W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB EM Validation Summary Report ⓘ

Jan 8, 2024 – 04:14 PM JST

|		
PDB ID	:	8K9T
EMDB ID	:	EMD-36997
Title	:	Cryo-EM structure of the products-bound PGAP1(Bst1)-S327A from Chaetonium thermophilum
Authors	:	Li, T.; Hong, J.; Qu, Q.; Li, D.
Deposited on	:	2023-08-01
Resolution	:	2.66 Å(reported)

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 2.66 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	Percentile Ranks	Value
Clashscore		6
Ramachandran outliers		0.1%
Sidechain outliers		1.1%

Worse   *Better*

▌Percentile relative to all structures
◫ Percentile relative to all EM structures

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	1447	56% 10% 34%
2	B	272	· 99%

2   Entry composition   ⓘ

There are 9 unique types of molecules in this entry. The entry contains 7747 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,MCherry protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	961	7520	4882	1269	1336	33	0	0

There are 34 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
A	327	ALA	SER	engineered mutation	UNP G0S652
A	1185	GLY	-	linker	UNP G0S652
A	1186	THR	-	linker	UNP G0S652
A	1187	LEU	-	linker	UNP G0S652
A	1188	GLU	-	linker	UNP G0S652
A	1189	VAL	-	linker	UNP G0S652
A	1190	LEU	-	linker	UNP G0S652
A	1191	PHE	-	linker	UNP G0S652
A	1192	GLN	-	linker	UNP G0S652
A	1193	GLY	-	linker	UNP G0S652
A	1194	PRO	-	linker	UNP G0S652
A	1195	LYS	-	linker	UNP G0S652
A	1196	LEU	-	linker	UNP G0S652
A	1197	GLU	-	linker	UNP G0S652
A	1198	PHE	-	linker	UNP G0S652
A	1345	SER	TRP	engineered mutation	UNP A0A366VY15
A	1363	VAL	ILE	engineered mutation	UNP A0A366VY15
A	1365	TYR	GLN	engineered mutation	UNP A0A366VY15
A	1399	ARG	ILE	engineered mutation	UNP A0A366VY15
A	1434	SER	-	expression tag	UNP A0A366VY15
A	1435	ALA	-	expression tag	UNP A0A366VY15
A	1436	HIS	-	expression tag	UNP A0A366VY15
A	1437	HIS	-	expression tag	UNP A0A366VY15
A	1438	HIS	-	expression tag	UNP A0A366VY15
A	1439	HIS	-	expression tag	UNP A0A366VY15

Continued on next page...
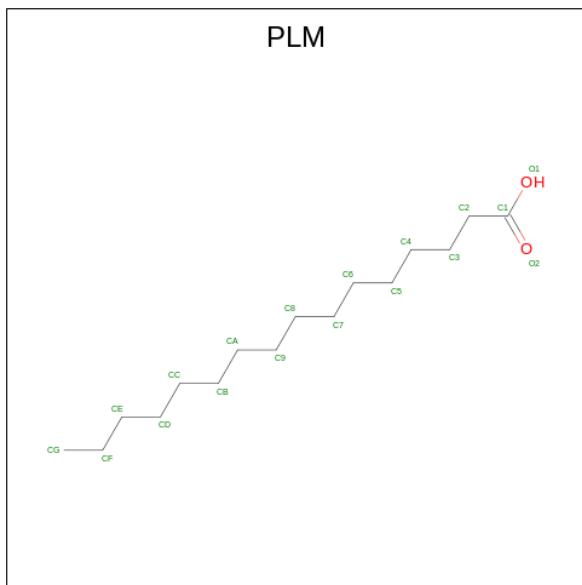
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1440	HIS	-	expression tag	UNP A0A366VY15
A	1441	HIS	-	expression tag	UNP A0A366VY15
A	1442	HIS	-	expression tag	UNP A0A366VY15
A	1443	HIS	-	expression tag	UNP A0A366VY15
A	1444	HIS	-	expression tag	UNP A0A366VY15
A	1445	HIS	-	expression tag	UNP A0A366VY15

- Molecule 2 is a protein called Green fluorescent protein,Complement decay-accelerating factor.
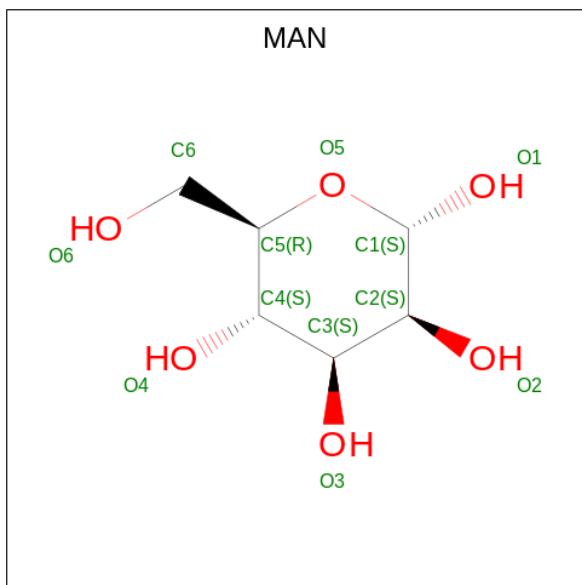
Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
2	B	4	24	13	4	7	0	0

- Molecule 3 is PALMITIC ACID (three-letter code: PLM) (formula: $C_{16}H_{32}O_2$).

