Full wwPDB X-ray Structure Validation Report (i)

Sep 16, 2021 – 02:02 pm BST

PDB ID	:	7O09
Title	:	Crystal structure of the human METTL3-METTL14 complex bound to Compound 7 (ADO_AC_074)
Authors	:	Bedi, R.K.; Dolbois, A.; Caflisch, A.
Deposited on	:	2021-03-25
Resolution	:	1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

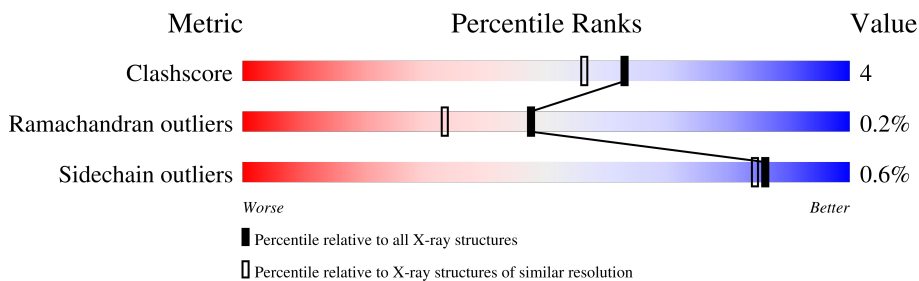|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	Percentile Ranks	Value
Clashscore		4
Ramachandran outliers		0.2%
Sidechain outliers		0.6%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	246	73% 7% 20%
2	B	290	75% 8% 17%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3855 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 N6-adenosine-methyltransferase catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	198	Total 1563	C 1006	N 266	O 282	S 9	0	2	0

There are 19 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
A	335	MET	-	initiating methionine	UNP Q86U44
A	336	GLY	-	expression tag	UNP Q86U44
A	337	HIS	-	expression tag	UNP Q86U44
A	338	HIS	-	expression tag	UNP Q86U44
A	339	HIS	-	expression tag	UNP Q86U44
A	340	HIS	-	expression tag	UNP Q86U44
A	341	HIS	-	expression tag	UNP Q86U44
A	342	HIS	-	expression tag	UNP Q86U44
A	343	SER	-	expression tag	UNP Q86U44
A	344	SER	-	expression tag	UNP Q86U44
A	345	GLY	-	expression tag	UNP Q86U44
A	346	ARG	-	expression tag	UNP Q86U44
A	347	GLU	-	expression tag	UNP Q86U44
A	348	ASN	-	expression tag	UNP Q86U44
A	349	LEU	-	expression tag	UNP Q86U44
A	350	TYR	-	expression tag	UNP Q86U44
A	351	PHE	-	expression tag	UNP Q86U44
A	352	GLN	-	expression tag	UNP Q86U44
A	353	GLY	-	expression tag	UNP Q86U44

- Molecule 2 is a protein called N6-adenosine-methyltransferase non-catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	241	Total 1899	C 1223	N 320	O 342	S 14	0	4	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	106	MET	-	initiating methionine	UNP Q9HCE5

- Molecule 3 is 6-[4-[6-[(4,4-dimethylpiperidin-1-yl)methyl]pyridin-3-yl]-1-oxa-4,9-diazaspiro [5.5]undecan-9-yl]-N-(phenylmethyl)pyrimidin-4-amine (three-letter code: UXK) (formula: $C_{32}H_{43}N_7O$) (labeled as "Ligand of Interest" by depositor).
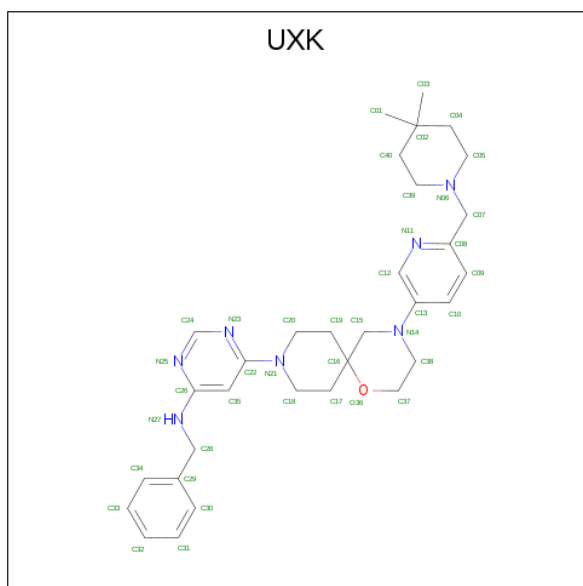
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	40	32	7	1	0	0

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).

