



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 17, 2023 – 05:11 am GMT

PDB ID : 2YI9  
Title : Structure of the RNA polymerase VP1 from Infectious Pancreatic Necrosis Virus in complex with magnesium  
Authors : Graham, S.C.; Sarin, L.P.; Bahar, M.W.; Myers, R.A.; Stuart, D.I.; Bamford, D.H.; Grimes, J.M.  
Deposited on : 2011-05-11  
Resolution : 2.20 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 : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

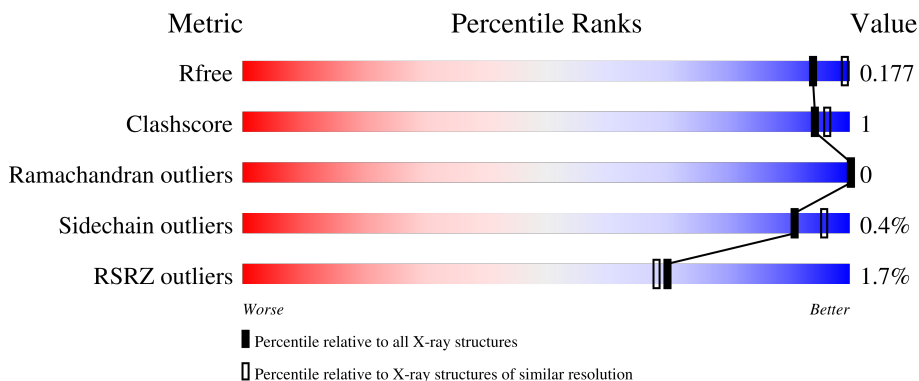
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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))
$R_{free}$	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	799	 2% 93%
1	B	799	 2% 93%
1	C	799	 % 92%
1	D	799	 2% 92%
1	E	799	 % 93%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 33270 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 RNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	771	6043	3836	1018	1162	27	0	4	0
1	B	771	6046	3836	1018	1165	27	0	6	0
1	C	771	6050	3840	1019	1164	27	0	5	0
1	D	771	6056	3843	1021	1165	27	0	7	0
1	E	771	6070	3853	1020	1170	27	0	7	0

There are 45 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	791	LYS	-	expression tag	UNP P22173
A	792	THR	-	expression tag	UNP P22173
A	793	GLY	-	expression tag	UNP P22173
A	794	HIS	-	expression tag	UNP P22173
A	795	HIS	-	expression tag	UNP P22173
A	796	HIS	-	expression tag	UNP P22173
A	797	HIS	-	expression tag	UNP P22173
A	798	HIS	-	expression tag	UNP P22173
A	799	HIS	-	expression tag	UNP P22173
B	791	LYS	-	expression tag	UNP P22173
B	792	THR	-	expression tag	UNP P22173
B	793	GLY	-	expression tag	UNP P22173
B	794	HIS	-	expression tag	UNP P22173
B	795	HIS	-	expression tag	UNP P22173
B	796	HIS	-	expression tag	UNP P22173
B	797	HIS	-	expression tag	UNP P22173
B	798	HIS	-	expression tag	UNP P22173
B	799	HIS	-	expression tag	UNP P22173
C	791	LYS	-	expression tag	UNP P22173

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	792	THR	-	expression tag	UNP P22173
C	793	GLY	-	expression tag	UNP P22173
C	794	HIS	-	expression tag	UNP P22173
C	795	HIS	-	expression tag	UNP P22173
C	796	HIS	-	expression tag	UNP P22173
C	797	HIS	-	expression tag	UNP P22173
C	798	HIS	-	expression tag	UNP P22173
C	799	HIS	-	expression tag	UNP P22173
D	791	LYS	-	expression tag	UNP P22173
D	792	THR	-	expression tag	UNP P22173
D	793	GLY	-	expression tag	UNP P22173
D	794	HIS	-	expression tag	UNP P22173
D	795	HIS	-	expression tag	UNP P22173
D	796	HIS	-	expression tag	UNP P22173
D	797	HIS	-	expression tag	UNP P22173
D	798	HIS	-	expression tag	UNP P22173
D	799	HIS	-	expression tag	UNP P22173
E	791	LYS	-	expression tag	UNP P22173
E	792	THR	-	expression tag	UNP P22173
E	793	GLY	-	expression tag	UNP P22173
E	794	HIS	-	expression tag	UNP P22173
E	795	HIS	-	expression tag	UNP P22173
E	796	HIS	-	expression tag	UNP P22173
E	797	HIS	-	expression tag	UNP P22173
E	798	HIS	-	expression tag	UNP P22173
E	799	HIS	-	expression tag	UNP P22173

