



wwPDB EM Validation Summary Report ⓘ

Jan 8, 2024 – 04:12 PM JST

PDB ID : 8K9Q
EMDB ID : EMD-36995
Title : Cryo-EM structure of the GPI inositol-deacylase (PGAP1/Bst1) from *Chaetomium thermophilum*
Authors : Hong, J.; Li, T.; Qu, Q.; Li, D.
Deposited on : 2023-08-01
Resolution : 2.84 Å(reported)

This is a wwPDB EM Validation Summary 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/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 : **FAILED**
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

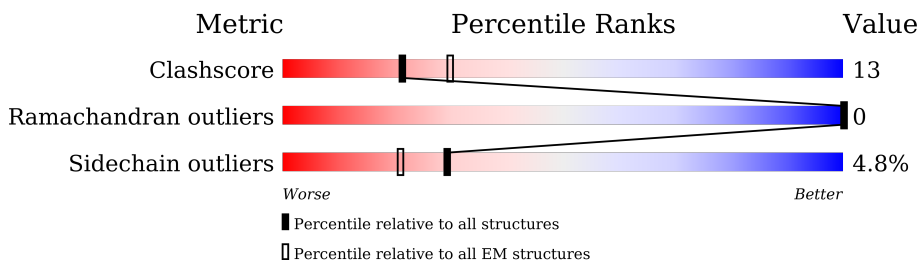
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 2.84 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)	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 ≥ 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	1469	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 7549 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.

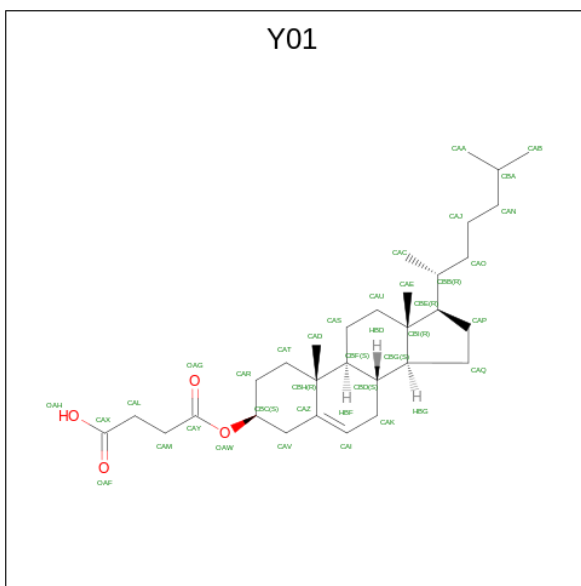
- Molecule 1 is a protein called GPI inositol-deacylase, fused thermostable green fluorescent protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	927	7340	4759	1238	1308	35	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	initiating methionine	UNP G0S652
A	0	GLY	-	expression tag	UNP G0S652
A	1	SER	-	expression tag	UNP G0S652

- Molecule 2 is CHOLESTEROL HEMISUCCINATE (three-letter code: Y01) (formula: $C_{31}H_{50}O_4$).



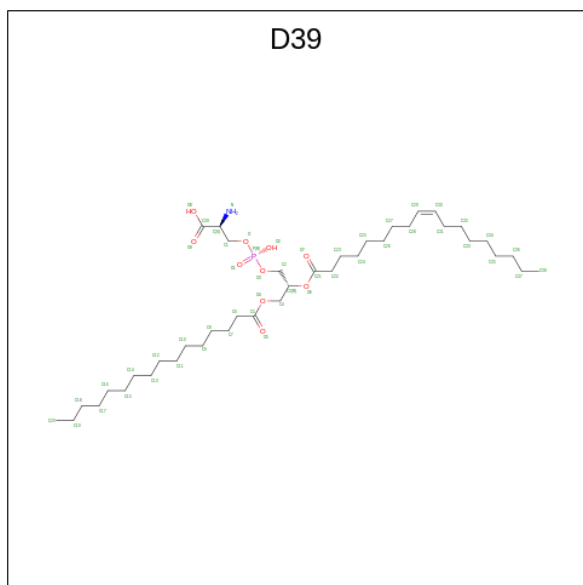
Mol	Chain	Residues	Atoms			AltConf
2	A	1	Total	C	O	0
2	A	1	35	31	4	0
2	A	1	Total	C	O	0
2	A	1	35	31	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			AltConf
			Total	C	O	
2	A	1	35	31	4	0

