wwPDB X-ray Structure Validation Summary Report (i)

May 16, 2020 – 06:42 pm BST

PDB ID	:	1XCT
Title	:	Complex HCV core-Fab 19D9D6-Protein L mutant (D55A, L57H, Y64W) in space group P21212
Authors	:	Menez, R.; Housden, N.G.; Harrison, S.; Jolivet-Reynaud, C.; Gore, M.G.; Stura, E.A.
Deposited on	:	2004-09-03
Resolution	:	3.05 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 (i) symbol.

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

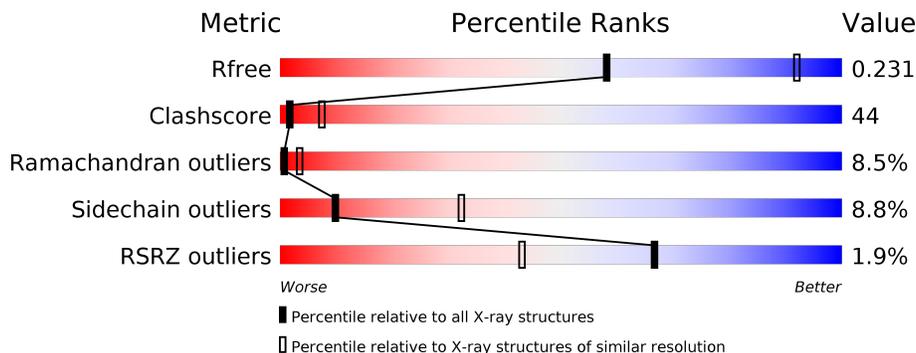|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
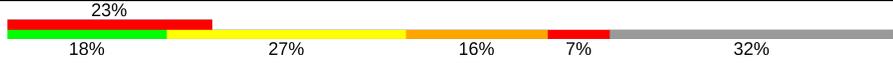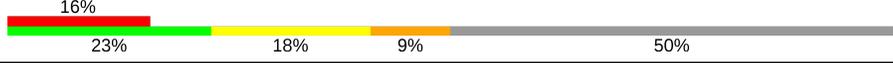*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.05 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.
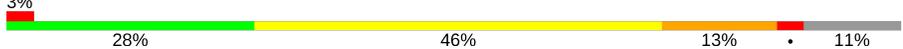
Metric	Percentile Ranks	Value
Rfree		0.231
Clashscore		44
Ramachandran outliers		8.5%
Sidechain outliers		8.8%
RSRZ outliers		1.9%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1754 (3.10-3.02)
Clashscore	141614	1864 (3.10-3.02)
Ramachandran outliers	138981	1794 (3.10-3.02)
Sidechain outliers	138945	1793 (3.10-3.02)
RSRZ outliers	127900	1713 (3.10-3.02)

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 <=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	P	44	23% / 18% 27% 16% 7% 32%
1	Q	44	16% / 23% 18% 9% 50%
2	A	220	44% 51% 5%
2	C	220	38% 50% 10% •
3	B	218	43% 45% 11%
3	D	218	37% 52% 11%

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	L	80	38% 44% 6% 13%
4	M	80	3% 28% 46% 13% 11%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8552 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 Capsid protein C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	P	30	Total 225	C 141	N 48	O 36		0	0	0
1	Q	22	Total 155	C 98	N 32	O 25		0	0	0

- Molecule 2 is a protein called Monoclonal antibody 19D9D6 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	220	Total 1708	C 1063	N 291	O 346	S 8	0	0	0
2	C	220	Total 1708	C 1063	N 291	O 346	S 8	0	0	0

- Molecule 3 is a protein called Monoclonal antibody 19D9D6 Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	218	Total 1660	C 1058	N 270	O 325	S 7	0	0	0
3	D	218	Total 1654	C 1052	N 270	O 325	S 7	0	0	0

