



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2026 – 03:54 AM UTC

PDB ID : 5YU5 / pdb_00005yu5
Title : Crystal structure of shaft pilin spaD from Lactobacillus rhamnosus GG
Authors : Chaurasia, P.; Pratap, S.; Palva, A.; von Ossowski, I.; Krishnan, V.
Deposited on : 2017-11-20
Resolution : 2.27 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

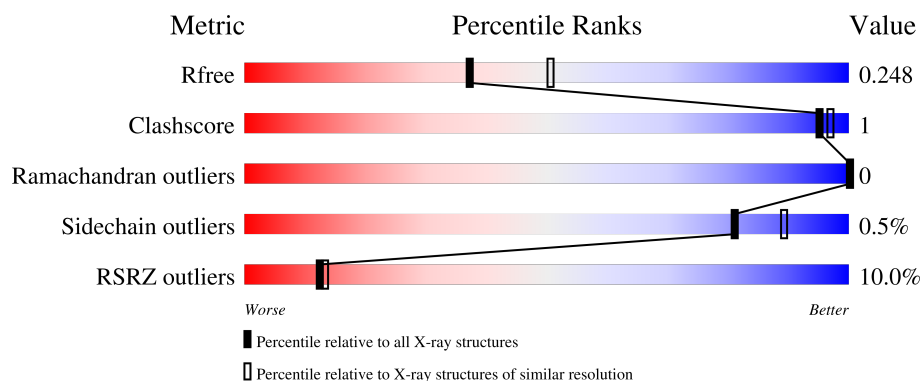
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.27 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}	180053	9078 (2.30-2.26)
Clashscore	190562	9802 (2.30-2.26)
Ramachandran outliers	187476	9690 (2.30-2.26)
Sidechain outliers	187428	9691 (2.30-2.26)
RSRZ outliers	180081	9085 (2.30-2.26)

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	465	11% 92% 5%
1	B	465	8% 94%

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 7222 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 Pilus assembly protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	444	Total	C	N	O	S	0	0	0
			3420	2157	578	677	8			
1	B	448	Total	C	N	O	S	0	0	0
			3443	2171	580	684	8			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	29	MET	-	expression tag	UNP A0A179XFF5
A	30	GLY	-	expression tag	UNP A0A179XFF5
A	31	ARG	-	expression tag	UNP A0A179XFF5
A	32	ASP	-	expression tag	UNP A0A179XFF5
A	33	PRO	-	expression tag	UNP A0A179XFF5
A	34	ASN	-	expression tag	UNP A0A179XFF5
A	35	SER	-	expression tag	UNP A0A179XFF5
A	486	LEU	-	expression tag	UNP A0A179XFF5
A	487	GLU	-	expression tag	UNP A0A179XFF5
A	488	HIS	-	expression tag	UNP A0A179XFF5
A	489	HIS	-	expression tag	UNP A0A179XFF5
A	490	HIS	-	expression tag	UNP A0A179XFF5
A	491	HIS	-	expression tag	UNP A0A179XFF5
A	492	HIS	-	expression tag	UNP A0A179XFF5
A	493	HIS	-	expression tag	UNP A0A179XFF5
B	29	MET	-	expression tag	UNP A0A179XFF5
B	30	GLY	-	expression tag	UNP A0A179XFF5
B	31	ARG	-	expression tag	UNP A0A179XFF5
B	32	ASP	-	expression tag	UNP A0A179XFF5
B	33	PRO	-	expression tag	UNP A0A179XFF5
B	34	ASN	-	expression tag	UNP A0A179XFF5
B	35	SER	-	expression tag	UNP A0A179XFF5
B	486	LEU	-	expression tag	UNP A0A179XFF5
B	487	GLU	-	expression tag	UNP A0A179XFF5
B	488	HIS	-	expression tag	UNP A0A179XFF5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	489	HIS	-	expression tag	UNP A0A179XFF5
B	490	HIS	-	expression tag	UNP A0A179XFF5
B	491	HIS	-	expression tag	UNP A0A179XFF5
B	492	HIS	-	expression tag	UNP A0A179XFF5
B	493	HIS	-	expression tag	UNP A0A179XFF5

