Accurately determining single molecule trajectories of molecular motion on surfaces

Kevin Claytor,[1] Saumyakanti Khatua,[1] Jason M. Guerrero,[1] Alexei Tcherniak,[1] James M. Tour,[1] and Stephan Link[1,2,a)]
[1]*Department of Chemistry, Rice University, 6100 Main Street, Houston, Texas 77005, USA*
[2]*Department of Electrical and Computer Engineering, Rice University, 6100 Main Street, Houston, Texas 77005, USA*

(Received 25 January 2009; accepted 17 March 2009; published online 28 April 2009)

This paper presents a method for simultaneously determining multiple trajectories of single molecules from sequential fluorescence images in the presence of photoblinking. The tracking algorithm is computationally nondemanding and does not assume a model for molecular motion, which allows one to determine correct trajectories even when a distribution of movement speeds is present. We applied the developed procedure to the important problem of monitoring surface motion of single molecules under ambient conditions. By limiting the laser exposure using sample scanning confocal microscopy, long-time trajectories have been extracted without the use of oxygen scavengers for single fluorescent molecules. Comparison of the experimental results to simulations showed that the smallest diffusion constants extracted from the trajectories are limited by detector shot noise giving error in locating the positions of the individual molecules. The simulations together with the single molecule trajectories and distributions of diffusion constants allowed us therefore to distinguish between mobile and immobile molecules. Because the analysis algorithm only requires a time series of images, the procedure presented here can be used in conjunction with various imaging methodologies to study a wide range of diffusion processes. © *2009 American Institute of Physics*. [DOI: 10.1063/1.3118982]

I. INTRODUCTION

Single molecule studies have become useful and pervasive as a powerful tool to measure the dynamics of individual molecules rather than an ensemble average.[1–4] By imaging single molecules, the averages, deviations, and all moments of the distributions can be obtained. This allows one to verify many more predictions than can be achieved with ensemble measurements.[1–4] The wealth of information that can be gathered from studying individual molecules has therefore made single molecule techniques relevant to many fields. In chemistry, single molecule spectroscopy has been used to probe the diffusion of individual molecules in a variety of environments such as phospholipid membranes,[5] nanoporous networks of sol-gel glasses,[6] on[7] and near[8] patterned surfaces, and at solution-crystal interfaces.[9] Biophysics also quickly adapted single molecule microscopy techniques to such challenging problems as measuring the microrheology in cells[10] and the motion of individual motor proteins.[11–13]

Single molecule techniques that measure the dynamics of individual molecules and particles include both optical and nonoptical methods. Scanning tunneling microscopy (STM) and atomic force microscopy (AFM) are the primary nonoptical tools that have been applied to probe surface diffusion.[1,14–16] STM is particularly useful because it offers topological information of the substrate with atomic scale resolution in addition to the molecules under investigation.[1,17] AFM has the advantage that it can also operate on nonconductive surfaces, but its spatial resolution is limited compared to STM.[1] Optical single molecule methods are predominantly based on fluorescence, which provides little or no direct structural information. However, because of the noninvasive nature of optical spectroscopies, single particle tracking[18–20] and single molecule fluorescence correlation spectroscopy[21,22] have been widely applied to three-dimensional diffusion in liquids.

Fluorescence imaging requires that the single molecules to be studied emit light after photoexcitation. In the absence of intrinsic molecular fluorescence, chemical functionalization with luminescent tags such as dyes and quantum dots has become a standard procedure in order to render weakly or nonfluorescent molecules visible.[7,19,23] However, a significant problem is that dyes and quantum dots intermittently cease to emit.[3,24] This effect is known as photoblinking and makes individual molecules invisible for a fraction of an image or even for several frames. In addition, photochemical reactions cause permanent photobleaching of dyes thereby limiting the total time single molecules can be investigated.[3] While adding oxygen scavengers can minimize photoblinking and extend the lifetimes of fluorescent dyes in solution,[3,25] slower molecular diffusion processes, as encountered on surfaces under ambient conditions, require the acquisition of long-time trajectories in the presence of photoblinking.

To improve the quality of single molecule trajectories, much work focused on the precision and accuracy to which a

a)Author to whom correspondence should be addressed. Electronic mail: slink@rice.edu.

molecule can be localized[26–31] and assigned to a trajectory.[19,20,32] Different microscopy schemes allow both a fast-time series of images to be obtained by wide-field imaging, such as total internal reflection, or permit a long-time series to be measured by limiting illumination using scanning techniques, such as near-field scanning microscopy and confocal microscopy.[3] Despite sufficient time resolution, simultaneously identifying the correct trajectories of multiple molecules for long acquisition times in the presence of blinking and at high concentrations is still a nontrivial task.[19,20] Multiple-hypothesis tracking seeks solutions to this problem by globally optimizing the trajectories for all molecules in space and time.[33] Assuming models for the amount and type of movement can, however drastically reduce computation time[34] and also account for photoblinking.[18] In order to further increase the reliability by which single molecule trajectories can be extracted from a time series of images, recent advances have been made by considering a subset of the full combinatorial problem[19] or by taking a Bayesian or maximum likelihood approach.[35,36]

Presented here is a procedure that is computationally nondemanding and reliably extracts long-time fluorescent single molecule trajectories, even when interrupted by blinking. Rather than considering a full global optimization problem in space and time, we tracked molecules by considering nearby molecules in subsequent frames without assuming a model for molecular motion. This concept was extended, so that the optimal trajectory was identified, even when multiple molecules were present in the frame. The algorithm is independent of the technique used to capture images and may be applied to other microscopy methods that capture two-dimensional single molecule images as a function of time as long as the particle density and frame rate are adjusted to avoid particle movement in excess of the average nearest neighbor distance in between two successive images. We applied the developed procedure to the surface motion of dye-tagged nanocars, which are molecular machines designed to roll on flat surfaces.[37] By using sample scanning confocal microscopy, long-time trajectories have been extracted without the use of oxygen scavengers. Furthermore, comparison to stationary dye molecules and large fluorescent polystyrene beads together with simulations showed that the smallest diffusion constants measured are determined by noise only and, therefore, care has to be taken when interpreting surface motion from single molecule fluorescence images.

