



# wwPDB X-ray Structure Validation Summary Report

May 13, 2020 – 03:43 am BST

PDB ID : 1H5S  
Title : Thymidyltransferase complexed with TMP  
Authors : Rosano, C.; Zuccotti, S.; Bolognesi, M.  
Deposited on : 2001-05-25  
Resolution : 2.30 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)  
Xtrriage (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) : 2.11

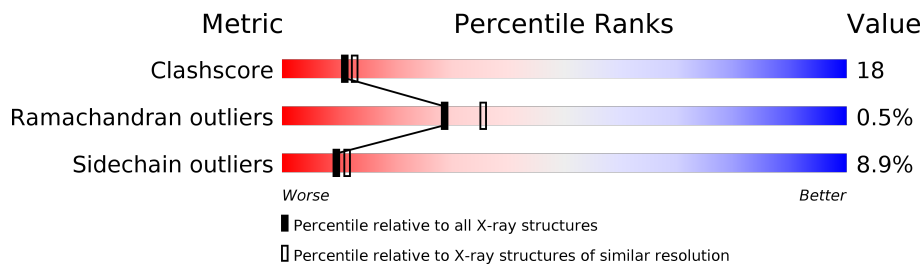
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)

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 was not executed.

Mol	Chain	Length	Quality of chain
1	A	293	
2	B	293	
3	C	293	
4	D	293	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 9667 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 Glucose-1-phosphate thymidyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	290	2276	1461	377	428	10	26	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	185	LEU	GLN	conflict	UNP P37744
A	217	LEU	MET	conflict	UNP P37744

- Molecule 2 is a protein called Glucose-1-phosphate thymidyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	291	2280	1461	380	429	10	13	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	217	LEU	MET	conflict	UNP P37744
B	288	GLN	MET	conflict	UNP P37744

- Molecule 3 is a protein called Glucose-1-phosphate thymidyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	290	2277	1459	378	429	11	13	0	0

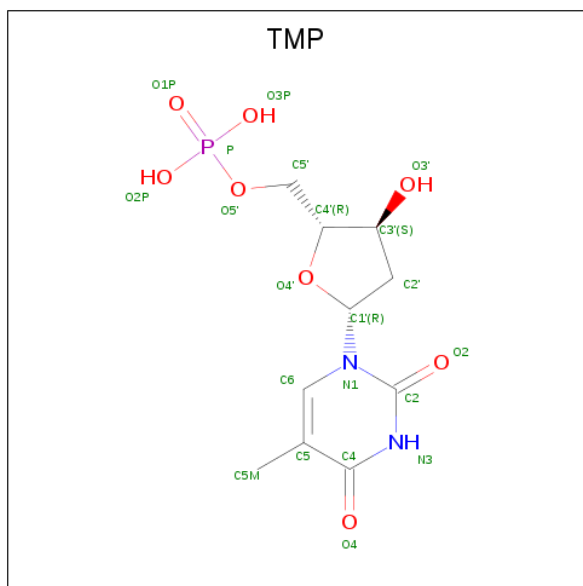
- Molecule 4 is a protein called Glucose-1-phosphate thymidyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	290	2274	1458	377	429	10	13	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	102	ALA	GLY	conflict	UNP P37744
D	217	LEU	MET	conflict	UNP P37744

- Molecule 5 is THYMIDINE-5'-PHOSPHATE (three-letter code: TMP) (formula:  $C_{10}H_{15}N_2O_8P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
5	A	1	Total	C	N	O	P	0	0
			21	10	2	8	1		
5	A	1	Total	C	N	O	P	0	0
			21	10	2	8	1		
5	B	1	Total	C	N	O	P	0	0
			21	10	2	8	1		
5	B	1	Total	C	N	O	P	0	0
			21	10	2	8	1		
5	C	1	Total	C	N	O	P	0	0
			21	10	2	8	1		
5	C	1	Total	C	N	O	P	0	0
			21	10	2	8	1		
5	D	1	Total	C	N	O	P	0	0
			21	10	2	8	1		
5	D	1	Total	C	N	O	P	0	0
			21	10	2	8	1		

- Molecule 6 is water.

