



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 12:24 pm BST

PDB ID : 5XRQ  
Title : Crystal structure of human monoclonal antibody H3v-47  
Authors : Zhang, H.; Willson, I.A.  
Deposited on : 2017-06-09  
Resolution : 2.60 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

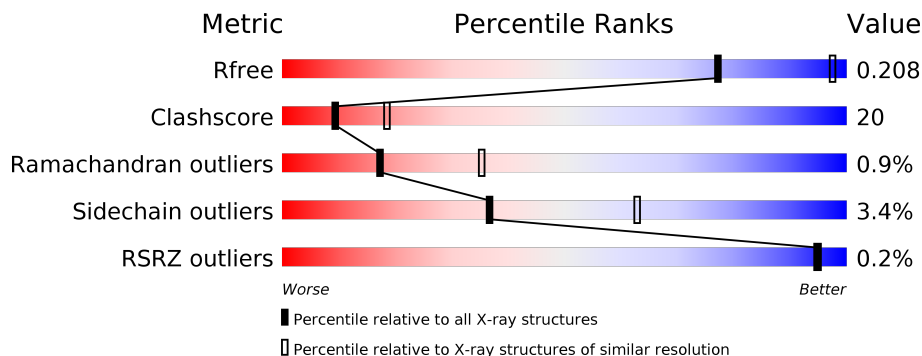
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	236	
1	H	236	
2	B	214	
2	L	214	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 6639 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 H3v-47 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	H	220	Total 1647	C 1039	N 275	O 326	S 7	0	0	0
1	A	219	Total 1638	C 1034	N 273	O 324	S 7	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	217	HIS	-	expression tag	UNP Q6N089
H	218	HIS	-	expression tag	UNP Q6N089
H	219	HIS	-	expression tag	UNP Q6N089
H	220	HIS	-	expression tag	UNP Q6N089
H	221	HIS	-	expression tag	UNP Q6N089
H	222	HIS	-	expression tag	UNP Q6N089
A	217	HIS	-	expression tag	UNP Q6N089
A	218	HIS	-	expression tag	UNP Q6N089
A	219	HIS	-	expression tag	UNP Q6N089
A	220	HIS	-	expression tag	UNP Q6N089
A	221	HIS	-	expression tag	UNP Q6N089
A	222	HIS	-	expression tag	UNP Q6N089

- Molecule 2 is a protein called Fab H3v-47 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	L	213	Total 1629	C 1009	N 282	O 333	S 5	0	1	0
2	B	212	Total 1619	C 1003	N 280	O 331	S 5	0	1	0

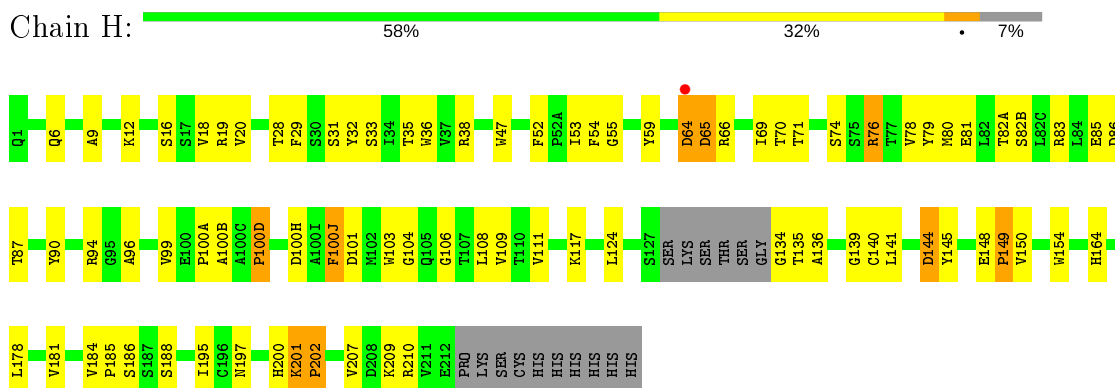
- Molecule 3 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
3	H	17	Total 17	O 17	0	0
3	L	30	Total 30	O 30	0	0
3	A	19	Total 19	O 19	0	0
3	B	40	Total 40	O 40	0	0

