Single-Molecule DNA Analysis

J. William Efcavitch and John F. Thompson

Helicos BioSciences Corporation, Cambridge, Massachusetts, 02139;
email: jwefcavitch@helicosbio.com, jthompson@helicosbio.com

Annu. Rev. Anal. Chem. 2010. 3:109–28

First published online as a Review in Advance on February 24, 2010

The *Annual Review of Analytical Chemistry* is online at anchem.annualreviews.org

This article's doi:
10.1146/annurev.anchem.111808.073558

1936-1327/10/0719-0109$20.00

Key Words

dynamics, G-quadruplex, Holliday junctions, nanopores, next-generation sequencing

Abstract

The ability to detect single molecules of DNA or RNA has led to an extremely rich area of exploration of the single most important biomolecule in nature. In cases in which the nucleic acid molecules are tethered to a solid support, confined to a channel, or simply allowed to diffuse into a detection volume, novel techniques have been developed to manipulate the DNA and to examine properties such as structural dynamics and protein-DNA interactions. Beyond the analysis of the properties of nucleic acids themselves, single-molecule detection has enabled dramatic improvements in the throughput of DNA sequencing and holds promise for continuing progress. Both optical and nonoptical detection methods that use surfaces, nanopores, and zero-mode waveguides have been attempted, and one optically based instrument is already commercially available. The breadth of literature related to single-molecule DNA analysis is vast; this review focuses on a survey of efforts in molecular dynamics and nucleic acid sequencing.

1. INTRODUCTION

Before we review single-molecule DNA studies, it is important to determine why single-molecule studies are necessary, given that, in some cases, they are more complex to carry out than are ensemble-based studies. Why go to the trouble of examining single molecules when it is generally so much easier to observe populations? Obviously, a tremendous understanding of many DNA properties can be gained through examination of bulk or averaged properties, but detailed insight into many properties and processes simply cannot be attained by looking at an average DNA molecule behaving in an average fashion. The dynamics of DNA moving from an inactive state or complex to an active state cannot be examined directly when only a tiny fraction of molecules are in that state at any given time. DNA achieves its complex structure and function in a highly regulated and orchestrated manner, with different molecules behaving in very nonaverage ways. Therefore, it is necessary to look at the variations in DNA behavior that are attainable via individual molecules to observe how the dynamic functioning of DNA occurs. The very critical area of DNA dynamics and metastable states can only be inferred from observations of initial and final states. Only single-molecule studies allow a smooth transition from one state to the other.

In one single-molecule review (1), the authors asked, "When can a single-molecule result be trusted?" Perhaps an equally legitimate question would be "When can an ensemble-of-molecules result be trusted?" One answer to both questions is that neither type of result may, on its own and in all cases, provide a full picture of how DNA performs its normal functions. Many DNA sequences are involved in intricate functions that simply cannot be described by average properties, yet the extremes of those behaviors determine how those functions can occur. When both single-molecule and ensemble behaviors can be understood with respect to one another, the results can be integrated to show how individual behavior can generate bulk properties.

There are a variety of techniques available for examination of the properties of single molecules of DNA. However, each of these techniques has its own set of limitations, in terms of throughput, time resolution, and modifications to DNA properties, that may affect results. These limitations are incurred in order to visualize the molecule or to keep it physically or temporally separate from other molecules long enough for reliable detection to be performed. These perturbations must be carefully examined to ensure that they do not significantly affect the property being measured. For example, if a DNA molecule is attached to a surface at multiple points, dynamic motion and many protein interactions may occur very differently because the DNA can no longer freely rotate or diffuse in solution. Such a restriction may be useful when one is trying to distinguish different kinetic properties such as one-dimensional diffusion versus hopping from one segment to another, but one should enter into such limitations knowingly. Attachment to any surface or large object may have an impact on many properties, as may covalent or noncovalent attachment of dye molecules that would allow fluorescence detection. Additionally, some single-molecule techniques not only provide the ability to observe individual molecules but also allow one to manipulate the molecules by stretching, rotating, or pulling the DNA in different directions with varying levels of force.

In a typical single-molecule experiment, there is a means to localize the DNA being examined to a discrete volume so that its properties can be distinguished from all the other molecules in the sample. Additionally, depending on the time resolution of the experiment or the type of data being collected, it may be important to keep track of individual molecules and their history. In other experiments, it is unnecessary to know which molecule is being detected: Only the knowledge that the signal is from a single molecule is required. Localization to a specific space for examination of specific molecules is generally performed via attachment to a fixed surface (including beads, glass, or similar media) or via confinement to a channel with passage through the channel controlled

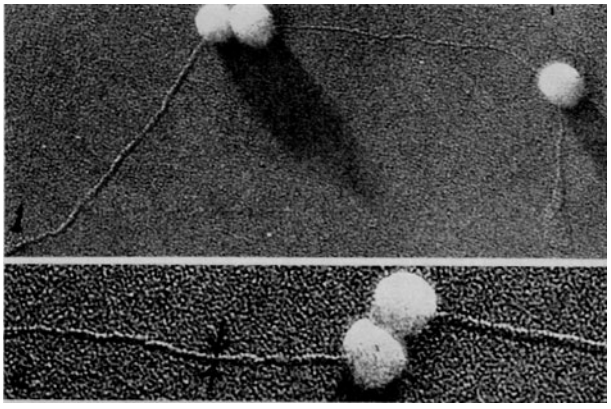

Figure 1

Single molecules of DNA. One of the first single-molecule studies of DNA utilized electron microscopy to visualize individual molecules of double-stranded DNA in the presence of polystyrene beads. When the DNA was pretreated in various ways, the differences between single- and double-stranded DNA could be readily observed. Reprinted from Reference 2.

by flow, electrochemical gradients, lasers, or other means. Attachment to a surface or a bead can be done covalently via specific or nonspecific chemical moieties or via protein–small molecule interactions; antibody-antigen and biotin-streptavidin interactions are the most common. If the localization occurs within a living cell, the detected component must be localized to a motion-restricted compartment or attached to a large structure such as a membrane to avoid diffusion. If it is not necessary to track individual molecules, they can be left freely diffusing in solution, and detection of single molecules can be accomplished by keeping molecules at a low enough physical or temporal density such that only a single molecule is excited and/or detected at a given time. However, this can only be done if the background signal is low and can be readily distinguished from the desired signal.

The power of single-molecule DNA studies has been recognized since the early days of DNA study. The obvious importance of observing single molecules was evidenced by the title of one of the first efforts in the area: "Visualization of Individual Macromolecules with the Electron Microscope" (2). In fact, some of the earliest single-molecule studies showing electron micrographs of DNA next to polystyrene beads (**Figure 1**) are remarkably similar in appearance to DNA tethered to beads in contemporary visualizations. Through the ability to see single molecules and observe (*a*) how they varied across the whole population and (*b*) how they responded to treatments such as denaturation, investigators gained important insights into DNA structure and behavior. Although our knowledge of DNA has advanced substantially in the intervening half century, the challenges remain the same: How can we select individual molecules and analyze their behavior so that we may better understand the structure and dynamics of the entire population, and how can we predict functional outcomes of any perturbations we may introduce?

The field of single-molecule DNA studies has grown so large that it is no longer possible to comprehensively review all of it. Therefore, we discuss only a small number of representative areas of study and use the lessons learned from them as exemplars for other, similar fields of study. There are many other reviews that cover different aspects of single-molecule DNA studies that may be referred to for deeper coverage of specific areas. These reviews include those that address the whole range of technologies, with little discussion of the resultant biology (1, 3), and more restricted and in-depth reviews that cover technologies allowing simultaneous manipulation

and viewing of DNA (4) [or only a subset of those technologies (5, 6)]; technologies compatible with cellular viewing (7); and technologies applied only to gene expression (8), chromatin (9), or recombination (10). There are also reviews that address particular detection methodologies (11) without regard to manipulation technologies. Because no single review can cover all these areas in depth, we discuss certain functions in which DNA plays a role and how those functions are productively addressed by single-molecule techniques. These areas range from simple conformational changes to complex, multicomponent systems; they include G-quadruplexes, DNA bending and looping, Holliday junctions (HJ), and DNA-replication machinery. Also, we discuss the application of single-molecule sequencing and the unique benefits derived therefrom. Because of space constraints, we cite only a small fraction of the available literature, but we note that the research cited here is, obviously, built on a rich foundation of older work.

HJ: Holliday junctions

FRET: Förster resonance energy transfer

SPR: surface plasmon resonance

AFM: atomic force microscopy

2. SINGLE-MOLECULE DNA DYNAMICS

2.1. Folding of G-Quadruplex DNA Structures

G-quadruplexes are a class of structures of high biological interest due to their involvement in chromosome telomeres, chromosome stability, gene expression, and cancer. They have been extensively studied via bulk techniques, which have generated a wealth of solution structures that are providing valuable insight. However, the nature of these structures highlights their dynamic nature. How stable are they, and what factors may affect their folding? Can their structure and dynamics be exploited to develop therapeutics for cancer or senescence? These questions cannot be answered by simply examining static structures; rather, single-molecule techniques are needed. Förster resonance energy transfer (FRET) studies were among the earliest efforts undertaken to understand G-quadruplex dynamics in freely diffusing molecules (12). Simple distance measurements and how they changed as a function of cation and competing oligonucleotides allowed the confirmation of earlier structural findings, but this study design allowed the tracking of individual molecules only for very short times.