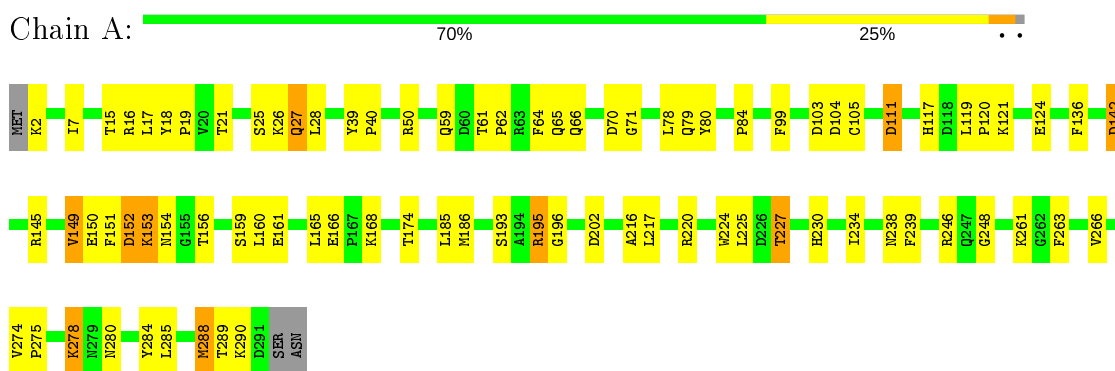
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	A	111	Total 111	O 111	0	0
6	B	71	Total 71	O 71	0	0
6	C	109	Total 109	O 109	0	0
6	D	101	Total 101	O 101	0	0

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

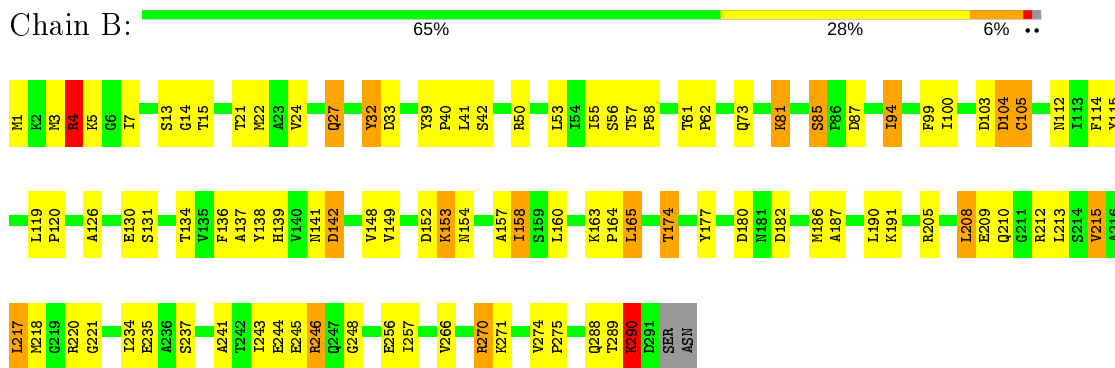
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according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

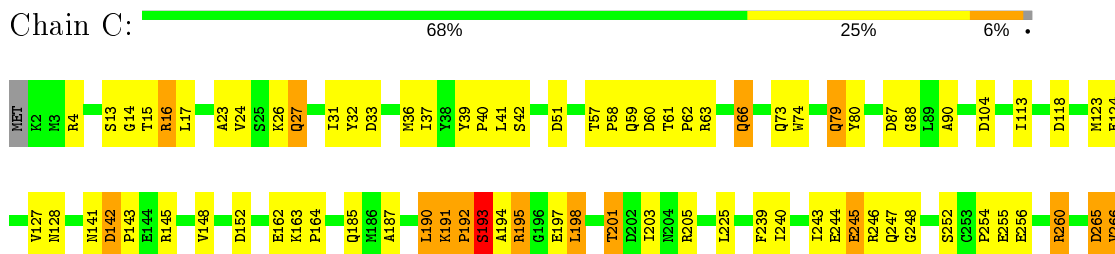
- Molecule 1: Glucose-1-phosphate thymidyltransferase 1



- Molecule 2: Glucose-1-phosphate thymidyltransferase 1



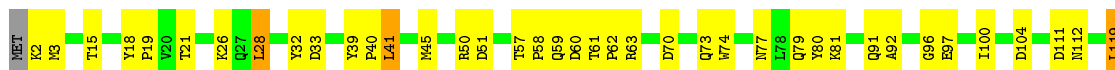
- Molecule 3: Glucose-1-phosphate thymidyltransferase 1





● Molecule 4: Glucose-1-phosphate thymidyltransferase 1

Chain D: 69% 25% 5%



## 4 Data and refinement statistics

Xtrriage (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, $\alpha$ , $\beta$ , $\gamma$	73.10Å 119.27Å 81.11Å 90.00° 112.66° 90.00°	Depositor
Resolution (Å)	12.00 – 2.30	Depositor
% Data completeness (in resolution range)	99.0 (12.00-2.30)	Depositor
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.176 , 0.257	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	9667	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	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: TMP

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	0.97	4/2323 (0.2%)	1.48	16/3145 (0.5%)
2	B	0.92	1/2327 (0.0%)	0.99	7/3150 (0.2%)
3	C	1.04	1/2324 (0.0%)	1.20	12/3145 (0.4%)
4	D	1.03	2/2321 (0.1%)	1.20	12/3144 (0.4%)
All	All	0.99	8/9295 (0.1%)	1.23	47/12584 (0.4%)

