Perspective: Reaches of chemical physics in biology

Martin Gruebele and D. Thirumalai

Citation: The Journal of Chemical Physics **139**, 121701 (2013); doi: 10.1063/1.4820139
View online: http://dx.doi.org/10.1063/1.4820139
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/139/12?ver=pdfcov
Published by the AIP Publishing

Perspective: Reaches of chemical physics in biology

Martin Gruebele[1,a)] and D. Thirumalai[2,a)]
[1]*Departments of Chemistry and Physics, and Center for Biophysics and Computational Biology, University of Illinois, Urbana, Illinois 61801, USA*
[2]*Department of Chemistry and Institute for Physical Sciences and Technology, University of Maryland, College Park, Maryland 20742, USA*

(Received 14 July 2013; accepted 20 August 2013; published online 25 September 2013)

Chemical physics as a discipline contributes many experimental tools, algorithms, and fundamental theoretical models that can be applied to biological problems. This is especially true now as the molecular level and the systems level descriptions begin to connect, and multi-scale approaches are being developed to solve cutting edge problems in biology. In some cases, the concepts and tools got their start in non-biological fields, and migrated over, such as the idea of glassy landscapes, fluorescence spectroscopy, or master equation approaches. In other cases, the tools were specifically developed with biological physics applications in mind, such as modeling of single molecule trajectories or super-resolution laser techniques. In this introduction to the special topic section on chemical physics of biological systems, we consider a wide range of contributions, all the way from the molecular level, to molecular assemblies, chemical physics of the cell, and finally systems-level approaches, based on the contributions to this special issue. Chemical physicists can look forward to an exciting future where computational tools, analytical models, and new instrumentation will push the boundaries of biological inquiry. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4820139]

I. INTRODUCTION

Cellular processes encompass a bewildering array of time and length scales.[1] Quantitative methods, rooted in physical principles, have played a major role in describing cellular functions for a long time.[2,3] Indeed, some of the challenges in unearthing the principles of biology using physical methods were laid out nearly 70 years ago by Delbrück.[4] However, more recently the needs, urgency, and the extent of awareness in developing advanced experimental techniques, and novel theoretical and computational methods to tackle a broad range of biological problems are on the rise. In part this is driven by the hope that by understanding the information content of genome sequences within the context of cellular function we could not only qualitatively transform the practice of medicine but also enable us to make novel class of biomimetic materials that match the dexterity of biological molecules. In order to realize this and other goals it is necessary to bring to bear concepts from many different fields, as nature itself seems to have utilized many of them in the course of evolution to replicate and adapt to ever changing environmental demands.

The quest to understand the working of biological systems at the molecular level began in earnest with the discovery of the structure of DNA,[5] although the role physical chemistry was to play in determination and structure of proteins was already evident in the remarkable studies announcing the organization of α-helices and β-sheets.[6] In the decade starting in the mid-1950s stunning discoveries were made, which have set the stage for researching subjects ranging from molecular to system biology. This period inaugurated a great era in biology and resulted in the discovery of the central dogma of molecular biology,[7,8] firmed up the well-documented structure function relationship in the aftermath of determination of the structures of myoglobin and hemoglobin,[9,10] ushered in the notion that sequence of a protein determines the folded structure,[11] introduced the tenets of system biology,[12] and produced a basis for understanding allosteric effects in proteins.[13] Reflecting, in part, on this extraordinary pace of development, prompted Crick to marvel at the notion that conceptually a vast majority of problems in molecular biology were in principle essentially solved,[14] with protein folding problem being an exception. Inspired by the beauty of biological systems, approaches in chemical physics were used to produce truly spectacular successes, especially in the realm of molecular description of biological systems ranging from self-assembly of proteins and RNA to functions of molecular machines and enzyme.[15–22]

Although the view expressed by Crick was overly optimistic, it is fair to say that these developments raised new challenges, which were initially mostly under the purview of biologists. However, in the mid-1980s physicists, chemists, engineers, and more recently material scientists brought to bear an amazing array of methodologies to understand a large range of biological systems quantitatively. The result is the emergence of several interdisciplinary areas of inquiry with ever expanding set of problems, whose solutions now demand rigorous foundations commonly associated with the physical sciences. The collection of articles in this special issue shows that, in the current golden era of quantitative approaches to biological problems, the techniques of chemical physics already play a central role. It is not surprising that chemical physics should be vital to study biology, which differs from

[a)]Authors to whom correspondence should be addressed. Electronic addresses: mgruebel@illinois.edu, FAX: (001) 217 244 3186 and thirum@umd.edu, FAX: (001) 301 314 9121.

physics and chemistry in two unique aspects, namely, evolution and replication (information transfer from one generation to another). Both of these aspects involve chemical reactions carried out by the molecules of life (DNA, proteins, RNA) in a seemingly organized but noisy environment. Therefore, applications of chemical reaction rate theories, effects of molecular fluctuations in chemical reactions, statistical mechanics principles of self-assembly, and information transfer on mesoscopic scales are needed in describing cellular processes. These are problems that the chemical physics community treats on lengths from sub nanometer to micron level.

The modern age in which chemical physics started to play an important role began over 20 years ago with the focus on solving how protein molecules fold. A decade or so afterwards the chemical physics community has become greatly engaged in contributing to this endeavor by bringing new tools and ideas into the fray. It is amusing that theoretical ideas meant to illustrate fundamental aspects of physical systems, without ostensible connections to biology, have become workhorses in a variety of areas in biology. Two examples illustrate this trend. One is the application of two-dimensional infrared spectroscopy, invented by the late Hochstrasser. Using this impressive method, it has been shown that amyloid fibrils from Aβ contain approximately 1.5 water molecules per monomer. Surprisingly, these water molecules are localized in a hydrophobic pocket![23] On the theoretical side, the chemical Langevin equation, studied extensively by Gillespie,[24] has become a standard way to introduce noise in signaling networks in which low copy numbers exaggerate the role of stochastic fluctuations. Both developments took place within the chemical physics community without concern for application to biology, and yet they play a key role in analyzing many unrelated biological systems. These examples and others documented here are reminders of the far reaches of the principles of chemical physics in biology. It is clear that this era has just begun with more to come as this century unfolds.

II. BIOLOGICAL PHYSICS AT THE MOLECULAR LEVEL

Following the synthesis by the ribosome, most of the proteins fold spontaneously, execute their intended functions, and are then degraded, as depicted in the upper right corner of Fig. 1. Even though these processes occur routinely, protein and RNA folding still are astounding examples of molecular self-assembly. Although Anfinsen[11] demonstrated that proteins reversibly reach the folded state spontaneously the quest to understand how proteins fold began in earnest only in the 1980s. By generalizing concepts in the physics of disordered systems and polymer physics, and computations using highly simplified models a conceptual framework of protein folding emerged.[25–28] It was realized that by considering the phases of polymers constructed from random sequences one could understand the special role evolution has played in synthesizing foldable sequences. Naturally, evolved proteins are special because their folding landscape is "funneled," which implies that under folding conditions $\nabla F_{i\to N} \gg \nabla F_{i\to j}$ where $\nabla F_{i\to N}$ is the gradient in the multidimensional folding landscape connecting an arbitrary state i to the native state, N. The key insights, highlighted in a number recent reviews,[29–31] showed that the self-assembly of proteins can only be understood in terms of statistical description, which ironically was already echoed in a prescient monograph by Schrodinger.