FRET was also used both to determine the number of folding intermediates and to measure their lifetimes as a function of salt and temperature (13). Later FRET studies further resolved folding intermediates using a variety of substituted oligonucleotides (14, 15), but movement between species remained uncertain due to the inability to follow individual molecules over time. More recent experiments have used surface-attached oligonucleotides so that the lifetime of individual folded states could be followed as a function of solution conditions and added ligand (16). Surface plasmon resonance (SPR) was also used to generate distances, although some researchers were concerned that some transition conditions did not match solution data, suggesting an issue with surface immobilization (17). The extensibility of longer, surface-immobilized telomere mimics were studied through use of flow techniques (18). Although technically this was a single-molecule study, the results were garnered from G-quadruplex structures that were present in multiple copies on each molecule, making their individual behavior uncertain. Other techniques that have more recently been applied to the G-quadruplex folding/unfolding equilibrium are nanopores and atomic force microscopy (AFM) (19, 20). Because the folded structures cannot pass through the narrow opening, it is possible to determine their stability. AFM has been used to compare the force required to rupture different G-quadruplex sequences at varying concentrations of salt; each molecule could be followed for long periods of time. Thus, no single technique has provided a complete picture of G-quadruplex folding; however, the combination of approaches has added to the overall picture of the dynamics of the structure, perhaps making these structures more amenable to therapeutic intervention.

2.2. Holliday Junctions

Another class of DNA folding, which is at the tertiary level rather than the secondary level, is that of four-stranded HJ. Such structures naturally occur as intermediates in a variety of DNA-recombination pathways, but they are frequently studied as synthetic structures with fixed junctions or only slightly mobile junctions to maintain a stable, consistent structure for study. Although classically drawn as extended two-dimensional structures, they are typically folded more compactly in the presence of physiological concentrations of salt. Ensemble studies of such structures have provided many details about the static properties of HJ but very few data on the key dynamic behavior of the these structures. The ease with which these molecules can be constructed synthetically, with a variety of options for placing fluorescent tags and modifications for surface attachment, has made this area ideal for single-molecule studies.

Single-molecule FRET of oligonucleotides attached to a surface has been used to examine many aspects of HJ structure and dynamics. Early single-molecule FRET studies examined immobile junctions and characterized the nature of the folding and dynamics of the different isoforms (21, 22). Because the DNA was surface attached, individual molecules could be studied for periods of time that were long relative to the DNA dynamics; hence, structural lifetimes could be followed. Even more detailed information about overall structure could be generated through use of three-color FRET, in which multiple, simultaneous distance measurements can be made (23). Good agreement was found between the ensemble-observed equilibrium ratios and the lifetimes as observed with single molecules. These studies provided valuable information about the nature of this specific reaction, although there were caveats of surface attachment and the very short length of the DNA required by the system. Following these studies, the dynamics of branch migration in molecules capable of only a single–base pair migration and multiple potential folding pathways were tested (24). The authors therefore followed the dependence of the dynamics of branch migration as a function of sequence, and they showed that these dynamics were much faster than predicted on the basis of ensemble measurements. A similar approach, but with a longer mobile junction (25), found that the transition between structures was too fast to measure at low magnesium concentrations, although multiple long-lived structures were detectable at higher concentrations. Intermediates that moved the junction by only a single base pair were not observed because transitions over the complete range of homology occurred too rapidly to distinguish. Through use of numerous immobile junctions and junctions with single–base pair nonhomologies, additional kinetic insight was achieved (26). Much longer regions of homology and overlap have been examined via AFM (27), which allows kilobase-scale movements to be monitored over time. These experiments have the advantage of being in vivo–like in terms of DNA lengths, but there is a potential drawback of surface or interface interactions needed to immobilize the DNA for viewing and occur over its extended length. A combination of approaches can yield even finer detail, as evidenced by the use of optical trapping in combination with FRET (28). Varying amounts of force applied along different helical orientations has allowed both the stability of different folding structures and the kinetics of the transition between them to be determined (**Figure 2**).

2.3. DNA Bending and Looping

The next layer of complexity beyond simple conformational changes is reached when protein binding and protein-binding energies are added to the mix. DNA bending can be induced by a range of factors ranging from simple sequence context to bulges to protein binding. These structural features are critical in that they can facilitate a number of more complex interactions and functions. DNA has been traditionally viewed as a long, linear molecule, and the concept of

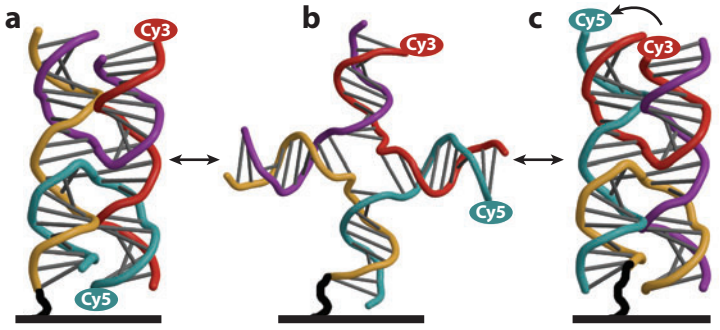

Figure 2

Conformational changes in Holliday junctions. The subtle changes in alignment of helices in a Holliday junction can be detected by judicious placement of donor and acceptor dye molecules and observing how Förster resonance energy transfer changes as a function of the manner in which the molecules are manipulated (21). Single molecules starting in (*a*) one conformational state can pass rapidly through (*b*) an intermediate state and then into (*c*) a second stable state. Figure reprinted from Reference 112 with permission.

sharp bends in the double-helical structure is at odds with this simplistic view. However, the level of compaction within a cell or nucleus reinforces the necessity of invoking some type of folding. To determine the extent of DNA bending or compaction, either the change in distance between two fixed points on a DNA helix or the change in force required to achieve a given distance is generally measured. However, there are multiple ways of performing such measurements; each has advantages and disadvantages. For example, DNA compaction by nonspecific DNA-binding proteins from bacteria has been examined with magnetic tweezers. Through examination of the force required to achieve a normal DNA length, multiple binding modes were detected (29). However, because the binding was nonspecific and of uncertain stoichiometry, little detail was extracted about the behavior of individual proteins and their local impact on DNA structure. Experiments on a similar system have also been carried out with an optical tweezer system (30). This system was set up to detect interactions between adjacent DNA molecules and the potential transfer of protein molecules from one to the other.

Instead of using nonspecifically binding proteins, investigators utilized a structurally similar but specific binding protein, integration host factor (IHF), to observe the effects at a single binding site (31). The change in end-to-end distance of a single molecule was measured via an evanescent wave–scattering method to detect a bead bound to one end of the DNA. The Brownian motion of the bead can also be measured both with and without the protein bound to determine the change in DNA length between the two states. Because single molecules of both DNA and IHF are being observed (albeit indirectly via the bead), the on and off rates of single molecules can be determined. Each bead and DNA molecule are in their own microenvironment, so different tensions can be generated, and the effect of such tensions on binding can be observed. To capture both specific and nonspecific interactions, complex mixtures of the same proteins have been examined using AFM imaging (32). In these experiments, single molecules of lambda DNA were examined for complex DNA folding as a function of multiple proteins. Although the contributions of individual proteins were not measured, the resulting images show the complexity of the structures that can be generated with simple binding proteins (see, for example, figure 4 of Reference 32).

In addition to DNA-compaction proteins, two other classes of bacterial proteins are frequently examined for their effect on DNA folding: transcription factors, such as the lac repressor, and

restriction enzymes. In both cases, these proteins can generate looped DNA that is best studied with single-molecule techniques. Furthermore, it is not just the impact on DNA structure that is of interest but also the mode of binding and how an individual protein can find its specific binding site among so many incorrect sites. With both classes of proteins, there is generally a second DNA-binding site that allows looping of the DNA. For example, investigators followed the kinetics of loop formation induced by the restriction enzymes NaeI and NarI by observing changes in the Brownian motion of tethered particles (33). The various kinetic steps in the reaction could be characterized because each molecule was tracked individually rather than grouped as an ensemble with protein molecules in different intermediate processes (binding, dimerization/looping, catalysis, and dissociation) at different times. Proper characterization of the intermediate steps was verified by varying protein concentration and inhibiting the process through use of calcium as an inhibitor of catalysis. Similarly, the enzyme SfiI was examined via tethered particle motion (TPM)—with the addition of a protein mutated to be catalytically inactive—and a DNA molecule with either two binding sites or only one recognition site to allow observation of nonspecific binding (34). The enzyme EcoRV was examined via an alternative system for distance measurements, SPR with gold particles (35). This technique has the advantage of not being photobleachable, thus allowing longer-term tracking of individual molecules. In contrast, the donor dye in FRET assays can be destroyed by the high laser intensities needed and may become unable to report on a given molecule over long periods of time. However, the much larger size of the gold particles relative to FRET dyes raises concerns about the impact of the particle on the properties of the protein-DNA complex. This system also has a different optimal distance range relative to that of FRET. Sensitive, time-resolved distance measurements have shown that bending can occur simultaneously with binding. Binding can also be separated from cleavage by omission of the cation during the binding phase.

