



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 01:43 PM UTC

PDB ID : 5XH6 / pdb_00005xh6
Title : Crystal structure of the Acidaminococcus sp. BV3L6 Cpf1 RVR variant in complex with crRNA and target DNA (TATA PAM)
Authors : Nishimasu, H.; Yamano, T.; Ishitani, R.; Nureki, O.
Deposited on : 2017-04-19
Resolution : 2.00 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 ⓘ 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
Mogul : 2022.3.0, CSD as543be (2022)
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

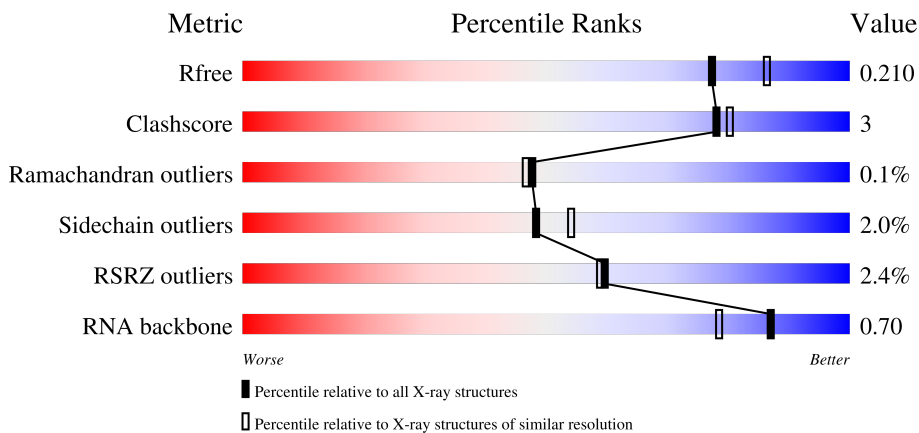
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)
RNA backbone	3983	1000 (2.36-1.64)

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	1310	 3% 90% 7%
2	B	43	 86% 7% 7%
3	C	34	 68% 21% 12%
4	D	10	 80% 20%

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 12831 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 CRISPR-associated endonuclease Cpf1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1283	10455	6704	1772	1951	28	0	13	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP U2UMQ6
A	-1	SER	-	expression tag	UNP U2UMQ6
A	0	HIS	-	expression tag	UNP U2UMQ6
A	542	ARG	SER	engineered mutation	UNP U2UMQ6
A	548	VAL	LYS	engineered mutation	UNP U2UMQ6
A	552	ARG	ASN	engineered mutation	UNP U2UMQ6

- Molecule 2 is a RNA chain called crRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	40	829	370	141	279	39	0	0	0

- Molecule 3 is a DNA chain called Target DNA strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	C	30	611	291	111	179	30	0	0	0

- Molecule 4 is a DNA chain called Non-target DNA strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	D	10	199	97	35	58	9	0	0	0

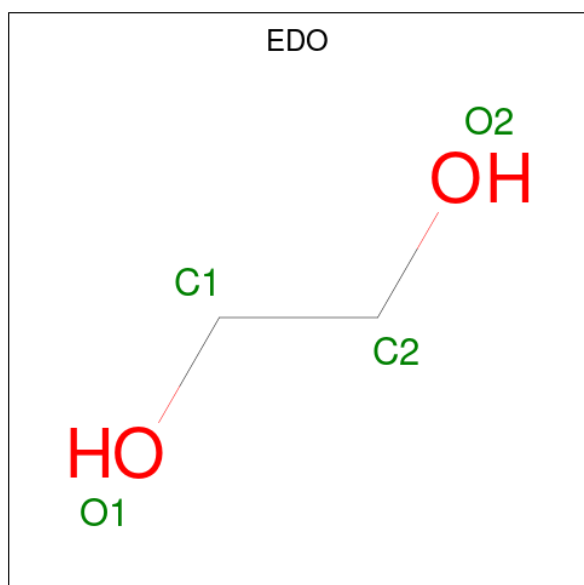
- Molecule 5 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Na 1 1	0	0
5	B	3	Total Na 3 3	0	0
5	C	1	Total Na 1 1	0	0

