



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 5, 2023 – 12:39 am GMT

PDB ID : 1UMY  
Title : BHMT from rat liver  
Authors : Gonzalez, B.; Pajares, M.A.; Sanz-Aparicio, J.  
Deposited on : 2003-09-02  
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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

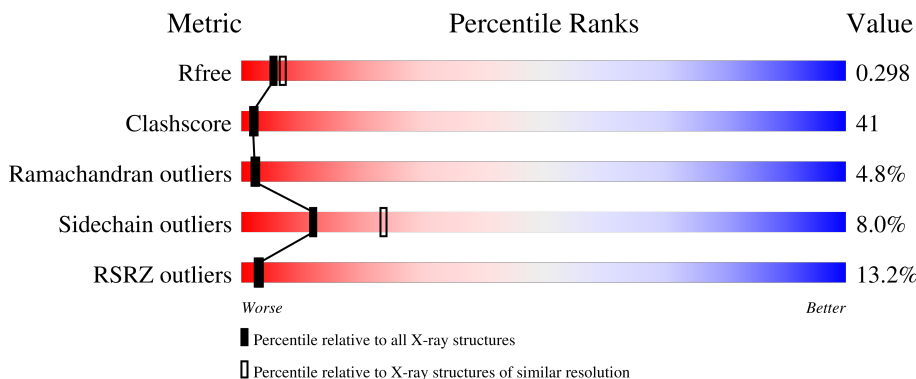
# 1 Overall quality at a glance

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))
$R_{free}$	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (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	407	
1	B	407	
1	C	407	
1	D	407	

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	BME	A	1400	-	X	-	-
3	BME	B	1400	-	X	-	-

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 12353 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 Betaine--homocysteine S-methyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	375	2908	1847	512	533	16	0	0	1
1	B	381	2956	1876	519	545	16	0	0	1
1	C	381	2964	1884	522	542	16	0	0	1
1	D	382	2975	1893	524	542	16	0	0	1

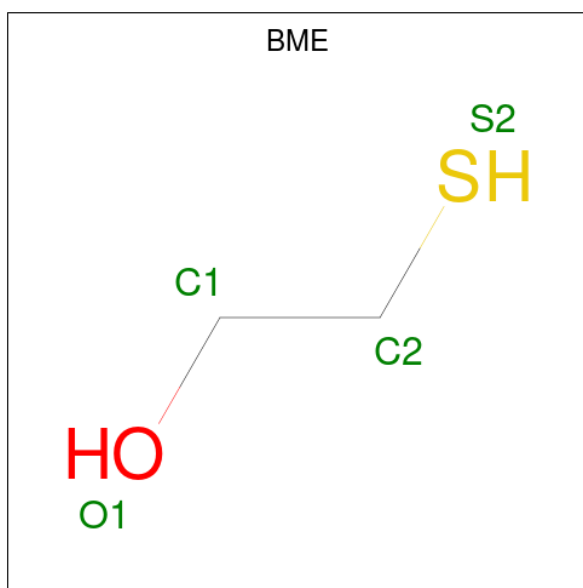
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	247	GLN	HIS	conflict	UNP O09171
A	248	PRO	ALA	conflict	UNP O09171
B	247	GLN	HIS	conflict	UNP O09171
B	248	PRO	ALA	conflict	UNP O09171
C	247	GLN	HIS	conflict	UNP O09171
C	248	PRO	ALA	conflict	UNP O09171
D	247	GLN	HIS	conflict	UNP O09171
D	248	PRO	ALA	conflict	UNP O09171

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Zn 1	0	0
2	B	1	Total 1	Zn 1	0	0
2	C	1	Total 1	Zn 1	0	0
2	D	1	Total 1	Zn 1	0	0

- 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	0	0
			4	2	1	1		
3	B	1	Total	C	O	S	0	0
			4	2	1	1		

