



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Feb 13, 2017 – 12:06 PM EST

PDB ID : 5KEP
EMDB ID: : EMDB-6620
Title : High resolution cryo-EM maps of Human Papillomavirus 16 reveal L2 location and heparin-induced conformational changes
Authors : Guan, J.; Bywaters, S.M.; Brendle, S.A.; Ashley, R.E.; Makhov, A.M.; Conway, J.F.; Christensen, N.D.; Hafenstein, S.
Deposited on : 2016-06-10
Resolution : 4.30 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
Mogul : unknown
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) : rb-20028442

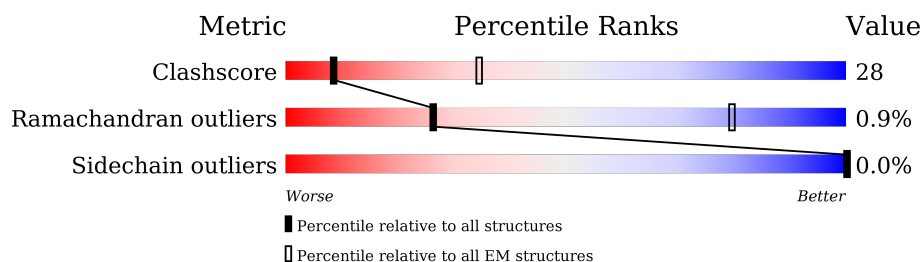
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.30 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	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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	483	47% 49% ..
1	B	483	52% 42% . .
1	C	483	52% 43% ...
1	D	483	54% 41% . .
1	E	483	56% 41% .
1	F	483	55% 41% ..

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 22267 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.

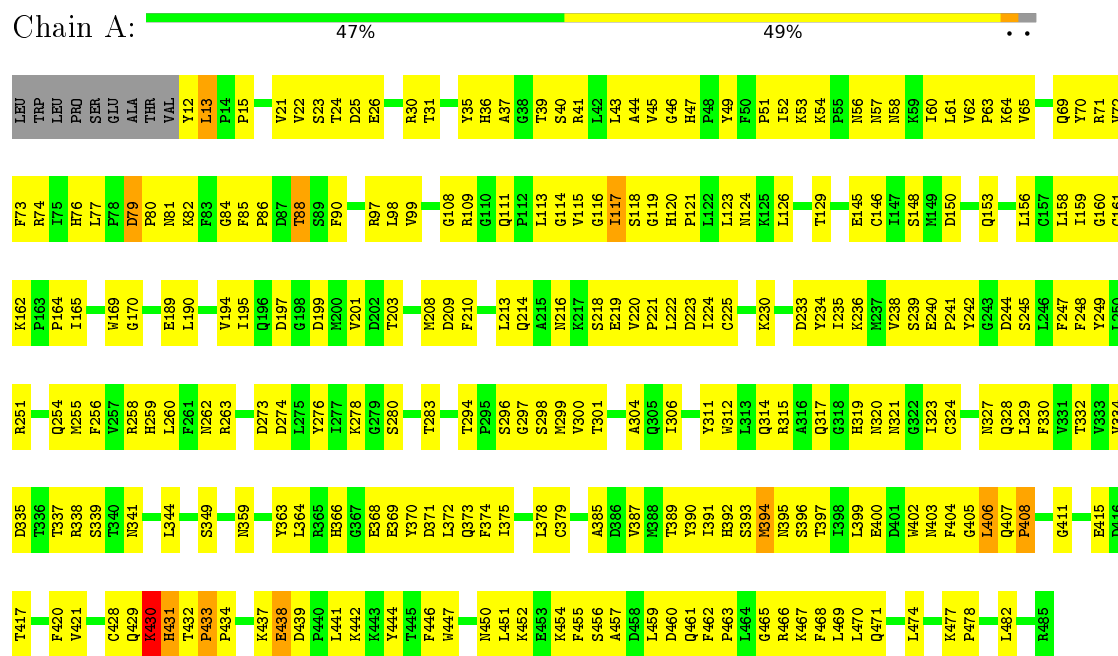
- Molecule 1 is a protein called Major capsid protein L1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	474	Total	C	N	O	S	0	0
			3725	2380	627	697	21		
1	B	463	Total	C	N	O	S	0	0
			3638	2321	612	684	21		
1	C	472	Total	C	N	O	S	0	0
			3705	2368	621	695	21		
1	D	471	Total	C	N	O	S	0	0
			3700	2365	620	694	21		
1	E	482	Total	C	N	O	S	0	0
			3785	2422	633	709	21		
1	F	473	Total	C	N	O	S	0	0
			3714	2373	622	698	21		

