



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 11:47 PM UTC

PDB ID : 5N2U / pdb_00005n2u
Title : Influenza D virus nucleoprotein
Authors : Labaronne, A.; Ruigrok, R.W.H.; Crepin, T.
Deposited on : 2017-02-08
Resolution : 2.40 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

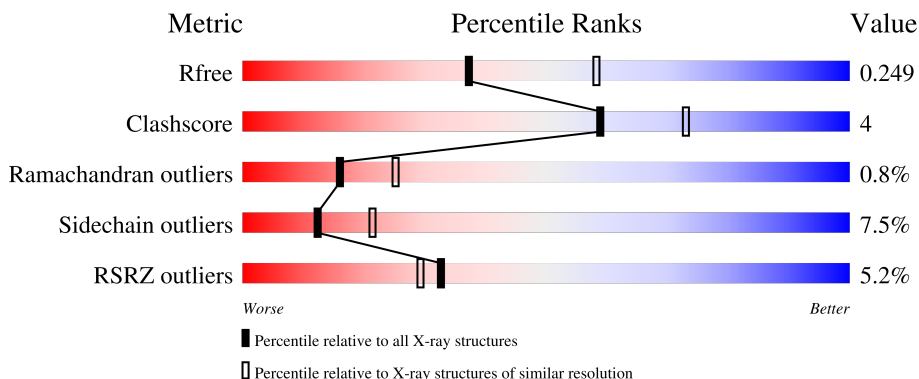
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	4912 (2.40-2.40)
Clashscore	190562	5391 (2.40-2.40)
Ramachandran outliers	187476	5320 (2.40-2.40)
Sidechain outliers	187428	5321 (2.40-2.40)
RSRZ outliers	180081	4916 (2.40-2.40)

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	552	7% 65% 18% • 15%
1	B	552	4% 67% 15% • 17%
1	C	552	3% 72% 11% • 15%
1	D	552	4% 74% 11% • 14%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14881 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	469	Total	C	N	O	S	0	0	0
			3674	2329	635	684	26			
1	B	460	Total	C	N	O	S	0	0	0
			3591	2279	614	672	26			
1	C	467	Total	C	N	O	S	0	0	0
			3649	2314	628	681	26			
1	D	476	Total	C	N	O	S	0	0	0
			3732	2367	647	692	26			

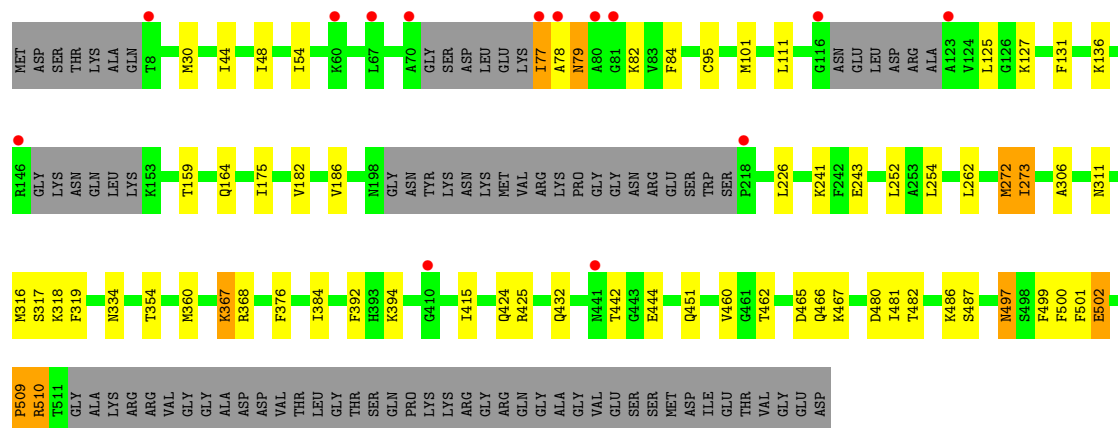
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	45	Total	O	0	0
			45	45		
2	B	56	Total	O	0	0
			56	56		
2	C	86	Total	O	0	0
			86	86		
2	D	48	Total	O	0	0
			48	48		

GLY
ALA
ASP
ASP
VAL
THR
LEU
GLY
THR
SER
GLN
PRO
LYS
LYS
ARG
GLY
ARG
GLN
GLY
ALA
GLY
VAL
GLU
SER
SER
SER
MET
ASP
ILE
GLU
THR
VAL
GLY
GLU
ASP

