The physics of life: one molecule at a time

Mark C. Leake

The esteemed physicist Erwin Schrödinger, whose name is associated with the most notorious equation of quantum mechanics, also wrote a brief essay entitled ‘What is Life?’, asking: ‘How can the events in space and time which take place within the spatial boundary of a living organism be accounted for by physics and chemistry?’ The 60þ years following this seminal work have seen enormous developments in our understanding of biology on the molecular scale, with physics playing a key role in solving many central problems through the development and application of new physical science techniques, biophysical analysis and rigorous intellectual insight. The early days of single-molecule biophysics research was centred around molecular motors and biopolymers, largely divorced from a real physiological context. The new generation of single-molecule bioscience investigations has much greater scope, involving robust methods for understanding molecular-level details of the most fundamental biological processes in far more realistic, and technically challenging, physiological contexts, emerging into a new field of ‘single-molecule cellular biophysics’.

Here, I outline how this new field has evolved, discuss the key active areas of current research and speculate on where this may all lead in the near future.

1. Introduction

Feynman, celebrated physicist and bongo-drums enthusiast, gave a lecture in 1959 viewed by nanotechnologists of the future as a prophecy imagining perfectly their own field. The title was ‘There’s plenty of room at the bottom’, and it discussed a potential future to control and manipulate machines and store information on a length scale tens of thousands times smaller than that of the everyday ‘macroscopic’ world [1]. It was a clarion call to engineers and scientists to establish a new discipline, later coined nanotechnology [2]. Feynmann alluded to this small scale as relevant to that of biological systems and how cells could function at this scale to perform ‘all kinds of marvellous things’. We now know that this fundamental minimal unit is the single biological molecule. It is not to say that atoms comprising these molecules do not matter, nor subatomic particles that make up the individual atoms, nor smaller still the quarks of which the subatomic particles are composed. The point is, in general, we do not need to refer to length scales smaller than single molecules to understand most biological processes.

Technological developments in experimental biological physics have been the primary driving force in establishing the field of single-molecule biophysics, and even though the discipline in its modern form is only a human generation in age, it is clear that at the often prickly interfaces between the physical and the life sciences and at the scale of the single biological molecule, many of the most fundamental questions concerning cellular systems are being addressed. This field is evolving into a new discipline of single-molecule cellular biophysics [3]. It is manifested not only in investigations at the single-molecule level using live cells as the test system, i.e. in vivo single-molecule studies, but also in some highly ingenious single-molecule studies in vitro that, although divorced from the native physiological context, have a very high level of complexity either in the make up of the experimental components studied or in the combinatorial
single-molecule biophysics methods used, which greatly enhance the physiological relevance of the data obtained.

2. The establishment of single-molecule biophysics

(a) Why bother with single molecules?
An experimental method that uses single-molecule biophysics gives us information on the position of a biomolecule in space at a given time or will allow the control and/or measurement of forces exerted by/on that molecule [4], or sometimes both. However, these approaches, despite being established for over two decades in dedicated scientific research laboratories around the world, are still technically challenging because they operate in a regime dominated by stochastic thermal fluctuations of water solvent molecules whose characteristic energy scale, that of $k_B T$ where $k_B$ is Boltzmann’s constant and $T$ the absolute temperature measured in Kelvin, is comparable to energy transitions involved in molecular processes in biology. Forces are characterized by the piconewton (pN) scale, and the length scale of molecules and complexes is of the order of a few nanometers (nm), two orders of magnitude smaller than the wavelength of visible light (figure 1).

Why should we wish to perform such experiments that, as a rule, require measurements of tiny signals in environments of significant noise, in all but rare cases suffering from poor yields and, traditionally, being not remotely high-throughput? There already exist many robust bulk ensemble average methods that illuminate several aspects of structure and function of cellular systems involved in molecular processes in biology. Forces are characterized by the piconewton (pN) scale, and the length scale of molecules and complexes is of the order of a few nanometers (nm), two orders of magnitude smaller than the wavelength of visible light (figure 1).

