Intrinsically Disordered Proteins at the Nano-scale

T. Ehm[1,2], H. Shinar[1], S. Meir[1], A. Sekhon[1], V. Sethi[1], I. L. Morgan[3], G. Rahamim[1], O. A. Saleh[3,4] and R. Beck[1]

[1] The School of Physics and Astronomy, The Center for Nanoscience and Nanotechnology, and the Center for Physics and Chemistry of Living Systems, Tel Aviv University, Israel

[2] Faculty of Physics and Center for NanoScience, Ludwig-Maximilians-Universität, München, Germany

[3] BMSE Program, University of California, Santa Barbara CA 93110

[4] Materials Department, University of California, Santa Barbara CA 93110

E-mail: roy@tauex.tau.ac.il

Abstract.

The human proteome is enriched in proteins that do not fold into a stable 3D structure. These intrinsically disordered proteins (IDPs) spontaneously fluctuate between a large number of configurations in their native form. Remarkably, the disorder does not lead to dysfunction as with denatured folded proteins. In fact, unlike denatured proteins, recent evidence strongly suggests that multiple biological functions stem from such structural plasticity. Here, focusing on the nanometer length-scale, we review the latest advances in IDP research and discuss some of the future directions in this highly promising field.

Keywords: Intrinsically disordered proteins, nanoscopic characterization, single-molecule force spectroscopy, SAXS, FRET, NMR, FCS

1. Introduction

The conventional paradigm in structural biology draws a direct connection between the amino-acid sequence of a protein and a singular 3D structure. The unique structure is considered essential to the protein's biological function by permitting only specific interactions.

In the last two decades it has been recognized that up to 40-50% of the proteome does not fit this simplified convention (Fig. 1). Instead, intrinsically disordered proteins and regions (IDP/IDR) fluctuate between a large number of conformations while still retaining their biological functions [1–5]. An IDR is usually defined as an unstructured amino-acid stretch as part of a (folded) protein, and an IDP as a complete protein that does not fold to a stable 3D structure [2–5]. For brevity, we will also use the IDP term to describe proteins having IDRs.

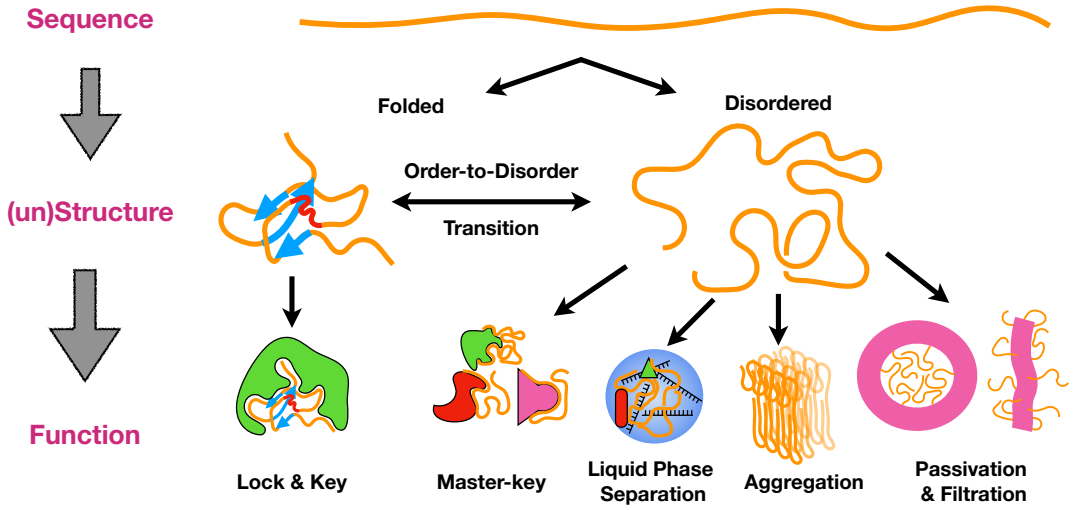

Figure 1. Contemporary sequence-function paradigm. The folded and disordered conformations, and the transition between the two, lead to biological function.

Typically, most IDP sequences are rich in structure-breaking charged and polar amino acids, and depleted in order-promoting hydrophobic residues (Fig. 2, and Refs. [2, 4–8]). Indeed, computer-based methods exploit the amino-acid propensity and sequence as a sign of a disorder [9]. Generally, IDPs fall into three distinct compositional classes that reflect the fraction of charged versus polar residues: polar, polyampholytes, and polyelectrolytes [10, 11]. In addition, the balance between solvent mediated intra-chain attraction and repulsion directly influences the IDP's compactness. The compactness, in-turn, determines the accessibility to interact with other biomolecules. Therefore, modeling the inter- and intra-molecular IDP interactions, as well as the resulting ensemble of conformations, is an extremely complex and relevant problem in the research of nano-scale systems.

It is possible to characterize an IDP's ensemble average parameters such as the end-to-end distance (R_{ee}) or the radius of gyration (R_G) distributions. In addition, the IDP's length, charged amino-acid distribution, and propensity to form transient bridges dictate the ensemble physical properties and can be connected to polymer physics theories. This could lead to a new language that relates the biological function to basic physical parameters. Thus, IDP research has the opportunity to gain a novel insight into the biological sequence-to-function problem (Fig. 1).

2. Biological function

The structural plasticity of IDPs is the key to their function. This plasticity, almost by definition, limits the strength of the IDPs' interactions with other biomolecules and the environment. Interactions larger than the thermal energy would ultimately lead to a specific configuration, which conflicts with the IDPs' large ensemble of conformations [13, 14]. Therefore, we must ask ourselves: does structural plasticity compete with

