

Review

Roles, Characteristics, and Analysis of Intrinsically Disordered Proteins: A Minireview

Frederik Lermyte 

Department of Chemistry, Technical University of Darmstadt, Alarich-Weiss-Straße 4, 64287 Darmstadt, Germany; frederik.lermyte@tu-darmstadt.de

Received: 22 October 2020; Accepted: 26 November 2020; Published: 30 November 2020



Abstract: In recent years, there has been a growing understanding that a significant fraction of the eukaryotic proteome is intrinsically disordered, and that these conformationally dynamic proteins play a myriad of vital biological roles in both normal and pathological states. In this review, selected examples of intrinsically disordered proteins are highlighted, with particular attention for a few which are relevant in neurological disorders and in viral infection. Next, the underlying causes for intrinsic disorder are discussed, along with computational methods used to predict whether a given amino acid sequence is likely to adopt a folded or unfolded state in solution. Finally, biophysical methods for the analysis of intrinsically disordered proteins will be discussed, as well as the unique challenges they pose in this context due to their highly dynamic nature.

Keywords: intrinsically disordered protein; structural biology; biophysics; mass spectrometry

1. Introduction

In the conventional view of molecular biology, proteins fold to adopt a well-defined three-dimensional structure in order to fulfil their biological function. This folding is driven by thermodynamics, so as to maximize interactions such as hydrogen bonds and salt bridges, while shielding hydrophobic residues from the aqueous environment in the cell (in the standard case of a soluble protein). Intrinsically disordered proteins (IDPs) defy this paradigm by adopting a broad range of transient conformations which are similar in free energy, and with no great kinetic barriers to overcome when transitioning between them. Based on the sequence-based prediction of disorder (see Section 3.1), it was estimated in the early 2000s that one-third of eukaryotic proteins have intrinsically disordered regions (IDRs) more than 30 residues long [1,2]. It is important to note in this context that this does not mean that 30% of eukaryotic proteins possess no higher-order structure, as a spectrum exists between fully structured and fully disordered proteins. Many of the proteins possessing IDRs, for instance, have a mostly folded structure with local disorder, or have folded domains which are connected by disordered linkers (e.g., antibodies) [3]. For the sake of simplicity, in the rest of this review the term “IDP” will be used as a catch-all term for proteins that contain significant IDRs, regardless of whether they also have domains with a stable folded structure. Considering the frequent occurrence of IDPs, as well as their important biological roles (both normal and pathological), it is not surprising that this field has garnered significant interest in recent years, and several excellent reviews have been published [4–12].

A first class of IDPs is notable not for their normal biological function, but for their pathological role in degenerative amyloid diseases. The most prominent example is alpha-synuclein, associated with Parkinson’s disease, multiple system atrophy, and dementia with Lewy bodies [13,14]. However, prion protein (PrP) and amyloid beta also exhibit a significant disorder [15–18]. The transition to a more ordered (often beta-sheet rich), aggregation-prone conformational state leads to oligomerization and eventual fibril formation. The underlying process is not fully understood, although there are

indications that binding to certain ligands or trace metal ions might play a role [19–23]. A related IDP that is relevant to neurodegeneration is tau. First discovered in 1975 and shown to be essential for microtubule assembly [24], hyperphosphorylated isoforms of this protein are the main constituents of the neurofibrillary tangles observed in the brain of Alzheimer’s disease patients and in other so-called “tauopathies” [25,26].

The archetypal example of a functional IDP is the tumor suppression protein p53, sometimes referred to as the “Guardian of the Genome” [27]. By interacting with a myriad of different binding partners, this protein plays several vital roles to maintain genomic stability and promote DNA repair, thereby reducing the frequency of (potentially cancer-causing) mutations [28]. Interestingly, one of the key pathways for this involves another IDP, p21, which in turn is able to bind to and inhibit cyclin-dependent kinase 2, thereby arresting the cell cycle and allowing enough time for damaged DNA to be repaired before it is passed on to the next generation of cells through cell division [29]. Conversely, certain mutant p53 variants are not only unable to facilitate DNA repair, but actively promote tumor progression and metastasis, again through a variety of interactions with other proteins [30]. Another manner in which IDPs interact with the genome is through regulation of gene expression, as silencing of methylated DNA—most often on the cytosine C₅ carbon of a cytosine/guanine dinucleotide (CpG)—is mediated through several methyl-CpG binding domain (MBD) proteins [31], which contain long IDRs [32–35]. The most studied of these proteins is MeCP2, the gene for which is located on the X chromosome. Certain mutations in this gene are lethal in males and linked with Rett syndrome, a progressive neurodevelopmental disorder, in females [36].

Paradoxically, IDPs play a vital role in ensuring that other proteins are folded correctly, since many chaperone proteins exhibit some level of intrinsic disorder [37]. For example, in the bacterial GroEL–GroES complex—a 21-mer composed of 14 GroEL and 7 GroES monomers—a disordered 23-residue C-terminal portion of GroEL faces the central cavity of the complex in which folding occurs. Removing this disordered tail has been shown to lead to a dramatic deterioration in the chaperone function [38]. Similarly, IDRs are found in human small heat shock proteins, produced as part of the cellular stress response, such as Hsp22 and α B-crystallin [39,40]. The examples provided here serve to illustrate the ubiquitous nature and critical biological importance of IDPs. In the rest of this review, the underlying causes for intrinsic disorder will be discussed, both from a physicochemical perspective (i.e., the thermodynamic factors that cause a protein to not adopt a well-defined structure in the solution), as well as an evolutionary one (i.e., how their intrinsic disorder allows these proteins to fulfil their biological role). Finally, a selection of computational, as well as experimental methods for investigating IDPs will be considered.

2. General Characteristics of IDPs

2.1. Physicochemical (Sequence) Characteristics

As mentioned in the Introduction, protein folding is driven by thermodynamics. Hydrogen bonding, dipole interactions, and salt bridges all result in enthalpic stabilization; however, it should be noted that similarly favorable interactions are generally possible with water molecules in the environment. Meanwhile, the entropy of the protein backbone is reduced upon restriction of conformational freedom, but for most folded proteins, this is offset by the ability to shield hydrophobic side chains in the interior of the protein structure, which increases the entropy of the surrounding water molecules. Given these factors, it should come as no surprise that IDPs are generally characterized by a small number of hydrophobic residues, and a large number of charged residues [41,42].

Another common characteristic of IDPs is that they tend to have an uncompensated net charge at physiological pH. Two important metrics used to assess the (dis)ordered nature of protein sequences are the fraction of charged residues (FCR) and net charge per residue (NCPR) [42,43]. FCR is simply the number of residues that have a net charge—either positive or negative—at pH 7 (i.e., D, E, K, R) divided by the total number of amino acid residues in the sequence. NCPR is the net charge (i.e., the number of

positively charged residues minus the number of negatively charged residues) divided by the total number of residues. As stated previously, charged side chains are to some extent a double-edged sword in the context of protein folding—on the one hand, they allow for favorable interactions with water molecules (which favors unfolding), but salt bridges between oppositely charged side chains can also stabilize a folded structure. For this reason, Das and Pappu developed a more sophisticated metric that reflects the linear sequence distribution of positively and negatively charged residues [44]. This charge mixing metric is essentially calculated by evaluating the FCR and NCPR not just for the entire protein, but also across a sliding window of five or six residues (the final step in the calculation involves averaging the results for both window sizes). In this manner, the degree of charge asymmetry in all of these small sub-sequences is calculated, summed, and normalized to the extreme case of a sequence with the same length, FCR, and NCPR, but with perfect charge separation, i.e., all charged residues consolidated into a positive and negative “block”, positioned at both sequence termini. In this manner, a parameter κ is obtained, which varies from 0 for perfectly mixed sequences, to 1 for sequences where positive and negative charges are perfectly segregated. Note that artefacts can occur in this calculation, for example, if there are very few charged residues, which can lead to calculated κ values greater than 1. Sequences with low κ values are more likely to be disordered, as many transient interactions with nearby oppositely charged residues are possible, whereas sequences with oppositely charged blocks are more likely to form hairpin-like conformations where these blocks can interact [44].

In Figure 1, two relevant parameters for five of the examples of IDPs discussed in the Section 1 are shown in a 2D plot (top-left panel; blue dots). On the vertical axis, the fraction of residues incorporated in disordered regions is shown, as predicted by the PONDR-VLXT algorithm (see Section 3.1) [45]. On the horizontal axis, the charge mixing parameter κ is shown. For reference, these two parameters were also calculated for six globular proteins (data points in red). As expected, most of the proteins with significant disordered regions displayed in Figure 1 indeed have κ values below 0.3, with the exception being alpha-synuclein ($\kappa = 0.42$) which has a positively charged N terminus and a negatively charged C terminus, which are known to interact with one another to form transient tertiary structures [46,47]. Five of the six globular proteins, however, also have κ values below 0.3 (range: 0.13–0.40), indicating that while this parameter is useful for comparing permutants of a single sequence, it alone is not sufficient for classifying a protein (region) as ordered or disordered. The more sophisticated prediction method (PONDR), however, was able to accurately classify the globular proteins as possessing less disorder than the IDPs in this small data set. In the other panels of Figure 1, the PONDR score per residue is displayed for each of the five IDPs considered, with regions scoring high (corresponding to a prediction of a more disordered region) displayed in red. Above the graphs are (partial) structures for these proteins, obtained through cryo-electron microscopy (EM), X-ray diffraction (XRD), or nuclear magnetic resonance (NMR). In most cases, structures for the full-length proteins are not available, illustrating the analytical challenge that IDPs pose (discussed further in Section 3.2). In these cases, sequence regions which are not part of the crystal structure are displayed as a dotted black, rather than solid grey, line in the graph.

