



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 05:55 AM UTC

PDB ID : 4DQ9 / pdb_00004dq9
Title : Crystal structure of the minor pseudopilin EPSH from the type II secretion system of *Vibrio cholerae*
Authors : Raghunathan, K.; Vago, F.S.; Grindem, D.; Ball, T.; Wedemeyer, W.J.; Arvidson, D.N.
Deposited on : 2012-02-15
Resolution : 1.59 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

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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

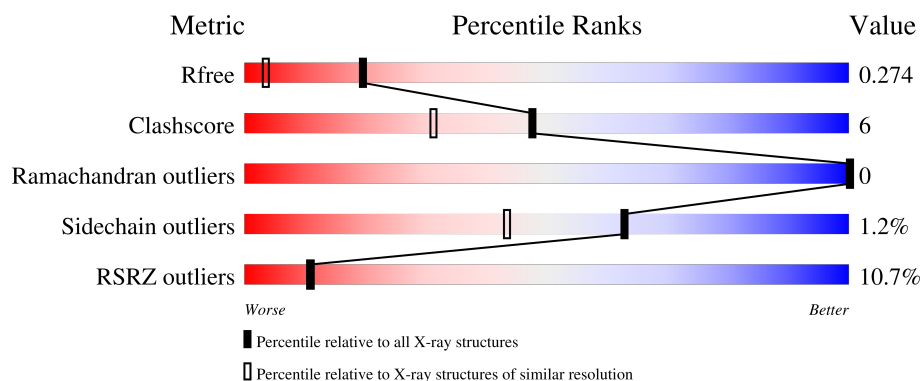
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.59 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}	180053	4673 (1.60-1.60)
Clashscore	190562	4931 (1.60-1.60)
Ramachandran outliers	187476	4831 (1.60-1.60)
Sidechain outliers	187428	4830 (1.60-1.60)
RSRZ outliers	180081	4672 (1.60-1.60)

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 ≥ 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	170	8% 75% 10% • 12%
1	B	170	11% 77% 6% 16%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2618 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 General secretion pathway protein H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	149	Total	C	N	O	S	0	6	0
			1225	773	210	240	2			
1	B	142	Total	C	N	O	S	0	2	0
			1136	717	192	225	2			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	189	LEU	-	expression tag	UNP P45774
A	190	GLU	-	expression tag	UNP P45774
A	191	HIS	-	expression tag	UNP P45774
A	192	HIS	-	expression tag	UNP P45774
A	193	HIS	-	expression tag	UNP P45774
A	194	HIS	-	expression tag	UNP P45774
A	195	HIS	-	expression tag	UNP P45774
A	196	HIS	-	expression tag	UNP P45774
B	189	LEU	-	expression tag	UNP P45774
B	190	GLU	-	expression tag	UNP P45774
B	191	HIS	-	expression tag	UNP P45774
B	192	HIS	-	expression tag	UNP P45774
B	193	HIS	-	expression tag	UNP P45774
B	194	HIS	-	expression tag	UNP P45774
B	195	HIS	-	expression tag	UNP P45774
B	196	HIS	-	expression tag	UNP P45774

- Molecule 2 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Na	0	0
			2	2		

- Molecule 3 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total 1	Cl 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	149	Total 149	O 149	0	0
4	B	105	Total 105	O 105	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	53.39Å 71.11Å 84.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.94 – 1.59 27.94 – 1.59	Depositor EDS
% Data completeness (in resolution range)	96.3 (27.94-1.59) 96.3 (27.94-1.59)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.10 (at 1.59Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.201 , 0.249 0.227 , 0.274	Depositor DCC
R_{free} test set	2128 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	23.9	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 27.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	2618	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.23% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: CL, NA

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	1.44	4/1251 (0.3%)	1.29	1/1691 (0.1%)
1	B	1.31	0/1159	1.12	1/1566 (0.1%)
All	All	1.38	4/2410 (0.2%)	1.21	2/3257 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	171	ALA	C-O	-6.18	1.17	1.24
1	A	34	ILE	C-O	5.22	1.30	1.24
1	A	159	LYS	N-CA	5.16	1.52	1.46
1	A	158	PRO	C-O	5.15	1.29	1.23

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	157	PHE	N-CA-C	6.94	114.47	108.22
1	B	150	THR	O-C-N	5.47	125.74	121.47

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	171	ALA	Mainchain,Peptide

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	1225	0	1186	21	0
1	B	1136	0	1093	6	0
2	A	2	0	0	0	0
3	B	1	0	0	0	0
4	A	149	0	0	1	0
4	B	105	0	0	0	0
All	All	2618	0	2279	26	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.

