



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2026 – 06:46 AM UTC

PDB ID : 8H56 / pdb_00008h56
Title : Crystal structure of Rep' of porcine circovirus type 2
Authors : Guan, S.Y.; Song, Y.F.
Deposited on : 2022-10-12
Resolution : 2.33 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

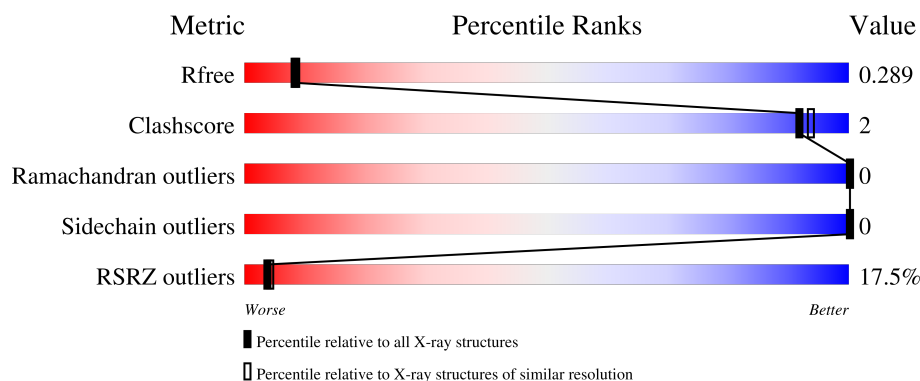
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.33 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}	180053	3031 (2.36-2.32)
Clashscore	190562	3127 (2.36-2.32)
Ramachandran outliers	187476	3095 (2.36-2.32)
Sidechain outliers	187428	3095 (2.36-2.32)
RSRZ outliers	180081	3033 (2.36-2.32)

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	186	15% 77% 5% 17%
1	B	186	15% 79% • 17%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2555 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 Isoform Rep' of Replication-associated protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	154	Total	C	N	O	S	0	0	0
			1241	776	219	238	8			
1	B	154	Total	C	N	O	S	0	0	0
			1241	776	219	238	8			

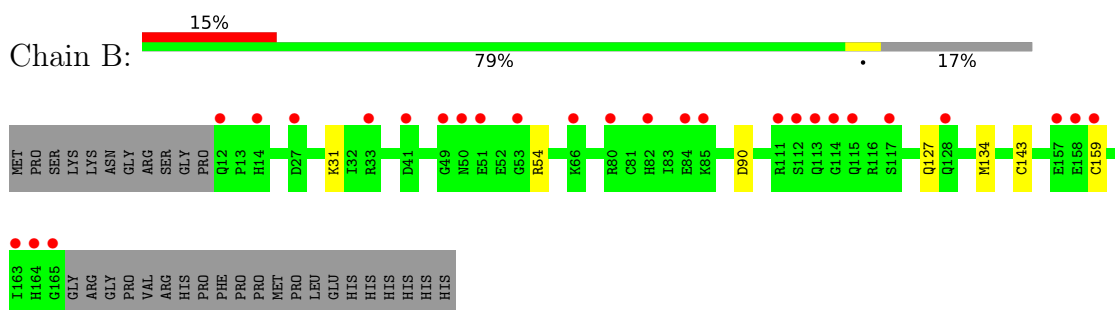
There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	34	ASP	GLU	conflict	UNP Q8BB16
A	35	LEU	PRO	conflict	UNP Q8BB16
A	179	LEU	-	expression tag	UNP Q8BB16
A	180	GLU	-	expression tag	UNP Q8BB16
A	181	HIS	-	expression tag	UNP Q8BB16
A	182	HIS	-	expression tag	UNP Q8BB16
A	183	HIS	-	expression tag	UNP Q8BB16
A	184	HIS	-	expression tag	UNP Q8BB16
A	185	HIS	-	expression tag	UNP Q8BB16
A	186	HIS	-	expression tag	UNP Q8BB16
B	34	ASP	GLU	conflict	UNP Q8BB16
B	35	LEU	PRO	conflict	UNP Q8BB16
B	179	LEU	-	expression tag	UNP Q8BB16
B	180	GLU	-	expression tag	UNP Q8BB16
B	181	HIS	-	expression tag	UNP Q8BB16
B	182	HIS	-	expression tag	UNP Q8BB16
B	183	HIS	-	expression tag	UNP Q8BB16
B	184	HIS	-	expression tag	UNP Q8BB16
B	185	HIS	-	expression tag	UNP Q8BB16
B	186	HIS	-	expression tag	UNP Q8BB16

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	40	Total 40	O 40	0	0
2	B	33	Total 33	O 33	0	0

