wwPDB EM Validation Summary Report

Feb 20, 2025 – 07:49 AM EST

PDB ID	:	2ZZ9
Title	:	Structure of aquaporin-4 S180D mutant at 2.8 A resolution by electron crystallography
Authors	:	Tani, K.; Mitsuma, T.; Hiroaki, Y.; Kamegawa, A.; Nishikawa, K.; Tanimura, Y.; Fujiyoshi, Y.
Deposited on	:	2009-02-06
Resolution	:	2.80 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB/EMDB 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 (i) 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 (1)) were used in the production of this report:

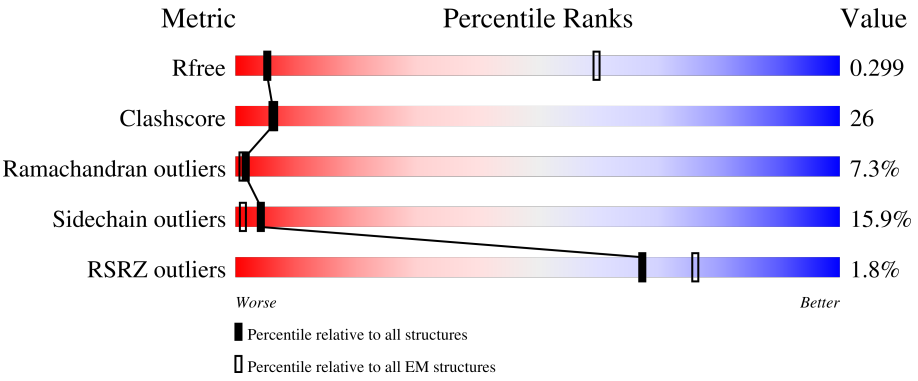
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.80 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	301	% 38% 28% 7% • 26%

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
2	PEE	A	401	X	-	-	-
2	PEE	A	402	X	-	-	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	PEE	A	403	X	-	-	-
2	PEE	A	404	X	-	-	-
2	PEE	A	405	X	-	-	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1867 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.

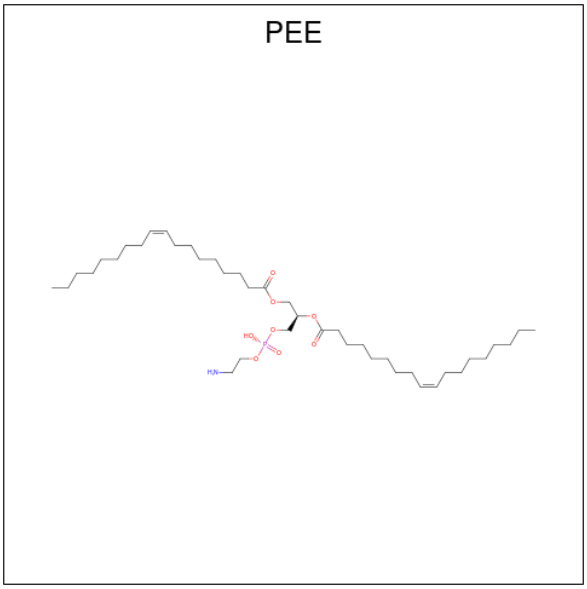
- Molecule 1 is a protein called Aquaporin-4.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	223	Total	C	N	O	S	0	0
			1642	1079	265	286	12		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	180	ASP	SER	engineered mutation	UNP P47863

- Molecule 2 is 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (three-letter code: PEE) (formula: $C_{41}H_{78}NO_8P$).



Mol	Chain	Residues	Atoms					AltConf
2	A	1	Total	C	N	O	P	0
			44	34	1	8	1	
2	A	1	Total	C	N	O	P	0
			49	39	1	8	1	
2	A	1	Total	C	N	O	P	0
			49	39	1	8	1	

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf
2	A	1	Total	C	N	O	P	0
			34	24	1	8	1	
2	A	1	Total	C	N	O	P	0
			35	25	1	8	1	

- Molecule 3 is water.

