

Review

# Functional Implications of Dynamic Structures of Intrinsically Disordered Proteins Revealed by High-Speed AFM Imaging

Toshio Ando

Nano Life Science Institute, Kanazawa University, Kanazawa 920-1192, Japan; tando@staff.kanazawa-u.ac.jp

**Abstract:** The unique functions of intrinsically disordered proteins (IDPs) depend on their dynamic protean structure that often eludes analysis. High-speed atomic force microscopy (HS-AFM) can conduct this difficult analysis by directly visualizing individual IDP molecules in dynamic motion at sub-molecular resolution. After brief descriptions of the microscopy technique, this review first shows that the intermittent tip-sample contact does not alter the dynamic structure of IDPs and then describes how the number of amino acids contained in a fully disordered region can be estimated from its HS-AFM images. Next, the functional relevance of a dumbbell-like structure that has often been observed on IDPs is discussed. Finally, the dynamic structural information of two measles virus IDPs acquired from their HS-AFM and NMR analyses is described together with its functional implications.

**Keywords:** intrinsically disordered proteins; IDPs; intrinsically disordered regions; IDRs; high-speed AFM; dynamic structure; bioimaging



**Citation:** Ando, T. Functional Implications of Dynamic Structures of Intrinsically Disordered Proteins Revealed by High-Speed AFM Imaging. *Biomolecules* **2022**, *12*, 1876. <https://doi.org/10.3390/biom12121876>

Academic Editors: Elisar Barbar, Nathalie Sibille and Vladimir N. Uversky

Received: 28 November 2022

Accepted: 11 December 2022

Published: 14 December 2022

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



**Copyright:** © 2022 by the author. 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

In physiological conditions, intrinsically disordered proteins (IDPs) are devoid of a well-defined three-dimensional structure entirely or partly. And yet, they can carry out biological functions and often work in the regulation of transcription and translation and in cellular signal transduction [1,2]. Therefore, IDPs are more abundant in eukaryotic cells than in prokaryotic cells. Interestingly, IDPs are relatively abundant in viruses [3,4]; Sonia Longhi and colleagues first found a viral IDP in the measles virus [5]. This abundance relative to the total number of viral proteins is possibly because viruses have highjacked the cellular machinery of their hosts. The intrinsically disordered regions (IDRs) contained in IDPs are highly protean. The residues in an IDR responsible for weak intramolecular interactions are widely distributed in space and time. Therefore, IDRs can dynamically sample multiple conformations, including fully disordered tail-like structures, loosely folded conformations, and incompletely or fully folded secondary structures. The residues responsible for target binding are also distributed in space and time, enabling an IDP to bind multiple targets by structural adaptation, resulting in different folded structures by coupled folding, as exemplified in the binding of an IDP (p53) to different structured proteins [6]. Therefore, IDPs can be multifunctional. The coupled folding of an IDR induced by binding to structured proteins [7] may imply that strong binding requires folding of the IDR. However, dimerization or target binding of IDRs without folding has been observed [8,9]. Their dissociation constants ( $K_d$ ) are in the order of 10  $\mu$ M. A recent study has suggested that two highly disordered IDPs can bind each other with much higher affinity while retaining their structural disorder and highly dynamic character [10,11]. Moreover, the flexible and adaptable nature of IDRs bestows on them high sensitivity to environmental changes and post-translational modifications. That is, the structure of IDPs and their inter-molecular interactions can be altered by minute changes in the intracellular environment and post-translational modifications. The biological importance of such high sensitivity is typically exemplified by the formation of liquid droplet-like condensates (membrane-less organelles) by liquid-liquid phase separation (LLPS) [12–16]. Thus, the

flexible, dynamic structural nature of IDRs is the basis for IDPs' distinct functions and functional mechanisms that cannot be achieved by structured proteins.

As such, elucidating the conformational ensemble and dynamics of IDPs is essential for understanding their functional mechanisms. However, various techniques developed and used for studying the structure of structured proteins are often not instrumental in the structural analysis of IDPs. The two major structural biological techniques, X-ray crystallography and cryogenic electron microscopy (cryo-EM), do not have practical applications for IDPs because IDRs hamper crystallization and are too thin to be visualized by cryo-EM. Two ensemble-averaging methods, small-angle X-ray scattering (SAXS) [17,18] and solution NMR spectroscopy [19], have mainly been used for structural analysis of IDPs (for their approximate performances, see Table 1). Various SAXS-derived data, such as radius of gyration ( $R_g$ ) and Kratky plot, provide information about the overall structural properties of IDPs. However, distinguishing individual structures of an IDP coexisting in solution and quantifying their populations require modeling of structural ensembles. For modeling, several methods have been developed, such as the ensemble optimization method (EOM) [20,21], ILBOMD [22], and MultiFoXS [23]. However, the resolution of SAXS is low (usually ~5 nm) [18,24], and the unique determination of multiple structures and their populations from SAXS data alone is difficult. Therefore, SAXS analysis is usually combined with molecular dynamics simulations or NMR analysis. In contrast to SAXS, NMR cannot determine the overall structure of IDPs. Instead, NMR determines high-resolution local structures of folded regions, transiently appearing structured regions, and residues interacting with targets using two-dimensional  $^1\text{H}$ - $^{15}\text{N}$ / $^1\text{H}$ - $^{13}\text{C}$  heteronuclear single quantum correlation spectra. This also means that NMR can specify disordered regions in an IDP that do not interact with other residues. Importantly, NMR allows model-based population analysis for multiple conformational states. A recent study proposed an improved NMR method for ensemble analysis based on solvent paramagnetic relaxation enhancement, enabling accurate, high-sensitive detection of low populations of residual structure [25]. However, NMR is still poor at distinguishing fully disordered and loosely folded structures. Loosely folded structures are often identified wrongly as fully disordered. Besides, there is an upper limit of analyzable molecular weight (~40 kDa). For larger IDPs, their deletion constructs with smaller molecular weights have to be subjected to NMR analysis. However, the analysis in this way does not necessarily guarantee that the revealed structures of segments are identical to those in the full-length protein. This is because intramolecular interactions within an IDP facilitated by a flexible IDR can alter the IDP structure, as has been demonstrated for Atg13 [26,27]. Moreover, it is hard to determine the rates of transitions between different conformational states that occur within >~100 ms, although the fast motion of individual residues can be detected by NMR relaxation analysis without information on structural coordinates [28].

**Table 1.** Comparison of four techniques used in structural analysis of proteins.

	Data Collection Time	Sample Amount	Spatial Resolution	Dynamics	Molecular Weight limitation
NMR <sup>(a)</sup>	days–month	2–100 mg	~0.2 nm	ps–ms <sup>(b)</sup>	<40 kDa
SAXS	s–h <sup>(c)</sup>	0.1–10 mg	1–10 nm <sup>(d)</sup>	s–min	none
HS-AFM	~10 s	0.1 µg	2–3 nm (XY) 0.1 nm (Z)	20–100 ms <sup>(e)</sup>	none
AFM	min	0.1 µg	2–3 nm (XY) 0.1 nm (Z)	>30 s	none

<sup>(a)</sup> Solution NMR spectroscopy; <sup>(b)</sup> Obtainable dynamics do not necessarily include structural coordinates; <sup>(c)</sup> Largely depends on the power of X-ray source (long exposure to X-ray damages proteins); <sup>(d)</sup> Usually ~5 nm. Revealed molecular shapes are model dependent; <sup>(e)</sup> 20–100 ms for IDPs and ~100 ms for structured proteins.

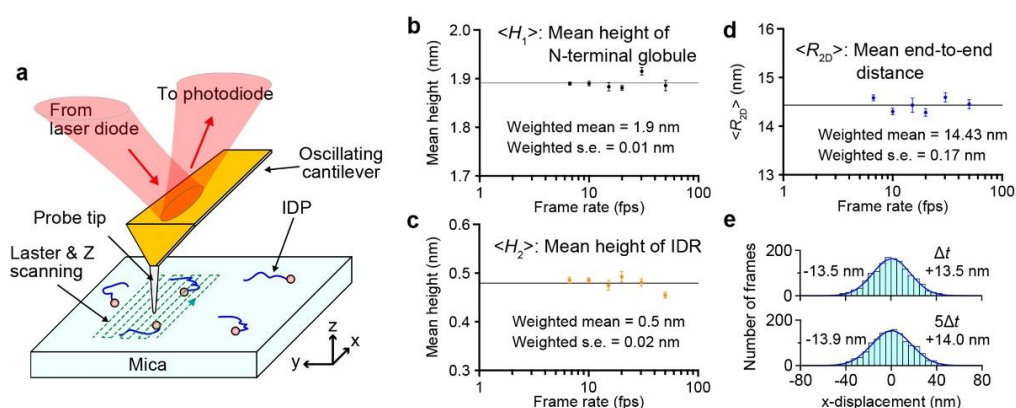
As a complement to the limited capability of these ensemble-averaging methods, we need a single-molecule method that can observe the dynamic protean structure of

individual IDP molecules, even at a spatial resolution lower than NMR. Conventional atomic force microscopy (AFM) was previously used to characterize IDPs. However, its target was aggregated assemblies of an IDP or constructs of structured domains within IDPs, such as the nonstructural protein 3 (a protease with folded structure, unlike the name of this protein) from the hepatitis C virus [29] and the DNA binding domain of p53 [30]. Crystal structures of these domains have been obtained. Single molecules of IDPs with disordered regions of >20 residues have never been successfully imaged with conventional AFM. Therefore, we have explored the potential of high-speed AFM (HS-AFM) in the dynamic structural analysis of IDPs. This relatively young microscopy has been successfully used to visualize the dynamic behavior of structured proteins during their functional activity at 2–3 nm lateral and ~0.1 nm vertical spatial resolution and at ~100 ms temporal resolution under near-physiological conditions, without any chemical labeling (Table 1), as demonstrated for a variety of structured proteins [31–35]. IDPs have also been successfully visualized by HS-AFM: e.g., facilitates chromatin transcription (FACT) protein [36,37], the archaeal Hef protein [38], and the bacterial flagellar hook-length control protein FliK [39]. These studies have demonstrated the capability of HS-AFM to visualize the thin structure of IDRs in dynamic motion. Importantly, as demonstrated in a recent study [40], HS-AFM can determine approximately the number of amino acids contained in a fully disordered IDR. Thus, HS-AFM can delineate and quantify the dynamic structure of IDPs, as demonstrated for several IDPs, including viral proteins [40]. Moreover, function-related phenomena of IDPs have also been visualized in recent HS-AFM studies: e.g., ion-dependent structural transitions of a stator ring of the bacterial flagellar motor between unfolded and folded states [41], amyloid fibril formation from  $\alpha$ -synuclein [42] and Sup35 [43], the assembly process of translation factors around the ribosomal stalk complex [44] and the formation of the droplet-like pre-autophagosomal structure (PAS) via LLPS [14]. In this review, some recent HS-AFM studies on IDPs, including two measles virus IDPs and those with a dumbbell-like shape, are described after a brief introduction of HS-AFM and verification of the absence of AFM-tip disturbance to IDP structure.

