



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 06:52 am GMT

PDB ID : 3EEO  
Title : 2.9A crystal structure of methyl-isocitrate lyase from Burkholderia pseudomallei  
Authors : Seattle Structural Genomics Center for Infectious Disease (SSGCID)  
Deposited on : 2008-09-28  
Resolution : 2.90 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  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

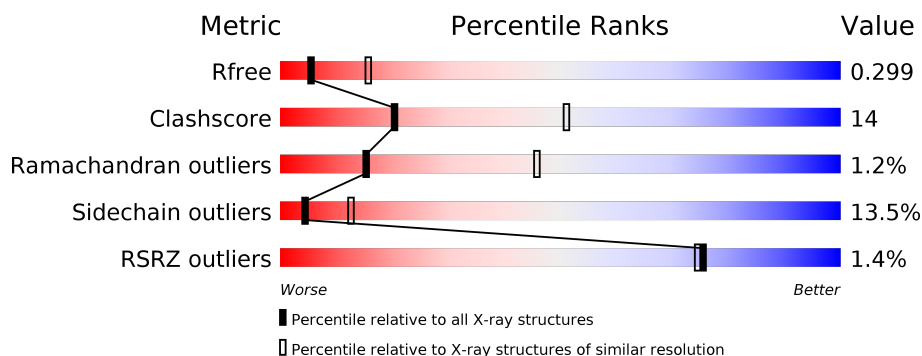
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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))
$R_{free}$	100719	1586 (2.90-2.90)
Clashscore	112137	1807 (2.90-2.90)
Ramachandran outliers	110173	1768 (2.90-2.90)
Sidechain outliers	110143	1770 (2.90-2.90)
RSRZ outliers	101464	1596 (2.90-2.90)

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	298	
1	B	298	
1	C	298	
1	D	298	
1	E	298	
1	F	298	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	298	<div><div></div><div>70%21%5%</div><div></div></div>
1	H	298	<div><div>%</div><div></div><div>63%28%5%</div><div></div></div>
1	I	298	<div><div></div><div>70%22%5%</div><div></div></div>
1	J	298	<div><div></div><div>67%24%5%</div><div></div></div>
1	K	298	<div><div></div><div>69%22%</div><div></div></div>
1	L	298	<div><div></div><div>62%29%5%</div><div></div></div>
1	M	298	<div><div>4%</div><div></div><div>62%28%6%</div><div></div></div>
1	N	298	<div><div>%</div><div></div><div>59%31%5%</div><div></div></div>
1	O	298	<div><div>8%</div><div></div><div>60%30%7%</div><div></div></div>
1	P	298	<div><div>5%</div><div></div><div>63%27%6%</div><div></div></div>

