W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 14, 2020 – 07:27 pm BST

PDB ID	:	3B4S
Title	:	Crystal structure of a LuxT domain from Vibrio parahaemolyticus RIMD 2210633
Authors	:	Tan, K.; Zhou, M.; Gu, M.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)
Deposited on	:	2007-10-24
Resolution	:	3.10 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
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

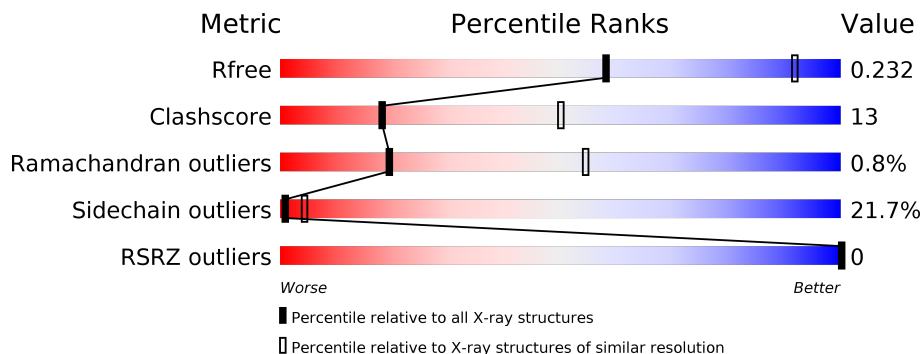The following versions of software and data (see references (i)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

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 3.10 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
Rfree		0.232
Clashscore		13
Ramachandran outliers		0.8%
Sidechain outliers		21.7%
RSRZ outliers		0

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))
R_{free}	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 >=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	94	66% ... 28% ..
1	B	94	68% ... 22% ... 6% .
1	C	94	52% ... 39% ... 6% .
1	D	94	59% ... 31% ... 6% .
1	E	94	66% ... 20% ... 11% .
1	F	94	69% ... 19% ... 6% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	94	64% 23% 10%
1	H	94	49% 33% 13%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 5906 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 Protein LuxT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	92	Total 745	C 481	N 124	O 137	Se 3	0	0	0
1	B	91	Total 740	C 478	N 123	O 136	Se 3	0	0	0
1	C	92	Total 745	C 481	N 124	O 137	Se 3	0	0	0
1	D	90	Total 732	C 474	N 122	O 133	Se 3	0	0	0
1	E	91	Total 740	C 478	N 123	O 136	Se 3	0	0	0
1	F	90	Total 732	C 474	N 122	O 133	Se 3	0	0	0
1	G	91	Total 740	C 478	N 123	O 136	Se 3	0	0	0
1	H	90	Total 732	C 474	N 122	O 133	Se 3	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	60	SER	-	EXPRESSION TAG	UNP Q87J33
A	61	ASN	-	EXPRESSION TAG	UNP Q87J33
A	62	ALA	-	EXPRESSION TAG	UNP Q87J33
B	60	SER	-	EXPRESSION TAG	UNP Q87J33
B	61	ASN	-	EXPRESSION TAG	UNP Q87J33
B	62	ALA	-	EXPRESSION TAG	UNP Q87J33
C	60	SER	-	EXPRESSION TAG	UNP Q87J33
C	61	ASN	-	EXPRESSION TAG	UNP Q87J33
C	62	ALA	-	EXPRESSION TAG	UNP Q87J33
D	60	SER	-	EXPRESSION TAG	UNP Q87J33
D	61	ASN	-	EXPRESSION TAG	UNP Q87J33
D	62	ALA	-	EXPRESSION TAG	UNP Q87J33
E	60	SER	-	EXPRESSION TAG	UNP Q87J33

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	61	ASN	-	EXPRESSION TAG	UNP Q87J33
E	62	ALA	-	EXPRESSION TAG	UNP Q87J33
F	60	SER	-	EXPRESSION TAG	UNP Q87J33
F	61	ASN	-	EXPRESSION TAG	UNP Q87J33
F	62	ALA	-	EXPRESSION TAG	UNP Q87J33
G	60	SER	-	EXPRESSION TAG	UNP Q87J33
G	61	ASN	-	EXPRESSION TAG	UNP Q87J33
G	62	ALA	-	EXPRESSION TAG	UNP Q87J33
H	60	SER	-	EXPRESSION TAG	UNP Q87J33
H	61	ASN	-	EXPRESSION TAG	UNP Q87J33
H	62	ALA	-	EXPRESSION TAG	UNP Q87J33

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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.

