The C-terminal Domain of the Measles Virus Nucleoprotein Is Intrinsically Disordered and Folds upon Binding to the C-terminal Moiety of the Phosphoprotein* 

Sonia Longhi‡‡, Véronique Receveur-Bréchot‡, David Karlin‡‡, Kenth Johansson‡‡, Hervé Darbon‡, David Bhella**, Robert Yeo**, Stéphanie Finet‡‡, and Bruno Canard‡

From the ‡Architecte et Fonction des Macromolécules Biologiques, UMR 6988 CNRS et Université Aix-Marseille 1 et II, ESIL, Campus de Luminy, 13288 Marseille Cedex 9, France, **ESIL Virology Unit, Institute of Virology, Church St., Glasgow G11 5JR, United Kingdom, and ‡‡European Synchrotron Radiation Facility, BP 220, 38043 Grenoble Cedex, France

The nucleoprotein of measles virus consists of an N-terminal moiety, N CORE, resistant to proteolysis and a C-terminal moiety, N TAIL, hypersensitive to proteolysis and not visible as a distinct domain by electron microscopy. We report the bacterial expression, purification, and characterization of measles virus N TAIL. Using nuclear magnetic resonance, circular dichroism, gel filtration, dynamic light scattering, and small angle x-ray scattering, we show that N TAIL is not structured in solution. Its sequence and spectroscopic and hydrodynamic properties indicate that N TAIL belongs to the pre-molten globule subfamily within the class of intrinsically disordered proteins. The same epitopes are exposed in N TAIL and within the nucleoprotein, which rules out dramatic conformational changes in the isolated N TAIL domain compared with the full-length nucleoprotein. Most unstructured proteins undergo some degree of folding upon binding to their partners, a process termed "induced folding." We show that N TAIL is able to bind its physiological partner, the phosphoprotein, and that it undergoes such an unstructured-to-structured transition upon binding to the C-terminal moiety of the phosphoprotein. The presence of flexible regions at the surface of the viral nucleocapsid would enable plastic interactions with several partners, whereas the gain of structure arising from induced folding would lead to modulation of these interactions. These results contribute to the study of the emerging field of natively unfolded proteins.

Measles virus (MV) is an enveloped RNA virus within the Morbillivirus genus of the Paramyxoviridae family. Its nonsegmented, negative sense, single-stranded RNA genome is packaged by the viral nucleoprotein (N) within a helical nucleocapsid. Transcription and replication are carried out on this (N-RNA) complex by the RNA-dependent RNA polymerase (L) associated with the phosphoprotein (P) (reviewed in Refs. 1 and 2). During genome replication, synthesis of viral RNA and encapsidation by N are concomitant (1, 2). However, N has the capacity to self-assemble on cellular RNA to form nucleocapsid-like particles in the absence of viral RNA and of any other viral protein (3–7). Therefore, a regulatory mechanism is necessary to prevent illegitimate self-assembly of N onto cellular RNA in the absence of ongoing genomic RNA synthesis. This role is played by P; association of P with the soluble, monomeric form of N (N*) prevents its illegitimate self-assembly (8–10). This soluble N*-P complex is the substrate used by the polymerase to initiate encapsidation of genomic RNA (10, 11). N forms complexes with P and with the PL complex during transcription and replication also in its self-assembled form (N*P*) (12–15).

Deletion analyses have shown that nucleoproteins of Paramyxoviridae are divided into two regions: a N-terminal moiety, N CORE, well conserved in sequence, and a hypervariable, C-terminal moiety, N TAIL (2) (see Fig. 1A). N CORE contains all of the regions necessary for self-assembly and RNA binding (16, 17). Within N CORE, deletion and mutational studies have shown that the central conserved region (MV; aa 258–357) and the aa 189–239 region, are involved in self-association and RNA binding (4, 19, 21). The precise location of RNA binding sites within N CORE is unknown. The RNA binding site is probably formed by maturation of N during encapsidation, as suggested by the observation that all N mutants displaying an impaired self-association are not able to package RNA (4, 19, 21).

Within nucleocapsids, N TAIL protrudes from the globular body of N CORE (22, 23). In Morbilliviruses and Respiroviruses, the region responsible for binding of nucleocapsids to P is located within N TAIL (13, 18, 20, 24). Consequently, N TAIL is required for N-RNA to act as a template for viral RNA synthe-
Induced Folding of the C-terminal Domain of MV N

During the last decade, it has been shown that a considerable number of proteins (>100) have little or no ordered rigid structure under physiological conditions. These proteins are referred to as natively unfolded or “intrinsically disordered or unstructured” (for reviews, see Refs. 30–32). Conformational analysis has shown that natively unfolded proteins do not represent a uniform family, but rather they can be subdivided into two structurally distinct groups. The first group consists of proteins with extended hydrodynamic dimensions typical of random coils with no (or little) ordered secondary structure. The second group comprises the so-called premolten globules, which are more compact (but still less compact than globular or molten globule proteins) and conserve some residual ordered secondary structure (31, 32).

It has recently been proposed that the protein structure-function paradigm should be better replaced by the “protein quartet model” (31). According to this model, proteins can exist in any of the four thermodynamic states of proteins (ordered, molten globule, premolten globule, and random coil), and function can arise from any of these conformations and transitions between them. The functional importance of disorder may reside in advantages of flexible structures in comparison with rigid structures. In particular, an increased plasticity would (i) allow to couple a high specificity with a low affinity, (ii) enable binding of numerous structurally distinct targets, and (iii) provide the ability to overcome steric restrictions, enabling larger interaction surfaces in protein-protein and protein-ligand complexes than those obtained with rigid partners (31–35).

In this paper, we report the bacterial expression, purification, and characterization of MV N TAIL. Using different, complementary biochemical and biophysical methods, we show that N TAIL is intrinsically disordered and belongs to the premolten globule subfamily. We also show that N TAIL is able to bind its physiological partner, P. Although there are intrinsically disordered proteins that carry out their function while being disordered from end to end (32), a majority folds in the presence of their physiological partner(s) (36, 37), a process known as “induced folding.” We show that N TAIL undergoes such an unstructured-to-structured transition upon binding to the C-terminal moiety of P (PCT).

EXPERIMENTAL PROCEDURES

Bacterial Strains and Media—The Escherichia coli strain XL1 Blue (Stratagene) was used for selection and amplification of DNA constructs. The E. coli strains BL21(DE3) and BL21 (Novagen) were used for expression of recombinant proteins. E. coli was grown in Luria-Bertani medium.

Chemicals and Antibodies—Restriction enzymes and T4 DNA ligase were purchased from NE Biolabs. Pfu polymerase was from Promega. Primers were purchased from Genosys. Nucleotide sequencing was carried out by Genome Express. The anti-hexahistidine tag monoclonal antibody (mAb) was purchased from Qiagen. The anti-N CI 105 2 and CI 25 (38, 39) and the anti-P 49.21 mAbs (40) were kindly provided by D. Gerlier.

