



# wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 09:52 PM GMT

PDB ID : 3OFK  
Title : Crystal structure of N-methyltransferase NodS from Bradyrhizobium japonicum WM9 in complex with S-adenosyl-l-homocysteine(SAH)  
Authors : Cakici, O.; Sikorski, M.; Stepkowski, T.; Bujacz, G.; Jaskolski, M.  
Deposited on : 2010-08-15  
Resolution : 1.85 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

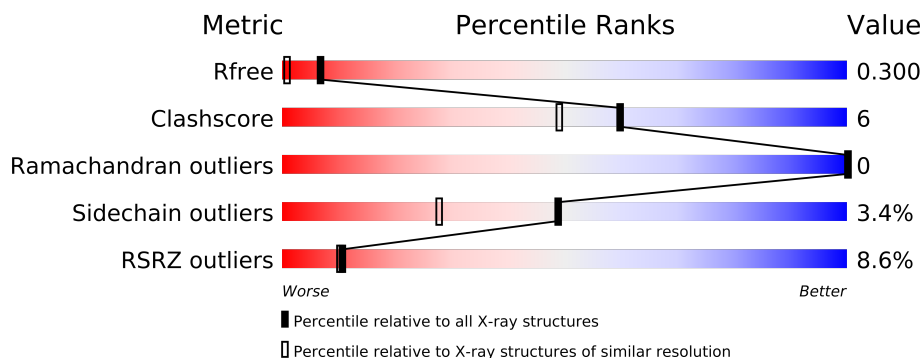
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.85 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))
$R_{free}$	66092	1269 (1.86-1.86)
Clashscore	79885	1470 (1.86-1.86)
Ramachandran outliers	78287	1451 (1.86-1.86)
Sidechain outliers	78261	1451 (1.86-1.86)
RSRZ outliers	66119	1269 (1.86-1.86)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	216	
1	B	216	
1	C	216	
1	D	216	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7102 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Nodulation protein S.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	204	Total	C	N	O	S	0	0	0
			1600	996	287	305	12			
1	B	196	Total	C	N	O	S	0	0	0
			1539	956	279	293	11			
1	C	202	Total	C	N	O	S	0	0	0
			1588	985	291	301	11			
1	D	181	Total	C	N	O	S	0	0	0
			1406	874	257	264	11			

There are 32 discrepancies between the modelled and reference sequences:

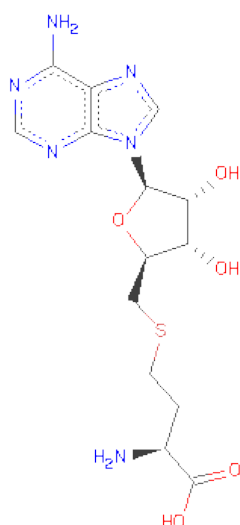
Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	GLY	-	EXPRESSION TAG	UNP Q9AQ22
A	-5	ILE	-	EXPRESSION TAG	UNP Q9AQ22
A	-4	ASP	-	EXPRESSION TAG	UNP Q9AQ22
A	-3	PRO	-	EXPRESSION TAG	UNP Q9AQ22
A	-2	PHE	-	EXPRESSION TAG	UNP Q9AQ22
A	-1	THR	-	EXPRESSION TAG	UNP Q9AQ22
A	0	MET	-	EXPRESSION TAG	UNP Q9AQ22
A	1	VAL	-	SEE REMARK 999	UNP Q9AQ22
B	-6	GLY	-	EXPRESSION TAG	UNP Q9AQ22
B	-5	ILE	-	EXPRESSION TAG	UNP Q9AQ22
B	-4	ASP	-	EXPRESSION TAG	UNP Q9AQ22
B	-3	PRO	-	EXPRESSION TAG	UNP Q9AQ22
B	-2	PHE	-	EXPRESSION TAG	UNP Q9AQ22
B	-1	THR	-	EXPRESSION TAG	UNP Q9AQ22
B	0	MET	-	EXPRESSION TAG	UNP Q9AQ22
B	1	VAL	-	SEE REMARK 999	UNP Q9AQ22
C	-6	GLY	-	EXPRESSION TAG	UNP Q9AQ22
C	-5	ILE	-	EXPRESSION TAG	UNP Q9AQ22
C	-4	ASP	-	EXPRESSION TAG	UNP Q9AQ22
C	-3	PRO	-	EXPRESSION TAG	UNP Q9AQ22
C	-2	PHE	-	EXPRESSION TAG	UNP Q9AQ22