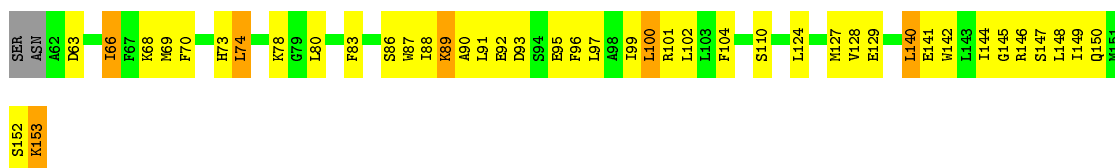- Molecule 1: Protein LuxT

Chain A:

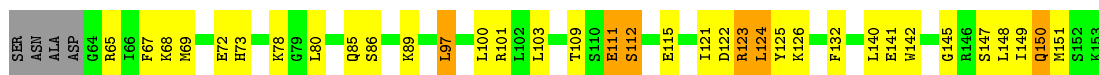
- Molecule 1: Protein LuxT

Chain B:

- Molecule 1: Protein LuxT

Chain C:

- Molecule 1: Protein LuxT

Chain D:

- Molecule 1: Protein LuxT

Chain E:

- Molecule 1: Protein LuxT

Chain F: 69% 19% 6%

SER ASN ALA ASP G64 F67 K68 M69 E72 E77 L80 K89 L97 L100 L103 V108 T109 S112 N119 G120 I121 D122 R123 L124 Y125 K126 M127 V128 E129 S130 Q131 S134 I144 I149 K153

● Molecule 1: Protein LuxT

Chain G: 64% 23% 10%

SER ASN ALA D63 G64 R65 I66 F67 K68 M69 F70 I71 E72 H73 L74 E75 L80 F83 S84 Q85 D93 S94 E95 F96 L97 A98 I99 L102 L103 F104 H105 R123 L124 Y125 V128 K138 E139 L140 E141 R146 I149 Q150 K153

● Molecule 1: Protein LuxT

Chain H: 49% 33% 13%

SER ASN ALA ASP G64 R65 K68 M69 F70 I71 E72 H73 L74 E75 F76 E77 K78 G79 L80 D81 Q85 S86 E95 F96 L97 A98 I99 L100 R101 L102 L103 H106 I107 V108 T109 S110 E111 S112 A113 H114 E115 N119 R123 L124 Y125 K126 M127 V128 E129 G136 D137 E138 E139

L140 I144 G145 R146 S147 L148 I149 Q150 K153
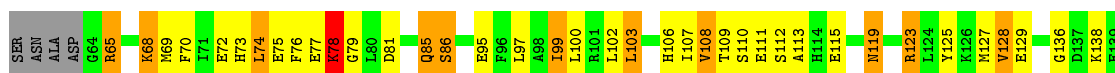
4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	147.44Å 147.44Å 382.49Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.22 – 3.10 48.22 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.22-3.10) 99.8 (48.22-3.10)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	3.33 (at 3.12Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.190 , 0.232 0.188 , 0.232	Depositor DCC
R_{free} test set	1490 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	71.4	Xtriage
Anisotropy	0.038	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 27.3	EDS
L-test for twinning[2]	$<|L|> = 0.48$, $<L^2> = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5906	wwPDB-VP
Average B, all atoms (Å^2)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.13% 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	0.88	1/758 (0.1%)	0.94	1/1009 (0.1%)
1	B	0.89	0/753	0.92	0/1002
1	C	0.86	0/758	0.86	0/1009
1	D	0.88	0/745	0.84	0/991
1	E	0.86	0/753	0.88	1/1002 (0.1%)
1	F	0.86	1/745 (0.1%)	0.88	0/991
1	G	0.83	0/753	0.91	1/1002 (0.1%)
1	H	0.83	0/745	0.86	0/991
All	All	0.86	2/6010 (0.0%)	0.89	3/7997 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	123	ARG	CG-CD	5.58	1.65	1.51
1	F	144	ILE	CA-CB	5.22	1.66	1.54

