W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 9, 2023 – 07:50 AM EDT

PDB ID	:	6W89
Title	:	Structure of DNMT3A (R882H) in complex with CGA DNA
Authors	:	Anteneh, H.; Song, J.
Deposited on	:	2020-03-20
Resolution	:	2.50 Å(reported)

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 ⓘ

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.
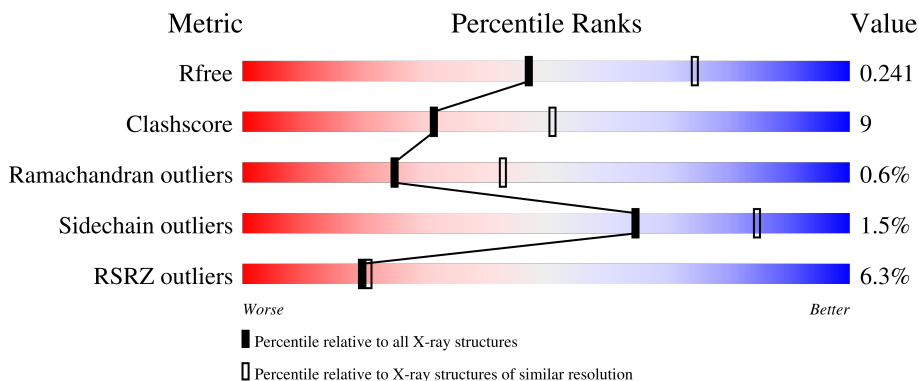
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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% 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	85% 15%
1	D	285	.% 90% 10%
1	G	285	87% 13%
1	J	285	.% 88% 12%
2	B	209	13% 67% 26% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	209	11% · 60% · 24% · 15%
2	H	209	17% · 63% · 32% · 6%
2	I	209	16% · 63% · 22% · 13%
3	E	25	4% · 64% · 28% · ·
3	F	25	64% · 36%
3	K	25	56% · 40% ·
3	L	25	60% · 36% ·

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 17145 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	285	Total 2284	C 1458	N 407	O 406	S 13	0	2	0
1	D	285	Total 2300	C 1469	N 413	O 405	S 13	0	3	0
1	G	284	Total 2289	C 1462	N 410	O 404	S 13	0	3	0
1	J	285	Total 2308	C 1474	N 414	O 407	S 13	0	4	0

There are 4 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
G	882	HIS	ARG	engineered mutation	UNP Q9Y6K1
J	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	194	Total 1464	C 953	N 241	O 267	S 3	0	1	0
2	C	178	Total 1275	C 823	N 215	O 235	S 2	0	0	0
2	H	197	Total 1477	C 966	N 242	O 265	S 4	0	0	0
2	I	182	Total 1314	C 851	N 224	O 237	S 2	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	24	Total	C	N	O	P	0	0	0
			489	235	88	143	23			
3	F	25	Total	C	N	O	P	0	0	0
			510	245	93	148	24			
3	K	25	Total	C	N	O	P	0	0	0
			510	245	93	148	24			
3	L	25	Total	C	N	O	P	0	0	0
			510	245	93	148	24			

- Molecule 4 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).

CIT
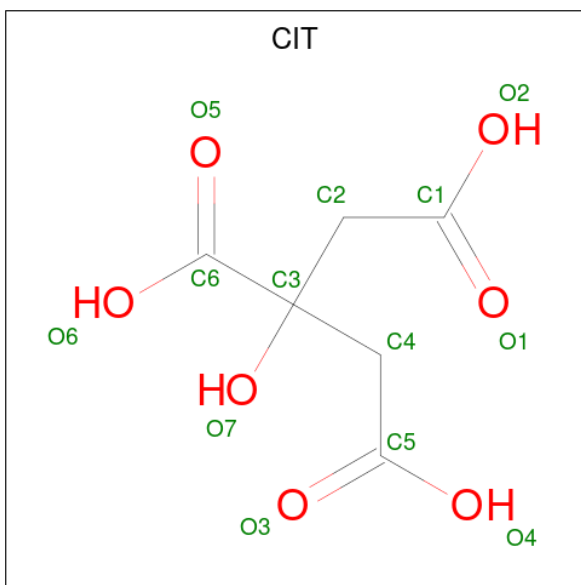
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			13	6	7		
4	D	1	Total	C	O	0	0
			13	6	7		
4	G	1	Total	C	O	0	0
			13	6	7		
4	J	1	Total	C	O	0	0
			13	6	7		

- Molecule 5 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: $C_{14}H_{20}N_6O_5S$).

SAH

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		
5	G	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
5	J	1	Total	C	N	O	S	0	0
			26	14	6	5	1		

- Molecule 6 is water.

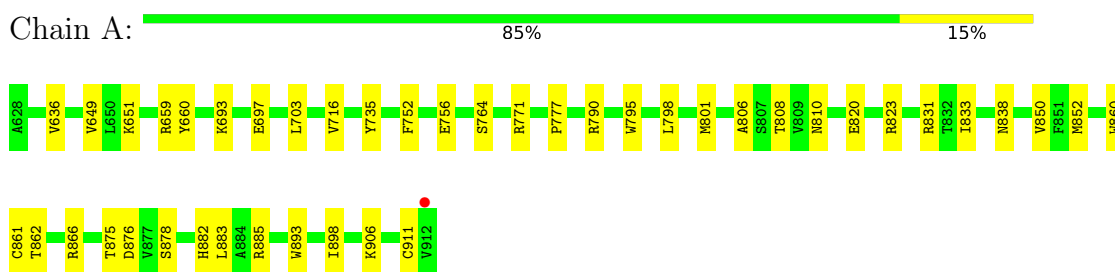| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	69	Total	O	0	0
			69	69		
6	B	12	Total	O	0	0
			12	12		
6	D	41	Total	O	0	0
			41	41		
6	C	3	Total	O	0	0
			3	3		
6	E	2	Total	O	0	0
			2	2		
6	F	2	Total	O	0	0
			2	2		
6	G	56	Total	O	0	0
			56	56		
6	H	4	Total	O	0	0
			4	4		

Continued on next page...

