



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 02:50 am BST

PDB ID : 3FA4
Title : Crystal structure of 2,3-dimethylmalate lyase, a PEP mutase/isocitrate lyase superfamily member, triclinic crystal form
Authors : Narayanan, B.C.; Herzberg, O.
Deposited on : 2008-11-14
Resolution : 2.18 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

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
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

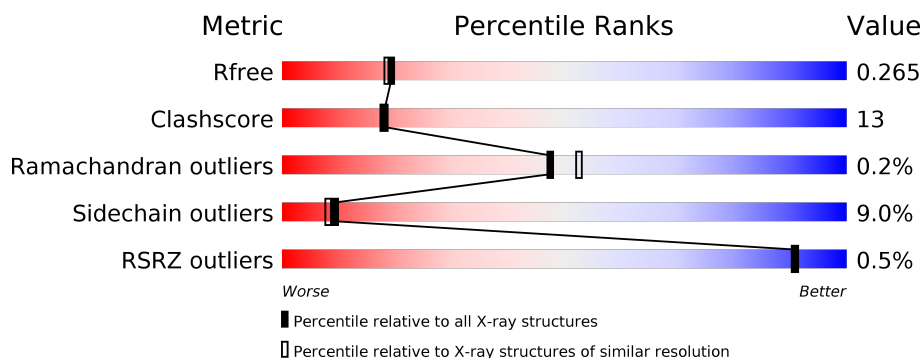
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.18 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	6864 (2.20-2.16)
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)
RSRZ outliers	127900	6738 (2.20-2.16)

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 ≥ 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	302	<div> <div></div> <div>70% 21% • 6%</div> </div>
1	B	302	<div> <div>%</div> <div>73% 21% • 5%</div> </div>
1	C	302	<div> <div>%</div> <div>69% 22% • 6%</div> </div>
1	D	302	<div> <div>%</div> <div>72% 19% • 6%</div> </div>
1	E	302	<div> <div></div> <div>70% 21% • • 5%</div> </div>
1	F	302	<div> <div>%</div> <div>73% 18% • • 6%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	302	<div><div></div><div>69%22%6%</div></div>
1	H	302	<div><div></div><div>65%26%7%</div></div>
1	I	302	<div><div></div><div>%67%21%6%7%</div></div>
1	J	302	<div><div></div><div>62%26%5%7%</div></div>
1	K	302	<div><div></div><div>72%16%7%</div></div>
1	L	302	<div><div></div><div>%69%23%5%</div></div>

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 26432 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 2,3-dimethylmalate lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	284	Total	C	N	O	S	0	0	0
			2111	1312	375	409	15			
1	B	288	Total	C	N	O	S	0	0	0
			2137	1330	379	413	15			
1	C	284	Total	C	N	O	S	0	0	0
			2112	1313	375	409	15			
1	D	284	Total	C	N	O	S	0	0	0
			2111	1312	375	409	15			
1	E	286	Total	C	N	O	S	0	0	0
			2118	1316	377	410	15			
1	F	285	Total	C	N	O	S	0	0	0
			2116	1315	376	410	15			
1	G	284	Total	C	N	O	S	0	0	0
			2107	1310	375	407	15			
1	H	282	Total	C	N	O	S	0	0	0
			2080	1293	369	403	15			
1	I	282	Total	C	N	O	S	0	0	0
			2100	1305	373	407	15			
1	J	280	Total	C	N	O	S	0	0	0
			2068	1284	369	400	15			
1	K	280	Total	C	N	O	S	0	0	0
			2072	1287	368	403	14			
1	L	288	Total	C	N	O	S	0	0	0
			2126	1320	384	407	15			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	1	Total	Mg	0	0
			1	1		
2	J	1	Total	Mg	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	1	Total 1	Mg 1	0	0
2	K	1	Total 1	Mg 1	0	0
2	E	1	Total 1	Mg 1	0	0
2	H	1	Total 1	Mg 1	0	0
2	B	1	Total 1	Mg 1	0	0
2	I	1	Total 1	Mg 1	0	0
2	C	1	Total 1	Mg 1	0	0
2	A	1	Total 1	Mg 1	0	0
2	L	1	Total 1	Mg 1	0	0
2	F	1	Total 1	Mg 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	116	Total 116	O 116	0	0
3	B	115	Total 115	O 115	0	0
3	C	109	Total 109	O 109	0	0
3	D	135	Total 135	O 135	0	0
3	E	104	Total 104	O 104	0	0
3	F	109	Total 109	O 109	0	0
3	G	124	Total 124	O 124	0	0
3	H	104	Total 104	O 104	0	0
3	I	38	Total 38	O 38	0	0

