



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 2, 2024 – 11:15 AM EST

PDB ID : 2C7B
Title : The Crystal Structure of EstE1, a New Thermophilic and Thermostable Carboxylesterase Cloned from a Metagenomic Library
Authors : Byun, J.-S.; Rhee, J.-K.; Kim, D.-U.; Oh, J.-W.; Cho, H.-S.
Deposited on : 2005-11-21
Resolution : 2.30 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

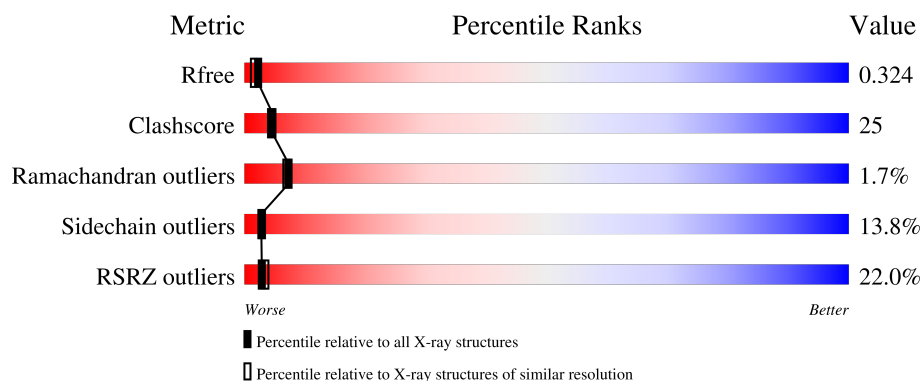
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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}	164625	5963 (2.30-2.30)
Clashscore	180529	6698 (2.30-2.30)
Ramachandran outliers	177936	6640 (2.30-2.30)
Sidechain outliers	177891	6640 (2.30-2.30)
RSRZ outliers	164620	5963 (2.30-2.30)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	311	
1	B	311	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4607 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 CARBOXYLESTERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	295	Total	C	N	O	S	Se	12	0	1
			2248	1436	388	419	1	4			
1	B	291	Total	C	N	O	S	Se	0	0	1
			2221	1418	384	414	1	4			

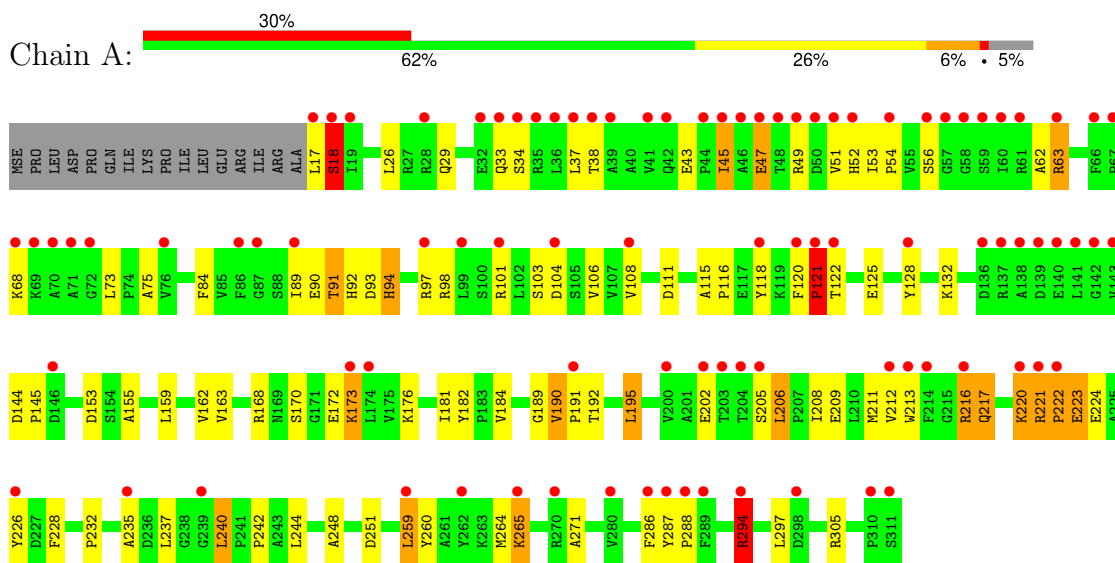
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	86	Total	O	0	0
			86	86		
2	B	52	Total	O	0	0
			52	52		

