



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 17, 2022 – 06:24 pm GMT

PDB ID : 8B7U  
Title : Bacterial chalcone isomerase H33A with taxifolin  
Authors : Hinrichs, W.; Palm, G.J.  
Deposited on : 2022-10-03  
Resolution : 2.80 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 types of validation reports are described at

<https://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 : 1.8.4, 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) : 2.31.2

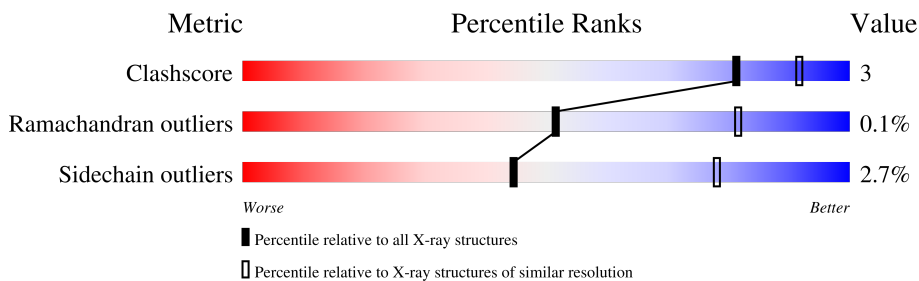
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

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  $\geq 3$ , 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	283	82% 7% 9%
1	B	283	86% 5% 8%
1	C	283	82% 8% 8%
1	D	283	86% 5% 8%
1	E	283	85% 5% 8%
1	F	283	87% 8%
1	G	283	84% 6% 9%
1	H	283	86% 5% 8%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain	
1	I	283	86%	5% • 8%
1	J	283	85%	5% • 8%
1	K	283	85%	6% • 8%
1	L	283	86%	5% • 8%
1	M	283	86%	5% • 8%
1	N	283	82%	8% • 8%
1	O	283	82%	8% • 9%
1	P	283	83%	7% • 9%
1	Q	283	84%	7% • 9%
1	R	283	82%	8% • 8%

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	CL	C	1005	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 39179 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 Chalcone isomerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	257	2112	1380	342	377	13	0	0	0
1	B	259	2122	1386	344	379	13	0	0	0
1	C	259	2122	1386	344	379	13	0	0	0
1	D	259	2122	1386	344	379	13	0	0	0
1	E	259	2122	1386	344	379	13	0	0	0
1	F	260	2127	1389	345	380	13	0	0	0
1	G	258	2117	1383	343	378	13	0	0	0
1	H	259	2122	1386	344	379	13	0	0	0
1	I	259	2122	1386	344	379	13	0	0	0
1	J	259	2122	1386	344	379	13	0	0	0
1	K	259	2122	1386	344	379	13	0	0	0
1	L	259	2122	1386	344	379	13	0	0	0
1	M	260	2127	1389	345	380	13	0	0	0
1	N	259	2122	1386	344	379	13	0	0	0
1	O	258	2117	1383	343	378	13	0	0	0
1	P	258	2117	1383	343	378	13	0	0	0

*Continued on next page...*

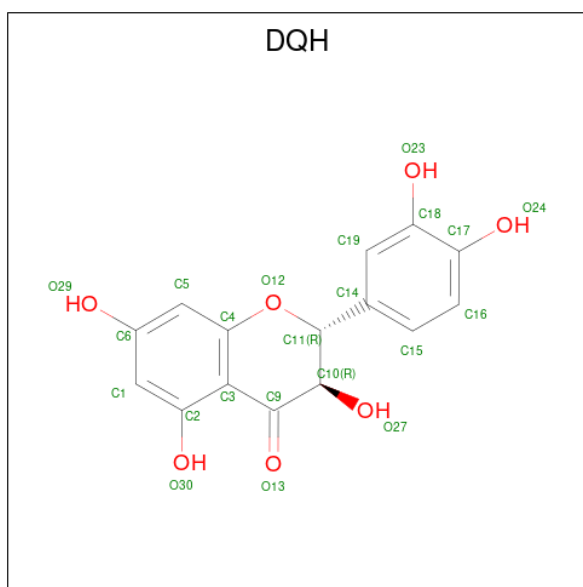
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	Q	258	2117	1383	343	378	13	0	0	0
1	R	259	2122	1386	344	379	13	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	33	ALA	HIS	engineered mutation	UNP V9P0A9
B	33	ALA	HIS	engineered mutation	UNP V9P0A9
C	33	ALA	HIS	engineered mutation	UNP V9P0A9
D	33	ALA	HIS	engineered mutation	UNP V9P0A9
E	33	ALA	HIS	engineered mutation	UNP V9P0A9
F	33	ALA	HIS	engineered mutation	UNP V9P0A9
G	33	ALA	HIS	engineered mutation	UNP V9P0A9
H	33	ALA	HIS	engineered mutation	UNP V9P0A9
I	33	ALA	HIS	engineered mutation	UNP V9P0A9
J	33	ALA	HIS	engineered mutation	UNP V9P0A9
K	33	ALA	HIS	engineered mutation	UNP V9P0A9
L	33	ALA	HIS	engineered mutation	UNP V9P0A9
M	33	ALA	HIS	engineered mutation	UNP V9P0A9
N	33	ALA	HIS	engineered mutation	UNP V9P0A9
O	33	ALA	HIS	engineered mutation	UNP V9P0A9
P	33	ALA	HIS	engineered mutation	UNP V9P0A9
Q	33	ALA	HIS	engineered mutation	UNP V9P0A9
R	33	ALA	HIS	engineered mutation	UNP V9P0A9

- Molecule 2 is (2R,3R)-2-(3,4-DIHYDROXYPHENYL)-3,5,7-TRIHYDROXY-2,3-DIHYDRO-4H-CHROMEN-4-ONE (three-letter code: DQH) (formula: C<sub>15</sub>H<sub>12</sub>O<sub>7</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 22 15 7	0	0
2	A	1	Total C O 22 15 7	0	0
2	B	1	Total C O 22 15 7	0	0
2	B	1	Total C O 22 15 7	0	0
2	C	1	Total C O 22 15 7	0	0
2	C	1	Total C O 22 15 7	0	0
2	D	1	Total C O 22 15 7	0	0
2	D	1	Total C O 22 15 7	0	0
2	E	1	Total C O 22 15 7	0	0
2	E	1	Total C O 22 15 7	0	0
2	F	1	Total C O 22 15 7	0	0
2	F	1	Total C O 22 15 7	0	0
2	G	1	Total C O 22 15 7	0	0
2	G	1	Total C O 22 15 7	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	H	1	Total	C	O	0	0
			22	15	7		
2	H	1	Total	C	O	0	0
			22	15	7		
2	I	1	Total	C	O	0	0
			22	15	7		
2	I	1	Total	C	O	0	0
			22	15	7		
2	J	1	Total	C	O	0	0
			22	15	7		
2	K	1	Total	C	O	0	0
			22	15	7		
2	K	1	Total	C	O	0	0
			22	15	7		
2	L	1	Total	C	O	0	0
			22	15	7		
2	L	1	Total	C	O	0	0
			22	15	7		
2	M	1	Total	C	O	0	0
			22	15	7		
2	M	1	Total	C	O	0	0
			22	15	7		
2	N	1	Total	C	O	0	0
			22	15	7		
2	N	1	Total	C	O	0	0
			22	15	7		
2	O	1	Total	C	O	0	0
			22	15	7		
2	O	1	Total	C	O	0	0
			22	15	7		
2	P	1	Total	C	O	0	0
			22	15	7		
2	P	1	Total	C	O	0	0
			22	15	7		
2	Q	1	Total	C	O	0	0
			22	15	7		
2	Q	1	Total	C	O	0	0
			22	15	7		
2	R	1	Total	C	O	0	0
			22	15	7		
2	R	1	Total	C	O	0	0
			22	15	7		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cl 1 1	0	0
3	B	1	Total Cl 1 1	0	0
3	C	3	Total Cl 3 3	0	0
3	D	2	Total Cl 2 2	0	0
3	E	2	Total Cl 2 2	0	0
3	F	3	Total Cl 3 3	0	0
3	G	4	Total Cl 4 4	0	0
3	H	2	Total Cl 2 2	0	0
3	I	2	Total Cl 2 2	0	0
3	J	1	Total Cl 1 1	0	0
3	K	5	Total Cl 5 5	0	0
3	L	3	Total Cl 3 3	0	0
3	M	3	Total Cl 3 3	0	0
3	N	5	Total Cl 5 5	0	0
3	O	1	Total Cl 1 1	0	0
3	P	2	Total Cl 2 2	0	0
3	Q	2	Total Cl 2 2	0	0
3	R	5	Total Cl 5 5	0	0

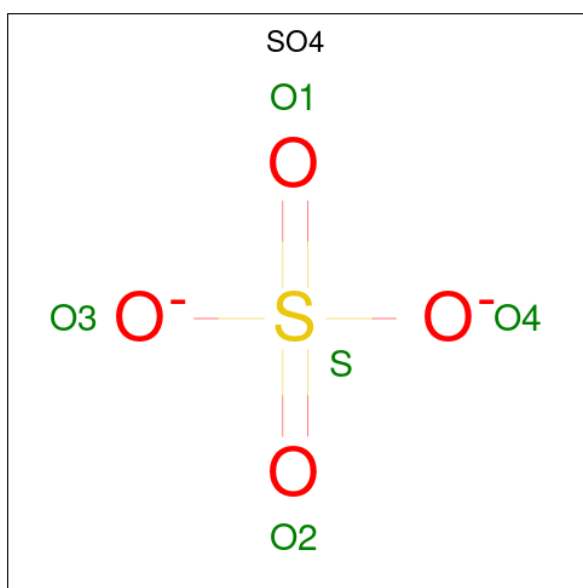
- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).





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

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	D	1	Total	O	S	0	0
			5	4	1		
5	F	1	Total	O	S	0	0
			5	4	1		
5	H	1	Total	O	S	0	0
			5	4	1		
5	I	1	Total	O	S	0	0
			5	4	1		
5	K	1	Total	O	S	0	0
			5	4	1		
5	R	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	K	0	0
			1	1		
6	C	1	Total	K	0	0
			1	1		
6	D	1	Total	K	0	0
			1	1		
6	F	1	Total	K	0	0
			1	1		
6	G	1	Total	K	0	0
			1	1		
6	H	1	Total	K	0	0
			1	1		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	7	Total	O	0	0
			7	7		
7	B	10	Total	O	0	0
			10	10		
7	C	12	Total	O	0	0
			12	12		
7	D	12	Total	O	0	0
			12	12		
7	E	9	Total	O	0	0
			9	9		

*Continued on next page...*

*Continued from previous page...*

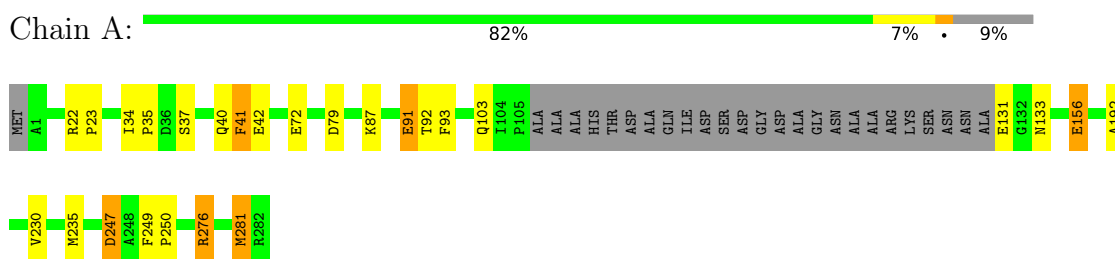
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
7	F	9	Total O 9 9	0	0
7	G	8	Total O 8 8	0	0
7	H	11	Total O 11 11	0	0
7	I	3	Total O 3 3	0	0
7	J	4	Total O 4 4	0	0
7	K	7	Total O 7 7	0	0
7	L	5	Total O 5 5	0	0
7	M	4	Total O 4 4	0	0
7	N	5	Total O 5 5	0	0
7	O	6	Total O 6 6	0	0
7	P	5	Total O 5 5	0	0
7	Q	5	Total O 5 5	0	0
7	R	6	Total O 6 6	0	0

### 3 Residue-property plots [i](#)

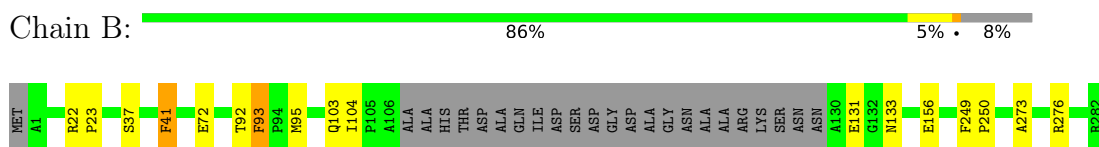
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded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

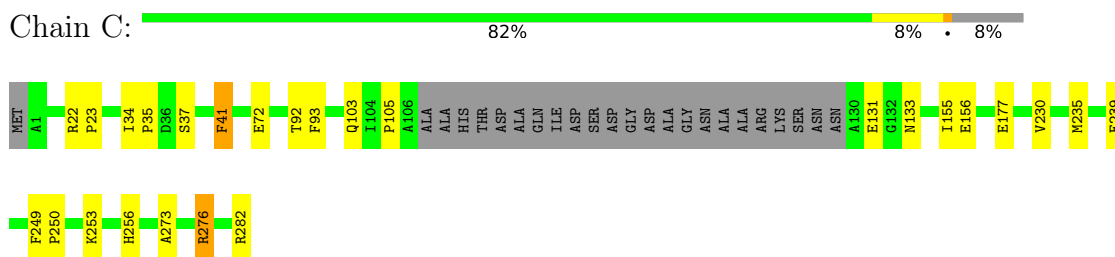
- Molecule 1: Chalcone isomerase



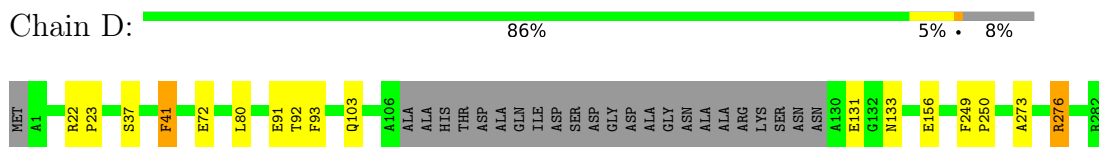
- Molecule 1: Chalcone isomerase



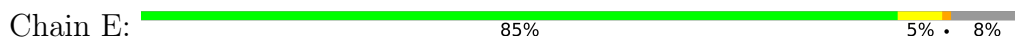
- Molecule 1: Chalcone isomerase



- Molecule 1: Chalcone isomerase




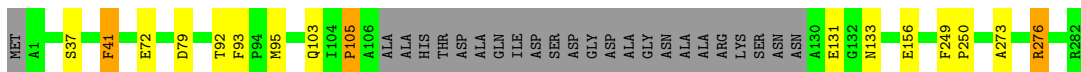
- Molecule 1: Chalcone isomerase




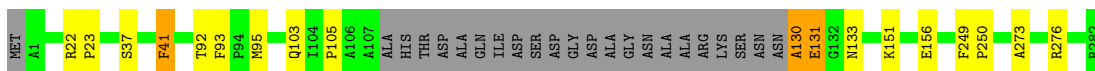


R282


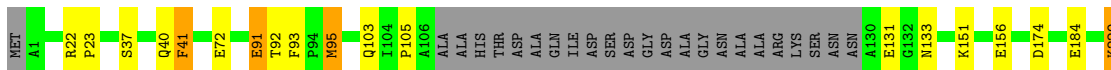
- Molecule 1: Chalcone isomerase

Chain L:  86% 5% • 8%


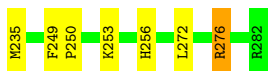
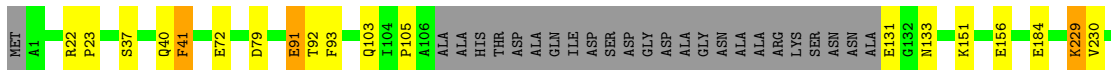
- Molecule 1: Chalcone isomerase

Chain M:  86% 5% • 8%


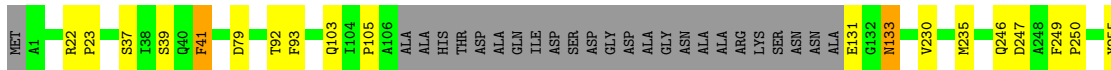
- Molecule 1: Chalcone isomerase

Chain N:  82% 8% • 8%


- Molecule 1: Chalcone isomerase

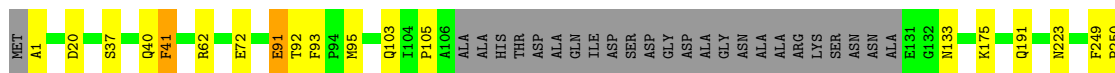
Chain O:  82% 8% • 9%

- Molecule 1: Chalcone isomerase

Chain P:  83% 7% • 9%

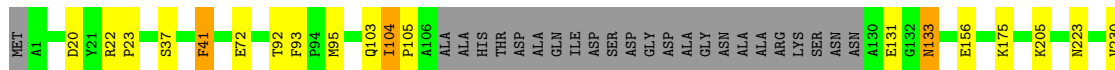
- Molecule 1: Chalcone isomerase

Chain Q:  84% 7% • 9%



- Molecule 1: Chalcone isomerase

Chain R:  82%  8%  8%



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	I 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	187.58Å 203.74Å 560.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.42 – 2.80	Depositor
% Data completeness (in resolution range)	99.4 (49.42-2.80)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 2.77Å)	Xtrriage
Refinement program	REFMAC 5.8.0352	Depositor
R, $R_{free}$	0.213 , 0.225	Depositor
Wilson B-factor (Å <sup>2</sup> )	48.4	Xtrriage
Anisotropy	0.139	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	39179	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.87% 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: CL, SO4, DQH, K, GOL

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	0.47	3/2182 (0.1%)	0.71	1/2966 (0.0%)
1	B	0.48	1/2192 (0.0%)	0.71	0/2980
1	C	0.52	4/2192 (0.2%)	0.72	0/2980
1	D	0.53	2/2192 (0.1%)	0.70	0/2980
1	E	0.47	2/2192 (0.1%)	0.71	0/2980
1	F	0.46	2/2197 (0.1%)	0.71	0/2987
1	G	0.45	1/2187 (0.0%)	0.70	0/2973
1	H	0.47	1/2192 (0.0%)	0.71	1/2980 (0.0%)
1	I	0.44	1/2192 (0.0%)	0.69	0/2980
1	J	0.44	2/2192 (0.1%)	0.69	0/2980
1	K	0.47	1/2192 (0.0%)	0.70	0/2980
1	L	0.42	1/2192 (0.0%)	0.69	0/2980
1	M	0.43	0/2197	0.68	0/2987
1	N	0.45	1/2192 (0.0%)	0.70	1/2980 (0.0%)
1	O	0.44	1/2187 (0.0%)	0.69	1/2973 (0.0%)
1	P	0.42	0/2187	0.69	0/2973
1	Q	0.47	1/2187 (0.0%)	0.71	0/2973
1	R	0.43	1/2192 (0.0%)	0.70	0/2980
All	All	0.46	25/39436 (0.1%)	0.70	4/53612 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
1	C	0	2
1	D	0	1
1	E	0	2

*Continued on next page...*

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	F	0	1
1	G	0	1
1	H	0	1
1	J	0	1
1	L	0	2
1	M	0	1
1	O	0	2
1	P	0	2
1	Q	0	2
1	R	0	2
All	All	0	22

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	72	GLU	CD-OE1	7.13	1.33	1.25
1	E	72	GLU	CD-OE1	6.77	1.33	1.25
1	D	72	GLU	CD-OE1	6.54	1.32	1.25
1	H	72	GLU	CD-OE1	6.41	1.32	1.25
1	O	72	GLU	CD-OE1	6.21	1.32	1.25

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	247	ASP	CB-CG-OD1	-7.96	111.14	118.30
1	O	229	LYS	CA-CB-CG	6.12	126.86	113.40
1	N	229	LYS	CA-CB-CG	6.11	126.85	113.40
1	A	247	ASP	CB-CG-OD1	-5.96	112.93	118.30

There are no chirality outliers.

