WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 11, 2021 – 02:32 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	:	2006-04-05
Resolution	:	2.10 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

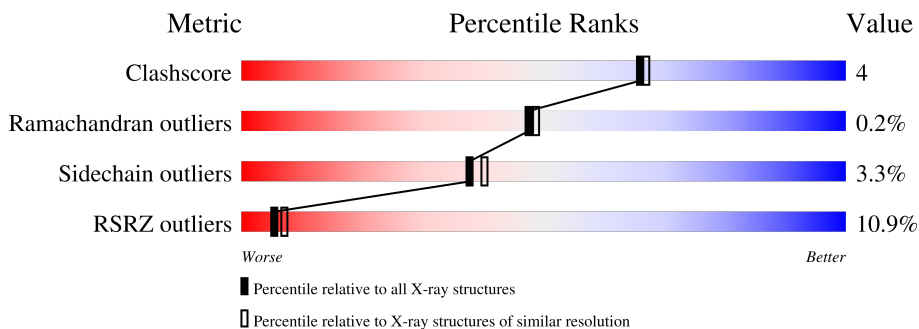|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 2.10 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	Percentile Ranks	Value
Clashscore		4
Ramachandran outliers		0.2%
Sidechain outliers		3.3%
RSRZ outliers		10.9%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (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 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 <=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	8% / 76% / 12% / 12%
1	B	139	6% / 71% / 10% / · / 18%
1	C	139	10% / 68% / 15% / 17%
1	D	139	12% / 74% / 9% / · / 16%

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 mutation	UNP P03012
A	53	MSE	MET	modified residue	UNP P03012
A	56	LYS	GLU	engineered mutation	UNP P03012
A	68	HIS	ARG	engineered mutation	UNP P03012
A	76	MSE	MET	modified residue	UNP P03012
A	101	SER	GLY	engineered mutation	UNP P03012
A	102	TYR	GLU	engineered mutation	UNP P03012
A	103	ILE	MET	engineered mutation	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 mutation	UNP P03012
B	53	MSE	MET	modified residue	UNP P03012
B	56	LYS	GLU	engineered mutation	UNP P03012
B	68	HIS	ARG	engineered mutation	UNP P03012
B	76	MSE	MET	modified residue	UNP P03012
B	101	SER	GLY	engineered mutation	UNP P03012

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	102	TYR	GLU	engineered mutation	UNP P03012
B	103	ILE	MET	engineered mutation	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 mutation	UNP P03012
C	53	MSE	MET	modified residue	UNP P03012
C	56	LYS	GLU	engineered mutation	UNP P03012
C	68	HIS	ARG	engineered mutation	UNP P03012
C	76	MSE	MET	modified residue	UNP P03012
C	101	SER	GLY	engineered mutation	UNP P03012
C	102	TYR	GLU	engineered mutation	UNP P03012
C	103	ILE	MET	engineered mutation	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 mutation	UNP P03012
D	53	MSE	MET	modified residue	UNP P03012
D	56	LYS	GLU	engineered mutation	UNP P03012
D	68	HIS	ARG	engineered mutation	UNP P03012
D	76	MSE	MET	modified residue	UNP P03012
D	101	SER	GLY	engineered mutation	UNP P03012
D	102	TYR	GLU	engineered mutation	UNP P03012
D	103	ILE	MET	engineered mutation	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

