Full wwPDB X-ray Structure Validation Report ⓘ

Mar 3, 2025 – 08:36 PM JST

PDB ID	:	8YK4
Title	:	Structure of SARS-CoV-2 RBD and antibody NT-108
Authors	:	Anraku, Y.; Kita, S.; Onodera, T.; Sato, A.; Tadokoro, T.; Adachi, Y.; Ito, S.; Suzuki, T.; Sasaki, J.; Shiwa, N.; Iwata, N.; Nagata, N.; Kazuki, Y.; Oshimura, M.; Sasaki, M.; Orba, Y.; Suzuki, T.; Sawa, H.; Hashiguchi, T.; Fukuhara, H.; Takahashi, Y.; Maenaka, K.
Deposited on	:	2024-03-04
Resolution	:	3.20 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:

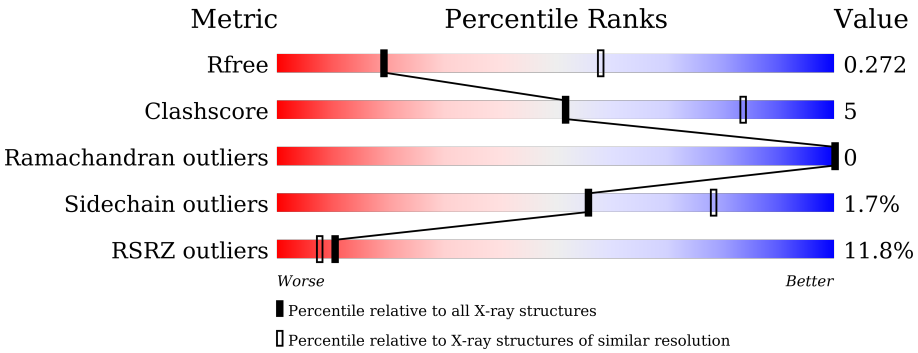
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.2

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 3.20 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}	164625	1370 (3.20-3.20)
Clashscore	180529	1497 (3.20-3.20)
Ramachandran outliers	177936	1479 (3.20-3.20)
Sidechain outliers	177891	1478 (3.20-3.20)
RSRZ outliers	164620	1371 (3.20-3.20)

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	A	282	5% 72% 9% 18%
2	C	215	15% 66% 17% 17%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3179 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 antibody NT-108, single chain Fv fragment.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	230	1761	1111	299	345	6	0	0	0

- Molecule 2 is a protein called Spike protein S1.

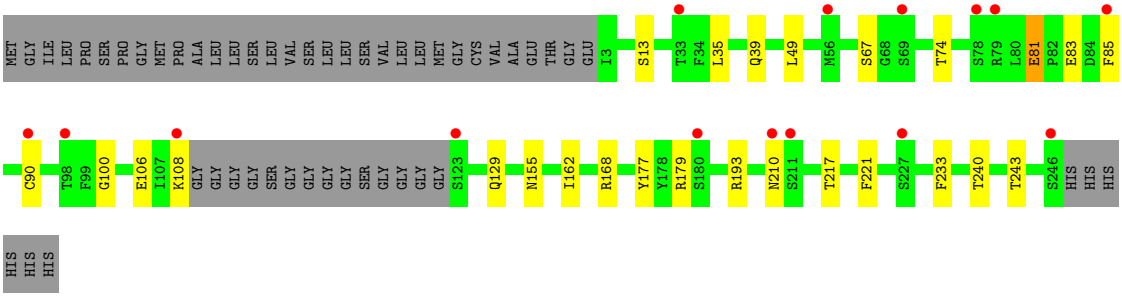
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	178	1418	908	237	267	6	0	0	0

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.

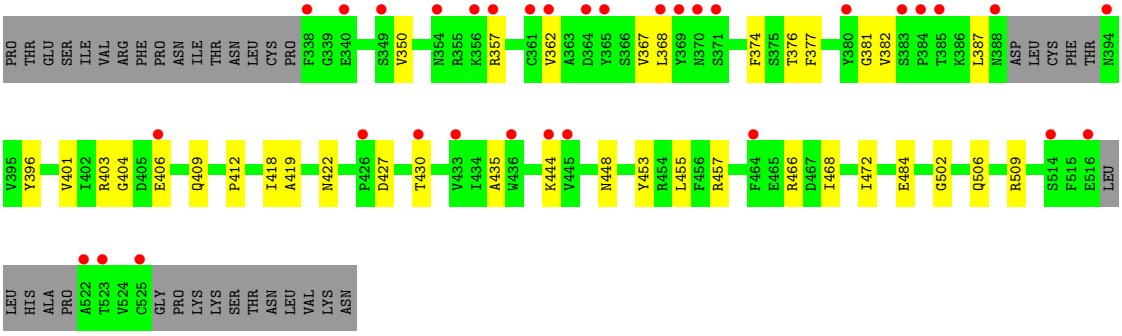
• Molecule 1: antibody NT-108, single chain Fv fragment

Chain A:



• Molecule 2: Spike protein S1

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	55.15Å 149.64Å 253.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.35 – 3.20 48.35 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.3 (48.35-3.20) 99.3 (48.35-3.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.83 (at 3.19Å)	Xtriage
Refinement program	PHENIX 1.20.1	Depositor
R, R_{free}	0.243 , 0.267 0.245 , 0.272	Depositor DCC
R_{free} test set	868 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	58.1	Xtriage
Anisotropy	0.662	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 47.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	3179	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.39% 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	A	0.24	0/1803	0.43	0/2452
2	C	0.24	0/1455	0.41	0/1974
All	All	0.24	0/3258	0.43	0/4426