## 2. HS-AFM Imaging in Amplitude Modulation Mode

In the amplitude modulation mode, the AFM cantilever oscillates in the Z-direction so that the sharp tip attached to the free end of the cantilever contacts ~~with~~ the sample surface intermittently. This intermittent contact can avoid sample dragging during lateral scanning of the sample stage relative to the tip. Upon tip-sample contact, the cantilever oscillation amplitude is reduced. To maintain the contact (tapping) force constant, feedback control moves the sample stage in the Z-direction so that the oscillation amplitude returns to its set point value. As a result, the motion of the sample stage traces the sample surface (the Z-motion direction is opposite to the sample height). Therefore, we can construct the sample surface topography using a computer from the signal that drives the sample stage in the Z-direction. In our HS-AFM system, all instrumental components, including the cantilever, the scanner, and all electronics, are optimized for highly frequent tip-sample contact, fast conversion of the cantilever deflection signal to the amplitude signal, fast data sampling, fast scanning of the sample stage, or fast feedback control. To minimize the tip-sample contact strength, the free oscillation amplitude  $A_0$  of a short cantilever (7–10  $\mu\text{m}$ ) with a spring constant of 0.1–0.2 N/m is set at 1–2 nm, and its set point is set at  $0.85\text{--}0.9 \times A_0$ , resulting in the cantilever's oscillation energy loss per tap is only a few  $k_B T$  on average ( $k_B$ , Boltzmann constant;  $T$ , the temperature in kelvin). The feedback control speed depends on the ratio of  $A_0$  to the sample height  $h_s$ ; i.e., the control speed increases with increasing  $A_0/h_s$ . Since the height of IDPs is generally small (the mean height of a fully disordered IDR is about 0.5 nm), the imaging speed can be increased up to 50 fps for many IDPs. Very recent improvements to our HS-AFM system have enabled ~70 fps even for structured proteins with  $h_s < 25$  nm, although this improved system has not yet been used for IDPs. For HS-AFM imaging of biological molecules, they have to be placed on a substrate surface (Figure 1a). HS-AFM imaging of IDPs requires the surface to be

hydrophilic and flat at the atomic level. Practically, bare mica that meets these requirements is probably the only choice. In general, IDRs only weakly attach to mica via electrostatic and van der Waals interactions and, therefore, exhibit fast Brownian motion on mica. When an IDR is disordered over the entire length, it diffuses too fast to be imaged even at 50 fps. In such a case, we need to fuse a folded protein (e.g., GFP) to the N- or C-terminus of the IDR as an anchor onto mica. More details of practical issues in HS-AFM imaging of IDPs are described in [45]. For more comprehensive descriptions of the fundamentals, techniques, and biological applications of HS-AFM, a book is available [46]. HS-AFM systems have already been commercialized by three manufacturers (RIBM, Bruker-JPK, and Oxford Instruments-Asylum). Short cantilevers with a spring constant of 0.1–0.2 N/m and a resonant frequency higher than 400 kHz in water are also commercially available from NanoWorld and RIBM.



**Figure 1.** Absence of tip disturbance to the structure of IDPs. **(a)** Schematic showing HS-AFM imaging of IDP molecules placed on mica surface. **(b–d)** Molecular features of PQBP-1 (1–214) measured by HS-AFM at various imaging rates (6.7–50 fps): **(b)** Mean height of N-terminal globule, **(c)** mean height of IDR, and **(d)** mean 2D end-to-end distance of IDR. **(e)** Distributions of C-terminal end displacements in the X-direction were observed for WT PQBP-1 during time periods of  $\Delta t$  and  $5\Delta t$ .  $\Delta t$  is one frame acquisition time (42 ms). The numbers shown on each panel are the mean displacements in the +X and –X directions. Note that the IDR within PQBP-1 is constantly disordered.

### 3. No Effect of Tip-Sample Contact on the Structure of IDPs

The tip-sample interaction is much weaker in liquids than in air because of the lack of capillarity forces in liquids [47]. Nevertheless, there still may be a concern that the physical contact between a protein molecule and an oscillating tip in an aqueous solution would have a strong impact on the structure and function of the molecule. The absence of such an impact has been verified for various structured proteins. For example, the HS-AFM images that captured the processive movement of myosin V on actin filaments at 7 fps showed that its movement velocity is identical to that measured by fluorescence microscopy in the same buffer solution condition [31]. The conformational changes of F<sub>1</sub>-ATPase filmed with HS-AFM at 12.5 fps for over 40 s were observed to continue without the changes being debilitated [32]. During this imaging, the molecules were tapped with the AFM tip tens of millions of times. However, the absence of a tip disturbance to the dynamic structure of IDPs cannot be verified straightforwardly because the dynamic conformational changes of IDPs occur thermally and stochastically.

Although the feedback control error is generally small for low height samples with  $h_s \leq A_0$ , the error becomes larger (and hence the tip-sample interaction force becomes stronger in both X- and Z-directions) with increasing X-scanning velocity. Therefore, the best way to examine the possible tip disturbance to IDPs is to compare the characteristic structural parameters of an IDR quantified from its images filmed at various scanning velocities. Polyglutamine tract-binding protein 1 (PQBP-1) and its deletion construct PQBP-1 (1–214) were subjected to this examination [40]. PQBP-1 has an N-terminal globular

domain (1–81) containing a WW domain (48–81) followed by a constantly disordered region (82–265). The three structural parameters of PQBP-1 (1–214) were measured from many images captured at six different imaging rates (6.7–50 fps): the height of the globule ( $H_1$ ), the height of the disordered region ( $H_2$ ), and the two-dimensional (2D) end-to-end distance between the N-terminal globule and the C-terminal end ( $R_{2D}$ ). The  $R_{2D}$  was calculated as  $R_{2D} = D - H_1/2 - H_2/2$ , where  $D$  is the direct distance between the N-terminal globule and the C-terminal end. Note that  $R_{2D}$  can be measured accurately despite the fact that all pixel data contained in an image are not acquired simultaneously but at different times due to raster scanning of the sample stage during one-frame acquisition. This is because the N-terminal globule is nearly stationary, although the C-terminal IDR moves fast on mica; the positions of the N-terminal globule and the C-terminal end can be considered as detected at the same time. As shown in Figure 1b–d, the mean values of these parameters, ( $H_1$ ), ( $H_2$ ), and ( $R_{2D}$ ), were independent of the imaging rate. Moreover, the displacement of the C-terminal end of WT PQBP-1 in the fast scan X-direction during a given time interval was measured relative to the N-terminal globule. The displacement was distributed symmetrically around the zero-mean displacement, irrespective of the time interval (Figure 1e). Note that in raster scanning, the sample surface is traced (from  $-X$  to  $+X$ ) and retraced (from  $+X$  to  $-X$ ) in the fast scan direction (Figure 1a), but only images captured in trace imaging are used. Therefore, the symmetric distribution of the relative displacement indicates no effect of the tip on the fast motion of the flexible IDR on mica. The time interval independency of the displacement distribution suggests that the shape of the IDR changes much faster than the image acquisition rate. The autocorrelation function of  $R_{2D}(t)$  was zero except for  $\Delta t = 0$ , even when the HS-AFM images were captured at 50 fps [40]. While the tip intermittently contacts the sample, the mica surface is always in contact with the sample. Therefore, the mica surface may have a non-negligible effect, especially on the highly flexible structure of IDRs. An IDR is confined to a 2D solution space on the mica surface, entailing ( $R_{2D}$ ) longer than the mean three-dimensional (3D) end-to-end distance in solution, ( $R_{3D}$ ). However, similar confinement can also occur in the crowded cytoplasm, nucleus, and viral capsid.

#### 4. Constancy of Flexibility/Rigidity of Fully Disordered Regions

Except for rare cases, chemically denatured polypeptide chains are known to have unique flexibility or rigidity, irrespective of the amino acid sequence and composition [48]. This constancy can be expressed by a power law for the relationship between the number of amino acids  $N_{aa}$  and the mean radius of gyration,  $\langle R_g \rangle$ : i.e.,  $\langle R_g \rangle = 0.1927 \text{ nm} \times N_{aa}^{0.598}$  (Figure 2a) [48]. In the 3D worm-like chain model for semi-flexible polymer chains, the mean square of  $R_g$ , i.e.,  $\langle R_g^2 \rangle$ , can be expressed as a function of the 3D persistence length  $P_{3D}$  as

$$\langle R_g^2 \rangle = \frac{P_{3D}L}{3} - P_{3D}^2 + \frac{2P_{3D}^3}{L} - \frac{2P_{3D}^4}{L^2} \times \left[ 1 - \text{Exp}\left(-\frac{L}{P_{3D}}\right) \right] \quad (1)$$

where  $L$  is the contour length of a polymer chain. In the case of polypeptide chains,  $L$  can be calculated approximately as  $L = 0.360 \text{ nm} \times (N_{aa} - 1)$ ; 0.360 nm is the average distance between two adjacent amino acids [49,50]. The  $\langle R_g \rangle$  data of chemically denatured polypeptide chains approximately followed Equation (1), suggesting that chemically denatured proteins behave like worm-like chains. From this fitting, the value of  $P_{3D}$  is determined to be 0.976 nm.

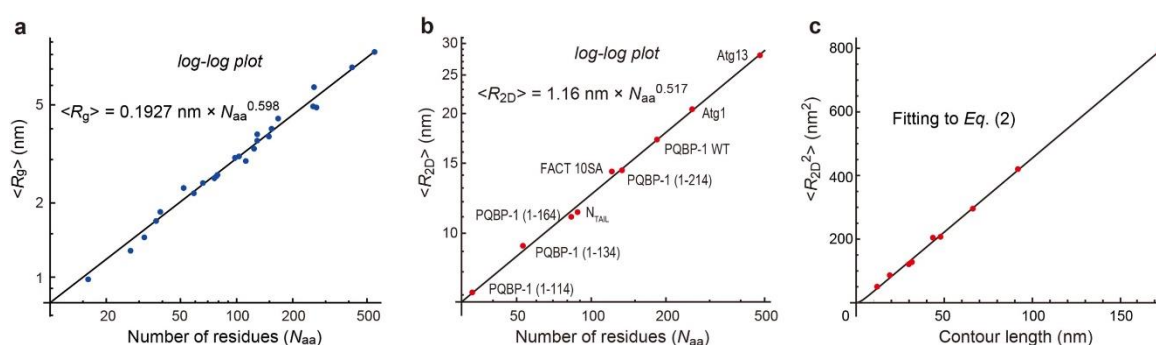
In contrast to chemically denatured polypeptides, the SAXS-measured  $\langle R_g \rangle$  values of 72 IDRs without permanent secondary structures in the native state have been reported not to follow a unique power law [51]. However, detailed structural analyses on some of these IDRs have identified regions that tend to form residual structures [52–55]. In contrast to SAXS, HS-AFM imaging can detect transiently appearing residual structures, and therefore we can select images of IDRs in the fully disordered state. All the log-log plots of mean 2D end-to-end distance  $\langle R_{2D} \rangle$  vs.  $N_{aa}$  (contained in a fully disordered IDR) measured by the HS-AFM imaging of nine different constructs of IDPs were on a single line representing a

power law of  $\langle R_{2D} \rangle = 1.16 \text{ nm} \times N_{aa}^{0.517}$  (Figure 2b) [40]. When worm-like chains freely equilibrate on the surface as in a 2D solution, the relationship between  $\langle R_{2D}^2 \rangle$  and the 2D persistence length ( $P_{2D}$ ) is expressed as [56,57]

$$\langle R_{2D}^2 \rangle = 4LP_{2D} \left[ 1 - \frac{2P_{2D}}{L} \left\{ 1 - \text{Exp} \left( -\frac{L}{2P_{2D}} \right) \right\} \right]. \quad (2)$$

Because the nine  $\langle R_{2D} \rangle$  data were well fitted to Equation (2), as shown in Figure 2c, these IDRs in the fully disordered state are uniquely characterized by the measured  $P_{2D}$  value of 1.16 nm. In addition, the fully disordered IDRs are also uniquely characterized by their mean diameter of approximately 0.5 nm. The unique power law enables the determination of  $N_{aa}$  contained in a fully disordered IDR from its HS-AFM images, as described later.