5 of 22 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	276	ARG	Sidechain
1	B	276	ARG	Sidechain
1	C	105	PRO	Peptide
1	C	276	ARG	Sidechain
1	D	276	ARG	Sidechain

## 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	2112	0	2045	18	0
1	B	2122	0	2055	11	0
1	C	2122	0	2055	15	0
1	D	2122	0	2055	10	0
1	E	2122	0	2055	11	0
1	F	2127	0	2060	4	0
1	G	2117	0	2050	8	0
1	H	2122	0	2055	8	0
1	I	2122	0	2055	10	0
1	J	2122	0	2055	12	0
1	K	2122	0	2055	10	0
1	L	2122	0	2055	7	0
1	M	2127	0	2060	9	0
1	N	2122	0	2055	22	0
1	O	2117	0	2050	18	0
1	P	2117	0	2050	11	0
1	Q	2117	0	2050	12	0
1	R	2122	0	2055	22	0
2	A	44	0	24	4	0
2	B	44	0	24	3	0
2	C	44	0	24	0	0
2	D	44	0	24	3	0
2	E	44	0	24	3	0
2	F	44	0	24	0	0
2	G	44	0	24	0	0
2	H	44	0	22	1	0
2	I	44	0	23	2	0
2	J	22	0	11	0	0
2	K	44	0	24	3	0
2	L	44	0	24	1	0
2	M	44	0	24	0	0
2	N	44	0	24	2	0
2	O	44	0	24	3	0
2	P	44	0	24	1	0
2	Q	44	0	24	2	0
2	R	44	0	24	0	0
3	A	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1	0	0	0	0
3	C	3	0	0	2	0
3	D	2	0	0	0	0
3	E	2	0	0	0	0
3	F	3	0	0	0	0
3	G	4	0	0	0	0
3	H	2	0	0	0	0
3	I	2	0	0	0	0
3	J	1	0	0	0	0
3	K	5	0	0	1	0
3	L	3	0	0	0	0
3	M	3	0	0	0	0
3	N	5	0	0	0	0
3	O	1	0	0	0	0
3	P	2	0	0	0	0
3	Q	2	0	0	0	0
3	R	5	0	0	1	0
4	B	6	0	8	0	0
4	C	6	0	8	0	0
5	B	5	0	0	0	0
5	C	5	0	0	0	0
5	D	5	0	0	0	0
5	F	5	0	0	0	0
5	H	5	0	0	0	0
5	I	5	0	0	0	0
5	K	5	0	0	0	0
5	R	5	0	0	0	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
6	F	1	0	0	0	0
6	G	1	0	0	0	0
6	H	1	0	0	0	0
7	A	7	0	0	0	0
7	B	10	0	0	0	0
7	C	12	0	0	0	0
7	D	12	0	0	0	0
7	E	9	0	0	1	0
7	F	9	0	0	0	0
7	G	8	0	0	0	0
7	H	11	0	0	0	0
7	I	3	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	J	4	0	0	0	0
7	K	7	0	0	2	0
7	L	5	0	0	0	0
7	M	4	0	0	0	0
7	N	5	0	0	0	0
7	O	6	0	0	0	0
7	P	5	0	0	0	0
7	Q	5	0	0	0	0
7	R	6	0	0	0	0
All	All	39179	0	37402	195	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 3.

The worst 5 of 195 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:E:91:GLU:HG2	2:E:1002:DQH:C2	1.85	1.06
1:A:91:GLU:HG2	2:A:1002:DQH:C2	1.92	0.99
1:Q:91:GLU:HG2	2:Q:1002:DQH:C2	1.94	0.97
1:J:192:ALA:HB2	1:N:184:GLU:OE2	1.65	0.96
1:B:93:PHE:CE2	2:B:1002:DQH:H5	2.01	0.96

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	253/283 (89%)	250 (99%)	3 (1%)	0	100 100
1	B	255/283 (90%)	251 (98%)	4 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	255/283 (90%)	251 (98%)	4 (2%)	0	100	100
1	D	255/283 (90%)	249 (98%)	6 (2%)	0	100	100
1	E	255/283 (90%)	250 (98%)	5 (2%)	0	100	100
1	F	256/283 (90%)	250 (98%)	6 (2%)	0	100	100
1	G	254/283 (90%)	249 (98%)	5 (2%)	0	100	100
1	H	255/283 (90%)	251 (98%)	4 (2%)	0	100	100
1	I	255/283 (90%)	249 (98%)	6 (2%)	0	100	100
1	J	255/283 (90%)	250 (98%)	4 (2%)	1 (0%)	34	66
1	K	255/283 (90%)	251 (98%)	4 (2%)	0	100	100
1	L	255/283 (90%)	251 (98%)	3 (1%)	1 (0%)	34	66
1	M	256/283 (90%)	251 (98%)	4 (2%)	1 (0%)	34	66
1	N	255/283 (90%)	250 (98%)	4 (2%)	1 (0%)	34	66
1	O	254/283 (90%)	250 (98%)	4 (2%)	0	100	100
1	P	254/283 (90%)	250 (98%)	3 (1%)	1 (0%)	34	66
1	Q	254/283 (90%)	249 (98%)	4 (2%)	1 (0%)	34	66
1	R	255/283 (90%)	250 (98%)	5 (2%)	0	100	100
All	All	4586/5094 (90%)	4502 (98%)	78 (2%)	6 (0%)	51	81

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	J	105	PRO
1	M	131	GLU
1	P	105	PRO
1	Q	105	PRO
1	L	105	PRO

### 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	224/240 (93%)	216 (96%)	8 (4%)	35	69
1	B	224/240 (93%)	219 (98%)	5 (2%)	52	83
1	C	224/240 (93%)	219 (98%)	5 (2%)	52	83
1	D	224/240 (93%)	219 (98%)	5 (2%)	52	83
1	E	224/240 (93%)	218 (97%)	6 (3%)	44	78
1	F	224/240 (93%)	219 (98%)	5 (2%)	52	83
1	G	224/240 (93%)	219 (98%)	5 (2%)	52	83
1	H	224/240 (93%)	219 (98%)	5 (2%)	52	83
1	I	224/240 (93%)	219 (98%)	5 (2%)	52	83
1	J	224/240 (93%)	218 (97%)	6 (3%)	44	78
1	K	224/240 (93%)	218 (97%)	6 (3%)	44	78
1	L	224/240 (93%)	218 (97%)	6 (3%)	44	78
1	M	224/240 (93%)	218 (97%)	6 (3%)	44	78
1	N	224/240 (93%)	215 (96%)	9 (4%)	31	65
1	O	224/240 (93%)	217 (97%)	7 (3%)	40	74
1	P	224/240 (93%)	218 (97%)	6 (3%)	44	78
1	Q	224/240 (93%)	217 (97%)	7 (3%)	40	74
1	R	224/240 (93%)	217 (97%)	7 (3%)	40	74
All	All	4032/4320 (93%)	3923 (97%)	109 (3%)	44	78

5 of 109 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	93	PHE
1	M	105	PRO
1	Q	95	MET
1	K	156	GLU
1	L	131	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	R	256	HIS
1	R	133	ASN
1	P	256	HIS
1	P	133	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	Q	191	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 98 ligands modelled in this entry, 53 are monoatomic - leaving 45 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$
2	DQH	O	1002	-	24,24,24	0.67	0	36,36,36	1.01	4 (11%)
2	DQH	D	1002	-	24,24,24	0.64	0	36,36,36	1.61	5 (13%)
2	DQH	E	1002	-	24,24,24	1.12	3 (12%)	36,36,36	1.76	8 (22%)
2	DQH	G	1002	-	24,24,24	0.87	1 (4%)	36,36,36	1.13	4 (11%)
2	DQH	Q	1001	-	24,24,24	0.69	0	36,36,36	1.01	3 (8%)
2	DQH	B	1002	-	24,24,24	0.76	0	36,36,36	1.17	2 (5%)
5	SO4	B	1004	-	4,4,4	0.19	0	6,6,6	0.13	0
2	DQH	F	1002	-	24,24,24	1.05	3 (12%)	36,36,36	1.63	8 (22%)
2	DQH	P	1002	-	24,24,24	0.72	1 (4%)	36,36,36	0.89	0
2	DQH	H	1002	-	24,24,24	1.15	3 (12%)	36,36,36	1.47	7 (19%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	DQH	D	1001	-	24,24,24	0.64	0	36,36,36	0.98	3 (8%)
2	DQH	K	1002	-	24,24,24	0.79	1 (4%)	36,36,36	1.55	8 (22%)
2	DQH	I	1001	-	24,24,24	0.70	0	36,36,36	0.95	2 (5%)
2	DQH	J	1001	-	24,24,24	0.55	0	36,36,36	0.81	1 (2%)
2	DQH	R	1001	-	24,24,24	0.64	0	36,36,36	1.09	2 (5%)
4	GOL	C	1003	-	5,5,5	0.48	0	5,5,5	0.70	0
2	DQH	L	1001	-	24,24,24	0.57	0	36,36,36	1.08	2 (5%)
2	DQH	P	1001	-	24,24,24	0.70	0	36,36,36	0.93	1 (2%)
2	DQH	M	1002	-	24,24,24	0.75	1 (4%)	36,36,36	1.16	1 (2%)
5	SO4	D	1003	-	4,4,4	0.18	0	6,6,6	0.19	0
2	DQH	I	1002	-	24,24,24	0.62	0	36,36,36	1.58	9 (25%)
5	SO4	F	1003	-	4,4,4	0.29	0	6,6,6	0.13	0
2	DQH	C	1002	-	24,24,24	1.01	1 (4%)	36,36,36	1.46	7 (19%)
2	DQH	K	1001	-	24,24,24	0.63	0	36,36,36	0.99	2 (5%)
2	DQH	A	1002	-	24,24,24	0.89	1 (4%)	36,36,36	1.33	4 (11%)
2	DQH	N	1002	-	24,24,24	0.67	0	36,36,36	1.02	2 (5%)
5	SO4	H	1003	-	4,4,4	0.29	0	6,6,6	0.11	0
2	DQH	L	1002	-	24,24,24	0.78	0	36,36,36	1.01	4 (11%)
2	DQH	O	1001	-	24,24,24	0.85	1 (4%)	36,36,36	1.07	2 (5%)
5	SO4	R	1003	-	4,4,4	0.31	0	6,6,6	0.08	0
2	DQH	C	1001	-	24,24,24	0.89	2 (8%)	36,36,36	1.03	2 (5%)
2	DQH	M	1001	-	24,24,24	0.71	1 (4%)	36,36,36	0.98	3 (8%)
4	GOL	B	1003	-	5,5,5	0.13	0	5,5,5	0.32	0
2	DQH	G	1001	-	24,24,24	0.73	0	36,36,36	1.22	4 (11%)
2	DQH	E	1001	-	24,24,24	0.88	1 (4%)	36,36,36	1.23	4 (11%)
2	DQH	R	1002	-	24,24,24	0.61	0	36,36,36	0.93	2 (5%)
2	DQH	N	1001	-	24,24,24	0.58	0	36,36,36	0.95	1 (2%)
2	DQH	B	1001	-	24,24,24	0.75	0	36,36,36	0.99	3 (8%)
2	DQH	A	1001	-	24,24,24	0.81	0	36,36,36	0.98	1 (2%)
2	DQH	F	1001	-	24,24,24	0.60	0	36,36,36	0.91	1 (2%)
2	DQH	H	1001	-	24,24,24	0.69	0	36,36,36	1.10	3 (8%)
5	SO4	I	1003	-	4,4,4	0.32	0	6,6,6	0.17	0
5	SO4	K	1003	-	4,4,4	0.23	0	6,6,6	0.13	0
5	SO4	C	1004	-	4,4,4	0.32	0	6,6,6	0.12	0
2	DQH	Q	1002	-	24,24,24	0.87	0	36,36,36	1.04	3 (8%)

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	DQH	O	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	D	1002	-	-	2/4/20/20	0/3/3/3
2	DQH	E	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	G	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	Q	1001	-	-	4/4/20/20	0/3/3/3
2	DQH	B	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	F	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	P	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	H	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	D	1001	-	-	4/4/20/20	0/3/3/3
2	DQH	K	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	I	1001	-	-	0/4/20/20	0/3/3/3
2	DQH	J	1001	-	-	3/4/20/20	0/3/3/3
2	DQH	R	1001	-	-	4/4/20/20	0/3/3/3
4	GOL	C	1003	-	-	2/4/4/4	-
2	DQH	L	1001	-	-	4/4/20/20	0/3/3/3
2	DQH	P	1001	-	-	4/4/20/20	0/3/3/3
2	DQH	M	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	I	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	C	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	K	1001	-	-	3/4/20/20	0/3/3/3
2	DQH	A	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	N	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	L	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	O	1001	-	-	4/4/20/20	0/3/3/3
2	DQH	C	1001	-	-	4/4/20/20	0/3/3/3
2	DQH	M	1001	-	-	4/4/20/20	0/3/3/3
4	GOL	B	1003	-	-	4/4/4/4	-
2	DQH	G	1001	-	-	4/4/20/20	0/3/3/3
2	DQH	E	1001	-	-	4/4/20/20	0/3/3/3
2	DQH	R	1002	-	-	0/4/20/20	0/3/3/3
2	DQH	N	1001	-	-	4/4/20/20	0/3/3/3
2	DQH	B	1001	-	-	2/4/20/20	0/3/3/3
2	DQH	A	1001	-	-	4/4/20/20	0/3/3/3
2	DQH	F	1001	-	-	3/4/20/20	0/3/3/3

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	DQH	H	1001	-	-	4/4/20/20	0/3/3/3
2	DQH	Q	1002	-	-	0/4/20/20	0/3/3/3

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	1002	DQH	C3-C9	2.63	1.52	1.46
2	O	1001	DQH	O13-C9	2.59	1.25	1.22
2	E	1002	DQH	C3-C2	2.59	1.45	1.41
2	E	1002	DQH	C3-C9	2.40	1.52	1.46
2	H	1002	DQH	O13-C9	2.39	1.25	1.22

The worst 5 of 118 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	1002	DQH	C14-C11-C10	-6.07	101.90	114.85
2	E	1002	DQH	O23-C18-C17	4.86	131.41	118.45
2	E	1002	DQH	C2-C3-C9	4.45	125.38	120.61
2	M	1002	DQH	C14-C11-C10	-4.42	105.40	114.85
2	F	1002	DQH	C2-C3-C9	4.21	125.12	120.61

There are no chirality outliers.

5 of 71 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	1003	GOL	O1-C1-C2-C3
4	B	1003	GOL	C1-C2-C3-O3
4	C	1003	GOL	O1-C1-C2-C3
4	B	1003	GOL	O2-C2-C3-O3
4	C	1003	GOL	O1-C1-C2-O2

There are no ring outliers.

14 monomers are involved in 28 short contacts:

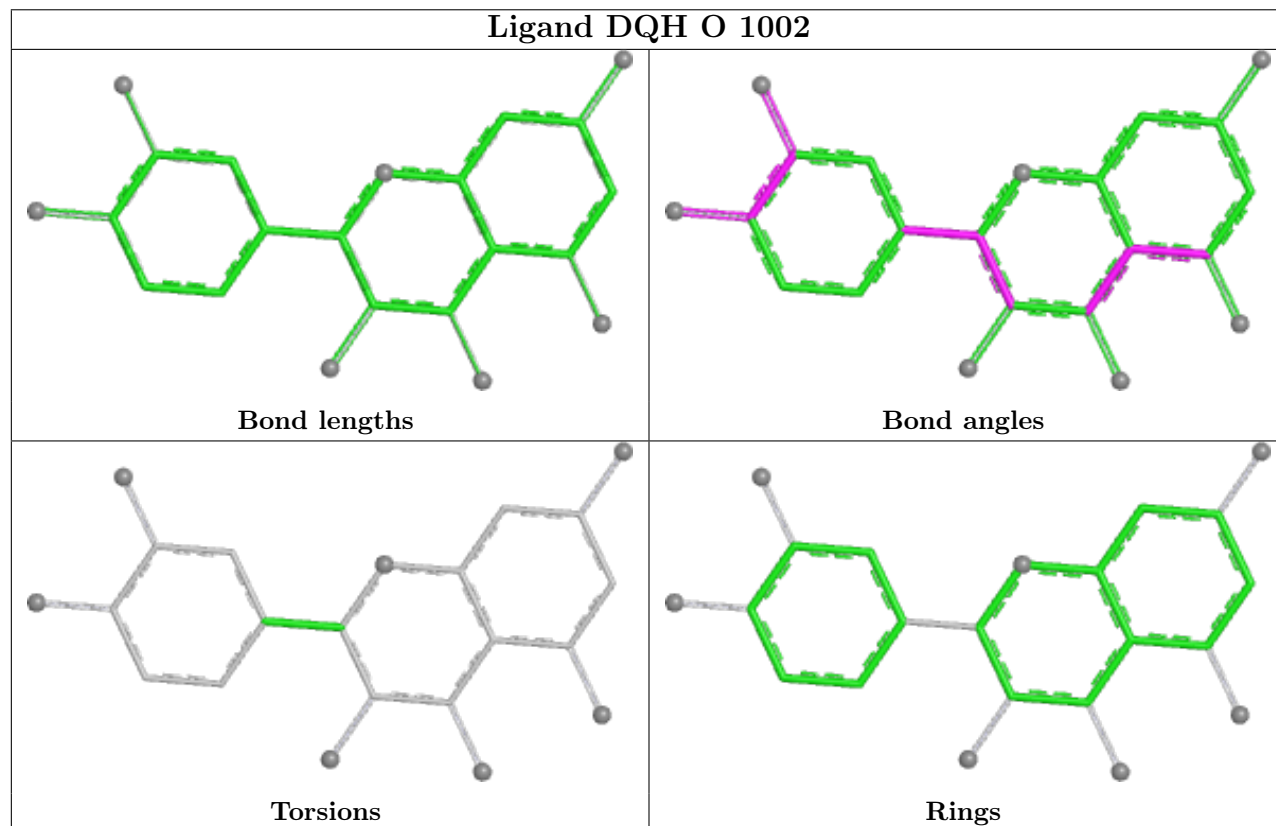
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	O	1002	DQH	2	0
2	D	1002	DQH	3	0
2	E	1002	DQH	3	0
2	B	1002	DQH	3	0
2	K	1002	DQH	3	0
2	L	1001	DQH	1	0

