USING THE FEATURES OF BROWNIAN MOTION TO CHARACTERIZE THE NUCLEAR PORE COMPLEX, MOLECULAR ROBOTS, AND ANTIMONY-DOPED TIN OXIDE

By

Nicole Anne Michelotti

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2013

Doctoral Committee:
Professor Nils G. Walter, Co-Chair
Professor Duncan Steel, Co-Chair
Professor Robert T. Kennedy
Associate Professor Jennifer Ogilvie

For from the greatness and the beauty of created things their original author, by analogy, is seen.

~Wisdom 13:5

To Cameron, whose constant support throughout graduate school has given me the strength to persevere.

ACKNOWLEDGEMENTS

I would like to start by thanking Nils Walter for welcoming a physicist into his lab. I not only appreciated his mentoring and expertise, but also his prioritizing making the lab a sociable, and therefore an enjoyable, place to work. It takes a special adviser to let you take time off from lab work in order to decorate cakes for thesis defenses or cook turkey for a lab Thanksgiving celebration, and I will always treasure those memories.

Transitioning from working with physicists to working with other natural scientists was a culture shock for me. I learned invaluable lessons concerning interdisciplinary research from this cultural immersion. I went from having very little experimental, let alone wet-lab, experience to gaining expertise in various experimental techniques. It is because of the patience of the members of the Walter lab, past and present, and their eagerness to assist me when I struggled, that overcoming this cultural hurdle was manageable. It has been a pleasure and an honor to work with so many talented scientists. I would especially like to thank Alex Johnson-Buck, with whom I collaborated the most, and who always had insightful feedback for my research.

The research I did was a part of a large collaboration. Among all the members of that collaboration, I would especially like to thank Steven Taylor from Milan Stojanovic's lab at Columbia University, Jeanette Nangreave from Hao Yan's lab at Arizona State University, and Alex Volosin from Don Seo's lab at Arizona State University. This dissertation is a testament to their work as well as mine.

I would like to thank my committee members: Duncan Steel, Bob Kennedy, and Jennifer Ogilvie. I would like to thank Duncan in particular. He was my first adviser in graduate school, and his advice ultimately led me to work in Nils's lab. I have returned to his office multiple times over the years for further advice on life and have always appreciated his wisdom.

I had many mentors before graduate school who helped me immensely during the initial stages of my physics career. In particular, I would like to thank Jean Zamin, who was an exceptional high school physics teacher whose passion was contagious. I would also like to thank Stephen Pordes and Vic Scarpine, who were my advisers during my first summer internship at Fermilab, who introduced me to computer programming, and who gave me the confidence I needed to work independently. The skills they equipped me with were invaluable for graduate research and beyond.

I would like to thank all the friends that I made during graduate school, especially my former apartment-mate, Lauren Van Wassenhove, and former housemates, Jenn McCarthy and Kira Landenberger, who shared so many wonderful memories with me. During graduate school, I was the member of multiple groups that were filled with inspiring people: the UMS Choral Union, the physics book club, and the chem prayer group. I'd like to thank the many friends I made in these different groups and their support throughout my graduate school career.

As the child of two physicists with graduate-level degrees, it is easy for me to take my intellectual upbringing for granted. My parents raised me to think critically with a sensitivity to truth, not only in the sciences, but across the disciplines from religion to art to politics. I am very grateful for their life lessons as well as their unconditional love and support. I am grateful to my entire extended family, especially on my father's side which is filled with doctors and lawyers who were all children of a man who only had a third grade education. My extended family, on

both my mother's side and my father's side, taught me the value of goals, work, and perseverance. I thank my siblings who have been so supportive throughout these past six years, especially my older brother, Paul, who generously gave his time and talent to create the image that led to the spider project being posted on blogs and news sources around the world. Had it not been for his selfless contribution to the spider project, the project would not have received nearly the amount of publicity that it did.

Finally, I'd like to thank my fiancé, Cameron. Cameron was there by my side throughout graduate school. We worked on our homework together, days and nights, during the first year. Two summers later, he attended my practice prelim presentation, after which we sat down at a restaurant where he gave me feedback and a pep talk. We persuaded each other to stay in graduate school after our prelims, at a time when we were both questioning our next steps, and consequently ended up remaining best friends and dating a couple years later, followed by getting engaged. Cameron encouraged me to seriously discern what I would like to do next in my career, which has led me to my new position with Tim McKay in physics education research. I am incredibly thankful for the influence Cameron has had in my life thus far, and look forward to our future years together.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT:
USING THE FEATURES OF BROWNIAN MOTION TO CHARACTERIZE THE NUCLEAR PORE COMPLEX, MOLECULAR ROBOTS, AND ANTIMONY-DOPED TIN OXIDE

By

Nicole Anne Michelotti

Co-Chairs: Nils G. Walter and Duncan Steel

Brownian motion is the apparently random motion of small particles in a solution that results from the bombardment of molecules within the solution. The theoretical understanding of this motion was developed by Einstein in the early 1900s. Since then, features of Brownian motion, such as the fact that Brownian motion can be modeled using a random walk, or the fact that ensemble mean squared displacement (MSD) can be used to determine a diffusion coefficient and type of diffusive behavior, have been utilized to characterize a vast array of systems that are both naturally occurring and synthetic.

In this thesis, I characterize three different types of systems using features of Brownian motion: naturally occurring nuclear pore complexes, synthetic molecular robots that are based on naturally occurring bipedal molecular walkers, and synthetic conductive nanoporous antimony-doped tin oxide (ATO). For the nuclear pore complex, the diffusion of particles through each region of the complex was modeled using a random walk in order to help determine the relative diffusion coefficients of the three regions. For the molecular robots, the movement of the robots

was modeled using a more advanced random-walk simulation that utilizes the Gillespie algorithm; the movement of the robots was evaluated based on the MSDs, and the results were used to characterize the directional bias in the walking mechanism of the robots. For the ATO, fluorescent particles were monitored as they underwent Brownian motion while diffusing through the nanopores; MSDs were used to determine that these particles are embedded in the nanopores and that the diffusion coefficient depended in an unexpected way on the potential applied across the material.

CHAPTER 1:
TECHNIQUES FOR AND EXAMPLES OF CHARACTERIZING MATERIALS USING BROWNIAN MOTION[1]

1.1 The Nature of Brownian Motion

Brownian motion is the apparently random movement of particles within a fluid. Botanist Robert Brown was the first to scientifically study this motion (in pollen grains he observed floating in still water) and bring the question of the source of the movement to the scientific forefront, which is why the behavior is associated with his name[1]. However, he could not explain the source of this movement. Over a half-century later, Einstein proposed a quantitative and qualitative explanation that became our current understanding of Brownian motion[2]. For brevity and clarity, I will derive the equation from a different starting point[3] than the one Einstein used, but the same crucial assumptions are implemented in both derivations.

A particle in a fluid in the absence of an applied external force will still experience two forces: The first is a drag force that is linearly proportional to the particle's velocity, *v*, and the second is an assumed fluctuating force from the heat bath that gives rise to the random movement, F(t). In one dimension, we obtain:

$$m\frac{dv}{dt} = F(t) - bv$$

Multiplying both sides of the equation by the coordinate x, we obtain:

$$mx\frac{dv}{dt} = mx\frac{d\dot{x}}{dt} = m\left[\frac{d}{dt}(x\dot{x}) - \dot{x}^2\right] = -bx\dot{x} + xF(t)$$

[1] Portions were adapted from Michelotti, N., de Silva, C., Johnson-Buck, A.E., Manzo, A.J., Walter, N.G. (2010) A bird's eye view: tracking slow nanometer-scale movements of single molecular nano-assemblies. Methods Enzymol. 475, p. 121-148.

Taking the ensemble average of both sides, we get:

$$m\left\langle\left[\frac{d}{dt}(x\dot{x})-\dot{x}^2\right]\right\rangle = m\frac{d}{dt}\langle x\dot{x}\rangle - m\langle\dot{x}^2\rangle = -b\langle x\dot{x}\rangle + \langle xF(t)\rangle$$

Because the fluctuating force is random and *x* and *F(t)* are independent, <*xF(t)*> goes to zero. For one dimension, the equipartition theorem shows that a particle will have an energy of $\frac{1}{2}m\langle\dot{x}^2\rangle = \frac{1}{2}kT$. Thus, we obtain:

$$m\frac{d}{dt}\langle x\dot{x}\rangle - kT = -b\langle x\dot{x}\rangle$$

We substitute in $\langle x\dot{x}\rangle = \frac{1}{2}\frac{d\langle x^2\rangle}{dt}$:

$$m\frac{d}{dt}\frac{d\langle x^2\rangle}{dt} + b\frac{d\langle x^2\rangle}{dt} - 2kT = 0$$

The solution to this differential equation gives:

$$\frac{d\langle x^2\rangle}{dt} = C_1 e^{-\frac{b}{m}t} + \frac{2kT}{b}$$

Under the initial condition that $x = 0$ at $t = 0$, we obtain $C_1 = -kT/b$. Putting in our value for C_1, we get:

$$\frac{d\langle x^2\rangle}{dt} = \frac{2kT}{b}\left(1 - e^{-\frac{b}{m}t}\right)$$

Taking the integral with respect to time gives:

$$\langle x^2\rangle = \frac{2kT}{b}\left(t + \frac{m}{b}e^{-\frac{b}{m}t} + C_2\right)$$

From our initial condition, $C_2 = \frac{2kT}{b}\frac{m}{b}$. Substituting in for C_2, we obtain our equation for the mean squared displacement:

$$\langle x^2\rangle = \frac{2kT}{b}\left(t + \frac{m}{b}e^{-\frac{b}{m}t} + \frac{m}{b}\right)$$

In the limiting case where $t >> m/b$, we obtain the more familiar equation:

$$\langle x^2 \rangle = \frac{2kT}{b} t$$

A common form of this equation defines the diffusion coefficient as $D = \frac{kT}{b}$, but Stoke's Law gives an explicit value to the drag coefficient, b, which is intuitively related to the size of the particle and the viscosity of the fluid the particle is in. Namely, $b = 6\pi\eta a$ where η is the viscosity of the fluid and a is the radius of the particle. The validity of this equation was experimentally verified by Perrin a few years after Einstein published his work[4].

I go through the exercise of explicitly deriving this equation here to illustrate a few points. First, the derivation and experimental proof of this equation had great historical significance in establishing the atomic theory. The substitution based on the equipartition theorem assumes the existence of atoms, which was still a controversial claim at the time, and Einstein proposed the validity of his derivation to be a test of the atomic theory of matter[2]. The explanation of random movement as the bombardment of atomic particles was thus confirmed, solidifying the atomic nature of matter.

Second, using the derivation, it is easier to discern for a given application when the equation is valid. In particular, we made an assumption that the ensemble average of the time-dependent fluctuating force will vanish. This is only valid at an ensemble level. Throughout this dissertation, we make efforts to ensure that there is a statistically significant number of particles used to characterize a system.

A generalized version of the mean-squared-displacement (MSD) equation is still commonly used to characterize particles that undergo Brownian motion[5-7].

$$\langle x^2 \rangle = \alpha D t^{\gamma}$$

Here, α is twice the number of dimensions to which the movement is confined. γ provides additional information concerning the particle behavior. Subdiffusive behavior, which is

exemplified by confined particles, corresponds to $0 < \gamma < 1$. For random Brownian motion, $\gamma = 1$. Superdiffusive behavior corresponds to the case for $\gamma > 1$. All diffusive processes are determined by the diffusive particles' environments, and are commonplace in nature[8]. There are various modifications that the MSD fitting equation can have depending on the system under study, such as an additive velocity term for particles exposed to a continuous directional drift (directed transport)[7,9].

1.2 Single-Molecule Techniques Used to Observe Brownian Motion

During the last few decades, there has been remarkable growth in the use of fluorescence spectroscopy in biophysical studies. Fluorescence-based tools are now being employed to understand the properties and dynamics of proteins[9-12] and nucleic acids[13-16]. They are implemented in cutting-edge applications in medical and clinical chemistry for high-throughput screening and detection[17,18], as well as in cellular imaging for characterizing the localization and movement of intracellular components[19-22]. Continued advances in fluorescence techniques and instrumentation have fueled applications of fluorescence spectroscopy to more detailed characterization of biomolecules. For example, there has been rapid expansion in the use of nucleotide analogs as fluorescent probes[23] to detect and characterize single nucleic acid molecules in real-time. Fluorescence resonance energy transfer (FRET) has emerged as a particularly powerful tool to probe distances[24-26], conformational changes[27-29], and dynamics[30,31] of macromolecules on the order of 1–10 nm. To measure larger distances of typically 10 nm or more, super-accuracy single-particle fluorescence tracking has proven useful[32-35].

The accuracy of particle tracking techniques is determined in part by the finite resolution of light microscopy. When light passes through a lens with a circular aperture as it does in a single-molecule fluorescence microscope, the focused light emitted from a point-like source

forms a diffraction pattern known as an Airy pattern. The radius of the bright central region, called the Airy disk, can be approximated as $\lambda/(2 \times NA)$, where λ is the wavelength of the light source and NA is the numerical aperture of the lens. Within this radius, according to Rayleigh's criterion, no features may be resolved[36]. When imaging in the visible spectrum with a typical NA of 1.2, the radius of the Airy disk, and therefore the optical resolution limit, is ~ 250 nm. However, so-termed super-resolution (or super-accuracy) methods have been developed in the last few years that overcome this optical resolution barrier and bring the localization accuracy of a single particle down to the low nanometer range[21,22,37-44]. It should be noted that we are using the term "super-resolution" here in its broader sense for all techniques that localize and track over time one or more single-molecule emitters at better than the diffraction limit of accuracy[21], while the more narrow sense of the term refers only to imaging techniques that resolve many closely spaced emitters by observing only few of them at a time over a time series of images[43].

Low-nanometer localization accuracy can allow for tracking slowly moving molecular devices in real-time. Fluorescent single-particle tracking entails using the fluorescence emission from a point-like source to accurately determine the emitting particle's location, typically over a span of time ranging from milliseconds to minutes that depends on the speed of particle motion. One of the first examples of fluorescent particle tracking was accomplished by Barak and Webb in their study of the diffusion of intensely fluorescent (~ 45 fluorophores) low density lipoprotein (LDL)–receptor complexes along human fibroblasts in which they were able to observe the movement of as few as one to three "molecules" in a given region[45]. By using low temperatures, single chromophores were first optically detected in solids by Moerner and Kador[46]. Advancements in single-molecule fluorescence techniques allowed for single-molecule detection in more biologically relevant conditions with ever-improving localization accuracy. For example,

near-field scanning optical microscopy (NSOM) brought the tracking error down to ~ 14 nm in solution at room temperature[47]. The high temporal as well as spatial resolution often required for single-particle tracking was accomplished by Schmidt et al who successfully combined a low localization error of ~30 nm with a high time resolution of 40 ms, enabling them to study the diffusion of single phospholipids in a phospholipid membrane[48].

Experimental fluorescent single-particle tracking called for the development of its analytical counterpart. Using a maximum likelihood estimation analysis, Bobroff developed a quantitative method for analyzing the statistical error in position measurements made with light and particle signals, taking into consideration the measurement signal, noise distribution, and instrument resolution. This was done particularly for a Gaussian signal[49]. Based on Bobroff's least-squares fitting approach, Webb and coworkers derived a simple equation for calculating the standard error of the mean of the position measurements (σ_μ) that depends on the instrumentation parameters and features of the Gaussian fit:

$$\sigma_{\mu_i} = \sqrt{\left(\frac{{s_i}^2}{N} + \frac{a^2/12}{N} + \frac{8\pi {s_i}^4 b^2}{a^2 N^2}\right)}$$

where s_i is the standard deviation of the Gaussian distribution of the *i*th index that indicates either the *x*- or *y*-direction, *N* is the number of photons, *a* is the pixel size, and *b* is the standard deviation of the background. The first term (${s_i}^2/N$) arises from photon noise, the second term represents the effect of the finite pixel size of the camera, and the third term arises from the background signal of the sample[44,50]. Webb and coworkers were able to determine, according to this equation, the position of stationary beads with ~2 nm localization precision[44].

These advances in single-molecule fluorescence imaging and analysis laid a firm foundation for the development of fluorescence imaging with one-nanometer accuracy (FIONA) and related techniques. Developed by Selvin and coworkers, FIONA enabled the localization of

singly fluorophore-labeled myosin V motor proteins along microtubules with typically 3 nm precision (<1.5 nm precision for the brighter molecules) and 0.5 s temporal resolution, using a total internal reflection fluorescence (TIRF) microscope at room temperature. This accomplishment was achieved by maximizing the number of photons collected (to ~ 5000–10,000 photons) while optimizing the camera pixel size (to 86 nm) and minimizing the background noise (to a standard deviation of ~ 11 photons)[35].

We used FIONA, which was previously used to study ATP-fueled biological motor proteins[33,35], to study synthetic nanowalkers, termed spiders, walking along pseudo-one-dimensional tracks (chapter 3)[51], and to study single fluorescent dyes diffusing through nanopores (chapter 4). Both of these exhibit Brownian motion.

1.3 Monte Carlo Simulations: Theoretical Counterpart to Single-Molecule Experiments

Monte Carlo simulations provide a means for modeling heterogeneous behaviors of single molecules. The output is generated through numerical input of random numbers, which is ideal for stochastic systems. All the simulations presented here are based on the random walk model, which is commonly used to describe Brownian motion.

In the basic random walk model, an initial position is determined. After each given time point, the object undergoing the random walk takes a step away from the previous position, in a random direction, with a predetermined step-size. The random walk ends after a specified number of steps have been taken. Variations of this fundamental model are incorporated throughout this thesis, including a variable-step-size random walk (chapter 4), a variable-step-size confined random walk (chapter 2). The random walk for the system presented in chapter 3 is dependent on chemical reactions; therefore, the Gillespie Algorithm[52,53], which numerically

simulates the time evolution of a chemical system, is introduced in the Monte Carlo simulations (chapter 3).

1.4 Utilizing Brownian Motion to Characterize Materials

The bombardment of molecules that leads to Brownian motion may seem to be an undesirable obstacle for objects at the nanometer scale. One can imagine that the reckless motion might hinder a natural nanowalker such as kinesin from maintaining its footing along its microtubule track, or prevent molecular motors like RNA polymerase from being able to successfully stabilize on DNA and fulfill their function. However, it is suspected that Brownian motion is actually exploited by the cell to help enhance performance of intracellular motors. It is proposed that the speed of the kinesin protein is increased by the jostling[54], and constant movement of proteins around or along the DNA helps them to be exposed to many regions, which allows them to find their specific target sequences[55,56].

Just as nature capitalizes on this inherent property of nanoscale systems, so too do experimenters take advantage of this nanoscale feature to study properties of nanomaterials that are otherwise hidden. This thesis presents three ways in which the features of Brownian motion aid scientific understanding across various systems.

1.4.1 Characterization of Intracellular Mechanisms: Nuclear Pore Complex

It is well-established that all known living organisms are composed of cells. And yet, many details of the microscopic world of the cell remain a mystery. By fluorescently labeling specific components inside the cell and using single-particle tracking techniques to study the behavior of their Brownian motion, great insights into their working mechanisms can be

achieved. This technique has been used to examine nuclear trafficking[57], transcripts[58], and microRNA assembly[59].

In chapter 2, we demonstrate the use of Brownian motion to improve our understanding of the kinetics involved in the pathway that the messenger RNA:protein complex uses to transport through the nuclear pore complex (NPC)[60]. Previous studies proposed that the diffusion coefficient of particles was increased during transport through the center of the complex compared with the nucleoplasmic or cytoplasmic sides of the NPC[61]. Given the spaghetti-like filament composition of the central region (Figure 1.1), this conflicts with the conjecture that the filaments provide an obstacle for diffusing species. However, previous results lacked the high-speed resolution that SPEED microscopy provides and did not provide Monte Carlo simulations to help interpret the data. We found that experimental evidence and simulations suggest that the diffusion coefficient through the central region is significantly decreased when compared with the periphery.

1.4.2 Characterization of Nanowalkers Inspired by Nature: Spiders

Cellular nanowalkers, including kinesin and myosin, are capable of accomplishing directed steps along their one-dimensional microtubule tracks. In an attempt to better understand how this processive behavior is generated from stochastic binding events, experiments have sought to create their own nanowalker that would "walk" with biased Brownian motion. These DNA-based nanowalkers, comprising two single-stranded DNA (ssDNA) "legs," traverse tracks composed of ssDNA strands complementary to the legs. They move in an experimentally controlled direction using thermodynamically favored strand displacement (or exchange) by an ssDNA fuel strand as an energy source (Figure 2A).

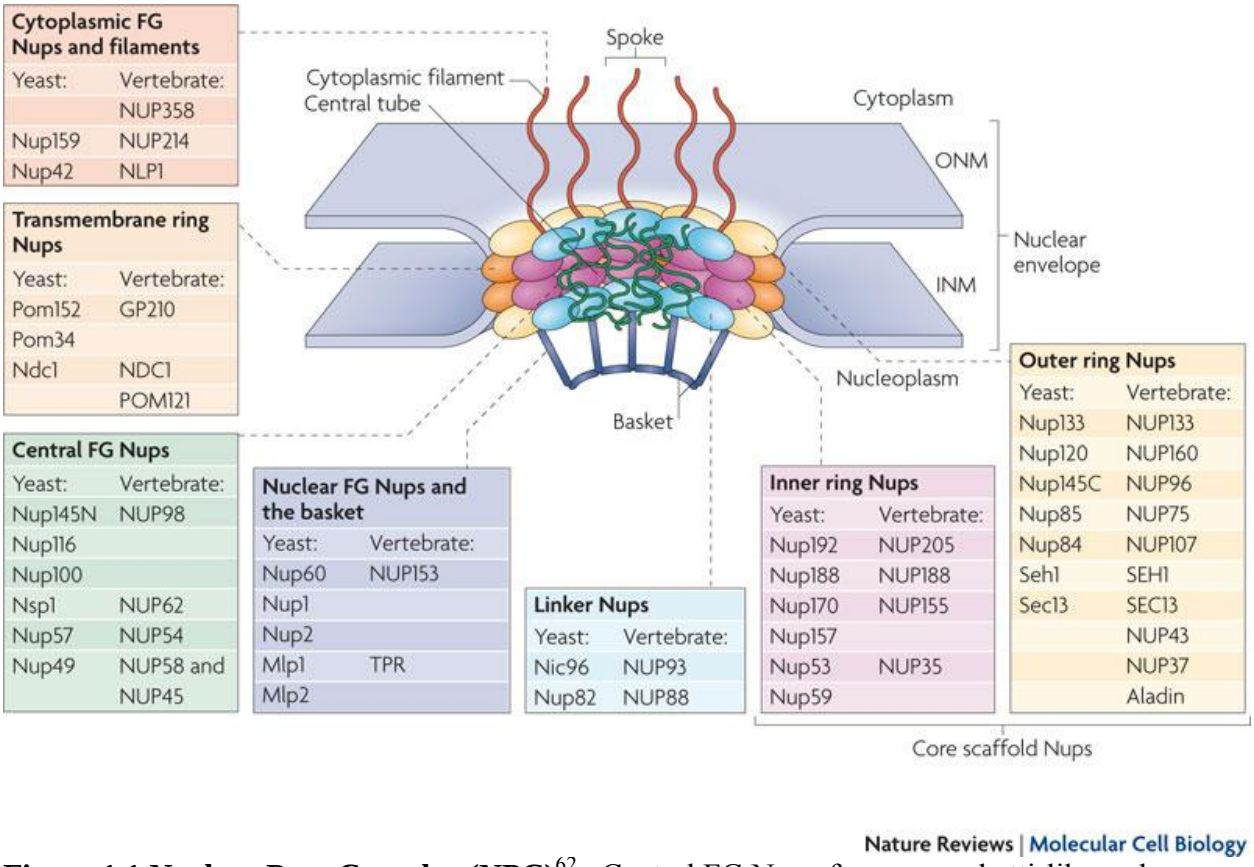


Figure 1.1 Nuclear Pore Complex (NPC)[62]. Central FG Nups form a spaghetti-like web, making it difficult for large particles to freely diffuse through the NPC. Reprinted by permission from Macmillan Publishers Ltd: *Nat. Rev. Mol. Cell Biol.* (Ref. 62), copyright 2010.

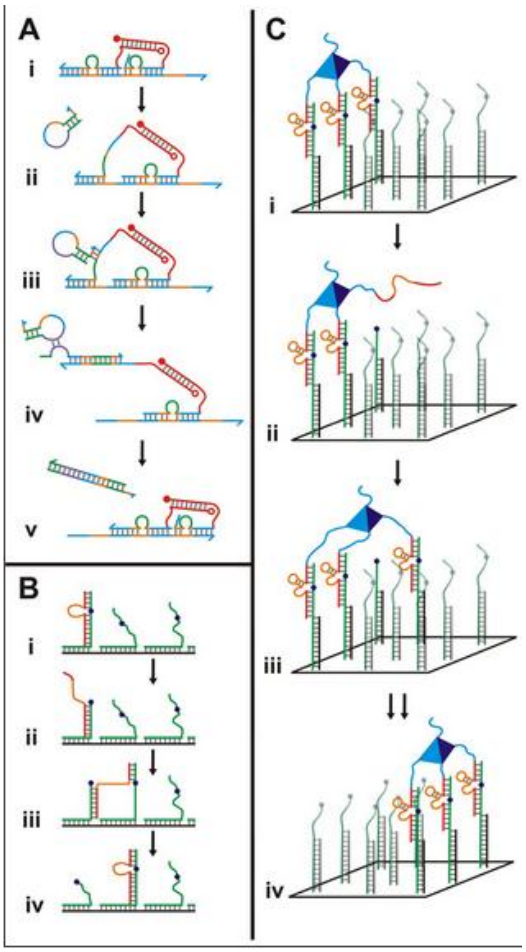


Figure 1.2 Example mechanisms for processive movement of DNA-based nanowalkers. (**A**) Biped nanowalker from Green et al.[63] that utilizes fuel consisting of two complementary DNA hairpins. Colors represent complementary sequences. (i) The competition between identical feet to bind to the track permits the exposure of a toehold region in the left foot (ii). A hairpin hybridizes to the toehold region (iii), displacing the left foot from the track. A second, complementary hairpin hybridizes to the first hairpin (iv) to form a waste product, allowing the foot to rebind to the track with equal probability to the left or right. (**B**) Single-stranded deoxyribozyme-based nanowalker from Tian et al.[64]. (i) The 10–23 deoxyribozyme (red with orange active site) is able to cleave its substrate (green) at a specific site (purple) in the presence of Mg^{2+}. The shorter end dissociates from its product (ii) and hybridizes to a neighboring strand (iii). Via displacement of the cleavage product by neighboring substrates, the deoxyribozyme progresses along the track (iv). (**C**) Spider moving along a three-substrate-wide origami track Lund et al.[51]. The spider is composed of a streptavidin body, a capture leg, spacers, all shown in blue, and three 8–17 deoxyribozyme legs (red binding arms with orange active site). (i) The deoxyribozyme spider legs hybridize to substrates (green) that are attached to the origami scaffold *via* hybridization to staple overhangs (black). In the presence of Zn^{2+}, each leg of the spider cleaves its substrate, dissociates from its products (ii) and hybridizes to a neighboring strand (iii). The greater affinity of the leg for the substrate than the cleavage products makes it energetically favorable for the strand to bind to the full substrate, generating a biased-random walk from the cleaved towards the uncleaved substrate (iv). Reprinted from Ref. 83, copyright 2010.

Due to slow kinetics, strand displacement is not an ideal source of energy for molecular walkers. Biological motor proteins have velocities on the order of ~ 1 μm/s *in vitro* under saturating ATP conditions[65,66], while these synthetic nanowalkers are limited by the kinetics of unwinding one DNA duplex and forming another, both ~15–50 base pairs in length[63,67-71], leading to estimated velocities of ~ 10 nm/h[70]. In addition, while it has been predicted that they have the ability to traverse longer tracks, as yet they have only been shown to accomplish a few successive steps along short tracks approximated to be typically on the lower end of tens of nanometers[63,68].

The catalytic power of deoxyribozymes, or DNAzymes, offers an attractive alternative to strand displacement for driving locomotion. DNAzymes are DNA sequences with the ability to site-specifically cleave chimeric DNA–RNA substrates in the presence of an appropriate divalent metal ion cofactor. Tian *et al.*[72] were the first to incorporate DNAzymes into their nanowalker in the form of the 10–23 DNAzyme[73]. As it cleaves an oligonucleotide on its track, the DNAzyme dissociates from the shorter cleavage product and is able to progress along the track via displacement of the still bound longer product portion by an adjacent substrate strand (Figure 2B)[64]. The speed of movement may still be limited by the kinetics of strand displacement despite the shorter cleavage product, and the processivity limited by the risk of complete dissociation of the single leg from its track.

To overcome these limitations, Stojanovic and coworkers recently developed a polypedal DNAzyme-based nanowalker dubbed a "spider"[74]. Spiders consist of a streptavidin "body" bound to multiple biotinylated 8–17-based DNAzyme[75] "legs." The spider's multivalent binding allows it to remain securely bound to the surface even as individual DNAzymes cleave their substrate and detach, and therefore allows a large number of substrate sites to be visited and

cleaved by a single spider before it dissociates from the surface[74]. Once a leg cleaves its bound substrate, it can more rapidly dissociate from the 10-nucleotide-long products and bind another 18-nucleotide-long substrate in the vicinity. Mathematical modeling of this system suggests that these properties will result in spiders undergoing biased movement on a substrate-field, avoiding sites they have previously visited (Figure 2C)[76,77].

In chapter 3, we employ single-particle tracking techniques to obtain real-time movement information on fluorescently-labeled spiders as they walk in a biased motion along pseudo-one-dimensional DNA origami tracks[51]. We then further optimize the walking mechanism using insights from Monte Carlo simulations that take experimentally determined rate constants as input.

1.4.3. Characterization of Synthetic Material: Antimony-Doped Tin Oxide

Synthetic materials play an important role in our everyday lives, especially our technologies. There is a constant need for materials that meet the demands of our ever-more-sophisticated technologies. Porous materials, such as metal-organic frameworks or mesoporous silica thin films, are of particular interest since pores increase the specific surface area of a material, which is often cost-effective since it enables greater functionality with a smaller volume[78,79]. While microscopic techniques such as SEM may be used to observe the surface of these materials, the inner region within the pores is more difficult to probe. The behavior of the Brownian motion of single fluorophores confined by the pores gives rise to understanding the inner structure of these materials, including the size of the pores and connectivity between the pores[80-82].

In chapter 4, we use this technique to give insight into the structure and electrical properties of Antimony-doped Tin Oxide (ATO). ATO is a transparent nanoporous material that

is conductive, enabling it to be used as an electrode with a high specific surface area, while also enabling fluorescent imaging. Single fluorophores were tracked as they diffused within the pores. The pores demonstrate high connectivity, as shown by the confined two-dimensional Brownian motion of diffusing fluorophores.

1.5 References

1 Brown, R. A Brief Account of Microscopical Observations Made in the Months of June, July, and August, 1827 on the Particles Contained in the Pollen of Plants and on the General Existence of Active Molecules in Organic and Inorganic Bodies. *unpublished* (1827).

2 Einstein, A. On the Motion of Small Particles Suspended in a Stationary Liquid, as Required by the Molecular Kinetic Theory of Heat. *Annalen der Physik* **17**, 549-560 (1905).

3 Reif, F. *Fundamentals of statistical and thermal physics*. (McGraw-Hill, 1965).

4 Perrin, J. Brownian Motion and Molecular Reality. *Annales de Chime et de Physique* **18**, 1-114 (1909).

5 Kusumi, A., Sako, Y. & Yamamoto, M. Confined Lateral Diffusion of Membrane-Receptors as Studied by Single-Particle Tracking (Nanovid Microscopy) - Effects of Calcium-Induced Differentiation in Cultured Epithelial-Cells. *Biophysical Journal* **65**, 2021-2040 (1993).

6 Qian, H., et al. Single Particle Tracking: Analysis of diffusion and flow in two-dimensional systems. *Biophys J* **60**, 910-921 (1991).

7 Saxton, M. J. & Jacobson, K. Single-particle tracking: Applications to membrane dynamics. *Annual Review of Biophysics and Biomolecular Structure* **26**, 373-399 (1997).

8 Klafter, J. & Sokolov, I. M. Anomalous diffusion spreads its wings. *Phys World* **18**, 29-32 (2005).

9 Giepmans, B. N. G., Adams, S. R., Ellisman, M. H. & Tsien, R. Y. Review - The fluorescent toolbox for assessing protein location and function. *Science* **312**, 217-224 (2006).

10 Schuler, B. & Eaton, W. A. Protein folding studied by single-molecule FRET. *Curr Opin Struct Biol* **18**, 16-26, doi:10.1016/j.sbi.2007.12.003 (2008).

11 Min, W. *et al.* Fluctuating enzymes: lessons from single-molecule studies. *Acc Chem Res* **38**, 923-931, doi:10.1021/ar040133f (2005).

12 Shi, J., Dertouzos, J., Gafni, A. & Steel, D. Application of single-molecule spectroscopy in studying enzyme kinetics and mechanism. *Methods in enzymology* **450**, 129-157, doi:10.1016/S0076-6879(08)03407-1 (2008).

13 Ditzler, M. A., Aleman, E. A., Rueda, D. & Walter, N. G. Focus on function: single molecule RNA enzymology. *Biopolymers* **87**, 302-316, doi:10.1002/bip.20819 (2007).

14 Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C. & Ha, T. Advances in single-molecule fluorescence methods for molecular biology. *Annual review of biochemistry* **77**, 51-76, doi:10.1146/annurev.biochem.77.070606.101543 (2008).

15 Pljevaljcic, G. & Millar, D. P. Single-molecule fluorescence methods for the analysis of RNA folding and ribonucleoprotein assembly. *Methods in enzymology* **450**, 233-252, doi:10.1016/S0076-6879(08)03411-3 (2008).

16 Zhao, R. & Rueda, D. RNA folding dynamics by single-molecule fluorescence resonance energy transfer. *Methods* **49**, 112-117, doi:10.1016/j.ymeth.2009.04.017 (2009).

17 Gribbon, P. *et al.* Experiences in implementing uHTS--cutting edge technology meets the real world. *Current drug discovery technologies* **1**, 27-35 (2004).

18 Hintersteiner, M. & Auer, M. Single-bead, single-molecule, single-cell fluorescence: technologies for drug screening and target validation. *Annals of the New York Academy of Sciences* **1130**, 1-11, doi:10.1196/annals.1430.055 (2008).

19 Lippincott-Schwartz, J., Altan-Bonnet, N. & Patterson, G. H. Photobleaching and photoactivation: following protein dynamics in living cells. *Nat Rev Mol Cell Bio*, S7-S14 (2003).
20 Michalet, X. *et al.* Quantum dots for live cells, in vivo imaging, and diagnostics. *Science* **307**, 538-544 (2005).
21 Moerner, W. E. & Kador, L. Optical-Detection and Spectroscopy of Single Molecules in a Solid. *Physical Review Letters* **62**, 2535-2538 (1989).
22 Walter, N. G., Huang, C. Y., Manzo, A. J. & Sobhy, M. A. Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nat Methods* **5**, 475-489, doi:10.1038/nmeth.1215 (2008).
23 Rist, M. J. & Marino, J. P. Fluorescent nucleotide base analogs as probes of nucleic acid structure, dynamics and interactions. *Current Organic Chemistry* **6**, 775-793 (2002).
24 Deniz, A. A. *et al.* Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Forster distance dependence and subpopulations. *P Natl Acad Sci USA* **96**, 3670-3675 (1999).
25 Stryer, L. & Haugland, R. P. Energy transfer: a spectroscopic ruler. *Proc Natl Acad Sci U S A* **58**, 719-726 (1967).
26 Ha, T. *et al.* Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc Natl Acad Sci U S A* **93**, 6264-6268 (1996).
27 Kim, H. D. *et al.* Mg2+-dependent conformational change of RNA studied by fluorescence correlation and FRET on immobilized single molecules. *P Natl Acad Sci USA* **99**, 4284-4289 (2002).
28 Truong, K. & Ikura, M. The use of FRET imaging microscopy to detect protein-protein interactions and protein conformational changes in vivo. *Curr Opin Struc Biol* **11**, 573-578 (2001).
29 Weiss, S. Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nat Struct Biol* **7**, 724-729 (2000).
30 Al-Hashimi, H. M. & Walter, N. G. RNA dynamics: it is about time. *Curr Opin Struc Biol* **18**, 321-329, doi:Doi 10.1016/J.Sbi.2008.04.004 (2008).
31 Zhuang, X. W. *et al.* Correlating structural dynamics and function in single ribozyme molecules. *Science* **296**, 1473-1476 (2002).
32 Barak, L. S. & Webb, W. W. Diffusion of Low-Density Lipoprotein-Receptor Complex on Human-Fibroblasts. *J Cell Biol* **95**, 846-852 (1982).
33 Churchman, L. S., Okten, Z., Rock, R. S., Dawson, J. F. & Spudich, J. A. Single molecule high-resolution colocalization of Cy3 and Cy5 attached to macromolecules measures intramolecular distances through time. *Proc Natl Acad Sci U S A* **102**, 1419-1423 (2005).
34 Gordon, M. P., Ha, T. & Selvin, P. R. Single-molecule high-resolution imaging with photobleaching. *P Natl Acad Sci USA* **101**, 6462-6465 (2004).
35 Yildiz, A. *et al.* Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science* **300**, 2061-2065 (2003).
36 Hecht, E. *Optics*. 4th edn, (Addison-Wesley, 2002).
37 Betzig, E. & Chichester, R. J. Single Molecules Observed by near-Field Scanning Optical Microscopy. *Science* **262**, 1422-1425 (1993).

