Seminars in Cell & Developmental Biology 37 (2015) 26–34

Contents lists available at ScienceDirect

Seminars in Cell & Developmental Biology

journal homepage: www.elsevier.com/locate/semcdb

Review

Probing protein disorder and complexity at single-molecule resolution

Taehyung Lee, Crystal R. Moran-Gutierrez, Ashok A. Deniz *

Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, United States

ARTICLE INFO

Article history:
Available online 8 October 2014

Keywords:
Intrinsically disordered proteins
Single-molecule
Protein Folding
Protein dynamics
Allostery

ABSTRACT

A substantial fraction of the human proteome encodes disordered proteins. Protein disorder is associated with a variety of cellular functions and misfunction, and is therefore of clear import to biological systems. However, disorder lends itself to conformational flexibility and heterogeneity, rendering proteins which feature prominent disorder difficult to study using conventional structural biology methods. Here we discuss a few examples of how single-molecule methods are providing new insight into the biophysics and complexity of these proteins by avoiding ensemble averaging, thereby providing direct information about the complex distributions and dynamics of this important class of proteins. Examples of note include characterization of isolated IDPs in solution as collapsed and dynamic species, detailed insight into complex IDP folding landscapes, and new information about how tunable regulation of structure-mediated binding cooperativity and consequent function can be achieved through protein disorder. With these exciting advances in view, we conclude with a discussion of a few complementary and emerging single-molecule efforts of particular promise, including complementary and enhanced methodologies for studying disorder in proteins, and experiments to investigate the potential role for IDP-induced phase separation as a critical functional element in biological systems.

Contents

1. Introduction

Proteins are involved in myriad cellular and developmental roles, including architecture, chemical reactions, selective transport across biological membranes, and interaction and regulation of biomolecular networks and signaling cascades. To date, static 3D structural characterization of large ensembles of highly ordered proteins has dominated structural biology, and has provided much insight into protein function. Despite this success, intrinsic disorder is now understood to be a critical and ubiquitous contributor to protein function, leading to a substantial revision in the classic 3 Dstructure–function paradigm, and highlighting the need for investigational approaches not limited to well-behaved and structurally robust proteins [1–6]. Biophysicists have long recognized that to a greater or lesser extent, proteins are in general dynamic and flexible species. However, intrinsically disordered regions in proteins, whether local or global (IDRs and IDPs respectively), encode a much greater degree of these features, and require both new perspective

* Corresponding author.
E-mail address: deniz@scripps.edu (A.A. Deniz).

http://dx.doi.org/10.1016/j.semcdb.2014.09.027
1084-9521/

and new tools for successful investigation. The physics of this disorder could confer a number of biologically significant functional advantages on these systems, and therefore not only require, but merit careful study. Additionally, a number of disease-linked amyloid forming proteins are disordered in their monomeric-unbound states, suggesting a potential and important link between disorder and aberrant misfolding. Therefore, a detailed biophysical understanding of these paradigm-shifting proteins is important for both fundamental protein science and a more precise understanding of cellular function and disease, despite the inherent challenges in studying such conformationally complex and dynamic species.

Expanding the experimental potential for understanding IDP biophysics has been a significant opportunity afforded through some exciting advances in single-molecule detection methods over the past few decades [7–10]. Capitalizing on improvements in relevant technologies, biophysical single-molecule experiments based on force, fluorescence and other methods began appearing in the 1980s [11–16]. These methods fundamentally altered our views of molecular complexity and opened the door to more direct tests of mechanistic models by avoiding the averaging and loss of information that are necessary in ensemble experiments to achieve high signal-to-noise data. Single-molecule methods have already been used to probe the complex conformational distributions, dynamics, interactions, and aggregation propensities of IDPs, with much success. Early application of single-molecule techniques to IDPs began appearing in the literature in the mid-late 2000s, with investigation of conformational features, dynamics and interactions of amyloidogenic IDPs. Also, and of particular note for aggregation-prone members of this protein class, single-molecule experiments utilize very low molecular concentrations, avoiding the confounding effect of unwanted aggregation or molecular interaction. Several studies have followed since on these and other types of IDPs, and have broadened our understanding of the biophysics of proteins and the systems in which they function.

Discussed below is a sampling of some of the important biological questions being answered with single-molecule experiments, presented in three broad classes of structural and functional complexity: (i) the conformational features and dynamics of *monomeric* IDPs, (ii) interaction of IDPs with *binding partners* and concomitant folding, and (iii) more complex behavior of IDPs, with a specific focus on *binding-modulated function* by interaction with multiple partners. Biophysical features at each of these levels are expected to offer critical insight into biological function, and single-molecule investigation is helping to shed light on each of these levels of molecular and folding complexity. Lastly, we also discuss a few complementary and emerging directions for the utility of single-molecule methods in the effort to study disordered protein systems.

