Full wwPDB X-ray Structure Validation Report ⓘ

Oct 10, 2023 – 11:22 PM EDT

PDB ID	:	6W8J
Title	:	Structure of DNMT3A (R882H) in complex with CAG DNA
Authors	:	Anteneh, H.; Song, J.
Deposited on	:	2020-03-20
Resolution	:	2.44 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 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:

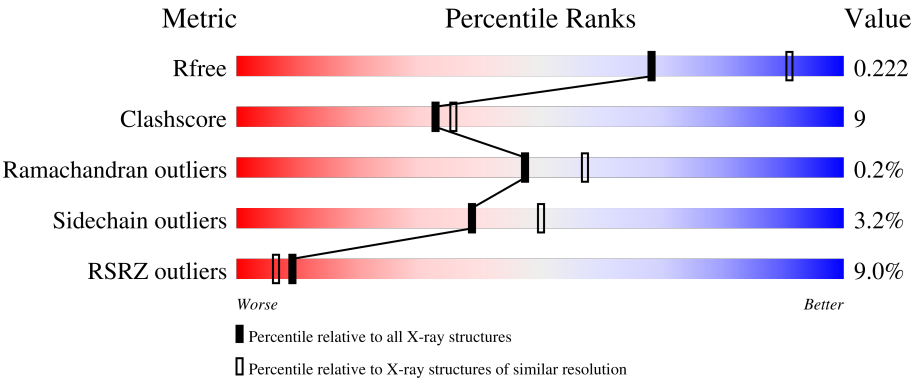
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.44 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	1564 (2.46-2.42)
Clashscore	141614	1631 (2.46-2.42)
Ramachandran outliers	138981	1617 (2.46-2.42)
Sidechain outliers	138945	1617 (2.46-2.42)
RSRZ outliers	127900	1547 (2.46-2.42)

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% 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	285	2% 84% 16%
1	D	285	2% 83% 16%
2	B	209	11% 70% 21% 8%
2	C	209	25% 63% 21% 15%
3	E	25	64% 36%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	25	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8500 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 3A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	284	Total	C	N	O	S	0	2	0
			2287	1460	410	404	13			
1	D	285	Total	C	N	O	S	0	3	0
			2297	1467	417	400	13			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	882	HIS	ARG	engineered mutation	UNP Q9Y6K1
D	882	HIS	ARG	engineered mutation	UNP Q9Y6K1

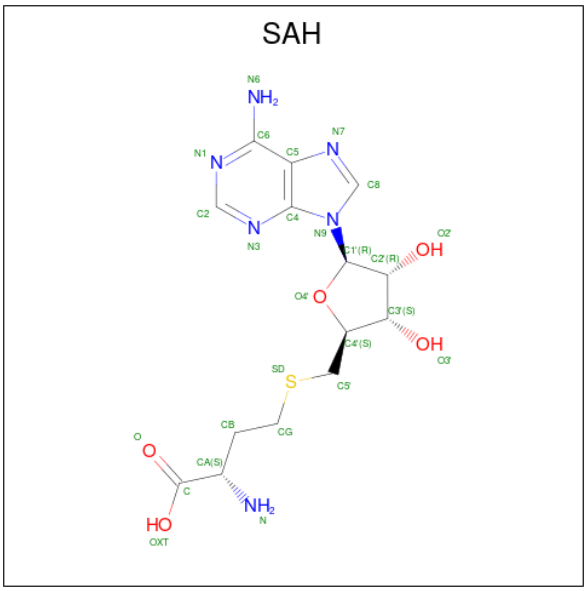
- Molecule 2 is a protein called DNA (cytosine-5)-methyltransferase 3-like.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	193	Total	C	N	O	S	0	1	0
			1469	956	248	262	3			
2	C	178	Total	C	N	O	S	0	0	0
			1263	820	214	227	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	25	Total	C	N	O	P	0	0	0
			493	235	88	146	24			
3	F	24	Total	C	N	O	P	0	0	0
			488	235	88	142	23			

- Molecule 4 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: $C_{14}H_{20}N_6O_5S$).



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

- Molecule 5 is water.

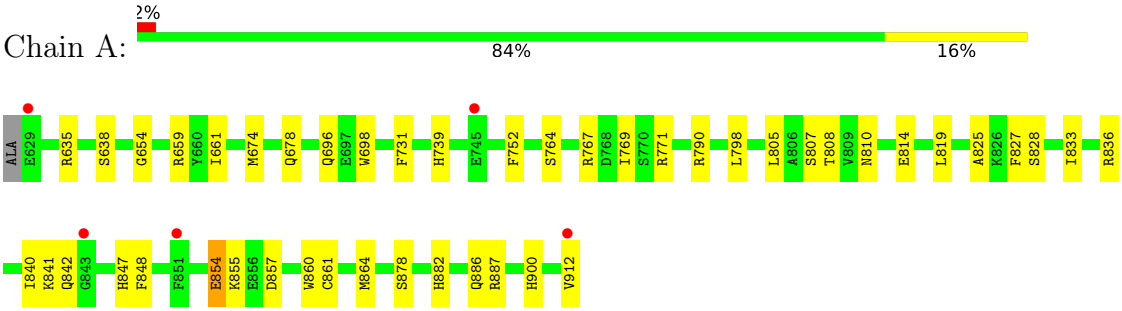
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	72	Total O 72 72	0	0
5	B	14	Total O 14 14	0	0
5	C	6	Total O 6 6	0	0
5	D	44	Total O 44 44	0	0
5	E	10	Total O 10 10	0	0
5	F	5	Total O 5 5	0	0

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 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.

