



HAL
open science

Dissecting Henipavirus W proteins conformational and fibrillation properties: contribution of their N- and C-terminal constituent domains

Giulia Pesce, Frank Gondelaud, Denis Ptchelkine, Christophe Bignon, Patrick Fourquet, Sonia Longhi

► To cite this version:

Giulia Pesce, Frank Gondelaud, Denis Ptchelkine, Christophe Bignon, Patrick Fourquet, et al.. Dissecting Henipavirus W proteins conformational and fibrillation properties: contribution of their N- and C-terminal constituent domains. FEBS Journal, 2024, Online ahead of print. 10.1111/febs.17239 . hal-04683726

HAL Id: hal-04683726

<https://hal.science/hal-04683726v1>

Submitted on 2 Sep 2024


HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Dissecting *Henipavirus* W proteins conformational and fibrillation properties: contribution of their N- and C-terminal constituent domains

Giulia Pesce¹, Frank Gondelaud¹, Denis Ptchelkine¹, Christophe Bignon¹, Patrick Fourquet² and Sonia Longhi¹ 

¹ Laboratoire Architecture et Fonction des Macromolécules Biologiques (AFMB), UMR 7257, Centre National de la Recherche Scientifique (CNRS) and Aix Marseille University, France

² INSERM, Centre de Recherche en Cancérologie de Marseille (CRCM), Centre National de la Recherche Scientifique (CNRS), Marseille Protéomique, Institut Paoli-Calmettes, Aix Marseille University, France

Keywords

amyloid-like fibrils and fibrillation; Congo Red and Thioflavin T binding assays; intrinsically disordered proteins/regions; negative-staining transmission electron microscopy; small-angle X-ray scattering

Correspondence

S. Longhi, Lab. AFMB, UMR 7257, CNRS and Aix Marseille University, 163, Avenue de Luminy, Case 932, 13288 Marseille, France

Tel: +33 (0) 4 13 94 95 17

E-mail: sonia.longhi@univ-amu.fr

(Received 12 March 2024, revised 7 May 2024, accepted 23 July 2024)

doi:10.1111/febs.17239

The Nipah and Hendra viruses are severe human pathogens. In addition to the P protein, their P gene also encodes the V and W proteins that share with P their N-terminal intrinsically disordered domain (NTD) and possess distinct C-terminal domains (CTDs). The W protein is a key player in the evasion of the host innate immune response. We previously showed that the W proteins are intrinsically disordered and can form amyloid-like fibrils. However, structural information on W CTD (CTD_W) and its potential contribution to the fibrillation process is lacking. In this study, we demonstrate that CTD_{WS} are disordered and able to form dimers mediated by disulfide bridges. We also show that the NTD and the CTD_W interact with each other and that this interaction triggers both a gain of secondary structure and a chain compaction within the NTD. Finally, despite the lack of intrinsic fibrillogenic properties, we show that the CTD_W favors the formation of fibrils by the NTD both *in cis* and *in trans*. Altogether, the results herein presented shed light on the molecular mechanisms underlying *Henipavirus* pathogenesis and may thus contribute to the development of targeted therapies.

Introduction

Nipah (NiV) and Hendra (HeV) viruses are zoonotic biosafety level 4 (BL4) pathogens within the *Henipavirus* genus (Paramyxoviridae family, Mononegavirales order) that cause severe acute respiratory illness and fatal encephalitis in humans [1]. Despite the threat they pose to human health, neither antiviral therapies nor vaccines are available. The asymptomatic natural

reservoir of these viruses is bats of the *Pteropus* genus (also called flying foxes).

HeV was isolated in 1994 in Brisbane, Australia, where it caused a fatal respiratory and neurological disease in horses. Human exposure to body fluids, tissues, or feces of infected animals resulted in the transmission of the disease, with a human case fatality rate

Abbreviations

BL4, biosafety level 4; CD, circular dichroism; CTD_W, C-terminal domain of W protein; DTT, dithiothreitol; HeV, Hendra Virus; IDR/P, intrinsically disordered region/protein; IFN, interferon; IMAC, immobilized metal affinity chromatography; MS, mass spectrometry; NF-κB, nuclear factor kappa-light-chain-enhancer of activated B cells; NiV, Nipah Virus; NTD, N-terminal domain; pI, isoelectric point; PMG, premolten globule; RC, random coil; RdRp, RNA-dependent RNA polymerase; RNA, ribonucleic acid; SAXS, small-angle X-ray scattering; SDS/PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; SEC, size exclusion chromatography; STAT, signal transducer and activator of transcription; TEM, transmission electron microscopy; TFE, trifluoroethanol; ThT, thioflavin T; TLR3, toll-like receptor 3; TRX, thioredoxin; ZnFD, zinc-finger domain.

of 57%. NiV first outbreak occurred in Malaysia in 1998, causing a respiratory and neurological disease in pigs and encephalitis in humans. Since this first event, NiV outbreaks were observed in subsequent years in Bangladesh, India, and Philippines, with a case fatality rate comprised between 40% and 75%. An additional threat posed by NiV is its ability to spread from human-to-human, contributing to the spillover in densely populated regions. NiV is therefore on the WHO priority list of pathogens with pandemic potential requiring urgent research.

Henipaviruses are enveloped viruses whose genome consists in a single-stranded, non-segmented RNA molecule with negative polarity encoding six structural proteins (Fig. 1A,B). The genome of Henipaviruses is encapsidated by the nucleoprotein (N) into a helical nucleocapsid that is the substrate of the RNA-dependent-RNA polymerase (RdRp). The RdRp, responsible for transcription and replication, consists of the L protein in complex with the phosphoprotein (P) that enables recruitment of L on the nucleocapsid template. *Henipaviruses* N and P proteins display long intrinsically disordered regions (IDRs) [2–4], that is, regions devoid of stable secondary and tertiary structures [5].

The *Henipavirus* P protein consists of a long N-terminal intrinsically disordered domain (NTD) and a C-terminal region that contains both structured and disordered regions (Fig. 1B) [6]. In NiV and HeV, the P gene also encodes for the V and W proteins. The latter result

from the addition of either one (V protein) or two (W protein) non-templated guanosines at the editing site of the P messenger. As shown in Fig. 1B, the editing site is located at the end of the P gene region encoding the long, intrinsically disordered NTD. Therefore, P, V, and W share the same NTD and possess unique C-terminal domains (herein referred to as CTD_P, CTD_V, and CTD_W).

The V and the W proteins play a key role in the evasion of the host antiviral innate immune response [7–9]. Beyond inhibiting type I interferon (IFN) induction [10], V and W antagonize IFN-mediated antiviral response through binding to STAT proteins. P, V, and W proteins all block the cellular response to IFN by binding to STAT1 through their disordered NTD and preventing tyrosine phosphorylation of STAT1, a post-translation modification that is necessary for STAT1 activation [11–13]. In addition, the W protein localizes into the nucleus and retains therein unphosphorylated STAT1 [12]. Beyond STAT1 and STAT2, NiV P, V, and W proteins also bind to STAT4 through their NTD [14]. Furthermore, *Henipavirus* V and W proteins inhibit chemokine production *in vitro* and modulate the inflammatory response *in vivo* [15]. NiV W inhibits TLR3 signaling and activation of the NF- κ B pathway [10,16,17]. Finally, *Henipavirus* W proteins bind to 14-3-3 proteins via their CTD, and this interaction modulates several cellular processes, including apoptosis [18] and inhibition of the NF- κ B-induced proinflammatory response [19].

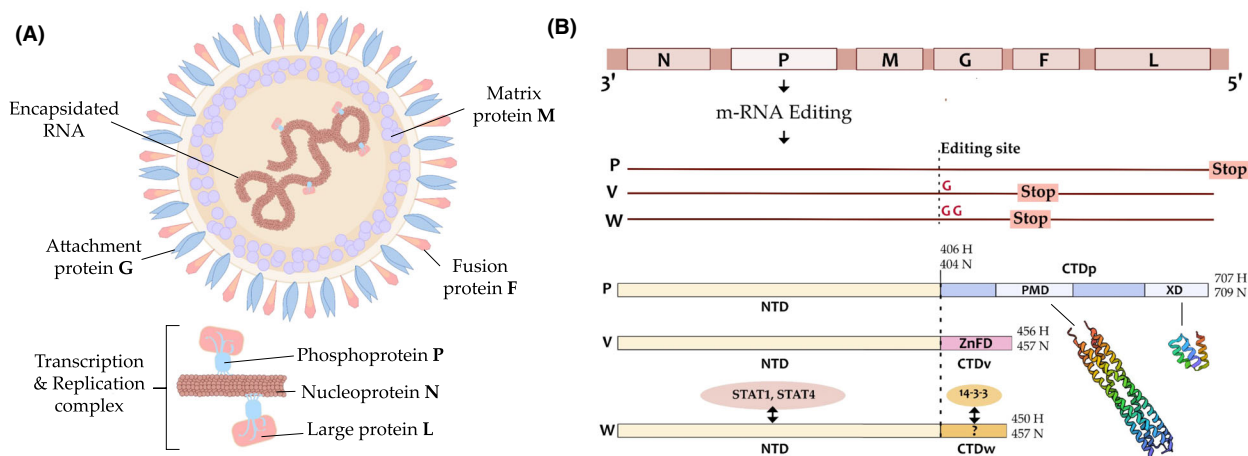


Fig. 1. *Henipavirus* virion structure and genome organization. (A) Schematic representation of a *Henipavirus* particle (top) and of the transcription and replication complex formed by the viral RNA genome encapsidated by the nucleoprotein (N) and by the RNA-dependent-RNA-polymerase (RdRp), composed of the large (L) protein and the phosphoprotein (P) (bottom). (B) Schematic representation of the *Henipavirus* genome, consisting of a single-stranded negative-sense RNA molecule containing six genes encoding for the nucleoprotein (N), the phosphoprotein (P), the matrix protein (M), the attachment glycoprotein (G), the fusion protein (F), and the large protein (L). The Nipah virus (NiV) and Hendra virus (HeV) P gene encodes also three additional non-structural proteins, V, W, and C. V and W result from P mRNA editing. The structures of the NiV P multimerization domain (PMD) (PDB code 4N5B) and of the NiV P X domain (XD) (as extracted from PDB code 7PNO) are also shown.

We previously characterized V and W proteins of NiV and HeV and showed that they are prevalently disordered [20,21]. Contrary to the CTD_V, which folds into a zinc-finger conformation (ZnFD) [20], no direct structural information is available on the isolated CTD_W. Our previous preliminary bioinformatic analysis suggested a disordered nature of the CTD_W [21], and the extended conformation it adopts upon binding to importin $\alpha 1$ or $\alpha 3$ [22] or to 14-3-3 isoform σ [18] supports the hypothesis that it may be disordered in isolation. However, the disordered nature of the isolated CTD_W has never been demonstrated experimentally.

We previously reported that the full-length W proteins form amyloid-like fibrils that we hypothesized might have a role in impairing the host innate immune and inflammatory response [21]. The ability of the W proteins to form fibrils is very likely conferred by the PNT3 region (aa 200–310) that we previously identified to be amyloidogenic [23,24]. In addition, we showed that the formation of fibrils by the W proteins critically depends on the redox state of cysteine residues [25]. This latter finding, and the presence of a conserved cysteine in both CTD_{WS} (at position 421 and 419 in the NiV and HeV W sequences), suggests a potential involvement of the CTD_W in W fibrillation, prompting a direct experimental characterization and assessment of its possible dimerization ability in order to unravel the precise redox mechanisms of W fibrillation.

In light of the molecular partnership established by the CTD_W and of its involvement in counteracting the host antiviral response, investigating the conformational properties of this domain and its contribution to the fibrillation process of the W protein is a prerequisite toward rational antiviral approaches. We herein have combined various biochemical and biophysical approaches to shed light on the structural properties of the HeV and NiV CTD_W and on its contribution to the overall conformational properties and fibrillation abilities of the W proteins. Our results revealed that the *Henipavirus* CTD_{WS} are intrinsically disordered and non-fibrillogenic *per se* and that the NTD and the CTD_W interact with each other. For both viruses, this interaction not only affects the secondary structure content and compactness of the W protein but also influences its fibrillation abilities.

Results and Discussion

Disorder predictions of the HeV and NiV CTD_W

Our previous preliminary *in silico* analysis hinted at a disordered nature of the HeV and NiV CTD_W [21].

Here, we performed a more in-depth bioinformatic analysis using MeDor, a metaserver for disorder prediction that combines the prediction of multiple disorder predictors [26]. As shown in Fig. 2, the various disorder predictors yielded conflicting results, resulting in an overall ambiguous prediction. Indeed, the presence of hydrophobic clusters in the HCA plot suggests an at least partly folded conformation, while ESpritz-D and DisEMBL HOTLOOPS predict the two CTD_{WS} as disordered. In addition, the predicted disorder content is not the same for the two CTD_{WS}, with NiV CTD_W being predicted to be more disordered than its HeV counterpart (cf. Consensus 1 in the two panels of Fig. 2). The ambiguity in disorder predictions, together with the differences in predicted disorder in the two CTD_{WS}, calls for a direct experimental assessment of their conformational properties. To this end, we undertook their expression and purification in *Escherichia coli*.

Expression and purification of the CTD_W from both HeV and NiV

Since the expression and purification of the HeV and NiV W hexahistidine-tagged proteins driven by a pDEST17 construct was successful [21], we first attempted at producing the two CTD_{WS} using the same strategy; however, they were not expressed under these conditions and could not be recovered. To circumvent this technical difficulty, we used a cleavable thioredoxin (TRX) solubility tag and generated a construct expressing a TRX-CTD_W fusion protein carrying a Tobacco Etch virus (TEV) protease cleavage site between the two and with a C-terminal 6His-Tag, the latter being necessary for interaction studies (Fig. 3A). Using this system, both CTD_{WS} could be expressed and purified in high amount via a three-step purification process. The first step was an immobilized metal affinity chromatography (IMAC), followed by an overnight TEV protease cleavage to remove the TRX tag. Ion exchange chromatography (IEC) was used to separate the TEV protease and the TRX solubility tag from the cleaved CTD_W (Fig. 3B). The purified CTD_{WS} of both NiV and HeV show an apparent molecular mass of ~ 11 kDa (Fig. 3B), in contrast with the expected molecular mass of ~ 6 kDa, as calculated from the amino acid sequence. However, for both CTD_{WS}, mass spectrometry (MS) analysis of peptides resulting from tryptic digestion confirmed the identity of the purified proteins (Fig. 4A,B). Anomalous migration had previously been reported for *Henipavirus* NTD [2], V [20], and W proteins [21] and is a characteristic feature of IDPs or IDRs.

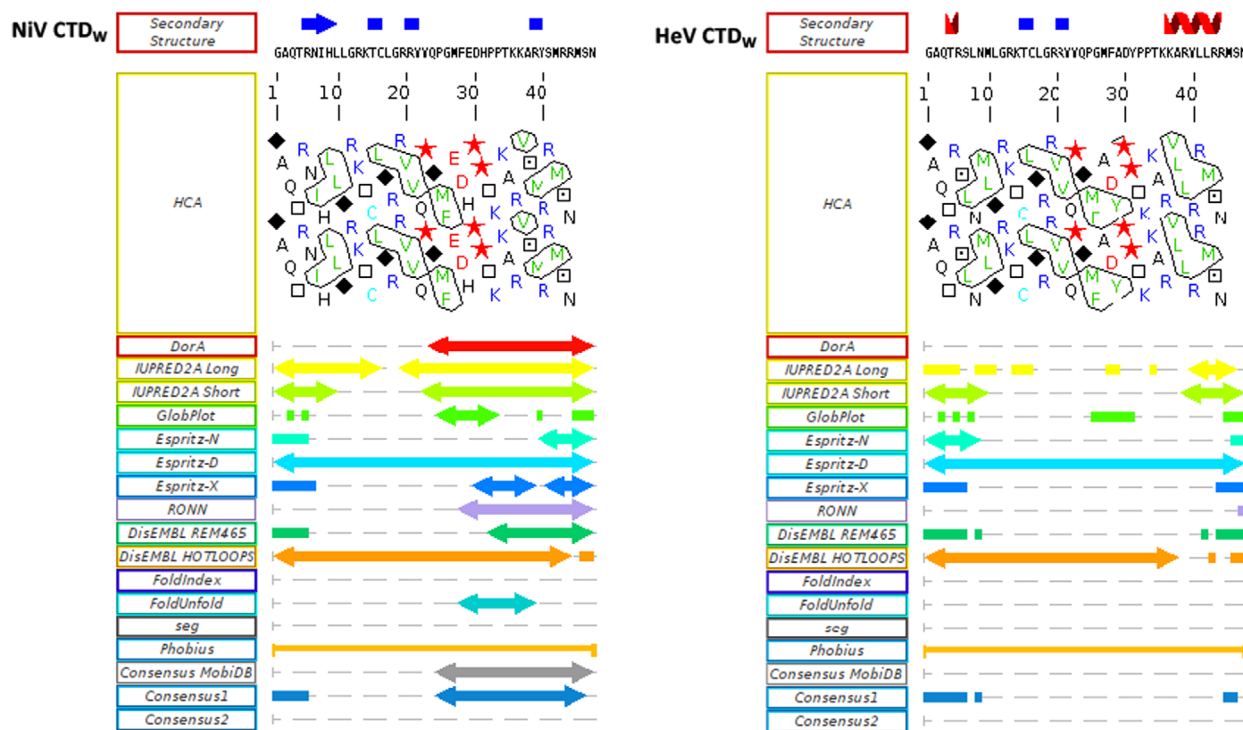


Fig. 2. Analysis of disorder content of Nipah virus (NiV) and Hendra virus (HeV) W protein C-terminal domain (CTD_W) as provided by MeDor. The amino acid sequence is represented as a single, continuous horizontal line below the predicted secondary structure elements. Below the sequence is shown the hydrophobic cluster analysis (HCA) plot and the predicted regions of disorder that are represented by bi-directional arrows. In the HCA plot, prolines are shown as red stars, glycines as black diamonds, threonines as squares, and serines as dotted squares. Hydrophobic residues are shown in green. Basic and acidic residues are shown in blue and red, respectively. The consensus disorder prediction generated by MobiCD-Lite is shown (ConsensusMobiDB), as are the two consensus generated by MeDor. Consensus 1 corresponds to regions predicted as disordered by more than half of the predictors, while Consensus 2 corresponds to regions predicted as disordered by all predictors.