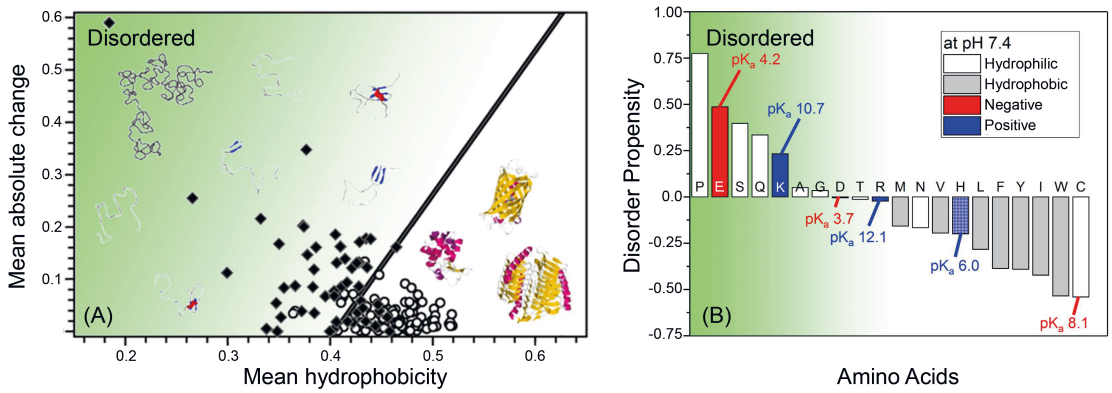

Figure 2. The role of charge, hydrophobicity and amino acid residue in IDPs. (A) Nearly 250 folded (the open circles) and nearly 90 natively unfolded proteins (the black diamonds) demonstrate that IDPs are charged and hydrophilic. The solid line empirically separates between IDPs and compact globular proteins. The panel is adapted from Ref. [5]. (B) The contribution of each amino acid in promoting disorder. Disordered propensity is evaluated from the fractional difference of amino acids composition of IDPs in the DisProt database and a completely ordered proteins from the protein-database (PDB). The panel is adapted from Ref. [12].

specificity?

Distinct from many structural proteins, IDPs can interact with multiple different partners. In analogy to the lock-and-key concept for structural proteins, IDPs serve as "master-keys", each capable of opening many different "locks" [1, 3, 7]. However, given IDP's prevalence, several other functions have been reported including cellular signaling pathways, a regulator for protein-protein and protein-DNA networks or folding into ordered structures upon binding to other cellular counterparts, to name a few (Fig. 1 and Refs. [15–18]).

In several cases a disorder-to-order transition of IDPs is the ensemble's conformational response to various stimuli. Such a response can lead to collapse or expansion of the IDP, e.g. globule-to-coil or collapse transitions [7, 10, 19]. Recently, high-speed AFM measurements demonstrated constantly folded and disordered regions as well as disorder-to-order transitions in IDPs [19].

Unraveling the factors governing the IDPs' ensemble conformation and their disorder-to-order transition is crucial to understand their biological functionalities [20]. An illustrative example is the regulation of glucose homeostasis by human pancreatic glucokinase (GCK) enzyme. GCK catalyzes glucose catabolism in the pancreas. At low glucose concentration, IDR of GCK associates with glucose and undergoes a disordered-to-ordered transition. Following glucose release, the IDR undergoes a order-to-disorder transition on millisecond time-scale. This "time delay loop" allows slow turnover and kinetic cooperativity of the enzyme. At high glucose concentration, the delay loop is bypassed and excess glucose is catabolized [21].

Obviously, the function of an IDP is driven by its amino-acid sequence, with

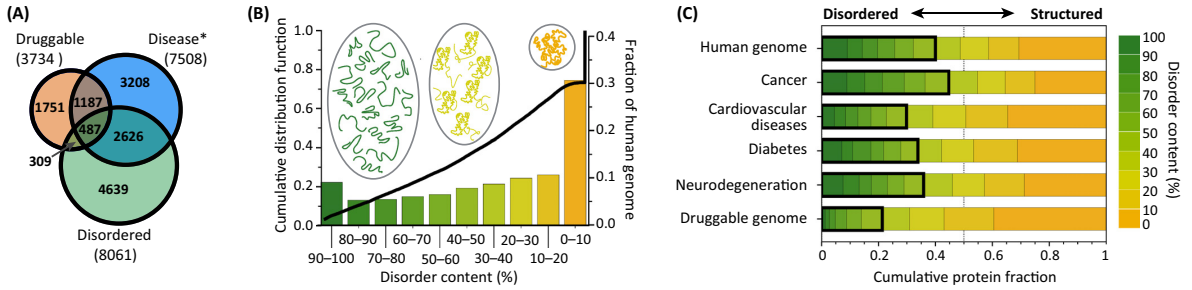

Figure 3. Disordered proteins link to human diseases. (A) Venn diagram of three types of proteins: those that interact with drugs; those that are related to disease; and those that are disordered. (B) The fraction of humans proteins encoded by the human genome (right axis) binned according to their content of structural disorder. The black line is the cumulative distribution function (left axis). (C) Amount of disorder in different protein categories. The figure is adapted from Ref. [22].

mutations resulting in numerous diseases (Fig. 3). Known examples are the disordered regions found in the Alzheimer and Parkinson associated proteins, Amyloid-β and α-synuclein, that can form toxic oligomers, amyloid fibrils, and other types of aggregates [23,24]. Other diseases, such as the recently emerged novel coronavirus (SARS-CoV-19), are intimately related to protein disorder. The SARS-CoV-19 genome sequence encodes an IDR in its nucleocapsid, an essential structural component that binds to RNA and interacts with several proteins in its multitasking role [25,26].

IDPs are also present in many biological receptors and play a functional role that is often poorly understood [27,28]. For example, the role of the disordered acidic domains of the Toc159 chloroplast preprotein, which binds to transit peptides, is still unknown [27]. Many RNA binding proteins (RBP) also contain disordered regions. There, the disordered domains play a central role in regulatory processes (e.g., stability) [29]. One fascinating class of IDPs comprises neuronal-specific cytoskeleton proteins, including neurofilament (NF) proteins, α-internexin, vimentin, microtubule-associated protein 2 (MAP2), and tau [6,30–34]. A key motif displayed by these proteins involves long IDRs mediating the assembly of filaments into a hydrogel network [6,34,35]. This network is responsible, for example, for defining the structure, size, and mechanics of axons, with direct effects on electrical conduction [31,34,36–38]. Interestingly, the basic structural properties of these networks can be understood using the polymer-physics theoretical arsenal. Nonetheless, the full functional behavior is sequence-specific, and so can not be coarse-grained [6,34,37].

To cope with the enormous traffic of molecular exchange across the nuclear envelope, nuclear pore complexes (NPCs) have evolved to exchange cargo rapidly and with high selectivity. NPCs are composed of 30 different nucleoporins that assemble into sub-complexes to build the NPC's distinct modules. Though it is not entirely clear how the cytoplasmic filaments are anchored to the NPC core, but the central, disordered segments of Nup145N, Nup100 and Nup116 are known integral parts of this interaction [39,40].

