



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Mar 2, 2017 – 11:22 am GMT

PDB ID : 3IYJ
EMDB ID: : EMD-5155
Title : Bovine papillomavirus type 1 outer capsid
Authors : Wolf, M.; Garcea, R.L.; Grigorieff, N.; Harrison, S.C.
Deposited on : 2009-12-15
Resolution : 4.20 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : recalc29047

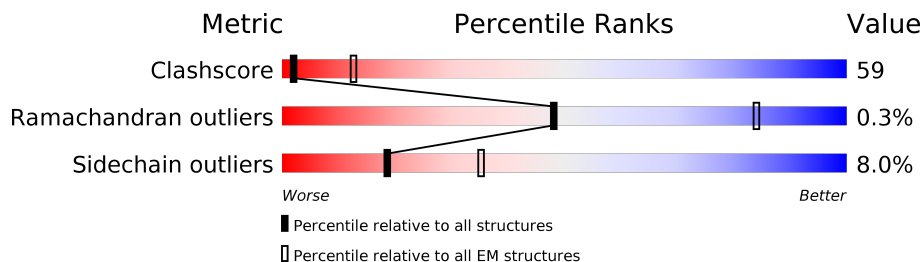
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.20 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	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	495	36% 57% 5% .
1	B	495	34% 55% 5% 5%
1	C	495	35% 57% 5% .
1	D	495	37% 55% 5% .
1	E	495	38% 54% 5% .
1	F	495	37% 54% . 5%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 22614 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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.

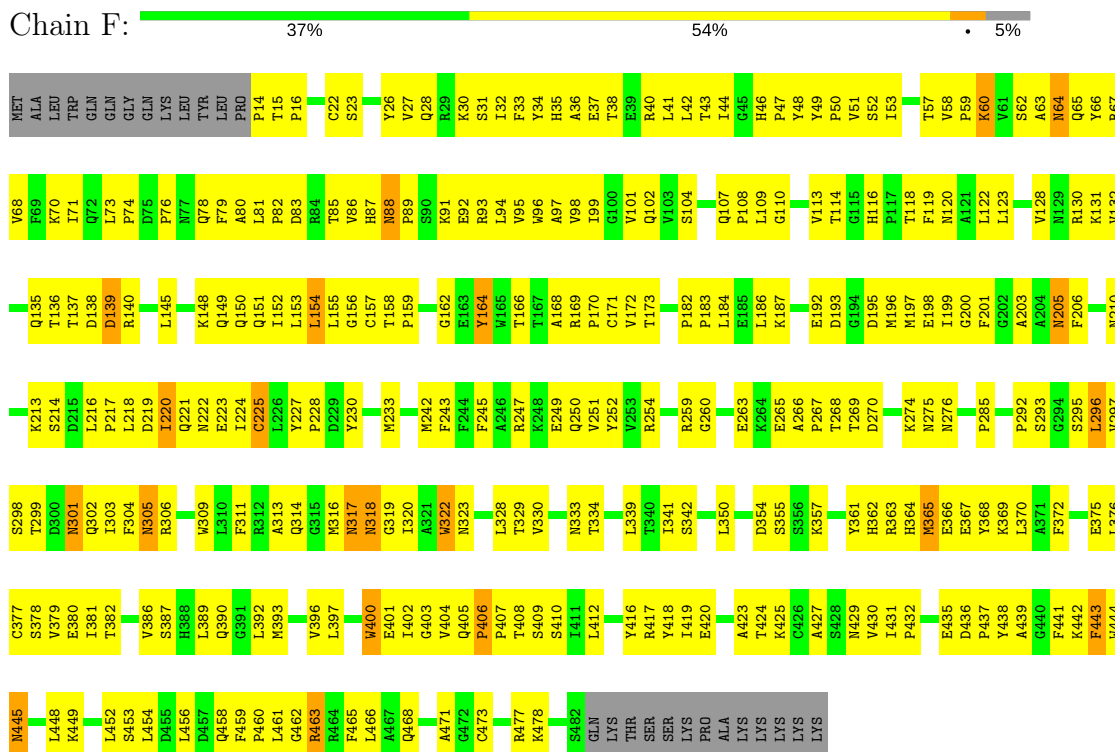
- Molecule 1 is a protein called Major capsid protein L1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	F	469	Total	C	N	O	S	0	0
			3701	2345	639	702	15		
1	A	481	Total	C	N	O	S	0	0
			3803	2414	656	718	15		
1	B	469	Total	C	N	O	S	0	0
			3701	2345	639	702	15		
1	C	481	Total	C	N	O	S	0	0
			3803	2414	656	718	15		
1	D	481	Total	C	N	O	S	0	0
			3803	2414	656	718	15		
1	E	481	Total	C	N	O	S	0	0
			3803	2414	656	718	15		