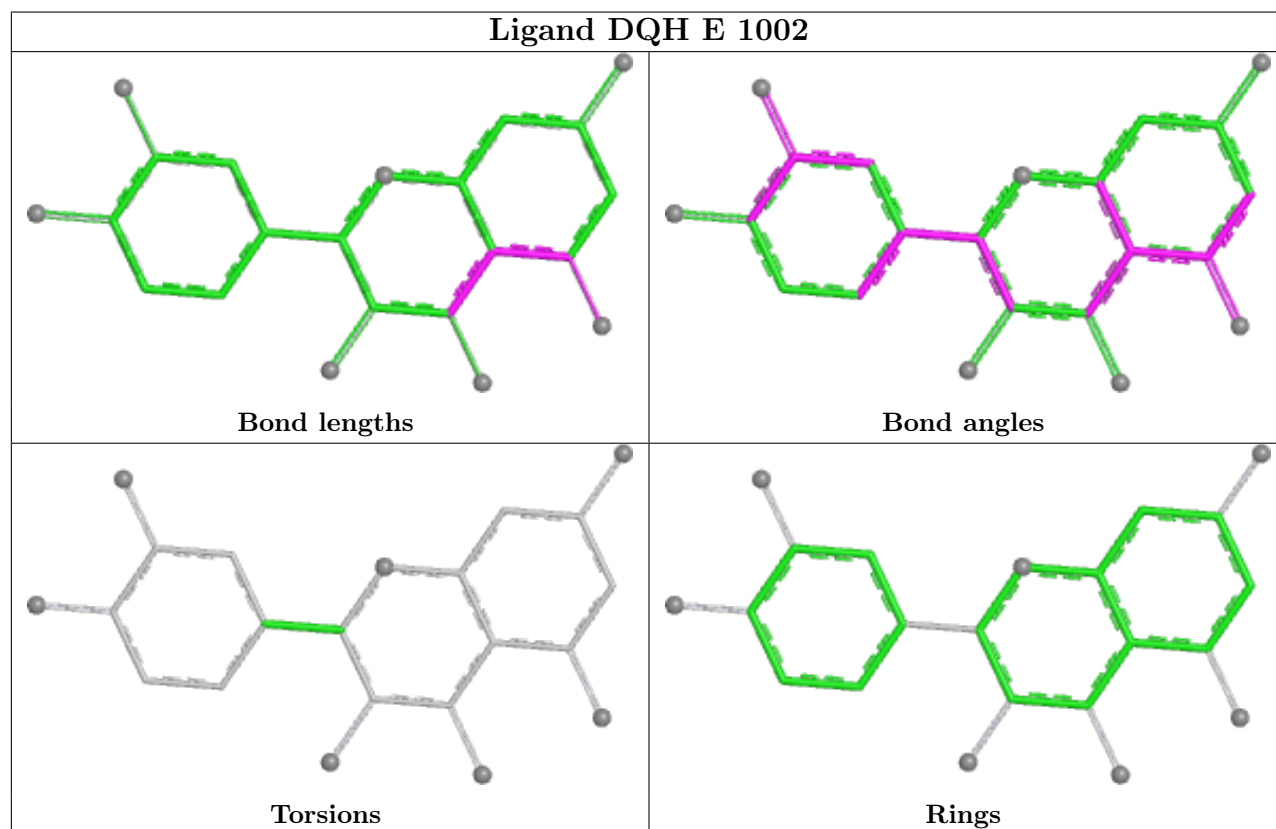
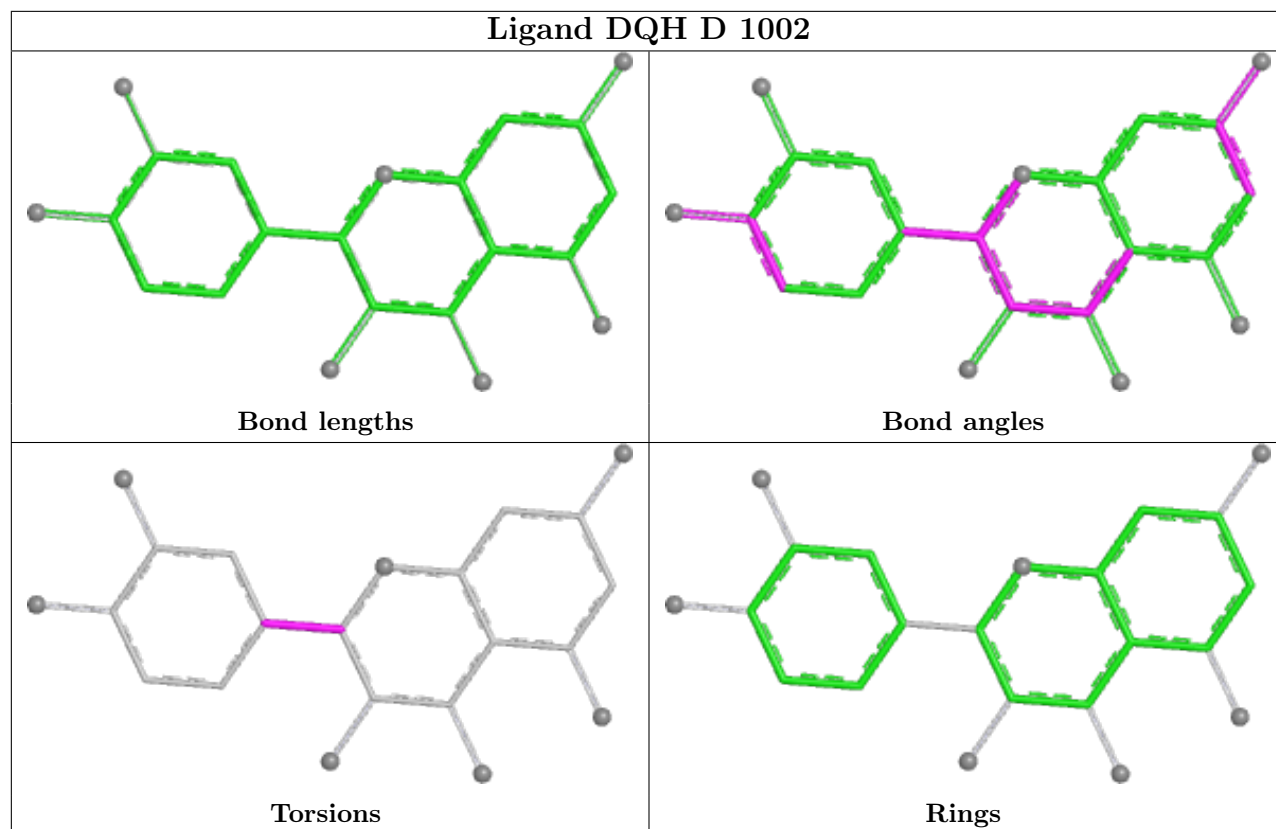
*Continued on next page...*

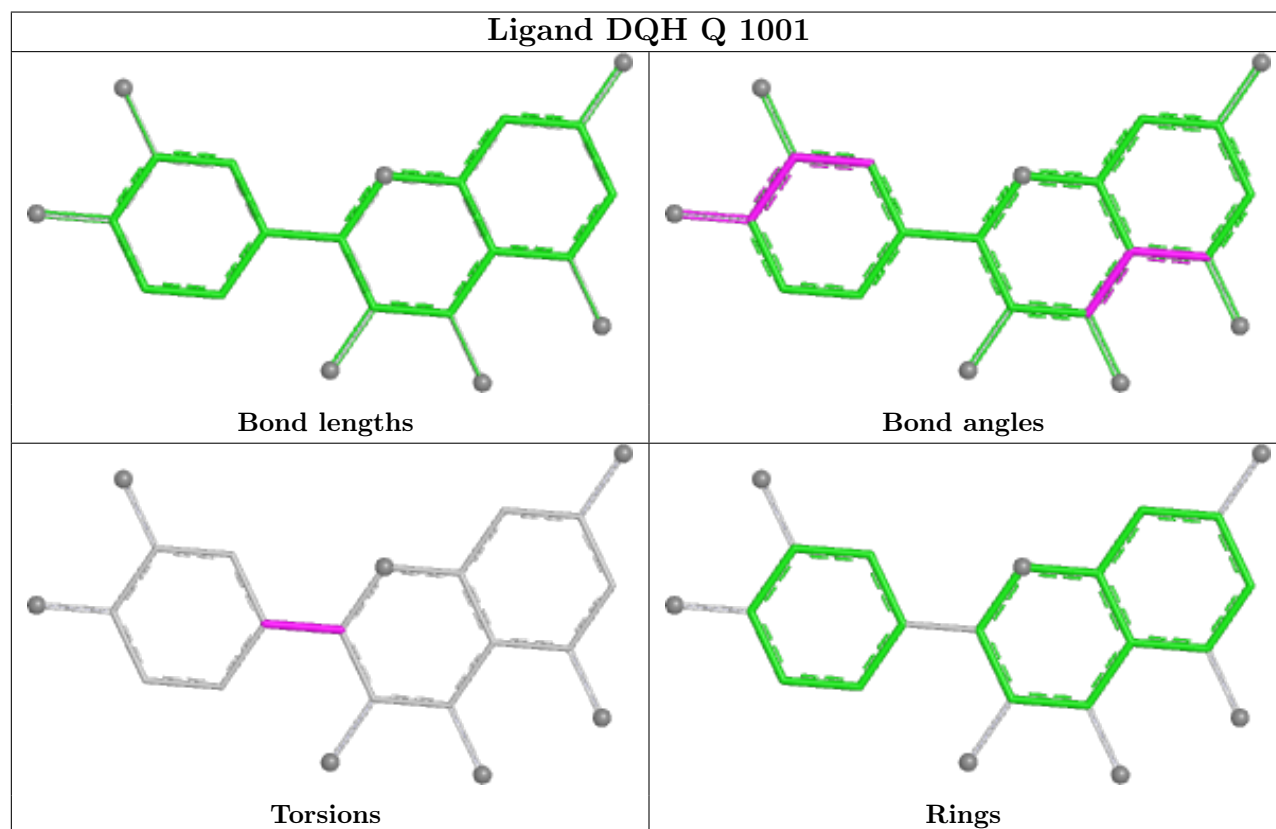
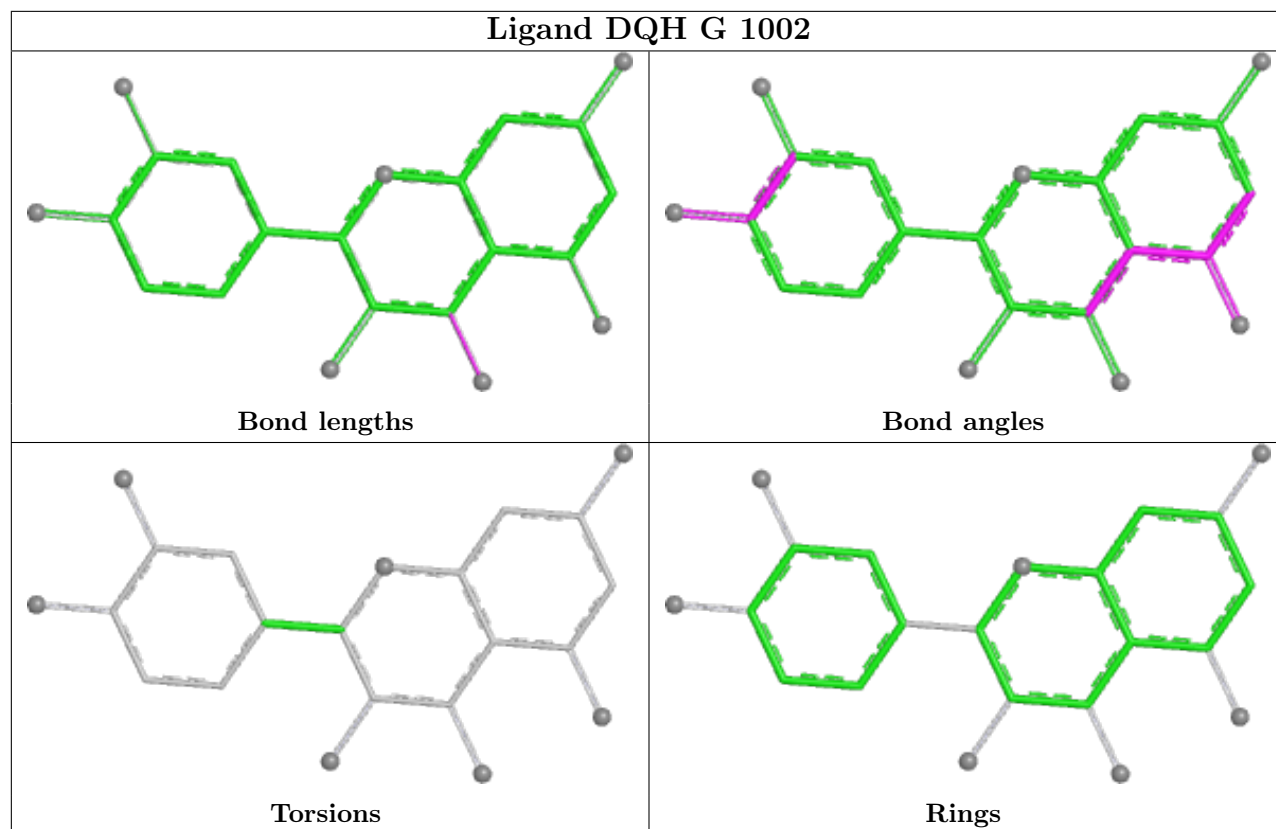
*Continued from previous page...*

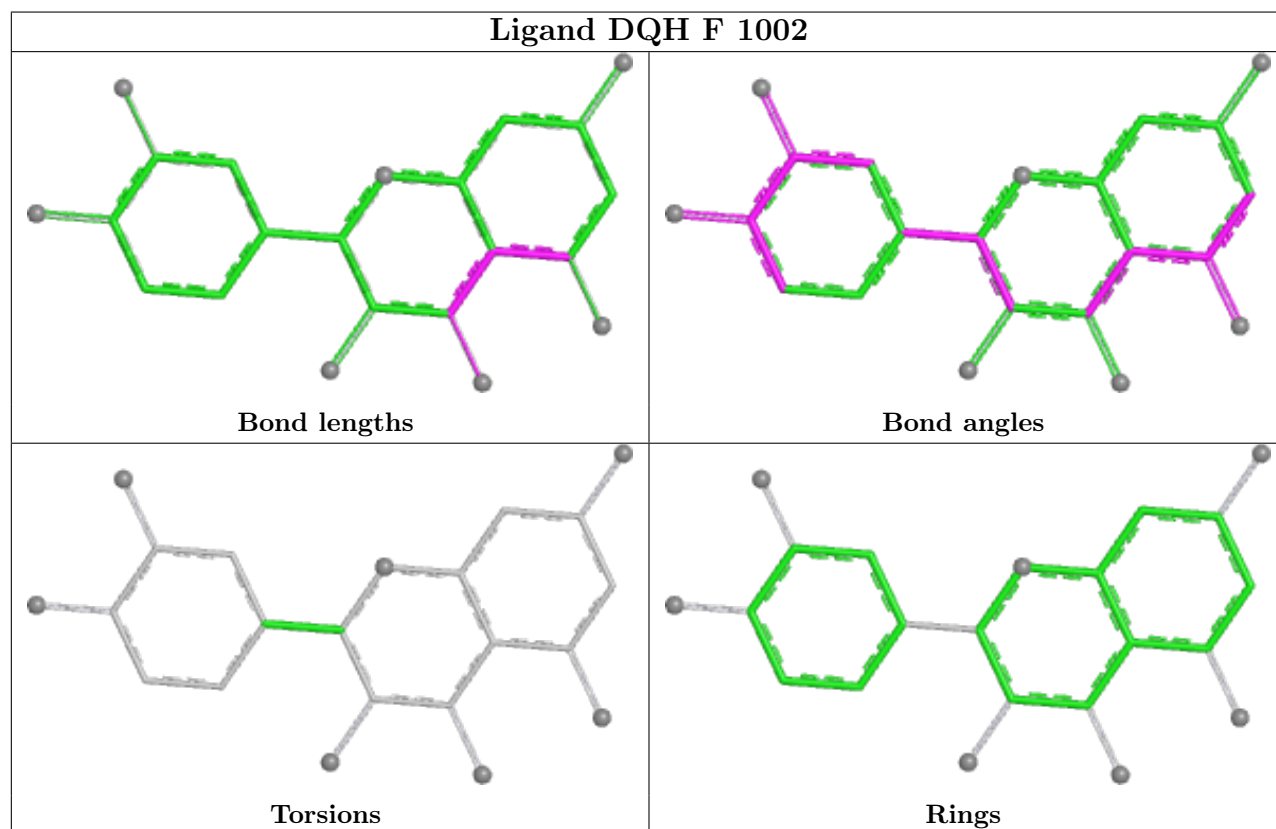
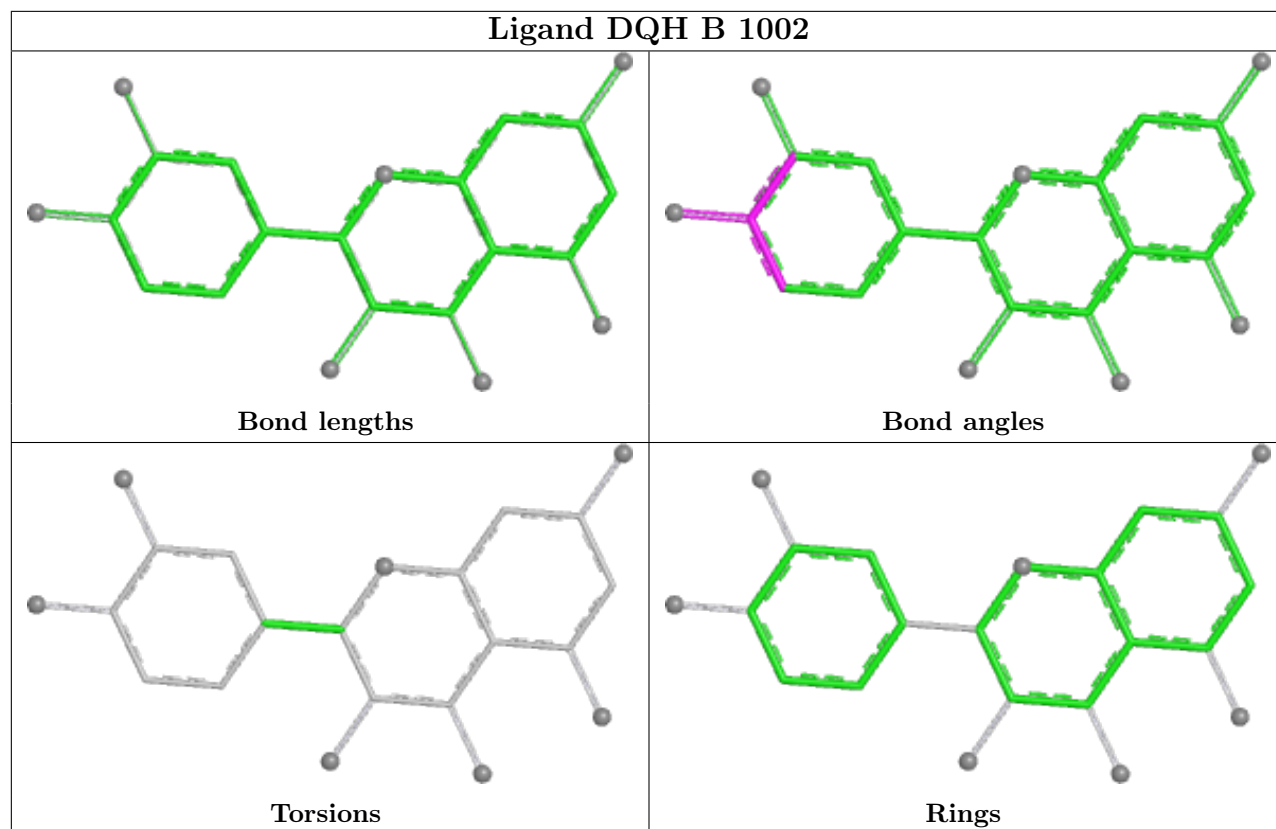
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	P	1001	DQH	1	0
2	I	1002	DQH	2	0
2	A	1002	DQH	3	0
2	N	1002	DQH	2	0
2	O	1001	DQH	1	0
2	A	1001	DQH	1	0
2	H	1001	DQH	1	0
2	Q	1002	DQH	2	0

