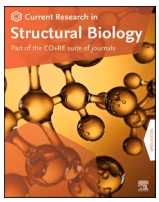
Current Research in Structural Biology 7 (2024) 100138

Contents lists available at ScienceDirect

Current Research in Structural Biology

journal homepage: www.journals.elsevier.com/current-research-in-structural-biology



Experimental methods to study the structure and dynamics of intrinsically disordered regions in proteins

Snigdha Maiti [**,1], Aakanksha Singh, Tanisha Maji, Nikita V. Saibo, Soumya De [*]

School of Bioscience, Indian Institute of Technology Kharagpur, Kharagpur, WB, 721302, India

ARTICLE INFO

Handling Editor: Dr A Wlodawer

Keywords:
NMR spectroscopy
Circular dichroism (CD) spectroscopy
Fourier-transform infrared spectroscopy (FTIR) spectroscopy
Small angle x-ray scattering (SAXS)
Fluorescence spectroscopy
Atomic force microscopy (AFM)

ABSTRACT

Eukaryotic proteins often feature long stretches of amino acids that lack a well-defined three-dimensional structure and are referred to as intrinsically disordered proteins (IDPs) or regions (IDRs). Although these proteins challenge conventional structure-function paradigms, they play vital roles in cellular processes. Recent progress in experimental techniques, such as NMR spectroscopy, single molecule FRET, high speed AFM and SAXS, have provided valuable insights into the biophysical basis of IDP function. This review discusses the advancements made in these techniques particularly for the study of disordered regions in proteins. In NMR spectroscopy new strategies such as ^{13}C detection, non-uniform sampling, segmental isotope labeling, and rapid data acquisition methods address the challenges posed by spectral overcrowding and low stability of IDPs. The importance of various NMR parameters, including chemical shifts, hydrogen exchange rates, and relaxation measurements, to reveal transient secondary structures within IDRs and IDPs are presented. Given the high flexibility of IDPs, the review outlines NMR methods for assessing their dynamics at both fast (ps-ns) and slow (μs-ms) timescales. IDPs exert their functions through interactions with other molecules such as proteins, DNA, or RNA. NMR-based titration experiments yield insights into the thermodynamics and kinetics of these interactions. Detailed study of IDPs requires multiple experimental techniques, and thus, several methods are described for studying disordered proteins, highlighting their respective advantages and limitations. The potential for integrating these complementary techniques, each offering unique perspectives, is explored to achieve a comprehensive understanding of IDPs.

1. Introduction

The field of structural biology commenced its journey with the groundbreaking discovery of the DNA double helix structure (Watson and Crick, 1953), followed by the revelation of the myoglobin protein structure (Kendrew et al., 1958). Initially, Emil Fischer's "lock and key" model for enzyme catalysis built the foundation to understand the link between an enzyme and its function (Fischer, 1894). Subsequently, it became evident that the structure of a protein is determined by its amino acid sequence (Schechter et al., 1970). The relation between amino acid sequence and folding of the protein into its biologically active conformation was studied extensively following denaturation and refolding of several proteins (Mirsky and Pauling, 1936) while later the refolding study and amino acid sequence of the protein ribonuclease ultimately led a Nobel Prize winning work (Sela et al., 1957). The advent of X-ray crystallography transformed our ability to explore protein structure at the atomic level, affirming the core principle that an amino acid sequence governs the three-dimensional shape of a protein and, consequently, its function.

The exceptions to the paradigm of protein structure and function relationship emerged with the discovery of intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs) within proteins. IDPs lack stable three-dimensional structures, and IDRs are extended stretches of amino acids within a protein without any stable secondary or tertiary structure. Proteins display a wide spectrum of structural characteristics. In extreme instances, certain proteins are completely disordered (FlgM, 4 E-BP1, p21, p27, and p57). Conversely, others take on a modular form, showcasing a combination of folded regions and long segments of intrinsically disordered regions (IDRs). Seminal works in the early 2000s highlighted the prevalence and

* Corresponding author.
** Corresponding author.
E-mail addresses: maitisnigdha@gmail.com (S. Maiti), somde@iitkgp.ac.in (S. De).
[1] Present address: Department of Structural Biology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA.

https://doi.org/10.1016/j.crstbi.2024.100138
Received 7 November 2023; Received in revised form 12 March 2024; Accepted 15 March 2024
Available online 21 March 2024

importance of IDPs and IDRs, prompting a shift away from the conventional structure-function paradigm (Wright and Dyson, 1999; Uversky et al., 2000; Dunker et al., 2001; Tompa, 2002).

The absence of folded structures in proteins was recognized as early as the 1950s, but the lack of suitable experimental methods hindered their study for decades. Locally disorder regions were first noticed in X-ray crystallography data represented as regions with missing electron density (Arnone et al., 1971) as they are highly dynamic and fail to diffract X-rays. Over time, a large fraction of the protein crystal structures deposited in the Protein Data Bank (PDB) were reported to have regions of missing electron density (Gall et al., 2007). The crystal structure of tobacco mosaic virus (TMV) coat protein showed a 25-residue long lysine rich loop region with missing electron density, which is one of the earliest reported functionally important IDR (Champness et al., 1976) that undergoes disorder to order transition upon binding to RNA during viral assembly. However, most missing regions in crystal structures are ambiguous and short and a direct conclusion cannot be drawn for a region to be disordered just from missing data in crystal structures. Unambiguous data obtained from NMR spectroscopy and other techniques, such as circular dichroism (CD) spectroscopy, initiated the development of a manually curated IDP database DisProt, which currently holds 2649 proteins annotated as disordered (Quaglia et al., 2022).

In recent years, a substantial body of well-crafted reviews and textbooks has been published, highlighting several key advancements in experimental techniques for studying IDPs (Dyson and Wright, 2005; Mittag and Forman-Kay, 2007; Dunker et al., 2008; Dunker et al., 2008b; Uversky et al., 2008; Wright and Dyson, 2009, 2015; Metallo, 2010; Turoverov et al., 2010; Uversky and Dunker, 2010; Babu et al., 2011; Uversky, 2011, 2013, 2019; Tompa, 2012; Jensen et al., 2013; Kosol et al., 2013; Jensen et al., 2014; Lee et al., 2014; Theillet et al., 2014; Babu, 2016; Skriver et al., 2023). These works illuminate emerging trends and pivotal insights that have significantly influenced the field, making them invaluable resources for researchers and professionals in protein structural biology and dynamics.

1.1. Abundance of disordered proteins

An intrinsically disordered protein (IDP) refers to a polypeptide chain that lacks a well-defined three-dimensional structure on its own. On the other hand, intrinsically disordered regions (IDRs) represent segments or regions within a protein that do not fold into a specific three-dimensional structure. Studies have revealed that long disordered regions (>30 residues) are present across all three domains of life, with archaea, bacteria, and eukaryotes exhibiting percentages of 2%, 4%, and 33% respectively (Ward et al., 2004). More than one-tenth of eukaryotic proteins are completely disordered (Bogatyreva et al., 2006; Pancsa and Tompa, 2012). Computational and bioinformatic analyses have emphasized the prevalence of IDPs and IDRs in the proteome, particularly in complex eukaryotic organisms, which possess a higher proportion of disordered regions compared to archaea and bacteria. This suggests that IDRs may be involved in more complex signaling and regulatory pathways (Dunker et al., 2000, 2001, 2002; Iakoucheva et al., 2002; Pancsa and Tompa, 2012). It is estimated that 44% of human proteins contain long IDRs (Oates et al., 2013). These findings highlight the importance of IDPs and IDRs in understanding the complexity of biological systems.

1.2. Functional significance of IDPs

Over the past few years, significant discoveries have demonstrated the crucial biological functions of intrinsically disordered regions and proteins (Wright and Dyson, 2015). The disordered regions within proteins are commonly associated with various signaling and regulatory processes, including differentiation, transcriptional and translational regulation, RNA processing, cell cycle control, storage of small molecules, and the regulation of large multiprotein complex assembly (Dyson and Wright, 2005; Xie et al., 2007; Uversky et al., 2008). Given their vital involvement in numerous cellular pathways, malfunctions of these proteins are frequently associated with various human diseases such as developmental disorders, neurodegenerative diseases, cardiovascular diseases, type II diabetes, and cancer (Iakoucheva et al., 2002; Dyson and Wright, 2005; Xie et al., 2007; Galea et al., 2008; Uversky et al., 2008; Uversky, 2013; Knowles et al., 2014).

1.3. IDPs are involved in various diseases

IDPs have been implicated in various aspects of cancer development and progression. Additionally, IDPs are involved in signaling pathways related to cell proliferation, invasion, and metastasis. For instance, the tumor suppressor protein p53 plays a crucial role in regulating cell cycle arrest and apoptosis. p53 relies extensively on its disordered regions to facilitate and regulate interactions with other proteins (Oldfield et al., 2008). Breast cancer type 1 susceptibility protein (BRCA1) has been implicated in a variety of different cancers (Venkitaraman, 2002). Over 50 proteins such as p53, BRCA2, the oncogenes c-Myc, JunB, Rad50 and Rad51 which encompass a range of DNA damage sensors, DNA repair proteins, and signal transducers, interact with BRCA1. Most of these interactions take place within the IDR region of BRCA1. Neurodegenerative diseases like Alzheimer's, Parkinson's, and Huntington's are associated with the aggregation of specific IDPs. For example, the amyloid-beta protein in Alzheimer's and the alpha-synuclein protein in Parkinson's are intrinsically disordered and form toxic aggregates (Giasson et al., 2001; Klimov and Thirumalai, 2003; Shigemitsu and Hiroaki, 2018). Other diseases, where intrinsic disorder in protein components was reported are family of polyQ diseases (Zoghbi and Orr, 2000), prion disease (Prusiner, 2001), multiple system atrophy (Dev, 2003).

2. Techniques to study IDPs and IDRs

The biophysical and biochemical techniques that are used to study folded proteins are also being used to study IDPs. However, several innovations have been made in these techniques to account for the fundamental differences between folded proteins and IDPs. Folded proteins have a unique native state and have a hierarchy of structures: secondary structures, loops, super-secondary structures, tertiary structure and for some proteins quaternary structure. This structural hierarchy has guided the structural and dynamic studies of folded proteins. On the other hand, IDPs lack such structural hierarchy and instead display a continuum of ensemble of conformations ranging from random coil like extended structures to compact globular and molten globule structures, that are characterized by varying degrees of transient or stable secondary structures (Uversky, 2002; Dyson and Wright, 2005; Lee et al., 2014). Studies of IDPs in the past few decades have identified several important biophysical properties, which can be grouped as global properties and local properties. Here we provide a list of these properties and the experimental techniques that can be used to probe these properties of IDPs. In the next sections, we discuss these experimental techniques and provide examples of IDPs studied by them.

Global properties of IDPs provide insights into the overall shape and structure of the protein.

1. Compactness: The hydrodynamic volume of proteins serves as a crucial indicator of their shape, whether they are compact, partially swollen or completely extended. Notably, native and unfolded conformations of globular proteins exhibit distinct molecular mass dependencies of their hydrodynamic radii, known as the Stokes radius. Experiments that can be used to measure the compactness of an IDP are limited proteolysis, size exclusion chromatography (SEC), mass spectrometry, NMR spectroscopy, small angle x-ray scattering (SAXS), small angle neutron scattering (SANS), fluorescence based

experiments such as fluorescence correlation spectroscopy (FCS), and fluorescence (or Förster) resonance energy transfer (FRET).
2. Globularity: This parameter reflects the degree of globularity, indicating the presence or absence of a tightly packed core within the protein molecule.
   Hydrogen-exchange experiments using NMR spectroscopy or mass spectrometry can be used to measure globularity of a protein.
3. Conformational Ensemble: IDRs exist as dynamic ensembles of interconverting conformations rather than adopting a single well-defined structure. The composition of this ensemble is also influenced by factors such as temperature, pH, and the presence of binding partners.
   Experiments such as circular dichroism (CD) spectroscopy, fluorescence based experiments and NMR spectroscopy can be used to probe the conformations of an IDP.

Local properties of IDPs provide information at the residue level of the protein.

1. Secondary structure content: Unlike folded proteins, IDRs lack stable secondary structures like α-helices or β-sheets. However, they may transiently adopt elements of secondary structure, albeit with lower stability and shorter lifetimes.
   Experiments such as CD, FTIR, NMR spectroscopy and mass spectrometry can probe the presence of transient secondary structures.
2. Degree of flexibility: Different regions within disordered proteins may exhibit varying degrees of rigidity and flexibility, resulting in conformational heterogeneity.
   Experiments such as limited proteolysis, mass spectrometry and NMR spectroscopy can probe flexibility of an IDP.
3. Local rigid segments without a secondary structure: Despite lacking a stable structure, IDRs often contain short linear motifs or regions with transient secondary structures, such as helices or coils, which play roles in molecular recognition and interaction specificity.

Local rigid segments can be determined by mass spectrometry and NMR spectroscopy.

It is important to note that IDPs, unlike folded proteins, cannot be described by a single structure. The structural and dynamic characteristics of an IDP can only be described by an ensemble of structures, which cannot be determined by any single experiment. Thus, IDPs can be best studied using a combination of complementary techniques. In the following sections we have described several experimental techniques that are commonly used to study IDPs.

2.1. Studies of disordered proteins by NMR spectroscopy

Nuclear Magnetic Resonance (NMR) spectroscopy presents a distinctive approach for investigating protein structure and dynamics in solution. This technique has proven to be exceptionally advantageous in the study of IDPs, as it allows for direct analysis at the atomic level (Konrat, 2014; Dyson and Wright, 2021; Saibo et al., 2022). Over the past few years, there has been a significant surge in experimental research focused on IDRs. This can be largely attributed to the continuous advancements in NMR technology, including the introduction of ultra-high field magnets, the application of multidimensional NMR spectroscopy, and the use of direct carbon detection methods (Cook et al., 2018; Sahu et al., 2014; Zhang et al., 1994; Brutscher et al., 2015; Jensen et al., 2009, 2014).

2.1.1. Backbone chemical shift assignment by standard methods

In general, multidimensional heteronuclear (^{1}H, ^{15}N, and ^{13}C) spectra are acquired for proteins. Collecting a 2D- ^{1}H–^{15}N HSQC spectrum, which is undoubtedly one of the most versatile NMR experiments, is frequently the first step in NMR study for a protein. This 2D experiment generates a fingerprint spectrum for the protein, where all the backbone and side chain N–H pairs show up as cross peaks. A fully structured protein generates a well resolved 2D spectrum. However, disordered regions lack unique structure, which reduces the chemical shift dispersion, especially in the proton dimension and the spectral resolution suffers greatly. In the absence of structure, long IDRs generate a crowded spectrum, where the ^{15}N–^{1}H^{N} cross peaks show up in the 7.5–8.5 ppm range in the ^{1}H-dimension. This spectral crowding along with sharp peaks is the first hint that the protein has a long disordered region. Assignments of the ^{1}H–^{15}N HSQC spectra are done by multiple 3D experiments, using a^{13}C/^{15}N double labeled protein (Sattler, 1999). For proteins with disordered regions as long as ~50–60 residues, attached with one or more folded domains, this sequential backbone assignment method is a standard way to proceed. However, repetitive sequences, such as poly glutamine, arginine-glycine-glycine, glycine--serine, or poly prolines, which are quite prevalent in eukaryotic IDRs make the standard assignment more complicated. Compared to ^{1}H^{N} chemical shift, the dispersion range of ^{15}N atom is notably better, therefore the HNN experiment, which exploits the dispersion range of ^{15}N dimension, is effective in backbone assignment of IDRs (Bhavesh et al., 2001; Panchal et al., 2001; Chakraborty et al., 2012).

2.1.2. Backbone assignment using 13C detection

In the case of very long IDRs, ^{15}N-HSQC experiments can be of limited use due to signal overlap as well as drop in peak intensity due to fast exchange of the exposed backbone amide with bulk water. The highly abundant proline residues in IDRs are also not detectable in ^{15}N-HSQC due to the lack of the amide proton. These issues can be eliminated by using alternative experiments which use direct detection of ^{13}C atoms in place of ^{1}H and expand the possibility to study IDRs by NMR (Bertini et al., 2011; Nováček et al., 2011; Bermel et al., 2012; Felli and Pierattelli, 2014; Gibbs and Kriwacki, 2018). The 2D^{13}C–^{15}N CON and ^{13}C–^{15}N CAN experiments are very helpful in these cases, which show amide nitrogen – carbonyl carbon and amide nitrogen – alpha carbon cross-peaks, respectively (Goradia et al., 2015; Felli et al., 2021). Sequential assignments can be done using different 3D experiments such as (HACA)N(CA)CON and (HACA)N(CA)NCO, which are highly effective for repetitive sequence assignments (Bastidas et al., 2015), along with the 3D experiments C_CCCON (Bermel et al., 2006) and 3D C_H(CC)CON (O'Hare et al., 2009) for side chain assignments.

