wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 08:53 PM EDT

PDB ID	:	5J96
Title	:	Crystal structure of Slow Bee Paralysis Virus at 3.4A resolution
Authors	:	Kalynych, S.; Levdansky, Y.; Palkova, L.; Plevka, P.
Deposited on	:	2016-04-08
Resolution	:	3.41 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:

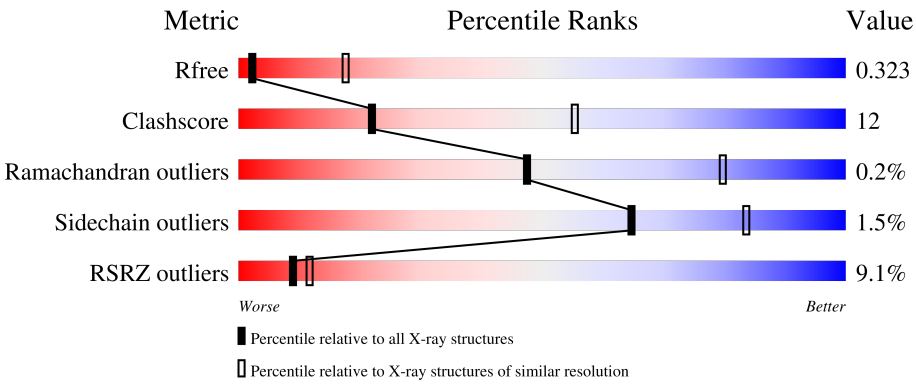
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.41 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	1486 (3.50-3.34)
Clashscore	141614	1572 (3.50-3.34)
Ramachandran outliers	138981	1534 (3.50-3.34)
Sidechain outliers	138945	1535 (3.50-3.34)
RSRZ outliers	127900	1395 (3.50-3.34)

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	266	3%; 58%; 31%; • ; 10%
2	B	261	3%; 74%; 22%; • •
3	C	430	15%; 69%; 27%; •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7029 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 VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	239	1896	1216	327	345	8	0	0	0

- Molecule 2 is a protein called VP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	252	1954	1260	316	364	14	0	0	0

- Molecule 3 is a protein called Genome polyprotein.

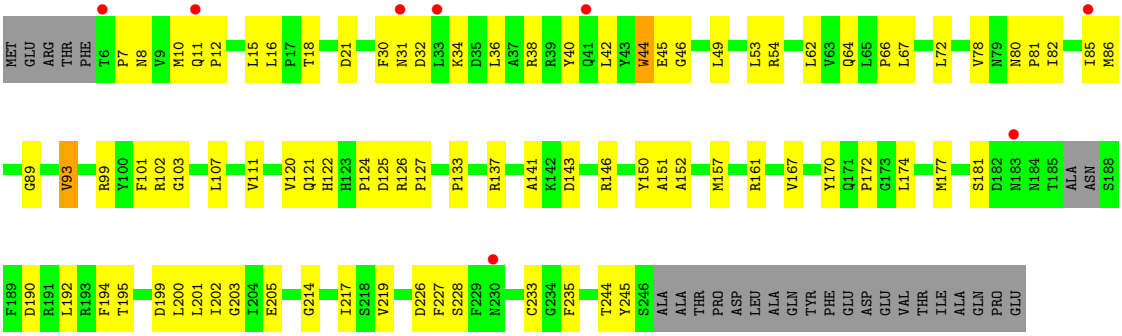
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	411	3179	2056	527	588	8	0	0	0

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.

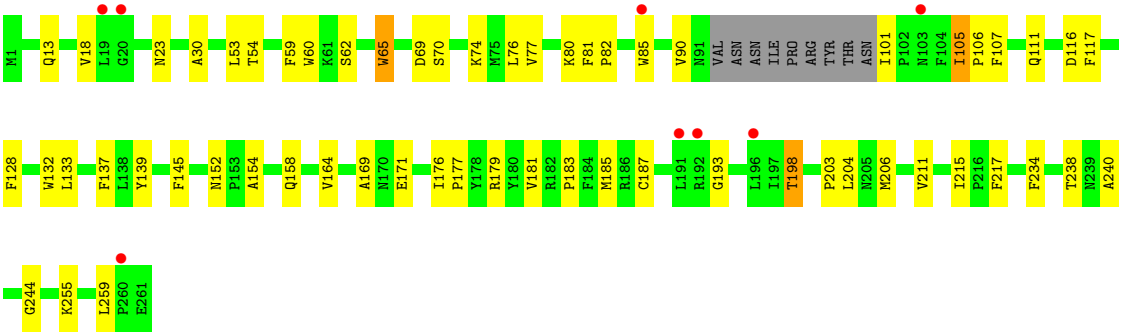
- Molecule 1: VP1

Chain A:



