

Electrokinetic trapping at the one nanometer limit

Alexander P. Fields^a and Adam E. Cohen^{b,1}

^aBiophysics Program and ^bDepartments of Chemistry and Chemical Biology and of Physics, Harvard University, 12 Oxford Street, Cambridge, MA 02138

Edited* by W. E. Moerner, Stanford University, Stanford, CA, and approved April 14, 2011 (received for review March 3, 2011)

Anti-Brownian electrokinetic traps have been used to trap and study the free-solution dynamics of large protein complexes and long chains of DNA. Small molecules in solution have thus far proved too mobile to trap by any means. Here we explore the ultimate limits on trapping single molecules. We developed a feedback-based anti-Brownian electrokinetic trap in which classical thermal noise is compensated to the maximal extent allowed by quantum measurement noise. We trapped single fluorophores with a molecular weight of <1 kDa and a hydrodynamic radius of 6.7 Å for longer than one second, in aqueous buffer at room temperature. This achievement represents an 800-fold decrease in the mass of objects trapped in solution, and opens the possibility to trap and manipulate any soluble molecule that can be fluorescently labeled. To illustrate the use of this trap, we studied the binding of unlabeled RecA to fluorescently labeled single-stranded DNA. Binding of RecA induced changes in the DNA diffusion coefficient, electrophoretic mobility, and brightness, all of which were measured simultaneously and on a molecule-by-molecule basis. This device greatly extends the size range of molecules that can be studied by room temperature feedback trapping, and opens the door to further studies of the binding of unmodified proteins to DNA in free solution.

A longstanding challenge in single-molecule spectroscopy has been to observe a small molecule in solution for an extended time, without surface tethering or other mechanical immobilization. Stable observation becomes more difficult as the particle decreases in size, because smaller objects diffuse more quickly, in accordance with the Stokes–Einstein relation. Gold nanoparticles as small as 18 nm in diameter, corresponding to a mass of 35 MDa, have been trapped using laser tweezers (1). Below this size laser tweezers fail because the trapping force is proportional to the volume of the trapped object. Real-time feedback provides an alternate strategy, and has been used to trap single atoms in vacuum (2). The anti-Brownian electrokinetic (ABEL) trap uses feedback to suppress Brownian motion in solution and can confine particles as small as the 800 kDa complex of the chaperonin GroEL (3–5). A 104 kDa protein, allophycocyanin, was recently studied in an ABEL trap in which the viscosity was increased with 50% glycerol to slow the Brownian motion (6). Past attempts to trap small-molecule fluorophores in aqueous solution resulted in transient confinement, but not stable trapping (3). Small-molecule fluorophores are the tiniest objects that one can conceive of trapping in aqueous solution. If a particular fluorophore can be trapped, then so too can any molecule to which it is attached.

Laser tweezers and the ABEL trap both confine small objects in solution, but the two technologies enable different kinds of measurements. The optical forces of laser tweezers enable precise (subnanometer) localization of the trapped object, and permit application of precisely calibrated point forces for the purpose of force spectroscopy (7). The ABEL feedback strategy enables trapping of smaller objects, including individual molecules, but tracking imprecision, algorithmic errors, and feedback latency permit residual Brownian motion, typically with an amplitude of several hundred nanometers. The feedback forces in the ABEL trap are applied as body forces to all molecules in the solution, so this trap is not appropriate for force spectroscopy. The ABEL trap is primarily suited to nonperturbative observation of single-molecule dynamics (8).

One goal of single-molecule trapping is to study binding interactions in solution that may be perturbed by surface tethering. An obstacle to reaching this goal is that if both species are labeled, at least one must be at a concentration incommensurate with single-molecule fluorescence for dissociation constants greater than approximately 0.1 nM. An alternative strategy is to detect binding via the influence of an unlabeled molecule on the photophysical or transport properties of a labeled and trapped binding partner. Binding of RecA to DNA has been extensively studied in bulk (9), structurally (10), and at the single-molecule level (11–13). Here we report observations of changes in transport coefficients of freely diffusing short DNA oligos upon binding of RecA.

The design of our ABEL trap is illustrated in Fig. 1. The position of an object is tracked in real-time via fluorescence, and electrokinetic feedback is applied via two orthogonal pairs of platinum electrodes to cancel Brownian motion in the plane. Out-of-plane motion is physically restricted by the fused silica walls of the nanofabricated sample cell. When necessary, the walls are chemically passivated to minimize surface interactions with the analyte. Fully three-dimensional electrokinetic trapping, which avoids problems of surface interactions, has been demonstrated on 0.6 μm polystyrene beads, but not yet for single molecules (14).

Our optical detection scheme uses a scanned laser and a single-photon counting detector, with spatial information derived from the location of the laser at the instant each photon is detected (3, 15). The position of the laser is specified by a field-programmable gate array (FPGA), similar to the strategy outlined in a recent theoretical article (16). The key innovation that improves our trap's performance is a statistically rigorous real-time tracking and feedback system, implemented on the FPGA. This system operates at the quantum limit imposed by the finite information carried by each fluorescence photon. The LabVIEW FPGA code is publicly available (17).

Each detected photon specifies the particle's location with a precision limited by the standard deviation of the Gaussian laser spot, measured to be 360 nm. Averaging over many photons improves localization precision, but neglects the motion of the particle between photon detections. A more sophisticated strategy is to construct a “running average” in which recently detected photons are weighted more heavily than those detected earlier. A Kalman filter (18, 19), schematized in the upper left of Fig. 1, appropriately weights the information from each photon (see *Methods* for details) and is simple enough to implement in the FPGA and to run at high speed. Thus the tracking system operates close to the physical limits imposed by diffusion, diffraction, and the finite rate of photon detection events.

