

## Article

# Structural Analysis of the Menangle Virus P Protein Reveals a Soft Boundary between Ordered and Disordered Regions

Melissa N. Webby<sup>1,2</sup>, Nicole Herr<sup>1</sup>, Esther M. M. Bulloch<sup>1</sup>, Michael Schmitz<sup>3</sup>, Jeremy R. Keown<sup>1,4</sup> , David C. Goldstone<sup>1</sup>  and Richard L. Kingston<sup>1,\*</sup>

- <sup>1</sup> School of Biological Sciences, University of Auckland, Auckland 1010, New Zealand; melissa.webby@bioch.ox.ac.uk (M.N.W.); nicole.herr@auckland.ac.nz (N.H.); e.bulloch@auckland.ac.nz (E.M.M.B.); jeremy@strubi.ox.ac.uk (J.R.K.); d.goldstone@auckland.ac.nz (D.C.G.)
- <sup>2</sup> Department of Biochemistry, University of Oxford, Oxford OX1 2JD, UK
- <sup>3</sup> School of Chemical Sciences, University of Auckland, Auckland 1010, New Zealand; m.schmitz@auckland.ac.nz
- <sup>4</sup> Division of Structural Biology, University of Oxford, Oxford OX1 2JD, UK
- \* Correspondence: rl.kingston@auckland.ac.nz

**Abstract:** The paramyxoviral phosphoprotein (P protein) is the non-catalytic subunit of the viral RNA polymerase, and coordinates many of the molecular interactions required for RNA synthesis. All paramyxoviral P proteins oligomerize via a centrally located coiled-coil that is connected to a downstream binding domain by a dynamic linker. The C-terminal region of the P protein coordinates interactions between the catalytic subunit of the polymerase, and the viral nucleocapsid housing the genomic RNA. The inherent flexibility of the linker is believed to facilitate polymerase translocation. Here we report biophysical and structural characterization of the C-terminal region of the P protein from Menangle virus (MenV), a bat-borne paramyxovirus with zoonotic potential. The MenV P protein is tetrameric but can dissociate into dimers at sub-micromolar protein concentrations. The linker is globally disordered and can be modeled effectively as a worm-like chain. However, NMR analysis suggests very weak local preferences for alpha-helical and extended beta conformation exist within the linker. At the interface between the disordered linker and the structured C-terminal binding domain, a gradual disorder-to-order transition occurs, with X-ray crystallographic analysis revealing a dynamic interfacial structure that wraps the surface of the binding domain.

**Keywords:** Paramyxoviruses; RNA-dependent RNA polymerase; pararubulaviruses; intrinsically disordered proteins; protein self-association; conformational exchange; small-angle X-ray scattering; analytical ultracentrifugation; solution NMR spectroscopy; X-ray crystallography



**Citation:** Webby, M.N.; Herr, N.; Bulloch, E.M.M.; Schmitz, M.; Keown, J.R.; Goldstone, D.C.; Kingston, R.L. Structural Analysis of the Menangle Virus P Protein Reveals a Soft Boundary between Ordered and Disordered Regions. *Viruses* **2021**, *13*, 1737. <https://doi.org/10.3390/v13091737>

Academic Editors: Vernon Ward and John A. Taylor

Received: 28 July 2021

Accepted: 28 August 2021

Published: 31 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The Paramyxoviruses are a large family of negative-sense, non-segmented, single-stranded RNA viruses. Paramyxoviruses enter the host via the respiratory tract; infect most vertebrate species; and cause serious disease in both humans and animals. Among the well-established human pathogens in this family are the measles virus, mumps virus, and the four pathogenic human parainfluenza viruses (HPIV1–HPIV4).

The widespread distribution of paramyxoviruses creates the potential for cross-species transmission. Bats are now recognized as a particularly important animal reservoir for paramyxoviruses [1,2]. Pteropodid bats are the natural host for Hendra and Nipah viruses, which have spilled over into human and animal populations numerous times in the past three decades [3,4]. While Hendra and Nipah viruses cause no apparent disease in bats, they are highly pathogenic when transmitted to other animal populations, where they can cause fatal encephalitis [5]. Bats host many other paramyxoviruses with zoonotic potential. Among these is Menangle virus (MenV) [6,7], a pararubulavirus that was first identified following a novel disease outbreak in domesticated pigs [8].

All paramyxoviruses require three proteins to facilitate virally-directed RNA synthesis; one protein to package, protect and organize the RNA genome, and two proteins that together form the viral RNA-dependent RNA polymerase (RdRp). The genome is packaged by the nucleocapsid protein (N protein), resulting in formation of a helical protein-RNA complex termed the nucleocapsid [9–14]. The viral RdRp [15–18] performs transcription and genome replication sequentially, using the nucleocapsid as a template [19]. The large protein (L protein) is the catalytic subunit of the RdRp, facilitating synthesis and chemical modification of RNA, while the phosphoprotein (P protein) is the non-catalytic subunit of the RdRp, enabling its processive translocation along the viral nucleocapsid, and modulating the interactions between the nucleocapsid protein and RNA.

The P protein effectively functions as a flexible hub that coordinates binding activities that are integral to polymerase function. In all paramyxoviruses, the phosphoprotein self-associates via a centrally located coiled-coil, while the regions N-terminal and C-terminal to the coiled-coil are predominantly disordered, and involved in binding both viral and cellular proteins. Structural analysis of the coiled-coil has established that it exists as a tetramer [16,20–23]. The four helices contributing to the tetramer generally run parallel, giving the coiled-coil four-fold rotational symmetry along its central axis [16,20,21,23]. In the case of the Mumps virus P protein, however, the coiled-coil is an antiparallel “dimer-of-dimers” with two-fold rotational symmetry perpendicular to the central axis [22,24]. As a result, the chain termini of the Mumps P protein are not all juxtaposed. Hence, some variation in the exact mode of oligomerization can clearly be tolerated [16,22,24], though the functional implications are presently unknown. To date, only a single structure of an intact paramyxoviral P/L complex has been determined, from parainfluenza virus 5 [16]. In this case, the C-terminus of the coiled-coil is the point of attachment with the L protein, and the body of the coiled-coil projects out a considerable distance from the catalytic subunit [16].

It is the C-terminal region of the phosphoprotein, encompassing the coiled-coil and downstream sequences, that allows the polymerase to engage with the nucleocapsid, and is critical for processive RNA synthesis. Following the coiled-coil, there is a flexible linker that varies widely in length and composition between paramyxoviruses [25], and then a structured binding domain (the foot domain or X domain) that interacts with both the L protein [16,26] and the RNA-bound N protein [27–31]. In all cases the binding domain forms a compact bundle of three alpha-helices [30,32–35], though in the rubulaviruses, the binding domains can be unstable, and exist natively in a partially structured state [34–36]. Different faces of the three-helix bundle are implicated in binding of L and N [16,26]. The C-terminal binding domains are expected to be spatially adjacent, as a result of tetramerization of the P protein, and the tethering activity of the linker, and likely bind the RNA-associated N protein in a multivalent fashion.

It is not known how the various binding events mediated by the C-terminal region of the P protein are coordinated to facilitate polymerase translocation, the threading of the RNA genome through the active site of the polymerase, and the transient displacement of the nucleocapsid protein from the genome. However, it seems likely that extension and compaction of the dynamic linker must occur to facilitate RdRp movement. Polymerase translocation must involve the repeated attachment and release of the binding domain to the N protein subunits of the nucleocapsid, and possibly also the L protein.

Here we report structural and biophysical analysis of the C-terminal region of the MenV P protein. MenV belongs to the pararubulavirus genus, for which there has been little prior characterization of the P protein. We investigate the oligomerization state of the MenV P protein in solution, as well as the structural characteristics of the flexible linker that connects the coiled-coil and the C-terminal binding domain. The results obtained broaden our understanding of the organization of the paramyxoviral P protein, an essential component of the paramyxoviral replication machinery.

## 2. Materials and Methods

### 2.1. Protein Production and Purification

DNA encoding the Menangle virus P protein (UniProt ID Q91MK1; Genbank Accession AF326114), codon-optimized for expression in *Escherichia coli*, was commercially synthesized (Invitrogen GeneArt, Thermo Fisher Scientific). Bacterial expression plasmids enabling the production of six truncated MenV P proteins ('CC-L-BD'=P<sub>209-388</sub>; 'CC-L'=P<sub>209-336</sub>; 'CC'=P<sub>209-271</sub>; 'L-BD'=P<sub>267-388</sub>; 'L'=P<sub>267-328</sub>; 'BD'=P<sub>337-388</sub>) were generated using standard molecular biology procedures (Table S1). These constructs appended protease-cleavable N-terminal affinity tags to the P protein to facilitate purification. Additional expression plasmids were generated for the crystallographic analysis (Table S2). These constructs enabled production of the final part of the MenV P linker and the binding domain, fused to the C-terminus of maltose-binding protein (MBP).

All proteins were produced in *E. coli* BL21(DE3) (Stratagene). Transformed bacteria were cultured in LB media, supplemented with 1% (*v/v*) glycerol and selective antibiotics. Cultures were shaken in baffled flasks at 37 °C until the absorbance at 600 nm was ~0.6. Gene expression was induced with the addition of 0.5 mM IPTG, and the cultures subsequently maintained at 18 °C for 18 h, before pelleting the bacteria using centrifugation.

Bacterial cell pellets were subsequently resuspended in a construct-specific lysis buffer (Table S3), supplemented with protease inhibitors (cOmplete EDTA-Free Protease Inhibitor Cocktail, Roche) and lysed by passage through a cell disruptor. Cellular debris was removed using high speed centrifugation. Proteins were isolated by affinity chromatography, using appropriate resin, wash and elution buffers (Table S3). Where required, the affinity tag was cleaved by incubation with Tobacco Etch Virus (TEV) protease (4–18 °C for ~18 h), either pre-elution (GST fusion proteins) or post-elution (poly-histidine fusion proteins). Subsequently, ion exchange chromatography and size exclusion chromatography (SEC) were used sequentially to purify all proteins (Table S3). Proteins in SEC elution buffer (Table S3) were spin concentrated and final protein concentrations were determined from absorbance at 280 nm, with molar absorption coefficients estimated from amino acid composition [37]. The linker domain (P<sub>267-328</sub>) contained no side chain chromophores absorbing in the near UV, and therefore its concentration was determined using the biuret assay, as previously described [38].

### 2.2. SEC-MALLS

For analysis using size exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALLS), proteins in SEC elution buffer (Table S3) were loaded at varying concentrations onto either Superdex 75 (L-BD; L; BD) or Superdex 200 (CC-L-BD) SEC columns (Cytiva). The columns were attached to a Dionex HPLC equipped with a PSS SLD7000 7-angle multi-angle laser light scattering detector and Shodex RI-101 differential refractive index detector. The mass averaged molar mass of each sample was determined from the refractive index and light scattering measurements using the PSS WinGPC UniChrom software, under the assumption of Rayleigh scattering. A constant refractive index increment ( $\partial n/\partial c$ ) of 0.186 was used to estimate all protein concentrations. The MALLS detector was calibrated using a bovine serum albumin solution (2 mg/mL).

### 2.3. Large Zone SEC

For large zone size exclusion experiments, a Tricon 5/5 Superdex 200 column was equilibrated at ambient temperature (20–22 °C) with a buffer containing 12.5 mM MOPS/KOH and 250 mM NaCl. The column void volume ( $V_0$ ) and included volume ( $V_1$ ) were determined using blue dextran (fully excluded standard, eluting at  $V_0$ ) and tyrosine (fully included standard, eluting at  $V_0 + V_1$ ). Solutions of CC-L-BD at varying concentrations (2–90  $\mu$ M) were loaded onto the column until a plateau of constant protein concentration was achieved, as monitored by UV absorbance (Figure S1A). The midpoint of the leading edge ( $V_e$ ) was subsequently determined by numerical integration of the area under the

peak [39]. From the midpoint, a mass average partition coefficient ( $\bar{\sigma}$ ) for the protein was calculated:

$$\bar{\sigma} = \frac{(V_E - V_0)}{V_I} \quad (1)$$

The change in mass average partition coefficient with protein concentration was modeled under assumption of a dimer–tetramer equilibrium. In this case, the experimental data should be described by [39,40]:

$$\bar{\sigma} = f_{P2}\sigma_{P2} + (1 - f_{P2})\sigma_{P4} \quad (2)$$

where  $\sigma_{P2}$  and  $\sigma_{P4}$  are the partition coefficients of the dimer and tetramer respectively, and  $f_{P2}$  is the fraction dimer:

$$f_{P2} = \frac{c_{P2}}{c_{P2} + c_{P4}} = \frac{2[P2]}{2[P2] + 4[P4]} = \frac{2[P2]}{[P_{Total}]} \quad (3)$$

where  $c_{P2}$  and  $c_{P4}$  are the mass concentrations of the dimer and tetramer;  $[P2]$  and  $[P4]$  are the molar concentrations of the dimer and tetramer; and  $[P_{total}]$  is the total molar protein concentration.

