# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 16, 2021 – 02:20 pm BST

| | | |
|---|---|---|
| PDB ID | : | 7O2C |
| Title | : | X-RAY STRUCTURE OF SMYD3 IN COMPLEX WITH the benzodiazepine-based probe BAY-6035 |
| Authors | : | Steuber, H. |
| Deposited on | : | 2021-03-30 |
| Resolution | : | 1.52 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 ⓘ symbol.

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

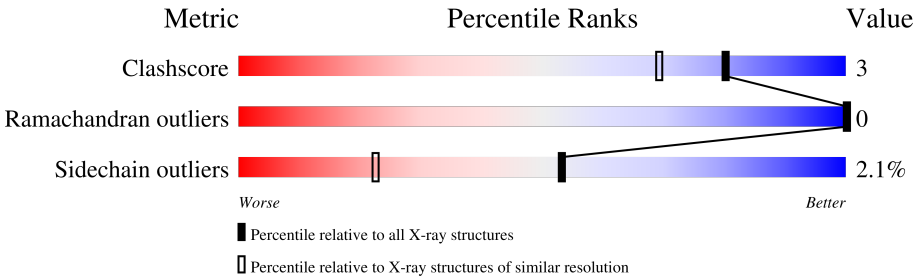
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| Xtriage (Phenix) | : | NOT EXECUTED |
| EDS | : | NOT EXECUTED |
| buster-report | : | 1.1.7 (2018) |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| 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

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

The reported resolution of this entry is 1.52 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) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| Clashscore | 141614 | 4249 (1.54-1.50) |
| Ramachandran outliers | 138981 | 4148 (1.54-1.50) |
| Sidechain outliers | 138945 | 4146 (1.54-1.50) |

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 | 428 | 91% 8% . |

# 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4031 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.

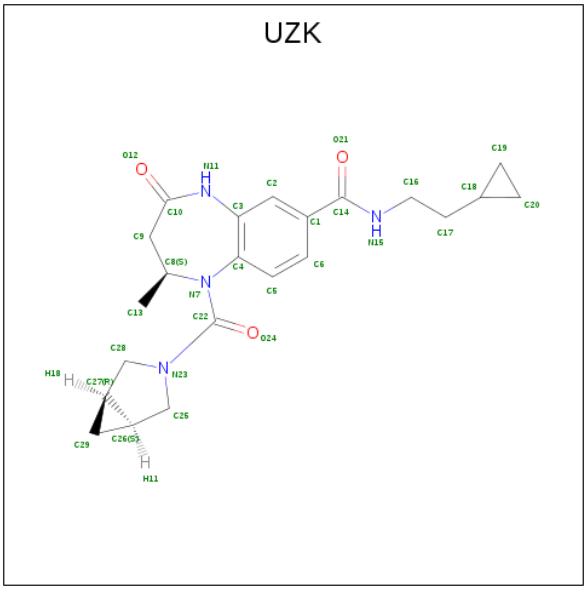
- Molecule 1 is a protein called Histone-lysine N-methyltransferase SMYD3.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 427 | Total | C | N | O | S | 0 | 7 | 0 |
| | | | 3428 | 2164 | 594 | 632 | 38 | | | |

There are 2 discrepancies between the modelled and reference sequences:

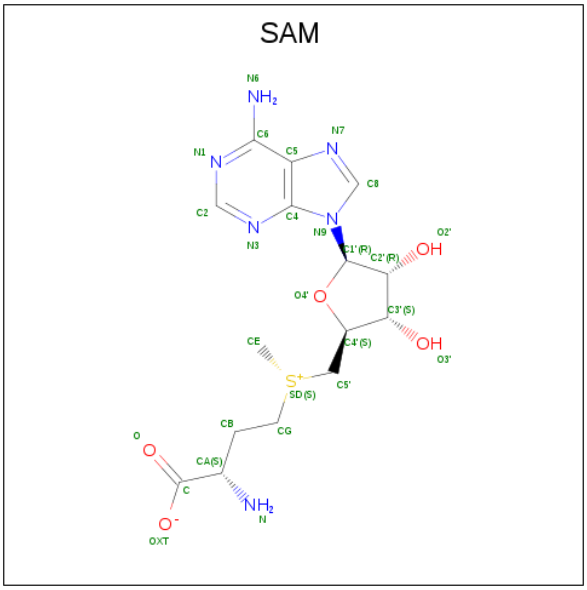
| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| A | 13 | ASN | LYS | variant | UNP Q9H7B4 |
| A | 140 | ARG | LYS | variant | UNP Q9H7B4 |

