



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 11:26 AM UTC

PDB ID : 6KL2 / pdb_00006kl2
Title : Structure of the N-terminal domain of Middle East respiratory syndrome coronavirus nucleocapsid protein
Authors : Hou, M.H.; Wang, Y.S.; Lin, S.M.; Hsu, J.N.
Deposited on : 2019-07-29
Resolution : 2.63 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 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

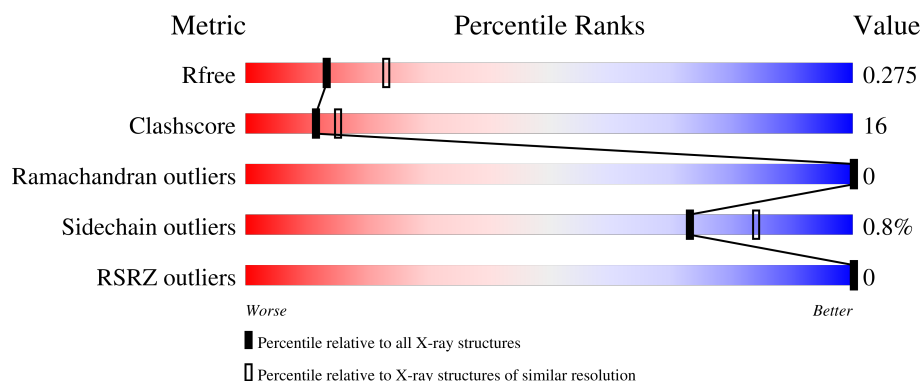
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

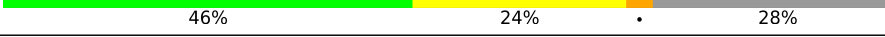
The reported resolution of this entry is 2.63 Å.

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}	180053	2053 (2.66-2.62)
Clashscore	190562	2097 (2.66-2.62)
Ramachandran outliers	187476	2066 (2.66-2.62)
Sidechain outliers	187428	2066 (2.66-2.62)
RSRZ outliers	180081	2052 (2.66-2.62)

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 $\leq 5\%$. 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	144	
1	B	144	
1	C	144	
1	D	144	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3546 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 Nucleoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	116	Total	C	N	O	S	0	0	0
			909	582	160	166	1			
1	B	118	Total	C	N	O	S	0	0	0
			922	589	163	169	1			
1	C	104	Total	C	N	O		0	0	0
			807	518	141	148				
1	D	104	Total	C	N	O		0	0	0
			807	518	141	148				

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	HIS	-	expression tag	UNP A0A0D3MU65
A	23	HIS	-	expression tag	UNP A0A0D3MU65
A	24	HIS	-	expression tag	UNP A0A0D3MU65
A	25	HIS	-	expression tag	UNP A0A0D3MU65
A	26	HIS	-	expression tag	UNP A0A0D3MU65
A	27	HIS	-	expression tag	UNP A0A0D3MU65
A	28	SER	-	expression tag	UNP A0A0D3MU65
A	29	SER	-	expression tag	UNP A0A0D3MU65
A	30	GLY	-	expression tag	UNP A0A0D3MU65
A	31	LEU	-	expression tag	UNP A0A0D3MU65
A	32	VAL	-	expression tag	UNP A0A0D3MU65
A	33	PRO	-	expression tag	UNP A0A0D3MU65
A	34	ARG	-	expression tag	UNP A0A0D3MU65
A	35	GLY	-	expression tag	UNP A0A0D3MU65
A	36	SER	-	expression tag	UNP A0A0D3MU65
A	37	HIS	-	expression tag	UNP A0A0D3MU65
A	38	MET	-	expression tag	UNP A0A0D3MU65
B	22	HIS	-	expression tag	UNP A0A0D3MU65
B	23	HIS	-	expression tag	UNP A0A0D3MU65
B	24	HIS	-	expression tag	UNP A0A0D3MU65
B	25	HIS	-	expression tag	UNP A0A0D3MU65

