



wwPDB EM Validation Summary Report ⓘ

May 21, 2024 – 10:10 AM JST

PDB ID : 8H8C
EMDB ID : EMD-34542
Title : Type VI secretion system effector RhsP in its post-autoproteolysis and dimeric 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 : 3.36 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

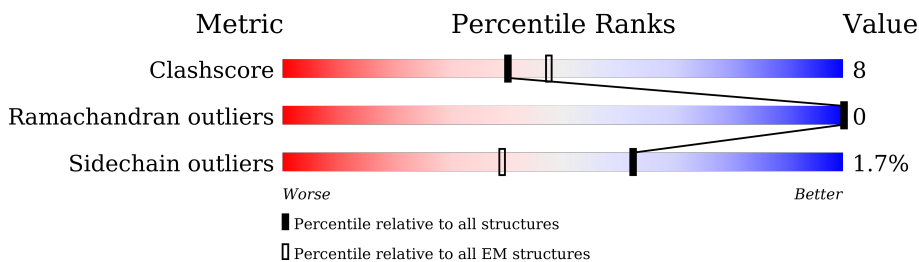
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.36 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	
1	B	1152	
2	C	250	
2	D	250	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 16940 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	B	964	Total	C	N	O	S	0	0
			7860	4955	1361	1535	9		
1	A	964	Total	C	N	O	S	0	0
			7860	4955	1361	1535	9		

There are 42 discrepancies between the modelled and reference sequences:

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

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
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	C	75	Total	C	N	O	S	0	0
			610	376	108	123	3		
2	D	75	Total	C	N	O	S	0	0
			610	376	108	123	3		

There are 2 discrepancies between the modelled and reference sequences:

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

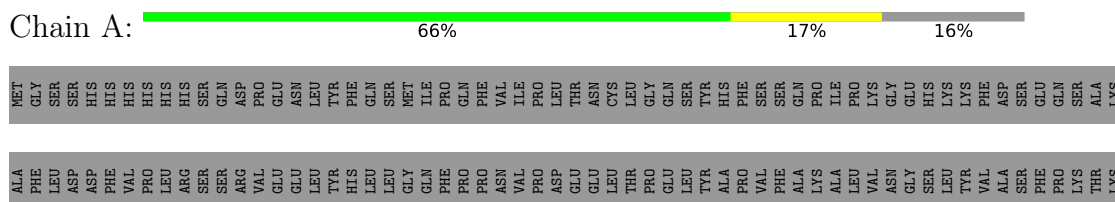
3 Residue-property plots

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 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. 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 1: Putative Rhs-family protein



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

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	212763	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 BIOQUANTUM (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	0/8054	0.43	0/10904
1	B	0.24	0/8054	0.43	0/10904
2	C	0.23	0/617	0.39	0/827
2	D	0.23	0/617	0.40	0/827
All	All	0.24	0/17342	0.43	0/23462

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	7860	0	7412	129	0
1	B	7860	0	7412	117	0
2	C	610	0	589	3	0
2	D	610	0	589	9	0
All	All	16940	0	16002	249	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 8.

The worst 5 of 249 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:204:LEU:HB2	1:B:269:LEU:HD22	1.41	1.01
1:B:239:TYR:CD1	1:B:246:SER:HB3	2.02	0.95
1:B:239:TYR:CD1	1:B:246:SER:CB	2.55	0.89
1:B:205:LYS:H	1:B:211:LYS:HG2	1.36	0.89
1:B:204:LEU:HB2	1:B:269:LEU:CD2	2.11	0.80

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%)	877 (91%)	83 (9%)	0	100	100
1	B	960/1152 (83%)	866 (90%)	94 (10%)	0	100	100
2	C	71/250 (28%)	66 (93%)	5 (7%)	0	100	100
2	D	71/250 (28%)	65 (92%)	6 (8%)	0	100	100
All	All	2062/2804 (74%)	1874 (91%)	188 (9%)	0	100	100

There are no Ramachandran outliers to report.

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	836/1004 (83%)	822 (98%)	14 (2%)	60	80
1	B	836/1004 (83%)	819 (98%)	17 (2%)	55	78

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	C	67/218 (31%)	67 (100%)	0	100	100
2	D	67/218 (31%)	67 (100%)	0	100	100
All	All	1806/2444 (74%)	1775 (98%)	31 (2%)	62	80

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1008	PHE
1	A	783	ARG
1	A	207	VAL
1	A	1065	ASP
1	A	702	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	373	HIS
1	A	737	ASN
1	A	510	GLN
1	A	622	GLN
1	A	772	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.