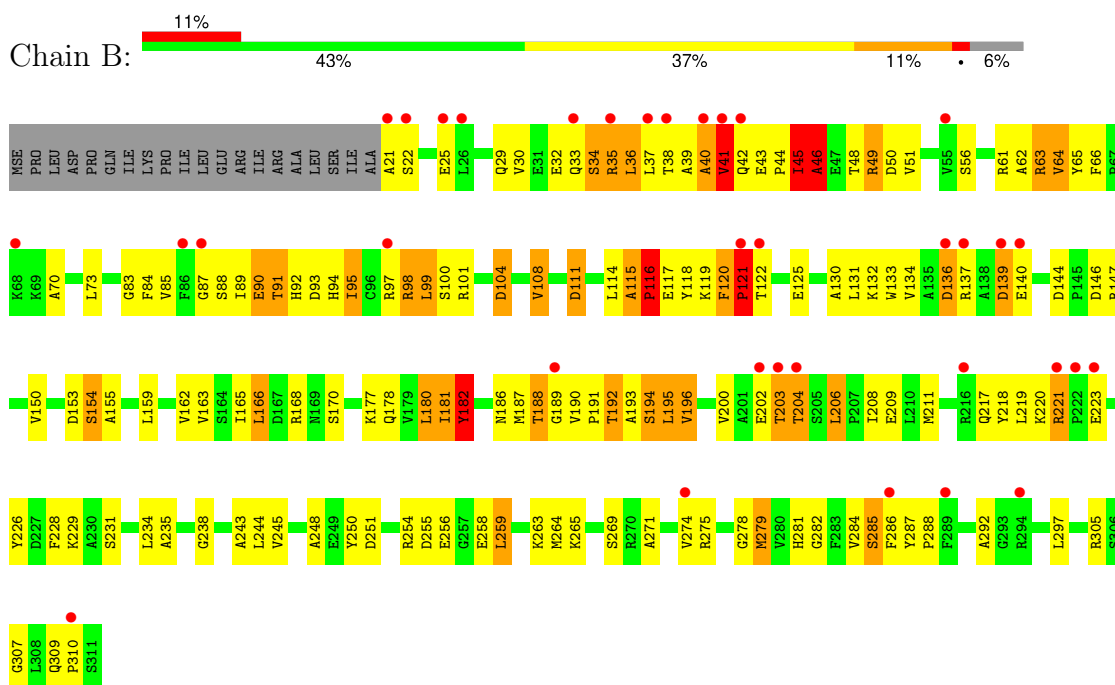
3 Residue-property plots

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. 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. 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: CARBOXYLESTERASE



• Molecule 1: CARBOXYLESTERASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	73.71Å 73.71Å 234.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30 20.00 – 2.30	Depositor EDS
% Data completeness (in resolution range)	76.1 (20.00-2.30) 99.2 (20.00-2.30)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.76 (at 2.10Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.234 , 0.261 0.329 , 0.324	Depositor DCC
R_{free} test set	1508 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å ²)	17.5	Xtriage
Anisotropy	0.441	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 25.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.80	EDS
Total number of atoms	4607	wwPDB-VP
Average B, all atoms (Å ²)	8.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.62% 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	0.59	1/2294 (0.0%)	0.95	3/3114 (0.1%)
1	B	0.99	14/2267 (0.6%)	1.48	27/3077 (0.9%)
All	All	0.82	15/4561 (0.3%)	1.25	30/6191 (0.5%)

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	1

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	115	ALA	C-N	21.80	1.75	1.34
1	B	203	THR	C-N	-11.89	1.06	1.34
1	B	114	LEU	C-N	9.98	1.56	1.34
1	B	98	ARG	C-N	-9.69	1.11	1.34
1	B	120	PHE	C-N	7.26	1.48	1.34

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	116	PRO	CA-N-CD	-33.87	64.08	111.50
1	B	121	PRO	CA-N-CD	-32.45	66.07	111.50
1	A	120	PHE	C-N-CD	-23.48	68.93	120.60
1	B	203	THR	O-C-N	-14.45	99.57	122.70
1	A	120	PHE	C-N-CA	13.16	177.27	122.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	203	THR	Mainchain

5.2 Too-close contacts

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	2248	0	2251	86	0
1	B	2221	0	2214	146	1
2	A	86	0	0	3	0
2	B	52	0	0	2	0
All	All	4607	0	4465	227	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 25.

The worst 5 of 227 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:115:ALA:C	1:B:116:PRO:N	1.75	1.37
1:B:243:ALA:HB3	1:B:264:MSE:HE1	1.37	1.07
1:B:115:ALA:CA	1:B:116:PRO:N	2.20	1.03
1:B:264:MSE:HE2	1:B:271:ALA:HB2	1.39	1.02
1:A:265:LYS:HZ1	1:A:271:ALA:N	1.59	1.01

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:B:139:ASP:N	1:B:221:ARG:NH2[6_455]	1.63	0.57

5.3 Torsion angles

5.3.1 Protein backbone

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	293/311 (94%)	280 (96%)	9 (3%)	4 (1%)	9	9
1	B	289/311 (93%)	262 (91%)	21 (7%)	6 (2%)	5	5
All	All	582/622 (94%)	542 (93%)	30 (5%)	10 (2%)	7	7

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	121	PRO
1	A	222	PRO
1	B	40	ALA
1	B	41	VAL
1	B	46	ALA

5.3.2 Protein sidechains

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	233/244 (96%)	204 (88%)	29 (12%)	4	4
1	B	230/244 (94%)	195 (85%)	35 (15%)	2	2
All	All	463/488 (95%)	399 (86%)	64 (14%)	3	3

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	219	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	223	GLU
1	A	265	LYS
1	A	259	LEU
1	B	244	LEU

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

Mol	Chain	Res	Type
1	A	29	GLN
1	A	33	GLN
1	A	309	GLN
1	B	186	ASN
1	B	309	GLN

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 oligosaccharides 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	3

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	115:ALA	C	116:PRO	N	1.75
1	B	98:ARG	C	99:LEU	N	1.11
1	B	203:THR	C	204:THR	N	1.06

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

Warning: The R factor obtained from EDS is 0.3282, which does not match the depositor's R factor of 0.2338. Please interpret the results in this section carefully.

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	291/311 (93%)	1.76	92 (31%) 1 1	6, 16, 31, 42	3 (1%)
1	B	287/311 (92%)	1.00	35 (12%) 10 11	0, 0, 0, 0	0
All	All	578/622 (92%)	1.38	127 (21%) 3 3	0, 7, 28, 42	3 (0%)

The worst 5 of 127 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	311	SER	10.7
1	A	33	GLN	6.8
1	A	202	GLU	6.2
1	A	17	LEU	6.0
1	A	220	LYS	6.0

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.