



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 10:44 PM GMT

PDB ID : 5C76
Title : ATP-driven lipid-linked oligosaccharide flippase PglK in apo-inward facing state (2)
Authors : Perez, C.; Gerber, S.; Locher, K.P.
Deposited on : 2015-06-24
Resolution : 3.94 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
A user guide is available at
<http://wwpdb.org/validation/2016/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
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20026982

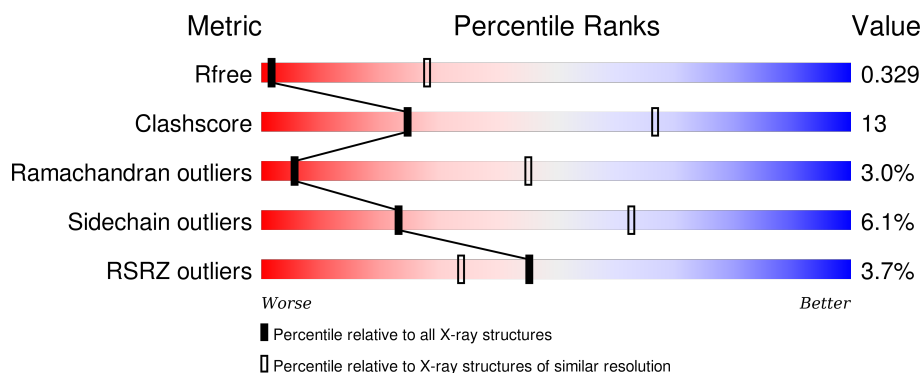
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.94 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}	91344	1007 (4.34-3.54)
Clashscore	102246	1042 (4.30-3.58)
Ramachandran outliers	100387	1000 (4.30-3.58)
Sidechain outliers	100360	1021 (4.32-3.56)
RSRZ outliers	91569	1011 (4.34-3.54)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	564	<div> <div>2%</div> <div>64%</div> <div>33%</div> <div>.</div> </div>
1	B	564	<div> <div>4%</div> <div>65%</div> <div>32%</div> <div>..</div> </div>
1	C	564	<div> <div>4%</div> <div>65%</div> <div>31%</div> <div>.</div> </div>
1	D	564	<div> <div>5%</div> <div>67%</div> <div>30%</div> <div>.</div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 18228 atoms, of which 0 are hydrogens and 0 are deuteriums.

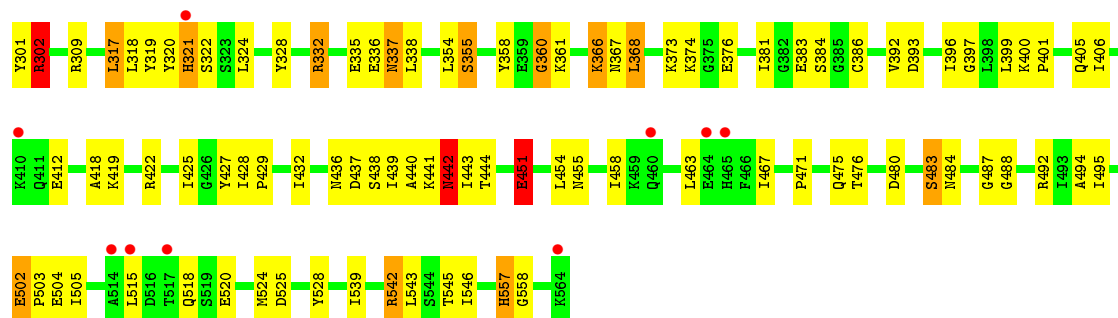
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called WlaB protein.