Through consideration of mass balance and the chemical equilibrium, the dimer concentration and hence the fraction dimer can be found as the solution to the quadratic equation:

$$f_{P2} = \frac{-1 + \sqrt{1 + 4K[P_{total}]}}{2K[P_{total}]} \quad (4)$$

where  $K$  is the equilibrium association constant.

#### 2.4. Analytical Ultracentrifugation

All analytical ultracentrifugation (AUC) experiments, on construct CC-L-BD in SEC elution buffer (Table S3), were conducted at 20 °C in a Beckman Coulter model XL-I centrifuge equipped with absorbance optics. Sample and reference solutions were transferred into quartz-windowed cells with a double-sector centerpiece and loaded into a Beckman Coulter 8-place An-50 Ti Rotor. Sedimentation velocity experiments were carried out at 50,000 rpm, with protein sedimentation followed by measuring absorbance at 275 nm. Data were analyzed with the program SEDFIT to derive sedimentation coefficient distributions  $c(s)$ , and molar mass distributions  $c(M)$  [41,42]. The program SEDNTERP [43] was used to estimate the required values for the buffer viscosity (0.01036 cP), buffer density (1.00949 g/mL), and protein partial specific volume (0.750 mL/g). Sedimentation equilibrium experiments were carried out at three protein concentrations (1 mM, 2 mM, and 4 mM) and two rotor speeds (10,000 and 13,000 rpm). Data were collected at a wavelength of 228 nm, with a step size of 0.001 cm, and averaging of 25 measurements per step. The program SEDPHAT [44,45] was used for comparative model fitting, assuming either a single sedimenting species, or a variety of schemes for reversible self-association (with the molar mass of the monomer as a fixed parameter [46,47]). Models were compared by analysis of the global goodness of fit (chi-squared values) and by inspection of model residuals for systematic deviations.

#### 2.5. Small Angle X-ray Scattering (SAXS)

##### 2.5.1. Data Collection and Model Free Analysis

SAXS data (Table 1) on proteins CC-L-BD, L and L-BD in SEC elution buffer (Table S3) were collected at the Australian synchrotron SAXS/WAXS beamline over a  $q$  range 0.01–0.6 Å<sup>-1</sup> (where  $q$  is the amplitude of the momentum transfer vector). Data were collected from protein solutions in a 1.5 mm capillary at a temperature of 10 °C, under continuous flow conditions. The program ScatterBrain (Australian Synchrotron) was used to obtain circularly averaged 1D scattering profiles. Baseline subtraction and model-free data analysis was performed using components of the ATSAS software package [48]. The generation of Guinier and Kratky plots was performed with PRIMUS [49], while calcula-

tion of the particle distance distribution function  $P(r)$  via indirect Fourier transform was performed with GNOM [50].

**Table 1.** Small angle X-ray scattering (SAXS) data collection parameters.

	Protein		
	"CC-L-BD" MenV P <sub>209-388</sub>	"L-BD" MenV P <sub>267-388</sub>	"L" MenV P <sub>267-328</sub>
Instrument	Australian Synchrotron SAXS/WAXS beamline, Pilatus 1M detector	Australian Synchrotron SAXS/WAXS beamline, Pilatus 1M detector	Australian Synchrotron SAXS/WAXS beamline, Pilatus 1M detector
Wavelength (Å)	1.0332	1.0332	1.0332
q range (Å <sup>-1</sup> )	0.01–0.60	0.01–0.60	0.01–0.60
Exposure time (s) <sup>a</sup>	1	2	1
Protein concentration range (g/L)	0.1–4.8	0.35–11.1	0.45–3.62
Buffer composition	12.5 mM MOPS/KOH pH 7.0, 250 mM NaCl	12.5 mM MOPS/KOH pH 7.0, 150 mM NaCl	12.5 mM Tris/HCl pH 8.5, 150 mM NaCl
Temperature (°C)	10	10	10
SASDB accession codes	SASDLG9	SASDLH9	SASDLJ9

<sup>a</sup> Data were collected under continuous flow conditions.

The relative scattering intensity at zero angle  $I(0)$  was estimated via standard Guinier analysis [51]. The Guinier plots were linear, and  $I(0)$  scaled linearly with protein concentration in all cases (Figure S2) suggesting the absence of any significant inter-particle interference. This also established that the tetrameric CC-L-BD does not measurably dissociate over the studied concentration range (8–240  $\mu\text{M}$ ), reinforcing the hydrodynamic analysis performed previously.

### 2.5.2. Fit of a Polymer Physics Model to the Data

In the absence of inter-particle interference, the scattered intensity  $I(q)$ , normalized by the scattering at zero angle  $I(0)$ , is equal to the form factor of the scatterer  $P(q)$ .

$$P(q) = \frac{I(q)}{I(0)} \quad (5)$$

An analytic approximation for the form factor of a Kratky–Porod (worm-like) chain was given by Sharp and Bloomfield [52] and has been used for fitting SAXS data for intrinsically disordered proteins and other polymers [53–58]:

$$P_{SB}(q) = \frac{2}{x^2}(x - 1 + e^{-x}) + \frac{b}{L}\left(\frac{4}{15} + \frac{7}{15x} - \left(\frac{11}{15} + \frac{7}{15x}\right)e^{-x}\right) \quad (6)$$

where  $x = \frac{q^2 L b}{6}$ .

The expression (6) is valid only for an infinitely thin chain. The effects of finite chain thickness can be modeled through inclusion of a corrective term into the form factor [59], involving the mean square radius of gyration of the cross section  $R_c^2$ :

$$P(q) = P_{SB} \exp\left(\frac{-q^2 R_c^2}{2}\right) \quad (7)$$

The parameter  $R_c$  is not independent of the contour length  $L$ , the two being related by the volume of the protein [56]. This leads to:

$$R_c^2 = \frac{v_p M}{N_A 2\pi L} \quad (8)$$

where  $v_p$  is the partial specific volume of the protein,  $M$  its molar mass, and  $N_A$  is Avogadro's constant. Together, equations (5)–(8) constitute a model of the small angle X-ray scattering by a worm-like chain, involving three unknown parameters ( $I(0)$ ,  $L$  and  $b$ ) which can be fit to the experimental data  $I(q)$ . For SAXS analysis, the protein partial specific volume was estimated using a group additivity scheme [60].

SAXS model fitting was carried out using pro Fit (QuantumSoft), using a combination of Monte Carlo and Levenberg–Marquardt algorithms. Fit of the model was excellent over the  $q$  range 0–0.35  $\text{\AA}^{-1}$ , with some systematic divergence between data and model observed if the fit was extended to higher scattering angles.

## 2.6. NMR Spectroscopy

The construct L-BD (MenV P<sub>267–388</sub>) was characterized using solution NMR spectroscopy. Details of sample preparation, and the experimental procedures used for  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  NMR chemical shift assignment have been fully detailed in a companion paper [61]. The chemical shifts of MenV P<sub>267–388</sub> were deposited in the Biological Magnetic Resonance Data Bank (<http://www.bmrb.wisc.edu>) under the accession number 27634.

Acquisition and processing of 2D  $^1\text{H}$   $^{15}\text{N}$  HSQC and 3D  $^{15}\text{N}$ -edited NOESY-HSQC spectra was previously described [61]. Using a  $^{15}\text{N}$  labeled L-BD sample in pH 7 phosphate buffer [61], a 2D ARTSY-J experiment [62] was executed to enable estimation of  $^3J_{\text{HN-H}\alpha}$  couplings. The spectra were acquired on a Bruker Avance 600 MHz spectrometer equipped with an inverse triple-resonance pulsed-field gradient cryoprobe. The interleaved time domain data matrix consists of  $2 \times 1024^*$  ( $^1\text{H}$ , acquisition time 136.4 ms)  $\times$   $256^*$  ( $^{15}\text{N}$ , acquisition time 29.0 ms) complex data points. The total dephasing time was set to 40 ms. Squared sine bell functions were used to apodize the data in both dimensions, which were also zero filled prior to Fourier transformation, to yield a final data matrix size of  $4096$  ( $^1\text{H}$ )  $\times$   $1024$  ( $^{15}\text{N}$ ) real points. Spectral processing was performed using NMRPipe [63], and the quantitation of spectral peak intensities was carried out with PINT [64]. The  $^3J_{\text{HN-H}\alpha}$  couplings were estimated from signal intensities in the paired ARTSY-J spectra, for which  $^3J_{\text{HN-H}\alpha}$  dephasing was either omitted, or included, with uncertainties estimated from the noise in the spectral base-plane [62]. Results are presented in Table S4.

The Vadar webserver [65] was used to derive the torsion angle  $\Phi$  from crystallographic model coordinates. The associated  $^3J_{\text{HN-H}\alpha}$  coupling constants were calculated using a Karplus relation, parameterized as follows [66]:

$$^3J_{\text{HN-H}\alpha} = 7.97 \cos^2(\phi - 60^\circ) - 1.26 \cos(\phi - 60^\circ) + 0.63 \quad (9)$$

## 2.7. X-ray Crystallography

Crystallization experiments were performed using several MBP-MenV P fusion proteins, in which the final part of the MenV P linker and the binding domain were fused to the C-terminus of MBP (Table S2). Concentrated proteins (120–160 mg/mL in 12.5 mM Tris/HCl pH 8.5, 150 mM NaCl, 5 mM maltose) were subject to crystallization trials [67] at 18 °C using the sitting-drop vapor diffusion method. Large single crystals of MBP-MenV P<sub>329–388</sub> were obtained in several conditions (Table 2). In preparation for data collection, crystals were harvested using a small fiber loop and transferred into appropriate cryo-protective solutions (Table 2) before being flash frozen by direct immersion in liquid nitrogen. X-ray diffraction data were collected using the screenless oscillation method, using synchrotron and laboratory X-ray sources. Crystals were maintained at 110K in a cold gas stream throughout data collection. Integration and scaling of X-ray diffraction data was carried out with the program HKL2000 [68]. Phases were determined by the method of molecular replacement, using the program Phaser [69], independently positioning MBP and a previously determined structure for the MenV P binding domain (residues 337–388,

PDB Accession Code 4KYC) [35]. The programs Coot [70] and REFMAC [71] were used to perform model building and refinement, respectively. Statistics associated with the X-ray diffraction data and atomic models are collated in Table 2.

