



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 02:49 pm BST

PDB ID : 5USL  
Title : Structure of vaccinia virus D8 protein bound to human Fab vv304  
Authors : Zajonc, D.M.  
Deposited on : 2017-02-13  
Resolution : 2.90 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](mailto: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

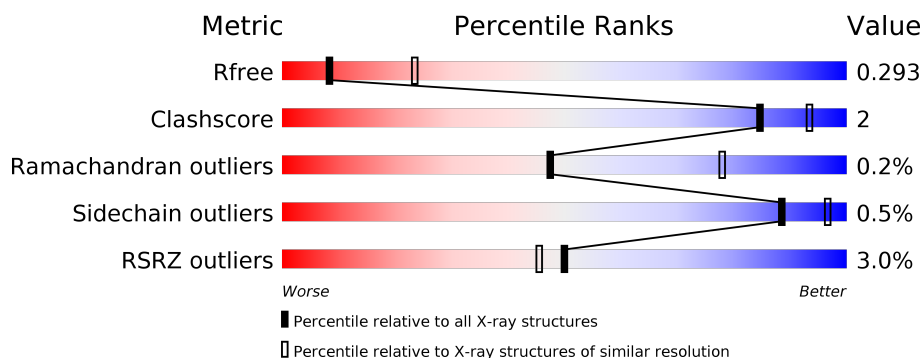
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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  $\geq 3$ , 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	241	<div> <div>91%</div> <div>6%</div> </div>
1	X	241	<div> <div>90%</div> <div>6%</div> </div>
2	B	224	<div> <div>4%</div> <div>90%</div> <div>6%</div> </div>
2	H	224	<div> <div>4%</div> <div>88%</div> <div>7%</div> <div>5%</div> </div>
3	C	211	<div> <div>4%</div> <div>94%</div> <div>5%</div> </div>
3	L	211	<div> <div>5%</div> <div>91%</div> <div>9%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10022 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 IMV membrane protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	X	227	Total	C	N	O	0	0	0
			1854	1199	304	351			
1	A	226	Total	C	N	O	0	0	0
			1857	1202	305	350			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	236	HIS	-	expression tag	UNP Q1M1K6
X	237	HIS	-	expression tag	UNP Q1M1K6
X	238	HIS	-	expression tag	UNP Q1M1K6
X	239	HIS	-	expression tag	UNP Q1M1K6
X	240	HIS	-	expression tag	UNP Q1M1K6
X	241	HIS	-	expression tag	UNP Q1M1K6
A	236	HIS	-	expression tag	UNP Q1M1K6
A	237	HIS	-	expression tag	UNP Q1M1K6
A	238	HIS	-	expression tag	UNP Q1M1K6
A	239	HIS	-	expression tag	UNP Q1M1K6
A	240	HIS	-	expression tag	UNP Q1M1K6
A	241	HIS	-	expression tag	UNP Q1M1K6

- Molecule 2 is a protein called Fab vv304 Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	213	Total	C	N	O	S	0	0	0
			1596	1014	265	311	6			
2	B	211	Total	C	N	O	S	0	0	0
			1571	997	260	308	6			

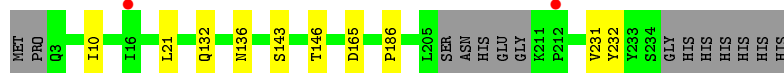
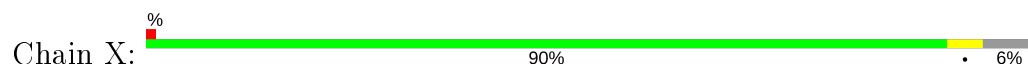
- Molecule 3 is a protein called Fab vv304 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	210	Total	C	N	O	S	0	0	0
			1572	982	259	325	6			
3	C	210	Total	C	N	O	S	0	0	0
			1572	981	258	327	6			

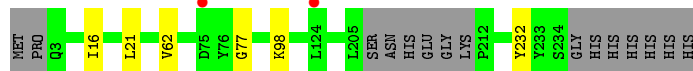
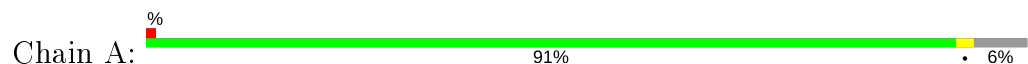
### 3 Residue-property plots [i](#)

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 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.

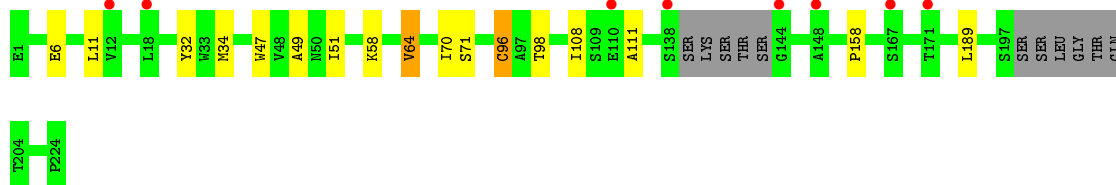
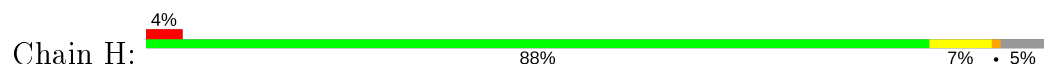
- Molecule 1: IMV membrane protein