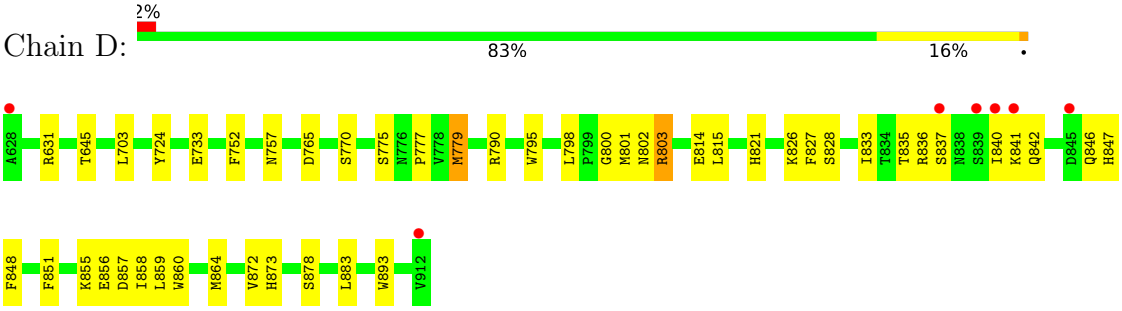
• Molecule 1: DNA (cytosine-5)-methyltransferase 3A

Chain A:



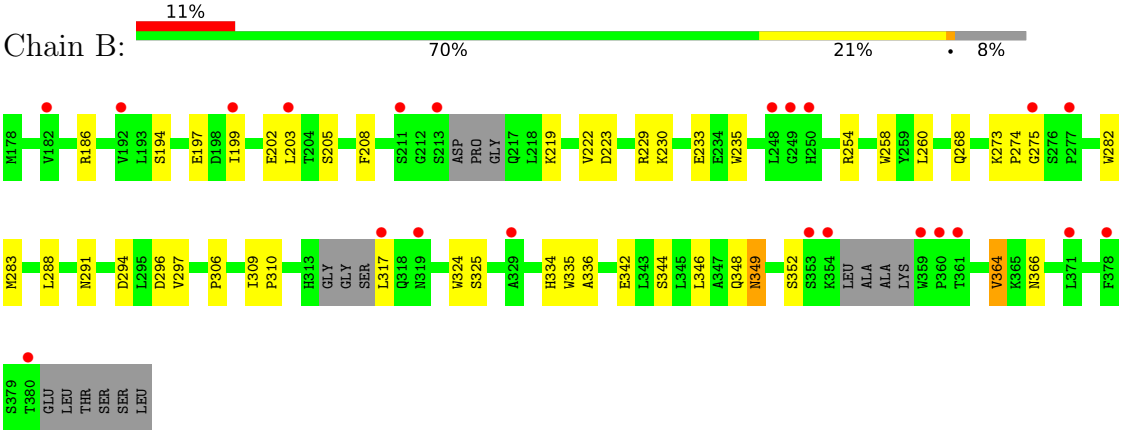
• Molecule 1: DNA (cytosine-5)-methyltransferase 3A

Chain D:

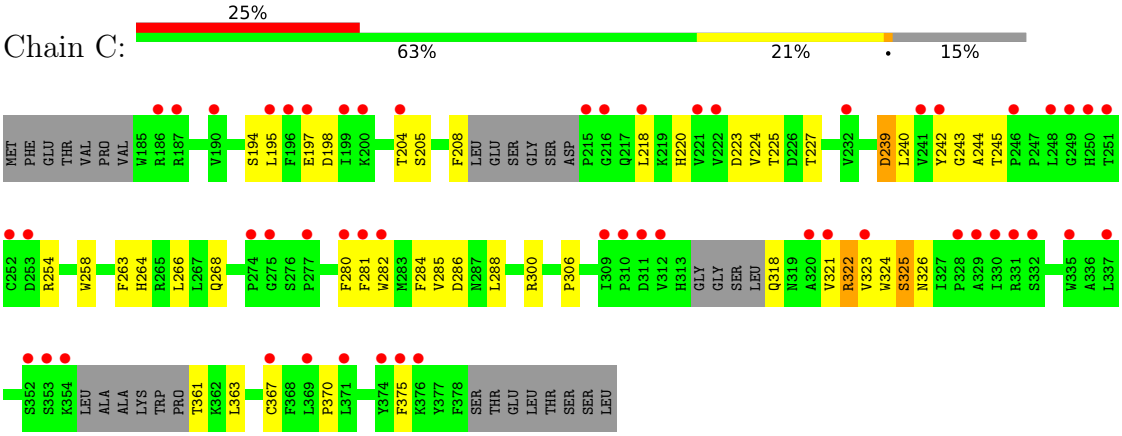


• Molecule 2: DNA (cytosine-5)-methyltransferase 3-like

Chain B:



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



• Molecule 3: CAG DNA (25-MER)

• Molecule 3: CAG DNA (25-MER)