Continued from previous page...

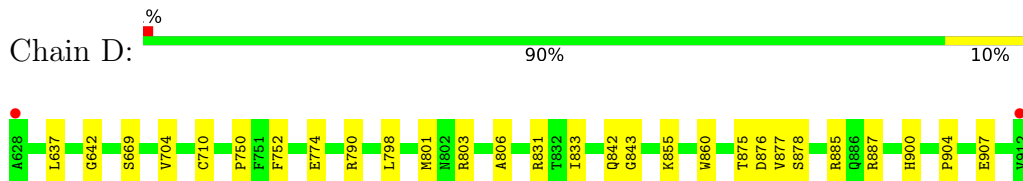| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	J	51	Total 51	O 51	0	0
6	I	7	Total 7	O 7	0	0
6	K	5	Total 5	O 5	0	0
6	L	7	Total 7	O 7	0	0

3 Residue-property plots ⓘ

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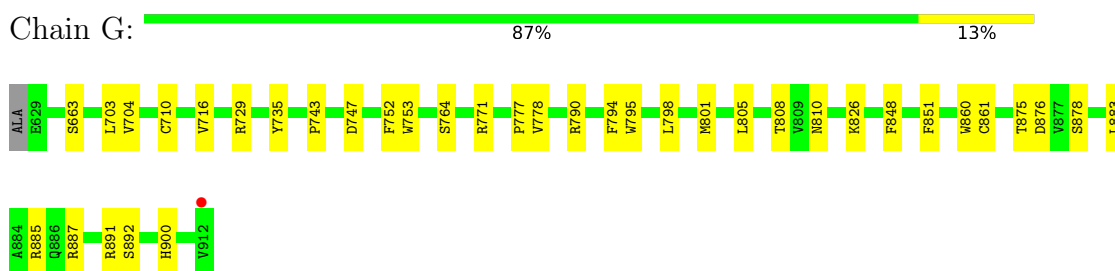
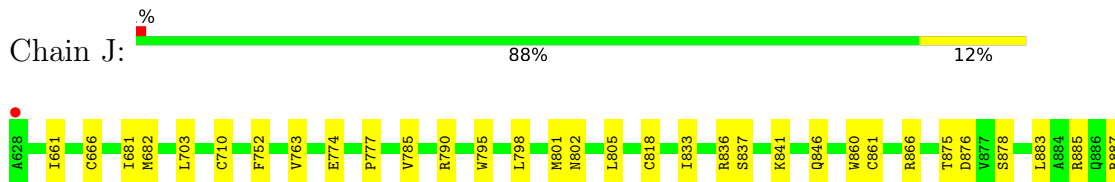• Molecule 1: DNA (cytosine-5)-methyltransferase 3A

Chain A:

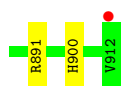• Molecule 1: DNA (cytosine-5)-methyltransferase 3A

Chain D:

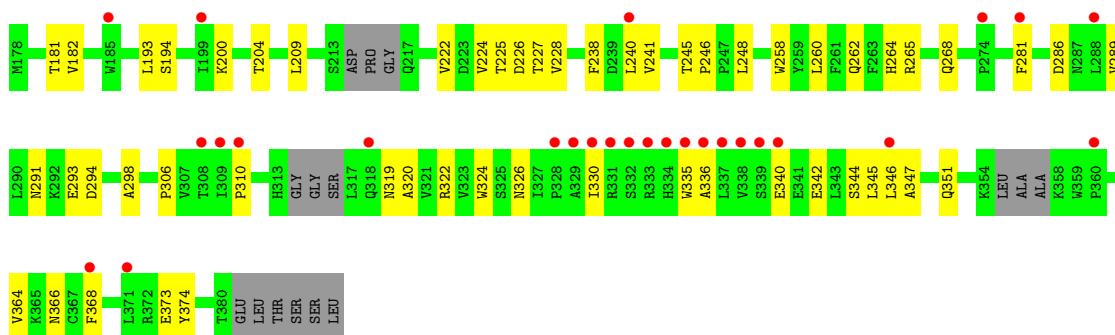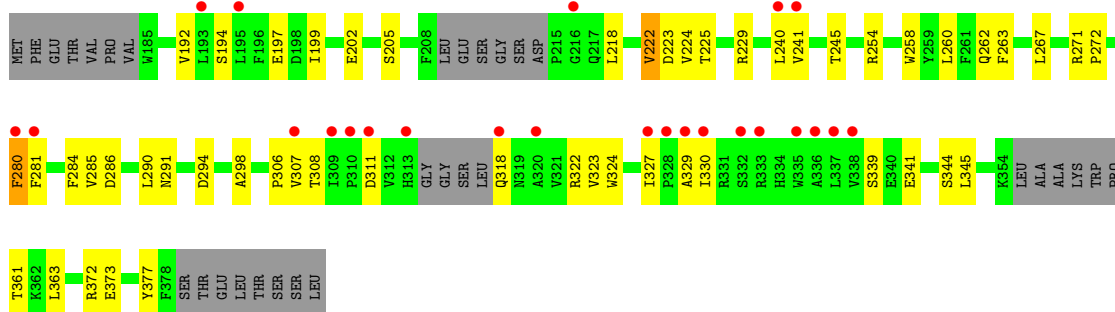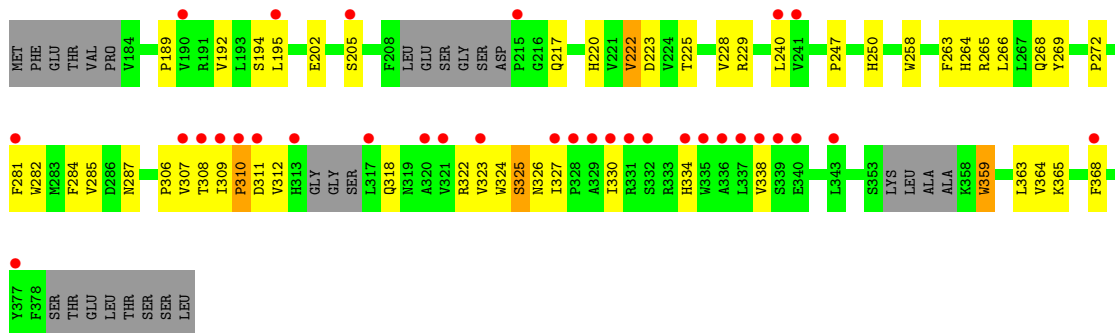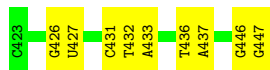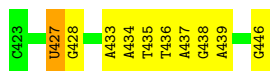
• Molecule 1: DNA (cytosine-5)-methyltransferase 3A

