



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 12:27 am GMT

PDB ID : 3V97  
Title : Crystal structure of bifunctional methyltransferase YcbY (RlmLK) from Escherichia coli, SAH binding  
Authors : Su, X.D.; Wang, K.T.  
Deposited on : 2011-12-23  
Resolution : 2.20 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

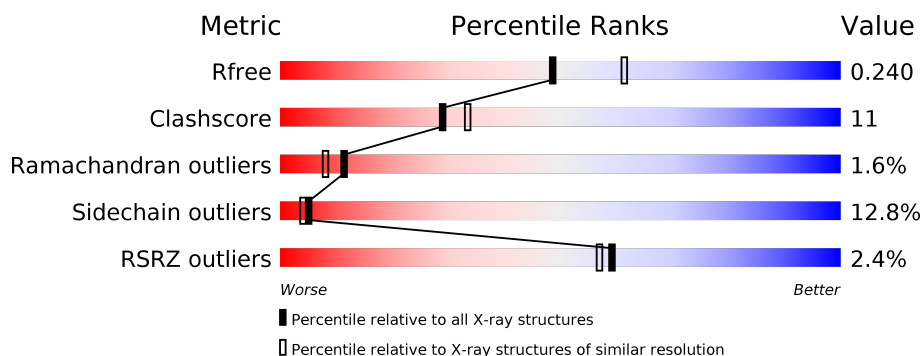
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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))
$R_{free}$	100719	4002 (2.20-2.20)
Clashscore	112137	4730 (2.20-2.20)
Ramachandran outliers	110173	4656 (2.20-2.20)
Sidechain outliers	110143	4657 (2.20-2.20)
RSRZ outliers	101464	4033 (2.20-2.20)

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	703	<div> <div>%</div> <div> <div></div> <div>70%</div> <div>19%</div> <div>5%</div> <div>5%</div> </div> </div>
1	B	703	<div> <div>4%</div> <div> <div></div> <div>69%</div> <div>16%</div> <div>5%</div> <div>9%</div> </div> </div>

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	GOL	A	803	-	X	-	-
3	GOL	A	804	-	-	X	X
3	GOL	B	804	-	-	-	X

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 10697 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.

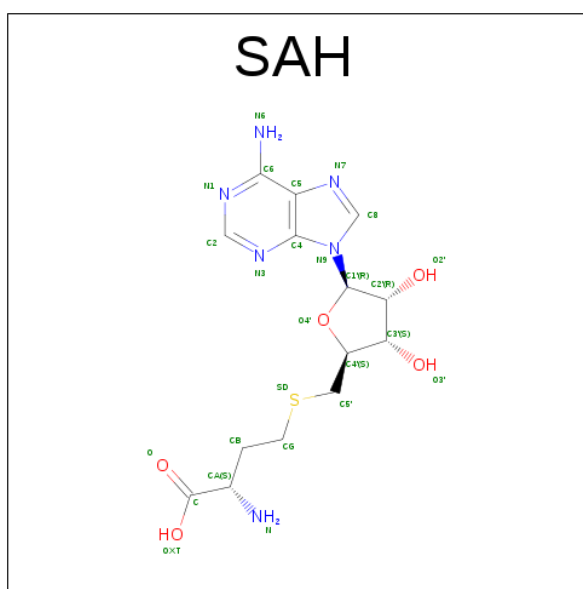
- Molecule 1 is a protein called Ribosomal RNA large subunit methyltransferase L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	667	Total	C	N	O	S	0	0	0
			5139	3262	901	954	22			
1	B	641	Total	C	N	O	S	0	0	0
			4898	3106	857	914	21			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	ALA	-	EXPRESSION TAG	UNP P75864
B	0	ALA	-	EXPRESSION TAG	UNP P75864

- 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).



