



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 10, 2017 – 04:02 PM EST

PDB ID : 5LNQ  
Title : Crystal structure of SCP2 thiolase from *Leishmania mexicana*. Complex of the C123A mutant with acetoacetyl-CoA.  
Authors : Harijan, R.K.; Kiema, T.-R.; Wierenga, R.K.  
Deposited on : 2016-08-06  
Resolution : 1.98 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20028442  
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) : rb-20028442

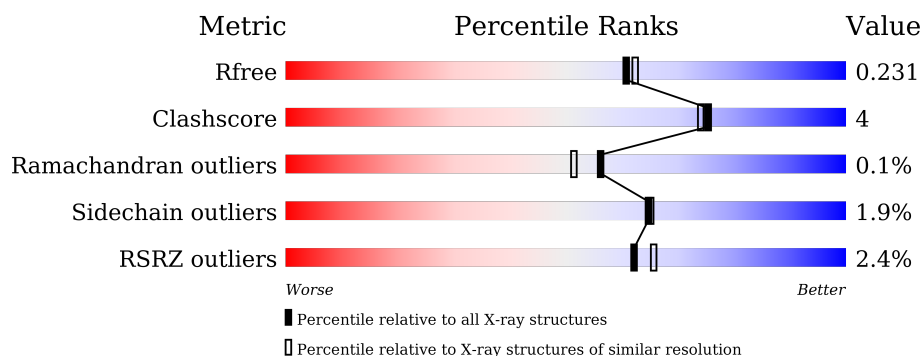
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.98 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	8664 (2.00-1.96)
Clashscore	102246	9905 (2.00-1.96)
Ramachandran outliers	100387	9792 (2.00-1.96)
Sidechain outliers	100360	9791 (2.00-1.96)
RSRZ outliers	91569	8679 (2.00-1.96)

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	457	<div> <div>3%</div> <div>87% 7% • 6%</div> </div>
1	B	457	<div> <div>87% 7% • 6%</div> </div>
1	C	457	<div> <div>3%</div> <div>87% 6% • 6%</div> </div>
1	D	457	<div> <div>3%</div> <div>85% 8% 6%</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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	CAA	A	501[A]	-	-	-	X
2	CAA	A	501[B]	-	-	-	X
2	CAA	D	501[A]	-	-	-	X
2	CAA	D	501[B]	-	-	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 14016 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 3-ketoacyl-CoA thiolase-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	430	Total	C	N	O	S	0	1	0
			3229	2014	566	623	26			
1	B	430	Total	C	N	O	S	0	1	0
			3229	2014	566	623	26			
1	C	430	Total	C	N	O	S	0	1	0
			3229	2014	566	623	26			
1	D	430	Total	C	N	O	S	0	1	0
			3229	2014	566	623	26			

There are 180 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-15	HIS	-	expression tag	UNP E9AW84
A	-14	HIS	-	expression tag	UNP E9AW84
A	-13	HIS	-	expression tag	UNP E9AW84
A	-12	HIS	-	expression tag	UNP E9AW84
A	-11	HIS	-	expression tag	UNP E9AW84
A	-10	HIS	-	expression tag	UNP E9AW84
A	-9	SER	-	expression tag	UNP E9AW84
A	-8	SER	-	expression tag	UNP E9AW84
A	-7	GLY	-	expression tag	UNP E9AW84
A	-6	LEU	-	expression tag	UNP E9AW84
A	-5	VAL	-	expression tag	UNP E9AW84
A	-4	PRO	-	expression tag	UNP E9AW84
A	-3	ARG	-	expression tag	UNP E9AW84
A	-2	GLY	-	expression tag	UNP E9AW84
A	-1	SER	-	expression tag	UNP E9AW84
A	0	HIS	-	expression tag	UNP E9AW84
A	6	LEU	MET	conflict	UNP E9AW84
A	12	ALA	THR	conflict	UNP E9AW84
A	31	PRO	LYS	conflict	UNP E9AW84
A	34	ILE	VAL	conflict	UNP E9AW84
A	45	LYS	GLN	conflict	UNP E9AW84