TPM: tethered particle motion

Investigators have also used TPM to examine lac repressor–induced looping so as to understand the impact of bending on loop formation and to assess the different intermediate steps of binding (36). Through use of mutated proteins, the flexibility of the protein hinge was shown to be important in loop formation. Additional characterization of this system as a function of protein concentration and intersite spacing has also been carried out (37). Through examination of discrete binding sites in individual molecules, an understanding of the different intermediates in binding and looping can be obtained. The flexibility of the protein allows loops to be formed even when binding sites are out of phase or spaced very closely together. The ability to see individual molecules go through these transitions is key to understanding the kinetic intermediates. Although it is not yet possible to visualize living cells at this level of detail, stroboscopic illumination of lac repressor has been used to follow binding in living cells (38), thereby allowing connections to be made between isolated proteins observed in vitro with their in vivo counterparts.

2.4. DNA Replication

In more complicated situations, such as the molecular machines involved in DNA replication and other complex systems, all the components of DNA folding, protein binding, and DNA looping and bending interact to provide a highly regulated and efficient set of interactions that require analysis of the individual components as well as of the assembled macromolecules. There are numerous excellent examples of elegant systems such as transcription, recombination, and nucleosome dynamics, but we focus solely on DNA replication.

DNA replication consists of many different activities and interactions. For example, one early single-molecule experiment utilized varying amounts of force to toggle a DNA polymerase between extension and exonuclease activity (39). Because single- and double-stranded DNA respond

TIRF: total internal reflection fluorescence

differently to an externally applied force, the energetically favored reaction can be made to vary. In addition to polymerase activity, many other activities are required for efficient replication. To replicate double-stranded DNA, helicases are generally needed to open up the double helix for polymerization. Through use of optical trapping to apply force, studies of T7 helicase, which is a member of the ring-shaped class of helicases, showed that it plays an active role in DNA unwinding (40). Also required for complete replication are processivity factors. These factors can fall into the class of sliding clamp proteins that keep the polymerase properly positioned with respect to the DNA end being polymerized. FRET studies have examined interactions between the clamp and DNA and have observed the diffusion of the protein along DNA in real time (41) by placing one fluor on the protein and one on an oligonucleotide that, upon binding to a single-stranded DNA, serves as a mimic for a replication fork. Diffusion of the *Escherichia coli* processivity factor was measured on naked, single-stranded DNA or on DNA coated with a single-stranded DNA–binding protein, thus providing valuable information about how multiple components of the system interact. The diffusion rate was assumed to be constant along the length of the DNA, but other single-molecule studies have shown that this is not necessarily the case (42). Using stretched DNAs and a different processivity factor (UL42) labeled with a quantum dot, investigators have found that the rate of diffusion along DNA is much more rapid than had been determined via ensemble techniques because the protein spends much more time on the ends of the DNA than on internal segments. However, these studies are subject to caveats, as the stretched DNA may behave differently than unstretched DNA. For instance, the binding of nucleotides to a polymerase bound to DNA with an inactive 3′ end was examined through use of nanopores (43, 44). After a partially double-stranded DNA was captured, an electric field was applied to accelerate capture of the Klenow fragment of DNA polymerase in an alpha-hemolysin (αHL) nanopore. This complex was then queried with varying concentrations of nucleotide triphosphates to detect binding events and their kinetics. By sequentially assembling different parts of the ternary complex, one can assess each part for its effect on the kinetic lifetimes of the bound states. Kinetics of binding and dissociation can be determined over orders of magnitude of component concentration. Examination of even higher-order complexes and active replication has been accomplished through use of SPR and flow control of reactants to control DNA synthesis. A multiprotein complex can be assembled in the absence of magnesium, the unbound protein components can be removed to prevent reassociation, and the reaction can be initiated with nucleotides and magnesium. For the examination of an even larger spectrum of events (45), ternary complexes of DNA polymerase and cofactors were assembled on replication forks on lambda DNA attached to a glass surface; they were then manipulated with magnetic beads at the other end. The degree of processivity was determined as a function of added primase (DnaG). The unbound polymerase synthesized an average of 10.9 kb in a single event (less than observed in ensemble studies, possibly due to a lack of distinction between multiple events), which then dropped to less than 4 kb upon the addition of DnaG. This effect was independent of whether or not the primase actually synthesized opposite-strand primers. Similarly, complete replisomes for both T7 and *E. coli* polymerase were examined via a total internal reflection fluorescence (TIRF) visualization system and dyes to light up the lagging DNA strand during synthesis (46). Both speed of synthesis and processivity were monitored as laminar flow stretched out each DNA molecule. In addition to leading-strand synthesis, lagging-strand synthesis and resolution of loops have been examined with T7 replisomes (47). Both the dependence of the size and number of loops on various conditions and the determination of the factors affecting release of the loops have been characterized. Similar studies with TIRF visualization were carried out with DNA attached to a lipid bilayer rather than to glass (48). Leading-strand synthesis speed and processivity were examined both with and without lagging-strand synthesis, and large variations among individual molecules were observed.

3. SINGLE-MOLECULE DNA SEQUENCING

3.1. Background

SBS: sequencing by synthesis

Ever since some of the earliest demonstrations of single-molecule detection in solution (49, 50), investigators have hoped to apply this approach to the sequencing of individual DNA or RNA molecules. In some respects, this desire may be driven by the simple realization that life is encoded in one cell by one molecule and thus that performing single-cell sequencing would automatically lead to the ability to detect one nucleotide at a time, regardless of whether that nucleotide is a monomer within the intact genome or a monomer that has been sequentially removed from the genomic polymer. Jett et al. (51) discussed aspects of the possibility of real-time detection of fluorescently labeled nucleotides in a flowing aqueous stream for DNA sequencing, and Ishikawa et al. (52) discussed the prospect of detection of a fluorophore on a surface for DNA sequencing. Investigators considered applying single-molecule approaches to the large-scale sequencing of nucleic acid (NA) molecules soon after the successful demonstration of single-molecule detection; this shift was probably driven by the daunting realization that genome sequences are large in size and ubiquitous in biology. Even before the now-popular concept of deep and broad sequencing for medical purposes or comparative genomics was widely accepted, the initiation of the human genome project in October, 1990, drove the realization that the then–currently available automated sequencers would be economically challenged to acquire the large amounts of sequence data of even one genome. Thus, barely three years after a prominent description of automated fluorescence-based sequencing was published (53), investigators came to realize that a dramatically different technology, rather than a simple increase in the throughput and parallelization of electrophoresis-based sequencing methods, was required. Single-molecule-based approaches seem to have been identified as the most promising technology even before the advent of the $1000-genome goal set by the National Institutes of Health in 2004.

Much of the history of single–DNA molecule sequencing was driven by the essential dichotomy between schemes based on indirect detection of the nucleotides in a DNA polymer versus those based on direct detection. Once the obstacles to optically detecting single fluorophores had been overcome, a salient problem of single-molecule DNA- or RNA-sequencing technology emerged. This challenge consists of the chemistry- and biochemistry-related issues associated with incorporating a fluorescent label on each and every base, then removing the label from each and every base in a near-quantitative manner for many consecutive nucleotides. As such, optically based single-molecule DNA–sequencing methods resemble classical Sanger sequencing in that they use sequencing by synthesis (SBS) but avoid the inherent need for ensembles of molecules and the technical challenges of the massive parallelization required of electrophoretic separations (54, 55). Other next-generation technologies have arisen in the past few years that are based on the optical detection of ensembles of molecules and that use several different molecular processes to decode the sequence, including ligation (56), pyrosequencing (57), hybridization (58), and SBS (59). Several groups have attempted to advance technologies based on nonoptical or direct detection of the individual nucleotides in a single DNA or RNA molecule, and the following sections address efforts in both approaches.

3.2. Historical Efforts

In several papers, Keller and coworkers (51) proposed using a novel approach to the sequencing of long strands of DNA. Their technique involves the progressive release of individually labeled nucleotide bases into a flowing stream of an aqueous liquid and the subsequent real-time detection of a single fluorophore in a highly focused, volume-limited laser spot. The approach promises very

fast readout and a length of read limited only by the ability to label and degrade in a coordinated fashion. Efforts to bring this approach to fruition have mainly focused on the salient problems of (*a*) incorporating a fluorophore in a quantitative fashion at each and every nucleotide of a long fragment of DNA and (*b*) subsequently releasing the labeled nucleotides by an exonuclease (60, 61). The complexities and cost of parallel fluid flow, laser excitation, and detection must be addressed so that the required throughput for genome-scale sequencing can be achieved. So far there have been no reports of sequencing of any useful length, and there appear to be no commercial efforts arising from this approach.