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	26	HIS	-	expression tag	UNP A0A0D3MU65
B	27	HIS	-	expression tag	UNP A0A0D3MU65
B	28	SER	-	expression tag	UNP A0A0D3MU65
B	29	SER	-	expression tag	UNP A0A0D3MU65
B	30	GLY	-	expression tag	UNP A0A0D3MU65
B	31	LEU	-	expression tag	UNP A0A0D3MU65
B	32	VAL	-	expression tag	UNP A0A0D3MU65
B	33	PRO	-	expression tag	UNP A0A0D3MU65
B	34	ARG	-	expression tag	UNP A0A0D3MU65
B	35	GLY	-	expression tag	UNP A0A0D3MU65
B	36	SER	-	expression tag	UNP A0A0D3MU65
B	37	HIS	-	expression tag	UNP A0A0D3MU65
B	38	MET	-	expression tag	UNP A0A0D3MU65
C	22	HIS	-	expression tag	UNP A0A0D3MU65
C	23	HIS	-	expression tag	UNP A0A0D3MU65
C	24	HIS	-	expression tag	UNP A0A0D3MU65
C	25	HIS	-	expression tag	UNP A0A0D3MU65
C	26	HIS	-	expression tag	UNP A0A0D3MU65
C	27	HIS	-	expression tag	UNP A0A0D3MU65
C	28	SER	-	expression tag	UNP A0A0D3MU65
C	29	SER	-	expression tag	UNP A0A0D3MU65
C	30	GLY	-	expression tag	UNP A0A0D3MU65
C	31	LEU	-	expression tag	UNP A0A0D3MU65
C	32	VAL	-	expression tag	UNP A0A0D3MU65
C	33	PRO	-	expression tag	UNP A0A0D3MU65
C	34	ARG	-	expression tag	UNP A0A0D3MU65
C	35	GLY	-	expression tag	UNP A0A0D3MU65
C	36	SER	-	expression tag	UNP A0A0D3MU65
C	37	HIS	-	expression tag	UNP A0A0D3MU65
C	38	MET	-	expression tag	UNP A0A0D3MU65
D	22	HIS	-	expression tag	UNP A0A0D3MU65
D	23	HIS	-	expression tag	UNP A0A0D3MU65
D	24	HIS	-	expression tag	UNP A0A0D3MU65
D	25	HIS	-	expression tag	UNP A0A0D3MU65
D	26	HIS	-	expression tag	UNP A0A0D3MU65
D	27	HIS	-	expression tag	UNP A0A0D3MU65
D	28	SER	-	expression tag	UNP A0A0D3MU65
D	29	SER	-	expression tag	UNP A0A0D3MU65
D	30	GLY	-	expression tag	UNP A0A0D3MU65
D	31	LEU	-	expression tag	UNP A0A0D3MU65
D	32	VAL	-	expression tag	UNP A0A0D3MU65
D	33	PRO	-	expression tag	UNP A0A0D3MU65

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	34	ARG	-	expression tag	UNP A0A0D3MU65
D	35	GLY	-	expression tag	UNP A0A0D3MU65
D	36	SER	-	expression tag	UNP A0A0D3MU65
D	37	HIS	-	expression tag	UNP A0A0D3MU65
D	38	MET	-	expression tag	UNP A0A0D3MU65

