



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 23, 2024 – 12:08 PM JST

PDB ID : 8XI1  
Title : Cellodextrin phosphorylase from Clostridium thermocellum mutant - all cysteine residues were substituted with serines  
Authors : Kuga, T.; Sunagawa, N.; Igarashi, K.  
Deposited on : 2023-12-19  
Resolution : 1.37 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

<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)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

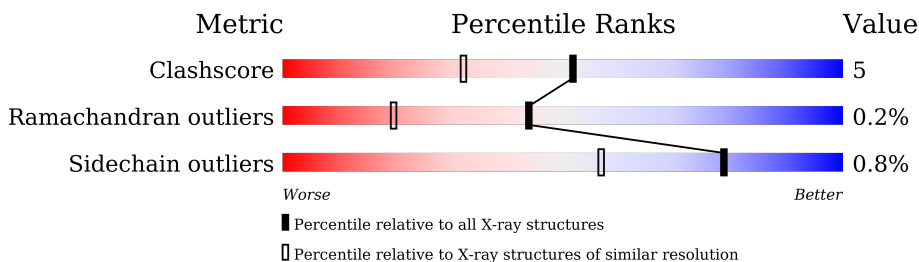
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.37 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	180529	4183 (1.40-1.36)
Ramachandran outliers	177936	4116 (1.40-1.36)
Sidechain outliers	177891	4115 (1.40-1.36)

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 was not executed.

Mol	Chain	Length	Quality of chain
1	A	992	
1	B	992	

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
5	ACT	B	1001	-	-	X	-
5	ACT	B	1004	-	-	X	-

## 2 Entry composition [i](#)

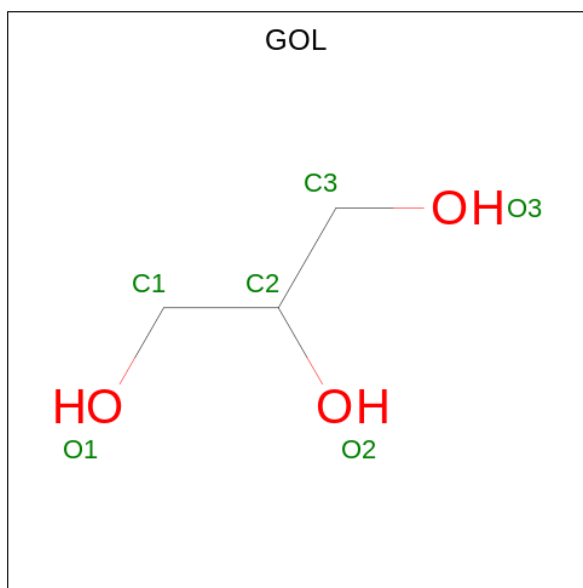
There are 6 unique types of molecules in this entry. The entry contains 33275 atoms, of which 14445 are hydrogens and 0 are deuteriums.

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 cellodextrin phosphorylase variant.

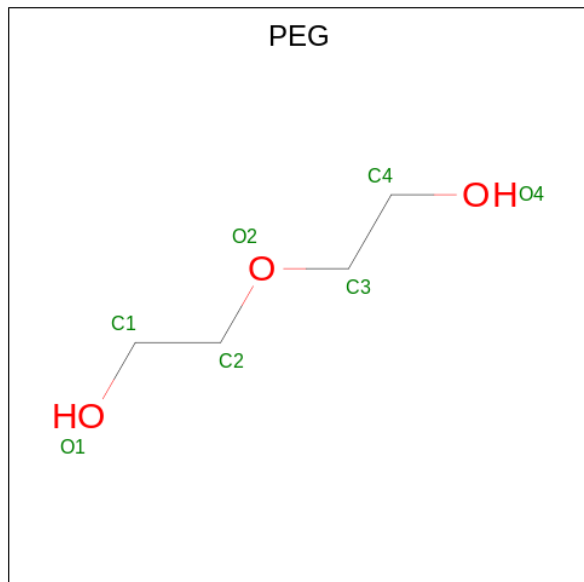
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	984	15293	5185	7175	1353	1556	24	177	31	0
1	B	984	15357	5195	7219	1362	1560	21	171	33	0

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	A	1	14	3	8	3	0	0
2	A	1	14	3	8	3	0	0
2	B	1	14	3	8	3	0	0
2	B	1	14	3	8	3	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
3	A	1	17	4	10	3	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Cl	0	0
			2	2		
4	B	2	Total	Cl	0	0
			2	2		

- Molecule 5 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	H	O	0	0
			7	2	3	2		
5	B	1	Total	C	H	O	0	0
			7	2	3	2		
5	B	1	Total	C	H	O	0	0
			7	2	3	2		

- Molecule 6 is water.