TIR: total internal reflection

3.3. Optical Sequencing by Synthesis

It was not until 2003 that Braslavsky et al. (62) described a single-molecule SBS approach that determined short sequences with unlabeled spacers from template molecules localized on a support. Their approach brought together several concepts to solve problems of single-molecule SBS on a solid support created by the use of fluorescently labeled deoxynucleotide triphosphates at concentrations of approximately several hundred nanomolars. Those concepts include (*a*) the use of FRET to reduce noise from nonspecifically bound fluorophores, (*b*) the use of polyelectrolyte-coated surfaces to further minimize nonspecifically bound fluorophores, and (*c*) the use of TIR illumination to decrease the background fluorescence associated with glass and the aqueous medium. With only two of the four bases labeled with Cy-5, the authors demonstrated the acquisition of a fingerprint of five bases spanning a length of 13 nt. By adding an automated scanning stage, the authors extrapolated the performance of their system to 12 million individual templates in a 25-mm square (roughly 1.2×10^9 strands would provide 10 times the coverage of a whole human genome).

The pioneering work of Quake and coworkers (62, 63) was advanced and commercialized by Helicos BioSciences Corporation, which introduced the world's first single-molecule DNA sequencer early in 2008 (**Figure 3**). Instead of using the FRET system employed by Quake's team, the Helicos group of Harris and coworkers (64) opted to use a single Cy-5 fluorophore excited at 647 nm. The fluorophore was removed by chemical cleavage after each imaging step, in contrast to the laser-mediated photobleaching approach practiced by the Quake group (62). The authors reported that ~280,000 individual molecules, representing fragments of the 7424-nt genome, were subjected to an SBS process that produced an average and median read length of 23 nt. However, the sequencing of each molecule was asynchronous, so some molecules actually provided reads that were 28–34 nt in length. In this early study, more than one fluorescent nucleotide could be inserted per homopolymer region, making an accurate length determination of the homopolymers difficult. To solve this limitation of the chemistry, Bowers et al. (65) developed a set of nucleotide analogs that limited enzymatic incorporation to one per cycle so that homopolymers of any length that are spanned by the length of a read could be accurately sequenced. Single-molecule SBS with reads measuring 33–35 nt in length are useful for resequencing of genomes greater than several megabases. The medical resequencing of human DNA and the comparative sequencing of different model species' DNA can be performed with single-molecule sequencing, as demonstrated by the publication of the first ever single-molecule sequence of a human genome through use of a HeliScope^{TM} single-molecule DNA sequencer (66). This technology has also been applied to the quantification of expressed mRNA, which allowed highly reproducible counting of individual complementary DNA copies of transcripts from a yeast model system (67). In a further effort to reduce any bias introduced by complementary DNA generation, Ozsolak et al. (68) used single-molecule SBS for the direct sequencing of RNA templates in a DNA-primed RNA:DNA hybrid.

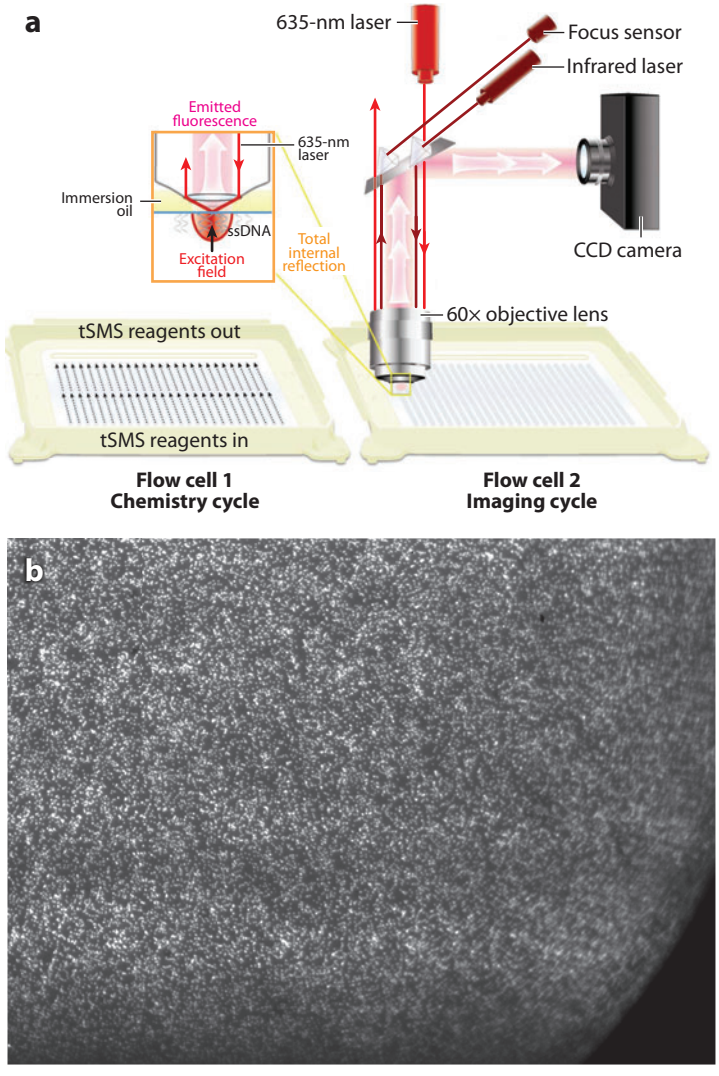

Figure 3

Single-molecule sequencing. (*a*) Schematic of the two–flow cell detection layout of the HeliScope™ single-molecule DNA sequencer. One flow cell is being imaged, and the other is simultaneously undergoing sequencing by synthesis chemistry. (*b*) Single-molecule image of fluorescently labeled primer-template duplexes immobilized on the surface of one channel of a 50-channel flow cell. The image represents 25% of one field of view of ~31,000 single-molecule objects. Used with permission from Helicos BioSciences Corporation. Abbreviations: CCD, charge-coupled device; ssDNA, single-stranded DNA; tSMS™, true single-molecule sequencing.

3.4. Optical Real-Time Sequencing by Synthesis

Another approach to single-molecule SBS, undertaken by Webb and colleagues (69), essentially returns to the concept of confining the illumination volume to a zeptoliter volume through the use of 100-nm-diameter zero-mode waveguides consisting of subwavelength holes. This approach addresses the throughput issues of classical scanning near-field approaches (70) by parallelizing

the detection volumes via arrays of these holes. This single-molecule detection technology was then coupled with an anchored polymerase and gamma-phosphate-linked nucleotide analogs to allow real-time monitoring of fluorophore release while the polymerase incorporated nucleotides in a contiguous SBS (71). The scheme depends upon the ability to discriminate between a labeled nucleotide analog that is transiently occupying the hole by diffusion and a nucleotide that is transiently occupying the hole during enzyme-mediated incorporation. A report from a Pacific Biosciences–based research group (72) showed the four-color sequencing of a 150-nt circular template molecule and identified 1008 base incorporations, suggesting that the circular template used was sequenced ~14 times. Raw error rates of ~20% were reported; however, by sequencing the circular template multiple times in a rolling-circle fashion, the authors reportedly achieved a consensus error rate of 99.3% after oversampling each base. Presumably, large genomes would have to be sheared and circularized by ligation for use in this process; a 10-fold coverage of a human genome was estimated by the authors to require ~140,000 waveguides.

STS: scanning tunneling spectroscopy

STM: scanning tunneling microscopy

3.5. Nonoptical Real-Time

Nonoptical, real-time sequencing comprises (*a*) direct detection and identification of the purine and pyrimidine bases without the need for modification with a reporter and (*b*) real-time detection in a single step. Currently, there are two single-molecule approaches that promise to meet these criteria. One is nanopore-based single-molecule DNA sequencing, and the other is scanning tunneling spectroscopy (STS). Both methods involve scanning an intact DNA molecule one base at a time and detecting either the change in conductance across a nanopore or the difference in conductance among the individual bases themselves.

