Full wwPDB X-ray Structure Validation Report ⓘ

Jan 15, 2024 – 08:08 PM JST

PDB ID	:	8K3K
Title	:	The crystal structure of nanobody Nb4 in complex with receptor binding domain (RBD) of BA.1 Spike protein
Authors	:	Wang, H.Y.; Xu, W.Q.
Deposited on	:	2023-07-16
Resolution	:	2.43 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references ⓘ) were used in the production of this report:

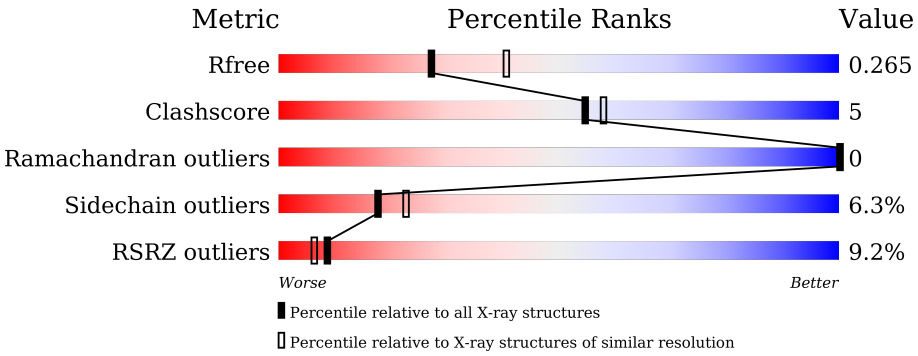
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.43 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	1564 (2.46-2.42)
Clashscore	141614	1631 (2.46-2.42)
Ramachandran outliers	138981	1617 (2.46-2.42)
Sidechain outliers	138945	1617 (2.46-2.42)
RSRZ outliers	127900	1547 (2.46-2.42)

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 >=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 <=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	E	248	9% 64% 13% • 22%
2	D	153	6% 80% 13% • 7%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2696 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
			Total	C	N	O	S			
1	E	194	1556	1005	261	282	8	0	0	0

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	339	ASP	GLY	variant	UNP P0DTC2
E	371	LEU	SER	variant	UNP P0DTC2
E	373	PRO	SER	variant	UNP P0DTC2
E	375	PHE	SER	variant	UNP P0DTC2
E	417	ASN	LYS	variant	UNP P0DTC2
E	440	LYS	ASN	variant	UNP P0DTC2
E	446	SER	GLY	variant	UNP P0DTC2
E	477	ASN	SER	variant	UNP P0DTC2
E	478	LYS	THR	variant	UNP P0DTC2
E	484	ALA	GLU	variant	UNP P0DTC2
E	493	LYS	GLN	variant	UNP P0DTC2
E	496	SER	GLY	variant	UNP P0DTC2
E	498	ARG	GLN	variant	UNP P0DTC2
E	501	TYR	ASN	variant	UNP P0DTC2
E	505	HIS	TYR	variant	UNP P0DTC2
E	532	GLY	-	expression tag	UNP P0DTC2
E	533	THR	-	expression tag	UNP P0DTC2
E	534	GLY	-	expression tag	UNP P0DTC2
E	535	LEU	-	expression tag	UNP P0DTC2
E	536	ASN	-	expression tag	UNP P0DTC2
E	537	ASP	-	expression tag	UNP P0DTC2
E	538	ILE	-	expression tag	UNP P0DTC2
E	539	PHE	-	expression tag	UNP P0DTC2
E	540	GLU	-	expression tag	UNP P0DTC2
E	541	ALA	-	expression tag	UNP P0DTC2
E	542	GLN	-	expression tag	UNP P0DTC2
E	543	LYS	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	544	ILE	-	expression tag	UNP P0DTC2
E	545	GLU	-	expression tag	UNP P0DTC2
E	546	TRP	-	expression tag	UNP P0DTC2
E	547	HIS	-	expression tag	UNP P0DTC2
E	548	GLU	-	expression tag	UNP P0DTC2
E	549	GLY	-	expression tag	UNP P0DTC2
E	550	SER	-	expression tag	UNP P0DTC2
E	551	LEU	-	expression tag	UNP P0DTC2
E	552	GLU	-	expression tag	UNP P0DTC2
E	553	VAL	-	expression tag	UNP P0DTC2
E	554	LEU	-	expression tag	UNP P0DTC2
E	555	PHE	-	expression tag	UNP P0DTC2
E	556	GLN	-	expression tag	UNP P0DTC2
E	557	GLY	-	expression tag	UNP P0DTC2
E	558	PRO	-	expression tag	UNP P0DTC2
E	559	HIS	-	expression tag	UNP P0DTC2
E	560	HIS	-	expression tag	UNP P0DTC2
E	561	HIS	-	expression tag	UNP P0DTC2
E	562	HIS	-	expression tag	UNP P0DTC2
E	563	HIS	-	expression tag	UNP P0DTC2
E	564	HIS	-	expression tag	UNP P0DTC2
E	565	HIS	-	expression tag	UNP P0DTC2
E	566	HIS	-	expression tag	UNP P0DTC2

