



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 1, 2025 – 04:35 PM JST

PDB ID : 9LNJ / pdb\_00009lnj  
Title : Crystal structure of VANC21 in complex with its target DNA (5-bromouridine substituted).  
Authors : Tanaka, Y.; Shihoya, W.; Hirano, H.; Itoh, Y.; Kakutani, T.; Nureki, O.  
Deposited on : 2025-01-21  
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](mailto: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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
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)	:	2.46

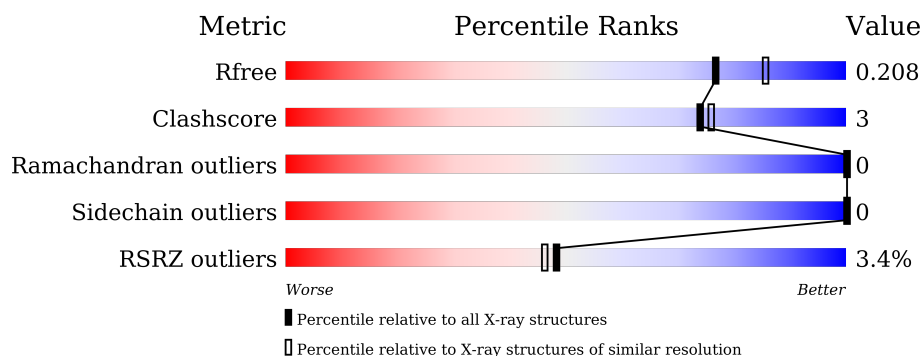
# 1 Overall quality at a glance

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))
$R_{free}$	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

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  $\geq 3$ , 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	383	<div> <div>3%</div> <div> <div></div> <div>85%</div> <div>9%</div> <div>5%</div> </div> </div>
2	B	13	<div> <div>100%</div> </div>
3	C	13	<div> <div>100%</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 3689 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 VANC21.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	362	2959	1904	505	527	23	0	4	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	256	MET	-	initiating methionine	UNP O80464
A	257	GLY	-	expression tag	UNP O80464
A	258	HIS	-	expression tag	UNP O80464
A	259	HIS	-	expression tag	UNP O80464
A	260	HIS	-	expression tag	UNP O80464
A	261	HIS	-	expression tag	UNP O80464
A	262	HIS	-	expression tag	UNP O80464
A	263	HIS	-	expression tag	UNP O80464
A	264	HIS	-	expression tag	UNP O80464
A	265	HIS	-	expression tag	UNP O80464
A	266	HIS	-	expression tag	UNP O80464
A	267	HIS	-	expression tag	UNP O80464
A	268	SER	-	expression tag	UNP O80464
A	269	SER	-	expression tag	UNP O80464
A	270	GLY	-	expression tag	UNP O80464
A	271	GLU	-	expression tag	UNP O80464
A	272	ASN	-	expression tag	UNP O80464
A	273	LEU	-	expression tag	UNP O80464
A	274	TYR	-	expression tag	UNP O80464
A	275	PHE	-	expression tag	UNP O80464
A	276	GLN	-	expression tag	UNP O80464
A	277	GLY	-	expression tag	UNP O80464
A	278	SER	-	expression tag	UNP O80464
A	279	HIS	-	expression tag	UNP O80464

- Molecule 2 is a DNA chain called VANC21 binding DNA fragment (motif strand).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	13	Total	Br	C	N	O	P	0	0	0
			260	2	125	44	77	12			

- Molecule 3 is a DNA chain called VANC21 binding DNA fragment (complementary strand).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	13	Total	Br	C	N	O	P	0	0	0
			267	1	128	51	75	12			

- Molecule 4 is ZINC ION (CCD ID: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Zn	0	0
			1	1		

