



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2020 – 01:34 PM BST

PDB ID : 6U91  
Title : Crystal structure of DNMT3B(Q772R)-DNMT3L in complex with CpGpT DNA  
Authors : Gao, L.; Song, J.  
Deposited on : 2019-09-06  
Resolution : 3.00 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 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) : 1.13  
EDS : 2.13.1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

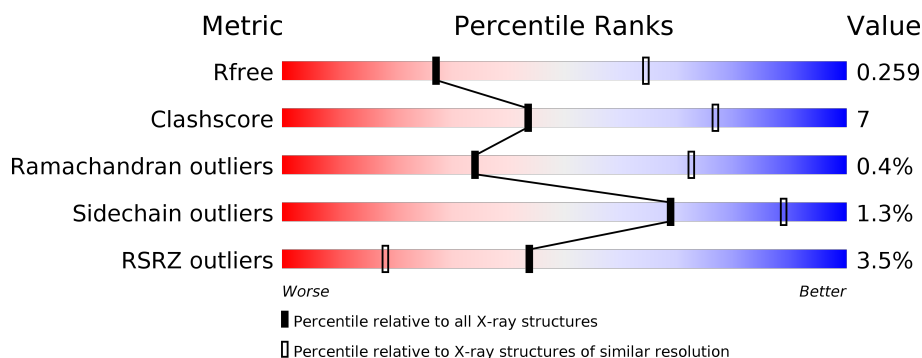
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
$R_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	291	<div> <div></div> <div>83%16%.</div> </div>
1	D	291	<div> <div></div> <div>82%17%.</div> </div>
2	B	209	<div> <div>9%</div> <div>76%16%8%</div> </div>
2	C	209	<div> <div>8%</div> <div>74%18%8%</div> </div>
3	E	25	<div> <div></div> <div>56%40%. .</div> </div>
3	F	25	<div> <div></div> <div>56%40%. .</div> </div>

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 8622 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 DNA (cytosine-5)-methyltransferase 3B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	291	Total	C	N	O	S	0	1	0
			2305	1478	406	409	12			
1	D	291	Total	C	N	O	S	0	1	0
			2314	1486	406	410	12			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	772	ARG	GLN	engineered mutation	UNP Q9UBC3
D	772	ARG	GLN	engineered mutation	UNP Q9UBC3

- Molecule 2 is a protein called DNA (cytosine-5)-methyltransferase 3-like.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	193	Total	C	N	O	S	0	0	0
			1457	952	247	255	3			
2	C	192	Total	C	N	O	S	0	0	0
			1439	935	246	254	4			

- Molecule 3 is a DNA chain called CpGpT DNA (25-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	25	Total	C	N	O	P	0	0	0
			510	245	93	148	24			
3	F	25	Total	C	N	O	P	0	0	0
			510	245	93	148	24			

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

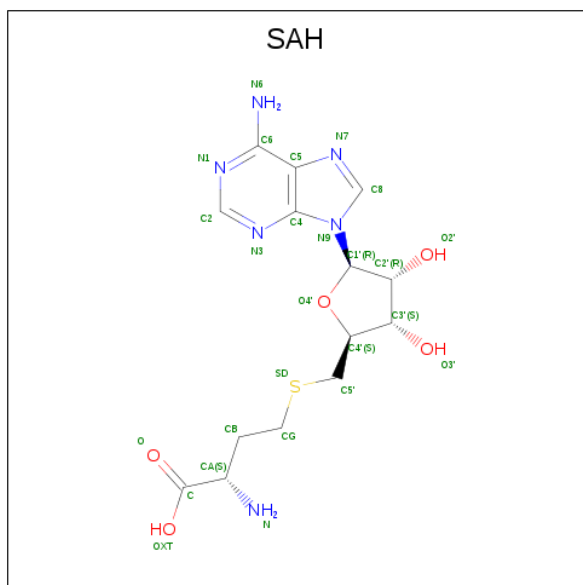
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Mg	0	0
			2	2		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	1	Total	Mg	0	0
			1	1		

- Molecule 5 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula:  $C_{14}H_{20}N_6O_5S$ ) (labeled as "Ligand of Interest" by author).



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

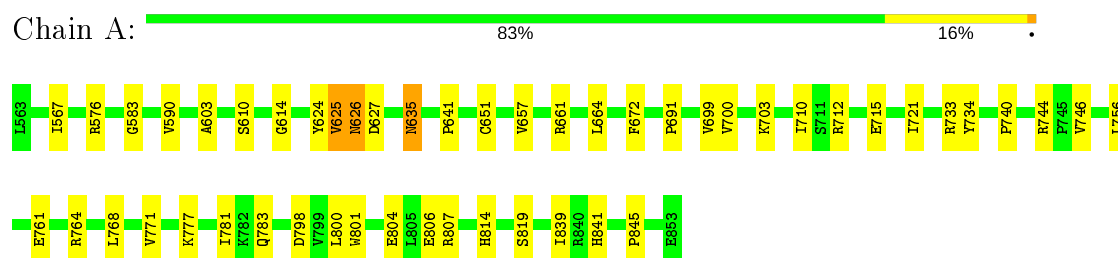
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	16	Total	O	0	0
			16	16		
6	B	1	Total	O	0	0
			1	1		
6	C	1	Total	O	0	0
			1	1		
6	D	12	Total	O	0	0
			12	12		
6	E	1	Total	O	0	0
			1	1		
6	F	1	Total	O	0	0
			1	1		