2. Structural features and dynamics of monomeric IDPs

Investigation of protein disorder begins conceptually with intrinsic structural propensity in monomers, arising from a defining sequence of amino acids. Higher order interactions and structural features can be thought of as functions of this basic state.

In an elegant 2006 study, polyglutamine ("poly-Q") was investigated by Crick et al. [17] as a model of the Huntington's disease-causing protein, huntingtin, using fluorescence correlation spectroscopy (FCS) to determine the scaling relationship between poly-Q chain length and molecular diffusion times. FCS is a near-single-molecule resolution method to measure and analyze fluorescence fluctuations in a subfemtoliter detection volume (achieved through confocal detection) from molecules diffusing freely in solution [18]. Fluorescence intensity data are subjected to correlation analysis to identify molecular events ranging from molecular diffusion (as relatively slow decays) to conformational dynamics at very rapid time scales.

Poly-Q was studied by labeling the cysteine residue in Gly-Gln_N-Cys-Lys_2 peptides (where N = chain length) with the bright fluorescent dye Alexa 488 by a maleimide moiety. These experiments showed a monotonic increase in diffusion time with chain length, with no perceptible change in the trend at or around the disease-critical tract length of $N = 35$. Even more significantly from a polymer physics perspective, the scaling of diffusion times as a function of chain length revealed a slope ν of 0.32 ± 0.02 (Fig. 1A), which indicated a polymer in poor solvent. From this result, the authors concluded that poly-Q is poorly solvated and compacted in aqueous solution, an unexpected result for a protein with minimal hydrophobicity.

Soon after these results were published, single-molecule investigation of the yeast protein Sup35 corroborated compaction as a feature in IDPs, but did so via direct measurement of intramolecular distance across its amyloid-determining NM region (for *N*-terminus and *m*iddle), which is disordered in the native state [19]. Förster resonance energy transfer at single-molecule resolution (termed smFRET), is a powerful method for studying structural features in biological molecules, and is especially well-suited for investigating conformationally heterogeneous IDPs. In smFRET, energy is transferred in a non-radiative and distance-dependent manner between appropriate donor and acceptor fluorophores [9,20,21]. In a diffusing format, fluorescence intensities are recorded for each molecule that traverses a subfemtoliter detection volume and FRET efficiency (E_{FRET}) values are calculated, providing a sensitive measure of inter-dye distance across molecular distances, and an indication of the protein's conformational state.

In the NM experiments, thousands of protein monomers labeled with amyloid region-flanking donor and acceptor dyes were examined individually using smFRET. The resulting E_{FRET} values, plotted as histograms, revealed a population enriched in high E_{FRET} states with a mean value of 0.8 (Fig. 1B), which corresponds to an inter-dye distance of ~43 Å. Compared to denaturing conditions, where protein expansion resulted in a dye separation of ~63 Å ($E_{FRET} \approx 0.3$), these data provided direct evidence of relative proximity of the NM spanning dyes, consistent with a population of compact monomers.

Further, it was noted that the observed E_{FRET} peak was quite narrowly focused around the peak center. Such a narrow peak indicates either of two possible scenarios: (i) a uniform and highly stable population of conformers (unexpected for an IDP, but consistent with a stably folded structure), or (ii) rapid fluctuation of monomers relative to the detection timescale of 0.5 ms. Additional smFRET experiments utilized guanidinium hydrocholoride to denature the protein molecules and showed a progressive, non-cooperative decrease in E_{FRET} consistent with a population lacking stable structural elements, suggesting that rapid conformational rearrangement was likely responsible. To provide further support for the conclusions drawn from these single-molecule results, specific FCS experiments were designed and utilized, identifying fluorescence decays in the 20–300 ns time scale (Fig 1C), and confirming that the monomer population was indeed an ensemble of compact and rapidly fluctuating structures. In these NM experiments, single-molecule methods proved to be a powerful investigational tool for direct examination of a protein known to be structurally complex and heterogeneous, both in their ability to offer an information-rich snapshot into the population in various conditions, and also in their ability to inform and direct other biophysical methodologies (FCS).

Subsequent single-molecule studies have provided further insight into the polymer physics of IDPs [22–25]. In one study using smFRET, Müller-Späth and Soranno et al. [24] showed that two IDPs with high net charge show scaling behavior indicating expanded conformations as compared to unfolded states of folded proteins