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	1	1
2	B	1	1
All	All	2	2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	290	LYS	C-N	28.13	1.98	1.34
4	D	290	LYS	C-N	24.05	1.89	1.34
2	B	290	LYS	C-N	-19.94	0.88	1.34
1	A	290	LYS	C-N	15.96	1.70	1.34
4	D	290	LYS	CA-CB	-15.49	1.19	1.53

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	290	LYS	O-C-N	-56.21	32.76	122.70
4	D	290	LYS	O-C-N	-38.56	61.01	122.70

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	290	LYS	O-C-N	-38.11	61.72	122.70
1	A	124	GLU	CA-CB-CG	10.69	136.91	113.40
2	B	290	LYS	CA-CB-CG	10.58	136.68	113.40

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	290	LYS	CA
2	B	290	LYS	CA

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	196	GLY	Peptide
2	B	290	LYS	Mainchain

## 5.2 Too-close contacts

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	2276	0	2288	67	0
2	B	2280	0	2283	98	0
3	C	2277	0	2283	100	0
4	D	2274	0	2276	90	0
5	A	42	0	26	0	0
5	B	42	0	26	1	0
5	C	42	0	26	6	0
5	D	42	0	26	1	0
6	A	111	0	0	5	1
6	B	71	0	0	2	0
6	C	109	0	0	9	1
6	D	101	0	0	10	0
All	All	9667	0	9234	330	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 18.

The worst 5 of 330 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:153:LYS:N	1:A:153:LYS:HD3	1.46	1.24
4:D:266:VAL:HG13	4:D:289:THR:HG23	1.25	1.17
3:C:16:ARG:N	3:C:16:ARG:HD2	1.62	1.14
3:C:16:ARG:H	3:C:16:ARG:CD	1.58	1.09
3:C:195:ARG:O	3:C:195:ARG:HG2	1.31	1.08

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)
6:A:2048:HOH:O	6:C:2069:HOH:O[1_554]	2.12	0.08

## 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	288/293 (98%)	277 (96%)	11 (4%)	0	100	100
2	B	289/293 (99%)	276 (96%)	11 (4%)	2 (1%)	22	26
3	C	288/293 (98%)	276 (96%)	9 (3%)	3 (1%)	15	17
4	D	288/293 (98%)	280 (97%)	7 (2%)	1 (0%)	41	50
All	All	1153/1172 (98%)	1109 (96%)	38 (3%)	6 (0%)	29	35

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	290	LYS
3	C	193	SER
4	D	144	GLU
3	C	192	PRO
2	B	32	TYR

### 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	244/247 (99%)	225 (92%)	19 (8%)	12	16
2	B	243/247 (98%)	219 (90%)	24 (10%)	8	9
3	C	244/247 (99%)	220 (90%)	24 (10%)	8	9
4	D	243/247 (98%)	223 (92%)	20 (8%)	11	14
All	All	974/988 (99%)	887 (91%)	87 (9%)	9	11

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	266	VAL
3	C	66	GLN
4	D	165	LEU
2	B	270	ARG
3	C	16	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	154	ASN
2	B	230	HIS
4	D	27	GLN
2	B	139	HIS
4	D	170	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

8 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	TMP	D	1292	-	19,22,22	1.47	3 (15%)	23,33,33	1.72	4 (17%)
5	TMP	C	1292	-	19,22,22	1.46	2 (10%)	23,33,33	3.38	12 (52%)
5	TMP	B	1292	-	19,22,22	1.45	3 (15%)	23,33,33	2.03	6 (26%)
5	TMP	A	1292	-	19,22,22	1.28	2 (10%)	23,33,33	1.95	3 (13%)
5	TMP	D	1293	-	19,22,22	1.35	2 (10%)	23,33,33	3.09	8 (34%)
5	TMP	B	1293	-	19,22,22	1.20	1 (5%)	23,33,33	3.29	10 (43%)
5	TMP	C	1293	-	19,22,22	1.36	3 (15%)	23,33,33	2.05	5 (21%)
5	TMP	A	1293	-	19,22,22	1.08	1 (5%)	23,33,33	2.05	7 (30%)

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
5	TMP	D	1292	-	-	3/7/22/22	0/2/2/2
5	TMP	C	1292	-	-	3/7/22/22	0/2/2/2
5	TMP	B	1292	-	-	4/7/22/22	0/2/2/2
5	TMP	A	1292	-	-	3/7/22/22	0/2/2/2
5	TMP	D	1293	-	-	0/7/22/22	0/2/2/2
5	TMP	B	1293	-	-	3/7/22/22	0/2/2/2
5	TMP	C	1293	-	-	3/7/22/22	0/2/2/2
5	TMP	A	1293	-	-	0/7/22/22	0/2/2/2

