Structure and dynamics of the unassembled nucleoprotein of rabies virus in complex with its phosphoprotein chaperone module

Francine Gérard, Jean-Marie Bourhis, Caroline Mas, Anaïs Branchard, Duc Duy Vu, Sylvia Varhoshkova, Cedric Leyrat, Marc Jamin

viruses

MDPI

Article

Structure and Dynamics of the Unassembled Nucleoprotein of Rabies Virus in Complex with Its Phosphoprotein Chaperone Module

Francine C. A. Gérard [1,†,‡], Jean-Marie Bourhis [1,†], Caroline Mas [2], Anaïs Branchard [1], Duc Duy Vu [1], Sylvia Varhoshkova [1], Cédric Leyrat [3,*] and Marc Jamin [1,*]

[1] Institut de Biologie Structurale (IBS), Université Grenoble Alpes, CEA, CNRS, 71 Avenue des Martyrs, 38000 Grenoble, France
[2] Integrated Structural Biology Grenoble (ISBG), Université Grenoble Alpes, CNRS, CEA, EMBL, 71 Avenue des Martyrs, 38000 Grenoble, France
[3] Institut de Génomique Fonctionnelle, Université de Montpellier, CNRS, INSERM, 34094 Montpellier, France
* Correspondence: cedric.leyrat@igf.cnrs.fr (C.L.); marc.jamin@ibs.fr (M.J.)
† These authors contributed equally to this work.
‡ Current address: Laboratory of Molecular Microbiology and Structural Biochemistry, CNRS UMR5086, Université de Lyon, 69367 Lyon, France.

Abstract: As for all non-segmented negative RNA viruses, rabies virus has its genome packaged in a linear assembly of nucleoprotein (N), named nucleocapsid. The formation of new nucleocapsids during virus replication in cells requires the production of soluble N protein in complex with its phosphoprotein (P) chaperone. In this study, we reconstituted a soluble heterodimeric complex between an armless N protein of rabies virus (RABV), lacking its N-terminal subdomain ($N_{NT\text{-}ARM}$), and a peptide encompassing the N^0 chaperon module of the P protein. We showed that the chaperone module undergoes a disordered−order transition when it assembles with N^0 and measured an affinity in the low nanomolar range using a competition assay. We solved the crystal structure of the complex at a resolution of 2.3 Å, unveiling the details of the conserved interfaces. MD simulations showed that both the chaperon module of P and RNA-mediated polymerization reduced the ability of the RNA binding cavity to open and close. Finally, by reconstituting a complex with full-length P protein, we demonstrated that each P dimer could independently chaperon two N^0 molecules.

Keywords: rabies virus; *Mononegavirales*; phosphoprotein; nucleocapsid assembly; X-ray crystallography; small-angle X-ray scattering; molecular dynamics simulation

Citation: Gérard, F.C.A.; Bourhis, J.-M.; Mas, C.; Branchard, A.; Vu, D.D.; Varhoshkova, S.; Leyrat, C.; Jamin, M. Structure and Dynamics of the Unassembled Nucleoprotein of Rabies Virus in Complex with Its Phosphoprotein Chaperone Module. *Viruses* **2022**, *14*, 2813. https://doi.org/10.3390/v14122813

Academic Editor: Akira Ono

Received: 1 November 2022
Accepted: 7 December 2022
Published: 16 December 2022

1. Introduction

Rabies is a zoonotic, incurable brain disease that, despite the availability of an efficient vaccine and post-exposure prophylaxis, continues to kill tens of thousands of people every year according to WHO [1]. Its etiological agent, the eponymous rabies virus (RABV), is a prototypic member of the *Rhabdoviridae*, a large family of nonsegmented negative-sense RNA viruses that includes vesicular stomatitis virus (VSV) [2], Chandipura virus [3], and several underexplored viruses infecting various animals and plants [4]. Because these viruses share a similar organization of their genome and of their virion as well as similar strategies for the transcription, replication, and encapsidation of their genome with other negative-sense RNA viruses, the family *Rhabdoviridae* is classified into the order *Mononegavirales* (*MNV*) with, inter alia, the families *Paramyxoviridae* (including measles virus, mumps virus, parainfluenza viruses, and the zoonotic Nipah and Hendra viruses), *Pneumoviridae* (including respiratory syncytial virus and metapneumovirus), and *Filoviridae* (including the zoonotic Ebola and Marburg viruses) [5,6]. The order *Mononegavirales* is now classified with all other known segmented and nonsegmented negative-sense RNA viruses in the phylum *Negarnaviricota* [6].

The RABV genome of 11.9 kb comprises five transcription units, which encode five structural proteins: the nucleoprotein (N), the phosphoprotein (P), the matrix protein (M), the surface glycoprotein (G), and the large RNA-dependent RNA polymerase (L-RdRP). The multiplication of the virus in appropriate host cells involves (i) the transcription of the genomic RNA to generate messenger RNAs, (ii) the replication of the genome through the intermediate production of positive-sense antigenomic RNAs, which are subsequently used as a template to generate new genomic RNAs, and (iii) the packaging of both genomic and antigenomic RNAs into linear homopolymers of N, forming helical and flexible ribonucleoprotein complexes, named nucleocapsids (NCs) [7,8]. A unique molecular machine involving three viral proteins (N-, P-, and L-RdRP) carries out these three processes within cytoplasmic membrane-less viral factories, called Negri's bodies [9]. In particular, the assembly of novel NCs is an essential step in the infectious cycle of these viruses because only the genomic and antigenomic RNAs packaged in NC can serve as template for RNA synthesis by the polymerase complex (composed of P- and L-RdRP) [10–12]. Indeed, if L-RdRP catalyzes RNA synthesis, as well as the processing of mRNAs, including the synthesis and methylation of the 5′ cap and the synthesis of the 3′ poly-A tail, it is unable to proceed efficiently on naked RNAs and requires the presence of both P and N for processive transcription and replication [10]. Thus, genome replication requires the production of the unassembled RNA-free N protein, named N^0, in the form a complex named the N^0−P complex [13,14], in which P prevents the polymerization of N and the illegitimate assembly with cellular RNA until N is transferred to nascent viral RNA [15].

RABV P is an essential multi-functional protein with a modular architecture [16], which acts both as a cofactor of the L-RdRP and as a chaperone of N^0, but also hijacks cellular components [17] and counteracts the host antiviral responses [18]. RABV P forms dimers [19]; each protomer consists of a long N-terminal intrinsically disordered region (P_{NTR}, aa 1–90) and a C-terminal region (P_{CTR}) consisting of two folded domains, the multimerization domain, P_{MD} (aa 91–131), and the C-terminal domain (aa 186–297), P_{CTD}, which are connected by a flexible linker (aa 132–185) [20–22]. The P protein works as a hub, in which structural and functional modules interact independently with different partners; the extremity of P_{NTR} consists of the chaperon module (P_{CM}—aa 1–40) that binds N^0 [23], while most of the remaining part of P_{NTR} binds L-RdRP (residues 40–88) [24], whereas P_{CTD} binds to the polymeric N−RNA complexes (residues 185–297) [21]. The N-terminal segment of P_{NTR} is sufficient for keeping N in a monomeric soluble form [23,25–28], but it is also believed that the N^0−P complex is recruited at the site of RNA synthesis by the attachment of P_{CTD} to the template [29]. Different modules or motifs of RABV P bind the dynein LC8 [17], the focal adhesion kinase [30], the ribosomal protein L9 [31], the mitochondrial complex I [32], IRF-3 [33], STAT1 [34], PML [33,35], nuclear import and export factors [36], BECN1 [37], and Cdc37/Hsp90 complex [38]. RABV N is composed of two globular domains that are connected by a hinge, creating an RNA binding groove [39]. It also comprises two subdomains, an N-terminal extension, named the $N_{NT\text{-}ARM}$, and an internal loop in the C-terminal domain, named the $N_{CT\text{-}ARM}$, that dock to neighboring subunits in the polymeric form and thereby stabilize the assembled NC [39].

Previously, working with VSV, we developed a strategy to reconstitute a core N^0−P complex between a nucleoprotein truncated of its N-terminal arm (VSV $N_{NT\text{-}ARM}$: aa 1–21), VSV $N_{\Delta 21}$, and a peptide of 60 residues (VSV P_{60}) derived from P that comprised the chaperone module (P_{CM}) and we solved the crystal structure of the complex [40]. The reconstituted VSV $N_{\Delta 21}{}^0$−P_{60} complex was monomeric in solution, but assembled into RNA-free 10-mers circular complexes in the crystal, resembling the recombinant circular NC-like complexes [41]. In the crystal structure, P_{CM} extended in a groove of the C-terminal domain and formed a long α-helix bound at the interface of the N- and C-terminal domains of N. The structure unveiled a mechanism for preventing the assembly of N by interfering with the binding of the $N_{NT\text{-}ARM}$ from the adjacent N_{i-1} protomer and with the binding of $N_{CT\text{-}ARM}$ from the N_{i+1} protomer [40]. The protrusion of the C-terminal part in the RNA

binding groove of the N protein suggested a direct interference with RNA binding [40]. We also showed that no additional interaction was detected between full-length P and N^0, that the phosphorylation of P_{NTR} by casein kinases (Ser60, Thr62, and Ser64) had no effect on the interaction with N^0 and that dimeric P could bind and chaperone one or two N^0 molecules depending of the concentration of N^0 in the assembly conditions [42]. The structure of the N^0−P complex has since then been obtained for several other RNA viruses in the families *Paramyxoviridae* [43–45], *Pneumoviridae* [46], and *Filoviridae* [26,28,47], showing that the interference of P_{CM} with the $N_{NT\text{-}ARM}$ and $N_{CT\text{-}ARM}$ binding is a conserved feature within the order *Mononegavirales*, although the structure and position of P_{CM} on the surface of N varied between the viruses. In VSV crystal structure, the N^0 protein was in a close conformation, unable in this state to accommodate the incoming RNA molecules, whereas, by contrast, in the N^0−P complex of these other viruses, the N^0 molecule was trapped in an open conformation, unable to strongly bind RNA, but ready to grasp an incoming RNA molecule.

The comparison of the structure of the VSV $N_{\Delta 21}{}^0$−P_{60} complex [40] with that of the VSV N−RNA complex [41] and with those of the N^0−P complexes of other viruses [26,28,43–47] raised several questions. First, despite the binding of P_{CM} and the truncation of the $N_{NT\text{-}ARM}$, the VSV $N_{\Delta 21}$ protein assembled in a polymeric form in the crystal, suggesting that P_{60} was not sufficient to maintain the N^0 protein in its monomeric state. Second, in the $N_{\Delta 21}{}^0$−P_{60} crystal structure, the N protein was in its closed conformation, which was identical to its RNA-bound conformation, whereas in the structure of the other N^0−P complexes, the N protein was in an open conformation [26,28,43–47]. This suggested the occurrence of a hinge motion in N allowing the insertion and release of the RNA molecule, but it raised the possibility that the assembly of VSV N into circular complexes had induced the closure of the protein. Third, in the VSV $N_{\Delta 21}{}^0$−P_{60} complex, the N-terminal part of P_{CM} was bound in same groove of N_{CTD} as the $N_{NT\text{-}ARM}$ and in the same orientation. In other known N^0−P structures, the chaperone module (P_{CM}) adopted an orientation opposite to that of the N-terminal arm of N ($N_{NT\text{-}ARM}$), against leaving a doubt that the assembly of N could influenced the interaction with P_{CM} [26,28,43–47].

To address these questions, we studied the structure and the dynamics of the RABV N^0−P complex. We used the same strategy for reconstituting the N^0−P core complex than for VSV and NiV [40,43] that involved the independent expression and purification of truncated N and of a P fragment, their mixing, and the purification of the reconstituted complex. We hypothesized that a slightly longer disordered region at the C-terminus of the P peptide might prevent the assembly of N^0 into rings by further masking the surface of N with the flexible part of the P peptide. We characterized the structure of the complex in a solution by SEC-MALLS and SEC-SAXS and solved the crystal structure. We implemented a fluorescence anisotropy assay to monitor the binding kinetics and equilibrium of the peptide. To better understand the hinge motion mechanism in the N protein and how the binding of P or the assembly into a polymeric N−RNA complex affected this conformational change, we performed molecular dynamics simulations with the different forms of N.

2. Materials and Methods

2.1. Bioinformatics

The amino acid sequences of Lyssavirus phosphoproteins were retrieved from the Uniprot database [48] and multiple sequence alignments (MSA) were performed with Clustal Omega using default parameters [49]. The calculation of the D-score, which provided a consensus prediction of 13 disorder algorithms available through webservers and defined the boundaries of structured domains, was performed as previously described [20]. The interface between $N_{D23}{}^0$ and P_{68} was analyzed with PDBePISA [50], structural alignment with PDBeFOLD [51], and computational alanine scanning mutagenesis was performed with FoldX [52].

2.2. Constructs

The plasmid (pET22b(+)) containing the gene of full-length RABV P (CVS-11 strain) (UniProt P22363) fused to a C-terminal two-amino linker and 6xHis-tag for expression in bacteria was previously described [20]. The cDNA encoding residues from 1 to 68 of RABV phosphoprotein were cloned into the pET28a (Novagen, Darmstadt, Germany) vector using NcoI and XhoI restriction sites in a frame with a downstream 6xHis-tag and a two amino-acid linker (Glu-Leu). A synthetic cDNA (Geneart, Regensburg, Germany) encoding the first 42 residues of RABV phosphoprotein with a cysteine substitution at position 41 (G41C), an N-terminal 6xHis tag, a SUMO tag and a tobacco etch virus protease (TEV) cleavage site were cloned into the pET22b expression vector (Novagen) using the NcoI and XhoI restriction sites. Point mutations were introduced in this construct by site-directed mutagenesis using the QuickChange II protocol (Agilent, Santa Clara, CA, USA). The cDNA encoding residues from 24 to 450 of rabies virus (strain CVS-11) nucleoprotein (UniProt Q8JXF6) were cloned into the pETM-40 (EMBL) vector using NcoI and XhoI restriction sites in frame with the upstream MalE gene encoding the maltose binding protein (MBP) and a TEV cleavage site. All the plasmids were verified using standard dideoxy sequencing.

2.3. Protein Expression and Purification

All the constructs were transformed into *E. coli* BL21 (DE3). The cells were grown at 37 °C in Luria Bertoni medium containing ampicillin 100 μg/mL (for pET22b constructs) or kanamycin 50 μg/mL (for pET28a and pETM-40 constructs) until the optical density at 600 nm reached 0.6. The protein expression was then induced by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and cells were further grown at 18 °C for 18 h. The cells were harvested by centrifugation and the pellet was suspended in buffer A (50 mM Tris-HCl buffer at pH 7.5 containing 300 mM NaCl, and 0.2 mM Tris(2-carboxyethyl)phosphine (TCEP)) and supplemented with an EDTA-free Complete Protease Inhibitor Cocktail (Roche, Bâle, Switzerland). The cells were disrupted by sonication and the crude extract was cleared by centrifugation at 35,000× *g* at 4 °C for 30 min. For proteins containing a 6xHis-tag (RABV P_{68}, P_{42}, and P_{FL}), the supernatant was loaded onto an Ni−NTA Agarose (Qiagen, Hilden, Germany) column equilibrated in buffer A. The column was washed with ten column volumes of buffer B (50 mM Tris-HCl buffer at pH 7.5 containing 1 M NaCl, 15 mM imidazole, and 0.2 mM TCEP) and the protein was eluted with buffer C (50 mM Tris-HCl buffer at pH 7.5 containing 150 mM NaCl, 400 mM imidazole, and 0.2 mM TCEP) and dialyzed overnight against buffer D at 4 °C (50 mM Tris-HCl buffer at pH 7.5 containing 150 mM NaCl and 0.2 mM TCEP). For RABV P_{42} containing a TEV site, the overnight dialysis against buffer D was performed in the presence of the TEV protease at an approximate weight ratio of 100:2 (fusion protein/TEV). After concentration with Vivaspin 20 (3 kDa molecular weight cut-off (MWCO), Sartorius, Göttingen, Germany) or with Amicon (molecular mass cutoff, 10 kDa) (Millipore, Burlington, MA, USA) concentrators, the protein was loaded on a S75 Superdex column (Cytiva, Malborough, MA, USA) equilibrated in buffer D. For $N_{\Delta 23}$−MBP, the supernatant was loaded onto an amylose resin column (New England BioLabs, Ipswich, MA, USA) equilibrated in buffer A. The column was washed with ten column volumes of buffer B and the protein was eluted with buffer E (50 mM Tris-HCl pH 7.5 containing 150 mM NaCl, 0.2 mM TCEP, 50 mM arginine, 50 mM glutamate, and 250 mM maltose). The fractions containing the protein were pooled and concentrated to 3 mg/mL with Amicon concentrators (molecular mass cutoff, 30 kDa) (Millipore), the protein was loaded on a S200 Superdex column (Cytiva) equilibrated in buffer D at 4 °C. Each protein sample was analyzed by SEC-MALLS (Multi-Angle Static Light Scattering—Dawn Heleos II, Wyatt, Santa Barbara, CA, USA) and the concentration was measured by on-line refractometry (Optilab T-rex refractometer, Wyatt).