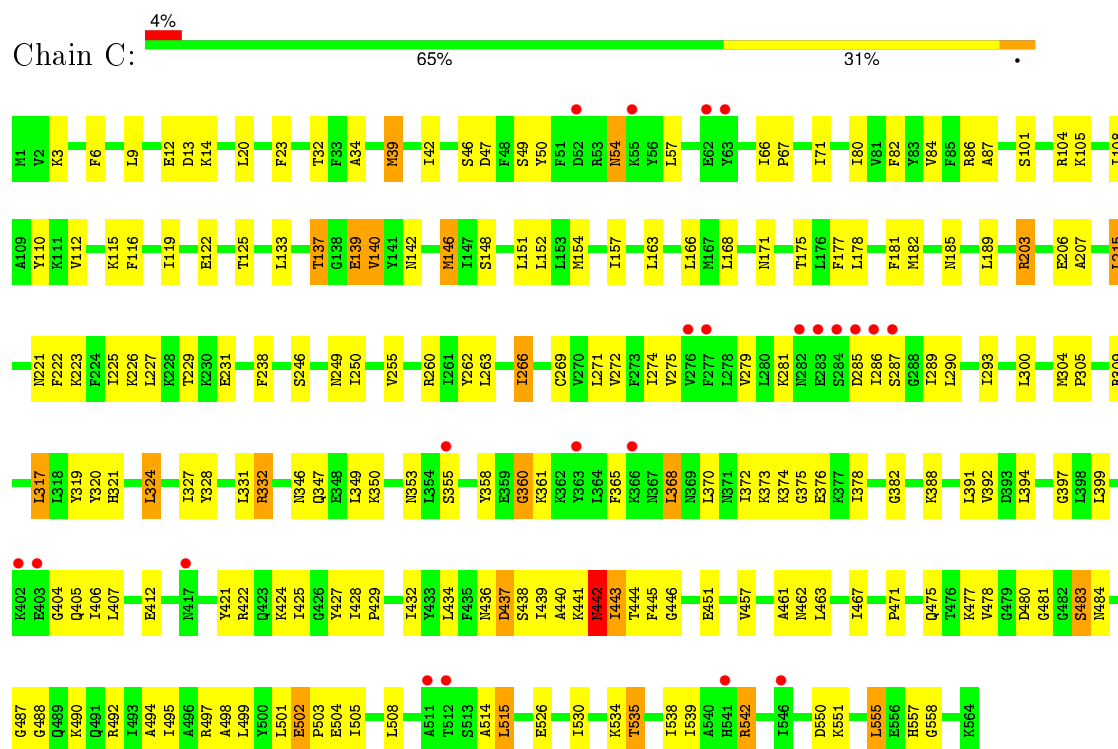
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	564	Total	C	N	O	S	0	0	0
			4557	3000	732	811	14			
1	B	564	Total	C	N	O	S	0	0	0
			4557	3000	732	811	14			
1	C	564	Total	C	N	O	S	0	0	0
			4557	3000	732	811	14			
1	D	564	Total	C	N	O	S	0	0	0
			4557	3000	732	811	14			

There are 12 discrepancies between the modelled and reference sequences:

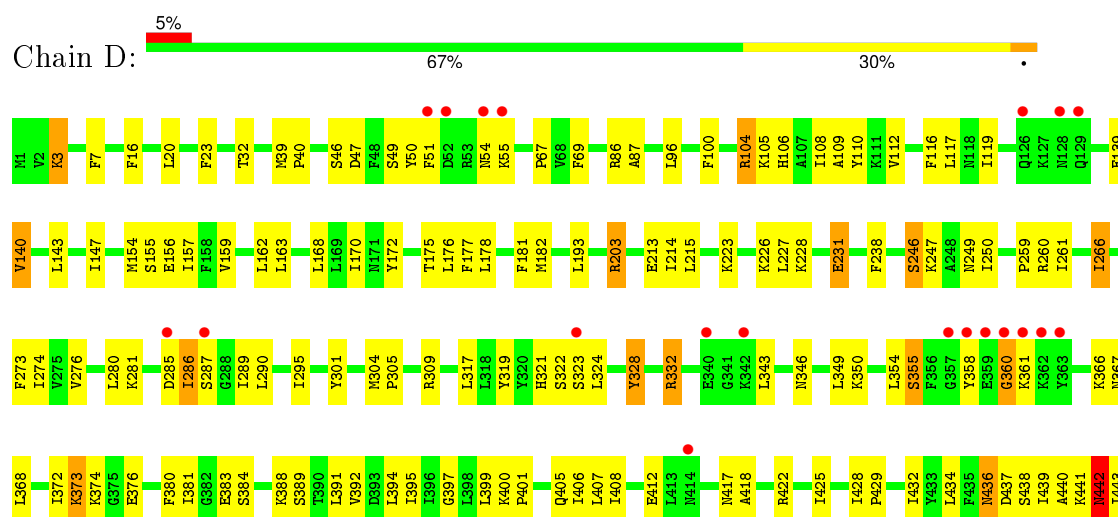
Chain	Residue	Modelled	Actual	Comment	Reference
A	2	VAL	LEU	conflict	UNP O86150
A	105	LYS	TYR	conflict	UNP O86150
A	510	GLN	GLU	conflict	UNP O86150
B	2	VAL	LEU	conflict	UNP O86150
B	105	LYS	TYR	conflict	UNP O86150
B	510	GLN	GLU	conflict	UNP O86150
C	2	VAL	LEU	conflict	UNP O86150
C	105	LYS	TYR	conflict	UNP O86150
C	510	GLN	GLU	conflict	UNP O86150
D	2	VAL	LEU	conflict	UNP O86150
D	105	LYS	TYR	conflict	UNP O86150
D	510	GLN	GLU	conflict	UNP O86150

