



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 24, 2020 – 05:51 am BST

PDB ID : 4E4G  
Title : Crystal structure of putative Methylmalonate-semialdehyde dehydrogenase from *Sinorhizobium meliloti* 1021  
Authors : Malashkevich, V.N.; Bhosle, R.; Toro, R.; Hillerich, B.; Gizzi, A.; Garforth, S.; Kar, A.; Chan, M.K.; Laffuer, J.; Patel, H.; Matikainen, B.; Chamala, S.; Lim, S.; Celikgil, A.; Villegas, G.; Evans, B.; Zenchek, W.; Love, J.; Fiser, A.; Khafizov, K.; Seidel, R.; Bonanno, J.B.; Almo, S.C.; New York Structural Genomics Research Consortium (NYSGRC)  
Deposited on : 2012-03-12  
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

<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

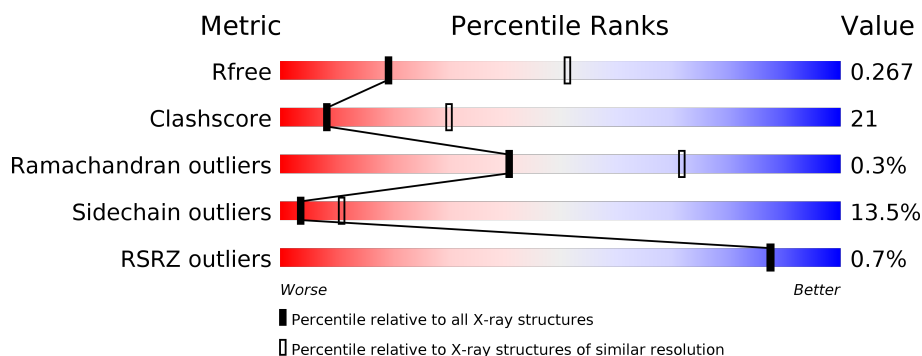
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 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))
$R_{free}$	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (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 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	521	
1	B	521	
1	C	521	
1	D	521	
1	E	521	
1	F	521	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	521	<div><div></div><div>58%30%5%7%</div></div>
1	H	521	<div><div></div><div>56%33%•7%</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 29653 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 Methylmalonate-semialdehyde dehydrogenase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	488	Total	C	N	O	S	Se	0	0	0
			3681	2325	640	693	5	18			
1	B	486	Total	C	N	O	S	Se	0	0	0
			3665	2316	638	689	5	17			
1	C	486	Total	C	N	O	S	Se	0	0	0
			3664	2315	637	689	5	18			
1	D	488	Total	C	N	O	S	Se	0	0	0
			3681	2325	640	693	5	18			
1	E	486	Total	C	N	O	S	Se	0	0	0
			3665	2316	638	689	5	17			
1	F	484	Total	C	N	O	S	Se	0	0	0
			3644	2301	635	686	5	17			
1	G	485	Total	C	N	O	S	Se	0	0	0
			3660	2312	636	688	5	19			
1	H	485	Total	C	N	O	S	Se	0	0	0
			3653	2307	637	687	5	17			

