



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 05:21 am BST

PDB ID : 4U2H  
Title : The crystal structure of apo CalE6, a methionine gamma lyase from *Micromonospora echinospora*  
Authors : Song, H.G.; Xu, R.  
Deposited on : 2014-07-17  
Resolution : 1.85 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 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	:	2.11
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)	:	2.11

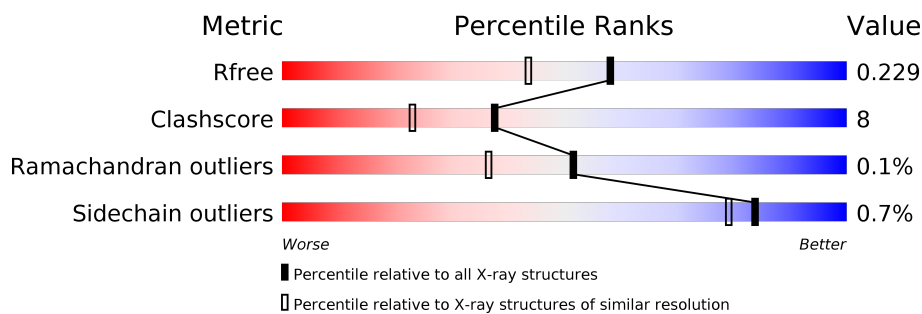
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	389	84% 9% 7%
1	B	389	84% 10% 6%
1	C	389	79% 13% 7%
1	D	389	84% 10% 6%
1	E	389	83% 12% 6%
1	F	389	83% 11% 7%
1	G	389	84% 9% 7%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	389	<div><div></div><div>80%</div><div>13%</div><div>6%</div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 23976 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 CalE6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	362	Total	C	N	O	S	0	3	0
			2660	1673	478	500	9			
1	B	365	Total	C	N	O	S	0	1	0
			2674	1683	479	504	8			
1	C	360	Total	C	N	O	S	0	3	0
			2670	1679	483	500	8			
1	D	366	Total	C	N	O	S	0	0	0
			2672	1681	481	502	8			
1	E	366	Total	C	N	O	S	0	1	0
			2697	1694	486	509	8			
1	F	362	Total	C	N	O	S	0	2	0
			2651	1670	474	498	9			
1	G	361	Total	C	N	O	S	0	1	0
			2656	1670	482	495	9			
1	H	364	Total	C	N	O	S	0	1	0
			2644	1667	472	496	9			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	382	LEU	-	expression tag	UNP Q8KNG3
A	383	GLU	-	expression tag	UNP Q8KNG3
A	384	HIS	-	expression tag	UNP Q8KNG3
A	385	HIS	-	expression tag	UNP Q8KNG3
A	386	HIS	-	expression tag	UNP Q8KNG3
A	387	HIS	-	expression tag	UNP Q8KNG3
A	388	HIS	-	expression tag	UNP Q8KNG3
A	389	HIS	-	expression tag	UNP Q8KNG3
B	382	LEU	-	expression tag	UNP Q8KNG3
B	383	GLU	-	expression tag	UNP Q8KNG3
B	384	HIS	-	expression tag	UNP Q8KNG3
B	385	HIS	-	expression tag	UNP Q8KNG3
B	386	HIS	-	expression tag	UNP Q8KNG3

