



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 30, 2021 – 12:07 pm BST

PDB ID : 7OAO  
Title : Nanobody C5 bound to RBD  
Authors : Naismith, J.H.; Mikolajek, H.  
Deposited on : 2021-04-19  
Resolution : 1.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

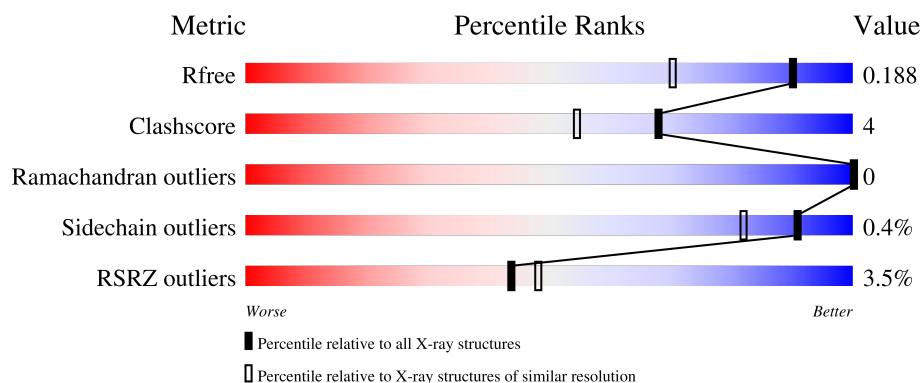
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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 of 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	EEE	210	<div> <div style="width: 100%; height: 10px; background-color: red;"></div> <div style="display: flex; justify-content: space-between; align-items: center;"> <span>%</span> <span>87%</span> <span>7%</span> <span>7%</span> </div> </div>
2	FFF	128	<div> <div style="width: 100%; height: 10px; background-color: red;"></div> <div style="display: flex; justify-content: space-between; align-items: center;"> <span>6%</span> <span>87%</span> <span>8%</span> <span>5%</span> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2800 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 Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	EEE	196	Total	C	N	O	S	0	3	0
			1562	1004	258	292	8			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
EEE	533	LYS	-	expression tag	UNP P0DTC2
EEE	534	HIS	-	expression tag	UNP P0DTC2
EEE	535	HIS	-	expression tag	UNP P0DTC2
EEE	536	HIS	-	expression tag	UNP P0DTC2
EEE	537	HIS	-	expression tag	UNP P0DTC2
EEE	538	HIS	-	expression tag	UNP P0DTC2
EEE	539	HIS	-	expression tag	UNP P0DTC2

- Molecule 2 is a protein called C5 nanobody.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	FFF	121	Total	C	N	O	S	0	6	0
			930	583	159	180	8			

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	EEE	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	EEE	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	EEE	202	Total 202	O 202	0	0
5	FFF	88	Total 88	O 88	0	0

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


These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

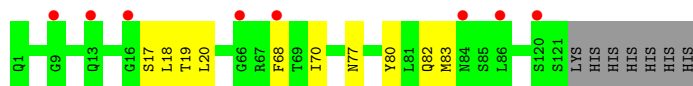
- Molecule 1: Spike protein S1

Chain EEE: 



- Molecule 2: C5 nanobody

Chain FFF: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.22Å 154.33Å 28.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	51.44 – 1.50 41.70 – 1.50	Depositor EDS
% Data completeness (in resolution range)	99.4 (51.44-1.50) 99.4 (41.70-1.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.65 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.152 , 0.186 0.154 , 0.188	Depositor DCC
$R_{free}$ test set	2555 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.4	Xtriage
Anisotropy	0.499	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	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	0.97	EDS
Total number of atoms	2800	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.21% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ACT, NAG

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	EEE	0.60	0/1615	0.73	0/2200
2	FFF	0.69	0/964	0.80	0/1305
All	All	0.64	0/2579	0.76	0/3505

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	EEE	1562	0	1491	9	0
2	FFF	930	0	926	10	0
3	EEE	14	0	13	0	0
4	EEE	4	0	3	0	0
5	EEE	202	0	0	2	0
5	FFF	88	0	0	1	0
All	All	2800	0	2433	19	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 4.

