



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 11, 2021 – 04:17 PM JST

PDB ID : 7EEA  
Title : Cyanophage Pam1 tailspike receptor-binding domain  
Authors : Zhang, J.T.; Jiang, Y.L.; Zhou, C.Z.  
Deposited on : 2021-03-18  
Resolution : 2.67 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 : 2.23.2  
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

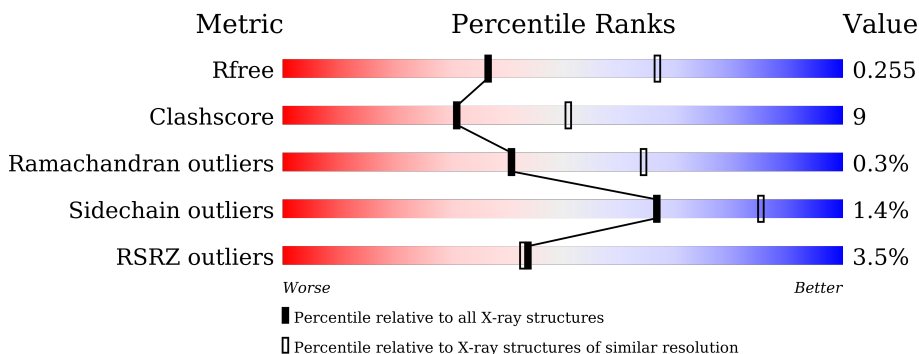
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.67 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	3863 (2.70-2.66)
Clashscore	141614	4210 (2.70-2.66)
Ramachandran outliers	138981	4141 (2.70-2.66)
Sidechain outliers	138945	4141 (2.70-2.66)
RSRZ outliers	127900	3780 (2.70-2.66)

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  $\geq 3$ , 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	664	 77% 20% .
1	B	664	 76% 20% ..
1	C	664	 80% 17% .

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14646 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 Short-tailed cyanophage tailspike receptor-binding domain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	646	4809	2971	871	954	13	0	0	1
1	B	643	4790	2961	869	947	13	0	0	0
1	C	646	4811	2972	872	954	13	0	0	0

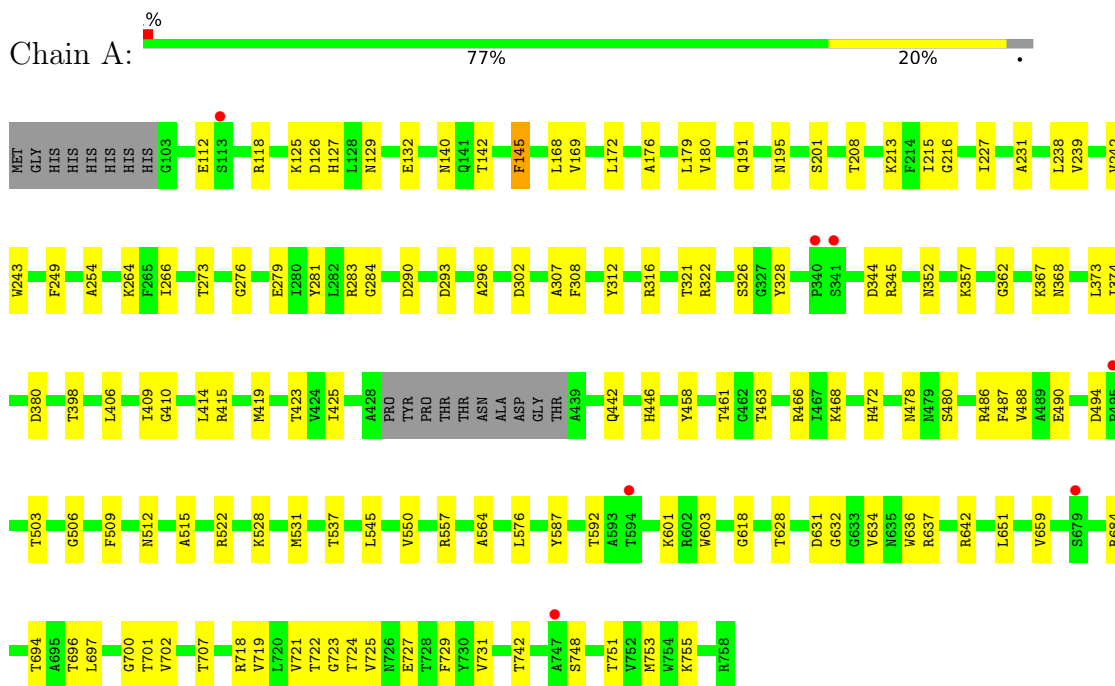
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	85	Total 85	O 85	0	0
2	B	76	Total 76	O 76	0	0
2	C	75	Total 75	O 75	0	0

