wwPDB X-ray Structure Validation Summary Report (i)

Jun 23, 2024 – 08:38 AM EDT

PDB ID	:	6I7P
Title	:	Crystal structure of the full-length Zika virus NS5 protein (Human isolate Z1106033)
Authors	:	Ferrero, D.S.; Ruiz-Arroyo, V.M.; Soler, N.; Uson, I.; Verdaguer, N.
Deposited on	:	2018-11-16
Resolution	:	3.98 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

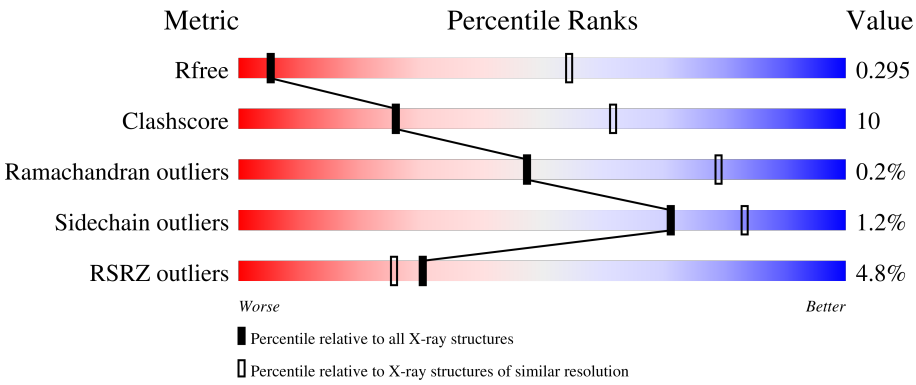
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.98 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	1039 (4.26-3.70)
Clashscore	141614	1099 (4.26-3.70)
Ramachandran outliers	138981	1061 (4.26-3.70)
Sidechain outliers	138945	1053 (4.26-3.70)
RSRZ outliers	127900	1021 (4.30-3.66)

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	914	6% 73% 24% •
1	B	914	7% 75% 21% •
1	C	914	7% 76% 21% •
1	D	914	2% 72% 25% •
1	E	914	2% 72% 24% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	914	4% 73% 23%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	PO4	D	1003	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 42817 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 NS5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	883	7097	4463	1286	1302	46	0	0	0
1	B	882	7090	4459	1285	1300	46	0	0	0
1	C	883	7097	4463	1286	1302	46	0	0	0
1	D	883	7097	4463	1286	1302	46	0	0	0
1	E	883	7097	4463	1286	1302	46	0	0	0
1	F	883	7097	4463	1286	1302	46	0	0	0

There are 66 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP A0A1B2ZC85
A	904	GLY	-	expression tag	UNP A0A1B2ZC85
A	905	SER	-	expression tag	UNP A0A1B2ZC85
A	906	SER	-	expression tag	UNP A0A1B2ZC85
A	907	SER	-	expression tag	UNP A0A1B2ZC85
A	908	HIS	-	expression tag	UNP A0A1B2ZC85
A	909	HIS	-	expression tag	UNP A0A1B2ZC85
A	910	HIS	-	expression tag	UNP A0A1B2ZC85
A	911	HIS	-	expression tag	UNP A0A1B2ZC85
A	912	HIS	-	expression tag	UNP A0A1B2ZC85
A	913	HIS	-	expression tag	UNP A0A1B2ZC85
B	0	MET	-	initiating methionine	UNP A0A1B2ZC85
B	904	GLY	-	expression tag	UNP A0A1B2ZC85
B	905	SER	-	expression tag	UNP A0A1B2ZC85
B	906	SER	-	expression tag	UNP A0A1B2ZC85
B	907	SER	-	expression tag	UNP A0A1B2ZC85
B	908	HIS	-	expression tag	UNP A0A1B2ZC85

Continued on next page...

Continued from previous page...

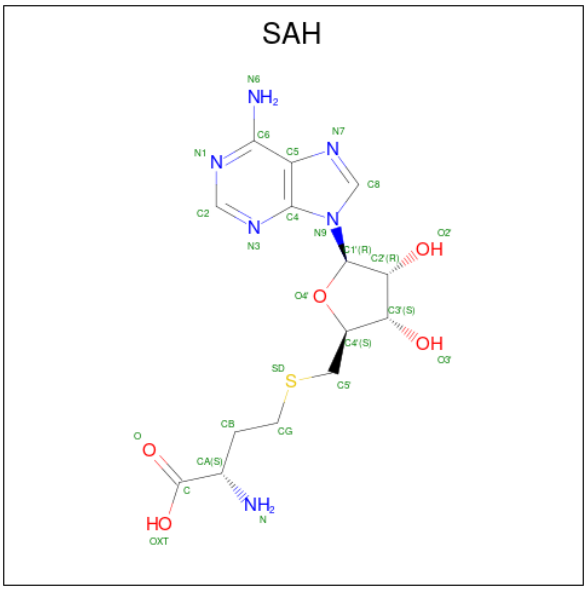
Chain	Residue	Modelled	Actual	Comment	Reference
B	909	HIS	-	expression tag	UNP A0A1B2ZC85
B	910	HIS	-	expression tag	UNP A0A1B2ZC85
B	911	HIS	-	expression tag	UNP A0A1B2ZC85
B	912	HIS	-	expression tag	UNP A0A1B2ZC85
B	913	HIS	-	expression tag	UNP A0A1B2ZC85
C	0	MET	-	initiating methionine	UNP A0A1B2ZC85
C	904	GLY	-	expression tag	UNP A0A1B2ZC85
C	905	SER	-	expression tag	UNP A0A1B2ZC85
C	906	SER	-	expression tag	UNP A0A1B2ZC85
C	907	SER	-	expression tag	UNP A0A1B2ZC85
C	908	HIS	-	expression tag	UNP A0A1B2ZC85
C	909	HIS	-	expression tag	UNP A0A1B2ZC85
C	910	HIS	-	expression tag	UNP A0A1B2ZC85
C	911	HIS	-	expression tag	UNP A0A1B2ZC85
C	912	HIS	-	expression tag	UNP A0A1B2ZC85
C	913	HIS	-	expression tag	UNP A0A1B2ZC85
D	0	MET	-	initiating methionine	UNP A0A1B2ZC85
D	904	GLY	-	expression tag	UNP A0A1B2ZC85
D	905	SER	-	expression tag	UNP A0A1B2ZC85
D	906	SER	-	expression tag	UNP A0A1B2ZC85
D	907	SER	-	expression tag	UNP A0A1B2ZC85
D	908	HIS	-	expression tag	UNP A0A1B2ZC85
D	909	HIS	-	expression tag	UNP A0A1B2ZC85
D	910	HIS	-	expression tag	UNP A0A1B2ZC85
D	911	HIS	-	expression tag	UNP A0A1B2ZC85
D	912	HIS	-	expression tag	UNP A0A1B2ZC85
D	913	HIS	-	expression tag	UNP A0A1B2ZC85
E	0	MET	-	initiating methionine	UNP A0A1B2ZC85
E	904	GLY	-	expression tag	UNP A0A1B2ZC85
E	905	SER	-	expression tag	UNP A0A1B2ZC85
E	906	SER	-	expression tag	UNP A0A1B2ZC85
E	907	SER	-	expression tag	UNP A0A1B2ZC85
E	908	HIS	-	expression tag	UNP A0A1B2ZC85
E	909	HIS	-	expression tag	UNP A0A1B2ZC85
E	910	HIS	-	expression tag	UNP A0A1B2ZC85
E	911	HIS	-	expression tag	UNP A0A1B2ZC85
E	912	HIS	-	expression tag	UNP A0A1B2ZC85
E	913	HIS	-	expression tag	UNP A0A1B2ZC85
F	0	MET	-	initiating methionine	UNP A0A1B2ZC85
F	904	GLY	-	expression tag	UNP A0A1B2ZC85
F	905	SER	-	expression tag	UNP A0A1B2ZC85
F	906	SER	-	expression tag	UNP A0A1B2ZC85

Continued on next page...

Continued from previous page...

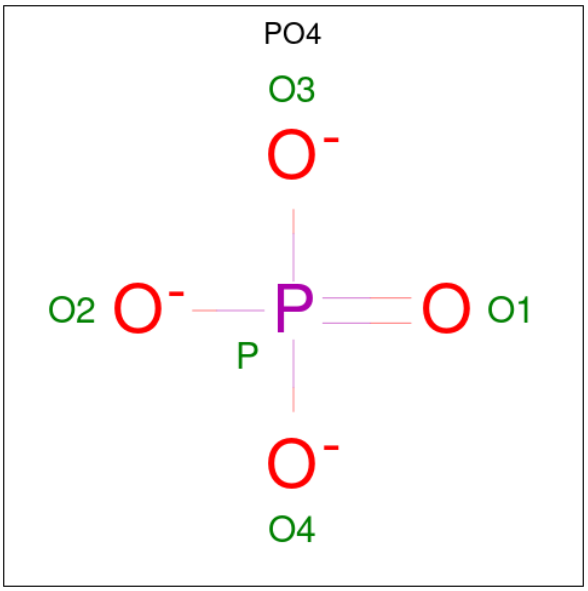
Chain	Residue	Modelled	Actual	Comment	Reference
F	907	SER	-	expression tag	UNP A0A1B2ZC85
F	908	HIS	-	expression tag	UNP A0A1B2ZC85
F	909	HIS	-	expression tag	UNP A0A1B2ZC85
F	910	HIS	-	expression tag	UNP A0A1B2ZC85
F	911	HIS	-	expression tag	UNP A0A1B2ZC85
F	912	HIS	-	expression tag	UNP A0A1B2ZC85
F	913	HIS	-	expression tag	UNP A0A1B2ZC85

- Molecule 2 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: $C_{14}H_{20}N_6O_5S$).