Recent work showed that IDRs could be primed for targeting. The early oncogene product E7 of the human papilloma virus HPV16 has an IDR that contains two transient helical pre-structured motifs that overlap with important target binding moieties (E2F-mimic motif and a pRb-binding LXCXE). These transient strictures can explain its target-binding behavior with multiple partners [41].

Recently, much attention has been drawn to membrane-less organelles, which are liquid-like condensates formed reversibly by dynamic self-assembly of proteins, mostly IDPs, and RNA through a liquid-liquid phase separation (LLPS) [14, 16, 42]. This behavior has raised many questions: Does the high density within such condensates enable certain protein functions? Is it the proteins' way of defending themselves against degradation? Or is LLPS just an unwanted outcome of the IDPs' sequence, which leads to transient interactions with multiple partners? The answers to these questions are most likely different for alternative organelles compositions.

There is little understanding of the IDP's sequence-encoded mechanism that drives this phase separation. Regardless, constituents of membrane-less organelles often contain multiple repetitive sequences, facilitating multivalent, weak interactions with their partners to form the condensates. As mentioned, such interactions are frequent for IDPs. The sizes of membrane-less organelles are in the range of sub μm to 10 μm and thus can be observed with optical microscopy. For example, the disordered regions of Ddx4 [43], LAF-1 [44], FUS [45], which are rich in glycine, form droplets of 1μm.

Even folded proteins that have IDRs, such as hnRNPA1 [46] and TDP-43 [47], form droplets, designated as stress granules [48]. Assembly and disassembly of these granules is a highly regulated process involving the IDR interactions with multiple proteins and mRNAs. Recently, using multi-bait engineered ascorbate peroxidase (APEX) proximity labeling technique, over a hundred proteins were discovered that either promote assembly or induced disassembly of stress granules [16]. Importantly, APEX revealed that most of these proteins are indeed disordered. Such studies are extremely relevant for in-depth understanding of the compositional changes in stress granules' proteins in neurodegenerative disorders.

3. Artificial IDPs and polymer physics

The ability of disordered proteins to engage in a variety of manners often results in a rich ensemble of phase transition behaviors [49, 50]. These phase behaviors can be easily manipulated by slightly changing the physio-chemical properties of the IDPs [51]. Recently, the Chilkoti group studied a simplified version of artificial IDPs, all made of a repetitive octamer peptide sequence originated from Drosophilia melanogaster Rec-1 resilin [52]. Interestingly, changing a single amino acid in the repeat, or increasing the number of repeats, results in a drastic and robust change in the LLPS temperature. The IDP sequence also controls the temperature-ramp hysteresis phenomena. Notably, this hysteresis changes when reversing the amino acid sequence originating from N- to C-terminal [53]. These findings demonstrate the ability to fine-tune the phase behaviors

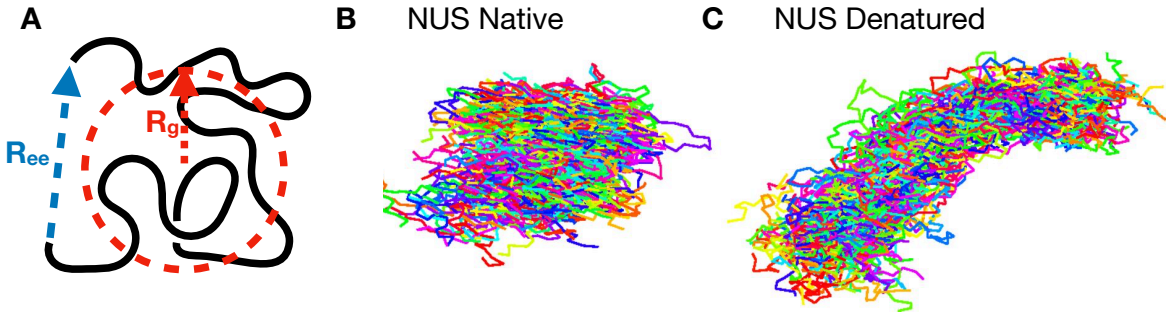

Figure 4. (A) Ensemble structural characterization of IDP that includes the end to end distance (R_{ee}), typically measured by FRET, and radius of gyration (R_g), measured by SAXS. Representative ensembles for NUS protein in its (B) native condition and (C) in denatured condition. Both ensembles recapture FRET and SAXS data at these conditions and show larger asphericity in denatured conditions. Panels (B) and (C) are adapted from Ref. [54].

and indicate the significance of IDP's sequence, as in structural proteins.

Given IDPs' structural plasticity, it is tempting to investigate them as polymers, at least to first order (Fig. 4). By doing so, the power of statistical-mechanics can be harnessed to quantify the volumetric dimension of IDPs. Then, coarse-grained structural descriptors are obtained from the scaling laws of polymer physics [55]. Specifically, the degree of IDP collapse or expansion can be quantified in terms of chain's length and its relation to the ensemble's radius of gyration: $R_g \propto N^{\nu}$. Here, N is the number of monomers, and ν is the Flory scaling exponent [6, 55].

Qualitatively, ν determines the balance between intra-chain and solvent-chain interactions, also referred to as solvent quality. For good ($\nu \approx 0.6$), θ (ν=0.5), or poor ($\nu \approx 0.3$) solvents the conditions are respectively less, equal, or favorable intra-chain interactions with respect to solvent/chain ones. Thus, the value of the scaling exponent ν can act as a physical descriptor for the collapse transition as a function of different environmental conditions [56]. In the following, we will introduce several techniques targeting the nanoscopic dimensions of IDPs.

4. Nanoscopic techniques

The detection and characterization of disordered proteins and regions are of great interest in order to determine their function. IDPs' structural plasticity dictate techniques and analyses targeting representative parameters of the protein's conformational ensemble [57]. Naturally, we will not be able to discuss the entire breadth of techniques; instead, we will focus on some of those targeting the nano-scale regime, which is, in most cases, the relevant length scale.

4.1. Simulations

To further understand the conformational dynamics of IDPs, advanced molecular dynamics (MD) simulations are used. However, the inherently large number of degrees-of-freedom, the inaccuracies in the simulation models, and the need for long simulations have stymied progress. To overcome these limitations, alternative approaches have been developed. For example, a hierarchical approach [58] based on the "flexible-meccano" model [59] can generate representative and meaningful starting ensembles. Alternatively, in the "sample-and-select" approach, sub ensembles are derived from a broader distribution of molecular conformations based on experimental NMR and small-angle X-ray scattering (SAXS) data [60–63].

