Full wwPDB X-ray Structure Validation Report ⓘ

Apr 12, 2022 – 07:10 pm BST

PDB ID	:	7PNO
Title	:	C terminal domain of Nipah Virus Phosphoprotein fused to the Ntail alpha more of the Nucleoprotein.
Authors	:	Bourhis, J.M.; Yabukaski, F.; Tarbouriech, N.; Jamin, M.
Deposited on	:	2021-09-07
Resolution	:	2.79 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 following versions of software and data (see references ⓘ) were used in the production of this report:

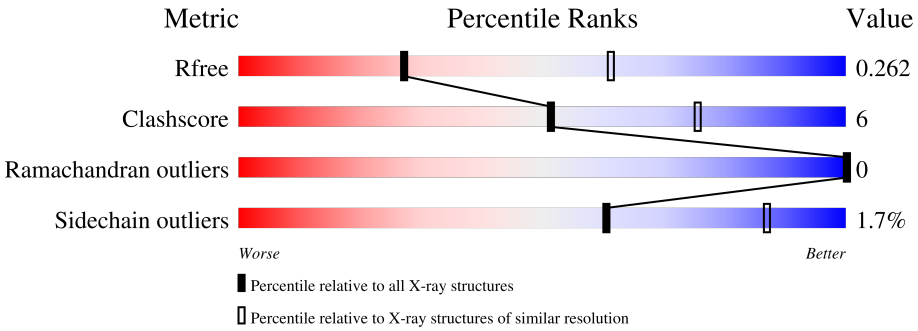
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.27
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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 2.79 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.



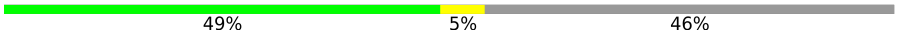
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

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%

Mol	Chain	Length	Quality of chain
1	A	55	82% 7% 11%
1	C	55	73% 16% 11%
1	E	55	73% 16% 11%
1	G	55	73% 16% 11%
1	I	55	71% 16% • 11%
1	K	55	75% 13% • 11%
1	M	55	71% 16% • 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	39	
2	D	39	
2	F	39	
2	H	39	
2	J	39	
2	L	39	
2	N	39	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3798 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 Phosphoprotein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	49	Total	C	N	O	0	0	0
			395	241	71	83			
1	C	49	Total	C	N	O	0	0	0
			395	241	71	83			
1	E	49	Total	C	N	O	0	0	0
			395	241	71	83			
1	G	49	Total	C	N	O	0	0	0
			395	241	71	83			
1	I	49	Total	C	N	O	0	0	0
			395	241	71	83			
1	K	49	Total	C	N	O	0	0	0
			395	241	71	83			
1	M	49	Total	C	N	O	0	0	0
			395	241	71	83			

- Molecule 2 is a protein called alpha MoRE of Nipah virus Nucleoprotein tail.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	23	Total	C	N	O	0	0	0
			161	96	33	32			
2	D	24	Total	C	N	O	0	0	0
			166	99	34	33			
2	F	23	Total	C	N	O	0	0	0
			160	96	33	31			
2	H	21	Total	C	N	O	0	0	0
			150	90	31	29			
2	J	14	Total	C	N	O	0	0	0
			108	66	23	19			
2	L	18	Total	C	N	O	0	0	0
			128	77	27	24			
2	N	20	Total	C	N	O	0	0	0
			146	88	30	28			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total O 2 2	0	0
3	C	3	Total O 3 3	0	0
3	E	2	Total O 2 2	0	0
3	G	1	Total O 1 1	0	0
3	I	2	Total O 2 2	0	0
3	K	2	Total O 2 2	0	0
3	M	2	Total O 2 2	0	0

3 Residue-property plots (i)

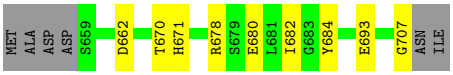
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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: Phosphoprotein

Chain A:

• Molecule 1: Phosphoprotein

Chain C:



• Molecule 1: Phosphoprotein

Chain E:

• Molecule 1: Phosphoprotein

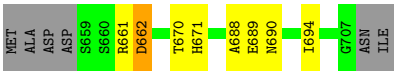
Chain G:

• Molecule 1: Phosphoprotein

Chain I:

• Molecule 1: Phosphoprotein

Chain K:

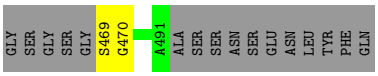


• Molecule 1: Phosphoprotein

Chain M:

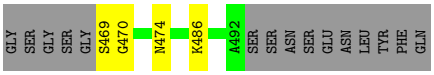
• Molecule 2: alpha MoRE of Nipah virus Nucleoprotein tail

Chain B:



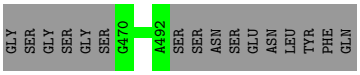
• Molecule 2: alpha MoRE of Nipah virus Nucleoprotein tail