Author contributions: A.P.F. and A.E.C. designed research; A.P.F. performed research; A.P.F. and A.E.C. contributed new reagents/analytic tools; A.P.F. and A.E.C. analyzed data; and A.P.F. and A.E.C. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

¹To whom correspondence should be addressed. E-mail: cohen@chemistry.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1103554108/-DCSupplemental.

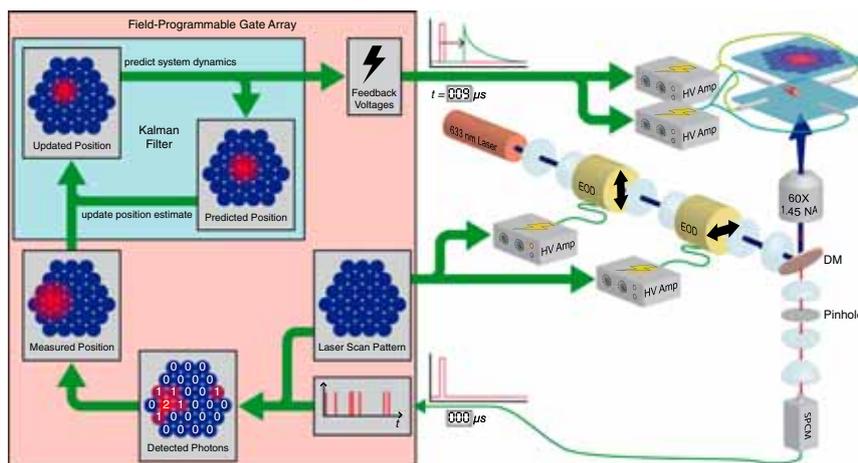


Fig. 1. Instrumentation. Two electrooptic deflectors (EODs) scan light from a 633 nm HeNe laser among a set of 27 discrete points with a dwell time of $3.1 \mu\text{s}$ per point. Fluorescence emitted by a fluorophore in the sample cell (*Top Right*) is separated from the illumination by a dichroic mirror (DM) and detected by an avalanche photodiode single-photon counting module (SPCM). A Kalman filter implemented on a field-programmable gate array incorporates the information from each photon detection into a running estimate of the fluorophore position, and generates appropriate feedback voltages that are amplified and applied to the sample cell via four platinum electrodes. The latency of the feedback loop (between photon detection and voltage response) is $9 \mu\text{s}$.

Results and Discussion

In the absence of feedback, fluorophores of Alexa 647 diffused across the laser scan pattern with an average residence time of 2 ms (*Fig. 2A, Top*). When feedback was applied, fluorophores that diffused into the trap were quickly pushed to the trap center and held for an average of 800 ms prior to photobleaching, photobleaching, or escape (*Fig. 2A, Middle*), corresponding to the collection of an average of 37,000 photons per event. Some single fluorophores were trapped for as long as 10 s, yielding as many as 450,000 photons. The time-averaged illumination was uniform throughout the region explored by the molecule, so residual molecular motion did not lead to brightness fluctuations. The fluorescence intensity was constant during each event and from one event to the next, and every event ended with a quantal step to background fluorescence, establishing that the trapped species contained only one fluorophore. Molecule-by-molecule analysis of diffusion coefficients (see below) yielded a narrow distribution peaked around $325 \mu\text{m}^2/\text{s}$, which matched the value obtained in bulk (20) and confirmed that every event corresponded to a free fluorophore. Occasional short-lived positive intensity spikes during trapping events signified the approach of a second fluorophore near the trap; the Brownian motion of the fluorophores was uncorrelated, so after a few milliseconds one diffused away. Segments of dsDNA (30 bp), doubly labeled with Alexa 647 showed two clear photobleaching steps, each equal in intensity to that of a trapped single fluorophore (*Fig. 2A, Bottom*), further confirming that the objects trapped in the free dye sample were single fluorophores. *Fig. 2B* shows in red a time-averaged CCD image of a series of trapped single fluorophores (*Movie S1*). Displayed in blue is the time-averaged laser scan pattern.

To determine the loss mechanisms from the trap, we studied the trapping time of single fluorophores as a function of laser power (*Fig. 2C*). A trapping event was considered to end when the fluorescence dropped to background for longer than $300 \mu\text{s}$. The mean trapping time was nonmonotonic in laser power, indicating that trapping time was limited by photon statistics and diffusional escape at low power, and by photobleaching or photobleaching at high power. The trap was typically operated under conditions to maximize mean trapping time, in which case rates of diffusional escape and photobleaching or bleaching were approximately equal.

Photobleaching rates in the ABEL trap were higher than in typical surface-tethered experiments because high count rates were required to achieve stable feedback. Furthermore, we interpreted every blinking event as the end of a single-molecule

trajectory. This procedure contributed to the shorter reported observation time in the ABEL trap compared to surface-tethered