- Molecule 2 is (2S)-1-[[(1R,5S)-3-azabicyclo[3.1.0]hexan-3-yl]carbonyl]-N-(2-cyclopropylethyl)-2-methyl-4-oxidanylidene-3,5-dihydro-2H-1,5-benzodiazepine-7-carboxamide (three-letter code: UZK) (formula: $C_{22}H_{28}N_4O_3$) (labeled as "Ligand of Interest" by depositor).




| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|
| 2 | A | 1 | Total | C | N | O | 0 | 0 |
| | | | 29 | 22 | 4 | 3 | | |

- Molecule 3 is S-ADENOSYLMETHIONINE (three-letter code: SAM) (formula: $C_{15}H_{22}N_6O_5S$).




| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|---|
| 3 | A | 1 | Total | C | N | O | S | 0 | 0 |
| | | | 27 | 15 | 6 | 5 | 1 | | |

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 4 | A | 3 | Total | Zn | 0 | 0 |
| | | | 3 | 3 | | |

- Molecule 5 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 5 | A | 544 | Total | O | 0 | 0 |
| | | | 544 | 544 | | |

# 3 Residue-property plots (i)

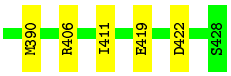
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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.

Note EDS was not executed.

- Molecule 1: Histone-lysine N-methyltransferase SMYD3

Chain A:





# 4 Data and refinement statistics (i)

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

| Property | Value | Source |
|---|---|---|
| Space group | P 21 21 21 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 61.50Å 66.17Å 107.17Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 41.64 – 1.52 | Depositor |
| % Data completeness<br>(in resolution range) | 99.5 (41.64-1.52) | Depositor |
| $R_{merge}$ | (Not available) | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| Refinement program | REFMAC 5.8.0258 | Depositor |
| R, $R_{free}$ | 0.158 , 0.225 | Depositor |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| Total number of atoms | 4031 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 20.0 | wwPDB-VP |

# 5 Model quality

## 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, SAM, UZK

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 | 0.88 | 2/3514 (0.1%) | 0.99 | 5/4735 (0.1%) |

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 |
|---|---|---|---|
| 1 | A | 0 | 1 |

All (2) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 1 | A | 230 | GLU | CD-OE2 | -6.93 | 1.18 | 1.25 |
| 1 | A | 189 | GLU | CD-OE1 | 5.08 | 1.31 | 1.25 |

All (5) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 1 | A | 155 | MET | CG-SD-CE | -8.11 | 87.23 | 100.20 |
| 1 | A | 239 | TYR | CB-CG-CD1 | 6.15 | 124.69 | 121.00 |
| 1 | A | 272 | ASP | CB-CG-OD2 | -5.55 | 113.31 | 118.30 |
| 1 | A | 14 | ARG | CB-CG-CD | -5.48 | 97.35 | 111.60 |
| 1 | A | 107 | ARG | NE-CZ-NH2 | -5.03 | 117.78 | 120.30 |

There are no chirality outliers.

All (1) planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|---|---|---|---|---|
| 1 | A | 2 | GLU | 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 | 3428 | 0 | 3416 | 18 | 0 |
| 2 | A | 29 | 0 | 0 | 1 | 0 |
| 3 | A | 27 | 0 | 22 | 0 | 0 |
| 4 | A | 3 | 0 | 0 | 0 | 0 |
| 5 | A | 544 | 0 | 0 | 4 | 2 |
| All | All | 4031 | 0 | 3438 | 19 | 2 |

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 3.

