wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 07:51 AM EDT

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*
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 (1)) were used in the production of this report:

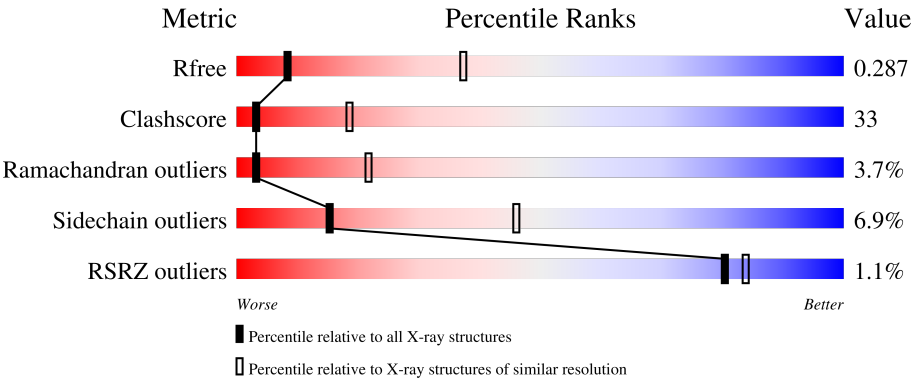
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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}	130704	1558 (3.42-3.30)
Clashscore	141614	1627 (3.42-3.30)
Ramachandran outliers	138981	1599 (3.42-3.30)
Sidechain outliers	138945	1598 (3.42-3.30)
RSRZ outliers	127900	1507 (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 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	1095	% 42% 50% 6%

2 Entry composition (i)

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.

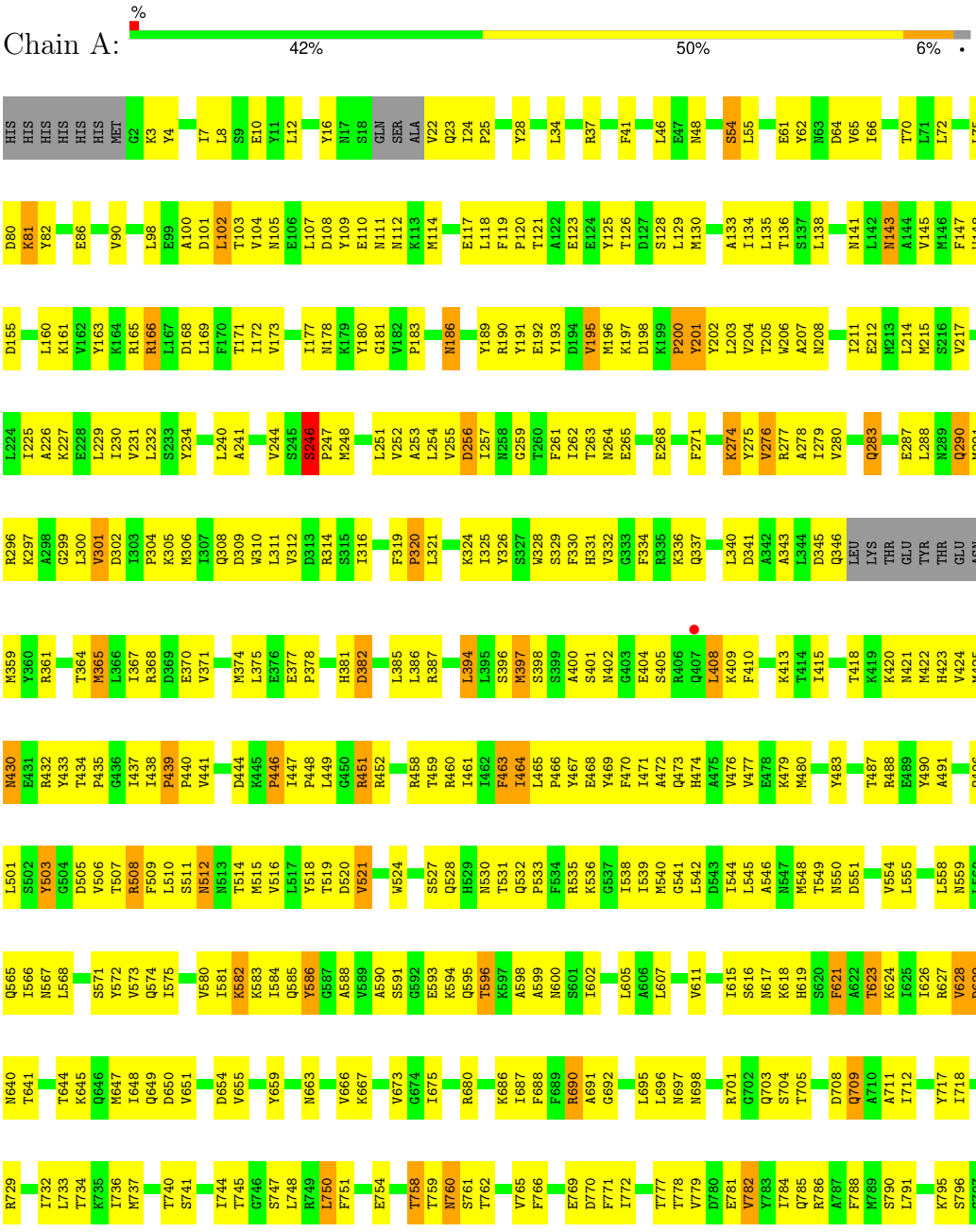
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1071	8741	5604	1457	1642	38	0	10	0