A. Biased and coarse-grained models

A particularly important application of the energy landscape perspective is that many aspects of folding could be

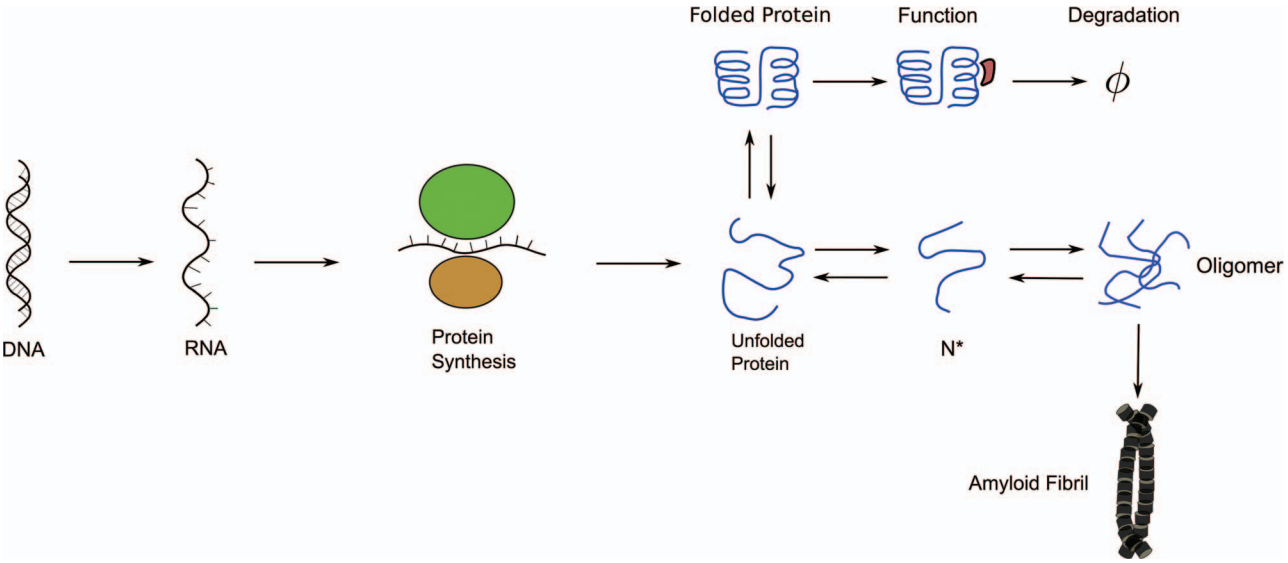

FIG. 1. The complex pathways followed by a polypeptide chain. As expressed succinctly in the central dogma expression of the gene produces RNA, which is translated by the ribosome to yield a polypeptide chain. In the normal function the unfolded protein folds spontaneously, executes the designed function, and is ultimately degraded as indicated in the upper right corner. If the folding channel does not operate as planned the unfolded protein can form N*, an aggregation-prone species, which can then form toxic oligomers eventually resulting in the form of insoluble amyloid fibrils, as displayed in the lower right hand corner.

described by energy functions that emphasize interactions that stabilize the native states.[29] Such biased energy functions, used in the lattice models and more realistic off-lattice models, have been remarkably successful not only in the context of protein folding[32] but also in providing insights into the functions of motors[33] and transcription initiation.[34] However, only recently rigorous methods for devising CG models have been developed. Lu *et al*.[35] have used force-matching method to construct CG potentials that reproduce distribution functions obtained in detailed simulations using a formalism based on Yuon-Born-Green equations familiar in the context of classical many body systems.

The anisotropic network model, introduced to describe gels and polymer rubber networks, is an example of a CG model that has been remarkably successful in describing global dynamics of proteins. In the present issue, Gur *et al*.[36] used the anisotropic model to establish that it can capture the dynamics obtained in all atom simulations for the well-studied protein Bovine Pancreatic Trypsin Inhibitor (BPTI) and an archaeal aspartate transporter, GltPh. The CG simulations also suggest that the native topology has evolved so that functional substates, which are low frequency excitations, are easily accessible. This finding supports the notion that the topology of the native state rather than sequences which are conserved to a greater extent.

B. Folding and function

Although the folding landscape of foldable sequences is smooth without deep kinetic traps, it has been realized that the requirements for folding and function could be different, thus creating a conflict or frustration in certain regions of the proteins. It has been shown that the folding kinetics of the designed Top7 could be considerably more complex than evolved proteins of comparable complexity. Truong *et al*.[37] use tools specifically developed to understand folding of natural proteins to highlight the differences between designed and evolved proteins. These authors also consider the robustness of the designed proteins to the number of amino acids used in the code (nature used 20 residue types), which is an important issue in the synthesis of artificial proteins.

C. Conformational heterogeneity

There are a number of proteins that remain disordered (Intrinsically Disordered Proteins or IDPs) but acquire structure only upon association with a partner.[38] IDPs, as well as unfolded conformations of globular proteins are molecularly heterogeneous. Although this is appreciated qualitatively, addressing the extent of heterogeneity quantitatively remains a challenge just as it is in glassy systems. Lyle *et al*.[39] propose a new method that sheds light on this issue. They develop a new order parameter which distinguishes polymers that fold into a structurally homogeneous ensemble from those that "merely" undergo a coil-to-globule collapse. Devising such novel ways of quantifying the nature of heterogeneous behavior in these systems will be necessary in studying coupled-folding and binding in IDPs.

D. Watching unfolding one protein at a time

Single molecule pulling experiments in which mechanical force is applied to two points on the protein have given an unprecedented view of the energy landscape of proteins. In general, application of force enhances the unfolding rates (slip bond behavior) as anticipated by the Bell model. However, one of the intriguing observations is that in some cases force can prolong the life of a folded or bound state.[40] The structural basis of this behavior, referred to as catch bond, is not fully understood. In an interesting paper Kreuzer and Elber[41] show using all atom molecular dynamics simulations that a simple helix exhibits catch bond behavior force at low forces before making a transition to the slip bond behavior. The findings are linked to complex network of connected states in this simple system. In a number of AFM experiments force-unfolding experiments are performed on poly-protein constructs, and the distribution of unfolding times is usually analyzed by assuming that each protein is an independent module. Kononova *et al*.[42] simulated a dimer composed of non-interacting WW domains. At low forces the two domains behaved independently. However, at high forces mechanical forces break the symmetry of the two identical domains, thus inducing a topological coupling. This finding could be important in biology because a large number of multi-domain proteins are found in cells.