All (19) 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:154[A]:PHE:CD2 | 1:A:155:MET:SD | 2.85 | 0.70 |
| 1:A:390:MET:HE1 | 1:A:422:ASP:N | 2.08 | 0.68 |
| 1:A:352[A]:THR:HG21 | 5:A:949:HOH:O | 1.95 | 0.66 |
| 1:A:318:GLU:HG2 | 5:A:788:HOH:O | 2.05 | 0.56 |
| 1:A:348:LEU:O | 1:A:352[A]:THR:HG23 | 2.05 | 0.56 |
| 2:A:501:UZK:C4 | 2:A:501:UZK:C28 | 2.85 | 0.54 |
| 1:A:2:GLU:CB | 1:A:3:PRO:O | 2.59 | 0.50 |
| 1:A:390:MET:HE1 | 1:A:422:ASP:CA | 2.42 | 0.50 |
| 1:A:35:PRO:HG2 | 1:A:38:TYR:CD2 | 2.47 | 0.50 |
| 1:A:332[A]:ASP:OD2 | 1:A:358:TYR:OH | 2.27 | 0.48 |
| 1:A:264:PHE:CE1 | 1:A:268:THR:HG21 | 2.49 | 0.47 |
| 1:A:49:CYS:HB2 | 1:A:71:CYS:HB3 | 1.95 | 0.47 |
| 1:A:6:VAL:HG12 | 1:A:235:LEU:HD11 | 2.00 | 0.44 |
| 1:A:84:LYS:HD3 | 5:A:995:HOH:O | 2.16 | 0.44 |
| 1:A:64:GLN:N | 1:A:64:GLN:HE21 | 2.17 | 0.42 |
| 1:A:419:GLU:HG3 | 5:A:609:HOH:O | 2.19 | 0.42 |
| 1:A:154[A]:PHE:CE2 | 1:A:155:MET:SD | 3.14 | 0.41 |
| 1:A:22:THR:HB | 1:A:23:PRO:HD2 | 2.02 | 0.41 |

*Continued on next page...*

*Continued from previous page...*

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|---|---|---|---|
| 1:A:17:GLY:HA2 | 1:A:203:LEU:O | 2.21 | 0.40 |

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|---|---|---|---|
| 5:A:676:HOH:O | 5:A:1066:HOH:O[3_554] | 1.92 | 0.28 |
| 5:A:638:HOH:O | 5:A:1066:HOH:O[1_565] | 1.98 | 0.22 |

## 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 | 432/428 (101%) | 422 (98%) | 10 (2%) | 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 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 | 382/382 (100%) | 373 (98%) | 9 (2%) | 49 19 |

All (9) residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 14 | ARG |
| 1 | A | 50 | ASP |
| 1 | A | 51[A] | ARG |
| 1 | A | 51[B] | ARG |
| 1 | A | 60 | MET |
| 1 | A | 64 | GLN |
| 1 | A | 156 | ARG |
| 1 | A | 406 | ARG |
| 1 | A | 411 | ILE |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 16 | ASN |
| 1 | A | 64 | GLN |
| 1 | A | 382 | HIS |
| 1 | A | 424 | 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 monosaccharides in this entry.

### 5.6 Ligand geometry

Of 5 ligands modelled in this entry, 3 are monoatomic - leaving 2 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 | $\#\|Z\|>2$ |
| 2 | UZK | A | 501 | - | 32,33,33 | 2.45 | 14 (43%) | 37,49,49 | 3.47 | 19 (51%) |
| 3 | SAM | A | 502 | - | 21,29,29 | 0.68 | 0 | 18,42,42 | 0.86 | 1 (5%) |

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 | UZK | A | 501 | - | - | 0/15/49/49 | 0/4/5/5 |
| 3 | SAM | A | 502 | - | - | 2/8/33/33 | 0/3/3/3 |

All (14) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 2 | A | 501 | UZK | C3-C4 | -7.20 | 1.32 | 1.40 |
| 2 | A | 501 | UZK | C22-N7 | -4.22 | 1.31 | 1.41 |
| 2 | A | 501 | UZK | C5-C4 | 4.06 | 1.46 | 1.39 |
| 2 | A | 501 | UZK | C8-N7 | 3.58 | 1.51 | 1.48 |
| 2 | A | 501 | UZK | C29-C27 | 3.47 | 1.60 | 1.49 |
| 2 | A | 501 | UZK | C28-C27 | 3.47 | 1.58 | 1.52 |
| 2 | A | 501 | UZK | C2-C3 | 3.04 | 1.44 | 1.39 |
| 2 | A | 501 | UZK | C25-C26 | -2.79 | 1.48 | 1.52 |
| 2 | A | 501 | UZK | C3-N11 | -2.44 | 1.38 | 1.42 |
| 2 | A | 501 | UZK | C27-C26 | 2.44 | 1.58 | 1.51 |
| 2 | A | 501 | UZK | C6-C5 | 2.36 | 1.43 | 1.38 |
| 2 | A | 501 | UZK | C2-C1 | 2.33 | 1.42 | 1.39 |
| 2 | A | 501 | UZK | O24-C22 | 2.14 | 1.26 | 1.22 |
| 2 | A | 501 | UZK | C1-C14 | 2.10 | 1.54 | 1.50 |

