



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 12:08 PM GMT

PDB ID : 1M4T  
Title : Biosynthetic thiolase, Cys89 butyrylated  
Authors : Kursula, P.; Ojala, J.; Lambeir, A.-M.; Wierenga, R.K.  
Deposited on : 2002-07-03  
Resolution : 1.77 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

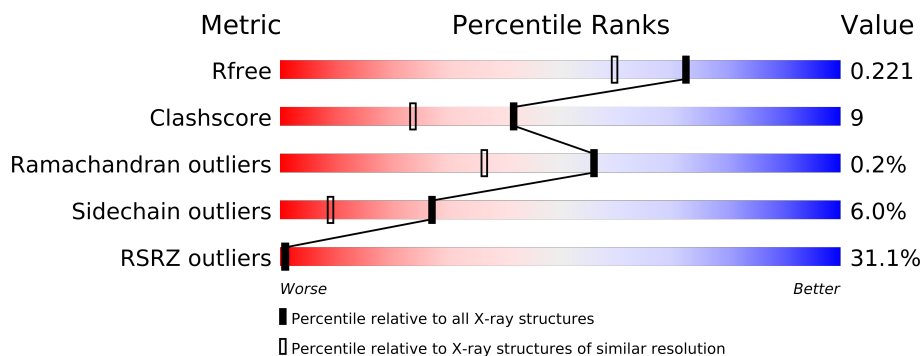
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.77 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))
$R_{free}$	66092	4987 (1.80-1.76)
Clashscore	79885	6152 (1.80-1.76)
Ramachandran outliers	78287	6074 (1.80-1.76)
Sidechain outliers	78261	6073 (1.80-1.76)
RSRZ outliers	66119	4990 (1.80-1.76)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	392	
1	B	392	
1	C	392	
1	D	392	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	SO4	A	9722	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 12421 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Acetyl-CoA acetyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	390	Total	C	N	O	S	0	1	0
			2827	1756	510	540	21			
1	B	390	Total	C	N	O	S	0	1	0
			2827	1756	510	540	21			
1	C	390	Total	C	N	O	S	0	1	0
			2822	1752	510	539	21			
1	D	390	Total	C	N	O	S	0	1	0
			2822	1752	510	539	21			

There are 12 discrepancies between the modelled and reference sequences:

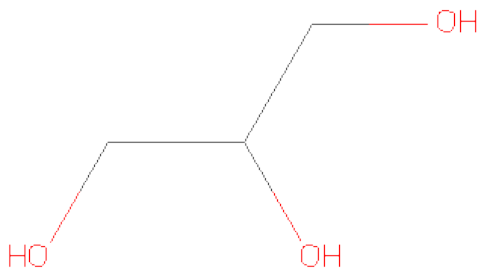
Chain	Residue	Modelled	Actual	Comment	Reference
A	10	ALA	-	INSERTION	UNP P07097
A	89	CY4	CYS	MODIFIED RESIDUE	UNP P07097
A	129	ARG	ALA	CONFLICT	UNP P07097
B	10	ALA	-	INSERTION	UNP P07097
B	89	CY4	CYS	MODIFIED RESIDUE	UNP P07097
B	129	ARG	ALA	CONFLICT	UNP P07097
C	10	ALA	-	INSERTION	UNP P07097
C	89	CY4	CYS	MODIFIED RESIDUE	UNP P07097
C	129	ARG	ALA	CONFLICT	UNP P07097
D	10	ALA	-	INSERTION	UNP P07097
D	89	CY4	CYS	MODIFIED RESIDUE	UNP P07097
D	129	ARG	ALA	CONFLICT	UNP P07097

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 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
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is water.

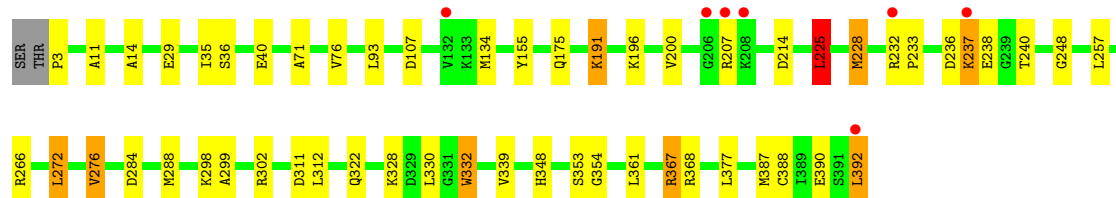
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	403	Total	O	0	0
			403	403		
4	B	406	Total	O	0	0
			406	406		
4	C	144	Total	O	0	0
			144	144		
4	D	138	Total	O	0	0
			138	138		