- Molecule 3 is (2 {S})-2-azanyl-3-[[[(2 {R})-3-hexadecanoyloxy-2-[({Z})-octadec-9-enoyl]oxy-propoxy]-oxidanyl-phosphoryl]oxy-propanoic acid (three-letter code: D39) (formula: C₄₀H₇₆NO₁₀P).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
3	A	1	52	40	1	10	1	0
3	A	1	52	40	1	10	1	0

GLU SER
 GLU ASP
 GLY ILE
 GLY ILE
 ALA HIS
 ALA TRP
 LEU MET
 LEU TYR
 PRO ARG
 PHE CYS
 PHE ASP
 TYR PHE
 TYR ASP
 ILE THR
 LEU TYR
 THR ILE
 THR THR
 PRO ARG
 ALA PHE
 ALA PHE
 GLN GLY
 TYR THR
 GLY ASP
 ASP THR
 ASN THR
 PHE ASN
 ARG PRO
 ARG PRO
 ALA ALA
 ALA PHE
 ASN ASN
 GLY THR
 THR THR
 LYS THR
 TYR TYR
 TYR MET
 PRO PRO
 ASP ASP
 TYR TYR
 TYR PHE
 LYS LYS
 GLU GLU
 GLN GLN
 ALA ALA
 TYR TYR
 GLU GLU
 LYS LYS
 TYR TYR
 VAL VAL
 SER SER
 TRP TRP
 GLU GLU
 ARG ARG
 MET MET
 THR THR
 TYR TYR
 GLU GLU
 ASP ASP
 VAL VAL
 THR THR
 GLU GLU
 LYS LYS
 MET MET
 TYR TYR
 VAL VAL
 SER SER
 TRP TRP
 GLU GLU
 ARG ARG
 VAL VAL
 LEU LEU
 TYR TYR
 GLY GLY
 HIS HIS
 VAL VAL
 ARG ARG
 LEU LEU
 TYR TYR
 GLU GLU
 HIS HIS
 HIS HIS
 ALA ALA
 GLU GLU
 ALA ALA
 TYR TYR
 ASP ASP
 LYS LYS
 HIS HIS
 HIS HIS
 HIS HIS
 HIS HIS
 HIS HIS
 HIS HIS

SER ASP
 ASP ASP
 THR THR
 MET MET
 TYR TYR
 GLU GLU
 CYS CYS
 ASP ASP
 PHE PHE
 TYR TYR
 PHE PHE
 THR THR
 TYR TYR
 GLU GLU
 ILE ILE
 ARG ARG
 PHE PHE
 LYS LYS
 ASP ASP
 THR THR
 ASN ASN
 PHE PHE
 ASN ASN
 PRO PRO
 PRO PRO
 VAL VAL
 MET MET
 THR THR
 GLN GLN
 LYS LYS
 LYS LYS
 ILE ILE
 THR THR
 PRO PRO
 LEU LEU
 LYS LYS
 TRP TRP
 PHE PHE
 LYS LYS
 TYR TYR
 THR THR
 GLU GLU
 LYS LYS
 TYR TYR
 VAL VAL
 ASP ASP
 VAL VAL
 THR THR
 MET MET
 LEU LEU
 LEU LEU
 LEU LEU
 GLU GLU

GLY GLY
 GLY GLY
 HIS HIS
 TYR TYR
 ARG ARG
 CYS CYS
 ASP ASP
 PHE PHE
 LYS LYS
 THR THR
 THR THR
 TYR TYR
 LYS LYS
 ALA ALA
 LYS LYS
 LYS LYS
 VAL VAL
 ARG ARG
 LEU LEU
 PRO PRO
 ASP ASP
 ALA ALA
 TRP TRP
 ALA ALA
 HIS HIS
 GLU GLU
 VAL VAL
 ASP ASP
 HIS HIS
 ARG ARG
 ILE ILE
 GLU GLU
 ILE ILE
 LEU LEU
 SER SER
 HIS HIS
 HIS HIS
 LYS LYS
 ASP ASP
 LYS LYS
 TYR TYR
 ASN ASN
 LYS LYS
 VAL VAL
 ARG ARG
 LEU LEU
 TYR TYR
 GLU GLU
 HIS HIS
 HIS HIS
 ALA ALA
 GLU GLU
 ALA ALA
 ARG ARG
 TYR TYR
 SER SER
 GLY GLY
 GLY GLY
 SER SER
 GLY GLY
 GLY GLY

