



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 11, 2024 – 08:04 AM EDT

PDB ID : 9CEL  
Title : Juvenimicin Thioesterase  
Authors : Akey, D.L.; Smith, J.S.; Choudhary, V.  
Deposited on : 2024-06-26  
Resolution : 2.27 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) : 1.20.1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.002 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.38.3

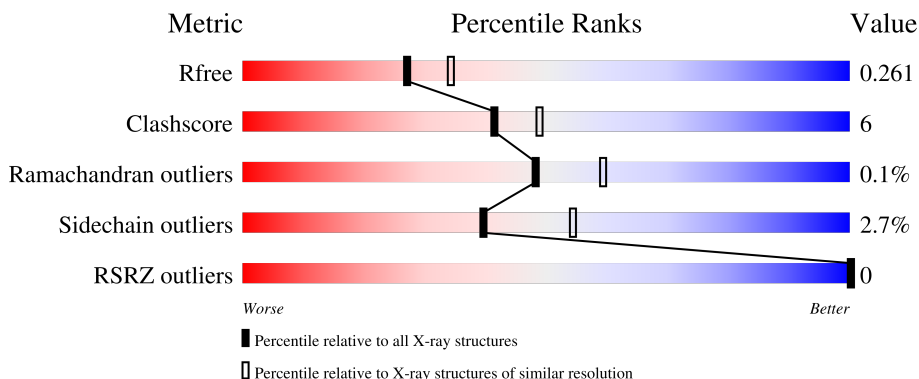
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.27 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	8487 (2.30-2.26)
Clashscore	180529	9437 (2.30-2.26)
Ramachandran outliers	177936	9341 (2.30-2.26)
Sidechain outliers	177891	9342 (2.30-2.26)
RSRZ outliers	164620	8487 (2.30-2.26)

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  $\leq 5\%$ . 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	271	
1	B	271	
1	C	271	
1	D	271	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7700 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 Type I PKS module 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	246	1793	1135	319	330	9	0	0	0
1	B	250	1835	1161	328	337	9	0	0	0
1	C	257	1891	1199	335	347	10	0	0	0
1	D	261	1929	1225	343	351	10	0	0	0

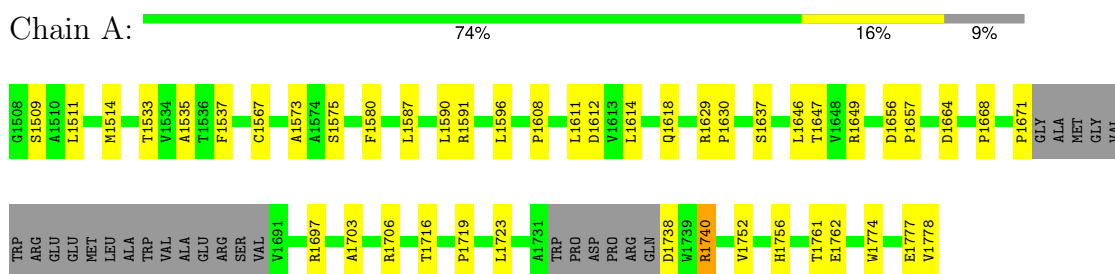
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	66	Total	O	0	0
			66	66		
2	B	68	Total	O	0	0
			68	68		
2	C	62	Total	O	0	0
			62	62		
2	D	56	Total	O	0	0
			56	56		

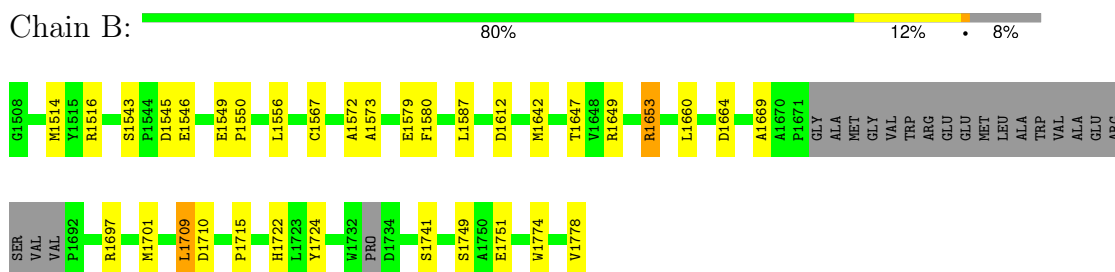
### 3 Residue-property plots

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 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.