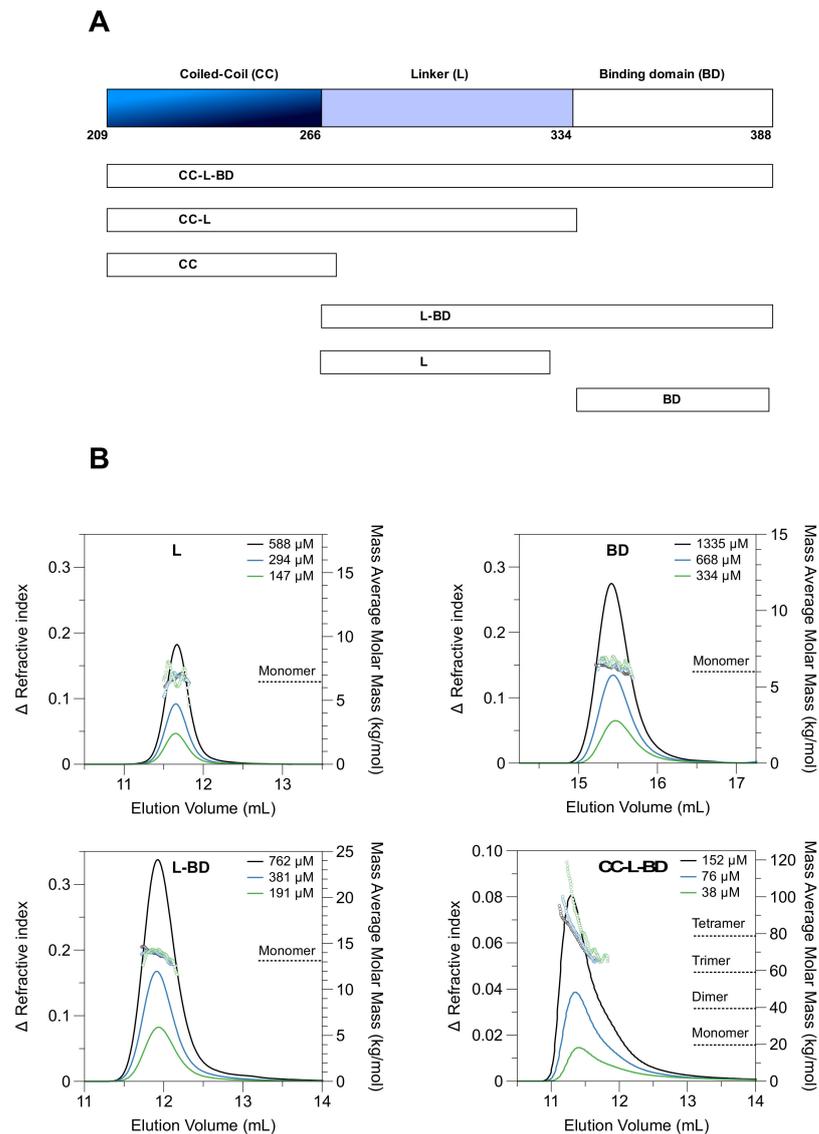
**Table 2.** MBP-MenV P<sub>329–388</sub>: Protein crystallization conditions and statistics associated with X-ray diffraction data and atomic models.

	Crystal Form I	Crystal Form II
Crystallization Conditions		
Protein concentration ( $\mu\text{M}$ )	3300	3300
Crystallization Method	Sitting Drop Vapor diffusion	Sitting Drop Vapor diffusion
Reservoir Solution	1.65 M Ammonium Sulphate 0.20 M Malic acid/KOH pH 5.5	20% ( <i>w/v</i> ) PEG 5000 monomethyl ether 0.2 M PIPES/KOH pH 6.7 0.1 M Proline
Temperature ( $^{\circ}\text{C}$ )	18	18
X-ray diffraction data		
Cryoprotectant	1.65 M Ammonium Sulphate 0.20 M Malic acid/KOH pH 5.5 1M Lithium Sulphate 5 mM Maltose	20% ( <i>w/v</i> ) PEG 5000 monomethyl ether 0.2 M PIPES/KOH pH 6.7 0.1 M Proline 20% ( <i>v/v</i> ) Ethylene glycol 5 mM Maltose
Space group	P2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Unit cell dimensions	a = 57.6, b = 70.1, c = 111.7 Å $\beta = 96.5^{\circ}$	a = 67.6 b = 77.4, c = 79.1 Å
X-ray source	Australian Synchrotron Beamline MX1	Rigaku MicroMax-007 HF Rotating Copper Anode
X-ray wavelength (Å)	0.95370	1.54179
Sample Temperature (K)	110	110
Data resolution limits (Å) <sup>a</sup>	59.26–1.31 (1.33–1.31)	31.24–1.55 (1.57–1.55)
Number of unique observations <sup>a</sup>	149703 (307)	59991 (2017)
Mean Redundancy <sup>a</sup>	3.2 (1.5)	6.8 (5.6)
Completeness (%) <sup>a</sup>	70.7 (2.9)	97.4 (66.5)
R <sub>measure</sub> <sup>a</sup>	0.061 (0.857)	0.026 (0.204)
R <sub>merge</sub> <sup>a</sup>	0.051 (0.607)	0.026 (0.166)
CC1/2 <sup>a</sup>	(0.783)	(0.982)
Crystallographic models		
Number of protein molecules in the asymmetric unit	2 × 431 residues	1 × 431 residues
R <sub>work</sub> /R <sub>free</sub> <sup>b</sup>	14.9/19.6	16.7/19.0
Total number of protein atoms	6998	3569
Number of water molecules	1321	505
Other ligands	Maltose, Sulphate	Maltose, Ethylene Glycol, PIPES, Proline
Disorder Model	Individual Anisotropic B-factors	Individual Isotropic B-factors
Mean total isotropic B-factor, all protein atoms (Å <sup>2</sup> ):	11.7	15.0
Bulk Solvent model	Mask	Mask
RMSD from ideal geometry: Bond lengths (Å)/Bond angles ( $^{\circ}$ )	0.009/1.495	0.010/1.398
Residues in Favoured/Allowed regions of Ramachandran plot (%) <sup>c</sup>	99/1	99/1
PDB ID	7KD4	7KD5

<sup>a</sup> Numbers in parentheses are for the highest resolution shell <sup>b</sup> Calculated from a randomly selected 5% of observations omitted from all model refinement. <sup>c</sup> Defined by the MolProbity web-server.

### 3. Results

To enable biophysical and structural analysis of the MenV P protein C-terminal region (residues 209–388), six truncated variants were bacterially expressed and purified (Figure 1A, Table S1), representing all contiguous combinations of the coiled-coil (CC), the flexible linker (L), and the binding domain (BD). The boundaries of the coiled-coil domain were inferred from prior crystallographic analysis of the mumps virus phosphoprotein [22]. Two of the variants, encompassing the coiled-coil alone (CC), or the coiled-coil and the linker (CC-L), formed large soluble aggregates, as assessed by size exclusion chromatography (SEC) and dynamic light scattering (DLS) (data not shown). These constructs were excluded from subsequent analysis.



**Figure 1.** The MenV P coiled-coil drives tetramer formation. **(A)** A schematic showing the truncated proteins employed in this study; **(B)** SEC-MALLS data for four individual proteins: P<sub>267–328</sub> “L”; P<sub>337–388</sub> “BD”; P<sub>267–388</sub> “L-BD”; P<sub>209–388</sub> “CC-L-BD” as indicated. Solid lines show the SEC elution profile (change in refractive index versus elution volume) at the indicated protein loading concentrations. Hollow circles show the mass averaged molar mass estimates resulting from MALLS. Proteins L, BD and L-BD were analyzed using Superdex 75 media (Cytiva), and the plots employ a common vertical scale for refractive index. Protein CC-L-BD was analyzed using Superdex 200 media (Cytiva), and the plot employs a differing vertical scale for refractive index. Molar mass estimates are displayed over the full width at half maximum peak height, in all cases.

### 3.1. The Coiled-Coil Drives Tetramer Formation

The four tractable truncated variants (CC-L-BD, L-BD, L, BD) were characterized using SEC-MALLS to assess their oligomeric state. The three variants lacking the coiled-coil (L, BD, L-BD) had nearly Gaussian elution profiles (Figure 1B), with neither the profile shape nor elution volume dependent on the protein loading concentration in the  $\mu\text{M}$  range. The mass average molar masses were nearly constant across each peak, and consistent with each protein being monomeric. In contrast, the variant carrying the coiled-coil domain (CC-L-BD) had an asymmetric elution profile (Figure 1B). The mass average molar mass at the leading edge of the peak was correspondent with a tetramer but declined significantly across the trailing edge of the peak. Although there was no large shift in the mean elution volume as the protein concentration was varied ( $\mu\text{M}$  concentration range), the peak asymmetry is itself characteristic of a protein undergoing rapid and reversible self-association [72]. The SEC-MALLS data therefore suggested that the coiled-coil domain drives formation of a P protein tetramer, which is capable of dissociation under *in vitro* conditions.

### 3.2. A Dimer–Tetramer Equilibrium Exists at Sub-Micromolar Protein Concentrations

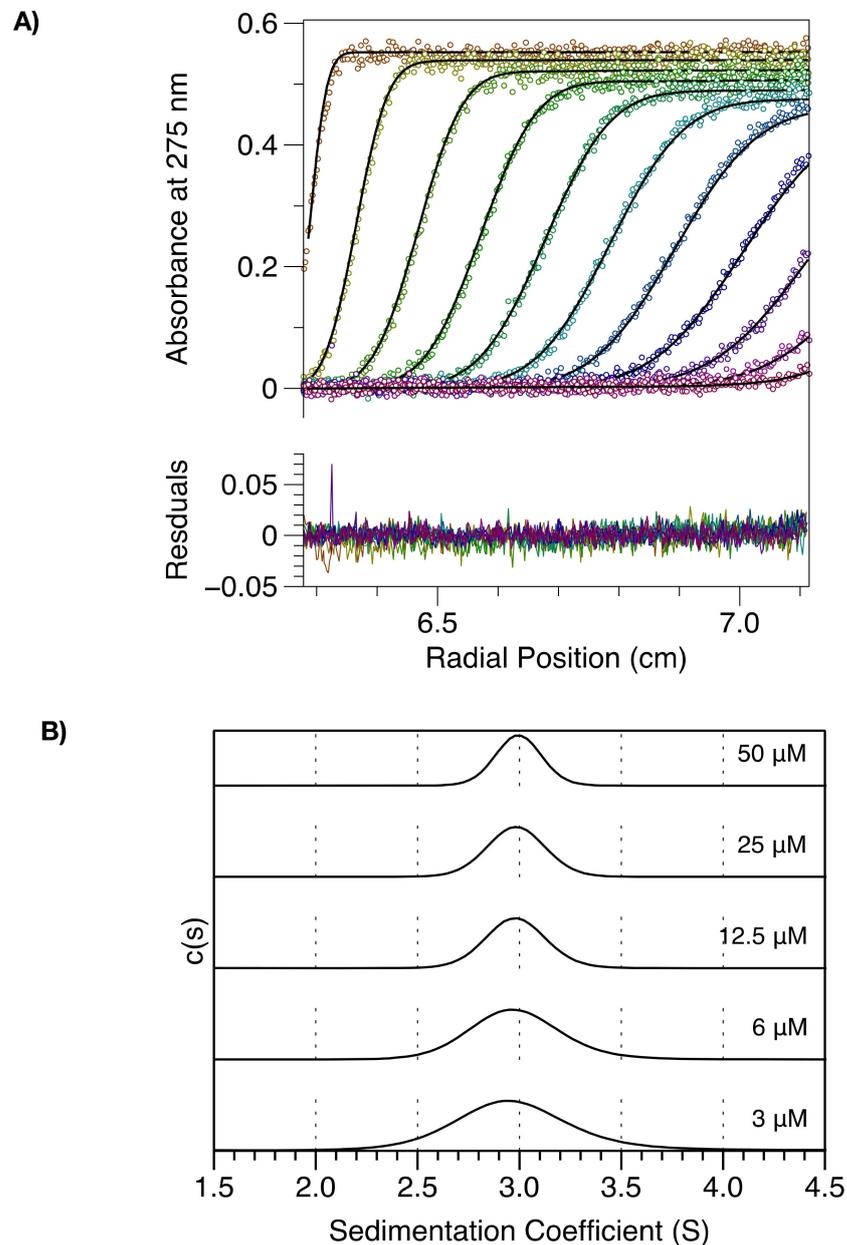
To further investigate self-association of the MenV P protein, sedimentation velocity analytical ultracentrifugation (SV-AUC) experiments were carried out on CC-L-BD. Protein sedimentation was followed by measuring absorbance at 275 nm and the data were analyzed in SEDFIT [41,42] to derive sedimentation coefficient distributions,  $c(s)$ , and molar mass distributions,  $c(M)$ , for the protein at varying concentration (3–50  $\mu\text{M}$ ).