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.
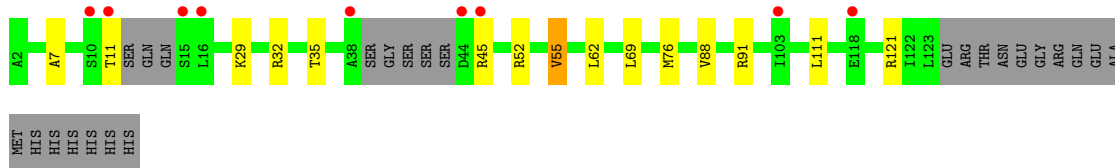
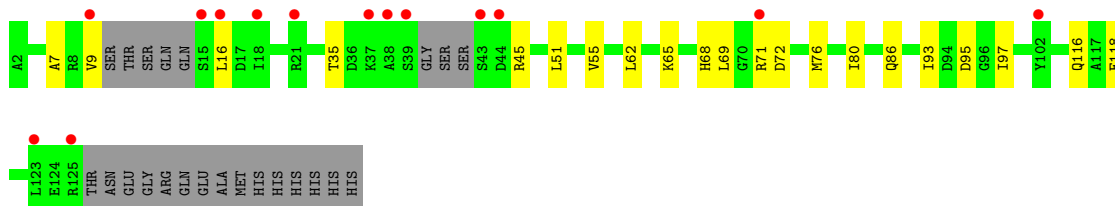- Molecule 1: Transposon gamma-delta resolvase

Chain A:
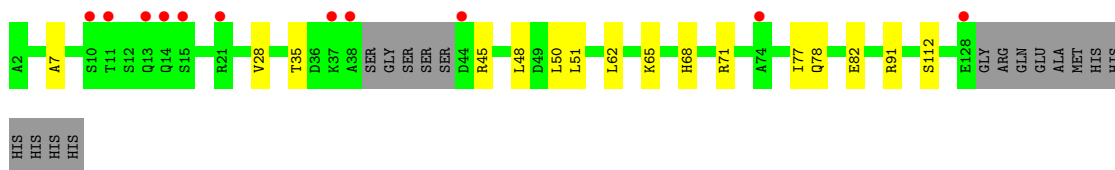
- Molecule 1: Transposon gamma-delta resolvase

Chain B:

- Molecule 1: Transposon gamma-delta resolvase

Chain C:

- Molecule 1: Transposon gamma-delta resolvase

Chain D:

GLU ALA MET 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, α, β, γ	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
$< I/\sigma(I) >$ [1]	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.	wwPDB-VP
Wilson B-factor (Å^2)	33.2	Xtriage
Anisotropy	0.508	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 54.9	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.019 for -h,l,k	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3817	wwPDB-VP
Average B, all atoms (Å^2)	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.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

All (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
1:A:7:ALA:HB3	1:A:35:THR:HG22	1.89	0.52
1:D:72:ASP:HB2	1:D:116:GLN:HE22	1.75	0.51
1:A:62:LEU:HD23	1:A:91:ARG:HB3	1.92	0.50
1:C:65:LYS:HB2	1:C:68:HIS:CD2	2.48	0.49
1:B:62:LEU:HD22	1:B:91:ARG:HB3	1.95	0.48
1:A:77:ILE:HG21	1:A:112:SER:HB3	1.96	0.47
1:C:7:ALA:HB3	1:C:35:THR:HG22	1.95	0.47
1:C:45:ARG:HE	1:C:71:ARG:HH21	1.62	0.46
1:D:117:ALA:HA	1:D:120:GLN:HE21	1.81	0.46
1:D:52:ARG:HH12	1:D:82:GLU:CD	2.19	0.46
1:A:45:ARG:HD2	1:A:48:LEU:HD23	1.98	0.45
1:C:95:ASP:O	1:C:97:ILE:HD12	2.17	0.45
1:D:55:VAL:HG13	1:D:88:VAL:HG21	1.97	0.45
1:B:55:VAL:HG13	1:B:88:VAL:HG21	1.99	0.45
1:C:76:MSE:HE3	1:C:80:ILE:HD12	2.00	0.44
1:C:72:ASP:HB2	1:C:116:GLN:OE1	2.18	0.44
1:B:45:ARG:HE	1:B:52:ARG:HH12	1.65	0.44
1:C:118:GLU:HG2	2:C:172:HOH:O	2.17	0.44
1:C:9:VAL:HG13	1:C:16:LEU:HG	2.00	0.43
1:B:69:LEU:HB3	1:B:76:MSE:HG3	2.00	0.42
1:B:29:LYS:HB2	1:B:32:ARG:HG3	2.02	0.41
1:A:65:LYS:HB2	1:A:68:HIS:CD2	2.55	0.41
1:C:86:GLN:O	1:C:86:GLN:HG3	2.21	0.41
1:A:28:VAL:HG23	1:A:62:LEU:HD12	2.03	0.41
1:C:45:ARG:HE	1:C:71:ARG:NH2	2.19	0.40