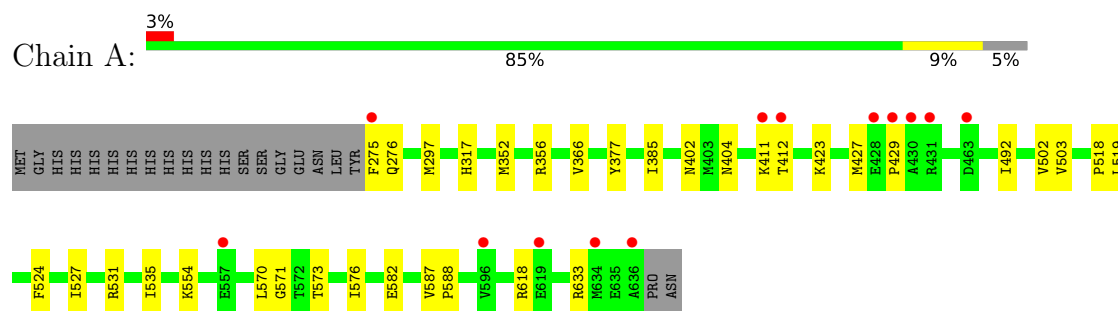
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	150	Total	O	0	0
			150	150		
5	B	19	Total	O	0	0
			19	19		
5	C	33	Total	O	0	0
			33	33		

### 3 Residue-property plots [i](#)

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: VANC21



- Molecule 2: VANC21 binding DNA fragment (motif strand)



There are no outlier residues recorded for this chain.

- Molecule 3: VANC21 binding DNA fragment (complementary strand)



There are no outlier residues recorded for this chain.

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	52.35Å 77.16Å 59.00Å 90.00° 93.30° 90.00°	Depositor
Resolution (Å)	40.26 – 2.00 40.26 – 2.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (40.26-2.00) 99.9 (40.26-2.00)	Depositor EDS
$R_{merge}$	0.54	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.27 (at 2.00Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, $R_{free}$	0.176 , 0.208 0.176 , 0.208	Depositor DCC
$R_{free}$ test set	1605 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	30.7	Xtriage
Anisotropy	0.193	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 49.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3689	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.17% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: 3D1, ZN, BRU

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	0.42	0/3049	0.61	0/4116
2	B	0.51	0/225	0.75	0/344
3	C	0.54	0/257	0.78	0/395
All	All	0.44	0/3531	0.64	0/4855

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

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	2959	0	2946	20	0
2	B	260	0	144	0	0
3	C	267	0	146	0	0
4	A	1	0	0	0	0
5	A	150	0	0	1	0
5	B	19	0	0	0	0
5	C	33	0	0	0	0
All	All	3689	0	3236	20	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 20 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:423:LYS:O	1:A:427:MET:HG3	2.03	0.59
1:A:297:MET:HE3	1:A:503:VAL:HG13	1.89	0.54
1:A:317:HIS:HB3	1:A:429:PRO:HG3	1.93	0.51
1:A:385:ILE:HG21	1:A:518:PRO:HA	1.94	0.49
1:A:524:PHE:O	1:A:531:ARG:HB2	2.14	0.48

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	364/383 (95%)	357 (98%)	7 (2%)	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	323/338 (96%)	323 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such



sidechains are listed below:

Mol	Chain	Res	Type
1	A	568	HIS
1	A	603	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

3 non-standard protein/DNA/RNA residues are modelled in this entry.

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	$\# Z  > 2$
2	BRU	B	12	3,2	18,21,22	0.16	0	26,30,33	0.26	0
2	BRU	B	13	2	18,21,22	0.22	0	26,30,33	0.31	0
3	BRU	C	13	3	18,21,22	0.21	0	26,30,33	0.30	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
2	BRU	B	12	3,2	-	0/7/21/22	0/2/2/2
2	BRU	B	13	2	-	1/7/21/22	0/2/2/2
3	BRU	C	13	3	-	0/7/21/22	0/2/2/2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	13	BRU	C4'-C5'-O5'-P

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	362/383 (94%)	-0.03	13 (3%) 46 44	15, 36, 66, 100	4 (1%)
2	B	10/13 (76%)	-1.07	0 100 100	21, 25, 35, 44	0
3	C	11/13 (84%)	-0.82	0 100 100	24, 30, 34, 41	0
All	All	383/409 (93%)	-0.08	13 (3%) 48 46	15, 35, 66, 100	4 (1%)

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	430	ALA	5.3
1	A	429	PRO	4.9
1	A	275	PHE	3.9
1	A	412	THR	3.7
1	A	636	ALA	3.3