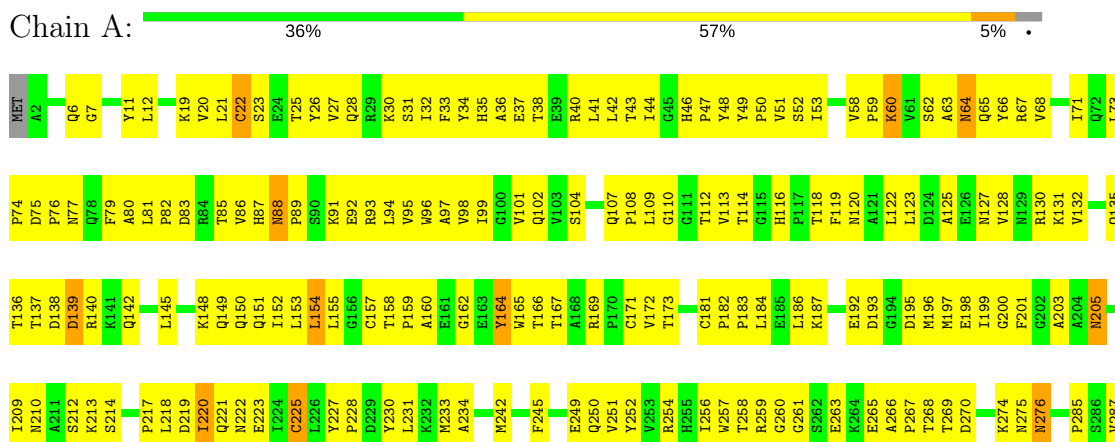
3 Residue-property plots

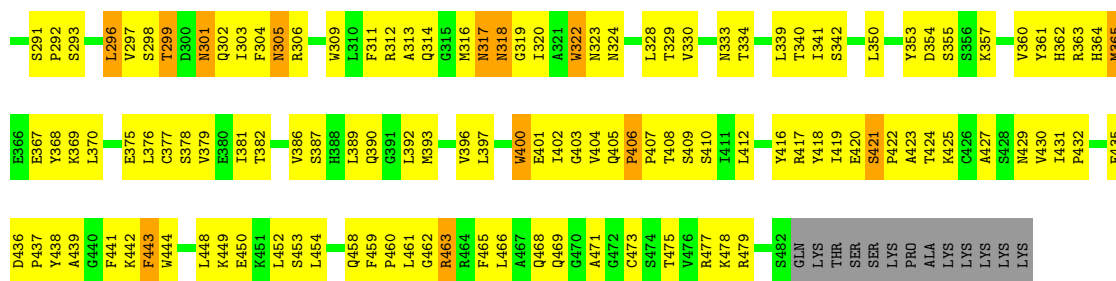
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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. 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: Major capsid protein L1



