Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2021 – 12:03 am BST

PDB ID	:	6Z0F
Title	:	Crystal structure of the membrane pseudokinase YukC/EssB from Bacillus subtilis T7SS
Authors	:	Tassinari, M.; Bellinzoni, M.; Alzari, P.M.; Fronzes, R.; Gubellini, F.
Deposited on	:	2020-05-08
Resolution	:	2.55 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see references ⓘ) were used in the production of this report:

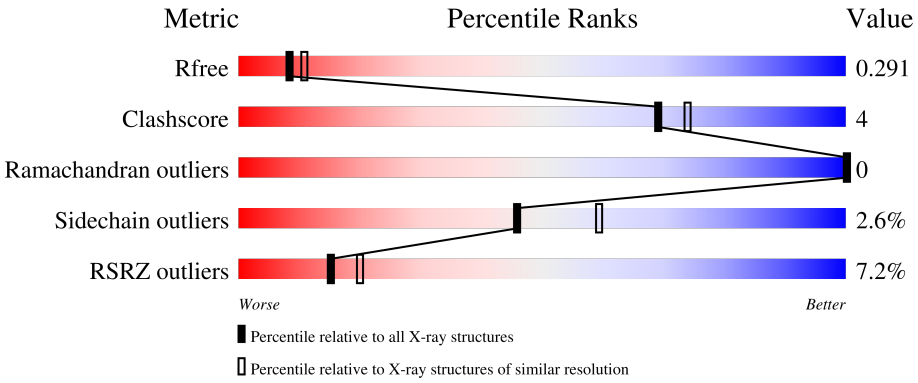
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.18
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.18

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.55 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	1279 (2.58-2.54)
Clashscore	141614	1327 (2.58-2.54)
Ramachandran outliers	138981	1312 (2.58-2.54)
Sidechain outliers	138945	1312 (2.58-2.54)
RSRZ outliers	127900	1269 (2.58-2.54)

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	423	6% 83% 7% 9%
1	B	423	7% 78% 9% 13%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5835 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 ESX secretion system protein YukC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	384	Total	C	N	O	S	0	0	0
			2996	1932	480	582	2			
1	B	369	Total	C	N	O	S	0	0	0
			2831	1828	454	547	2			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	414	TRP	-	expression tag	UNP P71070
A	415	SER	-	expression tag	UNP P71070
A	416	HIS	-	expression tag	UNP P71070
A	417	PRO	-	expression tag	UNP P71070
A	418	GLN	-	expression tag	UNP P71070
A	419	PHE	-	expression tag	UNP P71070
A	420	GLU	-	expression tag	UNP P71070
A	421	LYS	-	expression tag	UNP P71070
A	422	ALA	-	expression tag	UNP P71070
A	423	ALA	-	expression tag	UNP P71070
B	414	TRP	-	expression tag	UNP P71070
B	415	SER	-	expression tag	UNP P71070
B	416	HIS	-	expression tag	UNP P71070
B	417	PRO	-	expression tag	UNP P71070
B	418	GLN	-	expression tag	UNP P71070
B	419	PHE	-	expression tag	UNP P71070
B	420	GLU	-	expression tag	UNP P71070
B	421	LYS	-	expression tag	UNP P71070
B	422	ALA	-	expression tag	UNP P71070
B	423	ALA	-	expression tag	UNP P71070

- Molecule 2 is water.

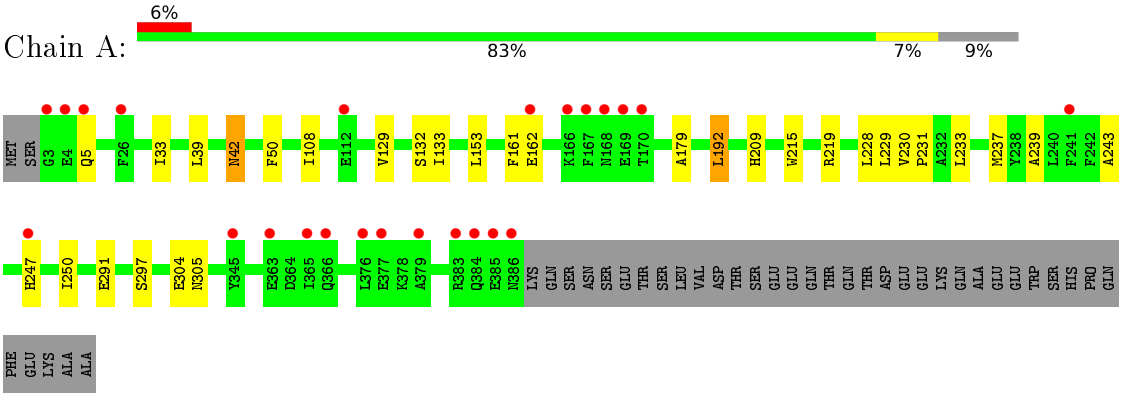
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	6	Total O 6 6	0	0
2	B	2	Total O 2 2	0	0

3 Residue-property plots (i)

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.