MS analysis revealed that NiV and HeV CTD_W have average molecular masses of 6054.6 and 5998.8 Da, respectively (Fig. 3C,D), a value that exceeds by ~42 Da the expected mass value (6017.94 and 5961.96 Da, respectively). This shift in the molecular mass is consistent with protein carbamylation, a modification that can occur when proteins are purified in the presence of urea [27]. MS analysis of the peptides resulting from tryptic digestion of the CTD_W confirmed carbamylation of residues Lys34 and, preferentially, of Lys33.

For both CTD_{WS}, MS analysis revealed the presence of an additional, intense peak (Fig. 3C,D), with a mass close to that expected for a dimeric form.

Oligomerization state of the CTD_W from size exclusion chromatography (SEC)

To obtain information on the oligomerization state of the CTD_W, we used analytical SEC. The SEC profile

of both proteins features a major peak at ~16 mL and an additional peak (of very low intensity in the case of HeV) at ~14 mL (Fig. 5A).

We suspected these peaks to correspond to a dimeric (~14 mL) and monomeric (~16 mL) form of the protein. The co-existence of a monomeric and dimeric form was confirmed by SDS/PAGE in both reducing and non-reducing conditions (Fig. 5B). In the absence of DTT, two bands are visible for both NiV and HeV CTD_W, at ~11 kDa and ~16 kDa, with the higher one disappearing upon addition of DTT (Fig. 5B). These findings are consistent with MS data that revealed the presence of both monomeric and dimeric species. As expected, the addition of DTT causes the peak at ~14 mL to disappear in the SEC elution profile (Fig. 5A). Results thus indicate that both CTD_{WS} undergo disulfide-driven dimerization via their unique cysteine residue (C421 and C419 in the NiV and HeV W sequences) (UniProtKB P0C1C7, P0C1C6). The apparently higher dimerization propensity of NiV

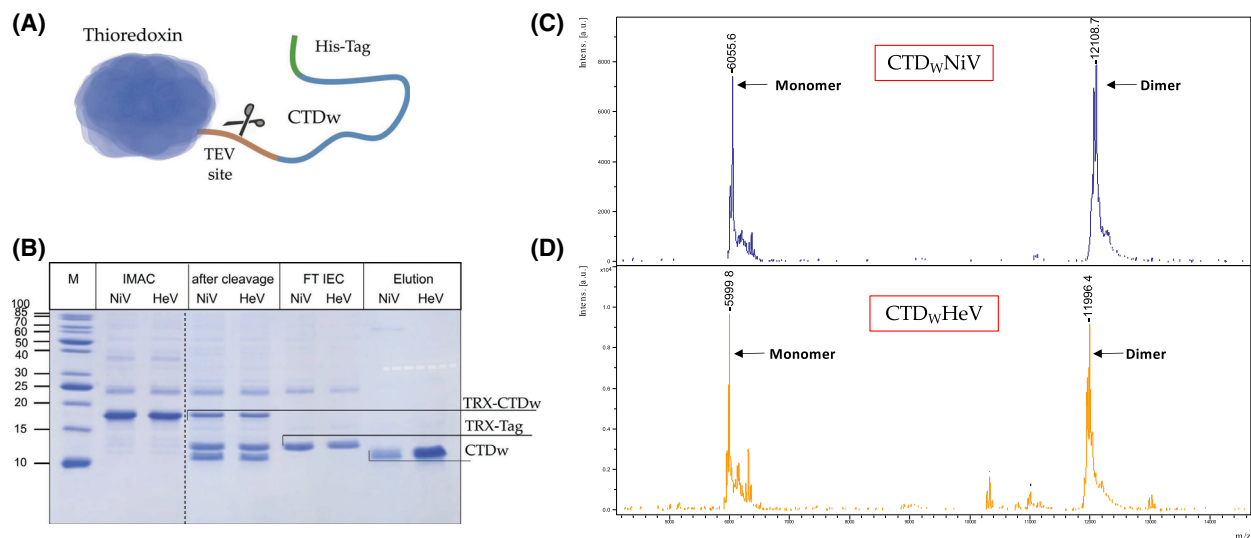


Fig. 3. Purification of the W protein C-terminal domains (CTD_W). (A) Schematic representation of the recombinant CTD_W consisting, from N to C terminus, of a solubility tag (Thioredoxin, TRX), a Tobacco Etch Virus (TEV) protease cleavage site, the CTD_W, and a 6His tag. (B) SDS/PAGE analysis of the purification of the CTD_W. M, molecular Markers; IMAC, eluent from Immobilized Metal Affinity Chromatography (IMAC); after cleavage: sample after overnight incubation with TEV protease, showing the presence of three main bands corresponding to uncleaved recombinant TRX-CTD_W, TRX tag, and CTD_W; FT IEC: Flow-through (i.e., unretained fraction) from Ion Exchange Chromatography (IEC); Elution: eluent from IEC. Data are representative of at least five independent analyses. The dashed line marks where the gel was cut to remove an irrelevant lane. (C, D) MALDI-TOF MS result for Nipah virus (NiV) (C) and Hendra virus (HeV) (D) CTD_W. An additional peak corresponding to the mass expected for a dimer is also detectable. Data are the result of a single experiment.

(A) NiV CTD_W 64.71% coverage

GAQTRNIHLLGRKTCLGRRVVQPGMFEDHPPTKKARVSMRRMSNYHHHHHH

(B) HeV CTD_W 64.71% coverage

GAQTRSLNMLGRKTCLGRRVVQPGMFADYPPTKKARVLLRRMSNYHHHHHH

(C) NiV NTD 80.54% coverage

GDKLELVNDGLNIIDFIQKNQKEIQKTYGRSSIQQPSIKDQTKAWEDFLQCTSGESEQVEGGMSKDDGDVERNLED
 SSTSPDTGTIGKRVSNTRDWAEGSDDIQLDPVVTDVVYHDHGECTGYGFTSSPERGWSDYTSGANNNGNVLVSDAK
 MLSYAPEIAVSKEDRETDLVHLENKLSTTGLNPTAVPFTLRNLSDPAKDSPVIAEHYYGLGVKEQNVGPQTSRNVNL
 DSIKLYTSDDEEADQLEFEDEFAGSSSEVIVGISPEDEEPSVVGKPNESIGRTIEGQSIIRDNLQAKDNKSTDVPGA
 GPKDSAVKEEPPQKRLPMLAEFEFCSGSEDPPIIRELLKENSLINCQQGKDAQPPYHWSIERSISPDKTEIVNGAVQT
 ADRQRPGTPMPKSRGPIPK

(D) HeV NTD 87.87% coverage

GDKLDLVNDGLDIIDFIQKNQKEIQKTYGRSSIQQPSIKDQTKAWEDFLQCTSGEHEQAEGGMPKNDGGTEGRNVED
 LSSVTSSDGTIGQRVSNTRAWAEDPDDIQLDPMVTDVVYHDHGECTGHGSPSSPERGWSYHMSGTHDGNVRAVPT
 KVLPNAKTTVPEEVREIDLIGLEDKFAAGLNPAAVFVFPKNQSTPTEPPVPIPEYYYSGRRGDLKSKSPPRGNVNL
 DSIKIYTSDDDENQLEYEDEFAGSSSEVIVGISPEDEEPSVVGKPNESIGRTIEGQSIIRDNLQAKDNKSTDVPGA
 SLMQDSCRRGGVPKRLPMLSEEFECSSGDDPIIQELEREGSHPGGSLRLREPPQSSGNSRNQPDRLKTDGAASPGG
 VQRPGTPMPKSRIMPIPK

Fig. 4. Results of peptide mass fingerprint (PMF) of the Nipah virus (NiV) (A, C) and Hendra virus (HeV) (B, D) W C-terminal domain (CTD_W) (A, B) and untagged N-terminal domain (NTD) (C, D). The peaks obtained have been compared to the theoretical digestion peak list of the expected protein. Expected peptides found in the spectrum are represented in green. The sequence coverage (%) is indicated. Data are the result of a single experiment.

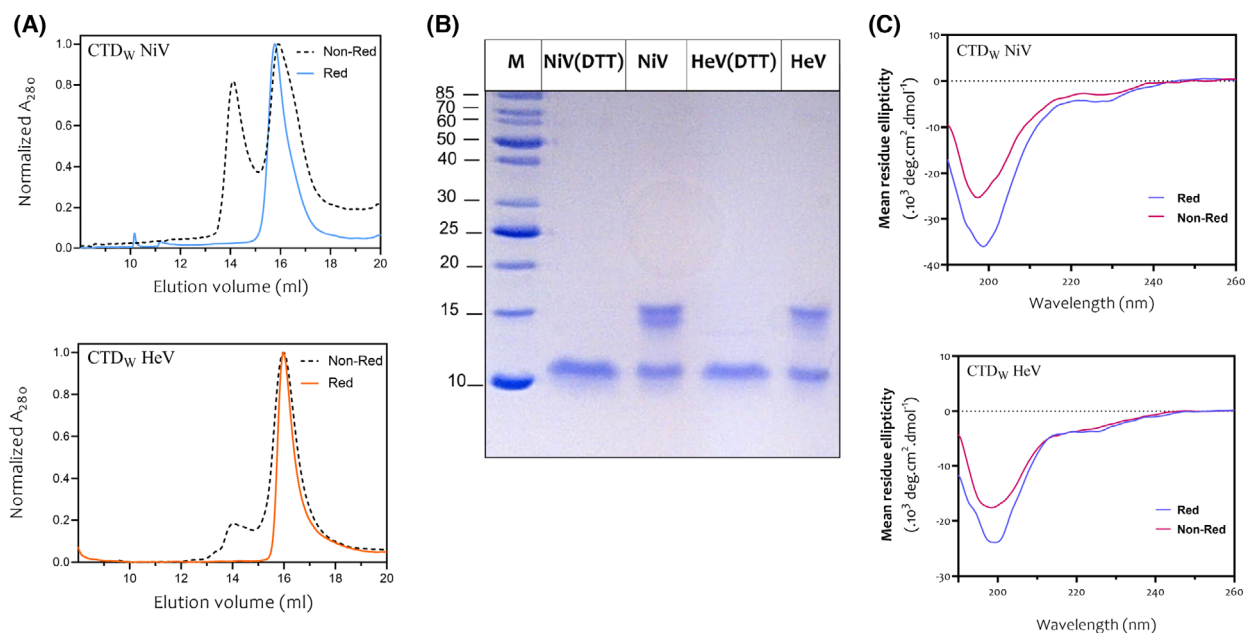


Fig. 5. Influence of the RedOx state of the W protein C-terminal domains (CTD_Ws) on their conformational properties. (A) Size exclusion chromatography (SEC) profile of a CTD_W sample devoid of (black dotted curves) or containing DTT (blue and orange curves). Data are representative of three independent experiments. (B) SDS/PAGE analysis of the CTD_W in the presence or absence of DTT. Data are representative of three independent analyses. (C) CD spectra of CTD_W samples previously incubated either in the presence (red) or absence (non-red) of DTT. Spectra are representative of three independent experiments.

CTD_W compared to HeV CTD_W is noteworthy in light of the high (82%) sequence identity between the two domains.

Far UV circular dichroism (CD) studies of the CTD_W

In order to obtain insights into the secondary structure content of the CTD_W, far-UV CD spectra were recorded on samples incubated in reducing or non-reducing conditions (Fig. 5C). All spectra show a prominent negative peak at ~200 nm, low ellipticity at 190 nm, and no strong negative peak above 205 nm (typical of α -helix and β -sheet) [28]. These features are consistent with a predominantly disordered conformation for both CTD_Ws, in agreement with the results provided by SDS/PAGE that revealed an anomalous migration typical of disordered proteins. In addition, the redox state was found to affect the disorder content to some extent, as judged from the reduced amplitude of the peak at 200 and increased ellipticity at 190 nm in the non-reduced samples at comparable absorption values (Fig. 5C). These results, that mirror those obtained on the full-length W proteins [21], thus reflect an at least partial disorder-to-order transition upon cysteine oxidation, consistent with a role of

disulfide bridge formation in secondary structure stabilization.

Far-UV CD spectroscopy also allows distinguishing between IDPs adopting a random coil (RC)-like state and those adopting a premolten globule (PMG)-like state (i.e., a disordered state endowed with some residual fluctuating secondary and tertiary structure) based on the ratio of the ellipticity values at 200/222 nm [29]. According to this ratio, both CTD_Ws fall within the random coil-like region (Coil) of the plot (Fig. 6A).