ACT

OXT

O⁻

C

CH3

O

O

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total 4	C 2	O 2	0	0

- Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	137	Total 137	O 137	0	0
5	B	212	Total 212	O 212	0	0

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.
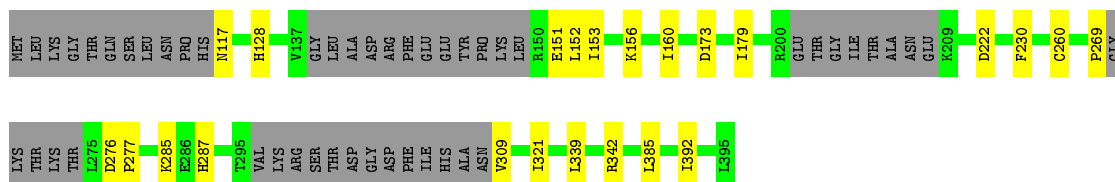
Note EDS was not executed.

• Molecule 1: N6-adenosine-methyltransferase catalytic subunit

Chain A:
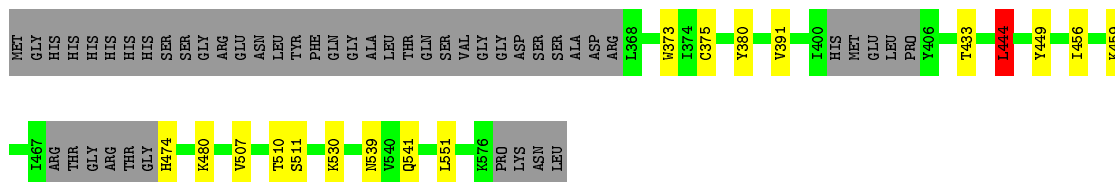
• Molecule 2: N6-adenosine-methyltransferase non-catalytic subunit

Chain B:

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	63.96Å 63.96Å 226.01Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.63 – 1.80	Depositor
% Data completeness (in resolution range)	99.9 (44.63-1.80)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PHENIX 1.19.1_4122	Depositor
R, R_{free}	0.193 , 0.225	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3855	wwPDB-VP
Average B, all atoms (Å²)	34.0	wwPDB-VP

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: UXK, ACT

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	0.40	0/1608	0.69	2/2192 (0.1%)
2	B	0.41	0/1955	0.64	0/2655
All	All	0.40	0/3563	0.66	2/4847 (0.0%)

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	444[A]	LEU	CA-CB-CG	9.08	136.18	115.30
1	A	444[B]	LEU	CA-CB-CG	9.08	136.18	115.30

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	1563	0	1504	14	0
2	B	1899	0	1786	16	0
3	A	40	0	0	0	0
4	B	4	0	3	0	0
5	A	137	0	0	4	0
5	B	212	0	0	6	0
All	All	3855	0	3293	29	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 4.