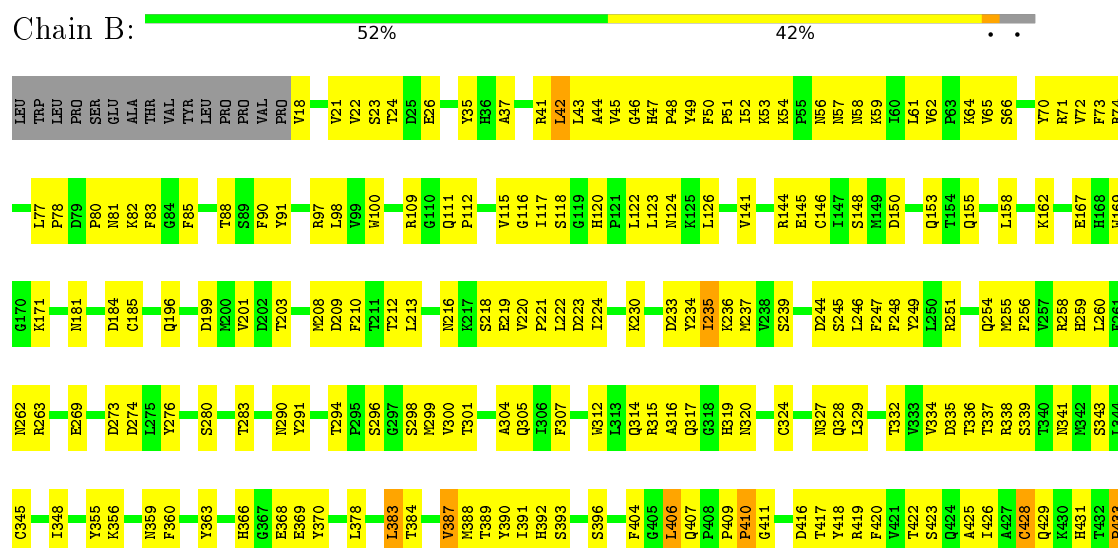
3 Residue-property plots

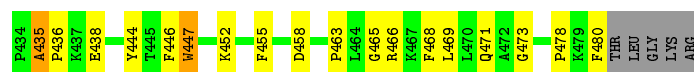
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric 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