Continued on next page...

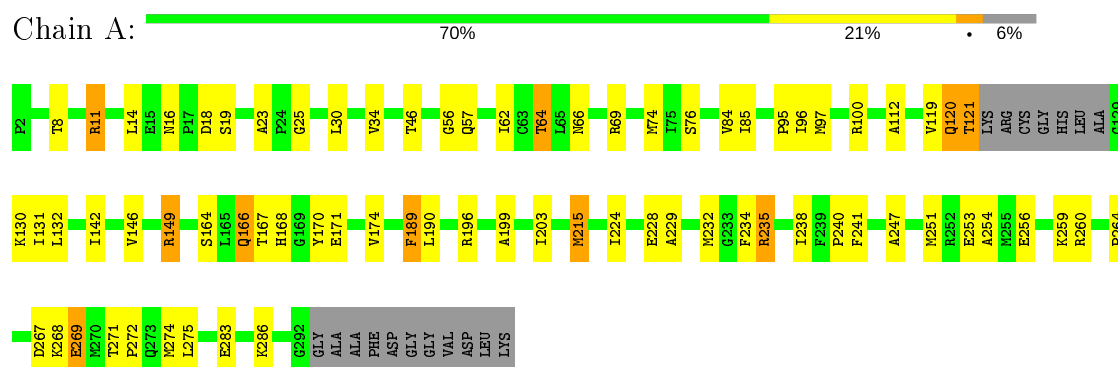
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	J	83	Total 83	O 83	0	0
3	K	61	Total 61	O 61	0	0
3	L	64	Total 64	O 64	0	0

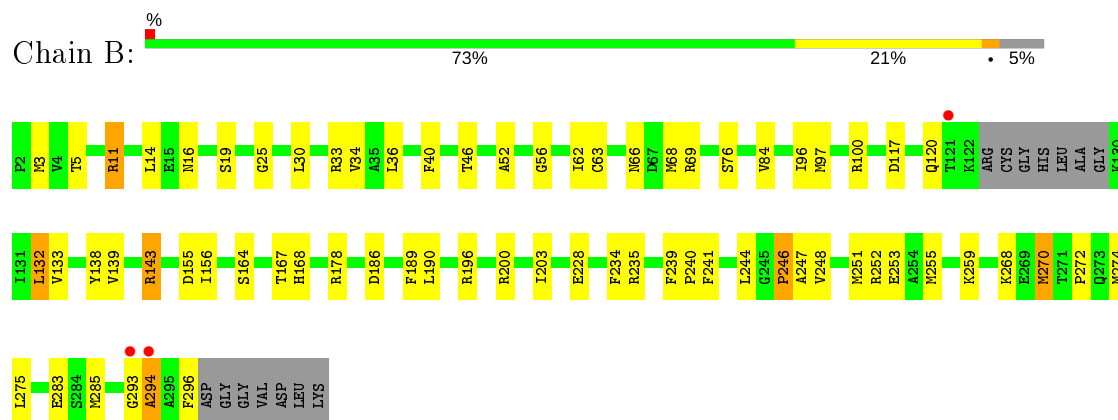
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 ($\text{RSRZ} > 2$). 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.

