



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 05:05 am BST

PDB ID : 2R1A  
Title : Crystal structure of the periplasmic lipopolysaccharide transport protein LptA (YhbN), trigonal form  
Authors : Suits, M.D.L.; Polissi, A.; Jia, Z.; Montreal-Kingston Bacterial Structural Genomics Initiative (BSGI)  
Deposited on : 2007-08-22  
Resolution : 3.26 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](mailto: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 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	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

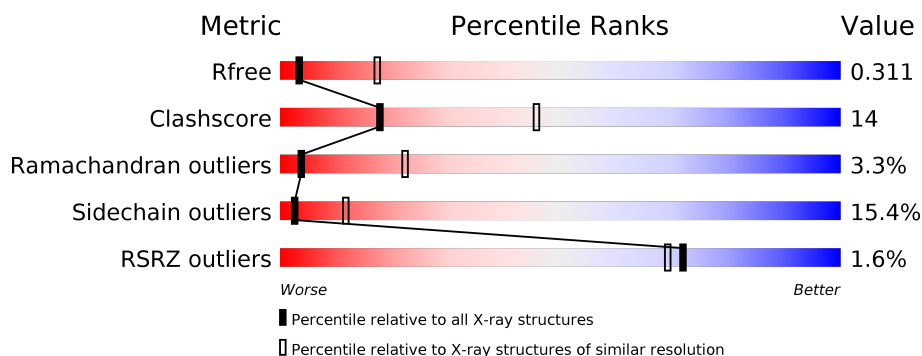
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.26 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	1191 (3.30-3.22)
Clashscore	141614	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)
RSRZ outliers	127900	1154 (3.30-3.22)

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  $\geq 3$ , 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	159	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, green 1%, green 58%, yellow 58%, yellow 80%, orange 80%, orange 82%, grey 82%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>58%</span> <span>22%</span> <span>••</span> <span>18%</span> </div> </div>
1	B	159	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 2%, green 2%, green 47%, yellow 47%, yellow 81%, orange 81%, orange 83%, grey 83%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>2%</span> <span>47%</span> <span>34%</span> <span>••</span> <span>16%</span> </div> </div>
1	C	159	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, green 1%, green 60%, yellow 60%, yellow 80%, orange 80%, orange 85%, grey 85%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>60%</span> <span>20%</span> <span>5%</span> <span>15%</span> </div> </div>
1	D	159	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, green 1%, green 61%, yellow 61%, yellow 82%, orange 82%, orange 86%, grey 86%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>61%</span> <span>21%</span> <span>•</span> <span>14%</span> </div> </div>
1	E	159	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, green 1%, green 55%, yellow 55%, yellow 79%, orange 79%, orange 83%, grey 83%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>55%</span> <span>24%</span> <span>••</span> <span>17%</span> </div> </div>
1	F	159	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 4%, green 4%, green 45%, yellow 45%, yellow 75%, orange 75%, orange 83%, grey 83%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>4%</span> <span>45%</span> <span>30%</span> <span>8%</span> <span>17%</span> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	159	<div><div></div><div>67%13%5%15%</div></div>
1	H	159	<div><div>2%<div></div></div><div>60%18%•18%</div></div>

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 7620 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 Protein yhbN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	131	Total	C	N	O	S	0	0	0
			927	578	154	191	4			
1	B	134	Total	C	N	O	S	0	0	0
			956	594	160	198	4			
1	C	135	Total	C	N	O	S	0	0	0
			989	617	163	206	3			
1	D	136	Total	C	N	O	S	0	0	0
			923	572	156	193	2			
1	E	132	Total	C	N	O	S	0	0	0
			918	569	155	192	2			
1	F	132	Total	C	N	O	S	0	0	0
			938	584	155	196	3			
1	G	135	Total	C	N	O	S	0	0	0
			980	609	165	203	3			
1	H	131	Total	C	N	O	S	0	0	0
			921	572	152	195	2			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	9	Total	O	0	0
			9	9		
2	B	8	Total	O	0	0
			8	8		
2	C	2	Total	O	0	0
			2	2		
2	D	8	Total	O	0	0
			8	8		
2	E	14	Total	O	0	0
			14	14		
2	F	9	Total	O	0	0
			9	9		

*Continued on next page...*

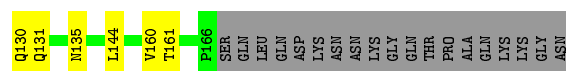
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	14	Total	O	0	0
			14	14		
2	H	4	Total	O	0	0
			4	4		

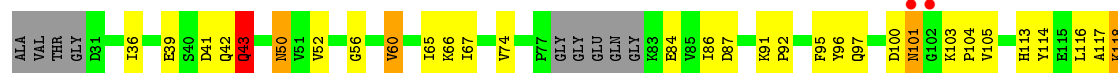


- Molecule 1: Protein vhbN

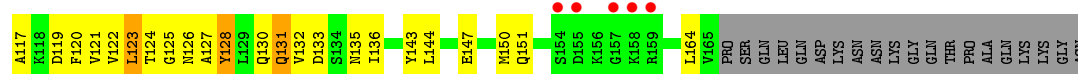
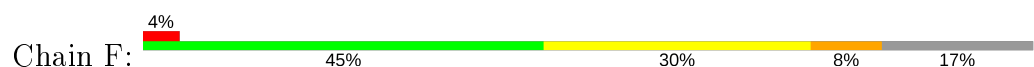