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	1293	TMP	O4-C4	4.19	1.35	1.24
5	A	1292	TMP	O4-C4	4.10	1.34	1.24
5	D	1293	TMP	O4-C4	4.02	1.34	1.24
5	B	1293	TMP	O4-C4	3.97	1.34	1.24
5	C	1292	TMP	O4-C4	3.67	1.33	1.24

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	1293	TMP	O5'-P-O1P	-9.20	80.67	106.47
5	D	1293	TMP	O5'-P-O1P	-8.24	83.35	106.47
5	B	1293	TMP	O2P-P-O5'	-7.73	86.17	106.73
5	A	1292	TMP	C4-N3-C2	7.60	121.56	115.14
5	C	1292	TMP	C4-N3-C2	7.17	121.19	115.14

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	D	1292	TMP	C5'-O5'-P-O2P
5	D	1292	TMP	C5'-O5'-P-O3P
5	C	1292	TMP	O4'-C4'-C5'-O5'
5	C	1292	TMP	O4'-C1'-N1-C6
5	B	1292	TMP	C5'-O5'-P-O2P

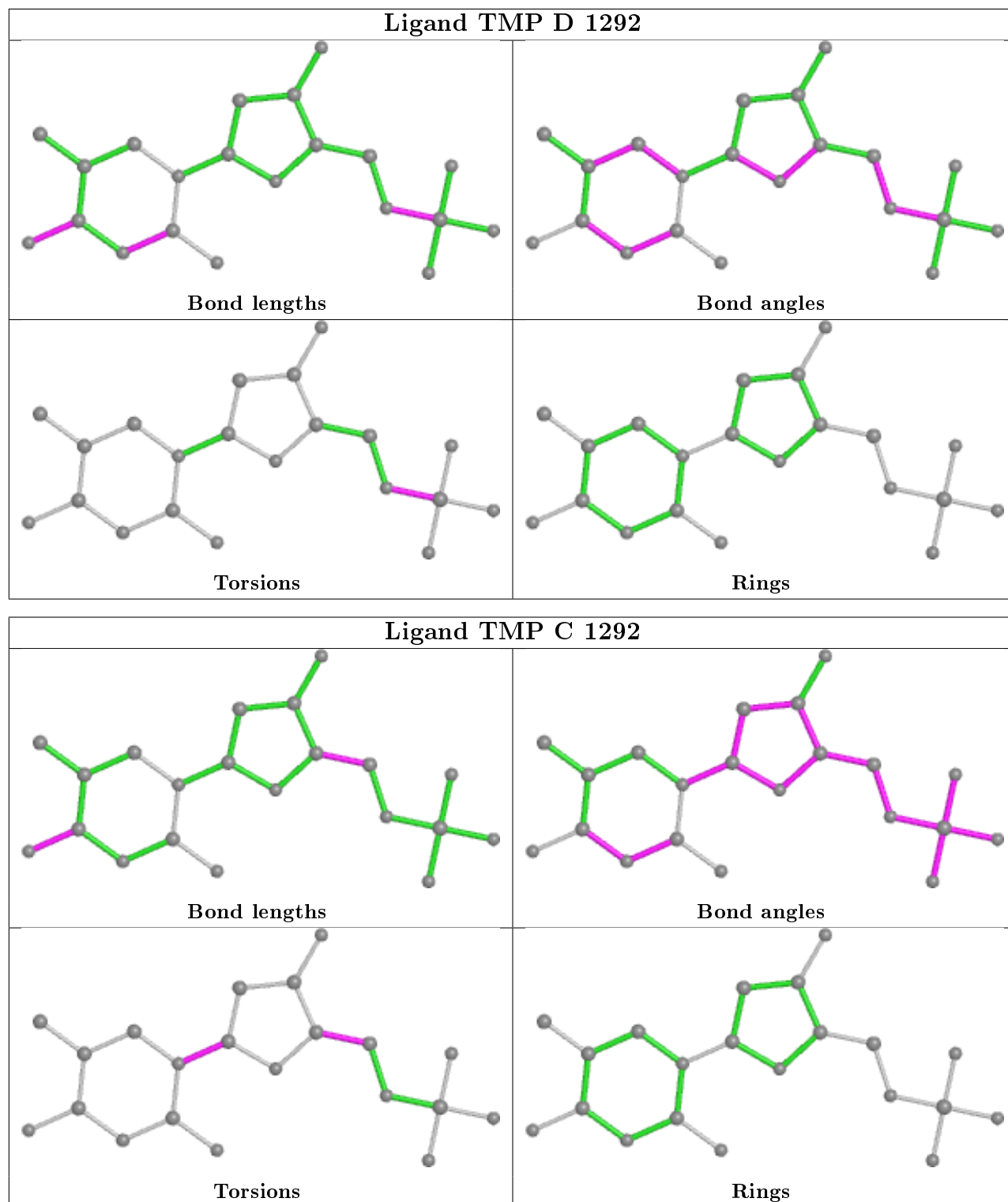
There are no ring outliers.