3.6. Scanning Tunneling Spectroscopy

Direct observation of the atomic-level structure of NAs is a concept that came quickly on the heels of the first demonstrations of scanning tunneling microscopy (STM) in 1981; several very early reports described attempts at such observations (73, 74). It was later recognized that simple imaging of NA molecules was subject to artifacts and that it would not be suitable for the type of resolution necessary to distinguish the linear sequences of DNA and RNA molecules (75). The realization that DNA is capable of charge transport led to the application of an STM probe to measure the conductance of individual molecules in an aqueous solution (76). This set the stage for the use of the tunneling-current variation involving applied voltage to construct current-voltage (*I-V*) curves, which provide a type of electronic spectrum of the observed molecules (77, 78) to improve molecular-level discrimination of the nucleobases. Several groups have attempted to apply the concept of an STS spectrum (*I-V* curves) to differentiate the individual bases of a single DNA molecule on a surface (79–81), with limited success. Recently, Tanaka & Kawai (82) obtained a partial sequence of a stretched single-stranded M13 DNA molecule. In this study, the authors detected only the sequence of guanine residues in a span of 139 nucleotides (**Figure 4**). Advances in the deposition of intact M13 DNA molecules (7249 bases) onto Cu surfaces were key to the ability to partially sequence such a long stretch of bases. The tunneling-current resolution of guanine residues is thought to arise from the fact that guanine's ionization potential is the lowest of the four bases (83). Significant challenges remain for this approach, given that the requirement to obtain a conductivity (*dI/dV*) spectrum takes approximately 1h for the ~200-μm field of observation; in contrast, STM imaging without spectroscopy would allow a significant improvement in throughput. Perhaps a more serious limitation on the use of this technology for

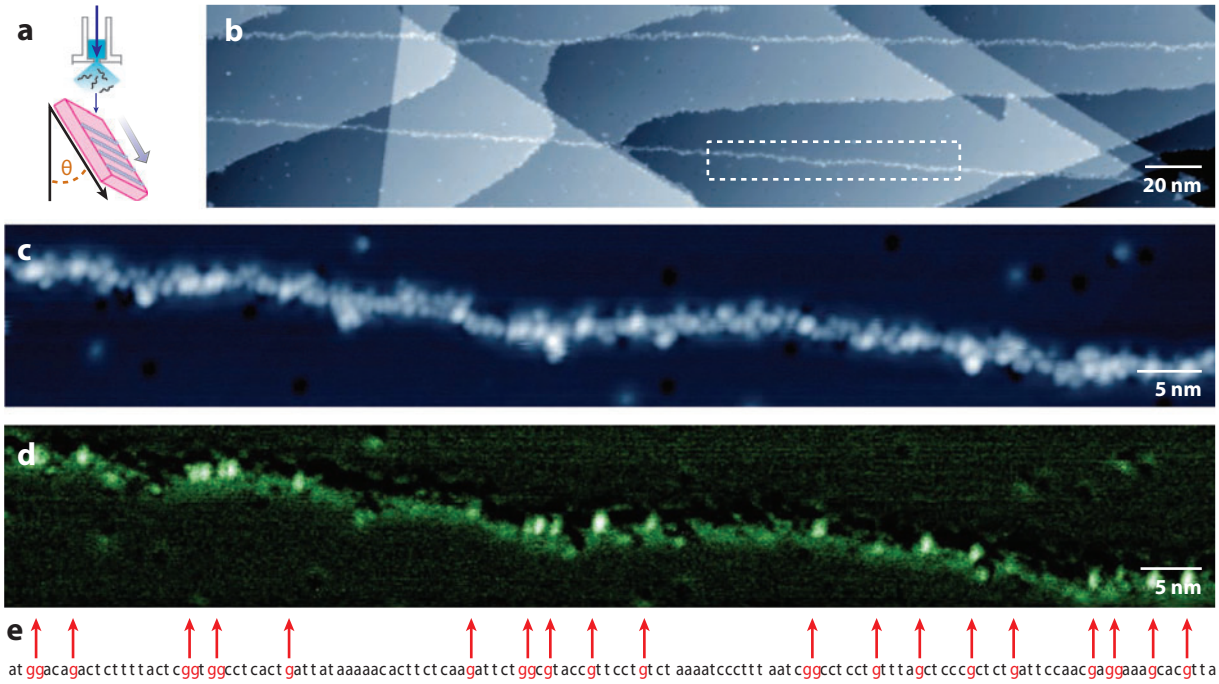

Figure 4

Deposition and scanning tunneling microscopy (STM) analysis of single-stranded M13mp18 DNA molecules on Cu(111). (*a*) Schematic illustration of the oblique injection method. DNA strands tend to align more perpendicular than horizontal to the flow (injection) direction. The Cu(111) substrate is inclined at ~45° to the aqueous solution of DNA, which is introduced from a pulse valve. (*b*) Typical wide-area image of an M13mp18 DNA molecule (voltage, 2 V; current, 5 pA; width, 400 nm). Atomic steps in the Cu(111) substrate form a staircase surface structure. Sections of an M13mp18 DNA molecule are visualized as linear adsorbed material running from top left to bottom right. (*c*) An enlarged view of the rectangular region enclosed by the white dashed line in panel *b* (voltage, −2 V; current, 5 pA; width, 100 nm). (*d*) A conductivity (*dI/dV*) map of the same region as in panel *c* (voltage, −1.5 V; current, 20 pA; width, 100 nm). To maximize the detection of the density of states of guanine, the measurements were made under slightly lower bias conditions than in panel *c*. (*e*) Part of the base sequence of M13mp18 obtained from a database (the sequence of bases at positions 5322 through 5461). To facilitate comparison with the STM data, the guanine sites are indicated by red characters and are connected by red arrows to the corresponding parts of the image. Data taken from Tanaka & Kawai (82).

DNA sequencing is that there is no clear understanding as to whether a differentiating tunneling current can be utilized for discriminating the other three nucleotides of DNA.

3.7. Nanopore-Based Sequencing

The second nonoptical real-time method we review is based on the original proposal (84) that a strand of DNA or RNA can be transported through the lumen of a naturally occurring protein nanopore under the influence of an applied voltage. It has been reported (85) that the size of a 7.2-kbp DNA molecule was experimentally determined, by transport through a silicone nitride nanopore, to be 6.9 + 0.6 kbp. The simplicity of using intact, unmodified NAs in a very parallelizable system is attractive.

The best-studied biological nanopore is the *Staphylococcus aureus* toxin αHL (86). This toxin has an internal diameter of 1.4 nm at its narrowest and a total pore length of approximately 10 nm, and it has been extensively studied for use as a generalized small-molecule sensor. The pore diameter

is just sufficient to accommodate the passage of a single-stranded DNA molecule. A nanopore is embedded in a synthetic lipid bilayer, and the deviations in the ion current across the pore is measured to indicate the stochastic blockage of the current as each base interacts with the wall of the pore. Current levels on the order of picoamperes and dwell times on the order of milliseconds are the expected values to be measured experimentally as the analyte transits the length of the lumen. It was quickly realized that, although current modulations can be detected during the translocation of NA molecules, the length of the nanopore is too long to distinguish the passage of a single nucleotide because ~10–20 nt can be accommodated by this length (87–90). Additionally, the speed of translocation of a single-stranded DNA molecule through a nanopore (1–20 μs per base) makes discrimination of the individual nucleotide-mediated ionic current modulation difficult to measure (88, 91). Nevertheless, the properties of nanopores as biosensors have enabled several interesting single-molecule studies involving nucleotide composition (92), DNA-protein complexes (93), and DNA polymerase activity (94, 95). Single-base substitutions can apparently be determined if the DNA strand is immobilized within the protein nanopore (96). The degree to which the basic αHL pore can be modified to create or enhance specificity of detection is notably extensive. Site-directed mutagenesis and protein-engineering techniques have allowed modification of the internal charge of a pore (97) and the covalent coupling of molecular adapters such as beta-cyclodextrin (98) to the barrel of the protein pore. The former modification was used to increase the number of positively charged amino acids in the lumen of αHL. The frequency of DNA-translocation events thereby increased by almost an order of magnitude, and the threshold voltage for DNA translocation decreased by 50 mV from the 120 mV required for the wild-type protein. The realization that a nanopore could not be used to discriminate single nucleotides of a NA and hence be useful for DNA sequencing triggered the pursuit of several novel concepts, such as tunneling (99) and base-specific modification (100), for increasing the level of discrimination as the NA translocates through the pore.

In a remarkable example of déjà vu, perhaps the most successful demonstration to date of the potential for DNA sequencing using nanopores involves a return to Keller's concept (51) of using exonuclease to progressively remove nucleotides one at a time and to subsequently identify them during translocation through a pore (101). Unambiguous assignment frequencies of 99.4%, 90.3%, 90.9%, and 99.99% for the deoxynucleotide monophosphates of guanine, thymine, adenine, and cytidine, respectively, were observed, and a specific ionic current was detected for 5-methylcytosine, a naturally occurring species associated with transcriptional silencing (102–104). Although the progress in and understanding of the capabilities of nanopores as chemical sensors are substantial, several challenges remain to be solved before this technology is ready for availability as a high-throughput DNA sequencer. As more of the cost of the sequencing technology moves from the instrumentation and reagents to the protein nanopore–sensor device itself, it remains to be seen whether the cost of making such exquisitely engineered protein devices follows the same trend as have more classical discrete electronic components.