Chain G:

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

Chain J:

R891 H900 V912

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

13%

Chain B: 67% 26% 7%

M178 T181 V182 W185 L193 S194 I199 K200 T204 L209 S213 ASP PRO GLY Q217 V222 D223 V224 T225 T227 V228 F238 D239 L240 V241 T245 P246 P247 L248 W258 Y259 L260 F261 Q262 F263 H264 R265 Q268 P274 F281 D286 N287 L288 V289

L290 N291 K292 E293 D294 A298 P306 V307 T308 I309 P310 H313 GLY GLY SER L317 Q318 N319 A320 V321 R322 V323 W324 S325 N326 I327 P328 A329 I330 R331 S332 R333 H334 A336 L337 V338 S339 E340 E341 E342 L343 S344 L345 L346 A347 Q351 K354 LEU ALA ALA K358 W359 P360

V364 K365 N366 C367 F368 L371 R372 E373 Y374 T380 GLU LEU THR SER SER LEU

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

11%

Chain C: 60% 24% · 15%

MET PHE GLU THR VAL PRO VAL W185 V192 L193 S194 L195 F196 E197 D198 I199 E202 S205 F208 LEU GLU SER GLY SER ASP P215 G216 Q217 L218 V222 D223 V224 T225 R229 L240 V241 T245 R254 W258 Y259 L260 F261 Q262 F263 L267 R271 P272

F280 F281 F284 V285 D286 L290 N291 D294 A298 P306 V307 T308 I309 P310 D311 V312 H313 GLY GLY SER LEU Q318 N319 A320 V321 R322 V323 W324 I327 P328 A329 I330 R331 S332 R333 H334 W335 A336 L337 V338 S339 E340 E341 S344 L345 K354 LEU ALA ALA LYS TRP PRO

T361 K362 L363 R372 E373 Y377 F378 SER THR GLU LEU THR SER SER LEU

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

17%

Chain H: 63% 32% 6%

M178 P183 R187 L193 S194 I199 K200 K201 E202 L203 T204 S205 L206 G207 F208 E210 S211 G212 P215 V221 V224 T226 T227 W235 L240 V241 Y242 P246 P247 L248 G249 C252 S257 W258 Y259 L260 H264 Q268 P274 G275 F280 F281

V285 L288 V289 L290 N291 K292 E293 D294 L295 A298 P306 V307 T308 I309 P310 D311 V312 HIS GLY GLY SER L317 Q318 N319 A320 V321 R322 V323 W324 S325 N326 I327 P328 A329 I330 R331 S332 R333 H334 W335 A336 L337 V338 S339 E340 E341 E342 L343 S344 L345 S352 S353 K354 LEU ALA

A357 T361 K362 L363 V364 F368 L371 F375 F378 S379 T380 GLU LEU THR SER SER LEU

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

16%

Chain I: 63% 22% · 13%

MET PHE GLU THR VAL PRO V184 P189 V190 R191 V192 L193 S194 L195 E202 S205 F208 LEU GLU SER GLY SER ASP P215 G216 Q217 H220 V221 V222 D223 V224 T225 V228 R229 L240 V241 P247 H250 W258 F263 H264 R265 L266 L267 Q268 Y269 P272

F281 W282 M283 F284 V285 D286 N287 P306 V307 T308 I309 P310 D311 V312 H313 GLY GLY SER L317 Q318 N319 A320 V321 R322 V323 W324 S325 N326 I327 P328 A329 I330 R331 S332 R333 H334 W335 A336 L337 V338 S339 E340 L343 S353 LYS LEU ALA ALA K358 W359 L363 V364 K365 F368

Y377 F378 SER THR GLU LEU THR SER LEU

• **Molecule 3: CGA DNA (25-MER)**

Chain E:

4%
64%   28%   • •

C423 U427 G428 A433 A434 T435 T436 A437 G438 G446 DG

• **Molecule 3: CGA DNA (25-MER)**

Chain F:

64%   36%

C423 G426 U427 C431 T432 A433 T436 A437 G446 G447

• **Molecule 3: CGA DNA (25-MER)**

Chain K:

56%   40%   •

C423 U427 G428 A433 A434 T435 T436 A437 G438 A439 G446 G447

• **Molecule 3: CGA DNA (25-MER)**

Chain L:

60%   36%   •

C423 G426 U427 C431 T432 A433 A434 T435 T436 A437 G446 G447

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	187.56Å 187.56Å 82.40Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	37.72 – 2.50 39.53 – 2.50	Depositor EDS
% Data completeness (in resolution range)	98.5 (37.72-2.50) 98.5 (39.53-2.50)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.65 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.15.2_3472	Depositor
R, R_{free}	0.196 , 0.240 0.199 , 0.241	Depositor DCC
R_{free} test set	2014 reflections (1.82%)	wwPDB-VP
Wilson B-factor (Å^2)	56.2	Xtriage
Anisotropy	0.729	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 58.7	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.487 for -h,-k,l 0.023 for h,-h-k,-l 0.023 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	17145	wwPDB-VP
Average B, all atoms (Å^2)	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.01% 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, CIT