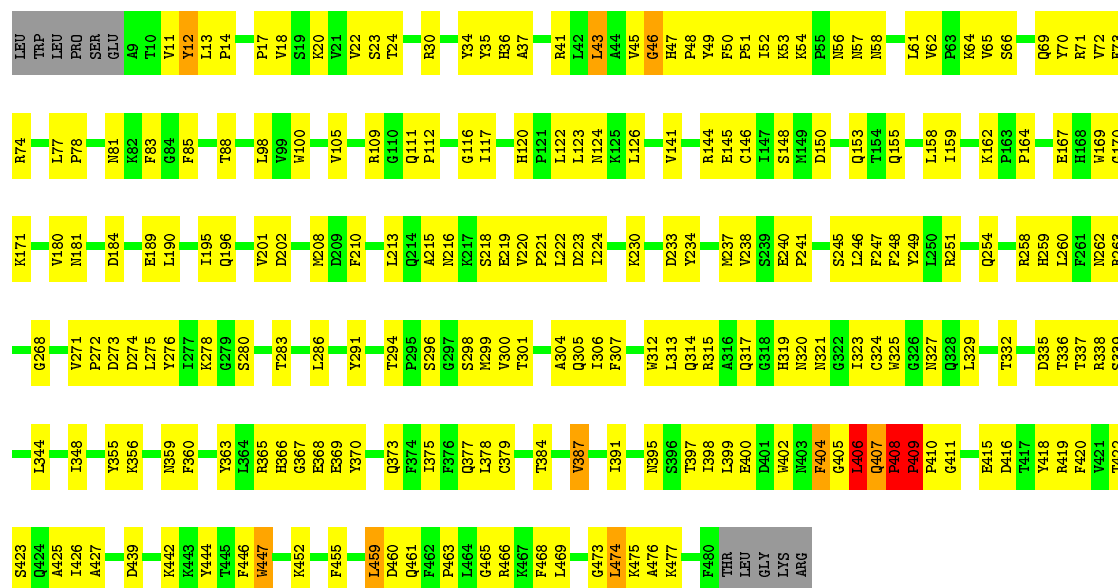
• Molecule 1: Major capsid protein L1





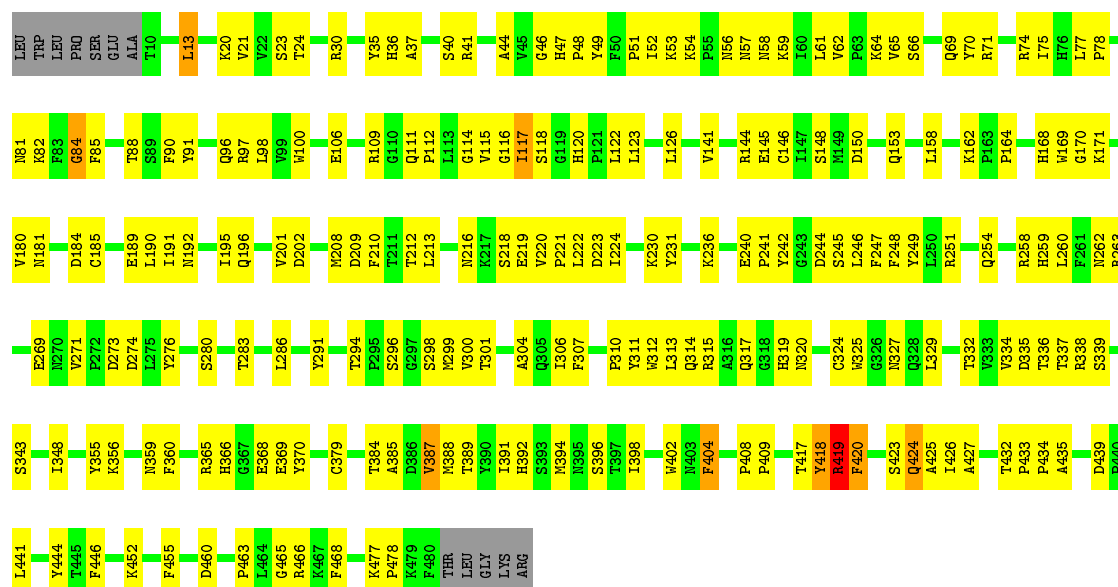
• Molecule 1: Major capsid protein L1

Chain C: 52% 43%



• Molecule 1: Major capsid protein L1

Chain D: 54% 41%



• Molecule 1: Major capsid protein L1

Chain E: 56% 41%

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	57556	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	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	0.50	0/3827	0.66	4/5209 (0.1%)
1	B	0.62	4/3736 (0.1%)	0.69	5/5083 (0.1%)
1	C	0.66	4/3807 (0.1%)	0.70	5/5185 (0.1%)
1	D	0.65	3/3802 (0.1%)	0.73	6/5178 (0.1%)
1	E	0.63	1/3890 (0.0%)	0.65	5/5299 (0.1%)
1	F	0.60	2/3816 (0.1%)	0.68	6/5197 (0.1%)
All	All	0.61	14/22878 (0.1%)	0.69	31/31151 (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	A	0	11
1	B	0	6
1	C	0	4
1	D	0	7
1	E	0	11
1	F	0	3
All	All	0	42

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	435	ALA	C-N	-10.16	1.15	1.34
1	F	404	PHE	CE1-CZ	-9.70	1.19	1.37
1	C	46	GLY	C-N	-8.66	1.14	1.34
1	B	447	TRP	CB-CG	-8.18	1.35	1.50
1	E	407	GLN	C-O	7.07	1.36	1.23

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	D	419	ARG	NE-CZ-NH2	17.08	128.84	120.30
1	B	42	LEU	CA-CB-CG	15.47	150.89	115.30
1	B	383	LEU	CB-CG-CD1	-10.49	93.16	111.00
1	E	405	GLY	N-CA-C	-9.03	90.52	113.10
1	F	428	CYS	CA-CB-SG	-8.46	98.76	114.00

There are no chirality outliers.

5 of 42 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	13	LEU	Peptide
1	A	15	PRO	Peptide
1	A	408	PRO	Peptide
1	A	415	GLU	Peptide
1	A	88	THR	Peptide

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	3725	0	3648	231	0
1	B	3638	0	3551	221	0
1	C	3705	0	3621	251	0
1	D	3700	0	3616	201	0
1	E	3785	0	3706	211	0
1	F	3714	0	3628	215	0
All	All	22267	0	21770	1225	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 28.

The worst 5 of 1225 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (\AA)	Clash overlap (\AA)
1:C:404:PHE:HB2	1:C:405:GLY:HA2	1.34	1.08
1:C:400:GLU:HA	1:C:404:PHE:CZ	1.97	1.00

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:97:ARG:HH22	1:D:404:PHE:HB2	1.26	0.98
1:A:46:GLY:HA3	1:A:65:VAL:HB	1.47	0.95
1:B:44:ALA:HB3	1:B:368:GLU:HB2	1.46	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 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	472/483 (98%)	431 (91%)	38 (8%)	3 (1%)	30	74
1	B	461/483 (95%)	430 (93%)	28 (6%)	3 (1%)	26	71
1	C	470/483 (97%)	431 (92%)	33 (7%)	6 (1%)	15	60
1	D	469/483 (97%)	430 (92%)	35 (8%)	4 (1%)	21	67
1	E	480/483 (99%)	441 (92%)	35 (7%)	4 (1%)	24	70
1	F	471/483 (98%)	440 (93%)	25 (5%)	6 (1%)	15	60
All	All	2823/2898 (97%)	2603 (92%)	194 (7%)	26 (1%)	26	67

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	408	PRO
1	A	117	ILE
1	C	406	LEU
1	D	117	ILE
1	F	422	THR

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 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	412/420 (98%)	411 (100%)	1 (0%)	95	97
1	B	402/420 (96%)	402 (100%)	0	100	100
1	C	410/420 (98%)	410 (100%)	0	100	100
1	D	410/420 (98%)	410 (100%)	0	100	100
1	E	419/420 (100%)	419 (100%)	0	100	100
1	F	411/420 (98%)	411 (100%)	0	100	100
All	All	2464/2520 (98%)	2463 (100%)	1 (0%)	100	100

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	406	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	D	1
1	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	46:GLY	C	47:HIS	N	1.14
1	D	435:ALA	C	436:PRO	N	1.14