wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 09:31 PM EDT

PDB ID	:	3UA3
Title	:	Crystal Structure of Protein Arginine Methyltransferase PRMT5 in complex with SAH
Authors	:	Sun, L.; Wang, M.; Lv, Z.; Yang, N.; Liu, Y.; Bao, S.; Gong, W.; Xu, R.M.
Deposited on	:	2011-10-20
Resolution	:	3.00 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:

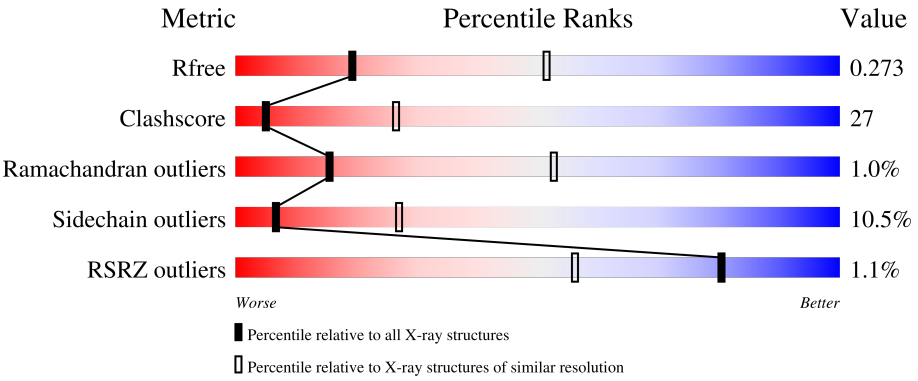
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 3.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments 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	745	% 47% 37% 5% 11%
1	B	745	% 47% 35% 6% 13%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10737 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.

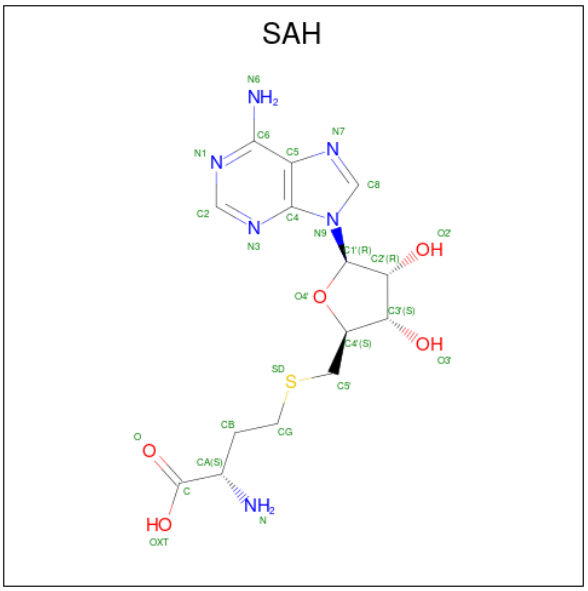
- Molecule 1 is a protein called Protein arginine N-methyltransferase 5.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	662	Total	C	N	O	S	Se	1	0	0
			5314	3400	901	993	6	14			
1	B	651	Total	C	N	O	S	Se	1	0	0
			5225	3350	881	974	6	14			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MSE	-	EXPRESSION TAG	UNP P46580
A	-1	ALA	-	EXPRESSION TAG	UNP P46580
A	0	SER	-	EXPRESSION TAG	UNP P46580
A	735	LEU	-	EXPRESSION TAG	UNP P46580
A	736	GLU	-	EXPRESSION TAG	UNP P46580
A	737	HIS	-	EXPRESSION TAG	UNP P46580
A	738	HIS	-	EXPRESSION TAG	UNP P46580
A	739	HIS	-	EXPRESSION TAG	UNP P46580
A	740	HIS	-	EXPRESSION TAG	UNP P46580
A	741	HIS	-	EXPRESSION TAG	UNP P46580
A	742	HIS	-	EXPRESSION TAG	UNP P46580
B	-2	MSE	-	EXPRESSION TAG	UNP P46580
B	-1	ALA	-	EXPRESSION TAG	UNP P46580
B	0	SER	-	EXPRESSION TAG	UNP P46580
B	735	LEU	-	EXPRESSION TAG	UNP P46580
B	736	GLU	-	EXPRESSION TAG	UNP P46580
B	737	HIS	-	EXPRESSION TAG	UNP P46580
B	738	HIS	-	EXPRESSION TAG	UNP P46580
B	739	HIS	-	EXPRESSION TAG	UNP P46580
B	740	HIS	-	EXPRESSION TAG	UNP P46580
B	741	HIS	-	EXPRESSION TAG	UNP P46580
B	742	HIS	-	EXPRESSION TAG	UNP P46580

