



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 12:54 am BST

PDB ID : 4CUJ
Title : Structure of Salmonella D-Lactate Dehydrogenase
Authors : Attarataya, J.; Zaccai, N.R.; Shoemark, D.K.; Brady, R.L.
Deposited on : 2014-03-19
Resolution : 1.88 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

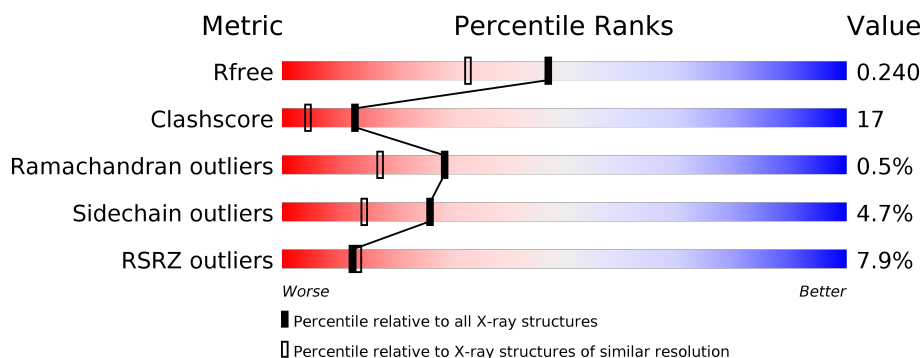
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.88 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	9470 (1.90-1.86)
Clashscore	141614	10282 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)
RSRZ outliers	127900	9303 (1.90-1.86)

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	338	<div> <div>5%</div> <div> <div></div> <div>86%</div> <div>9%</div> <div>...</div> </div> </div>
1	B	338	<div> <div>5%</div> <div> <div></div> <div>84%</div> <div>11%</div> <div>...</div> </div> </div>
1	C	338	<div> <div>5%</div> <div> <div></div> <div>86%</div> <div>8%</div> <div>...</div> </div> </div>
1	D	338	<div> <div>16%</div> <div> <div></div> <div>71%</div> <div>19%</div> <div>7%</div> <div>...</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10915 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-LACTATE DEHYDROGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	330	Total	C	N	O	S	0	0	0
			2567	1630	438	484	15			
1	B	330	Total	C	N	O	S	0	0	0
			2567	1630	438	484	15			
1	C	329	Total	C	N	O	S	0	0	0
			2558	1624	436	483	15			
1	D	329	Total	C	N	O	S	0	0	0
			2558	1624	436	483	15			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	330	LYS	-	expression tag	UNP Q8Z780
A	331	SER	-	expression tag	UNP Q8Z780
A	332	SER	-	expression tag	UNP Q8Z780
A	333	HIS	-	expression tag	UNP Q8Z780
A	334	HIS	-	expression tag	UNP Q8Z780
A	335	HIS	-	expression tag	UNP Q8Z780
A	336	HIS	-	expression tag	UNP Q8Z780
A	337	HIS	-	expression tag	UNP Q8Z780
A	338	HIS	-	expression tag	UNP Q8Z780
A	273	VAL	ASN	conflict	UNP Q8Z780
B	330	LYS	-	expression tag	UNP Q8Z780
B	331	SER	-	expression tag	UNP Q8Z780
B	332	SER	-	expression tag	UNP Q8Z780
B	333	HIS	-	expression tag	UNP Q8Z780
B	334	HIS	-	expression tag	UNP Q8Z780
B	335	HIS	-	expression tag	UNP Q8Z780
B	336	HIS	-	expression tag	UNP Q8Z780
B	337	HIS	-	expression tag	UNP Q8Z780
B	338	HIS	-	expression tag	UNP Q8Z780
B	273	VAL	ASN	conflict	UNP Q8Z780
C	330	LYS	-	expression tag	UNP Q8Z780

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	331	SER	-	expression tag	UNP Q8Z780
C	332	SER	-	expression tag	UNP Q8Z780
C	333	HIS	-	expression tag	UNP Q8Z780
C	334	HIS	-	expression tag	UNP Q8Z780
C	335	HIS	-	expression tag	UNP Q8Z780
C	336	HIS	-	expression tag	UNP Q8Z780
C	337	HIS	-	expression tag	UNP Q8Z780
C	338	HIS	-	expression tag	UNP Q8Z780
C	273	VAL	ASN	conflict	UNP Q8Z780
D	330	LYS	-	expression tag	UNP Q8Z780
D	331	SER	-	expression tag	UNP Q8Z780
D	332	SER	-	expression tag	UNP Q8Z780
D	333	HIS	-	expression tag	UNP Q8Z780
D	334	HIS	-	expression tag	UNP Q8Z780
D	335	HIS	-	expression tag	UNP Q8Z780
D	336	HIS	-	expression tag	UNP Q8Z780
D	337	HIS	-	expression tag	UNP Q8Z780
D	338	HIS	-	expression tag	UNP Q8Z780
D	273	VAL	ASN	conflict	UNP Q8Z780

