



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 13, 2020 – 05:26 PM BST

PDB ID : 4MVY  
Title : Trypanosoma brucei methionyl-tRNA synthetase in complex with inhibitor 1-{3-[(3,5-dichlorobenzyl)amino]propyl}-3-(3-hydroxyphenyl)urea (Chem 1387)  
Authors : Koh, C.Y.; Kim, J.E.; Wetzel, A.B.; de van der Schueren, W.J.; Shibata, S.; Liu, J.; Zhang, Z.; Fan, E.; Verlinde, C.L.M.J.; Hol, W.G.J.  
Deposited on : 2013-09-24  
Resolution : 2.31 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.14.4.dev1  
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) : 2.14.4.dev1

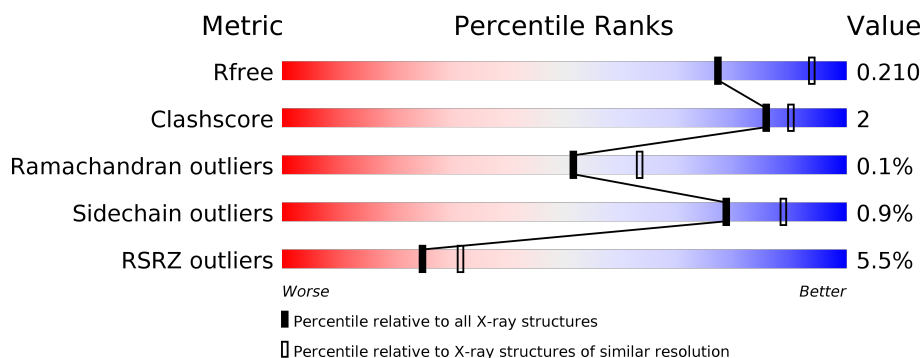
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.31 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	5974 (2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	5855 (2.34-2.30)

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 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  $\leq 5\%$ . 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	542	<div> <div>5%</div> <div> <div></div> <div>92%</div> <div>5% .</div> </div> </div>
1	B	542	<div> <div>6%</div> <div> <div></div> <div>94%</div> <div>. .</div> </div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 9184 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 Methionyl-tRNA synthetase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	531	Total	As	C	N	O	S	0	2	0
			4268	1	2747	722	787	11			
1	B	530	Total	As	C	N	O	S	0	4	0
			4263	1	2746	725	780	11			

There are 22 discrepancies between the modelled and reference sequences:

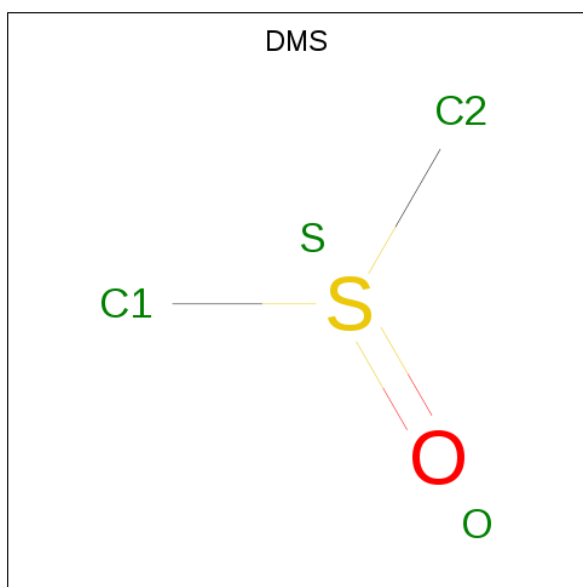
Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	EXPRESSION TAG	UNP Q38C91
A	-3	PRO	-	EXPRESSION TAG	UNP Q38C91
A	-2	GLY	-	EXPRESSION TAG	UNP Q38C91
A	-1	SER	-	EXPRESSION TAG	UNP Q38C91
A	0	MET	-	EXPRESSION TAG	UNP Q38C91
A	309	THR	ALA	CONFLICT	UNP Q38C91
A	452	ALA	LYS	ENGINEERED MUTATION	UNP Q38C91
A	453	ARG	LYS	ENGINEERED MUTATION	UNP Q38C91
A	454	ALA	GLU	ENGINEERED MUTATION	UNP Q38C91
A	499	VAL	ALA	CONFLICT	UNP Q38C91
A	503	ASN	SER	CONFLICT	UNP Q38C91
B	-4	GLY	-	EXPRESSION TAG	UNP Q38C91
B	-3	PRO	-	EXPRESSION TAG	UNP Q38C91
B	-2	GLY	-	EXPRESSION TAG	UNP Q38C91
B	-1	SER	-	EXPRESSION TAG	UNP Q38C91
B	0	MET	-	EXPRESSION TAG	UNP Q38C91
B	309	THR	ALA	CONFLICT	UNP Q38C91
B	452	ALA	LYS	ENGINEERED MUTATION	UNP Q38C91
B	453	ARG	LYS	ENGINEERED MUTATION	UNP Q38C91
B	454	ALA	GLU	ENGINEERED MUTATION	UNP Q38C91
B	499	VAL	ALA	CONFLICT	UNP Q38C91
B	503	ASN	SER	CONFLICT	UNP Q38C91