*Continued on next page...*

*Continued from previous page...*

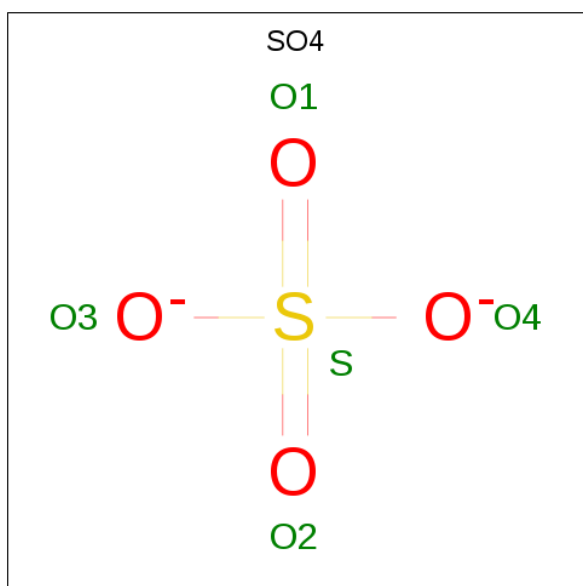
Chain	Residue	Modelled	Actual	Comment	Reference
B	387	HIS	-	expression tag	UNP Q8KNG3
B	388	HIS	-	expression tag	UNP Q8KNG3
B	389	HIS	-	expression tag	UNP Q8KNG3
C	382	LEU	-	expression tag	UNP Q8KNG3
C	383	GLU	-	expression tag	UNP Q8KNG3
C	384	HIS	-	expression tag	UNP Q8KNG3
C	385	HIS	-	expression tag	UNP Q8KNG3
C	386	HIS	-	expression tag	UNP Q8KNG3
C	387	HIS	-	expression tag	UNP Q8KNG3
C	388	HIS	-	expression tag	UNP Q8KNG3
C	389	HIS	-	expression tag	UNP Q8KNG3
D	382	LEU	-	expression tag	UNP Q8KNG3
D	383	GLU	-	expression tag	UNP Q8KNG3
D	384	HIS	-	expression tag	UNP Q8KNG3
D	385	HIS	-	expression tag	UNP Q8KNG3
D	386	HIS	-	expression tag	UNP Q8KNG3
D	387	HIS	-	expression tag	UNP Q8KNG3
D	388	HIS	-	expression tag	UNP Q8KNG3
D	389	HIS	-	expression tag	UNP Q8KNG3
E	382	LEU	-	expression tag	UNP Q8KNG3
E	383	GLU	-	expression tag	UNP Q8KNG3
E	384	HIS	-	expression tag	UNP Q8KNG3
E	385	HIS	-	expression tag	UNP Q8KNG3
E	386	HIS	-	expression tag	UNP Q8KNG3
E	387	HIS	-	expression tag	UNP Q8KNG3
E	388	HIS	-	expression tag	UNP Q8KNG3
E	389	HIS	-	expression tag	UNP Q8KNG3
F	382	LEU	-	expression tag	UNP Q8KNG3
F	383	GLU	-	expression tag	UNP Q8KNG3
F	384	HIS	-	expression tag	UNP Q8KNG3
F	385	HIS	-	expression tag	UNP Q8KNG3
F	386	HIS	-	expression tag	UNP Q8KNG3
F	387	HIS	-	expression tag	UNP Q8KNG3
F	388	HIS	-	expression tag	UNP Q8KNG3
F	389	HIS	-	expression tag	UNP Q8KNG3
G	382	LEU	-	expression tag	UNP Q8KNG3
G	383	GLU	-	expression tag	UNP Q8KNG3
G	384	HIS	-	expression tag	UNP Q8KNG3
G	385	HIS	-	expression tag	UNP Q8KNG3
G	386	HIS	-	expression tag	UNP Q8KNG3
G	387	HIS	-	expression tag	UNP Q8KNG3
G	388	HIS	-	expression tag	UNP Q8KNG3

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	389	HIS	-	expression tag	UNP Q8KNG3
H	382	LEU	-	expression tag	UNP Q8KNG3
H	383	GLU	-	expression tag	UNP Q8KNG3
H	384	HIS	-	expression tag	UNP Q8KNG3
H	385	HIS	-	expression tag	UNP Q8KNG3
H	386	HIS	-	expression tag	UNP Q8KNG3
H	387	HIS	-	expression tag	UNP Q8KNG3
H	388	HIS	-	expression tag	UNP Q8KNG3
H	389	HIS	-	expression tag	UNP Q8KNG3

- 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		
2	C	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	D	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		
2	G	1	Total	O	S	0	0
			5	4	1		
2	G	1	Total	O	S	0	0
			5	4	1		
2	G	1	Total	O	S	0	0
			5	4	1		
2	H	1	Total	O	S	0	0
			5	4	1		