2.1.3. Fast data acquisition for disordered proteins

Extensive improvements in recent times have been made to reduce experimental time and increase the resolution of the NMR experiments, which helps in studying aggregation-prone and degradation-prone IDRs. Non-uniform sampling (NUS) and band selective excitation short transient (BEST) techniques are the two methods that have been successfully used to reduce experimental time while studying different proteins (Schanda et al., 2006; Waudby and Christodoulou, 2012). In traditional NMR, signals are acquired discreetly in incremental steps for each dimension. However, in NUS, a subset of data points is deliberately not collected. NUS involves sparsely collecting data in the indirect dimension, either by skipping or randomly selecting points in the FID, in order to conserve time. Subsequently, during the processing stage, the missing data is reconstructed. Mathematical algorithms like MDD (Orekhov and Jaravine, 2011), IST (Hyberts et al., 2012), SMILE (Ying et al., 2017), and deep neural networks (Hansen, 2019) are used to reconstruct the missing data points. BEST is a technique that reduces NMR experiment duration by reducing the time interval between scans. BEST selectively excites amide protons without perturbing aliphatic protons, allowing for shorter recycle delays and faster spin polarization restoration (Schanda et al., 2006). Recently, for the assignment of backbone residues of a disordered protein α-synuclein, NUS-BEST combined pulses have been used to collect 3D experiments like HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO which has been shown to greatly reduce the NMR data acquisition time (Rao Kakita et al., 2018).

2.1.4. Segmental isotope labeling to reduce spectral crowding

Segmental isotope labeling enables the targeted investigation of a particular region within a protein. In eukaryotic proteins, roughly one-third feature extensive intrinsically disordered regions (IDRs) alongside one or more folded domains (Bogatyreva et al., 2006). To comprehensively grasp the interplay between the disordered region and the folded domain and elucidate the complete structure-function relationship, it is imperative to analyze these regions within the framework of the entire protein (Tompa, 2012).

However, spectral crowding due to signals from the folded domain of the protein can be challenging. To address this issue, methods such as expressed protein ligation (EPL), protein trans-splicing (PTS), or sortase-mediated ligation (SML) are used to eliminate signals originating from the folded portion of the protein, enabling better analysis of the disordered region.

In the EPL and PTS methods, the disordered region fused with intein enzyme is expressed with ^{13}C and ^{15}N labeling, while the folded part is expressed without isotope enrichment (Flavell and Muir, 2009). The purified proteins are combined to form a functional intein enzyme, which facilitates a splicing reaction, joining the disordered and folded segments of the target protein while excising the intein itself.

In the SML method, the enzyme sortase is used to form a peptide bond between the LPXTG motif located at the C-terminus of the disordered region and the poly-glycine bridge present in the cell wall of gram-positive bacteria (Mao et al., 2004). Here the disordered region is expressed with ^{15}N and ^{13}C labeling, while the folded domain is expressed without isotope enrichment. Once purified, these proteins are linked together through the action of the sortase enzyme. In all three methods (EPL, PTS, and SML), it is possible to attach the disordered region to the C-terminus of a folded domain by interchanging the constructs.

Each method has its advantages and disadvantages. EPL is not recommended when dealing with folded domains that contain disulfide linkages. PTS can encounter challenges related to expression and solubility, depending on the specific target sequences attached to the split inteins. In the SML method, a non-native sequence (LTXTGGG) is introduced between the disordered and folded regions as part of the ligation process.

2.1.5. Structural information from chemical shifts

Although intrinsically disordered regions do not form structured segments like the classical secondary and super-secondary structures observed in folded proteins, the disordered regions contain multiple segments that drive the IDR function. These segments are described as short linear motifs (SLiMs) and molecular recognition features (MoRFs) based on their propensity to form structure upon binding and the length of the binding site (Hsu et al., 2013). SLiMs and MoRFs are the functional elements of IDRs, which serve as interaction sites with other co-factors (small molecules, DNA, RNA, or proteins). These elements have distinct structural and dynamical features, which aid in the function of IDRs. Detailed structural and dynamic characterization of these elements help elucidate their roles in the functions of disordered proteins.

Local chemical environments significantly influence chemical shifts in proteins, especially of backbone atoms ($^{1}H^{N}$, ^{15}N, $^{13}C'$, $^{1}H^{\alpha}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$), which are sensitive to the φ and ψ torsion angles (Cornilescu et al., 1999; Wishart and Sykes, 1994). Hydrogen bonding and electrostatic interactions further affect $^{1}H^{N}$ and ^{15}N chemical shifts (Asakura et al., 1995). Secondary chemical shifts (Δδ), i.e., deviations from random coil chemical shifts, provide information on the protein backbone structure. Random coil shifts are derived from short peptide experiments (Kjaergaard and Poulsen, 2011; Schwarzinger et al., 2000) or data in the Biological Magnetic Resonance Data Bank (De Simone et al., 2009; Tamiola et al., 2010). However, chemical shifts are also sensitive to neighboring residues (De Simone et al., 2009; Tamiola et al., 2010), pH, and temperature (Jensen et al., 2013; Kjaergaard et al., 2010; Kjaergaard and Poulsen, 2011), necessitating careful selection of a random coil chemical shift library for analysis.

Intrinsically disordered regions (IDRs) exhibit multiple conformational states with population-weighted average backbone chemical shifts. Various methods are employed to identify distinct conformations by analyzing these chemical shifts. The Neighbor-Corrected Structural Propensity Calculator (ncSPC) algorithm utilizes a random coil repository, considering neighboring residues and experimental chemical shifts to estimate α-helix and β-extended conformation populations (Tamiola et al., 2010). Alternatively, the RefDB database (H. Zhang et al., 2003) offers amino acid-specific chemical shifts for random coil, α-helical, and β-extended structures, allowing manual population calculation. Polyproline II conformations, common in IDRs, can be predicted using the δ2D method, an extension of ncSPC (Camilloni et al., 2012). Consequently, backbone chemical shifts facilitate the determination of transient structural tendencies in IDRs.

Localized secondary structure and transient long-range intramolecular interactions in IDRs can also be detected using residual dipolar coupling (RDC) (Clore et al., 1998; Jensen et al., 2009; Salmon et al., 2010; Tjandra and Bax, 1997). Another widely used technique for detecting long range contacts is paramagnetic relaxation enhancement (PRE) (Clore et al., 2007; Gillespie and Shortle, 1997). This technique is also useful to detect preformed binding sites that are sparsely populated and short-lived (Fisher and Stultz, 2011).

2.1.6. Residual dipolar coupling (RDC)

Residual dipolar coupling (RDC) is a useful tool for studying conformational sampling of disordered regions in proteins. In a magnetic field, nuclei in close proximity influence the relaxation of each other via dipolar coupling, which depends on the angle between the internuclear vector and the magnetic field. Normally, the isotropic tumbling of molecules in solution leads to an average dipolar coupling of zero. However, weakly aligning media like lipid bicelles, filamentous phages, or stretched polyacrylamide gels can introduce bias in the tumbling, resulting in measurable RDC (Tjandra and Bax, 1997; Clore et al., 1998).

In IDPs and IDRs, transiently formed secondary structural elements tend to orient themselves in alignment with the magnetic field. This orientation results in positive RDC values for α-helical conformations and negative RDC values for β-extended conformations. Additionally, RDC can provide insights into transient long-range order in disordered regions, going beyond local structural propensities (Jensen et al., 2009; Salmon et al., 2010). RDC has been successfully employed in conjunction with SAXS to generate ensemble structures of the unfolded domain of phosphoprotein P from Sendai virus (SeV) (Bernadó et al., 2005) and intrinsically disordered N-terminal transactivation domain (TAD) of tumor suppressor p53 (Wells et al., 2008).

Initially, RDC-based studies of disordered regions faced a major drawback due to their dynamic nature and change shape over time. This posed challenges in obtaining accurate structural information using traditional RDC analysis. The drawback has been overcome by recent algorithms that compute individual alignment tensors from coordinates at each simulation step (Esteban-Martín et al., 2010). This advancement allows for more accurate ensemble structures of IDPs using RDCs.

2.1.7. Paramagnetic relaxation enhancement (PRE)

Disordered regions often exhibit transient long-distance interactions. While nuclear Overhauser effect (NOE) is ideal for studying close-range (<6 Å) interactions between protons, transient contacts between more distant regions can be obtained from paramagnetic relaxation enhancement (PRE) measurements (Gillespie and Shortle, 1997; Bertoncini et al., 2005; Clore et al., 2007). PRE involves attaching a paramagnetic probe to a cysteine sidechain of the protein, which increases the relaxation rates of proximal nuclei (Kawale and Burmann, 2023; Vinogradova and Qin, 2011). Nitroxide stable labels or metal chelators like EDTA-Mn2+ with unpaired electrons are commonly used for the paramagnetic atom.

PRE generates site-specific enhancement of the longitudinal (R_1) and transverse (R_2) relaxation rates in a distance-dependent manner ($1/r^6$) and provides long-range distance restraints (25–35 Å) due to the large magnetic moment of the electron in the probe (Clore and Iwahara, 2009; Kawale and Burmann, 2023; Vinogradova and Qin, 2011). It offers insights into dynamic motions at microsecond timescales and is sensitive to rapidly exchanging conformational states (Kawale and Burmann, 2023). PRE is also useful for studying low population intermediates in macromolecular interactions, especially those involving disordered regions (Vinogradova and Qin, 2011).

Despite being an informative NMR tool, PRE has limitations, including the need for a covalently attached tag and the inherent rotation or flexibility of the paramagnetic atom, leading to time-average distances over all possible conformations (Su and Otting, 2010; Vinogradova and Qin, 2011).

2.1.8. Determination of structural ensemble of disordered regions

As IDRs and IDPs are highly dynamic and sample a wide variety of conformations, a single three-dimensional structure of the protein is misleading. An ensemble of structures capturing all the major conformations is more appropriate (Kosol et al., 2013; Salvi et al., 2019). Computationally generated conformations are screened by experimental data to generate an ensemble of IDR conformations (Jensen et al., 2014). More experimental data results in a better ensemble description of IDRs. NMR parameters such as chemical shifts, hydrogen exchange rates, ^{15}N-relaxation measurements, J-coupling, and residual dipolar couplings (RDC) have been used to determine the existence of local transient secondary structures, while paramagnetic relaxation (PRE) enhancement has been used to study transient long-range contacts. Many such IDR ensemble structures are deposited in the Protein Ensemble Database (Lazar et al., 2021). However, it is important to note that the experimentally determined restraints are much less compared to the number of conformations sampled by an IDR. Thus, the determination of an ensemble of IDR structures is an underdetermined problem.

Collecting multiple complementary experimental data can help with better ensemble structures. Along with NMR spectroscopy, various other biophysical techniques such as CD spectroscopy, Fourier transform infrared (FTIR) spectroscopy, analytical size exclusion chromatography (SEC), small-angle X-ray scattering (SAXS), atomic force microscopy (AFM) and fluorescence resonance energy transfer (FRET) can also be employed. Analytical SEC can be used to obtain the hydrodynamic properties of the protein. SAXS data provide the radius of gyration of the protein and can detect the presence of higher-order oligomeric states. FRET is used to assess the compactness of the protein and detect long-range interactions.

2.1.9. Fast timescale (ps-ns) dynamics of disordered proteins

An inherent advantage of NMR spectroscopy lies in its ability to investigate the dynamic behavior of proteins across various timescales at atomic-level resolution. Intrinsically disordered sequences, in particular, exhibit a wide range of dynamics. These include rapid librational motions on the picosecond timescale, torsion angle dynamics as well as segmental and chain-like motions occurring on the nanosecond timescale, binding interactions in microsecond to millisecond timescale, and in some cases folding upon binding, which can range from millisecond to much slower timescales (Khan et al., 2015; Abyzov et al., 2016; Narayanan et al., 2017; Salvi et al., 2017; Maiti et al., 2019). NMR spectroscopy offers a distinctive method for investigating the dynamic properties that occur across a wide range of time scales, spanning from picoseconds to days. (Palmer, 2001; Kempf and Loria, 2002; Sekhar and Kay, 2019).

Backbone ^{15}N spin relaxation parameters are sensitive to the local and global motions of proteins in the fast ps - ns timescale. These relaxation processes are caused by the fluctuations in the local magnetic field due to the random tumbling motions of a molecule. The fluctuations can be expressed by time dependent rotational correlation function C($\boldsymbol{\tau}$)

$$C(\tau) = \frac{1}{5} e^{-|\tau|/\tau_c}$$

where $\boldsymbol{\tau}_C$ is the correlation time. C($\boldsymbol{\tau}$) decays exponentially within a few nanoseconds. For flexible disordered regions and small proteins, it decays faster, while it decays slower for larger proteins. Fourier transformation of correlation function C($\boldsymbol{\tau}$) provides spectral density function J (ω), which quantitates the amount of motion present in the protein ensemble at a given frequency ω.

$$J(\omega) = \frac{2}{5} \frac{S^2 \tau_C}{1 + \omega^2 \tau_C^2}$$

For proteins the most widely studied spin relaxation processes are the backbone amide relaxation where ^{15}N longitudinal relaxation rate (R_1), the ^{15}N transverse relaxation rate (R_2), and the cross-correlation relaxation heteronuclear nuclear Overhauser effect ({^{1}H}-^{15}N-NOE) are determined. All these parameters are calculated using 2D^{15}N-HSQC detected experiments by modulating time delay while a particular relaxation is ongoing. These relaxation parameters are used to provide a quantitative measure of the amplitude and time scales of the motion of the amide nitrogen for each residue. The R_1, R_2, and {^{1}H}-^{15}N-NOE are typically sensitive to fast time scale dynamics (pico-nano second motion) although R_2 is also sensitive to slower exchange (micro-millisecond). Direct analysis of relaxation data can readily provide local and global conformational properties of a protein. The theoretical {^{1}H}-^{15}N-NOE values range from 0.8 (highly rigid residue) to as low as −4 (highly flexible residue) and offer quick identification of rigid regions and highly flexible disorder regions in a protein. One of the initial reports highlighting the presence of a long disordered region in Bcl-xL protein was studied using NMR relaxation studies specifically from negative {^{1}H}-^{15}N-NOE values (Muchmore et al., 1996).

These three parameters, measured for each backbone amide, depend on the spectral density functions J (ω) at five frequencies 0, ω_N, ω_H, ω_H - ω_N and $\omega_H + \omega_N$, where ω_N and ω_H are the Larmor frequencies of ^{15}N and ^{1}H nuclei, respectively. As ω_H is 10-fold larger than ω_N, the spectral density functions at the higher frequencies i.e. J (ω_H - ω_N), J (ω_H) and J ($\omega_H + \omega_N$) can be merged to a single function of J (0.87ωH) (Farrow et al., 1995). This analysis of relaxation data is known as reduced spectral density analysis. The slower motion present in the molecule contributes to the J (0) function whereas faster motions in the molecule contribute to J (0.87ω_H) function.

Another approach is to use the ^{15}N relaxation data for Lipari-Szabo model free formalism to analyze molecular tumbling which gives global correlation time (τ_m) and order parameter (S^2) along with another correlation time (τ_e) related to internal motion, for each residue (Lipari and Szabo, 1982b; Lipari and Szabo, 1982). This analysis is highly popular for the dynamic analysis of folded globular protein. However, this approach starts with an assumption that internal motion (<50ps) of the bond vector (N–H) is independent of the global correlation time (few ns) and can be separated from each other. This assumption does not hold true for highly disordered proteins as the global and the internal motions are coupled for IDRs. For folded protein, the rotational correlation time can be obtained from the ratio of R_2/R_1. However, for IDRs transient structural propensity of residues can be obtained where residues with secondary structural propensity show elevated R_2/R_1 and disorder residues with no secondary structural propensity show low R_2/R_1 value along with low to negative {^{1}H}-^{15}N-NOE value. Another model for describing relaxation data of IDRs is the segmental motion model (Farrow et al., 1997) that replaces global over segmental correlation time that varies through the polypeptide chain and provides persistence length of the segmental chain motion. Torsion angles (φ and ψ) of amino acids in a disordered region are sensitive to its neighbors. This makes a polypeptide chain not as flexible as a freely linked chain, and the amino acid composition influences its flexibility. Determination

of persistence length is one way to evaluate the global flexibility of a disordered region; the longer the persistence length, the stiffer the polypeptide chain. Notably, Schwalbe et al. have demonstrated that the transverse relaxation rate (R_2) can be fitted to an equation to determine the persistence length of a polypeptide chain.

$$R_2(i) = R_{int} \sum_{J=i}^{N} e^{\frac{|i-j|}{\lambda_0}}$$

where Rint is the intrinsic relaxation rate, which depends on the temperature and viscosity of the solution, λ0 is the persistence length of the polypeptide chain (in numbers of residues) and N is the total chain length of the polypeptide. Generally, for unfolded disordered proteins λ0 is found to be 7 (Farrow et al., 1997; Wirmer et al., 2006).

2.1.10. Identification of short linear motifs (SLiMs)

Disordered regions in proteins possess small linear motifs, which play a crucial role in facilitating their interactions with cofactors. The eukaryotic linear motif database provides a listing of approximately 4138 short linear motifs (SLiMs), and it is believed that there could be around 10^5 such motifs within the eukaryotic proteome (Tompa et al., 2014). However, identifying these motifs in IDRs using bioinformatics methods is challenging due to the high mutation rate and extremely poor sequence conservation observed in IDRs. Additionally, the lack of knowledge about binding partners and the weak interactions between IDRs and their partner molecules further complicate the experimental identification of these linear motifs.

In certain cases, such as the Drosophila transcription factors Sex combs reduced (SCR) and deformed (DFD), it has been demonstrated that these functional motifs exhibit dynamic rigidity and can be identified from the ^{15}N backbone relaxation data. The flexibility of individual residues in SCR and DFD was determined using R_1, R_2, and heteronuclear {^1H}-^{15}N-NOE measurements. Through reduced spectral density analysis, varying degrees of flexibility were revealed in the disordered regions of these proteins (Maiti et al., 2019).