To investigate possible structural transitions that the CTD_W might undergo, NiV and HeV CTD_W spectra were also recorded on non-reduced samples containing increasing concentration of the secondary structure stabilizer trifluoroethanol (TFE) (Fig. 4B,C). The titration with TFE clearly triggers a progressive increase of the signal at 208 and 222 nm and a shift of the minimum from 198 to 205 nm for NiV (Fig. 6B) and from 200 to 208 nm for HeV (Fig. 6C) CTD_W. These data, together with the characteristic increase at 190 nm, indicate a gain of α -helicity in the presence of increasing concentrations of TFE (Fig. 6B–G). Quantitative analysis of the data, as obtained by spectral deconvolution, revealed a decrease in the disorder content from ~80% to ~25% for NiV (Fig. 6D) and from ~70% to ~20% for HeV (Fig. 6E) CTD_W upon increasing the TFE

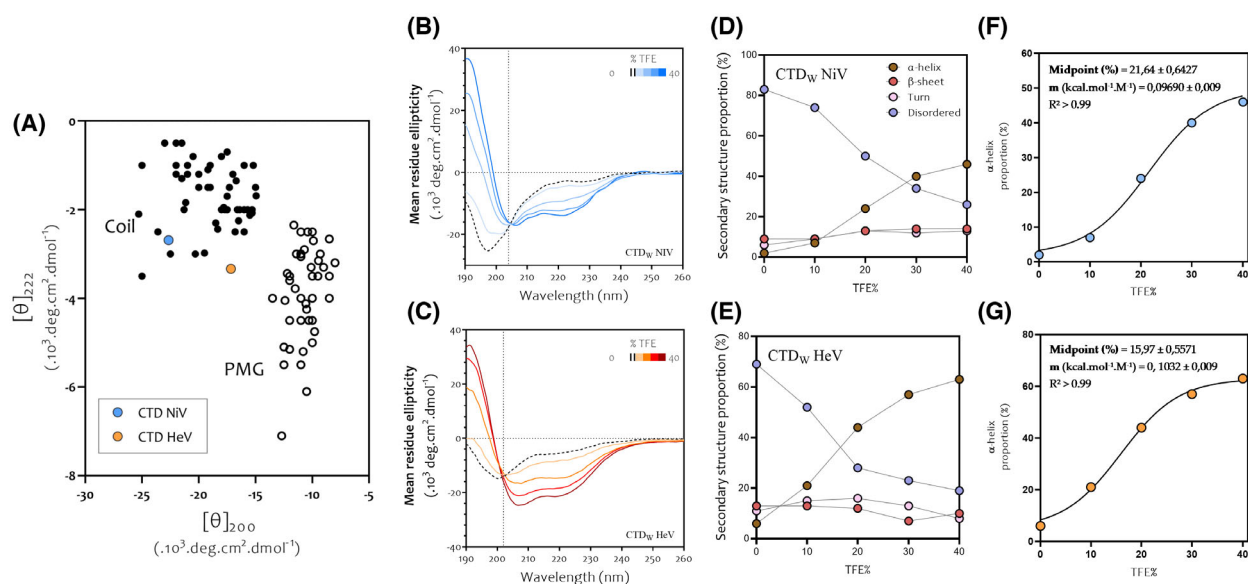


Fig. 6. Secondary structure content of the W protein C-terminal domains (CTD_Ws) from far-UV circular dichroism (CD) studies. (A) Plot of the molar residue ellipticity (MRE) at 222 nm and at 200 nm of a set of well-characterized unfolded, random coil-like (Coil, full circles) or premolten globule-like (PMG, empty circles) proteins (with data replotted from [29] with permission, Copyright © 2002 The Protein Society). The position of Nipah virus (NiV) (in blue) and Hendra virus (HeV) CTD_W (in orange) is shown. (B, C) Far-UV CD spectra of NiV (B) and HeV (C) CTD_W in the presence of increasing concentration of Trifluoroethanol (TFE). Data are representative of three individual experiments. The vertical line indicates the isodichroic point. (D, E) Secondary structure content at increasing concentration of TFE (% v/v) for NiV (D) and HeV (E) CTD_W. (F, G) Transition diagrams of NiV (F) and HeV (G) CTD_W. The α -helical content of the two domains, as derived from the CD spectra, was plotted as a function of the TFE concentration (% v/v), and data were fitted to a sigmoidal curve (corresponding to a two-state transition). The midpoint of transition, the m -value, and the quality of the fitting (R^2) are shown.

concentration from 0% to 40%. This observation, along with the higher α -helical propensity of HeV CTD_W at each point of the TFE titration (Fig. 6D–G), supports a higher folding propensity of HeV CTD_W compared to its NiV counterpart. For both CTD_W, the spectra display an isodichroic point at ~ 203 nm, indicative of a two-state transition [30] (Fig. 6B,C). By plotting the percentage of α -helix as a function of TFE concentration and by fitting the data to a sigmoidal curve (corresponding to a two-state transition), it is possible to extrapolate the midpoint of transition and the m -value (Fig. 6F,G). The m -value is a measure of the osmolyte-dependent cooperativity of disorder-to-order (or order-to-disorder) transitions and measures the efficacy of an osmolyte to either forcing folding or unfolding. Since TFE induces a disorder-to-order transition, driving the equilibrium to a folded state, it acts as a protective osmolyte, characterized by a positive m -value. In agreement, m -values of 0.0969 ± 0.009 kcal·mol⁻¹·M⁻¹ and of 0.1032 ± 0.009 kcal·mol⁻¹·M⁻¹ were obtained for NiV and HeV CTD_W, respectively (Fig. 6F,G). While the m -value is very close between the two CTD_Ws, the midpoint differs significantly, being at $21.64 \pm 0.64\%$ of TFE for NiV and $15.97 \pm 0.56\%$ for HeV (Fig. 6F,G).

Small-angle X-ray scattering (SAXS) studies of CTD_W

To achieve further structural insights on both CTD_Ws, we performed small-angle X-ray scattering (SAXS) studies coupled to SEC (SEC-SAXS). In order to obtain a monodisperse solution and to analyze the monomeric form, experiments were carried out in reducing conditions. The resulting SEC profiles show a single and symmetrical peak for both samples (Fig. 7A,B). The deconvoluted scattering curves (Fig. 7C,D) revealed linearity in the Guinier region (i.e., for $qR_g < 1$) (Fig. 7E,F), with no indication of protein aggregation. The R_g value derived from Guinier analysis is 20.1 ± 0.18 Å for HeV and 21.5 ± 0.12 Å for NiV CTD_W. These values are close to those derived from the $P(r)$ (22.1 ± 0.08 for NiV CTD_W and 21.7 ± 0.1 Å for HeV CTD_W) (Table 1). The pairwise distance distribution functions, $P(r)$, have an elongated and non-symmetrical shape, typical of disordered proteins (Fig. 7G) with a maximum intramolecular distance (D_{\max}) of 78 Å for NiV CTD_W and 74 Å for HeV CTD_W (Table 1). The theoretical R_g values for natively folded, or IDPs, or unfolded proteins of the same length were calculated using Flory's power law (see

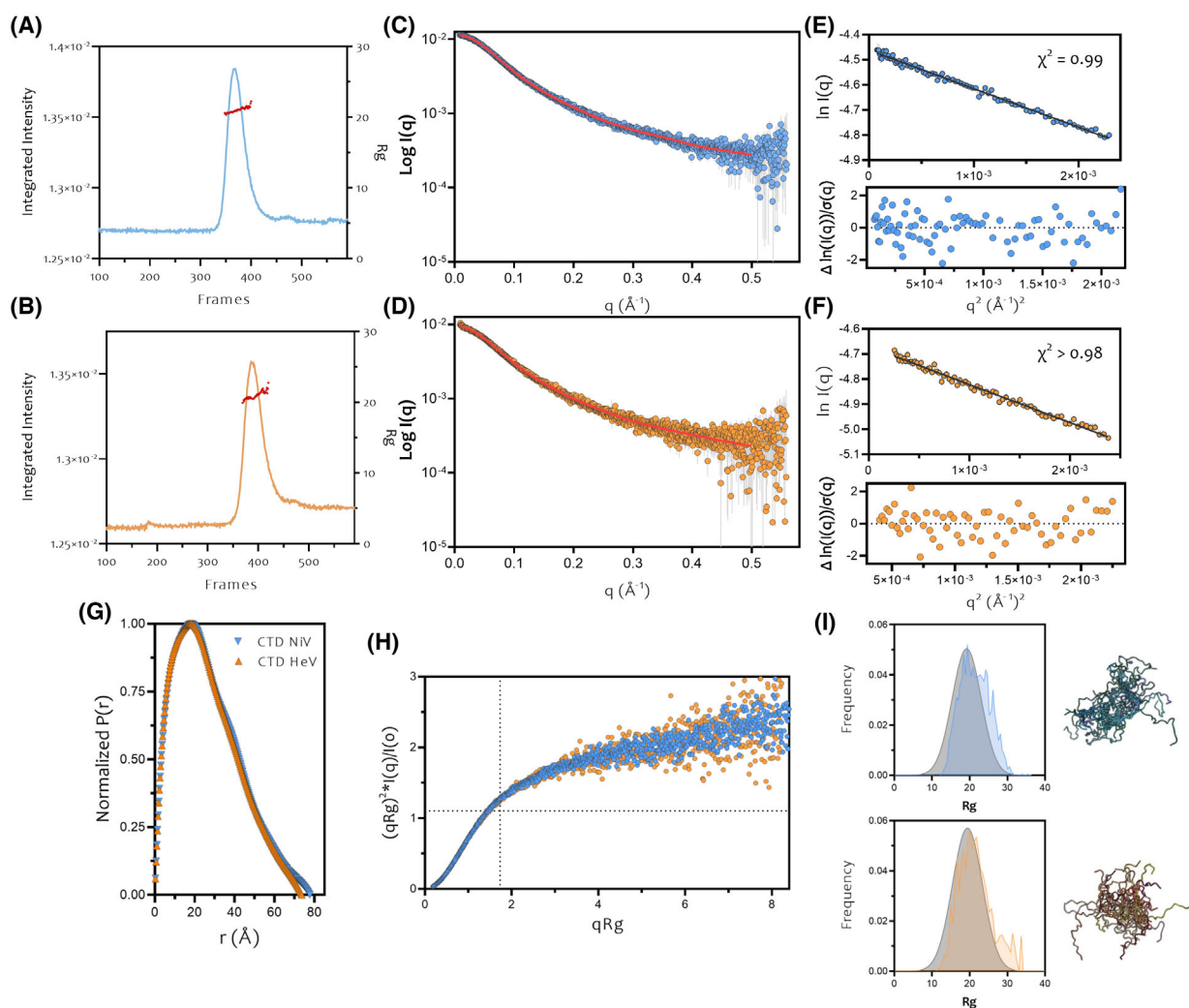


Fig. 7. Conformational properties of the W protein C-terminal domains (CTD_W) from size exclusion chromatography coupled to small-angle X-ray scattering (SEC-SAXS) studies. In blue and orange are shown results pertaining Nipah virus (NiV) CTD_W and Hendra virus (HeV) CTD_W, respectively. Elution peak and R_g distribution for NiV (A) and HeV (B) CTD_W. (C, D). Scattering curves of NiV and HeV CTD_W with solid line corresponding to the curve back-calculated from the Ensemble Optimization Method (EOM) ensemble, full circles to experimental data and gray line to experimental errors. (E, F) Guinier plots (top) showing linearity in the Guinier region and plot of residuals (bottom). The χ^2 of the fit is shown in the top panel. (G) Pairwise distance distribution function for HeV and NiV CTD_W. (H) Normalized Kratky plots showing no clear maximum, consistent with a disordered nature of the sample. (I) Left panel: R_g distribution of the initial ensemble (in gray) and of the selected ensembles of NiV and HeV CTD_W as obtained using EOM. Right panel: conformational ensembles (20 conformers) of NiV and HeV CTD_W as generated by EOM. Data are representative of a single experiment.

Eqn 8 in **Materials and methods**), and they correspond to 11.5 Å for the natively folded state (R_g^{NF}), 19.8 Å for an IDP (R_g^{IDP}), and 23.4 Å for a denatured form (R_g^{U}). The experimentally determined R_g values are close to, and even slightly larger than those expected for IDPs possessing the same number of residues, confirming a highly disordered nature. The R_g -based CI (see Eqn 9 in **Materials and methods**) of NiV CTD_W is 0.162, slightly lower than that of HeV CTD_W (0.196) (see Table 1), a finding in agreement with the higher content in predicted

disorder of the former (Fig. 2). The overall SAXS parameters inferred for both CTD_Ws are shown in Table 1. The shape of the normalized Kratky plots (Fig. 7H), and in particular the lack of a clear maximum, confirms the disordered nature of both CTD_W.

Finally, to achieve a realistic, ensemble description of the CTD_W, we used the program EOM 3.0. For both proteins, the scattering curve back-calculated from the selected ensemble (Fig. 7C,D solid lines) fits well the experimental SAXS data as judged from both

Table 1. Size exclusion chromatography coupled to small-angle X-ray scattering (SEC-SAXS) parameters. $I(0)$: Intensity at zero angle as determined from Guinier approximation; R_g Guinier: Radius of gyration (R_g) as obtained from Guinier approximation; $R_g P(r)$: R_g as obtained from pairwise distance distribution; $R_g Ave^{EOM}$: average value in the final Ensemble Optimization Method (EOM ensemble); D_{max} : maximal intramolecular distance from $P(r)$; $R_{flex,pool}$: flexibility index in the initial pool; $R_{flex,ens}$: flexibility index in the final ensemble; χ^2 : quality of the fit between experimental and back-calculated data from the EOM ensemble; P -value: quality of the fit between experimental and back-calculated data from the EOM ensemble, as provided by CORMAP. R_g^{NF} : R_g expected for a natively folded protein based on Flory's equation (see Eqn 8). R_g^{IDP} : R_g expected for an Intrinsically Disordered Protein (IDP) based on Flory's Eqn (8); R_g^U : R_g expected for a chemically denatured protein based on Flory's equation (see Eqn 8). CI, compaction index. SEC-SAXS data can be retrieved from the SASBDB [46] under codes SASDUQ3 and SASDUR3 for the set of data of Hendra virus (HeV) CTD_W and of Nipah virus (NiV) CTD_W, respectively.

	CTD _W NiV	CTD _W HeV
$I(0)$	$0.012 \pm 2.2 \times 10^{-5}$	$0.0093 \pm 3.3 \times 10^{-5}$
R_g Guinier (Å)	21.5 ± 0.12	20.1 ± 0.18
$R_g P(r)$ (Å)	22.1 ± 0.08	21.7 ± 0.1
$R_g Ave^{EOM}$ (Å)	21.6	21.4
D_{max} (Å)	78	74
$R_{flex,pool}$ (%)	86.41	85.52
$R_{flex,ens}$ (%)	86.20	85.03
χ^2	0.980	1.005
P -value CORMAP	0.408	0.986
R_g^{NF} (Å)	11.5	11.5
R_g^{IDP} (Å)	19.8	19.8
R_g^U (Å)	23.4	23.4
CI	0.162	0.196

χ^2 and CORMAP ($\chi^2 = 0.98$, $P = 0.408$ for NiV and $\chi^2 = 1.002$, $P = 0.986$ for HeV CTD_W) (Table 1). For both CTD_Ws, the resulting final R_g distributions (Fig. 7I) are unimodal and close to that of the initial pool, confirming that the conformational ensembles consist of randomly distributed and non-compact flexible conformations. To quantify the flexibility of the system, we used the R_{flex} parameter, which is equal to 100% for fully flexible systems. The R_{flex} was estimated to be 86.20% for NiV and 85.03% for HeV CTD_W, to be compared to 86.41% and 85.52% for their respective initial pools. These data, together with previous evidence from SDS/PAGE and CD analyses, unambiguously confirm the disordered nature of the NiV and HeV CTD_W.

Fibrillation abilities of the CTD_W

We previously documented the ability of the *Henipavirus* W proteins [21] and, more specifically, of their

PNT3 region (aa 200–310 of P/V/W) to fibrillate [23,24]. Here, we investigated whether the isolated CTD_W could also fibrillate *per se* and/or could contribute to the process. To this end, we carried out binding assays with Thioflavin T (ThT), a well-known amyloid-specific dye whose fluorescence increases upon binding to amyloids [31]; we compared the ThT binding abilities of the CTD_W to those of the NiV and HeV W proteins and of their NTD. To this end, we purified the W proteins as previously described [21], and the NiV and HeV NTD via a two-step protocol (Fig. 8A). The NiV and HeV NTD were purified in large amounts from both the soluble and insoluble fraction of *E. coli* through IMAC and IEC (Fig. 8A, right panel). As previously observed for pDEST14-encoded forms [2], in SDS/PAGE they show an apparent molecular mass of ~60 kDa, in contrast with the expected molecular mass of ~45 kDa (Fig. 8A) and consistent with their disordered nature, as already documented [2].

No ThT fluorescence increase was observed in the presence of the CTD_W, consistent with lack of intrinsic fibrillogenic properties (Fig. 8B), while increasing fluorescence was observed for both W proteins and their NTD, in line with the already shown fibrillation ability of the two [25]. Negative-staining transmission electron microscopy (TEM) studies unambiguously showed that the CTD_Ws, even after a prolonged incubation (72 h) at 37 °C at a concentration as high as 100 μM, are unable to fibrillate (Fig. 8C). Note that binding assays to Congo Red, another dye currently used to detect amyloids, yielded false-positive results, that is, a significant shift in the Congo Red absorption spectrum (data not shown), reflecting Congo Red aspecific binding to non-fibrillar cationic species [32,33].

Interaction between the NTD and the CTD_W and mutual impact on their conformational properties

We next investigated what role the CTD_W might play in the context of the full-length W protein and whether it could affect the conformational properties of the NTD.

