



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 24, 2023 – 08:56 AM EDT

PDB ID : 5UDO  
Title : Crystal structure of the coiled-coil domain from Listeria Innocua Phage Integrase (Tetragonal Form II)  
Authors : Gupta, K.; Yuan, J.B.; Sharp, R.; Van Duyne, G.D.  
Deposited on : 2016-12-28  
Resolution : 2.54 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 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 : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

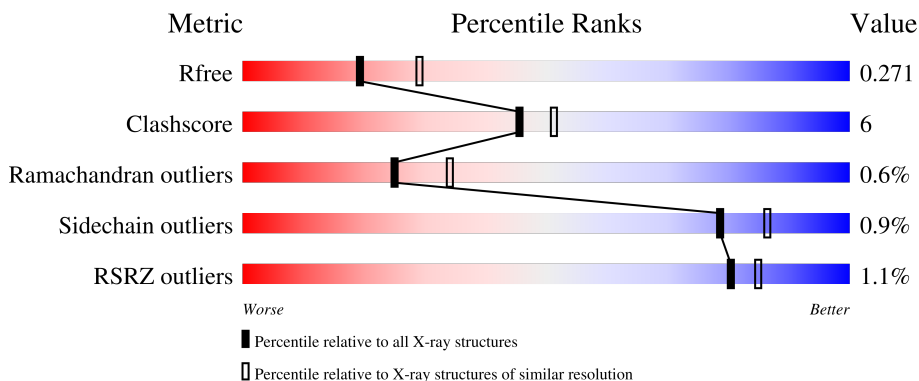
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


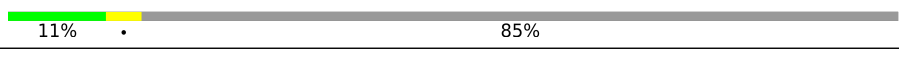

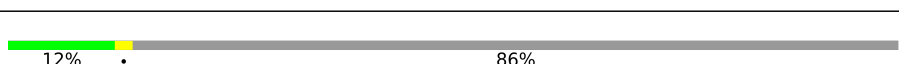

The reported resolution of this entry is 2.54 Å.

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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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	328	 13% . 85%
1	B	328	 11% . 85%
1	C	328	 11% . 87%
1	D	328	 11% . 85%
1	E	328	 12% . 86%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	328	 11% 85%
1	G	328	 12% 87%
1	H	328	 15% 85%

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 3881 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 A118 serine integrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	48	398	251	64	81	2	0	0	0
1	B	49	406	257	65	82	2	0	0	0
1	C	44	359	228	58	71	2	0	0	0
1	D	48	398	252	64	80	2	0	0	0
1	E	47	392	249	63	78	2	0	0	0
1	F	48	398	252	64	80	2	0	0	0
1	G	42	353	224	57	70	2	0	0	0
1	H	48	398	251	64	81	2	0	0	0