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	123	ARG	NE-CZ-NH1	6.79	123.69	120.30
1	A	102	LEU	CB-CG-CD1	-5.91	100.95	111.00
1	E	91	LEU	CA-CB-CG	5.74	128.51	115.30

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	745	0	730	18	0
1	B	740	0	725	13	0
1	C	745	0	730	29	0
1	D	732	0	721	28	0
1	E	740	0	725	21	0
1	F	732	0	721	21	0
1	G	740	0	725	22	0
1	H	732	0	721	45	0
All	All	5906	0	5798	157	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 13.

The worst 5 of 157 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:G:153:LYS:HA	1:G:153:LYS:HE3	1.29	1.10
1:A:119:ASN:ND2	1:A:122:ASP:H	1.68	0.91
1:C:83:PHE:O	1:C:86:SER:HB3	1.71	0.90
1:F:69:MSE:HA	1:F:72:GLU:HG2	1.54	0.89
1:B:119:ASN:HD22	1:B:119:ASN:C	1.78	0.85

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	90/94 (96%)	84 (93%)	5 (6%)	1 (1%)	14	46
1	B	89/94 (95%)	84 (94%)	5 (6%)	0	100	100
1	C	90/94 (96%)	83 (92%)	6 (7%)	1 (1%)	14	46
1	D	88/94 (94%)	79 (90%)	7 (8%)	2 (2%)	6	28

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	89/94 (95%)	85 (96%)	4 (4%)	0	100	100
1	F	88/94 (94%)	78 (89%)	10 (11%)	0	100	100
1	G	89/94 (95%)	84 (94%)	5 (6%)	0	100	100
1	H	88/94 (94%)	81 (92%)	5 (6%)	2 (2%)	6	28
All	All	711/752 (94%)	658 (92%)	47 (7%)	6 (1%)	19	54

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	78	LYS
1	D	112	SER
1	H	112	SER
1	A	77	GLU
1	D	115	GLU

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	78/77 (101%)	63 (81%)	15 (19%)	1	6
1	B	78/77 (101%)	57 (73%)	21 (27%)	0	1
1	C	78/77 (101%)	61 (78%)	17 (22%)	1	4
1	D	77/77 (100%)	60 (78%)	17 (22%)	1	4
1	E	78/77 (101%)	61 (78%)	17 (22%)	1	4
1	F	77/77 (100%)	62 (80%)	15 (20%)	1	6
1	G	78/77 (101%)	63 (81%)	15 (19%)	1	6
1	H	77/77 (100%)	59 (77%)	18 (23%)	1	3
All	All	621/616 (101%)	486 (78%)	135 (22%)	1	4

5 of 135 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	111	GLU
1	E	84	SER
1	H	97	LEU
1	D	112	SER
1	D	126	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	85	GLN
1	F	85	GLN
1	H	73	HIS
1	D	73	HIS
1	G	85	GLN

5.3.3 RNA (i)

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 carbohydrates 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, 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	89/94 (94%)	-0.30	0	100	100	46, 59, 81, 96	0
1	B	88/94 (93%)	-0.33	0	100	100	45, 62, 80, 104	0
1	C	89/94 (94%)	-0.37	0	100	100	56, 64, 80, 100	0
1	D	87/94 (92%)	-0.14	0	100	100	53, 70, 108, 113	0
1	E	88/94 (93%)	-0.27	0	100	100	49, 65, 86, 91	0
1	F	87/94 (92%)	-0.33	0	100	100	51, 73, 97, 100	0
1	G	88/94 (93%)	-0.51	0	100	100	50, 70, 91, 102	0
1	H	87/94 (92%)	-0.20	0	100	100	49, 70, 99, 109	0
All	All	703/752 (93%)	-0.31	0	100	100	45, 67, 95, 113	0

There are no RSRZ outliers to report.

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.