



# Full wwPDB X-ray Structure Validation Report ⓘ

Apr 18, 2022 – 04:02 PM EDT

PDB ID : 6D7R  
Title : Crystal structure of an inactivate variant of the quorum-sensing master regulator HapR from protease-deficient non-O1, non-O139 *Vibrio cholerae* strain V2  
Authors : Cruite, J.T.; Succo, P.; Raychaudhuri, S.; Kull, F.J.  
Deposited on : 2018-04-25  
Resolution : 2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 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	:	2.27
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)	:	2.27

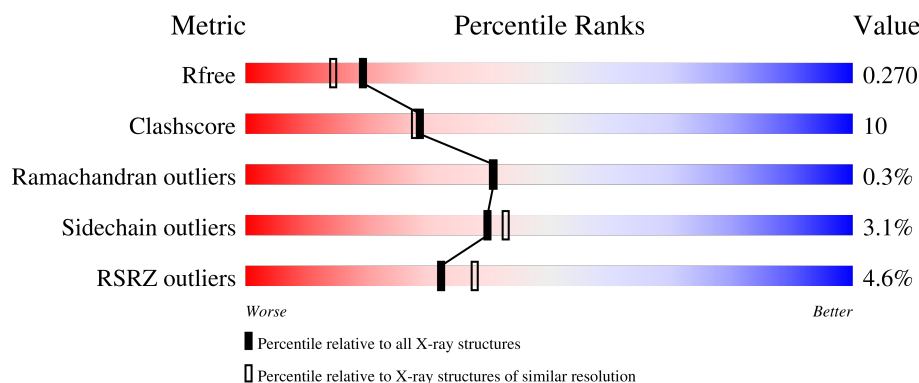
# 1 Overall quality at a glance

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))
$R_{free}$	130704	5197 (2.10-2.10)
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  $\geq 3$ , 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	223	<div> <div>5%</div> <div> <div></div> <div>69%</div> <div>18%</div> <div>•</div> <div>12%</div> </div> </div>
1	B	223	<div> <div>3%</div> <div> <div></div> <div>68%</div> <div>20%</div> <div></div> <div>12%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3323 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 Hemagglutinin/protease regulatory protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	197	Total	C	N	O	S	0	0	0
			1622	1037	279	294	12			
1	B	197	Total	C	N	O	S	0	0	0
			1622	1037	279	294	12			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP B2CKP3
A	-18	GLY	-	expression tag	UNP B2CKP3
A	-17	SER	-	expression tag	UNP B2CKP3
A	-16	SER	-	expression tag	UNP B2CKP3
A	-15	HIS	-	expression tag	UNP B2CKP3
A	-14	HIS	-	expression tag	UNP B2CKP3
A	-13	HIS	-	expression tag	UNP B2CKP3
A	-12	HIS	-	expression tag	UNP B2CKP3
A	-11	HIS	-	expression tag	UNP B2CKP3
A	-10	HIS	-	expression tag	UNP B2CKP3
A	-9	SER	-	expression tag	UNP B2CKP3
A	-8	SER	-	expression tag	UNP B2CKP3
A	-7	GLY	-	expression tag	UNP B2CKP3
A	-6	LEU	-	expression tag	UNP B2CKP3
A	-5	VAL	-	expression tag	UNP B2CKP3
A	-4	PRO	-	expression tag	UNP B2CKP3
A	-3	ARG	-	expression tag	UNP B2CKP3
A	-2	GLY	-	expression tag	UNP B2CKP3
A	-1	SER	-	expression tag	UNP B2CKP3
A	0	HIS	-	expression tag	UNP B2CKP3
A	39	ASP	GLY	conflict	UNP B2CKP3
B	-19	MET	-	initiating methionine	UNP B2CKP3
B	-18	GLY	-	expression tag	UNP B2CKP3
B	-17	SER	-	expression tag	UNP B2CKP3
B	-16	SER	-	expression tag	UNP B2CKP3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-15	HIS	-	expression tag	UNP B2CKP3
B	-14	HIS	-	expression tag	UNP B2CKP3
B	-13	HIS	-	expression tag	UNP B2CKP3
B	-12	HIS	-	expression tag	UNP B2CKP3
B	-11	HIS	-	expression tag	UNP B2CKP3
B	-10	HIS	-	expression tag	UNP B2CKP3
B	-9	SER	-	expression tag	UNP B2CKP3
B	-8	SER	-	expression tag	UNP B2CKP3
B	-7	GLY	-	expression tag	UNP B2CKP3
B	-6	LEU	-	expression tag	UNP B2CKP3
B	-5	VAL	-	expression tag	UNP B2CKP3
B	-4	PRO	-	expression tag	UNP B2CKP3
B	-3	ARG	-	expression tag	UNP B2CKP3
B	-2	GLY	-	expression tag	UNP B2CKP3
B	-1	SER	-	expression tag	UNP B2CKP3
B	0	HIS	-	expression tag	UNP B2CKP3
B	39	ASP	GLY	conflict	UNP B2CKP3