PLM

Mol	Chain	Residues	Atoms			AltConf
			Total	C	O	
3	A	1	18	16	2	0
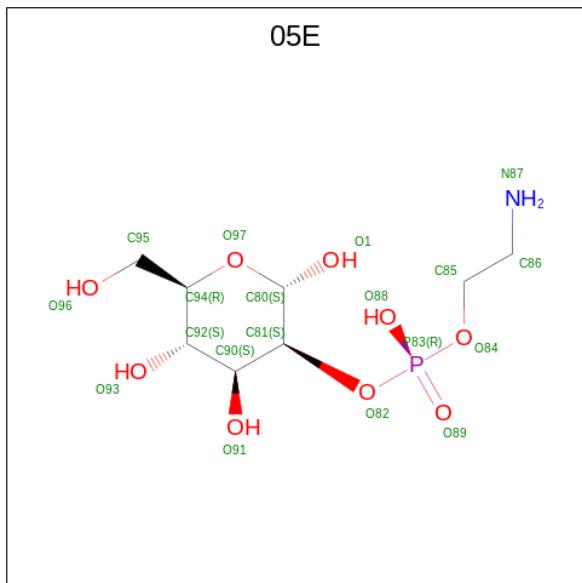
- Molecule 4 is alpha-D-mannopyranose (three-letter code: MAN) (formula: $C_6H_{12}O_6$) (labeled as "Ligand of Interest" by depositor).

MAN

Mol	Chain	Residues	Atoms			AltConf
4	A	1	Total 11	C 6	O 5	0
4	A	1	Total 11	C 6	O 5	0

- Molecule 5 is 2-azanylethyl [(2 {S},3 {S},4 {S},5 {S},6 {R})-6-(hydroxymethyl)-2,4,5-tri s(oxidanyl)oxan-3-yl] hydrogen phosphate (three-letter code: 05E) (formula: $C_8H_{18}NO_9P$) (labeled as "Ligand of Interest" by depositor).

05E
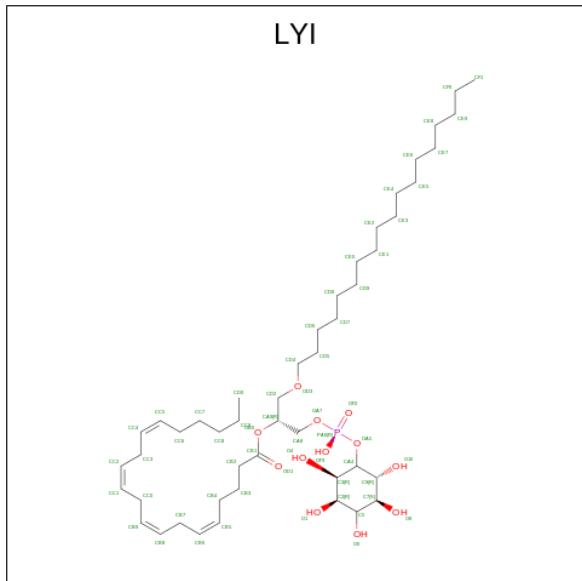
Mol	Chain	Residues	Atoms					AltConf
5	A	1	Total 18	C 8	N 1	O 8	P 1	0

- Molecule 6 is 2-amino-2-deoxy-alpha-D-glucopyranose (three-letter code: PA1) (formula: $C_6H_{13}NO_5$) (labeled as "Ligand of Interest" by depositor).
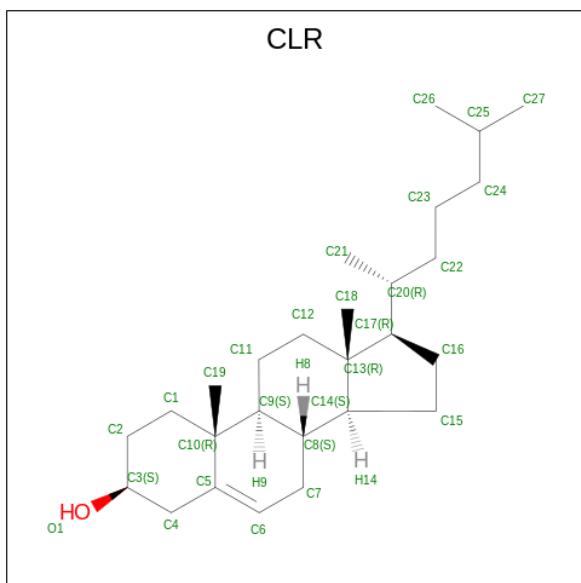
Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
6	A	1	11	6	1	4	0

- Molecule 7 is [(2 {R})-1-octadecoxy-3-[oxidanyl-[(2 {R},3 {R},5 {S},6 {R})-2,3,4,5,6-penta kis(oxidanyl)cyclohexyl]oxy-phosphoryl]oxy-propan-2-yl] (5 {Z},8 {Z},11 {Z},14 {Z})-icosa-5,8,11,14-tetraenoate (three-letter code: LYI) (formula: $C_{47}H_{85}O_{12}P$) (labeled as "Ligand of Interest" by depositor).
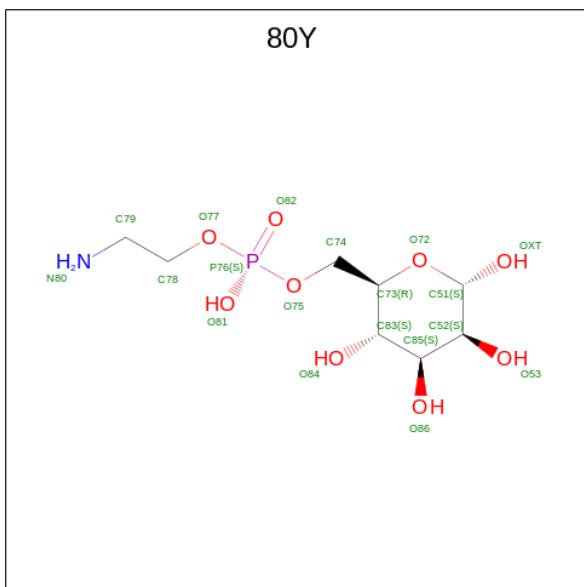
Mol	Chain	Residues	Atoms				AltConf
			Total	C	O	P	
7	A	1	60	47	12	1	0

- Molecule 8 is CHOLESTEROL (three-letter code: CLR) (formula: $C_{27}H_{46}O$).