II. EXPERIMENTAL

All data were taken using a home-built sample scanning confocal microscope as depicted in Fig. 1(a). A 532 nm frequency doubled diode pumped laser (Coherent, Verdi V6) was used as an excitation source. The sample was placed on a *XYZ* piezoelectric scanning stage (Physik Instrumente, P-517.3CL) mounted on an inverted microscope (Zeiss, Axiovet 200). The position of the stage was controlled by a surface probe microscope controller (RHK Technology, SPM 1000). The excitation light was circularly polarized and focused to a diffraction limited spot of 250–300 nm with a 100× objective (Zeiss, Fluar with numerical aperture of 1.3).

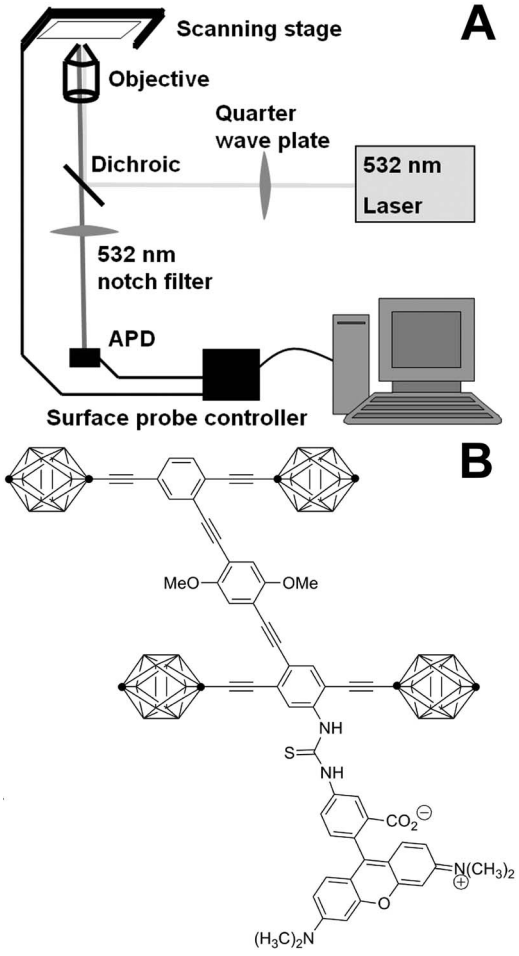

FIG. 1. (a) Schematic of the setup. A 532 nm laser was used for excitation. Fluorescence was collected by the objective and sent to an APD. A dichroic beam splitter and filter removed the 532 nm excitation light. To form an image, the sample was scanned and the fluorescence intensity stored using a surface probe controller with a photon counting board interfaced to a computer. (b) Chemical structure of the TRITC-tagged nanocar.

The fluorescence was captured by the same objective and sent back along the beam path to a dichroic filter (Chroma R532RDC) and then a 532 nm notch filter (CVI), which transmit the fluorescent light and filter out the scattered excitation light, respectively. An avalanche photodiode (APD) (Perkin Elmer SPCM-AQR-15) was used as the detector. The signal from the APD was sent to the photon counter board of the surface probe microscope controller.

Samples of individual molecules were isolated on a coverslip by drop casting or spin casting (3500 rpm for 90 s) on cleaned coverslips from solution with concentrations of 10^{-10}–10^{-12} mol/L. Molecules studied here include tetramethylrhodamine isothiocyanate (TRITC) (Invitrogen) and a nanocar consisting of four carborane "wheels" linked to a stiff carbon backbone "chassis" to which TRITC was attached,[23,38] as shown in Fig. 1(b). The coverslips were cleaned by sonication in acetone for 15 min followed by plasma cleaning at 200 mtorr for 30 s. All measurements were carried out at room temperature and in air. Typically, 10×10 μm^2 images consisting of 128×128 pixels were captured with an average dwell time of 1 ms per pixel. As the stage was moved from pixel to pixel line by line, the total

exposure time for each fluorescent molecule was only a few milliseconds for each image and individual molecules continued to fluoresce for many minutes at an average excitation power of 1 kW/cm^2. For comparison and calibration, samples of fluorescent 100 nm polystyrene beads (Invitrogen, FluoSpheres) with absorption/emission maxima of 540/560 nm and intensity equivalent to 7400 fluorescent molecules were similarly prepared and imaged using an average power of 1 W/cm^2. All data analyses were carried out using MATLAB R2006A.

III. RESULTS AND DISCUSSION

Figure 2 shows four images of single TRITC labeled nanocars selected from a sequence of fluorescence images. These images exhibit the four types of movement of interest: stationary, moving, stationary and blinking, and moving and blinking. When a molecule blinks, it temporarily ceases to fluoresce.[3] Consequently, for one entire frame, or, if the off time is less than the line time of 128 ms, then for one line of the scan, the molecule is not imaged. Of the several moving molecules in Fig. 2, movement can range up to a micrometer. Quantifying this movement by creating single molecule trajectories, despite blinking and lack of an absolute reference location is the main goal of the procedure detailed below.

The first step in creating a single molecule trajectory is identifying the molecules in each frame. This was done by locating those pixels in the image with intensities that were greater than a specified background threshold value. The local maxima of this collection of pixels were then determined and fit to a Gaussian.[26] Here, we determined the centroid of the Gaussian using Eq. (1), while the standard deviation s_x of the Gaussian is given by Eq. (2).