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	THR	-	EXPRESSION TAG	UNP Q9AQ22
C	0	MET	-	EXPRESSION TAG	UNP Q9AQ22
C	1	VAL	-	SEE REMARK 999	UNP Q9AQ22
D	-6	GLY	-	EXPRESSION TAG	UNP Q9AQ22
D	-5	ILE	-	EXPRESSION TAG	UNP Q9AQ22
D	-4	ASP	-	EXPRESSION TAG	UNP Q9AQ22
D	-3	PRO	-	EXPRESSION TAG	UNP Q9AQ22
D	-2	PHE	-	EXPRESSION TAG	UNP Q9AQ22
D	-1	THR	-	EXPRESSION TAG	UNP Q9AQ22
D	0	MET	-	EXPRESSION TAG	UNP Q9AQ22
D	1	VAL	-	SEE REMARK 999	UNP Q9AQ22

- 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		

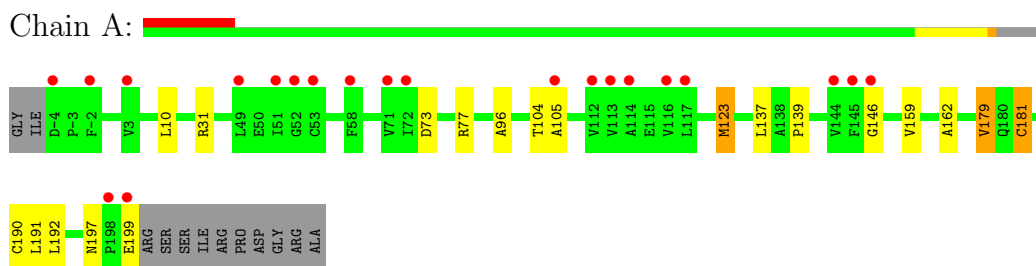
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	170	Total 174	O 174	0	4
3	B	238	Total 244	O 244	0	6
3	C	259	Total 262	O 262	0	3
3	D	179	Total 185	O 185	0	6

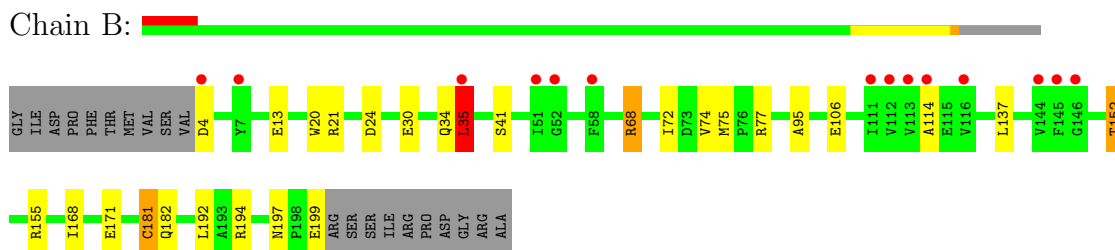
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

