



Full wwPDB X-ray Structure Validation Report ⓘ

Dec 12, 2022 – 03:28 pm GMT

PDB ID : 7QFJ
Title : Crystal structure of S-layer protein SlpX from *Lactobacillus acidophilus*, domain II (aa 194-362)
Authors : Sagmeister, T.; Pavkov-Keller, T.; Buhlheller, C.; Baek, M.; Read, R.; Baker, D.
Deposited on : 2021-12-06
Resolution : 2.50 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 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 : 2.31.3
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.31.3

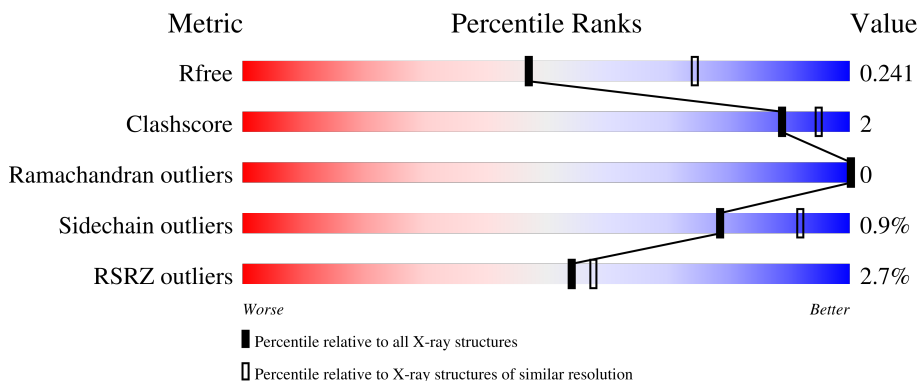
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 $\leq 5\%$. 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	189	 87% 9%
1	B	189	 87% 5% 8%
1	C	189	 86% 5% 9% 2%
1	D	189	 89% 7%
1	E	189	 85% 5% 10% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	189	 <p>A horizontal bar chart representing the quality of the chain. The bar is divided into three segments: a red segment on the left labeled '5%', a large green segment in the middle labeled '89%', and a small grey segment on the right labeled '9%'. A small black dot is visible at the end of the grey segment.</p>

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 8004 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 SlpX.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	172	1293	792	221	280	0	0	0
1	B	174	1305	798	224	283	0	0	0
1	C	172	1293	792	221	280	0	0	0
1	D	176	1321	808	226	287	0	0	0
1	E	170	1283	784	220	279	0	0	0
1	F	172	1293	792	221	280	0	0	0

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	182	MET	-	initiating methionine	UNP Q5FLN0
A	183	GLY	-	expression tag	UNP Q5FLN0
A	363	LEU	-	expression tag	UNP Q5FLN0
A	364	GLU	-	expression tag	UNP Q5FLN0
A	365	HIS	-	expression tag	UNP Q5FLN0
A	366	HIS	-	expression tag	UNP Q5FLN0
A	367	HIS	-	expression tag	UNP Q5FLN0
A	368	HIS	-	expression tag	UNP Q5FLN0
A	369	HIS	-	expression tag	UNP Q5FLN0
A	370	HIS	-	expression tag	UNP Q5FLN0
B	182	MET	-	initiating methionine	UNP Q5FLN0
B	183	GLY	-	expression tag	UNP Q5FLN0
B	363	LEU	-	expression tag	UNP Q5FLN0
B	364	GLU	-	expression tag	UNP Q5FLN0
B	365	HIS	-	expression tag	UNP Q5FLN0
B	366	HIS	-	expression tag	UNP Q5FLN0
B	367	HIS	-	expression tag	UNP Q5FLN0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	368	HIS	-	expression tag	UNP Q5FLN0
B	369	HIS	-	expression tag	UNP Q5FLN0
B	370	HIS	-	expression tag	UNP Q5FLN0
C	182	MET	-	initiating methionine	UNP Q5FLN0
C	183	GLY	-	expression tag	UNP Q5FLN0
C	363	LEU	-	expression tag	UNP Q5FLN0
C	364	GLU	-	expression tag	UNP Q5FLN0
C	365	HIS	-	expression tag	UNP Q5FLN0
C	366	HIS	-	expression tag	UNP Q5FLN0
C	367	HIS	-	expression tag	UNP Q5FLN0
C	368	HIS	-	expression tag	UNP Q5FLN0
C	369	HIS	-	expression tag	UNP Q5FLN0
C	370	HIS	-	expression tag	UNP Q5FLN0
D	182	MET	-	initiating methionine	UNP Q5FLN0
D	183	GLY	-	expression tag	UNP Q5FLN0
D	363	LEU	-	expression tag	UNP Q5FLN0
D	364	GLU	-	expression tag	UNP Q5FLN0
D	365	HIS	-	expression tag	UNP Q5FLN0
D	366	HIS	-	expression tag	UNP Q5FLN0
D	367	HIS	-	expression tag	UNP Q5FLN0
D	368	HIS	-	expression tag	UNP Q5FLN0
D	369	HIS	-	expression tag	UNP Q5FLN0
D	370	HIS	-	expression tag	UNP Q5FLN0
E	182	MET	-	initiating methionine	UNP Q5FLN0
E	183	GLY	-	expression tag	UNP Q5FLN0
E	363	LEU	-	expression tag	UNP Q5FLN0
E	364	GLU	-	expression tag	UNP Q5FLN0
E	365	HIS	-	expression tag	UNP Q5FLN0
E	366	HIS	-	expression tag	UNP Q5FLN0
E	367	HIS	-	expression tag	UNP Q5FLN0
E	368	HIS	-	expression tag	UNP Q5FLN0
E	369	HIS	-	expression tag	UNP Q5FLN0
E	370	HIS	-	expression tag	UNP Q5FLN0
F	182	MET	-	initiating methionine	UNP Q5FLN0
F	183	GLY	-	expression tag	UNP Q5FLN0
F	363	LEU	-	expression tag	UNP Q5FLN0
F	364	GLU	-	expression tag	UNP Q5FLN0
F	365	HIS	-	expression tag	UNP Q5FLN0
F	366	HIS	-	expression tag	UNP Q5FLN0
F	367	HIS	-	expression tag	UNP Q5FLN0
F	368	HIS	-	expression tag	UNP Q5FLN0
F	369	HIS	-	expression tag	UNP Q5FLN0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	370	HIS	-	expression tag	UNP Q5FLN0


- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	41	Total O 41 41	0	0
2	B	49	Total O 49 49	0	0
2	C	33	Total O 33 33	0	0
2	D	54	Total O 54 54	0	0
2	E	21	Total O 21 21	0	0
2	F	18	Total O 18 18	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: SlpX

Chain A:  87% 9%




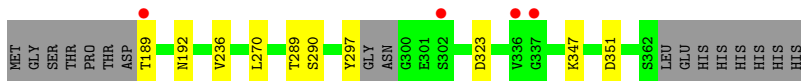
- Molecule 1: SlpX

Chain B:  87% 5% 8%




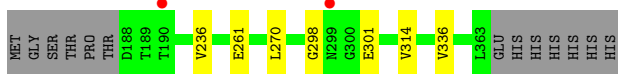
- Molecule 1: SlpX

Chain C:  86% 5% 9%




- Molecule 1: SlpX

Chain D:  89% 7%

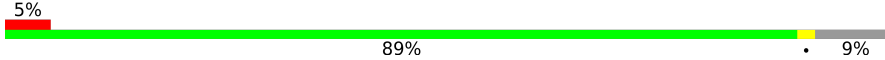


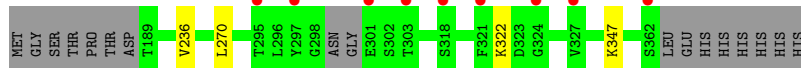
- Molecule 1: SlpX

Chain E:  85% 5% 10%



- Molecule 1: SlpX

Chain F:  5% 89% 9%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	65.68Å 143.02Å 73.44Å 90.00° 96.70° 90.00°	Depositor
Resolution (Å)	46.03 – 2.50 46.03 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.9 (46.03-2.50) 99.9 (46.03-2.50)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.09 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.206 , 0.242 0.205 , 0.241	Depositor DCC
R_{free} test set	2290 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å ²)	40.7	Xtriage
Anisotropy	0.073	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	8004	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.68% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

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	0.40	0/1312	0.73	0/1787
1	B	0.39	0/1325	0.69	0/1806
1	C	0.41	0/1312	0.75	1/1787 (0.1%)
1	D	0.44	1/1341 (0.1%)	0.70	0/1828
1	E	0.39	0/1302	0.73	0/1773
1	F	0.39	0/1312	0.68	0/1787
All	All	0.41	1/7904 (0.0%)	0.72	1/10768 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	261	GLU	CD-OE2	6.91	1.33	1.25

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	289	THR	OG1-CB-CG2	-5.03	98.44	110.00

There are no chirality outliers.

There are no planarity outliers.

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	1293	0	1240	7	0
1	B	1305	0	1250	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1293	0	1240	5	0
1	D	1321	0	1265	6	0
1	E	1283	0	1225	7	0
1	F	1293	0	1240	4	0
2	A	41	0	0	0	0
2	B	49	0	0	1	0
2	C	33	0	0	0	0
2	D	54	0	0	0	0
2	E	21	0	0	0	0
2	F	18	0	0	0	0
All	All	8004	0	7460	29	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 2.

