



# wwPDB X-ray Structure Validation Summary Report

Oct 3, 2023 – 06:19 AM EDT

PDB ID : 6PUA  
Title : The 2.0 Å Crystal Structure of the Type B Chloramphenicol Acetyltransferase from *Vibrio cholerae*  
Authors : Kim, Y.; Maltseva, N.; Stam, J.; Joachimiak, A.; Center for Structural Genomics of Infectious Diseases (CSGID)  
Deposited on : 2019-07-18  
Resolution : 2.00 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 : **FAILED**  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : **FAILED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

## 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 2.00 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 5475 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 Chloramphenicol acetyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	210	1693	1080	288	314	11	0	3	0
1	B	210	1665	1064	280	310	11	0	0	0
1	C	206	1650	1055	275	309	11	0	2	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q9KMN1
A	-1	ASN	-	expression tag	UNP Q9KMN1
A	0	ALA	-	expression tag	UNP Q9KMN1
B	-2	SER	-	expression tag	UNP Q9KMN1
B	-1	ASN	-	expression tag	UNP Q9KMN1
B	0	ALA	-	expression tag	UNP Q9KMN1
C	-2	SER	-	expression tag	UNP Q9KMN1
C	-1	ASN	-	expression tag	UNP Q9KMN1
C	0	ALA	-	expression tag	UNP Q9KMN1

- Molecule 2 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 8 6 2	0	0
2	A	1	Total C O 8 6 2	0	0
2	B	1	Total C O 8 6 2	0	0
2	B	1	Total C O 8 6 2	0	0
2	C	1	Total C O 8 6 2	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).

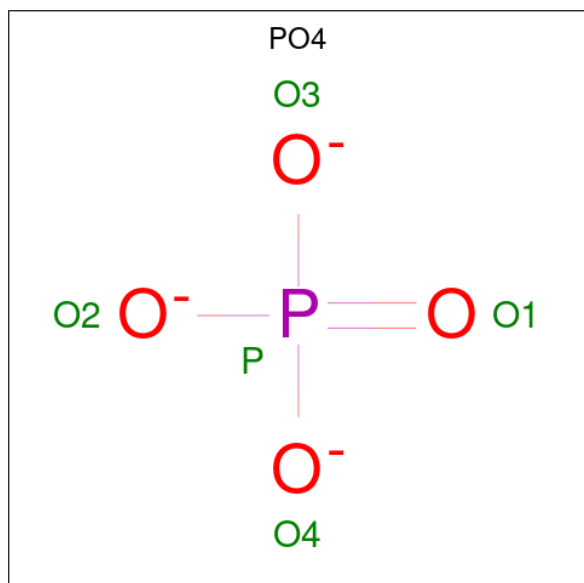


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	C	1	Total C O 4 2 2	0	0
3	C	1	Total C O 4 2 2	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Cl 1 1	0	0

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	P	0	0
			5	4	1		
5	B	1	Total	O	P	0	0
			5	4	1		
5	B	1	Total	O	P	0	0
			5	4	1		
5	C	1	Total	O	P	0	0
			5	4	1		
5	C	1	Total	O	P	0	0
			5	4	1		
5	C	1	Total	O	P	0	0
			5	4	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	145	Total	O	0	0
			145	145		
6	B	103	Total	O	0	0
			103	103		
6	C	108	Total	O	0	0
			108	108		

MolProbity and EDS failed to run properly - this section is therefore empty.

### 3 Data and refinement statistics i

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.06Å 100.06Å 128.04Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.42 – 2.00	Depositor
% Data completeness (in resolution range)	99.5 (39.42-2.00)	Depositor
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.51 (at 2.00Å)	Xtrriage
Refinement program	PHENIX 1.13_2998	Depositor
R, $R_{free}$	0.163 , 0.188	Depositor
Wilson B-factor (Å <sup>2</sup> )	37.4	Xtrriage
Anisotropy	0.190	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.018 for -h,-k,l	Xtrriage
Total number of atoms	5475	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.78% 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.