There are 184 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-24	MSE	-	EXPRESSION TAG	UNP Q92RW4
A	-23	HIS	-	EXPRESSION TAG	UNP Q92RW4
A	-22	HIS	-	EXPRESSION TAG	UNP Q92RW4
A	-21	HIS	-	EXPRESSION TAG	UNP Q92RW4
A	-20	HIS	-	EXPRESSION TAG	UNP Q92RW4
A	-19	HIS	-	EXPRESSION TAG	UNP Q92RW4
A	-18	HIS	-	EXPRESSION TAG	UNP Q92RW4
A	-17	SER	-	EXPRESSION TAG	UNP Q92RW4
A	-16	SER	-	EXPRESSION TAG	UNP Q92RW4
A	-15	GLY	-	EXPRESSION TAG	UNP Q92RW4
A	-14	VAL	-	EXPRESSION TAG	UNP Q92RW4
A	-13	ASP	-	EXPRESSION TAG	UNP Q92RW4
A	-12	LEU	-	EXPRESSION TAG	UNP Q92RW4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	GLY	-	EXPRESSION TAG	UNP Q92RW4
A	-10	THR	-	EXPRESSION TAG	UNP Q92RW4
A	-9	GLU	-	EXPRESSION TAG	UNP Q92RW4
A	-8	ASN	-	EXPRESSION TAG	UNP Q92RW4
A	-7	LEU	-	EXPRESSION TAG	UNP Q92RW4
A	-6	TYR	-	EXPRESSION TAG	UNP Q92RW4
A	-5	PHE	-	EXPRESSION TAG	UNP Q92RW4
A	-4	GLN	-	EXPRESSION TAG	UNP Q92RW4
A	-3	SER	-	EXPRESSION TAG	UNP Q92RW4
A	-2	MSE	-	EXPRESSION TAG	UNP Q92RW4
B	-24	MSE	-	EXPRESSION TAG	UNP Q92RW4
B	-23	HIS	-	EXPRESSION TAG	UNP Q92RW4
B	-22	HIS	-	EXPRESSION TAG	UNP Q92RW4
B	-21	HIS	-	EXPRESSION TAG	UNP Q92RW4
B	-20	HIS	-	EXPRESSION TAG	UNP Q92RW4
B	-19	HIS	-	EXPRESSION TAG	UNP Q92RW4
B	-18	HIS	-	EXPRESSION TAG	UNP Q92RW4
B	-17	SER	-	EXPRESSION TAG	UNP Q92RW4
B	-16	SER	-	EXPRESSION TAG	UNP Q92RW4
B	-15	GLY	-	EXPRESSION TAG	UNP Q92RW4
B	-14	VAL	-	EXPRESSION TAG	UNP Q92RW4
B	-13	ASP	-	EXPRESSION TAG	UNP Q92RW4
B	-12	LEU	-	EXPRESSION TAG	UNP Q92RW4
B	-11	GLY	-	EXPRESSION TAG	UNP Q92RW4
B	-10	THR	-	EXPRESSION TAG	UNP Q92RW4
B	-9	GLU	-	EXPRESSION TAG	UNP Q92RW4
B	-8	ASN	-	EXPRESSION TAG	UNP Q92RW4
B	-7	LEU	-	EXPRESSION TAG	UNP Q92RW4
B	-6	TYR	-	EXPRESSION TAG	UNP Q92RW4
B	-5	PHE	-	EXPRESSION TAG	UNP Q92RW4
B	-4	GLN	-	EXPRESSION TAG	UNP Q92RW4
B	-3	SER	-	EXPRESSION TAG	UNP Q92RW4
B	-2	MSE	-	EXPRESSION TAG	UNP Q92RW4
C	-24	MSE	-	EXPRESSION TAG	UNP Q92RW4
C	-23	HIS	-	EXPRESSION TAG	UNP Q92RW4
C	-22	HIS	-	EXPRESSION TAG	UNP Q92RW4
C	-21	HIS	-	EXPRESSION TAG	UNP Q92RW4
C	-20	HIS	-	EXPRESSION TAG	UNP Q92RW4
C	-19	HIS	-	EXPRESSION TAG	UNP Q92RW4
C	-18	HIS	-	EXPRESSION TAG	UNP Q92RW4
C	-17	SER	-	EXPRESSION TAG	UNP Q92RW4
C	-16	SER	-	EXPRESSION TAG	UNP Q92RW4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-15	GLY	-	EXPRESSION TAG	UNP Q92RW4
C	-14	VAL	-	EXPRESSION TAG	UNP Q92RW4
C	-13	ASP	-	EXPRESSION TAG	UNP Q92RW4