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	453	LEU	-	expression tag	UNP Q928V6
A	454	GLU	-	expression tag	UNP Q928V6
A	455	HIS	-	expression tag	UNP Q928V6
A	456	HIS	-	expression tag	UNP Q928V6
A	457	HIS	-	expression tag	UNP Q928V6
A	458	HIS	-	expression tag	UNP Q928V6
A	459	HIS	-	expression tag	UNP Q928V6
A	460	HIS	-	expression tag	UNP Q928V6
B	453	LEU	-	expression tag	UNP Q928V6
B	454	GLU	-	expression tag	UNP Q928V6
B	455	HIS	-	expression tag	UNP Q928V6
B	456	HIS	-	expression tag	UNP Q928V6
B	457	HIS	-	expression tag	UNP Q928V6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	458	HIS	-	expression tag	UNP Q928V6
B	459	HIS	-	expression tag	UNP Q928V6
B	460	HIS	-	expression tag	UNP Q928V6
C	453	LEU	-	expression tag	UNP Q928V6
C	454	GLU	-	expression tag	UNP Q928V6
C	455	HIS	-	expression tag	UNP Q928V6
C	456	HIS	-	expression tag	UNP Q928V6
C	457	HIS	-	expression tag	UNP Q928V6
C	458	HIS	-	expression tag	UNP Q928V6
C	459	HIS	-	expression tag	UNP Q928V6
C	460	HIS	-	expression tag	UNP Q928V6
D	453	LEU	-	expression tag	UNP Q928V6
D	454	GLU	-	expression tag	UNP Q928V6
D	455	HIS	-	expression tag	UNP Q928V6
D	456	HIS	-	expression tag	UNP Q928V6
D	457	HIS	-	expression tag	UNP Q928V6
D	458	HIS	-	expression tag	UNP Q928V6
D	459	HIS	-	expression tag	UNP Q928V6
D	460	HIS	-	expression tag	UNP Q928V6
E	453	LEU	-	expression tag	UNP Q928V6
E	454	GLU	-	expression tag	UNP Q928V6
E	455	HIS	-	expression tag	UNP Q928V6
E	456	HIS	-	expression tag	UNP Q928V6
E	457	HIS	-	expression tag	UNP Q928V6
E	458	HIS	-	expression tag	UNP Q928V6
E	459	HIS	-	expression tag	UNP Q928V6
E	460	HIS	-	expression tag	UNP Q928V6
F	453	LEU	-	expression tag	UNP Q928V6
F	454	GLU	-	expression tag	UNP Q928V6
F	455	HIS	-	expression tag	UNP Q928V6
F	456	HIS	-	expression tag	UNP Q928V6
F	457	HIS	-	expression tag	UNP Q928V6
F	458	HIS	-	expression tag	UNP Q928V6
F	459	HIS	-	expression tag	UNP Q928V6
F	460	HIS	-	expression tag	UNP Q928V6
G	453	LEU	-	expression tag	UNP Q928V6
G	454	GLU	-	expression tag	UNP Q928V6
G	455	HIS	-	expression tag	UNP Q928V6
G	456	HIS	-	expression tag	UNP Q928V6
G	457	HIS	-	expression tag	UNP Q928V6
G	458	HIS	-	expression tag	UNP Q928V6
G	459	HIS	-	expression tag	UNP Q928V6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	460	HIS	-	expression tag	UNP Q928V6
H	453	LEU	-	expression tag	UNP Q928V6
H	454	GLU	-	expression tag	UNP Q928V6
H	455	HIS	-	expression tag	UNP Q928V6
H	456	HIS	-	expression tag	UNP Q928V6
H	457	HIS	-	expression tag	UNP Q928V6
H	458	HIS	-	expression tag	UNP Q928V6
H	459	HIS	-	expression tag	UNP Q928V6
H	460	HIS	-	expression tag	UNP Q928V6

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	115	Total O 115 115	0	0
2	B	101	Total O 101 101	0	0
2	C	87	Total O 87 87	0	0
2	D	94	Total O 94 94	0	0
2	E	102	Total O 102 102	0	0
2	F	92	Total O 92 92	0	0
2	G	86	Total O 86 86	0	0
2	H	102	Total O 102 102	0	0









THR  
VAL  
ASP  
PHE  
ASN  
SER  
LEU  
GLU  
PHE  
ARG  
GLU  
LYS  
GLN  
LEU  
TYR  
LEU  
LYS  
SER  
LEU  
ILE  
ASN  
LYS  
ILE  
TYR  
ILE  
ASP  
GLY  
GLN  
VAL  
THR  
ILE  
GLU  
TRP  
LEU  
LEU  
GLU  
HIS  
HIS  
HIS  
HIS  
HIS

## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	98.02Å 98.02Å 52.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.85 – 2.54 49.01 – 2.53	Depositor EDS
% Data completeness (in resolution range)	97.5 (35.85-2.54) 94.7 (49.01-2.53)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.97 (at 2.54Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.252 , 0.265 0.260 , 0.271	Depositor DCC
$R_{free}$ test set	800 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	79.0	Xtrriage
Anisotropy	0.069	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , 76.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.470 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3881	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	92.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.30% 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	0.22	0/403	0.32	0/539
1	B	0.21	0/411	0.32	0/550
1	C	0.21	0/364	0.35	0/488
1	D	0.21	0/403	0.32	0/539
1	E	0.22	0/397	0.33	0/531
1	F	0.21	0/403	0.35	0/539
1	G	0.22	0/358	0.33	0/478
1	H	0.21	0/403	0.31	0/539
All	All	0.21	0/3142	0.33	0/4203