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Mg 1 1	0	0
2	B	1	Total Mg 1 1	0	0
2	C	1	Total Mg 1 1	0	0
2	D	1	Total Mg 1 1	0	0
2	E	1	Total Mg 1 1	0	0

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total K 1 1	0	0
3	B	1	Total K 1 1	0	0
3	C	1	Total K 1 1	0	0
3	D	1	Total K 1 1	0	0
3	E	1	Total K 1 1	0	0

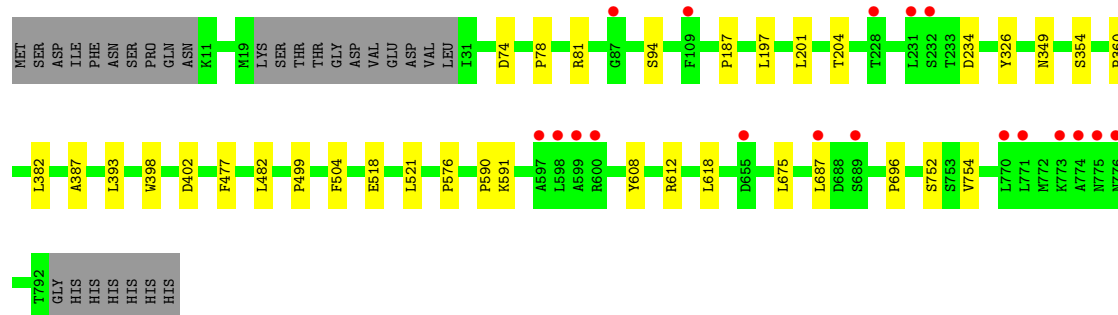
- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Cl 1 1	0	0
4	B	1	Total Cl 1 1	0	0
4	C	1	Total Cl 1 1	0	0
4	D	1	Total Cl 1 1	0	0
4	E	2	Total Cl 2 2	0	0

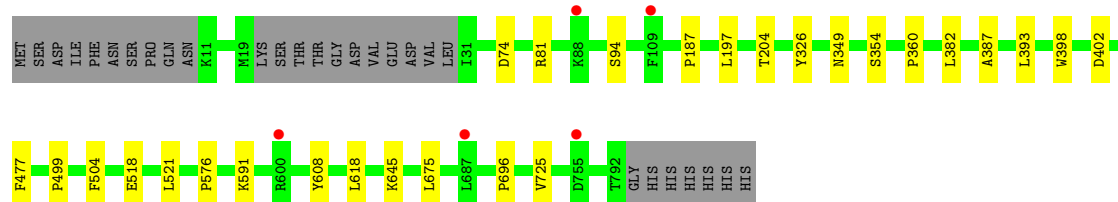
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	527	Total O 527 527	0	0
5	B	591	Total O 591 591	0	0
5	C	570	Total O 570 570	0	0
5	D	632	Total O 632 632	0	0
5	E	669	Total O 669 669	0	0





- Molecule 1: RNA-DIRECTED RNA POLYMERASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.05Å 183.87Å 244.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.21 – 2.20 32.04 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (32.21-2.20) 100.0 (32.04-2.20)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.50 (at 2.20Å)	Xtrriage
Refinement program	BUSTER 2.9.2	Depositor
R, $R_{free}$	0.160 , 0.181 0.159 , 0.177	Depositor DCC
$R_{free}$ test set	3069 reflections (1.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.5	Xtrriage
Anisotropy	0.197	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 48.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	33270	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.18% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, K, 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	0.51	0/6190	0.58	0/8425
1	B	0.52	0/6199	0.59	0/8438
1	C	0.52	0/6200	0.59	0/8438
1	D	0.53	0/6212	0.59	0/8455
1	E	0.56	0/6226	0.60	0/8470
All	All	0.53	0/31027	0.59	0/42226

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	6043	0	5998	16	0
1	B	6046	0	5994	15	0
1	C	6050	0	6009	16	0
1	D	6056	0	6011	18	0
1	E	6070	0	6042	15	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1	0	0	0	0
2	E	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
4	E	2	0	0	0	0
5	A	527	0	0	1	0
5	B	591	0	0	1	0
5	C	570	0	0	0	0
5	D	632	0	0	0	0
5	E	669	0	0	1	0
All	All	33270	0	30054	78	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.

