



wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Aug 17, 2017 – 04:34 PM EDT

PDB ID : 2YMN
EMDB ID: : EMD-2209
Title : Organization of the Influenza Virus Replication Machinery
Authors : Moeller, A.; Kirchdoerfer, R.N.; Potter, C.S.; Carragher, B.; Wilson, I.A.
Deposited on : unknown
Resolution : 20.00 Å(reported)
Based on PDB ID : 2IQH

This is a wwPDB/EMDatabank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029824

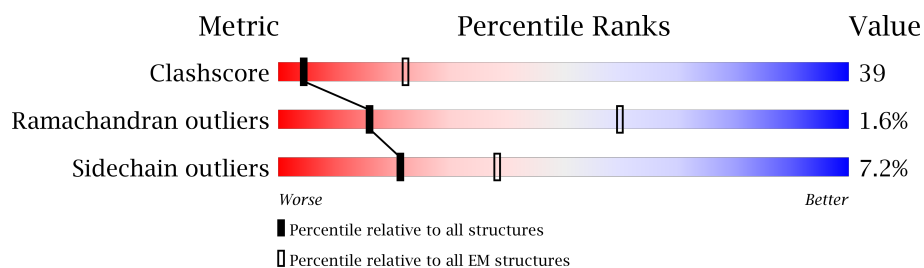
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 20.00 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)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	498	
1	B	498	
1	C	498	
1	D	498	
1	E	498	
1	F	498	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 18994 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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.

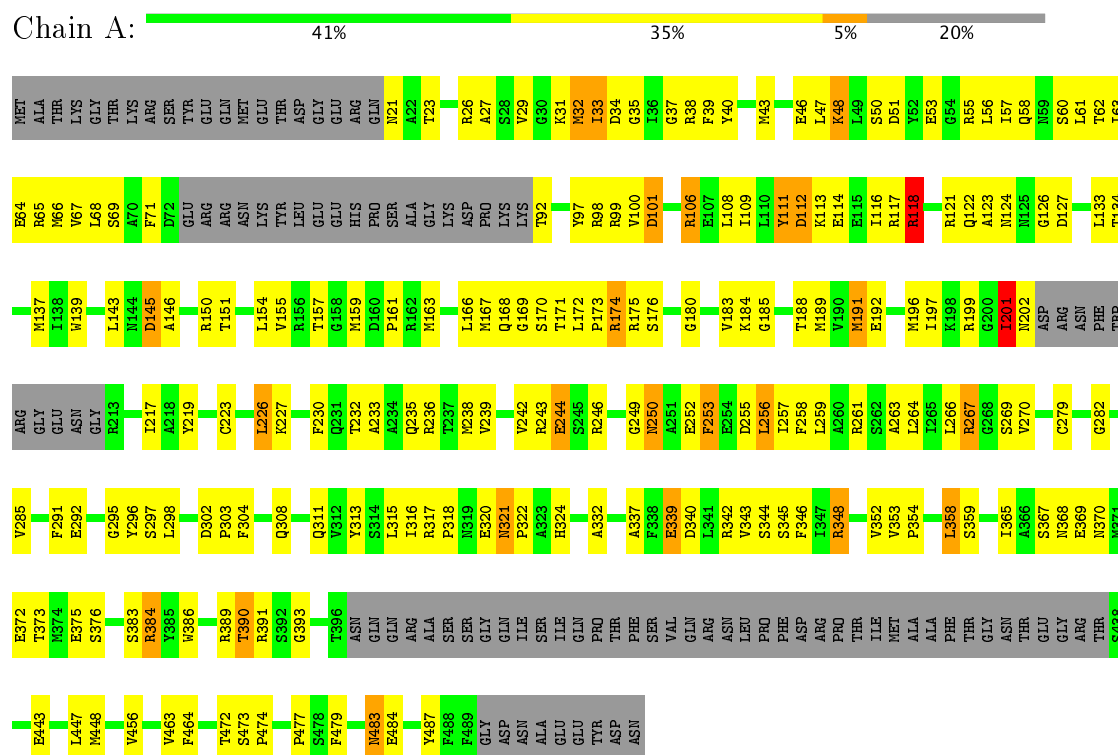
- Molecule 1 is a protein called NUCLEOPROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	399	Total	C	N	O	S	0	0
			3151	1954	586	586	25		
1	B	399	Total	C	N	O	S	0	0
			3151	1954	586	586	25		
1	C	409	Total	C	N	O	S	0	0
			3239	2007	605	602	25		
1	D	399	Total	C	N	O	S	0	0
			3151	1954	586	586	25		
1	E	399	Total	C	N	O	S	0	0
			3151	1954	586	586	25		
1	F	399	Total	C	N	O	S	0	0
			3151	1954	586	586	25		