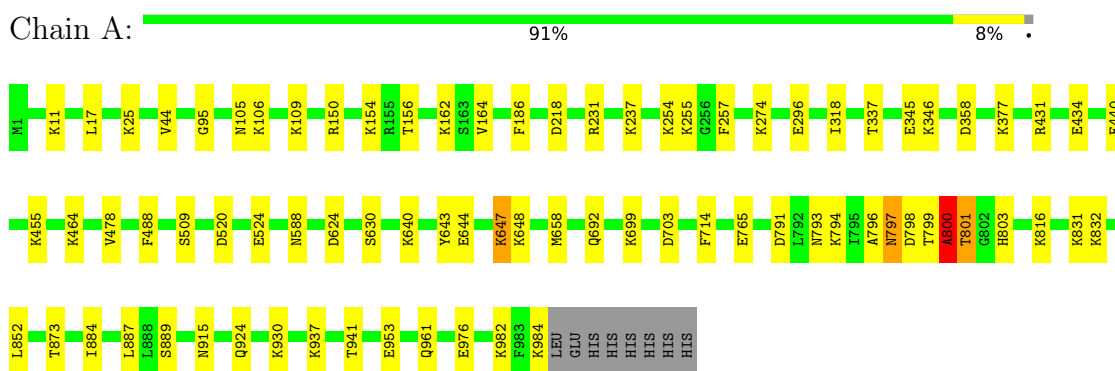
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1274	Total	O	0	0
			1274	1274		
6	B	1253	Total	O	0	0
			1253	1253		

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

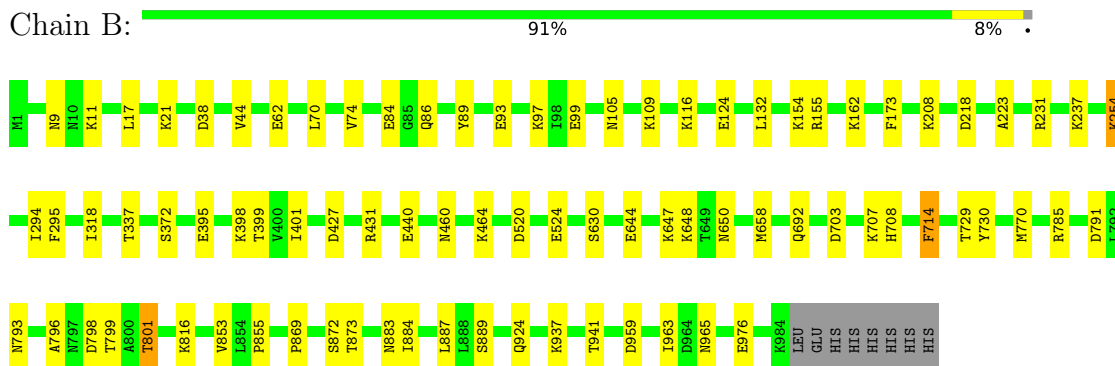
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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 was not executed.

- Molecule 1: cellodextrin phosphorylase variant



- Molecule 1: cellodextrin phosphorylase variant



## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	83.44Å 89.05Å 88.99Å 98.88° 110.55° 110.67°	Depositor
Resolution (Å)	43.76 – 1.37	Depositor
% Data completeness (in resolution range)	93.5 (43.76-1.37)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	PHENIX 1.21rc1_5127	Depositor
R, $R_{free}$	0.174 , 0.200	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	33275	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	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: GOL, ACT, CL, PEG

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	0.43	0/8295	0.65	0/11219
1	B	0.43	0/8314	0.65	0/11241
All	All	0.43	0/16609	0.65	0/22460

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	2
1	B	0	1
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	797	ASN	Peptide
1	A	800	ALA	Peptide
1	B	798	ASP	Peptide

### 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	8118	7175	7952	75	0
1	B	8138	7219	7986	84	0
2	A	12	16	16	2	0
2	B	12	16	16	0	0
3	A	7	10	10	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
5	B	12	9	9	5	0
6	A	1274	0	0	37	1
6	B	1253	0	0	44	1
All	All	18830	14445	15989	161	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 5.

The worst 5 of 161 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:B:62[B]:GLU:OE2	6:B:1101:HOH:O	1.87	0.91
1:B:62[B]:GLU:OE1	6:B:1102:HOH:O	1.89	0.89
1:A:254:LYS:HD2	6:A:2038:HOH:O	1.79	0.81
1:A:643:TYR:O	1:A:647[A]:LYS:HE2	1.81	0.81
1:A:643:TYR:O	1:A:647[B]:LYS:HE2	1.81	0.81

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:1983:HOH:O	6:B:1632:HOH:O[1_454]	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	1013/992 (102%)	979 (97%)	32 (3%)	2 (0%)	44	20
1	B	1016/992 (102%)	983 (97%)	32 (3%)	1 (0%)	48	21
All	All	2029/1984 (102%)	1962 (97%)	64 (3%)	3 (0%)	44	21

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	801	THR
1	B	801	THR
1	A	800	ALA

### 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	881/859 (103%)	873 (99%)	8 (1%)	75	52
1	B	883/859 (103%)	872 (99%)	11 (1%)	67	40
All	All	1764/1718 (103%)	1745 (99%)	19 (1%)	79	43

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	254[B]	LYS
1	B	692	GLN
1	B	714	PHE
1	B	372[B]	SER
1	B	155[A]	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	367	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	458	GLN
1	B	817	HIS

