wwPDB X-ray Structure Validation Summary Report (i)

Dec 5, 2022 – 05:22 pm GMT

PDB ID	:	7QEL
Title	:	Structure of the mouse 8-oxoguanine DNA Glycosylase mOGG1 in complex with ligand TH011247
Authors	:	Davies, J.R.; Scaletti, E.R.; Stenmark, P.
Deposited on	:	2021-12-03
Resolution	:	2.50 Å(reported)

This is a wwPDB X-ray Structure 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/XrayValidationReportHelp
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:

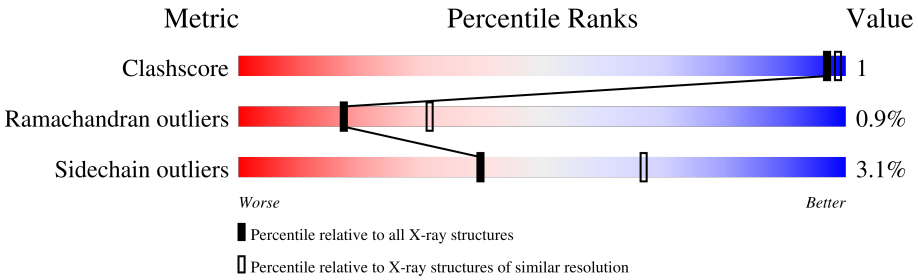
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 2.50 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	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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 failed to run properly.

Mol	Chain	Length	Quality of chain
1	AAA	318	95%
1	BBB	318	92%
1	CCC	318	88% 9%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7333 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.

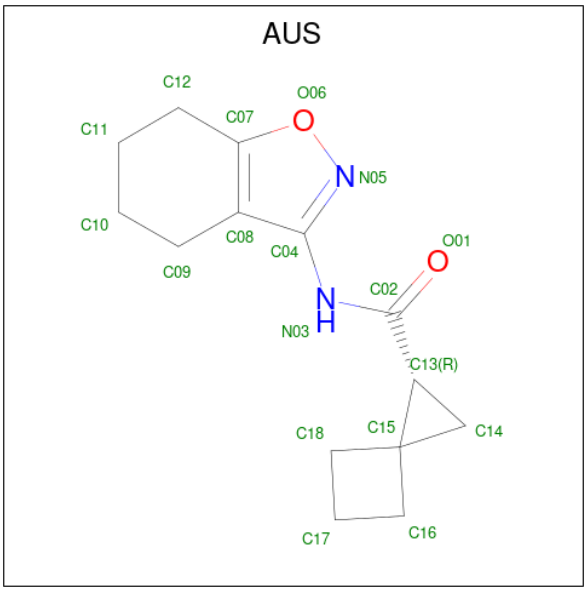
- Molecule 1 is a protein called N-glycosylase/DNA lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	312	Total	C	N	O	S	0	0	0
			2447	1560	443	433	11			
1	BBB	306	Total	C	N	O	S	0	0	0
			2340	1501	416	412	11			
1	CCC	311	Total	C	N	O	S	0	1	0
			2370	1514	423	422	11			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	8	GLY	-	expression tag	UNP O08760
AAA	10	HIS	SER	conflict	UNP O08760
BBB	8	GLY	-	expression tag	UNP O08760
BBB	10	HIS	SER	conflict	UNP O08760
CCC	8	GLY	-	expression tag	UNP O08760
CCC	10	HIS	SER	conflict	UNP O08760

- Molecule 2 is (2 {R})- {N}-(4,5,6,7-tetrahydro-1,2-benzoxazol-3-yl)spiro[2.3]hexane-2-carboxamide (three-letter code: AUS) (formula: $C_{14}H_{18}N_2O_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total C N O 18 14 2 2	0	0
2	BBB	1	Total C N O 18 14 2 2	0	0
2	CCC	1	Total C N O 18 14 2 2	0	0

- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	1	Total Ni 1 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	75	Total O 75 75	0	0
4	BBB	22	Total O 22 22	0	0
4	CCC	24	Total O 24 24	0	0

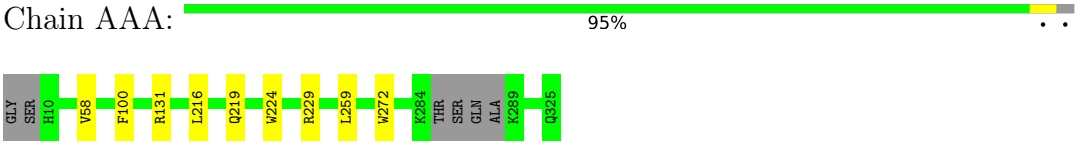
3 Residue-property plots (i)

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 failed to run properly.

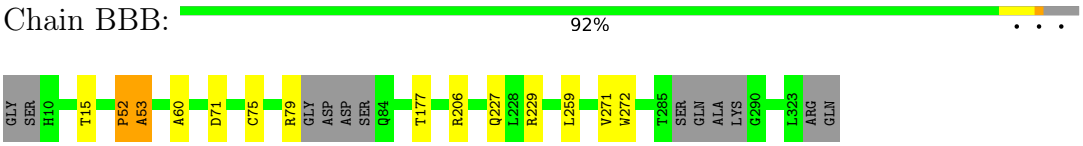
- Molecule 1: N-glycosylase/DNA lyase

Chain AAA:



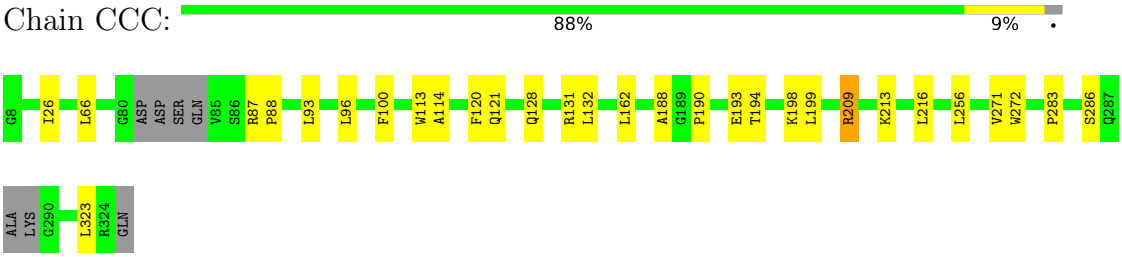
- Molecule 1: N-glycosylase/DNA lyase

Chain BBB:



- Molecule 1: N-glycosylase/DNA lyase

Chain CCC:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	81.35Å 81.40Å 168.41Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.73 – 2.50	Depositor
% Data completeness (in resolution range)	100.0 (40.73-2.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.98 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.252 , 0.295	Depositor
Wilson B-factor (Å^2)	51.6	Xtriage
Anisotropy	0.067	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.008 for k,h,-l	Xtriage
Total number of atoms	7333	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.41% 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: AUS, NI

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	AAA	0.64	0/2517	0.70	0/3433
1	BBB	0.66	0/2407	0.71	0/3291
1	CCC	0.66	0/2437	0.71	0/3331
All	All	0.66	0/7361	0.70	0/10055

There are no bond length outliers.

There are no bond angle outliers.

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	AAA	2447	0	2357	3	0
1	BBB	2340	0	2218	2	0
1	CCC	2370	0	2222	8	0
2	AAA	18	0	0	0	0
2	BBB	18	0	0	0	0
2	CCC	18	0	0	0	0
3	AAA	1	0	0	0	0
4	AAA	75	0	0	0	0
4	BBB	22	0	0	0	0
4	CCC	24	0	0	0	0
All	All	7333	0	6797	13	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 1.