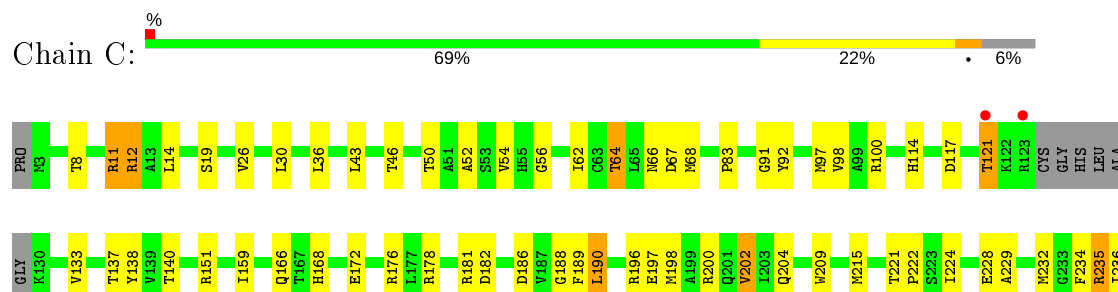
• Molecule 1: 2,3-dimethylmalate lyase

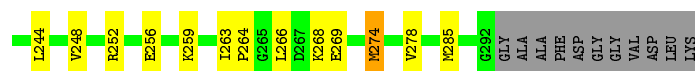


• Molecule 1: 2,3-dimethylmalate lyase

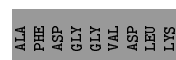
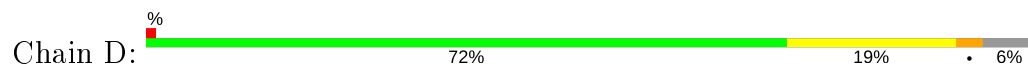


• Molecule 1: 2,3-dimethylmalate lyase

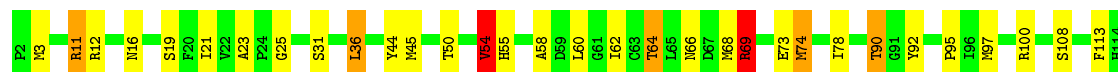




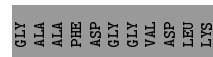
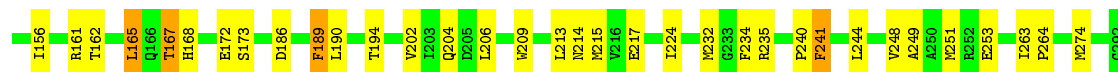
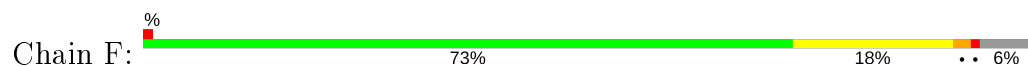
- Molecule 1: 2,3-dimethylmalate lyase