38 Heintzmann, R., Jovin, T. M. & Cremer, C. Saturated patterned excitation microscopy - a concept for optical resolution improvement. *Journal of the Optical Society of America a-Optics Image Science and Vision* **19**, 1599-1609 (2002).
39 Hess, S. T., Girirajan, T. P. K. & Mason, M. D. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal* **91**, 4258-4272 (2006).
40 Hofmann, M., Eggeling, C., Jakobs, S. & Hell, S. W. Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins. *P Natl Acad Sci USA* **102**, 17565-17569 (2005).
41 Qu, X. H., Wu, D., Mets, L. & Scherer, N. F. Nanometer-localized multiple single-molecule fluorescence microscopy. *P Natl Acad Sci USA* **101**, 11298-11303 (2004).
42 Rust, M. J., Bates, M. & Zhuang, X. W. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods* **3**, 793-795 (2006).
43 Huang, B., Bates, M. & Zhuang, X. Super-resolution fluorescence microscopy. *Annual review of biochemistry* **78**, 993-1016, doi:10.1146/annurev.biochem.77.061906.092014 (2009).
44 Thompson, R. E., Larson, D. R. & Webb, W. W. Precise nanometer localization analysis for individual fluorescent probes. *Biophys J* **82**, 2775-2783, doi:10.1016/S0006-3495(02)75618-X (2002).
45 Barak, L. S. & Webb, W. W. Diffusion of low density lipoprotein-receptor complex on human fibroblasts. *J Cell Biol* **95**, 846-852 (1982).
46 Moerner, W. E. & Kador, L. Optical detection and spectroscopy of single molecules in a solid. *Phys Rev Lett* **62**, 2535-2538 (1989).
47 Betzig, E. & Chichester, R. J. Single molecules observed by near-field scanning optical microscopy. *Science* **262**, 1422-1425, doi:10.1126/science.262.5138.1422 (1993).
48 Schmidt, T., Schutz, G. J., Baumgartner, W., Gruber, H. J. & Schindler, H. Imaging of single molecule diffusion. *Proc Natl Acad Sci U S A* **93**, 2926-2929 (1996).
49 Bobroff, N. Position measurement with a resolution and noise-limited instrument. *Rev. Sci. Instrum* **57**, 1152-1157 (1986).
50 Yildiz, A., Tomishige, M., Vale, R. D. & Selvin, P. R. Kinesin walks hand-over-hand. *Science* **303**, 676-678, doi:10.1126/science.1093753 (2004).
51 Lund, K. *et al.* Molecular robots guided by prescriptive landscapes. *Nature* **465**, 206-210, doi:Doi 10.1038/Nature09012 (2010).
52 Gillespie, D. T. A General Method for Numerically Simulating the Stochastic Tim Evolution of Coupled Chemical Reactions. *Journal of Computational Physics* **22**, 403-434 (1976).
53 Gillespie, D. T. Exact Stochastic Simulation of Coupled Chemical Reactions. *Journal of Physical Chemistry* **81**, 2340-2361 (1977).
54 Amos, L. A. Molecular motors: not quite like clockwork. *Cell Mol. Life Sci.* **65**, 509-515 (2008).
55 Halford, S. E. & Marko, J. F. How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res* **32**, 3040-3052, doi:10.1093/nar/gkh624 (2004).
56 Kolomeisky, A. B. Physics of protein-DNA interactions: mechanisms of facilitated target search. *Physical chemistry chemical physics : PCCP* **13**, 2088-2095, doi:10.1039/c0cp01966f (2011).

57 Babcock, H. P., Chen, C. & Zhuang, X. Using single-particle tracking to study nuclear trafficking of viral genes. *Biophys J* **87**, 2749-2758, doi:10.1529/biophysj.104.042234 (2004).
58 Itzkovitz, S. & van Oudenaarden, A. Validating transcripts with probes and imaging technology. *Nat Methods* **8**, S12-19, doi:10.1038/nmeth.1573 (2011).
59 Pitchiaya, S., Androsavich, J. R. & Walter, N. G. Intracellular single molecule microscopy reveals two kinetically distinct pathways for microRNA assembly. *EMBO reports* **13**, 709-715, doi:10.1038/embor.2012.85 (2012).
60 Ma, J. Three-dimensional mapping with ultrahigh spatiotemporal accuracy unveils the selectivity mechanism of mRNA export through the nuclear pore *Nature Communications*, in review (2013).
61 Grunwald, D. & Singer, R. H. In vivo imaging of labelled endogenous beta-actin mRNA during nucleocytoplasmic transport. *Nature* **467**, 604-607, doi:10.1038/nature09438 (2010).
62 Strambio-De-Castillia, C., Niepel, M. & Rout, M. P. The nuclear pore complex: bridging nuclear transport and gene regulation. *Nature reviews. Molecular cell biology* **11**, 490-501, doi:10.1038/nrm2928 (2010).
63 Green, S. J., Bath, J. & Turberfield, A. J. Coordinated chemomechanical cycles: a mechanism for autonomous molecular motion. *Phys Rev Lett* **101**, 238101 (2008).
64 Tian, Y., He, Y., Chen, Y., Yin, P. & Mao, C. D. Molecular devices - A DNAzyme that walks processively and autonomously along a one-dimensional track. *Angewandte Chemie-International Edition* **44**, 4355-4358 (2005).
65 King, S. J. & Schroer, T. A. Dynactin increases the processivity of the cytoplasmic dynein motor. *Nature cell biology* **2**, 20-24, doi:10.1038/71338 (2000).
66 Kural, C., Balci, H. & Selvin, P. R. Molecular motors one at a time: FIONA to the rescue. *Journal of physics. Condensed matter : an Institute of Physics journal* **17**, S3979-3995, doi:10.1088/0953-8984/17/47/023 (2005).
67 Bath, J., Green, S. J., Allen, K. E. & Turberfield, A. J. Mechanism for a directional, processive, and reversible DNA motor. *Small* **5**, 1513-1516, doi:10.1002/smll.200900078 (2009).
68 Omabegho, T., Sha, R. & Seeman, N. C. A bipedal DNA Brownian motor with coordinated legs. *Science* **324**, 67-71, doi:10.1126/science.1170336 (2009).
69 Sherman, W. B., Seeman, N.C. A precisely controlled DNA biped walking device. *Nano Lett* **4**, 1203-1207 (2004).
70 Shin, J. S. & Pierce, N. A. A synthetic DNA walker for molecular transport. *Journal of the American Chemical Society* **126**, 10834-10835, doi:10.1021/ja047543j (2004).
71 Yin, P., Yan, H., Daniell, X. G., Turberfield, A. J. & Reif, J. H. A unidirectional DNA walker that moves autonomously along a track. *Angewandte Chemie* **43**, 4906-4911, doi:10.1002/anie.200460522 (2004).
72 Tian, Y., He, Y., Chen, Y., Yin, P., Mao, C.D. Molecular devices - A DNAzyme that walks processively and autonomously along a one-dimensional track. *Angew. Chem. In. Ed* **44**, 4355-4358 (2005).
73 Santoro, S. W. & Joyce, G. F. A general purpose RNA-cleaving DNA enzyme. *Proc Natl Acad Sci U S A* **94**, 4262-4266 (1997).

74 Pei, R. *et al.* Behavior of polycatalytic assemblies in a substrate-displaying matrix. *Journal of the American Chemical Society* **128**, 12693-12699, doi:10.1021/ja058394n (2006).
75 Li, J., Zheng, W., Kwon, A. H. & Lu, Y. In vitro selection and characterization of a highly efficient Zn(II)-dependent RNA-cleaving deoxyribozyme. *Nucleic Acids Res* **28**, 481-488 (2000).
76 Antal, T. & Krapivsky, P. L. Molecular spiders with memory. *Physical Review E* **76**, - (2007).
77 Samii, L., Linke, H., Zuckermann, M. J. & Forde, N. R. Biased motion and molecular motor properties of bipedal spiders. *Physical review. E, Statistical, nonlinear, and soft matter physics* **81**, 021106 (2010).
78 Sharma, S., Volosin, A.M., Schmitt, D., Seo, D. Preparation and Electrochemical Properties of Nanoporous Transparent Antimony-Doped Tin Oxide (ATO) Coatings. *J. Mater. Chem. A.* **1**, 699-706 (2013).
79 Volosin, A. M., Sharma, S., Traverse, C., Newman, N., Seo, D. One-pot synthesis of highly mesoporous antimony-doped tin oxide from interpenetrating inorganic/organic networks. *J. Mater. Chem.* **21**, 13232-13240 (2011).
80 Liao, Y., Yang, S. K., Koh, K., Matzger, A. J. & Biteen, J. S. Heterogeneous single-molecule diffusion in one-, two-, and three-dimensional microporous coordination polymers: directional, trapped, and immobile guests. *Nano Lett* **12**, 3080-3085, doi:10.1021/nl300971t (2012).
81 Kirstein, J. *et al.* Exploration of nanostructured channel systems with single-molecule probes. *Nature materials* **6**, 303-310, doi:10.1038/nmat1861 (2007).
82 Zurner, A., Kirstein, J., Doblinger, M., Brauchle, C. & Bein, T. Visualizing single-molecule diffusion in mesoporous materials. *Nature* **450**, 705-708, doi:10.1038/nature06398 (2007).
83 Michelotti, N., de Silva, C., Johnson-Buck, A. E., Manzo, A. J. & Walter, N. G. A Bird's Eye View: Tracking Slow Nanometer-Scale Movements of Single Molecular Nano-Assemblies. *Method Enzymol* **475**, 121-148, doi:Doi 10.1016/S0076-6879(10)75006-0 (2010).

CHAPTER 2:
MONTE CARLO SIMULATIONS AID IN INTERPRETING THE DIFFUSIVE BEHAVIOR OF PARTICLES TRANSPORTED THROUGH THE NUCLEAR PORE COMPLEX[2,3]

2.1 Introduction

In eukaryotic cells, the flow of genetic information from DNA to mRNA to protein encounters a physical impediment, as the mRNA has to pass through the nuclear envelope (NE) membrane. The membrane-bound nuclear pore complex (NPC) acts as a gateway that mediates selective nucleocytoplasmic transport of individual mRNA molecules. Arguably the largest nanomachine of the cell, each NPC is ~200 nm in length and ~50 nm in inner diameter, and is known to consist of 8, 16 or 32 copies of ~30 different nucleoporins (Nups)[6]. Over one third of all Nups, natively unfolded and rich in phenylalanine-glycine (FG) repeats, form the selective gateway for nucleocytoplasmic transport[10-12]. Nuclear mRNA export through this gateway includes the following: assembling with protein cofactors to form an mRNP, trafficking through the NPC, and releasing into the cytoplasm for subsequent translation[1,2]. In human cells, the transport receptor protein Tap forms a heterodimer with cofactor p15 to facilitate mRNP export by interacting with the FG Nups[13,14]. After shuttling through the NPC, mRNPs are directionally dissociated into the cytoplasm by the essential mRNP export factors Gle1, IP6 and DDX at the location of Nup214[2,15,16]. Electron microscopy has provided a high-resolution, but static picture

[2]Reproduced in part from Jiong Ma, Zhen Liu, Nicole Michelotti, Ram Veerapaneni, Sethuramasundaram Pitchiaya, John R. Androsavich, Nils G. Walter and Weidong Yang. Nature communications, under review.

[3] Nicole Michelotti programmed the Monte Carlo simulations software. Nicole Michelotti and Jiong Ma performed computations simulations. Jiong Ma and Weidong Yang designed and performed experiments. Ram Veerapaneni, Sethuramasundaram Pitchiaya, John R. Androsavich, Nils Walter, and Weidong Yang established cell lines. Jiong Ma, Zhen Liu, and Weidong Yang conducted data analysis of single-molecule tracking and 3D simulations.

of some of the mRNP translocation steps[17]. Recently, the real-time imaging of dynamic nuclear export events of single mRNPs has become possible through super-registration wide-field epi-fluorescence and light sheet fluorescence microscopy[7-9]. Despite this current level of understanding, however, most dynamic features of nuclear export remain obscure.

We previously developed a single-molecule fluorescence imaging approach, termed single-point edge-excitation sub-diffraction (SPEED) microscopy, and used it to capture transient interactions between transiting molecules and the NPC in human cells that had escaped previous observation[11,18-20]. Combined with a deconvolution algorithm, SPEED microscopy further enabled us to distinguish the distinct 3D transport routes between passive and facilitated transport for the first time under real-time trafficking through the native intact NPCs, which helped to refine the current understanding of the nuclear transport mechanism[18-20]. The implementation of Monte Carlo simulations aided in the interpretation of these results. Here we have expanded SPEED microscopy to measure the kinetics and map the 3D transport pathways of single mRNPs through NPCs in living cells, achieving a spatiotemporal super-accuracy of 8 nm and 2 ms (Appendix C). With an ~3-fold higher spatial and 10-fold higher temporal resolution than any previous measurements[7-9], and with the first 3D view of the mRNP transport route, critical details of the dynamic features and selectivity of nuclear mRNP export are finally coming into focus.

2.2 Materials and Methods

Cell Culture and Transport Conditions. Our genetically engineered HeLa cell line expresses stably GFP conjugated POM121 and transiently mCherry-tagged firefly luciferase mRNA. In detail, POM121, a structural Nup anchored at the central scaffold of the NPC, was genetically fused to GFP. Firefly Luciferase mRNA was extended by 24 MS2 stem loops on the 3' end, and

each loop was the binding target of an MS2 coating protein (MCP) dimer fused to mCherry. pMCP-EGFP, a plasmid bearing the open reading frame (ORF) of the MCP fused to EGFP and an SV-40 nuclear localization signal (NLS), and pSL-MS2_24x, a plasmid bearing 24 copies of the MS2 stem loops, were received as generous gifts from Robert Singer (Albert Einstein College of Medicine). The GFP ORF in pMCP-GFP was replaced with PCR amplified mCherry ORF bearing AgeI and BsrGI restriction enzyme sites, to result in pMCP-mCherry. A Luciferase reporter plasmid bearing the MS2 stem loops was created in two steps. First, the ORF of IF2, PCR amplified using a forward primer bearing SbfI and EcoRI restriction enzyme sites and a reverse primer bearing a NotI restriction enzyme site, was cloned into the corresponding SbfI and NotI sites in plasmid pmiR-GLO (Promega). 24 copies of the MS2 stem loop from pSL-MS2_24x were then cloned into the EcoRI-NotI restriction enzyme sites of the resultant plasmid, pmG-IF2, to generate pmG-MS2. We had to resort to this two-step procedure because linearized plasmids containing the MS2 stem loops often recombined with the genome of the bacterial competent cells we were using, thus resulting in clones bearing smaller plasmids. Clones containing the MS2 stem loops were created in SURE2 bacterial cells (Stratagene) to minimize recombination of the MS2 repeats with the bacterial genome. The NLS was added to MCP to promote its import into the nucleus and enhance sufficient labeling of mRNAs even at the desired low MCP expression level. Extensive control experiments showed that eight copies of POM121 in the selected NPCs were labeled with GFPs (Figure C.1) and each mRNP complex imaged in our experiments was tagged with ~10 GFP or mCherry molecules (Figure C.2; this number of fluorescent protein molecules per mRNP was also expected due to using a tagging procedure similar to the one pioneered by Rob Singer). The fluorescence images of GFP and mCherry were used to localize the centroid of NPC and track mRNP export in living cells,

respectively. Freshly split HeLa cells were grown overnight on coverslips in DMEM (Grand Island, NY) supplemented with 10 mg/mL Streptomycin, 100 U/ml Penicillin, 10% Newborn Calf Serum, at 37°C, in a 5% CO_2 incubator.

Single molecule imaging and tracking experiments of the nuclear export of labeled mRNPs were conducted in living cells, whereas the nucleocytoplasmic transport of Alexa Flour 647 labeled Tap-p15 and GFP was imaged in digitonin-permeabilized cells. For the experiments conducted in permeabilized cells, flow chambers were constructed with a top coverslip and two lines of silicone grease as spacers. Cells were washed with transport buffer (20 mM HEPES-KOH, 110 mM KOAc, 5 mM NaOAc, 2 mM $Mg(OAc)_2$, 1 mM EGTA, pH 7.3), permeabilized for two min with 40 μg/mL digitonin in transport buffer, and washed again with transport buffer supplemented with 1.5% polyvinylpyrrolidone (PVP; 360 kDa). PVP was included in all transport buffer solutions after digitonin treatment to prevent osmotic swelling of nuclei.

SPEED Microscopy. The SPEED microscope consists of an Olympus IX81 equipped with a 1.4 N.A. 100x oil-immersion apochromatic objective (UPLSAPO 100XO, Olympus), a 35 mW 633 nm HeNe laser (Melles Griot), an on-chip multiplication gain CCD camera (Cascade 128^+, Roper Scientific), two 120 mW ArKr tunable ion lasers (Melles Griot), and the Slidebook software package (Intelligent Imaging Innovations) for data acquisition and processing. An optical chopper (Newport) was used to generate an on-off mode of laser excitation. GFP, YFP, mCherry, tdTomato and Alexa Fluor 647 were excited by 488 nm, 568 nM and 633 nm lasers, respectively. To minimize alignment problems in dual color measurements, both green and red fluorescence emissions were collected by the same objective, filtered by a dichroic filter (Di01-

R405/488/561/635-25x36, Semrock) and an emission filter (NF01-405/488/561/635-25X5.0, Semrock), and imaged by an identical Cascade 128^{+} CCD camera successively.

Proteins and Labeling. GFP and human $Tap_{\Delta NLS}$-p15 heterodimer proteins were expressed in *E. coli* and purified by Ni-NTA Superflow (Qiagen), MonoQ, and Superdex 200 (Amersham) chromatography. The solvent accessible cysteines on the proteins were labeled with a 20-fold molar excess of Alexa647 maleimide dye (Invitrogen) for 2 h at room temperature in 50 mM sodium phosphate, 150 mM NaCl, pH 7.5. Reactions were quenched with 2-mercaptoethanol and the products were dialyzed to remove the free dye. The labeling ratio was about 2 dye molecules per TapΔNLS/p15 heterodimer, and 1 dye per GFP.

Copy Number of MCP-FP per mRNP and Size of Firefly Luciferase mRNPs. To determine the copy number of MCP-FP molecules per mRNP, we quantified the fluorescence of a GFP and that of NPCs with eight copies of GFP-POM121 *in vitro*, and GFP-MCP-mRNP *in vivo*, each at the single molecule level. As shown in Figure C.2, each mRNP detected in our cell system has 9-10 copies of GFP-MCP, which agrees well with previous reports in a similar MS2-mRNP cell system[7].

The mRNA with the 24 MS2 loops has a length of ~3.3 kb (about 1.5 MDa) and each attached GFP-MCP is 39 kDa. With multiple endogenous proteins bound to the mRNA, the final average mass of the mRNPs is estimated to be 4-5 MDa, with a corresponding diameter of up to ~25 nm when assuming a roughly spherical shape.

Localization of the NE and Orientation of a Single NPC on the NE. The position of the nuclear envelope (NE) was determined at super-accuracy by fitting the fluorescence of GFP-POM121 as follows. The pixel intensities within a row or a column approximately perpendicular to the NE were fit with a Gaussian. The peak position of the Gaussian for a particular set of pixel intensities was considered the NE position for that row and column. The peak positions of a series of such Gaussians were then fit with a second-degree polynomial, yielding the orientation of the NE within the entire image.

The following rules were then used to select a single NPC and determine its orientation, which needs to be perpendicular to the NE on the equator of the nucleus and to the y direction of the Cartesian coordinates (x, y) in the CCD camera: (i) To focus on a GFP-NPC with eight copies of GFP-POM121, individual GFP-NPCs on the NE were selected when their fluorescence intensity was ~8-fold that of a single GFP; (ii) we chose a fluorescent NPC on the equator of the nucleus such that the tangent of the NE at the location of this NPC was parallel to the y-direction of the Cartesian coordinates (x, y) in the CCD camera; and (iii) we examined the ratio of Gaussian widths in the long and short axes of the chosen GFP-NPC fluorescence spot, which needed to fall between 1.74 and 1.82. Within this range, an illuminated NPC only has a free angle of 1.4^{o} to the perpendicular direction to the NE[19].

Localization Precisions of Isolated Fluorescent Spots. The localization precision for fluorescent NPCs, immobile and moving fluorescent molecules was defined as how precisely the central point of each detected fluorescent diffraction-limited spot was determined. For immobile molecules and fluorescent NPCs, the fluorescent spot was fitted with a 2D symmetrical and elliptical Gaussian, respectively, and the localization precision was determined by the standard

deviation of multiple measurements of the central point. By contrast, for moving molecules, the fluorescent spot was fitted with a 2D elliptical Gaussian function, and the localization precision (σ) was determined as: $\sigma = \sqrt{\frac{s^2}{N} + \frac{a^2}{12N} + \frac{8\pi s^4 b^2}{a^2 N^2}}$, where N is the number of collected photons, a is the effective pixel size of the detector, b is the standard deviation of the background in photons per pixel, and s is the standard deviation of the point spread function[31].

To verify the precision as obtained from the standard deviation of multiple measurements and the above equation, both methods were used to determine the localizations of immobile fluorescent molecules (Alexa Fluor 647-labeled-GFP absorbed on the surface of a cover-slip) and fluorescent NPCs. A total of 230 immobile GFP-labeled molecules were measured, and the two methods yielded a difference of 0.5 ± 0.1 nm.

In our measurements, typically 3,000 signal photons versus 15 noise photons were collected from 10 mCherry-labeled mRNP molecules with a detection frame rate of 500 Hz in living cells. Correspondingly, the localization precision was ~8 nm for the immobile single mRNPs and ~10 nm for moving mRNPs. Due to inevitable vibration of NPCs on the NE of living cells, the localization precision of the NPC centroid was 5-11 nm. Additionally, based on the detection of 230 immobile Alexa647-labeled GFP molecules adsorbed on a coverslip, the system error (σ_{sys}) of aligned red and green fluorescence channels was determined to 3.0 ± 0.1 nm. Therefore, the overall tracking precision for mCherry-fused mRNP export through the GFP-labeled NPC in living cells was ~11-15 nm.

Monte Carlo Simulations of Particles Moving through the NPC. It is challenging to directly measure the diffusion coefficients of particles transiting through each sub-region of the NPC in living cells, that is, the nucleoplasmic side, the central region and the cytoplasmic side.

Therefore, Monte Carlo computational simulations were employed to assess the dynamic features of mRNPs moving through these sub-regions of the NPC.

Home-built Monte Carlo simulation software (MATLAB) was programmed to track the position of N non-interacting particles experiencing unbiased diffusion, modeled using a random walk algorithm, throughout the NPC over time. The software requires user input to specify multiple parameters: the size of each sub-region of the NPC, the number of particles to be tracked, the maximum step-size a given particle can take during each iteration, the number of steps to be tracked, the fractional decrease in the maximum step-size for particles in the central region, and the uncertainty in position determination (optional). The initial position for each particle was chosen at random. Each subsequent position of a given particle was determined by adding a random number between the negative maximum steps-size specified and the maximum step-size specified to each of the three dimensions, with a different random number determined for each dimension. If the position chosen was a boundary, another position was chosen at random, meaning there will be greater error in position determination near the boundaries than in the open areas. To test the impact this error along the edge would have on the results, various step-sizes were tested. A larger step-size should exhibit a greater error. The results were found to be independent of the step-size used, indicating that this error along the border is negligible. If the particle was in the central region, the maximum step-size was decreased by an amount specified by the user. A histogram was used to present of the distribution of particles across a two-dimensional projection of each (3D) sub-region of the NPC. Parameters were chosen so that this histogram resembles the histograms generated for the experimental data (Figure 2.2).

If the uncertainty in the position determination of the particle was selected to be incorporated into the simulation, the position of the particle was modified after each step; instead

of the final position of the particle being the exact location specified by the random numbers chosen for each dimension, the position assigned the particle was randomly chosen from within the area specified by the diameter of uncertainty input. A second histogram of the distribution of particles given the uncertainty across each sub-region of the NPC was computed.

Deconvolution and 3D Image Processing. The detailed deconvolution process used to obtain the 3D spatial probability density maps of particles transiting through the NPC was described in our previous publications[11, 18-20]. In short, the 3D spatial locations of molecules transiting through the NPC can be considered in either Cartesian (x, y, z) or cylindrical (x, r, θ) coordinates. In microscopic imaging, the observed 2D spatial distribution of particle localizations is a projection of its actual 3D spatial locations onto the x, y-plane. The underlying 3D spatial distributions can be recovered by projection of the measured Cartesian (x, y) coordinates back onto the simplified cylindrical (r, x) coordinates, based on the expected cylindrically symmetric distribution along the θ direction of the nuclear pore[19]. The resulting 3D, surface-rendered visualizations shown in figures were generated with Amira 5.2 (Visage Imaging).

Standard Error. Experimental measurements are reported as mean ± standard error of the mean unless otherwise noted.

2.3 Results

For firefly luciferase mRNPs we found that 34 ± 5% of all nuclear mRNP export events were successful (i.e., started from the nucleus, interacted with the NPC, and arrived in the cytoplasm), while the remaining 66 ± 5% events were aborted (i.e., started from and returned to the nucleus after interacting with the NPC), consistent with previous reports[7,9]. The mRNP:NPC

interaction times were similar at 11 ± 2 ms and 12 ± 2 ms for successful and abortive export events, respectively (Figure 2.1f-g). To test the generality of our observations, the export efficiency and time for similar-sized, YFP-labeled β-actin mRNPs were measured under the same experimental conditions, and were found to be the same as those for firefly luciferase mRNPs (Table 2.1). Strikingly, the export times obtained for both mRNPs were more than 10-fold shorter than those previously reported for YFP-labeled β-actin mRNPs[7,8]. We found that this difference was at least in part related to the technical advances of SPEED microscopy, including: i) Faster detection speed. Our detection time resolution of 2 ms per frame, 10-times faster than that previously achieved, enabled us to capture faster mRNA export events that may have escaped earlier observation. Indeed, when we lowered the detection resolution to 20 ms per frame, we found that the transport time was resolved to 21 ± 1 ms for β-actin mRNPs and 16 ± 2 ms for firefly luciferase mRNPs (Table 2.1). ii) Capability of localizing a single NPC on the NE with high localization precision. SPEED microscopy only illuminates a single NPC embedded in the NE of living cells. The centroid position of the imaged single NPC is localized with 1-3 nm precision in permeabilized cells and 5-11 nm in living cells. By contrast, wide-field epi-fluorescence microscopy simultaneously excites multiple proximal GFP-NPCs on the NE (Figure C.6). The resulting, heavily overlapping diffraction-limited fluorescent spots of GFP-NPCs form a continuous fluorescent NE, complicating an accurate determination of the NE middle plane or the centroid of individual NPCs on the NE (Figure C.6). In addition, many slowly moving or even immobile mRNPs reside around the periphery of the NE on either the nuclear or cytoplasmic side (Figure C.7). With a less accurately determined NE middle plane or NPC centroid, one would inevitably have difficulty to determine whether some mRNPs move inside or outside an NPC, and thus one may include some mRNPs that do not undergo export through the

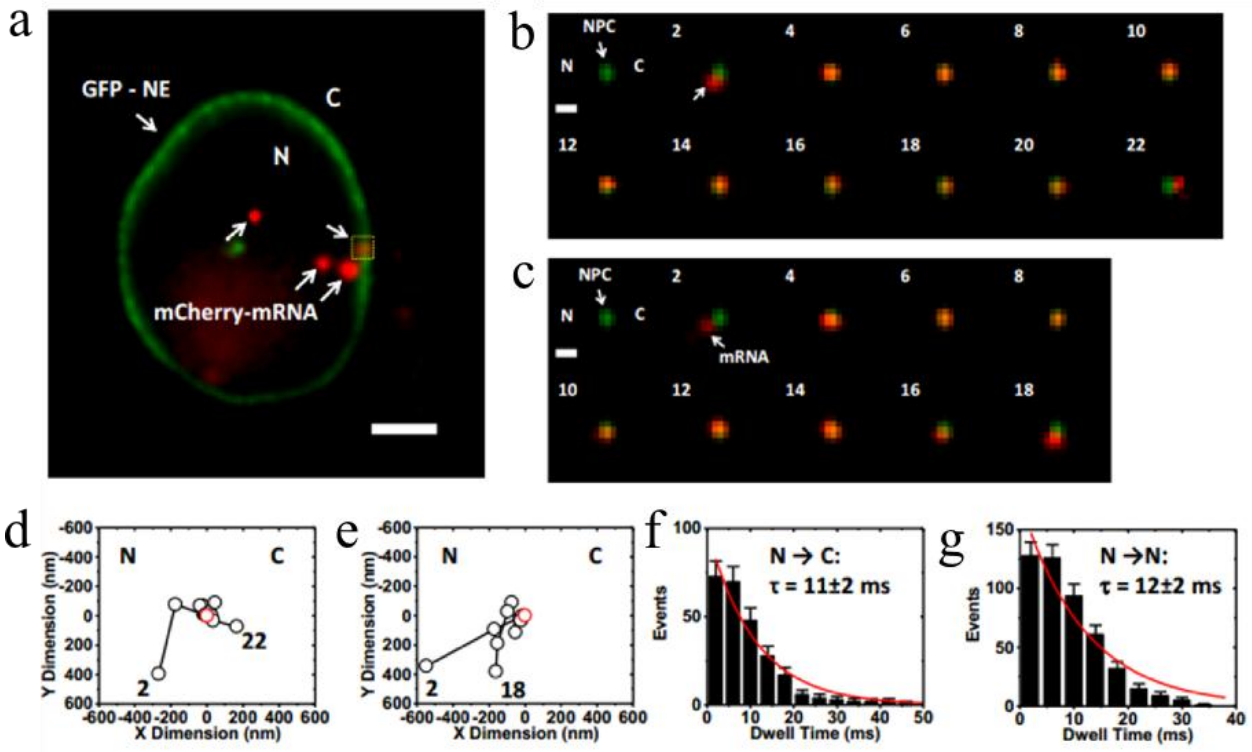


Figure 2.1. Tracking single mRNPs through single NPCs. (**a**) Single mRNPs in the nucleus and on the NE imaged by wide-field epi-fluorescence. Four mCherry-tagged mRNP complexes (red spots) were observed, three in the nucleus and one on the GFP-NE (green rim). Scale bar: 5 □m. C, cytoplasm. N, nucleoplasm. (**b**) A typical successful single mRNP export event captured by SPEED microscopy. A single mCherry-tagged mRNP (red spot) started from the nucleus, interacted with a single GFP-tagged NPC (green spot), and arrived in the cytoplasm. Numbers denote time in milliseconds. (**c**) A typical abortive single mRNP export event. A single mCherry-tagged mRNP (red spot) started from nucleus, interacted with a single GFP-tagged NPC (green spot), and returned to the nucleus. (**d**, **e**) Single-particle tracks (black open dots) and the centroid of the NPC (red open dot) were acquired by 2D Gaussian fitting to point spread functions in a series of images for either the successful event in (b) or the abortive event in (c). (**f**) Export time distribution of successful mRNP export events. Fitting with a mono-exponential decay function (red line) yielded the indicated total export time. (**g**) Export time distribution of abortive mRNP export events.

Cell type	RNA Type (3.3 kb)	Label	Detecting Single mRNAs on or in	Detection Time per Frame (ms)	Dwell Time (ms)	Efficiency (%)
MEF	β-actin	YFP-MCP	NE	20	21±1	32±5
HeLa	Firefly luciferase	mCherry-MCP	NE	20	16±2	31±5
MEF	β-actin	YFP-MCP	Single NPC	2	12±1	36±5
HeLa	Firefly luciferase	mCherry-MCP	Single NPC	2	11±2	34±5
HeLa	Firefly luciferase	mCherry-MCP	Nucleus, cytoplasm and NPC-deficit NE	200	t_1=620±130 t_2=65,000±21,000	n/a

Table 2.1 Intracellular dwell times and export efficiencies for single labeled β-actin and firefly luciferase mRNPs exported through single NPCs in the NE, as well as the detected two dwell times (t_1 and t_2) in the cytoplasm and the nucleus. n/a, not applicable.

NPC. To explore the above problems, we measured the dwell times of mRNPs outside the nucleus, including mRNPs that interacted with NPC-deficient regions on the NE, or moved within the nucleus or cytoplasm. Their dwell times were determined to be as long as hundreds of milliseconds or even seconds (Table 2.1).

Altogether 2,379 spatial localizations from export events of firefly luciferase mRNPs through 20 NPCs from 20 cells were superimposed to obtain the 2D spatial distribution map of Figure 2.2a, and identified a void space at the position of the NE, as expected. The binned histogram of mRNP locations along the NPC axis indicated more localizations on both the nucleoplasmic (-120 nm to -20 nm) and cytoplasmic sides (20 nm to 120 nm) compared to the NPC center (-20 nm to 20 nm), a distribution that could be fitted with two Gaussian functions (Figure 2.2b). These concentrated localizations of mRNPs on either side of the NPC were previously observed and assigned to cofactor assembly and disassembly processes, respectively[3]. However, we observed similarly clustered localizations for passively diffusing GFP and a $Tap_{\Delta NLS}$-p15 complex (here, the transport receptor Tap lacks an NLS to prevent complex formation with NLS-related transport receptors, such as Importins α and $\beta 1$[21]), neither of which requires such (dis)assembly (Figure 2.2d-h). We therefore conclude that assembly and disassembly of mRNPs unlikely cause the two localization clusters observed here.

To gain further insight into the diffusive behavior of translocating mRNPs, we utilized Monte Carlo simulations under the geometric constraints of the NPC, and applied one of three alternative sets of constraints: (i) in the "constant" model mRNPs move through the entire NPC with a constant diffusion coefficient; (ii) in the "slow-fast-slow" model higher diffusion coefficients (faster movements) apply at the NPC center; and (iii) in the "fast-slow-fast" model lower diffusion coefficients (slower movements) apply at the NPC center (Appendix C).

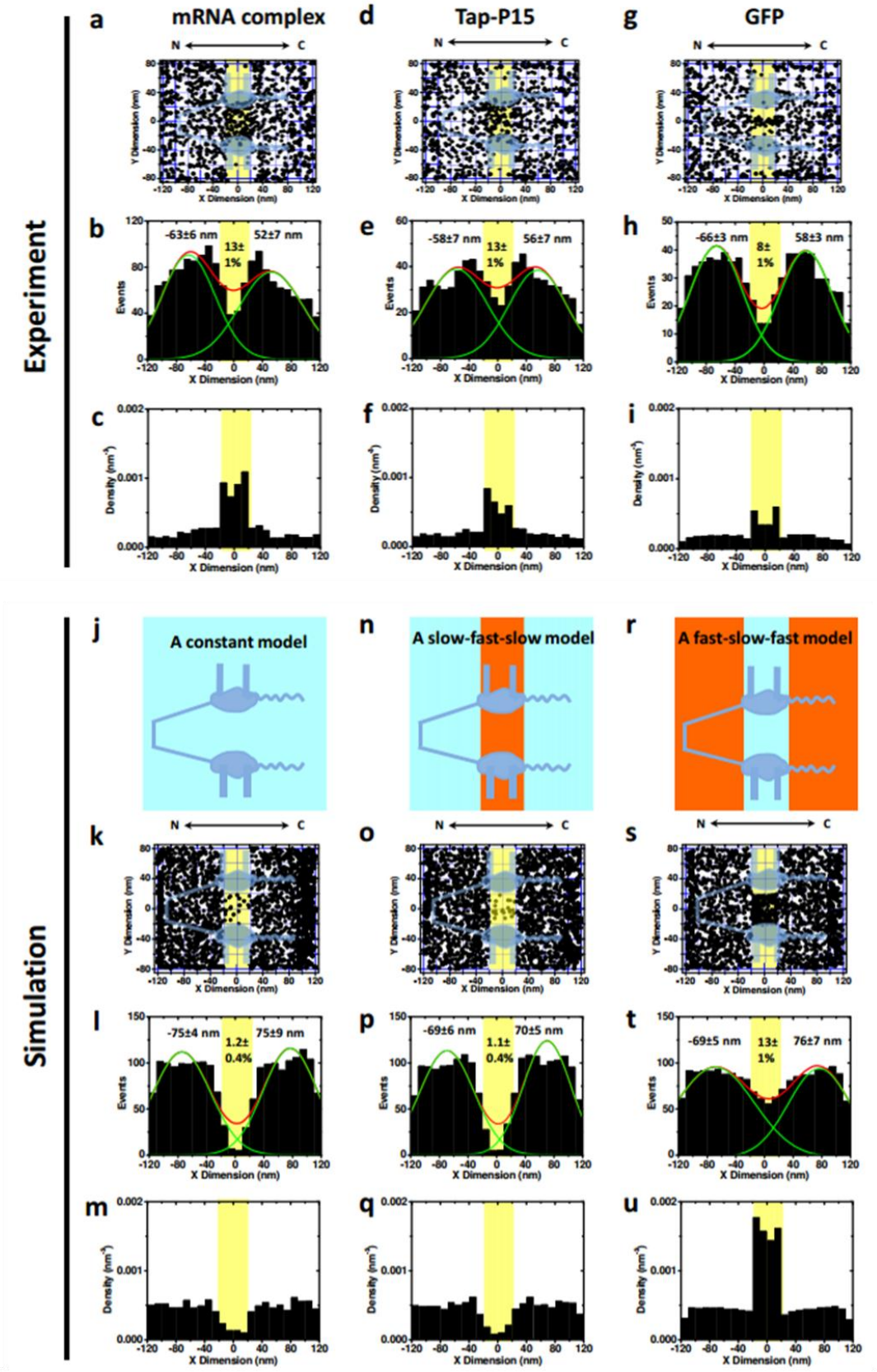

Experiment
Simulation
a
mRNA complex
d
Tap-P15
g
GFP
N
C
Y Dimension (nm)
X Dimension (nm)
b
e
h
Events
-63±6 nm
13±
1%
52±7 nm
-58±7 nm
56±7 nm
-66±3 nm
8±
58±3 nm
c
f
i
Density (nm⁻¹)
j
A constant model
n
A slow-fast-slow model
r
A fast-slow-fast model
k
o
s
l
-75±4 nm
1.2±
0.4%
75±9 nm
p
-69±6 nm
1.1±
70±5 nm
t
-69±5 nm
76±7 nm
m
q
u

Figure 2.2. 2D spatial distribution of molecular locations in the NPC. (**a**) Experimentally determined 2D spatial locations of mRNPs in the NPC. A schematic of the NPC (light blue) is superimposed and the central region of the NPC (-20 nm to 20 nm) is highlighted in yellow. N, nucleoplasmic side of the NPC; C, cytoplasmic side of the NPC. (**b**) Gaussian fitting of the binned histogram of spatial mRNP locations along the NPC axis (x dimension) produced two major peaks on either side of the NPC. Approximately 13% of total localizations fall in the central channel of the NPC (-20 nm to 20 nm). (**c**) Density of spatial locations along X dimension. (**d** to **f**) The 2D spatial locations, binned histogram of spatial locations along the NPC axis, and corresponding spatial density for $Tap_{\Delta NLS}$-p15. (**g** to **i**) The 2D spatial locations, binned histogram of spatial locations along the NPC axis, and corresponding spatial density for GFP. (**j**) Diagram of the simulated "constant" model, where mRNPs move through the entire NPC with a constant diffusion coefficient (cyan). (**k** to **m**) The predicted 2D spatial locations, binned histogram of spatial locations along the NPC axis, and corresponding spatial density from Monte Carlo simulations. The central region of the NPC (-20 nm to 20 nm) is highlighted in yellow. (**n** to **q**) Diagram of the simulated "slow-fast-slow" model, where mRNPs diffuse through either side (cyan) and accelerate in the central NPC channel (orange), with the 2D spatial locations, binned histogram of spatial locations along the NPC axis, and corresponding spatial density predicted from Monte Carlo simulations. (**r** to **u**) Diagram of the simulated "fast-slow-fast" model, where mRNPs move faster on either side (orange) and decelerate in the central NPC channel (cyan), with the 2D spatial locations, binned histogram of spatial locations along the NPC axis, and corresponding spatial density predicted from Monte Carlo simulations.