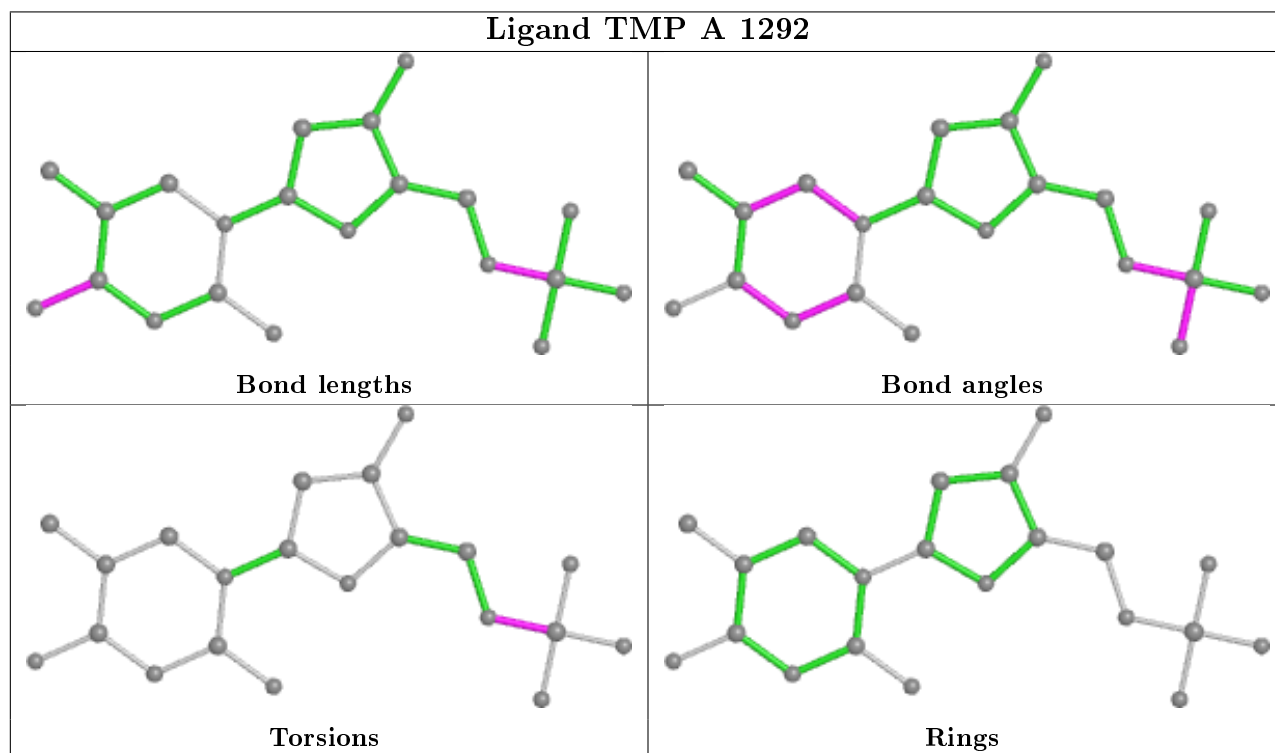
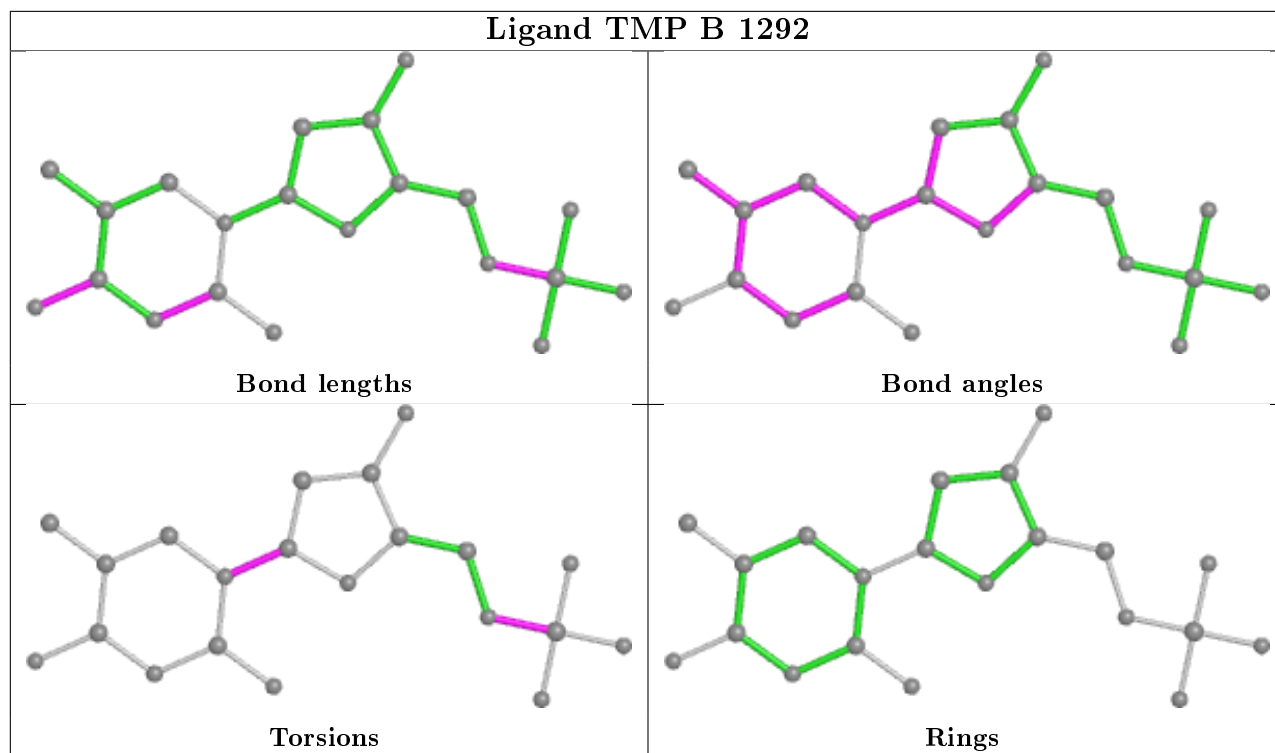
4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	1292	TMP	1	0
5	C	1292	TMP	5	0
5	B	1292	TMP	1	0
5	C	1293	TMP	1	0