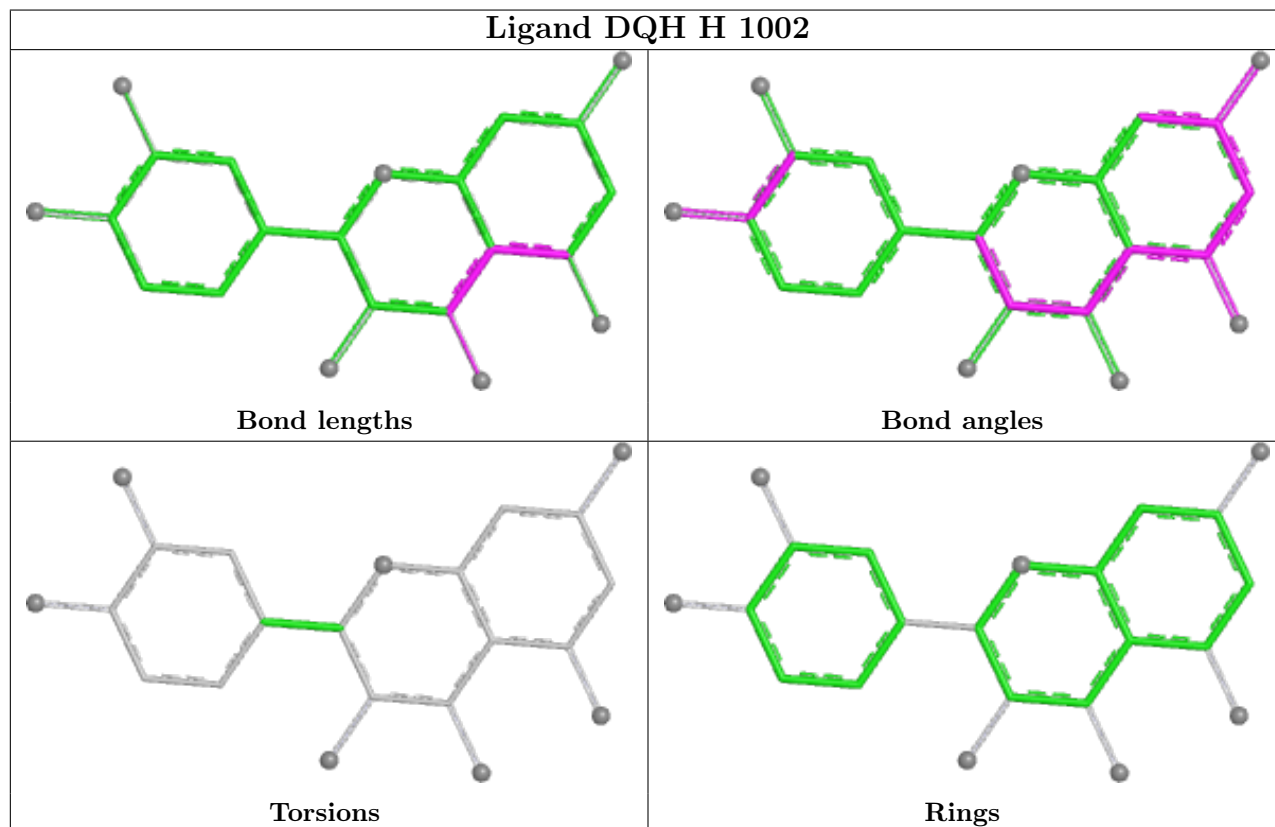
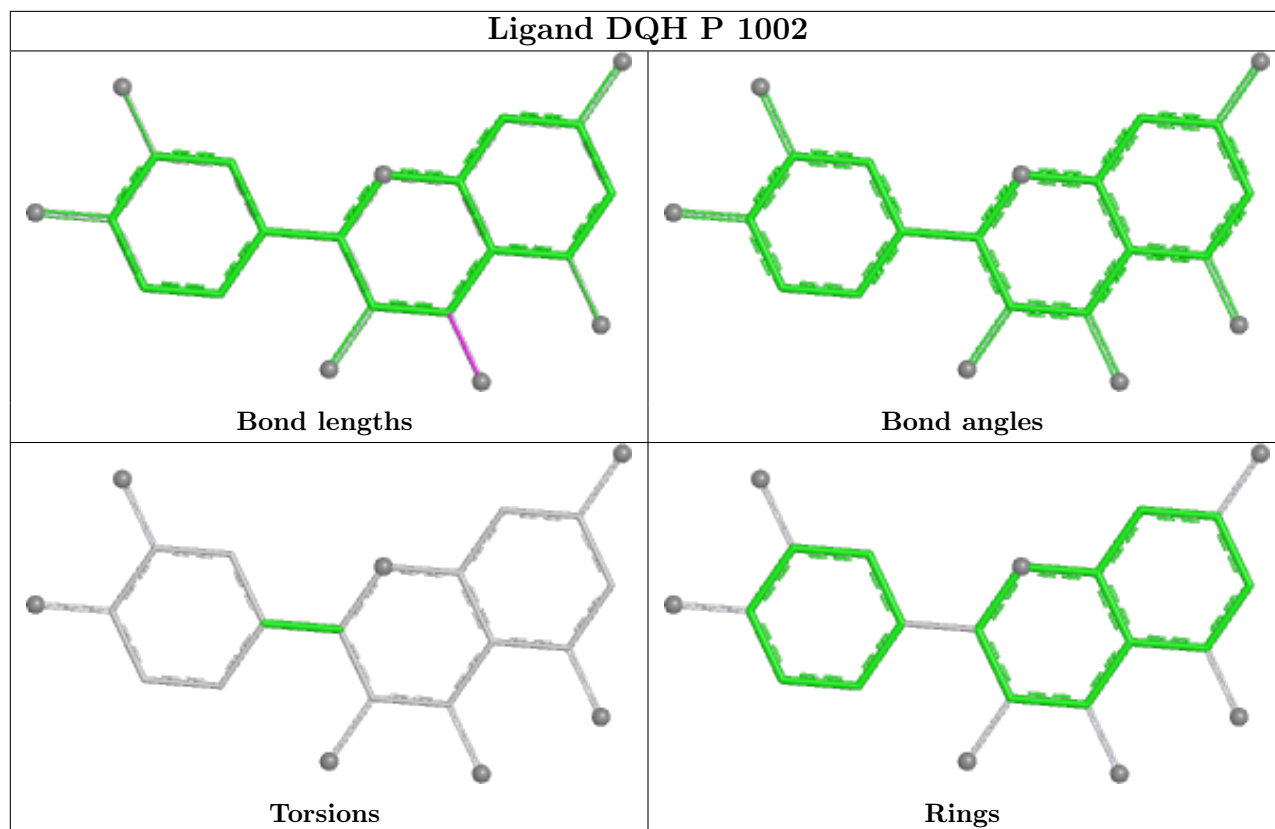
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.



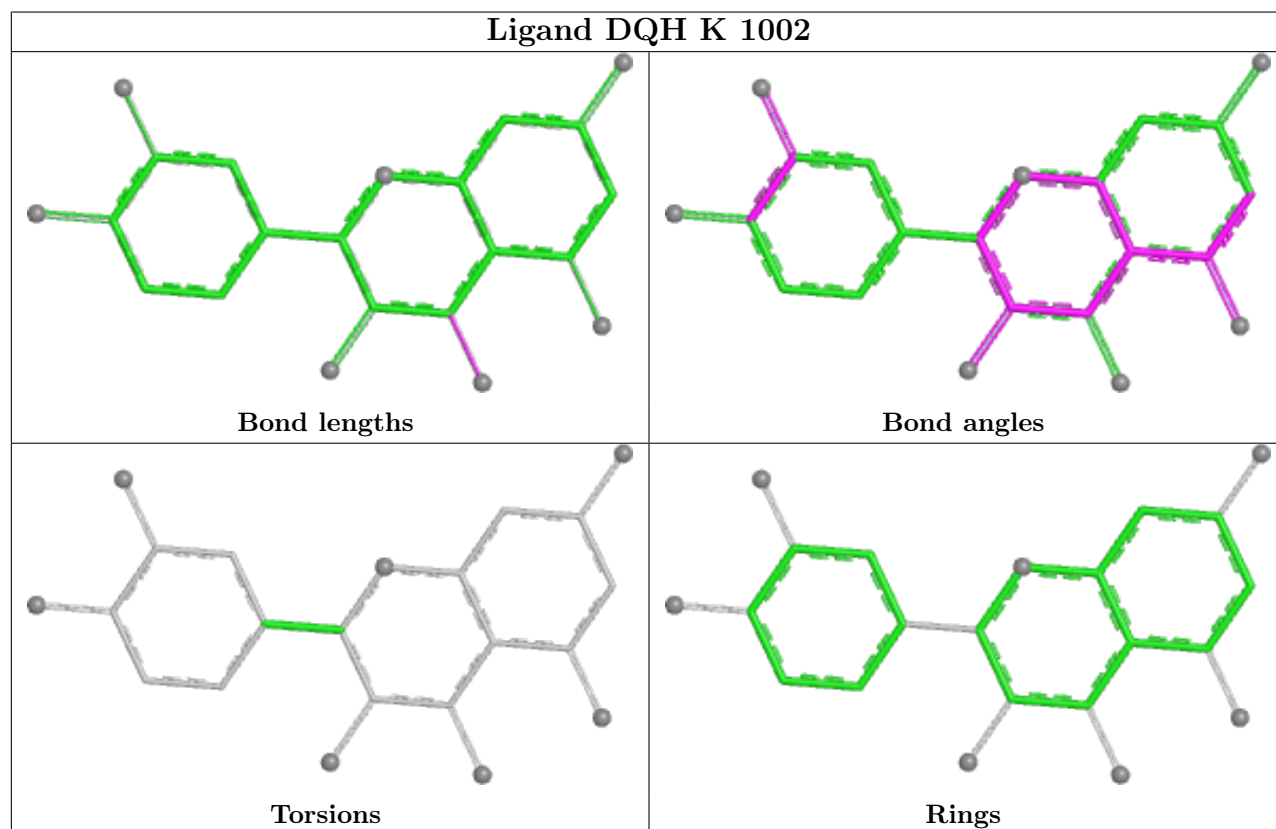
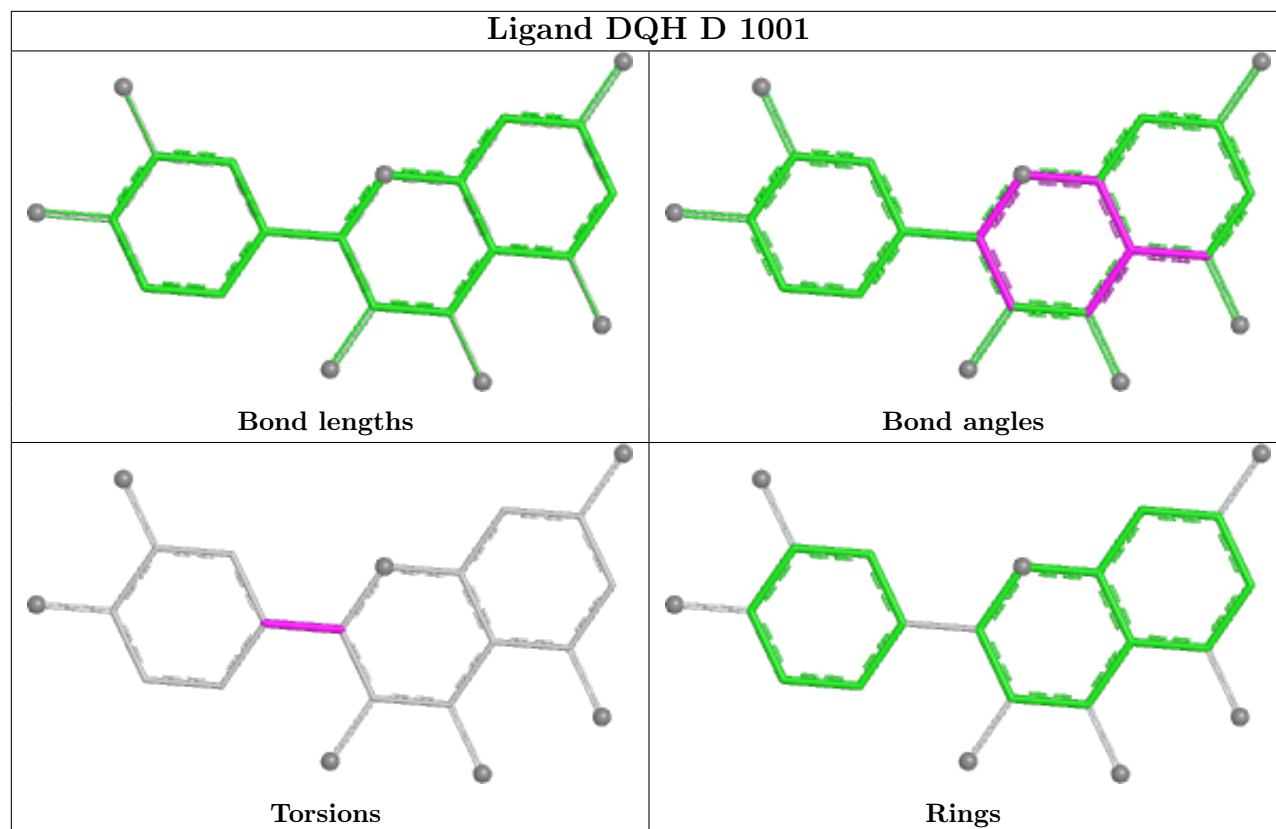


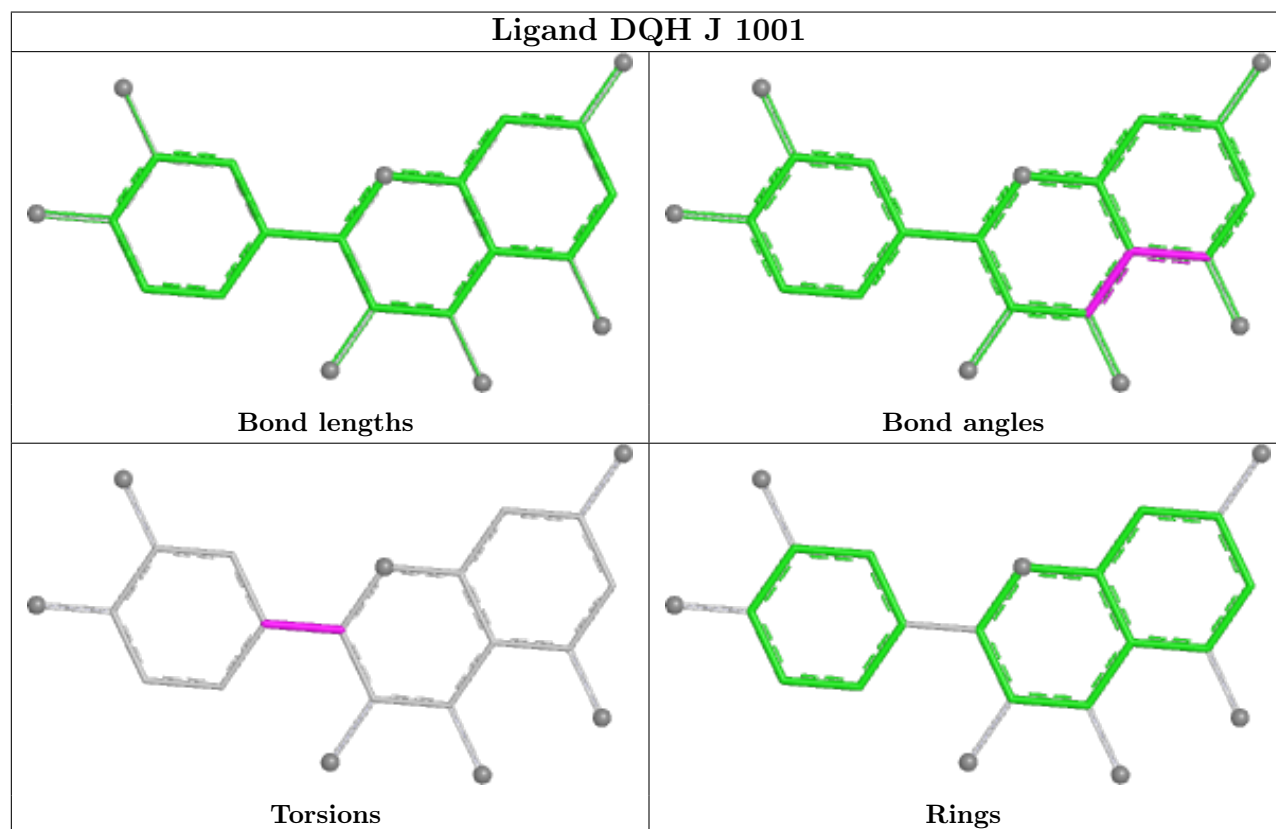
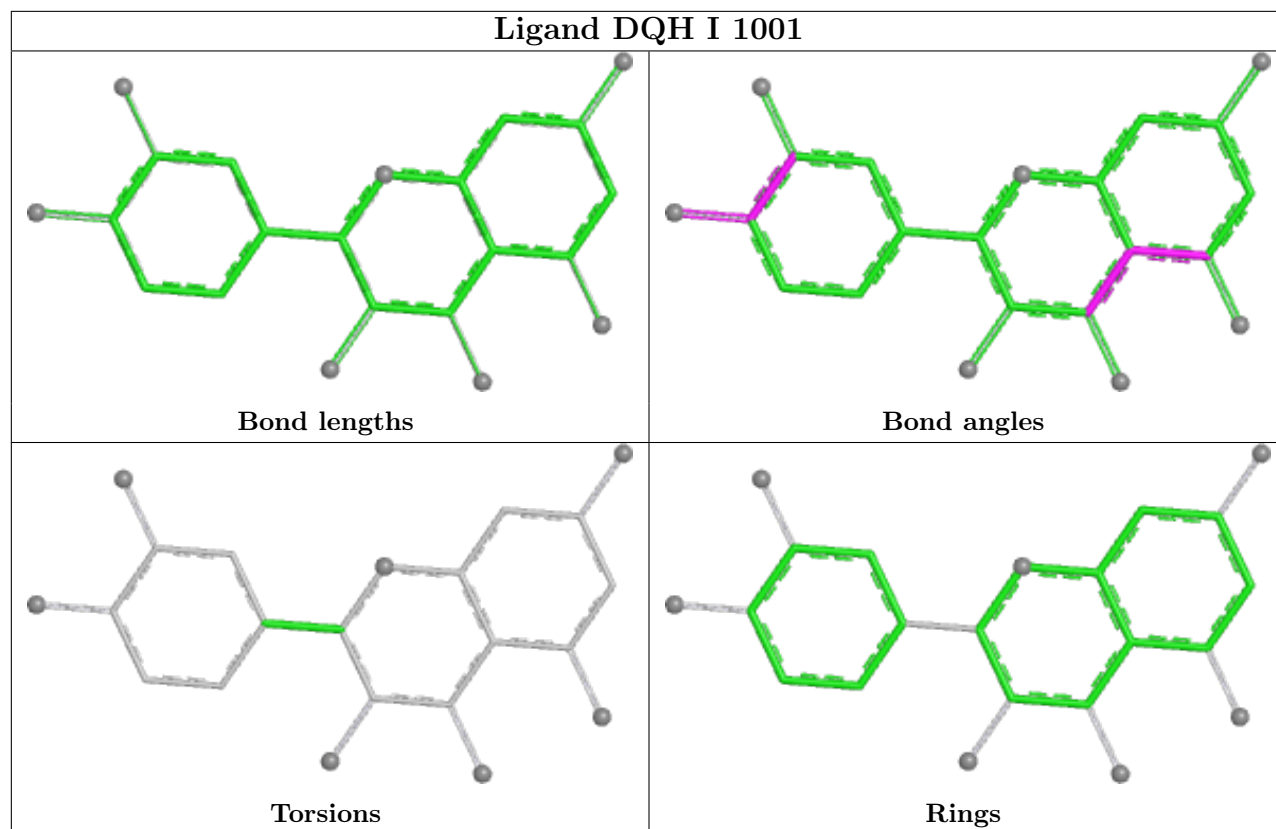