4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	205.94Å 205.94Å 89.37Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.35 – 2.44 43.35 – 2.45	Depositor EDS
% Data completeness (in resolution range)	100.0 (43.35-2.44) 100.0 (43.35-2.45)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.82 (at 2.45Å)	Xtriage
Refinement program	PHENIX 1.15.2_3472	Depositor
R, R_{free}	0.183 , 0.223 0.183 , 0.222	Depositor DCC
R_{free} test set	1996 reflections (3.82%)	wwPDB-VP
Wilson B-factor (Å^2)	64.4	Xtriage
Anisotropy	0.055	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 66.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.023 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8500	wwPDB-VP
Average B, all atoms (Å^2)	98.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.37% 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: PYO, 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	0.28	0/2342	0.45	0/3166
1	D	0.25	0/2355	0.45	0/3181
2	B	0.26	0/1513	0.46	0/2076
2	C	0.30	0/1298	0.57	3/1783 (0.2%)
3	E	0.53	0/530	0.95	0/815
3	F	0.59	0/525	0.98	1/807 (0.1%)
All	All	0.32	0/8563	0.57	4/11828 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	322	ARG	NE-CZ-NH1	-8.20	116.20	120.30
2	C	322	ARG	NE-CZ-NH2	7.80	124.20	120.30
3	F	435	DT	OP1-P-O3'	5.09	116.41	105.20
2	C	322	ARG	CA-CB-CG	5.07	124.54	113.40

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	A	2287	0	2237	33	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	2297	0	2274	31	0
2	B	1469	0	1313	28	0
2	C	1263	0	1053	37	0
3	E	493	0	270	5	0
3	F	488	0	268	9	0
4	A	26	0	19	0	0
4	D	26	0	19	0	0
5	A	72	0	0	2	0
5	B	14	0	0	2	0
5	C	6	0	0	3	0
5	D	44	0	0	1	0
5	E	10	0	0	0	0
5	F	5	0	0	0	0
All	All	8500	0	7453	136	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 9.