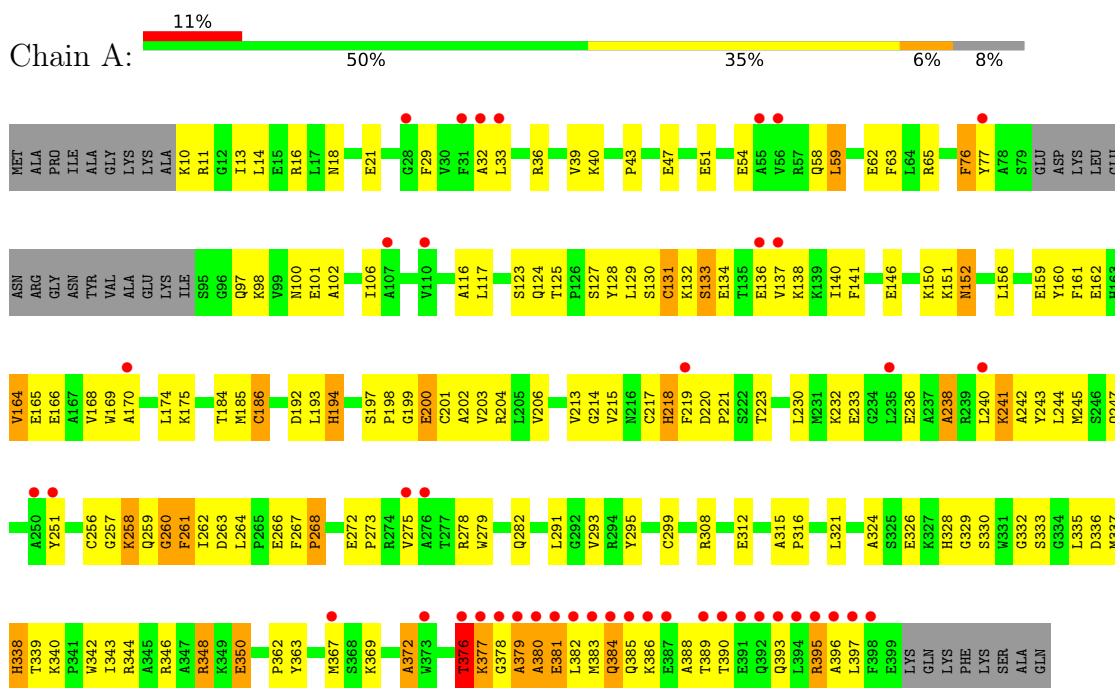
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	140	Total	O	0	0
			140	140		
4	B	149	Total	O	0	0
			149	149		
4	C	107	Total	O	0	0
			107	107		
4	D	142	Total	O	0	0
			142	142		

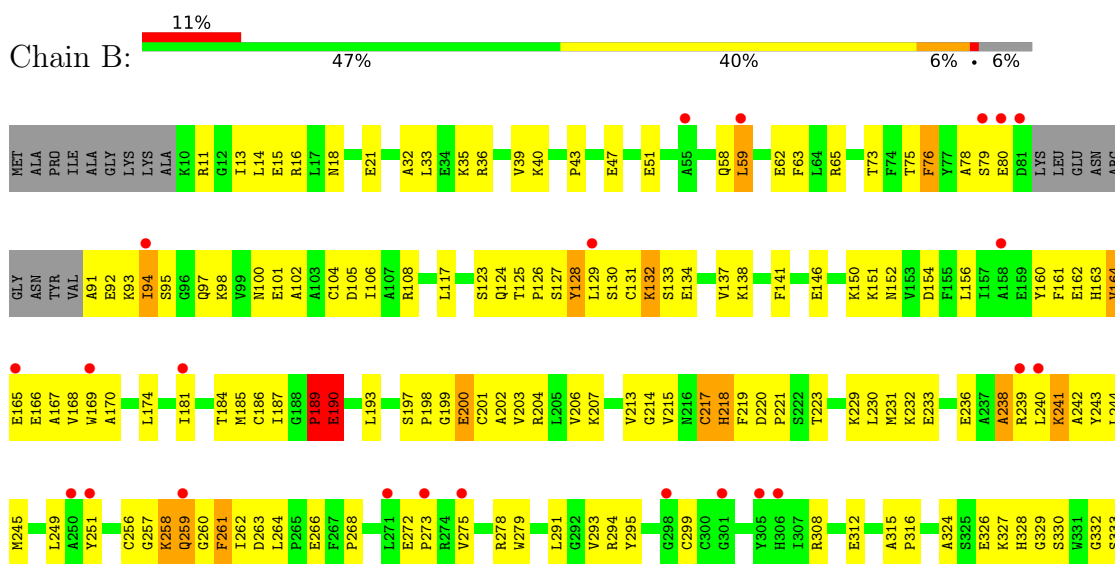
### 3 Residue-property plots [i](#)