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	0.40	0/2342	0.54	0/3168
1	D	0.37	0/2364	0.52	0/3193
1	G	0.35	0/2353	0.50	0/3180
1	J	0.37	0/2372	0.51	0/3204
2	B	0.34	0/1508	0.53	0/2071
2	C	0.37	0/1311	0.54	0/1802
2	H	0.41	0/1523	0.58	0/2093
2	I	0.36	0/1352	0.55	0/1859
3	E	0.64	0/526	0.98	0/808
3	F	0.68	0/550	0.98	0/846
3	K	0.64	0/550	1.00	0/846
3	L	0.66	0/550	0.98	0/846
All	All	0.42	0/17301	0.62	0/23916

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	A	2284	0	2230	41	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	2300	0	2281	22	0
1	G	2289	0	2255	27	0
1	J	2308	0	2283	30	0
2	B	1464	0	1289	34	0
2	C	1275	0	1069	40	0
2	H	1477	0	1309	37	0
2	I	1314	0	1127	39	0
3	E	489	0	271	11	0
3	F	510	0	279	5	0
3	K	510	0	279	13	0
3	L	510	0	279	10	0
4	A	13	0	5	1	0
4	D	13	0	5	3	0
4	G	13	0	5	0	0
4	J	13	0	5	1	0
5	A	26	0	19	0	0
5	D	26	0	19	0	0
5	G	26	0	19	1	0
5	J	26	0	19	1	0
6	A	69	0	0	10	0
6	B	12	0	0	2	0
6	C	3	0	0	0	0
6	D	41	0	0	2	0
6	E	2	0	0	1	0
6	F	2	0	0	0	0
6	G	56	0	0	3	0
6	H	4	0	0	0	0
6	I	7	0	0	1	0
6	J	51	0	0	6	0
6	K	5	0	0	3	0
6	L	7	0	0	0	0
All	All	17145	0	15047	283	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.

The worst 5 of 283 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:660:TYR:HA	6:A:1105:HOH:O	1.57	1.03
2:H:285:VAL:HG12	2:H:323:VAL:HG22	1.50	0.93

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:281:PHE:HD2	2:C:327:ILE:HG12	1.33	0.91
2:C:285:VAL:HG23	2:C:323:VAL:HG22	1.50	0.91
1:A:636:VAL:HB	6:A:1105:HOH:O	1.74	0.87

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	285/285 (100%)	279 (98%)	6 (2%)	0	100	100
1	D	286/285 (100%)	280 (98%)	6 (2%)	0	100	100
1	G	285/285 (100%)	280 (98%)	5 (2%)	0	100	100
1	J	287/285 (101%)	281 (98%)	6 (2%)	0	100	100
2	B	187/209 (90%)	174 (93%)	10 (5%)	3 (2%)	9	17
2	C	170/209 (81%)	156 (92%)	13 (8%)	1 (1%)	25	43
2	H	191/209 (91%)	171 (90%)	17 (9%)	3 (2%)	9	17
2	I	174/209 (83%)	157 (90%)	12 (7%)	5 (3%)	4	6
All	All	1865/1976 (94%)	1778 (95%)	75 (4%)	12 (1%)	25	43

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	222	VAL
2	C	222	VAL
2	H	212	GLY
2	H	215	PRO
2	I	222	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%)	239 (100%)	1 (0%)	91	97
1	D	246/250 (98%)	245 (100%)	1 (0%)	91	97
1	G	243/250 (97%)	239 (98%)	4 (2%)	62	84
1	J	246/250 (98%)	245 (100%)	1 (0%)	91	97
2	B	140/191 (73%)	138 (99%)	2 (1%)	67	86
2	C	110/191 (58%)	107 (97%)	3 (3%)	44	71
2	H	139/191 (73%)	133 (96%)	6 (4%)	29	53
2	I	119/191 (62%)	115 (97%)	4 (3%)	37	63
All	All	1483/1764 (84%)	1461 (98%)	22 (2%)	65	85

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

Mol	Chain	Res	Type
2	H	211	SER
1	J	861	CYS
2	H	342	GLU
2	I	194	SER
2	C	339	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	G	810	ASN
1	G	900	HIS
2	I	217	GLN
2	I	287	ASN
2	I	319	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

4 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	F	427	1,3	16,20,21	1.99	4 (25%)	22,28,31	1.14	1 (4%)
3	PYO	E	427	1,3	16,20,21	1.96	4 (25%)	22,28,31	1.16	1 (4%)
3	PYO	K	427	1,3	16,20,21	2.00	4 (25%)	22,28,31	1.04	1 (4%)
3	PYO	L	427	1,3	16,20,21	2.02	4 (25%)	22,28,31	1.03	1 (4%)

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	F	427	1,3	-	4/7/25/26	0/2/2/2
3	PYO	E	427	1,3	-	4/7/25/26	0/2/2/2
3	PYO	K	427	1,3	-	4/7/25/26	0/2/2/2
3	PYO	L	427	1,3	-	4/7/25/26	0/2/2/2

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	K	427	PYO	C6-C5	6.34	1.49	1.35
3	E	427	PYO	C6-C5	6.33	1.49	1.35
3	L	427	PYO	C6-C5	6.19	1.49	1.35
3	F	427	PYO	C6-C5	6.12	1.49	1.35
3	L	427	PYO	C2-N1	2.91	1.46	1.40

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	427	PYO	C5-C4-N3	-4.72	118.50	124.29

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	427	PYO	C5-C4-N3	-4.37	118.93	124.29
3	K	427	PYO	C5-C4-N3	-4.17	119.17	124.29
3	L	427	PYO	C5-C4-N3	-3.70	119.75	124.29

There are no chirality outliers.