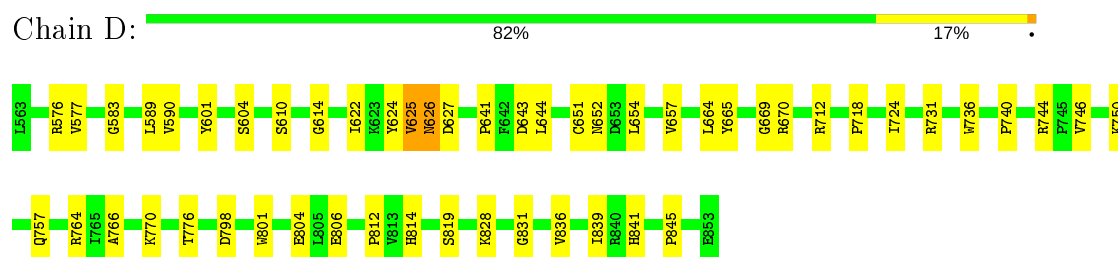
### 3 Residue-property plots [i](#)

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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.

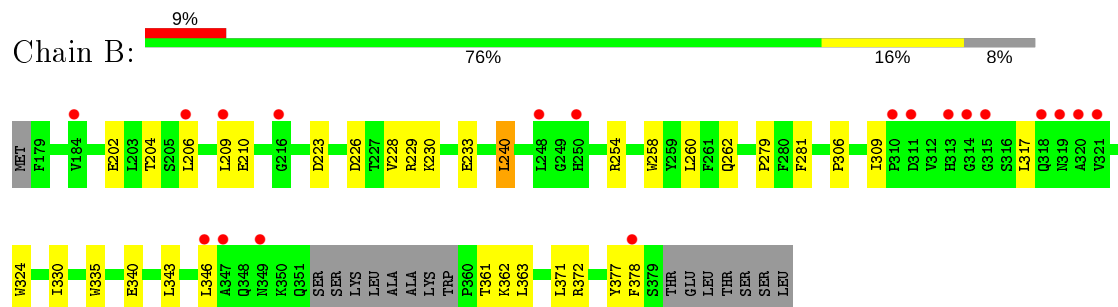
- Molecule 1: DNA (cytosine-5)-methyltransferase 3B



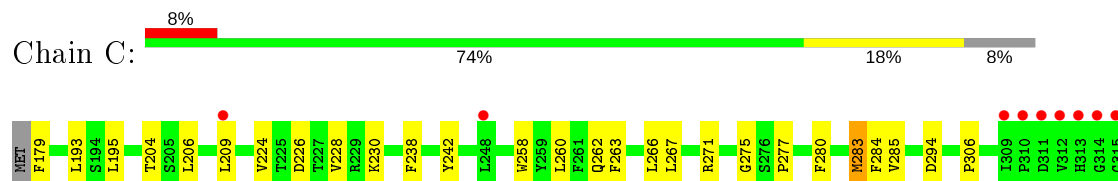
- Molecule 1: DNA (cytosine-5)-methyltransferase 3B

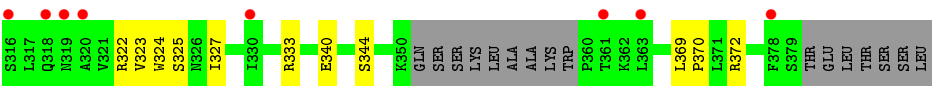


- Molecule 2: DNA (cytosine-5)-methyltransferase 3-like



- Molecule 2: DNA (cytosine-5)-methyltransferase 3-like

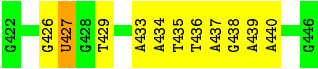




● Molecule 3: CpGpT DNA (25-MER)



● Molecule 3: CpGpT DNA (25-MER)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	190.34Å 190.34Å 49.64Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.59 – 3.00 47.59 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (47.59-3.00) 99.7 (47.59-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.24 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.12_2829, PHENIX 1.12_2829	Depositor
R, $R_{free}$	0.215 , 0.259 0.214 , 0.259	Depositor DCC
$R_{free}$ test set	2018 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	65.3	Xtriage
Anisotropy	0.490	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 34.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.036 for -h,-k,l 0.440 for h,-h-k,-l 0.032 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	8622	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.48% 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.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, SAH, PYO

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	0.26	0/2359	0.44	0/3193
1	D	0.26	0/2369	0.43	0/3206
2	B	0.26	0/1505	0.49	0/2063
2	C	0.26	0/1486	0.44	0/2039
3	E	0.50	0/550	0.90	0/846
3	F	0.51	0/550	0.91	0/846
All	All	0.30	0/8819	0.53	0/12193

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
1	D	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	625	VAL	Peptide
1	D	625	VAL	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	2305	0	2282	30	0
1	D	2314	0	2293	34	0
2	B	1457	0	1307	22	0
2	C	1439	0	1273	20	0
3	E	510	0	278	8	0
3	F	510	0	279	8	0
4	A	2	0	0	0	0
4	D	1	0	0	0	0
5	A	26	0	19	0	0
5	D	26	0	19	0	0
6	A	16	0	0	0	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
6	D	12	0	0	0	0
6	E	1	0	0	0	0
6	F	1	0	0	0	0
All	All	8622	0	7750	110	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 7.

