



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 6, 2021 – 09:04 PM GMT

PDB ID : 6XUX  
Title : Crystal structure of Megabody Mb-Nb207-cYgjK\_NO  
Authors : Steyaert, J.; Uchanski, T.; Fischer, B.  
Deposited on : 2020-01-21  
Resolution : 1.90 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  
Xtriage (Phenix) : 1.13  
EDS : 2.16  
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) : 2.16

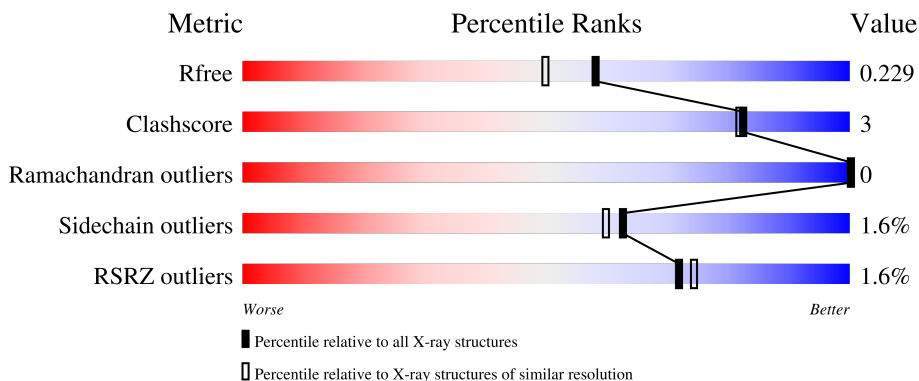
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

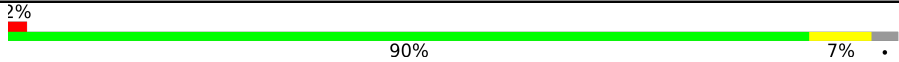
The reported resolution of this entry is 1.90 Å.

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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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 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	A	907	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7672 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 Nanobody,Glucosidase YgjK,Glucosidase YgjK,Nanobody.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	879	7021	4455	1210	1333	23	0	4	0

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	310	ALA	-	linker	UNP P42592
A	311	SER	-	linker	UNP P42592
A	312	GLY	-	linker	UNP P42592
A	313	GLY	-	linker	UNP P42592
A	314	GLY	-	linker	UNP P42592
A	315	SER	-	linker	UNP P42592
A	316	GLY	-	linker	UNP P42592
A	317	GLY	-	linker	UNP P42592
A	318	GLY	-	linker	UNP P42592
A	319	GLY	-	linker	UNP P42592
A	320	SER	-	linker	UNP P42592
A	321	GLY	-	linker	UNP P42592
A	322	GLY	-	linker	UNP P42592
A	323	GLY	-	linker	UNP P42592
A	324	GLY	-	linker	UNP P42592
A	325	SER	-	linker	UNP P42592
A	326	GLY	-	linker	UNP P42592
A	898	HIS	-	expression tag	PDB 6XUX
A	899	HIS	-	expression tag	PDB 6XUX
A	900	HIS	-	expression tag	PDB 6XUX
A	901	HIS	-	expression tag	PDB 6XUX
A	902	HIS	-	expression tag	PDB 6XUX
A	903	HIS	-	expression tag	PDB 6XUX
A	904	GLU	-	expression tag	PDB 6XUX
A	905	PRO	-	expression tag	PDB 6XUX
A	906	GLU	-	expression tag	PDB 6XUX
A	907	ALA	-	expression tag	PDB 6XUX

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Ca 1	0	0

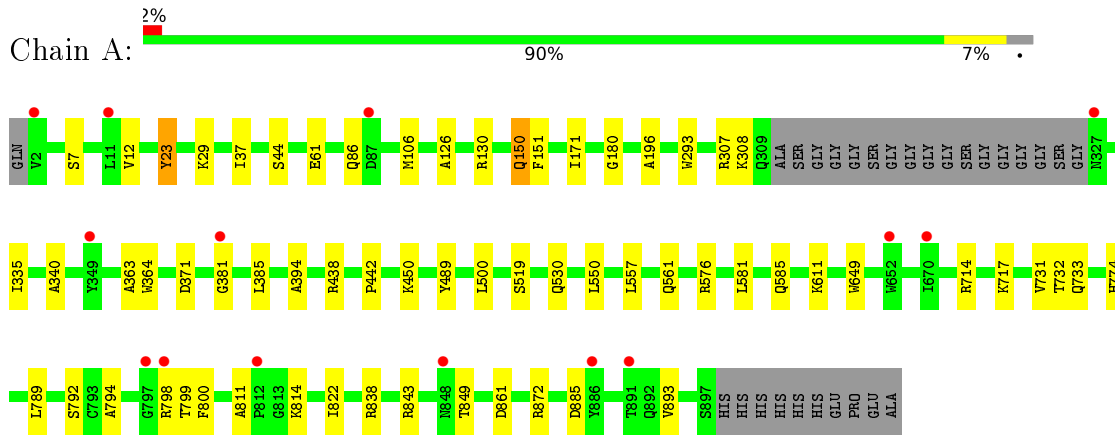
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	650	Total 650	O 650	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: Nanobody,Glucosidase YgjK,Glucosidase YgjK,Nanobody



