



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2024 – 11:45 AM EST

PDB ID : 3MIX
Title : Crystal structure of the cytosolic domain of B. subtilis FlhA
Authors : Bange, G.; Kuemmerer, N.; Bozkurt, G.; Wild, K.; Sinning, I.
Deposited on : 2010-04-12
Resolution : 2.30 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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 : 2.36
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) : 2.36

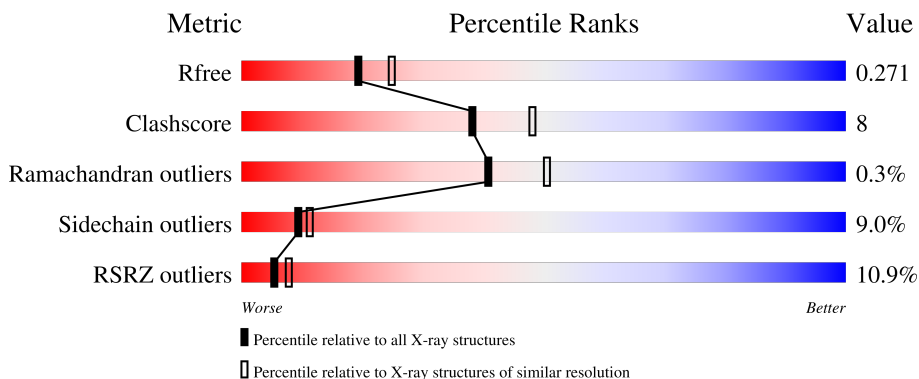
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 of 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	382	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2588 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
			Total	C	N	O	S			
1	A	313	2468	1572	411	481	4	0	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	296	MET	-	expression tag	UNP P35620
A	297	GLY	-	expression tag	UNP P35620
A	298	HIS	-	expression tag	UNP P35620
A	299	HIS	-	expression tag	UNP P35620
A	300	HIS	-	expression tag	UNP P35620
A	301	HIS	-	expression tag	UNP P35620
A	302	HIS	-	expression tag	UNP P35620

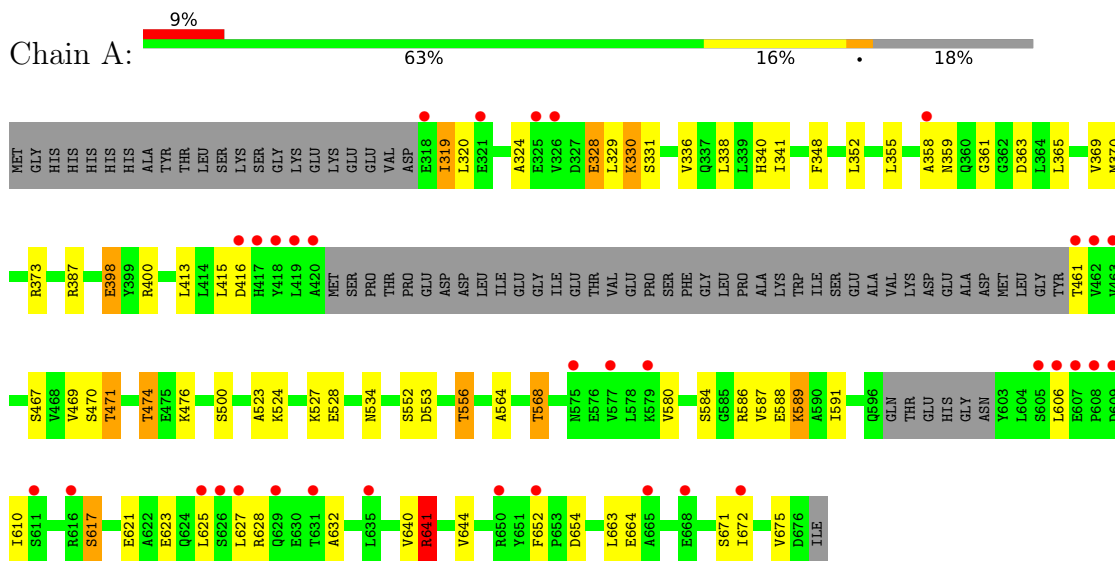
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	120	Total	O	0	0
			120	120		

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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	P 65	Depositor
Cell constants a, b, c, α , β , γ	84.18Å 84.18Å 118.16Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.83 – 2.30 34.83 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.3 (34.83-2.30) 99.3 (34.83-2.30)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	12.07 (at 2.29Å)	Xtrriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.216 , 0.269 0.221 , 0.271	Depositor DCC
R_{free} test set	1056 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	23.2	Xtrriage
Anisotropy	0.072	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 44.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.057 for h,-h-k,-l	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	2588	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.24% 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	0.56	1/2500 (0.0%)	0.63	2/3391 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	523	ALA	CA-CB	5.58	1.64	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	641	ARG	NE-CZ-NH1	-6.67	116.97	120.30
1	A	641	ARG	NE-CZ-NH2	5.10	122.85	120.30

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	2468	0	2544	38	1
2	A	120	0	0	2	1
All	All	2588	0	2544	38	1

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 8.