2.4. Peptide Labelling with a Fluorescent Dye

A solution of concentrated His_6-SUMO-TEV-$P_{42\text{-}G41C}$ protein in buffer D was incubated for 20 min with a 10-fold molar excess of TCEP and flushed for 1 min with nitrogen

gas (N_2). A 10-fold molar excess of 5-carboxyfluorescein (5-FAM) maleimide (Thermo Scientific, Waltham, MA, USA) dissolved in dimethylsulfoxyde (DMSO) was added dropwise and the mixture was incubated overnight in the dark. A 10-fold molar excess of reduced glutathione was added and the solution was loaded on a PD10 desalting column (Econo-Pac® 10DG columns, Bio-Rad, Hercules, CA, USA) and eluted with buffer D. The protein was cleaved with the TEV protease at an approximate weight ratio of 100:2 (fusion protein/TEV) and overnight incubation at 4 °C to remove the tag. The labeling efficiency was assessed using liquid chromatography coupled to electrospray ionization mass spectrometry (ISBG facility, Grenoble). The labeled peptide $P_{1–42,\,G41C}$-FAM was analyzed by SEC-MALLS (Multi-Angle Static Light Scattering—Dawn Heleos II, Wyatt) and the concentration was measured by on-line refractometry (Optilab T-rex refractometer, Wyatt).

2.5. Reconstitution of the $N_{\Delta 23}{}^0$–P_{68}, $N_{\Delta 23}{}^0$–$P_{42\text{-}FAM}$, and $N_{\Delta 23}{}^0$–P_{FL} Complexes

The purified MBP-TEV-$N_{\Delta 23}$ was mixed with an excess of purified peptide (P_{68}–His_6, P_{42}–FAM) or of purified P_{FL} and the mixture was incubated for 2 h at 4 °C. The MBP-tag was cleaved with the TEV protease at an approximate weight ratio of 100:2 (fusion protein/TEV) and overnight incubation at 4 °C. The solution was concentrated and loaded onto a S200 Superdex (Cytiva) column coupled to a short amylose resin (NEB) column equilibrated in 50 mM HEPES buffer at pH 7 containing 150 mM NaCl, 50 mM arginine, 50 mM glutamate, and 0.2 mM TCEP to completely remove cleaved MBP.

2.6. Size Exclusion Chromatography and Multiangle Laser Light Scattering (SEC-MALLS)

The size exclusion chromatography (SEC) combined with on-line detection using multiangle laser light scattering (MALLS) and refractometry is a method for measuring the absolute molecular mass of a particle in a solution that is independent of its dimensions and shape [53]. The SEC was performed at 20 °C with a Superdex 200 or Superdex 75 column (Cytiva) equilibrated with a 20 mM Tris-HCl buffer containing 150 mM NaCl and 0.2 mM TCEP and a flow rate of 0.5 mL.min^{-1}. The column was calibrated with globular standard proteins of known hydrodynamic radius (R_h) [54]. MALLS detection was performed on-line with a DAWN-HELEOS II detector (Wyatt Technology, Santa Barbara, CA, USA) using a laser emitting at 690 nm and protein concentration was measured on-line using differential refractive index measurements with an Optilab T-rEX detector (Wyatt Technology) and a refractive index increment, dn/dc, of 0.185 mL.g^{-1}. The weight-averaged molecular mass was calculated using the ASTRA software (Wyatt Technology).

2.7. Competition Anisotropy Binding Assay

In the kinetic assay, the $N_{\Delta 23}{}^0$−$P_{42\text{-}FAM}$ complex at 100 nM in buffer F was mixed with the unlabeled $P_{42\text{-}G41C}$ at a final concentration of 15 µM and fluorescence anisotropy was measured with a spectrofluorimeter (Quantamaster QM4CW Photon Technology International, Birmingham, NJ, USA) with polarizers and thermostated at 20 °C. The excitation and emission wavelengths were set at 495 nm and 520 nm, respectively. In the equilibrium binding assay, the $N_{\Delta 23}{}^0$−$P_{42\text{-}FAM}$ complex (100 nM) was mixed with serially diluted unlabeled peptides in black 96-well plates. The samples were then incubated for 24 h at room temperature and fluorescence anisotropy was measured using a plate reader (Clariostar, BMG LABTECH, Ortenberg, Germany) at 20 °C using an excitation filter at 485 nm and an emission filter at 535 nm. All the experiments were carried out in triplicate. The curve fitting was performed with the software program Dynafit [55].

2.8. X-ray Crystallography

The crystallization screenings of the $N_{\Delta 23}$−P_{68} complex were carried out at the High Throughput Crystallization Laboratory of the EMBL Grenoble Outstation (HTX Lab.). The crystals of the RABV $N_{\Delta 23}$−P_{68} complex were obtained at 20 °C in (0.02 M sodium/potassium phosphate and 0.1 M Bis Tris propane at pH 6.5 containing 20% *w*/*v* PEG 3350). The diffraction data were collected on the ID23-1 beamline at the ESRF (Greno-

ble, France). The $N_{\Delta 23}-P_{68}$ complex crystallized in space group $P2_12_12$ with one complex per asymmetric unit. The single crystals were harvested from the drop, briefly soaked in the reservoir solution supplemented with 25% glycerol and flash frozen in liquid nitrogen at 100 K before data collection. The data were processed using the program XDS [56] and scaled with the program Scala from the ccp4 suite [57]. The structure was solved by molecular replacement with the program Phaser [58] using a protomer of N extracted from the N−RNA crystal structure (2GTT.pdb) [39] as a search model. The overall structure was refined to a resolution of 2.3 Å using Coot [59] and Refmac5 [60] and Buster [61]. The quality of the model was checked with PROCHECK [62]. The data collection and refinement statistics are summarized in Table 1.

Table 1. Data collection and refinement statistics.

	$N_{\Delta 23}{}^0$–$P_{1\text{-}68}$ (Molecular Replacement)	**N_{11}–RNA (Re-Refinement)**
Data Collection		
Space group	$P2_12_12_1$	$P2_12_12$
Cell dimensions		
a, *b*, *c* (Å)	39.6, 73.6, 154.3	270.4, 281, 236.9 [b]
α, β, γ (Å)	90.0, 90.0, 90.0	90.0, 90.0, 90.0 [b]
Resolution (Å) [a]	50.0–2.3 (2.44–2.30) [a]	100–3.49 (3.63–3.49) [b]
R_{merge} (%) [a]	12.7% (52.9%)	Not reported [b]
I/σI [a]	12.31 (3.76)	12.15 (3.36) [b]
Completeness (%) [a]	99.8 (99.5)	99.9 (99.9) [b]
Redundancy	7.3	Not reported [b]
Refinement		
Resolution (Å) [a]	42.17–2.30 (2.36–2.30)	49.75–3.49 (3.62–3.49)
Numbers of reflections [a]	19,735 (1397)	226,495 (21,091)
R_{work}/R_{free} (%) [a]	16.2/18.7	22.8/25.2 (26.9/31.3 [c])
Numbers of atoms		
Protein	3243	73,326
RNA	-	4103
Water	266	0
B factors		
Protein	36.2	106.9
RNA	-	105.9
Water	46.3	-
r.m.s deviations		
Bond lengths (Å)	0.010	0.005
Bond angles (°)	1.03	0.87
Ramachandran statistics		
Ramachandran favored (%)	97.3	96.25
Ramachandran allowed (%)	2.7	3.60
Ramachandran outliers (%)	0.0	0.15

[a] Values in parentheses are for highest-resolution shell. [b] Values are reproduced from the original deposition (2GTT.pdb). [c] R_{work}/R_{free} of the original pdb entry 2GTT are indicated in parentheses.

We re-refined the circular RABV N_{11}−RNA crystal structure (PDB ID 2GTT) in order to correct a frameshift error present in the original entry. The error was detected by comparing the structure of one N subunit from the N_{11}−RNA crystal structure with the N structure from the newly solved $N_{\Delta 23}{}^0-P_{68}$ crystal structure (8B8V.pdb) and by performing MD simulations of an N−RNA trimer extracted from the N_{11}−RNA crystal structure

(2GTT.pdb), which showed significant unfolding of the N-terminal domains within a few hundred ns of MD. The N_{11}−RNA structure was re-refined using a combination of manual rebuilding in Coot [59] based on the new $N_{\Delta 23}{}^{0}$−P_{68} crystal structure, molecular dynamics flexible fitting (MDFF) in ISOLDE [63], and repetitive rounds of restrained refinement in PHENIX [64] and Autobuster [61]. The new coordinates have been deposited in the PDB under the code 8FFR.

2.9. Small Angle X-ray Scattering (SAXS) and SEC-SAXS Experiments

The small angle X-ray scattering (SAXS) experiments were performed on the BioSAXS beamline BM29 and former beamline ID14-3 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France and on the SWING beamline at SOLEIL, Paris, France. For direct SAXS experiments, the scattering from the buffer alone was measured before and after each sample measurement and used for background subtraction. The 1D scattering profiles were generated and buffer subtraction was carried out by the automated data processing pipeline available at the different beamlines. For SEC-SAXS experiments, the samples were loaded onto a Superdex^{TM} 200 increase 5/150 GL equilibrated with 50 mM HEPES buffer at pH 7.5 containing 150 mM NaCl and 0.5 mM TCEP. The program Foxtrot [65] was used to integrate and subtract the SEC-SAXS frames. All the data were analyzed with the program PRIMUS from the ATSAS 3.0.0 package [66]. The radius of gyration was determined with the program PRIMUS according to the Guinier approximation at low q values and the molecular weights were estimated based on the invariant V_c and R_g values [67].

For isolated RABV P_{68}, an ensemble of 10,000 conformers was generated using the software Flexible-Meccano [68] and sidechains were added using SCWRL4 [69]. For the $N_{\Delta 23}{}^{0}$−P_{68} complex, an ensemble of 6639 models was built by extracting snapshots of explicit solvent MD trajectories (see Methods—MD simulations), using a time step of 1 ns. For the $N_{\Delta 23}{}^{0}$−P_{68} complex, several ensembles of 10,000 models were built with the software programs RANCH [70], PD2 [71], and SCWRL4 [69]. The theoretical SAXS patterns were calculated with the program CRYSOL [72] and ensemble optimization fitting was performed with GAJOE [70]. The optimum selected ensemble size and relative weights of the models were determined automatically by GAJOE.

2.10. Molecular Dynamics (MD) Simulations

The classical explicit solvent MD simulations were used to study the conformational dynamics of the N protein in its free form (N^0), in complex with P_{68} (N^0−P_{68}), and in its RNA-bound form (N−RNA). The N^0−P_{68} MD trajectories were additionally used to provide an ensemble of models suitable for fitting the SAXS data using EOM. The N^0 and N^0−P_{68} models were based on the $N_{\Delta 23}{}^{0}$−P_{68} crystal structure (missing residues in N^0 and P_{68} were constructed as random chains) and the simulated N−RNA model corresponded to a trimer of RNA-bound N protein extracted from the revised version of the 11-mer N−RNA ring crystal structure. All 3 systems were simulated in GROMACS [73] using an amber99SBws forcefield [74], which was designed to reproduce the properties of intrinsically disordered proteins. At the beginning of each simulation, the protein was immersed in a box of TIP4P2005 water, with a minimum distance of 1.0 nm between protein atoms and the edges of the box. The genion tool was used to add 150 mM NaCl. The long-range electrostatics were treated with the particle-mesh Ewald summation. The bond lengths were constrained using the P-LINCS algorithm. The integration time step was 5 fs. The v-rescale thermostat and the Parrinello–Rahman barostat were used to maintain a temperature of 300 K and a pressure of 1 atm. Each system was energy minimized using 1000 steps of steepest descent and equilibrated for 500 ps with restrained protein heavy atoms prior to production simulations; 4 or 5 independent MD trajectories were calculated for each system, representing a total aggregated simulation time of 3.1 µs (N^0), 6.6 µs (N^0−$P_{1\text{-}68}$), and 2.5 µs (N−RNA). RMSD, RMSF, and principal component analysis (PCA) were calculated using GROMACS routines.

3. Results

3.1. The N-Terminal Chaperone Module of RABV P Is Intrinsically Disordered

In a recent study, we have showed by nuclear magnetic resonance (NMR) spectroscopy and small-angle X-ray scattering (SAXS) that the entire N-terminal region of RABV P (from aa 1 to 90) is intrinsically disordered both in isolation and within the full-length dimeric P protein [17]. However, multiple sequence alignments revealed that the region encompassing residues from 1 to 60 was conserved within the genus Lyssavirus (Figure 1A) and our meta-analysis of disorder predicted that the first 55 residues were structured (Figure 1B). Together with the previous studies conducted with rabies virus and other viruses, these results suggested that the N-terminal end contained a molecular recognition element (MoRE) involved in chaperoning the unassembled, RNA-free nucleoprotein [23,25,40,43,75,76]. On the basis of this information, we generated two constructs that encompassed the N-terminal 42 or 68 residues of P; the former was fused to N-terminal cleavable 6-histidine and SUMO tags (His6-SUMO-P_{42}), while the latter was fused to a C-terminal two-amino acid linker (LE) and a 6-histidine tag ($P_{68}-His_6$) (Figure 1C). We purified both peptides using affinity chromatography on Ni−NTA resin and size-exclusion chromatography (SEC) on a Superdex S75 column. Each peptide eluted from the SEC column in a single peak and its weight-average molecular mass (M_w) determined from static light scattering and refractometry measurements (SEC-MALLS) was in agreement with the theoretical molecular masses of the respective monomers (Figure 1D). For RABV P_{68}, we recorded SAXS curves for scattering vector (q) values ranging from 0.05 to 4.0 nm^{-1} over a concentration range from 3.3 to 11.0 $mg.mL^{-1}$ (Figure 1E). The Guinier plots at low q values ($q.R_g < 1.3$) were linear (Figure 1F) and showed only a slight dependence on a protein concentration indicating the absence of aggregation but the presence of some interparticle interactions at the highest concentration (11 $mg.mL^{-1}$) (Figure 1G). To optimize the signal-to-noise ratio and eliminate any possible structure factor contribution, we merged the curves at the different concentrations [77] (Figure 1H) and used the software GAJOE to select ensembles of a limited number of conformers from a large initial ensemble that reproduced the experimental SAXS data [70]. The initial ensemble of the all-atom models generated with Flexible-Meccano [68] displayed a Gaussian distribution of R_g and D_{max} values as expected for a random coil (Figure 1I,J). The selected ensembles of the 12 conformers adequately reproduced the SAXS curve ($\chi^2 = 0.824$) (Figure 1H), while the R_g and D_{max} distributions of the selected ensembles were comparable to those of the initial ensemble, confirming that P_{68} obeyed random coil statistics (Figure 1I,J). Figure 1K shows a representative ensemble of 12 conformers.

3.2. The Chaperone Module of RABV P Blocks N in a Monomeric Closed Conformation

We expressed a truncated form of RABV N lacking its N-terminal arm ($N_{NT\text{-}ARM}$—aa 1–23) in fusion with a TEV cleavable N-terminal MBP tag (Figure 2A) (MBP-$N_{\Delta 23}$). We purified the protein using affinity chromatography on an amylose resin. We then mixed the preparations of MBP-$N_{\Delta 23}$ and purified $P_{68}-His_6$, cleaved the MBP tag with the TEV protease, and purified the ${N_{\Delta 23}}^0-P_{68}$ complex using SEC on a Superdex S200 column. We showed by SEC-MALLS that the ${N_{\Delta 23}}^0-P_{68}$ complex eluted as a single, symmetric peak from the Superdex S200 column (Figure 2B). The M_w value of 58 ± 3 kDa obtained from light scattering and refractometry measurements was in agreement with the theoretical molecular mass of 56.7 kDa calculated for the heterodimer (Figure 2B). The M_w was constant throughout the chromatographic peak with M_w/M_n ratio (where M_n is the number-average molecular mass) of 1.001, indicating a monodisperse species.