At all concentrations examined, CC-L-BD migrated as a single species with a mean sedimentation coefficient of  $\sim 3$  S (Figure 2). At the lowest protein concentrations examined ( $\leq 6$   $\mu\text{M}$ ) some peak broadening is apparent in the  $c(s)$  distribution, and there is a very slight reduction in the mass averaged sedimentation coefficient, obtained through peak integration [42]. The molar mass of the single species observed in  $c(M)$  distributions (not shown) was 78 kg/mol, consistent with the expected molar mass of the tetramer (78.8 kg/mol). These data establish that at protein concentrations  $> 3$   $\mu\text{M}$ , the MenV P protein exists predominantly in the tetrameric state, but are not conclusive with respect to tetramer dissociation. Hence, several additional experiments were performed.

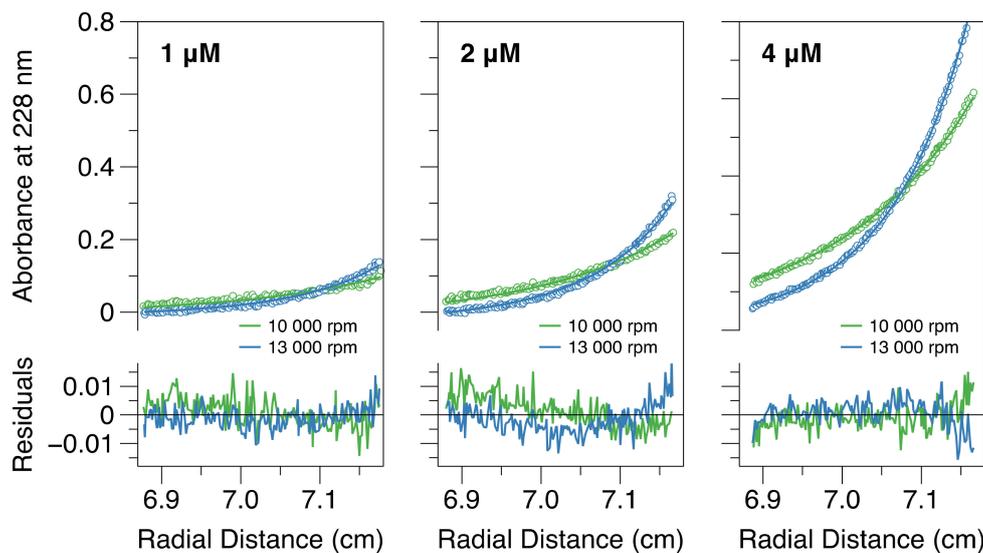
Initially, large zone SEC analysis [39,40] was carried out on CC-L-BD (Figure S1). In this experiment, large volumes of protein were loaded onto a miniaturized SEC column [73,74], sufficient to establish a “plateau region” of constant protein concentration (2–90 mM) (Figure S1A). Under these conditions, the position of the leading boundary of the elution profile accurately reflects the equilibrium distribution of species at the plateau concentration [39,40]. The leading boundary position increased significantly at the lowest plateau protein concentrations tested (2  $\mu\text{M}$ ), suggesting that the tetramer begins to dissociate at these concentrations. However, the range of protein concentrations that could be explored was limited by the sensitivity of the HPLC absorbance optics. Hence, it was not possible to meaningfully fit the experimental data to recover the partition coefficients of the species present, nor the associated equilibrium constant. The data are, however, satisfactorily explained by a dimer–tetramer equilibrium model (Equations (2)–(4)) with physically reasonable values for the parameters (Figure S1B).

Subsequently, sedimentation equilibrium analytical ultracentrifugation (SE-AUC) experiments were carried out on CC-L-BD. This technique reports on protein mass rather than size, and hence is more straightforward to interpret than large zone SEC. SE-AUC data were collected at two speeds, and three protein concentrations (1 mM, 2 mM and 4 mM), with the equilibrium distribution of the protein in the AUC cell measured via absorbance at 228 nm. Comparative model analysis established that a reversible dimer–tetramer equilibrium provided the best overall description of the data. When the data were fit to this model (Figure 3), the equilibrium dissociation constant ( $K_D$ ) was estimated as 58 nM. However confidence surface analysis [75] established that the point estimate is loosely bounded from above. This results from the relatively high protein concentrations

employed in the experiment, which were again dictated by the instrumental sensitivity. If a dimer–tetramer equilibrium exists in solution, with  $K_D = 58$  nM and a total protein concentration of 1  $\mu$ M, the fraction dimer would be 0.2. Thus, only the initial stage of tetramer dissociation can be followed experimentally within the micromolar concentration range, which is consistent with the results obtained using all techniques.



**Figure 2.** MenV P exists as a tetramer in the micro-molar concentration range. The figure shows sedimentation velocity analytical ultra-centrifugation data and its direct modeling using the program SEDFIT. (A) Time-dependent absorbance scans of CC-L-BD (50  $\mu$ M concentration) undergoing sedimentation (hollow circles), together with the fit model (solid black lines). Every fourth scan is shown, with the model residuals displayed in the bottom panel. (B) Sedimentation coefficient distributions,  $c(s)$ , as a function of protein concentration.

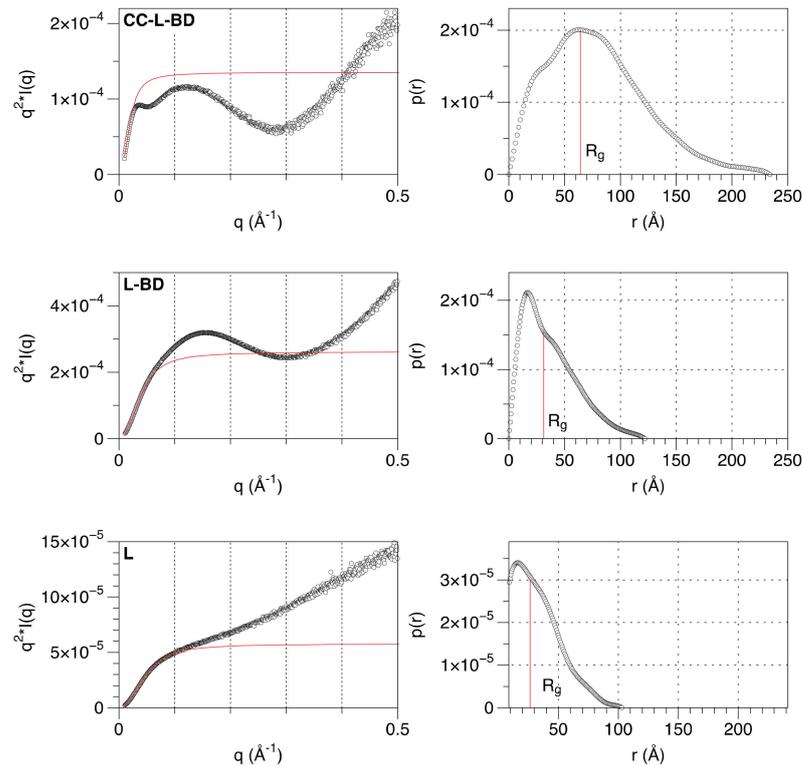


**Figure 3.** A dimer–tetramer equilibrium becomes apparent at low micromolar protein concentrations. The figure shows sedimentation equilibrium analytical ultra-centrifugation data (hollow circles) for CC-L-BD at three different protein concentrations and two different rotor speeds, and its global fit to a rapid reversible dimer–tetramer self-association model (solid lines). Model residuals are shown below the main plots.

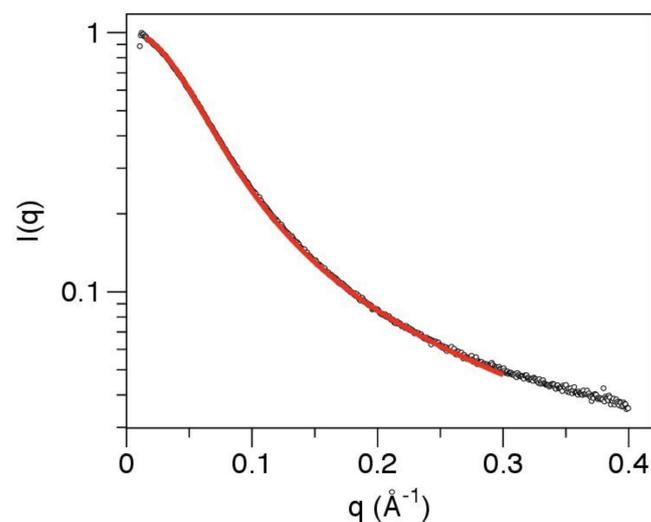
### 3.3. The Linker Sequence Can Be Modeled as A Worm-like Chain and Exhibits Only Weak Local Conformational Preferences

The structural properties of the linker that connects the coiled-coil and the binding domain were further investigated. Small angle X-ray scattering (SAXS) data were collected for three of the truncated variants (CC-L-BD, L-BD, L) (Table 1). Model-free analysis of the SAXS data confirms the expected organization of the MenV P protein C-terminal region. The Kratky plot (Figure 4) for the entire C-terminal region (CC-L-BD) has a bell-shaped appearance, consistent with the major scattering contribution coming from the two structured domains (CC and BD) positioned at either end of the chain. In contrast, the Kratky plot for the isolated linker (L) increases monotonically with the scattering angle, which is characteristic of unfolded proteins [76]. The Kratky plot for L-BD has a hybrid appearance, reflecting its nearly equal partition between structured and unstructured regions. From the pair distance distribution function  $P(r)$  (Figure 4), the radius of gyration ( $R_g$ ) of the linker is estimated as 26.4 Å by numerical integration. For comparison, the radius of gyration of an intrinsically disordered protein of this chain length (62 residues) is predicted to be 22.7 Å, using a power law expression developed for chemically denatured proteins [77]. Hence, the linker is globally disordered.

If the linker were a completely flexible Gaussian chain with no local rigidity, a Kratky plot would be expected to plateau at high scattering angles [78], which it does not (Figure 4). A more realistic model allows for the stiffness of the polypeptide chain. The SAXS data were therefore fit to a polymer physics model for chain scattering, following established approaches [55–58]. The worm-like chain (Kratky–Porod) model fit to the data (Equations 5–8) allows for local rigidity in the polypeptide backbone and incorporates a correction for finite chain thickness. Despite the apparent complexity of the model, it involves only three independent parameters: the scattered intensity at zero angle  $I(0)$ , the contour length  $L$  (the length of the chain at maximum possible extension) and the statistical length  $b$  (which is twice the persistence length, and a measure of local chain rigidity). Although the analytic expression describing the scattering of the worm-like chain (Equation 6) involves approximations which are strictly valid only for long chains ( $L/b > 10$ ), and at low scattering angles ( $qb < 3$ ) [53,54], the model fits the data very well over the  $q$  range 0.015–0.30 Å<sup>−1</sup> (Figure 5).



**Figure 4.** Model-free analysis of SAXS data confirms that the linker is globally disordered. The panels on the left show the Kratky plots ( $I \cdot q^2$  vs.  $q$ ) for CC-L-BD (Top), L-BD (middle) and L (bottom). For comparison, the Kratky plot expected of a completely flexible, infinitely thin Gaussian chain, with the same overall radius of gyration ( $R_g$ ) is shown with a solid red curve. The form factor of the Gaussian chain is given by the Debye function  $P(q) = \frac{2}{x^2}(x - 1 + e^{-x})$ , where  $x = (q \cdot R_g)^2$  [78]. The panels on the right show the correspondent pair distance distribution function  $P(r)$ . The pair distance distribution functions were calculated by indirect Fourier transformation of the experimental SAXS profiles, as implemented in the program GNOM [50]. The radius of gyration, calculated by numerical integration of the pair distribution functions, is indicated.



**Figure 5.** SAXS data for the linker can be fit to a worm-like-chain (Kratky–Porod) model, providing estimates for the contour length and persistence length. Hollow circles show the SAXS data for the linker on a log-linear scale. The red line shows the fit of the model over the  $q$  range  $0.015\text{--}0.30 \text{ \AA}^{-1}$ . Best fit model parameters: contour length  $192 \text{ \AA}$  and persistence length  $13.5 \text{ \AA}$ .

The SAXS analysis therefore confirms that the linker has an extended conformation with no significantly populated tertiary structure. It can be modeled effectively as a worm-like chain, with a mean persistence length of 13.5 Å and contour length of 192 Å (Figure 5). However, this does not establish that the sequence is uniformly disordered across its entire length. Indeed the mean persistence length is at the upper end of the range observed for intrinsically disordered proteins, with values of 9–10 Å expected for a fully disordered protein [57].