We first probed the interaction between the two domains. To this end, we used a pull-down assay in which a His-tagged form of the CTD_W was immobilized onto an IMAC resin (preloaded with Nickel ions) and an untagged form of the NTD was added. The latter form was obtained by TEV protease digestion of the product obtained from IMAC, resulting in removal of the vector-encoded stretch encompassing the histidine tag, followed by an IEC step (Fig. 8A, left panel). MS analysis of the peptides resulting from tryptic

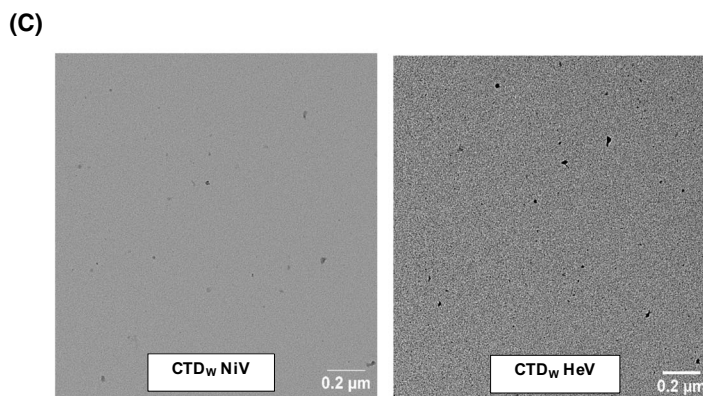
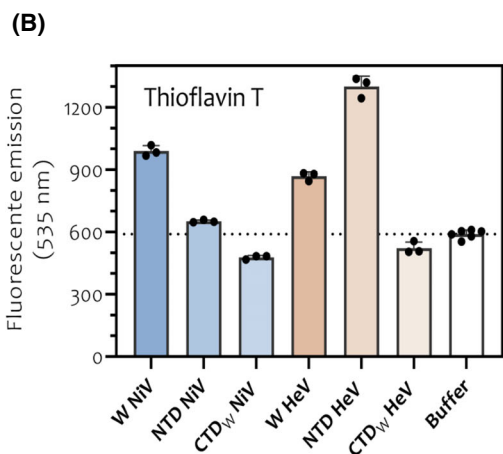
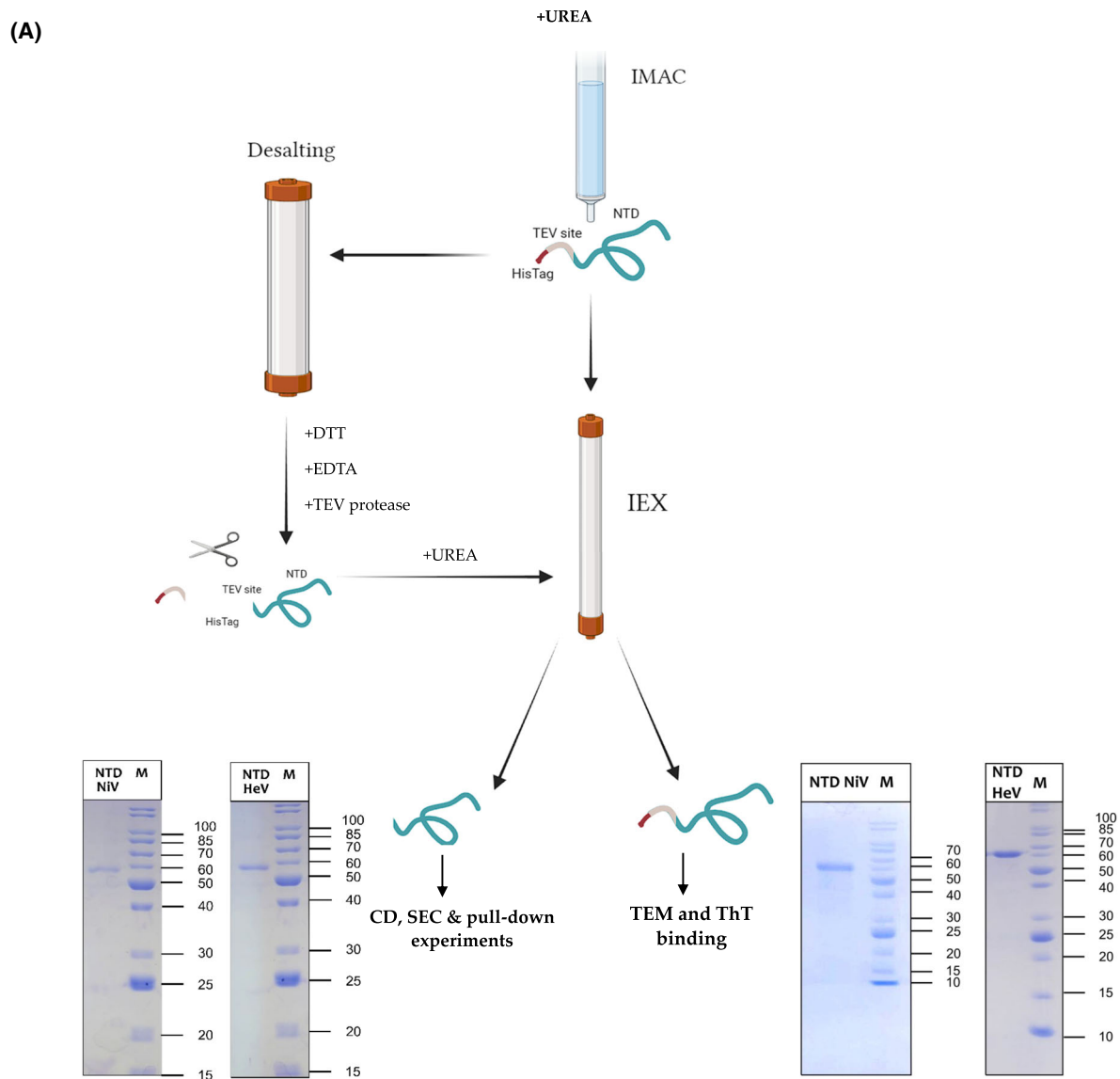


Fig. 8. W protein N-terminal domain (NTD) purification pipeline and fibrillation abilities of the two W C-terminal domains (CTD_Ws). (A) Purification of untagged and tagged NTDs. Schematic representation of the two purification processes used for Nipah virus (NiV) and Hendra virus (HeV) NTD and SDS/PAGE analysis of the untagged (bottom left) and tagged (bottom right) NTDs. Data are representative of two independent analyses. (B) Fluorescence emission at 535 nm of Thioflavin T (ThT) in the presence of full-length W, NTD, and CTD_W from both NiV and HeV. The error bar corresponds to the standard deviation, with $n=3$ for protein samples in both panels and $n=5$ for the buffer. (C) Negative-staining transmission electron microscopy (TEM) micrographs of NiV and HeV CTD_W (at 100 μ M) after 72 h of incubation at 37 °C. Scale bar = 0.2 μ m.

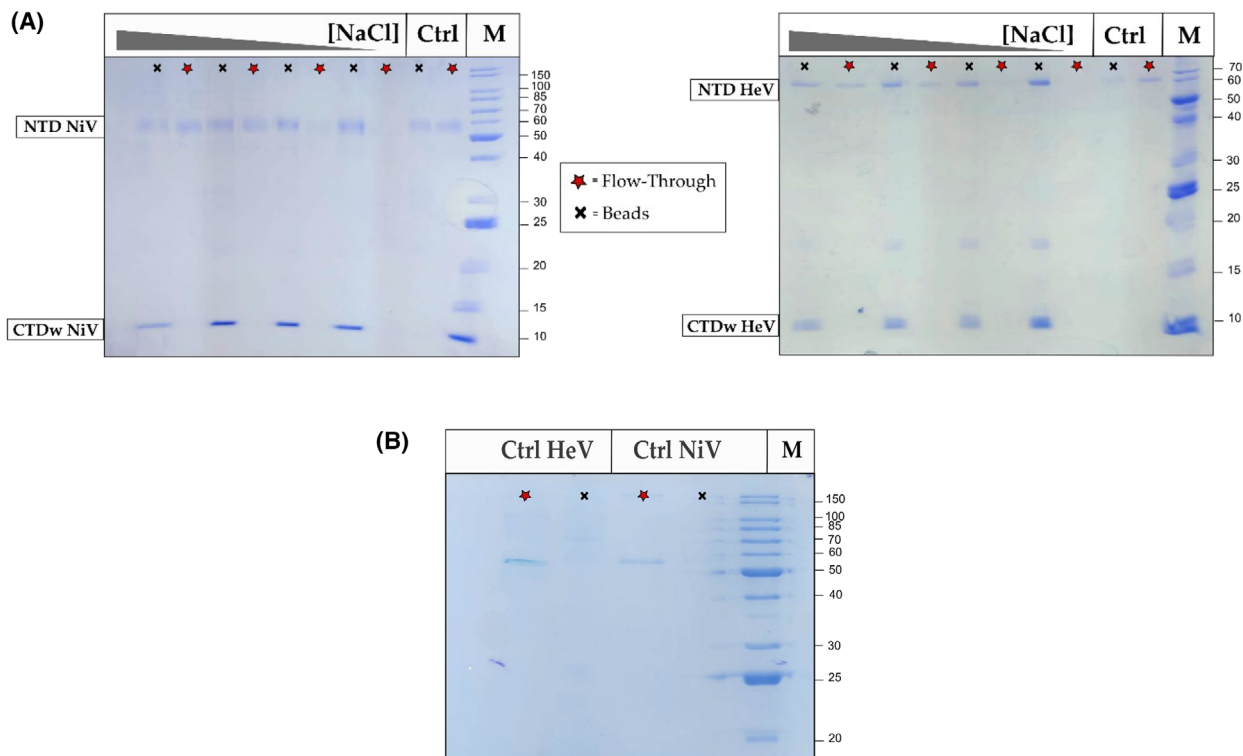


Fig. 9. Interaction between the W protein N-terminal domain (NTD) and the C-terminal domain (CTD_W) from pull-down assays. Data are representative of three independent experiments. (A) SDS/PAGE analysis of the pull-down assay between untagged NTD and tagged CTD_W for both Nipah virus (NiV) and Hendra virus (HeV). Red stars correspond to lanes loaded with the unretained fraction (Flow-Through) and black crosses to beads. M, molecular markers; Ctrl, control without CTD_W, at NaCl = 50 mM, [NaCl] = samples with decreasing salt concentration from left to right, that is, 1 M, 500, 200, and 50 mM. (B) SDS/PAGE analysis of the pull-down assay of untagged NTD and tagged CTD_W in the presence of 1 M NaCl for both NiV and HeV.

digestion confirmed the identity of the purified untagged NTD proteins (Fig. 4C,D). The resulting product (Fig. 8A, left panel) was used in all the experiments described in this section. In this assay, any interaction between the two domains would result in the NTD being retained on the beads, while in the absence of interaction, the NTD would only be found in the unretained fraction. Moreover, given that the NTD and the CTD_W have different isoelectric points (~ 4.6 and ~ 11.6 , respectively), we tested whether this interaction was affected by salt, and used four different NaCl concentrations (50 mM, 200 mM, 500 mM, and

1 M). As a control, the NTD was incubated with the resin in the absence of the CTD_W at the lowest salt concentration (50 mM). The results of the assay are shown in Fig. 9A.

For both HeV and NiV, some non-specific interaction between the NTD and the resin (control of both experiments) can be observed (see faint NTD band in lanes loaded with beads in Fig. 9A). However, as soon as the CTD_W is added to the system, the NTD is fully retained on the resin at the two lowest salt concentrations (50 and 200 mM). At higher salt concentrations, the interaction is partly impaired, as judged from the

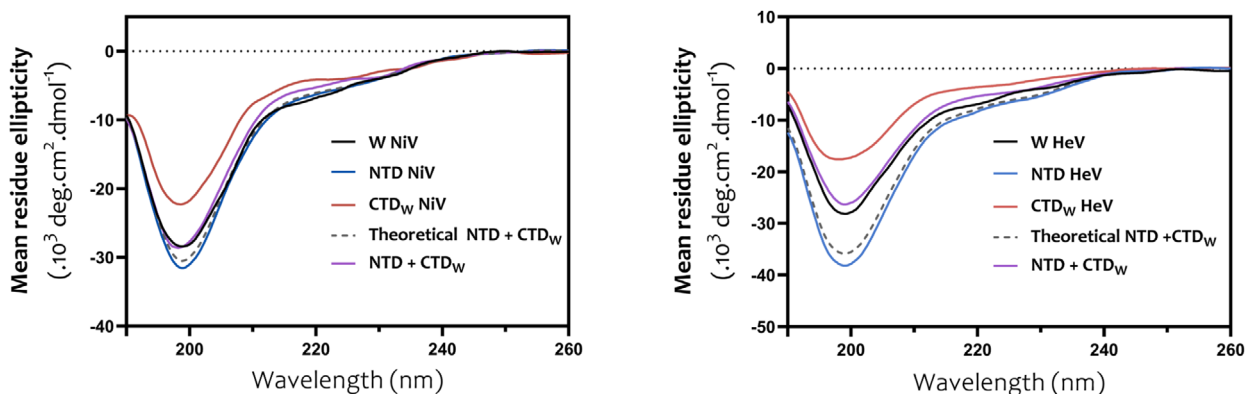


Fig. 10. Contribution of the W C-terminal domain (CTD_W) to the secondary structure content of the W proteins. Far-UV circular dichroism (CD) spectra of the full-length W proteins and of their isolated domains, along with the spectra of equimolar mixtures of their CTD and N-terminal domain (NTD) either calculated (Theoretical) or experimentally obtained for Nipah virus (NiV) (on the left) and Hendra virus (HeV) (on the right). All the spectra present maximum intensity at ~ 200 nm. Data are representative of three independent experiments.

reappearance of the NTD band in the Flow-Through at 500 mM and 1 M NaCl. However, not all the NTD is washed out from the resin, indicating that, even if weaker, there is still an interaction taking place (Fig. 9A). An additional control with the NTD alone in the presence of 1 M salt enabled ruling out that the interaction observed at higher salt concentrations does not arise from non-specific interactions with the resin (Fig. 9B). The salt dependence of the interaction between the NTD and the CTD_W suggests that the interaction is of electrostatic nature.

After having demonstrated an interaction between the NTD and the CTD_W, our next objective was to assess how this interaction might affect the secondary structure content of these two domains. Specifically, we were interested in assessing possible effects both *in trans* (i.e., when mixing the NTD and the CTD_W) and *in cis* (i.e., in the context of the full-length W protein). To this end, we used far-UV CD spectroscopy. For both NiV and HeV, the CD spectra of the W proteins and of their isolated constituent domains (NTD and CTD_W) were compared to the CD spectrum of a mixture containing equimolar amounts of CTD_W and NTD (20 μ M each). Additionally, a theoretical average spectrum, corresponding to the spectrum of the mixture expected in case no rearrangement of the secondary structure takes place, was calculated (see [Materials and methods](#), Eqn 6).

For both NiV and HeV, all the spectra exhibit a pronounced peak at ~ 200 nm and a very low ellipticity at 190 nm, indicating a prevalently disordered nature (Fig. 10). In spite of this common feature, there is a noticeable difference among the spectra of the NTD, of the CTD_W and of the W proteins. In particular, the CD

spectra of the two NTD (in blue in Fig. 10) display the highest intensity at ~ 200 nm, while the spectra with the lowest intensity peak are those of the two CTD_Ws (in red in Fig. 10). The spectra of the W proteins (in black in Fig. 10) are situated between the spectra of the NTDs and of the two CTD_Ws, indicating that the W proteins are less disordered than their isolated NTDs. This observation is consistent with an ordering effect brought by the CTD_W in the context of the W protein (i.e., when the NTD and the CTD_W are *in cis*).

For both NiV and HeV, the theoretical spectrum of the mixture, as expected in the absence of any secondary structure rearrangement (dashed gray line in Fig. 10), is very close to the spectrum of the NTD. However, the spectra obtained for equimolar NTD + CTD_W mixtures (in purple in Fig. 10) deviate from the corresponding theoretical spectra (dashed line in Fig. 10) and are very close to the spectra of the full-length W proteins, indicating that the CTD_W exerts an impact on the NTD also when the two domains are *in trans*.

These findings indicate that the interactions established by the NTD and the CTD_W lead to a rearrangement in the secondary structure content of the individual domains and that the resulting overall gain of structure observed *in cis* and *in trans* is comparable.

