W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 06:49 PM GMT

PDB ID : 1CN3
Title : INTERACTION OF POLYOMAVIRUS INTERNAL PROTEIN VP2 WITH
MAJOR CAPSID PROTEIN VP1 AND IMPLICATIONS FOR PARTICIPA-
TION OF VP2 IN VIRAL ENTRY
Authors : Chen, X.; Stehle, T.; Harrison, S.C.
Deposited on : 1999-05-24
Resolution : 2.20 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : NOT EXECUTED
EDS : NOT EXECUTED
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
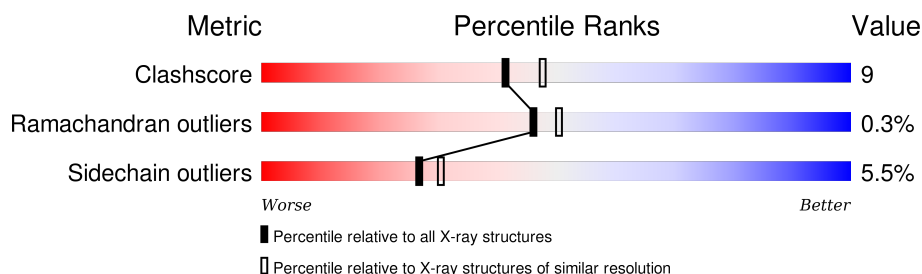Validation Pipeline (wwPDB-VP) : trunk26865

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.20 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		9
Ramachandran outliers		0.3%
Sidechain outliers		5.5%

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	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)

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 >=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 <=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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	283	81% 17% •
1	B	283	79% 18% •
1	C	283	80% 17% •
1	D	283	80% 16% •
1	E	283	78% 19% •
2	F	29	31% 41% 28%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 11261 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 COAT PROTEIN VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	283	Total 2213	C 1400	N 371	O 431	S 11	0	0	0
1	B	283	Total 2213	C 1400	N 371	O 431	S 11	0	0	0
1	C	283	Total 2213	C 1400	N 371	O 431	S 11	0	0	0
1	D	283	Total 2213	C 1400	N 371	O 431	S 11	0	0	0
1	E	283	Total 2213	C 1400	N 371	O 431	S 11	0	0	0

- Molecule 2 is a protein called FRAGMENT OF COAT PROTEIN VP2.

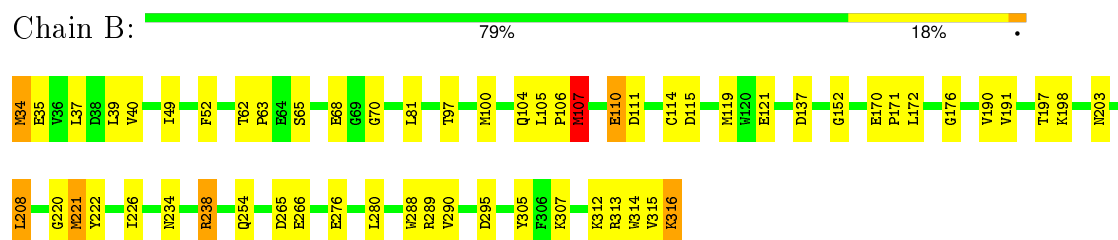| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	29	Total 196	C 124	N 36	O 35	S 1	0	0	0