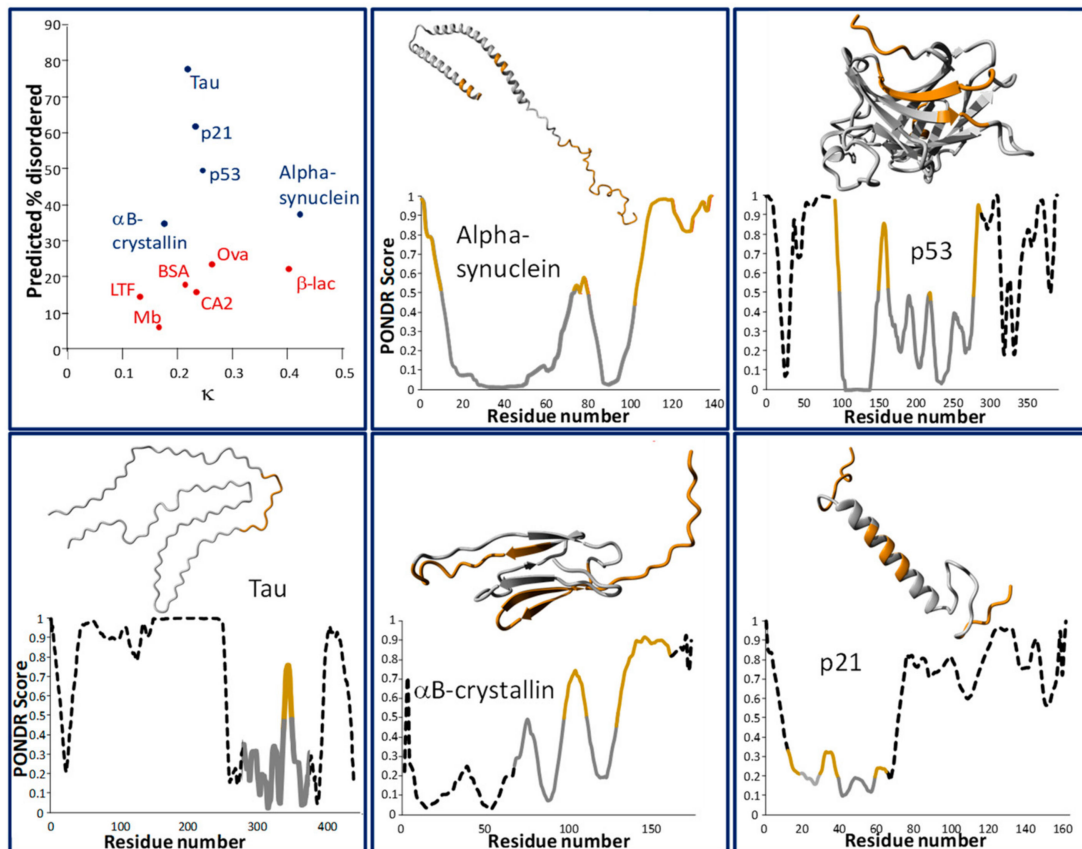


Figure 1. (Top-left) Predicted fraction of residues in disordered regions versus charge mixing parameter κ for five intrinsically disordered (Tau, p21, p53, α B-crystallin, and alpha-synuclein) and six globular (myoglobin, lactotransferritin, carbonic anhydrase 2, bovine serum albumin, chicken ovalbumin, and bovine β -lactoglobulin) proteins. (Other panels) show results for the five intrinsically disordered proteins (IDPs) in detail, displaying (partial, except for alpha-synuclein) experimentally obtained structures and the PONDNR (Predictor of Natural Disordered Regions) score per residue. Note that for several of these structures, a complex with interacting proteins or ligands (not shown) was analyzed rather than the monomeric IDP, resulting in a more ordered structure than might be expected. Regions predicted to be disordered by the PONDNR algorithm (score > 0.5) are shown in gold in both the graphs and structures, except for p21, where a lower cut-off (0.2) was used to highlight the local maxima due to none of the parts of the sequence which are included in the experimentally observed structure scoring as “disordered”. PONDNR scores for regions which are not part of the experimentally obtained structure are shown as a black dotted line in the graphs, regardless of their value. Protein Data Bank (PDB) identifiers for the structures are as follows—tau: 6TJO (441-residue isoform; structure obtained by cryo-EM of filaments); p21: 6P8H (XRD; measured as part of a complex with cyclin-dependent kinase 4 and cyclin D1); p53: 1TUP (XRD; measured as part of a complex with DNA); α B-crystallin: 3L1G (XRD); α SN: 1XQ8 (solution NMR; micelle-bound). Unless otherwise noted, the unmodified human sequence variant of these proteins was used in the calculations. Protein structure visualizations were generated using YASARA View [48].

To more systematically investigate the lack of clear correlation between the κ value and degree of disorder seen for the small data set in Figure 1, three important parameters—NCPR, FCR, and κ —were calculated for (1) all sequences in the DisProt database of disordered proteins (release 2020_06) [49–51], (2) all human proteins in the SwissProt database, and (3) all *E. coli* (strain K12) proteins in SwissProt. As discussed earlier, approximately 30% of human proteins have significant disordered regions, whereas for *E. coli*, this is less than 5% [2]. Of course, 100% of the sequences in DisProt are intrinsically disordered. The result of this calculation is shown in Figure 2. In this figure, each protein with a mass

up to 50 kDa (6656 Disprot sequences, 20,342 human proteins, and 4518 *E. coli* proteins) is represented by a data point in three ((parameter) vs. mass) plots, where the three aforementioned parameters are shown in the first, second, and third row, respectively. Interestingly, while the disordered sequences seem to have a greater diversity of NCPR and FCR values—especially at low masses—it is clear from comparing the three columns that caution is warranted when using only these simple parameters for protein classification.

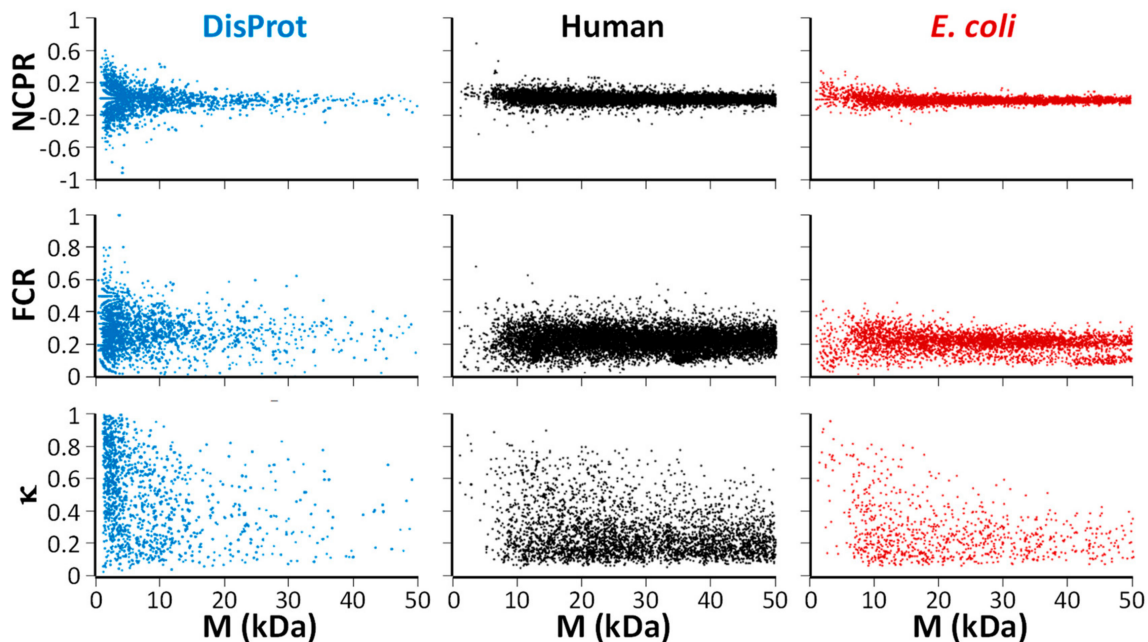


Figure 2. Net charge per residue (NCPR; **first row**), fraction of charged residues (FCR; **second row**), and κ (**third row**) for proteins with masses up to 50 kDa for (**first column**; in blue) sequences in the DisProt database (all intrinsically disordered), (**second column**; in black) human proteins (ca. 30% disordered), and (**third column**; in red) *E. coli* (<5% disordered).

Of note is recent work by Mittag and Pappu et al., in which it was found that, in addition to the patterning of charged residues, patterning of aromatic residues in intrinsically disordered prion-like domains determines their tendency to undergo a liquid-liquid phase separation (LLPS) [52]. This process is increasingly understood to be important for the formation of membraneless organelles such as nucleoli and stress granules [53]. However, it may also play a role in disease, such as neurodegenerative disorders. For example, it has been shown that amyotrophic lateral sclerosis-related mutations in TDP-43 lead to a reduced capacity to undergo LLPS [54]. Conversely, the concentration of proteins such as tau, alpha-synuclein, or huntingtin in phase-separated droplets may provide favorable nucleation conditions for potentially pathological fibril formation [55]. Combining NMR, small-angle X-ray scattering (SAXS), and all-atom simulations, a “stickers-and-spacers” model was developed, in which aromatic residues are the “stickers” that drive intra- and intermolecular interactions. A parameter Ω_{aro} was defined to quantify how uniformly aromatic residues are distributed along the sequence and it was found that, for the prion-like domain of heterogeneous nuclear ribonucleoprotein A1, the aromatic residues are more uniformly spaced (lower Ω_{aro}) than expected by chance ($p < 0.0001$). Other proteins known to undergo LLPS were shown to have similarly high degrees of uniformity in their distribution of aromatic residues along the sequence. Higher values of Ω_{aro} for these proteins were associated with aggregation, rather than phase separation. Thermodynamically, this is explained by the energetic stabilization resulting from the “clumping” of stretches rich in aromatic residues being sufficient to overcome the stabilization provided by solubilization of the more hydrophilic spacers.

Given the importance of charged residues in determining whether a protein adopts a well-defined structure, it can be expected that post-translational modifications—particularly ones such as phosphorylation, which converts a neutral residue to a negatively charged one—could have a significant conformational effect [13]. In one recent example, Mittag and Pappu et al. studied the effect of phosphorylation on the intrinsically disordered region of the *S. cerevisiae* transcription factor Ash1 (residues 420–500) [56]. Remarkably, of the 81 residues in this protein, 10 are possible phosphorylation sites, 17 are charged (16 positive, one negative; FCR = 0.21, NCPR = 0.19, κ = 0.79), and 12 are prolines. Using SAXS, multi-dimensional NMR, and all-atom Monte Carlo simulations, it was found that the global conformational properties of Ash1^{420–500} did not change upon multiple phosphorylation. It was concluded that the conformational behavior in this case could be rationalized by the linear sequence patterning of prolines and charged residues. Note that, while the value for κ is very high, this is due to the fact that nearly all (94%) of the charged residues in the sequence are positive, which illustrates a limitation of the use of this parameter. Interestingly, enhanced R_2 relaxation rates were observed in NMR after phosphorylation, indicating a less dynamic central region (around residues 450–460). The experimental methods used in this work unfortunately could not probe the transient interactions that led to this less dynamic behavior in detail.

