WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB EM Validation Summary Report ⓘ

Feb 20, 2025 – 08:40 AM EST

PDB ID	:	8SJX
Title	:	Structure of lens aquaporin-0 array in sphingomyelin/cholesterol bilayer (2SM:1Chol)
Authors	:	Chiu, P.-L.; Walz, T.
Deposited on	:	2023-04-18
Resolution	:	2.50 Å(reported)
Based on initial model	:	2B6O

The following versions of software and data (see references ⓘ) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.4

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
ELECTRON CRYSTALLOGRAPHY

The reported resolution of this entry is 2.50 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.
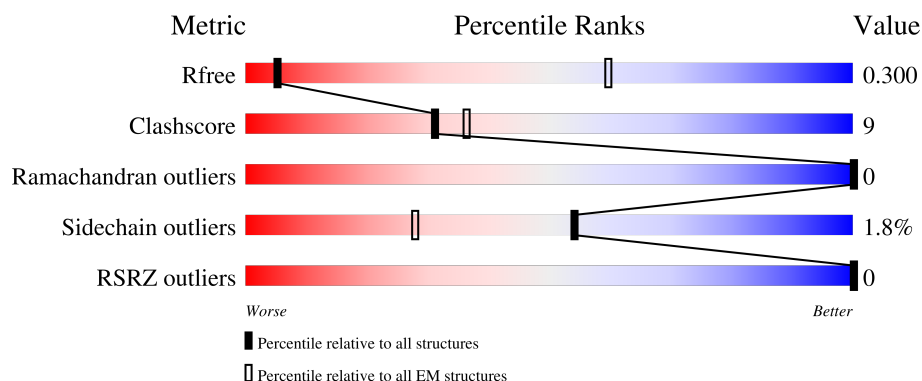
Metric	Whole archive (#Entries)	EM structures (#Entries)
R_{free}	164678	53
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RSRZ outliers	164674	54

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	263	67% 16% • 16%

2   Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 2010 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 Lens fiber major intrinsic protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	220	1663	1107	276	275	5	0	0

- Molecule 2 is [(E,2S,3R)-2-(hexadecanoylamino)-3-oxidanyl-octadec-4-enyl] 2-(trimethyl azaniumyl)ethyl phosphate (three-letter code: HWP) (formula: $C_{39}H_{79}N_2O_6P$) (labeled as "Ligand of Interest" by depositor).

HWP
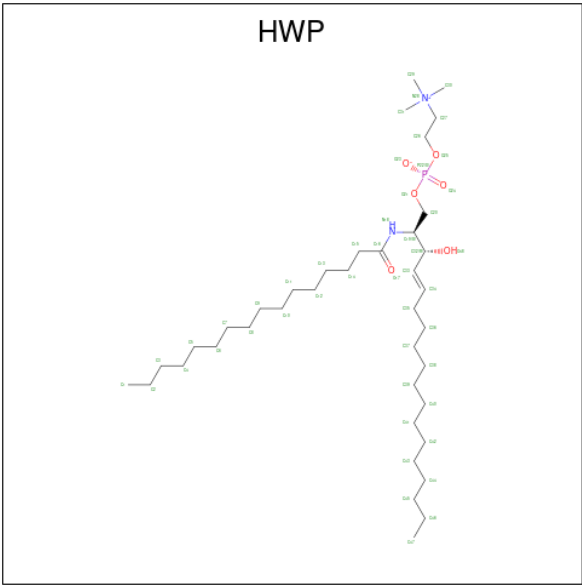
Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
2	A	1	43	34	2	6	1	0
2	A	1	46	37	2	6	1	0
2	A	1	48	39	2	6	1	0
2	A	1	45	36	2	6	1	0
2	A	1	43	34	2	6	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf
2	A	1	Total	C	N	O	P	0
			41	32	2	6	1	
2	A	1	Total	C	N	O	P	0
			42	33	2	6	1	