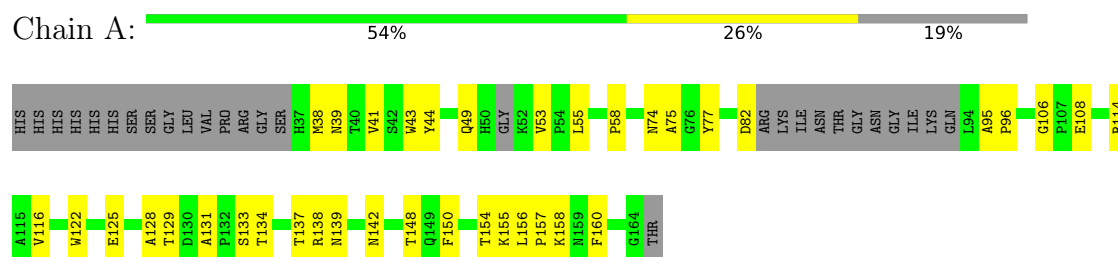
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	23	Total O 23 23	0	0
2	B	33	Total O 33 33	0	0
2	C	25	Total O 25 25	0	0
2	D	20	Total O 20 20	0	0

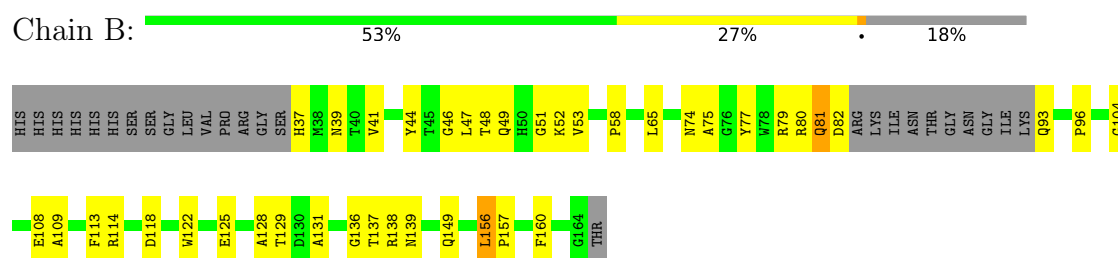
3 Residue-property plots

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.

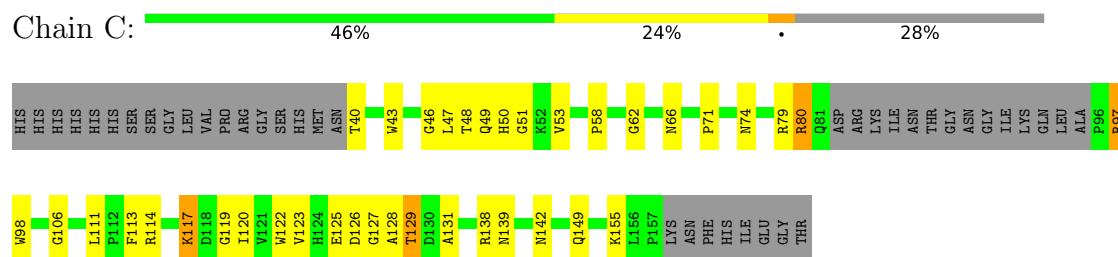
• Molecule 1: Nucleoprotein



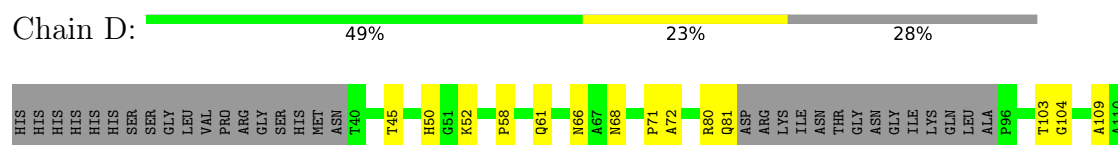
• Molecule 1: Nucleoprotein



• Molecule 1: Nucleoprotein



• Molecule 1: Nucleoprotein



L111	P112	F113	R114	A115	V116	K117	D118	G119	I120	V121	W122	E125	T129	S133	T134	R138	A145	T148	T154	K155	L156	P157	LYS	ASN	PHE	HIS	ILE	GLU	THR
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	35.60Å 109.64Å 91.99Å 90.00° 101.23° 90.00°	Depositor
Resolution (Å)	23.47 – 2.63 23.47 – 2.63	Depositor EDS
% Data completeness (in resolution range)	92.9 (23.47-2.63) 92.8 (23.47-2.63)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.36 (at 2.64Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.280 , 0.290 (Not available) , 0.275	Depositor DCC
R_{free} test set	1915 reflections (9.25%)	wwPDB-VP
Wilson B-factor (Å ²)	43.1	Xtriage
Anisotropy	0.145	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 13.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.438 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	3546	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.89% 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 [i](#)

