

**The structure of the nucleoprotein of Influenza D shows that all *Orthomyxoviridae* nucleoproteins have a similar NP<sub>CORE</sub>, with or without a NP<sub>TAIL</sub> for nuclear transport**

Amélie DONCHET<sup>1</sup>, Justine OLIVA<sup>2</sup>, Alice LABARONNE<sup>1</sup>, Laura TENGO<sup>1</sup>, Myriam MILOUDI<sup>1</sup>, Francine C. A. GERARD<sup>1</sup>, Caroline MAS<sup>3</sup>, Guy SCHOEHN<sup>1</sup>, Rob W. H. RUIGROK<sup>1</sup>, Mariette DUCATEZ<sup>2</sup> and Thibaut CREPIN<sup>1\*</sup>

<sup>1</sup> Institut de Biologie Structurale (IBS), Univ. Grenoble Alpes, CEA, CNRS, 38044 Grenoble, France

<sup>2</sup> IHAP, Université de Toulouse, INRA, ENVT, Toulouse, France

<sup>3</sup> Integrated Structural Biology Grenoble (ISBG) - UMS 3518 (CNRS-CEA-UJF-EMBL), 38044 Grenoble, France

**Supplementary information**

## Supplementary Figure 1

**Sequence alignment of NP from representative Influenza strains.** The sequences used are the same as in Table 1. The figure has been drawn using ESPript <sup>1</sup>. The upper and lower secondary structures correspond to the X-ray structures of D/NP (PDB entry 5N2U) and A/NP (PDB entry 2IQH), respectively. The stars indicate the phosphorylation sites identified for A/NP (closed symbol) and B/NP (open symbol) by Hutchinson *et al.* (red)<sup>2</sup>, Li *et al.* (green)<sup>3</sup> and Monda *et al.* (blue)<sup>4</sup>.

## Supplementary Figure 2

**Effect of trimethylamine N-oxide (TMAO) on the folding of D/NP<sub>TAIL</sub>.** Far UV-circular dichroism spectra of either D/NP<sub>TAIL</sub> alone (purple) or in presence of TMAO (0.5 M in yellow, 1M in red, 2M in blue and 3 M in green) were collected. Because of the high background induced by the presence of TMAO around 200 nm, spectra were cut when the voltage (right bottom insert) became too high (> 700 V).

## References

- 1 Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-324, doi:10.1093/nar/gku316 (2014).
- 2 Hutchinson, E. C. *et al.* Mapping the phosphoproteome of influenza A and B viruses by mass spectrometry. *PLoS Pathog* **8**, e1002993, doi:10.1371/journal.ppat.1002993 (2012).
- 3 Li, Y. *et al.* Phosphorylation and dephosphorylation of threonine 188 in nucleoprotein is crucial for the replication of influenza A virus. *Virology* **520**, 30-38, doi:10.1016/j.virol.2018.05.002 (2018).
- 4 Mondal, A. *et al.* Influenza virus recruits host protein kinase C to control assembly and activity of its replication machinery. *Elife* **6**, doi:10.7554/eLife.26910 (2017).
- 5 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21, doi:10.1107/S09074444090042073 (2010).

**Supplementary Table 1: Crystallographic data collection and structure refinement statistics.** Values in parentheses are for highest resolution shell.

Crystal Parameters	D/NP
<b>Data collection</b>	
Beamline	ID30A3 - ESRF
Wavelength (Å)	0.9677
Space group	P1
Cell Dimensions a, b, c (Å)	75.17, 85.17, 103.39
$\alpha, \beta, \gamma$ (°)	91.20, 101.94, 101.02
Molecules per asymmetric unit	4
Wilson B-factor (Å <sup>2</sup> )	51.4
Resolution (Å)	40 - 2.4 (2.48 - 2.4)
Observed Reflections	191 888 (18 695)
Unique Reflections	87 824 (8 505)
Completeness (%)	91.3 (94.1)
R <sub>sym</sub> (I) <sup>(a)</sup> (%)	9.3 (84.7)
<I/σ(I)>	7.78 (1.05)
<b>Model quality indicators</b>	
R <sub>cryst</sub> (%)	20.2 (21.9)
R <sub>free</sub> (%)	25.2 (23.8)
Non-hydrogen atoms	14 646
Water atoms	235
Average B-factors (Å <sup>2</sup> )	55.7
B-factors per chains (Å <sup>2</sup> )	59.8, 54.3, 52.8, 56.3
B-factors water (Å <sup>2</sup> )	49.5
Ramachandran favoured / allowed (%)	94.0 / 5.1
rms deviations, bonds (Å) / angles (°)	0.009 / 1.15
Clashscore <sup>(b)</sup>	4.31
MolProbity score <sup>(b)</sup>	2.28

<sup>(a)</sup>  $R_{\text{sym}} = \frac{\sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl,j}}$

<sup>(b)</sup> calculated using MolProbity<sup>5</sup>

D/NP .....  
 C/NP .....  
 toad/NP MATQEQVAKAAKLF AEASATVEKAITPPIEEKKTKRAATQIEKYRKYIYEQKLMDDQYPAALAKLNRLEGKSPNTPEDMP  
 B/NP ..... MSNMDIDGINTGTIDKAPE  
 A/NP .....

