

Review

Biophysical and Integrative Characterization of Protein Intrinsic Disorder as a Prime Target for Drug Discovery

Shuqi Luo ¹, Samuel Wohl ², Wenwei Zheng ^{3,*} and Sichun Yang ^{1,4,*} 

¹ Center for Proteomics and Department of Nutrition, School of Medicine, Case Western Reserve University, Cleveland, OH 44106, USA

² Department of Physics, Arizona State University, Tempe, AZ 85287, USA

³ College of Integrative Sciences and Arts, Arizona State University, Mesa, AZ 85212, USA

⁴ Case Comprehensive Cancer Center, Case Western Reserve University, Cleveland, OH 44106, USA

* Correspondence: wenweizheng@asu.edu (W.Z.); sichun.yang@case.edu (S.Y.)

Abstract: Protein intrinsic disorder is increasingly recognized for its biological and disease-driven functions. However, it represents significant challenges for biophysical studies due to its high conformational flexibility. In addressing these challenges, we highlight the complementary and distinct capabilities of a range of experimental and computational methods and further describe integrative strategies available for combining these techniques. Integrative biophysics methods provide valuable insights into the sequence–structure–function relationship of disordered proteins, setting the stage for protein intrinsic disorder to become a promising target for drug discovery. Finally, we briefly summarize recent advances in the development of new small molecule inhibitors targeting the disordered N-terminal domains of three vital transcription factors.

Keywords: protein intrinsic disorder; integrative biophysics; drug discovery



Citation: Luo, S.; Wohl, S.; Zheng, W.; Yang, S. Biophysical and Integrative Characterization of Protein Intrinsic Disorder as a Prime Target for Drug Discovery. *Biomolecules* **2023**, *13*, 530. <https://doi.org/10.3390/biom13030530>

Academic Editor: Brigita Urbanc

Received: 10 February 2023

Revised: 7 March 2023

Accepted: 10 March 2023

Published: 14 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Intrinsic disorder in proteins is becoming important due to its prevalence in the human proteome and its roles in cellular signaling in normal and abnormal cells [1]. The amino acid sequence of these intrinsically disordered proteins (IDPs) presents a challenge, as it lacks a well-defined 3D structure and is highly flexible. The demand for their functional and disease-driven understanding is beyond simple sequence-based bioinformatic analysis. An in-depth understanding requires adding the “structure” component to the disorder–function relationship typically expected for structurally folded proteins.

The high flexibility of IDPs prompts a revisit of available biophysical tools. For simplicity, we categorize these tools into experimental and computational methods before discussing their synergistic integration. These techniques often complement one another, fostering the growth of integrative biophysics through a combination of experiments and computations.

The biophysical studies of IDPs present opportunities for their potential application in drug discovery due to their links to various diseases. Targeting the disorder itself, rather than its upstream or downstream coregulator proteins, has been found to be viable, with multiple successful examples reported. This short review concludes by providing an overview of the status of targeting the intrinsic disorder in the N-terminal domains (NTDs) of three key transcription factors: p53, androgen receptor (AR), and estrogen receptor (ER).

2. Experimental Biophysical Techniques

IDPs vary in molecular size from tens to over one thousand amino acids weighing more than 100 kDa [2]. Each IDP may require different biophysical techniques for structural analysis due to varying chain lengths. Nuclear magnetic resonance (NMR) spectroscopy is particularly useful for studying IDPs but is limited to small proteins, e.g., those under

25 kDa. Some other techniques have no size limit, making them suitable for larger proteins, but they may not provide the same level of detail and amino acid coverage as NMR. As such, we begin this section with general techniques before discussing NMR-specific tools, some of which are illustrated in Figure 1.

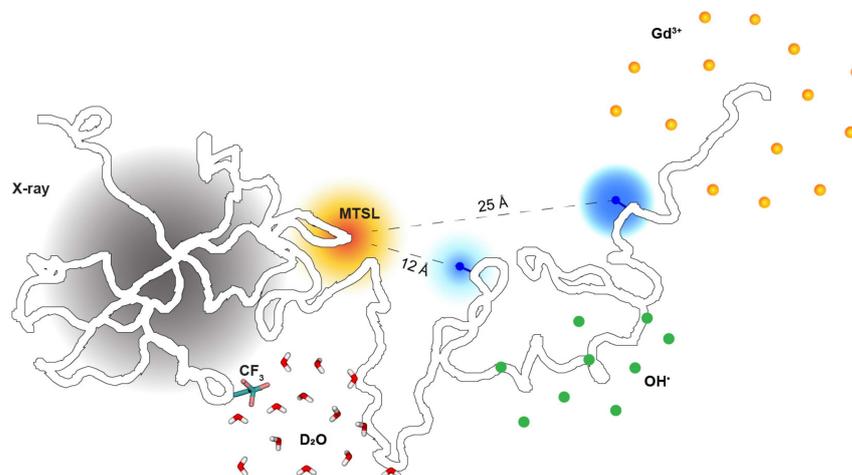


Figure 1. A schematic diagram of various biophysical techniques probing the structural properties of intrinsically disordered proteins (IDPs). A large gray circle illustrates the 2D pattern from small-angle X-ray scattering to provide information about the global conformation and the pairwise distance distribution between atoms within the protein. MTSL (large orange circle): a nitroxide spin label; blue dots: observed NMR-active nuclei whose signal intensity is attenuated (blue circles) to monitor specific distances from the paramagnetic spin-label, typically within the range of 12–25 Å. Gd^{3+} represents gadodiamide. OH: hydroxyl radicals; CF_3 : a trifluoromethyl tag; D_2O : heavy/deuterated water. These agents act as a probe to detect site-specific information about solvent accessibility at the peptide or single-residue level.

2.1. Global Conformations via Small-Angle X-ray Scattering (SAXS)

SAXS has been a primary tool for studying the relationship between the overall radius of gyration (R_g) and polypeptide length N (i.e., number of amino acids) for a broad range of unfolded and natively disordered proteins without size limits [3,4]. A power law typically describes this relationship as $R_g \sim N^v$, where v is a scaling component [5]. Exceptions may occur, particularly for proteins with a high percentage of hydrophobic amino acids [6,7], although this power law remains a proper first-order estimate from a polymer perspective.

The two critical parameters in analyzing an experimental SAXS intensity profile $I(q)$ are the physical R_g and v , where q is the X-ray scattering distance in the reciprocal space (i.e., the amplitude of momenta transfer during the scattering). R_g is typically determined from the low- q region (e.g., $q \cdot R_g < 1.3$), while v is calculated from the high- q region (e.g., $q \cdot R_g \sim 3-10$ [8,9]). The v value of an IDP can range from 0.45 (for compact disorder) to 0.5 (for modest disorder) to 0.65 (for expanded disorder) [10]. This range of disorder behaviors is often illustrated in a Kratky plot of $q^2 \cdot I(q)$ vs. $q \cdot R_g$, distinguishing folded proteins with a bell-like shape from IDPs that level off to reach a plateau at high- q regions [9]. Other parameters, such as the Porod volume that integrates over the entire q region, can assist in, e.g., processing raw scattering data, but their full utilization has yet to be fully explored [11,12].

Individual IDPs are known for their high flexibility with a range/ensemble of conformations in solution. SAXS provides an ensemble-averaged representation of the distribution of distances between all pairs of atoms, a well-known approach for non-biological systems [13]. For proteins, by using the GNOM method to transform the entire $I(q)$ profile, the pair distance distribution function can be determined [14], providing a global view of IDP conformations in aqueous equilibrium.

The high flexibility of IDPs can pose challenges for accurate SAXS data acquisition, as they tend to aggregate or show heterogeneity. Size-exclusion chromatography-coupled SAXS (SEC-SAXS) is a step-forward solution to eliminate unwanted species that may be present in the standard flow cell setup [15,16]. However, a higher protein concentration is required for SEC-SAXS due to dilution from SEC elution. With improvements in synchrotron light sources and increased X-ray brightness, protein concentration is becoming less of a concern, and the desire for accuracy and reliability is often given priority. Thus, SEC-SAXS is the preferred method for obtaining accurate scattering information in biological applications when feasible.

2.2. Site-Specific Solvent Accessibility through the Lens of Three Labeling Techniques

Several biophysical techniques are available to probe the solvent exposure of specific residues in a polypeptide chain. These methods typically involve labeling, quantification, and structural mapping. To show their common features and differences, we describe three exemplary techniques: H/D exchange (HDX) [17–19], hydroxyl radical protein footprinting (HRPF) [20–25], and D₂O-induced fluorine chemical shift perturbations (DFCS) [6,26,27]. These labeling techniques are particularly useful for studying large proteins that NMR cannot analyze.

HDX and HRPF share similar concepts but have distinct features. HDX utilizes excess deuterated (D₂O) buffer to exchange amide hydrogens, while HRPF relies on X-ray hydrolysis [20,28] or laser photolysis [22,29] to generate hydroxyl radicals that can irreversibly and covalently modify the sidechains of individual amino acids. The efficiency of HRPF labeling is based on the diffusion of labeling agents within a short time frame (e.g., milliseconds). The labeling site is a key difference between the methods: HDX focuses on backbone amino hydrogens and HRPF on sidechains. Both techniques often involve a dose–response process at different time windows, followed by the quenching of exchange or reaction before protein digestion by proteases into small peptic peptides.

The DFCS technique uses fluorine labeling by attaching a trifluoromethyl (–CF₃) tag, typically from 3-Bromo-1,1,1-trifluoroacetone (BTFA), to cysteine sidechains [30–34]. This method is beneficial for proteins without native cysteine residues, as the tag can be placed at any position that is mutated to cysteine. However, it can be nontrivial for proteins with multiple cysteines, as it requires identifying individual labeling sites. The process can also be labor-intensive and time-consuming if multiple sites are needed for individual characterization. Each site needs a new protein construct with a cysteine mutation, as we demonstrated for 12 sites of fluorine labeling [6]. Because isotopic D₂O water causes a change in fluorine chemical shift (up to 0.2 ppm), the fluorine tag acts as a probe to evaluate its local solvent environment at varying D₂O concentrations.

Labeling quantification is conducted using liquid chromatography coupled with tandem mass spectrometry for HRPF and most HDX experiments, typically at the level of peptides. Advances in HRPF have enabled a single-residue description, taking advantage of the hydrodynamic difference between individual labeled sites separated by chromatography elution [35,36]. Furthermore, sample delivery has been improved using a liquid injection jet without a container [37]. Other advanced options include time-resolved HRPF, either using a rapid-mixing stopped-flow system [38,39] or a rapid-relaxation temperature jump setup [40], which has been demonstrated to study the kinetics of protein–protein binding, e.g., at the (sub-)millisecond or even microsecond timescale, providing information beyond the ensemble-averaged thermodynamic properties afforded by standard HRPF measurements. DFCS quantification is more straightforward and involves recording fluorine chemical shift spectra and identifying fluorine peaks. The rate/slope of these peaks changes as a function of D₂O concentration report information on the exposure of the fluorine-tagged site to deuterated solvent.

Structural mapping can be achieved from the HDX rate, HRPF rate, and DFCS slope. The protection factor (PF) method has been well established for structural analysis using the HDX rate [18]. A similar PF analysis has been introduced for the HRPF rate, which accounts

for variations among different amino acid types in their intrinsic rate at a peptide or single-residue level [41]. Unlike HDX and HRPF, the DFCS slope allows direct comparison between various labeling sites, utilizing the same fluorine tag uniformly [42].

The final amino acid position coverage varies among techniques due to labeling efficiency, location of sites, and protease digestion. In the case of a 184-residue protein [6], the HDX data provide excellent coverage at the peptide level. However, due to high solvent exposure, the averaging-out across all amino acids within each peptide cannot yield a meaningful description. In contrast, HRPF effectively characterizes the solvent exposure of 16 amino acids (out of 184), demonstrating that some residues are well protected from the solvent despite the intrinsic disorder [6].

2.3. Probing Single Pairwise Distances between Amino Acids

The distance between a specific pair of amino acids can be probed via amino acid labeling. These methods include Förster resonance energy transfer (FRET) [43], double electron–electron resonance (DEER) [44,45], and photoinduced electron transfer (PET) [46]. The major difference between these distance-related methods is the relation between the experimental signal and the distance between the pair of labeled/specific amino acids. Different methods are often most sensitive to different distance regimes. Therefore, they are often applied in other contexts but can sometimes be complementary, considering the wide distance distribution between two amino acids within a conformational ensemble of an IDP. However, interpreting the physical variables from these methods can be non-trivial due to the heterogeneous conformations in IDP ensembles.

FRET. The FRET method covalently links a pair of a donor and an acceptor dye at a specific amino acid site of interest along the chain. The donor dye is optically excited, and the excited energy can either be emitted as a photon or transferred to an acceptor dye. The energy transfer efficiency E is related to the distance r between the pair of dyes if the dye can rapidly experience different orientations within time scales of the donor lifetime. The physical interpretation of the FRET signal can be captured by the Förster equation $E(r) = [1 + (r/R_0)^6]^{-1}$, where R_0 is the Förster radius [47], a value intrinsic to a given set of dyes. This value determines the optimal distance range for FRET measurements. By varying the type of dye, R_0 can range from approximately 40 to 70 Å. Such a distance regime reasonably covers the averaging end-to-end distance of a 100-residue IDP with a size close to a random coil. If multiple pair labeling positions are affordable, FRET can also provide distances between more than one pair of amino acids [10,48]. This information sheds light on the conformational tendencies of various regions of an IDP and reveals scaling behavior [10] and heteropolymeric properties [49].

DEER. The DEER method, a type of electron paramagnetic resonance (EPR) spectroscopy, measures the dipole–dipole couplings between two unpaired electron spins. The spin labels can be introduced as labels on specific amino acids far apart in the sequence. DEER measurements have a distance dependence of r^{-3} in contrast to the r^{-6} dependence in FRET and are most sensitive to distances of 20–80 Å [44]. More specifically, the distance distribution can be obtained through methods such as Tikhonov regularization [50].

PET. In contrast, the PET method does not require a label attached to a specific amino acid; instead, the quenching happens between two naturally occurring amino acids, tryptophan and cysteine [46,51]. These two amino acids are not commonly seen in an IDP sequence, suggesting that it is often impossible to directly measure an IDP's conformation without mutations. PET studies often involve mutating an aromatic amino acid to tryptophan and serine to cysteine, which minimizes the modification impact [52]. The rate of PET decays exponentially as a function of the distance between two amino acids, typically less than 8 Å [53]. This indicates that if PET is applied alone for an IDP, the sequence separation between two amino acids of interest should not exceed 40 residues. This restraint poses a challenge for the PET application, considering the typical length of an IDP. However, due to the growing interest in capturing transient specific interactions within IDPs, PET could be an alternative method focusing on these short-range distances of interest.

For example, PET studies of p53-NTD have revealed a kinetic slowdown of long-range loop closure between two amino acids (e.g., V31 and W53) due to phosphorylation [54].

2.4. Versatile NMR Techniques

For proteins amenable to chemical shift assignments, NMR is a premier tool for in-depth investigations beyond analyzing the transient secondary structure and chemical shift perturbation [55–58]. High protein concentrations, typically above 200 μM , are required for resonance assignments. However, for 2D NMR spectra such as heteronuclear single-quantum coherence (HSQC), a lower concentration of around 30–100 μM is generally sufficient to produce adequate signal-to-noise ratios in a reasonable acquisition time, enabling its broad application to IDPs.

The sampling temperature for NMR data acquisition is an important distinction between disordered and folded proteins. For folded proteins, higher temperatures (e.g., room temperature) are commonly used to enable fast rotational diffusion for sharp resonances due to the restricted mobility of structured regions. On the other hand, lower temperatures (e.g., 4–10 $^{\circ}\text{C}$) are favored for high-quality 2D NMR spectra of highly flexible IDPs because of the increase in line broadening caused by amide hydrogen exchange with the solvent, particularly for solvent-exposed residues.

The low protein concentration requirement for 2D NMR spectra (e.g., HSQC) is a crucial advantage in the studies of highly flexible IDPs. The reduced concentration minimizes interference from nonspecific intermolecular interactions and enables the focus on intramolecular dynamics. Furthermore, this allows using a wide range of NMR techniques to study IDPs. The most informative NMR experiments for IDPs include assessments of backbone dynamics using relaxation measurements, long-range interactions using paramagnetic relaxation enhancements (PRE), and backbone solvent accessibility using solvent-PRE.

^{15}N relaxation. Backbone dynamics can be probed through ^{15}N relaxation measurements (longitudinal R_1 and transverse R_2) by monitoring the intensity decays of individual amino acids [59,60]. This method has been used to study both unfolded and disordered proteins. One approach uses $R_{1\rho}$, the ^{15}N longitudinal relaxation rate in the rotating frame, with longer relaxation delays to account for the relatively slow ^{15}N -relaxation of disordered protein [61]. Applications include identifying residual structural features, such as hydrophobic clustering, and locating regions of restricted backbone mobility as indicated by large R_2/R_1 ratios [62–64].