A recent computational work [64] developed a related strategy: experimental SAXS data on a few select IDPs was used, through a fitting approach, to develop a set of sequence-specific residue interaction parameters in a coarse-grained simulation. These parameters were then used to investigate the conformational ensembles of a range of IDPs. The authors found the IDPs displayed apparent polymeric (swollen chain, $\nu = 0.6$) behavior when considering R_g vs. N. Yet, the simulated structural ensembles showed significant heterogeneity: certain IDPs showed elevated intra-segment contacts and transient sub-segment compactions. However, this was not universal: other IDPs in the study behaved as homogeneous random-walk chains. The authors concluded that sequence-specific structure could be hidden behind polymeric exponents [64].

4.2. SAXS

A common and powerful technique suited for IDP structural characterization is SAXS. The key advantage of the technique is the ability to measure the ensemble conformations while kept in solution without the need for crystallization or external tagging. A simple SAXS analysis immediately provides R_g and the pair distance distribution [65, 66]. Additionally, traditional Kratky plots ($q^2I(q)$ vs. q), directly obtained from the SAXS profiles of IDPs, provide a qualitative picture of the presence of globular or unfolded conformations [67, 68]. Here, I is the scattering intensity, and q is the scattered wave-vector proportional to the scattering angle.

Moreover, the scattering profile can be used to extract the Flory exponent (ν) mentioned above [54, 69]. More advanced analysis techniques are regularly used to extract the ensemble's dominant conformations from the SAXS signal (Fig. 4). The ensemble optimization method [70, 71], and molecular form-factor [56] are just a few examples for such techniques resulting in representative conformations that fit the experimental data. It is worthwhile to note that the combination of SAXS with other complementary methods (Fig. 4) such as simulations, NMR, and Förster resonance energy transfer (FRET) provides a more comprehensive picture of IDPs, and their conformations in solution [6, 37, 54, 57, 65, 72, 73].

4.3. FRET

The distance-dependent dipole-dipole interactions between probes bound to the side chains of IDPs provide the basis for determining long-range intra-molecular distances between selected sites. Experiments based on single-molecule [74] or ensemble [75] measurements are characterized by $1-10$ nm distance resolution range and an ability to recover distributions of intra-molecular distances in the transient ensembles of IDP. For example, it was shown that α-synuclein deviates from ideal chain behavior at segments labeled in the NAC domain and N-termini using the ensemble method. It was suggested that those conformations bias might be related to the initiation of amyloid transition [75]. Another example, using a single-molecule method, showed that the previously mentioned scaling exponents (ν) in water strongly depend on the sequence composition. Two of the total examined IDPs did not reach the θ-point under any solvent conditions. This may reflects their biological function need for an expanded state optimized for interactions with cellular partners [74].

4.4. NMR

NMR spectroscopy offers a unique platform in deciphering IDP dynamics and structure [76]. The traditional NMR hydrogen chemical shift signal is, in many cases, insufficient for IDP characterization. Fortunately, new tools have been developed to overcome low signal to noise ratio difficulties such as paramagnetic relaxation enhancement (PRE) [77] and proline based 2D H-N residue correlations [78].

In PRE, the introduction of paramagnetic spin labels in proteins affects the chemical shift and transverse relaxation rate signal between the unpaired electron and NMR active nuclei on the basis of the distance between them [77]. For example, the PRE signal and ^{15}N relaxation data were analyzed to quantify the interaction between the IDP osteopontin and heparin [79]. There, it was found that on binding with heparin, osteopontin largely remains in a disordered state and undergoes structural/dynamical adaption which is mainly mediated by electrostatic interactions.

A recent development using in-cell NMR spectroscopy provides the opportunity to explore the structural plasticity of an IDP in its native environment [80, 81]. Both eukaryotic and prokaryotic IDPs have been investigated using in-cell NMR such as α-synuclein, prokaryotic ubiquitin-like protein, Pup, tubulin-related neuronal protein, tau, FG-Nups in nucleoporins, and the negative regulator of flagellin synthesis, FlgM. These studies revealed that the cellular conformational dynamics may differ significantly from those observed in-vitro [82].

4.5. FCS

Due to diffusion and the protein structural plasticity, the emission of fluorescently-labeled molecules fluctuate. In fluorescence correlation spectroscopy (FCS), these fluctuations are recorded within an illuminated confocal volume [83–85]. The timescale

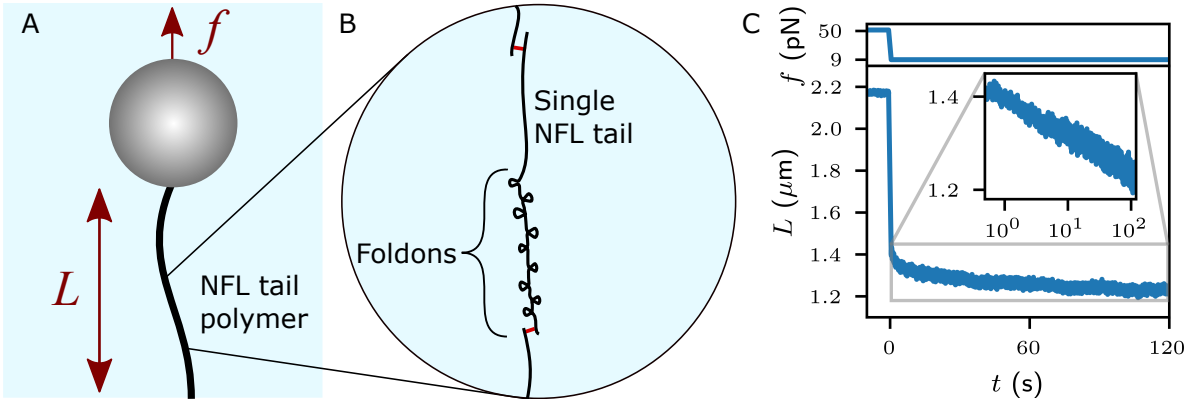

Figure 5. Mechanical perturbation reveals subtle internal IDP structure. (A) Under constant tension f, the length L of a polymer of NFLt is measured. (B) A single NFLt contains multiple, independent regions with internal structure (foldons). (C) An initial large tension breaks apart internal structure in the NFL tails. After rapidly lowering the tension, at constant tension, the NFL polymer shows a logarithmic decrease in length over time t, indicating multiple regions of internal structure are reforming. Panel (C) is adapted from Ref. [88].

for conformational fluctuation (i.e., chemical kinetics within the biomolecule) lies in the nano- to microseconds range, whereas for translational diffusion from microsecond to milliseconds. Information about the diffusion coefficient and chemical kinetics can be inferred from the fluorescence fluctuations autocorrelation and cross-correlation signal. Furthermore, from the diffusion coefficient, it is possible to determine the molecular size and hydrodynamic radius of the investigated biomolecules.