GLY SER
 SER SER
 ALA ALA
 TRP TRP
 SER SER
 HIS HIS
 PRO PRO
 GLN GLN
 PHE PHE
 GLU GLU
 LYS LYS
 GLY GLY
 GLY GLY
 GLY GLY
 SER SER
 GLY GLY
 GLY GLY
 GLY GLY
 SER SER
 GLY GLY
 GLY GLY
 TRP TRP
 ALA ALA
 SER SER
 HIS HIS
 PRO PRO
 GLN GLN
 PHE PHE
 GLU GLU
 LYS LYS
 GLY GLY
 SER SER
 HIS HIS
 HIS HIS
 HIS HIS
 HIS HIS
 HIS HIS
 HIS HIS
 HIS HIS

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	407921	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	60.0	Depositor
Minimum defocus (nm)	1200	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: Y01, D39

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	0.72	0/7530	0.87	0/10269

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	7340	0	7394	191	0
2	A	105	0	147	19	0
3	A	104	0	0	0	0
All	All	7549	0	7541	198	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 13.

The worst 5 of 198 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:271:VAL:HG11	1:A:292:LEU:HD21	1.19	1.13

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:689:LEU:HD13	1:A:694:ILE:HG13	1.32	1.09
1:A:689:LEU:HD13	1:A:694:ILE:CG1	1.86	1.06
1:A:174:PRO:HB2	1:A:627:PHE:HZ	1.19	1.03
1:A:689:LEU:CD1	1:A:694:ILE:HG13	1.87	1.02

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	915/1469 (62%)	897 (98%)	18 (2%)	0	100 100

There are no Ramachandran outliers to report.

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	819/1257 (65%)	780 (95%)	39 (5%)	25 49

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	811	LYS
1	A	927	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	817	ARG
1	A	826	PHE
1	A	1078	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1047	ASN
1	A	1056	HIS
1	A	1069	HIS
1	A	1057	ASN
1	A	786	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](#)

5 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	$\# Z > 2$
2	Y01	A	1502	-	38,38,38	0.56	0	57,57,57	1.24	5 (8%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	Y01	A	1503	-	38,38,38	0.44	0	57,57,57	0.44	0
3	D39	A	1504	-	50,51,51	0.40	0	54,58,58	0.51	0
2	Y01	A	1501	-	38,38,38	0.51	0	57,57,57	0.54	0
3	D39	A	1505	-	50,51,51	0.50	0	54,58,58	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	Y01	A	1502	-	-	14/19/77/77	0/4/4/4
2	Y01	A	1503	-	-	11/19/77/77	0/4/4/4
3	D39	A	1504	-	-	28/57/57/57	-
2	Y01	A	1501	-	-	13/19/77/77	0/4/4/4
3	D39	A	1505	-	-	39/57/57/57	-

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1502	Y01	CAQ-CBG-CBI	3.04	107.50	103.84
2	A	1502	Y01	CBG-CBI-CBE	2.88	103.48	100.07
2	A	1502	Y01	CAU-CBI-CBG	-2.88	102.81	107.27
2	A	1502	Y01	CAP-CAQ-CBG	-2.64	99.90	105.13
2	A	1502	Y01	CBF-CBD-CBG	-2.61	105.59	109.09

There are no chirality outliers.

5 of 105 torsion outliers are listed below:

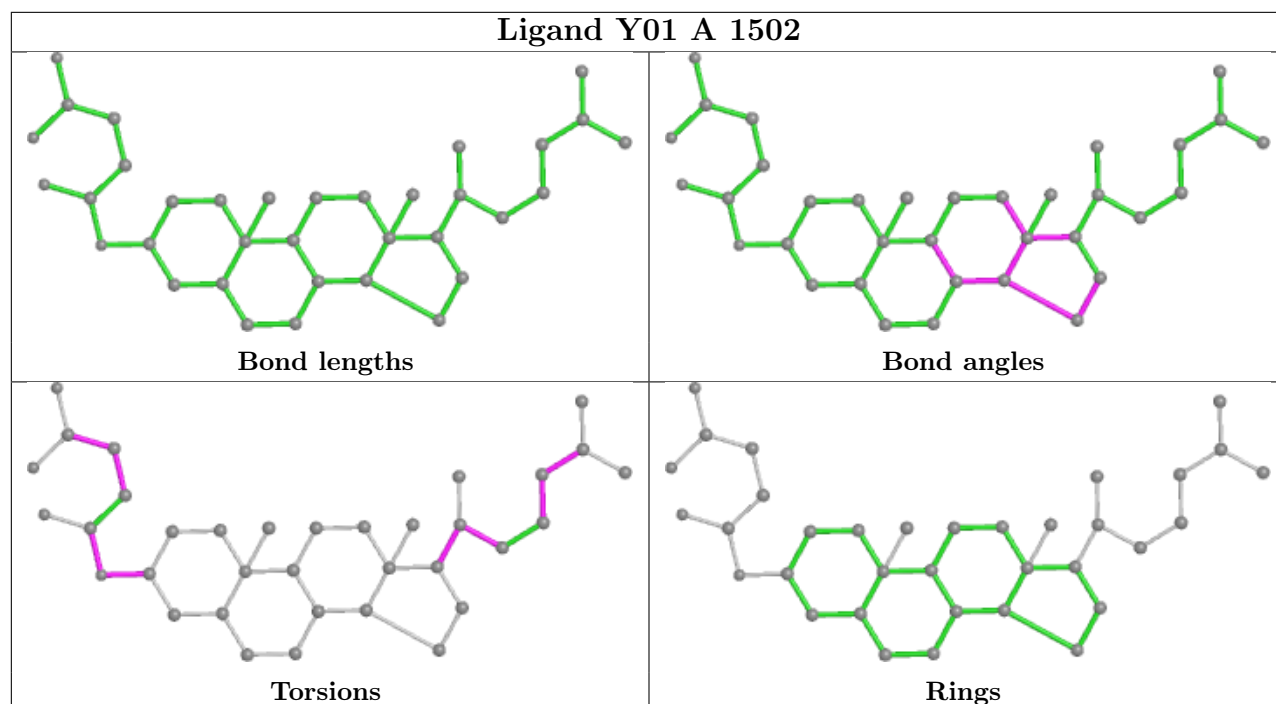
Mol	Chain	Res	Type	Atoms
2	A	1501	Y01	OAG-CAY-OAW-CBC
2	A	1502	Y01	CAO-CBB-CBE-CBI
2	A	1502	Y01	CAC-CBB-CBE-CBI
2	A	1502	Y01	CAV-CBC-OAW-CAY
2	A	1503	Y01	CAV-CBC-OAW-CAY