C	-12	LEU	-	EXPRESSION TAG	UNP Q92RW4
C	-11	GLY	-	EXPRESSION TAG	UNP Q92RW4
C	-10	THR	-	EXPRESSION TAG	UNP Q92RW4
C	-9	GLU	-	EXPRESSION TAG	UNP Q92RW4
C	-8	ASN	-	EXPRESSION TAG	UNP Q92RW4
C	-7	LEU	-	EXPRESSION TAG	UNP Q92RW4
C	-6	TYR	-	EXPRESSION TAG	UNP Q92RW4
C	-5	PHE	-	EXPRESSION TAG	UNP Q92RW4
C	-4	GLN	-	EXPRESSION TAG	UNP Q92RW4
C	-3	SER	-	EXPRESSION TAG	UNP Q92RW4
C	-2	MSE	-	EXPRESSION TAG	UNP Q92RW4
D	-24	MSE	-	EXPRESSION TAG	UNP Q92RW4
D	-23	HIS	-	EXPRESSION TAG	UNP Q92RW4
D	-22	HIS	-	EXPRESSION TAG	UNP Q92RW4
D	-21	HIS	-	EXPRESSION TAG	UNP Q92RW4
D	-20	HIS	-	EXPRESSION TAG	UNP Q92RW4
D	-19	HIS	-	EXPRESSION TAG	UNP Q92RW4
D	-18	HIS	-	EXPRESSION TAG	UNP Q92RW4
D	-17	SER	-	EXPRESSION TAG	UNP Q92RW4
D	-16	SER	-	EXPRESSION TAG	UNP Q92RW4
D	-15	GLY	-	EXPRESSION TAG	UNP Q92RW4
D	-14	VAL	-	EXPRESSION TAG	UNP Q92RW4
D	-13	ASP	-	EXPRESSION TAG	UNP Q92RW4
D	-12	LEU	-	EXPRESSION TAG	UNP Q92RW4
D	-11	GLY	-	EXPRESSION TAG	UNP Q92RW4
D	-10	THR	-	EXPRESSION TAG	UNP Q92RW4
D	-9	GLU	-	EXPRESSION TAG	UNP Q92RW4
D	-8	ASN	-	EXPRESSION TAG	UNP Q92RW4
D	-7	LEU	-	EXPRESSION TAG	UNP Q92RW4
D	-6	TYR	-	EXPRESSION TAG	UNP Q92RW4
D	-5	PHE	-	EXPRESSION TAG	UNP Q92RW4
D	-4	GLN	-	EXPRESSION TAG	UNP Q92RW4
D	-3	SER	-	EXPRESSION TAG	UNP Q92RW4
D	-2	MSE	-	EXPRESSION TAG	UNP Q92RW4
E	-24	MSE	-	EXPRESSION TAG	UNP Q92RW4
E	-23	HIS	-	EXPRESSION TAG	UNP Q92RW4
E	-22	HIS	-	EXPRESSION TAG	UNP Q92RW4
E	-21	HIS	-	EXPRESSION TAG	UNP Q92RW4
E	-20	HIS	-	EXPRESSION TAG	UNP Q92RW4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-19	HIS	-	EXPRESSION TAG	UNP Q92RW4
E	-18	HIS	-	EXPRESSION TAG	UNP Q92RW4
E	-17	SER	-	EXPRESSION TAG	UNP Q92RW4
E	-16	SER	-	EXPRESSION TAG	UNP Q92RW4
E	-15	GLY	-	EXPRESSION TAG	UNP Q92RW4
E	-14	VAL	-	EXPRESSION TAG	UNP Q92RW4
E	-13	ASP	-	EXPRESSION TAG	UNP Q92RW4
E	-12	LEU	-	EXPRESSION TAG	UNP Q92RW4
E	-11	GLY	-	EXPRESSION TAG	UNP Q92RW4
E	-10	THR	-	EXPRESSION TAG	UNP Q92RW4
E	-9	GLU	-	EXPRESSION TAG	UNP Q92RW4
E	-8	ASN	-	EXPRESSION TAG	UNP Q92RW4
E	-7	LEU	-	EXPRESSION TAG	UNP Q92RW4
E	-6	TYR	-	EXPRESSION TAG	UNP Q92RW4
E	-5	PHE	-	EXPRESSION TAG	UNP Q92RW4
E	-4	GLN	-	EXPRESSION TAG	UNP Q92RW4
E	-3	SER	-	EXPRESSION TAG	UNP Q92RW4
E	-2	MSE	-	EXPRESSION TAG	UNP Q92RW4
F	-24	MSE	-	EXPRESSION TAG	UNP Q92RW4
F	-23	HIS	-	EXPRESSION TAG	UNP Q92RW4
F	-22	HIS	-	EXPRESSION TAG	UNP Q92RW4
F	-21	HIS	-	EXPRESSION TAG	UNP Q92RW4
F	-20	HIS	-	EXPRESSION TAG	UNP Q92RW4