As shown in Figure 2.2j-t, histograms along the NPC axis for each model reveal major clusters of locations on either side of the NPC, as also observed experimentally. These observations suggest that the existence of such clusters does not reflect the assembly and disassembly of mRNPs, but simply results from the geometric constraints imposed by the NPC that allows for more spatial dwellings outside its narrow center. In addition, our analysis revealed that only the "fast-slow-fast" model accurately reproduces the spatial densities in the NPC center experimentally observed for transiting mRNPs, $Tap_{\Delta NLS}$-p15 and GFP (compare Figure 2.2t-u, to Figure 2.2c, f, i). Therefore, both experimental measurements and simulations suggest that mRNPs, $Tap_{\Delta NLS}$-p15 and GFP are decelerated when passing through the NPC center, opposite to a previous proposal based on intracellular single particle tracking[7], but consistent with the previous suggestion of a physical/entropy barrier in the crowded NPC center[22,23].

Our recent studies have shown that the determination of nucleocytoplasmic transport pathways based solely on 1D or 2D spatial localization are incomplete[11,18-20]. Following well-established deconvolution procedures[18-20], we therefore converted our 2D spatial distributions of mRNPs, $Tap_{\Delta NLS}$-p15 and GFP in the NPC (Figure 2.2a) into their corresponding 3D probability density maps (Figure 2.3). Comparison of the 3D maps revealed that mRNPs translocate through the periphery around a rarely entered central axial channel of the NPC before they dissociate on the cytoplasmic side. Strikingly, mRNPs and $Tap_{\Delta NLS}$-p15 show major overlap between their export pathways, particularly in the NPC center and on the nucleoplasmic side (Figure 2.3d), consistent with the fact that Tap-p15 mediates the interaction of mRNPs with the FG Nups on the nucleoplasmic face and in the central region of the NPC[2]. The diameter of the unoccupied channel at the cross-section of the NPC center is 10 ± 1 nm for mRNPs and 15 ± 1 nm for Tap-p15, indicating that the larger mRNPs still do not occupy the entire volume of the NPC. By

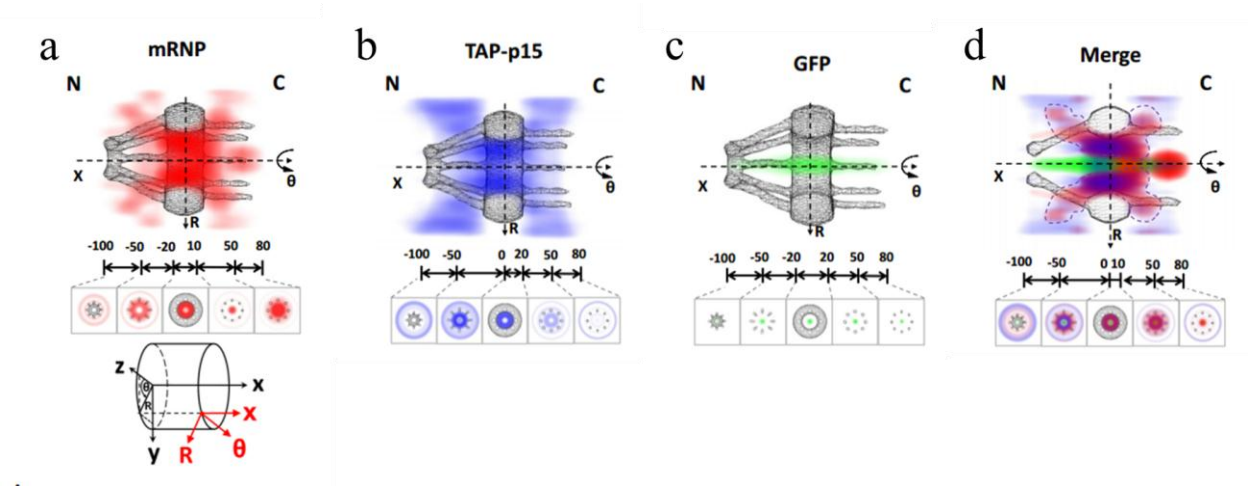


Figure 2.3 3D spatial probability density maps (3D transport routes) for mRNP, TapΔNLS-p15 and GFP. (**a**) 3D spatial density map of mRNPs (red, deeper shade indicates higher density), generated using a 2D-to-3D deconvolution algorithm, is shown in both a cut-away and a cross-section view superimposed on the NPC architecture (grey). Five regions with distinct spatial location distributions for mRNPs are marked with relative distances (in nm) from the centroid of the NPC. Cartesian (x, y, z) and cylindrical (x, R, θ) coordinate systems are shown. N, nucleoplasmic side of the NPC; C, cytoplasmic side of the NPC. (**b**) Cut-away and cross-section views of the 3D spatial density map for the Tap-p15 heterodimer (blue, deeper shade indicates higher density) superimposed on the NPC architecture (grey). (**c**) Cut-away and cross-section views of the 3D spatial density map for passively diffusing GFP molecules (green, deeper shade indicates higher density) superimposed on the NPC architecture (grey). (**d**) Merged maps of the three transport routes through the NPC. The regions overlapping between mRNP and $Tap_{\Delta NLS}$-p15 are outlined with dashed lines.

contrast, this central axial channel is utilized by the passively diffusing GFP (Figure 2.3c). In agreement with previous conclusions[18], passively diffusing molecules such as GFP (or mCherry) take a path consistent with size confinement in the central channel of the NPC[24], whereas mRNPs and Tap-p15 conquer the central dense FG-repeat barrier anchored at the central channel wall. Both types of interaction, however, lead to the observed decelerated diffusion.

After passing through the central NPC region, mRNPs start to broadly diffuse around the cytoplasmic fibrils and invade the passive diffusion channel on the cytoplasmic side (Figure 2.3a), differing thus from the $Tap_{\Delta NLS}$-p15 paths. This difference is consistent with the previous proposal that DDX and other transport cofactors disassemble the mRNP:Tap-p15 complex from Nup214 on the cytoplasmic side of the NPC[25,26], a process that may not be a required for Tap-p15 alone. We propose that mRNPs dissociate from the NPC in this zone, 50-80 nm into the cytoplasmic side.

To ask what impact Tap-p15 association has on the export efficiency of mRNPs, we determined the export efficiency of $Tap_{\Delta NLS}$-p15 alone to be ~50%, consistent with the transport efficiencies of other receptors such as Crm1 and Importin β1 (Table C.1)[18,19]. Moreover, the NPC acquiesces only ~50% transport efficiency for either passively diffusing molecules such as GFP, or for Importin β1-facilitated translocation of cargo complex (Table C.1)[18,19,27]. That is, only half of all incoming molecules successfully pass through the NPC under normal physiological conditions during both passive and facilitated transport. Although the process of transport in living cells may be more complex than that in permeabilized cells, previous studies have shown similar nucleocytoplasmic transport times and efficiencies for either cargo-bound or cargo-free transport receptors when comparing permeabilized and living HeLa cells[27, 30]. We therefore propose that only an additional ~16%, rather than ~66%, of all mRNP export attempts

(corresponding to ~32% of all abortive events) are rejected by the NPC's central channel. The lower abortive probability and shorter export time determined here suggest a more efficient mRNP nuclear export in living cells than previously thought, with important consequences for our understanding of the dynamics of genetic information flow in eukaryotic cell viability and growth.

Finally, to understand the mechanism of mRNP selectivity, we compared the spatial distributions of successful and abortive mRNP export events (Figure 2.4). The superimposed trajectories of 38 successful mRNP export events showed a peak centered around 47 nm into the cytoplasmic side. The deepest penetrations of these events inevitably move beyond the central region (-20 to 20 nm) and reach the cytoplamic side of the NPC (20 to 120 nm). By contrast, the spatial locations of 50 abortive export events form a peak around -115 nm on the nuclear side. In addition, the localizations of 80% of these mRNPs are restricted to the nuclear side and the central region (-120 to 20 nm). These results suggest that mRNPs only gain a high probabilty to complete their nuclear export once they transverse the NPC center. Notably, recent studies have shown that protein cargo molecules larger than the cut-off size of passive diffusion (>40-60 kDa) are similarly restricted by the NPC central region[11,18,19,27], suggesting that this region generally serves as the critical selection barrier for all actively translocated cargo molecules.

2.4 Discussion

Exploiting the enhanced spatiotemporal super-accuracy afforded by SPEED microscopy, we here have characterized the kinetics, 3D routes and selectivity mechanism for the export of mRNPs through the NPC. First, we have demonstrated that the central FG Nups of the NPC function as the major selective barrier for nuclear export of mRNPs. Only about 34% of all mRNPs entering the NPC from the nucleus successfully complete their export after accessing

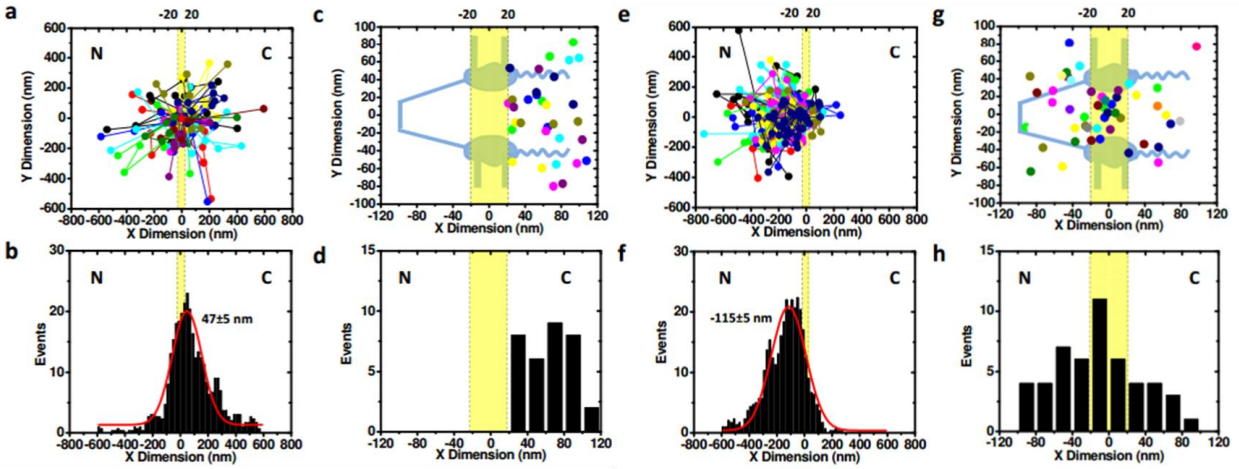


Figure 2.4 Single-molecule trajectories of successful and abortive mRNP export events. (**a**) Superimposed single-molecule trajectories of 38 successful mRNP export events. The central region of the NPC (-20 nm to 20 nm) is highlighted in yellow. N, nucleoplasm; C, cytoplasm. (**b**) The binned histogram of spatial locations for the successful mRNP export events along the NPC axis was fitted by a Gaussian function to yield a peak at position 47 ± 5 nm on the cytoplasmic side. (**c**) Spatial locations of the last stage of the successful mRNP export events shown in (a). (**d**) Binned histogram of locations in (c). (**e**) Superimposed single-molecule trajectories of 50 abortive mRNP export events. (**f**) The binned histogram of spatial locations for the abortive mRNP export events in (e) was fitted by a Gaussian function to yield a peak at position -115 ± 5 nm on the nucleoplasmic side. (**g**) Spatial locations of the deepest penetration steps for the abortive mRNP export events shown in (e). (**h**) Binned histogram of locations in (g).

this central selective barrier, whereas the remaining mRNPs abort their export after interacting with the FG filaments in the central and cytoplasmic regions of the NPC. Accordingly, we found that mRNPs adopt a fast-slow-fast diffusion pattern through the NPC, moving faster on either end and slower in the central region of the NPC, in direct opposition to the previous proposal that mRNP diffusion is attenuated in both peripheral regions of the NPC, but not its center[3,7]. Second, mRNPs only need ~12 ms for transit through the NPC, 10-times shorter than previously suggested[7]. This much faster nuclear export was observed for both firefly luciferase and β-actin mRNP in living cells (Table 1), the latter chosen for direct comparison with the previous work that found ~10-fold slower export times for this mRNA/cell line combination[7]. Third, reconstructions of the 3D export routes in living cells revealed that large mRNP complexes primarily interact with the periphery of the nucleoplasmic side and the center of the NPC, only rarely occupying the central axial channel utilized for passive diffusion, and eventually dissociate on the cytoplasmic side. The latter results invoke a multilane traffic model wherein the signal-independent passive diffusion of small molecules and the transport-receptor facilitated diffusion of large mRNP complexes follow distinct transport paths through the NPC[18-20, 28-29]. SPEED microscopy has thus proven to be a powerful tool to study nucleocytoplasmic transport of proteins and mRNPs *in vitro* and *in vivo*[18-20]. In the future, the method can be expanded to other types and/or sizes of RNPs, or to the observation of conformational changes associated with their transport though the NPC.

2.5 Acknowledgements

We thank Yoshihiro Yoneda for the plasmids of $Tap_{\Delta NLS}$-p15, David Gruenwald for YFP-labeled β-actin mRNA MEF cells, Rob Singer for the gift of several plasmids, and Allen Nicholson for

his critical reading. The project was supported by grants from the National Institutes of Health (NIH GM094041 and GM097037 to W.Y. and GM081025 to N.G.W.).

2.6 References

1. Kohler, A. & Hurt, E. Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell. Biol.* **8**, 761-773 (2007).
2. Carmody, S. R. & Wente, S. R. mRNA nuclear export at a glance. *J. Cell Sci.* **122**, 1933-1937 (2009).
3. Grunwald, D., Singer, R. H. & Rout, M. Nuclear export dynamics of RNA–protein complexes. *Nature* **475**, 333-341 (2011).
4. Stewart, M. Nuclear export of mRNA. *Trends Biochem. Sci.* **9**, 569-578 (1999).
5. Beach, D., Salmon, E. & Bloom, K. Localization and anchoring of mRNA in budding yeast *Current Biology* **9**, 569-578 (1999).
6. Rout, M. et al. The Yeast Nuclear Pore Complex: Composition, Architecture, and Transport Mechanism. *J. Cell Biol.* **148**, 635-651 (2000).
7. Grunwald, D. & Singer, R. H. In vivo imaging of labeled endogenous β-actin mRNA during nucleocytoplasmic transport. *Nature* **467**, 604-607 (2010).
8. Mor, A. *et al.* Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. *Nat. Cell Biol.* **12**, 543-552 (2010).
9. Siebrasse, J. P. *et al.* Nuclear export of single native mRNA molecules observed by light sheet fluorescence microscopy. *Proc. Natl. Acad. Sci. USA* **109**, 9426-9431 (2012).
10. Beck, M. *et al.* Nuclear pore complex structure and dynamics revealed by cryo-electron tomography. *Science* **306**, 1387-1390 (2004).
11. Yang, W. "Natively Unfolded" Nucleoporins in Nucleocytoplasmic Transport: Clustered or Evenly Distributed? *Nucleus* **2**, 10-16 (2010).
12. Patel, S., Belmont, B., Sante, J. & Rexach, M. Natively Unfolded Nucleoporins Gate Protein Diffusion across the Nuclear Pore Complex. *Cell* **129**, 83-96 (2007).
13. Guzik, B. W. *et al.* NXT1 (p15) is a crucial cellular cofactor in TAP-dependent export of intron-containing RNA in mammalian cells. *Mol. Cell. Biol.* **21**, 2545-2554 (2001).
14. Braun, I. C. *et al.* Overexpression of TAP/p15 Heterodimers Bypasses Nuclear Retention and Stimulates Nuclear mRNA Export. *J. Biol. Chem.* **276**, 20536-20543 (2001).
15. Dossani, Z. Y. *et al.* Structure of the C-terminus of the mRNA export factor Dbp5 reveals the interaction surface for the ATPase activator Gle1. *Proc. Natl. Acad. Sci. USA* **106**, 16251-16256 (2009).
16. Kendirgi, F. *et al.* Interaction between the Shuttling mRNA Export Factor Gle1 and the Nucleoporin hCG1: A Conserved Mechanism in the Export of Hsp70 mRNA. *Mol. Biol. Cell.* **16**, 4304-4315 (2005).
17. Bjork, P. & Wieslander, L. Nucleocytoplasmic mRNP export is an integral part of mRNP biogenesis. *Chromosoma* **120**, 23-38 (2011).
18. Ma, J., Goryaynov, A., Sarma, A. & Yang, W. Self-regulated viscous channel in the nuclear pore complex. *Proc. Natl. Acad. Sci. USA* **109**, 7326-7331 (2012).
19. Ma, J. & Yang, W. Three-Dimensional Distribution of Transient Interactions in the Nuclear Pore Obtained by Single-Molecule Snapshots. *Proc. Natl. Acad. Sci. USA* **107**, 7305-7310 (2010).
20. Goryaynov, A., Ma, J. & Yang, W. Single-Molecule Studies of Nucleocytoplasmic Transport: from One Dimension to Three Dimensions, *Integr. Biol.* 4, 10-21 (2012).
21. Katahira, J. *et al.* Complex Formation between Tap and p15 Affects Binding to FG-repeat Nucleoporins and Nucleocytoplasmic Shuttling. *J. Biol. Chem.* **277**, 9242-9246 (2002).

22. Rout, M. P., Aitchison, J. D., Magnasco, M. O. & Chait, B. T. Virtual gating and nuclear transport: the hole picture. *Trends Cell Biol.* **13**, 622-628 (2003).
23. Macara, I. Transport into and out of the Nucleus. *Microbiol. Mol. Biol. Rev.* **65**, 570-594 (2001).
24. Cui, S. T. Molecular self-diffusion in nanoscale cylindrical pores and classical Fick's law predictions. *J. Chem. Phys.* **123**, 054706-054709 (2005).
25. Paulillo, S. M. *et al.* Nucleoporin domain topology is linked to the transport status of the nuclear pore complex. *J. Mol. Biol.* **351**, 784-798 (2005).
26. Walther, T. C. *et al.* The cytoplasmic filaments of the nuclear pore complex are dispensable for selective nuclear protein import. *J. Cell. Biol.* **158**, 63-77 (2002).
27. Yang, W. & Musser, S. M. Nuclear transport time and efficiency are dependent on importin beta concentrations. *J. Cell. Biol.* **174**, 951-961 (2006).
28. Naim, B. *et al.* Passive and facilitated transport in nuclear pore complexes is largely uncoupled. *J. Biol. Chem.* **282**, 3881–3888 (2007).
29. Fiserova, J., Richards, S. A., Wente, S. R. & Goldberg, M. W. Facilitated transport and diffusion take distinct spatial routes through the nuclear pore complex. *J. Cell Sci.* **123**, 2773–2780(2010).
30. Dange, T., Grünwald, D., Grünwald, A., Peters, R., Kubitscheck, U. Autonomy and robustness of translocation through the nuclear pore complex: A single-molecule study. *J. Cell Biol.* **183**, 77–86 (2008).

CHAPTER 3:
STUDYING THE WALKING MECHANISM FOR MOLECULAR ROBOTS GUIDED BY PRESCRIPTIVE LANDSCAPES[4,5]

3.1 Introduction

Traditional robots[1] rely on computing to coordinate sensing and actuating components and to store internal representations of their goals and environment. Moving robotics to the single-molecule level is possible in principle, but requires facing the limited ability of individual molecules to store complex information and programs. One strategy to overcome this problem is to use systems that can obtain complex behavior from the interaction of simple robots with their environment[2-4]. An initial step in this direction was the development of DNA walkers[5], which have transitioned from being non-autonomous systems[6,7] to being capable of directed but brief motion on one-dimensional tracks[8-11]. Herein, we demonstrate elementary robotic behaviors from the interaction between a random walker incorporating deoxyribozymes[12] and a precisely defined environment. Single-molecule microscopy observations confirm that such walkers achieve directionality by sensing and modifying their environment, tracks of substrate molecules

[4] Reproduced in part from Lund, K.; Manzo, A. J.; Dabby, N.; Michelotti, N.; Johnson-Buck, A.; Nangreave, J.; Taylor, S.; Pei, R.; Stojanovic, M. N.; Walter, N. G.; Winfree, E.; Yan, H. Nature 465, 206–210. Copyright Nature Publishing Group, 2010.

[5] Nicole Michelotti (majority) and Anthony J. Manzo performed single-particle tracking experiments and RMSD analysis. Anthony J. Manzo, Nicole Michelotti, and Alexander Johnson-Buck collaborated in devising single-particle tracking experiments and trajectory classification criteria. The unpublished addendum was done by Nicole Michelotti with origami samples provided by Jeanette Nangreave and Shuoxing Jiang. Alexander Johnson-Buck performed displacement and velocity analysis of single-particle tracking experiments, fluorescence assays of spider leg cleavage, and Monte Carlo simulations of spider walking. AFM experiments were performed by Kyle Lund (majority), Jeanette Nangreave, and Nadine Dabby. Analysis of AFM experiments was performed by Nadine Dabby, Kyle Lund, Jeanette Nangreave, and Steven Taylor. Spiders were synthesized, purified, and their integrity confirmed and monitored by Steven Taylor. SPR experiments were performed by Renjun Pei. The contents of the manuscript on which this chapter is based were produced with collaboration from all the above individuals as well as Nils G. Walter, Hao Yan, Milan N. Stojanovic, and Erik Winfree.

laid out on a two-dimensional DNA origami landscape[13]. Given appropriately designed origami, these molecular robots autonomously carry out sequences of actions such as "start", "follow", "turn", and "stop". We anticipated that this strategy will result in more complex robotic behavior at the molecular level if additional control mechanisms are incorporated. One example might be interactions between multiple molecular robots leading to collective behavior[14,15]; another might be the ability to read and transform secondary cues on the landscape as a means of implementing Turing-universal algorithmic behavior[2,16,17].

Our walkers, called molecular spiders, comprise a streptavidin protein "body" and three catalytic "legs". The legs are adapted from the 8-17 DNA enzyme that binds and cleaves oligodeoxynucleotide (henceforth "oligonucleotide") substrates with a single ribose moiety (Figure 3.1a,b) into two shorter products that have lower affinities for the enzyme[18]. The difference in affinities between substrate and product ensure that a spider's interactions with a layer of immobilized substrate and/or product sites can be modeled using a simple 'memory' principle[19]: each leg moves independently from sites to accessible neighboring sites, but if a leg is on a site not visited before, it will stay longer on average. Put biochemically, a deoxyribozyme attached to a site that was previously converted to a product will dissociate faster, whereas it will stick longer on the substrates and eventually cleave them. Because spiders have multiple legs, a single dissociated leg will quickly reattach to nearby product or substrate. It follows that the body of a spider positioned at the interface between products and substrates will move toward the substrate region, because after cleaving, each leg will explore neighboring sites until it finds another substrate to bind to for longer. On a linear track of substrates, this mechanism predicts a deviation from an otherwise random walk process, yielding directional movement as the substrates are cleaved. Previously engineered using "burnt bridge"

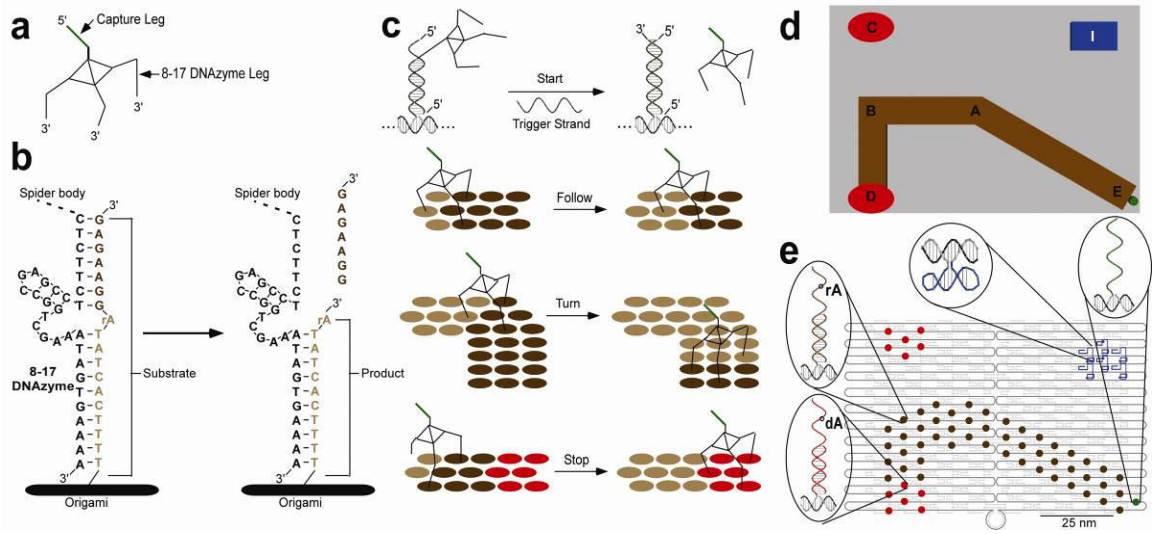


Figure 3.1 Deoxyribozyme based molecular walker and origami prescriptive landscape schematics. a, The NICK3.4A$_{3+1}$ spider consists of a streptavidin core that displays a 20 base ssDNA that positions the spider at the start (green), and three deoxyribozyme legs. **b**, The 8-17 deoxyribozyme cleaves its substrate at an RNA base creating two shorter products (seven and eleven bases). Dissociation from these products allows legs to associate with the next substrate. **c**, Spider actions: after release by a 27-base ssDNA trigger, the spider follows the substrate track, turns, and continues to a stop site (red). **d**, Schematic of the DNA origami landscape with positions A-E labeled; track EABD is shown. **e**, A representative origami landscape shows the start position (green), the substrate track (brown), stop and control sites (red), and a topographical marker (blue).

mechanisms[6-9,11] and Brownian ratchets found in nature[20] render revisiting the same path impossible, but our spiders will perform Brownian walks on product sites until they again encounter substrate.

In analogy to the reactive planning used in simple robots[4], the sensor-actuator feedback afforded when legs sense and modify nearby oligonucleotides allows us to design prescriptive landscapes that direct the spiders' motion along a predefined path (Figure 3.1c,d). Prescriptive landscapes were constructed using the DNA origami scaffolding technique[13]. The scaffold consists of a 7249-nucleotide single-stranded DNA folded by 202 distinct staple strands into a rectangular shape roughly 65x90x2 nm in size and with 6-nm feature resolution (Figure 3.1e). Each staple can be extended on its 5′ end with probes that recruit substrates, products, goal and control DNA strands[21].

We designed pseudo-one-dimensional tracks on origami of about spider width (three adjacent rows of substrates; Figure 3.1d). Tracks are coded by a sequence of points (A, B, C, D, E) such that on an ABD landscape the spider starts at A and passes through B before ending at D. Staples were modified to position a START oligonucleotide, used to position a spider at the start of the experiment, that is complementary to a TRIGGER oligonucleotide used to release the spider[22] (the "start" action); substrate TRACK probes to capture the 5′ extension on substrates forming the TRACK (directing the "follow" and "turn" actions); STOP probes, complementary to the 5′ extension on STOP strands (non-chimeric and uncleavable analogs of the substrate) that do not influence directional movement but trap spiders to prevent them from walking backwards after completing the track (the "stop" action); CONTROL probes (identical to the STOP probes, but disconnected from the track), used to assess the extent to which free-floating spiders are captured directly from solution; and MARKER oligonucleotides based on inert dumbbell

hairpins, aiding in origami classification within atomic force microscopy (AFM) images (Figure 3.1e). To position spiders at START sites, we replaced one of the four catalytic legs of the NICK-4.4A[12] spider with a tethering oligonucleotide (Figures D.1-4) partially complementary to the START oligonucleotide.

To estimate the efficiency of spider motion directed by the TRACK, we defined and tested four paths with no (EAC), one (ABD), or two (EABD, EABC) turns (Figure 3.2 and Figures D.8, 11, 14, 17). The basic experimental procedure involves assembling the origami; attaching the spider to the START site; adding TRACK, STOP, and CONTROL strands to complete the landscape; and initiating an experiment by releasing the spider through addition of TRIGGER and 1 mM Zn^{2+} cofactor[23] (Figures D.6 and D.25). We sampled the origami solution before and after spider release, and imaged individual samples by AFM to determine the locations of spiders. We scored only "face-up" origami (substrates projected away from mica) to avoid artifacts, using procedures that minimize readout bias.

3.2 Materials and Methods

ABBREVIATIONS

iSp18 is a hexa-ethyleneglycol internal spacer; Bio is biotin; and BioTEG is biotin-tetra-ethyleneglycol.

PREPARATION OF SPIDERS

Materials and Instrumentation for the Preparation and Characterization of NICK3.4A+1 and $NICK_{3.4A+1}$*• (Cy3)*$_3$**.** Synthesis and purification of the modified DNA strands used to construct **NICK3.4A+1** and **NICK$_{3.4A+1}$• (Cy3)$_3$** were carried out by Integrated DNA Technologies

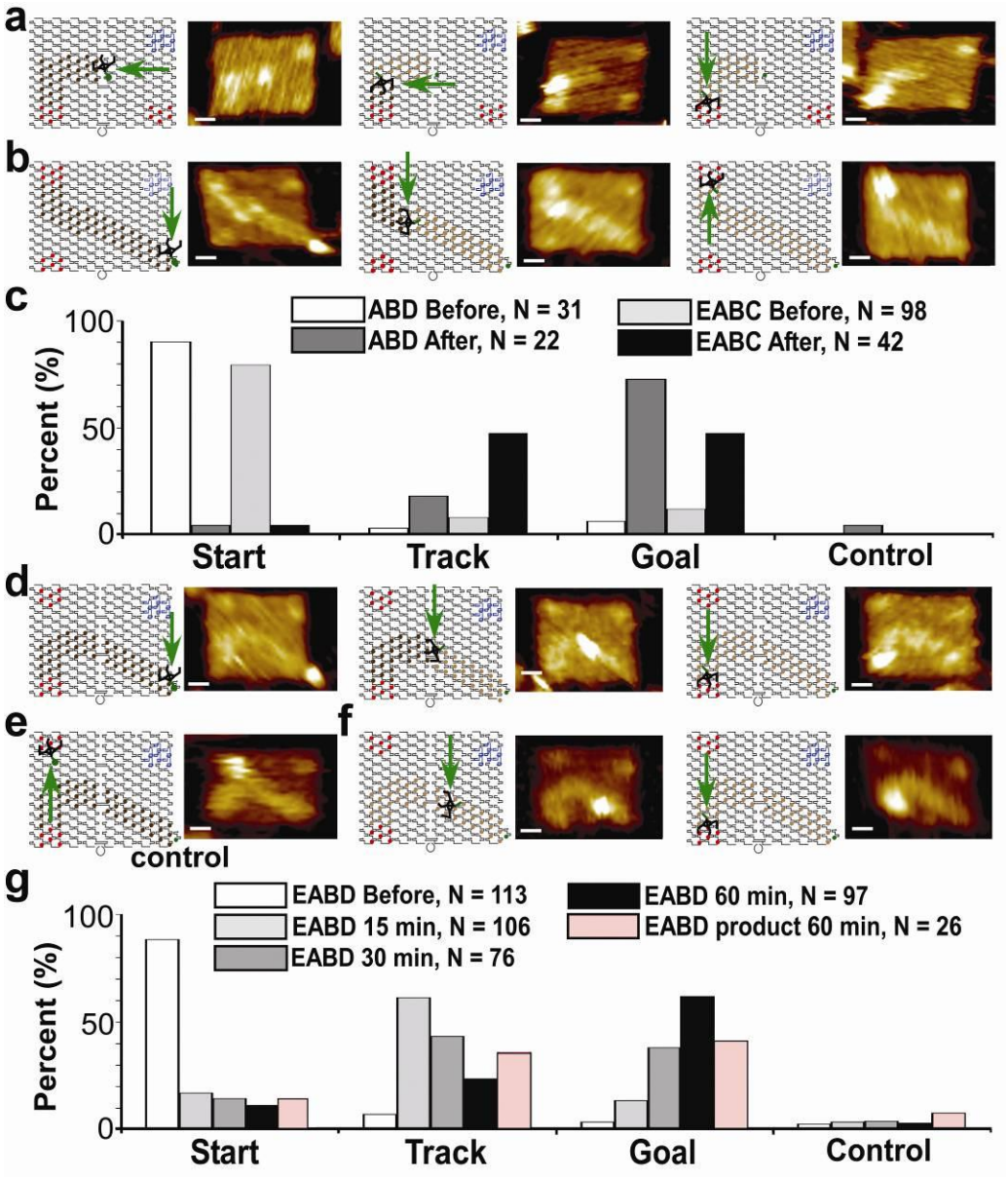


Figure 3.2 Results of spider movement along three tracks with schematics and AFM images of the spider at the start, on the track, and at the stop site. a, ABD track. **b**, EABC track. **c**, Graph of ABD and EABC spider statistics before and 30 minutes after release. **d**, EABD track. **e**, EABD track with spider on control. **f**, EABD product-only track. **g**, Graph of the EABD spider statistics before, and 15, 30 and 60 min after release, and 60 min after release on the EABD product-only track. All AFM images are 144 x 99.7 nm, the scale bar is 20 nm. Legend text indicates the number of origami with a single spider that were counted for the given sample.

(Coralville, IA) and used as received. Streptavidin was obtained from Pierce, product number 21125 (Rockford, IL). IE-HPLC purification was performed using a Shimadzu LC-6AD pump equipped with a Shimadzu SPD-M10A PDA detector, with separation carried out on an anion exchange TSKgel DEAE-NPR column, 4.6x50 mm (IDxL) (Tosoh Biosciences). Concentrations of oligonucleotides were determined on an Amersham Biosciences Ultrospec 3300 pro UV/visible spectrophotometer.

*Assembly of NICK3.4A+1***.** *Part A*; capture leg [5′ - GCC GAG AAC CTG ACG CAA GT/iSp18//iSp18//3Bio/ - 3′] (**C**) (47 nmoles in 10 mL of 10 mM HEPES,150 mM NaCl, pH 7.4) was added drop-wise to a stirred solution of streptavidin (STV) (5 mg, 94 nmoles in 1 mL of 10 mM K3PO4,pH6.5). The desired one-to-one conjugate product ("STV-(**C**)1") was purified by ion exchange (IE) HPLC (Figure D.1 for details of purification conditions). *Part B*; deoxyribozyme leg [5′ - /5BioTEG//iSp18//iSp18/TCT CTT CTC CGA GCC GGT CGA AAT AGT GAA AA - 3′] (**L**) (100 μM, in water) was titrated into the isolated 1:1 conjugate HPLC fraction from 'Part A' above, until all three remaining biotin binding sites of the 1:1 conjugate "STV-(**C**)1" were occupied by **L** to give the final desired product "STV-(**C**)1(**L**)3" i.e. **NICK3.4A+1**. The titration was monitored by IE-HPLC, and was deemed complete when a slight excess of **L** was observed with no intermediate species, i.e. no "STV-(**C**)1(**L**)1" or "STV-(**C**)1(**L**)2", present, see Figure D.2. The assembly was purified by IE-HPLC (see Figure D.2 for details of purification conditions) and the volume of the elutent reduced (by centrifugation) to give a final concentration of 2.3 μM, as determined by absorbance at 260 nm. Characterization of the assembly was carried out by IE-HPLC and PAGE (Figures D.1, 2, 3 and 4). The assembly was stable at -20 °C for at least six months.

*Assembly of NICK$_{3.4A+1}$• (Cy3)$_{3}$**. Part A and part B* were carried out in identical fashion to the assembly of **NICK3.4A+1** above, except "(C)" was [5′ – /5Cy3/GCC GAG AAC CTG ACG CAA GT/iSp18//iSp18//3Bio/ - 3′] and triethanolamine (20 mM) was used in place of HEPES and TRIS for the assembly and HPLC purification respectively. *Part C;* the volume of **NICK$_{3.4A+1}$• (Cy3)$_{1}$**, fraction isolated by HPLC, was concentrated to 1 mL (0.834 nmoles) and Cy3 Mono NHS ester (20 nmoles) (PA13101, Lot number 359269, GE Healthcare) dissolved in DMSO added to the solution containing the assembly (giving a total DMSO concentration of 10%). The resulting mixture was incubated at room temperature overnight, protected from light. Excess dye was separated from the **NICK$_{3.4A+1}$• (Cy3)$_{3}$** product by gel filtration (PD-10 column, 17-0851-01, lot 367770, GE Healthcare). Ratio of dye to streptavidin-DNA assembly was obtained by determining concentrations at 550 nm (ε_{max} 150,000 $M^{-1}cm^{-1}$) and 260 nm (ε_{max} 1,220,000 $M^{-1}cm^{-1}$) respectively.