To distinguish between rigid and flexible segments within IDRs, the relative contribution of lower frequency (J (0) and J (ω_N)) as well as higher frequency (J ($0.87\omega_H$)) components was examined. In rigid segments, the spectral density function was primarily influenced by the lower frequency components, while in flexible segments, the higher frequency component also had a considerable impact. The parameter J (0)*J (ω_N)/J ($0.87\omega_H$) was proposed as a sensitive indicator to identify and characterize rigid and flexible segments within an IDR. This parameter helps in understanding the dynamic behavior of IDRs and assists in further elucidating their functional implications.

Furthermore, the product J (0)*J (ω_N)/J ($0.87\omega_H$) was substituted with the equivalent product R_1R_2/(1-NOE) to determine residue-wise rigidity. Residues with significantly higher R_1R_2/(1-NOE) values than the disordered region average were identified as rigid segments. This simple method successfully revealed functionally important rigid segments in the disordered regions of SCR and DFD, which interact with the partner transcription factor extradenticle. Another group studying sequence-dependent backbone dynamics of IDPs, performed long MD simulations on these HOX factors SCR and DFD, revealing that the identified SLiM is anchored by local π–π stacking between Tyr and Trp residue and can also occasionally form long range hydrogen bonds or cation-π bonds with the structured DNA binding domain (Dey et al., 2022). These interactions explain the local enhancement in R_2 rate in the identified rigid segment.

To investigate the sensitivity of the developed R_1R_2/(1-NOE) parameter to measure residue-wise rigidity, a series of mutations in the linker region had been designed to modify the flexibility of the flexible linkers in the IDR of HOX transcription factor SCR (Maiti and De, 2021). The R_1R_2/(1-NOE) parameter is compared with order parameter (S^2) for the wildtype and the SCR mutants to establish any correlation between these two methods and to check the robustness of each parameter in detecting variation of flexibility in IDR of SCR. The R_1R_2/(1-NOE) is correlated to order parameter (S^2) by an exponential growth function, suggesting that for IDRs the parameter R_1R_2/(1-NOE) can be utilized as a measure of residue-wise rigidity, analogous to the order parameter (S^2) for folded protein. While the widely recognized order parameter estimation is effective in detecting rigid segments, residue-wise rigidity R_1R_2/(1-NOE) proves to be a much more straightforward approach with a superior dynamic range to identify prospective SLiMs.

2.1.11. Slow timescale (μs-ms) dynamics of disordered proteins

Intrinsically disordered regions encompass various functional motifs, including molecular recognition features (MoRFs). These elements typically consist of 10–70 amino acid residues and undergo structural transformation upon binding. To monitor such changes, NMR titration-based assays are employed, where the unlabeled partners (proteins or nucleic acids) are titrated against the ^{15}N-labeled IDR. At each titration point, 2D^{15}N HSQC or HMQC spectra are obtained, enabling the identification of interacting residues through chemical shift perturbation (Algamal et al., 2017).

Numerous cases of binding-induced folding of IDRs have been reported. For instance, the disordered region of p21 undergoes folding upon binding to cyclin-dependent kinase (CDK) (Kriwacki et al., 1996). Likewise, the phosphorylated KID domain of CREB protein folds upon binding to the KIX domain of CBP (Radhakrishnan et al., 1997), and the N-terminal disordered transactivation domain of p53 forms a helix after binding to Mdm2 (Chi et al., 2005).

These binding-induced folding events within IDRs occur on a slower timescale, ranging from microseconds to milliseconds. To study such phenomena, various experiments like relaxation dispersion (Palmer, 2001), CEST (Vallurupalli et al., 2012), and lineshape analysis (De et al., 2022) are conducted. Recently, chemical exchange saturation transfer (CEST) and dark-state exchange saturation transfer (DEST) techniques have been developed, proving valuable for studying sparsely populated conformational states and transiently formed complexes with conformational exchanges on the millisecond timescale (Ahmed et al., 2021).

The binding of IDRs to partner molecules can involve multiple steps. For example, the disordered pKID domain forms a complex with KIX through an intermediate (Sugase et al., 2007). Analyzing multiple site exchange of relaxation dispersion data has led to the development of a general pathway (Schneider et al., 2015). Moreover, lineshape analysis of titration data can accommodate more complex binding schemes, providing both thermodynamic and kinetic parameters of complex formation. For instance, the six-state binding and catalysis model involving the disordered cytoplasmic tail of the amyloid precursor protein (APP) and the enzyme Pin1 were analyzed using lineshape analysis (De et al., 2012; Greenwood et al., 2011).

2.1.12. Hydrogen exchange of disordered proteins

The Hydrogen/Deuterium (H/D) exchange technique, in combination with two-dimensional (2D) NMR spectroscopy, is highly effective for studying residual structures in unfolded proteins. This method tracks the NMR signals of peptide amide (NH) groups using ^1H–^{15}N heteronuclear single quantum coherence (HSQC) spectra in ^{15}N-labeled proteins (Boral et al., 2020). Intrinsically disordered proteins (IDPs) are known for their dynamic nature, and HDX-NMR allows for the examination of hydrogen exchange rates across different protein regions, providing insights into IDP dynamics. Despite their lack of stable structures, IDPs may contain transient secondary structures such as alpha-helices or beta-sheets, which HDX-NMR can identify.

HDX-NMR is also useful in studying IDP interactions with other biomolecules like proteins, nucleic acids, or small molecules. It helps in understanding structural changes upon binding and characterizing binding interfaces. Amide solvent exchange rates offer critical insights into the structure and dynamics of unfolded proteins. Protection factors, which compare exchange data to ideal random-coil values, indicate structural elements like nascent helices or hydrophobic clusters.

The ^{15}NH/D-SOLEXSY experiment, which uses nitrogen spin to probe solvent exchange, measures exchange rates without the complications associated with protons (Chevelkov et al., 2010). This method revealed that α-Synuclein in neurons remains disordered within cells, while FlgM from *Salmonella typhimurium* shows mixed characteristics in different environments (Smith et al., 2015). However, it has limitations due to sensitivity issues and complex data analysis.

The Relax-EXSY experiment, a simplified version of H/D-SOLEXSY, offers increased sensitivity and simplicity (Lopez et al., 2014). It does not require ^{13}C labeling and avoids lengthy coherence transfers, thus enhancing sensitivity. This method is efficient in measuring amide exchange rates for various amino acids, as demonstrated in the TauF4 fragment study, by using different H_2O/D_2O ratios to improve accuracy (Lopez et al., 2014).

2.2. Fluorescence spectroscopy

Fluorescence spectroscopy allows the analysis of biomolecules carrying intrinsic fluorophores (occurring naturally such as tryptophan) and extrinsic fluorophores (dyes and probes). The fluorescence lifetime of a fluorophore, as well as its quantum yield, is crucial to determine the interaction time of a molecule with its binding partners or its diffusion time in a particular environment. Fluorescence-based techniques such as Fluorescence Resonance Energy Transfer (FRET), fluorescence anisotropy, and Fluorescence Correlation Spectroscopy (FCS) are used to study IDPs. These methods help illuminate the folding dynamics of IDPs, disorder-order conformational transitions, attributes of the disordered regions on the IDP, and changes in configuration upon interaction with binding partners (Lakowicz, 2006; Nag et al., 2022).

2.2.1. Intrinsic protein fluorescence

Intrinsic fluorescence of a protein emerges from its aromatic amino acid residues i.e. tryptophan, tyrosine, and phenylalanine. The indole group of tryptophan is the predominant source of both UV absorbance and emission. The indole group is a solvent-sensitive fluorophore. Its emission spectra reveal the location of a tryptophan residue as solvent exposed or buried within the core of a protein. Tryptophan present in a hydrophobic environment has emission between 330 nm and 335 nm (Lakowicz, 2006). An emission spectrum between 340 nm and 345 nm is observed for partially exposed residues of tryptophan near the surface of a protein. While completely exposed tryptophan residues, such as in an IDP, show maximum emission at 350 nm–355 nm. An IDP interacting with a binding partner can alter the location of a tryptophan residue from solvent exposed to partially buried interface and result in a blue-shift in the emission maxima of the protein.

Using intrinsic tryptophan fluorescence, the structure of the C-terminal domain of caldesmon (CaD136), an actin-binding protein found in chicken gizzard, was studied (Permyakov et al., 2003). This protein plays a role in smooth muscle contraction, non-muscle motility, and cytoskeleton generation. The tryptophan fluorescence spectrum for CaD136 peaks at 350 nm, which suggests their solvent-exposed nature. Furthermore, CD spectroscopy and SEC analysis of the C-terminal domain confirms the lack of a secondary structure. On binding to CaM, CaD136 shows a blueshift in the fluorescence emission from 350 nm to 333 nm. Additional experiments show that CaM binds to a hydrophobic pocket created by tryptophan on CaD136 which leads to the burial of the indole rings, inducing partial folding of the protein structure (Permyakov et al., 2003).

2.2.2. Extrinsic fluorescence probes

Often IDPs lack tryptophan or tyrosine residues, which results in the absence of intrinsic fluorescence, and interact with molecules that are devoid of intrinsic fluorescence. In such cases, extrinsic fluorescent probes are used to characterize IDPs and their interactions with other molecules. Both covalent and noncovalent fluorophores are commercially available. Covalent fluorophores come with a varied range of reactive groups that can couple with amines, sulfhydryl, and histidine side chains in proteins. Dansyl Chloride, a solvent-sensitive fluorophore, is used to label proteins, particularly for anisotropy measurements (Lakowicz, 2006).

An extrinsic fluorophore, 8-Anilino-1-naphthalenesulfonic acid (ANS), binds exclusively to the hydrophobic patches on proteins. Thus, it binds better to the expanded states of proteins, such as molten globule, as compared to the compact folded state, where hydrophobic residues are buried in the core of the protein. This was demonstrated for carbonic anhydrase, β-lactamase, and α-lactalbumin (Semisotnov et al., 1991). ANS was also used to show that the native dimeric state of BpUreG, an essential protein for the *in vivo* activation of urease, is disordered even though it maintains a significant amount of tertiary structure (Neyroz et al., 2006). Its transition to molten globule state followed by pre-molten globule state, upon addition of GuHCl, was also detected by ANS fluorescence. Using ANS binding the ribosomal protein Toxoplasma P2 (TgP2) was shown to have a monomeric molten globule state at acidic pH compared to the multimeric state with reduced hydrophobic surface area at neutral pH (Mishra et al., 2015).

2.2.3. Fluorescence correlation spectroscopy

Fluorescence correlation spectroscopy (FCS) is a technique used to observe a single molecule or a group of molecules, bearing fluorophores, in solution (Lakowicz, 2006). It has been used to study the compactness of several IDPs. Beta casein (β-CN) is an IDP and forms one of the components of milk. FCS measurements showed that the β-CN structure samples both extended (hydrodynamic radius $r_H = 34.5 \pm 7$ nm) and collapsed ($r_H = 1$ nm) states in solution, whereas inside a pluronic copolymer (F127) it experiences 10–16 times higher translational friction were observed (Pramanik et al., 2021). For α-synuclein a distribution of diffusion coefficients was obtained from repetitive FCS measurements, with short sampling time, revealing the disappearance of monomeric α-synuclein and the nucleation of oligomerization. In this experiment, Alexa488-labeled α-synuclein in traces was probed with different concentrations of excess unlabeled α-synuclein. The disappearance of fast-diffusing species was noticed during the lag phase of oligomerization. With the increase in aggregation time, a higher form of oligomer represented by heterogeneous distribution was observed. A distinct intermediate distribution was spotted between the sharp monomer peak and a wider distribution of higher oligomers, with an increase in concentration, having a diffusion coefficient of 78 ± 15 $\mu m^2/s$. This confers that a gradual decrease in the monomeric state and the appearance of intermediate states along with heterogeneous higher oligomers is concentration-dependent, followed by conformational changes prior to β-structure formation (Nath et al., 2010).

2.2.4. Fluorescence anisotropy

Fluorescence anisotropy accounts for the orientational changes in a molecule in between excitation and emission spectra. In fluorescence anisotropy, light emitted by a fluorophore in a solution has intensities along different axes of polarization. In a sample when excited by a polarized light, molecules containing fluorophores whose transition moments align parallel to that of the electric vector of the incident radiation have the highest probability of getting excited (Lakowicz, 2006; Nag et al., 2022). For a sample excited by vertically polarized light, the probability of excitation light absorbed by a fluorophore is proportional to cos2 θ, where θ is the angle made by the absorption dipole with the z-axis. Therefore, excitation with polarized light provides a population of excited fluorophores that are partially aligned with the z-axis. The intensity of the emission from the fluorophore is measured through a polarizer situated in front of a detector either placed parallelly or perpendicularly to the incident beam. The observed intensities, depending on the position of the emission polarizer are denoted as I|| and I⊥. These intensity values are used to calculate anisotropy (r) and the equation is as follows,

$$r = (I_{\parallel} - I_{\perp}) / (I_{\parallel} + 2I_{\perp})$$

Henceforth, fluorescence anisotropy is dependent on the rotation of the fluorophore dipole and is associated with the rotational correlation which in turn is dependent on the hydrodynamic radii of the molecules. Change in the hydrodynamic radius of a molecule due to association or dissociation with alike or different molecules consequently influences its anisotropy (Lakowicz, 2006).

Fluorescence anisotropy is often used to study molecular dynamics, particularly to monitor order-to-disorder transitions. Disordered regions in ligands can increase binding affinity and increase the rate of association by allowing maximum functional ligand conformations during random ligand-substrate interplay before binding. PKIα is a 75 amino acids long isoform of an inhibitor of cAMP-dependent protein kinase (cAPK), housing disordered regions. Fluorescence anisotropy was used to study the binding dynamics of PKIα to cAPK. Mutations were made at tentative interacting sites and three full-length single-site mutants were obtained (V3C, S28C, and S59C) which were further conjugated with fluorescein iodoacetamide via the side chains of the substituted cysteines. Anisotropy decay for each mutant in solution reported similar results mostly due to the local R-carbonyl backbone movements. Although, a small population of anisotropy decay showed slower, whole-body tumbling which affirmed that PKIα is mostly disordered at all these selected regions. The binding of the mutants at the catalytic site of cAPK lowered the rate of whole-body tumbling for the three mutants. The anisotropy decay profiles for both FI-V3C-PKIα and FI-S28C-PKIα on interacting with cAPK remarkably reduced contributions to the fast decay. In hindsight, FI-S59C- PKIα rendered unaffected on interacting with cAPK. Hence, using anisotropy one can not only identify flexible or disordered regions on IDPs but also estimate binding sites on it (Hauer et al., 1999).

2.2.5. Single molecule FRET

Single molecule fluorescence (or Förster) resonance energy transfer (smFRET) is a remarkably sensitive method capable of probing intramolecular distances between polypeptide chains ranging from ~2–10 nm, which is particularly useful to study structural compactness and distance dynamics in disordered proteins (Schuler et al., 2016). This method is amenable to a varied range of protein concentrations, reaction conditions, and the size of the proteins or protein complexes. smFRET can also be combined with other fluorescence techniques to investigate the structural diversity of an IDP further. With smFRET, a broad timescale of dynamics can be covered (Schuler and Hofmann, 2013).

SmFRET has been used to probe IDP aggregates, binding kinetics of disordered proteins, conformational changes upon binding and formation of biomolecular condensates. Diffusion-based smFRET hinders protein aggregation while acquiring real-time measurements under favorable conditions. This helps in the characterization of the dynamic nature and the different transitional states of proteins prone to aggregation such as α-Synuclein in Parkinson's disease and β-amyloid and tau in Alzheimer's disease (Tompa et al., 2009; Trexler and Rhoades, 2010). In another study, smFRET in association with ensemble smFRET could decipher conformational packing along with its allying subunits mandate for the formation or the dissolution of a secondary structure of a particular domain of nucleophosmin, responsible for oligomerization (Banerjee et al., 2016). A more recent discovery concerning IDPs is their ability to facilitate liquid-liquid phase separation; these IDPs are highly prone to form bio-condensates that exist as membrane-less organelles *in vivo*. smFRET techniques, such as TIRF-FRET (Total internal reflection fluorescence microscope-FRET), have been used to study the underlying dynamics of IDPs leading to LLPS. TIRF-FRET allows the real-time tracking of labeled molecules thereby enabling conformational sampling within each droplet (Elbaum-Garfinkle et al., 2015). Using this technique the viscosity and dynamics of the disordered P granule protein LAF-1 from *C. elegans* was studied. The intrinsically disordered RNA helicase LAF-1 phase separates into liquid droplets promoting P granule formation in *C. elegans*. On binding to RNA, the viscosity of the droplet as well as the dynamics are altered. Interaction with RNA decreases the viscosity and increases internal molecular dynamics of LAF-1 droplets. (Elbaum-Garfinkle, S et al., 2015).

2.3. Mass spectrometry

Mass spectrometry has been widely used to identify multiple conformations as well as interaction motifs in IDPs. Mass spectrometry can also be used to detect protium-deuterium exchange of the backbone amides similar to NMR spectroscopy (Balasubramaniam and Komives, 2013). The advantage of mass spectrometry over NMR spectroscopy is that no isotope labeling, such as ^{15}N, is required, less amount of protein is used and proteins of any size can be analyzed. The disadvantage is lower resolution and absence of rate of exchange information. Both electrospray ionization (ESI) (Banerjee and Mazumdar, 2012) and MALDI-TOF (Mandell et al., 1998) can be used to measure amide hydrogen exchange.