## 2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 34623 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.

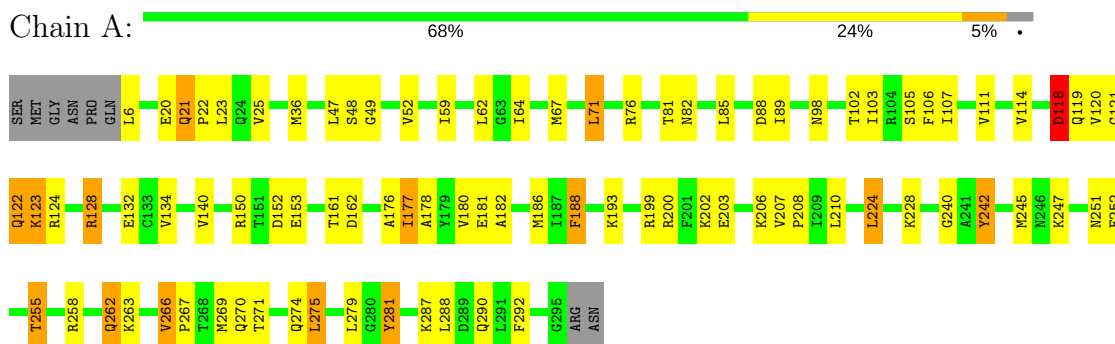
- Molecule 1 is a protein called Methylisocitrate lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	290	Total	C	N	O	S	0	0	0
			2183	1379	376	416	12			
1	B	286	Total	C	N	O	S	0	0	0
			2154	1362	370	410	12			
1	C	287	Total	C	N	O	S	0	0	0
			2162	1366	372	412	12			
1	D	287	Total	C	N	O	S	0	0	0
			2162	1368	371	411	12			
1	E	288	Total	C	N	O	S	0	0	0
			2170	1372	373	413	12			
1	F	286	Total	C	N	O	S	0	0	0
			2154	1362	370	410	12			
1	G	287	Total	C	N	O	S	0	0	0
			2162	1368	371	411	12			
1	H	288	Total	C	N	O	S	0	0	0
			2170	1372	373	413	12			
1	I	289	Total	C	N	O	S	0	0	0
			2179	1377	375	415	12			
1	J	288	Total	C	N	O	S	0	0	0
			2171	1373	373	413	12			
1	K	286	Total	C	N	O	S	0	0	0
			2154	1362	370	410	12			
1	L	286	Total	C	N	O	S	0	0	0
			2154	1362	370	410	12			
1	M	287	Total	C	N	O	S	0	0	0
			2162	1366	372	412	12			
1	N	287	Total	C	N	O	S	0	0	0
			2162	1368	371	411	12			
1	O	288	Total	C	N	O	S	0	0	0
			2170	1372	373	413	12			
1	P	286	Total	C	N	O	S	0	0	0
			2154	1362	370	410	12			

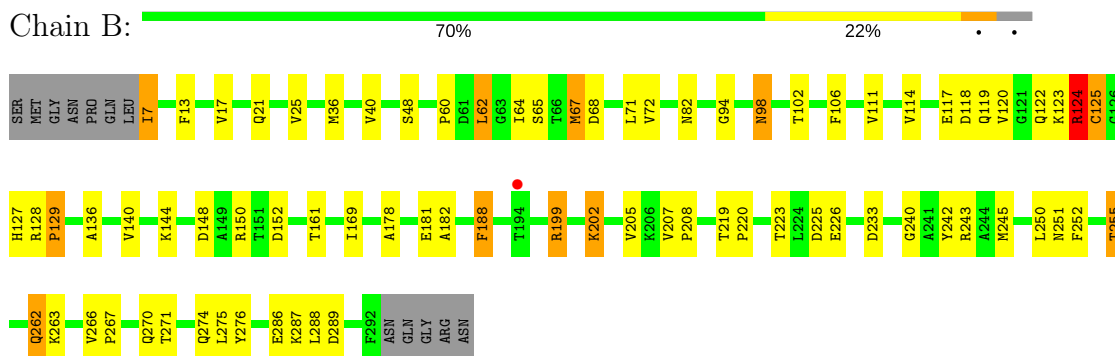
### 3 Residue-property plots [i](#)

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 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.