- Molecule 4 is a protein called Protein L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	70	Total 547	C 346	N 90	O 110	S 1	0	0	0
4	M	71	Total 556	C 352	N 92	O 111	S 1	0	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	P	15	Total 15	O 15	0	0
5	Q	10	Total 10	O 10	0	0
5	A	91	Total 91	O 91	0	0
5	B	83	Total 83	O 83	0	0
5	L	21	Total 21	O 21	0	0
5	C	46	Total 46	O 46	0	0
5	D	57	Total 57	O 57	0	0
5	M	16	Total 16	O 16	0	0

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 the various outlier classes displayed in the second graphic. The second graphic 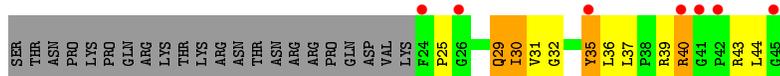shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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 C

Chain P:

- Molecule 1: Capsid protein C

Chain Q:

- Molecule 2: Monoclonal antibody 19D9D6 Light chain

Chain A:

- Molecule 2: Monoclonal antibody 19D9D6 Light chain

Chain C:

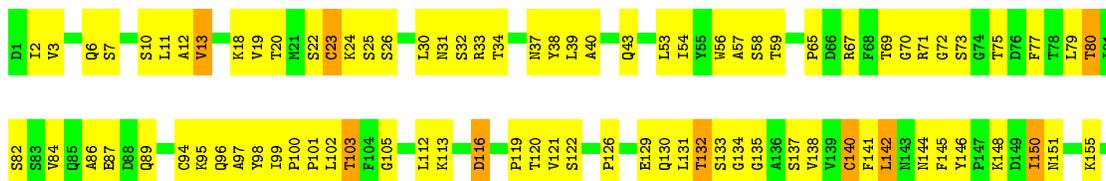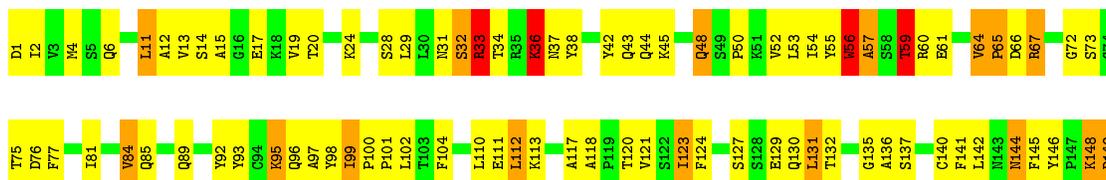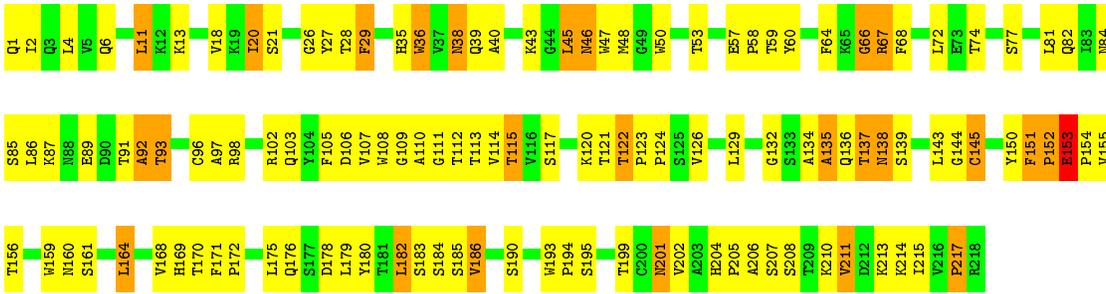I150 N151 V152 K153 W154 K155 I156 D157 G158 S159 E160 R161 V165 L166 N167 S168 W169 Q172 D173 S174 K175 D176 S177 T178 Y179 S180 M181 S182 S183 T184 L185 T186 L187 T188 K189 D190 E191 Y192 E193 R194 H195 N196 S197 Y198 T199 C200 E201 A202 T203 H204 K205 T206 I211 V212 K213 S214 F215

N216 R217 N218 E219 C220

• Molecule 3: Monoclonal antibody 19D9D6 Heavy chain

Chain B: 43%   45%   11%

Q1 I2 Q3 L4 V5 Q6 L11 K12 K13 V18 K19 I20 S21 G26 Y27 T28 F29 H35 W36 V37 N38 Q39 A40 K43 G44 L45 N46 W47 M48 G49 W50 T53 E57 P58 T59 Y60 F64 K65 G66 R67 F68 L72 E73 T74 S77 L81 Q82 I83 N84

