



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 4, 2021 – 12:09 PM JST

PDB ID : 7ET0
Title : Crystal structure of the complex formed by Wolbachia cytoplasmic incompatibility factors CinA and CinB from wPip
Authors : Xiao, Y.J.; Wang, W.; Chen, X.; Ji, X.Y.; Yang, H.T.
Deposited on : 2021-05-12
Resolution : 2.20 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 : 2.23.2
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) : 2.23.2

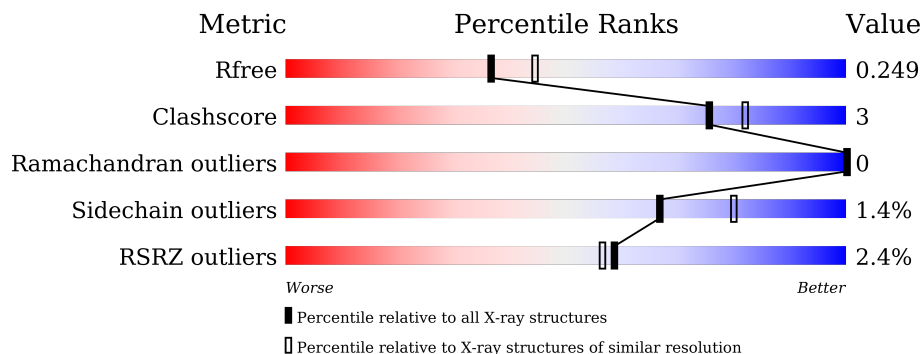
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	740	 % 81% 9% 10%
1	C	740	 3% 83% 9% 8%
2	B	453	 % 81% 8% 11%
2	D	453	 3% 82% 7% 11%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 17117 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 Bacteria factor B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	665	5208	3346	868	978	16	0	1	0
1	C	684	5364	3447	886	1015	16	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	733	LEU	-	expression tag	UNP B3CP74
A	734	GLU	-	expression tag	UNP B3CP74
A	735	HIS	-	expression tag	UNP B3CP74
A	736	HIS	-	expression tag	UNP B3CP74
A	737	HIS	-	expression tag	UNP B3CP74
A	738	HIS	-	expression tag	UNP B3CP74
A	739	HIS	-	expression tag	UNP B3CP74
A	740	HIS	-	expression tag	UNP B3CP74
C	733	LEU	-	expression tag	UNP B3CP74
C	734	GLU	-	expression tag	UNP B3CP74
C	735	HIS	-	expression tag	UNP B3CP74
C	736	HIS	-	expression tag	UNP B3CP74
C	737	HIS	-	expression tag	UNP B3CP74
C	738	HIS	-	expression tag	UNP B3CP74
C	739	HIS	-	expression tag	UNP B3CP74
C	740	HIS	-	expression tag	UNP B3CP74

- Molecule 2 is a protein called Bacteria factor A.

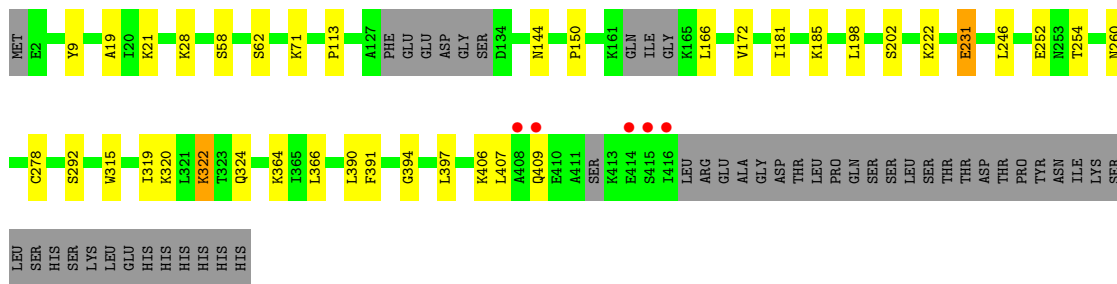
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	405	3239	2063	535	626	15	0	1	0
2	D	404	3180	2019	528	619	14	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