## 4 Model quality [i](#)

### 4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.3 Torsion angles [i](#)

#### 4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

#### 4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

#### 4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 4.6 Ligand geometry [i](#)

Of 22 ligands modelled in this entry, 1 is monoatomic - leaving 21 for Mogul analysis.

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
3	EDO	A	302	-	3,3,3	0.51	0	2,2,2	0.15	0
3	EDO	B	302	-	3,3,3	0.49	0	2,2,2	0.25	0
2	MPD	C	304	-	7,7,7	0.57	0	9,10,10	0.75	0
3	EDO	B	301	-	3,3,3	0.45	0	2,2,2	0.37	0
3	EDO	B	303	-	3,3,3	0.50	0	2,2,2	0.32	0
2	MPD	A	307	-	7,7,7	0.37	0	9,10,10	1.17	1 (11%)
5	PO4	B	306	-	4,4,4	0.86	0	6,6,6	0.53	0
5	PO4	C	302	-	4,4,4	0.92	0	6,6,6	0.39	0
3	EDO	A	303	-	3,3,3	0.55	0	2,2,2	0.16	0
5	PO4	A	309	-	4,4,4	0.84	0	6,6,6	0.48	0
3	EDO	A	305	-	3,3,3	0.48	0	2,2,2	0.31	0
2	MPD	A	301	-	7,7,7	0.23	0	9,10,10	0.40	0
5	PO4	C	306	-	4,4,4	0.87	0	6,6,6	0.48	0
3	EDO	A	304	-	3,3,3	0.51	0	2,2,2	0.20	0
5	PO4	B	304	-	4,4,4	0.88	0	6,6,6	0.35	0
2	MPD	B	305	-	7,7,7	0.30	0	9,10,10	0.53	0
3	EDO	C	303	-	3,3,3	0.45	0	2,2,2	0.43	0
3	EDO	A	306	-	3,3,3	0.48	0	2,2,2	0.32	0
2	MPD	B	307	-	7,7,7	0.27	0	9,10,10	0.23	0
5	PO4	C	301	-	4,4,4	0.89	0	6,6,6	0.49	0
3	EDO	C	305	-	3,3,3	0.47	0	2,2,2	0.37	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. <sup>1,2</sup> means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	MPD	B	305	-	-	3/5/5/5	-
3	EDO	A	304	-	-	0/1/1/1	-
3	EDO	B	303	-	-	0/1/1/1	-
3	EDO	C	303	-	-	0/1/1/1	-
3	EDO	A	302	-	-	0/1/1/1	-
3	EDO	A	306	-	-	0/1/1/1	-
3	EDO	C	305	-	-	1/1/1/1	-
3	EDO	B	302	-	-	1/1/1/1	-
2	MPD	C	304	-	-	2/5/5/5	-
2	MPD	A	307	-	-	3/5/5/5	-

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	MPD	B	307	-	-	0/5/5/5	-
3	EDO	A	303	-	-	0/1/1/1	-
3	EDO	B	301	-	-	1/1/1/1	-
3	EDO	A	305	-	-	1/1/1/1	-
2	MPD	A	301	-	-	0/5/5/5	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	A	307	MPD	CM-C2-C1	-2.73	104.89	110.57

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	301	EDO	O1-C1-C2-O2
2	A	307	MPD	O2-C2-C3-C4
2	B	305	MPD	O2-C2-C3-C4
2	C	304	MPD	C2-C3-C4-O4
2	A	307	MPD	C1-C2-C3-C4

There are no ring outliers.

No monomer is involved in short contacts.

#### 4.7 Other polymers [i](#)

There are no such residues in this entry.

#### 4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 5 Fit of model and data

### 5.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

### 5.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

### 5.4 Ligands

EDS failed to run properly - this section is therefore empty.

### 5.5 Other polymers

EDS failed to run properly - this section is therefore empty.