



wwPDB X-ray Structure Validation Summary Report ⓘ

May 24, 2020 – 11:25 pm BST

PDB ID : 1GV1
Title : Structural Basis for Thermophilic Protein Stability: Structures of Thermophilic and Mesophilic Malate Dehydrogenases
Authors : Dalhus, B.; Sarinen, M.; Sauer, U.H.; Eklund, P.; Johansson, K.; Karlsson, A.; Ramaswamy, S.; Bjork, A.; Synstad, B.; Naterstad, K.; Sirevag, R.; Eklund, H.
Deposited on : 2002-02-04
Resolution : 2.50 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

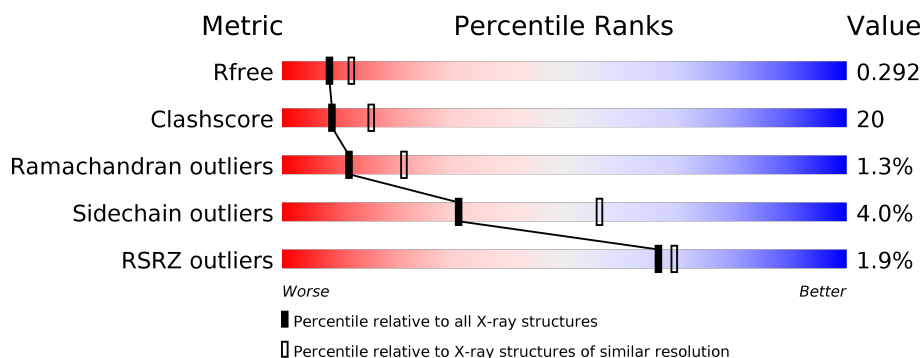
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	310	<div> <div>%</div> <div> <div></div> <div>66%</div> <div>30%</div> <div>• •</div> </div> </div>
1	B	310	<div> <div>3%</div> <div> <div></div> <div>59%</div> <div>32%</div> <div>• 6%</div> </div> </div>
1	C	310	<div> <div>%</div> <div> <div></div> <div>63%</div> <div>32%</div> <div>• •</div> </div> </div>
1	D	310	<div> <div>2%</div> <div> <div></div> <div>60%</div> <div>31%</div> <div>• 6%</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9395 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 MALATE DEHYDROGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	305	Total	C	N	O	S	0	0	0
			2293	1451	391	436	15			
1	B	290	Total	C	N	O	S	0	0	0
			2172	1378	367	415	12			
1	C	305	Total	C	N	O	S	0	0	0
			2293	1451	391	436	15			
1	D	292	Total	C	N	O	S	0	0	0
			2187	1386	370	418	13			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	227	ALA	SER	conflict	UNP P80039
A	229	ALA	GLY	conflict	UNP P80039
B	227	ALA	SER	conflict	UNP P80039
B	229	ALA	GLY	conflict	UNP P80039
C	227	ALA	SER	conflict	UNP P80039
C	229	ALA	GLY	conflict	UNP P80039
D	227	ALA	SER	conflict	UNP P80039
D	229	ALA	GLY	conflict	UNP P80039

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	135	Total	O	0	0
			135	135		
2	B	78	Total	O	0	0
			78	78		
2	C	115	Total	O	0	0
			115	115		
2	D	122	Total	O	0	0
			122	122		

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.

Chain A:

Sequence logo for Chain A. The y-axis represents information content in bits (0.00 to 0.15). The x-axis shows positions 1 to 300. A color scale at the top indicates conservation levels: red (0.00-0.02), yellow (0.02-0.04), green (0.04-0.06), blue (0.06-0.08), and dark blue (0.08-0.15). A legend at the top right shows the color scale for bits. A red dot is at position 100, and a blue dot is at position 200.