PRE. The PRE method allows for determining long-range distances between amino acids, typically in the range of 12–25 \AA [65–67] (illustrated in Figure 1), compared to NOE-derived interproton distances of less than 6 \AA [68]. The technique involves attaching a nitroxide spin label, such as MTSL, to a cysteine residue (either native or introduced by mutagenesis) via a disulfide bond. The spin label enhances the transverse relaxation rates of nearby amino acids, resulting in line-broadening due to dipole–dipole interactions between the spin label and NMR-active nuclei. This paramagnetic effect follows an inverse sixth power of the distance between the label and the observed residue, as demonstrated in folded proteins or complexes [69,70]. For many disordered proteins, this PRE method is beneficial in identifying long-range interactions between amino acids [71–76]. It provides information on many amino acid pairs simultaneously, compared to methods that monitor the distance between a single pair of amino acids. An interesting expansion is double spin-labeling, also referred to as paramagnetic relaxation interference [77–79], where the two paramagnetic sites enable accurate triangulation of individual amino acids of interest by probing the collective effect of spin labels on these amino acids without severe intensity attenuation, practically outside the intermediate surroundings of the spin labels.

Solvent-PRE. The solvent accessibility of specific residues can be probed without covalently attaching probes to the protein's amino acids by using highly soluble paramagnetic agents such as Gadodiamide, also known as Omniscan. These agents diffuse freely and rapidly and enhance the transverse relaxation rates of nearby nuclei, such as amide hydrogens [80,81]. The rate of enhancement is proportional to the concentration of

paramagnetic agents in the solution, making it an adequate measure of solvent accessibility. This method has been used to characterize folded proteins [81] and has been applied to disordered proteins showing low solvent exposure of a native-like beta-hairpin and overall high solvent exposure for the rest of the denatured ubiquitin [82,83]. Compared to non-NMR methods, this solvent-PRE method significantly improves the detection of backbone solvent accessibility with a higher amino acid coverage, as many residues are resolved by 2D NMR spectra through either proton or carton detection [84]. A new development uses two differently charged co-solutes (cationic and anionic or neutral) as free-diffusion paramagnetic agents [85–87]. These co-solutes are used to determine an effective per-residue electrostatic potential by utilizing an inverse sixth power of the distance between the paramagnetic co-solutes and the observed residue, particularly useful in characterizing the electrostatics of well-defined ligand-binding cavities, highly charged DNA-binding surfaces, or electrostatics-driven protein–protein interfaces.

3. Theoretical and Computational Biophysical Techniques

Experiments are often accompanied by theoretical and computational methods in various forms. This section explores four aspects of this collaboration between computation and experiment, arranged by ease of application. These aspects include sequence-based predictors for distinguishing IDPs from folded proteins, polymer models for interpreting experimental measurements, molecular simulations and modeling techniques that are parameterized via experimental data, and ensemble-fitting that integrates experiments and computations.

3.1. Prediction from the IDP's Primary Amino Acid Sequence

In the 1990s, while investigating proteins involved in transcription, it was observed that the minimum requirement for functional amino acid sequences often included highly acidic contents and negatively charged amino acids [88,89]. Given the repulsive interactions involved in short regions of tens of amino acids, it was difficult to imagine these regions could fold into a well-defined three-dimensional structure, as often seen in folded proteins. Increasing numbers of sequence segments without a definable “structure” led to attempts at sequence-level classifications between disordered and structured regions of proteins. In 2000, Uversky and Dunker introduced a diagram using two sequence-based descriptors, mean net charge and mean net hydrophobicity, and found that known folded proteins and IDPs often occupy different regions of this two-dimensional diagram [90]. This work demonstrated qualitatively that the physical properties of individual amino acids could be used to predict the general structural preference of IDPs, despite some exceptions where IDPs cross the folded–disordered boundary [6].

Sequence descriptors. Investigations have been carried out on various amino acid properties to determine their suitability for IDP prediction. The first type of sequence descriptor focuses on charged amino acids. The fraction of charged amino acids often affects the contribution of other sequence descriptors based on charged amino acids to the overall conformational preference. For instance, IDPs with a high content of charged residues depend primarily on the arrangement of their charged amino acids [91]. When determining the overall attractive or repulsive interactions, one can look at more detailed charge-relevant sequence descriptions such as net charge per residue (NCPR) or a more complex fraction of positively/negatively charged amino acids, which are often thought to be more effective in predicting the conformational preference of an IDP [92]. The second type of sequence descriptor is based on the hydrophobicity of amino acids, and several hydrophobicity scales are available for the 20 amino acids [93–95]. In addition, secondary structure preference [96,97] and solvent accessibility [98] of amino acids can be used as inputs to predict disordered regions. Furthermore, a fraction of different types of amino acids can be used, and in some cases, a specific type of amino acid has been found to be important to the conformational preference of an IDP [99–101]. However, when considering an n-gram language model (e.g., the fraction of an n-amino-acid pattern) in a protein

sequence, there may be many such sequence descriptors, and a machine learning method is often needed to achieve predictive power [102]. Figure 2A provides an example using two representative sequence descriptors introduced by Uversky [90]: the absolute value of the net charge per residue $|\langle q \rangle|$ and the amino acid hydrophobicity per residue $\langle H \rangle$ [95]. The folded proteins used here were obtained from the TOP2018 database [92] (excluding the regions that cannot be assigned a secondary structure with DSSP software [103]), while disordered proteins were obtained from the DisProt database [104] (with a criterion for a chain length of longer than 30 amino acids). As shown in Figure 2A, most folded proteins are located within the border of this Uversky-proposed boundary line, while disordered proteins occupy a broad range of space. With the increasing number of IDPs, there can be quite a few getting close to the well-folded protein regime, suggesting these proteins have similar sequence properties, at least in terms of the two sequence descriptors used.

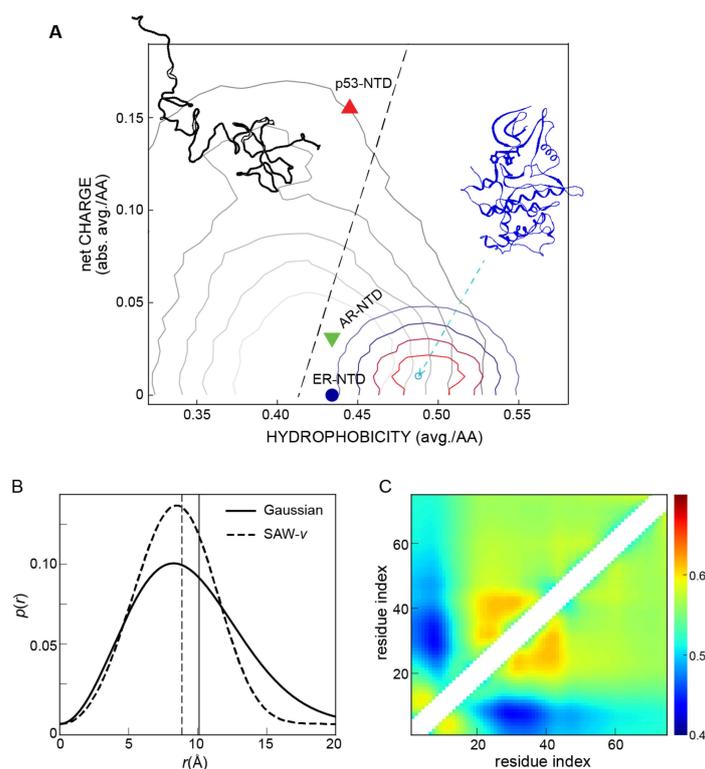


Figure 2. Theoretical and computational methods for studying IDPs. (A) Folded vs. disordered proteins in a two-dimensional plot of hydrophobicity and net charge of individual proteins. Blue: a representative folded structure of CDK7 kinase (PDB id: 7B5O [105]); colored lines: a contour plot derived from a large set of high-resolution folded protein structures available from the Protein Data Bank via the Top2018 database [106]. Black: a representative disordered structure (PDB-Dev id: PDBDEV_00000027 and SASDB id: SASDEE2); gray lines: a contour plot derived from a set of disordered proteins via the DisProt database [104]. Dashed line: a boundary line of $\langle H \rangle = (|\langle q \rangle| + 1.151)/2.785$ as proposed by Uversky [90] where $\langle H \rangle$ is the average hydrophobicity per amino acid [95] and $|\langle q \rangle|$ is the absolute value of average net charge per amino acid. Most folded proteins are within the border of this boundary line, while disordered proteins spread over a broad range of space. Red triangle: p53-NTD (M1-V97; UniPort id: P0463; $\langle H \rangle = 0.445$ and $|\langle q \rangle| = 0.155$); green triangle: AR-NTD (M1-K558; UniPort id: P10275; $\langle H \rangle = 0.434$ and $|\langle q \rangle| = 0.031$); blue circle: ER-NTD (M1-Y184; UniPort id: P03372; $\langle H \rangle = 0.434$ and $|\langle q \rangle| = 0.000$). (B) Polymer models for deriving single-distance distributions indicate that the interpretation of experimental measurement is highly model-dependent by converting a FRET efficiency of 0.2 from a 100 amino-acid peptide to the distance distribution function $p(r)$ [107]. (C) Pairwise scaling exponent map from coarse-grained simulations for the N-terminal disordered region of E-cadherin [49].

Machine learning methods. Several machine learning methods, from simple linear regression to more sophisticated approaches such as support vector machines and deep learning (artificial neural networks with multiple layers), can combine existing sequence descriptors to predict the disordered sequences. With the increasing degrees of freedom (sequence descriptors) and increasing training datasets (e.g., DisProt [108], IDEAL [109], MobiDB [110], and solved PDB structures), deep learning has become a commonly used method for this purpose. A recent assessment testing 43 predictors found that machine learning methods and specifically deep learning methods outperform physicochemical methods [111]. However, predicting disordered regions for binding can still be challenging. Due to the ease of applying these predictors, most of which have existing web interfaces [112–115], one can always try several methods and increase the confidence level of determining the disordered region when facing a new sequence. However, the contribution of a specific sequence descriptor to the prediction or the underlying sequence grammar can often be challenging to access due to the hidden layers of deep learning methods.

Short sequence regions. Significant efforts have been devoted to exploring short sequence regions that facilitate specific interactions between disordered regions and various biomolecules [116]. Two major categories of these regions are molecular recognition fragments (MoRFs) and short linear motifs (SLiMs). MoRF can undergo a disorder-to-order transition upon binding to their partner and can be predicted using various methods with sequence lengths ranging from 5 to 25 amino acids [117–121]. On the other hand, SLiMs are short sequence patches each containing 3 to 15 amino acids that are often found within the disordered regions of diverse proteins and can be highly conserved [116,122–125]. Such sequence conservation, e.g., within low-complexity disordered regions, suggests potential coevolution with binding partners for specific functions. In this case, sequence-based algorithms have been developed to predict binding regions within an IDP that interact with other proteins [126,127], nucleic acids [128], and even lipids [129]. Databases such as DIBS [130] and FuzDB [131] can be used for this purpose or as a training dataset for their algorithm development. Recent evidence has suggested coevolution between SLiMs and linkers for a particular IDP [132], indicating that flanking regions with less-conserved sequences in IDPs might also affect interactions between these short sequence regions and their binding partners [133], although this realization remains to be validated on a case-by-case basis.

Patterning of sequence descriptors. One can also investigate the patterning of existing sequence descriptors, which might provide additional physical insights. It has been shown that charge patterning, for example, plays a significant role in determining individual chain configurations [134–136]. By considering the patterning of even hydrophobic amino acids, often thought of as secondary to charged interactions, predictions of global properties such as polymer scaling exponent and radius of gyration are further improved [137]. Charge patterning can also be applied to understand the interactions between two IDPs dominated by the patterning of the charged amino acids [138]. More interestingly, the charge block idea has been realized for some critical IDP functions [139,140].

It should be noted that structure prediction methods such as AlphaFold v2.0 [141,142] can be used to distinguish folded and disordered regions. In addition, sequence-based algorithms can be extended to predict other factors such as prion-like domains [143,144], liquid–liquid phase separation [145–147], protein aggregation [148,149], and mutual synergistic protein folding [150], with an increasing number of experimental measurements serving as the training data set. A clear advantage of sequence-based predictors is their ease of use. Many predictors come with a web interface, making them accessible to quickly analyze new sequences of interest before more complex computational and experimental techniques are applied. Therefore, it is recommended to use sequence-based predictors before using any other computational/theoretical methods for IDPs. Even though there is ongoing interest in developing computational models for both folded and disordered proteins, most of the methods described here only apply to IDPs.

3.2. Polymer Models for Interpreting Experimental Measurements

Experimental measurements typically correspond to averaged physical variables from an ensemble of diverse conformations. Without a physics-based model, it is nontrivial to convert the experimental measurements directly. For instance, an experimental measurement that provides the distance between two amino acids still requires a distance distribution function to connect the experimental signal and the distance r . This can be performed through various methods, ranging from polymer models with analytical equations for $p(r)$ to all-atom explicit solvent simulations. This section briefly describes polymer models, often the first step for interpreting experimental data.

Gaussian chain. When looking at the sizes of IDPs measured using SAXS (i.e., R_g), FRET (i.e., distance R), and dynamic light scattering (DLS) or pulsed-field gradient NMR (i.e., hydrodynamic radius R_h), IDPs of varying chain lengths N were found to be close to the scaling behavior of a random coil as R_g , R_h or $R \sim N^\nu$, where ν is the scaling exponent [3,4,151]. Therefore, a Gaussian chain model [5] is often used for analyzing, e.g., FRET data and helping convert the FRET efficiency into the distance between the pair labeling positions. The distance distribution function $P(r)$ of the model can be written as

$$P(r) = \left(\frac{3}{2\pi}\right)^{3/2} \frac{4\pi}{R} \left(\frac{r}{R}\right)^2 \exp\left[-\frac{3}{2}\left(\frac{r}{R}\right)^2\right]$$

where R is the root mean squared distance of all the conformations in the ensemble and r is the distance between a specific pair of amino acids for one conformation. Then, R can be obtained by minimizing $|\int E(r)P(r)dr - E_{expt}|$, where $E(r)$ is the Förster equation [47] describing the FRET efficiency as a function of the distance and E_{expt} as the experimentally measured FRET efficiency. However, for one specific IDP, the scaling exponent can differ from 0.5. A FRET investigation that labeled multiple pair positions on different proteins revealed that six IDPs exhibit scaling exponents between approximately 0.45 and 0.65 [10]. It has been noted that the Gaussian chain model tends to overestimate the R value interpreted from FRET efficiency when the specific IDP is closer in behavior to an excluded volume chain [152].

Self-avoiding walk. A more general polymer model other than the Gaussian chain model is the self-avoiding walk (a polymer which cannot cross itself) model, in which the distance distribution $P(r)$ can be adjusted according to the scaling exponent, referred to as a SAW- ν model, and the $P(r)$ can be written as

$$P(r, \nu) = A \frac{4\pi}{R} \left(\frac{r}{R}\right)^{2+g} \exp\left[-B\left(\frac{r}{R}\right)^\delta\right]$$

where R is the root mean squared distance, A and B are obtained from the conditions $1 = \int_0^\infty P(r)dr$ and $R^2 = \int_0^\infty r^2 P(r)dr$, and the exponents are given by $g \approx (\gamma - 1)/\nu$ [153], $\gamma = 1.1615$ [154], and $\delta = \frac{1}{1-\nu}$ [155]. Then the scaling exponent ν can be obtained by minimizing $|\int E(r)P(r, \nu)dr - E_{expt}|$ by the restraint of $R \propto N^\nu$. We show in Figure 2B, with provided FRET efficiency of 0.2 of a 100 amino-acid peptide, that the $P(r)$ reconstructed using a Gaussian chain and the SAW- ν model can be quite different, suggesting such data analysis is model-dependent. The results from the SAW- ν model have been in close agreement with the all-atom explicit solvent simulations [107], which are usually computationally demanding to generate. The SAW- ν model can also be applied to other experimental methods to provide the distance between two specific amino acids. For instance, for PET and PRE, one can easily replace the $E(r)$ in the previous minimization with the equation corresponding to the experimental signal and the distance, and then the $P(r)$ can be obtained using a different experimental method. In addition, methods that

provide R_g from SAXS data can be compared with the methods that provide distance R via the relation between R_g and R [156],

$$\lambda = \frac{R^2}{R_g^2} = \frac{2(\gamma + 2\nu)(\gamma + 2\nu + 1)}{\gamma(\gamma + 1)}$$

where $\gamma = 1.1615$ [154]. SAXS can be analyzed similarly with a higher-order correction factor to the original Guinier analysis [157].