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

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	O	P			
3	A	1	5	4	1		0	0
			Total	O	P			
3	A	1	5	4	1		0	0
			Total	O	P			
3	B	1	5	4	1		0	0
			Total	O	P			
3	B	1	5	4	1		0	0
			Total	O	P			
3	C	1	5	4	1		0	0
			Total	O	P			
3	D	1	5	4	1		0	0
			Total	O	P			
3	D	1	5	4	1		0	0
			Total	O	P			
3	D	1	5	4	1		0	0
			Total	O	P			
3	D	1	5	4	1		0	0
			Total	O	P			
3	E	1	5	4	1		0	0
			Total	O	P			
3	E	1	5	4	1		0	0
			Total	O	P			
3	F	1	5	4	1		0	0
			Total	O	P			
3	F	1	5	4	1		0	0
			Total	O	P			
3	F	1	5	4	1		0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	P		
3	F	1	5	4	1	0	0

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

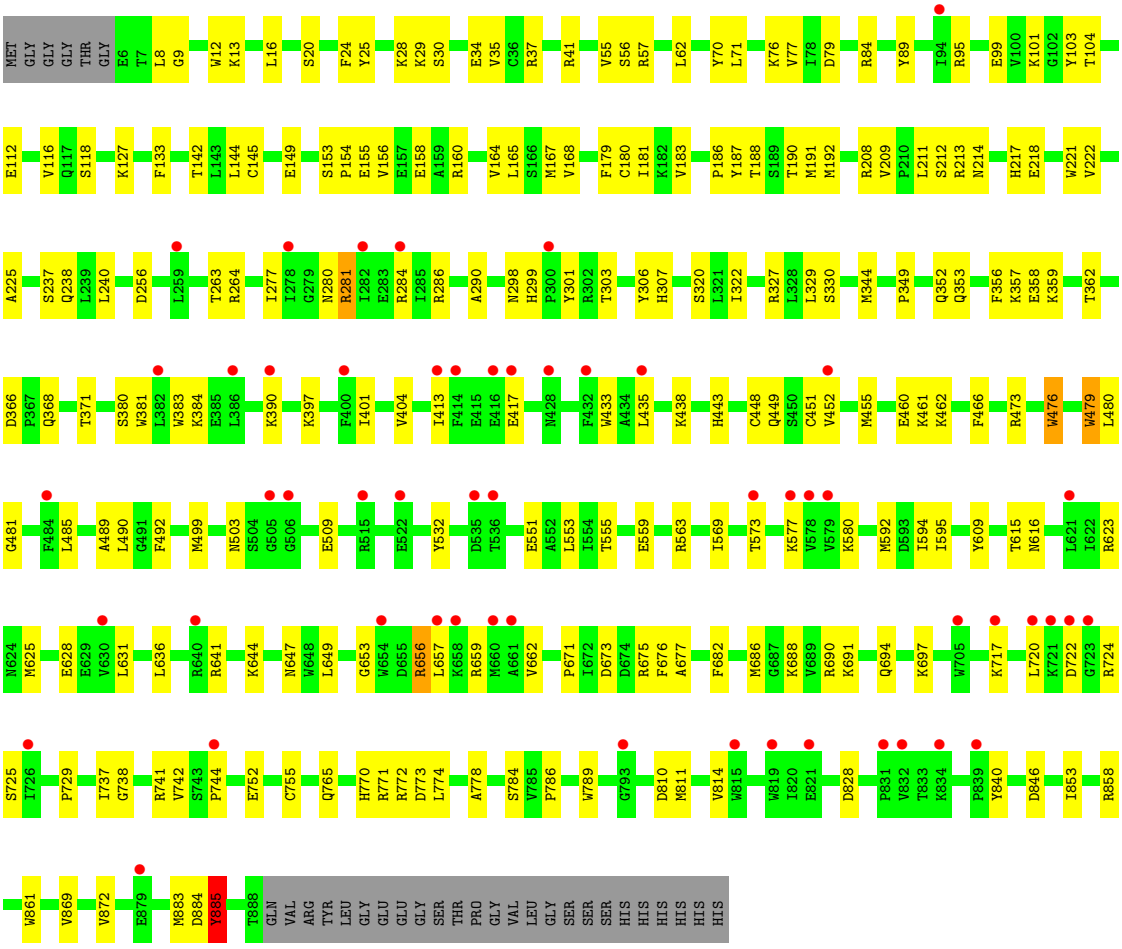
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
4	A	2	2	2	0	0
4	B	1	1	1	0	0
4	C	2	2	2	0	0
4	D	2	2	2	0	0
4	E	2	2	2	0	0
4	F	2	2	2	0	0

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.

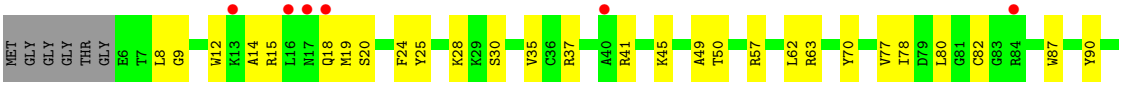
- Molecule 1: NS5

Chain A:



- Molecule 1: NS5

Chain B:



K101 G102 Y103 T104 L115 V116 Q117 G120 I123 L126 F133 C140 D141 T142 C145 D146 I147 S153 V156 E157 E158 A159 R160 R163 V164 L165 V168 W171 R175 I181 L184 C185 P186 Y187 R197 L198 Q199 R200 L206 V207 R208 N214

H217 E218 M219 Y220 A225 K226 S227 K231 S237 Q238 L239 M244 E254 R264 I278 R281 I282 E283 R284 H289 T292 W293 F294 W304 S317 S320 N323 L328 I342 A343 M344 P349 Y350 G351 Q352 Q353 R354 K357 E358 D366

P367 L382 W383 L386 G387 C395 T396 K397 F400 K403 V404 A412 I413 F414 E415 E416 E417 V427 N428 W433 V436 R440 E441 R445 G446 E447 C448 Q449 S450 C451 V452 Y453 N454 M455 K458 R459 E460 K461 K462 G471 W476 W479 L480 G481 A482

E488 A489 L490 L493 W498 G512 L513 Q514 R515 L516 M523 G529 A533 T536 I543 E551 T555 N556 Q557 M558 E559 R563 A566 L567 A568 I569 I570 Y574 Q575 N576 K577 V578 V579 K580 V581 L582 R583 P584 G588 K589 T590 V591 M592 I595

S596 R597 T613 F614 T615 N616 E628 E629 V630 L631 E632 M633 R641 K644 V645 T646 N647 W648 L649 G653 R656 L657 K658 R659 V662 C667 K670 D673 D674 R675 R681 F682 D685 K688 V689 R690 K697 N704 E707 H713 H714 L720

K721 D722 G723 R724 C730 R731 H732 Q733 G738 V742 S743 P744 L756 A757 M763 W764 Q765 Y768 F769 H770 D773 A778 S783 S784 V790 G793 G801 K802 G803 E804 W805 W819 N823 D828 L841 S851 R858 T868 V869 V872

M883 S887 THR GLN VAL ARG TYR LEU GLY GLU GLU GLY SER THR PRO GLY VAL LEU GLY SER SER SER HIS HIS HIS HIS HIS HIS

- Molecule 1: NS5

Chain C:

MET GLY GLY GLY THR GLY E6 W12 K13 A14 R15 L16 N17 Q18 M19 S20 F24 Y25 S26 Y27 K28 K29 S30 E34 V35 C36 R37 E38 V55 S56 V66 Y70 L71 Y74 I78 D79 L80 G81 C82 G83 R84 G85 G120 I123 V130 D131 V132 F133

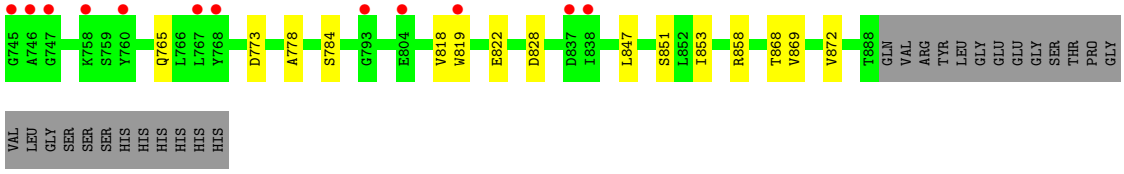
A137 C140 C145 D146 S153 P154 E155 V156 E157 E158 T161 L162 R163 V164 L165 S166 M167 V168 W171 R175 P176 G177 A178 F179 C180 V183 L184 C185 P186 Y187 T188 M191 M192 E193 T194 L195 E196 R197 L198 Q199 L206 V209 P210 L211 N214 V222 S223