All (20) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 2 | A | 501 | UZK | C3-C4-N7 | 11.29 | 126.29 | 119.26 |
| 2 | A | 501 | UZK | C5-C4-N7 | -8.61 | 110.96 | 120.48 |
| 2 | A | 501 | UZK | O12-C10-N11 | -8.15 | 112.98 | 121.09 |
| 2 | A | 501 | UZK | O12-C10-C9 | 5.21 | 129.02 | 120.90 |
| 2 | A | 501 | UZK | C29-C27-C28 | 4.44 | 123.14 | 116.01 |
| 2 | A | 501 | UZK | O24-C22-N23 | -4.15 | 114.67 | 123.80 |
| 2 | A | 501 | UZK | C16-N15-C14 | -4.12 | 112.68 | 122.08 |
| 2 | A | 501 | UZK | C26-C25-N23 | 3.83 | 108.65 | 102.84 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 2 | A | 501 | UZK | C2-C3-C4 | 3.16 | 123.01 | 119.67 |
| 2 | A | 501 | UZK | O21-C14-C1 | -2.63 | 116.24 | 120.94 |
| 2 | A | 501 | UZK | C6-C5-C4 | -2.63 | 113.79 | 119.19 |
| 2 | A | 501 | UZK | C5-C6-C1 | 2.62 | 123.83 | 120.78 |
| 2 | A | 501 | UZK | C5-C4-C3 | 2.49 | 122.72 | 119.63 |
| 2 | A | 501 | UZK | C4-N7-C8 | -2.27 | 115.28 | 117.71 |
| 2 | A | 501 | UZK | O21-C14-N15 | 2.14 | 126.87 | 122.61 |
| 2 | A | 501 | UZK | C13-C8-N7 | 2.14 | 112.31 | 110.76 |
| 2 | A | 501 | UZK | C16-C17-C18 | -2.11 | 108.97 | 114.32 |
| 3 | A | 502 | SAM | C5-C6-N6 | 2.09 | 123.52 | 120.35 |
| 2 | A | 501 | UZK | C1-C2-C3 | -2.08 | 115.88 | 119.70 |
| 2 | A | 501 | UZK | C25-N23-C28 | -2.05 | 108.76 | 112.81 |

There are no chirality outliers.

All (2) torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 3 | A | 502 | SAM | CB-CG-SD-CE |
| 3 | A | 502 | SAM | CB-CG-SD-C5' |

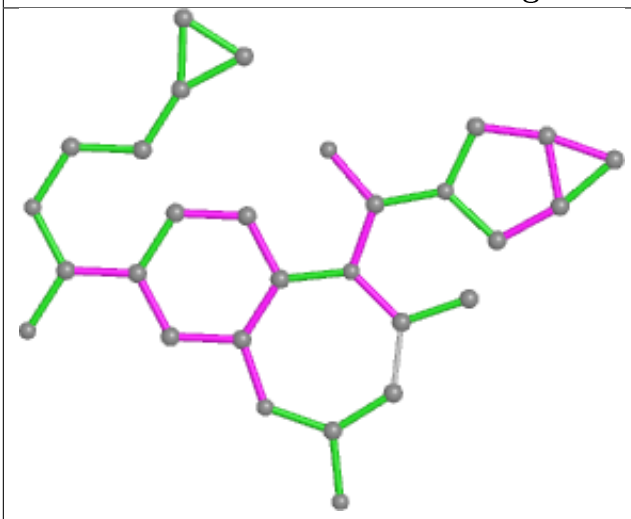
There are no ring outliers.

1 monomer is involved in 1 short contact:

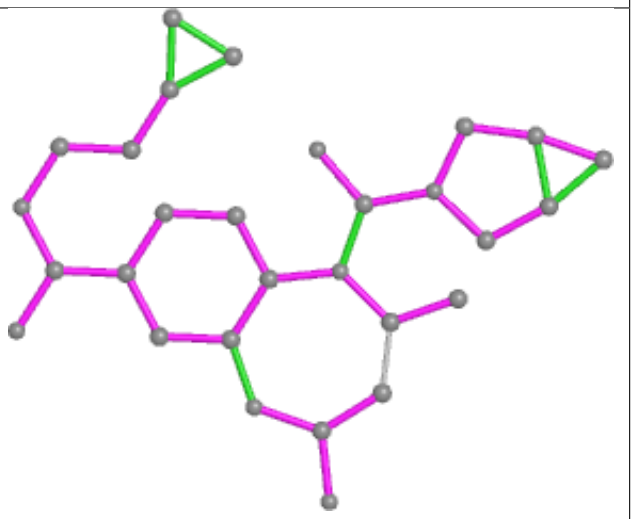
| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|---|---|---|---|---|---|
| 2 | A | 501 | UZK | 1 | 0 |

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.

