



wwPDB EM Validation Summary Report ⓘ

Mar 31, 2025 – 04:52 PM JST

PDB ID : 6J99 / pdb_00006j99
EMDB ID : EMD-9783
Title : Cryo-EM structure of human DOT1L in complex with an H2B-monoubiquitinated nucleosome
Authors : Yao, T.; Huang, J.
Deposited on : 2019-01-22
Resolution : 4.10 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 : 0.0.1.dev117
Mogul : 1.8.5 (274361), CSD as541be (2020)
MolProbity : 4.02b-467
buster-report : 1.1.7 (2018)
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.42

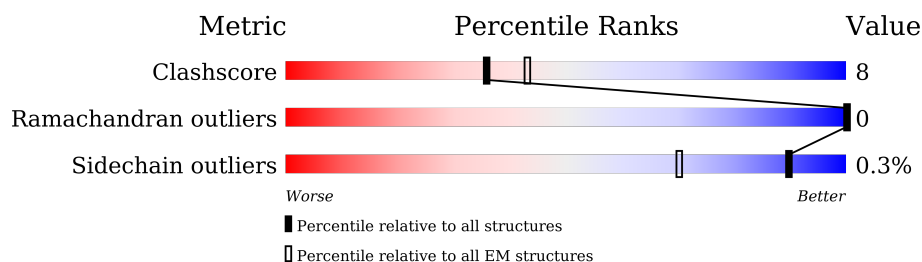
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.10 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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	136	50% 21% 29%
1	E	136	51% 20% 29%
2	B	103	66% 17% 17%
2	F	103	71% 8% 21%
3	C	130	65% 18% 17%
3	G	130	66% 15% 19%
4	D	123	61% 15% 24%
4	H	123	63% 15% 23%
5	I	147	78% 19% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	J	147	 76% 22% ..
7	L	76	 53% 47%
8	K	416	 57% 21% 22%

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 15266 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 Histone H3.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	97	Total	C	N	O	S	0	0
			802	506	155	138	3		
1	E	97	Total	C	N	O	S	0	0
			802	506	155	138	3		

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	86	Total	C	N	O	S	0	0
			698	440	141	116	1		
2	F	81	Total	C	N	O	S	0	0
			648	410	126	111	1		

- Molecule 3 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	C	108	Total	C	N	O	0	0
			834	525	165	144		
3	G	105	Total	C	N	O	0	0
			809	510	158	141		

- Molecule 4 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	94	Total	C	N	O	S	0	0
			738	463	134	138	3		
4	H	95	Total	C	N	O	S	0	0
			742	466	133	140	3		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	MET	-	initiating methionine	UNP P02281

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	29	THR	SER	engineered mutation	UNP P02281
D	117	CYS	LYS	engineered mutation	UNP P02281
H	0	MET	-	initiating methionine	UNP P02281
H	29	THR	SER	engineered mutation	UNP P02281
H	117	CYS	LYS	engineered mutation	UNP P02281

- Molecule 5 is a DNA chain called DNA (144-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
5	I	144	Total	C	N	O	P	0	0
			2935	1393	536	862	144		

- Molecule 6 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	145	Total	C	N	O	P	0	0
			2990	1415	559	871	145		

- Molecule 7 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	L	76	Total	C	N	O	S	0	0
			603	379	105	117	2		

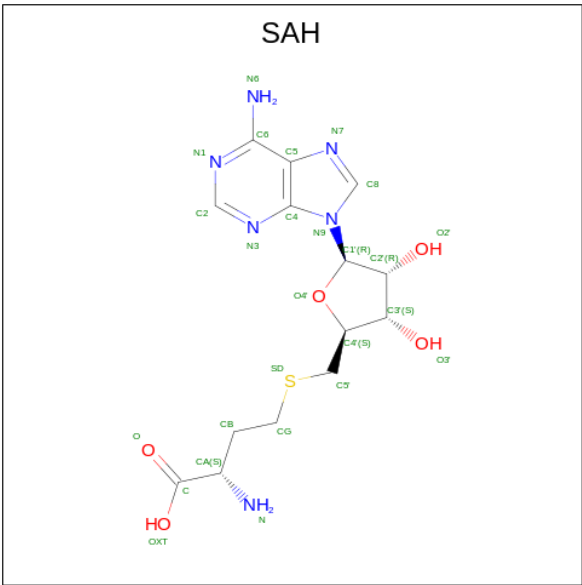
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	76	CYS	GLY	engineered mutation	UNP P62979