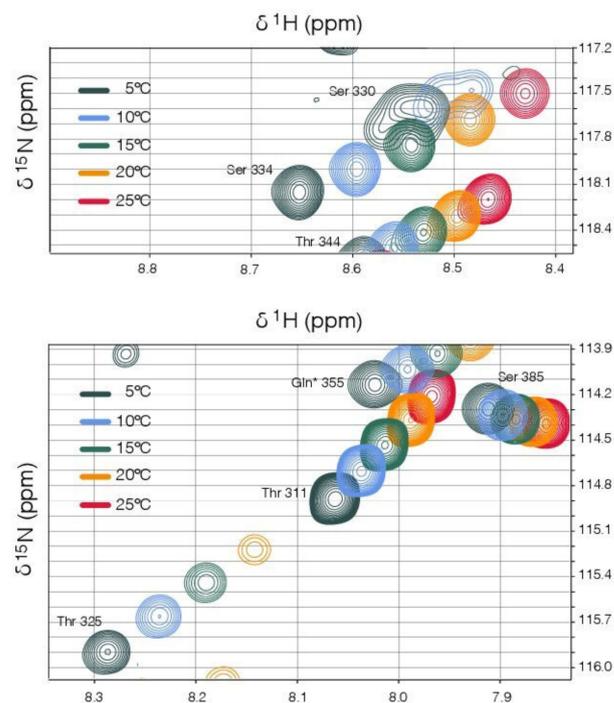
Sequence analysis suggests the potential for localized and transient structure formation within the linker. Relative to folded proteins, IDPs are typically depleted in hydrophobic residues, and enriched in charged and polar residues [79,80]. However, this is not true of the MenV P linker sequence. The linker has the mean hydrophobicity (0.48 assessed using the Kyte–Doolittle hydrophobicity scale [81]) expected of a folded and globular protein [79]. A significant complement of hydrophobic residues, distributed throughout the linker sequence, could promote localized hydrophobically-driven collapse of the polypeptide chain.

In addition, the sequence of the linker is not hypervariable across the para and orthorubulaviruses, and contains three conserved motifs (Figure S3). The consensus sequences of these motifs are TTIKIMDPG (aa 267–275 in MenV P); KKxFKEVPVVVSGP (aa 286–299 in MenV P) and IxLDELARP (aa 312–320 in MenV P), where x is any amino acid. These motifs are most likely related to binding activities of the P protein and are potential loci for coupled binding and folding events.

Finally, the linker sequence is predicted to have some local conformational preferences. In particular, application of the GOR IV algorithm [82]—the standard information theoretic method for secondary structure prediction—identifies several regions (Figure S4A) with propensity for either alpha-helical conformation (residues 280–289 and 309–317) or extended beta-conformation (residues 267–272 and 294–298).

To experimentally investigate the local structural characteristics of the linker we used solution NMR spectroscopy. A companion paper [61] describes the near complete NMR chemical shift assignment of the construct encompassing both the linker and the binding domain (L-BD). As expected, backbone amide resonances associated with the linker (L) showed very limited chemical shift dispersion in the proton dimension, consistent with global disorder, while those associated with the binding domain (BD) were widely dispersed, consistent with overwhelming population of the folded state [61]. The propensity for structure formation in the linker was assessed using the ncSPC algorithm (Figure S4B), based on the weighted deviation of the chemical shifts of five backbone nuclei ( $^1\text{H}\alpha$ ,  $^{13}\text{CO}$ ,  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}\beta$ ,  $^{15}\text{N}$ ) from reference values specific for intrinsically disordered proteins [83]. The results are fairly concordant with secondary structure prediction (Figure S4A) and suggest that very weak local structural preferences for extended beta-conformation or alpha-helical conformation do exist, with these preferences alternating along the length of the linker.

A striking feature of the NMR spectra is the very strong temperature-dependent line broadening that occurs in several regions of the linker. Particularly affected are resonances associated with residues 277–288, a region proximal to the coiled-coil predicted to have alpha-helical propensity (Figure S2), and residues 325–333, the region immediately adjacent to the binding domain. In each case the resonances broaden with increasing temperature (5–25 °C). For the region adjacent to the binding domain, the line broadening is so extensive that some resonances can no longer be detected at the higher temperatures (Figure 6). This suggests that temperature-dependent conformational exchange processes are occurring in both these regions, though it is not possible to infer either the timescale or the exact nature of the underlying conformational changes from the line broadening alone [84,85].



**Figure 6.** Solution NMR data suggests residues adjacent to the binding domain (residues 325–333) are undergoing temperature-dependent conformational exchange. Selected regions of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra, collected at temperatures between 5 and 25 °C, are overlaid in the two panels, with identical iso-contours displayed. As expected [86], the amide chemical shifts track fairly linearly with temperature in all cases. For the residues T325 (bottom panel) and S330 (top panel), which are typical of this region, the backbone amide peak intensities reduce markedly with temperature, with the S330 resonance broadened beyond the detection limit at temperatures >15 °C.

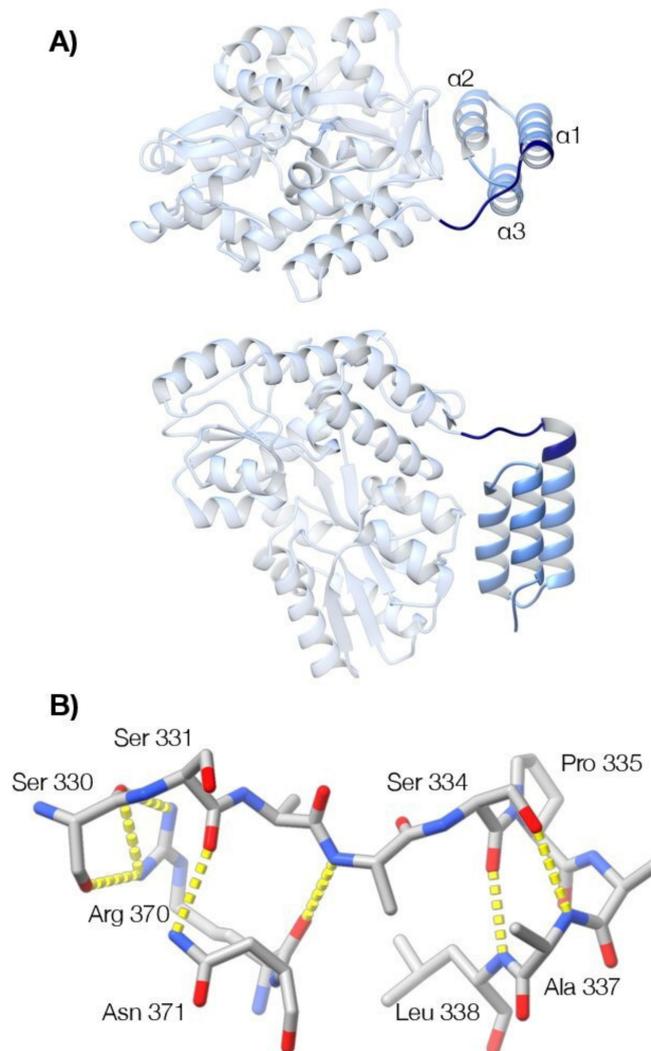
Overall, the NMR analysis confirmed that the linker is not uniformly disordered. It contains localized segments with very weak conformational preferences, and there is evidence for conformational exchange occurring in some regions. A comprehensive investigation of polypeptide chain dynamics using nuclear spin relaxation is ongoing.

#### 3.4. The Structural Transition between the Disordered Linker and the Ordered Binding Domain Is Not Abrupt

The NMR-based evidence for conformational exchange in the sequence immediately preceding the binding domain prompted us to reinvestigate the 3D structure of this region using X-ray crystallography. A series of constructs were created for crystallographic analysis in which the originally characterized binding domain of MenV P (amino acids 337–388, PDB ID 4KYC [35]) was N-terminally extended by 7, 8, 10, or 16 residues. As in the prior crystallographic analysis, each of these proteins was fused to a non-cleavable, modified MBP tag, optimized for crystallization [87]. For one of these proteins (MBP-P<sub>329–388</sub>), crystals were obtained in multiple conditions and used for crystallographic structure determination (Table 2).

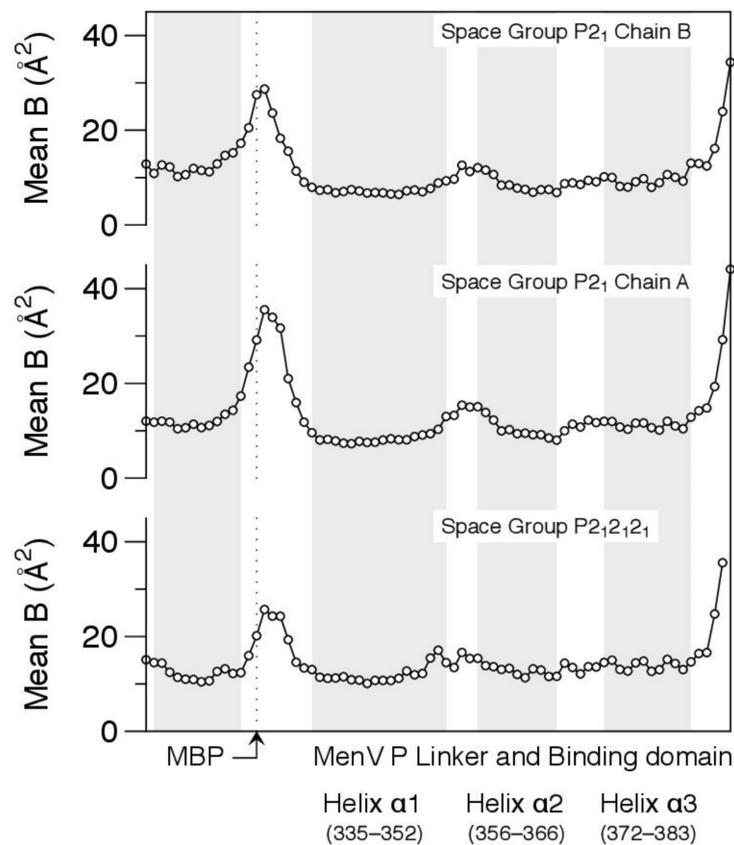
X-ray diffraction data were collected to ~1.3 Å resolution for crystal form one (space group P2<sub>1</sub>; two molecules in the asymmetric unit) and 1.55 Å resolution for crystal form two (space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>; one molecule in the asymmetric unit). The structures were determined by the method of molecular replacement, providing a total of three instances of the molecule. In each case, helix  $\alpha$ 1 of the BD is N-terminally extended beyond Ser339, the previously identified N-terminal residue of the domain (Figure 7). Helix  $\alpha$ 1 is capped at its N-terminus by Pro335, which orients the preceding residues (329–334) such that they form a clasp or latch that runs across the top of the three-helix bundle. Superimposition of the three molecular models shows that this latch is equivalently positioned in each case

with only minor variation in the side-chain hydrogen bonding that mediates formation of the latch. In particular, Arg370 and Asn371, located within the loop connecting helices  $\alpha 2$  and  $\alpha 3$  of the BD, interact with Ser330 and Ser331 in the latch sequence, through their side chain guanidinium and carboxamide groups (Figure 7).



**Figure 7.** Crystallographic analysis reveals a latch preceding the canonical three-helix bundle of the binding domain. **(A)** The structure of MenV P<sub>329-388</sub> (blue) fused to MBP (grey), in ribbon representation (P<sub>2</sub><sub>1</sub>2<sub>1</sub>2<sub>1</sub> crystal form, PDB ID 7KD5). Two orthogonal views are shown. The previously determined structure of the binding domain (residues 339–388) is displayed in light blue, while the newly identified structure (residues 329–338) is displayed in navy blue. **(B)** Selected residues that mediate formation of the latch and the N-terminal cap of helix  $\alpha 1$  are shown in stick representation (P<sub>2</sub><sub>1</sub>2<sub>1</sub>2<sub>1</sub> crystal form, PDB ID 7KD5). Hydrogen bonding interactions are shown by the yellow dashed lines. The program UCSF ChimeraX was used to prepare the figures [88].