The peroxisome proliferator-activated receptor γ coactivator-1α (PGC-1α) has an intrinsically disordered N-terminal activation domain. Hydrogen exchange by mass spectrometry (HX-MS) studies on PGC-1α revealed a disordered-to-ordered transition of the isolated activation domain of PGC-1α upon binding to the ligand-binding domain of estrogen-related receptor γ (ERRγLBD) (Devarakonda et al., 2011). Isolated PGC-1α was fully deuterated within seconds of hydrogen exchange. In the presence of the ligand-binding domain of the nuclear receptor PPARγ, the nuclear receptor box 2 region of PGC-1α becomes substantially protected against hydrogen exchange. Similar protection was detected in complexes between PGC-1α and the ligand-binding domains of two other nuclear receptors: RORγ and ERRγ, and the corresponding interaction regions were identified.

Ion mobility mass spectrometry (IM-MS) was used to investigate stepwise conformational changes of human MT2A as a function of the number of Cd^{2+} ions bound (Chen et al., 2014). Human metallothionein-2A (MT) is an intrinsically disordered protein that acts as an *in vivo* metal reservoir and transporter. ApoMT samples several conformational states ranging from globular-like compact to extended conformations. In the initial steps of metal addition (number of bound ions = 1−3), the metal ions bind to different sites of the protein resulting in distinct conformations. For more than 4 bound ions, the conformational changes are associated with the β domain, which becomes more ordered upon binding to metal ions, and converges to a dumbbell-shaped conformation.

2.4. Limited proteolysis

Limited proteolysis is a simple and rapid technique to dissect structural features of a protein (Fontana et al., 2012). It can be used to identify disordered segments, rigid domains and also to study the unfolding dynamics of a protein. The choice of proteases can define the nature and location of cleavage resulting in the isolated regions. Most commonly used proteases in limited proteolysis are trypsin, subtilisin, proteinase K, chymotrypsin and thermolysin (Fontana et al., 2012). The enzyme substrate ratio used is typically 1:100 or 1:50. For IDPs, lower ratios of 1:1000 to 1:10,000 are also used as the rate of proteolytic degradation is faster due to the easier accessibility and proteolysis of disordered regions. The time and temperature are controlled depending upon the nature of the objective for the experiment.

Usually, the proteolyzed fragments are analyzed by SDS-PAGE. Based on the type of proteases used, the boundaries of the cleavage site can help to identify rigid domains in IDPs. The reaction mixtures can also be separated and analyzed by reverse phase-HPLC; and the molecular mass of the fragments determined by mass spectrometry (Gao et al., 2005). Limited proteolysis helps identify whether an IDP is completely disordered, or has one or more folded domains. Coupled with mass spectrometry, precise domain boundaries can be identified (Gao et al.,

2005).

High throughput screening using limited proteolysis and mass spectrometry were performed to identify folded domains and disordered regions in 400 proteins (Gao et al., 2005). Limited proteolysis was used to predict intrinsically disordered regions of human proIGF-1Ea, proIGF-1Eb and proIGF-1Ec. The E-domain was found to carry disordered residues and the Eb-domain was also shown to carry two predicted molecular recognition features (MoRFs) and two phosphorylation sites (Annibalini et al., 2018). Using limited proteolysis, the local unfolding of the chain region 82–97 surrounding the helix F in the holo protein sperm-whale apomyoglobin (apoMb) was detected (Picotti et al., 2004).

2.5. Circular dichroism (CD) spectroscopy

The usefulness of circular dichroism (CD) spectroscopy lies in its ability to probe the secondary structures of proteins in solution in a rapid manner, using small amounts of sample and no additional labeling (Drake, 1986). Far-UV CD spectrum is sensitive to the peptide bond environment and reveals important characteristics of the protein's secondary structure. The near-UV CD spectrum (>250 nm) of proteins provides information on the tertiary structure. IDPs can form transient and dynamic secondary structures (Uversky and Dunker, 2010), such as random coils, helical regions and β-turns which can co-exist in the same polypeptide chain. CD excels in capturing these subtle structural changes, offering a window into the diverse conformational heterogeneity of IDPs. Changing the chemical environment of these persistent and transient structures can help stabilize them as was observed for the activation domain of the ACTR (activator for thyroid hormone and retinoid receptors), the cytosolic C−terminal distal tail of the human hoNHE1cdt (sodium–proton exchanger) and the Spd1 (S−phase delayed protein) where increase in temperature caused loss of structure in all helical elements and changes in the PPII content (Kjaergaard et al., 2010). For α-synuclein, addition of TFE causes an increase in **α**-helix content (Micsonai et al., 2015). Analyzing the difference between the CD spectra of a protein in different conditions can help identify the structural elements and conformational transitions in the IDPs.

CD spectroscopy provides quick and real-time data acquisition, allowing monitoring of the changes in protein structure in a matter of minutes. It also requires relatively small sample volumes in comparison to NMR spectroscopy. Its sensitivity to changes in the protein environment such as the pH, temperature, and ligand binding allows the study of effects of various conditions on IDPs. CD can be adapted for high-throughput studies which allows the screening of multiple sample conditions or samples rapidly. While CD spectroscopy offers these advantages, it is essential to note that a combination of techniques is needed to gain a comprehensive understanding of IDPs.

2.6. UV–visible absorption spectroscopy

UV–Visible spectroscopy is one of the oldest analytical techniques that measures the intensity of light in UV (100–400 nm) and visible (400–800 nm) regions as a function of wavelength (Akash and Rehman, 2020). UV–Visible spectroscopy has been widely used to study IDPs by analyzing the absorbance spectra of aromatic amino acids such as tryptophan, tyrosine, and phenylalanine (Nag et al., 2022). The absorption spectra of these aromatic amino acids present in IDPs is different from the absorbance spectra of ordered proteins due to high flexibility of the IDPs and more solvent exposure of these residues in IDPs. UV–Visible spectroscopy can be very useful in characterizing the intrinsic disorder of a protein. UV absorption properties can be used to conformationally probe protein folding and unfolding (Mignon et al., 2022). The process of protein folding and unfolding can be analyzed by calculating the foldedness index (FI), calculated as $FI = (A^{280}/A^{275}) + (A^{280}/A^{258})$ (Mignon et al., 2022). The FI decreases upon denaturation as protein absorbance shifts from 280 to 258 nm, diminishes the contribution at 280 nm.

FI has been used to characterize the degree of disorder of an epigenetic factor which plays important roles in chromatin remodeling and acts as oncogenic agents. The double plant homeodomain finger 3 (DPF3) protein exists in two isoforms, DPF3a and DPF3b. The two isoforms are different in their sequence length and composition at the C-terminal region. Both the isoforms of DPF3 belong to the class of IDPs. However, DPF3b is less disordered in comparison to DPF3a because it has an ordered C-terminal region. FI studies on both the isoforms of DPF3 display interesting absorption properties, exhibiting a maximum absorption spectra at ~258–260 nm. This results in a low FI value of ~1.4 and ~1.7 for DPF3a and DPF3b, respectively, suggesting that both the isoforms are highly disordered. The lower FI value for DPF3a indicates that it is more disordered than DPF3b (Mignon et al., 2022).

2.7. Size exclusion chromatography

Size exclusion chromatography (SEC) also known as molecular exclusion chromatography or gel permeation chromatography is a common chromatography technique used for separating biomolecules, such as proteins, as per their hydrodynamic dimensions. IDPs, due to their extended conformations, have larger hydrodynamic radii compared to folded proteins. Since SEC primarily separates analytes based on their hydrodynamic radii, it is a useful technique to determine the degree of compactness of an IDP (Uversky, 2012).

SEC of the C-terminal domain of chicken gizzard caldesmon (CaD136) revealed that its hydrodynamic dimension (Stokes radius) at neutral pH closely matched the dimensions under denaturing conditions using 6 M Guanidine hydrochloride (Permyakov et al., 2003). This confirmed that CaD136 remains unfolded even in the absence of denaturants. In another study, hydrodynamics dimensions of three synucleins, α, β and γ, were measured by SEC. β-synuclein was found to have a hydrodynamic radius resembling that of a completely unfolded polypeptide. On the other hand, α-synuclein and γ-synuclein were found to be more compact under the same conditions (Uversky et al., 2002). A comparative study on the degree of compactness for all three proteins was done based on their V_h^{coil}/V_h ratios, where V_h^{coil} corresponds to the hydrodynamic volume of the random coil of a certain molecular mass and V_h refers to the measured hydrodynamic volume and it was concluded that β-synuclein sampled a more extended conformation, under physiological pH, than α-synuclein and γ-synuclein.

Similarly, in cGMP phosphodiesterase (PDE) is a tetramer comprising two homologous catalytic subunits, α and β, and two inhibitory γ subunits. To further characterize the conformation of the γ subunit, the hydrodynamic radius (R_S) of PDEγ was measured under neutral pH both in the absence and in the presence of 8 M urea. The SEC profile for both conditions showed single symmetric elution peaks; the positions of these peaks were utilized to obtain the R_S value. It was observed that the hydrodynamic dimensions under non-denaturing conditions were close to the measurements determined under denaturing conditions. The R_s values resembled that of a fully unfolded protein with a molecular mass of 9.6 KDa. Hence it was established that PDEγ is natively unfolded which makes this subunit target subject involved in various functional interactions (Uversky et al., 2002).

2.8. Raman spectroscopy

Raman spectroscopy has been used to study the secondary composition and conformational heterogeneity of several IDPs as a label-free, real-time, and non-invasive detection method. The IDP α-synuclein was analyzed by three-component band fitting of the amide-I region of its Raman spectra (Maiti et al., 2004). The signature bands at 1653 cm^{-1} for α-helix, 1667 cm^{-1} for β-strand, and 1674−1685 cm^{-1} for extended β-strand and polyproline II structure were used to fit the Raman spectra of α-synuclein and identify the presence, or change in relative contribution, of these secondary structural elements. A similar approach has been used to study the secondary structure composition and

conformational heterogeneity of the onco-suppressor protein p53 in different conditions (Signorelli et al., 2017). In this case, the amide-I bands in the Raman spectra were analyzed by principal component analysis to discriminate the spectral band of the disordered region of p53 and its folded DNA-binding domain. Raman spectroscopy has also been used to study the impact of aggregation promoting conditions, cofactors and post-translational modifications on the formation of tau fibrils (Devitt et al., 2021). Unique signatures in the Raman spectra were identified for tau fibrils using heparin or RNA cofactors and for phosphorylated tau.

Less sample requirement and fast data acquisition make Raman spectroscopy an attractive method for the studies of IDR. Its biggest limitation is the low sensitivity due to weak Raman scattering but the detection can be improved with installation of highly optimized instrumentation (Bumbrah and Sharma, 2016).

2.9. Fourier transform infrared spectroscopy

Fourier transform infrared (FTIR) spectroscopy has found valuable applications in studying IDPs. This is another non-destructive method that requires a small amount of protein sample and provides unique insights into the structural features, conformational changes, and interaction dynamics of IDPs in solution. Random coil structures and α-helices typically exhibit peaks around 1645-1655 cm^{-1} and 1655-1665 cm^{-1} in the Amide-I band, respectively. Integrating the area under specific peaks gives an estimate of relative content of random coil and α-helical structures in the IDP. Antiparallel β-sheets exhibit a distinct peak in the lower wavenumber range of Amide-I band (1615-1635 cm^{-1}). This peak is often sharper and more defined for antiparallel β-sheets. The parallel β-sheets exhibit peaks between 1635 and 1660 cm^{-1} in the Amide I band, which are broader and less intense compared to the antiparallel β-sheets. β-turns and loops can contribute to peaks around 1670-1690 cm^{-1} in the Amide I band. These peaks are generally less intense than those of α-helices and β-sheets.

FTIR deconvolution was used to quantify the secondary structure content of the IDP tau in solution (Rauch et al., 2017). This study revealed the dominance of random coil structure and minor ***β***-sheet content. In the IDP Sic1, FTIR revealed that the isolated C-terminal KID of Sic1 retains dynamic helical structure and populates collapsed states of different compactness (Brocca et al., 2011). FTIR spectroscopy also monitors spectral changes upon folding, providing real-time information. FTIR was used to investigate the folding of the IDP p27(Kip1). It was observed that p27(Kip1) showed changes in the Amide-I band region indicating the formation of ***β***-sheet structure upon binding (Sivakolundu et al., 2005). FTIR spectra can also be recorded at different temperatures, revealing thermal-induced changes. FTIR spectroscopy study of IDP CRL1 (*Candida rugosa* lipase 1) temperature-induced structural transitions indicated a transition from a random coil to both α-helical and ***β***-sheet-rich structure at lower temperatures (Natalello et al., 2005). FTIR spectroscopy elucidated the binding mechanism of IDP Androgen receptor (AR) to its partner protein RAP74. The spectral changes revealed the conformational transition of the AF1 domain of AR upon binding, from a disordered structure to a more helical conformation (Kumar et al., 2004).

FTIR spectroscopy can be performed in aqueous as well as non-aqueous environments. Its aqueous compatibility allows environments resembling physiological conditions, enabling the study of IDPs in biologically relevant conditions. It also allows for the study of protein films that can be obtained by evaporating protein solutions. This can be done using Attenuated Total reflection (ATR). When an infrared beam passes through a high refractive index material (usually a crystal of zinc, selenide, diamond, or germanium) with a flat surface upon which the sample is placed, it undergoes multiple internal reflections at the interface between the crystal and the sample (Goormaghtigh et al., 1999). With each reflection, an evanescent wave is generated in the sample, which penetrates a short distance into the sample. ATR has been used to investigate the influence of electrostatic charges on the aggregation propensity of the N-terminus moiety of the measles virus phosphoprotein PNT (Tedeschi et al., 2017).

2.10. Small-angle X-ray scattering

Small-angle X-ray scattering (SAXS) is a powerful tool for characterizing the structure and dynamics of both ordered and disordered proteins in solution with nanometer resolution (Bernadó and Svergun, 2012). SAXS provides data on size, oligomeric state, and shape for proteins and complexes ranging from a few kilodaltons to gigadaltons (Kikhney and Svergun, 2015). It is particularly valuable for studying intrinsically disordered proteins (IDPs) due to its speed, minimal sample requirements, and high precision (Bernadó and Svergun, 2012). In a simple SAXS experiment, macromolecules in a solution are exposed to X-rays, and the scattered X-ray intensity is measured at different angles (Bernadó and Svergun, 2012). By subtracting the solvent's scattering, information about the sample's size and shape is obtained (Kikhney and Svergun, 2015). This technique is particularly useful for studying flexible and polydisperse macromolecules, where the scattering intensity is a sum of contributions from different components in the sample.

Partially folded proteins like Factor H, p53, Prion protein, Alix, TIAR, and Replication Protein A show dual behavior in SAXS experiments due to the presence of both folded and disordered regions (Aslam and Perkins, 2001; Lima et al., 2006; Tidow et al., 2007; Pretto et al., 2010; Shi et al., 2010; Kim et al., 2011). IDPs are highly sensitive to their local environment, and SAXS is extensively used to investigate structural changes in response to environmental factors. SAXS has been successfully used to examine how temperature, pH, ionic strength, reducing agents, post-translational modifications (e.g., phosphorylation, sumoylation, glycosylation), the presence of specific ions, and point mutations impact the structural properties of IDPs (Bernadó and Svergun, 2012). SAXS experiments have also been used to study the folding upon binding of IDPs. In the case of p53, a DNA-binding protein, SAXS has revealed insights into its tetrameric assembly and how it encloses DNA upon binding (Tidow et al., 2007). Additionally, SAXS serves as a valuable tool for validating computational models of unstructured proteins. However, it has limitations, including potential sample damage by intense X-ray beams in synchrotrons and requires highly pure monodisperse samples (90–95%) for getting reliable results (Bernadó and Svergun, 2012; Kikhney and Svergun, 2015).

2.11. Small-angle neutron scattering

Small angle neutron scattering (SANS) parallels SAXS but employs neutrons instead of X-rays. Neutrons interact with atomic nuclei in the sample, and the resultant scattering pattern reveals structural details. SANS uses isotopic labeling, typically deuterium, for enhanced contrast. It uniquely elucidates protein-solute interactions, solvent variations between unfolded and folded proteins, influence of side-chain flexibility on high-angle scattering, and deviations from the random-coil model of IDPs (Chen et al., 2023).

The conformational changes in α-synuclein (αSN) upon its capture by PbaB, a homotetrameric protein that acts as both a proteasome activator and a molecular chaperone, was studied by SANS (Sugiyama et al., 2014). SANS profile of non-deuterated αSN in 100% D_2O was collected in presence and absence of 75% deuterated PbaB tetramer. For free αSN the radius of gyration R_g was determined as 39.1 ± 1.4 Å. SANS also indicated the presence of a partially folded structure in αSN. Interestingly, after interaction of αSN with PbaB tetramer no partially folded structure of αSN was detected and the R_g value dercreased to 36.0 ± 1.6 Å. This indicates that αSN loses its residual ordered structure when it interacts with the PbaB tetramer while leaving the remaining originally unstructured region still flexible (Sugiyama et al., 2014).

2.12. High-speed atomic force microscopy (HS-AFM)

High-speed atomic force microscopy (HS-AFM) has revolutionized our ability to capture real-time, sub-molecular structural dynamics of individual molecules. This technology provides spatial resolution at the sub-molecular level and temporal precision of under 100 ms, all achieved without the necessity for labeling. Its application to study IDRs was first shown for the protein known as the Facilitates Chromatin Transcription (FACT) protein (Miyagi et al., 2008). Subsequently, HS-AFM has been applied to quantitatively analyze the dynamic structural characteristics of several IDPs. Disordered proteins tend to diffuse on mica surfaces rapidly due to their flexibility, which makes it difficult to visualize with HS-AFM. To address this challenge, the diffusion of the protein is slowed down by optimizing conditions such as pH and ionic strength or introducing solubility or affinity tags and facilitating visualization in HS-AFM (Kodera et al., 2021).

