wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 16, 2023 – 07:30 PM EDT

PDB ID	:	2B4I
Title	:	Crystal Structure of the Rhesus Rotavirus VP5 Antigen Domain Trimer
Authors	:	Yoder, J.D.; Dormitzer, P.R.
Deposited on	:	2005-09-24
Resolution	:	2.00 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 ⓘ 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:

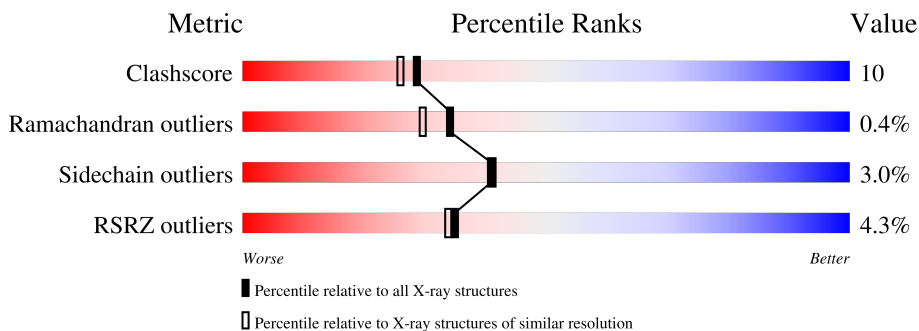|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.00 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	Percentile Ranks	Value
Clashscore		10
Ramachandran outliers		0.4%
Sidechain outliers		3.0%
RSRZ outliers		4.3%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	254	.% 82% 8% · 9%
1	B	254	6% 72% 19% · 8%
1	C	254	5% 69% 21% · 9%

2   Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 6030 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 Outer capsid protein VP4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	232	Total	C	N	O	S	0	0	0
			1842	1171	312	352	7			
1	B	234	Total	C	N	O	S	0	0	0
			1858	1180	316	355	7			
1	C	232	Total	C	N	O	S	0	0	0
			1842	1171	312	352	7			

There are 63 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	226	MET	-	expression tag	UNP Q91HI9
A	227	GLY	-	expression tag	UNP Q91HI9
A	228	SER	-	expression tag	UNP Q91HI9
A	229	SER	-	expression tag	UNP Q91HI9
A	230	HIS	-	expression tag	UNP Q91HI9
A	231	HIS	-	expression tag	UNP Q91HI9
A	232	HIS	-	expression tag	UNP Q91HI9
A	233	HIS	-	expression tag	UNP Q91HI9
A	234	HIS	-	expression tag	UNP Q91HI9
A	235	HIS	-	expression tag	UNP Q91HI9
A	236	SER	-	expression tag	UNP Q91HI9
A	237	SER	-	expression tag	UNP Q91HI9
A	238	GLY	-	expression tag	UNP Q91HI9
A	239	LEU	-	expression tag	UNP Q91HI9
A	240	VAL	-	expression tag	UNP Q91HI9
A	241	PRO	-	expression tag	UNP Q91HI9
A	242	ARG	-	expression tag	UNP Q91HI9
A	243	GLY	-	expression tag	UNP Q91HI9
A	244	SER	-	expression tag	UNP Q91HI9
A	245	HIS	-	expression tag	UNP Q91HI9
A	246	MET	-	expression tag	UNP Q91HI9
B	226	MET	-	expression tag	UNP Q91HI9
B	227	GLY	-	expression tag	UNP Q91HI9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	228	SER	-	expression tag	UNP Q91HI9
B	229	SER	-	expression tag	UNP Q91HI9
B	230	HIS	-	expression tag	UNP Q91HI9
B	231	HIS	-	expression tag	UNP Q91HI9
B	232	HIS	-	expression tag	UNP Q91HI9
B	233	HIS	-	expression tag	UNP Q91HI9
B	234	HIS	-	expression tag	UNP Q91HI9
B	235	HIS	-	expression tag	UNP Q91HI9
B	236	SER	-	expression tag	UNP Q91HI9
B	237	SER	-	expression tag	UNP Q91HI9
B	238	GLY	-	expression tag	UNP Q91HI9
B	239	LEU	-	expression tag	UNP Q91HI9
B	240	VAL	-	expression tag	UNP Q91HI9
B	241	PRO	-	expression tag	UNP Q91HI9
B	242	ARG	-	expression tag	UNP Q91HI9
B	243	GLY	-	expression tag	UNP Q91HI9
B	244	SER	-	expression tag	UNP Q91HI9
B	245	HIS	-	expression tag	UNP Q91HI9
B	246	MET	-	expression tag	UNP Q91HI9
C	226	MET	-	expression tag	UNP Q91HI9
C	227	GLY	-	expression tag	UNP Q91HI9
C	228	SER	-	expression tag	UNP Q91HI9
C	229	SER	-	expression tag	UNP Q91HI9
C	230	HIS	-	expression tag	UNP Q91HI9
C	231	HIS	-	expression tag	UNP Q91HI9
C	232	HIS	-	expression tag	UNP Q91HI9
C	233	HIS	-	expression tag	UNP Q91HI9
C	234	HIS	-	expression tag	UNP Q91HI9
C	235	HIS	-	expression tag	UNP Q91HI9
C	236	SER	-	expression tag	UNP Q91HI9
C	237	SER	-	expression tag	UNP Q91HI9
C	238	GLY	-	expression tag	UNP Q91HI9
C	239	LEU	-	expression tag	UNP Q91HI9
C	240	VAL	-	expression tag	UNP Q91HI9
C	241	PRO	-	expression tag	UNP Q91HI9
C	242	ARG	-	expression tag	UNP Q91HI9
C	243	GLY	-	expression tag	UNP Q91HI9
C	244	SER	-	expression tag	UNP Q91HI9
C	245	HIS	-	expression tag	UNP Q91HI9
C	246	MET	-	expression tag	UNP Q91HI9