FCS is also sensitive to fluorophore quenching reactions. These measurements can provide information about the internal dynamics of the IDPs [86]. Fluorescence self-quenching of tetramethyl rhodamine, which is chemically labeled at two different residual positions, was analyzed to study the conformational kinetics of unfolded intestinal fatty acid-binding protein [87] and amyloid-forming yeast prion protein, Sup35 [86].

4.6. Single-molecule force spectroscopy

Single-molecule force spectroscopy (SMFS) uses the thermodynamic effects of applied tension to gain insight into the nanoscopic conformation of IDPs. SMFS consists of tethering an IDP between a static surface and a force probe and measuring the chain extension, with nanometer accuracy, as a function of force. This nanomechanical assay is a high-precision differential measurement of the effect of varying perturbations (tensions) on IDP interactions and conformations. For example, AFM-based experiments on several amyloid precursor IDPs show a sawtooth pattern that indicates the mechanical unfolding of multiple different conformations [89]. Similarly, optical tweezer experiments on α-synuclein show it has several marginally stable and rapidly fluctuating subsegments [90].

A noteworthy recent work demonstrated the power of SMFS-based perturbation

in revealing more subtle IDP structural behaviors [88]. In the work, a polyprotein of the disordered Neurofilament light-chain tail (NFLt) domain was subjected to force-quench experiments, in which a large tension pulled the chain straight, and was followed by a sudden jump to a small tension that permitted structure in the IDP to re-form (Fig. 5). The experiment was carried out in a magnetic tweezer, an SMFS instrument that offers high-stability in applied force, allowing the tracking of NFLt dynamics over long periods. The experiment revealed a glassy, non-exponential relaxation in which the chain extension decreased logarithmically with time for many minutes. This behavior, extraordinary for its slow speed, was attributed to multiple, small collapse events within a single NFLt domain, with each collapse changing chain extension by only ≈ 1 nm. This experiment highlights both the high sensitivity of SMFS, as well as the rich and physically-complex structural dynamics of IDPs.

4.7. IDP imaging and quantification of their inter-molecular interaction

Intermolecular interactions of biomolecules play an essential role in the proper function and growth of biological cells [91–93]. These interactions can be specific or non-specific. Non-specific interactions are transient and weak, mainly governed by steric repulsion or ionic bridging. IDPs' function is largely dominated by these weak interactions. This was demonstrated, for example, in a recent AFM-based SMFS study, in which the interactions between nuclear transport factors and disordered FG repeats were found to display unexpected complexities due to weak, multivalent contacts between the interacting partners [94]. Moreover, using cryogenic electron microscopy, the collective behavior of IDPs inside an artificial DNA origami ring was imaged [95]. These results mimic the NPCs environment to show their structural and transport properties.

Improved protein-protein interaction resolution can be achieved using bulk, low-throughput spectroscopic methods such as FRET, surface plasmon resonance (SPR), and nuclear magnetic resonance (NMR), to name only a few. For example, the interaction between the disordered region in transcription factors Sp1 and TAF4 was estimated by NMR to be $K_d = 69\mu$M [96]. IDP Ntr2 modulates the RNA Helicase Brr2 with $K_d = 14\mu$M and 7μM determined by SPR and NMR, respectively [97]. SPR was also used to detect and characterize IDP–ligand interactions with tau protein as a model IDP [98].

Another recently introduced technique aiming to overcome the technical difficulties to measure weak and transient protein-protein interaction is the nanoparticle mobility assay [13]. There, nanoparticles grafted with IDPs were imaged while diffusing over a surface grafted with a second set of IDPs. A similar approach was used to study strong DNA-DNA interactions [99], delocalized long-range polymer-surface interactions [100] and bulk-mediated diffusion on supported lipid-bilayers [101]. By carefully analyzing the particles' diffusive nature, the authors detected an altered interaction caused by a single mutation on the polypeptide sequence [13].

5. Summary and perspective

Two decades of research places IDPs at the forefront of proteomics [1]. Apparently, structural plasticity is valuable primarily because it enables the IDPs to adjust to their environment. In addition, IDPs are also used in various nano-biomedical technologies. For example, *PASylation* technology involves conjugating pharmacologically active compounds with natively disordered biosynthetic polymers made of the small L-amino acids Proline, Alanine and/or Serine [102]. This technology may serve as an alternative to PEGylation, which is currently widely used to sterically stabilized and prevent aggregation of nanoparticles or therapeutic proteins and peptides but is often accompanied by immunogenicity and lack of biodegradability [103, 104]. IDP have the benefit of possible interactions with many alternative binding partners. Therefore, unlike specific folded proteins, incorporating IDPs may help to target multiple binding partnerss as ligands attached to nanostructures. IDPs can also serve to modify the adhesion properties of surfaces. For example, a study on barnacle adhesive IDPs on silica surfaces showed modified protein-surface affinities, adhesive activities and kinetic adsorption rates [105].

Proteins are increasingly proposed as suitable candidates for the fabrication of biological computers and other biomolecular electronic devices, mainly due to their interesting structure-related intrinsic electrical properties [106, 107]. So far, mostly ordered proteins have been proposed, but IDP can also be used for "proteotronics". For example, in analogy to probing DNA polymerase activity between a pair of electrodes, one could envision a similar system with IDPs capable of binding to alternative partners [108].Recent examples demonstrated that bio-nanowires could be fabricated from peptides, permitting long-range electron transport [109]. Here, we imagine that the structural plasticity of IDPs could be used to connect nearby wires through self-assembly, or the IDP order-disorder transition could act as an electronic switch.