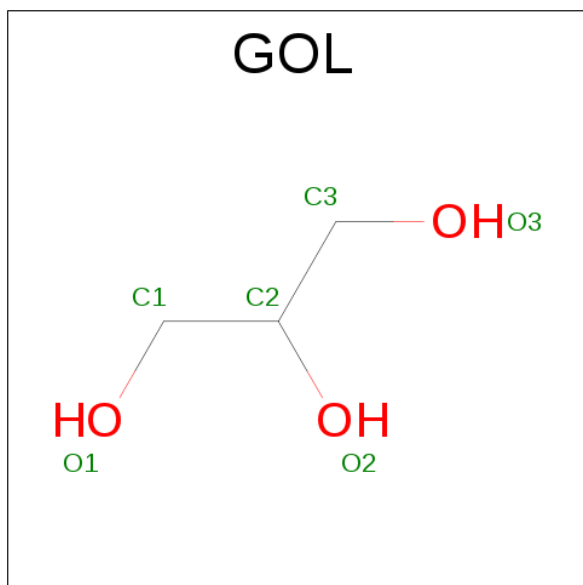
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			26	14	6	5	1		

*Continued on next page...*

Continued from previous page...

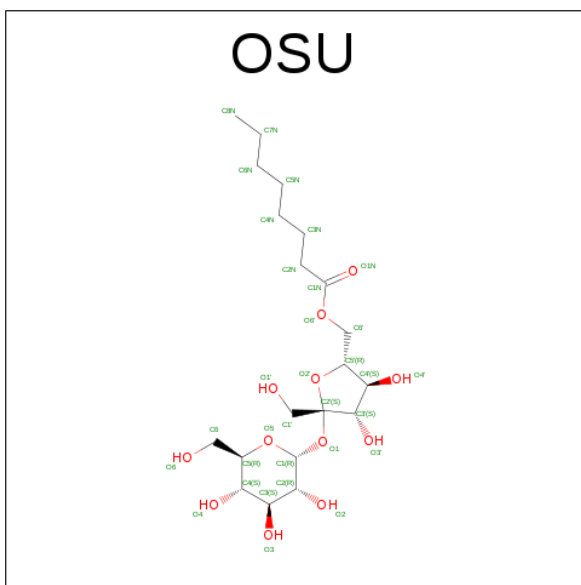
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
2	B	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
2	B	1	Total	C	N	O	S	0	0
			26	14	6	5	1		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is N-OCTANOYLSUCROSE (three-letter code: OSU) (formula:  $C_{20}H_{36}O_{12}$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 31	C 19	O 12	0	0
4	B	1	Total 30	C 18	O 12	0	0

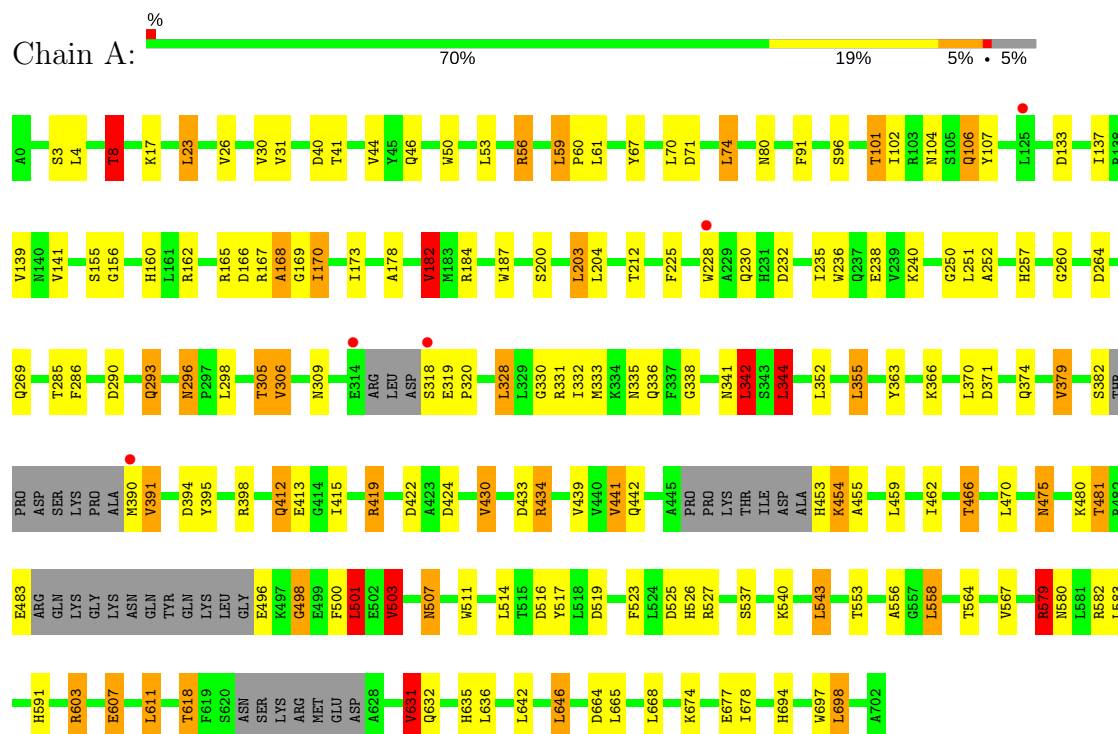
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	270	Total O 270 270	0	0
5	B	201	Total O 201 201	0	0