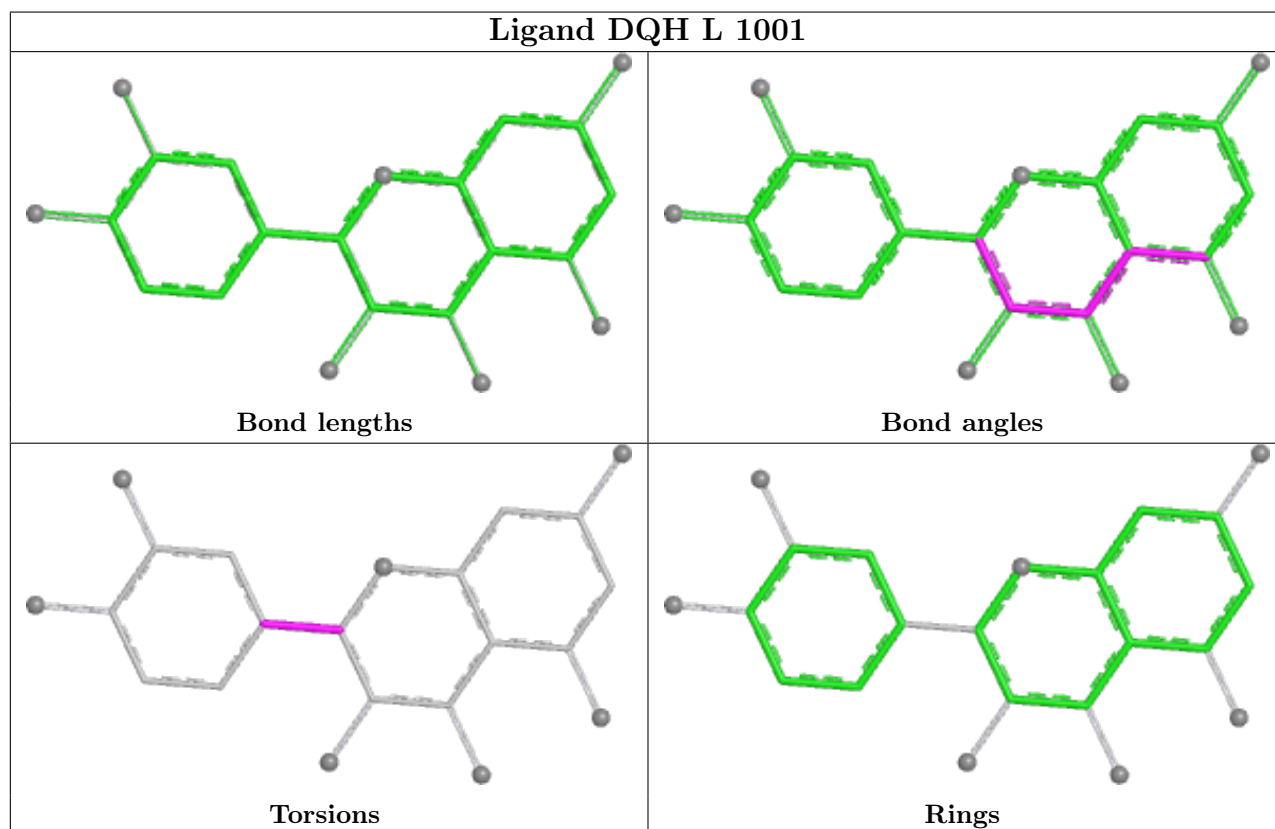
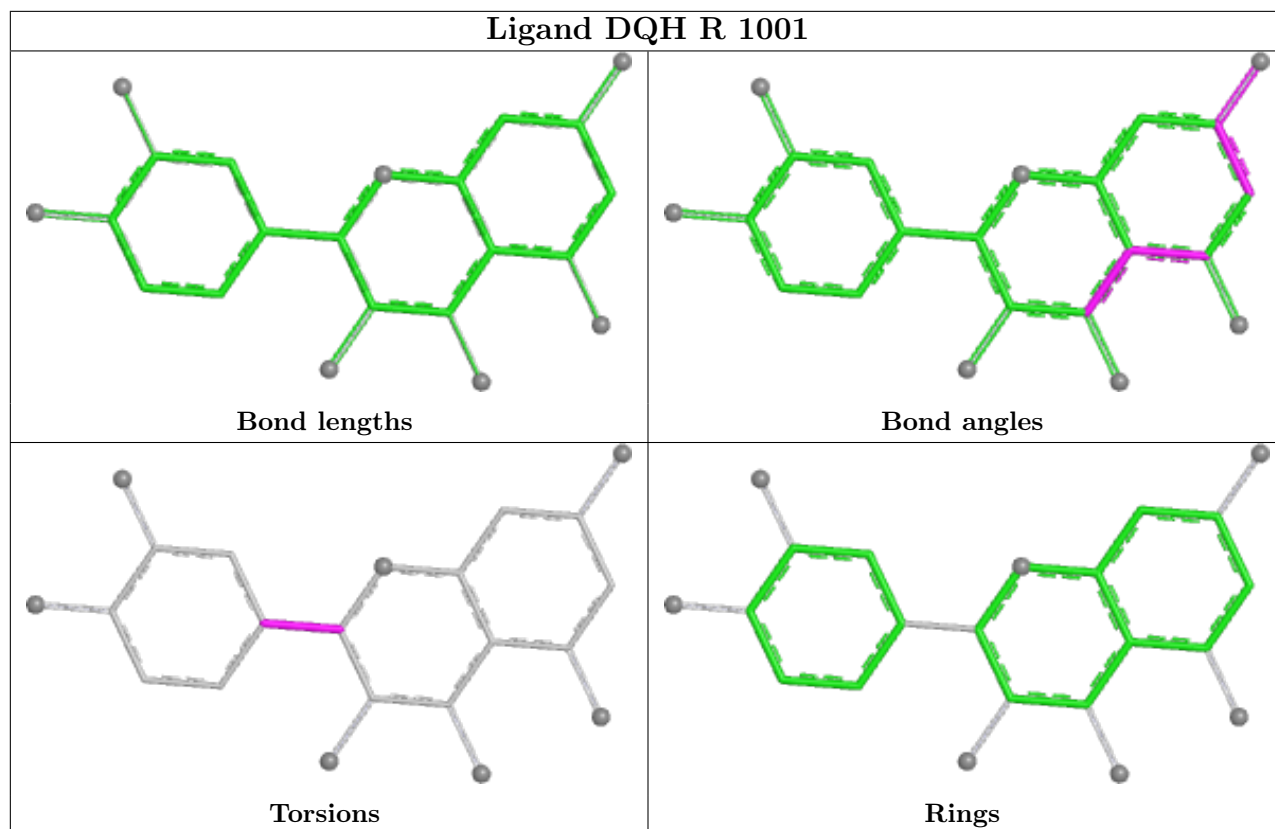


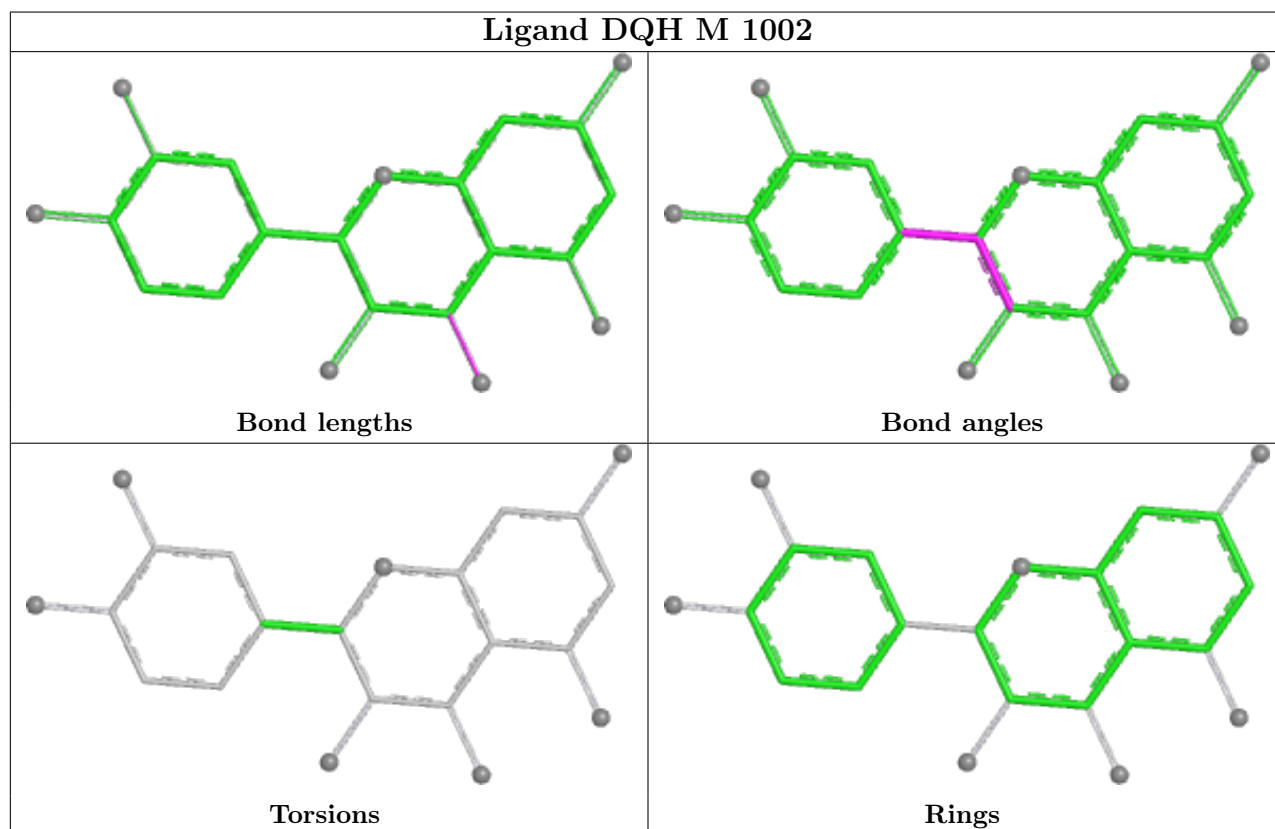
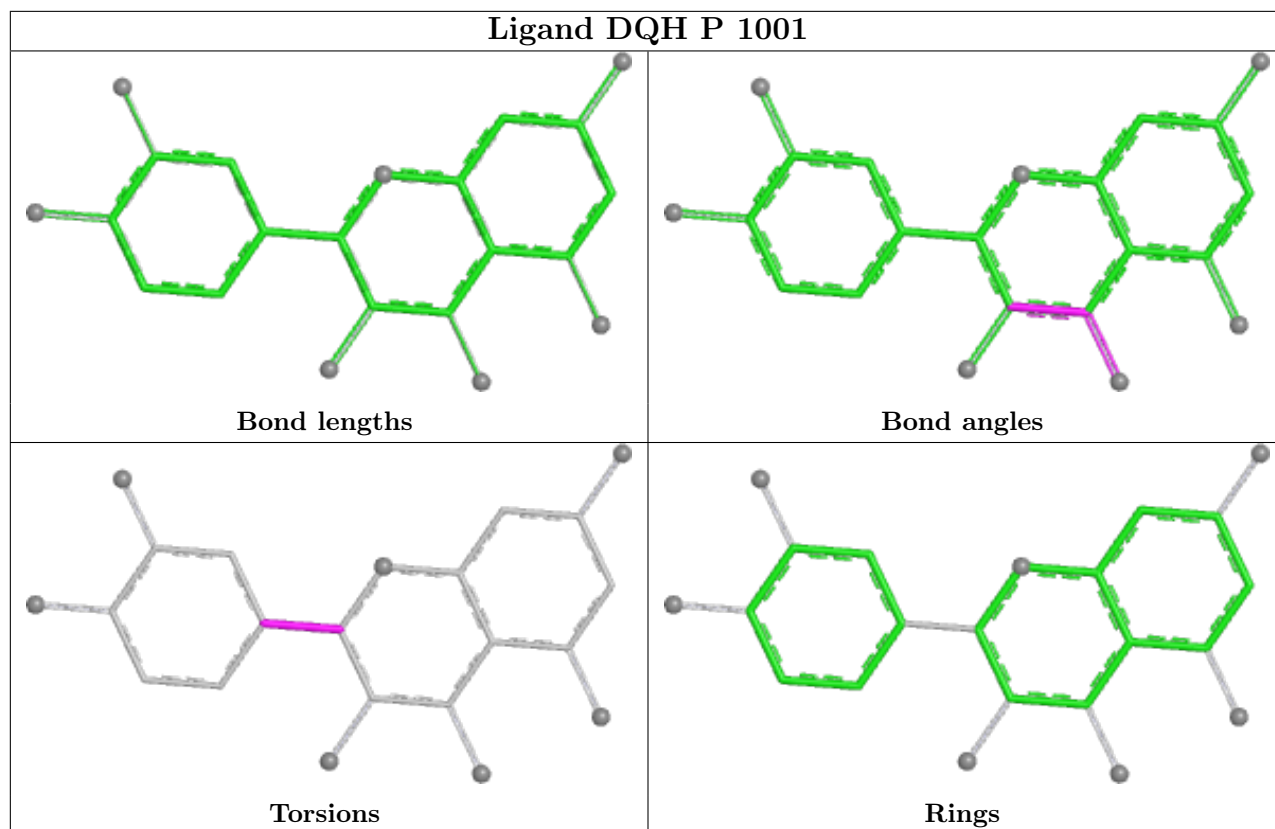


