



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 14, 2019 – 08:53 PM EDT

PDB ID : 6RCU  
Title : PfRH5 bound to monoclonal antibodies R5.004 and R5.016  
Authors : Alanine, D.W.G.; Draper, S.J.; Higgins, M.K.  
Deposited on : 2019-04-11  
Resolution : 4.00 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 : 2.4  
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) : 2.4

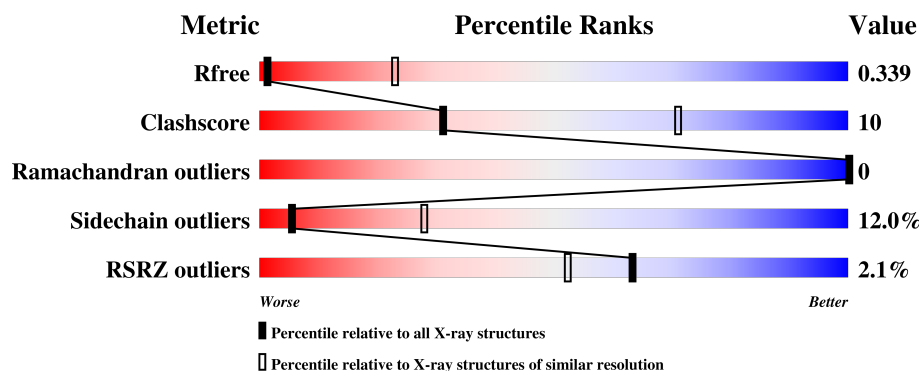
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.00 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	1008 (4.38-3.62)
Clashscore	122126	1012 (4.34-3.66)
Ramachandran outliers	120053	1000 (4.36-3.64)
Sidechain outliers	120020	1023 (4.38-3.62)
RSRZ outliers	108989	1107 (4.40-3.60)

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. 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	501	
2	B	232	
3	C	219	
4	D	219	
5	E	464	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9142 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 Reticulocyte binding protein homologue 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	304	Total	C	N	O	S	0	0	0
			2583	1666	436	465	16			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	38	GLN	ASN	conflict	UNP Q8IFM5
A	216	ALA	THR	conflict	UNP Q8IFM5

- Molecule 2 is a protein called R5.004 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	222	Total	C	N	O	S	495	1	0
			1658	1045	276	330	7			

- Molecule 3 is a protein called R5.004 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	211	Total	C	N	O	S	501	0	0
			1560	973	265	318	4			

- Molecule 4 is a protein called R5.016 light chain.

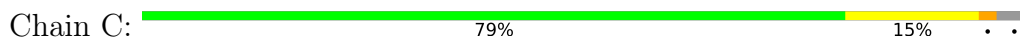
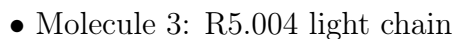
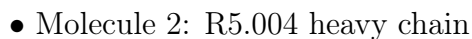
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	210	Total	C	N	O	S	333	0	0
			1618	1012	272	329	5			

- Molecule 5 is a protein called R5.016 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	228	Total	C	N	O	S	570	0	0
			1723	1086	285	343	9			



- Molecule 1: Reticulocyte binding protein homologue 5





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	235.15Å 58.78Å 116.79Å 90.00° 106.41° 90.00°	Depositor
Resolution (Å)	46.92 – 4.00 46.92 – 4.01	Depositor EDS
% Data completeness (in resolution range)	98.0 (46.92-4.00) 98.0 (46.92-4.01)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.76 (at 4.00Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.285 , 0.339 0.285 , 0.339	Depositor DCC
$R_{free}$ test set	644 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	113.6	Xtriage
Anisotropy	0.562	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 142.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	9142	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	155.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% 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	0.36	0/2638	0.69	0/3537
2	B	0.35	0/1698	0.72	0/2312
3	C	0.35	0/1600	0.64	0/2189
4	D	0.27	0/1653	0.50	0/2245
5	E	0.33	0/1766	0.60	0/2408
All	All	0.33	0/9355	0.64	0/12691

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	A	2583	0	2602	65	0
2	B	1658	0	1620	34	0
3	C	1560	0	1506	21	0
4	D	1618	0	1566	20	0
5	E	1723	0	1665	14	0
All	All	9142	0	8959	141	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 10.

The worst 5 of 141 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:B:8:GLY:O	2:B:114:THR:HG23	1.61	0.98
3:C:49:ILE:HG23	3:C:54:GLN:O	1.75	0.87
2:B:6:GLN:HE22	2:B:20:VAL:HG12	1.45	0.81
1:A:241:PRO:HG3	1:A:409:THR:HG21	1.68	0.76
1:A:155:TYR:HD1	1:A:156:ASN:H	1.34	0.75

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	300/501 (60%)	283 (94%)	17 (6%)	0	100	100
2	B	221/232 (95%)	201 (91%)	20 (9%)	0	100	100
3	C	209/219 (95%)	194 (93%)	15 (7%)	0	100	100
4	D	208/219 (95%)	193 (93%)	15 (7%)	0	100	100
5	E	224/464 (48%)	210 (94%)	14 (6%)	0	100	100
All	All	1162/1635 (71%)	1081 (93%)	81 (7%)	0	100	100

There are no Ramachandran outliers to report.

### 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	294/486 (60%)	235 (80%)	59 (20%)	1	10

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	186/196 (95%)	164 (88%)	22 (12%)	6	28
3	C	176/183 (96%)	158 (90%)	18 (10%)	8	33
4	D	185/192 (96%)	165 (89%)	20 (11%)	7	32
5	E	193/409 (47%)	188 (97%)	5 (3%)	49	74
All	All	1034/1466 (70%)	910 (88%)	124 (12%)	5	28

5 of 124 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	491	ASN
2	B	63	LYS
4	D	106	ILE
1	A	500	VAL
2	B	13	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	6	GLN
2	B	39	GLN
3	C	39	HIS

### 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	304/501 (60%)	-0.34	0 100 100	66, 104, 152, 220	0
2	B	153/232 (65%)	0.20	13 (8%) 11 8	78, 131, 266, 301	0
3	C	146/219 (66%)	-0.00	1 (0%) 87 81	84, 129, 216, 233	0
4	D	169/219 (77%)	0.27	4 (2%) 59 47	91, 158, 224, 248	0
5	E	151/464 (32%)	0.14	1 (0%) 87 81	86, 128, 230, 242	0
All	All	923/1635 (56%)	-0.01	19 (2%) 63 52	66, 122, 228, 301	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	179	SER	4.9
2	B	166	LEU	4.5
2	B	158	THR	3.9
2	B	167	THR	3.2
4	D	136	LEU	3.0

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.