



## wwPDB EM Validation Summary Report ⓘ

Oct 23, 2023 – 12:09 PM JST

PDB ID : 8GY6  
EMDB ID : EMD-34370  
Title : Structure of SARS-CoV-2 RNA-dependent RNA polymerase with gossypol binding  
Authors : Wang, W.; Ren, M.; Li, F.  
Deposited on : 2022-09-21  
Resolution : Not provided

This is a wwPDB EM 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/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:

EMDB validation analysis : 0.0.1.dev70  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
MolProbity : 4.02b-467  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

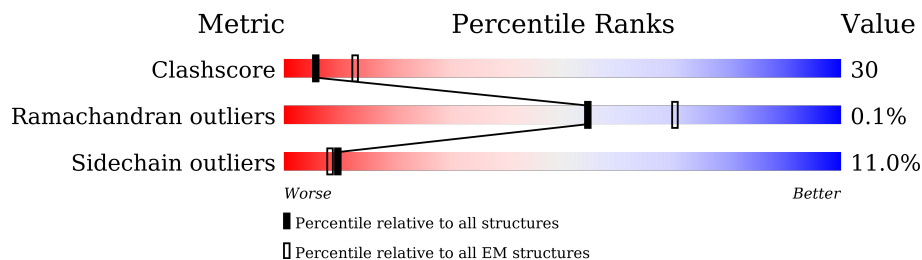
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is unknown.

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)
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	932	
2	B	198	
2	D	198	
3	C	83	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 8587 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 RNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	859	6903	4424	1154	1279	46	0	0

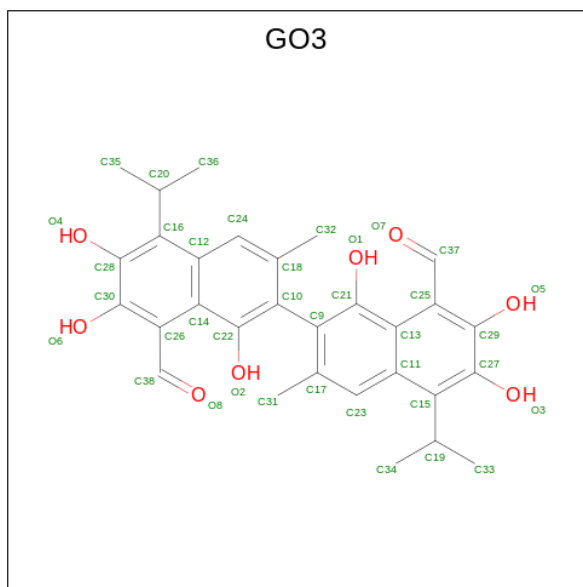
- Molecule 2 is a protein called Non-structural protein 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	113	858	543	142	166	7	0	0
2	D	43	315	199	51	60	5	0	0

- Molecule 3 is a protein called Non-structural protein 7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	56	435	275	72	84	4	0	0

- Molecule 4 is Gossypol (three-letter code: GO3) (formula: C<sub>30</sub>H<sub>30</sub>O<sub>8</sub>) (labeled as "Ligand of Interest" by depositor).