** The number of Cy3 dyes per spider is an average. This particular protocol sometimes produced an average of four Cy3 dyes per spider molecule, hence such spiders will be notated in the text as **NICK$_{3.4A+1}$• (Cy3)$_{4}$**.

SURFACE PLASMON RESONANCE (SPR)

*Materials and Instrumentation for SPR Experiments***.** Immunopure avidin was purchased from Pierce (Rockford, USA). We used a Biacore X system and commercially available Biacore SA sensor chips, and Biacore C1 sensor chips, from GE Healthcare (Piscataway, USA). 1× HBS buffer (10 mM HEPES, pH7.4 with 150 mM NaCl) was employed as running buffer.

Preparation of Substrates on pseudo-2D Hydrogel Matrix Surfaces for SPR. A 20 μM solution of cleavable substrates (5′-BioTEG-TTTTTTTTCACTATrAGGAAGAG, "r" precedes a ribonucleotide) was applied to both channels of the SA sensor chip (carboxymethylated dextran matrix pre-immobilized with streptavidin) for 16 min at 5 μL/min, followed by a 60 s wash with 4 M urea and 15 mM EDTA in both channels to remove any nonspecifically adsorbed materials. The quantity of substrates adsorbed was calculated by the change in measured mass as described[24].

Preparation of Substrates on 2D Monolayer Surfaces for SPR. Avidin was covalently bound to the C1 sensor chip surface (a carboxymethylated monolayer) via amino groups using the following protocol. The carboxymethylated surface was first activated at a flow rate of 5 μL/min by using a 7 min injection pulse of an aqueous solution containing N-hydroxysuccinimide (NHS, 0.05 M) and N-ethyl-N`-(dimethylaminopropyl) carbodiimide (EDC, 0.2 M). Next, an 80 μL injection of 1 mg/mL avidin (in 1× HBS) was flowed over the activated surfaces of both channels for 40 min at 2 μL/min. The remaining activated sites on the chip surfaces were blocked with a 35 μL injection of an ethanolamine hydrochloride solution (1 M, pH 8.5). Then, a 20 μM solution of cleavable substrate was applied to both channels of C1 sensor chip for 20 min at 4 μL/min, followed by a 60 s wash with 4 M urea and 15 mM EDTA. Based on the average SPR responses for avidin (~2,010 RU, 0.03 pmole/mm^2) and substrate (450 RU, 0.056 pmole/mm^2), there are two substrates bound for each avidin molecule. The average intersubstrate distance is 5.5 nm.

SPR Monitoring of Dissociation of NICK$_{3.4A+1}$ Spider on Non-cleavable Substrate and Product Surfaces. The non-cleavable substrate analog (substrate in which rA was substituted with A) or product surfaces were prepared in a similar manner to the preparation of substrate on 2D monolayer surfaces. The spider was loaded to channel 2, with channel 1 serving as a negative control. We calculated the ratio of spider to non-cleavable substrate or product by measuring the change in SPR response units (RU) after the spider was flowed onto the chip, then used the equation: ratio (spider/S or P)= M_w (S or P) × RU(spider)/[M_w (spider) × RU(S or P)] (Figure D.5). Monitoring the dissociation of the spider was performed in 1× TA-Mg buffer (40 mM Tris, 20 mM acetic acid, 12.5 mM Magnesium acetate) with 1 mM $ZnCl_2$.

We could not directly measure the dissociation rate of spiders from cleavable substrate because 1) dissociation of the cleavage product from the surface accounts for the vast majority of the SPR response, and 2) the ratio of substrate to cleavage product changes with time, so the dissociation rate of spiders is not constant. Therefore, we instead monitored the SPR response to obtain the dissociation rate of spider on non-cleavable substrate, and product. We observed that over the course of 30 min >92% of spiders remained on a product covered surface and over the course of 60 min 86% remained bound (Figure D.5). These percentages represent an upper-bound on spider dissociation from our tracks (which will be a mixture of substrates and products as the spider walks over it). So we estimate an upper-bound for the dissociation rate as less than 8-14 % over the time scale of our experiments on AFM and fluorescence microscopy.

SPR Monitoring of Cleavage of Substrates by NICK$_{3.4A+1}$ Spider. Spiders (0.8~6.3 nM in 1× HBS buffer) were loaded only on channel 2 at 5 μL/min, with channel 1 used as a negative control. The amount of spider applied was controlled by adjusting concentrations and the reaction times

of spiders in the loading solution. Monitoring the cleavage of the substrate was initiated by switching to 1× TA-Mg buffer with 1 mM $ZnCl_2$ or 1× HBS buffer with 1 mM $ZnCl_2$ with the Biacore X system 'Working Tools Wash'. Product formation in real time was measured through the decrease in mass, using the formula 1,000 RU = 1 $ng \cdot mm^{-2}$. Rates of cleavage were determined from the approximately linear region of the product release curves during the initial 10% of substrates cleaved. On the 2D monolayer surface, real-time processivity of spiders was measured to be ~79% (percentage of total substrate cleaved over the course of the experiment) at a 1:291 ratio of spider (17.8 RU) to substrate (448.4 RU) with a cleavage rate of 1.42 min^{-1} per spider. On the pseudo-2D matrix surface, spiders showed a real-time processivity ~86% of total substrate cleaved at a 1:990 ratio of spider (26 RU) to substrate (2,222 RU) with a cleavage rate of 2.81 min^{-1} per spider (Figure D.6).

PREPARATION OF SPIDER-ORIGAMI ARRAYS

Assembly of Spider-Origami Arrays for Atomic Force Microscopy (AFM). The spider arrays (Figures D.7-24) consist of M13mp18 viral DNA (New England Biolabs) and 202 ssDNA staples (Integrated DNA Technologies, see Figure D.7 for DNA sequences). The arrays were annealed in 1× TA-Mg Buffer (40 mM Tris, 20 mM acetic acid, 12.5 mM Mg^{2+}, pH 7.6) using a 1:3 ratio of M13 to staple strands and a final concentration of 10 nM (M13). The arrays were annealed in two hours from 94 ^{o}C to 25 ^{o}C using an Eppendorf PCR machine (Eppendorf). The **NICK3.4A+1 or NICK$_{3.4A+1}$• (Cy3)$_3$** were then added to the arrays at a 1:1 ratio of START strand to spider and left at room temp overnight. Because origami folding is sensitive to stoichiometry, we expect that some fraction of origami are missing the START strand and are thus unable to position a spider before the TRACK is deposited. The substrate strand and

CONTROL strand were then added at a 1:1 (for initial ABD, EABC and Before EABD samples) or 1:3 (for 15, 30 and 60 minute EABD samples) ratio of staple probes to substrate or CONTROL and allowed to bind overnight at room temperature (20 ^{o}C to 24 ^{o}C). We observed (by AFM) a larger percentage of apparently unbroken TRACKS when excess substrate was added. In the presence of excess substrate there is a low probability that a spider leg may bind to a free floating substrate or STOP strand that would deter or inhibit interactions with the TRACK. Note that the 8-17 deoxyribozyme has reduced but non-negligible activity in TA-Mg buffer (relative to maximal activity with Zn^{2+}; see PAGE Activity Assays, below), suggesting that spiders bound at START may cleave immediately neighboring substrates during the overnight incubation. Since spiders undergo (unbiased) walks on product tracks with little dissociation, this possibility is not a concern. To minimize stacking interactions that can cause aggregation of origami, the staples on the left and right edges of the origami were removed. Schematics of the assembled origami landscapes are shown in Figures D.8, 11, 14, and 17.

Modification of Spider-Origami Arrays for Fluorescence Microscopy. To make the origami arrays compatible with fluorescence microscopy, we returned 4 of the removed staples to the corners of the origami. In order to affix the origami to slides for analysis, we divided the corresponding staples into two strands so that we could affix biotin labels onto the 5′ end that is antiparallel to staple probes (as in Figures D.11b, 14b and 17b). We modified the CONTROL strand by adding a Cy5 fluorophore to its 3' end, which resulted in 6 Cy5 fluorophores labeling the STOP position. On all landscapes, CONTROL staples were replaced with staples lacking the non-cleave-able substrate probes. The EAC landscape used in both fluorescence microscopy and AFM experiments lacked a CONTROL site. In addition, the EAC arrays for fluorescence

microscopy were annealed in 5× SSC buffer (75 mM sodium citrate, pH 7.0, 750 mM NaCl), and the EABC and EABD arrays in 1× TA-Mg buffer. Fluorescence microscopy was also performed for origami arrays containing a truncated substrate TRACK, or product TRACK. The product strand is 8 nucleotides shorter than the full length substrate and includes only the sequence 5′ of the RNA base. The resulting 31 oligonucleotides have the same sequence as the corresponding portion of the full length cleavable substrate. All other assembly details for origami arrays for fluorescence microscopy including DNA concentrations, relative strand ratios, and binding conditions were unchanged.

ATOMIC FORCE MICROSCOPY

AFM Imaging. "Before" samples were deposited on mica without the addition of TRIGGER or $ZnCl_2$. "After" samples were prepared by releasing the spider from the START strand through the addition of a 27-base TRIGGER strand, immediately followed by the addition of 10mM $ZnCl_2$ to a final concentration of 1 mM. Spiders were allowed to traverse the product or substrate TRACK array in solution for 15, 30, or 60 min (depending on the experiment) at room temperature before the origami were deposited on mica. Samples (2 μL) were deposited onto a freshly cleaved mica surface (Ted Pella, Inc.) and left to adsorb for 3 min. Buffer (1× TA-Mg, 400 μL) was added to the liquid cell and the sample was scanned in tapping mode on a Pico-Plus AFM (Molecular Imaging, Agilent Technologies) with NP-S tips (Veeco, Inc.). Each sample was scanned for 2-3 hrs before being discarded (therefore "30 minutes after" means that the sample spent 30 minutes in solution followed by up to 3 hours on mica). Note that the reduced but non-negligible deoxyribozyme cleavage rate in TA-Mg raises the possibility that spiders could move during the this imaging period; however, given the apparent difficulty of spider movement on mica-bound origami even in the presence of Zn^{2+} (see AFM Imaging for Movie) and the

consistent trends in the time-lapse experiments (Figure 3.2), we conclude that very little movement takes place during the imaging period. All imaging by AFM was carried out at room temperature.

AFM Imaging for Movie. The sample (2 μL) was deposited onto a freshly cleaved mica surface and left to bind for 2 min. Then 1 μL of TRIGGER strand was added to the sample on the surface and after 2 min 270 μL of buffer and 30 μL of 10mM $ZnCl_2$ was added to the sample cell. The four images were taken over a 26-minute time frame with about 10 min between the saving of each scan. (It should be noted that many prior and subsequent attempts were made to capture another AFM movie using various optimizations of our buffer, and protocol, without success.) Although we were only able to capture one movie, reported in Figure 3.3, we are convinced that it is not an artifact. The origami with the moving spider is substrate face-up while the three origami in the same image are substrate side down (see below for a discussion of how the face of the origami affects spider analysis). As a result spiders on the three adjacent origami are stationary over the time course of the movie. In addition the spider's motion follows the TRACK in each frame (therefore it is not randomly diffusing, because it neither moves backwards nor off the TRACK). If the AFM tip were merely pushing the spider forward we would not expect the spider to turn in the transition from frame 3 to frame 4.

AFM Time Lapse Experiments. There is one seeming contradiction in our report that we would like to address here. If we were to suggest (as we do in Figure 3.3) that the spider can walk on origami deposited on mica, then how could we expect to obtain viable statistics from time lapse experiments imaged for up to 3 hours? We assume that under these conditions, most spiders get

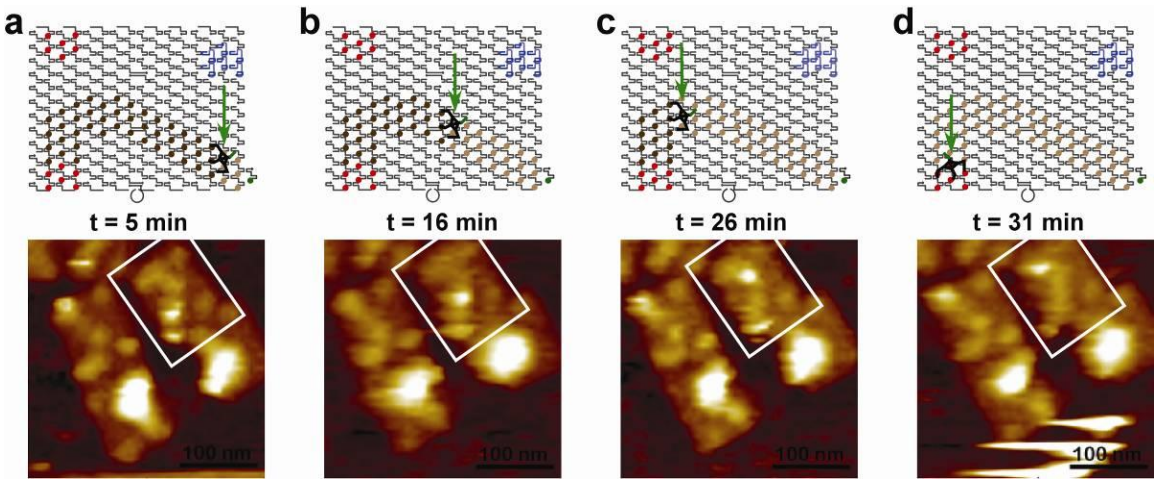


Figure 3.3 AFM movie of spider movement. a, b, c, d, Schematics and AFM images of the spider moving along the EABD track at 5 min (**a**), 16 min (**b**), 26 min, (**c**) and 31 min (**d**) after trigger was added. AFM images are 300 x 300 nm and the scale bar is 100 nm.

stuck on the origami, while some small percentage of spiders are able to continue moving. We find that we can differentiate between samples deposited at 15 minutes from those deposited at 30 and 60 minutes (Table D.1). These results help to explain why obtaining the AFM movie was so difficult.

Statistical Analysis of AFM images. We divided our flattened AFM images into 1 x 1 um images and numbered them. Within each of these images, we assigned a roman character to each origami (thus each origami we analyzed could be uniquely identified by a number and letter (i.e. "EABD Before 1e", or "EABD 30 min 3a"; Figures D.9, 10, 12, 13, 15, 16, 18, 19). The origami arrays were classified by the following criteria: orientation (is the origami "face-up" or "face-down"?), number of spiders (0,1, multiple), location of spiders (START, TRACK, STOP, CONTROL), image quality (do imaging errors or sample impurities make the classification difficult?). This process was conducted independently by three people, for each data set excluding the EABD 15 minute and EABD 60 minute data sets, which were conducted by two people. The classifications were then compared: if two or more people agreed on the origami classification it was held, otherwise the origami was discarded from further analysis. By this method, we sought to ensure that our results are neither subjective nor irreproducible. While it is possible that some putative spiders were actually image artifacts or molecular contaminants, it is unlikely that this inaccuracy in our measurements could affect the main trends in our data or the qualitative conclusions we drew from them.

An origami that is "face up" is one that displays its substrates and spiders on the face opposite the mica; an origami that is "face down" displays its substrates on the face that rests on the mica (Figures D.20 and D.21). Orientation was determined by landscape asymmetries in the

positions of the TRACK and marker. By analyzing the statistics of origami classification, we concluded that the probability of an origami landing on one face or the other was approximately equal. However, we discovered that "face down" origami appeared to have a larger number of spiders at the STOP. We conducted a double-blind study in which 6 researchers were given an AFM image of origami and asked to classify these according to our criteria. We discovered that in the absence of spiders, all "face-up" origami were classified as vacant while a significant portion of "face-down" origami were classified as displaying a spider at the STOP site, when in fact there was none (Figures D.22). Due to this "false positive" effect, we did not count "face-down" origami in our statistics. Approximately 50% of "face-up" origami were unoccupied by any spiders, and between 0 and 7% displayed more than one spider on the TRACK. Because the quantity of multiply occupied origami was small compared to the quantity of unoccupied and singly-occupied origami, we only considered singly-occupied origami to simplify our analysis (Figure 3.2).

Experimental results for all four landscapes with substrate TRACKS showed that the fraction of spiders at the START diminishes with a concomitant increase in spiders observed on the STOP positions (Figure 3.2c,g, and Figure D.23). Our shortest track (ABD, spanning 48 nm) efficiently delivers spiders to the STOP, with less than 20% of spiders on the TRACK after 30 min (Figure 3.2c). If the TRACK was omitted on the ABD landscape, spiders were equally distributed between the STOP and CONTROL sites after 30 min, implying that the track is needed for efficient delivery to the STOP site (Figure D.24). On longer TRACKS (such as EABD, spanning ~ 90 nm) ~15% of spiders are delivered to the STOP within 15 min after release. Longer incubation times (30 and 60 min) increase the efficacy of delivering spiders to the STOP to up to 70%, (Figure 3.2c,g). Even at 60 min, however, we observed between 10-15%

of spiders still on the TRACK. This outcome could be attributed to the distribution of spider velocities resulting from the stochastic nature of individual walks and possibly from backward steps onto product, initiating an unbiased random walk on product. We observed no significant difference in the efficacy of "turn right" and "turn left" actions (paths EABD and EABC, respectively) 30 min after release (Figure 3.2c,g).

PAGE ACTIVITY ASSAYS OF $NICK_{3.4A+1}$• $(Cy3)_4$.

The cleavage activity of spiders under various conditions in bulk solution was tested as follows. Reactions were initiated by combining **$NICK_{3.4A+1}$• $(Cy3)_4$** (34 nM) with 4A substrate (5'-/5bio//iSp18//iSp18//iSp18//TTT TTT TTT TTC ACT AT(rA) GGA AGA G-Cy5, 34 nM) in the presence of either 1× SSC (15 mM sodium citrate (Mallinckrodt Inc.), pH 7.0, 150 mM NaCl) or 1× TA-Mg, and 0, 1, 2 or 10 mM $ZnSO_4$ (all reported concentrations are final). All reactions were supplemented with an oxygen scavenger system (1× OSS) consisting of 25 nM protocatechuate dioxygenase, 2.5 mM protocatechuate, and 1 mM Trolox as described[25]. Reactions (10 μL) were quenched after 0, 5, or 30 minutes with 2.5 μL of 0.25 M EDTA, and characterized by denaturing PAGE (Figure D.25a). Fluorescence from Cy5 and Cy3 was detected on a Typhoon 9410 Variable Mode Imager (Amersham Biosciences) and the fraction of cleaved substrate quantified in ImageQuant 5.2 (Molecular Dynamics). Substrate was cleaved at least five-fold more slowly in 1× SSC + 1 mM Zn^{2+} than in 1× TA-Mg + 1 mM Zn^{2+}, while cleavage in 1× SSC + 10 mM Zn^{2+} was only about two-fold slower than in 1× TA-Mg + 1mM Zn^{2+} (Figure D.25b). The maximal extent of cleavage is also about 9-fold lower in SSC + 1 mM Zn^{2+} than in TA-Mg + 1 mM Zn^{2+}, consistent with a significant fraction of inactive deoxyribozyme-substrate complexes. This discrepancy among buffers is likely due to partial

complexation of Zn^{2+} ions by citrate: from a direct Zn^{2+} concentration measurement in buffer using the low-affinity (30 μM) indicator dye Newport Green PDX (Molecular Probes) we estimate the free Zn^{2+} concentration in SSC buffer to be approximately 3-fold lower than in TA-Mg buffer at 1 mM total Zn^{2+}. Nevertheless, these assays demonstrate that spiders are active under the buffer conditions used in Single Molecule Fluorescence Microscopy imaging (see below). A limiting factor for increasing the Zn^{2+} concentration above 1 mM is slow spontaneous $Zn(OH)_2$ precipitation at the near-neutral pH used in our studies. We therefore varied the buffer conditions in our Single Molecule Fluorescence Microscopy imaging experiments between 1× SSC with 0-10 mM $ZnSO_4$, 1× HBS buffer (10 mM HEPES, 150 mM NaCl, pH 7.4) with 0-5 mM $ZnSO_4$, and 1× TA-Mg, carefully monitoring (and avoiding) any $Zn(OH)_2$ precipitation.

Further studies have examined the buffer-dependence of 8-17 spider leg cleavage activity in **NICK3.4A+1**; cleavage rates varied from 0.25-1.5 min^{-1} in TA-Mg, TA-Mg with 1 mM Zn^{2+}, and HEPES with 1 mM Zn^{2+} (Taylor, Pei, Stojanovic, unpublished results). In particular, the non-negligible cleavage rate in TA-Mg with no Zn^{2+} has implications for the AFM experiments prior to adding TRIGGER, as discussed above. Finally, these solution-based cleavage assays and the SPR assays (discussed above and in Figure D.6), while useful for detection of cleavage activity under various conditions, may not be in quantitative agreement with the cleavage rate at the surface of an origami tile, where the locally high density of substrates and other surface effects may have a large impact on the rate-limiting step of this reaction.

SUPER-RESOLUTION PARTICLE TRACKING WITH FLUORESCENCE MICROSCOPY

Overview. For more facile real-time observation of the movement of individual spiders along tracks, we applied super-resolution imaging by total internal reflection fluorescence (TIRF)

video microscopy[26]. Four biotin molecules were attached to the underside of the origami to facilitate its immobilization on the avidin-conjugated quartz slide. Experiments were performed using EAC, EABC, and EABD tracks. Spiders were covalently labeled with on average 2.3 Cy3 fluorophores (λ_{ex} 568 nm), and the STOP position was labeled with 6 Cy5 fluorophores (λ_{ex} 672 nm). This labeling scheme allowed us to colocalize spider position relative to its STOP using two-color single-molecule high-resolution colocalization (SHREC)[27] and monitor their relative movement by single-particle tracking[28]. In a typical experiment, spider-loaded tracks were incubated with TRIGGER in the absence of Zn^{2+} ions and then immobilized on the slide (Figure D.26). Within 20 min of commencing fluorescence imaging, we added Zn^{2+} to promote spider movement via substrate cleavage. As the 8-17 deoxyribozyme's activity depends sensitively on buffer conditions[29], we optimized our conditions for a combination of best catalytic activity and SMFM imaging quality, obtaining best results from SSC or HEPES with increased Zn^{2+} concentrations and no Mg^{2+} (see below and Figure D.6 and D.25). The position of a spider on its origami path relative to the START was extracted over time by fitting the diffraction-limited point-spread functions (PSFs) to two-dimensional Gaussians in an up to 80-min sequence of wide-field images (time resolution 15-30 s) with a precision (standard deviation) of 10-30 nm. We controlled for focal drift and developed a consistent set of criteria to distinguish moving spiders from stationary ones as detailed below and in Figures D.27 and D.28.

Preparation of Avidin-Coated Microscope Slides. Two 1-mm holes were drilled in each microscope slide (fused silica) to allow for buffer exchange. The slides were immersed in boiling "piranha" solution (5% (v/v) ammonium hydroxide, 14% (v/v) hydrogen peroxide) for at least 20 min, then sonicated for 30 min in 1 M KOH, and flamed for several seconds with a

propane torch. The slides were then aminosilanized by immersing them in a 5% (v/v) solution of 3-aminopropyltriethoxysilane (Sigma-Aldrich) in acetone for 1 h, rinsed with acetone, and dried for 1 h at 80 °C. A layer of the bifunctional crosslinking agent *para*-phenylene diisothiocyanate (PDITC) was covalently coupled to the aminosilanized surface by incubating the slides for 2 h in a solution of 0.2 % (w/v) PDITC, 10% (v/v) pyridine in *N*,*N*-dimethylformamide (spectroscopic grade). The slides were rinsed thoroughly with methanol and acetone, and 70 μL of 0.5 mg/mL avidin (Sigma-Aldrich) was added to each slide, covered with a glass coverslip (VWR), and allowed to incubate for 2 h at room temperature in a closed container above a water bath to avoid drying out. The coverslips were removed, and the slides were washed thoroughly with deionized water, followed by 1 M NaCl plus 40 mM NaOH, and again washed with deionized water, then dried under nitrogen. A flow channel about 2-3 mm wide was made between the drilled holes with two strips of double-sided tape, a coverslip was placed on the tape, and the edges were sealed with Epoxy glue (Hardman Adhesives). A schematic of the resulting surface structure is shown in Figure D.26. Slides were stored in an evacuated desiccator at 4 °C for up to four weeks.

Fluorescence Microscopy. For the EAC, EABC, and EABD tracks**,** spider-origami complexes at 10 nM in the annealing buffer were combined with an equal volume of 1 μM to 10 μM TRIGGER strand in water and incubated for 30-60 min on ice in the absence of Zn^{2+} ions. The mixture was then diluted to 10 pM in the imaging buffer: for the EAC track, 1× SSC or HBS; for EABC and EABD tracks,1× TA-Mg. (Note that for the EABC and EABD tracks, the reduced but non-negligible deoxyribozyme cleavage rate in TA-Mg raises the concern that pre-incubation with TRIGGER may allow some spider movement prior to imaging; however, the reduced

temperature would be expected to inhibit such movement.) All buffers used for fluorescence imaging were supplemented with 1×-5× OSS to reduce the rate of signal loss through fluorophore photobleaching. The spider-origami complexes were immobilized on avidin-coated microscope slides for imaging.

Samples were imaged at room temperature by a prism-based total internal reflection fluorescence microscope with a 1.2 NA 60× water objective (IX71, Olympus). Cy3 excitation was provided by a 532-nm green laser (ultra-compact diode-pumped Nd:YAG laser GCL-025-S, CrystaLaser) and Cy5 excitation by a 638-nm red diode laser (Coherent CUBE 635-25C, Coherent Inc.). The Cy3 and Cy5 emission signals were separated by a dichroic mirror with a cutoff of 610 nm (Chroma) and projected side by side onto an ICCD camera chip (iPentamax HQ Gen III, Roper Scientific, Inc.). Relay lenses matched the microscope image with the camera focal plane and the IX71 internal 1.6x magnifier (final effective pixel length 133 nm) was used during collection of all traces except EAC 1 and 2 (Figure D.29) in which no magnifier was used (effective pixel length 196 nm). The donor channel image was passed through a band pass filter (HQ580/60m, Chroma) and the acceptor channel was passed through a long pass filter (HQ655LP, Chroma). A cleanup filter (z640/20, Chroma) was placed at the output of the red laser to reject any extraneous or infrared light. A Newport ST-UT2 vibration isolation table was used in all experiments. After introducing imaging buffer without oxygen scavenger to the slide flow channel, a small fluorescent background was observed; this was bleached briefly by exposing the slide to excitation light from both lasers until the background stabilized. The origami sample with oxygen scavenger was then introduced into the sample channel in the dark, allowed to incubate for 2-10 minutes, and the excess flushed out with fresh imaging buffer. The sample was imaged at room temperature with excitation from both lasers using a 1- to 2.5-s

signal integration time and a 12.5- to 27.5-s delay (i.e., 2-4 frames per min). After 0-20 min of imaging, depending on the experiment, the same imaging buffer containing or lacking $ZnSO_4$ was introduced into the flow channel, and the sample was imaged for an additional 60-70 min. For substrate-covered EAC tracks in SSC, the concentration of $ZnSO_4$ introduced was either 0 or between 1 and 10 mM. For substrate-covered EABD and EABC tracks, the $ZnSO_4$ concentration was 0 or 1 mM $ZnSO_4$. For substrate-covered EAC tracks in 1× HBS, the $ZnSO_4$ concentration was 0 or 5 mM.

EAC track origami with product-covered tracks were prepared and imaged as described above for the substrate-covered EAC track origami in SSC buffer. Due to concerns about releasing spiders from the START prematurely on product tracks where the walk is independent of cleavage activity, experiments were also conducted in which the spider-origami assemblies were not incubated with TRIGGER 30-60 min prior to imaging, as described above, but instead SSC imaging buffer containing 1 mM $ZnSO_4$ and 10 μM TRIGGER was added to the sample channel 10-15 min before imaging. In both types of experiments, $ZnSO_4$ was not introduced until immediately prior to imaging by fluorescence microscopy.

Fitting and Filtering of Particle Tracking Data. Point spread functions (PSFs) of fluorescence emission from individual spiders and origami were imaged by fluorescence microscopy, and their relative positions tracked through time by fitting 2-D Gaussian functions to the PSFs. First, PSFs from Cy3 (spider) and Cy5 (origami) were imaged on spectrally separated halves of the ICCD camera using WinView32 software (Roper Scientific, Inc.). PSFs were identified in the ICCD output and paired with their corresponding partner using methods described previously[30], resulting in intensity traces such as those shown in Figures D.29 and D.30 that reflect the total

photon count per movie frame for each PSF over time. The Cy3 and Cy5 channels were registered with a locally weighted mean mapping[27] using fluorescent beads that appear in both channels (Fluospheres, red fluorescent (580/605), 0.2 □m, Molecular Probes FluoSpheres F8801), to establish with ~50-nm accuracy that the Cy3 PSF in each pair was located within 200 nm of its Cy5 partner. To ensure adequate signal intensity and duration for tracking, candidate PSF pairs were only included in the analysis if they met both of the following criteria:

1. Cy3 and Cy5 signal of more than 1,000 photon counts per frame for at least 25 min (1-33% of all pairs fulfilled this criterion per experiment)
2. No erratic signal intensities such as from excessive blinking or nearby unresolved PSFs (23-95% of all remaining pairs fulfilled this criterion per experiment)

Traces that were discarded based on low or absent signal intensity from either Cy3 or Cy5 likely resulted from incompletely labeled spider-origami complexes, fragmented or disassembled complexes, or other fluorescent contaminants. Each PSF in the remaining pairs (0.4-22% of all candidate pairs) was fit, frame-by-frame, with a two-dimensional Gaussian function (Figure D.27d) of the form:

$$f(x,y) = z_0 + A\exp\left\{\frac{1}{2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\} \qquad (1)$$

The position values μ_x and μ_y from Gaussian fitting of each Cy5 PSF (Figure D.27 e, g) were subtracted, frame-by-frame, from those of its corresponding Cy3 partner. The resulting difference trajectory was then plotted against time for each Cy3-Cy5 pair to show the motion of each spider relative to its Cy5-labeled STOP position (Figure D.27 f, h). This subtraction served as a necessary internal drift control since, as shown in Figure D.27, there was often significant

drift through the x-y plane in the course of a typical 30-80 min movie. Brief aberrant position measurements, such as those caused by transient binding of nearby fluorescent contaminants, were identified by a large distance from the median position (> 3 standard deviations in the x or y direction) or sudden displacements of >100 nm within a single frame, and removed. Focal drift throughout an experiment, if severe, sometimes resulted in an apparent motion of Cy3 relative to Cy5 (data not shown). This focal drift was evident visually from the original video image as well as from very asymmetric PSF shapes during Gaussian fitting. Such traces were also discarded.

Probable moving spiders were selected using the following criteria:

1. Relative motion of Cy3 and Cy5 > 45 nm, corresponding to 2-3 times the standard deviation in individual position measurements (33-44% of all fitted pairs fulfilled this criterion per experiment)
2. No discontinuities in position, i.e., sudden jumps in position of 45 nm or greater (89-100% of all fitted pairs fulfilled this criterion per experiment)
3. Apparent movement < 45 nm prior to zinc addition (88-100% of all fitted pairs fulfilled this criterion per experiment)

This process is illustrated in Figure D.28 for representative experiments from the EAC, EABC, and EABD constructs. The resulting spiders (22-39% of all fitted traces) are included in Figure D.29. Examples of spider trajectories that did not satisfy all three of these criteria are also shown in Figure D.29 (EAC Tier 2) and 30 (all stationary spiders observed on the EAC track in HBS). A statistical summary of this filtering process for the EAC, EABD, and EABC tracks is presented in Table D.3.

Representation of Spider Trajectories. To smooth the trajectories for presentation, a 16-frame rolling average was applied separately to the trajectories of Cy3 and Cy5 before subtracting them for drift correction (black line in Figures D.29b and D.30b). The error bars shown in Figures D.29b and D.30b are the standard deviations of the raw trajectory from the temporally corresponding points in the smoothed trajectory. For ease of viewing in Fig. 3.4a and Figure D.31, the trajectories were instead smoothed with an 8-frame rolling average followed by a 4-frame sequential average before drift correction.

Measurement of Displacement. Net displacement was determined as follows for motion of each spider on the EAC track. An initial position (x_0, y_0) was defined as the arithmetic mean of the first 16 position measurements after the time t_{zinc} at which $ZnSO_4$ or control buffer lacking zinc ions was added (t = 0 min in Figures D.29 and D.30). For traces containing data prior to t_{zinc}, the initial position was instead calculated as the mean of the 16 position measurements centered on t_{zinc} (i.e., the interval from frame -7 to frame 8, where t_{zinc} occurs at frame 0). The center time coordinate of this averaged initial position (x_{start}, y_{start}) was designated t_{start} (i.e., the interval from frame -7 to frame 8, where t_{zinc} occurs at frame 0). The distance of (x_{start}, y_{start}) from each subsequent position measurement (x_i, y_i) was then calculated to obtain the spider's net displacement over time (green line, Figures D.29c and D.30c). As has been noted in similar distance determinations[27], these displacement measurements are artificially increased when equal to or less than the noise level (hence why displacement typically does not equal 0 nm near t = 0 min). Therefore, an analogous displacement vs. time curve was calculated from the smoothed trajectory (black line in Figures D.29b and D.30b) and was plotted as a black line in Figures D.29c and D.30c. This smoothed displacement has a value of zero at t_{start}, resulting in a

systematic deviation from the noise-inflated raw curve at low displacements. The time of stopping t_{stop} was defined as the time coordinate of the first local maximum in the smoothed displacement curve that approaches within 20 nm of the global maximum in the smoothed displacement curve (considering only the interval from t_{zinc} to the end of the trace). The value of 20 nm is a typical standard deviation in our position measurements. The total net displacement d (inset box, Figures D.29b and D.30b) was then defined as the smoothed displacement value at t_{stop}. The time of travel Δt was defined as the difference $t_{stop} - t_{start}$, and the mean magnitude of velocity was calculated for each EAC spider as $v = d/\Delta t$ (box, Figures D.29b and D.30b). The resulting displacement vs. time plots are shown in Figures D.29c and D.30c.

Interpretation. In some traces we observed movement before addition of Zn^{2+}; we could not determine whether these represented spiders walking in the absence of Zn^{2+} or were due to other causes. We also observed several moving traces that exhibit net displacements significantly smaller than others, which similarly is consistent with spiders having finished (part of) their tracks early, taking the wrong direction after walking in the absence of Zn^{2+}, prematurely stopping or stalling on the track, and/or taking backward steps onto product. These issues are discussed in more detail below. In the following, we enumerate all independent lines of evidence that these time traces represent genuine walking spiders:

1. The highest density of PSFs we observed in each channel with $\geq$ 1,000 photon counts over at least 25 minutes was 0.03 □m-2. Given this density, the probability that a Cy3 and Cy5 PSF will colocalize to within 200 nm of one another by coincidence is 0.9%31. However, in each experiment we observe that, on average, 31% of PSFs in one channel

are colocalized with a PSF in the other channel. This strongly suggests that the majority of signals originate from spiders bound to origami.

2. We find most of the trajectories longer than 45 nm to be consistent in length and shape with a progressive walk on the respective track design (Figures D.28 and D.29b). In particular the trajectories observed on the EAC track in SSC buffer, which are nearly linear and often stop nearly 100 nm from the starting position (EAC 1, 2, 4, 5, 9, 12, 13, and 15), are in good agreement with expectations based on the track design.
3. Comparison of experiments to negative controls (such as in Figures D.28d and D.30), rules out instrument drift as the sole source of the observed spider motion.
4. Moving and non-moving spiders are seen alongside each other in experiments conducted in the presence of both Zn^{2+} ion and release strand (Figure D.28a-c), providing further fiduciary markers and a strong argument against instrument drift as the cause for movement.
5. Ensemble MSD (Figure D.32a) and RMSD plots (Figure 3.4c) of the 15 Tier 1 EAC spiders (Figure D.29) are consistent with an approximately 100-nm walk across the prescribed linear substrate track.

Especially when considered in combination with the results from our AFM studies, the fluorescence microscopy data are most consistent with processive walking of individual spiders on DNA origami. While the stopping distances are not strong evidence (filtering precluded walks shorter than 45nm, and photobleaching may have precluded having many walks longer than 100nm), this interpretation is confirmed by control experiments lacking zinc in the buffer and on product tracks, as discussed below.