IDPs' large repertoire has been implicated in numerous diseases, making them potential targets for therapeutic intervention. Disease-associated IDPs can serve as potential targets for drugs modulating protein-protein interaction networks that participate in both the "one to many" and "many to one" interaction [110]. The relation between IDPs and polymer- and statistical-physics is continuously revisited. Naïvely, IDP *function* could have been understood from coarse metrics such as the IDP's length, total charge, net hydrophobicity, etc. However, as with folded proteins, the sequence of amino acids, and not just the net composition, does dictate the resulting function, at least to some extent. Yet building a coarse-grained model for IDPs is a highly complicated and fascinating problem due to their unique properties. The most fundamental difficulty is somewhat technical - the large ensemble of IDP conformations is often impossible to sample with modern computers [111]. Another problem evolves from the extended amount of electrostatic interactions that frequently exist in IDP systems. Those interactions tend to span over a broad range of length and time scales, and hence are hard to quantify by a finite force field [112], although new

IDP-specific force fields have been suggested [113–116]. Nevertheless, further research is highly needed, especially since many phenomena involving IDPs are still not fully understood [117].

Acknowledgments

This work was supported by the National Science Foundation under Grant No. 1715627, the United States-Israel Bi-national Science Foundation under Grant No. 2016696, and the Israel Science Foundation under Grants No. 1454/19.

References

[1] Wright P E and Dyson H J 1999 *Journal of Molecular Biology* **293** 321–331
[2] Uversky V N 2014 *Chemical Reviews* **114** 6557–6560
[3] Dunker A K, Lawson J D, Brown C J, Williams R M, Romero P, Oh J S, Oldfield C J, Campen A M, Ratliff C M, Hipps K W *et al.* 2001 *Journal of Molecular Graphics and Modelling* **19** 26–59
[4] Uversky V N and Dunker A K 2010 *Biochimica Et Biophysica Acta-Proteins and Proteomics* **1804** 1231–1264
[5] Uversky V N 2011 *The International Journal of Biochemistry & Cell Biology* **43** 1090–1103
[6] Kornreich M, Avinery R, Malka-Gibor E, Laser-Azogui A and Beck R 2015 *FEBS letters* **589** 2464–76
[7] Uversky V N 2019 *Frontiers in Physics* **7** 10
[8] Uversky V N, Gillespie J R and Fink A L 2000 *Proteins: Structure, Function, and Bioinformatics* **41** 415–427
[9] Ferron F, Longhi S, Canard B and Karlin D 2006 *Proteins: Structure, Function, and Bioinformatics* **65** 1–14
[10] Das R K, Ruff K M and Pappu R V 2015 *Current opinion in structural biology* **32** 102–112
[11] Mao A H, Lyle N and Pappu R V 2013 *Biochemical Journal* **449** 307–318
[12] Radivojac P, Iakoucheva L M, Oldfield C J, Obradovic Z, Uversky V N and Dunker A K 2007 *Biophysical Journal* **92** 1439–1456
[13] Chakraborty I, Rahamim G, Avinery R, Roichman Y and Beck R 2019 *Nano Letters* **19** 6524–6534
[14] Shin Y and Brangwynne C P 2017 *Science* **357**
[15] Wright P E and Dyson H J 2015 *Nature Reviews Molecular Cell Biology* **16** 18–29
[16] Marmor Kollet H, Siany A, Kedersha N, Knafo N, Rivkin N, Danino Y M, Olender T, Cohen N, Moens T, Higginbottom A *et al.* 2020 *Molecular Cell*
[17] Dunker A K, Cortese M S, Romero P, Iakoucheva L M and Uversky V N 2005 *The FEBS journal* **272** 5129–5148
[18] Gianni S, Dogan J and Jemth P 2016 *Current Ppinion in Structural Biology* **36** 18–24
[19] Kodera N, Noshiro D, Dora S K, Mori T, Habchi J, Blocquel D, Gruet A, Dosnon M, Salladini E, Bignon C *et al.* 2020 *Nature Nanotechnology* 1–9
[20] Zhou J, Oldfield C J, Yan W, Shen B and Dunker A K 2019 *Protein Science* **28** 1652–1663
[21] Larion M, Salinas R K, Bruschweiler-Li L, Miller B G and Brüschweiler R 2012 *PLoS Biology* **10** e1001452
[22] Heller G T, Sormanni P and Vendruscolo M 2015 *Trends in Biochemical Sciences* **40** 491–496
[23] Carballo-Pacheco M and Strodel B 2017 *Protein Science* **26** 174–185
[24] Coskuner-Weber O and Uversky V N 2018 *International journal of molecular sciences* **19** 336
[25] Ceraolo C and Giorgi F M 2020 *Journal of Medical Virology* **92** 522–528

