Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

How does solvation in the cell affect protein folding and binding?

Caitlin M Davis[1,2], Martin Gruebele[1,2,3] and Shahar Sukenik[1]

The cellular environment is highly diverse and capable of rapid changes in solute composition and concentrations. Decades of protein studies have highlighted their sensitivity to solute environment, yet these studies were rarely performed *in situ*. Recently, new techniques capable of monitoring proteins in their natural context within a live cell have emerged. A recurring theme of these investigations is the importance of the often-neglected cellular solvation environment to protein function. An emerging consensus is that protein processes in the cell are affected by a combination of steric and non-steric interactions with this solution. Here we explain how protein surface area and volume changes control these two interaction types, and give recent examples that highlight how even mild environmental changes can alter cellular processes.

Addresses
[1] Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[2] Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[3] Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Corresponding author: Gruebele, Martin (mgruebel@illinois.edu)

Current Opinion in Structural Biology 2018, **48**:23–29

This review comes from a themed issue on **Folding and binding**

Edited by **Stefano Gianni** and **Amedeo Caflisch**

http://dx.doi.org/10.1016/j.sbi.2017.09.003

Introduction

The cell's interior is a dynamic and heterogenous environment: in it, metabolites, ions, small solutes and macromolecules such as proteins and nucleic acids are unevenly distributed, and are constantly in flux due to both internal and external changes. Internally, a cell may morph and shift shape during motility or as part of its replication cycle, causing volume changes, internal water movement and concentration gradients [1–3]. Membrane bound organelles display distinct solution compositions from other cellular regions [4•], and membrane-less, phase-separated regions within the cytoplasm create regions of high solute concentration [5]. The cell cycle involves the breakdown of the nuclear envelope, releasing large amounts of charged species that were previously contained in the nucleus [6]. Externally, single-cell organisms, plants, and insects are susceptible to environmental changes in osmotic pressure, temperature, and water content [7–9]. Multi-celled organisms capable of homeostasis provide a more stable environment for most of their cells, but certain cell types (e.g. kidney, gut, skin) are nonetheless exposed to environmental changes [10]. In addition, pathological conditions such as fevers, diabetes, and other metabolic diseases or therapies will induce various environmental stresses even in multi-cell organisms [11].

Even *in vitro*, proteins display a remarkable sensitivity to their solvation environment, as evolution for function in many cases seems to favor marginally stable proteins [12]. It is therefore reasonable to expect proteins to be sensitive to the dynamic cellular environment. While studies of protein function in the cell are technically challenging to perform and difficult to interpret, they are necessary to advance our understanding of how proteins interact with their natural environment. Specifically, live-cell NMR and fluorescence microscopy have emerged as two complementary techniques that can detect protein dynamics within the cellular environment. NMR reports on protein dynamics of isotopically labeled overexpressed proteins at the single atom level [13,14••]. Microscopy of fluorescently tagged proteins provides high temporal resolution of protein dynamics, and reveals the context of their function within the cell, though such experiments suffer from limited structural resolution [15,16]. Such studies complement *in vitro* studies with well-defined crowders or cell lysates.

Here we consider recent findings from NMR and fluorescence microscopy, and present a framework with which to understand how the cellular environment affects protein processes. It is important to note that many factors beyond solute composition alter protein dynamics inside the cell. Perhaps most importantly, we will not discuss interactions with chaperones or other post-transcriptional regulating proteins, such as kinases [17], and refer the interested reader to other reviews in this issue [18]. We begin by discussing the composition of the cell's internal solution. Next we present a model for translating protein thermodynamics *in vitro* to the cellular solvation environment. Finally, we highlight experimental data that shows how altering the environment in which a protein is observed affects the studied process.

What does the cellular environment look like?

On average, ~60–70% of a cell's volume is composed of water [19], with the rest being a combination of electrolytes, small organic molecules and metabolites, nucleic acids, and proteins [20]. The relative concentrations of many molecular species have been characterized using biochemical assays [21] to determine multi-cell averages. Absolute numbers in a single cell depend on organism, cell type, and volume [22], and even then different assays give concentrations that can vary by over an order of magnitude [23]. With the exception of cases like acidity in lysosomes, little is known about how these solutes (e.g. Mg^{2+} [24]) are distributed spatially in the cell. Localization of proteins is a well-known phenomenon, but only recently have experiments revealed differences in smaller solute composition between cytoplasm [25], membrane-bound organelles [4•,26] and membrane-less microenvironments within the cytoplasm [5].