Single molecule spectroscopy using Förster resonance energy transfer (FRET) detection is also making headway into understanding the structure of denatured proteins.[40] A case in point is the contribution by Schuler *et al*. in this issue, who measure the size of unfolded cold shock protein in denaturants at physiological pH and at low pH. Despite the much higher net charge of the unfolded protein at low pH (about 14 *e* of extra charge per molecule), the polypeptide chain remains as compact as at neutral pH. The authors postulate that newly neutralized acidic side chains allows for additional hydrophobic interactions. Organization by additional hydrogen bonds and counterion screening could also contribute to the observed effect. It is very clear from these and other recent ensemble measurements[43,44] that unfolded state structure is critical for determining the ease of the folding process.

III. FROM MOLECULES TO INTERACTIONS

Cellular functions are invariably carried out by interaction between the molecules of life. In addition, cells contain a complex cytoskeletal network involving a number of different proteins. As predicted by the central dogma of molecular biology after gene transcription resulting in the production of mRNA, proteins are synthesized by translating the coding sequence. The translational machinery involves the ribosome that reads the message. Typically, proteins execute the intended functions and are subsequently degraded by the proteosomes. In some cases proteins do not fold correctly, and associate among themselves leading to a number of deposition diseases such Alzheimer's and prion disorders.[45] The imbalance between protein synthesis and degradation results in the failure of proteostasis, which is suspected to be the cause of large number of diseases. Even though the biology associated

with maintenance of proteostasis has not been fully clarified it is clear that biophysical methods that give rise to aberrant inter protein interactions will play a major role in elucidating the molecular basis of unwarranted inter interactions between proteins. Given the tremendous biological importance of these processes it is not surprising that many approaches are being used to understand the molecular details of protein-protein association.

A. Phases of protein solutions

The importance of understanding the forces that drive protein association takes on added importance in light of the discussion given above. As a first step in this endeavor rigorous experiments and theories are needed, as illustrated in two articles in this issue, which characterize the phases of well-defined protein systems under *in vitro* conditions. Ketchum *et al*.[46] study the crystallization of hematin, a process involved in the detoxification of heme released in malaria-infected erythrocytes. In this case it has not been clear whether crystallization occurs in an aqueous medium or if lipid support is needed to obtain ordered structures. Through insightful experimental studies they show that lipid bodies present in the vacuoles must play a role in facilitating crystallization of hematin. The study rules out the possibility that the crystals could grow in aqueous medium.

In a related work, Wang *et al*.[47] provide detailed phase transition studies of different IgG proteins, whose aggregation is related to deposition diseases. The study gives several insights into the generic behavior of this system. The critical concentrations of all the IgG proteins are roughly similar whereas their critical temperatures vary greatly. Both these studies serve as benchmark for developing microscopic theories to understand phase transitions in proteins, and could serve as the basis for understanding protein aggregation in general.

B. Importance of monomer fluctuations in determining aggregation

In addition to experiments, theory and simulations have played an important role in contributing to the factors that are responsible for protein aggregation. In this vein, it is useful to develop analytical models that could provide global insights into the phases of proteins. One of the key insights from theory and simulations is that fluctuations in the spectrum of monomers provide fundamental insights into the proteins to aggregate.[48] A formulation by Weber and Pande,[49] based on heteropolymer models that take secondary structural element into account supports this point of view, which further underscores the need to characterize not only the most stable states of proteins but also excitations (or aggregation prone states) around the global minimum.

C. Depolymerization of microtubules

Microtubules (MTs) are integral parts of the cytoskeleton framework and are found in the cytoplasm. They are found in eukaryotic cells and are formed from polymerization of tubulin dimers. MTs are involved in a number of cellular functions by serving as polar tracks for transport of cargo and are also involved in mitosis. These hollow long stiff polymers are highly dynamic and undergo dynamic instability at the plus end by a depolymerizing mechanism in which the tubulin dimers are actively removed. The mechanism of MT depolymerization is under intense scrutiny and is difficult to probe using standard computer simulation techniques. The power of CG modes is further illustrated in an article by Theisen *et al*.,[50] in which they probe the effect of mechanical force on MT in an attempt to understand how the tubulin dimers are dislodged.[51] The key finding is that the MT severing proteins have to interact with two points on the MT, which facilitates rupture of tubulin. The effective confinement due to the presence of multiple protofilament essentially reduces the entropic barrier, thus speeding up depolymerization of MT.

D. Nucleic acid-protein interactions

Interactions between DNA and proteins drive many biological processes such as transcription initiation (more generally gene expression) and DNA repair. In the process of interactions with proteins at specific DNA sites the structure of DNA is perturbed.[52] Despite being of great importance basic questions such as the flexibility of short DNA molecules, the dependence on the DNA sequence, and time scales of protein-DNA recognition have not been quantitatively answered. It has been difficult to isolate the events involved in recognition of DNA proteins and the resulting distortion of DNA structure. For example, in the formation of transcription bubble due to interaction of DNA and RNA polymerase it is still not settled if DNA binding causes transcription bubble formation or the other way around. In this issue, Vivas *et al*.[53] resolve the DNA bending from the binding process using the protein Integration Host Factor (IHF), which is involved in chromosomal compaction and DNA recombination. IHF binds to specific DNA sites and induces sharp turns. Using laser temperature jump experiments these authors have measure the relaxation kinetics of association as well as their dependence on salt concentration. They suggest that their methodology is a general way of probing the kinetics of protein-DNA interactions and could provide the needed role of salt effects and role hydration plays in the association and disassociation process.

The recognition that RNA is not merely a passive transmitter of information in the genetic code but plays an active role in a number of cellular processes has prompted an intense effort to understand how they fold.[54–56] Due to the presence of easily accessible excitations around the native state the energy landscape of RNA is more rugged than proteins. As a consequence, in cellular conditions it is suspected that RNA chaperones assist in their folding. By generalizing the iterative annealing mechanism introduced to describe the role of bacterial chaperones in protein-assisted folding, Hyeon and Thirumalai have developed a theory to quantitatively describe both passive and active roles of RNA chaperones.[57] The resulting theory gives a quantitative account of experiments.