### 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 oligosaccharides in this entry.

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

Of 12 ligands modelled in this entry, 4 are monoatomic - leaving 8 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	GOL	B	1002	-	5,5,5	0.53	0	5,5,5	0.68	0
5	ACT	B	1003	-	3,3,3	1.32	0	3,3,3	1.25	0
5	ACT	B	1004	-	3,3,3	1.16	0	3,3,3	1.34	0
2	GOL	A	1000	-	5,5,5	0.31	0	5,5,5	0.55	0
2	GOL	A	1001	-	5,5,5	0.28	0	5,5,5	0.31	0
5	ACT	B	1001	-	3,3,3	1.21	0	3,3,3	1.11	0
3	PEG	A	1002	-	6,6,6	0.27	0	5,5,5	0.27	0
2	GOL	B	1000	-	5,5,5	0.39	0	5,5,5	0.49	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	GOL	B	1002	-	-	4/4/4/4	-
2	GOL	A	1000	-	-	1/4/4/4	-
2	GOL	A	1001	-	-	0/4/4/4	-
3	PEG	A	1002	-	-	1/4/4/4	-
2	GOL	B	1000	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

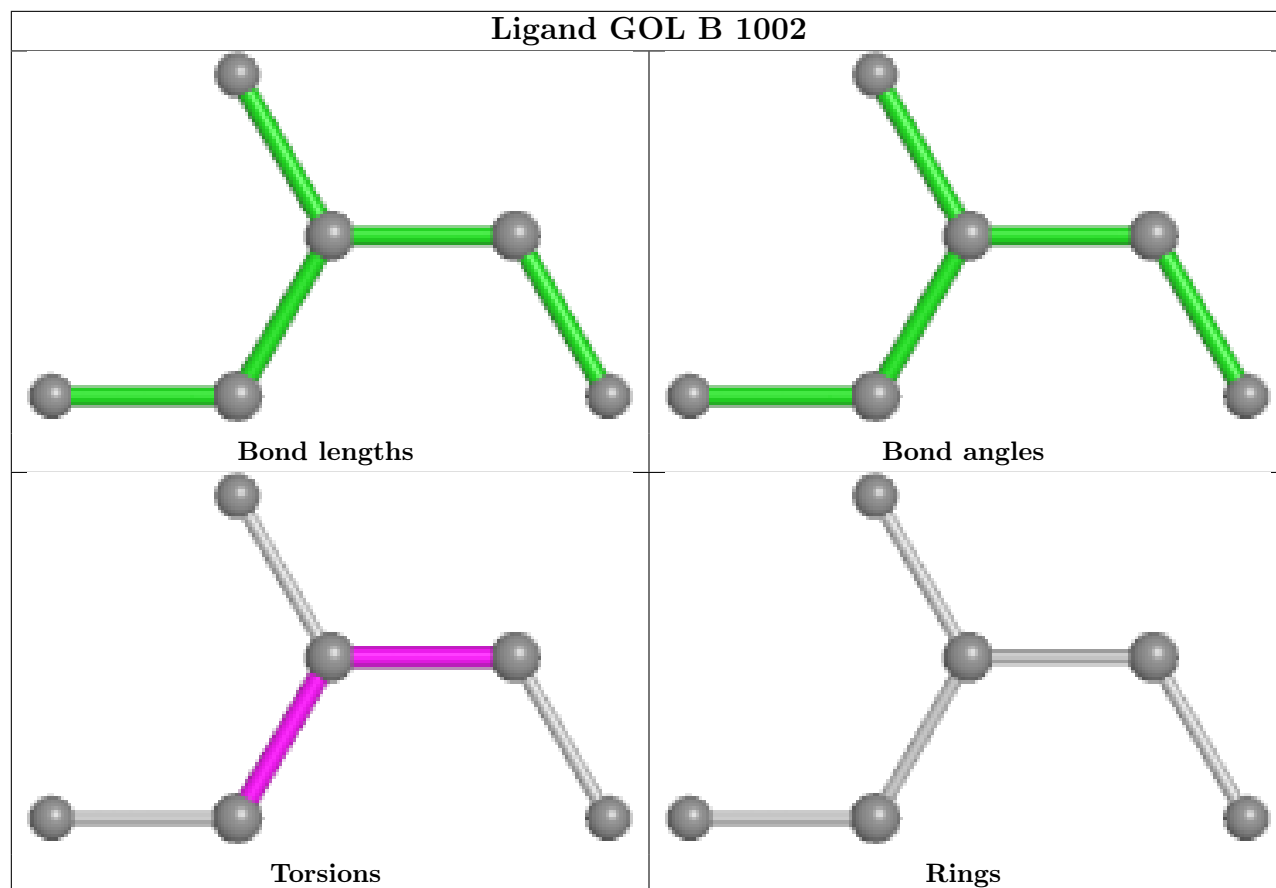
Mol	Chain	Res	Type	Atoms
2	B	1000	GOL	O1-C1-C2-C3
2	B	1002	GOL	O1-C1-C2-C3
2	B	1000	GOL	O1-C1-C2-O2
2	B	1002	GOL	O1-C1-C2-O2
2	B	1002	GOL	O2-C2-C3-O3