- Molecule 6 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Cl 1 1	0	0

- Molecule 7 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0

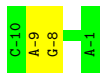
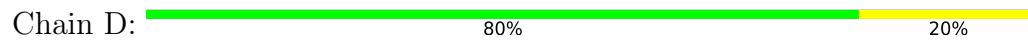
Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	C	1	Total C O 4 2 2	0	0
7	C	1	Total C O 4 2 2	0	0
7	C	1	Total C O 4 2 2	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	486	Total O 486 486	0	0
8	B	95	Total O 95 95	0	0
8	C	65	Total O 65 65	0	0
8	D	13	Total O 13 13	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	81.16Å 133.68Å 199.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.99 – 2.00 39.99 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.9 (39.99-2.00) 97.9 (39.99-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.07 (at 2.00Å)	Xtrriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	0.175 , 0.210 0.177 , 0.210	Depositor DCC
R_{free} test set	7068 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å ²)	40.2	Xtrriage
Anisotropy	0.430	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 44.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	12831	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, NA, CL

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	0.28	0/10735	0.46	0/14524
2	B	0.32	0/925	0.55	0/1440
3	C	0.34	0/684	0.57	0/1052
4	D	0.32	0/222	0.62	0/340
All	All	0.29	0/12566	0.48	0/17356

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	10455	0	10322	59	0
2	B	829	0	416	1	0
3	C	611	0	338	4	0
4	D	199	0	115	1	0
5	A	1	0	0	0	0
5	B	3	0	0	0	0
5	C	1	0	0	0	0
6	A	1	0	0	0	0
7	A	44	0	66	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	16	0	24	1	0
7	C	12	0	18	0	0
8	A	486	0	0	9	0
8	B	95	0	0	0	0
8	C	65	0	0	1	0
8	D	13	0	0	0	0
All	All	12831	0	11299	65	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 3.

The worst 5 of 65 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:983[C]:MET:SD	1:A:990:VAL:HG21	2.14	0.88
4:D:-9:DA:H2'	4:D:-8:DG:C8	2.13	0.83
1:A:900:PRO:HB3	1:A:1282:LYS:HE3	1.68	0.76
1:A:341:ARG:NH2	8:A:1506:HOH:O	2.22	0.72
1:A:124:ARG:HH21	1:A:145:GLN:HB3	1.55	0.71

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	1290/1310 (98%)	1268 (98%)	21 (2%)	1 (0%)	48 46

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	600	ASP

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	1136/1172 (97%)	1114 (98%)	22 (2%)	50 56

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	665	ASP
1	A	989	VAL
1	A	941	GLN
1	A	1014	GLN
1	A	268	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	744	GLN
1	A	969	GLN
1	A	941	GLN
1	A	974	GLN
1	A	293	HIS

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	38/43 (88%)	2 (5%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	-16	U
2	B	-6	G

There are no RNA pucker outliers to report.

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 24 ligands modelled in this entry, 6 are monoatomic - leaving 18 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	$\# Z > 2$
7	EDO	A	1403	-	3,3,3	0.50	0	2,2,2	0.33	0
7	EDO	B	105	-	3,3,3	0.46	0	2,2,2	0.43	0
7	EDO	A	1406	-	3,3,3	0.39	0	2,2,2	0.66	0
7	EDO	A	1411	-	3,3,3	0.39	0	2,2,2	0.59	0
7	EDO	C	103	-	3,3,3	0.38	0	2,2,2	0.48	0
7	EDO	A	1404	-	3,3,3	0.47	0	2,2,2	0.26	0
7	EDO	A	1413	-	3,3,3	0.42	0	2,2,2	0.38	0
7	EDO	A	1410	-	3,3,3	0.39	0	2,2,2	0.60	0
7	EDO	B	106	-	3,3,3	0.44	0	2,2,2	0.32	0
7	EDO	A	1405	-	3,3,3	0.44	0	2,2,2	0.30	0
7	EDO	B	107	-	3,3,3	0.49	0	2,2,2	0.29	0
7	EDO	A	1409	-	3,3,3	0.47	0	2,2,2	0.29	0
7	EDO	A	1407	-	3,3,3	0.60	0	2,2,2	0.15	0
7	EDO	A	1412	-	3,3,3	0.49	0	2,2,2	0.19	0
7	EDO	A	1408	-	3,3,3	0.45	0	2,2,2	0.56	0
7	EDO	B	104	-	3,3,3	0.49	0	2,2,2	0.33	0
7	EDO	C	102	-	3,3,3	0.44	0	2,2,2	0.41	0
7	EDO	C	104	-	3,3,3	0.35	0	2,2,2	0.41	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
7	EDO	A	1403	-	-	0/1/1/1	-
7	EDO	B	105	-	-	0/1/1/1	-
7	EDO	A	1406	-	-	0/1/1/1	-
7	EDO	A	1411	-	-	0/1/1/1	-
7	EDO	C	103	-	-	0/1/1/1	-
7	EDO	A	1404	-	-	0/1/1/1	-
7	EDO	A	1413	-	-	0/1/1/1	-
7	EDO	A	1410	-	-	0/1/1/1	-
7	EDO	B	106	-	-	0/1/1/1	-
7	EDO	A	1405	-	-	0/1/1/1	-
7	EDO	B	107	-	-	0/1/1/1	-
7	EDO	A	1409	-	-	0/1/1/1	-
7	EDO	A	1407	-	-	0/1/1/1	-
7	EDO	A	1412	-	-	0/1/1/1	-
7	EDO	A	1408	-	-	1/1/1/1	-
7	EDO	B	104	-	-	0/1/1/1	-
7	EDO	C	102	-	-	1/1/1/1	-
7	EDO	C	104	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	C	102	EDO	O1-C1-C2-O2
7	A	1408	EDO	O1-C1-C2-O2