Four dimensional coordinates (x, y, z, t) can be deduced from the acquired HS-AFM images for different portions of an IDP such as the highest positions of the globular domains, the mid-sections as well as the terminals. Although positions along a dynamic IDP are difficult to specify accurately, regions flanking at both terminals are imaged with great precision. These positional coordinates are required to obtain both time-series data as well as the two-dimensional distance between the two ends of an IDP (R_{2D}). IDPs in their fully disordered state tend to follow the power law, $R_{2D} = (1.16 \pm 0.057\ \text{nm}) \times N_{aa}\ 0.52 \pm 0.009$, where N_{aa} is the total number of amino acids contained in that disordered region. From this law R_{2D} can be accurately determined even if there is no protein structure available for the IDP (Kodera and Ando, 2022).

HS-AFM images of various disordered proteins, including PQBP-1, Atg1, and Atg13, on mica revealed two categories. PQBP-1-like IDRs remained constantly disordered, while Atg1 and Atg13-like IDRs transitioned between ordered and disordered states (Kodera et al., 2021; Yamamoto et al., 2016). This method quantitatively described the dynamic structure of proteins like Sic1 and MeV phosphoprotein N-terminal domain (PNT) and allowed for the determination of structural transition rates (Kodera et al., 2021). HS-AFM has also been employed in identifying post-transcriptional modifications within intrinsically disordered proteins (IDPs), particularly alterations in length and height ratios (Hashimoto et al., 2013). Furthermore, it has garnered significant attention in the study of biomolecular condensates, shedding light on the initial forces driving liquid-liquid phase separation and providing insights into the surface properties of these droplets (Kodera and Ando, 2022).

2.13. Computational methods

Common experimental techniques such as CD spectroscopy, SAXS, smFRET, NMR, X-ray crystallography, and cryo-electron microscopy offer indirect insights into IDP conformations and may suffer from missing data (Liu et al., 2019; Ramanathan et al., 2021). In contrast, computational approaches play a crucial role in understanding IDPs, encompassing their properties, evolutionary relationships, and the connection between sequence, structure, function, and energy (Ramanathan et al., 2021). A holistic understanding of IDPs is best achieved through the combination of experimental and computational methods.

Bioinformatics presents a cost-effective means to predict disordered proteins and regions based on amino acid sequences. The Disprot database (https://disprot.org/), a curated resource, is frequently utilized for developing disorder prediction tools, aiding in the identification of disorder residues in protein sequence, and predicting regions that exhibit disorder in specific binding interactions (Aspromonte et al., 2024). Over the past two decades, numerous servers and predictors have emerged to provide the propensity of amino acids to be disordered in proteins or entire proteomes (Liu et al., 2019). Some routinely used predictors include PONDR (Romero et al., 2001), PONDR VLXT (Romero et al., 2001), PONDR VSL2 (Obradovic et al., 2005), PONDR VL3 (Peng et al., 2005), PONDR FIT (Xue et al., 2010), IUPred (Dosztányi et al., 2005), FoldIndex (Prilusky et al., 2005), MobiDB (Piovesan et al., 2021), PrDOS (Ishida and Kinoshita, 2007), MetaDisorder (Kozlowski and Bujnicki, 2012), and DisEMBL (Linding et al., 2003). Additionally, specialized predictors like ANCHOR (Dosztányi et al., 2009), MoRFpred (Disfani et al., 2012), MoRFchibi_web (Malhis et al., 2016), DISOPRED3 (Jones and Cozzetto, 2015), DRNAPred (Yan and Kurgan, 2017), DisoRDPbind (Peng et al., 2017), and PPRInt (Kumar et al., 2008) have been developed to identify motifs within disordered regions that can bind to other proteins or nucleic acids. Recent advancements in algorithm development have significantly enhanced the accuracy of disorder predictions.

Artificial intelligence (AI)-based methods have revolutionized predictive capabilities, yielding results closely aligned with experimental data (Gao et al., 2020). Machine learning techniques, such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), Random Forest, and Logistic Regression, are employed for classification tasks. Sequence labeling benefits from methods like Recurrent Neural Networks (RNN), Conditional Random Fields (CRF), and Convolutional Neural Networks (CNN) (Liu et al., 2019). A standard network, such as deep neural networks, when trained using various algorithms, possesses the capability to interpret multiple concealed layers within high-dimensional data (Larrañaga et al., 2006). This neural network-based approach enables the examination of diverse conformations of IDRs/IDPs, along with their emergent characteristics, through a multiscale analysis (Ramanathan et al., 2021). Notably, AlphaFold, a deep learning-based model, excels in protein structure prediction and ranks highly in evaluations (Senior et al., 2020). Low confidence regions in AlphaFold's predictions are found to align with disordered regions, indicating that over 30% of protein regions exhibit disorder (Ruff and Pappu, 2021). Machine learning also aids in modeling protein flexibility and homology prediction.

With the growing computational resources available, protein characterization has become more accessible. Molecular Dynamics (MD) simulations are widely used to explore the structural conformations, transitions, energetics, and interactions of proteins with various partners, ranging from small molecules to large macromolecules. MD simulations replicate cellular environments, shedding light on the structural fluctuations and dynamic ensembles of IDPs (Kumar et al., 2021; Ramanathan et al., 2021). These simulations provide mechanical insights into coupled folding and binding processes. IDP conformations are determined by energy profiles, rather than being completely flexible (Babu, 2016). Various simulation approaches, including all-atom MD, Monte Carlo (MC), coarse-grained (CG), and replica-exchange molecular dynamics (REMD), have been employed to investigate IDP conformations under different conditions. Multiple force fields have been developed for both globular and disordered proteins, including OPLS, CHARMM, GROMOS, AMBER, and Drude, with advanced and modified versions (Kumar et al., 2021). Graphical representations of parameters such as root mean square deviation (RMSD), radius of gyration (R_g), and root mean square fluctuations (RMSF) offer detailed information on conformational changes, compactness, and residue flexibility. Simulation trajectories generate multidimensional data that can be analyzed using statistical parameters, facilitating a better understanding of IDP secondary structures across frames and the analysis of conformational dynamics and structural transitions.

3. Conclusion and future prospects

In this review, we have delved into significant progress made in experimental techniques aimed at investigating the structural and dynamic aspects of IDRs and IDPs. To gain a comprehensive understanding of these proteins, which exhibit a highly heterogeneous nature and conformational fluctuations spanning various time scales, it is imperative to adopt a multiparametric approach rather than relying on a single

tool or method. While notable advancements have been made in this field over the past 15 years, there remains ample room for further improvement. An essential enhancement that could greatly facilitate progress in this area is the development of computational force fields that can accurately simulate realistic structures for both folded and disordered sequences. Presently, existing force fields tend to excel in one category but face difficulties when applied to proteins containing extensive disordered regions alongside folded domains. Another significant challenge lies in the design of motif discovery algorithms, primarily due to the lack of effective multiple sequence alignment tools capable of efficiently aligning IDRs. An intriguing avenue for exploration is the oligomerization properties of IDPs, ranging from simple one-to-one complexes to more complex phenomena like driving phase separation and the formation of aggregates or amyloids. Identifying and characterizing all the species involved, as well as elucidating their functional or pathological roles, presents an ongoing challenge. IDPs present appealing opportunities for drug development due to their close association with a broad spectrum of diseases. Nonetheless, their lack of well-defined binding sites, highly dynamic and transient interactions, and limited site-specific structural data present significant challenges in therapeutically targeting IDPs. The exploration of innovative approaches such as PROTACs (proteolysis-targeting chimeras) and molecular glues designed to target IDRs for either protein degradation or stabilization holds promise for the development of novel therapeutic interventions in diseases linked to disordered proteins.

CRediT authorship contribution statement

Snigdha Maiti: Conceptualization, Writing – review & editing. **Aakanksha Singh:** Writing – review & editing. **Tanisha Maji:** Writing – review & editing. **Nikita V. Saibo:** Writing – review & editing. **Soumya De:** Conceptualization, Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgement

This work and salary support for Tanisha Maji are funded by Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India (CRG/2020/004984). Nikita V. Saibo acknowledges a NET-JRF fellowship NET-JRF fellowship (Joint CSIR-UGC NET JUNE 2016).

References

Abyzov, A., Salvi, N., Schneider, R., Maurin, D., Ruigrok, R.W.H., Jensen, M.R., Blackledge, M., 2016. Identification of dynamic modes in an intrinsically disordered protein using temperature-dependent NMR relaxation. J. Am. Chem. Soc. 138 (19), 6240–6251. https://doi.org/10.1021/jacs.6b02424.

Ahmed, R., Huang, J., Akimoto, M., Shi, T., Melacini, G., 2021. Atomic resolution map of hierarchical self-assembly for an amyloidogenic protein probed through thermal 15 N–R 2 correlation matrices. J. Am. Chem. Soc. 143 (12), 4668–4679. https://doi.org/10.1021/jacs.0c13289.

Akash, M.S.H., Rehman, K., 2020. Ultraviolet-visible (UV-VIS) spectroscopy. Essentials of Pharmaceutical Analysis 1–222. https://doi.org/10.1007/978-981-15-1547-7.

Algamal, M., Ahmed, R., Jafari, N., Ahsan, B., Ortega, J., Melacini, G., 2017. Atomic-resolution map of the interactions between an amyloid inhibitor protein and amyloid β (Aβ) peptides in the monomer and protofibril states. J. Biol. Chem. 292 (42), 17158–17168. https://doi.org/10.1074/jbc.M117.792853.

Annibalini, G., Contarelli, S., De Santi, M., Saltarelli, R., Di Patria, L., Guescini, M., Villarini, A., Brandi, G., Stocchi, V., Barbieri, E., 2018. The intrinsically disordered E-domains regulate the IGF-1 prohormones stability, subcellular localisation and secretion. Sci. Rep. 8 (1) https://doi.org/10.1038/s41598-018-27233-3.

Arnone, A., Bier, C.J., Cotton, F.A., Day, V.W., Hazen, E.E., Richardson, D.C., Richardson, J.S., Yonath, A., 1971. A high resolution structure of an inhibitor complex of the extracellular nuclease of Staphylococcus aureus. J. Biol. Chem. 246 (7), 2302–2316. https://doi.org/10.1016/s0021-9258(19)77221-4.

Asakura, T., Taoka, K., Demura, M., Williamson, M.P., 1995. The relationship between amide proton chemical shifts and secondary structure in proteins. J. Biomol. NMR 6 (3), 227–236. https://doi.org/10.1007/BF00197804.

Aslam, M., Perkins, S.J., 2001. Folded-back solution structure of monomeric factor H of human complement by synchrotron X-ray and neutron scattering, analytical ultracentrifugation and constrained molecular modelling. J. Mol. Biol. 309 (5), 1117–1138. https://doi.org/10.1006/jmbi.2001.4720.

Aspromonte, M.C., Nugnes, M.V., Quaglia, F., Bouharoua, A., Tosatto, S.C., Piovesan, D., 2024. DisProt in 2024: improving function annotation of intrinsically disordered proteins. Nucleic Acids Res. 52 (D1), D434–D441. https://doi.org/10.1093/nar/gkad928.

Babu, M.M., 2016. The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. Biochem. Soc. Trans. 44 (5), 1185–1200. https://doi.org/10.1042/BST20160172. Portland Press Ltd.

Babu, M.M., van der Lee, R., de Groot, N.S., Gsponer, J., 2011. Intrinsically disordered proteins: regulation and disease. Curr. Opin. Struct. Biol. 21 (3), 432–440. https://doi.org/10.1016/j.sbi.2011.03.011.

Balasubramaniam, D., Komives, E.A., 2013. Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. Biochim. Biophys. Acta Protein Proteonomics 1834 (6), 1202–1209.

Banerjee, S., Mazumdar, S., 2012. Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. International journal of analytical chemistry 2012.

Banerjee, P.R., Mitrea, D.M., Kriwacki, R.W., Deniz, A.A., 2016. Asymmetric modulation of protein order–disorder transitions by phosphorylation and partner binding. Angew. Chem. 128 (5), 1707–1711. https://doi.org/10.1002/ange.201507728.

Bastidas, M., Gibbs, E.B., Sahu, D., Showalter, S.A., 2015. A primer for carbon-detected NMR applications to intrinsically disordered proteins in solution. Concepts Magn. Reson. 44 (1), 54–66. https://doi.org/10.1002/cmr.a.21327.

Bermel, W., Bertini, I., Chill, J., Felli, I.C., Haba, N., Kumar, M.V.V., Pierattelli, R., 2012. Exclusively heteronuclear 13C-detected amino-acid-selective NMR experiments for the study of intrinsically disordered proteins (IDPs). Chembiochem 13 (16), 2425–2432. https://doi.org/10.1002/cbic.201200447.

Bermel, W., Bertini, I., Felli, I.C., Lee, Y.M., Luchinat, C., Pierattelli, R., 2006. Protonless NMR experiments for sequence-specific assignment of backbone nuclei in unfolded proteins. J. Am. Chem. Soc. 128 (12), 3918–3919. https://doi.org/10.1021/ja0582206.

Bernadó, P., Svergun, D.I., 2012. Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. Mol. Biosyst. 8 (1), 151–167. https://doi.org/10.1039/C1MB05275F.

Bernadó, P., Blanchard, L., Timmins, P., Marion, D., Ruigrok, R.W.H., Blackledge, M., 2005. A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. Proc. Natl. Acad. Sci. USA 102 (47), 17002–17007. https://doi.org/10.1073/pnas.0506202102.

Bertini, I., Gonnelli, L., Luchinat, C., Mao, J., Nesi, A., 2011. A new structural model of Aβ40 fibrils. J. Am. Chem. Soc. 133 (40), 16013–16022.

Bertoncini, C.W., Jung, Y.-S., Fernandez, C.O., Hoyer, W., Griesinger, C., Jovin, T.M., Zweckstetter, M., 2005. Release of long-range tertiary interactions potentiates aggregation of natively unstructured α-synuclein. Proc. Natl. Acad. Sci. USA 102 (5), 1430–1435. https://doi.org/10.1073/pnas.0407146102.

Bhavesh, N.S., Panchal, S.C., Hosur, R.V., 2001. An efficient high-throughput resonance assignment procedure for structural genomics and protein folding research by NMR. Biochemistry 40 (49), 14727–14735. https://doi.org/10.1021/bi015683p.

Bogatyreva, N.S., Finkelstein, A.V., Galzitskaya, O.V., 2006. Trend of amino acid composition of proteins of different taxa. J. Bioinf. Comput. Biol. 4 (2), 597–608. https://doi.org/10.1142/S0219720006002016.

Boral, S., Maiti, S., Basak, A.J., Lee, W., De, S., 2020. Structural, dynamic, and functional characterization of a DnaX mini-intein derived from Spirulina platensis provides important insights into Intein-mediated catalysis of protein splicing. Biochemistry 59 (50), 4711–4724.

Brocca, S., Testa, L., Sobott, F., Šamalikova, M., Natalello, A., Papaleo, E., Lotti, M., De Gioia, L., Doglia, S.M., Alberghina, L., Grandori, R., 2011. Compaction properties of an intrinsically disordered protein: Sic1 and its kinase-inhibitor domain. Biophys. J. 100 (9), 2243–2252. https://doi.org/10.1016/j.bpj.2011.02.055.

Brutscher, B., Felli, I.C., Gil-Caballero, S., Hošek, T., Kümmerle, R., Piai, A., et al., 2015. NMR methods for the study of instrinsically disordered proteins structure, dynamics, and interactions: general overview and practical guidelines. Intrinsically Disordered Proteins Studied by NMR Spectroscopy, pp. 49–122.

Bumbrah, G.S., Sharma, R.M., 2016. Raman spectroscopy – basic principle, instrumentation and selected applications for the characterization of drugs of abuse. Egypt. J. Food Sci. 6 (3), 209–215. https://doi.org/10.1016/j.ejfs.2015.06.001.

Camilloni, C., De Simone, A., Vranken, W.F., Vendruscolo, M., 2012. Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. Biochemistry 51 (11), 2224–2231. https://doi.org/10.1021/bi3001825.

Chakraborty, S., Paul, S., Hosur, R.V., 2012. Simultaneous acquisition of 13C α- 15N and 1H- 15N- 15N sequential correlations in proteins: application of dual receivers in 3D HNN. J. Biomol. NMR 52 (1), 5–10. https://doi.org/10.1007/s10858-011-9596-z.