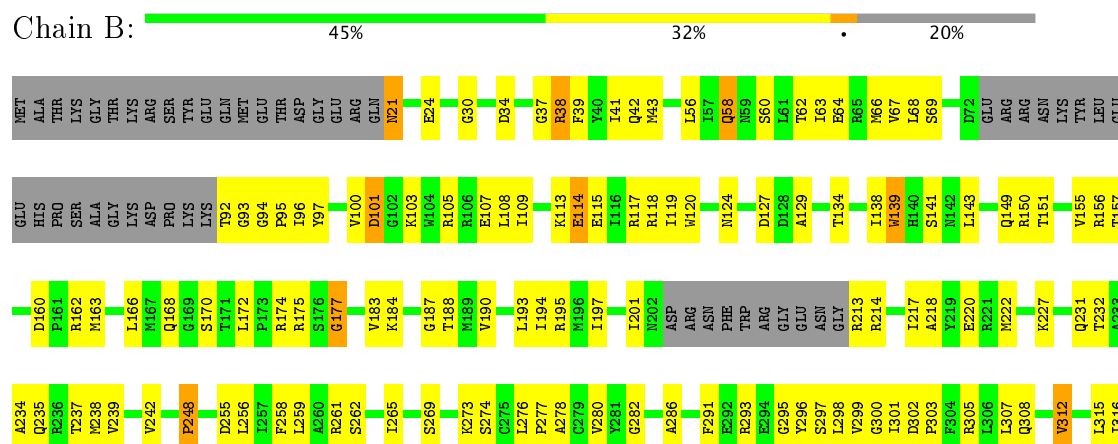
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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. 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. 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.