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	161	Total 161	O 161	0	0
2	B	134	Total 134	O 134	0	0
2	C	193	Total 193	O 193	0	0

3   Residue-property plots ⓘ

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.
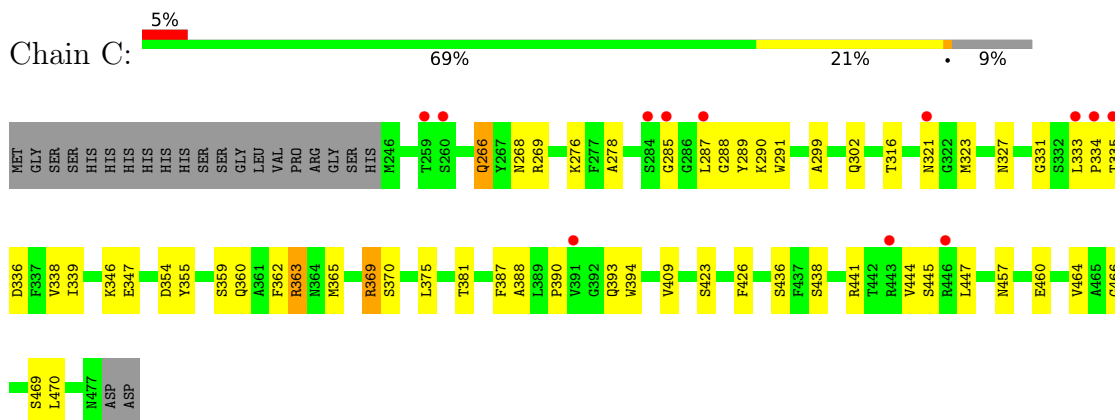
• Molecule 1: Outer capsid protein VP4
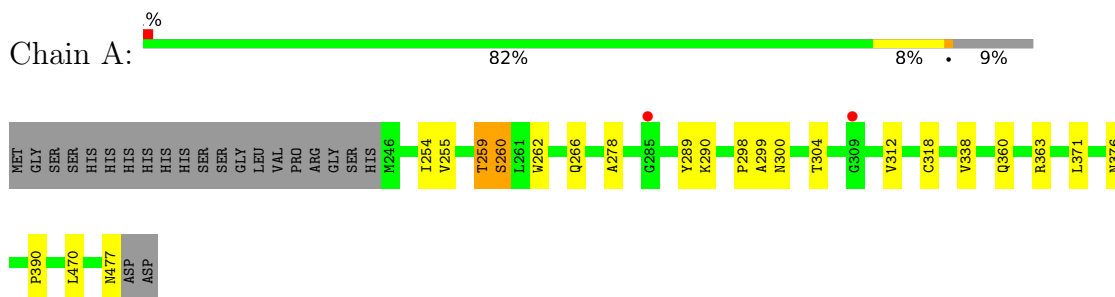
• Molecule 1: Outer capsid protein VP4
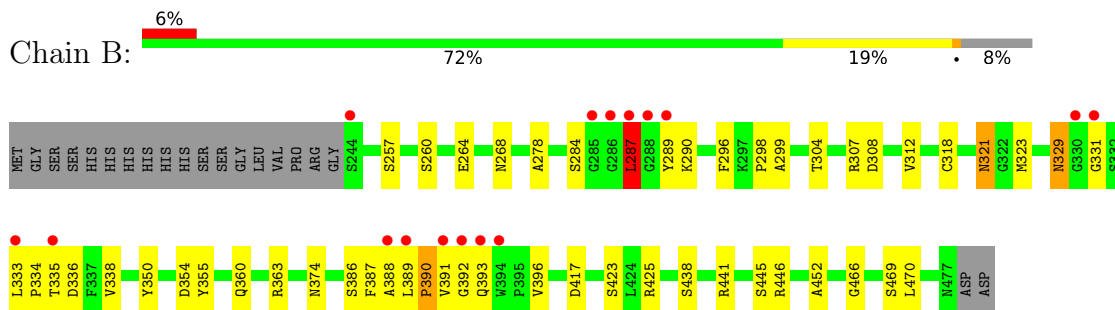
• Molecule 1: Outer capsid protein VP4