5 of 16 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	K	427	PYO	C2'-C1'-N1-C6
3	F	427	PYO	C2'-C1'-N1-C6
3	L	427	PYO	C2'-C1'-N1-C6

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	427	PYO	1	0
3	K	427	PYO	1	0
3	L	427	PYO	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

8 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	CIT	A	1001	-	12,12,12	1.08	0	17,17,17	1.62	3 (17%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
4	CIT	D	1001	-	12,12,12	1.15	0	17,17,17	1.85	4 (23%)
4	CIT	G	1001	-	12,12,12	1.00	0	17,17,17	1.45	2 (11%)
5	SAH	A	1002	-	24,28,28	1.18	2 (8%)	25,40,40	1.78	5 (20%)
5	SAH	J	1002	-	24,28,28	1.24	2 (8%)	25,40,40	1.73	6 (24%)
5	SAH	G	1002	-	24,28,28	1.18	3 (12%)	25,40,40	1.66	5 (20%)
4	CIT	J	1001	-	12,12,12	1.00	0	17,17,17	7.66	9 (52%)
5	SAH	D	1002	-	24,28,28	1.27	2 (8%)	25,40,40	1.67	5 (20%)

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	CIT	A	1001	-	-	2/16/16/16	-
4	CIT	D	1001	-	-	3/16/16/16	-
4	CIT	G	1001	-	-	3/16/16/16	-
5	SAH	A	1002	-	-	2/11/31/31	0/3/3/3
5	SAH	J	1002	-	-	5/11/31/31	0/3/3/3
5	SAH	G	1002	-	-	3/11/31/31	0/3/3/3
4	CIT	J	1001	-	-	6/16/16/16	-
5	SAH	D	1002	-	-	5/11/31/31	0/3/3/3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	1002	SAH	C2-N3	4.16	1.38	1.32
5	A	1002	SAH	C2-N3	4.02	1.38	1.32
5	J	1002	SAH	C2-N3	4.00	1.38	1.32
5	G	1002	SAH	C2-N3	3.92	1.38	1.32
5	D	1002	SAH	C2-N1	2.80	1.39	1.33

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	J	1001	CIT	O7-C3-C6	13.31	127.56	108.86
4	J	1001	CIT	O7-C3-C2	-13.27	78.35	109.40
4	J	1001	CIT	O7-C3-C4	-13.26	78.38	109.40
4	J	1001	CIT	C4-C3-C6	-13.06	82.04	110.11

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	J	1001	CIT	C2-C3-C6	-12.21	83.86	110.11

There are no chirality outliers.

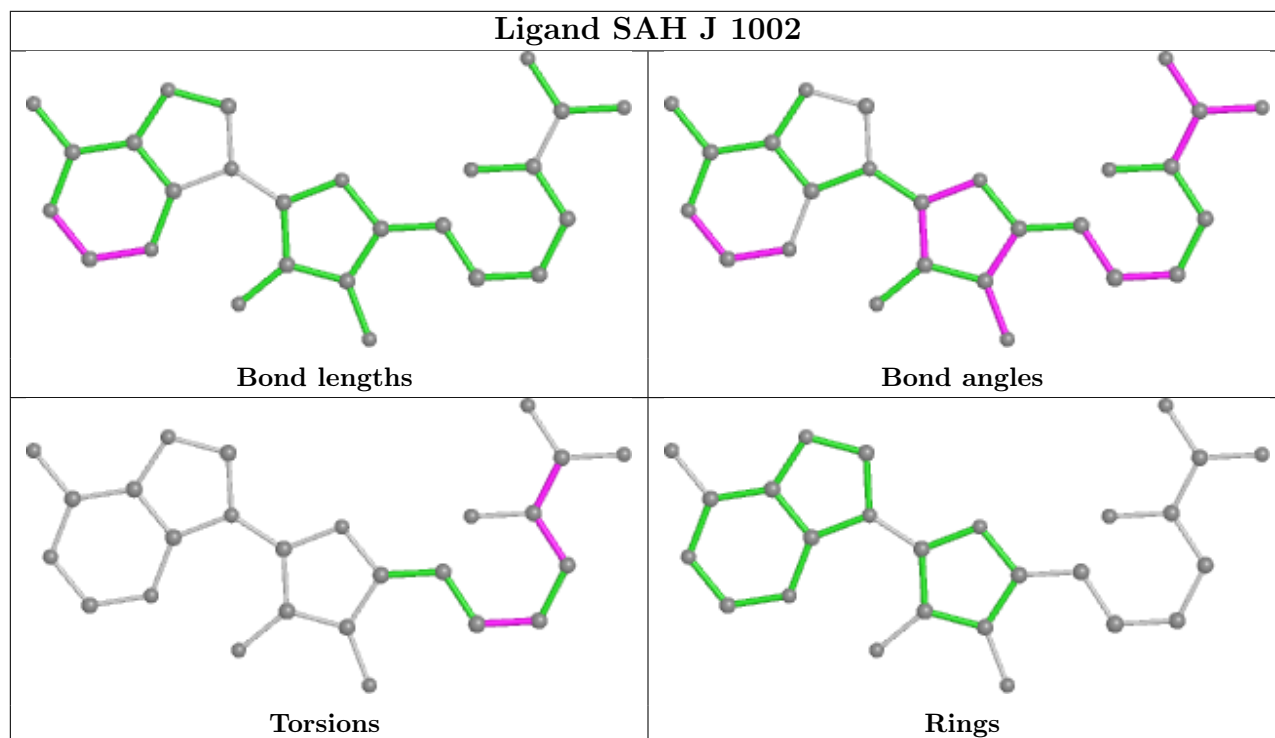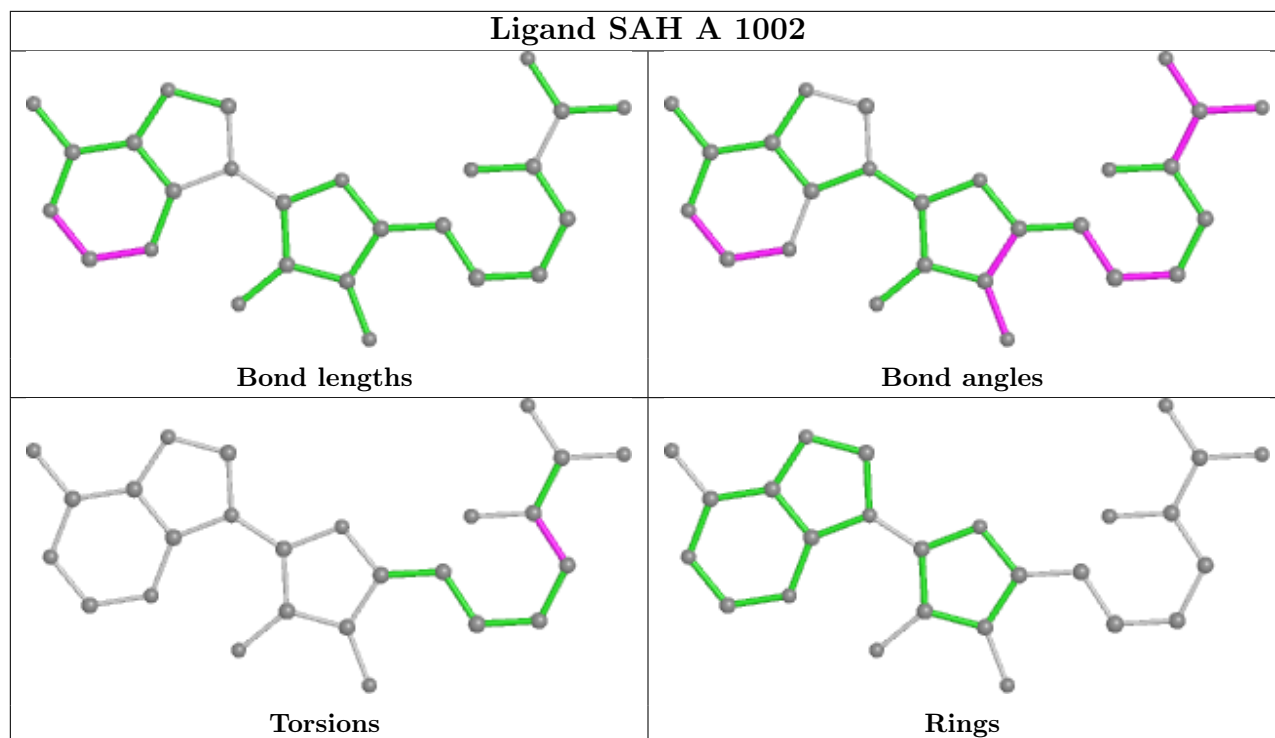5 of 29 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	1001	CIT	O7-C3-C4-C5
4	D	1001	CIT	C6-C3-C4-C5
4	J	1001	CIT	C1-C2-C3-O7
4	J	1001	CIT	O7-C3-C4-C5
4	J	1001	CIT	C6-C3-C4-C5