The worst 5 of 13 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:CCC:132:LEU:HD21	1:CCC:256:LEU:HG	1.69	0.73
1:CCC:193:GLU:HB2	1:CCC:209:ARG:HH12	1.59	0.67
1:BBB:229:ARG:HG3	1:BBB:259:LEU:HA	1.90	0.53
1:AAA:100:PHE:O	1:AAA:131:ARG:HD3	2.09	0.52
1:CCC:113:TRP:HE3	1:CCC:120:PHE:CZ	2.31	0.48

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	AAA	308/318 (97%)	294 (96%)	14 (4%)	0	100	100
1	BBB	300/318 (94%)	277 (92%)	19 (6%)	4 (1%)	12	21
1	CCC	306/318 (96%)	283 (92%)	19 (6%)	4 (1%)	12	21
All	All	914/954 (96%)	854 (93%)	52 (6%)	8 (1%)	17	31

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	CCC	199	LEU
1	BBB	60	ALA
1	BBB	71	ASP
1	BBB	53	ALA
1	CCC	286	SER

5.3.2 Protein sidechains

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	AAA	251/266 (94%)	248 (99%)	3 (1%)	71 88
1	BBB	232/266 (87%)	224 (97%)	8 (3%)	37 63
1	CCC	233/266 (88%)	222 (95%)	11 (5%)	26 49
All	All	716/798 (90%)	694 (97%)	22 (3%)	40 67

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	CCC	162	LEU
1	CCC	209	ARG
1	CCC	198	LYS
1	CCC	216	LEU
1	BBB	177	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 (i)

Of 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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	AUS	BBB	401	-	15,21,21	2.64	4 (26%)	11,32,32	1.67	1 (9%)
2	AUS	CCC	401	-	15,21,21	2.66	4 (26%)	11,32,32	1.46	1 (9%)
2	AUS	AAA	401	-	15,21,21	2.52	4 (26%)	11,32,32	1.59	1 (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	AUS	BBB	401	-	-	0/6/31/31	0/4/4/4
2	AUS	CCC	401	-	-	0/6/31/31	0/4/4/4
2	AUS	AAA	401	-	-	0/6/31/31	0/4/4/4

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	CCC	401	AUS	C12-C07	7.04	1.55	1.50
2	BBB	401	AUS	C12-C07	6.77	1.55	1.50
2	AAA	401	AUS	C12-C07	6.44	1.55	1.50
2	BBB	401	AUS	C02-N03	5.17	1.47	1.35
2	CCC	401	AUS	C02-N03	5.12	1.47	1.35

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	BBB	401	AUS	C13-C02-N03	4.66	119.23	114.38
2	AAA	401	AUS	C13-C02-N03	4.19	118.74	114.38
2	CCC	401	AUS	C13-C02-N03	4.05	118.59	114.38

There are no chirality outliers.

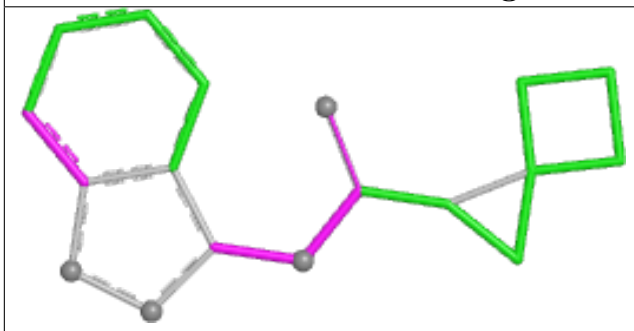
There are no torsion outliers.

There are no ring outliers.