- 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		

- Molecule 3 is water.

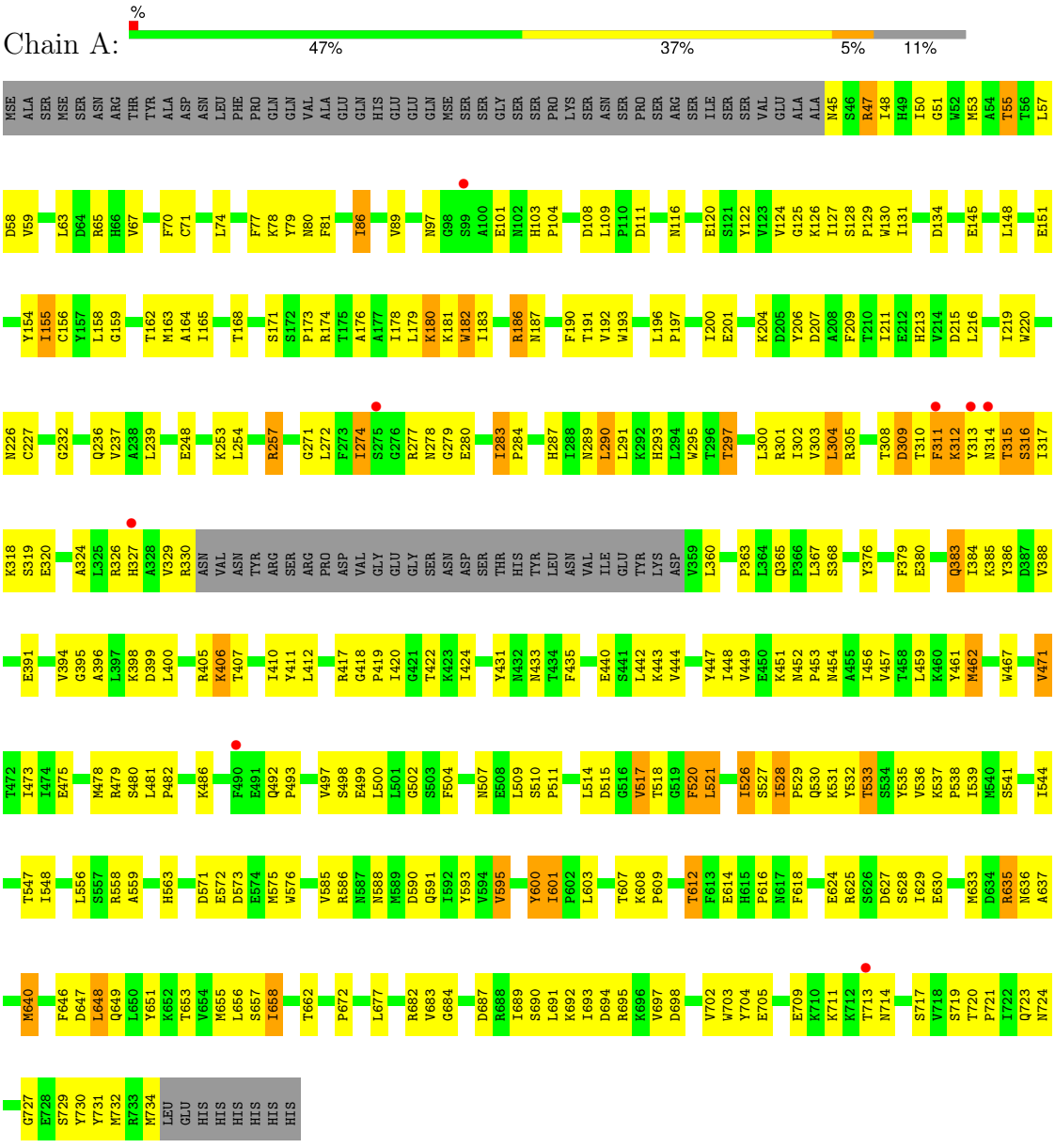
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	79	Total	O	0	0
			79	79		
3	B	67	Total	O	0	0
			67	67		

