wwPDB X-ray Structure Validation Summary Report (i)

Feb 10, 2024 – 06:59 PM EST

PDB ID	:	2R7T
Title	:	Crystal Structure of Rotavirus SA11 VP1/RNA (UGUGAACC) Complex
Authors	:	Lu, X.; Harrison, S.C.; Tao, Y.J.; Patton, J.T.; Nibert, M.L.
Deposited on	:	2007-09-10
Resolution	:	3.00 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 (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:

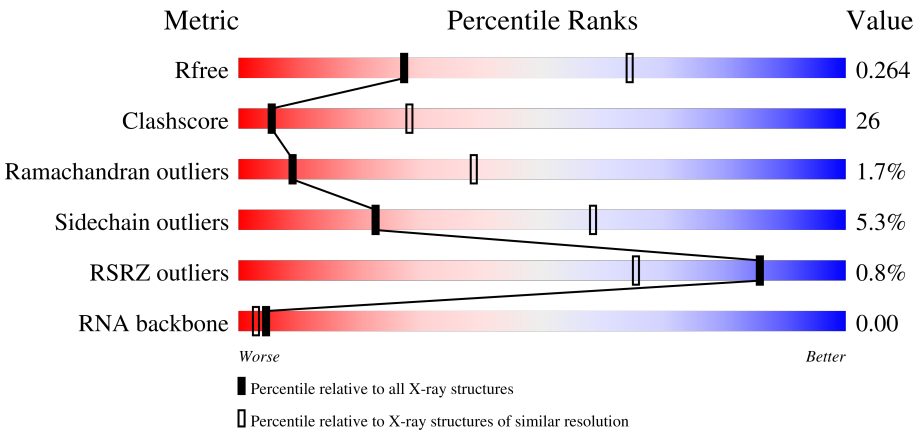
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)
RNA backbone	3102	1173 (3.30-2.70)

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	X	8	88% 12%
2	A	1095	% 49% 45% · ·

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 8846 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 RNA chain called RNA (5'-R(*UP*GP*UP*GP*AP*AP*CP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	X	7	147	67	27	47	6	0	0	0

- Molecule 2 is a protein called RNA-dependent RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	1073	8699	5579	1448	1634	38	0	0	0

3 Residue-property plots (i)

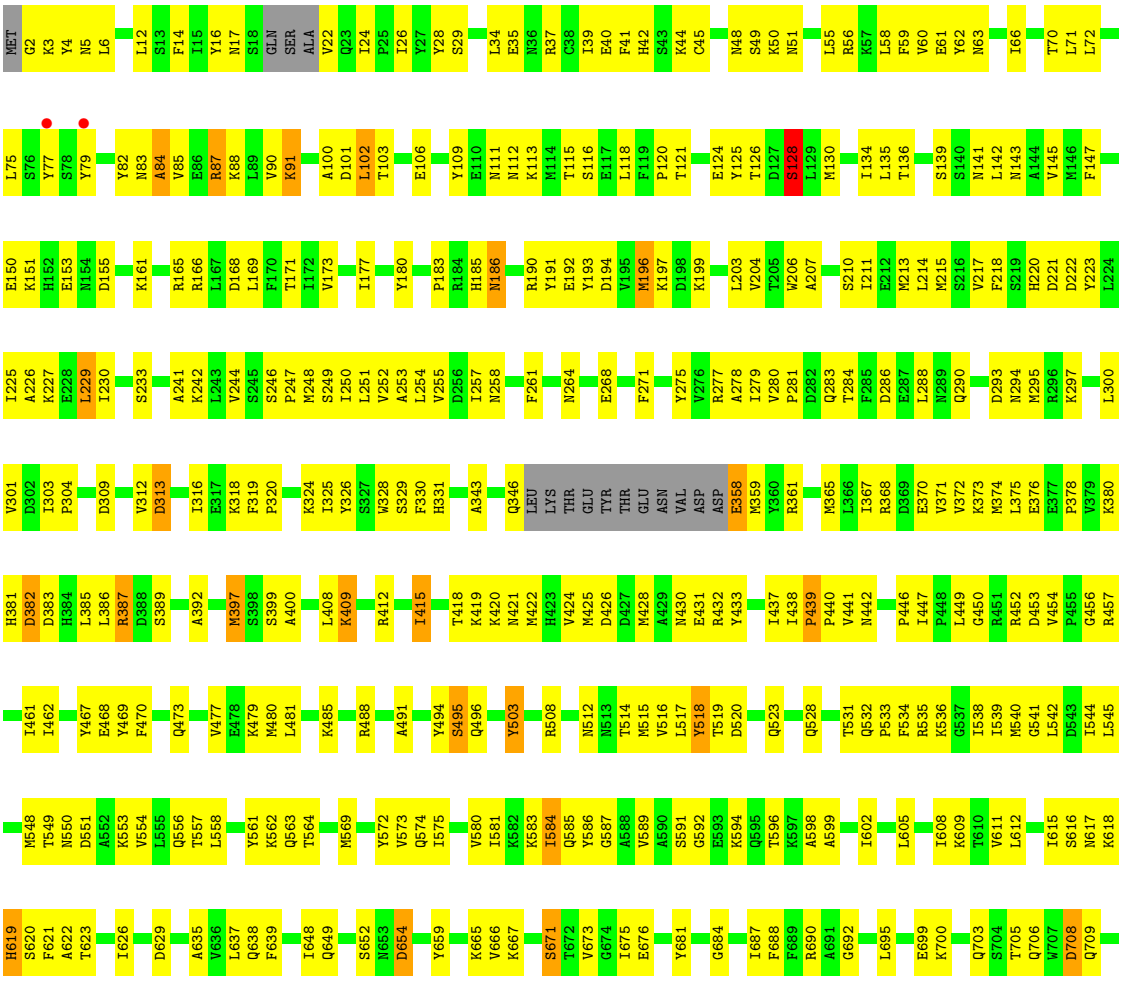
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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.

