WORLDWIDE PDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Nov 11, 2014 – 08:07 AM EST

PDB ID : 4RK4

Title : Crystal structure of LacI family transcriptional regulator from Lactobacillus casei, Target EFI-512911, with bound glucose

Authors : Patskovsky, Y.; Toro, R.; Bhosle, R.; Al Obaidi, N.; Chamala, S.; Attonito, J.D.; Scott Glenn, A.; Chowdhury, S.; Lafleur, J.; Siedel, R.D.; Hillerich, B.; Love, J.; Whalen, K.L.; Gerlt, J.A.; Almo, S.C.; Enzyme Function Initiative (EFI)

Deposited on : 2014-10-11

Resolution : 1.32 Å(reported)

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 1.32 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		0.165
Clashscore		2
Ramachandran outliers		0.4%
Sidechain outliers		0
RSRZ outliers		4.1%

Worse   *Better*

▌Percentile relative to all X–ray structures

▯Percentile relative to X–ray structures of similar resolution
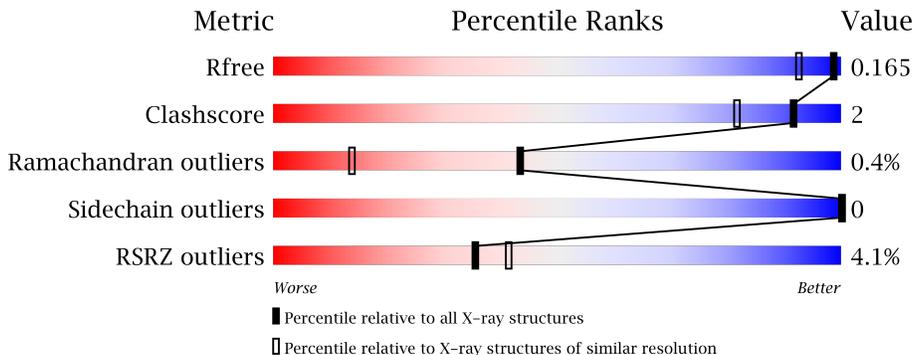
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1157 (1.36-1.28)
Clashscore	79885	1260 (1.36-1.28)
Ramachandran outliers	78287	1218 (1.36-1.28)
Sidechain outliers	78261	1218 (1.36-1.28)
RSRZ outliers	66119	1157 (1.36-1.28)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	283	

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 2392 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Transcriptional regulator, LacI family.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	270	Total 2095	C 1333	N 369	O 388	S 5	0	9	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	51	GLN	-	EXPRESSION TAG	UNP Q036L9
A	52	SER	-	EXPRESSION TAG	UNP Q036L9
A	53	MET	-	EXPRESSION TAG	UNP Q036L9

- Molecule 2 is SUGAR (ALPHA-D-GLUCOSE) (three-letter code: GLC) (formula: $C_6H_{12}O_6$).
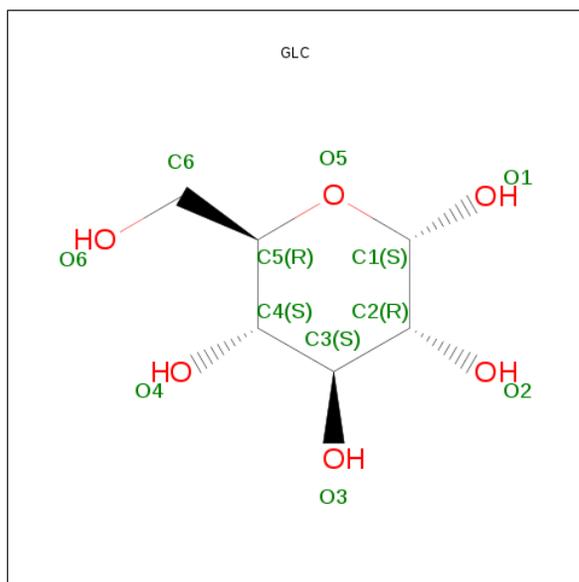
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 12	C 6	O 6	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Na 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	284	Total 284	O 284	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Transcriptional regulator, LacI family