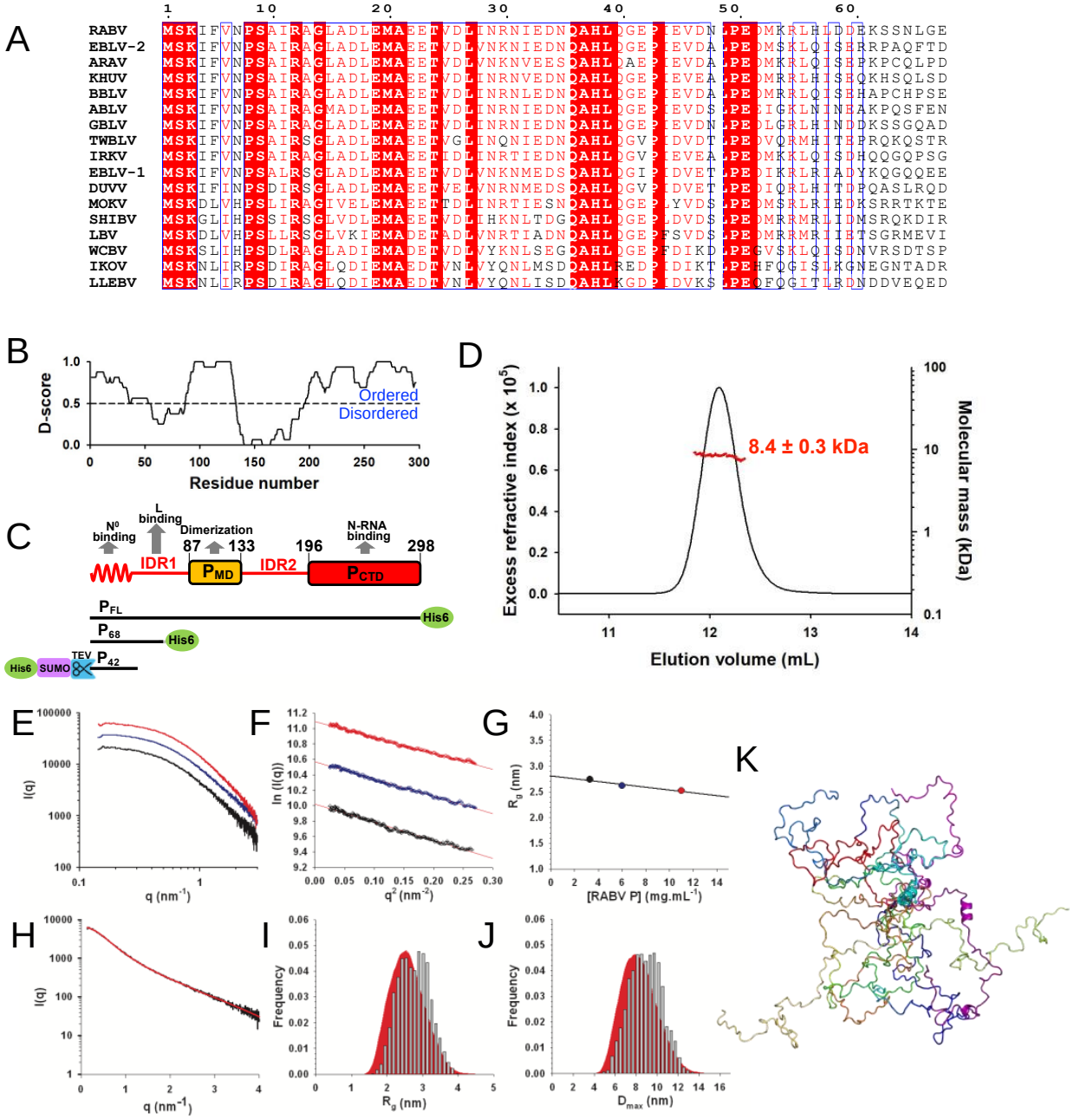

Figure 1. The N^0 chaperone module of RABV phosphoprotein. (**A**) Multiple sequence alignment of Lyssavirus P_{NTR} region. Members of the Lyssavirus genus and their UniProt accession number: RABV—rabies virus CVS−11 strain P22363, EBLV−1—European bat lyssavirus 1 A4UHP9, EBLV−2—European bat lyssavirus A4UHQ4, ARAV—Aravan virus Q6X1D7, KHUV—Khujand virus Q6X1D3, BBLV—Bokeloh bat lyssavirus U3MZL8, ABLV—Australian bat lyssavirus Q9QSP3, GBLV—Gannoruwa virus A0A1J0RI70, TWBLV—Taiwan bat lyssavirus A0A3P8MNG3, IRKV—Irkut virus Q5VKP5, DUKV—Duvenhage virus O56774, MOKV—Mokola virus P0C569, SHIBV—Shimoni bat virus D4NRJ9, LBV—Lagos bat virus D4NRK4, WCBV—West Caucasian bat virus Q5VKP1, IKOV—Ikoma virus J5JWQ7, LLEBV—Lleida virus A0A1I9RG27. (**B**) D-score. A consensus disordered prediction (D-score) was calculated as described in [20]. The threshold to distinguish between the ordered and disordered region was arbitrarily set at 0.5. The shaded areas indicate the positions of the known folded dimerization domain (P_{MD}) and NC-binding C-terminal domain (P_{CTD}) (see panel (**C**)). (**C**) Schematic representation of RABV phosphoprotein modular organization and constructs. The upper part shows the structural organization of the phosphoprotein. Boxes indicate the localization of folded domains, undulated lines the localization of predicted MoRE and lines the

localization of intrinsically disordered regions (IDR 1−2). The grey arrows indicate the location of the regions associated with functions in RNA synthesis. The lower part shows the three constructs of the phosphoprotein used in this study, indicating the position of the tags and cleavage sites. (**D**) Size exclusion chromatography and multiple-angle laser light scattering (SEC-MALLS) of RABV P_{68}. The elution was monitored on-line using multi-angle laser light scattering and differential refractometry. The line shows the chromatograms monitored using differential refractive index measurements. The red crosses indicate the molecular mass across the elution peak calculated from static light scattering and refractive index, and the numbers indicate the weight-averaged molecular mass (kDa) with standard deviations. (**E**–**K**) SAXS experiments and modeling. SAXS profiles recorded at 3.3 mg.mL^{-1} (in black), 6.0 mg.mL^{-1} (in blue), and 11.0 mg.mL^{-1} (in red) are shown in direct plot (**E**) and Guinier plots (**F**). (**G**) R_g values obtained from the Guinier approximation. (**H**) Merged SAXS curve (in black). The red line shows the fit obtained by the EOM method with an ensemble of 12 conformers (see panels (**J**,**K**)). (**I**) R_g distribution. The shaded area shows the R_g distribution of the initial ensemble. The bars show the R_g distribution of selected ensembles. (**J**) D_{max} distribution. The shaded area shows the D_{max} distribution of the initial ensemble. The bars show the D_{max} distribution of selected ensembles. (**K**) Representative ensemble of 12 conformers selected by GAJOE. Each conformer is shown in a different color.

The $N_{\Delta 23}{}^0-P_{68}$ complex crystallized with a single molecule in the asymmetric unit and the crystals diffracted up to 2.3 Å resolution (Table 1). The structure was solved by molecular replacement using the structure of an N protomer extracted from the circular N−RNA complex (PDB code: 2GTT) [39]. The $N_{\Delta 23}{}^0-P_{68}$ complex contained no RNA and an analysis with the webserver PDBePISA revealed no intermolecular interactions between neighboring heterodimers besides the crystal contacts (Complex Formation Significance Score (CSS) = 1.0 for the interaction between $N_{\Delta 23}{}^0$ and P_{68} and CSS = 0.0 for all other contacts), demonstrating the absence of N polymerization in the crystal (Figure S1A). The $N_{\Delta 23}$ molecules arranged themselves side by side in linear arrays in the crystal and their N-terminal domains contact each other (Figure S1B), but without forming a tight complex as in NC (Figure S1C). The $N_{\Delta 23}$ exhibited the typical *Mononegavirales* nucleoprotein fold with two globular domains, N_{NTD} and N_{CTD}, connected by a hinge (Figure 2C). In the absence of neighboring N subunits, the entire $N_{CT\text{-}ARM}$ (aa 350–400) was not visible in the crystal structure and a 5-residue loop in N_{NTD} (aa 155–159) was also absent from the crystal structure, suggesting that it was flexible in the absence of RNA as expected, since this loop interacted with one of the bound nucleotides in the N−RNA complex [39]. Two parts of P_{68} were clearly visible in the electronic density; residues from 4 to 15 were bound in a groove of N_{CTD} with residues from 9 to 12 forming a helical turn, while residues from 20 to 39 formed a long α helix that docked at the interface between N_{NTD} and N_{CTD} (Figure 2C). The four residue-connecting loops (aa 16–19) were less well-defined in the electron density map, suggesting that it was slightly flexible in the crystal (Figure 2C). The complex was stabilized by multiple interchain interactions. The binding of the P_{68} N-terminal moiety involved three H-bonds and four salt-bridges and buried a surface area of 1190 Å^2, whereas the binding of the P_{68} C-terminal moiety involved thirteen H-bonds and three salt bridges and buried a surface area of 1617 Å^2 (Supplementary Table S1). The interface regions of N and some interface residues of P appeared to be conserved among the Lyssavirus (Figure 2D,E).

The structural alignments using PDBeFold showed that N_{CTD} in our refined $N_{\Delta 23}{}^0-P_{68}$ structure was similar to that in the N−RNA complex [39] (RMSD = 0.84 Å) (Figure 2F), while the N_{NTD} exhibited some differences with residues from 35 to 103 and from 189 to 200 being out-of-register in the structural alignment despite an RMSD of 0.95 Å (Figure 2G). An analysis of the original electron density map of the N−RNA complex [39] revealed that the absence of density for residues from 105 to 118 and 185 to 188 and the low resolution of density had resulted in the incorrect positioning of the latter region relative the remaining N_{NTD} (see below).

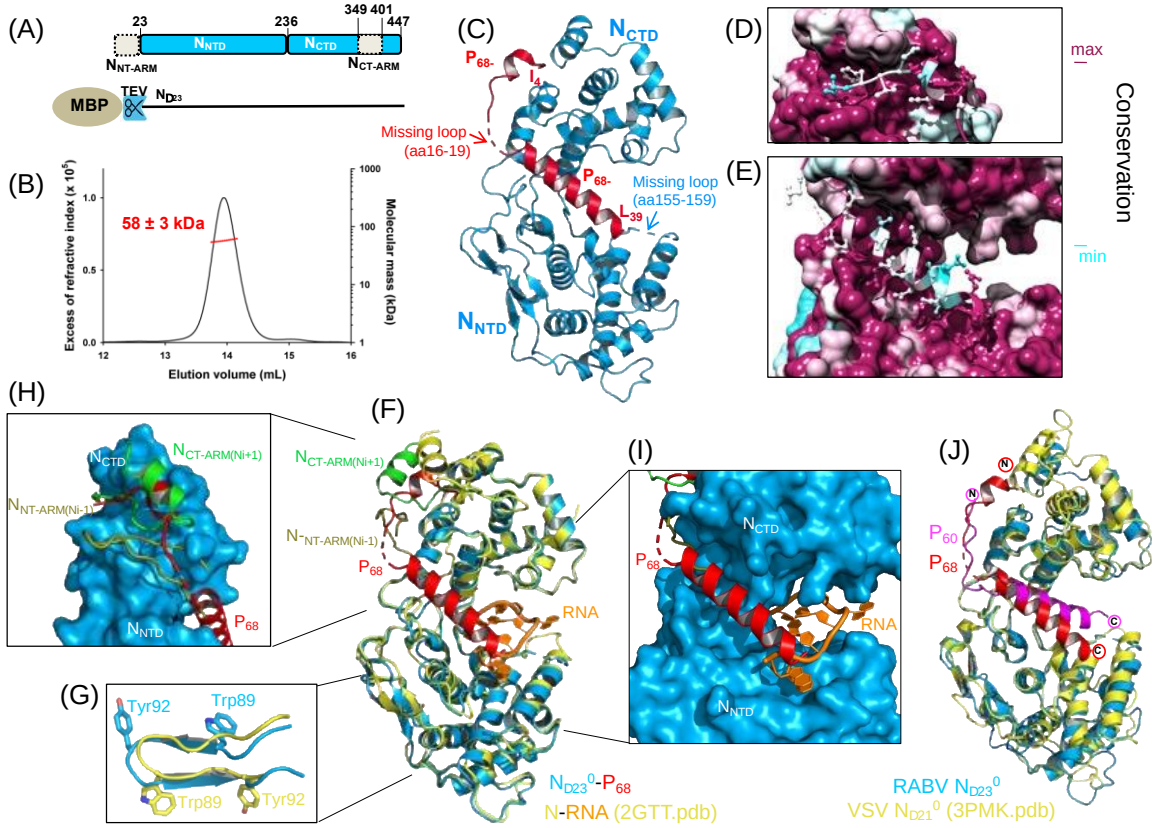

Figure 2. Crystal structure of RABV $N_{\Delta 23}{}^{0}-P_{68}$ core complex. (**A**) Schematic representation of RABV nucleoprotein modular organization and construct. The upper part shows the structural organization of the nucleoprotein. Boxes indicate the localization of folded domains, dotted boxes the localization of the exchanging subdomains $N_{NT\text{-}ARM}$ and $N_{CT\text{-}ARM}$. The lower part shows the construct of the nucleoprotein used in this study. (**B**) SEC MALLS of the reconstituted complex. The elution was monitored on-line using multi-angle laser light scattering and differential refractometry. The line shows the chromatograms monitored by differential refractive index measurements. The red crosses indicate the molecular mass across the elution peak calculated from static light scattering and the numbers indicate the weight-averaged molecular mass (kDa) with standard deviations. Theoretical mass of the heterodimer = 48,208.5 Da ($N_{\Delta 23}$) + 8488.4 Da (P_{68}) = 56,696.9 Da. (**C**) Crystal structure of RABV $N_{\Delta 23}{}^{0}-P_{68}$ core complex (8b8v.pdb) in cartoon representation. $N_{\Delta 23}{}^{0}$ is shown in blue and P_{68} in red. The dotted line in P_{68} shows the position of the four residue loop that is less-well defined in the electron density map, while the dotted lines in N show the position of the missing loops in N_{NTD} and the $N_{CT\text{-}ARM}$. The N- and C-terminal residues of P_{68} are indicated. (**D**) Conservation of the $N^{0}-P$ interface. View of the region of interaction of residues 4 to 15 of P_{CM}. The complex is shown with surface and conservation representations for $N_{\Delta 23}$ and with stick-and-ball representations for P_{68}. The conservation in N derived from multiple sequence alignment is displayed on the surface of NiV N: blue low-level conservation, <20%; maroon, high-level conservation, >80%. The side chains of conserved residues in the P N-terminal region are shown in stick representation. (**E**) Conservation of the $N^{0}-P$ interface. View of the region of interaction of residues from 20 to 39 of P_{CM} with same color scheme and representation as in panel (**D**). (**F**) Superposition of RABV $N^{0}-P$ and N−RNA structures. The structures were superposed by aligning the C-terminal domains of both structures. (**G**) Out-of-register construction in N_{NTD}. Close-up view of a small region of N_{NTD} illustrating the change of register of residues from 27 to 130 between the original circular $N_{11}-RNA$ structure (2GTT.pdb) shown in yellow and the $N_{\Delta 23}{}^{0}-P_{68}$ structure (8B8V.pdb) shown in blue. (**H**) Interference of P_{CM} binding with $N_{NT\text{-}ARM}$ and $N_{CT\text{-}ARM}$ binding. Close-up view of the interference between P_{CM} (in red) and the $N_{NT\text{-}ARM}$ from the $N_{i\text{-}1}$ subunit (in olive) and $N_{CT\text{-}ARM}$ of the Ni+1 subunit (in green). (**I**) Interference of P_{CM} binding with RNA binding. Close-up view of the interference between P_{CM} in (red) and RNA (in orange). (**J**) Superposition of RABV $N_{\Delta 23}{}^{0}-P_{68}$ and VSV $N_{\Delta 21}{}^{0}-P_{60}$ complex.