- Molecule 1: Isoform Rep' of Replication-associated protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	73.12Å 73.12Å 133.85Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.00 – 2.33 46.00 – 2.33	Depositor EDS
% Data completeness (in resolution range)	99.4 (46.00-2.33) 99.6 (46.00-2.33)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.97 (at 2.34Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.253 , 0.283 0.254 , 0.289	Depositor DCC
R_{free} test set	876 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å ²)	47.5	Xtriage
Anisotropy	0.169	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 43.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.027 for -h,-k,l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	2555	wwPDB-VP
Average B, all atoms (Å ²)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.17% 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	0.71	0/1269	1.18	4/1709 (0.2%)
1	B	0.70	0/1269	1.14	2/1709 (0.1%)
All	All	0.70	0/2538	1.16	6/3418 (0.2%)

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	54	ARG	N-CA-C	6.69	119.58	109.95
1	A	107	CYS	CA-C-N	-6.28	112.47	122.65
1	A	107	CYS	C-N-CA	-6.28	112.47	122.65
1	B	54	ARG	N-CA-C	5.96	118.27	109.69
1	B	90	ASP	CA-CB-CG	5.76	118.36	112.60
1	A	90	ASP	CA-CB-CG	5.14	117.74	112.60

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	1241	0	1193	6	0
1	B	1241	0	1193	4	0
2	A	40	0	0	0	0
2	B	33	0	0	1	0
All	All	2555	0	2386	8	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 2.

All (8) 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:B:31:LYS:NZ	2:B:201:HOH:O	2.33	0.61
1:A:117:SER:C	1:A:119:LEU:H	2.11	0.58
1:A:140:CYS:SG	1:B:143:CYS:SG	3.09	0.46
1:A:117:SER:C	1:A:119:LEU:N	2.76	0.44
1:A:119:LEU:HD21	1:A:144:ARG:HG2	1.99	0.43
1:B:127:GLN:C	1:B:134:MET:HE3	2.44	0.43
1:A:163:ILE:HD12	1:B:159:CYS:HB3	2.02	0.41
1:A:127:GLN:C	1:A:134:MET:HE3	2.45	0.41

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	152/186 (82%)	146 (96%)	6 (4%)	0	100	100
1	B	152/186 (82%)	149 (98%)	3 (2%)	0	100	100
All	All	304/372 (82%)	295 (97%)	9 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains

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	137/165 (83%)	137 (100%)	0	100	100
1	B	137/165 (83%)	137 (100%)	0	100	100
All	All	274/330 (83%)	274 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	14	HIS
1	A	164	HIS
1	B	115	GLN

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 oligosaccharides 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	154/186 (82%)	0.99	27 (17%) 4 4	25, 43, 90, 113	0
1	B	154/186 (82%)	1.06	27 (17%) 4 4	28, 41, 89, 116	0
All	All	308/372 (82%)	1.02	54 (17%) 4 4	25, 42, 90, 116	0

All (54) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	140	CYS	7.3
1	B	51	GLU	5.4
1	B	165	GLY	4.6
1	A	165	GLY	4.3
1	A	80	ARG	4.3
1	B	115	GLN	4.2
1	B	113	GLN	4.1
1	A	85	LYS	3.7
1	B	53	GLY	3.7
1	B	112	SER	3.5
1	B	50	ASN	3.5
1	B	14	HIS	3.5
1	A	117	SER	3.5
1	B	111	ARG	3.4
1	B	114	GLY	3.3
1	B	49	GLY	3.2
1	A	111	ARG	3.2
1	B	12	GLN	3.1
1	B	80	ARG	3.0
1	A	163	ILE	3.0
1	A	51	GLU	3.0
1	A	162	THR	2.9
1	B	117	SER	2.9
1	A	118	ASP	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	84	GLU	2.8
1	A	164	HIS	2.7
1	B	159	CYS	2.7
1	B	158	GLU	2.7
1	A	41	ASP	2.6
1	B	41	ASP	2.6
1	A	45	VAL	2.6
1	A	113	GLN	2.5
1	B	163	ILE	2.5
1	A	158	GLU	2.5
1	A	23	ASN	2.5
1	A	53	GLY	2.5
1	B	164	HIS	2.4
1	B	157	GLU	2.4
1	A	55	THR	2.3
1	A	12	GLN	2.3
1	B	27	ASP	2.3
1	A	52	GLU	2.3
1	A	87	LYS	2.3
1	B	66	LYS	2.3
1	A	50	ASN	2.2
1	A	54	ARG	2.2
1	A	49	GLY	2.2
1	B	82	HIS	2.2
1	A	77	LEU	2.1
1	B	85	LYS	2.1
1	A	128	GLN	2.1
1	B	128	GLN	2.1
1	A	115	GLN	2.1
1	B	33	ARG	2.0

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 oligosaccharides 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.