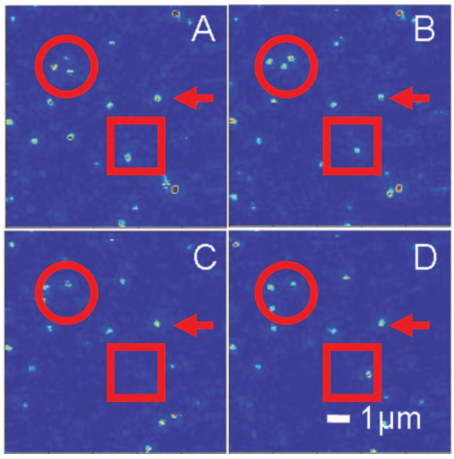

FIG. 2. (Color) (a)–(d) are images of nanocars acquired at times 0, 51, 179, and 204 s, respectively. This series shows the typical behaviors and challenges of molecule identification and tracking. The red arrow points to a stationary, nonblinking molecule. The circle highlights three molecules that move relative to each other. The boxed molecule shows both movement and blinking. Diffusion analysis for the boxed molecule and the molecule to which the arrow is pointing is shown below.

$$x_0 = \left(\sum_{x=x_i-\delta}^{x_i+\delta} \left(\sum_{y=y_i-\delta}^{y_i+\delta} I_{x,y} \right) x \right) \Big/ \left(\sum_{x=x_i-\delta}^{x_i+\delta} \left(\sum_{y=y_i-\delta}^{y_i+\delta} I_{x,y} \right) \right), \tag{1}$$

$$s_x = \sqrt{\left(\sum_{x=x_i-\delta}^{x_i+\delta} \left(\sum_{y=y_i-\delta}^{y_i+\delta} I_{x,y} \right) (x-x_0)^2 \right) \Big/ \left(\sum_{x=x_i-\delta}^{x_i+\delta} \left(\sum_{y=y_i-\delta}^{y_i+\delta} I_{x,y} \right) - 1 \right)}. \tag{2}$$

In both equations, the centroid location along x is given by x_0, the local maxima by x_i and y_i, the intensity at pixel (x,y) by $I_{x,y}$, and δ denotes the size of the subset of the image over which these values are computed. The value δ was chosen such that only 25% of the points a distance δ away from the local maxima were above the threshold. The sum over the other index assures that all values over which the molecule extends are considered. For the centroid location and standard deviation along the y direction, the indices were reversed. For subsequent analysis, the location of the centroid in x and y, i.e., x_0 and y_0, was recorded as the location of the molecule. Molecules that could not be described well by a Gaussian, such as strongly asymmetric intensity distributions, possibly due to two closely spaced molecules, were removed. The brightest molecules with intensities above a second threshold were also deleted, as they were likely to be aggregates of multiple fluorescent molecules. The precision of locating single TRITC molecules was typically about 100 nm for our experimental conditions.

Once the locations of the molecules in the first frame were determined, the same molecules needed to be identified in subsequent frames and the positions assembled into trajectories. The association of a molecule in one frame with one in another was done through a variable parameter defined here as the search radius. In the first frame a circle of this radius r was drawn around each molecule. If the centroid of another molecule was found within this circle in the next frame then that molecule was considered to be the same as in the first frame. Figure 3 illustrates what happens as the search radius increased. At first no associations were made, as r was too small for even small fluctuations in the location of the molecules. As r increased more associations become possible. Finally when r was large enough to cause overlap of the search areas, multiple associations complicated the assignment.

Given the random distribution of intermolecular distances, even small values of the search radius will in general cause inevitable overlap and hence multiple associations. However, additional correct associations can be extracted

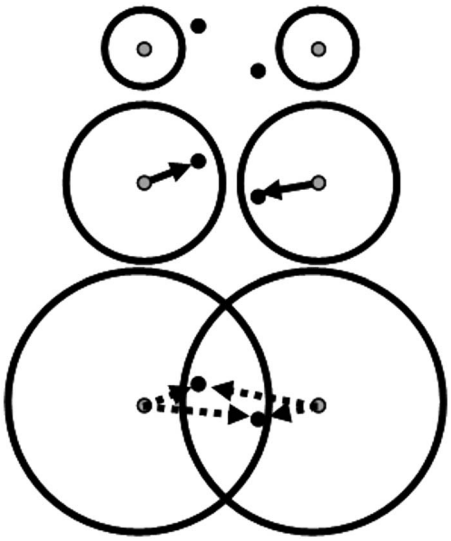

FIG. 3. An illustration of search radius variation. Initially (top) no associations were made; then as the search radius increased (middle) an increasing number of associations became possible. This continued until the search areas started to overlap (bottom), which complicated association.

from these cases as well. Consider Fig. 4, which illustrates the possibilities of two molecules with overlapping search regions. Here we see that if only one molecule entered the overlapping area, both could still be properly identified (case B). If, however, both molecules entered the overlapping area, neither one could be properly identified (case C). In general, if an overlapping area existed, which contained only one molecule, then this molecule was properly identified. However, if the molecule appeared in an overlapping area and had one or more partners, then it could not be properly associated. Once a molecule was identified and associated correctly, it and the partner it was associated with were removed from the list of possibilities for other identified molecules and the overlapping regions were reconsidered to determine if other molecules could be correctly paired from successive images. This process continued until no more molecule pairs could be assigned and the remaining molecules were discarded from the analysis. The only assumption made here was that the movement between frames for a particle pair was less than the nearest neighbor spacing. While no assumptions about the type, magnitude, or distribution of

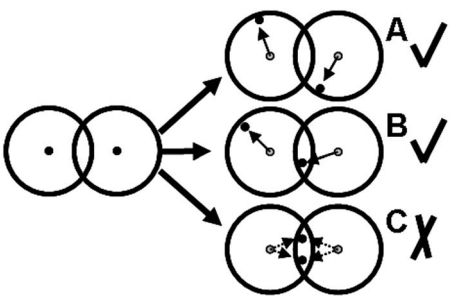

FIG. 4. Two molecules were identified in one frame and a search radius was drawn around them. The molecules could then move into an overlapping region in the next frame. In cases (a) and (b), but not (c), it was possible to uniquely determine which molecules are associated with the previous molecule (shown in gray). These cases were generalized for an arbitrary number of molecules.

single molecule movements were made, this condition is similar to other algorithms[18,34] and can be directly tested within our method, and then corrected for by adjusting the particle density and acquisition speed as shown below.