3.8. Optical Mapping

Although technically not a single-fluorophore detection technology, the optical mapping approach designed by Schwartz et al. (105) describes the visualization and physical mapping of very large fragments of single molecules of DNA fixed to a surface. Following the adsorption of high-molecular-weight fragments to a surface in a linear orientation, the fragments are subjected to restriction digestion. The intercalation of multiple fluorophores (e.g., YOYO-1) along the length of the molecules allows the contour length of the stretched molecules to be determined. Also, because the restriction fragments are generated from one contiguous molecule, they are

automatically ordered in the proper sequence, allowing a low-resolution sequence map to be obtained without prior knowledge of the primary sequence (106). Use of the integrated intensity of the intercalated fluorophores (instead of a tracing routine) to determine the contour length reportedly allows fragment-length estimation and makes said estimation robust to artifacts of DNA deposition. This method is useful in the comparative genomics of bacterial strains because it is sensitive to the insertions, deletions, inversions, and translocations that characterize many species differences (107, 108). The success of this methodology depends on the effective and high-yield adsorption of single molecules (~500 kb) onto a surface in an elongated fashion and with little to no overlap between individual molecules (109). The use of shear flow from capillary action (rather than from mechanical pumping) in microchannels 8 μm high × 100 μm wide × 10 mm long formed by poly(dimethyl-siloxane) replicas on modified glass surfaces is reportedly sufficient to allow the highly reproducible deposition of molecules up to 3.9 Mb in size or ~1.3 mm in length. Image acquisition has been automated such that a surface area of ~10 mm^2 can be scanned in less than 1 h; automation is required because the optical mapping of a complex genome such as rice (382.17 Mb) requires the analysis of 260,205 individual DNA molecules (110). The optical system used to collect these data employs an argon ion laser–illuminated Zeiss 135M microscope with a Roper Scientific, Inc., cooled charge-coupled device digital camera (109); therefore, an individual laboratory's cost of performing optical mapping would be relatively low. The longer-term question is whether or not optical mapping will continue to play a role in genomics research as the cost of DNA sequencing fueled by single-molecule technologies falls.

3.9. Nongenome Applications of Single-Molecule Sequencing

Our review has focused on genomic applications of sequencing, but there are myriad other applications of sequencing (111). Methods that use small segments of sequence to provide quantitative analysis for measuring gene expressing, assessing copy number variation, and determining metagenomic diversity are particularly well served by single-molecule approaches because they avoid the potential bias and nonlinearity problems associated with amplification. Additionally, many genomic DNA-sequencing projects are undertaken with poor-quality or degraded DNA; such projects could be better carried out with single-molecule approaches because the DNA does not need to be copied prior to sequencing.

4. SUMMARY

Use of single molecules has required the development of new technology for detection and analysis, but these extra efforts have rewarded us with valuable information generated by the study of all individual DNA molecules and not merely an average of molecules. In addition to observing how a DNA molecule behaves, many methodologies also allow manipulation of the molecule so that behavior under extreme conditions can be tested. Many of these manipulations cannot be carried out on a population of molecules, so single-molecule approaches are the only alternative. Similarly, although single-molecule sequencing can generate the same type of data as ensemble sequencing, it surpasses those approaches by (*a*) allowing tiny quantities of DNA to be studied, (*b*) avoiding the pitfalls of amplification, and (*c*) providing the high throughput required for the ubiquitous acquisition of genomic data.

DISCLOSURE STATEMENT

Both authors are employees and stockholders of Helicos BioSciences Corporation.

LITERATURE CITED

1. Kapanidis AN, Strick T. 2009. Biology, one molecule at a time. *Trends Biochem. Sci.* 34:234–43
2. Hall CE. 1956. Visualization of individual macromolecules with the electron microscope. *Proc. Natl. Acad. Sci. USA* 42:801–6
3. Walter NG, Huang CY, Manzo AJ, Sobhy MA. 2008. Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nat. Methods* 5:475–89
4. van Mameren J, Peterman EJ, Wuite GJ. 2008. See me, feel me: methods to concurrently visualize and manipulate single DNA molecules and associated proteins. *Nucleic Acids Res.* 36:4381–89
5. Moffitt JR, Chemla YR, Smith SB, Bustamante C. 2008. Recent advances in optical tweezers. *Annu. Rev. Biochem.* 77:205–28
6. Neuman KC, Nagy A. 2008. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods* 5:491–505
7. Xie XS, Choi PJ, Li GW, Lee NK, Lia G. 2008. Single-molecule approach to molecular biology in living bacterial cells. *Annu. Rev. Biophys.* 37:417–44
8. Raj A, van Oudenaarden A. 2009. Single-molecule approaches to stochastic gene expression. *Annu. Rev. Biophys.* 38:255–70
9. Pondugula S, Kladde MP. 2008. Single-molecule analysis of chromatin: changing the view of genomes one molecule at a time. *J. Cell. Biochem.* 105:330–37
10. Finkelstein IJ, Greene EC. 2008. Single molecule studies of homologous recombination. *Mol. Biosyst.* 4:1094–104
11. Roy R, Hohng S, Ha T. 2008. A practical guide to single-molecule FRET. *Nat. Methods* 5:507–16
12. Ying L, Green JJ, Li H, Klenerman D, Balasubramanian S. 2003. Studies on the structure and dynamics of the human telomeric G quadruplex by single-molecule fluorescence resonance energy transfer. *Proc. Natl. Acad. Sci. USA* 100:14629–34
13. Lee JY, Okumus B, Kim DS, Ha T. 2005. Extreme conformational diversity in human telomeric DNA. *Proc. Natl. Acad. Sci. USA* 102:18938–43
14. Okamoto K, Sannohe Y, Mashimo T, Sugiyama H, Terazima M. 2008. G-quadruplex structures of human telomere DNA examined by single molecule FRET and BrG-substitution. *Bioorg. Med. Chem.* 16:6873–79
15. Lee JY, Kim DS. 2009. Dramatic effect of single-base mutation on the conformational dynamics of human telomeric G-quadruplex. *Nucleic Acids Res.* 37:3625–34
16. Jena PV, Shirude PS, Okumus B, Laxmi-Reddy K, Godde F, et al. 2009. G-quadruplex DNA bound by a synthetic ligand is highly dynamic. *J. Am. Chem. Soc.* 131:12522–23
17. Zeng ZX, Zhao Y, Hao YH, Tan Z. 2005. Tetraplex formation of surface-immobilized human telomere sequence probed by surface plasmon resonance using single-stranded DNA binding protein. *J. Mol. Recognit.* 18:267–71
18. Pomerantz AK, Moerner WE, Kool ET. 2008. Visualization of long human telomere mimics by single-molecule fluorescence imaging. *J. Phys. Chem. B* 112:13184–87
19. Shim JW, Tan Q, Gu LQ. 2009. Single-molecule detection of folding and unfolding of the G-quadruplex aptamer in a nanopore nanocavity. *Nucleic Acids Res.* 37:972–82
20. Lynch S, Baker H, Byker SG, Zhou D, Sinniah K. 2009. Single molecule force spectroscopy on G-quadruplex DNA. *Chemistry* 15:8113–16
21. McKinney SA, Déclais AC, Lilley DM, Ha T. 2003. Structural dynamics of individual Holliday junctions. *Nat. Struct. Biol.* 10:93–97
22. Joo C, McKinney SA, Lilley DM, Ha T. 2004. Exploring rare conformational species and ionic effects in DNA Holliday junctions using single-molecule spectroscopy. *J. Mol. Biol.* 341:739–51
23. Hohng S, Joo C, Ha T. 2004. Single-molecule three-color FRET. *Biophys. J.* 87:1328–37
24. McKinney SA, Freeman AD, Lilley DM, Ha T. 2005. Observing spontaneous branch migration of Holliday junctions one step at a time. *Proc. Natl. Acad. Sci. USA* 102:5715–20
25. Karymov M, Daniel D, Sankey OF, Lyubchenko YL. 2005. Holliday junction dynamics and branch migration: single-molecule analysis. *Proc. Natl. Acad. Sci. USA* 102:8186–91