F	-19	HIS	-	EXPRESSION TAG	UNP Q92RW4
F	-18	HIS	-	EXPRESSION TAG	UNP Q92RW4
F	-17	SER	-	EXPRESSION TAG	UNP Q92RW4
F	-16	SER	-	EXPRESSION TAG	UNP Q92RW4
F	-15	GLY	-	EXPRESSION TAG	UNP Q92RW4
F	-14	VAL	-	EXPRESSION TAG	UNP Q92RW4
F	-13	ASP	-	EXPRESSION TAG	UNP Q92RW4
F	-12	LEU	-	EXPRESSION TAG	UNP Q92RW4
F	-11	GLY	-	EXPRESSION TAG	UNP Q92RW4
F	-10	THR	-	EXPRESSION TAG	UNP Q92RW4
F	-9	GLU	-	EXPRESSION TAG	UNP Q92RW4
F	-8	ASN	-	EXPRESSION TAG	UNP Q92RW4
F	-7	LEU	-	EXPRESSION TAG	UNP Q92RW4
F	-6	TYR	-	EXPRESSION TAG	UNP Q92RW4
F	-5	PHE	-	EXPRESSION TAG	UNP Q92RW4
F	-4	GLN	-	EXPRESSION TAG	UNP Q92RW4
F	-3	SER	-	EXPRESSION TAG	UNP Q92RW4
F	-2	MSE	-	EXPRESSION TAG	UNP Q92RW4
G	-24	MSE	-	EXPRESSION TAG	UNP Q92RW4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	-23	HIS	-	EXPRESSION TAG	UNP Q92RW4
G	-22	HIS	-	EXPRESSION TAG	UNP Q92RW4
G	-21	HIS	-	EXPRESSION TAG	UNP Q92RW4
G	-20	HIS	-	EXPRESSION TAG	UNP Q92RW4
G	-19	HIS	-	EXPRESSION TAG	UNP Q92RW4
G	-18	HIS	-	EXPRESSION TAG	UNP Q92RW4
G	-17	SER	-	EXPRESSION TAG	UNP Q92RW4
G	-16	SER	-	EXPRESSION TAG	UNP Q92RW4
G	-15	GLY	-	EXPRESSION TAG	UNP Q92RW4
G	-14	VAL	-	EXPRESSION TAG	UNP Q92RW4
G	-13	ASP	-	EXPRESSION TAG	UNP Q92RW4
G	-12	LEU	-	EXPRESSION TAG	UNP Q92RW4
G	-11	GLY	-	EXPRESSION TAG	UNP Q92RW4
G	-10	THR	-	EXPRESSION TAG	UNP Q92RW4
G	-9	GLU	-	EXPRESSION TAG	UNP Q92RW4
G	-8	ASN	-	EXPRESSION TAG	UNP Q92RW4
G	-7	LEU	-	EXPRESSION TAG	UNP Q92RW4
G	-6	TYR	-	EXPRESSION TAG	UNP Q92RW4
G	-5	PHE	-	EXPRESSION TAG	UNP Q92RW4
G	-4	GLN	-	EXPRESSION TAG	UNP Q92RW4
G	-3	SER	-	EXPRESSION TAG	UNP Q92RW4
G	-2	MSE	-	EXPRESSION TAG	UNP Q92RW4
H	-24	MSE	-	EXPRESSION TAG	UNP Q92RW4
H	-23	HIS	-	EXPRESSION TAG	UNP Q92RW4
H	-22	HIS	-	EXPRESSION TAG	UNP Q92RW4
H	-21	HIS	-	EXPRESSION TAG	UNP Q92RW4
H	-20	HIS	-	EXPRESSION TAG	UNP Q92RW4
H	-19	HIS	-	EXPRESSION TAG	UNP Q92RW4
H	-18	HIS	-	EXPRESSION TAG	UNP Q92RW4
H	-17	SER	-	EXPRESSION TAG	UNP Q92RW4
H	-16	SER	-	EXPRESSION TAG	UNP Q92RW4
H	-15	GLY	-	EXPRESSION TAG	UNP Q92RW4
H	-14	VAL	-	EXPRESSION TAG	UNP Q92RW4
H	-13	ASP	-	EXPRESSION TAG	UNP Q92RW4
H	-12	LEU	-	EXPRESSION TAG	UNP Q92RW4
H	-11	GLY	-	EXPRESSION TAG	UNP Q92RW4
H	-10	THR	-	EXPRESSION TAG	UNP Q92RW4
H	-9	GLU	-	EXPRESSION TAG	UNP Q92RW4
H	-8	ASN	-	EXPRESSION TAG	UNP Q92RW4
H	-7	LEU	-	EXPRESSION TAG	UNP Q92RW4
H	-6	TYR	-	EXPRESSION TAG	UNP Q92RW4
H	-5	PHE	-	EXPRESSION TAG	UNP Q92RW4