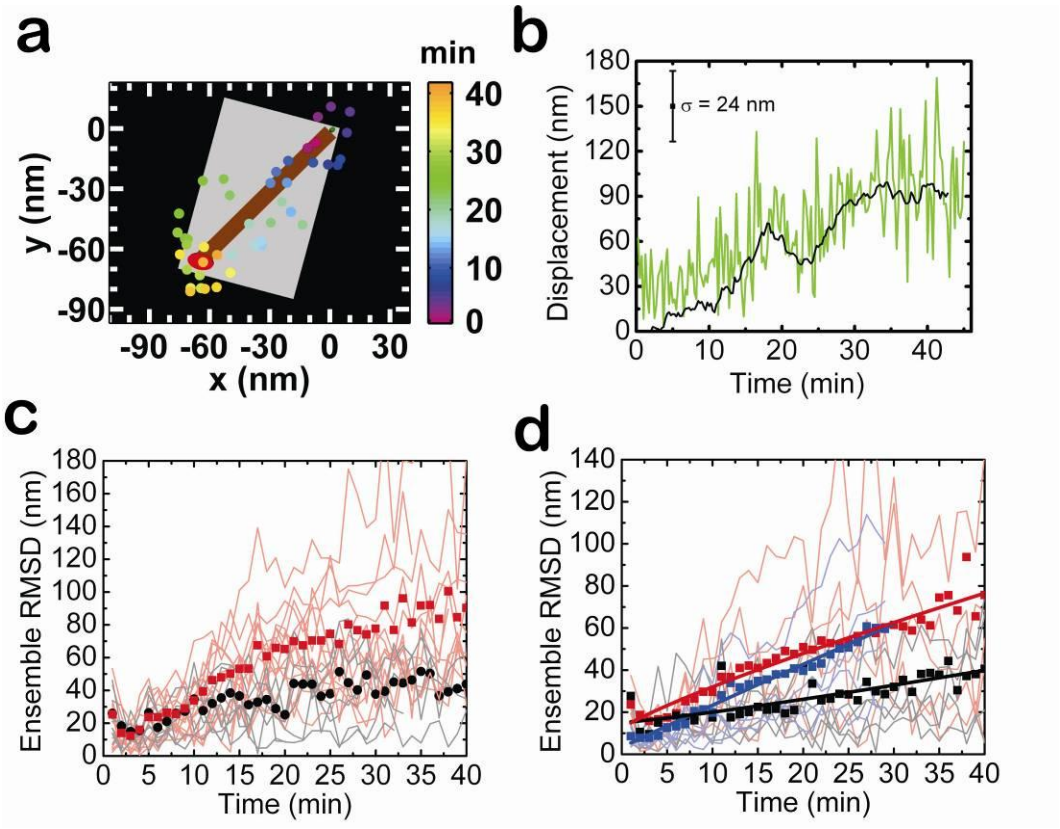


Figure 3.4 Spiders imaged on origami tracks in real-time using super-resolution TIRF microscopy. a, Position-time trajectory of a selected spider (EAC 2, Cy3-labeled) on the EAC substrate track. The position as a function of time is represented by color-coded dots. A small green dot represents the START and a large red oval represents the Cy5-labeled STOP site. $ZnSO_4$ was added at time zero. **b,** Displacement of the spider trajectory in panel **a** from its initial position as a function of time. The green line represents displacement calculated using averaged position measurements of 1 min intervals, and the black line represents the displacement from a rolling 4-min average. **c**, Ensemble root mean square displacement (RMSD) of exemplary spiders on the EAC substrate track in the presence (red, corresponding to the 15 Tier 1 Spiders in Figure D.29) and absence (black, 7 spiders) of Zn^{2+}, with the corresponding displacements used to calculate each ensemble RMSD for each buffer condition (similarly colored line graphs). **d,** Ensemble RMSD for spiders on EAC tracks satisfying simple filtering criteria. Curves are shown for spiders on EAC substrate track (red, 85 spiders), EAC product track with TRIGGER introduced to the sample 10-15 min before imaging (blue, 18 spiders), and EAC product track with TRIGGER introduced 30-60 min before imaging (black, 29 spiders). EAC substrate and 10-15 min trigger product RMSD plots are fit to a power law function, and the EAC 30-60 min trigger product RMSD is fit to a straight line. Individual displacements are shown with colors corresponding to the respective ensemble RMSD plots. All Figure 3.4 data were obtained in SSC buffer.

The large percent of spiders moving less than 45 nm (22-67 % of all PSF pairs fit to Gaussians in a given experiment) likely results from some combination of the following: 1) immobile contaminants that fluoresce in both channels, thus having the appearance of a colocalized Cy3-Cy5 pair; 2) a substantial fraction of inactive or slowly cleaving spider legs, especially in SSC + 1 mM Zn^{2+}, 3) failed or delayed release of a spider from the START position, 4) spiders binding initially at the STOP instead of the START position (though precautions against this were taken in the assembly of origami-spider complexes), or 5) undirected, random diffusion of a spider on previously cleaved or damaged substrate. As we cannot distinguish between these possibilities, the estimated percent of non-moving spiders must be taken as an upper bound.

Most of the trajectories from the EAC track show clearly biased, generally linear motion with few or no discontinuities in displacement (Figure D.28). However, some spiders exhibit non-monotonically increasing displacements with time (e.g. EAC 5) that could have resulted, for example, from spiders taking steps backwards onto cleavage product. Furthermore, some trajectories exhibit unexplained irregularities in the 2-D motion trajectory, displacement or velocity measurements (Figure D.29, EAC 16-19). These issues are described in the captions above the respective traces in Figure D.29. For instance, the net displacement values of EAC 16 and 18 at $t > 0$ are less than 45 nm and therefore less reliable. Some putative spiders (EAC 17, 18, and 19) also show significant displacement before addition of Zn^{2+} at $t = 0$. Slight focal drift or an instrumental perturbation might have resulted in an apparent displacement between the PSFs[32] in these traces, particularly during the addition of Zn^{2+}-containing buffer. Although this addition was performed slowly and carefully (generally at a rate of 1 mL/min or less), it occasionally brought the image out of focus. Such slight focal drift could also affect the

measurement of net displacement values. For example, similar influences could have given rise to the few trajectories that showed an apparent net displacement larger than the track length of 110 nm (e.g., EAC 10). Alternatively, although the 8-17 deoxyribozyme legs are inactive in 1× SSC lacking $ZnSO_4$ (see Figure D.25), the spider might still exhibit slow diffusion on a surface of substrate. It is also possible that some origami assemblies exhibit rotational dynamics relative to the slide that contribute to the observed motion of PSFs. Finally, the calculation of net displacement for some spiders is likely biased by early photobleaching which may prevent observation of the entire trajectory of the spider (see, for example, EAC 3).

Calculation of Ensemble Mean Square Displacement and Root Mean Square Displacement. To characterize the ensemble behavior of spiders, ensemble mean square displacement (MSD, Figure D.32) and root mean square displacement (RMSD, Figure 3.4c-d) versus time plots were generated. To calculate the individual displacements plotted in figure 3.4c-d, and used to calculate the MSD and RMSD, an initial position (x_{start},y_{start}) was first calculated as the arithmetic mean of the 16 points of the raw trajectory closest to the time at which $ZnSO_4$ or control buffer lacking zinc ions was added, t_{zinc} (i.e., the interval from frame -7 to frame 8 if data were taken before t_{zinc}, or the interval from frame 1 to frame 16, if data acquisition began at t_{zinc}). Trajectory data were averaged separately for Cy3 and Cy5 in (sequential) one-minute intervals, and the averaged trajectory of Cy5 was subtracted from that of Cy3 to correct for microscope stage drift. Each displacement value was then calculated as the distance of each averaged position (x_i, y_i) from the initial position (x_{start},y_{start}). This same procedure was applied to data acquired at the rate of two, three, and four frames per minute. The displacement of each spider for each 1-minute time interval was squared and then averaged across all spiders within a

given dataset to yield the ensemble MSD as a function of time. The square root of the ensemble MSD was calculated for each time interval to yield the ensemble RMSD as a function of time. Note that outliers were removed from the raw data as described above (Filtering and Fitting of Particle Tracking Data).

For comparison with Tier 1 EAC spiders, MSD and RMSD versus time plots were also generated from the 7 EAC spiders in a no-zinc control experiment on the EAC substrate track in 1× SSC. These spiders were subjected to the same selection criteria as the Tier 1 EAC spiders except that they were not required to move $\geq$45 nm for inclusion in the MSD plot (by this criterion, no moving spiders were observed in this control). Both of these MSD plots are shown in Figure D.32a, and the RMSD plots shown in Fig. 3.4c.

In an attempt to determine the relative impact of substrate cleavage on the motion of EAC spiders (compared to that of the presence of START and STOP sites), control experiments were conducted on EAC tracks covered with cleavage product instead of substrate in 1× SSC and 1 mM $ZnSO_4$. For consistency, identical experimental procedures were applied, including addition of Zn^{2+} immediately prior to imaging (although product walks are not expected to be zinc-dependent). To reduce the risk of bias in comparing these two types of experiments, we employed a less stringent set of selection criteria than those described above. Specifically, all spider trajectories with Cy3 and Cy5 signal intensity above an arbitrary cutoff were retained. Individual data points in a trajectory were discarded if the ellipticity E exceeded 0.3 ($E = 1 - w_{minor}/w_{major}$, where w_{minor} and w_{major} are the full widths at half maximum along the major and minor axes of the fitted 2-D Gaussian function, respectively). Position measurements greater than three standard deviations from the median of all position measurements within a trace in either the x- or y-direction (or 500 nm from the position of the spider when zinc was

added, whichever is smaller) were regarded as outliers and discarded. An application of these more inclusive criteria first to our substrate-covered track data resulted in 85 traces that were converted to the ensemble MSD and RMSD versus time plots described above; the results are shown in Figure D.32b and Figure 3.4d (see Table D.4 for full statistics of selection based on these criteria). The roughly twofold difference in steepness from and less pronounced curvature than the substrate MSD plot in Figure D.32a are likely due to the inclusion of a larger number of slow-moving or stationary spiders in Figure D.32b.

We found that the ensemble MSD versus time plot for the product-covered linear EAC track in 1× SSC and 1 mM $ZnSO_4$, generated using the more inclusive selection criteria above, dramatically depends on whether the TRIGGER was added 10-15 min (short incubation) or 30-60 min before the experiment (long incubation, similar to our protocol for the linear substrate track). In the former case, the MSD plot of 18 spiders increases non-linearly with a concave up slope curvature greater than that seen for the substrate track, while in the latter case, linear behavior with a much shallower slope is observed in an MSD plot of 29 spiders (Figure D.32b). Since Zn^{2+} is not predicted to be required for diffusive walking on a product surface, a long pre-incubation with TRIGGER is expected to allow many spiders to prematurely walk and possibly be captured by the STOP site prior to the onset of imaging, resulting in a much lower net displacement over the time window of observation. However, when the TRIGGER is added 10-15 min before the experiment, the spider release from the START position may become rate-limiting to effect an initial delay followed by Brownian diffusion of the released spiders along the track. This possibility prevents a direct comparison of the MSD plot of the latter experiment with that of the linear substrate track in Figure D.32b (see also discussion of Monte Carlo simulations below). We therefore conclude that we cannot distinguish the behavior of spiders on

substrate- and product-covered tracks with confidence from these experiments except insofar as they respond differently to pre-incubation with the TRIGGER.

As an additional control, MSD versus time plots (Figure D.32c) were created for the EAC spiders in HBS buffer shown in Figure D.30. The MSD plot begins with the addition of HBS buffer containing 0 mM (EAC 1-21H) or 5 mM (EAC 1-16HZ) $ZnSO_4$. As for the MSD of the Tier 1 EAC substrate track spiders observed in 1× SSC, the presence of zinc increases the slope of the MSD versus time plot for spiders in 1× HBS, suggesting that the movement of spiders on the EAC substrate track is zinc-dependent in these buffers.

Monte Carlo simulations of spiders on EAC track. To aid in the interpretation of our experimental results, Monte Carlo simulations of simplified models of spiders walking on EAC tracks were conducted as follows. The spider consists of three legs, each of which can exist in an unbound state or bind a specific substrate or product within a 2-dimensional array based on the EAC track dimensions. The three legs are constrained to bind substrates within 10 nm (an estimated effective leg span) of all other bound legs, and can bind any such substrate with equal probability as long as that substrate is not already bound by another leg. The spider's body position is taken as the arithmetic mean of the positions of all legs bound to the substrate array.

At each time step of the simulation, each leg acts independently to perform one or more of the following actions:

- If bound to a substrate, it can cleave it or not.
- If it is bound to a substrate or product, it can dissociate or remain bound.
- If it is unbound, it binds a substrate or product within 10 nm of other bound legs within the same timestep of the simulation.

Each of the first two of these actions has an associated probability P_i that can be related to an effective first-order rate constant k_i according to $P_i = (1-\exp(-k_i*t))$, where t is the length of a timestep, chosen here as 1 second. There are thus three adjustable probability parameters: the probability of cleaving a bound substrate (P_{cleave}), of dissociating from a bound substrate ($P_{off,substrate}$), and of dissociating from a bound product ($P_{off,product}$). The legspan is a fourth adjustable parameter. Note that, for simplicity, it is assumed that hybridization to a new site is instantaneous compared to a timestep, and independent of whether that substrate has been cleaved.

At the beginning of each simulation, the spider is positioned with all three legs bound to substrates (or products) within 10 nm of one end of the track corresponding to the START position. At the opposite end of the track are six non-cleavable substrates which constitute the STOP site. The spider's legs are then allowed to freely cleave, dissociate, and bind substrates and products. The legs must remain within 10 nm of the START end until the spider is released by a TRIGGER event which occurs with a probability $P_{release}$. Each simulation ran for 35 min (2100 time steps).

Probabilities for cleavage of and dissociation from substrates were determined from effective first-order rate constants: k_{cleave} = 1 min^{-1}, $k_{off,substrate}$ = 0.002 s^{-1}, $k_{off,product}$ = 0.2 s^{-1}. These rate constants are within one order of magnitude of those determined by bulk fluorescence experiments in 10 mM HEPES, 150 mM NaCl and 1 mM ZnCl2, pH 7.4 (data not shown), preserve the experimentally observed ratio $k_{off,substrate}/k_{off,product}$, and are compatible with sound principles of nucleic acid thermodynamics and kinetics[33].

Ensemble MSD versus time curves (Figure D.32d) were calculated based on the current spider's body position relative to its first observed position, and are an average over either 20, 80

(thin lines), or 1,000 (thick lines) simulated spider trajectories. When allowed to walk on a linear EAC track containing cleavable substrates, simulated spiders yielded an MSD plot with positive (concave up) curvature, similar to the slightly positive curvature seen in the experimental plots (Figure D.32a-c). In contrast, when walking on an EAC track with cleavage product, the spiders yield an MSD curve that first increases linearly, then asymptotically approaches a maximum value corresponding to the STOP position (spiders undergoing an undirected walk are still able to bind and become trapped at the non-cleavable STOP). However, positive MSD curvature is also obtained from a product track simulation if the spider is released from the START with a half-life of 10 min (rather than immediately), so we cannot rule out a lag phase as contributing to the positive curvature of the experimentally observed ensemble MSD versus time plots. Furthermore, simulations of only 20-80 spiders yielded a fairly broad range of MSD behaviors, often obscuring the idealized curvature and slope; this suggests that our experimental MSD plots, constructed from 85 or fewer spiders, may not represent the fully converged behavior of the system.

The qualitative features of MSD versus time plots generated from these simulations, such as curvature (linear or concave-up) and maximal extent of increase, are robust to variations in k_{off}, and k_{cleave} at least one order of magnitude about their experimentally observed values. However, the precise values of the parameters can affect the slope of the MSD versus time plots. The relative slopes of product and substrate walks are quite sensitive to the effective legspan parameter. Furthermore, the introduction of unequal association probabilities for substrate and product can affect the slope and curvature of these plots. Still, these simulations show that the observed ensemble MSD versus time behavior for the linear EAC substrate track is consistent

with the proposed mechanism of spider locomotion based on reasonable kinetic parameters, even if other mechanisms cannot be conclusively ruled out.

3.3 Results

In all samples imaged before spider release, 30-40% of the assembled origami carry at least one spider, 80-95% of which are singly occupied, and of these 80-90% bound their spider at the START position (Table D.1 and Figure D.9, 10, 12, 13, 15, 16, 18, and 19). Upon adding TRIGGER, all four landscapes with substrate tracks showed that the fraction of spiders at the START diminishes with a concomitant increase in spiders observed on the STOP sites (Figure 3.2c,g and Figure D.23). A spider's ability to reach the STOP sites decreased with increased TRACK length and with decreased time of incubation in solution. In time-lapse experiments on a long path (EABD, spanning ~ 90 nm) we observed a gradual increase of up to 70% of spiders on STOP sites within 60 min (Figure 3.2c,g). A short path (ABD, ~ 48 nm) was completed to the same extent within 30 min.

We captured a series of AFM images of a spider moving along an origami track (Figure 3.3). The rate of spider movement (~90 nm over 30 min, with approximately 6 nm per three parallel cleavage events) was consistent with the processive cleavage rates (~1 min^{-1}) of spiders on a two-dimensional surface as obtained by surface plasmon resonance (SPR; Figure D.6). More systematic sequential imaging proved difficult due to mica's inhibitory effects on the spider.

To confirm that spiders can indeed traverse product tracks by means of an unbiased random walk, we tested spiders using EABD origami in which the substrate was replaced by product on the TRACK. Spiders still reached the STOP sites, albeit more slowly (Figure 3.2f,g),

as expected from purely Brownian spider movement even if individual steps are somewhat faster[19].

If all three legs simultaneously dissociate before any leg reattaches, a spider could "jump" by completely dissociating from the origami and subsequently reattaching elsewhere at random. Evidence against frequent jumping (or an excess of spiders in solution during the initial assembly stage) comes from the low level of spider occupancy at CONTROL sites in both substrate and product track experiments (Figure 3.2c,e,g) and the stable proportions of unoccupied and multiply-occupied origami (Table D.1; both before and after the addition of trigger, 5-10% of origami displayed more than one spider on its track). In contrast, when spiders were released on ABD landscapes with no TRACK strands, after 30 min we observed an equal distribution between STOP and CONTROL sites (Figure D.24 and Table D.2), as expected for a process that involves spider dissociation from, and random rebinding to, the origami.

In independent ensemble experiments using SPR to monitor spider attachment and with a constant flow passing over the surface, we observed that up to 15% of spiders dissociated from a non-origami, two-dimensional product-covered surface within 60 min (Figure D.5). When using similar surfaces but covered with substrate, spiders showed an average processivity of ~200 substrates before being removed by flow (Figures D.5 and D.6). Together, these results rule out that spiders move predominantly by jumping; there is insufficient jumping even on product tracks to explain the 50-70% occupation of the STOP sites after walks on ABD, EABC, and EABD substrate tracks.

For a more facile real-time observation of the movement of individual spiders, we applied particle tracking by super-resolution total-internal-reflection fluorescence (TIRF) video microscopy[26]. Four biotin molecules were attached to the underside of the origami for

immobilization on the avidin-coated quartz slide. Spiders were covalently labeled with on average 2.3 Cy3 fluorophores, and STOP sites were labeled with 6 Cy5 fluorophores. The labeling allowed us to monitor changes in spider position relative to the STOP site by two-color fluorescent particle tracking[34,35]. In a typical experiment, spider-loaded tracks were incubated with TRIGGER and immobilized on the slide (Figure D.26), and then Zn^{2+} was added to promote spider movement by means of substrate cleavage. Recognizing that the activity of the 8-17 DNA enzyme depends on buffer conditions[23], we obtained the best results from SSC or HEPES with increased Zn^{2+} concentrations but without Mg^{2+} (Figure D.6 and D.25).

Our resolution was not sufficient to detect turns reliably, so we focused on EAC landscapes. Individual particle traces showed a distribution of behaviors that may result from variations across molecules, idiosyncrasies of the sample preparation, the stochastic nature of the observed process, photobleaching, and/or instrument measurement error (Figure 3.4a,b, Figures D.29-31, Table D.3). Despite this variability, traces of moving particles commonly showed net displacements between 60 and 140 nm and their mean velocity varied between 1 and 6 nm/min; within error, these values are consistent with the track length (~90 nm) and the deoxyribozyme cleavage rate (~1 min^{-1} per leg), respectively.

Tests with and without Zn^{2+} and/or TRIGGER, both on substrate and product tracks, yielded root mean squared (r.m.s.) displacement plots of the particle traces that in each case varied as expected on the basis of the behavior of spiders on origai tracks, despite the inherent noise associated with single-particle tracking over tens-of-nanometer length scales and tens-of-minute time scales (Figure 3.4c,d). For instance, r.m.s. displacement plots indicated substantially more movement on substrate tracks in the presence of Zn^{2+} and trigger than in their individual absence (Figure 3.4c, Figures D.30-32 and Table D.4). On product tracks, results

were consistent with an unbiased random walk, independent of the presence of Zn^{2+}. When product tracks were pre-incubated with TRIGGER 30-60 min prior to addition of Zn^{2+} and the onset of imaging (as were substrate tracks), little or no movement was observed (Figures 3.4d), consistent with spiders having been released and having moved toward or reaching the STOP sites prior to imaging. In contrast, when TRIGGER and Zn^{2+} were both added shortly prior to imaging, substantial movement was observed (Figure 3.4d), consistent with our AFM results for spiders on product tracks (Figure 3.2f,g) and with Monte Carlo simulations of spider movement (Figure D.32).

3.4 Discussion

The results of our single-molecule experiments are consistent with random DNA-based walkers guided by their landscapes over distances as great as 100 nm, for up to 50 cleavage steps, at speeds of roughly 3 nm/min. We note, however, that the distance over which a spider can move is limited by dissociation and backtracking, and any increase in processivity comes at the cost of a lower speed[12]. Other limitations arise from the current mechanism consuming substrate that must be recharged to sustain directed movement, and from spiders being subject to the stochastic uncertainty as to whether each individual robot can accomplish its task (compare with the notions of "faulty" behavior in robotics and "yield" in chemistry). Furthermore, in comparison with protein-based walkers using solution phase fuels[36], our walkers are not as fast, efficient, or powerful. As candidates for molecular robots, however, they offer the advantages of programmability[5,10,37-39], predictable biophysical behavior[5], and designable landscapes[13]. The ability to obtain programmed behavior from the interaction of simple molecular robots with a complex modifiable environment suggests that by exploiting stochastic local rules and

programming the environment, we can effectively circumvent the limitations that molecular construction places on the complexity of robotic behavior at the nanoscale.

3.5 Addendum: Further Characterization of the Spider Walking Mechanism

3.5.1 Introduction

The development and experimental results demonstrating nanorobots that generate their own bias attracted the attention of theorists who desired to learn more about the source and effectiveness of this biasing mechanism[40-45]. The models, however, involved over-simplified one-dimensional tracks and limited the legs to choosing their next location via a hand-over-hand or inchworm model. Furthermore, the rate constants used in the simulations were estimated based on values in the literature that did not use the same experimental conditions or same nucleic acid sequences, and the simulations were never directly compared with experimental results. Here, we use experimentally determined rate constants in Monte Carlo simulations to further illuminate the walking mechanism of the spider. We also introduce two novel experimental techniques developed to provide an efficient means for obtaining a statistically significant amount of data to experimentally verify the results of these simulations.

3.5.2 Materials and Methods

All oligonucleotides were obtained from IDT unless otherwise specified.

Monte Carlo Simulations Based on the Gillespie Algorithm. Home-programmed Monte Carlo simulations were used to model the walking behavior of the spiders along the track based on experimentally determined rate constants. The user specified the number of spiders, the spider's leg span, the time for which the spider walks, the time resolution, and the kinetic rate constants that impact the spider walking mechanism: $k_{bind,sub}$, $k_{bind,prod}$, k_{cleave}, $k_{dissoc,sub}$, and $k_{dissoc,\ prod}$ (see

Figure 3.5a). The dimensions of the origami track in the model were determined by the distance between the extended staple strands along the origami used to form the track. The three legs were initially placed on three substrates at the beginning of the track. The body was modeled to be located at the average of the three positions. The time interval in which the next reaction occurs and the type of reaction (leg binding to substrate, leg binding to product, cleavage, leg dissociating from substrate, and leg dissociating from product) that will occur was determined by the Gillespie Algorithm[46,47]. Once the reaction was chosen, a leg was determined at random to undergo the reaction. If the reaction was for the leg to bind to either a substrate or a product, the substrate or product was selected at random and tested to see if it was within the leg span of the spider. If not, another was selected. If no substrate or product was within the leg span, a new reaction was selected at random. If all three legs have dissociated without another binding, the time at which the spider dissociates is noted and the program terminates (no spiders have fully dissociated with the current parameters). After the spiders' trajectories were simulated, ensemble MSD plots were generated.

Streptavidin-coated Biotinylated-PEGylated Slide Preparation. Five slides were prepared at a time. Each slide contained two holes to act as outlets for fluids. For cleaning, slides were boiled in water for 30 min to remove the epoxy resin and double-sided tape used to adhere the coverslip. Slides were then sonicated in alconox for 1 h. Slides were rinsed thoroughly in water to remove alconox and sonicated in 1 M KOH for 20 min. They were left in a base piranha solution (80 mL water, 20 mL ammonium hydroxide, 20 mL hydrogen peroxide) for 30 min, rinsed with water, and flamed using a propane torch.

The slides were silanized in a nonaqueous environment using a 5% (v/v) solution of 3-aminopropyltriethoxysilane (Sigma-Aldrich) in acetone. The solution was mixed thoroughly initially and then mixed again halfway through the one-hour incubation. Slides were rinsed in water and dried with nitrogen. The slides were placed in a humid environment, namely a closed pipette box with water under the pipette tip holder, to ensure they would not dry during the next incubation step. 21 mg of biotin-PEG-SVA(Laysan Bio Inc) and 70 mg mPEG-Succinimidyl Valerate (Laysan Bio Inc) were dissolved in 320 µL 100 mM sodium bicarbonate, a buffer prepared immediately before use. The solution was placed in a centrifuge for 1 min at 10,000 rpm and then in a SpeedVac for 10 min to remove oxygen. 70 µL of the solution was placed on each slide. A coverslip was placed on top of the solution and allowed to incubate for 3 h or overnight.

Slides were rinsed with water and dried with nitrogen, and placed back on the pipette tip holder. 10 mg Disulfosuccinimidyltartrate (DST; Soltec Ventures) was dissolved in 350 µL 1 M sodium bicarbonate prepared immediately before use. The solution was centrifuged for 1 min at 10,000 rpm and placed in the SpeedVac for 10 min. 70 µL of the solution was placed on each slide. A coverslip was placed on top of the solution, and the DST solution incubated on the slide for 30 min.

The slides were rinsed thoroughly with water and dried with nitrogen. Two pieces of double-sided Scotch tape were placed on the slide to create two boundaries for the fluidic channel between the two holes. A coverslip was placed on top of the tape and epoxy resin was applied along all the sides of the coverslip to adhere the coverslip to the slide surface. These slides were stored at room temperature in a dark and dry location.

Immediately prior to running an experiment, the slides were hydrated with T50 buffer (10 mM Tris-HCl, 50 mM NaCl, pH 8) followed by 1 mg/mL biotinylated-BSA for 10 min. After the biotinylated-BSA was flushed out with T50 buffer, 0.2 mg/mL streptavidin was added to the slide surface to incubate for 10 min. T50 buffer was then flushed through followed by 1XHBS (10 mM HEPES, 150 mM NaCl, pH 7.4).

Binding and Dissociation Rate Constant Determination. Binding rate constants to both product and substrates were determined with the biotinylated-origami immobilized on streptavidin-coated PEG surfaces. The origami tiles used for this assay were saturated with overhangs (187 strands). 300 pM origami tiles were incubated on the slide surface for 15 min. Excess was flushed out with 1XHBS, and either 20 nM Cy5-labeled substrate [5- GAT GAA TGG TGG GTG AGA GGT TTT TCA CTA TAG GAA GAG T/3Cy5Sp/] or 20 nM Cy5-labeled product [5- GAT GAA TGG TGG GTG AGA GGT TTT TCA CTA TAA /3Cy5Sp/ -3] was introduced to the flow channel for 30 min. OSS (5 nM protocatechuate (PCA), 50 nM protocatechuate dioxygenase (PCD), and 2 mM Trolox) was flowed through the channel. The slide was imaged using the prism-based TIRF setup, described earlier in this chapter, and the 638-nm laser. 1XHBS was flowed through the channel and the Cy5-labeled substrate or product strands on the origami were photobleached. During imaging, either 25 nM, 50 nM or 75 nM Cy3-labeled leg [5’ - /5Cy3/TCT CTT CTC CGA GCC GGT CGA AAT AGT GAA AA] in OSS was flowed onto the slide immediately after the second frame of the movie was collected (see next paragraph for imaging details), and the time course of the binding was determined by observing the increase in fluorescence intensity over time. For the dissociation rate, 50 nM Cy3-labeled leg was added and allowed to incubate for 30 min. After the second frame of the movie was

captured, 75 nM unlabeled leg in OSS was flowed through the channel, and the decrease of fluorescence intensity for each origami was monitored (Figure 3.6a).

For the real-time imaging of the binding and dissociation of the Cy3-labeled leg, a 1.0 ND filter was placed in front of a 50 mW 532-nm laser, and the MCP on the ICCD camera was set to 50. The slide was imaged with a 0.5-s integration time followed by a 29.5-s shuttered period, resulting in a temporal resolution of 30 s.

Data from the binding and dissociation rate constant experiments were analyzed using home-programmed software (MATLAB). The increase in intensity of each individual origami was normalized to each origami's peak intensity. The intensity of all the origami at each time point was then averaged together and fit to a single exponential.

Cleavage Rate Constant Determination. Eppendorf tubes were filled with 5 μL water and 3 μL 250 mM EDTA. For the multiple-turnover cleavage rate constant determinations, 5.5 μL of 1.21 μM Cy3-labeled spider was allowed to incubate with 5.5 μL 10 μM Cy5-labeled substrate for at least 5 min. For the single-turnover cleavage rate constant determinations, 5.5 μL 10 μM Cy5-labeled substrate incubated with 5.5 μL 50 μM Cy3-labeled leg for at least 5 min. 2X reaction buffer was prepared consisting of twice the concentration of the specified metal ion in 2XHBS. For the zero time-point, 1 μL of the 2X reaction buffer was added to the EDTA solution followed by 1 μL of either the labeled spider-substrate complex or the labeled leg-substrate complex. For the remaining time points, the spider-substrate or leg-substrate complexes were added to the 2X reaction buffer in a 1:1 ratio. The cleavage reaction was stopped by chelating the divalent metal ion at each time point of interest by removing 2 μL of the reaction volume and mixing it with the EDTA solution.

20% denaturing PAGE and a denaturing loading buffer were used to separate product from substrate and quantify cleavage. Each time point was run in a separate lane. The band containing product as well as the band containing substrate were quantified using ImageQuant (GE Healthcare) for each time point. The fraction cleaved was determined using the following equation: fraction cleaved = product/(product + substrate). The multiple-turnover rate constant was determined by taking the slope from fitting the first four data points of the fraction cleaved with a line (Figure 3.8d). The single-turnover rate was determined by fitting the fraction cleaved over time with an exponential.

Cleavage Rate Constant of Spiders Cleaving Track. 300 pM of the racetrack origami (see Appendix A) was introduced to the flow channel of a streptavidin-coated PEG slide for 15 min. The origami was flushed out with 1XHBS. 1 nM Cy3-labeled spider was flowed onto the slide and allowed to incubate for 5 min. Excess spider was flushed off with 1XHBS. 20 nM Cy5-labeled substrate in OSS was introduced to the fluidic channel. As the control, instead of the substrate, an all-DNA strand was introduced with the same sequence as the substrate, making it noncleavable because it lacks the RNA moiety necessary for cleavage to occur. After the second frame of imaging, 10 μM release strand and the specified metal ion in OSS were introduced into the fluidic channel.

The sample was illuminated using a 50 mW 532-nm green laser and 5 mW 638-nm red laser. A 0.04 ND filter was placed in front of both lasers. The MCP of the ICCD camera was set to 70. The illumination was shuttered with a 0.5-s integration time and 119.5 s shuttered period, resulting in 2 min temporal resolution.

3.5.3 Results

It was proposed that the bias in the spider's walk is due to the leg's greater affinity for substrate than product, primarily because the dissociation rate constant of the leg from the substrate ($k_{dissoc,sub}$) is lower than the dissociation rate constant of the leg from the product ($k_{dissoc,prod}$)[45]. To test this proposed walking mechanism, multiple models were implemented. The master-equation approach describes the imposed self-bias, but is limited to one dimension and can only describe ensemble-averaged behaviors, leaving no sense of the heterogeneity in the spider behavior[48,49]. The Monte Carlo simulation approach incorporates the stochastic, nondeterministic nature of the system into the model and permits analyzing a distribution of behaviors while still obtaining the same ensemble results as the master equation. In the published simulations, the binding rates for both leg binding to product ($k_{bind,prod}$) and leg binding to substrate ($k_{bind,sub}$) were either approximated to be infinite[44,45] or were assumed to be equal[40,41], a reasonable assumption given the binding rate constants were orders of magnitude faster than the other potential types of reactions and the difference between the length of the product and substrate was only seven nucleotides. However, unpublished ensemble experimental results determining the rate constants involved in the spider walking mechanism (Figure 3.5a) found that $k_{bind,prod}$ is 360-fold higher than $k_{bind,sub}$ [50] (Table 3.1). This indicates a greater affinity for product than was previously assumed. But $k_{dissoc,prod}$ was also found to be 100-fold higher than $k_{dissoc,sub}$. Therefore, while the difference between the two binding rate constants indicates a greater affinity for product than substrate, the dissociation rate constants indicate a greater affinity for substrate over product. The walking mechanism thus turned out to be more complicated than the previous models suggested. This required the development of a new model.

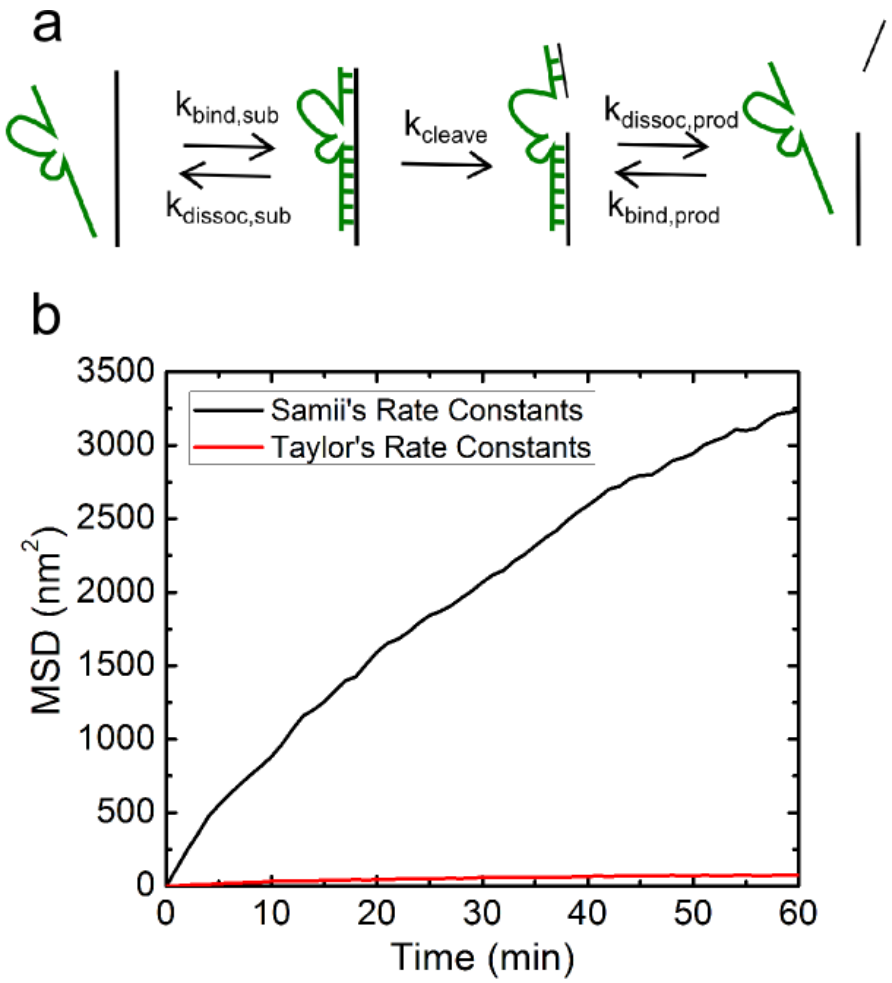


Figure 3.5 Monte Carlo Simulations. (**a**) The kinetic rate constants used as input in the Monte Carlo Simulations. (**b**) Ensemble MSD from simulated results generated from the trajectory of 1000 spiders walking along their own origami tracks using kinetic rate constants estimated by Samii[40,41] and Taylor's experimentally determined rate constants[50] as input (Table 3.1).

	$k_{bind,sub}$ ($M^{-1}s^{-1}$)	$k_{bind,prod}$ ($M^{-1}s^{-1}$)	$K_{dissoc,sub}$ (s^{-1})	$K_{dissoc,prod}$ (s^{-1})	k_{cleave} (s^{-1})
Samii (estimated) [40,41]	20	20	0.035	0.046	0.055
Taylor (1 mM $ZnSO_4$)[50]	2,200	790,000	0.0014	0.14	0.0058
Michelotti (10 mM $MgCl_2$)	98,118	359,316	0.0003	0.0017	0.0106

Table 3.1 Kinetic rate constants for the spider walking mechanism.

Monte Carlo simulations were used to determine the effect of this newly found difference in $k_{bind,sub}$ and $k_{bind,prod}$. The simulation incorporated the Gillespie Algorithm[46,47] to stochastically determine the next type of reaction to occur and the time at which it would occur. As opposed to assuming a strictly one-dimensional track, the simulation incorporated the geometry track of 3 substrates wide and 16 substrates long with each separated by either 5.8 or 3.75 nm, depending on the dimension. Using the rate constants estimated by Samii[40,41] as input for the simulation (Table 3.1), 52% of the spiders were able to successfully complete the track within 1 h. From the ensemble mean square displacement (MSD) plot, we obtained a diffusion coefficient of 47 nm^2/min (Figure 3.5b). The plot also indicated confinement (i.e., the alpha parameter was less than one) with an alpha parameter of 0.7. In the simulation, once the spider reaches the end of the track, motion ceases, because the origami track is designed to capture the spider with noncleavable strands complementary to the legs once it reaches the end of the track. Therefore, this confinement is most likely due to the spider being captured at the end of the track.