Another advantage of using the SAW- ν model is that it provides the scaling exponent in addition to R . The scaling exponent sometimes tells more than just the size of an IDP. For instance, in the case of liquid–liquid phase separation, a strong correlation was found between the critical temperature of phase separation and the theta-solvent temperature at which the scaling exponent is 0.5 [158]. It is important to note that the scaling exponent is only well-defined for a homopolymer, and the polymer models discussed assume that the IDP being studied is a homopolymer. This assumption is acceptable for some IDPs with low-complexity sequences and weak nonspecific interactions. However, growing evidence suggests specific IDPs exhibit transient interactions between pairs of amino acids [159,160]. Further work may be required, such as incorporating a new term into the current polymer model or using more sophisticated simulation models.

3.3. Molecular Simulations and Modeling Methods

Computational simulation and modeling techniques require experimental data for parameterization and calibration and can be computationally demanding. However, once established, these techniques can be applied to a wide range of systems beyond those used for parameterization. Techniques include all-atom explicit or implicit solvent simulations and coarse-grained modeling, which differ in the level of detail they provide for amino acids and water molecules. Choosing the proper simulation method requires finding a balance between detail and feasibility. All simulations rely on experimental data for parameterization or validation of results. Low-resolution models typically have fewer free parameters and require less experimental input. This approach may be appealing due to their physical intuition for understanding the underlying mechanisms. However, they may lack the detail to capture experimental measurements accurately. Higher-resolution models have more free parameters and thus require more experimental data, but they may not be easily transferable to new proteins without verification. Other approaches, such as Rosetta [161] and AlphaFold [142], can be used to model disordered regions that may be partially structured but not discussed here.

All-atom simulations. All-atom explicit solvent simulations offer the highest resolution and may be able to describe specific residue interactions that can be lost in lower-resolution coarse-grained models, such as hydrogen bonds, salt bridge, cation- π , π - π interactions, and general hydrophobic/van der Waals interactions [162]. Since water molecules are explicitly represented, all-atom explicit solvent simulations can also lead naturally into discussing the dynamics of IDPs rather than just the averaging equilibrium properties [163]. For instance, end-to-end chain relaxation time from all-atom simulations have been found in close agreement with that estimated from the FRET experiment [164]. However, one major challenge of using all-atom explicit solvent simulations is the accuracy of the force field [165–169]. Since an IDP lacks a nonlocal tertiary structure, this enlarges minor inaccuracies of local secondary structure preference and amino acid interactions of old force fields. Recent attempts to improve the accuracy of all-atom force fields rely primarily on implementing better dihedral potentials for reproducing secondary structure propensities [170] and fine-tuning protein–solvent interactions for capturing the sizes of IDPs [165]. Another option is to use implicit instead of explicit solvent [171]. Implicit solvent can be problematic when simulating interactions between charged amino acids at physiological ionic strength. This issue can be solved by introducing explicit ions such as the ABSINTH model [172,173],

which can be a good alternative between an all-atom explicit solvent model and more reduced coarse-grained models.

However, all-atom models are challenging for simulating more complex phenomena with more than one chain in the simulation, such as folding upon binding [174], liquid–liquid phase separation [162], or aggregation [175] due to the high computational cost for obtaining trajectories with sufficient time scales. There can be a few possible ways to overcome sampling difficulties. One option is the use of advanced sampling methods. Replica exchange molecular dynamics (REMD) can be applied to IDPs [176,177] and have been applied to the p53 disordered region [178]. Despite the high computational demands, this REMD method can be combined with other advanced sampling techniques, such as Gaussian-accelerated molecular dynamics (GaMD) [179,180], to enhance its capabilities further. Combining REMD and GaMD has been applied for the ER disordered region [6]. Collective-variable-based methods are commonly applied to folded proteins [181–183]; however, their application to IDPs remains to be seen due to the lack of obvious collective variables for IDP dynamics. Other approaches to accelerate simulations include using specialized supercomputers such as Anton [184] or implementing GPU-assisted versions of molecular dynamics packages [185–187].

Coarse-grained simulations. Coarse-grained (CG) models further reduce the complexity of amino acids in addition to just implicit solvents. Resolution varies greatly across CG models according to their intended use, from several CG beads for each residue (e.g., AWSEM [188–190], and flexible-meccano [191]), to one CG bead per residue, to one CG bead for the entire domain. One needs to choose an appropriate resolution depending on the problem of interest. A model with a resolution of one bead per residue could be a good balance between reducing computational cost and achieving amino acid specificity. Here we briefly describe one example, the HPS model [192]. There are three different types of interactions of local bonded interactions, electrostatics, and short-range pairwise interactions. The electrostatic interactions are modeled using a Coulombic term with Debye–Hückle electrostatic screening [193] to account for salt concentrations. The short-range pairwise interactions account for protein–protein and protein–solvent interactions and are based on the amino acid hydrophathy scale [99]. In the current HPS model, the Ashbaugh–Hatch functional form is used for the short-range pairwise interactions [194], but other functional forms can be used to consider the nonbonded interactions in addition to the electrostatic interactions between charged amino acids [195–198]. The overall interaction strength of this pairwise interaction term and amino-acid-specific parameters (e.g., hydrophathy scales) can be optimized with the experimental data of IDPs [190,199–201]. This term can also be temperature-dependent on accounting for the upper and lower critical solution temperatures [202] and salt-dependent to account for the salting-out effect at high salt concentrations [203]. Additional angle and dihedral terms can be introduced to capture the residue-specific secondary structure propensities of the chain [204,205]. The entire framework is flexible and easy to re-optimize with growing experimental measurements [190,199–201] and can be extended to biomolecules such as nucleic acids [206]. The HPS model lacks specific interactions such as hydrogen bond, salt bridge, cation- π , and sp^2/π interactions and often underestimates specific strong interactions that might exist in a particular IDP. However, this model is usually sufficient to capture interactions between charged amino acids. As shown in Figure 2C, the HPS model can correctly capture the attractive interactions (blue in the scaling exponent map) between charged amino acids within the N-terminal region of the disordered E-cadherin protein. These interactions lead to salt-induced expansion of the first 40 amino acids in contrast to the salt-induced collapse of the other regions of the protein seen in the FRET measurement.

3.4. Computational Strategies for Combining Multiple Experimental Measurements

Simulation models that are parameterized with experimental data are often considered transferable. However, when applied to a new system of interest, they may not always match the latest experimental data, requiring further improvement. Re-optimizing the

model with new experimental data is a straightforward solution, but this is often performed with CG models due to fewer built-in free parameters. Optimizing all-atom models to match a new set of experimental data can be time-consuming. Nonetheless, two alternatives include biased simulations with experimental data as restraints [207,208] and ensemble fitting that reweights the conformations of existing ensembles to best fit experimental data [152,209–214]. As integrative biophysics approaches are emerging [215], both methods are critical for IDP characterization by integrating these various experimental inputs, as depicted in Figure 3.



Figure 3. A schematic diagram for integrative biophysics combining various biophysical datasets. Complementary restraints from experimental studies of small-angle X-ray scattering, site-specific solvent accessibility, and various NMR techniques, as well as computations, are fed into ensemble-fitting machinery to generate a comprehensive picture for the ensemble structures of highly flexible biomolecules such as intrinsically disordered proteins. Reprinted with permission from Ref. [215].

Central to integrative biophysics is the development of “bridges” and “connectors” between computations and experiments. These tools allow experimental measurements to be calculated from the conformations of IDPs obtained using computational methods. Linking to SAXS data includes model-free coarse-grained computing [216–218] and atomistic-level modeling [219,220]. HDX and HRPf analysis mainly utilize protection factor analyses [18,41] to connect the solvent accessibility surface area. For NMR measurements, PRE data can be analyzed via ensemble averaging over inverse sixth power of distances [72], while solvent-PRE data can be analyzed via grid-based surface volume calculations [81,221,222]. FRET efficiencies can be calculated from a distance between the two labels using the Förster equation [47], and PET rates can be estimated using an exponential function to the distance through all-atom modeling [53]. DEER spectra are typically converted to distance distributions before being applied to computational methods [50,223–225]. Such tools play a critical role in advancing integrative data analysis.

Experiment-restrained simulations. Experiment-restrained simulation methods have succeeded in exploring new conformations or leading simulations toward conformational changes of interest. However, unlike folded proteins, these methods are often not straightforward for IDPs. IDP measurements are the results of ensemble-averaged features, making it difficult to determine how to design the biasing potential for simulations. This ambiguity makes biased IDP simulations rely on the time-consuming processes of replica averag-

ing [208], maximizing entropy, and extensive iterations [226–229]. Examples include modeling strategies with experimental restraints from SAXS [230,231], DEER [232], HRPF [233], FRET [234], and NMR observables [207,208,235,236]. Additionally, the overall results are influenced by the accuracy of the physics-based model used. Improvements in all-atom force fields and growing sources of experimental data are expected to alleviate some of these concerns.

Ensemble fitting. As an alternative approach, ensemble fitting directly incorporates experimental bias into ranking and scoring candidate structures obtained from computations. This ensemble approach is achieved by post-processing an ensemble of these putative conformations, where weights are assigned to individual conformations and then adjusted to best fit experimental observables. One prolific example of ensemble fitting is the combination of SAXS data with various docking and modeling algorithms. This SAXS-assisted method has been implemented and applied to various research areas, including protein–protein interactions [237–243], high-order structures [213,214,244–247], protein dynamics [248–257], RNA dynamics [258–261], and the study of IDPs [6,262–265].

Ensemble fitting is frequently combined with multiple experimental data types to obtain a complete picture of protein behavior. By combining SAXS or FRET data about global conformations with site-specific information on solvent accessibility (e.g., HRPF and DFCS) or NMR distance data (e.g., PRE), insight has been gained into the behavior of IDPs [72,212,214,264,266–269]. As shown in our recent publication [42], the amino acid contact map of the ER disordered domain can be obtained through ensemble fitting of data from SAXS, HRPF, and DFCS, which reveals previously unknown nonlocal contacts.

It is still an open question regarding how to best proceed with ensemble fitting to meet all experimental measurements. Two different strategies have been employed to prepare the basis set of initial conformations for fitting. One involves using a large pool of candidate structures for maximum entropy analysis [270], while the other requires minimizing the number of conformational clusters before fitting [250]. The first strategy involves handling a large number of structures and applying the maximum entropy principle to prevent overfitting, recognizing that some structures share similar experimental observables. The second strategy conducts conformational clustering before fitting and requires well-defined collective variables that can separate the pool of structures into distinct clusters, serving as a basis set of conformations for ensemble fitting. A combined approach has been attempted using a modest number of conformations and the maximum entropy method [6,271] that has successfully made predictions that were subsequently validated. If the initial pool of structures captures the majority of target conformations, then conformational clustering based on experimental observables before fitting could be considered a method of choice instead of imposing a statistical bias on the fly, where a minimum number of distinct conformations is utilized as a de facto basis set to avoid a potential issue of double-counting in ensemble fitting (i.e., using both a non-equal probability weight for individual conformations and an entropy penalty for the overall density of conformations, simultaneously). Nevertheless, this assertion requires further investigation in future studies.

4. Targeting Protein Intrinsic Disorder as a New Frontier of Drug Discovery

IDPs are emerging as a promising class of targets for small molecule binding [272–274]. Notable examples of these ligands include 10058-F4/sAJM589 targeting the transcription factor c-Myc [275], EGCG against p53-NTD [276], Fasudil against α -synuclein [277], 10074-G5 against A β 42 [278], SJ403 against p27-Kip1 [279], EPI against AR-NTD [280], CLR01 as a molecular tweezer against the disordered protein–protein interface of Cdc25C [281], and NSC635437 against the fusion oncoprotein EWS-FLI1 [282], some of which are depicted in Figure 4.

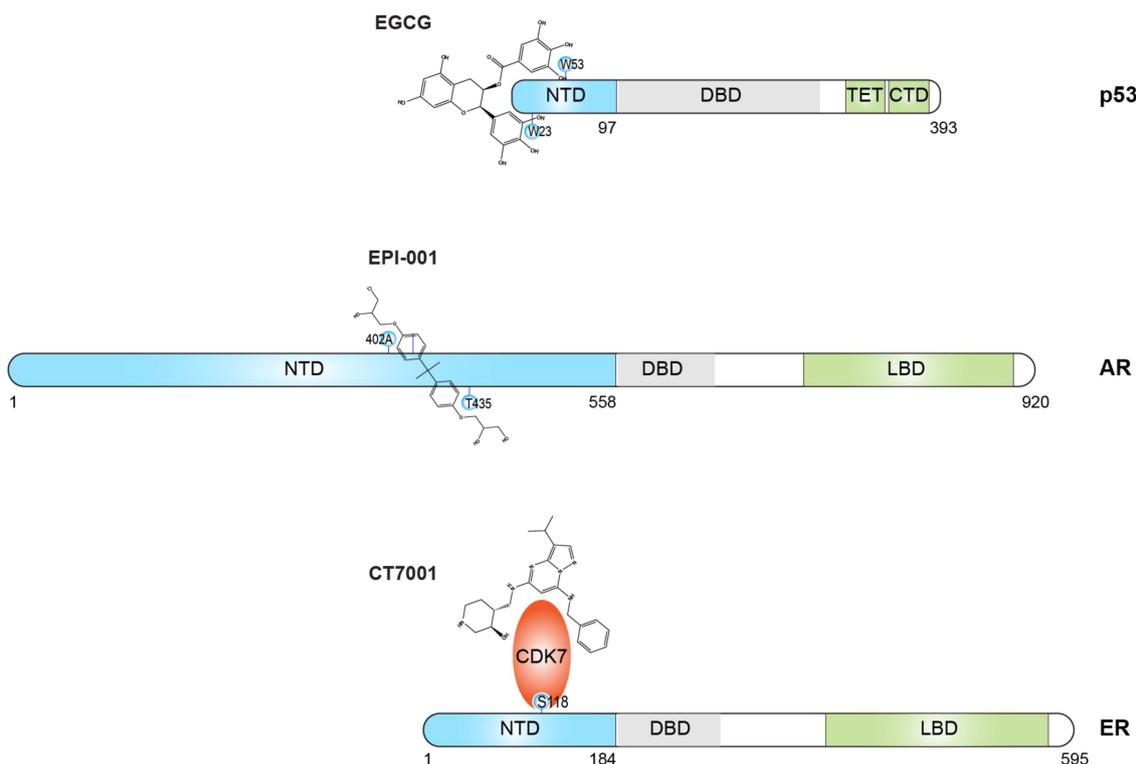


Figure 4. Small molecule targeting against protein intrinsic disorder as a new frontier in drug discovery. NTD: N-terminal domain; DBD: DNA-binding domain; LBD: ligand-binding domain; TET: tetramerization domain; CTD: C-terminal domain. AR: androgen receptor; ER: estrogen receptor. EGCG, a compound found in green tea, has been shown to interact with two specific amino acids of p53-NTD, W23 and W53 [276]; EPI-001, isolated from marine sponges [283], has been found to interact with A402 and T435 residues of AR-NTD [284]. CT7001 has been demonstrated to bind the ATP-binding site of CDK7 (PDB id: 7B5O [91] and 6Z4X [285]), a serine/threonine kinase that can phosphorylate Ser118 within the ER-NTD [286].

p53-NTD has been extensively studied using various biophysical techniques, including SAXS [287], PRE [288], solvent-PRE [84], and PET coupled with fluorescence correlation spectroscopy [54], as well as computations [178,287]. Knowledge accumulated over the years has not only aided in understanding the binding mechanism between EGCG and p53-NTD [276], but also provided the molecular basis for finding new binders.

AR-NTD was among the first IDPs selected as a therapeutic target for drug development [280,289]. Unlike p53-NTD, small molecule binders were identified before biophysical data of AR-NTD binding became available. EPI-001, one of the early compounds, was isolated from marine sponges and found to inhibit AR-NTD transcriptional activity [283,290]. Subsequent binding characterization included chemical shift perturbation analysis [288] as well as computational modeling [291]. Despite the wealth of biophysical data available for p53-NTD, a comprehensive structural ensemble of p53-NTD (either in the absence or presence of EGCG) is not currently available that explicitly accounts for the diverse experimental restraints. Encouragingly, chemical shift perturbations have identified specific amino acids that are well separated in their primary amino acids for ligand–protein interactions, as indicated in Figure 4.

Studies targeting ER-NTD have lagged behind AR-NTD, and there is currently no small molecule inhibitor that directly binds ER-NTD. While ER-NTD and AR-NTD belong to the same nuclear receptor superfamily, ER-NTD is shorter (184 amino acids) than AR-NTD (558 amino acids) [292]. Despite being shorter than AR-NTD, it is longer than p53-NTD (97 amino acids) [293], as illustrated in Figure 4. Counterintuitively, the structural information for ER-NTD is limited [294,295] compared to AR-NTD and p53-NTD, whose

chemical shifts have been mostly assigned. Maintaining protein stability and homogeneity has posed difficulties in conducting NMR studies on ER-NTD. However, non-NMR studies have provided early insights into its inner workings, including SAXS, HRPF, and DFCS data and computational studies (5, 28).

