



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 01:36 PM GMT

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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

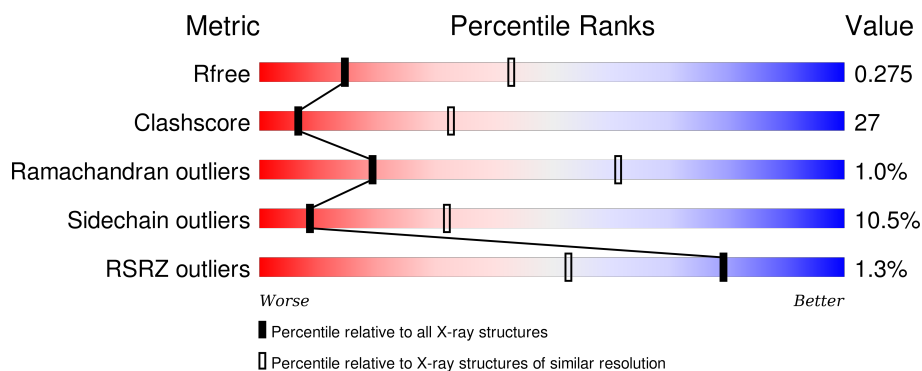
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	745	<div> <div></div> <div> <div></div> <div>47%</div> <div>37%</div> <div>5%</div> <div>11%</div> </div> </div>
1	B	745	<div> <div></div> <div> <div></div> <div>47%</div> <div>35%</div> <div>6%</div> <div>13%</div> </div> </div>

## 2 Entry composition

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.

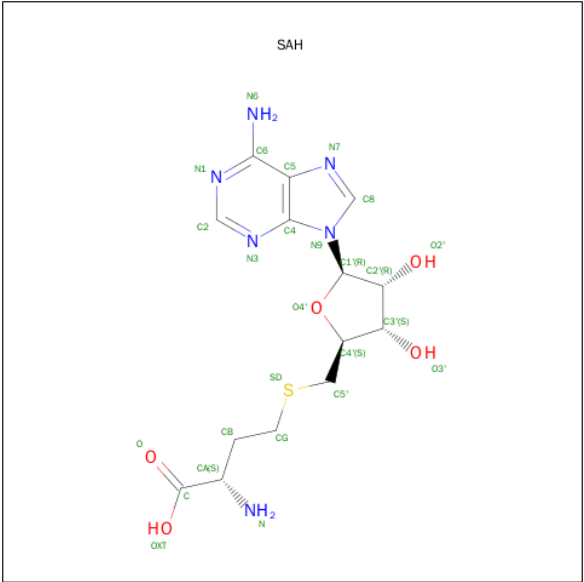
- Molecule 1 is a protein called Protein arginine N-methyltransferase 5.

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<sub>14</sub>H<sub>20</sub>N<sub>6</sub>O<sub>5</sub>S).