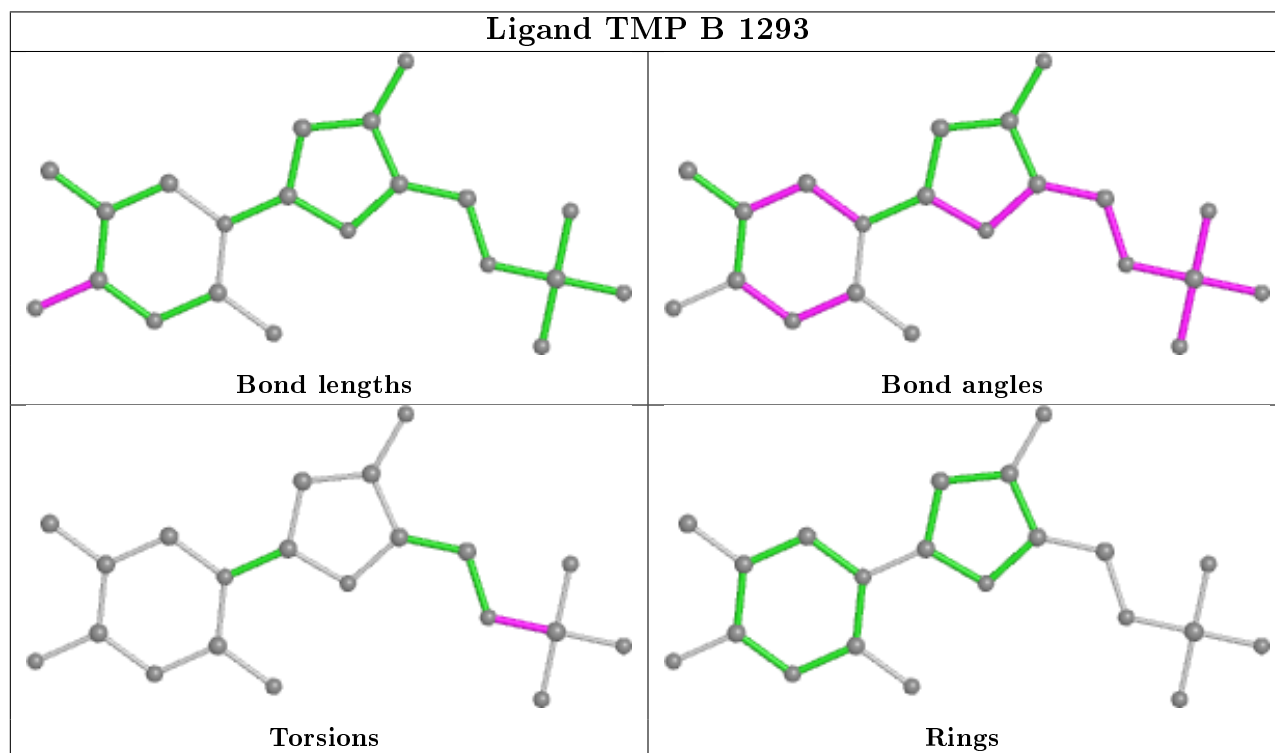
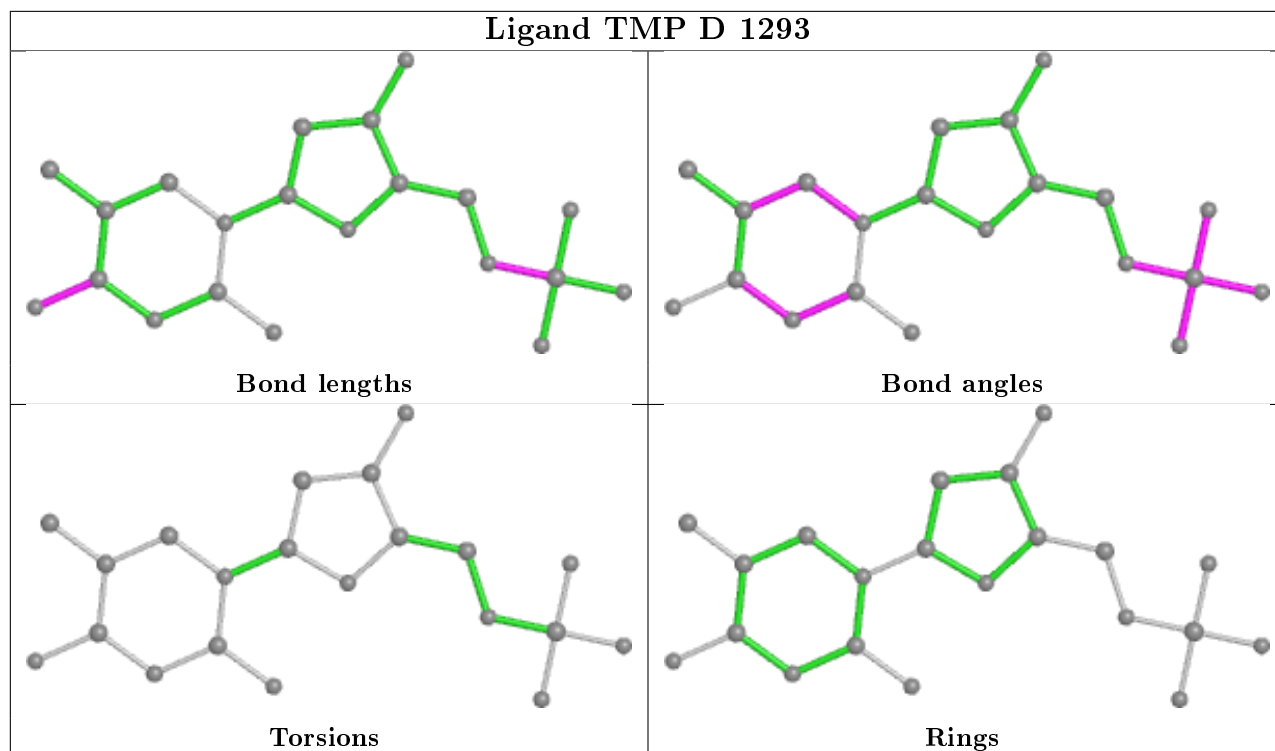
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