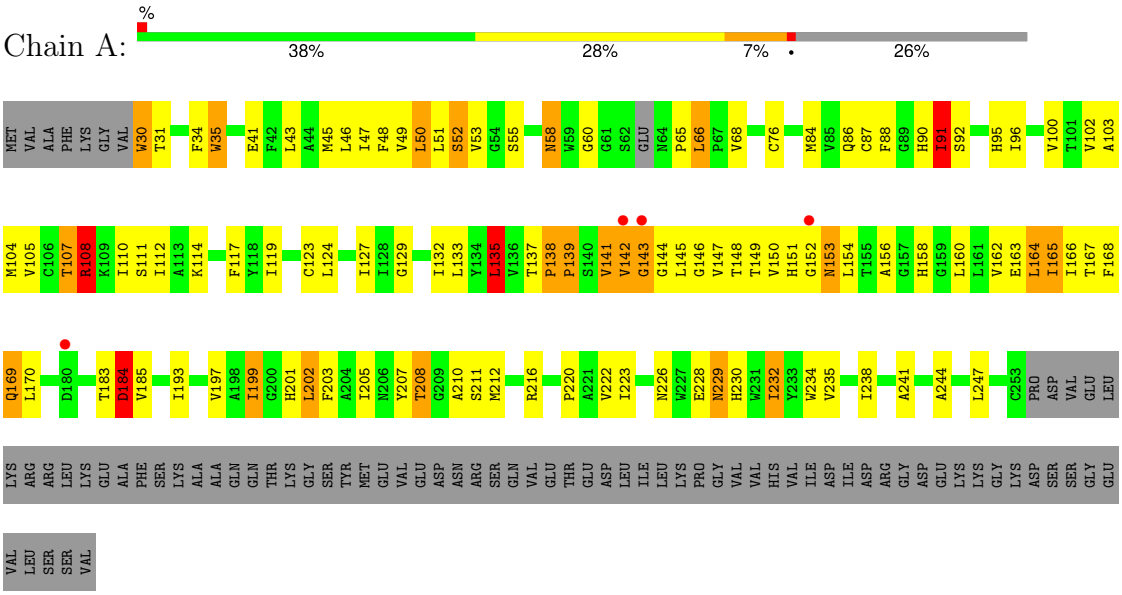
Mol	Chain	Residues	Atoms		AltConf
3	A	14	Total	O	0
			14	14	

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: Aquaporin-4

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 4 21 2	Depositor
Cell constants a, b, c, α, β, γ	69.00Å 69.00Å 160.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 2.80 10.00 – 2.80	Depositor EDS
% Data completeness (in resolution range)	81.9 (10.00-2.80) 84.3 (10.00-2.80)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	0.20	Depositor
$< I/\sigma(I) >$ [1]	1.96 (at 2.78Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.231 , 0.288 0.245 , 0.299	Depositor DCC
R_{free} test set	481 reflections (5.61%)	wwPDB-VP
Wilson B-factor (Å^2)	49.3	Xtriage
Anisotropy	1.033	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.47 , 147.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	1867	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.47% 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: PEE

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	0.56	0/1683	0.70	2/2296 (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	164	LEU	CA-CB-CG	7.25	131.98	115.30
1	A	135	LEU	CA-CB-CG	5.56	128.10	115.30

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	1642	0	1646	88	0
2	A	211	0	292	15	0
3	A	14	0	0	0	0
All	All	1867	0	1938	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 26.

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:58:ASN:HD21	1:A:66:LEU:H	1.02	0.96
1:A:151:HIS:CE1	1:A:153:ASN:HA	2.00	0.94
1:A:49:VAL:O	1:A:53:VAL:HB	1.70	0.91
1:A:151:HIS:CG	1:A:152:GLY:H	1.91	0.88
1:A:138:PRO:CB	1:A:139:PRO:CD	2.55	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	219/301 (73%)	176 (80%)	27 (12%)	16 (7%)	1 2

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	65	PRO
1	A	66	LEU
1	A	142	VAL
1	A	184	ASP
1	A	138	PRO

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	170/243 (70%)	143 (84%)	27 (16%)	2 7

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	154	LEU
1	A	165	ILE
1	A	232	ILE
1	A	164	LEU
1	A	169	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	206	ASN
1	A	226	ASN
1	A	229	ASN
1	A	58	ASN
1	A	32	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 oligosaccharides 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	PEE	A	405	-	34,34,50	5.34	6 (17%)	37,39,55	4.21	8 (21%)
2	PEE	A	402	-	48,48,50	4.72	8 (16%)	51,53,55	3.58	5 (9%)
2	PEE	A	404	-	33,33,50	5.59	6 (18%)	36,38,55	4.12	8 (22%)
2	PEE	A	401	-	43,43,50	4.96	8 (18%)	46,48,55	4.28	6 (13%)
2	PEE	A	403	-	48,48,50	4.59	8 (16%)	51,53,55	3.48	5 (9%)

