



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 21, 2024 – 06:25 PM EDT

PDB ID : 3C5M  
Title : Crystal structure of oligogalacturonate lyase (VPA0088) from *Vibrio parahaemolyticus*. Northeast Structural Genomics Consortium Target VpR199  
Authors : Forouhar, F.; Abashidze, M.; Seetharaman, J.; Janjua, H.; Mao, L.; Xiao, R.; Owens, L.A.; Wang, D.; Baran, M.C.; Acton, T.B.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)  
Deposited on : 2008-01-31  
Resolution : 2.60 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
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

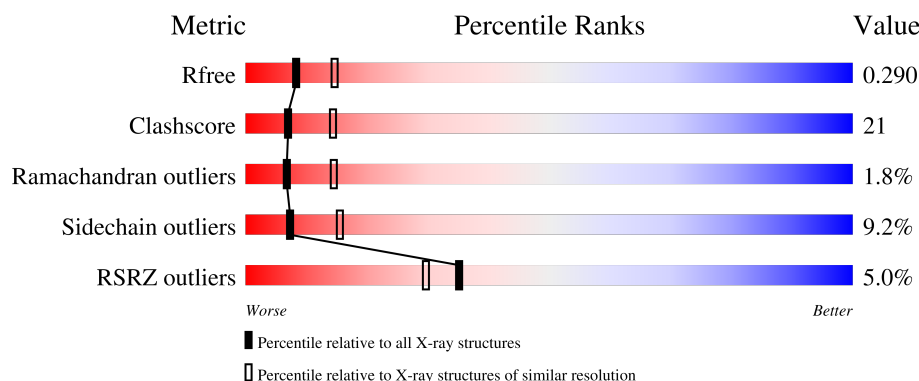
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	3775 (2.60-2.60)
Clashscore	180529	4181 (2.60-2.60)
Ramachandran outliers	177936	4129 (2.60-2.60)
Sidechain outliers	177891	4129 (2.60-2.60)
RSRZ outliers	164620	3775 (2.60-2.60)

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	396	
1	B	396	
1	C	396	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9292 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 Oligogalacturonate lyase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	376	Total	C	N	O	S	Se	0	0	0
			3022	1919	494	593	9	7			
1	B	378	Total	C	N	O	S	Se	0	0	0
			3036	1929	496	595	9	7			
1	C	376	Total	C	N	O	S	Se	0	0	0
			3022	1919	494	593	9	7			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	31	ILE	VAL	SEE REMARK 999	UNP Q87K10
A	389	LEU	-	expression tag	UNP Q87K10
A	390	GLU	-	expression tag	UNP Q87K10
A	391	HIS	-	expression tag	UNP Q87K10
A	392	HIS	-	expression tag	UNP Q87K10
A	393	HIS	-	expression tag	UNP Q87K10
A	394	HIS	-	expression tag	UNP Q87K10
A	395	HIS	-	expression tag	UNP Q87K10
A	396	HIS	-	expression tag	UNP Q87K10
B	31	ILE	VAL	SEE REMARK 999	UNP Q87K10
B	389	LEU	-	expression tag	UNP Q87K10
B	390	GLU	-	expression tag	UNP Q87K10
B	391	HIS	-	expression tag	UNP Q87K10
B	392	HIS	-	expression tag	UNP Q87K10
B	393	HIS	-	expression tag	UNP Q87K10
B	394	HIS	-	expression tag	UNP Q87K10
B	395	HIS	-	expression tag	UNP Q87K10
B	396	HIS	-	expression tag	UNP Q87K10
C	31	ILE	VAL	SEE REMARK 999	UNP Q87K10
C	389	LEU	-	expression tag	UNP Q87K10
C	390	GLU	-	expression tag	UNP Q87K10
C	391	HIS	-	expression tag	UNP Q87K10
C	392	HIS	-	expression tag	UNP Q87K10

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	393	HIS	-	expression tag	UNP Q87K10
C	394	HIS	-	expression tag	UNP Q87K10
C	395	HIS	-	expression tag	UNP Q87K10
C	396	HIS	-	expression tag	UNP Q87K10

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Mn 1 1	0	0
2	B	1	Total Mn 1 1	0	0
2	C	1	Total Mn 1 1	0	0