*Continued on next page...*



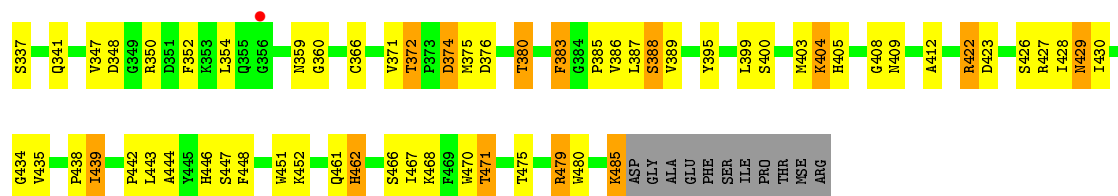
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	-4	GLN	-	EXPRESSION TAG	UNP Q92RW4
H	-3	SER	-	EXPRESSION TAG	UNP Q92RW4
H	-2	MSE	-	EXPRESSION TAG	UNP Q92RW4

- Molecule 2 is water.

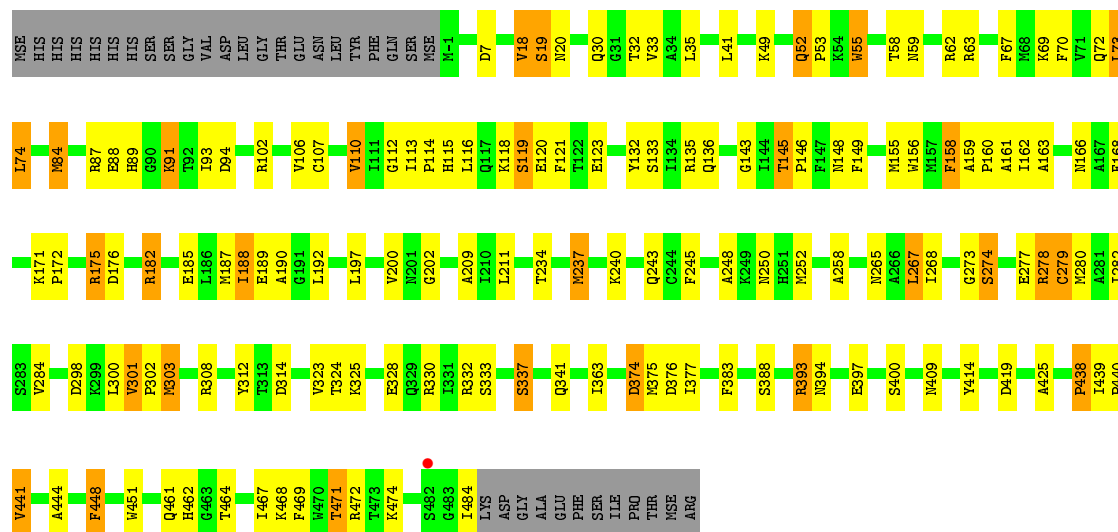
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	58	Total	O	0	0
			58	58		
2	B	44	Total	O	0	0
			44	44		
2	C	51	Total	O	0	0
			51	51		
2	D	35	Total	O	0	0
			35	35		
2	E	33	Total	O	0	0
			33	33		
2	F	35	Total	O	0	0
			35	35		
2	G	42	Total	O	0	0
			42	42		
2	H	42	Total	O	0	0
			42	42		





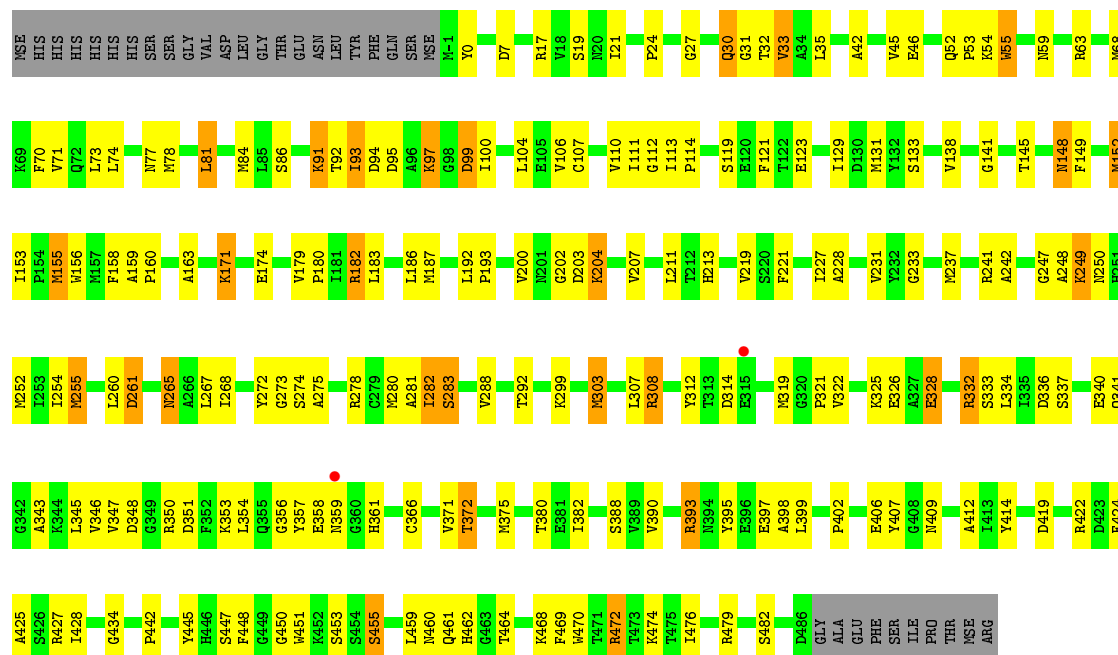
• Molecule 1: Methylmalonate-semialdehyde dehydrogenase