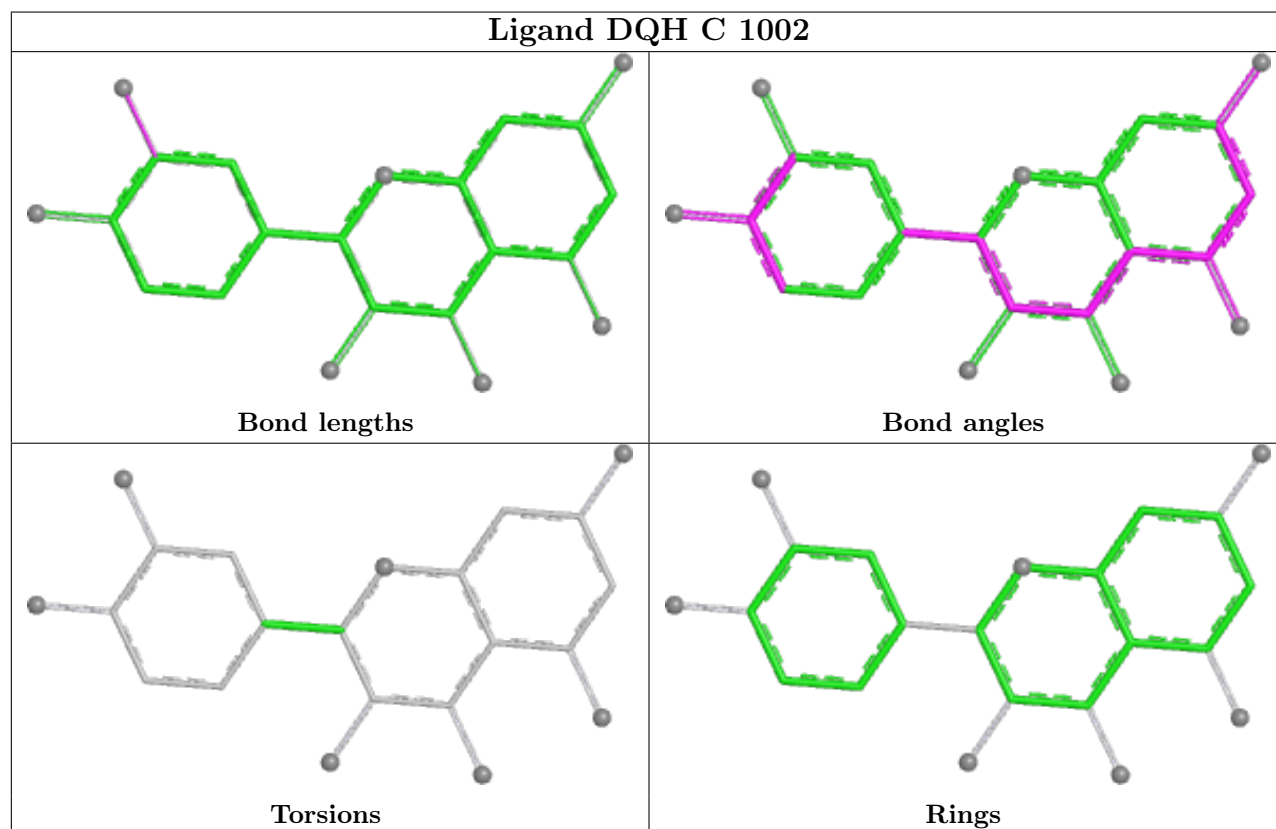
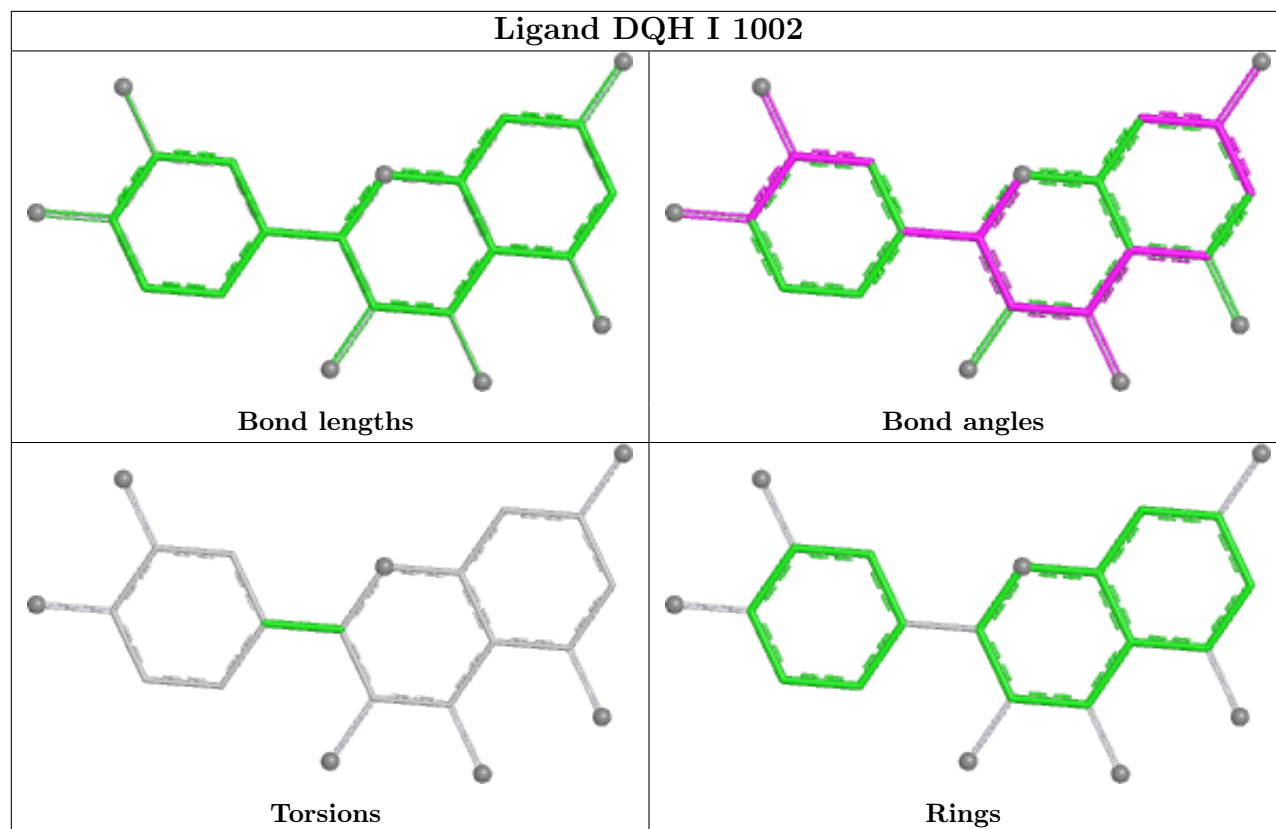


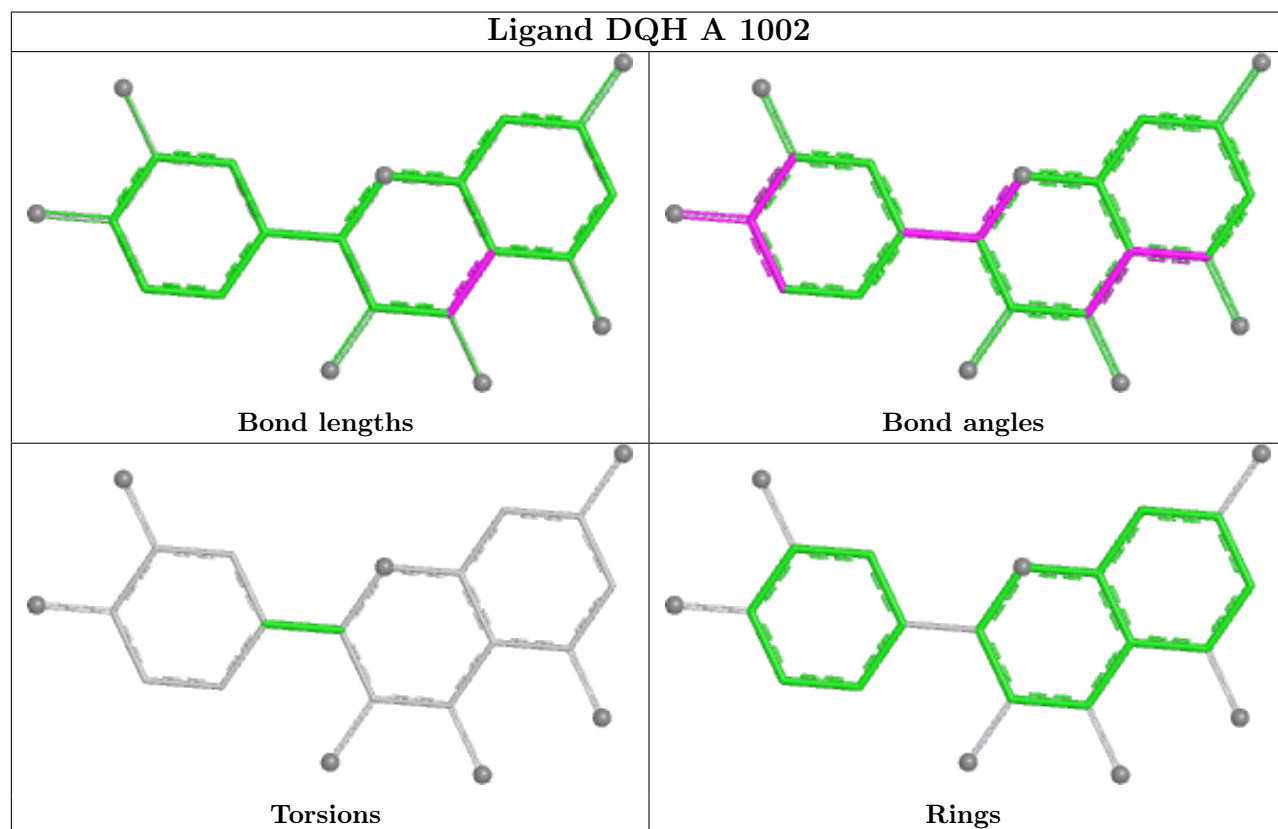
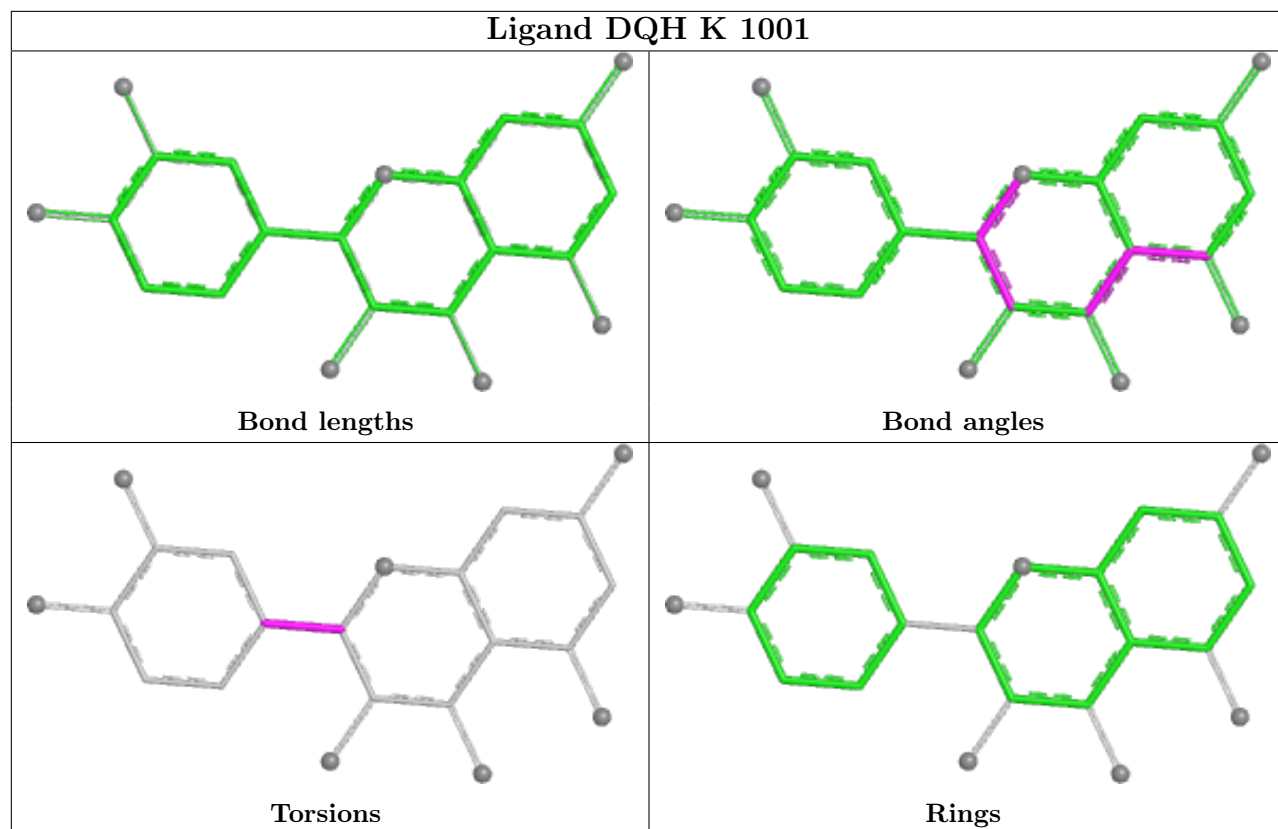


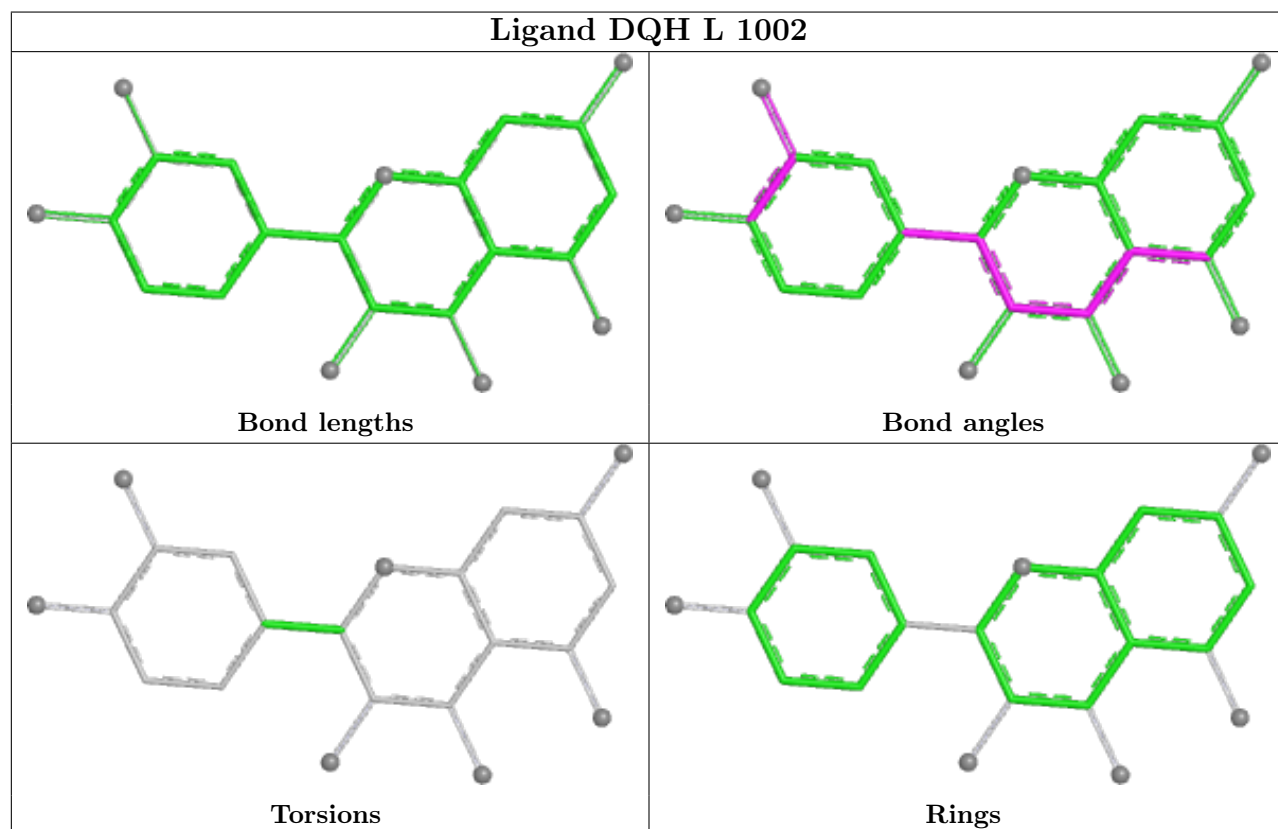
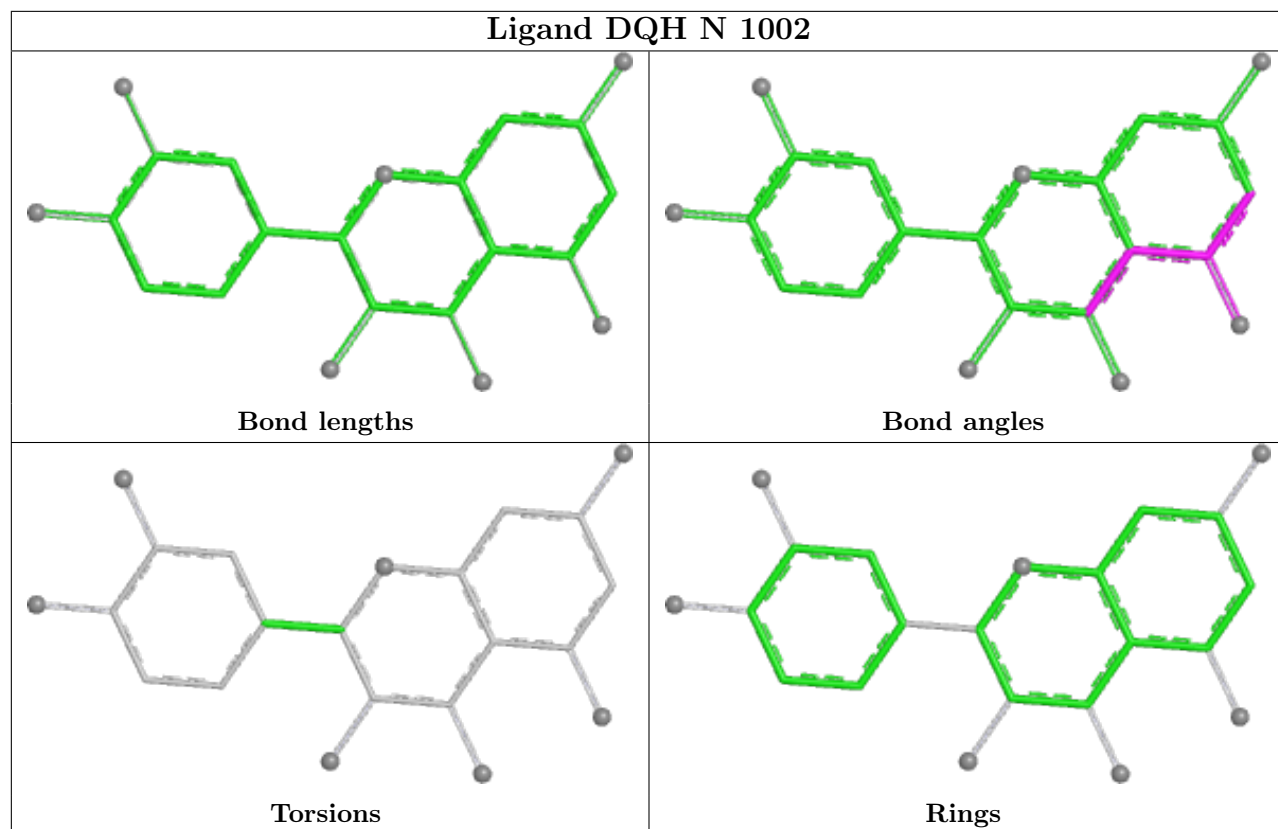