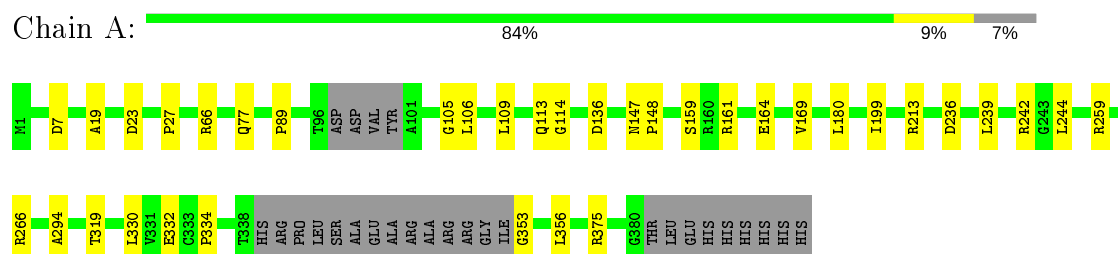
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	345	Total	O	0	0
			345	345		
3	B	331	Total	O	0	0
			331	331		
3	C	329	Total	O	0	0
			329	329		
3	D	282	Total	O	0	0
			282	282		
3	E	329	Total	O	0	0
			329	329		
3	F	300	Total	O	0	0
			300	300		
3	G	337	Total	O	0	0
			337	337		
3	H	319	Total	O	0	0
			319	319		

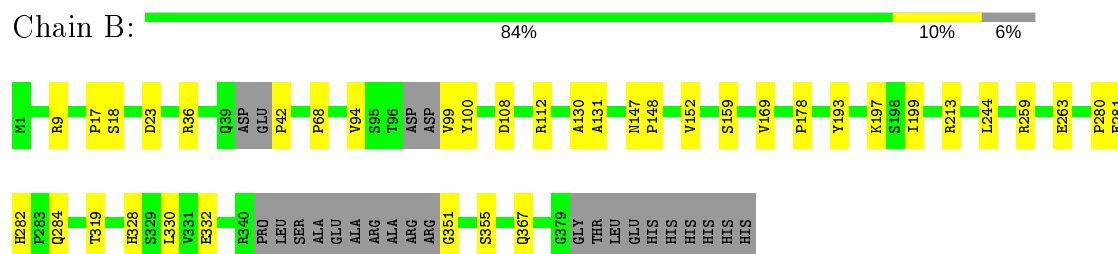
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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. 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. 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.