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	93.12Å 94.03Å 97.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.00 33.08 – 2.00	Depositor EDS
% Data completeness (in resolution range)	94.8 (50.00-2.00) 90.1 (33.08-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	3.62 (at 2.00Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.222 , 0.252 0.240 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	33.0	Xtriage
Anisotropy	0.550	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 43.9	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.015 for -h,l,k 0.018 for -l,-k,-h 0.014 for k,h,-l 0.024 for k,l,h 0.024 for l,h,k	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6030	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.84% 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	0.41	0/1887	0.68	0/2562
1	B	0.45	0/1904	0.71	0/2585
1	C	0.44	1/1887 (0.1%)	0.72	0/2562
All	All	0.43	1/5678 (0.0%)	0.70	0/7709

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	334	PRO	N-CD	5.05	1.54	1.47

There are no bond angle outliers.

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	1842	0	1776	14	0
1	B	1858	0	1788	54	0
1	C	1842	0	1776	38	0
2	A	161	0	0	0	0
2	B	134	0	0	3	0
2	C	193	0	0	5	0
All	All	6030	0	5340	105	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 10.

The worst 5 of 105 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:B:289:TYR:CE1	1:B:390:PRO:HD2	2.02	0.94
1:B:290:LYS:HG2	1:B:338:VAL:HB	1.52	0.88
1:B:289:TYR:OH	1:B:390:PRO:HB2	1.73	0.88
1:B:289:TYR:CE2	1:B:390:PRO:HG2	2.10	0.86
1:B:331:GLY:HA2	2:B:612:HOH:O	1.78	0.83

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	230/254 (91%)	221 (96%)	9 (4%)	0	100	100
1	B	232/254 (91%)	215 (93%)	15 (6%)	2 (1%)	17	11
1	C	230/254 (91%)	215 (94%)	14 (6%)	1 (0%)	34	30
All	All	692/762 (91%)	651 (94%)	38 (6%)	3 (0%)	34	30

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	287	LEU
1	C	285	GLY
1	B	390	PRO

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	
1	A	202/221 (91%)	197 (98%)	5 (2%)	47	49
1	B	204/221 (92%)	198 (97%)	6 (3%)	42	43
1	C	202/221 (91%)	195 (96%)	7 (4%)	36	35
All	All	608/663 (92%)	590 (97%)	18 (3%)	41	41

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	347	GLU
1	C	460	GLU
1	C	369	ARG
1	B	323	MET
1	C	327	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	329	ASN
1	C	249	GLN
1	C	376	ASN
1	C	268	ASN
1	C	302	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 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	A	232/254 (91%)	-0.13	2 (0%)	84 83	25, 35, 50, 67	0
1	B	234/254 (92%)	0.32	16 (6%)	17 16	27, 39, 76, 100	0
1	C	232/254 (91%)	0.25	12 (5%)	27 26	24, 36, 65, 82	0
All	All	698/762 (91%)	0.15	30 (4%)	35 34	24, 37, 67, 100	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	286	GLY	10.7
1	B	288	GLY	5.4
1	B	285	GLY	5.0
1	C	285	GLY	4.6
1	C	287	LEU	3.9

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.