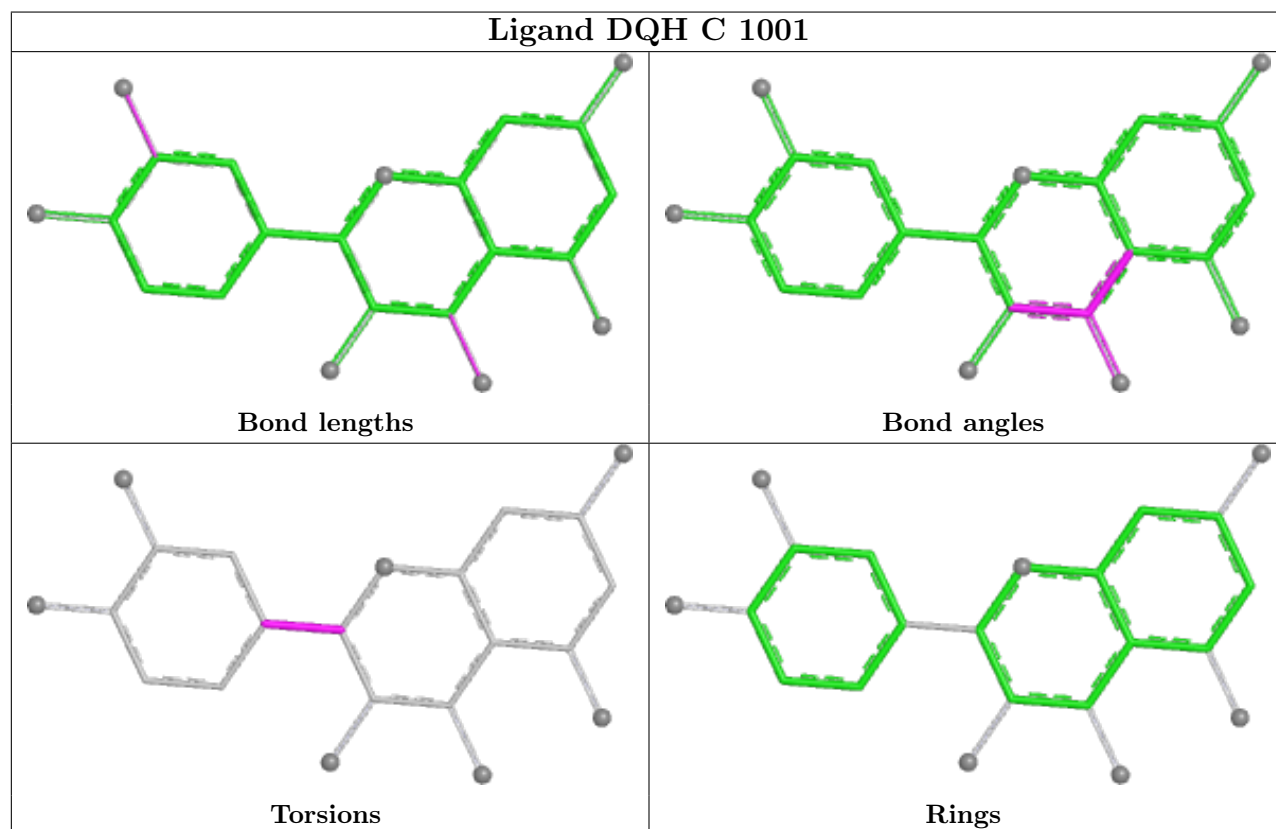
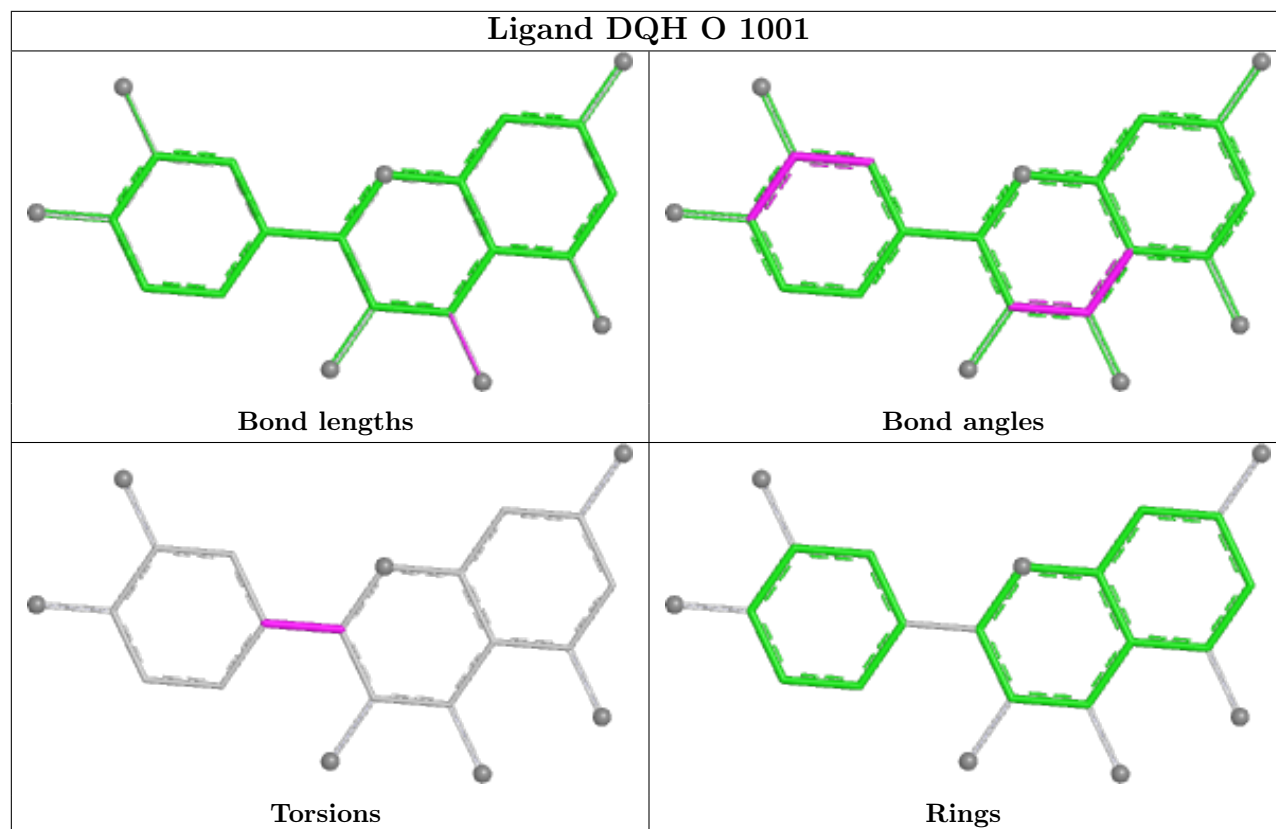




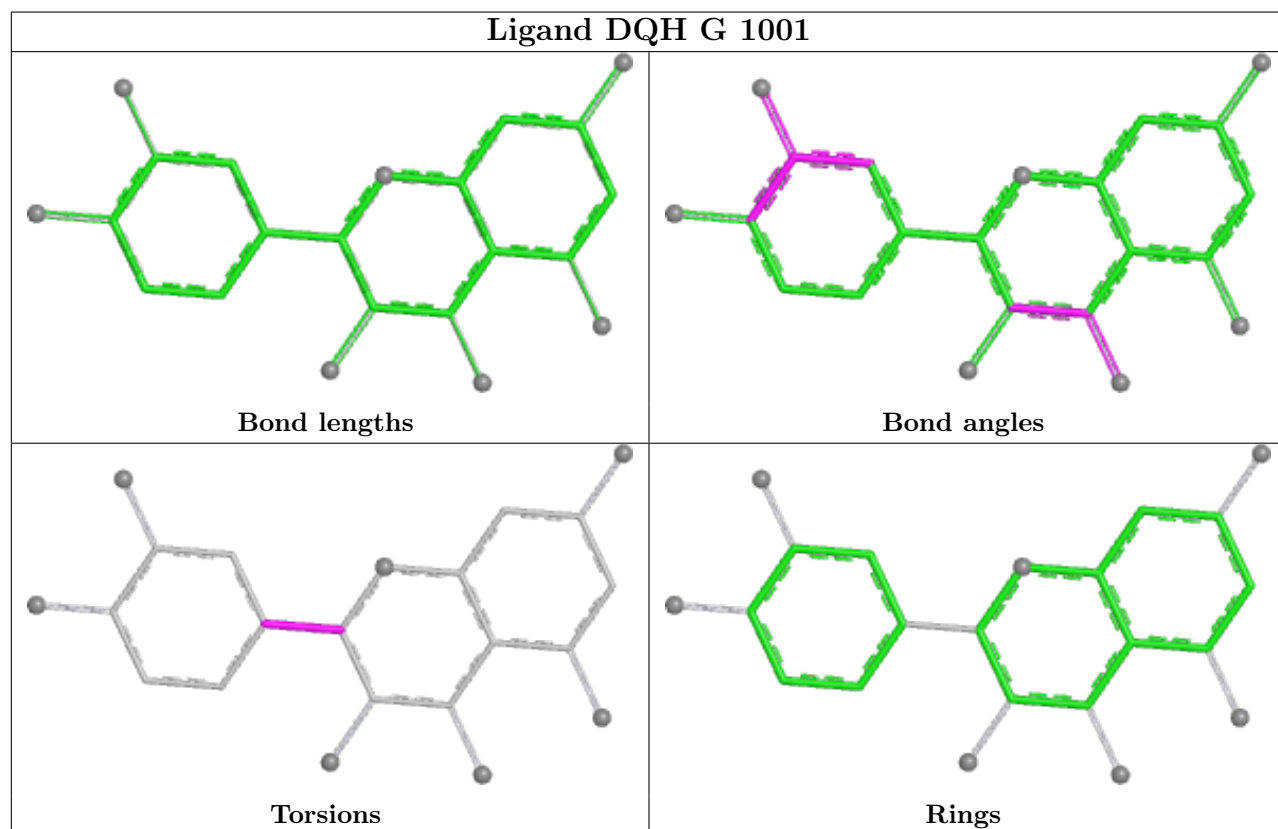
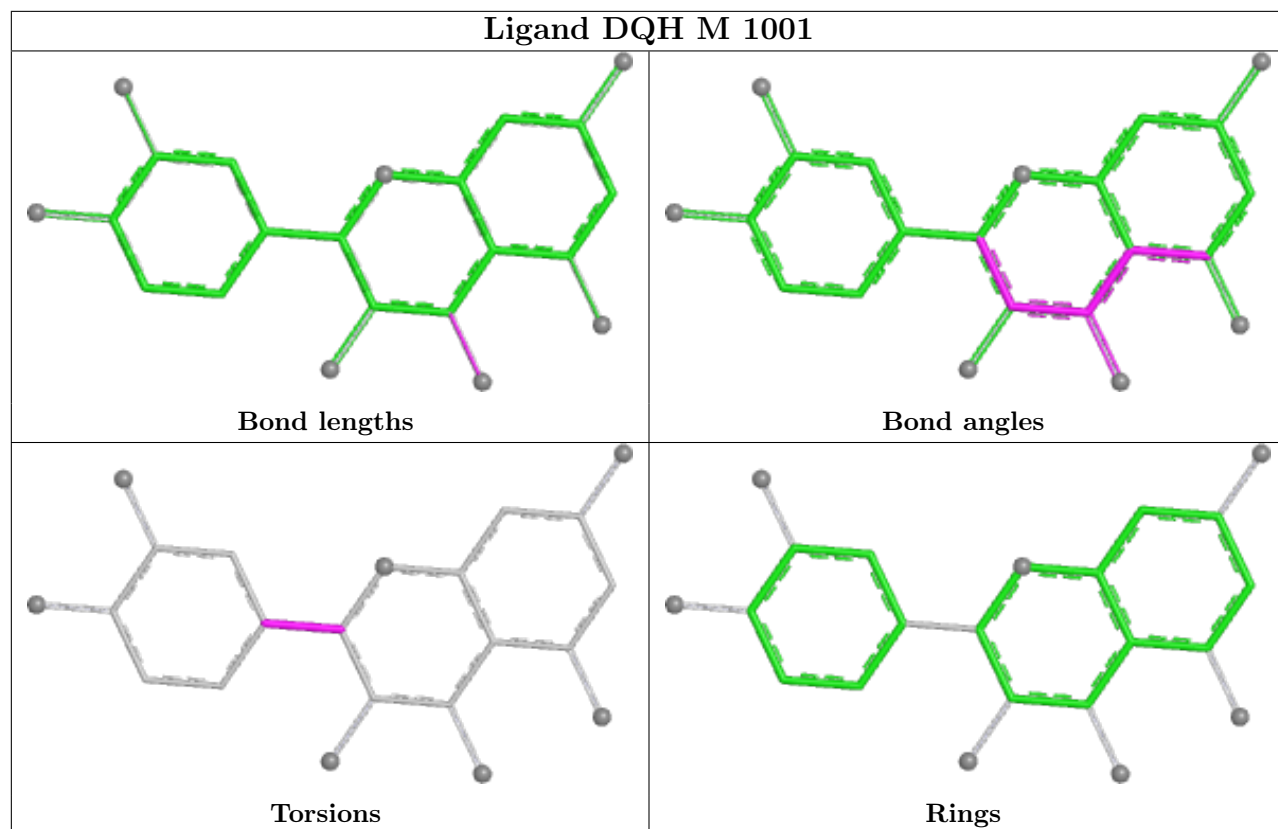


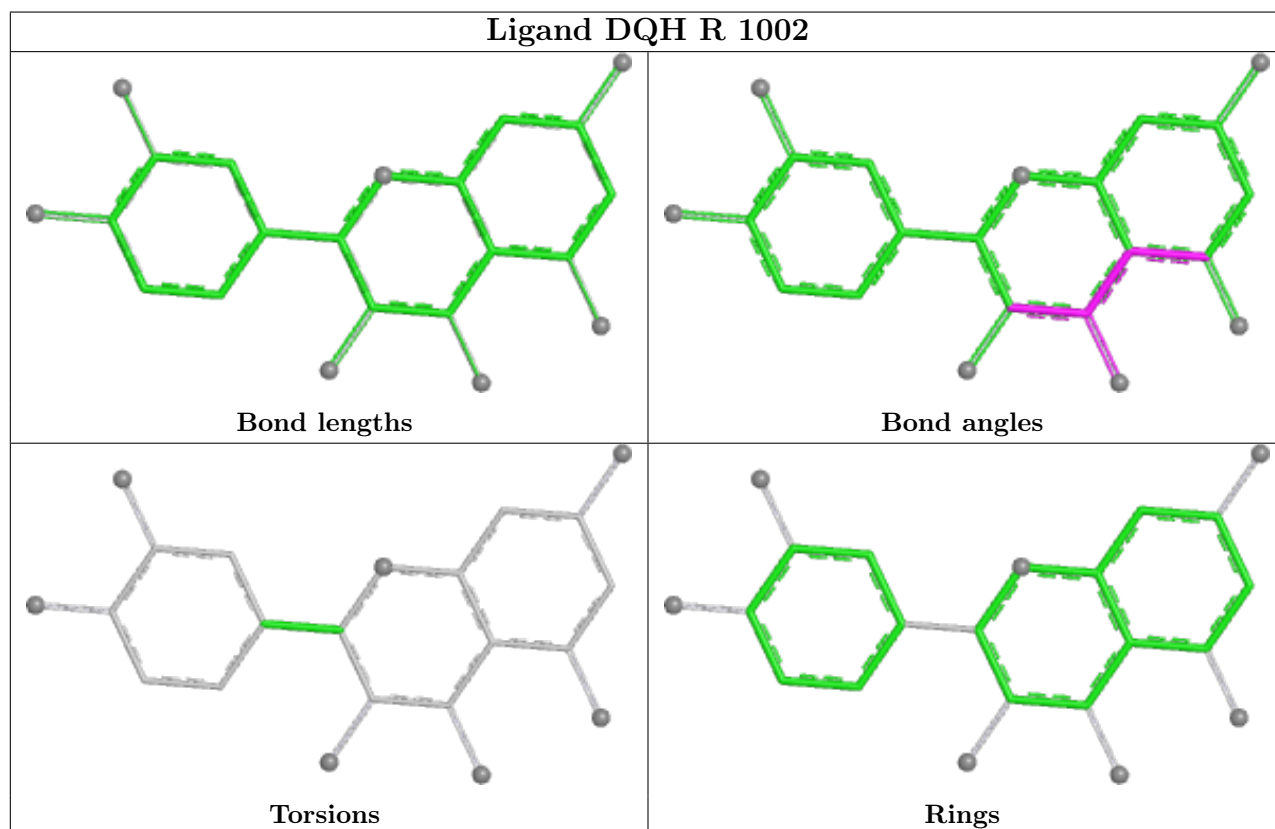
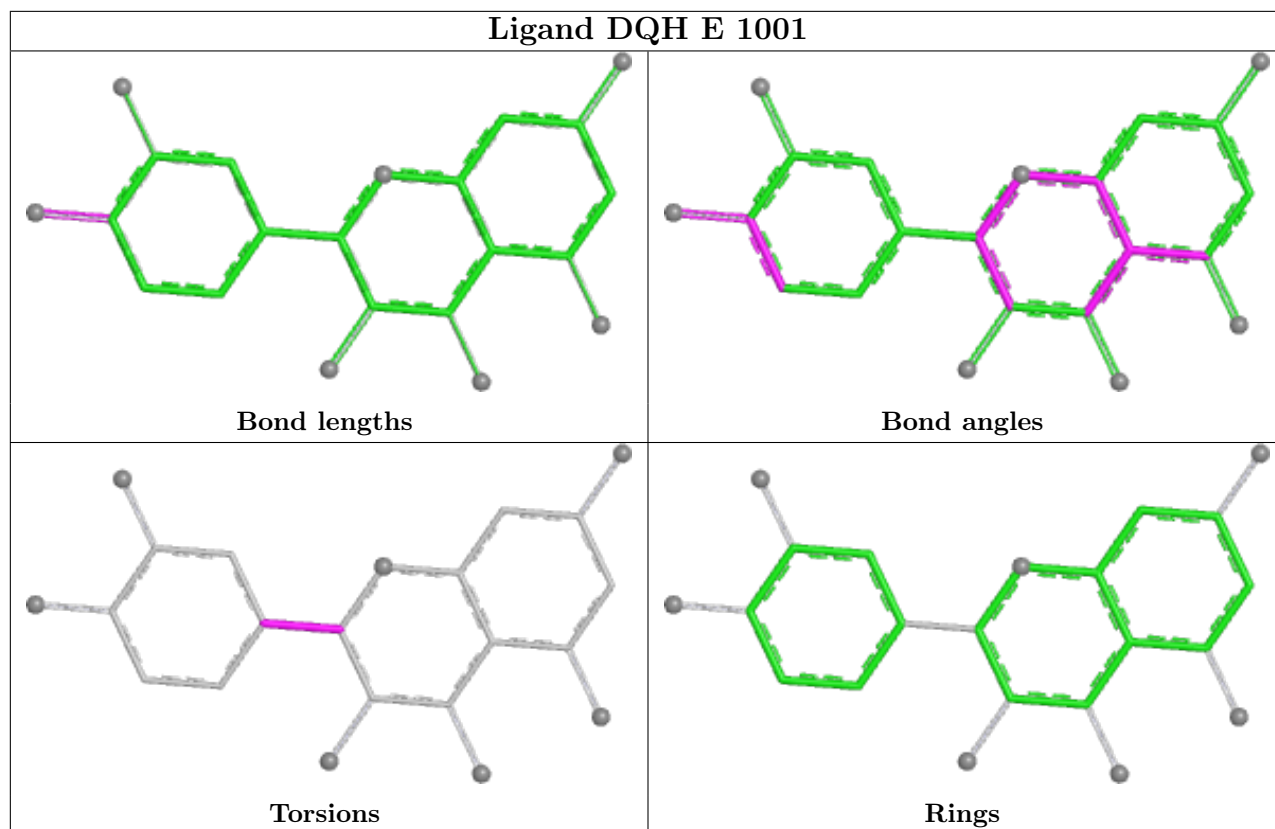