26. Karymov MA, Bogdanov A, Lyubchenko YL. 2008. Single molecule fluorescence analysis of branch migration of Holliday junctions: effect of DNA sequence. *Biophys. J.* 95:1239–47
27. Lushnikov AY, Bogdanov A, Lyubchenko YL. 2003. DNA recombination: Holliday junctions dynamics and branch migration. *J. Biol. Chem.* 278:43130–34
28. Hohng S, Zhou R, Nahas MK, Yu J, Schulten K, et al. 2007. Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the Holliday junction. *Science* 318:279–83
29. Skoko D, Wong B, Johnson RC, Marko JF. 2004. Micromechanical analysis of the binding of DNA-bending proteins HMGB1, NHP6A, and HU reveals their ability to form highly stable DNA-protein complexes. *Biochemistry* 43:13867–74
30. Dame RT, Noom MC, Wuite GJ. 2006. Bacterial chromatin organization by H-NS protein unraveled using dual DNA manipulation. *Nature* 444:387–90
31. Dixit S, Singh-Zocchi M, Hanne J, Zocchi G. 2005. Mechanics of binding of a single integration-host-factor protein to DNA. *Phys. Rev. Lett.* 94:118101
32. Maurer S, Fritz J, Muskhelishvili G. 2009. A systematic in vitro study of nucleoprotein complexes formed by bacterial nucleoid-associated proteins revealing novel types of DNA organization. *J. Mol. Biol.* 387:1261–76
33. Van Den Broek B, Vanzi F, Normanno D, Pavone FS, Wuite GJ. 2006. Real-time observation of DNA looping dynamics of Type IIE restriction enzymes NaeI and NarI. *Nucleic Acids Res.* 34:167–74
34. Laurens N, Bellamy SR, Harms AF, Kovacheva YS, Halford SE, Wuite GJ. 2009. Dissecting protein-induced DNA looping dynamics in real time. *Nucleic Acids Res.* 37:5454–64
35. Reinhard BM, Sheikholeslami S, Mastroianni A, Alivisatos AP, Liphardt J. 2007. Use of plasmon coupling to reveal the dynamics of DNA bending and cleavage by single EcoRV restriction enzymes. *Proc. Natl. Acad. Sci. USA* 104:2667–72
36. Vanzi F, Broggio C, Sacconi L, Pavone FS. 2006. Lac repressor hinge flexibility and DNA looping: single molecule kinetics by tethered particle motion. *Nucleic Acids Res.* 34:3409–20
37. Han L, Garcia HG, Blumberg S, Towles KB, Beausang JF, et al. 2009. Concentration and length dependence of DNA looping in transcriptional regulation. *PLoS One* 4:e5621
38. Elf J, Li GW, Xie XS. 2007. Probing transcription factor dynamics at the single-molecule level in a living cell. *Science* 316:1191–94
39. Wuite GJ, Smith SB, Young M, Keller D, Bustamante C. 2000. Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. *Nature* 404:103–6
40. Johnson DS, Bai L, Smith BY, Patel SS, Wang MD. 2007. Single-molecule studies reveal dynamics of DNA unwinding by the ring-shaped T7 helicase. *Cell* 129:1299–309
41. Laurence TA, Kwon Y, Johnson A, Hollars CW, O'Donnell M, et al. 2008. Motion of a DNA sliding clamp observed by single molecule fluorescence spectroscopy. *J. Biol. Chem.* 283:22895–906
42. Komazin-Meredith G, Mirchev R, Golan DE, van Oijen AM, Coen DM. 2008. Hopping of a processivity factor on DNA revealed by single-molecule assays of diffusion. *Proc. Natl. Acad. Sci. USA* 105:10721–26
43. Wilson NA, Abu-Shumays R, Gyarfas B, Wang H, Lieberman KR, et al. 2009. Electronic control of DNA polymerase binding and unbinding to single DNA molecules. *Am. Chem. Soc. Nano* 3:995–1003
44. Hurt N, Wang H, Akeson M, Lieberman KR. 2009. Specific nucleotide binding and rebinding to individual DNA polymerase complexes captured on a nanopore. *J. Am. Chem. Soc.* 131:3772–78
45. Tanner NA, Hamdan SM, Jergic S, Loscha KV, Schaeffer PM, et al. 2008. Single-molecule studies of fork dynamics in *Escherichia coli* DNA replication. *Nat. Struct. Mol. Biol.* 15:170–76
46. Tanner NA, Loparo JJ, Hamdan SM, Jergic S, Dixon NE, van Oijen AM. 2009. Real-time single-molecule observation of rolling-circle DNA replication. *Nucleic Acids Res.* 37:e27
47. Hamdan SM, Loparo JJ, Takahashi M, Richardson CC, van Oijen AM. 2009. Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature* 457:336–39
48. Yao NY, Georgescu RE, Finkelstein J, O'Donnell ME. 2009. Single-molecule analysis reveals that the lagging strand increases replisome processivity but slows replication fork progression. *Proc. Natl. Acad. Sci. USA* 106:13236–41
49. Peck K, Stryer L, Glazer AH, Mathies RA. 1989. Single-molecule fluorescence detection: autocorrelation criterion and experimental realization with phycoerythrin. *Proc. Natl. Acad. Sci. USA* 86:4087–91

50. Shera EB, Seitzinger N, Davis LM, Keller RA, Soper SA. 1990. Detection of single fluorescent molecules. *Chem. Phys. Lett.* 174:553–57
51. Jett JH, Keller RA, Martin JC, Marrone BL, Moyzis RK, et al. 1989. High-speed DNA sequencing: an approach based upon fluorescence detection of single molecules. *J. Biomol. Struct. Dyn.* 5:301–9
52. Ishikawa M, Hirano K, Hayakawa T, Hosoi S, Brenner S. 1994. Single-molecule detection by laser-induced fluorescence technique with a position-sensitive photon-counting apparatus. *Jpn. J. Appl. Phys.* 33:1571–76
53. Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, et al. 1986. Fluorescence detection in automated DNA sequence analysis. *Nature* 321:674–79
54. Blazej RG, Kumaresan P, Mathies RA. 2006. Microfabricated bioprocessor for integrated nanoliter-scale Sanger DNA sequencing. *Proc. Natl. Acad. Sci. USA* 103:7240–45
55. Fredlake CP, Hert DG, Root BE, Barron AE. 2008. Polymer systems designed specifically for DNA sequencing by microchip electrophoresis: a comparison with commercially available materials. *Electrophoresis* 29:4652–62
56. Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, et al. 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309:1728–32
57. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–80
58. Pihlak A, Baurén G, Hersoug E, Lönnerberg P, Metsis A, Linnarsson S. 2008. Rapid genome sequencing with short universal tiling probes. *Nat. Biotechnol.* 26:676–84
59. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, et al. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456:53–59
60. Werner JH, Cai H, Jett JH, Reha-Krantz L, Keller RA, Goodwin PM. 2003. Progress towards single-molecule DNA sequencing: a one color demonstration. *J. Biotechnol.* 102:1–14
61. Werner JH, Cai H, Keller RA, Goodwin PM. 2005. Exonuclease I hydrolyzes DNA with a distribution of rates. *Biophys. J.* 88:1403–12
62. Braslavsky I, Hebert B, Kartalov E, Quake SR. 2003. Sequence information can be obtained from single DNA molecules. *Proc. Natl. Acad. Sci. USA* 100:3960–64
63. Kartalov EP, Unger MA, Quake SR. 2003. Polyelectrolyte surface interface for single-molecule fluorescence studies of DNA polymerase. *Biotechniques* 34:505–10
64. Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, et al. 2008. Single-molecule DNA sequencing of a viral genome. *Science* 320:106–9
65. Bowers J, Mitchell J, Beer E, Buzby PR, Causey M, et al. 2009. Virtual terminator nucleotides for next-generation DNA sequencing. *Nat. Methods* 6:593–95
66. Pushkarev D, Neff NF, Quake SR. 2009. Single-molecule sequencing of an individual human genome. *Nat. Biotechnol.* 27:847–50
67. Lipson D, Raz T, Kieu A, Jones DR, Giladi E, et al. 2009. Quantification of the yeast transcriptome by single-molecule sequencing. *Nat. Biotechnol.* 27:652–58
68. Ozsolak F, Platt AR, Jones DR, Reifenberger JG, Sass LE, et al. 2009. Direct RNA sequencing. *Nature* 461:814–18
69. Levene MJ, Korlach J, Turner SW, Foquet M, Craighead HG, Webb WW. 2003. Zero-mode waveguides for single-molecule analysis at high concentrations. *Science* 299:682–86
70. Betzig E, Trautman JK. 1992. Near-field optics: microscopy, spectroscopy, and surface modification beyond the diffraction limit. *Science* 257:189–95
71. Korlach J, Levene M, Foquet M, Turner SW, Craighead HG, Webb WW. 2003. *Single molecule DNA sequence profiling in zero-mode waveguides using gamma-phosphate linked nucleotide analogs*. Presented at Annu. Meet. Biophys. Soc., 46th, San Antonio
72. Eid J, Fehr A, Gray J, Luong K, Lyle J, et al. 2009. Real-time DNA sequencing from single polymerase molecules. *Science* 323:133–38
73. Amrein M, Stasiak A, Gross H, Stoll E, Travaglini G. 1988. Scanning tunneling microscopy of recA-DNA complexes coated with a conducting film. *Science* 240:514–15
74. Beebe TP, Wilson TE, Ogletree DF, Katz JE, Balhorn R, et al. 1989. Direct observation of native DNA structures with the scanning tunneling microscope. *Science* 243:370–72