Blinking further complicated making associations. If the molecule blinks off for only the duration of one pixel or one line during the image acquisition, the resulting effect is to increase the error in locating the molecule according to Eq. (2). This behavior will be treated as another source of noise in later analysis and the term "blinking" will refer only to a molecule ceasing to fluoresce for an entire frame. If a molecule blinked off in a frame, it is subsequently ignored. Likewise if the molecule photobleached and never re-entered a fluorescent state, it too was ignored and the trajectory was terminated at that point. If, however, a molecule blinked back on, then a molecule that could be associated with it was searched for in previous frames. If one was found and it was not already associated with another molecule, then they were associated and the trajectory was continued from the new location of the molecule.

What value of the search radius then gives the correct associations for which a trajectory can be extracted? To find this optimal search radius, the procedure above was repeated for various values of the search radius. In order to correctly associate molecules as shown in Figs. 3 and 4, we note that the search radius must contain values that are a fraction of the interparticle spacing. A metric that was found useful in looking at the trajectories produced was the association efficiency. This was defined as the fraction of molecules associated into trajectories divided by the total number of molecules identified in all frames for a given value of the search radius. For small search radii (see Fig. 3, top), variations in the centroid position due to molecule movement or noise caused few associations to be made. For large search radii there was considerable overlap in the search regions (Fig. 3, bottom) and many pairs could not be associated and were discarded, resulting in a decrease in efficiency to zero. Somewhere between these two extremes, the efficiency reached a peak.

To verify this concept, we measured 100 nm diameter fluorescent polystyrene beads at different surface coverages as shown in Figs. 5(a) and 5(b). Figures 5(c) and 5(d) are the corresponding efficiency plots for this stationary, nonblinking sample. Red circles identified beads that met the criteria for locating single particles and were used in the analysis, including the efficiency plots. Beads circled in yellow were identified, but did not meet the criteria imposed, and were ignored in all of the analysis. The search radius was varied in steps of 0.5 pixels, starting at 0 pixels. The anticipated behavior is indeed observed experimentally, with the efficiency starting at 0 for a search radius of 0 pixels, increasing to a maximum, and then decreasing back to zero. The search radius corresponding to the maximum efficiency is expected to produce the correct trajectories and the trajectories found with this value of search radius were used for later analysis.

The plots of association efficiency versus search radius were also found to provide qualitative information as to the behavior of the system and parameters, which could be optimized. For the polystyrene bead sample in Fig. 5, it can be

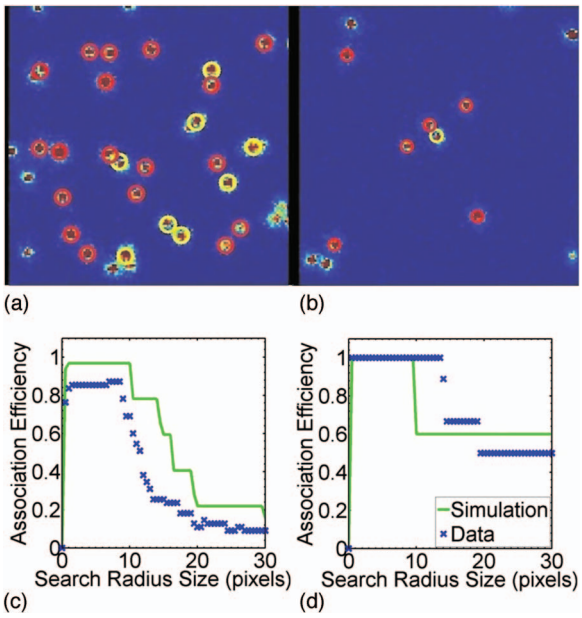

FIG. 5. (Color) Fluorescence images for high (a) and low (b) density polystyrene beads. Beads that were identified but did not meet the cutoff values have a yellow circle around them, while beads that were retained for further analysis have red circles. Beads too close to the edge were excluded from identification. The association efficiency vs search radius plots for the corresponding image sequences are shown in (c) and (d), respectively. Also included are simulations (light, solid lines) for beads with similar intensity, size, and density per image.

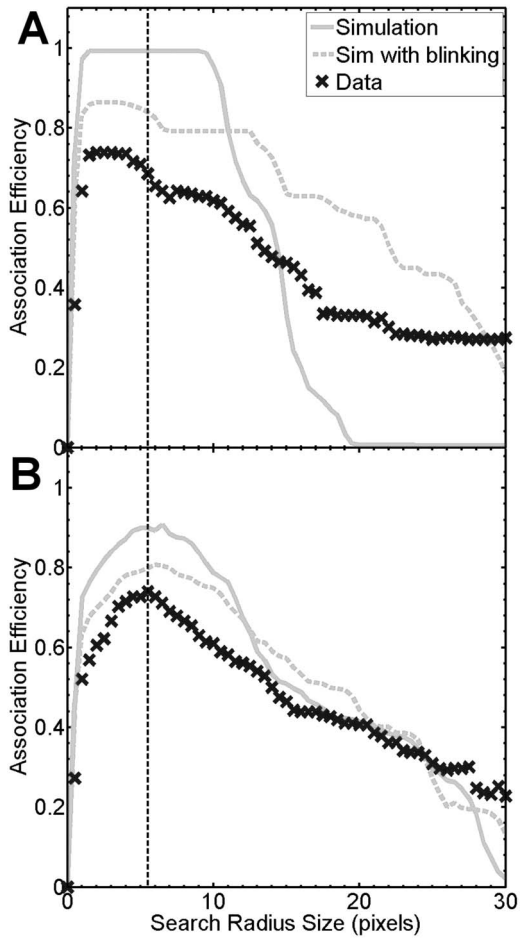

FIG. 6. Association efficiency for a time series of fluorescence images collected for TRITC (a) and TRITC tagged nanocars (b). The peak of the association efficiency for the nanocars was at a search radius of 5.5 pixels (vertical dotted line), which was significantly greater than the peak for TRITC. This effect was observed even though only 3 out of 12 nanocars were moving in the corresponding image. Both samples were of comparable density.

seen that there is a rapid increase to an efficiency of 1. The large number of fluorescent molecules per bead eliminates particle blinking, as it is unlikely that all fluorescent molecules blink off at the same time. The absence of blinking together with a larger brightness and hence increased signal-to-noise (S/N) ratio allow a more precise location of the center of the particle. Consistent with the absence of movement for the large polystyrene beads on a coverslip surface at room temperature, a small search radius captured the next position of each bead. The efficiency starts at zero, however, as shot noise caused the centroid location to fluctuate.

