WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 3, 2025 – 03:13 pm GMT

PDB ID	:	8S7L
Title	:	Crystal structure of a double mutant of VirB8-like OrfG central and C-terminal domains of Streptococcus thermophilus ICESt3 (Gram positive conjugative type IV secretion system).
Authors	:	Favier, F.; Didierjean, C.; Maffo-Woulefack, R.; Douzi, B.; Leblond-Bourget, N.
Deposited on	:	2024-03-03
Resolution	:	2.60 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*
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:

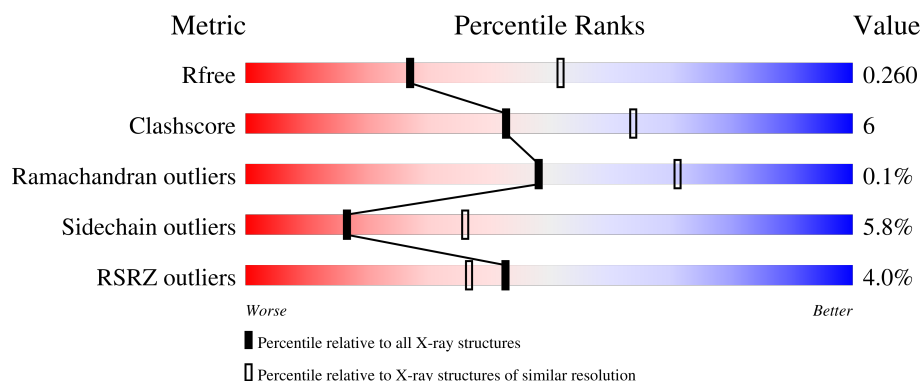|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41

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 2.60 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	Percentile Ranks	Value
Rfree		0.260
Clashscore		6
Ramachandran outliers		0.1%
Sidechain outliers		5.8%
RSRZ outliers		4.0%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	3775 (2.60-2.60)
Clashscore	180529	4181 (2.60-2.60)
Ramachandran outliers	177936	4129 (2.60-2.60)
Sidechain outliers	177891	4129 (2.60-2.60)
RSRZ outliers	164620	3775 (2.60-2.60)

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 >=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 <=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	268	3% 71% 16% 10%
1	B	268	5% 69% 17% 11%
1	C	268	3% 69% 17% 10%

WORLDWIDE

wwPDB

PROTEIN DATA BANK

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 11398 atoms, of which 5486 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 Putative transfer protein.

| Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	241	Total	C	H	N	O	S	0	0	0
			3823	1271	1840	316	393	3			
1	B	238	Total	C	H	N	O	S	0	0	0
			3777	1257	1819	312	386	3			
1	C	240	Total	C	H	N	O	S	0	0	0
			3798	1264	1827	314	390	3			

There are 6 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
A	204	CYS	ALA	engineered mutation	UNP Q70CA4
A	264	CYS	ALA	engineered mutation	UNP Q70CA4
B	204	CYS	ALA	engineered mutation	UNP Q70CA4
B	264	CYS	ALA	engineered mutation	UNP Q70CA4
C	204	CYS	ALA	engineered mutation	UNP Q70CA4
C	264	CYS	ALA	engineered mutation	UNP Q70CA4

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.
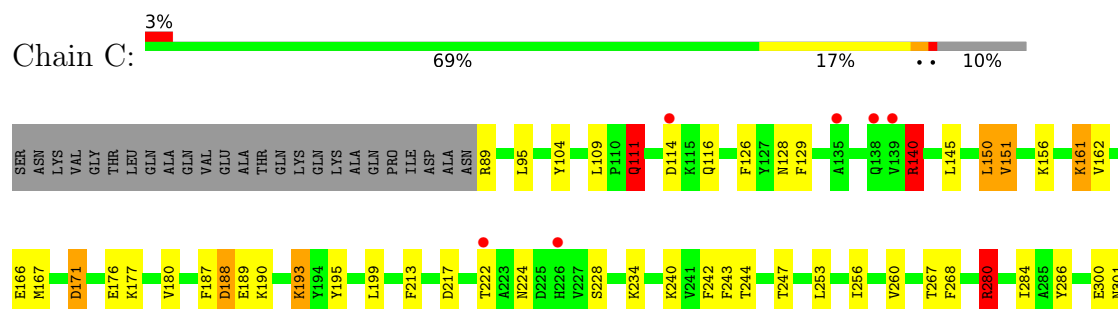
• Molecule 1: Putative transfer protein
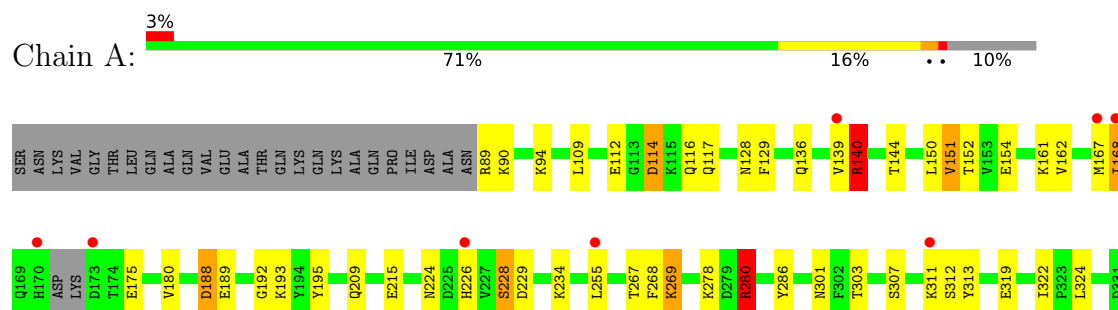
• Molecule 1: Putative transfer protein
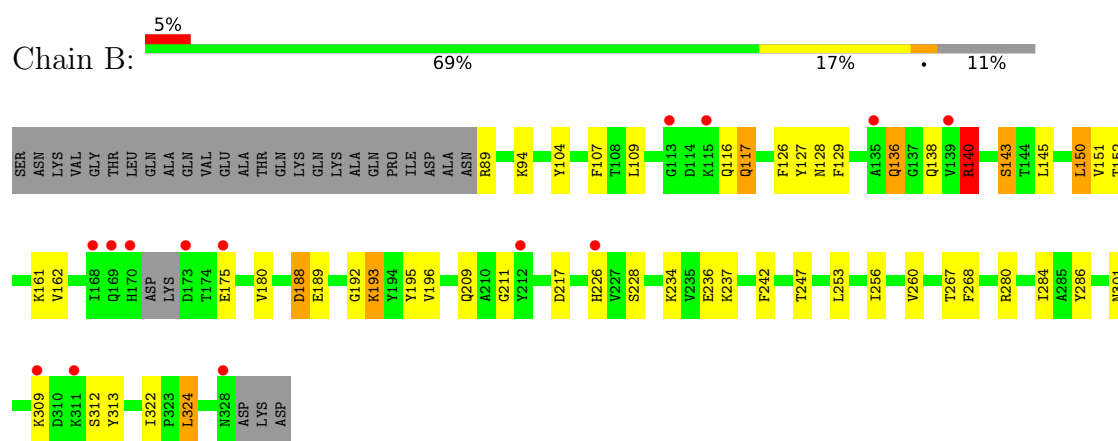
• Molecule 1: Putative transfer protein