- Molecule 3 is CHOLESTEROL (three-letter code: CLR) (formula: $C_{27}H_{46}O$) (labeled as "Ligand of Interest" by depositor).
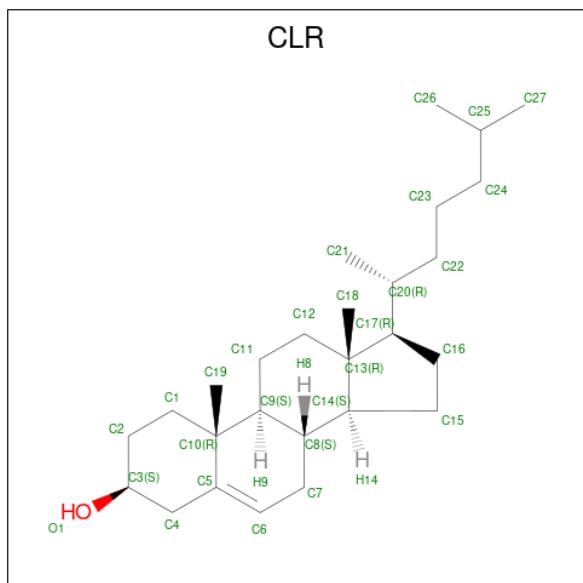
| Mol	Chain	Residues	Atoms			AltConf
3	A	1	Total	C	O	0
			28	27	1	

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		AltConf
4	A	11	Total	O	0
			11	11	

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: Lens fiber major intrinsic protein
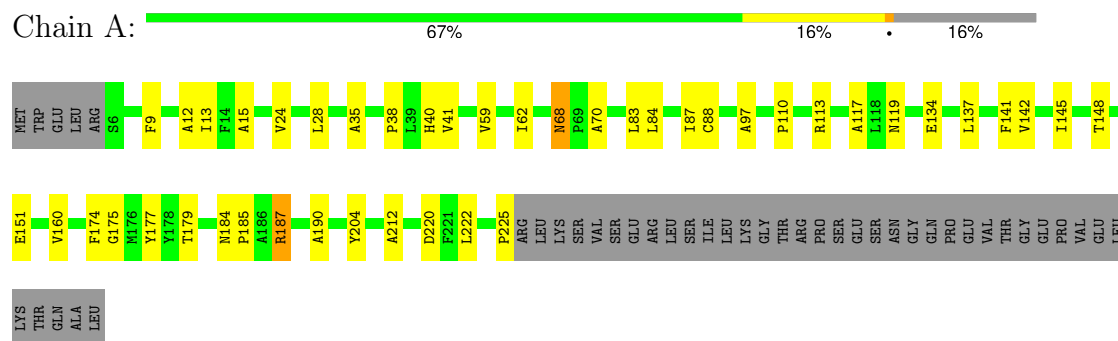
4   Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 4 2 2	Depositor
Cell constants a, b, c, α, β, γ	65.50Å 65.50Å 200.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	2.50 – 2.50 2.50 – 2.50	Depositor EDS
% Data completeness (in resolution range)	86.7 (2.50-2.50) 0.0 (2.50-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.35 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.260 , 0.286 0.273 , 0.300	Depositor DCC
R_{free} test set	6 reflections (14.29%)	wwPDB-VP
Wilson B-factor (Å^2)	47.5	Xtriage
Anisotropy	0.306	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 100.0	EDS
L-test for twinning[2]	$< |L| > = 0.44, < L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.00	EDS
Total number of atoms	2010	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.27% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: HWP, 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	0.29	0/1713	0.51	2/2344 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	222	LEU	CA-CB-CG	6.44	130.12	115.30
1	A	83	LEU	CB-CG-CD2	-5.73	101.26	111.00

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	1663	0	1679	28	0
2	A	308	0	0	2	0
3	A	28	0	46	3	0
4	A	11	0	0	0	0
All	All	2010	0	1725	32	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 9.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

The worst 5 of 32 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:87:ILE:HG23	3:A:308:CLR:H25	1.70	0.74
1:A:137:LEU:HD11	1:A:175:GLY:HA3	1.78	0.64
1:A:84:LEU:HA	1:A:87:ILE:HD12	1.83	0.60
2:A:302:HWP:O21	2:A:302:HWP:O48	2.18	0.60
1:A:68:ASN:HD21	1:A:184:ASN:HA	1.76	0.50

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	218/263 (83%)	208 (95%)	10 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	166/205 (81%)	163 (98%)	3 (2%)	54	78

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	68	ASN
1	A	151	GLU
1	A	187	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	68	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 oligosaccharides in this entry.