While no small molecule directly targets ER-NTD, efforts are underway to develop small molecule inhibitors that target its coregulatory proteins. As illustrated in Figure 4, CDK7 is an upstream protein kinase that activates ER-NTD by phosphorylating serine at position 118 [286], and CDK7 inhibitors have been developed to reduce ER-NTD activity [296,297]. One such inhibitor, CT7001, is currently undergoing clinical trials for the therapeutics of ER-positive breast cancer (phase 2) and castrate-resistant prostate cancer (phase 1 as of February 2023) [298]. However, the multifaceted role of CDK7 as an activation initiator for multiple proteins involved in transcription and cell cycle regulation can lead to cellular toxicity and off-target effects. As we gain more knowledge about ER-NTD at the molecular level, a more direct strategy is approaching to target ER-NTD for the discovery of novel binders.

These examples represent one approach of targeting protein intrinsic disorder at specific protein regions or post-translational modifications to shift the equilibrium of disordered conformations. Another strategy involves using small molecules or peptides that mimic binding partner proteins to alter the disordered protein–protein interface [299–302]. Notably, a significant portion of such protein–protein interactions is mediated by so-called short linear motifs (SLiMs), commonly found within disordered regions [303]. Typically, each SLiM is a small polypeptide stretch consisting of 3 to 15 residues [122,125] and can be grouped into six distinctive classes via the eukaryotic linear motif (ELM) database [304,305], including the LIG class for covering the function of protein interactions with ligand proteins, MOD for post-translational modifications such as phosphorylation, CLV for proteolytic cleavage, TRG for subcellular targeting, DEG for degradation with protein polyubiquitylation, and DOC for classic docking of enzyme recruitment [306]. The classification of SLiMs into distinct classes provides a comprehensive understanding of the multifaceted functions that many SLiMs can carry out, even within the same disordered protein sequence. For instance, the discrimination between LIG and MOD classes is an excellent example of how protein phosphorylation sites are distinct from protein–protein interaction sites; the fact that phosphorylation sites are categorized as MOD motifs and not LIG motifs indicates that they may not directly participate in protein–protein interactions. This distinction has been demonstrated through the analysis of ELM search results of, e.g., p53-NTD, AR-NTD, and ER-NTD, where their SLiMs spread over the amino acid sequence with little overlap. Given that IDPs often engage in promiscuous interactions with a vast array of partner proteins [1,307], a thorough investigation of the IDP of interest is important to identify whether a particular SLiM dominates over or coordinates with others in order to fully understand its potential as a drug target.

5. Perspectives: Chaotic Life of Protein Intrinsic Disorder at a Crossroads

Intrinsic disorder in proteins imposes difficulties for biophysical studies and challenges the conventional structure–function paradigm learned from folded proteins. As IDPs are critical in many biological processes, such as transcription and signaling, understanding the inner workings of IDPs requires innovative use of available biophysical tools and a proactive approach combining complementary techniques. The limited yet growing knowledge provides a new perspective on the IDPs' sequence–structure–function relationship, thereby allowing for the study of protein intrinsic disorder to find new binders against important therapeutic targets.

Author Contributions: Conceptualization, S.L., S.W., W.Z. and S.Y.; writing, S.L., S.W., W.Z. and S.Y.; supervision, W.Z. and S.Y.; funding acquisition, W.Z. and S.Y. All authors have read and agreed to the published version of the manuscript.

Funding: S.Y. received support from the National Institutes of Health (R01GM114056 and R03CA241977) as well as the Case Cancer accelerator award via the NCI (P30CA043703). W.Z. acknowledges the support from the National Science Foundation (MCB-2015030) and the National Institutes of Health (R35GM146814).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank the three anonymous reviewers for providing invaluable feedback that improved the scope and depth of this brief review.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Wright, P.E.; Dyson, H.J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 18–29. [[CrossRef](#)] [[PubMed](#)]
2. van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R.J.; Daughdrill, G.W.; Dunker, A.K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D.T.; et al. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **2014**, *114*, 6589–6631. [[CrossRef](#)] [[PubMed](#)]
3. Bernado, P.; Svergun, D.I. Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Mol. Biosyst.* **2012**, *8*, 151–167. [[CrossRef](#)] [[PubMed](#)]
4. Kohn, J.E.; Millett, I.S.; Jacob, J.; Zagrovic, B.; Dillon, T.M.; Cingel, N.; Dothager, R.S.; Seifert, S.; Thiyagarajan, P.; Sosnick, T.R.; et al. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12491–12496. [[CrossRef](#)]
5. Flory, P.J. *Principles of Polymer Chemistry*; Cornell University Press: Ithaca, NY, USA; London, UK, 1953.
6. Peng, Y.; Cao, S.; Kiselar, J.; Xiao, X.; Du, Z.; Hsieh, A.; Ko, S.; Chen, Y.; Agrawal, P.; Zheng, W.; et al. A metastable contact and structural disorder in the estrogen receptor transactivation domain. *Structure* **2019**, *27*, 229–240.e4. [[CrossRef](#)]
7. Belorusova, A.; Osz, J.; Petoukhov, M.V.; Peluso-Iltis, C.; Kieffer, B.; Svergun, D.I.; Rochel, N. Solution behavior of the intrinsically disordered n-terminal domain of retinoid x receptor alpha in the context of the full-length protein. *Biochemistry* **2016**, *55*, 1741–1748. [[CrossRef](#)]
8. Johansen, D.; Trewheella, J.; Goldenberg, D.P. Fractal dimension of an intrinsically disordered protein: Small-angle X-ray scattering and computational study of the bacteriophage lambda n protein. *Protein Sci.* **2011**, *20*, 1955–1970. [[CrossRef](#)]
9. Riback, J.A.; Bowman, M.A.; Zmyslowski, A.M.; Knoverek, C.R.; Jumper, J.M.; Hinshaw, J.R.; Kaye, E.B.; Freed, K.F.; Clark, P.L.; Sosnick, T.R. Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water. *Science* **2017**, *358*, 238–241. [[CrossRef](#)]
10. Hofmann, H.; Soranno, A.; Borgia, A.; Gast, K.; Nettels, D.; Schuler, B. Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 16155–16160. [[CrossRef](#)]
11. Koch, M.H.J.; Vachette, P.; Svergun, D.I. Small-angle scattering: A view on the properties, structures and structural changes of biological macromolecules in solution. *Q. Rev. Biophys.* **2003**, *36*, 147–227. [[CrossRef](#)]
12. Putnam, C.D.; Hammel, M.; Hura, G.L.; Tainer, J.A. X-ray solution scattering. (saxs) combined with crystallography and computation: Defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **2007**, *40*, 191–285. [[CrossRef](#)]
13. Chandler, D. *Introduction to Modern Statistical Mechanics*; Oxford University Press: New York, NY, USA, 1987.
14. Svergun, D.I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **1992**, *25*, 495–503. [[CrossRef](#)]
15. Yang, S. Methods for saxs-based structure determination of biomolecular complexes. *Adv. Mater.* **2014**, *26*, 7902–7910. [[CrossRef](#)]
16. Perez, J.; Nishino, Y. Advances in X-ray scattering: From solution saxs to achievements with coherent beams. *Curr. Opin. Struct. Biol.* **2012**, *22*, 670–678. [[CrossRef](#)]
17. Englander, S.W.; Kallenbach, N.R. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* **1983**, *16*, 521–655. [[CrossRef](#)]
18. Bai, Y.; Milne, J.S.; Mayne, L.; Englander, S.W. Primary structure effects on peptide group hydrogen exchange. *Proteins* **1993**, *17*, 75–86. [[CrossRef](#)]
19. Goswami, D.; Devarakonda, S.; Chalmers, M.J.; Pascal, B.D.; Spiegelman, B.M.; Griffin, P.R. Time window expansion for hdx analysis of an intrinsically disordered protein. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1584–1592. [[CrossRef](#)]
20. Xu, G.; Chance, M.R. Hydroxyl radical-mediated modification of proteins as probes for structural proteomics. *Chem. Rev.* **2007**, *107*, 3514–3543. [[CrossRef](#)]

21. Ralston, C.Y.; Sharp, J.S. Structural investigation of therapeutic antibodies using hydroxyl radical protein footprinting methods. *Antibodies* **2022**, *11*, 71. [[CrossRef](#)]
22. Hambly, D.M.; Gross, M.L. Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 2057–2063. [[CrossRef](#)]
23. Johnson, D.T.; Jones, L.M. Hydroxyl radical protein footprinting for analysis of higher order structure. *Trends Biochem. Sci.* **2022**, *47*, 989–991. [[CrossRef](#)] [[PubMed](#)]
24. McKenzie-Coe, A.; Montes, N.S.; Jones, L.M. Hydroxyl radical protein footprinting: A mass spectrometry-based structural method for studying the higher order structure of proteins. *Chem. Rev.* **2022**, *122*, 7532–7561. [[CrossRef](#)] [[PubMed](#)]
25. Sharp, J.S.; Chea, E.E.; Misra, S.K.; Orlando, R.; Popov, M.; Egan, R.W.; Holman, D.; Weinberger, S.R. Flash oxidation. (fox) system: A novel laser-free fast photochemical oxidation protein footprinting platform. *J. Am. Soc. Mass Spectrom.* **2021**, *32*, 1601–1609. [[CrossRef](#)] [[PubMed](#)]
26. Kitevski-LeBlanc, J.L.; Prosser, R.S. Current applications of 19f nmr to studies of protein structure and dynamics. *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, *62*, 1–33. [[CrossRef](#)]
27. Chrisman, I.M.; Nemetchev, M.D.; de Vera, I.M.S.; Shang, J.; Heidari, Z.; Long, Y.; Reyes-Caballero, H.; Galindo-Murillo, R.; Cheatham, T.E., 3rd; Blayo, A.L.; et al. Defining a conformational ensemble that directs activation of ppargamma. *Nat. Commun.* **2018**, *9*, 1794. [[CrossRef](#)]
28. Chance, M.R.; Farquhar, E.R.; Yang, S.; Lodowski, D.T.; Kiselar, J. Protein footprinting: Auxiliary engine to power the structural biology revolution. *J. Mol. Biol.* **2020**, *432*, 2973–2984. [[CrossRef](#)]
29. Liu, X.R.; Zhang, M.M.; Gross, M.L. Mass spectrometry-based protein footprinting for higher-order structure analysis: Fundamentals and applications. *Chem. Rev.* **2020**, *120*, 4355–4454. [[CrossRef](#)]
30. Liu, J.J.; Horst, R.; Katritch, V.; Stevens, R.C.; Wuthrich, K. Biased signaling pathways in beta(2)-adrenergic receptor characterized by f-19-nmr. *Science* **2012**, *335*, 1106–1110. [[CrossRef](#)]
31. Didenko, T.; Liu, J.J.; Horst, R.; Stevens, R.C.; Wuthrich, K. Fluorine-19 nmr of integral membrane proteins illustrated with studies of gpcrs. *Curr. Opin. Struct. Biol.* **2013**, *23*, 740–747. [[CrossRef](#)]
32. Matei, E.; Gronenborn, A.M. (19)f paramagnetic relaxation enhancement: A valuable tool for distance measurements in proteins. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 150–154. [[CrossRef](#)]
33. Evanics, F.; Kitevski, J.L.; Bezonova, I.; Forman-Kay, J.; Prosser, R.S. F-19 nmr studies of solvent exposure and peptide binding to an sh3 domain. *BBA Gen. Subjects* **2007**, *1770*, 221–230. [[CrossRef](#)]
34. Gerig, J.T. Fluorine nmr of proteins. *Prog. Nucl. Mag. Res. Sp.* **1994**, *26*, 293–370. [[CrossRef](#)]
35. Kaur, P.; Kiselar, J.; Yang, S.; Chance, M.R. Quantitative protein topography analysis and high-resolution structure prediction using hydroxyl radical labeling and tandem-ion mass spectrometry. (ms). *Mol. Cell Proteom.* **2015**, *14*, 1159–1168. [[CrossRef](#)]
36. Kiselar, J.; Chance, M.R. High-resolution hydroxyl radical protein footprinting: Biophysics tool for drug discovery. *Annu. Rev. Biophys.* **2018**, *47*, 315–333. [[CrossRef](#)]
37. Gupta, S.; Chen, Y.; Petzold, C.J.; DePonte, D.P.; Ralston, C.Y. Development of container free sample exposure for synchrotron X-ray footprinting. *Anal. Chem.* **2020**, *92*, 1565–1573. [[CrossRef](#)]
38. Shcherbakova, I.; Mitra, S.; Beer, R.H.; Brenowitz, M. Fast fenton footprinting: A laboratory-based method for the time-resolved analysis of DNA, rna and proteins. *Nucleic Acids Res.* **2006**, *34*, e48. [[CrossRef](#)]
39. Gupta, S.; Celestre, R.; Petzold, C.J.; Chance, M.R.; Ralston, C. Development of a microsecond X-ray protein footprinting facility at the advanced light source. *J. Synchrotron Radiat.* **2014**, *21*, 690–699. [[CrossRef](#)]
40. Chen, J.; Rempel, D.L.; Gross, M.L. Temperature jump and fast photochemical oxidation probe submillisecond protein folding. *J. Am. Chem. Soc.* **2010**, *132*, 15502–15504. [[CrossRef](#)]
41. Huang, W.; Ravikumar, K.M.; Chance, M.R.; Yang, S. Quantitative mapping of protein structure by hydroxyl radical footprinting-mediated structural mass spectrometry: A protection factor analysis. *Biophys. J.* **2015**, *108*, 107–115. [[CrossRef](#)]
42. Zheng, W.; Du, Z.; Ko, S.B.; Wickramasinghe, N.P.; Yang, S. Incorporation of d(2)o-induced fluorine chemical shift perturbations into ensemble-structure characterization of the eralpha disordered region. *J. Phys. Chem. B* **2022**, *126*, 9176–9186. [[CrossRef](#)]
43. Schuler, B.; Soranno, A.; Hofmann, H.; Nettels, D. Single-molecule fret spectroscopy and the polymer physics of unfolded and intrinsically disordered proteins. *Annu. Rev. Biophys.* **2016**, *45*, 207–231. [[CrossRef](#)]
44. Drescher, M. Epr in protein science: Intrinsically disordered proteins. *Top. Curr. Chem.* **2012**, *321*, 91–119.
45. Schiemann, O.; Heubach, C.A.; Abdullin, D.; Ackermann, K.; Azarkh, M.; Bagryanskaya, E.G.; Drescher, M.; Endeward, B.; Freed, J.H.; Galazzo, L.; et al. Benchmark test and guidelines for deer/peldor experiments on nitroxide-labeled biomolecules. *J. Am. Chem. Soc.* **2021**, *143*, 17875–17890. [[CrossRef](#)]
46. Lapidus, L.J.; Eaton, W.A.; Hofrichter, J. Measuring the rate of intramolecular contact formation in polypeptides. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7220–7225. [[CrossRef](#)] [[PubMed](#)]
47. Förster, T. Zwischenmolekulare energiewanderung und fluoreszenz. *Ann. Phys.* **1948**, *6*, 55–75. [[CrossRef](#)]
48. Trexler, A.J.; Rhoades, E. Single molecule characterization of alpha-synuclein in aggregation-prone states. *Biophys. J.* **2010**, *99*, 3048–3055. [[CrossRef](#)] [[PubMed](#)]
49. Wiggers, F.; Wohl, S.; Dubovetskiy, A.; Rosenblum, G.; Zheng, W.; Hofmann, H. Diffusion of a disordered protein on its folded ligand. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2106690118. [[CrossRef](#)]