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	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%)	17	12
1	D	111/139 (80%)	110 (99%)	1 (1%)	0	100	100
All	All	447/556 (80%)	437 (98%)	9 (2%)	1 (0%)	47	49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	69	LEU

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	103/114 (90%)	100 (97%)	3 (3%)	42	46
1	B	95/114 (83%)	91 (96%)	4 (4%)	30	30
1	C	97/114 (85%)	95 (98%)	2 (2%)	53	59
1	D	99/114 (87%)	95 (96%)	4 (4%)	31	32
All	All	394/456 (86%)	381 (97%)	13 (3%)	38	40

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	50	LEU
1	A	51	LEU
1	A	71	ARG
1	B	11	THR
1	B	55	VAL
1	B	111	LEU
1	B	121	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	51	LEU
1	C	55	VAL
1	D	45	ARG
1	D	51	LEU
1	D	55	VAL
1	D	66	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	68	HIS
1	A	78	GLN
1	B	19	GLN
1	C	120	GLN
1	D	116	GLN
1	D	120	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 monosaccharides 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 (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	119/139 (85%)	0.62	11 (9%)	9	11	36, 41, 54, 59	0
1	B	111/139 (79%)	0.47	9 (8%)	12	15	34, 41, 48, 55	0
1	C	113/139 (81%)	0.88	14 (12%)	4	5	31, 42, 51, 57	0
1	D	114/139 (82%)	0.89	16 (14%)	2	3	28, 41, 57, 64	0
All	All	457/556 (82%)	0.71	50 (10%)	5	7	28, 41, 53, 64	0

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	38	ALA	10.0
1	D	126	THR	7.0
1	D	15	SER	6.2
1	D	14	GLN	6.1
1	A	38	ALA	6.0
1	D	13	GLN	5.6
1	C	39	SER	5.4
1	C	18	ILE	5.1
1	C	21	ARG	4.9
1	C	16	LEU	4.6
1	D	16	LEU	4.5
1	A	11	THR	4.4
1	A	44	ASP	4.3
1	A	14	GLN	4.2
1	C	9	VAL	4.1
1	B	10	SER	4.1
1	D	18	ILE	4.0
1	B	15	SER	3.9
1	C	125	ARG	3.8
1	C	71	ARG	3.8
1	D	125	ARG	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	9	VAL	3.7
1	C	15	SER	3.6
1	B	11	THR	3.3
1	D	122	ILE	3.2
1	A	37	LYS	3.2
1	D	127	ASN	3.1
1	C	43	SER	3.0
1	D	21	ARG	2.8
1	B	16	LEU	2.8
1	D	37	LYS	2.7
1	B	103	ILE	2.6
1	A	21	ARG	2.6
1	C	102	TYR	2.5
1	C	123	LEU	2.5
1	A	13	GLN	2.5
1	B	118	GLU	2.5
1	C	37	LYS	2.5
1	C	44	ASP	2.4
1	D	44	ASP	2.4
1	D	71	ARG	2.4
1	A	74	ALA	2.4
1	B	38	ALA	2.4
1	D	25	ASP	2.2
1	A	10	SER	2.2
1	A	128	GLU	2.2
1	D	70	GLY	2.2
1	B	45	ARG	2.1
1	A	15	SER	2.1
1	B	44	ASP	2.1

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 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.