



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 22, 2024 – 12:13 PM EDT

PDB ID : 5H7Z  
Title : Apo structure of immunity protein TplEi of T6SS from *Pseudomonas aeruginosa*  
Authors : Gao, X.P.; Mu, Z.X.; Cui, S.  
Deposited on : 2016-11-21  
Resolution : 3.06 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
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
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)	:	2.37.1

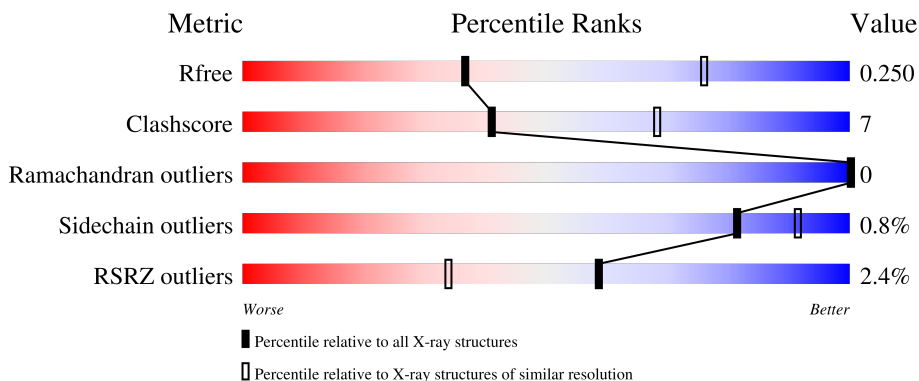
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.06 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	1754 (3.10-3.02)
Clashscore	141614	1864 (3.10-3.02)
Ramachandran outliers	138981	1794 (3.10-3.02)
Sidechain outliers	138945	1793 (3.10-3.02)
RSRZ outliers	127900	1713 (3.10-3.02)

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	376	<div> <div>%</div> <div> <div></div> <div>70%</div> <div>8%</div> <div>21%</div> </div> </div>
1	B	376	<div> <div>3%</div> <div> <div></div> <div>70%</div> <div>10%</div> <div>20%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SO4	A	404	-	-	-	X
2	SO4	A	406	-	-	-	X
2	SO4	A	407	-	-	-	X
2	SO4	B	404	-	-	-	X
2	SO4	B	405	-	-	-	X

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 9476 atoms, of which 4642 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 Uncharacterized protein.

Mol	Chain	Residues	Atoms								ZeroOcc	AltConf	Trace
1	A	296	Total	C	H	N	O	S	Se		0	0	0
			4670	1501	2305	414	443	2	5				
1	B	302	Total	C	H	N	O	S	Se		0	0	0
			4746	1525	2337	423	454	2	5				

There are 42 discrepancies between the modelled and reference sequences:

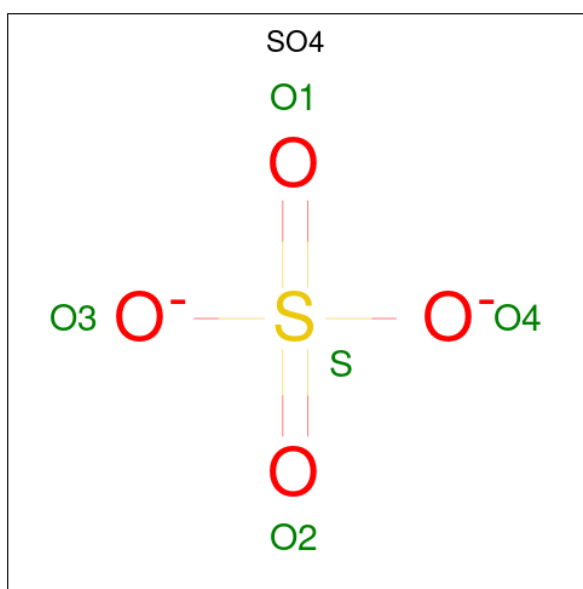
Chain	Residue	Modelled	Actual	Comment	Reference
A	5	MSE	-	expression tag	UNP Q9I3K3
A	6	GLY	-	expression tag	UNP Q9I3K3
A	7	SER	-	expression tag	UNP Q9I3K3
A	8	SER	-	expression tag	UNP Q9I3K3
A	9	HIS	-	expression tag	UNP Q9I3K3
A	10	HIS	-	expression tag	UNP Q9I3K3
A	11	HIS	-	expression tag	UNP Q9I3K3
A	12	HIS	-	expression tag	UNP Q9I3K3
A	13	HIS	-	expression tag	UNP Q9I3K3
A	14	HIS	-	expression tag	UNP Q9I3K3
A	15	SER	-	expression tag	UNP Q9I3K3
A	16	SER	-	expression tag	UNP Q9I3K3
A	17	GLY	-	expression tag	UNP Q9I3K3
A	18	LEU	-	expression tag	UNP Q9I3K3
A	19	VAL	-	expression tag	UNP Q9I3K3
A	20	PRO	-	expression tag	UNP Q9I3K3
A	21	ARG	-	expression tag	UNP Q9I3K3
A	22	GLY	-	expression tag	UNP Q9I3K3
A	23	SER	-	expression tag	UNP Q9I3K3
A	24	HIS	-	expression tag	UNP Q9I3K3
A	25	MSE	-	expression tag	UNP Q9I3K3
B	5	MSE	-	expression tag	UNP Q9I3K3
B	6	GLY	-	expression tag	UNP Q9I3K3
B	7	SER	-	expression tag	UNP Q9I3K3
B	8	SER	-	expression tag	UNP Q9I3K3