IV. BIOMIMETICS

The molecular interactions discussed in Secs. II and III can be used to tailor systems that mimic nature, but include synthetic ingredients. The area of molecular biomimetics provides model systems for chemical physical studies, and has applications in drug delivery, tissue scaffolding, and many practical biomedical applications.[58]

A. Modified proteins and peptides

Modified proteins and peptides of various types are an important class of molecular biomimetics. We can think of these as human-engineered post-translational modifications, broadening the already large palette of possibilities offered by natural proteins. For example, *trans-cis* isomerizing groups such as *N*-methylthioacetamide can be added to peptides to make them optically switchable and induce folding or unfolding reactions.[59] Artificial proteins that replace amide linkages between amino acids by esters allow for the fundamental study of backbone hydrogen bonding.[60,61] PEGylation has a long history in drug delivery, as well as in fundamental research trying to understand how glycosylation and similar processes affect the dynamics an solvation of proteins.[62] As the example of PEG illustrates, biomimetic polymers are of course not limited to peptide-like systems. A good example from the realm of nucleic acids is peptide nucleic acid, wherein the phosphate-deoxyribose backbone is replaced by an amino acid.[63] Like many such hybrid systems, it maintains the ability of molecular recognition, while having very different metabolic and stability properties from the original nucleic acid.

B. Biomimetic membranes

Biomimetic membranes are another large class of systems where evolution can be replaced by engineering.[64] Such membranes can have useful properties for separations, electrochemical processes, and as catalytic supports.[65] Much of the chemical physics of transport phenomena in these systems remains to be elucidated, creating an active field. In this issue, artificial membranes are represented by the work of Toca-Herrera and co-workers.[66] They were able to grow large crystalline domains of the protein SbpA, a bacterial coat protein, on poly-lactide films. The domain size could be tuned via the polymer support's glass transition temperature. Such experiments probe the maximum size of defect-free biomimetic films that can be created. A highly active area in chemical physics is currently coupled to electron-proton transport.[67] Liquid membranes (e.g., quinone impregnated nitrocellulose) allow proton-diffusion limited coupled redox reactions, and should be scalable to nanometer thickness over current proof-of-principle designs.[65]

C. Biomimetic assembly

Biomimetic assembly is not limited to two-dimensional structures, and computational research in this area is also very active. An example in this issue is the modeling of clathrin basket formation by Muthukumar and Nossal.[68] Clathrins are proteins that in the natural cellular environment form coated vesicles.[69] Indeed, one of the earliest functions assigned to the ubiquitous hsp70 class of chaperones was to aid in the assembly/disassembly of clathrin coats.[70] The kinetic model study in this issue, based on ideas from micellization theory, shows that a critical threshold concentration of clathrin is required before three-dimensional cage structures build up, a good example of the highly nonlinear behavior that often accompanies the assembly of complex structures. A currently very active example of such nonlinearity is the nucleation of amyloid fibrils, and the paper by Muschol and co-workers[71] makes the distinction between growth processes that proceed through oligomers and those that are oligomer-free. Regarding the question of oligomer *vs*. fibril toxicity, the pendulum has currently swung in favor of oligomers as the toxic species, so physico-kinetic studies of such nucleation events are important, as in the example of the very general combinatorial analysis of stochastic nucleation by Chou and co-workers in this issue.[72]

D. Challenges of biomimetics in chemical physics

The major challenges of molecular biomimetics that can be addressed by chemical physics are twofold: in the modeling of kinetically complex assembly phenomena, especially cases under kinetic control where simple thermodynamic considerations (e.g., knowing the ground state of the system) are insufficient. In that regard, much could be learned from a related area of chemical physics that of glassy dynamics. The second challenge is in the development of characterization methods that can be used to unravel the structure and energetics of complex biomimetic assemblies. This area includes the extension of single molecule spectroscopies to the few-molecule and molecular assembly level, while maintaining the detailed information accessible to single molecule studies.

V. BIOLOGICAL PHYSICS OF THE CELL

A different direction from biomimetics, in which chemical physics scales up beyond molecule-molecule interactions, is the structure, thermodynamics, and dynamics within the cell. The living cell is the biological link between molecular level detail, and the systems-level at which much biology is described. Chemical physics can contribute at both levels, and indeed, many of the techniques developed by chemical physicists are equally applicable at the molecular and systems level.

A. Master equation models

A good example of such dual-use is the Pauli master equation formalism. Master equations were originally implemented to describe quantum transport processes probabilistically.[73,74] A system can be in many states, and the time-varying state occupation probabilities are described by coupled differential equations that resemble the first order differential equations of chemical kinetics, $\partial \mathbf{P}/\partial \mathrm{t} = \mathrm{A}\mathbf{P}(t)$. Both

Markovian (time independent rate coefficients as in chemical kinetics) and non-Markovian processes can be described. Master equations can be equally well applied to molecular systems such as chemical kinetics, and to highly coarse-grained systems, where the propagator may depend on an extended past history of the system, or $\partial \mathbf{P}/\partial t = \int dt' A(t, t')\mathbf{P}(t')$. At the cell-level, this issue contains several examples of such kinetic models. Teo and Schulten develop a comprehensive particle-based Markov model of how molecules to be transported to and from the cell approach and pass through channels.[75] They apply the model to a practical case, diffusion of ions to and their capture by mechanosensitive channels on the bacterial surface. A similar theme is picked up by Berezhovskii and Szabo,[76] who elucidate how ligand diffusion at cell surfaces affects their interaction with the multiple signaling receptors available at the surface. As a nice illustration of how "portable" such chemical physics models can be, their formalism is also applicable to a macromolecule that interacts with small molecules via multiple identical binding sites.

B. Coarse graining

One of the most important concepts in chemical physics of the cell is coarse-graining. The need is both practical and very fundamental. On the practical side, it is not possible to computationally simulate at the atomistic level—at least not yet—entire cells or even organelles or very large macromolecular assemblies. The ribosome and viruses are currently at the upper limit. On the fundamental side, self-organized systems are often hierarchical, and coarse-graining shows up as a natural variable in analytical theory descriptions of such networks. Both sides of the equation show up in this issue. Bowman *et al*.[77] discuss how to use Bayesian modeling to decide which coarse-grained representation provides the most faithful description of the underlying fine-grained dynamics. Among other conclusions, they find that the original Perron analysis works better than some more recently developed algorithms. Kravats *et al*. use coarse-grained dynamics to study how an alpha helical model protein interacts with ring-shaped ATPase "nanomachines" that unfold and assist in translocation of proteins.[78] In particular, they distinguish handed (clockwise and counter-clockwise) from random interactions of loops of the complex that protrude into the opening space where the substrate protein interacts. The authors suggest a lower size limit on the ring assemblies that can remain active.