Champness, J.N., Bloomer, A.C., Bricogne, G., Butler, P.J.G., Klug, A., 1976. The structure of the protein disk of tobacco mosaic virus to 5 Å resolution. Nature 259 (5538), 20–24. https://doi.org/10.1038/259020a0.
Chen, S.H., Chen, L., Russell, D.H., 2014. Metal-induced conformational changes of human metallothionein-2A: a combined theoretical and experimental study of metal-free and partially metalated intermediates. J. Am. Chem. Soc. 136 (26), 9499–9508.
Chen, S.H., Weiss, K.L., Stanley, C., Bhowmik, D., 2023. Structural characterization of an intrinsically disordered protein complex using integrated small-angle neutron scattering and computing. Protein Sci. 32 (10), 1–17. https://doi.org/10.1002/pro.4772.
Chevelkov, V., Xue, Y., Krishna Rao, D., Forman-Kay, J.D., Skrynnikov, N.R., 2010. 15NH/D-SOLEXSY experiment for accurate measurement of amide solvent exchange rates: application to denatured drkN SH3. J. Biomol. NMR 46 (3), 227–244. https://doi.org/10.1007/s10858-010-9398-8.
Chi, S.W., Lee, S.H., Kim, D.H., Ahn, M.J., Kim, J.S., Woo, J.Y., Torizawa, T., Kainosho, M., Han, K.H., 2005. Structural details on mdm2-p53 interaction. J. Biol. Chem. 280 (46), 38795–38802. https://doi.org/10.1074/jbc.M508578200.
Clore, G.M., Iwahara, J., 2009. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. Chem. Rev. 109 (9), 4108–4139. https://doi.org/10.1021/cr900033p.
Clore, G.M., Starich, M.R., Gronenborn, A.M., 1998. Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses. J. Am. Chem. Soc. 120 (40), 10571–10572. https://doi.org/10.1021/ja982592f.
Clore, G.M., Tang, C., Iwahara, J., 2007. Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. Curr. Opin. Struct. Biol. 17 (5), 603–616. https://doi.org/10.1016/j.sbi.2007.08.013.
Cook, E.C., Usher, G.A., Showalter, S.A., 2018. The use of 13C direct-detect NMR to characterize flexible and disordered proteins. In: Methods in Enzymology, first ed., Vol. 611. Elsevier Inc. https://doi.org/10.1016/bs.mie.2018.08.025.
Cornilescu, G., Delaglio, F., Bax, A., 1999. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J. Biomol. NMR 13 (3), 289–302. https://doi.org/10.1023/A:1008392405740.
De Simone, A., Cavalli, A., Hsu, S.T.D., Vranken, W., Vendruscolo, M., 2009. Accurate random coil chemical shifts from an analysis of loop regions in native states of proteins. J. Am. Chem. Soc. 131 (45), 16332–16333. https://doi.org/10.1021/ja904937a.
De, S., Greenwood, A.I., Acevedo, L.A., Korson, N.E., Nicholson, L.K., 2022. Lineshape analysis as a tool for probing functional motions at biological interfaces. In: NMR Spectroscopy for Probing Functional Dynamics at Biological Interfaces. Royal Society of Chemistry, London, pp. 82–121.
De, S., Greenwood, A.I., Rogals, M.J., Kovrigin, E.L., Lu, K.P., Nicholson, L.K., 2012. Complete thermodynamic and kinetic characterization of the isomer-specific interaction between Pin1-WW domain and the amyloid precursor protein cytoplasmic tail phosphorylated at Thr668. Biochemistry 51 (43), 8583–8596. https://doi.org/10.1021/bi3008214.
Dev, K., 2003. Part II: α-synuclein and its molecular pathophysiological role in neurodegenerative disease. Neuropharmacology 45 (1), 14–44. https://doi.org/10.1016/S0028-3908(03)00140-0.
Devarakonda, S., Gupta, K., Chalmers, M.J., Hunt, J.F., Griffin, P.R., Van Duyne, G.D., Spiegelman, B.M., 2011. Disorder-to-order transition underlies the structural basis for the assembly of a transcriptionally active PGC-1α/ERRγ complex. Proc. Natl. Acad. Sci. USA 108 (46), 18678–18683.
Devitt, G., Crisford, A., Rice, W., Weismiller, H.A., Fan, Z., Commins, C., Hyman, B.T., Margittai, M., Mahajan, S., Mudher, A., 2021. Conformational fingerprinting of tau variants and strains by Raman spectroscopy. RSC Adv. 11 (15), 8899–8915. https://doi.org/10.1039/D1RA00870F.
Dey, S., MacAinsh, M., Zhou, H.-X., 2022. Sequence-dependent backbone dynamics of intrinsically disordered proteins. J. Chem. Theor. Comput. 18 (10), 6310–6323. https://doi.org/10.1021/acs.jctc.2c00328.
Disfani, F.M., Hsu, W.L., Mizianty, M.J., Oldfield, C.J., Xue, B., Keith Dunker, A., Uversky, V.N., Kurgan, L., 2012. MoRFpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order transitioning binding regions in proteins. Bioinformatics 28 (12), 75–83. https://doi.org/10.1093/bioinformatics/bts209.
Dosztányi, Z., Csizmok, V., Tompa, P., Simon, I., 2005. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics 21 (16), 3433–3434. https://doi.org/10.1093/bioinformatics/bti541.
Dosztányi, Z., Mészáros, B., Simon, I., 2009. ANCHOR: web server for predicting protein binding regions in disordered proteins. Bioinformatics 25 (20), 2745–2746. https://doi.org/10.1093/bioinformatics/btp518.
Drake, A.F., 1986. Polarisation modulation-the measurement of linear and circular dichroism. J. Phys. E Sci. Instrum. 19 (3), 170–181. https://doi.org/10.1088/0022-3735/19/3/002.
Dunker, A.K., Brown, C.J., Lawson, J.D., Iakoucheva, L.M., Obradović, Z., 2002. Intrinsic disorder and protein function. Biochemistry 41 (21), 6573–6582. https://doi.org/10.1021/bi012159+.
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., Hipps, K.W., Ausio, J., Nissen, M.S., Reeves, R., Kang, C.H., Kissinger, C.R., Bailey, R.W., Griswold, M.D., Chiu, W., Garner, E.C., Obradovic, Z., 2001. Intrinsically disordered protein. J. Mol. Graph. Model. 19 (1), 26–59. https://doi.org/10.1016/S1093-3263(00)00138-8.
Dunker, A.K., Obradovic, Z., Romero, P., Garner, E.C., Brown, C.J., 2000. Intrinsic protein disorder in complete genomes. Genome Informatics. Workshop on Genome Informatics 11, 161–171. https://doi.org/10.11234/gi1990.11.161.
Dunker, A.K., Oldfield, C.J., Meng, J., Romero, P., Yang, J.Y., Chen, J.W., Vacic, V., Obradovic, Z., Uversky, V.N., 2008a. The unfoldomics decade: an update on intrinsically disordered proteins. BMC Genom. 9 (Suppl. 2) https://doi.org/10.1186/1471-2164-9-S2-S1.
Dunker, A.K., Silman, I., Uversky, V.N., Sussman, J.L., 2008b. Function and structure of inherently disordered proteins. Curr. Opin. Struct. Biol. 18 (6), 756–764. https://doi.org/10.1016/j.sbi.2008.10.002.
Dyson, H.J., Wright, P.E., 2005. Intrinsically unstructured proteins and their functions. Nat. Rev. Mol. Cell Biol. 6 (3), 197–208. https://doi.org/10.1038/nrm1589.
Dyson, H.J., Wright, P.E., 2021. NMR illuminates intrinsic disorder. Curr. Opin. Struct. Biol. 70, 44–52. https://doi.org/10.1016/j.sbi.2021.03.015.
Elbaum-Garfinkle, S., Kim, Y., Szczepaniak, K., Chen, C.C.H., Eckmann, C.R., Myong, S., Brangwynne, C.P., 2015. The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. Proc. Natl. Acad. Sci. U.S.A. 112 (23), 7189–7194. https://doi.org/10.1073/pnas.1504822112.
Esteban-Martín, S., Fenwick, R.B., Salvatella, X., 2010. Refinement of ensembles describing unstructured proteins using NMR residual dipolar couplings. J. Am. Chem. Soc. 132 (13), 4626–4632. https://doi.org/10.1021/ja906995x.
Farrow, N.A., Zhang, O., Szabo, A., Torchia, D.A., Kay, L.E., 1995. Spectral density function mapping using 15 N relaxation data exclusively. J. Biomol. NMR 6, 153–162.
Farrow, N.A., Zhang, Ouwen, Forman-Kay, J.D., Kay, L.E., 1997. Characterization of the backbone dynamics of folded and denatured states of an SH3 domain. Biochemistry 36 (9), 2390–2402. https://doi.org/10.1021/bi962548h.
Felli, I.C., Pierattelli, R., 2014. Novel methods based on 13C detection to study intrinsically disordered proteins. J. Magn. Reson. 241 (1), 115–125. https://doi.org/10.1016/j.jmr.2013.10.020.
Felli, I.C., Bermel, W., Pierattelli, R., 2021. Exclusively heteronuclear NMR experiments for the investigation of intrinsically disordered proteins: focusing on proline residues. Magnetic Resonance 2 (1), 511–522. https://doi.org/10.5194/mr-2-511-2021.
Fischer, E., 1894. Einfluss der Configuration auf die Wirkung der Enzyme. Ber. Dtsch. Chem. Ges. 27, 2985–2993.
Fisher, C.K., Stultz, C.M., 2011. Constructing ensembles for intrinsically disordered proteins. Curr. Opin. Struct. Biol. 21 (3), 426–431. https://doi.org/10.1016/j.sbi.2011.04.001.
Flavell, R.R., Muir, T.W., 2009. Expressed protein ligation (EPL) in the study of signal transduction, ion conduction, and chromatin biology. Accounts Chem. Res. 42 (1), 107–116.
Fontana, A., de Laureto, P.P., Spolaore, B., Frare, E., 2012. Identifying disordered regions in proteins by limited proteolysis. Intrinsically Disordered Protein Analysis: Volume 2, Methods and Experimental Tools 297–318.
Galea, C.A., Wang, Y., Sivakolundu, S.G., Kriwacki, R.W., 2008. Regulation of cell division by intrinsically unstructured proteins: intrinsic flexibility, modularity, and signaling conduits. Biochemistry 47 (29), 7598–7609. https://doi.org/10.1021/bi8006803.
Gall, T. Le, Romero, P.R., Cortese, M.S., Uversky, V.N., Dunker, A.K., 2007. Intrinsic disorder in the protein data bank. J. Biomol. Struct. Dyn. 24 (4), 325–341. https://doi.org/10.1080/07391102.2007.10507123.
Gao, X., Bain, K., Bonanno, J.B., Buchanan, M., Henderson, D., Lorimer, D., Marsh, C., Reynes, J.A., Sauder, J.M., Schwinn, K., Thai, C., Burley, S.K., 2005. High-throughput limited proteolysis/mass spectrometry for protein domain elucidation. J. Struct. Funct. Genom. 6 (2–3), 129–134. https://doi.org/10.1007/s10969-005-1918-5.
Gao, W., Mahajan, S.P., Sulam, J., Gray, J.J., 2020. Deep learning in protein structural modeling and design. Patterns 1 (9). https://doi.org/10.1016/j.patter.2020.100142. Cell Press.
Giasson, B.I., Murray, I.V.J., Trojanowski, J.Q., Lee, V.M.Y., 2001. A hydrophobic stretch of 12 amino acid residues in the middle of α-synuclein is essential for filament assembly. J. Biol. Chem. 276 (4), 2380–2386. https://doi.org/10.1074/jbc.M008919200.
Gibbs, E.B., Kriwacki, R.W., 2018. Direct detection of carbon and nitrogen nuclei for high-resolution analysis of intrinsically disordered proteins using NMR spectroscopy. Methods 138–139, 39–46. https://doi.org/10.1016/j.ymeth.2018.01.004.
Gillespie, J.R., Shortle, D., 1997. Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. J. Mol. Biol. 268 (1), 158–169. https://doi.org/10.1006/jmbi.1997.0954.
Goormaghtigh, E., Raussens, V., Ruysschaert, J.-M., 1999. Attenuated total reflection infrared spectroscopy of proteins and lipids in biological membranes. Biochim. Biophys. Acta Rev. Biomembr. 1422 (2), 105–185. https://doi.org/10.1016/S0304-4157(99)00004-0.
Goradia, N., Wiedemann, C., Herbst, C., Görlach, M., Heinemann, S.H., Ohlenschläger, O., Ramachandran, R., 2015. An approach to NMR assignment of intrinsically disordered proteins. ChemPhysChem 16 (4), 739–746. https://doi.org/10.1002/cphc.201402872.
Greenwood, A.I., Rogals, M.J., De, S., Lu, K.P., Kovrigin, E.L., Nicholson, L.K., 2011. Complete determination of the Pin1 catalytic domain thermodynamic cycle by NMR lineshape analysis. J. Biomol. NMR 51 (1–2), 21–34. https://doi.org/10.1007/s10858-011-9538-9.
Hansen, D.F., 2019. Using deep neural networks to reconstruct non-uniformly sampled NMR spectra. J. Biomol. NMR 73 (10–11), 577–585. https://doi.org/10.1007/s10858-019-00265-1.

Hashimoto, M., Kodera, N., Tsunaka, Y., Oda, M., Tanimoto, M., Ando, T., Morikawa, K., Tate, S.I., 2013. Phosphorylation-coupled intramolecular dynamics of unstructured regions in chromatin remodeler FACT. Biophys. J. 104 (10), 2222–2234. https://doi.org/10.1016/j.bpj.2013.04.007.

Hauer, J.A., Taylor, S.S., Johnson, D.A., 1999. Binding-dependent Disorder−Order transition in PKIα: a fluorescence anisotropy study. Biochemistry 38 (21), 6774–6780.

Hsu, W., Oldfield, C.J., Xue, B., Meng, J., Huang, F., Romero, P., Uversky, V.N., Dunker, A.K., 2013. Exploring the binding diversity of intrinsically disordered proteins involved in one-to-many binding. Protein Sci. 22 (3), 258–273. https://doi.org/10.1002/pro.2207.

Hyberts, S.G., Milbradt, A.G., Wagner, A.B., Arthanari, H., Wagner, G., 2012. Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. J. Biomol. NMR 52 (4), 315–327. https://doi.org/10.1007/s10858-012-9611-z.

Iakoucheva, L.M., Brown, C.J., Lawson, J.D., Obradović, Z., Dunker, A.K., 2002. Intrinsic disorder in cell-signaling and cancer-associated proteins. J. Mol. Biol. 323 (3), 573–584. https://doi.org/10.1016/S0022-2836(02)00969-5.

Ishida, T., Kinoshita, K., 2007. PrDOS: prediction of disordered protein regions from amino acid sequence. Nucleic Acids Res. 35 (Suppl. 2) https://doi.org/10.1093/nar/gkm363.

Jensen, M.R., Markwick, P.R.L., Meier, S., Griesinger, C., Zweckstetter, M., Grzesiek, S., Bernadó, P., Blackledge, M., 2009. Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings. Structure 17 (9), 1169–1185. https://doi.org/10.1016/j.str.2009.08.001.

Jensen, M.R., Ruigrok, R.W.H., Blackledge, M., 2013. Describing intrinsically disordered proteins at atomic resolution by NMR. Curr. Opin. Struct. Biol. 23 (3), 426–435. https://doi.org/10.1016/j.sbi.2013.02.007.

Jensen, M.R., Zweckstetter, M., Huang, J.R., Blackledge, M., 2014. Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy. Chem. Rev. 114 (13), 6632–6660. https://doi.org/10.1021/cr400688u.

Jones, D.T., Cozzetto, D., 2015. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. Bioinformatics 31 (6). https://doi.org/10.1093/bioinformatics/btu744.

Kawale, A.A., Burmann, B.M., 2023. Advanced NMR spectroscopy methods to study protein structure and dynamics. In: Advanced Spectroscopic Methods to Study Biomolecular Structure and Dynamics. Elsevier, pp. 125–152. https://doi.org/10.1016/B978-0-323-99127-8.00010-6.

Kempf, J.G., Loria, J.P., 2002. Protein dynamics from solution NMR: theory and applications. Cell Biochem. Biophys. 37 (3), 187–211. https://doi.org/10.1385/CBB:37:3:187.

Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H., Phillips, D.C., 1958. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature 181 (4610), 662–666. https://doi.org/10.1038/181662a0.

Khan, S.N., Charlier, C., Augustyniak, R., Salvi, N., Déjean, V., Bodenhausen, G., Lequin, O., Pelupessy, P., Ferrage, F., 2015. Distribution of pico- and nanosecond motions in disordered proteins from nuclear spin relaxation. Biophys. J. 109 (5), 988–999. https://doi.org/10.1016/j.bpj.2015.06.069.

Kikhney, A.G., Svergun, D.I., 2015. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. FEBS (Fed. Eur. Biochem. Soc.) Lett. 589 (19PartA), 2570–2577. https://doi.org/10.1016/j.febslet.2015.08.027.

Kim, H.S., Wilce, M.C.J., Yoga, Y.M.K., Pendini, N.R., Gunzburg, M.J., Cowieson, N.P., Wilson, G.M., Williams, B.R.G., Gorospe, M., Wilce, J.A., 2011. Different modes of interaction by TIAR and HuR with target RNA and DNA. Nucleic Acids Res. 39 (3), 1117–1130. https://doi.org/10.1093/nar/gkq837.

Kjaergaard, M., Nørholm, A., Hendus−Altenburger, R., Pedersen, S.F., Poulsen, F.M., Kragelund, B.B., 2010. Temperaturedependent structural changes in intrinsically disordered proteins: formation of α−helices or loss of polyproline II? Protein Sci. 19 (8), 1555–1564. https://doi.org/10.1002/pro.435.

Kjaergaard, M., Poulsen, F.M., 2011. Sequence correction of random coil chemical shifts: Correlation between neighbor correction factors and changes in the Ramachandran distribution. J. Biomol. NMR 50 (2), 157–165. https://doi.org/10.1007/s10858-011-9508-2.

Klimov, D.K., Thirumalai, D., 2003. Dissecting the assembly of Aβ16–22 amyloid peptides into antiparallel β sheets. Structure 11 (3), 295–307.

Knowles, T.P.J., Vendruscolo, M., Dobson, C.M., 2014. The amyloid state and its association with protein misfolding diseases. Nat. Rev. Mol. Cell Biol. 15 (6), 384–396. https://doi.org/10.1038/nrm3810.