Next, we investigated the impact of the NTD-CTD_W interaction on the hydrodynamic properties of the W proteins. To this end, for both NiV and HeV, we analyzed the SEC profile of the W proteins and of their NTDs. Since we wanted to infer information on possible compaction effects brought by the CTD_W in the context of the W protein, we focused our analysis on the peak corresponding to the monomeric species and disregarded any additional higher molecular mass

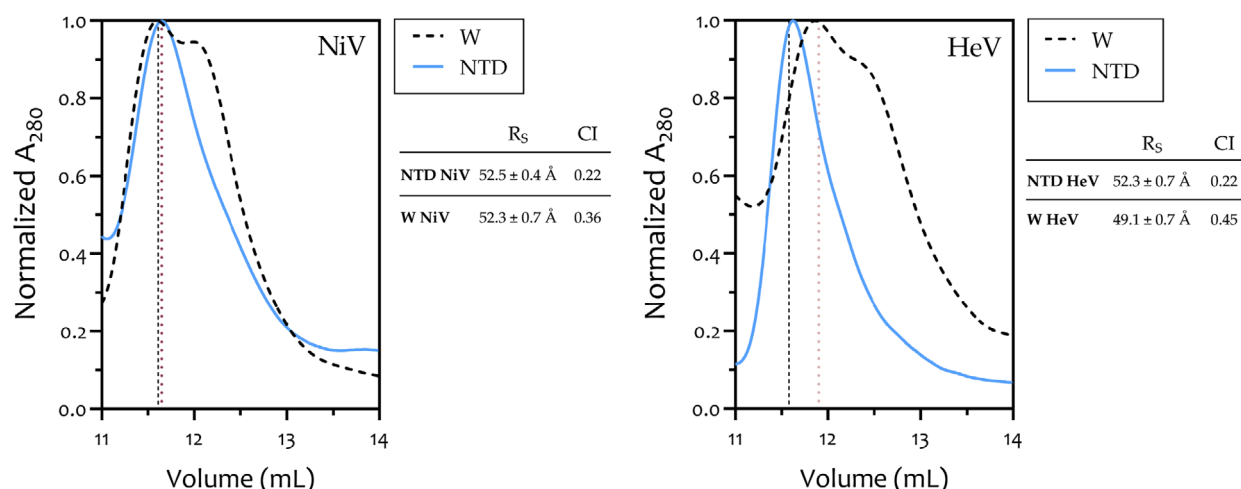


Fig. 11. Contribution of the W protein C-terminal domain (CTD_W) to compaction properties of the W proteins. Size exclusion chromatography (SEC) profiles of the Nipah virus (NiV) (left panel) and Hendra virus (HeV) (right panel) W proteins (dashed black line) and of their constituent N-terminal domains (NTDs) (continuous blue line). For the sake of clarity, only the portion of the chromatogram corresponding to the peak of the monomeric species is shown. The vertical gray and black dashed lines highlight the elution volume of the W and NTD peaks, respectively. The shoulder visible for both W proteins corresponds to a more compact intramolecularly disulfide-bridged species, as described in detail in Ref. [25]. For both W proteins and for their NTD, the relative inferred Stokes radii (R_S) and Compaction Index (CI) values are shown. Data are representative of three independent experiments.

peaks. In the case of NiV, the W protein and the NTD display a very similar elution profile with a peak at ~ 11.65 mL (Fig. 11), thereby leading to very close inferred R_S values ($52.3 \pm 0.7 \text{ \AA}$ for W and $52.5 \pm 0.4 \text{ \AA}$ for NTD) (Fig. 11). In the case of HeV, the profile of the W protein is slightly different with respect to that of its NTD, with derived R_S values of $49.1 \pm 0.7 \text{ \AA}$ and of $52.3 \pm 0.7 \text{ \AA}$, respectively. Analysis of the CI revealed that the W proteins are comparatively more compact than their constituent NTDs. Based on the above-described pull-down assay and CD results, it can be hypothesized that the interaction between the NTD and the CTD_W causes the latter to fold back on the NTD, leading to W protein compaction.

Since the CI of the NTD from both viruses is the same, the higher compactness of the HeV W protein is ascribable either to the higher CI of the HeV CTD_W or to a higher compaction effect brought by the HeV CTD_W compared to that of NiV, or to a combination of both.

Impact of CTD_W on W fibrillation

Having established the interaction between the two domains and having shown that the CTD_W does not fibrillate *per se*, we next addressed the question as to whether the CTD_W plays a regulatory role in W fibrillation.

To this end, we used TEM and ThT assay to monitor fibril formation at time 0 (T0) and after 72 h (T72) of incubation at 37 °C. To enable a consistent T0 analysis, samples stored at -80 °C were thawed and then incubated for 30 min in the presence of 5 mM DTT prior to being buffer exchanged in HBS, thus preventing the formation of oxidation-induced aggregates. We analyzed the isolated CTD_W and NTD, along with the W proteins and compared results to those obtained with equimolar NTD + CTD_W mixtures.

Results obtained at T0 are shown in Fig. 12. As previously shown (Fig. 8C), no fibrils were observed in micrographs of NiV and HeV CTD_W samples (Fig. 12A,B). In the case of NiV, no fibrils could be detected for the NTD either (Fig. 12A). By contrast, some fibrillar structures were found to be already present for both the NiV W protein and the NiV NTD + CTD_W mixtures (Fig. 12A). In ThT binding assays of NiV samples, the highest fluorescence signal was observed for the W protein, while the signal obtained with the NTD is close to that of the control (buffer), as is the one of the NTD + CTD_W mixture (Fig. 12C). As already observed (Fig. 8B), the signal obtained with the NiV CTD_W sample is even lower than that of the buffer (Fig. 12C).

In contrast with NiV, TEM analysis of HeV NTD at T0 revealed the presence of several very short fibrils or fibril precursors (Fig. 12B), supporting a higher fibrillogenic potential of HeV NTD compared to NiV

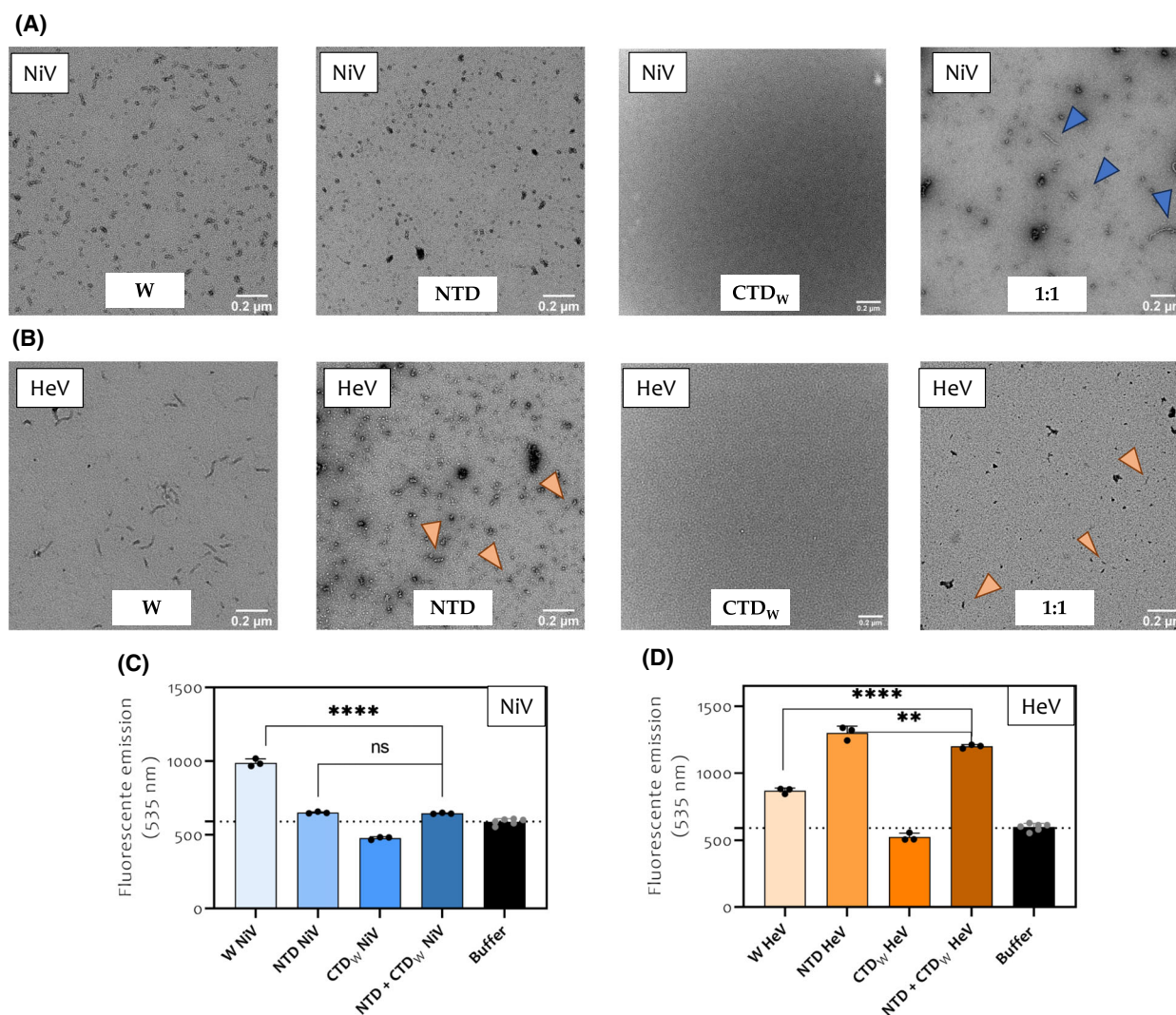


Fig. 12. Fibrillation abilities of Nipah virus (NiV) (A, C) and Hendra virus (HeV) (B, D) W proteins, of their constituent N-terminal and C-terminal domains (NTD and CTD_W), and of equimolar NTD+CTD_W mixtures at T0. Data are representative of three independent experiments. (A, B) Transmission Electron Microscopy (TEM) micrograph of W, NTD, CTD_W, and of an equimolar NTD + CTD_W mixture. Small fibrils are indicated by blue (NiV) and orange (HeV) arrowheads. (C, D) Thioflavin T binding assay. Scale bar = 0.2 μm. The error bar corresponds to the standard deviation. The asterisks indicate statistically significant difference obtained with one-way ANOVA test. ns, not significant ($P > 0.05$). **** $P < 0.0001$; ** $0.01 > P < 0.001$.

NTD. In line with this trend, fibrils of HeV W were found to be longer than those formed by NiV W (cf. Fig. 12A,B). Whether this finding is related to the higher compactness of HeV W deserves further investigations. For the HeV NTD+CTD_W mixture, small fibrils with dimensions similar to those observed for HeV NTD are detectable (Fig. 12B). In ThT binding assays of HeV samples, the highest signal is observed for the NTD. As in the case of NiV, the NTD+CTD_W mixture yields a signal close to that of the NTD (Fig. 12D).

Results obtained for NiV samples after 72 h of incubation are shown in Fig. 13. Fibrillar structures can be observed in the micrographs of W, NTD, and of the NTD+CTD_W mixture (Fig. 13A). These results are mirrored in the ThT binding assay, where all these conditions correspond to a strong signal (Fig. 13C). Long fibrils are prominent in the W and NTD+CTD_W samples, while NTD samples show shorter, more numerous fibrils (Fig. 13A).

The same pattern is observed in HeV samples after 72 h, with W and NTD+CTD_W samples displaying

long fibrils and NTD samples showing shorter ones (Fig. 13B). In agreement with these results, a strong signal is observed in ThT binding assays for W, NTD, and the NTD + CTD_W mixture (Fig. 13D), with the NTD yielding a signal higher than that of W and mixture, as already observed at T0.

At T0, NiV fibrils are found only in micrographs of W and NTD + CTD_W mixtures, but not in those of the NTD alone (Fig. 13A), supporting a role for the CTD_W, both *in cis* and *in trans*, in fibril formation. For HeV, at T0 the micrographs revealed not only that the NTD forms fibrils already at T0, suggesting a higher fibrillogenic potential compared to NiV, but also that the NTD + CTD_W mixture has an intermediate phenotype with respect to W and NTD (Fig. 12B). In fact, while at T0 for the W protein short fibrillar structures can be observed, in the NTD + CTD_W mixture the observed fibrils are even shorter (Fig. 12B) and less abundant, supporting a negative impact on nucleation when the CTD_W is provided *in trans*. The ThT binding assay results (Fig. 12D) appear to support these observations, with the fluorescence of the NTD + CTD_W mixture being slightly lower than that of the NTD alone.

At longer incubation times (T72) for both NiV and HeV, the fibrils observed in the NTD + CTD_W mixtures (Fig. 13A,B,E) show a morphology and dimensions similar to those of the W protein, rather than those of the NTD alone. This supports a role of the CTD_W, both *in cis* and *in trans*, in fibril elongation for both NiV and HeV. Interestingly, the micrographs of the NTD + CTD_W mixtures show fibrils longer than those observed for the W protein (Fig. 13A,B,E). The results suggest that the kinetics of fibril formation are accelerated when the CTD_W is provided *in trans* compared to when it is provided *in cis*, albeit to varying degrees. Furthermore, and again in support of a negative effect on nucleation when the CTD_W is provided *in trans*, the HeV fibrils observed for the NTD + CTD_W mixture are less numerous than those observed for W and NTD. This observation is also reflected in the ThT results at both T0 and T72 (Figs 12D and 13D), where, although W, NTD, and NTD + CTD_W all produce high signals, the highest fluorescence is recorded for the NTD alone. Finally, while at T0, the ThT binding assays of NiV samples show a lower signal for both NTD and NTD + CTD_W mixture compared to W (Fig. 12C), after 72 h of incubation they all yield a strong signal, coherent with the formation of fibrils as observed in the respective micrographs (Fig. 12A,C).

Altogether, the results described above support the conclusion that the CTD_W, although devoid of intrinsic fibrillogenic properties, favors fibril formation in both NiV and HeV.

Conclusions

This study provides insights into the conformational properties of the C-terminal domain of the *Henipavirus* W proteins (CTD_{WS}). By conducting a detailed biophysical characterization, we demonstrate the disordered nature of both NiV and HeV CTD_W. Specifically, using CD and SEC-SAXS (Figs 5C, 6, and 7), we showed that the CTD_W of both viruses is intrinsically disordered, with the HeV CTD_W being slightly more compact than that of NiV (Table 1). Further investigation into the redox state of the CTD_W from both viruses also revealed that it dimerizes via the formation of a disulfide bridge (Figs 3C,D and 5A,B), with the dimerization being accompanied by a gain of structure (Fig. 5C).

The CTD_W was found to be both unable to bind to ThT (Fig. 8B) and to fibrillate (Fig. 8C). Pull-down experiments unveiled that the NTD and the CTD_W interact with each other (Fig. 9), with such interaction being salt-dependent and thus of electrostatic nature. CD studies revealed a gain of secondary structure by the NTD brought by the CTD_W, both when the latter is provided *in cis* (i.e., in the context of the W protein) and *in trans* (i.e., when added in equimolar amounts to the NTD) (Fig. 10). Additionally, SEC analysis showed that the W proteins are more compact than their isolated NTD (Fig. 11), consistent with a compaction effect on the NTD brought by the CTD_W. These studies also unveiled that the HeV W is more compact than its NiV counterpart (Fig. 11), likely reflecting a higher compactness of the HeV CTD_W combined with a higher compaction effect brought by the CTD_W of HeV compared to NiV.

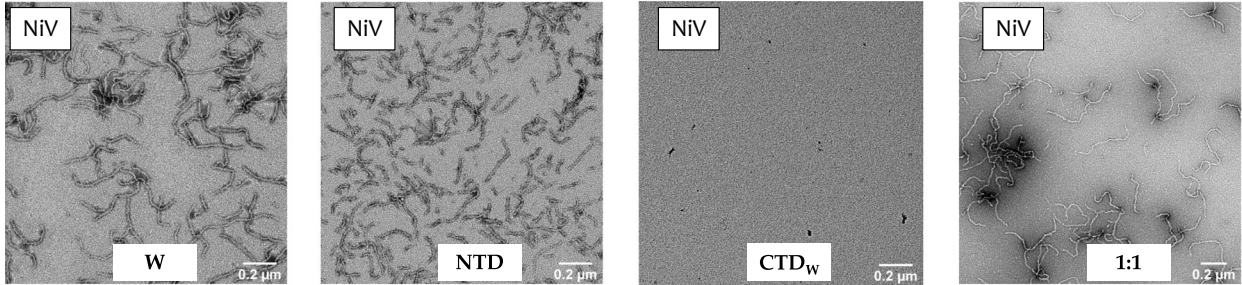
ThT binding assays, coupled to TEM studies, highlighted the impact of the CTD_W on the fibrillation abilities of the NTD. Results show that the CTD_W has a significant effect on the fibrillation properties of the NTD (Figs 12 and 13). Specifically, following a prolonged incubation under native conditions, the CTD_W was found to favor NTD fibrillation. Furthermore, in the case of HeV, the addition of CTD_W, both *in cis* and *in trans*, was found to exert a negative effect on nucleation (Figs 12 and 13).