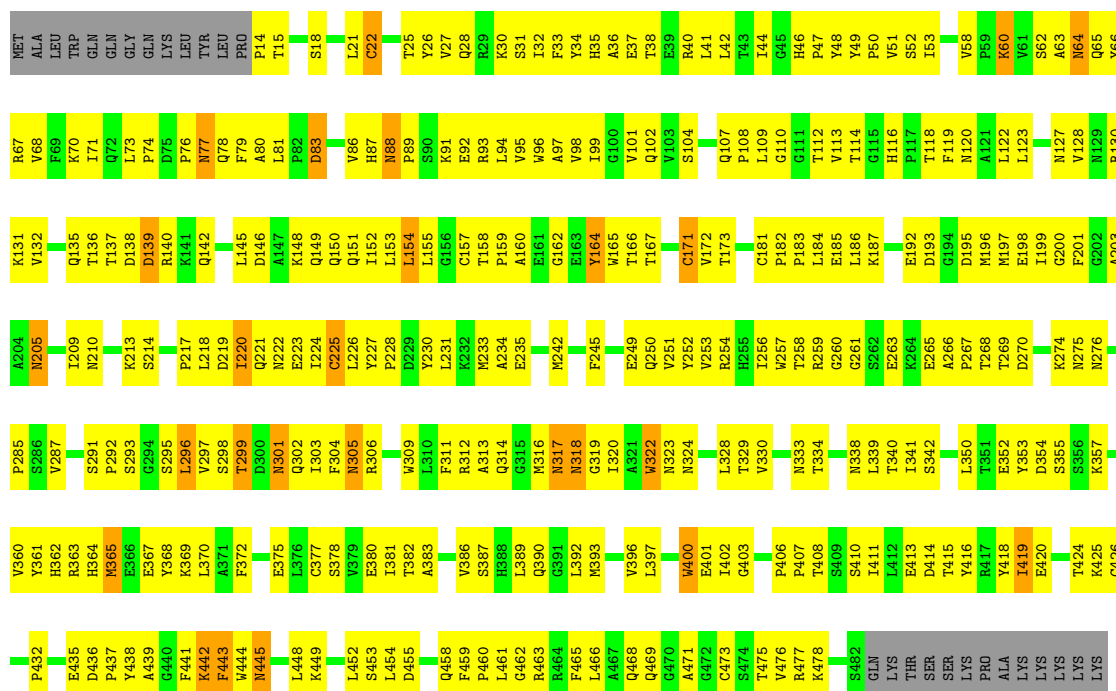
• Molecule 1: Major capsid protein L1





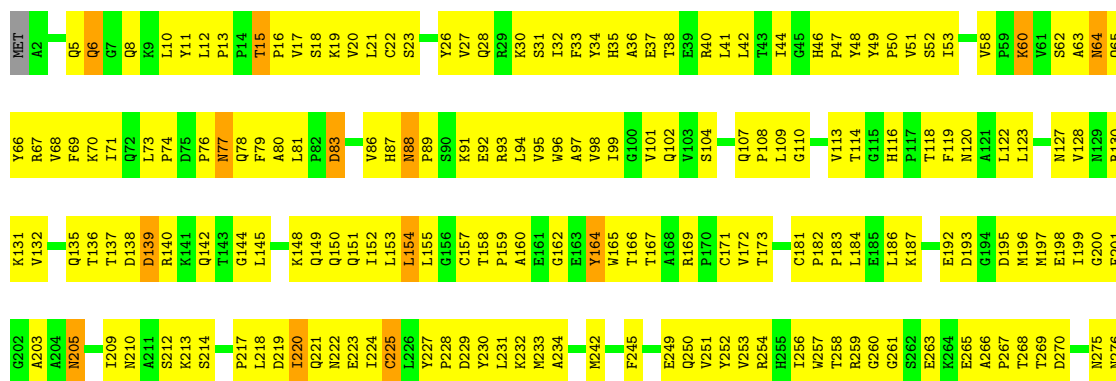
• Molecule 1: Major capsid protein L1

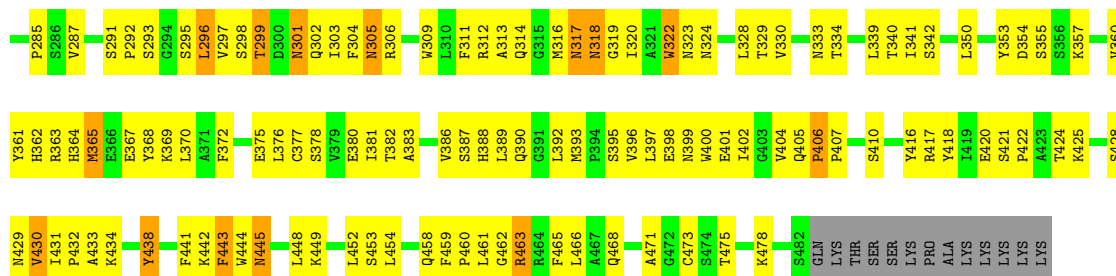
Chain B: 34% 55% 5% 5%



• Molecule 1: Major capsid protein L1

Chain C: 35% 57% 5%





L448	M365	P292
K449	E366	S293
	E367	G294
L452	Y368	G295
S453	K369	L296
L454	L370	V297
D455	A371	S298
L456	F372	T299
		D300
F459	E375	N301
P460	L376	Q302
L461	C377	L303
G462	S378	F304
R463	V379	N305
K464	G380	R306
F465	I381	
L466	T382	
A467	V386	W309
A471	H388	L310
G472	L389	F311
C473	Q390	R312
S474	G391	A313
T475	L392	Q314
V476	G393	G315
R477	P394	M316
K478	S395	N317
	G396	G319
	V397	T320
	L397	A321
	E398	W322
	N399	N323
	W400	N324
	E401	
	Q405	L328
	P406	T329
		V330
	S410	N333
	I411	T334
	L412	
	E413	N338
	D414	L339
	L415	T340
	Y416	L341
		S342
	I419	L350
	E420	
	C426	Y353
	A427	D354
	S428	S355
	P432	S356
		K357
	F441	V360
	K442	Y361
	F443	H362
	W444	R363
	Y445	T364

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	CTFILT3 WITH INDIVIDUAL PARTICLE ADJUSTMENT	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	25.00	Depositor
Minimum defocus (nm)	1800.00	Depositor
Maximum defocus (nm)	2900.00	Depositor
Magnification	56588	Depositor
Image detector	KODAK SO163 FILM	Depositor

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	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.51	0/3896	0.77	4/5296 (0.1%)
1	B	0.50	0/3790	0.74	1/5150 (0.0%)
1	C	0.50	0/3896	0.76	1/5296 (0.0%)
1	D	0.49	0/3896	0.75	1/5296 (0.0%)
1	E	0.50	0/3896	0.76	1/5296 (0.0%)
1	F	0.50	0/3790	0.74	1/5150 (0.0%)
All	All	0.50	0/23164	0.75	9/31484 (0.0%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	7	GLY	N-CA-C	5.75	127.49	113.10
1	E	205	ASN	N-CA-C	-5.55	96.01	111.00
1	A	205	ASN	N-CA-C	-5.49	96.17	111.00
1	B	205	ASN	N-CA-C	-5.46	96.27	111.00
1	C	205	ASN	N-CA-C	-5.45	96.29	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	3803	0	3748	536	0
1	B	3701	0	3644	464	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3803	0	3748	525	0
