



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:46 PM GMT

PDB ID : 4JFA  
Title : Crystal Structure of Plasmodium falciparum Tryptophanyl-tRNA synthetase  
Authors : Khan, S.; Garg, A.; Manickam, Y.; Sharma, A.  
Deposited on : 2013-02-28  
Resolution : 2.60 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

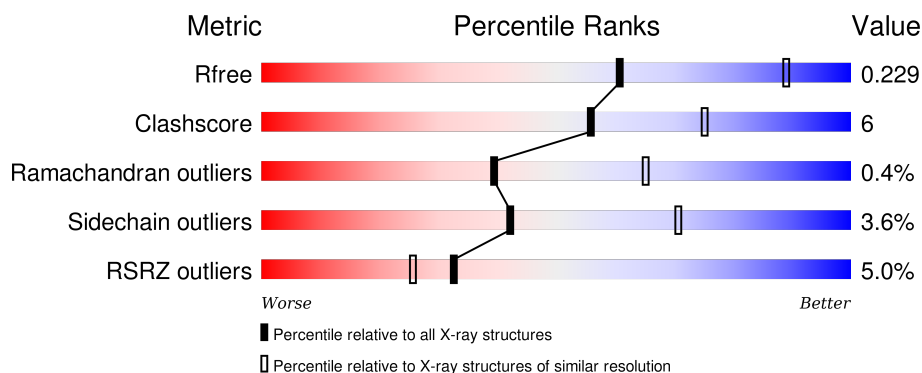
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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	420	<div> <div>3%</div> <div>72% 13% • 14%</div> </div>
1	B	420	<div> <div>4%</div> <div>75% 12% • 12%</div> </div>
1	C	420	<div> <div>2%</div> <div>68% 14% • 17%</div> </div>
1	D	420	<div> <div>8%</div> <div>70% 14% • 15%</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11674 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 Tryptophan–tRNA ligase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	362	Total	C	N	O	S	0	0	0
			2891	1870	479	531	11			
1	B	368	Total	C	N	O	S	0	2	0
			2958	1914	489	543	12			
1	C	348	Total	C	N	O	S	0	1	0
			2754	1783	455	505	11			
1	D	357	Total	C	N	O	S	0	0	0
			2758	1782	459	506	11			

There are 60 discrepancies between the modelled and reference sequences:

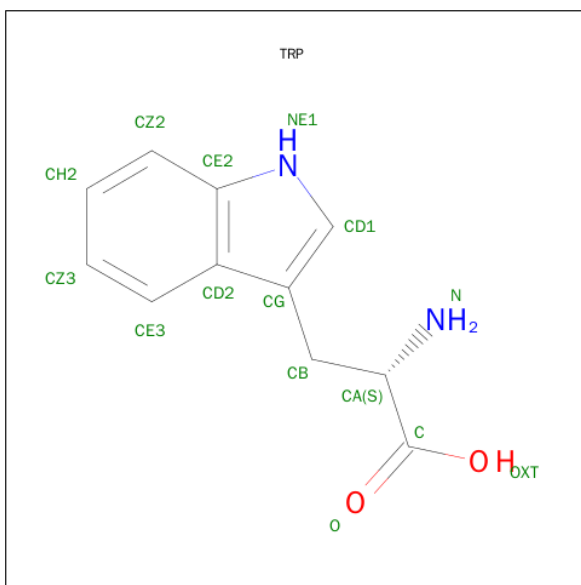
Chain	Residue	Modelled	Actual	Comment	Reference
A	226	MET	-	EXPRESSION TAG	UNP Q8IDW3
A	227	ALA	-	EXPRESSION TAG	UNP Q8IDW3
A	633	GLU	-	EXPRESSION TAG	UNP Q8IDW3
A	634	ASN	-	EXPRESSION TAG	UNP Q8IDW3
A	635	LEU	-	EXPRESSION TAG	UNP Q8IDW3
A	636	TYR	-	EXPRESSION TAG	UNP Q8IDW3
A	637	PHE	-	EXPRESSION TAG	UNP Q8IDW3
A	638	GLN	-	EXPRESSION TAG	UNP Q8IDW3
A	639	GLY	-	EXPRESSION TAG	UNP Q8IDW3
A	640	HIS	-	EXPRESSION TAG	UNP Q8IDW3
A	641	HIS	-	EXPRESSION TAG	UNP Q8IDW3
A	642	HIS	-	EXPRESSION TAG	UNP Q8IDW3
A	643	HIS	-	EXPRESSION TAG	UNP Q8IDW3
A	644	HIS	-	EXPRESSION TAG	UNP Q8IDW3
A	645	HIS	-	EXPRESSION TAG	UNP Q8IDW3
B	226	MET	-	EXPRESSION TAG	UNP Q8IDW3
B	227	ALA	-	EXPRESSION TAG	UNP Q8IDW3
B	633	GLU	-	EXPRESSION TAG	UNP Q8IDW3
B	634	ASN	-	EXPRESSION TAG	UNP Q8IDW3
B	635	LEU	-	EXPRESSION TAG	UNP Q8IDW3
B	636	TYR	-	EXPRESSION TAG	UNP Q8IDW3

