



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 10, 2017 – 12:24 PM EDT

PDB ID : 2GM5  
Title : An activated, truncated gamma-delta resolvase tetramer  
Authors : Kamtekar, S.; Ho, R.S.; Li, W.; Steitz, T.A.  
Deposited on : unknown  
Resolution : 2.10 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

<http://wwpdb.org/validation/2016/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) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

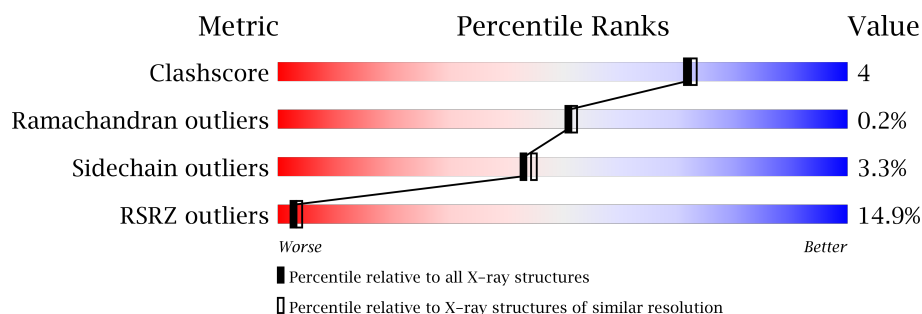
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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))
Clashscore	112137	4788 (2.10-2.10)
Ramachandran outliers	110173	4740 (2.10-2.10)
Sidechain outliers	110143	4741 (2.10-2.10)
RSRZ outliers	101464	4275 (2.10-2.10)

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. 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	139	<div> <div>13%</div> <div> <div></div> <div>76%</div> <div>12%</div> <div>12%</div> </div> </div>
1	B	139	<div> <div>7%</div> <div> <div></div> <div>71%</div> <div>10%</div> <div>•</div> <div>18%</div> </div> </div>
1	C	139	<div> <div>14%</div> <div> <div></div> <div>68%</div> <div>15%</div> <div>17%</div> </div> </div>
1	D	139	<div> <div>14%</div> <div> <div></div> <div>74%</div> <div>9%</div> <div>•</div> <div>16%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3817 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 Transposon gamma-delta resolvase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	122	Total	C	N	O	Se	0	0	0
			955	597	173	182	3			
1	B	114	Total	C	N	O	Se	0	0	0
			887	560	159	165	3			
1	C	116	Total	C	N	O	Se	0	0	0
			906	570	164	169	3			
1	D	117	Total	C	N	O	Se	0	0	0
			922	579	168	172	3			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	ALA	ARG	ENGINEERED	UNP P03012
A	53	MSE	MET	MODIFIED RESIDUE	UNP P03012
A	56	LYS	GLU	ENGINEERED	UNP P03012
A	68	HIS	ARG	ENGINEERED	UNP P03012
A	76	MSE	MET	MODIFIED RESIDUE	UNP P03012
A	101	SER	GLY	ENGINEERED	UNP P03012
A	102	TYR	GLU	ENGINEERED	UNP P03012
A	103	ILE	MET	ENGINEERED	UNP P03012
A	106	MSE	MET	MODIFIED RESIDUE	UNP P03012
A	135	HIS	-	EXPRESSION TAG	UNP P03012
A	136	HIS	-	EXPRESSION TAG	UNP P03012
A	137	HIS	-	EXPRESSION TAG	UNP P03012
A	138	HIS	-	EXPRESSION TAG	UNP P03012
A	139	HIS	-	EXPRESSION TAG	UNP P03012
A	140	HIS	-	EXPRESSION TAG	UNP P03012
B	2	ALA	ARG	ENGINEERED	UNP P03012
B	53	MSE	MET	MODIFIED RESIDUE	UNP P03012
B	56	LYS	GLU	ENGINEERED	UNP P03012
B	68	HIS	ARG	ENGINEERED	UNP P03012
B	76	MSE	MET	MODIFIED RESIDUE	UNP P03012
B	101	SER	GLY	ENGINEERED	UNP P03012