[26] Cubuk J, Alston J J, Incicco J J, Singh S, Stuchell-Brereton M D, Ward M D, Zimmerman M I, Vithani N, Griffith D, Wagoner J A, Bowman G R, Hall K B, Soranno A and Holehouse A S 2020 *bioRxiv*
[27] Richardson L G, Jelokhani-Niaraki M and Smith M D 2009 *BMC Biochemistry* **10** 1–8
[28] Chino H, Hatta T, Natsume T and Mizushima N 2019 *Molecular Cell* **74** 909–921
[29] Basu S and Bahadur R P 2016 *Cellular and Molecular Life Sciences* **73** 4075–4084
[30] Pregent S, Lichtenstein A, Avinery R, Laser-Azogui A, Patolsky F and Beck R 2015 *Nano letters* **15** 3080–7
[31] Laser-Azogui A, Kornreich M, Malka-Gibor E and Beck R 2015 *Current Opinion in Cell Biology* **32** 92–101
[32] Safinya C R, Deek J, Beck R, Jones J B, Leal C, Ewert K K and Li Y 2013 *Liquid crystals* **40** 1748–1758
[33] Guharoy M, Szabo B, Martos S C, Kosol S and Tompa P 2013 *Cytoskeleton* **70** 550–571
[34] Chung P J, Song C, Deek J, Miller H P, Li Y, Choi M C, Wilson L, Feinstein S C and Safinya C R 2016 *Nature Communications* **7** 12278
[35] Kornreich M, Malka-Gibor E, Zuker B, Laser-Azogui A and Beck R 2016 *Physical Review Letters* **117** 148101
[36] Chang L and Goldman R D 2004 *Nature Reviews Molecular Cell Biology* **5** 601–613
[37] Malka-Gibor E, Kornreich M, Laser-Azogui A, Doron O, Zingerman-Koladko I, Harapin J, Medalia O and Beck R 2017 *Biophysical Journal* **112** 892–900 (*Preprint* `1609.05546`)
[38] Beck R, Deek J, Jones J B and Safinya C R 2010 *Nature Materials* **9** 40–46
[39] Schwartz T U 2016 *Journal of Molecular Biology* **428** 1986–2000
[40] Lemke E A 2016 *Journal of Molecular Biology* **428** 2011–2024
[41] Lee C, Kim D H, Lee S H, Su J and Han K H 2016 *BMB reports* **49** 431
[42] Brangwynne C P, Eckmann C R, Courson D S, Rybarska A, Hoege C, Gharakhani J, Julicher F and Hyman A A 2009 *Science* **324** 1729–1732
[43] Nott T J, Petsalaki E, Farber P, Jervis D, Fussner E, Plochowietz A, Craggs T D, Bazett-Jones D P, Pawson T, Forman-Kay J D and Baldwin A J 2015 *Molecular Cell* **57** 936–947
[44] Elbaum-Garfinkle S, Kim Y, Szczepaniak K, Chen C C H, Eckmann C R, Myong S and Brangwynne C P 2015 *Proceedings of the National Academy of Sciences* **112** 7189–7194
[45] Patel A, Lee H O, Jawerth L, Maharana S, Jahnel M, Hein M Y, Stoynov S, Mahamid J, Saha S, Franzmann T M, Pozniakovski A, Poser I, Maghelli N, Royer L A, Weigert M, Myers E W, Grill S, Drechsel D, Hyman A A and Alberti S 2015 *Cell* **162** 1066–1077
[46] Molliex A, Temirov J, Lee J, Coughlin M, Kanagaraj A P, Kim H J, Mittag T and Taylor J P 2015 *Cell* **163** 123–133
[47] Conicella A E, Zerze G H, Mittal J and Fawzi N L 2016 *Structure* **24** 1537–1549
[48] Buchan J R and Parker R 2009 *Molecular Cell* **36** 932–941
[49] Dignon G L, Zheng W, Kim Y C, Best R B and Mittal J 2018 *PLoS Computational Biology* **14** e1005941
[50] Zhou H X, Nguemaha V, Mazarakos K and Qin S 2018 *Trends in Biochemical Sciences* **43** 499–516
[51] Jacoby G, Asher M S, Ehm T, Ionita I A, Shinar H, Azoulay-Ginsburg S, Danino D, Kozlov M M, Amir R J and Beck R 2021 Order from disorder with intrinsically disordered peptide amphiphiles (*Preprint* `arXiv 2103.10131`)
[52] Dzuricky M, Rogers B A, Shahid A, Cremer P S and Chilkoti A 2020 *Nature Chemistry* **12** 814–825
[53] Quiroz F G, Li N K, Roberts S, Weber P, Dzuricky M, Weitzhandler I, Yingling Y G and Chilkoti A 2019 *Science Advances* **5** eaax5177
[54] Fuertes G, Banterle N, Ruff K M, Chowdhury A, Mercadante D, Koehler C, Kachala M, Girona G E, Milles S, Mishra A *et al.* 2017 *Proceedings of the National Academy of Sciences* **114** E6342–E6351
[55] Wang Z G 2017 *Macromolecules* **50** 9073–9114

[56] 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 and Sosnick T R 2017 *Science* **358** 238–241
[57] Lazar T, Martínez-Pérez E, Quaglia F, Hatos A, Chemes L B, Iserte J A, Méndez N A, Garrone N A, Saldaño T E, Marchetti J, Rueda A J V, Bernadó P, Blackledge M, Cordeiro T N, Fagerberg E, Forman-Kay J D, Fornasari M S, Gibson T J, Gomes G N W, Gradinaru C C, Head-Gordon T, Jensen M R, Lemke E A, Longhi S, Marino-Buslje C, Minervini G, Mittag T, Monzon A M, Pappu R V, Parisi G, Ricard-Blum S, Ruff K M, Salladini E, Skepö M, Svergun D, Vallet S D, Varadi M, Tompa P, Tosatto S C E and Piovesan D 2020 *Nucleic Acids Research*
[58] Pietrek L M, Stelzl L S and Hummer G 2019 *Journal of Chemical Theory and Computation* **16** 725–737
[59] Bernadó P, Bertoncini C W, Griesinger C, Zweckstetter M and Blackledge M 2005 *Journal of the American Chemical Society* **127** 17968–17969
[60] Jensen M R, Zweckstetter M, Huang J r and Blackledge M 2014 *Chemical Reviews* **114** 6632–6660
[61] Fisher C K and Stultz C M 2011 *Current Opinion in Structural Biology* **21** 426–431
[62] Mittag T and Forman-Kay J D 2007 *Current Opinion in Structural Biology* **17** 3–14
[63] Allison J R, Varnai P, Dobson C M and Vendruscolo M 2009 *Journal of the American Chemical Society* **131** 18314–18326
[64] Baul U, Chakraborty D, Mugnai M L, Straub J E and Thirumalai D 2019 *The Journal of Physical Chemistry B* **123** 3462–3474
[65] Kornreich M, Avinery R and Beck R 2013 *Current Opinion in Biotechnology* **24** 716–723
[66] Manalastas-Cantos K, Konarev P, Hajizadeh N R, Kikhney A, Petoukhov M, Molodenskiy D, Panjkovich A, Mertens H, Gruzinov A, Borges C *et al.* 2020 *Journal of Applied Crystallography* **54**
[67] Bernadó P and Svergun D I 2012 *Molecular Biosystems* **8** 151–167
[68] Almagor L, Avinery R, Hirsch J A and Beck R 2013 *Biophysical Journal* **104** 2392–2400
[69] Hammouda B 2012 *Macromolecular Theory and Simulations* **21** 372–381
[70] Bernadó P, Mylonas E, Petoukhov M V, Blackledge M and Svergun D I 2007 *Journal of the American Chemical Society* **129** 5656–5664
[71] Tria G, Mertens H D, Kachala M and Svergun D I 2015 *International Union of Crystallography Journal* **2** 207–217
[72] Storm I M, Kornreich M, Hernandez-Garcia A, Voets I K, Beck R, Cohen Stuart M A, Leermakers F A and De Vries R 2015 *The Journal of Physical Chemistry B* **119** 4084–4092
[73] Storm I M, Kornreich M, Voets I K, Beck R, de Vries R, Stuart M A C and Leermakers F A 2016 *Soft Matter* **12** 8004–8014
[74] Hofmann H, Soranno A, Borgia A, Gast K, Nettels D and Schuler B 2012 *Proceedings of the National Academy of Sciences of the United States of America* **109** 16155–60
[75] Grupi A and Haas E 2011 *Journal of Molecular Biology* **405** 1267–1283
[76] Konrat R 2014 *Journal of Magnetic Resonance* **241** 74–85
[77] Otting G 2010 *Annual review of biophysics* **39** 387–405
[78] Solyom Z, Schwarten M, Geist L, Konrat R, Willbold D and Brutscher B 2013 *Journal of Biomolecular NMR* **55** 311–321
[79] Platzer G, Schedlbauer A, Chemelli A, Ozdowy P, Coudevylle N, Auer R, Kontaxis G, Hartl M, Miles A J, Wallace B A *et al.* 2011 *Biochemistry* **50** 6113–6124
[80] Luchinat E and Banci L 2016 *Journal of Biological Chemistry* **291** 3776–3784
[81] Barbieri L, Luchinat E and Banci L 2016 *Nature Protocols* **11** 1101–1111
[82] Sciolino N, Burz D S and Shekhtman A 2019 *Proteomics* **19** 1800055
[83] Magde D, Elson E and Webb W W 1972 *Physical Review Letters* **29** 705
[84] Elson E L and Magde D 1974 *Biopolymers: Original Research on Biomolecules* **13** 1–27
[85] Elson E L 2011 *Biophysical Journal* **101** 2855–2870
[86] Mukhopadhyay S, Krishnan R, Lemke E A, Lindquist S and Deniz A A 2007 *Proceedings of the National Academy of Sciences* **104** 2649–2654