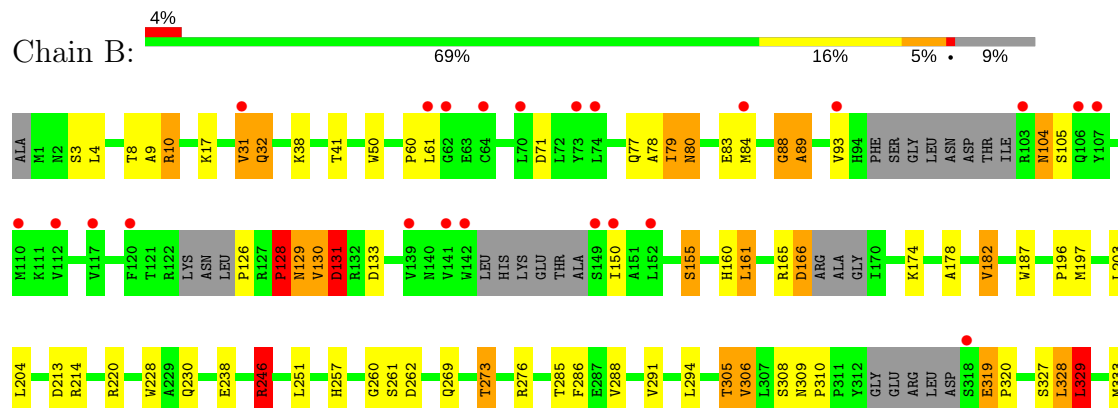
### 3 Residue-property plots

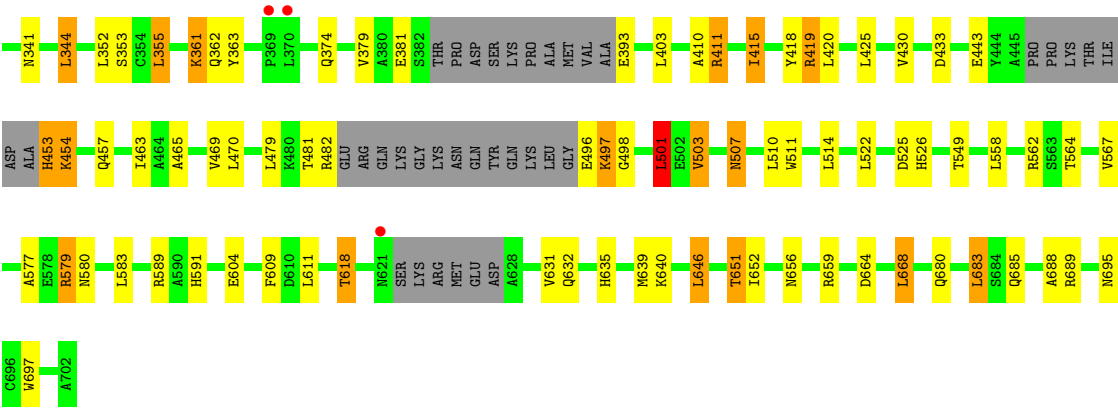
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Ribosomal RNA large subunit methyltransferase L



