wwPDB X-ray Structure Validation Summary Report (i)

May 24, 2020 – 06:38 am BST

PDB ID	:	1VL6
Title	:	Crystal structure of NAD-dependent malic enzyme (TM0542) from Thermotoga maritima at 2.61 A resolution
Authors	:	Joint Center for Structural Genomics (JCSG)
Deposited on	:	2004-07-13
Resolution	:	2.61 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

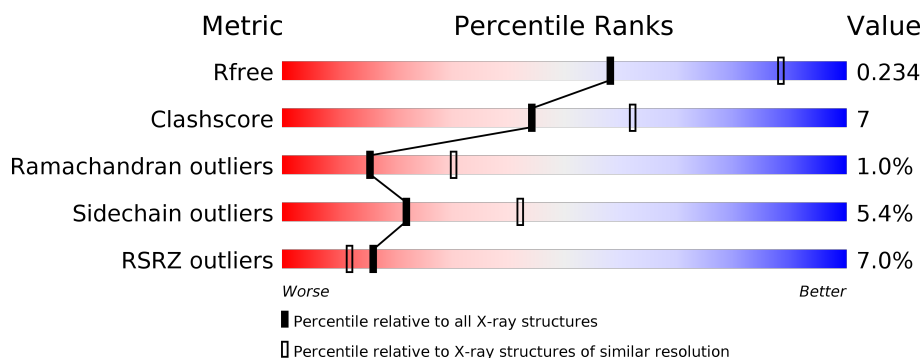|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 2.61 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	3797 (2.64-2.60)
Clashscore	141614	4168 (2.64-2.60)
Ramachandran outliers	138981	4093 (2.64-2.60)
Sidechain outliers	138945	4093 (2.64-2.60)
RSRZ outliers	127900	3731 (2.64-2.60)

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 <=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	388	
1	B	388	
1	C	388	
1	D	388	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 11180 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 oxidoreductase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	377	Total	C	N	O	S	Se	0	2	0
			2879	1846	481	541	5	6			
1	B	355	Total	C	N	O	S	Se	0	1	0
			2600	1664	434	492	5	5			
1	C	377	Total	C	N	O	S	Se	0	0	0
			2813	1796	468	538	5	6			
1	D	377	Total	C	N	O	S	Se	0	1	0
			2844	1824	470	539	5	6			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MSE	-	LEADER SEQUENCE	UNP Q9WZ12
A	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZ12
A	-9	SER	-	LEADER SEQUENCE	UNP Q9WZ12
A	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZ12
A	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZ12
A	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZ12
A	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
A	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
A	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
A	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
A	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
A	0	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
A	1	VAL	MET	SEE REMARK 999	UNP Q9WZ12
A	6	VAL	ILE	SEE REMARK 999	UNP Q9WZ12
A	84	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
A	147	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
A	272	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
A	322	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
A	336	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
A	362	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
B	-11	MSE	-	LEADER SEQUENCE	UNP Q9WZ12

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZ12
B	-9	SER	-	LEADER SEQUENCE	UNP Q9WZ12
B	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZ12
B	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZ12
B	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZ12
B	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
B	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
B	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
B	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
B	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
B	0	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
B	1	VAL	MET	SEE REMARK 999	UNP Q9WZ12
B	6	VAL	ILE	SEE REMARK 999	UNP Q9WZ12
B	84	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
B	147	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
B	272	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
B	322	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
B	336	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
B	362	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
C	-11	MSE	-	LEADER SEQUENCE	UNP Q9WZ12
C	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZ12
C	-9	SER	-	LEADER SEQUENCE	UNP Q9WZ12
C	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZ12
C	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZ12
C	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZ12
C	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
C	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
C	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
C	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
C	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
C	0	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
C	1	VAL	MET	SEE REMARK 999	UNP Q9WZ12
C	6	VAL	ILE	SEE REMARK 999	UNP Q9WZ12
C	84	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
C	147	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
C	272	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
C	322	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
C	336	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
C	362	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
D	-11	MSE	-	LEADER SEQUENCE	UNP Q9WZ12
D	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZ12
D	-9	SER	-	LEADER SEQUENCE	UNP Q9WZ12

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZ12
D	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZ12
D	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZ12
D	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
D	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
D	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
D	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
D	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
D	0	HIS	-	LEADER SEQUENCE	UNP Q9WZ12
D	1	VAL	MET	SEE REMARK 999	UNP Q9WZ12
D	6	VAL	ILE	SEE REMARK 999	UNP Q9WZ12
D	84	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
D	147	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
D	272	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
D	322	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
D	336	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12
D	362	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ12

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	29	Total 29	O 29	0	0
2	B	1	Total 1	O 1	0	0
2	C	2	Total 2	O 2	0	0
2	D	12	Total 12	O 12	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 (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: malate oxidoreductase

- Molecule 1: malate oxidoreductase

- Molecule 1: malate oxidoreductase

F100 P101 S114 E120 P121 I126 V151 D154 G158 V162 L168 L171 T174 A191 L202 D211 D220 C224 I232 T236 R240 L241 S242 G243 D244 L245 E246 T247 A248 L249 A252 V258 P266 I269 M272 S273 R274

K275 P276 V277 I278 N283 I288 A293 R294 E295 A296 G297 A298 F299 I300 V301 R305 S306 D307 N310 N313 N314 L315 G320 S330 N335 M336 L337 L338 S339 A340 V341 P352 E353 I356 P357 M362 K363 V364 A371 A376