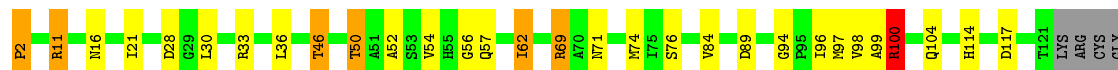
- Molecule 1: 2,3-dimethylmalate lyase

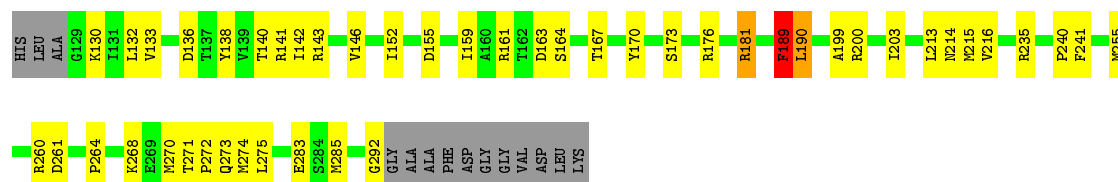


- Molecule 1: 2,3-dimethylmalate lyase



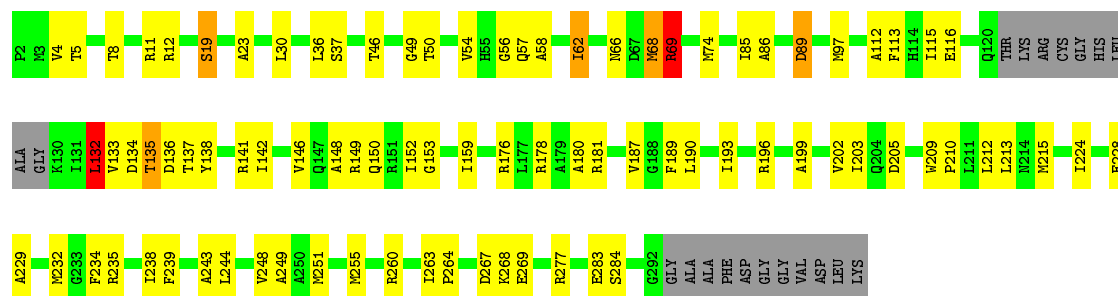
- Molecule 1: 2,3-dimethylmalate lyase





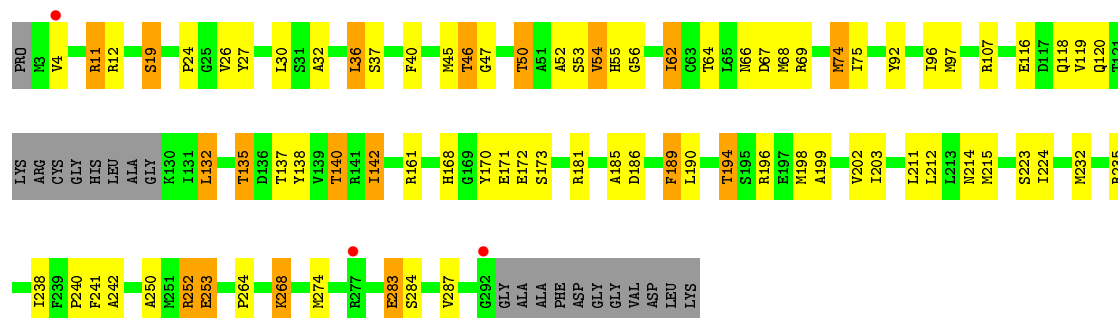
• Molecule 1: 2,3-dimethylmalate lyase

Chain H: 65% 26% 7%



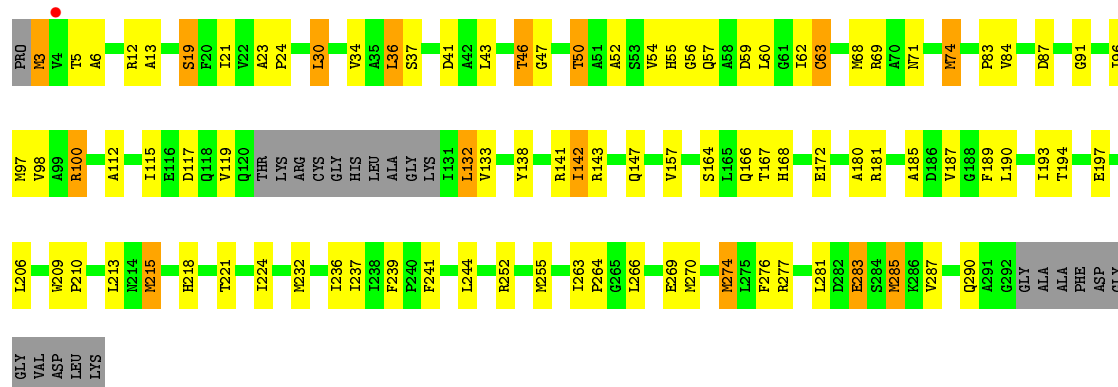
• Molecule 1: 2,3-dimethylmalate lyase

Chain I: 67% 21% 6% 7%



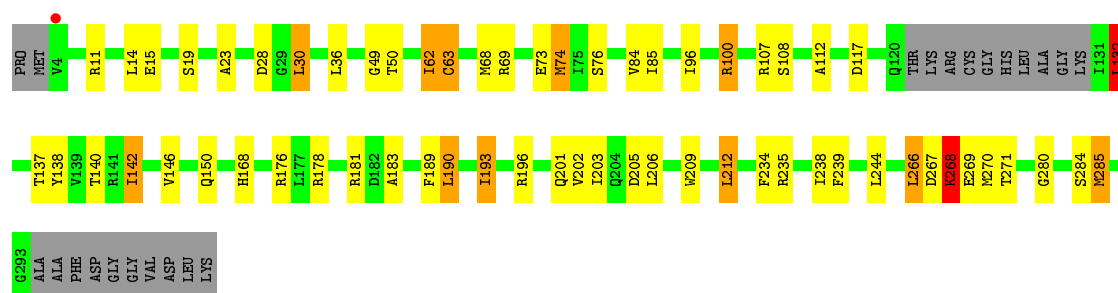
• Molecule 1: 2,3-dimethylmalate lyase

Chain J: 62% 26% 5% 7%



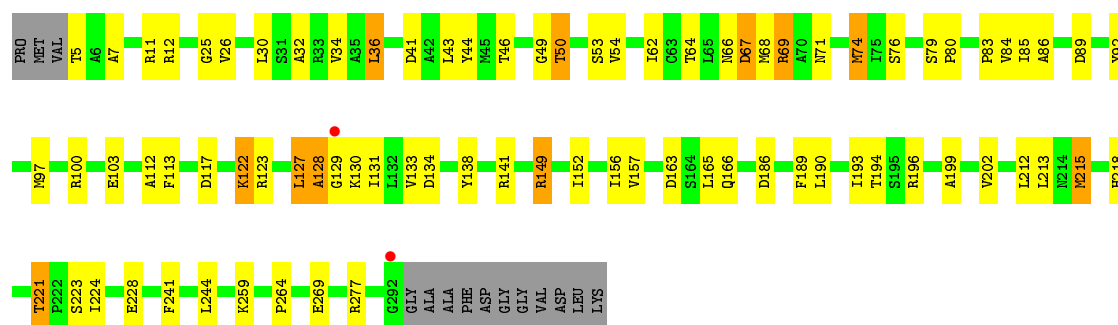
• Molecule 1: 2,3-dimethylmalate lyase

Chain K:  72% 16% 7%



• Molecule 1: 2,3-dimethylmalate lyase

Chain L:  69% 23% 5%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	79.09 Å 115.47 Å 115.75 Å 119.66° 90.71° 96.28°	Depositor
Resolution (Å)	46.83 – 2.18 46.81 – 2.18	Depositor EDS
% Data completeness (in resolution range)	93.2 (46.83-2.18) 91.8 (46.81-2.18)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.33 (at 2.18 Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.200 , 0.265 0.202 , 0.265	Depositor DCC