- Molecule 1: Ribosomal RNA large subunit methyltransferase L





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.65Å 140.83Å 102.87Å 90.00° 102.30° 90.00°	Depositor
Resolution (Å)	49.91 – 2.20 49.91 – 2.20	Depositor EDS
% Data completeness (in resolution range)	96.1 (49.91-2.20) 96.2 (49.91-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.86 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.186 , 0.234 0.194 , 0.240	Depositor DCC
$R_{free}$ test set	4970 reflections (5.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	42.5	Xtriage
Anisotropy	0.117	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 43.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10697	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.96% 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, 0.333 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: GOL, SAH, OSU

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	1.14	8/5243 (0.2%)	1.09	31/7108 (0.4%)
1	B	0.96	3/4990 (0.1%)	1.01	17/6763 (0.3%)
All	All	1.05	11/10233 (0.1%)	1.05	48/13871 (0.3%)

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	B	0	1
All	All	0	2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	238	GLU	CG-CD	6.60	1.61	1.51
1	A	238	GLU	CG-CD	6.59	1.61	1.51
1	A	631	VAL	CB-CG1	-6.17	1.39	1.52
1	A	556	ALA	CA-CB	6.11	1.65	1.52
1	A	200	SER	CA-CB	5.19	1.60	1.52

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	246	ARG	NE-CZ-NH1	11.48	126.04	120.30
1	B	589	ARG	NE-CZ-NH1	9.07	124.83	120.30
1	B	246	ARG	NE-CZ-NH2	-8.20	116.20	120.30
1	A	579	ARG	NE-CZ-NH1	-7.93	116.33	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	589	ARG	NE-CZ-NH2	-7.77	116.41	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	498	GLY	Peptide
1	B	453	HIS	Peptide

## 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	5139	0	4995	117	1
1	B	4898	0	4719	94	0
2	A	52	0	38	0	0
2	B	52	0	38	2	0
3	A	12	0	16	5	0
3	B	12	0	16	3	0
4	A	31	0	27	4	0
4	B	30	0	25	3	0
5	A	270	0	0	16	0
5	B	201	0	0	12	0
All	All	10697	0	9874	219	1

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 11.

The worst 5 of 219 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)
4:A:805:OSU:O5	4:A:805:OSU:C1	1.64	1.42
4:B:805:OSU:O5	4:B:805:OSU:C1	1.68	1.40
1:B:341:ASN:HB2	5:B:1040:HOH:O	1.38	1.21
1:B:88:GLY:HA2	1:B:89:ALA:HB2	1.44	1.00
1:A:412:GLN:HG3	5:A:1089:HOH:O	1.62	0.99

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:71:ASP:OD1	1:A:331:ARG:NH1[1_455]	2.16	0.04

## 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	655/703 (93%)	630 (96%)	20 (3%)	5 (1%)	22	21
1	B	621/703 (88%)	586 (94%)	19 (3%)	16 (3%)	6	3
All	All	1276/1406 (91%)	1216 (95%)	39 (3%)	21 (2%)	11	8

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	170	ILE
1	B	10	ARG
1	B	32	GLN
1	B	129	ASN
1	B	131	ASP

### 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	517/581 (89%)	449 (87%)	68 (13%)	5	4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	488/581 (84%)	427 (88%)	61 (12%)	5	4
All	All	1005/1162 (86%)	876 (87%)	129 (13%)	5	4

5 of 129 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	607	GLU
1	B	61	LEU
1	B	583	LEU
1	A	618	THR
1	A	665	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	591	HIS
1	A	694	HIS
1	B	676	GLN
1	A	685	GLN
1	B	22	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