The electron density associated with the residues within the latch is unambiguous (Figure 7), but relatively weak, and this is reflected in the atomic displacement parameters of the structural models (Figure 8). In all cases, the main chain isotropic B-factors of residues within the latch are significantly higher than values in the bordering helices of MBP and the BD, suggesting that the structure is quite dynamic.



**Figure 8.** Crystallographic analysis suggests the latch preceding the binding domain is dynamic. Mean isotropic displacement parameters (B-factors) for main chain atoms are shown for the final helix of MBP and MenV P<sub>329–388</sub>, which incorporates the last part of the linker and the binding domain. Positions of the alpha helical secondary structures, as defined by the DSSP algorithm [89], are indicated with gray background shading.

### 3.5. Crystallographic Model for the Interfacial Region Is Consistent with the NMR Observations

The structure observed crystallographically at the N-terminal boundary of the binding domain is consistent with the NMR observations. At the beginning of helix  $\alpha 1$ , the characteristic  $i+3$  and  $i+4$  hydrogen bonding interactions of the alpha helix bring the backbone amide group of Ser334 in close proximity to the side chain methyl group of Ala337 (Figure 7). The existence of this close contact in the solution structure was confirmed by detection of the correspondent Nuclear Overhauser Effect (NOE) (Figure S5). The presence of a canonical capping structure on helix  $\alpha 1$ , involving the side chain hydroxyl group of Ser 334, and the backbone amide group of Ala 337 (Figure 7) is also robustly predicted from the NMR chemical shifts alone, using the MICS algorithm [90]. Hence, the NMR data clearly establishes the existence of the N-terminal cap on the first helix of the binding domain, directing the polypeptide chain across the top of the binding domain, as observed in the crystal structures (Figure 7).

Crystallographic analysis suggests the latch wrapping the surface of the binding domain is likely to be highly mobile (Figure 8), and this is supported by the solution NMR data. In addition to the severe temperature-dependent line broadening of the resonances in this region (Figure 6), suggestive of conformational exchange, we were unable to detect any NOES between the residues in the latch and the body of the binding domain, despite the close contacts observed in the crystallographic models. A caveat is that the most diagnostic NOES in this region involve the rapidly exchangeable protons from the guanidinium group of Arg370 (Figure 7).

However, measurement of  $^3J_{\text{HN-H}\alpha}$  coupling constants provides additional evidence that the latch residues are undergoing conformational averaging in solution. The  $^3J_{\text{HN-H}\alpha}$  coupling constants can be quantitatively related to the backbone torsion angle  $\varphi$  via a Karplus-type relation [91], and were measured experimentally for almost all residues in L-BD (Table S4). For residues within the helices of the structured binding domain, variance between the experimentally measured  $^3J_{\text{HN-H}\alpha}$  values, and those calculated from the crystal structure using the Karplus relation is small (RMSD 0.99 Hz; calculations performed for chain A, space group P2<sub>1</sub>) However, for the residues located in the latch (residues 329–334), the variance between experimental  $^3J_{\text{HN-H}\alpha}$  values and those calculated from the crystal structure is much larger (RMSD 2.53 Hz), supporting a high degree of intra-chain mobility in this region. Overall, we hypothesize that the conformation of the latch observed in the crystals is only partially populated in solution, and in exchange with the unstructured state. This hypothesis is consistent with the solution NMR data, though not strictly established by it.

#### 4. Discussion

The paramyxoviral P protein is an integral component of the viral RdRp and its C-terminal region enables the RdRp to engage with the viral nucleocapsid. We have investigated the biophysical and structural characteristics of the C-terminal region of the P protein from Menangle virus, a bat borne pararubulavirus. The major focus of our study was to establish the oligomerization state of the MenV P protein, and the structural characteristics of the flexible linker that connects the centrally located coiled-coil to the C-terminal binding domain.

Experimental analysis confirms that the coiled-coil alone drives tetramerization of the MenV P protein (Figures 1 and 2). However, the MenV P protein is not a constitutive tetramer in solution. Using a combination of SEC-MALLS (Figure 1), large zone SEC (Figure S1) and sedimentation velocity and equilibrium AUC (Figures 2 and 3), we showed that as the total protein concentration drops into the nano-molar range, the tetramer begins to dissociate into dimers. From the sedimentation equilibrium AUC experiments (Figure 3), we estimate the dissociation constant  $K_D$  for the dimer–tetramer equilibrium as ~58 nM. These data suggest that the MenV P protein will likely function as an obligate tetramer in vivo, especially given that RNA synthesis appears to take place in inclusion bodies [92–96], where the local concentration of the P protein would be very high.

Nonetheless, the existence of a measurable dimer–tetramer equilibrium has some structural implications. It suggests that the tetrameric coiled-coil of the MenV P protein is unlikely to be arranged in a parallel configuration with 4-fold rotational symmetry, like the majority of paramyxoviral P proteins characterized to date [16,20,21,23]. Instead, the MenV P coiled-coil is more likely a parallel or antiparallel “dimer-of-dimers”, with only 2-fold rotational symmetry, as this allows for dimer dissociation. The mumps virus P protein is configured in this way [22,24] and its self-associative behavior was previously characterized using sedimentation velocity AUC experiments [22]. At a fixed concentration of 1 mg/mL (~70  $\mu$ M), only a single tetrameric species was apparent, however that is consistent with our own results for the MenV P CC domain at similar concentrations (Figure 2).

There has been prior characterization of the linker region from the P proteins of Sendai virus (genus *Respirovirus*) [97,98], Nipah virus (genus *Henipavirus*) [23] and measles virus (genus *Morbillivirus*) [99] using NMR spectroscopy, sometimes in combination with SAXS. In all cases the linker is globally disordered. For Sendai virus, the C-terminal half of the ~87 residue linker was analyzed, and was found to be highly flexible, with local conformational sampling effectively dictated by the amino acid type at each position in the linker. These findings were based on both analysis of structural preferences using residual dipolar couplings [97] and analysis of chain dynamics using nuclear spin relaxation [98]. In the case of Nipah Virus, NMR chemical shift analysis suggested that there are weak local preferences for  $\alpha$ -helical conformation within the central region of the ~70 amino acid linker [23]. For measles virus, ensemble modeling based on NMR chemical shifts suggested

that the ~80 amino acid linker has two non-contiguous regions with weak preferences for  $\alpha$ -helical conformation, but is otherwise largely devoid of significantly populated secondary structure [99].

In Menangle virus, the ~62 residue linker is also globally disordered and lacks any tertiary structure. This is readily apparent from model-free analysis of the SAXS data (Figure 4). Fit of the SAXS data to a worm-like chain model (Figure 5) returns a mean persistence length of 13.5 nm for the linker, slightly larger than the value (9–10 Å) expected of a fully disordered protein [57]. This may reflect either the weak secondary structural preferences indicated by NMR chemical shift measurements (Figure S4), and/or the chain-stiffening effects of the six proline residues that are distributed throughout the linker. The contour length of the linker (192 Å) clearly allows for multivalent attachment of the binding domains of the P protein to the nucleocapsid (the inter-subunit distances within the nucleocapsid being of the order of 50 Å [9–11,13,14]). However, it is presently unknown if this is required for polymerase function.

Using X-ray crystallography we determined the structure of residues 329–388 of MenV P, encompassing the final part of the linker and the binding domain (Figure 7). The structural analysis showed that helix  $\alpha 1$  of the binding domain is N-terminally extended, relative to other paramyxoviral P proteins [30,32–34]. Helix  $\alpha 1$  is capped at its N-terminus by a proline residue (P335), which directs the polypeptide chain across the top of the three-helix bundle, forming a latch-like structure. This latch is held in place by H-bonding interactions between Arg370 and Asn371, located in a loop connecting  $\alpha 2$  and  $\alpha 3$  of the NBD, and Ser 330 and 331, located within the latch (Figure 7), though the crystallography suggests a high degree of intrachain mobility in the latch residues (Figure 8). Based on the NMR observations, the latch cannot be the only populated state in solution, and the crystallographic model must capture just one of the conformational states accessible to the protein. Hence, the transition between the disordered linker and the ordered binding domain is not abrupt, and the interfacial region is best described as partially structured. The existence of a partially structured interface between the linker and the C-terminal binding domain has not previously been observed, however this feature may be restricted to MenV and its most immediate clade, as even among the ortho- and para-rubulaviruses, the key residues involved in formation of the latch (Figure 7) are not highly conserved (Figure S3).

Several lines of evidence suggest that the linker region may act as more than a simple tether for the C-terminal binding domains during polymerase translocation. Firstly, in the structure of the parainfluenza virus 5 P/L complex determined by Cryo-EM [16], a single binding domain is found attached to the surface of the L protein, on the face opposite the entrance and exit channels for the template RNA. The binding domain and coiled-coil of the P protein are connected by diffuse yet compact density that is assigned to the linker. While the density was not directly interpretable, the compacted linker appears to wrap the surface of the L protein, and it is possible that the contacts could involve specific sequence motifs within the linker. In a different study, investigating the determinants of liquid–liquid phase separation by the measles virus N and P proteins, very weak binding was detected between residues within the linker, and the RNA bound N protein, which appeared to modulate the physical properties of the phase separated droplets [99]. Hence, it may be that the linker, while intrinsically disordered, and traditionally viewed as a spacer or tether, is the locus for several ancillary binding events during RNA synthesis. Our work has outlined some of the basic physical properties of the C-terminal region of the MenV P protein, which will facilitate understanding of such binding events, as well as the extension and compaction of the linker which appears necessary for polymerase function.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/v13091737/s1>.

**Author Contributions:** Conceptualization, M.N.W., N.H. and R.L.K.; Methodology, M.N.W., N.H., E.M.M.B., M.S., J.R.K., D.C.G. and R.L.K.; Software, R.L.K.; Validation, M.N.W., N.H., E.M.M.B., M.S., J.R.K., D.C.G. and R.L.K.; Formal Analysis, M.N.W., N.H., E.M.M.B., J.R.K., D.C.G. and R.L.K.;

Investigation, M.N.W., N.H., M.S., J.R.K. and R.L.K.; Resources, R.L.K.; Data Curation, R.L.K.; Writing—Original Draft Preparation, M.N.W., N.H. and R.L.K.; Writing—Review and Editing, M.N.W. and R.L.K.; Visualization, M.N.W., N.H. and R.L.K.; Supervision, R.L.K., E.M.M.B. and D.C.G.; Project Administration, R.L.K.; Funding Acquisition, R.L.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by a RSNZ Marsden grant, and additionally supported by University of Auckland doctoral scholarships (to M.N.W. and N.H.). Access to analytical ultracentrifugation instrumentation was enabled by the Maurice Wilkins Centre, a Centre of Research Excellence funded by the NZ government. This research was undertaken in part using the macromolecular crystallography beamlines at the Australian Synchrotron, a part of ANSTO, with access enabled by the NZ government.

**Data Availability Statement:** The SAXS data have been deposited with the SASBDB (Accession codes SASDLG9, SASDLH9, and SASDLJ9). The crystallographic models and the processed diffraction data have been deposited with the PDB (Accession codes 7KD4 and 7KD5).

