



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:33 AM GMT

PDB ID : 2R7O  
Title : Crystal Structure of VP1 apoenzyme of Rotavirus SA11 (N-terminal hexahistidine-tagged)  
Authors : Lu, X.; Harrison, S.C.; Tao, Y.J.; Patton, J.T.; Nibert, M.L.  
Deposited on : 2007-09-09  
Resolution : 3.35 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  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

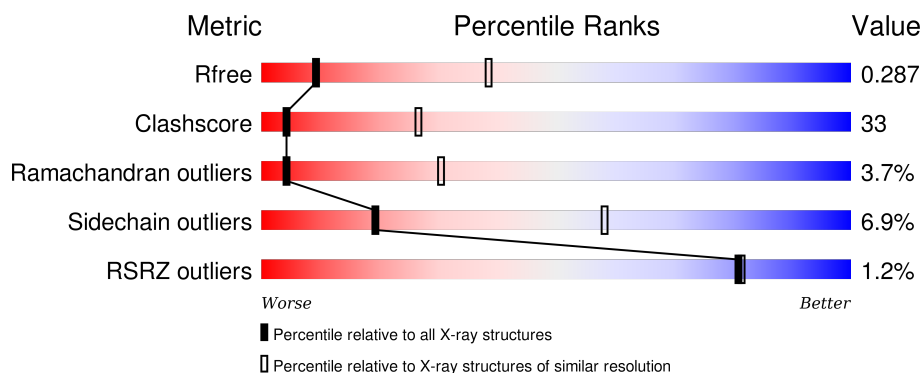
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

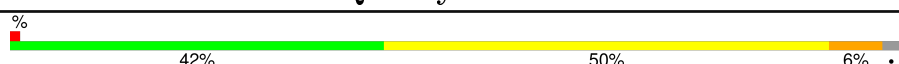
The reported resolution of this entry is 3.35 Å.

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))
$R_{free}$	91344	1005 (3.42-3.30)
Clashscore	102246	1076 (3.42-3.30)
Ramachandran outliers	100387	1059 (3.42-3.30)
Sidechain outliers	100360	1058 (3.42-3.30)
RSRZ outliers	91569	1010 (3.42-3.30)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	1095	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 8741 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.

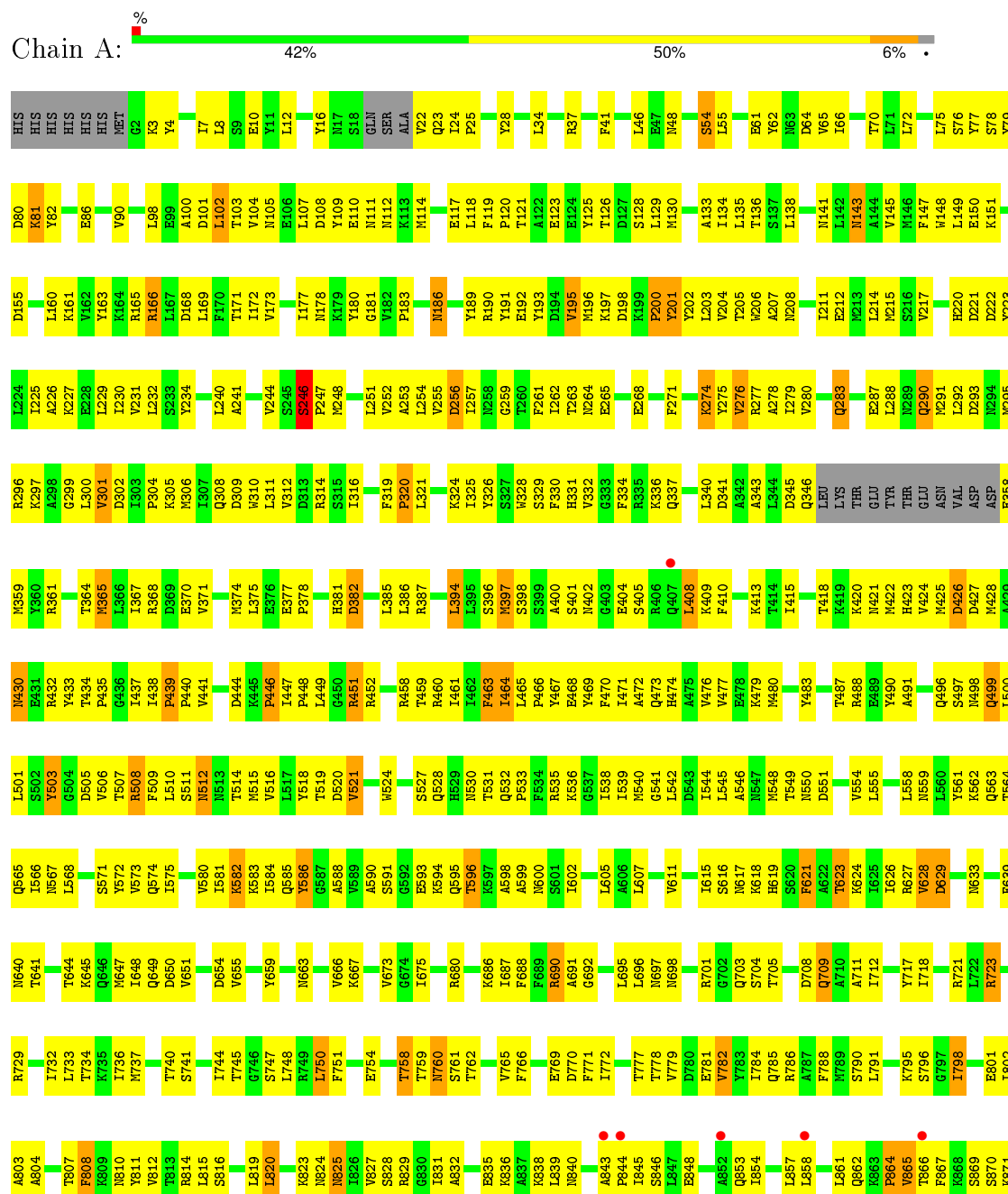
- Molecule 1 is a protein called RNA-dependent RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1071	Total	C	N	O	S	0	10	0
			8741	5604	1457	1642	38			

