WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Apr 9, 2024 – 02:18 PM JST

PDB ID	:	8XMS
Title	:	Crystal structure of Porcine Circovirus type II Rep ATPase domain
Authors	:	Guan, S.Y.; Song, Y.F.
Deposited on	:	2023-12-28
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 (i) 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 (i)) were used in the production of this report:

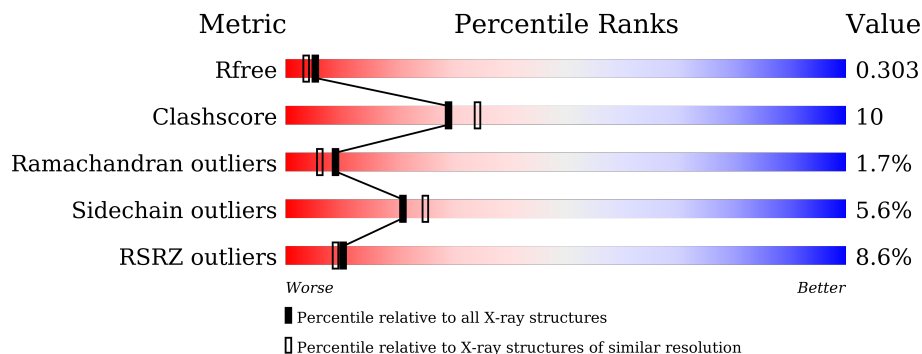|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.20 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	Percentile Ranks	Value
Rfree		0.303
Clashscore		10
Ramachandran outliers		1.7%
Sidechain outliers		5.6%
RSRZ outliers		8.6%

Worse   *Better*

▌Percentile relative to all X-ray structures

▯Percentile relative to X-ray structures of similar resolution

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 <=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	163	5% 71% 13% ·· 13%
1	B	163	7% 75% 10% · 12%
1	C	163	7% 74% 10% · 14%
1	D	163	9% 69% 15% · 13%
1	E	163	6% 74% 12% · 13%
1	F	163	11% 69% 13% ·· 15%

WORLDWIDE wwPDB PROTEIN DATA BANK

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 7076 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 Replication-associated protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	142	Total	C	N	O	S	0	0	0
			1154	749	191	210	4			
1	B	144	Total	C	N	O	S	0	0	0
			1166	756	193	213	4			
1	C	140	Total	C	N	O	S	0	0	0
			1141	742	189	206	4			
1	D	142	Total	C	N	O	S	0	0	0
			1153	749	190	210	4			
1	E	141	Total	C	N	O	S	0	0	0
			1149	747	189	209	4			
1	F	139	Total	C	N	O	S	0	0	0
			1135	739	188	204	4			

There are 66 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP Q8BB16
A	154	GLY	-	expression tag	UNP Q8BB16
A	155	SER	-	expression tag	UNP Q8BB16
A	156	LEU	-	expression tag	UNP Q8BB16
A	157	GLU	-	expression tag	UNP Q8BB16
A	158	HIS	-	expression tag	UNP Q8BB16
A	159	HIS	-	expression tag	UNP Q8BB16
A	160	HIS	-	expression tag	UNP Q8BB16
A	161	HIS	-	expression tag	UNP Q8BB16
A	162	HIS	-	expression tag	UNP Q8BB16
A	163	HIS	-	expression tag	UNP Q8BB16
B	1	MET	-	initiating methionine	UNP Q8BB16
B	154	GLY	-	expression tag	UNP Q8BB16
B	155	SER	-	expression tag	UNP Q8BB16
B	156	LEU	-	expression tag	UNP Q8BB16
B	157	GLU	-	expression tag	UNP Q8BB16
B	158	HIS	-	expression tag	UNP Q8BB16

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	159	HIS	-	expression tag	UNP Q8BB16
B	160	HIS	-	expression tag	UNP Q8BB16
B	161	HIS	-	expression tag	UNP Q8BB16
B	162	HIS	-	expression tag	UNP Q8BB16
B	163	HIS	-	expression tag	UNP Q8BB16
C	1	MET	-	initiating methionine	UNP Q8BB16
C	154	GLY	-	expression tag	UNP Q8BB16
C	155	SER	-	expression tag	UNP Q8BB16
C	156	LEU	-	expression tag	UNP Q8BB16
C	157	GLU	-	expression tag	UNP Q8BB16
C	158	HIS	-	expression tag	UNP Q8BB16
C	159	HIS	-	expression tag	UNP Q8BB16
C	160	HIS	-	expression tag	UNP Q8BB16
C	161	HIS	-	expression tag	UNP Q8BB16
C	162	HIS	-	expression tag	UNP Q8BB16
C	163	HIS	-	expression tag	UNP Q8BB16
D	1	MET	-	initiating methionine	UNP Q8BB16
D	154	GLY	-	expression tag	UNP Q8BB16
D	155	SER	-	expression tag	UNP Q8BB16
D	156	LEU	-	expression tag	UNP Q8BB16
D	157	GLU	-	expression tag	UNP Q8BB16
D	158	HIS	-	expression tag	UNP Q8BB16
D	159	HIS	-	expression tag	UNP Q8BB16
D	160	HIS	-	expression tag	UNP Q8BB16
D	161	HIS	-	expression tag	UNP Q8BB16
D	162	HIS	-	expression tag	UNP Q8BB16
D	163	HIS	-	expression tag	UNP Q8BB16
E	1	MET	-	initiating methionine	UNP Q8BB16
E	154	GLY	-	expression tag	UNP Q8BB16
E	155	SER	-	expression tag	UNP Q8BB16
E	156	LEU	-	expression tag	UNP Q8BB16
E	157	GLU	-	expression tag	UNP Q8BB16
E	158	HIS	-	expression tag	UNP Q8BB16
E	159	HIS	-	expression tag	UNP Q8BB16
E	160	HIS	-	expression tag	UNP Q8BB16
E	161	HIS	-	expression tag	UNP Q8BB16
E	162	HIS	-	expression tag	UNP Q8BB16
E	163	HIS	-	expression tag	UNP Q8BB16
F	1	MET	-	initiating methionine	UNP Q8BB16
F	154	GLY	-	expression tag	UNP Q8BB16
F	155	SER	-	expression tag	UNP Q8BB16
F	156	LEU	-	expression tag	UNP Q8BB16