C. Hydrodynamics

When dynamics are coarse grained, chemical and mechanical phenomena begin to mix, and descriptions that mix concepts from mechanics and molecular dynamics come into play. This is illustrated in this issue by the work of Skolnick and co-workers,[79] who investigate the role of hydrodynamics in whole-cell simulations. At the most elementary level of dynamics, there are only microcanonical (energy conserving) collisions. Once a system containing water and many macromolecules, like the cell, is coarse-grained, concepts such as local viscosity and hydrodynamic flow emerge. Hydrodynamics must be added to the description when macromolecules interact with water molecules, which in turn can interact with other macromolecules. For example, macromolecules that are not at all in contact with one another can transfer linear and angular momentum via intervening water molecules that are not treated explicitly. Instead, a quasi-mechanical description can be invoked. The specific problem tackled by Ando *et al*. is that Stokesian dynamics, while an accurate representation of hydrodynamic interaction in cells over long distances, is computationally expensive, scaling as the cube of the number of particles. They propose a modification where long distance interactions are treated diagonally in the propagation matrix,[79] reducing the effort to linear in the number of particles, and they compare the two methods to map out the limitations of the faster algorithm. Correlated motions of macromolecules over long distances in the cell will still require the full Stokesian dynamics.

D. Whole cell simulations and experiments

Whole cell simulations are now coming online, combining a variety of coarse graining tools. For example, Elcock and co-workers, and Luthey-Schulten and co-workers have simulated the cytoplasm of bacteria, including components from small proteins to ribosomes, as well as adding the cell membrane and nucleic acid in some cases.[80,81] In the near future, such simulations will scale up from procaryotic to eucaryotic cells. Awaiting whole cell simulations is a new generation of experiments that look at protein folding in different organelles of the cell,[82] super-resolution structure of the cell such as the recently discovered "skeleton" of neurons[83] or time-resolved super-resolution dynamics in cell membranes,[84] *in vivo* characterization of motors and other cellular machinery,[85] and cryoelectron microscopy of cellular machinery. Many areas, from neurobiology to developmental biology will benefit from the development of these tools.

E. Atomistic simulations

All that is not to say that molecular dynamics simulations are not forging ahead to provide minimally coarse-grained descriptions of macromolecular machinery in cells. With the size of ribosomes and viruses reached,[86,87] atomistic simulations are moving to the length scale of cellular machinery (see Fig. 2 showing an HIV viral capsid simulation). In this issue, Whitford and Sanbonmatsu use targeted molecular dynamics to reveal different pathways by which tRNAse interacts with the ribosome during translocation.[88] Such results can provide insight for future experimental control of ribosome dynamics, allowing protein manufacture to be experimentally controlled via dynamics, as opposed to just mRNA sequence.

F. Single molecule experiments and modeling

This section would be amiss not to mention the application of single molecule techniques to the cell. This approach can now connect the molecular level to the cellular level of

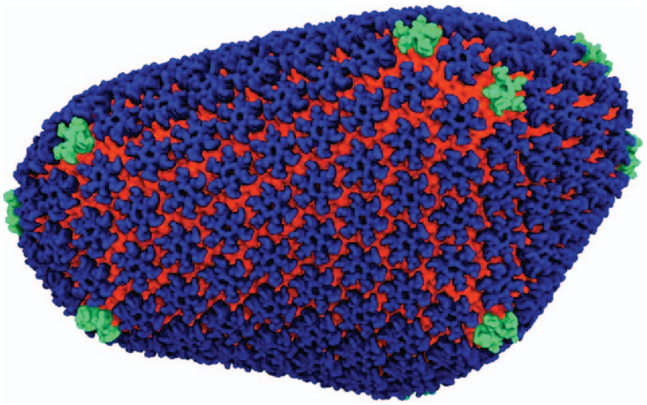

FIG. 2. The viral capsid of the AIDS-causing virus (HIV-1) unveiled at the atomic level by a combination of experimental techniques and computation. The capsid protects the viral RNA, reverse-transcriptase, and other auxiliary proteins essential to the infective cycle of the virus. The capsid has to be stable enough to protect its content, yet also brittle upon a chemical trigger, to release its content after the capsid enters a cell. The architecture of the virus follows a theorem of Euler according to which a fully enclosed encasing can be realized through hexagonal and pentagonal elements where one needs exactly 12 pentagons (some visible in green), but can adopt any number of hexagons (here 216, shown in blue). The number of hexagons determines the size, the distribution of pentagons the shape. The capsid is made of only protein CA that can accommodate a distribution of surface curvatures. There are about 1300 CAs in the capsid. The image is based on an atomic model derived through crystallography and nuclear magnetic resonance (NMR) structure analysis of isolated CA dimers, pentamers and hexamers, through electron microscopy of hexameric surfaces, and through data-guided molecular dynamics simulations (pdb code 3J3Q). Capsid structure and dynamics pose a new challenge to chemical physics. The model and a first molecular dynamics simulation (involving 64×10^6 atoms) of a solvated capsid have been reported in Ref. 86.

dynamics, as molecular machinery is investigated by fluorescence labeling, and *in vivo* single molecule techniques are becoming more widely applied. Initial experiments in low copy number gene expression[89] are now complemented by a wide range of techniques using photo-localization, total internal reflection, and other optical and kinetics tools to achieve single molecule sensitivity in the background of a living cell.[90] Chemical physics will contribute here in many ways, from adaptation of sensitive laser spectroscopy tools, to modeling of single molecule dynamics needed to interpret increasingly complex experiments.[91,92]

VI. PHYSICAL PRINCIPLES OF BIOLOGICAL ORGANIZATION

As mentioned in Sec. V, many of the algorithms developed in chemical physics can be applied at different levels of the molecule to systems biology hierarchy. This issue contains a number of papers that study model systems of relevance to molecular transport (e.g., Langevin-based modeling), self-replication, global dynamics and stability of biological networks, systems biology networks, and other areas where chemical physics can make conceptual contributions.

A. Stability of complex systems

Living systems rely on coupling and feedback to control processes, resulting in complex networks that often operate near the edge of stability or large fluctuations. For example, single molecule FRET studies reveal that some protein-protein interactions can be fine-tuned from positive to negative cooperativity to control downstream signaling very sensitively.[52] A good example in this issue is the article by Ghosh and co-workers,[93] which considers two competing reactions, $A + B \rightarrow AB$ and $A + C \rightarrow AC$. B and C compete for the same binding partner A. It is clear from basic statistical mechanics that such systems are subject to strong fluctuations when there are few copies of A, B, and C around. What they show is that large fluctuations remain when the copy numbers of B and C go to infinity, and only A is small. This is observed only when the two reactions compete—a single reaction shows smooth dynamics if just one of its two components has large copy number. This type of behavior has been long known from the chemical physics of phase transitions, where the merger of two distinct free energy wells (first order transition) into a single well (critical transition) can produce fluctuations on a macroscopic scale. Such amplification of fluctuations also plays an important role in biological systems, from gene regulation to embryogenesis. Wu and Wang probe this question of instability with a statistical field theory. In their view, spatio-temporal diffusion processes can be characterized by a Lyapunov functional that allows the stable points of the diffusing system to be identified globally.[94] The formalism could be applied to any types of populations subject to stochastic dynamics, from interacting diffusing molecules in a cell, to populations in an ecosystem, another case where molecular-level physics concepts can be applied on completely different time- and length scales.[95]