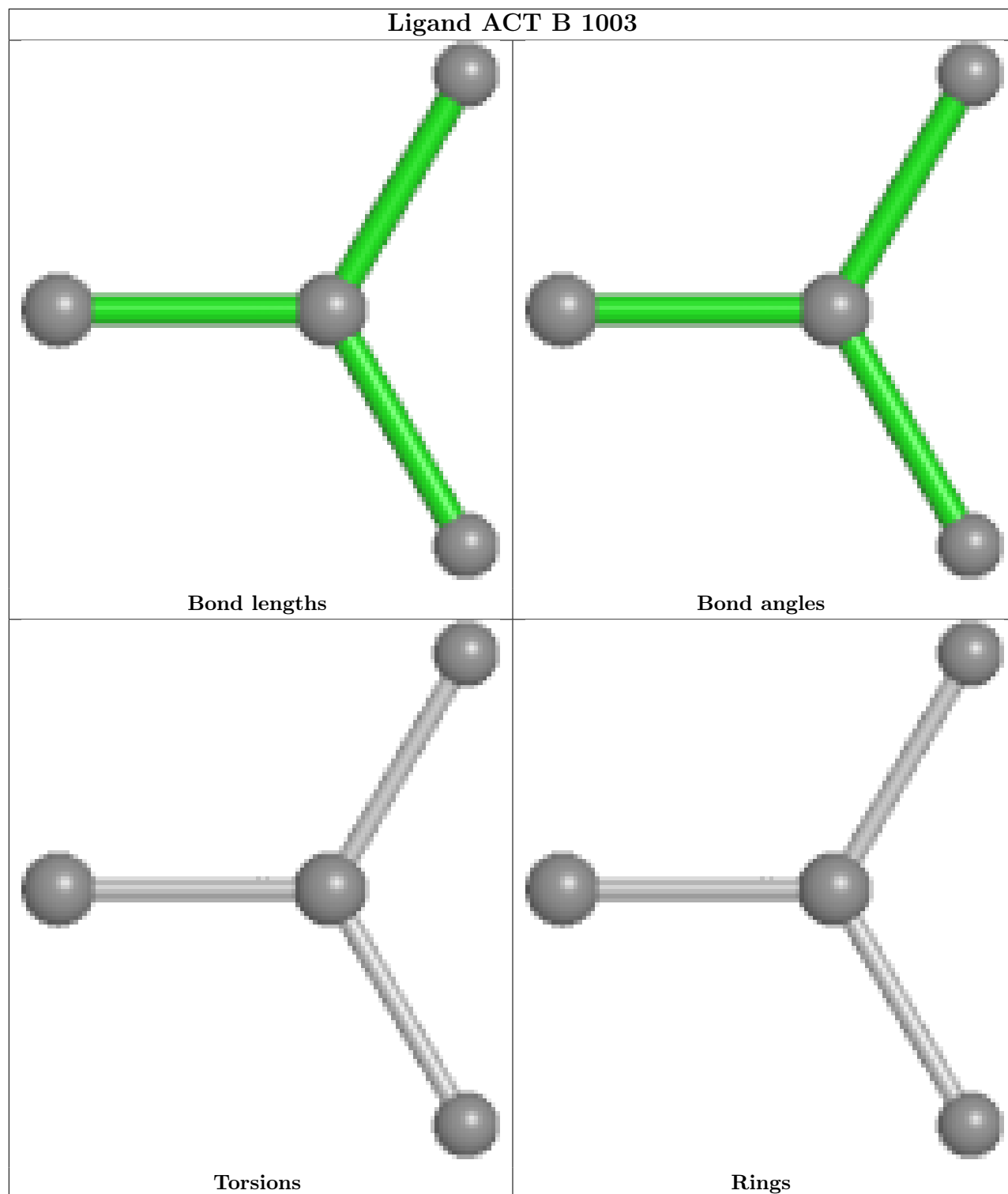
There are no ring outliers.