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	59	GLY	HIS	conflict	UNP E9AW84
A	61	MET	ILE	conflict	UNP E9AW84
A	63	HIS	ASN	conflict	UNP E9AW84
A	64	SER	VAL	conflict	UNP E9AW84
A	65	GLY	SER	conflict	UNP E9AW84
A	67	SER	LYS	conflict	UNP E9AW84
A	69	ARG	LYS	conflict	UNP E9AW84
A	71	GLY	ASP	conflict	UNP E9AW84
A	107	SER	ALA	conflict	UNP E9AW84
A	112	MET	ILE	conflict	UNP E9AW84
A	123	ALA	CYS	engineered mutation	UNP E9AW84
A	134	MET	THR	conflict	UNP E9AW84
A	139	SER	ALA	conflict	UNP E9AW84
A	145	THR	VAL	conflict	UNP E9AW84
A	157	SER	THR	conflict	UNP E9AW84
A	176	GLN	LYS	conflict	UNP E9AW84
A	229	THR	SER	conflict	UNP E9AW84
A	268	VAL	ALA	conflict	UNP E9AW84
A	277	ARG	LYS	conflict	UNP E9AW84
A	327	ILE	VAL	conflict	UNP E9AW84
A	356	GLU	ASP	conflict	UNP E9AW84
A	368	ASP	GLU	conflict	UNP E9AW84
A	398	ILE	VAL	conflict	UNP E9AW84
A	410	GLU	GLY	conflict	UNP E9AW84
B	-15	HIS	-	expression tag	UNP E9AW84
B	-14	HIS	-	expression tag	UNP E9AW84
B	-13	HIS	-	expression tag	UNP E9AW84
B	-12	HIS	-	expression tag	UNP E9AW84
B	-11	HIS	-	expression tag	UNP E9AW84
B	-10	HIS	-	expression tag	UNP E9AW84
B	-9	SER	-	expression tag	UNP E9AW84
B	-8	SER	-	expression tag	UNP E9AW84
B	-7	GLY	-	expression tag	UNP E9AW84
B	-6	LEU	-	expression tag	UNP E9AW84
B	-5	VAL	-	expression tag	UNP E9AW84
B	-4	PRO	-	expression tag	UNP E9AW84
B	-3	ARG	-	expression tag	UNP E9AW84
B	-2	GLY	-	expression tag	UNP E9AW84
B	-1	SER	-	expression tag	UNP E9AW84
B	0	HIS	-	expression tag	UNP E9AW84
B	6	LEU	MET	conflict	UNP E9AW84
B	12	ALA	THR	conflict	UNP E9AW84

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	31	PRO	LYS	conflict	UNP E9AW84
B	34	ILE	VAL	conflict	UNP E9AW84
B	45	LYS	GLN	conflict	UNP E9AW84
B	59	GLY	HIS	conflict	UNP E9AW84
B	61	MET	ILE	conflict	UNP E9AW84
B	63	HIS	ASN	conflict	UNP E9AW84
B	64	SER	VAL	conflict	UNP E9AW84
B	65	GLY	SER	conflict	UNP E9AW84
B	67	SER	LYS	conflict	UNP E9AW84
B	69	ARG	LYS	conflict	UNP E9AW84
B	71	GLY	ASP	conflict	UNP E9AW84
B	107	SER	ALA	conflict	UNP E9AW84
B	112	MET	ILE	conflict	UNP E9AW84
B	123	ALA	CYS	engineered mutation	UNP E9AW84
B	134	MET	THR	conflict	UNP E9AW84
B	139	SER	ALA	conflict	UNP E9AW84
B	145	THR	VAL	conflict	UNP E9AW84
B	157	SER	THR	conflict	UNP E9AW84
B	176	GLN	LYS	conflict	UNP E9AW84
B	229	THR	SER	conflict	UNP E9AW84
B	268	VAL	ALA	conflict	UNP E9AW84
B	277	ARG	LYS	conflict	UNP E9AW84
B	327	ILE	VAL	conflict	UNP E9AW84
B	356	GLU	ASP	conflict	UNP E9AW84
B	368	ASP	GLU	conflict	UNP E9AW84
B	398	ILE	VAL	conflict	UNP E9AW84
B	410	GLU	GLY	conflict	UNP E9AW84
C	-15	HIS	-	expression tag	UNP E9AW84
C	-14	HIS	-	expression tag	UNP E9AW84
C	-13	HIS	-	expression tag	UNP E9AW84
C	-12	HIS	-	expression tag	UNP E9AW84
C	-11	HIS	-	expression tag	UNP E9AW84
C	-10	HIS	-	expression tag	UNP E9AW84
C	-9	SER	-	expression tag	UNP E9AW84
C	-8	SER	-	expression tag	UNP E9AW84
C	-7	GLY	-	expression tag	UNP E9AW84
C	-6	LEU	-	expression tag	UNP E9AW84
C	-5	VAL	-	expression tag	UNP E9AW84
C	-4	PRO	-	expression tag	UNP E9AW84
C	-3	ARG	-	expression tag	UNP E9AW84
C	-2	GLY	-	expression tag	UNP E9AW84
C	-1	SER	-	expression tag	UNP E9AW84

