



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 30, 2024 – 10:43 PM EST

PDB ID : 1KEZ  
Title : Crystal Structure of the Macrocyclic-forming Thioesterase Domain of Erythromycin Polyketide Synthase (DEBS TE)  
Authors : Tsai, S.-C.; Miercke, L.J.W.; Krucinski, J.; Gokhale, R.; Chen, J.C.-H.; Foster, P.G.; Cane, D.E.; Khosla, C.; Stroud, R.M.  
Deposited on : 2001-11-19  
Resolution : 2.80 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](mailto: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  
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

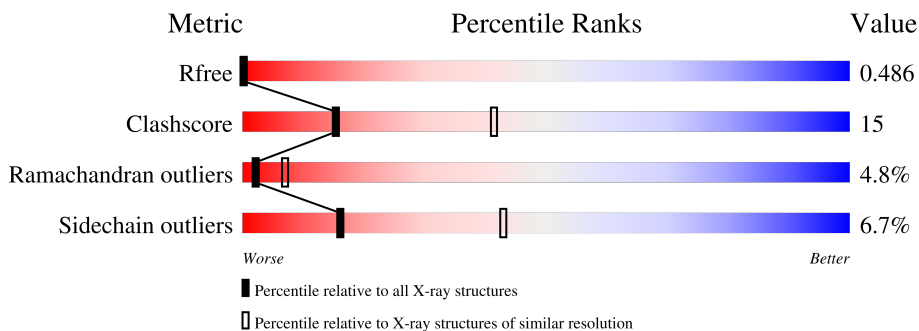
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

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  $\geq 3$ , 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	300	61% 23% 5% 11%
1	B	300	59% 24% 6% 11%
1	C	300	61% 27% • 11%

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 6426 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 ERYTHRONOLIDE SYNTHASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	267	2011	1263	355	384	9	0	0	0
1	B	267	2011	1263	355	384	9	0	0	0
1	C	267	2011	1263	355	384	9	0	0	0

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	cloning artifact	UNP Q03133
A	2	ALA	-	cloning artifact	UNP Q03133
A	3	SER	-	cloning artifact	UNP Q03133
A	284	SER	-	expression tag	UNP Q03133
A	285	SER	-	expression tag	UNP Q03133
A	286	VAL	-	expression tag	UNP Q03133
A	287	ASP	-	expression tag	UNP Q03133
A	288	LYS	-	expression tag	UNP Q03133
A	289	LEU	-	expression tag	UNP Q03133
A	290	ALA	-	expression tag	UNP Q03133
A	291	ALA	-	expression tag	UNP Q03133
A	292	ALA	-	expression tag	UNP Q03133
A	293	LEU	-	expression tag	UNP Q03133
A	294	GLU	-	expression tag	UNP Q03133
A	295	HIS	-	expression tag	UNP Q03133
A	296	HIS	-	expression tag	UNP Q03133
A	297	HIS	-	expression tag	UNP Q03133
A	298	HIS	-	expression tag	UNP Q03133
A	299	HIS	-	expression tag	UNP Q03133
A	300	HIS	-	expression tag	UNP Q03133
B	1	MET	-	cloning artifact	UNP Q03133
B	2	ALA	-	cloning artifact	UNP Q03133
B	3	SER	-	cloning artifact	UNP Q03133