Chain C: 65% 23% 6% 7%

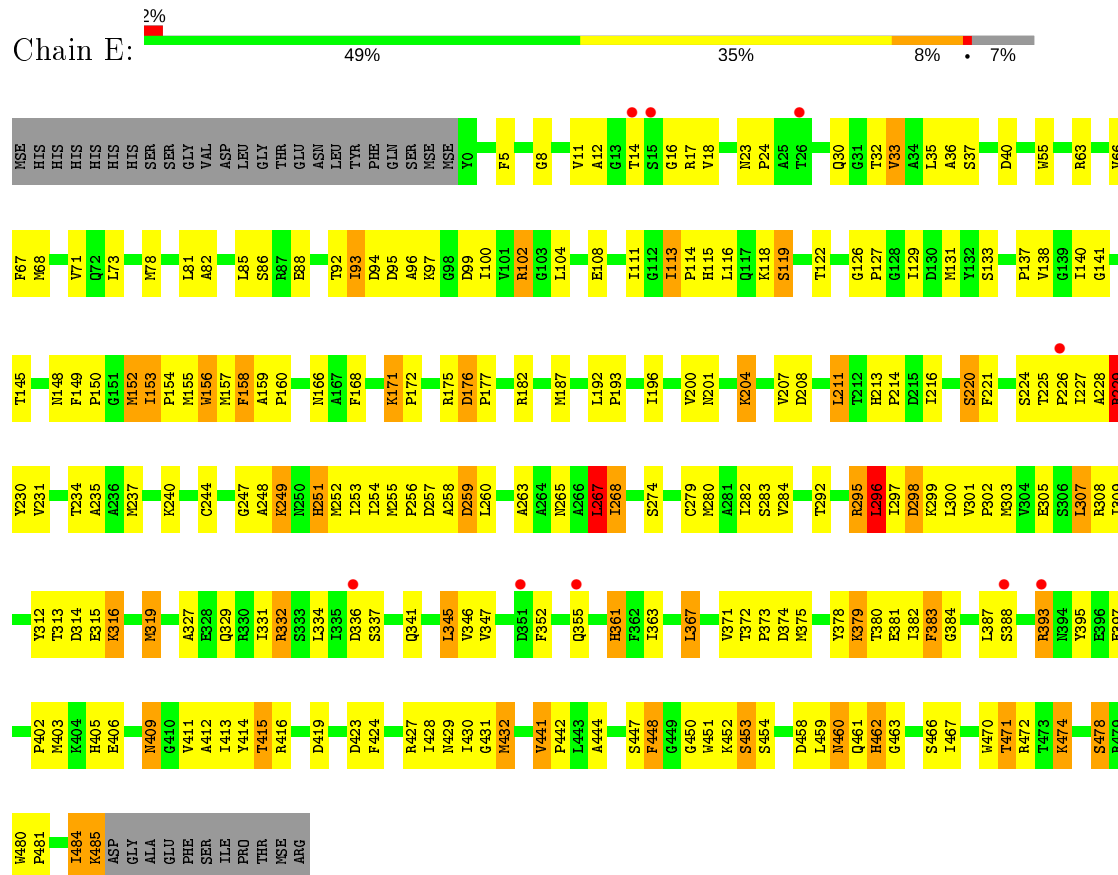


• Molecule 1: Methylmalonate-semialdehyde dehydrogenase

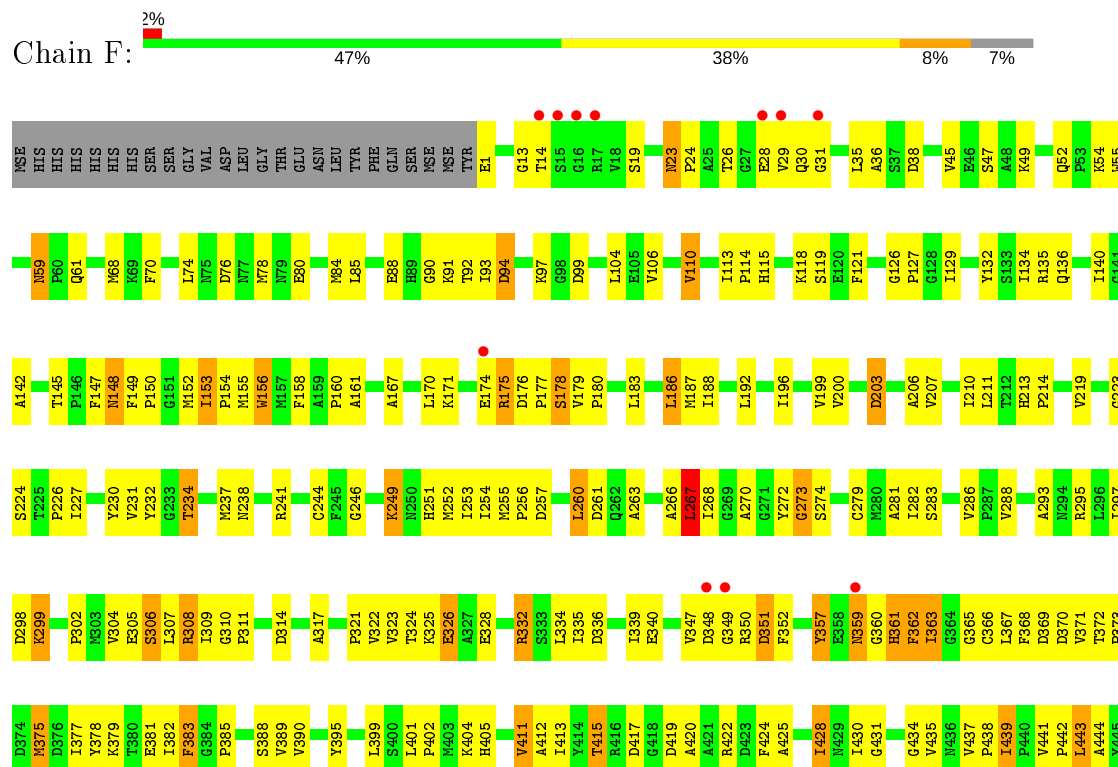
Chain D: 56% 32% 5% 6%



• Molecule 1: Methylmalonate-semialdehyde dehydrogenase



• Molecule 1: Methylmalonate-semialdehyde dehydrogenase





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	197.44Å 171.28Å 159.69Å 90.00° 124.03° 90.00°	Depositor
Resolution (Å)	19.97 – 2.90 19.97 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.6 (19.97-2.90) 98.9 (19.97-2.90)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.06 (at 2.88Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.188 , 0.269 0.188 , 0.267	Depositor DCC
$R_{free}$ test set	4795 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.1	Xtriage
Anisotropy	0.059	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 45.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	29653	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 38.02 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.9244e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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	0.60	1/3742 (0.0%)	0.72	1/5043 (0.0%)
1	B	0.60	2/3726 (0.1%)	0.74	4/5022 (0.1%)
1	C	0.61	1/3725 (0.0%)	0.75	2/5021 (0.0%)
1	D	0.63	1/3742 (0.0%)	0.73	3/5043 (0.1%)
1	E	0.60	1/3726 (0.0%)	0.77	7/5022 (0.1%)
1	F	0.75	8/3704 (0.2%)	0.77	5/4993 (0.1%)
1	G	0.61	2/3720 (0.1%)	0.75	1/5012 (0.0%)
1	H	0.62	2/3713 (0.1%)	0.73	1/5004 (0.0%)
All	All	0.63	18/29798 (0.1%)	0.75	24/40160 (0.1%)

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	357	TYR	CG-CD2	13.39	1.56	1.39
1	F	357	TYR	CE1-CZ	9.63	1.51	1.38
1	F	451	TRP	CD2-CE2	6.00	1.48	1.41
1	C	55	TRP	CD2-CE2	5.79	1.48	1.41
1	A	312	TYR	CE1-CZ	-5.69	1.31	1.38