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	HIS	-	expression tag	UNP E9AW84
C	6	LEU	MET	conflict	UNP E9AW84
C	12	ALA	THR	conflict	UNP E9AW84
C	31	PRO	LYS	conflict	UNP E9AW84
C	34	ILE	VAL	conflict	UNP E9AW84
C	45	LYS	GLN	conflict	UNP E9AW84
C	59	GLY	HIS	conflict	UNP E9AW84
C	61	MET	ILE	conflict	UNP E9AW84
C	63	HIS	ASN	conflict	UNP E9AW84
C	64	SER	VAL	conflict	UNP E9AW84
C	65	GLY	SER	conflict	UNP E9AW84
C	67	SER	LYS	conflict	UNP E9AW84
C	69	ARG	LYS	conflict	UNP E9AW84
C	71	GLY	ASP	conflict	UNP E9AW84
C	107	SER	ALA	conflict	UNP E9AW84
C	112	MET	ILE	conflict	UNP E9AW84
C	123	ALA	CYS	engineered mutation	UNP E9AW84
C	134	MET	THR	conflict	UNP E9AW84
C	139	SER	ALA	conflict	UNP E9AW84
C	145	THR	VAL	conflict	UNP E9AW84
C	157	SER	THR	conflict	UNP E9AW84
C	176	GLN	LYS	conflict	UNP E9AW84
C	229	THR	SER	conflict	UNP E9AW84
C	268	VAL	ALA	conflict	UNP E9AW84
C	277	ARG	LYS	conflict	UNP E9AW84
C	327	ILE	VAL	conflict	UNP E9AW84
C	356	GLU	ASP	conflict	UNP E9AW84
C	368	ASP	GLU	conflict	UNP E9AW84
C	398	ILE	VAL	conflict	UNP E9AW84
C	410	GLU	GLY	conflict	UNP E9AW84
D	-15	HIS	-	expression tag	UNP E9AW84
D	-14	HIS	-	expression tag	UNP E9AW84
D	-13	HIS	-	expression tag	UNP E9AW84
D	-12	HIS	-	expression tag	UNP E9AW84
D	-11	HIS	-	expression tag	UNP E9AW84
D	-10	HIS	-	expression tag	UNP E9AW84
D	-9	SER	-	expression tag	UNP E9AW84
D	-8	SER	-	expression tag	UNP E9AW84
D	-7	GLY	-	expression tag	UNP E9AW84
D	-6	LEU	-	expression tag	UNP E9AW84
D	-5	VAL	-	expression tag	UNP E9AW84
D	-4	PRO	-	expression tag	UNP E9AW84