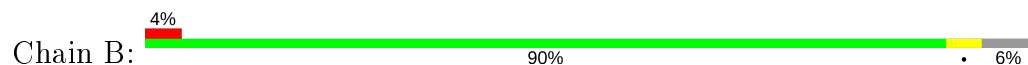
- Molecule 1: IMV membrane protein



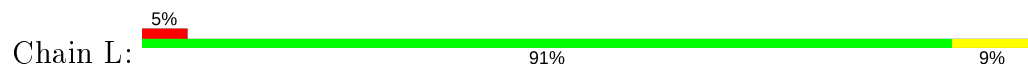
- Molecule 2: Fab vv304 Heavy chain

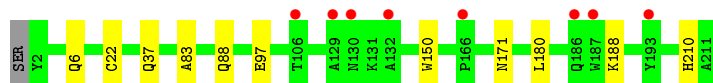


- Molecule 2: Fab vv304 Heavy chain



- Molecule 3: Fab vv304 Light chain





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.98Å 117.98Å 104.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.81 – 2.90 47.81 – 2.90	Depositor EDS
% Data completeness (in resolution range)	94.7 (47.81-2.90) 94.7 (47.81-2.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.81 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.247 , 0.297 0.246 , 0.293	Depositor DCC
$R_{free}$ test set	1537 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	54.1	Xtriage
Anisotropy	0.094	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 19.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.065 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	10022	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.09 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.1269e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical 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	0.26	0/1907	0.44	0/2590
1	X	0.26	0/1904	0.44	0/2590
2	B	0.26	0/1604	0.46	0/2188
2	H	0.26	0/1631	0.45	0/2223
3	C	0.26	0/1612	0.46	0/2204
3	L	0.27	0/1612	0.47	0/2203
All	All	0.26	0/10270	0.45	0/13998

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	1857	0	1821	5	0
1	X	1854	0	1800	6	0
2	B	1571	0	1521	5	0
2	H	1596	0	1573	12	0
3	C	1572	0	1490	5	0
3	L	1572	0	1497	12	0
All	All	10022	0	9702	45	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 2.

The worst 5 of 45 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)
3:L:188:LYS:HA	3:L:210:HIS:CE1	2.20	0.77
2:H:51:ILE:HD12	2:H:58:LYS:HD3	1.85	0.58
2:H:98:THR:OG1	2:H:111:ALA:O	2.23	0.55
1:A:21:LEU:HD11	1:A:232:TYR:CB	2.38	0.53
2:H:64:VAL:O	2:H:64:VAL:HG12	2.08	0.53

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	222/241 (92%)	207 (93%)	14 (6%)	1 (0%)	29	61
1	X	223/241 (92%)	211 (95%)	12 (5%)	0	100	100
2	B	201/224 (90%)	195 (97%)	6 (3%)	0	100	100
2	H	207/224 (92%)	199 (96%)	7 (3%)	1 (0%)	29	61
3	C	208/211 (99%)	198 (95%)	10 (5%)	0	100	100
3	L	208/211 (99%)	198 (95%)	10 (5%)	0	100	100
All	All	1269/1352 (94%)	1208 (95%)	59 (5%)	2 (0%)	47	78

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	64	VAL
1	A	77	GLY

### 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	209/222 (94%)	209 (100%)	0	100	100
1	X	207/222 (93%)	207 (100%)	0	100	100
2	B	174/189 (92%)	173 (99%)	1 (1%)	86	96
2	H	178/189 (94%)	177 (99%)	1 (1%)	86	96
3	C	176/178 (99%)	175 (99%)	1 (1%)	86	96
3	L	176/178 (99%)	173 (98%)	3 (2%)	60	86
All	All	1120/1178 (95%)	1114 (100%)	6 (0%)	88	96

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	L	119	LEU
3	C	171	ASN
3	L	208	MET
3	L	110	GLN
2	B	19	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
3	L	196	GLN
3	C	196	GLN
3	C	171	ASN
3	L	37	GLN
3	C	172	ASN

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	226/241 (93%)	0.19	2 (0%) 84 84	37, 54, 69, 91	0
1	X	227/241 (94%)	0.08	2 (0%) 84 84	39, 52, 67, 99	0
2	B	211/224 (94%)	0.42	8 (3%) 40 36	44, 65, 81, 91	0
2	H	213/224 (95%)	0.37	8 (3%) 40 36	42, 58, 87, 101	0
3	C	210/211 (99%)	0.46	8 (3%) 40 36	38, 57, 76, 84	0
3	L	210/211 (99%)	0.36	11 (5%) 27 23	38, 53, 75, 97	0
All	All	1297/1352 (95%)	0.31	39 (3%) 50 45	37, 56, 78, 101	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	117	VAL	3.5
3	L	209	ALA	3.2
3	C	130	ASN	3.1
3	L	211	ALA	3.0
1	X	16	ILE	3.0

### 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.