All (110) 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)
2:B:343:LEU:HA	2:B:346:LEU:HG	1.55	0.88
2:B:240:LEU:HD13	2:B:371:LEU:HD11	1.64	0.77
2:C:294:ASP:OD1	1:D:670:ARG:NH1	2.22	0.71
2:B:340:GLU:HA	2:B:343:LEU:HG	1.71	0.71
2:B:330:ILE:HA	2:B:335:TRP:HH2	1.57	0.69
1:A:744:ARG:NH2	1:A:841:HIS:O	2.29	0.65
3:E:433:DA:H2"	3:E:434:DA:H5"	1.79	0.63
2:B:240:LEU:HD12	2:B:281:PHE:HB2	1.83	0.61
2:C:228:VAL:HG12	2:C:230:LYS:H	1.66	0.61
3:F:436:DT:H2"	3:F:437:DA:C8	2.36	0.60
1:D:664:LEU:HA	1:D:669:GLY:HA3	1.84	0.60

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:206:LEU:O	2:B:372:ARG:NH2	2.36	0.59
1:D:744:ARG:NH2	1:D:841:HIS:O	2.35	0.59
1:A:715:GLU:OE2	2:B:229:ARG:NH1	2.37	0.58
1:D:731:ARG:NH2	1:D:831:GLY:O	2.33	0.58
1:D:590:VAL:HG12	1:D:836:VAL:HG13	1.86	0.58
2:C:195:LEU:HD21	2:C:266:LEU:HD12	1.84	0.58
1:D:576:ARG:HH21	1:D:641:PRO:HB2	1.67	0.58
1:D:625:VAL:O	1:D:626:ASN:O	2.21	0.57
1:A:781:ILE:O	1:A:783:GLN:NE2	2.34	0.57
1:A:761:GLU:OE1	1:A:807:ARG:NH2	2.34	0.56
1:A:764:ARG:NH2	1:A:798:ASP:OD2	2.39	0.56
2:B:306:PRO:HB3	2:B:324:TRP:CE2	2.41	0.55
2:C:285:VAL:HG22	2:C:323:VAL:HG22	1.88	0.55
2:B:228:VAL:HG12	2:B:229:ARG:H	1.71	0.54
1:D:576:ARG:O	1:D:643:ASP:N	2.39	0.54
3:F:433:DA:H2"	3:F:434:DA:H5"	1.89	0.53
2:C:340:GLU:O	2:C:344:SER:N	2.41	0.53
1:A:819:SER:HA	1:D:801:TRP:CE2	2.44	0.53
1:A:625:VAL:O	1:A:626:ASN:O	2.27	0.52
3:E:436:DT:H2"	3:E:437:DA:C8	2.45	0.52
1:D:590:VAL:HG21	1:D:839:ILE:HG22	1.92	0.52
1:A:651:CYS:N	3:F:427:PYO:HC5	2.26	0.51
1:D:651:CYS:N	3:E:427:PYO:HC5	2.25	0.51
1:D:625:VAL:HG23	1:D:625:VAL:O	2.11	0.51
1:D:665:TYR:O	1:D:670:ARG:NH2	2.43	0.51
2:B:230:LYS:HA	2:B:233:GLU:HB2	1.92	0.51
1:A:777:LYS:HG3	3:F:429:DT:H71	1.93	0.51
2:C:204:THR:HG22	2:C:209:LEU:HB2	1.92	0.51
1:D:577:VAL:HG12	1:D:644:LEU:HB3	1.91	0.51
1:D:601:TYR:HD1	1:D:622:ILE:HG23	1.76	0.50
2:B:223:ASP:OD1	2:B:254:ARG:NH1	2.44	0.50
2:C:275:GLY:O	2:C:277:PRO:HD3	2.11	0.50
2:B:204:THR:HG22	2:B:209:LEU:HB2	1.94	0.50
2:C:306:PRO:HB3	2:C:324:TRP:CE2	2.48	0.49
1:A:576:ARG:HH21	1:A:641:PRO:HB2	1.78	0.49
2:B:258:TRP:CE2	2:B:262:GLN:HG3	2.47	0.49
1:D:614:GLY:HA3	1:D:624:TYR:OH	2.13	0.48
3:E:434:DA:H2"	3:E:435:DT:H5"	1.96	0.48
2:B:254:ARG:HD3	2:B:258:TRP:CE3	2.49	0.47
1:D:718:PRO:HB3	1:D:736:TRP:CE2	2.49	0.47
2:B:210:GLU:HG3	2:B:378:PHE:CE2	2.48	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:664:LEU:HD11	1:A:710:ILE:HG13	1.95	0.47
3:E:423:DC:H2"	3:E:424:DA:C8	2.50	0.47
1:A:721:ILE:HG23	1:A:746:VAL:HG21	1.95	0.46
2:C:242:TYR:HD1	2:C:283:MET:HG2	1.80	0.46
2:C:267:LEU:O	2:C:271:ARG:HG3	2.15	0.46