All (26) 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:174[B]:ASN:OD1	1:A:176[B]:THR:HG23	1.91	0.69
1:A:173[B]:GLU:C	1:A:175[B]:GLY:N	2.51	0.65
1:A:172[B]:GLN:CG	1:A:176[B]:THR:OG1	2.45	0.64
1:A:174[B]:ASN:CG	1:A:176[B]:THR:HG23	2.24	0.63
1:A:173[B]:GLU:C	1:A:175[B]:GLY:H	2.11	0.58
1:A:174[B]:ASN:OD1	1:A:176[B]:THR:CG2	2.54	0.54
1:A:94:LYS:HB2	1:A:94:LYS:NZ	2.24	0.51
1:A:174[B]:ASN:ND2	1:A:176[B]:THR:HG23	2.25	0.51
1:B:99:LEU:HD12	1:B:99:LEU:C	2.37	0.49
1:A:172[B]:GLN:HG2	1:A:176[B]:THR:OG1	2.12	0.49
1:A:176[A]:THR:C	1:A:177:LEU:HD13	2.37	0.48
1:A:175[A]:GLY:O	1:A:177:LEU:HD22	2.12	0.48
1:A:45:LEU:HD23	1:A:45:LEU:C	2.39	0.47
1:A:56:ASP:OD2	1:B:56:ASP:OD2	2.33	0.47
1:A:173[B]:GLU:O	1:A:174[B]:ASN:C	2.57	0.47
1:B:179:LEU:O	1:B:180:LEU:HD23	2.15	0.46
1:B:40:TYR:CZ	1:B:44:LEU:HD11	2.51	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:178[B]:ARG:HH11	1:A:178[B]:ARG:HD3	1.63	0.44
1:A:175[A]:GLY:C	1:A:177:LEU:HD13	2.43	0.43
1:A:67:ARG:HD3	1:A:90:GLN:CD	2.43	0.43
1:B:167:TRP:CE2	1:B:182:PRO:HD3	2.55	0.42
1:A:29:LYS:HG3	4:A:386:HOH:O	2.20	0.42
1:A:44:LEU:HD23	1:A:177:LEU:HD23	2.02	0.41
1:A:172[B]:GLN:HG3	1:A:174[B]:ASN:ND2	2.35	0.41
1:A:37:GLN:O	1:A:41:GLN:HG3	2.21	0.41
1:B:101:PHE:HA	1:B:153:THR:O	2.20	0.41

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	151/170 (89%)	150 (99%)	1 (1%)	0	100	100
1	B	138/170 (81%)	137 (99%)	1 (1%)	0	100	100
All	All	289/340 (85%)	287 (99%)	2 (1%)	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 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	131/149 (88%)	127 (97%)	4 (3%)	35	13
1	B	121/149 (81%)	121 (100%)	0	100	100
All	All	252/298 (85%)	248 (98%)	4 (2%)	63	33

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	94	LYS
1	A	172[A]	GLN
1	A	172[B]	GLN
1	A	177	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	47	ASN
1	B	90	GLN
1	B	172	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 ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	149/170 (87%)	0.48	13 (8%) 16 16	15, 22, 38, 63	6 (4%)
1	B	142/170 (83%)	0.78	18 (12%) 8 7	17, 28, 52, 86	2 (1%)
All	All	291/340 (85%)	0.62	31 (10%) 11 11	15, 25, 52, 86	8 (2%)

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	126	PHE	5.4
1	A	125	MET	4.4
1	A	135	GLN	4.3
1	B	107	ALA	4.0
1	B	176	THR	3.5
1	B	173	GLU	3.3
1	B	108	TRP	3.3
1	B	186	ASP	3.2
1	A	124	GLU	3.1
1	A	175[A]	GLY	3.1
1	A	176[A]	THR	3.0
1	A	174[A]	ASN	2.8
1	B	185	SER	2.7
1	B	165	GLU	2.7
1	A	121	PHE	2.7
1	A	122	ASP	2.7
1	B	182	PRO	2.7
1	B	174	ASN	2.6
1	B	30	ASP	2.5
1	B	177	LEU	2.4
1	B	27	SER	2.4
1	A	183	GLY	2.4
1	B	129	GLU	2.3
1	B	183	GLY	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	175	GLY	2.2
1	A	165	GLU	2.2
1	B	28	VAL	2.2
1	B	109	GLN	2.1
1	A	173[A]	GLU	2.1
1	B	122	ASP	2.1
1	A	77	LYS	2.1

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 oligosaccharides 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, 95th 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(Å ²)	Q<0.9
2	NA	A	201	1/1	0.96	0.06	25,25,25,25	0
2	NA	A	202	1/1	0.97	0.05	28,28,28,28	0
3	CL	B	201	1/1	0.98	0.13	28,28,28,28	0

6.5 Other polymers [i](#)

There are no such residues in this entry.