S227 N228 T229 L241 M244 S261 G262 T263 V266 V267 K276 I277 I278 R281 I282 W293 F294 F295 N298 H299 P300 Y301 Y306 S309 T314 S317 A318 S319 S320 L321 I322 R327 L328 L329 I342 A343 M344 T345 D346 P349 Q352 Q353 R354

D366 P367 Q368 T371 W383 G387 P392 G411 A412 I413 F414 E415 E416 E417 V424 V427 F432 W433 V436 R440 E447 V452 Y453 N454 M455 R459 E460 K461 K462 Q463 K468 A469 K470 G471 S472 R473 A474 I475 W476 Y477 M478 W479 A482 L485

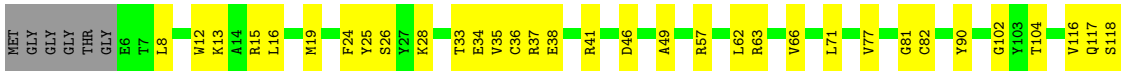
E488 A489 L490 G491 F492 S504 G505 G506 L513 Q514 R515 L516 G517 Y518 M523 S524 R525 A533 D534 F546 L553 I554 T555 N556 Q557 M558 E559 R563 N576 K577 V578 V579 K580 V581 P584 A585 E586 K587 T590 V591 M592 D593 I594 I595 G602 L611

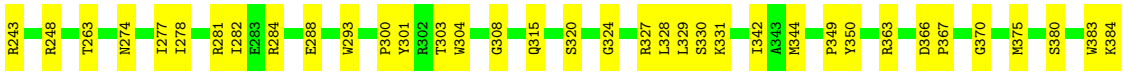
Q620 E629 V630 L631 W637 K644 V645 T646 N647 W648 L649 W654 D655 R656 L657 K658 R659 M660 A661 G664 C667 V668 V669 D673 D674 R675 F682 Q694 K697 D703 N704 W705 E706 N716 K717 L720 R724 S725 I726 V727 Q733 G738 P744

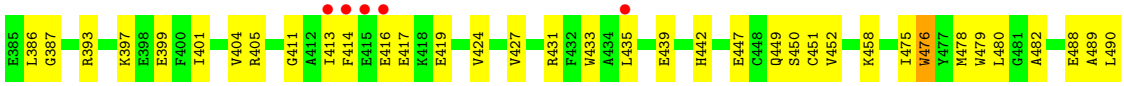


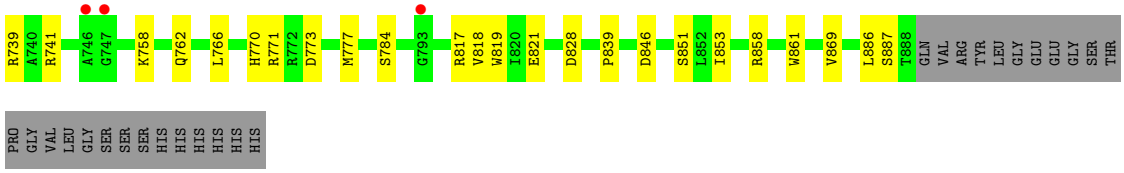
- Molecule 1: NS5

Chain D:



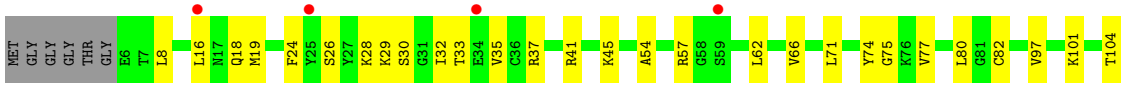


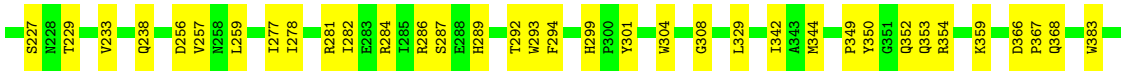


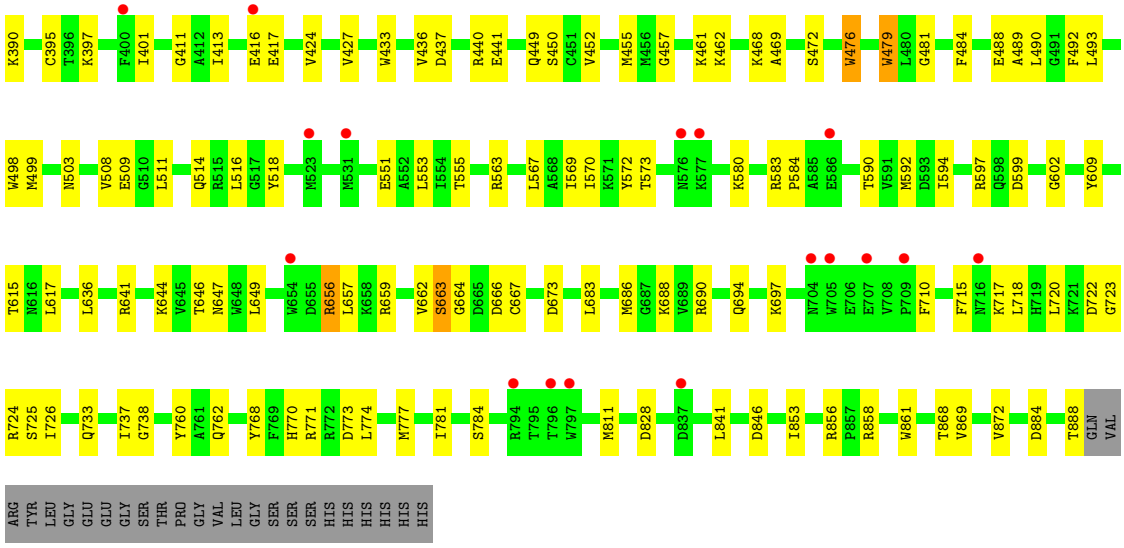


- Molecule 1: NS5

Chain E:

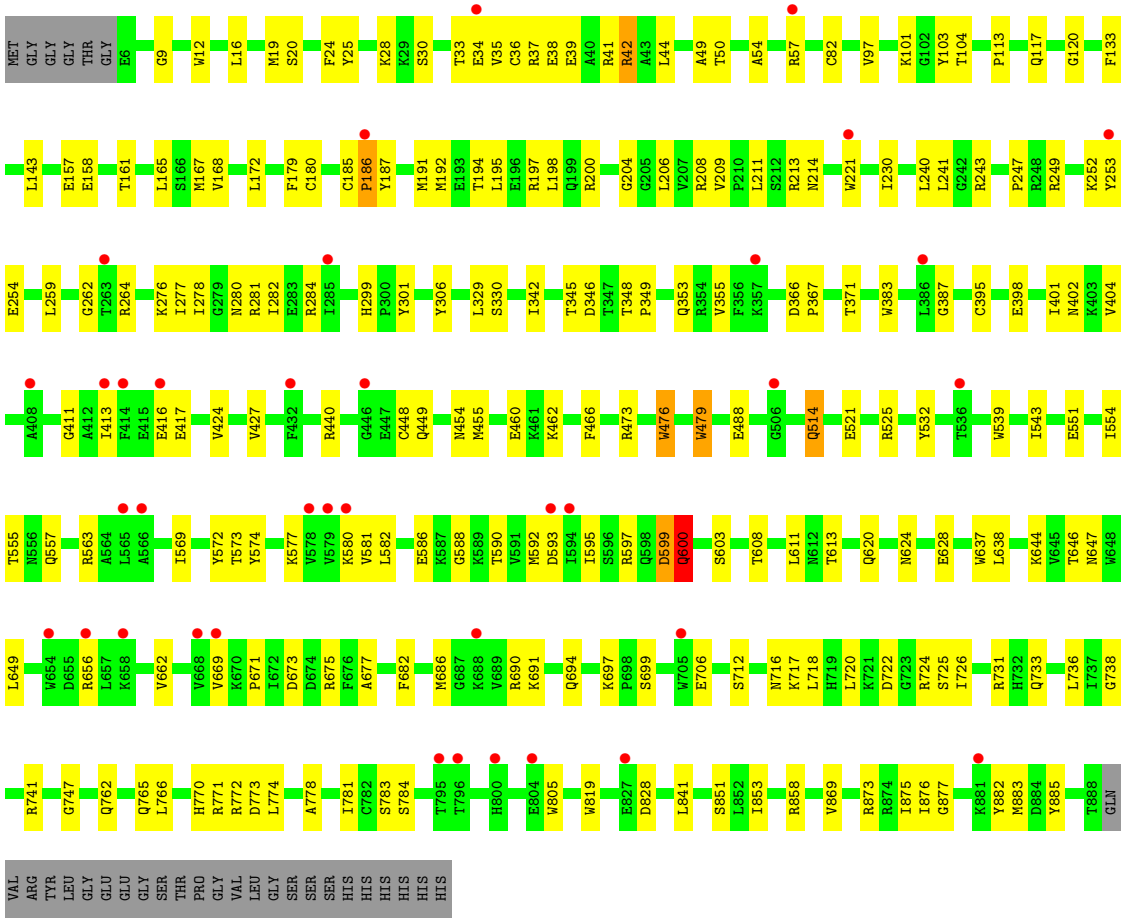






- Molecule 1: NS5

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	191.06Å 192.06Å 407.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.52 – 3.98 49.52 – 3.98	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.52-3.98) 98.2 (49.52-3.98)	Depositor EDS
R_{merge}	0.96	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.69 (at 4.00Å)	Xtriage
Refinement program	PHENIX (1.14_3260: ???)	Depositor
R, R_{free}	0.280 , 0.280 0.283 , 0.295	Depositor DCC
R_{free} test set	6430 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	77.2	Xtriage
Anisotropy	0.054	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 63.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.28$, $< L^2 > = 0.11$	Xtriage
Estimated twinning fraction	0.217 for k,h,-l	Xtriage
F_o,F_c correlation	0.77	EDS
Total number of atoms	42817	wwPDB-VP
Average B, all atoms (Å^2)	101.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.69% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, SAH, PO4