The concentration of beads also had an effect on the association efficiency. As seen in Figs. 5(c) and 5(d), the efficiency for the high concentration bead sample returned to zero much faster than for the low concentration sample, as the radius increased. Because the density of objects in the image was higher, overlap of the search areas happened at a smaller radius and more beads were hence discarded. The difference in concentration was also responsible for the size of the steps in the decay. When an object was discarded for a higher density image it contributed a smaller fraction of the total and the step size of the decay was therefore smaller.

Next, we applied this procedure to single TRITC molecules that were stationary, but blinked. We measured TRITC on glass coverslips exposed to air [Fig. 6(a)] as well as immersed in a polyvinyl alcohol (PVA) matrix (not shown). The sharp rise in Fig. 6(a) was again due to the stationary behavior of TRITC, but it was not as steep as in Fig. 5. This was due to the reduced S/N ratio contributing to additional uncertainty in the centroid location. Furthermore, when the molecule blinked off, the trajectory might not be recovered when it subsequently blinked back on again, reducing the number of molecules included in the trajectories and consequently lowered the association efficiency compared to the beads.

Figure 6(b) shows the results for TRITC tagged nanocars. In a previous study, we were able to establish that the nanocars move under these experimental conditions as much as a micrometer.[23] Because of the movement, a larger search radius was needed to associate the moving molecules in a trajectory and the peak in the association efficiency occurred at a higher search radius value. The peak roughly corresponded to the movement magnitude per frame and was significantly larger than for the stationary TRITC molecules. At the same time, for an unknown behavior, the association efficiency versus search radius plots gave a quick qualitative picture if conditions were met to observe the diffusion of single molecules despite the presence of blinking and the absence of a stationary reference point.

To verify our interpretation, we simulated time sequences of images that contained a similar number of molecules with comparable brightness and size as well as degree of movement and blinking. Fluorescence images with arbi-

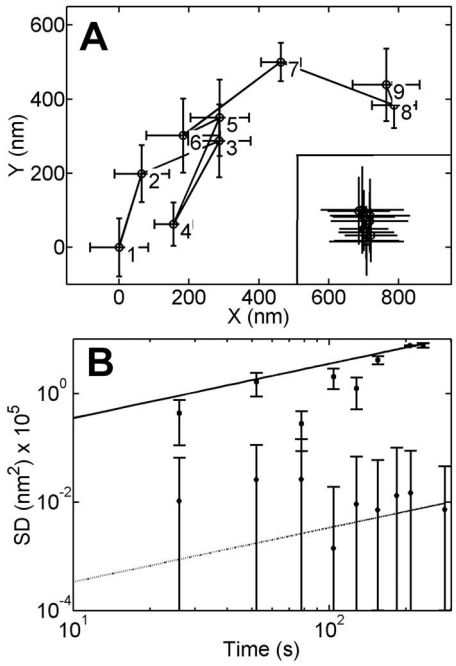

FIG. 7. (a) Trajectory of a moving nanocar from the data in Fig. 2. Inset: A stationary nanocar from the same data set on the same scale. The moving molecule corresponds to the boxed molecule, while the stationary molecule is the molecule to which the arrow is pointing in Fig. 2. The centroid of the stationary molecule stayed within the error bars of the entire trajectory, while the moving molecule clearly moved a greater distance than the error. (b) Squared displacement SD vs time. A linear fit yielded single molecule diffusion constants of $d=875$ nm^2 s^{-1} and 0.841 nm^2 s^{-1} for the moving (solid) and stationary (dashed) molecules, respectively.

trarily placed molecules were created by first randomly choosing the centroid locations of the molecules and then plotting two-dimensional Gaussian functions on the 128 $\times$ 128 image grid. Next, the same analysis was performed on the simulated data and the results are included in Figs. 5 and 6. For best comparison, parameters were chosen to approximately match the data; for example, in all simulations, the number of molecules generated for the simulation was identical to the number of molecules identified in the experimental image frame. Detector shot noise was also added. All major trends with respect to concentration, blinking, and movement were accurately reproduced in the simulated data. TRITC was simulated first without blinking and then blinking was simulated stochastically with each molecule given a probability of 0.10 for completely blinking off during a frame. The blinking accounted for the decrease from optimal association and the slower decay with increasing search radius. However, blinking did not change the location of the association peak [Fig. 6(a)]. For the nanocars, blinking even at a higher probability of blinking off of 0.15 per frame had less of an effect compared to movement by 25% of the molecules per frame [Fig. 6(b)].

Trajectories were constructed from the centroid positions of the molecules for which unique association were made for the search radius corresponding to the maximum efficiency and hence the greatest number of molecules in trajectories. Figure 7(a) shows the trajectories of the boxed molecule and the molecule to which the arrow is pointing in Fig. 2. The error bars denote the error in centroid position and were primarily restricted by the number of photon counts for the molecule and the background counts. The error in centroid position is given by Eq. (3).[26]

$$\mathrm{SE}_i = \frac{(s_i)^2 + a^2/12}{N} + \frac{4\sqrt{\pi}s_i^3 b^2}{aN^2}. \tag{3}$$

Here SE_i is the standard error in locating the centroid with $i=x$ or y, s_i is the standard deviation of the centroid given by Eq. (2), a is the area of the pixel, b is the standard deviation of the background counts near the molecule, and N refers to the intensity counts of the molecule.