All (29) 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:375:CYS:SG	5:A:810:HOH:O	2.29	0.89
2:B:276:ASP:OD2	5:B:501:HOH:O	1.99	0.81
2:B:342[B]:ARG:NH2	5:B:503:HOH:O	2.15	0.78
2:B:309:VAL:N	5:B:505:HOH:O	2.23	0.70
2:B:222:ASP:OD2	5:B:502:HOH:O	2.14	0.65
1:A:444[A]:LEU:HD13	1:A:449:TYR:HB2	1.79	0.65
1:A:474:HIS:N	5:A:703:HOH:O	2.29	0.64
1:A:391:VAL:HB	1:A:530:LYS:HG2	1.82	0.61
2:B:276:ASP:OD1	2:B:277:PRO:HD2	2.00	0.60
1:A:373:TRP:HB2	1:A:551:LEU:HD13	1.87	0.56
1:A:456[A]:ILE:HG12	2:B:285:LYS:HE2	1.87	0.56
2:B:151:GLU:O	2:B:153:ILE:N	2.32	0.54
2:B:385:LEU:HD21	2:B:392:ILE:HD11	1.92	0.50
1:A:459:LYS:HB3	1:A:507:VAL:HG23	1.93	0.50
1:A:433:THR:HG23	5:A:807:HOH:O	2.11	0.49
1:A:539:ASN:O	1:A:541:GLN:HG2	2.13	0.48
2:B:260[B]:CYS:SG	2:B:287:HIS:CE1	3.08	0.47
2:B:117:ASN:N	5:B:516:HOH:O	2.48	0.46
2:B:128:HIS:NE2	2:B:269:PRO:HG3	2.31	0.46
2:B:179:ILE:HG23	5:B:509:HOH:O	2.13	0.46
2:B:151:GLU:C	2:B:153:ILE:H	2.18	0.45
1:A:380:TYR:OH	5:A:701:HOH:O	2.19	0.44
2:B:156:LYS:O	2:B:160:ILE:HG13	2.18	0.44
2:B:230:PHE:CE1	2:B:339:LEU:HD22	2.53	0.44
1:A:444[A]:LEU:HD13	1:A:449:TYR:CB	2.49	0.43
1:A:510:THR:O	1:A:511:SER:HB2	2.18	0.42
2:B:321:ILE:HD12	2:B:321:ILE:HA	1.94	0.41
1:A:456[A]:ILE:HG23	1:A:480:LYS:HB2	2.02	0.41
1:A:459:LYS:HB3	1:A:507:VAL:CG2	2.51	0.41

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 X-ray entries followed by that with respect to entries of similar resolution.

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	194/246 (79%)	189 (97%)	5 (3%)	0	100	100
2	B	235/290 (81%)	227 (97%)	7 (3%)	1 (0%)	34	21
All	All	429/536 (80%)	416 (97%)	12 (3%)	1 (0%)	47	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	152	LEU

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 X-ray entries followed by that with respect to entries of similar resolution.

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	166/215 (77%)	164 (99%)	2 (1%)	71	65
2	B	194/259 (75%)	193 (100%)	1 (0%)	88	87
All	All	360/474 (76%)	357 (99%)	3 (1%)	86	78

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

Mol	Chain	Res	Type
1	A	444[A]	LEU
1	A	444[B]	LEU
2	B	173	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are

no such sidechains identified.

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)

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	#$|Z| > 2$
3	UXK	A	601	-	44,45,45	1.66	6 (13%)	57,64,64	1.77	15 (26%)
4	ACT	B	401	-	1,3,3	8.28	1 (100%)	0,3,3	0.00	-

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
3	UXK	A	601	-	-	1/17/53/53	0/6/6/6

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	401	ACT	CH3-C	8.28	1.59	1.48
3	A	601	UXK	O36-C37	-4.67	1.37	1.44
3	A	601	UXK	C26-N27	3.20	1.41	1.36
3	A	601	UXK	C17-C16	-3.20	1.48	1.52
3	A	601	UXK	C01-C02	3.17	1.59	1.53
3	A	601	UXK	C38-N14	2.58	1.50	1.46
3	A	601	UXK	C20-N21	-2.36	1.43	1.46

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	601	UXK	O36-C37-C38	3.78	116.57	111.53
3	A	601	UXK	C28-N27-C26	3.58	130.45	123.40
3	A	601	UXK	C35-C26-N25	-3.18	118.32	122.75
3	A	601	UXK	C09-C08-N11	-3.09	117.72	122.17
3	A	601	UXK	C10-C13-C12	-2.99	112.40	118.07
3	A	601	UXK	C35-C22-N21	2.89	125.73	122.29
3	A	601	UXK	C19-C16-C17	-2.74	106.01	109.56
3	A	601	UXK	C37-O36-C16	2.73	118.83	114.70
3	A	601	UXK	C12-N11-C08	2.50	121.02	117.82
3	A	601	UXK	C35-C26-N27	2.48	124.90	120.86
3	A	601	UXK	C29-C28-N27	2.47	120.19	113.77
3	A	601	UXK	C40-C39-N06	-2.41	108.86	111.23
3	A	601	UXK	C09-C10-C13	2.36	123.42	120.32
3	A	601	UXK	C39-N06-C05	-2.28	103.70	108.83
3	A	601	UXK	O36-C16-C19	-2.03	101.93	108.69

There are no chirality outliers.