- Molecule 8 is a protein called Histone-lysine N-methyltransferase, H3 lysine-79 specific.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	K	324	Total	C	N	O	S	0	0
			2639	1684	451	492	12		

- Molecule 9 is S-ADENOSYL-L-HOMOCYSTEINE (CCD ID: SAH) (formula: C₁₄H₂₀N₆O₅S).



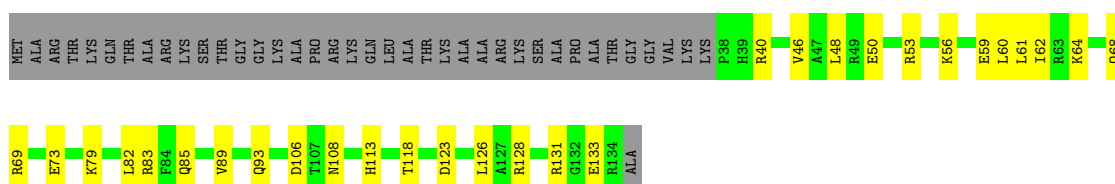
Mol	Chain	Residues	Atoms					AltConf
9	K	1	Total	C	N	O	S	0
			26	14	6	5	1	

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.

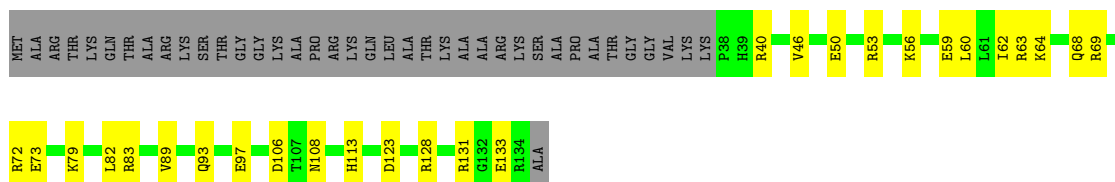
- Molecule 1: Histone H3

Chain A: 



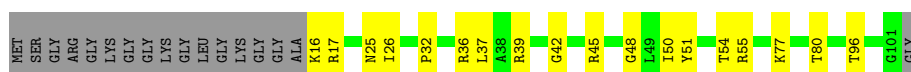
- Molecule 1: Histone H3

Chain E: 



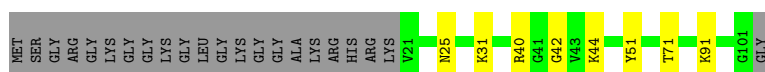
- Molecule 2: Histone H4

Chain B: 



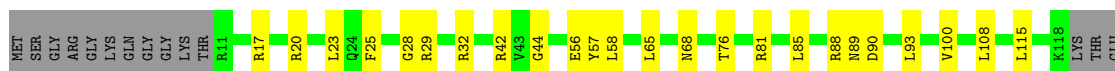
- Molecule 2: Histone H4

Chain F: 



- Molecule 3: Histone H2A

Chain C: 



SER
SER
LYS
LYS
ALA
LYS
SER
LYS

• Molecule 3: Histone H2A

Chain G:  66% 15% 19%

MET SER GLY ARG GLY GLN GLY THR ARG ALA K13 A14 K15 T16 R17 S18 L23 Q24 F25 G28 R29 R42 V43 Q44 T76 T79 P80 R81 H82 L85 N89 T101 N10 L116 P117 LYS LYS THR GLU SER SER LYS SER ALA LYS LYS

• Molecule 4: Histone H2B 1.1

Chain D:  61% 15% 24%

MET ALA SER SER PRO ALA PRO LYS LYS GLY THR LYS LYS VAL THR THR GLN LYS LYS ASP GLY LYS ARG R27 K28 T29 R30 K31 E32 S33 Y37 V45 E68 N81 K82 R83 S84 T85 R89 L99 E102 L103 A104 K105 H106 E110 K113

A114 S120 ALA LYS


• Molecule 4: Histone H2B 1.1

Chain H:  63% 15% 23%

MET ALA LYS SER PRO ALA PRO LYS LYS GLY THR LYS LYS VAL THR THR GLN LYS LYS ASP GLY LYS ARG K28 E32 S33 Y34 Y37 G50 R69 E73 R76 S84 T87 E90 T93 A94 V95 L98 L99 P100 G101 E102


H106 A114 K122

• Molecule 5: DNA (144-MER)

Chain I:  78% 19% ..