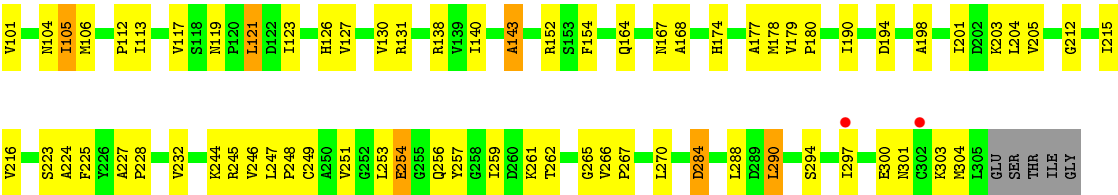
Position	Amino Acid	Information Content (bits)
1	M1	0.02
2	K2	0.02
3	I3	0.02
4	G9	0.02
5	F17	0.02
6	L28	0.02
7	V29	0.02
8	L30	0.02
9	L31	0.02
10	V34	0.02
11	E35	0.02
12	G36	0.02
13	I37	0.02
14	P38	0.02
15	Q39	0.02
16	G40	0.02
17	K41	0.02
18	D44	0.02
19	M45	0.02
20	Y46	0.02
21	E47	0.02
22	T48	0.02
23	G49	0.02
24	G52	0.02
25	L53	0.02
26	G60	0.02
27	S61	0.02
28	M62	0.02
29	D63	0.02
30	Y64	0.02
31	A65	0.02
32	G78	0.02
33	L79	0.02
34	P80	0.02
35	R81	0.02
36	K82	0.02
37	M85	0.02
38	T86	0.02
39	R87	0.02
40	E88	0.02
41	D89	0.02
42	L90	0.02
43	L91	0.02
44	M92	0.02
45	K93	0.02
46	N94	0.02
47	I97	0.02
48	V101	0.02
49	N111	0.02
50	P112	0.02
51	I113	0.02
52	V117	0.02
53	S118	0.02
54	M119	0.02
55	P120	0.02
56	L121	0.02
57	H126	0.02
58	R138	0.02
59	M142	0.02
60	A143	0.02
61	R150	0.02
62	F151	0.02
63	R152	0.02
64	S153	0.02
65	F154	0.02
66	M157	0.02
67	I166	0.02
68	L171	0.02
69	H174	0.02
70	A177	0.02
71	M178	0.02
72	V179	0.02
73	P180	0.02
74	V181	0.02
75	V182	0.02
76	V187	0.02
77	P191	0.02
78	D194	0.02
79	L195	0.02
80	L196	0.02
81	P197	0.02
82	L204	0.02
83	V205	0.02
84	I215	0.02
85	L219	0.02
86	K220	0.02
87	Q221	0.02
88	F225	0.02
89	Y226	0.02
90	A227	0.02
91	P229	0.02
92	V232	0.02
93	V236	0.02
94	V240	0.02
95	L241	0.02
96	D242	0.02
97	R245	0.02
98	V246	0.02
99	L247	0.02
100	L253	0.02
101	Q256	0.02
102	Y257	0.02
103	G258	0.02
104	L259	0.02
105	V266	0.02
106	P267	0.02
107	V268	0.02
108	R272	0.02
109	E280	0.02
110	L281	0.02
111	M282	0.02
112	L283	0.02
113	D284	0.02
114	Q285	0.02
115	L290	0.02
116	S294	0.02
117	V298	0.02
118	D299	0.02
119	E300	0.02
120	M301	0.02
121	C302	0.02
122	K303	0.02
123	M304	0.02
124	L305	0.02
125	GLU	0.02
126	SER	0.02
127	THR	0.02
128	ILE	0.02
129	GLY	0.02

Chain B:

Chain C:

63% 32%

M1 K2 I3 A16 F17 R18 I19 K22 A25 R26 E27 L28 V29 V33 V34 E35 G36 I37 P38 Q39 L43 D44 M45 Y46 E47 I48 Q49 P50 I53 T59 G60 S61 M62 D63 Y64 A65 D66 L79 P80 R81 K82 P83 G84 M85 P88 E89 L90 L91 M92 K93 P100