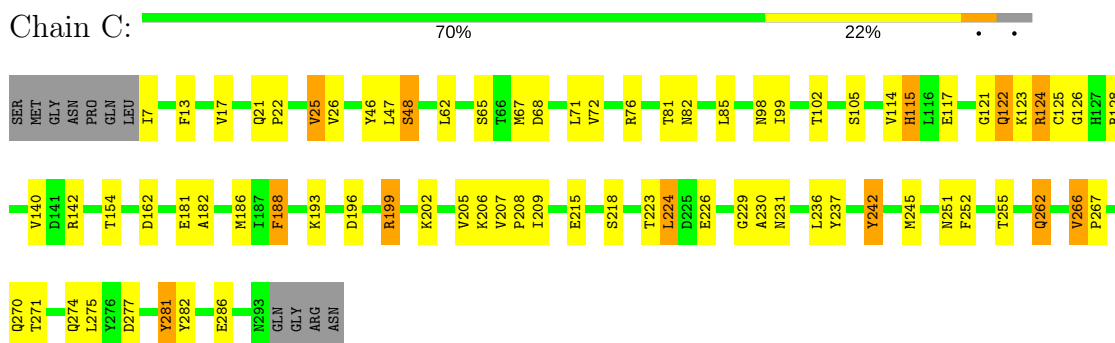
- Molecule 1: Methylisocitrate lyase



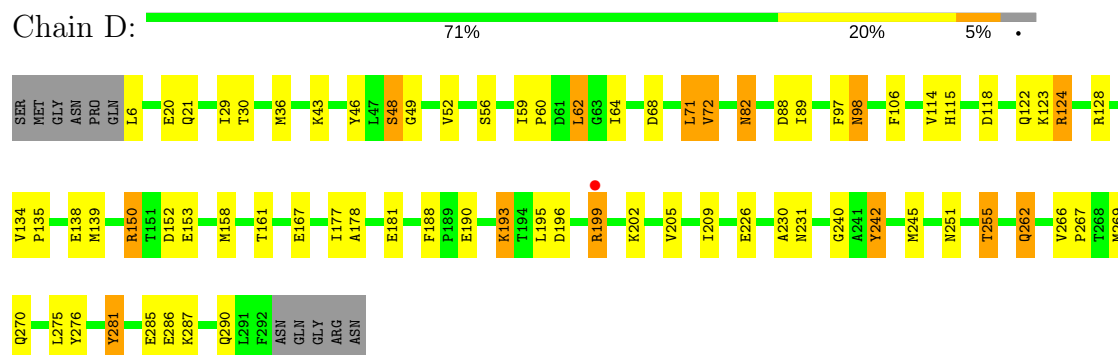
- Molecule 1: Methylisocitrate lyase



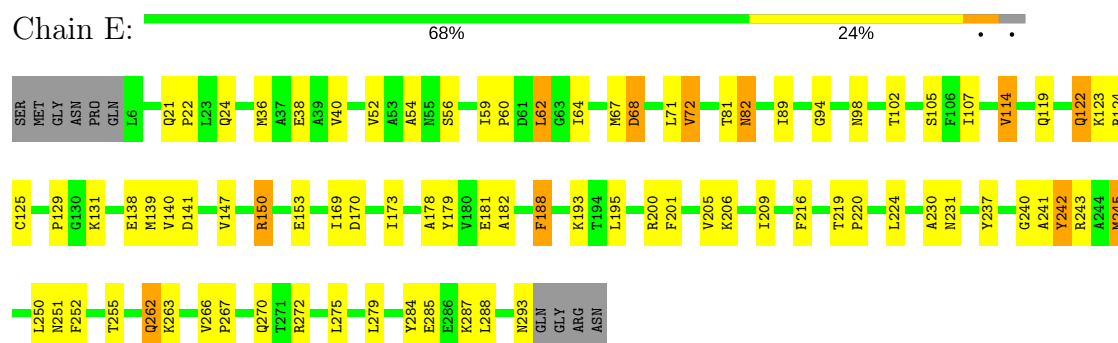
- Molecule 1: Methylisocitrate lyase



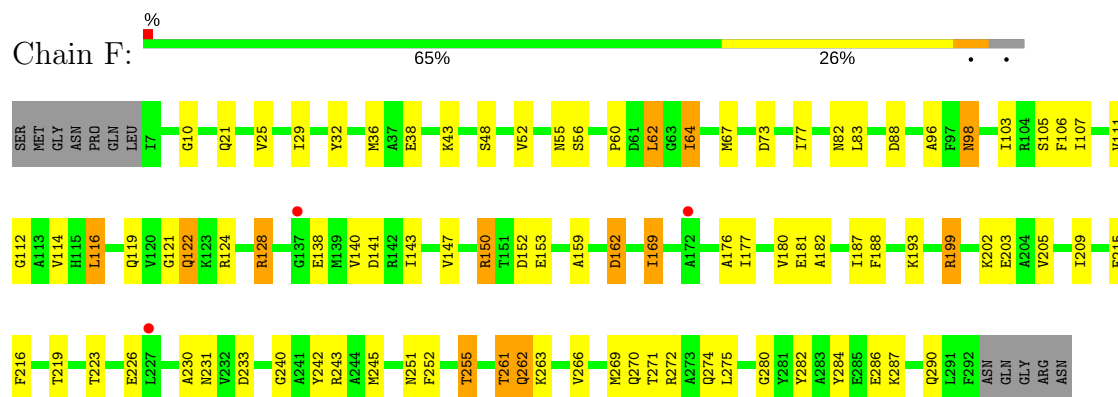
- Molecule 1: Methylisocitrate lyase



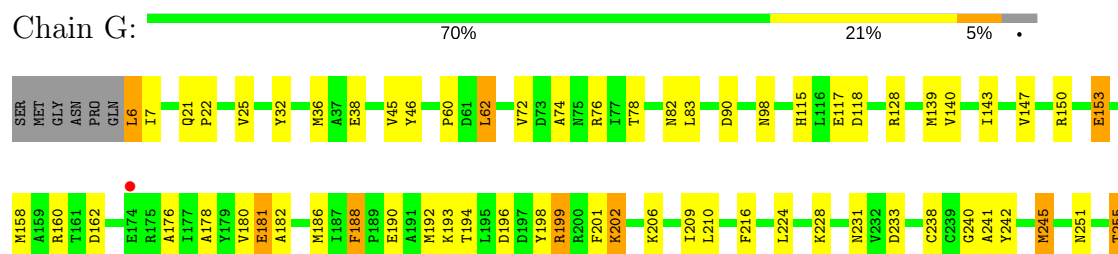
- Molecule 1: Methylisocitrate lyase

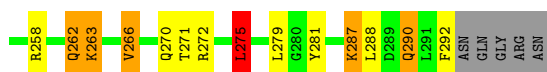


- Molecule 1: Methylisocitrate lyase

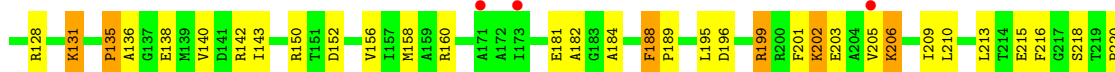