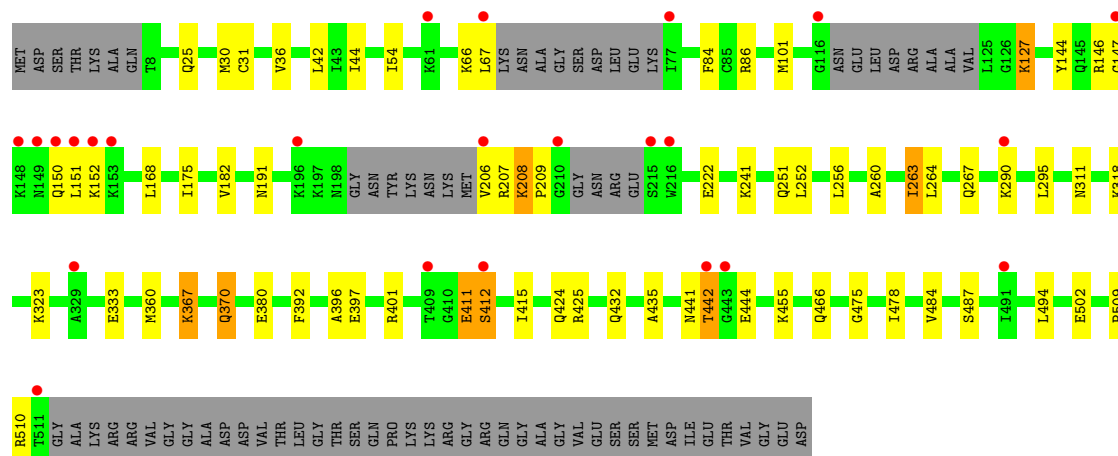
● Molecule 1: Nucleoprotein

Chain C:  3% 72% 11% 15%



● Molecule 1: Nucleoprotein

Chain D:  4% 74% 11% 14%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	75.17Å 85.17Å 103.39Å 91.20° 101.94° 101.02°	Depositor
Resolution (Å)	39.71 – 2.40 39.71 – 2.40	Depositor EDS
% Data completeness (in resolution range)	91.4 (39.71-2.40) 91.4 (39.71-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.44 (at 2.39Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.202 , 0.252 (Not available) , 0.249	Depositor DCC
R_{free} test set	4331 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å ²)	42.8	Xtriage
Anisotropy	0.195	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 46.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	14881	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.61% 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	0.84	3/3736 (0.1%)	1.45	22/5016 (0.4%)
1	B	0.83	3/3649 (0.1%)	1.43	15/4898 (0.3%)
1	C	0.80	0/3708	1.41	17/4976 (0.3%)
1	D	0.80	0/3793	1.38	16/5088 (0.3%)
All	All	0.82	6/14886 (0.0%)	1.42	70/19978 (0.4%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	101	MET	SD-CE	-7.45	1.60	1.79
1	B	278	MET	SD-CE	-6.49	1.63	1.79
1	A	53	MET	SD-CE	-5.92	1.64	1.79
1	A	278	MET	SD-CE	-5.70	1.65	1.79
1	B	360	MET	SD-CE	-5.26	1.66	1.79

The worst 5 of 70 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	263	ILE	N-CA-CB	7.95	118.16	111.64
1	B	113	ASP	CA-CB-CG	7.49	120.09	112.60
1	B	388	SER	CA-C-N	6.76	134.45	121.54
1	B	388	SER	C-N-CA	6.76	134.45	121.54
1	A	374	GLY	CA-C-N	6.65	130.09	120.38

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	3674	0	3715	54	0
1	B	3591	0	3631	42	0
1	C	3649	0	3701	24	0
1	D	3732	0	3789	21	0
2	A	45	0	0	0	0
2	B	56	0	0	0	0
2	C	86	0	0	0	0
2	D	48	0	0	0	0
All	All	14881	0	14836	129	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 4.

The worst 5 of 129 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:125:LEU:HB3	1:A:128:ARG:HB2	1.64	0.78
1:B:272:MET:HE1	1:B:274:MET:HG3	1.70	0.74
1:A:82:LYS:HD2	1:A:101:MET:HE2	1.72	0.71
1:A:442:THR:HG21	1:B:256:LEU:HD13	1.73	0.70
1:B:360:MET:HA	1:B:394:LYS:HB2	1.77	0.66

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	461/552 (84%)	429 (93%)	27 (6%)	5 (1%)	11 18
1	B	448/552 (81%)	412 (92%)	32 (7%)	4 (1%)	14 22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	457/552 (83%)	440 (96%)	14 (3%)	3 (1%)	18	28
1	D	465/552 (84%)	440 (95%)	22 (5%)	3 (1%)	21	32
All	All	1831/2208 (83%)	1721 (94%)	95 (5%)	15 (1%)	16	25

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	147	GLY
1	A	467	LYS
1	D	412	SER
1	A	59	ASP
1	B	409	THR

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	395/462 (86%)	366 (93%)	29 (7%)	13	22
1	B	388/462 (84%)	355 (92%)	33 (8%)	10	16
1	C	394/462 (85%)	369 (94%)	25 (6%)	16	29
1	D	403/462 (87%)	371 (92%)	32 (8%)	11	19
All	All	1580/1848 (86%)	1461 (92%)	119 (8%)	12	21

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	415	ILE
1	D	397	GLU
1	C	243	GLU
1	D	367	LYS
1	D	502	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	289	ASN
1	D	459	GLN
1	D	311	ASN
1	D	393	HIS
1	D	488	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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	D	1
1	B	1
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	197:LYS	C	198:ASN	N	4.81

Continued on next page...

Continued from previous page...

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	502:GLU	C	503:PHE	N	3.25
1	A	392:PHE	C	393:HIS	N	2.99

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	469/552 (84%)	0.66	37 (7%)	18 15	36, 55, 91, 114	0
1	B	460/552 (83%)	0.49	22 (4%)	35 32	35, 51, 81, 108	0
1	C	467/552 (84%)	0.35	14 (2%)	52 48	32, 49, 77, 103	0
1	D	476/552 (86%)	0.48	24 (5%)	34 30	33, 52, 84, 109	0
All	All	1872/2208 (84%)	0.49	97 (5%)	33 29	32, 52, 84, 114	0

The worst 5 of 97 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	124	VAL	7.1
1	A	123	ALA	6.5
1	A	122	ALA	6.2
1	A	147	GLY	5.7
1	D	215	SER	5.4

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.