



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Feb 8, 2018 – 07:59 AM EST

PDB ID : 6BVN
EMDB ID: : EMD-7295
Title : Cryo-EM Structure of Hepatitis B virus T=3 capsid in complex with the fluorescent allosteric modulator HAP-TAMRA
Authors : Schlicksup, C.; Wang, J.C.; Zlotnick, A.
Deposited on : 2017-12-13
Resolution : 4.00 Å(reported)
Based on PDB ID : 5D7Y

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 : 1.7.2 (RC1), CSD as538be (2017)
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) : rb-20030736

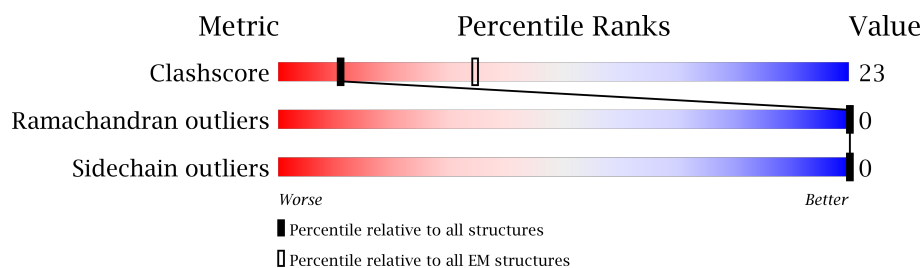
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.00 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	150	61% 35% .
1	B	150	55% 41% . .
1	C	150	58% 38% .

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3517 atoms, of which 17 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.

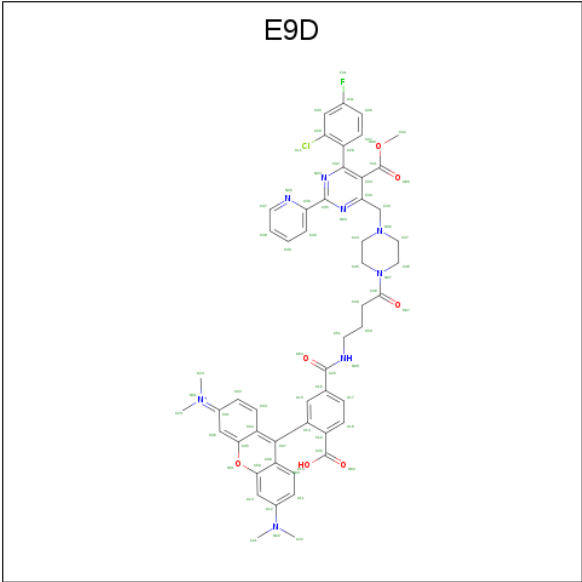
- Molecule 1 is a protein called Capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	C	144	Total	C	N	O	S	0	0
			1144	746	188	208	2		
1	B	144	Total	C	N	O	S	0	0
			1144	746	188	208	2		
1	A	144	Total	C	N	O	S	0	0
			1144	746	188	208	2		

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	48	ALA	CYS	engineered mutation	UNP P03147
C	61	ALA	CYS	engineered mutation	UNP P03147
C	107	ALA	CYS	engineered mutation	UNP P03147
C	150	CYS	-	expression tag	UNP P03147
B	48	ALA	CYS	engineered mutation	UNP P03147
B	61	ALA	CYS	engineered mutation	UNP P03147
B	107	ALA	CYS	engineered mutation	UNP P03147
B	150	CYS	-	expression tag	UNP P03147
A	48	ALA	CYS	engineered mutation	UNP P03147
A	61	ALA	CYS	engineered mutation	UNP P03147
A	107	ALA	CYS	engineered mutation	UNP P03147
A	150	CYS	-	expression tag	UNP P03147

- Molecule 2 is Heteroaryldihydropyrimidine tetramethylrodamine (three-letter code: E9D) (formula: C₅₁H₄₉ClFN₈O₇).

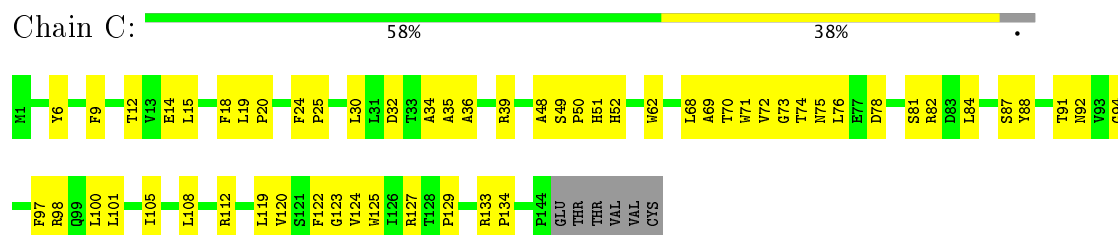


Mol	Chain	Residues	Atoms							AltConf
			Total	C	Cl	F	H	N	O	
2	B	1	85	51	1	1	17	8	7	0