All (1) torsion outliers are listed below:

| Mol	Chain	Res	Type	Atoms
3	A	601	UXK	C08-C07-N06-C05

There are no ring outliers.

No monomer is involved in short contacts.

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 UXK A 601

Bond lengths

Bond angles

Torsions

Rings
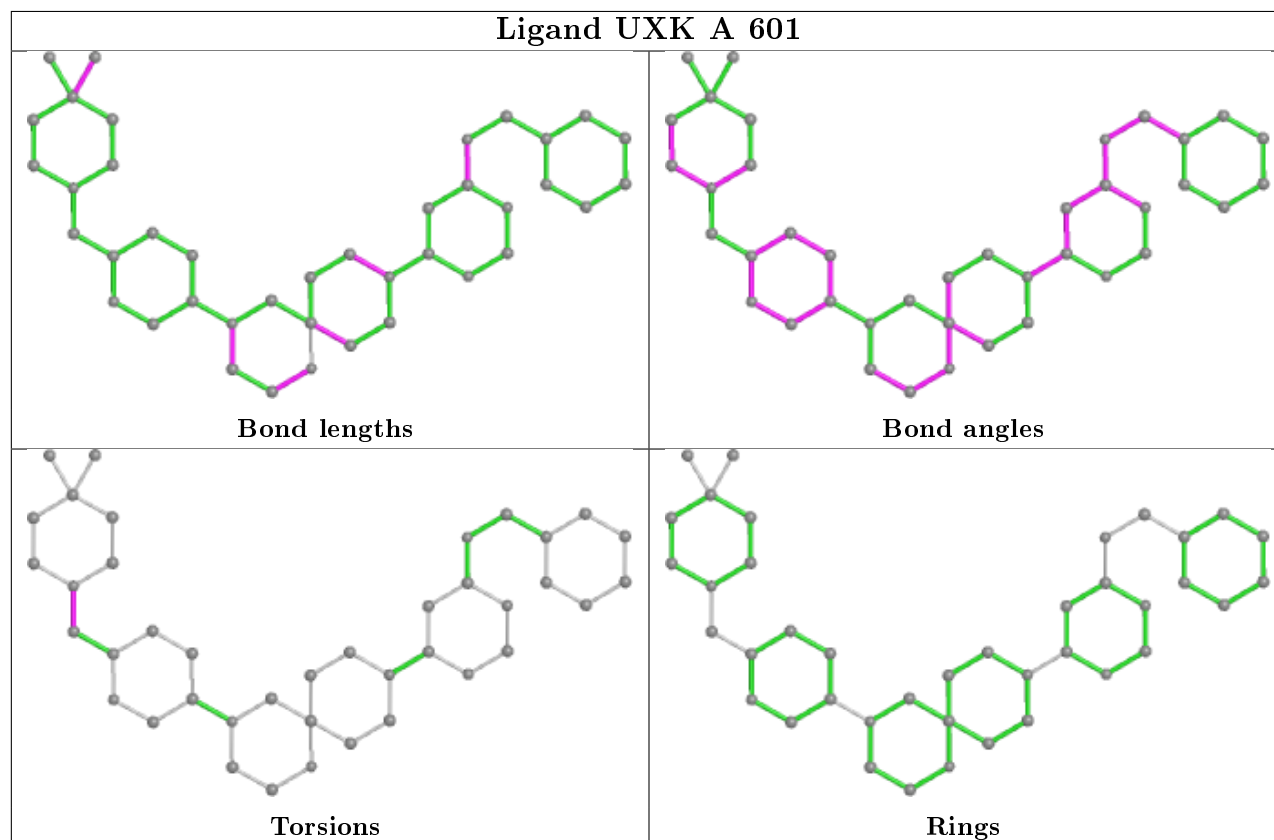
5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.