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	152.74Å 89.19Å 98.42Å 90.00° 112.31° 90.00°	Depositor
Resolution (Å)	91.05 – 2.60 91.05 – 2.60	Depositor EDS
% Data completeness (in resolution range)	98.6 (91.05-2.60) 98.6 (91.05-2.60)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.88 (at 2.62Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.202 , 0.256 0.205 , 0.260	Depositor DCC
R_{free} test set	1819 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å^2)	60.3	Xtriage
Anisotropy	0.300	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 61.3	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	11398	wwPDB-VP
Average B, all atoms (Å^2)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.26% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0.65	1/2029 (0.0%)	1.34	17/2749 (0.6%)
1	B	0.64	0/2004	1.32	11/2716 (0.4%)
1	C	0.61	1/2018 (0.0%)	1.28	10/2737 (0.4%)
All	All	0.63	2/6051 (0.0%)	1.31	38/8202 (0.5%)

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	A	0	3
1	B	0	3
1	C	0	4
All	All	0	10

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	300	GLU	CD-OE1	5.78	1.32	1.25
1	A	319	GLU	CD-OE2	5.11	1.31	1.25

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	226	HIS	CB-CA-C	9.15	128.69	110.40
1	A	226	HIS	CB-CA-C	8.50	127.40	110.40
1	A	150	LEU	CB-CG-CD2	7.20	123.23	111.00
1	B	313	TYR	N-CA-CB	-7.03	97.95	110.60
1	B	161	LYS	CB-CA-C	6.84	124.09	110.40

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	140	ARG	Sidechain
1	A	224	ASN	Sidechain
1	A	280	ARG	Sidechain
1	B	136	GLN	Peptide
1	B	140	ARG	Sidechain

5.2 Too-close contacts ⓘ

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	1983	1840	1898	22	0
1	B	1958	1819	1877	24	1
1	C	1971	1827	1884	29	0
All	All	5912	5486	5659	70	1

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 70 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:217:ASP:HA	1:B:324:LEU:HD11	1.61	0.82
1:C:280:ARG:HH11	1:C:280:ARG:HG3	1.45	0.82
1:A:112:GLU:HB3	1:A:114:ASP:OD1	1.93	0.68
1:A:301:ASN:HB3	1:A:322:ILE:HG13	1.77	0.66
1:B:301:ASN:HB3	1:B:322:ILE:HG13	1.79	0.64

All (1) 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:175:GLU:OE2	1:B:175:GLU:OE2[2_555]	2.19	0.01

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	237/268 (88%)	219 (92%)	17 (7%)	1 (0%)	30	52
1	B	234/268 (87%)	219 (94%)	15 (6%)	0	100	100
1	C	238/268 (89%)	224 (94%)	14 (6%)	0	100	100
All	All	709/804 (88%)	662 (93%)	46 (6%)	1 (0%)	48	71

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	228	SER

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	220/242 (91%)	207 (94%)	13 (6%)	16	35
1	B	217/242 (90%)	205 (94%)	12 (6%)	18	38
1	C	218/242 (90%)	205 (94%)	13 (6%)	16	35
All	All	655/726 (90%)	617 (94%)	38 (6%)	17	36

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	150	LEU
1	C	228	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	151	VAL
1	C	177	LYS
1	C	307	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	116	GLN
1	B	111	GLN
1	B	136	GLN
1	C	116	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6   Fit of model and data ⓘ

6.1   Protein, DNA and RNA chains ⓘ

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, 95th 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^2)	Q<0.9
1	A	241/268 (89%)	0.27	8 (3%)	49	43	61, 81, 119, 153	0
1	B	238/268 (88%)	0.44	14 (5%)	29	24	61, 85, 131, 152	0
1	C	240/268 (89%)	-0.04	7 (2%)	54	48	63, 82, 125, 158	0
All	All	719/804 (89%)	0.22	29 (4%)	43	37	61, 82, 125, 158	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	170	HIS	4.9
1	B	170	HIS	4.7
1	B	169	GLN	4.2
1	C	139	VAL	4.1
1	B	168	ILE	3.3

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

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

6.3   Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4   Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.