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: Protein arginine N-methyltransferase 5

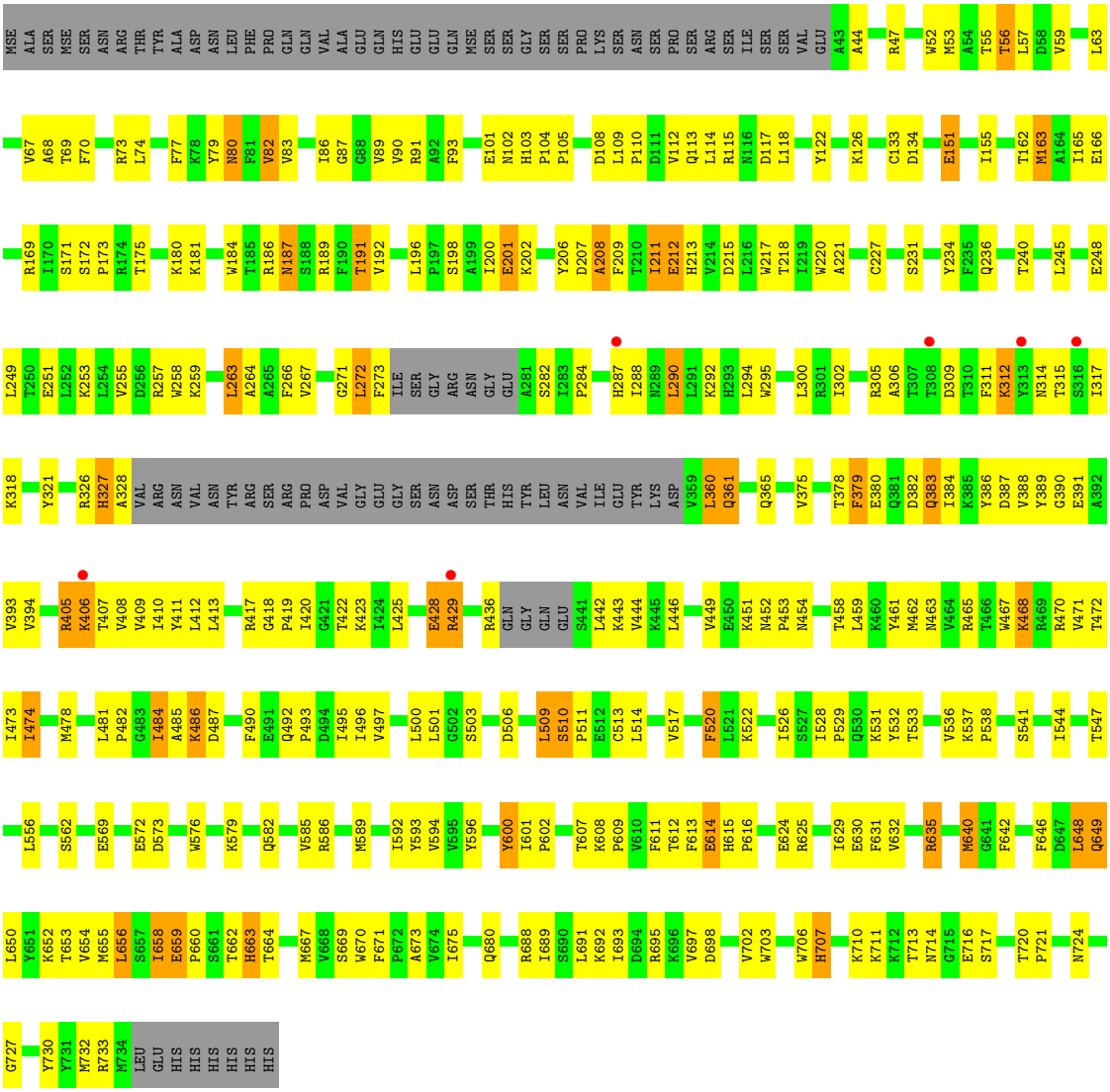
Chain A:



• Molecule 1: Protein arginine N-methyltransferase 5

Chain B:


%
47%
35%
6%
13%


MSE A43 R47 W52 M53 A54 T55 T56 L57 D58 V59 L63
V67 A68 T69 F70 R73 L74 F77 K78 Y79 N80 F81 V82 V83 I86 G87 G88 V89 V90 R91 A92 F93 E101 N102 H103 P104 P105 D108 L109 P110 D111 V112 Q113 L114 R115 N116 D117 L118 Y122 K126 C133 D134 E151 I155 T162 M163 A164 I165 E166
R169 I170 S171 S172 P173 R174 T175 K180 K181 W184 T185 R186 N187 S188 R189 F190 T191 V192 L196 P197 S198 A199 I200 E201 K202 Y206 D207 A208 F209 T210 I211 E212 H213 Y214 D215 L216 W217 T218 I219 W220 A221 C227 S231 Y234 F235 Q236 T240 L245 E248
L249 T250 E251 L252 K253 L254 V255 D256 R257 W258 K259 L263 A264 A265 F266 V267 G271 L272 F273 ILE SER GLY ARG ASN GLY GLU A281 S282 I283 P284 H287 I288 N289 L290 L291 K292 H293 L294 W295 L300 R301 I302 R305 A306 T307 T308 D309 T310 F311 K312 Y313 N314 T315 S316 I317
K318 Y321 R326 H327 A328 VAL ARG ASN VAL ASN TYR ARG SER ARG PRO ASP VAL GLY GLU GLY SER ASN ASP SER THR HIS TYR LEU ASN VAL ILE GLU TYR LYS ASP V359 L360 Q361 Q365 V375 T378 F379 E380 Q381 D382 Q383 I384 K385 Y386 D387 V388 Y389 G390 E391 A392
V393 V394 R405 K406 T407 V408 V409 I410 Y411 L412 L413 R417 G418 P419 I420 G421 T422 K423 I424 L425 E428 R429 R436 GLN GLY GLN GLU S441 L442 K443 V444 K445 L446 V449 E450 K451 N452 P453 N454 T458 L459 K460 Y461 M462 N463 V464 R465 T466 W467 K468 R469 R470 V471 T472
I473 I474 M478 L481 P482 G483 I484 A485 K486 D487 F490 E491 Q492 P493 D494 I495 I496 V497 L500 L501 G502 S503 D506 L509 S510 P511 E512 C513 L514 V517 F520 L521 K522 I526 S527 I528 P529 Q530 K531 Y532 T533 V536 K537 P538 S541 I544 T547
L556 S562 E569 E572 D573 W576 K579 Q582 V585 R586 M589 I592 Y593 V594 V595 Y596 Y600 I601 P602 T607 K608 P609 V610 F611 T612 F613 E614 H615 P616 E624 R625 I629 E630 F631 V632 R635 M640 G641 F642 F646 D647 L648 Q649
L650 Y651 K652 T653 V654 M655 L656 S657 I658 E659 P660 S661 T662 H663 T664 M667 V668 S669 W670 F671 P672 A673 V674 I675 Q680 R688 I689 S690 L691 K692 I693 D694 R695 K696 V697 D698 V702 W703 W706 H707 K710 K711 K712 T713 N714 G715 E716 S717 T720 P721 N724
G727 Y730 Y731 M732 R733 M734 LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	100.38Å 129.38Å 149.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.95 – 3.00 29.95 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.95-3.00) 100.0 (29.95-3.00)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.41 (at 3.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.2_432)	Depositor
R, R_{free}	0.223 , 0.281 0.219 , 0.273	Depositor DCC
R_{free} test set	1828 reflections (4.62%)	wwPDB-VP
Wilson B-factor (Å^2)	38.8	Xtriage
Anisotropy	0.014	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 69.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	10737	wwPDB-VP
Average B, all atoms (Å^2)	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.62% 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: 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	0.31	0/5433	0.51	0/7360
1	B	0.30	0/5342	0.50	0/7237
All	All	0.31	0/10775	0.50	0/14597

