Full wwPDB X-ray Structure Validation Report ⓘ

May 25, 2020 – 07:15 am BST

PDB ID	:	5M46
Title	:	Alpha-amino epsilon-caprolactam racemase (ACLR) from Rhizobacterium freirei
Authors	:	Frese, A.; Sutton, P.W.; Turkenburg, J.P.; Grogan, G.
Deposited on	:	2016-10-18
Resolution	:	1.62 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ①) were used in the production of this report:

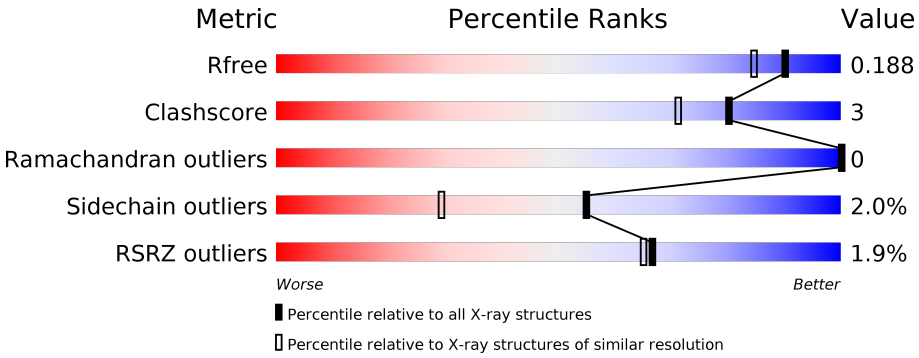
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.62 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	4693 (1.64-1.60)
Clashscore	141614	5002 (1.64-1.60)
Ramachandran outliers	138981	4888 (1.64-1.60)
Sidechain outliers	138945	4887 (1.64-1.60)
RSRZ outliers	127900	4609 (1.64-1.60)

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	440	2% 90% 7%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3383 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.

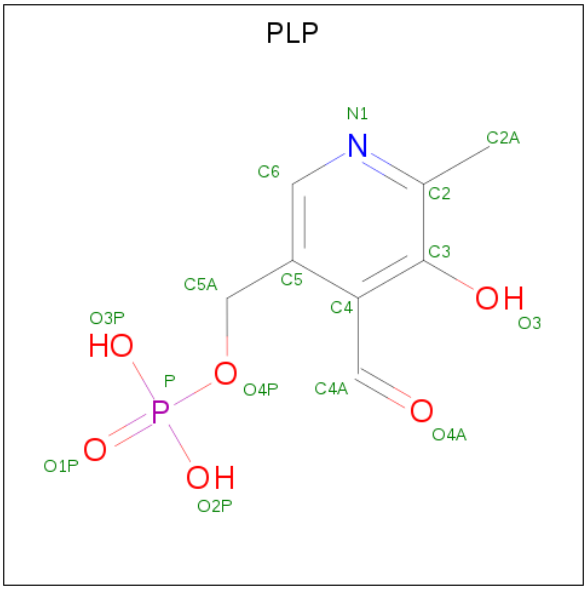
- Molecule 1 is a protein called Aminotransferase class-III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	429	3109	1968	553	576	12	0	2	0

There are 4 discrepancies between the modelled and reference sequences:

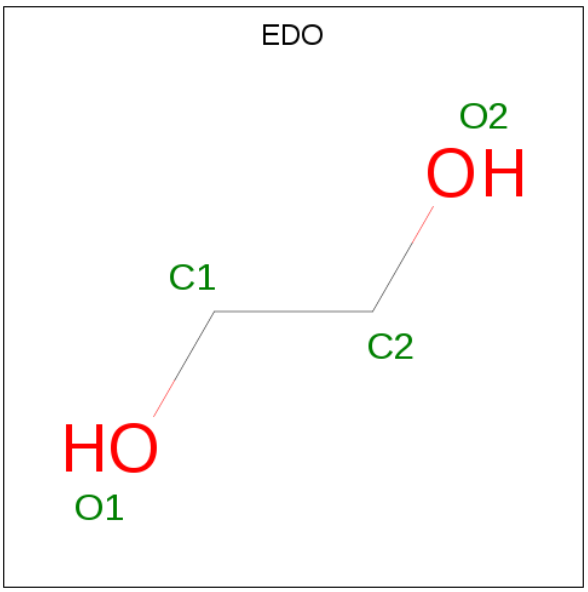
Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLN	-	expression tag	UNP N6UXY4
A	-2	GLY	-	expression tag	UNP N6UXY4
A	-1	PRO	-	expression tag	UNP N6UXY4
A	0	ALA	-	expression tag	UNP N6UXY4

- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: $C_8H_{10}NO_6P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	15	8	1	5	1	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0

- Molecule 4 is water.