Plasmids—The plasmid pQE32/NTAIL, encoding residues 400–525 of MV N (strain Edmonston B) with a hexahistidine tag fused to its N terminus was a kind gift of C. Rabourdin-Combe. The plasmids pET21a/P CT and pET21a/P encoding the MV P protein (strain Edmonston B) with or without an N-terminal hexahistidine tag, respectively, have been already described (41). The PCT gene construct, encoding the C-terminal moiety of P (PCT) (residues 231–507 of P) with a hexahistidine tag fused to its N terminus, was obtained by PCR, using pET21/P CT as template. Forward primer (5′-GGCCCATATGGCCATCACTCATCATCATGCGCGCCACGAGCGGAGATTAC-3′) was designed to introduce an NdeI restriction site fused to the N terminus of PCT. Reverse reverse primer (5′-GGCGGATCCATCTATCTTCTATTAC-3′) was designed to introduce an EcoRI site after the P stop codon. The PCR product was digested with NdeI and EcoRI, purified, and ligated between the NdeI and EcoRI sites of pET21a to yield pET21a/PCT.

The plasmid pET21a/N FLAG-H6, encoding the N-terminal domain of P (PNT) (residues 1–290 of P) with an N-terminal hexahistidine tag, has been already described (41).

The plasmid pET21a/N FLAG-H6, which encodes MV N with an N-terminal FLAG sequence (DYKD DDK DDYK) and a C-terminal hexahistidine tag, has already been described (4).

Expression of N TAIL and PCT—E. coli strains BL21 and BL21(DE3) were used for the expression of pET21/N TAIL and of pET21a-derivative constructs, respectively. Since MV N and P genes contain several rare arginine codons (AGG) that are used with a very low frequency in E. coli, co-expression of N and P constructs with the plasmid pUBSS20 (43), which supplies the corresponding rare tRNA, was carried out. Cultures were grown overnight to saturation in Luria-Bertani (LB) medium containing 100 μg/ml ampicillin and 50 μg/ml kanamycin. An aliquot of the overnight culture was diluted 1:25 in LB medium and grown at 37 °C. At 0.007 of OD, isopropyl β-thiogalactopyranoside was added to a final concentration of 0.2 mm, and the cells were grown at 37 °C for 3 h. The induced cells were harvested, washed, and collected by centrifugation. The resulting pellets were frozen at −20 °C.

Expression of tagged N, PNT, and untagged P was carried out as described in Refs. 25 and 48.

Purification of N TAIL and PCT—The pellet containing N TAIL was resuspended in 5 volumes (v/v) buffer A (50 mm sodium phosphate, pH 8, 300 mm NaCl, 10 mm imidazole, 10 mm phenylmethylsulfonyl fluoride (PMSF)) supplemented with 0.1 mg/ml lysozyme, 10 μg/ml DNase I, and protease inhibitor mixture (Sigma) (50 μg/ml of cells). After a 20-min incubation with gentle agitation, the cells were disrupted by sonication (using a 750-watt sonicator and four cycles of 30 s each at 60% power output). The lysate was clarified by centrifugation at 30,000 × g for 30 min. Starting from a 1-liter culture, the clarified supernatant was incubated for 1 h with gentle shaking with 4 ml of cleaving Sepharose 4B resin (Amersham Biosciences), previously equilibrated in buffer A. The resin was washed with buffer A, and N TAIL was eluted in buffer A containing 250 mm imidazole. Eluates were analyzed by SDS-PAGE for the presence of N TAIL. The fractions containing N TAIL were combined; diluted with 5 volumes of 50 mm sodium phosphate, pH 8, 1 mm PMSF, and incubated for 1 h with gentle shaking with 4 ml of Sepharose Q Fast Flow resin (Amersham Biosciences). The flow-through was recovered, filtered onto 0.8-μm filters (Pall Gelman Laboratories), and concentrated using a Centricron Plus-20 concentrator (molecular mass cut-off, 5000 Da) (Millipore Corp.). The protein was then loaded onto a Superdex 75 HR 10/30 column (Amersham Biosciences) and eluted in either 10 mm sodium phosphate, pH 8, or 10 mm Tris/HCl, pH 8, 5 mm EDTA. The protein was stored at −20 °C in the presence of 10% glycerol.

The pellet containing PCT was resuspended in 5 volumes (v/v) buffer B (10 mm Tris/HCl, pH 8, 300 mm NaCl, 10 mm imidazole, 10% glycerol, 1 mm PMSF) supplemented with 0.1 mg/ml lysozyme, 10 μg/ml DNase I, and protease inhibitor mixture (Sigma) (50 μg/ml of cells). After a 20-min incubation with gentle agitation, the lysate was clarified as described above. Starting from a 1-liter culture, the clarified supernatant was incubated for 1 h with gentle shaking with 4 ml of Tiron-Sepharose Fast Flow resin (Amersham Biosciences). The flow-through was recovered, filtered onto 0.8-μm filters (Pall Gelman Laboratories), and concentrated using a Centricron Plus-20 concentrator (molecular mass cut-off, 5000 Da) (Millipore Corp.). The eluate from the affinity chromatography step was then loaded onto a Superdex 200 HR 10/30...
column (Amerham Biosciences) and eluted in 10 mM Tris/Cl, pH 8, 300 mM NaCl, 5 mM EDTA, 10% glycerol, and 1 mM PMSF. The protein was stored at −20 °C.

Purification of PNT and of N were carried out as described in (25, 48). The 43-kDa N-terminal fragment of N, NCORE, was obtained by limited trypsin digestion of N as described in Ref. 25, followed by gel filtration onto a Superdex 200 HR 10/30 column (Amerham Biosciences). NCORE was eluted with 50 mM Tris/Cl, pH 7.5, 1 mM PMSF and stored at −20 °C in the presence of 20% glycerol.

All purification steps, except for gel filtrations, were carried out at 4 °C.

Apparent molecular mass of proteins eluted from gel filtration columns was deduced from a calibration carried out with low and high molecular weight gel filtration kits (Amersham Biosciences). Theoretical Stokes radii (Rs) of a native (RN) and fully unfolded (RU) protein with a molecular mass (MM) (in daltons) were calculated according to Ref. 44: log(Rn) = 0.369 × log(MM) − 0.254, and log(RU) = 0.533 × log(MM) − 0.682. The hydrodynamic volume (V) was calculated from the Stokes radius as V = 4/3πRs3. The theoretical hydrodynamic volumes of a native (VN), natively unfolded random coil (VNrandom), and natively unfolded premonoblob globule (VNpreglobule) of n residues were calculated according to Ref. 31: log(VN) = (−2.197 ± 0.037) + (1.072 ± 0.015) × log n, log(VNrandom) = (1.997 ± 0.078) + (1.498 ± 0.035) × log n, and log(VNpreglobule) = (2.33 ± 0.12) + (1.234 ± 0.047) × log n.

**Determination of Protein Concentration**—Protein concentrations were determined using the molar absorption coefficients ε (mg/ml cm) at 280 nm as obtained using the program ProtParam at the EXPASY server (available on the World Wide Web at www.expasy.ch/tools), or the Bio-Rad protein assay reagent (Bio-Rad).

**Mass Spectrometry**—Mass analysis of PCT was performed using a Voyager DE FT mass spectrometer (PerSeptive Biosystems). Samples (0.7 µl containing 15 pmol) were mixed with an equal volume of sinapinic acid matrix solution, spotted on the target, and then dried at room temperature for 10 min. The mass standard was aposyoglobin.

