



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 02:12 am BST

PDB ID : 3C6A  
Title : Crystal Structure of the RB49 gp17 nuclease domain  
Authors : Sun, S.; Rossmann, M.G.  
Deposited on : 2008-02-04  
Resolution : 1.16 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](mailto: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

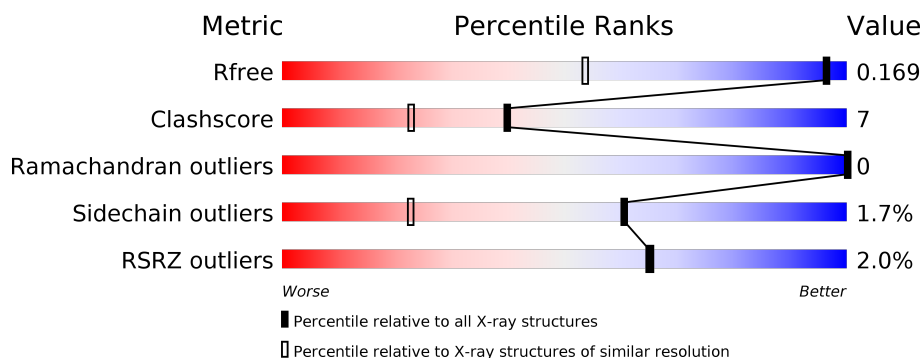
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.16 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	1758 (1.20-1.12)
Clashscore	141614	1832 (1.20-1.12)
Ramachandran outliers	138981	1768 (1.20-1.12)
Sidechain outliers	138945	1768 (1.20-1.12)
RSRZ outliers	127900	1724 (1.20-1.12)

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  $\geq 3$ , 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	232	<div> <div>2%</div> <div> <div></div> <div>72%</div> <div>13%</div> <div>•</div> <div>15%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2024 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 Terminase large subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	198	1635	1051	262	312	10	0	9	0

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	333	MET	-	EXPRESSION TAG	UNP Q9T1C3
A	334	GLY	-	EXPRESSION TAG	UNP Q9T1C3
A	335	SER	-	EXPRESSION TAG	UNP Q9T1C3
A	336	SER	-	EXPRESSION TAG	UNP Q9T1C3
A	337	HIS	-	EXPRESSION TAG	UNP Q9T1C3
A	338	HIS	-	EXPRESSION TAG	UNP Q9T1C3
A	339	HIS	-	EXPRESSION TAG	UNP Q9T1C3
A	340	HIS	-	EXPRESSION TAG	UNP Q9T1C3
A	341	HIS	-	EXPRESSION TAG	UNP Q9T1C3
A	342	HIS	-	EXPRESSION TAG	UNP Q9T1C3
A	343	SER	-	EXPRESSION TAG	UNP Q9T1C3
A	344	SER	-	EXPRESSION TAG	UNP Q9T1C3
A	345	GLY	-	EXPRESSION TAG	UNP Q9T1C3
A	346	LEU	-	EXPRESSION TAG	UNP Q9T1C3
A	347	VAL	-	EXPRESSION TAG	UNP Q9T1C3
A	348	PRO	-	EXPRESSION TAG	UNP Q9T1C3
A	349	ARG	-	EXPRESSION TAG	UNP Q9T1C3
A	350	GLY	-	EXPRESSION TAG	UNP Q9T1C3
A	351	SER	-	EXPRESSION TAG	UNP Q9T1C3
A	352	HIS	-	EXPRESSION TAG	UNP Q9T1C3
A	353	MET	-	EXPRESSION TAG	UNP Q9T1C3
A	354	LEU	-	EXPRESSION TAG	UNP Q9T1C3
A	355	GLU	-	EXPRESSION TAG	UNP Q9T1C3
A	356	ASP	-	EXPRESSION TAG	UNP Q9T1C3
A	357	PRO	-	EXPRESSION TAG	UNP Q9T1C3
A	358	MET	-	EXPRESSION TAG	UNP Q9T1C3

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Mg 1	0	0

- Molecule 3 is water.

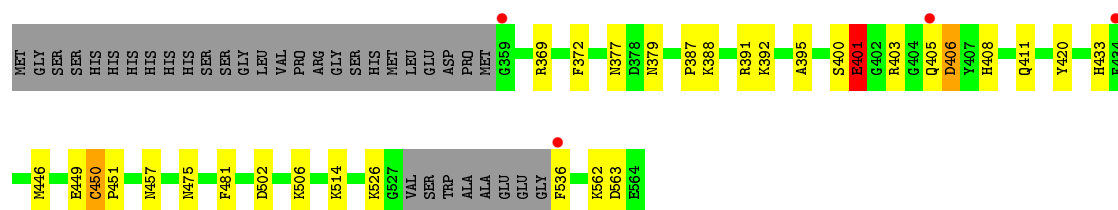
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	388	Total 388	O 388	0	0

### 3 Residue-property plots [i](#)

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: Terminase large subunit

Chain A: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	52.57Å 125.16Å 37.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 1.16 40.25 – 1.16	Depositor EDS
% Data completeness (in resolution range)	94.5 (10.00-1.16) 99.4 (40.25-1.16)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.37 (at 1.16Å)	Xtriage
Refinement program	SHELX, SHELXL-97	Depositor
R, $R_{free}$	0.134 , 0.171 0.133 , 0.169	Depositor DCC
$R_{free}$ test set	4300 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	11.5	Xtriage
Anisotropy	0.236	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 64.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	2024	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.85% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG

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	0.90	2/1692 (0.1%)	1.24	12/2275 (0.5%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	450	CYS	CB-SG	-17.40	1.52	1.82
1	A	450	CYS	CA-CB	11.48	1.79	1.53

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	481	PHE	CB-CG-CD2	-8.29	114.99	120.80
1	A	401	GLU	OE1-CD-OE2	7.46	132.26	123.30
1	A	406[A]	ASP	CB-CG-OD1	7.24	124.81	118.30
1	A	406[B]	ASP	CB-CG-OD1	7.24	124.81	118.30
1	A	563	ASP	CB-CG-OD2	-6.97	112.03	118.30

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	1635	0	1650	24	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	1	0	0	0	0
3	A	388	0	0	10	1
All	All	2024	0	1650	24	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 7.

The worst 5 of 24 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:450:CYS:CB	1:A:450:CYS:CA	1.79	1.61
1:A:400:SER:HB2	1:A:406[B]:ASP:OD1	1.84	0.78
1:A:536:PHE:HB3	3:A:907:HOH:O	1.84	0.78
1:A:450:CYS:SG	1:A:450:CYS:CA	2.71	0.76
1:A:526[B]:LYS:HE3	3:A:865:HOH:O	1.90	0.71

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)
3:A:947:HOH:O	3:A:948:HOH:O[1_554]	2.16	0.04

## 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	204/232 (88%)	199 (98%)	5 (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	186 / 204 (91%)	183 (98%)	3 (2%)	62 25

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	379	ASN
1	A	388	LYS
1	A	401	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	379	ASN
1	A	408	HIS

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	198/232 (85%)	0.06	4 (2%) 65 65	8, 15, 37, 71	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	536	PHE	7.2
1	A	405	GLN	5.1
1	A	434	PHE	3.4
1	A	359	GLY	2.8

### 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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	MG	A	1	1/1	1.00	0.04	19,19,19,19	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.