Using the rate constants determined by Taylor[50] as input for the simulation (Table 3.1), no spiders reached the end of the track within 1 h. The diffusion coefficient determined from the ensemble MSD plot was 2 nm^2/min (Figure 3.5b), over 20 times lower than the diffusion coefficient determined using Samii's estimated rate constants. The spider experienced much more confinement under these rate constants with an alpha parameter of 0.5. Because the spider never reaches the end of its track, there must be an alternative source of this confinement. One contribution to this confinement is the barrier imposed by the track, but this confinement is also present under Samii's kinetic rate constants. The other contribution is the high binding affinity to product. If it is more favorable for the spider to bind to product than substrate, as product is generated, it will be more likely to bind to the product instead of progressing toward the

substrate. Based on the results of these simulations, in order to improve spider mobility, the $k_{bind,sub}$ and $k_{bind,prod}$ would need to be made more comparable so that the bias becomes based solely on the dissociation rate constants. To this end, the source of the 100-fold difference between these two binding rate constants must first be determined.

To find the source of the discrepancy between $k_{bind,sub}$ and $k_{bind,prod}$, we sought to determine if this discrepancy existed under the experimental conditions used for the real-time fluorescence tracking. Specifically, because the real-time fluorescence tracking experiments were single-molecule experiments, they were performed at low concentrations (pM-nM) and on origami tiles immobilized on a slide surface. Taylor's assays for determining the rate constants, on the other hand, were performed in solution with higher concentrations[50]. These factors may impact the kinetic rate constants. With this in mind, the experiments to determine the kinetic rate constants were monitored using single-molecule fluorescence techniques to observe the increase in fluorescence as Cy3-labeled leg hybridized to substrate or product strands on origami tiles. The rate of this increase in fluorescence was used as an indicator of the fraction of leg bound to the surface at given time points and was used to determine the rate of binding and dissociation (Figure 3.6a).

This single-molecule method for determining the rate constants assumes a linear relationship between the fluorescence intensity and the number of legs bound to substrate. To test the validity of this assumption, the binding rate constant was determined using multiple ratios of labeled-to-unlabeled substrate (Figure 3.6b). The technique had a 14% uncertainty in determining the rate constant using different ratios of labeled-to-unlabeled. The rate constant determined did not exhibit an overall trend that was dependent on the ratio of labeled-to-unlabeled substrate. Therefore, if there does not appear to be a nonlinear relationship between

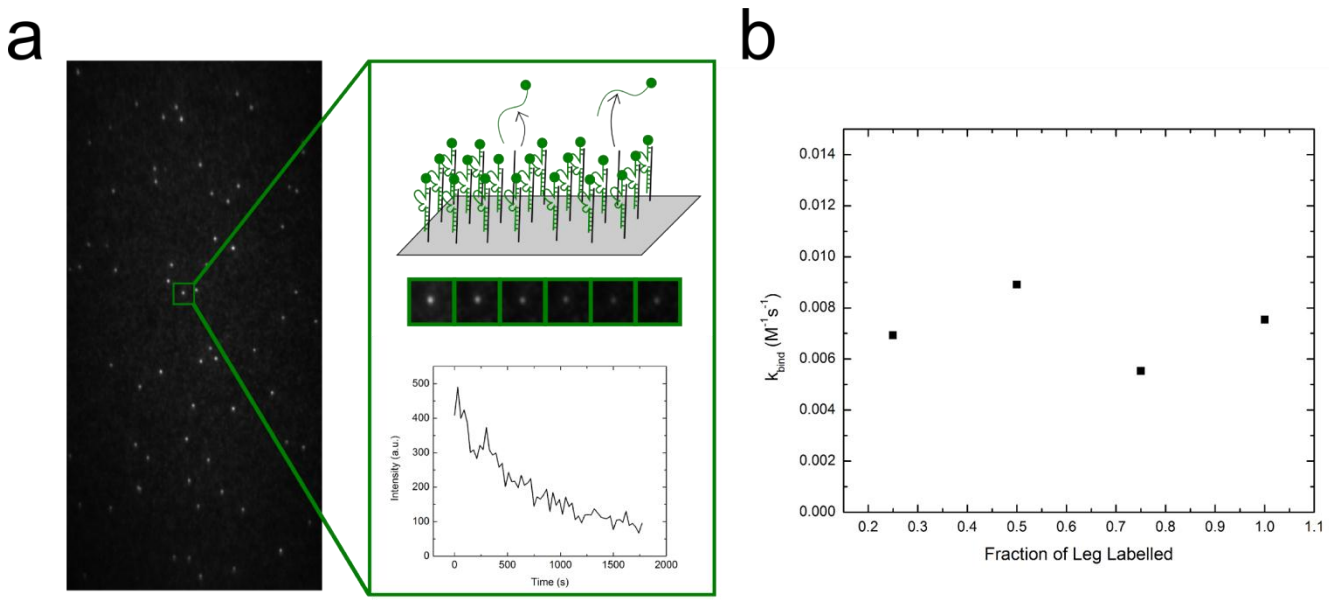


Figure 3.6 Method for determining rate constants. (**a**) For each origami in a field of view, the dissociation rate is experimentally determined by monitoring the loss in fluorescence over time. These individual dissociation curves are normalized and averaged together over all the origami in a field of view (see Figure 3.9c-d). The binding rate constant is determined in a similar way, by monitoring the increase in fluorescence over time instead of the decrease. (**b**) Binding rate constant determined for 50 nM leg given different fractions of labeled-to-unlabeled leg.

the fluorescence intensity and the number of legs bound that affects the rate constant experimentally determined.

In the presence of 1 mM $ZnSO_4$ in 1XHBS (i.e., under the original reaction buffer conditions[45]), Cy3-labeled legs tended to form large aggregates that crashed to the slide surface as indicated by a bright fluorescent spot appearing in a single time-point that was not colocalized with an origami tile (Figure 3.7a). In lower concentrations of $ZnSO_4$, this effect was a rare occurrence, revealing that the binding rate constant was dependent on zinc concentration (Figure 3.7b). At very low concentrations of $ZnSO_4$ (~250 μM), the binding rate constant increased when compared with the case where a divalent metal ion was absent. However, as the concentration increased, binding became impaired. Lowering the $ZnSO_4$ concentration in experiments would be a suboptimal solution to improving the binding rate constant, however, because lower concentrations of zinc would also reduce the cleavage rate constant.

To overcome this challenge of binding while still maintaining a high cleavage rate, $MgCl_2$ was chosen as a potential substitute cofactor. The cleavage rate constants for various concentrations of $MgCl_2$ were determined in solution using ensemble conditions. The cleavage rate constants in 1 mM $ZnSO_4$ and for multiple concentrations of $MgCl_2$ were determined under both single-turnover (i.e., with an excess of legs over substrates; Figure 3.8a-b) and multiple-turnover (i.e., with an excess of substrates over legs; Figure 3.8c-d) conditions. Although magnesium was determined to be a less effective cofactor for the zinc-selected 8-17 deoxyribozyme used for the leg[51], increasing the magnesium concentration can increase the cleavage rate constant to the point that it surpasses that at 1 mM $ZnSO_4$ (Figure 3.8). While 20 mM $MgCl_2$ exhibited the highest cleavage rate constant (0.93 min^{-1}), its extent of cleavage after 1 h was low (65%) under multiple-turnover conditions, similar to lower concentrations of $MgCl_2$

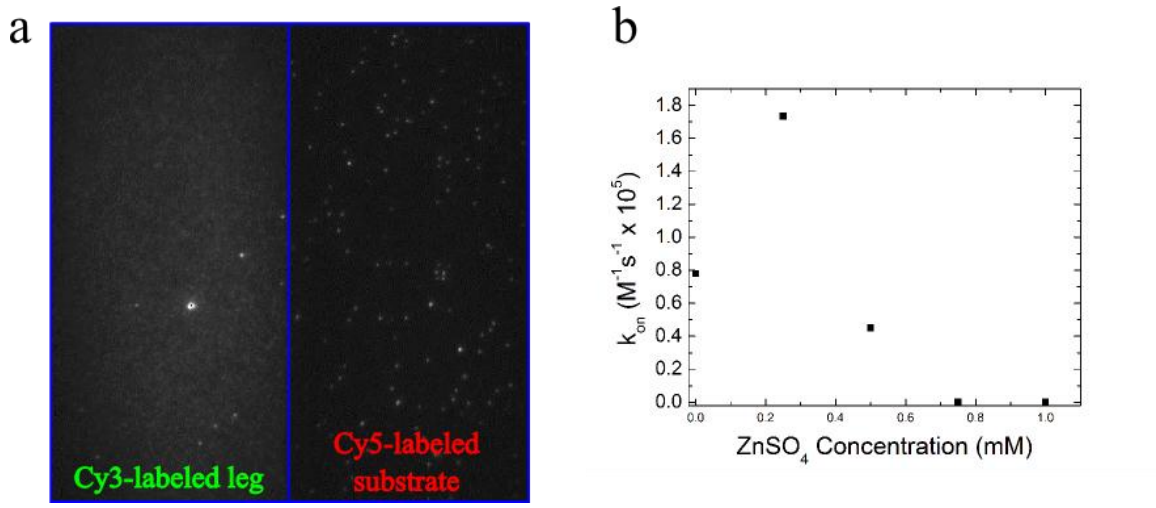


Figure 3.7 Zinc-dependent effect on leg-substrate binding rate. a. Large aggregates of Cy3-labeled leg nonspecifically crash to the slide surface. These aggregates of leg are not colocalized with the Cy5-labeled substrate on the origami. **b.** The binding rate constant as a function of $ZnSO_4$ concentration.

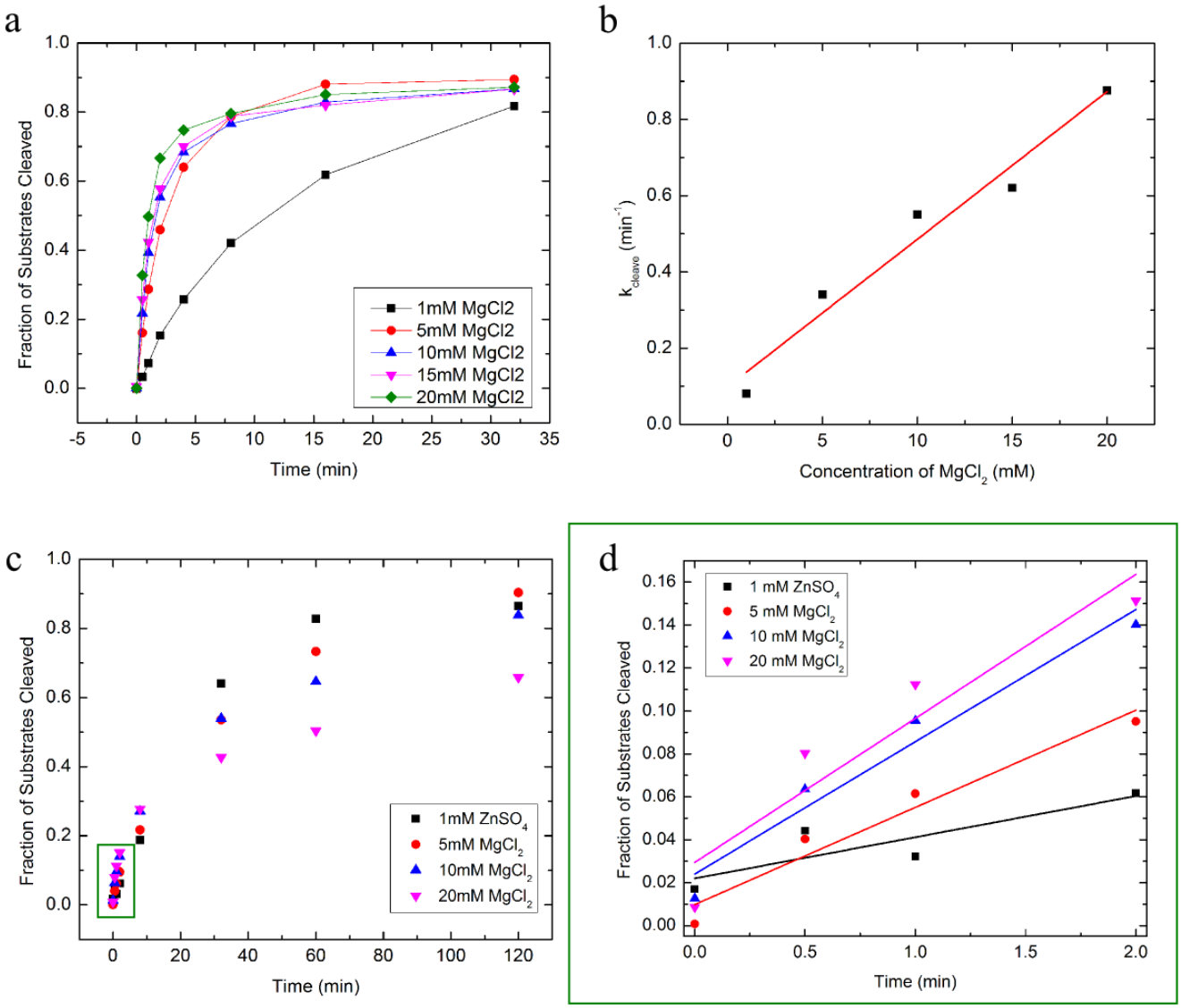


Figure 3.8 Cleavage rates for varying concentrations of divalent metal ions. **a.** Fraction of substrates cleaved under single-turnover conditions. The single-turnover cleavage rate constants for the various $MgCl_2$ concentrations, computed by fitting the fractions cleaved to exponentials, were $k_{1\ mM\ Mg} = 0.10\ min^{-1}$, $k_{5\ mM\ Mg} = 0.36\ min^{-1}$, $k_{10\ mM\ Mg} = 0.60\ min^{-1}$, $k_{15\ mM\ Mg} = 0.68\ min^{-1}$ $k_{20\ mM\ Mg} = 0.93\ min^{-1}$. **b.** The cleavage rate is a linear function of the concentration of $MgCl_2$. **c.** Fraction of substrates cleaved under multiple-turnover conditions. **d.** The first four data points of c, boxed in green, were fit linearly to determine the multiple-turnover rate constants: $k_{1\ mM\ Zn} = 0.019\ min^{-1}$, $k_{5\ mM\ Mg} = 0.045\ min^{-1}$, $k_{10\ mM\ Mg} = 0.062\ min^{-1}$, $k_{20\ mM\ Mg} = 0.067\ min^{-1}$.

(Figure 3.8c-d). Because the leg's ability to bind to its substrate, cleave the substrate, dissociate, and bind to a new substrate (the entire process of which is captured by the multiple-turnover rate and extent of cleavage as opposed to the single-turnover condition, which only captures the leg's ability to bind and cleave) is essential to the spider walking mechanism, this limited extent of cleavage at 20 mM $MgCl_2$ must be taken into consideration when choosing what concentration to use. 10 mM $MgCl_2$ was found to be the best compromise between cleavage rate (0.60 min^{-1}) and ability to continuously dissociate and rebind to another substrate (extent of cleavage of 83% after 1h). It is worth noting that in the presence of 1 mM $ZnSO_4$ the total amount of cleavage leveled off after an hour with an extent of cleavage of 86%. This stopping short of 100% is consistent with aggregate formation at the single-molecule level. The leg's ability to bind to its substrates during the ensemble experiments and not during the single-molecule binding experiments in the presence of $ZnSO_4$ may be due to the excess exposure to substrates. As cleavage occurs and the excess of substrates diminishes, the legs are more likely to aggregate, limiting access to fresh substrate and leading to premature cessation of cleavage.

$k_{bind,sub}$, $k_{bind,prod}$, $k_{dissoc,sub}$, and $k_{dissoc,prod}$ for 10 mM $MgCl_2$ were determined using the same single-molecule technique described for 1 mM ZnSO4 (Table 3.1; Figure 3.9). There was a ~4-fold difference between $k_{bind,prod}$ and $k_{bind,sub}$, in 10 mM $MgCl_2$,which was better for the speed of the spider than the 100-fold difference in 1 mM $ZnSO_4$. However, there was also only a ~6-fold difference between $k_{dissoc,sub}$ and $k_{dissoc,prod}$ in 10 mM $MgCl_2$, which is not as beneficial as the 360-fold difference between these two rate constants in 1 mM $ZnSO_4$. According to Monte Carlo simulations, the kinetic rate constants for 10 mM $MgCl_2$ were slightly more favorable to generate directional bias in the spider walking mechanism (Figure 3.10). The diffusion coefficient for 10 mM $MgCl_2$ was 1.5-times higher than in 1 mM $ZnSO_4$ (3.4 nm^2/min) and

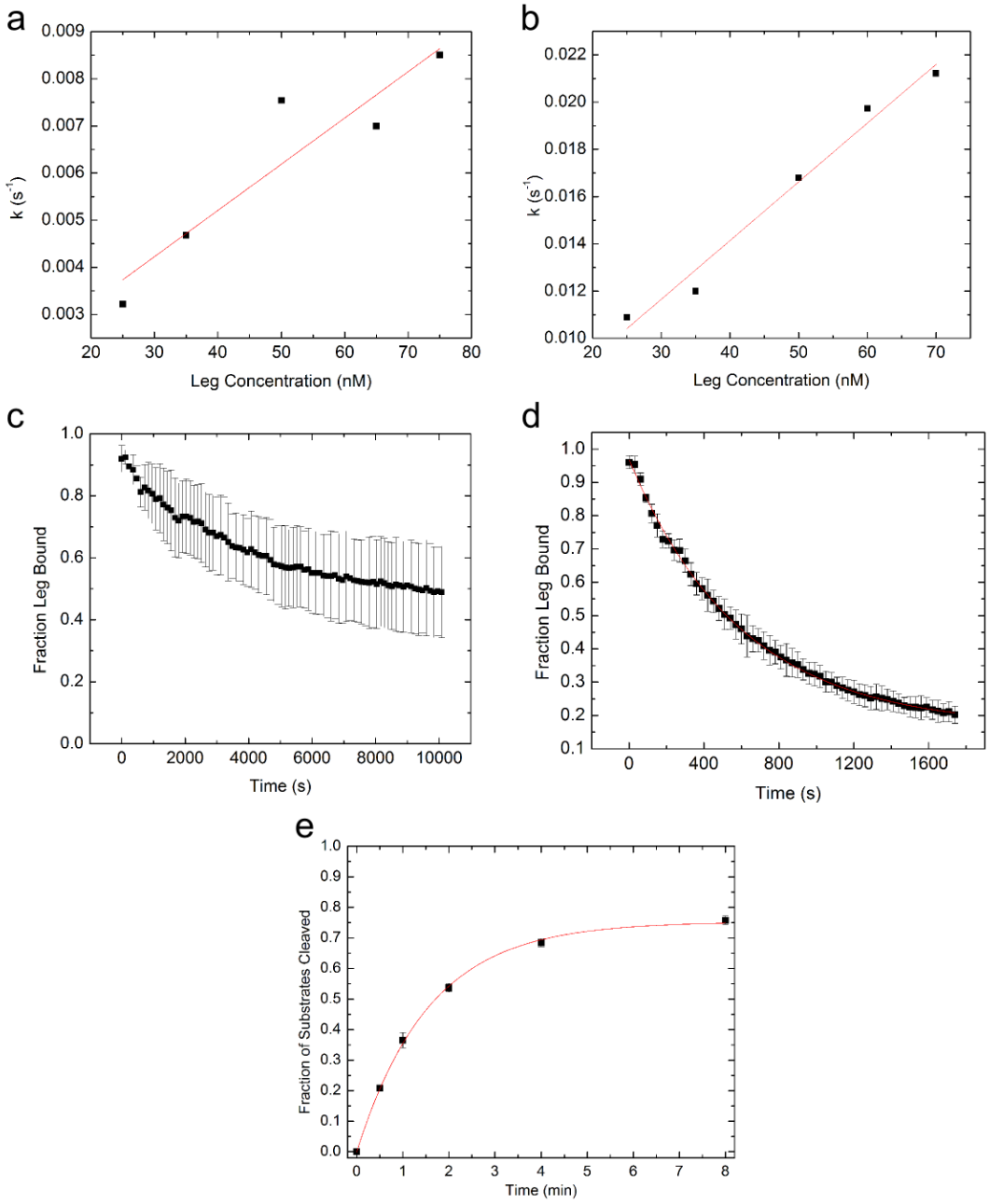


Figure 3.9 Rate constants for 10 mM MgCl2. The (**a**) binding rate constant for the substrate ($k_{bind,sub}$) and the (**b**) binding rate constant for the product ($k_{bind,prod}$) are determined by the slope of the observed rate constants as a function of concentration of leg used in each experiment. The (**c**) dissociation rate of the substrate ($k_{dissoc,sub}$) and the (**d**) dissociation rate of the product ($k_{dissoc,prod}$) are determined by averaging three trials, fitting the data to an exponential, and taking the exponent. The error bars are the standard deviation for each time point. The (e) cleavage rate (k_{cleave}) is determined by averaging three trials, fitting the average fraction cleaved over time to an exponential, and taking the exponent. The error bars are the standard deviation for each time point.

demonstrated less confinement with an alpha parameter of 0.6 (Figure 3.10a). After 3 h, one spider given the 10 mM $MgCl_2$ rate constants completed the track while no spiders completed the track under the conditions for 1 mM $ZnSO_4$. The mean substrates cleaved of the 42 total substrates per origami after 3 h, as determined by a Gaussian fit to the distribution of substrates cleaved per origami, was 11.2 (27%) under 1 mM ZnSO4 conditions and 13.1 (31%) under 10 mM MgCl2 conditions.

We would like to test these simulated results experimentally. However, as seen from the limited number of trajectories previously reported[45], obtaining a statistically significant number of spider trajectories is not trivial. This low yield is primarily due to an insufficient number of spiders bound to their origami tracks. To increase this number, a large excess of spiders may be added to the origami. This results in a large number of spiders nonspecifically binding to the slide surface, making it difficult to determine which spiders are bound to the origami track. Therefore, two novel single-molecule fluorescence methods for monitoring the spider movement were developed that relied on monitoring the changes in the substrate track as it is cleaved by the spider, instead of monitoring the movement of the spider itself. These techniques are meant to increase the number of trajectories observed from one per experiment to tens to thousands.

The first method involved observing the fluorescent signal decrease on each origami as a result of the Cy5-labeled substrates being cleaved by the spider (Figure 3.11). The rate at which the spider cleaves the substrate in 10 mM $MgCl_2$ (0.081 min^{-1}) was 2.8-times faster than that of the spider in 1 mM $ZnSO_4$ (0.029 min^{-1}). In the presence of 10 mM $MgCl_2$, the initial cleavage seemed faster and the extent of cleavage leveled off sooner than it did with the spider in 1 mM $ZnSO_4$. This is the opposite trend from that demonstrated by the simulations (Figure 3.10). However, after 3 h, the fraction of substrate cleaved (as determined by taking one minus the

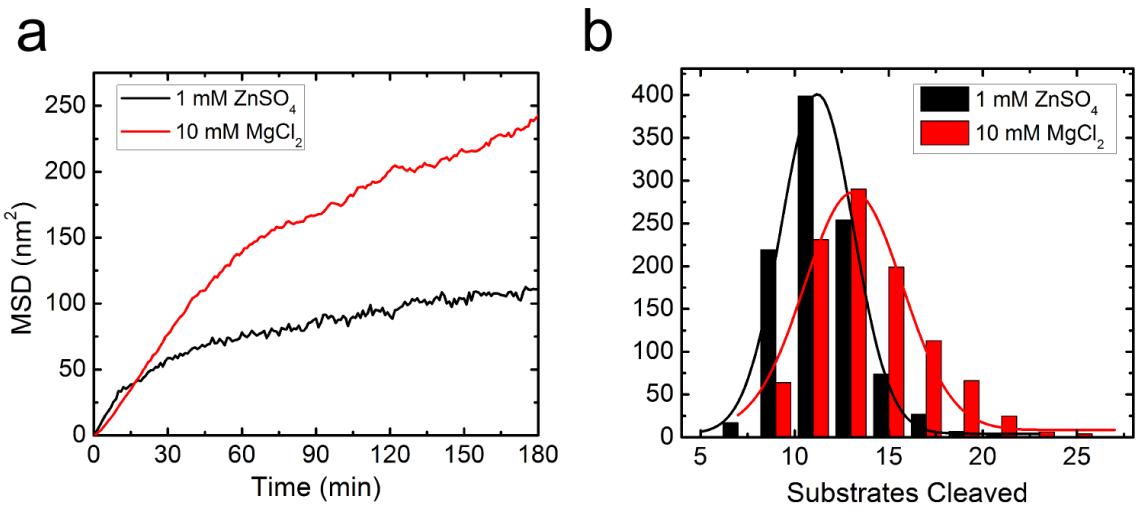


Figure 3.10 Simulated dependence of spider mobility on the type of divalent metal ion. (**a**) Monte Carlo simulation results using the rate constants measured by Taylor[50] in the presence of 1 mM $ZnSO_4$ and the rate constants determined from the data in figure 3.9 with 10 mM $MgCl_2$ (Table 3.1). (**b**) The predicted number of substrates cleaved per origami given the experimentally determined rate constants for 1 mM $ZnSO_4$ and 10 mM $MgCl_2$.

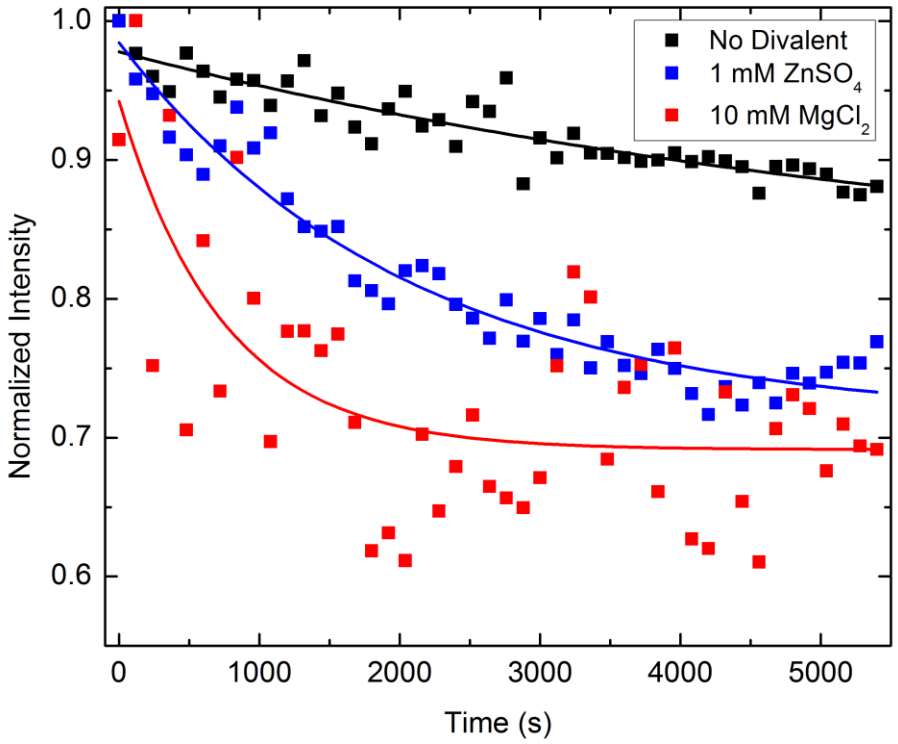


Figure 3.11 Real-time observation of spiders cleaving their Cy5-labeled substrates by observing the decrease in Cy5 intensity over time. The observed rates were $k_{none} = 0.009\ min^{-1}$, $k_{1mMZnSO4} = 0.029\ min^{-1}$, $k_{10mMMgCl2} = 0.081\ min^{-1}$.

value of the fraction of substrate still remaining) was 0.29 for 1 mM $ZnSO_4$ and 0.31 for 10 mM $MgCl_2$, consistent with the simulated results that estimate 0.27 for 1 mM ZnSO4 and 0.31 for 10 mM $MgCl_2$ for the fraction of substrates cleaved.

Magnesium is known to affect the photophysical properties of the fluorophores, and therefore may contribute to the experimental results. As a control to determine how significant this contribution was, the same experiment was performed in 10 mM $MgCl_2$ in the absence of spiders. The rate constant associated with a decrease in fluorescence intensity from $MgCl_2$ and photobleaching (0.009 ± 0.014 min^{-1}) in the control cannot account for rate constants obtained in the experiments with the spiders in the presence of divalent metal ions that are significantly faster.

While this technique has the advantage of yielding real-time measurements for 50-100 origami per experiment and providing rate constants for the spider walking along its track, it has several limitations. First, after the initial burst of cleavage, it is difficult to distinguish the substrates being cleaved from photobleaching. Second, even if nonspecific binding was not a problem, the Cy3-labeled spider and Cy5-labeled track undergo sufficient FRET, making it difficult to detect the Cy3-labeled spider and determine which origami are occupied by the spider. The unoccupied origami will affect the rate constant by making the exponential less steep since cleavage does not occur on unoccupied origami. Therefore, without having a way of distinguishing which origami were occupied, we assumed the same number of origami were occupied in each experiment in order to determine a ratio between the rate constants found in different experiments. The spider cleavage rate constants reported thus provide a lower bound for the actual spider cleavage rate constant.

A complementary technique was developed in which the spider cleaved the Cy5-labeled substrates along its origami track in solution. Afterward, the origami was immobilized on the slide surface. The peak Cy5-intensity from each origami was detected from several fields of view. This permits the user to scan over several fields of view, enabling the ability to generate large data sets within a single experiment. Histograms of these intensities were generated as a means for determining the distribution of substrates cleaved on each origami (Figure 3.12a). To determine the uncertainty involved in this method, three different histogram distributions were randomly generated from the origami from 39 different fields of view on 37 different slides (Figure 3.12b) to provide a standard deviation for the mean of the three histograms, which should ideally all have the same mean. The standard deviation of the mean of these three Gaussians was 76 photoelectrons. This measured standard deviation is what we will use for the error in determining the mean.

The first observation of interest in using this method is that the number of origami analyzed has increased by an order of magnitude (3,000-5,000) from the previous method. The next observation is that the histogram of the intensity of the substrates on the origami in the absence of a divalent metal ion and the histogram in 10 mM $MgCl_2$ each appear to be a superposition of Gaussians. By fitting each of the distributions with two Gaussians, we can filter out the unoccupied origami (Figure 3.12c). The mean intensity for 1 mM $ZnSO_4$ is higher than that for 1 mM $MgCl_2$ and no divalent due to aggregation of origami (if the origami are aggregated, this means that there will be more Cy5-labeled substrates within a diffraction-limited region, resulting in a higher perceived intensity for each localization), which is consistent with the large aggregates of leg that formed in the presence of 1 mM $ZnSO_4$. For 10 mM $MgCl_2$, 48% of the origami were unoccupied, as determined by the ratio in amplitudes of the two Gaussians.

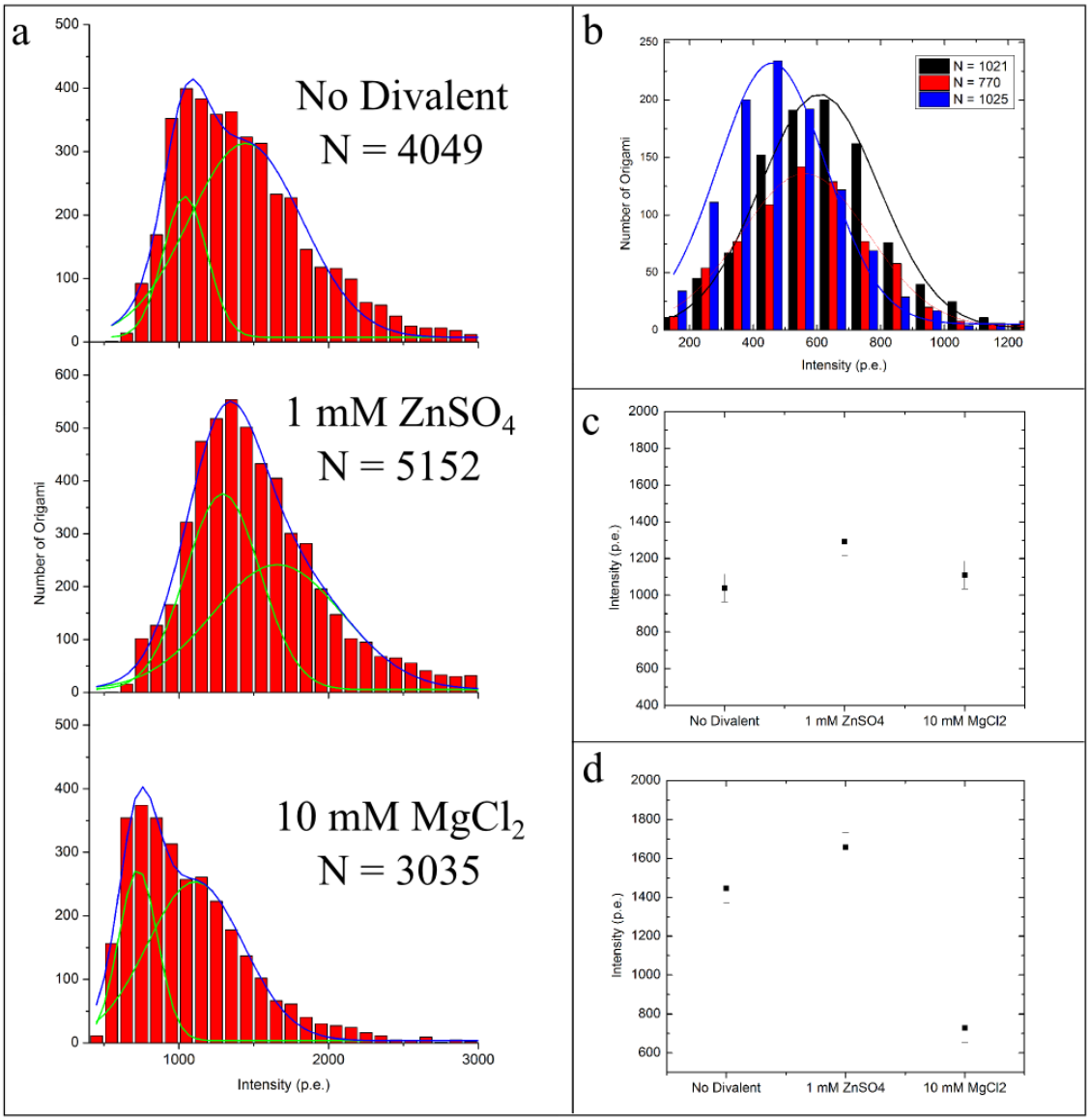


Figure 3.12 Determination of the extent of cleavage from the spider's walk along origami in solution. **a.** Histograms of the intensity from the remaining Cy5-labeled substrates on each origami. **b.** Histograms of the intensity of origami from 37 different slides with the same number of fluorescently-labeled substrates. Three groups of origami were formed at random so that each group had approximately the same number of origami. The central position for each Gaussian was 608, 561, and 459 p.e., giving a standard deviation of 76. This number is used to form the error bars in (c) and (d). **c.** The central peak intensity of the Gaussian fit from (a) attributed to unoccupied origami. These values were 1040, 1290, and 1110 p.e. for no divalent, 1 mM $ZnSO_4$, and 10 mM $MgCl_2$ respectively. **d.** The central peak intensity of the Gaussian fit from (a) attributed to something other than unoccupied origami. These values were 1440, 1660, and 730 p.e. for no divalent, 1 mM $ZnSO_4$ and 10 mM $MgCl_2$ respectively.

The mean of the second Gaussian provides information concerning the aggregation in the case of 1 mM $ZnSO_4$ and no divalent, or the number of cleaved substrates in the case of 10 mM $MgCl_2$ (Figure 3.12d). We can determine the number of substrates cleaved by taking one minus the mean of the cleaved-substrates Gaussian divided by the mean of the unbound-origami Gaussian. Doing this, we find that 34% of the substrates are cleaved, a value comparable to the 31% predicted by the Monte Carlo simulations.

This technique has the advantage of building up large data sets well over 1,000 data points within a single experiment, but is only a snapshot and therefore lacks the real-time measurements of the other technique.

3.5.4 Discussion

Here, we have shown the importance of including the difference between $k_{bind,sub}$ and $k_{bind,prod}$ in simulations, a difference that had previously been neglected in simulations. This effect radically slows the spider's progression to the end of its track. The rate constants to favor forward movement and decrease aggregation may be somewhat improved by replacing the divalent metal ion used. However, on average, it is estimated that given these rate constants, only ~30% of the track will be cleaved after three hours. This was experimentally confirmed using two different single-molecule fluorescence imaging techniques.

While spiders may not be efficient at generating their own bias due to their high affinity for product, their ability to manipulate their environments where they walk by cleaving their substrates may be applicable to nanopatterning, provided they have an external source to provide the bias (Appendix B).

3.6 REFERENCES

1 Siegwart, R. & Nourbakhsh, I. R. *Introduction to Autonomous Mobile Robots*. (MIT Press, 2004).

2 Turing, A. M. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society Series* **2**, 230-265 (1936).

3 Braitenberg, V. *Vehicles: Experiments in Synthetic Psychology*. (MIT Press, 1984).

4 Brooks, R. A. Intelligence without representation. *Artificial Intelligence* **47**, 139–159 (1991).

5 Bath, J. & Turberfield, A. DNA nanomachines. *Nature Nanotechnology* **2**, 275-284 (2007).

6 Sherman, W. B. & Seeman, N. C. A precisely controlled DNA biped walking device. *Nano Letters* **4**, 1203-1207 (2004).

7 Shin, J. S. & Pierce, N. A. A synthetic DNA walker for molecular transport. *Journal of the American Chemical Society* **126**, 10834-10835 (2004).

8 Tian, Y., He, Y., Chen, Y., Yin, P. & Mao, C. A DNAzyme that walks processively and autonomously along a one-dimensional track. *Angew. Chem. Int. Edn* **44**, 4355-4358 (2005).

9 Bath, J., Green, S. J. & Turberfield, A. J. A free-running DNA motor powered by a nicking enzyme. *Angew. Chem. Int. Edn* **44**, 4358-4361 (2005).

10 Yin, P., Choi, H., Calvert, C. R. & Pierce, N. A. Programming biomolecular self-assembly pathways. *Nature* **451**, 318-322 (2008).

11 Omabegho, T., Sha, R. & Seeman, N. C. A Bipedal DNA Brownian Motor with Coordinated Legs. *Science* **324**, 67-71 (2009).