S85 L86 K87 N88 E89 D90 T91 A92 T93 C96 A97 R98 R102 Q103 Y104 F105 D106 V107 W108 G109 A110 G111 T112 T113 V114 T115 V116 S117 K120 T121 T122 P123 P124 S125 V126 L129 G132 S133 A134 A135 Q136 T137 N138 S139 L143 G144 C145 Y150 F151 P152 E153 P154 V155

T156 W159 N160 S161 L164 V168 H169 T170 F171 P172 L175 Q176 S177 D178 L179 Y180 T181 L182 S183 S184 S185 V186 S190 W193 P194 S195 T199 C200 N201 V202 A203 H204 P205 A206 S207 S208 T209 K210 V211 D212 K213 K214 I215 V216 P217 R218

• Molecule 3: Monoclonal antibody 19D9D6 Heavy chain

Chain D: 37%   52%   11%

Q1 I2 Q3 L4 V5 Q6 S7 G8 P9 E10 L11 K12 K13 E16 T17 V18 K19 I20 S21 C22 K23 A24 S25 G26 Y27 T28 F29 T30 D31 H35 N38 Q39 W47 M48 G49 W50 V51 N52 T53 E54 T55 G56 E57 P58 T59 Y60 A61 F64 K65 G66 R67 F70 T74

S75 A76 S77 T78 A79 Y80 L81 Q82 I83 N84 S85 L86 K87 D90 T91 Y94 F95 C96 A97 R98 F99 R102 Q103 Y104 F105 D106 V107 G111 T112 T113 V114 T115 V116 S117 S118 A119 K120 T121 T122 P123 P124 S125 V126 Y127 P128 L129 S133 A134 A135 Q136 T137 N138 S139 M140

L143 G144 C145 G149 Y150 F151 P152 E153 P154 V157 T158 W159 N160 S161 G162 S163 L164 V168 H169 T170 F171 P172 A173 V174 L175 Q176 S177 D178 L179 Y180 T181 L182 S185 V186 S190 W193 P194 S195 E196 T197 V198 T199 C200 N201 H204 P205 A206 S207 S208 T209 K210 V211 D212

K213 K214 I215 R218

• Molecule 4: Protein L

Chain L: 38%   44%   6%   13%

MET ASN ILE LYS PHE ALA GLY LYS GLU LYS T13 P14 E15 E16 P17 K18 E19 E20 V21 T22 I23 K24 V25 N26 L27 I28 T36 A37 E38 T42 F43 E44 E45 A46 T47 A48 Y51 R52 A55 L56 H57 A58 K59 V60 N61 G62 D67 L68 E69 D70 N73 H74 M75 N76

I77 K78 F79 A80 G81 K82

• Molecule 4: Protein L

Chain M: 3%   28%   46%   13%   11%

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	129.50Å 222.48Å 43.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.05 19.96 – 3.04	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-3.05) 89.4 (19.96-3.04)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.12 (at 3.04Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.193 , 0.329 0.205 , 0.231	Depositor DCC
R_{free} test set	1161 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å^2)	54.2	Xtriage
Anisotropy	0.244	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.24 , 55.4	EDS
L-test for twinning[2]	$<|L|> = 0.48$, $<L^2> = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	8552	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	#$|Z|$ >5	RMSZ	#$|Z|$ >5
1	P	0.56	0/230	1.06	0/307
1	Q	0.33	0/158	0.62	0/211
2	A	0.39	0/1745	0.64	0/2366
2	C	0.36	0/1745	0.65	0/2366
3	B	0.41	0/1707	0.68	0/2335
3	D	0.36	0/1700	0.63	0/2326
4	L	0.37	0/558	0.67	0/751
4	M	0.43	0/567	0.63	0/762
All	All	0.39	0/8410	0.67	0/11424

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 (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	P	225	0	228	32	0
1	Q	155	0	161	15	0
2	A	1708	0	1658	118	0
2	C	1708	0	1658	182	0
3	B	1660	0	1615	133	0
3	D	1654	0	1608	165	0
4	L	547	0	522	50	0
4	M	556	0	535	69	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	91	0	0	4	0
5	B	83	0	0	4	0
5	C	46	0	0	2	0
5	D	57	0	0	2	0
5	L	21	0	0	2	0
5	M	16	0	0	2	0
5	P	15	0	0	0	0
5	Q	10	0	0	0	0
All	All	8552	0	7985	710	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 44.