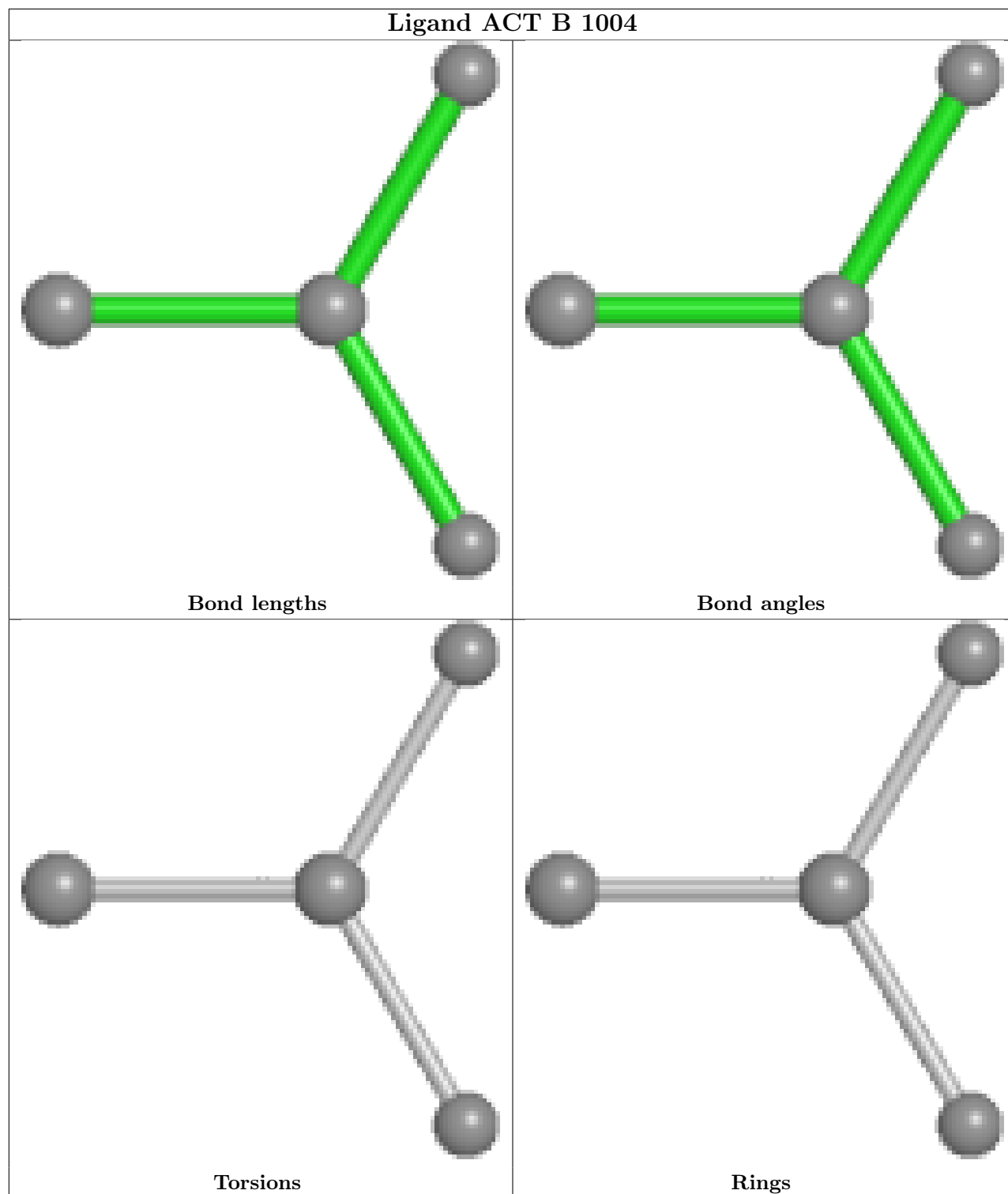
4 monomers are involved in 7 short contacts:

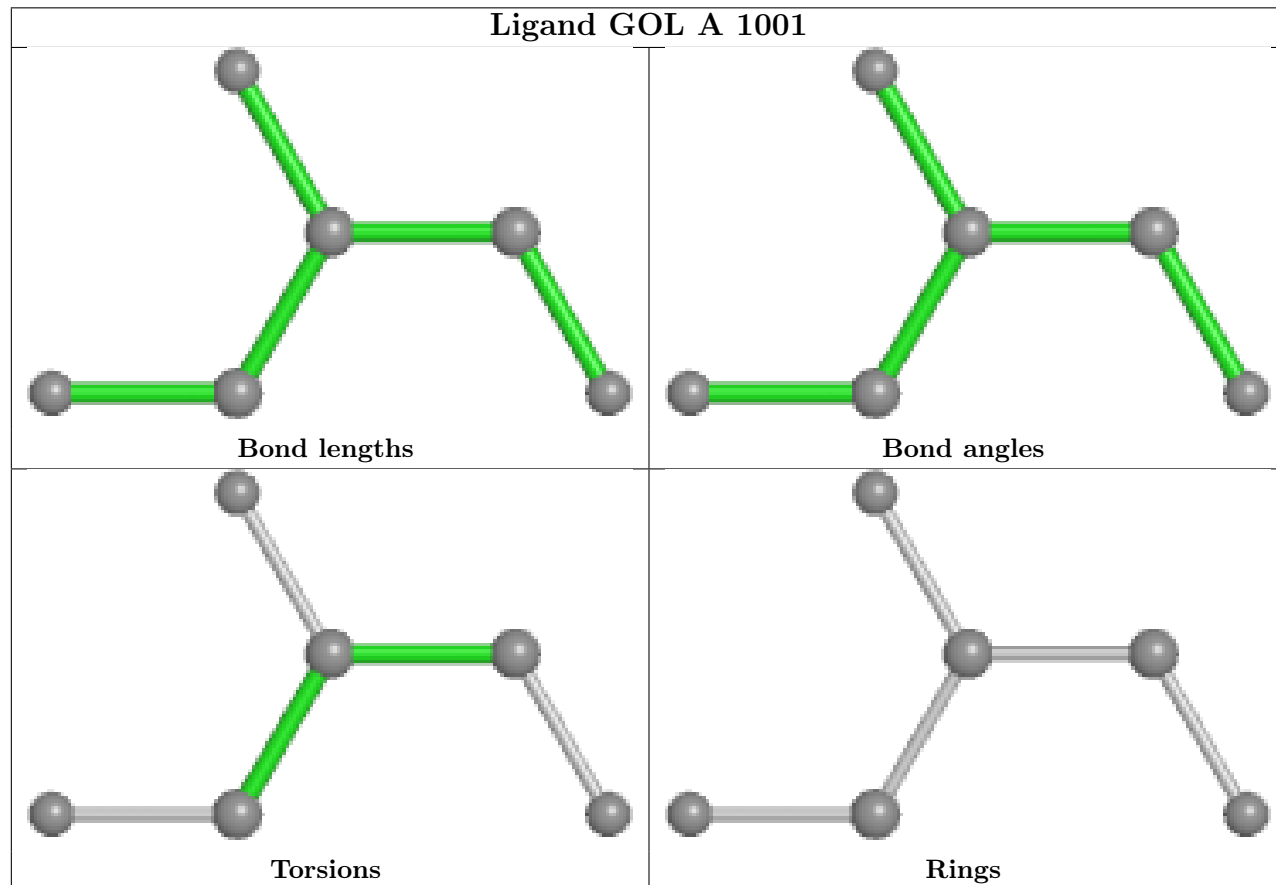
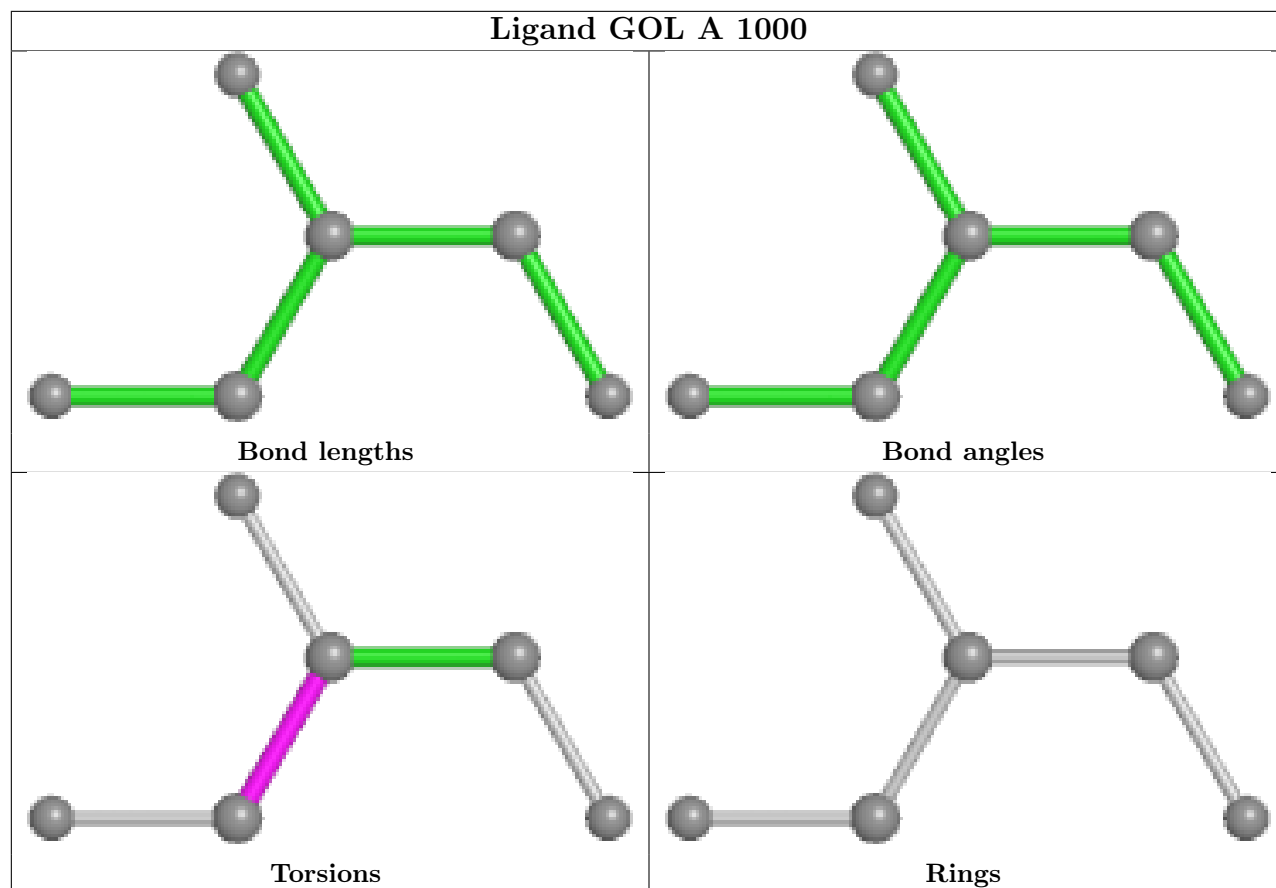
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	1004	ACT	3	0
2	A	1000	GOL	1	0
2	A	1001	GOL	1	0
5	B	1001	ACT	2	0

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.