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	5314	0	5259	294	0
1	B	5225	0	5173	285	0
2	A	26	0	19	3	0
2	B	26	0	19	0	0
3	A	79	0	0	3	0
3	B	67	0	0	6	0
All	All	10737	0	10470	572	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 27.

The worst 5 of 572 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:74:LEU:HB3	1:A:79:TYR:HE1	1.22	1.02
1:A:635:ARG:HH11	1:A:635:ARG:HG2	1.23	1.02
1:B:57:LEU:HD12	1:B:306:ALA:HB1	1.43	0.96
1:A:449:VAL:HG21	1:A:481:LEU:HD11	1.48	0.94
1:A:422:THR:HG22	1:A:467:TRP:HE1	1.32	0.94

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	658/745 (88%)	579 (88%)	72 (11%)	7 (1%)	14	50
1	B	643/745 (86%)	581 (90%)	56 (9%)	6 (1%)	17	55
All	All	1301/1490 (87%)	1160 (89%)	128 (10%)	13 (1%)	15	53

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	312	LYS
1	A	97	ASN
1	A	182	TRP
1	A	315	THR
1	B	212	GLU

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	587/645 (91%)	527 (90%)	60 (10%)	7 28
1	B	577/645 (90%)	515 (89%)	62 (11%)	6 26
All	All	1164/1290 (90%)	1042 (90%)	122 (10%)	7 27

5 of 122 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	694	ASP
1	B	648	LEU
1	B	191	THR
1	B	640	MSE
1	B	707	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	377	ASN
1	B	383	GLN
1	B	707	HIS
1	B	615	HIS
1	B	649	GLN

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)

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	$\#\|Z\| > 2$
2	SAH	A	743	-	23,28,28	1.24	3 (13%)	22,40,40	1.90	3 (13%)
2	SAH	B	743	-	23,28,28	1.24	3 (13%)	22,40,40	1.81	3 (13%)

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	A	743	-	-	2/11/31/31	0/3/3/3
2	SAH	B	743	-	-	3/11/31/31	0/3/3/3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	743	SAH	C2-N3	4.00	1.38	1.32
2	A	743	SAH	C2-N3	3.91	1.38	1.32
2	A	743	SAH	C2-N1	2.57	1.38	1.33
2	B	743	SAH	C2-N1	2.55	1.38	1.33
2	A	743	SAH	OXT-C	-2.27	1.23	1.30

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	743	SAH	N3-C2-N1	-6.35	120.06	128.67
2	B	743	SAH	N3-C2-N1	-6.18	120.28	128.67
2	A	743	SAH	C5'-SD-CG	-4.29	89.53	102.26
2	B	743	SAH	C5'-SD-CG	-3.57	91.65	102.26
2	A	743	SAH	OXT-C-O	-2.88	117.55	124.08

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	743	SAH	C-CA-CB-CG
2	B	743	SAH	N-CA-CB-CG
2	B	743	SAH	CB-CG-SD-C5'
2	A	743	SAH	CB-CG-SD-C5'
2	A	743	SAH	C-CA-CB-CG