As in other *Mononegavirales* N^0-P complexes, the P chaperone module prevented the polymerization of N by interfering with the binding of the $N_{NT\text{-}ARM}$ and $N_{CT\text{-}ARM}$ subdomains from adjacent N subunits. The N-terminal part of RABV P_{68} competed with the $N_{CT\text{-}ARM}$ of the N*i*+1 protomer, whereas the C-terminal part competed with the $N_{NT\text{-}ARM}$ of the N*i*–1 protomer (Figure 2H). Additionally, in the RABV ${N_{\Delta 23}}^0-P_{68}$ complex, the RNA binding groove in the N protein adopted a close conformation as in the N−RNA complex [39] and in VSV ${N_{\Delta 21}}^0-P_{60}$ structure [40] (Figure 2F). The structural alignment with VSV ${N_{\Delta 21}}^0-P_{60}$ structure showed that the overall orientation of P_{68} relative to N was similar to that of P_{60} in VSV complex, although its exact position on the surface of N, notably its C-terminal part, was slightly different (Figure 2J). Finally, as in the VSV ${N_{\Delta 21}}^0-P_{60}$, the RABV P_{68} protruded into the RNA binding groove of N and interfered directly with RNA binding (Figure 2I and see below).

3.3. The Chaperone Module Forms a Compact but Fuzzy Complex with ${N_{\Delta 23}}^0$

To obtain structural information about the disordered parts of the RABV ${N_{\Delta 23}}^0-P_{68}$ complex, we turned to SAXS and SEC-SAXS. We injected a sample of ${N_{\Delta 23}}^0-P_{68}$ on a Superdex 200 Increase column and collected SAXS data at regular intervals along the elution peak (Figure 3A). We also collected SAXS curves in batch mode over a concentration range from 3.0 to 10.0 $mg.mL^{-1}$ (scattering vector (q) values ranging from 0.06 to 5.3 nm^{-1}) (Figure 3B). The shapes of the scattering curves were independent of protein concentration (Figure 3B). The Guinier plots at low q values ($q.R_g < 1.4$) were linear (Figure 3C) and the R_g value calculated by using Guinier approximation was constant throughout the chromatographic peak in the SEC-SAXS experiment (Figure 3A) and showed no dependence on the protein concentration (Figure 3D), thus indicating the absence of aggregation or intermolecular interaction. The dimensionless Kratky plots reached a plateau near 1.2 for $q.R_g$ values near 1.7, which is indicative of a globular structure (Figure 3E). Again, in an effort to optimize the signal-to-noise ratio and minimize the possible structure factor contribution, we merged the curves at the different concentrations [77] (Figure 3F) and used the software GAJOE to further explore the conformational diversity of the complex with the Ensemble Optimization Method (EOM). [70]. We generated an initial ensemble of ~6600 atomic models of the ${N_{\Delta 23}}^0-P_{68}$ complex by molecular dynamics simulations (see below) and used the software GAJOE to select sub-ensembles of three conformers that reproduced the experimental SAXS curve ($\chi^2_{exp} = 0.95$, Figure 3F). The R_g and D_{max} distributions of these selected sub-ensembles indicated the presence of two main populations of the ${N_{\Delta 23}}^0-P_{68}$ complex (Figure 3G,H). One population (50%) corresponded to conformers where the C-terminal part of P_{68} was located in the RNA binding groove and bound to the surface of N (conformers 1 and 2 in Figure 3I). This is in agreement with the presence in the flexible part of P_{68} (aa 40–68) of eight acidic residues (Asp + Glu) for only three basic (Lys + Arg) and the presence of basic residues in the RNA binding groove of N. This insertion in the RNA binding groove could also contribute to the chaperon activity of P by competing with RNA molecules (Figure S2). The other population (50%) corresponded to conformers where the C-terminal part of P_{68} extended in the solvent (conformers three in Figure 3I) and was rather extended in comparison with the statistical distribution of the initial ensemble (Figure 3G,H).

3.4. A Slow Off-Rate Sets the Affinity of RABV P Chaperone Module for N^0 in the Low Nanomolar Range

The interaction between RABV ${N_{\Delta 23}}^0$ and P_{68} resisted separation by size exclusion chromatography (Figure 2B), suggesting by a rule of thumb that the dissociation constant was less than 1 µM. However, our inability to prepare ${N_{\Delta 23}}^0$ protein in an isolated, monomeric form prevented the easy measurement of the affinity for P_{68}. In order to quantify this interaction in solution, we developed a competition fluorescence anisotropy assay to monitor the reaction (Figure 4A). We generated, expressed, and purified a variant of P_{42} where we replaced the glycine residue at position 41 by a cysteine ($P_{42\text{-}G41C}$) and we

chemically labeled this cysteine with fluorescein (FAM). We showed by mass spectrometry that the labeling was complete at more than 95% (Figure S3). We reconstituted the complex with the fluorescently labeled peptide ($N_{\Delta 23}{}^{0}-P_{42-G41C}$*), purified it by SEC, and confirmed by SEC-MALLS analysis that it formed a heterodimer. At a concentration of 100 nM of $N_{\Delta 23}{}^{0}-P_{42\text{-}G41C}$* complex, fluorescence anisotropy was near 0.24, whereas that of the isolated peptide $P_{42\text{-}G41C}$* was near 0.07 (Figure 4B). We measured the dissociation kinetics at 20 °C in the absence and presence of an unlabeled competing peptide. In the absence of competing peptide, the anisotropy value remained stable for more than 24 h (Figure 4B), supporting a low nanomolar range affinity constant. In the presence of a large excess of unlabeled competing P_{42} or $P_{42\text{-}G41C}$ (final concentrations of 15 µM), we observed a monotonous decrease in fluorescence anisotropy revealing that the fluorescent peptide dissociated from $N_{\Delta 23}{}^{0}$ and that the system evolved toward a new equilibrium (Figure 4B). In a first approach, assuming an irreversible first-order dissociation process, we fitted the dissociation curve in the presence of unlabeled peptide with a single exponential equation and obtained a k_{off} value of 0.0027 ± 0.0001 min^{-1} (half time = 256 min). In a second approach, we assumed similar k_{on} and k_{off} rate constants for $P_{42\text{-}G41C}$* and $P_{42\text{-}G41C}$ and fitted the dissociation curve in the presence of the competitor peptide to a reversible exchange mechanism (Figure 4A) by using a numerical integration method implemented in the software Dynafit [55] and obtained k_{off} = 0.0027 ± 0.0001 min^{-1} (45.10^{-6} s^{-1}) and k_{on} = (5.7 ± 0.1) 10^{-3} nM^{-1} min^{-1} (95,000 M^{-1} s^{-1}). The dissociation constant K_d of 0.50 ± 0.03 nM was readily calculated from these rate constants (k_{off}/k_{on}). However, in the conditions used in our experiment, the kinetics were dominated by the off-rate as shown by the identical k_{off} value obtained in both fitting procedures. Thus, the value of k_{on} should be considered a lower bound value, as no significant difference in the goodness of fit was found when larger values were imposed and the K_d value of 0.50 nM represented an upper bound value. Considering an upper limit for a bimolecular protein–protein interaction on-rate (k_{on}) near 5 × 10^5 M^{-1} s^{-1} [78], a lowest bound estimate for K_d value of 0.09 nM was obtained.

We then measured equilibrium competition binding curves by incubating the FAM-labeled $N_{\Delta 23}{}^{0}-P_{42\text{-}G41C}$* complex with serially diluted concentrations of unlabeled $P_{42\text{-}G41C}$ and P_{68}. The optimal incubation time required for the system to reach equilibrium, mainly depending on the off-rate (k_{off}), was set by a rule of thumb at 24 h (5 × 0.693/k_{off} = 1284 min = 21 h). The curve provided a measure of the relative affinity of unlabeled peptide to labeled peptide. All things being equal, we expected that the half-saturation concentration should be that of the initial concentration of the labeled complex ($C_{1/2}$ = 100 nM indicating that in a solution containing 100 nM labeled peptide and 100 nM unlabeled peptide, the protein was 50% saturated by the labeled peptide). The curve obtained with unlabeled $P_{42\text{-}G41C}$ had a midpoint ($C_{1/2}$) around 300 nM, whereas that with P_{68} had a midpoint near 1000 nM (Figure 4C), revealing that the FAM moiety slightly contributed to the stability of the complex (previous attempts with P_{42} labeled with FAM at its N-terminus revealed an even stronger stabilization) and that the C-terminal part of P_{68} slightly reduced the stability of the complex. If we assumed K_d* = 0.5 nM for the labeled peptides, fitting the equilibrium competition binding curves by numerical integration with the software Dynafit [55] yielded K_d values of 4.5 ± 0.5 nM (ΔΔG = +1.3 kcal/mol) for $P_{42\text{-}G41C}$ and of 8.2 ± 0.6 nM (ΔΔG = +1.7 kcal/mol) for P_{68} (Figure 4C and Table 2).

To further characterize how individual residues in P_{68} contributed to the binding affinity for $N_{\Delta 23}{}^{0}$, we analyzed the structure with FoldX, a knowledge-based algorithm for predicting hot spots in protein–protein interfaces [52]. Four residues were predicted to play a key role in the stability of the complex with ΔΔG > 2 kcal/mol for Ala replacement: Phe5, Pro8, Val25, and Ile28, whereas seven other residues could be considered as milder contributors with 2 kcal/mol > ΔΔG > 1 kcal/mol: Ile4, Arg12, Glu22, Thr24, Leu27, Ile32, and Asn35. To test these predictions and validate our structural model, we generated, produced, and purified alanine variants for four of these residues (Phe5Ala, Pro8Ala, Arg12Ala, and Ile28Ala), including as a control the variant Arg12Ala, which was not predicted to be a

hot spot and measured equilibrium competition binding curves (Figure 4D and Table 2). The curve obtained with the mutant $P_{42,G41C,R12A}$ was close to that obtained with $P_{42,G41C}$, with a K_d value of 2.9 ± 0.6 nM. By contrast, the three other mutant peptides were less efficient than the wt peptides for displacing the FAM-labeled peptide and, consequently, the equilibrium binding curves were shifted to higher peptide concentrations, confirming the role of these residue side chains in stabilizing the complex. The results confirmed the importance of these key residues in the stabilization of the complex and the order of experimental ΔΔG values reflects the predictions of FoldX (Figure 4D and Table 2).

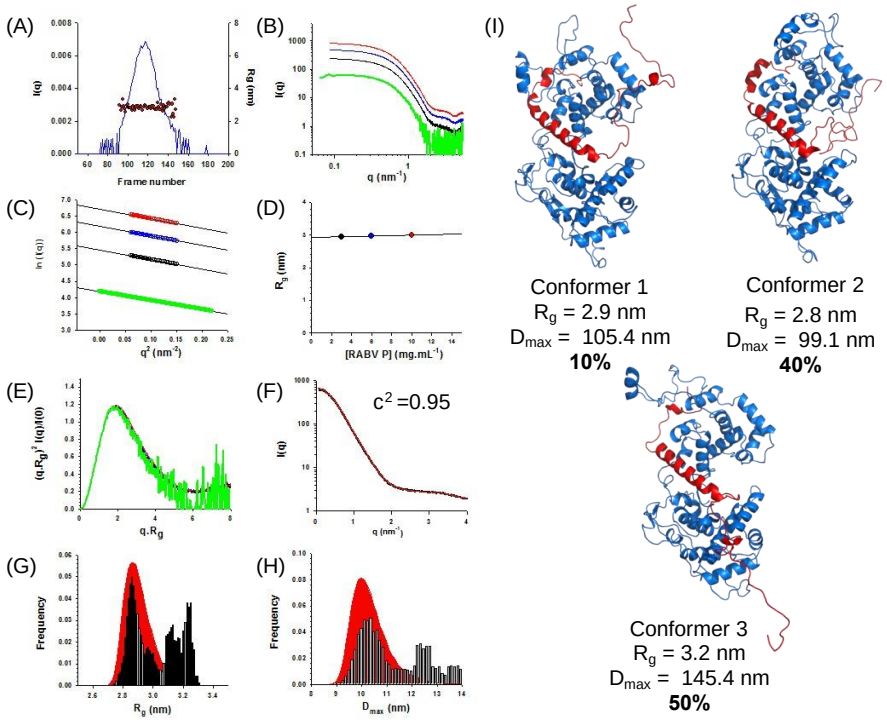

Figure 3. SAXS and SEC-SAXS of RABV $N_{\Delta 23}{}^{0}-P_{68}$ core complex. (**A**) SEC-SAXS elution profile and R_g across the elution peak; 50 µL of the $N_{\Delta 23}{}^{0}-P_{68}$ sample were injected onto a Superdex 200 column and monitored on-line by SAXS. The black line shows the intensity at zero angle (I_0), which is proportional to both the MM and concentration. The red dots indicate the values of the radius of gyration calculated from the Guinier approximation at the different time intervals. (**B**) SAXS profiles at different protein concentrations. SAXS profiles were recorded in batch mode at 3 mg.mL^{-1} (in black), 6 mg.mL^{-1} (in blue), and 11 mg.mL^{-1} (in red). The curve in green was obtained by averaging the individual profiles recorded throughout the SEC elution peak shown in Panel (**A**). (**C**) Guinier plots at different protein concentrations. Same color scheme as in panel (**B**). (**D**) Rg at different protein concentrations. The R_g value were calculated from the Guinier plot (panel (**C**)) for the profiles recorded in batch mode. (**E**) Dimensionless Kratky plots at different protein concentrations. Same color scheme as in panel (**B**). (**F**) Merged curve and conformational ensemble modeling by the ensemble optimization method (EOM). The black line shows the scattering profile obtained by merging segments of the profiles obtained at different protein concentrations (panel (**B**)). The red line shows a back-calculated scattering curve for a selected ensemble of three conformers measured in different proportions and shown in panel (**I**) (χ^2 = 0.95). (**G**) R_g distribution. The red area shows the R_g distribution calculated for the initial ensemble of conformers, whereas the black bars show the R_g distribution of the selected ensemble that fit the experimental SAXS data (panel (**F**)). (**H**) D_{max} distribution. The red area shows the D_{max} distribution calculated for the initial ensemble of conformers, whereas the black bars show the D_{max} distribution of the selected ensemble that fit the experimental SAXS data. (**I**) Representative ensemble of conformers. N is shown in blue and P_{68} including the eight C-terminal residues of the linker and His-tag are shown in red (cartoon representation).

Table 2. Dissociation constants determined by fitting the equilibrium competition binding curves. The K_D values were determined by fitting the curves in Figure 4D to the competition model shown in Figure 4A and fixing K_D^* = 0.5 nM. The $\Delta\Delta G^0$ values were calculated by reference to the K_D^* value for FAM labeled P_{42} using $\Delta\Delta G^0 = -RT \ln (K_D^*/K_D)$ and $\Delta\Delta G_{calc}$ are the values predicted by FoldX from the crystal structure.

