



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 05:00 pm BST

PDB ID : 3A5I
Title : Structure of the cytoplasmic domain of FlhA
Authors : Imada, K.; Saijo-Hamano, Y.; Shimada, M.; Namba, K.
Deposited on : 2009-08-07
Resolution : 2.80 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

<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

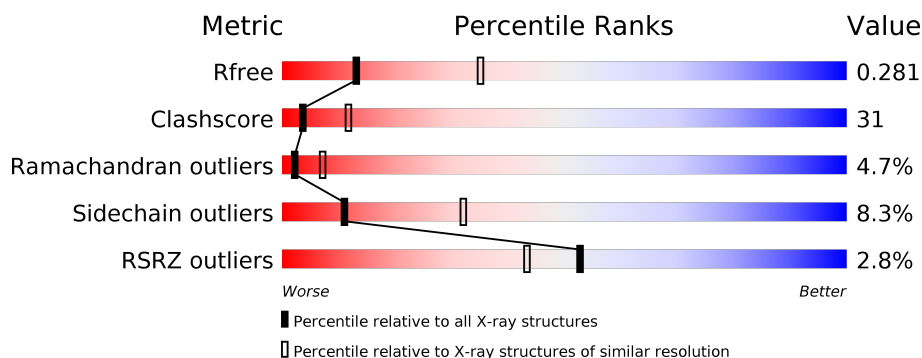
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 $\leq 5\%$. 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	389	<div> <div>3%</div> <div> <div></div> <div>40%</div> <div>41%</div> <div>7%</div> <div>11%</div> </div> </div>
1	B	389	<div> <div>2%</div> <div> <div></div> <div>49%</div> <div>34%</div> <div>5%</div> <div>12%</div> </div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 5336 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 Flagellar biosynthesis protein flhA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	346	Total	C	N	O	S	0	0	0
			2681	1695	478	499	9			
1	B	342	Total	C	N	O	S	0	0	0
			2655	1681	472	493	9			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	304	MET	-	EXPRESSION TAG	UNP P40729
A	305	GLY	-	EXPRESSION TAG	UNP P40729
A	306	HIS	-	EXPRESSION TAG	UNP P40729
A	307	HIS	-	EXPRESSION TAG	UNP P40729
A	308	HIS	-	EXPRESSION TAG	UNP P40729
A	309	HIS	-	EXPRESSION TAG	UNP P40729
A	310	HIS	-	EXPRESSION TAG	UNP P40729
A	311	HIS	-	EXPRESSION TAG	UNP P40729
A	312	HIS	-	EXPRESSION TAG	UNP P40729
A	313	HIS	-	EXPRESSION TAG	UNP P40729
A	314	HIS	-	EXPRESSION TAG	UNP P40729
A	315	HIS	-	EXPRESSION TAG	UNP P40729
A	316	SER	-	EXPRESSION TAG	UNP P40729
A	317	SER	-	EXPRESSION TAG	UNP P40729
A	318	GLY	-	EXPRESSION TAG	UNP P40729
A	319	HIS	-	EXPRESSION TAG	UNP P40729
A	320	ILE	-	EXPRESSION TAG	UNP P40729
A	321	ASP	-	EXPRESSION TAG	UNP P40729
A	322	ASP	-	EXPRESSION TAG	UNP P40729
A	323	ASP	-	EXPRESSION TAG	UNP P40729
A	324	ASP	-	EXPRESSION TAG	UNP P40729
A	325	LYS	-	EXPRESSION TAG	UNP P40729
A	326	HIS	-	EXPRESSION TAG	UNP P40729
A	327	MET	-	EXPRESSION TAG	UNP P40729
B	304	MET	-	EXPRESSION TAG	UNP P40729

Continued on next page...