The worst 5 of 710 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:A:196:ASN:HD21	2:A:218:ASN:HB3	1.19	1.03
4:M:19:GLU:HG2	4:M:40:LYS:HD2	1.42	1.02
3:D:157:VAL:HG12	3:D:158:THR:H	1.25	1.01
2:A:6:GLN:HG2	2:A:23:CYS:SG	2.07	0.94
2:C:19:VAL:HG12	2:C:81:ILE:HB	1.48	0.93

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 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	P	28/44 (64%)	12 (43%)	7 (25%)	9 (32%)	0	0
1	Q	20/44 (46%)	9 (45%)	8 (40%)	3 (15%)	0	0
2	A	218/220 (99%)	191 (88%)	24 (11%)	3 (1%)	11	36

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	218/220 (99%)	174 (80%)	28 (13%)	16 (7%)	1	5
3	B	216/218 (99%)	181 (84%)	17 (8%)	18 (8%)	1	4
3	D	216/218 (99%)	152 (70%)	40 (18%)	24 (11%)	0	2
4	L	68/80 (85%)	57 (84%)	7 (10%)	4 (6%)	1	8
4	M	69/80 (86%)	46 (67%)	11 (16%)	12 (17%)	0	0
All	All	1053/1124 (94%)	822 (78%)	142 (14%)	89 (8%)	1	3

5 of 89 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	P	18	ARG
1	P	23	LYS
1	P	24	PHE
1	P	42	PRO
2	A	57	ALA

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	P	21/37 (57%)	16 (76%)	5 (24%)	0	2
1	Q	14/37 (38%)	10 (71%)	4 (29%)	0	1
2	A	195/195 (100%)	179 (92%)	16 (8%)	11	35
2	C	195/195 (100%)	174 (89%)	21 (11%)	6	22
3	B	187/187 (100%)	165 (88%)	22 (12%)	5	18
3	D	186/187 (100%)	180 (97%)	6 (3%)	39	68
4	L	54/62 (87%)	52 (96%)	2 (4%)	34	64
4	M	55/62 (89%)	51 (93%)	4 (7%)	14	40
All	All	907/962 (94%)	827 (91%)	80 (9%)	10	32

5 of 80 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	114	VAL
3	B	186	VAL
3	D	122	THR
3	B	115	THR
3	B	153	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	46	ASN
2	C	31	ASN
4	M	26	ASN
4	L	35	GLN
2	C	44	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 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(Å²)	Q<0.9
1	P	30/44 (68%)	0.95	10 (33%)	0	0	23, 146, 199, 200	0
1	Q	22/44 (50%)	1.48	7 (31%)	0	0	122, 162, 198, 200	0
2	A	220/220 (100%)	-0.83	0	100	100	10, 37, 77, 143	0
2	C	220/220 (100%)	-0.57	0	100	100	15, 56, 125, 173	0
3	B	218/218 (100%)	-0.87	0	100	100	10, 32, 78, 189	0
3	D	218/218 (100%)	-0.66	1 (0%)	91	79	15, 56, 118, 191	0
4	L	70/80 (87%)	-0.61	0	100	100	12, 49, 146, 199	0
4	M	71/80 (88%)	-0.56	2 (2%)	53	28	12, 44, 167, 200	0
All	All	1069/1124 (95%)	-0.62	20 (1%)	66	43	10, 45, 140, 200	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	M	13	THR	3.7
1	P	27	GLY	3.5
1	Q	26	GLY	3.3
1	P	19	PRO	3.1
1	P	22	VAL	3.1

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.