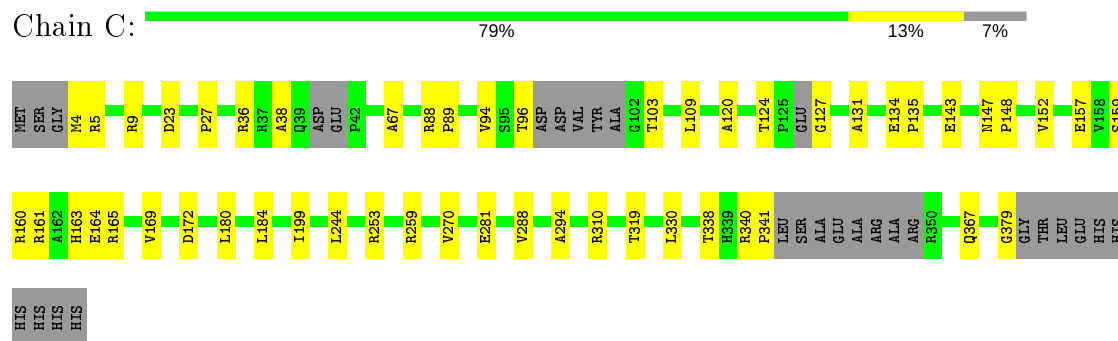
#### • Molecule 1: CalE6



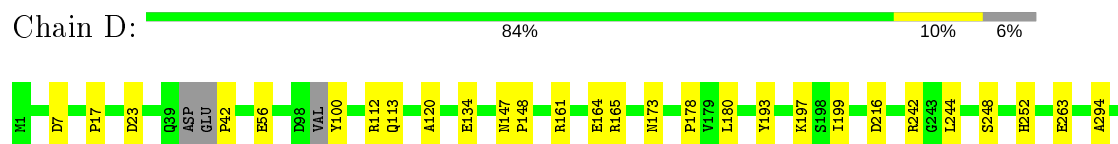
#### • Molecule 1: CalE6



#### • Molecule 1: CalE6



#### • Molecule 1: CalE6

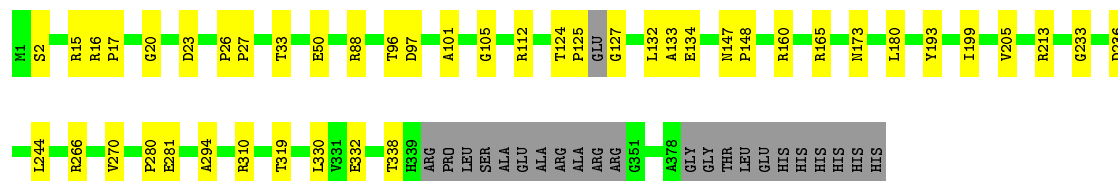






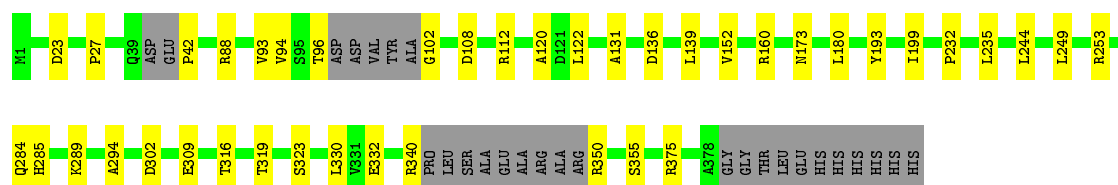
• Molecule 1: CalE6

Chain E: 83% 12% 6%



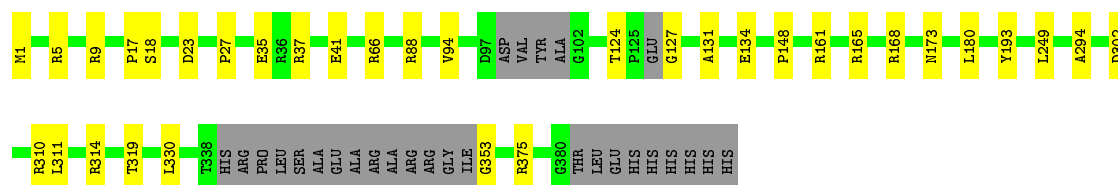
• Molecule 1: CalE6

Chain F: 83% 11% 7%



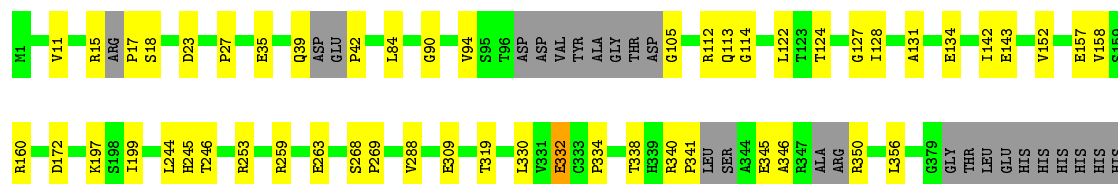
• Molecule 1: CalE6

Chain G: 84% 9% 7%



• Molecule 1: CalE6

Chain H: 80% 13% 6%



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	145.20Å 145.40Å 346.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.30 – 1.85 49.25 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.30-1.85) 97.1 (49.25-1.85)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.57 (at 1.86Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.209 , 0.232 0.206 , 0.229	Depositor DCC
$R_{free}$ test set	15586 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.7	Xtriage
Anisotropy	0.110	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 30.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	0.286 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	23976	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.04% 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: SO4

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	0.39	0/2711	0.53	0/3702
1	B	0.39	0/2725	0.54	0/3721
1	C	0.43	0/2721	0.57	1/3712 (0.0%)
1	D	0.39	0/2724	0.53	0/3722
1	E	0.42	0/2749	0.56	0/3755
1	F	0.40	0/2701	0.55	0/3687
1	G	0.40	0/2706	0.55	0/3692
1	H	0.49	0/2693	0.57	0/3676
All	All	0.41	0/21730	0.55	1/29667 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	4	MET	CA-CB-CG	5.54	122.72	113.30

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	2660	0	2624	38	0
1	B	2674	0	2635	33	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2670	0	2640	49	0
1	D	2672	0	2627	39	0
1	E	2697	0	2660	43	0
1	F	2651	0	2600	37	0
1	G	2656	0	2631	43	0
1	H	2644	0	2589	51	0
2	A	10	0	0	0	0
2	B	10	0	0	0	0
2	C	10	0	0	0	0
2	D	10	0	0	1	0
2	E	10	0	0	0	0
2	F	10	0	0	0	0
2	G	15	0	0	0	0
2	H	5	0	0	0	0
3	A	345	0	0	29	1
3	B	331	0	0	25	1
3	C	329	0	0	34	1
3	D	282	0	0	30	1
3	E	329	0	0	27	1
3	F	300	0	0	23	1
3	G	337	0	0	34	1
3	H	319	0	0	32	2
All	All	23976	0	21006	328	5

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 8.