*Continued on next page...*

*Continued from previous page...*

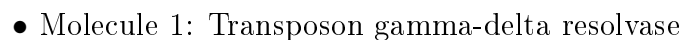
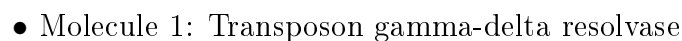
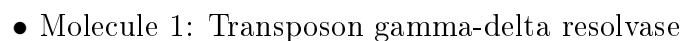
Chain	Residue	Modelled	Actual	Comment	Reference
B	102	TYR	GLU	ENGINEERED	UNP P03012
B	103	ILE	MET	ENGINEERED	UNP P03012
B	106	MSE	MET	MODIFIED RESIDUE	UNP P03012
B	135	HIS	-	EXPRESSION TAG	UNP P03012
B	136	HIS	-	EXPRESSION TAG	UNP P03012
B	137	HIS	-	EXPRESSION TAG	UNP P03012
B	138	HIS	-	EXPRESSION TAG	UNP P03012
B	139	HIS	-	EXPRESSION TAG	UNP P03012
B	140	HIS	-	EXPRESSION TAG	UNP P03012
C	2	ALA	ARG	ENGINEERED	UNP P03012
C	53	MSE	MET	MODIFIED RESIDUE	UNP P03012
C	56	LYS	GLU	ENGINEERED	UNP P03012
C	68	HIS	ARG	ENGINEERED	UNP P03012
C	76	MSE	MET	MODIFIED RESIDUE	UNP P03012
C	101	SER	GLY	ENGINEERED	UNP P03012
C	102	TYR	GLU	ENGINEERED	UNP P03012
C	103	ILE	MET	ENGINEERED	UNP P03012
C	106	MSE	MET	MODIFIED RESIDUE	UNP P03012
C	135	HIS	-	EXPRESSION TAG	UNP P03012
C	136	HIS	-	EXPRESSION TAG	UNP P03012
C	137	HIS	-	EXPRESSION TAG	UNP P03012
C	138	HIS	-	EXPRESSION TAG	UNP P03012
C	139	HIS	-	EXPRESSION TAG	UNP P03012
C	140	HIS	-	EXPRESSION TAG	UNP P03012
D	2	ALA	ARG	ENGINEERED	UNP P03012
D	53	MSE	MET	MODIFIED RESIDUE	UNP P03012
D	56	LYS	GLU	ENGINEERED	UNP P03012
D	68	HIS	ARG	ENGINEERED	UNP P03012
D	76	MSE	MET	MODIFIED RESIDUE	UNP P03012
D	101	SER	GLY	ENGINEERED	UNP P03012
D	102	TYR	GLU	ENGINEERED	UNP P03012
D	103	ILE	MET	ENGINEERED	UNP P03012
D	106	MSE	MET	MODIFIED RESIDUE	UNP P03012
D	135	HIS	-	EXPRESSION TAG	UNP P03012
D	136	HIS	-	EXPRESSION TAG	UNP P03012
D	137	HIS	-	EXPRESSION TAG	UNP P03012
D	138	HIS	-	EXPRESSION TAG	UNP P03012
D	139	HIS	-	EXPRESSION TAG	UNP P03012
D	140	HIS	-	EXPRESSION TAG	UNP P03012

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	44	Total 44	O 44	0	0
2	B	36	Total 36	O 36	0	0
2	C	39	Total 39	O 39	0	0
2	D	28	Total 28	O 28	0	0



- Molecule 1: Transposon gamma-delta resolvase



GLU
GLY
ARG
GLN
GLU
ALA
MET
HIS
HIS
HIS
HIS
HIS
HIS

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	53.33Å 98.48Å 100.41Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.14 – 2.10 47.10 – 2.10	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.14-2.10) 99.7 (47.10-2.10)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.24 (at 2.10Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.215 , 0.255 0.218 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	33.2	Xtriage
Anisotropy	0.508	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 45.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.019 for -h,l,k	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3817	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.39% 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	0.34	0/957	0.51	0/1275
1	B	0.37	0/888	0.54	0/1181
1	C	0.35	0/907	0.52	0/1205
1	D	0.34	0/923	0.50	0/1227
All	All	0.35	0/3675	0.52	0/4888

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	955	0	995	7	0
1	B	887	0	935	6	0
1	C	906	0	952	11	0
1	D	922	0	966	6	0
2	A	44	0	0	0	0
2	B	36	0	0	0	0
2	C	39	0	0	1	0
2	D	28	0	0	0	0
All	All	3817	0	3848	30	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 4.

The worst 5 of 30 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:C:62:LEU:HB3	1:C:93:ILE:HD11	1.78	0.66
1:D:116:GLN:HE21	1:D:119:ARG:HH12	1.44	0.65
1:A:78:GLN:HE21	1:A:82:GLU:HG3	1.67	0.59
1:B:7:ALA:HB3	1:B:35:THR:HG22	1.89	0.55
1:D:76:MSE:HE3	1:D:77:ILE:HD13	1.89	0.54

There are no symmetry-related clashes.

## 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	118/139 (85%)	115 (98%)	3 (2%)	0	100	100
1	B	108/139 (78%)	105 (97%)	3 (3%)	0	100	100
1	C	110/139 (79%)	107 (97%)	2 (2%)	1 (1%)	20	14
1	D	111/139 (80%)	110 (99%)	1 (1%)	0	100	100
All	All	447/556 (80%)	437 (98%)	9 (2%)	1 (0%)	51	52

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	69	LEU

### 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	103/114 (90%)	100 (97%)	3 (3%)	48	51
1	B	95/114 (83%)	91 (96%)	4 (4%)	34	33
1	C	97/114 (85%)	95 (98%)	2 (2%)	59	64
1	D	99/114 (87%)	95 (96%)	4 (4%)	36	36
All	All	394/456 (86%)	381 (97%)	13 (3%)	43	45

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	111	LEU
1	B	121	ARG
1	D	51	LEU
1	B	55	VAL
1	D	45	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	19	GLN
1	D	120	GLN
1	C	120	GLN
1	A	78	GLN
1	D	116	GLN

### 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 carbohydrates 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 [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	119/139 (85%)	1.03	18 (15%) 3 4	36, 41, 54, 59	0
1	B	111/139 (79%)	0.80	10 (9%) 10 13	34, 41, 48, 55	0
1	C	113/139 (81%)	1.25	20 (17%) 2 2	31, 42, 51, 57	0
1	D	114/139 (82%)	1.13	20 (17%) 2 2	28, 41, 57, 64	0
All	All	457/556 (82%)	1.05	68 (14%) 3 4	28, 41, 53, 64	0

The worst 5 of 68 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	126	THR	8.2
1	C	38	ALA	7.7
1	D	14	GLN	6.7
1	D	15	SER	6.3
1	A	38	ALA	6.1

### 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 carbohydrates 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.