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	0.39	2/7262 (0.0%)	0.58	3/9813 (0.0%)
1	B	0.36	0/7255	0.54	0/9803
1	C	0.37	0/7262	0.55	0/9813
1	D	0.46	3/7262 (0.0%)	0.59	1/9813 (0.0%)
1	E	0.40	1/7262 (0.0%)	0.58	2/9813 (0.0%)
1	F	0.42	4/7262 (0.1%)	0.61	5/9813 (0.1%)
All	All	0.40	10/43565 (0.0%)	0.58	11/58868 (0.0%)

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	F	0	1
All	All	0	2

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	663	SER	C-N	-14.88	1.06	1.33
1	D	886	LEU	C-N	13.81	1.65	1.34
1	F	875	ILE	C-N	9.58	1.56	1.34
1	A	884	ASP	C-N	8.66	1.53	1.34
1	D	887	SER	C-N	8.37	1.53	1.34

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	F	599	ASP	O-C-N	-13.97	100.35	122.70
1	A	885	TYR	O-C-N	-12.07	103.39	122.70
1	F	599	ASP	C-N-CA	11.24	149.80	121.70
1	F	599	ASP	CA-C-N	9.71	138.57	117.20
1	E	663	SER	O-C-N	7.79	136.44	123.20

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	885	TYR	Mainchain
1	F	600	GLN	Mainchain

5.2 Too-close contacts (i)

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	7097	0	6997	179	0
1	B	7090	0	6993	139	0
1	C	7097	0	6998	131	0
1	D	7097	0	6997	147	0
1	E	7097	0	6998	147	0
1	F	7097	0	6998	159	0
2	A	26	0	19	1	0
2	B	26	0	19	1	0
2	C	26	0	19	3	0
2	D	26	0	19	1	0
2	E	26	0	19	1	0
2	F	26	0	19	0	0
3	A	10	0	0	0	0
3	B	10	0	0	0	0
3	C	5	0	0	0	0
3	D	20	0	0	3	0
3	E	10	0	0	0	0
3	F	20	0	0	2	0
4	A	2	0	0	0	0
4	B	1	0	0	0	0
4	C	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	2	0	0	0	0
4	E	2	0	0	0	0
4	F	2	0	0	0	0
All	All	42817	0	42095	891	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 10.

The worst 5 of 891 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:417:GLU:HB2	1:A:479:TRP:CH2	1.45	1.50
1:F:143:LEU:HD23	1:F:172:LEU:HD21	1.27	1.17
1:B:412:ALA:HA	1:B:417:GLU:OE1	1.45	1.16
1:A:359:LYS:NZ	1:A:473:ARG:CD	2.10	1.15
1:A:417:GLU:CB	1:A:479:TRP:CH2	2.33	1.11

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	881/914 (96%)	836 (95%)	45 (5%)	0	100	100
1	B	880/914 (96%)	837 (95%)	42 (5%)	1 (0%)	51	83
1	C	881/914 (96%)	832 (94%)	47 (5%)	2 (0%)	47	79
1	D	881/914 (96%)	829 (94%)	49 (6%)	3 (0%)	41	74
1	E	881/914 (96%)	834 (95%)	45 (5%)	2 (0%)	47	79
1	F	881/914 (96%)	835 (95%)	45 (5%)	1 (0%)	51	83
All	All	5285/5484 (96%)	5003 (95%)	273 (5%)	9 (0%)	47	79

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	36	CYS
1	D	451	CYS
1	B	186	PRO
1	C	186	PRO
1	E	186	PRO

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	756/779 (97%)	749 (99%)	7 (1%)	78	87
1	B	755/779 (97%)	748 (99%)	7 (1%)	78	87
1	C	756/779 (97%)	748 (99%)	8 (1%)	73	84
1	D	756/779 (97%)	747 (99%)	9 (1%)	71	83
1	E	756/779 (97%)	744 (98%)	12 (2%)	62	79
1	F	756/779 (97%)	746 (99%)	10 (1%)	69	81
All	All	4535/4674 (97%)	4482 (99%)	53 (1%)	71	83

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	479	TRP
1	E	462	LYS
1	F	514	GLN
1	D	514	GLN
1	E	281	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 49 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	503	ASN
1	E	18	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	576	ASN
1	D	762	GLN
1	E	323	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 32 ligands modelled in this entry, 11 are monoatomic - leaving 21 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	$\#\|Z\| > 2$
3	PO4	B	1003	-	4,4,4	0.97	0	6,6,6	0.45	0
3	PO4	D	1004	-	4,4,4	0.92	0	6,6,6	0.52	0
2	SAH	F	1001	-	23,28,28	1.28	2 (8%)	22,40,40	1.74	5 (22%)
2	SAH	D	1001	-	23,28,28	1.35	3 (13%)	22,40,40	2.04	4 (18%)
3	PO4	F	1003	-	4,4,4	0.95	0	6,6,6	0.37	0
2	SAH	A	1001	-	23,28,28	1.20	3 (13%)	22,40,40	1.87	4 (18%)
2	SAH	C	1001	-	23,28,28	1.25	3 (13%)	22,40,40	1.70	3 (13%)
3	PO4	E	1003	-	4,4,4	0.96	0	6,6,6	0.47	0
3	PO4	D	1003	-	4,4,4	0.98	0	6,6,6	0.51	0
3	PO4	F	1002	-	4,4,4	0.94	0	6,6,6	0.47	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	SAH	E	1001	-	23,28,28	1.35	3 (13%)	22,40,40	2.02	4 (18%)
3	PO4	C	1002	-	4,4,4	0.93	0	6,6,6	0.50	0
2	SAH	B	1001	-	23,28,28	1.35	3 (13%)	22,40,40	1.95	5 (22%)
3	PO4	A	1003	-	4,4,4	0.86	0	6,6,6	0.60	0
3	PO4	D	1002	-	4,4,4	0.91	0	6,6,6	0.44	0
3	PO4	E	1002	-	4,4,4	0.91	0	6,6,6	0.54	0
3	PO4	B	1002	-	4,4,4	0.95	0	6,6,6	0.50	0
3	PO4	A	1002	-	4,4,4	0.92	0	6,6,6	0.41	0
3	PO4	F	1005	-	4,4,4	0.91	0	6,6,6	0.50	0
3	PO4	F	1004	-	4,4,4	0.91	0	6,6,6	0.50	0
3	PO4	D	1005	-	4,4,4	0.85	0	6,6,6	0.54	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
2	SAH	B	1001	-	-	7/11/31/31	0/3/3/3
2	SAH	D	1001	-	-	7/11/31/31	0/3/3/3
2	SAH	A	1001	-	-	7/11/31/31	0/3/3/3
2	SAH	C	1001	-	-	4/11/31/31	0/3/3/3
2	SAH	F	1001	-	-	0/11/31/31	0/3/3/3
2	SAH	E	1001	-	-	3/11/31/31	0/3/3/3

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	1001	SAH	C2-N3	4.48	1.39	1.32
2	B	1001	SAH	C2-N3	4.45	1.39	1.32
2	D	1001	SAH	C2-N3	4.41	1.38	1.32
2	F	1001	SAH	C2-N3	4.39	1.38	1.32
2	C	1001	SAH	C2-N3	4.29	1.38	1.32

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	1001	SAH	N3-C2-N1	-6.13	120.35	128.67
2	A	1001	SAH	N3-C2-N1	-6.13	120.35	128.67
2	D	1001	SAH	N3-C2-N1	-6.01	120.52	128.67

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	1001	SAH	N3-C2-N1	-5.85	120.74	128.67
2	B	1001	SAH	N3-C2-N1	-5.67	120.98	128.67

There are no chirality outliers.