All (29) 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:D:314:VAL:CG2	1:D:336:VAL:HG11	2.00	0.92
1:A:253:THR:HG21	1:A:271:THR:OG1	1.75	0.86
1:E:322:LYS:HE2	1:F:347:LYS:HZ3	1.49	0.77
1:D:314:VAL:HG23	1:D:336:VAL:HG11	1.68	0.74
1:E:322:LYS:HE2	1:F:347:LYS:NZ	2.03	0.72
1:D:314:VAL:CG2	1:D:336:VAL:CG1	2.69	0.71
1:D:314:VAL:HG21	1:D:336:VAL:CG1	2.27	0.64
1:A:253:THR:CG2	1:A:271:THR:OG1	2.52	0.55
1:C:189:THR:HB	1:C:192:ASN:ND2	2.23	0.53
1:A:253:THR:CG2	1:A:271:THR:CB	2.89	0.50
1:A:297:TYR:HA	1:A:323:ASP:OD1	2.13	0.48
1:E:322:LYS:HD3	1:F:347:LYS:HZ1	1.80	0.47
1:D:298:GLY:O	1:D:301:GLU:N	2.50	0.46
1:C:347:LYS:HE3	1:C:351:ASP:OD2	2.16	0.45
1:A:253:THR:CG2	1:A:271:THR:HB	2.47	0.45
1:E:298:GLY:HA3	1:E:301:GLU:O	2.17	0.45
1:E:361:THR:O	1:E:362:SER:CB	2.64	0.45
1:A:253:THR:HG21	1:A:271:THR:CB	2.47	0.44
1:C:189:THR:HB	1:C:192:ASN:HD22	1.82	0.44
1:E:297:TYR:HA	1:E:323:ASP:OD1	2.17	0.44
1:A:357:ILE:HD11	1:B:303:THR:HG21	1.99	0.44
1:C:297:TYR:HA	1:C:323:ASP:OD1	2.17	0.43
1:B:334:LYS:HD3	2:B:442:HOH:O	2.18	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:236:VAL:HG11	1:E:270:LEU:CD1	2.49	0.43
1:F:236:VAL:HG11	1:F:270:LEU:CD1	2.49	0.43
1:D:236:VAL:HG11	1:D:270:LEU:CD1	2.49	0.42
1:C:236:VAL:HG11	1:C:270:LEU:CD1	2.50	0.42
1:B:297:TYR:HA	1:B:323:ASP:OD1	2.19	0.42
1:B:236:VAL:HG11	1:B:270:LEU:CD1	2.51	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	168/189 (89%)	166 (99%)	2 (1%)	0	100	100
1	B	172/189 (91%)	169 (98%)	3 (2%)	0	100	100
1	C	168/189 (89%)	165 (98%)	3 (2%)	0	100	100
1	D	174/189 (92%)	171 (98%)	3 (2%)	0	100	100
1	E	166/189 (88%)	163 (98%)	3 (2%)	0	100	100
1	F	168/189 (89%)	166 (99%)	2 (1%)	0	100	100
All	All	1016/1134 (90%)	1000 (98%)	16 (2%)	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	150/165 (91%)	148 (99%)	2 (1%)	69	87
1	B	151/165 (92%)	148 (98%)	3 (2%)	55	79
1	C	150/165 (91%)	149 (99%)	1 (1%)	84	94
1	D	153/165 (93%)	153 (100%)	0	100	100
1	E	149/165 (90%)	148 (99%)	1 (1%)	84	94
1	F	150/165 (91%)	149 (99%)	1 (1%)	84	94
All	All	903/990 (91%)	895 (99%)	8 (1%)	78	92

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	301	GLU
1	A	334	LYS
1	B	242	LYS
1	B	254	SER
1	B	322	LYS
1	C	290	SER
1	E	320	THR
1	F	322	LYS

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

Mol	Chain	Res	Type
1	B	202	GLN
1	D	233	ASN
1	F	317	ASN

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 monosaccharides 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

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	172/189 (91%)	-0.02	0 100 100	24, 36, 68, 103	0
1	B	174/189 (92%)	0.16	1 (0%) 89 90	24, 38, 68, 94	0
1	C	172/189 (91%)	0.18	4 (2%) 60 63	25, 48, 86, 119	0
1	D	176/189 (93%)	0.13	2 (1%) 80 82	23, 35, 77, 110	0
1	E	170/189 (89%)	0.33	12 (7%) 16 16	29, 53, 99, 120	0
1	F	172/189 (91%)	0.41	9 (5%) 27 29	29, 55, 93, 126	0
All	All	1036/1134 (91%)	0.20	28 (2%) 54 58	23, 43, 87, 126	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	297	TYR	5.1
1	E	347	LYS	3.8
1	F	321	PHE	3.6
1	D	190	THR	3.2
1	C	189	THR	3.1
1	F	301	GLU	3.1
1	E	302	SER	3.0
1	F	324	GLY	3.0
1	C	337	GLY	2.8
1	F	327	VAL	2.8
1	E	312	ASN	2.8
1	E	335	THR	2.8
1	E	300	GLY	2.7
1	F	318	SER	2.5
1	D	299	ASN	2.5
1	F	362	SER	2.4
1	E	332	GLN	2.4
1	F	295	THR	2.4
1	E	328	TYR	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	336	VAL	2.3
1	C	302	SER	2.2
1	E	340	SER	2.2
1	E	189	THR	2.2
1	E	334	LYS	2.2
1	B	190	THR	2.1
1	E	304	ILE	2.1
1	F	303	THR	2.0
1	E	190	THR	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 [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.