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	9	HIS	-	expression tag	UNP Q9I3K3
B	10	HIS	-	expression tag	UNP Q9I3K3
B	11	HIS	-	expression tag	UNP Q9I3K3
B	12	HIS	-	expression tag	UNP Q9I3K3
B	13	HIS	-	expression tag	UNP Q9I3K3
B	14	HIS	-	expression tag	UNP Q9I3K3
B	15	SER	-	expression tag	UNP Q9I3K3
B	16	SER	-	expression tag	UNP Q9I3K3
B	17	GLY	-	expression tag	UNP Q9I3K3
B	18	LEU	-	expression tag	UNP Q9I3K3
B	19	VAL	-	expression tag	UNP Q9I3K3
B	20	PRO	-	expression tag	UNP Q9I3K3
B	21	ARG	-	expression tag	UNP Q9I3K3
B	22	GLY	-	expression tag	UNP Q9I3K3
B	23	SER	-	expression tag	UNP Q9I3K3
B	24	HIS	-	expression tag	UNP Q9I3K3
B	25	MSE	-	expression tag	UNP Q9I3K3

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		



- Molecule 1: Uncharacterized protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	182.23Å 182.23Å 78.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	31.25 – 3.06 31.25 – 3.06	Depositor EDS
% Data completeness (in resolution range)	96.9 (31.25-3.06) 97.5 (31.25-3.06)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.28 (at 3.06Å)	Xtriage
Refinement program	PHENIX 1.9_1690	Depositor
R, $R_{free}$	0.203 , 0.245 0.215 , 0.250	Depositor DCC
$R_{free}$ test set	1236 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	63.7	Xtriage
Anisotropy	0.334	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 42.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.014 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	9476	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.15% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.36	0/2421	0.46	0/3269
1	B	0.43	0/2466	0.51	1/3331 (0.0%)
All	All	0.40	0/4887	0.48	1/6600 (0.0%)

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	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	169	ALA	C-N-CD	5.70	140.37	128.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	141	ARG	Sidechain
1	B	43	PHE	Peptide

### 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	2365	2305	2307	32	1
1	B	2409	2337	2346	45	2
2	A	35	0	0	0	0
2	B	25	0	0	1	0
All	All	4834	4642	4653	67	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 7.

The worst 5 of 67 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:326:SER:O	1:B:328:ARG:NH1	1.63	1.28
1:B:179:PHE:CD2	1:B:194:GLU:OE1	1.89	1.24
1:B:194:GLU:OE2	1:B:284:LEU:HD22	1.59	1.01
1:A:241:MSE:HE3	1:B:238:PHE:CG	2.04	0.93
1:B:179:PHE:HD2	1:B:194:GLU:OE1	1.38	0.91

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)
1:B:87:GLU:OE2	1:B:213:ARG:NH1[6_555]	2.17	0.03
1:A:228:ASP:OD2	1:B:98:ASN:H[6_555]	1.60	0.00

## 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.

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	292/376 (78%)	278 (95%)	14 (5%)	0	100	100
1	B	300/376 (80%)	287 (96%)	13 (4%)	0	100	100
All	All	592/752 (79%)	565 (95%)	27 (5%)	0	100	100

There are no Ramachandran outliers to report.

### 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	253/313 (81%)	251 (99%)	2 (1%)	81	91
1	B	258/313 (82%)	256 (99%)	2 (1%)	81	91
All	All	511/626 (82%)	507 (99%)	4 (1%)	81	91

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	123	LYS
1	A	259	TYR
1	B	168	ARG
1	B	194	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

12 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	SO4	B	401	-	4,4,4	0.14	0	6,6,6	0.04	0
2	SO4	A	402	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	A	407	-	4,4,4	0.14	0	6,6,6	0.07	0
2	SO4	A	406	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	A	404	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	B	403	-	4,4,4	0.14	0	6,6,6	0.04	0
2	SO4	B	404	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	A	405	-	4,4,4	0.15	0	6,6,6	0.10	0
2	SO4	B	402	-	4,4,4	0.14	0	6,6,6	0.10	0
2	SO4	A	403	-	4,4,4	0.14	0	6,6,6	0.08	0
2	SO4	A	401	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	B	405	-	4,4,4	0.14	0	6,6,6	0.04	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	405	SO4	1	0

## 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	291/376 (77%)	-0.15	3 (1%) 82 63	27, 52, 88, 122	0
1	B	297/376 (78%)	0.19	11 (3%) 41 20	37, 66, 115, 178	0
All	All	588/752 (78%)	0.02	14 (2%) 59 34	27, 60, 104, 178	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	302	ASN	7.3
1	B	301	GLU	4.4
1	B	299	SER	3.8
1	B	300	ASN	3.8
1	B	32	SER	3.7

### 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](#)

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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	SO4	A	406	5/5	0.63	0.49	92,102,116,141	0
2	SO4	A	404	5/5	0.74	0.51	63,77,89,123	0
2	SO4	A	407	5/5	0.75	0.50	95,99,126,140	0
2	SO4	B	402	5/5	0.76	0.31	68,84,93,115	0
2	SO4	B	405	5/5	0.76	0.54	90,106,140,145	0
2	SO4	B	404	5/5	0.79	0.46	98,99,120,130	0
2	SO4	B	403	5/5	0.86	0.25	102,114,124,129	0
2	SO4	A	403	5/5	0.89	0.18	50,75,91,98	0
2	SO4	A	402	5/5	0.90	0.34	63,84,92,108	0
2	SO4	B	401	5/5	0.92	0.28	90,101,115,125	0
2	SO4	A	405	5/5	0.93	0.18	56,57,69,86	0
2	SO4	A	401	5/5	0.96	0.14	74,79,82,83	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.