One factor which is often neglected is the fact that the cell is a much more crowded environment than the *in vitro* samples typically used in biophysical or structural biology studies, and this can affect the conformational dynamics of proteins. In this context, the possibility to perform *in-cell* NMR must be mentioned, as discussed further in Section 3.2. Schuler et al. have explored the dynamics of intrinsically disordered proteins (C- and N-terminal segments of prothymosin α , activator for thyroid hormones and retinoid receptors (ACTR), and the N-terminal domain of the HIV-1 integrase) in the presence of high volume fractions (up to more than 30%) of polyethylene glycol (PEG) of different average lengths (PEG200 up to PEG35000) [57,58]. Using single-molecule Förster resonance energy transfer (FRET) spectroscopy, they found that energy transfer efficiency increased—indicating compaction of the conformation—upon either an increase in the PEG volume fraction, or chain length. They were able to explain these findings quantitatively using a modified Flory-Huggins theory of polymer solution. This compaction in an environment that more faithfully mimics that of the cell has important implications for the *in vivo* behavior of IDPs, including a reduced capture radius for binding targets, and an increased diffusion coefficient. It should be noted, however, that many transient interactions occur *in cellula*, both specific and non-specific, and for a truly realistic modelling of this environment, an excluded volume effect alone is insufficient [59].

2.2. Evolutionary Characteristics

Having considered the underlying physicochemical causes of why an IDP does not adopt a unique 3D structure, it is useful to also consider the functional reason. The primary benefit of the conformational flexibility of IDPs is generally that it enables a certain level of binding promiscuity—indeed many IDPs are “hubs” in interaction networks and are able to bind/interact with several different targets through an induced fit/conformational selection mechanism, rather than the interactions of more rigid proteins, which resemble more closely the traditional “lock-and-key” mechanism [9,10]. Chaperone proteins provide one example of this, as they need to ensure the proper folding of a variety of substrate proteins. However, the archetypal example of a protein with an extremely broad range of binding partners is p53, as mentioned in the Introduction. This protein displays enormous binding promiscuity, as reviewed in 2016 by Uversky, who showed that the interactome of p53 comprises hundreds of partners [28]. The reason for this exceptional promiscuity—even compared to other IDPs—is that p53 occurs in several different proteoforms: Not only is it able to form homotetramers, but alternative splicing leads to nine relatively common isoforms, and 60 of the 393 residues—many of which are located in intrinsically disordered regions—in this protein can be post-translationally modified. By these mechanisms, hundreds of p53 proteoforms can be produced for specific functions.

Other than the role of IDPs as interaction hubs, a somewhat less commonly cited possibility, which is of significant current interest, is one proposed by Uversky et al. in which intrinsically disordered regions in virus capsids are related to transmission pathways. Specifically, they proposed that capsids with low levels of disorder form a robust protective shell for virions, allowing them to remain infectious outside the body [60–63]. Conversely, shells with higher levels of disorder were suggested to be characteristic of viruses that rely on airborne transmission. In early 2020, they applied this model to the nucleocapsid (N) and membrane (M) proteins of a range of human and animal coronaviruses, including SARS-CoV-2, responsible for the ongoing Covid-19 pandemic [64,65]. Based on this analysis, they concluded that SARS-CoV-2 spreads through both respiratory and faecal-oral pathways, and that virions are sufficiently robust that an infected body is likely to shed large numbers of infectious particles. While current mitigation strategies are primarily focused on preventing airborne transmission through aerosols [66], strategies such as regular handwashing are still promoted throughout the world to prevent fomite-based transmission [67]. It is likely that both pathways play a role to some extent in real-life scenarios, and it will be interesting to see how well this prediction holds up as more data become available over time.

Viruses also provide insight into the selection pressure for disordered regions in proteins which are able to engage multiple targets. An ability to engage the human analogue of the protein used for cell entry in the normal animal reservoir of a virus allows initial entry into the human population, and a certain degree of disorder could be expected to be beneficial in this context. Calculating the charge mixing parameter κ and performing sequence analysis with the PONDR-VLXT algorithm [45] for the key cell entry protein of a range of viruses (spike protein of the seven known human coronaviruses, envelope glycoprotein 120 of human immunodeficiency virus 1, glycoprotein D of herpes simplex virus 1, capsid protein VP1 of human coxsackievirus A21, and viral protein 1 of human rhinovirus 14) reveals evidence of significant levels of disorder in these proteins (see Figure 3). With the exception of MERS-CoV and HCoV-NL63, all 11 viral proteins in this data set had κ values below 0.3. The PONDR algorithm predicted that intrinsically disordered regions comprise up to 35% (in HRV14 and CA21) of these sequences. Interestingly, while coronavirus spike proteins show less overall disorder according to the PONDR analysis than the other four examples, for the betacoronaviruses (SARS-CoV-1, SARS-CoV-2, MERS-CoV, HCoV-OC43, and HCoV-HKU1) the receptor-binding domain (RBD) comprises the most disordered sequence region.

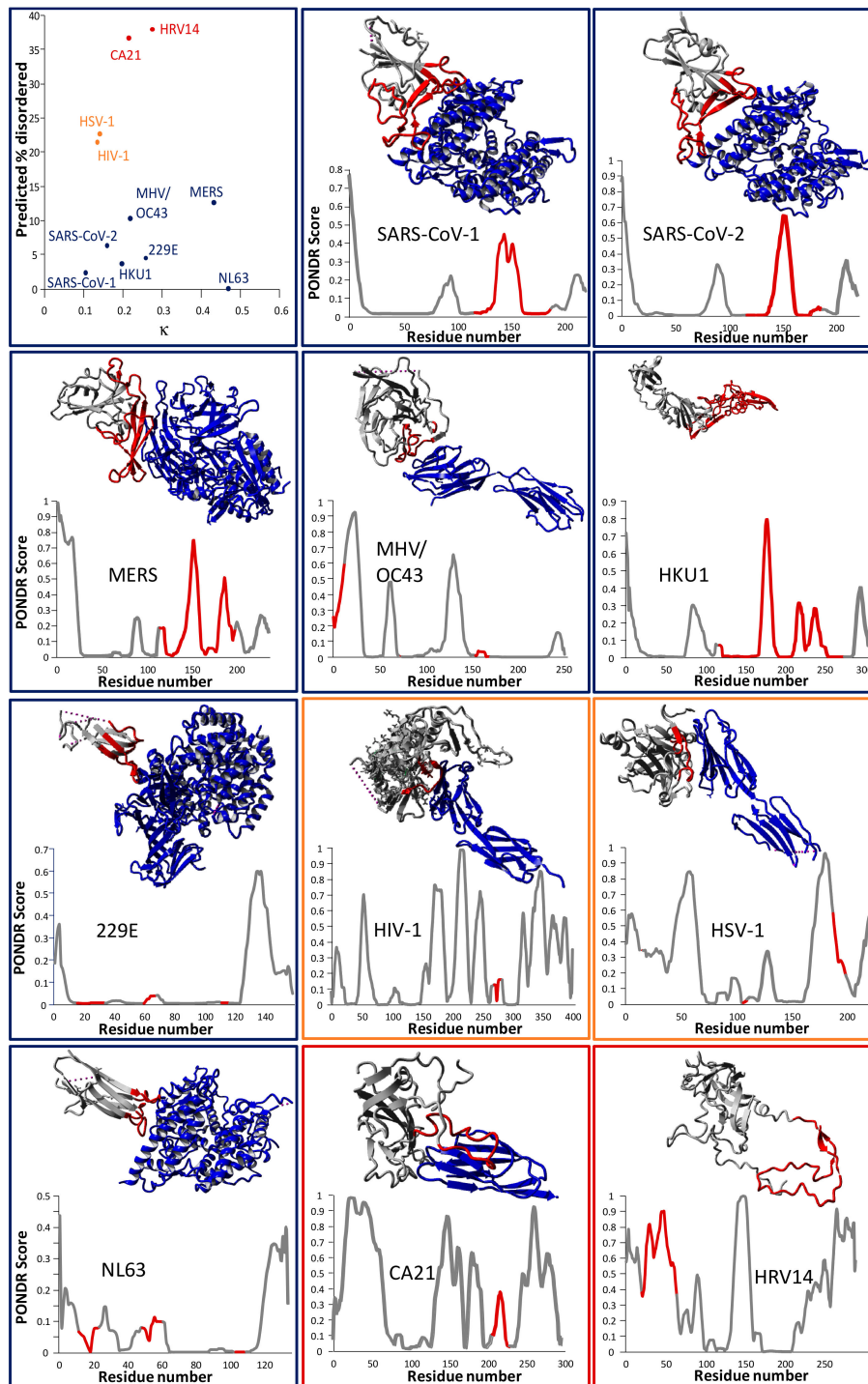


Figure 3. (Top-left) The same analysis as in Figure 1 was performed for the key cell entry proteins of 11 human viruses, i.e., all seven human coronaviruses, HIV-1, HSV-1, CA21, and HRV14. Interestingly, the coronavirus spike proteins (in blue; targeting various receptors) are less disordered overall than the other four viral proteins, but span a range of κ values, while HIV-1/HSV-1 (orange; targeting CD4 [68] and nectin-1 [69], respectively) and CA21/HRV14 (red; both targeting ICAM-1) [70,71] seem to cluster together in the 2D plot. In the (other panels), crystal structures of the viral proteins are shown in the complex with their receptors (except for HCoV-HKU1 and HRV14, for which no structure of this complex was readily available), together with PONDR scores per residue. Border colors of these frames match the dots in the top-left panel. The receptor-binding domain (RBD) is shown in red in both the graphs and the structures. No crystal structure was readily available for the HCoV-OC43 spike protein.

Therefore, the spike protein of the closely related (71% genome identity) murine hepatitis virus A59 (MHV) is shown [72]. PDB identifiers for the structures are as follows—SARS-CoV-1: 2AJF (XRD); SARS-CoV-2: 6LZG (XRD); CA21: 1Z7Z (cryo-EM); HCoV-NL63: 3KBH (XRD); MHV: 3R4D (XRD); HSV-1: 3SKU (XRD); MERS-CoV: 4L72 (XRD); HCoV-HKU1: 5KWB (XRD); HCoV-229E: 6ATK (XRD); HRV14: 1HRV (XRD); HIV-1: 1GC1 (XRD). Grey dashed lines present in some of the structures represent short stretches of the sequence which were not resolved. Protein structure visualizations were generated using YASARA View [48].

