wwPDB X-ray Structure Validation Summary Report (i)

Nov 14, 2023 – 11:22 AM JST

PDB ID	:	5Z1Z
Title	:	The apo-structure of D-lactate dehydrogenase from Escherichia coli
Authors	:	Furukawa, N.; Miyanaga, A.; Nakajima, M.; Taguchi, H.
Deposited on	:	2017-12-28
Resolution	:	1.97 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

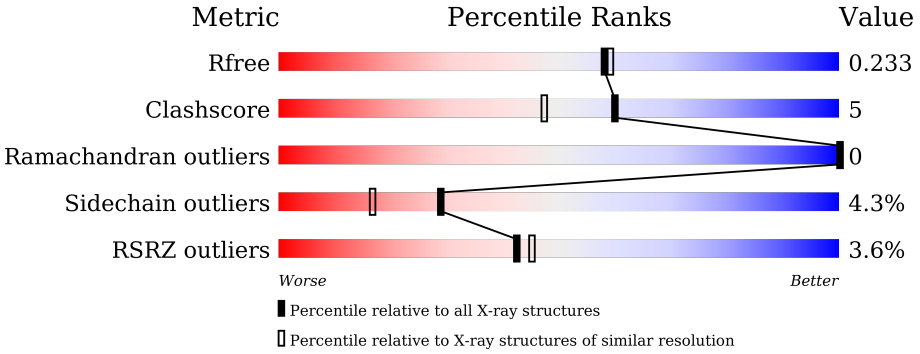
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.97 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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 of 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	329	2%, 87%, 8% • •
1	B	329	4%, 88%, 7% • 5%
1	C	329	%, 85%, 9% • •
1	D	329	6%, 85%, 8% • 5%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9901 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.

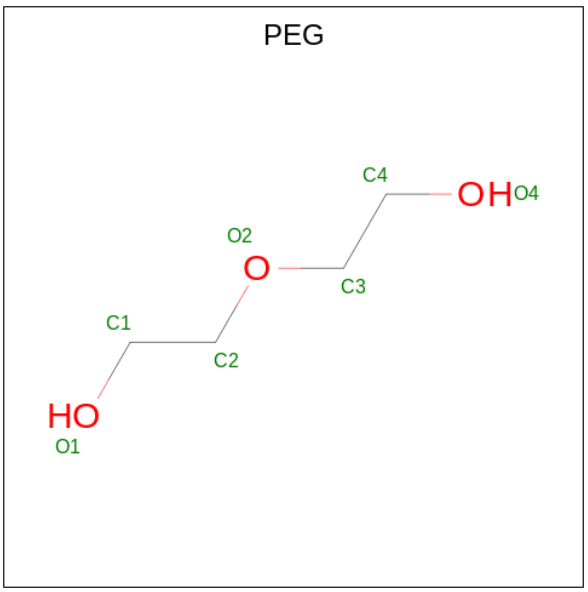
- Molecule 1 is a protein called D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	315	2446	1550	416	464	16	0	0	0
1	B	314	2439	1545	415	463	16	0	0	0
1	C	316	2454	1554	417	467	16	0	0	0
1	D	311	2406	1524	406	460	16	0	0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
2	A	1	1	1	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	7	4	3	0	0

- Molecule 4 is water.

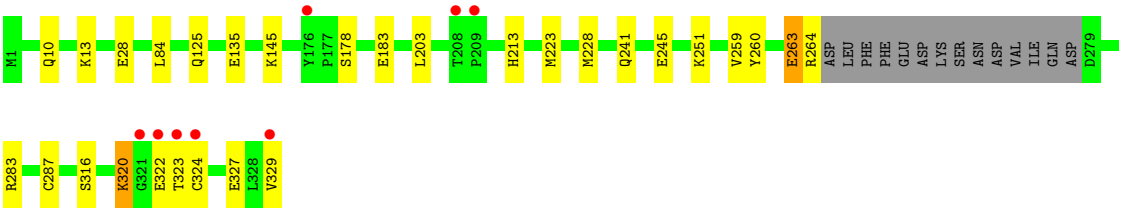
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	47	47	47	0	0
			Total	O		
4	B	41	41	41	0	0
			Total	O		
4	C	31	31	31	0	0
			Total	O		
4	D	29	29	29	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

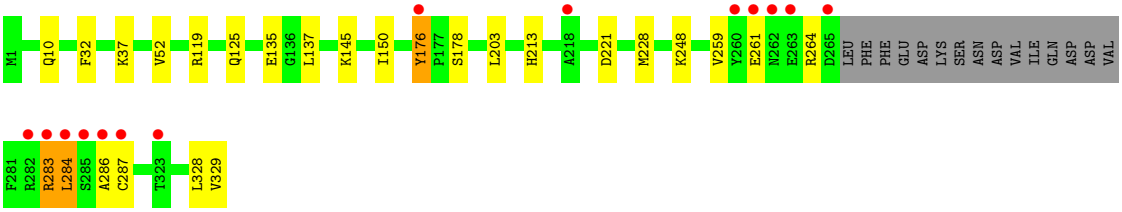
• Molecule 1: D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding

Chain A:



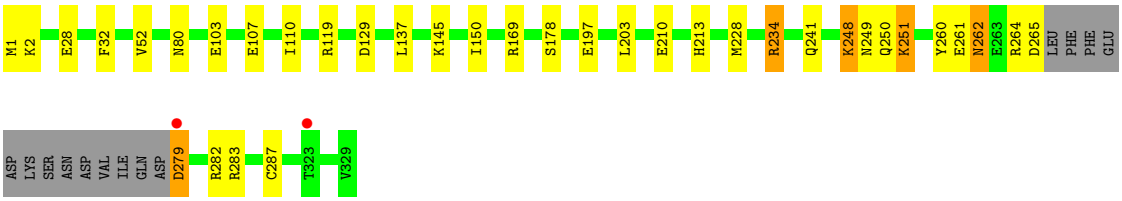
• Molecule 1: D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding

Chain B:



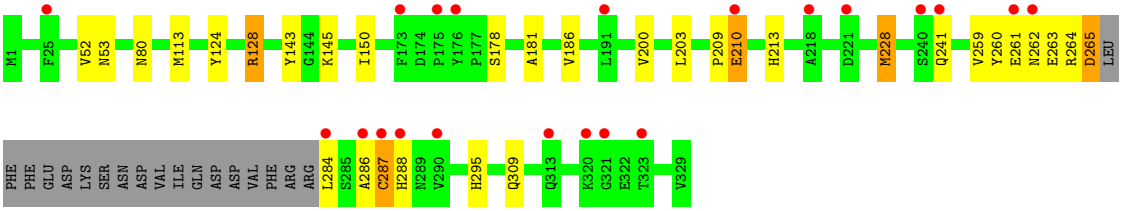
• Molecule 1: D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding

Chain C:



• Molecule 1: D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding

Chain D:


M1
F25
V52
N53
N80
M113
Y124
R128
Y143
G144
K145
I150
F173
D174
P175
Y176
P177
S178
A181
V186
L191
V200
L203
P209
E210
H213
A218
D221
M228
S240
Q241
V259
Y260
E261
N262
E263
R264
D265
LEU
PHE
PHE
GLU
ASP
LYS
SER
ASN
ASP
VAL
ILE
GLN
ASP
ASP
VAL
PHE
ARG
ARG
L284
S285
A286
C287
H288
N289
V290
H295
Q309
Q313
K320
G321
E322
T323
V329

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	139.60Å 150.49Å 62.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.24 – 1.97 48.20 – 1.97	Depositor EDS
% Data completeness (in resolution range)	96.3 (48.24-1.97) 96.3 (48.20-1.97)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.75 (at 1.97Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.204 , 0.227 0.212 , 0.233	Depositor DCC
R_{free} test set	4550 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	29.2	Xtriage
Anisotropy	0.047	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 44.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	9901	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.59% 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)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, PEG

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	0.90	0/2487	0.85	0/3356
1	B	0.89	0/2480	0.89	0/3346
1	C	0.81	0/2495	0.83	0/3367
1	D	0.83	0/2446	0.85	0/3302
All	All	0.86	0/9908	0.86	0/13371

There are no bond length outliers.

There are no bond angle outliers.

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	2446	0	2447	23	0
1	B	2439	0	2438	22	0
1	C	2454	0	2451	26	0
1	D	2406	0	2403	32	0
2	A	1	0	0	0	0
3	A	7	0	10	3	0
4	A	47	0	0	0	0
4	B	41	0	0	0	0
4	C	31	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	29	0	0	2	0
All	All	9901	0	9749	89	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 5.

The worst 5 of 89 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:D:260:TYR:HB2	1:D:263:GLU:OE1	1.62	0.99
1:A:260:TYR:HB2	1:A:263:GLU:CD	1.88	0.93
1:C:279:ASP:HB3	1:C:282:ARG:HD2	1.50	0.93
1:A:260:TYR:HB2	1:A:263:GLU:CG	1.99	0.92
1:B:283:ARG:NH1	1:B:287:CYS:SG	2.43	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	311/329 (94%)	306 (98%)	5 (2%)	0	100	100
1	B	310/329 (94%)	303 (98%)	7 (2%)	0	100	100
1	C	312/329 (95%)	306 (98%)	6 (2%)	0	100	100
1	D	307/329 (93%)	297 (97%)	10 (3%)	0	100	100
All	All	1240/1316 (94%)	1212 (98%)	28 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	262/276 (95%)	250 (95%)	12 (5%)	27	14
1	B	261/276 (95%)	252 (97%)	9 (3%)	37	25
1	C	263/276 (95%)	249 (95%)	14 (5%)	22	10
1	D	258/276 (94%)	248 (96%)	10 (4%)	32	19
All	All	1044/1104 (95%)	999 (96%)	45 (4%)	29	17

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	234	ARG
1	D	80	ASN
1	C	241	GLN
1	C	261	GLU
1	D	145	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	PEG	A	402	-	6,6,6	0.35	0	5,5,5	0.61	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PEG	A	402	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	402	PEG	O2-C3-C4-O4
3	A	402	PEG	C1-C2-O2-C3

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	402	PEG	3	0

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	315/329 (95%)	0.26	8 (2%)	57	59	16, 32, 59, 76	0
1	B	314/329 (95%)	0.30	14 (4%)	33	35	15, 33, 65, 96	0
1	C	316/329 (96%)	0.33	2 (0%)	89	90	20, 38, 68, 79	0
1	D	311/329 (94%)	0.49	21 (6%)	17	19	21, 39, 67, 98	0
All	All	1256/1316 (95%)	0.35	45 (3%)	42	45	15, 35, 66, 98	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	287	CYS	6.1
1	B	286	ALA	5.6
1	B	262	ASN	5.1
1	D	286	ALA	5.1
1	D	262	ASN	4.6

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 monosaccharides 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
3	PEG	A	402	7/7	0.89	0.14	43,48,51,53	0
2	MG	A	401	1/1	0.97	0.05	39,39,39,39	0

6.5 Other polymers (i)

There are no such residues in this entry.