Mass analysis of tryptic fragments of NTAIL was carried out by digesting (0.25 µg of trypsin) 1 µg of purified NTAIL, obtained after separation onto 12% SDS-PAGE. The experimental mass values of the tryptic fragment determined as described above (before calibration with theoretical values found in protein data bases (available on the World Wide Web at www.matrixscience.com and www.expasy.ch/cgi-bin/findmod0.pl)). The mass standards were either autolytic trypsin peptides or keratin peptides.

**Dynamic Light Scattering (DLS)**—Dynamic light scattering experiments were performed with a Dynapro MTCST-200 (Protein Solutions) at 20 °C. All samples were filtered prior to the measurements (Millex syringe filters, 0.22 µm; Millipore Corp.). The hydrodynamic radius was deduced from translational diffusion coefficients using the Stokes-Einstein equation. Diffusion coefficients were inferred from the analysis of the decay of the scattered intensity autocorrelation function. All calculations were performed using the software provided by the manufacturer.

**Two-dimensional X-ray scattering (2D-XS)**—A sample containing purified NTAIL at a concentration of 0.5 mg in 10 mM sodium phosphate, pH 7, was used for the acquisition of a NOEY spectrum on a DRX500 Bruker spectrometer at 300 K with 2048 complex points in the directly acquired dimension and 512 complex points in the indirectly detected dimension. Solution suppression was achieved by the WATERGATE 3:9:19 pulse (45). The data were processed using the XNMR software; they were multiplied by a sine-squared bell and zero-filled to 1 K in the first dimension prior to Fourier transformation.

**Circular Dichroism**—The CD spectra were recorded on a Jasco 810 diode array spectrophotometer using thick quartz cells in 10 mM sodium phosphate, pH 7, at 20 °C. Structural variations of NTAIL were measured as a function of changes in the initial CD spectrum upon the addition of either increasing concentrations of 2,2,2-trifluoroethanol (TFE) (Fluka) or different amounts of PCT, PNT, or lysozyme (Sigma). CD spectra were measured between either 185 and 260 nm or 190 and 260 nm at 0.015 °C. The molar ellipticities (ε) were averaged from three independent measurements. Protein concentrations ranging from 0.1 to 0.15 mg/ml were used when recording spectra of protein mixtures. In this case, mean ellipticity values per residue (εav) were calculated as [ε] = 3300 Å² cm²/dmole, where I represents path length (0.1 cm), n is the number of residues, m is the molecular mass in daltons, and c is the protein concentration expressed in mg/ml. Numbers of residues (n) are 139 for NTAIL, 236 for PNT, 284 for PCT, and 129 for lysozyme, whereas m values are 16,804 Da for NTAIL, 24,800 Da for PNT, 30,800 Da for PCT, and 14,300 Da for lysozyme. Protein concentrations of 0.1 mg/ml were used when recording individual spectra. Protein concentrations ranging from 0.1 to 0.15 mg/ml were used when recording spectra of protein mixtures. In this case, mean ellipticity values per residue (εav) were calculated as [ε] = 3300 Å² cm²/dmole, where I represents path length (0.1 cm), n is the number of residues, m is the molecular mass in daltons, and c1 or c2 is protein concentration expressed in mg/ml for each of the two proteins in the mixture. The theoretical average ellipticity values per residue (θav) expected, assuming that neither any unstructured-to-structured transition nor any secondary structure rearrangement occurs, were calculated as follows:

θav = (θ2 × n2 + θ1 × n1)/ (n1 + n2), where θ1 and θ2 correspond to the measured mean ellipticity values per residue, n1 and n2 to the number of residues for each of the two proteins, and R to the excess molar ratio of protein 2. The α-helical content was derived from the ellipticity at 220 nm as described in Ref. 46.

**Small Angle X-ray Scattering (SAXS)**—SAXS experiments were carried out on beamline ID2 (European Synchrotron Radiation Facility, Grenoble, France). Samples of purified NTAIL (in buffer containing 10 mM Tris/Cl, pH 8, 5 mM EDTA, with 10% glycerol as radiative scavenger) were filtered prior each measurement (Millex syringe filters (0.22 µm); Millipore). The wavelength was 1.0 Å, and the sample-to-detector distance was 3.0 and 1.0 m, leading to scattering vectors q ranging from 0.02 to 0.20 Å−1 and 0.05 to 0.40 Å−1, respectively. The scattering vector is defined as q = 4πsinθ/λ, where λ is the scattering angle. The detector was an X-ray image intensified optically coupled to a FReLoN CCD camera developed by the European Synchrotron Radiation Facility. 40 successive frames of 0.7 s with a 5-s pause between each frame were recorded for each sample. The protein solution was evaporated using the vapor absorption capillary between each frame. Thus, no protein solution was irradiated longer than 0.7 s. Each frame was then carefully inspected to check for possible bubble formation or radiation-induced aggregation. No such effect was observed, and individual frames could then be averaged. Absolute calibration was made with a Laponol sample. A series of measurements at different protein concentrations ranging from 1.8 to 9 mg/ml were performed to check for interparticle interaction. Background scattering was measured before or after each protein sample using the buffer solution and then subtracted from the protein scattering patterns after proper normalization and correction from detector response. All of the experiments were carried out at 20 °C.

**EM Studies**—5 µl of protein suspension (at ~0.1 mg/ml) was loaded onto a glow-discharged carbon-coated support grid for 30 s, washed in a 50-µl droplet of distilled H2O, and then negatively stained with 2% ammonium molybdate solution (pH 7.5). The structures formed by N and NCORE were imaged in a JEOL 1200 EX II transmission electron microscope and recorded on Kodak SO163 film at ×30,000 magnification under low electron dose conditions.

To make accurate measurements of nucleocapsid dimensions, micrographs were digitized, on a Dünvegan Hi-Scan drum scanner (Düvegan SA, Lausanne, Switzerland), at a raster step size of 10 µm/pixel, corresponding to 3.4 Å/pixel in the specimen. 200 top views of nucleocapsid rings were selected by cross-correlation of scanned micrographs against a ring model of the appropriate dimensions and excised into individual images using the SPIDER image processing package. Rings were centered by cross-correlation against a ring model, and radial density profiles were calculated for each image. Average radial density profiles were calculated and plotted using ORIGIN (Microcal).

**Immunoprecipitation (IP)**—Studies of NTAIL—IP experiments were carried out using the anti-N mAbs Cl 25 and Cl 105 and bacterial lysates expressing either tagged N and N TAIL or no recombinant protein. The different bacterial cultures were induced as described above, and 5-mI aliquots were pelleted and frozen at −20 °C.

The different aliquots were individually resuspended in 500 µl of buffer C (10 mM Tris/HCl, pH 8, 0.1% Igepal, 150 mM NaCl) and sonicated using a 750-watt sonicator and three cycles of 7 s at 35% (output). The lysed samples were centrifuged at 16,000 × g at 4 °C, and the supernatants, containing the soluble fractions, were recovered. Soluble fractions (50 µl) were mixed with 4 µg of either Cl 25 or Cl 105 anti-N mAb, 400 µl of buffer C were added in order to increase the volume during the binding step, and the mixtures
were incubated with 20 μl of Protein G-Sepharose resin (Amersham Biosciences) pre-equilibrated with buffer C. After a 1-h incubation with gentle agitation at 20 °C, the mixtures were centrifuged, and the resin was washed twice with 20 volumes of buffer C. It was then mixed with an equal volume of 2× Laemmli sample buffer, boiled, and loaded onto a 12% SDS-PAGE gel.