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	64.42Å 85.82Å 117.50Å 90.00° 104.61° 90.00°	Depositor
Resolution (Å)	19.92 – 2.50 19.97 – 2.50	Depositor EDS
% Data completeness (in resolution range)	88.0 (19.92-2.50) 88.1 (19.97-2.50)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.75 (at 2.50Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.216 , 0.305 0.208 , 0.292	Depositor DCC
R_{free} test set	3815 reflections (9.60%)	wwPDB-VP
Wilson B-factor (Å ²)	24.9	Xtriage
Anisotropy	0.794	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 54.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.024 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	9395	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% 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 [i](#)

5.1 Standard geometry [i](#)

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	0.36	0/2323	0.64	1/3144 (0.0%)
1	B	0.34	0/2199	0.60	0/2978
1	C	0.36	0/2323	0.64	0/3144
1	D	0.37	0/2214	0.62	0/2998
All	All	0.36	0/9059	0.63	1/12264 (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	A	113	ILE	N-CA-C	-5.12	97.19	111.00

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	2293	0	2381	94	0
1	B	2172	0	2252	98	0
1	C	2293	0	2381	91	0
1	D	2187	0	2265	101	0
2	A	135	0	0	19	0
2	B	78	0	0	9	0
2	C	115	0	0	11	0
2	D	122	0	0	11	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	9395	0	9279	371	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 20.

The worst 5 of 371 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:86:THR:HB	1:A:89:ASP:HB2	1.32	1.07
1:B:90:LEU:HA	1:B:93:LYS:HD3	1.48	0.95
1:C:254:GLU:H	1:C:256:GLN:HE22	1.02	0.94
1:D:121:LEU:HD13	1:D:143:ALA:HB2	1.47	0.92
1:D:220:LYS:HG2	1:D:221:GLN:HE21	1.37	0.89

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	303/310 (98%)	274 (90%)	25 (8%)	4 (1%)	12	21
1	B	286/310 (92%)	261 (91%)	18 (6%)	7 (2%)	6	9
1	C	303/310 (98%)	277 (91%)	23 (8%)	3 (1%)	15	28
1	D	288/310 (93%)	265 (92%)	22 (8%)	1 (0%)	41	61
All	All	1180/1240 (95%)	1077 (91%)	88 (8%)	15 (1%)	12	21

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	285	GLN
1	A	80	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	143	ALA
1	B	225	PHE
1	C	121	LEU

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	251/255 (98%)	245 (98%)	6 (2%)	49	74
1	B	237/255 (93%)	229 (97%)	8 (3%)	37	63
1	C	251/255 (98%)	239 (95%)	12 (5%)	25	48
1	D	239/255 (94%)	226 (95%)	13 (5%)	22	42
All	All	978/1020 (96%)	939 (96%)	39 (4%)	31	56

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	83	PRO
1	C	223	SER
1	D	270	LEU
1	C	100	GLU
1	C	105	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	111	ASN
1	C	301	ASN
1	D	282	ASN
1	C	126	HIS
1	C	256	GLN

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](#)

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 [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, 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	305/310 (98%)	-0.29	3 (0%) 82 84	7, 24, 49, 71	0
1	B	290/310 (93%)	-0.11	10 (3%) 45 48	10, 32, 53, 91	0
1	C	305/310 (98%)	-0.26	4 (1%) 77 79	7, 23, 51, 76	0
1	D	292/310 (94%)	-0.22	6 (2%) 63 66	7, 24, 50, 95	0
All	All	1192/1240 (96%)	-0.22	23 (1%) 66 69	7, 25, 52, 95	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	297	ILE	5.7
1	B	298	VAL	3.8
1	C	84	GLY	3.2
1	C	302	CYS	3.2
1	A	78	GLY	3.2

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.