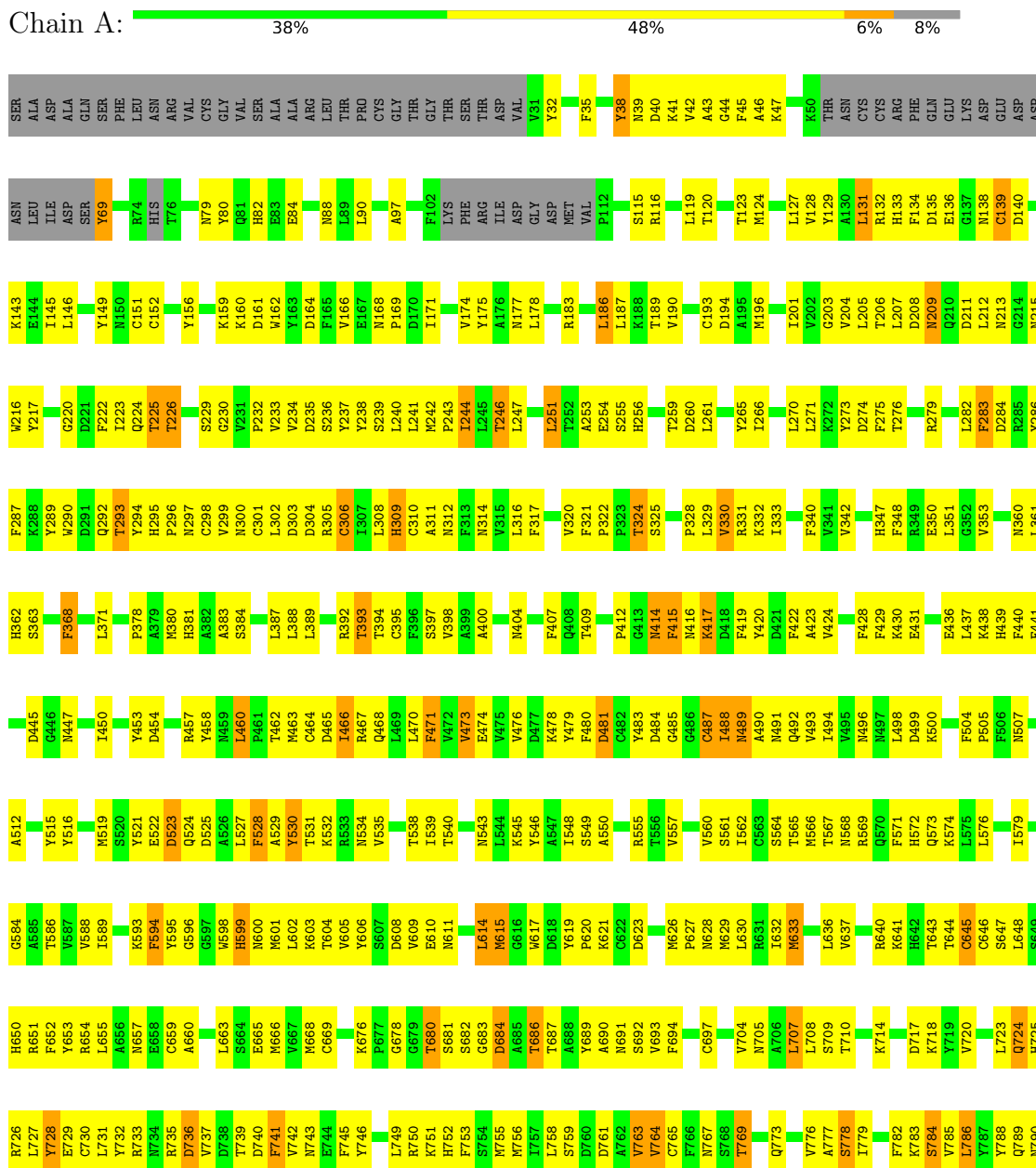


Mol	Chain	Residues	Atoms			AltConf
4	A	1	Total	C	O	0
			38	30	8	
4	A	1	Total	C	O	0
			38	30	8	

### 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: RNA-directed RNA polymerase





## 4 Experimental information

Property	Value	Source
EM reconstruction method	TOMOGRAPHY	Depositor
Imposed symmetry	POINT, Not provided	
Number of tilted images used	1224147	Depositor
Resolution determination method	Not provided	
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	50.0	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor

## 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: GO3

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	0/7078	0.53	0/9609
2	B	0.28	0/870	0.53	0/1189
2	D	0.26	0/316	0.45	0/429
3	C	0.29	0/438	0.46	0/590
All	All	0.40	0/8702	0.52	0/11817

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	6903	0	6648	457	0
2	B	858	0	857	53	0
2	D	315	0	318	12	0
3	C	435	0	459	17	0
4	A	76	0	0	14	0
All	All	8587	0	8282	514	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 30.



The worst 5 of 514 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:593:LYS:HE3	1:A:594:PHE:CE1	1.94	1.02
4:A:1001:GO3:O1	4:A:1001:GO3:C32	2.12	0.97
1:A:116:ARG:HD2	1:A:217:TYR:HB2	1.53	0.89
2:B:101:ASP:HA	2:B:104:ASN:HD22	1.42	0.83
1:A:892:HIS:CE1	1:A:916:TRP:CZ2	2.67	0.83

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 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	849/932 (91%)	702 (83%)	146 (17%)	1 (0%)	51	51
2	B	109/198 (55%)	88 (81%)	21 (19%)	0	100	100
2	D	39/198 (20%)	37 (95%)	2 (5%)	0	100	100
3	C	54/83 (65%)	51 (94%)	3 (6%)	0	100	100
All	All	1051/1411 (74%)	878 (84%)	172 (16%)	1 (0%)	54	51

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	610	GLU

### 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	749/823 (91%)	661 (88%)	88 (12%)	5	5
2	B	95/167 (57%)	82 (86%)	13 (14%)	3	3
2	D	36/167 (22%)	34 (94%)	2 (6%)	21	21
3	C	53/77 (69%)	53 (100%)	0	100	100
All	All	933/1234 (76%)	830 (89%)	103 (11%)	10	6

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	643	THR
1	A	730	CYS
2	B	185	ILE
1	A	645	CYS
1	A	692	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	698	GLN
1	A	789	GLN
2	D	109	ASN
1	A	724	GLN
1	A	790	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 monosaccharides in this entry.