Variants	K_D (nM)	$\Delta\Delta G^0$ (kcal.mol^{-1})	$\Delta\Delta G_{calc}$ (kcal.mol^{-1}) from FoldX
$P_{42,G41C}$	4.5 ± 0.5	1.3 ± 0.1	
$P_{42,G41C,F5A}$	150 ± 40	3.3 ± 0.9	4.4
$P_{42,G41C,P8A}$	80 ± 20	3.0 ± 0.8	2.4
$P_{42,G41C,R12A}$	2.9 ± 0.6	1.0 ± 0.2	1.7
$P_{42,G41C,I28A}$	2700 ± 1100	5.0 ± 2.0	3.5
P_{68}	8.2 ± 0.6	1.7 ± 0.1	

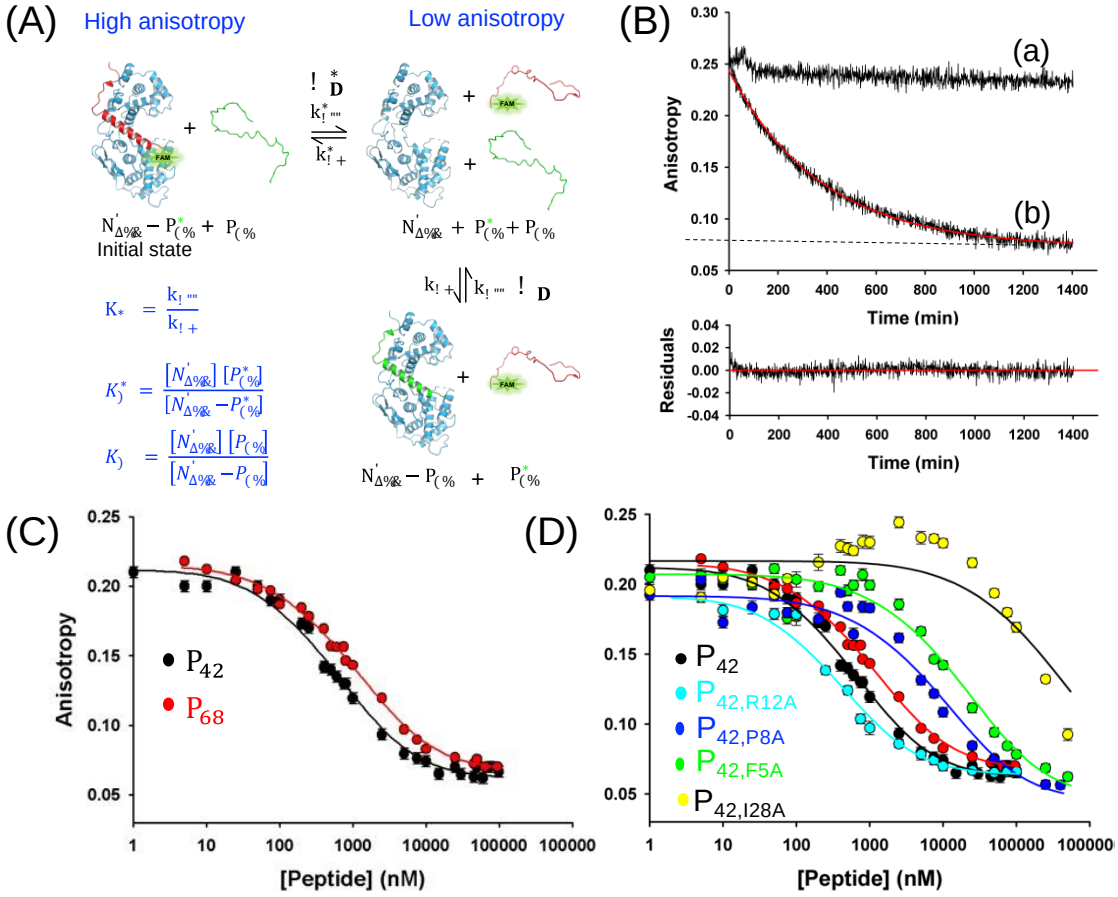

Figure 4. Competitive binding assay using fluorescence anisotropy. (**A**) Schematic representation of the competition assay and the mathematical models. (**B**) Dissociation kinetics; 200 µL of $N_{\Delta 23}{}^0-P_{68}$ at a concentration of 100 nM in 20 mM Tris/HCl pH 7.5, 150 mM NaCl at 20 °C. The upper panel shows the anisotropy variation in the absence (a) and presence (b) of a competing unlabeled peptide. The average anisotropy of P_{68}−FAM alone is indicated by a dotted line. The red line shows the fit obtained with DYNAFIT. The lower panel shows the plot of the residuals for the fit obtained with DYNAFIT and shown in the upper panel. (**C**) Equilibrium binding curves with WT peptides. The black circles are for P_{42} and the red circles for P_{68}. The lines show the fits obtained with DYNAFIT and the parameters shown in Table 2. (**D**) Equilibrium binding curves with mutants. The black circles are for P_{42} and are shown as reference. The light blue, dark blue, green, and yellow circles are for the mutant R12A, P8A, F5A, and I28A, respectively, and the lines show the fits obtained with DYNAFIT and the parameters shown in Table 2.

3.5. Binding of the Chaperone Module of RABV P and RNA-Mediated Polymerization of N Restrict Motions in N and Stabilize the Protein in Distinct Conformational States

In known structures of the N^0-P_{CM} complex, N was locked either in an open (*Paramyxoviridae*, *Pneumoviridae*, and *Filoviridae*) [26,28,43–47,79] or a closed conformation (*Rhabdoviridae*) ([40] and this paper). In *Rhabdoviridae*, the closed conformation seen in the N^0-P_{CM} complex is almost identical to the conformation observed in the polymeric, RNA bound state [39,41]. We hypothesized that P_{CM} interfered with the dynamics of the N protein as part of the chaperoning mechanism and, in particular, perturbed the opening and closing of the RNA binding groove [43]. To test this model, we used $N_{\Delta 23}{}^0$ alone (extracted from the crystal structure of $N_{\Delta 23}{}^0-P_{68}$), $N_{\Delta 23}{}^0-P_{68}$, and N_3-RNA, a set of three adjacent subunits extracted from the circular $N_{11}-RNA$ complex (2GTT.pdb), in which we built the flexible parts to generate full-length N molecules. We performed molecular dynamics (MD) simulations to evaluate the effects of P_{68}, RNA binding, and multimerization on the dynamics of N protein (Figure S4A–F). The $N_{\Delta 23}{}^0$ and $N_{\Delta 23}{}^0-P_{68}$ systems were stable over the course of the simulations, although high flexibility was observed in loop regions, particularly in the C-terminal loop ($N_{CT\text{-}ARM}$) (Figure S4G). For $N_{\Delta 23}{}^0-P_{68}$, the C-terminal part of P_{68} (res from 42 to 68 + tag) remained flexible and sampled a wide range of conformations (Figure S4H), consistent with the SAXS-based ensemble analysis (Figure 3). In the MD simulations of the N_3-RNA system, the C-terminal domains remained stable as previously found [80], but the N-terminal domains started to unfold in the first 100 ns of the simulation (Figure S5), reflecting possible packing defects due to the incorrect construction of the protein segment from residue from 135 to 200 (see above). We thus re-refined the $N_{11}-RNA$ model against the original dataset using more recent software [61,63,64] and the model of N extracted from the $N_{\Delta 23}{}^0-P_{68}$ structure (built from higher resolution data), which yielded an improved model of the circular undecameric N−RNA complex (Figure 5 and Table 1). Using three adjacent N subunits (N_3-RNA) from the re-refined N−RNA complex structure, the MD simulations showed that the N protein structure was stable over several hundred ns in multiple independent trajectories (Figure S4F). The root mean square deviation (RMSD) values stabilized around 0.6–0.8 nm after an initial rise during the first 100 ns.

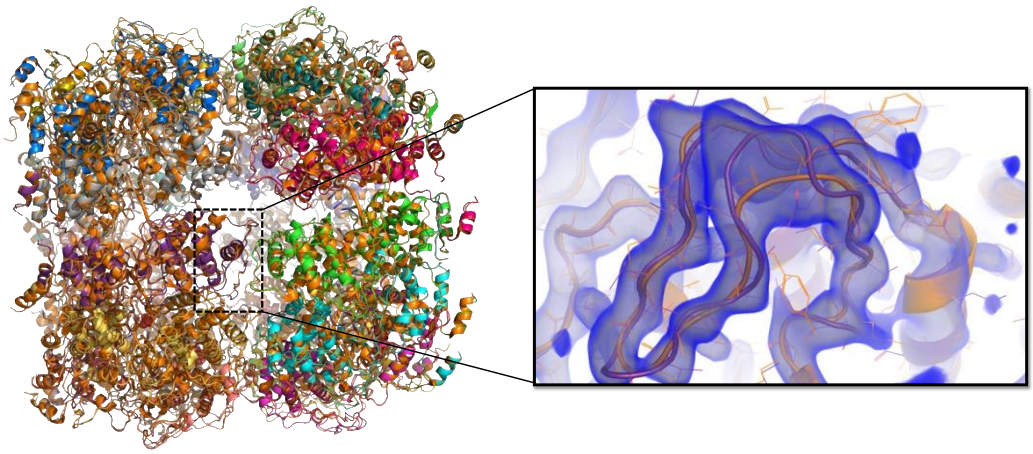

Figure 5. Re-refinement of the circular polymeric $N_{11}-RNA$ complex crystal structure. The different N subunits of the re-refined model are shown in cartoon representation with different colors for each chain and superimposed onto the original model (PDB ID 2GTT, shown in orange). The close-up view shows an excerpt of the region that was rebuilt (in orange), with the experimental 2Fo-Fc map shown in volume representation in blue and contoured at 1σ (drawn with PyMOL).

In order to understand how P_{CM} binding, RNA binding, and multimerization modulated the conformational landscape of the N protein, we performed a principal component analysis (PCA) on the MD simulation datasets, using the N residues common to all three systems (Figure 6A). The principal components analysis (PCA) revealed that about 45% of

the variance is accounted by the first two principal components, while about 70% of the variance was accounted by the first eight components (Figure 6B). The first component (PC1) represented a twist or shear motion of N_{NTD} relative to N_{CTD} around the hinge region, whereas the second component (PC2) corresponded mainly to the opening−closing movement of the RNA binding groove also around the hinge region (Figure 6C). The 2D projections of the two first principal components indicated that N^0 alone was more dynamics than N^0 in complex with P_{68} or than in the N−RNA complex, as evidenced by the larger basin sampled during MD (Figure 6A). Strikingly, we found that N^0 mainly populated a region in PC2 space characterized by $-4 \leq PC2 \leq 0$ (Figure 6A, left panel), which corresponded to a much more open state relative to N^0-P_{68} or N−RNA ($-1 \leq PC2 \leq 3$) (Figure 6A, middle and right panels). N^0 also sampled a large range of values for PC1 ($-4 \leq PC1 \leq 2$) with a main basin centered around PC1 ≈ from −2 to −1, while PC1 values for N−RNA were comprised in a shorter range between −3 and 0. This suggested that RNA binding and multimerization moderately impacted interdomain twisting of N as both systems showed PC1 values roughly centered around the same value (Figure 6C). By contrast, the PC1 values for N^0-P_{68} were mainly comprised between 0 and 3, implying that P_{68} binding twisted the relative orientation of N_{NTD} and N_{CTD} through its intermolecular contacts in the hinge region (Figure 6C).

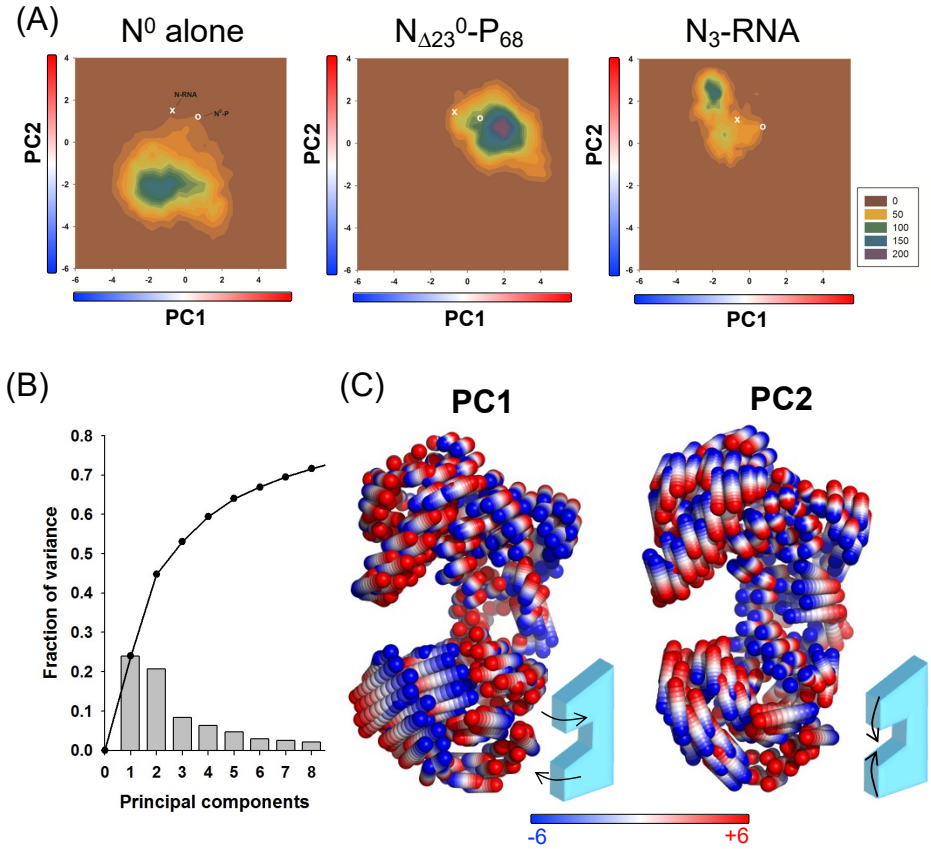

Figure 6. Molecular dynamic simulations. (**A**) Principal component analysis (PCA); 2D projections of the first two principal components (PCs) calculated for the N protein from MD simulations run with N^0 alone (left panel), $N_{\Delta 23}-P_{68}$ (middle panel) and N_3-RNA (right panel). The data are shown as a 2D histogram with the density distribution color scheme indicated on the right. The value calculated from the crystal structures of the $N_{11}-$RNA complex and $N_{\Delta 23}-P_{68}$ complex are shown for reference on each graph as a white cross and a white circle, respectively. (**B**) Fraction of the variance captured by each PC (histogram) and cumulative contributions of the first eight PCs. (**C**) Collective motions of N captured by PC1 and PC2. The motions are illustrated as linear interpolations between the extreme projections of the structures onto the PCs. Each cylinder thus describes the path of a Cα atom between its extreme positions (on a red–white–blue color scale).

In conclusion, the presence of P_{68}, as well as RNA and multimerization, restricted the dynamics of N and induced the closure of the RNA binding groove. The P_{68} binding further stabilized the N protein in a twisted conformation relative to N^0 or N−RNA. These data were consistent with an encapsidation mechanism in which the release of P_{CM} unlocked the opening−closing movements in the N protein, allowing the protein to grab the RNA.

3.6. Dimeric Full-Length RABV P Can Chaperone Two N^0 Molecules

The RABV phosphoprotein forms dimers [17,19,20,22]. To determine whether each chain in a dimeric P molecule was able to chaperone one N^0 molecule as previously shown for VSV P [42], we reconstituted the complex with full-length P (P_{FL}) and purified the $N_{\Delta 23}{}^0-P_{FL}$ complex by SEC on a Superdex 200 column. We performed a SEC-SAXS experiment by injecting 50 µL of a sample of $N_{\Delta 23}{}^0-P_{FL}$ into a Superdex 200 Increase column and monitored the elution by SAXS. We collected scattering profiles for q values ranging from 0.085 nm^{-1} to 3.0 nm^{-1} at regular time intervals during the elution. The complex eluted as a single peak, but the radius of gyration (R_g) determined by using the Guinier approximation at small scattering vector (q) values ($q_{max}.R_g < 1.3$) varied across the elution peak (Figure 7A) and the scattering intensity profiles in the front part (F1) and the tail part (F2) of the peak were slightly different (Figure 7B, upper panel), suggesting heterogeneity of the eluting species. The differential scattering intensity profile (Figure 7B, lower panel) indicated that the main differences between F1 and F2 were in the low and medium q ranges. The difference at low q was likely due to both differences in protein concentration (the protein concentration was not measured during elution and the profile could not be normalized) and in the molecular mass of the eluting particle (molecular heterogeneity). The negative band between 0.2 and 0.8 nm^{-1} in the differential scattering intensity profile corresponding to distances between 1 and 5 nm in the real space could result from the presence of an additional N molecule.

The Guinier plots for F1 and F2 were linear and we determined R_g values of 5.07 ± 0.02 nm for F1 and 4.69 ± 0.02 nm for F2 (Figure S6A,B) from the slope and molecular mass (MM) values of 150 ± 10 kDa for F1 and 120 ± 10 kDa for F2 from the invariant V_c. The MM determined for F1 was closed to the theoretical mass of 165 kDa calculated for a 2N−2P complex and that for F2 was closed to the theoretical mass of 116 kDa calculated for a 1N−2P complex, also supporting the presence of complexes were each dimer of P bound two (F1) or one (F2) N molecules.

To further investigate this hypothesis, we generated two ensembles of physically accessible conformers of the $N_{\Delta 23}{}^0-P_{FL}$ complex, one ensemble with one N attached to each chain of P (2N−2P) and one with a single N attached to one of the P chain (1N−2P). We then used the GAJOE program for selecting sub-ensembles that reproduced the experimental SAXS curves at positions F1 and F2 in the chromatogram. The curve in the front part of the elution peak (F1) was better reproduced with 2N−2P models (reduced χ^2 = 1.9) than with 1N−2P models (reduced χ^2 = 2.7) (Figure 7C and Figure S6C). A representative sub-ensemble of four different conformers in different proportions is shown in Figure 7D. The curve in the tail part of the elution peak (F2) was better reproduced with 1N−2P models (reduced χ^2 = 1.9) than with 2N−2P models (reduced χ^2 = 19.9) (Figure 7F and Figure S6D). A representative sub-ensemble of three different conformers in different proportions is shown in Figure 7G. The selected ensembles also revealed that both forms of the complex (2N−2P and 1N−2P) were rather compact, with most conformers in the selected ensembles with R_g value smaller than the mean R_g of the initial distribution (Figure 7E,H). Surprisingly, the R_g values of the complexes at positions F1 and F2 were very similar to the R_g value of 4.9 ± 0.1 nm determined previously for RABV P alone [17,19]. Furthermore, we calculated that a 2N−2P complex, where the two N molecules were assembled side-by-side as in the N−RNA complex, would introduce features in the scattering intensity curve that were not observed in our experiments, providing evidence that both N^0 molecules remained independent within the complex.