- Molecule 2 is a protein called Nanobody Nb4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	143	Total	C	N	O	S	0	1	0
			1121	701	192	223	5			

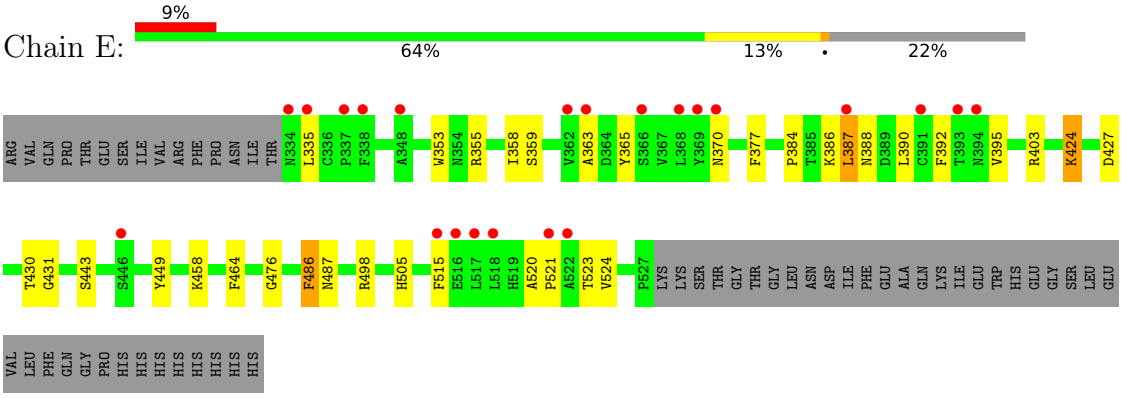
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	10	Total	O	0	0
			10	10		
3	D	9	Total	O	0	0
			9	9		

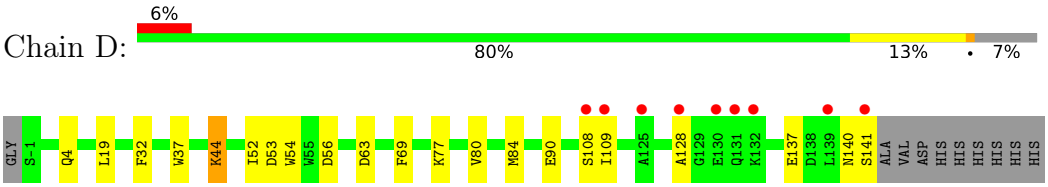
3 Residue-property plots (i)

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



• Molecule 2: Nanobody Nb4



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	76.11Å 76.11Å 251.45Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	35.37 – 2.43 35.37 – 2.43	Depositor EDS
% Data completeness (in resolution range)	99.2 (35.37-2.43) 92.9 (35.37-2.43)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.24 (at 2.42Å)	Xtriage
Refinement program	PHENIX 1.19.2-4158	Depositor
R, R_{free}	0.224 , 0.269 0.223 , 0.265	Depositor DCC
R_{free} test set	1704 reflections (10.01%)	wwPDB-VP
Wilson B-factor (Å^2)	58.7	Xtriage
Anisotropy	0.010	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 51.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	2696	wwPDB-VP
Average B, all atoms (Å^2)	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.45% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	E	0.27	0/1603	0.54	0/2182
2	D	0.27	0/1150	0.63	3/1561 (0.2%)
All	All	0.27	0/2753	0.58	3/3743 (0.1%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	128	ALA	CB-CA-C	6.65	120.07	110.10
2	D	108	SER	N-CA-C	5.80	126.65	111.00
2	D	128	ALA	N-CA-C	-5.68	95.67	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	E	1556	0	1487	18	0
2	D	1121	0	1031	9	0
3	D	9	0	0	0	0
3	E	10	0	0	0	0
All	All	2696	0	2518	27	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 5.