However, the substrate surface can have an additional effect arising from 2D frictional forces locally exerting from the surface against the fast 2D Brownian motion of a highly flexible IDR chain [58]. In the absence of this effect,  $\langle R_{2D} \rangle$  is identical to  $2\sqrt{3} \times \langle R_g \rangle$ , where  $\langle R_g \rangle$  is the mean radius of gyration in 3D solution (not on the substrate surface). The comparison of our  $\langle R_{2D} \rangle$  values for the nine different IDP constructs with SAXS-measured  $\langle R_g \rangle$  values of four PQBP-1 IDR constructs and 10 tau protein constructs without large extended domains [55] quantitatively revealed that IDRs are swelled by the frictional force from mica by a factor of 1.24, i.e.,  $\langle R_{2D} \rangle = 1.24 \times 2\sqrt{3} \times \langle R_g \rangle$ . Thus, we could derive a power law for  $\langle R_g \rangle$  as  $\langle R_g \rangle = 0.248 \text{ nm} \times N_{aa}^{0.532}$  and  $P_{3D} = 0.765 \text{ nm}$ . In literature, we can find many SAXS-measured  $\langle R_g \rangle$  data of IDPs that follow this power law [59–69], although the absence of residual structures in these IDPs is unknown. Note that the value of  $P_{3D}$  (0.765 nm) is smaller than that of chemically denatured polypeptides (0.976 nm), indicating that chemically denatured polypeptides are more extended than fully disordered IDRs. This distinct difference can be arguably explained simply by the fact that denaturants such as urea and guanidinium chloride are good solvents. Therefore, the polypeptide chain is extended because the interaction between the chain backbone and a good solvent is favorable. In contrast, chain-chain interactions are repulsive, although different interpretations may be possible [70].

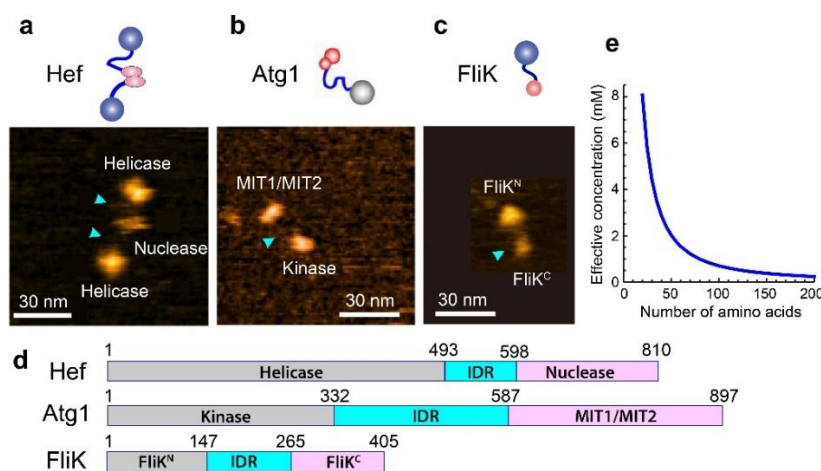


**Figure 2.** Relationship between molecular dimensions and the number of amino acid residues. (a,b) Power laws are found in chemically denatured proteins (a) and fully disordered regions within IDPs (b). The data in (a) are from the SAXS study by Kohn et al. [48]. The mean  $R_{2D}$  values,  $\langle R_{2D} \rangle$ , in (b) are those for fully disordered IDRs observed in the indicated protein constructs. (c) The mean square of  $R_{2D}$  values,  $\langle R_{2D}^2 \rangle$ , of the nine IDP constructs shown in (b) are fitted to Equation (2), from which the value of  $P_{2D}$  is estimated to be 1.16 nm. The number of residues in (b) is those contained in the fully disordered IDRs.

## 5. IDPs with Dumbbell Shape

IDPs randomly selected for HS-AFM imaging often showed a dumbbell-like shape (i.e., two structured domains are linked via an IDR): e.g., the archaeal Hef (helicase-associated endonuclease for fork-structured DNA) protein [38], bacterial FliK protein [39] and yeast

Atg1 [40] (Figure 3a–c). In addition, other methods found various dumbbell-shaped IDPs: e.g., the regulatory subunits of the cAMP-dependent protein kinase [71], cardiac troponin C [72], the DNA repair factor CtIP [73], CAHS (cytoplasmic abundant heat soluble) protein [74], and Ca<sup>2+</sup>-unbound calmodulin [75]. In these IDPs, the N-terminal domain (NTD) and the C-terminal domain (CTD) often possess distinct functions. For example, Hef has a helicase domain at the N-terminus and a nuclease domain at the C-terminus and dimerizes via the nuclease domain. The helicase activity is dramatically stimulated by fork-structured DNAs, while the nuclease specifically cleaves nicked, flapped, and fork-structured DNAs [76,77]. From these activities and phenotype analyses of mutant strains, Hef functions mainly in stalled replication fork repair [78]. In the case of Atg1, the sole kinase in all Atg proteins, the NTD is a kinase domain, while the CTD is the tandem MIT (microtubule interacting and transport) domain that binds the MIT-interacting motif (MIM) in Atg13. The NTD in FliK works as a length ruler of the bacterial flagella hook, while the CTD functions to switch secretion from the hook protein to the filament protein in the type-III flagellar export apparatus [79]. The domain diagrams of these dumbbell-shaped IDPs are shown in Figure 3d. A previous NMR analysis of the FliK<sup>N</sup> fragment (1–147) indicated a largely unfolded structure, whereas the FliK<sup>C</sup> fragment (204–370) is structured, except for its N-terminal region 204–252 [80]. The HS-AFM images of dumbbell-shaped FliK showed larger (mean height, 2.7 nm) and smaller (mean height, 2.0 nm) globules connected by a flexible linker (Figure 3c). From HS-AFM images of deletion constructs of FliK, the larger and smaller globules were identified as the NTD and CTD, respectively [39]. However, their size difference is inconsistent with the domain arrangement in FliK (Figure 3d), suggesting that the NTD is loosely folded. There may be two possibilities for the disagreement between HS-AFM and NMR results of the structure of FliK<sup>N</sup>; (i) the loosely folded structure of the FliK<sup>N</sup> fragment could not be detected but was detected by NMR as unstructured, or (ii) the FliK<sup>N</sup> fragment is unfolded.



**Figure 3.** Dumbbell-shaped IDPs observed by HS-AFM. (a–c) Clips from HS-AFM images and their schematized molecular features of the three IDPs. The light blue arrowheads point to IDRs. (d) Domain diagrams of the three IDPs. (e) Effective concentration of CTD (NTD) around NTD (CTD) as a function of the number of amino acids contained in the fully disordered IDR connecting NTD and CTD.

The common structural feature in dumbbell-shaped IDPs must have multiple functional benefits. In the repair of stalled replication fork by Hef, the fork structure has to be altered by the helicase, and DNA segments have to be removed by the nuclease. Although other proteins are also involved in the repair process, it is evident that the helicase and nuclease activities must proceed in harmony. The flexible IDR linker connecting the helicase and nuclease domains can localize these functional sites more efficiently and facilitates their harmonized actions compared to the case where separated helicase and nuclease proteins

work independently. In the case of Atg1, its C-terminal MIT1/MIT2 domain can bind the dephosphorylated MIM domain located in the C-terminal disordered region of Atg13. Therefore, starvation-induced dephosphorylation of Atg13 triggers its binding to Atg1 and then recruits Atg17-29-31 and Atg11 to form the Atg1 complex (Atg1-Atg13-Atg17-Atg29-Atg31) that works as a scaffold for PAS formation [14,26]. The Atg13-Atg17 interaction can occur when regions close to the MIM domain are dephosphorylated. Although the reaction timing after PAS or autophagosome formation is not well known, Atg1 kinase phosphorylates Atg13 and Atg11 [27,81], which is considered to dissolve a large complex formed by the PAS complex and other Atg proteins, and regulates selective autophagy [81,82]. When Atg1 phosphorylates the Atg13 IDR by the N-terminal kinase, Atg1 also grasps this IDR using the C-terminal MIT1/MIT2 domain. Thus, the N- and C-terminal domains of Atg1 need to get in close proximity to each other, which can be facilitated by the long flexible linker connecting the two interaction sites.

Another functional benefit of the dumbbell-like structure is that this structure allows and facilitates dynamic interactions between the two distantly separated functional domains. Even when the  $K_d$  of intramolecular interactions between distantly separated domains is large, the effective concentration ( $C_{\text{eff}}$ ) of one domain around the other can be significantly high, enabling their frequent interactions. The  $C_{\text{eff}}$  can be estimated as a function of  $\langle R_{3D} \rangle$ , as follows. A fully disordered IDR is considered to behave as a worm-like chain. In this case, the distribution function of  $R_{3D}$  is given by

$$P(R_{3D})dR_{3D} = 4\pi R_{3D}^2 \left[ \frac{3}{2\pi \langle R_{3D}^2 \rangle} \right]^{3/2} \exp \left[ -\frac{3}{2} \frac{R_{3D}^2}{\langle R_{3D}^2 \rangle} \right] dR_{3D}. \quad (3)$$

The two globular domains linked via such an IDR associate with each other when they approach within a certain range of distance. However, when we consider the actual dimensions of the two globules,  $R_{3D}$  becomes a complex function of the distance between the two globules. To avoid this complexity, we assume that NTD-CTD association occurs when  $R_{3D}$  becomes  $R_0$  or shorter (i.e.,  $R_{3D} \leq R_0$ ). In this approximation, the probability of finding the CTD in the volume ( $R_{3D} \leq R_0$ ) around the NTD is given by

$$p = \int_0^{R_0} P(R_{3D})dR_{3D} \quad (4)$$

Therefore, the  $C_{\text{eff}}$  of the CTD in the volume in close proximity to the NTD is given as

$$C_{\text{eff}} = \alpha \times p \times 10^{-3} / \left( \frac{4\pi}{3} R_0^3 \cdot N_A \right) \quad (5)$$

where  $N_A$  is the Avogadro constant and  $\alpha$  is a steric hindrance effect factor by which the NTD-CTD interaction frequency is reduced. Considering the NTD surface available for CTD binding, the steric hindrance factor is approximately 0.5. Supposing that the IDR contains 100 amino acids, its  $\langle R_{3D} \rangle$  value is 7.25 nm, under which Equation (4) provides  $p = 3.56 \times 10^{-3}$ . Therefore, the effective concentration of the CTD around the NTD is approximately estimated to be  $C_{\text{eff}} = 0.71$  mM (the  $N_{\text{aa}}$  dependence of  $C_{\text{eff}}$  is shown in Figure 3e). Thus, the two terminal domains in a dumbbell-shaped IDP can have a large chance to interact with each other even when their  $K_d$  value is in the order of mM.