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.45Å 84.46Å 140.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.14 – 1.90 42.14 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (42.14-1.90) 100.0 (42.14-1.90)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.93 (at 1.89Å)	Xtrriage
Refinement program	PHENIX 1.12_2829	Depositor
R, $R_{free}$	0.189 , 0.230 0.188 , 0.229	Depositor DCC
$R_{free}$ test set	3098 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.2	Xtrriage
Anisotropy	0.931	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 41.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7672	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% 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:  
CA

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	0.71	0/7220	0.77	0/9808

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	7021	0	6737	35	0
2	A	1	0	0	0	0
3	A	650	0	0	9	0
All	All	7672	0	6737	35	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 3.

All (35) 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:12:VAL:HG22	1:A:789:LEU:HD22	1.59	0.82

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:106:MET:SD	3:A:1566:HOH:O	2.43	0.77
1:A:442:PRO:HG3	3:A:1381:HOH:O	1.97	0.64
1:A:714:ARG:HB3	1:A:774:HIS:CE1	2.34	0.62
1:A:340:ALA:HB2	1:A:371:ASP:OD2	2.06	0.55
1:A:381:GLY:HA2	1:A:394:ALA:O	2.06	0.54
1:A:126:ALA:O	1:A:130:ARG:HG3	2.06	0.54
1:A:838:ARG:NH2	1:A:861:ASP:OD2	2.35	0.54
1:A:150:GLN:HB3	1:A:196:ALA:HB1	1.90	0.53
1:A:798:ARG:O	1:A:800:PHE:CE2	2.64	0.51
1:A:581:LEU:O	1:A:585:GLN:HG3	2.11	0.50
1:A:44:SER:HB3	3:A:1181:HOH:O	2.12	0.50
1:A:363:ALA:HA	1:A:442:PRO:O	2.12	0.49
1:A:530:GLN:NE2	3:A:1126:HOH:O	2.46	0.49
1:A:335:ILE:HD11	1:A:611:LYS:HG3	1.94	0.48
1:A:86:GLN:O	3:A:1101:HOH:O	2.20	0.48
1:A:649:TRP:HB2	1:A:717:LYS:HB3	1.96	0.47
1:A:86:GLN:HG2	3:A:1600:HOH:O	2.14	0.46
1:A:335:ILE:HD11	1:A:611:LYS:CG	2.46	0.45
1:A:872:ARG:NH1	1:A:885:ASP:OD2	2.51	0.44
1:A:733:GLN:NE2	3:A:1141:HOH:O	2.51	0.44
1:A:822:ILE:HD13	1:A:843:ARG:HB2	2.00	0.44
1:A:450:LYS:NZ	3:A:1138:HOH:O	2.50	0.43
1:A:364:TRP:HA	1:A:385:LEU:HD12	2.00	0.43
1:A:731:VAL:HG23	1:A:732:THR:HG23	2.00	0.43
1:A:794:ALA:HA	1:A:849:THR:HG22	2.00	0.43
1:A:561:GLN:NE2	3:A:1150:HOH:O	2.53	0.42
1:A:519:SER:HA	1:A:557:LEU:O	2.20	0.42
1:A:438:ARG:CZ	1:A:585:GLN:HG2	2.50	0.41
1:A:61:GLU:CD	1:A:61:GLU:H	2.24	0.41
1:A:7:SER:HB2	1:A:792:SER:OG	2.20	0.41
1:A:171:ILE:HG22	1:A:171:ILE:O	2.20	0.41
1:A:811:ALA:HB3	1:A:814:LYS:HG3	2.02	0.41
1:A:151:PHE:CE2	1:A:180:GLY:HA3	2.56	0.41
1:A:23:TYR:CD1	1:A:37:ILE:HD11	2.56	0.40

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	879/907 (97%)	847 (96%)	32 (4%)	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	732/742 (99%)	720 (98%)	12 (2%)	62 60

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	23	TYR
1	A	29	LYS
1	A	150	GLN
1	A	293	TRP
1	A	307	ARG
1	A	308	LYS
1	A	489	TYR
1	A	500	LEU
1	A	550	LEU
1	A	576	ARG
1	A	799	THR
1	A	893	VAL

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

Mol	Chain	Res	Type
1	A	150	GLN
1	A	569	GLN
1	A	882	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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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	A	879/907 (96%)	0.13	14 (1%) 72 74	13, 24, 43, 68	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	798	ARG	5.5
1	A	797	GLY	3.8
1	A	2	VAL	3.2
1	A	812	PRO	3.1
1	A	349	TYR	3.1
1	A	11	LEU	2.9
1	A	848	ASN	2.7
1	A	670	ILE	2.4
1	A	891	THR	2.3
1	A	327	ASN	2.3
1	A	381	GLY	2.2
1	A	652	TRP	2.2
1	A	87	ASP	2.1
1	A	886	TYR	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 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
2	CA	A	1001	1/1	0.97	0.07	16,16,16,16	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.