- Molecule 1: Methylisocitrate lyase





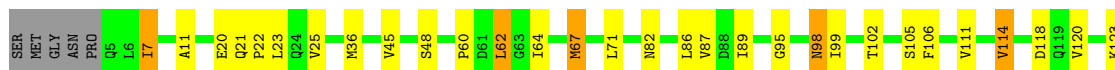
• Molecule 1: Methylisocitrate lyase



• Molecule 1: Methylisocitrate lyase

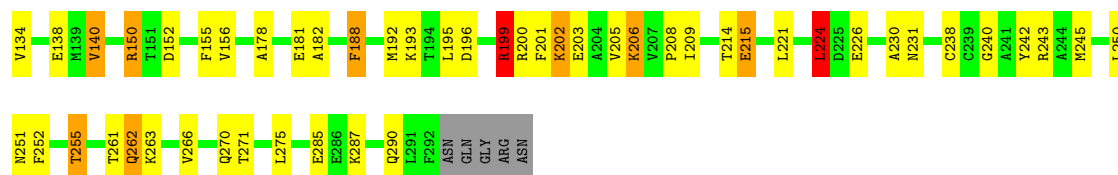


• Molecule 1: Methylisocitrate lyase



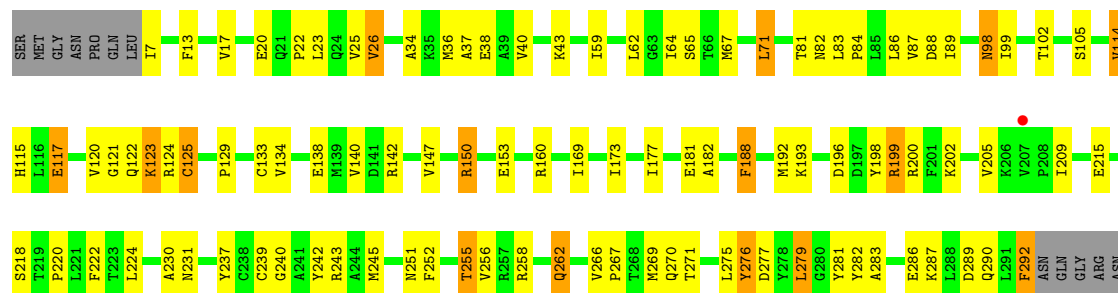
• Molecule 1: Methylisocitrate lyase





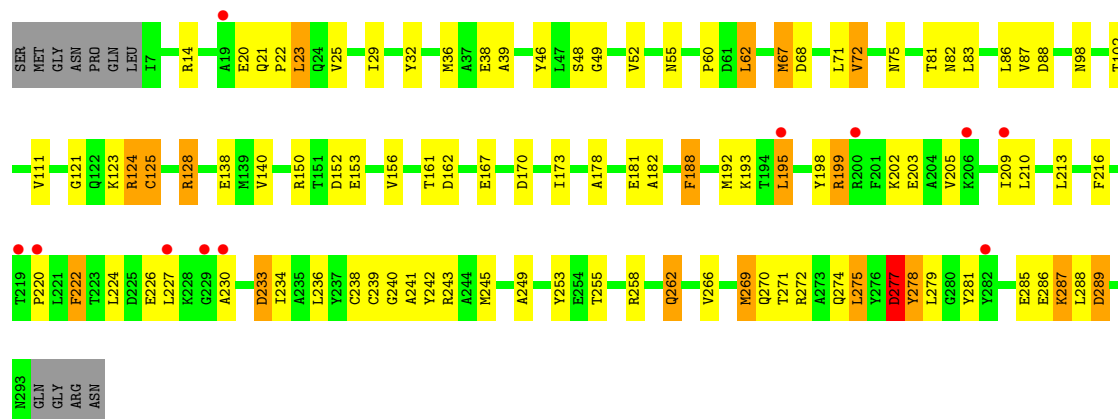
• Molecule 1: Methylisocitrate lyase

Chain L: 62% 29% 5%



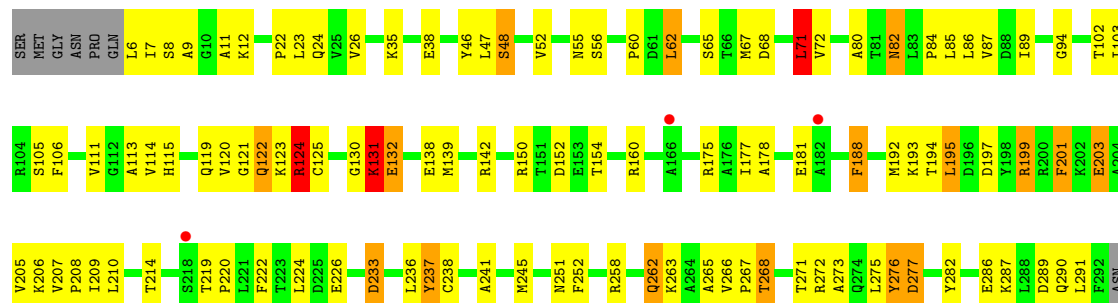
• Molecule 1: Methylisocitrate lyase