**Binding of NT** from P—5-ml aliquots of bacterial cultures expressing either NT or P induced as described above were harvested, and the cellular pellets were frozen at −20 °C.

An NT and a P aliquot were individually resuspended in 500 μl of buffer D (10 mM Tris/HCl, pH 8, 20 mM imidazole, 0.5 M NaCl) and sonicated (using a 750-watt sonicator and three cycles of 7 s at 35% power output). The lysed aliquots were clarified by centrifugation at 16 000 × g for 10 min at 4 °C, and the supernatants, containing the soluble fractions, were recovered. 250 μl of each of the two soluble fractions were mixed and immediately incubated with 50 μl of chelating Sepharose fast flow resin preloaded with Ni2+ ions (Amersham Biosciences), pre-equilibrated with buffer D. After a 1-h incubation at 20 °C with gentle agitation, the mixture was centrifuged, and the resin was washed five times with 10 volumes of buffer D. It was then mixed with 50 μl of 2× Laemmli sample buffer, boiled, and loaded onto a 12% SDS-PAGE gel.

**Sequence Analysis and Secondary Structure Predictions**—The accession number for NT in the Swiss-Prot database is P04851. Deviation in amino acid composition of NT and NCORE were computed using the average amino acid frequencies of the SWISS-PROT data base (as obtained from the World Wide Web at http://expasy.org/sprot) as the reference value. The mean net charge (R) of a protein is determined as the absolute value of the difference between the number of positively and negatively charged residues divided by the total number of amino acid residues. It was calculated using the program ProtParam at the EXPASY server (available on the World Wide Web at www.expasy.ch/tools). The mean hydrophobicity (H) is the sum of normalized hydrophobicities of individual residues divided by the total number of amino acid residues minus 4 residues (to take into account fringe effects in the calculation of hydrophobicity). Individual hydrophobicities were determined using the ProtScale program at the EXPASY server (available on the World Wide Web at www.expasy.ch/tools), using the options “Hphob/Kyte & Doolittle,” a window size of 5, and normalizing the scale from 0 to 1. The values computed for individual residues were then exported to a spreadsheet, summed, and divided by the total number of residues minus 4 to yield H. Hboundary was computed as described by Uversky (31): $H_{\text{boundary}} = \frac{\left| R \right| + 1.15}{2.785}$.

Secondary structure predictions were carried out using the JPRED program (available on the World Wide Web at on. ebi.ac.uk/servers/jpred.html) (52).

**RESULTS**

MV N is a modular protein composed of two regions: a well-conserved N-terminal moiety, NCORE (aa 1–399) and a hyper-variable C-terminal moiety, NT (aa 400–525) (see Fig. 1A). The sequence variability and the sensitivity to proteolysis of NT, (4, 28) are both hallmarks of intrinsic disorder (29, 53). We have thus analyzed the sequence properties of NT in order to check whether this domain possesses the features typical of intrinsically disordered proteins.

**Domain Organization of NT and Sequence Properties of NT**—The amino acid composition of NT and the deviation from the average values in the SWISS-PROT data base are shown in Fig. 1A. NT has a peculiar composition, being depleted in all “order-promoting” residues (Trp, Cys, Phe, Tyr, Ile, Leu) and enriched in most “disorder-promoting” residues (Arg, Gln, Ser, and Glu), as already described for intrinsically disordered proteins (32) (Fig. 1A). Moreover, NT is predicted to be natively unfolded by the PONDR predictor of naturally disordered regions (54), as well by the method based on the mean hydrophobicity/mean net charge ratio (55) (data not shown). Finally, NT has little predicted secondary structure, a feature that has been recently noticed in protein regions with “no ordered regular structure” (56).

In order to directly investigate the structural properties of NT, we have expressed, purified, and characterized the NT domain alone.

**Expression and Purification of NT**—Most NT was recovered from the soluble fraction of bacterial lysates (Fig. 1B, lane SN). NT was purified to homogeneity (>95%) in three steps: immobilized metal affinity chromatography (IMAC), inverse ion exchange chromatography, and gel filtration (Fig. 1B). The identity of the recombinant product was confirmed by mass spectrometry analysis of the tryptic fragments obtained after digestion of purified NT. As shown in Fig. 1B, NT displays an abnormally slow migration in SDS-PAGE even after heat denaturation. In particular, it migrates with an apparent molecular mass (MM) of 20 kDa, whereas the expected MM is 15 kDa. This abnormal behavior has already been
observed in the case of the intrinsically disordered MV PNT (41) as well as in other intrinsically disordered proteins and can most likely be ascribed to their unusual sequence composition (30).

Size Exclusion Chromatography of NTAIL—NTAIL (15 kDa) is eluted from the gel filtration column as a peak corresponding to a 36-kDa globular protein (Fig. 1C). The same profile is obtained regardless of the nature of the buffer used and the presence of different NaCl concentrations (not shown), thus excluding the possibility that the observed behavior could be ascribed to nonspecific interactions of NTAIL with the column matrix. The sharpness of the peak indicates the presence of a well defined species. The elution volume of a protein from a gel filtration column depends on its hydrodynamic properties. The hydrodynamic radius of a protein (Stokes radius ($R_S$)) can be deduced from its apparent MM (as seen by gel filtration) (44). The apparent MM of 36 kDa measured for NTAIL corresponds to a Stokes radius of 27 ± 3 Å. On the other hand, the theoretical Stokes radius of a monomeric, native (RTAIL) or fully unfolded (Rf) protein can be estimated according to the equations described in Ref. 44. In the case of NTAIL, $R_f N = 19 Å$ and $R_S U = 35 Å$. Therefore, the Stokes radius experimentally measured for NTAIL (27 Å) is not compatible with a monomeric, globular protein. Rather, such a large value of the Stokes radius can be attributed either to dimerization or to an extended conformation.

DLS Studies on NTAIL—The hydrodynamic radius of proteins can also be derived by dynamic light scattering studies, since diffusion coefficients of proteins depend upon their size and shape. Globular proteins differ notably from fully or partly unstructured proteins in their hydrodynamic properties, and an empirical relationship between number of residues and hydrodynamic radius $R_g$ has been established for native, globular proteins and for fully denatured proteins (57). The hydrodynamic radius measured for NTAIL by DLS is $R_g = 30 ± 2 Å$, in either 10 mM sodium phosphate, pH 7, or 10 mM Tris/HCl, pH 8.75, mM NaCl at a concentration of 1.25 mg/ml. This value is consistent (within the error bars) with the $R_g$ value measured by gel filtration and corresponds to a hydrodynamic volume ($V$) of ~97,000 Å³. According to the equations described to the authors described by Uversky (31), the theoretical $V$ expected for a native, globular protein composed of 139 residues is about 31,000 Å³, whereas for a fully denatured protein it would be close to 180,000 Å³. Thus, the $V$ of NTAIL calculated by DLS is about 3 times larger than the value expected for a globular protein and half of that expected for a denatured, fully unfolded protein. Rather, these hydrodynamic properties are consistent with the hypothesis that NTAIL is a native premolten globule (31).