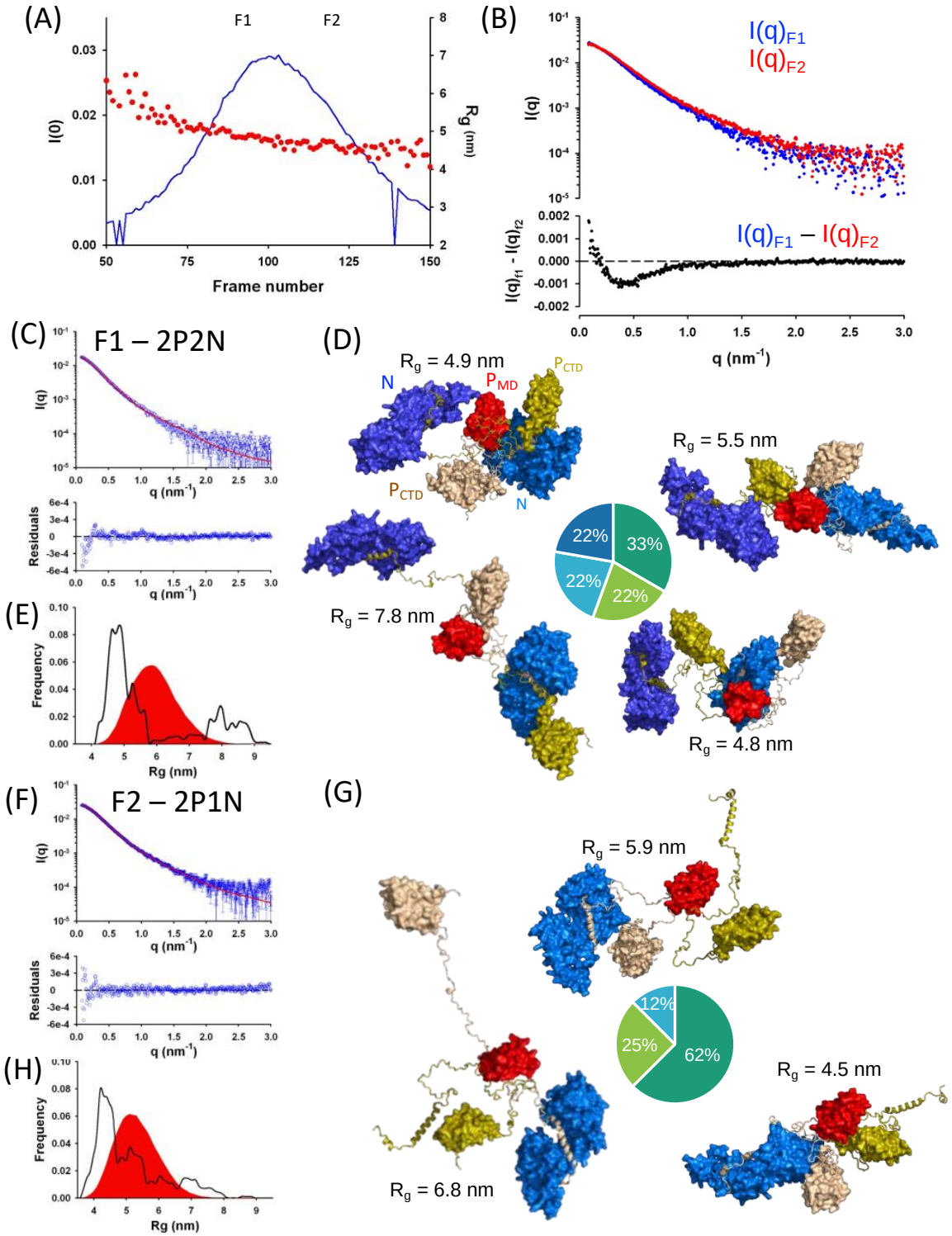

Figure 7. SEC-SAXS of $N_{\Delta 23}{}^{0}-P_{FL}$ complex. (**A**) SEC-SAXS elution profile and R_g across the elution peak; 50 µL of $N_{\Delta 23}{}^{0}-P_{FL}$ sample were injected onto a Superdex 200 column and monitored on-line by SAXS. The blue line shows the intensity at zero angle (I_0), which is proportional to both MM and concentration. The red dots indicate the values of the radius of gyration calculated from the Guinier approximation at the different time intervals. The shaded areas (labeled F1 and F2) show the frames that were averaged and used for analysis in the next panels. (**B**) Average SAXS profiles at two different positions in the SEC profile. The upper part shows the curves obtained by averaging the individual profiles recorded across the two fractions (F1 and F2) of the SEC elution peak shown in Panel (**A**). The lower part shows the difference scattering profile ($I(q)_{F1}$–$I(q)_{F2}$). (**C**) Conformational ensemble

modeling by the ensemble optimization method (EOM). The upper panel show the experimental F1scattering profile (blue circles) The red line shows back-calculated scattering curve for a selected ensemble of 4 2P−2N conformers shown in panel (**E**) (χ^2 = 1.95). The lower panel shows the plot of the residuals (blue circles). (**D**) Representative ensemble of conformers that reproduce the curve at position F1. The pie chart indicates the fraction of each conformer used in the calculated curve. The dimerization domain of P (P_{MD}) is shown in surface representation in red. The rest of the chains of P are shown in wheat and olive and the C-terminal domains (P_{CTD}) are shown in surface representation, while the intrinsically disordered regions are shown as cartoons. The N^0 molecules are shown in two shades of blue in surface representation. (**E**) R_g distribution. The red area shows the R_g distribution calculated for the initial ensemble of conformers (2P−2N), whereas the black line shows the R_g distribution of the selected ensembles that fit the experimental SAXS data (panel (**C**)). (**F**) Conformational ensemble modeling by the ensemble optimization method (EOM). The upper panel show the experimental F1 scattering profile (blue circles) The red line shows back-calculated scattering curve for a selected ensemble of three 2P−1N conformers shown in panel H (χ^2 = 1.99). The lower panel shows the plot of the residuals (blue circles). (**G**) Representative ensemble of conformers that reproduce the curve at position F2. The pie chart indicates the fraction of each conformer used in the calculated curve. The P protein is shown as in panel E and the N^0 molecule is shown in blue in surface representation. (**H**) R_g distribution. The red area shows the R_g distribution calculated for the initial ensemble of conformers (2P−1N), whereas the black line shows the R_g distribution of the selected ensembles that fit the experimental SAXS data (panel (**F**)).

4. Discussion

4.1. The N-Terminal Region of P Contains an Autonomous Chaperone Module (P_{CM})

As previously demonstrated for RABV and other viruses, the N-terminal intrinsically disordered region of P contains the chaperon module (P_{CM}) that autonomously maintains the N protein (N^0) in an RNA-free, monomeric, and soluble form [23,25–28,40,43,44,81]. As found for VSV and NiV, an isolated peptide comprising the RABV chaperone module is sufficient for reconstituting a heterodimeric complex with an armless N^0 (RABV $N_{\Delta 23}$ lacking the $N_{NT\text{-}ARM}$) [40,43]. However, if P_{CM} is sufficient to keep N^0 in a soluble, RNA-free monomeric form, it is possible that NC assembly in cells requires additional functional modules of P. Indeed, exogenous peptides encompassing P_{CM} from different viruses, including RABV, have been found to inhibit viral replication in cultured cells, whereas they were able to solubilize N in the cytoplasm [24,28,43,79,81], suggesting that their antiviral activity is based on the sequestration of the N protein in unproductive complexes or in a wrong cellular compartment. The C-terminal part of P attaches to the NC [21,80] and the N-terminal region binds to the L-RdRP [82], and either one of these interactions could be necessary to recruit the unassembled N^0 molecules near the site of RNA synthesis and/or to trigger NC assembly. The dimerization of P through its multimerization domain (P_{MD}) [22] was found to be unessential for the activity of the polymerase [83], although it was essential for formation of the membraneless compartments where viral replication occurs [9]. Full-length RABV P might thus be required to recruit N^0 within the biological condensates, in which viral RNA replication occurs. Finally, full-length RABV P protein forms dimers [19,22], where each chaperone module behaves independently of the other allowing the binding of one or two N^0 molecules per P dimer depending on the concentration of N present at the time of the reconstitution of the complex [42].

Surprisingly, the overall dimensions of the $N_{\Delta 23}{}^0-P_{FL}$ complex are not different from those of isolated P_{FL} in similar experimental conditions [17,19]. The scattering intensity at the zero angle at different positions in the chromatographic peak clearly indicates that one or two N^0 molecules were bound to the P dimer, but the radii of gyration are similar. The SAXS data also clearly show that the two bound N proteins are not preassembled within the N^0-P complex, but rather remain independent of each other. It is worthwhile noting that if we assume a generic partial specific volume of 0.73 $mL.gr^{-1}$ for both RABV P and N proteins [84] and consider the volume defined by the radius of gyration, the isolated P

protein occupies 17% of the volume (R_g = 4.9 nm, MM = 68.4 kDa), whereas the 2N−2P complex occupies 36% of the volume (R_g = 5.1 nm, MM = 165 kDa) (Figure 8). Similar calculations indicate that VSV P and NiV P occupy 13% and 7% of the volume defined by the radius of gyration, respectively [85,86]. The attachment of N molecules on P thus lead to an increase in density with almost no change in the dimensions of the protein complex as compared to P protein alone.

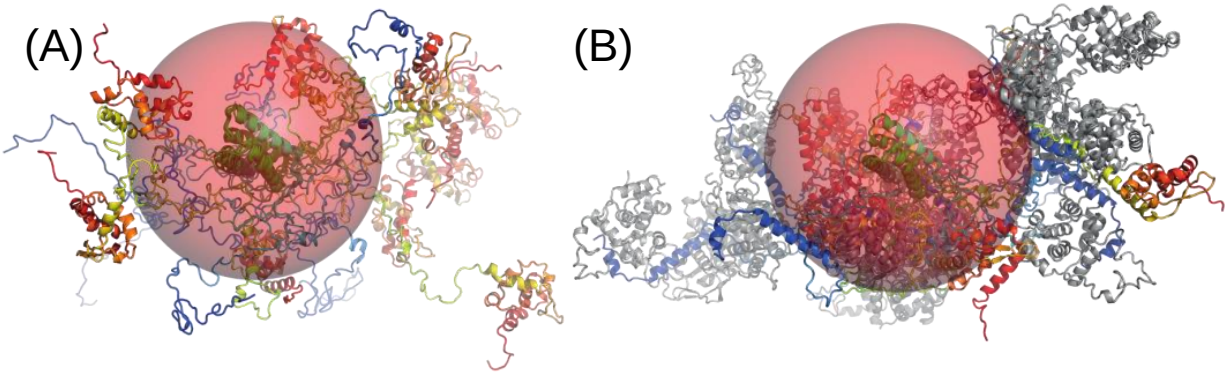

Figure 8. Comparison of ensemble models of P_{FL} and of the 2P−2N $N_{\Delta 23}{}^0$−P_{FL} complex. (**A**) Full-length P dimers. Representative ensemble of the five conformers of dimeric P_{FL} selected by GAJOE to reproduce SAXS data [17]. Each P protomer is colored from N-terminus (blue) to C-terminus (red) and the conformers were aligned by superimposing their individual dimerization domain. The semi-transparent red sphere has a radius of 4.9 nm corresponding to the radius of gyration determined from SAXS data. (**B**) 2P−2N $N_{\Delta 23}{}^0$−P_{FL} complex. Representative ensemble of the four conformers of the 2P−2N $N_{\Delta 23}{}^0$−P_{FL} complex selected by GAJOE to reproduce SAXS data (Figure 7E). Each P protomer is colored from N-terminus (blue) to C-terminus (red), the $N_{\Delta 23}$ molecules are shown in gray and the conformers were aligned by superimposing their individual dimerization domain. The semi-transparent red sphere has a radius of 5.1 nm corresponding to the radius of gyration determined from SAXS data.

Additionally, as found for other viruses [26,76], the truncated N protein ($N_{\Delta 23}$) was unable to assemble into NC even in the presence of RNA. The assembly of N into polymeric N−RNA complexes could only be reproduced in vitro with full-length N, demonstrating that the $N_{NT\text{-}ARM}$ is required for NC assembly (Bourhis and Jamin, unpublished data), although the mechanism of NC assembly and the detailed role played by the $N_{NT\text{-}ARM}$ and $N_{CT\text{-}ARM}$ remain poorly understood. We hypothesize that for an incoming N protein to incorporate into a nascent NC, the binding of its $N_{NT\text{-}ARM}$ onto the last already incorporated N subunit could be required either for anchoring the new N protein and allowing the assembly or for stabilizing the complex once it is formed or both.

4.2. Thermodynamic Control of NC Assembly

The interaction between $N_{\Delta 23}{}^0$ and P_{68} is strong with a dissociation constant in the low nanomolar range. The dissociation constant of the FAM-labeled peptide was below 1 nM and that of the unlabeled peptide measured by competition assay was 4 nM. This strong affinity is consistent with the large amount of surface area that becomes buried upon association with P_{68} (~2800 Å^2). The binding of the N-terminal part of P_{CM} (aa 4–15) onto $N_{\Delta 23}{}^0$ buries about 1200 Å^2, whereas the binding of the C-terminal part (aa 20–38) buries about 1600 Å^2. For protein–protein and protein−peptide complexes that bury more than 2000 Å^2 upon assembly, an average value of 4.0 cal/mol/Å^2 was estimated for the $\Delta G_{binding}$ [87], which for the burial of 2800 Å^2 would predict a $\Delta G_{binding}$ of 11.2 kcal.mol^{-1} and a dissociation constant of 6 nM (at 20 °C), which is in good agreement with our measurements. By comparison, a dissociation constant of 1.1 nM has been determined for the interaction between Ebola virus N^0 and a peptide of VP35 that contains the chaperone module, whose association buried a surface area of ~2400 Å^2. These figures place RABV

N^0-P complex among the protein–protein complexes that bury the largest surface area upon assembly and have a high binding affinity [87]. This strong interaction might be required for assembling the N^0-P complex in the infected cells at the earliest stages of the virus replication cycle, when only small amounts (low concentrations) of viral proteins are present. The nucleoprotein has a strong tendency to nonspecifically assemble onto cellular RNA and thus chaperoning by the P protein must occur as soon as N synthesis begins. However, the N^0-P complex is not the final step in the production line of the viral N protein. N^0 must be subsequently transferred to nascent viral RNA to form new viral NCs. To drive the assembly of NC, the N−RNA must probably be as stable or even more stable than the N^0-P complex and it is noteworthy that binding of the $N_{NT\text{-}ARM}$ of the N_{i-1} subunit buries ~1500 $Å^2$, while the binding of the $N_{CT\text{-}ARM}$ of the N_{i+1} subunit buries ~1700 $Å^2$, totaling ~3200 $Å^2$ per N subunit upon NC assembly, which translates into predicted $\Delta G_{binding}$ value of 12.8 $kcal.mol^{-1}$ and a dissociation constant of 0.3 nM (at 20 °C). The assembly of NC from the N^0-P complex could thus simply be driven by thermodynamics, the reaction evolving toward the most stable state of the system.

The molecular dynamics simulations revealed that the C-terminal part of P_{68} can transiently insert into the RNA binding groove of N perturbing/preventing the insertion of an RNA molecule and thus provide a putative additional contribution to the N^0 chaperon activity of P. This part of RABV P_{NTD} is rich in negatively charged amino acid and is thus suitable for inserting in the positively charged RNA binding groove. However, this part of the peptide is not visible in the crystal structure, indicating either that the interaction is weak and transient or that the peptide binds in different conformations in the RNA binding groove. In the meantime, the presence of this region destabilized the complex, possibly because the decrease in conformational entropy is not compensated by the formation of additional stabilizing interactions.