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	332	ARG	NE-CZ-NH2	-8.76	115.92	120.30
1	A	84	MSE	N-CA-CB	-6.81	98.34	110.60
1	E	259	ASP	CB-CG-OD1	-6.46	112.49	118.30
1	F	349	GLY	N-CA-C	6.34	128.95	113.10
1	F	332	ARG	NE-CZ-NH2	-6.33	117.13	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	3681	0	3652	107	0
1	B	3665	0	3639	161	0
1	C	3664	0	3635	117	0
1	D	3681	0	3652	140	0
1	E	3665	0	3639	225	1
1	F	3644	0	3617	253	1
1	G	3660	0	3630	128	0
1	H	3653	0	3630	148	0
2	A	58	0	0	2	0
2	B	44	0	0	2	0
2	C	51	0	0	3	0
2	D	35	0	0	1	0
2	E	33	0	0	1	0
2	F	35	0	0	2	0
2	G	42	0	0	2	0
2	H	42	0	0	5	0
All	All	29653	0	29094	1224	1

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 21.

The worst 5 of 1224 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:F:70:PHE:CE1	1:F:187:MSE:SE	2.30	1.33
1:H:110:VAL:CG1	1:H:163:ALA:HB2	1.78	1.13
1:F:126:GLY:HA3	1:F:129:ILE:HB	1.28	1.12
1:F:68:MSE:HE1	1:H:71:VAL:HG11	1.14	1.11
1:F:152:MSE:HE2	1:F:153:ILE:HD11	1.29	1.09

All (1) 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:E:298:ASP:O	1:F:299:LYS:NZ[4_446]	2.14	0.06

## 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	486/521 (93%)	457 (94%)	28 (6%)	1 (0%)	47	78
1	B	484/521 (93%)	455 (94%)	28 (6%)	1 (0%)	47	78
1	C	484/521 (93%)	457 (94%)	25 (5%)	2 (0%)	34	66
1	D	486/521 (93%)	457 (94%)	28 (6%)	1 (0%)	47	78
1	E	484/521 (93%)	451 (93%)	31 (6%)	2 (0%)	34	66
1	F	482/521 (92%)	450 (93%)	31 (6%)	1 (0%)	47	78
1	G	483/521 (93%)	456 (94%)	25 (5%)	2 (0%)	34	66
1	H	483/521 (93%)	452 (94%)	28 (6%)	3 (1%)	25	58
All	All	3872/4168 (93%)	3635 (94%)	224 (6%)	13 (0%)	41	71

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	462	HIS
1	A	462	HIS
1	C	462	HIS
1	D	462	HIS
1	E	462	HIS

### 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	383/391 (98%)	347 (91%)	36 (9%)	8	26
1	B	381/391 (97%)	330 (87%)	51 (13%)	4	11
1	C	381/391 (97%)	339 (89%)	42 (11%)	6	19
1	D	383/391 (98%)	327 (85%)	56 (15%)	3	9
1	E	381/391 (97%)	318 (84%)	63 (16%)	2	7
1	F	379/391 (97%)	310 (82%)	69 (18%)	1	5
1	G	381/391 (97%)	328 (86%)	53 (14%)	3	10
1	H	380/391 (97%)	339 (89%)	41 (11%)	6	20
All	All	3049/3128 (98%)	2638 (86%)	411 (14%)	4	11

5 of 411 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	14	THR
1	E	361	HIS
1	H	119	SER
1	E	73	LEU
1	E	211	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	243	GLN
1	D	250	ASN
1	G	359	ASN
1	C	341	GLN
1	C	460	ASN

### 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	470/521 (90%)	-0.62	1 (0%) 95 95	17, 36, 55, 71	0
1	B	469/521 (90%)	-0.54	1 (0%) 95 95	19, 38, 58, 76	0
1	C	468/521 (89%)	-0.70	1 (0%) 95 95	15, 28, 49, 79	0
1	D	470/521 (90%)	-0.48	2 (0%) 92 93	15, 43, 75, 93	0
1	E	469/521 (90%)	-0.10	9 (1%) 66 65	21, 57, 81, 97	0
1	F	467/521 (89%)	-0.04	11 (2%) 59 56	29, 59, 87, 107	0
1	G	466/521 (89%)	-0.63	1 (0%) 95 95	18, 35, 53, 70	0
1	H	468/521 (89%)	-0.54	1 (0%) 95 95	18, 38, 57, 74	0
All	All	3747/4168 (89%)	-0.46	27 (0%) 87 87	15, 40, 73, 107	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	359	ASN	3.4
1	F	29	VAL	3.3
1	H	16	GLY	2.9
1	E	226	PRO	2.9
1	F	17	ARG	2.8

### 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 ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.