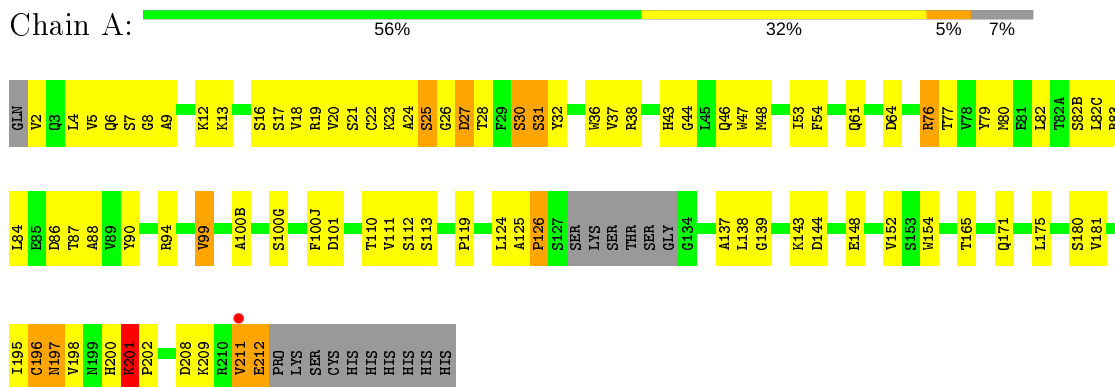
### 3 Residue-property plots [i](#)

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 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.

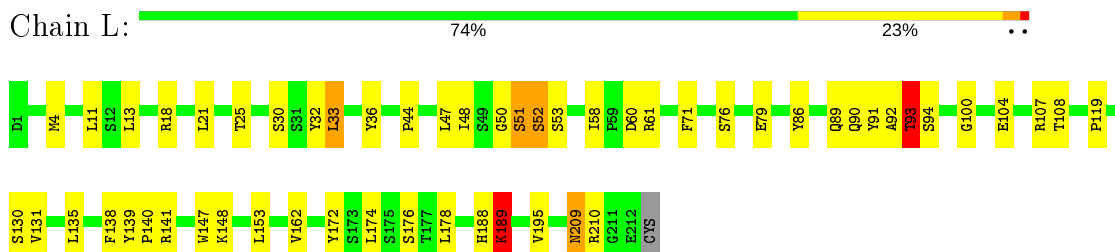
- Molecule 1: Fab H3v-47 heavy chain



- Molecule 1: Fab H3v-47 heavy chain

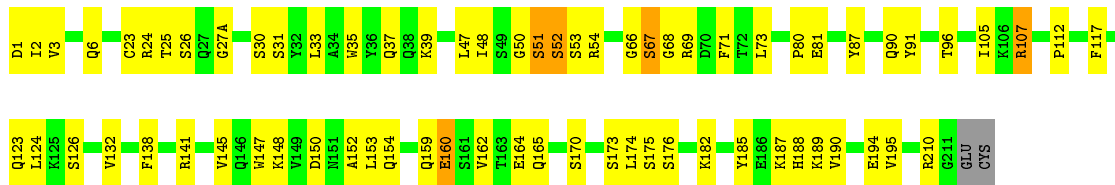


- Molecule 2: Fab H3v-47 light chain



- Molecule 2: Fab H3v-47 light chain

Chain B:  66% 30% ..