5.6　Ligand geometry ⓘ

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	$\#|Z| > 2$
2	HWP	A	305	-	41,42,47	1.47	7 (17%)	47,50,55	1.58	8 (17%)
2	HWP	A	307	-	40,41,47	1.48	8 (20%)	46,49,55	1.07	2 (4%)
2	HWP	A	302	-	44,45,47	1.43	5 (11%)	50,53,55	1.07	2 (4%)
3	CLR	A	308	-	31,31,31	0.41	0	48,48,48	0.78	0
2	HWP	A	301	-	41,42,47	1.40	5 (12%)	47,50,55	1.39	6 (12%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
2	HWP	A	304	-	43,44,47	1.20	6 (13%)	49,52,55	1.47	2 (4%)
2	HWP	A	306	-	39,40,47	1.46	6 (15%)	45,48,55	1.24	5 (11%)
2	HWP	A	303	-	46,47,47	1.28	5 (10%)	52,55,55	1.13	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
2	HWP	A	305	-	-	22/47/47/52	-
2	HWP	A	307	-	-	14/46/46/52	-
2	HWP	A	302	-	-	19/50/50/52	-
3	CLR	A	308	-	-	9/10/68/68	0/4/4/4
2	HWP	A	301	-	-	26/47/47/52	-
2	HWP	A	304	-	-	11/49/49/52	-
2	HWP	A	306	-	-	14/45/45/52	-
2	HWP	A	303	-	-	18/52/52/52	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	302	HWP	C20-C19	4.95	1.60	1.51
2	A	305	HWP	P22-O25	4.56	1.77	1.59
2	A	306	HWP	P22-O25	4.56	1.77	1.59
2	A	303	HWP	P22-O25	4.51	1.77	1.59
2	A	301	HWP	C20-C19	4.46	1.59	1.51

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	304	HWP	C32-C33-C34	-7.22	109.73	124.69
2	A	305	HWP	C11-C10-C9	-5.15	88.35	114.37
2	A	305	HWP	C13-C14-C15	-4.43	96.87	113.13
2	A	301	HWP	C32-C33-C34	-4.30	115.78	124.69
2	A	306	HWP	C35-C34-C33	-3.42	110.23	125.47

There are no chirality outliers.