*Continued on next page...*

*Continued from previous page...*

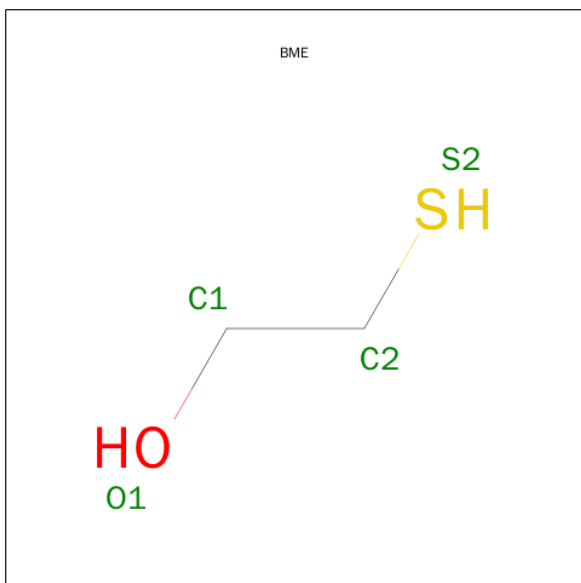
Chain	Residue	Modelled	Actual	Comment	Reference
B	637	PHE	-	EXPRESSION TAG	UNP Q8IDW3
B	638	GLN	-	EXPRESSION TAG	UNP Q8IDW3
B	639	GLY	-	EXPRESSION TAG	UNP Q8IDW3
B	640	HIS	-	EXPRESSION TAG	UNP Q8IDW3
B	641	HIS	-	EXPRESSION TAG	UNP Q8IDW3
B	642	HIS	-	EXPRESSION TAG	UNP Q8IDW3
B	643	HIS	-	EXPRESSION TAG	UNP Q8IDW3
B	644	HIS	-	EXPRESSION TAG	UNP Q8IDW3
B	645	HIS	-	EXPRESSION TAG	UNP Q8IDW3
C	226	MET	-	EXPRESSION TAG	UNP Q8IDW3
C	227	ALA	-	EXPRESSION TAG	UNP Q8IDW3
C	633	GLU	-	EXPRESSION TAG	UNP Q8IDW3
C	634	ASN	-	EXPRESSION TAG	UNP Q8IDW3
C	635	LEU	-	EXPRESSION TAG	UNP Q8IDW3
C	636	TYR	-	EXPRESSION TAG	UNP Q8IDW3
C	637	PHE	-	EXPRESSION TAG	UNP Q8IDW3
C	638	GLN	-	EXPRESSION TAG	UNP Q8IDW3
C	639	GLY	-	EXPRESSION TAG	UNP Q8IDW3
C	640	HIS	-	EXPRESSION TAG	UNP Q8IDW3
C	641	HIS	-	EXPRESSION TAG	UNP Q8IDW3
C	642	HIS	-	EXPRESSION TAG	UNP Q8IDW3
C	643	HIS	-	EXPRESSION TAG	UNP Q8IDW3
C	644	HIS	-	EXPRESSION TAG	UNP Q8IDW3
C	645	HIS	-	EXPRESSION TAG	UNP Q8IDW3
D	226	MET	-	EXPRESSION TAG	UNP Q8IDW3
D	227	ALA	-	EXPRESSION TAG	UNP Q8IDW3
D	633	GLU	-	EXPRESSION TAG	UNP Q8IDW3
D	634	ASN	-	EXPRESSION TAG	UNP Q8IDW3
D	635	LEU	-	EXPRESSION TAG	UNP Q8IDW3
D	636	TYR	-	EXPRESSION TAG	UNP Q8IDW3
D	637	PHE	-	EXPRESSION TAG	UNP Q8IDW3
D	638	GLN	-	EXPRESSION TAG	UNP Q8IDW3
D	639	GLY	-	EXPRESSION TAG	UNP Q8IDW3
D	640	HIS	-	EXPRESSION TAG	UNP Q8IDW3
D	641	HIS	-	EXPRESSION TAG	UNP Q8IDW3
D	642	HIS	-	EXPRESSION TAG	UNP Q8IDW3
D	643	HIS	-	EXPRESSION TAG	UNP Q8IDW3
D	644	HIS	-	EXPRESSION TAG	UNP Q8IDW3
D	645	HIS	-	EXPRESSION TAG	UNP Q8IDW3