### 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 errors 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.

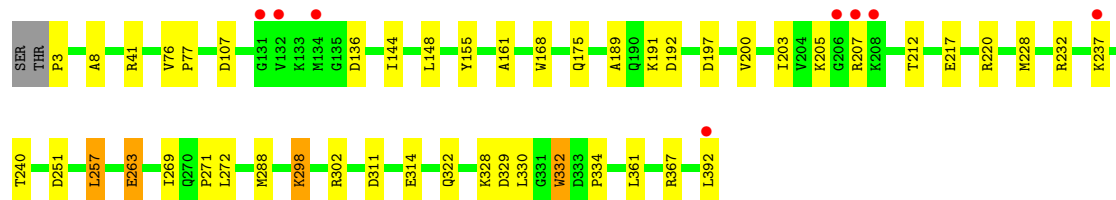
#### • Molecule 1: Acetyl-CoA acetyltransferase

Chain A: 



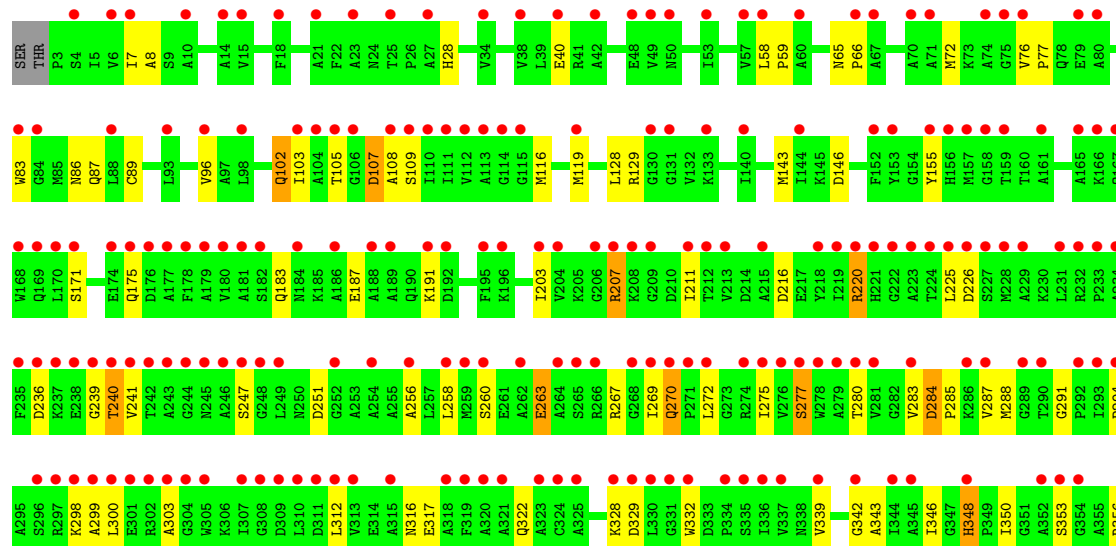
#### • Molecule 1: Acetyl-CoA acetyltransferase

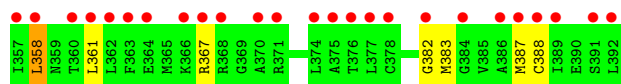
Chain B: 



#### • Molecule 1: Acetyl-CoA acetyltransferase

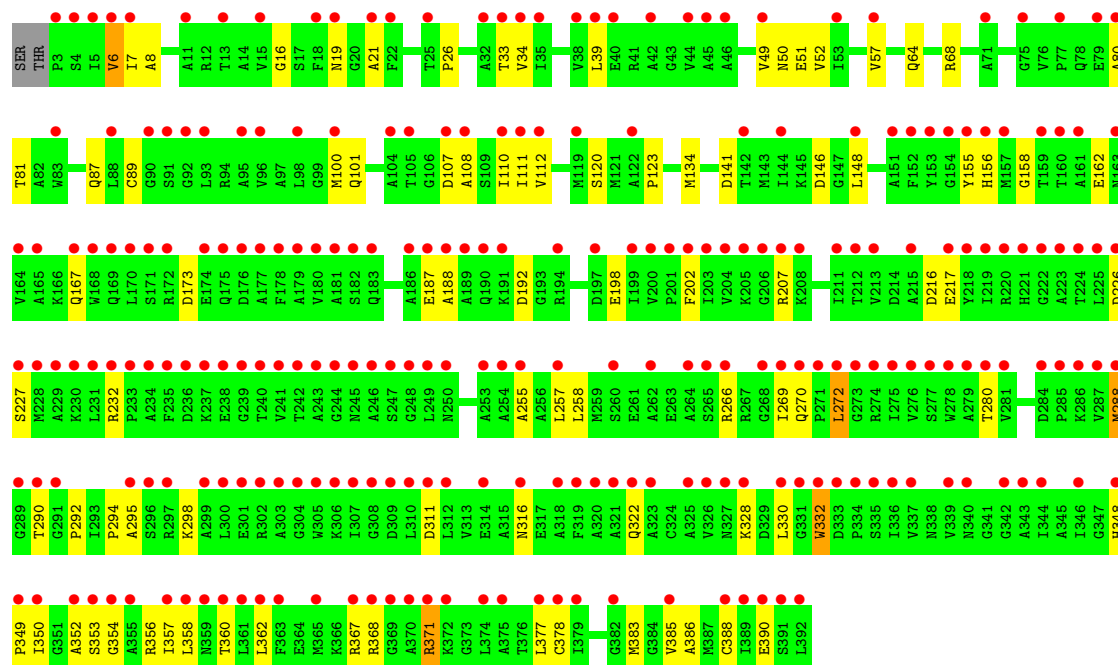
Chain C: 