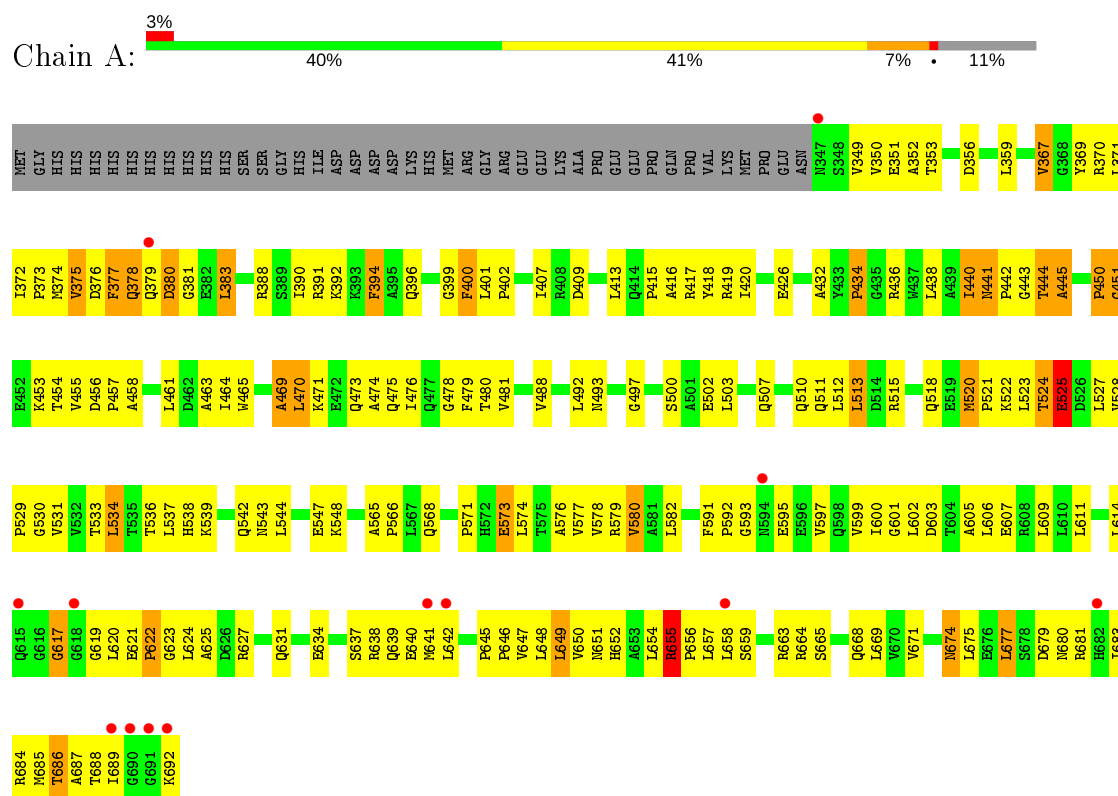
Continued from previous page...

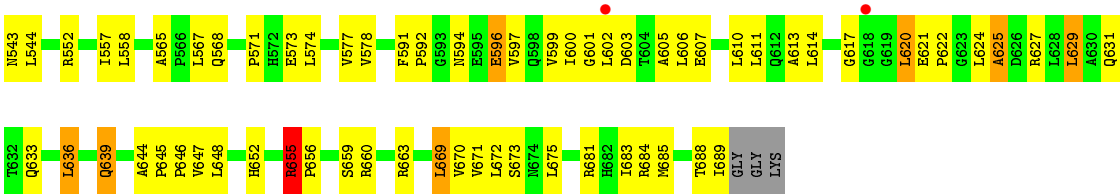
Chain	Residue	Modelled	Actual	Comment	Reference
B	305	GLY	-	EXPRESSION TAG	UNP P40729
B	306	HIS	-	EXPRESSION TAG	UNP P40729
B	307	HIS	-	EXPRESSION TAG	UNP P40729
B	308	HIS	-	EXPRESSION TAG	UNP P40729
B	309	HIS	-	EXPRESSION TAG	UNP P40729
B	310	HIS	-	EXPRESSION TAG	UNP P40729
B	311	HIS	-	EXPRESSION TAG	UNP P40729
B	312	HIS	-	EXPRESSION TAG	UNP P40729
B	313	HIS	-	EXPRESSION TAG	UNP P40729
B	314	HIS	-	EXPRESSION TAG	UNP P40729
B	315	HIS	-	EXPRESSION TAG	UNP P40729
B	316	SER	-	EXPRESSION TAG	UNP P40729
B	317	SER	-	EXPRESSION TAG	UNP P40729
B	318	GLY	-	EXPRESSION TAG	UNP P40729
B	319	HIS	-	EXPRESSION TAG	UNP P40729
B	320	ILE	-	EXPRESSION TAG	UNP P40729
B	321	ASP	-	EXPRESSION TAG	UNP P40729
B	322	ASP	-	EXPRESSION TAG	UNP P40729
B	323	ASP	-	EXPRESSION TAG	UNP P40729
B	324	ASP	-	EXPRESSION TAG	UNP P40729
B	325	LYS	-	EXPRESSION TAG	UNP P40729
B	326	HIS	-	EXPRESSION TAG	UNP P40729
B	327	MET	-	EXPRESSION TAG	UNP P40729