Understanding the conformational dynamics of these proteins may have important public health implications—not only is this manuscript being written during the Covid-19 pandemic (caused by SARS-CoV-2), but SARS-CoV-1 caused a major outbreak in 2003 [73]. MERS-CoV, while rare, is highly lethal [73], and genomic analysis of the commonly occurring HCoV-OC43 indicates that bovine-to-human zoonosis occurred around 1890, suggesting that the 1889–1890 pandemic—often attributed to H2N2 influenza—was caused by the first introduction of this virus to an immunologically naive human population [74]. The importance of understanding the structure and dynamics of the spike protein is also highlighted by neutralizing spike protein-reactive antibodies having been shown to persist in recovered Covid-19 patients on a timescale of at least seven months, and by several ongoing vaccine development projects which target this protein [75–77]. The HIV/AIDS pandemic has been ongoing since the 1980s, and continues to claim hundreds of thousands of lives each year [78].

3. Analysis of IDPs

3.1. Computational Methods for Sequence-Based Prediction of Disorder

There is an obvious need for algorithms that can predict, based on a protein sequence, whether that protein is likely to be ordered or disordered under physiological conditions, and in the latter case, which regions will exhibit the greatest degree of disorder. Simple, global physicochemical parameters such as FCR, NCP, and κ are trivial to calculate, but dozens of computational algorithms have been developed over the years. Many of these were recently reviewed by Liu et al. [79]. Broadly speaking, these algorithms can be classified based on their complexity. The simplest ones use physicochemical properties, discussed in Section 2.1, to predict the disorder from first principles. An example of this is FoldIndex [80], which works by evaluating the mean net charge and hydrophobicity as defined by Uversky et al. [41] across a sliding window in order to identify “regional” folding propensities throughout a given sequence. IUPred [81–83] is designed along somewhat different principles and predicts disorder based on pairwise intramolecular interaction energies between amino acid residues in a sequence. The assumption in this case is that if the stabilization from such interactions is insufficient to offset the reduction in entropy that results from folding, then the protein is likely to adopt a disordered state. GlobPlot [84] is another relatively straightforward algorithm, and defines a propensity for each of the 20 common amino acid residues to be in an ordered or disordered sequence region. This is again evaluated across a sliding window. TopIDP [85] is built following similar principles and ranks amino acid residues as (W, F, Y, I, M, L, V, N, C, T, A, G, R, D, H, Q, K, S, E, P) from order- to disorder-promoting.

A more advanced class of algorithms uses machine learning to distinguish ordered from disordered sequences. While these have the benefit of being based on training sets containing empirical data, results are not as straightforward to rationalize as for the physicochemical methods. To improve the chances of obtaining accurate predictions from these somewhat opaque algorithms, “meta” methods make up a third category and incorporate several of the predictors from the first two categories and “fuse” these to form a consensus prediction. The PONDR series of algorithms utilizes a neural network to predict disorder. Interestingly, during the development of these algorithms, the authors found that slightly different versions of the algorithm yielded more accurate disorder predictions in sequences of different chain lengths [45,86,87]. Other popular machine-learning algorithms for disorder prediction include DISOPRED [88] and DISOPRED2 [2], developed at University College London, and DisEMBL [89],

developed at the European Molecular Biology Laboratory. Very recently, a new neural network-based method called ODINPred was introduced, with initial results indicating that it may be able to outperform many older algorithms in a head-to-head comparison. The PONDR family of algorithms also includes a meta method, i.e., PONDR-FIT [90]. This combines machine-learning algorithms PONDR-VLXT, PONDR-VL3, and PONDR-VSL2 with physicochemical algorithms IUPred, FoldIndex, and TopIDP. After evaluating a sequence with these six algorithms in parallel, a consensus prediction is generated across a sliding window. The integration of these six outputs is achieved through another neural network, which was shown to result in more informative consensus outputs than simple vote-counting or a linear combination of scores from independent algorithms. DISOPRED3 [91] is a meta-predictor which was developed using similar principles, and builds on the DISOPRED2 algorithm.

Finally, moving beyond the general prediction of order or disorder, computational methods—specifically, molecular dynamics simulations—can be used to supplement experimental results and obtain more detailed structural insights. This has the benefit of allowing the probing of systems and timescales which are not (easily) experimentally accessible, and can therefore be a valuable tool in the study of IDPs [92–94]. Care must be taken in such studies, however, as it has been shown that results of such MD studies can depend strongly on how the simulation was set up. Specifically, the choice of force field can have a strong effect on the outputs, as shown by Rauscher et al. [95]. In their work, these authors showed that the CHARMM 22* force field was the best option overall, based on a comparison to SAXS and NMR data for a disordered arginine/serine peptide [96]; however, it remains to be seen to what extent this result can be generalized to other IDPs.

3.2. Experimental Methods for Structural Characterization of IDPs

Selected methods of particular importance are highlighted in this section. For a more comprehensive overview, we refer to the excellent recent review by Longhi et al. [97]. The two most commonly used methods in structural biology are XRD and NMR. XRD requires that samples are in a crystalline state, i.e., that the molecules in the sample adopt a repeating pattern of well-defined conformations. This is complicated due to the tendency of IDPs to dynamically sample a large conformational space, and although it is possible to obtain crystal structures from pure IDPs [98], crystals are often derived from either smaller fragments that tend to adopt a well-defined structure, or from noncovalent complexes with binding partners that induce folding [99,100]. A method that circumvents the need for large monocrystals is micro-electron diffraction (microED), pioneered by Eisenberg and Gonen [101,102]. In this cryo-EM method, data can rapidly be obtained from crystals which are smaller than the wavelength of visible light and hence “invisible”. Typically, data sets from multiple nanocrystals are averaged to obtain higher-quality structures. This approach was first demonstrated on an 11-residue peptide derived from the non-amyloid beta component (NAC) region of alpha-synuclein [101]. Other structures solved by this method which are relevant for gaining insight into IDPs include a hexa- and heptapeptide derived from the amyloid core of the Sup35 prion protein [103] and tau-derived peptide VQIVYK [104]. A different X-ray based method, SAXS, is able to provide a degree of insight into the behavior of IDPs in solution. More precisely, the size distribution of proteins in solution can be obtained from this method, providing a measure of the compactness of the solution structure [105].

Multidimensional NMR spectroscopy of IDPs is challenging due to the rapid interconversion between transient conformational states, leading to a severe spectral overlap. The chemical exchange of amide protons with surrounding water molecules is also common and leads to poor signal-to-noise ratios [106]. Performing the experiment at low temperatures can ameliorate this issue to some degree. Alternatively, detection of other NMR-active nuclei (particularly ^{13}C and ^{15}N) can also overcome the problem of amide proton exchange, and has the benefit of increased chemical shift dispersion compared to ^1H -NMR [107,108]. Isotopic enrichment is often required for this, and therefore, homonuclear ^{13}C - ^{13}C coupling is a potential issue. In particular, coupling between the carbonyl carbon and amide nitrogen provides information on sequential residues, and is also able to probe prolines, which lack an amide proton altogether.

An elegant NMR-based study was carried out by Veglia et al., who, using a combination of solid-state NMR and solution NMR chemical exchange saturation transfer, studied the interaction of alpha-synuclein with the membranes of synaptic-like vesicles, and showed that the N-terminal, C-terminal, and central NAC regions of this IDP have distinct functions in this interaction [109]. In this work, the N-terminus adopted an alpha-helical structure in these experiments and anchored the protein to the vesicle, while the C-terminus remained largely unstructured. Interestingly, the behavior of the terminal regions was rather insensitive to the lipid composition of the membrane, while the NAC region modulated the interaction strength in a lipid-selective fashion [109]. Given the increased awareness in recent years of the importance of interactions of amyloidogenic disordered peptides and proteins with membranes in several neurodegenerative diseases, this type of work could provide valuable insights into disease mechanisms [110–112]. Further information on the use of NMR for the *in vitro* study of IDPs can be found in the excellent review by Konrat [106]. NMR measurements can also be carried out *in cellula*, which requires the introduction of protein that has been enriched in NMR-active nuclei. This can be accomplished in several ways, for example, the microinjection of exogenously produced protein, or diffusion into the cell through the membrane after treatment with pore-forming toxins. Alternatively, overexpression of proteins of interest can be induced in the cells in which the analysis is to be performed. As this overexpression will outpace normal protein production, transferring the cells to an isotopically-labelled medium at this stage, results in a significant concentration of isotopically-labelled protein of interest. For further information on these methods, the reader is directed to the review by Selenko et al. [113].

Mass spectrometry (MS) offers intriguing avenues for studying the structure of proteins in the gas phase. Electrospray ionization (ESI) is most commonly used for structural studies. Generally speaking, for globular, water-soluble proteins, if they are in their most native-like state immediately prior to the ionization process (in practice, this means ESI from aqueous solutions at physiological pH) this will result in compact, folded protein ions with a narrow distribution of charge states [114]. In contrast to this “native MS” approach, denaturing the protein in solution (e.g., by addition of an organic solvent or lowering the solution pH) results in extended gas-phase conformations with a broad charge state distribution, and possessing higher average charge states than under native conditions. The exact underlying reason for this behavior is still actively being investigated, although it has been proposed that folded and unfolded proteins are transferred into the gas phase through different ESI mechanisms [115]. IDPs possess a relatively extended structure under physiological conditions, and therefore, exhibit a characteristic behavior in MS. Specifically, they consistently generate broad, “non-native-like” charge state distributions under “native MS” conditions [116–122]. These broad charge state distributions correspond to a tremendous conformational heterogeneity, and it has been shown that IDPs are in fact able to sample an even greater conformational space in the gas phase than in solution [123]. This was accomplished through a combination of molecular dynamics simulations, native MS, and ion mobility (IM) spectrometry [3]. This third technique relies on the gas-phase separation of ions based on their electrophoretic mobility through an inert gas, providing a measure of their rotationally averaged collision cross-section. Recently, Barran et al. investigated the IM-MS behavior of three permutants of the C-terminal IDR of the protein p27^{Kip1}, with values for the charge mixing parameter κ of 0.14, 0.27 (wild-type), and 0.56. They found that a higher value for κ resulted in a more compact gas-phase conformation (smaller collision cross-section) and a shift of the charge state distribution to produce fewer, and lower charge states, consistent with the principles outlined above [124].

Beyond the global (un)folding state from charge state distributions and IM measurements, there are several methods available to obtain structural details through MS. One class of methods involves labelling the protein in the solution in a conformation-sensitive manner, after which the conventional MS/MS analysis maps the number and position of labels. One of the most classic experiments is hydrogen-deuterium exchange [125], in which amide hydrogen atoms are exchanged for deuterium after diluting the protein solution in deuterium oxide. The kinetics of this exchange depend on the

protein conformation and dynamics. This creates practical problems for fully unfolded states, as the exchange goes to essentially 100% in milliseconds. However, this method is exquisitely sensitive for probing transient local folding or stabilization, for example, by binding of the IDP to an interaction partner [121,126]. A more direct approach is also possible, in which no solution labelling is performed, but MS/MS is performed of proteins ionized under native-like conditions. For both folded proteins and IDPs, this often results in a fragmentation pattern which is dependent on the protein charge state and conformation, so that structural information can be inferred from this pattern [121,127–129].

