



wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 06:45 am BST

PDB ID : 6CWT
Title : Hepatitis B core-antigen in complex with Fab e21
Authors : Eren, E.; Steven, A.C.; Wingfield, P.T.
Deposited on : 2018-03-30
Resolution : 3.15 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

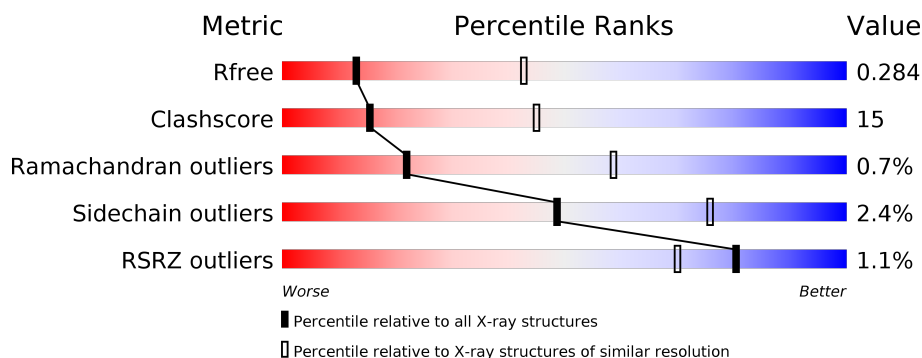
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.15 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1665 (3.20-3.12)
Clashscore	141614	1804 (3.20-3.12)
Ramachandran outliers	138981	1770 (3.20-3.12)
Sidechain outliers	138945	1769 (3.20-3.12)
RSRZ outliers	127900	1616 (3.20-3.12)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	233	<div> <div>%</div> <div> <div></div> <div>70%</div> <div>21%</div> <div>8%</div> </div> </div>
1	C	233	<div> <div>%</div> <div> <div></div> <div>63%</div> <div>22%</div> <div>14%</div> </div> </div>
2	B	216	<div> <div></div> <div> <div></div> <div>63%</div> <div>33%</div> <div>..</div> </div> </div>
2	D	216	<div> <div></div> <div> <div></div> <div>58%</div> <div>37%</div> <div>..</div> </div> </div>
3	E	149	<div> <div>%</div> <div> <div></div> <div>57%</div> <div>25%</div> <div>17%</div> </div> </div>
3	F	149	<div> <div>2%</div> <div> <div></div> <div>62%</div> <div>18%</div> <div>..</div> <div>19%</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8078 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 Fab e21 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	214	Total	C	N	O	S	0	0	0
			1557	982	260	310	5			
1	C	200	Total	C	N	O	S	0	0	0
			1466	928	244	289	5			

- Molecule 2 is a protein called Fab e21 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	214	Total	C	N	O	S	0	0	0
			1588	985	270	328	5			
2	D	208	Total	C	N	O	S	0	0	0
			1541	952	263	321	5			

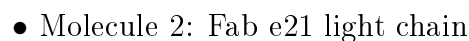
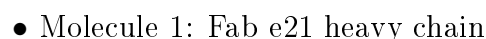
- Molecule 3 is a protein called Capsid protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	124	Total	C	N	O	S	0	0	0
			969	637	160	170	2			
3	F	121	Total	C	N	O	S	0	0	0
			957	629	157	169	2			