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



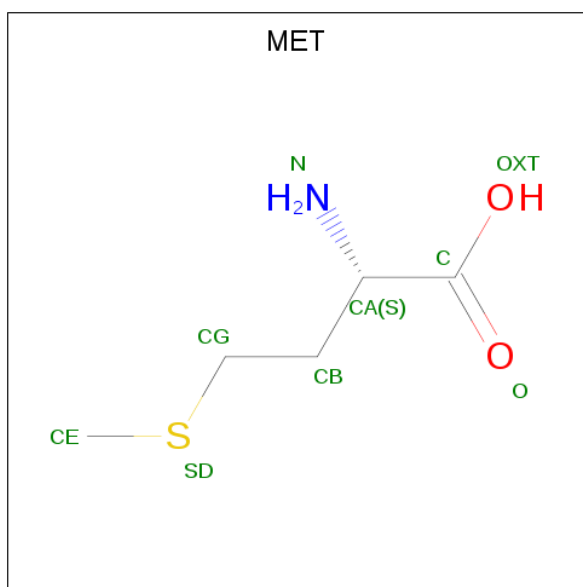
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C<sub>2</sub>H<sub>6</sub>OS).



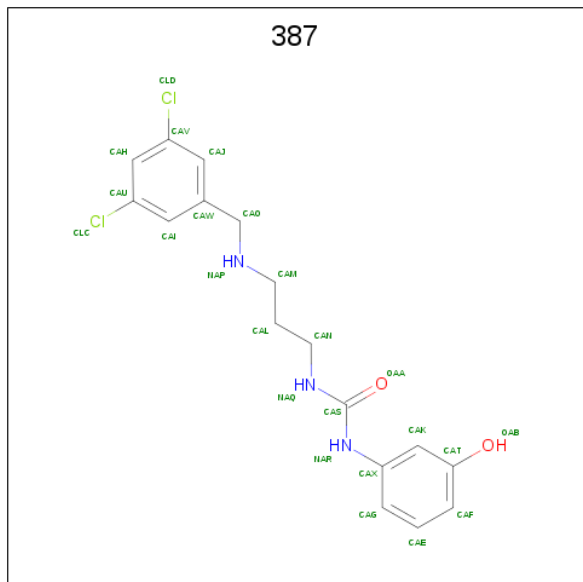
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	O	S	0	0
			4	2	1	1		
3	A	1	Total	C	O	S	0	0
			4	2	1	1		
3	B	1	Total	C	O	S	0	0
			4	2	1	1		

- Molecule 4 is METHIONINE (three-letter code: MET) (formula:  $C_5H_{11}NO_2S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			9	5	1	2	1		

- Molecule 5 is 1-{3-[(3,5-dichlorobenzyl)amino]propyl}-3-(3-hydroxyphenyl)urea (three-letter code: 387) (formula: C<sub>17</sub>H<sub>19</sub>Cl<sub>2</sub>N<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total	C	Cl	N	O	0	0
			24	17	2	3	2		

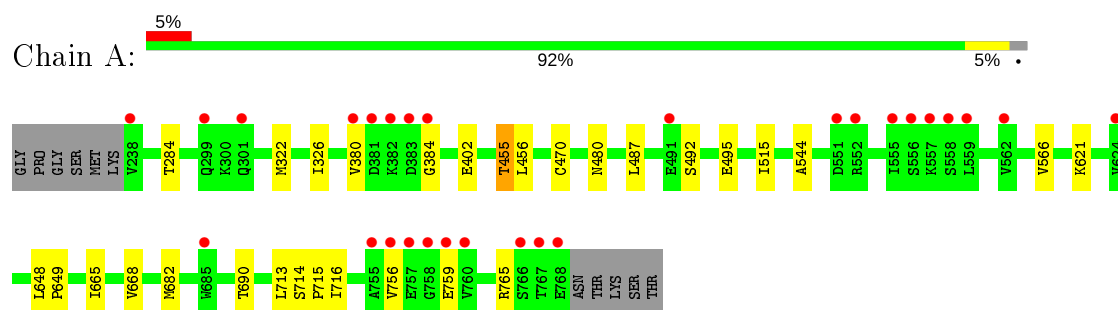
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	287	Total	O	0	0
			287	287		
6	B	273	Total	O	0	0
			273	273		