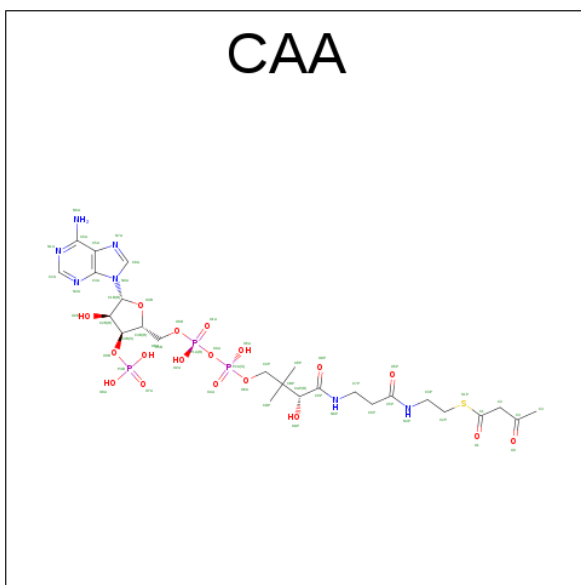
*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	-3	ARG	-	expression tag	UNP E9AW84
D	-2	GLY	-	expression tag	UNP E9AW84
D	-1	SER	-	expression tag	UNP E9AW84
D	0	HIS	-	expression tag	UNP E9AW84
D	6	LEU	MET	conflict	UNP E9AW84
D	12	ALA	THR	conflict	UNP E9AW84
D	31	PRO	LYS	conflict	UNP E9AW84
D	34	ILE	VAL	conflict	UNP E9AW84
D	45	LYS	GLN	conflict	UNP E9AW84
D	59	GLY	HIS	conflict	UNP E9AW84
D	61	MET	ILE	conflict	UNP E9AW84
D	63	HIS	ASN	conflict	UNP E9AW84
D	64	SER	VAL	conflict	UNP E9AW84
D	65	GLY	SER	conflict	UNP E9AW84
D	67	SER	LYS	conflict	UNP E9AW84
D	69	ARG	LYS	conflict	UNP E9AW84
D	71	GLY	ASP	conflict	UNP E9AW84
D	107	SER	ALA	conflict	UNP E9AW84
D	112	MET	ILE	conflict	UNP E9AW84
D	123	ALA	CYS	engineered mutation	UNP E9AW84
D	134	MET	THR	conflict	UNP E9AW84
D	139	SER	ALA	conflict	UNP E9AW84
D	145	THR	VAL	conflict	UNP E9AW84
D	157	SER	THR	conflict	UNP E9AW84
D	176	GLN	LYS	conflict	UNP E9AW84
D	229	THR	SER	conflict	UNP E9AW84
D	268	VAL	ALA	conflict	UNP E9AW84
D	277	ARG	LYS	conflict	UNP E9AW84
D	327	ILE	VAL	conflict	UNP E9AW84
D	356	GLU	ASP	conflict	UNP E9AW84
D	368	ASP	GLU	conflict	UNP E9AW84
D	398	ILE	VAL	conflict	UNP E9AW84
D	410	GLU	GLY	conflict	UNP E9AW84

- Molecule 2 is ACETOACETYL-COENZYME A (three-letter code: CAA) (formula: C<sub>25</sub>H<sub>40</sub>N<sub>7</sub>O<sub>18</sub>P<sub>3</sub>S).





Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	S	0	1
			108	50	14	36	6	2		
2	B	1	Total	C	N	O	P	S	0	0
			54	25	7	18	3	1		
2	D	1	Total	C	N	O	P	S	0	1
			108	50	14	36	6	2		