Overall, this study enhances our comprehension of the structural properties of the *Henipavirus* W proteins and highlights the intricate interplay between their NTD and CTD_W. In light of the speculated role of W fibrils in counteracting the antiviral response [21,23,24,34], unraveling the impact of the CTD_W on NTD fibril formation has direct relevance to the elucidation of the molecular mechanisms underlying *Henipavirus* pathogenesis and may thus contribute to the development of targeted therapies. It also constitutes

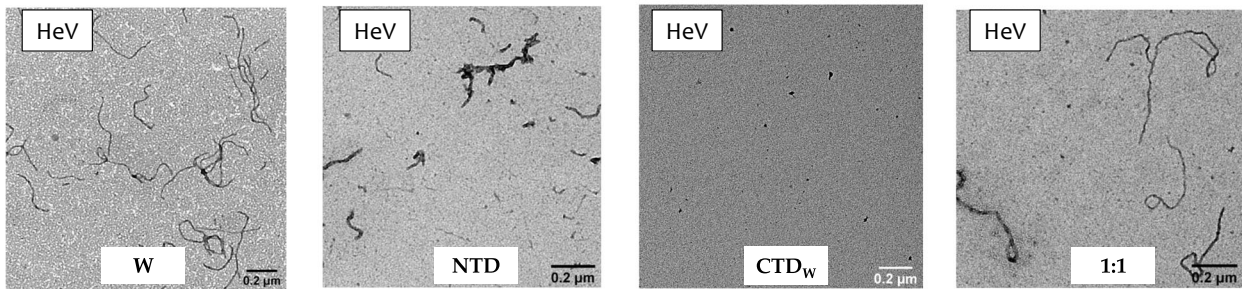
an asset for future studies aimed at specifically investigating the impact of the CTD_W and of its dimerization on W fibril formation in the context of transfected

and/or infected cells and paves the way toward new appealing research avenues focused on the inhibition of W fibrillation.

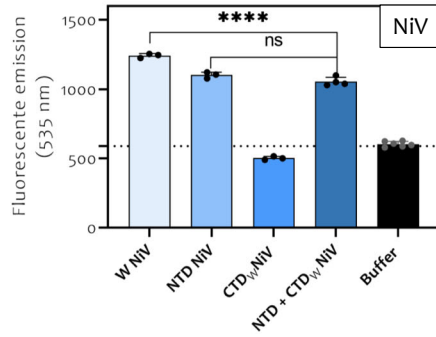
(A)



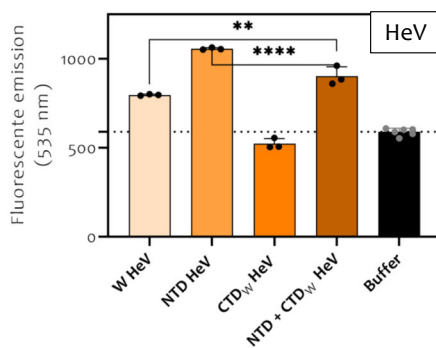
(B)



(C)



(D)



(E)

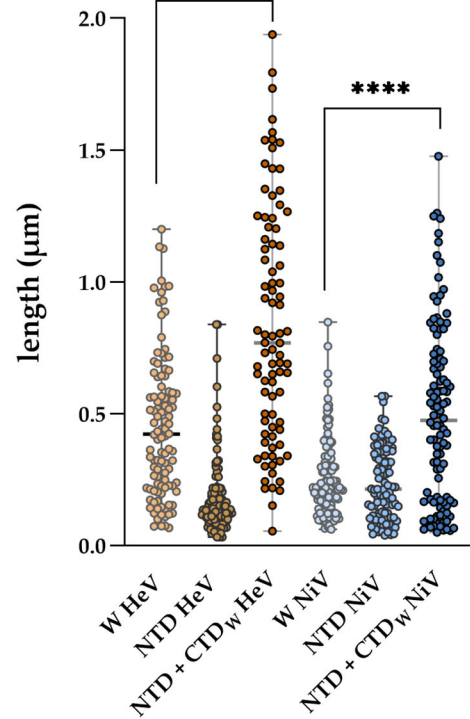


Fig. 13. Fibrillation abilities of Nipah virus (NiV) (A, C) and Hendra virus (HeV) (B, D) W proteins, of their constituent N-terminal and C-terminal domains (NTD and CTD_W), and of equimolar NTD+CTD_W mixtures at T72. Data are representative of three independent experiments. (A, B) Transmission Electron Microscopy (TEM) micrograph of W, NTD, CTD_W, and of an equimolar NTD+CTD_W mixture (C, D) Thioflavin T binding assays. Scale bar = 0.2 μm. The error bar corresponds to the standard deviation. The asterisks indicate statistically significant difference obtained with one-way ANOVA test. ns, not significant ($P > 0.05$). **** $P < 0.0001$; ** $0.01 > P < 0.001$. (E) Fibril contour length distributions at T72 for W, NTD, and NTD+CTD_W from both HeV and NiV. Asterisks indicate a statistically significant difference with P -value < 0.0001 (Student's t -test). The micrograph of HeV CTD_W shown in panel B is the same control experiment as that shown in panel C of Fig. 8.

Materials and methods

Disorder predictions

Disorder predictions were obtained using a modified in-house version of the M_ETaserver of DisORder MeDor [26]. MeDor allows fast, simultaneous analysis of a query sequence by multiple predictors. It provides a graphical interface with a unified view of the output of multiple disorder predictors. MeDor includes a prediction of transmembrane regions (Phobius, <https://phobius.sbc.su.se/instructions.html>) and of secondary structure prediction (SSP), based on the StrBioLib library of the Pred2ary program, Hydrophobic Cluster Analysis (HCA), DorA (an unpublished predictor developed in the AFMB lab that uses size and abundance of hydrophobic clusters in the HCA plot to predict disorder), MoreRONN, FoldUnfold, FoldIndex, and MobiDB-lite. MobiDB-lite is a metapredictor that uses eight different disorder predictor (GlobPlot, three versions of ESpritz, two versions of IUPred, and two versions of DisEMBL) to derive a consensus that is refined to remove short disordered regions and keep only those that consist of at least 20 consecutive residues predicted as disordered. Beyond displaying the consensus prediction of MobiDB-lite (Consensus MobiDB), MeDor also generates two consensus disorder predictions, one corresponding to regions predicted as disordered by the majority of predictors (Consensus 1) and one corresponding to regions predicted as disordered by all the predictors (Consensus 2).

Generation of the W CTD and NTD expression constructs

The various constructs were obtained by PCR followed by Gateway[®] cloning (Invitrogen, Carlsbad, CA, USA). Whatever the construct, the amplicons were digested with DpnI (New England Biolabs, Ipswich, MA, USA) to remove the methylated DNA template and then transferred into the pDONR vector (Invitrogen) through a Gateway[®] BP clonase-mediated recombination reaction (Invitrogen). From the resulting entry vector, the sequence of interest was then transferred into either the pDEST170I or the pDEST14 vector using Gateway[®] LR clonase (Invitrogen). Unless differently specified, the constructs encoding the NTD and the CTD of the NiV and HeV W proteins were

PCR-amplified using the pDEST170I/NiV W and pDEST170I/HeV W plasmids as templates [21]. These plasmids enable the expression, under the control of the T7 promoter, of the HeV W (UniProt code [O55777](#)) and NiV W (UniProt code [Q997F2](#)) proteins with an N-terminal, vector-encoded MSYYHHHHHHLESTSLYKKAGS amino acid stretch followed by a TEV protease cleavage site (ENLYFQG) [21].

The strategy used to generate the HeV NTD construct is described in Fig. 14A. For the generation of the NiV NTD construct, we used the overlap PCR extension technique as shown in Fig. 14B, as the protocol used for the generation of the HeV NTD construct failed to yield the desired construct after repeated attempts. The pDEST14/NiV NTD construct, which codes for the NiV NTD with a C-terminal His-tag, has already been described [2]. The resulting NTD constructs (named pDEST170I/HeV NTD and pDEST170I/NiV NTD) allow the production of a recombinant protein in which the NTD is preceded by the same stretch of 22 vector-encoded residues of the W constructs and by a TEV protease cleavage site. The 22 vector-encoded residues can be cleaved off by TEV protease digestion to yield the NTD bearing only a non-native N-terminal glycine residue.

For the CTD_W constructs, we first generated constructs in pDEST170I by PCR, using the pDEST170I/NiV W and pDEST170I/HeV W constructs as templates. Primers were designed to amplify the CTD-encoding region, with forward primer starting with an attB1 sequence and reverse primer starting with an attB2 sequence. Subsequently, we generated CTD_W constructs encoding a fusion protein consisting of TRX, of a TEV protease cleavage site, and of the CTD_W followed by a non-native tyr residue (to enable detection of the protein at 280 nm) and by a C-terminal His-tag. To this end, we used the overlap PCR extension technique as detailed in Fig. 14C. The resulting constructs (named pDEST14/TRX-HeV CTD_W and pDEST14/TRX-NiV CTD_W) encode TRX-TEV-CTD_W-Y-His6 fusion proteins in which the TRX tag can be cleaved off by TEV protease digestion thus leading to a recombinant protein consisting of the CTD_W with no additional non-native N-terminal residues and with a C-terminal YHHHHHH tag.

The sequence of all the primers used to generate the various constructs is provided in Table 2. The sequence of all the constructs was checked by sequencing (Genewiz,

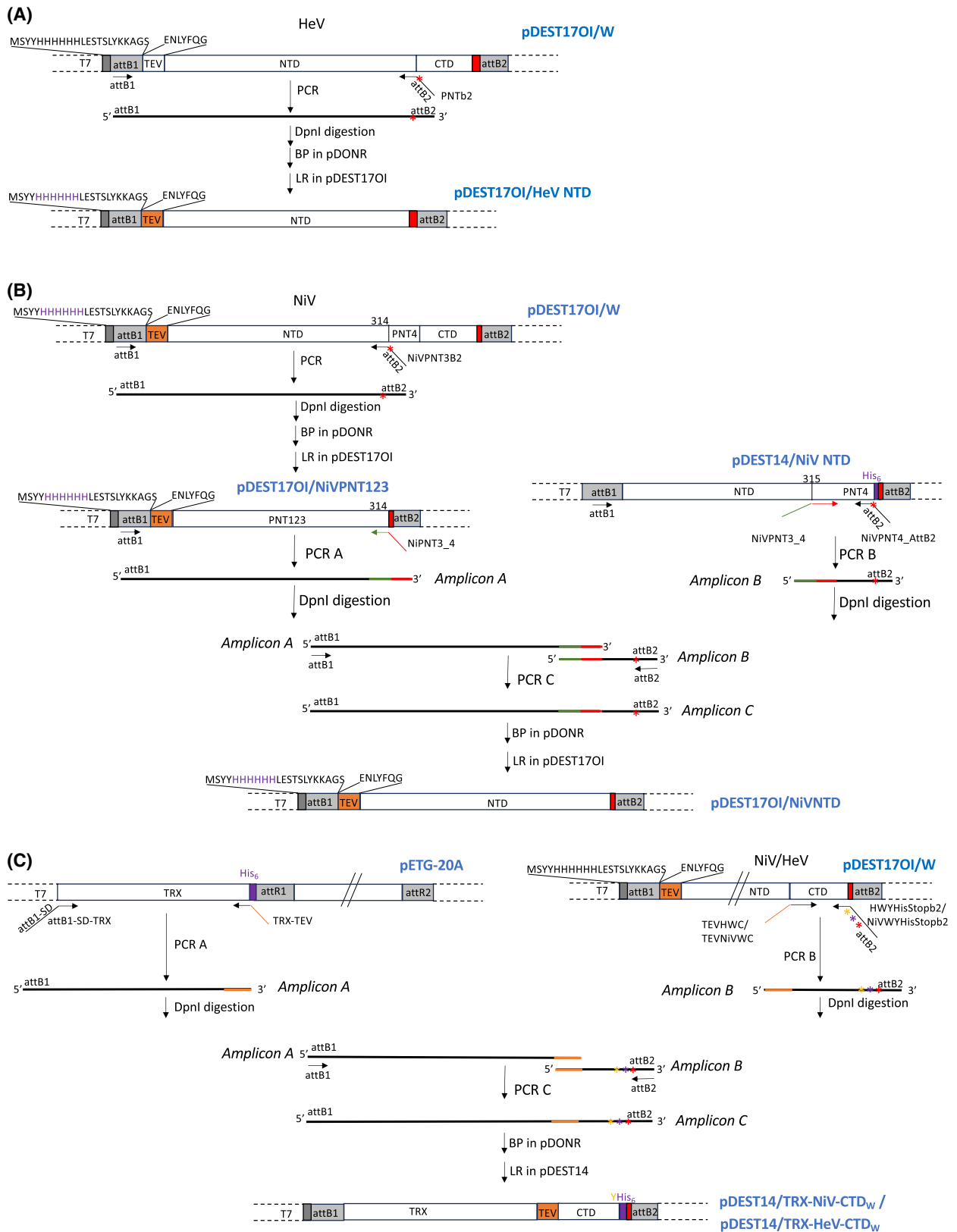


Fig. 14. Schematic representation of the strategy used to generate the various constructs. Procedure used to generate the Hendra virus (HeV) (A) and Nipah virus (NiV) (B) N-terminal domain (NTD) constructs and the Thioredoxin (TRX)-C-terminal domain (CTD) constructs from both viruses (C). Overlapping regions between amplicons A and B are shown in color. The red asterisk corresponds to two TAA stop codons, while the yellow and the purple asterisks correspond to a tyrosine-encoding codon and to a hexahistidine tag-encoding region, respectively. The vertical purple and red bars correspond to the hexahistidine tag and to the two stop codons, respectively.

Azenta Life Sciences, Leipzig, Germany) and found to conform to expectations.

Expression and purification of the recombinant proteins

The expression of all recombinant proteins was carried out in the *E. coli* strain T7pRos. All cultures were grown overnight to saturation in LB medium containing 100 $\mu\text{g}\cdot\text{mL}^{-1}$ ampicillin (Amp) and 34 $\mu\text{g}\cdot\text{mL}^{-1}$ chloramphenicol (Cam). An aliquot of the overnight culture was diluted 1/10 into 2 L of TB medium supplemented with Amp and Cam and grown at 37 °C. When the optical density at 600 nm (OD₆₀₀) reached 0.6–0.8, isopropyl β -D-thiogalactopyranoside was added to a final concentration of 0.5 mM, and the cultures were grown 5 h at 37 °C. The induced cells were harvested, washed, and collected by centrifugation (4000 \times g, 10 min).

Bacterial pellets were resuspended in 50 mL of buffer A1 (50 mM Hepes, 300 mM NaCl, 20 mM Imidazole, pH 7.2), supplemented with 6 M urea in the case of bacteria expressing the NTDs, and frozen at –20 °C. Upon thawing, the cells were disrupted by sonication with four cycles of 5 min each with intervals of 1.5 min at an amplitude 40%. The lysate was then centrifuged at 14 000 g for 40 min to discard all insoluble material. The supernatant was purified by IMAC using IMAC Sepharose 6 Fast Flow resin (Cytiva, Marlborough, MA, USA). The purification was performed in batch using 20 mL of resin for each protein. The resin was equilibrated in buffer A1. After a one-hour incubation, the resin was washed five times with one volume of buffer A1 and the elution was carried out in buffer A2 (50 mM Hepes, 150 mM NaCl, 300 mM Imidazole, pH 7.2) supplemented with 6 M urea in the case of the NTDs to prevent sample precipitation. The eluents from IMAC were analyzed by SDS/PAGE, and the fractions containing the protein of interest were pooled. In the case of the NTDs, the sample was desalted using a HiPrep 26/10 desalting column (Cytiva) in buffer HBS (20 mM Hepes, 150 mM NaCl, pH 7.2). DTT 5 mM, EDTA 1 mM, and TEV protease (at a final concentration of 0.01 $\text{mg}\cdot\text{mL}^{-1}$) were then added to either the desalted NTD samples or to the CTD_W samples obtained after IMAC (final volume of ~25 mL). TEV protease cleavage was performed overnight at room temperature with stirring at 60 r.p.m. At the end of the cleavage, urea was added to the digested samples to a final concentration of 6 M. IEC was then used to remove the cleaved

tag and the TEV protease. Prior to IEC, all samples were diluted in water so as to reduce the NaCl concentration to ~50 mM. Because all samples were diluted by approximately one third, the resulting final concentration of Hepes was ~17 mM. The CTDs were then loaded onto a HiPrep SP HP 16/10 (Cytiva), while the NTDs were loaded on a HiPrep DEAE FF 16/10 column (Cytiva), both previously equilibrated in buffer B1 (20 mM Hepes, 50 mM NaCl, 6 M Urea, pH 7.2). The proteins were eluted using a linear gradient of buffer B2 (20 mM Hepes, 1 M NaCl, 6 M Urea, pH 7.2). Purification of a tagged form of the NTDs was carried out as described omitting the TEV protease digestion step. For all proteins, the fractions from IEC containing the protein of interest were concentrated at ~10 $\text{mg}\cdot\text{mL}^{-1}$ and stored at –80 °C. Both IMAC and IEC were performed at room temperature (RT).

Prior to all ensuing analyses, and unless differently stated, urea was removed from the protein samples and the buffer exchanged to HBS (20 mM Hepes, 150 mM NaCl, pH 7.2) using a Sephadex G-25 medium column (Cytiva). Unless differently indicated, tagged forms of the NTDs were used.