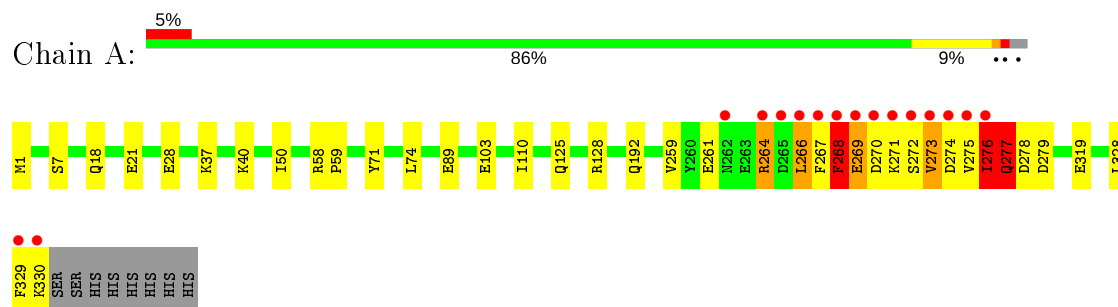
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	216	Total O 216 216	0	0
2	B	164	Total O 164 164	0	0
2	C	148	Total O 148 148	0	0
2	D	137	Total O 137 137	0	0

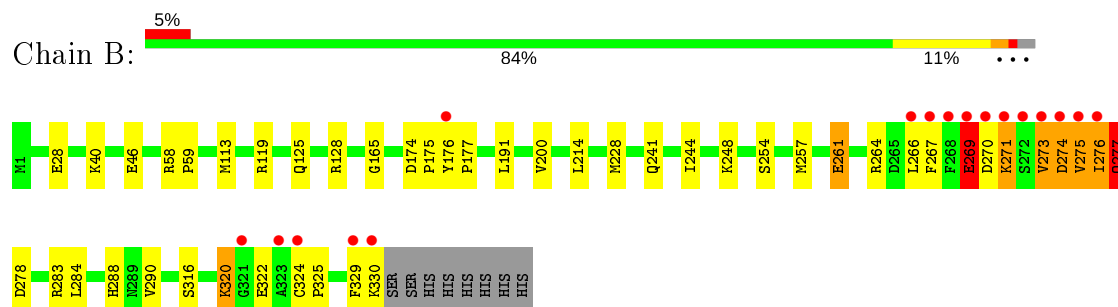
3 Residue-property plots [i](#)

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.

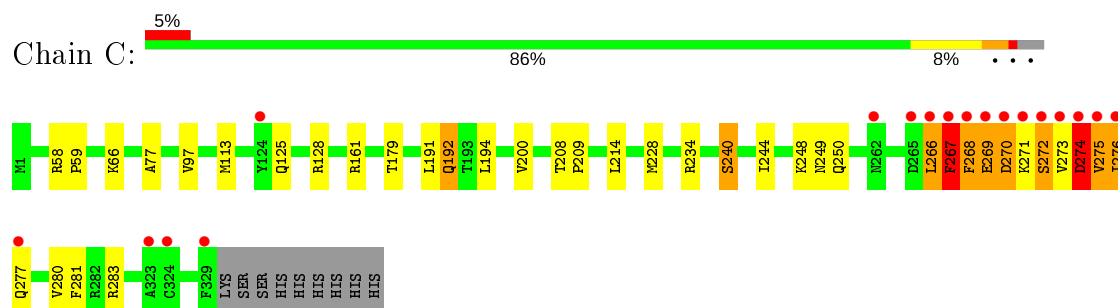
• Molecule 1: D-LACTATE DEHYDROGENASE



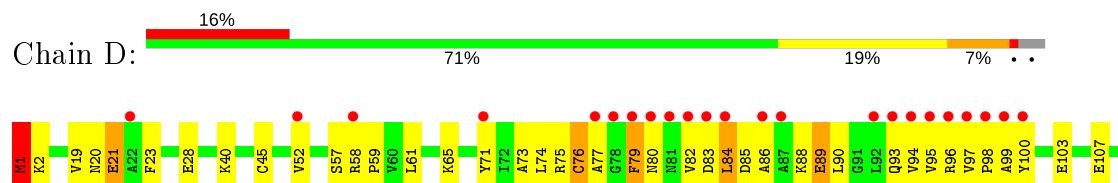
• Molecule 1: D-LACTATE DEHYDROGENASE

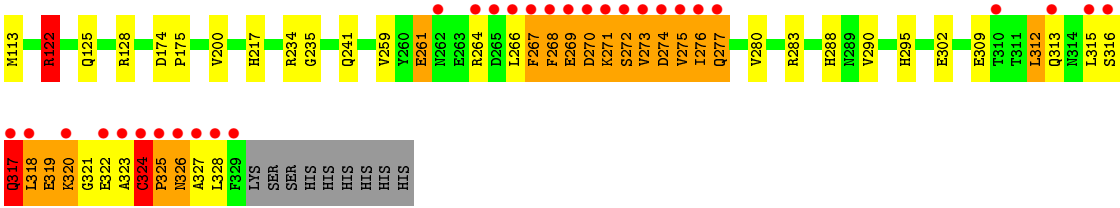


• Molecule 1: D-LACTATE DEHYDROGENASE



• Molecule 1: D-LACTATE DEHYDROGENASE





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	75.68Å 137.21Å 145.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	72.79 – 1.88 52.46 – 1.88	Depositor EDS
% Data completeness (in resolution range)	99.9 (72.79-1.88) 99.9 (52.46-1.88)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.81 (at 1.88Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.191 , 0.233 0.207 , 0.240	Depositor DCC
R_{free} test set	6158 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	19.3	Xtriage
Anisotropy	0.017	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 49.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10915	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.63% 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