1	D	3803	0	3748	448	0
1	E	3803	0	3748	451	0
1	F	3701	0	3646	486	0
All	All	22614	0	22282	2640	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 59.

The worst 5 of 2640 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:E:153:LEU:CD1	1:E:330:VAL:HG22	1.43	1.47
1:B:153:LEU:CD1	1:B:330:VAL:HG22	1.43	1.46
1:D:87:HIS:CE1	1:D:94:LEU:HG	1.52	1.44
1:D:88:ASN:ND2	1:D:89:PRO:HD2	1.29	1.44
1:C:153:LEU:CD1	1:C:330:VAL:HG22	1.46	1.43

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 PDB entries followed by that with respect to all EM entries.

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	479/495 (97%)	436 (91%)	42 (9%)	1 (0%)	51	84
1	B	467/495 (94%)	430 (92%)	36 (8%)	1 (0%)	51	84
1	C	479/495 (97%)	435 (91%)	41 (9%)	3 (1%)	28	70
1	D	479/495 (97%)	444 (93%)	35 (7%)	0	100	100
1	E	479/495 (97%)	438 (91%)	39 (8%)	2 (0%)	38	76
1	F	467/495 (94%)	425 (91%)	41 (9%)	1 (0%)	51	84

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	2850/2970 (96%)	2608 (92%)	234 (8%)	8 (0%)	48 81

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	406	PRO
1	A	406	PRO
1	C	406	PRO
1	C	430	VAL
1	E	432	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 PDB entries followed by that with respect to all EM entries.

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	417/430 (97%)	388 (93%)	29 (7%)	18 53
1	B	407/430 (95%)	372 (91%)	35 (9%)	12 45
1	C	417/430 (97%)	382 (92%)	35 (8%)	13 46
1	D	417/430 (97%)	382 (92%)	35 (8%)	13 46
1	E	417/430 (97%)	381 (91%)	36 (9%)	12 45
1	F	407/430 (95%)	379 (93%)	28 (7%)	18 53
All	All	2482/2580 (96%)	2284 (92%)	198 (8%)	18 48

5 of 198 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	443	PHE
1	C	296	LEU
1	E	301	ASN
1	C	6	GLN
1	C	88	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 115 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	318	ASN
1	C	142	GLN
1	E	250	GLN
1	B	338	ASN
1	C	64	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 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.