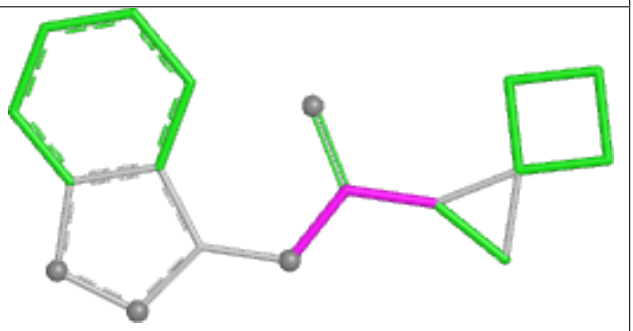
No monomer is involved in short contacts.

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.

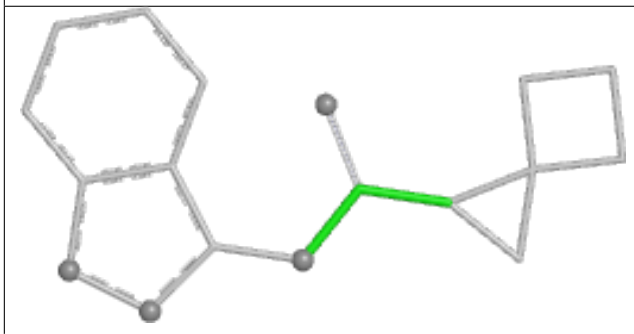
Ligand AUS BBB 401



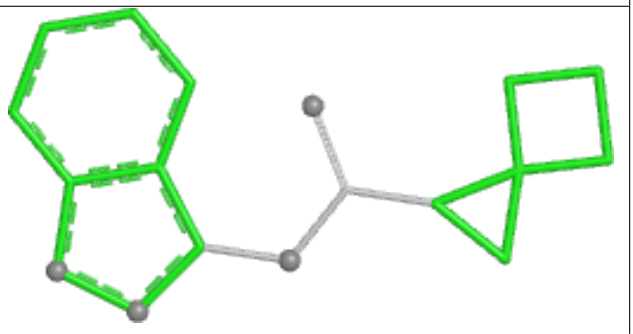
Bond lengths



Bond angles

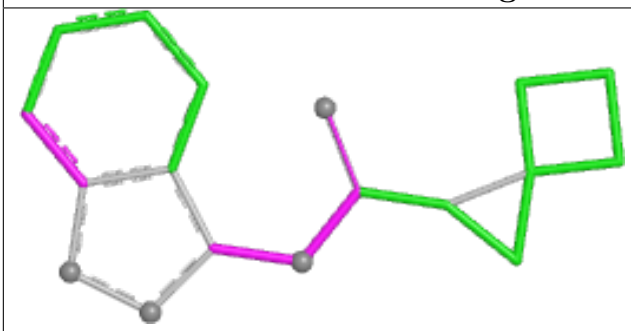


Torsions

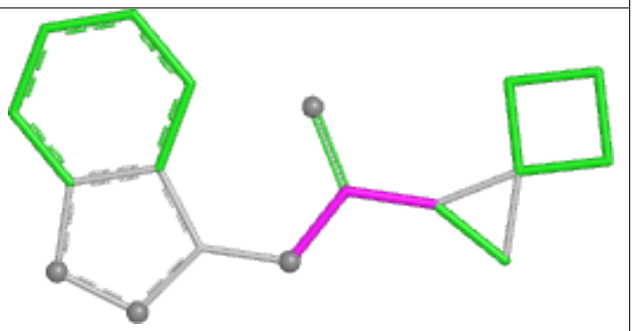


Rings

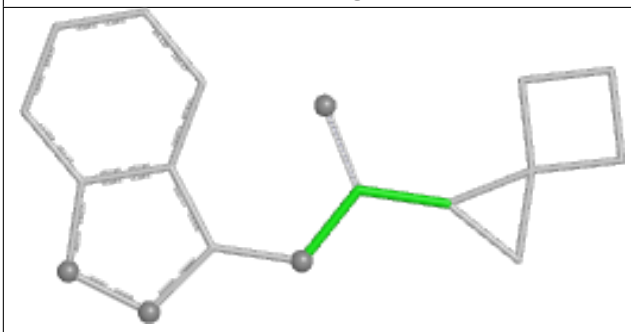
Ligand AUS CCC 401



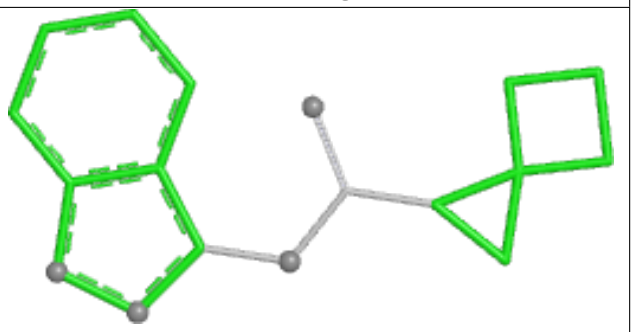
Bond lengths



Bond angles

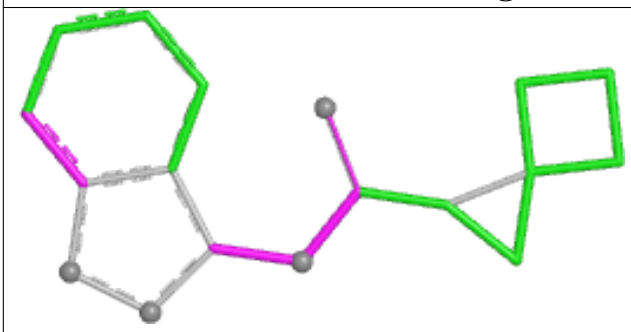


Torsions

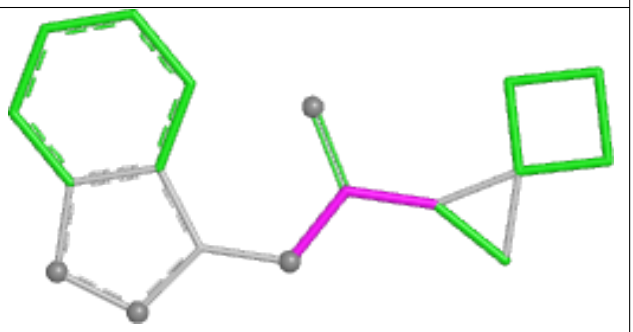


Rings