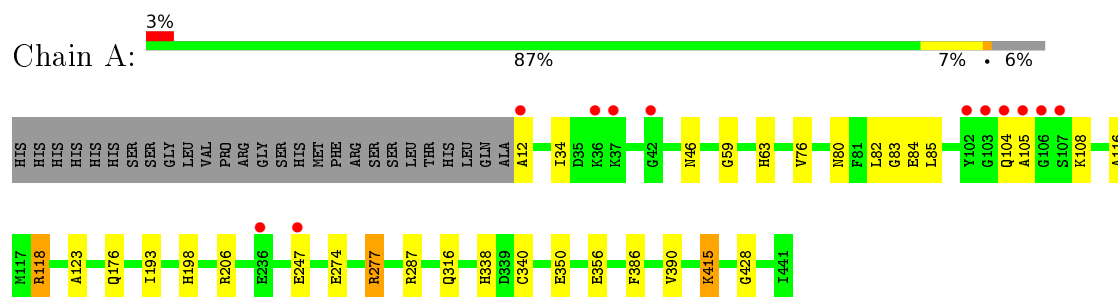
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	231	Total	O	0	0
			231	231		
3	B	233	Total	O	0	0
			233	233		
3	C	183	Total	O	0	0
			183	183		
3	D	183	Total	O	0	0
			183	183		

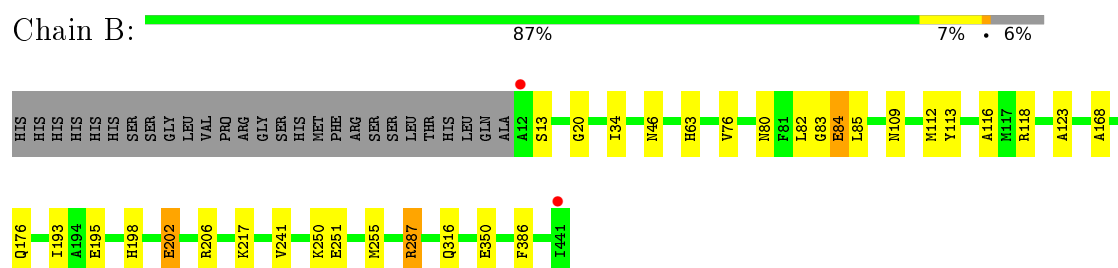
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

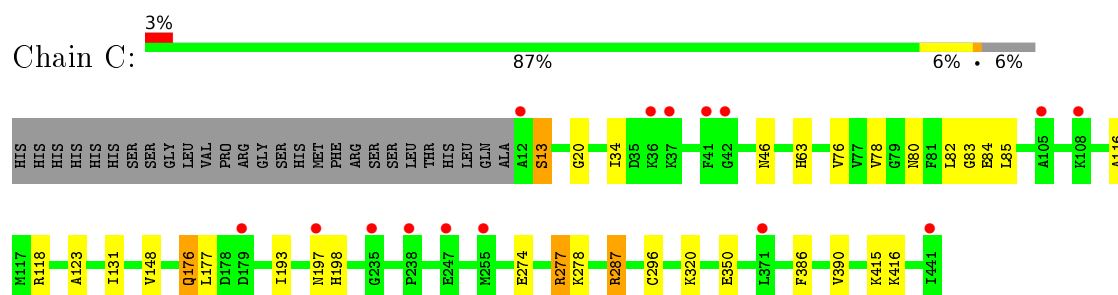
- Molecule 1: 3-ketoacyl-CoA thiolase-like protein



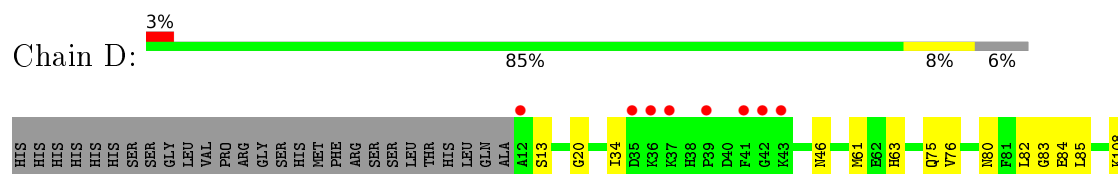
- Molecule 1: 3-ketoacyl-CoA thiolase-like protein