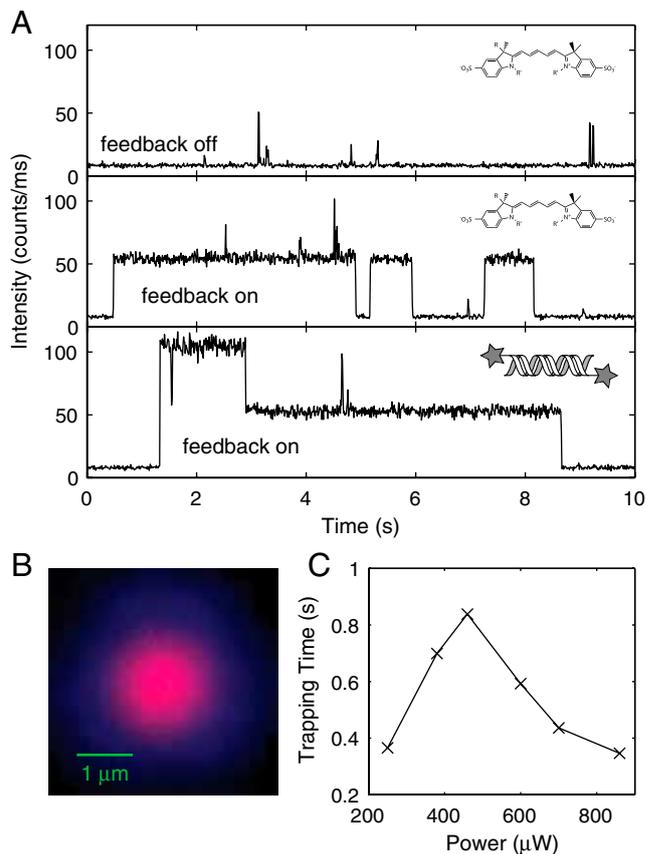


Fig. 2. Trapped molecules. (A) In the absence of applied feedback, the fluorescence of Alexa 647 molecules (chemical structure inset) showed brief bursts averaging 2 ms in duration (*Top*). When feedback was applied, molecular residence time was greatly enhanced (*Middle*). Trapping of 30 bp dsDNA doubly labeled with Alexa 647 showed two-step photobleaching (*Bottom*). In the structure of Alexa 647 (compound 9 in ref. 21), *R* refers to *N*-hydroxysuccinimide hexan-6-ylate and *R'* refers to 3-sulfonatopropyl. (B) Time-averaged image of a series of trapped Alexa 647 molecules (red) merged in software with an image of the laser scan pattern (blue). (C) Mean trapping time as a function of laser power, showing the balance between diffusional escape and photobleaching or bleaching.

experiments where one typically averages over blinks. The distribution of trapping times had a long tail, possibly due to variations in rates of photobleaching and photoblinking caused by variations in the concentration of oxygen and triplet quenchers.

The photon-by-photon recording of each trapping event enabled quantitative determination of the spatial trajectory, transport coefficients and photophysical properties of each molecule, with a precision far beyond that of any other single-molecule technique. We developed a maximum-likelihood assumed density filter (ADF) to perform these calculations (see *Methods*). We applied the ADF to trajectories of single trapped molecules of Alexa 647 to determine the strength and relaxation time of the trap (Fig. 3*A* and *B*). The fluorophore was tightly constrained to the center of the trap, with a rms deviation of 416 nm, well within the approximately 5 μm diameter of the laser scan.

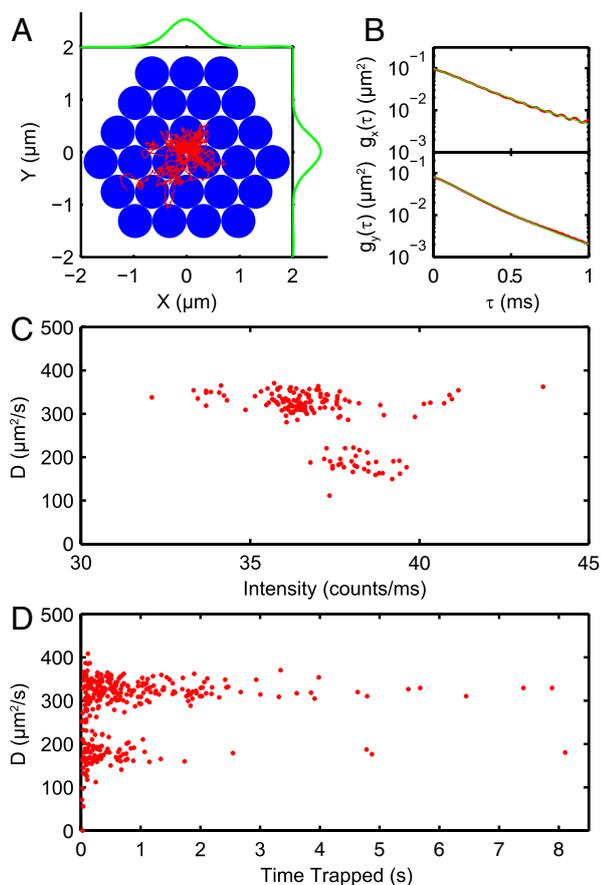


Fig. 3. Performance of the ABEL trap. (*A*) First 10 ms of the reconstructed trajectory of a single trapped molecule of Alexa 647 (red), plotted over a cartoon of the 27-point scan pattern (blue). Time-averaged probability densities are plotted along each axis (green). The mean precision with which each point in the trajectory was localized is 240 nm, which is less than the width of the laser spot (360 nm) because the estimate includes information from several photons. The molecule was confined with a rms deviation of about 416 nm, well within the approximately 5 μm diameter of the laser scan. From the equipartition theorem, we calculate effective spring constants of 0.022 pN/ μm and 0.026 pN/ μm in the x and y directions, respectively. (*B*) Time-autocorrelation of the position of the molecule along each axis (red), with fit to a single exponential (green). Relaxation times are $\tau_x = 290 \mu\text{s}$ and $\tau_y = 240 \mu\text{s}$. (*C-D*) Determination of single-molecule diffusion coefficients. Molecules were trapped from a mixture of free Alexa 647 and 30 nt ssDNA singly labeled with Alexa 647. The fitted diffusion coefficient for each molecule is plotted versus the brightness of the molecule (*C*) or the duration for which it was trapped (*D*). The two diffusion peaks have mean values of $178 \pm 2 \mu\text{m}^2/\text{s}$ and $325 \pm 2 \mu\text{m}^2/\text{s}$ (SEM), corresponding to the ssDNA and free dye species, respectively.

To test our ability to measure diffusion coefficients of small objects, we studied a mixture of free Alexa 647 dye and 30 nt ssDNA, singly labeled with Alexa 647. Many single molecules were sequentially trapped until photobleaching. Based on their intensity or trapping time alone, the two species in the sample were indistinguishable, but they were clearly resolved by their diffusion coefficients (Fig. 3*C* and *D*). From experiments in which only one species was trapped, we associate the lower diffusion coefficient with the DNA species and the greater one with free Alexa 647.

