



wwPDB EM Validation Summary Report ⓘ

May 21, 2024 – 10:06 AM JST

PDB ID : 8H8A
EMDB ID : EMD-34540
Title : Type VI secretion system effector RhsP in its post-autoproteolysis and monomeric form
Authors : Tang, L.; Dong, S.Q.; Rasheed, N.; Wu, H.W.; Zhou, N.K.; Li, H.D.; Wang, M.L.; Zheng, J.; He, J.; Chao, W.C.H.
Deposited on : 2022-10-22
Resolution : 3.25 Å(reported)

This is a wwPDB EM 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/EMValidationReportHelp>

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:

EMDB validation analysis : **FAILED**
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36.2

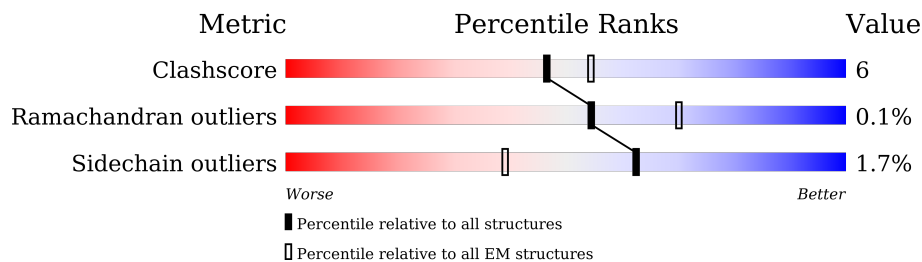
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.25 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)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	1152	
2	B	250	

2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 8454 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Putative Rhs-family protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	964	7844	4941	1360	1534	9	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	initiating methionine	UNP Q87PI5
A	-19	GLY	-	expression tag	UNP Q87PI5
A	-18	SER	-	expression tag	UNP Q87PI5
A	-17	SER	-	expression tag	UNP Q87PI5
A	-16	HIS	-	expression tag	UNP Q87PI5
A	-15	HIS	-	expression tag	UNP Q87PI5
A	-14	HIS	-	expression tag	UNP Q87PI5
A	-13	HIS	-	expression tag	UNP Q87PI5
A	-12	HIS	-	expression tag	UNP Q87PI5
A	-11	HIS	-	expression tag	UNP Q87PI5
A	-10	SER	-	expression tag	UNP Q87PI5
A	-9	GLN	-	expression tag	UNP Q87PI5
A	-8	ASP	-	expression tag	UNP Q87PI5
A	-7	PRO	-	expression tag	UNP Q87PI5
A	-6	GLU	-	expression tag	UNP Q87PI5
A	-5	ASN	-	expression tag	UNP Q87PI5
A	-4	LEU	-	expression tag	UNP Q87PI5
A	-3	TYR	-	expression tag	UNP Q87PI5
A	-2	PHE	-	expression tag	UNP Q87PI5
A	-1	GLN	-	expression tag	UNP Q87PI5
A	0	SER	-	expression tag	UNP Q87PI5

- Molecule 2 is a protein called C-terminal peptide from Putative Rhs-family protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	75	610	376	108	123	3	0	0

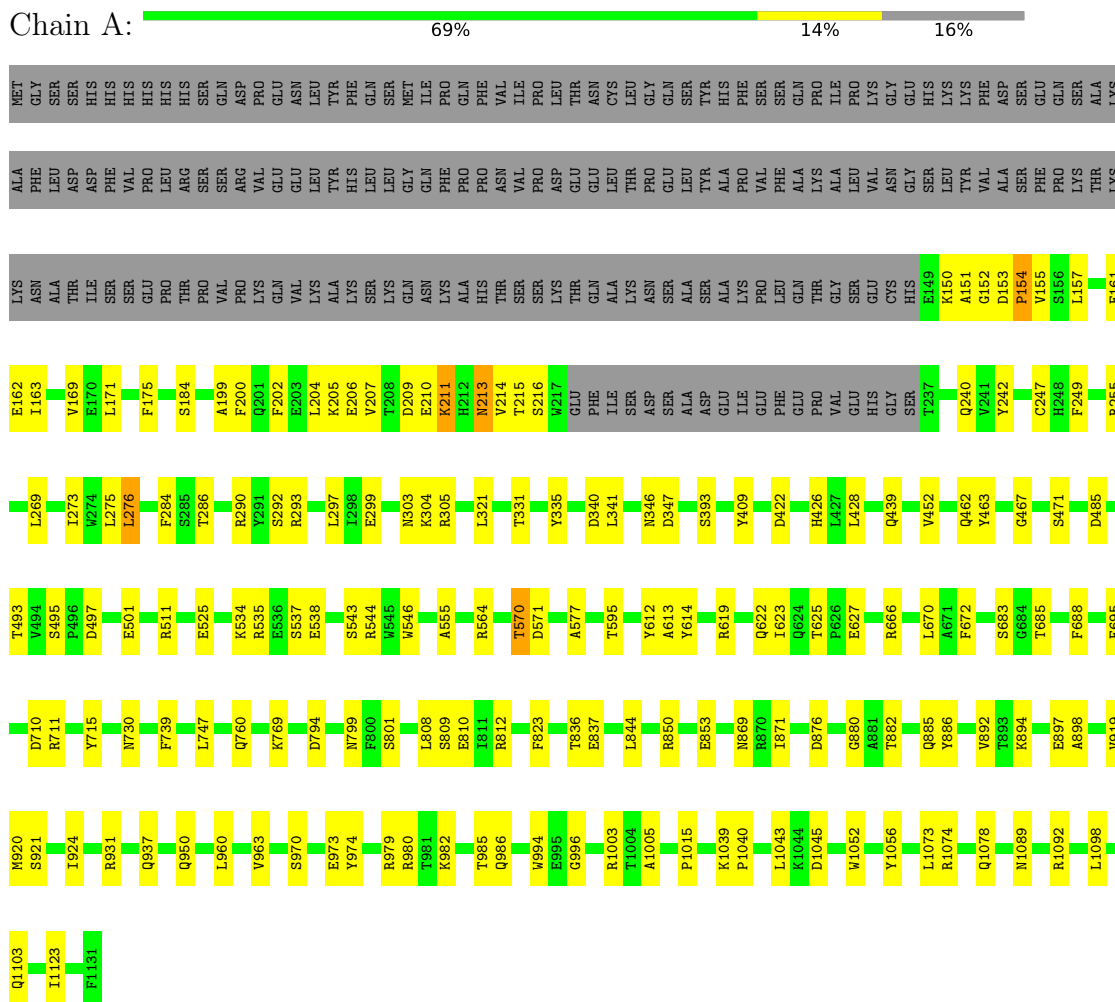
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1354	ALA	HIS	engineered mutation	UNP Q87PI5

