



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 26, 2021 – 04:12 PM JST

PDB ID : 7CMX
Title : Isocitrate lyase from Bacillus cereus ATCC 14579
Authors : Kim, K.; Ki, D.; Lee, S.H.
Deposited on : 2020-07-29
Resolution : 1.79 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	:	2.22
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)	:	2.22

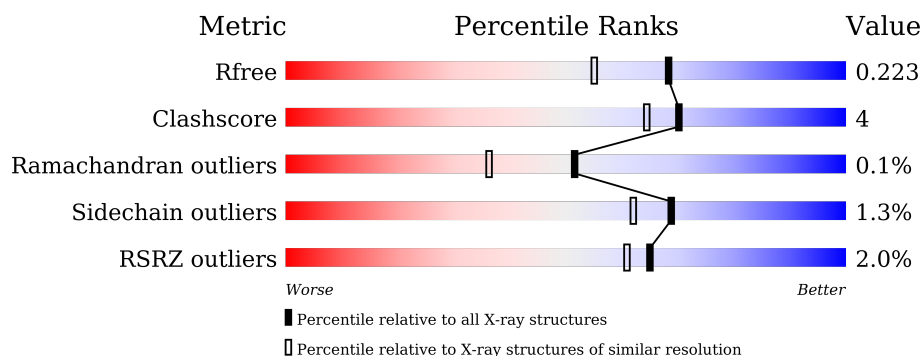
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.79 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 $\leq 5\%$. 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	433	 3% 91% 6% .
1	B	433	 2% 91% 6% .
1	C	433	 2% 88% 8% . .
1	D	433	 2% 85% 12% .

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14088 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 Isocitrate lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	422	Total	C	N	O	S	0	0	0
			3277	2060	567	639	11			
1	B	422	Total	C	N	O	S	0	0	0
			3277	2060	567	639	11			
1	C	422	Total	C	N	O	S	0	0	0
			3277	2060	567	639	11			
1	D	421	Total	C	N	O	S	0	0	0
			3268	2054	565	638	11			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	426	LEU	-	expression tag	UNP Q81GQ9
A	427	GLU	-	expression tag	UNP Q81GQ9
A	428	HIS	-	expression tag	UNP Q81GQ9
A	429	HIS	-	expression tag	UNP Q81GQ9
A	430	HIS	-	expression tag	UNP Q81GQ9
A	431	HIS	-	expression tag	UNP Q81GQ9
A	432	HIS	-	expression tag	UNP Q81GQ9
A	433	HIS	-	expression tag	UNP Q81GQ9
B	426	LEU	-	expression tag	UNP Q81GQ9
B	427	GLU	-	expression tag	UNP Q81GQ9
B	428	HIS	-	expression tag	UNP Q81GQ9
B	429	HIS	-	expression tag	UNP Q81GQ9
B	430	HIS	-	expression tag	UNP Q81GQ9
B	431	HIS	-	expression tag	UNP Q81GQ9
B	432	HIS	-	expression tag	UNP Q81GQ9
B	433	HIS	-	expression tag	UNP Q81GQ9
C	426	LEU	-	expression tag	UNP Q81GQ9
C	427	GLU	-	expression tag	UNP Q81GQ9
C	428	HIS	-	expression tag	UNP Q81GQ9
C	429	HIS	-	expression tag	UNP Q81GQ9
C	430	HIS	-	expression tag	UNP Q81GQ9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	431	HIS	-	expression tag	UNP Q81GQ9
C	432	HIS	-	expression tag	UNP Q81GQ9
C	433	HIS	-	expression tag	UNP Q81GQ9
D	426	LEU	-	expression tag	UNP Q81GQ9
D	427	GLU	-	expression tag	UNP Q81GQ9
D	428	HIS	-	expression tag	UNP Q81GQ9
D	429	HIS	-	expression tag	UNP Q81GQ9
D	430	HIS	-	expression tag	UNP Q81GQ9
D	431	HIS	-	expression tag	UNP Q81GQ9
D	432	HIS	-	expression tag	UNP Q81GQ9
D	433	HIS	-	expression tag	UNP Q81GQ9