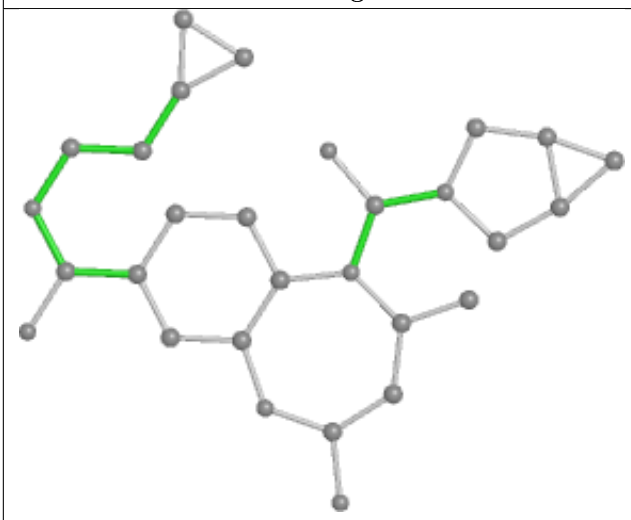
## Ligand UZK A 501



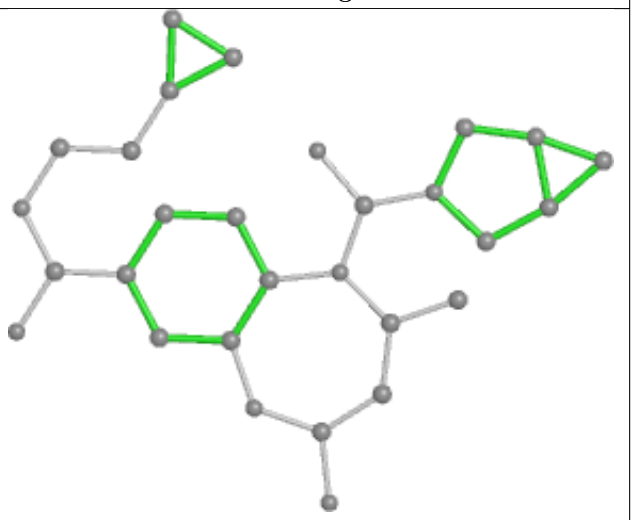
Bond lengths



Bond angles

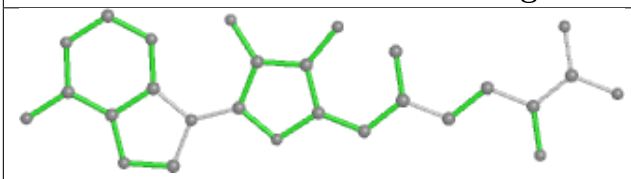


Torsions

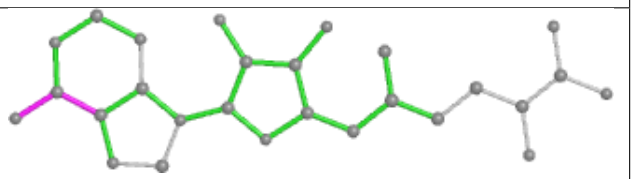


Rings

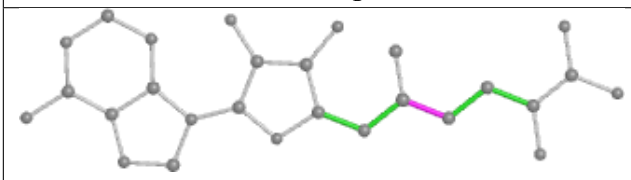
## Ligand SAM A 502



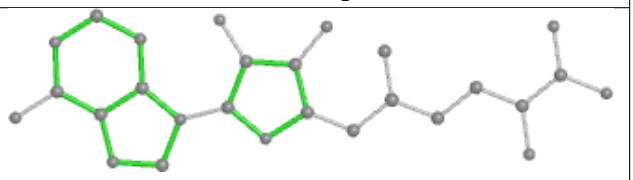
Bond lengths



Bond angles



Torsions



Rings

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

# 6 Fit of model and data

## 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

## 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

## 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

## 6.4 Ligands

EDS was not executed - this section is therefore empty.

## 6.5 Other polymers

EDS was not executed - this section is therefore empty.