Chain D:



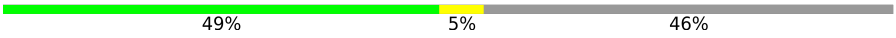
• Molecule 2: alpha MoRE of Nipah virus Nucleoprotein tail

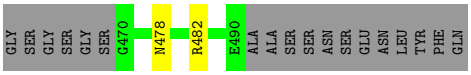
Chain F:



• Molecule 2: alpha MoRE of Nipah virus Nucleoprotein tail

Chain H:





• Molecule 2: alpha MoRE of Nipah virus Nucleoprotein tail

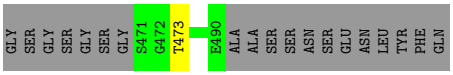
Chain J:

• Molecule 2: alpha MoRE of Nipah virus Nucleoprotein tail

Chain L:

• Molecule 2: alpha MoRE of Nipah virus Nucleoprotein tail

Chain N:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	60.46Å 131.97Å 156.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	23.94 – 2.79 23.94 – 2.79	Depositor EDS
% Data completeness (in resolution range)	99.5 (23.94-2.79) 99.7 (23.94-2.79)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.73 (at 2.80Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.203 , 0.257 0.206 , 0.262	Depositor DCC
R_{free} test set	808 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	55.1	Xtriage
Anisotropy	0.722	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	3798	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.92% 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	0.62	0/396	0.77	0/532
1	C	0.76	0/396	0.85	0/532
1	E	0.79	0/396	0.88	0/532
1	G	0.74	0/396	0.86	1/532 (0.2%)
1	I	0.72	0/396	0.86	0/532
1	K	0.71	0/396	0.83	1/532 (0.2%)
1	M	0.82	0/396	0.94	0/532
2	B	0.59	0/160	0.68	0/212
2	D	0.69	0/165	0.77	0/219
2	F	0.56	0/159	0.67	0/211
2	H	0.65	0/149	0.82	0/197
2	J	0.49	0/107	0.65	0/142
2	L	0.71	0/127	0.73	0/169
2	N	0.73	0/145	0.88	0/192
All	All	0.71	0/3784	0.83	2/5066 (0.0%)

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	K	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	662	ASP	CB-CG-OD1	5.29	123.06	118.30
1	K	662	ASP	CB-CG-OD1	5.13	122.92	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	K	689	GLU	Peptide

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	395	0	396	3	0
1	C	395	0	396	9	0
1	E	395	0	396	7	0
1	G	395	0	396	5	0
1	I	395	0	396	6	0
1	K	395	0	396	6	0
1	M	395	0	396	11	0
2	B	161	0	171	1	0
2	D	166	0	176	3	0
2	F	160	0	171	0	0
2	H	150	0	161	1	0
2	J	108	0	121	2	0
2	L	128	0	139	0	0
2	N	146	0	158	0	0
3	A	2	0	0	0	0
3	C	3	0	0	0	0
3	E	2	0	0	0	0
3	G	1	0	0	0	0
3	I	2	0	0	0	0
3	K	2	0	0	0	0
3	M	2	0	0	0	0
All	All	3798	0	3869	43	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 6.

All (43) 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:I:661:ARG:NH2	1:I:688:ALA:O	2.18	0.75