DA DT C3 A7 A8 T9 C17 C18 C28 A29 A30 T31 T32 A39 G44 G50 G55 G56 A61 G83 C84 T87 A91 A98 G99 G100 C109 A113 A120 T127 G128 T129 C140 A146 DT

• Molecule 6: DNA (145-MER)

Chain J:  76% 22% ..

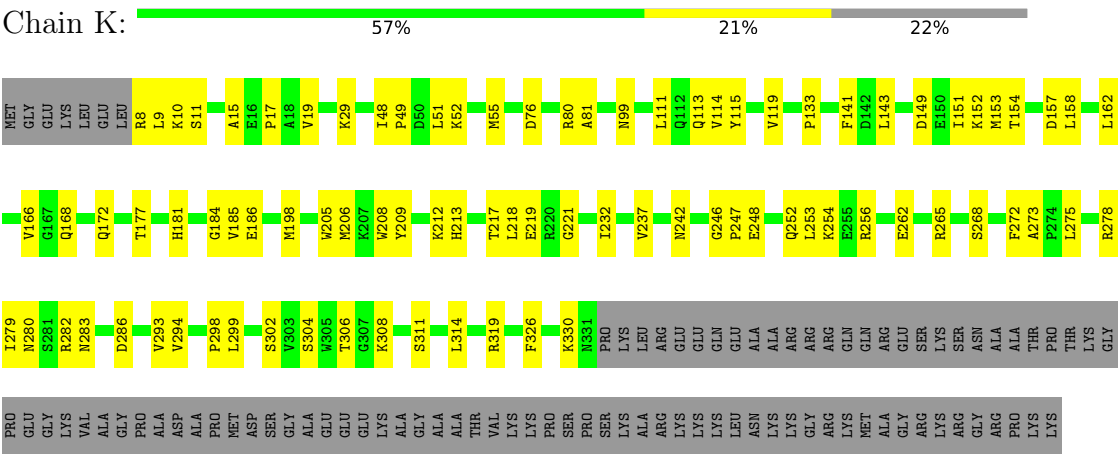
DA T2 T7 G8 G30 A31 G32 A40 C47 T50 T51 T58 A59 A60 T79 A80 D81 C81 G82 T83 G84 A91 G92 G95 T96 C108 T109 A110 C111 G112 G122 C123 G124 G125 C126 C132 A146 DT

• Molecule 7: Ubiquitin

Chain L:  53% 47%

M1 Q2 I3 F4 V5 K6 T7 K11 T14 L15 P19 T22 I23 K27 T30 Q31 D32 A80 C81 E34 K33 G35 I36 P37 P38 D39 R42 L43 E51 D52 G53 E54 T55 L56 S57 D58 Y59 N60 I61 Q62 K63 E64 L67 V70 R74 G75 C76

● Molecule 8: Histone-lysine N-methyltransferase, H3 lysine-79 specific



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	151545	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	0.98	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (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: SAH

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	0.39	0/814	0.57	0/1092
1	E	0.39	0/814	0.57	0/1092
2	B	0.42	0/706	0.58	0/943
2	F	0.43	0/655	0.56	0/878
3	C	0.39	0/844	0.55	0/1138
3	G	0.34	0/819	0.52	0/1106
4	D	0.41	0/749	0.53	0/1008
4	H	0.35	0/753	0.51	0/1012
5	I	0.93	0/3289	1.10	7/5069 (0.1%)
6	J	0.93	0/3357	1.04	2/5184 (0.0%)
7	L	0.31	0/609	0.61	0/819
8	K	0.36	0/2708	0.59	0/3672
All	All	0.66	0/16117	0.83	9/23013 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
4	H	0	1
8	K	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	I	39	DA	O4'-C4'-C3'	-6.06	102.08	104.50
5	I	44	DG	O4'-C1'-N9	5.97	112.18	108.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	I	109	DC	O4'-C4'-C3'	-5.74	102.20	104.50
5	I	56	DC	O4'-C1'-N1	5.68	111.98	108.00
5	I	140	DC	O4'-C4'-C3'	-5.52	102.29	104.50