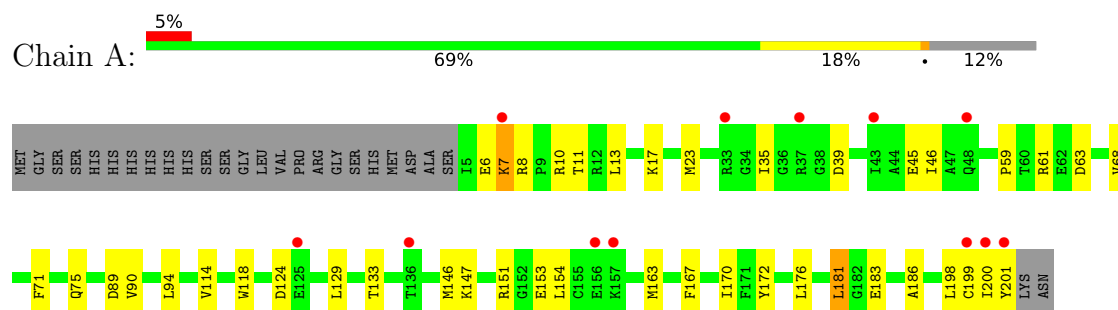
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	37	Total O 37 37	0	0
2	B	42	Total O 42 42	0	0

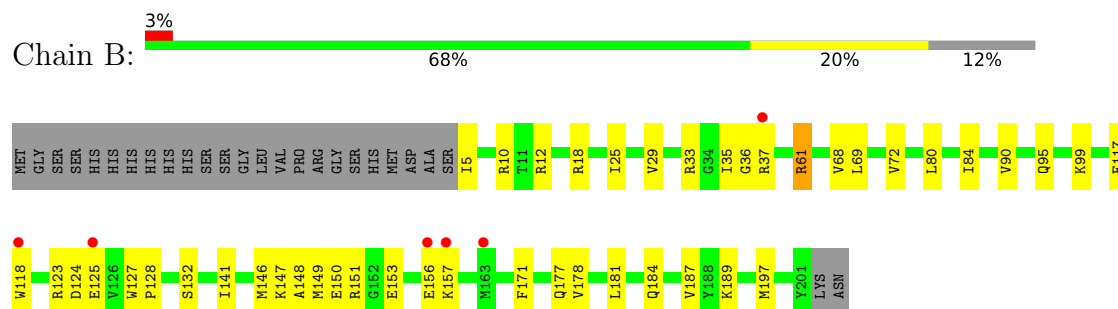
### 3 Residue-property plots [i](#)

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: Hemagglutinin/protease regulatory protein