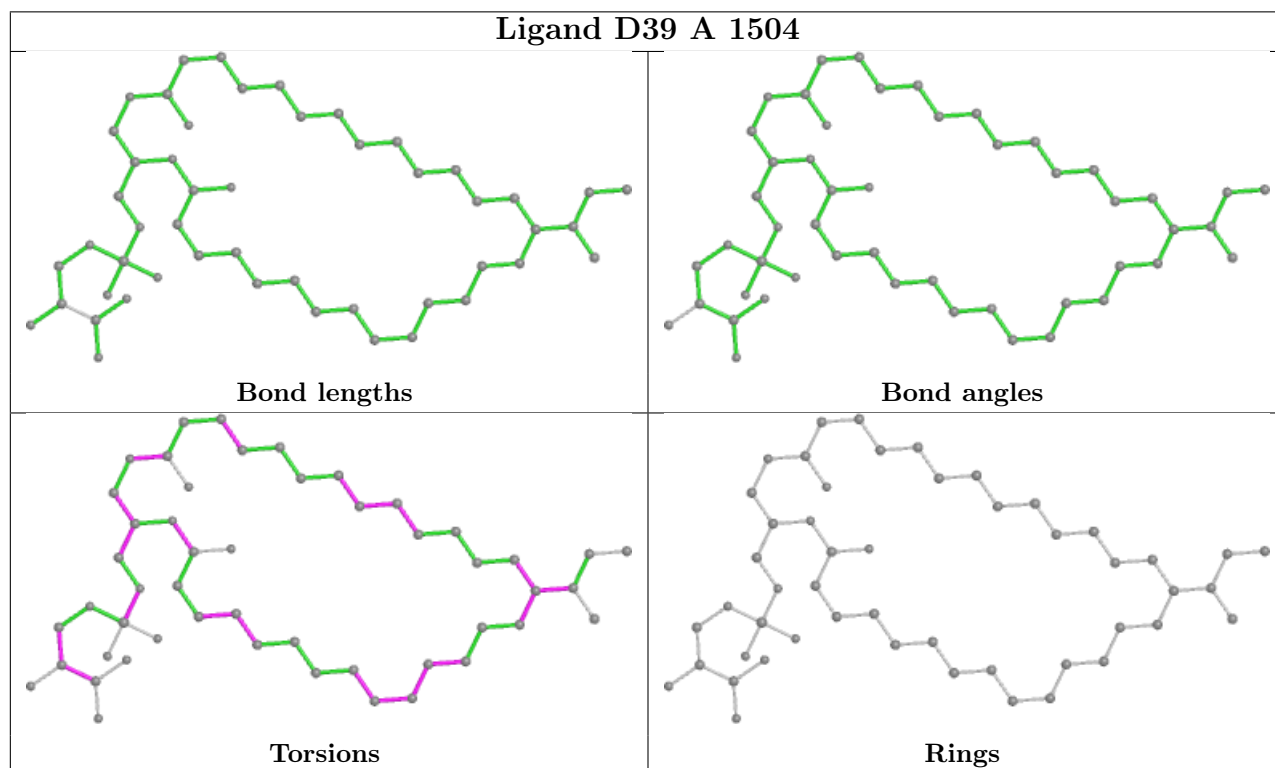
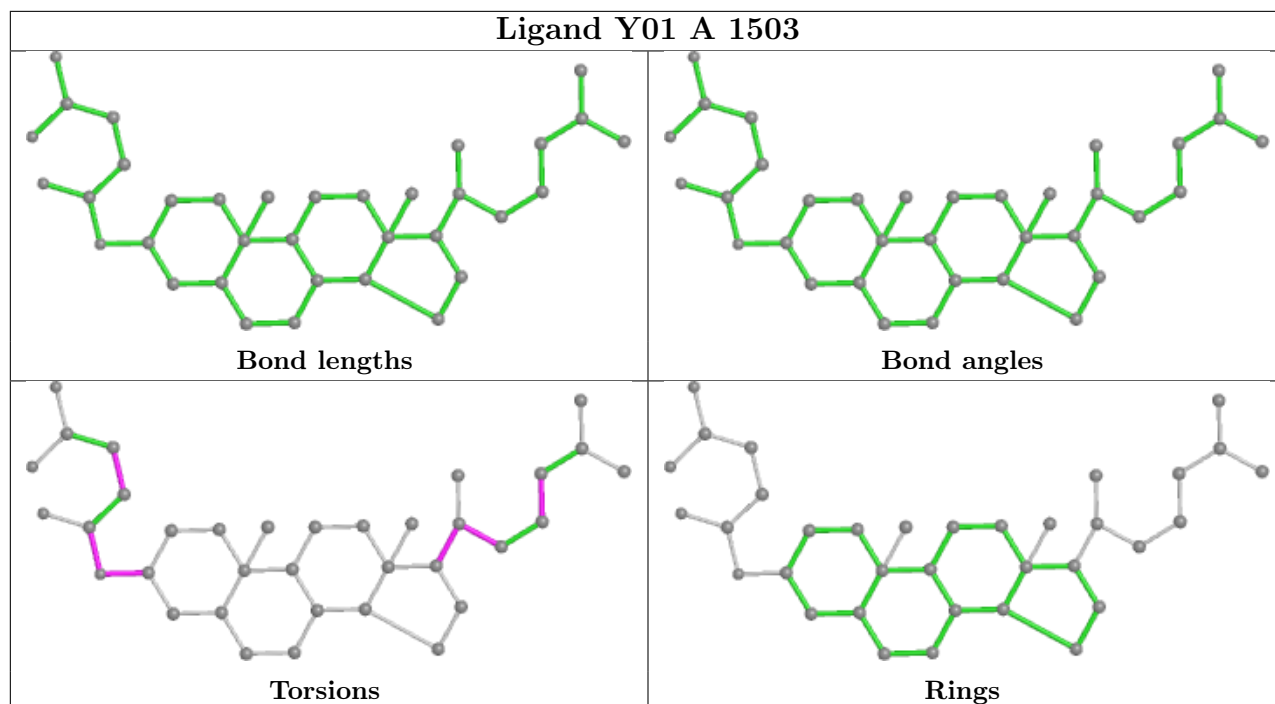
There are no ring outliers.