The HeV and NiV W proteins were expressed and purified as tagged forms as already described [21]. We did not attempt at removing the vector-encoded tag as previous attempts at cleaving it off failed.

Protein concentrations were calculated using the theoretical absorption coefficients at 280 nm as obtained using the program ProtParam at the EXPASY server (<http://web.expasy.org/protparam/>).

Mass spectrometry—intact protein mass analysis

Protein masses were determined on purified solution samples. One microliter of protein at ~5 $\text{mg}\cdot\text{mL}^{-1}$ was mixed with 4 μL of sinapinic acid matrix solution at 20 $\text{mg}\cdot\text{mL}^{-1}$ in 0.3% trifluoroacetic acid/acetonitrile (50 : 50 v/v), following the procedure described in Ref. [21], using a maximum accelerating potential of 25 kV and a 4000–20 000 m/z range (LP_12kDa method).

Mass spectrometry—peptide mass fingerprinting

As previously described in Ref. [23], protein bands were excised from SDS/PAGE and digested overnight at 37 °C with high-sequencing-grade trypsin (Promega, Madison, WI, USA) after cysteine reduction in the presence of 10 mM

Table 2. Sequence of the forward (F) and reverse (R) primers used to generate the various constructs. The name of the template used in the PCR and the name of the final construct are indicated. For the cloning strategies that used the overlap PCR extension technique, the name of the intermediate amplicon is indicated in brackets. Constructs are available upon request.

Primer name	Sequence (5'–3')	Template name	Name of the final construct (name of the intermediate amplicon)
attB1 (F)	ACAAGTTTGTACAAAAAAGCAGGCT	pDEST170/HeV W	pDEST170/HeV NTD
PNTb2 (R)	ACCACTTTGTACAAGAAAGCTGGGT		
attB1 (F)	ACAAGTTTGTACAAAAAAGCAGGCT	pDEST170/NIV W	pDEST170/NIV PNT123
NIVPNT3B2 (R)	ACCACTTTGTACAAGAAAGCTGGGT		
attB1 (F)	ACAAGTTTGTACAAAAAAGCAGGCT	pDEST17/NIV	pDEST170/NIV NTD
NIPNT3_4 (R)	CCGGGTGCCGGTCCAAAAGGACTCAGCCGTTAAAAGAGGAAC	PNT123	(Amplicon A)
NIVPNT3-4 (F)	GGAAGAGTTTGAATGTGCCGTTAAAGAGGAACCCGCCGAC	pDEST14/NIV NTD	pDEST170/NIV NTD
NIVPNT4_AttB2 (R)	ACCACTTTGTACAAGAAAGCTGGGTCTTATTTATTTTGGATGCCAGG CTTTTCCGGCATCCG		(Amplicon B)
attB1 (F)	ACAAGTTTGTACAAAAAAGCAGGCT	Amplicon	pDEST170/NIV NTD
attB2 (R)	ACCACTTTGTACAAGAAAGCTGGGT	A + Amplicon B	
attb1-SD-TRX (F)	GTACAAAAAGCAGGCTAAATAATTTGTTTAACTTTAAGAAGGAGATATACAT ATGAGCGATAAAAATATTCAC	pETG-20A	pDEST14/TRX-HeV CTD _W & pDEST14/TRX-NIV CTD _W
TRX-TEV (R)	CTGGAAGTACAGGTTTTCGCCAGAAACCAGAAACCAGCCGCGCAG		(Amplicon A)
TEVHWC (F)	GAAAACTGTACTTCCAGGGCCGCCAGACCCCGCAGCCTG		
HWYHisStoppb2 (R)	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAAATGATGATGATGATGATG GTAGTTGCTCATGCGGCCAGCAG	pDEST170/HeV W	pDEST14/TRX-HeV CTD _W (Amplicon B)
TEVNiWC (F)	GAAAACTGTACTTCCAGGGCCGCCAGACCCGAAATATC		
NiWYHisStoppb2 (R)	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAAATGATGATGATGATGATGATG GTAGTTGGACATTCGCCCATTTGAC	pDEST170/NIV W	pDEST14/TRX-NIV CTD _W (Amplicon B)

DTT and alkylation with 50 mM iodoacetamide. Peptide extracts were pooled and dried in a centrifugal vacuum system. Samples were reconstituted with 0.1% trifluoroacetic acid in 4% acetonitrile and analyzed by LC-MS/MS using an Orbitrap QExactive Plus Mass Spectrometer (Thermo Electron, Bremen, Germany) online with an Ultimate 3000RSLCnano chromatography system (Thermo Fisher Scientific, Sunnyvale, CA, USA). Peptides were separated on a Dionex Acclaim PepMap RSLC C18 column. First, peptides were concentrated and purified on a pre-column in solvent A (0.1% formic acid (FA) in 2% acetonitrile). In the second step, peptides were separated on a reverse-phase LC EASY-Spray C18 column from Dionex (PepMap RSLC C18, 50 cm × 75 μm I.D., 100 Å of pore size, 2 μm of particle size) at 300 nL·min⁻¹ flow rate. After column equilibration using 2.5% of solvent B (20% water–80% acetonitrile–0.1% FA), peptides were eluted by a two-step linear gradient (2.5–25% acetonitrile/H₂O; 0.1% FA for 40 min and 25–50% acetonitrile/H₂O; 0.1% FA for 10 min). For peptide ionization in the EASY-Spray nanosource, spray voltage was set at 1.9 kV and the capillary temperature at 250 °C. All samples were measured in a data-dependent acquisition mode. Each run was preceded by a blank MS run in order to monitor system background. The peptide masses were measured in a survey full scan (scan range 375–1500 *m/z*, with 70 K FWHM resolution at *m/z* = 400, target AGC value of 3.00 × 10⁶ and maximum injection time of 100 ms). Following the high-resolution full scan in the Orbitrap, the more abundant precursor ions were selected, fragmented in HCD cell, and measured in Orbitrap (normalized collision energy of 25%, activation time of 10 ms, target AGC value of 1.00 × 10³, intensity threshold 1.00 × 10⁴ maximum injection time 100 ms, isolation window 2 *m/z*, 17.5 K FWHM resolution, scan range 200–2000 *m/z*). Charge state screening was enabled to exclude precursors with 0 and 1 charge states. Dynamic exclusion was implemented with a repeat count of 1 and exclusion duration of 10 s.

The acquired raw data were processed using PROTEOME DISCOVERER (version 1.4.1.14; Thermo Fisher Scientific). Spectra were searched using SEQUEST (Thermo Fisher Scientific) against a homemade database comprising 20 150 human and 4306 *E. coli* sequences implemented with the expected HeV and NiV NTD and CTD_W sequences (Swiss-Prot—human—reviewed—170315_20150_UP_coli_171120_4306_Patrick230801_ID_Bandes.fasta). Search parameters were (a) trypsin; two miscleavages allowed; (b) mass tolerance of 6 p.p.m. for monoisotopic precursor ions and 0.8 p.p.m. for fragment ions from MS/MS; (c) Cys carbamidomethylation (+57.02146 Da) as a fixed modification and Met oxidation (+15.99491 Da) as variable modification; and (d) minimum peptide length of four residues. Only high-score peptides were selected. Proteins were identified with a false discovery rate (FDR) of 1%.

Analytical SEC

The W proteins and their constituent NTDs and CTDs were analyzed by analytical size exclusion chromatography (SEC) using either a Superdex 75 Increase 10/300 GL (GE Healthcare, Uppsala, Sweden) (CTD_{WS}) or a Superdex 200 increase 10/300 GL (GE Healthcare) (W and NTD) and HBS as elution buffer.

For experiments using the CTD_{WS}, 100 μL of purified protein at ~6 mg·mL⁻¹ was injected. For analysis of the reduced samples, samples stored at -80 °C were thawed and supplemented with 5 mM DTT; after ~15 min of incubation at RT, the sample was loaded onto a G-25 column (Cytiva), eluted in HBS, and injected immediately. For the non-reduced sample, the eluent from the G-25 was collected and allowed to oxidize for ~24 h at RT before being injected.

In the case of the NTD and W proteins, 100 μL at ~1.5 mg·mL⁻¹ was injected. Before being loaded, all the samples were desalted using a G-25 column (Cytiva), eluted in HBS, and incubated 1 h at RT before analysis. The hydrodynamic radii (Stokes radii, *R_S*) of the W and NTD proteins eluted from the SEC column were extrapolated from a calibration curve obtained using globular proteins of known *R_S* (Ferritin: 61 Å, Aldolase: 48.1 Å, Conalbumin: 36.4 Å, Carbonic anhydrase: 23 Å, RNase A: 16.4 Å, and Aprotinin: 13.5 Å).

The *R_S* (in Å) of a natively folded (*R_S*^{NF}) and fully unfolded state in urea (*R_S*^U) for a protein with known molecular mass (MM) (in Daltons) were calculated according to Ref. [35]:

$$\log (R_S^{\text{NF}}) = 0.357 \times (\log \text{MM}) - 0.204, \quad (1)$$

$$\log (R_S^{\text{U}}) = 0.521 \times (\log \text{MM}) - 0.649, \quad (2)$$

The compaction index (CI) was calculated according to Ref. [36]:

$$\text{CI} = (R_S^{\text{U}} - R_S^{\text{Obs}}) / (R_S^{\text{U}} - R_S^{\text{NF}}), \quad (3)$$

This parameter, which allows comparison between proteins of different lengths, varies between 0 and 1, with 0 meaning minimal compaction and 1 maximal compaction.

Far-UV CD

Far-UV CD spectra were measured using a Jasco J-810 dichrograph (Hachioji, Japan), flushed with N₂ and at room temperature. One-mm thick quartz cuvettes were used. Protein concentrations were ~0.1 mg·mL⁻¹. Spectra were measured between 260 and 190 nm with a scanning speed of 100 nm·min⁻¹ and a data pitch of 0.5 nm. Response time was set to 4 s and the bandwidth to 2 nm. CD spectra were recorded in 10 mM sodium phosphate pH

7.2 at 25 °C. Each spectrum is the average of 10 acquisitions.

For analysis of the reduced CTD_W samples, a sample conserved at −80 °C was thawed and supplemented with 5 mM DTT. After 30 min of incubation at RT, the reduced sample was desalted using a G-25 column (Cytiva) and immediately analyzed by CD. The non-reduced sample was directly desalted from the −80 °C protein stock solution without adding DTT and directly analyzed by CD.

Far-UV CD spectra were also recorded for the NiV and HeV W proteins, on their NTDs, as well as on equimolar NTD + CTD_W mixtures. For the latter, a sample containing the NTD (untagged form) and the CTD_W in equimolar amounts (20 μM : 20 μM) was prepared and desalted using a G-25 column (Cytiva). The mixture was then diluted to obtain a final concentration of ~0.1 mg·mL^{−1}.

Far-UV CD spectra of CTD_W samples were also recorded in the presence of increasing concentrations of TFE (from 10 to 40% v/v). To this end, CTD_W samples were buffer exchanged in sodium phosphate 10 mM at pH 7.2 at a final concentration of ~1 mg mL^{−1}. The sample was then diluted in either TFE-free buffer or buffer supplemented with TFE to obtain the desired TFE concentration (0%, 10%, 20%, 30%, or 40% v/v) and a final protein concentration of ~0.1 mg mL^{−1}.

In all cases, the spectrum of the buffer was subtracted from the protein spectrum. Spectra were smoothed using PRISM.

Mean molar ellipticity values per residue (MRE) of single proteins were calculated as:

$$[\theta] = 3300m\Delta A/lcn, \quad (4)$$

where m is the mass of the protein in Daltons, l the path length (0.1 cm), c the concentration of the sample in mg·mL^{−1}, and n the number of residues. Molecular mass values and number of residues of the HeV and NiV NTD, CTD_W and of the W proteins are indicated in Table 3.

DICHHROWEB (<http://dichroweb.cryst.bbk.ac.uk/html/home.shtml>) [37] was used to analyze the experimental data in the 190–260 nm range using unsmoothed, subtracted

Table 3. Molecular mass and number of residues of the proteins used in circular dichroism (CD) studies. Note that in the case of the W proteins, we used the molecular mass and n values corresponding to a form in which the initial methionine has been cleaved off, as experimentally shown in Ref. [21].

Protein	MW (Daltons)	Number of residues
NTD HeV	44 319	404
NTD NiV	44 434	406
CTD _W HeV	5962	51
CTD _W NiV	6018	51
W HeV	52 705	476
W NiV	52 875	478

spectra. The content in the various types of secondary structure was estimated using the CDSSTR deconvolution algorithm with the reference protein set 7.

Mean molar ellipticity values per residue for mixtures containing two proteins, that is, NTD (protein 1) and CTD (protein 2), were calculated as:

$$[\theta] = 3300\Delta A/\{[(C_1 n_1)/m_1] + (C_2 n_2/m_2)\} l, \quad (5)$$

where n_1 and n_2 are the number of residues of protein 1 and protein 2, m_1 and m_2 are the molecular masses in Daltons, and C_1 and C_2 are the protein concentrations, expressed in mg·mL^{−1}, of each of the two proteins in the mixture.

The theoretical mean molar ellipticity values *per* residue of mixtures, expected if neither any disorder-to-order transition nor any secondary structure rearrangement occur, were calculated as:

$$[\theta]_{Ave} = [([\theta]_1 n_1) + ([\theta]_2 n_2 R)]/(n_1 + n_2 R), \quad (6)$$

where $[\theta]_1$ and $[\theta]_2$ correspond to the measured mean ellipticity values per residue of the two proteins in the mixture, n_1 and n_2 are the number of residues of each of the two proteins, and R is the excess molar ratio of protein 2.

Small-angle X-ray scattering (SAXS)

SEC-SAXS data of both CTD_{WS} were collected at SOLEIL (Gif-sur-Yvette, France) on the SWING beamline. Data collection was carried out as described in Ref. [21]. The calibration was made with water. Samples were at 9 mg·mL^{−1} in HBS supplemented with 1 M urea. Prior to analysis, 5 mM of DTT was added and samples were incubated 30 min at room temperature. Subsequently, for both samples, 70 μL was injected onto a BioSec 3–100 SEC column (Agilent, Santa Clara, CA, USA). The elution buffer was HBS. The flow rate was 0.2 mL·min^{−1}, and the temperature was 20 °C. The exposure was in continuous mode, with 1 frame·s^{−1} (990 ms exposure and 10 ms dead time). Data reduction and frames subtraction were done manually with the beamline software FOXTROT. Buffer blank frames were taken in the dead volume of the column. Deconvoluted data corresponding to the CTD_W were compared with the CORMAP test and averaged.

The final scattering intensities were analyzed using the ATSAS program package [38] and the BIOXTAS RAW program. Linearity in the Guinier region showed no sign of sample aggregation. The radius of gyration (R_g) was estimated at low angles ($q < 1.12/R_g$ for HeV and $q < 1.15/R_g$ for NiV) according to the Guinier approximation [39,40]:

$$\ln I(q) = \ln I_0 - (q^2 R_g^2)/3. \quad (7)$$

The pair-distance distribution function, $P(r)$, from which the D_{max} and the R_g were estimated, was then computed

using GNOM and manually adjusted until a good CorMap *P*-value ($\alpha > 0.01$) was obtained [41].

The theoretical R_g value (in Å) expected for various conformational states was calculated using Flory's equation:

$$R_g = R_0 N^\nu, \quad (8)$$

where N is the number of amino acid residues, R_0 a constant, and ν a scaling factor. For IDPs, $R_0 = 2.54 \pm 0.01$ and $\nu = 0.522 \pm 0.01$ [42], for chemically denatured (U) proteins $R_0 = 2.23 \pm 0.58$ and $\nu = 0.598 \pm 0.028$ [42], and for natively folded (NF) proteins $R_0 = \sqrt{(3/5)} \times 4.75$ and $\nu = 0.29$ [43]. The R_g -based CI can be calculated as follows [36]:

$$CI = (R_g^U - R_g^{obs}) / (R_g^U - R_g^{NF}), \quad (9)$$

where R_g^{obs} is the experimentally observed value and R_g^U and R_g^{NF} are the reference values calculated for a fully unfolded (U) and natively folded (NF) form. As in the case of the R_S -based CI, this index increases with increasing compaction (where 1 corresponds to the most compact form). The flexibility of the proteins was assessed with the dimensionless Kratky plot ($(qR_g)^2 I(q)/I_0$ vs qR_g).

Models of the two CTD_{WS} as conformational ensembles were obtained with Ensemble Optimization Method (EOM) 2.1 [44] using the default parameters and random coil conformers. For both the CTD_{WS}, the amino acid sequence provided as input to EOM was that of the recombinant CTD_W proteins as resulting from TEV protease cleavage and including the YHHHHHH C-terminal stretch. Using EOM 2.1, systematic quantification of the flexibility was made using the R_{flex} [44]. Experimental error-independent goodness-of-fit was also confirmed using the software CORMAP [45].