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	H	100	PRO	Peptide
8	K	304	SER	Peptide

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	802	0	841	24	0
1	E	802	0	841	20	0
2	B	698	0	752	18	0
2	F	648	0	693	7	0
3	C	834	0	895	24	0
3	G	809	0	864	18	0
4	D	738	0	760	17	0
4	H	742	0	765	15	0
5	I	2935	0	1615	21	0
6	J	2990	0	1628	31	0
7	L	603	0	631	24	0
8	K	2639	0	2582	54	0
9	K	26	0	17	3	0
All	All	15266	0	12884	203	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 8.

The worst 5 of 203 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)
7:L:30:ILE:O	7:L:34:GLU:HB2	1.73	0.88
8:K:133:PRO:HB2	9:K:501:SAH:H8	1.72	0.71
3:G:29:ARG:HH22	4:H:32:GLU:HB3	1.56	0.70
1:A:123:ASP:OD1	1:E:113:HIS:NE2	2.20	0.68
3:G:28:GLY:HA3	6:J:30:DG:H3'	1.75	0.67

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	95/136 (70%)	88 (93%)	7 (7%)	0	100	100
1	E	95/136 (70%)	88 (93%)	7 (7%)	0	100	100
2	B	84/103 (82%)	80 (95%)	4 (5%)	0	100	100
2	F	79/103 (77%)	76 (96%)	3 (4%)	0	100	100
3	C	106/130 (82%)	98 (92%)	8 (8%)	0	100	100
3	G	103/130 (79%)	96 (93%)	7 (7%)	0	100	100
4	D	92/123 (75%)	89 (97%)	3 (3%)	0	100	100
4	H	93/123 (76%)	90 (97%)	3 (3%)	0	100	100
7	L	74/76 (97%)	66 (89%)	8 (11%)	0	100	100
8	K	322/416 (77%)	280 (87%)	42 (13%)	0	100	100
All	All	1143/1476 (77%)	1051 (92%)	92 (8%)	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	85/111 (77%)	84 (99%)	1 (1%)	67	79
1	E	85/111 (77%)	84 (99%)	1 (1%)	67	79
2	B	72/79 (91%)	72 (100%)	0	100	100
2	F	67/79 (85%)	67 (100%)	0	100	100
3	C	85/102 (83%)	85 (100%)	0	100	100
3	G	83/102 (81%)	82 (99%)	1 (1%)	67	79
4	D	81/103 (79%)	81 (100%)	0	100	100
4	H	81/103 (79%)	81 (100%)	0	100	100
7	L	69/69 (100%)	69 (100%)	0	100	100
8	K	290/360 (81%)	290 (100%)	0	100	100
All	All	998/1219 (82%)	995 (100%)	3 (0%)	90	92

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

Mol	Chain	Res	Type
1	A	64	LYS
1	E	64	LYS
3	G	110	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
8	K	315	HIS
8	K	234	ASN
3	G	110	ASN
3	G	38	ASN
8	K	126	ASN

5.3.3 RNA ⓘ

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](#)

1 ligand is 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$
9	SAH	K	501	8	24,28,28	1.16	2 (8%)	25,40,40	1.75	6 (24%)

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
9	SAH	K	501	8	-	4/11/31/31	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	K	501	SAH	C2-N3	3.13	1.37	1.32
9	K	501	SAH	OXT-C	-2.15	1.23	1.30

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	K	501	SAH	N3-C2-N1	-5.17	120.59	128.68
9	K	501	SAH	C5'-SD-CG	-3.95	90.41	102.27

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	K	501	SAH	OXT-C-O	-2.61	118.17	124.09
9	K	501	SAH	OXT-C-CA	2.23	120.97	113.38
9	K	501	SAH	C4-C5-N7	-2.18	107.12	109.40

There are no chirality outliers.

