



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 2, 2023 – 05:55 AM EDT

PDB ID : 6N3G  
Title : Crystal structure of histone lysine methyltransferase SmyD2 in complex with polyethylene glycol  
Authors : Perry, E.; Spellmon, N.; Brunzelle, J.; Yang, Z.  
Deposited on : 2018-11-15  
Resolution : 2.43 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](mailto: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 ⓘ 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:

MolProbity : **FAILED**  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : **FAILED**  
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) : 2.35.1

## 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.43 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 3633 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.

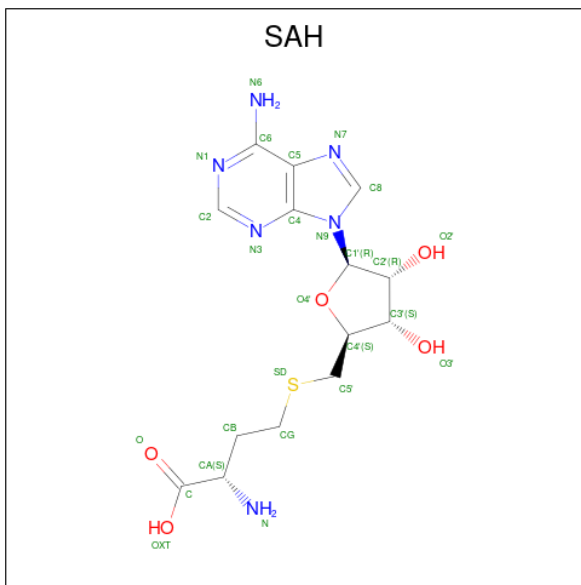
- Molecule 1 is a protein called N-lysine methyltransferase SMYD2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	430	3471	2201	593	647	30	0	1	0

There is a discrepancy between the modelled and reference sequences:

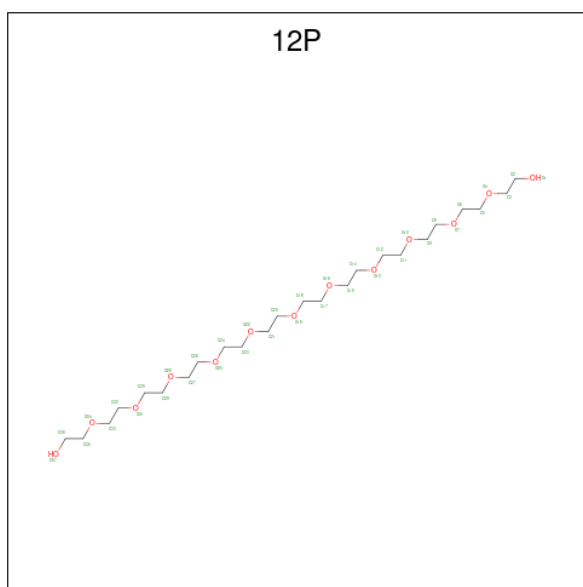
Chain	Residue	Modelled	Actual	Comment	Reference
A	165	GLU	GLY	conflict	UNP Q9NRG4

- Molecule 2 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula:  $C_{14}H_{20}N_6O_5S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	A	1	26	14	6	5	1	0	0

- Molecule 3 is DODECAETHYLENE GLYCOL (three-letter code: 12P) (formula:  $C_{24}H_{50}O_{13}$ ).

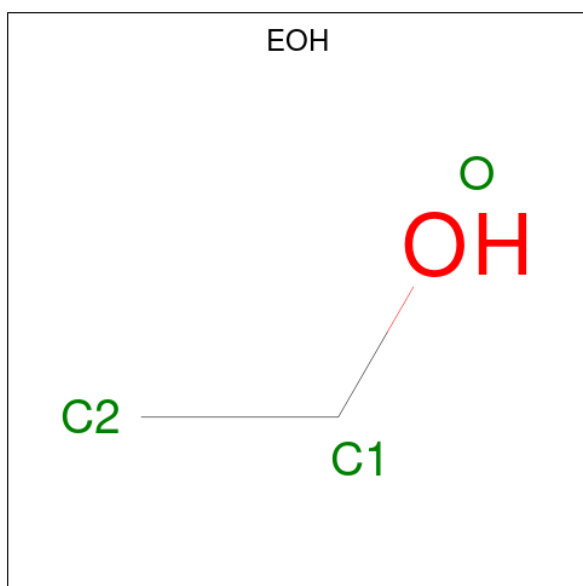


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			37	24	13		

- Molecule 4 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Ni	0	0
			2	2		

- Molecule 5 is ETHANOL (three-letter code: EOH) (formula: C<sub>2</sub>H<sub>6</sub>O).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			3	2	1		
5	A	1	Total	C	O	0	0
			3	2	1		
5	A	1	Total	C	O	0	0
			3	2	1		
5	A	1	Total	C	O	0	0
			3	2	1		

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	3	Total	Zn	0	0
			3	3		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	79	Total	O	0	0
			79	79		

MolProbity and EDS failed to run properly - this section is therefore empty.