4.3. The Chaperone Module of P and the $N_{NT\text{-}ARM}$ of the Adjacent N Protein Controls the Opening and Closing Movement in N

A comparison with the other known structures of the N^0-P_{CM} core complex from different viruses belonging to different families of the *Mononegavirales* revealed many similarities in the complex architecture and in the mechanism of action of P_{CM}, but also some differences [29,40,43]. Globally, all these complexes have a contiguous segment of about 30–35 residues located at or near the N-terminal end of the P protein that binds mainly or solely to N_{CTD} and directly competes with the $N_{NT\text{-}ARM}$ and $N_{CT\text{-}ARM}$ of adjacent subunits. In the *Paramyxoviridae* (NiV, MeV, hPIV3, PIV5), *Pneumoviridae* (hMPV), and *Filoviridae* (EBOV and MARV), P_{CM} binds exclusively to N_{CTD}, in the same groove as the $N_{NT\text{-}ARM}$ subdomain of the N_{i-1} subunit in the polymeric N−RNA complex. In these complexes, the two protein segments (P_{CM} and $N_{NT\text{-}ARM}$) adopt opposite orientations although they bind in the same groove of the N protein [26,43–46]. In some cases, P_{CM} extends to the binding of the $N_{CT\text{-}ARM}$ of the N_{i+1} subunit and thus also directly interferes with its binding to the N_i subunit. In the *Rhabdoviridae* (VSV and RABV), one part of P_{CM} also binds to N_{CTD}, in the same groove as the $N_{NT\text{-}ARM}$ of the N_{i-1} subunit and the $N_{CT\text{-}ARM}$ of the N_{i+1} subunit, while the major part of the chaperone module forms an α-helix that binds at the interface between N_{NTD} and N_{CTD} (Figure 2). In the *Rhabdoviridae*, we confirm here that in RABV N^0-P complex, as previously shown in the VSV complex, the polypeptide chain of P_{CM} follows the same orientation than that of the $N_{NT\text{-}ARM}$ [40].

The first crystal structures of polymeric, circular N−RNA complexes revealed that the RNA molecule was embedded in its protein shell, with some bases facing the protein away from the solvent [39,41,88]. This directly suggested that the N protein must open and close, with the N_{NTD} and N_{CTD} moving relative to each other by rotation around a hinge located in their connecting region. This motion not only allows the insertion of the RNA molecule during the assembly of the NC but also provides access to all the nucleotide bases during the passage of the polymerase. Another striking difference observed by comparing the N^0-P_{CM} complexes of different viruses was the conformation of the N

protein. In the N^0-P_{CM} complexes of *Rhabdoviridae* (VSV and RABV), the N protein is locked in a closed conformation, causing the RNA binding site to be inaccessible to an incoming RNA molecule and suggesting that the protein must open to accommodate an RNA molecule. When initially observed in the VSV N^0-P_{CM} structure, the polymeric assembly of N raised the possibility that the closure of the RNA binding groove was a consequence of the polymerization [40]. In the new crystal structure of the RABV N^0-P_{CM} complex, the N protein is unassembled but is also trapped in a similar closed conformation, demonstrating that this state of the protein is independent of the polymerization and is an intrinsic property of the protein–protein complex as confirmed by our principal component analysis of MD simulations.

By contrast, in the N^0-P_{CM} complexes of *Paramyxoviridae*, *Pneumoviridae*, and *Filoviridae*, the N protein is blocked in an open conformation, in which the RNA-binding groove is too widely open to allow the formation of all the bonds evidenced between the protein and the RNA molecule within the N−RNA complexes [43–46,88–90]. In these cases, the protein is rather in a conformation ready to receive an incoming RNA molecule. This finding with a visual analysis of the NiV N^0-P_{CM} complex led us to propose that the binding of P_{CM} perturbs the dynamics of the N protein by linking together different subdomains that must move relative to one another when the protein switches between its open and close conformations [29,43]. The MD simulations reported in this paper provide evidence that the N protein has a natural tendency to undergo large conformational changes, where the N_{NTD} and N_{CTD} move relative to each other in different directions. The binding of P_{CM} but also of the $N_{NT\text{-}ARM}$ and RNA onto the surface of N restricts these movements in the protein. The binding of P_{CM} blocks N in open or close conformation depending on the viral family, preventing the positioning of an RNA molecule in the groove between N_{NTD} and N_{CTD}. The binding of RNA, $N_{NT\text{-}ARM}$, and also $N_{CT\text{-}ARM}$ blocks N in a close conformation that shields the RNA from the solvent, cellular proteins, and the viral polymerase. Different hypotheses have been put forward to explain the opening of the N protein, involving the intervention of P or L to trigger this process. The simulations suggest that the release of P_{CM} and of the $N_{NT\text{-}ARM}$ from the surface of N would allow the protein to undergo its intrinsic opening−closing movements and allow the insertion or release of the RNA molecule.

5. Conclusions

On the basis of the structural architecture and tentative mechanism of action of the N^0-P_{CM} complex, viruses can currently be separated into two different groups, with one group including the *Rhabdoriviridae* and another group including *Paramyxoviridae*, *Pneumoviridae*, and *Filoviridae*, which likely reflect evolutionary relationships. The formation of the N^0-P complex is likely an ancestral feature of the *Mononegavirales*, which has evolved by divergence in the different families, conserving the essential properties and showing that the great sequence−structure space diversity of the polypeptide chain allied with the high capacity to mutate of these viruses can create important variations around the same "theme". A striking example is the inversion of the direction of binding of the P_{CM} on N^0.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/v14122813/s1, Figure S1: Absence of N polymerization in the crystal of the $N_{\Delta 23}{}^0-P_{68}$ complex. Figure S2: Computational models extracted from MD simulations showing transient interactions between the C-terminal flexible part of P_{68} and $N_{\Delta 23}$. Figure S3: Assessment of the labeling of P_{42} by FAM using mass spectrometry. Figure S4: Unfolding of the nucleoprotein in molecular dynamics simulations revealed defects in the structural model. Figure S5: Movements in $N_{\Delta 23}$ alone, in $N_{\Delta 23}$ in complex with P_{68}, and in assembled full-length N in complex with RNA. Figure S6: SEC-SAXS of $N_{\Delta 23}{}^0-P_{FL}$ complex. Table S1: Polar contacts between N_{D23} and P_{68} in the crystal. Hydrogen bonds and salt bridges were identified with the server PISA.

Author Contributions: F.C.A.G., J.-M.B., C.L. and M.J. conceived the experiments; F.C.A.G., J.-M.B., C.M., A.B., D.D.V., S.V. and C.L. performed the experiments; F.C.A.G., J.-M.B., C.M., A.B., D.D.V., S.V., C.L. and M.J. analyzed and interpreted the data; J.-M.B., C.L. and M.J. wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported by grants from the French Agence Nationale de la Recherche to M.J. (ANR BSV8-2012—NNViPol) and from the Fond de la Recherche Médicale (FRM) to M.J. (grant "Equipe 2017" DEQ20170336754). D.D.V. was supported by a M2 scholarship from the Graduate School (EUR) CBH—Chemistry, Biology, and Health and S.V. was supported by a scholarship from the LABEX GRAL. F.C.A.G. was supported by a post-doctoral fellowship from the ANR program (ANR BSV8-2012—NNViPol). We acknowledge the European Synchrotron Radiation Facility for the allocation of beamtime on ID14-3, BM29, and ID23-1 and the Synchrotron SOLEIL for the allocation of beamtime on SWING. This work used the platforms of the Grenoble Instruct-ERIC center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-0005-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 8B8V.pdb and the re-refined structure has been deposited under the accession code 8FFR.pdb.

Acknowledgments: The authors thank A.A. Albertini (I2BC, Université Paris-Saclay) and R.W. Ruigrok for providing access to the original dataset of diffraction data collected for the N_{11}-RNA complex.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dietzschold, B.; Li, J.; Faber, M.; Schnell, M. Concepts in the Pathogenesis of Rabies. *Future Virol.* **2008**, *3*, 481–490. [CrossRef] [PubMed]
2. Rodriguez, L.L.; Rodrı, L.L. Emergence and Re-Emergence of Vesicular Stomatitis in the United States. *Virus Res.* **2002**, *85*, 211–219. [CrossRef] [PubMed]
3. Basak, S.; Mondal, A.; Polley, S.; Mukhopadhyay, S.; Chattopadhyay, D. Reviewing Chandipura: A Vesiculovirus in Human Epidemics. *Biosci. Rep.* **2007**, *27*, 275–298. [CrossRef] [PubMed]
4. Hogenhout, S.A.; Redinbaugh, M.G.; Ammar, E.D. Plant and Animal Rhabdovirus Host Range: A Bug's View. *Trends Microbiol.* **2003**, *11*, 264–271. [CrossRef]
5. Pringle, C.R. The Order Mononegavirales—Current Status. *Arch. Virol.* **1997**, *142*, 2321–2326.
6. Kuhn, J.H.; Adkins, S.; Alioto, D.; Alkhovsky, S.V.; Amarasinghe, G.K. Taxonomic Update for Phylum *Negarnaviricota* (*Riboviria: Orthornavirae*), Including the Large Orders *Bunyavirales* and *Mononegavirales*. *Arch. Virol.* **2020**, *165*, 3023–3072. [CrossRef]
7. Ruigrok, R.W.H.; Crepin, T.; Kolakofsky, D. Nucleoproteins and Nucleocapsids of Negative-Strand RNA Viruses. *Curr. Opin. Microbiol.* **2011**, *14*, 504–510. [CrossRef]
8. Iseni, F.; Barge, A.; Baudin, F.; Blondel, D.; Ruigrok, R.W. Characterization of Rabies Virus Nucleocapsids and Recombinant Nucleocapsid-like Structures. *J. Gen. Virol.* **1998**, *79*, 2909–2919. [CrossRef]
9. Nikolic, J.; Le Bars, R.; Lama, Z.; Scrima, N.; Lagaudrière-Gesbert, C.; Gaudin, Y.; Blondel, D. Negri Bodies Are Viral Factories with Properties of Liquid Organelles. *Nat. Commun.* **2017**, *8*, 58. [CrossRef]
10. Morin, B.; Rahmeh, A.A.; Whelan, S.P. Mechanism of RNA Synthesis Initiation by the Vesicular Stomatitis Virus Polymerase. *EMBO J.* **2012**, *31*, 1320–1329. [CrossRef]
11. Morin, B.; Liang, B.; Gardner, E.; Ross, R.A.; Whelan, S.P.J. An In Vitro RNA Synthesis Assay for Rabies Virus Defines Ribonucleoprotein Interactions Critical for Polymerase Activity. *J. Virol.* **2017**, *91*, e01508-16. [CrossRef]
12. Arnheiter, H.; Davis, N.L.; Wertz, G.; Schubert, M.; Lazzarini, R.A. Role of the Nucleocapsid Protein in Regulating Vesicular Stomatitis Virus RNA Synthesis. *Cell* **1985**, *41*, 259–267. [CrossRef]
13. Masters, P.S.; Banerjee, A.K. Complex Formation with Vesicular Stomatitis Virus Phosphoprotein NS Prevents Binding of Nucleocapsid Protein N to Nonspecific RNA. *J. Virol.* **1988**, *62*, 2658–2664. [CrossRef]
14. Howard, M.; Wertz, G. Vesicular Stomatitis Virus RNA Replication: A Role for the NS Protein. *J. Gen. Virol.* **1989**, *70*, 2683–2694. [CrossRef]
15. Curran, J.; Latorre, P.; Kolakofsky, D. Translational Gymnastics on the Sendai Virus P/C MRNA. In *Seminars in VIROLOGY 8*; Academic Press: Cambridge, MA, USA, 1998; pp. 351–357.

16. Leyrat, C.; Ribeiro, E.A.; Gérard, F.C.; Ivanov, I.; Ruigrok, R.W.H.; Jamin, M. Structure, Interactions with Host-Cell and Functions of Rhabdovirus Phosphoprotein. *Future Virol.* **2011**, *6*, 465–481. [CrossRef]
17. Jespersen, N.E.; Leyrat, C.; Gérard, F.C.; Bourhis, J.M.; Blondel, D.; Jamin, M.; Barbar, E. The LC8-RavP Ensemble Structure Evinces a Role for LC8 in Regulating Lyssavirus Polymerase Functionality. *J. Mol. Biol.* **2019**, *431*, 4959–4977. [CrossRef]
18. Blondel, D.; Maarifi, G.; Nisole, S.; Chelbi-Alix, M.K. Resistance to Rhabdoviridae Infection and Subversion of Antiviral Responses. *Viruses* **2015**, *7*, 3675–3702. [CrossRef]
19. Gerard, F.C.A.; Ribeiro, E.D.A.; Albertini, A.A.V.; Gutsche, I.; Zaccai, G.; Ruigrok, R.W.H.; Jamin, M. Unphosphorylated Rhabdoviridae Phosphoproteins Form Elongated Dimers in Solution. *Biochemistry* **2007**, *46*, 10328–10338. [CrossRef]
20. Gerard, F.C.A.; Ribeiro, E.d.A.; Leyrat, C.; Ivanov, I.; Blondel, D.; Longhi, S.; Ruigrok, R.W.H.; Jamin, M. Modular Organization of Rabies Virus Phosphoprotein. *J. Mol. Biol.* **2009**, *388*, 978–996. [CrossRef]
21. Mavrakis, M.; McCarthy, A.A.; Roche, S.; Blondel, D.; Ruigrok, R.W. Structure and Function of the C-Terminal Domain of the Polymerase Cofactor of Rabies Virus. *J. Mol. Biol.* **2004**, *343*, 819–831. [CrossRef]
22. Ivanov, I.; Crépin, T.; Jamin, M.; Ruigrok, R.W.H. Structure of the Dimerization Domain of the Rabies Virus Phosphoprotein. *J. Virol.* **2010**, *84*, 3707–3710. [CrossRef] [PubMed]
23. Mavrakis, M.; Mehouas, S.; Real, E.; Iseni, F.; Blondel, D.; Tordo, N.; Ruigrok, R.W. Rabies Virus Chaperone: Identification of the Phosphoprotein Peptide That Keeps Nucleoprotein Soluble and Free from Non-Specific RNA. *Virology* **2006**, *349*, 422–429. [CrossRef] [PubMed]
24. Castel, G.; Chteoui, M.; Caignard, G.; Prehaud, C.; Mehouas, S.; Real, E.; Jallet, C.; Jacob, Y.; Ruigrok, R.W.H.; Tordo, N. Peptides That Mimic the Amino-Terminal End of the Rabies Virus Phosphoprotein Have Antiviral Activity. *J. Virol.* **2009**, *83*, 10808–10820. [CrossRef]
25. Curran, J.; Marq, J.B.; Kolakofsky, D. An N-Terminal Domain of the Sendai Paramyxovirus P Protein Acts as a Chaperone for the NP Protein during the Nascent Chain Assembly Step of Genome Replication. *J. Virol.* **1995**, *69*, 849. [CrossRef] [PubMed]
26. Kirchdoerfer, R.N.; Abelson, D.M.; Li, S.; Wood, M.R.; Saphire, E.O. Assembly of the Ebola Virus Nucleoprotein from a Chaperoned VP35 Complex. *Cell Rep.* **2015**, *12*, 140–149. [CrossRef] [PubMed]
27. Chen, M.; Ogino, T.; Banerjee, A.K. Interaction of Vesicular Stomatitis Virus P and N Proteins: Identification of Two Overlapping Domains at the N Terminus of P That Are Involved in N0-P Complex Formation and Encapsidation of Viral Genome RNA. *J. Virol.* **2007**, *81*, 13478–13485. [CrossRef]
28. Leung, D.W.; Borek, D.; Luthra, P.; Binning, J.M.; Anantpadma, M.; Liu, G.; Harvey, I.B.; Su, Z.; Endlich-Frazier, A.; Pan, J.; et al. An Intrinsically Disordered Peptide from Ebola Virus VP35 Controls Viral RNA Synthesis by Modulating Nucleoprotein-RNA Interactions. *Cell Rep.* **2015**, *11*, 376–389. [CrossRef]
29. Jamin, M.; Yabukarski, F. Non-Segmented Negative-Sense RNA Viruses—Structural Data Brings New Insights into Nucleocapsid Assembly. *Adv. Virus Res.* **2016**, *97*, 143–185.
30. Fouquet, B.; Nikolic, J.; Larrous, F.; Bourhy, H.; Wirblich, C.; Lagaudrière-Gesbert, C.; Blondel, D. Focal Adhesion Kinase Is Involved in Rabies Virus Infection through Its Interaction with Viral Phosphoprotein P. *J. Virol.* **2014**, *89*, 1640–1651. [CrossRef]
31. Li, Y.; Dong, W.; Shi, Y.; Deng, F.; Chen, X.; Wan, C.; Zhou, M.; Zhao, L.; Fu, Z.F.; Peng, G. Rabies Virus Phosphoprotein Interacts with Ribosomal Protein L9 and Affects Rabies Virus Replication. *Virology* **2016**, *488*, 216–224. [CrossRef]
32. Kammouni, W.; Wood, H.; Saleh, A.; Appolinario, C.M.; Fernyhough, P.; Jackson, A.C. Rabies Virus Phosphoprotein Interacts with Mitochondrial Complex I and Induces Mitochondrial Dysfunction and Oxidative Stress. *J. Neurovirol.* **2015**, *21*, 370–382. [CrossRef]
33. Chelbi-Alix, M.K.; Vidy, A.; El Bougrini, J.; Blondel, D. Rabies Viral Mechanisms to Escape the IFN System: The Viral Protein P Interferes with IRF-3, Stat1, and PML Nuclear Bodies. *J. Interferon Cytokine Res.* **2006**, *26*, 271–280. [CrossRef]
34. Vidy, A.; Chelbi-Alix, M.; Blondel, D. Rabies Virus P Protein Interacts with STAT1 and Inhibits Interferon Signal Transduction Pathways. *J. Virol.* **2005**, *79*, 14411–14420. [CrossRef]
35. Blondel, D.; Regad, T.; Poisson, N.; Pavie, B.; Harper, F.; Pandolfi, P.P.; De Thé, H.; Chelbi-Alix, M.K. Rabies Virus P and Small P Products Interact Directly with PML and Reorganize PML Nuclear Bodies. *Oncogene* **2002**, *21*, 7957–7970. [CrossRef]
36. Oksayan, S.; Wiltzer, L.; Rowe, C.L.; Blondel, D.; Jans, D.A.; Moseley, G.W. A Novel Nuclear Trafficking Module Regulates the Nucleocytoplasmic Localization of the Rabies Virus Interferon Antagonist, P Protein. *J. Biol. Chem.* **2012**, *287*, 28112–28121. [CrossRef]
37. Liu, J.; Wang, H.; Gu, J.; Deng, T.; Yuan, Z.; Hu, B.; Xu, Y.; Yan, Y.; Zan, J.; Liao, M.; et al. BECN1-Dependent CASP2 Incomplete Autophagy Induction by Binding to Rabies Virus Phosphoprotein. *Autophagy* **2017**, *13*, 739–753. [CrossRef]
38. Xu, Y.; Liu, F.; Liu, J.; Wang, D.; Yan, Y.; Ji, S.; Zan, J.; Zhou, J. The Co-Chaperone Cdc37 Regulates the Rabies Virus Phosphoprotein Stability by Targeting to Hsp90AA1 Machinery. *Sci. Rep.* **2016**, *6*, 27123. [CrossRef]
39. Albertini, A.A.V.; Wernimont, A.K.; Muziol, T.; Ravelli, R.B.G.; Clapier, C.R.; Schoehn, G.; Weissenhorn, W.; Ruigrok, R.W.H. Crystal Structure of the Rabies Virus Nucleoprotein-RNA Complex. *Science* **2006**, *313*, 360–363. [CrossRef]
40. Leyrat, C.; Yabukarski, F.; Tarbouriech, N.; Ribeiro, E.A., Jr.; Jensen, M.R.; Blackledge, M.; Ruigrok, R.W.; Jamin, M. Structure of the Vesicular Stomatitis Virus N-P Complex. *PLoS Pathog.* **2011**, *7*, e1002248. [CrossRef]
41. Green, T.J.; Zhang, X.; Wertz, G.W.; Luo, M. Structure of the Vesicular Stomatitis Virus Nucleoprotein-RNA Complex. *Science* **2006**, *313*, 357–360. [CrossRef]