R_{free} test set	8603 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	34.7	Xtriage
Anisotropy	0.115	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 38.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.159 for -h,-k-l,l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	26432	wwPDB-VP
Average B, all atoms (Å ²)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.23% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: MG

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	1.09	1/2141 (0.0%)	1.04	5/2898 (0.2%)
1	B	1.11	1/2168 (0.0%)	1.03	2/2935 (0.1%)
1	C	1.08	2/2141 (0.1%)	1.08	8/2898 (0.3%)
1	D	1.12	2/2141 (0.1%)	1.06	5/2898 (0.2%)
1	E	1.07	4/2148 (0.2%)	1.08	8/2907 (0.3%)
1	F	1.02	0/2146	1.03	4/2905 (0.1%)
1	G	1.16	4/2137 (0.2%)	1.08	8/2893 (0.3%)
1	H	1.09	2/2110 (0.1%)	1.06	4/2861 (0.1%)
1	I	0.93	1/2129 (0.0%)	0.96	1/2882 (0.0%)
1	J	0.97	1/2097 (0.0%)	0.95	3/2841 (0.1%)
1	K	0.96	1/2101 (0.0%)	0.94	3/2846 (0.1%)
1	L	0.92	1/2157 (0.0%)	0.96	4/2922 (0.1%)
All	All	1.05	20/25616 (0.1%)	1.02	55/34686 (0.2%)