● Molecule 1: Acetyl-CoA acetyltransferase

Chain D:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	84.35Å 79.31Å 147.32Å 90.00° 93.97° 90.00°	Depositor
Resolution (Å)	20.00 – 1.77 37.17 – 1.77	Depositor EDS
% Data completeness (in resolution range)	99.2 (20.00-1.77) 80.8 (37.17-1.77)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.16 (at 1.77Å)	Xtriage
Refinement program	REFMAC 5.1.19	Depositor
R, $R_{free}$	0.206 , 0.242 0.207 , 0.221	Depositor DCC
$R_{free}$ test set	8648 reflections (4.85%)	DCC
Wilson B-factor (Å <sup>2</sup> )	17.3	Xtriage
Anisotropy	0.574	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 56.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.38$ , $\langle L^2 \rangle = 0.21$	Xtriage
Outliers	0 of 187197 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	12421	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.69% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.



## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, SO4, CY4

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	1.02	1/2861 (0.0%)	1.02	13/3861 (0.3%)
1	B	1.00	3/2861 (0.1%)	0.99	9/3861 (0.2%)
1	C	0.57	3/2861 (0.1%)	0.77	7/3861 (0.2%)
1	D	0.52	2/2861 (0.1%)	0.77	7/3861 (0.2%)
All	All	0.81	9/11444 (0.1%)	0.90	36/15444 (0.2%)

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	277	SER	CB-OG	11.53	1.57	1.42
1	C	109	SER	C-O	6.63	1.35	1.23
1	D	272	LEU	C-O	6.45	1.35	1.23
1	C	260	SER	CB-OG	6.26	1.50	1.42
1	B	189	ALA	CA-CB	5.49	1.64	1.52

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	107	ASP	CB-CG-OD2	10.02	127.31	118.30
1	A	284	ASP	CB-CG-OD1	8.02	125.52	118.30
1	A	266	ARG	NE-CZ-NH2	-7.81	116.40	120.30
1	A	107	ASP	CB-CG-OD2	7.15	124.73	118.30
1	A	302	ARG	NE-CZ-NH2	-6.96	116.82	120.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2827	0	2833	40	0
1	B	2827	0	2833	33	0
1	C	2822	0	2826	66	0
1	D	2822	0	2826	70	0
2	A	10	0	0	0	0
2	B	10	0	0	0	0
3	A	6	0	7	1	0
3	B	6	0	8	1	0
4	A	403	0	0	23	2
4	B	406	0	0	16	1
4	C	144	0	0	43	0
4	D	138	0	0	42	0
All	All	12421	0	11333	207	2

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 9.

The worst 5 of 207 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:387:MET:SD	4:A:9864:HOH:O	1.90	1.23
1:D:378:CYS:SG	4:D:499:HOH:O	2.02	1.17
4:C:491:HOH:O	1:D:383:MET:SD	2.07	1.12
1:D:388:CYS:SG	4:D:514:HOH:O	2.09	1.08
1:D:377:LEU:HD11	4:D:471:HOH:O	1.60	1.01

All (2) 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	Distance(Å)	Clash(Å)
4:A:400:HOH:O	4:A:407:HOH:O[2_655]	2.03	0.17
4:A:417:HOH:O	4:B:440:HOH:O[2_645]	2.19	0.01

## 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	388/392 (99%)	376 (97%)	12 (3%)	0	100	100
1	B	388/392 (99%)	373 (96%)	15 (4%)	0	100	100
1	C	388/392 (99%)	372 (96%)	14 (4%)	2 (0%)	38	18
1	D	388/392 (99%)	367 (95%)	20 (5%)	1 (0%)	50	30
All	All	1552/1568 (99%)	1488 (96%)	61 (4%)	3 (0%)	56	36