1:D:724:ILE:HD11	1:D:770:LYS:HG3	1.96	0.46
1:D:589:LEU:HD21	1:D:812:PRO:HD3	1.97	0.46
1:A:583:GLY:O	1:A:610:SER:HB3	2.15	0.46
1:A:614:GLY:HA3	1:A:624:TYR:OH	2.16	0.46
1:D:746:VAL:HG12	1:D:841:HIS:NE2	2.30	0.46
3:E:445:DT:H2"	3:E:446:DG:C8	2.51	0.46
2:C:193:LEU:HB2	2:C:238:PHE:CE1	2.52	0.45
2:C:258:TRP:CE2	2:C:262:GLN:HG3	2.51	0.45
1:A:699:VAL:HA	1:A:733:ARG:HD2	1.98	0.45
1:D:764:ARG:NH2	1:D:798:ASP:OD2	2.49	0.45
1:A:801:TRP:CE2	1:D:819:SER:HA	2.52	0.45
1:A:764:ARG:HD3	1:A:804:GLU:OE2	2.17	0.45
1:D:583:GLY:O	1:D:610:SER:HB3	2.17	0.44
1:D:652:ASN:N	1:D:652:ASN:OD1	2.50	0.44
2:C:226:ASP:HA	1:D:712:ARG:HE	1.82	0.44
1:D:757:GLN:HG2	1:D:766:ALA:HB3	1.99	0.44
1:A:635:ASN:OD1	1:A:635:ASN:N	2.50	0.44
1:D:740:PRO:O	1:D:845:PRO:HG3	2.16	0.44
1:A:700:VAL:HA	1:A:734:TYR:CE2	2.52	0.44
2:B:202:GLU:N	2:B:202:GLU:OE1	2.48	0.44
2:B:317:LEU:HD13	2:B:346:LEU:HD22	2.00	0.44
2:C:206:LEU:O	2:C:372:ARG:NH1	2.49	0.44
1:A:740:PRO:O	1:A:845:PRO:HG3	2.18	0.43
1:D:651:CYS:H	3:E:427:PYO:HC5	1.83	0.43
1:A:603:ALA:O	1:A:625:VAL:HG22	2.17	0.43
1:A:661:ARG:NH1	1:A:703:LYS:HG3	2.33	0.43
2:C:263:PHE:CE2	2:C:284:PHE:HB2	2.53	0.43
2:C:224:VAL:HG11	2:C:266:LEU:HD11	2.01	0.43
2:B:309:ILE:HG21	2:B:363:LEU:O	2.19	0.43
1:D:654:LEU:HD21	1:D:664:LEU:HB3	2.01	0.43
1:D:776:THR:HG22	1:D:828:LYS:HA	2.01	0.42
1:A:657:VAL:HB	3:F:426:DG:H1'	2.00	0.42
1:A:664:LEU:HD21	1:A:672:PHE:CD2	2.55	0.42
1:A:567:ILE:HD11	1:A:691:PRO:HG3	2.00	0.42
2:B:240:LEU:HA	2:B:281:PHE:O	2.19	0.42
2:B:361:THR:OG1	2:B:362:LYS:N	2.53	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:756:LEU:HD11	1:A:800:LEU:HD21	2.01	0.42
2:C:369:LEU:N	2:C:370:PRO:HD2	2.35	0.42
1:D:750:LYS:HB2	1:D:750:LYS:HE2	1.88	0.42
3:F:439:DA:H2''	3:F:440:DA:C8	2.55	0.41
1:D:806:GLU:OE2	1:D:814:HIS:N	2.51	0.41
2:B:279:PRO:HB2	2:B:281:PHE:CE2	2.55	0.41
2:C:325:SER:OG	2:C:327:ILE:HG12	2.21	0.41
1:A:712:ARG:HE	2:B:226:ASP:HA	1.84	0.41
3:F:438:DG:H2''	3:F:439:DA:C8	2.56	0.41
1:D:764:ARG:HD3	1:D:804:GLU:OE2	2.21	0.41
1:D:657:VAL:HB	3:E:426:DG:H1'	2.01	0.41
1:A:590:VAL:HG21	1:A:839:ILE:HB	2.03	0.41
2:C:271:ARG:HG2	2:C:280:PHE:CE2	2.56	0.40
2:C:322:ARG:HH12	2:C:340:GLU:HG2	1.86	0.40
1:A:806:GLU:OE2	1:A:814:HIS:N	2.53	0.40
1:A:771:VAL:HG13	1:A:783:GLN:HE22	1.87	0.40
3:F:434:DA:H1'	3:F:435:DT:H5'	2.03	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	290/291 (100%)	279 (96%)	9 (3%)	2 (1%)	22	60
1	D	290/291 (100%)	284 (98%)	4 (1%)	2 (1%)	22	60
2	B	189/209 (90%)	179 (95%)	10 (5%)	0	100	100
2	C	188/209 (90%)	179 (95%)	9 (5%)	0	100	100
All	All	957/1000 (96%)	921 (96%)	32 (3%)	4 (0%)	34	72