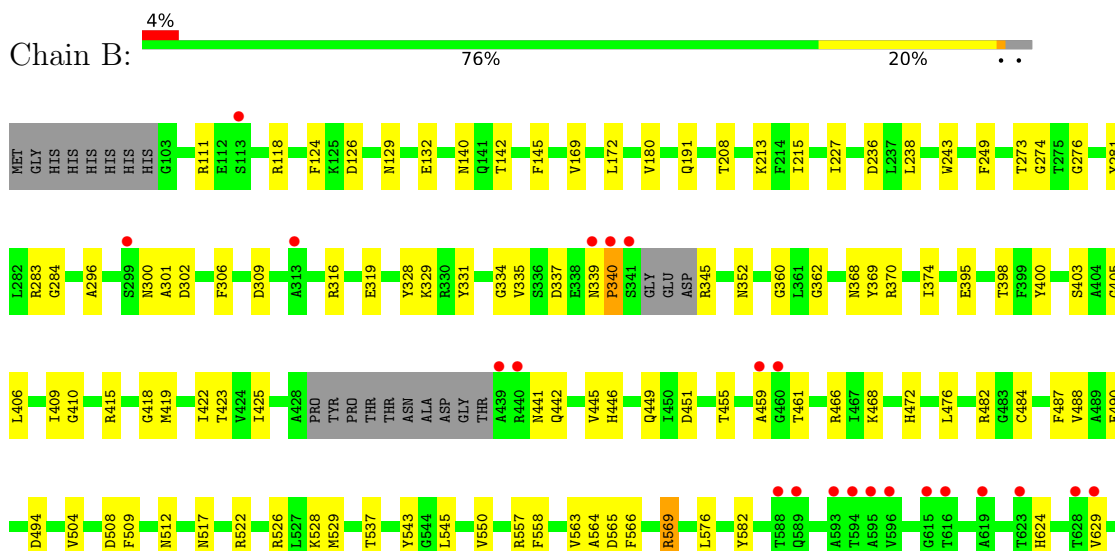
### 3 Residue-property plots

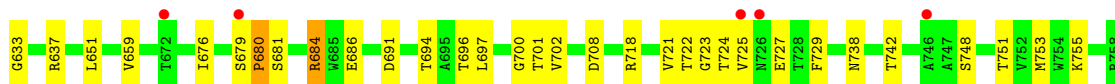
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Short-tailed cyanophage tailspike receptor-binding domain

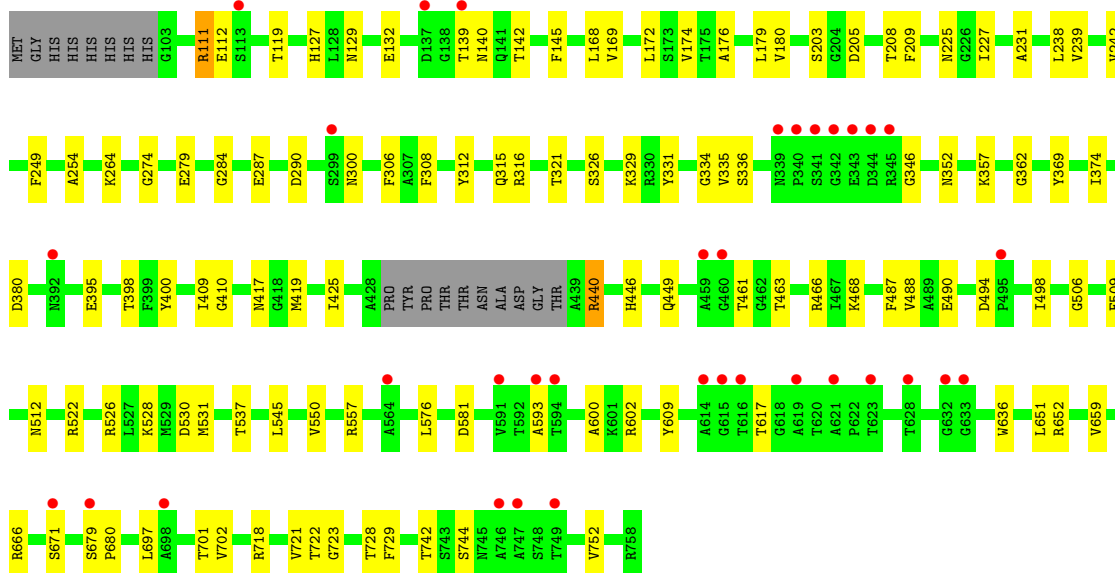
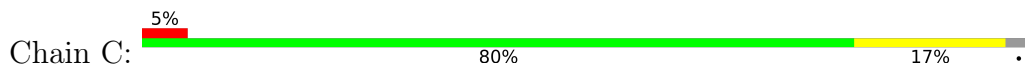