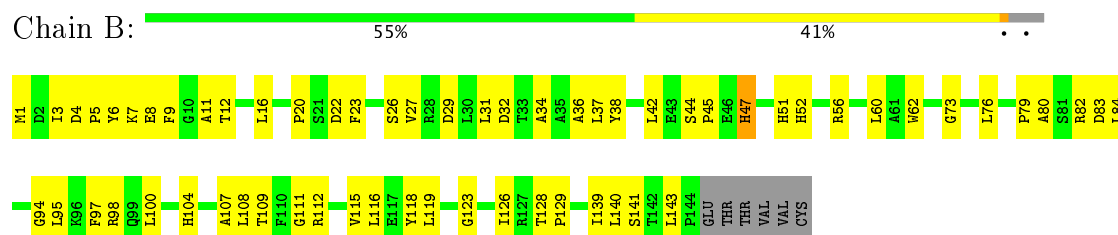
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

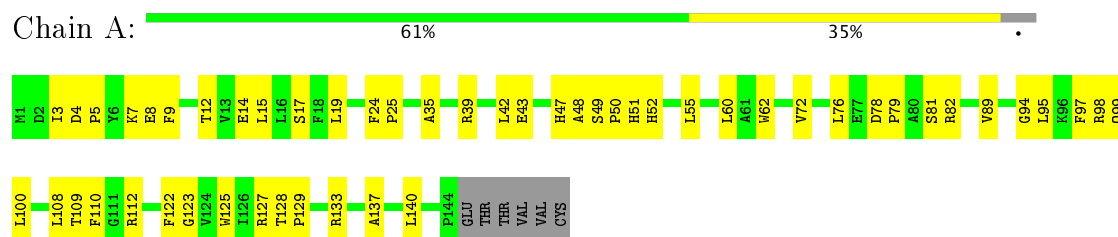
- Molecule 1: Capsid protein



- Molecule 1: Capsid protein



- Molecule 1: Capsid protein



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	16008	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	33	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: E9D

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	0.36	0/1180	0.43	0/1620
1	B	0.38	0/1180	0.47	0/1620
1	C	0.37	0/1180	0.43	0/1620
All	All	0.37	0/3540	0.44	0/4860

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	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	47	HIS	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	1144	0	1128	49	0
1	B	1144	0	1128	72	0
1	C	1144	0	1128	55	0
2	B	68	17	0	2	0
All	All	3500	17	3384	158	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 23.

The worst 5 of 158 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:25:PRO:HG2	1:C:30:LEU:HD12	1.55	0.86
1:C:81:SER:HA	1:C:84:LEU:HD12	1.63	0.79
1:B:3:ILE:HG12	1:A:43:GLU:HG2	1.64	0.78
1:C:34:ALA:HB2	1:C:105:ILE:HD13	1.65	0.78
1:A:4:ASP:HB3	1:A:7:LYS:HB2	1.68	0.75

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	142/150 (95%)	138 (97%)	4 (3%)	0	100	100
1	B	142/150 (95%)	135 (95%)	7 (5%)	0	100	100
1	C	142/150 (95%)	134 (94%)	8 (6%)	0	100	100
All	All	426/450 (95%)	407 (96%)	19 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	124/130 (95%)	124 (100%)	0	100	100
1	B	124/130 (95%)	124 (100%)	0	100	100
1	C	124/130 (95%)	124 (100%)	0	100	100
All	All	372/390 (95%)	372 (100%)	0	100	100

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

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	52	HIS

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](#)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The

Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	E9D	B	201	-	69,75,75	2.64	22 (31%)	92,108,108	2.07	28 (30%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	E9D	B	201	-	-	0/41/63/63	0/8/8/8

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	201	E9D	C43-N06	-7.66	1.32	1.47
2	B	201	E9D	C47-N06	-6.01	1.30	1.47
2	B	201	E9D	C44-N06	-5.45	1.32	1.47
2	B	201	E9D	C32-N04	-3.08	1.29	1.34
2	B	201	E9D	C33-C32	-3.01	1.35	1.41

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	201	E9D	C44-C45-N07	-5.13	98.76	110.44
2	B	201	E9D	C30-C29-C28	-4.78	118.56	121.83
2	B	201	E9D	C18-C19-C20	-4.78	112.88	120.23
2	B	201	E9D	C27-C28-C32	-4.34	111.90	119.56
2	B	201	E9D	O01-C09-C13	-3.88	111.56	116.11

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	201	E9D	2	0

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.