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	PEE	A	405	-	1/1/4/8	22/38/38/54	-
2	PEE	A	402	-	1/1/4/8	26/52/52/54	-
2	PEE	A	404	-	1/1/4/8	20/36/36/54	-
2	PEE	A	401	-	1/1/4/8	21/47/47/54	-
2	PEE	A	403	-	1/1/4/8	33/52/52/54	-

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	402	PEE	C11-C10	30.44	2.39	1.50
2	A	404	PEE	C11-C10	30.44	2.39	1.50
2	A	401	PEE	C11-C10	30.20	2.38	1.50
2	A	403	PEE	C11-C10	29.51	2.36	1.50
2	A	405	PEE	C11-C10	29.28	2.35	1.50

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	402	PEE	O4-C10-C11	-23.12	33.37	123.78
2	A	404	PEE	O4-C10-C11	-23.08	33.52	123.78
2	A	401	PEE	O4-C10-C11	-23.03	33.72	123.78
2	A	403	PEE	O4-C10-C11	-22.90	34.22	123.78
2	A	405	PEE	O4-C10-C11	-22.80	34.62	123.78

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	401	PEE	C2
2	A	402	PEE	C2
2	A	403	PEE	C2
2	A	404	PEE	C2
2	A	405	PEE	C2

5 of 122 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	401	PEE	O4-C10-O2-C2
2	A	401	PEE	C4-O4P-P-O3P
2	A	401	PEE	C4-O4P-P-O2P
2	A	401	PEE	O4P-C4-C5-N
2	A	401	PEE	O5-C30-O3-C3

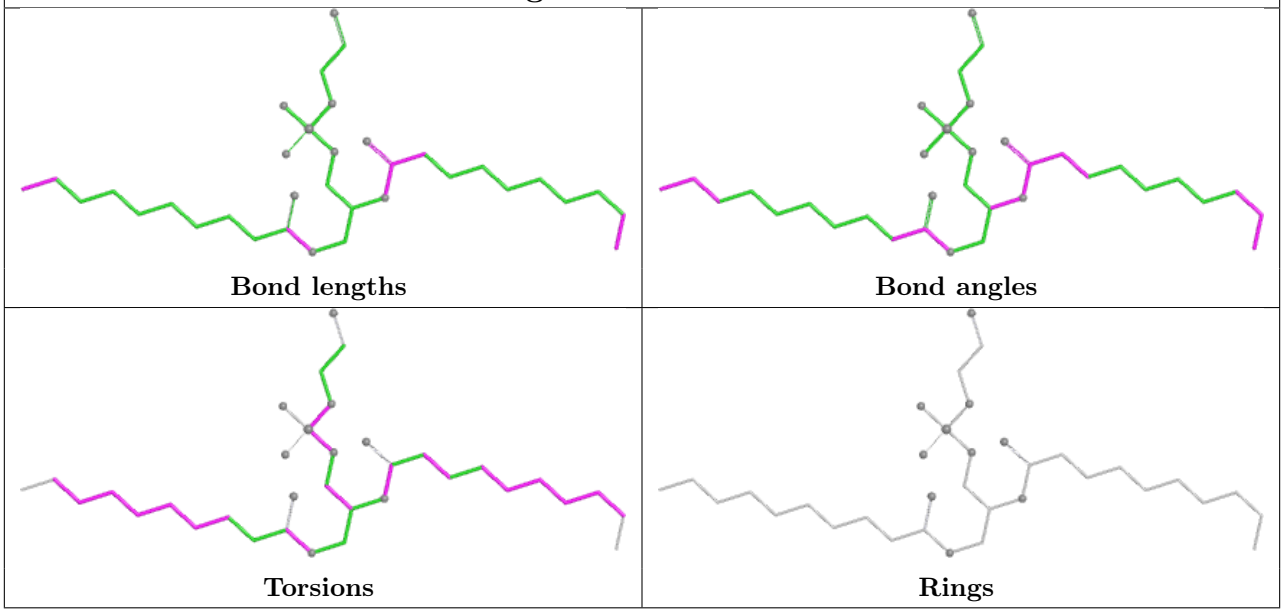
There are no ring outliers.