CD Studies on NTAIL—The far-UV CD spectrum of NTAIL at neutral pH is typical of an unstructured protein, as seen from its large negative ellipticity at 198 nm and very low ellipticity at 185 nm (Fig. 2B). However, the observed ellipticity values at 220 and 222 nm (~9700 and ~2400 degrees cm² dmol⁻¹, respectively) (Fig. 2B) are consistent with the existence of some residual secondary structure, typical of the premolten globule state (39). Therefore, the spectroscopic and hydrodynamic parameters, together with the peculiar sequence properties of NTAIL, converge to show that NTAIL is mostly unstructured in solution and that it belongs to the class of natively unfolded or intrinsically disordered/unstructured proteins (37, 38, 42, 61, 62).

The majority of intrinsically disordered proteins folds upon binding to their physiological partner(s) (37). The solvent TFE mimics the hydrophobic environment experienced by proteins in protein-protein interactions and is therefore widely used as a probe to discover regions that have a propensity to undergo an induced folding (58). Thus, we recorded CD spectra of NTAIL in the presence of increasing concentrations of TFE (Fig. 2B). NTAIL shows an increasing gain of $α$-helicity upon the addition of TFE, as indicated by the characteristic maximum at 190 nm and minima at 208 and 222 nm (Fig. 2B). Most unstructured-to-structured transitions take place in the presence of 20% TFE, a concentration at which the $α$-helix content is estimated to be about 22% (using the ellipticity at 220 nm). The gain of
α-helicity induced by TFE, although already reported for other proteins, including MV PNT (41), is not a general rule; for instance, the acidic activator domain of GCN4 forms little or no α-helix in TFE concentrations as high as 30% and folds mostly as β-sheets in 50% TFE (59). In the case of N TAIL, the use of TFE reveals a clear α-helix forming potential within this protein domain. Interestingly, this observation is in agreement with the secondary structure prediction made by the JPRED program (52), which predicts an α-helix (residues 489–504) as the only secondary structure element within N TAIL.

SAXS Studies—SAXS studies are particularly well adapted to study flexible, low compactness or even extended macromolecules in solution. They provide low resolution structural data and give access to the mean particle size (radius of gyration, \( R_g \)) as well as to the maximal intramolecular distance (\( D_{\text{MAX}} \)). These two parameters give information on the degree of compactness of the molecule, and the latter gives an idea of the maximal degree of extension reached by the molecule in solution. Fig. 3A shows the Guinier plots of the SAXS data of N TAIL at four different concentrations. Each curve can be well approximated by a straight line in the Guinier region (\( q R_g < 1.0 \)). The slope gives the value of the radius of gyration, \( R_g \), whereas the intercept of the straight line gives the \( I(0) \), which is proportional to the MM of the scatterer. The \( R_g \) extrapolated at zero concentration is 27.5 ± 0.7 Å, and the deduced MM is 15,250 Da, which is in perfect agreement with the expected MM for a monomeric form (15,300 Da). A similar \( R_g \) value (26.9 ± 0.8 Å) has been obtained using the method based on the Debye function (60, 61). According to data already available in the literature (57, 62–64), the expected \( R_g \) of N TAIL would be 15 Å for a globular protein and around 35–38 Å for a denatured, fully unfolded protein. Therefore, the observed \( R_g \) indicates that N TAIL is not globular. However, the protein is more compact than a random coil, suggesting that it possesses some residual structure.

A very useful method to describe the structural properties of a molecule is the Kratky plot. In particular, one can from the shape of this plot infer the conformation adopted by the molecule. The Kratky plot of a globular protein has a typical bell shape with a clear maximum. For a completely unfolded protein or in a premolten globule conformation, no such maximum can be observed, and the curve displays a plateau (65). The Kratky plot of N TAIL displays a bump at \( q R_g < 1.0 \) followed by a plateau for \( q > 0.15 \) Å⁻¹ (see Fig. 3B). The absence of a maximum clearly indicates that N TAIL is not globular and does not possess a tightly packed core. However, the observed bump may be indicative of some residual structure.

The distance distribution function, deduced from the scattering intensities of N TAIL (data not shown), has a bell shape, with a maximum dimension \( D_{\text{MAX}} \) of 120–130 Å. This value, while being indicative of an extended, nonglobular conformation, is lower than expected for a random coil (62, 63).

All in all, the hydrodynamic properties of N TAIL inferred from gel filtration, DLS and SAXS, indicate that N TAIL adopts a typical nonglobular, premolten globule conformation in solution. Accordingly, it is more compact than a random coil and retains some residual secondary structure, in agreement with the CD studies. In conclusion, the spectroscopic and hydrodynamic properties of N TAIL, together with its peculiar sequence composition, converge to show that it is a natively unfolded protein, belonging to the premolten globule subfamily.

Comparison of N and N CORE by EM Studies—N produced in MV-infected cells (66) as well as recombinant MV N (4–6) forms nucleocapsids with a characteristic herringbone structure. In order to investigate whether the N TAIL domain affects the conformation of N within nucleocapsids, we have analyzed by EM the nucleocapsid-like structures formed by either MV N or by a truncated form devoid of the N TAIL moiety (N CORE). N CORE was obtained by limited proteolysis of purified N as described in Ref. 4, followed by gel filtration (data not shown).

EM analysis of full-length N reveals the presence of typical nucleocapsid-like, herringbone structures ranging in length from 20 to 80 nm (Fig. 4A, left panel, arrows). However, most of the nucleocapsid-like particles appear as rings with a mean diameter of 20 nm (Fig. 4A, left panel, feathered arrows). These rings correspond to shorter nucleocapsids viewed along their

Fig. 3. Small angle x-ray scattering experiments on N TAIL. A, Guinier plot of N TAIL in 10 mM Tris, pH 8, 5 mM EDTA, 10% glycerol, at different protein concentrations: 9 mg/ml (filled triangles), 6.7 mg/ml (open squares), 4.5 mg/ml (filled circles), and 1.8 mg/ml (crosses). The slope of the straight lines (shown as a thick line) gives the value of \( R_g \). The regression lines were fitted to the data within \( q R_g < 1.0 \). B, Kratky plot of the scattered intensity of N TAIL at 9 mg/ml. The arrow indicates the bump at \( q = 0.08 \) Å⁻¹.
changes take place in the isolated NTAIL domain compared with of N to P (18, 20, 26). To determine whether the isolated NTAIL domain was able to bind P, we tested the ability of bacterial lysate, thus confirming the specificity of the IP (see Fig. 5B, lane Nt + P). When both lysates were mixed, the IMAC resin selectively pulled down N TAIL expressed alone but not the untagged P expressed alone (Fig. 5B, lanes Nt and P). When both lysates were mixed, the IMAC resin selectively pulled down N TAIL, together with an additional protein that has the same electrophoretic mobility of P (Fig. 5B, lane Nt + P). The identity of this protein band was confirmed by Western blotting using an anti-P mAb (data not shown), thus proving that the N TAIL domain on its own is able to bind its physiological partner in bacterial lysates.

P, the Physiological Partner of NTAIL—As already mentioned, a large majority of natively unfolded proteins or protein domains characterized so far undergo disorder-to-order transitions in the presence of their physiological partner(s) (32, 33, 35, 37). The availability of the physiological partner of NTAIL enables us to check for the occurrence of such unstructured-to-structured transitions.