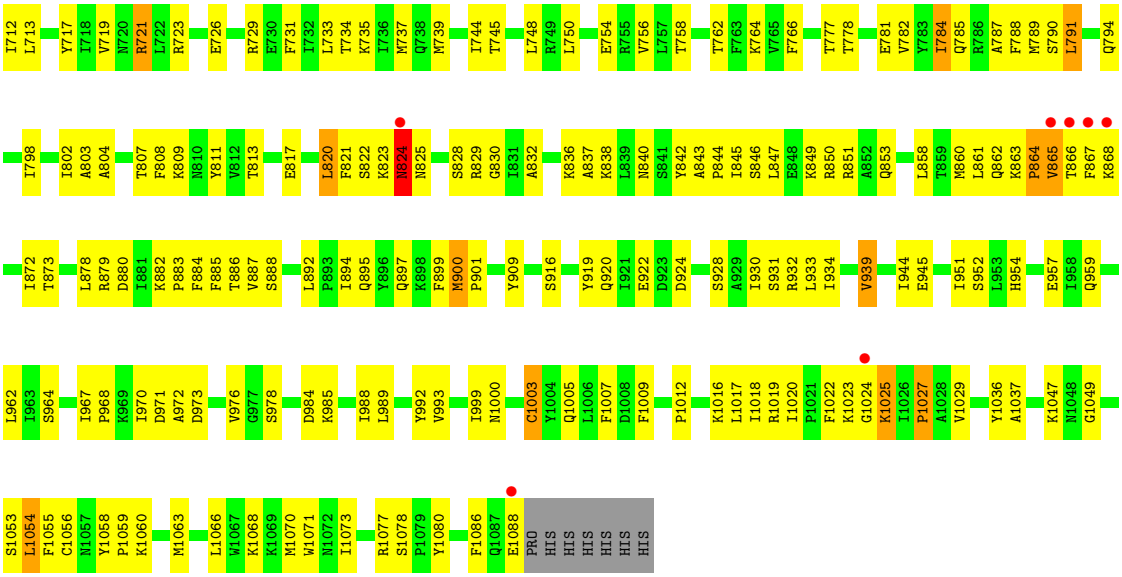
• Molecule 1: RNA (5'-R(*UP*GP*UP*GP*AP*AP*CP*C)-3')