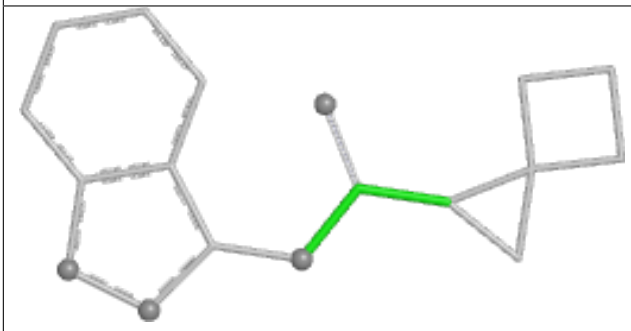
Ligand AUS AAA 401



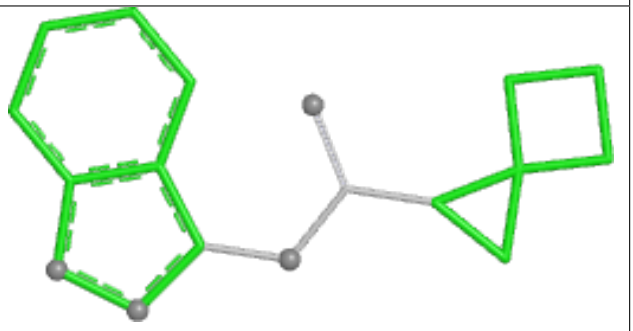
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers

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

6.1 Protein, DNA and RNA chains

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

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

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

6.3 Carbohydrates

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

6.4 Ligands

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

6.5 Other polymers

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