12 Pei, R. *et al.* Behavior of Polycatalytic Assemblies in a Substrate-Displaying Matrix. *J. Am. Chem. Soc* **128**, 12693-12699 (2006).

13 Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 298-302 (2006).

14 Rus, D., Butler, Z., Kotay, K. & Vona, M. Self-reconfiguring robots. *Communications of the ACM* **45**, 39-45 (2002).

15 Bonabeau, E., Dorigo, M. & Theraulaz, G. *Swarm Intelligence: From Natural to Artificial Systems*. (Oxford University Press, 1999).

16 Bennett, C. H. The thermodynamics of computation—a review. *International Journal of Theoretical Physics* **21**, 905-940 (1982).

17 Von Neumann, J. *Theory of Self-Reproducing Automata*, ed. Burks, A.W. (University of Illinois Press, Urbana, IL, 1966).

18 Santoro, S. W. & Joyce, G. F. A general purpose RNA-cleaving DNA enzyme. *PNAS* **94**, 4262-4266 (1997).

19 Antal, T. & Krapivsky, P. L. Molecular spiders with memory *Phys. Rev. E* **76**, 021121-021129 (2007).

20 Saffarian, S., Collier, I. E., Marmer, B. L., Elson, E. L. & Goldberg, G. Interstitial Collagenase is a Brownian Ratchet Driven by Proteolysis of Collagen. *Science* **306**, 108-111 (2004).

21 Ke, Y., Lindsay, S., Chang, Y., Liu, Y. & Yan, H. Self-Assembled Water-Soluble Nucleic Acid Probe Tiles for Label-Free RNA Hybridization Assays. *Science* **319**, 180-183 (2008).
22 Yurke, B., Turberfield, A. J., Mills, A. P., Simmel, F. C. & Neumann, J. L. A DNA-fuelled molecular machine made of DNA. *Nature* **406**, 605-608 (2000).
23 Li, J., Zheng, W., Kwon, A. H. & Lu, Y. In vitro selection and characterization of a highly efficient Zn (II)-dependent RNA-cleaving deoxyribozyme. *Nucleic Acids Research* **28**, 481-488 (2000).
24 Pei, R. *et al.* Behavior of Polycatalytic Assemblies in a Substrate-Displaying Matrix. *J. Am. Chem. Soc.* **128**, 12693-12699 (2006).
25 Aitken, C. E., Marshall, R. A. & Pulglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys. J.* **94**, 1826-1835 (2008).
26 Walter, N. G., Huang, C.-Y., Manzo, A. J. & Sobhy, M. A. Do-it-yourself guide: How to use the modern single-molecule toolkit. *Nat. Methods* **5**, 475-489 (2008).
27 Churchman, L. S., Okten, Z., Rock, R. S., Dawson, J. F. & Spudich, J. A. Single molecule high-resolution colocalization of Cy3 and Cy5 attached to macromolecules measures intramolecular distances through time. *Proc. Natl. Acad. Sci. USA* **102**, 1419–1423 (2005).
28 Yildiz, A. & Selvin, P. R. Fluorescence imaging with one manometer accuracy: Application to molecular motors. *Acc. Chem. Res.* **38**, 574-582 (2005).
29 Li, J., Zheng, W., Kwon, A. H. & Lu, Y. In vitro selection and characterization of a highly efficient Zn (II)-dependent RNA-cleaving deoxyribozyme. *Nucleic Acids Res.* **28**, 481-488 (2000).
30 Schutz, G. J., Trabesinger, W. & Schmidt, T. Direct observation of ligand colocalization on individual receptor molecules. *Biophys. J.* **74**, 2223-2226 (1998).
31 Schellart, N. in *Image Analysis: Methods and Applications, 2nd ed.* (ed Hader Donat-P) 359-361 (CRC Press, 2001).
32 Toprak, E. *et al.* Defocused orientation and position imaging (DOPI) of myosin V. *Proc. Natl. Acad. Sci. USA* **103**, 6495-6499 (2006).
33 Wetmur, J. G. & Fresco, J. DNA Probes: Applications of the Principles of Nucleic Acid Hybridization. *Crit. Rev. Biochem. Mol. Biol.* **26**, 227-259 (1991).
34 Churchman, L. S., Okten, Z., Rock, R. S., Dawson, J. F. & Spudich, J. A. Single molecule high-resolution colocalization of Cy3 and Cy5 attached to macromolecules measures intramolecular distances through time. *Proc. Natl. Acad. Sci.* **102**, 1419–1423 (2005).
35 Yildiz, A. & Selvin, P. R. Fluorescence imaging with one nanometer accuracy: application to molecular motors. *Accounts of chemical research* **38**, 574-582 (2005).
36 Hess, H. Toward Devices Powered by Biomolecular Motors. *Science* **312**, 860-861 (2006).
37 Adleman, L. Molecular computation of solutions to combinatorial problems. *Science* **266**, 1021-1024 (1994).
38 Stojanovic, M. N. & Stefanovic, D. A deoxyribozyme-based molecular automaton. *Nature Biotechnology* **21**, 1069-1074 (2003).
39 Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-Free Nucleic Acid Logic Circuits. *Science* **314**, 1585-1588 (2006).

40 Samii, L., Linke, H., Zuckermann, M. J. & Forde, N. R. Biased motion and molecular motor properties of bipedal spiders. *Physical review. E, Statistical, nonlinear, and soft matter physics* **81**, 021106 (2010).
41 Samii, L. *et al.* Time-dependent motor properties of multipedal molecular spiders. *Physical review. E, Statistical, nonlinear, and soft matter physics* **84**, 031111 (2011).
42 Antal, T. & Krapivsky, P. L. Molecular spiders with memory. *Physical Review E* **76**, - (2007).
43 Antal, T., Krapivsky, P. L. & Mallick, K. Molecular spiders in one dimension. *J Stat Mech-Theory E*, - (2007).
44 Semenov, O., Olah, M. J. & Stefanovic, D. Mechanism of diffusive transport in molecular spider models. *Physical review. E, Statistical, nonlinear, and soft matter physics* **83**, 021117 (2011).
45 Lund, K. *et al.* Molecular robots guided by prescriptive landscapes. *Nature* **465**, 206-210, doi:Doi 10.1038/Nature09012 (2010).
46 Gillespie, D. T. A General Method for Numerically Simulating the Stochastic Tim Evolution of Coupled Chemical Reactions. *Journal of Computational Physics* **22**, 403-434 (1976).
47 Gillespie, D. T. Exact Stochastic Simulation of Coupled Chemical Reactions. *Journal of Physical Chemistry* **81**, 2340-2361 (1977).
48 Antal, T. & Krapivsky, P. L. Molecular spiders with memory. *Physical review. E, Statistical, nonlinear, and soft matter physics* **76**, 021121 (2007).
49 Antal, T., Krapivsky, P. L. & Mallick, K. Molecular Spiders in One Dimension. *Journal of statistical mechanics* **2007**, P08027, doi:10.1088/1742-5468/2007/08/P08027 (2007).
50 Taylor, S. Unpublished. (2005).
51 Li, Y., Liu, Y. & Breaker, R. R. In vitro selection of deoxyribozymes with DNA capping activity. *Nucleic acids symposium series*, 237-238 (1999).

CHAPTER 4:
OBSERVING CHARGED PARTICLES EMBEDDED IN THE NANOPORES OF ANTIMONY-DOPED TIN OXIDE

4.1 Introduction

Transparent conducting oxides (TCOs) have been of great interest in recent years. Their conductive and transparent properties enable them to be used as electrodes for devices that may be imaged using common visible light microscopy techniques[1]. These materials have been used for biosensors[2], electrochemistry[3,4], photoelectrochemical solar cell fabrication[5], and electrooptical devices[6].

Antimony-doped tin oxide (ATO) is a porous TCO. Its nanoporous nature results in a high specific surface area, giving rise to high loading capacities for electrosensitive materials, such as dye-sensitized solar cells or photosynthetic inorganic-bio hybrid systems[1]. Its expense is comparable to that of Indium Tin Oxide (ITO), and it maintains high conductivity and optical transparency upon use[7].

Our collaborators recently developed and characterized a one-pot ATO synthesis procedure[7]. While the characterization of the ATO resulting from this synthesis was more thorough than previous analyses of ATO materials, it still had its limitations[1]. Visualization of the pores was performed using electron microscopy techniques, which only revealed surface features and did not disclose information concerning the connectivity of the pores. Ensemble electrochemical experimental techniques such as cyclic voltammetry do not reveal the behavior

of nanomaterials encapsulated within the pores upon confinement and the application of electric fields.

The transparent property of ATO lends itself to the utilization of single-molecule fluorescence microscopy techniques for further characterization. Here, charged fluorescent particles were introduced to the nanopores and observed in real-time as they diffuse about the labyrinth in the presence and absence of an electric field. Mapping the trajectories of individual particles and determining their diffusion coefficients provides information concerning the connectivity and geometry of the nanopores, while the interaction of these charged particles with an electric field of varying strength provides information concerning how embedded electrosensitive materials may be affected by the presence of an electric field.

4.2 Materials and Methods

ATO-coated Coverslip Fabrication. In a typical preparation of a precursor solution, 12 mg of $SbCl_3$ (Alfa Aesar, 99.9%) was dissolved in 1.20 g of absolute ethanol (Koptec). Then 0.280 g of $SnCl_4 \cdot xH_2O$ (Alfa Aesar, 98%) was dissolved in this same solution with stirring. A few drops of distilled water amounting to about 60 mg were added to the precursor solution followed by 0.150 g of polyethylene glycol (Fluka, reacted with bisphenol A digycidyl ether, MW Avg. 17,500). The precursor solution was then sonicated for 10 minutes to provide a cloudy white solution. Then 0.136 g of resorcinol (Sigma Aldrich, 99%) was added to the precursor solution which was then sonicated again for 2 minutes. Then 0.200 g of formaldehyde solution (Sigma Aldrich, 37 wt. % with 10-15% methanol as stabilizer) was added to the precursor solution with stirring. Finally, 0.55 g of epichlorohydrin (Fluka, 99%) was added to the precursor solution with stirring. The precursor solution became transparent and colorless approximately 1 min after the addition of the epichlorohydrin. After 20 min had elapsed since the addition of

epichlorohydrin, the aged precursor solution was then used to prepare thin coatings on the desired substrate.

Substrates were prepared by masking with Scotch tape to create a border of approximately 1 mm of tape on three edges of the substrate, leaving the fourth edge unmasked.

The coatings were prepared on the substrates using the doctor-blade method. A small drop of the aged precursor solution iwas placed on the masked edge of the substrate opposite the unmasked edge. The drop was then spread across the masked substrate towards the unmasked edge using the side of a glass Pasteur pipette in one fluid motion. Approximately 20 s after applying the coating, the coated substrate was then placed in a bath of mineral oil. Many coated substrates could be prepared in this way during the approximately 10-min window after the 20-min aging step and before the gelation of the precursor solution. The coated substrates in the mineral oil baths were left at room temperature for about 24 h, after which the oil baths with the coated substrates were placed in an oven at 70°C for 3 days to promote polymerization of the resorcinol and formaldehyde monomers. The oil baths with the coated substrates were removed to room temperature, after which the coated substrates were removed from the oil baths, rinsed with hexanes, and separated from the Scotch tape masks. The coated substrates were then heated in an ashing furnace at a rate of 100 °C/h from room temperature to 500 °C and held for 10 h followed by ambient cooling to room temperature to provide mesoporous ATO coatings.

Objective-type TIRF Microscopy. An Olympus IX81 microscope, with a 60 x 1.45 NA oil-immersion objective (Olympus), and EM-CCD camera (Evolve, Photometrics) were used for imaging. A solid state laser with wavelength 640 nm (100 mW) was directed through an acousto-optical tunable filter, split into different fiber-optic cables and then coupled to the cell-

TIRF module. A net power of ~8 mW was achieved. Laser beams were focused on the back-focal plane of the objective and made to travel parallel to the optic axis such that changing the distance from the optic axis controlled incident angle at the dish-media interface[8].

Fluorescent Probes. 25 mg Cy5 Hydrazide (GE Healthcare) was resuspended with 1.53 µL DMSO (Sigma-Aldrich 276855) and 170 µL water. The Cy5 Hydrazide was further diluted in water. During imaging, 16.8 pM Cy5 Hydrazide was used with 20% (v/v) Glycerol (Fisher BP 229-4) to slow diffusion, and an oxygen scavenging system[9] (OSS) to prolong fluorophore lifetime (5 nM protocatechuate (PCA), 50 nM protocatechuate dioxygenase (PCD), and 2 mM Trolox) in 1XHBS (10 mM HEPES, 150 mM NaCl, pH 7.4).

Applying Electric Potential Across Coverslip. Silver epoxy (MG Chemicals 8331-14G) was used to adhere two insulated flexible copper wires (Belden 8523) to opposite ends of a coverslip. The wires were then connected to a low-voltage power supply (Elenco Precision XP-660 Triple Regulated Power Supply).

Analysis. All analysis was done using home-built MATLAB programs. Gaussians were fit to point spread functions (PSFs) in every frame of every movie that passed an intensity filter. PSFs contiguous across frames were determined to be the same particle diffusing. The PSFs were tracked individually to determine their ensemble MSD ($< r(t)^2 > = 4Dt^{\alpha}$) or individual time-averaged MSDs (TAMSDs) ($< \Delta r(\tau)^2 > = \frac{1}{N-\tau}\sum_{i=1}^{N-\tau}|r_{i+\tau} - r_i|^2$).[10] The sum of the PSFs is presented to display the nanopores.

4.3 Results

Cy5-Hydrazides were successfully incorporated into the network of nanopores in the ATO-coated coverslip. Qualitatively, the nanopores exhibited high connectivity with a number of the trajectories overlapping (Figure 4.1a-c). Quantitatively, the ensemble mean square displacement (MSD) from 140 diffusing particles (Figure 4.1d) was used to calculate the diffusion coefficient and the alpha parameter by fitting the data to the equation:

$$< r(t)^2 > = 4Dt^{\alpha}$$

where r(t) is the displacement of the particle from its initial position, D is the diffusion coefficient, and the alpha parameter helps characterize the motion as subdiffusive ($\alpha < 1$), random ($\alpha = 1$), or superdiffusive ($\alpha > 1$). The MSD yielded a diffusion coefficient of 6.72 μm^2/min. The alpha parameter of 0.6, which is less than one, indicates confined diffusion; the confinement is presumably due to the walls of the nanopores.

To verify that the slowed diffusion is a consequence of nanopore confinement and not from surface interactions, the diffusion of Cy5-Hydrazides was observed on a couple of different types of surfaces. First, the diffusion on a nonconductive glass coverslip was observed, since it was the platform upon which the ATO was deposited. The dye either nonspecifically bound to the glass or momentarily bound and diffused back into solution (Figure 4.2). The second surface type tested was Indium Tin Oxide (ITO), which has similar properties to ATO in that it is conductive and transparent, but it does not have the nanoporous feature of ATO. Therefore, any diffusive behavior of the dye on ITO that resembles the diffusive behavior of the dye in ATO must be due to an interaction with the conductive surface rather than the confinement of the dye within nanopores. The dye upon ITO displayed a behavior similar to that of the dye on the glass

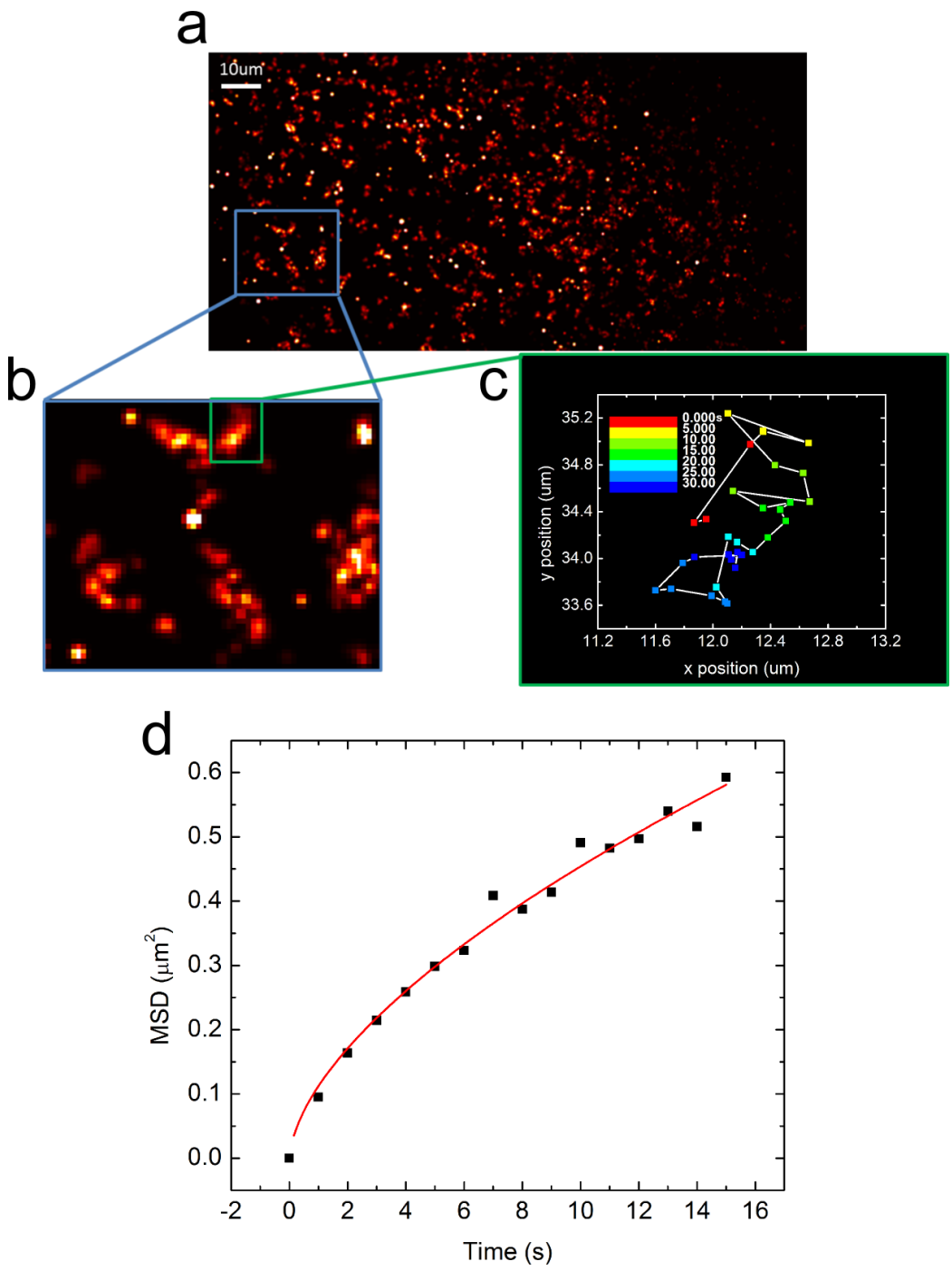


Figure 4.1 STORM reconstruction of particles diffusing. (**a**) A super-high-resolution image of the time course of all the particles diffusing through the ATO nanopores. (**b**) A zoomed-in portion gives a better sense of how this can be used to begin developing a map of the structure of the nanopores. (**c**) We can track individual particles over time to further characterize the pores by generating individual TAMSD (Figure 4.3) and **(d**) Ensemble MSD plots. From fitting the MSD (MSD = $4Dt^{\alpha}$), we obtained a diffusion coefficient of D = 0.112 $\mu m^2/s$ and $\alpha = 0.61$.

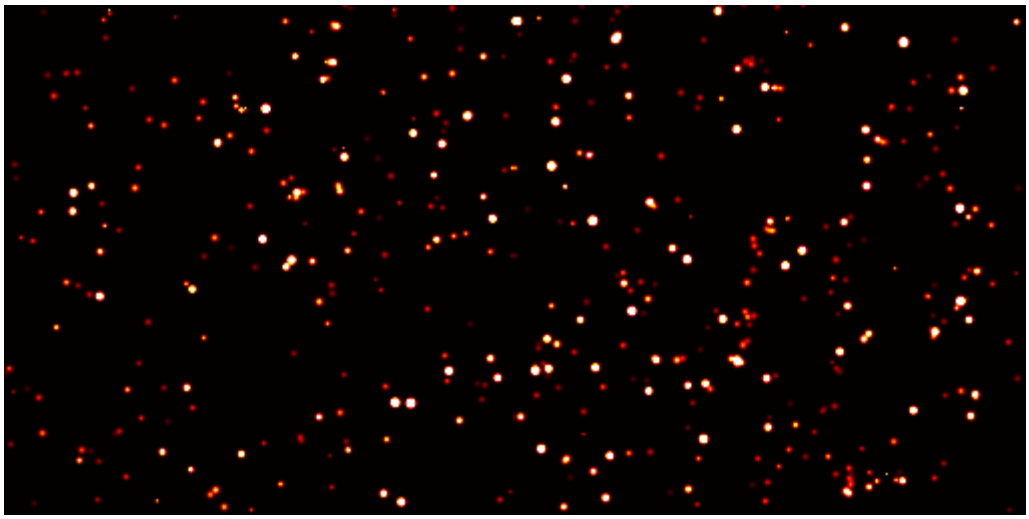


Figure 4.2 STORM reconstruction of particles on a glass coverslip without ATO.

coverslip (data not shown). Therefore, the diffusive behavior observed on ATO is most likely that of the dye diffusing within the nanoporous network.

The diffusion of the particles is heterogeneous as seen from the TAMSD of individual particles (Figure 4.3a). While it may be tempting to explain this variety of behavior as reflective of surface anomalies, 2D random walk Monte Carlo simulations suggest that such randomness of Brownian diffusion, i.e., the random bombardment by neighboring (invisible) molecules (Figure 4.4a). While these trends may indeed be the result of surface features, we cannot say with any level of certainty whether or not they are. Therefore, instead of focusing on individual TAMSD results as has been used previously for analysis[11], we will restrict our analysis to ensemble-averaged trends.

Because ATO could potentially be used as an electrode for embedded electroactive materials, the behavior was analyzed in the presence of an electric field. The alpha parameter from the ensemble MSDs maintained a value of 0.69 ± 0.05 (Figure 4.5a,c). This is consistent with the diffusing particles being embedded in the nanopores, and confined by the physical normal force of the wall. This also implies that the electric field is not significantly biasing the movement, as the alpha parameter would be greater than 1 if that were the case. This is further confirmed by $\langle x^2 \rangle$ and $\langle y^2 \rangle$ MSD plots and histograms (Figure 4.6), which exhibit no preference in directional movement. The diffusion coefficient, on the other hand, is significantly influenced by the presence of the electric field, increasing drastically when the potential reaches a high enough value and then dropping above ~10 V (Figure 4.5a-b).

This diffusive behavior is contrary to expectation. The Cy5-Hydrazide, with a charge of -1e, should experience a linear force from the presence of the electric field in the direction opposite from that of the field, and thus should exhibit a directional bias in its motion. It is also

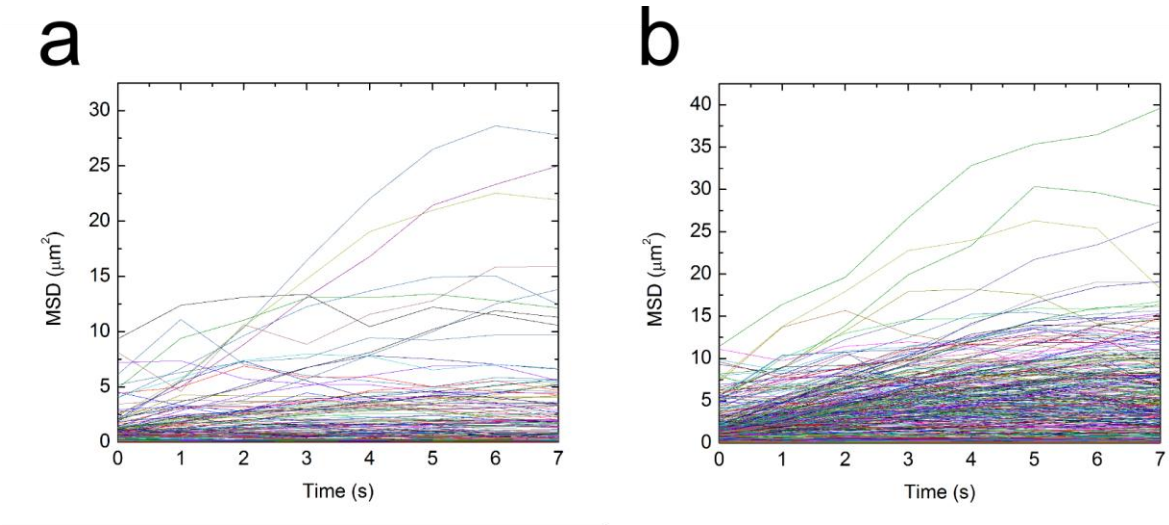


Figure 4.3 TAMSD plots for particles when a potential of (a) 0V or (b) 4V is applied.

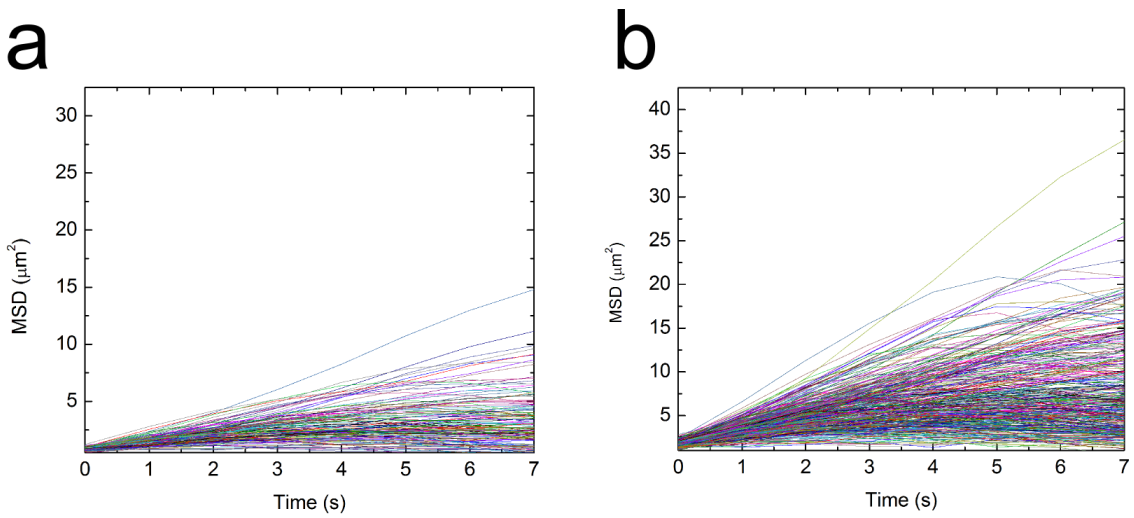


Figure 4.4 Simulated TAMSDs of individual particles undergoing random Brownian diffusion. Parameters are chosen so that these plots resemble the distribution of TAMSDs from experimental data for an applied potential of (**a**) 0V and (**b**) 4V (compare with Figure 4.3).

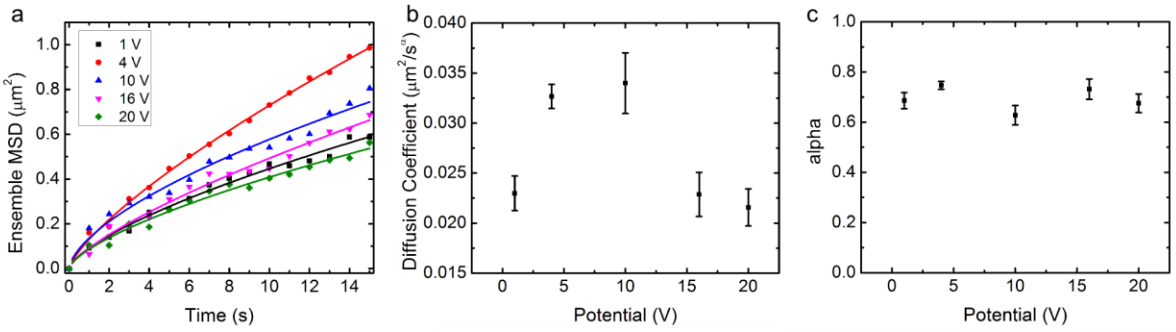


Figure 4.5 Dependence of ensemble MSDs on potential. (**a**) Ensemble MSDs of diffusing particles when various potentials are applied across the slide surface. When fit to the equation $MSD = 4Dt^{\alpha}$, the values for (**b**) the diffusion coefficient and (**c**) alpha are obtained for each potential.

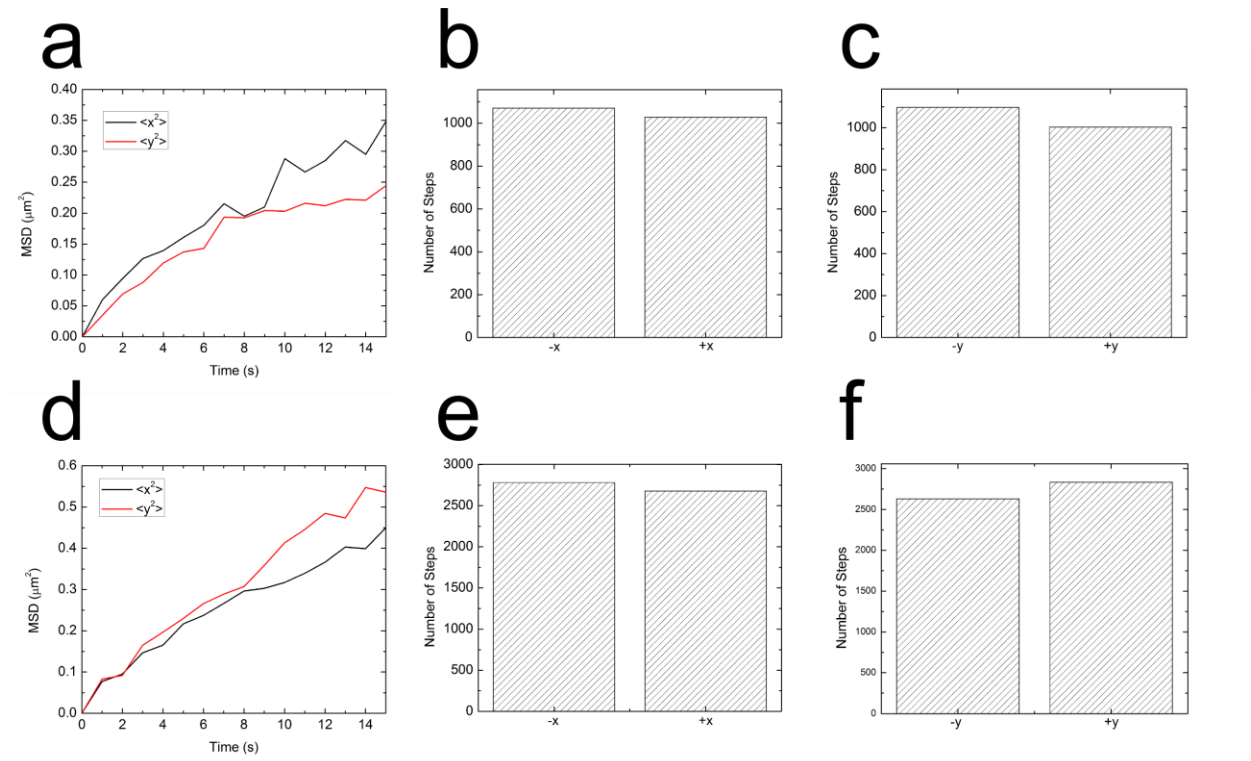


Figure 4.6 Tests of Directional Bias. (**a**) $\langle x^2 \rangle$ and $\langle y^2 \rangle$, (**b**) histograms representing the number of steps in the +x and –x directions, and (**c**) histograms representing the number of steps in the +y and –y directions for 0V. Similarly, (**d**) $\langle x^2 \rangle$ and $\langle y^2 \rangle$, (**e**) histograms representing the number of steps in the +x and –x directions, and (**f**) histograms representing the number of steps in the +y and –y directions for 4V.

unclear why the diffusion coefficient would have a potential threshold after which the diffusion coefficient decreases instead of increasing.

To investigate this unexpected behavior, the resistivity across the ATO as well as across a coverslip with buffer on it were measured. The resistance per length for the ATO and buffer were 28.7 $M\Omega cm^{-1}$ and 40 $k\Omega cm^{-1}$, respectively. There is thus a three-order-of-magnitude difference between the resistances of the two materials. Assuming the material is ohmic, a potential being applied across the ATO in contact with the buffer is analogous to two resistors connected in parallel in a simple circuit with a common voltage source. The current will be divided between the two materials in accordance with the equations $I_{buffer} = V/R_{buffer}$ and $I_{ATO} = V/R_{ATO}$. Hence, the current passing through the buffer should be about three orders of magnitude greater than the current passing through the ATO.

But there is another factor to consider. The buffer, being a fluid, contains ions that may move freely in the presence of an electric field. As the field causes the ions to be drawn toward the two leads, the current through the buffer will decrease. Therefore, the current and consequently the resistance in the buffer are expected to be functions of time.

We proceeded to test the current passing through each material. To determine the current passing through the buffer that was used for observing the diffusing Cy5-Hydrazides (i.e., 1XHBS), buffer was placed on a nonconductive glass coverslip with two wires connected to it such that the buffer completed the circuit between the two electrodes. The current was recorded over time for each potential with three trials for each potential (Figure 4.7). The buffer was replaced with fresh buffer after each trial. As the current passed through the buffer, the silver epoxy electrodes were damaged, making the three trials inconsistent. For instance, at 16 V, the rate constant determined from the first repetition increased by 250% from the original

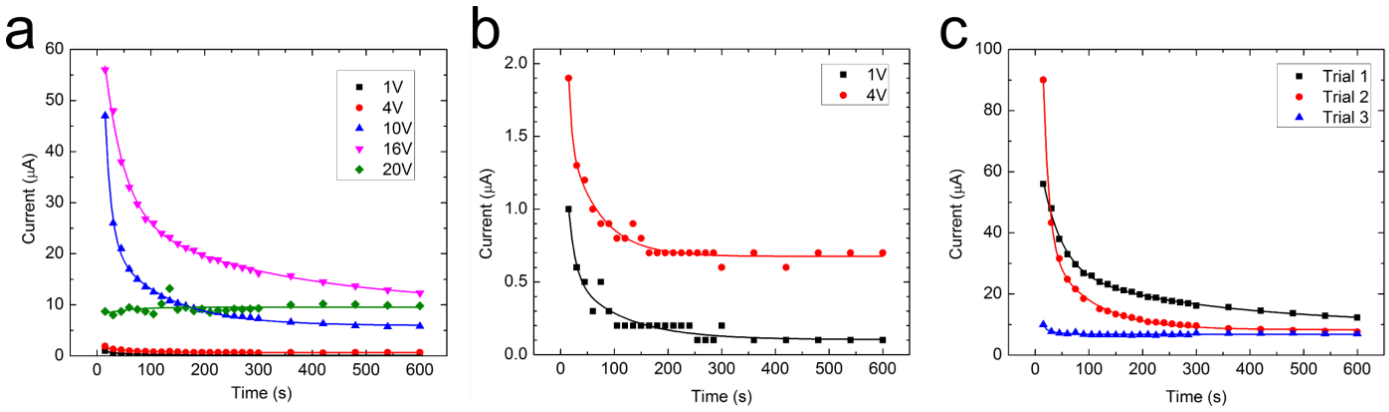


Figure 4.7 Time dependence of current through buffer on a nonconductive surface. (**a**) The current as a function of time for given potentials. (**b**) A magnified version of the lowest two potentials from (a). (**c**) Impact of damaged electrodes on measured current from 3 consecutive trials at 16 V.

measurement, and the second repetition increased by 700% from the original measurement (Figure 4.7c). Therefore, only the first trial was analyzed and presented here (Figure 4.7a-b). The current as a function of time was fit with a double-exponential, as each case exhibited a fast phase and a slow phase. The rate constants for the fast phase were $k_{buffer,fast,1V} = 0.0797\ s^{-1}$, $k_{buffer,fast,4V} = 0.1781\ s^{-1}$, $k_{buffer,fast,10V} = 0.0863\ s^{-1}$, $k_{buffer,fast,16V} = 0.0274\ s^{-1}$. The rate constants for the slow phase were $k_{buffer,slow,1V} = 0.0101\ s^{-1}$, $k_{buffer,slow,4V} = 0.0170\ s^{-1}$, $k_{buffer,slow,10V} = 0.\ 0094\ s^{-1}$, $k_{buffer,slow,16V} = 0.0034\ s^{-1}$. As most of the current decay occurred within the first 15 s and was not recorded, an accurate rate constant could not be determined from the data for 20 V. Also, because the first 15 s were not used for the other exponential decays, it is very likely that each fast-phase rate constant recorded above is a mix of the fast and slow phase, since the bulk of the fast phase, especially for the higher potentials, is not captured by the data. The currents leveled off at $I_{buffer,1V} = 0.1\ \mu A$, $I_{buffer,4V} = 0.7\ \mu A$, $I_{buffer,10V} = 5.8\ \mu A$, $I_{buffer,16V} = 12.3\ \mu A$, $I_{buffer,20V} = 9.8\ \mu A$.

Next, a potential was applied across the ATO in the absence of buffer. As suspected, the current did not change over the 10 min the current was monitored and increased linearly as a function of voltage. The currents were as follows: $I_{ATO,1V} = 0\mu A$, $I_{ATO,4V} = 0.1\mu A$, $I_{ATO,10V} = 0.2\mu A$, $I_{ATO,16V} = 0.3\mu A$, $I_{ATO,20V} = 0.4\mu A$. It should be noted that despite the exponential decay in the current through the buffer, the current through the buffer still levels off at a higher value than the current going through the ATO by itself, indicating that the resistance after 10 min remains higher in the ATO, and the majority of the current will favor the buffer over the ATO.

