Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in
Structural Biology

# A glass menagerie of low complexity sequences

Randal Halfmann[1,2]

Remarkably simple proteins play outsize roles in the execution of developmental complexity within biological systems. Sequence information determines structure and hence function, so how do low complexity sequences fulfill their functions? Recent discoveries are raising the curtain on a new dimension of the sequence-structure paradigm. In it, function derives not from the structures of individual proteins, but instead, from dynamic material properties of entire ensembles of the proteins acting in unison through phase changes. These phases include liquids, one-dimensional crystals, and — as elaborated herein — even glasses. The peculiar thermodynamics of glass-like protein assemblies, in particular, illuminate new principles of information flow through and, at times, orthogonal to the central dogma of molecular biology.

**Addresses**
[1] Stowers Institute for Medical Research, 1000 E. 50th St., Kansas City, MO 64110, USA
[2] Department of Molecular and Integrative Physiology, University of Kansas Medical Center, 3901 Rainbow Boulevard, Kansas City, KS 66160, USA

Corresponding author: Halfmann, Randal (rhn@stowers.org)

**Current Opinion in Structural Biology** 2016, **38**:18–25

This review comes from a themed issue on **Sequences and topology**

Edited by **L Aravind** and **Elizabeth Meiering**

For a complete overview see the Issue and the Editorial

Available online 31st May 2016

http://dx.doi.org/10.1016/j.sbi.2016.05.002

0959-440X/© 2016 Elsevier Ltd. All rights reserved.

## Introduction

Proteins are the most sophisticated macromolecules. Each performs a finely tuned function that is defined by the protein's unique three-dimensional structure, which is in turn specified by its unique amino acid sequence. The sequence-structure paradigm is too simplistic, however, for the approximately one third of eukaryotic proteins with long regions that are devoid of regular structure [1•]. Such 'intrinsically disordered regions' or IDRs each exist as a cloud of interconverting structures, no one of which alone suffices to carry out protein function. Rather, function manifests over relatively large spatiotemporal scales that include entire ensembles of structures. Disorder even allows certain proteins known as prions to switch between distinct structural states at such low frequency that multiple generations of cells come and go in only one state or the other.

Enzymatic and structural activities of proteins require precise three-dimensional geometries. IDRs rarely perform those tasks directly. Instead, they organize and regulate them. Disorder abounds in eukaryotic signaling pathways and also features prominently in transcriptional and post-transcriptional regulatory machinery. Disordered proteins interact broadly with other proteins, serve as hubs in protein interaction networks, and in the process orchestrate the assembly of massive supramolecular complexes [1•,2].

How, with an aversion to structure, do IDRs assemble some of the largest structures in the cell? The answer is deceptively simple. As if for water vapor condensing into dew droplets, the proteins coalesce out of the bulk cellular milieu into their own liquid phases. Unlike structured macromolecular complexes, individual polypeptides remain disordered within the liquid protein droplets, fleeting between self-solvated, energetically comparable intermolecular conformations. Attached globular domains and interacting macromolecules are pulled in along with their disordered partners. The compartmentalization of proteins within the liquid phase facilitates regulatory processes, including signaling, transcription, mRNA processing, and nucleation of cytoskeletal polymers [2,3]. The droplet-dependent localization of these activities enforces cell polarity, symmetry breaking, and cell differentiation. Intriguingly, droplets have been observed to solidify, or 'mature', over time, and this may be a basis for both functional and pathological differentiation of droplet activities [4•,5••,6••,7••,8].

As for any fundamental cellular process, misregulation of phase behavior can be catastrophic. Accelerated maturation or accumulation of mRNA granules is now heavily implicated in the pathogenesis of amyotrophic lateral sclerosis (ALS) and related proteopathies [2,4•,5••,6••,9•]. Disease-causing mutations in the low complexity proteins that drive granule condensation accelerate a time-dependent reduction in their fluidity. The mutations cause pathology not by inducing aggregation *de novo*, but rather, by subtly altering pre-existing phase behaviors. These dynamics mirror those long observed in classic pathological phase transitions in the form of protein aggregation [10•]. The IDRs behind the pathogenesis of ALS, Alzheimer's, Parkinson's, and prion diseases, among others, progress *in vitro* from disordered monomers through 'molten', or liquid like, oligomeric states en route to semicrystalline fibrillar aggregates known as amyloids [10•].

The progressive reduction in fluidity that characterizes both physiological and pathological phases suggests a simple underlying principle. Extrapolating from an extensive literature in polymer physics and protein folding, I propose that that principle is vitrification — the transition of a supercritical liquid to a non-equilibrium amorphous solid, or more precisely, a glass. Herein I discuss the biophysical bases for vitrification of low complexity sequences (LCSs), and outline its theoretical relationship to protein function and malfunction.