Chain M: 4% 62% 28% 6%



• Molecule 1: Methylisocitrate lyase

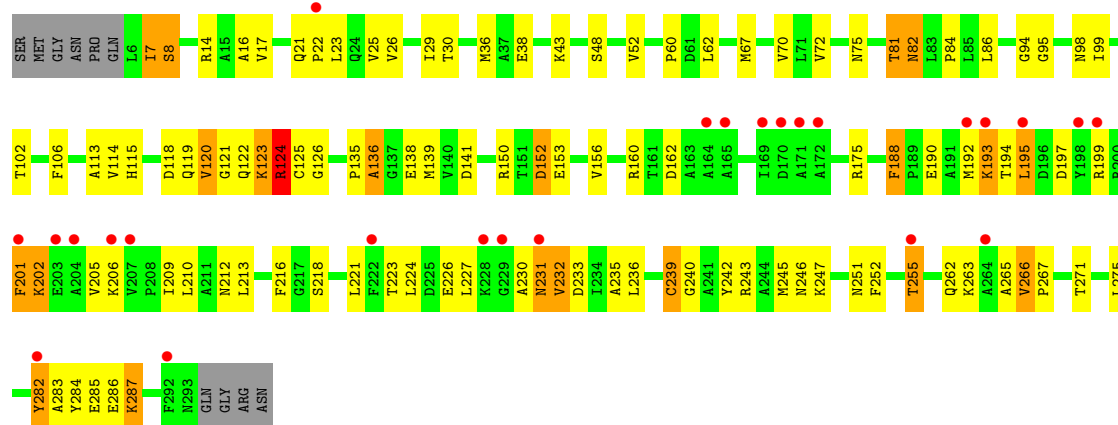
Chain N: 59% 31% 5%



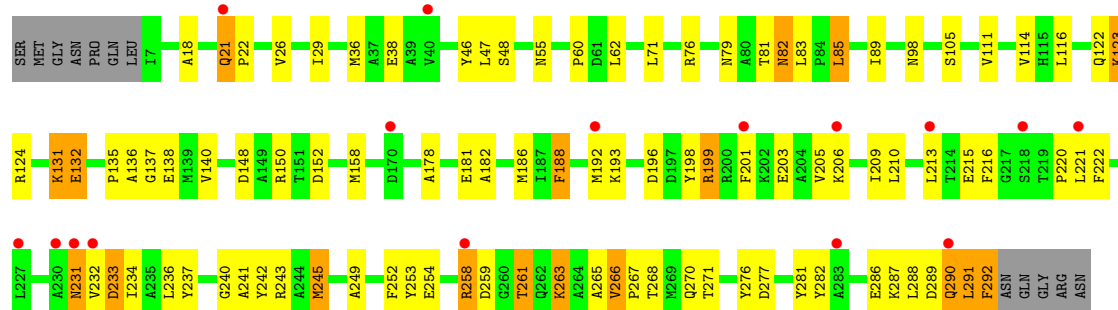


GLN  
GLY  
ARG  
ASN

• Molecule 1: Methylisocitrate lyase



• Molecule 1: Methylisocitrate lyase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	163.81Å 172.40Å 179.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.00 – 2.90 46.49 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.5 (48.00-2.90) 98.5 (46.49-2.90)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.49 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, $R_{free}$	0.237 , 0.299 0.237 , 0.299	Depositor DCC
$R_{free}$ test set	5514 reflections (5.23%)	DCC
Wilson B-factor (Å <sup>2</sup> )	53.2	Xtriage
Anisotropy	0.069	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 27.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	34623	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.61% 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

