



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 21, 2025 – 01:52 PM JST

PDB ID : 9LCW  
Title : Grimontia hollisae thermostable direct hemolysin K88A mutant in complex with 1-nt long 3'-overhang dsDNA  
Deposited on : 2025-01-05  
Resolution : 2.71 Å (reported)

**This wwPDB validation report is for manuscript review**

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

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)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

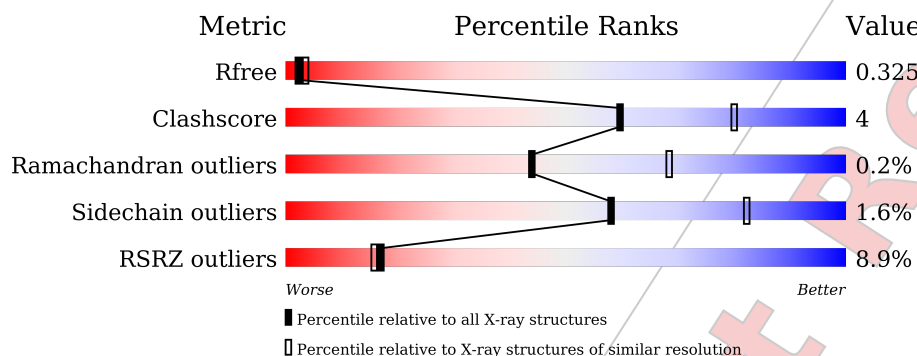
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.71 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))
$R_{free}$	164625	3333 (2.70-2.70)
Clashscore	180529	3684 (2.70-2.70)
Ramachandran outliers	177936	3633 (2.70-2.70)
Sidechain outliers	177891	3633 (2.70-2.70)
RSRZ outliers	164620	3333 (2.70-2.70)

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	165	<div> <div>7%</div> <div>85%</div> <div>7% • 7%</div> </div>
1	B	165	<div> <div>7%</div> <div>84%</div> <div>9% • 7%</div> </div>
1	C	165	<div> <div>9%</div> <div>78%</div> <div>15% • 7%</div> </div>
1	D	165	<div> <div>10%</div> <div>87%</div> <div>5% • 7%</div> </div>
2	E	11	<div> <div>18%</div> <div>100%</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	11	

For Manuscript Review

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5221 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 Thermostable direct hemolysin-related.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	154	Total	C	N	O	S	0	0	0
			1183	748	187	243	5			
1	B	154	Total	C	N	O	S	0	0	0
			1190	750	188	247	5			
1	C	154	Total	C	N	O	S	0	0	0
			1176	743	185	243	5			
1	D	154	Total	C	N	O	S	0	0	0
			1164	738	186	235	5			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	88	ALA	LYS	engineered mutation	UNP A0A377HK12
B	88	ALA	LYS	engineered mutation	UNP A0A377HK12
C	88	ALA	LYS	engineered mutation	UNP A0A377HK12
D	88	ALA	LYS	engineered mutation	UNP A0A377HK12

- Molecule 2 is a DNA chain called DNA (5'-D(\*CP\*TP\*CP\*TP\*AP\*TP\*AP\*GP\*AP\*GP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	11	Total	C	N	O	P	0	0	0
			223	108	42	63	10			
2	F	11	Total	C	N	O	P	0	0	0
			223	108	42	63	10			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	22	Total	O	0	0
			22	22		

*Continued on next page...*

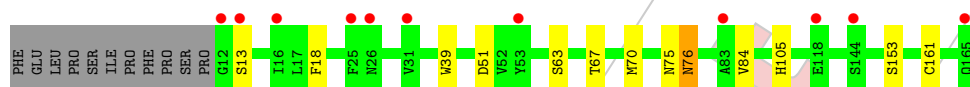
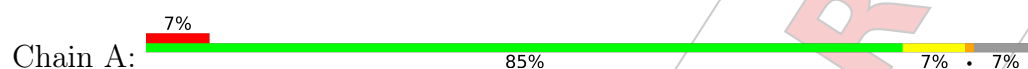
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	12	Total 12	O 12	0	0
3	C	11	Total 11	O 11	0	0
3	D	14	Total 14	O 14	0	0
3	E	2	Total 2	O 2	0	0
3	F	1	Total 1	O 1	0	0

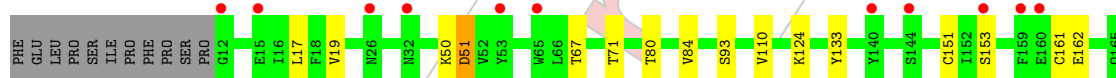
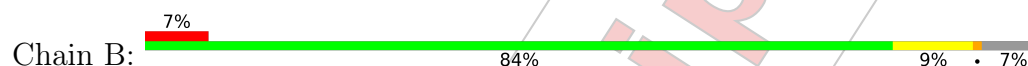
### 3 Residue-property plots [i](#)

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.

