W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB EM Validation Summary Report (i)

May 11, 2021 – 03:01 pm BST

PDB ID	:	7OEN
EMDB ID	:	EMD-12825
Title	:	Hepatitis B core protein mutant P5T with bound GSLLGRMKGA
Authors	:	Bottcher, B.; Makbul, C.
Deposited on	:	2021-05-03
Resolution	:	3.20 Å(reported)

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 3.20 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	Percentile Ranks	Value
Clashscore		7
Ramachandran outliers		0
Sidechain outliers		0.4%

Worse   *Better*

▌ Percentile relative to all structures

▯ Percentile relative to all EM structures
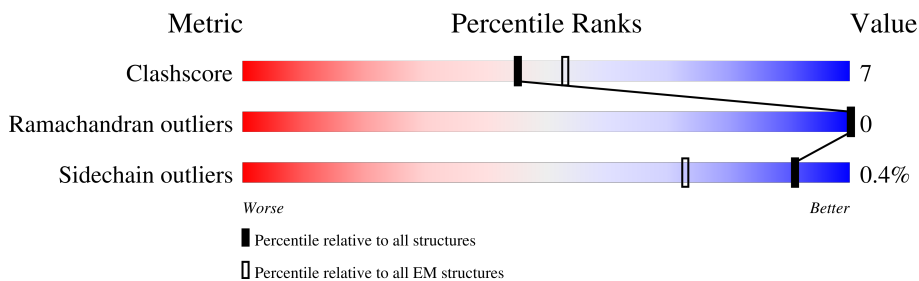
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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%

Mol	Chain	Length	Quality of chain
1	A	183	63% 15% 22%
1	B	183	67% 15% 18%
1	C	183	64% 14% 21%
1	D	183	65% 14% 21%
2	E	20	35% 65%
2	F	20	30% 70%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4704 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 Capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	B	150	Total	C	N	O	S	0	0
			1199	777	197	219	6		
1	A	143	Total	C	N	O	S	0	0
			1144	743	187	208	6		
1	C	144	Total	C	N	O	S	0	0
			1151	748	188	209	6		
1	D	144	Total	C	N	O	S	0	0
			1145	745	185	209	6		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	5	THR	PRO	engineered mutation	UNP P03146
A	5	THR	PRO	engineered mutation	UNP P03146
C	5	THR	PRO	engineered mutation	UNP P03146
D	5	THR	PRO	engineered mutation	UNP P03146

- Molecule 2 is a protein called GSLLGRMKGA.

| Mol	Chain	Residues	Atoms				AltConf	Trace
2	E	7	Total	C	N	O	0	0
			35	21	7	7		
2	F	6	Total	C	N	O	0	0
			30	18	6	6		

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. 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.
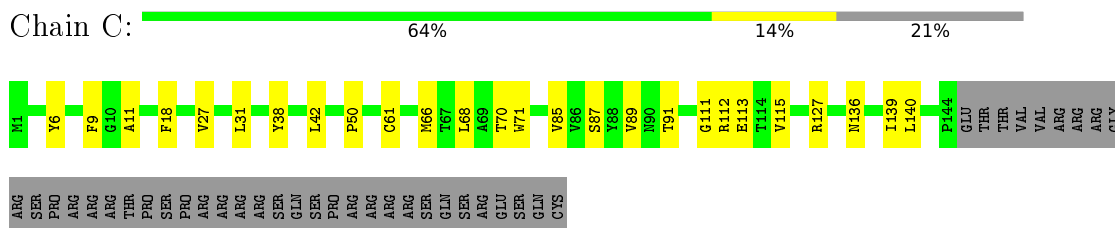
- Molecule 1: Capsid protein
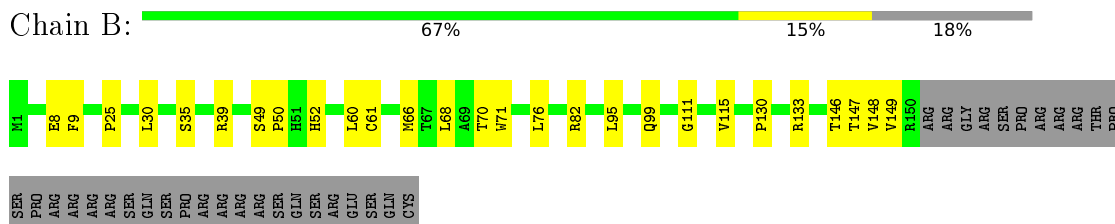
Chain B:

- Molecule 1: Capsid protein
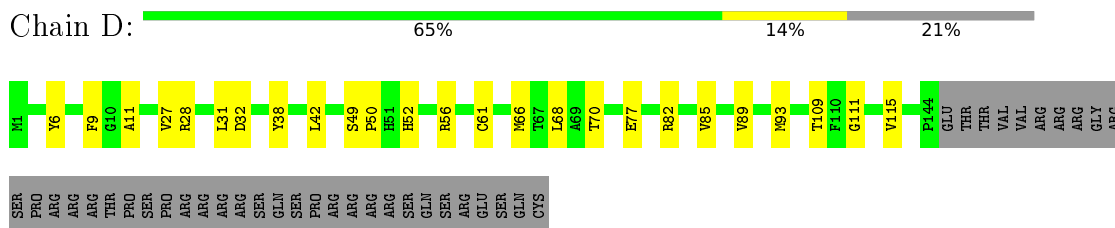
Chain A:
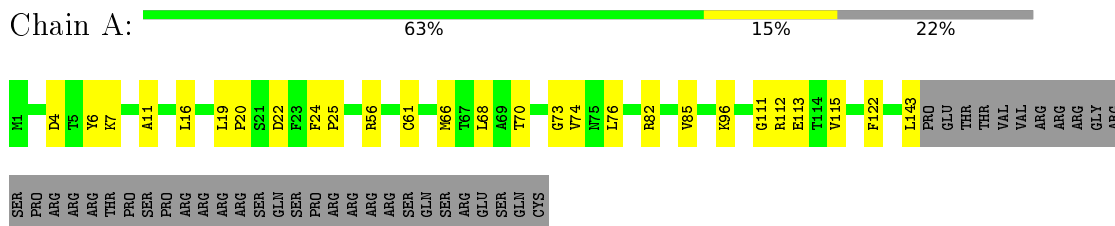
- Molecule 1: Capsid protein

Chain C:

- Molecule 1: Capsid protein

Chain D:

- Molecule 2: GSLLGRMKGA

Chain E:

35%   65%

X4 X10 UNK UNK UNK GLY SER LEU LEU GLY ARG MET LYS GLY ALA

● Molecule 2: GSLLGRMKGA

Chain F:

30%   70%

X4 X9 UNK UNK UNK GLY SER LEU LEU GLY ARG MET LYS GLY ALA

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	74555	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	49	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	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	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.54	0/1178	0.53	0/1612
1	B	0.57	0/1234	0.54	0/1690
1	C	0.54	0/1186	0.48	0/1624
1	D	0.53	0/1180	0.51	0/1617
All	All	0.55	0/4778	0.52	0/6543

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1144	0	1127	19	0
1	B	1199	0	1185	19	0
1	C	1151	0	1134	18	0
1	D	1145	0	1123	19	0
2	E	35	0	11	0	0
2	F	30	0	8	0	0
All	All	4704	0	4588	62	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 7.

The worst 5 of 62 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:A:66:MET:O	1:A:70:THR:HG23	1.87	0.74
1:C:112:ARG:NH2	1:C:113:GLU:OE2	2.22	0.72
1:D:49:SER:OG	1:D:52:HIS:ND1	2.23	0.71
1:C:66:MET:O	1:C:70:THR:HG23	1.94	0.68
1:C:71:TRP:CZ3	1:D:85:VAL:HG12	2.32	0.65

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	141/183 (77%)	138 (98%)	3 (2%)	0	100	100
1	B	148/183 (81%)	145 (98%)	3 (2%)	0	100	100
1	C	142/183 (78%)	136 (96%)	6 (4%)	0	100	100
1	D	142/183 (78%)	136 (96%)	6 (4%)	0	100	100
All	All	573/732 (78%)	555 (97%)	18 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	127/166 (76%)	125 (98%)	2 (2%)	62	84
1	B	134/166 (81%)	134 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	128/166 (77%)	128 (100%)	0	100	100
1	D	127/166 (76%)	127 (100%)	0	100	100
All	All	516/664 (78%)	514 (100%)	2 (0%)	91	95

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	16	LEU
1	A	22	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	47	HIS
1	D	75	ASN
1	D	99	GLN

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.