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:I:690:ASN:OD1	1:I:692:GLU:HG2	1.91	0.70
1:A:670:THR:OG1	1:A:671:HIS:HD2	1.78	0.67
2:D:469:SER:OG	2:D:470:GLY:HA3	1.96	0.65
1:E:679:SER:OG	1:M:707:GLY:HA3	1.97	0.64
2:B:469:SER:OG	2:B:470:GLY:HA3	2.02	0.60
1:C:693:GLU:OE2	2:D:486:LYS:NZ	2.26	0.58
1:A:670:THR:OG1	1:A:671:HIS:CD2	2.57	0.58
1:E:661:ARG:NH2	1:E:688:ALA:O	2.36	0.58
1:C:707:GLY:HA2	1:G:675:ARG:HE	1.71	0.56
1:K:661:ARG:NH2	1:K:688:ALA:O	2.42	0.52
1:C:707:GLY:HA3	1:G:679:SER:OG	2.08	0.52
1:C:678:ARG:O	1:C:682:ILE:HG12	2.08	0.52
1:K:690:ASN:O	1:K:694:ILE:HG12	2.10	0.52
1:M:661:ARG:NH2	1:M:688:ALA:O	2.43	0.51
1:M:670:THR:OG1	1:M:671:HIS:HD2	1.94	0.51
1:I:707:GLY:HA2	1:M:675:ARG:HE	1.76	0.50
1:K:670:THR:OG1	1:K:671:HIS:HD2	1.95	0.50
1:C:670:THR:OG1	1:C:671:HIS:HD2	1.95	0.49
1:I:678:ARG:O	1:I:682:ILE:HG12	2.13	0.49
1:M:665:LYS:HE2	1:M:682:ILE:HG23	1.95	0.48
1:E:670:THR:OG1	1:E:671:HIS:HD2	1.98	0.47
1:A:680:GLU:HG2	1:A:684:TYR:CE2	2.51	0.46
1:E:688:ALA:HA	1:E:693:GLU:OE1	2.16	0.46
1:K:671:HIS:HE1	1:M:662:ASP:OD2	1.99	0.46
1:M:670:THR:OG1	1:M:671:HIS:CD2	2.69	0.46
1:G:680:GLU:HG2	1:G:684:TYR:CE2	2.51	0.46
1:I:671:HIS:HE1	1:K:662:ASP:OD2	2.00	0.45
1:C:662:ASP:OD2	1:M:671:HIS:HE1	2.00	0.45
1:E:690:ASN:N	1:E:690:ASN:OD1	2.50	0.45
2:J:483:LEU:HD23	2:J:486:LYS:CE	2.48	0.44
2:J:483:LEU:HD23	2:J:486:LYS:HE2	2.00	0.43
1:C:670:THR:OG1	1:C:671:HIS:CD2	2.72	0.43
2:H:478:ASN:O	2:H:482:ARG:HD3	2.19	0.43
1:E:670:THR:OG1	1:E:671:HIS:CD2	2.73	0.41
1:C:680:GLU:HG2	1:C:684:TYR:CE2	2.56	0.41
1:I:680:GLU:HG2	1:I:684:TYR:CE2	2.56	0.41
1:M:665:LYS:HG2	1:M:682:ILE:HD13	2.02	0.41
2:D:474:ASN:HD21	1:G:687:LYS:HE2	1.87	0.40
1:K:670:THR:OG1	1:K:671:HIS:CD2	2.73	0.40
1:M:706:ASP:O	1:M:707:GLY:C	2.60	0.40
1:C:707:GLY:C	1:G:675:ARG:HH21	2.25	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:675:ARG:HE	1:M:707:GLY:HA2	1.85	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	47/55 (86%)	47 (100%)	0	0	100	100
1	C	47/55 (86%)	47 (100%)	0	0	100	100
1	E	47/55 (86%)	47 (100%)	0	0	100	100
1	G	47/55 (86%)	47 (100%)	0	0	100	100
1	I	47/55 (86%)	47 (100%)	0	0	100	100
1	K	47/55 (86%)	47 (100%)	0	0	100	100
1	M	47/55 (86%)	47 (100%)	0	0	100	100
2	B	21/39 (54%)	20 (95%)	1 (5%)	0	100	100
2	D	22/39 (56%)	22 (100%)	0	0	100	100
2	F	21/39 (54%)	20 (95%)	1 (5%)	0	100	100
2	H	19/39 (49%)	18 (95%)	1 (5%)	0	100	100
2	J	12/39 (31%)	11 (92%)	1 (8%)	0	100	100
2	L	16/39 (41%)	16 (100%)	0	0	100	100
2	N	18/39 (46%)	17 (94%)	1 (6%)	0	100	100
All	All	458/658 (70%)	453 (99%)	5 (1%)	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	45/50 (90%)	45 (100%)	0	100	100
1	C	45/50 (90%)	45 (100%)	0	100	100
1	E	45/50 (90%)	44 (98%)	1 (2%)	52	83
1	G	45/50 (90%)	42 (93%)	3 (7%)	16	43
1	I	45/50 (90%)	44 (98%)	1 (2%)	52	83
1	K	45/50 (90%)	45 (100%)	0	100	100
1	M	45/50 (90%)	44 (98%)	1 (2%)	52	83
2	B	16/28 (57%)	16 (100%)	0	100	100
2	D	16/28 (57%)	16 (100%)	0	100	100
2	F	15/28 (54%)	15 (100%)	0	100	100
2	H	15/28 (54%)	15 (100%)	0	100	100
2	J	12/28 (43%)	12 (100%)	0	100	100
2	L	13/28 (46%)	13 (100%)	0	100	100
2	N	15/28 (54%)	14 (93%)	1 (7%)	16	43
All	All	417/546 (76%)	410 (98%)	7 (2%)	60	87

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	706	ASP
1	G	661	ARG
1	G	692	GLU
1	G	706	ASP
1	I	661	ARG
1	M	706	ASP
2	N	473	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	671	HIS
2	B	478	ASN
1	C	671	HIS
2	D	474	ASN
1	E	671	HIS
1	G	671	HIS
1	I	671	HIS
1	I	702	ASN
2	J	474	ASN
1	K	671	HIS
1	M	671	HIS
1	M	702	ASN

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

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.