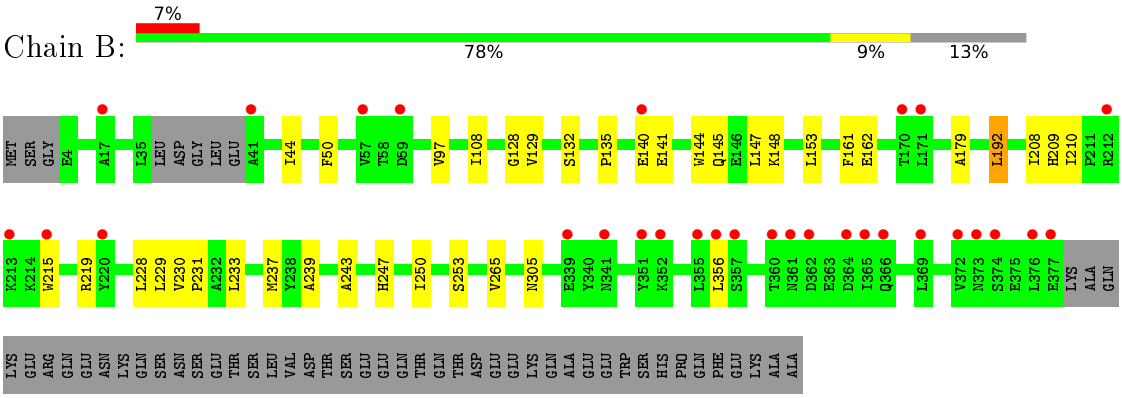
- Molecule 1: ESX secretion system protein YukC

Chain A:



- Molecule 1: ESX secretion system protein YukC

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	149.01Å 83.41Å 106.44Å 90.00° 108.16° 90.00°	Depositor
Resolution (Å)	29.42 – 2.55 29.42 – 2.55	Depositor EDS
% Data completeness (in resolution range)	65.0 (29.42-2.55) 65.0 (29.42-2.55)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.47 (at 2.54Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.235 , 0.269 0.253 , 0.291	Depositor DCC
R_{free} test set	1333 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	73.8	Xtriage
Anisotropy	0.031	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 47.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	5835	wwPDB-VP
Average B, all atoms (Å^2)	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% 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	0.29	0/3063	0.45	0/4176
1	B	0.29	0/2894	0.45	0/3949
All	All	0.29	0/5957	0.45	0/8125

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

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	2996	0	2797	24	0
1	B	2831	0	2623	30	0
2	A	6	0	0	0	0
2	B	2	0	0	0	0
All	All	5835	0	5420	41	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 4.

All (41) 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:179:ALA:HB1	1:A:192:LEU:HD11	1.61	0.82

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:215:TRP:CD2	1:B:208:ILE:HD13	2.30	0.67
1:B:179:ALA:HB1	1:B:192:LEU:HD11	1.81	0.63
1:A:230:VAL:HG23	1:A:231:PRO:HD3	1.81	0.62
1:B:230:VAL:HG23	1:B:231:PRO:HD3	1.83	0.60
1:A:215:TRP:CG	1:B:208:ILE:HD11	2.37	0.59
1:A:215:TRP:CD1	1:B:208:ILE:HD11	2.38	0.58
1:B:215:TRP:HE1	1:B:219:ARG:NH2	2.02	0.58
1:A:228:LEU:HD23	1:B:229:LEU:HD21	1.86	0.56
1:A:209:HIS:ND1	1:B:209:HIS:ND1	2.41	0.54
1:A:229:LEU:HD21	1:B:228:LEU:HD23	1.89	0.54
1:A:250:ILE:HD12	1:B:247:HIS:CE1	2.45	0.51
1:B:141:GLU:HG2	1:B:145:GLN:HE21	1.77	0.50
1:A:215:TRP:CE2	1:A:219:ARG:HD2	2.48	0.48
1:A:215:TRP:CE2	1:B:208:ILE:HD13	2.48	0.48
1:A:215:TRP:CG	1:B:208:ILE:CD1	2.96	0.48
1:A:230:VAL:CG2	1:A:231:PRO:HD3	2.44	0.48
1:A:39:LEU:HB3	1:A:42:ASN:HB3	1.96	0.48
1:B:208:ILE:HD12	1:B:210:ILE:HG12	1.96	0.47
1:A:215:TRP:CD2	1:B:208:ILE:CD1	2.97	0.47
1:A:247:HIS:CE1	1:B:250:ILE:HD12	2.49	0.46
1:A:247:HIS:ND1	1:B:247:HIS:ND1	2.61	0.46
1:B:230:VAL:CG2	1:B:231:PRO:HD3	2.46	0.45
1:A:239:ALA:HA	1:A:243:ALA:HB3	1.98	0.45
1:B:144:TRP:CE2	1:B:148:LYS:HD2	2.52	0.45
1:B:233:LEU:O	1:B:237:MET:HG2	2.17	0.45
1:B:239:ALA:HA	1:B:243:ALA:HB3	1.99	0.45
1:B:129:VAL:HB	1:B:132:SER:HB3	1.99	0.44
1:B:140:GLU:H	1:B:140:GLU:CD	2.20	0.44
1:A:233:LEU:O	1:A:237:MET:HG2	2.17	0.44
1:A:129:VAL:HB	1:A:132:SER:HB3	2.00	0.43
1:B:215:TRP:HE1	1:B:219:ARG:CZ	2.31	0.43
1:A:291:GLU:OE2	1:A:304:GLU:OE2	2.36	0.42
1:A:153:LEU:HD22	1:A:161:PHE:HB2	2.01	0.42
1:B:153:LEU:HD22	1:B:161:PHE:HB2	2.01	0.42
1:B:253:SER:HB2	1:B:265:VAL:HG13	2.02	0.42
1:A:250:ILE:HD12	1:B:247:HIS:HE1	1.84	0.41
1:B:128:GLY:HA3	1:B:135:PRO:HD2	2.02	0.41
1:A:33:ILE:HD12	1:A:133:ILE:HG12	2.02	0.41
1:B:97:VAL:HG21	1:B:147:LEU:HD13	2.02	0.41
1:B:215:TRP:NE1	1:B:219:ARG:NH2	2.69	0.40