..... MDSTKAO TPEEQR AKNAK T I L E N I Q I Y E R M C D L F G  
 ..... MSDRRQRNK T P D E Q R K A N A L I I N E N I E A Y I A I C K E V G  
 EERDENMSEEEDEEEIEAAEPTDRPEPATQEQPSTSRPQRKSGPETKEAAKQRNSDDGFANMIKMLGDLWVVLARFAS  
 EITSGTSGITTRPIIRPATLAPP SNKRTRNPSPERATTISEADVGRKTKQK T P T E I K K S V Y N M V V K L G E F Y N O M M V K A G  
 ..... MATKGTKRSEYEQMETDGERQ N A T E I R A S V G K M I D G I G R F Y I Q M C T E L K  
 .....  
 ☆ ☆ ☆

..... V S E . D D R L I E N S I S I E R M I R V V T D K K Y O D K . K L K N A . . . G S D L E K I A N A G K V F C R L I V E S T A G . . . . K C S A R L G M A L K P N  
 ..... L N G . D E M L I L E N G I A I E K A I R I C O D G K Y O E K . R E K K A R E A Q R A D S N F N A D S I G I R L V K R A G S G T N I T Y H A V V E L T S R R S R  
 ..... L T A P A E K D C L D N A I S I T K A I L H A S D T R K T E Y L K K E K G . . . . . G G G N D Y D V V G F N L F S V G N P . . . . . G M T V R R E T R K E A  
 ..... L N D D M E R N L I Q N A H A V E R I L L A A T D D K K T E F O K K N A R D V K E E K E E I D H N K T G G T F Y K M V R D D K I T Y F S P I R I T F L K E E  
 ..... L S D Y E G R . L I O N S L T I E R M V L S A F D E R R N K Y L E H P S . . . . . A G K . . . . . D P K K T G G F I Y R R V D G . . . . . K W R R E L I L Y D K E E  
 .....  
 → → → → →

..... V E A V L T D V L G N E L D R A A V L G K R M G F T A M F K S N L E E V L Y Q R G K N Q L K K R N S A E T F T I L S Q G A S L E A R F R P I M E K H L G V G T V  
 ..... I V Q I L K S H W G N E L N R A K I A G K R L G F S A L F A S N L E A I I Y Q R G R N A A R R N G S A E L F T I L T O G A G I E T R Y K W I M E K H I G I G V L  
 ..... L M D L V K V F Q N G K D L N N A L . . . . . I H L E L T K S Q Y G H A F Y T N T R P L K K L G L S N K Y R A L S K G Y D I P R K A G I T G M H A M G I A P L  
 ..... V K T M Y K T T . M G S D G F S G L N . . . . . H I M I G H S Q M N D V C Y O R S K A L K R V G L D P S L I S T F A G T I P R R S G A T G V A I K G G G T L  
 ..... L R R I W R Q A N N G D D A T A G L T . . . . . H M M I W H S N L N D A T F O R T R A L V R T G M D P R M C S L M Q G S T L P R R S G A A G A A V K G G T M  
 .....  
 ☆ ☆ ☆ A/NP Ser-165 ☆

..... V A S T K N T I A S K K N G N Y K N K M V R P P G N R E S W S P L E R E T I S F L N K R L F P G P M R O L C K K F E . . . . . Y L N D . . . . . Q E K Q L A L N L  
 ..... I A D A K G L L I N G K R E G K R G V D A N V K L R A G . T T G S P L E R A M Q G I E K A F P P L R A L A R R V V . . . . . K A N Y . . . . . N D A R E A L N V  
 ..... A S R A I G L I E S L P D R E . . . . . D P A S K G K K S A L Q K M L G V L V E K V S N P K Q K E V A Q M V A G M A T K R L G R N E K T E M T R A I E M  
 ..... V A E A I R F I G R A M A D R . . . . . G L L R . . . . . D I K A K T A Y E K I L N L K N K C S A P Q O K A L V D Q V I . . . . . G S R N P G I A D I E D L T L  
 ..... V M F L I R M I K R G I N D R . . . . . N F W R G E N G R R T R I A Y E R M C N I L K G K F Q T A A Q R T M V D Q V R . . . . . E S R N P G N A E F E D L I F  
 .....  
 → → → → →