Furthermore, there is a danger of lack of synchronicity in ensemble experiments. The issue here is that different molecules within a large population may be doing different things at different times; molecules may for example be in different conformations, but this single average parameter would look something close to the most stable of these many short-lived conformations still exist, which are used in different stages of motion and force generation. The mean conformation would look something close to the most stable of these many different conformations, but this single average parameter does not tell us a great deal about the behaviour of the other short-lived, but essential states. Bulk ensemble average analysis, irrespective of what experimental property is measured, cannot probe multiple states in a heterogeneous molecular system.

Also, temporal fluctuations in the molecules from a population result in broadening the distribution of a measured parameter from a bulk ensemble experiment, which can be difficult to interpret physiologically. These thermal fluctuations are driven by collisions from the surrounding water molecules (approx. $10^{9}$ per second—biological molecules are often described as existing in a thermal bath), which can drive biological molecules into different states. An ensemble experiment, this may broaden the measured value, making reliable inference difficult. In single-molecule measurements, these states can often be probed individually.

The principal reason for using novel physical methods and analyses for studying biological processes at the level of single molecules is the prevalence of molecular heterogeneity. One might suppose that the mean average property of approximately $10^{19}$ molecules (roughly the number of molecules in 1 µl of water, equivalent to $1/(18 \times 1000)$th of a mole), is the case for most bulk ensemble average techniques, is an adequate representation of the properties of any given single molecule. In some exceptional biological systems this is true; however, in general this is not the case. This is because single biological molecules usually exist in multiple states, intrinsically related to their biological functions. A state here is a measure of the energy locked into that molecule. For example, there are many molecules which exist in multiple spatial conformations, such as molecular motors, with each conformation having a characteristic energy state.

Although there may be a single conformation that is more stable than the others for these tiny molecular machines, several short-lived conformations still exist, which are used in different stages of motion and force generation. The mean conformation would look something close to the most stable of these many different conformations, but this single average parameter does not tell us a great deal about the behaviour of the other short-lived, but essential states. Bulk ensemble average analysis, irrespective of what experimental property is measured, cannot probe multiple states in a heterogeneous molecular system.

Figure 1. A schematic of the length scale of biological molecules and complexes in the context of larger macroscopic length-scale entities.
The real strength of single-molecule biophysics experiments is that these subpopulations of molecular states can be investigated. The importance to biology is that this multiple-state heterogeneity is actually an essential characteristic of the normal functioning of molecular machines; there is a fundamental instability in these molecules that allows them to switch between multiple states as part of their underlying physiological function.

A final point to note is that, although there is a wide range in concentration of biological molecules inside living cells, the actual number of molecules that are directly involved in any given biological process at any one time is generally low. Biological processes at this level can, therefore, be said to occur under minimal stoichiometry conditions in which just a few stochastic molecular events become important. In fact, it can often be these rarer, single-molecule events that may be the most significant to cellular processes, and so it becomes all the more important to investigate life at the level of single molecules, and many approaches developed from the physical sciences have now been established focused upon using single-molecule biophysics techniques to address fundamental biological questions [7].

(b) The first generation of single-molecule biophysics investigations

Single-molecule biophysics is still a youthful field, in the context of the traditional ‘core’ sciences. The first definitive single biological molecule investigations used pioneering electron microscopy techniques to produce metallic shadow replicas of large, filamentous molecules including DNA and a variety of proteins [8], using fixed samples in a vacuum. Single-particle detection began in non-biological samples, involving trapping single elementary particles in a gaseous phase in the form of a single electron [9] and later as a single atomic ion [10].

The first single-molecule biophysics investigation in which the surrounding medium included that one compound essential to all known forms of life, namely water, came with fluorescence imaging in the laboratory of Rotman [11], with the detection of single molecules of the enzyme β-galactosidase by chemically modifying one of its substrates to make it fluorescent and observing the emergence of these molecules during the enzyme-catalysed reaction inside microscopic droplets—although the sensitivity of detection at that time was not sufficiently high to monitor single fluorescent molecules directly, this particular assay used the fact that a single molecule of the β-galactosidase enzyme could generate several thousand product molecules that could be detected, and thereby indicate the presence of a single enzyme. Comparable observations were made in the laboratory of Thomas Hirshfeld over a decade later in aqueous solution without the need for microdroplets using the organic dye fluorescein, similar in structure to the fluorogenic component in the 1961 Rotman study, attached via antibodies to single globulin protein molecules, each with 80–100 individual fluorescein molecules bound [12]. The decade that followed involved marked developments in measurement sensitivity, including fluorescence detection of single molecules of a liquid-phase solution of the protein phycocyanin labelled with approximately 25 molecules of the orange organic dye rhodamine [13], as well as parallel developments in the detection of single molecules in solids using optical absorption of a non-biological sample [14].