10 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	801	-	20,28,28	1.41	4 (20%)	20,40,40	1.92	5 (25%)
2	SAH	A	802	-	20,28,28	1.01	1 (5%)	20,40,40	2.28	6 (30%)
3	GOL	A	803	-	5,5,5	2.11	2 (40%)	5,5,5	2.14	4 (80%)
3	GOL	A	804	-	5,5,5	0.72	0	5,5,5	1.12	0
4	OSU	A	805	-	32,32,33	5.17	9 (28%)	45,45,46	2.32	13 (28%)
2	SAH	B	801	-	20,28,28	1.41	4 (20%)	20,40,40	1.79	4 (20%)
2	SAH	B	802	-	20,28,28	1.32	3 (15%)	20,40,40	2.73	8 (40%)
3	GOL	B	803	-	5,5,5	0.92	0	5,5,5	0.93	0
3	GOL	B	804	-	5,5,5	0.82	0	5,5,5	1.07	0
4	OSU	B	805	-	31,31,33	5.50	9 (29%)	44,44,46	2.15	12 (27%)

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	801	-	-	0/7/31/31	0/3/3/3
2	SAH	A	802	-	-	0/7/31/31	0/3/3/3
3	GOL	A	803	-	-	0/4/4/4	0/0/0/0
3	GOL	A	804	-	-	0/4/4/4	0/0/0/0
4	OSU	A	805	-	-	0/21/60/61	0/2/2/2
2	SAH	B	801	-	-	0/7/31/31	0/3/3/3
2	SAH	B	802	-	-	0/7/31/31	0/3/3/3
3	GOL	B	803	-	-	0/4/4/4	0/0/0/0
3	GOL	B	804	-	-	0/4/4/4	0/0/0/0
4	OSU	B	805	-	-	0/20/59/61	0/2/2/2

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	805	OSU	C4-C5	-20.38	1.09	1.53
4	A	805	OSU	C4-C5	-20.04	1.10	1.53
4	B	805	OSU	C1-C2	-14.69	1.10	1.52
4	A	805	OSU	C1-C2	-13.62	1.13	1.52

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	801	SAH	C4-N3	-3.54	1.30	1.35

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	802	SAH	N3-C2-N1	-6.82	122.92	128.86
2	B	802	SAH	N3-C2-N1	-6.51	123.19	128.86
2	B	802	SAH	CB-CG-SD	-6.09	101.83	113.57
2	B	801	SAH	N3-C2-N1	-5.78	123.82	128.86
4	A	805	OSU	O2-C2-C3	-5.51	98.36	110.36

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	803	GOL	1	0
3	A	804	GOL	4	0
4	A	805	OSU	4	0
2	B	801	SAH	2	0
3	B	803	GOL	2	0
3	B	804	GOL	1	0
4	B	805	OSU	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	667/703 (94%)	-0.23	5 (0%) 87 86	25, 45, 73, 96	0
1	B	641/703 (91%)	0.05	26 (4%) 38 36	29, 57, 91, 125	0
All	All	1308/1406 (93%)	-0.09	31 (2%) 59 57	25, 50, 86, 125	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	150	ILE	5.6
1	B	64	CYS	3.9
1	B	61	LEU	3.7
1	B	93	VAL	3.6
1	B	107	TYR	3.4

### 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
3	GOL	A	804	6/6	0.67	0.21	4.21	63,67,68,70	0
3	GOL	B	804	6/6	0.74	0.21	3.87	59,65,66,72	0
3	GOL	A	803	6/6	0.84	0.15	1.61	35,43,43,49	0
3	GOL	B	803	6/6	0.88	0.15	0.33	56,61,63,66	0
2	SAH	A	802	26/26	0.96	0.12	0.26	33,40,44,45	0
4	OSU	B	805	30/32	0.94	0.14	0.19	37,44,60,62	0
2	SAH	B	801	26/26	0.96	0.12	0.11	44,50,64,64	0
2	SAH	A	801	26/26	0.98	0.10	-0.22	29,37,41,43	0
4	OSU	A	805	31/32	0.96	0.13	-0.28	27,36,54,58	0
2	SAH	B	802	26/26	0.96	0.11	-0.70	38,41,46,53	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.