5 of 28 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1001	SAH	N-CA-CB-CG
2	B	1001	SAH	C-CA-CB-CG
2	B	1001	SAH	OXT-C-CA-N
2	E	1001	SAH	CA-CB-CG-SD
2	C	1001	SAH	CA-CB-CG-SD

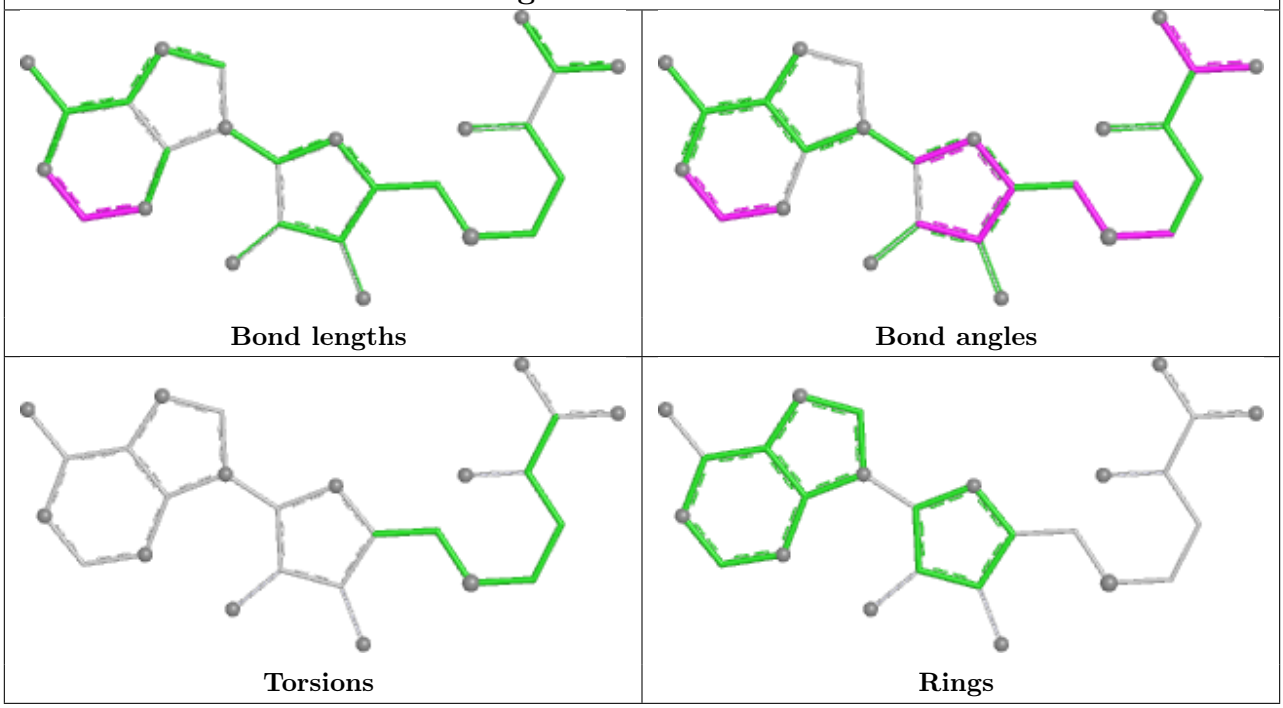
There are no ring outliers.

9 monomers are involved in 12 short contacts:

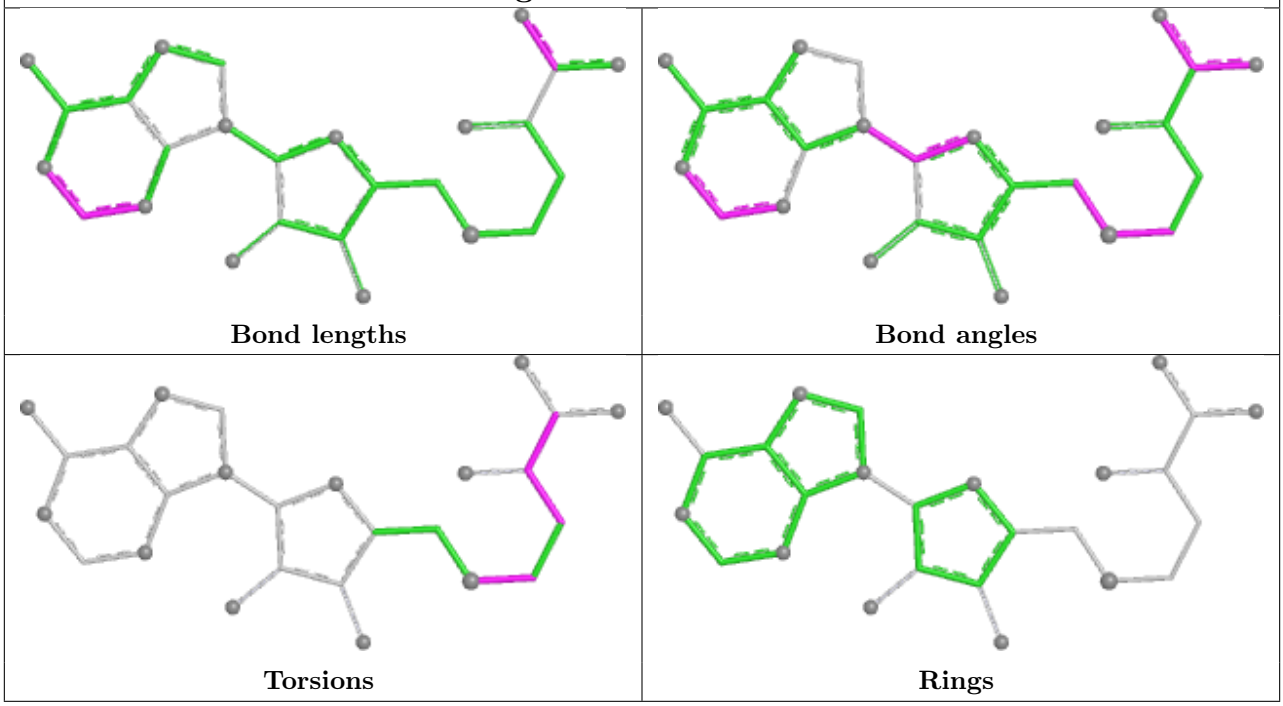
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1001	SAH	1	0
3	F	1003	PO4	1	0
2	A	1001	SAH	1	0
2	C	1001	SAH	3	0
3	D	1003	PO4	2	0
3	F	1002	PO4	1	0
2	E	1001	SAH	1	0
2	B	1001	SAH	1	0
3	D	1005	PO4	1	0

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.