50. Chiang, Y.W.; Borbat, P.P.; Freed, J.H. The determination of pair distance distributions by pulsed esr using tikhonov regularization. *J. Magn. Reson.* **2005**, *172*, 279–295. [[CrossRef](#)]
51. Buscaglia, M.; Kubelka, J.; Eaton, W.A.; Hofrichter, J. Determination of ultrafast protein folding rates from loop formation dynamics. *J. Mol. Biol.* **2005**, *347*, 657–664. [[CrossRef](#)]
52. Sizemore, S.M.; Cope, S.M.; Roy, A.; Ghirlanda, G.; Vaiana, S.M. Slow internal dynamics and charge expansion in the disordered protein cgrp: A comparison with amyl.lin. *Biophys. J.* **2015**, *109*, 1038–1048. [[CrossRef](#)]
53. Zerze, G.H.; Mittal, J.; Best, R.B. Diffusive dynamics of contact formation in disordered polypeptides. *Phys. Rev. Lett.* **2016**, *116*, 068102. [[CrossRef](#)]
54. Lum, J.K.; Neuweiler, H.; Fersht, A.R. Long-range modulation of chain motions within the intrinsically disordered transactivation domain of tumor suppressor p53. *J. Am. Chem. Soc.* **2012**, *134*, 1617–1622. [[CrossRef](#)]
55. Dyson, H.J.; Wright, P.E. Nmr illuminates intrinsic disorder. *Curr. Opin. Struct. Biol.* **2021**, *70*, 44–52. [[CrossRef](#)]
56. Prestel, A.; Bugge, K.; Staby, L.; Hendus-Altenburger, R.; Kragelund, B.B. Characterization of dynamic idp complexes by nmr spectroscopy. *Methods Enzymol.* **2018**, *611*, 193–226.
57. Williamson, M.P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16. [[CrossRef](#)] [[PubMed](#)]
58. Konrat, R. Nmr contributions to structural dynamics studies of intrinsically disordered proteins. *J. Magn. Reson.* **2014**, *241*, 74–85. [[CrossRef](#)] [[PubMed](#)]
59. Hansen, D.F.; Feng, H.; Zhou, Z.; Bai, Y.; Kay, L.E. Selective characterization of microsecond motions in proteins by nmr relaxation. *J. Am. Chem. Soc.* **2009**, *131*, 16257–16265. [[CrossRef](#)] [[PubMed](#)]
60. Kay, L.E.; Torchia, D.A.; Bax, A. Backbone dynamics of proteins as studied by 15n inverse detected heteronuclear nmr spectroscopy: Application to staphylococcal nuclease. *Biochemistry* **1989**, *28*, 8972–8979. [[CrossRef](#)]
61. Yuwen, T.; Skrynnikov, N.R. Proton-decoupled cpmg: A better experiment for measuring. (15)n r2 relaxation in disordered proteins. *J. Magn. Reson.* **2014**, *241*, 155–169. [[CrossRef](#)]
62. Klein-Seetharaman, J.; Oikawa, M.; Grimshaw, S.B.; Wirmer, J.; Duchardt, E.; Ueda, T.; Imoto, T.; Smith, L.J.; Dobson, C.M.; Schwalbe, H. Long-range interactions within a nonnative protein. *Science* **2002**, *295*, 1719–1722. [[CrossRef](#)]
63. Martin, E.W.; Holehouse, A.S.; Peran, I.; Farag, M.; Incicco, J.J.; Bremer, A.; Grace, C.R.; Soranno, A.; Pappu, R.V.; Mittag, T. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **2020**, *367*, 694–699. [[CrossRef](#)]
64. Yu, L.; Bruschweiler, R. Quantitative prediction of ensemble dynamics, shapes and contact propensities of intrinsically disordered proteins. *PLoS Comput. Biol.* **2022**, *18*, e1010036. [[CrossRef](#)] [[PubMed](#)]
65. Clore, G.M. Practical aspects of paramagnetic relaxation enhancement in biological macromolecules. *Methods Enzymol.* **2015**, *564*, 485–497.
66. Battiste, J.L.; Wagner, G. Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* **2000**, *39*, 5355–5365. [[CrossRef](#)]
67. Sjodt, M.; Clubb, R.T. Nitroxide labeling of proteins and the determination of paramagnetic relaxation derived distance restraints for nmr studies. *Bio. Protoc.* **2017**, *7*, e2207. [[CrossRef](#)]
68. Clore, G.M.; Gronenborn, A.M. Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. *Crit. Rev. Biochem. Mol. Biol.* **1989**, *24*, 479–564. [[CrossRef](#)]
69. Iwahara, J.; Tang, C.; Marius Clore, G. Practical aspects of. (1)h transverse paramagnetic relaxation enhancement measurements on macromolecules. *J. Magn. Reson.* **2007**, *184*, 185–195. [[CrossRef](#)] [[PubMed](#)]
70. Tang, C.; Iwahara, J.; Clore, G.M. Visualization of transient encounter complexes in protein-protein association. *Nature* **2006**, *444*, 383–386. [[CrossRef](#)]
71. Lietzow, M.A.; Jamin, M.; Dyson, H.J.; Wright, P.E. Mapping long-range contacts in a highly unfolded protein. *J. Mol. Biol.* **2002**, *322*, 655–662. [[CrossRef](#)] [[PubMed](#)]
72. Salmon, L.; Nodet, G.; Ozenne, V.; Yin, G.; Jensen, M.R.; Zweckstetter, M.; Blackledge, M. Nmr characterization of long-range order in intrinsically disordered proteins. *J. Am. Chem. Soc.* **2010**, *132*, 8407–8418. [[CrossRef](#)]
73. Senicourt, L.; le Maire, A.; Allemand, F.; Carvalho, J.E.; Guee, L.; Germain, P.; Schubert, M.; Bernado, P.; Bourguet, W.; Sibille, N. Structural insights into the interaction of the intrinsically disordered co-activator tif2 with retinoic acid receptor heterodimer. (rxr/rar). *J. Mol. Biol.* **2021**, *433*, 166899. [[CrossRef](#)] [[PubMed](#)]
74. Bertocini, C.W.; Jung, Y.S.; Fernandez, C.O.; Hoyer, W.; Griesinger, C.; Jovin, T.M.; Zweckstetter, M. Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 1430–1435. [[CrossRef](#)]
75. Mittag, T.; Marsh, J.; Grishaev, A.; Orlicky, S.; Lin, H.; Sicheri, F.; Tyers, M.; Forman-Kay, J.D. Structure/function implications in a dynamic complex of the intrinsically disordered sic1 with the cdc4 subunit of an scf ubiquitin ligase. *Structure* **2010**, *18*, 494–506. [[CrossRef](#)]
76. Mosure, S.A.; Munoz-Tello, P.; Kuo, K.-T.; MacTavish, B.; Yu, X.; Scholl, D.; Williams, C.C.; Strutzenberg, T.S.; Bass, J.; Brust, R.; et al. Structural basis of interdomain communication in pparγ. *bioRxiv* **2022**. [[CrossRef](#)]

77. Kurzbach, D.; Vanas, A.; Flamm, A.G.; Tarnoczi, N.; Kontaxis, G.; Maltar-Strmecki, N.; Widder, K.; Hinderberger, D.; Konrat, R. Detection of correlated conformational fluctuations in intrinsically disordered proteins through paramagnetic relaxation interference. *Phys. Chem. Chem. Phys.* **2016**, *18*, 5753–5758. [[CrossRef](#)] [[PubMed](#)]
78. Kurzbach, D.; Beier, A.; Vanas, A.; Flamm, A.G.; Platzner, G.; Schwarz, T.C.; Konrat, R. Nmr probing and visualization of correlated structural fluctuations in intrinsically disordered proteins. *Phys. Chem. Chem. Phys.* **2017**, *19*, 10651–10656. [[CrossRef](#)]
79. Kawasaki, R.; Tate, S.I. Impact of the hereditary p301I mutation on the correlated conformational dynamics of human tau protein revealed by the paramagnetic relaxation enhancement nmr experiments. *Int. J. Mol. Sci.* **2020**, *21*, 3920. [[CrossRef](#)]
80. Hocking, H.G.; Zangger, K.; Madl, T. Studying the structure and dynamics of biomolecules by using soluble paramagnetic probes. *Chemphyschem A Eur. J. Chem. Phys. Phys. Chem.* **2013**, *14*, 3082–3094. [[CrossRef](#)]
81. Gong, Z.; Gu, X.H.; Guo, D.C.; Wang, J.; Tang, C. Protein structural ensembles visualized by solvent paramagnetic relaxation enhancement. *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 1002–1006. [[CrossRef](#)]
82. Kooshapur, H.; Schwieters, C.D.; Tjandra, N. Conformational ensemble of disordered proteins probed by solvent paramagnetic relaxation enhancement. (spre). *Angew. Chem. Int. Ed. Engl.* **2018**, *57*, 13519–13522. [[CrossRef](#)]
83. Spreitzer, E.; Usluer, S.; Madl, T. Probing surfaces in dynamic protein interactions. *J. Mol. Biol.* **2020**, *432*, 2949–2972. [[CrossRef](#)]
84. Hartlmuller, C.; Spreitzer, E.; Gobl, C.; Falsone, F.; Madl, T. Nmr characterization of solvent accessibility and transient structure in intrinsically disordered proteins. *J. Biomol. Nmr.* **2019**, *73*, 305–317. [[CrossRef](#)] [[PubMed](#)]
85. Yu, B.; Pletka, C.C.; Pettitt, B.M.; Iwahara, J. De novo determination of near-surface electrostatic potentials by NMR. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2104020118. [[CrossRef](#)]
86. Toyama, Y.; Rangadurai, A.K.; Forman-Kay, J.D.; Kay, L.E. Mapping the per-residue surface electrostatic potential of caprin1 along its phase-separation trajectory. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2210492119. [[CrossRef](#)] [[PubMed](#)]
87. Rangadurai, A.K.; Toyama, Y.; Kay, L.E. Practical considerations for the measurement of near-surface electrostatics based on solvent paramagnetic relaxation enhancements. *J. Magn. Reson.* **2023**, *349*, 107400. [[CrossRef](#)]
88. Sigler, P.B. Transcriptional activation. Acid blobs and negative noodles. *Nature* **1988**, *333*, 210–212. [[CrossRef](#)] [[PubMed](#)]
89. Struhl, K. Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. *Cell* **1987**, *49*, 295–297. [[CrossRef](#)]
90. Uversky, V.N.; Gillespie, J.R.; Fink, A.L. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* **2000**, *41*, 415–427. [[CrossRef](#)] [[PubMed](#)]
91. Borgia, A.; Borgia, M.B.; Bugge, K.; Kissling, V.M.; Heidarsson, P.O.; Fernandes, C.B.; Sottini, A.; Soranno, A.; Buholzer, K.J.; Nettels, D.; et al. Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **2018**, *555*, 61–66. [[CrossRef](#)]
92. Mao, A.H.; Crick, S.L.; Vitalis, A.; Chicoine, C.L.; Pappu, R.V. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 8183–8188. [[CrossRef](#)]
93. Huang, F.; Oldfield, C.J.; Xue, B.; Hsu, W.L.; Meng, J.; Liu, X.; Shen, L.; Romero, P.; Uversky, V.N.; Dunker, A. Improving protein order-disorder classification using charge-hydrophobicity plots. *BMC Bioinform.* **2014**, *15*, S4. [[CrossRef](#)]
94. Kapcha, L.H.; Rosicky, P.J. A simple atomic-level hydrophobicity scale reveals protein interfacial structure. *J. Mol. Biol.* **2014**, *426*, 484–498. [[CrossRef](#)]
95. Kyte, J.; Doolittle, R.F. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105–132. [[CrossRef](#)]
96. Sormanni, P.; Camilloni, C.; Fariselli, P.; Vendruscolo, M. The s2d method: Simultaneous sequence-based prediction of the statistical populations of ordered and disordered regions in proteins. *J. Mol. Biol.* **2015**, *427*, 982–996. [[CrossRef](#)]
97. Muñoz, V.; Serrano, L. Elucidating the folding problem of helical peptides using empirical parameters. *Nat. Struct. Biol.* **1994**, *1*, 399–409. [[CrossRef](#)]
98. Petersen, B.; Petersen, T.N.; Andersen, P.; Nielsen, M.; Lundegaard, C. A generic method for assignment of reliability scores applied to solvent accessibility predictions. *BMC Struct. Biol.* **2009**, *9*, 51. [[CrossRef](#)]
99. Lin, Y.; Currie, S.L.; Rosen, M.K. Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J. Biol. Chem.* **2017**, *292*, 19110–19120. [[CrossRef](#)]
100. Mateos, B.; Conrad-Billroth, C.; Schiavina, M.; Beier, A.; Kontaxis, G.; Konrat, R.; Felli, I.C.; Pierattelli, R. The ambivalent role of proline residues in an intrinsically disordered protein: From disorder promoters to compaction facilitators. *J. Mol. Biol.* **2020**, *432*, 3093–3111. [[CrossRef](#)]
101. Cohan, M.C.; Shinn, M.K.; Lalmansingh, J.M.; Pappu, R.V. Uncovering non-random binary patterns within sequences of intrinsically disordered proteins. *J. Mol. Biol.* **2022**, *434*, 167373. [[CrossRef](#)]
102. He, B.; Wang, K.; Liu, Y.; Xue, B.; Uversky, V.N.; Dunker, A.K. Predicting intrinsic disorder in proteins: An overview. *Cell Res.* **2009**, *19*, 929–949. [[CrossRef](#)]
103. Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637. [[CrossRef](#)] [[PubMed](#)]
104. Quaglia, F.; Meszaros, B.; Salladini, E.; Hatos, A.; Pancsa, R.; Chemes, L.B.; Pajkos, M.; Lazar, T.; Pena-Diaz, S.; Santos, J.; et al. Disprot in 2022: Improved quality and accessibility of protein intrinsic disorder annotation. *Nucleic Acids Res.* **2022**, *50*, D480–D487. [[CrossRef](#)] [[PubMed](#)]
105. Greber, B.J.; Remis, J.; Ali, S.; Nogales, E. 2.5 Å-resolution structure of human cdk-activating kinase bound to the clinical inhibitor icec0942. *Biophys. J.* **2021**, *120*, 677–686. [[CrossRef](#)] [[PubMed](#)]