Chain X: 88% 12%

• Molecule 2: RNA-dependent RNA polymerase

Chain A: % 49% 45%




I712 L713 Y717 I718 V719 N720 R721 L722 R723 E726 R729 E730 F731 I732 L733 T734 K735 I736 M737 Q738 M739 I744 T745 L748 R749 L750 E754 R755 V756 L757 T758 T762 F763 K764 V765 F766 T777 T778 E781 V782 Y783 I784 Q785 R786 A787 F788 M789 S790 L791 Q794
I798 I802 A803 A804 T807 F808 K809 N810 Y811 V812 T813 E817 L820 F821 S822 K823 N824 N825 S828 R829 G830 I831 A832 K836 A837 K838 L839 N840 S841 Y842 A843 P844 I845 S846 L847 E848 K849 R850 R851 A852 Q853 L858 T859 M860 L861 Q862 K863 P864 V865 T866 F867 K868
I872 T873 L878 R879 D880 I881 K882 P883 F884 F885 T886 V887 S888 L892 P893 I894 Q895 Y896 Q897 K898 F899 M900 P901 Y909 S916 Y919 Q920 I921 E922 D923 D924 S928 A929 I930 S931 R932 L933 I934 V939 I944 E945 I951 S952 L953 H954 E957 I958 Q959
L962 I963 S964 I967 P968 K969 I970 D971 A972 D973 V976 G977 S978 D984 K985 I988 L989 Y992 V993 I999 N1000 C1003 Y1004 Q1005 L1006 F1007 D1008 F1009 P1012 K1016 L1017 I1018 R1019 I1020 P1021 F1022 K1023 G1024 K1025 I1026 P1027 A1028 V1029 Y1036 A1037 K1047 N1048 G1049
S1053 L1054 F1055 C1056 N1057 Y1058 P1059 K1060 M1063 L1066 W1067 K1068 K1069 M1070 W1071 N1072 I1073 R1077 S1078 P1079 Y1080 F1086 Q1087 E1088 PRO HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	76.65Å 112.53Å 143.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 3.00 40.60 – 3.00	Depositor EDS
% Data completeness (in resolution range)	87.4 (30.00-3.00) 87.5 (40.60-3.00)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.04 (at 3.01Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.235 , 0.280 0.220 , 0.264	Depositor DCC
R_{free} test set	2025 reflections (7.93%)	wwPDB-VP
Wilson B-factor (Å^2)	56.4	Xtriage
Anisotropy	0.424	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 38.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	8846	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.63% 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 (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	X	0.62	0/164	1.85	12/254 (4.7%)
2	A	0.44	1/8870 (0.0%)	0.61	1/11989 (0.0%)
All	All	0.45	1/9034 (0.0%)	0.66	13/12243 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	X	2	0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1003	CYS	CB-SG	-5.61	1.72	1.81

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	824	ASN	N-CA-C	10.30	138.82	111.00
1	X	1101	U	C2'-C3'-O3'	8.23	127.60	109.50
1	X	1105	A	C2'-C3'-O3'	7.73	126.51	109.50
1	X	1103	U	C2'-C3'-O3'	7.35	125.67	109.50
1	X	1102	G	C2'-C3'-O3'	7.23	125.42	109.50

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	X	1105	A	C3'
1	X	1106	A	C3'

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	X	147	0	77	15	0
2	A	8699	0	8793	456	0
All	All	8846	0	8870	463	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 26.

The worst 5 of 463 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)
2:A:385:LEU:HD23	2:A:479:LYS:HE2	1.33	1.07
2:A:520:ASP:HB3	2:A:667:LYS:HG2	1.46	0.96
2:A:101:ASP:OD1	2:A:103:THR:HG22	1.67	0.94
2:A:503:TYR:HB2	2:A:687:ILE:HD13	1.49	0.93
2:A:865:VAL:HG22	2:A:866:THR:H	1.31	0.93

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
2	A	1067/1095 (97%)	966 (90%)	83 (8%)	18 (2%)	9 39

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	382	ASP
2	A	822	SER
2	A	978	SER
2	A	1025	LYS
2	A	106	GLU

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
2	A	975/996 (98%)	923 (95%)	52 (5%)	22 58

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	619	HIS
2	A	784	ILE
2	A	1047	LYS
2	A	620	SER
2	A	671	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	574	GLN
2	A	1057	ASN
2	A	653	ASN
2	A	912	GLN
2	A	646	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	X	7/8 (87%)	6 (85%)	5 (71%)

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	X	1102	G
1	X	1103	U
1	X	1104	G
1	X	1105	A
1	X	1106	A

All (5) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	X	1101	U
1	X	1102	G
1	X	1103	U
1	X	1105	A
1	X	1106	A

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 (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	X	7/8 (87%)	0.08	0 100 100	45, 56, 71, 77	0
2	A	1073/1095 (97%)	-0.36	9 (0%) 86 65	13, 46, 75, 111	0
All	All	1080/1103 (97%)	-0.36	9 (0%) 86 65	13, 46, 75, 111	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	824	ASN	3.7
2	A	867	PHE	3.0
2	A	866	THR	3.0
2	A	1024	GLY	2.9
2	A	77	TYR	2.6

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.