Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 06:01 pm BST

PDB ID	:	6C5D
Title	:	N-terminal domain of Helicobacter pylori LlaJI.R1
Authors	:	Hosford, C.J.; Chappie, J.S.
Deposited on	:	2018-01-16
Resolution	:	1.97 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 following versions of software and data (see references ⓘ) were used in the production of this report:

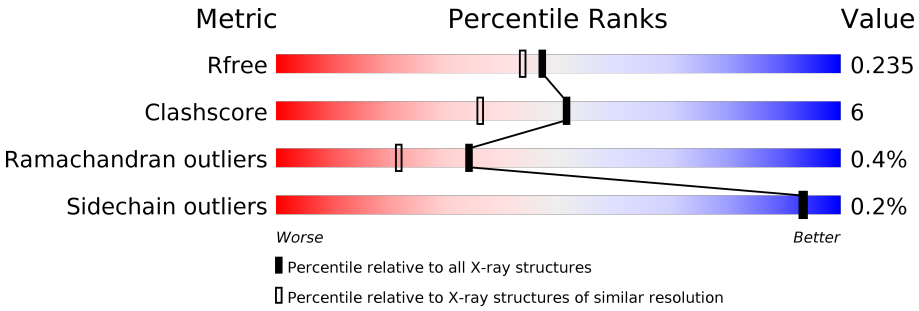
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.97 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)

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 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	144	75% 9% 16%
1	B	144	76% 15% 9%
1	C	144	76% 8% 16%
1	D	144	77% 14% 9%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4369 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 LlaJI.R1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	121	Total	C	N	O	Se	0	0	0
			990	636	168	184	2			
1	D	131	Total	C	N	O	Se	0	0	0
			1072	689	182	199	2			
1	B	131	Total	C	N	O	Se	0	0	0
			1072	689	182	199	2			
1	C	121	Total	C	N	O	Se	0	0	0
			990	636	168	184	2			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	138	GLU	-	expression tag	UNP Q9ZMQ4
A	139	HIS	-	expression tag	UNP Q9ZMQ4
A	140	HIS	-	expression tag	UNP Q9ZMQ4
A	141	HIS	-	expression tag	UNP Q9ZMQ4
A	142	HIS	-	expression tag	UNP Q9ZMQ4
A	143	HIS	-	expression tag	UNP Q9ZMQ4
A	144	HIS	-	expression tag	UNP Q9ZMQ4
D	138	GLU	-	expression tag	UNP Q9ZMQ4
D	139	HIS	-	expression tag	UNP Q9ZMQ4
D	140	HIS	-	expression tag	UNP Q9ZMQ4
D	141	HIS	-	expression tag	UNP Q9ZMQ4
D	142	HIS	-	expression tag	UNP Q9ZMQ4
D	143	HIS	-	expression tag	UNP Q9ZMQ4
D	144	HIS	-	expression tag	UNP Q9ZMQ4
B	138	GLU	-	expression tag	UNP Q9ZMQ4
B	139	HIS	-	expression tag	UNP Q9ZMQ4
B	140	HIS	-	expression tag	UNP Q9ZMQ4
B	141	HIS	-	expression tag	UNP Q9ZMQ4
B	142	HIS	-	expression tag	UNP Q9ZMQ4
B	143	HIS	-	expression tag	UNP Q9ZMQ4
B	144	HIS	-	expression tag	UNP Q9ZMQ4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	138	GLU	-	expression tag	UNP Q9ZMQ4
C	139	HIS	-	expression tag	UNP Q9ZMQ4
C	140	HIS	-	expression tag	UNP Q9ZMQ4
C	141	HIS	-	expression tag	UNP Q9ZMQ4
C	142	HIS	-	expression tag	UNP Q9ZMQ4
C	143	HIS	-	expression tag	UNP Q9ZMQ4
C	144	HIS	-	expression tag	UNP Q9ZMQ4

- Molecule 2 is water.

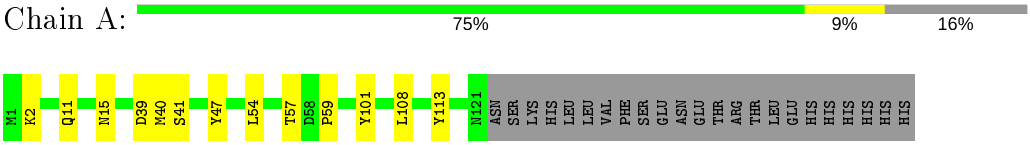
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	78	Total O 78 78	0	0
2	D	46	Total O 46 46	0	0
2	B	43	Total O 43 43	0	0
2	C	78	Total O 78 78	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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. 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. 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.