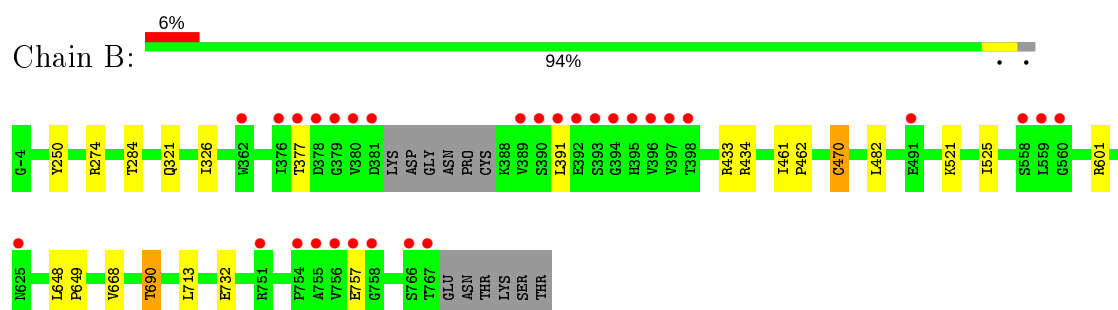
### 3 Residue-property plots [i](#)

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: Methionyl-tRNA synthetase



- Molecule 1: Methionyl-tRNA synthetase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.24Å 106.00Å 206.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.31 48.27 – 2.31	Depositor EDS
% Data completeness (in resolution range)	97.0 (30.00-2.31) 97.1 (48.27-2.31)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.37 (at 2.32Å)	Xtriage
Refinement program	REFMAC refmac_5.7.0032	Depositor
R, $R_{free}$	0.193 , 0.219 0.186 , 0.210	Depositor DCC
$R_{free}$ test set	4124 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.1	Xtriage
Anisotropy	0.372	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 31.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9184	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.11% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CAS, GOL, DMS, 387

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	0.40	0/4375	0.60	0/5945
1	B	0.40	0/4369	0.60	2/5934 (0.0%)
All	All	0.40	0/8744	0.60	2/11879 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	433	ARG	NE-CZ-NH2	-5.60	117.50	120.30
1	B	274	ARG	NE-CZ-NH2	-5.18	117.71	120.30

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	4268	0	4219	21	0
1	B	4263	0	4222	17	0
2	A	24	0	32	0	0
2	B	24	0	32	3	0
3	A	8	0	12	0	0
3	B	4	0	6	0	0
4	A	9	0	8	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	24	0	18	0	0
6	A	287	0	0	4	0
6	B	273	0	0	2	0
All	All	9184	0	8549	37	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 2.

