Full wwPDB X-ray Structure Validation Report (i)

May 7, 2025 – 10:34 AM EDT

PDB ID	:	9DAS / pdb_00009das
Title	:	Crystal structure of vWFA domain from large adhesion protein of Aeromonas hydrophila
Authors	:	Vance, T.D.R.; Ye, Q.; Davies, P.L.
Deposited on	:	2024-08-22
Resolution	:	1.40 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 (i) 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 (i)) were used in the production of this report:

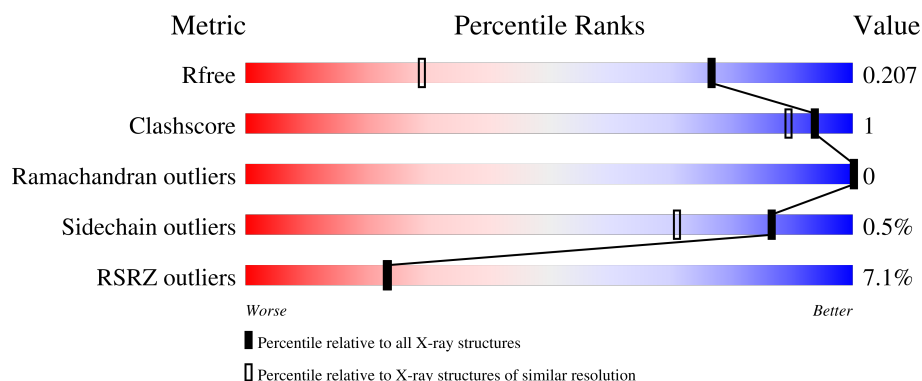|		
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

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 1.40 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		0.207
Clashscore		1
Ramachandran outliers		0
Sidechain outliers		0.5%
RSRZ outliers		7.1%

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}	164625	2247 (1.40-1.40)
Clashscore	180529	2446 (1.40-1.40)
Ramachandran outliers	177936	2398 (1.40-1.40)
Sidechain outliers	177891	2397 (1.40-1.40)
RSRZ outliers	164620	2246 (1.40-1.40)

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 <=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	301	3% / 88% · 9%
1	B	301	10% / 89% · 8%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 4306 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 Flagellin hook IN motif family.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	273	Total	C	N	O	S	0	0	0
			1999	1232	333	427	7			
1	B	277	Total	C	N	O	S	0	0	0
			2027	1249	338	433	7			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	4476	MET	-	initiating methionine	UNP A0KNW4
A	4477	ALA	-	expression tag	UNP A0KNW4
A	4478	SER	-	expression tag	UNP A0KNW4
A	4479	SER	-	expression tag	UNP A0KNW4
A	4480	HIS	-	expression tag	UNP A0KNW4
A	4481	HIS	-	expression tag	UNP A0KNW4
A	4482	HIS	-	expression tag	UNP A0KNW4
A	4483	HIS	-	expression tag	UNP A0KNW4
A	4484	HIS	-	expression tag	UNP A0KNW4
A	4485	HIS	-	expression tag	UNP A0KNW4
A	4486	SER	-	expression tag	UNP A0KNW4
A	4487	SER	-	expression tag	UNP A0KNW4
A	4488	GLY	-	expression tag	UNP A0KNW4
A	4489	LEU	-	expression tag	UNP A0KNW4
A	4490	VAL	-	expression tag	UNP A0KNW4
A	4491	PRO	-	expression tag	UNP A0KNW4
A	4492	ARG	-	expression tag	UNP A0KNW4
A	4493	GLY	-	expression tag	UNP A0KNW4
A	4494	SER	-	expression tag	UNP A0KNW4
A	4495	HIS	-	expression tag	UNP A0KNW4
A	4496	MET	-	expression tag	UNP A0KNW4
B	4476	MET	-	initiating methionine	UNP A0KNW4
B	4477	ALA	-	expression tag	UNP A0KNW4
B	4478	SER	-	expression tag	UNP A0KNW4
B	4479	SER	-	expression tag	UNP A0KNW4

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	4480	HIS	-	expression tag	UNP A0KNW4
B	4481	HIS	-	expression tag	UNP A0KNW4
B	4482	HIS	-	expression tag	UNP A0KNW4
B	4483	HIS	-	expression tag	UNP A0KNW4
B	4484	HIS	-	expression tag	UNP A0KNW4
B	4485	HIS	-	expression tag	UNP A0KNW4
B	4486	SER	-	expression tag	UNP A0KNW4
B	4487	SER	-	expression tag	UNP A0KNW4
B	4488	GLY	-	expression tag	UNP A0KNW4
B	4489	LEU	-	expression tag	UNP A0KNW4
B	4490	VAL	-	expression tag	UNP A0KNW4
B	4491	PRO	-	expression tag	UNP A0KNW4
B	4492	ARG	-	expression tag	UNP A0KNW4
B	4493	GLY	-	expression tag	UNP A0KNW4
B	4494	SER	-	expression tag	UNP A0KNW4
B	4495	HIS	-	expression tag	UNP A0KNW4
B	4496	MET	-	expression tag	UNP A0KNW4