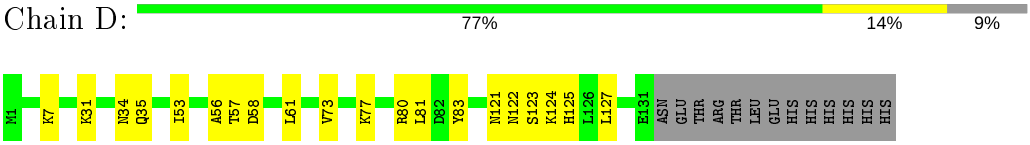
- Molecule 1: LlaJI.R1

Chain A:



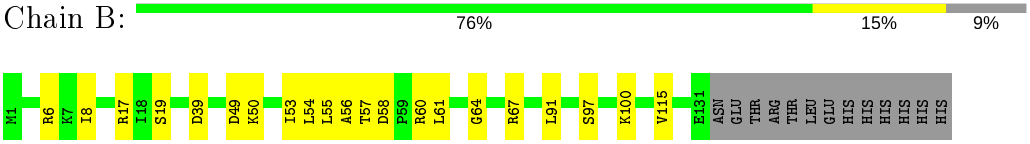
- Molecule 1: LlaJI.R1

Chain D:



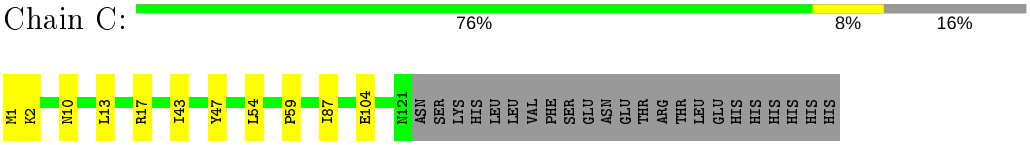
- Molecule 1: LlaJI.R1

Chain B:



- Molecule 1: LlaJI.R1

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	37.53Å 43.77Å 85.09Å 97.87° 93.86° 97.77°	Depositor
Resolution (Å)	83.97 – 1.97 83.97 – 1.97	Depositor EDS
% Data completeness (in resolution range)	93.6 (83.97-1.97) 93.9 (83.97-1.97)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.88 (at 1.97Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	(Not available) , (Not available) 0.206 , 0.235	Depositor DCC
R_{free} test set	1661 reflections (4.73%)	wwPDB-VP
Wilson B-factor (Å^2)	32.6	Xtriage
Anisotropy	0.395	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 48.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4369	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.12% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	0.74	0/1001	0.80	1/1343 (0.1%)
1	B	0.79	0/1085	0.76	0/1456
1	C	0.92	3/1001 (0.3%)	0.79	0/1343
1	D	0.66	0/1085	0.70	0/1456
All	All	0.78	3/4172 (0.1%)	0.76	1/5598 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	17	ARG	CZ-NH2	5.80	1.40	1.33
1	C	47	TYR	CE1-CZ	-5.45	1.31	1.38
1	C	17	ARG	CZ-NH1	5.17	1.39	1.33

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	39	ASP	CB-CG-OD2	5.09	122.88	118.30

There are no chirality outliers.

There are no planarity outliers.

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	990	0	1038	8	0
1	B	1072	0	1120	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	990	0	1038	8	0
1	D	1072	0	1120	20	0
2	A	78	0	0	1	0
2	B	43	0	0	1	2
2	C	78	0	0	0	0
2	D	46	0	0	0	2
All	All	4369	0	4316	47	2

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 6.