The seminal single-molecule biophysics work that came in the subsequent decade involved in vitro studies, experiments done, in effect, in the test tube. In the first instance, these investigations were driven by developments in a newly established technique of optical trapping, also known as laser or optical tweezer. The ability to trap particles using laser radiation pressure was reported by Ashkin [15], forefather of optical trapping, as early as 1970, though the modern form that results in a net optical force on refractive/dielectric particles of higher refractive index than the surrounding medium roughly towards the intensity maximum of a focused laser (figure 2a–c) was developed in the early 1980s by Ashkin et al. [16], and these optical force-transduction devices have since been applied with great diversity to study single-molecule biophysics [17,18].

Arguably, the key pioneering biophysical investigation involving optical trapping used only a relatively weak optical trap in combination with a very sensitive subnanometre precise detection technique called back focal plane interferometry [19], with micrometre-sized beads conjugated to molecules of the motor protein kinesin to monitor the displacement of single kinesin motors on a microtubule filament track, which indicated quantized stepping of each motor of a few nanometres consistent with the structural periodicity of kinesin binding sites on the microtubule [20]. This was followed by a study on another molecular motor of a type of myosin protein that was implicated in the generation of force during muscle contraction in its interaction with F-actin filaments [21]. This investigation used two independent optical traps to tether a single filament and lower it onto a third, surface-immobilized, bead which had been functionalized with the ‘motor-active’ part of the myosin molecule. This was the first study to clearly measure both the quantized nature of displacement and force of a single molecular motor to nm/pN precision.

Biopolymer molecules were also the source of seminal single-molecule biophysics investigations, using optical trapping to measure the mechanical properties by stretching molecules and observing how the forces that developed changed with end-to-end displacement. These were applied to both single- and double-stranded DNA [22] and RNA [23] nucleic acids (the latter study also investigating folding/unfolding transitions in the model RNA hairpin structural motif), as well as large modular proteins made up of repeating motifs of either the immunoglobulin or fibronectin family, including many proteins related to the class of giant muscle proteins known as titins [24–27].

A complementary technique of atomic force microscopy (AFM) also emerged at around the same time. Surface probe techniques originated through the seminal work of Gerd Binning using the scanning tunnelling microscope [28] that measured electron tunnelling between a sample surface and micrometre-sized probe tip (a quantum mechanical effect whose probability depended exponentially on the tunnelling distance involved) as a measure of the surface topography. This developed into AFM [29], in which a similar probe tip, typically composed of silicon nitride, detects primarily Van der Waals forces from a sample surface, allowing imaging of surface topography to subnanometre precision. AFM instead of imaging the surface uses a probe tip as a fishing-rod to clasp ends of molecules bound to a gold-coated surface and subsequently stretch them in retracting the tip away from the surface. This approach was used on
modular protein constructs of titin to demonstrate forced unfolding of individual immunoglobin modules. In doing so, this seminal paper showed evidence for a single-molecule ‘signature’—a physical measurement indicating that there is really a single molecule under investigation, as opposed to multiples or noise, and in the case of AFM this signature was a characteristic ‘sawtooth’ pattern of the molecular force–extension trace that indicated dramatic changes in molecular extension of approximately 20–30 nm whenever one of the immunoglobin modules made a forced transition from folded to unfolded conformations [30].

Developments in optical imaging, most importantly fluorescence microscopy, had an enormous impact on pushing single-molecule biophysics forward. These have included molecular interaction methods using single-molecule Förster resonance energy transfer (smFRET) in which energy can be transferred non-radiatively between differently coloured donor and acceptor dye molecules, each designed to be attached to biological structures which transiently interact as part of their biological function. FRET occurs provided that there is a suitable spectral overlap between the emission and absorption spectra, and the two molecules are both oriented appropriately and within less than approximately 10 nm of each other. The first clear report of smFRET measurements involved monitoring single-molecule assembly of the DNA double helix [31].