- Molecule 2 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	6	Total 6	Ca 6	0	0
2	B	6	Total 6	Ca 6	0	0

- Molecule 3 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).
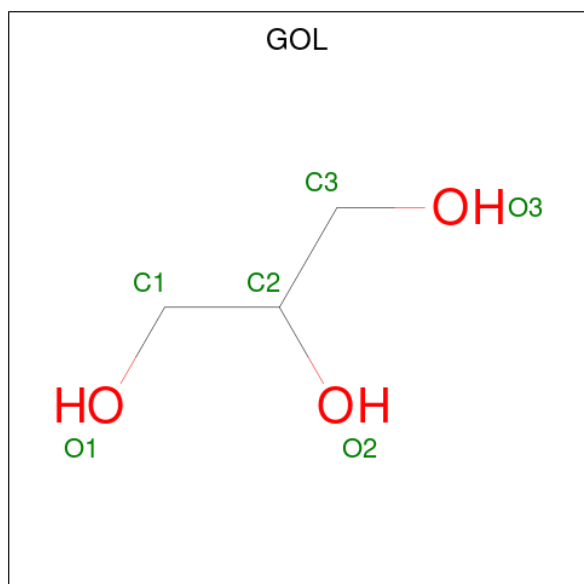
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total 6	C 3	O 3	0	0

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	158	Total 158	O 158	0	0
4	B	104	Total 104	O 104	0	0

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 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: Flagellin hook IN motif family

Chain A:

• Molecule 1: Flagellin hook IN motif family

Chain B:
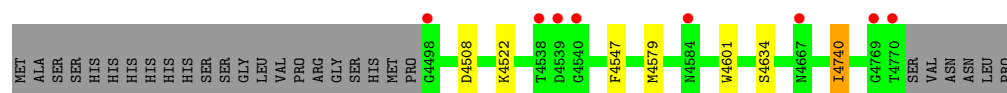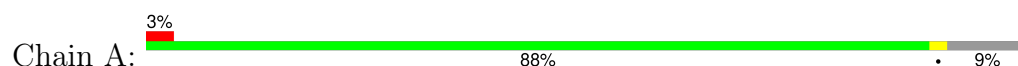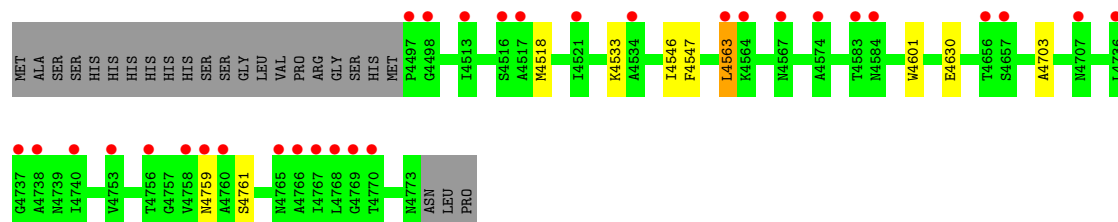
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	71.23Å 156.29Å 57.50Å 90.00° 99.23° 90.00°	Depositor
Resolution (Å)	45.97 – 1.40 45.97 – 1.40	Depositor EDS
% Data completeness (in resolution range)	98.4 (45.97-1.40) 98.4 (45.97-1.40)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.25 (at 1.41Å)	Xtriage
Refinement program	PHENIX (1.21_5207: 000)	Depositor
R, R_{free}	0.191 , 0.207 0.191 , 0.207	Depositor DCC
R_{free} test set	6060 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	16.9	Xtriage
Anisotropy	0.231	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 31.8	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4306	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.57% 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   ⓘ

5.1   Standard geometry   ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, CA

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	0.32	0/2027	0.57	1/2756 (0.0%)
1	B	0.21	0/2056	0.42	0/2796
All	All	0.27	0/4083	0.50	1/5552 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	4508	ASP	CA-CB-CG	6.38	118.98	112.60

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	1999	0	1897	4	0
1	B	2027	0	1927	6	0
2	A	6	0	0	0	0
2	B	6	0	0	0	0
3	B	6	0	8	1	0
4	A	158	0	0	1	0
4	B	104	0	0	0	0
All	All	4306	0	3832	10	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 1.