All (136) 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:C:240:LEU:HD23	2:C:281:PHE:HB2	1.57	0.86
1:D:840:ILE:HG21	1:D:859:LEU:H	1.48	0.78
1:A:842:GLN:HB2	1:A:847:HIS:HB2	1.68	0.74
1:D:835:THR:OG1	1:D:836[A]:ARG:NH1	2.22	0.73
2:C:195:LEU:HD21	2:C:266:LEU:HD12	1.72	0.71
1:D:840:ILE:HG21	1:D:858:ILE:HG13	1.74	0.68
2:C:361:THR:HG22	2:C:363:LEU:H	1.59	0.68
2:B:317:LEU:N	5:B:401:HOH:O	2.30	0.65
2:C:245:THR:HG22	2:C:286:ASP:HA	1.78	0.65
1:A:864:MET:HE1	1:A:886:GLN:HG3	1.79	0.64
1:A:790:ARG:NH2	1:A:833:ILE:O	2.30	0.64
1:D:842:GLN:H	1:D:847:HIS:HB3	1.63	0.63
1:A:842:GLN:HB2	1:A:847:HIS:CB	2.29	0.63
1:A:848:PHE:HB3	1:A:857:ASP:O	1.98	0.63
2:C:285:VAL:HB	2:C:323:VAL:HG22	1.82	0.62
2:C:281:PHE:HA	2:C:326:ASN:HD21	1.65	0.61
1:D:790:ARG:NH2	1:D:833:ILE:O	2.32	0.61
2:C:282:TRP:O	2:C:325:SER:HB2	2.01	0.60
2:B:202:GLU:N	2:B:202:GLU:OE2	2.34	0.60
3:E:436:DT:H2"	3:E:437:DA:C8	2.36	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:770:SER:HB3	1:D:795:TRP:HE1	1.66	0.60
2:B:283:MET:HG3	2:B:325:SER:HB2	1.85	0.59
1:A:878:SER:HA	1:D:860:TRP:CE2	2.38	0.59
2:C:195:LEU:HD23	2:C:224:VAL:HG11	1.85	0.59
3:F:435:DT:H4'	3:F:436:DT:OP1	2.01	0.58
2:B:291:ASN:ND2	2:B:294:ASP:OD2	2.35	0.58
1:A:840:ILE:HD11	1:A:886:GLN:HE22	1.68	0.58
2:B:194:SER:OG	2:B:197:GLU:O	2.23	0.57
1:A:805:LEU:HD13	1:A:900:HIS:CG	2.41	0.56
1:A:678:GLN:HB2	5:A:1168:HOH:O	2.05	0.56
1:A:840:ILE:HD11	1:A:886:GLN:NE2	2.22	0.55
2:B:273:LYS:O	2:B:275:GLY:N	2.40	0.55
3:F:436:DT:H2"	3:F:437:DA:C8	2.42	0.54
2:B:348:GLN:O	2:B:352:SER:OG	2.19	0.54
2:B:199:ILE:O	2:B:203:LEU:HG	2.08	0.54
1:A:861:CYS:SG	1:A:882:HIS:HB2	2.47	0.54
2:C:204:THR:HA	2:C:208:PHE:H	1.73	0.54
2:C:285:VAL:CG2	2:C:323:VAL:HG22	2.37	0.54
1:D:826:LYS:HE2	1:D:851:PHE:CE1	2.42	0.53
2:C:195:LEU:CD2	2:C:266:LEU:HD12	2.38	0.53
2:C:285:VAL:HG21	2:C:321:VAL:HG13	1.90	0.53
1:A:860:TRP:CE2	1:D:878:SER:HA	2.44	0.53
1:D:837:SER:HB3	1:D:883:LEU:HD21	1.90	0.53
3:E:438:DG:H2"	3:E:439:DA:H5"	1.89	0.53
1:D:826:LYS:NZ	1:D:856:GLU:OE1	2.34	0.52
2:C:318:GLN:N	5:C:401:HOH:O	2.41	0.52
2:B:306:PRO:HB3	2:B:324:TRP:CE2	2.45	0.51
1:A:840:ILE:HD12	1:A:840:ILE:H	1.74	0.51
2:C:285:VAL:CB	2:C:323:VAL:HG22	2.40	0.51
1:D:814:GLU:OE1	1:D:828:SER:HB2	2.10	0.51
1:D:826:LYS:HE2	1:D:851:PHE:CZ	2.45	0.51
1:A:752:PHE:HB3	1:A:798:LEU:HD23	1.92	0.50
2:B:254:ARG:HG2	2:B:258:TRP:CD2	2.46	0.50
2:C:375:PHE:HZ	5:C:403:HOH:O	1.95	0.49
1:D:840:ILE:HG23	1:D:859:LEU:HB2	1.94	0.49
2:C:306:PRO:HB3	2:C:324:TRP:CE2	2.47	0.49
1:A:836[A]:ARG:NH2	1:A:887[A]:ARG:HH21	2.10	0.49
1:A:827:PHE:CE2	1:A:842:GLN:HG3	2.47	0.49
2:C:197:GLU:OE1	2:C:244:ALA:HB3	2.13	0.49
1:A:819:LEU:HD11	1:A:825:ALA:HB2	1.95	0.49
1:D:800:GLY:O	1:D:803:ARG:HB3	2.13	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:435:DT:H2"	3:F:436:DT:H5"	1.94	0.48
1:D:645:THR:HG21	1:D:893:TRP:HB2	1.96	0.48
2:C:218:LEU:HD21	2:C:220:HIS:HB2	1.96	0.48
2:B:199:ILE:HD12	2:B:203:LEU:HD21	1.96	0.48
2:C:285:VAL:CG2	2:C:321:VAL:HG13	2.44	0.48
2:C:367:CYS:O	2:C:370:PRO:HD2	2.14	0.48
2:C:245:THR:HG21	2:C:286:ASP:OD1	2.14	0.47
2:B:282:TRP:O	2:B:325:SER:OG	2.26	0.47
3:F:438:DG:H2"	3:F:439:DA:C8	2.50	0.47
2:B:309:ILE:HG23	2:B:335:TRP:CE3	2.50	0.46
2:C:258:TRP:HD1	1:D:724:TYR:CZ	2.33	0.46
1:A:808:THR:OG1	1:A:810:ASN:OD1	2.31	0.46
2:C:195:LEU:HD12	2:C:242:TYR:O	2.15	0.46
2:C:245:THR:CG2	2:C:286:ASP:HA	2.43	0.46
2:C:264:HIS:O	2:C:268:GLN:HG2	2.16	0.46
2:C:285:VAL:HG21	2:C:321:VAL:CG1	2.45	0.46
1:D:815:LEU:HD11	1:D:859:LEU:HD11	1.97	0.46
1:D:840:ILE:HD11	1:D:864:MET:HE1	1.98	0.46
2:B:342:GLU:OE2	2:B:342:GLU:N	2.48	0.46
2:C:227:THR:HG21	2:C:266:LEU:HD11	1.97	0.46
1:D:841:LYS:HD3	1:D:847:HIS:H	1.79	0.46
2:B:294:ASP:HA	2:B:297:VAL:HG22	1.97	0.45
1:A:814:GLU:OE2	1:A:828:SER:HB3	2.16	0.45
1:A:841:LYS:HD2	1:A:847:HIS:N	2.31	0.45
2:C:243:GLY:O	2:C:284:PHE:HA	2.16	0.45
1:A:854:GLU:OE1	1:A:855:LYS:HG3	2.17	0.45
2:B:288:LEU:HD11	2:B:346:LEU:HD12	1.98	0.45
1:D:857:ASP:OD1	1:D:858:ILE:N	2.50	0.45
3:F:435:DT:H1'	3:F:436:DT:H5"	1.98	0.45
2:B:222:VAL:HG23	2:B:223:ASP:H	1.81	0.45
3:E:428:DA:H5'	3:E:428:DA:C8	2.52	0.45
2:B:310:PRO:HG2	2:B:336:ALA:HB3	1.98	0.45
1:A:764:SER:HB2	1:A:767:ARG:NH2	2.32	0.45
1:D:631[B]:ARG:H	1:D:631[B]:ARG:HG2	1.51	0.45
3:F:432:DT:H4'	3:F:433:DA:OP1	2.17	0.44
2:C:300:ARG:NH1	1:D:733:GLU:OE2	2.50	0.44
1:D:779:MET:C	1:D:779:MET:SD	2.96	0.44
2:B:230:LYS:HB2	5:B:408:HOH:O	2.17	0.44
2:C:288:LEU:HA	5:C:402:HOH:O	2.16	0.44
1:D:703:LEU:HA	1:D:752:PHE:O	2.18	0.44
1:D:798:LEU:O	1:D:801:MET:HG3	2.18	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:635:ARG:HD2	1:A:659:ARG:NH2	2.33	0.44
3:E:439:DA:H2"	3:E:440:DC:O4'	2.18	0.44
1:A:678:GLN:HG2	1:D:821:HIS:CG	2.53	0.43
2:C:263:PHE:CE2	2:C:284:PHE:HB2	2.54	0.43
2:B:229:ARG:O	2:B:233:GLU:HG3	2.19	0.43
2:C:194:SER:HA	2:C:242:TYR:O	2.18	0.43
1:A:696:GLN:HB3	5:A:1114:HOH:O	2.18	0.43
2:C:225:THR:HG22	2:C:258:TRP:HZ2	1.82	0.43
1:D:848:PHE:HB3	1:D:857:ASP:O	2.18	0.43
3:F:435:DT:C2'	3:F:436:DT:H5"	2.49	0.43
2:B:219:LYS:HE3	2:B:235:TRP:CE2	2.54	0.42
2:B:309:ILE:HA	2:B:310:PRO:HD3	1.90	0.42
2:B:203:LEU:HB3	2:B:208:PHE:HB2	2.01	0.42
1:A:840:ILE:CD1	1:A:886:GLN:HE22	2.32	0.42
2:B:364:VAL:HG12	2:B:364:VAL:O	2.19	0.42
2:C:239:ASP:HA	2:C:280:PHE:HD1	1.85	0.42
2:B:364:VAL:HA	2:B:366:ASN:OD1	2.19	0.42
2:C:285:VAL:HG21	2:C:323:VAL:HG22	2.00	0.42
1:A:739:HIS:CD2	2:B:268:GLN:HG2	2.55	0.42
2:C:285:VAL:HG23	2:C:322:ARG:C	2.40	0.42
1:A:836[B]:ARG:NH2	3:F:430:DT:O4	2.52	0.41
1:D:826:LYS:NZ	1:D:847:HIS:HE1	2.18	0.41
1:A:654:GLY:CA	1:A:912:VAL:HG22	2.49	0.41
1:A:661:ILE:CD1	1:A:698:TRP:HB3	2.50	0.41
1:A:840:ILE:HD12	1:A:840:ILE:N	2.34	0.41
2:B:186:ARG:H	2:B:186:ARG:HG2	1.77	0.41
2:B:349:ASN:HA	2:B:352:SER:HB2	2.02	0.41
1:D:872:VAL:HG13	1:D:873:HIS:CD2	2.55	0.41
3:E:432:DT:H1'	3:E:433:DA:C8	2.56	0.41
2:C:223:ASP:OD2	2:C:254:ARG:NH2	2.54	0.41
3:F:432:DT:H2"	3:F:433:DA:O5'	2.21	0.40
1:A:674:MET:HB3	5:D:1143:HOH:O	2.22	0.40
1:A:731:PHE:CE1	1:A:769:ILE:HG23	2.56	0.40
1:D:777:PRO:HB3	1:D:795:TRP:CE2	2.56	0.40