106. Williams, C.J.; Richardson, D.C.; Richardson, J.S. The importance of residue-level filtering and the top2018 best-parts dataset of high-quality protein residues. *Protein Sci.* **2022**, *31*, 290–300. [[CrossRef](#)] [[PubMed](#)]
107. Zheng, W.; Zerze, G.H.; Borgia, A.; Mittal, J.; Schuler, B.; Best, R.B. Inferring properties of disordered chains from fret transfer efficiencies. *J. Chem. Phys.* **2018**, *148*, 123329. [[CrossRef](#)]
108. Piovesan, D.; Tabaro, F.; Micetic, I.; Necci, M.; Quaglia, F.; Oldfield, C.J.; Aspromonte, M.C.; Davey, N.E.; Davidovic, R.; Dosztanyi, Z.; et al. Disprot 7.0: A major update of the database of disordered proteins. *Nucleic Acids Res.* **2017**, *45*, D219–D227. [[CrossRef](#)]
109. Fukuchi, S.; Amemiya, T.; Sakamoto, S.; Nobe, Y.; Hosoda, K.; Kado, Y.; Murakami, S.D.; Koike, R.; Hiroaki, H.; Ota, M. Ideal in 2014 illustrates interaction networks composed of intrinsically disordered proteins and their binding partners. *Nucleic Acids Res.* **2014**, *42*, D320–D325. [[CrossRef](#)]
110. Piovesan, D.; Del Conte, A.; Clementel, D.; Monzon, A.M.; Bevilacqua, M.; Aspromonte, M.C.; Iserle, J.A.; Orti, F.E.; Marino-Buslje, C.; Tosatto, S.C.E. Mobidb: 10 years of intrinsically disordered proteins. *Nucleic Acids Res.* **2023**, *51*, D438–D444. [[CrossRef](#)]
111. Necci, M.; Piovesan, D.; Predictors, C.; DisProt, C.; Tosatto, S.C.E. Critical assessment of protein intrinsic disorder prediction. *Nat. Methods* **2021**, *18*, 472–481. [[CrossRef](#)]
112. Hanson, J.; Paliwal, K.K.; Litfin, T.; Zhou, Y. Spot-disorder2: Improved protein intrinsic disorder prediction by ensembled deep learning. *Genom. Proteom. Bioinform.* **2019**, *17*, 645–656. [[CrossRef](#)]
113. Xue, B.; Dunbrack, R.L.; Williams, R.W.; Dunker, A.K.; Uversky, V.N. Ponderfit: A meta-predictor of intrinsically disordered amino acids. *BBA Proteins Proteom.* **2010**, *1804*, 996–1010. [[CrossRef](#)]
114. Hu, G.; Katuwawala, A.; Wang, K.; Wu, Z.; Ghadermarzi, S.; Gao, J.; Kurgan, L. Fldpnn: Accurate intrinsic disorder prediction with putative propensities of disorder functions. *Nat. Commun.* **2021**, *12*, 4438. [[CrossRef](#)]
115. Erdos, G.; Pajkos, M.; Dosztanyi, Z. Iupred3: Prediction of protein disorder enhanced with unambiguous experimental annotation and visualization of evolutionary conservation. *Nucleic Acids Res.* **2021**, *49*, W297–W303. [[CrossRef](#)]
116. Basu, S.; Kihara, D.; Kurgan, L. Computational prediction of disordered binding regions. *Comput. Struct. Biotechnol. J.* **2023**, *21*, 1487–1497. [[CrossRef](#)]
117. Oldfield, C.J.; Cheng, Y.; Cortese, M.S.; Romero, P.; Uversky, V.N.; Dunker, A.K. Coupled folding and binding with alpha-helix-forming molecular recognition elements. *Biochemistry* **2005**, *44*, 12454–12470. [[CrossRef](#)]
118. Xue, B.; Dunker, A.K.; Uversky, V.N. Retro-morfs: Identifying protein binding sites by normal and reverse alignment and intrinsic disorder prediction. *Int. J. Mol. Sci.* **2010**, *11*, 3725–3747. [[CrossRef](#)]
119. Sharma, R.; Raicar, G.; Tsunoda, T.; Patil, A.; Sharma, A. Opal: Prediction of morf regions in intrinsically disordered protein sequences. *Bioinformatics* **2018**, *34*, 1850–1858. [[CrossRef](#)]
120. Jones, D.T.; Cozzetto, D. Disopred3: Precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* **2015**, *31*, 857–863. [[CrossRef](#)]
121. Hanson, J.; Litfin, T.; Paliwal, K.; Zhou, Y. Identifying molecular recognition features in intrinsically disordered regions of proteins by transfer learning. *Bioinformatics* **2020**, *36*, 1107–1113. [[CrossRef](#)]
122. Krystkowiak, I.; Davey, N.E. Slimsearch: A framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. *Nucleic Acids Res.* **2017**, *45*, W464–W469. [[CrossRef](#)]
123. O'Brien, K.T.; Haslam, N.J.; Shields, D.C. Slimscape: A protein short linear motif analysis plugin for cytoscape. *BMC Bioinform.* **2013**, *14*, 224. [[CrossRef](#)] [[PubMed](#)]
124. Palopoli, N.; Lythgow, K.T.; Edwards, R.J. Qslimfinder: Improved short linear motif prediction using specific query protein data. *Bioinformatics* **2015**, *31*, 2284–2293. [[CrossRef](#)] [[PubMed](#)]
125. Kumar, M.; Michael, S.; Alvarado-Valverde, J.; Meszaros, B.; Samano-Sanchez, H.; Zeke, A.; Dobson, L.; Lazar, T.; Ord, M.; Nagpal, A.; et al. The eukaryotic linear motif resource: 2022 release. *Nucleic Acids Res.* **2022**, *50*, D497–D508. [[CrossRef](#)] [[PubMed](#)]
126. Meszaros, B.; Simon, I.; Dosztanyi, Z. Prediction of protein binding regions in disordered proteins. *PLoS Comput. Biol.* **2009**, *5*, e1000376. [[CrossRef](#)] [[PubMed](#)]
127. Wong, E.T.C.; Gsponer, J. Predicting protein-protein interfaces that bind intrinsically disordered protein regions. *J. Mol. Biol.* **2019**, *431*, 3157–3178. [[CrossRef](#)]
128. Peng, Z.; Kurgan, L. High-throughput prediction of rna, DNA and protein binding regions mediated by intrinsic disorder. *Nucleic Acids Res.* **2015**, *43*, e121. [[CrossRef](#)]
129. Katuwawala, A.; Zhao, B.; Kurgan, L. Disolippred: Accurate prediction of disordered lipid-binding residues in protein sequences with deep recurrent networks and transfer learning. *Bioinformatics* **2021**, *38*, 115–124. [[CrossRef](#)]
130. Schad, E.; Ficho, E.; Panca, R.; Simon, I.; Dosztanyi, Z.; Meszaros, B. Dibs: A repository of disordered binding sites mediating interactions with ordered proteins. *Bioinformatics* **2018**, *34*, 535–537. [[CrossRef](#)]
131. Miskei, M.; Antal, C.; Fuxreiter, M. Fuzdb: Database of fuzzy complexes, a tool to develop stochastic structure-function relationships for protein complexes and higher-order assemblies. *Nucleic Acids Res.* **2017**, *45*, D228–D235. [[CrossRef](#)]
132. Gonzalez-Foutel, N.S.; Glavina, J.; Borchers, W.M.; Safranchik, M.; Barrera-Vilarmau, S.; Sagar, A.; Estana, A.; Barozet, A.; Garrone, N.A.; Fernandez-Ballester, G.; et al. Conformational buffering underlies functional selection in intrinsically disordered protein regions. *Nat. Struct. Mol. Biol.* **2022**, *29*, 781–790. [[CrossRef](#)]
133. Bugge, K.; Brakti, I.; Fernandes, C.B.; Dreier, J.E.; Lundsgaard, J.E.; Olsen, J.G.; Skriver, K.; Kragelund, B.B. Interactions by disorder—a matter of context. *Front. Mol. Biosci.* **2020**, *7*, 110. [[CrossRef](#)]

134. Das, R.K.; Pappu, R.V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 13392–13397. [[CrossRef](#)]
135. Sawle, L.; Ghosh, K. A theoretical method to compute sequence dependent configurational properties in charged polymers and proteins. *J. Chem. Phys.* **2015**, *143*, 085101. [[CrossRef](#)]
136. Samanta, H.S.; Chakraborty, D.; Thirumalai, D. Charge fluctuation effects on the shape of flexible polyampholytes with applications to intrinsically disordered proteins. *J. Chem. Phys.* **2018**, *149*, 163323. [[CrossRef](#)]
137. Zheng, W.; Dignon, G.; Brown, M.; Kim, Y.C.; Mittal, J. Hydropathy patterning complements charge patterning to describe conformational preferences of disordered proteins. *J. Phys. Chem. Lett.* **2020**, *11*, 3408–3415. [[CrossRef](#)]
138. Amin, A.N.; Lin, Y.H.; Das, S.; Chan, H.S. Analytical theory for sequence-specific binary fuzzy complexes of charged intrinsically disordered proteins. *J. Phys. Chem. B* **2020**, *124*, 6709–6720. [[CrossRef](#)]
139. Yamazaki, H.; Takagi, M.; Kosako, H.; Hirano, T.; Yoshimura, S.H. Cell cycle-specific phase separation regulated by protein charge blockiness. *Nat. Cell Biol.* **2022**, *24*, 625–632. [[CrossRef](#)]
140. Lyons, H.; Veettil, R.T.; Pradhan, P.; Fornero, C.; De La Cruz, N.; Ito, K.; Eppert, M.; Roeder, R.G.; Sabari, B.R. Functional partitioning of transcriptional regulators by patterned charge blocks. *Cell* **2023**, *186*, 327–345.e28. [[CrossRef](#)]
141. Ruff, K.M.; Pappu, R.V. AlphaFold and implications for intrinsically disordered proteins. *J. Mol. Biol.* **2021**, *433*, 167208. [[CrossRef](#)]
142. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)]
143. Lancaster, A.K.; Nutter-Upham, A.; Lindquist, S.; King, O.D. Plaac: A web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics* **2014**, *30*, 2501–2502. [[CrossRef](#)] [[PubMed](#)]
144. Orlando, G.; Raimondi, D.; Tabaro, F.; Codice, F.; Moreau, Y.; Vranken, W.F. Computational identification of prion-like rna-binding proteins that form liquid phase-separated condensates. *Bioinformatics* **2019**, *35*, 4617–4623. [[CrossRef](#)] [[PubMed](#)]
145. Ibrahim, A.Y.; Khaodeuanepheng, N.P.; Amarasekara, D.L.; Correia, J.J.; Lewis, K.A.; Fitzkee, N.C.; Hough, L.E.; Whitten, S.T. Intrinsically disordered regions that drive phase separation form a robustly distinct protein class. *J. Biol. Chem.* **2023**, *299*, 102801. [[CrossRef](#)] [[PubMed](#)]
146. Vernon, R.M.; Chong, P.A.; Tsang, B.; Kim, T.H.; Bah, A.; Farber, P.; Lin, H.; Forman-Kay, J.D. Pi-pi contacts are an overlooked protein feature relevant to phase separation. *Elife* **2018**, *7*, e31486. [[CrossRef](#)] [[PubMed](#)]
147. Chu, X.; Sun, T.; Li, Q.; Xu, Y.; Zhang, Z.; Lai, L.; Pei, J. Prediction of liquid-liquid phase separating proteins using machine learning. *BMC Bioinform.* **2022**, *23*, 72. [[CrossRef](#)]
148. Vendruscolo, M.; Fuxreiter, M. Sequence determinants of the aggregation of proteins within condensates generated by liquid-liquid phase separation. *J. Mol. Biol.* **2022**, *434*, 167201. [[CrossRef](#)]
149. Hatos, A.; Tosatto, S.C.E.; Vendruscolo, M.; Fuxreiter, M. Fuzdrop on AlphaFold: Visualizing the sequence-dependent propensity of liquid-liquid phase separation and aggregation of proteins. *Nucleic Acids Res.* **2022**, *50*, W337–W344. [[CrossRef](#)]
150. Mentes, A.; Magyar, C.; Ficho, E.; Simon, I. Analysis of heterodimeric “mutual synergistic folding”-complexes. *Int. J. Mol. Sci.* **2019**, *20*, 5136. [[CrossRef](#)]
151. Marsh, J.A.; Forman-Kay, J.D. Sequence determinants of compaction in intrinsically disordered proteins. *Biophys. J.* **2010**, *98*, 2383–2390. [[CrossRef](#)]
152. Borgia, A.; Zheng, W.; Buholzer, K.; Borgia, M.B.; Schuler, A.; Hofmann, H.; Soranno, A.; Nettels, D.; Gast, K.; Grishaev, A.; et al. Consistent view of polypeptide chain expansion in chemical denaturants from multiple experimental methods. *J. Am. Chem. Soc.* **2016**, *138*, 11714–11726. [[CrossRef](#)]
153. des Cloizeaux, J. Langrangian theory for a self-avoiding random chain. *Phys. Rev. A* **1974**, *10*, 1665–1669. [[CrossRef](#)]
154. Le Guillou, J.C.; Zinn-Justin, J. Critical exponents for n-vector model in 3 dimensions from field-theory. *Phys. Rev. Lett.* **1977**, *39*, 95–98. [[CrossRef](#)]
155. Fisher, M.E. Shape of a self-avoiding walk or polymer chain. *J. Chem. Phys.* **1966**, *44*, 616–622. [[CrossRef](#)]
156. Witten, T.A.; Schäfer, L. Two critical ratios in polymer solutions. *J. Phys. A* **1978**, *11*, 1843–1854. [[CrossRef](#)]
157. Zheng, W.; Best, R.B. An extended guinier analysis for intrinsically disordered proteins. *J. Mol. Biol.* **2018**, *430*, 2540–2553. [[CrossRef](#)]
158. Dignon, G.L.; Zheng, W.; Best, R.B.; Kim, Y.C.; Mittal, J. Relation between single-molecule properties and phase behavior of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 9929–9934. [[CrossRef](#)]
159. Gruet, A.; Dosnon, M.; Blocquel, D.; Brunel, J.; Gerlier, D.; Das, R.K.; Bonetti, D.; Gianni, S.; Fuxreiter, M.; Longhi, S.; et al. Fuzzy regions in an intrinsically disordered protein impair protein-protein interactions. *FEBS J.* **2016**, *283*, 576–594. [[CrossRef](#)]
160. Staby, L.; Due, A.D.; Kunze, M.B.A.; Jorgensen, M.L.M.; Skriver, K.; Kragelund, B.B. Flanking disorder of the folded alpha-alpha-hub domain from radical induced cell death1 affects transcription factor binding by ensemble redistribution. *J. Mol. Biol.* **2021**, *433*, 167320. [[CrossRef](#)]
161. Wang, R.Y.; Han, Y.; Krassovsky, K.; Sheffler, W.; Tyka, M.; Baker, D. Modeling disordered regions in proteins using Rosetta. *PLoS ONE* **2011**, *6*, e22060. [[CrossRef](#)]
162. Zheng, W.; Dignon, G.L.; Jovic, N.; Xu, X.; Regy, R.M.; Fawzi, N.L.; Kim, Y.C.; Best, R.B.; Mittal, J. Molecular details of protein condensates probed by microsecond long atomistic simulations. *J. Phys. Chem. B* **2020**, *124*, 11671–11679. [[CrossRef](#)]
163. Zheng, W.; Hofmann, H.; Schuler, B.; Best, R.B. Origin of internal friction in disordered proteins depends on solvent quality. *J. Phys. Chem. B* **2018**, *122*, 11478–11487. [[CrossRef](#)] [[PubMed](#)]

164. Zheng, W.; Borgia, A.; Buholzer, K.; Grishaev, A.; Schuler, B.; Best, R.B. Probing the action of chemical denaturant on an intrinsically disordered protein by simulation and experiment. *J. Am. Chem. Soc.* **2016**, *138*, 11702–11713. [[CrossRef](#)]
165. Best, R.B.; Zheng, W.; Mittal, J. Balanced protein-water interactions improve properties of disordered proteins and non-specific protein association. *J. Chem. Theory Comput.* **2014**, *10*, 5113–5124. [[CrossRef](#)] [[PubMed](#)]
166. Piana, S.; Donchev, A.G.; Robustelli, P.; Shaw, D.E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B* **2015**, *119*, 5113–5123. [[CrossRef](#)] [[PubMed](#)]
167. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B.L.; Grubmuller, H.; MacKerell, A.D. Charmm36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71–73. [[CrossRef](#)]
168. Robustelli, P.; Piana, S.; Shaw, D.E. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E4758–E4766. [[CrossRef](#)]
169. Song, D.; Luo, R.; Chen, H.F. The idp-specific force field ff14idpsff improves the conformer sampling of intrinsically disordered proteins. *J. Chem. Inf. Model.* **2017**, *57*, 1166–1178. [[CrossRef](#)]
170. Best, R.B.; Zhu, X.; Shim, J.; Lopes, P.; Mittal, J.; Feig, M.; MacKerell, A.D., Jr. Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 dihedral angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273. [[CrossRef](#)]
171. Bottaro, S.; Lindorff-Larsen, K.; Best, R.B. Variational optimization of an all-atom implicit solvent force field to match explicit solvent simulation data. *J. Chem. Theory Comput.* **2013**, *9*, 5641–5652. [[CrossRef](#)]
172. Vitalis, A.; Pappu, R.V. Absinth: A new continuum solvation model for simulations of polypeptides in aqueous solutions. *J. Comput. Chem.* **2008**, *30*, 673–699. [[CrossRef](#)]
173. Choi, J.M.; Pappu, R.V. Improvements to the absinth force field for proteins based on experimentally derived amino acid specific backbone conformational statistics. *J. Chem. Theory Comput.* **2019**, *15*, 1367–1382. [[CrossRef](#)]
174. Robustelli, P.; Piana, S.; Shaw, D.E. Mechanism of coupled folding-upon-binding of an intrinsically disordered protein. *J. Am. Chem. Soc.* **2020**, *142*, 11092–11101. [[CrossRef](#)]
175. Strodel, B. Amyloid aggregation simulations: Challenges, advances and perspectives. *Curr. Opin. Struct. Biol.* **2021**, *67*, 145–152. [[CrossRef](#)]
176. Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics methods for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151. [[CrossRef](#)]
177. Liu, P.; Kim, B.; Friesner, R.A.; Berne, B.J. Replica exchange with solute tempering: A method for sampling biological systems in explicit water. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13749–13754. [[CrossRef](#)]
178. Liu, X.; Chen, J. Residual structures and transient long-range interactions of p53 transactivation domain: Assessment of explicit solvent protein force fields. *J. Chem. Theory Comput.* **2019**, *15*, 4708–4720. [[CrossRef](#)]
179. Miao, Y.; Feher, V.A.; McCammon, J.A. Gaussian accelerated molecular dynamics: Unconstrained enhanced sampling and free energy calculation. *J. Chem. Theory Comput.* **2015**, *11*, 3584–3595. [[CrossRef](#)]
180. Hamelberg, D.; Mongan, J.; McCammon, J.A. Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919–11929. [[CrossRef](#)]
181. Tribello, G.A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. Plumed 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604–613. [[CrossRef](#)]
182. Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562–12566. [[CrossRef](#)]
183. Torrie, G.M.; Valleau, J.P. Non-physical sampling distributions in monte-carlo free-energy estimation. *J. Comp. Phys.* **1977**, *23*, 187–199. [[CrossRef](#)]
184. Shaw, D.E.; Deneroff, M.M.; Dror, R.O.; Kuskin, J.S.; Larson, R.H.; Salmon, J.K.; Young, C.; Batson, B.; Bowers, K.J.; Chao, J.C.; et al. Anton, a special-purpose machine for molecular dynamics simulation. In *Isca'07: 34th Annual International Symposium on Computer Architecture, Conference Proceedings 1–12*; Assoc Computing Machinery: New York, NY, USA, 2007.
185. Pearlman, D.A.; Case, D.A.; Caldwell, J.W.; Ross, W.S.; Cheatham, T.E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. Amber, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Comm.* **1995**, *91*, 1–41. [[CrossRef](#)]
186. Hess, B.; Kutzner, C.; Van der Spoel, D.; Lindahl, E. Gromacs4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447. [[CrossRef](#)] [[PubMed](#)]
187. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.G.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable molecular dynamics with namd. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [[CrossRef](#)]
188. Chen, X.; Jin, S.; Chen, M.; Bueno, C.; Wolynes, P.G. The marionette mechanism of domain-domain communication in the antagonist, agonist, and coactivator responses of the estrogen receptor. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2216906120. [[CrossRef](#)]
189. Wu, H.; Wolynes, P.G.; Papoian, G.A. Awsem-idp: A coarse-grained force field for intrinsically disordered proteins. *J. Phys. Chem. B* **2018**, *122*, 11115–11125. [[CrossRef](#)]
190. Latham, A.P.; Zhang, B. Improving coarse-grained protein force fields with small-angle X-ray scattering data. *J. Phys. Chem. B* **2019**, *123*, 1026–1034. [[CrossRef](#)]