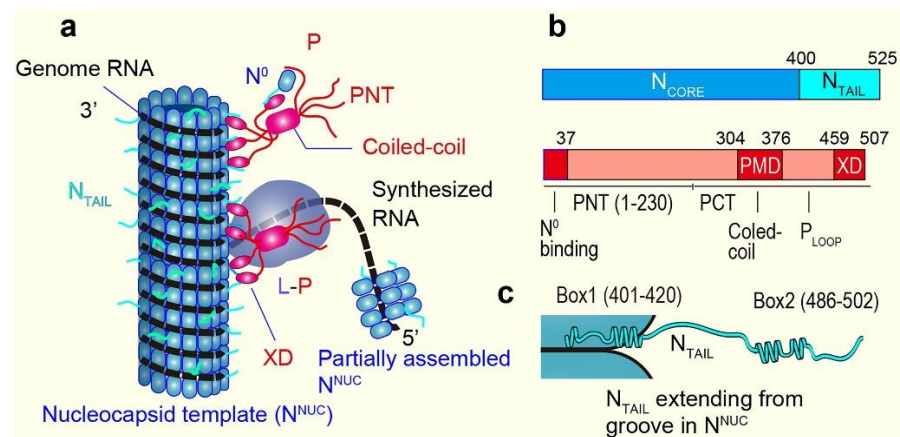
So far, only 2–3 IDPs with a property of “conformational malleability” have been found [26]. Atg13 is such a rarely found IDP. Although it is not dumbbell-shaped, its N-terminus folded HORMA domain (1–267) interacts with the C-terminus end region of the long IDR (268–740). The HORMA domain binds Atg9, a membrane protein existing on specific cytoplasmic vesicles (Atg9 vesicles) [83]. The C-terminal IDR contains the Atg1 binding site (460–521) and Atg17 binding sites (359–389 and 424–436) [27]. The folded HORMA domain becomes largely disordered when the C-terminus is fused to maltose-binding protein (MBP) [26,27]. The MBP-fusion impairs the PAS localization of



Atg9 in the cell. These observations indicate that transient NTD-CTD interactions are required for the formation of the stable and functional HORMA structure because the MBP moiety can interfere with NTD-CTD interactions. The stabilized structure of HORMA can be retained for a while even after the transient interaction is completely removed (i.e., malleability). Therefore, only occasional interactions are enough to keep the stable and functional HORMA structure. And yet, the functional role of this conformational malleability is still unknown. Although not yet clearly demonstrated, Flik possibly exhibits conformational malleability. That is, when the Flik<sup>N</sup>-Flik<sup>C</sup> interaction in *cis* is lost for a certain time duration, the Flik<sup>N</sup> domain (1–147) becomes fully disordered. When this domain is fully disordered, its contour length becomes ~53 nm, nearly identical to the maintained hook length of the bacterial flagellum (~55 nm), and thus, the Flik<sup>N</sup> domain can work as a ruler of hook length. Namely, the N-terminal end of FliK binds to the tip of the growing hook, while the disordered FliK<sup>N</sup> chain is moderately stretched in the narrow channel of the filament. When the hook grows up to ~55 nm, the Flik<sup>N</sup> chain is fully stretched and, thus, pulls the Flik<sup>C</sup> domain remaining in the cytoplasm. This pulling force changes the Flik<sup>C</sup>-secretion gate interaction to switch the secretion substrates from early proteins, including the hook protein, to late proteins, including the flagellin, as advocated previously [79,84].

## 6. Measles Virus IDPs

Before describing the HS-AFM study on two IDPs of measles virus, here I summarize their background. Comprehensive reviews on these IDPs and their functional roles are available elsewhere [85–88]. The measles virus is a nonsegmented, single-stranded, negative-sense RNA virus. Therefore, after virus entry into host cells, mRNAs coding for eight different virus proteins with six genes need to be synthesized from the genome RNA by using its own RNA-dependent RNA polymerase complex comprising the Large (L) protein and its cofactor, the phosphoprotein (P). This synthesis mode is later switched to the processive mode to produce cRNA for genome replication. The genome RNA is encapsidated by the nucleoprotein (N) to form a helical nucleocapsid with a diameter of ~18 nm and a helical pitch of ~6 nm, comprising 12.3 protomers per turn of the helix (Figure 4a), as suggested from the cryo-EM analysis of nucleocapsid-like particles [89]. This ribonucleoprotein complex is the substrate for both transcription and replication; the polymerase complex (L-P) cartwheels on the nucleocapsid template in these processes (Figure 4a). Each N binds six nucleotides, and therefore, the genome length (15,894 nucleotides) is exactly  $6 \times$  integer (2649). Monomeric N also exists in the cell as a soluble monomer which is referred to as N<sup>0</sup>, while the assembled form is referred to as N<sup>NUC</sup>. N<sup>0</sup> requires its chaperone, whose role is played by P and the V protein, which shares the first 231 amino acids with P [88]. The binding of P (and also V) blocks the polymerization of N. The tetrameric P tethers L onto the nucleocapsid template (Figure 4a). Both the N and P proteins contain large, disordered regions (Figure 4b). The N protein consists of the N-terminal structured domain (N<sub>core</sub>: 1–400) and the C-terminal disordered domain (N<sub>TAIL</sub>: 401–525). Note that the aforementioned nucleocapsid-like particles are formed without N<sub>TAIL</sub>; in the presence of N<sub>TAIL</sub>, the nucleocapsid is very flexible. The P protein consists of the N-terminal domain containing a large, disordered region (PNT: 1–230) and the C-terminal domain (PCT: 231–507) (Figure 4b). Transcription requires only PCT, whereas genome replication also requires PNT. The N-terminal PNT moiety (P<sub>N37</sub>: 1–37) binds N<sup>0</sup>, while the PCT region (304–376), referred to as PMD, is responsible for coiled-coil oligomerization of P and binding to L, and the C-terminal end region (459–507) referred to as the X domain (XD) is responsible for binding to N<sub>TAIL</sub> (the segment P<sub>Loop</sub> connecting PMD and XD is disordered). Thus, the formation of the N<sup>0</sup>-P complex is mediated by the dual PNT-N<sub>CORE</sub> and PCT-N<sub>TAIL</sub> interaction. In a different description of the domain organization of P, the region 1–304, larger than PNT, is referred to as P<sub>TAIL</sub>, which is mostly disordered except for P<sub>N37</sub> [88].



**Figure 4.** Measles virus IDPs, N and P. (a) Schematic showing complexes of N<sup>0</sup>-P, N<sup>NUC</sup>-P, and L-P RNA polymerase. The location of RNA is schematically represented for its visualization. Only a part of N<sub>TAIL</sub>s is shown. (b) Domain diagrams of N and P proteins. The light-color regions indicate IDRs. (c) Schematic showing N<sub>TAIL</sub> extending from the groove of nucleocapsid template.

There are fundamental questions to be solved about the measles viral transcription and replication processes. For example, (i) how does the polymerase switch between transcriptional and replicative modes? (ii) how are viral and cellular RNAs discriminated by the N protein? (iii) how can the L-P complex access the RNA which winds around the groove of the helical polymer of N protein? and (iv) why does the N<sup>0</sup>-P complex need to be formed before the N protein polymerizes? Various attempts to answer these questions and others have been carried out based on the structural features of N and P and their interactions. Some of the mechanistic proposals that answer these questions are summarized here. For more details, see reviews [85–88]. For instance, it has been proposed that the mode switch between transcription and replication is due to a conformational change of the nucleocapsid template caused by the folding of N<sub>TAIL</sub> induced by the interaction with XD [85]. However, a completely different mechanism has also been proposed for a negative-sense RNA virus (influenza virus). That is, there may be no mode switch, but nascent cRNA is degraded by host cell nucleases unless it is stabilized by newly synthesized N and P [90]. N<sub>TAIL</sub>-XD binding is also suggested to assist L in remaining in contact with the nucleocapsid template during successive synthesis of mRNAs [91]. The binding of P<sub>N37</sub> (and also V<sub>N37</sub>) to N<sup>0</sup> keeps N<sup>0</sup> in an open conformation ready to grasp viral RNA and is suggested to avoid non-specific binding of N to cellular RNA [92]. N<sub>TAIL</sub> is suggested to loosen the helically assembled ribonucleoprotein complex, facilitating the transcription and replication by L-P [93].

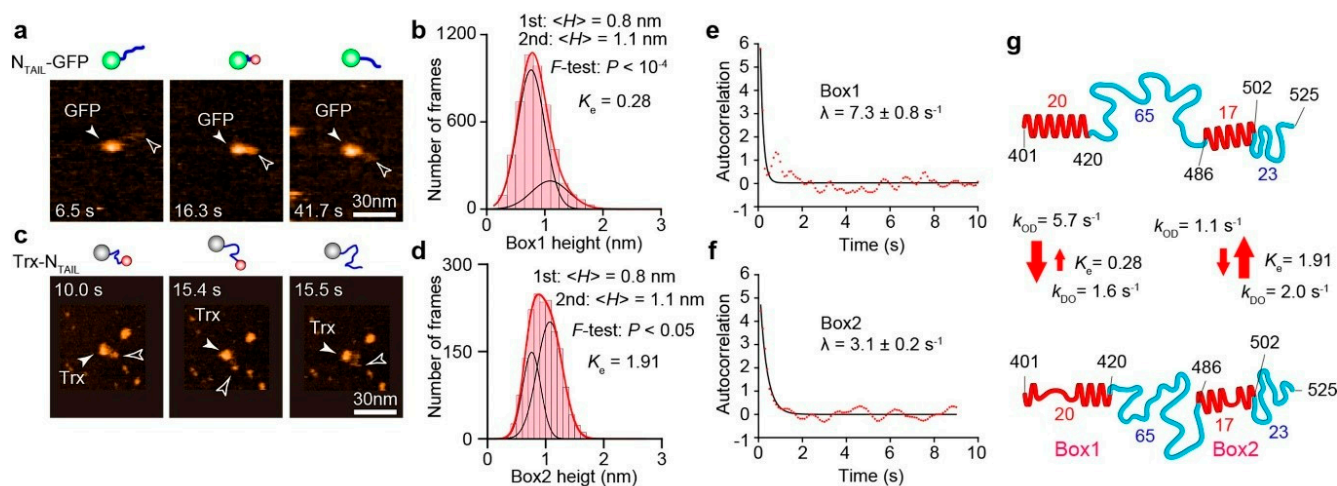
## 7. HS-AFM Study on N<sub>TAIL</sub>

Prior to the HS-AFM study on N<sub>TAIL</sub> (401–525) [40], its structural features have already been well characterized quantitatively by various methods. The two regions, called Box1 (401–420) and Box2 (486–502), were previously shown to contain a linear motif with  $\alpha$ -helical propensity ( $\alpha$ -MoRE) (Figure 4c) [93,94] undergoing transitions among completely unfolded and folded helical conformations [95–99]. The NMR-detected unfolded state of Box2 is short-lived ( $\mu$ s order) [86]. The binding of Box2 to XD in a triple  $\alpha$ -helical bundle is suggested to form a quadruple  $\alpha$ -helical bundle [100]. Box1 is located in the groove of N<sup>NUC</sup> (Figure 4c) [89,93]. In addition to Box1 and Box2, the region (517–525) is called Box3 because it is suggested to bind XD [101]. However, Box3 is not considered  $\alpha$ -MoRE because the secondary structure stabilizer, 2,2,2-trifluoroethanol (TFE), does not promote  $\alpha$ -helical folding within Box3 [101], unlike Box1 and Box2. Box3 possibly remains disordered even after binding to XD. Thus, the regions containing 88 amino acids in the two IDRs (421–485 and 503–525) are considered to be fully disordered. Since most of these structural features of