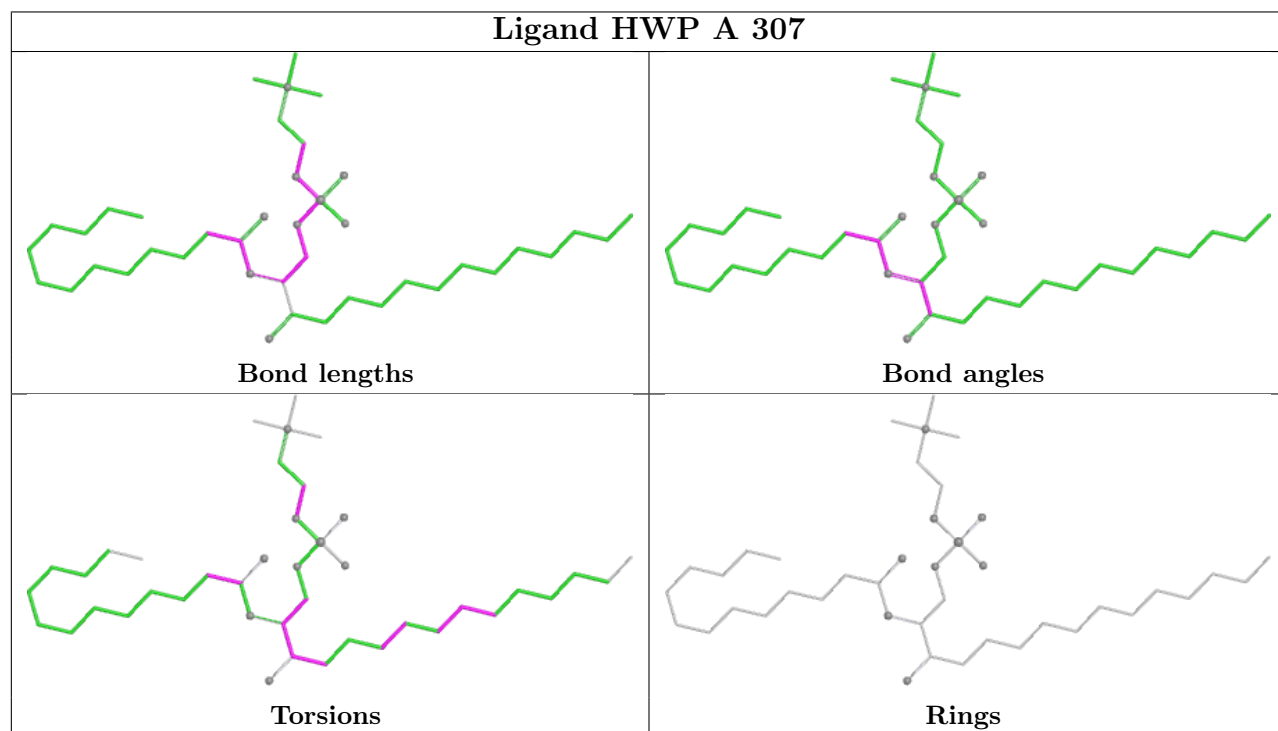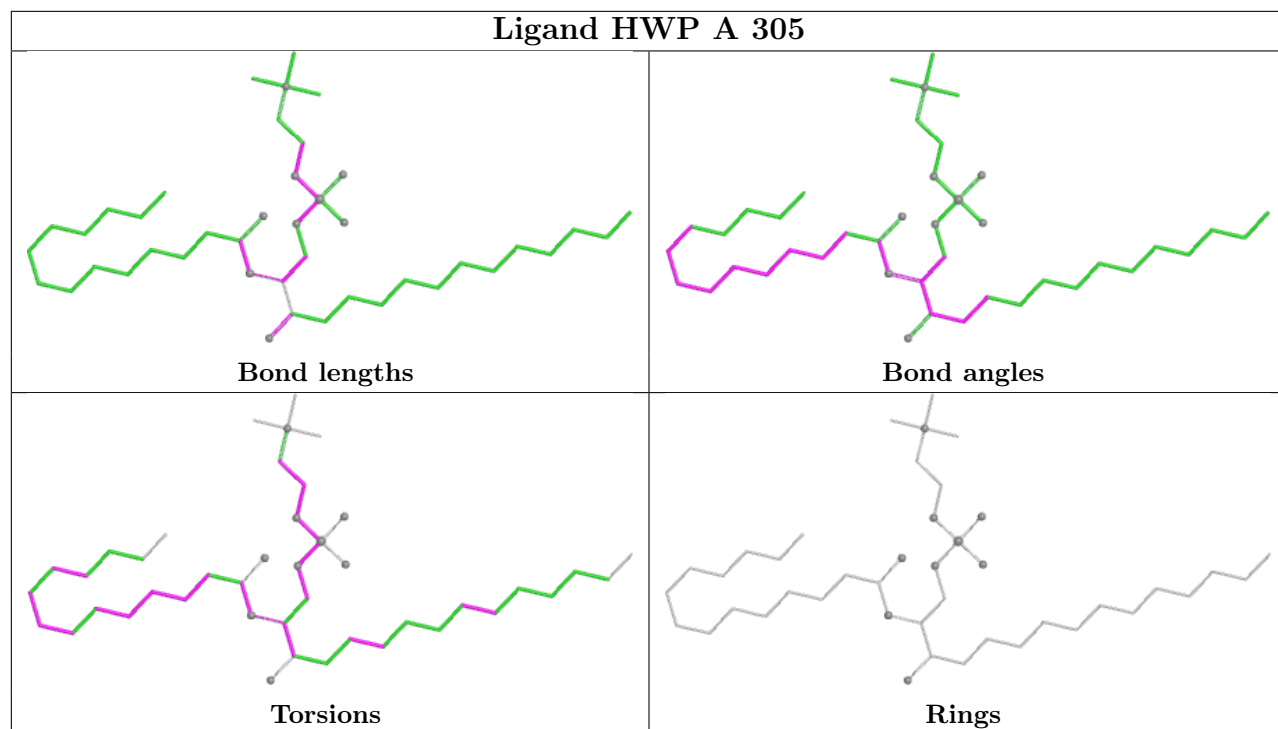5 of 133 torsion outliers are listed below:

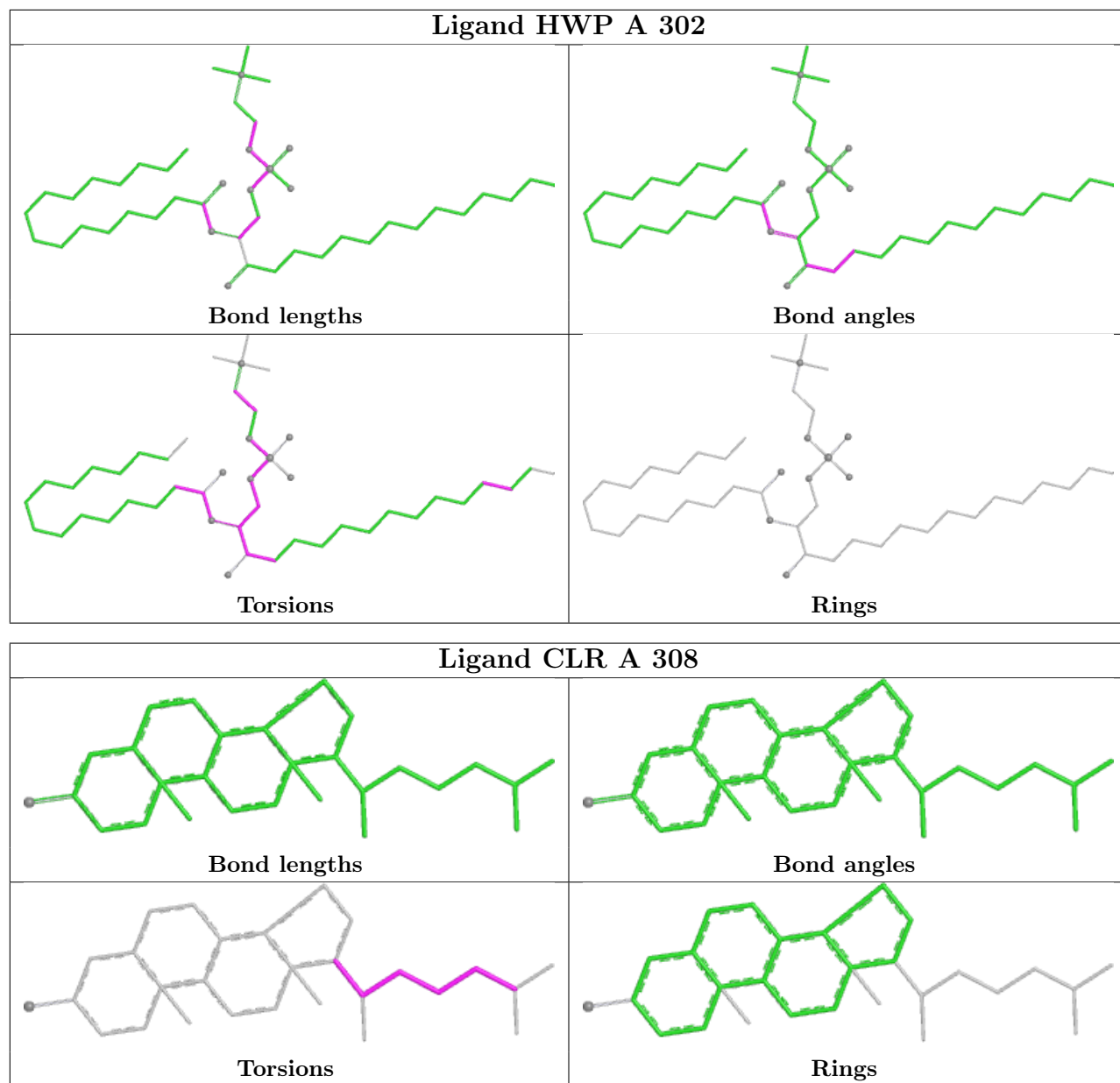| Mol	Chain	Res	Type	Atoms
2	A	301	HWP	C19-C20-O21-P22
2	A	301	HWP	O25-C26-C27-N28
2	A	301	HWP	C20-O21-P22-O23
2	A	301	HWP	C20-O21-P22-O25
2	A	301	HWP	C15-C16-N18-C19