These highly precise measurements of single-molecule diffusion coefficients enable quantitative comparison with bulk measurements and theory. The diffusion coefficient obtained in the ABEL trap for Alexa 647, $D_{\text{AF}} = 325 \pm 2 \mu\text{m}^2/\text{s}$ (SEM), is consistent with the value of $330 \mu\text{m}^2/\text{s}$ measured by two-focus fluorescence correlation spectroscopy (FCS) (20). The diffusion coefficient of the 30-mer ssDNA, $D_{30\text{-mer}} = 178 \pm 2 \mu\text{m}^2/\text{s}$, matches that predicted by the Zimm model for a polymer in good solvent, 166 $\mu\text{m}^2/\text{s}$, calculated with persistence length and rise per base taken from laser tweezers experiments (22) and no adjustable parameters (see *Methods* and ref. 23).

The diffusion coefficients can be converted to hydrodynamic radii via the Stokes–Einstein relation. The ABEL trap data yielded a population-average radius for Alexa 647 of $6.76 \pm 0.03 \text{ \AA}$ (SEM) and for 30-mer ssDNA of $12.3 \pm 0.2 \text{ \AA}$. The diffusion coefficient of each single molecule was determined more precisely as trapping time increased (Fig. 3*D*); we calculate a precision of roughly $20 \mu\text{m}^2/(\text{s}\sqrt{\text{Hz}})$ in determining D . Thus a single molecule of Alexa 647, trapped for 1 s, yielded an estimate of its hydrodynamic radius with a precision of 0.5 \AA (SD). The ability to measure hydrodynamic radii of single molecules in free solution so precisely may be useful for observing subtle conformational shifts.

Encouraged by our success in trapping and characterizing single fluorophores and small DNA molecules, we next studied the interaction of ssDNA with *Escherichia coli* RecA, a protein known to form helical filaments on ssDNA in the presence of ATP (9). The ssDNA sample was 60 nucleotides long, singly labeled at its 5' terminus with Alexa 647. In the absence of RecA, the ssDNA is expected to form a random coil with a contour length of 33.6 nm, a persistence length of 7.5 \AA , and a radius of gyration of 27 \AA (22). This length of DNA is sufficient to nucleate a RecA filament containing up to 20 monomers of RecA (24), but vastly shorter than the approximately 900 nm persistence length of the RecA filament (13). Binding of RecA is expected to convert the ssDNA from a random coil to a semi-rigid rod with a radius of roughly 4 nm and a rise of 5.1 \AA per nucleotide (10), corresponding to a total length of 30.6 nm and a hydrodynamic radius of 10.5 nm (see *Methods*).

Before studying the interaction of RecA with ssDNA in the ABEL trap, we first used conventional FCS to confirm binding of RecA to ssDNA and to study the ensemble-averaged effects of binding upon the diffusion coefficient and fluorescence brightness of the ssDNA. Addition of unlabeled RecA (1 μM) and ATP (1 mM) to a sample of ssDNA induced a 60% drop in the ensemble-averaged diffusion coefficient of the DNA (Fig. 4*A*), and a 40% increase in the average molecular brightness. The decrease in diffusion coefficient was consistent with a change in geometry from a compact random coil to an extended rod. The increase in brightness upon binding of RecA to fluorescently labeled ssDNA likely reflects changes in the chemical environment of the fluorophore, and is consistent with a previous report in which a different fluorophore was used (25). RecA in the absence of DNA had no detectable fluorescence. These FCS measurements provided no information on the underlying distributions of single-molecule brightness and diffusion coefficient, and provided no information about the ensemble-averaged or single-molecule values of the electrokinetic mobility.