All (37) 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:B:461:ILE:C	1:B:470:CAS:CE2	2.25	1.05
1:A:765[B]:ARG:HG2	1:A:765[B]:ARG:HH11	1.15	1.04
1:A:765[B]:ARG:HH11	1:A:765[B]:ARG:CG	1.87	0.86
1:B:321:GLN:HG3	6:B:1110:HOH:O	1.85	0.76
1:B:461:ILE:O	1:B:470:CAS:CE2	2.35	0.73
1:A:487:LEU:HD22	1:A:495:GLU:HG3	1.69	0.72
1:A:765[B]:ARG:HG2	1:A:765[B]:ARG:NH1	1.97	0.72
1:B:462:PRO:N	1:B:470:CAS:CE2	2.52	0.72
1:B:482:LEU:HD21	2:B:801:GOL:H31	1.84	0.60
1:A:765[B]:ARG:NH1	1:A:765[B]:ARG:CG	2.53	0.58
1:A:756:VAL:N	1:A:759:GLU:OE2	2.34	0.57
1:A:665:ILE:HG12	1:A:716:ILE:HD13	1.87	0.57
1:A:284:THR:HG22	1:A:326:ILE:HG21	1.86	0.56
1:A:668:VAL:HG11	1:A:713:LEU:HG	1.86	0.56
1:A:456:LEU:HD21	6:A:929:HOH:O	2.05	0.56
1:A:456:LEU:HD12	1:A:470:CAS:HB3	1.87	0.55
1:A:621:LYS:NZ	6:A:1018:HOH:O	2.40	0.54
6:A:1143:HOH:O	1:B:690:THR:CG2	2.55	0.54
1:B:732:GLU:HG2	6:B:973:HOH:O	2.07	0.53
1:A:402:GLU:OE2	1:A:455:THR:HG21	2.11	0.50
1:B:521[A]:LYS:HG3	1:B:525:ILE:HD12	1.95	0.49
1:B:668:VAL:HG11	1:B:713:LEU:HG	1.94	0.48
1:A:380:VAL:CG1	1:A:384:GLY:HA2	2.43	0.47
1:B:284:THR:HG22	1:B:326:ILE:HG21	1.97	0.45
1:A:322:MET:HE3	1:A:566:VAL:HG22	1.99	0.45
1:A:487:LEU:HD22	1:A:495:GLU:CG	2.43	0.44
1:A:714:SER:N	1:A:715:PRO:CD	2.81	0.44
1:A:515:ILE:O	1:A:544:ALA:HA	2.17	0.44
1:B:462:PRO:CA	1:B:470:CAS:CE2	2.97	0.43
1:B:482:LEU:HD21	2:B:801:GOL:C3	2.48	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:682:MET:O	1:B:434[B]:ARG:NH2	2.52	0.42
1:B:648:LEU:HB3	1:B:649:PRO:HD3	2.02	0.41
1:B:462:PRO:HA	1:B:470:CAS:CE2	2.51	0.41
6:A:1143:HOH:O	1:B:690:THR:HG23	2.18	0.41
1:A:648:LEU:HB3	1:A:649:PRO:HD3	2.02	0.41
1:B:601:ARG:HH22	2:B:802:GOL:H11	1.86	0.40
1:A:756:VAL:HB	1:A:759:GLU:HB3	2.03	0.40

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	530/542 (98%)	518 (98%)	12 (2%)	0	100	100
1	B	529/542 (98%)	518 (98%)	10 (2%)	1 (0%)	47	58
All	All	1059/1084 (98%)	1036 (98%)	22 (2%)	1 (0%)	51	63

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	757	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	460/468 (98%)	456 (99%)	4 (1%)	78	89
1	B	457/468 (98%)	453 (99%)	4 (1%)	78	89
All	All	917/936 (98%)	909 (99%)	8 (1%)	78	89

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	455	THR
1	A	480	ASN
1	A	492	SER
1	A	690	THR
1	B	250	TYR
1	B	377	THR
1	B	391	LEU
1	B	690	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
1	CAS	A	470	1	5,8,9	0.97	0	1,9,11	0.53	0
1	CAS	B	470	1	5,8,9	1.33	1 (20%)	1,9,11	0.11	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
1	CAS	A	470	1	-	0/0/7/9	-
1	CAS	B	470	1	-	0/0/7/9	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	470	CAS	AS-CE2	-2.26	1.90	1.96

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	470	CAS	1	0
1	B	470	CAS	5	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

13 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	GOL	A	801	-	5,5,5	0.43	0	5,5,5	0.67	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	DMS	A	805	-	3,3,3	0.51	0	3,3,3	0.53	0
4	MET	A	807	-	4,8,8	0.64	0	2,9,9	0.15	0
2	GOL	A	802	-	5,5,5	0.45	0	5,5,5	0.50	0
2	GOL	B	802	-	5,5,5	0.23	0	5,5,5	0.45	0
5	387	B	806	-	25,25,25	0.34	0	32,32,32	0.47	0
2	GOL	A	806	-	5,5,5	0.31	0	5,5,5	0.57	0
2	GOL	B	801	-	5,5,5	0.33	0	5,5,5	0.83	0
3	DMS	B	805	-	3,3,3	0.46	0	3,3,3	0.85	0
2	GOL	A	803	-	5,5,5	0.30	0	5,5,5	0.18	0
3	DMS	A	804	-	3,3,3	0.51	0	3,3,3	0.66	0
2	GOL	B	803	-	5,5,5	0.31	0	5,5,5	0.19	0
2	GOL	B	804	-	5,5,5	0.20	0	5,5,5	0.31	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	GOL	A	801	-	-	2/4/4/4	-
4	MET	A	807	-	-	0/4/8/8	-
2	GOL	A	802	-	-	3/4/4/4	-
2	GOL	B	802	-	-	2/4/4/4	-
5	387	B	806	-	-	0/13/13/13	0/2/2/2
2	GOL	A	806	-	-	4/4/4/4	-
2	GOL	B	801	-	-	0/4/4/4	-
2	GOL	A	803	-	-	0/4/4/4	-
2	GOL	B	803	-	-	0/4/4/4	-
2	GOL	B	804	-	-	3/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	801	GOL	C1-C2-C3-O3
2	A	802	GOL	O1-C1-C2-C3
2	A	806	GOL	O1-C1-C2-C3
2	B	802	GOL	C1-C2-C3-O3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	A	806	GOL	C1-C2-C3-O3
2	B	804	GOL	O1-C1-C2-C3
2	A	801	GOL	O2-C2-C3-O3
2	A	802	GOL	O1-C1-C2-O2
2	A	806	GOL	O1-C1-C2-O2
2	B	804	GOL	O2-C2-C3-O3
2	A	806	GOL	O2-C2-C3-O3
2	B	802	GOL	O2-C2-C3-O3
2	B	804	GOL	O1-C1-C2-O2
2	A	802	GOL	O2-C2-C3-O3