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

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	1761	0	1707	13	0
2	C	1418	0	1339	18	0
All	All	3179	0	3046	31	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 (31) 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:C:381:GLY:HA3	2:C:430:THR:HA	1.76	0.67
1:A:39:GLN:HB2	1:A:49:LEU:HD11	1.83	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:155:ASN:O	1:A:179:ARG:NH1	2.35	0.60
2:C:412:PRO:HB3	2:C:427:ASP:HA	1.84	0.60
2:C:350:VAL:HG11	2:C:418:ILE:HD12	1.85	0.58
2:C:401:VAL:HG22	2:C:509:ARG:HG2	1.89	0.55
2:C:357:ARG:NH2	2:C:396:TYR:OH	2.40	0.54
1:A:217:THR:HG23	1:A:243:THR:HA	1.94	0.49
1:A:90:CYS:O	1:A:100:GLY:N	2.44	0.49
1:A:67:SER:HG	1:A:74:THR:HG1	1.61	0.48
2:C:444:LYS:H	2:C:448:ASN:HB2	1.79	0.48
1:A:13:SER:HB3	1:A:108:LYS:HE3	1.94	0.48
2:C:376:THR:HB	2:C:435:ALA:HB3	1.97	0.47
2:C:403:ARG:HG2	2:C:404:GLY:H	1.82	0.45
2:C:368:LEU:HB3	2:C:374:PHE:HE2	1.81	0.45
1:A:177:TYR:HD2	1:A:179:ARG:HE	1.62	0.45
2:C:382:VAL:HG11	2:C:387:LEU:HD21	1.97	0.44
1:A:129:GLN:NE2	1:A:240:THR:HG22	2.32	0.44
1:A:193:ARG:HB2	1:A:210:ASN:HB3	2.00	0.42
2:C:466:ARG:HG2	2:C:468:ILE:HD11	2.01	0.42
2:C:502:GLY:O	2:C:506:GLN:HG3	2.19	0.42
1:A:162:ILE:HD11	1:A:233:PHE:CE1	2.55	0.42
1:A:162:ILE:O	1:A:221:PHE:N	2.41	0.41
2:C:418:ILE:HA	2:C:422:ASN:HD22	1.85	0.41
2:C:453:TYR:CE2	2:C:455:LEU:HD23	2.54	0.41
1:A:81:GLU:HB3	1:A:83:GLU:HG2	2.01	0.41
2:C:403:ARG:HB3	2:C:406:GLU:HG2	2.03	0.41
2:C:409:GLN:HB3	2:C:419:ALA:HB2	2.02	0.41
1:A:85:PHE:CD1	1:A:106:GLU:HA	2.56	0.41
2:C:367:VAL:HG13	2:C:368:LEU:HG	2.03	0.40
2:C:472:ILE:HD13	2:C:484:GLU:HB3	2.03	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	226/282 (80%)	212 (94%)	14 (6%)	0	100	100
2	C	172/215 (80%)	159 (92%)	13 (8%)	0	100	100
All	All	398/497 (80%)	371 (93%)	27 (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	194/228 (85%)	191 (98%)	3 (2%)	60	81
2	C	153/188 (81%)	150 (98%)	3 (2%)	50	75
All	All	347/416 (83%)	341 (98%)	6 (2%)	56	78

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	35	LEU
1	A	81	GLU
1	A	168	ARG
2	C	362	VAL
2	C	377	PHE
2	C	457	ARG

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 oligosaccharides 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^{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	A	230/282 (81%)	0.49	15 (6%) 26 18	30, 51, 84, 109	0
2	C	178/215 (82%)	1.24	33 (18%) 4 3	47, 92, 145, 198	0
All	All	408/497 (82%)	0.82	48 (11%) 10 8	30, 67, 135, 198	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	365	TYR	6.7
2	C	369	TYR	5.6
1	A	210	ASN	5.1
1	A	123	SER	4.6
1	A	211	SER	3.9
2	C	394	ASN	3.9
2	C	385	THR	3.8
2	C	522	ALA	3.8
2	C	361	CYS	3.7
1	A	78	SER	3.6
1	A	246	SER	3.6
2	C	338	PHE	3.6
1	A	33	THR	3.3
1	A	108	LYS	3.0
2	C	340	GLU	3.0
1	A	56	MET	3.0
2	C	362	VAL	2.9
2	C	383	SER	2.9
2	C	388	ASN	2.9
2	C	406	GLU	2.8
1	A	180	SER	2.8
1	A	98	THR	2.8
2	C	384	PRO	2.8
1	A	227	SER	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	90	CYS	2.7
2	C	349	SER	2.7
2	C	371	SER	2.7
2	C	426	PRO	2.6
2	C	368	LEU	2.6
2	C	364	ASP	2.6
2	C	514	SER	2.6
2	C	516	GLU	2.5
2	C	433	VAL	2.5
2	C	445	VAL	2.5
2	C	380	TYR	2.5
2	C	354	ASN	2.4
2	C	430	THR	2.4
2	C	525	CYS	2.3
2	C	370	ASN	2.3
2	C	464	PHE	2.3
1	A	85	PHE	2.2
1	A	79	ARG	2.2
2	C	356	LYS	2.2
2	C	523	THR	2.1
2	C	444	LYS	2.1
2	C	357	ARG	2.0
1	A	69	SER	2.0
2	C	436	TRP	2.0

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.