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 (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.
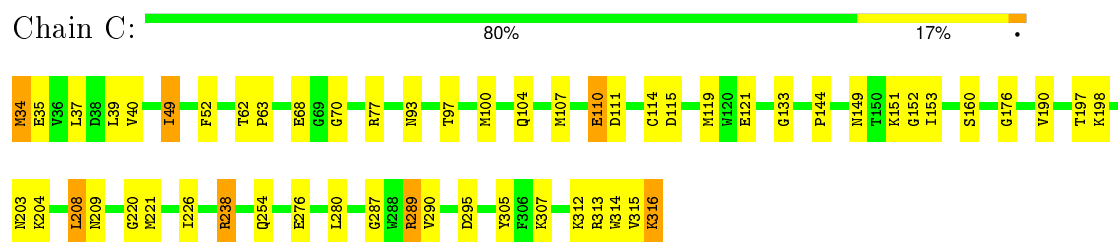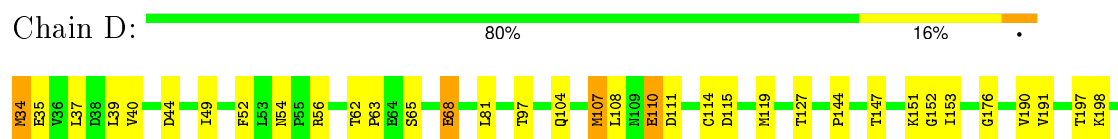
Note EDS was not executed.

- Molecule 1: COAT PROTEIN VP1
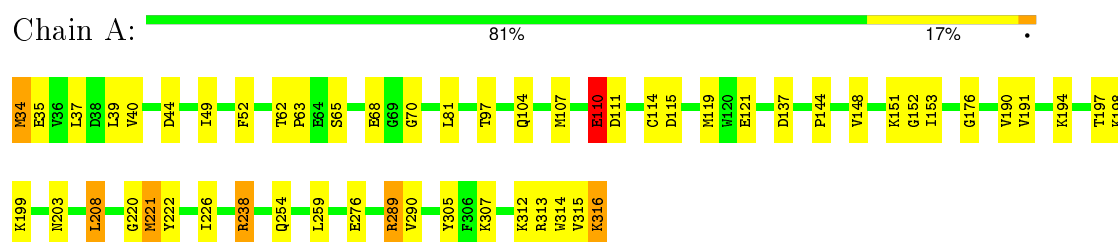
Chain A:

- Molecule 1: COAT PROTEIN VP1

Chain B:

- Molecule 1: COAT PROTEIN VP1

Chain C:

- Molecule 1: COAT PROTEIN VP1

Chain D:

N203 L208 G220 M221 Y222 E225 I226 R238 V252 L253 Q254 E276 R289 V290 D295 Y305 F306 K307 K312 R313 W314 V315 K316

● Molecule 1: COAT PROTEIN VP1

Chain E:

78%   19%   •

M34 E35 V36 L37 D38 L39 V40 I49 T62 P63 E64 S65 E68 G69 G70 L81 T97 Q104 M107 E110 D111 C114 D115 M119 W120 E121 T127 D137 P144 K151 G152 I153 V157 L172 Q175 G176 K186 E187 E188 G189 V190 V191

T197 K198 V207 L208 I211 S212 K213 G220 M221 Y222 I226 N234 R238 N242 Q254 E276 W288 R289 V290 D295 R304 Y305 F306 K307 L310 R311 K312 R313 W314 V315 K316
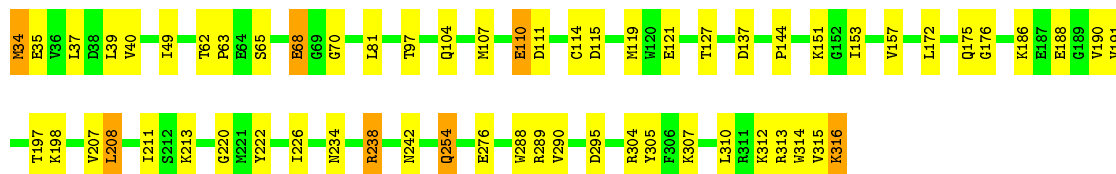
● Molecule 2: FRAGMENT OF COAT PROTEIN VP2

Chain F:

31%   41%   28%

G1 G4 G7 G8 A9 A10 S11 H2 Q13 R14 V15 T16 P17 D18 W19 M20 L21 P22 L23 I24 L25 G26 I27 Y28 G29

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	219.00Å 219.00Å 99.00Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.20	Depositor
% Data completeness (in resolution range)	95.0 (20.00-2.20)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	0.15	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.230 , 0.260	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	11261	wwPDB-VP
Average B, all atoms (Å²)	21.0	wwPDB-VP