Mol	Chain	Residues	Atoms			AltConf
8	A	1	Total	C	O	0
			28	27	1	
8	A	1	Total	C	O	0
			28	27	1	

- Molecule 9 is 2-azanylethyl [(2R,3S,4S,5S,6S)-3,4,5,6-tetrakis(oxidanyl)oxan-2-yl]methyl hydrogen phosphate (three-letter code: 80Y) (formula: $C_8H_{18}NO_9P$) (labeled as "Ligand of Interest" by depositor).
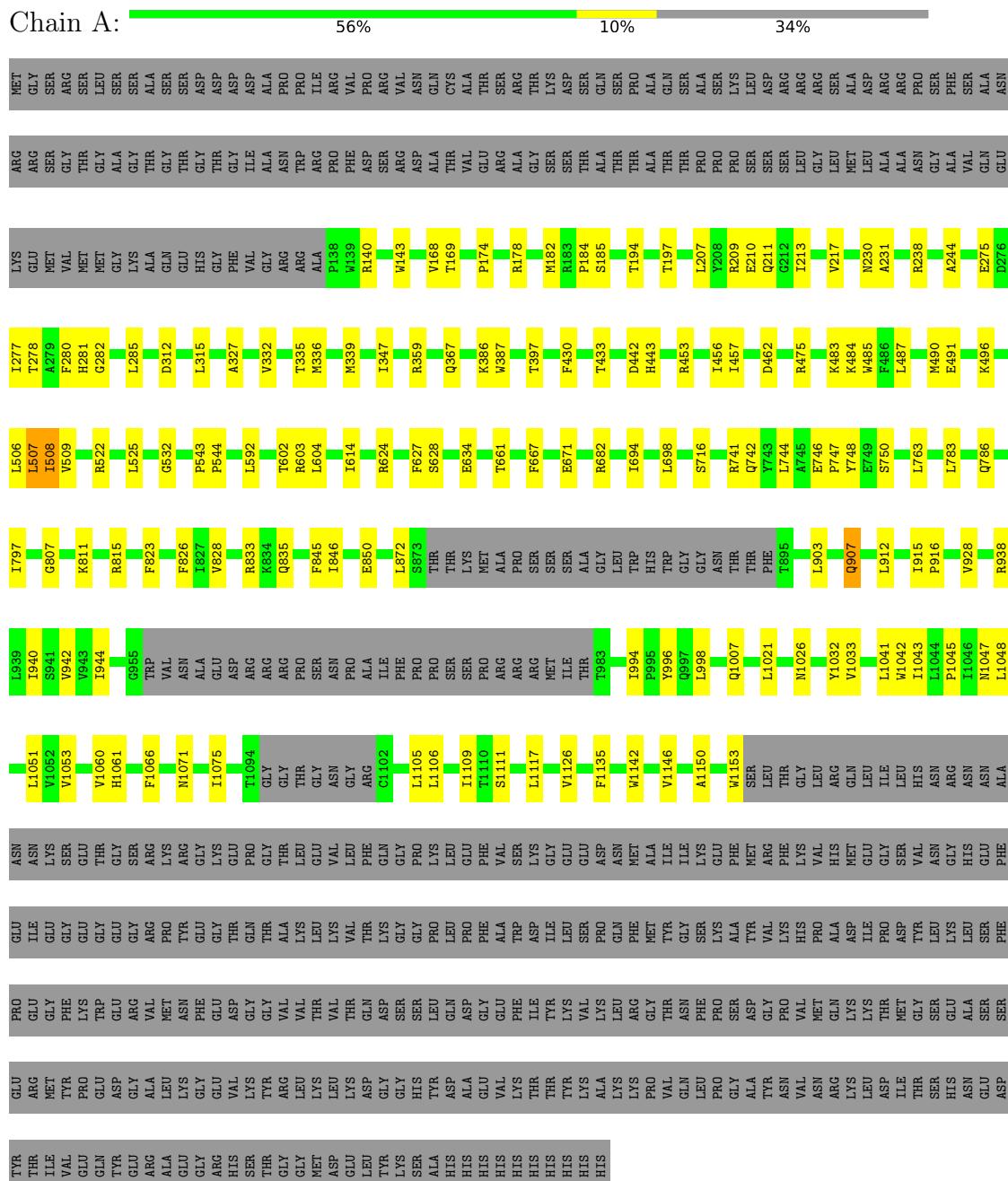
Mol	Chain	Residues	Atoms					AltConf
9	B	1	Total	C	N	O	P	0
			18	8	1	8	1	

3　Residue-property plots ⓘ

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: GPI inositol-deacylase,MCherry protein

Chain A:

• Molecule 2: Green fluorescent protein,Complement decay-accelerating factor

Chain B: • 99%

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

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

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

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

LYS GLY GLY GLY SER GLY GLY GLY SER GLY GLY SER ALA TRP SER HIS PRO GLN PHE GLU LYS GLY SER GLY SER PRO ASN LYS GLY SER G1 T2 T3 S4

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	353585	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 $(e^-/\text{Å}^2)$	50	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 05E, PA1, PLM, 80Y, LYI, MAN, CLR

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	0.42	0/7722	0.52	0/10549
2	B	0.53	0/23	0.93	0/30
All	All	0.42	0/7745	0.52	0/10579

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 ⓘ

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	7520	0	7512	93	0
2	B	24	0	24	3	0
3	A	18	0	31	1	0
4	A	22	0	19	1	0
5	A	18	0	0	0	0
6	A	11	0	10	1	0
7	A	60	0	0	0	0
8	A	56	0	92	10	0
9	B	18	0	0	0	0
All	All	7747	0	7688	97	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 6.