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	802	GOL	1	0
2	B	801	GOL	2	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

### 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	530/542 (97%)	-0.13	28 (5%) 26 33	19, 32, 72, 123	0
1	B	529/542 (97%)	-0.05	30 (5%) 23 30	19, 31, 92, 146	0
All	All	1059/1084 (97%)	-0.09	58 (5%) 25 31	19, 31, 81, 146	0

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	757	GLU	6.6
1	B	380	VAL	6.2
1	B	559	LEU	5.3
1	A	768	GLU	5.2
1	B	757	GLU	5.2
1	A	383	ASP	5.2
1	B	756	VAL	5.0
1	B	758	GLY	5.0
1	A	756	VAL	5.0
1	B	377	THR	4.9
1	A	384	GLY	4.8
1	A	767	THR	4.7
1	A	766	SER	4.7
1	A	556	SER	4.5
1	B	381	ASP	4.5
1	A	759	GLU	4.4
1	B	396	VAL	4.2
1	A	552	ARG	4.1
1	B	391	LEU	4.0
1	A	758	GLY	3.9
1	A	559	LEU	3.7
1	A	299	GLN	3.7
1	B	767	THR	3.7
1	A	238	VAL	3.7

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	379	GLY	3.5
1	B	389	VAL	3.5
1	B	755	ALA	3.4
1	B	378	ASP	3.3
1	A	558	SER	3.3
1	B	558	SER	3.3
1	B	395	HIS	3.2
1	A	760	VAL	3.0
1	B	393	SER	2.9
1	B	394	GLY	2.9
1	A	755	ALA	2.8
1	B	397	VAL	2.7
1	B	376	ILE	2.6
1	A	301	GLN	2.6
1	B	754	PRO	2.6
1	A	382	LYS	2.6
1	A	381	ASP	2.6
1	B	560	GLY	2.5
1	B	362	TRP	2.5
1	B	766	SER	2.5
1	A	557	LYS	2.5
1	A	380	VAL	2.4
1	B	491	GLU	2.4
1	B	398	THR	2.3
1	B	390	SER	2.3
1	A	491	GLU	2.3
1	A	624	VAL	2.3
1	A	685	TRP	2.3
1	B	751	ARG	2.2
1	B	625	ASN	2.2
1	A	562	VAL	2.2
1	B	392	GLU	2.2
1	A	551	ASP	2.1
1	A	555	ILE	2.1

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

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<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	B-factors( $\text{\AA}^2$ )	Q<0.9
1	CAS	A	470	9/10	0.98	0.08	26,29,37,38	3
1	CAS	B	470	9/10	0.98	0.11	25,27,32,33	3

### 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, 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	B-factors( $\text{\AA}^2$ )	Q<0.9
2	GOL	B	801	6/6	0.77	0.18	41,47,49,52	0
2	GOL	A	803	6/6	0.79	0.20	59,65,67,67	0
2	GOL	A	801	6/6	0.81	0.18	60,62,63,64	0
3	DMS	B	805	4/4	0.84	0.26	69,78,79,84	0
3	DMS	A	804	4/4	0.84	0.19	87,88,88,91	0
2	GOL	B	804	6/6	0.84	0.17	64,66,69,71	0
2	GOL	A	802	6/6	0.87	0.15	45,53,54,57	0
2	GOL	B	803	6/6	0.87	0.20	42,47,49,51	0
3	DMS	A	805	4/4	0.87	0.17	81,86,87,88	0
2	GOL	A	806	6/6	0.88	0.18	47,49,51,52	0
2	GOL	B	802	6/6	0.93	0.16	47,49,50,51	0
5	387	B	806	24/24	0.96	0.10	24,30,37,41	0
4	MET	A	807	9/9	0.97	0.12	24,25,26,27	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.