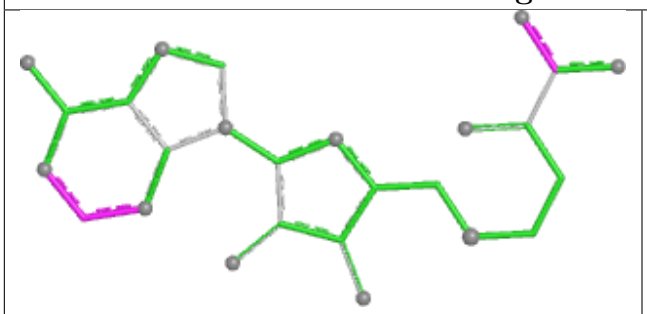
Ligand SAH F 1001



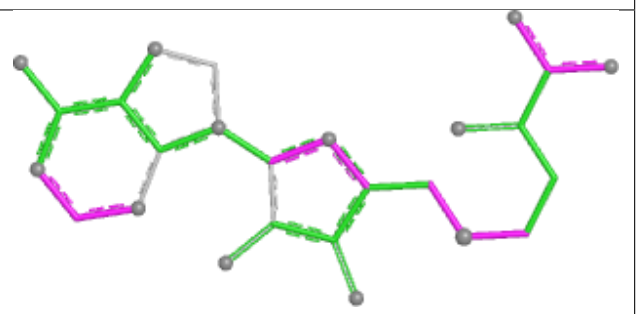
Ligand SAH D 1001



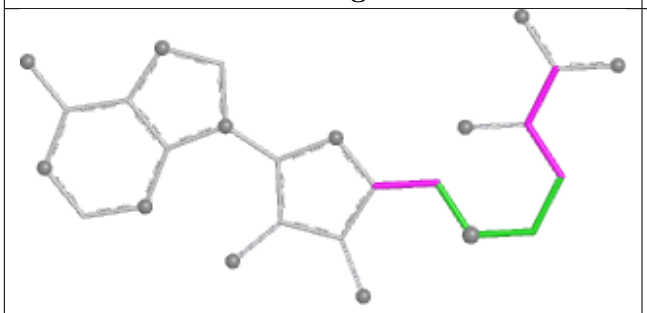
Ligand SAH A 1001



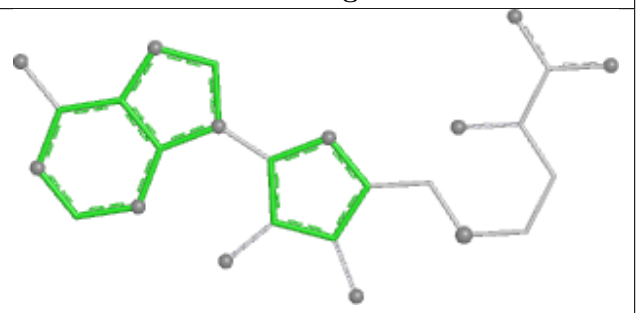
Bond lengths



Bond angles

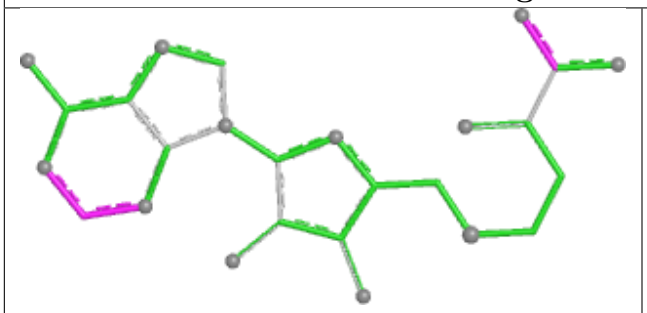


Torsions

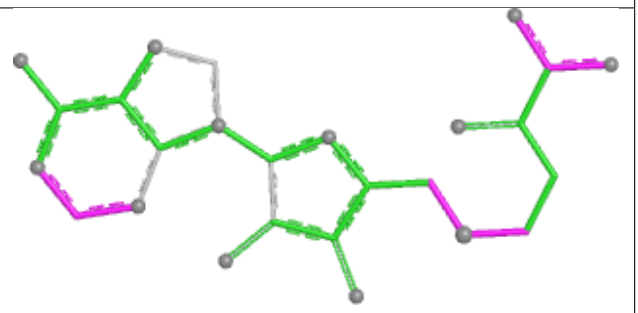


Rings

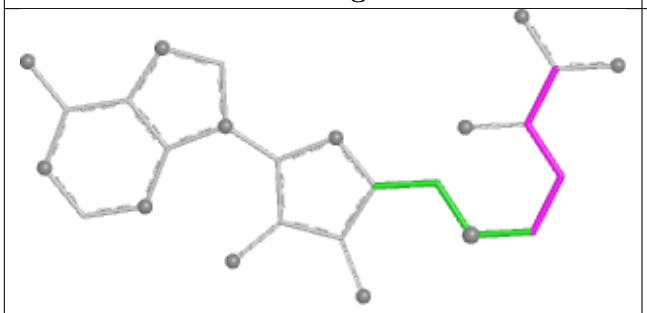
Ligand SAH C 1001



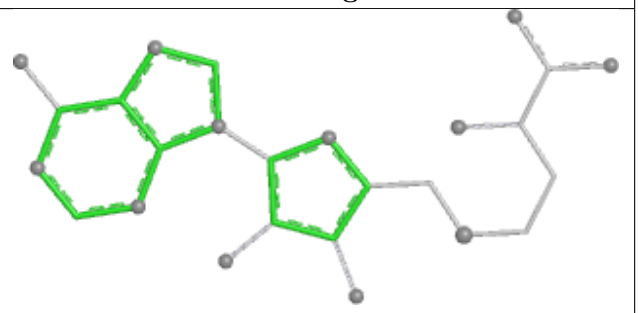
Bond lengths



Bond angles

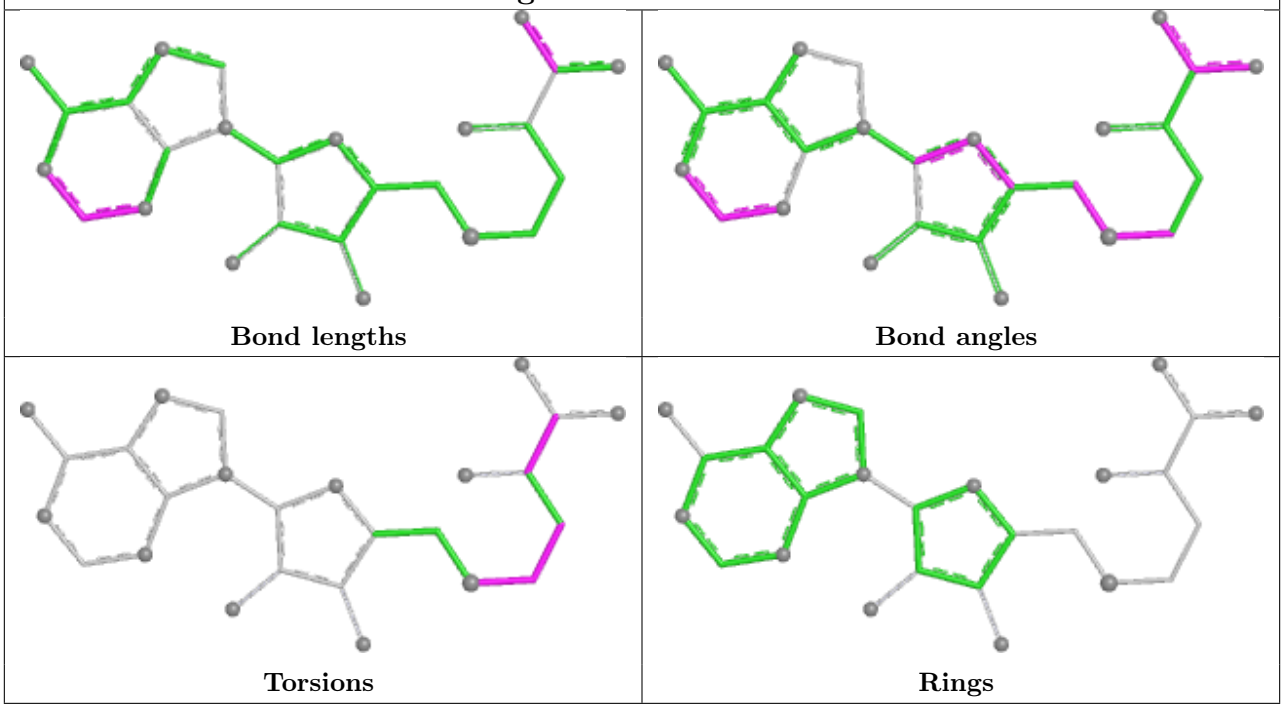


Torsions

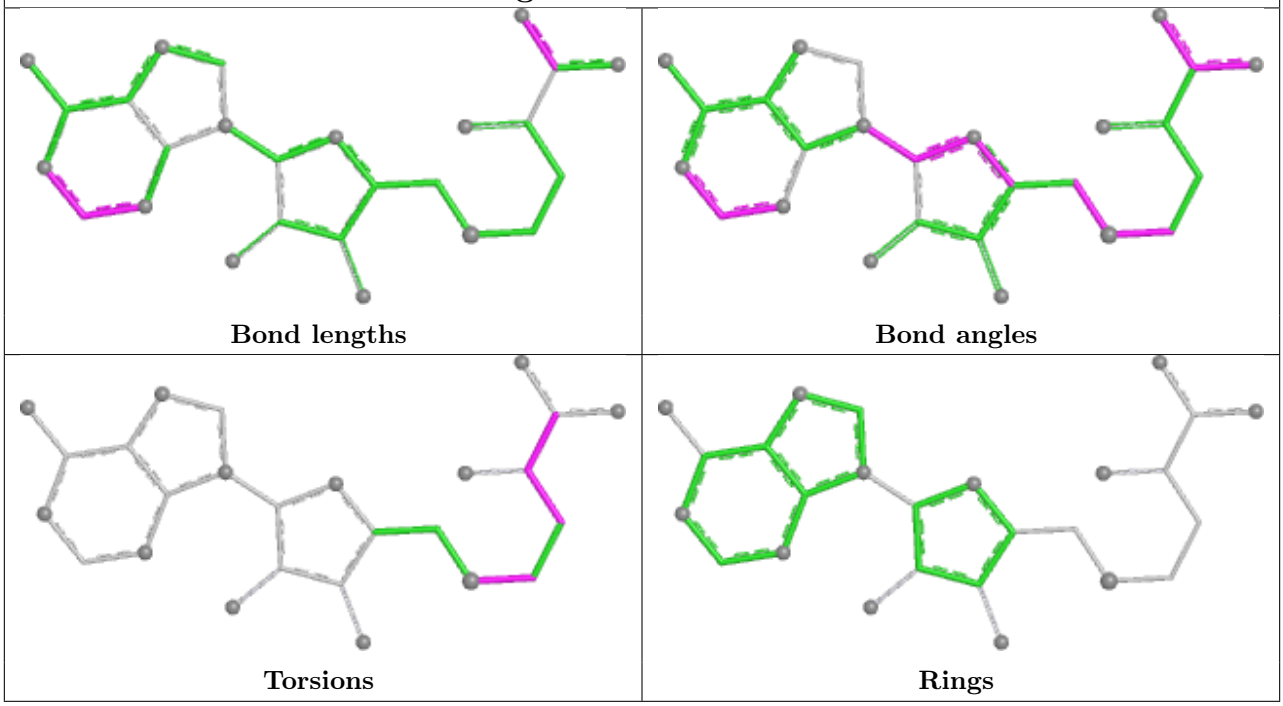


Rings

Ligand SAH E 1001



Ligand SAH B 1001



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	D	2
1	F	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	886:LEU	C	887:SER	N	1.65
1	F	876:ILE	C	877:GLY	N	1.18
1	F	599:ASP	C	600:GLN	N	1.16
1	D	663:SER	C	664:GLY	N	1.06