The P protein of MV consists of an N-terminal moiety (aa 1–230), PNT, responsible for binding to NCORE within the N°P complex (26), and of a C-terminal moiety (aa 231–507), PCT, responsible for binding to NTAIL in both N°P and N°N°C°P complexes (9, 69). We have recently shown that the PNT domain of MV P is intrinsically disordered when expressed alone (41) and probably also in the context of full-length P. Therefore, the interaction between NTAIL and P may result in a possible induced folding of PNT concomitantly to the possible induced folding of NTAIL. Since the purified PNT domain was already available (41), we purified the PCT moiety in order to monitor the possible structural transitions of NTAIL in the presence of either PNT or PCT.

Most bacterially expressed PCT is found in the soluble fraction of the bacterial lysate (data not shown). Purification consists of two steps, IMAC followed by gel filtration. The expected MM of the purified product (31 kDa) was confirmed by mass spectrometry. However, PCT migrates in SDS-PAGE as a 36-kDa protein (see Fig. 5C). This discrepancy may be due to the presence of a region (aa 231–307 of P) of peculiar composition (rich in Gly, Ser, Ala, and Glu) at the N terminus of PCT. Although PCT is eluted from the gel filtration column as a very broad peak (not shown), 1D and 2D gel studies have shown the presence of a major (97%) species with a MM of ~200 kDa. This MM may correspond either to a high degree of oligomerization or to a very elongated shape of the protein or to both, as in the case of Sendai virus (SeV) PCT (69).

The eluent from the gel filtration column (see Fig. 5C) contains several protein species. Western blot analysis with an anti-hexahistidine tag mAb (not shown) showed that the dif-

As discussed above, MV NTAIL, is required for stable binding of N to P (18, 20, 26). To determine whether the isolated NTAIL domain was able to bind P, we tested the ability of bacterial lysates expressing hexahistidine tagged NTAIL to pull down untagged P onto an IMAC resin. The IMAC resin selectively pulled down NTAIL expressed alone but not the untagged P expressed alone (Fig. 5B, lanes Nt and P). When both lysates were mixed, the IMAC resin selectively pulled down NTAIL.

axis, and their abundance suggests an intrinsic fragility in nucleocapsids formed by the bacterially expressed full-length nucleoprotein. Similar rings have also been described in Ref. 3.

Electron microscopy of N CORE, however (Fig. 4A, right panel), reveals a predominance of rigid helices up to 400 nm long, as previously described (22, 23). This indicates that the removal of the exposed NTAIL region appears to stabilize the interaction between turns of the helical nucleocapsid, either rendering it more robust or allowing short helices and rings to polymerize giving rise to longer helices. Radial density profiles, calculated from N and N CORE rings (Fig. 4B), combined with visual inspection of negative stain images, reveals no obvious extra density in N, in agreement with previous data (22, 23). This suggests that NTAIL is not visible as a distinct structural domain, because it is either poorly visualized due to its flexible nature or it is tightly associated with N CORE, as has been demonstrated for the C terminus of rabies virus N (67).

The intrinsic flexibility of NTAIL may sterically interfere with the formation of a stable interaction between successive turns of the helical nucleocapsid. The association of these rings gives rise to more rigid helical structures when imaged by negative stain EM. Therefore, the intrinsic disorder of NTAIL provides a structural explanation for the observed gain of rigidity of NTAIL-free nucleocapsids.

Epitope Exposure of NTAIL and Binding to P—In order to check whether the isolated NTAIL domain conserves a conformation similar to that adopted within the full-length N, we have performed IP studies using the anti-N Cl 25 and Cl 105 mAbs. These mAbs recognize the linear epitopes 457–476 and 515–525 of MV N, respectively. These mAbs are able to precipitate both N and NTAIL from bacterial lysates (see Fig. 5A, lanes N and Nt), whereas no such bands are observed with a control bacterial lysate, thus confirming the specificity of the IP (see Fig. 5A, lanes C). Thus, these two epitopes are equally accessible within the full-length protein and within the isolated NTAIL domain. This suggests that no dramatic conformational changes take place in the isolated NTAIL domain compared with the full-length N.

As discussed above, MV NTAIL is required for stable binding of N to P (18, 20, 26). To determine whether the isolated NTAIL domain was able to bind P, we tested the ability of bacterial lysates expressing hexahistidine tagged NTAIL to pull down untagged P onto an IMAC resin. The IMAC resin selectively pulled down NTAIL expressed alone but not the untagged P expressed alone (Fig. 5B, lanes Nt and P). When both lysates were mixed, the IMAC resin selectively pulled down NTAIL.

The identity of this protein band was confirmed by Western blotting using an anti-P mAb (data not shown), thus proving that the NTAIL domain on its own is able to bind its physiological partner in bacterial lysates.

P, the Physiological Partner of NTAIL—As already mentioned, a large majority of natively unfolded proteins or protein domains characterized so far undergo disorder-to-order transitions in the presence of their physiological partner(s) (32, 33, 35, 37). The availability of the physiological partner of NTAIL enables us to check for the occurrence of such unstructured-to-structured transitions.

The P protein of MV consists of an N-terminal moiety (aa 1–230), PNT, responsible for binding to N CORE within the N°P complex (26), and of a C-terminal moiety (aa 231–507), PCT, responsible for binding to NTAIL in both N°P and N°N°C°P complexes (9, 69). We have recently shown that the PNT domain of MV P is intrinsically disordered when expressed alone (41) and probably also in the context of full-length P. Therefore, the interaction between NTAIL and P may result in a possible induced folding of PNT concomitantly to the possible induced folding of NTAIL. Since the purified PNT domain was already available (41), we purified the PCT moiety in order to monitor the possible structural transitions of NTAIL in the presence of either PNT or PCT.

Most bacterially expressed PCT is found in the soluble fraction of the bacterial lysate (data not shown). Purification consists of two steps, IMAC followed by gel filtration. The expected MM of the purified product (31 kDa) was confirmed by mass spectrometry. However, PCT migrates in SDS-PAGE as a 36-kDa protein (see Fig. 5C). This discrepancy may be due to the presence of a region (aa 231–307 of P) of peculiar composition (rich in Gly, Ser, Ala, and Glu) at the N terminus of PCT. Although PCT is eluted from the gel filtration column as a very broad peak (not shown), 1D and 2D gel studies have shown the presence of a major (97%) species with a MM of ~200 kDa. This MM may correspond either to a high degree of oligomerization or to a very elongated shape of the protein or to both, as in the case of Sendai virus (SeV) PCT (69).

The eluent from the gel filtration column (see Fig. 5C) contains several protein species. Western blot analysis with an anti-hexahistidine tag mAb (not shown) showed that the dif-

3 S. Longhi, K. Johansson, D. Karlin, and B. Canard, unpublished data.

![Image](https://example.com/image.png)
different protein bands are degradation products. Among the different products, two major bands are present (see the arrows in Fig. 5C), the lower of which corresponds to a truncated form devoid of an 8-kDa C-terminal fragment (see arrow 2 in Fig. 5C). Mass spectrometry analysis of PCT allowed the precise identification of the cleavage site at position 436 of P, thus indicating that this region is flexible and exposed to the solvent.