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	382/423 (90%)	366 (96%)	16 (4%)	0	100	100
1	B	365/423 (86%)	347 (95%)	18 (5%)	0	100	100
All	All	747/846 (88%)	713 (95%)	34 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	299/374 (80%)	291 (97%)	8 (3%)	44	58
1	B	277/374 (74%)	270 (98%)	7 (2%)	47	61
All	All	576/748 (77%)	561 (97%)	15 (3%)	46	59

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	5	GLN
1	A	42	ASN
1	A	50	PHE
1	A	108	ILE
1	A	162	GLU
1	A	192	LEU
1	A	297	SER
1	A	305	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	44	ILE
1	B	50	PHE
1	B	108	ILE
1	B	162	GLU
1	B	192	LEU
1	B	305	ASN
1	B	356	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	42	ASN
1	B	11	ASN
1	B	145	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 monosaccharides 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, 95^{th} 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	384/423 (90%)	0.37	24 (6%)	20	25	54, 81, 136, 173	0
1	B	369/423 (87%)	0.52	30 (8%)	12	16	54, 88, 180, 212	0
All	All	753/846 (89%)	0.45	54 (7%)	15	20	54, 83, 145, 212	0

All (54) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	373	ASN	8.7
1	A	3	GLY	8.2
1	A	4	GLU	7.1
1	B	365	ILE	6.3
1	A	5	GLN	5.5
1	B	374	SER	5.4
1	B	366	GLN	5.3
1	B	362	ASP	5.2
1	A	169	GLU	5.1
1	A	386	ASN	5.0
1	B	360	THR	5.0
1	B	351	TYR	4.7
1	A	345	TYR	4.2
1	B	352	LYS	3.9
1	B	372	VAL	3.9
1	A	383	ARG	3.8
1	B	355	LEU	3.7
1	A	168	ASN	3.7
1	A	365	ILE	3.7
1	A	167	PHE	3.4
1	A	170	THR	3.3
1	B	57	VAL	3.3
1	B	357	SER	3.3
1	B	212	ARG	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	361	ASN	3.2
1	B	356	LEU	3.1
1	B	369	LEU	3.0
1	A	379	ALA	2.8
1	A	363	GLU	2.7
1	A	376	LEU	2.7
1	A	241	PHE	2.7
1	B	215	TRP	2.6
1	A	166	LYS	2.5
1	B	213	LYS	2.5
1	B	364	ASP	2.5
1	B	377	GLU	2.5
1	B	41	ALA	2.5
1	A	385	GLU	2.5
1	B	341	ASN	2.4
1	B	140	GLU	2.4
1	B	17	ALA	2.4
1	B	339	GLU	2.4
1	B	59	ASP	2.3
1	A	377	GLU	2.3
1	A	366	GLN	2.2
1	B	220	TYR	2.2
1	B	376	LEU	2.2
1	A	162	GLU	2.2
1	A	247	HIS	2.1
1	A	26	PHE	2.1
1	A	112	GLU	2.1
1	A	384	GLN	2.0
1	B	170	THR	2.0
1	B	171	LEU	2.0

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.