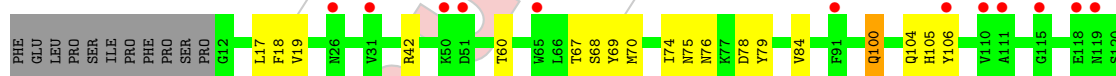
- Molecule 1: Thermostable direct hemolysin-related



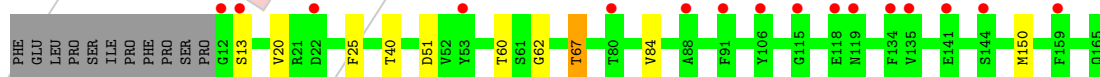
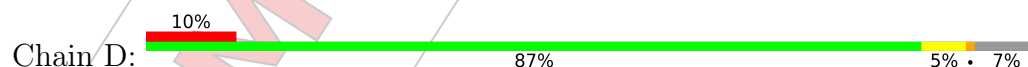
- Molecule 1: Thermostable direct hemolysin-related



- Molecule 1: Thermostable direct hemolysin-related

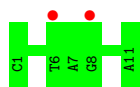


- Molecule 1: Thermostable direct hemolysin-related

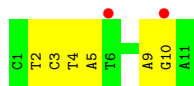
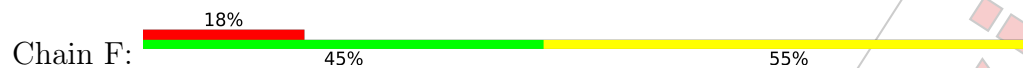


- Molecule 2: DNA (5'-D(\*CP\*TP\*CP\*TP\*AP\*TP\*AP\*GP\*AP\*GP\*A)-3')





- Molecule 2: DNA (5'-D(\*CP\*TP\*CP\*TP\*AP\*TP\*AP\*GP\*AP\*GP\*A)-3')



## 4 Data and refinement statistics [i](#)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.60Å 61.66Å 113.28Å 90.00° 106.06° 90.00°	Depositor
Resolution (Å)	29.78 – 2.71 29.78 – 2.71	Depositor EDS
% Data completeness (in resolution range)	73.0 (29.78-2.71) 73.0 (29.78-2.71)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.92 (at 2.72Å)	Xtriage
Refinement program	REFMAC 5.8.0415	Depositor
R, $R_{free}$	0.248 , 0.317 0.252 , 0.325	Depositor DCC
$R_{free}$ test set	1463 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	46.4	Xtriage
Anisotropy	0.046	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 32.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.53$ , $\langle L^2 \rangle = 0.36$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	5221	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.65% 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	0.38	0/1210	0.64	0/1647
1	B	0.36	0/1217	0.65	0/1657
1	C	0.38	0/1203	0.63	0/1640
1	D	0.36	0/1191	0.67	0/1624
2	E	0.54	0/250	0.93	0/384
2	F	0.66	0/250	0.97	0/384
All	All	0.40	0/5321	0.68	0/7336

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	42	ARG	Sidechain

### 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	1183	0	1073	6	0
1	B	1190	0	1078	9	0
1	C	1176	0	1055	15	0
1	D	1164	0	1050	4	0
2	E	223	0	126	0	0
2	F	223	0	126	3	0
3	A	22	0	0	0	0
3	B	12	0	0	0	0
3	C	11	0	0	0	0
3	D	14	0	0	0	0
3	E	2	0	0	0	0
3	F	1	0	0	0	0
All	All	5221	0	4508	37	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.

All (37) 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:60:THR:HG21	1:C:68:SER:OG	1.84	0.77
1:A:67:THR:HG22	1:A:84:VAL:HG22	1.77	0.65
1:B:151:CYS:HG	1:B:161:CYS:CB	2.14	0.59
1:C:78:ASP:O	1:C:100:GLN:NE2	2.25	0.58
1:C:153:SER:HB3	1:C:161:CYS:SG	2.45	0.56
1:A:13:SER:HA	1:A:63:SER:HA	1.87	0.56
1:A:39:TRP:CD1	1:A:105:HIS:HB3	2.42	0.55
1:B:50:LYS:O	1:B:51:ASP:HB2	2.07	0.54
1:C:18:PHE:HB3	1:C:70:MET:HE2	1.91	0.53
2:F:9:DA:H2'	2:F:10:DG:C8	2.44	0.53
1:B:151:CYS:SG	1:B:161:CYS:CB	2.99	0.51
1:D:67:THR:HG22	1:D:84:VAL:HG22	1.93	0.51
1:B:71:THR:HG23	1:B:80:THR:OG1	2.11	0.49
2:F:4:DT:H2'	2:F:5:DA:C8	2.48	0.48
1:C:151:CYS:HG	1:C:161:CYS:HG	0.93	0.47
1:B:17:LEU:CD2	1:B:19:VAL:HG23	2.45	0.47
1:D:13:SER:HB2	1:D:62:GLY:O	2.15	0.46
1:D:40:THR:HB	1:D:60:THR:HG21	1.96	0.46
1:C:67:THR:HG22	1:C:84:VAL:HG22	1.97	0.46
1:C:151:CYS:SG	1:C:161:CYS:HA	2.55	0.46
1:C:17:LEU:CD2	1:C:19:VAL:HG23	2.45	0.46
1:C:75:ASN:O	1:C:76:ASN:HB2	2.15	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:F:2:DT:H2''	2:F:3:DC:O4'	2.15	0.46
1:C:69:TYR:OH	1:C:105:HIS:CD2	2.68	0.46
1:C:122:PRO:HA	1:C:139:ALA:O	2.16	0.46
1:A:75:ASN:O	1:A:76:ASN:HB2	2.15	0.45
1:C:69:TYR:OH	1:C:105:HIS:HD2	1.99	0.45
1:B:67:THR:HB	1:B:110:VAL:HG11	1.99	0.44
1:B:133:TYR:HA	1:B:153:SER:O	2.18	0.43
1:C:121:ILE:HG13	1:C:122:PRO:HA	2.00	0.43
1:A:153:SER:HB3	1:A:161:CYS:SG	2.58	0.42
1:D:20:VAL:HA	1:D:150:MET:O	2.18	0.42
1:C:74:ILE:HD12	1:C:79:TYR:CE2	2.56	0.41
1:B:124:LYS:NZ	1:B:162:GLU:OE1	2.50	0.41
1:C:104:GLN:HB2	1:C:106:TYR:CE1	2.56	0.41
1:A:18:PHE:HB3	1:A:70:MET:HE2	2.02	0.40
1:B:67:THR:HG22	1:B:84:VAL:HG22	2.04	0.40