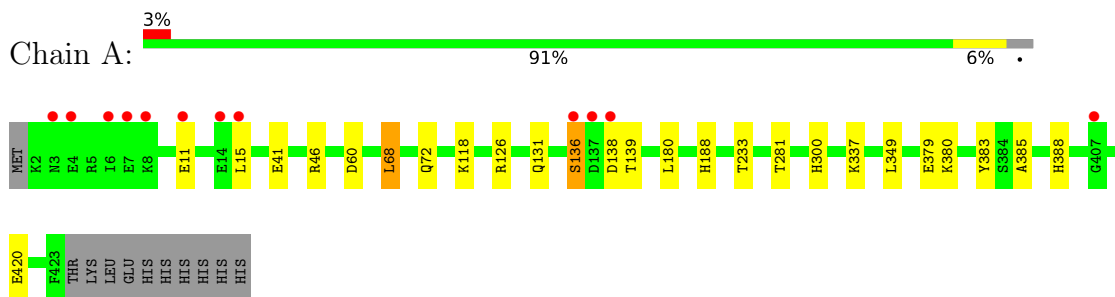
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	247	Total O 247 247	0	0
2	B	251	Total O 251 251	0	0
2	C	254	Total O 254 254	0	0
2	D	237	Total O 237 237	0	0

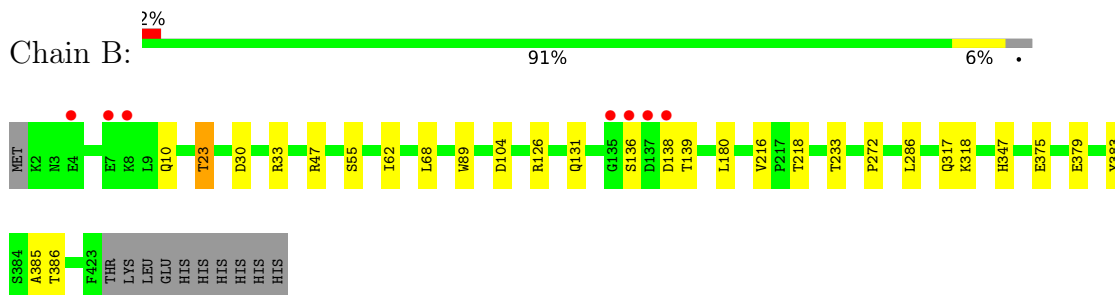
3 Residue-property plots [i](#)

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.