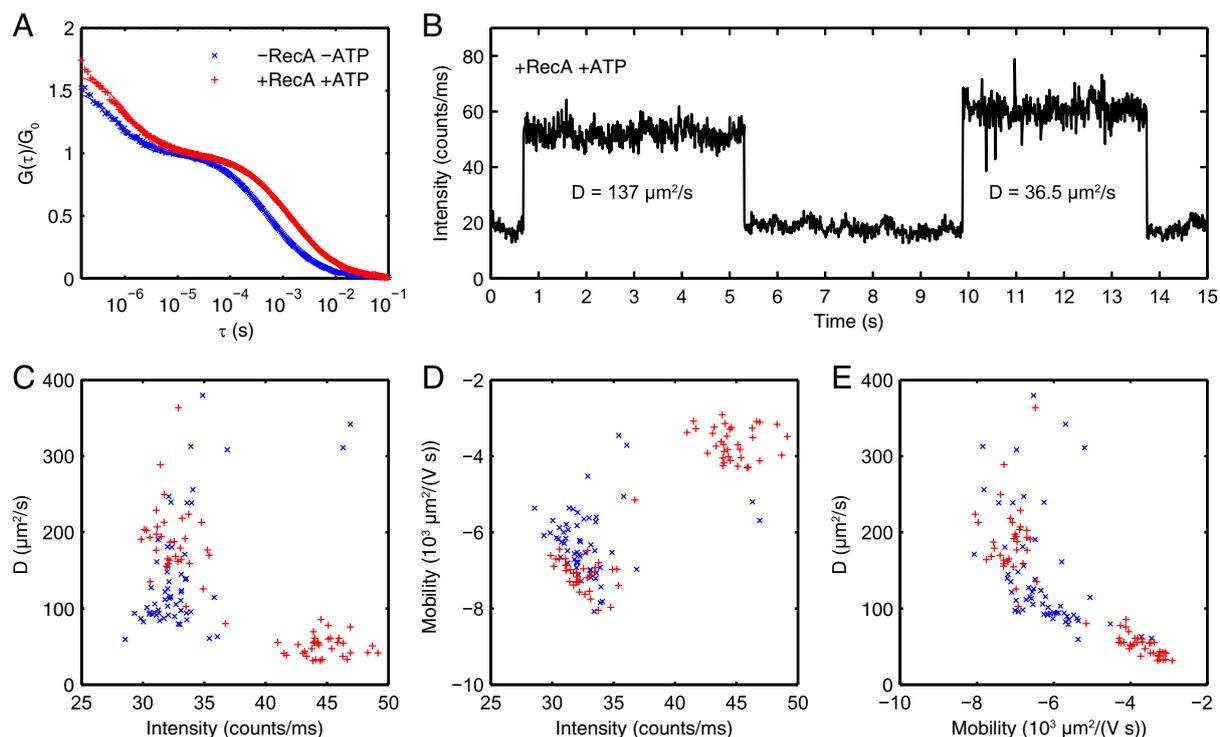


Fig. 4. RecA binding to single-stranded DNA. (A) Raw FCS data (points) and least-squares fits (lines) for a sample of 60-mer ssDNA labeled with Alexa 647 in the presence or absence of 1 μM RecA and 1 mM ATP. Plots are normalized to the fit value of G_0 , which neglects the triplet fraction (see *Methods* for fit function). The lowering of average diffusion coefficient upon binding of RecA is visible as a longer autocorrelation decay time. (B) Trapping 60-mer ssDNA molecules in the presence of RecA and ATP revealed two species, exemplified by this fluorescence timetrace. The first species, identified as bare ssDNA, was dimmer and had higher diffusion coefficient; the second, identified as RecA nucleoprotein filament, was brighter and diffused more slowly. (C–E) Three-parameter molecular profiling. (x) ssDNA without RecA, (+) ssDNA with RecA and ATP. Plotting the brightness, diffusion coefficient, and electrokinetic mobility of trapped ssDNA molecules reveals that binding of RecA induced changes in all three.

We next trapped single molecules of the fluorescently labeled ssDNA, first in the absence of RecA, and then in the presence of RecA (1 μM) and ATP (1 mM). Each trapped molecule was characterized simultaneously for its brightness, diffusion coefficient, and electrokinetic mobility (Fig. 4 B–E). These three parameters indicate different aspects of the molecular structure: brightness of the fluorophore is sensitive to the chemical environment at the 5' end of the ssDNA; diffusion coefficient is sensitive to hydrodynamic radius of the entire molecular complex; and electrokinetic mobility is sensitive to both charge and viscous drag. Thus each molecule was characterized with high precision in a multidimensional parameter space, allowing facile identification of heterogeneous subpopulations.

In the absence of RecA, we observed a homogeneous population of trapped molecules with diffusion coefficient $113 \pm 4 \mu\text{m}^2/\text{s}$ (SEM), a mobility of $-6.3 \pm 0.1 \times 10^3 \mu\text{m}^2/(\text{V s})$ and a mean molecular brightness of $32.0 \pm 0.2 \times 10^3$ photons/s, which we associate with bare ssDNA. Addition of RecA and ATP led to the appearance of a second subpopulation with diffusion coefficient of $50 \pm 2 \mu\text{m}^2/\text{s}$, mobility of $-3.65 \pm 0.07 \times 10^3 \mu\text{m}^2/(\text{V s})$ and a mean molecular brightness of $44.7 \pm 0.3 \times 10^3$ photons/s.

We compared the diffusion coefficients of the bare ssDNA and the nucleoprotein filament to theoretical predictions based on the expected geometries of these compounds. The measured diffusion coefficient of the ssDNA, $113 \pm 4 \mu\text{m}^2/\text{s}$, is in good agreement with the Zimm prediction of $124 \mu\text{m}^2/\text{s}$ for 60-mer ssDNA. The measured diffusion coefficient of the RecA-ssDNA nucleoprotein filaments, $50 \pm 2 \mu\text{m}^2/\text{s}$, is lower than that of the bare ssDNA, consistent with an increase in hydrodynamic radius upon RecA binding. However, the measured diffusion coefficient is significantly larger than expected from a rigid-rod model of the filament, $21 \mu\text{m}^2/\text{s}$. Several modifications to the rigid-rod geometry may account for this discrepancy, including: bent or curved

nucleoprotein structure [as seen in some electron microscope images (26)], incomplete coverage of the DNA by RecA, or multiple RecA domains separated by floppy dislocations. Distinguishing among these scenarios will require varying conditions such as the length of the DNA template and the RecA concentration. The discrepancy is unlikely to be due to transient dissociation of RecA monomers from the nucleoprotein filament, as transport coefficients were not affected by replacement of ATP with ATP γS , an analog known to reduce the rate of dissociation (27).

The decrease in absolute value mobility upon addition of RecA is consistent with a decrease in net charge or an increase in drag. The change in Stokes drag can be inferred from the change in diffusion coefficient, discussed above, but the relevance of Stokes drag to the electrophoretic mobility depends on the details of the ionic cloud around the complex. In the Hückel regime, in which the particle is much smaller than the Debye length of the buffer, the electrophoretic velocity is determined by a force balance between the Coulombic pull of the field and the Stokes drag on the moving particle. In the Smoluchowski regime, in which the particle is large compared to the Debye length, the electric and shear forces are both localized within the Debye layer and both grow proportionally to the size of the particle (28). Thus the electrophoretic mobility is independent of the size of the particle. In the case at hand, the size of the ssDNA-RecA complex is comparable to the Debye length of the buffer, estimated to be 1.5 nm. Thus, the molecule is between the Hückel and Smoluchowski regimes, where theoretical estimates of mobility are difficult. Free-solution electrophoretic mobilities of ssDNA (29) and DNA-protein complexes (30) have been measured and are similar to our single-molecule values, though the buffer composition in these experiments differed from ours, preventing quantitative comparison.

