



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 28, 2016 – 05:42 PM EST

PDB ID : 5C9O
Title : Crystal structure of recombinant PLL lectin from Photorhabdus luminescens at 1.5 Å resolution
Authors : Kumar, A.; Sykorova, P.; Demo, G.; Dobes, P.; Hyršl, P.; Wimmerova, M.
Deposited on : 2015-06-28
Resolution : 1.50 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 : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20028320
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-20028320

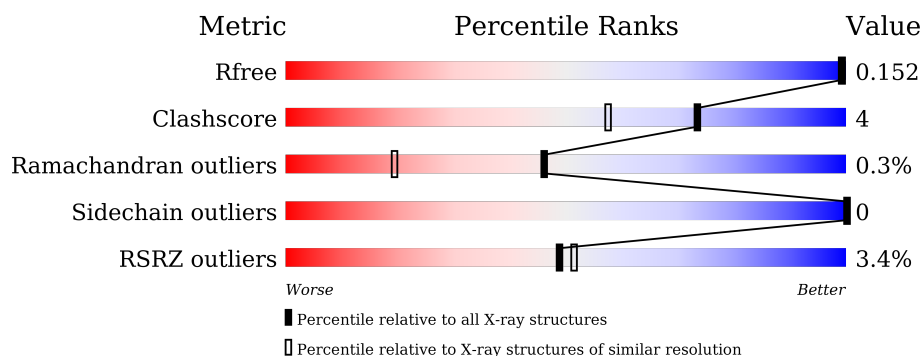
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.50 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}	91344	2072 (1.50-1.50)
Clashscore	102246	2274 (1.50-1.50)
Ramachandran outliers	100387	2218 (1.50-1.50)
Sidechain outliers	100360	2216 (1.50-1.50)
RSRZ outliers	91569	2075 (1.50-1.50)

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 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	381	<div> <div>3%</div> <div>80%</div> <div>13%</div> <div>7%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GOL	A	401	-	-	-	X
2	GOL	A	406	-	-	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GOL	A	407	-	-	-	X
2	GOL	A	408	-	-	-	X
2	GOL	A	409	-	-	-	X
2	GOL	A	411	-	-	-	X
2	GOL	A	412	-	-	-	X
2	GOL	A	414	-	-	-	X
2	GOL	A	415	-	-	-	X
2	GOL	A	416	-	-	-	X

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3344 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 PLL lectin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	354	Total	C	N	O	S	0	18	0
			2825	1784	497	540	4			

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	10	HIS	TYR	conflict	UNP Q7N8J0
A	16	GLU	ALA	conflict	UNP Q7N8J0
A	93	GLY	SER	conflict	UNP Q7N8J0
A	139	VAL	ALA	conflict	UNP Q7N8J0
A	142	SER	THR	conflict	UNP Q7N8J0
A	177	LEU	ILE	conflict	UNP Q7N8J0
A	225	ASN	GLY	conflict	UNP Q7N8J0
A	240	SER	ASN	conflict	UNP Q7N8J0
A	278	GLN	ARG	conflict	UNP Q7N8J0
A	298	HIS	GLN	conflict	UNP Q7N8J0
A	369	LEU	-	expression tag	UNP Q7N8J0
A	370	GLU	-	expression tag	UNP Q7N8J0
A	371	HIS	-	expression tag	UNP Q7N8J0
A	372	HIS	-	expression tag	UNP Q7N8J0
A	373	HIS	-	expression tag	UNP Q7N8J0
A	374	HIS	-	expression tag	UNP Q7N8J0
A	375	HIS	-	expression tag	UNP Q7N8J0
A	376	HIS	-	expression tag	UNP Q7N8J0
A	377	TRP	-	expression tag	UNP Q7N8J0
A	378	ARG	-	expression tag	UNP Q7N8J0
A	379	SER	-	expression tag	UNP Q7N8J0
A	380	GLY	-	expression tag	UNP Q7N8J0
A	381	CYS	-	expression tag	UNP Q7N8J0

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	1
			12	6	6		
2	A	1	Total	C	O	0	1
			12	6	6		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	405	Total	O	0	0
			405	405		