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		



6727	Q649	T547	I473	V393	K318	L249	R169	V67	WSE
L650	L650	L556	I474	V394	Y321	L250	R170	A68	ALA
T651	T651	L556	I474	V394	Y321	L250	R170	T69	SER
T652	T652	L556	I474	V394	Y321	L250	R170	F70	SER
T653	T653	L556	I474	V394	Y321	L250	R170	R73	ASN
T654	T654	L556	I474	V394	Y321	L250	R170	L74	ARG
T655	T655	L556	I474	V394	Y321	L250	R170		THR
T656	T656	L556	I474	V394	Y321	L250	R170	F77	TYR
T657	T657	L556	I474	V394	Y321	L250	R170	K180	ALA
T658	T658	L556	I474	V394	Y321	L250	R170	K181	ASP
T659	T659	L556	I474	V394	Y321	L250	R170	Y79	ASN
T660	T660	L556	I474	V394	Y321	L250	R170	D80	LEU
T661	T661	L556	I474	V394	Y321	L250	R170	H84	PHE
T662	T662	L556	I474	V394	Y321	L250	R170	T85	PRO
T663	T663	L556	I474	V394	Y321	L250	R170	R86	GLN
T664	T664	L556	I474	V394	Y321	L250	R170	N187	GLN
T665	T665	L556	I474	V394	Y321	L250	R170	V83	GLN
T666	T666	L556	I474	V394	Y321	L250	R170	I86	VAL
T667	T667	L556	I474	V394	Y321	L250	R170	G87	ALA
T668	T668	L556	I474	V394	Y321	L250	R170	G88	GLU
T669	T669	L556	I474	V394	Y321	L250	R170	V89	GLN
T670	T670	L556	I474	V394	Y321	L250	R170	V90	HIS
T671	T671	L556	I474	V394	Y321	L250	R170	R91	GLU
T672	T672	L556	I474	V394	Y321	L250	R170	A92	GLU
T673	T673	L556	I474	V394	Y321	L250	R170	F93	GLN
T674	T674	L556	I474	V394	Y321	L250	R170	E101	WSE
T675	T675	L556	I474	V394	Y321	L250	R170	M102	SER
T676	T676	L556	I474	V394	Y321	L250	R170	H103	GLY
T677	T677	L556	I474	V394	Y321	L250	R170	P104	SER
T678	T678	L556	I474	V394	Y321	L250	R170	P105	SER
T679	T679	L556	I474	V394	Y321	L250	R170	D207	PRO
T680	T680	L556	I474	V394	Y321	L250	R170	D108	LYS
T681	T681	L556	I474	V394	Y321	L250	R170	L109	SER
T682	T682	L556	I474	V394	Y321	L250	R170	P110	ASN
T683	T683	L556	I474	V394	Y321	L250	R170	D111	ASN
T684	T684	L556	I474	V394	Y321	L250	R170	V112	SER
T685	T685	L556	I474	V394	Y321	L250	R170	E212	PRO
T686	T686	L556	I474	V394	Y321	L250	R170	H213	SER
T687	T687	L556	I474	V394	Y321	L250	R170	D215	ARG
T688	T688	L556	I474	V394	Y321	L250	R170	L216	SER
T689	T689	L556	I474	V394	Y321	L250	R170	W217	ILE
T690	T690	L556	I474	V394	Y321	L250	R170	T218	SER
T691	T691	L556	I474	V394	Y321	L250	R170	I219	VAL
T692	T692	L556	I474	V394	Y321	L250	R170	W220	GLU
T693	T693	L556	I474	V394	Y321	L250	R170	A221	GLU
T694	T694	L556	I474	V394	Y321	L250	R170	K126	A43
T695	T695	L556	I474	V394	Y321	L250	R170	C227	A44
T696	T696	L556	I474	V394	Y321	L250	R170	G133	B47
T697	T697	L556	I474	V394	Y321	L250	R170	D134	
T698	T698	L556	I474	V394	Y321	L250	R170	E151	W52
T699	T699	L556	I474	V394	Y321	L250	R170	F235	W53
T700	T700	L556	I474	V394	Y321	L250	R170	I155	A54
T701	T701	L556	I474	V394	Y321	L250	R170	Q236	T55
T702	T702	L556	I474	V394	Y321	L250	R170	T240	T56
T703	T703	L556	I474	V394	Y321	L250	R170	L245	L57
T704	T704	L556	I474	V394	Y321	L250	R170	L245	W58
T705	T705	L556	I474	V394	Y321	L250	R170	L245	W59
T706	T706	L556	I474	V394	Y321	L250	R170	L245	L63
T707	T707	L556	I474	V394	Y321	L250	R170	L245	
T708	T708	L556	I474	V394	Y321	L250	R170	L245	
T709	T709	L556	I474	V394	Y321	L250	R170	L245	
T710	T710	L556	I474	V394	Y321	L250	R170	L245	
T711	T711	L556	I474	V394	Y321	L250	R170	L245	
T712	T712	L556	I474	V394	Y321	L250	R170	L245	
T713	T713	L556	I474	V394	Y321	L250	R170	L245	
T714	T714	L556	I474	V394	Y321	L250	R170	L245	
T715	T715	L556	I474	V394	Y321	L250	R170	L245	
T716	T716	L556	I474	V394	Y321	L250	R170	L245	
T717	T717	L556	I474	V394	Y321	L250	R170	L245	
T718	T718	L556	I474	V394	Y321	L250	R170	L245	
T719	T719	L556	I474	V394	Y321	L250	R170	L245	
T720	T720	L556	I474	V394	Y321	L250	R170	L245	
T721	T721	L556	I474	V394	Y321	L250	R170	L245	
T722	T722	L556	I474	V394	Y321	L250	R170	L245	
T723	T723	L556	I474	V394	Y321	L250	R170	L245	
T724	T724	L556	I474	V394	Y321	L250	R170	L245	