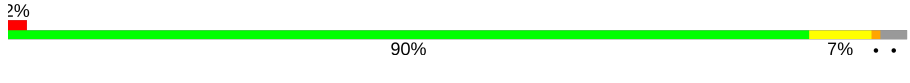
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	239	Total O 239 239	0	0

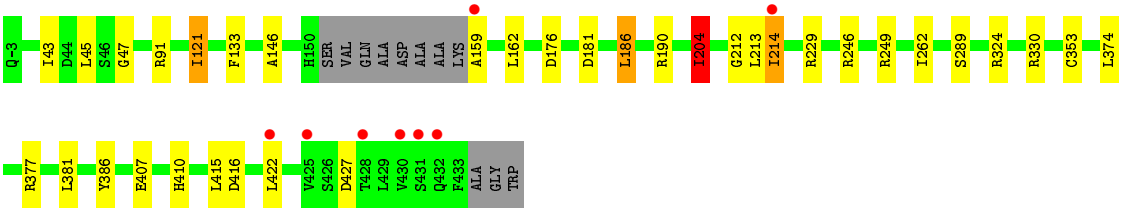
3 Residue-property plots (i)

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: Aminotransferase class-III

Chain A:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	87.61Å 76.53Å 58.57Å 90.00° 113.12° 90.00°	Depositor
Resolution (Å)	45.24 – 1.62 45.25 – 1.62	Depositor EDS
% Data completeness (in resolution range)	98.8 (45.24-1.62) 98.8 (45.25-1.62)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.83 (at 1.62Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.147 , 0.180 0.161 , 0.188	Depositor DCC
R_{free} test set	2152 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å^2)	16.9	Xtriage
Anisotropy	0.094	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 50.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	3383	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.92% 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

5.1 Standard geometry

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

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	1.11	5/3175 (0.2%)	1.12	17/4316 (0.4%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	289	SER	CB-OG	-6.69	1.33	1.42
1	A	407	GLU	CD-OE2	6.44	1.32	1.25
1	A	324	ARG	CZ-NH1	5.75	1.40	1.33
1	A	407	GLU	CD-OE1	5.66	1.31	1.25
1	A	386	TYR	CG-CD2	-5.54	1.31	1.39

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	416	ASP	CB-CG-OD1	9.11	126.50	118.30
1	A	249	ARG	NE-CZ-NH2	-7.79	116.40	120.30
1	A	324	ARG	NE-CZ-NH1	7.58	124.09	120.30
1	A	45	LEU	CB-CG-CD2	7.47	123.69	111.00
1	A	377	ARG	NE-CZ-NH1	7.28	123.94	120.30
1	A	91	ARG	NE-CZ-NH2	-7.27	116.67	120.30
1	A	324	ARG	NE-CZ-NH2	-7.19	116.70	120.30
1	A	416	ASP	CB-CG-OD2	-6.43	112.51	118.30
1	A	214	ILE	CG1-CB-CG2	-6.37	97.40	111.40
1	A	190	ARG	NE-CZ-NH1	6.21	123.41	120.30
1	A	229	ARG	NE-CZ-NH1	-5.83	117.38	120.30
1	A	181	ASP	CB-CG-OD1	5.36	123.12	118.30
1	A	204	ILE	CB-CG1-CD1	-5.33	98.97	113.90
1	A	186	LEU	CB-CG-CD2	5.28	119.98	111.00
1	A	246	ARG	NE-CZ-NH1	-5.25	117.67	120.30
1	A	121	ILE	CG1-CB-CG2	-5.05	100.30	111.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	204	ILE	CG1-CB-CG2	5.01	122.43	111.40

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	3109	0	3084	15	0
2	A	15	0	6	0	0
3	A	20	0	30	2	1
4	A	239	0	0	6	1
All	All	3383	0	3120	17	2

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 3.