B. Noisy dynamics in living systems

At the more molecular level, Metzler and co-workers present an extension of noisy random walk models that are used to explain anomalous diffusion of macromolecules.[96] They find that "wiggling in place" of particles, not described by the basic formalism, is taken into account when thermal fluctuations of the environment are added into the equation (Fig. 3). At the same time, the model maintains simplicity, a guiding ingredient in the best physico-chemical models. Yang and co-workers investigate an information metric for a related problem, that of Langevin transport of a particle subject to a noisy force.[97] Their approach quantifies how rich the parameter space needed to describe a given diffusion process really is, making use of eigenbasis decomposition that has been fruitfully applied to diffusive dynamics in other biological physics contexts also, such as low-barrier protein folding dynamics,[98] or the number of reaction coordinates required to describe dynamics of a polypeptide chain.[99] As one last example of self-organization at the molecular level, Chou and co-workers analyze stochastic self-assembly and nucleation.[72] They consider a "zero sum game" model where nucleation sites compete for a single resource of particles that can organize into clusters. The paper nicely illustrates how numerical analysis can go hand-in hand with analytical results, the latter describing important limiting cases of the dynamics, while the former bridge these limiting cases.

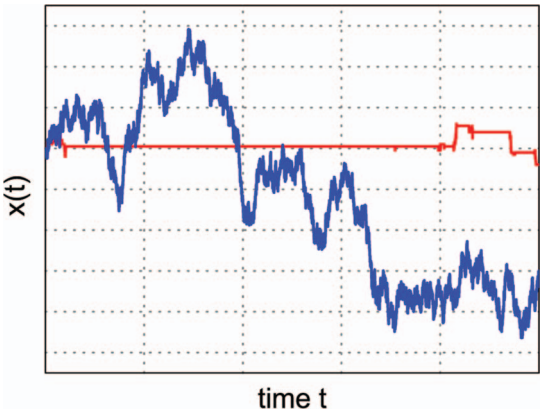

FIG. 3. A random walk from Ref. 96 in this issue. Continuous time random walks (CTRWs) are stochastic models for anomalous diffusion (AD) processes. AD is characterized by pronounced trapping periods, during which the particle cannot move. This may happen, for instance, when small tracer particles are successively caged in a matrix of semiflexible biopolymers, such as the ultrastructure in living cells. Such environments are intrinsically noisy: the matrix itself evolves in time and fluctuates thermally, rattling the particle during trapping periods. Noisy CTRWs can account for such rattling effects. The red line shows the case when the thermal noise only weakly disturbs the trapping events (horizontal plateaus between sudden changes in position). In the blue line strong noise completely changes the trajectory with the same initial condition, highlighting the importance of an evolving environment for the single particle dynamics.

C. Self-replication

One of the most important features that distinguishes life from the non-living is the ability to self-replicate.[100] Self-replication, as complex a phenomenon as it is, nonetheless is subject to basic chemical and physical laws, such as the second law of thermodynamics. In his contribution, England considers the entropy and heat production that must be minimally associated with replication.[101] As one might expect intuitively, more robust replicators require a greater expenditure of entropy for every replication than ones that are not durable. In this area of chemical physics applied to biology, statistical mechanics and information theory can work together closely to provide constraints on what is minimally required for a successful replicator. Ultimately, such analysis can inform the field of synthetic biology, which has many open questions, such as: what is the minimal number of chemical components that must constitute and autonomous self-replicator, given a specified environment of available components? What is a minimal number of genes (information) required to support replication and survival?[102] As researchers attempt to simplify life forms from the top down,[103] and manufacture artificial life from the bottom up,[104] such questions will not remain purely academic.

VII. OUTLOOK

Chemical physics research has woven a rich tapestry in biology, from fundamental principles such as energy landscapes or macromolecular structure, to computational models of dynamics, and experimental tools to examine biomolecules and cells in unprecedented detail. The role of chemical physics in biological inquiry is only going to increase: the molecular aspects of biology at the cellular and sub-cellular level have become even more important since the 1950s; at the same time, the systems level has benefited from many of the analysis tools originally developed and refined for applications in the physical sciences, such as random walk models and master equation approaches. Indeed, in many ways two branches of biology, at the molecular level and at the systems level, are coming closer together. Nowadays, it is not unusual to study entire bacterial ecosystems in relation to very specific molecules produced by the organisms that can act as messengers, toxins, attractors, or metabolic intake. The molecular machinery, the complex control of genetic networks (most such molecules are produced only in specific circumstances, never in "monoculture" in the Petri dish), the different folding properties of nucleic acids, proteins, and peptides (not all these molecules are small organics) in the cytoplasm and outside the cell, dynamics of transport, diffusion and binding, systems-level interaction of the bacteria, and many more such factors come into play. In addition, research in biology continues to produce major surprises, as revealed, for example, by the totally unanticipated roles of non-coding RNA molecules.[105] At all these levels, chemical physics approaches can make substantial contributions, and biology fosters new applications that were only speculation just a few years ago.

ACKNOWLEDGMENTS

M.G. was supported by funding from the National Science Foundation (Grant No. MCB 1019958) and the Center for Physics in Living Cells, and NSF Physics Frontier Center. D.T. acknowledges support from the National Science Foundation, Grant No. CHE 0910433. D.T. is grateful to Alexandra Thirumalai for useful comments. Figure 1 was prepared by Shaon Chakraborty, Fig. 2 was prepared by Juan Perilla and Klaus Schulten, and Fig. 3 by Ralf Metzler.