Kodera, N., Noshiro, D., Dora, S.K., Mori, T., Habchi, J., Blocquel, D., Gruet, A., Dosnon, M., Salladini, E., Bignon, C., Fujioka, Y., Oda, T., Noda, N.N., Sato, M., Lotti, M., Mizuguchi, M., Longhi, S., Ando, T., 2021. Structural and dynamics analysis of intrinsically disordered proteins by high-speed atomic force microscopy. Nat. Nanotechnol. 16 (2), 181–189. https://doi.org/10.1038/s41565-020-00798-9.

Kodera, N., Ando, T., 2022. Guide to studying intrinsically disordered proteins by high-speed atomic force microscopy. Methods 207, 44–56. https://doi.org/10.1016/j.ymeth.2022.08.008.

Konrat, R., 2014. NMR contributions to structural dynamics studies of intrinsically disordered proteins. J. Magn. Reson. 241 (1), 74–85. https://doi.org/10.1016/j.jmr.2013.11.011.

Kosol, S., Contreras-Martos, S., Cedeño, C., Tompa, P., 2013. Structural characterization of intrinsically disordered proteins by NMR spectroscopy. Molecules 18 (9), 10802–10828. https://doi.org/10.3390/molecules180910802.

Kozlowski, L.P., Bujnicki, J.M., 2012. MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins. BMC Bioinf. 13 (1) https://doi.org/10.1186/1471-2105-13-111.

Kriwacki, R.W., Hengst, L., Tennant, L., Reed, S.I., Wright, P.E., 1996. Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. Proc. Natl. Acad. Sci. U.S.A. 93 (21), 11504–11509. https://doi.org/10.1073/pnas.93.21.11504.

Kumar, R., Betney, R., Li, J., Thompson, E.B., McEwan, I.J., 2004. Induced α-helix structure in AF1 of the androgen receptor upon binding transcription factor TFIIF. Biochemistry 43 (11), 3008–3013. https://doi.org/10.1021/bi035934p.

Kumar, M., Gromiha, M.M., Raghava, G.P.S., 2008. Prediction of RNA binding sites in a protein using SVM and PSSM profile. Proteins: structure. Function and Genetics 71 (1), 189–194. https://doi.org/10.1002/prot.21677.

Kumar, P., Sharma, N., Kumar, A., Giri, R., 2021. Molecular dynamic simulation of intrinsically disordered proteins and relevant forcefields. In: Singh, S.K. (Ed.), Innovations and Implementations of Computer Aided Drug Discovery Strategies in Rational Drug Design. Springer, Singapore, pp. 317–333. https://doi.org/10.1007/978-981-15-8936-2_13.

Lakowicz, J.R. (Ed.), 2006. Principles of Fluorescence Spectroscopy. springer US, Boston, MA.

Larrañaga, P., Calvo, B., Santana, R., Bielza, C., Galdiano, J., Inza, I., Lozano, J.A., Armañanzas, R., Santafé, G., Pérez, A., Robles, V., 2006. Machine learning in bioinformatics. Briefings Bioinf. 7 (1), 86–112. https://doi.org/10.1093/bib/bbk007.

Lazar, T., Martínez-Pérez, E., Quaglia, F., Hatos, A., Chemes, L.B., Iserte, J.A., Méndez, N. A., Garrone, N.A., Saldaño, T.E., Marchetti, J., Rueda, A.J.V., Bernadó, P., Blackledge, M., Cordeiro, T.N., Fagerberg, E., Forman-Kay, J.D., Fornasari, M.S., Gibson, T.J., Gomes, G.N.W., et al., 2021. PED in 2021: a major update of the protein ensemble database for intrinsically disordered proteins. Nucleic Acids Res. 49 (D1), D404–D411. https://doi.org/10.1093/nar/gkaa1021.

Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D.T., Kim, P.M., Kriwacki, R.W., Oldfield, C.J., Pappu, R.V., Tompa, P., Uversky, V.N., Wright, P.E., Babu, M.M., 2014. Classification of intrinsically disordered regions and proteins. Chem. Rev. 114 (13), 6589–6631. https://doi.org/10.1021/cr400525m.

Lima, L.M.T.R., Cordeiro, Y., Tinoco, L.W., Marques, A.F., Oliveira, C.L.P., Sampath, S., Kodali, R., Choi, G., Foguel, D., Torriani, I., Caughey, B., Silva, J.L., 2006. Structural insights into the interaction between prion protein and nucleic acid. Biochemistry 45 (30), 9180–9187. https://doi.org/10.1021/bi060532d.

Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J., Russell, R.B., 2003. Protein disorder prediction: implications for structural proteomics. Structure 11 (11), 1453–1459. https://doi.org/10.1016/j.str.2003.10.002.

Lipari, G., Szabo, A., 1982b. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. J. Am. Chem. Soc. 104.

Lipari, G., Szabo, A., 1982a. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. J. Am. Chem. Soc. 104 (17), 4559–4570. https://doi.org/10.1021/ja00381a010.

Liu, Y., Wang, X., Liu, B., 2019. A comprehensive review and comparison of existing computational methods for intrinsically disordered protein and region prediction. Briefings Bioinf. 20 (1) https://doi.org/10.1093/bib/bbx126. Oxford University Press.

Lopez, J., Ahuja, P., Landrieu, I., Cantrelle, F.X., Huvent, I., Lippens, G., 2014. H/D exchange of a 15N labelled Tau fragment as measured by a simple Relax-EXSY experiment. J. Magn. Reson. 249, 32–37. https://doi.org/10.1016/j.jmr.2014.10.008.

Maiti, N.C., Apetri, M.M., Zagorski, M.G., Carey, P.R., Anderson, V.E., 2004. Raman spectroscopic characterization of secondary structure in natively unfolded proteins: α-synuclein. J. Am. Chem. Soc. 126 (8), 2399–2408.

Maiti, S., Acharya, B., Boorla, V.S., Manna, B., Ghosh, A., De, S., 2019. Dynamic studies on intrinsically disordered regions of two paralogous transcription factors reveal rigid segments with important biological functions. J. Mol. Biol. 431 (7), 1353–1369. https://doi.org/10.1016/j.jmb.2019.02.021.

Maiti, S., De, S., 2021. Identification of potential short linear motifs (SLiMs) in intrinsically disordered sequences of proteins by fast time-scale backbone dynamics. Journal of Magnetic Resonance Open 10–11 (December 2021), 100029. https://doi.org/10.1016/j.jmro.2021.100029.

Malhis, N., Jacobson, M., Gsponer, J., 2016. MoRFchibi SYSTEM: software tools for the identification of MoRFs in protein sequences. Nucleic Acids Res. 44 (W1), W488–W493. https://doi.org/10.1093/NAR/GKW409.

Mao, H., Hart, S.A., Schink, A., Pollok, B.A., 2004. Sortase-mediated protein ligation: a new method for protein engineering. J. Am. Chem. Soc. 126 (9), 2670–2671. https://doi.org/10.1021/ja039915e.

Mandell, J.G., Falick, A.M., Komives, E.A., 1998. Measurement of amide hydrogen exchange by MALDI-TOF mass spectrometry. Anal. Chem. 70 (19), 3987–3995.

Metallo, S.J., 2010. Intrinsically disordered proteins are potential drug targets. Curr. Opin. Chem. Biol. 14 (4), 481–488. https://doi.org/10.1016/j.cbpa.2010.06.169.

Micsonai, A., Wien, F., Kernya, L., Lee, Y.-H., Goto, Y., Réfrégiers, M., Kardos, J., 2015. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. Proc. Natl. Acad. Sci. USA 112 (24), E3095–E3103. https://doi.org/10.1073/pnas.1500851112.

Mignon, J., Mottet, D., Leyder, T., Uversky, V.N., Perpète, E.A., Michaux, C., 2022. Structural characterisation of amyloidogenic intrinsically disordered zinc finger protein isoforms DPF3b and DPF3a. Int. J. Biol. Macromol. 218 (June), 57–71. https://doi.org/10.1016/j.ijbiomac.2022.07.102.

Mirsky, A.E., Pauling, L., 1936. On the structure of native, denatured, and coagulated proteins. Proc. Natl. Acad. Sci. USA 22 (7), 439–447. https://doi.org/10.1073/pnas.22.7.439.

Mishra, P., Choudhary, S., Hosur, R.V., 2015. Ribosomal Protein P2 from apicomplexan parasite Toxoplasma gondii is intrinsically a molten globule. Biophys. Chem. 200, 27–33. https://doi.org/10.1016/j.bpc.2015.03.008.

Mittag, T., Forman-Kay, J.D., 2007. Atomic-level characterization of disordered protein ensembles. Curr. Opin. Struct. Biol. 17 (1), 3–14. https://doi.org/10.1016/j.sbi.2007.01.009.

Miyagi, A., Tsunaka, Y., Uchihashi, T., Mayanagi, K., Hirose, S., Morikawa, K., Ando, T., 2008. Visualization of intrinsically disordered regions of proteins by high-speed atomic force microscopy. ChemPhysChem 9 (13), 1859–1866. https://doi.org/10.1002/cphc.200800210.

Muchmore, S.W., Sattler, M., Liang, H., Meadows, R.P., Harlan, J.E., Yoon, H.S., Nettesheim, D., Chang, B.S., Thompson, C.B., Wong, S.-L., Ng, S.-C., Fesik, S.W., 1996. X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death. Nature 381 (6580), 335–341. https://doi.org/10.1038/381335a0.

Nag, N., Chetri, P.B., Uversky, V.N., Giri, R., Tripathi, T., 2022. Experimental methods to study intrinsically disordered proteins. In: Advances in Protein Molecular and Structural Biology Methods. Elsevier Inc. https://doi.org/10.1016/B978-0-323-90264-9.00031-3.

Narayanan, C., Bafna, K., Roux, L.D., Agarwal, P.K., Doucet, N., 2017. Applications of NMR and computational methodologies to study protein dynamics. Arch. Biochem. Biophys. 628, 71–80. https://doi.org/10.1016/j.abb.2017.05.002.

Natalello, A., Ami, D., Brocca, S., Lotti, M., Doglia, S.M., 2005. Secondary structure, conformational stability and glycosylation of a recombinant Candida rugosa lipase studied by Fourier-transform infrared spectroscopy. Biochem. J. 385 (2), 511–517. https://doi.org/10.1042/BJ20041296.

Nath, S., Meuvis, J., Hendrix, J., Carl, S.A., Engelborghs, Y., 2010. Early aggregation steps in α-synuclein as measured by FCS and FRET: evidence for a contagious conformational change. Biophys. J. 98 (7), 1302–1311.

Neyroz, P., Zambelli, B., Ciurli, S., 2006. Intrinsically disordered structure of Bacillus pasteurii UreG as revealed by steady-state and time-resolved fluorescence spectroscopy. Biochemistry 45 (29), 8918–8930.

Nováček, J., Zawadzka-Kazimierczuk, A., Papoušková, V., Žídek, L., Šanderová, H., Krásný, L., Koźmiński, W., Sklenář, V., 2011. 5D 13C-Detected experiments for backbone assignment of unstructured proteins with a very low signal dispersion. J. Biomol. NMR 50 (1), 1–11. https://doi.org/10.1007/s10858-011-9496-2.

O'Hare, B., Benesi, A.J., Showalter, S.A., 2009. Incorporating 1H chemical shift determination into 13C-direct detected spectroscopy of intrinsically disordered proteins in solution. J. Magn. Reson. 200 (2), 354–358. https://doi.org/10.1016/j.jmr.2009.07.014.

Oates, M.E., Romero, P., Ishida, T., Ghalwash, M., Mizianty, M.J., Xue, B., Dosztányi, Z., Uversky, V.N., Obradovic, Z., Kurgan, L., Dunker, A.K., Gough, J., 2013. D2P2: database of disordered protein predictions. Nucleic Acids Res. 41 (D1), 508–516. https://doi.org/10.1093/nar/gks1226.

Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P., Dunker, A.K., 2005. Exploiting heterogeneous sequence properties improves prediction of protein disorder. Proteins: structure. Function and Genetics 61 (Suppl. 7), 176–182. https://doi.org/10.1002/prot.20735.

Oldfield, C.J., Meng, J., Yang, J.Y., Qu, M.Q., Uversky, V.N., Dunker, A.K., 2008. Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. BMC Genom. 9 (Suppl. 1), 1–20. https://doi.org/10.1186/1471-2164-9-S1-S1.

Orekhov, V.Y., Jaravine, V.A., 2011. Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. Prog. Nucl. Magn. Reson. Spectrosc. 59 (3), 271–292. https://doi.org/10.1016/j.pnmrs.2011.02.002.

Palmer III, A.G., 2001. NMR probes of molecular dynamics: overview and comparison with other techniques. Annu. Rev. Biophys. Biomol. Struct. 30 (1), 129–155. https://doi.org/10.1146/annurev.biophys.30.1.129.

Panchal, S.C., Bhavesh, N.S., Hosur, R.V., 2001. Improved 3D triple resonance experiments , HNN and HN (C) N , for H N and 15 N sequential correlations in (13 C , 15 N) labeled proteins : application to unfolded proteins. Signals C, 135–147.

Pancsa, R., Tompa, P., 2012. Structural disorder in eukaryotes. PLoS One 7 (4). https://doi.org/10.1371/journal.pone.0034687.

Peng, K., Vucetic, S., Radivojac, P., Brown, C.J., Dunker, A.K., obradovic, z., 2005. Optimizing long intrinsic disorder predictors with protein evolutionary information. J. Bioinf. Comput. Biol. 3 (1), 35–60. https://doi.org/10.1142/S0219720005000886.

Peng, Z., Wang, C., Uversky, V.N., Kurgan, L., 2017. Prediction of disordered RNA, DNA, and protein binding regions using DisoRDPbind. In: Zhou, Y., Kloczkowski, A., Faraggi, E., Yang, Y. (Eds.), Prediction of Protein Secondary Structure. Springer, New York, pp. 187–203. https://doi.org/10.1007/978-1-4939-6406-2_14.

Permyakov, S.E., Millett, I.S., Doniach, S., Permyakov, E.A., Uversky, V.N., 2003. Natively unfolded C-terminal domain of caldesmon remains substantially unstructured after the effective binding to calmodulin. Proteins: Struct., Funct., Bioinf. 53 (4), 855.

Picotti, P., Marabotti, A., Negro, A., Musi, V., Spolaore, B., Zambonin, M., Fontana, A., 2004. Modulation of the structural integrity of helix F in apomyoglobin by single amino acid replacements. Protein Sci. 13 (6), 1572–1585. https://doi.org/10.1110/ps.04635304.

Piovesan, D., Necci, M., Escobedo, N., Monzon, A.M., Hatos, A., Mičetić, I., Quaglia, F., Paladin, L., Ramasamy, P., Dosztányi, Z., Vranken, W.F., Davey, N.E., Parisi, G., Fuxreiter, M., Tosatto, S.C.E., 2021. MobiDB: intrinsically disordered proteins in 2021. Nucleic Acids Res. 49 (D1), D361–D367. https://doi.org/10.1093/nar/gkaa1058.

Pramanik, U., Chakraborty, S., Bhattacharyya, K., Mukherjee, S., 2021. An intrinsically disordered protein in F127 hydrogel: fluorescence correlation spectroscopy and structural diversity of beta casein. Chem. Phys. Lett. 762, 138105.

Pretto, D.I., Tsutakawa, S., Brosey, C.A., Castillo, A., Chagot, M.E., Smith, J.A., Tainer, J. A., Chazin, W.J., 2010. Structural dynamics and single-stranded DNA binding activity of the three N-terminal domains of the large subunit of replication protein a from small angle X-ray scattering. Biochemistry 49 (13), 2880–2889. https://doi.org/10.1021/bi9019934.

Prilusky, J., Felder, C.E., Zeev-Ben-Mordehai, T., Rydberg, E.H., Man, O., Beckmann, J.S., Silman, I., Sussman, J.L., 2005. FoldIndex©: a simple tool to predict whether a given protein sequence is intrinsically unfolded. Bioinformatics 21 (16), 3435–3438. https://doi.org/10.1093/bioinformatics/bti537.

Prusiner, S.B., 2001. Neurodegenerative diseases and prions. N. Engl. J. Med. 344 (20), 1516–1526. https://doi.org/10.1056/NEJM200105173442006.

Quaglia, F., Mészáros, B., Salladini, E., Hatos, A., Pancsa, R., Chemes, L.B., Pajkos, M., Lazar, T., Peña-Díaz, S., Santos, J., Ács, V., Farahi, N., Fichó, E., Aspromonte, M.C., Bassot, C., Chasapi, A., Davey, N.E., Davidović, R., Dobson, L., et al., 2022. DisProt in 2022: improved quality and accessibility of protein intrinsic disorder annotation. Nucleic Acids Res. 50 (D1), D480–D487. https://doi.org/10.1093/nar/gkab1082.

Radhakrishnan, I., Pérez-Alvarado, G.C., Parker, D., Dyson, H.J., Montminy, M.R., Wright, P.E., 1997. Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. Cell 91 (6), 741–752. https://doi.org/10.1016/S0092-8674(00)80463-8.

Ramanathan, A., Ma, H., Parvatikar, A., Chennubhotla, S.C., 2021. Artificial intelligence techniques for integrative structural biology of intrinsically disordered proteins. Curr. Opin. Struct. Biol. 66, 216–224. https://doi.org/10.1016/j.sbi.2020.12.001. Elsevier Ltd.

Rao Kakita, V.M., Bopardikar, M., Kumar Shukla, V., Rachineni, K., Ranjan, P., Singh, J. S., Hosur, R.V., 2018. An efficient combination of BEST and NUS methods in multidimensional NMR spectroscopy for high throughput analysis of proteins. RSC Adv. 8 (32), 17616–17621. https://doi.org/10.1039/c8ra00527c.