..... M I D A S L I L K P Q V T H K M I M P W S M L A V K K Y A E M N K G S P S L E D L A A Y S G V R A F M A F N T A C Y M S K F T I G K G I V G D A E I M E N G  
 ..... I A E A S L L L K P Q I T N K M T M P W C M W L A A R L T L K D . . . . . E F A N F C A Y A G R R A F E V F N I A M E K I G I C S F Q G T I M N D D E I E S I  
 ..... I A Q S C C V F P P N S R N K V Y M P M C I Y S H L V P N F N T . . . . . Y N M M G Y E A Q A L F A M A P V T F V V K D G E . . . . . N P  
 ..... L A R S M V V R P S V A S K V V I P I S I Y A K I P Q L G . . . . . F N V E E Y S M V G Y E A M A L Y N M A T P V S I L R M G D . . . . . D A  
 ..... L A R S A L I L R G S V A H K S C I A C V Y G S A V A S G Y D . . . . . F E R E G S L V G I D E F R L L Q N S Q V Y S L I R P N E . . . . . N P  
 .....  
 ☆ ☆ ☆ → → → →

..... N D K M Q I T A M A C F G L A Y E D T G I V A A M I S Q P M K K R Y Q L K V G N F N P P E E G T I K G T S A G Y F H K . . . . . W A E F G N R L P F N S F G  
 ..... E D K A Q V I M M A C F G L A Y E D F S L V S A M V S H P L K L R N R M K I G N F R V G E K V S T V L S P L L R F T R . . . . . W A E F A Q R F A L Q A N T  
 ..... N P K S D L I C M S I L G C A Y E D M N V V S A I L G R K I K R S A L E M E V K G N S Y T E K D R T E G V V G K I V N L V N K P E N W A S L T S K G G G G T S N  
 ..... K P K S Q L F M S C F G A A Y E D L R V L S A L T G T E F K P R S A L K C K G F H V P A K E Q V E G M G A A L M S I K L . . . . . W A P M T R S G G N E V G G  
 ..... A H K S Q L V W M A C H S A A F E D L R V S S F I R G T K V V P R G K L S T R G V Q I A S N E N M E T M E S S T L E L R S R . . . . . W A I R T R S G G N T N Q Q  
 .....  
 → → → → →

exchange domain

..... T E S K Q I S N S G V F A V Q R P S T T N I Q R L A E L M A R N T G E T S D N F T Q L V Q K I R Q V G T F A D Q K A N L R E T G G Y I Y D I T D V T K S  
 ..... S R E G A Q I S N S A V F A V E R K I T T D V Q R V E L L N K V Q A H E D E F L Q T L Y K K V R E Q I S I I G R N K S E I K E F L G S S M Y D L N D Q E K F  
 ..... D R G S Q A E N V A T F G V E R V Q R I N K R D L L R G M G D G T S . R D D F I E D L R L A M M E L N . . . . . S E E K R T I P S G L K V H P M G D P T G E  
 ..... D G S S Q I S C S P V A V E R P I A L S K Q A V R R M L S M N T E G R D A D V K G N L K M M N D S . . . . . M A K K T S G N A F T G K K M Q I S D K N K T  
 ..... R A S S G Q I S I Q P T S V Q R N L P F D R F T I M A A F T G N T E G R T S D M R T E I I R L M E S A . . . . . R P E D V S . . . . . F Q G R G V E L S D E K A T  
 .....  
 ☆ ☆ ☆ A/NP Ser-407 ☆

..... N P K I P Q L G . . . . . G N S F F F E F T G S D V P R T G A K R R V G G A D D V T L G T S Q P K K R G R Q G A G V E S S M D I E T V G E D . . . . .  
 ..... N P L N F R S G . . . . . A H P F F F E F D P D Y N P I R V K R P K P I A K R N S N I S R L E E E G M D E N S E I G Q A K K M P L D Q L T S T S S N I P G K N  
 ..... R P V P I T A G . . . . . K M L F F F F G V T D I E . . . . .  
 ..... N P W E I Q I K Q T I P . N F F F G R D T A E D Y D D L D Y . . . . .  
 ..... S P E I V P S F D M S N E G S Y F F G D N A E E Y D N . . . . .  
 .....  
 ☆ → ☆

Supplementary Figure 2