All (10) 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:4759:ASN:HD22	1:B:4761:SER:H	1.47	0.61
1:A:4522:LYS:HE2	1:A:4579:MET:HB2	1.88	0.55
1:A:4634:SER:HB2	3:B:4807:GOL:H12	1.94	0.48
1:B:4630:GLU:HG2	1:B:4703:ALA:HB2	1.95	0.48
1:A:4740:ILE:HG23	4:A:4902:HOH:O	2.14	0.47
1:B:4547:PHE:HB2	1:B:4601:TRP:CE2	2.51	0.46
1:B:4546:ILE:HG13	1:B:4563:LEU:HG	2.02	0.42
1:A:4547:PHE:HB2	1:A:4601:TRP:CE2	2.56	0.40
1:B:4518:MET:HA	1:B:4518:MET:HE3	2.03	0.40
1:B:4533:LYS:HA	1:B:4563:LEU:HD13	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	271/301 (90%)	268 (99%)	3 (1%)	0	100	100
1	B	275/301 (91%)	271 (98%)	4 (2%)	0	100	100
All	All	546/602 (91%)	539 (99%)	7 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	218/243 (90%)	217 (100%)	1 (0%)	86	71
1	B	222/243 (91%)	221 (100%)	1 (0%)	86	71
All	All	440/486 (90%)	438 (100%)	2 (0%)	86	71

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4740	ILE
1	B	4563	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	4612	ASN
1	A	4739	ASN
1	B	4545	ASN
1	B	4573	GLN
1	B	4603	ASN
1	B	4662	GLN
1	B	4759	ASN

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 13 ligands modelled in this entry, 12 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
3	GOL	B	4807	-	5,5,5	0.29	0	5,5,5	0.57	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	B	4807	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	4807	GOL	O1-C1-C2-C3
3	B	4807	GOL	C1-C2-C3-O3
3	B	4807	GOL	O1-C1-C2-O2
3	B	4807	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	4807	GOL	1	0

5.7 Other polymers ⓘ

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	273/301 (90%)	0.07	8 (2%)	54	55	11, 16, 28, 42	0
1	B	277/301 (92%)	0.90	31 (11%)	11	10	15, 26, 40, 48	0
All	All	550/602 (91%)	0.49	39 (7%)	23	23	11, 20, 38, 48	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	4738	ALA	3.5
1	A	4667	ASN	3.4
1	B	4770	THR	3.3
1	B	4656	THR	3.2
1	B	4758	VAL	3.0
1	A	4769	GLY	3.0
1	B	4753	VAL	2.9
1	A	4540	GLY	2.9
1	B	4584	ASN	2.8
1	B	4497	PRO	2.8
1	B	4768	LEU	2.8
1	B	4498	GLY	2.8
1	A	4539	ASP	2.8
1	A	4584	ASN	2.7
1	B	4521	ILE	2.6
1	A	4770	THR	2.6
1	B	4583	THR	2.5
1	B	4517	ALA	2.5
1	B	4516	SER	2.4
1	B	4564	LYS	2.4
1	B	4756	THR	2.4
1	A	4538	THR	2.4
1	B	4759	ASN	2.4
1	B	4765	ASN	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	4769	GLY	2.3
1	B	4760	ALA	2.3
1	A	4498	GLY	2.3
1	B	4657	SER	2.3
1	B	4567	ASN	2.2
1	B	4707	ASN	2.2
1	B	4740	ILE	2.2
1	B	4767	ILE	2.2
1	B	4513	ILE	2.2
1	B	4563	LEU	2.1
1	B	4534	ALA	2.1
1	B	4736	LEU	2.1
1	B	4574	ALA	2.0
1	B	4766	ALA	2.0
1	B	4737	GLY	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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
3	GOL	B	4807	6/6	0.89	0.13	25,29,30,35	0
2	CA	B	4805	1/1	0.94	0.09	27,27,27,27	0
2	CA	B	4802	1/1	0.97	0.06	32,32,32,32	0
2	CA	B	4803	1/1	0.97	0.07	26,26,26,26	0
2	CA	A	4805	1/1	0.98	0.04	17,17,17,17	0
2	CA	B	4806	1/1	0.98	0.19	30,30,30,30	1
2	CA	A	4806	1/1	0.98	0.03	16,16,16,16	0
2	CA	A	4804	1/1	0.99	0.03	12,12,12,12	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CA	B	4801	1/1	0.99	0.03	15,15,15,15	0
2	CA	B	4804	1/1	0.99	0.05	17,17,17,17	0
2	CA	A	4802	1/1	1.00	0.02	13,13,13,13	0
2	CA	A	4803	1/1	1.00	0.05	18,18,18,18	0
2	CA	A	4801	1/1	1.00	0.02	12,12,12,12	0

6.5 Other polymers ⓘ

There are no such residues in this entry.