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	135.13Å 135.13Å 78.53Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.92 – 2.60 46.92 – 2.60	Depositor EDS
% Data completeness (in resolution range)	98.5 (46.92-2.60) 98.3 (46.92-2.60)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.60 (at 2.61Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.180 , 0.214 0.183 , 0.208	Depositor DCC
$R_{free}$ test set	2456 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.9	Xtriage
Anisotropy	0.631	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 39.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.026 for -h,-k,l 0.468 for h,-h-k,-l 0.027 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6639	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.57% of the height of the origin peak. No significant pseudotranslation is detected.*

<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 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	A	0.51	1/1677 (0.1%)	0.77	2/2287 (0.1%)
1	H	0.53	1/1686 (0.1%)	0.74	3/2299 (0.1%)
2	B	0.46	0/1656	0.74	4/2247 (0.2%)
2	L	0.47	0/1668	0.78	5/2263 (0.2%)
All	All	0.49	2/6687 (0.0%)	0.76	14/9096 (0.2%)

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	A	0	2
1	H	0	1
2	L	0	1
All	All	0	4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	202	PRO	N-CD	-9.35	1.34	1.47
1	A	76	ARG	CZ-NH1	-6.13	1.25	1.33

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	211	VAL	N-CA-C	-9.82	84.47	111.00
2	L	93	THR	N-CA-C	8.96	135.21	111.00
1	A	212	GLU	N-CA-C	8.89	134.99	111.00
2	B	52	SER	N-CA-CB	-8.84	97.25	110.50
1	H	65	ASP	N-CA-CB	-8.49	95.32	110.60

There are no chirality outliers.



All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	100(J)	PHE	Peptide
1	A	201	LYS	Mainchain
1	H	100(J)	PHE	Peptide
2	L	93	THR	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	1638	0	1598	90	0
1	H	1647	0	1609	72	0
2	B	1619	0	1575	55	1
2	L	1629	0	1586	43	1
3	A	19	0	0	3	0
3	B	40	0	0	4	1
3	H	17	0	0	2	1
3	L	30	0	0	2	0
All	All	6639	0	6368	254	2

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 20.

The worst 5 of 254 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:8:GLY:CA	1:A:201:LYS:NZ	1.80	1.44
1:A:8:GLY:HA3	1:A:201:LYS:NZ	1.33	1.42
1:H:201:LYS:HG3	1:H:202:PRO:HD3	1.09	1.07
1:A:8:GLY:HA2	1:A:201:LYS:NZ	1.69	1.04
1:A:100(G):SER:OG	2:B:91:TYR:O	1.79	0.99

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:164:GLU:OE1	3:B:303:HOH:O 3_454	2.06	0.14
2:L:107:ARG:NH2	3:H:305:HOH:O 3_564	2.14	0.06

### 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	215/236 (91%)	197 (92%)	15 (7%)	3 (1%)	11	22
1	H	216/236 (92%)	196 (91%)	16 (7%)	4 (2%)	8	15
2	B	211/214 (99%)	196 (93%)	15 (7%)	0	100	100
2	L	212/214 (99%)	203 (96%)	8 (4%)	1 (0%)	29	52
All	All	854/900 (95%)	792 (93%)	54 (6%)	8 (1%)	17	35

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	101	ASP
1	A	27	ASP
1	A	44	GLY
1	H	64	ASP
1	H	100(D)	PRO

#### 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	184/200 (92%)	176 (96%)	8 (4%)	29	54
1	H	185/200 (92%)	180 (97%)	5 (3%)	44	71
2	B	185/186 (100%)	177 (96%)	8 (4%)	29	54
2	L	186/186 (100%)	182 (98%)	4 (2%)	52	76
All	All	740/772 (96%)	715 (97%)	25 (3%)	37	63

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	31	SER
1	A	99	VAL
2	B	160	GLU
1	A	43	HIS
1	A	196	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	H	6	GLN
2	L	90	GLN
2	B	38	GLN
2	B	159	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, 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	219/236 (92%)	-0.28	1 (0%) 91   89	37, 64, 99, 117	0
1	H	220/236 (93%)	-0.21	1 (0%) 91   89	39, 65, 98, 171	0
2	B	212/214 (99%)	-0.30	0 100   100	29, 52, 81, 117	0
2	L	213/214 (99%)	-0.24	0 100   100	37, 57, 84, 102	0
All	All	864/900 (96%)	-0.26	2 (0%) 95   95	29, 60, 91, 171	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	64	ASP	4.6
1	A	211	VAL	2.1

### 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.