



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 04:35 am BST

PDB ID : 2GKI
Title : Heavy and light chain variable single domains of an anti-DNA binding antibody hydrolyze both double- and single-stranded DNAs without sequence specificity
Authors : Kim, Y.R.; Kim, J.S.; Lee, S.H.; Lee, W.R.; Sohn, J.N.; Chung, Y.C.; Shim, H.K.; Lee, S.C.; Kwon, M.H.; Kim, Y.S.
Deposited on : 2006-04-02
Resolution : 2.88 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

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	:	2.11
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)	:	2.11

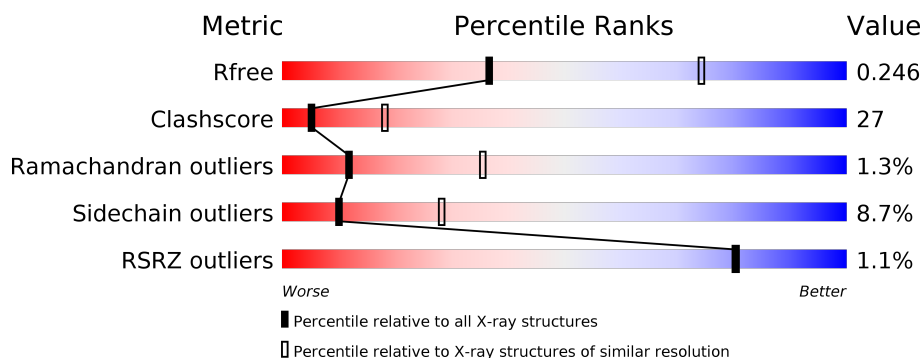
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.88 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	2691 (2.90-2.86)
Clashscore	141614	2947 (2.90-2.86)
Ramachandran outliers	138981	2868 (2.90-2.86)
Sidechain outliers	138945	2871 (2.90-2.86)
RSRZ outliers	127900	2629 (2.90-2.86)

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 ≥ 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	291	
1	B	291	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3754 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 nuclease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	235	Total	C	N	O	S	0	0	0
			1832	1157	303	361	11			
1	B	238	Total	C	N	O	S	0	0	0
			1852	1168	308	365	11			

There are 120 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	CLONING ARTIFACT	GB 8777863
A	2	LYS	-	CLONING ARTIFACT	GB 8777863
A	3	GLN	-	CLONING ARTIFACT	GB 8777863
A	4	SER	-	CLONING ARTIFACT	GB 8777863
A	5	THR	-	CLONING ARTIFACT	GB 8777863
A	6	ILE	-	CLONING ARTIFACT	GB 8777863
A	7	ALA	-	CLONING ARTIFACT	GB 8777863
A	8	LEU	-	CLONING ARTIFACT	GB 8777863
A	9	ALA	-	CLONING ARTIFACT	GB 8777863
A	10	LEU	-	CLONING ARTIFACT	GB 8777863
A	11	LEU	-	CLONING ARTIFACT	GB 8777863
A	12	PRO	-	CLONING ARTIFACT	GB 8777863
A	13	LEU	-	CLONING ARTIFACT	GB 8777863
A	14	LEU	-	CLONING ARTIFACT	GB 8777863
A	15	PHE	-	CLONING ARTIFACT	GB 8777863
A	16	THR	-	CLONING ARTIFACT	GB 8777863
A	17	PRO	-	CLONING ARTIFACT	GB 8777863
A	18	VAL	-	CLONING ARTIFACT	GB 8777863
A	19	THR	-	CLONING ARTIFACT	GB 8777863
A	20	LYS	-	CLONING ARTIFACT	GB 8777863
A	21	ALA	-	CLONING ARTIFACT	GB 8777863
A	22	ARG	-	CLONING ARTIFACT	GB 8777863
A	143	ARG	-	LINKER	GB 8777863
A	144	GLY	-	LINKER	GB 8777863
A	145	GLY	-	LINKER	GB 8777863

