



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 13, 2024 – 08:43 AM EST

PDB ID : 4ISS  
Title : SeMet-substituted Kluyveromyces lactis Allophanate Hydrolase  
Authors : Fan, C.; Xiang, S.  
Deposited on : 2013-01-17  
Resolution : 2.50 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) 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 : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

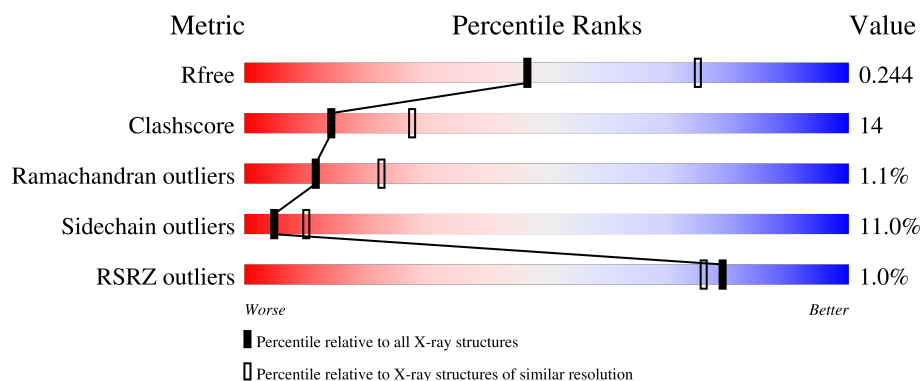
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
$R_{free}$	164625	5504 (2.50-2.50)
Clashscore	180529	6282 (2.50-2.50)
Ramachandran outliers	177936	6191 (2.50-2.50)
Sidechain outliers	177891	6193 (2.50-2.50)
RSRZ outliers	164620	5504 (2.50-2.50)

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  $\geq 3$ , 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	644	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 10%, orange 21%, yellow 5%, green 5%, grey 64%);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> <span>10%</span> <span>69%</span> <span>21%</span> <span>5%</span> <span>5%</span> </div> </div>
1	B	644	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 10%, orange 21%, yellow 5%, green 5%, grey 64%);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> <span>10%</span> <span>68%</span> <span>21%</span> <span>5%</span> <span>5%</span> </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	B	701	-	-	X	-

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 9768 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 Allophanate Hydrolase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	614	Total	C	N	O	S	Se	0	0	0
			4748	3051	786	897	11	3			
1	B	612	Total	C	N	O	S	Se	0	0	0
			4731	3041	784	893	11	2			

There are 46 discrepancies between the modelled and reference sequences:

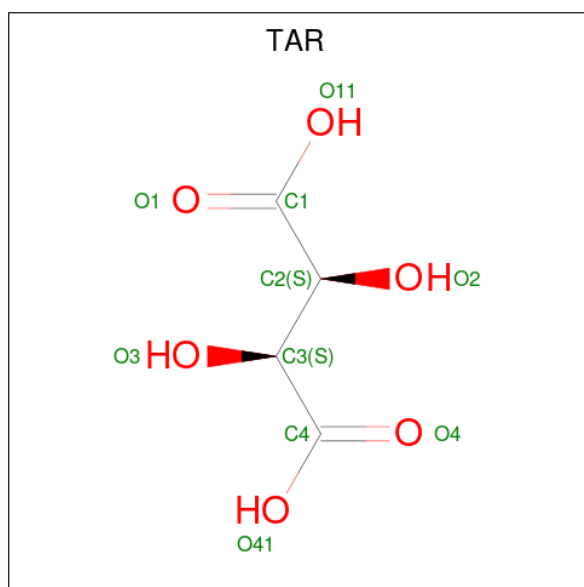
Chain	Residue	Modelled	Actual	Comment	Reference
A	-22	MSE	-	expression tag	UNP Q6CP22
A	-21	GLY	-	expression tag	UNP Q6CP22
A	-20	SER	-	expression tag	UNP Q6CP22
A	-19	SER	-	expression tag	UNP Q6CP22
A	-18	HIS	-	expression tag	UNP Q6CP22
A	-17	HIS	-	expression tag	UNP Q6CP22
A	-16	HIS	-	expression tag	UNP Q6CP22
A	-15	HIS	-	expression tag	UNP Q6CP22
A	-14	HIS	-	expression tag	UNP Q6CP22
A	-13	HIS	-	expression tag	UNP Q6CP22
A	-12	SER	-	expression tag	UNP Q6CP22
A	-11	SER	-	expression tag	UNP Q6CP22
A	-10	GLY	-	expression tag	UNP Q6CP22
A	-9	LEU	-	expression tag	UNP Q6CP22
A	-8	VAL	-	expression tag	UNP Q6CP22
A	-7	PRO	-	expression tag	UNP Q6CP22
A	-6	ARG	-	expression tag	UNP Q6CP22
A	-5	GLY	-	expression tag	UNP Q6CP22
A	-4	SER	-	expression tag	UNP Q6CP22
A	-3	HIS	-	expression tag	UNP Q6CP22
A	-2	MSE	-	expression tag	UNP Q6CP22
A	-1	ALA	-	expression tag	UNP Q6CP22
A	0	SER	-	expression tag	UNP Q6CP22
B	-22	MSE	-	expression tag	UNP Q6CP22
B	-21	GLY	-	expression tag	UNP Q6CP22