### 3 Data and refinement statistics i

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	151.84Å 151.84Å 53.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	75.92 – 2.43	Depositor
% Data completeness (in resolution range)	61.5 (75.92-2.43)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.78 (at 2.42Å)	Xtrriage
Refinement program	PHENIX 1.11.1-2575	Depositor
R, $R_{free}$	0.175 , 0.222	Depositor
Wilson B-factor (Å <sup>2</sup> )	38.2	Xtrriage
Anisotropy	0.014	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.019 for -k,-h,-l	Xtrriage
Total number of atoms	3633	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.72% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 4 Model quality [i](#)

### 4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.3 Torsion angles [i](#)

#### 4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

#### 4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

#### 4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 4.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 5 are monoatomic - leaving 7 for Mogul analysis.

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	EOH	A	506	-	2,2,2	0.51	0	1,1,1	0.24	0
5	EOH	A	507	-	2,2,2	0.54	0	1,1,1	0.19	0
3	12P	A	502	-	36,36,36	0.60	0	35,35,35	0.49	0
2	SAH	A	501	-	24,28,28	1.35	4 (16%)	25,40,40	1.89	5 (20%)
5	EOH	A	509	-	2,2,2	0.59	0	1,1,1	0.15	0
5	EOH	A	505	-	2,2,2	0.49	0	1,1,1	0.00	0
5	EOH	A	508	-	2,2,2	0.55	0	1,1,1	0.02	0

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	12P	A	502	-	-	19/34/34/34	-
2	SAH	A	501	-	-	5/11/31/31	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	SAH	C2-N3	3.75	1.38	1.32
2	A	501	SAH	C2-N1	2.56	1.38	1.33
2	A	501	SAH	C2'-C1'	-2.53	1.49	1.53
2	A	501	SAH	OXT-C	-2.15	1.23	1.30

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	SAH	N3-C2-N1	-5.57	119.97	128.68
2	A	501	SAH	CB-CG-SD	3.96	122.19	113.31
2	A	501	SAH	OXT-C-O	-3.30	116.60	124.09
2	A	501	SAH	C5'-SD-CG	-2.62	94.40	102.27
2	A	501	SAH	O4'-C4'-C3'	-2.39	100.38	105.11

There are no chirality outliers.

All (24) torsion outliers are listed below:

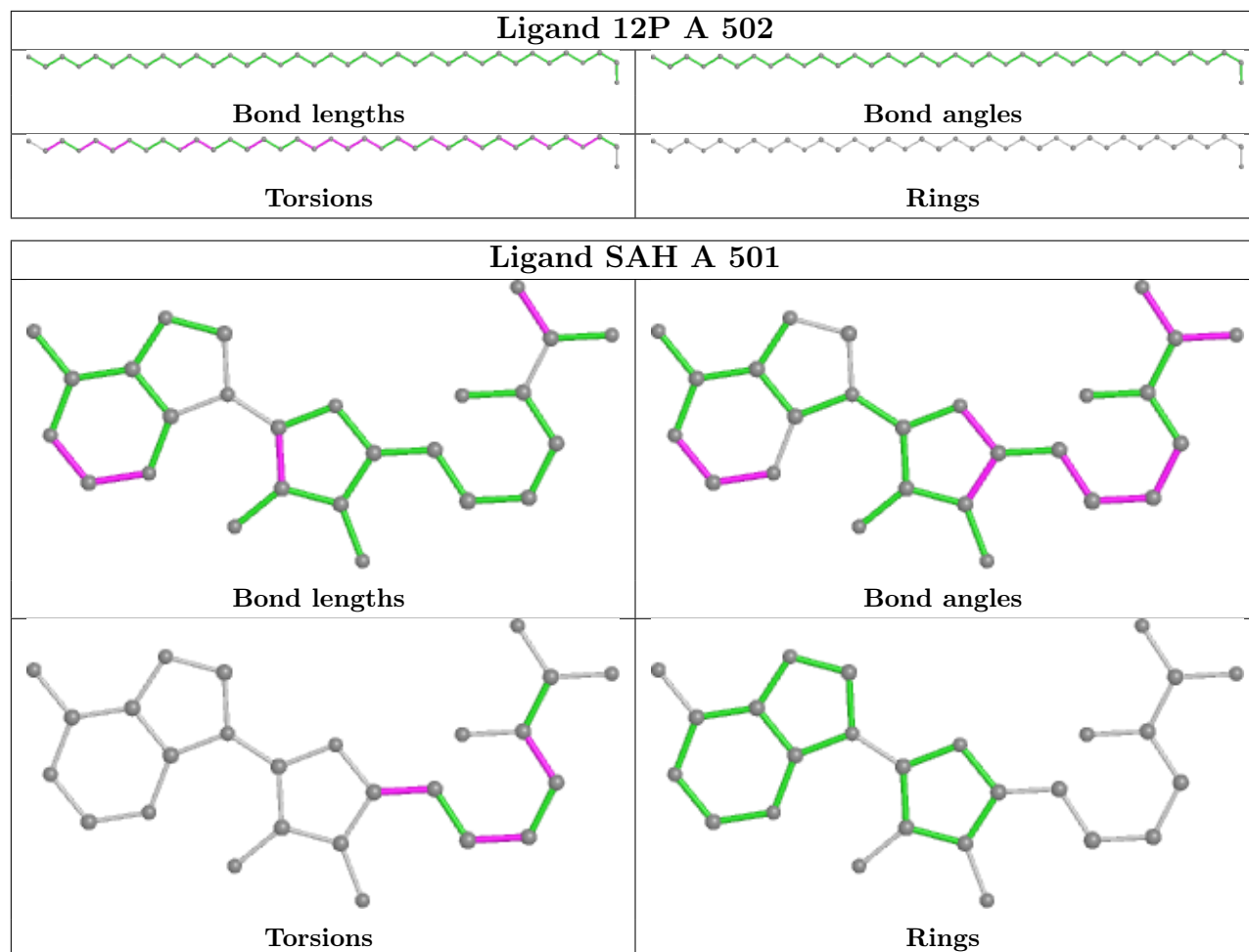


Mol	Chain	Res	Type	Atoms
3	A	502	12P	O28-C29-C30-O31
3	A	502	12P	O10-C11-C12-O13
3	A	502	12P	O22-C23-C24-O25
3	A	502	12P	O13-C14-C15-O16
3	A	502	12P	C5-C6-O7-C8
3	A	502	12P	O4-C5-C6-O7
3	A	502	12P	O1-C2-C3-O4
2	A	501	SAH	C-CA-CB-CG
3	A	502	12P	O19-C20-C21-O22
3	A	502	12P	C33-C32-O31-C30
3	A	502	12P	C12-C11-O10-C9
3	A	502	12P	C6-C5-O4-C3
2	A	501	SAH	N-CA-CB-CG
2	A	501	SAH	C3'-C4'-C5'-SD
3	A	502	12P	C27-C26-O25-C24
3	A	502	12P	O16-C17-C18-O19
3	A	502	12P	C26-C27-O28-C29
2	A	501	SAH	O4'-C4'-C5'-SD
3	A	502	12P	C20-C21-O22-C23
3	A	502	12P	C17-C18-O19-C20
3	A	502	12P	C32-C33-O34-C35
2	A	501	SAH	CB-CG-SD-C5'
3	A	502	12P	C21-C20-O19-C18
3	A	502	12P	C36-C35-O34-C33

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 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.



#### 4.7 Other polymers [\(i\)](#)

There are no such residues in this entry.

#### 4.8 Polymer linkage issues [\(i\)](#)

There are no chain breaks in this entry.

## 5 Fit of model and data [i](#)

### 5.1 Protein, DNA and RNA chains [i](#)

EDS failed to run properly - this section is therefore empty.

### 5.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS failed to run properly - this section is therefore empty.

### 5.3 Carbohydrates [i](#)

EDS failed to run properly - this section is therefore empty.

### 5.4 Ligands [i](#)

EDS failed to run properly - this section is therefore empty.

### 5.5 Other polymers [i](#)

EDS failed to run properly - this section is therefore empty.