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	626	ASN
1	D	626	ASN
1	A	627	ASP
1	D	627	ASP

### 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	A	242/252 (96%)	240 (99%)	2 (1%)	81	93
1	D	243/252 (96%)	242 (100%)	1 (0%)	91	97
2	B	137/191 (72%)	134 (98%)	3 (2%)	52	81
2	C	135/191 (71%)	131 (97%)	4 (3%)	41	75
All	All	757/886 (85%)	747 (99%)	10 (1%)	69	89

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	635	ASN
1	A	768	LEU
2	B	240	LEU
2	B	260	LEU
2	B	377	TYR
2	C	179	PHE
2	C	260	LEU
2	C	283	MET
2	C	333	ARG
1	D	604	SER

Some 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 ⓘ

2 non-standard protein/DNA/RNA residues 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	PYO	E	427	1,3	14,20,21	5.70	12 (85%)	16,28,31	3.48	4 (25%)
3	PYO	F	427	1,3	14,20,21	3.66	5 (35%)	16,28,31	3.14	3 (18%)

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	PYO	E	427	1,3	-	1/5/25/26	0/2/2/2
3	PYO	F	427	1,3	-	2/5/25/26	0/2/2/2

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	427	PYO	C2'-C3'	-10.19	1.25	1.53
3	E	427	PYO	C6-N1	9.14	1.47	1.35
3	F	427	PYO	C6-N1	8.88	1.46	1.35
3	E	427	PYO	O4'-C1'	6.69	1.50	1.41
3	F	427	PYO	C2-N3	6.40	1.50	1.38
3	E	427	PYO	O4'-C4'	-6.24	1.31	1.45
3	E	427	PYO	C2-N3	6.20	1.50	1.38
3	E	427	PYO	C3'-C4'	5.51	1.67	1.53
3	F	427	PYO	C6-C5	5.16	1.50	1.38
3	E	427	PYO	C6-C5	5.09	1.49	1.38
3	F	427	PYO	C4-N3	4.80	1.44	1.34
3	E	427	PYO	C4-N3	4.79	1.44	1.34
3	E	427	PYO	C2'-C1'	4.46	1.60	1.53
3	E	427	PYO	O2'-C2'	4.41	1.53	1.43
3	E	427	PYO	C5-C4	4.15	1.49	1.37
3	F	427	PYO	C5-C4	4.06	1.49	1.37

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	427	PYO	O3'-C3'	3.09	1.50	1.43

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	427	PYO	C4-N3-C2	12.81	120.74	114.42
3	F	427	PYO	C4-N3-C2	11.88	120.28	114.42
3	E	427	PYO	C3'-C2'-C1'	3.45	106.17	100.98
3	E	427	PYO	C5-C4-N3	-2.55	119.27	123.43
3	E	427	PYO	C6-C5-C4	2.38	120.30	117.27
3	F	427	PYO	C6-C5-C4	2.37	120.29	117.27
3	F	427	PYO	C5-C4-N3	-2.30	119.67	123.43

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	427	PYO	O4'-C1'-N1-C6
3	F	427	PYO	O4'-C1'-N1-C6
3	F	427	PYO	C2'-C1'-N1-C6

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	427	PYO	2	0
3	F	427	PYO	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

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
5	SAH	A	903	-	21,28,28	1.21	2 (9%)	20,40,40	1.70	2 (10%)
5	SAH	D	902	-	21,28,28	1.21	2 (9%)	20,40,40	1.74	2 (10%)

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
5	SAH	A	903	-	-	2/7/31/31	0/3/3/3
5	SAH	D	902	-	-	2/7/31/31	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	902	SAH	C2-N3	4.03	1.38	1.32
5	A	903	SAH	C2-N3	4.03	1.38	1.32
5	A	903	SAH	C2-N1	2.51	1.38	1.33
5	D	902	SAH	C2-N1	2.41	1.38	1.33

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	902	SAH	N3-C2-N1	-5.64	119.86	128.68
5	A	903	SAH	N3-C2-N1	-5.42	120.21	128.68
5	D	902	SAH	C5'-SD-CG	-4.01	90.24	102.27
5	A	903	SAH	C5'-SD-CG	-3.97	90.36	102.27

There are no chirality outliers.