- Molecule 1: Short-tailed cyanophage tailspike receptor-binding domain





• Molecule 1: Short-tailed cyanophage tailspike receptor-binding domain



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	123.58Å 241.27Å 176.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.44 – 2.67 19.44 – 2.67	Depositor EDS
% Data completeness (in resolution range)	97.2 (19.44-2.67) 97.2 (19.44-2.67)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.92 (at 2.67Å)	Xtrriage
Refinement program	PHENIX 1.14_3228	Depositor
R, $R_{free}$	0.210 , 0.255 0.210 , 0.255	Depositor DCC
$R_{free}$ test set	3583 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	44.4	Xtrriage
Anisotropy	0.708	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 45.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	14646	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.00% 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](#)

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	0.26	0/4893	0.48	0/6639
1	B	0.26	0/4873	0.48	0/6611
1	C	0.26	0/4895	0.48	0/6642
All	All	0.26	0/14661	0.48	0/19892

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	4809	0	4629	94	0
1	B	4790	0	4617	99	0
1	C	4811	0	4631	85	0
2	A	85	0	0	7	1
2	B	76	0	0	16	1
2	C	75	0	0	12	0
All	All	14646	0	13877	256	2

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 9.

The worst 5 of 256 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:127:HIS:O	1:C:111:ARG:NH1	1.63	1.30
1:A:512:ASN:H	1:A:537:THR:HG22	1.18	1.08
1:B:124:PHE:O	2:B:801:HOH:O	1.82	0.95
1:C:336:SER:O	2:C:801:HOH:O	1.85	0.95
1:B:738:ASN:OD1	2:B:802:HOH:O	1.84	0.94

All (2) 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)
2:A:868:HOH:O	2:A:876:HOH:O[3_654]	2.13	0.07
2:B:867:HOH:O	2:B:869:HOH:O[4_555]	2.13	0.07

## 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	642/664 (97%)	619 (96%)	22 (3%)	1 (0%)	47 71
1	B	637/664 (96%)	610 (96%)	23 (4%)	4 (1%)	25 47
1	C	642/664 (97%)	613 (96%)	29 (4%)	0	100 100
All	All	1921/1992 (96%)	1842 (96%)	74 (4%)	5 (0%)	41 64

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	564	ALA
1	B	564	ALA
1	B	633	GLY
1	B	340	PRO
1	B	680	PRO



### 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	492/507 (97%)	484 (98%)	8 (2%)	62 83
1	B	490/507 (97%)	483 (99%)	7 (1%)	67 85
1	C	492/507 (97%)	486 (99%)	6 (1%)	71 87
All	All	1474/1521 (97%)	1453 (99%)	21 (1%)	67 85

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	718	ARG
1	C	145	PHE
1	C	671	SER
1	C	440	ARG
1	C	112	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	449	GLN
1	B	473	GLN
1	C	358	ASN
1	C	442	GLN
1	C	449	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 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 [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	646/664 (97%)	-0.29	7 (1%) 80 81	32, 42, 60, 87	0
1	B	643/664 (96%)	-0.10	27 (4%) 36 34	36, 49, 76, 117	0
1	C	646/664 (97%)	-0.06	34 (5%) 26 24	34, 47, 84, 116	0
All	All	1935/1992 (97%)	-0.15	68 (3%) 44 43	32, 46, 70, 117	0

The worst 5 of 68 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	595	ALA	7.3
1	B	341	SER	7.2
1	A	341	SER	5.8
1	C	340	PRO	5.4
1	C	593	ALA	5.0

### 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.