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	240	THR
1	D	266	ARG
1	C	291	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	277/278 (100%)	261 (94%)	16 (6%)	28	9
1	B	277/278 (100%)	264 (95%)	13 (5%)	36	15
1	C	277/278 (100%)	257 (93%)	20 (7%)	21	5
1	D	277/278 (100%)	260 (94%)	17 (6%)	26	8
All	All	1108/1112 (100%)	1042 (94%)	66 (6%)	27	9

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	40	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	263	GLU
1	D	322	GLN
1	C	102	GLN
1	C	207	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	221	HIS
1	C	175	GLN
1	D	175	GLN
1	B	184	ASN
1	D	184	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

4 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	CY4	A	89	1	10,10,11	6.63	3 (30%)	9,11,13	2.81	4 (44%)
1	CY4	B	89	1	10,10,11	6.86	4 (40%)	9,11,13	3.04	4 (44%)
1	CY4	C	89	1	5,5,11	7.53	2 (40%)	2,5,13	0.90	0
1	CY4	D	89	1	5,5,11	8.30	2 (40%)	2,5,13	1.11	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	CY4	A	89	1	-	0/8/10/12	0/0/0/0
1	CY4	B	89	1	-	0/8/10/12	0/0/0/0
1	CY4	C	89	1	-	0/2/4/12	0/0/0/0
1	CY4	D	89	1	-	0/2/4/12	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	89	CY4	O-C	20.70	1.25	1.11
1	A	89	CY4	O-C	20.47	1.25	1.11
1	D	89	CY4	O-C	18.19	1.23	1.11
1	C	89	CY4	O-C	16.61	1.22	1.11
1	B	89	CY4	CB-SG	3.80	1.86	1.81

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	89	CY4	CB-SG-CA1	-6.72	96.31	100.95
1	B	89	CY4	CB-SG-CA1	-6.52	96.45	100.95
1	B	89	CY4	OA1-CA1-CA2	-4.17	120.30	123.95
1	B	89	CY4	OA1-CA1-SG	3.35	126.39	122.85
1	A	89	CY4	OA1-CA1-CA2	-3.19	121.15	123.95

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

6 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	$\# Z  > 2$
3	GOL	A	5394	-	5,5,5	1.04	1 (20%)	5,5,5	0.65	0
2	SO4	A	9720	-	4,4,4	0.46	0	6,6,6	0.31	0
2	SO4	A	9722	-	4,4,4	0.20	0	6,6,6	0.16	0
3	GOL	B	6394	-	5,5,5	0.92	1 (20%)	5,5,5	0.68	0
2	SO4	B	9719	-	4,4,4	0.21	0	6,6,6	0.17	0
2	SO4	B	9721	-	4,4,4	0.50	0	6,6,6	0.60	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
3	GOL	A	5394	-	-	0/4/4/4	0/0/0/0
2	SO4	A	9720	-	-	0/0/0/0	0/0/0/0
2	SO4	A	9722	-	-	0/0/0/0	0/0/0/0
3	GOL	B	6394	-	-	0/4/4/4	0/0/0/0
2	SO4	B	9719	-	-	0/0/0/0	0/0/0/0
2	SO4	B	9721	-	-	0/0/0/0	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	5394	GOL	O2-C2	-2.16	1.36	1.43
3	B	6394	GOL	O2-C2	-2.03	1.37	1.43

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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 ⓘ

### 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	390/392 (99%)	0.38	7 (1%) 65 65	5, 11, 28, 54	0
1	B	390/392 (99%)	0.38	8 (2%) 60 59	6, 11, 28, 55	0
1	C	390/392 (99%)	2.83	226 (57%) 0 0	2, 11, 28, 42	0
1	D	390/392 (99%)	3.50	246 (63%) 0 0	2, 13, 29, 45	0
All	All	1560/1568 (99%)	1.77	487 (31%) 1 1	2, 11, 29, 55	0

The worst 5 of 487 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	231	LEU	17.2
1	D	219	ILE	14.8
1	D	229	ALA	14.5
1	D	179	ALA	14.4
1	D	244	GLY	14.1

### 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. 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	CY4	A	89	11/12	0.13	1.46	9,12,30,31	0
1	CY4	B	89	11/12	0.12	1.14	8,10,26,28	0
1	CY4	D	89	6/12	0.14	-1.85	2,4,4,6	0
1	CY4	C	89	6/12	0.11	-2.55	8,10,11,15	0



### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	SO4	A	9722	5/5	0.22	2.60	74,74,75,75	0
3	GOL	B	6394	6/6	0.15	1.70	18,26,30,38	0
2	SO4	A	9720	5/5	0.12	1.11	45,50,50,51	0
2	SO4	B	9719	5/5	0.19	0.94	76,77,78,78	0
3	GOL	A	5394	6/6	0.12	0.34	20,31,33,33	0
2	SO4	B	9721	5/5	0.12	-0.93	47,48,50,51	0

### 6.5 Other polymers ⓘ

There are no such residues in this entry.