[87] Chattopadhyay K, Elson E L and Frieden C 2005 *Proceedings of the National Academy of Sciences* **102** 2385–2389
[88] Morgan I L, Avinery R, Beck R and Saleh O A 2020 *Biophysical Journal* **118** 337a–338a
[89] Hervas R, Oroz J, Galera-Prat A, Goni O, Valbuena A, Vera A M, Gomez-Sicilia A, Losada-Urzaiz F, Uversky V N, Menendez M, Laurents D V, Bruix M and Carrion-Vazquez M 2012 *PLoS Biol* **10** e1001335 ISSN 1545-7885 (Electronic) 1544-9173 (Linking) URL `https://www.ncbi.nlm.nih.gov/pubmed/22666178`
[90] Solanki A, Neupane K and Woodside M T 2014 *Phys Rev Lett* **112** 158103 ISSN 1079-7114 (Electronic) 0031-9007 (Linking) URL `https://www.ncbi.nlm.nih.gov/pubmed/24785077`
[91] Leckband D and Israelachvili J 2001 *Quarterly Reviews of Biophysics* **34** 105
[92] Wyttenbach T and Bowers M T 2007 *Annu. Rev. Phys. Chem.* **58** 511–533
[93] Musiani F and Giorgetti A 2017 Chapter two - protein aggregation and molecular crowding: Perspectives from multiscale simulations *Early Stage Protein Misfolding and Amyloid Aggregation* (*International Review of Cell and Molecular Biology* vol 329) ed Sandal M (Academic Press) pp 49 – 77
[94] Hayama R, Sorci M, Keating IV J J, Hecht L M, Plawsky J L, Belfort G, Chait B T and Rout M P 2019 *PLOS ONE* **14** e0217897
[95] Ketterer P, Ananth A N, Trip D S L, Mishra A, Bertosin E, Ganji M, van der Torre J, Onck P, Dietz H and Dekker C 2018 *Nature communications* **9** 1–8
[96] Hibino E, Inoue R, Sugiyama M, Kuwahara J, Matsuzaki K and Hoshino M 2016 *Protein Science* **25** 2006–2017
[97] Wollenhaupt J, Henning L M, Sticht J, Becke C, Freund C, Santos K F and Wahl M C 2018 *Biophysical Journal* **114** 788–799
[98] Vagrys D, Davidson J, Chen I, Hubbard R E and Davis B 2020 *International Journal of Molecular Sciences* **21** 5257
[99] Xu Q, Feng L, Sha R, Seeman N C and Chaikin P M 2011 *Physical Review Letters* **106** 228102
[100] Skaug M J, Mabry J N and Schwartz D K 2014 *Journal of the American Chemical Society* **136** 1327–1332
[101] Yoo J, Lee T S, Choi B, Shon M J and Yoon T Y 2016 *Journal of the American Chemical Society* **138** 14238–14241
[102] Binder U and Skerra A 2017 *Current Opinion in Colloid & Interface Science* **31** 10–17
[103] Gebauer M and Skerra A 2018 *Bioorganic & Medicinal Chemistry* **26** 2882–2887
[104] Schlapschy M, Binder U, Börger C, Theobald I, Wachinger K, Kisling S, Haller D and Skerra A 2013 *Protein Engineering, Design & Selection* **26** 489–501
[105] Wang X, Wang C, Xu B, Wei J, Xiao Y and Huang F 2018 *Applied Surface Science* **427** 942–949
[106] Alfinito E, Pousset J and Reggiani L 2015 Proteotronics: The emerging science of protein-based electronic devices *Journal of Physics: Conference Series* vol 647(1) (IOP Publishing) p 012002
[107] Talebi S, Daraghma S M, Subramaniam R T, Bhassu S, Gnana Kumar G and Periasamy V 2020 *ACS omega* **5** 7802–7808
[108] Zhang B, Deng H, Mukherjee S, Song W, Wang X and Lindsay S 2019 *ACS nano* **14** 1360–1368
[109] Ing N L, El-Naggar M Y and Hochbaum A I 2018 *The Journal of Physical Chemistry B* **122** 10403–10423
[110] Wang J, Cao Z, Zhao L and Li S 2011 *International Journal of Molecular Sciences* **12** 3205–3219
[111] Csizmok V, Follis A V, Kriwacki R W and Forman-Kay J D 2016 *Chemical Reviews* **116** 6424–6462
[112] Cragnell C, Rieloff E and Skepö M 2018 *Journal of Molecular Biology* **430** 2478–2492
[113] Knott M and Best R B 2014 *The Journal of Chemical Physics* **140** 05B603_1
[114] Sigalov A B 2016 *Biochimie* **125** 112–118
[115] Zhao Y, Cortes-Huerto R, Kremer K and Rudzinski J F 2020 *The Journal of Physical Chemistry B* **124** 4097–4113
[116] Schuler B, Soranno A, Hofmann H and Nettels D 2016 *Annual Review of Biophysics* **45** 207–231

[117] Bhattacharya S and Lin X 2019 *Biomolecules* **9** 146