SEC-SAXS data have been deposited in the Small Angle Scattering Biological Data Bank (SASBDB) [46] under codes SASDUQ3 and SASDUR3 for the set of data of HeV CTD_W and of NiV CTD_W, respectively. The SEC-SAXS-derived HeV and NiV CTD_W ensembles have been deposited within the Protein Ensemble Database (PED-DB, <https://proteinensemble.org/>) [47] under accession numbers PED00507 and PED00508, respectively.

Pull-down assay

One milliliter of IMAC Sepharose 6 Fast Flow resin (25% slurry) (Cytiva), previously equilibrated five volumes of either buffer C1 (20 mM Hepes, 20 mM imidazole, 50 mM NaCl pH 7.2) or C2 (buffer C1 containing 200 mM NaCl) or C3 (buffer C1 containing 500 mM NaCl) or C4 (buffer C1 containing 1 M NaCl), was incubated with 1 mL of CTD_W at 0.5 mg·mL⁻¹ in the corresponding and appropriate buffer for 1 h at room temperature with gentle shaking (20 r.p.m.).

To remove the excess of CTD_W, the resin was washed three times with 1 mL using the buffer used for equilibration

and binding. One milliliter of untagged NTD at 1 mg·mL⁻¹ was subsequently added to the mixture and incubated at 20 r.p.m. for 1 h at room temperature. After incubation, the supernatant (Flow-Through) was removed, and the resin was washed four times with 1 mL of the appropriate buffer. To evaluate possible non-specific NTD binding to the resin, two controls were done using NTD in buffer C1 or C4 without CTD_W. Ten microliter of the unretained fraction (Flow-Through) and of the washed resin (beads) were analyzed by 15% SDS/PAGE.

Th-T binding assays

ThT (Sigma-Aldrich, Saint Louis, MO, USA) was used to monitor fibril formation. A stock solution of ThT at 1 mM in HBS was used. The experiments were performed in HBS in the presence of either the isolated CTD_W, NTD, and W proteins at 20 μM or in the presence of both NTD and CTD_W in equimolar amounts (i.e., 20 μM–20 μM). For the analysis at T0 and T72, all samples were prepared as described above. The proteins were incubated in the presence of 50 μM ThT for 5 min (equilibration time) before following the aggregation process by fluorimetry. Fluorescence measurements were carried out on a Tecan microplate reader GENios Plus using a black 96-well plate with transparent flat bottom (655096; Greiner Bio-One, Les Ulis, France). ThT was excited at 440 nm (slitwidth: 5 nm), and fluorescence emission was recorded at 535 nm (slitwidth: 10 nm). As a control, the fluorescence of a sample containing solely ThT was monitored. All statistical analyses were performed with Prism One-way ANOVA test.

Negative-staining TEM

Negative-staining TEM studies were performed on protein samples in HBS containing either the isolated CTD_W, NTD, and W proteins at 20 μM or both NTD and CTD_W in equimolar amounts (i.e., 20 μM–20 μM). All the samples were diluted prior to analysis in HBS to a final concentration of 0.05 mg·mL⁻¹. Grid preparation and data collection were carried out as described in Ref. [21]. For the statistical analysis of the fibril contour length, IMAGEJ (<https://imagej.net/Welcome>) was used. Fibril length measures were taken for 100 fibrils for each condition.

Acknowledgements

We thank Aurélien Thureau (SOLEIL) for their help in recording SEC-SAXS data. We thank both the ESRF and the SOLEIL synchrotrons for beamtime allocation. We are also grateful to Gerlind Sulzenbacher (AFMB lab) for efficiently managing the AFMB BAG. We thank all the AFMB technical and support staff (Denis Patrat, Patricia Clamecy, Béatrice Rolland, Chantal

Falaschi and Fabienne Amalfitano). We thank the mass spectrometry facility of Marseille Proteomics (marseille-proteomique.univ-amu.fr), supported by IBISA (Infrastructures Biologie Santé et Agronomie), Plateforme Technologique Aix-Marseille, the Cancéropôle PACA, Région Sud-Alpes-Côte d'Azur, the Institut Paoli-Calmettes, the Centre de Recherche en Cancérologie de Marseille (CRCM), and Fonds Européen de Développement Régional and Plan Cancer. This work was carried out with the financial support of the Agence Nationale de la Recherche (ANR), specific project Heni-phase (ANR-21-CE11-0012-01). It was also partly supported by the French Infrastructure for Integrated Structural Biology (FRISBI) (ANR-10-INSB-0005) and by the Centre National de la Recherche Scientifique. FG is supported by a post-doctoral fellowship from the FRM (Fondation pour la Recherche Médicale). GP was supported by a joint doctoral fellowship from the AID (Agence Innovation Défense) and Aix-Marseille University. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. A CC-BY public copyright license has been applied by the authors to the present document and will be applied to all subsequent versions up to the author accepted manuscript arising from this submission, in accordance with the grant's open access conditions.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

SL conceived, designed, and supervised the study and acquired funding. CB generated all the bacterial expression constructs except for the one encoding NiV NTD that was generated by GP. GP carried out all the experiments except for MS analyses that were performed by PF. FG and DP provided assistance with SEC-SAXS and TEM studies, respectively. GP and SL analyzed and interpreted the data. GP generated the first draft of the manuscript and prepared all the figures except for Fig. 14 that was generated by SL. All the authors wrote and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Peer review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/febs.17239>.

Data availability statement

SEC-SAXS data have been deposited in the Small Angle Scattering Biological Data Bank (SASBDB) [46] under codes SASDUQ3 and SASDUR3 for the set of data of HeV CTD_W and of NiV CTD_W, respectively. The SEC-SAXS-derived HeV and NiV CTD_W ensembles have been deposited within the Protein Ensemble Database (PED-DB, <https://proteinsenble.org/>) [47] under accession numbers PED00507 and PED00508, respectively. Plasmids are available on request.

References

- Eaton BT, Mackenzie JS & Wang LF (2007) Henipaviruses. In *Fields Virology* (Fields BN, Knipe DM & Howley PM, eds), pp. 1587–1600. Lippincott-Raven, Philadelphia, PA.
- Habchi J, Mamelli L, Darbon H & Longhi S (2010) Structural disorder within Henipavirus nucleoprotein and phosphoprotein: from predictions to experimental assessment. *PLoS One* **5**, e11684.
- Jensen MR, Yabukarski F, Communie G, Condamine E, Mas C, Volchkova V, Tarbouriech N, Bourhis JM, Volchkov V, Blackledge M *et al.* (2020) Structural description of the Nipah virus phosphoprotein and its interaction with STAT1. *Biophys J* **118**, 2470–2488.
- Schiavina M, Salladini E, Murrall MG, Tria G, Felli IC, Pierattelli R & Longhi S (2020) Ensemble description of the intrinsically disordered N-terminal domain of the Nipah virus P/V protein from combined NMR and SAXS. *Sci Rep* **10**, 19574.
- Habchi J, Tompa P, Longhi S & Uversky VN (2014) Introducing protein intrinsic disorder. *Chem Rev* **114**, 6561–6588.
- Gondelaud F, Pesce G, Nilsson JF, Bignon C, Pchelkine D, Gerlier D, Mathieu C & Longhi S (2022) Functional benefit of structural disorder for the replication of measles, Nipah and Hendra viruses. *Essays Biochem* **66**, 915–934.
- Fontana JM, Bankamp B & Rota PA (2008) Inhibition of interferon induction and signaling by paramyxoviruses. *Immunol Rev* **225**, 46–67.
- Audsley MD & Moseley GW (2013) Paramyxovirus evasion of innate immunity: diverse strategies for common targets. *World J Virol* **2**, 57–70.
- Tsimbalyuk S, Cross EM, Hoad M, Donnelly CM, Roby JA & Forwood JK (2020) The intrinsically disordered W protein is multifunctional during Henipavirus infection, disrupting host signalling pathways and nuclear import. *Cells* **9**, 1913.
- Lo MK, Miller D, Aljofan M, Mungall BA, Rollin PE, Bellini WJ & Rota PA (2010) Characterization of the antiviral and inflammatory responses against Nipah

- virus in endothelial cells and neurons. *Virology* **404**, 78–88.
- 11 Rodriguez JJ, Parisien JP & Horvath CM (2002) Nipah virus V protein evades alpha and gamma interferons by preventing STAT1 and STAT2 activation and nuclear accumulation. *J Virol* **76**, 11476–11483.
 - 12 Shaw ML, Garcia-Sastre A, Palese P & Basler CF (2004) Nipah virus V and W proteins have a common STAT1-binding domain yet inhibit STAT1 activation from the cytoplasmic and nuclear compartments, respectively. *J Virol* **78**, 5633–5641.
 - 13 Hagmaier K, Stock N, Goodbourn S, Wang LF & Randall R (2006) A single amino acid substitution in the V protein of Nipah virus alters its ability to block interferon signalling in cells from different species. *J Gen Virol* **87**, 3649–3653.
 - 14 Keiffer TR, Ciancanelli MJ, Edwards MR & Basler CF (2020) Interactions of the Nipah virus P, V, and W proteins across the STAT family of transcription factors. *mSphere* **5**, e00449-20.
 - 15 Satterfield BA, Cross RW, Fenton KA, Agans KN, Basler CF, Geisbert TW & Mire CE (2015) The immunomodulating V and W proteins of Nipah virus determine disease course. *Nat Commun* **6**, 7483.
 - 16 Shaw ML, Cardenas WB, Zamarin D, Palese P & Basler CF (2005) Nuclear localization of the Nipah virus W protein allows for inhibition of both virus- and toll-like receptor 3-triggered signaling pathways. *J Virol* **79**, 6078–6088.
 - 17 Enchery F (2017) Etude de la modulation de la voie canonique d'activation de NF-κB par les protéines non structurales du virus Nipah. Thèse de Doctorat de l'Université de Lyon.
 - 18 Edwards MR, Hoad M, Tsimbalyuk S, Menicucci AR, Messaoudi I, Forwood JK & Basler CF (2020) Henipavirus W proteins interact with 14-3-3 to modulate host gene expression. *J Virol* **94**, e00373-20.
 - 19 Enchery F, Dumont C, Iampietro M, Pelissier R, Aurine N, Bloyet LM, Carbonelle C, Mathieu C, Journo C, Gerlier D *et al.* (2021) Nipah virus W protein harnesses nuclear 14-3-3 to inhibit NF-κB-induced proinflammatory response. *Commun Biol* **4**, 1292.
 - 20 Salladini E, Delauzun V & Longhi S (2017) The Henipavirus V protein is a prevalently unfolded protein with a zinc-finger domain involved in binding to DDB1. *Mol Biosyst* **13**, 2254–2267.
 - 21 Pesce G, Gondelaud F, Ptchelkine D, Nilsson JF, Bignon C, Cartalas J, Fourquet P & Longhi S (2022) Experimental evidence of intrinsic disorder and amyloid formation by the Henipavirus W proteins. *Int J Mol Sci* **23**, 923.
 - 22 Smith KM, Tsimbalyuk S, Edwards MR, Cross EM, Batra J, Soares da Costa TP, Aragao D, Basler CF & Forwood JK (2018) Structural basis for importin alpha 3 specificity of W proteins in Hendra and Nipah viruses. *Nat Commun* **9**, 3703.
 - 23 Nilsson JF, Baroudi H, Gondelaud F, Pesce G, Bignon C, Ptchelkine D, Chamieh J, Cottet H, Kajava AV & Longhi S (2022) Molecular determinants of fibrillation in a viral amyloidogenic domain from combined biochemical and biophysical studies. *Int J Mol Sci* **24**, 399.
 - 24 Salladini E, Gondelaud F, Nilsson J, Pesce G, Bignon C, Murrall MG, Horvat B, Fabre R, Pierattelli R, Kajava AV *et al.* (2021) Identification of a region in the common amino-terminal domain of Hendra virus P, V and W proteins responsible for phase transition and amyloid formation. *Biomolecules* **11**, 1324.
 - 25 Gondelaud F, Lalande A, Pesce G, Bignon C, Fourquet P, Ptchelkine D, Brouilly N, Lozach PY, Gerlier D, Mathieu C *et al.* (2024) Redox-dependent formation of a viral amyloid and functional impact. *bioRxiv*. doi: [10.1101/2024.01.22.576663](https://doi.org/10.1101/2024.01.22.576663) [PREPRINT].
 - 26 Lieutaud P, Canard B & Longhi S (2008) MeDor: a metaserver for predicting protein disorder. *BMC Genomics* **9**, S25.
 - 27 Stark GR, Stein WH & Morre S (1960) Reactions of the cyanate present in aqueous urea with amino acids and proteins. *J Biol Chem* **235**, 3177–3181.
 - 28 Chemes LB, Alonso LG, Noval MG & de Prat-Gay G (2012) Circular dichroism techniques for the analysis of intrinsically disordered proteins and domains. *Methods Mol Biol* **895**, 387–404.
 - 29 Uversky VN (2002) Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* **11**, 739–756.
 - 30 Scholtz JM, Qian H, York EJ, Stewart JM & Baldwin RL (1991) Parameters of helix-coil transition theory for alanine-based peptides of varying chain lengths in water. *Biopolymers* **31**, 1463–1470.
 - 31 LeVine H 3rd (1993) Thioflavine T interaction with synthetic Alzheimer's disease beta-amyloid peptides: detection of amyloid aggregation in solution. *Protein Sci* **2**, 404–410.
 - 32 Yakupova EI, Bobyleva LG, Vikhlyantsev IM & Bobylev AG (2019) Congo red and amyloids: history and relationship. *Biosci Rep* **39**, BSR20181415.
 - 33 Khurana R, Uversky VN, Nielsen L & Fink AL (2001) Is Congo red an amyloid-specific dye? *J Biol Chem* **276**, 22715–22721.
 - 34 Pesce G, Brocca S, Grandori R, Longhi S & Uversky AV (2023) Droplets of life: role of phase separation in virus replication and compartmentalization. In *Droplets of Life* (Uversky AV, ed.), pp. 567–615. Elsevier, Amsterdam.
 - 35 Uversky VN (2002) What does it mean to be natively unfolded? *Eur J Biochem* **269**, 2–12.
 - 36 Brocca S, Testa L, Sobott F, Samalikova M, Natalello A, Papaleo E, Lotti M, De Gioia L, Doglia SM,

- Alberghina L *et al.* (2011) Compaction properties of an intrinsically disordered protein: sic1 and its kinase-inhibitor domain. *Biophys J* **100**, 2243–2252.
- 37 Whitmore L & Wallace BA (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res* **32**, W668–W673.
- 38 Manalastas-Cantos K, Konarev PV, Hajizadeh NR, Kikhney AG, Petoukhov MV, Molodenskiy DS, Panjkovich A, Mertens HDT, Gruzinov A, Borges C *et al.* (2021) ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *J Appl Cryst* **54**, 343–355.
- 39 Guinier A (1939) La diffraction des rayons X aux tres petits angles; application a l'etude de phenomenes ultramicroscopiques. *Ann Phys (Paris)* **12**, 161–237.
- 40 Guinier A & Fournet F (1955) Small Angle Scattering of X-Rays. Wiley Interscience, New York, NY.
- 41 Svergun D (1992) Determination of the regularization parameters in indirect-transform methods using perceptual criteria. *J Appl Cryst* **25**, 495–503.
- 42 Bernado P & Blackledge M (2009) A self-consistent description of the conformational behavior of chemically denatured proteins from NMR and small angle scattering. *Biophys J* **97**, 2839–2845.
- 43 Wilkins DK, Grimshaw SB, Receveur V, Dobson CM, Jones JA & Smith LJ (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* **38**, 16424–16431.
- 44 Tria G, Mertens HDT, Kachala M & Svergun D (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 202–217.
- 45 Franke D, Jeffries CM & Svergun D (2015) Correlation map, a goodness-of-fit test for one-dimensional X-ray scattering spectra. *Nat Methods* **12**, 419–422.
- 46 Valentini E, Kikhney AG, Previtali G, Jeffries CM & Svergun DI (2015) SASBDB, a repository for biological small-angle scattering data. *Nucleic Acids Res* **43**, D357–D363.
- 47 Lazar T, Martínez-Pérez E, Quaglia F, Hatos A, Chemes LB, Iserte JA, Méndez NA, Garrone NA, Saldaño TE, Marchetti J *et al.* (2021) PED in 2021: a major update of the protein ensemble database for intrinsically disordered proteins. *Nucleic Acids Res* **49**, D404–D411.