[1] B. Alberts, A. Johnson, J. Lewis, and M. Raff, *Molecular Biology of the Cell* (Taylor & Francis, New York, 2007).
[2] A. M. Turing, Philos. Trans. R. Soc. London, Ser. B **237**, 37 (1952).
[3] L. Pauling, H. A. Itano, S. J. Singer, and I. C. Wells, Science **110**, 543 (1949).
[4] M. Delbrück, Trans. Conn. Acad. Arts Sci. **38**, 173 (1949) [reprinted in J. Cairns, G. S. Stent, and J. D. Watson, *Phage and the Origins of Molecular Biology*, expanded edition (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1992)].
[5] J. D. Watson and F. H. C. Crick, Nature (London) **171**, 737 (1953).
[6] L. Pauling, R. B. Corey, and H. R. Branson, Proc. Natl. Acad. Sci. U.S.A. **37**, 205 (1951).
[7] F. H. C. Crick, Nature (London) **227**, 561 (1970).
[8] F. H. C. Crick, Symp. Soc. Exp. Biol. **XII**, 139 (1958).
[9] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, and D. C. Phillips, Nature (London) **181**, 662 (1958).
[10] M. F. Perutz, W. Bolton, R. Diamond, H. Muirhead, and H. Watson, Nature (London) **203**, 687 (1964).
[11] C. B. Afinsen, Science **181**, 223 (1973).
[12] F. A. Jacob and J. Monod, J. Mol. Biol. **3**, 318 (1961).
[13] J. Monod, J. Wyman, and J. P. Changeus, J. Mol. Biol. **12**, 88 (1965).
[14] F. H. C. Crick, *What Mad Pursuit* (Basic Books, New York, 1990).
[15] W. A. Eaton and R. M. Hochstrasser, J. Chem. Phys. **49**, 985 (1968).
[16] M. Eigen, G. Maass, and D. Porschke, Angew. Chem., Int. Ed. **6**, 459 (1967).
[17] W. E. Moerner, J. Phys. Chem. B **106**, 910 (2002).
[18] A. Szabo and M. Karplus, J. Mol. Biol. **72**, 163 (1972).
[19] M. Levitt and A. Warshel, Nature (London) **253**, 694 (1975).
[20] S. J. Benkovic and S. Hammes-Schiffer, Science **312**, 208 (2006).