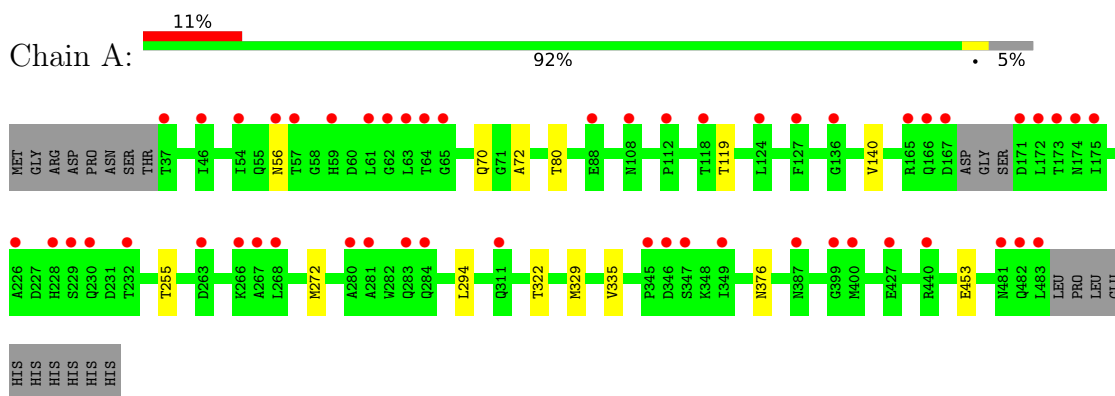
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	148	Total	O	0	0
			148	148		
2	B	211	Total	O	0	0
			211	211		

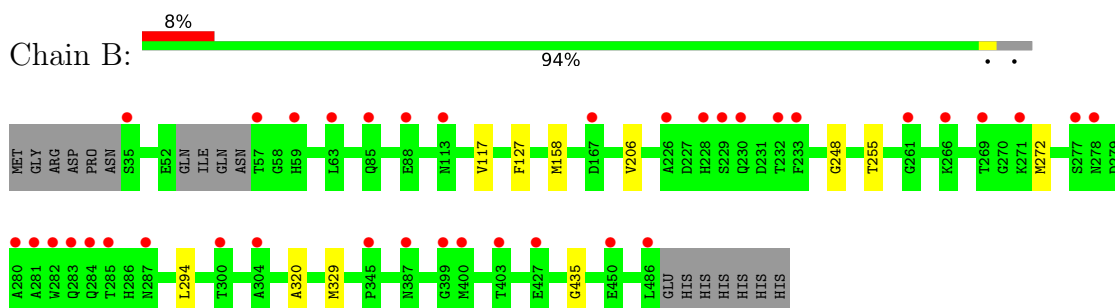
3 Residue-property plots [i](#)

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: Pilus assembly protein



- Molecule 1: Pilus assembly protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	54.41 Å 76.26 Å 400.70 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	200.00 – 2.27 200.00 – 2.27	Depositor EDS
% Data completeness (in resolution range)	50.1 (200.00-2.27) 50.1 (200.00-2.27)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.08 (at 2.27 Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.202 , 0.236 0.214 , 0.248	Depositor DCC
R_{free} test set	1995 reflections (2.55%)	wwPDB-VP
Wilson B-factor (Å ²)	39.1	Xtriage
Anisotropy	0.175	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 35.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	7222	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.32% 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	0.52	0/3486	0.72	0/4742
1	B	0.51	0/3510	0.72	0/4776
All	All	0.52	0/6996	0.72	0/9518

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 [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	3420	0	3340	9	0
1	B	3443	0	3364	7	0
2	A	148	0	0	0	0
2	B	211	0	0	0	0
All	All	7222	0	6704	16	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 1.

All (16) 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:80:THR:HA	1:A:140:VAL:HG12	1.59	0.84

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:80:THR:HA	1:A:140:VAL:CG1	2.23	0.67
1:A:255:THR:HG21	1:A:322:THR:HG23	1.77	0.65
1:B:272:MET:SD	1:B:294:LEU:HD11	2.49	0.52
1:A:72:ALA:N	1:A:119:THR:HG22	2.25	0.52
1:B:248:GLY:HA3	1:B:329:MET:HE2	1.94	0.48
1:A:272:MET:SD	1:A:294:LEU:HD11	2.53	0.48
1:B:255:THR:HG22	1:B:320:ALA:C	2.40	0.47
1:A:70:GLN:HA	1:A:119:THR:HG23	1.97	0.47
1:B:117:VAL:HG22	1:B:127:PHE:CZ	2.50	0.46
1:A:329:MET:HE1	1:A:335:VAL:HG21	1.99	0.44
1:B:248:GLY:HA3	1:B:329:MET:CE	2.48	0.44
1:B:206:VAL:HG13	1:B:435:GLY:CA	2.49	0.42
1:A:72:ALA:H	1:A:119:THR:HG22	1.83	0.41
1:A:329:MET:HE1	1:A:335:VAL:CG2	2.50	0.41
1:B:206:VAL:HG13	1:B:435:GLY:HA3	2.02	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	440/465 (95%)	428 (97%)	12 (3%)	0	100	100
1	B	444/465 (96%)	431 (97%)	13 (3%)	0	100	100
All	All	884/930 (95%)	859 (97%)	25 (3%)	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	367/387 (95%)	364 (99%)	3 (1%)	73	84
1	B	371/387 (96%)	370 (100%)	1 (0%)	86	92
All	All	738/774 (95%)	734 (100%)	4 (0%)	81	89