Despite the presence of RecA, many of the trapped molecules remained in the state associated with bare ssDNA. The presence of precisely two clearly resolved peaks in the multidimensional single-molecule distributions indicates highly cooperative binding of RecA to ssDNA, consistent with earlier measurements in which nucleoprotein filament formation was measured as a function of RecA concentration (31).

The unique ability to measure simultaneously the diffusion coefficient, electrokinetic mobility, and brightness of each trapped molecule, to high precision, allows species to be distinguished on the basis of size, charge, and photophysical properties, in free solution and in complex mixtures. Unlike fluctuation techniques such as FCS (32) and the photon-counting histogram (33), the ABEL trap characterizes individual molecules in isolation, allowing compilation of the full distribution of each measured parameter and enabling multiple parameters to be correlated at the single-molecule level. The long observation times per molecule in the ABEL trap provide significantly more precise information than is obtained from photon burst analysis (34, 35). Furthermore, the tracking data enables unambiguous separation of transport and photophysical dynamics, which are otherwise conflated in single-point confocal techniques.

Finally, the ABEL trap opens the possibility of observing molecular transitions as they occur; such transitions cannot be resolved in observations of passively diffusing molecules unless they occur on a timescale faster than the diffusion time (approximately 1–10 ms), but may be observed on timescales as long as seconds in the ABEL trap. In the present RecA data we observed occasional transitions indicative of RecA binding to an already trapped molecule of ssDNA; however, these transitions were too infrequent to merit detailed analysis.

In contrast to the ensemble-averaged FCS data, the ABEL trap data clearly discerned hidden heterogeneity in the sample of RecA-ssDNA. The multiparameter molecule-by-molecule data enabled a quantitative comparison to models of the RecA nucleoprotein filament. These capabilities are expected to be broadly useful in contexts beyond studying protein–DNA interactions.

Trapping small-molecule fluorophores in aqueous solution is the ultimate size limit of feedback trapping in solution. The median size of human proteins is 375 amino acids (36), corresponding to a diameter of approximately 4 nm (37). With previous trapping technology, only the largest of proteins and complexes, with diameter > 15 nm, could be trapped. The present ABEL trap extends this range to include all soluble proteins.

Methods

The optics, device design and fabrication, electronics, and sample preparation are described in *SI Text*.

Kalman Filter Feedback Algorithm. The Kalman filter is an algorithm that interprets the past record of detected photons and applied voltages to construct a Gaussian likelihood distribution for the location of the particle in the present. The estimate for the particle's position during the i th time bin, given all of the information recorded up to and including the j th time bin, is characterized by a mean $\hat{\mathbf{x}}_{ij}$ and variance $\hat{\mathbf{p}}_{ij}$. The filter operates recursively: To calculate a new estimate, the previous estimate is revised to account for the most recent observations and the expected motion of the particle.

The number of photons, n_k , detected during the laser's residence at the k th scan point are tallied and used to "update" the estimate according to

$$\hat{\mathbf{x}}_{k|k} = \frac{w^2 \hat{\mathbf{x}}_{k|k-1} + n_k \hat{\mathbf{p}}_{k|k-1} \mathbf{c}_k}{w^2 + n_k \hat{\mathbf{p}}_{k|k-1}}, \quad \hat{\mathbf{p}}_{k|k} = \frac{w^2 \hat{\mathbf{p}}_{k|k-1}}{w^2 + n_k \hat{\mathbf{p}}_{k|k-1}},$$

where \mathbf{c}_k and w are the mean and standard deviation, respectively, of the laser spatial intensity distribution during bin k . In this way, the information from photon arrivals is combined with the previous position estimate, with weighting coefficients determined by the uncertainty in each. The beam

profile and spot positions were measured prior to trapping experiments by scanning a small bead immobilized on a coverslip through the scan pattern using a piezo scanning stage (Thorlabs SCXYZ100).

When the laser's residence at the k th spot ends, a new estimate for the location of the particle is constructed, with mean and variance "predicted" according to

$$\hat{\mathbf{x}}_{k+1|k} = \hat{\mathbf{x}}_{k|k} + \mu \Delta t \mathbf{V}_k \quad \hat{\mathbf{p}}_{k+1|k} = \hat{\mathbf{p}}_{k|k} + 2D \Delta t,$$

where μ is the user-estimated electrokinetic mobility, \mathbf{V}_k is the voltage applied during bin k , D is the user-estimated diffusion constant, and Δt is the duration of the bin. At the beginning of the experiment, the estimated location of the particle is initialized at the trap center and the variance in this estimate is set to an arbitrary large value. The influence of these initial conditions decays after a few tens of microseconds. The Kalman filter is propagated each time the laser is moved to a new position (every 3.1 μ s).

Feedback voltages are calculated according to the equation

$$\mathbf{V}_{k+3} = -\frac{\hat{\mathbf{x}}_{k+3|k}}{\mu \Delta t},$$

where

$$\hat{\mathbf{x}}_{k+3|k} = \hat{\mathbf{x}}_{k+1|k} + \mu \Delta t (\mathbf{V}_{k+1} + \mathbf{V}_{k+2}).$$

This formula is used because the update step cannot be calculated immediately, so a delay of two bin periods is necessary before the feedback is applied to the sample. The voltage is capped at a maximum magnitude along each dimension (typically 50 V) to avoid nonlinear effects, sample heating, and degradation of the solution.

ADF Algorithm. The Kalman filter is an approximation to the optimal tracking strategy: it treats non-Gaussian probability distributions as Gaussian to allow calculations in real time. The ADF algorithm we developed for postprocessing is a recursive Bayesian estimator, which correctly handles background photons and Poisson-distributed shot noise. A derivation of the algorithm in the context of optical tracking and a discussion of its merits and limitations will be presented in a forthcoming manuscript (38, 39). An implementation in MATLAB is publicly available (17).