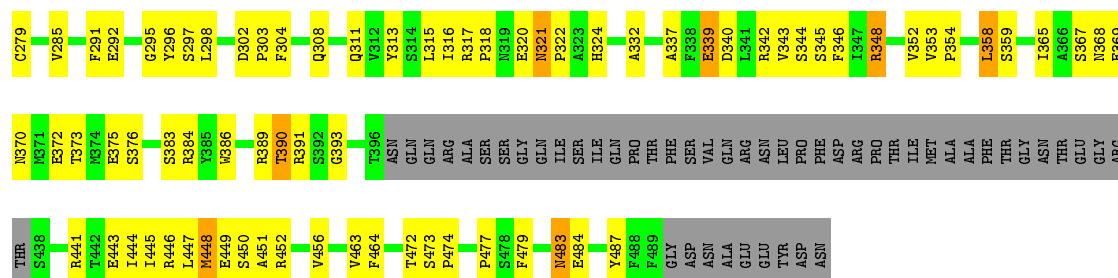
• Molecule 1: NUCLEOPROTEIN



• Molecule 1: NUCLEOPROTEIN

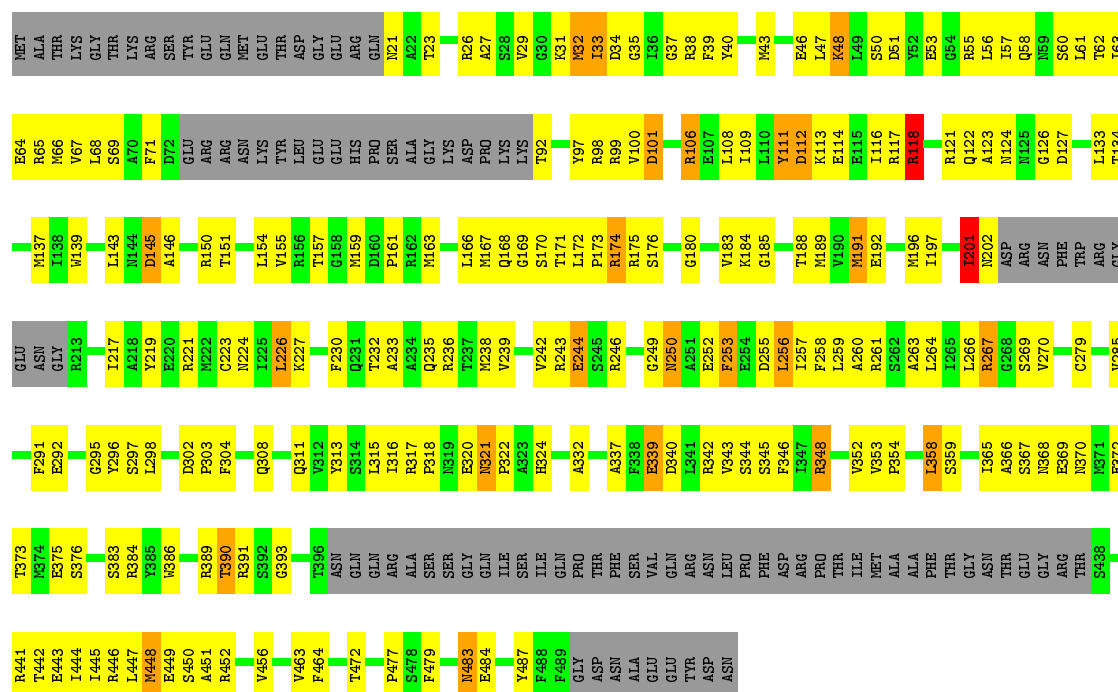






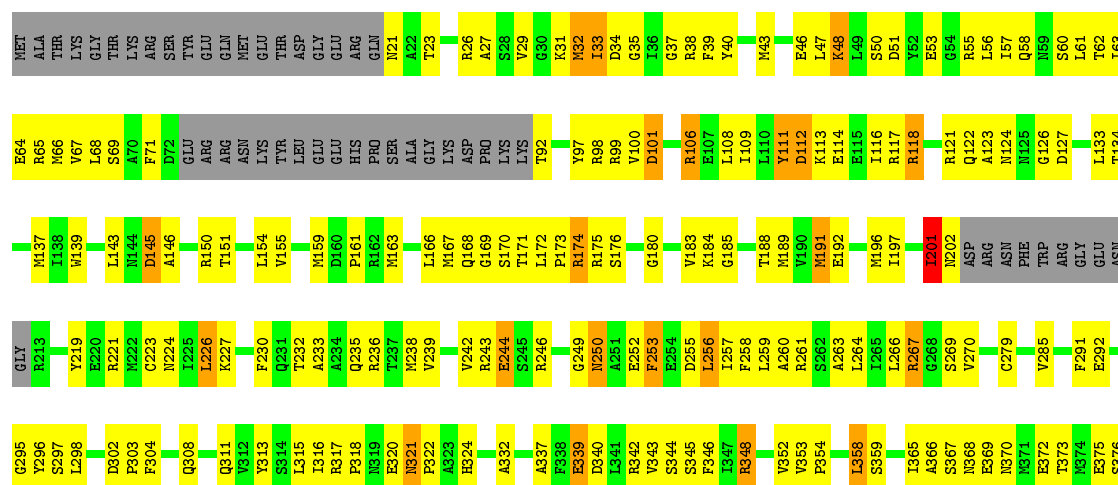
• Molecule 1: NUCLEOPROTEIN

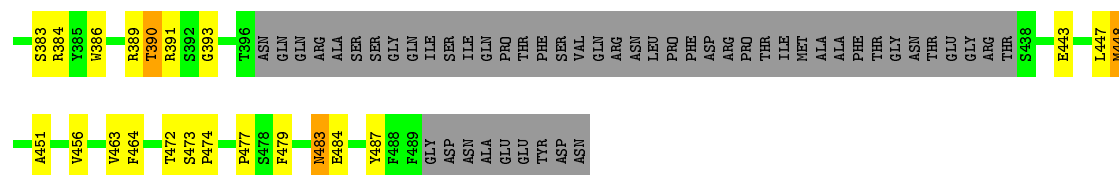
Chain E: 39% 36% 5% 20%



• Molecule 1: NUCLEOPROTEIN

Chain F: 40% 35% 5% 20%





4 Experimental information

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	HELICAL, twist=Not provided°, rise=Not provided Å, axial sym=Not provided	Depositor
Number of segments used	31573	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	WHOLEIMAGE	Depositor
Microscope	FEI TECNAI 20	Depositor
Voltage (kV)	120	Depositor
Electron dose ($e^-/\text{\AA}^2$)	15	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	50000	Depositor
Image detector	GENERIC CCD	Depositor

