wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:47 PM BST

PDB ID	:	3J6R
EMDB ID:	:	EMD-5932
Title	:	Electron cryo-microscopy of Human Papillomavirus Type 16 capsid
Authors	:	Cardone, G.; Moyer, A.L.; Cheng, N.; Thompson, C.D.; Dvoretzky, I.; Lowy, D.R.; Schiller, J.T.; Steven, A.C.; Buck, C.B.; Trus, B.L.
Deposited on	:	2014-03-20
Resolution	:	9.10 Å(reported)
Based on PDB ID	:	1DZL

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

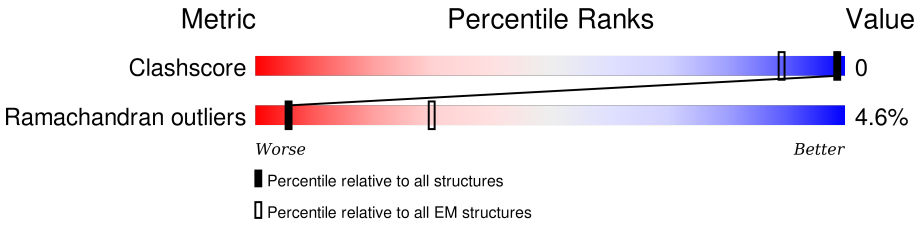
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 9.10 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

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	478	91% 8% •
1	B	478	90% 8% •
1	C	478	89% 10% •
1	D	478	90% 8% •
1	E	478	92% 8%
1	F	478	91% 8% •

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 17214 atoms, of which 5748 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	B	478	Total	C	H	N	O	0	0
			2869	956	958	478	477		
1	A	478	Total	C	H	N	O	0	0
			2869	956	958	478	477		
1	E	478	Total	C	H	N	O	0	0
			2869	956	958	478	477		
1	D	478	Total	C	H	N	O	0	0
			2869	956	958	478	477		
1	C	478	Total	C	H	N	O	0	0
			2869	956	958	478	477		
1	F	478	Total	C	H	N	O	0	0
			2869	956	958	478	477		

There are 18 discrepancies between the modelled and reference sequences:

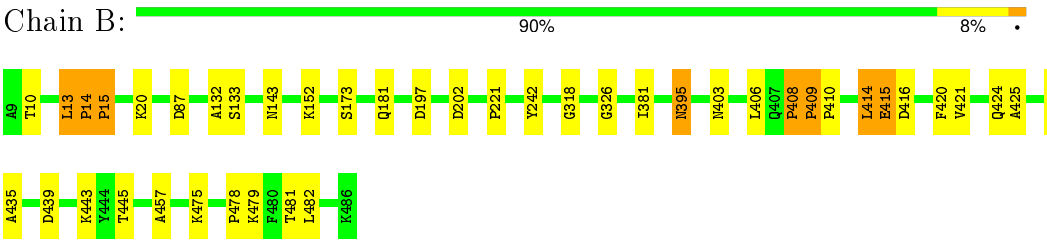
Chain	Residue	Modelled	Actual	Comment	Reference
B	177	GLN	ASN	CONFLICT	UNP Q4VRM0
B	181	GLN	ASN	CONFLICT	UNP Q4VRM0
B	472	LEU	ALA	CONFLICT	UNP Q4VRM0
A	177	GLN	ASN	CONFLICT	UNP Q4VRM0
A	181	GLN	ASN	CONFLICT	UNP Q4VRM0
A	472	LEU	ALA	CONFLICT	UNP Q4VRM0
E	177	GLN	ASN	CONFLICT	UNP Q4VRM0
E	181	GLN	ASN	CONFLICT	UNP Q4VRM0
E	472	LEU	ALA	CONFLICT	UNP Q4VRM0
D	177	GLN	ASN	CONFLICT	UNP Q4VRM0
D	181	GLN	ASN	CONFLICT	UNP Q4VRM0
D	472	LEU	ALA	CONFLICT	UNP Q4VRM0
C	177	GLN	ASN	CONFLICT	UNP Q4VRM0
C	181	GLN	ASN	CONFLICT	UNP Q4VRM0
C	472	LEU	ALA	CONFLICT	UNP Q4VRM0
F	177	GLN	ASN	CONFLICT	UNP Q4VRM0
F	181	GLN	ASN	CONFLICT	UNP Q4VRM0
F	472	LEU	ALA	CONFLICT	UNP Q4VRM0

3 Residue-property plots (i)

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.