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	0.69	0/2218	0.80	2/3005 (0.1%)
1	B	0.67	0/2189	0.78	0/2966
1	C	0.65	0/2197	0.78	0/2977
1	D	0.64	0/2197	0.74	0/2977
1	E	0.67	0/2205	0.78	0/2988
1	F	0.67	0/2189	0.75	2/2966 (0.1%)
1	G	0.64	0/2197	0.77	2/2977 (0.1%)
1	H	0.69	1/2205 (0.0%)	0.75	0/2988
1	I	0.65	0/2214	0.77	1/3000 (0.0%)
1	J	0.67	0/2206	0.77	0/2989
1	K	0.70	0/2189	0.77	2/2966 (0.1%)
1	L	0.68	1/2189 (0.0%)	0.77	2/2966 (0.1%)
1	M	0.74	0/2197	0.81	0/2977
1	N	0.75	0/2197	0.83	1/2977 (0.0%)
1	O	0.83	1/2205 (0.0%)	0.79	1/2988 (0.0%)
1	P	0.74	0/2189	0.75	0/2966
All	All	0.69	3/35183 (0.0%)	0.78	13/47673 (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
1	N	0	1
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	O	239	CYS	CB-SG	-10.33	1.64	1.82
1	H	239	CYS	CB-SG	-5.44	1.73	1.81

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	239	CYS	CB-SG	-5.28	1.73	1.81

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	279	LEU	CA-CB-CG	6.08	129.29	115.30
1	A	76	ARG	NE-CZ-NH2	-5.88	117.36	120.30
1	K	199	ARG	NE-CZ-NH1	5.65	123.13	120.30
1	L	71	LEU	CA-CB-CG	5.65	128.30	115.30
1	G	275	LEU	CA-CB-CG	5.61	128.19	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	121	GLY	Peptide
1	N	122	GLN	Peptide

## 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	2183	0	2179	59	0
1	B	2154	0	2151	59	0
1	C	2162	0	2157	50	0
1	D	2162	0	2162	57	0
1	E	2170	0	2168	65	0
1	F	2154	0	2151	69	0
1	G	2162	0	2162	60	0
1	H	2170	0	2168	69	0
1	I	2179	0	2176	55	0
1	J	2171	0	2170	64	0
1	K	2154	0	2151	59	0
1	L	2154	0	2151	64	0
1	M	2162	0	2157	92	0
1	N	2162	0	2162	95	0
1	O	2170	0	2168	91	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	P	2154	0	2151	81	0
All	All	34623	0	34584	969	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 14.

The worst 5 of 969 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:I:67:MET:HE3	1:I:102:THR:HA	1.23	1.15
1:L:199:ARG:HH11	1:L:199:ARG:HB2	1.14	1.07
1:I:123:LYS:HD2	1:I:124:ARG:H	1.12	1.07
1:I:67:MET:CE	1:I:102:THR:HA	1.85	1.05
1:G:199:ARG:HH11	1:G:199:ARG:HB2	0.88	1.03

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	288/298 (97%)	273 (95%)	13 (4%)	2 (1%)	25	60
1	B	284/298 (95%)	268 (94%)	12 (4%)	4 (1%)	13	41
1	C	285/298 (96%)	266 (93%)	16 (6%)	3 (1%)	17	48
1	D	285/298 (96%)	274 (96%)	9 (3%)	2 (1%)	25	60
1	E	286/298 (96%)	271 (95%)	13 (4%)	2 (1%)	25	60
1	F	284/298 (95%)	268 (94%)	14 (5%)	2 (1%)	25	60
1	G	285/298 (96%)	272 (95%)	13 (5%)	0	100	100
1	H	286/298 (96%)	272 (95%)	12 (4%)	2 (1%)	25	60

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	I	287/298 (96%)	274 (96%)	11 (4%)	2 (1%)	25	60
1	J	286/298 (96%)	274 (96%)	8 (3%)	4 (1%)	13	41
1	K	284/298 (95%)	272 (96%)	10 (4%)	2 (1%)	25	60
1	L	284/298 (95%)	270 (95%)	10 (4%)	4 (1%)	13	41
1	M	285/298 (96%)	262 (92%)	15 (5%)	8 (3%)	6	22
1	N	285/298 (96%)	267 (94%)	12 (4%)	6 (2%)	8	30
1	O	286/298 (96%)	269 (94%)	10 (4%)	7 (2%)	7	27
1	P	284/298 (95%)	271 (95%)	9 (3%)	4 (1%)	13	41
All	All	4564/4768 (96%)	4323 (95%)	187 (4%)	54 (1%)	15	46