Single-molecule FRET (Förster/fluorescence resonance energy transfer) is another popular method for structural analysis of IDPs [130,131]. An example was discussed earlier, in which the effect of molecular crowding on the conformational ensemble of IDPs was investigated [57,58]. Essentially, in this technique a protein is labelled at two sites with two different chromophores. This usually requires the introduction of two cysteine residues through mutation and recombinant expression. One of the chromophores, known as the donor, is excited by irradiation at its absorption maximum, and is able to transfer its energy to the second (acceptor) chromophore by dipole-dipole coupling. This then emits a photon at slightly lower energy than that used for excitation of the donor, and this emission is measured. The efficiency of this process drops off quickly as the distance between both chromophores increases (inverse sixth power relation), and therefore, FRET can be used to accurately measure pairwise distances between residues.

A conceptually somewhat similar technique to FRET is tryptophan-cysteine quenching [132]. In this approach, a tryptophan residue is excited to the triplet state using a laser pulse. In the absence of any quenchers, this state has a lifetime of around 40 μ s. Cysteine, however, is able to act as an efficient quencher and can induce decay to the singlet ground state in as little as 100 ns. Therefore, the measurement of the lifetime of the excited state can be used to measure the rate of intramolecular contact formation between the tryptophan and cysteine residues [132]. This method has been used by Lapidus et al. to study the aggregation of wild-type alpha-synuclein, as well as several Parkinson's disease-causing mutants. As this protein lacks both cysteine and tryptophan, a double mutant had to be engineered to enable these experiments [133,134].

4. Conclusions and Perspective

Intrinsically disordered proteins defy the conventional structure-function paradigm for how proteins operate, and represent some of the most conformationally dynamic biomolecules known. While they have been described as “mysterious” [12] and a “dark proteome” [135–137] due to the challenge they pose for structural biologists, it is clear that there is order to this chaos, and that their most accurate characterization might be that of “interaction specialists” [10]. Due to both their tendency to be involved in many different cellular processes, and the fact that they are balanced on a conformational knife's edge, mutations in these proteins often have important biomedical implications. This phenomenon has been captured in the D² (“disorder in disorders”) concept [138], and makes studying these proteins all the more important. Recent advances in structural biology technology for the condensed state, including the cryo-EM revolution, have made the analysis of these proteins more tractable. Approaches now exist to trap disordered proteins in a particular conformation—often as part of a complex—allowing information to be obtained through, e.g., X-ray crystallography. Meanwhile, mass spectrometry has emerged as a powerful method for IDP analysis in the gas phase, and this analysis is not inherently more difficult than for structured proteins. Undoubtedly, the coming years will see even more studies into the conformational ensembles and dynamics of intrinsically disordered proteins, providing new insights into their many, varied biological roles.

Funding: This work was funded by the LOEWE project TRABITA funded by the Ministry of Higher education, Research and the Arts (HMWK) of the state of Hesse. The APC was funded through support by the German Research Foundation and the Open Access Publishing Fund of the Technical University of Darmstadt.

Acknowledgments: This manuscript has benefited from insightful comments from the reviewers.

Conflicts of Interest: There are no conflicts of interest to declare.

References

1. Dunker, A.K.; Lawson, J.D.; Brown, C.J.; Williams, R.M.; Romero, P.; Oh, J.S.; Oldfield, C.J.; Campen, A.M.; Ratliff, C.M.; Higgs, K.W.; et al. Intrinsically disordered protein. *J. Mol. Graph. Model.* **2001**, *19*, 26–59. [[CrossRef](#)]
2. Ward, J.J.; Sodhi, J.S.; McGuffin, L.J.; Buxton, B.F.; Jones, D.T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **2004**, *337*, 635–645. [[CrossRef](#)] [[PubMed](#)]
3. Stuchfield, D.; Barran, P. Unique insights to intrinsically disordered proteins provided by ion mobility mass spectrometry. *Curr. Opin. Chem. Biol.* **2018**, *42*, 177–185. [[CrossRef](#)] [[PubMed](#)]
4. Receveur-Brechot, V.; Bourhis, J.M.; Uversky, V.N.; Canard, B.; Longhi, S. Assessing protein disorder and induced folding. *Proteins Struct. Funct. Bioinform.* **2006**, *62*, 24–45. [[CrossRef](#)] [[PubMed](#)]
5. Uversky, V.N. Intrinsically disordered proteins from A to Z. *Int. J. Biochem. Cell Biol.* **2011**, *43*, 1090–1103. [[CrossRef](#)]
6. Lee, V.D.R.; Buljan, M.; Lang, B.; Weatheritt, R.J.; Daughdrill, G.W.; Dunker, A.K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D.T.; et al. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **2014**, *114*, 6589–6631.
7. Oldfield, C.J.; Dunker, A.K. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* **2014**, *83*, 553–584. [[CrossRef](#)]
8. Habchi, J.; Tompa, P.; Longhi, S.; Uversky, V.N. Introducing Protein Intrinsic Disorder. *Chem. Rev.* **2014**, *114*, 6561–6588. [[CrossRef](#)]
9. Wright, P.E.; Dyson, H.J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell. Biol.* **2015**, *16*, 18–29. [[CrossRef](#)]
10. Tompa, P.; Schad, E.; Tantos, A.; Kalmar, L. Intrinsically disordered proteins: Emerging interaction specialists. *Curr. Opin. Struct. Biol.* **2015**, *35*, 49–59. [[CrossRef](#)]
11. Pauwels, K.; Lebrun, P.; Tompa, P. To be disordered or not to be disordered: Is that still a question for proteins in the cell? *Cell. Mol. Life Sci.* **2017**, *74*, 3185–3204. [[CrossRef](#)] [[PubMed](#)]
12. Uversky, V.N. Intrinsically Disordered Proteins and Their “Mysterious” (Meta)Physics. *Front. Phys. Lausanne* **2019**, *7*, 10. [[CrossRef](#)]
13. Kang, L.; Moriarty, G.M.; Woods, L.A.; Ashcroft, A.E.; Radford, S.E.; Baum, J. N-terminal acetylation of α -synuclein induces increased transient helical propensity and decreased aggregation rates in the intrinsically disordered monomer. *Protein Sci.* **2012**, *21*, 911–917. [[CrossRef](#)] [[PubMed](#)]
14. Kim, W.S.; Kagedal, K.; Halliday, G.M. Alpha-synuclein biology in Lewy body diseases. *Alzheimer’s Res. Ther.* **2014**, *6*, 73. [[CrossRef](#)] [[PubMed](#)]
15. Riek, R.; Hornemann, S.; Wider, G.; Billeter, M.; Glockshuber, R.; Wuthrich, K. NMR structure of the mouse prion protein domain PrP(121–231). *Nature* **1996**, *382*, 180–182. [[CrossRef](#)] [[PubMed](#)]
16. Donne, D.G.; Viles, J.H.; Groth, D.; Mehlhorn, I.; James, T.L.; Cohen, F.E.; Prusiner, S.B.; Wright, P.E.; Dyson, H.J. Structure of the recombinant full-length hamster prion protein PrP(29–231): The N terminus is highly flexible. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 13452–13457. [[CrossRef](#)]
17. Riek, R.; Hornemann, S.; Wider, G.; Glockshuber, R.; Wuthrich, K. NMR characterization of the full-length recombinant murine prion protein, mPrP(23–231). *FEBS Lett.* **1997**, *413*, 282–288. [[CrossRef](#)]
18. Maity, B.K.; Das, A.K.; Dey, S.; Moorthi, U.K.; Kaur, A.; Dey, A.; Surendran, D.; Pandit, R.; Kallianpur, M.; Chandra, B.; et al. Ordered and Disordered Segments of Amyloid- β Drive Sequential Steps of the Toxic Pathway. *ACS Chem. Neurosci.* **2019**, *10*, 2498–2509. [[CrossRef](#)]
19. Viles, J.H. Metal ions and amyloid fiber formation in neurodegenerative diseases. Copper, zinc and iron in Alzheimer’s, Parkinson’s and prion diseases. *Coord. Chem. Rev.* **2012**, *256*, 2271–2284. [[CrossRef](#)]
20. Faller, P.; Hureau, C.; la Penna, G. Metal Ions and Intrinsically Disordered Proteins and Peptides: From Cu/Zn Amyloid- β to General Principles. *Acc. Chem. Res.* **2014**, *47*, 2252–2259. [[CrossRef](#)]
21. Wongkongkathep, P.; Han, J.Y.; Choi, T.S.; Yin, S.; Kim, H.I.; Loo, J.A. Native Top-Down Mass Spectrometry and Ion Mobility MS for Characterizing the Cobalt and Manganese Metal Binding of α -Synuclein Protein. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 1870–1880. [[CrossRef](#)] [[PubMed](#)]
22. Lermyte, F.; Everett, J.; Brooks, J.; Bellingeri, F.; Billimoria, K.; Sadler, P.J.; O’Connor, P.B.; Telling, N.D.; Collingwood, J.F. Emerging Approaches to Investigate the Influence of Transition Metals in the Proteinopathies. *Cells* **2019**, *8*, 1231. [[CrossRef](#)] [[PubMed](#)]