All (4) torsion outliers are listed below:

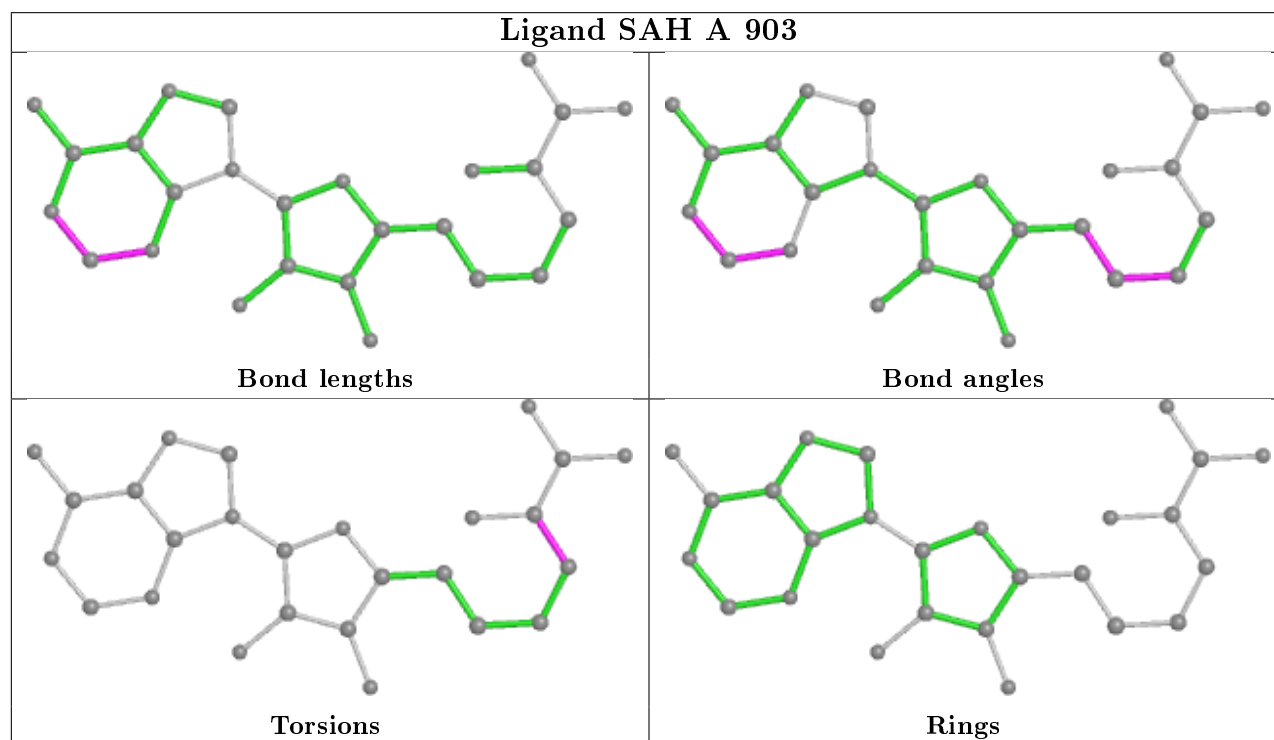
Mol	Chain	Res	Type	Atoms
5	A	903	SAH	N-CA-CB-CG
5	A	903	SAH	C-CA-CB-CG
5	D	902	SAH	N-CA-CB-CG
5	D	902	SAH	C-CA-CB-CG

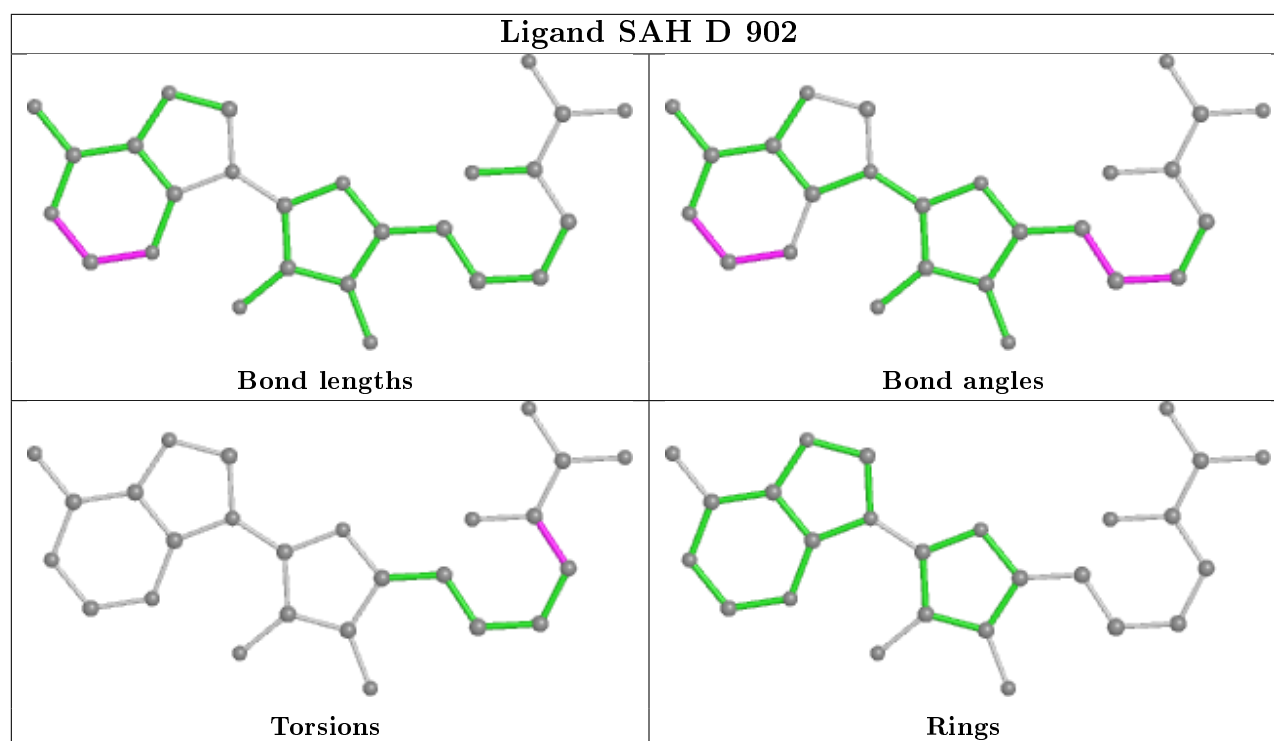


There are no ring outliers.