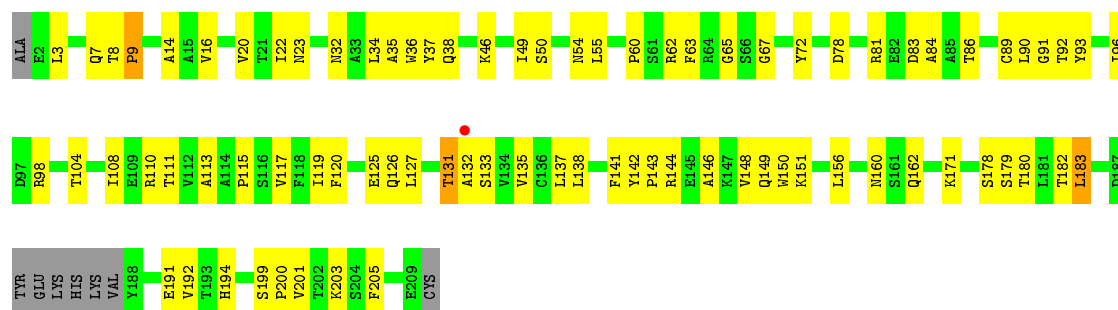
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	48	ALA	CYS	engineered mutation	UNP P03147
E	107	ALA	CYS	engineered mutation	UNP P03147
F	48	ALA	CYS	engineered mutation	UNP P03147
F	107	ALA	CYS	engineered mutation	UNP P03147

- Molecule 1: Fab e21 heavy chain

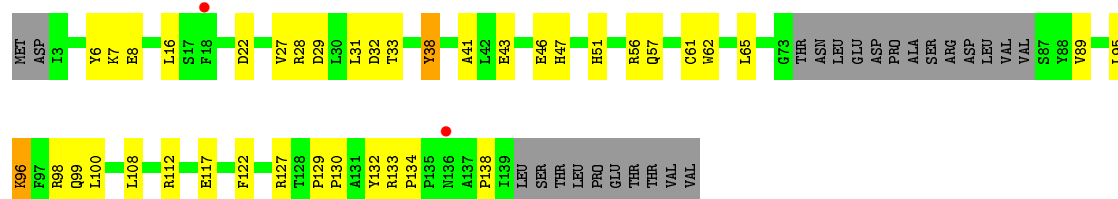


Chain D:  58% 37%



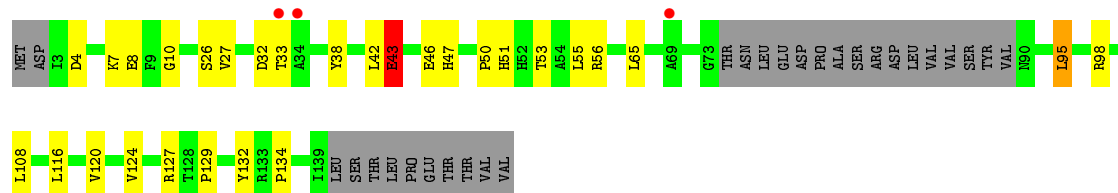
• Molecule 3: Capsid protein

Chain E:  57% 25% 17%



• Molecule 3: Capsid protein

Chain F:  62% 18% 19%



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	238.38Å 66.52Å 104.23Å 90.00° 110.30° 90.00°	Depositor
Resolution (Å)	39.98 – 3.15 39.98 – 3.15	Depositor EDS
% Data completeness (in resolution range)	97.0 (39.98-3.15) 91.5 (39.98-3.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.50 (at 3.12Å)	Xtriage
Refinement program	PHENIX (1.10.1 _2155: ???)	Depositor
R, R_{free}	0.231 , 0.283 0.231 , 0.284	Depositor DCC
R_{free} test set	2000 reflections (7.66%)	wwPDB-VP
Wilson B-factor (Å ²)	92.7	Xtriage
Anisotropy	0.259	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 56.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	8078	wwPDB-VP
Average B, all atoms (Å ²)	120.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.35% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	0.27	0/1596	0.52	0/2187
1	C	0.26	0/1502	0.50	0/2059
2	B	0.28	0/1618	0.57	1/2203 (0.0%)
2	D	0.28	0/1568	0.58	2/2132 (0.1%)
3	E	0.25	0/1001	0.46	0/1373
3	F	0.37	2/989 (0.2%)	0.46	1/1356 (0.1%)
All	All	0.28	2/8274 (0.0%)	0.52	4/11310 (0.0%)