23. Lermyte, F.; Everett, J.; Lam, Y.P.Y.; Wootton, C.A.; Brooks, J.; Barrow, M.P.; Telling, N.D.; Sadler, P.J.; O'Connor, P.B.; Collingwood, J.F. Metal Ion Binding to the Amyloid β Monomer Studied by Native Top-Down FTICR Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 2123–2134. [[CrossRef](#)] [[PubMed](#)]
24. Weingarten, M.D.; Lockwood, A.H.; Hwo, S.Y.; Kirschner, M.W. A protein factor essential for microtubule assembly. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 1858–1862. [[CrossRef](#)] [[PubMed](#)]
25. Braak, H.; Braak, E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol.* **1991**, *82*, 239–259. [[CrossRef](#)]
26. Alonso, A.; Zaidi, T.; Novak, M.; Grundke-Iqbal, I.; Iqbal, K. Hyperphosphorylation induces self-assembly of tau into tangles of paired helical filaments/straight filaments. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6923–6928. [[CrossRef](#)]
27. Lane, D.P. Cancer. p53, guardian of the genome. *Nature* **1992**, *358*, 15–16. [[CrossRef](#)]
28. Uversky, V.N. p53 Proteoforms and Intrinsic Disorder: An Illustration of the Protein Structure-Function Continuum Concept. *Int. J. Mol. Sci.* **2016**, *17*, 1874. [[CrossRef](#)]
29. Abbas, T.; Dutta, A. p21 in cancer: Intricate networks and multiple activities. *Nat. Rev. Cancer* **2009**, *9*, 400–414. [[CrossRef](#)]
30. Mantovani, F.; Collavin, L.; Sal, G.D. Mutant p53 as a guardian of the cancer cell. *Cell Death Differ.* **2019**, *26*, 199–212. [[CrossRef](#)]
31. Ballestar, E.; Wolffe, A.P. Methyl-CpG-binding proteins. Targeting specific gene repression. *Eur. J. Biochem.* **2001**, *268*, 1–6. [[CrossRef](#)]
32. Hite, K.C.; Kalashnikova, A.A.; Hansen, J.C. Coil-to-helix transitions in intrinsically disordered methyl CpG binding protein 2 and its isolated domains. *Protein Sci.* **2012**, *21*, 531–538. [[CrossRef](#)] [[PubMed](#)]
33. Hameed, U.F.; Lim, J.; Zhang, Q.; Wasik, M.A.; Yang, D.; Swaminathan, K. Transcriptional repressor domain of MBD1 is intrinsically disordered and interacts with its binding partners in a selective manner. *Sci. Rep.* **2014**, *4*, 4896. [[CrossRef](#)] [[PubMed](#)]
34. Desai, M.A.; Webb, H.D.; Sinanan, L.M.; Scarsdale, J.N.; Walavalkar, N.M.; Ginder, G.D.; Williams, D.C. An intrinsically disordered region of methyl-CpG binding domain protein 2 (MBD2) recruits the histone deacetylase core of the NuRD complex. *Nucleic Acids Res.* **2015**, *43*, 3100–3113. [[CrossRef](#)] [[PubMed](#)]
35. Kim, M.Y.; Na, I.; Kim, J.S.; Son, S.H.; Choi, S.; Lee, S.E.; Kim, J.H.; Jang, K.; Alterovitz, G.; Chen, Y.; et al. Rational discovery of antimetastatic agents targeting the intrinsically disordered region of MBD2. *Sci. Adv.* **2019**, *5*, eaav9810. [[CrossRef](#)] [[PubMed](#)]
36. Amir, R.E.; van den Veyver, I.B.; Wan, M.; Tran, C.Q.; Francke, U.; Zoghbi, H.Y. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **1999**, *23*, 185–188. [[CrossRef](#)] [[PubMed](#)]
37. Tompa, P.; Kovacs, D. Intrinsically disordered chaperones in plants and animals. *Biochem. Cell Biol.* **2010**, *88*, 167–174. [[CrossRef](#)]
38. Machida, K.; Kono-Okada, A.; Hongo, K.; Mizobata, T.; Kawata, Y. Hydrophilic Residues 526KNDAAAD531 in the Flexible C-terminal Region of the Chaperonin GroEL Are Critical for Substrate Protein Folding within the Central Cavity*. *J. Biol. Chem.* **2008**, *283*, 6886–6896. [[CrossRef](#)]
39. Kazakov, A.S.; Markov, D.I.; Gusev, N.B.; Levitsky, D.I. Thermally induced structural changes of intrinsically disordered small heat shock protein Hsp22. *Biophys. Chem.* **2009**, *145*, 79–85. [[CrossRef](#)]
40. Sudnitsyna, M.V.; Mymrikov, E.V.; Seit-Nebi, A.S.; Gusev, N.B. The role of intrinsically disordered regions in the structure and functioning of small heat shock proteins. *Curr. Protein Pept. Sci.* **2012**, *13*, 76–85. [[CrossRef](#)]
41. Uversky, V.N.; Gillespie, J.R.; Fink, A.L. Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins Struct. Funct. Genet.* **2000**, *41*, 415–427. [[CrossRef](#)]
42. Uversky, V.N. What does it mean to be natively unfolded? *Eur. J. Biochem.* **2002**, *269*, 2–12. [[CrossRef](#)] [[PubMed](#)]
43. Mao, A.H.; Crick, S.L.; Vitalis, A.; Chicoine, C.L.; Pappu, R.V. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 8183–8188. [[CrossRef](#)] [[PubMed](#)]
44. Das, R.K.; Pappu, R.V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 13392–13397. [[CrossRef](#)] [[PubMed](#)]

45. Romero, P.; Obradovic, Z.; Li, X.; Garner, E.C.; Brown, C.J.; Dunker, A.K. Sequence complexity of disordered protein. *Proteins* **2001**, *42*, 38–48. [[CrossRef](#)]
46. Dedmon, M.M.; Lindorff-Larsen, K.; Christodoulou, J.; Vendruscolo, M.; Christopher, M.D. Mapping Long-Range Interactions in α -Synuclein using Spin-Label NMR and Ensemble Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2005**, *127*, 476–477. [[CrossRef](#)] [[PubMed](#)]
47. Bertonecini, C.W.; Jung, Y.S.; Fernandez, C.O.; Hoyer, W.; Griesinger, C.; Jovin, T.M.; Zweckstetter, M. Mapping Long-Range Interactions in α -Synuclein using Spin-Label NMR and Ensemble Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 1430–1435. [[CrossRef](#)] [[PubMed](#)]
48. Krieger, E.; Vriend, G. YASARA View—Molecular graphics for all devices—From smartphones to workstation. *Bioinformatics* **2014**, *30*, 2981–2982. [[CrossRef](#)]
49. Uversky, V.N. Natively unfolded proteins: A point where biology waits for physics. *Protein Sci.* **2002**, *11*, 739–756. [[CrossRef](#)]
50. Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **2002**, *27*, 527–533. [[CrossRef](#)]
51. Dunker, A.K.; Brown, C.J.; Lawson, J.D.; Iakoucheva, L.M.; Obradovic, Z. Intrinsic disorder and protein function. *Biochemistry* **2002**, *41*, 6573–6582. [[CrossRef](#)] [[PubMed](#)]
52. Martin, E.W.; Holehouse, A.S.; Peran, I.; Farag, M.; Incicco, J.J.; Bremer, A.; Grace, C.R.; Soranno, A.; Pappu, R.V.; Mittag, T. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **2020**, *367*, 694–699. [[CrossRef](#)] [[PubMed](#)]
53. Alberti, S.; Gladfelter, A.; Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **2019**, *176*, 419–434. [[CrossRef](#)] [[PubMed](#)]
54. Conicella, A.E.; Zerze, G.H.; Mittal, J.; Fawzi, N.L. ALS Mutations Disrupt Phase Separation Mediated by α -Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure* **2016**, *24*, 1537–1549. [[CrossRef](#)] [[PubMed](#)]
55. Mathieu, C.; Pappu, R.V.; Taylor, J.P. Review Beyond aggregation: Pathological phase transitions in neurodegenerative disease. *Science* **2020**, *370*, 56–60. [[CrossRef](#)] [[PubMed](#)]
56. Martin, E.W.; Holehouse, A.S.; Grace, C.R.; Hughes, A.; Pappu, R.V.; Mittag, T. Sequence Determinants of the Conformational Properties of an Intrinsically Disordered Protein Prior to and upon Multisite Phosphorylation. *J. Am. Chem. Soc.* **2016**, *138*, 15323–15335. [[CrossRef](#)] [[PubMed](#)]
57. Soranno, A.; Koenig, I.; Borgia, M.B.; Hofmann, H.; Zosel, F.; Nettels, D.; Schuler, B. Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 4874–4879. [[CrossRef](#)] [[PubMed](#)]
58. Zosel, F.; Soranno, A.; Buholzer, K.J.; Nettels, D.; Schuler, B. Depletion interactions modulate the binding between disordered proteins in crowded environments. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 13480–13489. [[CrossRef](#)]
59. Mukherjee, S.K.; Gautam, S.; Biswas, S.; Kundu, J.; Chowdhury, P.K. Do Macromolecular Crowding Agents Exert Only an Excluded Volume Effect? A Protein Solvation Study. *J. Phys. Chem. B* **2015**, *119*, 14145–14156. [[CrossRef](#)]
60. Goh, G.K.M.; Dunker, A.K.; Uversky, V.N. Protein intrinsic disorder toolbox for comparative analysis of viral proteins. *BMC Genom.* **2008**, *9*, S4. [[CrossRef](#)]
61. Goh, G.K.M.; Dunker, A.K.; Uversky, V.N. Understanding Viral Transmission Behavior via Protein Intrinsic Disorder Prediction: Coronaviruses. *J. Pathog.* **2012**, *2012*, 738590. [[CrossRef](#)] [[PubMed](#)]
62. Goh, G.K.M.; Dunker, A.K.; Uversky, V.N. Prediction of Intrinsic Disorder in MERS-CoV/HCoV-EMC Supports a High Oral-Fecal Transmission. *PLoS. Curr.* **2013**, *5*. [[CrossRef](#)] [[PubMed](#)]
63. Goh, G.K.M.; Dunker, A.K.; Uversky, V.N. Shell disorder, immune evasion and transmission behaviors among human and animal retroviruses. *Mol. Biosyst.* **2015**, *11*, 2312–2323. [[CrossRef](#)] [[PubMed](#)]
64. Goh, G.K.M.; Dunker, A.K.; Foster, J.A.; Uversky, V.N. Shell disorder analysis predicts greater resilience of the SARS-CoV-2 (COVID-19) outside the body and in body fluids. *Microb. Pathog.* **2020**, *144*, 104177. [[CrossRef](#)] [[PubMed](#)]
65. Goh, G.K.M.; Dunker, A.K.; Foster, J.A.; Uversky, V.N. Rigidity of the Outer Shell Predicted by a Protein Intrinsic Disorder Model Sheds Light on the COVID-19 (Wuhan-2019-nCoV) Infectivity. *Biomolecules* **2020**, *10*, 331. [[CrossRef](#)]
66. Prather, K.A.; Marr, L.C.; Schooley, R.T.; McDiarmid, M.A.; Wilson, M.E.; Milton, D.K. Airborne transmission of SARS-CoV-2. *Science* **2020**, *6514*. [[CrossRef](#)]