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	284	SER	-	expression tag	UNP Q03133
B	285	SER	-	expression tag	UNP Q03133
B	286	VAL	-	expression tag	UNP Q03133
B	287	ASP	-	expression tag	UNP Q03133
B	288	LYS	-	expression tag	UNP Q03133
B	289	LEU	-	expression tag	UNP Q03133
B	290	ALA	-	expression tag	UNP Q03133
B	291	ALA	-	expression tag	UNP Q03133
B	292	ALA	-	expression tag	UNP Q03133
B	293	LEU	-	expression tag	UNP Q03133
B	294	GLU	-	expression tag	UNP Q03133
B	295	HIS	-	expression tag	UNP Q03133
B	296	HIS	-	expression tag	UNP Q03133
B	297	HIS	-	expression tag	UNP Q03133
B	298	HIS	-	expression tag	UNP Q03133
B	299	HIS	-	expression tag	UNP Q03133
B	300	HIS	-	expression tag	UNP Q03133
C	1	MET	-	cloning artifact	UNP Q03133
C	2	ALA	-	cloning artifact	UNP Q03133
C	3	SER	-	cloning artifact	UNP Q03133
C	284	SER	-	expression tag	UNP Q03133
C	285	SER	-	expression tag	UNP Q03133
C	286	VAL	-	expression tag	UNP Q03133
C	287	ASP	-	expression tag	UNP Q03133
C	288	LYS	-	expression tag	UNP Q03133
C	289	LEU	-	expression tag	UNP Q03133
C	290	ALA	-	expression tag	UNP Q03133
C	291	ALA	-	expression tag	UNP Q03133
C	292	ALA	-	expression tag	UNP Q03133
C	293	LEU	-	expression tag	UNP Q03133
C	294	GLU	-	expression tag	UNP Q03133
C	295	HIS	-	expression tag	UNP Q03133
C	296	HIS	-	expression tag	UNP Q03133
C	297	HIS	-	expression tag	UNP Q03133
C	298	HIS	-	expression tag	UNP Q03133
C	299	HIS	-	expression tag	UNP Q03133
C	300	HIS	-	expression tag	UNP Q03133

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	83	Total O 83 83	0	0

*Continued on next page...*

*Continued from previous page...*

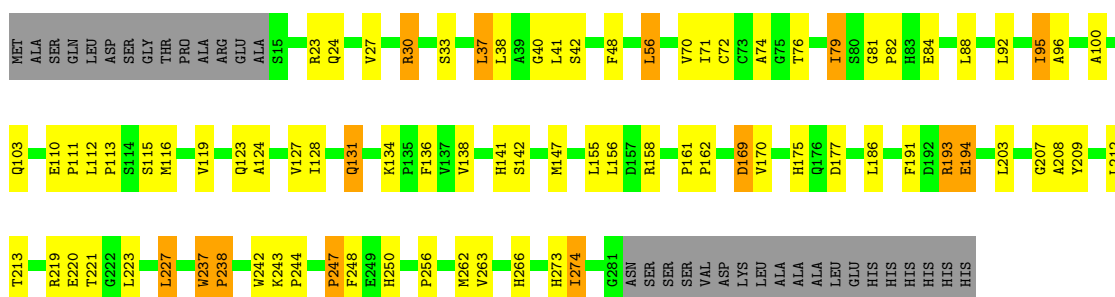
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
2	B	217	Total 217	O 217	0	0
2	C	93	Total 93	O 93	0	0

### 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.

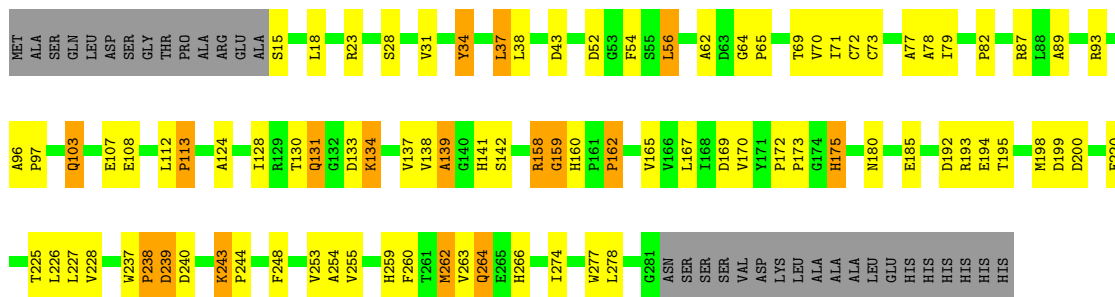
- Molecule 1: ERYTHRONOLIDE SYNTHASE

Chain A: 