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	A	284/285 (100%)	277 (98%)	7 (2%)	0	100 100
1	D	286/285 (100%)	278 (97%)	8 (3%)	0	100 100
2	B	186/209 (89%)	172 (92%)	12 (6%)	2 (1%)	14 15
2	C	170/209 (81%)	157 (92%)	13 (8%)	0	100 100
All	All	926/988 (94%)	884 (96%)	40 (4%)	2 (0%)	47 57

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	274	PRO
2	B	364	VAL

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	240/250 (96%)	236 (98%)	4 (2%)	60 73
1	D	243/250 (97%)	234 (96%)	9 (4%)	34 45
2	B	141/191 (74%)	135 (96%)	6 (4%)	29 38
2	C	105/191 (55%)	101 (96%)	4 (4%)	33 43
All	All	729/882 (83%)	706 (97%)	23 (3%)	39 50

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	638	SER
1	A	771	ARG
1	A	807	SER
1	A	854	GLU
2	B	205	SER
2	B	260	LEU
2	B	296	ASP
2	B	334	HIS
2	B	344	SER
2	B	349	ASN
2	C	198	ASP
2	C	205	SER
2	C	239	ASP
2	C	325	SER
1	D	757	ASN
1	D	765	ASP
1	D	775	SER
1	D	779	MET
1	D	802	ASN
1	D	803	ARG
1	D	827	PHE
1	D	846	GLN
1	D	855	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	853	ASN
1	A	886	GLN
2	B	319	ASN
2	C	220	HIS
1	D	692	GLN
1	D	847	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

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	16,20,21	2.07	4 (25%)	22,28,31	1.01	1 (4%)
3	PYO	F	427	1,3	16,20,21	1.94	4 (25%)	22,28,31	0.83	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	PYO	E	427	1,3	-	4/7/25/26	0/2/2/2
3	PYO	F	427	1,3	-	4/7/25/26	0/2/2/2

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	427	PYO	C6-C5	6.31	1.49	1.35
3	F	427	PYO	C6-C5	5.73	1.48	1.35
3	E	427	PYO	C5-C4	2.93	1.46	1.40
3	E	427	PYO	C2-N1	2.85	1.46	1.40
3	F	427	PYO	C2-N1	2.82	1.46	1.40
3	F	427	PYO	C5-C4	2.69	1.45	1.40
3	E	427	PYO	C4-N3	2.41	1.42	1.33
3	F	427	PYO	C4-N3	2.36	1.42	1.33

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	427	PYO	C5-C4-N3	-3.84	119.58	124.29

There are no chirality outliers.

All (8) torsion outliers are listed below:

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

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

2 ligands 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	$\#\|Z\| > 2$
4	SAH	A	1001	-	24,28,28	1.21	3 (12%)	25,40,40	1.61	4 (16%)
4	SAH	D	1001	-	24,28,28	1.19	3 (12%)	25,40,40	1.61	4 (16%)

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
4	SAH	A	1001	-	-	2/11/31/31	0/3/3/3
4	SAH	D	1001	-	-	4/11/31/31	0/3/3/3

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1001	SAH	C2-N3	4.04	1.38	1.32
4	D	1001	SAH	C2-N3	3.95	1.38	1.32
4	A	1001	SAH	C2-N1	2.46	1.38	1.33
4	D	1001	SAH	C2-N1	2.41	1.38	1.33
4	D	1001	SAH	OXT-C	-2.14	1.23	1.30
4	A	1001	SAH	OXT-C	-2.10	1.23	1.30