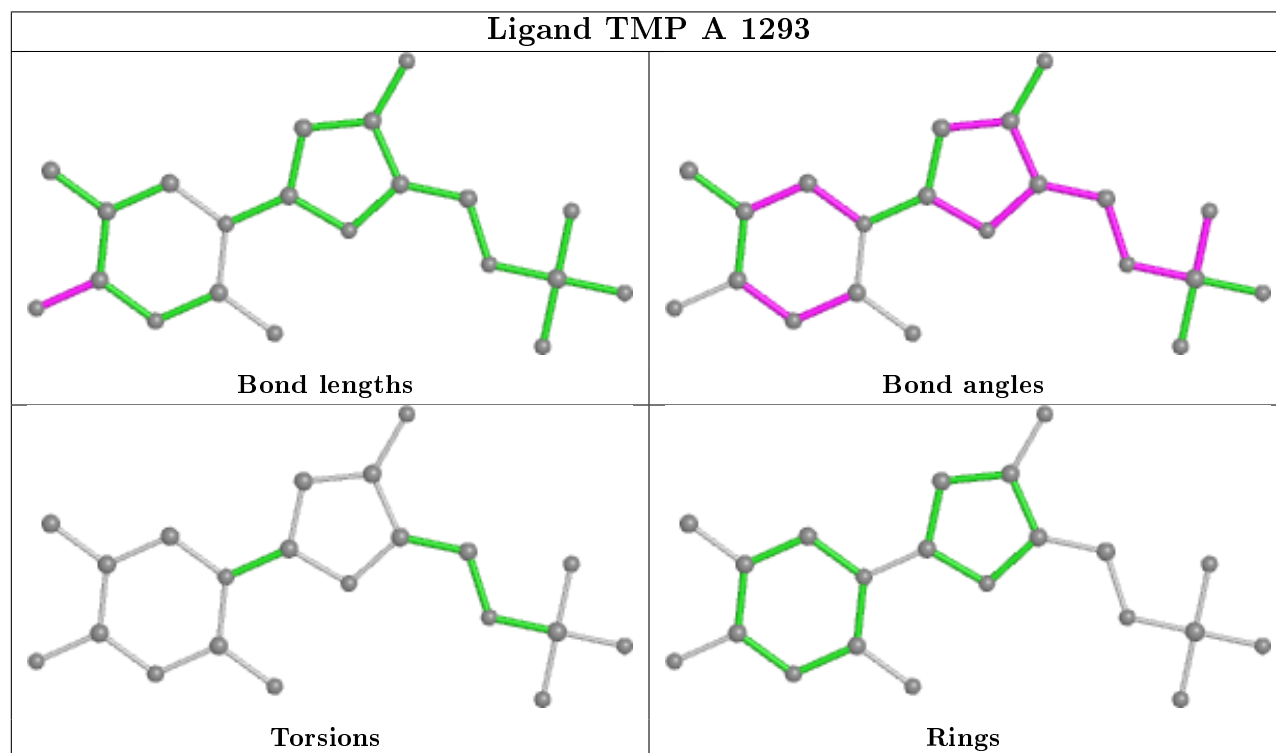
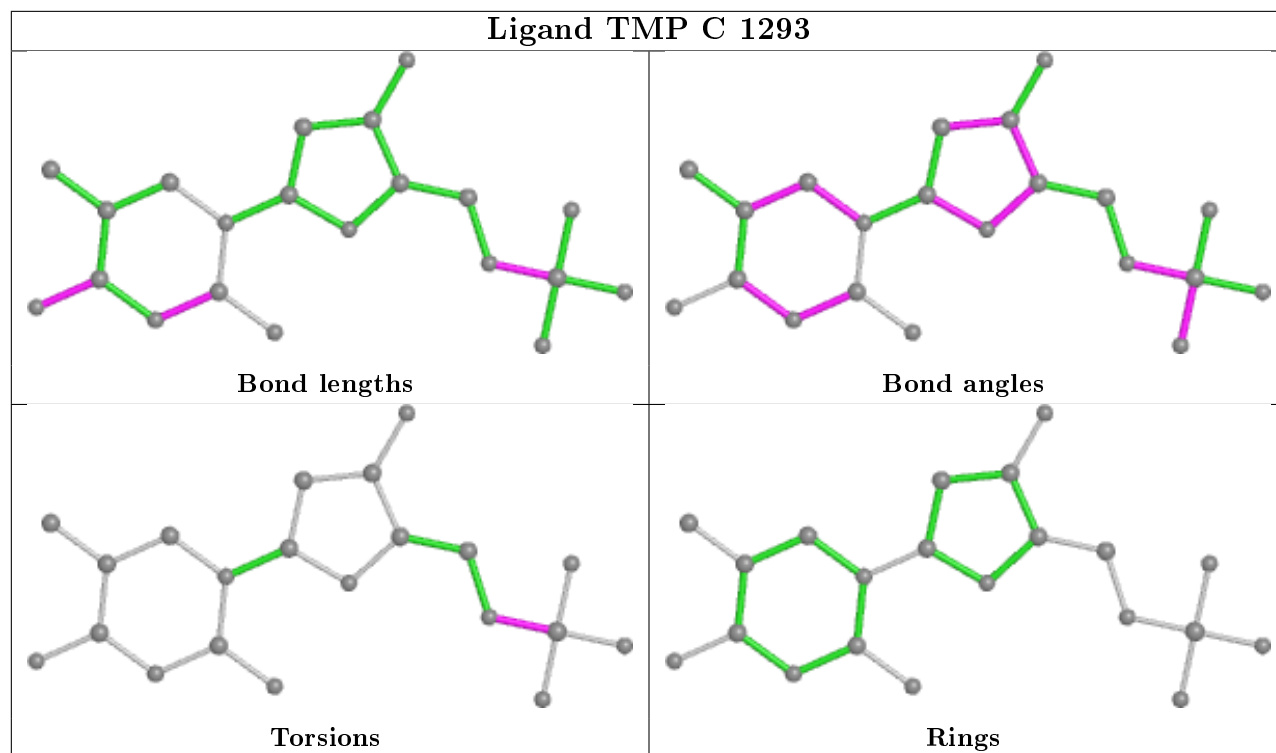
in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers 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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	B	1
1	A	1
4	D	1
3	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	290:LYS	C	291:ASP	N	1.98
1	D	290:LYS	C	291:ASP	N	1.89
1	A	290:LYS	C	291:ASP	N	1.70
1	B	290:LYS	C	291:ASP	N	0.88

## 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.