All (47) 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:D:73:VAL:HG12	1:D:77:LYS:HE3	1.55	0.87
1:A:57:THR:OG1	2:A:201:HOH:O	1.92	0.85
1:B:49:ASP:OD1	1:B:50:LYS:N	2.10	0.83
1:B:58:ASP:OD1	1:B:60:ARG:NH1	2.13	0.81
1:D:7:LYS:HD3	1:D:123:SER:O	1.93	0.69
1:D:73:VAL:CG1	1:D:77:LYS:HE3	2.23	0.68
1:D:77:LYS:HE2	1:D:83:TYR:O	1.96	0.66
1:B:8:ILE:HD11	1:B:91:LEU:HG	1.79	0.65
1:B:6:ARG:NH1	1:B:19:SER:O	2.28	0.63
1:A:108:LEU:HD23	1:A:113:TYR:CZ	2.34	0.62
1:D:124:LYS:HG3	1:D:125:HIS:H	1.64	0.62
1:D:56:ALA:O	1:D:58:ASP:N	2.34	0.61
1:C:13:LEU:HD12	1:C:87:ILE:HD11	1.84	0.59
1:D:7:LYS:HD3	1:D:123:SER:C	2.24	0.58
1:D:31:LYS:HG2	1:D:31:LYS:O	2.02	0.58
1:B:56:ALA:O	1:B:58:ASP:N	2.39	0.55
1:C:43:ILE:HG13	1:C:104:GLU:HB2	1.89	0.55
1:A:108:LEU:HD23	1:A:113:TYR:CE2	2.43	0.54
1:B:49:ASP:OD2	1:B:64:GLY:HA3	2.08	0.54
1:C:13:LEU:HD12	1:C:87:ILE:CD1	2.39	0.53
1:D:127:LEU:HB2	1:C:2:LYS:HB3	1.90	0.53
1:B:39:ASP:OD1	1:B:50:LYS:HE3	2.11	0.50
1:D:124:LYS:HB2	1:D:125:HIS:HD2	1.76	0.49
1:B:49:ASP:OD2	1:B:67:ARG:NH2	2.46	0.48
1:D:124:LYS:HB2	1:D:125:HIS:CD2	2.49	0.48
1:B:17:ARG:HD2	1:B:54:LEU:HD12	1.96	0.47
1:D:53:ILE:HD13	1:D:61:LEU:HD23	1.97	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:122:ASN:O	1:D:124:LYS:HG2	2.15	0.46
1:B:97:SER:OG	1:B:100:LYS:HD2	2.15	0.46
1:C:54:LEU:O	1:C:59:PRO:HA	2.16	0.46
1:A:41:SER:HA	1:A:47:TYR:O	2.15	0.46
1:D:124:LYS:HG3	1:D:125:HIS:N	2.30	0.45
1:B:115:VAL:HG23	2:B:215:HOH:O	2.17	0.44
1:B:17:ARG:HB3	1:B:61:LEU:O	2.18	0.43
1:A:108:LEU:CD2	1:A:113:TYR:CZ	3.01	0.43
1:A:54:LEU:O	1:A:59:PRO:HA	2.19	0.42
1:C:10:ASN:OD1	1:C:87:ILE:HD12	2.20	0.42
1:D:121:ASN:HB2	1:C:1:MSE:SE	2.68	0.42
1:D:77:LYS:HA	1:D:80:ARG:O	2.19	0.42
1:C:10:ASN:OD1	1:C:87:ILE:CD1	2.68	0.41
1:D:34:ASN:O	1:D:35:GLN:HB2	2.20	0.41
1:A:11:GLN:HG2	1:A:15:ASN:ND2	2.36	0.41
1:B:53:ILE:HG22	1:B:55:LEU:HD21	2.03	0.41
1:D:73:VAL:O	1:D:77:LYS:HG3	2.20	0.40
1:A:40:MSE:HA	1:A:101:TYR:O	2.21	0.40
1:D:53:ILE:HD13	1:D:61:LEU:CD2	2.51	0.40
1:D:77:LYS:O	1:D:81:LEU:HA	2.21	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:239:HOH:O	2:B:242:HOH:O[1_665]	1.96	0.24
2:D:237:HOH:O	2:B:242:HOH:O[1_665]	2.16	0.04

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	119/144 (83%)	117 (98%)	2 (2%)	0	100 100
1	B	129/144 (90%)	123 (95%)	5 (4%)	1 (1%)	19 9
1	C	119/144 (83%)	117 (98%)	2 (2%)	0	100 100
1	D	129/144 (90%)	124 (96%)	4 (3%)	1 (1%)	19 9
All	All	496/576 (86%)	481 (97%)	13 (3%)	2 (0%)	34 22

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	57	THR
1	B	57	THR

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	111/132 (84%)	110 (99%)	1 (1%)	78 77
1	B	121/132 (92%)	121 (100%)	0	100 100
1	C	111/132 (84%)	111 (100%)	0	100 100
1	D	121/132 (92%)	121 (100%)	0	100 100
All	All	464/528 (88%)	463 (100%)	1 (0%)	93 93

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	LYS

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

Mol	Chain	Res	Type
1	D	15	ASN

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

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.