All (38) 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:568:THR:HG21	1:A:664:GLU:H	1.32	0.93
1:A:328:GLU:O	1:A:330:LYS:HD3	1.70	0.92
1:A:527:LYS:HD2	2:A:87:HOH:O	1.70	0.90
1:A:552:SER:O	1:A:556:THR:HG23	1.77	0.83
1:A:564:ALA:O	1:A:568:THR:HG23	1.79	0.81
1:A:568:THR:HG22	1:A:663:LEU:HA	1.68	0.74
1:A:528:GLU:OE1	1:A:641:ARG:NH1	2.24	0.70
1:A:358:ALA:HA	1:A:361:GLY:O	1.98	0.64
1:A:470:SER:O	1:A:474:THR:HG23	1.97	0.64
1:A:467:SER:O	1:A:471:THR:HG23	1.98	0.63
1:A:398:GLU:HG2	2:A:119:HOH:O	2.00	0.62
1:A:470:SER:O	1:A:474:THR:CG2	2.48	0.60
1:A:369:VAL:O	1:A:373:ARG:HG3	2.02	0.60
1:A:580:VAL:O	1:A:671:SER:HA	2.02	0.59
1:A:625:LEU:HD22	1:A:632:ALA:HA	1.86	0.57
1:A:584:SER:HB2	1:A:675:VAL:O	2.05	0.56
1:A:528:GLU:CD	1:A:641:ARG:HH12	2.09	0.55
1:A:617:SER:O	1:A:621:GLU:HG2	2.07	0.55
1:A:628:ARG:HH22	1:A:672:ILE:HG12	1.73	0.54
1:A:324:ALA:HB2	1:A:338:LEU:HD21	1.92	0.52
1:A:330:LYS:HE2	1:A:331:SER:H	1.75	0.51
1:A:587:VAL:HG21	1:A:675:VAL:HG12	1.93	0.51
1:A:591:ILE:HD13	1:A:644:VAL:HG22	1.94	0.49
1:A:586:ARG:O	1:A:589:LYS:HG3	2.14	0.48
1:A:319:ILE:HG21	1:A:400:ARG:HH12	1.77	0.48
1:A:552:SER:O	1:A:556:THR:CG2	2.56	0.47
1:A:640:VAL:HG12	1:A:644:VAL:HG23	1.96	0.46
1:A:413:LEU:HD11	1:A:469:VAL:HB	1.97	0.45
1:A:319:ILE:HG21	1:A:400:ARG:NH1	2.32	0.45
1:A:348:PHE:HB2	1:A:352:LEU:HB2	1.99	0.44
1:A:623:GLU:HG3	1:A:627:LEU:HD13	2.00	0.43
1:A:524:LYS:NZ	1:A:553:ASP:OD1	2.41	0.43
1:A:524:LYS:HB3	1:A:556:THR:HG21	2.01	0.43
1:A:320:LEU:HD21	1:A:340:HIS:CD2	2.54	0.42
1:A:625:LEU:CD2	1:A:632:ALA:HA	2.48	0.42
1:A:652:PHE:CE1	1:A:654:ASP:OD1	2.72	0.42
1:A:606:LEU:HD23	1:A:610:ILE:HG21	2.03	0.40
1:A:568:THR:CG2	1:A:663:LEU:HA	2.45	0.40

All (1) 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)
1:A:341:ILE:O	2:A:115:HOH:O[5_555]	2.07	0.13

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	307/382 (80%)	300 (98%)	6 (2%)	1 (0%)	41 50

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	329	LEU

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	279/338 (82%)	254 (91%)	25 (9%)	9 11

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	319	ILE
1	A	328	GLU
1	A	330	LYS
1	A	336	VAL
1	A	355	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	359	ASN
1	A	363	ASP
1	A	365	LEU
1	A	370	MET
1	A	387	ARG
1	A	398	GLU
1	A	415	LEU
1	A	416	ASP
1	A	461	THR
1	A	471	THR
1	A	474	THR
1	A	476	LYS
1	A	500	SER
1	A	534	ASN
1	A	556	THR
1	A	568	THR
1	A	588	GLU
1	A	589	LYS
1	A	617	SER
1	A	641	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	417	HIS
1	A	479	GLN
1	A	511	ASN
1	A	534	ASN
1	A	561	GLN
1	A	661	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

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, 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	313/382 (81%)	0.52	34 (10%) 5 8	7, 42, 91, 114	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	609	ASP	9.6
1	A	608	PRO	6.8
1	A	461	THR	6.6
1	A	606	LEU	6.5
1	A	462	VAL	5.8
1	A	419	LEU	5.8
1	A	652	PHE	5.7
1	A	629	GLN	5.1
1	A	577	VAL	4.7
1	A	626	SER	3.9
1	A	579	LYS	3.9
1	A	358	ALA	3.8
1	A	326	VAL	3.7
1	A	463	VAL	3.7
1	A	420	ALA	3.6
1	A	418	TYR	3.6
1	A	616	ARG	3.4
1	A	631	THR	3.4
1	A	650	ARG	3.3
1	A	417	HIS	3.3
1	A	325	GLU	3.1
1	A	627	LEU	3.0
1	A	575	ASN	2.9
1	A	665	ALA	2.9
1	A	668	GLU	2.9
1	A	635	LEU	2.8
1	A	611	SER	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	625	LEU	2.7
1	A	607	GLU	2.7
1	A	605	SER	2.6
1	A	672	ILE	2.4
1	A	318	GLU	2.4
1	A	321	GLU	2.1
1	A	416	ASP	2.0

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 monosaccharides 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.