The worst 5 of 97 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:1146:VAL:O	1:A:1150:ALA:HB3	1.76	0.83
1:A:143:TRP:HB3	8:A:1507:CLR:H12	1.61	0.80
1:A:506:LEU:HB3	1:A:592:LEU:CD2	2.21	0.71
1:A:1045:PRO:HD3	8:A:1508:CLR:H241	1.74	0.69
1:A:506:LEU:HB3	1:A:592:LEU:HD23	1.74	0.69

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	953/1447 (66%)	934 (98%)	19 (2%)	0	100	100
2	B	2/272 (1%)	1 (50%)	0	1 (50%)	0	0
All	All	955/1719 (56%)	935 (98%)	19 (2%)	1 (0%)	54	69

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	2	THR

5.3.2 Protein sidechains ⓘ

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	824/1248 (66%)	816 (99%)	8 (1%)	76	86
2	B	3/219 (1%)	2 (67%)	1 (33%)	0	0
All	All	827/1467 (56%)	818 (99%)	9 (1%)	74	85

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1153	TRP
2	B	2	THR
1	A	508	ILE
1	A	509	VAL
1	A	903	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	907	GLN
1	A	1047	ASN

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

5.4   Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6   Ligand geometry ⓘ

9 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	05E	A	1504	6,4	18,18,19	0.69	0	22,25,27	0.87	1 (4%)
4	MAN	A	1503	5,9	11,11,12	0.54	0	15,15,17	1.23	3 (20%)
6	PA1	A	1505	5,7	11,11,12	0.74	0	12,15,17	1.55	2 (16%)
3	PLM	A	1501	-	17,17,17	0.60	0	17,17,17	0.51	0
4	MAN	A	1502	9	11,11,12	0.48	0	15,15,17	0.57	0
8	CLR	A	1507	-	31,31,31	0.22	0	48,48,48	0.36	0
9	80Y	B	101	2,4	18,18,19	0.66	0	24,25,27	0.90	1 (4%)
8	CLR	A	1508	-	31,31,31	0.28	0	48,48,48	0.50	0
7	LYI	A	1506	6	60,60,60	0.42	0	69,71,71	0.55	1 (1%)

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	05E	A	1504	6,4	-	7/12/29/32	0/1/1/1
4	MAN	A	1503	5,9	-	2/2/19/22	0/1/1/1
6	PA1	A	1505	5,7	-	2/2/19/22	0/1/1/1
3	PLM	A	1501	-	-	10/15/15/15	-
4	MAN	A	1502	9	-	0/2/19/22	0/1/1/1
8	CLR	A	1507	-	-	9/10/68/68	0/4/4/4
9	80Y	B	101	2,4	-	5/11/28/31	0/1/1/1
8	CLR	A	1508	-	-	7/10/68/68	0/4/4/4
7	LYI	A	1506	6	-	21/54/78/78	0/1/1/1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	1505	PA1	C4-C3-C2	-3.55	105.37	111.37
6	A	1505	PA1	C1-O5-C5	3.55	117.00	112.19
9	B	101	80Y	O53-C52-C85	-2.97	104.19	110.14
7	A	1506	LYI	O10-C9-C7	-2.82	103.82	110.35
4	A	1503	MAN	O2-C2-C3	-2.77	104.59	110.14

There are no chirality outliers.