$N_{\text{TAIL}}$  are already established, the HS-AFM imaging of  $N_{\text{TAIL}}$  constructs aimed to examine whether HS-AFM could faithfully detect these known structural features of  $N_{\text{TAIL}}$  and to determine dynamic structural transition rates that could not have been analyzed with other methods. As already included in Figure 2b,  $\langle R_{2D} \rangle = 11.6 \pm 4.4$  nm measured with the HS-AFM imaging of  $N_{\text{TAIL}}$ -GFP fusion (Figure 5a) approximately followed the power law,  $\langle R_{2D} \rangle = 1.16 \text{ nm} \times N_{\text{aa}}^{0.517}$  (GFP was used to impede the otherwise very fast diffusion of  $N_{\text{TAIL}}$  on mica). In this fusion construct, Box1 was always visible, whereas Box2 only occasionally appeared because it is positioned close to the large GFP moiety. The height distribution of Box1 was well fitted to a double Gaussian function with peaks at 0.8 nm and 1.1 nm (Figure 5b). The former is larger than the height of fully disordered regions ( $\sim 0.5$  nm), while the latter is identical to the diameter of a single  $\alpha$ -helix. The area ratio of the two Gaussian components ( $K_e = \text{higher state/lower state}$ ) was 0.28, indicating a propensity of Box1 to take partially disordered conformations. To observe dynamic structural transitions of Box2,  $N_{\text{TAIL}}$  was fused to thioredoxin (Trx) at the N-terminus of  $N_{\text{TAIL}}$  (Figure 5c). The height distribution of Box2 was also well fitted to a double Gaussian function with peaks at 0.8 nm and 1.1 nm (Figure 5d). However, Box2 tends to form an  $\alpha$ -helix compared to Box1, judging from its larger  $K_e$  value, 1.91. From these observations, we conclude  $N_{\text{TAIL}}$  is fully disordered except for Box1 and Box2. Since the measured  $R_{2D}$  values are widely distributed due to the highly flexible nature of the disordered regions, changes of  $R_{2D}$  resulting from the structural transitions of small Box1 and Box2 are masked by this wide distribution of  $R_{2D}$  and their preferred tendencies to take partially disordered and fully folded states, respectively. Nevertheless, the time-varying  $R_{2D}$  values are well correlated with the time-varying height values of Box1 and Box2, suggesting that the IDR length partially changes due to the transitions in these globules between the partially and fully folded states. One of the advantages of HS-AFM imaging of IDPs over other methods is that we can directly observe structural transitions and, therefore, determine the transition rates unless they are faster than  $\sim 50 \text{ s}^{-1}$ . To determine the transition rates of Box1 and Box2, the autocorrelation functions  $G(\tau)$  of their time-series height data  $H(t)$  were first calculated (Figure 5e,f) as

$$G(\tau) \equiv \frac{\sum_{n=1}^{N-h} [H(n\Delta t) - \hat{\mu}][H((n+h)\Delta t) - \hat{\mu}]}{\sum_{n=1}^N [H(n\Delta t) - \hat{\mu}]^2} \quad (6)$$

where  $\tau \equiv h\Delta t$  ( $h = 0, 1, 2, 3, \dots, N-1$ ) and  $\hat{\mu}$  is the mean value of  $H(n\Delta t)$  ( $n = 1, 2, 3, \dots, N$ ). When the order-to-disorder and disorder-to-order transition rates are expressed by  $k_{\text{OD}}$  and  $k_{\text{DO}}$ , respectively, the  $G(\tau)$  relaxation rate denoted by  $\lambda$  is expressed as  $\lambda = k_{\text{OD}} + k_{\text{DO}}$ . From the relationship,  $K_e = k_{\text{DO}}/k_{\text{OD}}$ ,  $k_{\text{OD}}$  and  $k_{\text{DO}}$  can be expressed as functions of  $\lambda$  and  $K_e$  as  $k_{\text{OD}} = \lambda/(1 + K_e)$  and  $k_{\text{DO}} = \lambda K_e/(1 + K_e)$ . In this way, the rate constants could be determined as  $k_{\text{OD}} = 5.7 \text{ s}^{-1}$  and  $k_{\text{DO}} = 1.6 \text{ s}^{-1}$  for Box1 and  $k_{\text{OD}} = 1.1 \text{ s}^{-1}$  and  $k_{\text{DO}} = 2.0 \text{ s}^{-1}$  for Box2. Thus, the dynamic structural features of  $N_{\text{TAIL}}$  are quantitatively delineated, as shown in Figure 5g. In both Box1 and Box2, the structural transitions are relatively slow, comparable to the elongation rate of the L-P polymerase (three nucleotides/s) [102]. Although this rate comparability may be just a coincidence,  $N_{\text{TAIL}}$  has to change its structure in concert with the motion of the L-P polymerase cartwheeling on the nucleocapsid template. For example, Box2 has to change its affinity for P (at XD) during this L-P polymerase motion. In contrast, Box1 has to be either disordered or folded (but likely the former) to loosen the helical nucleocapsid structure by interfering with the stacking of turns (see Figure 4a,c) [93,103] so that the polymerase reaction is facilitated.



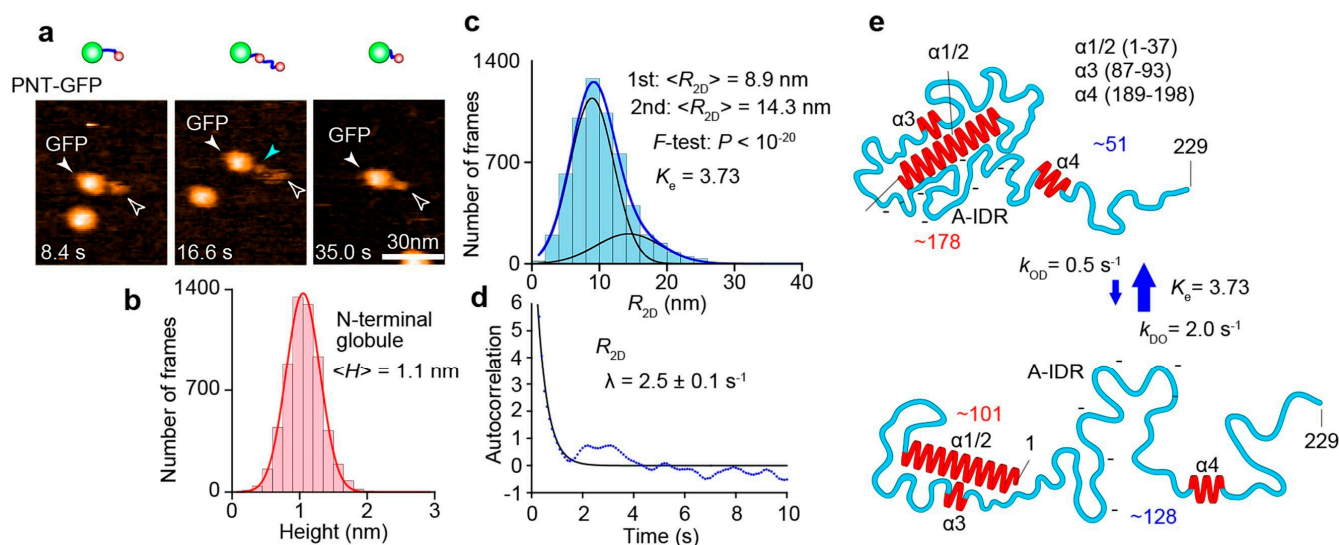
**Figure 5.** HS-AFM analysis of structure and dynamics of  $N_{TAIL}$ . (a) HS-AFM images and their schematics of  $N_{TAIL}$ -GFP. The open arrowheads point to the N-terminal end of  $N_{TAIL}$ . (b) Height distribution of the N-terminal globule (i.e., Box1). (c) HS-AFM images and their schematics of  $Trx-N_{TAIL}$ . The open arrowheads point to the C-terminal globule of  $N_{TAIL}$ . (d) Height distribution of the C-terminal globule (i.e., Box2). (e,f) Autocorrelation functions of time-series height data of Box1 (e) and Box2 (f). The values of  $\lambda$  represent the respective decay constants of the autocorrelation functions. (g) Schematics showing structural and dynamic features of  $N_{TAIL}$ . The top and bottom panels correspond to the more- and less-ordered states, respectively. The red arrows indicate the kinetic nature of changes in height of Box1 and Box2.

## 8. HS-AFM Study on PNT

When we started the HS-AFM imaging of PNT in 2012, nothing was known about its structure except for the resistance to proteolysis of the N-terminal 27–99 region of PNT in the presence of TFE [104]. In 2016, a crystal structure was revealed for a fused protein complex,  $P_{1-48}-N^0_{21-408}$ , that is devoid of the disordered regions of P and N [105]. This chimera protein is dimerized in the crystal, and  $P_{1-48}$  forms two  $\alpha$ -helices (longer  $\alpha 1$  and shorter  $\alpha 2$ ) connected with a short loop. The  $\alpha 1$  and  $\alpha 2$  bridge two adjacent  $N^0_{21-408}$  protomers, although this arrangement in the chimera proteins may not occur in the physiological case. It is also uncertain whether the N-terminal region, P (1–48), forms  $\alpha$ -helices without  $N^0$ . In 2018, the Blackledge group reported the NMR structures of  $P_{TAIL}$  (1–304) alone and the complex between the full length of N and  $P_{TAIL}$  [106]. In the NMR structure of  $P_{TAIL}$  alone, four regions were identified to have propensities to form  $\alpha$ -helices and thus named  $\alpha 1/2$  (1–37),  $\alpha 3$  (87–93), and  $\alpha 4$  (189–198). These three regions are linked by highly flexible IDRs. The IDR segment (125–170) that links  $\alpha 3$  and  $\alpha 4$  is rich in acidic residues and devoid of basic residues and, therefore, referred to as the acidic loop. In the  $N^0$ - $P_{TAIL}$  complex structure, a region containing  $\alpha 4$  (unstructured  $^{181}DVETA^{185}$  termed  $\delta + \alpha 4$ ) was identified as an additional N interaction site. This interaction occurs independently of the presence of  $\alpha 1/2$ . Its counterpart in N responsible for this additional interaction was identified to be a contiguous region 96–127 in  $N_{CORE}$ . This interaction occurs transiently with a time constant in the order of ms and  $K_d = 0.6$  mM. Although this  $K_d$  is large, the effective concentration  $C_{eff}$  of  $\delta\alpha 4$  around the  $N_{CORE}$  can be significantly enhanced by the  $\alpha 1/2$ - $N_{CORE}$  and  $N_{TAIL}$ -XD associations and the high flexibility of the linker segments, as mentioned in Section 5. An essential role of this  $\delta\alpha 4$ - $N_{CORE}$  interaction in the polymerase reaction was shown by a transcription assay. Since  $\alpha 3$  and the acidic loop are free from the other part of  $P_{TAIL}$  and not bound to  $N_{CORE}$ , the highly flexible motion of the long flexible segment between  $\alpha 1/2$  and  $\delta\alpha 4$  dynamically occupies a large space around  $N_{CORE}$ . This large space occupation potentially inhibits the interaction of RNA, N monomers, and other proteins with  $N_{CORE}$ , as discussed in this study [106]. As demonstrated in the HS-AFM study on a yeast prion protein Sup35, the highly flexible motion of long IDRs can have such a repelling effect [43].

However, despite these structural details revealed by the NMR analysis, the resistance of the N-terminal 27–99 region of PNT to proteolysis cannot be rationalized from the NMR data, suggesting the existence of a large-scale structure undetectable by NMR.