All (17) 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)
3:A:503:EDO:H11	4:A:656:HOH:O	1.74	0.88
1:A:133:PHE:CE1	1:A:204:ILE:HG23	2.25	0.71
1:A:176:ASP:OD1	4:A:602:HOH:O	2.10	0.70
1:A:330:ARG:NH1	4:A:603:HOH:O	2.17	0.64
1:A:43:ILE:HD11	1:A:410:HIS:CD2	2.35	0.61
1:A:212:GLY:C	1:A:214:ILE:HD12	2.23	0.58
1:A:214:ILE:HG22	1:A:214:ILE:O	2.03	0.58
3:A:503:EDO:C1	4:A:656:HOH:O	2.41	0.55
1:A:330:ARG:HD2	4:A:780:HOH:O	2.07	0.55
1:A:213:LEU:HD13	1:A:353:CYS:C	2.32	0.49
1:A:374:LEU:HB2	1:A:422:LEU:HD21	1.95	0.48
1:A:212:GLY:CA	1:A:214:ILE:HD12	2.44	0.48
1:A:47:GLY:HA3	4:A:727:HOH:O	2.14	0.47
1:A:146:ALA:O	1:A:159:ALA:N	2.49	0.46
1:A:121:ILE:HD11	1:A:262:ILE:HD12	1.99	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:133:PHE:HE1	1:A:204:ILE:HG23	1.76	0.44
1:A:212:GLY:HA3	1:A:214:ILE:CD1	2.51	0.41

All (2) 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)
4:A:612:HOH:O	4:A:613:HOH:O[4_455]	2.17	0.03
3:A:503:EDO:O1	3:A:504:EDO:O2[4_444]	2.18	0.02

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	427/440 (97%)	417 (98%)	10 (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	301/320 (94%)	295 (98%)	6 (2%)	55 29

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	162	LEU
1	A	186	LEU
1	A	204	ILE
1	A	381	LEU
1	A	415	LEU
1	A	427	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	391	ASN
1	A	410	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

6 ligands are modelled in this entry.

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	$\#\|Z\| > 2$
2	PLP	A	501	1	15,15,16	2.16	3 (20%)	20,22,23	1.36	3 (15%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	EDO	A	503	-	3,3,3	0.71	0	2,2,2	0.29	0
3	EDO	A	506	-	3,3,3	0.72	0	2,2,2	0.62	0
3	EDO	A	502	-	3,3,3	0.76	0	2,2,2	0.33	0
3	EDO	A	505	-	3,3,3	0.88	0	2,2,2	0.39	0
3	EDO	A	504	-	3,3,3	0.73	0	2,2,2	0.52	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
2	PLP	A	501	1	-	0/6/6/8	0/1/1/1
3	EDO	A	503	-	-	1/1/1/1	-
3	EDO	A	506	-	-	0/1/1/1	-
3	EDO	A	502	-	-	0/1/1/1	-
3	EDO	A	505	-	-	0/1/1/1	-
3	EDO	A	504	-	-	1/1/1/1	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	PLP	C3-C2	4.34	1.45	1.40
2	A	501	PLP	C5-C4	4.03	1.44	1.40
2	A	501	PLP	C3-C4	3.76	1.47	1.40

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	PLP	O4P-C5A-C5	2.82	114.73	109.35
2	A	501	PLP	O3-C3-C2	2.51	122.97	117.49
2	A	501	PLP	C2A-C2-N1	2.49	122.53	117.67

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	504	EDO	O1-C1-C2-O2
3	A	503	EDO	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	503	EDO	2	1
3	A	504	EDO	0	1

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

6.1 Protein, DNA and RNA chains

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, 95^{th} 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	429/440 (97%)	-0.55	8 (1%) 66 65	9, 15, 33, 62	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	428	THR	4.9
1	A	214	ILE	4.1
1	A	432	GLN	2.8
1	A	422	LEU	2.6
1	A	431	SER	2.5
1	A	425	VAL	2.3
1	A	159	ALA	2.2
1	A	430	VAL	2.2

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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, 95^{th} 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	EDO	A	503	4/4	0.86	0.90	258,259,265,273	0
3	EDO	A	504	4/4	0.90	0.19	34,35,36,39	0
3	EDO	A	505	4/4	0.93	0.08	25,27,28,30	0
3	EDO	A	506	4/4	0.93	0.09	25,27,33,34	0
3	EDO	A	502	4/4	0.94	0.08	19,21,22,23	0
2	PLP	A	501	15/16	0.98	0.06	10,11,15,15	0

6.5 Other polymers (i)

There are no such residues in this entry.