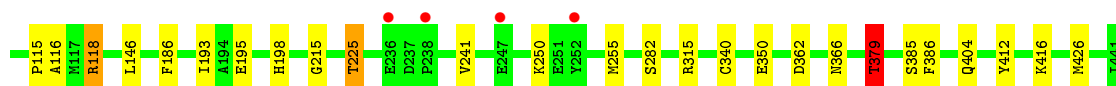


- Molecule 1: 3-ketoacyl-CoA thiolase-like protein



- Molecule 1: 3-ketoacyl-CoA thiolase-like protein





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.14Å 89.80Å 123.72Å 90.00° 104.76° 90.00°	Depositor
Resolution (Å)	119.63 – 1.98 56.03 – 1.98	Depositor EDS
% Data completeness (in resolution range)	99.1 (119.63-1.98) 99.1 (56.03-1.98)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.06 (at 1.98Å)	Xtriage
Refinement program	REFMAC 5.8.0131	Depositor
R, $R_{free}$	0.189 , 0.226 0.196 , 0.231	Depositor DCC
$R_{free}$ test set	5300 reflections (5.15%)	DCC
Wilson B-factor (Å <sup>2</sup> )	20.4	Xtriage
Anisotropy	0.108	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 39.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.044 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	14016	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.89% 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: CAA

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	0.79	0/3285	0.89	7/4429 (0.2%)
1	B	0.82	3/3285 (0.1%)	0.88	9/4429 (0.2%)
1	C	0.75	0/3285	0.91	8/4429 (0.2%)
1	D	0.77	0/3285	0.85	5/4429 (0.1%)
All	All	0.78	3/13140 (0.0%)	0.88	29/17716 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	251	GLU	CG-CD	6.02	1.60	1.51
1	B	84[A]	GLU	CD-OE2	5.43	1.31	1.25
1	B	84[B]	GLU	CD-OE2	5.43	1.31	1.25

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	287	ARG	NE-CZ-NH2	-8.82	115.89	120.30
1	C	287	ARG	NE-CZ-NH1	8.70	124.65	120.30
1	B	206	ARG	NE-CZ-NH1	8.62	124.61	120.30
1	A	118	ARG	NE-CZ-NH2	-8.51	116.05	120.30
1	C	277	ARG	NE-CZ-NH2	-8.46	116.07	120.30

There are no chirality outliers.

There are no planarity outliers.

### 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	3229	0	3208	20	2
1	B	3229	0	3208	21	0
1	C	3229	0	3208	15	2
1	D	3229	0	3208	29	0
2	A	108	0	71	16	0
2	B	54	0	36	5	0
2	D	108	0	72	15	0
3	A	231	0	0	3	0
3	B	233	0	0	6	0
3	C	183	0	0	1	0
3	D	183	0	0	2	0
All	All	14016	0	13011	103	2

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 4.

The worst 5 of 103 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)
2:A:501[A]:CAA:S1P	2:A:501[A]:CAA:H4'2	1.67	1.30
2:D:501[A]:CAA:H52A	2:D:501[A]:CAA:H8A	1.12	1.09
2:A:501[B]:CAA:CEP	2:A:501[B]:CAA:O9P	2.02	1.08
2:A:501[B]:CAA:H4'3	2:A:501[B]:CAA:S1P	1.90	1.07
1:D:340:CYS:SG	2:D:501[A]:CAA:H4'3	1.95	1.06

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	Interatomic distance (Å)	Clash overlap (Å)
1:A:105:ALA:O	1:C:197:ASN:CB[2_646]	2.07	0.13
1:A:105:ALA:O	1:C:197:ASN:O[2_646]	2.17	0.03

## 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	429/457 (94%)	422 (98%)	7 (2%)	0	100	100
1	B	429/457 (94%)	423 (99%)	6 (1%)	0	100	100
1	C	429/457 (94%)	419 (98%)	9 (2%)	1 (0%)	52	47
1	D	429/457 (94%)	422 (98%)	7 (2%)	0	100	100
All	All	1716/1828 (94%)	1686 (98%)	29 (2%)	1 (0%)	56	51

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	13	SER

### 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	339/362 (94%)	332 (98%)	7 (2%)	61	61
1	B	339/362 (94%)	335 (99%)	4 (1%)	78	80
1	C	339/362 (94%)	333 (98%)	6 (2%)	66	67
1	D	339/362 (94%)	330 (97%)	9 (3%)	52	50
All	All	1356/1448 (94%)	1330 (98%)	26 (2%)	65	65