The worst 5 of 78 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:402:ASP:OD2	1:A:518[B]:GLU:HG3	1.84	0.78
1:D:402:ASP:OD2	1:D:518[B]:GLU:HG3	1.84	0.77
1:E:402:ASP:OD2	1:E:518[B]:GLU:HG3	1.84	0.77
1:C:402:ASP:OD2	1:C:518[B]:GLU:HG3	1.84	0.77
1:B:402:ASP:OD2	1:B:518[B]:GLU:HG3	1.84	0.76

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	771/799 (96%)	755 (98%)	16 (2%)	0	100	100
1	B	773/799 (97%)	757 (98%)	16 (2%)	0	100	100
1	C	772/799 (97%)	757 (98%)	15 (2%)	0	100	100
1	D	774/799 (97%)	759 (98%)	15 (2%)	0	100	100
1	E	774/799 (97%)	759 (98%)	15 (2%)	0	100	100
All	All	3864/3995 (97%)	3787 (98%)	77 (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	658/692 (95%)	657 (100%)	1 (0%)	93	97
1	B	659/692 (95%)	657 (100%)	2 (0%)	92	97
1	C	660/692 (95%)	656 (99%)	4 (1%)	86	93
1	D	660/692 (95%)	657 (100%)	3 (0%)	88	94
1	E	665/692 (96%)	663 (100%)	2 (0%)	92	97
All	All	3302/3460 (95%)	3290 (100%)	12 (0%)	91	96

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	94	SER
1	D	326	TYR
1	E	326	TYR
1	D	687	LEU
1	C	94	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

Of 16 ligands modelled in this entry, 16 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 [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	771/799 (96%)	-0.49	18 (2%) 60 58	15, 27, 56, 76	0
1	B	771/799 (96%)	-0.54	14 (1%) 68 66	12, 24, 52, 72	0
1	C	771/799 (96%)	-0.52	11 (1%) 75 73	13, 24, 49, 69	0
1	D	771/799 (96%)	-0.48	18 (2%) 60 58	10, 22, 50, 78	0
1	E	771/799 (96%)	-0.64	5 (0%) 89 88	10, 19, 42, 68	0
All	All	3855/3995 (96%)	-0.54	66 (1%) 70 68	10, 23, 50, 78	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	232	SER	4.8
1	D	232	SER	4.8
1	D	600	ARG	3.9
1	A	775	ASN	3.8
1	A	776	ASN	3.6

### 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 monosaccharides 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, 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
2	MG	A	1793	1/1	0.95	0.05	28,28,28,28	0
2	MG	C	1793	1/1	0.95	0.04	27,27,27,27	0
2	MG	D	1793	1/1	0.96	0.03	24,24,24,24	0
2	MG	B	1793	1/1	0.98	0.04	30,30,30,30	0
2	MG	E	1793	1/1	0.99	0.02	26,26,26,26	0
3	K	A	1794	1/1	0.99	0.04	24,24,24,24	0
3	K	B	1794	1/1	0.99	0.03	23,23,23,23	0
3	K	C	1794	1/1	0.99	0.04	26,26,26,26	0
4	CL	A	1795	1/1	0.99	0.04	29,29,29,29	0
4	CL	C	1795	1/1	0.99	0.06	28,28,28,28	0
4	CL	D	1795	1/1	0.99	0.08	25,25,25,25	0
4	CL	E	1795	1/1	0.99	0.07	23,23,23,23	0
4	CL	E	1796	1/1	0.99	0.07	26,26,26,26	0
3	K	D	1794	1/1	1.00	0.03	18,18,18,18	0
4	CL	B	1795	1/1	1.00	0.05	25,25,25,25	0
3	K	E	1794	1/1	1.00	0.04	19,19,19,19	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.