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	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.92	1/3200 (0.0%)	0.92	8/4299 (0.2%)
1	B	0.97	3/3200 (0.1%)	0.92	3/4299 (0.1%)
1	C	0.84	4/3292 (0.1%)	0.88	4/4424 (0.1%)
1	D	0.92	1/3200 (0.0%)	0.92	6/4299 (0.1%)
1	E	0.92	1/3200 (0.0%)	0.92	7/4299 (0.2%)
1	F	0.92	1/3200 (0.0%)	0.92	6/4299 (0.1%)
All	All	0.92	11/19292 (0.1%)	0.91	34/25919 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	D	0	1
1	E	0	1
1	F	0	1
All	All	0	4

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	201	ILE	C-N	-13.48	1.03	1.34
1	C	202	ASN	C-N	-7.92	1.15	1.34
1	C	201	ILE	C-N	-7.91	1.15	1.34
1	B	312	VAL	CB-CG2	-5.83	1.40	1.52
1	C	247	ASN	CB-CG	5.77	1.64	1.51

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	201	ILE	O-C-N	-6.22	112.74	122.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	201	ILE	O-C-N	-6.22	112.75	122.70
1	A	201	ILE	O-C-N	-6.20	112.78	122.70
1	F	201	ILE	O-C-N	-6.19	112.79	122.70
1	C	201	ILE	O-C-N	-6.14	112.88	122.70

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	201	ILE	Mainchain
1	D	201	ILE	Mainchain
1	E	201	ILE	Mainchain
1	F	201	ILE	Mainchain

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	3151	0	3148	210	0
1	B	3151	0	3145	192	0
1	C	3239	0	3220	257	0
1	D	3151	0	3146	340	0
1	E	3151	0	3141	398	0
1	F	3151	0	3145	297	0
All	All	18994	0	18945	1463	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 39.

The worst 5 of 1463 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:447:LEU:HD21	1:F:367:SER:CB	1.13	1.61
1:D:446:ARG:H	1:E:368:ASN:CB	1.01	1.60
1:D:447:LEU:HD21	1:E:367:SER:CB	1.14	1.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:447:LEU:CD2	1:E:367:SER:HB2	1.27	1.59
1:E:446:ARG:H	1:F:368:ASN:CB	1.04	1.58

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 PDB entries followed by that with respect to all EM entries.

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	391/498 (78%)	339 (87%)	47 (12%)	5 (1%)	14	56
1	B	391/498 (78%)	341 (87%)	41 (10%)	9 (2%)	7	43
1	C	403/498 (81%)	332 (82%)	63 (16%)	8 (2%)	9	46
1	D	391/498 (78%)	339 (87%)	47 (12%)	5 (1%)	14	56
1	E	391/498 (78%)	339 (87%)	47 (12%)	5 (1%)	14	56
1	F	391/498 (78%)	339 (87%)	47 (12%)	5 (1%)	14	56
All	All	2358/2988 (79%)	2029 (86%)	292 (12%)	37 (2%)	16	51

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	203	ASP
1	C	395	ASN
1	C	486	SER
1	A	233	ALA
1	A	472	THR

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 PDB entries followed by that with respect to all EM

entries.

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	338/422 (80%)	313 (93%)	25 (7%)	16	48
1	B	338/422 (80%)	318 (94%)	20 (6%)	23	55
1	C	346/422 (82%)	319 (92%)	27 (8%)	15	46
1	D	338/422 (80%)	313 (93%)	25 (7%)	16	48
1	E	338/422 (80%)	313 (93%)	25 (7%)	16	48
1	F	338/422 (80%)	313 (93%)	25 (7%)	16	48
All	All	2036/2532 (80%)	1889 (93%)	147 (7%)	21	49

5 of 147 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	390	THR
1	D	174	ARG
1	F	244	GLU
1	C	446	ARG
1	D	33	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 63 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	309	ASN
1	D	224	ASN
1	F	319	ASN
1	C	319	ASN
1	C	324	HIS

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	C	2
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	201:ILE	C	202:ASN	N	1.15
1	C	202:ASN	C	203:ASP	N	1.15
1	B	201:ILE	C	202:ASN	N	1.03