5 of 63 torsion outliers are listed below:
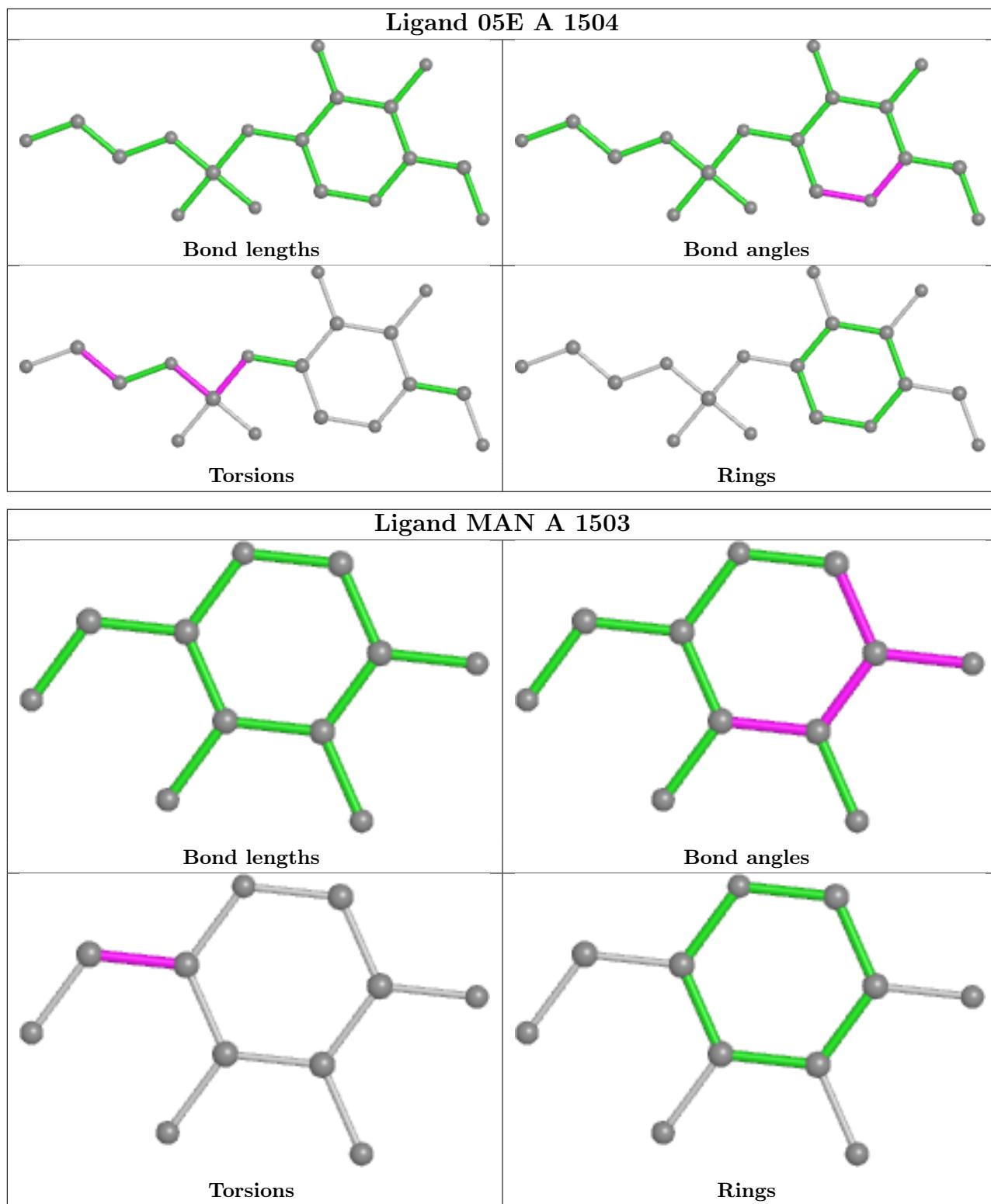
| Mol	Chain	Res	Type	Atoms
5	A	1504	05E	C81-O82-P83-O88
5	A	1504	05E	C85-O84-P83-O88
5	A	1504	05E	C85-O84-P83-O89
9	B	101	80Y	C78-O77-P76-O75
9	B	101	80Y	C78-O77-P76-O81
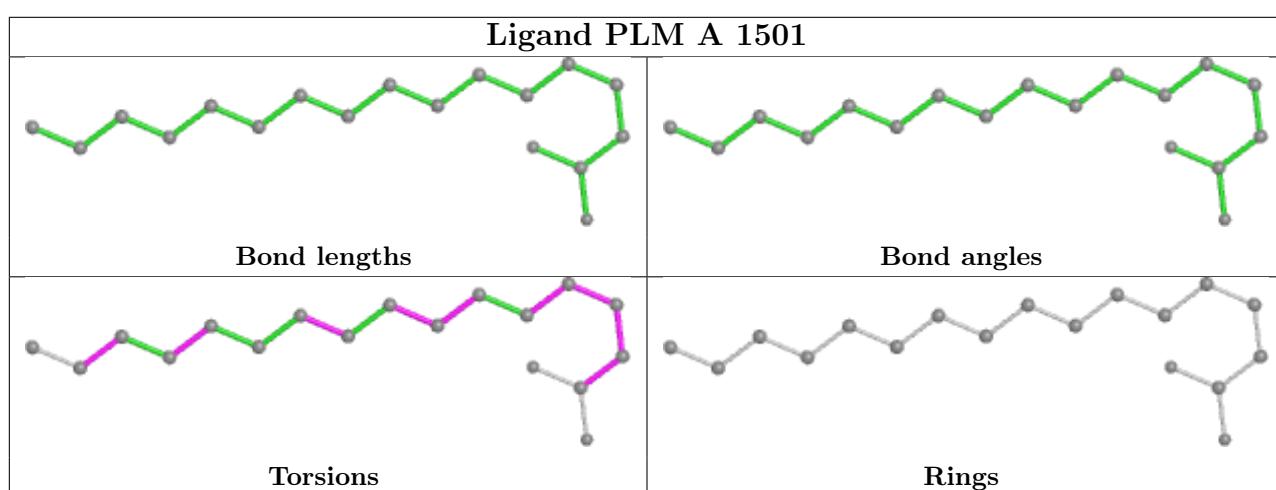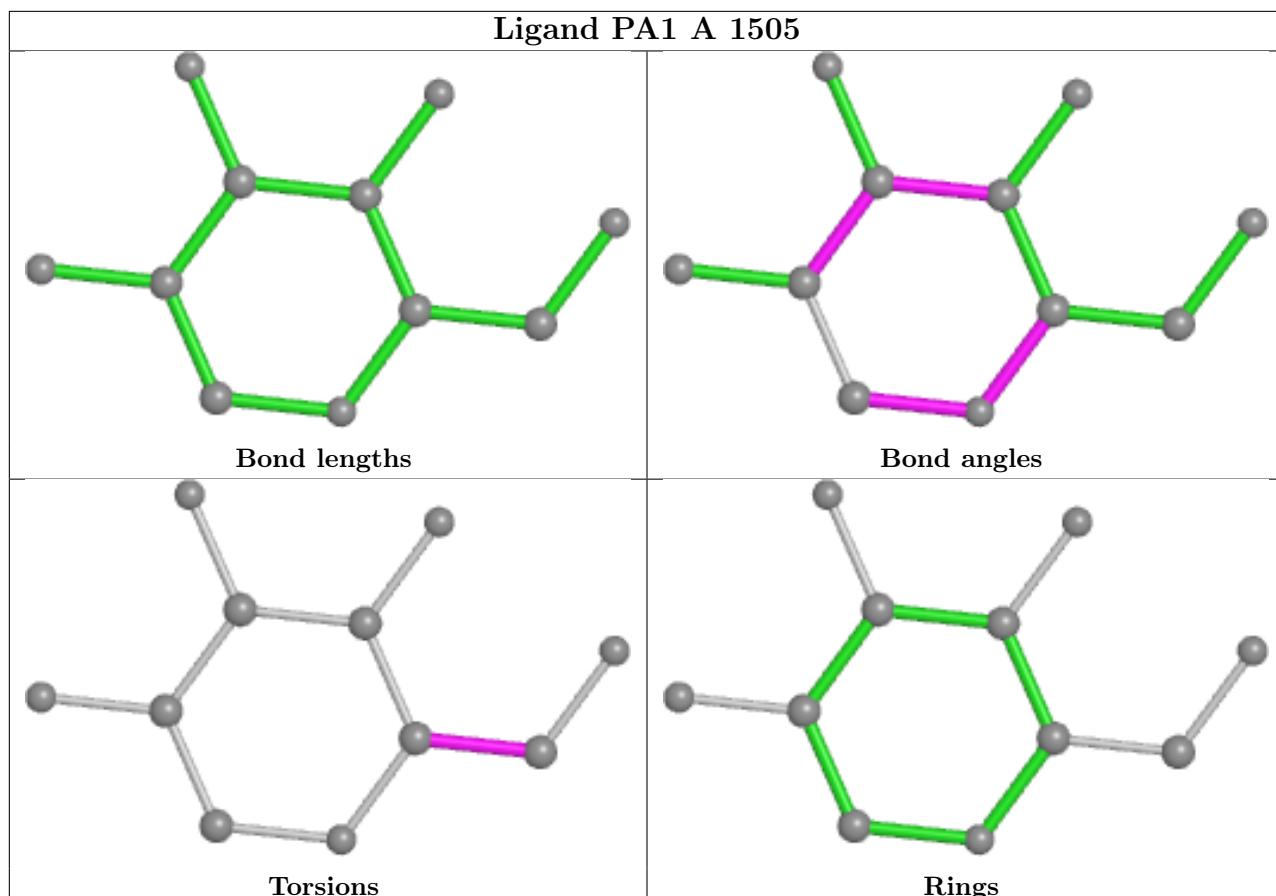
There are no ring outliers.