To assess under which conditions the algorithm failed to generate the correct trajectories, we analyzed known trajectories from simulated images with varying densities and movement magnitudes. It was found that while the movement magnitude was less than the nearest neighbor distance, the trajectories produced were almost always the correct ones and accounted for ~70% of the simulated particles, while incorrect trajectories accounted for only ~5% of the simulated particles. The remaining percentage of particles was discarded because those particles were in overlapping search areas and could not be associated. Once the movement magnitude exceeded the average nearest neighbor distance, correct trajectories were still made, however, they accounted for less than ~50% of the simulated particles. On the other hand, the number of incorrect trajectories did not increase significantly, as they accounted for ~10% of the simulated particles. This shows that the developed algorithm is aimed at minimizing incorrect associations of particle locations into trajectories at the expense of not accounting for the complete information available from all particles in each image frame.

From the single molecule trajectories, diffusion constants quantifying the degree of molecular motion were extracted by plotting the square of the displacement (SD) from the initial location against time t. This is shown for both trajectories in Fig. 7(b). The error of each SD data point, defined as s_r, was propagated from the error in locating the centroid, and is given by

$$s_r = \sqrt{x^2\mathrm{SE}_x^2 + y^2\mathrm{SE}_y^2}. \tag{4}$$

By fitting SD to a line through the origin, the two-dimensional single molecule diffusion constant d was obtained according to Eq. (5).

$$\mathrm{SD} = 4dt. \tag{5}$$

Because of the relatively long time between successive measurements of the molecular positions (~20 s) compared to the acquisition time (~500 ms), the measured diffusion constants could represent a minimum estimate of the true value. A scenario that was not captured by the current experiments and analysis is the possibility that the molecules may be moving substantially more in between frames. However, this is rather unlikely especially when considering that a broadening of the fluorescence spots in the images of the nanocars was not observed.

Performing the diffusion analysis for all the molecules captured yielded distributions of single molecule diffusion constants. Figure 8(a) shows these distributions for the mea-

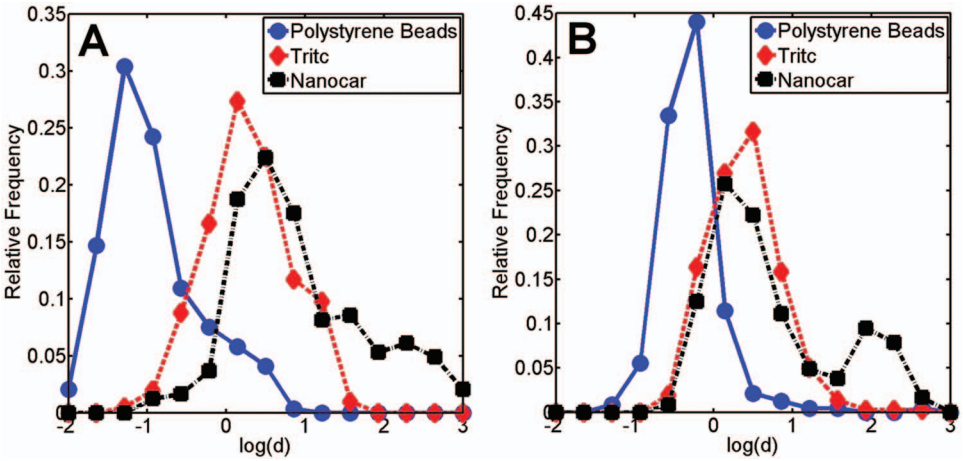

FIG. 8. (Color) Distributions of the $\log_{10}$ of the diffusion constant d. (a) Experimental distribution of diffusion constants for polystyrene beads, the fluorescent dye TRITC, and TRITC labeled nanocars. The bimodal distribution of the nanocars reflects the presence of moving and nonmoving molecules. (b) Simulated data for the three samples (see text for a full discussion of simulation parameters). The high S/N ratio of the beads allowed their positions to be accurately determined and was responsible for the low values of d. The immobile TRITC molecules had higher d values due only to noise, which introduced errors in locating the centroid positions. The simulated nanocars showed a bimodal distribution when assuming that $\frac{1}{4}$ of the nanocars were translating distances larger than the error of the centroid position.

sured beads, TRITC, and nanocars described above. To compare the diffusion constant distributions of all three samples, $\log(d)$ was plotted. Table I lists the average diffusion constants $\langle d\rangle$ for the studied samples. The beads had the lowest $\langle d\rangle = 0.12$ nm^2 s^{-1} due to their immobility and high S/N ratio. The TRITC molecules had the next lowest $\langle d\rangle = 1.89$ nm^2 s^{-1} as they had a lower S/N ratio than the beads. Unlike the beads and TRITC, the nanocars showed a bimodal distribution corresponding to (1) nanocars that translated over the glass surface having a high average diffusion constant of $\langle d\rangle = 220$ nm^2 s^{-1} and (2) nanocars that were stuck at surface dislocations and a small percentage of immobile TRITC tags with $\langle d\rangle = 4.05$ nm^2 s^{-1}. The overlap of the distribution with that of TRITC further confirms our assignment. This highlights the power of single molecule spectroscopy to measure distributions within heterogeneous samples. It should be noted that we verified that the average single molecule diffusion constants $\langle d\rangle$ were consistent with the diffusion constants D calculated by averaging SD instead and obtaining D by a mean squared displacement analysis according to MSD$=4Dt$.[19] However, the latter analysis did not allow us to distinguish between moving and stationary nanocars.

TABLE I. Average single molecule diffusion constants $\langle d\rangle$. The cutoff between the stationary and moving nanocars was determined from Fig. 8 to be $\log(d) = 1.75$. Those with a lower d were included in the average of the stationary nanocars, while those with a higher d were included in the average of the moving nanocars. The value in parenthesis for TRITC corresponds to the simulation run with the lower S/N ratio.