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-20	SER	-	expression tag	UNP Q6CP22
B	-19	SER	-	expression tag	UNP Q6CP22
B	-18	HIS	-	expression tag	UNP Q6CP22
B	-17	HIS	-	expression tag	UNP Q6CP22
B	-16	HIS	-	expression tag	UNP Q6CP22
B	-15	HIS	-	expression tag	UNP Q6CP22
B	-14	HIS	-	expression tag	UNP Q6CP22
B	-13	HIS	-	expression tag	UNP Q6CP22
B	-12	SER	-	expression tag	UNP Q6CP22
B	-11	SER	-	expression tag	UNP Q6CP22
B	-10	GLY	-	expression tag	UNP Q6CP22
B	-9	LEU	-	expression tag	UNP Q6CP22
B	-8	VAL	-	expression tag	UNP Q6CP22
B	-7	PRO	-	expression tag	UNP Q6CP22
B	-6	ARG	-	expression tag	UNP Q6CP22
B	-5	GLY	-	expression tag	UNP Q6CP22
B	-4	SER	-	expression tag	UNP Q6CP22
B	-3	HIS	-	expression tag	UNP Q6CP22
B	-2	MSE	-	expression tag	UNP Q6CP22
B	-1	ALA	-	expression tag	UNP Q6CP22
B	0	SER	-	expression tag	UNP Q6CP22

- Molecule 2 is D(-)-TARTARIC ACID (three-letter code: TAR) (formula: C<sub>4</sub>H<sub>6</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			10	4	6		

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



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

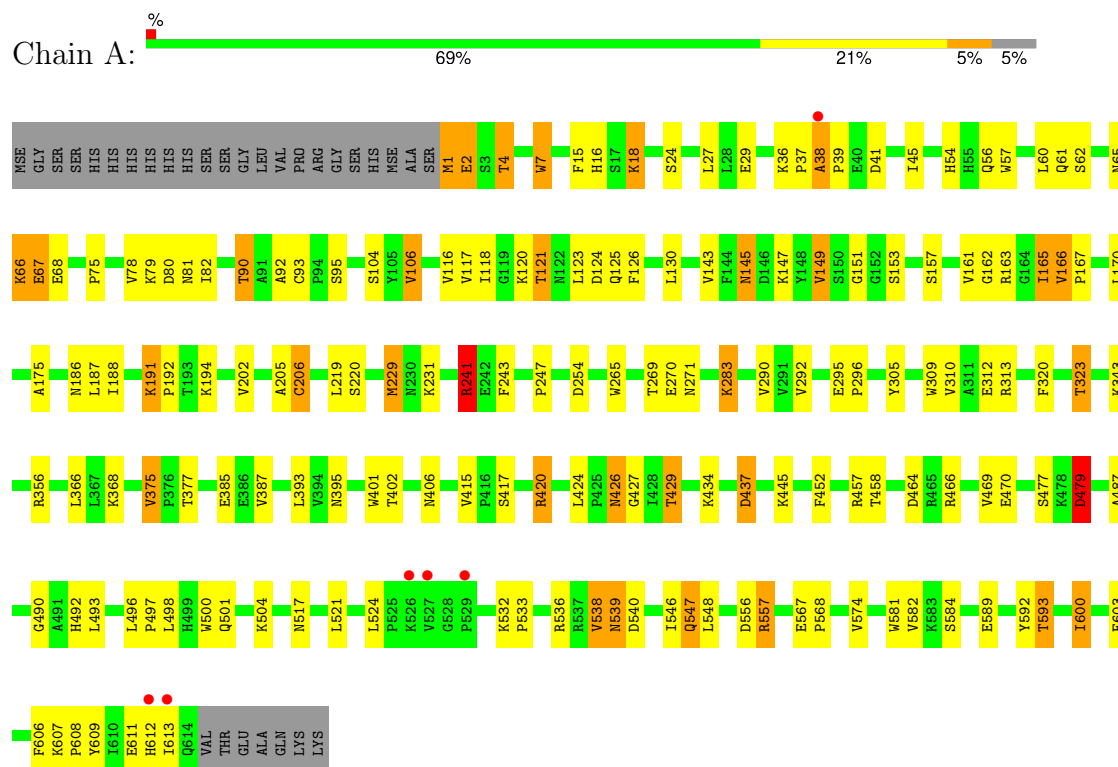
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	124	Total	O	0	0
			124	124		
4	B	149	Total	O	0	0
			149	149		