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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(Å^2)	Q<0.9
1	A	883/914 (96%)	0.47	54 (6%)	21	17	20, 104, 182, 261	0
1	B	882/914 (96%)	0.56	63 (7%)	16	13	26, 111, 191, 253	0
1	C	883/914 (96%)	0.54	62 (7%)	16	13	18, 108, 166, 222	0
1	D	883/914 (96%)	0.35	18 (2%)	65	56	17, 77, 131, 187	0
1	E	883/914 (96%)	0.36	21 (2%)	59	49	19, 86, 156, 221	0
1	F	883/914 (96%)	0.40	37 (4%)	36	29	27, 91, 149, 195	0
All	All	5297/5484 (96%)	0.45	255 (4%)	30	25	17, 96, 171, 261	0

The worst 5 of 255 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	578	VAL	7.4
1	A	723	GLY	6.9
1	B	446	GLY	6.8
1	B	452	VAL	6.4
1	A	720	LEU	5.9

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

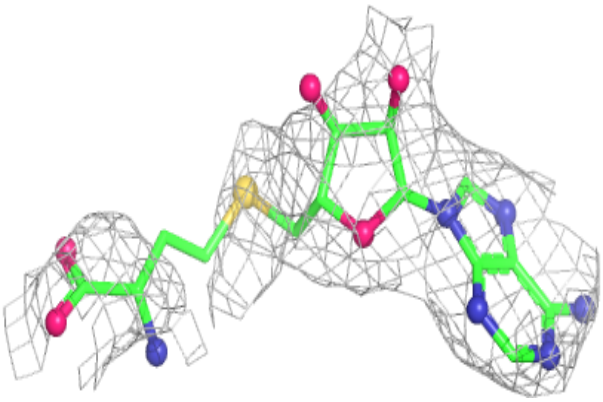
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

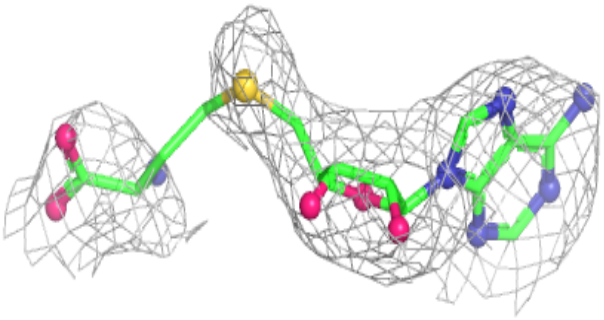
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	PO4	E	1003	5/5	0.39	0.24	168,181,230,239	0
3	PO4	F	1005	5/5	0.39	0.26	163,195,212,227	0
3	PO4	D	1002	5/5	0.67	0.24	123,143,182,184	0
3	PO4	C	1002	5/5	0.73	0.23	110,129,154,157	0
3	PO4	D	1004	5/5	0.73	0.17	156,179,206,209	0
3	PO4	B	1003	5/5	0.74	0.17	157,184,212,213	0
3	PO4	A	1002	5/5	0.74	0.26	127,148,153,161	0
3	PO4	D	1005	5/5	0.78	0.24	119,125,139,152	0
3	PO4	F	1003	5/5	0.81	0.18	103,138,171,194	0
3	PO4	D	1003	5/5	0.82	0.15	151,169,206,225	0
3	PO4	F	1004	5/5	0.89	0.19	79,90,144,173	0
2	SAH	B	1001	26/26	0.89	0.52	68,105,131,138	0
2	SAH	C	1001	26/26	0.90	0.39	43,74,103,125	0
3	PO4	B	1002	5/5	0.90	0.16	94,109,127,135	0
2	SAH	D	1001	26/26	0.92	0.41	31,85,110,128	0
2	SAH	F	1001	26/26	0.92	0.38	43,81,100,140	0
2	SAH	A	1001	26/26	0.93	0.34	44,65,92,141	0
3	PO4	E	1002	5/5	0.94	0.18	78,104,111,113	0
3	PO4	A	1003	5/5	0.95	0.15	88,95,130,137	0
2	SAH	E	1001	26/26	0.95	0.43	38,67,106,137	0
3	PO4	F	1002	5/5	0.95	0.18	70,78,116,131	0
4	ZN	A	1005	1/1	0.97	0.15	154,154,154,154	0
4	ZN	C	1003	1/1	0.97	0.22	68,68,68,68	0
4	ZN	E	1004	1/1	0.97	0.19	156,156,156,156	0
4	ZN	A	1004	1/1	0.98	0.20	73,73,73,73	0
4	ZN	C	1004	1/1	0.99	0.16	142,142,142,142	0
4	ZN	D	1006	1/1	0.99	0.20	44,44,44,44	0
4	ZN	B	1004	1/1	0.99	0.21	32,32,32,32	0
4	ZN	E	1005	1/1	0.99	0.22	10,10,10,10	0
4	ZN	F	1006	1/1	0.99	0.17	57,57,57,57	0
4	ZN	F	1007	1/1	0.99	0.16	118,118,118,118	0
4	ZN	D	1007	1/1	1.00	0.24	51,51,51,51	0

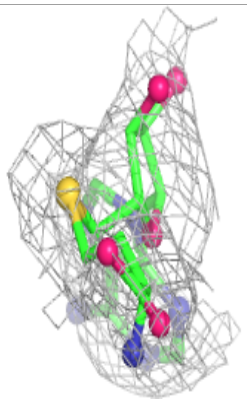
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around SAH B 1001:

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

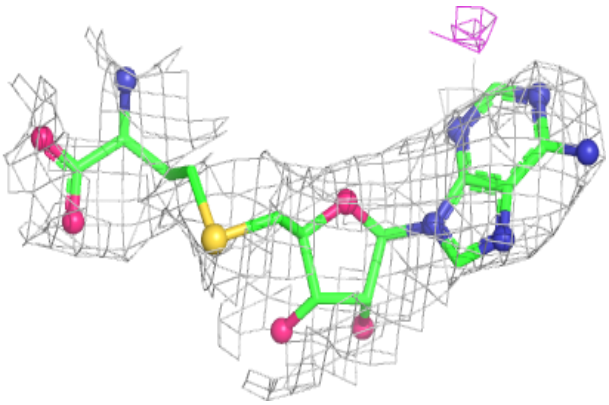


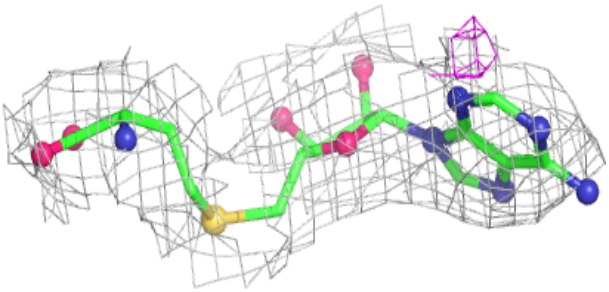


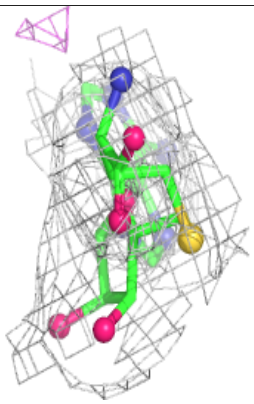


Electron density around SAH C 1001:

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

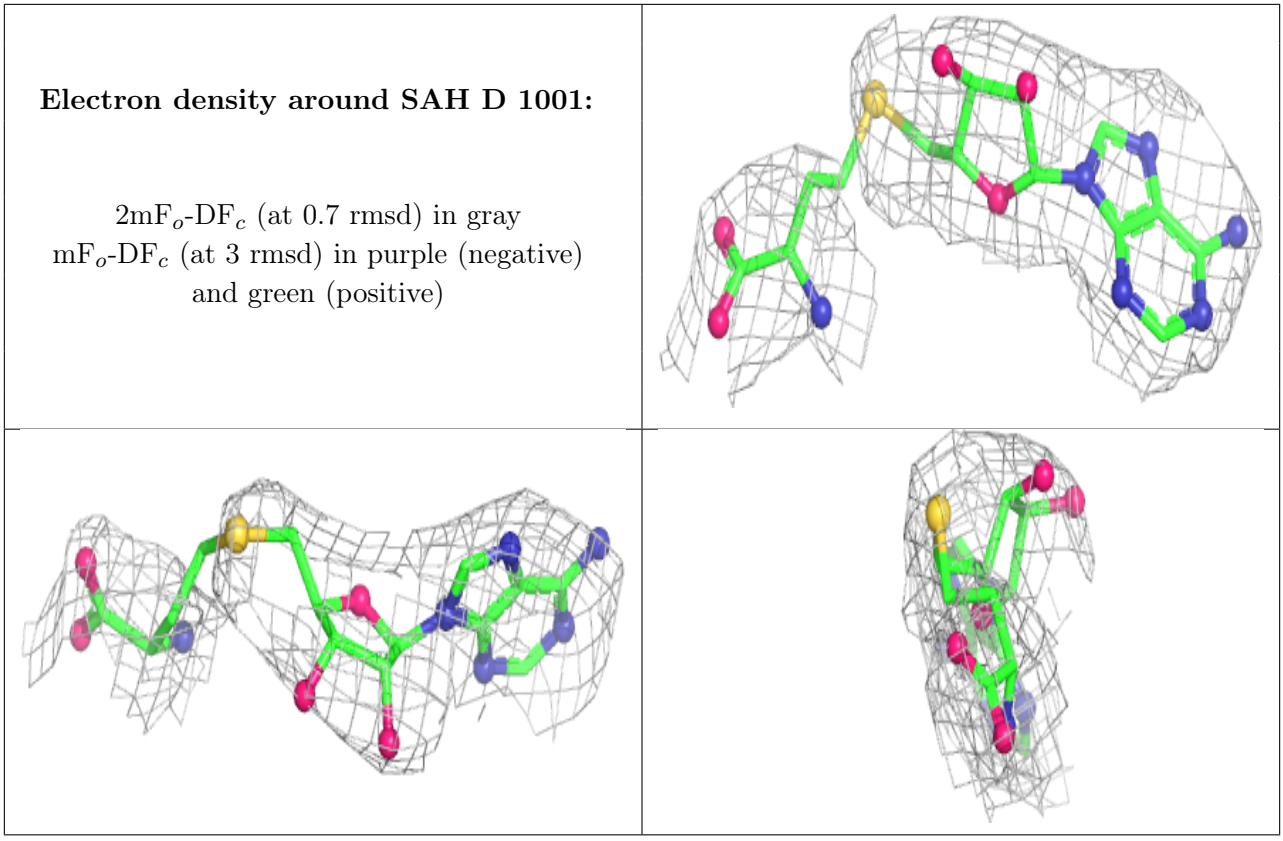






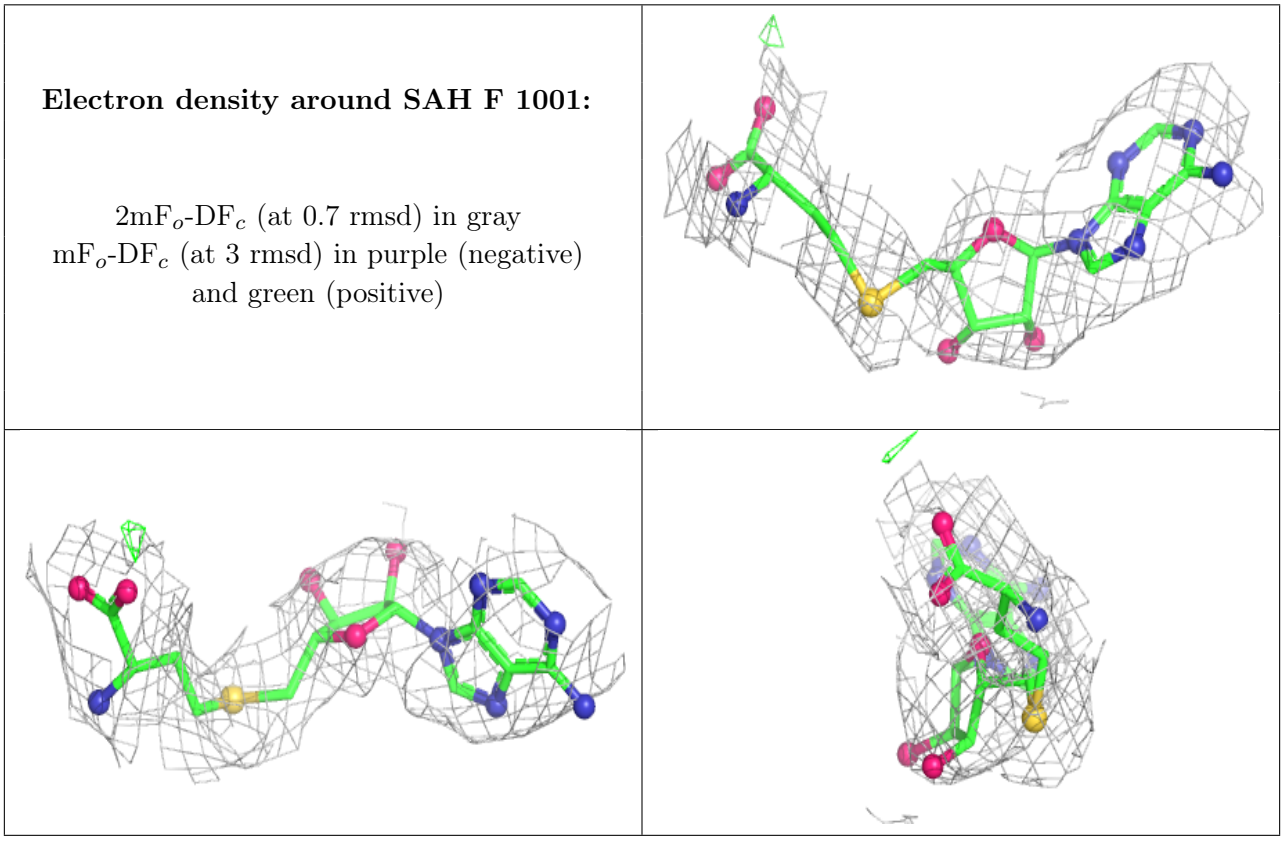
Electron density around SAH D 1001:

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



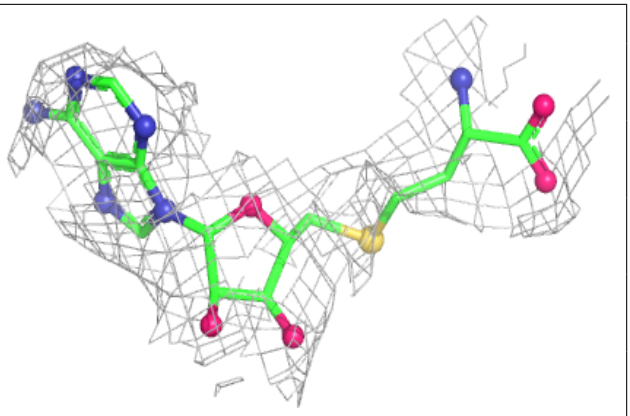
Electron density around SAH F 1001:

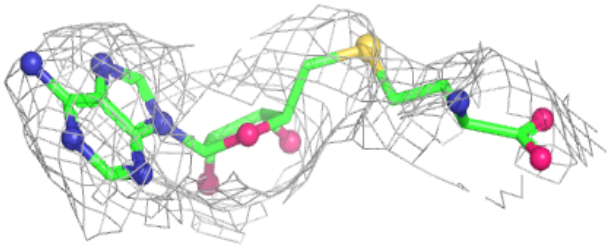
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

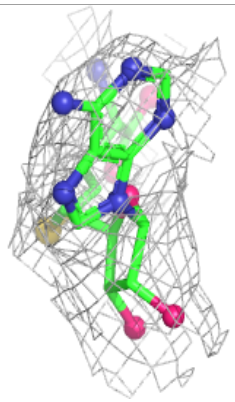


Electron density around SAH A 1001:

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

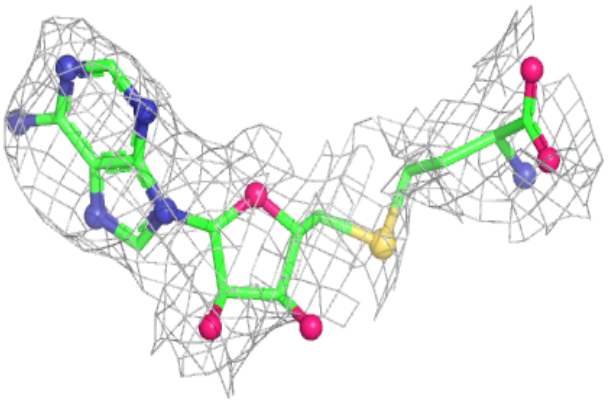


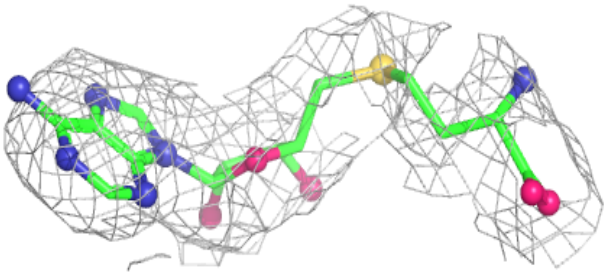


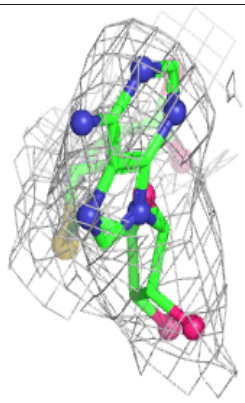


Electron density around SAH E 1001:

$2mF_o$-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 (i)

There are no such residues in this entry.