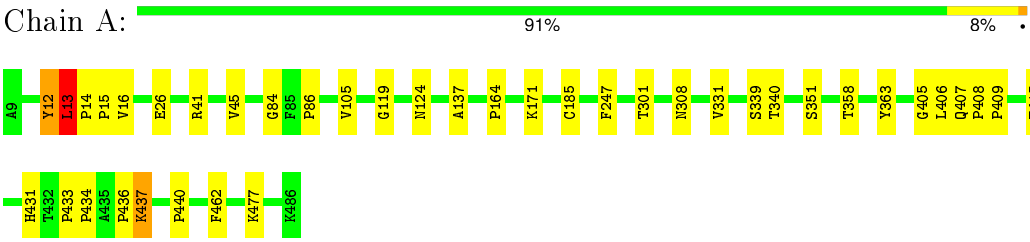
- Molecule 1: Major capsid protein L1

Chain B:



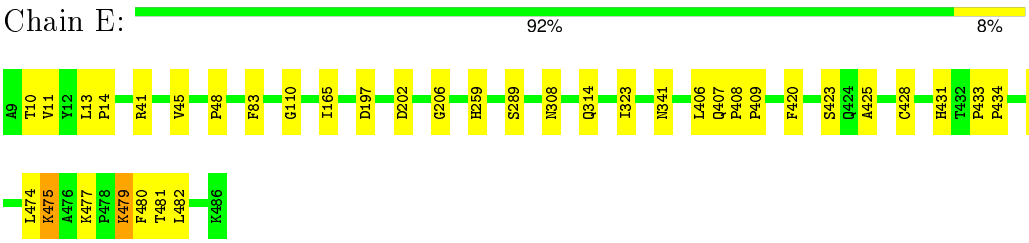
- Molecule 1: Major capsid protein L1

Chain A:



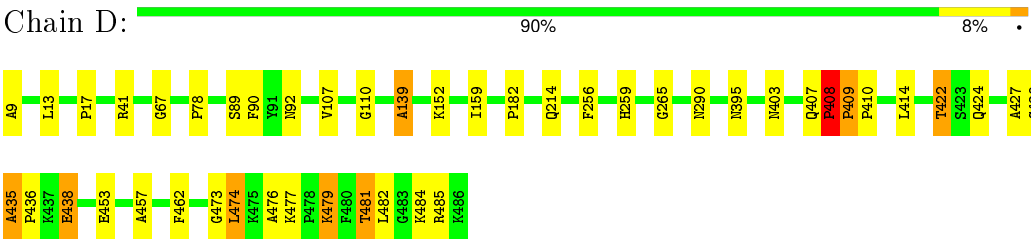
- Molecule 1: Major capsid protein L1

Chain E:



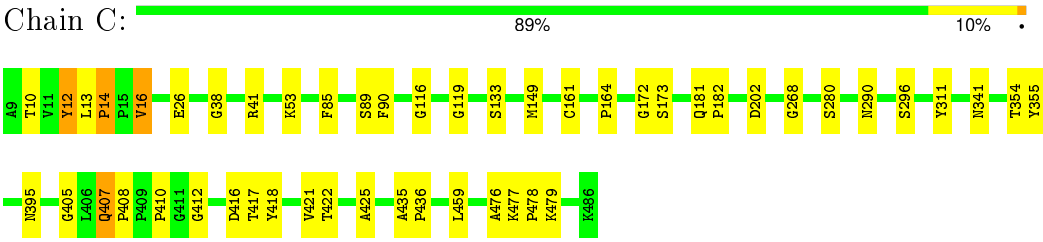
- Molecule 1: Major capsid protein L1

Chain D:



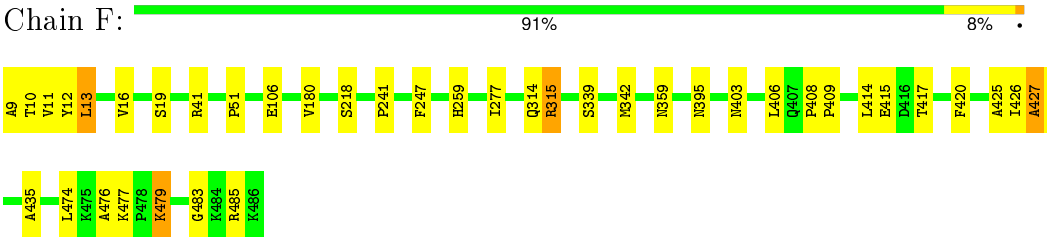
- Molecule 1: Major capsid protein L1

Chain C:



• Molecule 1: Major capsid protein L1

Chain F:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	5952	Depositor
Resolution determination method	FSC 0.333	Depositor
CTF correction method	each micrograph	Depositor
Microscope	FEI/PHILIPS CM200FEG	Depositor
Voltage (kV)	120	Depositor
Electron dose ($e^-/Å^2$)	Not provided	Depositor
Minimum defocus (nm)	537	Depositor
Maximum defocus (nm)	2175	Depositor
Magnification	38000	Depositor
Image detector	KODAK SO-163 FILM	Depositor

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	#$\|Z\|$ >2	RMSZ	#$\|Z\|$ >2
1	A	1.50	3/1910 (0.2%)	1.73	11/2386 (0.5%)
1	B	1.52	1/1910 (0.1%)	1.76	10/2386 (0.4%)
1	C	1.55	11/1910 (0.6%)	1.72	11/2386 (0.5%)
1	D	1.58	7/1910 (0.4%)	1.75	13/2386 (0.5%)
1	E	1.58	5/1910 (0.3%)	1.72	16/2386 (0.7%)
1	F	1.51	4/1910 (0.2%)	1.73	17/2386 (0.7%)
All	All	1.54	31/11460 (0.3%)	1.74	78/14316 (0.5%)

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	13
1	B	0	10
1	C	0	13
1	D	0	15
1	E	0	9
1	F	0	11
All	All	0	71

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	206	GLY	N-CA	-6.92	1.35	1.46
1	C	290	ASN	N-CA	-6.28	1.33	1.46
1	A	164	PRO	N-CA	-6.21	1.36	1.47
1	E	482	LEU	C-N	6.17	1.44	1.33
1	E	165	ILE	C-N	6.10	1.44	1.33

The worst 5 of 78 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	420	PHE	C-N-CA	8.66	143.34	121.70
1	B	15	PRO	C-N-CA	8.18	142.14	121.70
1	F	485	ARG	C-N-CA	8.07	141.88	121.70
1	F	427	ALA	O-C-N	-8.06	109.81	122.70
1	C	354	THR	O-C-N	7.33	134.42	122.70

There are no chirality outliers.

5 of 71 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	13	LEU	Peptide
1	B	14	PRO	Peptide
1	B	181	GLN	Peptide
1	B	408	PRO	Peptide
1	B	409	PRO	Peptide

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	1911	958	512	1	0
1	B	1911	958	512	1	0
1	C	1911	958	512	0	0
1	D	1911	958	512	0	0
1	E	1911	958	512	0	0
1	F	1911	958	512	1	0
All	All	11466	5748	3072	3	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 0.

All (3) 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:F:427:ALA:H	1:F:428:CYS:CA	2.27	0.48
1:A:12:TYR:H	1:A:13:LEU:CA	2.31	0.42
1:B:242:TYR:O	1:B:318:GLY:HA3	2.20	0.41

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 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	476/478 (100%)	403 (85%)	54 (11%)	19 (4%)	4	35
1	B	476/478 (100%)	395 (83%)	50 (10%)	31 (6%)	1	25
1	C	476/478 (100%)	401 (84%)	53 (11%)	22 (5%)	3	33
1	D	476/478 (100%)	404 (85%)	47 (10%)	25 (5%)	2	29
1	E	476/478 (100%)	417 (88%)	43 (9%)	16 (3%)	5	40
1	F	476/478 (100%)	405 (85%)	52 (11%)	19 (4%)	4	35
All	All	2856/2868 (100%)	2425 (85%)	299 (10%)	132 (5%)	5	33

5 of 132 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	15	PRO
1	B	87	ASP
1	B	133	SER
1	B	395	ASN
1	B	408	PRO

5.3.2 Protein sidechains

There are no protein residues with a non-rotameric sidechain to report in this entry.

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 carbohydrates 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.