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	146	GLY	-	LINKER	GB 8777863
A	147	GLY	-	LINKER	GB 8777863
A	148	SER	-	LINKER	GB 8777863
A	149	GLY	-	LINKER	GB 8777863
A	150	GLY	-	LINKER	GB 8777863
A	151	GLY	-	LINKER	GB 8777863
A	152	GLY	-	LINKER	GB 8777863
A	153	SER	-	LINKER	GB 8777863
A	154	GLY	-	LINKER	GB 8777863
A	155	GLY	-	LINKER	GB 8777863
A	156	GLY	-	LINKER	GB 8777863
A	157	GLY	-	LINKER	GB 8777863
A	158	SER	-	LINKER	GB 8777863
A	160	LEU	ILE	SEE REMARK 999	GB 8777865
A	231	SER	PRO	SEE REMARK 999	GB 8777865
A	272	HIS	-	CLONING ARTIFACT	GB 8777865
A	273	HIS	-	CLONING ARTIFACT	GB 8777865
A	274	HIS	-	CLONING ARTIFACT	GB 8777865
A	275	HIS	-	CLONING ARTIFACT	GB 8777865
A	276	HIS	-	CLONING ARTIFACT	GB 8777865
A	277	GLY	-	CLONING ARTIFACT	GB 8777865
A	278	LEU	-	CLONING ARTIFACT	GB 8777865
A	279	VAL	-	CLONING ARTIFACT	GB 8777865
A	280	PRO	-	CLONING ARTIFACT	GB 8777865
A	281	ARG	-	CLONING ARTIFACT	GB 8777865
A	282	GLY	-	CLONING ARTIFACT	GB 8777865
A	283	SER	-	CLONING ARTIFACT	GB 8777865
A	284	GLY	-	CLONING ARTIFACT	GB 8777865
A	285	ASP	-	CLONING ARTIFACT	GB 8777865
A	286	PRO	-	CLONING ARTIFACT	GB 8777865
A	287	LYS	-	CLONING ARTIFACT	GB 8777865
A	288	ALA	-	CLONING ARTIFACT	GB 8777865
A	289	ASP	-	CLONING ARTIFACT	GB 8777865
A	290	ASN	-	CLONING ARTIFACT	GB 8777865
A	291	LYS	-	CLONING ARTIFACT	GB 8777865
B	1	MET	-	CLONING ARTIFACT	GB 8777863
B	2	LYS	-	CLONING ARTIFACT	GB 8777863
B	3	GLN	-	CLONING ARTIFACT	GB 8777863
B	4	SER	-	CLONING ARTIFACT	GB 8777863
B	5	THR	-	CLONING ARTIFACT	GB 8777863
B	6	ILE	-	CLONING ARTIFACT	GB 8777863
B	7	ALA	-	CLONING ARTIFACT	GB 8777863

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	8	LEU	-	CLONING ARTIFACT	GB 8777863
B	9	ALA	-	CLONING ARTIFACT	GB 8777863
B	10	LEU	-	CLONING ARTIFACT	GB 8777863
B	11	LEU	-	CLONING ARTIFACT	GB 8777863
B	12	PRO	-	CLONING ARTIFACT	GB 8777863
B	13	LEU	-	CLONING ARTIFACT	GB 8777863
B	14	LEU	-	CLONING ARTIFACT	GB 8777863
B	15	PHE	-	CLONING ARTIFACT	GB 8777863
B	16	THR	-	CLONING ARTIFACT	GB 8777863
B	17	PRO	-	CLONING ARTIFACT	GB 8777863
B	18	VAL	-	CLONING ARTIFACT	GB 8777863
B	19	THR	-	CLONING ARTIFACT	GB 8777863
B	20	LYS	-	CLONING ARTIFACT	GB 8777863
B	21	ALA	-	CLONING ARTIFACT	GB 8777863
B	22	ARG	-	CLONING ARTIFACT	GB 8777863
B	143	ARG	-	LINKER	GB 8777863
B	144	GLY	-	LINKER	GB 8777863
B	145	GLY	-	LINKER	GB 8777863
B	146	GLY	-	LINKER	GB 8777863
B	147	GLY	-	LINKER	GB 8777863
B	148	SER	-	LINKER	GB 8777863
B	149	GLY	-	LINKER	GB 8777863
B	150	GLY	-	LINKER	GB 8777863
B	151	GLY	-	LINKER	GB 8777863
B	152	GLY	-	LINKER	GB 8777863
B	153	SER	-	LINKER	GB 8777863
B	154	GLY	-	LINKER	GB 8777863
B	155	GLY	-	LINKER	GB 8777863
B	156	GLY	-	LINKER	GB 8777863
B	157	GLY	-	LINKER	GB 8777863
B	158	SER	-	LINKER	GB 8777863
B	160	LEU	ILE	SEE REMARK 999	GB 8777865
B	231	SER	PRO	SEE REMARK 999	GB 8777865
B	272	HIS	-	CLONING ARTIFACT	GB 8777865
B	273	HIS	-	CLONING ARTIFACT	GB 8777865
B	274	HIS	-	CLONING ARTIFACT	GB 8777865
B	275	HIS	-	CLONING ARTIFACT	GB 8777865
B	276	HIS	-	CLONING ARTIFACT	GB 8777865
B	277	GLY	-	CLONING ARTIFACT	GB 8777865
B	278	LEU	-	CLONING ARTIFACT	GB 8777865
B	279	VAL	-	CLONING ARTIFACT	GB 8777865
B	280	PRO	-	CLONING ARTIFACT	GB 8777865

