



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 15, 2017 – 09:23 AM EDT

PDB ID : 2R7U  
Title : Crystal Structure of Rotavirus SA11 VP1/RNA (AAAAGCC) Complex  
Authors : Lu, X.; Harrison, S.C.; Tao, Y.J.; Patton, J.T.; Nibert, M.L.  
Deposited on : unknown  
Resolution : 3.10 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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	rb-20030345

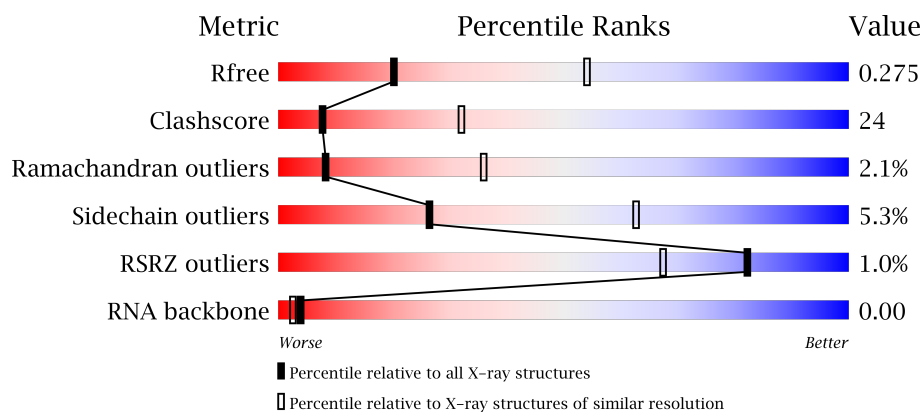
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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))
$R_{free}$	100719	1001 (3.12-3.08)
Clashscore	112137	1099 (3.12-3.08)
Ramachandran outliers	110173	1057 (3.12-3.08)
Sidechain outliers	110143	1057 (3.12-3.08)
RSRZ outliers	101464	1006 (3.12-3.08)
RNA backbone	2435	1112 (3.50-2.70)

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 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	X	7	
2	A	1095	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8803 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(\*AP\*A\*AP\*AP\*GP\*CP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	X	5	Total	C	N	O	P	0	0	0
			104	48	21	31	4			

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

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

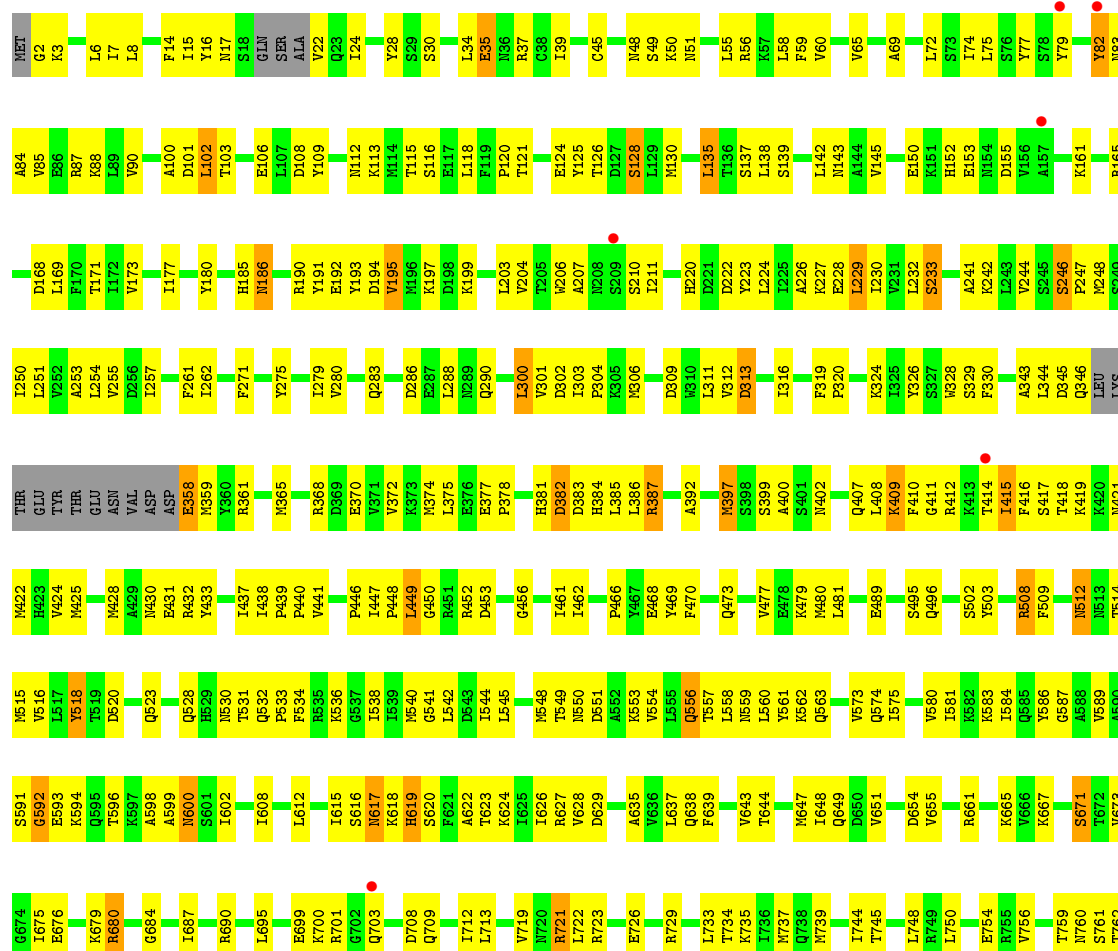
- Molecule 1: RNA (5'-R(\*AP\*A\*AP\*AP\*GP\*CP\*C)-3')