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1406	EDO	1	0
7	A	1404	EDO	1	0
7	A	1413	EDO	1	0
7	B	106	EDO	1	0

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	1283/1310 (97%)	0.12	33 (2%) 57 56	16, 49, 80, 121	13 (1%)
2	B	40/43 (93%)	-1.20	0 100 100	31, 39, 51, 88	0
3	C	30/34 (88%)	-0.51	0 100 100	35, 44, 100, 144	0
4	D	10/10 (100%)	0.04	0 100 100	45, 51, 123, 137	0
All	All	1363/1397 (97%)	0.07	33 (2%) 59 59	16, 49, 82, 144	13 (0%)

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	142	VAL	5.5
1	A	486	VAL	5.2
1	A	1010	ALA	4.5
1	A	526	TYR	4.3
1	A	997	PHE	4.2

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(\AA^2)	Q<0.9
7	EDO	A	1408	4/4	0.81	0.17	58,62,63,63	0
7	EDO	A	1404	4/4	0.85	0.21	35,43,59,64	0
7	EDO	A	1411	4/4	0.85	0.17	62,64,66,69	0
7	EDO	B	107	4/4	0.88	0.19	58,59,67,70	0
7	EDO	A	1412	4/4	0.90	0.14	65,66,71,71	0
7	EDO	B	106	4/4	0.90	0.14	55,60,62,64	0
7	EDO	A	1407	4/4	0.90	0.15	41,53,53,61	0
7	EDO	A	1413	4/4	0.92	0.15	65,70,76,79	0
7	EDO	B	104	4/4	0.92	0.12	53,55,58,60	0
5	NA	C	101	1/1	0.93	0.15	56,56,56,56	0
5	NA	B	102	1/1	0.93	0.14	58,58,58,58	0
7	EDO	A	1409	4/4	0.93	0.13	56,59,68,78	0
7	EDO	A	1405	4/4	0.94	0.09	47,49,51,57	0
7	EDO	A	1410	4/4	0.94	0.13	43,51,54,63	0
7	EDO	C	103	4/4	0.94	0.11	48,58,60,65	0
6	CL	A	1402	1/1	0.95	0.09	50,50,50,50	0
7	EDO	C	104	4/4	0.95	0.12	43,48,53,61	0
7	EDO	A	1406	4/4	0.96	0.10	46,49,52,69	0
7	EDO	C	102	4/4	0.96	0.10	48,56,57,62	0
5	NA	B	101	1/1	0.97	0.06	52,52,52,52	0
5	NA	B	103	1/1	0.97	0.05	56,56,56,56	0
7	EDO	B	105	4/4	0.98	0.08	40,40,50,53	0
7	EDO	A	1403	4/4	0.98	0.04	33,35,37,42	0
5	NA	A	1401	1/1	0.99	0.08	26,26,26,26	0

6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.