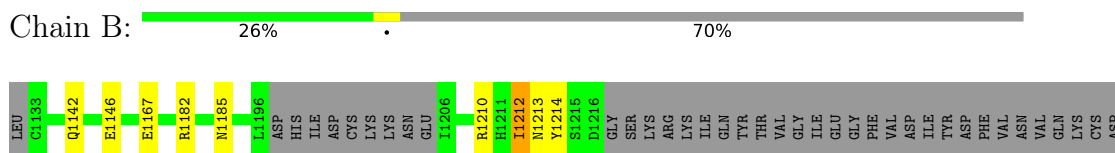
3 Residue-property plots

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. Residues are color-coded 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. 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: Putative Rhs-family protein



- Molecule 2: C-terminal peptide from Putative Rhs-family protein



ALA
GLN
VAL
TYR
ASP
THR
LYS
THR
SER
LYS
THR
VAL
GLY
GLY
ARG
LYS
ILE
TYR
ASP
ILE
ASN
SER
SER
GLU
PHE
PHE
ALA
GLY
LYS
THR
VAL
GLN
THR
ASP
GLY
LYS
GLY
GLY
ASP
MET
VAL
VAL
ARG
PHE
PHE
ASP
SER
SER
ASP
GLY
PHE
PRO
PHE
ASP
PHE
THR
THR
PRO
TYR
TYR
SER
VAL
LYS
LYS
LYS
THR
VAL
ARG
VAL
ILE
GLY
LEU
THR
GLN
GLY
ASP

MET
ALA
ASN
ASP
VAL
ASN
PRO
ASN
LEU
ALA
MET
ASN
ALA
TYR
ARG
ALA
VAL
LYS
ILE
THR
LYS
LYS
TYR
ASP
LYS
SER
SER
LYS
TYR
PHE
TYR
VAL
VAL
TRP
HIS
HIS
HIS
GLN
THR
ASP
GLY
LYS
LYS
THR
MET
ASN
MET
LEU
ILE
PHE
PRO
LYS
SER
SER
VAL
HIS
SER
SER
VAL
ARG
ASN
THR
GLY
GLY
VAL
VAL
ALA
ALA
THR
THR
GLY
GLY
ARG
SER
SER
VAL
ILE
LEU
GLN
HIS
HIS
ASN

LEU
LEU
ASN
PRO
ASN
LYS
LEU
ASN
TYR
SER
SER
PRO
PRO
GLU
GLU
LEU
LEU
VAL

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	368563	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TALOS ARCTICA	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{\AA}^2$)	60	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor

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	0.24	0/8035	0.43	0/10877
2	B	0.23	0/617	0.39	0/827
All	All	0.24	0/8652	0.43	0/11704

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	7844	0	7397	102	0
2	B	610	0	589	7	0
All	All	8454	0	7986	106	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 6.

The worst 5 of 106 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:153:ASP:HB3	1:A:154:PRO:HD3	1.60	0.84
1:A:290:ARG:HB3	1:A:299:GLU:HB3	1.72	0.71

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:152:GLY:HA3	1:A:155:VAL:HG22	1.72	0.70
1:A:462:GLN:HB3	1:A:471:SER:HB3	1.75	0.69
1:A:1078:GLN:HG2	1:A:1089:ASN:HB3	1.75	0.67

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 PDB entries followed by that with respect to all EM entries.

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	960/1152 (83%)	873 (91%)	86 (9%)	1 (0%)	51	82
2	B	71/250 (28%)	67 (94%)	4 (6%)	0	100	100
All	All	1031/1402 (74%)	940 (91%)	90 (9%)	1 (0%)	54	82

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	154	PRO

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 PDB entries followed by that with respect to all EM entries.

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	834/1004 (83%)	820 (98%)	14 (2%)	60	78
2	B	67/218 (31%)	66 (98%)	1 (2%)	65	80

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	901/1222 (74%)	886 (98%)	15 (2%)	62 78

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	284	PHE
1	A	931	ARG
1	A	286	THR
2	B	1212	ILE
1	A	619	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	868	GLN
1	A	1014	GLN
1	A	986	GLN
1	A	1071	ASN
1	A	460	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 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.