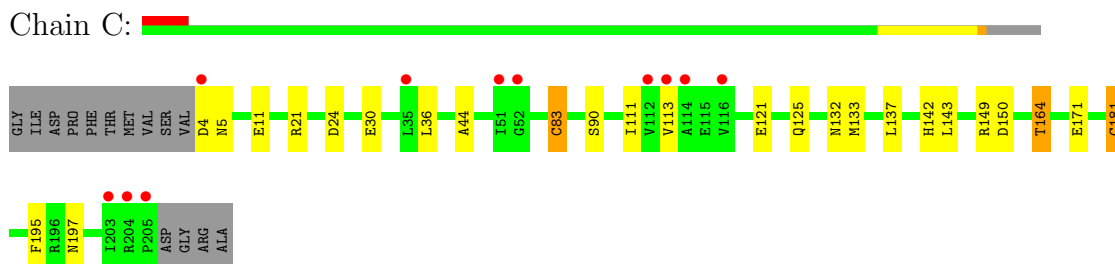
#### • Molecule 1: Nodulation protein S



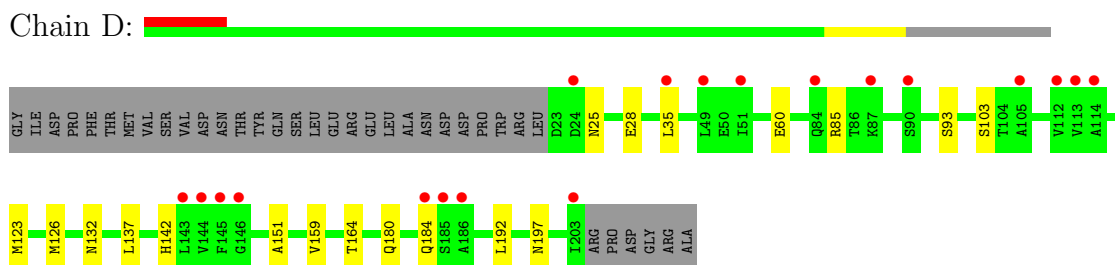
#### • Molecule 1: Nodulation protein S



#### • Molecule 1: Nodulation protein S



#### • Molecule 1: Nodulation protein S



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.01Å 143.30Å 75.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.70 – 1.85 29.70 – 1.85	Depositor EDS
% Data completeness (in resolution range)	100.0 (29.70-1.85) 94.5 (29.70-1.85)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.08 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, $R_{free}$	0.190 , 0.238 0.257 , 0.300	Depositor DCC
$R_{free}$ test set	1165 reflections (1.65%)	DCC
Wilson B-factor (Å <sup>2</sup> )	29.4	Xtriage
Anisotropy	0.166	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 17.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 71694 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7102	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.06% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 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	0.67	1/1629 (0.1%)	0.75	1/2210 (0.0%)
1	B	1.01	2/1566 (0.1%)	0.94	3/2123 (0.1%)
1	C	0.97	2/1616 (0.1%)	0.91	0/2190
1	D	0.72	0/1429	0.78	0/1934
All	All	0.86	5/6240 (0.1%)	0.85	4/8457 (0.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	83	CYS	CB-SG	-5.89	1.72	1.81
1	C	181	CYS	CB-SG	-5.59	1.72	1.81
1	B	181	CYS	C-N	-5.55	1.21	1.34
1	B	199	GLU	CD-OE2	5.51	1.31	1.25
1	A	181	CYS	CB-SG	-5.41	1.73	1.81

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	68	ARG	NE-CZ-NH1	5.53	123.07	120.30
1	B	194	ARG	NE-CZ-NH2	-5.33	117.63	120.30
1	B	35	LEU	CA-CB-CG	5.12	127.07	115.30
1	A	31	ARG	NE-CZ-NH1	5.11	122.86	120.30

There are no chirality outliers.

There are no planarity outliers.



## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1600	0	1573	17	0
1	B	1539	0	1510	14	0
1	C	1588	0	1565	29	0
1	D	1406	0	1399	14	0
2	A	26	0	19	0	0
2	B	26	0	19	0	0
2	C	26	0	19	0	0
2	D	26	0	19	0	0
3	A	174	0	0	1	0
3	B	244	0	0	4	0
3	C	262	0	0	11	0
3	D	185	0	0	5	0
All	All	7102	0	6123	74	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

The worst 5 of 74 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:159:VAL:HG22	3:D:231:HOH:O	1.50	1.08
1:C:133:MET:CE	1:C:143:LEU:HD11	1.89	1.02
1:C:133:MET:HE2	1:C:143:LEU:HD11	1.48	0.95
1:C:113:VAL:HG21	1:C:133:MET:CE	1.96	0.94
1:C:113:VAL:HG21	1:C:133:MET:HE1	1.50	0.93

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	202/216 (94%)	193 (96%)	9 (4%)	0	100	100
1	B	194/216 (90%)	188 (97%)	6 (3%)	0	100	100
1	C	200/216 (93%)	192 (96%)	8 (4%)	0	100	100
1	D	179/216 (83%)	173 (97%)	6 (3%)	0	100	100
All	All	775/864 (90%)	746 (96%)	29 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	172/181 (95%)	165 (96%)	7 (4%)	41	20
1	B	164/181 (91%)	158 (96%)	6 (4%)	45	24
1	C	170/181 (94%)	166 (98%)	4 (2%)	61	44
1	D	150/181 (83%)	145 (97%)	5 (3%)	50	29
All	All	656/724 (91%)	634 (97%)	22 (3%)	49	28

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

Mol	Chain	Res	Type
1	B	41	SER
1	B	182	GLN
1	D	180	GLN
1	B	77	ARG
1	B	152	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	184	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	184	GLN
1	C	5	ASN
1	D	142	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

4 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	217	-	28,28,28	1.13	2 (7%)	40,40,40	2.18	10 (25%)
2	SAH	B	217	-	28,28,28	1.12	4 (14%)	40,40,40	2.92	11 (27%)
2	SAH	C	217	-	28,28,28	1.16	2 (7%)	40,40,40	2.29	6 (15%)
2	SAH	D	217	-	28,28,28	1.03	2 (7%)	40,40,40	2.15	8 (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
2	SAH	A	217	-	-	0/15/31/31	0/1/3/3
2	SAH	B	217	-	-	0/15/31/31	0/1/3/3
2	SAH	C	217	-	-	0/15/31/31	0/1/3/3
2	SAH	D	217	-	-	0/15/31/31	0/1/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	217	SAH	C2-N3	3.89	1.39	1.32
2	A	217	SAH	C2-N3	3.74	1.39	1.32
2	C	217	SAH	C2-N3	2.89	1.37	1.32
2	A	217	SAH	C2-N1	2.87	1.39	1.33
2	B	217	SAH	C2-N3	2.78	1.37	1.32

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	217	SAH	N3-C2-N1	-14.66	116.46	128.71
2	C	217	SAH	N3-C2-N1	-11.46	119.13	128.71
2	D	217	SAH	N3-C2-N1	-9.53	120.74	128.71
2	A	217	SAH	N3-C2-N1	-9.28	120.95	128.71
2	A	217	SAH	N3-C4-N9	4.45	133.46	125.43

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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<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	204/216 (94%)	0.65	22 (10%) 6 6	29, 46, 70, 98	0
1	B	196/216 (90%)	0.49	14 (7%) 16 14	19, 35, 55, 67	0
1	C	202/216 (93%)	0.38	11 (5%) 25 23	19, 33, 57, 84	0
1	D	181/216 (83%)	0.55	19 (10%) 7 7	30, 42, 65, 84	0
All	All	783/864 (90%)	0.52	66 (8%) 11 11	19, 39, 64, 98	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3	VAL	4.7
1	C	205	PRO	4.3
1	D	112	VAL	4.2
1	B	51	ILE	4.1
1	D	114	ALA	4.0

### 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 carbohydrates 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. 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SAH	B	217	26/26	0.18	0.17	18,26,32,36	0
2	SAH	C	217	26/26	0.18	0.10	21,24,27,29	0
2	SAH	D	217	26/26	0.16	0.09	32,38,40,42	0
2	SAH	A	217	26/26	0.16	-0.11	26,32,35,37	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.