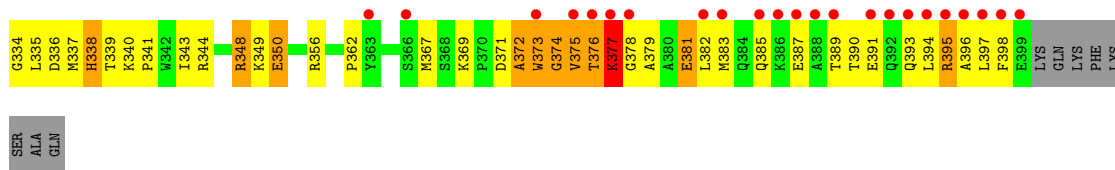
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Betaine-homocysteine S-methyltransferase 1

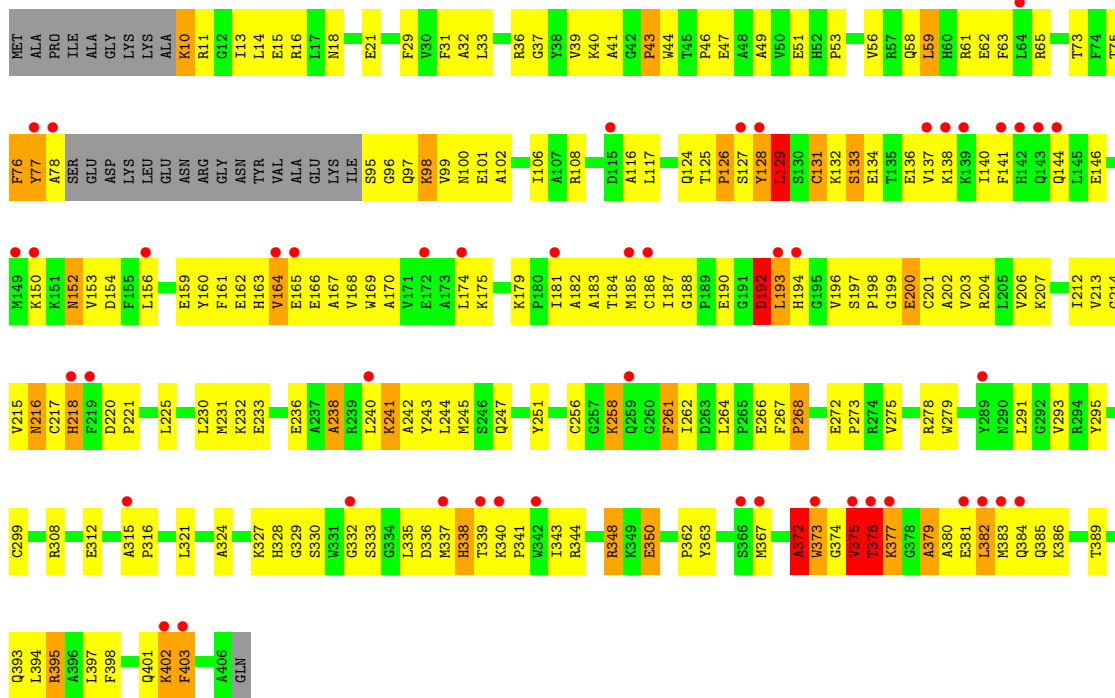


- Molecule 1: Betaine-homocysteine S-methyltransferase 1

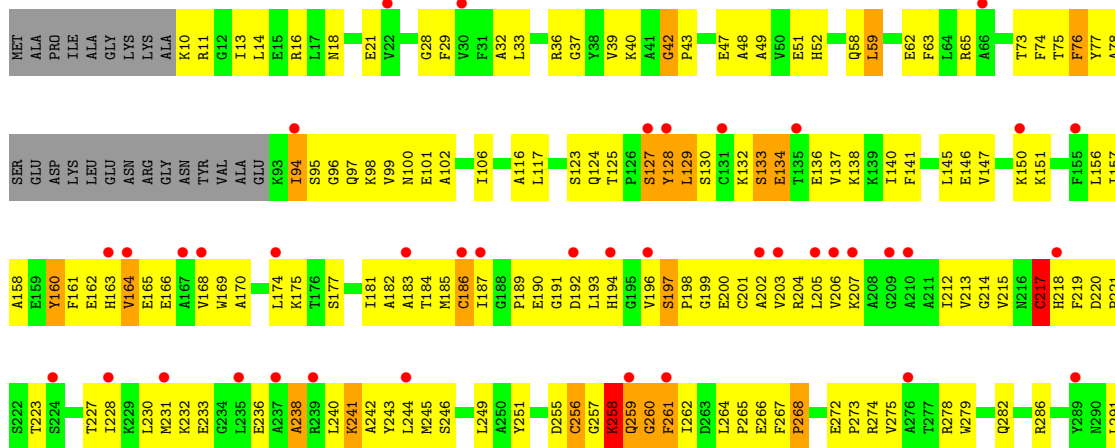


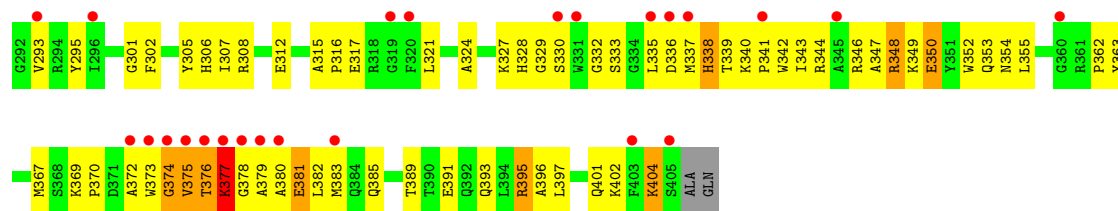


• Molecule 1: Betaine--homocysteine S-methyltransferase 1



• Molecule 1: Betaine--homocysteine S-methyltransferase 1







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.83Å 149.27Å 96.25Å 90.00° 92.92° 90.00°	Depositor
Resolution (Å)	48.40 – 2.50 48.43 – 2.44	Depositor EDS
% Data completeness (in resolution range)	97.1 (48.40-2.50) 92.1 (48.43-2.44)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.04 (at 2.45Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.253 , 0.291 0.263 , 0.298	Depositor DCC
$R_{free}$ test set	5643 reflections (10.14%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.5	Xtrriage
Anisotropy	0.563	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 56.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.042 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	12353	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.62% 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: ZN, 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	0.44	0/2979	0.64	0/4029
1	B	0.48	0/3027	0.65	1/4093 (0.0%)
1	C	0.52	2/3036 (0.1%)	0.67	3/4102 (0.1%)
1	D	0.50	0/3047	0.68	0/4116
All	All	0.49	2/12089 (0.0%)	0.66	4/16340 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	379	ALA	N-CA	6.12	1.58	1.46
1	C	375	VAL	C-N	-5.34	1.21	1.34

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	375	VAL	CB-CA-C	-6.38	99.28	111.40
1	C	372	ALA	CB-CA-C	-5.78	101.44	110.10
1	B	374	GLY	N-CA-C	-5.77	98.67	113.10
1	C	376	THR	O-C-N	5.31	131.20	122.70

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	2908	0	2862	208	0
1	B	2956	0	2906	232	0
1	C	2964	0	2925	276	0
1	D	2975	0	2942	314	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	4	0	5	0	0
3	B	4	0	5	0	0
4	A	140	0	0	19	0
4	B	149	0	0	25	0
4	C	107	0	0	32	0
4	D	142	0	0	67	0
All	All	12353	0	11645	973	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 41.