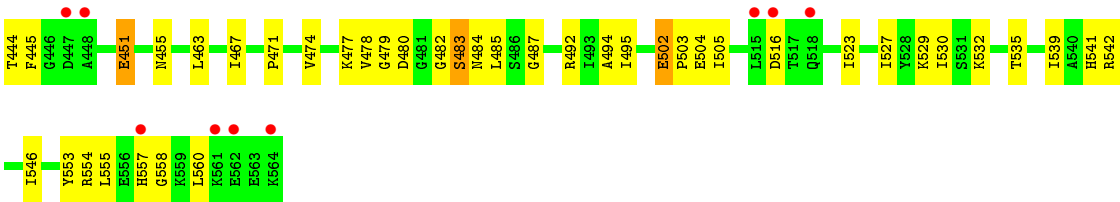


• Molecule 1: WlaB protein



• Molecule 1: WlaB protein





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	93.41Å 183.33Å 136.26Å 90.00° 106.70° 90.00°	Depositor
Resolution (Å)	29.82 – 3.94 29.82 – 3.94	Depositor EDS
% Data completeness (in resolution range)	86.2 (29.82-3.94) 86.3 (29.82-3.94)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.83 (at 3.98Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.275 , 0.329 0.274 , 0.329	Depositor DCC
R_{free} test set	1693 reflections (5.06%)	DCC
Wilson B-factor (Å ²)	166.3	Xtriage
Anisotropy	0.011	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , 111.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 33489 reflections	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	18228	wwPDB-VP
Average B, all atoms (Å ²)	208.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.97% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.29	0/4641	0.50	0/6246
1	B	0.28	0/4641	0.49	1/6246 (0.0%)
1	C	0.28	0/4641	0.48	0/6246
1	D	0.27	0/4641	0.47	0/6246
All	All	0.28	0/18564	0.48	1/24984 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	302	ARG	NE-CZ-NH2	5.11	122.86	120.30

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	4557	0	4764	141	0
1	B	4557	0	4764	130	0
1	C	4557	0	4764	131	0
1	D	4557	0	4764	128	0
All	All	18228	0	19056	476	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 13.

The worst 5 of 476 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:520:GLU:OE1	1:B:542:ARG:NH1	2.09	0.85
1:B:503:PRO:HB2	1:B:504:GLU:HG2	1.61	0.82
1:A:425:ILE:HG22	1:A:505:ILE:HB	1.62	0.81
1:C:503:PRO:HB2	1:C:504:GLU:HG2	1.62	0.81
1:D:425:ILE:HG22	1:D:505:ILE:HB	1.61	0.81

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	562/564 (100%)	478 (85%)	68 (12%)	16 (3%)	6	47
1	B	562/564 (100%)	478 (85%)	65 (12%)	19 (3%)	5	43
1	C	562/564 (100%)	478 (85%)	69 (12%)	15 (3%)	6	47
1	D	562/564 (100%)	477 (85%)	68 (12%)	17 (3%)	5	45
All	All	2248/2256 (100%)	1911 (85%)	270 (12%)	67 (3%)	5	45

5 of 67 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	321	HIS
1	A	436	ASN
1	A	439	ILE
1	A	442	ASN
1	A	443	ILE

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	505/505 (100%)	475 (94%)	30 (6%)	24	64
1	B	505/505 (100%)	472 (94%)	33 (6%)	21	61
1	C	505/505 (100%)	474 (94%)	31 (6%)	23	63
1	D	505/505 (100%)	476 (94%)	29 (6%)	25	65
All	All	2020/2020 (100%)	1897 (94%)	123 (6%)	23	63

5 of 123 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	437	ASP
1	C	139	GLU
1	D	328	TYR
1	B	451	GLU
1	B	557	HIS

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

Mol	Chain	Res	Type
1	A	218	ASN
1	B	241	GLN
1	D	329	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 carbohydrates 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, 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	564/564 (100%)	-0.21	13 (2%) 64 52	122, 177, 272, 322	0
1	B	564/564 (100%)	-0.22	20 (3%) 48 36	156, 204, 285, 320	0
1	C	564/564 (100%)	-0.07	22 (3%) 43 32	139, 215, 288, 326	0
1	D	564/564 (100%)	-0.04	29 (5%) 32 23	127, 218, 286, 328	0
All	All	2256/2256 (100%)	-0.14	84 (3%) 45 34	122, 202, 284, 328	0

The worst 5 of 84 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	52	ASP	8.3
1	B	564	LYS	7.5
1	C	541	HIS	7.5
1	B	287	SER	7.2
1	A	282	ASN	6.6

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 carbohydrates 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.