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1001	SAH	N3-C2-N1	-5.49	120.10	128.68
4	D	1001	SAH	N3-C2-N1	-5.46	120.14	128.68
4	D	1001	SAH	C5'-SD-CG	-3.12	92.91	102.27
4	A	1001	SAH	C5'-SD-CG	-2.93	93.47	102.27
4	A	1001	SAH	OXT-C-O	-2.85	117.62	124.09
4	D	1001	SAH	OXT-C-O	-2.70	117.96	124.09
4	A	1001	SAH	OXT-C-CA	2.29	121.17	113.38
4	D	1001	SAH	OXT-C-CA	2.19	120.85	113.38

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1001	SAH	N-CA-CB-CG
4	D	1001	SAH	N-CA-CB-CG
4	A	1001	SAH	CB-CG-SD-C5'
4	D	1001	SAH	CB-CG-SD-C5'
4	D	1001	SAH	O-C-CA-CB
4	D	1001	SAH	OXT-C-CA-CB

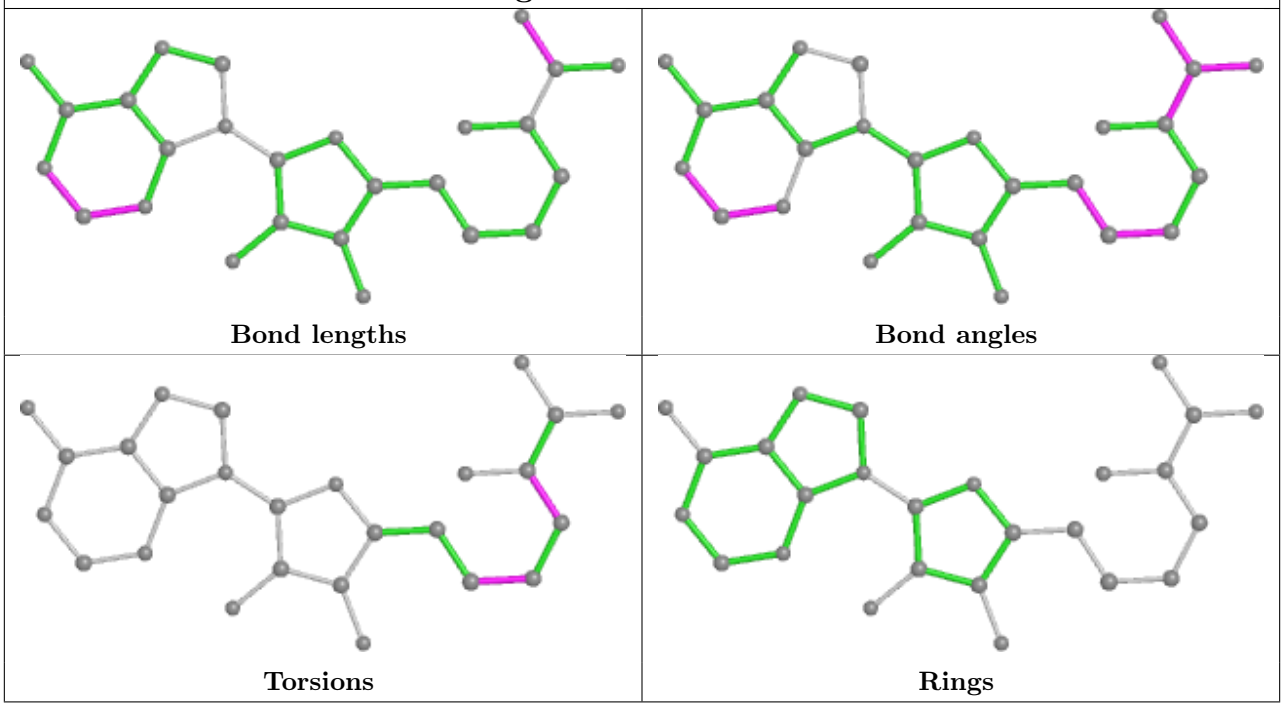
There are no ring outliers.

No monomer is involved in short contacts.

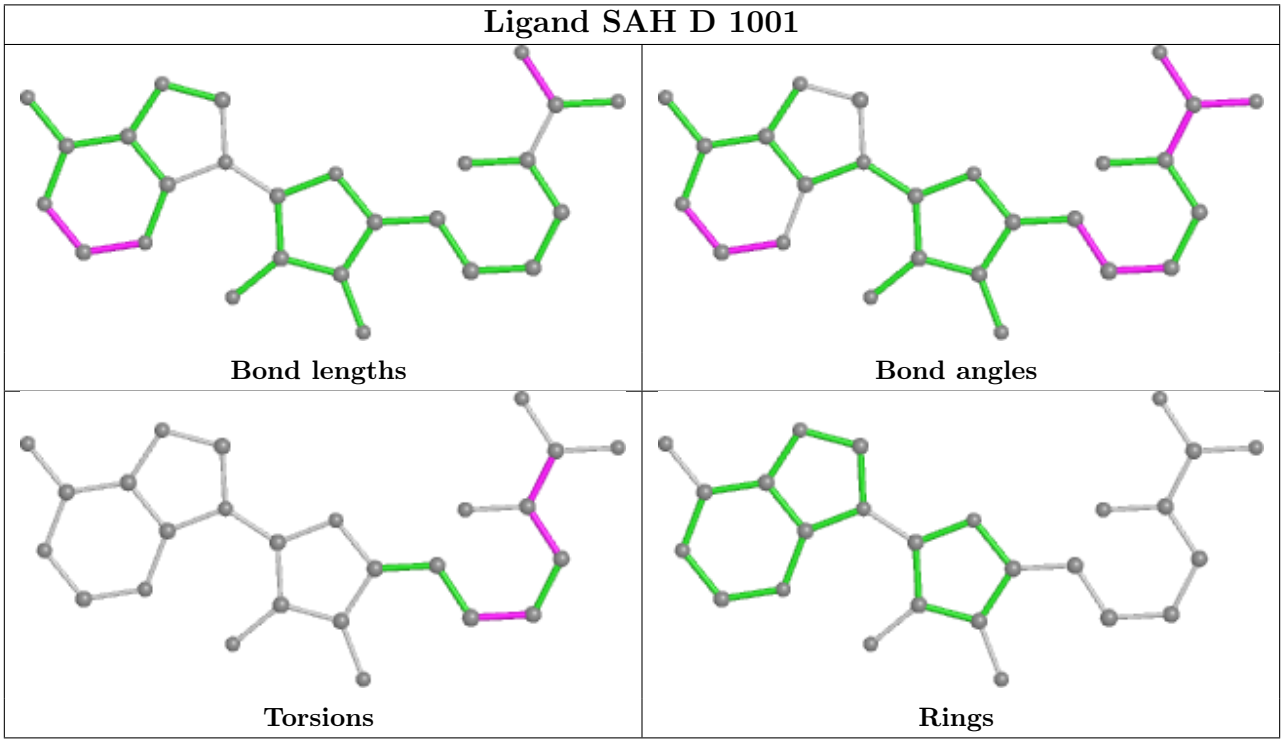
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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 SAH A 1001