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	C	0	1
2	B	0	3
All	All	0	4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	43	GLU	CD-OE1	-5.62	1.19	1.25
3	F	43	GLU	CD-OE2	-5.03	1.20	1.25

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	183	LEU	CA-CB-CG	7.66	132.92	115.30
2	B	90	LEU	CA-CB-CG	6.45	130.13	115.30
2	D	90	LEU	CA-CB-CG	5.59	128.17	115.30
3	F	95	LEU	CA-CB-CG	5.04	126.90	115.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	158	SER	Peptide
2	B	189	GLU	Peptide
2	B	8	THR	Peptide
1	C	152	PHE	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	1557	0	1517	38	0
1	C	1466	0	1422	40	0
2	B	1588	0	1518	50	0
2	D	1541	0	1474	80	0
3	E	969	0	929	28	0
3	F	957	0	925	22	0
All	All	8078	0	7785	234	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 15.

The worst 5 of 234 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)
2:D:110:ARG:NH2	2:D:113:ALA:HB2	1.07	1.38
2:D:110:ARG:NH2	2:D:113:ALA:CB	1.91	1.33
2:D:110:ARG:HE	2:D:142:TYR:HB2	1.23	1.02
2:D:110:ARG:NE	2:D:142:TYR:HB2	1.75	1.01
2:D:110:ARG:HH21	2:D:113:ALA:CB	1.74	0.94

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 X-ray entries followed by that with respect to entries of similar resolution.

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	212/233 (91%)	196 (92%)	15 (7%)	1 (0%)	29	65
1	C	194/233 (83%)	181 (93%)	13 (7%)	0	100	100
2	B	212/216 (98%)	187 (88%)	23 (11%)	2 (1%)	17	53
2	D	204/216 (94%)	181 (89%)	20 (10%)	3 (2%)	10	41
3	E	120/149 (80%)	105 (88%)	14 (12%)	1 (1%)	19	55
3	F	117/149 (78%)	106 (91%)	11 (9%)	0	100	100
All	All	1059/1196 (88%)	956 (90%)	96 (9%)	7 (1%)	22	59

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	8	THR
2	B	9	PRO
2	D	131	THR
2	D	9	PRO
2	D	201	VAL

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 X-ray entries followed by that with respect to entries of similar resolution.

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	172/195 (88%)	169 (98%)	3 (2%)	60	82
1	C	161/195 (83%)	156 (97%)	5 (3%)	40	70
2	B	173/180 (96%)	166 (96%)	7 (4%)	31	64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	169/180 (94%)	168 (99%)	1 (1%)	86	94
3	E	99/130 (76%)	96 (97%)	3 (3%)	41	71
3	F	100/130 (77%)	98 (98%)	2 (2%)	55	79
All	All	874/1010 (86%)	853 (98%)	21 (2%)	49	76

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	209	LYS
1	C	102	TYR
3	E	38	TYR
2	B	187	ASP
3	E	96	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	170	HIS
2	D	194	HIS
2	D	139	ASN
2	B	139	ASN
2	D	160	ASN

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 [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.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	214/233 (91%)	-0.18	2 (0%) 84 75	59, 107, 169, 192	0
1	C	200/233 (85%)	-0.12	3 (1%) 73 61	84, 120, 151, 179	0
2	B	214/216 (99%)	-0.21	1 (0%) 91 86	59, 95, 135, 188	0
2	D	208/216 (96%)	-0.03	1 (0%) 91 86	81, 123, 180, 246	0
3	E	124/149 (83%)	-0.03	2 (1%) 72 59	103, 141, 194, 279	0
3	F	121/149 (81%)	-0.04	3 (2%) 57 42	98, 134, 174, 260	0
All	All	1081/1196 (90%)	-0.11	12 (1%) 80 70	59, 119, 172, 279	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	140	GLY	4.7
1	C	141	THR	3.4
3	E	136	ASN	3.0
1	A	7	GLU	2.9
2	D	132	ALA	2.8

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.