Chain A:
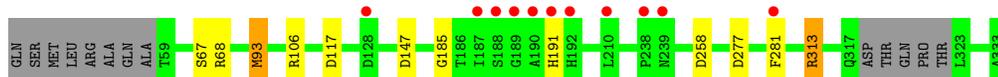
4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	108.85Å 108.85Å 125.89Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 1.32 34.28 – 1.32	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.00-1.32) 99.7 (34.28-1.32)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	1.88 (at 1.32Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.126 , 0.152 0.145 , 0.165	Depositor DCC
R_{free} test set	2115 reflections (3.16%)	DCC
Wilson B-factor (Å^2)	17.6	Xtriage
Anisotropy	0.072	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 46.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Outliers	0 of 67029 reflections	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	2392	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.59% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NA, GLC

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	1.15	4/2163 (0.2%)	1.15	9/2938 (0.3%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	67	SER	CB-OG	-7.84	1.32	1.42
1	A	281	PHE	CB-CG	-7.13	1.39	1.51
1	A	93	MET	CB-CG	6.06	1.70	1.51
1	A	67	SER	CA-CB	-5.44	1.44	1.52

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	93	MET	CG-SD-CE	-17.48	72.23	100.20
1	A	68	ARG	NE-CZ-NH2	-13.04	113.78	120.30
1	A	313[A]	ARG	NE-CZ-NH1	-9.70	115.45	120.30
1	A	313[B]	ARG	NE-CZ-NH1	-9.70	115.45	120.30
1	A	117	ASP	CB-CG-OD2	-8.32	110.81	118.30
1	A	68	ARG	NE-CZ-NH1	6.21	123.41	120.30
1	A	147	ASP	CB-CG-OD1	5.72	123.45	118.30
1	A	106	ARG	NE-CZ-NH2	-5.69	117.46	120.30
1	A	258	ASP	CB-CG-OD2	-5.53	113.32	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2095	0	2111	8	0
2	A	12	0	12	0	0
3	A	1	0	0	0	0
4	A	284	0	0	0	0
All	All	2392	0	2123	8	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

All (8) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:93:MET:CE	1:A:93:MET:SD	2.02	1.48
1:A:93:MET:CG	1:A:93:MET:CE	2.30	1.09
1:A:93:MET:CG	1:A:93:MET:HE3	2.06	0.85
1:A:93:MET:HE3	1:A:93:MET:HG2	1.58	0.85
1:A:93:MET:CB	1:A:93:MET:CE	2.80	0.59
1:A:185:GLY:HA3	1:A:191:HIS:HB2	1.91	0.51
1:A:93:MET:HE2	1:A:93:MET:HB2	2.02	0.40
1:A:313[B]:ARG:HD3	1:A:313[B]:ARG:HH21	1.74	0.40

There are no symmetry-related clashes.

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	275/283 (97%)	270 (98%)	4 (2%)	1 (0%)	43	14

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	277	ASP

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	222/226 (98%)	222 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	324	HIS

5.3.3   RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

5.6   Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	GLC	A	401	-	12,12,12	0.97	1 (8%)	17,17,17	1.16	1 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GLC	A	401	-	-	0/2/22/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	GLC	O5-C1	-2.34	1.38	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	401	GLC	O5-C1-C2	3.38	115.29	109.80

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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(Å²)	Q<0.9
1	A	270/283 (95%)	0.05	11 (4%)	35	41	17, 23, 43, 65	1 (0%)

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	190	ALA	6.9
1	A	210	LEU	5.9
1	A	192	HIS	4.9
1	A	191	HIS	3.6
1	A	238	PRO	3.6
1	A	188	SER	3.5
1	A	189	GLY	3.3
1	A	187	ILE	2.9
1	A	239	ASN	2.7
1	A	281	PHE	2.7
1	A	128	ASP	2.2

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 carbohydrates in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains.

The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	NA	A	402	1/1	0.05	-0.30	25,25,25,25	0
2	GLC	A	401	12/12	0.06	-0.48	17,18,21,22	0

6.5 Other polymers ⓘ

There are no such residues in this entry.