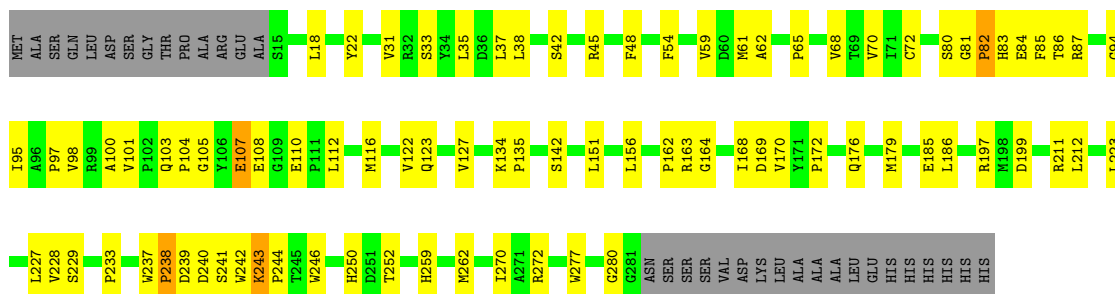
- Molecule 1: ERYTHRONOLIDE SYNTHASE

Chain B: 



- Molecule 1: ERYTHRONOLIDE SYNTHASE

Chain C: 



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	130.50Å 130.50Å 208.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.80 49.68 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.1 (30.00-2.80) 89.6 (49.68-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.03	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.43 (at 2.81Å)	Xtrriage
Refinement program	CNS, XTALVIEW	Depositor
R, $R_{free}$	0.254 , 0.279 0.476 , 0.486	Depositor DCC
$R_{free}$ test set	2278 reflections (4.48%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	79.6	Xtrriage
Anisotropy	0.107	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , -10.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.019 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.73	EDS
Total number of atoms	6426	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	71.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.44% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.35	0/2065	0.58	0/2821
1	B	0.40	0/2065	0.66	0/2821
1	C	0.35	0/2065	0.58	0/2821
All	All	0.37	0/6195	0.61	0/8463

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	2011	0	1925	64	0
1	B	2011	0	1925	66	0
1	C	2011	0	1925	50	0
2	A	83	0	0	12	0
2	B	217	0	0	26	0
2	C	93	0	0	2	0
All	All	6426	0	5775	176	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 15.

The worst 5 of 176 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:69:THR:HG23	1:B:97:PRO:HB2	1.53	0.90
1:C:107:GLU:HG3	1:C:108:GLU:N	1.98	0.78
1:A:71:ILE:HD13	1:A:136:PHE:HB2	1.67	0.76
1:A:169:ASP:HA	1:A:244:PRO:HG3	1.69	0.74
1:A:74:ALA:HB3	1:A:103:GLN:HG2	1.70	0.73

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	265/300 (88%)	219 (83%)	39 (15%)	7 (3%)	5 18
1	B	265/300 (88%)	189 (71%)	54 (20%)	22 (8%)	1 2
1	C	265/300 (88%)	218 (82%)	38 (14%)	9 (3%)	3 13
All	All	795/900 (88%)	626 (79%)	131 (16%)	38 (5%)	2 7

5 of 38 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	193	ARG
1	A	238	PRO
1	B	78	ALA
1	B	131	GLN
1	B	200	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	204/230 (89%)	189 (93%)	15 (7%)	13	37
1	B	204/230 (89%)	187 (92%)	17 (8%)	11	32
1	C	204/230 (89%)	195 (96%)	9 (4%)	28	61
All	All	612/690 (89%)	571 (93%)	41 (7%)	16	43

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	195	THR
1	C	186	LEU
1	B	199	ASP
1	C	87	ARG
1	C	199	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	103	GLN
1	B	141	HIS
1	B	264	GLN
1	A	47	HIS
1	A	24	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 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

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.