• Molecule 1: malate oxidoreductase

3%

Chain D:
84% 11% • •

MSE GLY SER ASP LYS ILE HIS HIS HIS HIS HIS H0 V1 D2 E5 V6 H7 K13 I14 R15 T16 A17 L18 P19 V20 L28 Y51 S55 T59 G79 P82 E85 G86 K87 D96 I97 D98 E109 E120 P121 D130 K135 R138 I139 L140

S144 D154 D155 T159 V163 L173 T174 E175 K176 I188 A209 N217 E218 N219 D220 P221 E222 T223 C224 L225 I232 A233 R234 I235 T236 R240 L241 T247 D253 R260 L264 I269 M272 I278 K328 R329 L366 K373 G374 S375

A376

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	143.96Å 143.96Å 163.43Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.12 – 2.61 47.12 – 2.61	Depositor EDS
% Data completeness (in resolution range)	98.5 (47.12-2.61) 98.5 (47.12-2.61)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.45 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.2.0003	Depositor
R, R_{free}	0.192 , 0.239 0.201 , 0.234	Depositor DCC
R_{free} test set	2919 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	60.8	Xtriage
Anisotropy	0.185	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 67.7	EDS
L-test for twinning[2]	$<|L|> = 0.53, <L^2> = 0.37$	Xtriage
Estimated twinning fraction	0.022 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11180	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.29% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $<|L|>$, $<L^2>$ 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	0.87	0/2926	0.93	12/3964 (0.3%)
1	B	0.77	1/2640 (0.0%)	0.89	4/3577 (0.1%)
1	C	0.70	1/2860 (0.0%)	0.86	7/3883 (0.2%)
1	D	0.80	0/2890	0.88	6/3917 (0.2%)
All	All	0.79	2/11316 (0.0%)	0.89	29/15341 (0.2%)

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	B	0	2
1	D	0	1
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	253	ASP	CG-OD2	5.88	1.38	1.25
1	C	51	TYR	CD2-CE2	5.50	1.47	1.39

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	155	ASP	CB-CG-OD2	7.85	125.37	118.30
1	A	155	ASP	CB-CG-OD1	-7.66	111.41	118.30
1	C	211	ASP	CB-CG-OD2	7.02	124.62	118.30
1	C	224	CYS	CA-CB-SG	-6.94	101.51	114.00
1	A	307	ASP	CB-CG-OD2	6.58	124.22	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	258	VAL	Peptide
1	B	283	ASN	Peptide
1	D	220	ASP	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	2879	0	2907	40	0
1	B	2600	0	2550	58	0
1	C	2813	0	2779	34	0
1	D	2844	0	2850	29	0
2	A	29	0	0	0	0
2	B	1	0	0	0	0
2	C	2	0	0	0	0
2	D	12	0	0	0	0
All	All	11180	0	11086	146	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 7.

The worst 5 of 146 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:B:362:MSE:CE	1:B:362:MSE:SE	2.14	1.46
1:A:100[A]:PHE:CE2	1:B:100[A]:PHE:CD1	2.50	0.99
1:A:100[A]:PHE:CE2	1:B:100[A]:PHE:CE1	2.51	0.98
1:A:100[A]:PHE:CD2	1:B:100[A]:PHE:CD1	2.58	0.92
1:C:272:MSE:HE1	1:C:278:ILE:HD11	1.52	0.91

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	377/388 (97%)	361 (96%)	16 (4%)	0	100	100
1	B	346/388 (89%)	311 (90%)	23 (7%)	12 (4%)	3	4
1	C	375/388 (97%)	357 (95%)	18 (5%)	0	100	100
1	D	376/388 (97%)	356 (95%)	18 (5%)	2 (0%)	29	50
All	All	1474/1552 (95%)	1385 (94%)	75 (5%)	14 (1%)	15	33

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	267	GLU
1	B	284	PRO
1	B	238	PRO
1	B	240	ARG
1	B	295	GLU

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	304/316 (96%)	288 (95%)	16 (5%)	22	43
1	B	265/316 (84%)	248 (94%)	17 (6%)	17	34
1	C	294/316 (93%)	278 (95%)	16 (5%)	22	42
1	D	297/316 (94%)	283 (95%)	14 (5%)	26	49
All	All	1160/1264 (92%)	1097 (95%)	63 (5%)	22	42

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	283	ASN
1	C	151	VAL
1	D	234	ARG
1	B	285	VAL
1	B	366	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	308	HIS
1	A	313	ASN
1	A	335	ASN
1	B	186	ASN
1	C	194	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)

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	371/388 (95%)	-0.05	3 (0%)	86	84	56, 64, 75, 86	0
1	B	350/388 (90%)	0.64	50 (14%)	2	1	41, 63, 73, 82	0
1	C	371/388 (95%)	0.57	37 (9%)	7	4	54, 63, 72, 90	0
1	D	371/388 (95%)	0.05	13 (3%)	44	37	55, 64, 75, 95	0
All	All	1463/1552 (94%)	0.30	103 (7%)	16	12	41, 64, 75, 95	0

The worst 5 of 103 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	243	GLY	9.8
1	C	0	HIS	7.3
1	B	300	ILE	6.9
1	B	244	ASP	5.8
1	C	269	ILE	5.8

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.