All (4) torsion outliers are listed below:

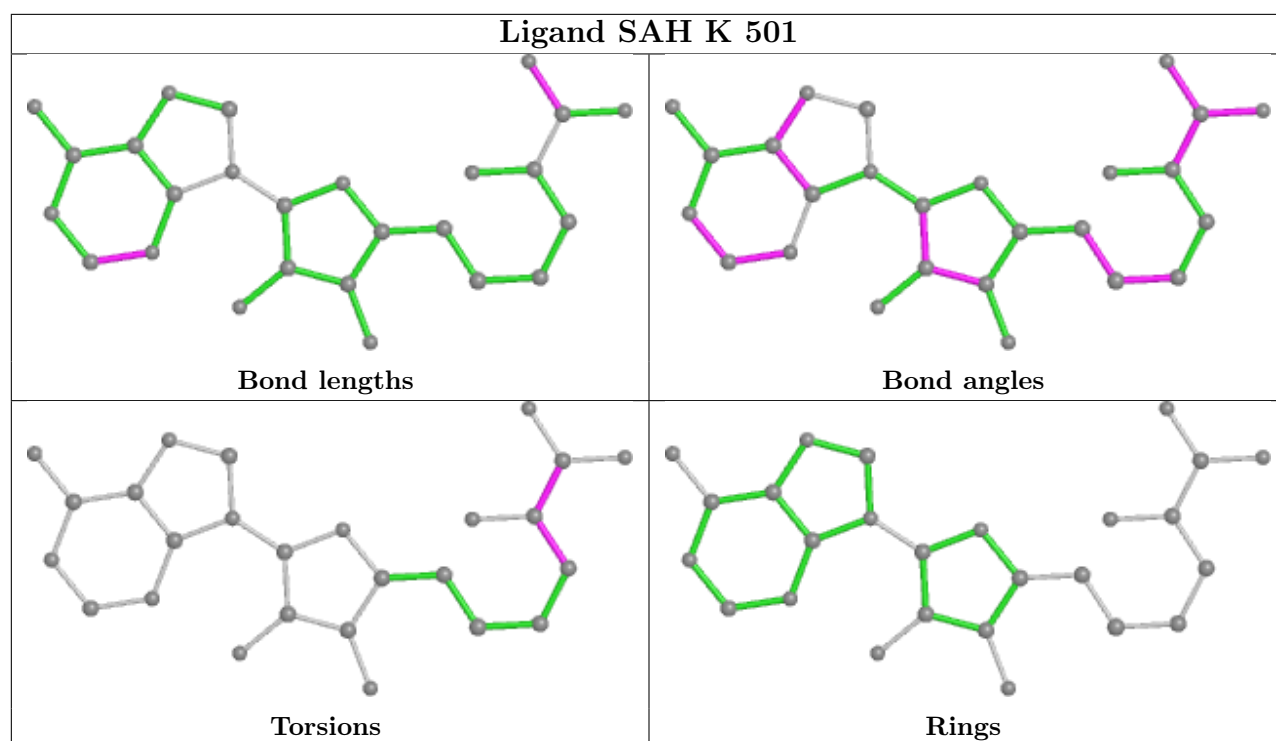
Mol	Chain	Res	Type	Atoms
9	K	501	SAH	N-CA-CB-CG
9	K	501	SAH	C-CA-CB-CG
9	K	501	SAH	O-C-CA-N
9	K	501	SAH	OXT-C-CA-N

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	K	501	SAH	3	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 [i](#)

There are no chain breaks in this entry.

6 Map visualisation

This section contains visualisations of the EMDB entry EMD-9783. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal standard-deviation projections (False-color)

This section was not generated.

6.5 Orthogonal surface views

This section was not generated.

6.6 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis ⓘ

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution ⓘ

This section was not generated.

7.2 Volume estimate versus contour level ⓘ

This section was not generated.

7.3 Rotationally averaged power spectrum ⓘ

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

This section was not generated.