There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1001	CIT	1	0
4	D	1001	CIT	3	0
5	J	1002	SAH	1	0
5	G	1002	SAH	1	0
4	J	1001	CIT	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 SAH A 1002

Bond lengths

Bond angles

Torsions

Rings

Ligand SAH J 1002

Bond lengths

Bond angles

Torsions

Rings

Ligand SAH G 1002

Bond lengths

Bond angles

Torsions

Rings

Ligand SAH D 1002

Bond lengths

Bond angles

Torsions

Rings
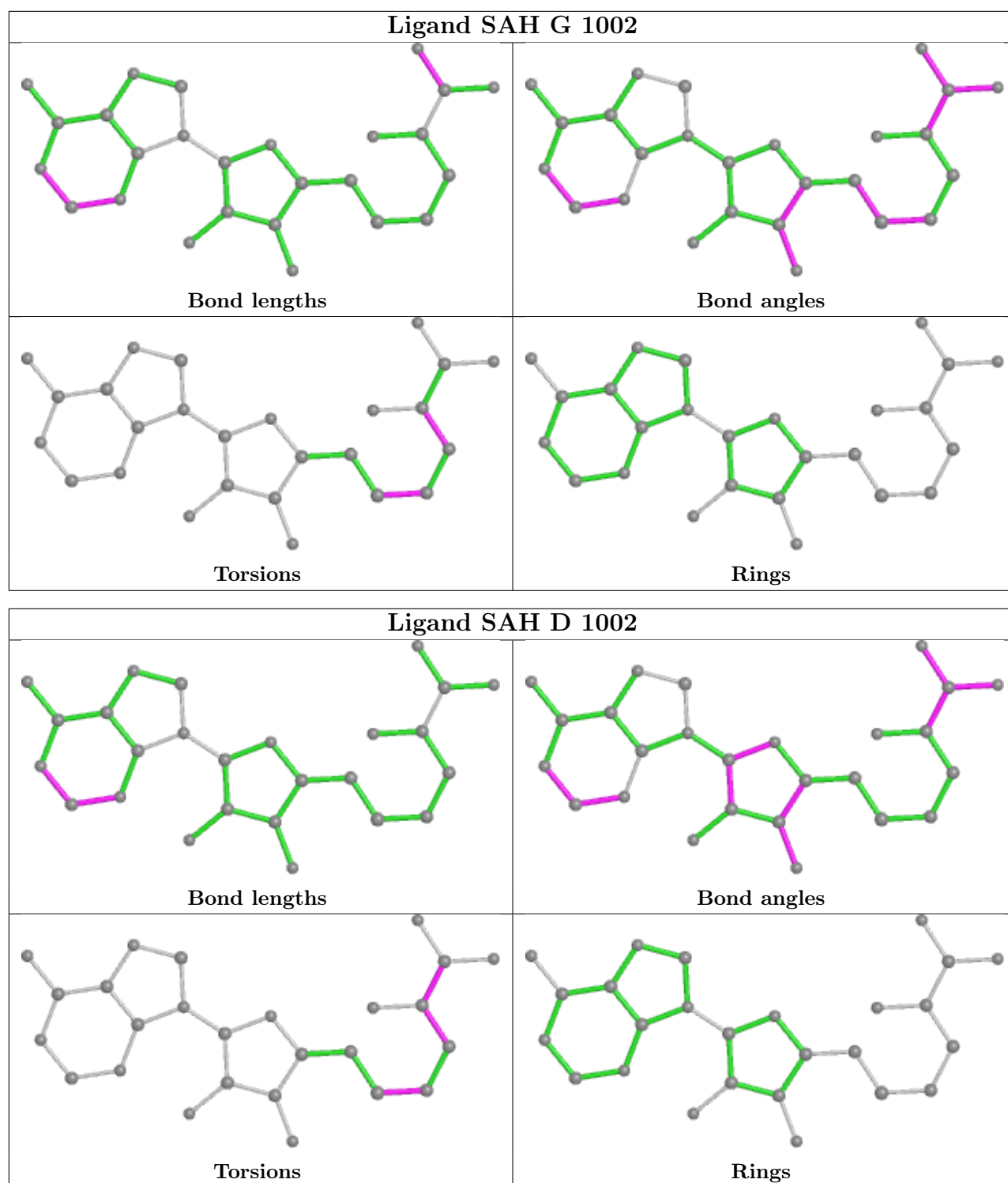
5.7 Other polymers (i)