Sample	$\langle d\rangle$ Experimental (nm^2 s^{-1})	$\langle d\rangle$ Simulation (nm^2 s^{-1})
Beads	0.12	0.52
TRITC	1.89	2.32 (7.06)
Nanocars (all)	8.41	5.52
Nanocars (stationary)	4.05	2.59
Nanocars (moving)	220	139

Due to noise in the signal that caused shifts in the centroid location of the molecule, the presented analysis will always result in a measurable diffusion constant even for molecules and particles that actually did not move. In fact, it is very unlikely that the large polystyrene beads diffuse freely on a glass coverslip. Experiments with TRITC embedded in a PVA matrix also showed a similar distribution as the TRITC in air, indicating that the TRITC molecules in air were not diffusing. To confirm these conclusions and to establish the role of noise as well as the smallest meaningful diffusion constant that could be resolved by this method, we calculated the distribution of diffusion constants from the same simulated images discussed above. The results are shown in Fig. 8(b). The simulation parameters included the intensity, noise, size of the molecule in pixels, and magnitude of the movement, and were the same as those used in the simulations in Figs. 5 and 6. For the simulated beads, the signal amplitude range was 50 to 80 counts/ms with random shot noise ($\sqrt{I}$) added for each pixel (S/N ratio=6.9 to 8.8). The TRITC molecules were assigned smaller signal amplitudes of 20 to 30 counts/ms and corresponding shot noise (S/N ratio=4.4 to 5.4). Background noise with a Gaussian distribution was also chosen to match that of the different samples. Consequently, the TRITC samples were generated with background intensities on average ~2.75 times that of the bead samples. The larger S/N ratio for the beads accounted for the difference in mean diffusion constants between these two samples (see Table I). By decreasing the signal amplitude to 15–20 counts/ms (S/N ratio=3.7 to 4.4), only 10% of the TRITC molecules were still identified, and the $\langle d\rangle$ value increased by a factor of 3, well less than the 50-fold increase in the nanocar movement. With a further increase in the noise molecules could no longer be identified reliably. In addition, because drift of the sample on the microscope translation stage may provide a false contribution to the measured diffusion constants, we searched for drift by analyzing the mean position of the stationary nanocars and

found a drift of approximately 11 nm over 10 image frames, which we included in the simulations for all samples. Including this drift had no effect on the simulated mean diffusion constant.

The simulation for the nanocars consisted of molecules with the same intensity and noise as used for TRITC, but $\frac{1}{4}$ of the molecules were allowed to move with interframe step sizes given by a Gaussian distribution with a mean of 117 nm and a standard deviation of 27 nm corresponding to a diffusion constant of 131 $nm^{-2}\ s^{-1}$. These values were determined from the experimental distribution of moving nanocar step sizes. The simulations produced a bimodal distribution consistent with the experimental results. The peak at lower d coincided with the one simulated for TRITC and must be the result of the population of nonmoving nanocars, while the peak at higher d was assigned to the moving nanocars. The means of the two populations differed by almost two orders of magnitude and were well outside the range that could be reached in the simulations by an increase in noise. In contrast with the beads and TRITC molecules, the simulations combined with the experimental data conclusively demonstrated that two distributions of the nanocars are present: 25% are moving while the rest are stationary. This was further in agreement with our previous studies, which showed that the nanocars are translating over a glass surface consistent with a wheel-like rolling mechanism.[23]

IV. CONCLUSIONS

We developed a method of tracking the translational movement of multiple single molecules on a surface by analyzing a time series of fluorescent images despite photoblinking and photobleaching of the fluorophores. The algorithm presented here does not assume the type or magnitude of the movement, but rather determines it from the data. However, as with other trajectory finding methods,[34] the movement between frames must be sufficiently small so that the molecules and particles do not pass each other. This means that the molecular movement must be less than the nearest neighbor distance, which can be controlled by adjusting the particle density and the image acquisition speed. The tracking algorithm is based on a local association procedure, which is computationally significantly less intensive than methods global in both space and time such as multiple-hypothesis tracking.[33] At the same time, the association efficiency versus search radius plots also provides a source of global optimization and a convenient parameter by which the single molecule density can be optimized for our algorithm and a given frame rate. We find the procedure to be most effective in situations that are moderately dense—enough particles are present that a pure local association routine would fail, but where global methods would be inefficient. Furthermore, as our algorithm discards associations it finds to be possibly incorrect, simulations at high sample densities still only give a small percentage of incorrect trajectories. This is because the algorithm is aimed at minimizing wrong trajectories while compromising on the total number of recovered trajectories.

We applied the developed procedure to the important problem of monitoring surface motion of single molecules under ambient conditions. By using confocal microscopy with limited exposure of the fluorescent dyes to the excitation laser, we were able to obtain long-time trajectories of single molecules and analyze them to determine distributions of diffusion constants. For the dye-tagged nanocars studied we observed two distributions of diffusion constants, which we assigned to moving and stationary molecules. This assignment was further supported by simulations, which allowed us to distinguish between diffusion constants based on molecular surface motion and those caused simply by fluctuations of the determined single molecule positions due to noise. We foresee that the combination of the developed tracking algorithm and sample scanning confocal microscopy used here will be particularly useful for measuring slow heterogeneous diffusion processes with multiple distributions, which requires collecting long-time trajectories of single molecules in the presence of photoblinking. However, our tracking algorithm is not dependent on the image acquisition method and may be applied to widefield imaging, or with slight modifications, to three dimensional imaging. Additionally with faster image acquisition speeds, trajectories from higher density environments can also be obtained.

V. FURTHER INFORMATION

An executable program of the developed MATLAB algorithm is available on our webpage at http://www.owlnet.rice.edu/~slink/. The source code can be obtained upon email request to slink@rice.edu.