Ligand SAH D 1001



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^{th} 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^2)	Q<0.9
1	A	284/285 (99%)	-0.07	5 (1%) 68 64	45, 70, 135, 190	0
1	D	285/285 (100%)	0.16	7 (2%) 57 53	43, 70, 125, 189	0
2	B	193/209 (92%)	0.48	22 (11%) 5 3	65, 113, 171, 194	0
2	C	178/209 (85%)	1.25	53 (29%) 0 0	83, 146, 196, 228	0
3	E	24/25 (96%)	-0.21	0 100 100	98, 113, 167, 187	0
3	F	23/25 (92%)	0.00	2 (8%) 10 7	70, 126, 195, 211	0
All	All	987/1038 (95%)	0.34	89 (9%) 9 6	43, 90, 171, 228	0

All (89) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	359	TRP	7.6
1	D	840	ILE	6.9
2	C	199	ILE	6.6
1	D	839	SER	6.0
1	D	912	VAL	5.7
2	C	331	ARG	5.6
2	C	215	PRO	5.5
2	B	248	LEU	5.4
2	C	190	VAL	5.2
2	C	375	PHE	5.2
1	A	912	VAL	5.1
2	B	192	VAL	4.9
2	C	275	GLY	4.8
2	B	360	PRO	4.8
2	C	251	THR	4.7
1	D	841	LYS	4.6
1	A	629	GLU	4.5
2	C	328	PRO	4.5
2	C	329	ALA	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	C	374	TYR	4.4
2	C	310	PRO	4.3
2	B	249	GLY	4.2
2	C	197	GLU	4.2
2	C	312	VAL	4.2
2	C	311	ASP	4.2
2	C	309	ILE	4.1
2	C	335	TRP	4.1
2	C	321	VAL	4.0
2	C	218	LEU	4.0
2	B	277	PRO	3.7
2	B	371	LEU	3.7
2	C	367	CYS	3.7
2	C	250	HIS	3.6
2	C	376	LYS	3.6
2	B	378	PHE	3.4
1	D	837	SER	3.4
1	A	843	GLY	3.3
2	B	250	HIS	3.3
2	C	248	LEU	3.3
2	C	369	LEU	3.3
1	A	851	PHE	3.2
2	C	320	ALA	3.2
2	C	330	ILE	3.2
2	B	361	THR	3.1
2	C	277	PRO	3.0
2	C	337	LEU	3.0
2	B	329	ALA	2.9
2	C	196	PHE	2.9
2	C	282	TRP	2.9
2	C	252	CYS	2.8
2	C	280	PHE	2.8
2	C	249	GLY	2.8
2	C	216	GLY	2.8
2	B	275	GLY	2.8
2	C	204	THR	2.7
2	C	274	PRO	2.7
2	B	203	LEU	2.7
2	C	353	SER	2.6
2	C	354	LYS	2.6
2	B	319	ASN	2.6
2	C	195	LEU	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	211	SER	2.5
2	C	332	SER	2.5
2	B	213	SER	2.5
2	B	354	LYS	2.4
2	C	242	TYR	2.4
2	C	281	PHE	2.4
2	C	371	LEU	2.4
2	B	182	VAL	2.4
2	C	186	ARG	2.4
2	C	187	ARG	2.4
2	C	200	LYS	2.3
2	C	221	VAL	2.3
2	C	241	VAL	2.3
2	C	246	PRO	2.3
1	D	845	ASP	2.3
2	B	380	THR	2.3
2	B	353	SER	2.2
2	B	317	LEU	2.2
3	F	439	DA	2.2
2	C	222	VAL	2.2
2	C	232	VAL	2.2
1	D	628	ALA	2.1
2	C	253	ASP	2.1
2	C	323	VAL	2.1
1	A	745	GLU	2.1
2	B	199	ILE	2.0
2	C	352	SER	2.0
3	F	440	DC	2.0

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

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^{th} 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^2)	Q<0.9
3	PYO	E	427	19/20	0.96	0.14	78,84,91,96	0
3	PYO	F	427	19/20	0.98	0.13	58,70,88,91	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