## Sequence complexity and phase behavior

A protein's function is the indirect manifestation of information contained within its amino acid sequence. The more complex, or less redundant that information, the more precisely a single polypeptide of that sequence can fold into a single conformation among the universe of possible conformations. In the parlance of energy landscapes, information determines topology; the steeply funneled landscape of a well-folded globular protein requires a highly complex sequence. Amino acid compositional biases and repeated sequence elements lower the information content such that the landscape flattens and is no longer dominated by one particular funnel. Low complexity sequences tend therefore not to 'fold' in the usual sense; they instead tend to interconvert between energetically comparable conformations [1•,11•]. The types of conformations within each ensemble are determined by the balance of attractive and repulsive interactions in the sequence: attractive interactions favor collapsed conformations while repulsive interactions favor more extended ones. Attractions dominate for most low complexity sequences due to dipole-dipole interactions in the polypeptide backbone. Hence the 'default' LCS ensemble — that is, in the absence of abundant repulsive interactions — appears to be a collapsed molten globule [12]. Amino acid side chains enhance or counteract this behavior according to their individual physicochemical properties [1•,13].

It is this cohesiveness of LCSs that drives phase behavior. Even transient, low affinity interactions between polypeptides raise the local concentrations of other potentially interacting moieties elsewhere in the sequence, leading the polypeptides to remain associated through avidity effects. But every intermolecular interaction displaces an energetically equivalent intramolecular interaction, rendering the unsatisfied moieties more likely themselves to find intermolecular partners. LCSs therefore tend to assemble cooperatively into polydisperse (*i.e.* lacking discrete stoichiometry) clusters, or phases, with critical concentrations for nucleation [1•,13]. The analogous chain reaction for structured proteins, called 'domain swapping', renders tandem repeat multidomain proteins particularly susceptible to form solid phases — that is, to aggregate [11•,14]. The intrinsic polyvalency of LCSs also explains their well-documented promiscuity: disordered proteins have more binding partners than structured proteins; interact aberrantly with other proteins, leading to dosage-sensitive proteotoxicity [15]; and are enriched in the insoluble fractions of proteomes [16].

In general, the higher the interaction valency of a sequence, the tighter its concentration window for assembly. Intracellular LCS-driven phases including RNA granules, the nuclear pore permeability barrier, Cajal bodies, and nucleoli, among others, exploit repetitive low-affinity sequence elements such as RG, SR, and FG dipeptides [1•,17••,18]. The strength of the polyvalent interactions determines the dynamics of conformational interconversion and hence the viscosity of the LCS phase. Weakly-interacting tandem repeats also characterize extracellular LCS assemblies such as elastin, resilin, and spider silks. The repeats not only drive their assembly through coacervation, they also endow the extraordinary adhesiveness and plasticity essential to their functions [1•,19].

Cells exploit the concentration thresholds for LCS assembly to control when and where phase separation occurs. They do so in part through post-translational modifications to LCS valency or interaction strength, including phosphorylation, glycosylation, and sumoylation [1•,2]. Unlike membrane-bound organelles, LCS droplets readily coalesce and dissolve even with small changes in the local subcellular environment. Droplets frequently nucleate through heterotypic interactions with other multivalent molecules. Most of the recently characterized phase forming LCS do so physiologically in complex with other multivalent proteins or nucleic acids, and the inclusion of such ligands *in vitro* greatly lowers concentration thresholds for liquid or gel coalescence [1•,2,13,20••].

## From LCS droplets are born amyloids and glasses

Although LCS phase separation can be likened to water vapor condensing into dew droplets, there is one essential difference. Liquid water is stable once it has formed, but liquid protein phases are not. The lowest energy state of supersaturated protein solutions is believed to be crystalline. Even though the proteins may first coalesce into liquids, such transitions are inherently metastable and ultimately give way to crystals [21]. That LCS droplets are indeed supersaturated is underscored by abundant recent observations that they solidify, or 'mature' over time both *in vitro* and *in vivo* [4•,5••,6••,7••,8,9•,22,23,24•,25].

Naturally, because disordered proteins lack a stable three-dimensional structure, they cannot readily crystallize in three dimensions. Their only structural regularity is the linkage of amino acids in the polypeptide backbone. Consequently, crystallization proceeds orthogonally to the backbone. The resulting one-dimensional paracrystals,

Download English Version:

https://daneshyari.com/en/article/8319686

Download Persian Version:

https://daneshyari.com/article/8319686

Daneshyari.com