191. Ozenne, V.; Bauer, F.; Salmon, L.; Huang, J.R.; Jensen, M.R.; Segard, S.; Bernado, P.; Charavay, C.; Blackledge, M. Flexible-meccano: A tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. *Bioinformatics* **2012**, *28*, 1463–1470. [[CrossRef](#)]
192. Dignon, G.L.; Zheng, W.W.; Kim, Y.C.; Best, R.B.; Mittal, J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput. Biol.* **2018**, *14*, e1005941. [[CrossRef](#)]
193. Debye, P.; Hückel, E. De la theorie des electrolytes. I. Abaissement du point de congelation et phenomenes associes. *Phys. Z.* **1923**, *24*, 185–206.
194. Ashbaugh, H.S.; Hatch, H.W. Natively unfolded protein stability as a coil-to-globule transition in charge/hydrophathy space. *J. Am. Chem. Soc.* **2008**, *130*, 9536–9542. [[CrossRef](#)]
195. Joseph, J.A.; Reinhardt, A.; Aguirre, A.; Chew, P.Y.; Russell, K.O.; Espinosa, J.R.; Garaizar, A.; Collepardo-Guevara, R. Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy. *Nat. Comput. Sci.* **2021**, *1*, 732–743. [[CrossRef](#)]
196. Wang, X.; Ramirez-Hinestrosa, S.; Dobnikar, J.; Frenkel, D. The lennard-jones potential: When. (not) to use it. *Phys. Chem. Chem. Phys.* **2020**, *22*, 10624–10633. [[CrossRef](#)]
197. Kim, Y.C.; Hummer, G. Coarse-grained models for simulations of multiprotein complexes: Application to ubiquitin binding. *J. Mol. Biol.* **2008**, *375*, 1416–1433. [[CrossRef](#)]
198. Ravikumar, K.M.; Huang, W.; Yang, S. Coarse-grained simulations of protein-protein association: An energy landscape perspective. *Biophys. J.* **2012**, *103*, 837–845. [[CrossRef](#)]
199. Dannenhoffer-Lafage, T.; Best, R.B. A data-driven hydrophobicity scale for predicting liquid-liquid phase separation of proteins. *J. Phys. Chem. B* **2021**, *125*, 4046–4056. [[CrossRef](#)]
200. Regy, R.M.; Thompson, J.; Kim, Y.C.; Mittal, J. Improved coarse-grained model for studying sequence dependent phase separation of disordered proteins. *Protein Sci.* **2021**, *30*, 1371–1379. [[CrossRef](#)]
201. Tesei, G.; Schulze, T.K.; Crehuet, R.; Lindorff-Larsen, K. Accurate model of liquid-liquid phase behavior of intrinsically disordered proteins from optimization of single-chain properties. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2111696118. [[CrossRef](#)]
202. Dignon, G.L.; Zheng, W.; Kim, Y.C.; Mittal, J. Temperature-controlled liquid-liquid phase separation of disordered proteins. *ACS Cent. Sci.* **2019**, *5*, 821–830. [[CrossRef](#)]
203. Wohl, S.; Jakubowski, M.; Zheng, W. Salt-dependent conformational changes of intrinsically disordered proteins. *J. Phys. Chem. Lett.* **2021**, *12*, 6684–6691. [[CrossRef](#)]
204. Rizuan, A.; Jovic, N.; Phan, T.M.; Kim, Y.C.; Mittal, J. Developing bonded potentials for a coarse-grained model of intrinsically disordered proteins. *J. Chem. Inf. Model.* **2022**, *62*, 4474–4485. [[CrossRef](#)] [[PubMed](#)]
205. Wang, H.; Wu, J.; Sterneke-Hoffmann, R.; Zheng, W.; Morman, C.; Luo, J. Multivariate effects of ph, salt, and zn(2+) ions on abeta(40) fibrillation. *Commun. Chem.* **2022**, *5*, 171. [[CrossRef](#)] [[PubMed](#)]
206. Regy, R.M.; Dignon, G.L.; Zheng, W.; Kim, Y.C.; Mittal, J. Sequence dependent phase separation of protein-polynucleotide mixtures elucidated using molecular simulations. *Nucleic Acids Res.* **2020**, *48*, 12593–12603. [[CrossRef](#)] [[PubMed](#)]
207. Best, R.B.; Vendruscolo, M. Determination of ensembles of protein structures consistent with nmr order parameters. *J. Am. Chem. Soc.* **2004**, *126*, 8090–8091. [[CrossRef](#)] [[PubMed](#)]
208. Lindorff-Larsen, K.; Best, R.B.; Depristo, M.A.; Dobson, C.M.; Vendruscolo, M. Simultaneous determination of protein structure and dynamics. *Nature* **2005**, *433*, 128–132. [[CrossRef](#)]
209. Jensen, M.R.; Salmon, L.; Nodet, G.; Blackledge, M. Defining conformational ensembles of intrinsically disordered and partially folded proteins directly from chemical shifts. *J. Am. Chem. Soc.* **2010**, *132*, 1270–1272. [[CrossRef](#)]
210. Kofinger, J.; Stelzl, L.S.; Reuter, K.; Allande, C.; Reichel, K.; Hummer, G. Efficient ensemble refinement by reweighting. *J. Chem. Theory Comput.* **2019**, *15*, 3390–3401. [[CrossRef](#)]
211. Brookes, D.H.; Head-Gordon, T. Experimental inferential structure determination of ensembles for intrinsically disordered proteins. *J. Am. Chem. Soc.* **2016**, *138*, 4530–4538. [[CrossRef](#)]
212. Gomes, G.W.; Krzeminski, M.; Namini, A.; Martin, E.W.; Mittag, T.; Head-Gordon, T.; Forman-Kay, J.D.; Gradinaru, C.C. Conformational ensembles of an intrinsically disordered protein consistent with nmr, saxs, and single-molecule fret. *J. Am. Chem. Soc.* **2020**, *142*, 15697–15710. [[CrossRef](#)]
213. Hsieh, A.; Lu, L.; Chance, M.R.; Yang, S. A practical guide to ispot modeling: An integrative structural biology platform. *Biol. Small Angle Scatt. Tech. Strateg. Tips* **2017**, *1009*, 229–238.
214. Huang, W.; Ravikumar, K.M.; Parisien, M.; Yang, S. Theoretical modeling of multiprotein complexes by ispot: Integration of small-angle X-ray scattering, hydroxyl radical footprinting, and computational docking. *J. Struct. Biol.* **2016**, *196*, 340–349. [[CrossRef](#)]
215. Yang, S.; Bernado, P. Integrative biophysics: Protein interaction and disorder. *J. Mol. Biol.* **2020**, *432*, 2843–2845. [[CrossRef](#)]
216. Tong, D.; Yang, S.; Lu, L. Accurate optimization of amino acid form factors for computing small-angle X-ray scattering intensity of atomistic protein structures. *J. Appl. Crystallogr.* **2016**, *49*, 1148–1161. [[CrossRef](#)]
217. Ravikumar, K.M.; Huang, W.; Yang, S. Fast-saxs-pro: A unified approach to computing saxs profiles of DNA, rna, protein, and their complexes. *J. Chem. Phys.* **2013**, *138*, 024112. [[CrossRef](#)]
218. Niebling, S.; Bjorling, A.; Westenhoff, S. Martini bead form factors for the analysis of time-resolved X-ray scattering of proteins. *J. Appl. Crystallogr.* **2014**, *47*, 1190–1198. [[CrossRef](#)]

219. Svergun, D.; Barberato, C.; Koch, M.H.J. Crysol—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **1995**, *28*, 768–773. [[CrossRef](#)]
220. Schneidman-Duhovny, D.; Hammel, M.; Tainer, J.A.; Sali, A. Accurate saxs profile computation and its assessment by contrast variation experiments. *Biophys. J.* **2013**, *105*, 962–974. [[CrossRef](#)]
221. Gong, Z.; Schwieters, C.D.; Tang, C. Theory and practice of using solvent paramagnetic relaxation enhancement to characterize protein conformational dynamics. *Methods* **2018**, *148*, 48–56. [[CrossRef](#)]
222. Schwieters, C.D.; Kuszewski, J.J.; Tjandra, N.; Clore, G.M. The xplor-nih nmr molecular structure determination package. *J. Magn. Reson.* **2003**, *160*, 65–73. [[CrossRef](#)]
223. Qi, Y.; Lee, J.; Cheng, X.; Shen, R.; Islam, S.M.; Roux, B.; Im, W. Charmm-gui deer facilitator for spin-pair distance distribution calculations and preparation of restrained-ensemble molecular dynamics simulations. *J. Comput. Chem.* **2020**, *41*, 415–420. [[CrossRef](#)]
224. Worswick, S.G.; Spencer, J.A.; Jeschke, G.; Kuprov, I. Deep neural network processing of deer data. *Sci. Adv.* **2018**, *4*, eaat5218. [[CrossRef](#)] [[PubMed](#)]
225. Islam, S.M.; Stein, R.A.; McHaourab, H.S.; Roux, B. Structural refinement from restrained-ensemble simulations based on epr/deer data: Application to t4 lysozyme. *J. Phys. Chem. B* **2013**, *117*, 4740–4754. [[CrossRef](#)] [[PubMed](#)]
226. Roux, B.; Weare, J. On the statistical equivalence of restrained-ensemble simulations with the maximum entropy method. *J. Chem. Phys.* **2013**, *138*, 084107. [[CrossRef](#)] [[PubMed](#)]
227. Dannenhoffer-Lafage, T.; White, A.D.; Voth, G.A. A direct method for incorporating experimental data into multiscale coarse-grained models. *J. Chem. Theory Comput.* **2016**, *12*, 2144–2153. [[CrossRef](#)]
228. White, A.D.; Voth, G.A. Efficient and minimal method to bias molecular simulations with experimental data. *J. Chem. Theory Comput.* **2014**, *10*, 3023–3030. [[CrossRef](#)]
229. Pitera, J.W.; Chodera, J.D. On the use of experimental observations to bias simulated ensembles. *J. Chem. Theory Comput.* **2012**, *8*, 3445–3451. [[CrossRef](#)]
230. Hermann, M.R.; Hub, J.S. Saxs-restrained ensemble simulations of intrinsically disordered proteins with commitment to the principle of maximum entropy. *J. Chem. Theory Comput.* **2019**, *15*, 5103–5115. [[CrossRef](#)]
231. Hub, J.S. Interpreting solution X-ray scattering data using molecular simulations. *Curr. Opin. Struct. Biol.* **2018**, *49*, 18–26. [[CrossRef](#)]
232. Shen, R.; Han, W.; Fiorin, G.; Islam, S.M.; Schulten, K.; Roux, B. Structural refinement of proteins by restrained molecular dynamics simulations with non-interacting molecular fragments. *PLoS Comput. Biol.* **2015**, *11*, e1004368. [[CrossRef](#)]
233. Biehn, S.E.; Lindert, S. Accurate protein structure prediction with hydroxyl radical protein footprinting data. *Nat. Commun.* **2021**, *12*, 341. [[CrossRef](#)]
234. Nath, A.; Sammalkorpi, M.; DeWitt, D.C.; Trexler, A.J.; Elbaum-Garfinkle, S.; O'Hern, C.S.; Rhoades, E. The conformational ensembles of alpha-synuclein and tau: Combining single-molecule fret and simulations. *Biophys. J.* **2012**, *103*, 1940–1949. [[CrossRef](#)]
235. Tang, C.; Gong, Z. Integrating non-nmr distance restraints to augment nmr depiction of protein structure and dynamics. *J. Mol. Biol.* **2020**, *432*, 2913–2929. [[CrossRef](#)]
236. Delhommel, F.; Gabel, F.; Sattler, M. Current approaches for integrating solution nmr spectroscopy and small-angle scattering to study the structure and dynamics of biomolecular complexes. *J. Mol. Biol.* **2020**, *432*, 2890–2912. [[CrossRef](#)]
237. Schindler, C.E.M.; de Vries, S.J.; Sasse, A.; Zacharias, M. Saxs data alone can generate high-quality models of protein-protein complexes. *Structure* **2016**, *24*, 1387–1397. [[CrossRef](#)]
238. Kozakov, D.; Hall, D.R.; Xia, B.; Porter, K.A.; Padhorny, D.; Yueh, C.; Beglov, D.; Vajda, S. The cluspro web server for protein-protein docking. *Nat. Protoc.* **2017**, *12*, 255–278. [[CrossRef](#)]
239. Xia, B.; Mamonov, A.; Leysen, S.; Allen, K.N.; Strelkov, S.V.; Paschalidis, I.; Vajda, S.; Kozakov, D. Accounting for observed small angle X-ray scattering profile in the protein-protein docking server cluspro. *J. Comput. Chem.* **2015**, *36*, 1568–1572. [[CrossRef](#)]
240. Jimenez-Garcia, B.; Bernado, P.; Fernandez-Recio, J. Structural characterization of protein-protein interactions with pydock saxs. *Methods Mol. Biol.* **2020**, *2112*, 131–144.
241. Vangone, A.; Oliva, R.; Cavallo, L.; Bonvin, A.M.J.J. Prediction of biomolecular complexes. In *From Protein Structure to Function with Bioinformatics*; Rigden, D.J., Ed.; Springer: Dordrecht, The Netherlands, 2017; pp. 265–292.
242. Karaca, E.; Bonvin, A.M.J.J. On the usefulness of ion-mobility mass spectrometry and saxs data in scoring docking decoys. *Acta Crystallogr. Sect. D* **2013**, *69*, 683–694. [[CrossRef](#)]
243. Schneidman-Duhovny, D.; Hammel, M.; Sali, A. Macromolecular docking restrained by a small angle X-ray scattering profile. *J. Struct. Biol.* **2011**, *173*, 461–471. [[CrossRef](#)]
244. Huang, W.; Peng, Y.; Kiselar, J.; Zhao, X.; Albaqami, A.; Mendez, D.; Chen, Y.; Chakravarthy, S.; Gupta, S.; Ralston, C.; et al. Multidomain architecture of estrogen receptor reveals interfacial cross-talk between its DNA-binding and ligand-binding domains. *Nat. Commun.* **2018**, *9*, 3520. [[CrossRef](#)]
245. Paissoni, C.; Jussupow, A.; Camilloni, C. Martini bead form factors for nucleic acids and their application in the refinement of protein-nucleic acid complexes against saxs data. *J. Appl. Crystallogr.* **2019**, *52*, 394–402. [[CrossRef](#)]
246. Pahari, S.; Liu, S.; Lee, C.H.; Akbulut, M.; Kwon, J.S. Saxs-guided unbiased coarse-grained monte carlo simulation for identification of self-assembly nanostructures and dimensions. *Soft Matter* **2022**, *18*, 5282–5292. [[CrossRef](#)] [[PubMed](#)]