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	15	GLU	CG-CD	5.95	1.60	1.51
1	J	63	CYS	CB-SG	-5.88	1.72	1.81
1	G	283	GLU	CG-CD	5.88	1.60	1.51
1	E	54	VAL	CB-CG2	-5.87	1.40	1.52
1	B	283	GLU	CG-CD	5.77	1.60	1.51

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	176	ARG	NE-CZ-NH1	9.53	125.07	120.30
1	C	235	ARG	NE-CZ-NH1	8.83	124.71	120.30
1	C	235	ARG	NE-CZ-NH2	-8.22	116.19	120.30
1	E	235	ARG	NE-CZ-NH1	8.17	124.39	120.30
1	L	149	ARG	NE-CZ-NH2	-8.01	116.30	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	2111	0	2110	71	0
1	B	2137	0	2131	79	0
1	C	2112	0	2107	61	0
1	D	2111	0	2110	53	0
1	E	2118	0	2113	68	0
1	F	2116	0	2112	64	0
1	G	2107	0	2106	53	0
1	H	2080	0	2057	72	0
1	I	2100	0	2099	72	0
1	J	2068	0	2046	85	0
1	K	2072	0	2053	48	0
1	L	2126	0	2119	74	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
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
2	I	1	0	0	0	0
2	J	1	0	0	0	0
2	K	1	0	0	0	0
2	L	1	0	0	0	0
3	A	116	0	0	3	0
3	B	115	0	0	4	0
3	C	109	0	0	3	0
3	D	135	0	0	2	0
3	E	104	0	0	3	0
3	F	109	0	0	6	0
3	G	124	0	0	6	0
3	H	104	0	0	7	0
3	I	38	0	0	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	J	83	0	0	10	0
3	K	61	0	0	2	0
3	L	64	0	0	4	0
All	All	26432	0	25163	678	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 13.

The worst 5 of 678 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:A:97:MET:CE	1:A:100:ARG:HD3	1.65	1.25
1:E:74:MET:CE	1:E:78:ILE:HD11	1.72	1.19
1:A:97:MET:HE2	1:A:100:ARG:HD3	1.19	1.18
1:A:62:ILE:HG22	1:D:97:MET:HE3	1.30	1.12
1:B:97:MET:HE1	1:B:100:ARG:HD3	1.34	1.09

There are no symmetry-related clashes.

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	280/302 (93%)	273 (98%)	7 (2%)	0	100	100
1	B	284/302 (94%)	270 (95%)	13 (5%)	1 (0%)	34	35
1	C	280/302 (93%)	273 (98%)	7 (2%)	0	100	100
1	D	280/302 (93%)	267 (95%)	13 (5%)	0	100	100
1	E	282/302 (93%)	276 (98%)	6 (2%)	0	100	100
1	F	281/302 (93%)	273 (97%)	8 (3%)	0	100	100
1	G	280/302 (93%)	274 (98%)	6 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	H	278/302 (92%)	265 (95%)	11 (4%)	2 (1%)	22	20
1	I	278/302 (92%)	268 (96%)	10 (4%)	0	100	100
1	J	276/302 (91%)	263 (95%)	12 (4%)	1 (0%)	34	35
1	K	276/302 (91%)	268 (97%)	6 (2%)	2 (1%)	22	20
1	L	286/302 (95%)	276 (96%)	10 (4%)	0	100	100
All	All	3361/3624 (93%)	3246 (97%)	109 (3%)	6 (0%)	47	52

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	294	ALA
1	H	268	LYS
1	J	119	VAL
1	K	268	LYS
1	H	89	ASP