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.38 Å 129.38 Å 149.31 Å 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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.41 (at 3.00 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.2_432)	Depositor
R, $R_{free}$	0.223 , 0.281 0.217 , 0.275	Depositor DCC
$R_{free}$ test set	1827 reflections (4.62%)	DCC
Wilson B-factor (Å <sup>2</sup> )	38.8	Xtriage
Anisotropy	0.014	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 69.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 39562 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	10737	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	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.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 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 [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	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%)	17	58
1	B	643/745 (86%)	581 (90%)	56 (9%)	6 (1%)	21	64
All	All	1301/1490 (87%)	1160 (89%)	128 (10%)	13 (1%)	19	61

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%)	9	33
1	B	577/645 (90%)	515 (89%)	62 (11%)	8	31
All	All	1164/1290 (90%)	1042 (90%)	122 (10%)	8	32

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

Mol	Chain	Res	Type
1	A	658	ILE
1	B	163	MSE
1	B	652	LYS
1	A	687	ASP
1	B	80	ASN

Some 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	195	GLN
1	B	377	ASN
1	B	615	HIS
1	B	116	ASN
1	B	551	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

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	-	20,28,28	1.08	2 (10%)	19,40,40	2.95	2 (10%)
2	SAH	B	743	-	20,28,28	1.10	2 (10%)	19,40,40	2.85	2 (10%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SAH	A	743	-	-	0/7/31/31	0/3/3/3
2	SAH	B	743	-	-	0/7/31/31	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	743	SAH	C2-N1	2.40	1.38	1.33
2	A	743	SAH	C2-N1	2.42	1.38	1.33
2	A	743	SAH	C2-N3	3.37	1.38	1.32
2	B	743	SAH	C2-N3	3.45	1.38	1.32

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	743	SAH	N3-C2-N1	-11.54	120.06	128.89
2	B	743	SAH	N3-C2-N1	-11.25	120.28	128.89
2	A	743	SAH	C5'-SD-CG	-4.29	89.53	102.41
2	B	743	SAH	C5'-SD-CG	-3.58	91.65	102.41

There are no chirality outliers.

There are no torsion outliers.

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

## 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 [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, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å <sup>2</sup> )	Q<0.9
1	A	648/745 (86%)	-0.40	8 (1%)	81 55	12, 78, 146, 239	0
1	B	637/745 (85%)	-0.22	9 (1%)	78 51	16, 88, 162, 216	13 (2%)
All	All	1285/1490 (86%)	-0.31	17 (1%)	79 53	12, 82, 154, 239	13 (1%)

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	313	TYR	5.0
1	B	429	ARG	3.8
1	B	308	THR	3.6
1	A	313	TYR	3.4
1	A	311	PHE	2.8

### 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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SAH	B	743	26/26	0.93	0.20	-0.23	85,97,121,203	0
2	SAH	A	743	26/26	0.96	0.14	-0.76	70,75,88,99	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.