The ADF projects each (posterior) estimate distribution onto a Gaussian shape parameterized by two-dimensional mean $\hat{\mathbf{x}}$ and two-by-two covariance matrix $\hat{\Sigma}$. The update equations become

$$\hat{\mathbf{x}}_{k|k} = \frac{1}{L_k} \sum_{m=0}^{\infty} l_m \hat{\chi}_m$$

$$\hat{\Sigma}_{k|k} = \left(\frac{1}{L_k} \sum_{m=0}^{\infty} l_m (\hat{\chi}_m \hat{\chi}_m^T + \hat{\Psi}_m) \right) - \hat{\mathbf{x}}_{k|k} \hat{\mathbf{x}}_{k|k}^T,$$

where

$$l_m \equiv e^{-B} \sum_{i=\max(n_k-m, 0)}^{n_k} \frac{(-1)^{m+i-n_k} B^i S^m}{(n_k - i)! i! (m + i - n_k)!} \frac{|\hat{\Psi}_m|^{1/2}}{|\hat{\Sigma}_{k|k-1}|^{1/2}}$$

$$\times e^{-(m(\hat{\mathbf{x}}_{k|k-1} - \mathbf{c}_k)^T (m \hat{\Sigma}_{k|k-1} + \mathbf{W})^{-1} (\hat{\mathbf{x}}_{k|k-1} - \mathbf{c}_k) / 2)}$$

$$\hat{\chi}_m \equiv \hat{\Psi}_m (\hat{\Sigma}_{k|k-1}^{-1} \hat{\mathbf{x}}_{k|k-1} + m \mathbf{W}^{-1} \mathbf{c}_k) \quad \hat{\Psi}_m \equiv (\hat{\Sigma}_{k|k-1}^{-1} + m \mathbf{W}^{-1})^{-1}.$$

B is the average number of background photons detected per spot residence period, S is the expected number of photons detected from a fluorophore positioned at the center of the Gaussian laser spot for an entire spot residence period, \mathbf{W} is the two-by-two covariance matrix of a Gaussian approximation to the laser spot shape, and other parameters are as defined previously. All vectors are treated as column vectors, with T or $^{-1}$ indicating matrix transposition or inversion, respectively. The likelihood of each data point is

$$L_k = \sum_{m=0}^{\infty} l_m.$$

The sums converge, and we truncate when the fractional changes due to additional terms are $<10^{-6}$.

The prediction equations are

$$\hat{\mathbf{x}}_{k+1|k} = \hat{\mathbf{x}}_{k|k} + \mu \Delta t \mathbf{V}_k \quad \hat{\Sigma}_{k+1|k} = \hat{\Sigma}_{k|k} + 2D \Delta t \mathbf{I} + \nu \mu^2 \Delta t^2 \mathbf{V}_k \mathbf{V}_k^T,$$

where \mathbf{I} is the two-by-two identity matrix, ν is the dimensionless ratio between the variance and the mean square of the effective mobility, and the other parameters are as defined previously. The effective mobility is treated as a Gaussian random variable to reflect the observation that molecules do not always respond identically to an applied voltage, perhaps due to spatial inhomogeneity of the field, or to unconstrained fluctuations in the vertical position of the particle in the trap and consequent changes in field strength and drag on the particle.

The overall log-likelihood of the entire data series is

$$\ln(\Lambda) = \sum_k \ln(L_k).$$

Maximum-likelihood parameter estimates are found by gradient ascent of this function. To convert the units of the calculated mobility parameter, the electric field strength was estimated as the applied voltage divided by the length of the trapping region (30 μm).

Diffusion Coefficient Calculation. Theoretical diffusion coefficients were calculated using the Zimm model presented in (23), specifically equation (4.84)

$$D_G = 0.2030 \frac{k_B T}{\sqrt{6} \eta_s R_g},$$

- Hansen PM, Bhatia VK, Harrit N, Oddershede L (2005) Expanding the optical trapping range of gold nanoparticles. *Nano Lett* 5:1937–1942.
- Kubanek A, et al. (2009) Photon-by-photon feedback control of a single-atom trajectory. *Nature* 462:898–901.
- Cohen AE, Moerner WE (2008) Controlling Brownian motion of single protein molecules and single fluorophores in aqueous buffer. *Opt Express* 16:6941–6956.
- Jiang Y, et al. (2008) Hardware-based anti-Brownian motion and application to ATP binding stoichiometry in multi-subunit enzymes. *Proc SPIE* 7038:1–12.
- Cohen AE, Moerner WE (2005) Method for trapping and manipulating nanoscale objects in solution. *Appl Phys Lett* 86:093109.
- Goldsmith RH, Moerner WE (2010) Watching conformational- and photodynamics of single fluorescent proteins in solution. *Nat Chem* 2:179–186.
- Greenleaf WJ, Woodside MT, Abbondanzieri EA, Block SM (2005) Passive all-optical force clamp for high-resolution laser trapping. *Phys Rev Lett* 95:208102.
- Fields AP, Cohen AE (2010) Anti-Brownian traps for studies on single molecules. *Methods Enzymol* 475:149–174.
- Lusetti SL, Cox MM (2002) The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu Rev Biochem* 71:71–100.
- Chen Z, Yang H, Pavletich NP (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* 453:489–494.
- Joo C, et al. (2006) Real-time observation of RecA filament dynamics with single monomer resolution. *Cell* 126:515–527.
- Galletto R, Amitani I, Baskin RJ, Kowalczykowski SC (2006) Direct observation of individual RecA filaments assembling on single DNA molecules. *Nature* 443:875–878.
- Hegner M, Smith SB, Bustamante C (1999) Polymerization and mechanical properties of single RecA-DNA filaments. *Proc Natl Acad Sci USA* 96:10109–10114.
- King JK (2009) Microfluidic device for the electrokinetic manipulation of single molecules. Master's thesis (University of Tennessee, Knoxville, TN).
- Enderlein J (2000) Tracking of fluorescent molecules diffusing within membranes. *Appl Phys B* 71:773–777.
- Wang Q, Moerner WE (2010) Optimal strategy for trapping single fluorescent molecules in solution using the ABEL trap. *Appl Phys B* 99:23–30.
- ABEL Trap Code Repository, Available at <https://www2.lsddiv.harvard.edu/labs/cohen/Research/TrapSingMol/Code/Code.htm>.
- Kalman RE (1960) A new approach to linear filtering and prediction problems. *Trans ASME J Basic Eng* 82:35–45.
- Welch G, Bishop G (2006) An introduction to the Kalman filter. *Technical Report* (University of North Carolina, Chapel Hill, NC), TR 95-041.
- Kapusta P (2010) Absolute diffusion coefficients: compilation of reference data for FCS calibration. *Application Note* (PicoQuant GmbH, Berlin), Available at http://www.picoquant.com/technote/appnote_diffusion_coefficients.pdf.
- Leung W, Cheung C, Yue S (2007) Modified carbocyanine dyes and their conjugates. *US Patent* 7,671,214.