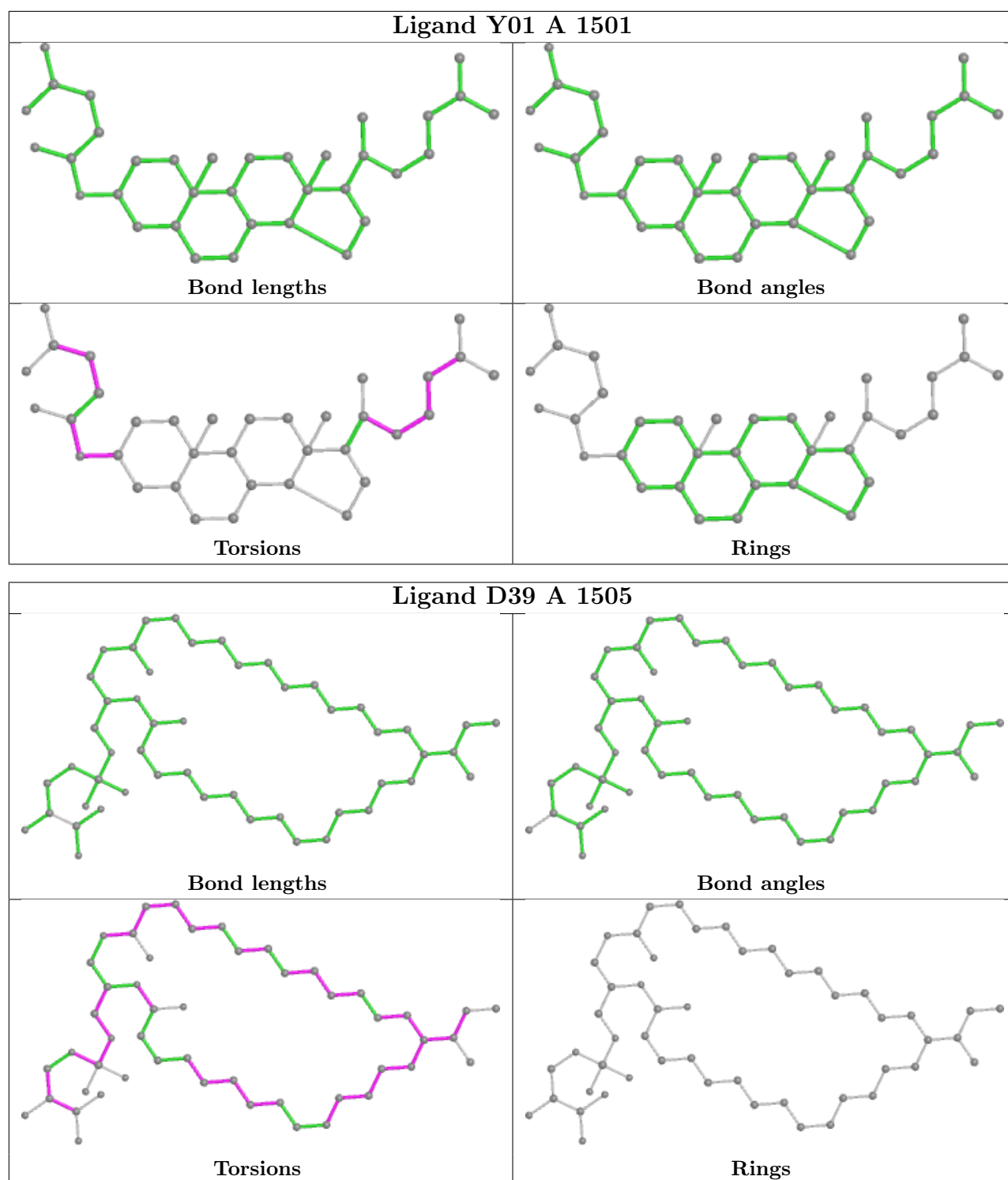
3 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1502	Y01	2	0
2	A	1503	Y01	8	0
2	A	1501	Y01	9	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

There are no chain breaks in this entry.