Fluorescence imaging was also applied to monitor rotation of single molecules of the rotary motor F1-ATPase by attachment of a rhodamine-tagged fluorescent filament of F-actin conjugated to the F1-ATPase rotor subunit, which not only demonstrated clear rotation of this vital biological machine responsible for the generation of the universal cellular fuel ATP, but also showed that the motion occurs in quantized angular units mirroring the symmetry of the enzyme’s atomic structure [32].

In another pioneering study, single-molecule fluorescent dye imaging was used to monitor the movement of tagged myosin molecules to show that they travelled along F-actin tracks in a hand-over-hand mechanism. This was the first study to show unconstrained walking of a single molecular motor, using nanometre-precise localization in the form of Gaussian fitting of the ‘point spread function’ image of each single fluorescent dye molecule, which the investigators denoted as fluorescence imaging with one nanometre accuracy or FIONA [33].

A seminal in vitro study that links to several key in vivo investigations involved the application of high-speed millisecond fluorescence imaging to monitor real-time diffusion of single lipid molecules labelled with an organic dye, expressed in an artificial lipid bilayer [34], thus acting as a mimic for real cell membranes. Here, investigators could track single molecules with an accuracy better than the optical resolution limit (approx. 200–300 nm) using a method that estimated the centre of the fuzzy diffraction-limited intensity image of single dye molecules to within a few tens of nanometres precision by using Gaussian fitting to the raw images (a method that was originally applied almost a decade earlier to determine the centre position of 190 nm diameter kinesin-coated beads conjugated to microtubules from non-fluorescence brightfield differential interference contrast images to within 1–2 nm precision [35]).

3. The ‘golden age’: the emergence of single-molecule ‘cellular’ biophysics

(a) Approaches that investigate living, functional cells

With so much exemplary single-molecule biophysics research performed in the test tube, a question that should be addressed is: why do we care about studying molecular
Figure 3. Schematics of (a) TIRF, (b) AFM and (c) SICM.

details in live-cell, or near live-cell, environments? Test tube environments are significantly more controllable, less contaminated and come associated with less measurement noise. The best answer is that cells are not test tubes. A test tube experiment is a much reduced version of the native biology containing only components that we think/hope are important. We now know definitively that even the simplest cells are not just bags of chemicals, but rather have localized processes in both space and time. Also, the effective numbers of molecules involved in many cellular processes are often low, sometimes just a few per cell, and these minimal stoichiometry conditions are not easy to reproduce in the test tube without incurring a significant reduction in physiological efficiency.

Single-molecule biophysics investigations in vivo are, however, technically very difficult. Here, fluorescence microscopy is an invaluable biophysical tool. It results in exceptionally high signal-to-noise ratios for determining the localization of molecules tagged with a fluorescent dye, but does so in a way that is relatively non-invasive compared with other single-molecule biophysics methods. This minimal perturbation to native physiology makes it a probe of choice in single-molecule biophysics studies in the living cell. Many of the improvements in our ability to detect single molecules have been driven by developments in the technology that allows photons to be efficiently collected from molecular report probes, several of which are fluorescent, including both ‘point’ detectors such as the photomultiplier tube and pixel arrays of the next-generation high-quantum efficiency cameras called electron multiply charge-coupled devices, and these comparative technologies are reviewed in Michalet et al. [36].

It was only as recently as the year 2000 that the first definitive single-molecule biophysics investigation involving a living sample was performed by Sako et al. [37] (in which the investigators performed single-molecule live-cell imaging on the cell membrane; here, the high-contrast imaging technique of total internal reflection fluorescence microscopy (TIRF; figure 3e) [38] was used to monitor fluorescently labelled epidermal growth factor ligands binding to membrane receptors), and by Byassee et al. [39] (in which the researchers performed single-molecule live-cell imaging inside the centre of a cell, using confocal microscopy to monitor fluorescently labelled transferrin molecules undergoing endocytosis).