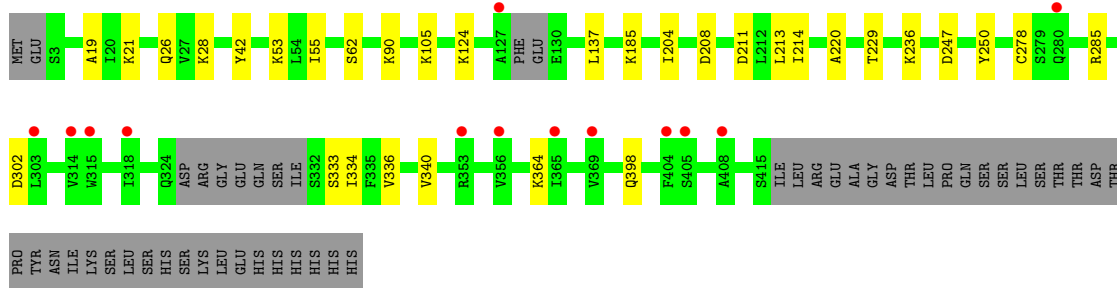
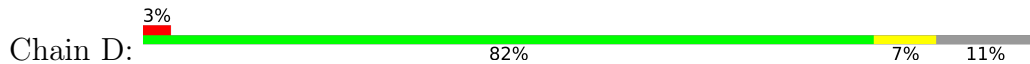
Chain	Residue	Modelled	Actual	Comment	Reference
B	446	LEU	-	expression tag	UNP B3CP73
B	447	GLU	-	expression tag	UNP B3CP73
B	448	HIS	-	expression tag	UNP B3CP73
B	449	HIS	-	expression tag	UNP B3CP73
B	450	HIS	-	expression tag	UNP B3CP73
B	451	HIS	-	expression tag	UNP B3CP73
B	452	HIS	-	expression tag	UNP B3CP73
B	453	HIS	-	expression tag	UNP B3CP73
D	446	LEU	-	expression tag	UNP B3CP73
D	447	GLU	-	expression tag	UNP B3CP73
D	448	HIS	-	expression tag	UNP B3CP73
D	449	HIS	-	expression tag	UNP B3CP73
D	450	HIS	-	expression tag	UNP B3CP73
D	451	HIS	-	expression tag	UNP B3CP73
D	452	HIS	-	expression tag	UNP B3CP73
D	453	HIS	-	expression tag	UNP B3CP73

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	51	Total O 51 51	0	0
3	B	43	Total O 43 43	0	0
3	C	22	Total O 22 22	0	0
3	D	10	Total O 10 10	0	0



● Molecule 2: Bacteria factor A



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	85.81Å 87.12Å 92.22Å 111.27° 90.97° 114.67°	Depositor
Resolution (Å)	46.33 – 2.20 46.33 – 2.19	Depositor EDS
% Data completeness (in resolution range)	93.2 (46.33-2.20) 93.3 (46.33-2.19)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.31 (at 2.20Å)	Xtrriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.219 , 0.248 0.219 , 0.249	Depositor DCC
R_{free} test set	2014 reflections (1.85%)	wwPDB-VP
Wilson B-factor (Å ²)	49.1	Xtrriage
Anisotropy	0.306	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 38.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	17117	wwPDB-VP
Average B, all atoms (Å ²)	62.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.95% 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

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	0.24	0/5299	0.41	0/7133
1	C	0.24	0/5456	0.41	0/7349
2	B	0.23	0/3292	0.36	0/4422
2	D	0.23	0/3231	0.36	0/4351
All	All	0.24	0/17278	0.39	0/23255

There are no bond length outliers.

There are no bond angle outliers.

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	5208	0	5176	37	0
1	C	5364	0	5305	36	0
2	B	3239	0	3175	20	0
2	D	3180	0	3061	17	0
3	A	51	0	0	1	0
3	B	43	0	0	0	0
3	C	22	0	0	0	0
3	D	10	0	0	0	0
All	All	17117	0	16717	103	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 3.

The worst 5 of 103 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:45:VAL:HG11	1:A:143:ILE:HD13	1.70	0.73
1:C:143:ILE:HG12	1:C:146:LYS:HE2	1.75	0.68
1:C:354:GLU:OE2	2:D:53:LYS:NZ	2.25	0.67
1:C:238:LYS:HA	1:C:243:LEU:HB2	1.77	0.66
1:C:653:GLN:OE1	1:C:656:ARG:NH1	2.31	0.64

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	656/740 (89%)	642 (98%)	14 (2%)	0	100	100
1	C	674/740 (91%)	660 (98%)	14 (2%)	0	100	100
2	B	398/453 (88%)	394 (99%)	4 (1%)	0	100	100
2	D	398/453 (88%)	394 (99%)	4 (1%)	0	100	100
All	All	2126/2386 (89%)	2090 (98%)	36 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	556/655 (85%)	553 (100%)	3 (0%)	88	94
1	C	574/655 (88%)	562 (98%)	12 (2%)	53	67
2	B	351/416 (84%)	345 (98%)	6 (2%)	60	74
2	D	339/416 (82%)	335 (99%)	4 (1%)	71	83
All	All	1820/2142 (85%)	1795 (99%)	25 (1%)	67	80

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	480	VAL
1	C	596	GLU
2	D	398	GLN
1	C	504	GLN
1	C	671	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

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	665/740 (89%)	0.00	11 (1%) 70 68	34, 59, 92, 126	0
1	C	684/740 (92%)	0.14	23 (3%) 45 43	40, 63, 100, 126	0
2	B	405/453 (89%)	-0.16	5 (1%) 79 77	35, 53, 82, 125	0
2	D	404/453 (89%)	0.09	13 (3%) 47 45	44, 63, 92, 117	0
All	All	2158/2386 (90%)	0.03	52 (2%) 59 56	34, 60, 95, 126	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	365	ILE	4.0
1	C	476	ARG	3.9
1	C	134	LEU	3.8
2	D	314	VAL	3.6
1	C	536	ARG	3.6

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.