



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 02:02 pm GMT

PDB ID : 3PP3  
Title : Epitope characterization and crystal structure of GA101 provide insights into the molecular basis for the type I / type II distinction of anti- CD20 antibodies  
Authors : Hopfner, K.-P.; Lammens, A.  
Deposited on : 2010-11-24  
Resolution : 2.51 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
Xtrriage (Phenix)	:	1.13
EDS	:	trunk30967
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30967

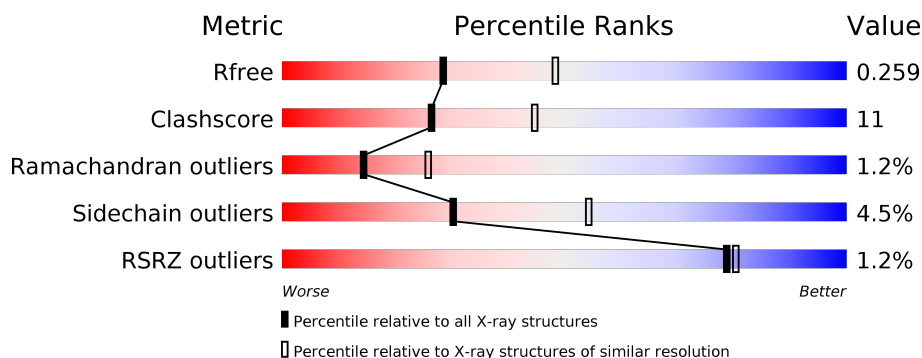
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.51 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))
$R_{free}$	111664	4155 (2.50-2.50)
Clashscore	122126	4827 (2.50-2.50)
Ramachandran outliers	120053	4735 (2.50-2.50)
Sidechain outliers	120020	4737 (2.50-2.50)
RSRZ outliers	108989	4058 (2.50-2.50)

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. 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	H	224	<div> <div>2%</div> <div> <div></div> <div>72%</div> <div>24%</div> <div>• •</div> </div> </div>
1	I	224	<div> <div>%</div> <div> <div></div> <div>74%</div> <div>21%</div> <div>• •</div> </div> </div>
2	K	219	<div> <div></div> <div> <div></div> <div>74%</div> <div>25%</div> <div>•</div> </div> </div>
2	L	219	<div> <div>%</div> <div> <div></div> <div>77%</div> <div>21%</div> <div>•</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7200 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 GA101 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	220	Total	C	N	O	S	50	0	0
			1657	1052	275	324	6			
1	I	219	Total	C	N	O	S	43	0	0
			1648	1046	273	323	6			

- Molecule 2 is a protein called GA101 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	219	Total	C	N	O	S	36	0	0
			1684	1060	278	339	7			
2	K	219	Total	C	N	O	S	18	0	0
			1684	1060	278	339	7			

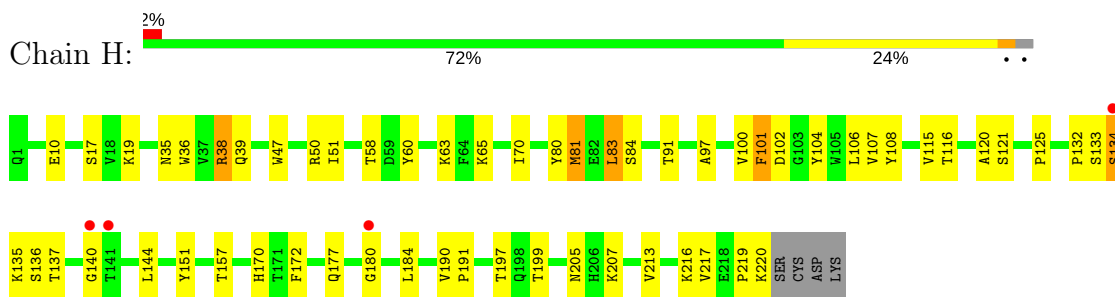
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	H	158	Total	O	0	0
			158	158		
3	L	95	Total	O	0	0
			95	95		
3	I	160	Total	O	0	0
			160	160		
3	K	114	Total	O	0	0
			114	114		

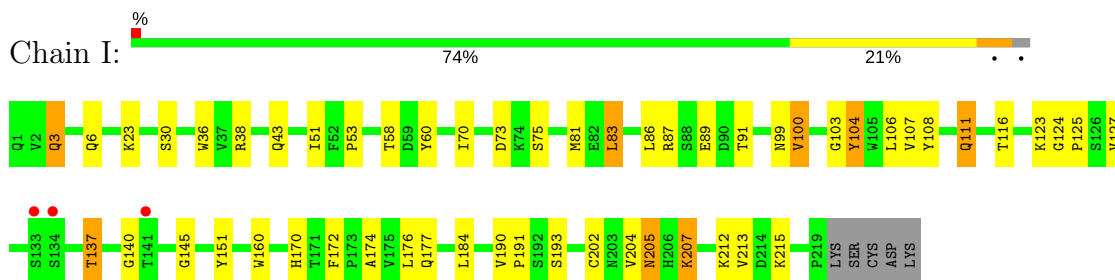
### 3 Residue-property plots

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.