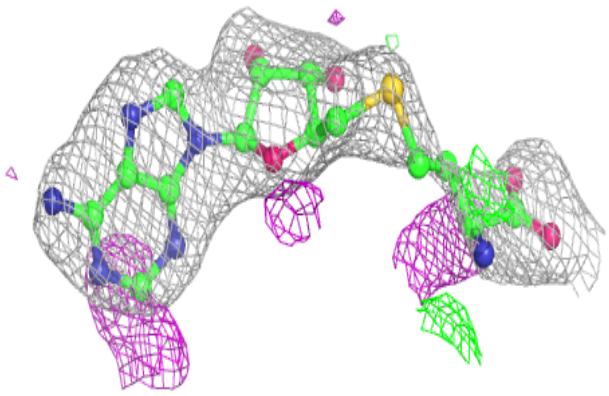
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

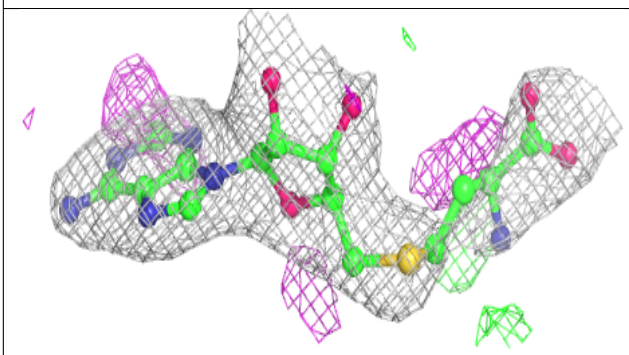
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	SAH	D	1001	26/26	0.94	0.18	54,80,110,112	0
4	SAH	A	1001	26/26	0.97	0.15	58,69,99,104	0

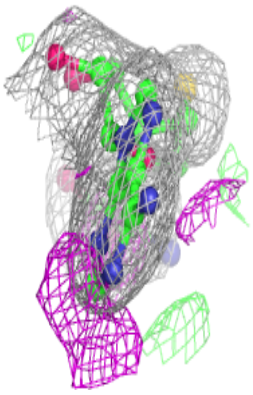
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around SAH D 1001:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

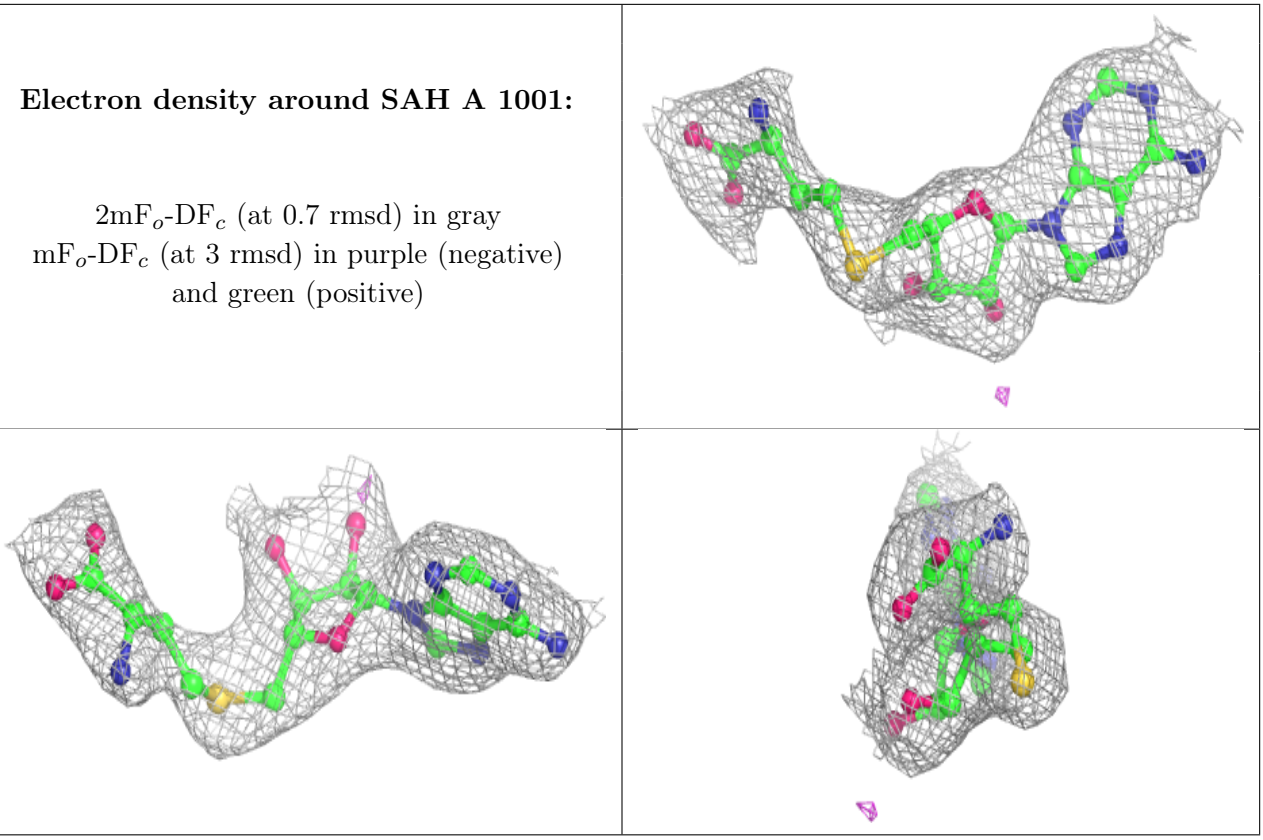






Electron density around SAH A 1001:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers

There are no such residues in this entry.