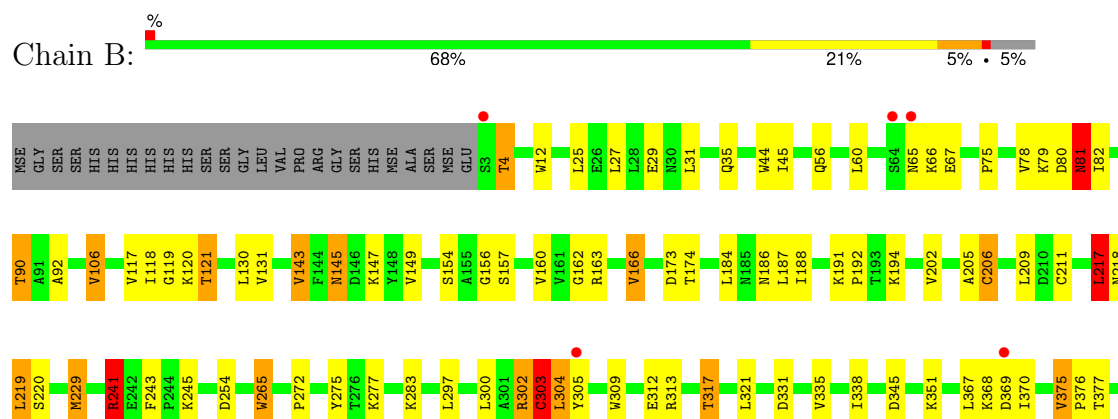
### 3 Residue-property plots [i](#)

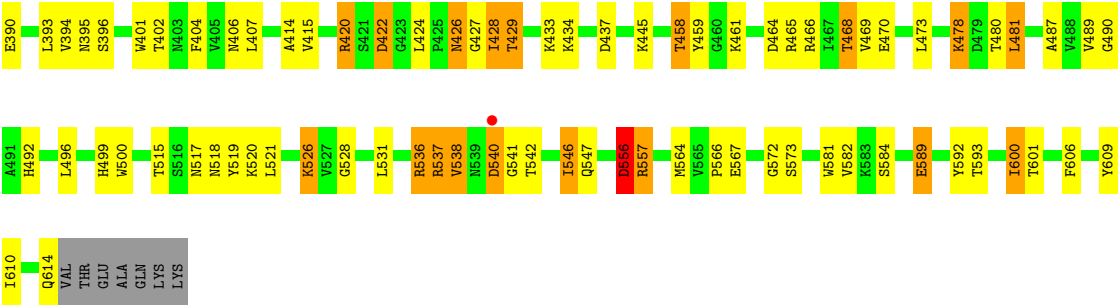
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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



#### • Molecule 1: Allophanate Hydrolase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.05Å 107.72Å 150.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.53 – 2.50 46.53 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.8 (46.53-2.50) 99.9 (46.53-2.50)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.18 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.176 , 0.247 0.177 , 0.244	Depositor DCC
$R_{free}$ test set	2715 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.2	Xtriage
Anisotropy	0.187	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 30.3	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	0.95	EDS
Total number of atoms	9768	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% 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: TAR, GOL