Significant developments have been made over the past decade in the field of live-cell super-resolution imaging [40], the ability to perform optical imaging in vivo at a spatial resolution better than that predicted from the Abbe optical resolution limit of approximately 0.61λ/NA, where λ is the detected wavelength for imaging and NA is the numerical aperture of the imaging system (typically set by the objective lens of the optical microscope of approx. 1.2–1.5), in particular, an ability to monitor functional molecular complexes with such precision [41,42]. There are several reviews that the reader can seek to discover the state of the art in regards to various super-resolution technologies, however in this Theme Issue super-resolution methods are reviewed in the context of a relatively new and highly promising technique called optical lock-in detection (OLID), which permits dramatic improvements to imaging contrast in native cellular imaging, far in excess of other competing super-resolution methods [43].

Recent developments in cellular single-molecule fluorescence imaging have included the ability to definitively count molecules that are involved in functional biological processes integrated in the cell membranes of live cells (e.g. to quantify multiple protein subunit components in relatively large molecular machines, such as the bacterial flagellar motor [44,45] or single ion channels [46]), and to combine counting with tracking of relatively mobile components around different spatial locations in the cell (such as molecular machines involved in protein translocation [47] and ATP fuel generation via oxidative phosphorylation [48,49]). The state of the art of our ability to image molecular components in cell membranes has led to substantial improvements to our understanding of their complex architecture, reviewed in two articles in this Theme Issue for model bacterial systems [50] as well focusing on putative zones of molecular confinement in the membrane, commonly referred to as lipid rafts [51]. By modifying the modes of fluorescence illumination, for example using narrowfield [34] or slitfield imaging [52], it has been possible to increase the excitation intensity in the vicinity of single cells to allow millisecond single-molecule imaging. This has permitted visualization of native components normally expressed in the cytoplasm of cells whose viscosity is 100–1000 times smaller than that of the cell membrane and so would be expected to diffuse at a faster rate by this same factor, allowing observation of gene expression bursts [53], regulation of transcription factors [54] and quantification of functional replicons components used in bacterial DNA replication machines [55].

Despite the central importance of fluorescence methods for single-molecule cellular imaging, there are also non-fluorescence detection techniques that can generate highly precise images. For example, scanning probe microscopy (SPM) techniques. These cover a range of experimental approaches allowing topographical detail from the surface of a sample to be obtained by laterally scanning a probe across the surface. There are more than 20 different types of SPM methods currently developed that measure a variety of physical parameters as the probe is placed in proximity to a sample surface, and the most popular to date has been
AFM (figure 3b). Klenerman et al. [56] review SPM techniques in the context of single-molecule precise imaging of the topographical details of live cells, namely probe-accessible features present on the cell membrane, and discusses in depth a relatively novel SPM approach of scanning ion conductance microscopy (SICM; figure 3c).

Another non-fluorescence technique that shows significant potential for single-molecule cellular biophysics is surface enhanced Raman scattering (SERS). Raman scattering is an inelastic process such that scattered photons from a sample have a marginally different frequency to those of the incident photons due primarily to vibrational energy transfer from the molecular orbitals in the sample, either resulting in a loss of energy from the photons (Stokes scattering) or, less commonly, a gain (anti-Stokes scattering). However, to detect the presence of a single molecule in a sample using Raman spectroscopy requires significant enhancement to the standard method used to acquire a scattering spectrum from a bulk, homogeneous sample. The most effective method uses surface enhancement, which is reviewed in Wang & Irudayaraj [57], involving placing the sample in a metallic particles, and in the vicinity of the surface the local electric field $E$ associated with the photons is enhanced by a factor $E'$.

The enhancement depends critically on the size/shape of the nanoparticles, but typically generates better measurement sensitivity by a factor of approximately $10^{14}$, particularly effective if the molecule itself is conjugated to the nanoparticle surface. This enhancement can be sufficient to detect single biomolecules.

(b) In vitro methods of high complexity

This is not to say that in vitro experiments are intrinsically bad and in vivo experiments are definitively good. Rather, they each provide complementary information.