The worst 5 of 973 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:258:LYS:HG2	1:D:259:GLN:H	0.93	1.09
1:C:126:PRO:HA	1:C:129:LEU:HB2	1.39	1.03
1:D:258:LYS:HG2	1:D:259:GLN:N	1.74	1.02
1:A:377:LYS:HG3	1:A:378:GLY:H	1.23	1.02
1:B:11:ARG:HH11	1:B:11:ARG:HB2	1.26	1.00

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	371/407 (91%)	320 (86%)	34 (9%)	17 (5%)	2	2
1	B	377/407 (93%)	322 (85%)	38 (10%)	17 (4%)	2	3
1	C	377/407 (93%)	320 (85%)	41 (11%)	16 (4%)	3	3
1	D	378/407 (93%)	322 (85%)	34 (9%)	22 (6%)	1	1
All	All	1503/1628 (92%)	1284 (85%)	147 (10%)	72 (5%)	2	2

5 of 72 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	152	ASN
1	A	260	GLY
1	A	338	HIS
1	A	376	THR
1	A	377	LYS

### 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	297/323 (92%)	277 (93%)	20 (7%)	16	31
1	B	302/323 (94%)	277 (92%)	25 (8%)	11	22
1	C	303/323 (94%)	276 (91%)	27 (9%)	9	19
1	D	304/323 (94%)	279 (92%)	25 (8%)	11	22
All	All	1206/1292 (93%)	1109 (92%)	97 (8%)	12	23

5 of 97 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	200	GLU
1	C	402	LYS
1	C	233	GLU
1	C	295	TYR
1	D	76	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 54 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	97	GLN
1	C	216	ASN
1	D	247	GLN
1	C	100	ASN
1	C	144	GLN

### 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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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
3	BME	A	1400	-	3,3,3	2.08	1 (33%)	1,2,2	2.50	1 (100%)
3	BME	B	1400	-	3,3,3	1.98	1 (33%)	1,2,2	2.68	1 (100%)

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
3	BME	A	1400	-	-	1/1/1/1	-
3	BME	B	1400	-	-	1/1/1/1	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1400	BME	O1-C1	-3.59	1.23	1.42
3	B	1400	BME	O1-C1	-3.38	1.24	1.42

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1400	BME	O1-C1-C2	2.68	121.40	110.83
3	A	1400	BME	O1-C1-C2	2.50	120.72	110.83

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1400	BME	O1-C1-C2-S2
3	B	1400	BME	O1-C1-C2-S2

There are no ring outliers.

No monomer is involved in short contacts.

## 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	375/407 (92%)	0.91	43 (11%) 4 4	31, 49, 65, 74	0
1	B	381/407 (93%)	0.89	46 (12%) 4 4	30, 49, 63, 70	24 (6%)
1	C	381/407 (93%)	0.75	48 (12%) 3 3	32, 52, 68, 73	0
1	D	382/407 (93%)	0.98	64 (16%) 1 1	32, 53, 69, 73	0
All	All	1519/1628 (93%)	0.88	201 (13%) 3 3	30, 51, 67, 74	24 (1%)

The worst 5 of 201 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	396	ALA	11.2
1	B	388	ALA	9.4
1	B	389	THR	8.7
1	B	376	THR	7.3
1	A	395	ARG	6.7

### 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
3	BME	B	1400	4/4	0.85	0.26	56,57,57,57	0
2	ZN	C	1406	1/1	0.87	0.15	154,154,154,154	0
3	BME	A	1400	4/4	0.92	0.17	59,62,62,68	0
2	ZN	B	1399	1/1	0.96	0.06	67,67,67,67	0
2	ZN	A	1399	1/1	0.97	0.11	72,72,72,72	0
2	ZN	D	1405	1/1	0.97	0.04	71,71,71,71	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.