Chain X: 



- Molecule 2: RNA-dependent RNA polymerase

Chain A: 



F763	K764	L847	R923	W1010
D770	F771	R851	D924	S1011
T777	T778	I854	S928	L1017
V779	D780	L857	A929	I1018
E781	V782	L858	I930	R1019
V783	I784	T859	S931	F1022
Q785	F788	L860	R932	K1023
M789	S790	L861	L933	G1024
L791	S796	Q862	S935	K1025
G797	I798	K863	Y939	I1026
A799	T866	P864	I944	P1027
I802	F867	E945	E945	A1028
A804	S796	Y948	I951	F1031
T807	T870	I951	S952	Y1036
F808	K871	S956	N956	A1037
K809	T873	Q959	I963	K1047
N810	L878	I963	S964	N1048
Y811	R879	S964	L965	G1049
V812	D880	I966	I966	S1053
T813	I881	I967	P968	L1054
E817	P883	F968	K969	F1055
L820	F884	D971	S978	C1056
F821	F885	D972	K985	N1057
S822	T886	A972	L989	Y1058
K823	S888	D973	E990	P1059
N824	L892	I976	S991	K1060
N825	F893	G977	Y992	M1063
S828	I894	S978	Y993	I1064
R829	Q895	K985	I999	K1065
G830	Y896	L989	G1002	K1069
I831	Q897	E990	Y1004	M1070
A832	K898	S991	Q1005	W1071
K836	F899	Y992	L1006	N1072
A837	P901	I999	F1007	I1073
K838	Q908	G1002	D1008	B1077
Y842	Y909	C1003	F1009	S1078
A843	Q912	Y1004		P1079
P844	S916	L1005		Y1080
	Y919	L1006		E1088
	E922	F1007		PRO
		F1009		HIS
				HIS
				HIS
				HIS
				HIS

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	76.74Å 112.15Å 144.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 3.10 30.00 – 3.10	Depositor EDS
% Data completeness (in resolution range)	92.4 (30.00-3.10) 92.5 (30.00-3.10)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.86 (at 3.11Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.234 , 0.289 0.225 , 0.275	Depositor DCC
$R_{free}$ test set	1688 reflections (7.86%)	DCC
Wilson B-factor (Å <sup>2</sup> )	61.7	Xtriage
Anisotropy	0.313	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 39.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	8803	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% 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

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	0.64	0/116	1.76	7/179 (3.9%)
2	A	0.45	1/8870 (0.0%)	0.66	2/11989 (0.0%)
All	All	0.45	1/8986 (0.0%)	0.69	9/12168 (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	-6.35	1.71	1.82

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	824	ASN	N-CA-C	9.66	137.08	111.00
1	X	1106	C	C2'-C3'-O3'	8.14	127.41	109.50
1	X	1104	A	C2'-C3'-O3'	7.77	126.60	109.50
1	X	1105	G	C2'-C3'-O3'	7.58	126.17	109.50
1	X	1104	A	C4'-C3'-C2'	6.17	108.77	102.60

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	X	1105	G	C3'
1	X	1106	C	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	104	0	57	10	0
2	A	8699	0	8793	421	0
All	All	8803	0	8850	427	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 24.

The worst 5 of 427 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:520:ASP:HB3	2:A:667:LYS:HG2	1.33	1.05
2:A:865:VAL:HG22	2:A:866:THR:H	1.20	1.05
2:A:101:ASP:OD1	2:A:103:THR:HG22	1.57	1.04
2:A:866:THR:HG22	2:A:867:PHE:H	1.29	0.96
2:A:385:LEU:HD23	2:A:479:LYS:HE2	1.49	0.95

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%)	967 (91%)	78 (7%)	22 (2%)	<b>8</b> 36



5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	106	GLU
2	A	382	ASP
2	A	822	SER
2	A	865	VAL
2	A	978	SER

### 5.3.2 Protein sidechains ⓘ

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%)	26 63

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

Mol	Chain	Res	Type
2	A	508	ARG
2	A	620	SER
2	A	1022	PHE
2	A	512	ASN
2	A	556	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	563	GLN
2	A	585	GLN
2	A	959	GLN
2	A	565	GLN
2	A	617	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	X	5/7 (71%)	4 (80%)	2 (40%)

All (4) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	X	1104	A
1	X	1105	G
1	X	1106	C
1	X	1107	C

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	X	1103	A
1	X	1105	G

## 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 carbohydrates 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, 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	X	5/7 (71%)	0.26	0 100 100	50, 56, 74, 74	0
2	A	1073/1095 (97%)	-0.35	11 (1%) 82 67	11, 46, 78, 120	0
All	All	1078/1102 (97%)	-0.34	11 (1%) 82 67	11, 46, 78, 120	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	824	ASN	3.9
2	A	867	PHE	3.3
2	A	1024	GLY	2.8
2	A	157	ALA	2.7
2	A	868	LYS	2.3

### 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.