As described below, the combination of the NMR structure of P<sub>TAIL</sub> and results from the HS-AFM imaging of PNT (1–229)-GFP fusion enriched our understanding of the dynamic structure of PNT and its functional relevance. The HS-AFM images of PNT-GFP (Figure 6a) showed an N-terminal globule connected to the GFP moiety via a flexible chain undergoing length changes. In addition, a small globule close to GFP could be occasionally detected, probably corresponding to  $\alpha 4$ . The mean height of the N-terminal globule remained constant at 1.1 nm (Figure 6b) while its lateral dimensions varied. The height of 1.1 nm corresponds to the diameter of a single  $\alpha$ -helix. Therefore, the N-terminal globule contains  $\alpha 1/2$  (and possibly also  $\alpha 3$ ) as the highest particle and other regions lower than  $\alpha 1/2$ . The  $R_{2D}$  value of the flexible chain connecting the GFP and the N-terminal small globule is distributed widely. Its distribution was well fitted to a double Gaussian function with peaks at 8.9 and 14.3 nm (Figure 6c). The mean height of the flexible chain was 0.4–0.5 nm in both metastable shorter and longer states. This suggests that the flexible chain is always fully disordered in both metastable states, except for the occasionally appearing  $\alpha 4$  ( $\alpha 3$  did not appear in the HS-AFM images as a distinct region). From the power law,  $\langle R_{2D} \rangle = 1.16 \text{ nm} \times N_{aa}^{0.517}$ , the  $N_{aa}$  contained in the shorter and longer chains were estimated to be 51 and 128, respectively. The presence of small  $\alpha 3$  and  $\alpha 4$  regions that may be undergoing structural transitions does not much affect this estimation. From these  $N_{aa}$  values, the number of amino acids contained in the N-terminal small globule was approximately estimated to be 178 and 101 for the shorter and longer chain states, respectively. From the number 101, the N-terminal small globule always contains  $\alpha 1/2$  (1–37),  $\alpha 3$  (87–93), and an IDR segment (38–86) connecting  $\alpha 1/2$  and  $\alpha 3$ . Therefore, these regions always form a compact structure, explaining no appearance of  $\alpha 3$  as a distinct region in the HS-AFM images. Importantly, this compact structure also explains the resistance to proteolysis of the N-terminal segment (27–99) containing the fully disordered IDR segment (38–86). The number 178 indicates that the structural transitions of PNT occur on a large scale, mainly by association and dissociation between the N-terminal compact structure and the acidic loop (125–170). From the degree of order propensity,  $K_e = 3.73$ , the larger compact structure, including the acidic loop, is formed more frequently than the smaller compact structure without the acidic loop. The  $k_{OD}$  and  $k_{DO}$  values of the observed global structural transitions were estimated as  $k_{OD} = 0.5 \text{ s}^{-1}$  and  $k_{DO} = 2.0 \text{ s}^{-1}$  from the correlation function of time-series  $R_{2D}$  data (Figure 6d) and the  $K_e$  value. Thus, the dynamic structural features of PNT are quantitatively delineated, as shown in Figure 6e. These rates are again comparable to the elongation rate of the L-P polymerase. In the larger compact form,  $\delta\alpha 4$  cannot reach its interaction site on N<sub>CORE</sub>.  $\alpha 1/2$  also may not be available for N<sub>CORE</sub> binding. Since the  $\delta\alpha 4$ -N<sub>CORE</sub> interaction is essential for RNA transcription, the docking and undocking of the acidic loop and the N-terminal structural unit (i.e., the smaller compact structure) must be regulated at certain steps in the transcription process. The cartwheeling L-P around the nucleocapsid may pull the acidic loop for its undocking, enabling smooth transcription. Since the N-terminal smaller and larger compact structures are not detected in the NMR analysis, some of the structural and dynamic features in the N<sup>0</sup>-P complex model constructed from NMR data have to be amended. However, this complex also needs to be investigated by using HS-AFM.



**Figure 6.** HS-AFM analysis of structure and dynamics of PNT. (a) HS-AFM images and their schematics of PNT-GFP. The open arrowheads point to the N-terminal globule of PNT, while the light-blue arrowhead points to a small globule appeared in the middle of disordered region (very likely corresponding to  $\alpha 4$ ). (b) Height distribution of the N-terminal globule. The value 1.1 nm is identical to the diameter of a single  $\alpha$ -helix. (c) Distribution of end-to-end distance of fully disordered IDR ( $R_{2D}$ ). (d) Autocorrelation function of time-series data of  $R_{2D}$ . The value of  $\lambda$  represents the decay constant of the autocorrelation function. (e) Schematics showing structural and dynamic features of PNT. The top and bottom panels correspond to the more- and less-ordered states, respectively. The numbers in red and blue represent the numbers of amino acids contained in the respective globules and fully disordered IDRs, respectively, estimated by using the power law and measured  $\langle R_{2D} \rangle$  values. The blue arrows indicate the kinetic nature of changes in  $R_{2D}$ .

## 9. Outlook

HS-AFM can visualize IDPs at the single molecule level and identify the structural states of their IDRs (e.g., fully disordered, partially folded, and fully folded states) and their dynamic transitions quantitatively. Importantly, the number of amino acids contained in a fully disordered IDR can be approximately estimated. In contrast, the ensemble NMR spectroscopy methods are good at the atomistic determination of local structures and interaction sites in IDPs but poor at detecting large-scale structures and identifying loosely folded regions. Thus, HS-AFM and NMR mutually complement each other. As such, the combination of these techniques can enrich our understanding of the dynamic structure of IDPs, as demonstrated in the studies of measles viral P protein. This review mainly highlighted the structure and dynamics of IDPs revealed by HS-AFM and their functional implications. However, a few HS-AFM studies have recently been carried out to observe the functional phenomena of IDPs, such as amyloid fibril formation and dissolution and the formation of liquid droplet-like condensates via LLPS. This new trend in HS-AFM applications will be expanded in the near future because tools to analyze the functional phenomena of IDPs are limited, while IDR-involved biological phenomena and diseases are emerging in succession. In this regard, the HS-AFM studies on measles virus IDPs also should be expanded to visualize dynamic processes, including N-P interactions, polymerization of N into N<sup>NUC</sup>, and even RNA synthesis. This line of studies may contribute to the development of new drugs against measles; although the measles vaccine is already available, many people still suffer and die from measles.

**Funding:** This research received no external funding.

**Acknowledgments:** The author acknowledges the wonderful collaboration with Sonia Longhi on the HS-AFM studies of measles virus IDPs.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Uversky, V.N.; Dunker, A.K. Understanding protein non-folding. *Biochim. Biophys. Acta* **2010**, *1804*, 1231–1264. [[CrossRef](#)]
2. Wright, P.E.; Dyson, H.J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 18–29. [[CrossRef](#)]
3. Kumar, N.; Kaushik, R.; Tennakoon, C.; Uversky, V.N.; Longhi, S.; Zhang, K.Y.J.; Bhatia, S. Comprehensive intrinsic disorder analysis of 6108 viral proteomes: From the extent of intrinsic disorder penetrance to functional annotation of disordered viral proteins. *J. Proteome Res.* **2021**, *20*, 2704–2713. [[CrossRef](#)]
4. Chen, J.W.; Romero, P.; Uversky, V.N.; Dunker, A.K. Conservation of intrinsic disorder in protein domains and families: I. A database of conserved predicted disordered regions. *J. Proteome Res.* **2006**, *5*, 879–887. [[CrossRef](#)]
5. Karlin, D.; Longhi, S.; Receveur, V.; Canard, B. The N-Terminal Domain of the Phosphoprotein of Morbilliviruses Belongs to the Natively Unfolded Class of Proteins. *Virology* **2002**, *296*, 251–262. [[CrossRef](#)]
6. Oldfield, C.J.; Meng, J.; Yang, J.Y.; Yang, M.Q.; Uversky, V.N.; Dunker, A.K. Flexible nets: Disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC Genom.* **2008**, *9* (Suppl. S1), S1. [[CrossRef](#)]
7. Dyson, H.J.; Wright, P.E. Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **2002**, *12*, 54–60. [[CrossRef](#)]
8. Sigalov, A.; Aivazian, D.; Stern, L. Homooligomerization of the cytoplasmic domain of the T cell receptor zeta chain and of other proteins containing the immunoreceptor tyrosine-based activation motif. *Biochemistry* **2004**, *43*, 2049–2061. [[CrossRef](#)]
9. Uversky, V.N. Unusual biophysics of intrinsically disordered proteins. *Biochim. Biophys. Acta Proteins Proteom.* **2013**, *1834*, 932–951. [[CrossRef](#)]
10. Schuler, B.; Borgia, A.; Borgia, M.B.; Heidarsson, P.O.; Holmstrom, E.D.; Nettels, D.; Sottini, A. Binding without folding —The biomolecular function of disordered polyelectrolyte complexes. *Curr. Opin. Struct. Biol.* **2020**, *60*, 66–76. [[CrossRef](#)]
11. Borgia, A.; Borgia, M.B.; Bugge, K.; Kissling, V.M.; Heidarsson, P.O.; Fernandes, C.B.; Sottini, A.; Soranno, A.; Buholzer, K.J.; Nettels, D.; et al. Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **2018**, *555*, 61–66. [[CrossRef](#)]
12. Xiang, S.; Kato, M.; Wu, L.C.; Lin, Y.; Ding, M.; Zhang, Y.; Yu, Y.; McKnight, S.L. The LC domain of hnRNPA2 adopts similar conformations in hydrogel polymers, liquid-like droplets, and nuclei. *Cell* **2015**, *163*, 829–839. [[CrossRef](#)]
13. Murthy, A.C.; Dignon, G.L.; Kan, Y.; Zerze, G.H.; Parekh, S.H.; Mittal, J.; Fawzi, N.L. Molecular interactions underlying liquid–liquid phase separation of the FUS low-complexity domain. *Nat. Struct. Mol. Biol.* **2019**, *26*, 637–648. [[CrossRef](#)]
14. Fujioka, Y.; Alam, J.M.; Noshiro, D.; Mouri, K.; Ando, T.; Okada, Y.; May, A.I.; Knorr, R.L.; Suzuki, K.; Ohsumi, Y.; et al. Phase separation organizes the site of autophagosome formation. *Nature* **2020**, *678*, 301–305. [[CrossRef](#)]
15. Brocca, S.; Grandori, R.; Longhi, S.; Uversky, V. Liquid–liquid phase separation by intrinsically disordered protein regions of viruses: Roles in viral life cycle and control of virus–host interactions. *Int. J. Mol. Sci.* **2020**, *21*, 9045. [[CrossRef](#)]
16. Uversky, V. (Ed.) *Droplets of Life: Membrane-Less Organelles, Biomolecular Condensates, and Biological Liquid-Liquid Phase Separation*, 1st ed.; Elsevier: Amsterdam, The Netherlands, 2022; *in press*.
17. Kikhney, A.G.; Svergun, D.I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* **2015**, *539*, 2570–2577. [[CrossRef](#)]
18. Bernado, P.; Svergun, D.I. Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Mol. Biosyst.* **2012**, *8*, 151–167. [[CrossRef](#)]
19. Camacho-Zarco, A.R.; Schnapka, V.; Guseva, S.; Abyzov, A.; Adamski, W.; Milles, S.; Jensen, M.R.; Zidek, L.; Salvi, N.; Blackledge, M. NMR Provides Unique Insight into the Functional Dynamics and Interactions of Intrinsically Disordered Proteins. *Chem. Rev.* **2022**, *122*, 9331–9356. [[CrossRef](#)]
20. Bernadó, P.; Mylonas, E.; Petoukhov, M.V.; Blackledge, M.; Svergun, D.I. Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* **2007**, *129*, 5656–5664. [[CrossRef](#)]
21. Tria, G.; Mertens, H.D.T.; Kachala, M.; Svergun, D.I. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCr* **2015**, *2*, 207–217. [[CrossRef](#)]
22. Pelikan, M.; Hura, G.L.; Hammel, M. Structure and flexibility within proteins as identified through small angle X-ray scattering. *Gen. Physiol. Biophys.* **2009**, *28*, 174–189. [[CrossRef](#)]
23. Schneidman-Duhovny, D.; Hammel, M.; Tainer, J.A.; Sali, A. FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Res.* **2016**, *44*, W424–W429. [[CrossRef](#)]
24. Putnam, C.D.; Hammel, M.; Hura, G.L.; Tainer, J.A. X-ray solution scattering (SAXS) combined with crystallography and computation: Defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **2007**, *40*, 191–285. [[CrossRef](#)]
25. Kooshapur, H.; Schwieters, C.D.; Tjandra, N. Conformational ensemble of disordered proteins probed by solvent paramagnetic relaxation enhancement (sPRE). *Angew. Chem. Int. Ed.* **2018**, *57*, 13519–13522. [[CrossRef](#)]
26. Kodera, N.; Ando, T. Visualization of intrinsically disordered proteins by high-speed atomic force microscopy. *Curr. Opin. Struct. Biol.* **2022**, *72*, 260–266. [[CrossRef](#)]