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	1.39	3/2612 (0.1%)	0.83	1/3527 (0.0%)
1	B	1.34	1/2612 (0.0%)	0.86	2/3527 (0.1%)
1	C	1.35	0/2603	0.86	2/3516 (0.1%)
1	D	1.32	2/2603 (0.1%)	0.99	6/3516 (0.2%)
All	All	1.35	6/10430 (0.1%)	0.89	11/14086 (0.1%)

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	A	0	3
1	B	0	2
1	C	0	1
1	D	0	6
All	All	0	12

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	325	PRO	N-CD	5.96	1.56	1.47
1	A	103	GLU	CD-OE1	-5.71	1.19	1.25
1	D	302	GLU	CD-OE1	-5.67	1.19	1.25
1	A	7	SER	CB-OG	-5.56	1.35	1.42
1	A	261	GLU	CD-OE2	-5.38	1.19	1.25

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	122	ARG	NE-CZ-NH2	-7.36	116.62	120.30
1	D	1	MET	CG-SD-CE	7.10	111.56	100.20
1	D	324	CYS	N-CA-C	6.65	128.97	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	268	PHE	N-CA-C	-6.01	94.78	111.00
1	D	321	GLY	N-CA-C	5.86	127.74	113.10

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	266	LEU	Peptide
1	A	276	ILE	Peptide
1	A	277	GLN	Peptide
1	B	269	GLU	Peptide
1	B	276	ILE	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	2567	0	2562	61	0
1	B	2567	0	2562	60	0
1	C	2558	0	2546	71	1
1	D	2558	0	2547	160	0
2	A	216	0	0	1	0
2	B	164	0	0	8	0
2	C	148	0	0	2	0
2	D	137	0	0	6	0
All	All	10915	0	10217	338	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 17.

The worst 5 of 338 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:C:240:SER:HB3	1:C:268:PHE:CD1	1.26	1.69
1:D:324:CYS:SG	1:D:326:ASN:ND2	1.95	1.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:240:SER:HB3	1:C:268:PHE:CE1	1.57	1.37
1:C:240:SER:CB	1:C:268:PHE:CD1	2.10	1.34
1:D:324:CYS:HB3	1:D:326:ASN:CG	1.48	1.34

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:C:66:LYS:NZ	1:C:270:ASP:O[4_445]	1.89	0.31

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	328/338 (97%)	320 (98%)	6 (2%)	2 (1%)	25	14
1	B	328/338 (97%)	324 (99%)	4 (1%)	0	100	100
1	C	327/338 (97%)	320 (98%)	6 (2%)	1 (0%)	41	30
1	D	327/338 (97%)	315 (96%)	9 (3%)	3 (1%)	17	7
All	All	1310/1352 (97%)	1279 (98%)	25 (2%)	6 (0%)	29	17

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	269	GLU
1	C	274	ASP
1	D	318	LEU
1	D	273	VAL
1	D	269	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	273/281 (97%)	266 (97%)	7 (3%)	46	36
1	B	273/281 (97%)	262 (96%)	11 (4%)	31	19
1	C	272/281 (97%)	261 (96%)	11 (4%)	31	19
1	D	272/281 (97%)	250 (92%)	22 (8%)	11	4
All	All	1090/1124 (97%)	1039 (95%)	51 (5%)	26	14

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	268	PHE
1	C	276	ILE
1	D	319	GLU
1	C	272	SER
1	D	1	MET

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

Mol	Chain	Res	Type
1	B	241	GLN
1	B	277	GLN
1	C	125	GLN
1	C	192	GLN
1	D	217	HIS

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	330/338 (97%)	0.25	16 (4%) 30 32	8, 19, 44, 120	1 (0%)
1	B	330/338 (97%)	0.49	17 (5%) 27 28	10, 25, 57, 130	0
1	C	329/338 (97%)	0.44	18 (5%) 25 26	10, 22, 51, 132	0
1	D	329/338 (97%)	0.85	53 (16%) 1 2	8, 23, 67, 146	1 (0%)
All	All	1318/1352 (97%)	0.51	104 (7%) 12 13	8, 22, 59, 146	2 (0%)

The worst 5 of 104 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	272	SER	16.6
1	A	273	VAL	15.9
1	B	276	ILE	15.5
1	C	268	PHE	13.7
1	B	275	VAL	12.9

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.