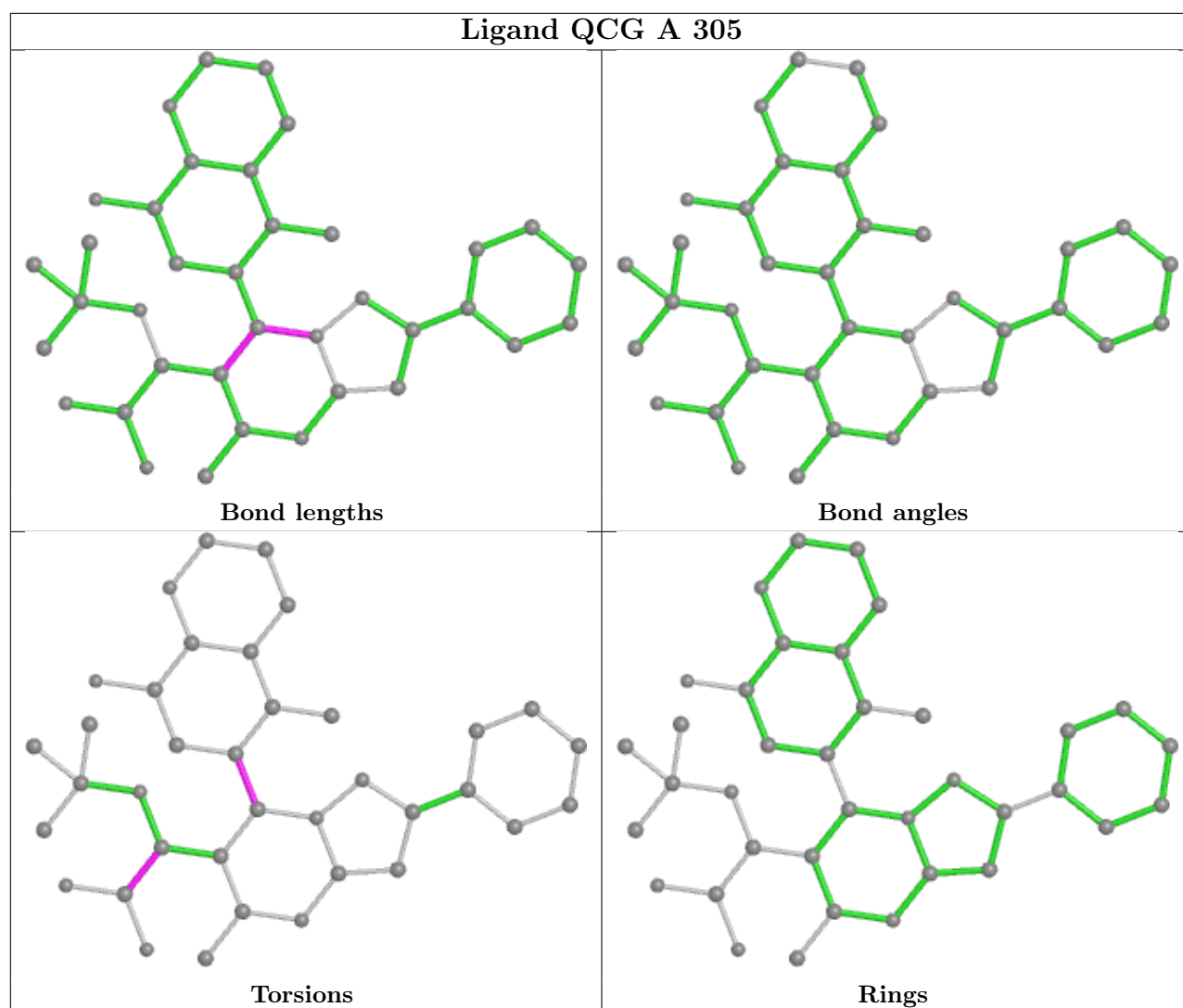
There are no ring outliers.

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.





4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

5 Fit of model and data

5.1 Protein, DNA and RNA chains

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

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

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

5.3 Carbohydrates

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

5.4 Ligands

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

5.5 Other polymers

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