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	0.77	5/4860 (0.1%)	0.93	9/6612 (0.1%)
1	B	0.83	6/4843 (0.1%)	0.96	15/6590 (0.2%)
All	All	0.80	11/9703 (0.1%)	0.95	24/13202 (0.2%)

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

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	254	ASP	CB-CG	-6.13	1.38	1.51
1	A	265	TRP	CD2-CE2	5.77	1.48	1.41
1	A	309	TRP	CD2-CE2	5.69	1.48	1.41
1	A	581	TRP	CD2-CE2	5.68	1.48	1.41
1	A	500	TRP	CD2-CE2	5.64	1.48	1.41

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	241	ARG	NE-CZ-NH1	8.53	124.56	120.30
1	A	188	ILE	CG1-CB-CG2	-7.83	94.18	111.40
1	B	241	ARG	NE-CZ-NH1	7.27	123.93	120.30
1	A	313	ARG	NE-CZ-NH2	-7.23	116.68	120.30
1	A	254	ASP	CB-CA-C	-6.95	96.50	110.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	539	ASN	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	4748	0	4774	123	0
1	B	4731	0	4756	144	0
2	A	10	0	4	1	0
3	B	6	0	8	5	0
4	A	124	0	0	6	0
4	B	149	0	0	13	0
All	All	9768	0	9542	264	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 14.

The worst 5 of 264 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:106:VAL:HG22	1:A:229:MSE:HE2	1.29	1.14
1:B:600:ILE:HD11	1:B:609:TYR:CE1	1.82	1.14
1:B:90:THR:HG22	1:B:92:ALA:H	1.15	1.10
1:A:38:ALA:HB1	1:A:39:PRO:HD3	1.36	1.07
1:B:186:ASN:HD21	1:B:458:THR:HG22	1.13	1.06

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	612/644 (95%)	563 (92%)	42 (7%)	7 (1%)	12	23
1	B	610/644 (95%)	580 (95%)	23 (4%)	7 (1%)	12	23
All	All	1222/1288 (95%)	1143 (94%)	65 (5%)	14 (1%)	12	23

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	38	ALA
1	B	66	LYS
1	B	303	CYS
1	A	206	CYS
1	A	479	ASP

### 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	524/544 (96%)	469 (90%)	55 (10%)	5	11
1	B	522/544 (96%)	462 (88%)	60 (12%)	4	9
All	All	1046/1088 (96%)	931 (89%)	115 (11%)	5	10

5 of 115 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	27	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	573	SER
1	B	217	LEU
1	B	567	GLU
1	B	478	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	58	ASN
1	B	357	GLN
1	B	186	ASN
1	B	395	ASN
1	A	406	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 oligosaccharides 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	TAR	A	701	-	9,9,9	1.26	0	12,12,12	1.72	3 (25%)
3	GOL	B	701	-	5,5,5	0.50	0	5,5,5	1.00	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	TAR	A	701	-	-	6/12/12/12	-
3	GOL	B	701	-	-	0/4/4/4	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	701	TAR	O2-C2-C3	-3.19	103.67	110.17
2	A	701	TAR	O11-C1-C2	3.11	121.96	113.31
2	A	701	TAR	O1-C1-C2	-2.19	115.80	121.62

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	701	TAR	O1-C1-C2-O2
2	A	701	TAR	O11-C1-C2-O2
2	A	701	TAR	C2-C3-C4-O4
2	A	701	TAR	C2-C3-C4-O41
2	A	701	TAR	O3-C3-C4-O41

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701	TAR	1	0
3	B	701	GOL	5	0

## 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 [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	611/644 (94%)	-0.55	6 (0%) 79 76	18, 36, 70, 116	0
1	B	610/644 (94%)	-0.63	6 (0%) 79 76	18, 33, 60, 103	0
All	All	1221/1288 (94%)	-0.59	12 (0%) 79 76	18, 34, 66, 116	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	3	SER	3.8
1	B	305	TYR	3.5
1	A	613	ILE	2.9
1	A	612	HIS	2.6
1	A	529	PRO	2.6

### 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 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	TAR	A	701	10/10	0.91	0.11	52,61,80,81	0
3	GOL	B	701	6/6	0.94	0.11	43,44,52,55	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.