### 6.2 Non-standard residues in protein, DNA, RNA chains [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	BRU	B	13	20/21	0.74	0.20	61,67,78,104	0
3	BRU	C	13	20/21	0.96	0.09	30,35,45,48	0
2	BRU	B	12	20/21	0.98	0.06	25,34,42,48	0

### 6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 6.4 Ligands

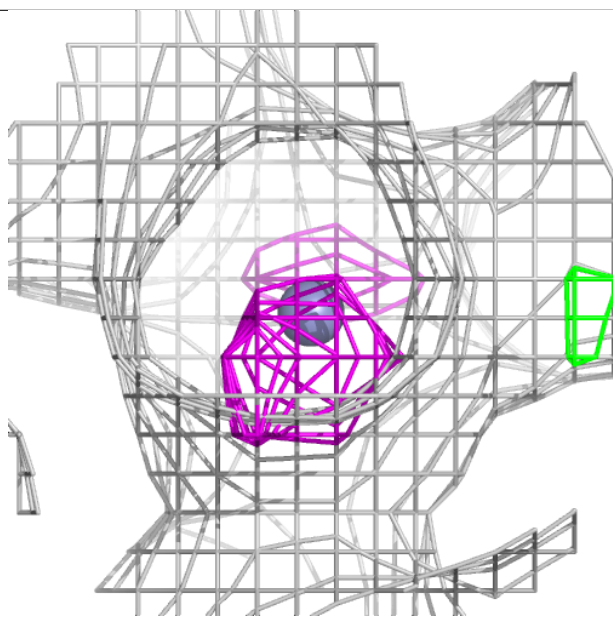
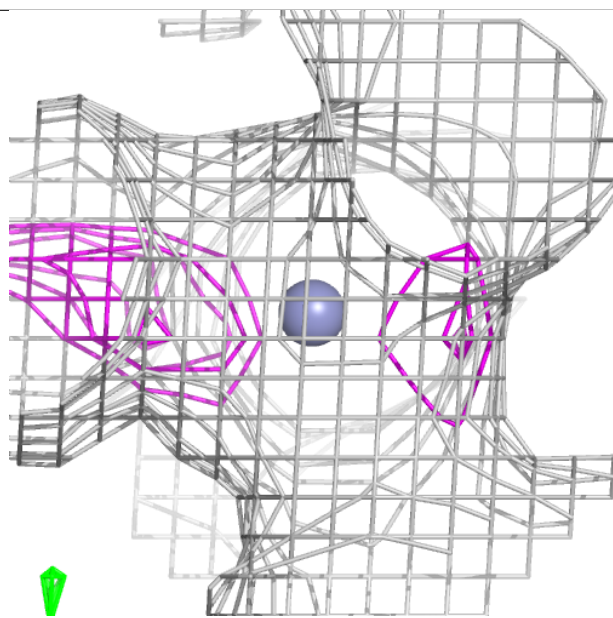
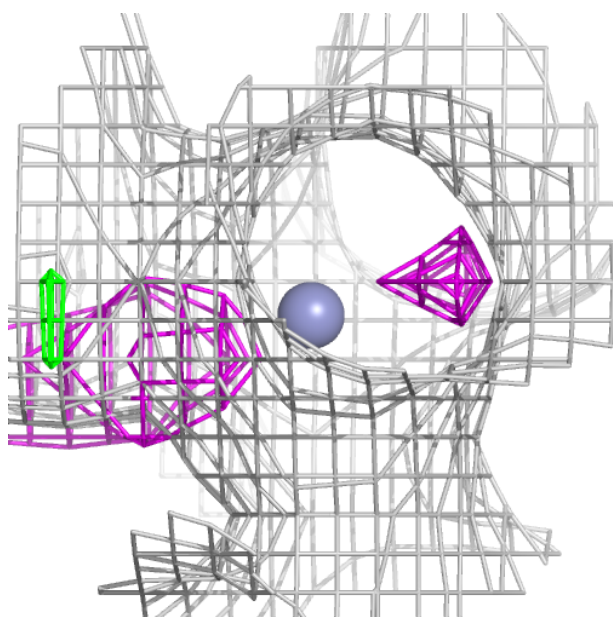
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ZN	A	701	1/1	0.99	0.06	41,41,41,41	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around ZN A 701:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers ⓘ

There are no such residues in this entry.