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

## References

1. Drexler, J.F.; Corman, V.M.; Müller, M.A.; Maganga, G.D.; Vallo, P.; Binger, T.; Gloza-Rausch, F.; Cottontail, V.M.; Rasche, A.; Yordanov, S.; et al. Bats host major mammalian paramyxoviruses. *Nat. Commun.* **2012**, *3*, 796. [[CrossRef](#)] [[PubMed](#)]
2. Smith, I.; Wang, L.-F. Bats and their virome: An important source of emerging viruses capable of infecting humans. *Curr. Opin. Virol.* **2013**, *3*, 84–91. [[CrossRef](#)]
3. Hauser, N.; Gushiken, A.; Narayanan, S.; Kottilil, S.; Chua, J. Evolution of Nipah Virus Infection: Past, Present, and Future Considerations. *Trop. Med. Infect. Dis.* **2021**, *6*, 24. [[CrossRef](#)]
4. Clayton, B.A. Nipah virus: Transmission of a zoonotic paramyxovirus. *Curr. Opin. Virol.* **2017**, *22*, 97–104. [[CrossRef](#)] [[PubMed](#)]
5. Ong, K.C.; Wong, K.T. Henipavirus Encephalitis: Recent Developments and Advances. *Brain Pathol.* **2015**, *25*, 605–613. [[CrossRef](#)]
6. Bowden, T.R.; Westenberg, M.; Wanga, L.-F.; Eaton, B.T.; Boyle, D.B. Molecular Characterization of Menangle Virus, a Novel Paramyxovirus which Infects Pigs, Fruit Bats, and Humans. *Virology* **2001**, *283*, 358–373. [[CrossRef](#)] [[PubMed](#)]
7. Barr, J.A.; Smith, C.; Marsh, G.A.; Field, H.; Wang, L.-F. Evidence of bat origin for Menangle virus, a zoonotic paramyxovirus first isolated from diseased pigs. *J. Gen. Virol.* **2012**, *93*, 2590–2594. [[CrossRef](#)]
8. Philbey, A.W.; Kirkland, P.D.; Ross, A.D.; Davis, R.J.; Gleeson, A.B.; Love, R.J.; Daniels, P.W.; Gould, A.R.; Hyatt, A.D. An Apparently New Virus (Family Paramyxoviridae) Infectious for Pigs, Humans, and Fruit Bats. *Emerg. Infect. Dis.* **1998**, *4*, 269–271. [[CrossRef](#)] [[PubMed](#)]
9. Alayyoubi, M.; Leser, G.P.; Kors, C.A.; Lamb, R.A. Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein–RNA complex. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E1792–E1799. [[CrossRef](#)]
10. Gutsche, I.; Desfosses, A.; Effantin, G.; Ling, W.L.; Haupt, M.; Ruigrok, R.W.H.; Sachse, C.; Schoehn, G. Structural virology. Near-atomic cryo-EM structure of the helical measles virus nucleocapsid. *Science* **2015**, *348*, 704–707. [[CrossRef](#)]
11. Song, X.; Shan, H.; Zhu, Y.; Hu, S.; Xue, L.; Chen, Y.; Ding, W.; Niu, T.; Gu, J.; Ouyang, S.; et al. Self-capping of nucleoprotein filaments protects the Newcastle disease virus genome. *eLife* **2019**, *8*. [[CrossRef](#)]
12. Ker, D.-S.; Jenkins, H.T.; Greive, S.J.; Antson, A.A. CryoEM structure of the Nipah virus nucleocapsid assembly. *PLoS Pathog.* **2021**, *17*, e1009740. [[CrossRef](#)] [[PubMed](#)]
13. Zinzula, L.; Beck, F.; Klumpe, S.; Bohn, S.; Pfeifer, G.; Bollschweiler, D.; Nagy, I.; Plitzko, J.M.; Baumeister, W. Cryo-EM structure of the cetacean morbillivirus nucleoprotein-RNA complex. *J. Struct. Biol.* **2021**, *213*, 107750. [[CrossRef](#)] [[PubMed](#)]
14. Zhang, N.; Shan, H.; Liu, M.; Li, T.; Luo, R.; Yang, L.; Qi, L.; Chu, X.; Su, X.; Wang, R.; et al. Structure and assembly of double-headed Sendai virus nucleocapsids. *Commun. Biol.* **2021**, *4*, 494. [[CrossRef](#)] [[PubMed](#)]
15. Fearn, R.; Plemper, R.K. Polymerases of paramyxoviruses and pneumoviruses. *Virus Res.* **2017**, *234*, 87–102. [[CrossRef](#)]
16. Abdella, R.; Aggarwal, M.; Okura, T.; Lamb, R.A.; He, Y. Structure of a paramyxovirus polymerase complex reveals a unique methyltransferase-CTD conformation. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 4931–4941. [[CrossRef](#)]
17. Liang, B. Structures of the Mononegavirales Polymerases. *J. Virol.* **2020**, *94*. [[CrossRef](#)]
18. Velthuis, A.J.W.T.; Grimes, J.M.; Fodor, E. Structural insights into RNA polymerases of negative-sense RNA viruses. *Nat. Rev. Microbiol.* **2021**, *19*, 303–318. [[CrossRef](#)]
19. Noton, S.L.; Fearn, R. Initiation and regulation of paramyxovirus transcription and replication. *Virology* **2015**, *479–480*, 545–554. [[CrossRef](#)]
20. Tarbouriech, N.; Curran, J.; Ruigrok, R.W.; Burmeister, W.P. Tetrameric coiled coil domain of Sendai virus phosphoprotein. *Nat. Struct. Biol.* **2000**, *7*, 777–781. [[CrossRef](#)]
21. Communie, G.; Crepin, T.; Maurin, D.; Jensen, M.R.; Blackledge, M.; Ruigrok, R.W.H. Structure of the Tetramerization Domain of Measles Virus Phosphoprotein. *J. Virol.* **2013**, *87*, 7166–7169. [[CrossRef](#)]

22. Cox, R.; Green, T.J.; Purushotham, S.; Deivanayagam, C.; Bedwell, G.J.; Prevelige, P.E.; Luo, M. Structural and Functional Characterization of the Mumps Virus Phosphoprotein. *J. Virol.* **2013**, *87*, 7558–7568. [[CrossRef](#)]
23. 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](#)]
24. Pickar, A.; Elson, A.; Yang, Y.; Xu, P.; Luo, M.; He, B. Oligomerization of Mumps Virus Phosphoprotein. *J. Virol.* **2015**, *89*, 11002–11010. [[CrossRef](#)]
25. Douglas, J.; Drummond, A.J.; Kingston, R.L. Evolutionary history of cotranscriptional editing in the paramyxoviral phosphoprotein gene. *Virus Evol.* **2021**, *7*, veab028. [[CrossRef](#)] [[PubMed](#)]
26. Du Pont, V.; Jiang, Y.; Plemper, R.K. Bipartite interface of the measles virus phosphoprotein X domain with the large polymerase protein regulates viral polymerase dynamics. *PLoS Pathog.* **2019**, *15*, e1007995. [[CrossRef](#)] [[PubMed](#)]
27. Longhi, S.; Receveur-Brechot, V.; Karlin, D.; Johansson, K.; Darbon, H.; Bhella, D.; Yeo, R.; Finet, S.; Canard, B. The C-terminal Domain of the Measles Virus Nucleoprotein Is Intrinsically Disordered and Folds upon Binding to the C-terminal Moiety of the Phosphoprotein. *J. Biol. Chem.* **2003**, *278*, 18638–18648. [[CrossRef](#)] [[PubMed](#)]
28. Kingston, R.L.; Hamel, D.J.; Gay, L.S.; Dahlquist, F.W.; Matthews, B.W. Structural basis for the attachment of a paramyxoviral polymerase to its template. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8301–8306. [[CrossRef](#)]
29. Houben, K.; Marion, D.; Tarbouriech, N.; Ruigrok, R.W.H.; Blanchard, L. Interaction of the C-Terminal Domains of Sendai Virus N and P Proteins: Comparison of Polymerase-Nucleocapsid Interactions within the Paramyxovirus Family. *J. Virol.* **2007**, *81*, 6807–6816. [[CrossRef](#)]
30. Communie, G.; Habchi, J.; Yabukarski, F.; Blocquel, D.; Schneider, R.; Tarbouriech, N.; Papageorgiou, N.; Ruigrok, R.W.H.; Jamin, M.; Jensen, M.R.; et al. Atomic Resolution Description of the Interaction between the Nucleoprotein and Phosphoprotein of Hendra Virus. *PLoS Pathog.* **2013**, *9*, e1003631. [[CrossRef](#)]
31. Sourimant, J.; Thakkar, V.D.; Cox, R.M.; Plemper, R.K. Viral evolution identifies a regulatory interface between paramyxovirus polymerase complex and nucleocapsid that controls replication dynamics. *Sci. Adv.* **2020**, *6*, eaaz1590. [[CrossRef](#)]
32. Johansson, K.; Bourhis, J.-M.; Campanacci, V.; Cambillau, C.; Canard, B.; Longhi, S. Crystal Structure of the Measles Virus Phosphoprotein Domain Responsible for the Induced Folding of the C-terminal Domain of the Nucleoprotein. *J. Biol. Chem.* **2003**, *278*, 44567–44573. [[CrossRef](#)] [[PubMed](#)]
33. Blanchard, L.; Tarbouriech, N.; Blackledge, M.; Timmins, P.; Burmeister, W.P.; Ruigrok, R.W.; Marion, D. Structure and dynamics of the nucleocapsid-binding domain of the Sendai virus phosphoprotein in solution. *Virology* **2004**, *319*, 201–211. [[CrossRef](#)] [[PubMed](#)]
34. Kingston, R.L.; Gay, L.S.; Baase, W.S.; Matthews, B.W. Structure of the Nucleocapsid-Binding Domain from the Mumps Virus Polymerase; an Example of Protein Folding Induced by Crystallization. *J. Mol. Biol.* **2008**, *379*, 719–731. [[CrossRef](#)]
35. Yegambaram, K.; Bulloch, E.; Kingston, R.L. Protein domain definition should allow for conditional disorder. *Protein Sci.* **2013**, *22*, 1502–1518. [[CrossRef](#)]
36. Hausrath, A.C.; Kingston, R.L. Conditionally disordered proteins: Bringing the environment back into the fold. *Cell. Mol. Life Sci.* **2017**, *74*, 3149–3162. [[CrossRef](#)] [[PubMed](#)]
37. Pace, C.N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **1995**, *4*, 2411–2423. [[CrossRef](#)]
38. Webby, M.N.; Sullivan, M.P.; Yegambaram, K.M.; Radjainia, M.; Keown, J.; Kingston, R.L. A method for analyzing the composition of viral nucleoprotein complexes, produced by heterologous expression in bacteria. *Virology* **2019**, *527*, 159–168. [[CrossRef](#)]
39. Valdes, R.; Ackers, G.K. Study of protein subunit association equilibria by elution gel chromatography. *Methods Enzymol.* **1979**, *61*, 125–142. [[CrossRef](#)]
40. Winzor, D.J. Analytical exclusion chromatography. *J. Biochem. Biophys. Methods* **2003**, *56*, 15–52. [[CrossRef](#)]
41. Schuck, P. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys. J.* **2000**, *78*, 1606–1619. [[CrossRef](#)]
42. Dam, J.; Schuck, P. Calculating Sedimentation Coefficient Distributions by Direct Modeling of Sedimentation Velocity Concentration Profiles. *Methods Enzymol.* **2004**, *384*, 185–212. [[CrossRef](#)]
43. Laue, T.M.; Shah, B.D.; Ridgeway, T.M.; Pelletier, S.L. Computer-Aided Interpretation of Analytical Sedimentation Data for Proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science*; Harding, S.E., Rowe, A.J., Horton, J.C., Eds.; Royal Society of Chemistry: London, UK, 1992.
44. Schuck, P. On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal. Biochem.* **2003**, *320*, 104–124. [[CrossRef](#)]
45. Vistica, J.; Dam, J.; Balbo, A.; Yikilmaz, E.; Mariuzza, R.A.; Rouault, T.A.; Schuck, P. Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. *Anal. Biochem.* **2004**, *326*, 234–256. [[CrossRef](#)]
46. Taylor, I.A.; Eccleston, J.F.; Rittinger, K.; Fu, H. Sedimentation Equilibrium Studies. *Methods Mol. Biol.* **2004**, *261*, 119–136. [[CrossRef](#)]
47. Schuck, P. Sedimentation Equilibrium Analytical Ultracentrifugation for Multicomponent Protein Interactions. In *Protein Interactions*; Schuck, P., Ed.; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2007; pp. 289–316.

