



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 28, 2024 – 06:56 am GMT

PDB ID : 4AKX  
Title : Structure of the heterodimeric complex ExoU-SpcU from the type III secretion system (T3SS) of *Pseudomonas aeruginosa*  
Authors : Gendrin, C.; Contreras-Martel, C.; Bouillot, S.; Elsen, S.; Lemaire, D.; Skoufias, D.A.; Huber, P.; Attree, I.; Dessen, A.  
Deposited on : 2012-02-29  
Resolution : 2.94 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	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

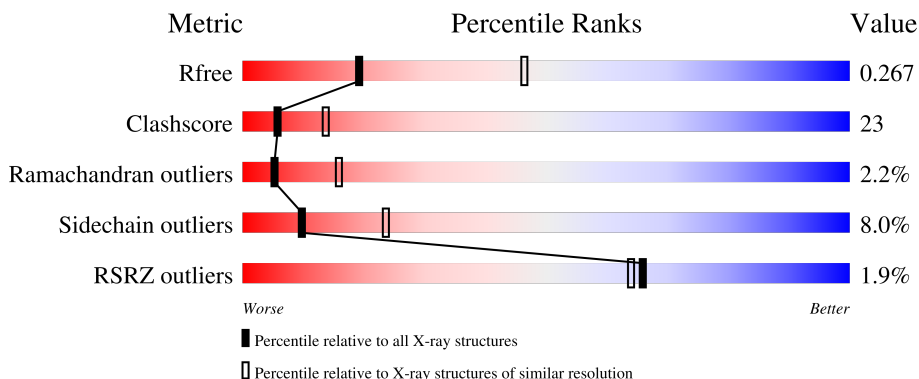
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.94 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	1067 (2.96-2.92)
Clashscore	180529	1122 (2.96-2.92)
Ramachandran outliers	177936	1075 (2.96-2.92)
Sidechain outliers	177891	1075 (2.96-2.92)
RSRZ outliers	164620	1067 (2.96-2.92)

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	127	
2	B	660	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5059 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 SPCU.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	121	Total	C	N	O	Se	0	1	0
			943	602	165	173	3			

- Molecule 2 is a protein called EXOU.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	534	Total	C	N	O	Se	0	1	0
			4078	2531	738	796	13			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	28	GLY	-	expression tag	UNP Q02IF1
B	29	SER	-	expression tag	UNP Q02IF1

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	9	Total	O	0	0
			9	9		
3	B	29	Total	O	0	0
			29	29		



T658	A660	R661	GLY	PHE	LEU	ARG	PHE	GLY	LYS	PRO	LEU	SER	S672	T673	M677	E685	PHE	THR
------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	154.89Å 52.80Å 121.81Å 90.00° 127.18° 90.00°	Depositor
Resolution (Å)	97.05 – 2.94 97.05 – 2.94	Depositor EDS
% Data completeness (in resolution range)	98.7 (97.05-2.94) 98.7 (97.05-2.94)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.00 (at 2.96Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.208 , 0.264 0.207 , 0.267	Depositor DCC
$R_{free}$ test set	840 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	81.0	Xtriage
Anisotropy	0.256	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 56.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.018 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5059	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	98.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.39% 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

### 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	0.74	1/962 (0.1%)	0.71	0/1312
2	B	0.66	2/4119 (0.0%)	0.64	2/5545 (0.0%)
All	All	0.68	3/5081 (0.1%)	0.66	2/6857 (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
2	B	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	327	SER	CB-OG	14.86	1.61	1.42
2	B	377	GLU	C-O	7.65	1.37	1.23
1	A	87	TRP	CD2-CE2	5.94	1.48	1.41

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	565	LEU	CB-CG-CD2	-5.88	101.00	111.00
2	B	99	LEU	CA-CB-CG	5.09	127.01	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	319	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	943	0	961	58	0
2	B	4078	0	4115	178	0
3	A	9	0	0	2	0
3	B	29	0	0	3	0
All	All	5059	0	5076	230	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 23.

The worst 5 of 230 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:1[A]:MSE:SE	1:A:1[A]:MSE:CE	2.14	1.45
1:A:21:LEU:HB2	1:A:33:MSE:HE3	1.21	1.20
1:A:119:GLU:HB3	1:A:120:PRO:HD3	1.35	1.08
2:B:221:GLU:HB3	2:B:222:PRO:HD2	1.35	1.06
2:B:453:HIS:CE1	2:B:457:ARG:HH11	1.73	1.05

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	119/127 (94%)	108 (91%)	9 (8%)	2 (2%)	7	21
2	B	519/660 (79%)	468 (90%)	39 (8%)	12 (2%)	5	15

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	638/787 (81%)	576 (90%)	48 (8%)	14 (2%)	5	16

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	94	VAL
2	B	319	PHE
2	B	153	SER
2	B	324	GLN
2	B	412	LEU

### 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	102/103 (99%)	95 (93%)	7 (7%)	13	30
2	B	440/517 (85%)	404 (92%)	36 (8%)	9	23
All	All	542/620 (87%)	499 (92%)	43 (8%)	10	24

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	451	LYS
2	B	619	TYR
2	B	470	LEU
2	B	531	LEU
2	B	626	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	471	GLN
2	B	499	GLN
2	B	648	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	575	GLN
2	B	623	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](#)

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	119/127 (93%)	-0.50	1 (0%) 82 81	50, 76, 130, 190	0
2	B	521/660 (78%)	-0.23	11 (2%) 63 61	33, 98, 175, 206	1 (0%)
All	All	640/787 (81%)	-0.28	12 (1%) 66 64	33, 94, 174, 206	1 (0%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	347	VAL	3.2
2	B	322	VAL	2.8
2	B	658	TYR	2.6
2	B	202	ILE	2.5
2	B	294	PRO	2.5

### 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.