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.





## 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å <sup>2</sup> )	Q<0.9
1	A	291/291 (100%)	-0.01	0	100 100	26, 51, 93, 112	0
1	D	291/291 (100%)	0.00	0	100 100	28, 52, 91, 115	0
2	B	193/209 (92%)	0.49	19 (9%)	7 2	59, 92, 174, 207	0
2	C	192/209 (91%)	0.54	17 (8%)	9 3	59, 95, 172, 207	0
3	E	24/25 (96%)	-0.40	0	100 100	67, 105, 140, 148	0
3	F	24/25 (96%)	-0.40	0	100 100	64, 106, 136, 143	0
All	All	1015/1050 (96%)	0.17	36 (3%)	44 18	26, 67, 153, 207	0

All (36) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	310	PRO	12.0
2	C	315	GLY	9.0
2	C	319	ASN	8.6
2	C	314	GLY	7.7
2	B	313	HIS	7.7
2	B	314	GLY	7.4
2	C	313	HIS	6.8
2	C	320	ALA	6.7
2	C	318	GLN	6.1
2	B	310	PRO	4.8
2	C	361	THR	4.6
2	B	250	HIS	4.6
2	B	315	GLY	4.2
2	B	349	ASN	4.0
2	B	319	ASN	4.0
2	C	316	SER	3.9
2	B	318	GLN	3.9
2	B	209	LEU	3.5
2	B	347	ALA	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	206	LEU	3.3
2	C	311	ASP	3.2
2	B	216	GLY	2.9
2	B	321	VAL	2.9
2	C	309	ILE	2.6
2	B	311	ASP	2.6
2	C	209	LEU	2.6
2	C	363	LEU	2.4
2	B	320	ALA	2.3
2	C	330	ILE	2.3
2	C	312	VAL	2.2
2	C	378	PHE	2.2
2	B	378	PHE	2.2
2	C	248	LEU	2.2
2	B	346	LEU	2.1
2	B	248	LEU	2.1
2	B	184	VAL	2.0

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	PYO	E	427	19/20	0.98	0.17	38,57,66,69	0
3	PYO	F	427	19/20	0.98	0.16	40,50,61,62	0

## 6.3 Carbohydrates ⓘ

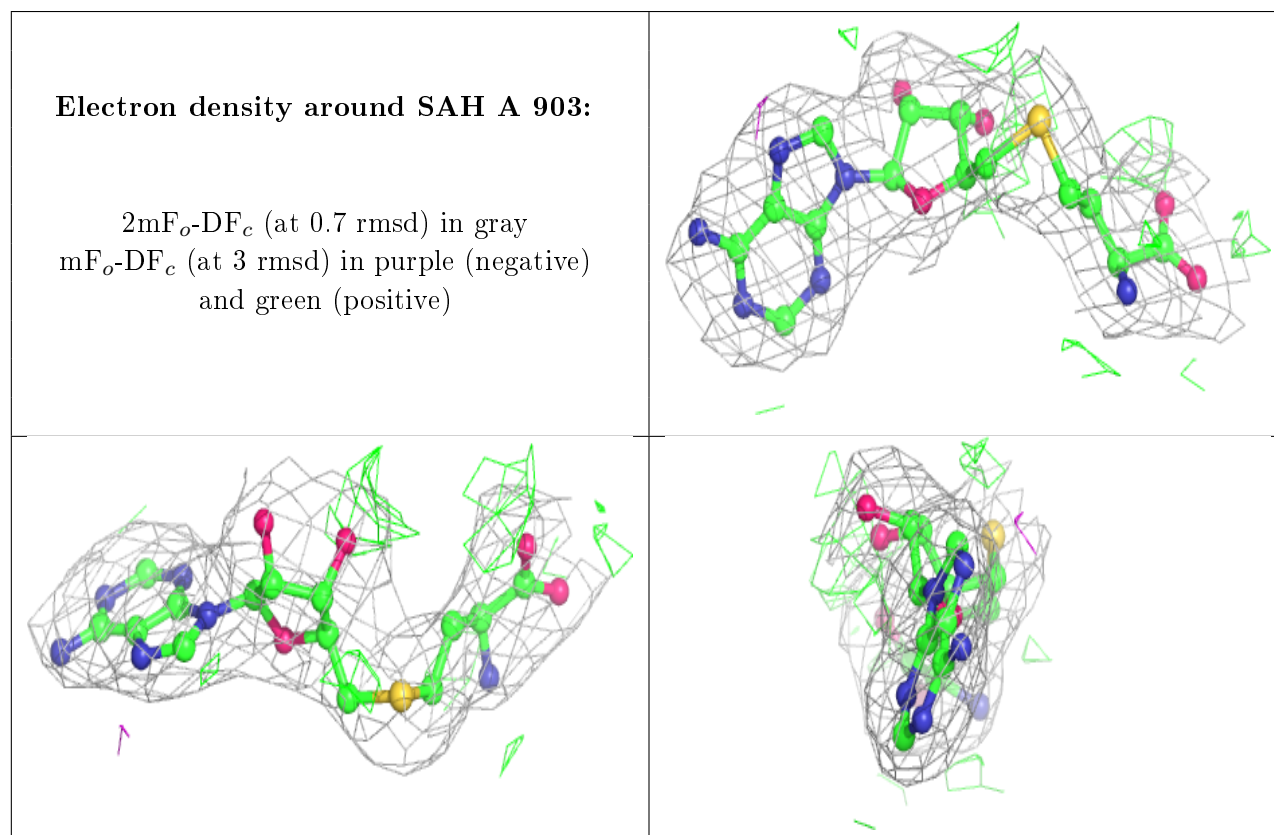
There are no monosaccharides in this entry.