75. Clemmer CR, Beebe TP Jr. 1991. Graphite: a mimic for DNA and other biomolecules in scanning tunneling microscope studies. *Science* 251:640–42
76. Xu B, Zhang P, Li X, Tao N. 2004. Direct conductance measurement of single molecules in aqueous solution. *Nano Lett.* 4:1105–8
77. Robinson AL. 1985. A spatially resolved surface spectroscopy. *Science* 229:1074–76
78. Iijima M, Kato T, Nakanishi S, Watanabe H, Kimura K, et al. 2005. STM/STS study of electron density of states at the bases sites in the DNA alternating copolymers. *Chem. Lett.* 34:1084–85
79. Yoshida Y, Nojima Y, Tanaka H, Kawai T. 2007. Scanning tunneling spectroscopy of single-strand deoxyribonucleic acid for sequencing. *J. Vac. Sci. Technol. B* 25:1071–123
80. Shapir E, Cohen H, Calzolari A, Cavazzoni C, Ryndyk DA, et al. 2008. Electronic structure of single DNA molecules resolved by traverse scanning tunneling spectroscopy. *Nat. Mater.* 7:68–74
81. Ryndyk DA, Shapir E, Porath D, Calozolari A, Di Felice R, Cuniberti G. 2009. Scanning tunneling spectroscopy of single DNA molecules. *Am. Chem. Soc. Nano* 3:1651–56
82. Tanaka H, Kawai T. 2009. Partial sequencing of a single DNA molecule with a scanning tunnelling microscope. *Nat. Nanotechnol.* 4:518–22
83. Yang X. Wang X-B, Vorpagel ER, Wang L-S. 2004. Direct experimental observation of the low ionization potentials of guanine in free oligonucleotides by using photoelectron spectroscopy. *Proc. Natl. Acad. Sci. USA* 101:17588–92
84. Kasianowicz JJ, Brandin E, Branton D, Deamer DW. 1996. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. USA* 93:13770–73
85. Fologea D, Brandin E, Uplinger J, Branton D, Li J. 2007. DNA conformation and base number simultaneously determined in a nanopore. *Electrophoresis* 28:3186–92
86. Bayley H, Cremer PS. 2001. Stochastic sensors inspired by biology. *Nature* 413:226–30
87. Akeson M, Branton D, Kasianowicz JJ, Brandin E, Deamer DW. 1999. Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys. J.* 77:3227–33
88. Meller A, Nivon L, Brandin E, Golovchenko J, Branton D. 2000. Rapid nanopore discrimination between single polynucleotide molecules. *Proc. Natl. Acad. Sci. USA* 97:1079–84
89. Meller A, Nivon L, Branton D. 2001. Voltage-driven DNA translocations through a nanopore. *Phys. Rev. Lett.* 86:3435–38
90. Ashkenasy N, Sanchez-Quesada J, Bayley H, Ghardiri MR. 2005. Recognizing a single base in an engineered protein nanopore equipped with a molecular adapter. *Angew. Chem. Int. Ed. Engl.* 44:1401–4
91. Fologea D, Uplinger J, Thomas B, McNabb DS, Li J. 2005. Slowing DNA translocation in a solid-state nanopore. *Nano Lett.* 5:1734–37
92. Deamer DW, Branton D. 2002. Characterization of nucleic acids by nanopore analysis. *Acc. Chem. Res.* 35:817–25
93. Hornblower B, Coombs A, Whitaker RD, Kolomeisky A, Picone SJ, et al. 2007. Single-molecule analysis of DNA-protein complexes using nanopores. *Nat. Methods* 4:315–17
94. Benner S, Chen RJA, Wilson NA, Abu-Shumays R, Hurt N, et al. 2007. Sequence-specific detection of individual DNA polymerase complexes in real time using a nanopore. *Nat. Nanotechnol.* 2:718–24
95. Cockroft SL, Chu J, Amorin M, Ghadiri MR. 2008. A single-molecule nanopore device detects DNA polymerase activity with single-nucleotide resolution. *J. Am. Chem. Soc.* 130:818–20
96. Purnell RF, Schmidt JJ. 2009. Discrimination of single base substitutions in a DNA strand immobilized in a biological nanopore. *Am. Chem. Soc. Nano* 3:2533–38
97. Maglia G, Restrepo MR, Mikhailova E, Bayley H. 2008. Enhanced translocation of single DNA molecules through α-hemolysin nanopores by manipulation of internal charge. *Proc. Natl. Acad. Sci. USA* 105:19720–25
98. Wu H-C, Astier Y, Maglia G, Mikhailova, Bayley H. 2007. Protein nanopores with covalently attached molecular adapters. *J. Am. Chem. Soc.* 129:16142–48
99. He J, Lin L, Zhang P, Lindsay S. 2007. Identification of DNA base pairing via tunnel-current decay. *Nano Lett.* 7:3854–58
100. Mitchell N, Howorka S. 2008. Chemical tags facilitate the sensing of individual DNA strands with nanopores. *Angew. Chem. Int. Ed. Engl.* 47:5565–68

101. Clarke J, Wu H-C, Jayasinghe L, Patel A, Reid S, Bayley H. 2009. Continuous base identification for single-molecule nanopore DNA sequencing. *Nat. Nanotechnol.* 4:265–70
102. Lister R, Ecker JR. 2009. Finding the fifth base: genome-wide sequencing of cytosine methylation. *Genome Res.* 19:959–66
103. Miranda TB, Jones PA. 2007. DNA methylation: the nuts and bolts of repression. *J. Cell. Physiol.* 213:384–90
104. Gopolakrishnan S, Van Emburgh BO, Robertson KD. 2008. DNA methylation in development and human disease. *Mutat. Res.* 647:30–80
105. Schwartz DC, Li X, Hernandez LI, Ramnarain SP, Huff EJ, Wang YK. 1993. Ordered restriction maps of *Saccharomyces cerevisiae* chromosomes constructed by optical mapping. *Science* 262:110–14
106. Valouev A, Schwartz DC, Zhou S, Waterman MS. 2006. An algorithm for assembly of ordered restriction maps from single DNA molecules. *Proc. Natl. Acad. Sci. USA* 103:15770–75
107. Zhou S, Kile A, Bechner M, Place M, Kvikstad E, et al. 2004. Single-molecule approach to bacterial genomic comparisons via optical mapping. *J. Bacteriol.* 186:7773–82
108. Shukla SK, Kislow J, Briska A, Henkhaus J, Dykes C. 2009. Optical mapping reveals a large genetic inversion between two methicillin-resistant *Staphyloccus aureas* strains. *J. Bacteriol.* 191:5717–23
109. Dimalanta ET, Lim A, Runnheim R, Lamers C, Churas C, et al. 2004. A microfluidic system for large DNA molecule arrays. *Anal. Chem.* 76:5293–301
110. Zhou S, Bechner MC, Place M, Churas CP, Pape L, et al. 2007. Validation of rice genome sequence by optical mapping. *BMC Genomics* 8:278–96
111. Kahvejian A, Quackenbush J, Thompson JF. 2008. What would you do if you could sequence everything? *Nat. Biotechnol.* 26:1125–33
112. Churchill ME. 2003. Watching flipping junctions. *Nat. Struct. Biol.* 10:73–75

Annual Review of Analytical Chemistry
Volume 3, 2010

Contents

An Editor's View of Analytical Chemistry (the Discipline)
Royce W. Murray 1

Integrated Microreactors for Reaction Automation: New Approaches to Reaction Development
Jonathan P. McMullen and Klavs F. Jensen 19

Ambient Ionization Mass Spectrometry
Min-Zong Huang, Cheng-Hui Yuan, Sy-Chyi Cheng, Yi-Tzu Cho, and Jentaie Shiea 43

Evaluation of DNA/Ligand Interactions by Electrospray Ionization Mass Spectrometry
Jennifer S. Brodbelt 67

Analysis of Water in Confined Geometries and at Interfaces
Michael D. Fayer and Nancy E. Levinger 89

Single-Molecule DNA Analysis
J. William Efcavitch and John F. Thompson 109

Capillary Liquid Chromatography at Ultrahigh Pressures
James W. Jorgenson 129

In Situ Optical Studies of Solid-Oxide Fuel Cells
Michael B. Pomfret, Jeffrey C. Owrutsky, and Robert A. Walker 151

Cavity-Enhanced Direct Frequency Comb Spectroscopy: Technology and Applications
Florian Adler, Michael J. Thorpe, Kevin C. Cossel, and Jun Ye 175

Electrochemical Impedance Spectroscopy
Byoung-Yong Chang and Su-Moon Park 207

Electrochemical Aspects of Electrospray and Laser Desorption/Ionization for Mass Spectrometry
Mélanie Abonnenc, Liang Qiao, BaoHong Liu, and Hubert H. Girault 231

Adaptive Microsensor Systems
Ricardo Gutierrez-Osuna and Andreas Hierlemann 255

Confocal Raman Microscopy of Optical-Trapped Particles in Liquids
Daniel P. Cherney and Joel M. Harris 277

Scanning Electrochemical Microscopy in Neuroscience
Albert Schulte, Michaela Nebel, and Wolfgang Schuhmann 299

Single-Biomolecule Kinetics: The Art of Studying a Single Enzyme
Victor I. Claessen, Hans Engelkamp, Peter C.M. Christianen, Jan C. Maan, Roeland J.M. Nolte, Kerstin Blank, and Alan E. Rowan 319

Chiral Separations
A.M. Stalcup 341

Gas-Phase Chemistry of Multiply Charged Bioions in Analytical Mass Spectrometry
Teng-Yi Huang and Scott A. McLuckey 365

Rotationally Induced Hydrodynamics: Fundamentals and Applications to High-Speed Bioassays
Gufeng Wang, Jeremy D. Driskell, April A. Hill, Eric J. Dufek, Robert J. Lipert, and Marc D. Porter 387

Microsystems for the Capture of Low-Abundance Cells
Udara Dharmasiri, Małgorzata A. Witek, Andre A. Adams, and Steven A. Soper 409

Advances in Mass Spectrometry for Lipidomics
Stephen J. Blanksby and Todd W. Mitchell 433

Indexes

Cumulative Index of Contributing Authors, Volumes 1–3 467

Cumulative Index of Chapter Titles, Volumes 1–3 470

Errata

An online log of corrections to *Annual Review of Analytical Chemistry* articles may be found at http://arjournals.annualreviews.org/errata/anchem.