## 5.6 Ligand geometry

2 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
4	GO3	A	1002	-	41,41,41	1.62	7 (17%)	64,64,64	2.62	22 (34%)
4	GO3	A	1001	-	41,41,41	1.62	7 (17%)	64,64,64	2.54	20 (31%)

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
4	GO3	A	1002	-	-	8/16/16/16	0/4/4/4
4	GO3	A	1001	-	-	8/16/16/16	0/4/4/4

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1001	GO3	C10-C9	3.92	1.58	1.50
4	A	1002	GO3	C10-C9	3.92	1.58	1.50
4	A	1001	GO3	C26-C38	3.89	1.53	1.45
4	A	1001	GO3	C25-C37	3.88	1.53	1.45
4	A	1002	GO3	C25-C37	3.87	1.53	1.45

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1002	GO3	C32-C18-C10	10.11	132.46	121.34
4	A	1001	GO3	C32-C18-C10	10.08	132.41	121.34
4	A	1001	GO3	C32-C18-C24	-6.33	105.21	120.34
4	A	1002	GO3	C32-C18-C24	-6.26	105.36	120.34
4	A	1002	GO3	C9-C10-C22	5.86	129.51	119.67

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

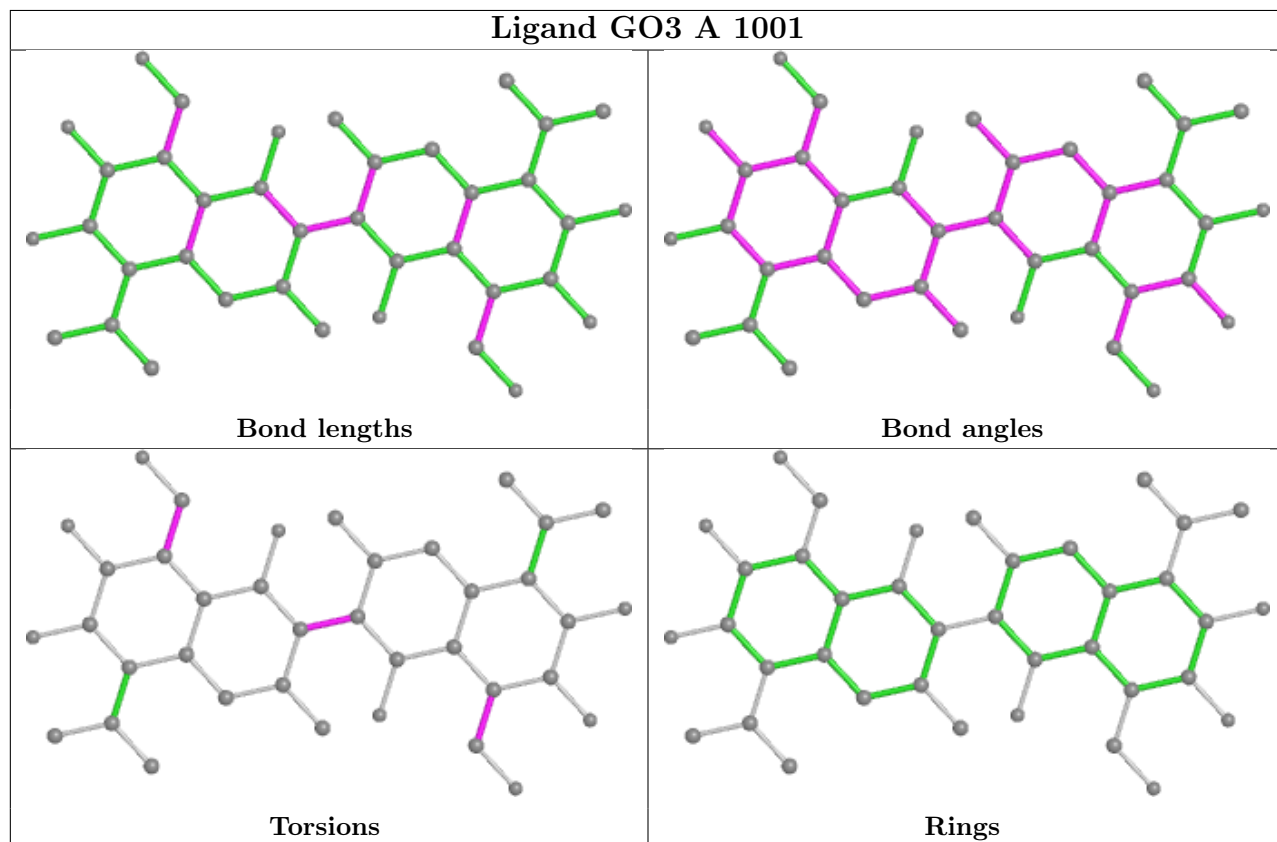
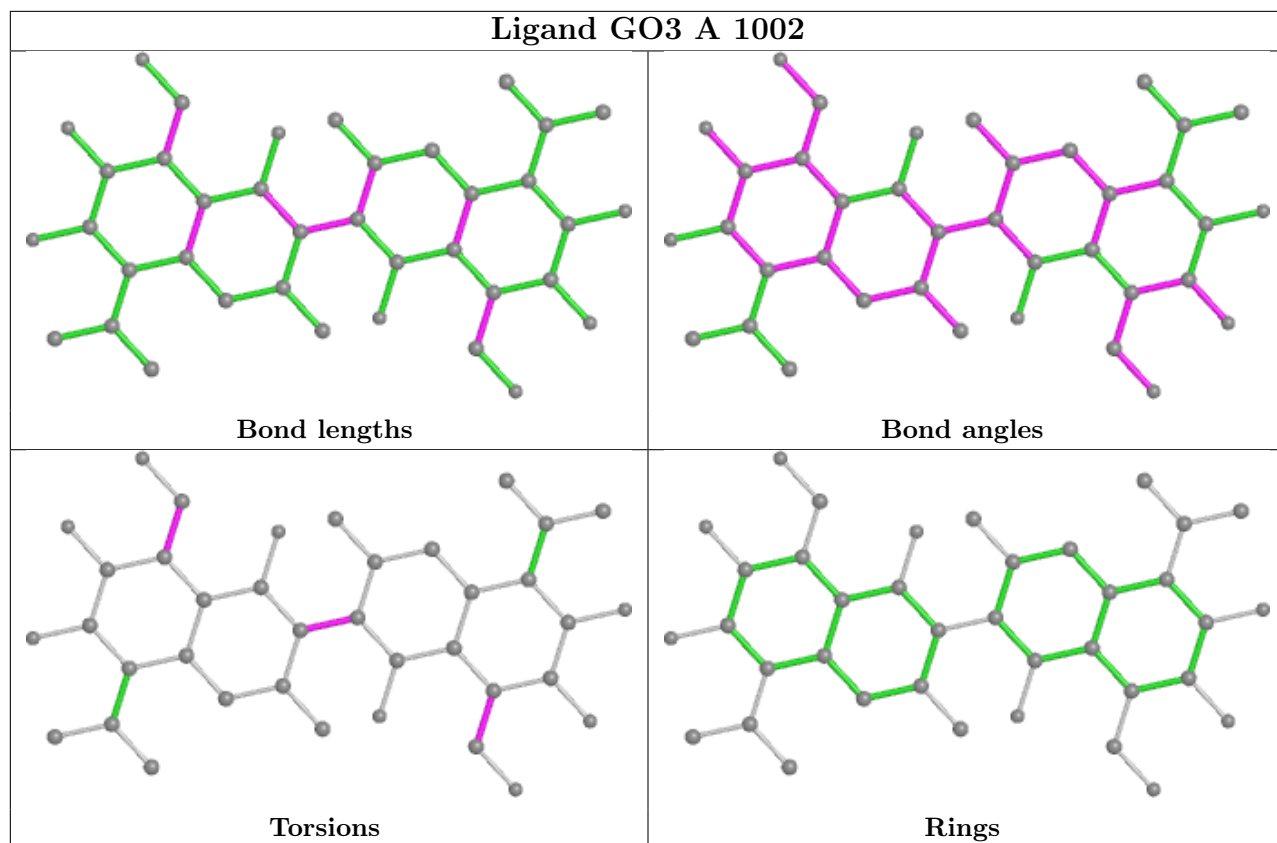
Mol	Chain	Res	Type	Atoms
4	A	1001	GO3	C18-C10-C9-C17
4	A	1001	GO3	C22-C10-C9-C17
4	A	1001	GO3	C18-C10-C9-C21
4	A	1001	GO3	C22-C10-C9-C21
4	A	1001	GO3	C13-C25-C37-O7

There are no ring outliers.

2 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1002	GO3	6	0
4	A	1001	GO3	8	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 Tomogram visualisation

This section contains visualisations of the EMDB entry EMD-34370. These allow visual inspection of the internal detail of the tomogram and identification of artifacts.

### 6.1 Orthogonal projections

This section was not generated.

### 6.2 Central slices

This section was not generated.

### 6.3 Largest variance slices

This section was not generated.

### 6.4 Orthogonal standard-deviation projections (False-color)

This section was not generated.

### 6.5 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

## 7 Tomogram analysis

This section contains the results of statistical analysis of the tomogram.

### 7.1 Map-value distribution

This section was not generated.



## 8 Map-model fit

This section was not generated.