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	157	GLU	-	expression tag	UNP Q8BB16
F	158	HIS	-	expression tag	UNP Q8BB16
F	159	HIS	-	expression tag	UNP Q8BB16
F	160	HIS	-	expression tag	UNP Q8BB16
F	161	HIS	-	expression tag	UNP Q8BB16
F	162	HIS	-	expression tag	UNP Q8BB16
F	163	HIS	-	expression tag	UNP Q8BB16

• Molecule 2 is water.

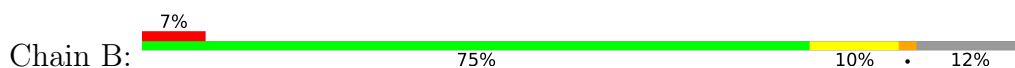| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	44	Total 44	O 44	0	0
2	B	35	Total 35	O 35	0	0
2	C	27	Total 27	O 27	0	0
2	D	25	Total 25	O 25	0	0
2	E	26	Total 26	O 26	0	0
2	F	21	Total 21	O 21	0	0

3 Residue-property plots ⓘ

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.
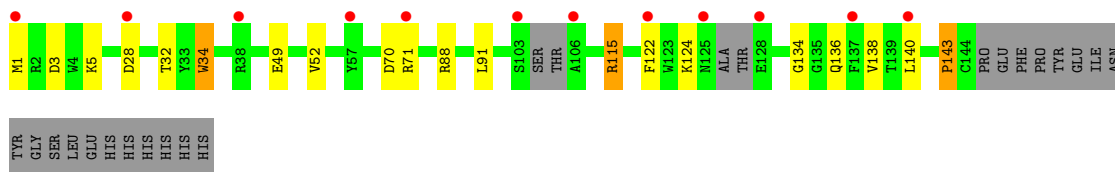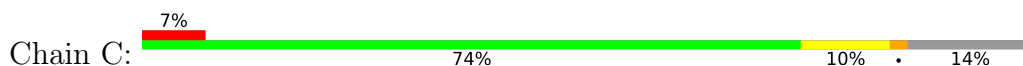
• Molecule 1: Replication-associated protein

Chain A:

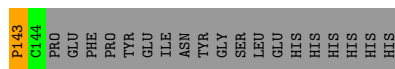• Molecule 1: Replication-associated protein
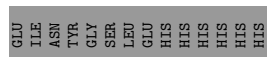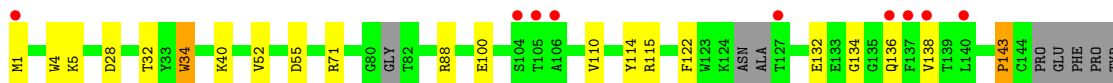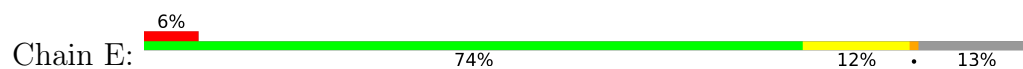
Chain B:

• Molecule 1: Replication-associated protein

Chain C:

• Molecule 1: Replication-associated protein

Chain D:

P143 C144 PRO GLU PHE PRO TYR GLU ILE ASN TYR GLY SER LEU GLU HIS HIS HIS HIS HIS HIS

• Molecule 1: Replication-associated protein

Chain E:

6%

74%

12%   •   13%

M1 W4 K5 D28 T32 Y33 W34 K40 V52 D55 R71 G80 GLY T82 R88 E100 S104 T105 A106 V110 Y114 R115 F122 W123 K124 ASN ALA T127 E132 E133 G134 G135 Q136 F137 V138 T139 L140 P143 C144 PRO GLU PHE PRO TYR