**Does NTAIL Undergo an Induced Folding in the Presence of PCT?**—To monitor possible structural transitions of NTAIL in the presence of PCT, we have used far-UV CD spectroscopy.

The CD spectrum of purified PCT (see Fig. 6A, gray line) is typical of a protein with a predominant $\alpha$-helical content, as seen by the positive ellipticity between 185 and 200 nm and by two minima at 208 and 222 nm. The $\alpha$-helical content (estimated as 37% based on the ellipticity at 220 nm) is in agreement with structural data on SeV PCT domains (70, 71). After mixing NTAIL with different molar excesses of PCT, the observed CD spectra of the mixtures differ from the corresponding theoretical average curves calculated from the individual spectra. Since the theoretical average curves correspond to the spectra that would be expected if no structural variations occur, deviations from these curves point out structural transitions. Most pronounced structural transitions are observed with an NTAIL/PCT molar ratio of 1:1.5 (Fig. 6A). When equimolar amounts are used, only slight deviations from the theoretical average curve are observed; the two spectra are almost perfectly superimposable except for the 185–195-nm region, in which a 30% increase of ellipticity is observed (data not shown). In the presence of molar excesses of PCT, a random coil to $\alpha$-helix transition can be observed, as indicated by the much more pronounced minima at 208 and 222 nm and by the higher ellipticity at 190 nm of the experimentally observed spectra compared with the corresponding theoretical average curves (Fig. 6A). Moreover, the $\alpha$-helical contents of the mixtures (ranging from 38 to 48% at different molar ratios, with the highest value being obtained with a 1.5-fold excess of PCT) are not only higher than those of the corresponding theoretical average curves, but are also higher than the $\alpha$-helical content of PCT alone. Gradually increasing the concentrations of PCT to molar ratios as high as 5 does not result in more drastic structural variations than those observed at a 1.5-fold molar excess of PCT (data not shown).

As a control, we recorded CD spectra of lysozyme either alone or mixed with NTAIL at different molar ratios. As shown in Fig. 6B (gray line), the lysozyme spectrum is typical of an $\alpha/\beta$ protein, in agreement with structural data (72). Mixing equimolar amounts of the two proteins does not result in any significant structural variations. The experimentally measured CD spectrum of the mixture is in fact superimposable on the theoretical average curve (see Fig. 6B), as expected if no interaction occurs between lysozyme and NTAIL. The same results were obtained at a 5-fold excess of lysozyme (data not shown).

As shown in Fig. 6C, no significant structural variations are observed with equimolar amounts of NTAIL and PNT, with all spectra being typical of unfolded proteins. The same results were obtained by gradually increasing the amounts of PNT to molar excesses as high as 5 (data not shown). These results can be accounted for by assuming either that the two proteins do not interact, in agreement with the results reported in the literature (9, 68), or that the interaction does not imply any concomitant unstructured-to-structured transition.

The quite good superimposition between experimental and theoretical curves observed when mixing NTAIL with either PNT or lysozyme (both at different molar excesses) confirms the significance of the deviation observed in the case of the NTAIL-PCT mixtures (see Fig. 6A). Therefore, these results indicate that NTAIL undergoes an induced folding upon binding to PCT.

A pioneering study has recently demonstrated that an intrinsically disordered protein region is in fact structured in living cells (73). FlgM, a 97-residue protein from *Salmonella typhimurium*, which regulates flagellar synthesis by binding the transcription factor $\alpha^{28}$, is fully unstructured in vitro (74). However, NMR performed on living *E. coli* overexpressing FlgM showed that the C terminus of FlgM is in fact structured in vivo, in the same manner as it becomes structured upon binding to $\alpha^{28}$ in vitro, whereas its N terminus remains unstructured. Furthermore, the same gain of structure is observed in vitro in the presence of >400 g/liter glucose (73).

In order to check whether such solute-induced structural transitions could be observed also in the case of NTAIL, we have recorded a CD spectrum of NTAIL in the presence of 440 g/liter glucose (Fig. 6D). As shown in Fig. 6D, the spectrum obtained is still characteristic of an unfolded protein, and is similar to the spectrum obtained in the absence of glucose. According to Dedmon et al. (73), these results suggest that NTAIL would belong to the class of intrinsically disordered proteins that “do not become structured at physiologically relevant solute concentrations” and that require a physiological partner to fold.

**DISCUSSION**

The high proteolytic sensitivity of NTAIL, together with its peculiar sequence properties, strongly suggests that it is mostly unstructured.

The hydrodynamic properties of NTAIL inferred from both gel filtration and DLS experiments are consistent with NTAIL being either a stable globular dimer or a monomeric, extended
protein. Using NMR and CD, we show that NTAIL is monomeric and mostly unstructured. However, its far-UV spectroscopic parameters indicate that NTAIL is not fully unfolded, but rather it conserves a residual secondary structure content typical of natively unfolded proteins with a premolten globule conformation (31).

The scattering patterns obtained by SAXS show that NTAIL is a nonglobular protein. However, it retains a certain degree of compactness, as seen by the $R_g$ and $D_{\text{max}}$ values. The structural properties of the polypeptide chain can also be determined by the $R_g/R$ ratio. This ratio should be $\left(\frac{3}{5}\right)^{1/2}$ for a globular protein, around 0.9 for a premolten globule, and -1.5 for a random coil (75). In the case of NTAIL, the $R_g$ and $R$ values obtained by DLS and SAXS, respectively, lead to a ratio of 1.1, which is consistent with a premolten globule. Moreover, the mean hydrodynamic volume inferred from gel filtration and DLS studies (~97,000 Å³) is in agreement with the value expected for a native premolten globule (about 94,000 Å³) (31).

In addition, NTAIL lies in a region of the hydrophobicity versus net charge plot, which is typical of native premolten globules (31). In particular, intrinsic coils are more distant from the border between rigid and intrinsically unstructured proteins than intrinsic premolten globules. In the case of NTAIL, the distance value from the border ($H_{\text{boundary}} - H$) is 0.05 (see under “Experimental Procedures”), which is consistent with the value expected for a native premolten globule (0.037 ± 0.033) (31).

Thus, NTAIL can be described as a nonglobular polypeptide chain, more compact than a random coil, containing a significant amount of residual structure. This residual structure may prevent the polypeptide chain from adopting multiple conformations. In agreement, the distribution of the conformations of NTAIL is narrow, as seen by the relative sharpness of the elution peak observed in gel filtration, compared with what would be expected for a fully denatured protein in the presence of chaotropic agents (urea, guanidinium chloride). The residual ordered secondary structure of NTAIL may arise from dynamic breaking and reforming of interactions. We can reasonably speculate that these residual interactions could enable a more efficient start of the folding process induced by a binding partner.

The intrinsic disorder of NTAIL provides a possible structural explanation for the observed increase in rigidity of NTAIL-free nucleocapsids. The removal of the NTAIL region by limited proteolysis leads to a more rigid nucleocapsid, as indicated by the fact that EM micrographs of NCORE mostly contain herringbone structures, as opposed to EM images of N, which mainly contain rings. The presence of flexible regions at the surface of the viral nucleocapsid would ideally enable the viral ribonucleoprotein complex to have plastic interactions with several partners, such as the polymerase complex (L-P) but also possibly components of the cell cytoskeleton (76, 77). Striking changes in nucleocapsid conformation in response to either
increasing salt concentrations or limited proteolysis have already been described (22, 23, 78) and point out the plasticity of its helical structure. This plasticity suggests that changes in the microenvironment of the nucleocapsid could affect its conformation during viral RNA transcription and replication, or during virus assembly, processes in which modifications in the coating of the nucleocapsid could be important.