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	176	GLN
1	C	386	PHE
1	D	386	PHE
1	C	287	ARG
1	C	320	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	216	ASN
1	B	359	HIS
1	D	198	HIS
1	B	319	GLN
1	C	38	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](#)

5 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	CAA	A	501[A]	-	47,56,56	1.00	3 (6%)	57,83,83	3.59	18 (31%)
2	CAA	A	501[B]	-	47,56,56	1.04	2 (4%)	57,83,83	3.61	12 (21%)
2	CAA	B	501	-	47,56,56	1.28	4 (8%)	57,83,83	2.08	15 (26%)
2	CAA	D	501[A]	-	47,56,56	1.22	5 (10%)	57,83,83	2.57	14 (24%)
2	CAA	D	501[B]	-	47,56,56	1.18	2 (4%)	57,83,83	1.69	6 (10%)

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	CAA	A	501[A]	-	-	0/50/71/71	0/3/3/3
2	CAA	A	501[B]	-	-	1/50/71/71	0/3/3/3
2	CAA	B	501	-	-	2/50/71/71	0/3/3/3
2	CAA	D	501[A]	-	-	2/50/71/71	0/3/3/3
2	CAA	D	501[B]	-	-	0/50/71/71	0/3/3/3

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	501[B]	CAA	C1-S1P	-5.70	1.68	1.76
2	B	501	CAA	C1-S1P	-4.61	1.70	1.76
2	A	501[B]	CAA	C1-S1P	-4.36	1.70	1.76
2	D	501[A]	CAA	C1-S1P	-3.09	1.72	1.76
2	A	501[A]	CAA	C1-S1P	-3.05	1.72	1.76

The worst 5 of 65 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501[B]	CAA	CEP-CBP-CAP	-14.09	83.47	109.17
2	A	501[A]	CAA	CEP-CBP-CAP	-13.21	85.07	109.17
2	A	501[B]	CAA	CDP-CBP-CAP	-12.74	85.93	109.17
2	A	501[A]	CAA	CDP-CBP-CAP	-10.24	90.49	109.17
2	D	501[A]	CAA	N3A-C2A-N1A	-9.12	121.70	128.87

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	501	CAA	O1-C1-S1P-C2P
2	B	501	CAA	C2-C1-S1P-C2P
2	A	501[B]	CAA	C4-C3-C2-C1
2	D	501[A]	CAA	O1-C1-S1P-C2P
2	D	501[A]	CAA	C2-C1-S1P-C2P

There are no ring outliers.

5 monomers are involved in 36 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501[A]	CAA	10	0
2	A	501[B]	CAA	6	0
2	B	501	CAA	5	0
2	D	501[A]	CAA	11	0
2	D	501[B]	CAA	4	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	430/457 (94%)	-0.03	12 (2%) 56 60	10, 18, 45, 132	0
1	B	430/457 (94%)	-0.10	2 (0%) 91 93	11, 18, 37, 65	0
1	C	430/457 (94%)	0.19	15 (3%) 48 52	11, 23, 48, 111	0
1	D	430/457 (94%)	0.06	12 (2%) 56 60	12, 22, 51, 87	0
All	All	1720/1828 (94%)	0.03	41 (2%) 62 66	10, 20, 44, 132	0

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	103	GLY	11.6
1	A	105	ALA	10.6
1	C	238	PRO	8.7
1	A	106	GLY	7.4
1	A	107	SER	5.5

### 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
2	CAA	D	501[A]	54/54	0.78	0.28	7.42	29,41,47,51	54
2	CAA	D	501[B]	54/54	0.78	0.28	7.24	17,122,137,146	54
2	CAA	A	501[A]	54/54	0.79	0.26	3.14	18,43,49,50	54
2	CAA	A	501[B]	54/54	0.79	0.26	3.11	14,147,214,218	54
2	CAA	B	501	54/54	0.94	0.11	0.18	19,25,41,50	0

## 6.5 Other polymers

There are no such residues in this entry.