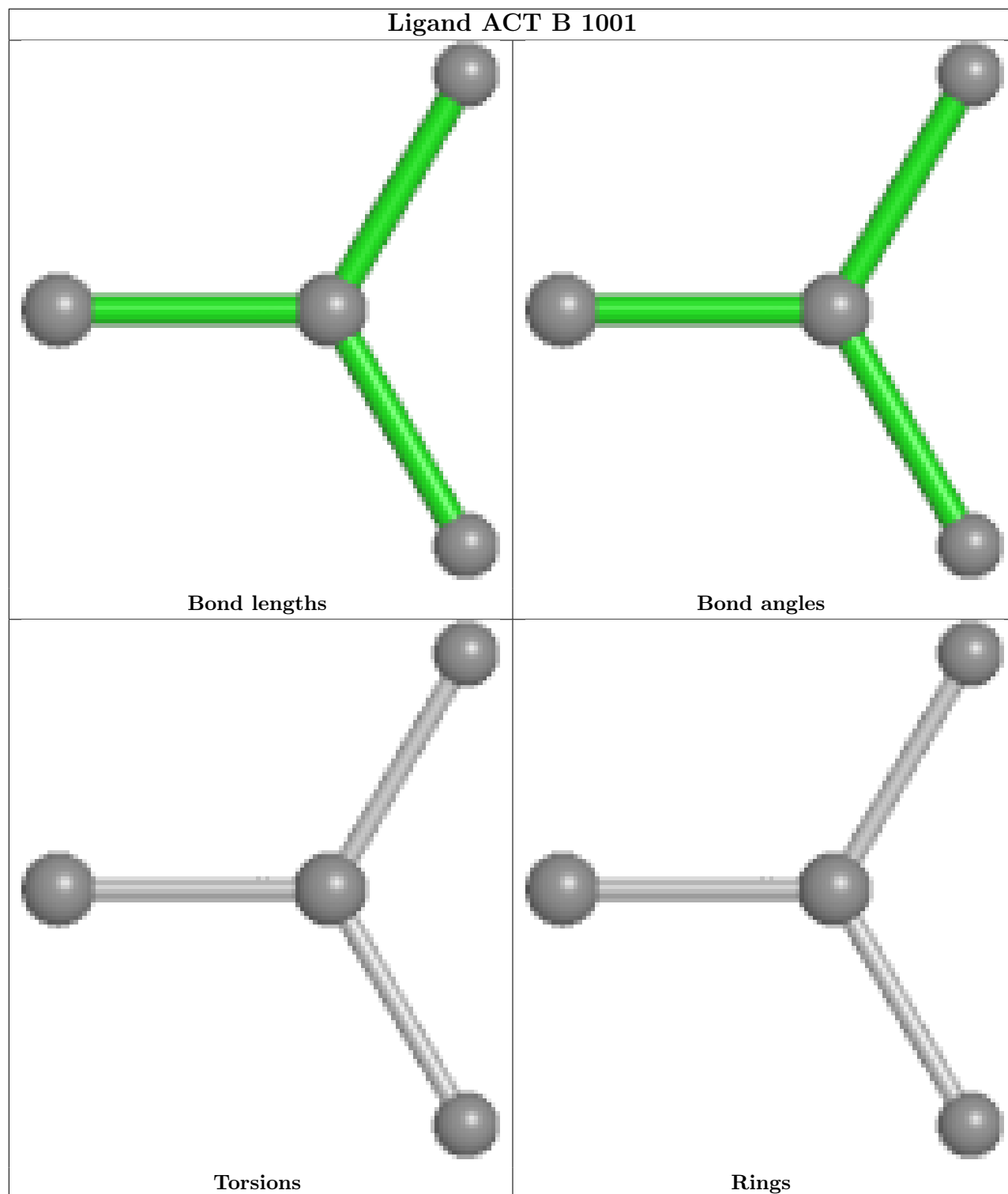


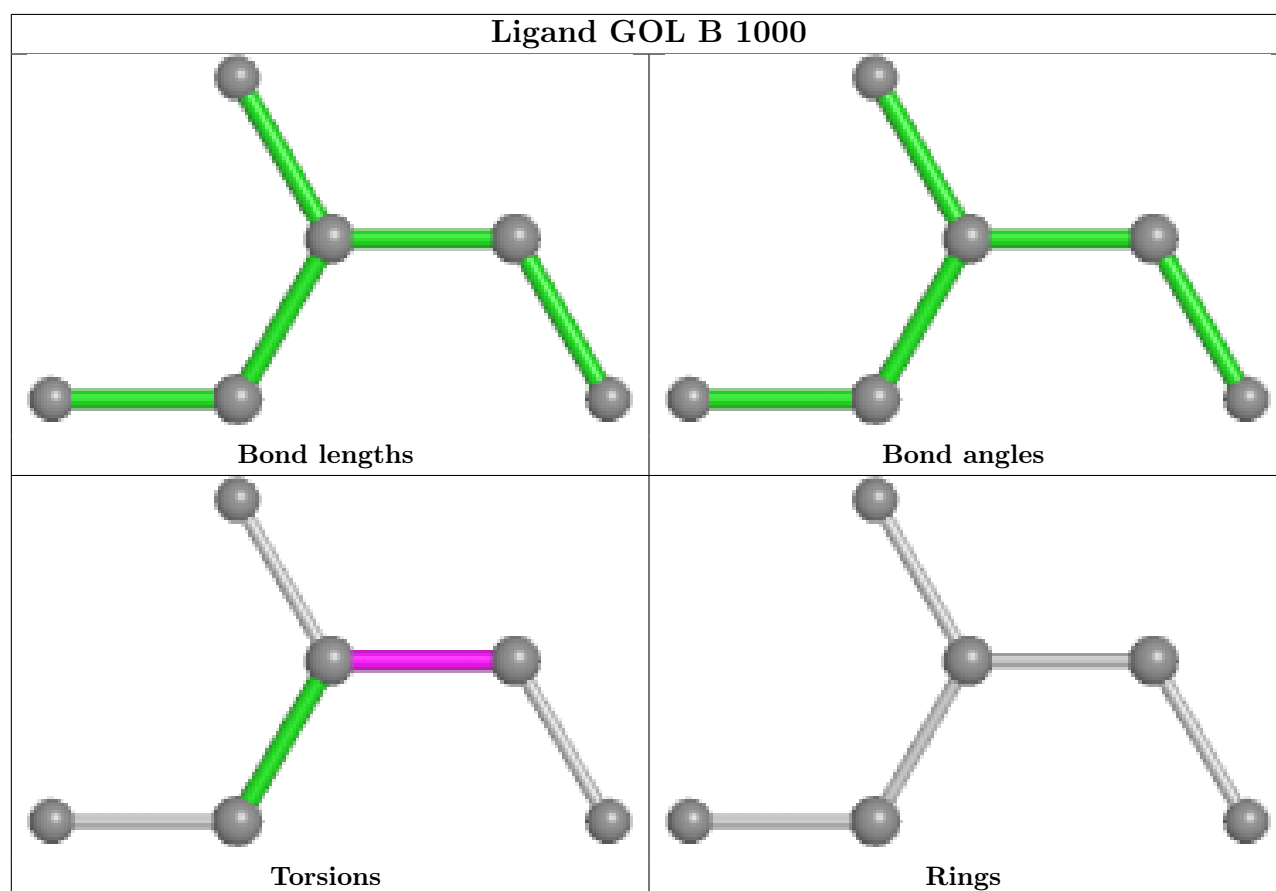
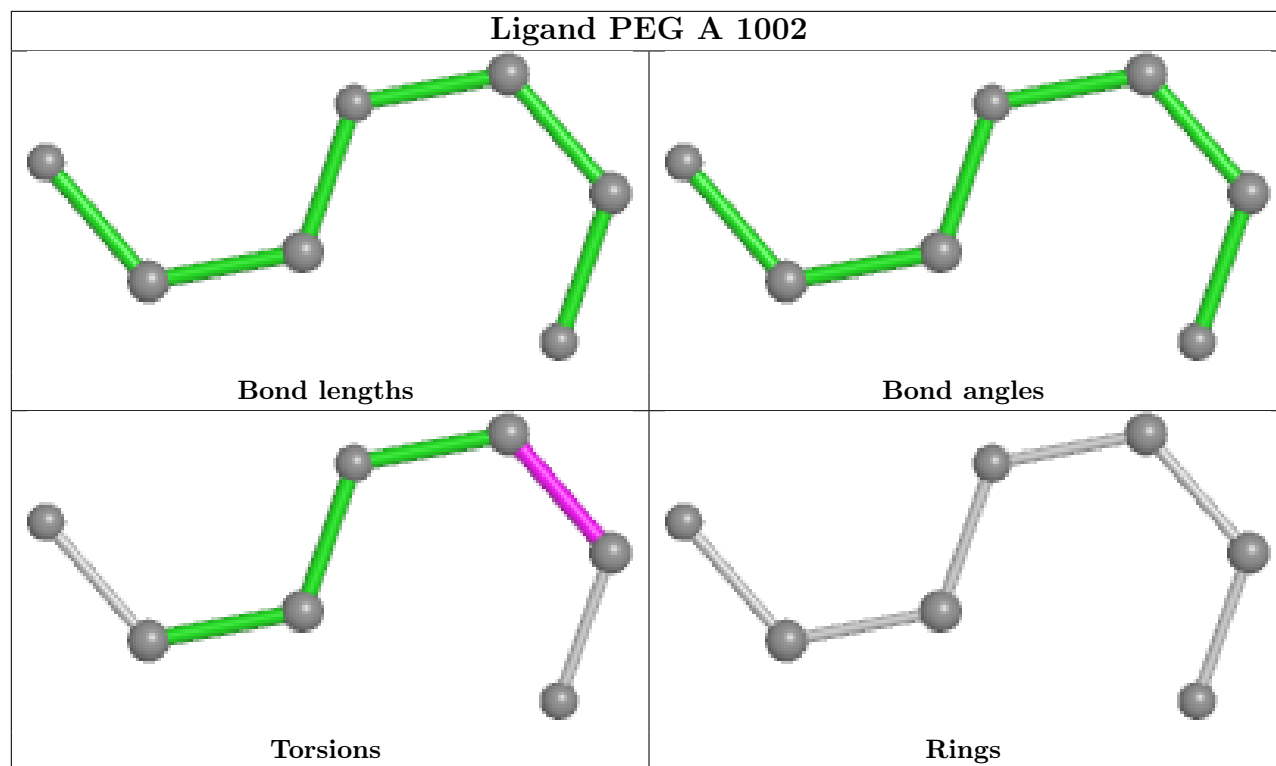












## 5.7 Other polymers [i](#)

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 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.