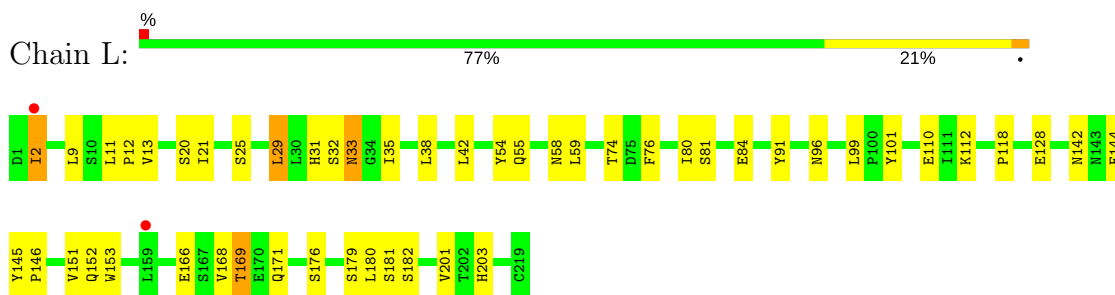
- Molecule 1: GA101 Fab heavy chain



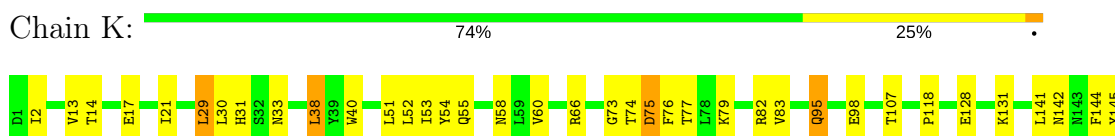
- Molecule 1: GA101 Fab heavy chain

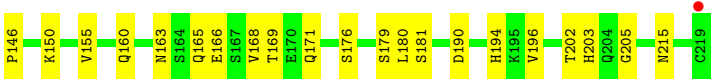


- Molecule 2: GA101 Fab light chain



- Molecule 2: GA101 Fab light chain





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	273.14Å 38.04Å 90.22Å 90.00° 98.89° 90.00°	Depositor
Resolution (Å)	25.00 – 2.51 34.97 – 2.51	Depositor EDS
% Data completeness (in resolution range)	97.6 (25.00-2.51) 97.6 (34.97-2.51)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.88 (at 2.51Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, $R_{free}$	0.206 , 0.269 0.198 , 0.259	Depositor DCC
$R_{free}$ test set	2000 reflections (6.36%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.0	Xtriage
Anisotropy	0.378	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 50.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7200	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	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 39.78 % 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 3.0178e-04. 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 [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	H	0.23	0/1700	0.42	0/2316
1	I	0.23	0/1691	0.40	0/2305
2	K	0.23	0/1721	0.42	0/2338
2	L	0.23	0/1721	0.41	0/2338
All	All	0.23	0/6833	0.41	0/9297

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	H	1657	0	1623	35	0
1	I	1648	0	1610	40	0
2	K	1684	0	1650	37	0
2	L	1684	0	1650	38	0
3	H	158	0	0	2	0
3	I	160	0	0	2	0
3	K	114	0	0	0	0
3	L	95	0	0	2	0
All	All	7200	0	6533	143	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 11.

The worst 5 of 143 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:I:111:GLN:HE21	1:I:111:GLN:H	1.17	0.91
1:H:60:TYR:HE2	1:H:70:ILE:HG12	1.43	0.84
2:K:66:ARG:HD2	2:K:82:ARG:O	1.78	0.83
1:I:60:TYR:HE2	1:I:70:ILE:HG12	1.49	0.77
1:I:6:GLN:H	1:I:111:GLN:HE22	1.29	0.77

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	H	218/224 (97%)	205 (94%)	8 (4%)	5 (2%)	7	11
1	I	217/224 (97%)	203 (94%)	11 (5%)	3 (1%)	12	22
2	K	217/219 (99%)	210 (97%)	6 (3%)	1 (0%)	31	51
2	L	217/219 (99%)	210 (97%)	6 (3%)	1 (0%)	31	51
All	All	869/886 (98%)	828 (95%)	31 (4%)	10 (1%)	14	26

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	140	GLY
1	H	134	SER
1	H	136	SER
1	H	135	LYS
1	I	104	TYR



### 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	H	185/189 (98%)	176 (95%)	9 (5%)	27	50
1	I	184/189 (97%)	173 (94%)	11 (6%)	21	39
2	K	194/194 (100%)	187 (96%)	7 (4%)	38	65
2	L	194/194 (100%)	187 (96%)	7 (4%)	38	65
All	All	757/766 (99%)	723 (96%)	34 (4%)	30	54

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	L	169	THR
1	I	43	GLN
2	K	95	GLN
1	I	23	LYS
1	H	177	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	43	GLN
1	I	205	ASN
2	K	203	HIS
1	I	99	ASN
1	I	111	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	H	216/224 (96%)	-0.19	4 (1%)	66 69	16, 28, 48, 77	11 (5%)
1	I	214/224 (95%)	-0.24	3 (1%)	75 77	13, 25, 43, 70	6 (2%)
2	K	219/219 (100%)	-0.25	1 (0%)	90 91	19, 29, 46, 88	7 (3%)
2	L	219/219 (100%)	-0.04	2 (0%)	84 85	21, 33, 54, 78	12 (5%)
All	All	868/886 (97%)	-0.18	10 (1%)	79 80	13, 29, 48, 88	36 (4%)

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	134	SER	4.9
1	H	180	GLY	4.5
1	H	140	GLY	4.2
2	L	2	ILE	3.5
2	L	159	LEU	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.