where η_s is the solvent viscosity and R_g is the radius of gyration. We calculated R_g using $\sqrt{6}R_g = \sqrt{N}b = \sqrt{L}b$, where b is the effective bond length (twice the persistence length), N is the number of effective bond segments, and L is the total contour length. For single-stranded DNA, we used an effective bond length of 1.5 nm and a contour length of 0.56 nm per nucleotide (22).

To calculate a theoretical diffusion coefficient of RecA bound to ssDNA, we applied the rigid-rod model of (40)

$$D_G = k_B T \frac{\ln(L/d + \gamma)}{3\pi\eta_s L},$$

where L and d are the rod length and diameter, respectively, and γ is an end-effect correction term. We used a rise per nucleotide of 0.51 nm and a nucleofilament diameter of 4 nm, based on the structure of (10), corresponding to $\gamma = 0.46$.

FCS Fits. FCS traces were fit to a 2D diffusion model with triplet state, adapted from (41)

$$G(\tau) = G_0 \frac{1 - F + F \exp(-\tau/\tau_F)}{1 - F} \frac{1}{1 + \tau/\tau_D} + G_\infty.$$

Nonlinear least-square five-parameter fits were performed using MATLAB.

ACKNOWLEDGMENTS. This project was supported in part by National Science Foundation (NSF) Grant CHE-0910824. Fabrication was performed at the Harvard Faculty of Arts and Sciences (FAS) Center for Nanoscale Systems, a member of the National Nanotechnology Infrastructure Network, supported by NSF Award ECS-0335765. Computations were run on the Odyssey cluster supported by the FAS Research Computing Group.

- Smith SB, Cui Y, Bustamante C (1996) Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. *Science* 271:795–799.
- Doi M, Edwards SF (1988) *The theory of polymer dynamics* (Oxford University Press, New York).
- Brenner SL, et al. (1987) RecA protein-promoted ATP hydrolysis occurs throughout RecA nucleoprotein filaments. *J Biol Chem* 262:4011–4016.
- Silver MS, Fersht AR (1982) Direct observation of complexes formed between RecA protein and a fluorescent single-stranded deoxyribonucleic acid derivative. *Biochemistry* 21:6066–6072.
- Heuser J, Griffith J (1989) Visualization of RecA protein and its complexes with DNA by quick-freeze/deep-etch electron microscopy. *J Mol Biol* 210:473–484.
- McEntee K, Weinstock GM, Lehman IR (1981) Binding of the RecA protein of *Escherichia coli* to single- and double-stranded DNA. *J Biol Chem* 256:8835–8844.
- Henry DC (1931) The cataphoresis of suspended particles. Part I. The equation of cataphoresis. *Proc R Soc London Ser A* 133:106–129.
- Stellwagen E, Stellwagen NC (2002) Determining the electrophoretic mobility and translational diffusion coefficients of DNA molecules in free solution. *Electrophoresis* 23:2794–2803.
- Foulds GJ, Eitzkorn FA (1998) A capillary electrophoresis mobility shift assay for protein-DNA binding affinities free in solution. *Nucleic Acids Res* 26:4304–4305.
- Menetski JP, Kowalczykowski SC (1985) Interaction of RecA protein with single-stranded DNA: Quantitative aspects of binding affinity modulation by nucleotide cofactors. *J Mol Biol* 181:281–295.
- Magde D, Elson E, Webb WW (1972) Thermodynamic fluctuations in a reacting system-measurement by fluorescence correlation spectroscopy. *Phys Rev Lett* 29:705–708.
- Chen Y, Müller JD, So PTC, Gratton E (1999) The photon counting histogram in fluorescence fluctuation spectroscopy. *Biophys J* 77:553–567.
- Shera EB, Seitzinger NK, Davis LM, Keller RA, Soper SA (1990) Detection of single fluorescent molecules. *Chem Phys Lett* 174:553–557.
- Deniz AA, et al. (1999) Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Förster distance dependence and subpopulations. *Proc Natl Acad Sci USA* 96:3670–3675.
- Brocchieri L, Karlin S (2005) Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Res* 33:3390–3400.
- Tyn MT, Gusek TW (1990) Prediction of diffusion coefficients of proteins. *Biotechnol Bioeng* 35:327–338.
- Minka TP (2001) A family of algorithms for approximate Bayesian inference. PhD thesis (Massachusetts Institute of Technology, Cambridge, MA).
- McHale K, Berglund AJ, Mabuchi H (2004) Bayesian estimation for species identification in single-molecule fluorescence microscopy. *Biophys J* 86:3409–3422.
- Tirado MM, de La Torre JG (1979) Translational friction coefficients of rigid, symmetric top macromolecules. application to circular cylinders. *J Chem Phys* 71:2581–2587.
- Lakowicz JR (2006) *Principles of Fluorescence Spectroscopy* (Springer, New York).