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

Mol	Chain	Res	Type
1	A	56	ASN
1	A	376	ASN
1	A	453	GLU
1	B	158	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	53	GLN
1	A	101	GLN
1	A	166	GLN
1	A	220	HIS
1	A	236	ASN
1	A	283	GLN
1	A	284	GLN
1	A	311	GLN
1	A	333	GLN
1	A	357	GLN
1	A	376	ASN
1	B	166	GLN
1	B	236	ASN
1	B	382	GLN
1	B	398	ASN
1	B	481	ASN

5.3.3 RNA ⓘ

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

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	444/465 (95%)	0.82	52 (11%) 9 10	30, 53, 84, 101	0
1	B	448/465 (96%)	0.63	37 (8%) 17 18	21, 46, 78, 98	0
All	All	892/930 (95%)	0.73	89 (9%) 12 13	21, 50, 82, 101	0

All (89) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	54	ILE	6.6
1	A	63	LEU	6.1
1	B	281	ALA	5.8
1	B	57	THR	5.4
1	A	483	LEU	5.3
1	A	171	ASP	4.9
1	A	481	ASN	4.9
1	A	166	GLN	4.9
1	A	167	ASP	4.8
1	B	63	LEU	4.5
1	A	37	THR	4.4
1	A	228	HIS	4.3
1	B	280	ALA	4.0
1	B	399	GLY	3.9
1	A	173	THR	3.9
1	B	283	GLN	3.9
1	A	65	GLY	3.7
1	A	174	ASN	3.7
1	B	284	GLN	3.7
1	A	266	LYS	3.5
1	B	427	GLU	3.4
1	B	285	THR	3.3
1	B	229	SER	3.2
1	B	278	ASN	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	112	PRO	3.1
1	B	400	MET	3.1
1	B	35	SER	3.1
1	A	399	GLY	3.1
1	B	345	PRO	3.1
1	B	387	ASN	3.0
1	A	59	HIS	3.0
1	B	230	GLN	3.0
1	A	56	ASN	3.0
1	A	345	PRO	3.0
1	A	400	MET	2.9
1	A	230	GLN	2.8
1	B	228	HIS	2.8
1	A	118	THR	2.8
1	A	283	GLN	2.8
1	A	64	THR	2.7
1	B	59	HIS	2.7
1	A	427	GLU	2.7
1	B	266	LYS	2.7
1	B	261	GLY	2.7
1	A	46	ILE	2.6
1	A	61	LEU	2.6
1	A	136	GLY	2.6
1	A	346	ASP	2.6
1	A	226	ALA	2.5
1	A	57	THR	2.5
1	A	232	THR	2.5
1	B	486	LEU	2.5
1	B	85	GLN	2.5
1	B	300	THR	2.4
1	B	167	ASP	2.4
1	A	268	LEU	2.4
1	A	263	ASP	2.4
1	B	88	GLU	2.4
1	B	287	ASN	2.4
1	A	124	LEU	2.4
1	A	267	ALA	2.3
1	B	233	PHE	2.3
1	A	347	SER	2.3
1	A	440	ARG	2.3
1	B	403	THR	2.3
1	A	62	GLY	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	281	ALA	2.3
1	A	482	GLN	2.3
1	A	284	GLN	2.2
1	B	450	GLU	2.2
1	A	108	ASN	2.2
1	A	387	ASN	2.2
1	A	280	ALA	2.2
1	B	304	ALA	2.2
1	A	229	SER	2.2
1	B	277	SER	2.2
1	B	226	ALA	2.2
1	B	232	THR	2.2
1	A	127	PHE	2.2
1	A	311	GLN	2.2
1	A	175	ILE	2.1
1	B	269	THR	2.1
1	A	172	LEU	2.1
1	B	282	TRP	2.1
1	B	271	LYS	2.0
1	A	88	GLU	2.0
1	B	113	ASN	2.0
1	A	165	ARG	2.0
1	A	349	ILE	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 oligosaccharides 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.