- Molecule 1: PLL lectin



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	71.79 Å 87.83 Å 158.34 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.17 – 1.50 38.40 – 1.50	Depositor EDS
% Data completeness (in resolution range)	99.9 (39.17-1.50) 99.9 (38.40-1.50)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.97 (at 1.50 Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.131 , 0.152 0.132 , 0.152	Depositor DCC
R_{free} test set	4023 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	11.7	Xtriage
Anisotropy	0.023	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 45.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3344	wwPDB-VP
Average B, all atoms (Å ²)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.61% 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](#)

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

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	1.36	12/2952 (0.4%)	1.28	15/4052 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	297	SER	CB-OG	6.89	1.51	1.42
1	A	228	ASP	CB-CG	-6.83	1.37	1.51
1	A	246	SER	CA-CB	-6.67	1.43	1.52
1	A	142[A]	SER	CA-CB	6.60	1.62	1.52
1	A	142[B]	SER	CA-CB	6.60	1.62	1.52

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	262	ASP	CB-CG-OD2	-11.79	107.69	118.30
1	A	69	PHE	CB-CG-CD1	8.05	126.44	120.80
1	A	236	ARG	NE-CZ-NH2	-7.63	116.48	120.30
1	A	364	ASP	CB-CG-OD2	-7.61	111.45	118.30
1	A	37	ARG	NE-CZ-NH1	7.32	123.96	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	299	ASN	Peptide

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	2825	0	2712	24	0
2	A	114	0	148	2	0
3	A	405	0	0	2	0
All	All	3344	0	2860	24	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.

The worst 5 of 24 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:88:GLU:OE2	1:A:100:HIS:HD2	1.75	0.70
1:A:39:GLU:OE2	1:A:51:HIS:HD2	1.74	0.70
1:A:202:ALA:O	1:A:204[B]:THR:HG23	1.92	0.68
1:A:184:GLU:OE2	1:A:196:HIS:HD2	1.76	0.68
1:A:136:GLU:OE2	1:A:148:HIS:HD2	1.77	0.67

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	370/381 (97%)	361 (98%)	7 (2%)	2 (0%)	34 10

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	215[A]	SER
1	A	215[B]	SER

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	311/316 (98%)	311 (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. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	148	HIS
1	A	339	HIS
1	A	196	HIS
1	A	100	HIS
1	A	176	HIS

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 [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

19 ligands are modelled in this entry.

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	GOL	A	401	-	5,5,5	1.15	1 (20%)	5,5,5	1.09	1 (20%)
2	GOL	A	402	-	5,5,5	1.04	0	5,5,5	0.77	0
2	GOL	A	403[A]	-	5,5,5	0.76	0	5,5,5	0.71	0
2	GOL	A	403[B]	-	5,5,5	0.95	0	5,5,5	0.49	0
2	GOL	A	404[A]	-	5,5,5	0.42	0	5,5,5	1.08	0
2	GOL	A	404[B]	-	5,5,5	0.51	0	5,5,5	1.60	1 (20%)
2	GOL	A	405	-	5,5,5	1.41	1 (20%)	5,5,5	1.26	1 (20%)
2	GOL	A	406	-	5,5,5	1.00	1 (20%)	5,5,5	1.10	0
2	GOL	A	407	-	5,5,5	0.92	0	5,5,5	1.60	1 (20%)
2	GOL	A	408	-	5,5,5	0.93	0	5,5,5	1.03	0
2	GOL	A	409	-	5,5,5	0.61	0	5,5,5	0.67	0
2	GOL	A	410	-	5,5,5	1.58	1 (20%)	5,5,5	1.96	1 (20%)
2	GOL	A	411	-	5,5,5	0.71	0	5,5,5	1.04	0
2	GOL	A	412	-	5,5,5	1.27	0	5,5,5	0.59	0
2	GOL	A	413	-	5,5,5	0.95	0	5,5,5	1.21	0
2	GOL	A	414	-	5,5,5	1.27	1 (20%)	5,5,5	1.48	1 (20%)
2	GOL	A	415	-	5,5,5	0.95	0	5,5,5	0.69	0
2	GOL	A	416	-	5,5,5	1.67	2 (40%)	5,5,5	1.54	1 (20%)
2	GOL	A	417	-	5,5,5	0.79	0	5,5,5	1.23	0