All (19) 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:EEE:364:ASP:O	1:EEE:367:VAL:HG12	1.82	0.78
2:FFF:70[A]:ILE:HD12	2:FFF:80:TYR:O	1.87	0.73
2:FFF:19:THR:OG1	2:FFF:82[A]:GLN:OE1	2.09	0.70
1:EEE:360:ASN:H	1:EEE:523:THR:HG23	1.70	0.56
2:FFF:68:PHE:CZ	2:FFF:83[B]:MET:CE	2.88	0.56
2:FFF:68:PHE:CZ	2:FFF:83[B]:MET:HE2	2.44	0.53
2:FFF:68:PHE:CE2	2:FFF:83[B]:MET:CE	2.94	0.51
2:FFF:82[A]:GLN:OE1	2:FFF:82[A]:GLN:HA	2.13	0.49
1:EEE:527:PRO:C	5:EEE:860:HOH:O	2.51	0.48
1:EEE:444:LYS:NZ	5:EEE:702:HOH:O	2.47	0.47
1:EEE:383:SER:O	1:EEE:387:LEU:HD13	2.16	0.45
1:EEE:354:ASN:O	1:EEE:398:ASP:HA	2.16	0.44
1:EEE:359:SER:HA	1:EEE:524:VAL:HG22	1.98	0.44
1:EEE:359:SER:HA	1:EEE:524:VAL:CG2	2.47	0.44
2:FFF:18[B]:LEU:HD12	2:FFF:18[B]:LEU:HA	1.82	0.42
1:EEE:392:PHE:CD1	1:EEE:515:PHE:HB3	2.54	0.42
2:FFF:20:LEU:HD11	2:FFF:83[A]:MET:HE1	2.03	0.41
2:FFF:82[A]:GLN:NE2	5:FFF:207:HOH:O	2.53	0.40

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	EEE	197/210 (94%)	192 (98%)	5 (2%)	0	100	100
2	FFF	125/128 (98%)	122 (98%)	3 (2%)	0	100	100
All	All	322/338 (95%)	314 (98%)	8 (2%)	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	EEE	172/183 (94%)	172 (100%)	0	100	100
2	FFF	104/105 (99%)	103 (99%)	1 (1%)	76	57
All	All	276/288 (96%)	275 (100%)	1 (0%)	91	82

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	FFF	77	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	EEE	601	1	14,14,15	0.32	0	17,19,21	1.12	1 (5%)
4	ACT	EEE	602	-	1,3,3	2.05	1 (100%)	0,3,3	-	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	EEE	601	1	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	EEE	602	ACT	CH3-C	2.05	1.51	1.48

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	EEE	601	NAG	O5-C5-C4	-2.24	105.39	110.83

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	EEE	196/210 (93%)	-0.18	3 (1%) 73 78	14, 23, 38, 65	0
2	FFF	121/128 (94%)	0.25	8 (6%) 18 19	15, 29, 49, 65	0
All	All	317/338 (93%)	-0.02	11 (3%) 44 48	14, 26, 45, 65	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	EEE	332	ILE	7.5
1	EEE	335	LEU	4.0
2	FFF	13	GLN	3.2
2	FFF	66	GLY	3.0
2	FFF	86	LEU	2.8
2	FFF	120	SER	2.5
2	FFF	84	ASN	2.4
2	FFF	9	GLY	2.4
2	FFF	16	GLY	2.3
2	FFF	68	PHE	2.3
1	EEE	333	THR	2.2

### 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 monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ACT	EEE	602	4/4	0.78	0.16	34,41,48,50	0
3	NAG	EEE	601	14/15	0.82	0.25	42,52,57,64	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.