- Molecule 1: Hemagglutinin/protease regulatory protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	43.47Å 85.71Å 112.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.43 – 2.10 19.43 – 2.10	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.43-2.10) 91.6 (19.43-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.00 (at 2.09Å)	Xtriage
Refinement program	PHENIX (1.12_2829)	Depositor
R, $R_{free}$	0.203 , 0.268 0.204 , 0.270	Depositor DCC
$R_{free}$ test set	2000 reflections (7.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.6	Xtriage
Anisotropy	0.458	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 60.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3323	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 19.75% 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

### 5.1 Standard geometry

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	0.40	0/1655	0.53	0/2236
1	B	0.41	0/1655	0.53	0/2236
All	All	0.40	0/3310	0.53	0/4472

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	1622	0	1616	33	0
1	B	1622	0	1616	31	0
2	A	37	0	0	5	0
2	B	42	0	0	4	0
All	All	3323	0	3232	63	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 10.

All (63) 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:B:90:VAL:HG21	1:B:148:ALA:HB2	1.43	1.00

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:36:GLY:O	1:B:61:ARG:NH2	2.17	0.76
1:A:183:GLU:OE1	2:A:301:HOH:O	2.08	0.70
1:A:63:ASP:OD2	2:A:302:HOH:O	2.09	0.69
1:A:35:ILE:HD12	1:A:114:VAL:HG13	1.79	0.65
1:B:146:MET:O	1:B:150:GLU:HG3	1.96	0.65
1:B:90:VAL:CG2	1:B:148:ALA:HB2	2.24	0.64
1:B:124:ASP:OD2	2:B:301:HOH:O	2.15	0.60
1:B:153:GLU:OE1	2:B:302:HOH:O	2.15	0.59
1:B:123:ARG:HE	1:B:124:ASP:N	2.01	0.58
1:A:118:TRP:CH2	1:A:133:THR:HG21	2.39	0.57
1:A:181:LEU:HD13	1:A:183:GLU:HB3	1.88	0.56
1:B:123:ARG:HE	1:B:124:ASP:H	1.53	0.56
1:A:7:LYS:HG2	1:A:45:GLU:HG2	1.89	0.55
1:B:69:LEU:HB3	1:B:118:TRP:CH2	2.41	0.55
1:B:90:VAL:HG21	1:B:148:ALA:CB	2.29	0.55
1:A:154:LEU:HD22	1:A:200:ILE:HB	1.89	0.53
1:A:181:LEU:CD1	1:A:183:GLU:H	2.21	0.53
1:B:69:LEU:HB3	1:B:118:TRP:CZ3	2.44	0.53
1:A:163:MET:HG3	1:B:197:MET:HE2	1.91	0.53
1:A:181:LEU:HD13	1:A:183:GLU:CB	2.39	0.53
1:A:147:LYS:HE3	1:A:151:ARG:HH21	1.73	0.52
1:A:153:GLU:O	1:A:201:TYR:HA	2.09	0.52
1:A:186:ALA:N	2:A:301:HOH:O	2.30	0.52
1:A:7:LYS:HA	1:A:46:ILE:HA	1.92	0.51
1:B:178:VAL:HB	1:B:187:VAL:HG11	1.92	0.51
1:A:181:LEU:HD13	1:A:183:GLU:H	1.76	0.51
1:A:170:ILE:HD13	1:A:198:LEU:HD11	1.92	0.50
1:B:80:LEU:HD22	1:B:84:ILE:HD11	1.92	0.50
1:A:172:TYR:CZ	1:A:176:LEU:HD21	2.47	0.50
1:B:95:GLN:O	1:B:99:LYS:HG2	2.12	0.49
1:B:123:ARG:HG3	1:B:125:GLU:H	1.77	0.48
1:A:89:ASP:HB2	1:A:153:GLU:OE2	2.14	0.48
1:A:59:PRO:HB2	2:A:302:HOH:O	2.15	0.47
1:B:189:LYS:N	1:B:189:LYS:HD3	2.29	0.47
1:A:6:GLU:O	1:A:46:ILE:HG23	2.16	0.46
1:A:23:MET:SD	1:A:68:VAL:HA	2.55	0.46
1:B:5:ILE:HD12	1:B:33:ARG:NH2	2.30	0.46
1:A:90:VAL:HG11	1:A:154:LEU:HD21	1.97	0.46
1:B:127:TRP:CG	1:B:128:PRO:HD3	2.50	0.46
1:B:149:MET:SD	1:B:156:GLU:HB3	2.57	0.45
1:B:68:VAL:O	1:B:72:VAL:HG13	2.17	0.45