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, 95th 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(Å²)	Q<0.9
1	A	285/285 (100%)	-0.02	1 (0%)	92	93	44, 63, 94, 159	0
1	D	285/285 (100%)	-0.08	2 (0%)	87	89	42, 61, 93, 137	0
1	G	284/285 (99%)	-0.05	1 (0%)	92	93	44, 63, 90, 133	0
1	J	285/285 (100%)	-0.09	2 (0%)	87	89	42, 61, 91, 148	0
2	B	194/209 (92%)	0.63	27 (13%)	2	2	68, 109, 172, 198	0
2	C	178/209 (85%)	0.55	24 (13%)	3	2	72, 135, 188, 200	0
2	H	197/209 (94%)	0.74	35 (17%)	1	1	30, 111, 176, 196	0
2	I	182/209 (87%)	0.87	33 (18%)	1	1	73, 140, 189, 205	0
3	E	23/25 (92%)	-0.46	1 (4%)	35	38	67, 124, 147, 154	0
3	F	24/25 (96%)	-0.66	0	100	100	84, 118, 159, 166	0
3	K	24/25 (96%)	-0.58	0	100	100	70, 127, 144, 187	0
3	L	24/25 (96%)	-0.61	0	100	100	81, 118, 155, 165	0
All	All	1985/2076 (95%)	0.20	126 (6%)	20	21	30, 80, 171, 205	0

The worst 5 of 126 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	I	337	LEU	20.0
2	I	338	VAL	14.3
2	H	332	SER	10.3
2	H	338	VAL	10.0
2	B	339	SER	9.1