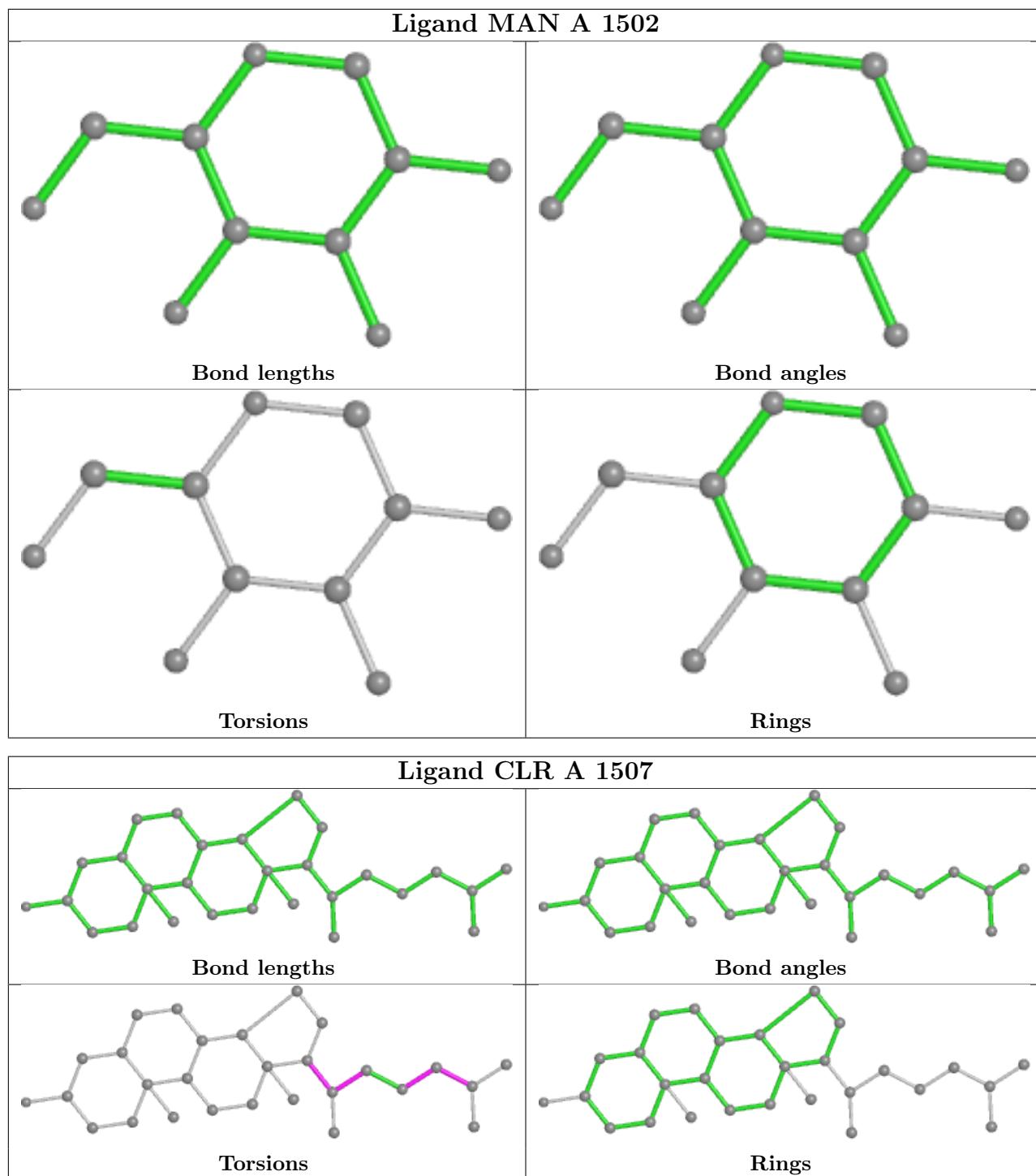
5 monomers are involved in 13 short contacts:
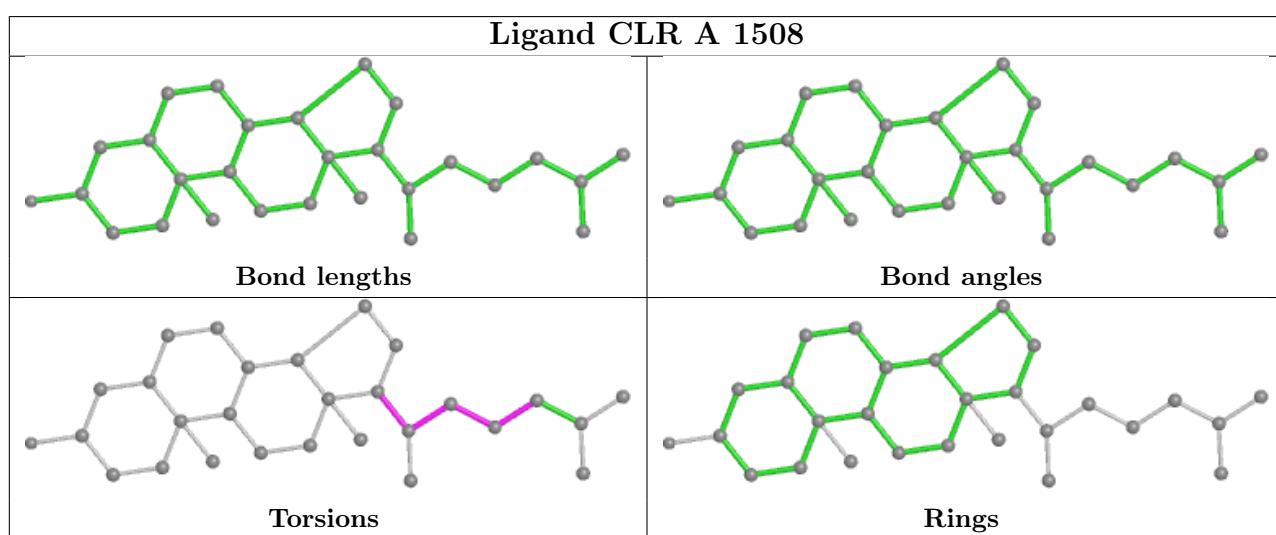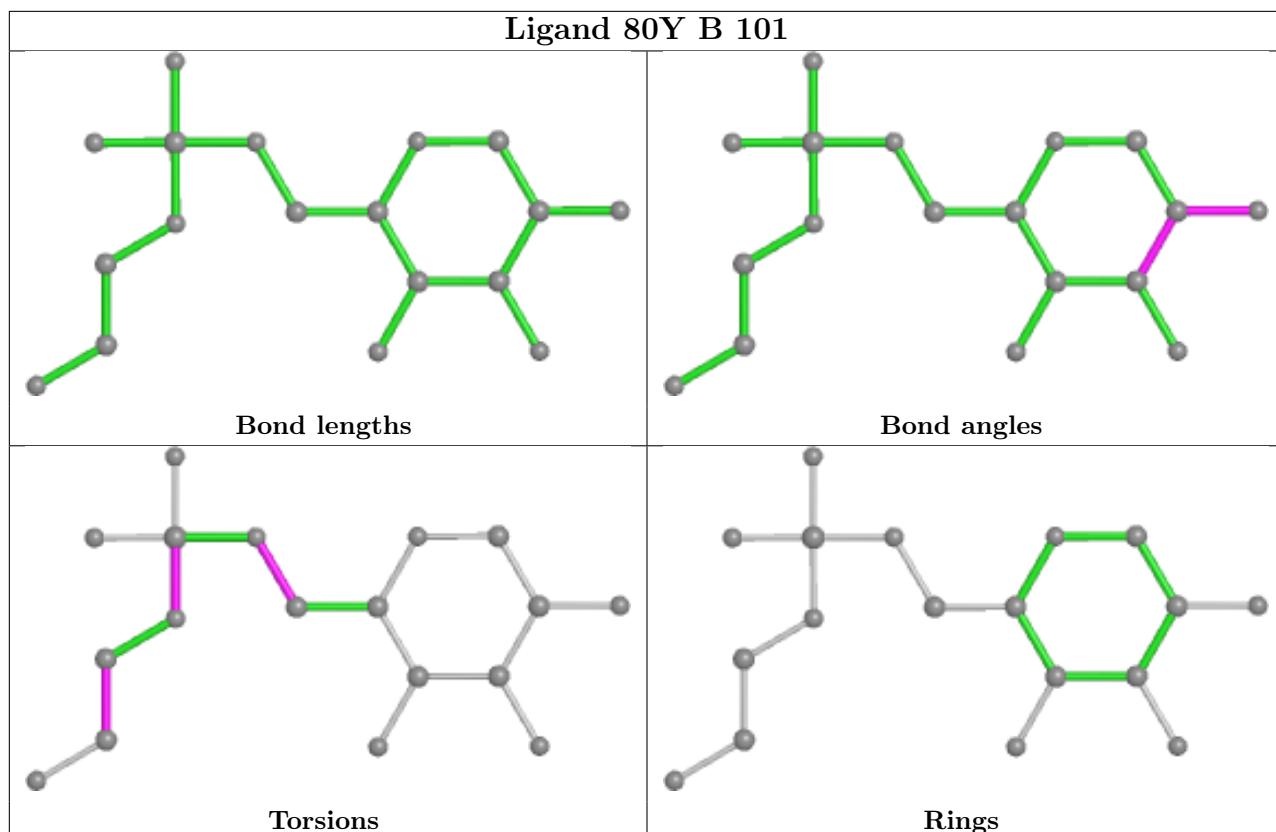
| Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1505	PA1	1	0
3	A	1501	PLM	1	0
4	A	1502	MAN	1	0
8	A	1507	CLR	4	0
8	A	1508	CLR	6	0
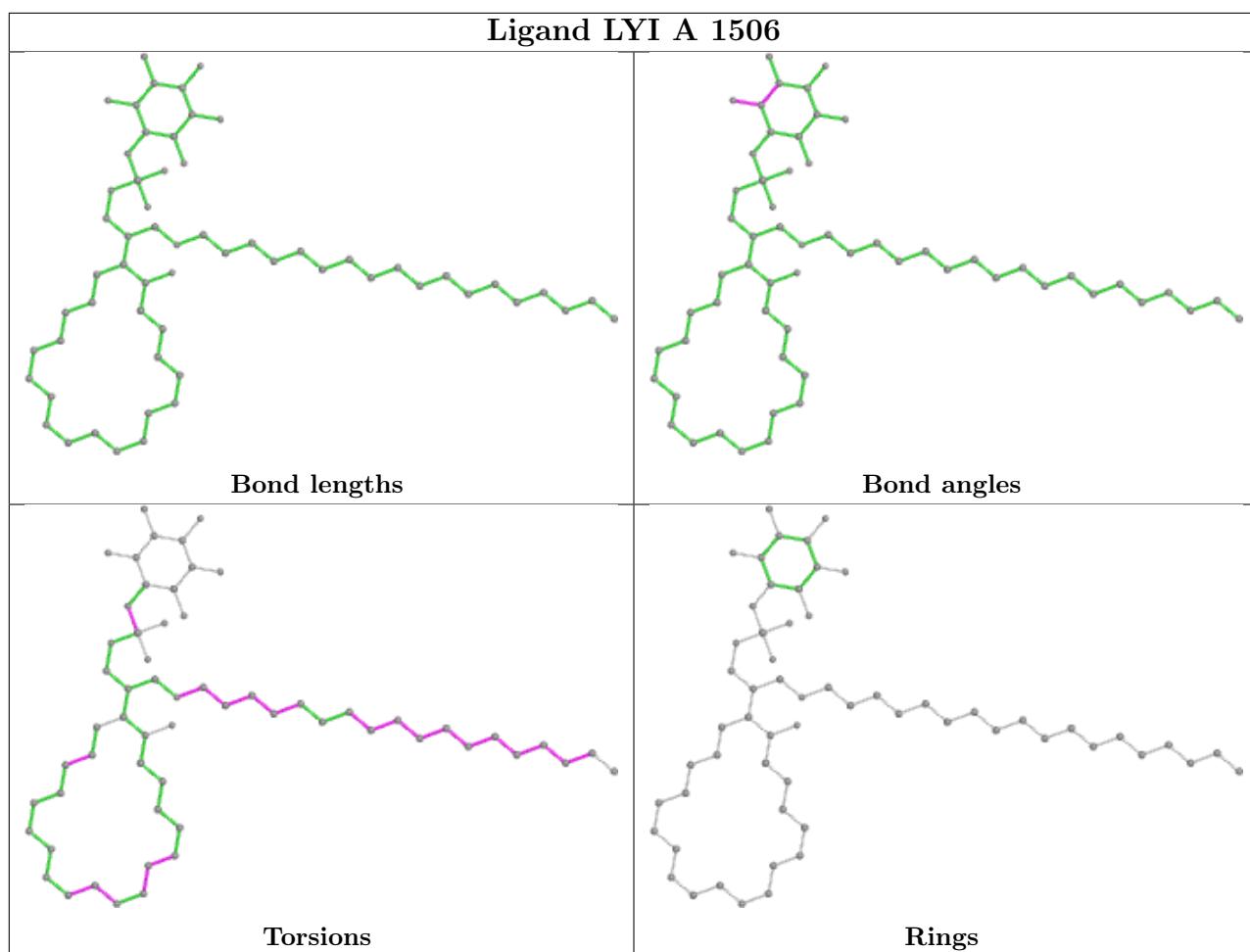
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 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.

Ligand 05E A 1504

Bond lengths

Bond angles

Torsions

Rings

Ligand MAN A 1503

Bond lengths

Bond angles

Torsions

Rings

Ligand PA1 A 1505

Bond lengths

Bond angles

Torsions

Rings

Ligand PLM A 1501

Bond lengths

Bond angles

Torsions

Rings

Ligand MAN A 1502

Bond lengths

Bond angles

Torsions

Rings

Ligand CLR A 1507

Bond lengths

Bond angles

Torsions

Rings

Ligand 80Y B 101

Bond lengths

Bond angles

Torsions

Rings

Ligand CLR A 1508

Bond lengths

Bond angles

Torsions

Rings

Ligand LYI A 1506

Bond lengths

Bond angles

Torsions

Rings

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.