5 monomers are involved in 15 short contacts:

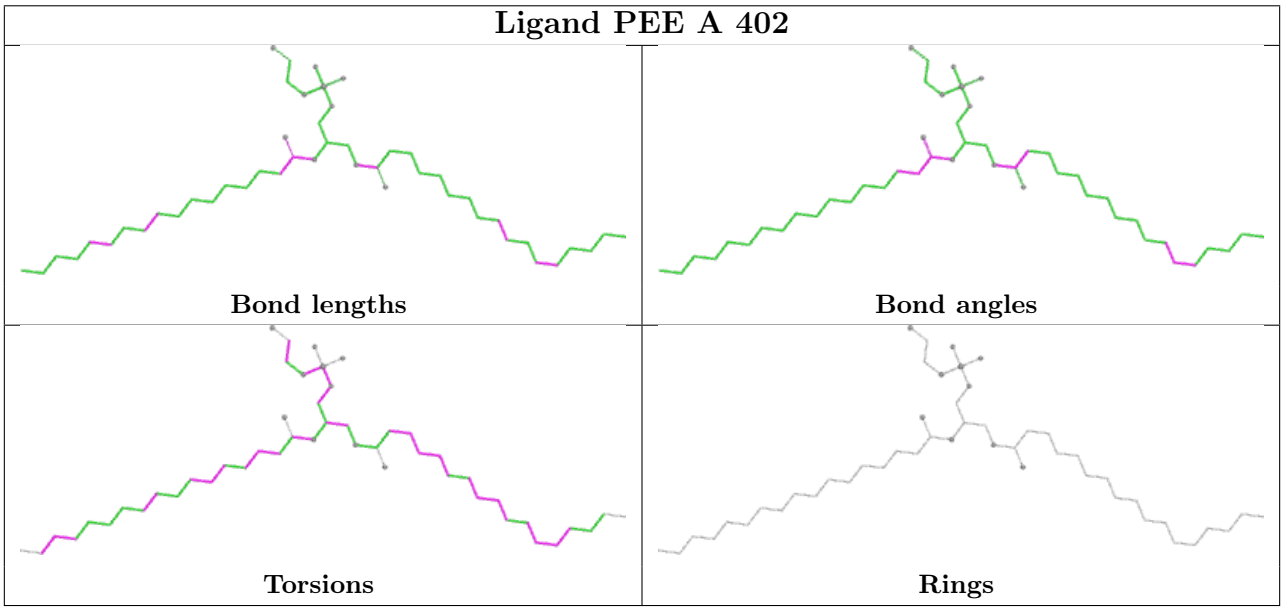
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	405	PEE	8	0
2	A	402	PEE	1	0
2	A	404	PEE	1	0
2	A	401	PEE	2	0
2	A	403	PEE	6	0

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.