5.1 Standard geometry [i](#)

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	0.76	1/941 (0.1%)	1.04	3/1289 (0.2%)
1	B	0.88	3/955 (0.3%)	1.08	5/1309 (0.4%)
1	C	0.85	1/837 (0.1%)	1.18	10/1150 (0.9%)
1	D	0.80	3/837 (0.4%)	1.05	4/1150 (0.3%)
All	All	0.83	8/3570 (0.2%)	1.09	22/4898 (0.4%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	53	VAL	CA-C	6.27	1.59	1.53
1	B	58	PRO	CA-C	6.10	1.55	1.51
1	B	81	GLN	CA-C	-5.39	1.46	1.52
1	C	66	ASN	C-O	-5.37	1.17	1.23
1	D	120	ILE	C-N	-5.22	1.27	1.33

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	53	VAL	N-CA-C	8.89	117.66	108.95
1	C	128	ALA	CA-C-N	7.87	134.74	122.94
1	C	128	ALA	C-N-CA	7.87	134.74	122.94
1	D	156	LEU	N-CA-C	-7.78	99.93	110.36
1	C	106	GLY	CA-C-N	7.45	127.09	119.56

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	909	0	854	31	0
1	B	922	0	866	28	0
1	C	807	0	762	35	0
1	D	807	0	762	18	0
2	A	23	0	0	2	0
2	B	33	0	0	6	0
2	C	25	0	0	8	1
2	D	20	0	0	0	1
All	All	3546	0	3244	107	1

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 16.

The worst 5 of 107 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:C:48:THR:HB	1:C:97:ARG:NE	1.47	1.28
1:C:48:THR:HA	1:C:97:ARG:HG2	1.28	1.08
1:C:48:THR:HB	1:C:97:ARG:HE	1.01	1.06
1:C:48:THR:CA	1:C:97:ARG:HG2	1.94	0.97
1:C:48:THR:CB	1:C:97:ARG:HE	1.79	0.94

All (1) 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)
2:C:216:HOH:O	2:D:216:HOH:O[2_556]	2.16	0.04

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	110/144 (76%)	107 (97%)	3 (3%)	0	100	100
1	B	114/144 (79%)	113 (99%)	1 (1%)	0	100	100
1	C	100/144 (69%)	98 (98%)	2 (2%)	0	100	100
1	D	100/144 (69%)	99 (99%)	1 (1%)	0	100	100
All	All	424/576 (74%)	417 (98%)	7 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	94/117 (80%)	93 (99%)	1 (1%)	65	79
1	B	95/117 (81%)	93 (98%)	2 (2%)	47	67
1	C	83/117 (71%)	83 (100%)	0	100	100
1	D	83/117 (71%)	83 (100%)	0	100	100
All	All	355/468 (76%)	352 (99%)	3 (1%)	73	84

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	158	LYS
1	B	37	HIS
1	B	149	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	73	GLN
1	C	50	HIS
1	D	81	GLN
1	D	139	ASN
1	D	142	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 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 [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	116/144 (80%)	-1.27	0 100 100	24, 35, 56, 65	0
1	B	118/144 (81%)	-1.22	0 100 100	23, 38, 59, 71	0
1	C	104/144 (72%)	-1.15	0 100 100	32, 43, 59, 66	0
1	D	104/144 (72%)	-1.16	0 100 100	28, 40, 58, 67	0
All	All	442/576 (76%)	-1.20	0 100 100	23, 39, 58, 71	0

There are no RSRZ outliers to report.

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 oligosaccharides 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.