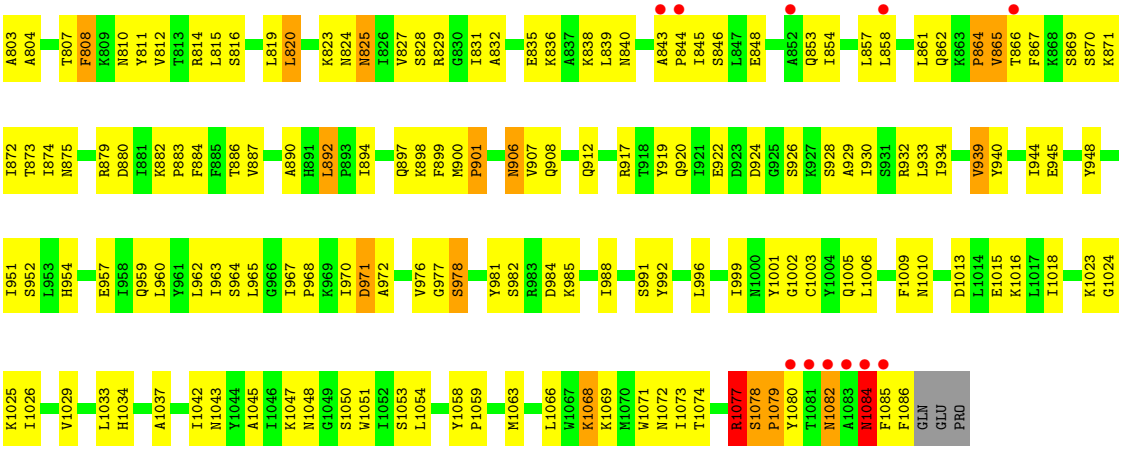
3 Residue-property plots (i)

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: RNA-dependent RNA polymerase

Chain A:




A803 A804 T807 F808 K809 N810 Y811 V812 T813 R814 L815 S816 L819 L820 K823 N824 N825 I826 V827 S828 R829 G830 I831 A832 E835 K836 A837 K838 L839 N840 A843 P844 I845 S846 L847 E848 A852 Q853 I854 L857 L858 L861 Q862 K863 P864 V865 T866 F867 K868 S869 S870 K871
I872 T873 I874 N875 R879 D880 I881 K882 P883 F884 F885 T886 V887 A890 H891 L892 P893 I894 Q897 K898 F899 M900 P901 N906 V907 Q908 Q912 R917 T918 Y919 Q920 I921 E922 D923 D924 G925 S926 K927 S928 A929 I930 S931 R932 L933 I934 V939 Y940 I944 E945 Y948
I951 S952 L953 H954 E957 I958 Q959 L960 Y961 L962 I963 S964 L965 C966 I967 P968 K969 I970 D971 A972 V976 G977 S978 Y981 S982 R983 D984 K985 I988 S991 Y992 L996 I999 N1000 Y1001 G1002 C1003 Y1004 Q1005 L1006 F1009 N1010 D1013 L1014 E1015 K1016 L1017 I1018 K1023 G1024
K1025 I1026 V1029 L1033 H1034 A1037 I1042 N1043 Y1044 A1045 I1046 K1047 N1048 G1049 S1050 W1051 I1052 S1053 L1054 Y1058 P1059 M1063 L1066 W1067 K1068 K1069 M1070 W1071 N1072 I1073 T1074 R1077 S1078 P1079 Y1080 T1081 N1082 A1083 N1084 F1085 F1086 GLN GLU PRO

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	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
$< I/\sigma(I) >$ [1]	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	1424 reflections (7.77%)	wwPDB-VP
Wilson B-factor (Å^2)	65.2	Xtriage
Anisotropy	0.385	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 49.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	8741	wwPDB-VP
Average B, all atoms (Å^2)	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.*

[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

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(^{o})	Ideal(^{o})
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

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%)	3 20

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%)	15 45

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

Mol	Chain	Res	Type
1	A	906	ASN
1	A	939	VAL
1	A	1078[A]	SER
1	A	365	MET
1	A	341	ASP

Sometimes 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	862	GLN
1	A	1043	ASN
1	A	906	ASN
1	A	1005	GLN
1	A	423	HIS

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)

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

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, 95^{th} 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	1071/1095 (97%)	-0.23	12 (1%) 80 84	4, 37, 95, 149	2 (0%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1085[A]	PHE	5.6
1	A	1081[A]	THR	5.4
1	A	852	ALA	5.0
1	A	1082[A]	ASN	4.5
1	A	1084[A]	ASN	4.1

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.