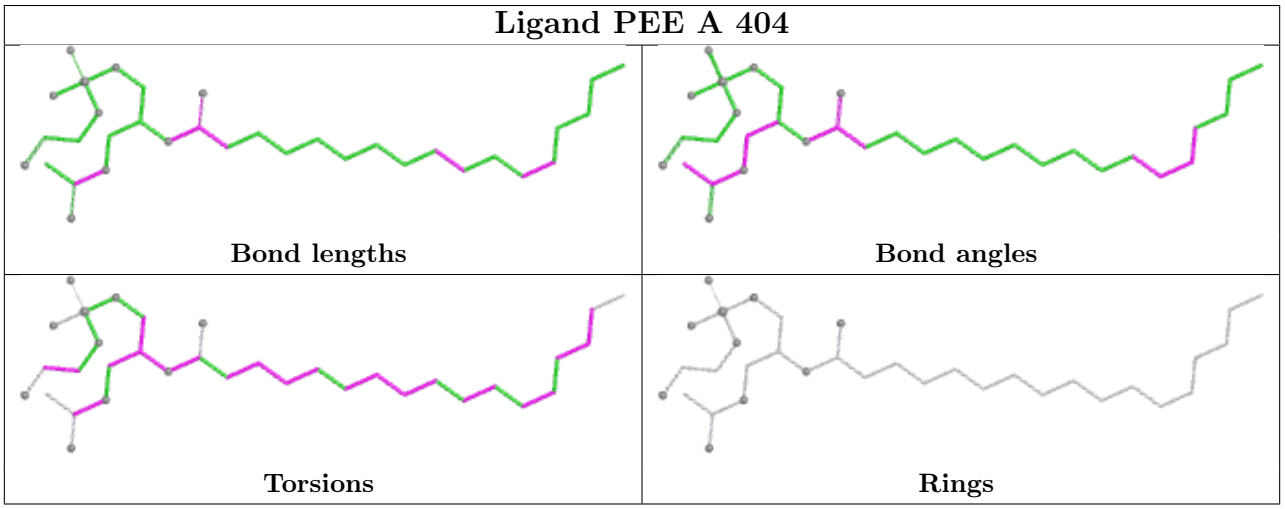
Ligand PEE A 405



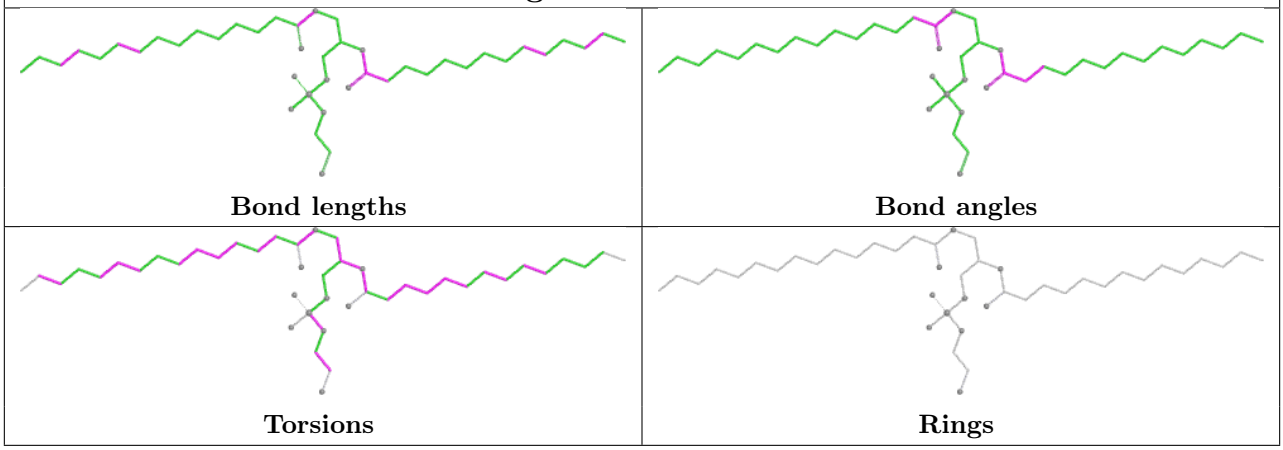
Ligand PEE A 402



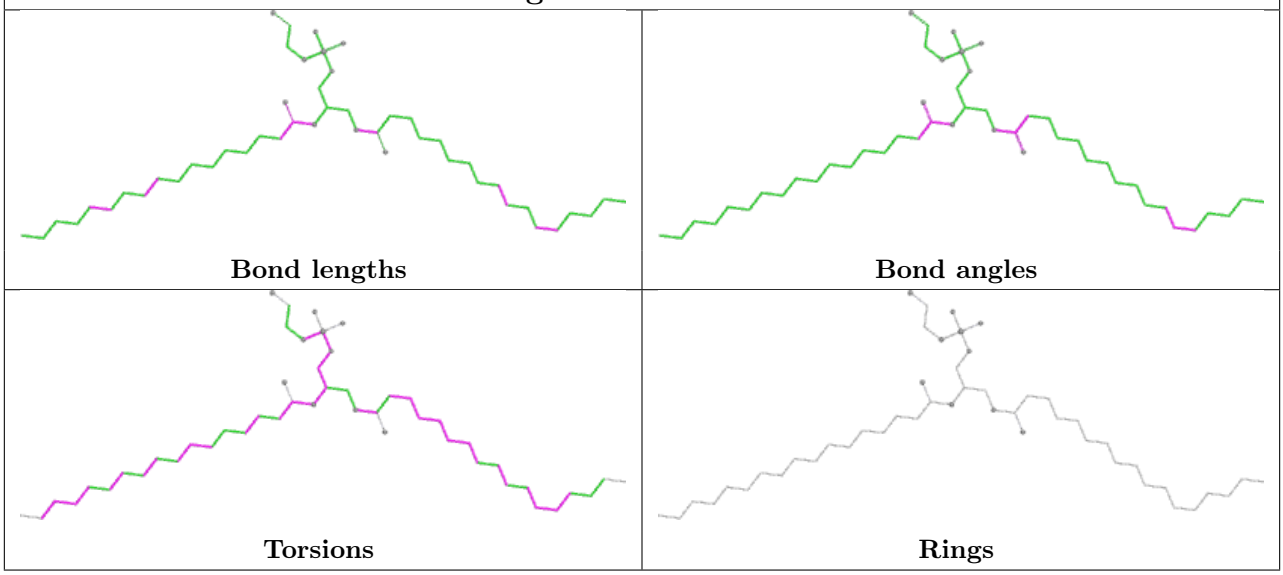
Ligand PEE A 404



Ligand PEE A 401



Ligand PEE A 403



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.