There are no ring outliers.
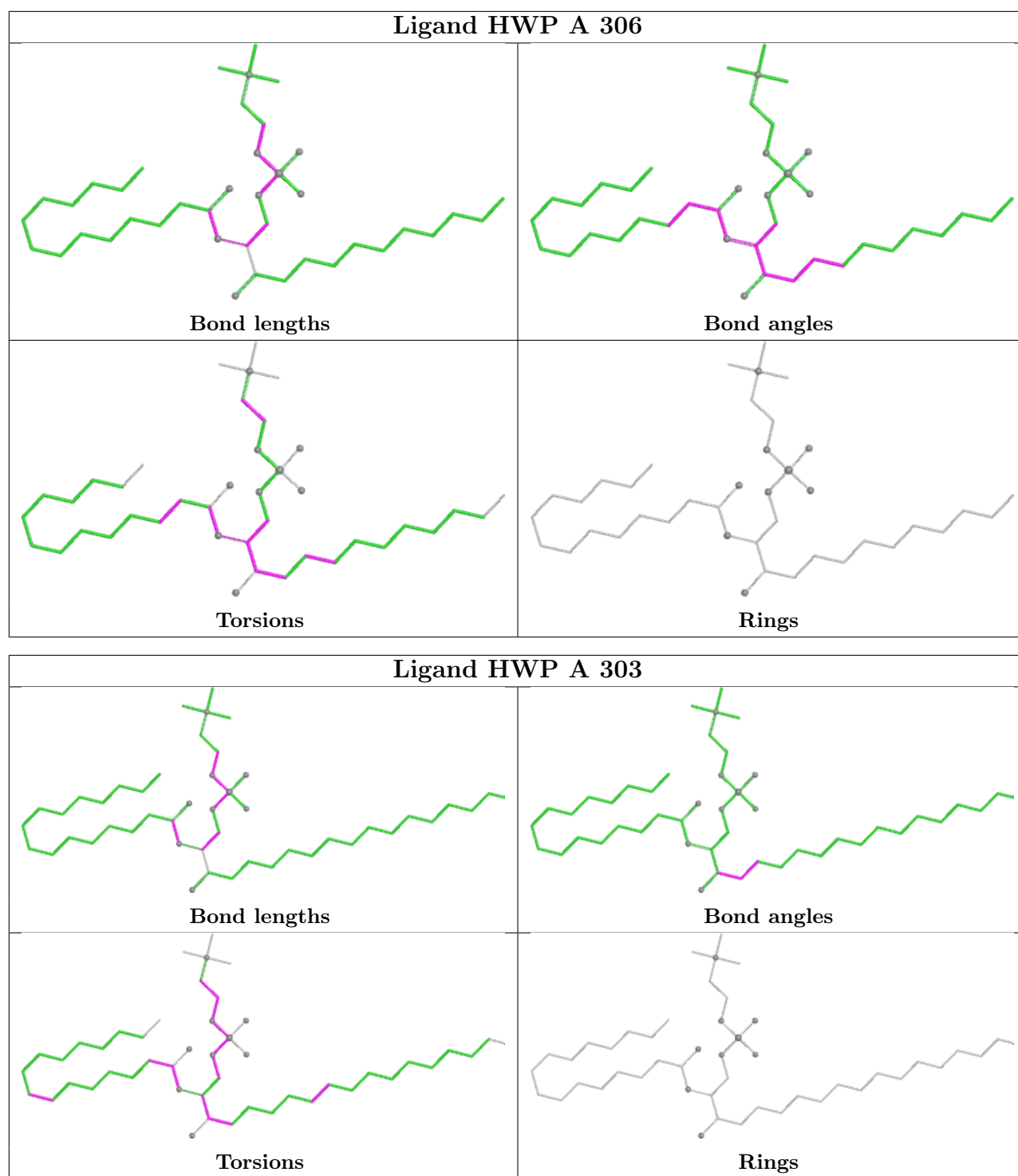
3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	302	HWP	1	0
3	A	308	CLR	3	0
2	A	301	HWP	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 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 HWP A 305

Bond lengths

Bond angles

Torsions

Rings

Ligand HWP A 307

Bond lengths

Bond angles

Torsions

Rings

Ligand HWP A 302

Bond lengths

Bond angles

Torsions

Rings

Ligand CLR A 308

Bond lengths

Bond angles

Torsions

Rings

Ligand HWP A 301

Bond lengths

Bond angles

Torsions

Rings

Ligand HWP A 304

Bond lengths

Bond angles

Torsions

Rings

Ligand HWP A 306

Bond lengths

Bond angles

Torsions

Rings

Ligand HWP A 303

Bond lengths

Bond angles

Torsions

Rings

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.