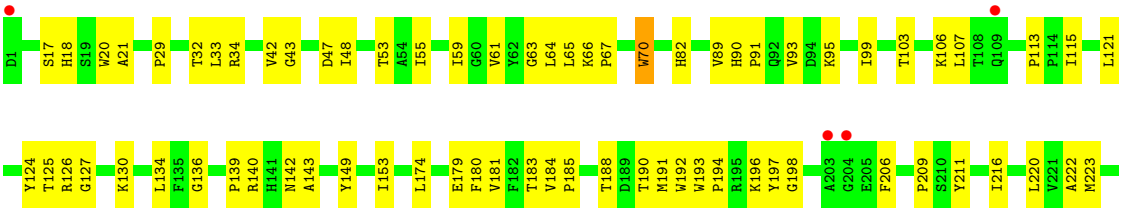
- Molecule 2: VP2

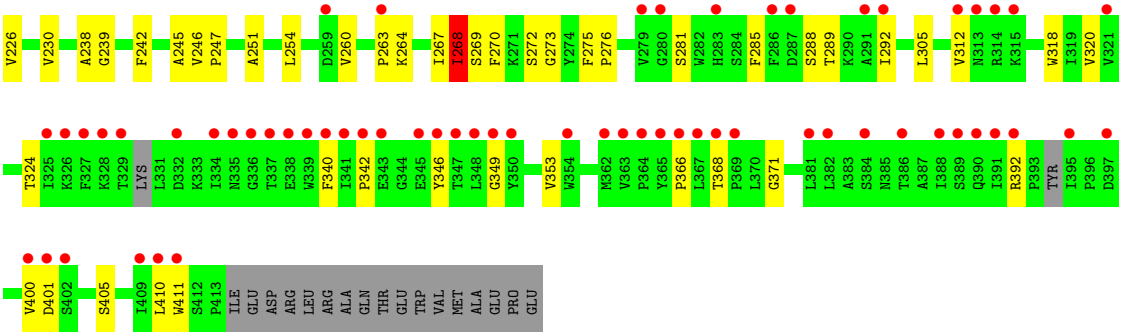
Chain B:



- Molecule 3: Genome polyprotein

Chain C:




V226
V230
A238
G239
F242
A245
V246
P247
A251
L254
D259
V260
P263
K264
I267
I268
S269
F270
K271
S272
G273
Y274
F275
P276
V279
G280
S281
W282
H283
S284
F285
F286
D287
S288
T289
K290
A291
I292
L305
V312
N313
R314
K315
W318
I319
V320
V321
T324
I325
K326
F327
K328
T329
LYS
L331
D332
K333
I334
N335
G336
T337
E338
W339
F340
I341
P342
E343
G344
E345
Y346
T347
L348
G349
Y350
V353
W354
M362
V363
P364
Y365
P366
L367
T368
P369
L370
G371
L381
L382
A383
S384
N385
T386
A387
I388
S389
Q390
I391
R392
P393
TYR
I395
P396
D397
V400
D401
S402
S405
I409
L410
W411
S412
P413
ILE
GLU
ASP
ARG
LEU
ARG
ALA
GLN
THR
GLU
TRP
VAL
MET
ALA
GLU
PRO
GLU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, α, β, γ	360.66Å 360.66Å 360.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	70.70 – 3.41 70.73 – 3.41	Depositor EDS
% Data completeness (in resolution range)	87.4 (70.70-3.41) 87.1 (70.73-3.41)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.20 (at 3.41Å)	Xtriage
Refinement program	PHENIX, CNS 1.3	Depositor
R, R_{free}	0.339 , 0.339 0.322 , 0.323	Depositor DCC
R_{free} test set	4570 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	66.9	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 5.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.026 for -l,-k,-h	Xtriage
F_o,F_c correlation	0.76	EDS
Total number of atoms	7029	wwPDB-VP
Average B, all atoms (Å^2)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.34% 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	A	0.27	0/1948	0.57	1/2655 (0.0%)
2	B	0.25	0/2004	0.49	0/2733
3	C	0.23	0/3274	0.46	1/4483 (0.0%)
All	All	0.25	0/7226	0.50	2/9871 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	264	LYS	N-CA-C	-5.34	96.57	111.00
1	A	30	PHE	N-CA-C	-5.08	97.29	111.00

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	1896	0	1832	67	0
2	B	1954	0	1865	43	0
3	C	3179	0	3107	89	0
All	All	7029	0	6804	163	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 12.

The worst 5 of 163 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)
3:C:91:PRO:HG3	3:C:125:THR:HG21	1.55	0.86
2:B:105:ILE:HG22	2:B:106:PRO:HD2	1.56	0.85
2:B:53:LEU:HB3	2:B:106:PRO:HB3	1.59	0.84
1:A:18:THR:HG21	3:C:185:PRO:HB3	1.62	0.81
2:B:90:VAL:HG11	2:B:255:LYS:HE3	1.65	0.78

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	235/266 (88%)	225 (96%)	10 (4%)	0	100	100
2	B	248/261 (95%)	236 (95%)	12 (5%)	0	100	100
3	C	405/430 (94%)	378 (93%)	25 (6%)	2 (0%)	29	65
All	All	888/957 (93%)	839 (94%)	47 (5%)	2 (0%)	47	80

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	268	ILE
3	C	349	GLY

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	203/230 (88%)	199 (98%)	4 (2%)	55	79
2	B	208/234 (89%)	204 (98%)	4 (2%)	57	80
3	C	344/366 (94%)	341 (99%)	3 (1%)	78	90
All	All	755/830 (91%)	744 (98%)	11 (2%)	65	83

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	198	THR
3	C	70	TRP
3	C	340	PHE
3	C	268	ILE
2	B	60	TRP

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

Mol	Chain	Res	Type
1	A	11	GLN
2	B	152	ASN

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 (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	A	239/266 (89%)	0.39	8 (3%)	46	46	45, 68, 91, 112	0
2	B	252/261 (96%)	0.39	8 (3%)	47	47	44, 63, 100, 109	0
3	C	411/430 (95%)	0.80	66 (16%)	1	2	38, 65, 119, 133	0
All	All	902/957 (94%)	0.58	82 (9%)	9	12	38, 65, 114, 133	0

The worst 5 of 82 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	291	ALA	6.3
3	C	340	PHE	5.7
3	C	335	ASN	5.0
3	C	279	VAL	4.7
2	B	103	ASN	4.7

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.