247. Ruan, H.; Kiselar, J.; Zhang, W.; Li, S.S.; Xiong, R.; Liu, Y.; Yang, S.; Lai, L. Integrative structural modeling of a multidomain polo-like kinase. *Phys. Chem. Chem. Phys.* **2020**, *22*, 27581–27589. [[CrossRef](#)] [[PubMed](#)]
248. Ekimoto, T.; Ikeguchi, M. Hybrid methods for modeling protein structures using molecular dynamics simulations and small-angle X-ray scattering data. In *Integrative Structural Biology with Hybrid Methods*; Nakamura, H., Kleywegt, G., Burley, S.K., Markley, J.L., Eds.; Springer: Singapore, 2018; pp. 237–258.
249. Bowerman, S.; Curtis, J.E.; Clayton, J.; Brookes, E.H.; Wereszczynski, J. Bees: Bayesian ensemble estimation from sas. *Biophys. J.* **2019**, *117*, 399–407. [[CrossRef](#)] [[PubMed](#)]
250. Yang, S.; Blachowicz, L.; Makowski, L.; Roux, B. Multidomain assembled states of hck tyrosine kinase in solution. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 15757–15762. [[CrossRef](#)]
251. Bernado, P.; Blackledge, M. Structural biology: Proteins in dynamic equilibrium. *Nature* **2010**, *468*, 1046–1048. [[CrossRef](#)]
252. Song, L.; Yang, L.; Meng, J.; Yang, S. Thermodynamics of hydrophobic amino acids in solution: A combined experimental-computational study. *J. Phys. Chem. Lett.* **2017**, *8*, 347–351. [[CrossRef](#)]
253. Antonov, L.D.; Olsson, S.; Boomsma, W.; Hamelryck, T. Bayesian inference of protein ensembles from saxs data. *Phys. Chem. Chem. Phys.* **2016**, *18*, 5832–5838. [[CrossRef](#)]
254. Jamros, M.A.; Oliveira, L.C.; Whitford, P.C.; Onuchic, J.N.; Adams, J.A.; Blumenthal, D.K.; Jennings, P.A. Proteins at work: A combined small angle X-ray scattering and theoretical determination of the multiple structures involved on the protein kinase functional landscape. *J. Biol. Chem.* **2010**, *285*, 36121–36128. [[CrossRef](#)]
255. Zhang, Y.; Wen, B.; Peng, J.; Zuo, X.; Gong, Q.; Zhang, Z. Determining structural ensembles of flexible multi-domain proteins using small-angle X-ray scattering and molecular dynamics simulations. *Protein Cell* **2015**, *6*, 619–623. [[CrossRef](#)]
256. Miyashita, O.; Gorba, C.; Tama, F. Structure modeling from small angle X-ray scattering data with elastic network normal mode analysis. *J. Struct. Biol.* **2011**, *173*, 451–460. [[CrossRef](#)]
257. Liu, Z.; Gong, Z.; Cao, Y.; Ding, Y.H.; Dong, M.Q.; Lu, Y.B.; Zhang, W.P.; Tang, C. Characterizing protein dynamics with integrative use of bulk and single-molecule techniques. *Biochemistry* **2018**, *57*, 305–313. [[CrossRef](#)]
258. Chen, Y.; Pollack, L. Saxs studies of rna: Structures, dynamics, and interactions with partners. *Wiley Interdiscip Rev. RNA* **2016**, *7*, 512–526. [[CrossRef](#)]
259. Yang, S.C.; Parisien, M.; Major, F.; Roux, B. Rna structure determination using saxs data. *J. Phys. Chem. B* **2010**, *114*, 10039–10048. [[CrossRef](#)]
260. Sun, L.Z.; Zhang, D.; Chen, S.J. Theory and modeling of rna structure and interactions with metal ions and small molecules. *Annu. Rev. Biophys.* **2017**, *46*, 227–246. [[CrossRef](#)]
261. Prajapati, J.D.; Onuchic, J.N.; Sanbonmatsu, K.Y. Exploring the energy landscape of riboswitches using collective variables based on tertiary contacts. *J. Mol. Biol.* **2022**, *434*, 167788. [[CrossRef](#)]
262. Bernado, P.; Mylonas, E.; Petoukhov, M.V.; Blackledge, M.; Svergun, D.I. Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* **2007**, *129*, 5656–5664. [[CrossRef](#)]
263. Sterckx, Y.G.; Volkov, A.N.; Vranken, W.F.; Kragelj, J.; Jensen, M.R.; Buts, L.; Garcia-Pino, A.; Jove, T.; Van Melderen, L.; Blackledge, M.; et al. Small-angle X-ray scattering-and nuclear magnetic resonance-derived conformational ensemble of the highly flexible antitoxin paa2. *Structure* **2014**, *22*, 854–865. [[CrossRef](#)]
264. Bernado, P.; Blanchard, L.; Timmins, P.; Marion, D.; Ruigrok, R.W.; Blackledge, M. A structural model for unfolded proteins from residual dipolar couplings and small-angle X-ray scattering. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17002–17007. [[CrossRef](#)]
265. Lin, X.; Roy, S.; Jolly, M.K.; Bocci, F.; Schafer, N.P.; Tsai, M.Y.; Chen, Y.; He, Y.; Grishaev, A.; Weninger, K.; et al. Page4 and conformational switching: Insights from molecular dynamics simulations and implications for prostate cancer. *J. Mol. Biol.* **2018**, *430*, 2422–2438. [[CrossRef](#)]
266. Receveur-Brechot, V.; Durand, D.H. How random are intrinsically disordered proteins? A small angle scattering perspective. *Curr. Protein Pept. Sci.* **2012**, *13*, 55–75. [[CrossRef](#)] [[PubMed](#)]
267. Marsh, J.A.; Forman-Kay, J.D. Ensemble modeling of protein disordered states: Experimental restraint contributions and validation. *Proteins* **2012**, *80*, 556–572. [[CrossRef](#)] [[PubMed](#)]
268. Gomes, G.-N.W.; Namini, A.; Gradinaru, C.C. Integrative conformational ensembles of sic1 using different initial pools and optimization methods. *Front. Mol. Biosci.* **2022**, *9*, 910956. [[CrossRef](#)] [[PubMed](#)]
269. Aznauryan, M.; Delgado, L.; Soranno, A.; Nettels, D.; Huang, J.R.; Labhardt, A.M.; Grzesiek, S.; Schuler, B. Comprehensive structural and dynamical view of an unfolded protein from the combination of single-molecule fret, nmr, and saxs. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E5389–E5398. [[CrossRef](#)]
270. Rozycki, B.; Kim, Y.C.; Hummer, G. Saxs ensemble refinement of esrt-iii chmp3 conformational transitions. *Structure* **2011**, *19*, 109–116. [[CrossRef](#)]
271. Manalastas, K.G.; Svergun, D.I. Molecular dissection of the intrinsically disordered estrogen receptor alpha-ntd. *Structure* **2019**, *27*, 207–208. [[CrossRef](#)]
272. Ambadipudi, S.; Zweckstetter, M. Targeting intrinsically disordered proteins in rational drug discovery. *Expert Opin. Drug Discov.* **2016**, *11*, 65–77. [[CrossRef](#)]
273. Ruan, H.; Sun, Q.; Zhang, W.; Liu, Y.; Lai, L. Targeting intrinsically disordered proteins at the edge of chaos. *Drug Discov. Today* **2019**, *24*, 217–227. [[CrossRef](#)]
274. Metallo, S.J. Intrinsically disordered proteins are potential drug targets. *Curr. Opin. Chem. Biol.* **2010**, *14*, 481–488. [[CrossRef](#)]

275. Choi, S.H.; Mahankali, M.; Lee, S.J.; Hull, M.; Petrassi, H.M.; Chatterjee, A.K.; Schultz, P.G.; Jones, K.A.; Shen, W. Targeted disruption of myc-max oncoprotein complex by a small molecule. *ACS Chem. Biol.* **2017**, *12*, 2715–2719. [CrossRef]
276. Zhao, J.; Blayney, A.; Liu, X.; Gandy, L.; Jin, W.; Yan, L.; Ha, J.H.; Canning, A.J.; Connelly, M.; Yang, C.; et al. Egcg binds intrinsically disordered n-terminal domain of p53 and disrupts p53-mdm2 interaction. *Nat. Commun.* **2021**, *12*, 986. [CrossRef]
277. Tatenhorst, L.; Eckermann, K.; Dambeck, V.; Fonseca-Ornelas, L.; Walle, H.; Lopes da Fonseca, T.; Koch, J.C.; Becker, S.; Tonges, L.; Bahr, M.; et al. Fasudil attenuates aggregation of alpha-synuclein in models of parkinson's disease. *Acta Neuropathol. Commun.* **2016**, *4*, 39. [CrossRef]
278. Heller, G.T.; Aprile, F.A.; Michaels, T.C.T.; Limbocker, R.; Perni, M.; Ruggeri, F.S.; Mannini, B.; Lohr, T.; Bonomi, M.; Camilloni, C.; et al. Small-molecule sequestration of amyloid-beta as a drug discovery strategy for alzheimer's disease. *Sci. Adv.* **2020**, *6*, eabb5924. [CrossRef]
279. Iconaru, L.I.; Das, S.; Nourse, A.; Shelat, A.A.; Zuo, J.; Kriwacki, R.W. Small molecule sequestration of the intrinsically disordered protein, p27(kip1), within soluble oligomers. *J. Mol. Biol.* **2021**, *433*, 167120. [CrossRef]
280. Myung, J.K.; Banuelos, C.A.; Fernandez, J.G.; Mawji, N.R.; Wang, J.; Tien, A.H.; Yang, Y.C.; Tavakoli, I.; Haile, S.; Watt, K.; et al. An androgen receptor n-terminal domain antagonist for treating prostate cancer. *J. Clin. Investig.* **2013**, *123*, 2948–2960. [CrossRef]
281. Bier, D.; Mittal, S.; Bravo-Rodriguez, K.; Sowislok, A.; Guillory, X.; Briels, J.; Heid, C.; Bartel, M.; Wettig, B.; Brunsveld, L.; et al. The molecular tweezer chr01 stabilizes a disordered protein-protein interface. *J. Am. Chem. Soc.* **2017**, *139*, 16256–16263. [CrossRef]
282. Erkizan, H.V.; Kong, Y.; Merchant, M.; Schlottmann, S.; Barber-Rotenberg, J.S.; Yuan, L.; Abaan, O.D.; Chou, T.H.; Dakshanamurthy, S.; Brown, M.L.; et al. A small molecule blocking oncogenic protein ews-flt1 interaction with rna helicase inhibits growth of ewing's sarcoma. *Nat. Med.* **2009**, *15*, 750–756. [CrossRef]
283. Andersen, R.J.; Mawji, N.R.; Wang, J.J.; Wang, G.; Haile, S.; Myung, J.K.; Watt, K.; Tam, T.; Yang, Y.C.; Banuelos, C.A.; et al. Regression of castrate-recurrent prostate cancer by a small-molecule inhibitor of the amino-terminus domain of the androgen receptor. *Cancer Cell* **2010**, *17*, 535–546. [CrossRef]
284. De Mol, E.; Fenwick, R.B.; Phang, C.T.; Buzon, V.; Szulc, E.; de la Fuente, A.; Escobedo, A.; Garcia, J.; Bertocini, C.W.; Estebanez-Perpina, E.; et al. Epi-001, a compound active against castration-resistant prostate cancer, targets transactivation unit 5 of the androgen receptor. *ACS Chem. Biol.* **2016**, *11*, 2499–2505. [CrossRef]
285. Peissert, S.; Schlosser, A.; Kendel, R.; Kuper, J.; Kisker, C. Structural basis for cdk7 activation by mat1 and cyclin h. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 26739–26748. [CrossRef]
286. Chen, D.; Riedl, T.; Washbrook, E.; Pace, P.E.; Coombes, R.C.; Egly, J.M.; Ali, S. Activation of estrogen receptor alpha by s118 phosphorylation involves a ligand-dependent interaction with tfiih and participation of cdk7. *Mol. Cell* **2000**, *6*, 127–137. [CrossRef] [PubMed]
287. Wells, M.; Tidow, H.; Rutherford, T.J.; Markwick, P.; Jensen, M.R.; Mylonas, E.; Svergun, D.I.; Blackledge, M.; Fersht, A.R. Structure of tumor suppressor p53 and its intrinsically disordered n-terminal transactivation domain. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5762–5767. [CrossRef] [PubMed]
288. Vise, P.; Baral, B.; Stancik, A.; Lowry, D.F.; Daughdrill, G.W. Identifying long-range structure in the intrinsically unstructured transactivation domain of p53. *Proteins* **2007**, *67*, 526–530. [CrossRef]
289. Sadar, M.D. Small molecule inhibitors targeting the "achilles' heel" of androgen receptor activity. *Cancer Res.* **2011**, *71*, 1208–1213. [CrossRef] [PubMed]
290. Sadar, M.D. Drugging the undruggable: Targeting the n-terminal domain of nuclear hormone receptors. *Adv. Exp. Med. Biol.* **2022**, *1390*, 311–326.
291. Zhu, J.; Salvatella, X.; Robustelli, P. Small molecules targeting the disordered transactivation domain of the androgen receptor induce the formation of collapsed helical states. *Nat. Commun.* **2022**, *13*, 6390. [CrossRef]
292. Lavery, D.N.; McEwan, I.J. Structure and function of steroid receptor af1 transactivation domains: Induction of active conformations. *Biochem. J.* **2005**, *391*, 449–464. [CrossRef]
293. Krois, A.S.; Dyson, H.J.; Wright, P.E. Long-range regulation of p53 DNA binding by its intrinsically disordered n-terminal transactivation domain. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E11302–E11310. [CrossRef]
294. Warnmark, A.; Wikstrom, A.; Wright, A.P.H.; Gustafsson, J.A.; Hard, T. The n-terminal regions of estrogen receptor alpha and beta are unstructured in vitro and show different tbp binding properties. *J. Biol. Chem.* **2001**, *276*, 45939–45944. [CrossRef]
295. Rajbhandari, P.; Finn, G.; Solodin, N.M.; Singarapu, K.K.; Sahu, S.C.; Markley, J.L.; Kadunc, K.J.; Ellison-Zelski, S.J.; Kariagina, A.; Haslam, S.Z.; et al. Regulation of estrogen receptor alpha n-terminus conformation and function by peptidyl prolyl isomerase pin1. *Mol. Cell Biol.* **2012**, *32*, 445–457. [CrossRef]
296. Patel, H.; Periyasamy, M.; Sava, G.P.; Bondke, A.; Slafer, B.W.; Kroll, S.H.B.; Barbazanges, M.; Starkey, R.; Ottaviani, S.; Harrod, A.; et al. Icc0942, an orally bioavailable selective inhibitor of cdk7 for cancer treatment. *Mol. Cancer Ther.* **2018**, *17*, 1156–1166. [CrossRef]
297. Sava, G.P.; Fan, H.; Coombes, R.C.; Buluwela, L.; Ali, S. Cdk7 inhibitors as anticancer drugs. *Cancer Metastasis Rev.* **2020**, *39*, 805–823. [CrossRef]
298. Limited, C.T. Modular Study to Evaluate ct7001 Alone in Cancer Patients with Advanced Malignancies. 2017. Available online: <https://ClinicalTrials.gov/show/NCT03363893> (accessed on 1 February 2023).
299. Sammak, S.; Zinzalla, G. Targeting protein-protein interactions. (ppis) of transcription factors: Challenges of intrinsically disordered proteins. (idps) and regions. (idrs). *Prog. Biophys. Mol. Biol.* **2015**, *119*, 41–46. [CrossRef]

300. Choudhary, S.; Lopus, M.; Hosur, R.V. Targeting disorders in unstructured and structured proteins in various diseases. *Biophys. Chem.* **2022**, *281*, 106742. [[CrossRef](#)]
301. Qiu, Y.; Li, X.; He, X.; Pu, J.; Zhang, J.; Lu, S. Computational methods-guided design of modulators targeting protein-protein interactions. (ppis). *Eur. J. Med. Chem.* **2020**, *207*, 112764. [[CrossRef](#)]
302. Martin, J.; Frezza, E. A dynamical view of protein-protein complexes: Studies by molecular dynamics simulations. *Front. Mol. Biosci.* **2022**, *9*, 970109. [[CrossRef](#)]
303. Tompa, P.; Davey, N.E.; Gibson, T.J.; Babu, M.M. A million peptide motifs for the molecular biologist. *Mol. Cell* **2014**, *55*, 161–169. [[CrossRef](#)] [[PubMed](#)]
304. Davey, N.E.; Van Roey, K.; Weatheritt, R.J.; Toedt, G.; Uyar, B.; Altenberg, B.; Budd, A.; Diella, F.; Dinkel, H.; Gibson, T.J. Attributes of short linear motifs. *Mol. Biosyst.* **2012**, *8*, 268–281. [[CrossRef](#)]
305. Kumar, M.; Gouw, M.; Michael, S.; Samano-Sanchez, H.; Pancsa, R.; Glavina, J.; Diakogianni, A.; Valverde, J.A.; Bukirova, D.; Calyseva, J.; et al. Elm-the eukaryotic linear motif resource in 2020. *Nucleic Acids Res.* **2020**, *48*, D296–D306. [[CrossRef](#)]
306. Dinkel, H.; Van Roey, K.; Michael, S.; Davey, N.E.; Weatheritt, R.J.; Born, D.; Speck, T.; Kruger, D.; Grebnev, G.; Kuban, M.; et al. The eukaryotic linear motif resource elm: 10 years and counting. *Nucleic Acids Res.* **2014**, *42*, D259–D266. [[CrossRef](#)]
307. Kim, M.; Park, J.; Bouhaddou, M.; Kim, K.; Rojc, A.; Modak, M.; Soucheray, M.; McGregor, M.J.; O’Leary, P.; Wolf, D.; et al. A protein interaction landscape of breast cancer. *Science* **2021**, *374*, eabf3066. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.