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	1.04	3/2266 (0.1%)	1.01	3/3088 (0.1%)
1	B	1.07	6/2266 (0.3%)	0.99	1/3088 (0.0%)
1	C	1.06	3/2266 (0.1%)	0.99	2/3088 (0.1%)
1	D	1.08	7/2266 (0.3%)	1.01	4/3088 (0.1%)
1	E	1.07	4/2266 (0.2%)	0.99	2/3088 (0.1%)
2	F	0.65	0/201	1.03	1/271 (0.4%)
All	All	1.06	23/11531 (0.2%)	1.00	13/15711 (0.1%)

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	221	MET	SD-CE	6.56	2.14	1.77
1	C	238	ARG	CZ-NH2	6.36	1.41	1.33
1	B	52	PHE	CE2-CZ	6.21	1.49	1.37
1	D	68	GLU	CD-OE2	6.17	1.32	1.25
1	C	68	GLU	CD-OE2	5.99	1.32	1.25

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	208	LEU	CA-CB-CG	-6.63	100.04	115.30
1	D	208	LEU	CA-CB-CG	-6.61	100.10	115.30
1	E	208	LEU	CA-CB-CG	-6.57	100.18	115.30
1	A	208	LEU	CA-CB-CG	-6.50	100.36	115.30
1	C	208	LEU	CA-CB-CG	-6.09	101.29	115.30

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	2213	0	2181	38	1
1	B	2213	0	2181	41	0
1	C	2213	0	2181	37	1
1	D	2213	0	2181	36	1
1	E	2213	0	2181	42	3
2	F	196	0	192	28	0
All	All	11261	0	11097	191	3

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 9.

The worst 5 of 191 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:C:221:MET:SD	1:C:221:MET:CE	2.05	1.44
1:B:107:MET:CE	1:B:107:MET:SD	2.06	1.43
1:B:221:MET:SD	1:B:221:MET:CE	2.08	1.40
1:D:221:MET:SD	1:D:221:MET:CE	2.10	1.38
1:A:221:MET:SD	1:A:221:MET:CE	2.14	1.35

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:149:ASN:OD1	1:E:186:LYS:NZ[3_564]	2.03	0.17
1:D:147:THR:CG2	1:E:188:GLU:OE2[3_564]	2.07	0.13
1:A:148:VAL:CG2	1:E:211:ILE:CD1[3_564]	2.15	0.05

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	281/283 (99%)	271 (96%)	9 (3%)	1 (0%)	39	42
1	B	281/283 (99%)	270 (96%)	11 (4%)	0	100	100
1	C	281/283 (99%)	272 (97%)	9 (3%)	0	100	100
1	D	281/283 (99%)	269 (96%)	12 (4%)	0	100	100
1	E	281/283 (99%)	270 (96%)	11 (4%)	0	100	100
2	F	27/29 (93%)	18 (67%)	5 (18%)	4 (15%)	0	0
All	All	1432/1444 (99%)	1370 (96%)	57 (4%)	5 (0%)	46	50

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	8	GLY
2	F	11	SER
2	F	10	ALA
2	F	23	LEU
1	A	110	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	
1	A	249/249 (100%)	236 (95%)	13 (5%)	29	33
1	B	249/249 (100%)	236 (95%)	13 (5%)	29	33
1	C	249/249 (100%)	236 (95%)	13 (5%)	29	33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	249/249 (100%)	236 (95%)	13 (5%)	29	33
1	E	249/249 (100%)	236 (95%)	13 (5%)	29	33
2	F	17/17 (100%)	13 (76%)	4 (24%)	1	0
All	All	1262/1262 (100%)	1193 (94%)	69 (6%)	27	30

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

Mol	Chain	Res	Type
1	C	190	VAL
1	C	316	LYS
1	E	290	VAL
1	C	208	LEU
1	C	276	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	242	ASN
1	E	254	GLN

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 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)

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS was not executed - this section will therefore be empty.

6.4 Ligands (i)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers (i)

EDS was not executed - this section will therefore be empty.