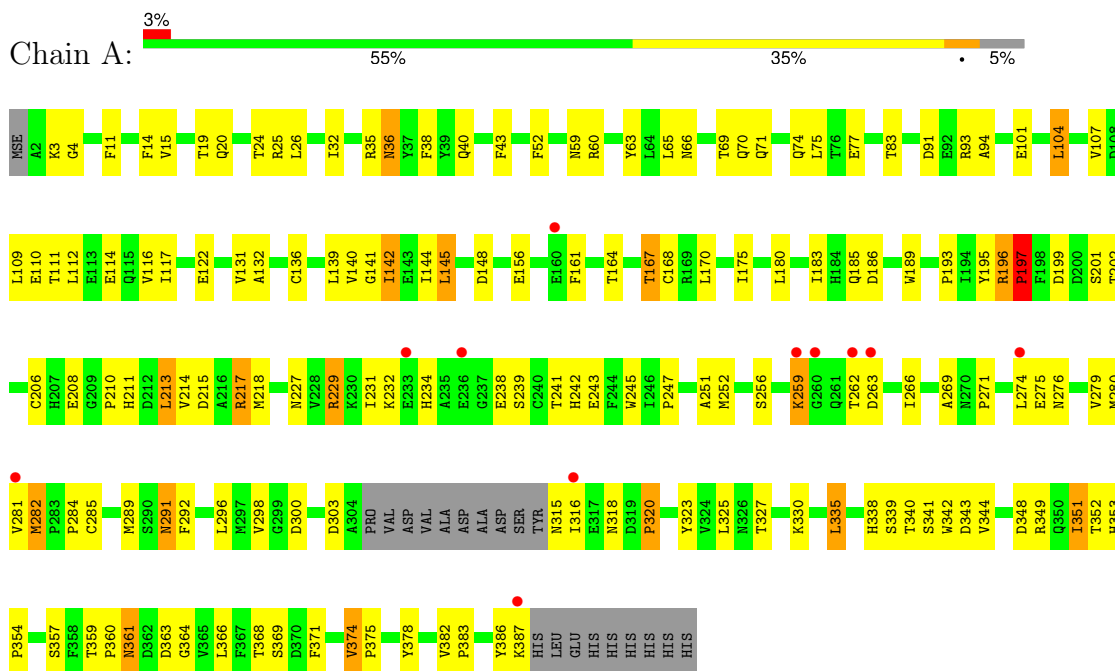
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	87	Total O 87 87	0	0
3	B	69	Total O 69 69	0	0
3	C	53	Total O 53 53	0	0