42. Yabukarski, F.; Leyrat, C.; Martinez, N.; Communie, G.; Ivanov, I.; Ribeiro, E.A.J.; Buisson, M.; Gérard, F.C.; Bourhis, J.-M.; Jensen, M.R.; et al. Ensemble Structure of the Highly Flexible Complex Formed between Vesicular Stomatitis Virus Unassembled Nucleoprotein and Its Phosphoprotein Chaperone. *J. Mol. Biol.* **2016**, *428*, 2671–2694. [CrossRef] [PubMed]
43. Yabukarski, F.; Lawrence, P.; Tarbouriech, N.; Bourhis, J.M.; Delaforge, E.; Jensen, M.R.; Ruigrok, R.W.; Blackledge, M.; Volchkov, V.; Jamin, M. Structure of Nipah Virus Unassembled Nucleoprotein in Complex with Its Viral Chaperone. *Nat. Struct. Mol. Biol.* **2014**, *21*, 754–759. [CrossRef] [PubMed]
44. Guryanov, S.G.; Liljeroos, L.; Kasaragod, P.; Kajander, T.; Butcher, S.J. Crystal Structure of the Measles Virus Nucleoprotein Core in Complex with an N-Terminal Region of Phosphoprotein. *J. Virol.* **2015**, *90*, 2849–2857. [CrossRef] [PubMed]
45. Aggarwal, M.; Leser, G.P.; Kors, C.A.; Lamb, R.A. Structure of the Paramyxovirus PIV5 Nucleoprotein in Complex with an Amino-Terminal Peptide of the Phosphoprotein. *J. Virol.* **2017**, *92*, e1304-17. [CrossRef]
46. Renner, M.; Bertinelli, M.; Leyrat, C.; Paesen, G.C.; Saraiva de Oliveira, L.F.; Huiskonen, J.T.; Grimes, J.M. Nucleocapsid Assembly in Pneumoviruses Is Regulated by Conformational Switching of the N Protein. *eLife* **2016**, *5*, e12627. [CrossRef] [PubMed]
47. Liu, B.; Dong, S.; Li, G.; Wang, W.; Liu, X. Structural Insight into Nucleoprotein Conformation Change Chaperoned by VP35 Peptide of Marburg Virus. *J. Virol.* **2017**, *91*, e00825-17. [CrossRef]
48. The UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, D480–D489. [CrossRef]
49. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539–544. [CrossRef]
50. Krissinel, E.; Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **2007**, *372*, 774–797. [CrossRef]
51. Krissinel, E.; Henrick, K. Secondary-Structure Matching (SSM), a New Tool for Fast Protein Structure Alignment in Three Dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2256–2268. [CrossRef]
52. Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L. The FoldX Web Server: An Online Force Field. *Nucleic Acids Res.* **2005**, *33*, W382–W388. [CrossRef]
53. Wyatt, P.J. Submicrometer Particle Sizing by Multiangle Light Scattering Following Fractionation. *J. Colloid Interface Sci.* **1998**, *197*, 9–20. [CrossRef]
54. Uversky, V.N. Use of Fast Protein Size-Exclusion Liquid Chromatography to Study the Unfolding of Proteins Which Denature through the Molten Globule. *Biochemistry* **1993**, *32*, 13288–13298. [CrossRef]
55. Kuzmic, P. DynaFit—A Software Package for Enzymology. *Methods Enzymol.* **2009**, *467*, 247–280. [CrossRef]
56. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132. [CrossRef]
57. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.W.; McCoy, A.; et al. Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235–242. [CrossRef]
58. McCoy, A.J.; Grosse-Kunstleve, R.W.; Adams, P.D.; Winn, M.D.; Storoni, L.C.; Read, R.J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674. [CrossRef]
59. Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132. [CrossRef]
60. Murshudov, G.N.; Skubak, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.A.; Winn, M.D.; Long, F.; Vagin, A.A. REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355–367. [CrossRef]
61. Smart, O.S.; Womack, T.O.; Flensburg, C.; Keller, P.; Paciorek, W.; Sharff, A.; Vonrhein, C.; Bricogne, G. Exploiting Structure Similarity in Refinement: Automated NCS and Target-Structure Restraints in BUSTER. *Acta Crystallogr. D Biol. Crystallogr.* **2012**, *68*, 368–380. [CrossRef]
62. Laskowski, R.A.; Macarthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Cryst.* **1993**, *26*, 283–291. [CrossRef]
63. Croll, T.I. ISOLDE: A Physically Realistic Environment for Model Building into Low-Resolution Electron-Density Maps Research Papers. *Acta Crystallogr. D Biol. Crystallogr.* **2018**, *74*, 519–530. [CrossRef] [PubMed]
64. Adams, P.D.; Afonine, P.V.; Bunkoczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213–221. [CrossRef] [PubMed]
65. David, G.; Pérez, J. Combined Sampler Robot and High-Performance Liquid Chromatography: A Fully Automated System for Biological Small-Angle X-Ray Scattering Experiments at the Synchrotron SOLEIL SWING Beamline. *J. Appl. Crystallogr.* **2009**, *42*, 892–900. [CrossRef]
66. Konarev, P.; Petoukhov, M.; Volchkov, V.; Svergun, D.I. ATSAS 2.1, a Program Package for Small-Angle Scattering Data Analysis. *J. Appl. Cryst.* **2006**, *39*, 277–286. [CrossRef]
67. Rambo, R.P.; Tainer, J. A Accurate Assessment of Mass, Models and Resolution by Small-Angle Scattering. *Nature* **2013**, *496*, 477–481. [CrossRef]

68. Ozenne, V.; Bauer, F.; Salmon, L.; Huang, J.R.; Jensen, M.R.; Segard, S.; Bernado, P.; Charavay, C.; Blackledge, M. Flexible-Meccano: A Tool for the Generation of Explicit Ensemble Descriptions of Intrinsically Disordered Proteins and Their Associated Experimental Observables. *Bioinformatics* **2012**, *28*, 1463–1470. [CrossRef]
69. Krivov, G.G.; Shapovalov, M.V.; Dunbrack, R.L., Jr. Improved Prediction of Protein Side-Chain Conformations with SCWRL4. *Proteins Struct. Funct. Bioinform.* **2009**, *77*, 778–795. [CrossRef]
70. Bernadó, P.; Mylonas, E.; Petoukhov, M.V.; Blackledge, M.; Svergun, D.I. Structural Characterization of Flexible Proteins Using Small-Angle X-Ray Scattering. *J. Am. Chem. Soc.* **2007**, *129*, 5656–5664. [CrossRef]
71. Moore, B.L.; Kelley, L.A.; Barber, J.; Murray, J.W.; MacDonald, J.T. High-Quality Protein Backbone Reconstruction from Alpha Carbons Using Gaussian Mixture Models. *J. Comput. Chem.* **2013**, *34*, 1881–1889. [CrossRef]
72. Svergun, D.; Barberato, C.; Koch, M.H. CRYSOL—A Program to Evaluate X-Ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Cryst.* **1995**, *28*, 768–773. [CrossRef]
73. Páll, S.; Zhmurov, A.; Bauer, P.; Abraham, M.; Lundborg, M.; Gray, A.; Hess, B. Heterogeneous Parallelization and Acceleration of Molecular Dynamics Simulations in GROMACS 1). *J. Phys. Chem.* **2020**, *153*, 134110. [CrossRef]
74. Best, R.B.; Zheng, W.; Mittal, J. Balanced Protein—Water Interactions Improve Properties of Disordered Proteins and Non-Speci Fi c Protein Association. *J. Chem. Theory Comput.* **2014**, *10*, 5113–5124. [CrossRef]
75. Leyrat, C.; Jensen, M.R.; Ribeiro, E.A.; Gérard, F.C.A.; Ruigrok, R.W.H.; Blackledge, M.; Jamin, M. The N0-Binding Region of the Vesicular Stomatitis Virus Phosphoprotein Is Globally Disordered but Contains Transient α-Helices. *Protein Sci.* **2011**, *20*, 542–556. [CrossRef]
76. Milles, S.; Jensen, M.R.; Communie, G.; Maurin, D.; Schoehn, G.; Ruigrok, R.W.H.; Blackledge, M. Self-Assembly of Measles Virus Nucleocapsid-like Particles: Kinetics and RNA Sequence Dependence. *Angew. Chem. Int. Ed.* **2016**, *55*, 9356–9360. [CrossRef]
77. Petoukhov, M.V.; Franke, D.; Shkumatov, A.V.; Tria, G.; Kikhney, A.G.; Gajda, M.; Gorba, C.; Mertens, H.D.T.; Konarev, P.V.; Svergun, D.I. New Developments in the ATSAS Program Package for Small-Angle Scattering Data Analysis. *J. Appl. Crystallogr.* **2012**, *45*, 342–350. [CrossRef]
78. Fersht, A.R. *Structure and Mechanism in Protein Science—A Guide to Enzyme Catalysis and Protein Folding*; Series in Structural Biology; World Scientific: Singapore, 2017; Volume 9.
79. Dong, X.; Wang, X.; Xie, M.; Wu, W.; Chen, Z. Structural Basis of Human Parainfluenza Virus 3 Unassembled Nucleoprotein in Complex with Its Viral Chaperone. *J. Virol.* **2021**, *96*, e0164821. [CrossRef]
80. Ribeiro, E.A.; Leyrat, C.; Gérard, F.C.; Albertini, A.A.; Falk, C.; Ruigrok, R.W.; Jamin, M. Binding of Rabies Virus Polymerase Cofactor to Recombinant Circular Nucleoprotein-RNA Complexes. *J. Mol. Biol.* **2009**, *394*, 558–575. [CrossRef]
81. Galloux, M.; Sourimant, J.; Richard, C.; England, P.; Moudjou, M.; Aumont-nicaise, M.; Fix, J. Identification and Characterization of the Binding Site of the Respiratory Syncytial Virus Phosphoprotein to RNA-Free Nucleoprotein. *J. Virol.* **2015**, *89*, 3484–3496. [CrossRef]
82. Horwitz, J.A.; Jenni, S.; Harrison, S.C.; Whelan, S.P.J. Structure of a Rabies Virus Polymerase Complex from Electron Cryo-Microscopy. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 2099–2107. [CrossRef]
83. Jacob, Y.; Real, E.; Tordo, N. Functional Interaction Map of Lyssavirus Phosphoprotein: Identification of the Minimal Transcription Domains. *J. Virol.* **2001**, *75*, 9613–9622. [CrossRef] [PubMed]
84. Durchschlag, H. Specific Volumes of Biological Macromolecules and Some Other Molecules of Biological Interest. In *Thermodynamic Data for Biochemistry and Biotechnology*; Springer-Verlag: Berlin/Heidelberg, Germany, 1986; pp. 45–128. [CrossRef]
85. Leyrat, C.; Schneider, R.; Ribeiro, E.A.; Yabukarski, F.; Yao, M.; Gérard, F.C.A.; Jensen, M.R.; Ruigrok, R.W.H.; Blackledge, M.; Jamin, M. Ensemble Structure of the Modular and Flexible Full-Length Vesicular Stomatitis Virus Phosphoprotein. *J. Mol. Biol.* **2012**, *423*, 182–197. [CrossRef] [PubMed]
86. Jensen, M.R.; Yabukarski, F.; Communie, G.; Condamine, E.; Mas, C.; Volchkova, V.; Tarbouriech, N.; Bourhis, J.M.; Volchkov, V.; Blackledge, M.; et al. Structural Description of the Nipah Virus Phosphoprotein and Its Interaction with STAT1. *Biophys. J.* **2020**, *118*, 2470–2488. [CrossRef] [PubMed]
87. Chen, J.; Sawyer, N.; Regan, L. Protein—Protein Interactions: General Trends in the Relationship between Binding Affinity and Interfacial Buried Surface Area. *Protein Sci.* **2013**, *22*, 510–515. [CrossRef]
88. Tawar, R.G.; Duquerroy, S.; Vonrhein, C.; Varela, P.F.; Damier-Piolle, L.; Castagne, N.; MacLellan, K.; Bedouelle, H.; Bricogne, G.; Bhella, D.; et al. Crystal Structure of a Nucleocapsid-like Nucleoprotein-RNA Complex of Respiratory Syncytial Virus. *Science* **2009**, *326*, 1279–1283. [CrossRef]
89. Gutsche, I.; Desfosses, A.; Effantin, G.; Ling, W.L.; Haupt, M.; Ruigrok, R.W.H.; Sachse, C.; Schoehn, G. Near-Atomic Cryo-EM Structure of the Helical Measles Virus Nucleocapsid. *Science* **2015**, *348*, 704–708. [CrossRef]
90. Ker, D.S.; Jenkins, H.T.; Greive, S.J.; Antson, A.A. CryoEM Structure of the Nipah Virus Nucleocapsid Assembly. *PLoS Pathog.* **2021**, *17*, 1–19. [CrossRef]