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	152/165 (92%)	142 (93%)	10 (7%)	0	100	100
1	B	152/165 (92%)	140 (92%)	11 (7%)	1 (1%)	19	42
1	C	152/165 (92%)	141 (93%)	11 (7%)	0	100	100
1	D	152/165 (92%)	140 (92%)	12 (8%)	0	100	100
All	All	608/660 (92%)	563 (93%)	44 (7%)	1 (0%)	44	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	51	ASP

### 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	128/150 (85%)	126 (98%)	2 (2%)	58	82
1	B	130/150 (87%)	129 (99%)	1 (1%)	79	91
1	C	126/150 (84%)	124 (98%)	2 (2%)	58	82
1	D	123/150 (82%)	120 (98%)	3 (2%)	44	73
All	All	507/600 (84%)	499 (98%)	8 (2%)	58	82

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	51	ASP
1	A	76	ASN
1	B	93	SER
1	C	100	GLN
1	C	150	MET
1	D	25	PHE
1	D	51	ASP
1	D	67	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 [i](#)

There are no oligosaccharides 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.

For Manuscript Review

## 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	154/165 (93%)	0.81	11 (7%) 23 21	25, 43, 69, 85	0
1	B	154/165 (93%)	0.78	11 (7%) 23 21	27, 42, 72, 85	0
1	C	154/165 (93%)	0.88	15 (9%) 15 14	26, 45, 73, 95	0
1	D	154/165 (93%)	0.88	16 (10%) 13 12	27, 46, 74, 89	0
2	E	11/11 (100%)	1.24	2 (18%) 4 4	64, 82, 98, 110	0
2	F	11/11 (100%)	1.64	2 (18%) 4 4	28, 89, 103, 114	0
All	All	638/682 (93%)	0.86	57 (8%) 17 16	25, 45, 77, 114	0

All (57) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	144	SER	4.9
1	C	118	GLU	4.5
1	B	144	SER	3.7
1	D	22	ASP	3.6
1	C	50	LYS	3.5
1	D	159	PHE	3.4
2	F	10	DG	3.3
1	D	12	GLY	3.3
1	A	118	GLU	3.1
1	C	26	ASN	3.1
1	C	110	VAL	3.0
1	D	118	GLU	3.0
1	A	25	PHE	3.0
1	A	26	ASN	2.9
1	B	15	GLU	2.9
2	F	6	DT	2.8
1	C	160	GLU	2.7
1	A	12	GLY	2.7
1	B	12	GLY	2.7

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	134	PHE	2.7
1	C	151	CYS	2.6
1	B	160	GLU	2.6
1	B	159	PHE	2.6
1	A	144	SER	2.5
1	A	165	GLN	2.5
1	C	115	GLY	2.5
1	A	16	ILE	2.4
1	D	144	SER	2.4
1	A	53	TYR	2.4
2	E	8	DG	2.4
1	C	65	TRP	2.4
1	C	91	PHE	2.4
1	B	53	TYR	2.4
1	B	32	ASN	2.4
1	C	51	ASP	2.3
1	B	153	SER	2.3
1	D	141	GLU	2.3
1	A	83	ALA	2.3
1	B	26	ASN	2.3
1	C	106	TYR	2.2
1	A	13	SER	2.2
1	D	13	SER	2.2
1	D	119	ASN	2.2
1	D	91	PHE	2.2
1	A	31	VAL	2.2
1	D	135	VAL	2.2
1	D	88	ALA	2.1
1	D	80	THR	2.1
1	C	31	VAL	2.1
1	D	106	TYR	2.1
1	C	111	ALA	2.1
1	C	119	ASN	2.1
1	D	115	GLY	2.1
1	B	140	TYR	2.0
1	B	65	TRP	2.0
2	E	6	DT	2.0
1	D	53	TYR	2.0

## 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 [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.

For Manuscript Review