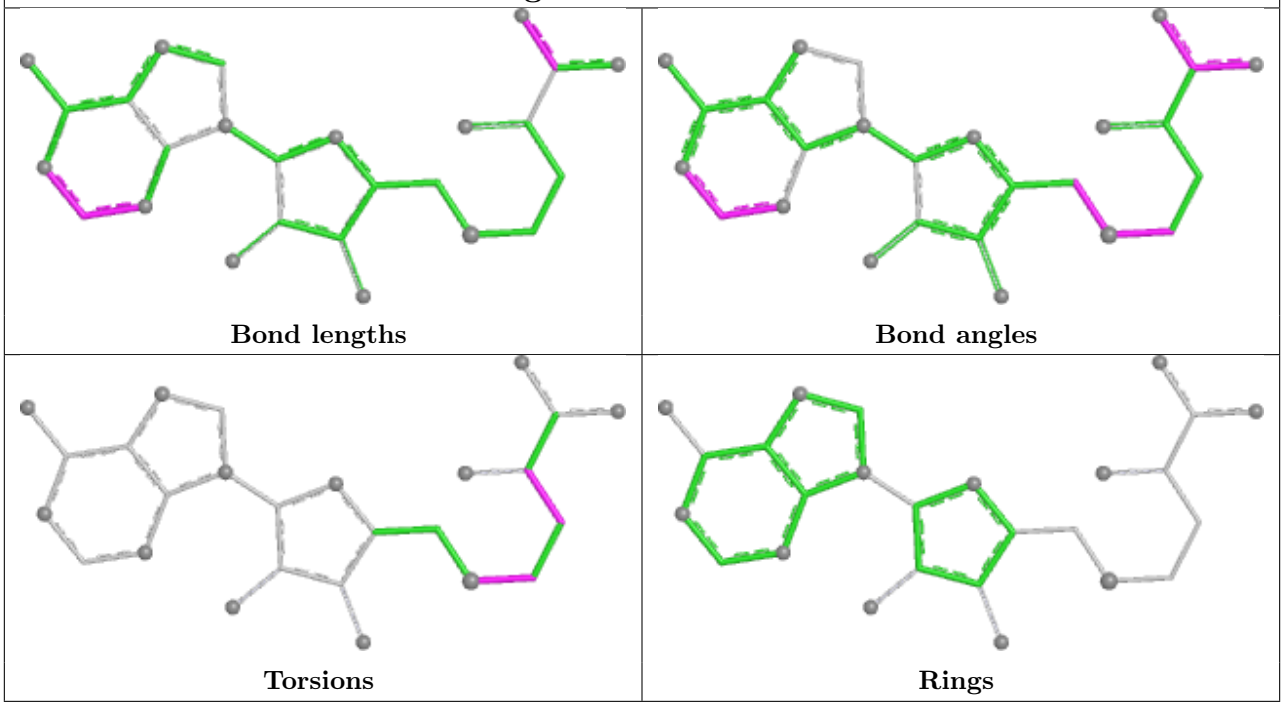
There are no ring outliers.

1 monomer is involved in 3 short contacts:

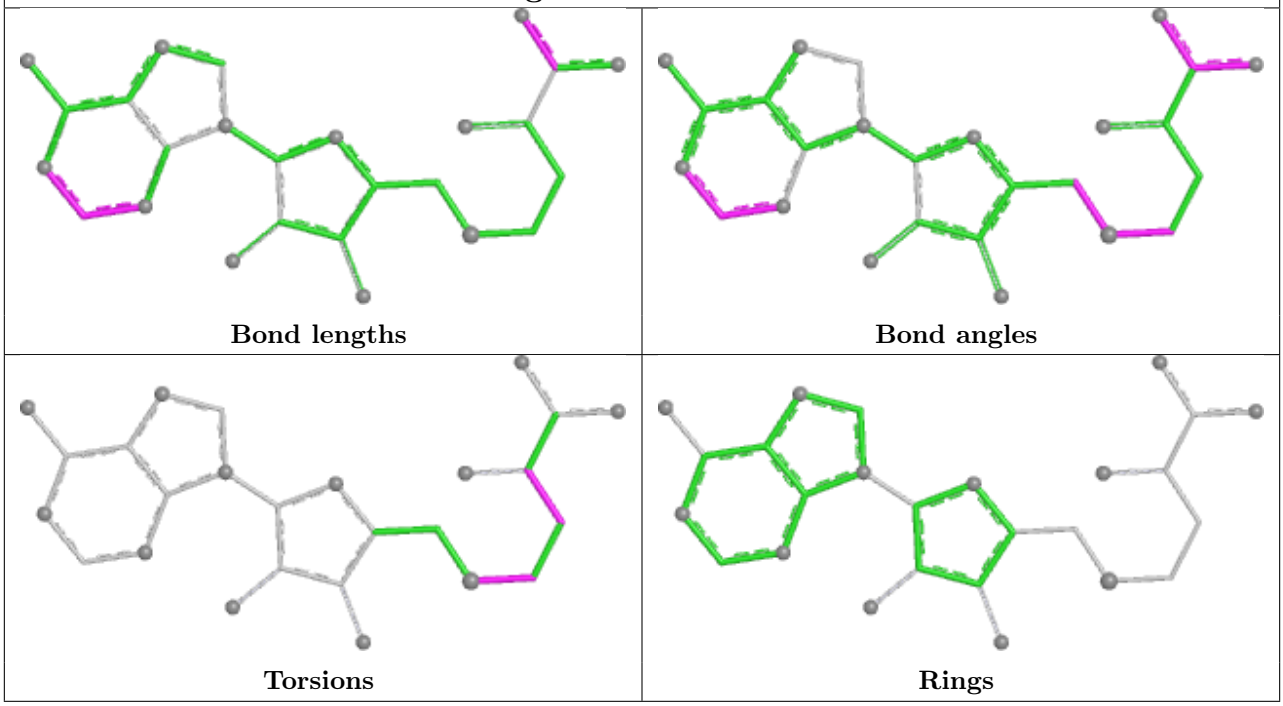
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	743	SAH	3	0

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 743



Ligand SAH B 743



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	648/745 (86%)	-0.42	8 (1%)	79	54	12, 78, 146, 239	0
1	B	637/745 (85%)	-0.24	6 (0%)	84	63	16, 88, 162, 216	13 (2%)
All	All	1285/1490 (86%)	-0.33	14 (1%)	80	56	12, 82, 154, 239	13 (1%)

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	313	TYR	4.9
1	B	429	ARG	3.7
1	B	308	THR	3.5
1	A	313	TYR	3.3
1	A	311	PHE	2.7

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

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

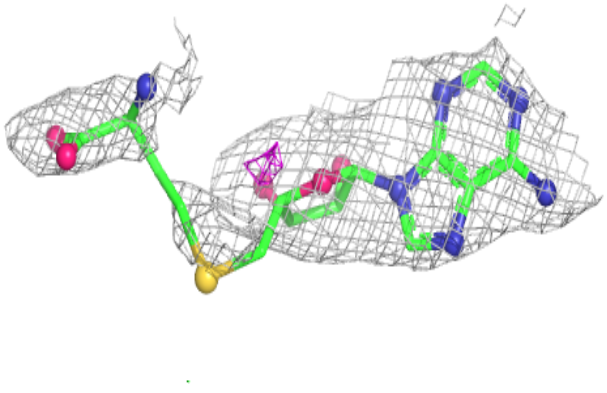
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

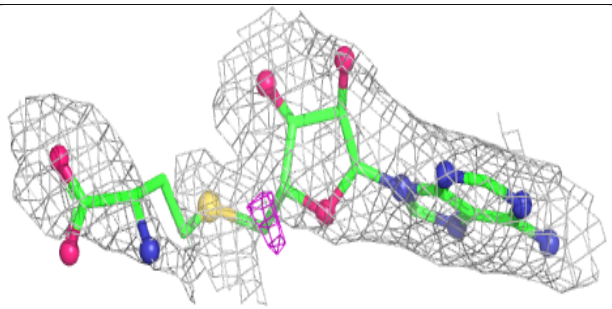
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SAH	B	743	26/26	0.93	0.20	85,97,121,203	0
2	SAH	A	743	26/26	0.96	0.14	70,75,88,99	0