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	GOL	A	401	-	-	0/4/4/4	0/0/0/0
2	GOL	A	402	-	-	0/4/4/4	0/0/0/0
2	GOL	A	403[A]	-	-	0/4/4/4	0/0/0/0
2	GOL	A	403[B]	-	-	0/4/4/4	0/0/0/0
2	GOL	A	404[A]	-	-	0/4/4/4	0/0/0/0
2	GOL	A	404[B]	-	-	0/4/4/4	0/0/0/0
2	GOL	A	405	-	-	0/4/4/4	0/0/0/0
2	GOL	A	406	-	-	0/4/4/4	0/0/0/0
2	GOL	A	407	-	-	0/4/4/4	0/0/0/0
2	GOL	A	408	-	-	0/4/4/4	0/0/0/0
2	GOL	A	409	-	-	0/4/4/4	0/0/0/0
2	GOL	A	410	-	-	0/4/4/4	0/0/0/0
2	GOL	A	411	-	-	0/4/4/4	0/0/0/0
2	GOL	A	412	-	-	0/4/4/4	0/0/0/0
2	GOL	A	413	-	-	0/4/4/4	0/0/0/0
2	GOL	A	414	-	-	0/4/4/4	0/0/0/0
2	GOL	A	415	-	-	0/4/4/4	0/0/0/0
2	GOL	A	416	-	-	0/4/4/4	0/0/0/0
2	GOL	A	417	-	-	0/4/4/4	0/0/0/0

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	405	GOL	C3-C2	-2.58	1.41	1.52
2	A	410	GOL	C1-C2	-2.49	1.42	1.52
2	A	406	GOL	O2-C2	-2.15	1.37	1.43
2	A	401	GOL	C1-C2	-2.01	1.44	1.52
2	A	416	GOL	O3-C3	2.22	1.52	1.42

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	407	GOL	O2-C2-C3	-3.20	93.09	108.47
2	A	416	GOL	O3-C3-C2	-2.76	95.98	109.97
2	A	401	GOL	O2-C2-C1	2.02	118.15	108.47
2	A	404[B]	GOL	O3-C3-C2	2.29	121.56	109.97
2	A	405	GOL	O1-C1-C2	2.41	122.18	109.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	409	GOL	1	0
2	A	415	GOL	1	0

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	354/381 (92%)	-0.22	12 (3%)	49 51	7, 10, 24, 75	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	250	SER	6.1
1	A	253	VAL	4.7
1	A	251	ASP	4.3
1	A	249	GLY	3.5
1	A	248	ILE	2.9

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](#)

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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	GOL	A	407	6/6	0.94	0.14	22.71	18,21,23,33	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	GOL	A	416	6/6	0.80	0.22	12.18	24,36,42,46	0
2	GOL	A	411	6/6	0.79	0.16	8.82	35,42,43,46	0
2	GOL	A	415	6/6	0.90	0.15	8.34	28,39,40,42	0
2	GOL	A	406	6/6	0.95	0.15	8.04	20,30,39,42	0
2	GOL	A	408	6/6	0.90	0.12	5.39	18,31,35,40	0
2	GOL	A	414	6/6	0.92	0.16	4.14	15,27,33,34	0
2	GOL	A	401	6/6	0.95	0.11	4.02	8,11,16,23	0
2	GOL	A	409	6/6	0.97	0.08	2.73	8,11,14,22	0
2	GOL	A	412	6/6	0.95	0.08	2.32	14,16,18,22	0
2	GOL	A	405	6/6	0.95	0.10	1.88	18,22,24,27	0
2	GOL	A	413	6/6	0.93	0.14	1.82	17,23,26,26	0
2	GOL	A	404[A]	6/6	0.94	0.12	1.66	13,15,17,18	6
2	GOL	A	404[B]	6/6	0.94	0.12	1.65	15,18,19,23	6
2	GOL	A	410	6/6	0.97	0.06	-0.46	11,15,16,17	0
2	GOL	A	402	6/6	0.98	0.06	-0.57	10,12,15,17	0
2	GOL	A	403[A]	6/6	0.99	0.06	-0.74	7,8,8,10	6
2	GOL	A	403[B]	6/6	0.99	0.06	-0.74	7,8,9,11	6
2	GOL	A	417	6/6	0.94	0.12	-	21,33,36,36	0

6.5 Other polymers ⓘ

There are no such residues in this entry.