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	217/228 (95%)	201 (93%)	16 (7%)	13	13
1	B	218/228 (96%)	203 (93%)	15 (7%)	15	15
1	C	216/228 (95%)	198 (92%)	18 (8%)	11	10
1	D	217/228 (95%)	202 (93%)	15 (7%)	15	15
1	E	216/228 (95%)	196 (91%)	20 (9%)	9	7
1	F	217/228 (95%)	200 (92%)	17 (8%)	12	11
1	G	216/228 (95%)	196 (91%)	20 (9%)	9	7
1	H	211/228 (92%)	193 (92%)	18 (8%)	10	9
1	I	216/228 (95%)	189 (88%)	27 (12%)	4	3
1	J	209/228 (92%)	181 (87%)	28 (13%)	4	2
1	K	210/228 (92%)	193 (92%)	17 (8%)	11	10

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	L	215/228 (94%)	193 (90%)	22 (10%)	7	6
All	All	2578/2736 (94%)	2345 (91%)	233 (9%)	9	8

5 of 233 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	50	THR
1	H	69	ARG
1	L	69	ARG
1	G	69	ARG
1	G	260	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	204	GLN
1	G	273	GLN
1	H	118	GLN
1	L	71	ASN

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](#)

Of 12 ligands modelled in this entry, 12 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

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, 95th 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 ²)	Q<0.9
1	A	284/302 (94%)	-0.39	0 100 100	30, 36, 43, 54	0
1	B	288/302 (95%)	-0.37	3 (1%) 82 82	29, 36, 44, 53	0
1	C	284/302 (94%)	-0.37	2 (0%) 87 88	30, 36, 43, 50	0
1	D	284/302 (94%)	-0.41	2 (0%) 87 88	31, 36, 43, 50	0
1	E	286/302 (94%)	-0.34	1 (0%) 94 94	31, 36, 43, 52	0
1	F	285/302 (94%)	-0.38	2 (0%) 87 88	28, 36, 44, 52	0
1	G	284/302 (94%)	-0.43	0 100 100	27, 36, 43, 49	0
1	H	282/302 (93%)	-0.40	0 100 100	29, 36, 43, 48	0
1	I	282/302 (93%)	-0.20	3 (1%) 80 80	30, 37, 43, 49	0
1	J	280/302 (92%)	-0.27	1 (0%) 92 92	29, 36, 42, 49	0
1	K	280/302 (92%)	-0.31	1 (0%) 92 92	31, 36, 42, 46	0
1	L	288/302 (95%)	-0.25	2 (0%) 87 88	30, 36, 43, 52	0
All	All	3407/3624 (94%)	-0.34	17 (0%) 91 91	27, 36, 43, 54	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	K	4	VAL	5.1
1	L	129	GLY	4.9
1	E	129	GLY	4.6
1	C	123	ARG	3.2
1	I	4	VAL	3.1

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

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. The B-factors column lists the minimum, median, 95th 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(Å ²)	Q<0.9
2	MG	L	401	1/1	0.90	0.07	48,48,48,48	0
2	MG	A	401	1/1	0.90	0.07	51,51,51,51	0
2	MG	J	401	1/1	0.92	0.06	47,47,47,47	0
2	MG	F	401	1/1	0.94	0.06	44,44,44,44	0
2	MG	D	401	1/1	0.94	0.05	40,40,40,40	0
2	MG	B	401	1/1	0.95	0.05	35,35,35,35	0
2	MG	I	401	1/1	0.97	0.05	42,42,42,42	0
2	MG	E	401	1/1	0.98	0.05	40,40,40,40	0
2	MG	C	401	1/1	0.98	0.04	46,46,46,46	0
2	MG	K	401	1/1	0.98	0.11	46,46,46,46	0
2	MG	H	401	1/1	0.98	0.09	43,43,43,43	0
2	MG	G	401	1/1	1.00	0.03	37,37,37,37	0

6.5 Other polymers [i](#)

There are no such residues in this entry.