- Molecule 2 is TRYPTOPHAN (three-letter code: TRP) (formula:  $C_{11}H_{12}N_2O_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			15	11	2	2		
2	B	1	Total	C	N	O	0	0
			15	11	2	2		
2	C	1	Total	C	N	O	0	0
			15	11	2	2		
2	D	1	Total	C	N	O	0	0
			15	11	2	2		

- Molecule 3 is BETA-MERCAPTOETHANOL (three-letter code: BME) (formula: C<sub>2</sub>H<sub>6</sub>OS).



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

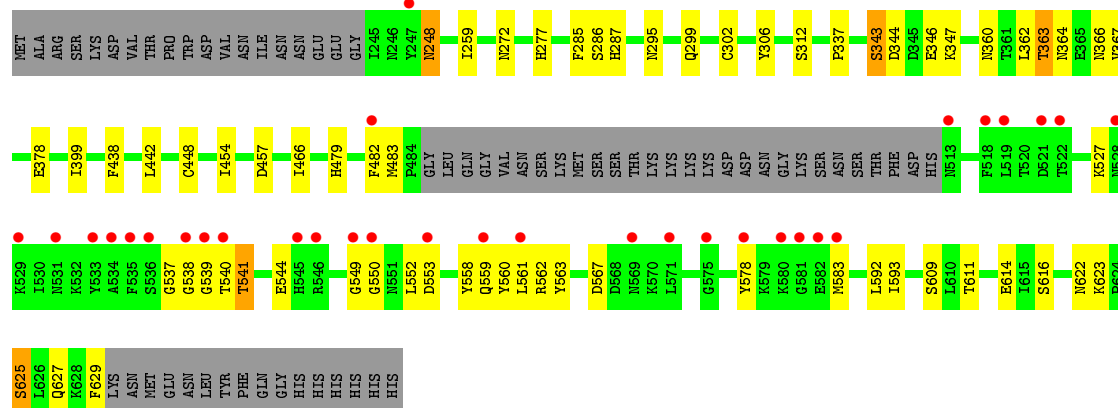
- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	D	1	Total K 1 1	0	0
4	C	1	Total K 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	74	Total O 74 74	0	0
5	B	79	Total O 79 79	0	0
5	C	43	Total O 43 43	0	0
5	D	43	Total O 43 43	0	0







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.18Å 213.25Å 101.95Å 90.00° 97.15° 90.00°	Depositor
Resolution (Å)	47.83 – 2.60 47.83 – 2.60	Depositor EDS
% Data completeness (in resolution range)	94.2 (47.83-2.60) 94.2 (47.83-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.18 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.191 , 0.229 0.193 , 0.229	Depositor DCC
$R_{free}$ test set	1275 reflections (2.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	42.1	Xtriage
Anisotropy	0.483	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 53.8	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	0 of 62543 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	11674	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.67% 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: K, BME

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	0/2964	0.45	0/4013
1	B	0.39	0/3039	0.47	0/4112
1	C	0.36	0/2827	0.51	2/3835 (0.1%)
1	D	0.31	0/2831	0.46	1/3845 (0.0%)
All	All	0.35	0/11661	0.47	3/15805 (0.0%)

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	C	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	541	THR	N-CA-C	-6.09	94.56	111.00
1	C	630	LYS	N-CA-C	-5.78	95.41	111.00
1	C	625	SER	N-CA-C	-5.40	96.41	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	626	LEU	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	2891	0	2781	36	0
1	B	2958	0	2835	27	0
1	C	2754	0	2602	42	0
1	D	2758	0	2507	34	0
2	A	15	0	9	0	0
2	B	15	0	9	0	0
2	C	15	0	9	0	0
2	D	15	0	9	1	0
3	A	4	0	5	1	0
3	B	4	0	5	0	0
3	D	4	0	5	1	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
5	A	74	0	0	0	0
5	B	79	0	0	0	0
5	C	43	0	0	1	0
5	D	43	0	0	0	0
All	All	11674	0	10776	138	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 6.