27. Yamamoto, H.; Fujioka, Y.; Suzuki, S.W.; Noshiro, D.; Suzuki, H.; Kondo-Kakuta, C.; Kimura, Y.; Hirano, H.; Ando, T.; Noda, N.N.; et al. The intrinsically disordered protein Atg13 mediates supramolecular assembly of autophagy initiation complexes. *Dev. Cell* **2016**, *38*, 86–99. [[CrossRef](#)]
28. Kleckner, I.R.; Foster, M.P. An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta* **2011**, *1814*, 942–968. [[CrossRef](#)]
29. Vega, S.; Neira, J.L.; Marcuello, C.; Lostao, A.; Abian, O.; Velazquez-Campoy, A. NS3 protease from hepatitis C virus: Biophysical studies on an intrinsically disordered protein domain. *Int. J. Mol. Sci.* **2013**, *14*, 13282–13306. [[CrossRef](#)]
30. Bizzarria, A.R.; Agostino, S.D.; Andolfi, L.; Cannistraro, S. A combined atomic force microscopy imaging and docking study to investigate the complex between p53 DNA binding domain and Azurin. *J. Mol. Recognit.* **2009**, *22*, 506–515. [[CrossRef](#)]
31. Kodera, N.; Yamamoto, D.; Ishikawa, R.; Ando, T. Video imaging of walking myosin V by high-speed atomic force microscopy. *Nature* **2010**, *468*, 72–76. [[CrossRef](#)]
32. Uchihashi, T.; Iino, R.; Ando, T.; Noji, H. High-speed atomic force microscopy reveals rotary catalysis of rotorless F<sub>1</sub>-ATPase. *Science* **2011**, *333*, 755–758. [[CrossRef](#)]
33. Ando, T.; Uchihashi, T.; Kodera, N. High-speed AFM and applications to biomolecular systems. *Annu. Rev. Biophys.* **2013**, *42*, 393–414. [[CrossRef](#)]
34. Ando, T.; Uchihashi, T.; Scheuring, S. Filming biomolecular processes by high-speed atomic force microscopy. *Chem. Rev.* **2014**, *114*, 3120–3188. [[CrossRef](#)]
35. Uchihashi, T.; Watanabe, Y.; Nakazaki, T.; Yamasaki, T.; Watanabe, H.; Maruno, T.; Ishii, K.; Uchiyama, S.; Song, C.; Murata, K.; et al. Dynamic structural states of ClpB involved in its disaggregation function. *Nat. Commun.* **2018**, *9*, 2147. [[CrossRef](#)]
36. Miyagi, A.; Tsunaka, Y.; Uchihashi, T.; Mayanagi, K.; Hirose, S.; Morikawa, K.; Ando, T. Visualization of intrinsically disordered regions of proteins by high-speed atomic force microscopy. *Chem. Phys. Chem.* **2008**, *9*, 1859–1866. [[CrossRef](#)]
37. Hashimoto, M.; Kodera, N.; Tsunaka, Y.; Oda, M.; Tanimoto, M.; Ando, T.; Morikawa, K.; Tate, S. Phosphorylation-coupled intramolecular dynamics of unstructured regions in chromatin remodeler FACT. *Biophys. J.* **2013**, *104*, 2222–2234. [[CrossRef](#)]
38. Ishino, S.; Yamagami, T.; Kitamura, M.; Kodera, N.; Mori, T.; Sugiyama, S.; Ando, T.; Goda, N.; Tenno, T.; Hiroaki, H.; et al. Multiple interactions of the intrinsically disordered region between the N-terminal helicase and C-terminal nuclease domains of the archaeal Hef protein. *J. Biol. Chem.* **2014**, *289*, 21627–21639. [[CrossRef](#)]
39. Kodera, N.; Uchida, K.; Ando, T.; Aizawa, S. Two-ball structure of the flagellar hook-length control protein FliK as revealed by high-speed atomic force microscopy. *J. Mol. Biol.* **2015**, *427*, 406–414. [[CrossRef](#)]
40. Kodera, N.; Noshiro, D.; Dora, S.K.; Mori, T.; Habchi, J.; Blocquel, D.; Gruet, A.; Dosnon, M.; Salladini, E.; Bignon, C.; et al. Structural and dynamics analysis of intrinsically disordered proteins by high-speed atomic force microscopy. *Nat. Nanotechnol.* **2021**, *16*, 181–189. [[CrossRef](#)]
41. Terahara, N.; Kodera, N.; Uchihashi, T.; Ando, T.; Namba, K.; Minamino, T. Na<sup>+</sup>-induced structural transition of MotPS for stator assembly of the Bacillus flagellar motor. *Sci. Adv.* **2017**, *3*, ea04119. [[CrossRef](#)]
42. Watanabe-Nakayama, T.; Nawa, M.; Konno, H.; Kodera, N.; Ando, T.; Teplow, D.B.; Ono, K. Self- and cross-seeding on alpha-synuclein fibril growth kinetics and structure observed by high-speed atomic force microscopy. *ACS Nano* **2020**, *14*, 9979–9989. [[CrossRef](#)]
43. Konno, H.; Watanabe-Nakayama, T.; Uchihashi, T.; Okuda, M.; Zhu, L.; Kodera, N.; Kikuchi, Y.; Ando, T.; Taguchi, H. Dynamics of oligomer and amyloid fibril formation by yeast prion Sup35 observed by high-speed atomic force microscopy. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 7831–7836. [[CrossRef](#)]
44. Imai, H.; Uchiumi, T.; Kodera, N. Direct visualization of translational GTPase factor pool formed around the archaeal ribosomal P-stalk by high-speed AFM. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 32386–32394. [[CrossRef](#)]
45. Kodera, N.; Ando, T. Guide to studying intrinsically disordered proteins by high-speed atomic force microscopy. *Methods* **2022**, *207*, 44–56. [[CrossRef](#)]
46. Ando, T. *High-Speed Atomic Force Microscopy in Biology*, 1st ed.; Springer: Berlin/Heidelberg, Germany, 2022; pp. 1–319.
47. Jang, J.; Schatz, G.C.; Ratner, M.A. Capillary force in atomic force microscopy. *J. Chem. Phys.* **2004**, *120*, 1157. [[CrossRef](#)]
48. Kohn, J.E.; Millett, I.S.; Jacob, J.; Zagrovic, B.; Dillon, T.M.; Cingel, N.; Dothager, R.S.; Seifert, S.; Thiagarajan, 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](#)]
49. 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](#)]
50. Cantor, C.R.; Schimmel, P.R. Configuration statistics of polymer chain. In *Biophysical Chemistry, Part III: The Behavior of Biological Macromolecules*; W. H. Freeman and Company: San Francisco, CA, USA, 1980; pp. 879–1018.
51. Cordeiro, T.N.; Herranz-Trillo, F.; Urbanek, A.; Estaña, A.; Cortés, J.; Sibille, N.; Bernadó, P. Structural characterization of highly flexible proteins by small-angle scattering. In *Biological Small Angle Scattering: Techniques, Strategies and Tips*; Chaudhuri, B., Muñoz, I., Qian, S., Urban, V., Eds.; Advances in Experimental Medicine and Biology; Springer: Singapore, 2017; Volume 1009, pp. 107–129.
52. 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](#)]



53. Nairn, K.M.; Lyons, R.E.; Mulder, R.J.; Mudie, S.T.; Cookson, D.J.; Lesieur, E.; Kim, M.; Lau, D.; Scholes, F.H.; Elvin, C.M. A synthetic resilin is largely unstructured. *Biophys. J.* **2008**, *95*, 3358–3365. [[CrossRef](#)]
54. Salmon, L.; Nodet, G.; Ozenne, V.; Yin, G.; Jensen, M.R.; Zweckstetter, M.; Blackledge, M. NMR characterization of long-range order in intrinsically disordered proteins. *J. Am. Chem. Soc.* **2010**, *132*, 8407–8418. [[CrossRef](#)]
55. Mylonas, E.; Hascher, A.; Bernadó, P.; Blackledge, M.; Mandelkow, E.; Svergun, D.I. Domain conformation of tau protein studied by solution small-angle X-ray scattering. *Biochemistry* **2008**, *47*, 10345–10353. [[CrossRef](#)]
56. Rivetti, C.; Guthold, M.; Bustamante, C. Scanning force microscopy of DNA deposited onto mica: Equilibration versus kinetic trapping studied by statistical polymer chain analysis. *J. Mol. Biol.* **1996**, *264*, 919–932. [[CrossRef](#)]
57. Baschnagel, J.; Meyer, H.; Wittmer, J.; Kulić, I.; Mohrbach, H.; Ziebert, F.; Nam, G.-M.; Lee, N.-K.; Johner, A. Semiflexible chains at surfaces: Worm-like chains and beyond. *Polymers* **2016**, *8*, 286. [[CrossRef](#)]
58. Kirk, J.; Ilg, P. Chain dynamics in polymer melts at flat surfaces. *Macromolecules* **2017**, *50*, 3703–3718. [[CrossRef](#)]
59. Lenton, S.; Seydel, T.; Nylander, T.; Holt, C.; Hartlein, M.; Teixeira, S.; Zaccai, G. Dynamic footprint of sequestration in the molecular fluctuations of osteopontin. *J. R. Soc. Interface* **2015**, *12*, 0506. [[CrossRef](#)]
60. Konno, T.; Tanaka, N.; Kataoka, M.; Takano, E.; Maki, M. A circular dichroism study of preferential hydration and alcohol effects on a denatured protein, pig calpastatin domain I. *Biochim. Biophys. Acta* **1997**, *1342*, 73–82. [[CrossRef](#)]
61. Gazi, A.D.; Bastaki, M.; Charova, S.N.; Gkougkouli, E.A.; Kapellios, E.A.; Panopoulos, N.J.; Kokkinidis, M. Evidence for a coiled-coil interaction mode of disordered proteins from bacterial type III secretion systems. *J. Biol. Chem.* **2008**, *283*, 34062–34068. [[CrossRef](#)]
62. Kjaergaard, M.; Nørholm, A.B.; Hendus-Altenburger, R.; Pedersen, S.F.; Poulsen, F.M.; Kragelund, B.B. Temperature-dependent structural changes in intrinsically disordered proteins: Formation of alpha-helices or loss of polyproline II? *Protein Sci.* **2010**, *19*, 1555–1564. [[CrossRef](#)]
63. Paz, A.; Zeev-Ben-Mordehai, T.; Lundqvist, M.; Sherman, E.; Mylonas, E.; Weiner, L.; Haran, G.; Svergun, D.I.; Mulder, F.A.A.; Sussman, J.L.; et al. Biophysical characterization of the unstructured cytoplasmic domain of the human neuronal adhesion protein neuroligin 3. *Biophys. J.* **2008**, *95*, 1928–1944. [[CrossRef](#)]
64. Ruskamo, S.; Chukhlieb, M.; Vahokoski, J.; Bhargav, S.P.; Liang, F.; Kursula, I.; Kursula, P. Juxtalin is an intrinsically disordered F-actin-binding protein. *Sci. Rep.* **2012**, *2*, 899. [[CrossRef](#)]
65. Wells, M.; Tidow, H.; Rutherford, T.J.; Markwick, P.; Jensen, M.R.; Mylonas, E.; Svergun, D.I.; Blackledge, M.; Fersht, A.R. Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5762–5767. [[CrossRef](#)]
66. Cragnell, C.; Durand, D.; Cabane, B.; Skepö, M. Coarse-grained modelling of the intrinsically disordered protein Histatin 5 in solution. Monte Carlo simulations in combination with SAXS. *Proteins* **2016**, *84*, 777–791. [[CrossRef](#)]
67. Fuertes, G.; Banterle, N.; Ruff, K.M.; Chowdhury, A.; Mercadante, D.; Koehler, C.; Kachala, M.; Girona, G.E.; Milles, S.; Mishra, A.; et al. Decoupling of size and shape fluctuations in heteropolymeric sequences reconciles discrepancies in SAXS vs. FRET measurements. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E6342–E6351. [[CrossRef](#)]
68. Borgia, A.; Zheng, W.; Buholzer, K.; Borgia, M.B.; Schüler, A.; Hofmann, H.; Soranno, A.; Nettels, D.; Gast, K.; Grishaev, A.; et al. Consistent view of polypeptide chain expansion in chemical denaturants from multiple experimental methods. *J. Am. Chem. Soc.* **2016**, *138*, 11714–11726. [[CrossRef](#)]
69. Moncoq, K.; Broutin, I.; Craescu, C.T.; Vachette, P.; Ducruix, A.; Durand, D. SAXS study of the PIR domain from the Grb14 molecular adaptor: A natively unfolded protein with a transient structure primer? *Biophys. J.* **2004**, *87*, 4056–4064. [[CrossRef](#)]
70. Holehouse, A.S.; Garai, K.; Lyle, N.; Vitalis, A.; Pappu, R.V. Quantitative assessments of the distinct contributions of polypeptide backbone amides versus side chain groups to chain expansion via chemical denaturation. *J. Am. Chem. Soc.* **2015**, *137*, 2984–2995. [[CrossRef](#)]
71. Vigil, D.; Blumenthal, D.K.; Heller, W.T.; Brown, S.; Canaves, J.M.; Taylor, S.S.; Trehwella, J. Conformational differences among solution structures of the type Ia, IIa and IIb protein kinase A regulatory subunit homodimers: Role of the linker regions. *J. Mol. Biol.* **2004**, *337*, 1183–1194. [[CrossRef](#)]
72. Abbott, M.B.; Gaponenko, V.; Abusamhadneh, E.; Finley, N.; Li, G.; Dvoretzky, A.; Rance, M.; Solaro, R.J.; Rosevear, P.R. Regulatory domain conformational exchange and linker region flexibility in cardiac troponin C bound to cardiac troponin I. *J. Biol. Chem.* **2000**, *275*, 20610–20617. [[CrossRef](#)]
73. Morton, C.R.; Rzechorzek, N.J.; Maman, J.D.; Kuramochi, M.; Sekiguchi, H.; Rambo, R.; Sasaki, Y.C.; Davies, O.R.; Pellegrini, L. Structural basis for the coiled-coil architecture of human CtIP. *Open Biol.* **2021**, *11*, 210060. [[CrossRef](#)]
74. Hesgrove, C.S.; Nguyen, K.H.; Biswas, S.; Childs, C.A.; Shradha, K.C.; Medina, B.X.; Alvarado, V.; Sukenik, S.; Yu, F.; Malferrari, M.; et al. Molecular Swiss army knives: Tardigrade CAHS proteins mediate desiccation tolerance through multiple mechanisms. *BioRxiv* **2021**. [[CrossRef](#)]
75. Jaiswal, S.; He, Y.; Lu, H.P. Probing functional conformation-state fluctuation dynamics in recognition binding between calmodulin and target peptide. *J. Chem. Phys.* **2022**, *156*, 055102. [[CrossRef](#)]
76. Komori, K.; Fujikane, R.; Shinagawa, H.; Ishino, Y. Novel endonuclease in Archaea cleaving DNA with various branched structure. *Genes Genet. Syst.* **2002**, *77*, 227–241. [[CrossRef](#)]