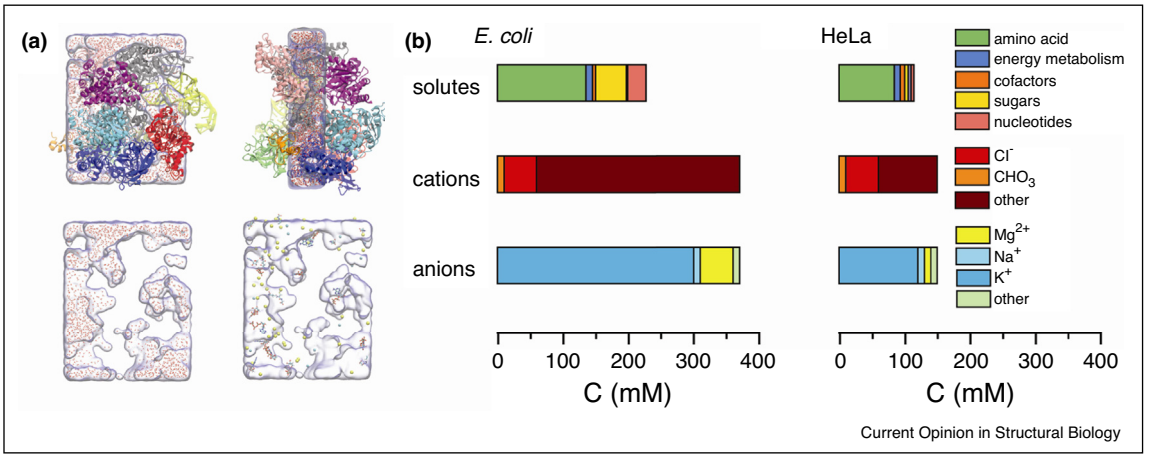
Recently, concentrations derived from these biological assays have been used to generate all-atom molecular dynamics simulations of the cytoplasm [32••,33]. These simulations provide a dynamic view of the interaction between the cell's internal solution and its proteome. Each protein in the cell is separated from other proteins by only a few layers of water (Figure 1a). Small molecules in the cellular milieu, which often exist in mM concentrations (Figure 1b), form transient interactions with protein surfaces. These solutes can alter the structure of interfacial water layers, effectively changing their interaction with protein surfaces [34]. Together with the limited volume of the cell's interior [35], these interactions change protein structure, activity, and interactions in cell compared to the dilute solutions of *in vitro* experiments.

Despite their highly confined heterogenous native environment, many proteins studied in idealized *in vitro* solutions, where water is abundant, solution properties are uniform, and binding partners are all but lacking, successfully recapitulate processes as they occur in the cell. This should not be taken to mean that protein studies can be conducted solely in the test tube; rather, it attests to the robustness of many cases of protein dynamics. Other proteins participate in new and unexpected behaviors depending on their environment, including phenomena such as protein moonlighting [36], intracellular phase separation [37], intrinsic disorder [38], and functional and pathological protein aggregation [39,40].

How does the cellular environment affect protein thermodynamics?

In this review, we highlight the interplay of steric and non-steric interactions and how they scale differently: volume versus surface area. Our definition of 'steric' is in line with molecular crowding, an idea pioneered by Minton and co-workers [41]. Steric interactions stem from the rigid volume a protein takes up in an already crowded environment composed of other rigid bodies

Figure 1



The composition of a cell. **(a)** A simulation box containing characteristic *E. coli* cytoplasmic concentrations [27] of proteins and nucleic acids within a slice of solvent (top). Removing the large macromolecules shows the sparsity of water molecules between adjacent proteins (bottom left), and the high concentration of small solutes and electrolytes in the regions surrounding macromolecular surfaces (bottom right). Image courtesy of Y. Zhang. **(b)** Characteristic concentrations of small molecules, metabolites (obtained using mass-spectrometry [4•,27]), and ions (obtained using biochemical methods [28–31]) inside *E. coli* (left) and HeLa cells (right). In both cases external osmolarity is ~0.3 Osm [30]. For reference, the total concentration of *all* proteins in the cell is in the 1–10 mM range. The reported concentrations include components that are bound to cellular species (i.e. not labile), and ion contributions in the 'other' category may be double-counted since they are also charged solutes. Note the large difference between the *E. coli* and HeLa absolute concentrations, which is in part due to differences in measurement methods, but also highlights the stark differences in the intracellular environments of different species.