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:200:ILE:HG13	1:A:201:TYR:CZ	2.51	0.45
1:B:69:LEU:HD13	1:B:118:TRP:CE2	2.52	0.44
1:A:11:THR:HG23	1:A:13:LEU:HG	1.99	0.43
1:B:5:ILE:N	2:B:309:HOH:O	2.50	0.43
1:B:10:ARG:HH22	1:B:18:ARG:NH2	2.16	0.43
1:A:124:ASP:N	2:A:307:HOH:O	2.38	0.43
1:A:8:ARG:N	1:A:46:ILE:O	2.51	0.43
1:A:147:LYS:HG2	1:A:151:ARG:NH2	2.33	0.43
1:B:25:ILE:O	1:B:29:VAL:HG23	2.18	0.43
1:B:123:ARG:HH21	1:B:124:ASP:HB2	1.84	0.43
1:A:94:LEU:HD22	1:A:167:PHE:CE1	2.55	0.42
1:A:167:PHE:HA	1:A:170:ILE:HD12	2.02	0.42
1:B:35:ILE:HG13	1:B:117:GLU:OE1	2.19	0.42
1:A:129:LEU:HD23	1:A:129:LEU:HA	1.86	0.42
1:B:141:ILE:HD13	1:B:141:ILE:HA	1.87	0.42
1:B:177:GLN:O	1:B:181:LEU:HB2	2.19	0.42
1:A:71:PHE:O	1:A:75:GLN:HG2	2.19	0.41
1:B:147:LYS:HE3	1:B:151:ARG:NH2	2.35	0.41
1:B:184:GLN:NE2	2:B:307:HOH:O	2.38	0.41
1:A:39:ASP:OD1	1:A:61:ARG:NH1	2.54	0.40
1:A:183:GLU:HG3	1:A:186:ALA:HB3	2.03	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	195/223 (87%)	186 (95%)	8 (4%)	1 (0%)	29	26
1	B	195/223 (87%)	192 (98%)	3 (2%)	0	100	100
All	All	390/446 (87%)	378 (97%)	11 (3%)	1 (0%)	41	41

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	10	ARG

### 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	179/201 (89%)	174 (97%)	5 (3%)	43	47
1	B	179/201 (89%)	173 (97%)	6 (3%)	37	39
All	All	358/402 (89%)	347 (97%)	11 (3%)	40	43

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	LYS
1	A	17	LYS
1	A	146	MET
1	A	181	LEU
1	A	199	CYS
1	B	12	ARG
1	B	37	ARG
1	B	61	ARG
1	B	132	SER
1	B	157	LYS
1	B	171	PHE

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 [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	197/223 (88%)	0.41	12 (6%) 21 26	22, 46, 88, 115	0
1	B	197/223 (88%)	0.12	6 (3%) 50 56	26, 44, 68, 94	0
All	All	394/446 (88%)	0.26	18 (4%) 32 38	22, 44, 80, 115	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	201	TYR	5.8
1	A	157	LYS	4.5
1	A	125	GLU	4.3
1	A	156	GLU	3.9
1	A	200	ILE	3.8
1	B	37	ARG	3.6
1	A	136	THR	3.5
1	A	33	ARG	3.3
1	A	37	ARG	3.1
1	A	48	GLN	3.1
1	B	156	GLU	3.0
1	B	125	GLU	3.0
1	B	157	LYS	2.5
1	B	163	MET	2.3
1	A	7	LYS	2.3
1	A	199	CYS	2.1
1	A	43	ILE	2.1
1	B	118	TRP	2.0

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.