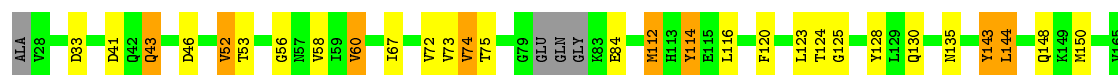
• Molecule 1: Protein yhbN



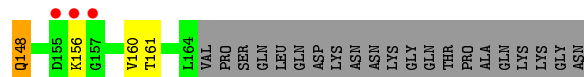
• Molecule 1: Protein yhbN



• Molecule 1: Protein yhbN



• Molecule 1: Protein yhbN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	146.21Å 146.21Å 186.81Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.98 – 3.26 29.98 – 3.25	Depositor EDS
% Data completeness (in resolution range)	83.6 (29.98-3.26) 97.3 (29.98-3.25)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.55 (at 3.24Å)	Xtriage
Refinement program	REFMAC 5.4.0065, PHENIX	Depositor
R, $R_{free}$	0.298 , 0.361 0.306 , 0.311	Depositor DCC
$R_{free}$ test set	1794 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	67.8	Xtriage
Anisotropy	0.935	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 62.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.033 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	7620	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 32.40 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.4591e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical 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	0.34	0/940	0.52	0/1283
1	B	0.38	0/971	0.54	0/1327
1	C	0.36	0/1004	0.53	0/1371
1	D	0.36	0/937	0.51	0/1285
1	E	0.36	0/932	0.57	0/1280
1	F	0.39	1/951 (0.1%)	0.58	0/1301
1	G	0.34	0/994	0.52	0/1356
1	H	0.34	0/934	0.50	0/1281
All	All	0.36	1/7663 (0.0%)	0.53	0/10484

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	132	VAL	CB-CG1	-5.18	1.42	1.52

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	927	0	798	25	0
1	B	956	0	831	35	0
1	C	989	0	899	28	0
1	D	923	0	762	20	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	918	0	758	25	0
1	F	938	0	817	36	0
1	G	980	0	882	16	0
1	H	921	0	783	17	0
2	A	9	0	0	1	0
2	B	8	0	0	1	0
2	C	2	0	0	0	0
2	D	8	0	0	0	0
2	E	14	0	0	1	0
2	F	9	0	0	1	0
2	G	14	0	0	1	0
2	H	4	0	0	2	0
All	All	7620	0	6530	191	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 14.

The worst 5 of 191 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:148:GLN:HG3	1:B:47:MET:HB2	1.48	0.95
1:F:98:MET:HA	1:F:99:GLN:CB	1.98	0.93
1:F:97:GLN:HG2	1:F:98:MET:H	1.43	0.83
1:B:148:GLN:HB3	1:C:47:MET:HB2	1.60	0.81
1:C:143:TYR:CE1	1:C:145:VAL:HA	2.15	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	125/159 (79%)	98 (78%)	23 (18%)	4 (3%)	4	23
1	B	130/159 (82%)	95 (73%)	31 (24%)	4 (3%)	4	24
1	C	131/159 (82%)	117 (89%)	12 (9%)	2 (2%)	10	39
1	D	132/159 (83%)	105 (80%)	19 (14%)	8 (6%)	1	10
1	E	128/159 (80%)	101 (79%)	19 (15%)	8 (6%)	1	9
1	F	128/159 (80%)	100 (78%)	24 (19%)	4 (3%)	4	24
1	G	131/159 (82%)	116 (88%)	14 (11%)	1 (1%)	19	52
1	H	127/159 (80%)	109 (86%)	15 (12%)	3 (2%)	6	28
All	All	1032/1272 (81%)	841 (82%)	157 (15%)	34 (3%)	4	22

5 of 34 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	156	LYS
1	D	77	PRO
1	D	119	ASP
1	E	118	LYS
1	F	99	GLN

### 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	88/135 (65%)	77 (88%)	11 (12%)	4	19
1	B	91/135 (67%)	73 (80%)	18 (20%)	1	5
1	C	101/135 (75%)	89 (88%)	12 (12%)	5	20
1	D	82/135 (61%)	72 (88%)	10 (12%)	5	20
1	E	82/135 (61%)	71 (87%)	11 (13%)	4	16
1	F	91/135 (67%)	65 (71%)	26 (29%)	0	1
1	G	98/135 (73%)	85 (87%)	13 (13%)	4	16
1	H	87/135 (64%)	77 (88%)	10 (12%)	5	22

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	720/1080 (67%)	609 (85%)	111 (15%)	<b>2</b> <b>12</b>

5 of 111 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	144	LEU
1	F	31	ASP
1	H	43	GLN
1	E	39	GLU
1	E	97	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	148	GLN
1	D	130	GLN
1	H	68	ASN
1	D	97	GLN
1	E	62	GLN

### 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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	131/159 (82%)	-0.06	1 (0%) 86 86	44, 64, 75, 77	0
1	B	134/159 (84%)	0.08	3 (2%) 62 59	54, 63, 74, 80	0
1	C	135/159 (84%)	-0.21	1 (0%) 87 88	32, 41, 61, 65	0
1	D	136/159 (85%)	-0.03	1 (0%) 87 88	47, 69, 76, 78	0
1	E	132/159 (83%)	0.11	2 (1%) 73 71	63, 71, 77, 80	0
1	F	132/159 (83%)	0.03	6 (4%) 33 31	46, 58, 68, 70	0
1	G	135/159 (84%)	-0.20	0 100 100	34, 46, 58, 60	0
1	H	131/159 (82%)	-0.09	3 (2%) 60 58	35, 54, 68, 71	0
All	All	1066/1272 (83%)	-0.05	17 (1%) 72 69	32, 60, 75, 80	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	77	PRO	3.6
1	F	49	GLY	2.6
1	H	155	ASP	2.6
1	F	155	ASP	2.5
1	B	77	PRO	2.4

### 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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.