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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: RNA-dependent RNA polymerase



K1025	I1026	Y1029	L1033	H1034	A1037	I1042	N1043	Y1044	A1045	I1046	K1047	M1048	G1049	S1050	W1051	I1052	S1053	L1054	Y1058	P1059	M1063	L1066	N1067	K1068	K1069	M1070	W1071	M1072	I1073	T1074	R1077	S1078	P1079	Y1080	T1081	M1082	A1083	M1084	F1085	F1086	GLN	GLU	PRO						
1951	1952	1953	1954	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1976	1977	1978	1981	1982	1983	1984	1985	1988	1991	1992	1996	1999	2000	2001	2002	2003	2004	2005	2006	2009	2010	2013	2014	2015	2016	2017	2018	2023	2024
1872	1873	1874	1875	1879	D880	1881	1882	1883	1884	1885	1886	1887	A890	1891	1892	1893	1894	Q897	K898	1899	1900	1901	1906	1907	1908	Q912	1917	1918	1919	1920	1921	1922	1923	1924	1925	1926	1927	1928	1929	1930	1931	1932	1933	1934	1939	1940	1944	1945	1948

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	76.90Å 112.08Å 142.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.35 40.48 – 3.35	Depositor EDS
% Data completeness (in resolution range)	84.1 (50.00-3.35) 84.1 (40.48-3.35)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.18 (at 3.32Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.235 , 0.318 0.222 , 0.287	Depositor DCC
$R_{free}$ test set	1175 reflections (7.62%)	DCC
Wilson B-factor (Å <sup>2</sup> )	65.2	Xtriage
Anisotropy	0.385	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 49.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 18336 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	8741	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.49% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

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

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	0.40	0/8914	0.62	5/12052 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	1082[A]	ASN	N-CA-C	6.33	128.10	111.00
1	A	1082[B]	ASN	N-CA-C	6.33	128.10	111.00
1	A	1077[A]	ARG	N-CA-C	5.58	126.07	111.00
1	A	1077[B]	ARG	N-CA-C	5.58	126.07	111.00
1	A	1085[A]	PHE	N-CA-C	5.15	124.90	111.00

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	8741	0	8825	584	0
All	All	8741	0	8825	584	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 33.

The worst 5 of 584 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:441:VAL:HB	1:A:447:ILE:HD11	1.45	0.96
1:A:385:LEU:HD23	1:A:479:LYS:HE2	1.48	0.95
1:A:651:VAL:O	1:A:655:VAL:HG23	1.66	0.94
1:A:8:LEU:HD12	1:A:737:MET:HG2	1.49	0.91
1:A:960:LEU:HA	1:A:963:ILE:HD12	1.51	0.90

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	1073/1095 (98%)	926 (86%)	105 (10%)	42 (4%)	4	29

5 of 42 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	54	SER
1	A	397	MET
1	A	401	SER
1	A	864	PRO
1	A	978	SER

### 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.

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
-----	-------	----------	-----------	----------	-------------

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	979/996 (98%)	910 (93%)	69 (7%)	19 57

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	451	ARG
1	A	621	PHE
1	A	1050	SER
1	A	499	GLN
1	A	518	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	565	GLN
1	A	703	GLN
1	A	1034	HIS
1	A	585	GLN
1	A	604	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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 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	1071/1095 (97%)	-0.21	13 (1%) 81 81	4, 37, 95, 149	2 (0%)

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1085[A]	PHE	5.7
1	A	1081[A]	THR	5.5
1	A	852	ALA	5.1
1	A	1082[A]	ASN	4.6
1	A	1084[A]	ASN	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 carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

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

There are no such residues in this entry.