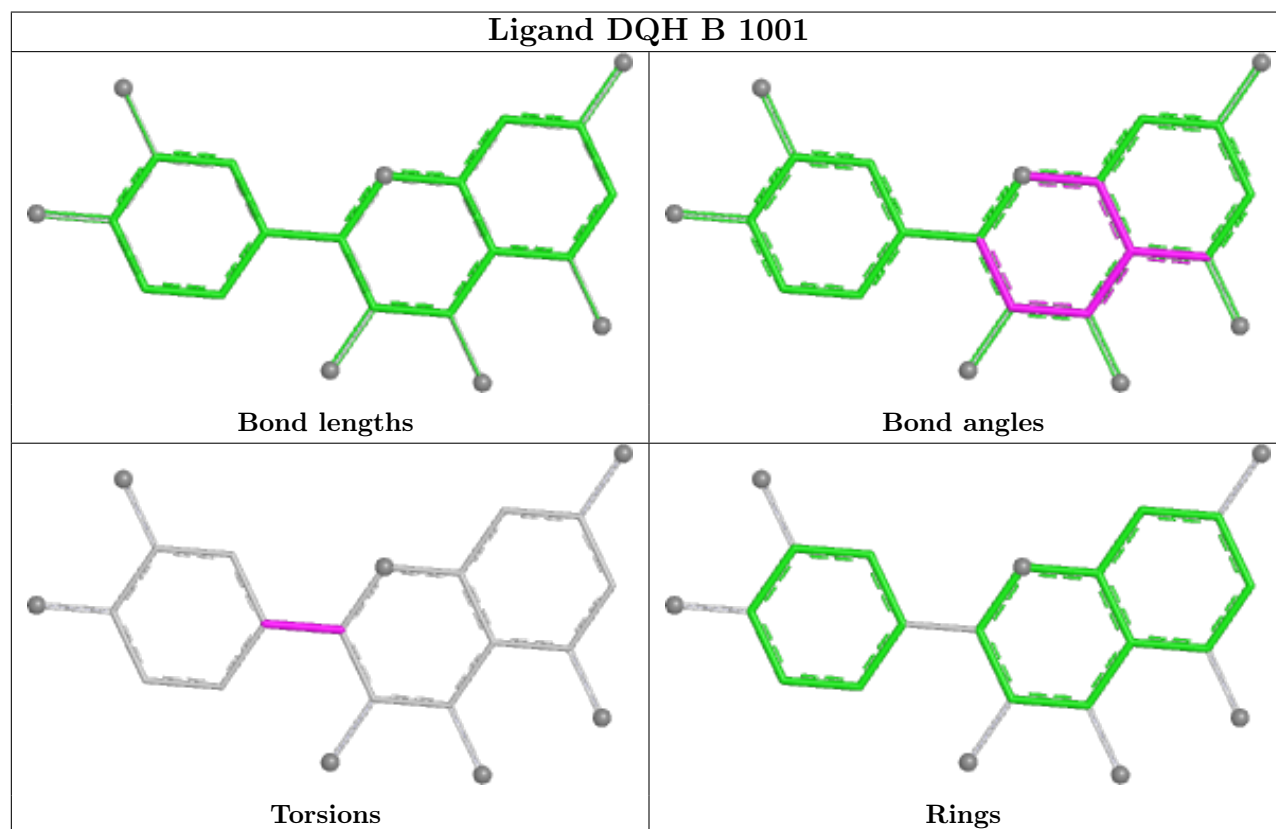
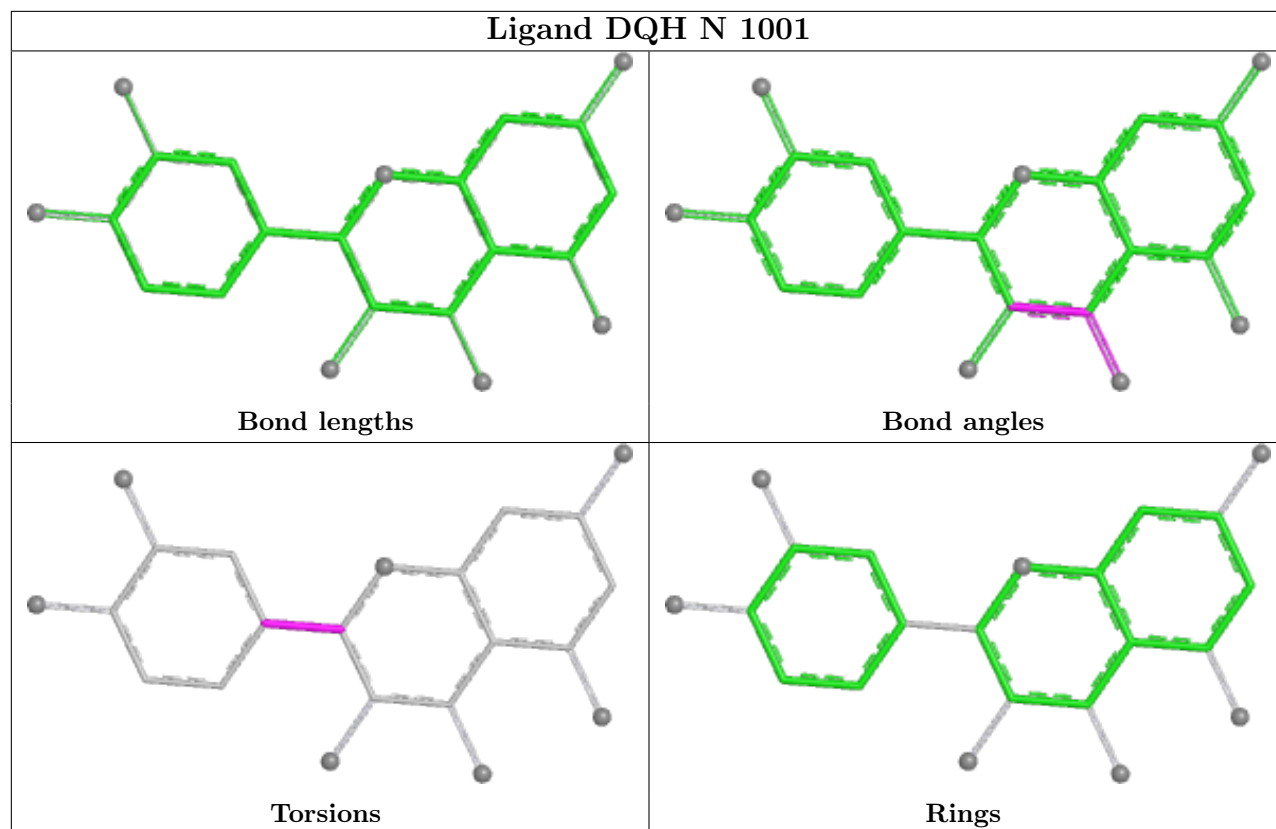


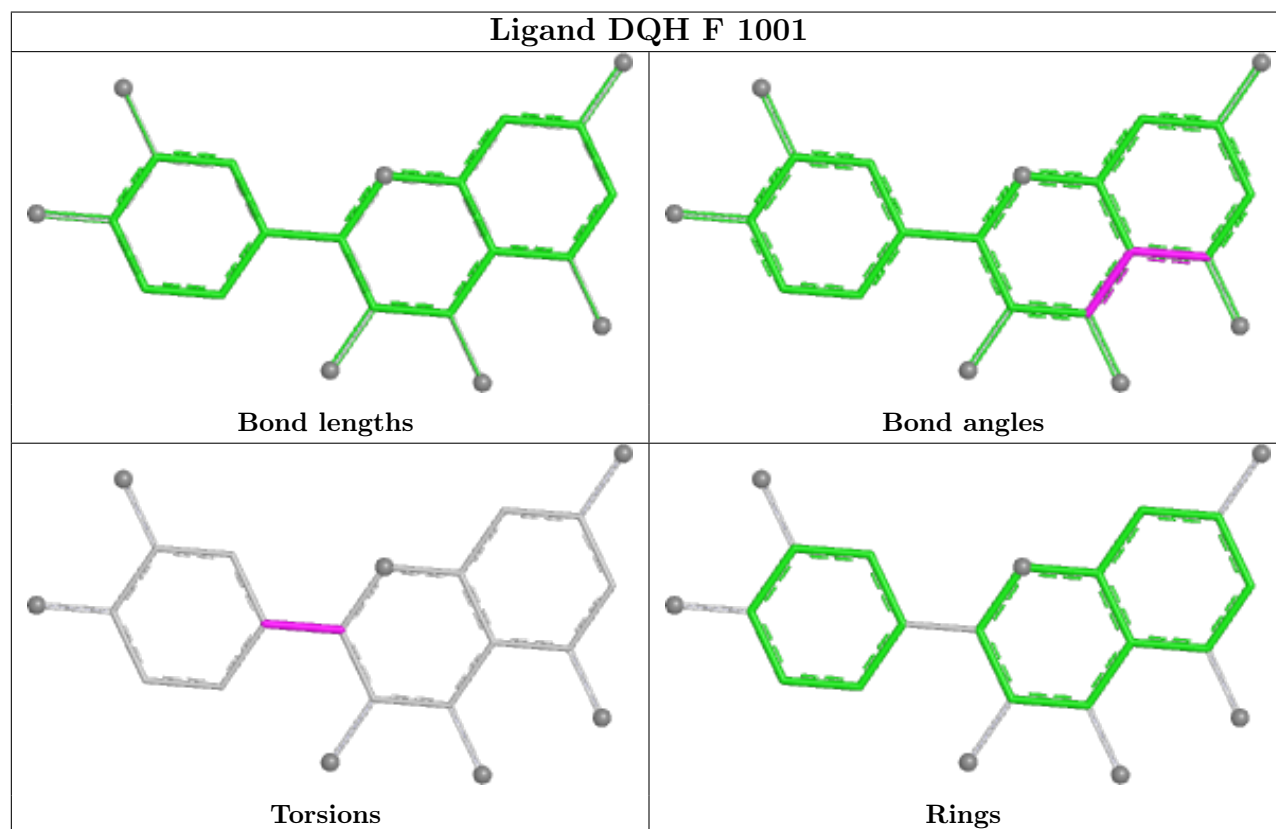
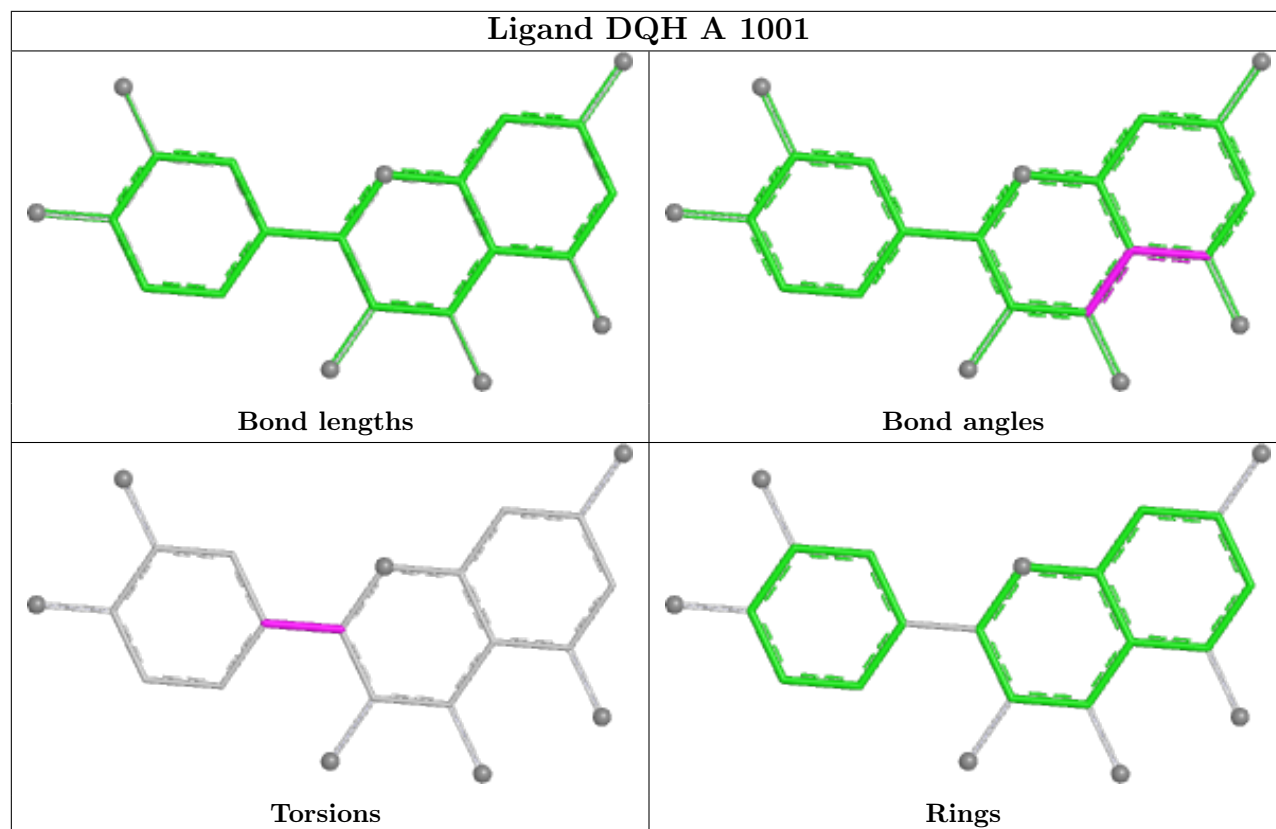


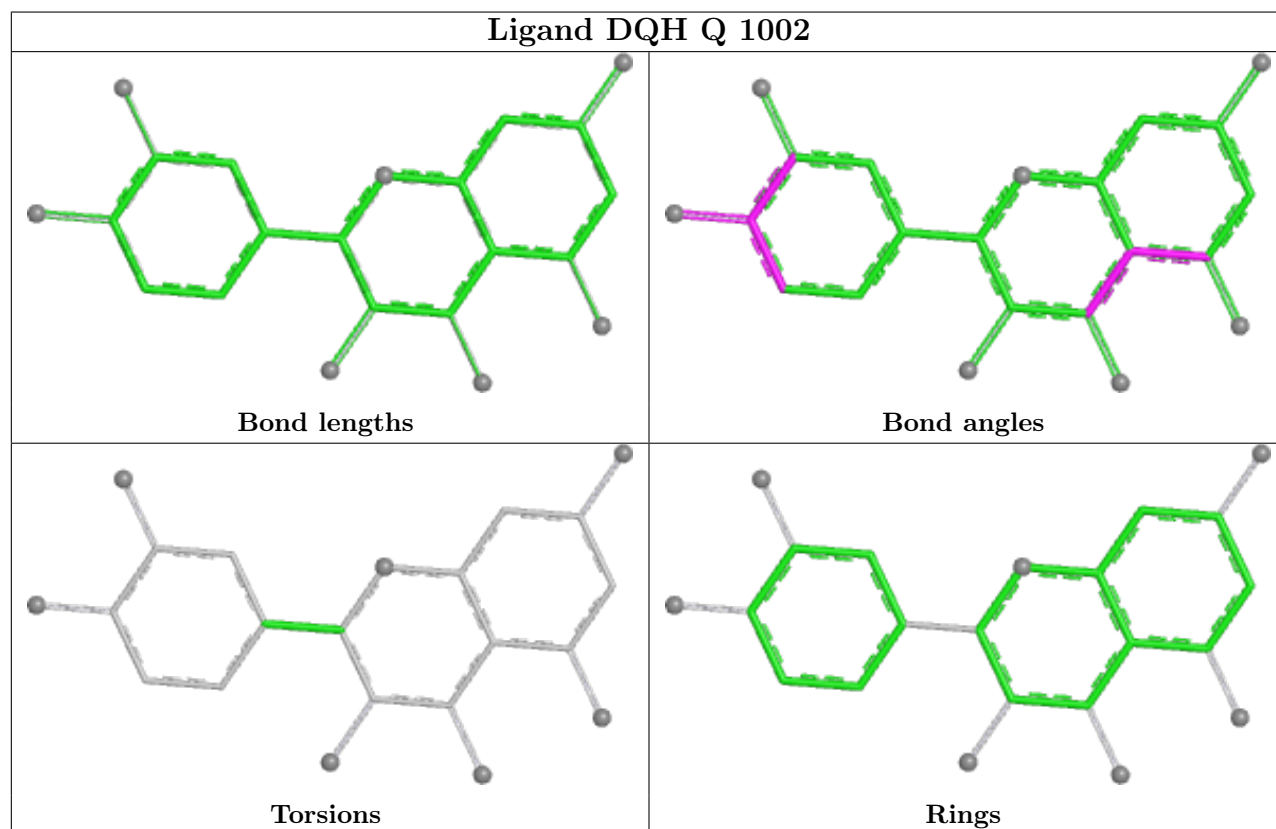
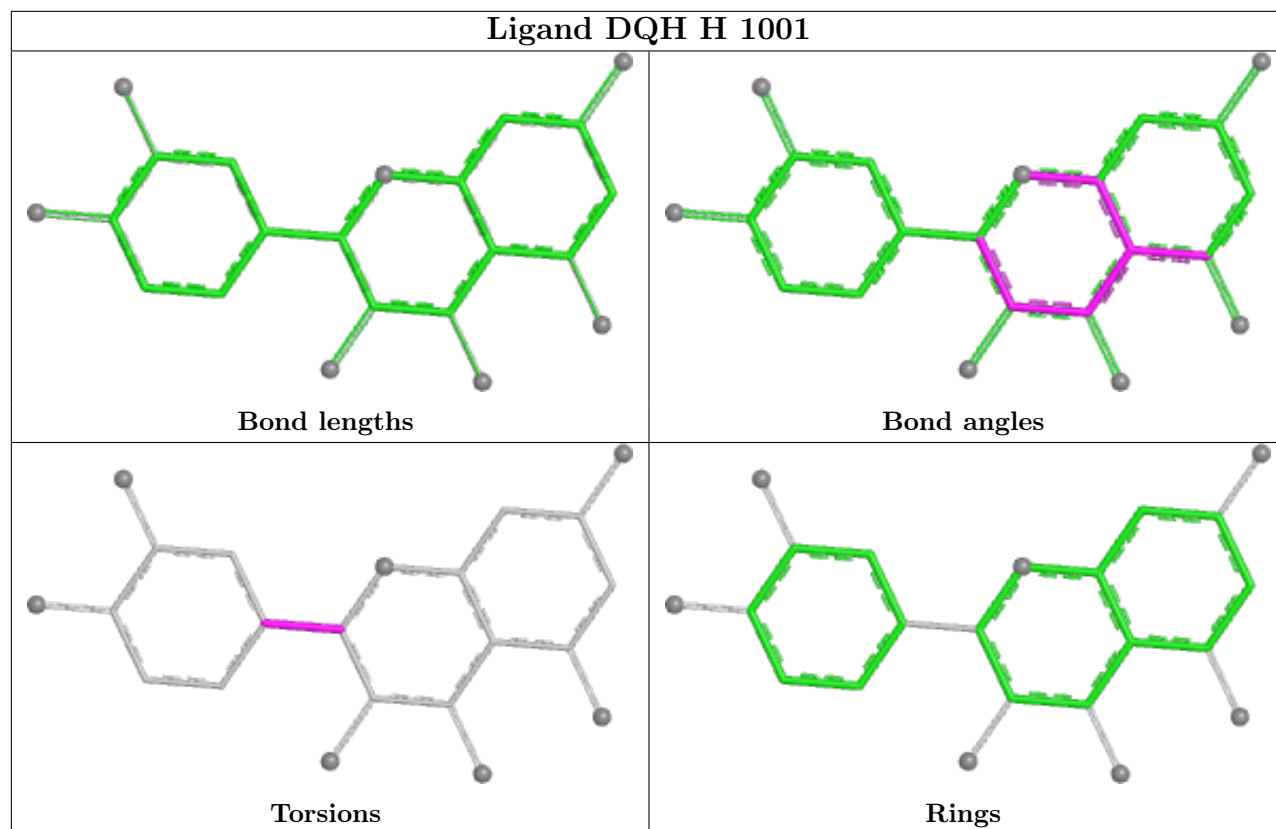












## 5.7 Other polymers [i](#)

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.