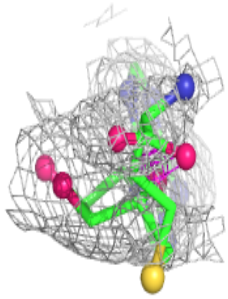
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around SAH B 743:

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

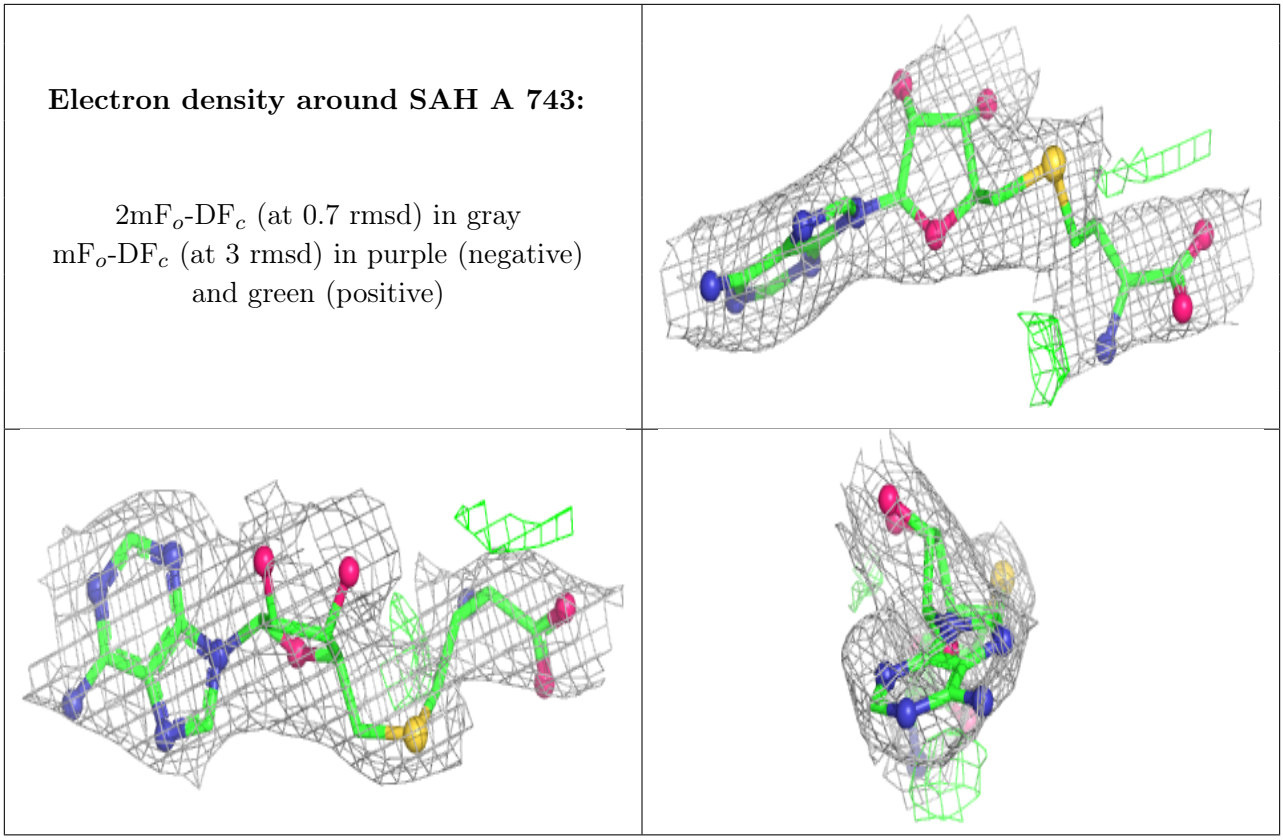






Electron density around SAH A 743:

$2mF_o$-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.