5 of 54 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	125	CYS
1	C	281	TYR
1	D	124	ARG
1	E	124	ARG
1	F	122	GLN

### 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	216/223 (97%)	186 (86%)	30 (14%)	4	12
1	B	213/223 (96%)	190 (89%)	23 (11%)	7	23
1	C	214/223 (96%)	184 (86%)	30 (14%)	4	12
1	D	214/223 (96%)	189 (88%)	25 (12%)	6	18
1	E	215/223 (96%)	191 (89%)	24 (11%)	7	21
1	F	213/223 (96%)	184 (86%)	29 (14%)	4	12
1	G	214/223 (96%)	182 (85%)	32 (15%)	3	10
1	H	215/223 (96%)	185 (86%)	30 (14%)	4	12

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	I	216/223 (97%)	187 (87%)	29 (13%)	4	13
1	J	215/223 (96%)	185 (86%)	30 (14%)	4	12
1	K	213/223 (96%)	185 (87%)	28 (13%)	5	14
1	L	213/223 (96%)	180 (84%)	33 (16%)	3	9
1	M	214/223 (96%)	187 (87%)	27 (13%)	5	15
1	N	214/223 (96%)	181 (85%)	33 (15%)	3	10
1	O	215/223 (96%)	184 (86%)	31 (14%)	4	11
1	P	213/223 (96%)	184 (86%)	29 (14%)	4	12
All	All	3427/3568 (96%)	2964 (86%)	463 (14%)	4	12

5 of 463 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	215	GLU
1	J	98	ASN
1	O	233	ASP
1	H	261	THR
1	I	131	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 88 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	290	GLN
1	I	122	GLN
1	O	246	ASN
1	H	98	ASN
1	H	270	GLN

### 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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 ⓘ

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	290/298 (97%)	0.01	0 100 100	26, 44, 60, 67	0
1	B	286/298 (95%)	0.01	1 (0%) 93 93	28, 45, 60, 67	0
1	C	287/298 (96%)	0.07	0 100 100	29, 46, 61, 69	0
1	D	287/298 (96%)	0.04	1 (0%) 93 93	32, 46, 61, 67	0
1	E	288/298 (96%)	0.03	0 100 100	32, 46, 62, 73	0
1	F	286/298 (95%)	0.13	3 (1%) 82 81	32, 48, 63, 71	0
1	G	287/298 (96%)	-0.03	1 (0%) 93 93	30, 46, 60, 68	0
1	H	288/298 (96%)	-0.01	3 (1%) 82 81	34, 47, 62, 70	0
1	I	289/298 (96%)	-0.05	1 (0%) 93 93	33, 46, 61, 70	0
1	J	288/298 (96%)	-0.14	0 100 100	33, 47, 62, 67	0
1	K	286/298 (95%)	0.06	0 100 100	32, 47, 62, 77	0
1	L	286/298 (95%)	0.16	1 (0%) 93 93	33, 48, 64, 74	0
1	M	287/298 (96%)	0.27	11 (3%) 41 35	33, 49, 65, 80	0
1	N	287/298 (96%)	0.21	3 (1%) 82 81	33, 50, 65, 72	0
1	O	288/298 (96%)	0.44	25 (8%) 11 8	34, 50, 71, 88	0
1	P	286/298 (95%)	0.37	16 (5%) 25 20	33, 49, 69, 78	0
All	All	4596/4768 (96%)	0.10	66 (1%) 75 74	26, 47, 63, 88	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	M	282	TYR	3.9
1	P	218	SER	3.4
1	O	222	PHE	3.4
1	O	164	ALA	3.3
1	P	232	VAL	3.3

## 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.