In vitro experiments are detached from a true physiological setting, but the level of environmental control is high. In vivo experiments are more demanding technically and are subject both to greater experimental noise and intrinsic biological variation—being in a native physiological environment is appealing at one level but offers difficulty in interpretation since there is a potential lack of control over other biological processes not directly under study, but that may influence the experimental results.

Next-generation in vitro single-molecule biophysics approaches are characterized by a much greater complexity than those involved in the early days of the field. Some of these often highly involved novel test tube approaches are discussed in Greene & Duzdevich [58], with a particular emphasis on a high-throughput single-molecule biophysics method to investigate the binding of proteins to DNA, called DNA curtains.

One particular focus of recent in vitro single-molecule experiments has been the FoF1-ATPase enzyme, a highly complex machine composed of two rotary molecular motors of the membrane-integrated Fo motor and the hydrophilic F1 motor, which are ultimately responsible for the generation of cellular ATP. Recent single-molecule biophysics approaches to investigate this vital, ubiquitous enzyme are reviewed in Börsch & Sielaff [59], with novel confirmation that the mechanism of nanoscale stepping of the F1 component elucidated in a thermophilic enzyme at room temperature, in which molecular rotation has been fuelled by the hydrolysis of ATP in the opposite direction to that involved during ATP manufacture, is shared by the mesophilic Escherichia coli F1 enzyme, suggesting that even in markedly different environments there are common modes of action to this ubiquitous, essential molecular machine [60].

(c) Novel automated and biocomputational techniques

Single-molecule biophysics experiments are often plagued with noise, with the effective signal-to-noise ratio being sometimes barely in excess of 1 and generally less than 10. This constitutes an enormous analytical challenge to reliably detect a true signal and not erroneously measure noise. Molecular events are often manifested as some form of transient step signal in a noisy time-series, for example, a motor protein might move via stepping along a molecular track. Thus, the challenge becomes one of reliable step-detection from noisy data. The aim is to assemble quantitative statistics of such step events in a fully objective, automated way.

Edge-preserving filtration of the raw, noisy data is often the first tool used, which preserves distinct edge events in time-series, such as the simple median filter, or better still the Chung–Kennedy filter that consists of two adjacent running windows whose output is the mean from the window possessing the smallest variance [26,27]—a step event may then be classed as ‘true’ on the basis of the change in the mean and variance between the two windows being above some pre-agreed threshold.

A significant issue with step-detection from a data time-series is that detection is sensitive to the level of threshold set. An alternative approach where all steps in a series are expected to be of the same size is to convert the time-series into a frequency-domain using a Fast Fourier transform and then detect the periodicity in the original trace by looking for a fundamental peak in the associated power spectrum, which has been used to good effect for the estimation of molecular stoichiometry using step-wise photobleaching of fluorescent proteins [44].

A recent improvement in objectifying single-molecule biophysics data is in how the distributions of single-molecule properties are rendered. Traditional approaches used histograms; however, these are highly sensitive to histogram bin size and position. A more general, objective approach uses kernel density estimation—data are convolved with a Gaussian whose width is the measurement error for that property in that particular experiment and whose height is normalized so that the area under the Gaussian is precisely one (i.e. one detected event), used to good effect in studying single-molecule architectures of the bacterial replisome [55].

Spatial dynamics of single molecules and complexes inside living cells is a feature of biological processes. However, owing to the low signal-to-noise ratio involved in cellular imaging experiments, the analysis of the motions of molecular complexes is non-trivial. Robson et al. [61] describe a novel method implementing a well-known weapon in the statistician’s armoury called Bayesian inference to robustly determine the underlying different modes of molecular diffusion relevant to live-cell imaging in both an objective and automated manner.

One of the biggest challenges to single-molecule biophysics is the traditionally low-throughput nature of experiments.
Singe-molecule cellular biophysics

Leake MC. 2012


Hall CE. 1956 Method for the observation of macromolecules with the electron microscope illustrated with micrographs of DNA. Biophys. Biochem. Cytol. 2, 625 – 628. (doi:10.1083/jcb.2.5.625)


References


37. Reyes-Lamotte R, Serratt DJ, Leake MC. 2010 Stoichiometry and architecture of active DNA
replication machinery in *Escherichia coli*. 