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, 95th 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	L	427	19/20	0.98	0.12	61,70,82,84	0
3	PYO	F	427	19/20	0.99	0.14	66,77,82,83	0
3	PYO	K	427	19/20	0.99	0.14	51,65,74,77	0
3	PYO	E	427	19/20	0.99	0.14	54,66,73,77	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, 95th 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	CIT	A	1001	13/13	0.78	0.19	83,94,106,107	0
4	CIT	D	1001	13/13	0.85	0.19	97,103,112,120	0
4	CIT	J	1001	13/13	0.85	0.18	87,104,109,109	0
4	CIT	G	1001	13/13	0.88	0.23	86,103,107,110	0
5	SAH	A	1002	26/26	0.97	0.12	45,63,67,72	0
5	SAH	G	1002	26/26	0.97	0.11	48,61,70,73	0
5	SAH	J	1002	26/26	0.97	0.13	44,59,66,72	0
5	SAH	D	1002	26/26	0.98	0.13	44,58,64,69	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.
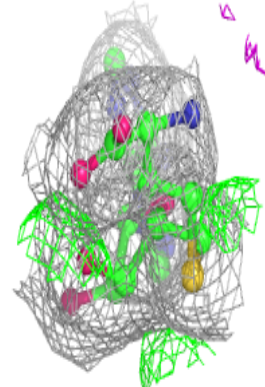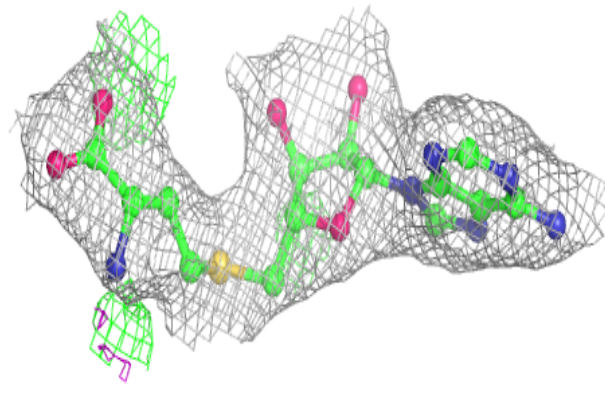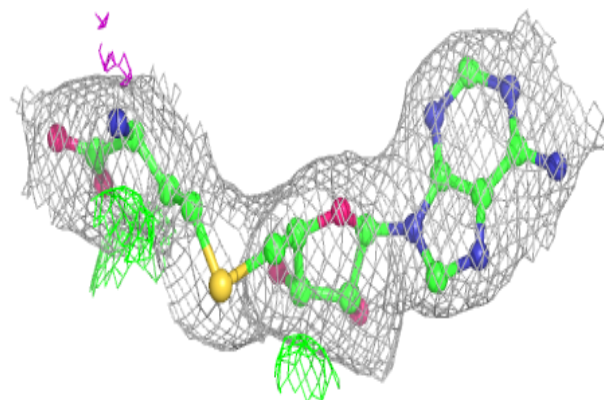
Electron density around SAH A 1002:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
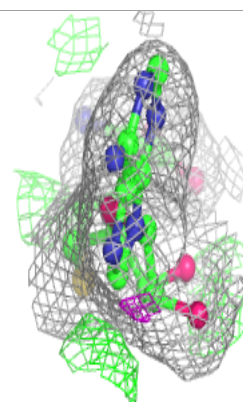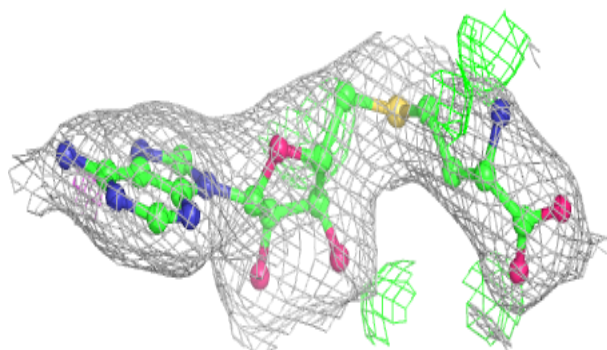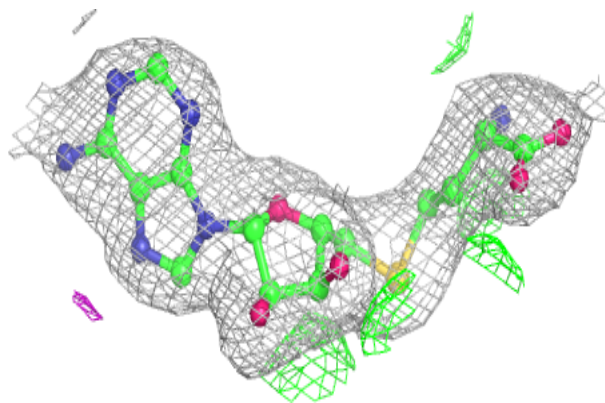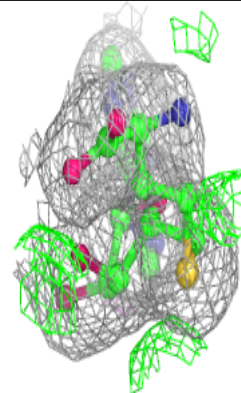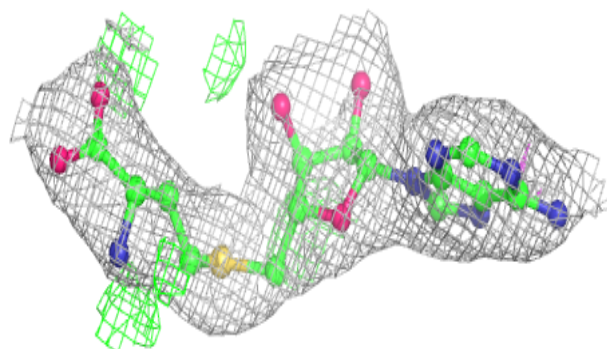and green (positive)
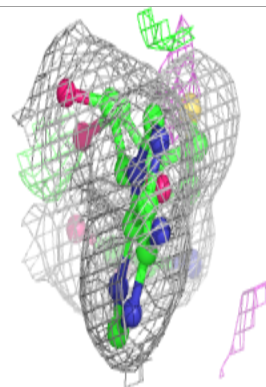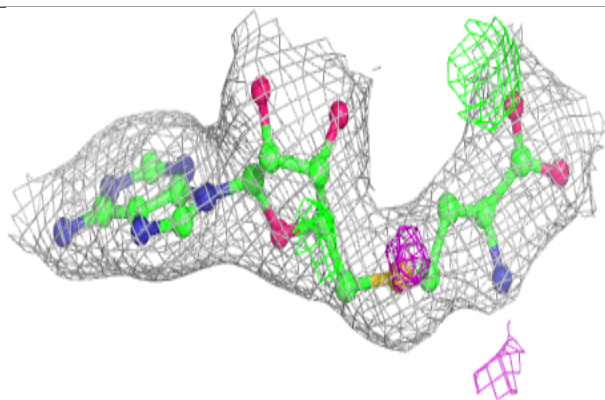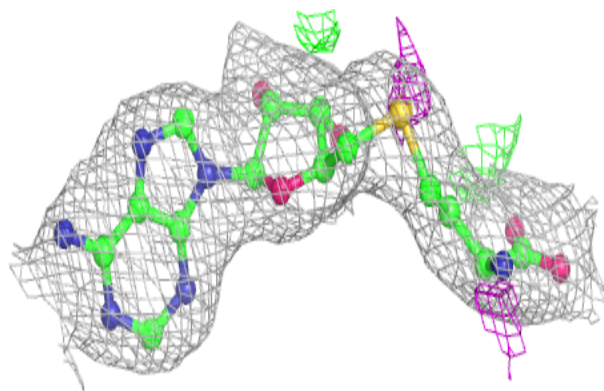
Electron density around SAH G 1002:

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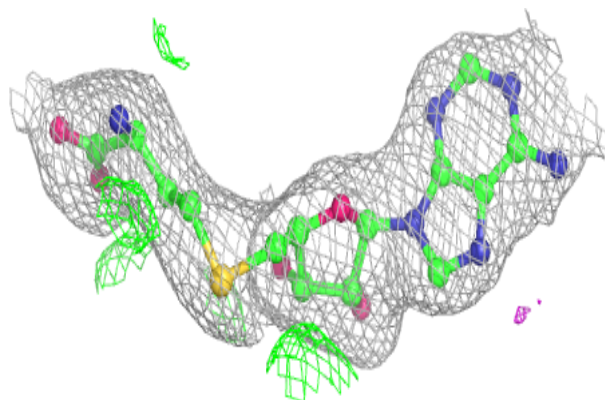and green (positive)

Electron density around SAH J 1002:

$2mF_o$-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 D 1002:

$2mF_o$-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 (i)

There are no such residues in this entry.