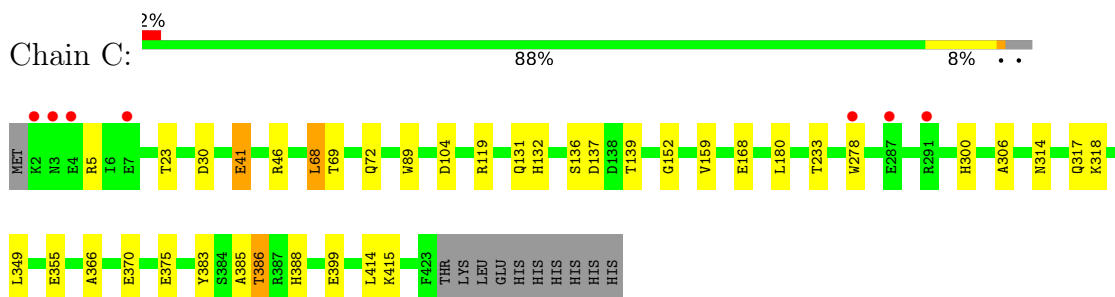
• Molecule 1: Isocitrate lyase



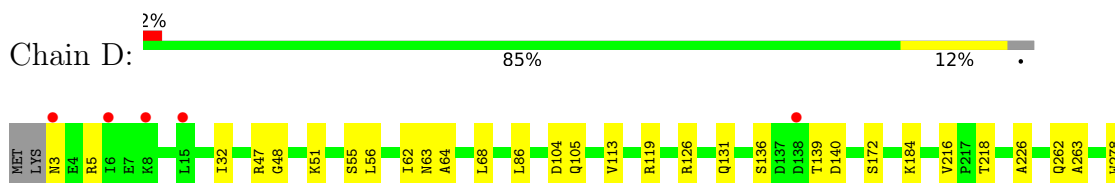
• Molecule 1: Isocitrate lyase



• Molecule 1: Isocitrate lyase



• Molecule 1: Isocitrate lyase





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	73.92Å 123.54Å 170.69Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.17 – 1.79 33.15 – 1.79	Depositor EDS
% Data completeness (in resolution range)	95.9 (33.17-1.79) 96.0 (33.15-1.79)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.61 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.175 , 0.217 0.184 , 0.223	Depositor DCC
R_{free} test set	7221 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å ²)	10.0	Xtriage
Anisotropy	0.119	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.42 , 45.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	14088	wwPDB-VP
Average B, all atoms (Å ²)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.16% 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	0.73	0/3345	0.84	1/4527 (0.0%)
1	B	0.72	0/3345	0.84	1/4527 (0.0%)
1	C	0.70	0/3345	0.83	1/4527 (0.0%)
1	D	0.72	0/3336	0.85	2/4516 (0.0%)
All	All	0.72	0/13371	0.84	5/18097 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	126	ARG	NE-CZ-NH2	-6.62	116.99	120.30
1	D	126	ARG	NE-CZ-NH2	-5.80	117.40	120.30
1	B	126	ARG	NE-CZ-NH2	-5.63	117.49	120.30
1	D	119	ARG	NE-CZ-NH2	-5.33	117.64	120.30
1	C	119	ARG	NE-CZ-NH2	-5.20	117.70	120.30

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	3277	0	3172	14	0
1	B	3277	0	3172	17	0
1	C	3277	0	3172	25	0
1	D	3268	0	3159	44	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	247	0	0	2	0
2	B	251	0	0	2	0
2	C	254	0	0	4	0
2	D	237	0	0	8	0
All	All	14088	0	12675	93	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 4.

The worst 5 of 93 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:168:GLU:HG2	2:C:505:HOH:O	1.36	1.22
1:D:375:GLU:HB3	1:D:386:THR:HG22	1.49	0.94
1:D:386:THR:HB	2:D:718:HOH:O	1.79	0.83
1:D:316:LYS:HE3	1:D:321:GLU:OE1	1.82	0.79
1:C:355:GLU:OE2	1:D:381:HIS:HE1	1.69	0.76

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	420/433 (97%)	414 (99%)	6 (1%)	0	100	100
1	B	420/433 (97%)	413 (98%)	6 (1%)	1 (0%)	47	33
1	C	420/433 (97%)	411 (98%)	9 (2%)	0	100	100
1	D	419/433 (97%)	411 (98%)	8 (2%)	0	100	100
All	All	1679/1732 (97%)	1649 (98%)	29 (2%)	1 (0%)	51	36

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	138	ASP

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	335/346 (97%)	331 (99%)	4 (1%)	71	65
1	B	335/346 (97%)	331 (99%)	4 (1%)	71	65
1	C	335/346 (97%)	328 (98%)	7 (2%)	53	42
1	D	334/346 (96%)	331 (99%)	3 (1%)	78	75
All	All	1339/1384 (97%)	1321 (99%)	18 (1%)	69	62

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	415	LYS
1	D	329	LYS
1	D	136	SER
1	C	5	ARG
1	C	386	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	188	HIS
1	C	373	GLN
1	D	388	HIS
1	C	317	GLN
1	C	388	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 monosaccharides 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	422/433 (97%)	-0.05	12 (2%) 53 47	5, 11, 26, 72	0
1	B	422/433 (97%)	-0.10	7 (1%) 70 66	5, 11, 25, 43	0
1	C	422/433 (97%)	0.01	7 (1%) 70 66	5, 12, 26, 66	0
1	D	421/433 (97%)	0.01	7 (1%) 70 66	6, 12, 28, 50	0
All	All	1687/1732 (97%)	-0.03	33 (1%) 65 61	5, 12, 26, 72	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	4	GLU	4.7
1	A	3	ASN	4.3
1	C	2	LYS	4.2
1	C	3	ASN	4.1
1	A	7	GLU	3.7

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.