67. Goldman, E. Exaggerated risk of transmission of COVID-19 by fomites. *Lancet Infect. Dis.* **2020**, *20*, 892–893. [[CrossRef](#)]
68. Kwong, P.D.; Wyatt, R.; Robinson, J.; Sweet, R.W.; Sodroski, J.; Hendrickson, W.A. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **1998**, *393*, 648–659. [[CrossRef](#)]
69. Giovine, P.D.; Settembre, E.C.; Bhargava, A.K.; Luftig, M.A.; Lou, H.; Cohen, G.H.; Eisenberg, R.J.; Krummenacher, C.; Carfi, A. Structure of Herpes Simplex Virus Glycoprotein D Bound to the Human Receptor Nectin-1. *PLoS Pathog.* **2011**, *7*, e1002277. [[CrossRef](#)]
70. Xiao, C.; Bator, C.M.; Bowman, V.D.; Rieder, E.; He, Y.; Hebert, B.; Bella, J.; Baker, T.S.; Wimmer, E.; Kuhn, R.J.; et al. Interaction of coxsackievirus A21 with its cellular receptor, ICAM-1. *J. Virol.* **2001**, *75*, 2444–2451. [[CrossRef](#)]
71. Grunert, H.P.; Wolf, K.U.; Langner, K.D.; Sawitzky, D.; Habermehl, K.O.; Zeichhardt, H. Internalization of human rhinovirus 14 into HeLa and ICAM-1-transfected BHK cells. *Med. Microbiol. Immunol.* **1997**, *186*, 1–9. [[CrossRef](#)]
72. Jean, J.R.; Jacomy, H.; Desforges, M.; Vabret, A.; Freymuth, F.; Talbot, P.J. Human respiratory coronavirus OC43: Genetic stability and neuroinvasion. *J. Virol.* **2004**, *78*, 8824–8834. [[CrossRef](#)]
73. Mahase, E. Coronavirus: Covid-19 has killed more people than SARS and MERS combined, despite lower case fatality rate. *BMJ* **2020**, *368*, m641. [[CrossRef](#)] [[PubMed](#)]
74. Vijgen, L.; Keyaerts, E.; Moes, E.; Thoelen, I.; Wollants, E.; Lemey, P.; Vandamme, A.M.; Ranst, M.V. Complete genomic sequence of human coronavirus OC43: Molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. *J. Virol.* **2005**, *79*, 1595–1604. [[CrossRef](#)] [[PubMed](#)]
75. Ripberger, T.J.; Uhrlaub, J.L.; Watanabe, M.; Wong, R.; Castaneda, Y.; Pizzato, H.A.; Thompson, M.R.; Bradshaw, C.; Weinkauff, C.C.; Bime, C.; et al. Orthogonal SARS-CoV-2 Serological Assays Enable Surveillance of Low-Prevalence Communities and Reveal Durable Humoral Immunity. *Immunity* **2020**, *53*, 725–733.
76. Wajnberg, A.; Amanat, F.; Firpo, A.; Altman, D.R.; Bailey, M.J.; Mansour, M.; McMahon, M.; Meade, P.; Mendu, D.R.; Muellers, K.; et al. Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. *Science* **2020**. [[CrossRef](#)]
77. Krammer, F. SARS-CoV-2 vaccines in development. *Nature* **2020**, *586*, 516–527. [[CrossRef](#)] [[PubMed](#)]
78. Fauci, A.S.; Lane, H.C. Four Decades of HIV/AIDS—Much Accomplished, Much to Do. *N. Engl. J. Med.* **2020**, *383*, 1–4. [[CrossRef](#)]
79. Liu, Y.; Wang, X.; Liu, B. A comprehensive review and comparison of existing computational methods for intrinsically disordered protein and region prediction. *Brief. Bioinform.* **2019**, *20*, 330–346. [[CrossRef](#)]
80. Prilusky, J.; Felder, C.E.; Zeev-Ben-Mordehai, T.; Rydberg, E.H.; Man, O.; Beckmann, J.S.; Silman, I.; Sussman, J.L. FoldIndex: A simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* **2005**, *21*, 3435–3438. [[CrossRef](#)]
81. Dosztanyi, Z.; Csizmok, V.; Tompa, P.; Simon, I. IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **2005**, *21*, 3433–3434. [[CrossRef](#)] [[PubMed](#)]
82. Dosztanyi, Z.; Csizmok, V.; Tompa, P.; Simon, I. The Pairwise Energy Content Estimated from Amino Acid Composition Discriminates between Folded and Intrinsically Unstructured Proteins. *J. Mol. Biol.* **2005**, *347*, 827–839. [[CrossRef](#)]
83. Dosztanyi, Z. Prediction of protein disorder based on IUPred. *Protein Sci.* **2018**, *27*, 331–340. [[CrossRef](#)] [[PubMed](#)]
84. Linding, R.; Russell, R.B.; Neduva, V.; Gibson, T.J. GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res.* **2003**, *31*, 3701–3708. [[CrossRef](#)] [[PubMed](#)]
85. Campen, A.; Williams, R.M.; Brown, C.J.; Meng, J.W.; Uversky, V.N.; Dunker, A.K. Protein intrinsic disorder and influenza virulence: The 1918 H1N1 and H5N1 viruses. *Protein Peptide Lett.* **2008**, *15*, 956–963. [[CrossRef](#)] [[PubMed](#)]
86. Peng, K.; Vucetic, S.; Radivojac, P.; Brown, C.J.; Dunker, A.K.; Obradovic, Z. Optimizing long intrinsic disorder predictors with protein evolutionary information. *J. Bioinform. Comput. Biol.* **2005**, *3*, 35–60. [[CrossRef](#)] [[PubMed](#)]
87. Peng, K.; Radivojac, P.; Vucetic, S.; Dunker, A.K.; Obradovic, Z. Length-dependent prediction of protein intrinsic disorder. *BMC Bioinform.* **2006**, *7*, 208. [[CrossRef](#)]

88. Ward, J.J.; McGuffin, L.J.; Bryson, K.; Buxton, B.F.; Jones, D.T. The DISOPRED server for the prediction of protein disorder. *Bioinformatics* **2004**, *20*, 2138–2139. [[CrossRef](#)]
89. Linding, R.; Jensen, L.J.; Diella, F.; Bork, P.; Gibson, T.J.; Russell, R.B. Protein disorder prediction: Implications for structural proteomics. *Structure* **2003**, *11*, 1453–1459. [[CrossRef](#)]
90. Xue, B.; Dunbrack, R.L.; Williams, R.W.; Dunker, A.K.; Uversky, V.N. PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. *BBA Proteins Proteom.* **2010**, *1804*, 996–1010. [[CrossRef](#)]
91. Jones, D.T.; Cozzetto, D. DISOPRED3: Precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* **2015**, *31*, 857–863. [[CrossRef](#)] [[PubMed](#)]
92. Sullivan, S.S.; Weinzierl, R.O.J. Optimization of Molecular Dynamics Simulations of c-MYC1-88—An Intrinsically Disordered System. *Life* **2020**, *10*, 109. [[CrossRef](#)]
93. Navarro-Paya, C.; Sanz-Hernandez, M.; de Simone, A. In Silico Study of the Mechanism of Binding of the N-Terminal Region of α Synuclein to Synaptic-Like Membranes. *Life* **2020**, *10*, 98. [[CrossRef](#)] [[PubMed](#)]
94. Sala, D.; Cosentino, U.; Ranaudo, A.; Greco, C.; Moro, G. Dynamical Behavior and Conformational Selection Mechanism of the Intrinsically Disordered Sic1 Kinase-Inhibitor Domain. *Life* **2020**, *10*, 110. [[CrossRef](#)] [[PubMed](#)]
95. Rauscher, S.; Gapsys, V.; Gajda, M.J.; Zweckstetter, M.; de Groot, B.L.; Grubmuller, H. Structural Ensembles of Intrinsically Disordered Proteins Depend Strongly on Force Field: A Comparison to Experiment. *J. Chem. Theory Comput.* **2015**, *11*, 5513–5524. [[CrossRef](#)] [[PubMed](#)]
96. Xiang, S.; Gapsys, V.; Kim, H.Y.; Bessonov, S.; Hsiao, H.H.; Mohlmann, S.; Klaukien, V.; Ficner, R.; Becker, S.; Urlaub, H.; et al. Phosphorylation drives a dynamic switch in serine/arginine-rich proteins. *Structure* **2013**, *21*, 2162–2174. [[CrossRef](#)] [[PubMed](#)]
97. Schramm, A.; Bignon, C.; Brocca, S.; Grandori, R.; Santambrogio, C.; Longhi, S. An arsenal of methods for the experimental characterization of intrinsically disordered proteins—How to choose and combine them? *Arch. Biochem. Biophys.* **2019**, *676*, 108055. [[CrossRef](#)]
98. Xing, Y.; Takemaru, K.; Liu, J.; Berndt, J.D.; Zheng, J.J.; Moon, R.T.; Xu, W. Crystal structure of a full-length beta-catenin. *Structure* **2008**, *16*, 478–487. [[CrossRef](#)]
99. de Genst, E.J.; Guilliams, T.; Wellens, J.; O’Day, E.M.; Waudby, C.A.; Meehan, S.; Dumoulin, M.; Hsu, S.T.; Cremades, N.; Verschuere, K.H.; et al. Structure and properties of a complex of α -synuclein and a single-domain camelid antibody. *J. Mol. Biol.* **2010**, *402*, 326–343. [[CrossRef](#)]
100. Abskharon, R.N.; Giachin, G.; Wohlkonig, A.; Soror, S.H.; Pardon, E.; Legname, G.; Steyaert, J. Probing the N-Terminal β -Sheet Conversion in the Crystal Structure of the Human Prion Protein Bound to a Nanobody. *J. Am. Chem. Soc.* **2014**, *136*, 937–944. [[CrossRef](#)]
101. Rodriguez, J.A.; Ivanova, M.I.; Sawaya, M.R.; Cascio, D.; Reyes, F.E.; Shi, D.; Sangwan, S.; Guenther, E.L.; Johnson, L.M.; Zhang, M.; et al. Structure of the toxic core of α -synuclein from invisible crystals. *Nature* **2015**, *525*, 486–490. [[CrossRef](#)] [[PubMed](#)]
102. Shi, D.; Nannenga, B.L.; de la Cruz, M.J.; Liu, J.; Sawtelle, S.; Calero, G.; Reyes, F.E.; Hattne, J.; Gonen, T. The collection of MicroED data for macromolecular crystallography. *Nat. Protoc.* **2016**, *11*, 895–904. [[CrossRef](#)] [[PubMed](#)]
103. Sawaya, M.R.; Rodriguez, J.; Cascio, D.; Collazo, M.J.; Shi, D.; Reyes, F.E.; Hattne, J.; Gonen, T.; Eisenberg, D.S. Ab initio structure determination from prion nanocrystals at atomic resolution by MicroED. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 11232–11236. [[CrossRef](#)] [[PubMed](#)]
104. de la Cruz, M.J.; Hattne, J.; Shi, D.; Seidler, P.; Rodriguez, J.; Reyes, F.E.; Sawaya, M.R.; Cascio, D.; Weiss, S.C.; Kim, S.K.; et al. Atomic-resolution structures from fragmented protein crystals with the cryoEM method MicroED. *Nat. Methods* **2017**, *14*, 399–402. [[CrossRef](#)] [[PubMed](#)]
105. 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**, *589*, 2570–2577. [[CrossRef](#)] [[PubMed](#)]
106. Konrat, R. NMR contributions to structural dynamics studies of intrinsically disordered proteins. *J. Magn. Reson.* **2014**, *241*, 74–85. [[CrossRef](#)] [[PubMed](#)]
107. Felli, I.C.; Pierattelli, R. Novel methods based on ^{13}C detection to study intrinsically disordered proteins. *J. Magn. Reson.* **2014**, *241*, 115–125. [[CrossRef](#)]
108. Chhabra, S.; Fischer, P.; Takeuchi, K.; Dubey, A.; Ziarek, J.J.; Boeszoermyeni, A.; Mathieu, D.; Bermel, W.; Davey, N.E.; Wagner, G.; et al. ^{15}N detection harnesses the slow relaxation property of nitrogen: Delivering