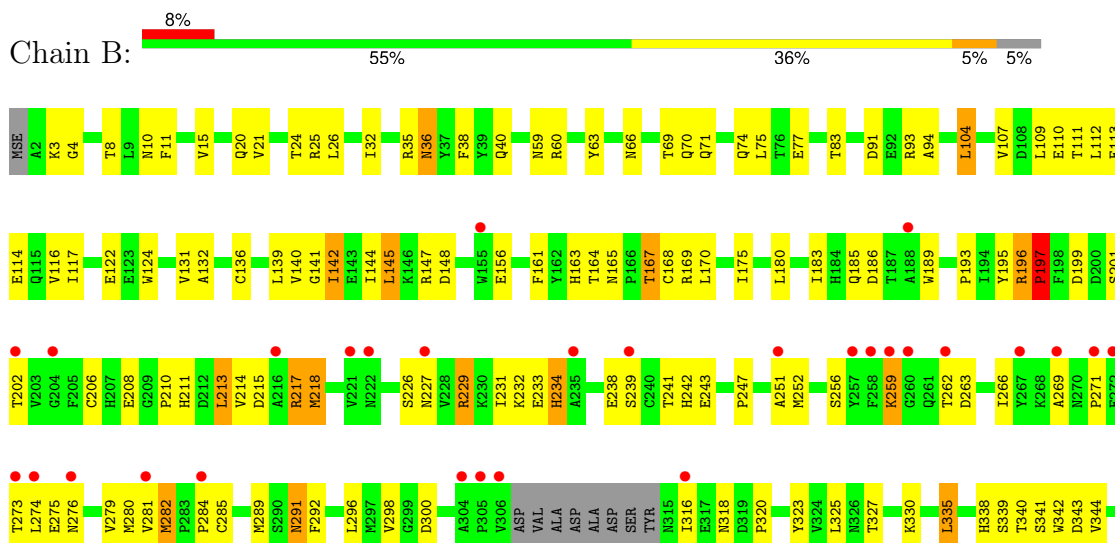
### 3 Residue-property plots

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: Oligogalacturonate lyase

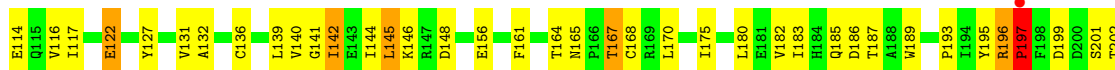


#### • Molecule 1: Oligogalacturonate lyase





● Molecule 1: Oligogalacturonate lyase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	115.23Å 115.23Å 209.95Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.96 – 2.60 19.96 – 2.60	Depositor EDS
% Data completeness (in resolution range)	89.8 (19.96-2.60) 96.7 (19.96-2.60)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.24 (at 2.61Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.234 , 0.277 0.249 , 0.290	Depositor DCC
$R_{free}$ test set	4856 reflections (9.73%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.0	Xtriage
Anisotropy	0.542	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 51.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.024 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	9292	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 39.21 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.2840e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MN

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	0.42	0/3093	0.59	0/4191
1	B	0.43	0/3108	0.60	0/4213
1	C	0.42	0/3093	0.59	0/4191
All	All	0.43	0/9294	0.59	0/12595

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

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	3022	0	2853	124	0
1	B	3036	0	2869	125	0
1	C	3022	0	2853	120	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	87	0	0	5	0
3	B	69	0	0	4	0
3	C	53	0	0	2	0
All	All	9292	0	8575	365	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 21.

The worst 5 of 365 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:38:PHE:HD1	1:A:289:MSE:HE3	1.38	0.89
1:A:374:VAL:HG22	1:A:375:PRO:HD2	1.56	0.87
1:C:38:PHE:HD1	1:C:289:MSE:HE3	1.38	0.86
1:C:213:LEU:H	1:C:213:LEU:HD23	1.42	0.85
1:B:107:VAL:HG22	1:B:114:GLU:HG2	1.59	0.84

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	372/396 (94%)	336 (90%)	30 (8%)	6 (2%)	8	17
1	B	374/396 (94%)	334 (89%)	33 (9%)	7 (2%)	6	13
1	C	372/396 (94%)	332 (89%)	33 (9%)	7 (2%)	6	13
All	All	1118/1188 (94%)	1002 (90%)	96 (9%)	20 (2%)	7	14

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	197	PRO
1	A	316	ILE
1	A	342	TRP
1	B	197	PRO
1	B	316	ILE

### 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	334/344 (97%)	303 (91%)	31 (9%)	7	15
1	B	336/344 (98%)	307 (91%)	29 (9%)	8	18
1	C	334/344 (97%)	302 (90%)	32 (10%)	7	14
All	All	1004/1032 (97%)	912 (91%)	92 (9%)	7	15

5 of 92 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	351	ILE
1	C	145	LEU
1	B	374	VAL
1	C	104	LEU
1	C	196	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 47 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	333	GLN
1	C	70	GLN
1	B	361	ASN
1	C	59	ASN
1	C	115	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 [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	369/396 (93%)	-0.05	11 (2%) 52 46	20, 44, 81, 98	0
1	B	371/396 (93%)	0.38	31 (8%) 18 15	17, 53, 86, 105	0
1	C	369/396 (93%)	0.02	14 (3%) 44 38	19, 46, 85, 100	0
All	All	1109/1188 (93%)	0.12	56 (5%) 35 30	17, 47, 85, 105	0

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	387	LYS	5.8
1	B	259	LYS	4.2
1	B	387	LYS	4.2
1	A	387	LYS	4.2
1	B	260	GLY	4.1

### 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 [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	MN	A	401	1/1	0.99	0.05	30,30,30,30	0
2	MN	B	401	1/1	0.99	0.03	20,20,20,20	0
2	MN	C	401	1/1	0.99	0.07	30,30,30,30	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.