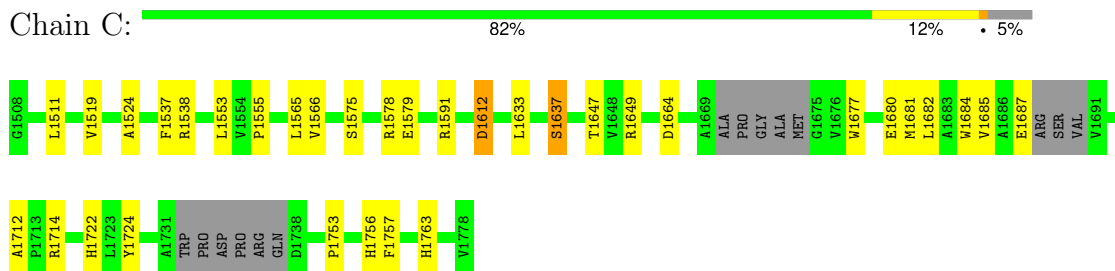
- Molecule 1: Type I PKS module 7



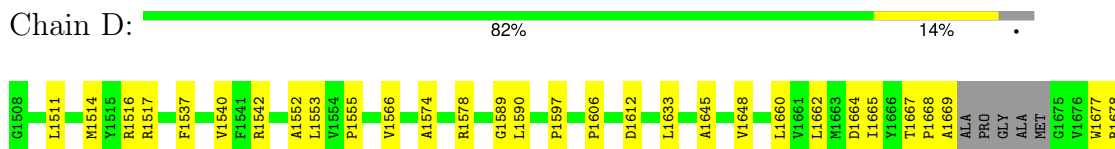
- Molecule 1: Type I PKS module 7



- Molecule 1: Type I PKS module 7



- Molecule 1: Type I PKS module 7



A1686	GLU	A1691
ARG		R1706
SER		L1709
VAL		A1712
		P1713
		W1732
		P1733
		D1734
		P1735
		A1736
		GLN
		D1738
		W1739
		R1740
		W1778

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.34Å 95.67Å 83.12Å 90.00° 93.67° 90.00°	Depositor
Resolution (Å)	45.62 – 2.27 45.62 – 2.27	Depositor EDS
% Data completeness (in resolution range)	97.8 (45.62-2.27) 98.1 (45.62-2.27)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.29 (at 2.27Å)	Xtrriage
Refinement program	PHENIX 1.21_5207	Depositor
R, $R_{free}$	0.201 , 0.262 0.204 , 0.261	Depositor DCC
$R_{free}$ test set	42914 reflections (4.55%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.7	Xtrriage
Anisotropy	0.349	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 38.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7700	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 72.96 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.0032e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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	0.44	0/1838	0.63	0/2519
1	B	0.43	0/1883	0.63	0/2580
1	C	0.41	0/1938	0.60	0/2653
1	D	0.42	0/1980	0.63	0/2713
All	All	0.42	0/7639	0.62	0/10465

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	1793	0	1770	30	0
1	B	1835	0	1804	24	0
1	C	1891	0	1859	24	0
1	D	1929	0	1894	21	0
2	A	66	0	0	3	0
2	B	68	0	0	7	0
2	C	62	0	0	2	0
2	D	56	0	0	3	0
All	All	7700	0	7327	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 6.

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:A:1647:THR:HG23	1:A:1657:PRO:HG2	1.64	0.78
1:A:1629:ARG:HD2	1:A:1630:PRO:HD2	1.71	0.72
1:B:1612:ASP:OD1	1:B:1649:ARG:NH2	2.24	0.70
1:B:1715:PRO:O	2:B:1801:HOH:O	2.08	0.70
1:A:1637:SER:HB3	1:A:1756:HIS:CE1	2.31	0.66

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	240/271 (89%)	231 (96%)	9 (4%)	0	100	100
1	B	244/271 (90%)	234 (96%)	10 (4%)	0	100	100
1	C	249/271 (92%)	237 (95%)	12 (5%)	0	100	100
1	D	253/271 (93%)	243 (96%)	9 (4%)	1 (0%)	30	36
All	All	986/1084 (91%)	945 (96%)	40 (4%)	1 (0%)	48	59

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	1668	PRO

### 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	177/197 (90%)	173 (98%)	4 (2%)	45	60
1	B	181/197 (92%)	175 (97%)	6 (3%)	33	46
1	C	186/197 (94%)	182 (98%)	4 (2%)	47	62
1	D	190/197 (96%)	184 (97%)	6 (3%)	34	47
All	All	734/788 (93%)	714 (97%)	20 (3%)	40	54

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	1516	ARG
1	D	1612	ASP
1	D	1678	ARG
1	D	1664	ASP
1	B	1664	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1763	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 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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	246/271 (90%)	-1.16	0 100 100	32, 50, 75, 93	0
1	B	250/271 (92%)	-1.16	0 100 100	35, 49, 71, 88	0
1	C	257/271 (94%)	-1.13	0 100 100	35, 51, 85, 108	0
1	D	261/271 (96%)	-1.11	0 100 100	33, 50, 92, 118	0
All	All	1014/1084 (93%)	-1.14	0 100 100	32, 50, 81, 118	0

There are no RSRZ outliers to report.

### 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.