Rauch, J.N., Olson, S.H., Gestwicki, J.E., 2017. Interactions between microtubule-associated protein tau (MAPT) and small molecules. Cold Spring Harbor Perspectives in Medicine 7 (7), a024034. https://doi.org/10.1101/cshperspect.a024034.

Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J., Dunker, A.K., 2001. Sequence complexity of disordered protein. Proteins: structure. Function and Genetics 42 (1), 38–48. https://doi.org/10.1002/1097-0134(20010101)42:1<38::AID-PROT50>3.0.CO;2-3.

Ruff, K.M., Pappu, R.V., 2021. AlphaFold and implications for intrinsically disordered proteins. J. Mol. Biol. 433 (20) https://doi.org/10.1016/j.jmb.2021.167208. Academic Press.

Sahu, D., Bastidas, M., Showalter, S.A., 2014. Generating NMR chemical shift assignments of intrinsically disordered proteins using carbon-detected NMR methods. Anal. Biochem. 449 (1), 17–25. https://doi.org/10.1016/j.ab.2013.12.005.

Saibo, N.V., Maiti, S., Acharya, B., De, S., 2022. Analysis of structure and dynamics of intrinsically disordered regions in proteins using solution NMR methods. In: Advances in Protein Molecular and Structural Biology Methods. Academic Press, pp. 535–550. https://doi.org/10.1016/B978-0-323-90264-9.00032-5.

Salmon, L., Nodet, G., Ozenne, V., Yin, G., Jensen, M.R., Zweckstetter, M., Blackledge, M., 2010. NMR characterization of long-range order in intrinsically disordered proteins. J. Am. Chem. Soc. 132 (24), 8407–8418. https://doi.org/10.1021/ja101645g.

Salvi, N., Abyzov, A., Blackledge, M., 2017. Atomic resolution conformational dynamics of intrinsically disordered proteins from NMR spin relaxation. Prog. Nucl. Magn. Reson. Spectrosc. 102–103, 43–60. https://doi.org/10.1016/j.pnmrs.2017.06.001.

Salvi, N., Abyzov, A., Blackledge, M., 2019. Solvent-dependent segmental dynamics in intrinsically disordered proteins. Sci. Adv. 5 (6) https://doi.org/10.1126/sciadv.aax2348.

Sattler, M., 1999. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog. Nucl. Magn. Reson. Spectrosc. 34 (2), 93–158. https://doi.org/10.1016/S0079-6565(98)00025-9.

Schanda, P., Van Melckebeke, H., Brutscher, B., 2006. Speeding up three-dimensional protein NMR experiments to a few minutes. J. Am. Chem. Soc. 128 (28), 9042–9043. https://doi.org/10.1021/ja062025p.

Schechter, A.N., Chen, R.F., Anfinsen, C.B., 1970. Kinetics of folding of staphylococcal nuclease. Science 167 (3919), 886–887. https://doi.org/10.1126/science.167.3919.886.

Schneider, R., Maurin, D., Communie, G., Kragelj, J., Hansen, D.F., Ruigrok, R.W.H., Jensen, M.R., Blackledge, M., 2015. Visualizing the molecular recognition trajectory of an intrinsically disordered protein using multinuclear relaxation dispersion NMR. J. Am. Chem. Soc. 137 (3), 1220–1229. https://doi.org/10.1021/ja511066q.

Schuler, B., Hofmann, H., 2013. Single-molecule spectroscopy of protein folding dynamics-expanding scope and timescales. Curr. Opin. Struct. Biol. 23 (1), 36–47. https://doi.org/10.1016/j.sbi.2012.10.008.

Schuler, B., Soranno, A., Hofmann, H., Nettels, D., 2016. Single-molecule FRET spectroscopy and the polymer physics of unfolded and intrinsically disordered proteins. Annu. Rev. Biophys. 45, 207–231. https://doi.org/10.1146/annurev-biophys-062215-010915.

Schwarzinger, S., Kroon, G.J., Foss, T.R., Wright, P.E., Dyson, H.J., 2000. Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRView. J. Biomol. NMR 18, 43–48.

Sekhar, A., Kay, L.E., 2019. An NMR view of protein dynamics in health and disease. Annu. Rev. Biophys. 48, 297–319. https://doi.org/10.1146/annurev-biophys-052118-115647.

Sela, Michael, White Jr., F.H., Anfinsen, C.B., 1957. Reductive cleavage of disulfide bridges in ribonuclease. Science 125, 691–692.
Semisotnov, G.V., Rodionova, N.A., Razgulyaev, O.I., Uversky, V.N., Gripas, A.F., Gilmanshin, R.I., 1991. Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. Biopolymers: Original Research on Biomolecules 31 (1), 119–128.
Senior, A.W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Žídek, A., Nelson, A.W.R., Bridgland, A., Penedones, H., Petersen, S., Simonyan, K., Crossan, S., Kohli, P., Jones, D.T., Silver, D., Kavukcuoglu, K., Hassabis, D., 2020. Improved protein structure prediction using potentials from deep learning. Nature 577 (7792), 706–710. https://doi.org/10.1038/s41586-019-1923-7.
Shi, X., Opi, S., Lugari, A., Restouin, A., Coursindel, T., Parrot, I., Perez, J., Madore, E., Zimmermann, P., Corbeil, J., Huang, M., Arold, S.T., Collette, Y., Morelli, X., 2010. Identification and biophysical assessment of the molecular recognition mechanisms between the human haemopoietic cell kinase Src homology domain 3 and ALG-2-interacting protein X. Biochem. J. 431 (1), 93–102. https://doi.org/10.1042/BJ20100314.
Shigemitsu, Y., Hiroaki, H., 2018. Common molecular pathogenesis of disease-related intrinsically disordered proteins revealed by NMR analysis. J. Biochem. 163 (1), 11–18. https://doi.org/10.1093/jb/mvx056. Oxford University Press.
Signorelli, S., Cannistraro, S., Bizzarri, A.R., 2017. Structural characterization of the intrinsically disordered protein p53 using Raman spectroscopy. Appl. Spectrosc. 71 (5), 823–832. https://doi.org/10.1177/0003702816651891.
Sivakolundu, S.G., Bashford, D., Kriwacki, R.W., 2005. Disordered p27Kip1 exhibits intrinsic structure resembling the Cdk2/cyclin A-bound conformation. J. Mol. Biol. 353 (5), 1118–1128. https://doi.org/10.1016/j.jmb.2005.08.074.
Skriver, K., Theisen, F.F., Kragelund, B.B., 2023. Conformational entropy in molecular recognition of intrinsically disordered proteins. Curr. Opin. Struct. Biol. 83 https://doi.org/10.1016/j.sbi.2023.102697. Elsevier Ltd.
Smith, A.E., Zhou, L.Z., Pielak, G.J., 2015. Hydrogen exchange of disordered proteins in Escherichia coli. Protein Sci. 24 (5), 706–713. https://doi.org/10.1002/pro.2643.
Su, X.-C., Otting, G., 2010. Paramagnetic labelling of proteins and oligonucleotides for NMR. J. Biomol. NMR 46 (1), 101–112. https://doi.org/10.1007/s10858-009-9331-1.
Sugase, K., Dyson, H.J., Wright, P.E., 2007. Mechanism of coupled folding and binding of an intrinsically disordered protein. Nature 447 (7147), 1021–1025. https://doi.org/10.1038/nature05858.
Sugiyama, M., Yagi, H., Yamaguchi, T., Kumoi, K., Hirai, M., Oba, Y., Sato, N., Porcar, L., Martel, A., Kato, K., 2014. Conformational characterization of a protein complex involving intrinsically disordered protein by small-angle neutron scattering using the inverse contrast matching method: a case study of interaction between α-synuclein and PbaB tetramer as a model chap. J. Appl. Crystallogr. 47 (1), 430–435. https://doi.org/10.1107/S1600576713033475.
Tamiola, K., Acar, B., Mulder, F.A.A., 2010. Sequence-specific random coil chemical shifts of intrinsically disordered proteins. J. Am. Chem. Soc. 132 (51), 18000–18003. https://doi.org/10.1021/ja105656t.
Tedeschi, G., Mangiagalli, M., Chmielewska, S., Lotti, M., Natalello, A., Brocca, S., 2017. Aggregation properties of a disordered protein are tunable by pH and depend on its net charge per residue. Biochim. Biophys. Acta Gen. Subj. 1861 (11), 2543–2550. https://doi.org/10.1016/j.bbagen.2017.09.002.
Theillet, F.X., Binolfi, A., Frembgen-Kesner, T., Hingorani, K., Sarkar, M., Kyne, C., Li, C., Crowley, P.B., Gierasch, L., Pielak, G.J., Elcock, A.H., Gershenson, A., Selenko, P., 2014. Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs). Chem. Rev. 114 (13), 6661–6714. https://doi.org/10.1021/cr400695p. American Chemical Society.
Tidow, H., Melero, R., Mylonas, E., Freund, S.M.V., Grossmann, J.G., Carazo, J.M., Svergun, D.I., Valle, M., Fersht, A.R., 2007. Quaternary structures of tumor suppressor p53 and a specific p53–DNA complex. Proc. Natl. Acad. Sci. USA 104 (30), 12324–12329. https://doi.org/10.1073/pnas.0705069104.
Tjandra, N., Bax, A., 1997. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science 278 (5340), 1111–1114. https://doi.org/10.1126/science.278.5340.1111.
Tompa, P., 2002. Intrinsically unstructured proteins. TRENDSin Biochemical Sciences 27 (10), 527–533.
Tompa, P., 2012. Intrinsically disordered proteins: a 10-year recap. Trends Biochem. Sci. 37 (12), 509–516. https://doi.org/10.1016/j.tibs.2012.08.004.
Tompa, P., Davey, N.E., Gibson, T.J., Babu, M.M., 2014. A Million peptide motifs for the molecular biologist. Mol. Cell 55 (2), 161–169. https://doi.org/10.1016/j.molcel.2014.05.032.
Tompa, P., Fuxreiter, M., Oldfield, C.J., Simon, I., Dunker, A.K., Uversky, V.N., 2009. Close encounters of the third kind: disordered domains and the interactions of proteins. Bioessays 31 (3), 328–335. https://doi.org/10.1002/bies.200800151.
Trexler, A.J., Rhoades, E., 2010. Single molecule characterization of α-synuclein in aggregation-prone states. Biophys. J. 99 (9), 3048–3055. https://doi.org/10.1016/j.bpj.2010.08.056.
Turoverov, K.K., Kuznetsova, I.M., Uversky, V.N., 2010. The protein kingdom extended: ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. Prog. Biophys. Mol. Biol. 102 (2–3), 73–84. https://doi.org/10.1016/j.pbiomolbio.2010.01.003.
Uversky, V.N., 2002. What does it mean to be natively unfolded? Eur. J. Biochem. 269 (1), 2–12.
Uversky, V.N., 2011. Intrinsically disordered proteins from A to Z. Int. J. Biochem. Cell Biol. 43 (8), 1090–1103. https://doi.org/10.1016/j.biocel.2011.04.001.
Uversky, V.N., 2012. Size-exclusion chromatography in structural analysis of intrinsically disordered proteins. Intrinsically Disordered Protein Analysis: Volume 2, Methods and Experimental Tools 179–194.
Uversky, V.N., 2013. Unusual biophysics of intrinsically disordered proteins. Biochim. Biophys. Acta Protein Proteonomics 1834 (5), 932–951. https://doi.org/10.1016/j.bbapap.2012.12.008.
Uversky, V.N., 2019. Intrinsically disordered proteins and their "Mysterious" (meta) physics. Frontiers in Physics 7 (FEB). https://doi.org/10.3389/fphy.2019.00010. Frontiers Media SA.
Uversky, V.N., Dunker, A.K., 2010. Understanding protein non-folding. Biochim. Biophys. Acta Protein Proteonomics 1804 (6), 1231–1264. https://doi.org/10.1016/j.bbapap.2010.01.017.
Uversky, V.N., Gillespie, J.R., Fink, A.L., 2000. Why are 'natively unfolded' proteins unstructured under physiologic conditions? Proteins: structure. Function and Genetics 41 (3), 415–427. https://doi.org/10.1002/1097-0134(20001115)41:3<415::AID-PROT130>3.0.CO;2-7.
Uversky, V.N., Li, J., Souillac, P., Millett, I.S., Doniach, S., Jakes, R., et al., 2002. Biophysical properties of the synucleins and their propensities to fibrillate: inhibition of α-synuclein assembly by β-and γ-synucleins. J. Biol. Chem. 277 (14), 11970–11978.
Uversky, V.N., Oldfield, C.J., Dunker, A.K., 2008. Intrinsically disordered proteins in human diseases: introducing the D 2 concept. Annu. Rev. Biophys. 37 (1), 215–246. https://doi.org/10.1146/annurev.biophys.37.032807.125924.
Vallurupalli, P., Bouvignies, G., Kay, L.E., 2012. Studying 'invisible' excited protein states in slow exchange with a major state conformation. J. Am. Chem. Soc. 134 (19), 8148–8161. https://doi.org/10.1021/ja3001419.
Venkitaraman, A.R., 2002. Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell 108 (2), 171–182. https://doi.org/10.1016/S0092-8674(02)00615-3.
Vinogradova, O., Qin, J., 2011. NMR as a unique tool in assessment and complex determination of weak protein–protein interactions. In: Zhu, G. (Ed.), NMR of Proteins and Small Biomolecules. Springer Berlin Heidelberg, pp. 35–45. https://doi.org/10.1007/128_2011_216.
Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F., Jones, D.T., 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J. Mol. Biol. 337 (3), 635–645. https://doi.org/10.1016/j.jmb.2004.02.002.
Watson, J., Crick, F., 1953. Molecular structure of nucleic acids A strcuture for deoxyribose nucleic acid. Nature 171, 737–738.
Waudby, C.A., Christodoulou, J., 2012. An analysis of NMR sensitivity enhancements obtained using non-uniform weighted sampling, and the application to protein NMR. J. Magn. Reson. 219, 46–52. https://doi.org/10.1016/j.jmr.2012.04.013.
Wells, M., Tidow, H., Rutherford, T.J., Markwick, P., Jensen, M.R., Mylonas, E., Svergun, D.I., Blackledge, M., Fersht, A.R., 2008. Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. Proc. Natl. Acad. Sci. USA 105 (15), 5762–5767. https://doi.org/10.1073/pnas.0801353105.
Wirmer, J., Peti, W., Schwalbe, H., 2006. Motional properties of unfolded ubiquitin: a model for a random coil protein. J. Biomol. NMR 35 (3), 175–186. https://doi.org/10.1007/s10858-006-9026-9.
Wishart, DavidS., Sykes, BrianD., 1994. The 13C Chemical-Shift Index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. J. Biomol. NMR 4 (2), 171–180. https://doi.org/10.1007/BF00175245.
Wright, P.E., Dyson, H.J., 1999. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. J. Mol. Biol. 293 (2), 321–331. https://doi.org/10.1006/jmbi.1999.3110.
Wright, P.E., Dyson, H.J., 2009. Linking folding and binding. Curr. Opin. Struct. Biol. 19 (1), 31–38. https://doi.org/10.1016/j.sbi.2008.12.003.
Wright, P.E., Dyson, H.J., 2015. Intrinsically disordered proteins in cellular signalling and regulation. Nat. Rev. Mol. Cell Biol. 16 (1), 18–29. https://doi.org/10.1038/nrm3920.
Xie, H., Vucetic, S., Iakoucheva, L.M., Oldfield, C.J., Dunker, A.K., Uversky, V.N., Obradovic, Z., 2007. Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. J. Proteome Res. 6 (5), 1882–1898. https://doi.org/10.1021/pr060392u.
Xue, B., Dunbrack, R.L., Williams, R.W., Dunker, A.K., Uversky, V.N., 2010. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. Biochim. Biophys. Acta Protein Proteonomics 1804 (4), 996–1010. https://doi.org/10.1016/j.bbapap.2010.01.011.
Yamamoto, H., Fujioka, Y., Suzuki, S.W., Noshiro, D., Suzuki, H., Kondo-Kakuta, C., Kimura, Y., Hirano, H., Ando, T., Noda, N.N., Ohsumi, Y., 2016. The intrinsically disordered protein Atg13 mediates supramolecular assembly of autophagy initiation complexes. Dev. Cell 38 (1), 86–99. https://doi.org/10.1016/j.devcel.2016.06.015.
Yan, J., Kurgan, L., 2017. DRNApred, fast sequence-based method that accurately predicts and discriminates DNA-and RNA-binding residues. Nucleic Acids Res. 45 (10) https://doi.org/10.1093/nar/gkx059.
Ying, J., Delaglio, F., Torchia, D.A., Bax, A., 2017. Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. J. Biomol. NMR 68 (2), 101–118. https://doi.org/10.1007/s10858-016-0072-7.
Zhang, H., Neal, S., Wishart, D.S., 2003. RefDB: a database of uniformly referenced protein chemical shifts. J. Biomol. NMR 25 (3), 173–195. https://doi.org/10.1023/A:1022836027055.
Zhang, Ouwen, Kay, L.E., Olivier, J.P., Forman-Kay, J.D., 1994. Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. J. Biomol. NMR 4 (6), 845–858. https://doi.org/10.1007/BF00398413.
Zoghbi, H.Y., Orr, H.T., 2000. Glutamine repeats and neurodegeneration. Annu. Rev. Neurosci. 23 (1), 217–247. https://doi.org/10.1146/annurev.neuro.23.1.217.