77. Komori, K.; Hidaka, M.; Horiuchi, T.; Fujikane, R.; Shinagawa, H.; Ishino, Y. Cooperation of the N-terminal helicase and C-terminal endonuclease activities of archaeal Hef protein in processing stalled replication forks. *J. Biol. Chem.* **2004**, *279*, 53175–53185. [[CrossRef](#)]
78. Fujikane, R.; Ishino, S.; Ishino, Y.; Forterre, P. Genetic analysis of DNA repair in the hyperthermophilic archaeon, *Thermococcus kodakaraensis*. *Genes Genet. Syst.* **2010**, *85*, 243–257. [[CrossRef](#)]
79. Minamino, T.; Gonzalez-Pedrajo, B.; Yamaguchi, K.; Aizawa, S.; Macnab, R.M. FliK, the protein responsible for flagellar hook Two-Ball Structure of FliK length control in *Salmonella*, is exported during hook assembly. *Mol. Microbiol.* **1999**, *34*, 295–304. [[CrossRef](#)]
80. Mizuno, S.; Amida, H.; Kobayashi, N.; Aizawa, S.; Tate, S. The NMR structure of FliK, the trigger for the switch of substrate specificity in the flagellar type III secretion apparatus. *J. Mol. Biol.* **2011**, *409*, 558–573. [[CrossRef](#)]
81. Yao, W.; Li, Y.; Chen, Y.; Chen, Y.; Xie, Y.; Ye, M.; Zhang, Y.; Chen, X.; Wu, X.; Feng, Y.; et al. Atg1-mediated Atg11 phosphorylation is required for selective autophagy by regulating its association with receptor proteins. *Autophagy* **2022**, *in press*. [[CrossRef](#)]
82. Kira, S.; Noguchi, M.; Araki, Y.; Oikawa, Y.; Yoshimori, T.; Miyahara, A.; Noda, T. Vacuolar protein Tag1 and Atg1–Atg13 regulate autophagy termination during persistent starvation in *S. cerevisiae*. *J. Cell Sci.* **2021**, *134*, jcs253682. [[CrossRef](#)]
83. Suzuki, S.W.; Yamamoto, H.; Oikawa, Y.; Kondo-Kakuta, C.; Kimura, Y.; Hirano, H.; Ohsumi, Y. Atg13 HORMA domain recruits Atg9 vesicles during autophagosome formation. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 3350–3355. [[CrossRef](#)]
84. Moriya, N.; Minamino, T.; Hughes, K.T.; Macnab, R.M.; Namba, K. The type III flagellar export specificity switch is dependent on FliK ruler and a molecular clock. *J. Mol. Biol.* **2006**, *359*, 466–477. [[CrossRef](#)]
85. Longhi, S. Nucleocapsid Structure and Function. In *Measles. Current Topics in Microbiology and Immunology*; Griffin, D.E., Oldstone, M.B.A., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; Volume 329, pp. 103–128.
86. Bourhis, J.-M.; Canard, B.; Longhi, S. Structural disorder within the replicative complex of measles virus: Functional implications. *Virology* **2006**, *344*, 94–110. [[CrossRef](#)]
87. Bignon, C.; Troilo, F.; Gianni, S.; Longhi, S. Modulation of measles virus NTAIL interactions through fuzziness and sequence features of disordered binding sites. *Biomolecules* **2019**, *9*, 8. [[CrossRef](#)]
88. Guseva, S.; Milles, S.; Blackledge, M.; Ruigrok, R.W.H. The nucleoprotein and phosphoprotein of measles virus. *Front. Microbiol.* **2019**, *10*, 1832. [[CrossRef](#)]
89. 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](#)]
90. Vreede, F.T.; Jung, T.E.; Brownlee, G.G. Model suggesting that replication of influenza virus Is regulated by stabilization of replicative intermediates. *J. Virol.* **2004**, *78*, 9568–9572. [[CrossRef](#)]
91. Bloyet, L.M.; Brunel, J.; Dosnon, M.; Hamon, V.; Erales, J.; Gruet, A.; Lazert, C.; Bignon, C.; Roche, P.; Longhi, S.; et al. Modulation of re-initiation of measles virus transcription at intergenic regions by PXD to Ntail binding strength. *PLoS Pathog.* **2016**, *12*, e1006058. [[CrossRef](#)]
92. Mavrikakis, M.; Iseni, F.; Mazza, C.; Schoehn, G.; Ebel, C.; Gentzel, M.; Franz, T.; Ruigrok, R.W. Isolation and characterisation of the rabies virus N0-P complex produced in insect cells. *Virology* **2003**, *305*, 406–414. [[CrossRef](#)]
93. Jensen, M.R.; Communie, G.; Ribeiro, E.A., Jr.; Martinez, N.; Desfosses, A.; Salmon, L.; Mollica, L.; Gabel, F.; Jamin, M.; Longhi, S.; et al. Intrinsic disorder in measles virus nucleocapsids. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9839–9844. [[CrossRef](#)]
94. Mohan, A.; Oldfield, C.J.; Radivojac, P.; Vacic, V.; Cortese, M.S.; Dunker, K.; Uversky, V.N. Analysis of molecular recognition features (MoRFs). *J. Mol. Biol.* **2006**, *362*, 1043–1059. [[CrossRef](#)]
95. Gely, S.; Lowry, D.F.; Bernard, C.; Jensen, M.R.; Blackledge, M.; Costanzo, S.; Bourhis, J.-M.; Darbon, H.; Daughdrill, G.; Longhi, S. Solution structure of the C-terminal X domain of the measles virus phosphoprotein and interaction with the intrinsically disordered C-terminal domain of the nucleoprotein. *J. Mol. Recognit.* **2010**, *23*, 435–447. [[CrossRef](#)]
96. Oldfield, C.J.; Cheng, Y.; Cortese, M.S.; Romero, P.; Uversky, V.N.; Dunker, A.K. Coupled folding and binding with  $\alpha$ -helix-forming molecular recognition elements. *Biochemistry* **2005**, *44*, 12454–12470. [[CrossRef](#)]
97. Morin, B.; Bourhis, J.-M.; Belle, V.; Woudstra, M.; Carrière, F.; Guigliarelli, B.; Fournel, A.; Longhi, S. Assessing induced folding of an intrinsically disordered protein by site-directed spin-labeling electron paramagnetic resonance spectroscopy. *J. Phys. Chem. B.* **2006**, *110*, 20596–20608. [[CrossRef](#)]
98. Belle, V.; Rouger, S.; Costanzo, S.; Liquière, E.; Strancar, J.; Guigliarelli, B.; Fournel, A.; Longhi, S. Mapping  $\alpha$ -helical induced folding within the intrinsically disordered C-terminal domain of the measles virus nucleoprotein by site-directed spin-labeling EPR spectroscopy. *Proteins* **2008**, *73*, 973–988. [[CrossRef](#)]
99. Longhi, S.; Belle, V.; Fournel, A.; Guigliarelli, B.; Carrière, F. Probing structural transitions in both structured and disordered proteins using site-directed spin-labeling EPR spectroscopy. *J. Pept. Sci.* **2011**, *17*, 315–328. [[CrossRef](#)]
100. 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](#)]
101. Bourhis, J.M.; Receveur-Bréchet, V.; Oglesbee, M.; Zhang, X.; Buccellato, M.; Darbon, H.; Canard, B.; Finet, S.; Longhi, S. The intrinsically disordered C-terminal domain of the measles virus nucleoprotein interacts with the C-terminal domain of the phosphoprotein via two distinct sites and remains predominantly unfolded. *Protein Sci.* **2005**, *14*, 1975–1992. [[CrossRef](#)]
102. Plumet, S.; Duprex, W.P.; Gerlier, D. Dynamics of viral RNA synthesis during measles virus infection. *J. Virol.* **2005**, *79*, 6900–6908. [[CrossRef](#)]

103. Jensen, M.R.; Houben, K.; Lescop, E.; Blanchard, L.; Ruigrok, R.W.H.; Blackledge, M. Quantitative conformational analysis of partially folded proteins from residual dipolar couplings: Application to the molecular recognition element of Sendai virus nucleoprotein. *J. Am. Chem. Soc.* **2008**, *130*, 8055–8061. [[CrossRef](#)]
104. Habchi, J.; Mamelli, L.; Darbon, H.; Longhi, S. Structural disorder within henipavirus nucleoprotein and phosphoprotein: From predictions to experimental assessment. *PLoS ONE* **2010**, *5*, e11684. [[CrossRef](#)]
105. Guryanov, S.G.; Liljeroos, L.; Kasaragod, P.; Kajander, T.; Butcher, S.J. Crystal structure of the measles virus nucleoprotein core in complex with an N-terminal region of phosphoprotein. *J. Virol.* **2016**, *90*, 2849–2857. [[CrossRef](#)]
106. Milles, S.; Jensen, M.R.; Lazert, C.; Guseva, S.; Ivashchenko, S.; Communie, G.; Maurin, D.; Gerlier, D.; Ruigrok, R.W.H.; Blackledge, M. An ultraweak interaction in the intrinsically disordered replication machinery is essential for measles virus function. *Sci. Adv.* **2018**, *4*, eaat7778. [[CrossRef](#)]