48. Petoukhov, M.V.; Franke, D.; Shkumatov, A.V.; Tria, G.; Kikhney, A.G.; Gajda, M.; Gorba, C.; Mertens, H.D.T.; Konarev, P.; Svergun, D.I. New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **2012**, *45*, 342–350. [[CrossRef](#)]
49. Konarev, P.; Volkov, V.V.; Sokolova, A.; Koch, M.H.J.; Svergun, D.I. PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **2003**, *36*, 1277–1282. [[CrossRef](#)]
50. Svergun, D. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **1992**, *25*, 495–503. [[CrossRef](#)]
51. Blanchet, C.E.; Svergun, D. Small-Angle X-Ray Scattering on Biological Macromolecules and Nanocomposites in Solution. *Annu. Rev. Phys. Chem.* **2013**, *64*, 37–54. [[CrossRef](#)] [[PubMed](#)]
52. Sharp, P.; Bloomfield, V.A. Light scattering from wormlike chains with excluded volume effects. *Biopolymers* **1968**, *6*, 1201–1211. [[CrossRef](#)] [[PubMed](#)]
53. Brulet, A.; Boue, F.; Cotton, J.P. About the Experimental Determination of the Persistence Length of Wormlike Chains of Polystyrene. *J. Phys. II Fr.* **1996**, *6*, 885–891. [[CrossRef](#)]
54. Pötschke, D. Analysis of the conformation of worm-like chains by small-angle scattering: Monte-Carlo simulations in comparison to analytical theory. *Macromol. Theory Simul.* **2000**, *9*, 345–353. [[CrossRef](#)]
55. Pérez, J.; Vachette, P.; Russo, D.; Desmadril, M.; Durand, D. Heat-induced unfolding of neocarzinostatin, a small all- $\beta$  protein investigated by small-angle X-ray scattering. *J. Mol. Biol.* **2001**, *308*, 721–743. [[CrossRef](#)] [[PubMed](#)]
56. Boze, H.; Marlin, T.; Durand, D.; Pérez, J.; Vernhet, A.; Canon, F.; Sarni-Manchado, P.; Cheynier, V.; Cabane, B. Proline-Rich Salivary Proteins Have Extended Conformations. *Biophys. J.* **2010**, *99*, 656–665. [[CrossRef](#)]
57. Receveur-Brechot, V.; Durand, D. How Random are Intrinsically Disordered Proteins? A Small Angle Scattering Perspective. *Curr. Protein Pept. Sci.* **2012**, *13*, 55–75. [[CrossRef](#)] [[PubMed](#)]
58. O'Brien, D.P.; Hernandez, B.; Durand, D.; Hourdel, V.; Sotomayor-Pérez, A.-C.; Vachette, P.; Ghomi, M.; Chamot-Rooke, J.; Ladant, D.; Brier, S.; et al. Structural models of intrinsically disordered and calcium-bound folded states of a protein adapted for secretion. *Sci. Rep.* **2015**, *5*, 14223. [[CrossRef](#)]
59. Rawiso, M.; Duplessix, R.; Picot, C. Scattering function of polystyrene. *Macromolecules* **1987**, *20*, 630–648. [[CrossRef](#)]
60. Häckel, M.; Hinz, H.J.; Hedwig, G.R. Partial molar volumes of proteins: Amino acid side-chain contributions derived from the partial molar volumes of some tripeptides over the temperature range 10–90 degrees C. *Biophys. Chem.* **1999**, *82*, 35–50. [[CrossRef](#)]
61. Herr, N.; Webby, M.N.; Bulloch, E.; Schmitz, M.; Kingston, R.L. NMR chemical shift assignment of the C-terminal region of the Menangle virus phosphoprotein. *Biomol. NMR Assign.* **2019**, *13*, 195–199. [[CrossRef](#)] [[PubMed](#)]
62. Roche, J.; Ying, J.; Shen, Y.; Torchia, D.A.; Bax, A. ARTSY-J: Convenient and precise measurement of  $^3\text{J}_{\text{HNH}\alpha}$  couplings in medium-size proteins from TROSY-HSQC spectra. *J. Magn. Reson.* **2016**, *268*, 73–81. [[CrossRef](#)]
63. Delaglio, F.; Grzesiek, S.; Vuister, G.W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995**, *6*, 277–293. [[CrossRef](#)]
64. Niklasson, M.; Otten, R.; Ahlner, A.; Andresen, C.; Schlagnitweit, J.; Petzold, K.; Lundström, P. Comprehensive analysis of NMR data using advanced line shape fitting. *J. Biomol. NMR* **2017**, *69*, 93–99. [[CrossRef](#)]
65. Willard, L.; Ranjan, A.; Zhang, H.; Monzavi, H.; Boyko, R.F.; Sykes, B.D.; Wishart, D.S. VADAR: A web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res.* **2003**, *31*, 3316–3319. [[CrossRef](#)] [[PubMed](#)]
66. Vogeli, B.; Ying, J.; Grishaev, A.; Bax, A. Limits on Variations in Protein Backbone Dynamics from Precise Measurements of Scalar Couplings. *J. Am. Chem. Soc.* **2007**, *129*, 9377–9385. [[CrossRef](#)] [[PubMed](#)]
67. Kingston, R.L.; Baker, H.M.; Baker, E.N. Search designs for protein crystallization based on orthogonal arrays. *Acta Crystallogr. D* **1994**, *50*, 429–440. [[CrossRef](#)] [[PubMed](#)]
68. Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **1997**, *276*, 307–326. [[CrossRef](#)]
69. McCoy, A.J.; Grosse-Kunstleve, R.W.; Adams, P.; Winn, M.D.; Storoni, L.C.; Read, R. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674. [[CrossRef](#)]
70. Emsley, P.; Lohkamp, B.; Scott, W.; Cowtan, K.D. Features and development of Coot. *Acta Crystallogr. D* **2010**, *66*, 486–501. [[CrossRef](#)]
71. Murshudov, G.N.; Skubák, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.; Winn, M.D.; Long, F.; Vagin, A.A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **2011**, *67*, 355–367. [[CrossRef](#)]
72. Stevens, F.J.; Schiffer, M. Computer simulation of protein self-association during small-zone gel filtration. Estimation of equilibrium constants. *Biochem. J.* **1981**, *195*, 213–219. [[CrossRef](#)]
73. Nenortas, E.; Beckett, D. Reduced-Scale Large-Zone Analytical Gel Filtration Chromatography for Measurement of Protein Association Equilibria. *Anal. Biochem.* **1994**, *222*, 366–373. [[CrossRef](#)]
74. Bailey, G.D.; Hyun, J.K.; Mitra, A.; Kingston, R.L. Proton-Linked Dimerization of a Retroviral Capsid Protein Initiates Capsid Assembly. *Structure* **2009**, *17*, 737–748. [[CrossRef](#)] [[PubMed](#)]
75. Schuck, P.; Radu, C.; Ward, E. Sedimentation equilibrium analysis of recombinant mouse FcRn with murine IgG1. *Mol. Immunol.* **1999**, *36*, 1117–1125. [[CrossRef](#)]
76. Rambo, R.P.; Tainer, J. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **2011**, *95*, 559–571. [[CrossRef](#)]

77. Kohn, J.E.; Millett, I.S.; Jacob, J.; Zagrovic, B.; Dillon, T.M.; Cingel, N.; Dothager, R.S.; Seifert, S.; Thiyagarajan, P.; Sosnick, T.R.; et al. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12491–12496. [[CrossRef](#)]
78. Feigin, L.A.; Svergun, D.I. *Structure Analysis by Small-Angle Neutron and X-ray Scattering*; Plenum Press: New York, NY, USA, 1987.
79. Uversky, V.N.; Gillespie, J.R.; Fink, A.L. Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins* **2000**, *41*, 415–427. [[CrossRef](#)]
80. Müller-Späth, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Rügger, S.; Reymond, L.; Nettels, D.; Schuler, B. Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14609–14614. [[CrossRef](#)]
81. Kyte, J.; Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105–132. [[CrossRef](#)]
82. Garnier, J.; Gibrat, J.-F.; Robson, B. GOR method for predicting protein secondary structure from amino acid sequence. *Methods Enzymol.* **1996**, *266*, 540–553. [[CrossRef](#)] [[PubMed](#)]
83. Tamiola, K.; Mulder, F.A.A. Using NMR chemical shifts to calculate the propensity for structural order and disorder in proteins. *Biochem. Soc. Trans.* **2012**, *40*, 1014–1020. [[CrossRef](#)] [[PubMed](#)]
84. Ishima, R.; Torchia, D.A. Estimating the time scale of chemical exchange of proteins from measurements of transverse relaxation rates in solution. *J. Biomol. NMR* **1999**, *14*, 369–372. [[CrossRef](#)]
85. Millet, O.; Loria, J.P.; Kroenke, C.D.; Pons, M.; Palmer, A.G., III. The Static Magnetic Field Dependence of Chemical Exchange Linebroadening Defines the NMR Chemical Shift Time Scale. *J. Am. Chem. Soc.* **2000**, *122*, 2867–2877. [[CrossRef](#)]
86. Baxter, N.J.; Williamson, M.P. Temperature dependence of <sup>1</sup>H chemical shifts in proteins. *J. Biomol. NMR* **1997**, *9*, 359–369. [[CrossRef](#)] [[PubMed](#)]
87. Moon, A.F.; Mueller, G.; Zhong, X.; Pedersen, L. A synergistic approach to protein crystallization: Combination of a fixed-arm carrier with surface entropy reduction. *Protein Sci.* **2010**, *19*, 901–913. [[CrossRef](#)]
88. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Meng, E.C.; Couch, G.S.; Croll, T.I.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **2021**, *30*, 70–82. [[CrossRef](#)] [[PubMed](#)]
89. Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637. [[CrossRef](#)]
90. Shen, Y.; Bax, A. Identification of helix capping and  $\beta$ -turn motifs from NMR chemical shifts. *J. Biomol. NMR* **2012**, *52*, 211–232. [[CrossRef](#)]
91. Salvador, P. Dependencies of J-Couplings upon Dihedral Angles on Proteins. *Annu. Rep. NMR Spectrosc.* **2014**, *81*, 185–227. [[CrossRef](#)]
92. Guseva, S.; Milles, S.; Jensen, M.R.; Schoehn, G.; Ruigrok, R.W.; Blackledge, M. Structure, dynamics and phase separation of measles virus RNA replication machinery. *Curr. Opin. Virol.* **2020**, *41*, 59–67. [[CrossRef](#)]
93. Zhou, Y.; Su, J.M.; Samuel, C.E.; Ma, D. Measles Virus Forms Inclusion Bodies with Properties of Liquid Organelles. *J. Virol.* **2019**, *93*. [[CrossRef](#)]
94. Ma, D.; George, C.X.; Nomburg, J.L.; Pfaller, C.K.; Cattaneo, R.; Samuel, C.E. Upon Infection, Cellular WD Repeat-Containing Protein 5 (WDR5) Localizes to Cytoplasmic Inclusion Bodies and Enhances Measles Virus Replication. *J. Virol.* **2018**, *92*. [[CrossRef](#)] [[PubMed](#)]
95. Fearn, R.; Young, D.F.; Randall, R. Evidence that the paramyxovirus simian virus 5 can establish quiescent infections by remaining inactive in cytoplasmic inclusion bodies. *J. Gen. Virol.* **1994**, *75*, 3525–3539. [[CrossRef](#)] [[PubMed](#)]
96. Su, J.M.; Wilson, M.Z.; Samuel, C.E.; Ma, D. Formation and Function of Liquid-Like Viral Factories in Negative-Sense Single-Stranded RNA Virus Infections. *Viruses* **2021**, *13*, 126. [[CrossRef](#)] [[PubMed](#)]
97. Bernado, P.; Blanchard, L.; Timmins, P.; Marion, D.; Ruigrok, R.W.H.; Blackledge, M. A structural model for unfolded proteins from residual dipolar couplings and small-angle X-ray scattering. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17002–17007. [[CrossRef](#)]
98. Houben, K.; Blanchard, L.; Blackledge, M.; Marion, D. Intrinsic Dynamics of the Partly Unstructured PX Domain from the Sendai Virus RNA Polymerase Cofactor P. *Biophys. J.* **2007**, *93*, 2830–2844. [[CrossRef](#)] [[PubMed](#)]
99. Guseva, S.; Milles, S.; Jensen, M.R.; Salvi, N.; Kleman, J.-P.; Maurin, D.; Ruigrok, R.W.H.; Blackledge, M. Measles virus nucleocapsid and phosphoproteins form liquid-like phase-separated compartments that promote nucleocapsid assembly. *Sci. Adv.* **2020**, *6*, eaaz7095. [[CrossRef](#)] [[PubMed](#)]