The worst 5 of 138 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:D:541:THR:CB	1:D:544:GLU:CB	2.37	1.01
1:A:530:ILE:HG22	1:A:589:LYS:HG3	1.66	0.76
1:B:546:ARG:O	1:B:546:ARG:NH1	2.20	0.74
1:A:248:ASN:OD1	1:A:258:LYS:CE	2.36	0.73
1:C:626:LEU:O	1:C:627:GLN:C	2.30	0.69

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	358/420 (85%)	349 (98%)	7 (2%)	2 (1%)	30	56
1	B	366/420 (87%)	360 (98%)	5 (1%)	1 (0%)	46	72
1	C	343/420 (82%)	335 (98%)	7 (2%)	1 (0%)	46	72
1	D	353/420 (84%)	338 (96%)	13 (4%)	2 (1%)	30	56
All	All	1420/1680 (84%)	1382 (97%)	32 (2%)	6 (0%)	39	65

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	537	GLY
1	A	539	GLY
1	C	625	SER
1	B	549	GLY
1	D	537	GLY

### 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	313/388 (81%)	301 (96%)	12 (4%)	40	68
1	B	318/388 (82%)	305 (96%)	13 (4%)	37	66
1	C	292/388 (75%)	284 (97%)	8 (3%)	52	79
1	D	274/388 (71%)	264 (96%)	10 (4%)	42	71
All	All	1197/1552 (77%)	1154 (96%)	43 (4%)	42	71

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	454	ILE
1	B	638	GLN
1	D	567	ASP
1	B	516	VAL
1	B	571	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	248	ASN
1	D	277	HIS

### 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](#)

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 for Mogul analysis.

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	TRP	A	701	-	12,16,16	0.68	0	7,22,22	0.95	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	BME	A	702	1	3,3,3	1.02	0	2,2,2	1.50	0
2	TRP	B	701	-	12,16,16	0.67	0	7,22,22	0.96	0
3	BME	B	702	1	3,3,3	0.23	0	2,2,2	0.35	0
2	TRP	C	701	-	12,16,16	0.69	0	7,22,22	0.94	0
2	TRP	D	701	-	12,16,16	0.67	0	7,22,22	1.03	0
3	BME	D	702	1	3,3,3	0.20	0	2,2,2	1.56	1 (50%)

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	TRP	A	701	-	-	0/3/8/8	0/2/2/2
3	BME	A	702	1	-	0/1/1/1	0/0/0/0
2	TRP	B	701	-	-	0/3/8/8	0/2/2/2
3	BME	B	702	1	-	0/1/1/1	0/0/0/0
2	TRP	C	701	-	-	0/3/8/8	0/2/2/2
2	TRP	D	701	-	-	0/3/8/8	0/2/2/2
3	BME	D	702	1	-	0/1/1/1	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	702	BME	C1-C2-S2	-2.20	100.58	112.58

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	702	BME	1	0
2	D	701	TRP	1	0
3	D	702	BME	1	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	362/420 (86%)	0.03	14 (3%) 43 35	19, 41, 100, 143	0
1	B	368/420 (87%)	0.14	18 (4%) 33 26	19, 43, 105, 136	0
1	C	348/420 (82%)	0.05	8 (2%) 64 57	21, 53, 101, 122	0
1	D	357/420 (85%)	0.30	32 (8%) 12 8	21, 61, 127, 149	0
All	All	1435/1680 (85%)	0.13	72 (5%) 32 26	19, 49, 110, 149	0

The worst 5 of 72 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	537	GLY	5.4
1	D	539	GLY	5.2
1	A	549	GLY	4.7
1	D	578	TYR	4.6
1	D	582	GLU	4.5

### 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	BME	A	702	4/4	0.90	0.21	1.74	51,62,64,93	0
3	BME	B	702	4/4	0.87	0.24	1.31	50,53,67,83	0
2	TRP	B	701	15/15	0.98	0.20	0.97	17,24,42,45	0
2	TRP	D	701	15/15	0.97	0.17	0.01	29,40,51,57	0
2	TRP	C	701	15/15	0.96	0.16	-0.17	24,36,49,58	0
3	BME	D	702	4/4	0.87	0.18	-0.22	63,70,72,89	0
2	TRP	A	701	15/15	0.98	0.16	-0.52	15,22,36,37	0
4	K	C	702	1/1	0.92	0.19	-	61,61,61,61	0
4	K	D	703	1/1	0.95	0.12	-	63,63,63,63	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.