ACKNOWLEDGMENTS

This work was supported by the Robert A. Welch Foundation (Grants Nos. C-1664 and C-1489), NSF NIRT (Grant No. ECCS-0708765), and the NSF Penn State MRSEC. S.L. acknowledges support through a 3M Nontenured Faculty Grant and the Energy and Environmental Systems Institute (EESI) at Rice University. We also thank Professor Jason Hafner for the use of his plasma cleaner, Professor Kevin Kelly for useful discussions, and RHK Technology for technical support with minimizing the acquisition time for a single image frame.

[1] W. Ho, J. Chem. Phys. **117**, 11033 (2002).
[2] C. Joo, H. Balci, Y. Ishitsuka, C. Buranachai, and T. Ha, Annu. Rev. Biochem. **77**, 51 (2008).
[3] W. E. Moerner and D. P. Fromm, Rev. Sci. Instrum. **74**, 3597 (2003).
[4] X. S. Xie and J. K. Trautman, Annu. Rev. Phys. Chem. **49**, 441 (1998).
[5] T. Schmidt, G. J. Shultz, W. Baumgartner, H. J. Gruber, and H. Schindler, Proc. Natl. Acad. Sci. U.S.A. **93**, 2926 (1996).
[6] C. Hellriegel, J. Kirstein, C. Braeuchle, V. Latour, T. Pigot, R. Olivier, S. Lacombe, R. Brown, V. Guieu, C. Payrastre, A. Izquierdo, and P. Mocho, J. Phys. Chem. B **108**, 14699 (2004).
[7] B. Takimoto, H. Nabika, and K. Murakoshi, Jpn. J. Appl. Phys., Part 1 **45**, 6039 (2006).
[8] E. Mei, A. Sharonov, F. Gaio, J. H. Ferris, and R. M. Hochstrasser, J. Phys. Chem. A **108**, 7339 (2004).
[9] G. Sazaki, M. Okada, T. Matsui, T. Watanabe, H. Higuchi, K. Tsukamoto, and K. Nakajima, Cryst. Growth Des. **8**, 2024 (2008).
[10] J. C. Crocker and B. D. Hoffman, Methods Cell Biol. **83**, 141 (2007).
[11] A. D. Mehta, M. Rief, J. A. Spudich, D. A. Smith, and R. M. Simmons, Science **283**, 1689 (1999).
[12] J. Gelles, B. J. Schnapp, and M. P. Sheetz, Nature (London) **331**, 450

(1988).
[13] A. Yildiz, J. N. Forkey, S. A. McKinney, T. Ha, Y. E. Goldman, and P. R. Selvin, Science **300**, 2061 (2003).
[14] B. S. Swartzentruber, Phys. Rev. Lett. **76**, 459 (1996).
[15] M. Schunack, T. R. Linderoth, F. Rosei, E. Laegsgaard, I. Stensgaard, and F. Besenbacher, Phys. Rev. Lett. **88**, 156102 (2002).
[16] T. T. Tsong, Prog. Surf. Sci. **67**, 235 (2001).
[17] Y. Shirai, A. J. Osgood, Y. M. Zhao, K. F. Kelly, and J. M. Tour, Nano Lett. **5**, 2330 (2005).
[18] S. Bonneau, M. Dahan, and L. D. Cohen, IEEE Trans. Image Process. **14**, 1384 (2005).
[19] K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. Schmid, and G. Danuser, Nat. Methods **5**, 695 (2008).
[20] A. Serge, N. Bertaux, H. Rigneault, and D. Marguet, Nat. Methods **5**, 687 (2008).
[21] J. Enderlein, I. Gregor, D. Patra, and J. Fitter, J. Fluoresc. **15**, 415 (2005).
[22] E. Haustein and P. Schwille, Annu. Rev. Biophys. Biomol. Struct. **36**, 151 (2007).
[23] S. Khatua, J. M. Guerrero, K. Claytor, G. Vives, A. B. Kolomeisky, J. M. Tour, and S. Link, ACS Nano **3**, 351 (2009).
[24] M. Kuno, D. P. Fromm, H. F. Hamann, A. Gallagher, and D. J. Nesbitt, J. Chem. Phys. **115**, 1028 (2001).
[25] C. E. Aitken, R. A. Marshall, and J. D. Pulglisi, Biophys. J. **94**, 1826 (2008).
[26] R. E. Thompson, D. R. Larson, and W. W. Webb, Biophys. J. **82**, 2775 (2002).
[27] X. Qu, D. Wu, L. Mets, and N. F. Scherer, Proc. Natl. Acad. Sci. U.S.A. **101**, 11298 (2004).
[28] A. Yildiz and P. R. Selvin, Acc. Chem. Res. **38**, 574 (2005).
[29] M. J. Rust, M. Bates, and X. W. Zhuang, Nat. Methods **3**, 793 (2006).
[30] J. S. Biteen, M. A. Thompson, N. K. Tselentis, G. R. Bowman, L. Shapiro, and W. E. Moerner, Nat. Methods **5**, 947 (2008).
[31] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, Science **313**, 1642 (2006).
[32] M. K. Cheezum, W. F. Walker, and W. H. Guilford, Biophys. J. **81**, 2378 (2001).
[33] D. B. Reid, IEEE Trans. Autom. Control **24**, 843 (1979).
[34] J. C. Crocker and D. G. Grier, J. Colloid Interface Sci. **179**, 298 (1996).
[35] A. Genovesio, T. Liedl, V. Emiliani, W. J. Parak, M. Coppey-Moisan, and J. C. Olivo-Marin, IEEE Trans. Image Process. **15**, 1062 (2006).
[36] J. W. Yoon, A. Bruckbauer, W. J. Fitzgerald, and D. Klenerman, Biophys. J. **94**, 4932 (2008).
[37] Y. Shirai, J. F. Morin, T. Sasaki, J. M. Guerrero, and J. M. Tour, Chem. Soc. Rev. **35**, 1043 (2006).
[38] J.-F. Morin, T. Sasaki, Y. Shirai, J. M. Guerrero, and J. M. Tour, J. Org. Chem. **72**, 9481 (2007).