GLU ILE ASN TYR GLY SER LEU GLU HIS HIS HIS HIS HIS HIS

• Molecule 1: Replication-associated protein

Chain F:

11%

69%

13%   ••   15%

M1 C17 G18 K19 A27 D28 T32 Y33 W34 R38 N39 K40 D43 V52 Y57 D64 R67 K79 G80 S89 E100 W101 Y102 SER SER THR A106 V110 Y114 R115 L120 V121 F122 W123 K124 N125 ALA THR E128 G134 G135 Q136 F137 V138

T139 L140 S141 P142 P143 C144 PRO GLU PHE PRO TYR GLU ILE ASN TYR GLY SER LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	67.92Å 103.93Å 76.85Å 90.00° 109.90° 90.00°	Depositor
Resolution (Å)	59.33 – 2.20 59.33 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (59.33-2.20) 100.0 (59.33-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.03 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0415	Depositor
R, R_{free}	0.289 , 0.319 0.283 , 0.303	Depositor DCC
R_{free} test set	2491 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å^2)	39.2	Xtriage
Anisotropy	0.020	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 38.6	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	7076	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.85% 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 (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	0.53	0/1195	0.83	0/1638
1	B	0.56	0/1208	0.85	0/1658
1	C	0.46	0/1181	0.78	0/1617
1	D	0.49	0/1194	0.84	0/1637
1	E	0.49	0/1189	0.82	0/1629
1	F	0.48	0/1175	0.78	0/1609
All	All	0.50	0/7142	0.82	0/9788

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 (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	1154	0	1107	29	0
1	B	1166	0	1120	14	0
1	C	1141	0	1094	17	0
1	D	1153	0	1108	25	0
1	E	1149	0	1104	29	0
1	F	1135	0	1089	23	0
2	A	44	0	0	1	0
2	B	35	0	0	1	0
2	C	27	0	0	4	0
2	D	25	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	26	0	0	2	0
2	F	21	0	0	6	0
All	All	7076	0	6622	130	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 10.

The worst 5 of 130 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:E:5:LYS:HA	1:E:71:ARG:HG2	1.17	1.14
1:D:5:LYS:HA	1:D:71:ARG:HG2	1.16	1.12
1:C:5:LYS:HA	1:C:71:ARG:HG2	1.35	1.08
1:E:132:GLU:HB3	1:E:136:GLN:OE1	1.54	1.06
1:E:5:LYS:CA	1:E:71:ARG:HG2	2.00	0.92

There are no symmetry-related clashes.

5.3   Torsion angles ⓘ

5.3.1   Protein backbone ⓘ

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	138/163 (85%)	129 (94%)	6 (4%)	3 (2%)	6	4
1	B	142/163 (87%)	131 (92%)	8 (6%)	3 (2%)	7	4
1	C	134/163 (82%)	125 (93%)	6 (4%)	3 (2%)	6	4
1	D	138/163 (85%)	130 (94%)	6 (4%)	2 (1%)	11	8
1	E	135/163 (83%)	127 (94%)	7 (5%)	1 (1%)	22	22
1	F	133/163 (82%)	124 (93%)	7 (5%)	2 (2%)	10	8
All	All	820/978 (84%)	766 (93%)	40 (5%)	14 (2%)	9	6

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	134	GLY
1	B	134	GLY
1	C	124	LYS
1	C	134	GLY
1	F	134	GLY

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	125/144 (87%)	116 (93%)	9 (7%)	14	15
1	B	126/144 (88%)	118 (94%)	8 (6%)	18	20
1	C	123/144 (85%)	117 (95%)	6 (5%)	25	31
1	D	125/144 (87%)	118 (94%)	7 (6%)	21	25
1	E	125/144 (87%)	121 (97%)	4 (3%)	39	50
1	F	122/144 (85%)	114 (93%)	8 (7%)	16	19
All	All	746/864 (86%)	704 (94%)	42 (6%)	21	25

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	137	PHE
1	F	38	ARG
1	D	140	LEU
1	E	100	GLU
1	F	100	GLU

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

Mol	Chain	Res	Type
1	A	125	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 monosaccharides 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 ⓘ

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(Å²)	Q<0.9
1	A	142/163 (87%)	0.49	8 (5%)	24	23	22, 35, 55, 82	0
1	B	144/163 (88%)	0.62	12 (8%)	11	10	19, 35, 69, 97	0
1	C	140/163 (85%)	0.71	12 (8%)	10	9	25, 42, 74, 120	0
1	D	142/163 (87%)	0.68	14 (9%)	7	6	25, 40, 81, 96	0
1	E	141/163 (86%)	0.62	9 (6%)	19	18	26, 40, 78, 107	0
1	F	139/163 (85%)	0.74	18 (12%)	3	3	31, 46, 72, 102	0
All	All	848/978 (86%)	0.64	73 (8%)	10	9	19, 40, 74, 120	0

The worst 5 of 73 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	105	THR	5.0
1	C	140	LEU	4.7
1	A	140	LEU	4.6
1	E	137	PHE	4.5
1	F	140	LEU	4.2

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 monosaccharides 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.