The worst 5 of 328 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:E:160:ARG:HD2	3:E:805:HOH:O	1.25	1.29
1:F:122:LEU:HG	3:F:797:HOH:O	1.33	1.27
1:B:281:GLU:CA	3:B:782:HOH:O	1.90	1.17
1:A:114:GLY:C	3:A:834:HOH:O	1.81	1.16
1:B:281:GLU:C	3:B:782:HOH:O	1.83	1.13

All (5) 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	Interatomic distance (Å)	Clash overlap (Å)
3:C:522:HOH:O	3:H:517:HOH:O[6_445]	1.82	0.38
3:H:776:HOH:O	3:H:776:HOH:O[2_455]	1.83	0.37
3:B:504:HOH:O	3:F:507:HOH:O[8_444]	1.90	0.30
3:A:523:HOH:O	3:E:509:HOH:O[8_544]	2.03	0.17
3:D:515:HOH:O	3:G:527:HOH:O[7_454]	2.10	0.10

## 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	359/389 (92%)	350 (98%)	9 (2%)	0	100	100
1	B	358/389 (92%)	349 (98%)	9 (2%)	0	100	100
1	C	353/389 (91%)	346 (98%)	7 (2%)	0	100	100
1	D	358/389 (92%)	352 (98%)	5 (1%)	1 (0%)	41	26
1	E	361/389 (93%)	352 (98%)	9 (2%)	0	100	100
1	F	356/389 (92%)	349 (98%)	7 (2%)	0	100	100
1	G	354/389 (91%)	346 (98%)	8 (2%)	0	100	100
1	H	353/389 (91%)	345 (98%)	6 (2%)	2 (1%)	25	12
All	All	2852/3112 (92%)	2789 (98%)	60 (2%)	3 (0%)	51	36

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	197[A]	LYS
1	H	197[B]	LYS
1	D	197	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	270/305 (88%)	268 (99%)	2 (1%)	84	79
1	B	272/305 (89%)	268 (98%)	4 (2%)	65	53
1	C	274/305 (90%)	273 (100%)	1 (0%)	91	89
1	D	271/305 (89%)	269 (99%)	2 (1%)	84	79
1	E	276/305 (90%)	275 (100%)	1 (0%)	91	89
1	F	267/305 (88%)	265 (99%)	2 (1%)	84	79
1	G	270/305 (88%)	269 (100%)	1 (0%)	91	89
1	H	265/305 (87%)	262 (99%)	3 (1%)	73	65
All	All	2165/2440 (89%)	2149 (99%)	16 (1%)	84	79

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	23	ASP
1	D	332	GLU
1	G	23	ASP
1	C	23	ASP
1	H	18	SER

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 ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

16 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	SO4	G	403	-	4,4,4	0.21	0	6,6,6	0.17	0
2	SO4	E	402	-	4,4,4	0.21	0	6,6,6	0.13	0
2	SO4	C	402	-	4,4,4	0.26	0	6,6,6	0.16	0
2	SO4	G	402	-	4,4,4	0.21	0	6,6,6	0.32	0
2	SO4	A	401	-	4,4,4	0.17	0	6,6,6	0.12	0
2	SO4	F	401	-	4,4,4	0.25	0	6,6,6	0.16	0
2	SO4	C	401	-	4,4,4	0.17	0	6,6,6	0.11	0
2	SO4	E	401	-	4,4,4	0.13	0	6,6,6	0.10	0
2	SO4	A	402	-	4,4,4	0.22	0	6,6,6	0.18	0
2	SO4	G	401	-	4,4,4	0.15	0	6,6,6	0.09	0
2	SO4	D	401	-	4,4,4	0.28	0	6,6,6	0.11	0
2	SO4	B	402	-	4,4,4	0.16	0	6,6,6	0.08	0
2	SO4	H	401	-	4,4,4	0.16	0	6,6,6	0.12	0
2	SO4	F	402	-	4,4,4	0.13	0	6,6,6	0.08	0
2	SO4	B	401	-	4,4,4	0.14	0	6,6,6	0.22	0
2	SO4	D	402	-	4,4,4	0.16	0	6,6,6	0.13	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	402	SO4	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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.