[21]M. Garcia-Vilaco, J. Gao, M. Karplus, and D. G. Truhlar, Science **303**, 186 (2004).
[22]H. Frauenfelder, S. Sligar, and P. G. Wolynes, Science **254**, 1598 (1991).
[23]Y. S. Kim, L. Liu, P. H. Axelsen, and R. M. Hochstrasser, Proc. Natl. Acad. Sci. U.S.A. **106**, 17751 (2009).
[24]D. T. Gillespie, J. Chem. Phys. **81**, 2340 (1977).
[25]J. D. Bryngelson and P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **84**, 7524 (1987).
[26]P. G. Wolynes, J. N. Onuchic, and D. Thirumalai, Science **267**, 1619 (1995).
[27]E. Shaknovich, Chem. Rev. **106**, 1559 (2006).
[28]K. A. Dill and H.-S. Chan, Nat. Struct. Biol. **4**, 10 (1997).
[29]J. N. Onuchic and P. G. Wolynes, Curr. Opin. Struct. Biol. **14**, 70 (2004).
[30]K. A. Dill, S. B. Ozkan, M. S. Shell, and T. R. Weikl, Annu. Rev. Biophys. **37**, 289 (2008).
[31]D. Thirumalai, E. P. O'Brien, G. Morrison, and C. Hyeon, Annu. Rev. Biophys. **39**, 159 (2010).
[32]D. Thirumalai, Z. Liu, E. P. O'Brien, and G. Reddy, Curr. Opin. Struct. Biol. **23**, 22 (2013).
[33]C. Hyeon and J. N. Onuchic, Proc. Natl. Acad. Sci. U.S.A. **104**, 2175 (2007).
[34]J. Chen, S. Darst, and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **107**, 12523 (2010).
[35]L. Lu, J. F. Dama, and G. A. Voth, J. Chem. Phys. **139**, 121906 (2013).
[36]M. Gur, E. Zomot, and I. Bahar, J. Chem. Phys. **139**, 121912 (2013).
[37]H. H. Truong, B. L. Kim, N. P. Schafer, and P. G. Wolynes, J. Chem. Phys. **139**, 121908 (2013).
[38]P. Wright and J. Dyson, Nat. Rev. Mol. Cell Biol. **6**, 197 (2005).
[39]N. Lyle, R. K. Das, and R. V. Pappu, J. Chem. Phys. **139**, 121907 (2013).
[40]K. A. Merchant, R. B. Best, J. M. Louis, I. V. Gopich, and W. A. Eaton, Proc. Natl. Acad. Sci. U.S.A. **104**, 1528 (2007).
[41]S. M. Kreuzer and R. Elber, J. Chem. Phys. **139**, 121902 (2013).
[42]O. Kononova, L. Jones, and V. Barsegov, J. Chem. Phys. **139**, 121913 (2013).
[43]V. A. Voelz, M. Jager, S. Yao, Y. Chen, L. Zhu, S. A. Waldauer, G. R. Bowman, M. Friedrichs, O. Bakajin, L. J. Lapidus, S. Weiss, and V. S. Pande, J. Am. Chem. Soc. **134**, 12565 (2012).
[44]J. Yao, H. J. Dyson, and P. E. Wright, FEBS Lett. **419**, 285 (1997).
[45]F. Chiti and C. M. Dobson, Annu. Rev. Biochem. **75**, 333 (2006).
[46]M. A. Ketchum, K. N. Olafson, E. V. Petrova, J. D. Rimer, and P. G. Vekilov, J. Chem. Phys. **139**, 121911 (2013).
[47]Y. Wang, A. Lomakin, R. F. Latypov, J. P. Laubach, T. Hideshima, P. G. Richardson, N. C. Munshi, K. C. Anderson, and G. B. Benedek, J. Chem. Phys. **139**, 121904 (2013).
[48]D. Thirumalai, R. Dima, and D. K. Lkimov, Curr. Opin. Struct. Biol. **13**, 146 (2003).
[49]J. K. Weber and V. S. Pande, J. Chem. Phys. **139**, 121917 (2013).
[50]K. E. Theisen, N. J. Desai, A. M. Volski, and R. I. Dima, J. Chem. Phys. **139**, 121926 (2013).
[51]T. Mitchison and M. Kirschner, Nature (London) **312**, 237 (1984).
[52]A. Robinson and A. M. van Oijen, Nat. Rev. Microbiol. **11**, 303 (2013).
[53]P. Vivas, Y. Velmurugu, S. V. Kuznetsov, P. A. Rice, and A. Ansari, J. Chem. Phys. **139**, 121927 (2013).
[54]D. Thirumalai, N. Lee, S. A. Woodson, and D. K. Klimov, Annu. Rev. Phys. Chem. **52**, 751 (2001).
[55]I. Tinoco and C. Bustamante, J. Mol. Biol. **293**, 271 (1999).
[56]S. J. Chen and K. A. Dill, Proc. Natl. Acad. Sci. U.S.A. **97**, 646 (2000).
[57]C. Hyeon and D. Thirumalai, J. Chem. Phys. **139**, 121924 (2013).
[58]M. Sarikaya, C. Tamerler, A. K. Y. Jen, K. Schulten, and F. Baneyx, Nature Mater. **2**, 577 (2003).
[59]J. Helbing, H. Bregy, J. Bredenbeck, R. Pfister, P. Hamm, R. Huber, J. Wachtveitl, L. De Vico, and M. Olivucci, J. Am. Chem. Soc. **126**, 8823 (2004).
[60]S. Deechongkit, H. Nguyen, E. T. Powers, P. E. Dawson, M. Gruebele, and J. W. Kelly, Nature (London) **430**, 101 (2004).
[61]M. Wang, T. E. Wales, and M. C. Fitzgerald, Proc. Natl. Acad. Sci. U.S.A. **103**, 2600 (2006).
[62]B. K. Pandey, M. S. Smith, C. Torgerson, P. B. Lawrence, S. S. Matthews, E. Watkins, M. L. Groves, M. B. Prigozhin, and J. L. Price, Bioconjugate Chem. **24**, 796 (2013).
[63]M. Egholm, P. E. Nielsen, O. Buchardt, and R. H. Berg, J. Am. Chem. Soc. **114**, 9677 (1992).
[64]G. Swiegers, *Bioinspiration and Biomimicry in Chemistry* (Wiley-VCH, Hoboken, NJ, 2012).
[65]N. M. Kocherginsky, M. G. Goldfeld, and I. S. Osak, J. Membr. Sci. **45**, 85 (1989).
[66]A. Lejardi, A. E. López, J. R. Sarasua, U. B. Sleytr, and J. L. Toca-Herrera, J. Chem. Phys. **139**, 121903 (2013).
[67]J. Y. Fang and S. Hammes-Schiffer, J. Chem. Phys. **106**, 8442 (1997).
[68]M. Muthukumar and R. Nossal, J. Chem. Phys. **139**, 121928 (2013).
[69]B. M. F. Pearse, Proc. Natl. Acad. Sci. U.S.A. **73**, 1255 (1976).
[70]C. Lagaudriere-Gesbert, S. L. Newmyer, T. F. Gregers, O. Bakke, and H. L. Ploegh, Proc. Natl. Acad. Sci. U.S.A. **99**, 1515 (2002).
[71]J. Foley, S. E. Hill, T. Miti, M. Mulai, M. Ciesla, R. Robeel, C. Persichilli, R. Raynes, S. Westerdeide, and M. Muschol, J. Chem. Phys. **139**, 121901 (2013).
[72]M. R. D'Orsogna, B. Zhao, B. Berenji, and T. Chou, J. Chem. Phys. **139**, 121918 (2013).
[73]A. Nordsieck, W. E. Lamb, and G. E. Uhlenbeck, Physica **7**, 344 (1940).
[74]W. Pauli and M. Fierz, Z. Phys. **106**, 572 (1937).
[75]I. Teo and K. Schulten, J. Chem. Phys. **139**, 121929 (2013).
[76]A. M. Berezhkovskii and A. Szabo, J. Chem. Phys. **139**, 121910 (2013).
[77]G. R. Bowman, L. Meng, and X. Huang, J. Chem. Phys. **139**, 121905 (2013).
[78]A. N. Kravats, S. Toddast-Navaei, R. J. Bucher, and G. Stan, J. Chem. Phys. **139**, 121921 (2013).
[79]T. Ando, E. Chow, and J. Skolnick, J. Chem. Phys. **139**, 121922 (2013).
[80]S. R. McGuffee and A. H. Elcock, PLOS Comput. Biol. **6**, e1000694 (2010).
[81]E. Roberts, J. E. Stone, and Z. Luthey-Schulten, J. Comput. Chem. **34**, 245 (2013).
[82]A. Dhar, K. Girdhar, D. Singh, H. Gelman, S. Ebbinghaus, and M. Gruebele, Biophys. J. **101**, 421 (2011).
[83]K. Xu, G. S. Zhong, and X. W. Zhuang, Science **339**, 452 (2013).
[84]Y. Ishitsuka, Y. M. Li, R. Fischer, N. Takeshita, and G. U. Nienhaus, Biophys. J. **104**, 652A (2013).
[85]B. H. Blehm, T. A. Schroer, K. M. Trybus, Y. R. Chemla, and P. R. Selvin, Proc. Natl. Acad. Sci. U.S.A. **110**, 3381 (2013).
[86]G. P. Zhao, J. R. Perilla, E. L. Yufenyuy, X. Meng, B. Chen, J. Y. Ning, J. Ahn, A. M. Gronenborn, K. Schulten, C. Aiken, and P. J. Zhang, Nature (London) **497**, 643 (2013).
[87]K. Y. Chan, L. G. Trabuco, E. Schreiner, and K. Schulten, Biopolymers **97**, 678 (2012).
[88]P. C. Whitford and K. Sanbonmatsu, J. Chem. Phys. **139**, 121919 (2013).
[89]J. Yu, J. Xiao, X. J. Ren, K. Q. Lao, and X. S. Xie, Science **311**, 1600 (2006).
[90]M. Coelho, N. Maghelli, and I. M. Tolic-Norrelykke, Integr. Biol. **5**, 748 (2013).
[91]P. Lenz, S. S. Cho, and P. G. Wolynes, Chem. Phys. Lett. **471**, 310 (2009).
[92]J. S. Cao, Chem. Phys. Lett. **327**, 38 (2000).
[93]T. Firman and K. Ghosh, J. Chem. Phys. **139**, 121915 (2013).
[94]W. Wu and J. Wang, J. Chem. Phys. **139**, 121920 (2013).
[95]T. Butler and N. Goldenfeld, Phys. Rev. E **80**, 030902(R) (2009).
[96]J.-H. Jeon, E. Barkai, and R. Metzler, J. Chem. Phys. **139**, 121916 (2013).
[97]K. R. Haas, H. Yang, and J.-W. Chu, J. Chem. Phys. **139**, 121931 (2013).
[98]G. Scott and M. Gruebele, J. Comput. Chem. **31**, 2428 (2010).
[99]O. M. Becker and M. Karplus, J. Chem. Phys. **106**, 1495 (1997).
[100]E. Schrödinger, *What is Life? The Physical Aspect of the Living Cell* (Cambridge University Press, Cambridge, 1944).
[101]J. England, J. Chem. Phys. **139**, 121923 (2013).
[102]E. V. Koonin, Annu. Rev. Genomics Hum. Genet. **1**, 99 (2000).
[103]C. M. Fraser, J. D. Gocayne, O. White, M. D. Adams, R. A. Clayton, R. D. Fleischmann, C. J. Bult, A. R. Kerlavage, G. Sutton, J. M. Kelley, R. D. Fritchman, J. F. Weidman, K. V. Small, M. Sandusky, J. Fuhrmann, D. Nguyen, T. R. Utterback, D. M. Saudek, C. A. Phillips, J. M. Merrick, J. F. Tomb, B. A. Dougherty, K. F. Bott, P. C. Hu, T. S. Lucier, S. N. Peterson, H. O. Smith, C. A. I. Hutchison, and J. C. Venter, Science **270**, 397 (1995).
[104]S. Rasmussen, L. H. Chen, M. Nilsson, and S. Abe, Artif. Life **9**, 269 (2003).
[105]C. C. Mello, Cell Death Differ. **14**, 2013 (2007).