All (27) 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:E:358:ILE:HG12	1:E:395:VAL:HG12	1.63	0.80
1:E:387:LEU:HA	1:E:390:LEU:HG	1.81	0.61
2:D:140:ASN:OD1	2:D:141:SER:N	2.38	0.57
1:E:384:PRO:HA	1:E:387:LEU:HD22	1.90	0.54
1:E:520:ALA:HB1	1:E:521:PRO:HD2	1.90	0.54
1:E:403:ARG:HD2	1:E:505:HIS:HA	1.91	0.53
2:D:52:ILE:HG12	2:D:53:ASP:O	2.09	0.53
1:E:486:PHE:HD1	1:E:487:ASN:H	1.57	0.52
2:D:90:GLU:N	2:D:90:GLU:OE1	2.42	0.51
2:D:69:PHE:CZ	2:D:84:MET:HG2	2.46	0.50
1:E:476:GLY:N	1:E:487:ASN:HB3	2.28	0.48
1:E:392:PHE:HB2	1:E:524:VAL:HB	1.96	0.47
1:E:431:GLY:HA2	1:E:515:PHE:CE2	2.50	0.46
2:D:44:LYS:N	2:D:44:LYS:HD2	2.31	0.46
2:D:37:TRP:CH2	2:D:80:VAL:HG13	2.50	0.45
1:E:486:PHE:HD1	1:E:487:ASN:N	2.15	0.45
1:E:365:TYR:CD1	1:E:365:TYR:N	2.86	0.44
1:E:523:THR:O	1:E:523:THR:OG1	2.29	0.44
1:E:353:TRP:HZ3	1:E:355:ARG:HH11	1.66	0.43
2:D:32:PHE:HB3	2:D:54:TRP:CG	2.54	0.43
2:D:52:ILE:HD11	2:D:56:ASP:HA	2.00	0.43
1:E:458:LYS:HE2	1:E:458:LYS:HB2	1.80	0.42
1:E:449:TYR:OH	1:E:498:ARG:NH2	2.53	0.41
2:D:19:LEU:HD12	2:D:19:LEU:HA	1.86	0.41
1:E:363:ALA:HB1	1:E:365:TYR:HE1	1.85	0.41
1:E:430:THR:HG23	1:E:515:PHE:HB2	2.02	0.41
1:E:424:LYS:HB2	1:E:424:LYS:HE2	1.67	0.41

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	E	192/248 (77%)	179 (93%)	13 (7%)	0	100	100
2	D	142/153 (93%)	137 (96%)	5 (4%)	0	100	100
All	All	334/401 (83%)	316 (95%)	18 (5%)	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	E	169/218 (78%)	157 (93%)	12 (7%)	14	17
2	D	116/124 (94%)	110 (95%)	6 (5%)	23	31
All	All	285/342 (83%)	267 (94%)	18 (6%)	18	23

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	335	LEU
1	E	359	SER
1	E	370	ASN
1	E	377	PHE
1	E	386	LYS
1	E	387	LEU
1	E	388	ASN
1	E	424	LYS
1	E	427	ASP
1	E	443	SER
1	E	464	PHE
1	E	486	PHE
2	D	4	GLN
2	D	44	LYS
2	D	63	ASP
2	D	77	LYS
2	D	109	ILE
2	D	137	GLU

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

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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^{th} 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^2)	Q<0.9
1	E	194/248 (78%)	0.54	22 (11%) 5 3	52, 79, 145, 167	0
2	D	143/153 (93%)	0.20	9 (6%) 20 16	50, 68, 115, 129	0
All	All	337/401 (84%)	0.39	31 (9%) 9 6	50, 73, 135, 167	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	518	LEU	7.1
1	E	369	TYR	5.1
1	E	515	PHE	4.9
2	D	141	SER	4.9
1	E	338	PHE	4.8
1	E	516	GLU	4.3
2	D	130	GLU	4.0
1	E	517	LEU	3.7
1	E	334	ASN	3.7
2	D	132	LYS	3.6
1	E	335	LEU	3.5
1	E	362	VAL	3.3
1	E	391	CYS	3.2
1	E	368	LEU	3.2
1	E	370	ASN	3.1
1	E	393	THR	3.0
1	E	366	SER	2.9
2	D	128	ALA	2.8
1	E	521	PRO	2.8
1	E	394	ASN	2.7
1	E	522	ALA	2.7
1	E	337	PRO	2.7
2	D	125	ALA	2.6
2	D	108	SER	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	363	ALA	2.6
2	D	139	LEU	2.6
1	E	348	ALA	2.5
2	D	131	GLN	2.4
1	E	387	LEU	2.4
2	D	109	ILE	2.1
1	E	446	SER	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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.