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 the various outlier classes 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: Flagellar biosynthesis protein flhA





4 Data and refinement statistics

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, α , β , γ	216.32Å 216.32Å 65.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.80 – 2.80 40.84 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.5 (40.80-2.80) 99.6 (40.84-2.80)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.57 (at 2.81Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.257 , 0.290 0.251 , 0.281	Depositor DCC
R_{free} test set	1863 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	73.5	Xtriage
Anisotropy	0.206	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 43.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.022 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	5336	wwPDB-VP
Average B, all atoms (Å ²)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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](#)

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	0.40	0/2727	0.65	0/3703
1	B	0.43	0/2701	0.67	1/3671 (0.0%)
All	All	0.42	0/5428	0.66	1/7374 (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	451	GLY	N-CA-C	5.07	125.77	113.10

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	2681	0	2740	211	0
1	B	2655	0	2715	139	0
All	All	5336	0	5455	339	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 31.

The worst 5 of 339 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:376:ASP:CB	1:A:379:GLN:HE21	1.36	1.37
1:A:373:PRO:HA	1:A:379:GLN:HE22	1.08	1.10
1:A:376:ASP:HB2	1:A:379:GLN:HE21	0.96	1.09
1:A:511:GLN:NE2	1:B:354:TRP:HE1	1.49	1.09
1:A:376:ASP:CB	1:A:379:GLN:NE2	2.19	1.03

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	344/389 (88%)	294 (86%)	29 (8%)	21 (6%)	1	4
1	B	340/389 (87%)	296 (87%)	33 (10%)	11 (3%)	4	13
All	All	684/778 (88%)	590 (86%)	62 (9%)	32 (5%)	2	7

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	445	ALA
1	A	450	PRO
1	A	451	GLY
1	A	470	LEU
1	A	525	GLU

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	289/328 (88%)	260 (90%)	29 (10%)	7	22
1	B	287/328 (88%)	268 (93%)	19 (7%)	16	44
All	All	576/656 (88%)	528 (92%)	48 (8%)	11	32

5 of 48 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	580	VAL
1	A	677	LEU
1	B	636	LEU
1	A	582	LEU
1	A	655	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	631	GLN
1	A	674	ASN
1	B	615	GLN
1	A	639	GLN
1	A	668	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 [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	346/389 (88%)	0.17	13 (3%) 40 30	37, 74, 123, 137	0
1	B	342/389 (87%)	0.09	6 (1%) 68 61	33, 72, 114, 130	0
All	All	688/778 (88%)	0.13	19 (2%) 53 43	33, 73, 122, 137	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	692	LYS	5.4
1	B	479	PHE	3.4
1	A	347	ASN	3.3
1	A	658	LEU	3.3
1	A	642	LEU	3.2

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.