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	398	0	385	6	0
1	B	406	0	396	8	0
1	C	359	0	345	6	0
1	D	398	0	390	12	0
1	E	392	0	385	6	0
1	F	398	0	390	9	0
1	G	353	0	344	3	0
1	H	398	0	385	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	115	0	0	0	0
2	B	101	0	0	2	0
2	C	87	0	0	0	0
2	D	94	0	0	2	0
2	E	102	0	0	2	0
2	F	92	0	0	2	1
2	G	86	0	0	1	1
2	H	102	0	0	0	0
All	All	3881	0	3020	39	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 6.

The worst 5 of 39 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:D:367:ASP:O	1:D:371:ASN:ND2	2.26	0.69
1:A:364:ARG:NH1	1:B:369:TYR:O	2.29	0.62
1:C:364:ARG:NH2	1:D:374:TYR:O	2.36	0.59
1:B:398:GLU:OE1	1:F:350:SER:N	2.36	0.59
1:G:354:LYS:NZ	2:G:506:HOH:O	2.36	0.58

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)
2:F:508:HOH:O	2:G:516:HOH:O[3_565]	2.08	0.12

## 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	46/328 (14%)	43 (94%)	3 (6%)	0	100	100
1	B	47/328 (14%)	43 (92%)	3 (6%)	1 (2%)	7	7
1	C	42/328 (13%)	39 (93%)	3 (7%)	0	100	100
1	D	46/328 (14%)	44 (96%)	1 (2%)	1 (2%)	6	7
1	E	45/328 (14%)	43 (96%)	2 (4%)	0	100	100
1	F	46/328 (14%)	45 (98%)	1 (2%)	0	100	100
1	G	40/328 (12%)	40 (100%)	0	0	100	100
1	H	46/328 (14%)	45 (98%)	1 (2%)	0	100	100
All	All	358/2624 (14%)	342 (96%)	14 (4%)	2 (1%)	25	34

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	396	GLN
1	D	396	GLN

### 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	43/302 (14%)	43 (100%)	0	100	100
1	B	44/302 (15%)	44 (100%)	0	100	100
1	C	39/302 (13%)	38 (97%)	1 (3%)	46	61
1	D	44/302 (15%)	44 (100%)	0	100	100
1	E	43/302 (14%)	43 (100%)	0	100	100
1	F	44/302 (15%)	42 (96%)	2 (4%)	27	37
1	G	39/302 (13%)	39 (100%)	0	100	100
1	H	43/302 (14%)	43 (100%)	0	100	100
All	All	339/2416 (14%)	336 (99%)	3 (1%)	78	86

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

Mol	Chain	Res	Type
1	C	359	HIS
1	F	351	LEU
1	F	366	PHE

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 [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	48/328 (14%)	-0.31	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	76, 103, 139, 147	0
1	B	49/328 (14%)	-0.24	1 (2%) <span style="border: 1px solid blue; padding: 2px;">65</span> <span style="border: 1px solid blue; padding: 2px;">72</span>	76, 100, 132, 162	0
1	C	44/328 (13%)	-0.30	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	78, 95, 111, 125	0
1	D	48/328 (14%)	-0.30	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	74, 90, 100, 101	0
1	E	47/328 (14%)	-0.31	1 (2%) <span style="border: 1px solid blue; padding: 2px;">63</span> <span style="border: 1px solid blue; padding: 2px;">70</span>	78, 98, 120, 127	0
1	F	48/328 (14%)	-0.25	2 (4%) <span style="border: 1px solid red; padding: 2px;">36</span> <span style="border: 1px solid red; padding: 2px;">42</span>	84, 99, 122, 131	0
1	G	42/328 (12%)	-0.22	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	80, 108, 128, 131	0
1	H	48/328 (14%)	-0.24	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	82, 91, 117, 135	0
All	All	374/2624 (14%)	-0.27	4 (1%) <span style="border: 1px solid blue; padding: 2px;">80</span> <span style="border: 1px solid blue; padding: 2px;">85</span>	74, 96, 128, 162	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	397	ILE	3.3
1	F	355	LEU	2.4
1	E	397	ILE	2.1
1	B	393	TYR	2.1

### 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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.