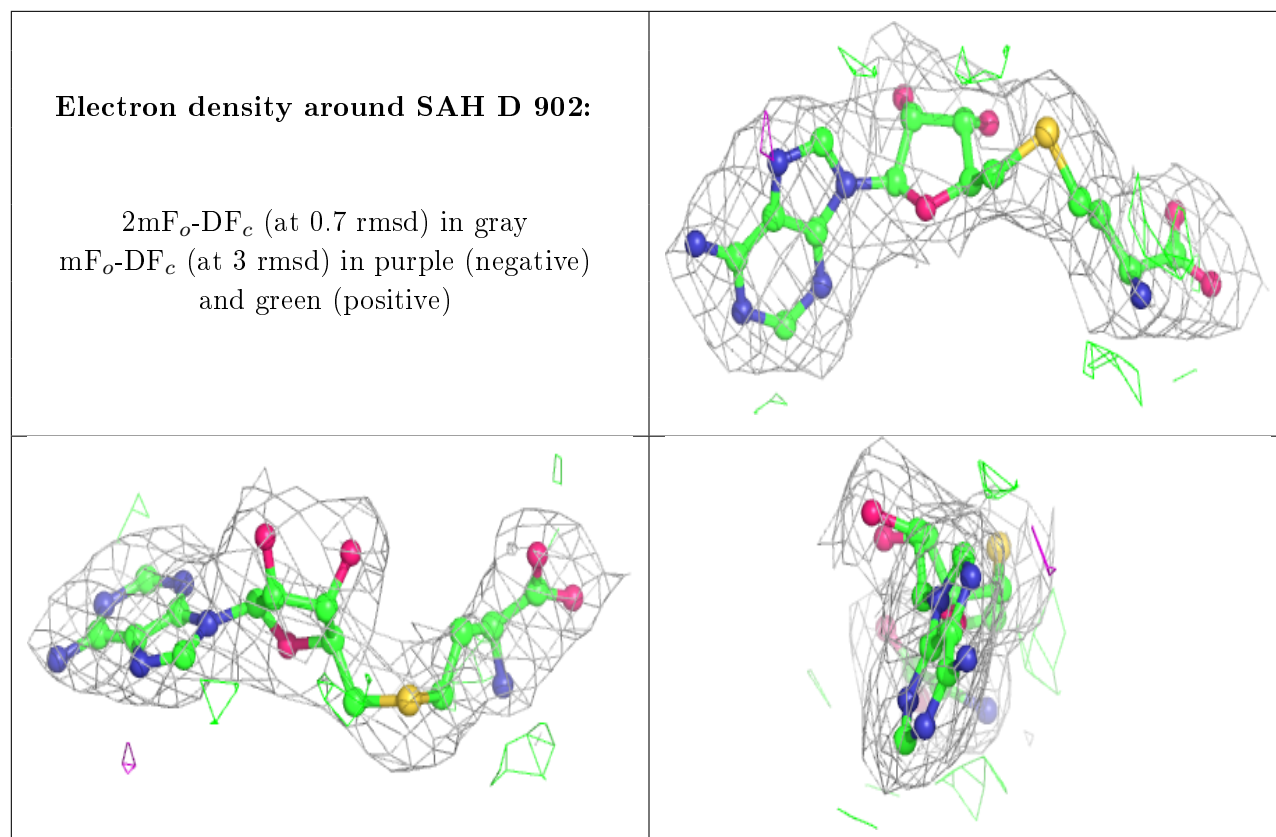
## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	MG	A	901	1/1	0.94	0.16	37,37,37,37	0
4	MG	D	901	1/1	0.94	0.23	59,59,59,59	0
5	SAH	A	903	26/26	0.97	0.19	24,41,60,66	0
4	MG	A	902	1/1	0.98	0.18	35,35,35,35	0
5	SAH	D	902	26/26	0.98	0.18	28,42,57,58	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.





## 6.5 Other polymers [i](#)

There are no such residues in this entry.