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	281	ARG	-	CLONING ARTIFACT	GB 8777865
B	282	GLY	-	CLONING ARTIFACT	GB 8777865
B	283	SER	-	CLONING ARTIFACT	GB 8777865
B	284	GLY	-	CLONING ARTIFACT	GB 8777865
B	285	ASP	-	CLONING ARTIFACT	GB 8777865
B	286	PRO	-	CLONING ARTIFACT	GB 8777865
B	287	LYS	-	CLONING ARTIFACT	GB 8777865
B	288	ALA	-	CLONING ARTIFACT	GB 8777865
B	289	ASP	-	CLONING ARTIFACT	GB 8777865
B	290	ASN	-	CLONING ARTIFACT	GB 8777865
B	291	LYS	-	CLONING ARTIFACT	GB 8777865

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	36	Total O 36 36	0	0
2	B	34	Total O 34 34	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α , β , γ	179.74Å 179.74Å 184.23Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.58 – 2.88 43.66 – 2.88	Depositor EDS
% Data completeness (in resolution range)	99.4 (49.58-2.88) 99.5 (43.66-2.88)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.16 (at 2.86Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.212 , 0.242 0.216 , 0.246	Depositor DCC
R_{free} test set	4017 reflections (10.04%)	wwPDB-VP
Wilson B-factor (Å ²)	53.4	Xtriage
Anisotropy	0.066	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 80.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	3754	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.98% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 ⓘ

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	0.35	0/1877	0.65	0/2535
1	B	0.36	0/1898	0.65	0/2563
All	All	0.36	0/3775	0.65	0/5098

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	1832	0	1764	99	1
1	B	1852	0	1779	97	0
2	A	36	0	0	0	0
2	B	34	0	0	2	1
All	All	3754	0	3543	193	2

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 27.

The worst 5 of 193 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:B:34:VAL:HG21	1:B:108:LEU:HD13	1.38	0.99

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:40:VAL:HG22	1:A:108:LEU:HD11	1.51	0.92
1:A:34:VAL:HG21	1:A:108:LEU:HD13	1.59	0.84
1:B:62:LYS:HB2	1:B:65:GLN:HG2	1.61	0.82
1:A:57:HIS:HD2	1:A:69:TRP:HE1	1.28	0.81

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:123:TYR:OH	1:A:123:TYR:OH[12_557]	1.75	0.45
2:B:308:HOH:O	2:B:308:HOH:O[9_766]	2.11	0.09

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	231/291 (79%)	211 (91%)	18 (8%)	2 (1%)	17	45
1	B	234/291 (80%)	206 (88%)	24 (10%)	4 (2%)	9	28
All	All	465/582 (80%)	417 (90%)	42 (9%)	6 (1%)	12	35

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	63	PRO
1	B	175	GLU
1	B	38	ALA
1	B	141	SER
1	A	130	ASP

5.3.2 Protein sidechains ⓘ

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	201/238 (84%)	183 (91%)	18 (9%)	9	26
1	B	203/238 (85%)	186 (92%)	17 (8%)	11	30
All	All	404/476 (85%)	369 (91%)	35 (9%)	10	28

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	254	GLN
1	B	33	LEU
1	B	258	HIS
1	A	258	HIS
1	A	273	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	243	GLN
1	B	57	HIS
1	B	189	ASN
1	A	201	GLN
1	B	185	GLN

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 [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 [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, 95th 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	235/291 (80%)	-0.27	1 (0%) 92 92	23, 44, 79, 140	0
1	B	238/291 (81%)	-0.08	4 (1%) 70 70	21, 49, 97, 140	0
All	All	473/582 (81%)	-0.17	5 (1%) 80 80	21, 46, 89, 140	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	191	ARG	2.5
1	B	108	LEU	2.2
1	B	34	VAL	2.2
1	B	258	HIS	2.0
1	B	257	TYR	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 carbohydrates 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.