- enhanced resolution for intrinsically disordered proteins. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E1710–E1719. [[CrossRef](#)]
109. Fusco, G.; de Simone, A.; Gopinath, T.; Vostrikov, V.; Vendruscolo, M.; Dobson, C.M.; Veglia, G. Direct observation of the three regions in α -synuclein that determine its membrane-bound behaviour. *Nat. Commun.* **2014**, *5*, 3827. [[CrossRef](#)]
 110. Fusco, G.; Chen, S.W.; Williamson, P.T.F.; Cascella, R.; Perni, M.; Jarvis, J.A.; Cecchi, C.; Vendruscolo, M.; Chiti, F.; Cremades, N.; et al. Structural basis of membrane disruption and cellular toxicity by α -synuclein oligomers. *Science* **2017**, *358*, 1440–1443. [[CrossRef](#)]
 111. Lautenschlager, J.; Stephens, A.D.; Fusco, G.; Strohl, F.; Curry, N.; Zacharopoulou, M.; Michel, C.H.; Laine, R.; Nespovitya, N.; Fantham, M.; et al. C-terminal calcium binding of α -synuclein modulates synaptic vesicle interaction. *Nat. Commun.* **2018**, *9*, 712. [[CrossRef](#)] [[PubMed](#)]
 112. Osterlund, N.; Moons, R.; Ilag, L.L.; Sobott, F.; Graslund, A. Native Ion Mobility-Mass Spectrometry Reveals the Formation of β -Barrel Shaped Amyloid- β Hexamers in a Membrane-Mimicking Environment. *J. Am. Chem. Soc.* **2019**, *141*, 10440–10450. [[CrossRef](#)] [[PubMed](#)]
 113. Theillet, F.X.; Binolfi, A.; Frembgen-Kesner, T.; Hingorani, K.; Sarkar, M.; Kyne, C.; Li, C.; Crowley, P.B.; Gierasch, L.; Pielak, G.J.; et al. Physicochemical Properties of Cells and Their Effects on Intrinsically Disordered Proteins (IDPs). *Chem. Rev.* **2014**, *114*, 6661–6714. [[CrossRef](#)] [[PubMed](#)]
 114. Leney, A.C.; Heck, A.J. Native Mass Spectrometry: What is in the Name? *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 5–13. [[CrossRef](#)] [[PubMed](#)]
 115. Konermann, L.; Ahadi, E.; Rodriguez, A.D.; Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **2013**, *85*, 2–9. [[CrossRef](#)] [[PubMed](#)]
 116. Kuprowski, M.C.; Konermann, L. Signal response of coexisting protein conformers in electrospray mass spectrometry. *Anal. Chem.* **2007**, *79*, 2499–2506. [[CrossRef](#)]
 117. Frimpong, A.K.; Abzalimov, R.R.; Uversky, V.N.; Kaltashov, I.A. Characterization of intrinsically disordered proteins with electrospray ionization mass spectrometry: Conformational heterogeneity of alpha-synuclein. *Proteins* **2010**, *78*, 714–722. [[CrossRef](#)]
 118. Testa, L.; Brocca, S.; Grandori, R. Charge-Surface Correlation in Electrospray Ionization of Folded and Unfolded Proteins. *Anal. Chem.* **2011**, *83*, 6459–6463. [[CrossRef](#)]
 119. Testa, L.; Brocca, S.; Santambrogio, C.; D’Urzo, A.; Habchi, J.; Longhi, S.; Uversky, V.N.; Grandori, R. Extracting structural information from charge-state distributions of intrinsically disordered proteins by non-denaturing electrospray-ionization mass spectrometry. *Intrinsically Disord. Proteins* **2013**, *1*, e25068. [[CrossRef](#)]
 120. Beveridge, R.; Covill, S.; Pacholarz, K.J.; Kalapothakis, J.M.; MacPhee, C.E.; Barran, P.E. A mass-spectrometry-based framework to define the extent of disorder in proteins. *Anal. Chem.* **2014**, *86*, 10979–10991. [[CrossRef](#)]
 121. Beveridge, R.; Phillips, A.S.; Denbigh, L.; Saleem, H.M.; MacPhee, C.E.; Barran, P.E. Relating gas phase to solution conformations: Lessons from disordered proteins. *Proteomics* **2015**, *15*, 2872–2883. [[CrossRef](#)] [[PubMed](#)]
 122. Santambrogio, C.; Natalello, A.; Brocca, S.; Ponzini, E.; Grandori, R. Conformational Characterization and Classification of Intrinsically Disordered Proteins by Native Mass Spectrometry and Charge-State Distribution Analysis. *Proteomics* **2019**, *19*, e1800060. [[CrossRef](#)] [[PubMed](#)]
 123. Borysik, A.J.; Kovacs, D.; Guharoy, M.; Tompa, P. Ensemble Methods Enable a New Definition for the Solution to Gas-Phase Transfer of Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2015**, *137*, 13807–13817. [[CrossRef](#)] [[PubMed](#)]
 124. Beveridge, R.; Migas, L.G.; Das, R.K.; Pappu, R.V.; Kriwacki, R.W.; Barran, P.E. Ion Mobility Mass Spectrometry Uncovers the Impact of the Patterning of Oppositely Charged Residues on the Conformational Distributions of Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2019**, *141*, 4908–4918. [[CrossRef](#)]
 125. Masson, G.R.; Burke, J.E.; Ahn, N.G.; Anand, G.S.; Borchers, C.; Brier, S.; Bou-Assaf, G.M.; Engen, J.R.; Englander, S.W.; Faber, J.; et al. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat. Methods* **2019**, *16*, 595–602. [[CrossRef](#)]
 126. Hansen, J.C.; Wexler, B.B.; Rogers, D.J.; Hite, K.C.; Panchenko, T.; Ajith, S.; Black, B.E. DNA binding restricts the intrinsic conformational flexibility of methyl CpG binding protein 2 (MeCP2). *J. Biol. Chem.* **2011**, *286*, 18938–18948. [[CrossRef](#)]

127. Chanthamontri, C.; Liu, J.; McLuckey, S.A. Charge State Dependent Fragmentation of Gaseous α -Synuclein Cations via Ion Trap and Beam-Type Collisional Activation. *Int. J. Mass Spectrom.* **2009**, *283*, 9–16. [[CrossRef](#)]
128. Phillips, A.S.; Gomes, A.F.; Kalapothakis, J.M.; Gillam, J.E.; Gasparavicius, J.; Gozzo, F.C.; Kunath, T.; MacPhee, C.; Barran, P.E. Early stages of insulin fibrillogenesis examined with ion mobility mass spectrometry and molecular modelling. *Analyst* **2015**, *140*, 3070–3081. [[CrossRef](#)]
129. Zhou, M.; Lantz, C.; Brown, K.A.; Ge, Y.; Tolic, L.P.; Loo, J.A.; Lermyte, F. Higher-order structural characterisation of native proteins and complexes by top-down mass spectrometry. *Chem. Sci.* **2020**. [[CrossRef](#)]
130. Miraglia, F.; Valvano, V.; Rota, L.; di Primio, C.; Quercioli, V.; Betti, L.; Giannaccini, G.; Cattaneo, A.; Colla, E. Alpha-Synuclein FRET Biosensors Reveal Early Alpha-Synuclein Aggregation in the Endoplasmic Reticulum. *Life* **2020**, *10*, 147. [[CrossRef](#)]
131. Visconti, L.; Malagrino, F.; Pagano, L.; Toto, A. Understanding the Mechanism of Recognition of Gab2 by the N-SH2 Domain of SHP2. *Life* **2020**, *10*, 85. [[CrossRef](#)] [[PubMed](#)]
132. Lapidus, L.J.; Eaton, W.A.; Hofrichter, J. Measuring the rate of intramolecular contact formation in polypeptides. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7220–7225. [[CrossRef](#)] [[PubMed](#)]
133. Acharya, S.; Saha, S.; Ahmad, B.; Lapidus, L.J. Effects of Mutations on the Reconfiguration Rate of α -Synuclein. *J. Phys. Chem. B* **2015**, *119*, 15443–15450. [[CrossRef](#)] [[PubMed](#)]
134. Ahmad, B.; Chen, Y.; Lapidus, L.J. Aggregation of α -synuclein is kinetically controlled by intramolecular diffusion. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 2336–2341. [[CrossRef](#)] [[PubMed](#)]
135. Kulkarni, P.; Uversky, V.N. Intrinsically Disordered Proteins and the Janus Challenge. *Proteomics* **2018**, *18*, 179. [[CrossRef](#)]
136. Singh, A.; Kumar, A.; Yadav, R.; Uversky, V.N.; Giri, R. Deciphering the dark proteome of Chikungunya virus. *Sci. Rep. UK* **2018**, *8*, 5822. [[CrossRef](#)]
137. Giri, R.; Bhardwaj, T.; Shegane, M.; Gehi, B.R.; Kumar, P.; Gadhave, K.; Oldfield, C.J.; Uversky, V.N. Understanding COVID-19 via comparative analysis of dark proteomes of SARS-CoV-2, human SARS and bat SARS-like coronaviruses. *Cell. Mol. Life Sci.* **2020**. [[CrossRef](#)]
138. Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Intrinsically disordered proteins in human diseases: Introducing the D2 concept. *Annu. Rev. Biophys.* **2008**, *37*, 215–246. [[CrossRef](#)] [[PubMed](#)]

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



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).