Most natively unfolded proteins or protein domains characterized so far under some degree of folding in the presence of a physiological partner (see Refs. 33, 35, and 37). After showing that NTAIL alone still conserves the ability to bind P, we have studied by far-UV spectroscopy the possible structural transitions of NTAIL in the presence of PCT and PNT. CD is the method of choice to monitor such structural variations, since a residual structural content as low as 10–20% can be determined by this method. Indeed, this method turned out to be sensitive enough to detect unstructured-to-structured transitions upon binding of NTAIL to PCT, whereas no structural transitions were detected in the presence of either PNT or lysozyme. In particular, binding of NTAIL to PCT results in a strong unstructured-to-structured transition rather than in minor local structural changes, as shown by the considerable gain of α-helicity of the mixture (48%) compared with PCT (37%). The gain of α-helicity of NTAIL observed in the presence of different molar excesses of PCT is in agreement with the strong α-helical propensity of NTAIL pointed out in experiments using TFE. The highest α-helical content has been obtained with a 1.5-fold excess of PCT. A possible interpretation for this result may come from analysis of purified PCT. As shown in Fig. 5C, purified PCT consists of two major bands: a full-length form and a truncated form devoid of the last 70 C-terminal residues, the relative abundance of which can be roughly estimated to be 70 and 30%, respectively. Biochemical data on SeV have shown that the extreme C terminus (aa 479–568) of PCT contains the region responsible for binding to N (79, 80). Accordingly, it is possible that in the purified PCT sample only the upper band may participate to the interaction with NTAIL. Therefore, a 1.5-fold excess of PCT may indeed correspond to an equimolar ratio between NTAIL and a form of PCT able to make a productive interaction. On the other hand, in the presence of lower amounts of PCT, such as those used in the equimolar mixture, only very slight structural transitions are observed, even if 70% of NTAIL is expected to be complexed to PCT. Thus, the gain of structure of NTAIL does not seem to take place gradually but rather follows a cooperative transition, where the midpoint transition is just below 1.5-fold molar excess of PCT. Such a cooperative transition is typical of protein folding/unfolding processes. Increasing the molar excess of PCT beyond 1.5-fold does not result in any stronger unstructured-to-structured transitions. This may reflect the formation of a 1:1 stoichiometric complex. Therefore, at concentrations of PCT above those required for the formation of such a complex, no more NTAIL would be available for binding to PCT, and saturation of the system would be achieved.

These results indicate that NTAIL folds upon binding to its physiological partner, P. However, the possibility that the observed disorder-to-order transitions may also concern flexible regions within PCT is an open question. In this regard, indications are expected to arise from further structural studies, in particular crystallographic studies of NTAIL-PCT complexes. In addition, we point out that the studies presented in this paper do not allow us to answer the question as to whether the gain of structure of NTAIL takes place in the PNT-P or in the P-P complex, or in both. Further studies are required to discriminate among these possibilities.

The intrinsic disorder is abundant within the replicative complex of MV, as indicated by experimental evidence accumulated on PNT and NTAIL. Furthermore, in this paper, we report biochemical evidence suggesting that MV P possesses an additional flexible region, as indicated by the protease sensitivity observed within PCT. These results are in agreement with biochemical and structural data on SeV P, indicating the presence of a flexible linker, sensitive to proteolysis within the 447–517 region (69–71).

One advantage of unstructured regions of proteins might be their ability to bind to a wide variety of structurally distinct substrates with a weak affinity (32, 35). The pattern of interactions of NTAIL is consistent with this hypothesis; NTAIL takes part in numerous interactions with different protein partners, including P (both within N-P and NNUC-P), the polymerase complex P-L, M (27), the interferon regulatory factor 3 (81), and possibly components of the cell cytoskeleton (76, 77). Moreover, NTAIL within viral nucleocapsids released from infected cells also binds to the human immunoglobulin G receptor FcγRII, provoking partial immunosuppression (82, 83). Finally, a conserved, hydrophobic patch at the extreme C terminus of Morbillivirus NTAIL has been described to bind to the heat-shock protein Hsp72, which modulates the level of viral RNA synthesis (25). This is coherent with the proposal that in addition to their function of promoting protein folding, some chaperones bind hydrophobic patches in intrinsically disordered proteins, modulating their accessibility and/or their sensitivity to proteolysis (32).

NTAIL plays very different roles during viral replication, being involved in RNA transcription, RNA replication, and viral assembly. One role of NTAIL in actively replicating nucleocapsids could be to put into contact several proteins within the replicative complex, such as the N-P and the P-L complexes. One can speculate that the gain of structure of NTAIL upon binding to its physiological partner could result in stabilization of the N-P complex. The increased rigidity of NTAIL would lead to a decrease in the number of conformations adopted by NTAIL in solution and thereby to an increased affinity for its target. At the same time, folding of NTAIL would result in a modification in the pattern of solvent-accessible regions, resulting in the shielding of specific regions of interaction. As a result, NTAIL would no longer be available for binding to its other partners.

This paper provides new perspectives in the study of disordered regions within the MV replicative complex; in particular, the gain of structure arising from the interaction with the physiological partner should lead to a protein conformation rigid enough to allow crystallization. The crystal structure determination of such complexes would represent an important step toward the understanding of the molecular mechanism of N and P.

Acknowledgments—We thank M. C. Trescol-Biémont and C. Rabourdin-Combe for kindly providing the plasmid encoding NTAIL. A. Mosbah for help with the NMR study, M. Desmadril for critical advice and D. Gerlier for the anti-N and anti-P monoclonal antibodies. We thank D. Gerlier and J. Curran for critical reading of the manuscript. We thank K. Dunker for useful discussions.

REFERENCES
1. Curran, J., and Kolakofsky, D. (1999) Adv. Virus Res. 54, 403–422
2. Lamb, R. A., and Kolakofsky, D. (2001) in Fields Virology (Fields, B. N., Knipe, D. M., and Howley, P. M., eds) 4th Ed., pp. 1305–1340, Lippincott-Raven, Philadelphia, PA
3. Warnes, A., Focks, A. R., Dowsett, A. B., Wilkinson, G. W., and Stephenson, J. R. (1995) Gene (Amst.) 160, 173–178
4. Karlin, D., Longhi, S., and Canard, B. (2002) Virology 302, 420–432
5. Focks, A. R., Stephenson, J. R., Warnes, A., Dowsett, A. B., Rima, B. K., and Wilkinson, G. W. (1993) J. Gen. Virol. 74, 1439–1444
6. Bhella, D., Ralph, A., Murphy, L. B., and Yeo, R. P. (2002) J. Gen. Virol. 83, 1831–1839
7. Spehner, D., Kirn, A., and Drillien, R. (1991) J. Virol. 65, 6296–6300
8. Spehner, D., Drillien, R., and Howley, P. M. (1997) Virology 232, 260–268
9. Huber, M., Cattaneo, R., Speelhude, P., Orwell, C., Norbury, E., Messeri, M.,