Finally, we studied the time dependence of the current as it passes through the ATO with buffer (Figure 4.8). It was ensured that the buffer did not touch the two electrodes so as to not damage them, resulting in consistent results across three trials. As opposed to the current

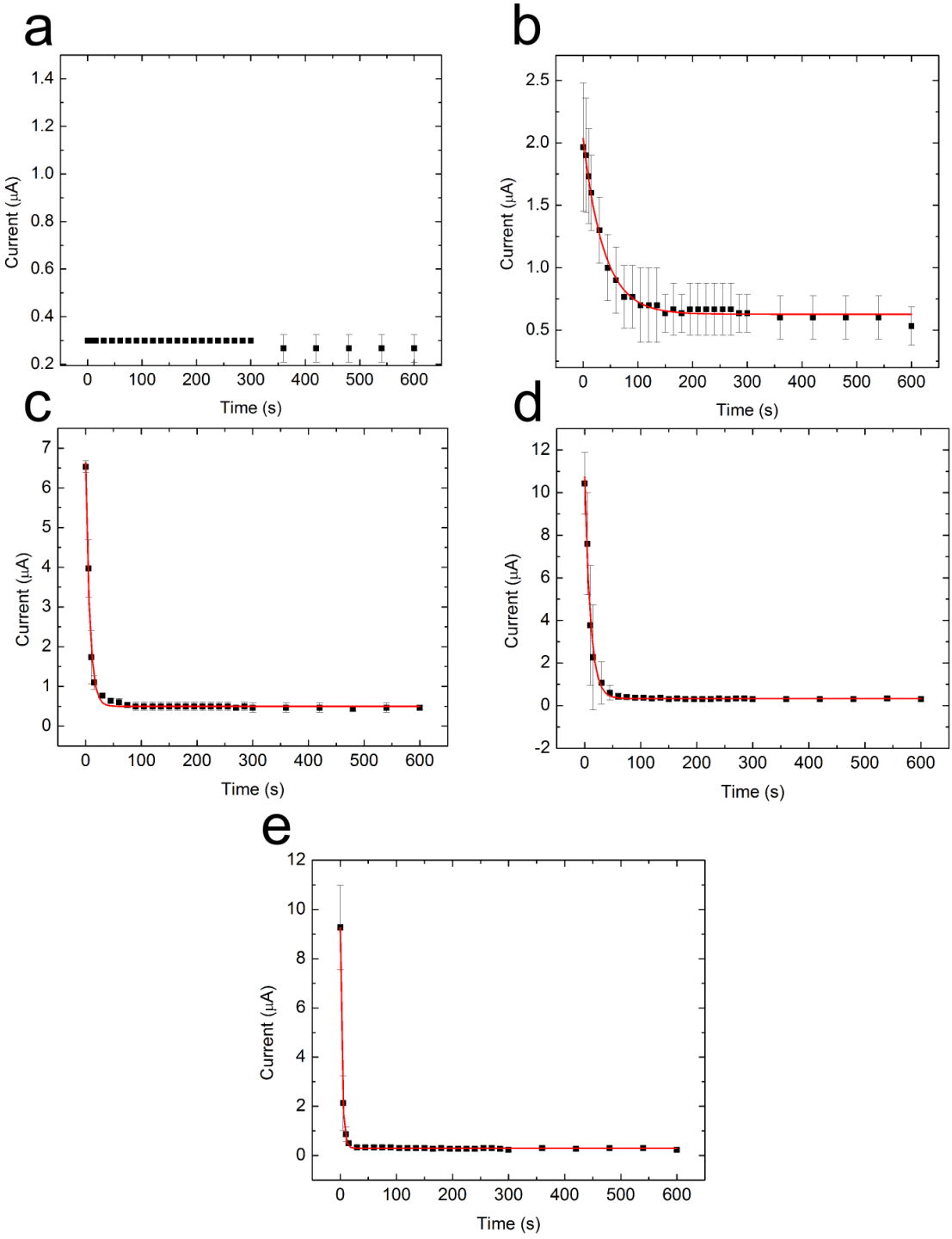


Figure 4.8 Time dependence of current through buffer-coated ATO when a potential of (**a**) 1 V, (**b**) 4 V, (**c**) 10 V, (**d**) 16 V, or (**e**) 20 V is applied.

passing through buffer alone, which requires two rate constants to describe the change in current, the change in current across ATO with buffer is fit well with a single-exponential. $k_{ATO+buffer,4V}$ = 0.0265 s^{-1}, $k_{ATO+buffer,10V}$ = 0.1384 s^{-1}, $k_{ATO+buffer,16V}$ = 0.0985 s^{-1}, $k_{ATO+buffer,20V}$ = 0.3084 s^{-1}. The decay in current in the 1 V case was so minimal that it could not be accurately determined from an exponential fit. The average value for which the current leveled off was $I_{ATO+buffer,1V}$ = 0.27 ± 0.06 μA, $I_{ATO+buffer,4V}$ = 0.54 ± 0.15 μA, $I_{ATO+buffer,10V}$ = 0.47 ± 0.06 μA, $I_{ATO+buffer,16V}$ = 0.30 ± 0.00 μA, $I_{ATO+buffer,20V}$ = 0.23 ± 0.06 μA (Figure 4.9). It is of interest to note that if the buffer was not replaced between measurements when increasing the current, as was the case for observing the diffusing Cy5-Hydrazide, the steep exponential decay was no longer observed and the values of the current measured were the leveled-off values. This implies that the exponential decay is due to the ions in the solution being drawn toward the leads and not returning to a uniform distribution throughout the solution in between observations. It is also of interest to note that none of the currents fall below the value of the current through the ATO by itself, except for the current at 20 V. The fact that the resistance of the "parallel circuit" is higher than that of ATO alone at 20 V is most likely due to bubbles formed by hydrolysis. The leveled-off values of the current follow the same trend as the diffusion coefficient for each potential (Figure 4.5b).

4.4 Discussion

To our knowledge, conductive porous materials have not previously been characterized using diffusing particles in the presence of an electric field. Here, we observed that while the alpha parameter, which provides information concerning the confinement of the particle, remained unchanged in the presence of an electric field, the diffusion coefficient was dependent on the potential applied and appeared to directly correspond to the resulting current. The force the particle experiences from the electric field would suggest that the particle should undergo

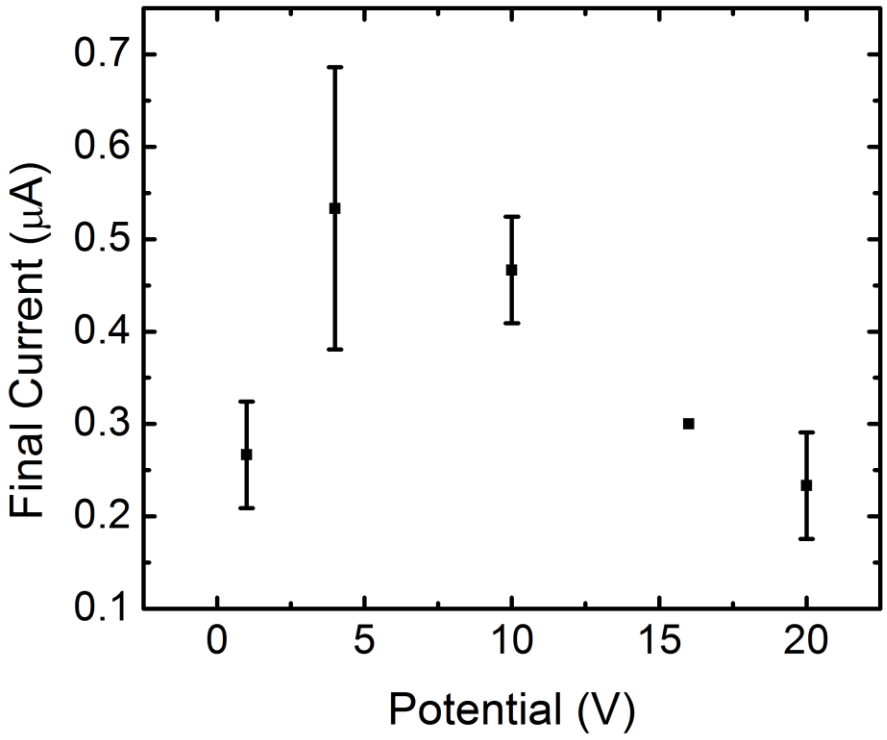


Figure 4.9 Value at which the current leveled off through ATO and buffer for each potential applied.

some directional bias in its diffusion, but no directional bias was detected. Given these results, we propose a model to interpret the causes of the diffusive motion in the nanoporous ATO material (Figure 4.10). The motion of the small charged dyes is influenced by three external forces: a normal force from the physical walls of the nanopores, a repulsive force from the negatively charged walls of the nanopores, and an external electrical force associated with the current, which passes primarily through the buffer. Even though the electric field is causing the particles to move more quickly, the repulsive force still plays a significant role in determining the direction of motion, causing the bias in motion generated by the electric field to be negligible.

ATO exhibits properties that are highly favorable for nanotechnology. It is possible to incorporate small particles of low charge into the nanopores, a conductive environment that has a high specific surface area. It is of interest to use biomaterials that require a buffer, such as DNA nanostructures, as guest molecules in these nanopores for electrochemical applications[12]. While the free salt ions of the buffer will move readily toward the leads, as was indicated by the exponential decrease in the current, the ions within the nanopores will remain confined in the nanopores, as was indicated by the presence of charged particles that remained diffusing in the pores. Also, while the resistance of the buffer and the resistance of the ATO initially differ by three orders of magnitude, once the current has leveled off, the resistances are of the same order of magnitude. This suggests that the ATO could serve as an electrode for biomaterials in a buffer environment, but it should be taken into account that the buffer will affect the electrical properties. Other factors that will affect the electrical properties of the system, such as water oxidation, must also be considered.

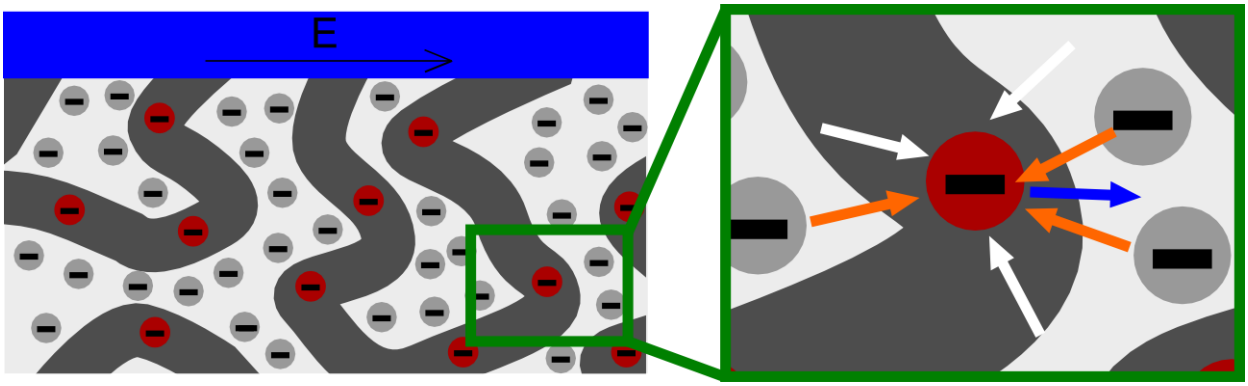


Figure 4.10 Model to interpret the diffusive motion of the negatively-charged dyes (red) within the nanopores (dark gray). There are three forces the influence the direction of movement for the particle: the repulsion force (orange) from the negatively-charged ATO (light gray), the normal force (white) from the walls of the nanopores, and the force generated by the electric field (blue).

4.5 References

1 Sharma, S., Volosin, A.M., Schmitt, D., Seo, D. Preparation and Electrochemical Properties of Nanoporous Transparent Antimony-Doped Tin Oxide (ATO) Coatings. *J. Mater. Chem. A.* **1**, 699-706 (2013).

2 Franke, M. E., Koplin, T. J. & Simon, U. Metal and metal oxide nanoparticles in chemiresistors: does the nanoscale matter? *Small* **2**, 36-50, doi:10.1002/smll.200500261 (2006).

3 Avila, L., Wirtz, M., Bunce, R. A. & Rivera, M. An electrochemical study of the factors responsible for modulating the reduction potential of putidaredoxin. *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry* **4**, 664-674 (1999).

4 Sadeh, A. *et al.* Sol-gel-derived composite antimony-doped, tin oxide-coated clay-silicate semitransparent and conductive electrodes. *Analytical chemistry* **79**, 5188-5195, doi:10.1021/ac070165r (2007).

5 Gratzel, M. Photoelectrochemical cells. *Nature* **414**, 338-344, doi:10.1038/35104607 (2001).

6 Hou, K., Puzzo, D., Helander, M.G., Lo, S.S., Bonifacio, L.D., Wang, W., Lu, Z, Scholes, G.D., Ozin, G.A. Dye-Anchored Mesoporous Antimony-Doped Tin Oxide Electrochemiluminescence Cell. *Cell. Adv. Mater.* **21**, 2492-2496 (2009).

7 Volosin, A. M., Sharma, S., Traverse, C., Newman, N., Seo, D. One-pot synthesis of highly mesoporous antimony-doped tin oxide from interpenetrating inorganic/organic networks. *J. Mater. Chem.* **21**, 13232-13240 (2011).

8 Pitchiaya, S., Androsavich, J. R. & Walter, N. G. Intracellular single molecule microscopy reveals two kinetically distinct pathways for microRNA assembly. *EMBO reports* **13**, 709-715, doi:10.1038/embor.2012.85 (2012).

9 Aitken, C. E., Marshall, R. A. & Puglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys J* **94**, 1826-1835 (2008).

10 Monnier, N. *et al.* Bayesian approach to MSD-based analysis of particle motion in live cells. *Biophys J* **103**, 616-626, doi:10.1016/j.bpj.2012.06.029 (2012).

11 Liao, Y., Yang, S. K., Koh, K., Matzger, A. J. & Biteen, J. S. Heterogeneous single-molecule diffusion in one-, two-, and three-dimensional microporous coordination polymers: directional, trapped, and immobile guests. *Nano Lett* **12**, 3080-3085, doi:10.1021/nl300971t (2012).

12 Simmons, C. R. *et al.* Size-selective incorporation of DNA nanocages into nanoporous antimony-doped tin oxide materials. *ACS nano* **5**, 6060-6068, doi:10.1021/nn2019286 (2011).

CHAPTER 5:
CONCLUSIONS AND FUTURE DIRECTIONS[6]

5.1 Conclusions

Brownian motion, the apparently random movement of small particles in a solution, is an established concept that modern-day scientists take for granted. But it is beneficial to remember that this movement puzzled the great minds of Einstein and his predecessors[1-3]. Thanks to their studies, we now have an explanation of Brownian motion based on the atomic theory of matter, and a great theoretical understanding of Brownian motion that equips us with a powerful tool to characterize a wide variety of systems, from those found in nature to synthetic systems.

Over a century after Brownian motion was first explained, we may no longer be perplexed by the random movement, but it is still an unintuitive process for us. With each new application, a different problem appears: it may be unclear how precisely to interpret the motion, or how to determine the number of particles that must be observed in order to have statistically significant results. In addition to using the features of Brownian motion to characterize various systems, this dissertation also demonstrates that Monte Carlo simulations are useful tools for building one's intuition for Brownian motion in novel applications.

In chapter 2, the software developed to help interpret the data of particle distribution across three regions of the nuclear pore complex (NPC) helped us conclude that in order to obtain the experimental histograms of the distribution of particles in each region, the particles diffusing through the central region must have a diffusion coefficient that was decreased by a

[6] Partially adapted from Michelotti, N., Johnson-Buck, A., Manzo, A.J., M. and Walter, N.G. (2012) Beyond DNA origami: The unfolding prospects of nucleic acid nanotechnology. *WIREs: Nanomed Nanobiotechnol*. **4**, p. 139-152.

factor of 7.8 for the mRNA-complex and GFP, and 12.9 for Tap-p15, compared with the cytoplasmic and nucleoplasmic regions. Similar distributions of particles had been obtained previously[4], but the conclusion was that the diffusion coefficient was increased through the central region. This would be an intuitive assessment of the data, since fewer particles in the central region would imply a shorter dwell time and thus a higher diffusion coefficient. However, this interpretation of the data neglects the differences in volume among the three regions. This volume difference is the greatest contributor to the observed bimodal distribution of particles and can thus not be neglected in the interpretation of the data.

While models may be helpful for building our intuition of systems, they can also be misleading if inadequate assumptions are made. Chapter 3 shows how using estimated rate constants instead of experimentally measured rate constants[5,6] in a simulation led to a predicted diffusion coefficient for the spider that was 20 times higher than that predicted when using experimentally determined values for the diffusion coefficients based on ensemble assays performed in solution. In performing our own single-molecule assay to determine the rate constants, we discovered that legs had a strong tendency to aggregate in the presence of 1 mM $ZnSO_4$. When the experimental conditions were changed so that magnesium was the cofactor, the diffusion coefficient was increased only by a factor of 1.5. Two single-molecule fluorescence experimental techniques were developed to test the accuracy of the simulations. The simulations predicted that 31% of the track of substrates would be cleaved in 10 mM $MgCl_2$. The two experimental techniques resulted in 31% and 34% of the track being cleaved, values that are comparable to the predicted values.

Single-molecule techniques provide an valuable method for observing the behavior of particles individually. It is tempting, however, to put too much emphasis on individual particles

and the heterogeneity of particles. Examples of studying Brownian motion in the literature have isolated particles with faster or slower diffusion coefficients as extracted from time-averaged MSD (TAMSD) plots for individual particles, or have tried to explain the heterogeneity as abnormalities in the surface upon which the particles are diffusing[7]. However, in chapter 4, we show using Monte Carlo simulations of a random walk that the TAMSD plots for individual molecules will be heterogeneous even in an ideal system, due to the stochasticity of the process. The heterogeneity has already been described by Einstein and thus requires no further explanation. Instead of focusing on individual particles, we looked at an ensemble of the particles to characterize the nanoporous antimony-doped tin oxide. It was found that the motion within the nanopores is confined, with an alpha parameter of 0.69 ± 0.05. The diffusion coefficient is dependent on the potential applied across the ATO. The relationship between the diffusion coefficient and the potential applied is consistent with the notion that the current passing through the system directly affects the diffusion coefficient.

Brownian motion may be a powerful tool for characterizing systems, but the data can be easily misinterpreted. The broader use of Monte Carlo simulations by experimentalists to interpret results would be beneficial.

5.2 Future

In chapter 2, the software for estimating the distribution of particles in the NPC was designed to capture the data in the same way as the experimental results. This gives a steady-state distribution that does not include temporal information. It was also designed so as to have the particles undergo random Brownian motion even though it is well-established that the transport through the central pore complex is directional. Using time-sensitive data acquired by those who have done single-particle tracking of various particles as they were being transported

through the central region, we could incorporate these rates into the simulations. We could also incorporate a bias by adding an effective force term to the simulation, $\frac{P_+}{P_-} = exp\left(-\frac{F\delta}{k_BT}\right)$ where P_+ is the probability of the particle moving toward the nucleoplasmic side, P_- is the probability of the particle moving toward the cytoplasmic side, F is the effective force applied to create this bias, and δ is the step size[5,6]. As a longer-term future direction, just as computational models are helpful for understanding an experimentlal system, synthetic models of nature can also be helpful. A DNA model of the central region of the NPC could be generated using the recently developed three-dimensional DNA tunnels[8]. By extending given strands in the DNA tunnel to generate overhangs that mimic the FG Nups in the central region, short DNA strands of ~7 nucleotides complementary to the overhangs diffusing in solution could be used to mimic particles diffusing through the NPC to model how the interaction between the FG Nups and the diffusing particles aids the bias in the transport of particles through the NPC. These oligonucleotides could then be attached to proteins to see how well they are able to chaperone proteins through the channel.

Chapter 3 characterized a DNAzyme-based nanowalker. Such directed walkers have already aided in the combinatorial assembly of different-sized gold nanoparticles[9] or the multistep synthesis of a small organic compound[10], in the former case aided by two-state DNA conformational switches (Figure 5.1). Two future applications for the spider, and the work that has been done toward these applications, are described in detail in Appendices A and B. The first is a system in which the outcome of the system is determined by the competition between two spiders. Two racetracks for two different spider species that share a common finish line are patterned on an origami tile. The finish line bears a fluorescently labeled "Signal" oligonucleotide that bears the sequence for the substrate of spider 1 on the 5' end, and a sequence

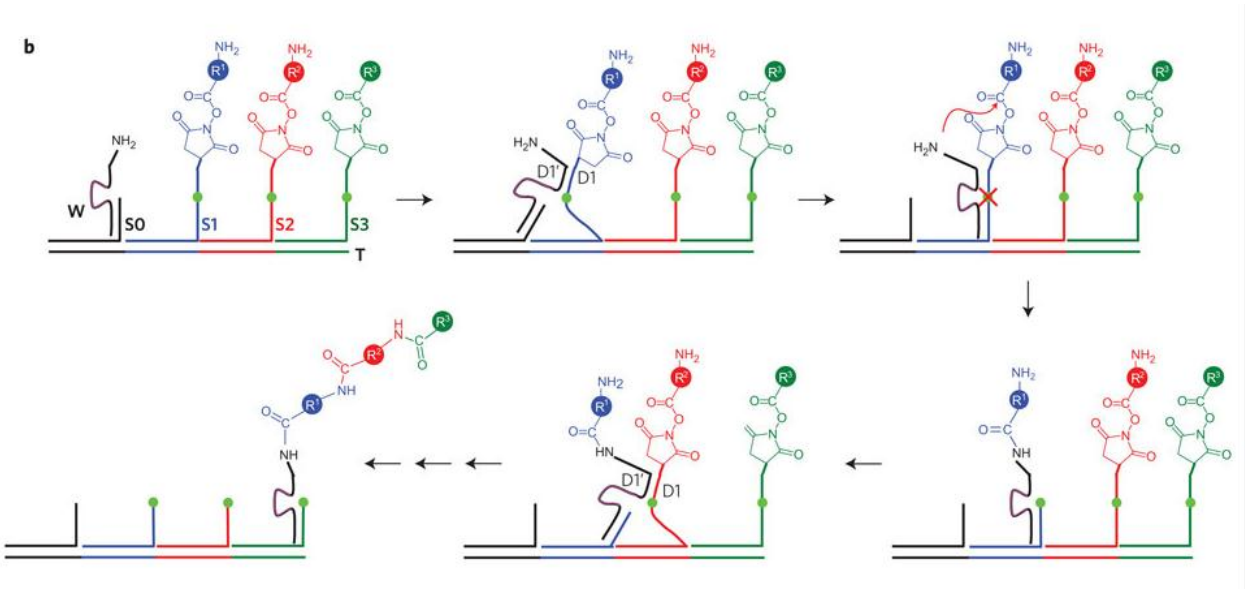


Figure 5.1 Multistep synthesis of an organic compound mediated by a deoxyribozyme. Reprinted by permission from Macmillan Publishers Ltd: *Nat. Nano.* (Ref. 10), copyright 2010.

that is fully complementary to the leg of spider 2 on the 3' end. If spider 1 reaches the end before spider 2, the Signal is cleaved and diffuses in solution, while if spider 2 reaches the end first, it remains tethered to the surface by the leg of spider 2 (Appendix A).

The second application is the MacGyver nanopatterner that exploits the product left behind by the spider as it cleaves its substrate in order to pattern a surface with nanometer precision. The motion of spiders along a substrate-coated surface can be biased by applying a voltage across the fluidic channel. Thus, the trail that the spider traverses can be controlled remotely. The trail can be detected by borrowing probing techniques from the biosensor community. The MacGyver nanopatterner only requires the use of materials found around a typical wet lab, greatly reducing the cost compared to standard lithographic techniques used for microfabrication (Appendix B).

Chapter 4 characterizes a conductive nanoporous material that could be used as an electrode for nanoelectronic applications. Three-dimensional static DNA structures have already been incorporated in these conductive nanopores[11]. Using DNA origami as a template, bioelectroactive enzyme cascades, such as the GOx-HRP and methanol dehydrogenase cascade, could be incorporated into the ATO nanopores to observe the fundamentals of the reaction. As another potential application, conductive static structures patterned via DNA could be embedded in the nanopores to be used in plasmonic applications. These nanoparticles placed in close proximity (spacing <2.5 times the diameter of the nanoparticle) undergo strong near-field coupling with field enhancements[12]. However, the optimal field enhancement is dependent on the geometry of the nanoparticle array, requiring precise nanoparticle placement and spacing[12,13]. To this end, gold nanoparticles have been assembled on DNA origami triangles[14,15]. While DNA origami tiles are advantageous in their rigidity and addressability, they are limited in length

(~100 nm) unless concatenated, in turn limiting the propagation of the plasmon resonance signal. In the future, longer scaffolds such as the nanopeapods[16] or DNA origami nanotubes[17] with lengths on the order of tens of microns, may be paired with top-down microfabrication techniques to construct larger nanoparticle arrays with precision and spacing on the order of tens of nanometers. Applications will likely include communication systems that merge electronics and photonics (optics) at the nanoscale[13].

5.3 References

1 Einstein, A. On the Motion of Small Particles Suspended in a Stationary Liquid, as Required by the Molecular Kinetic Theory of Heat. *Annalen der Physik* **17**, 549-560 (1905).

2 Perrin, J. Brownian Motion and Molecular Reality. *Annales de Chime et de Physique* **18**, 1-114 (1909).

3 Brown, R. A Brief Account of Microscopical Observations Made in the Months of June, July, and August, 1827 on the Particles Contained in the Pollen of Plants and on the General Existence of Active Molecules in Organic and Inorganic Bodies. *unpublished* (1827).

4 Grunwald, D. & Singer, R. H. In vivo imaging of labelled endogenous beta-actin mRNA during nucleocytoplasmic transport. *Nature* **467**, 604-607, doi:10.1038/nature09438 (2010).

5 Samii, L. *et al.* Time-dependent motor properties of multipedal molecular spiders. *Physical review. E, Statistical, nonlinear, and soft matter physics* **84**, 031111 (2011).

6 Samii, L., Linke, H., Zuckermann, M. J. & Forde, N. R. Biased motion and molecular motor properties of bipedal spiders. *Physical review. E, Statistical, nonlinear, and soft matter physics* **81**, 021106 (2010).

7 Liao, Y., Yang, S. K., Koh, K., Matzger, A. J. & Biteen, J. S. Heterogeneous single-molecule diffusion in one-, two-, and three-dimensional microporous coordination polymers: directional, trapped, and immobile guests. *Nano Lett* **12**, 3080-3085, doi:10.1021/nl300971t (2012).

8 Ke, Y., Ong, L. L., Shih, W. M. & Yin, P. Three-dimensional structures self-assembled from DNA bricks. *Science* **338**, 1177-1183, doi:10.1126/science.1227268 (2012).

9 Gu, H., Chao, J., Xiao, S. J. & Seeman, N. C. A proximity-based programmable DNA nanoscale assembly line. *Nature* **465**, 202-205, doi:10.1038/nature09026 (2010).

10 He, Y. & Liu, D. R. Autonomous multistep organic synthesis in a single isothermal solution mediated by a DNA walker. *Nat Nanotechnol* **5**, 778-782, doi:10.1038/nnano.2010.190 (2010).

11 Simmons, C. R. *et al.* Size-selective incorporation of DNA nanocages into nanoporous antimony-doped tin oxide materials. *ACS nano* **5**, 6060-6068, doi:10.1021/nn2019286 (2011).

12 Tan, S. J., Campolongo, M. J., Luo, D. & Cheng, W. Building plasmonic nanostructures with DNA. *Nat Nanotechnol* **6**, 268-276, doi:10.1038/nnano.2011.49 (2011).

13 Ozbay, E. Plasmonics: merging photonics and electronics at nanoscale dimensions. *Science* **311**, 189-193, doi:10.1126/science.1114849 (2006).

14 Hung, A. M. *et al.* Large-area spatially ordered arrays of gold nanoparticles directed by lithographically confined DNA origami. *Nat Nanotechnol* **5**, 121-126, doi:10.1038/nnano.2009.450 (2010).

15 Ding, B. *et al.* Gold nanoparticle self-similar chain structure organized by DNA origami. *Journal of the American Chemical Society* **132**, 3248-3249, doi:10.1021/ja9101198 (2010).

16 Lo, P. K. *et al.* Loading and selective release of cargo in DNA nanotubes with longitudinal variation. *Nature chemistry* **2**, 319-328, doi:10.1038/nchem.575 (2010).

17 Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302 (2006).

APPENDIX A:
SPIDER RACES

A.1 Introduction

We recently demonstrated the concept of molecular robotics with a synthetic DNA-based nanowalker, dubbed a "spider," composed of a streptavidin protein "body" attached to three biotinylated deoxyribozyme legs, moving along a one-dimensional track of chimeric DNA-RNA substrates positioned on a DNA origami[1,2]. By cleaving its substrates, the spider weakens the binding affinity between its legs and previously visited sites, resulting in a biased, processive, random walk towards fresh substrate. Additional components are now being incorporated into the spider world to increase its versatility and complexity in behavior. For example, we have designed and characterized a second spider (Spider 2) that walks on an orthogonal substrate, meaning that the leg of Spider 2 is not complementary to the substrate of Spider 1. Monte Carlo simulation results based on kinetic parameters from the spider legs suggest that changes in the buffer conditions can favor the movement of one spider at the expense of the other, allowing us to bias the outcome of the race. This advance will allow us to devise a controlled "spider race" towards a common finish line; depending on the predetermined winner, a fluorescent marker will either be released or not. While this track and components have been designed and fabricated, there are still several hurdles ahead in this spider race.

A.2 Materials and Methods

Design of Spider 2 leg. There were two potential sequence designs for the legs of Spider 2: the first was an 8-17 deoxyribozyme selected for by Schlosser et al[3] (Leg 2A). The second was based on Leg 2A, but was modified using Visual OMP (DNA Software) predictions in order to minimize undesired secondary structure and incorrect binding and to have the same binding energies as the leg of Spider 1 (Leg 2B).

Multiple-Turnover Cleavage Rate Assay. All oligonucleotides were ordered from IDT except the Signal strand that was ordered from Keck. The cleavage rate constant was determined for three different deoxyribozyme legs: The leg of Spider 1 (Leg 1) [5'- CTC TTC TCC GAG CCG GTC GAA ATA GTG AAA A], Leg 2A [5'- GTA ACG CAC TGT CAG CGA CTC GAA ATC TCT CTC], and Leg 2B [5'-CCA TCC TGT CAG CGA CTC GAA TTT CCC AC]. The sequences of their respective substrates were Substrate 1 [5'- Cy5/TTT TTC ACT ATrA GGA AGA G], Substrate 2A [5'- Cy5/AGA GAG AGA TrGG GTG CGT TAC], and Substrate 2B [5'-Cy5/AGT GGG AAA rGGG GAT GG].

1 μM of each deoxyribozyme leg incubated with 20 μM of its respective Cy5-labeled substrate for 5 min. For Leg 1, the reaction buffer, which will be referred to as Buffer 1, was 10 mM HEPES, 150 mM NaCl, 1 mM $ZnSO_4$, pH 7.4. For Legs 2A and 2B, the reaction buffer, Buffer 2, was 400 mM KCl, 100 mM NaCl, 7.5 mM $MgCl_2$, 7.5 mM $MnCl_2$, 50 mM HEPES-KOH, pH 7.0. Before adding the reaction buffer, a fraction of the leg-substrate complex was placed in EDTA followed by $ZnSO_4$ as a zero time point. The reaction buffer optimized for each specific deoxyribozyme was added to the oligos, diluting their concentrations to 0.5 μM leg and 10 μM substrate during the reaction. The reaction was stopped at various time points (0.5, 1, 2, 4, 8, 16, 32 min.) by placing 2 μL of the reaction into 8 μL of 75 mM EDTA.

The cleavage rate constants were determined by quantifying bands from a 20% denaturing Polyacrylamide Agarose Gel Electrophoresis (PAGE) gel. The gels were scanned for fluorescence signal using the Typhoon fluorescence scanner (GE Healthcare Life Sciences). The bands were quantified using ImageQuant (GE Healthcare Life Sciences). The fraction cleaved was calculated by taking the volume of the lower band (the product) and dividing it by the volume of the lower band plus the upper band (the product + the substrate).

Track Design. The track design is based on Rothemund's rectangular origami[4]. Overhang sequences were added to the 5' end of the staple strands to generate a template for complementary strands to bind, to create tracks with nanometer-scale precision. The new sequences are as follows: Start 1 [5' – GAT GTC TAC TTG CGT CAG GTT CTC GGC - staple], Start 2 [5' – GTG GTG TGA CTT GAG ACT TGT GAT GTG - staple], Substrate 1 [5'- (GAT GAA TGG TGG GTG AGA GG)T TTT TCA CTA TrAG GAA GA], Substrate 2 [5' - (TGT GGA TTG GTT GAA CTG TT)A GAG AGA GAT rGGG TGC GTT AC], Goal 1 [5' – (CTG GCT CAA CGA ACT GAA CC)T TTT TCA CTA TAG GAA GAG], Goal 2 [5' – (GAG ACT GAG AGA CAA AGG TA)G AGA GAG ATG GGT GCG TTA C], Signal [5' – (AAG TAT CGT ACG ACT AGA C)TT TTT CAC TAT rAGG AAG AGG AGA GAG ATT TCG AGT CGC TGA CAG TGC GTT ACA TG/3AmMO]. The sequences in parenthesis designate the portion of the sequence complementary to the overhang used to attach the sequences to the origami.

Monte Carlo Simulations. Monte Carlo simulations were written in MATLAB (Mathworks) to determine the probability of the spider race outcome given the various buffer conditions. The

logic of the simulation was based on probabilities of reactions occurring generated from experimentally determined rate constants[5]. The potential reactions were binding to substrate ($k = \infty$), cleaving substrate (k varies as specified), dissociating from substrate ($k = 0.0014\ s^{-1}$), binding to product ($k = \infty$), and dissociating from product ($k = 0.14\ s^{-1}$). The dissociation rate constants were assumed to be the same for Leg 2. The probability of a given reaction occurring at a given time point is determined using the appropriate rate constants in the equation $P(t) = 1\text{-}e^{-kt}$. The possible reactions were binding to substrate, cleaving substrate, dissociating from substrate, binding to product, and dissociating from product. In the Monte Carlo simulation, if the given probability of a reaction occurring was above a randomly generated number between 0 and 1, and if the legs were in a position where that outcome could occur, the reaction took place. The probability of the leg binding to substrate or product was assumed to be one such that when a leg dissociated from a substrate or product, it bound to a substrate or product at random without having to go through an additional iteration of the program. It should be noted that the Gillespie Algorithm was not implemented in these simulations.

A.3 Results

In order for two spiders to interact to accomplish a task, a track needed to be designed in which their environments were interwoven. We chose to design a track in which the outcome of a competition between two spiders would generate a given output (Figure A.1A). Each spider consists of a streptavidin protein body bound to three biotinylated deoxyribozyme legs and one Capture strand, a DNA sequence used to immobilize the spider on the origami. Thus, the Capture strands were made complementary to their respective Start strands located on the origami (the Capture strand of Spider 1 and Spider 2 would be used to immobilize on Start 1 and Start 2, respectively). Substrate 1 and Substrate 2 that make up the spider's race tracks consist of

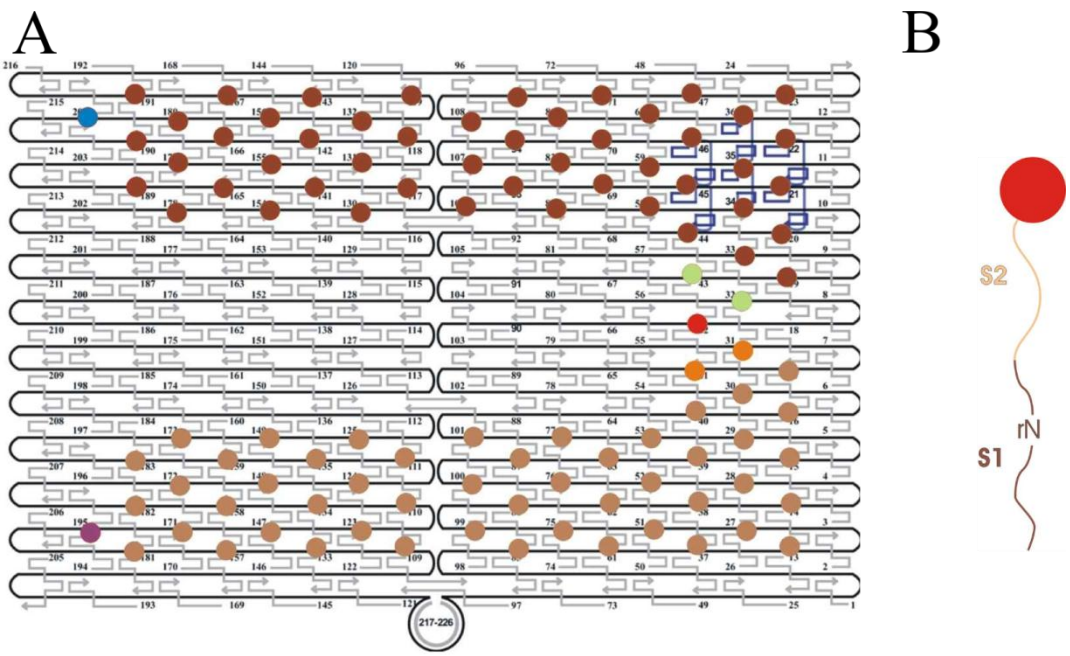


Figure A.1 Spider racetrack schematic. (A) Spider 1 and Spider 2 begin their race at Start 1 (blue) and Start 2 (purple) respectively. Each spider's track is composed of its substrate, Substrate 1 (dark brown) and Substrate 2 (peach). At the end of the track are noncleavable substrates that lack the RNA base necessary for cleavage, Goal 1 (green) and Goal 2 (orange), and the Signal (red). (B) The 5' end of the Signal strand contains the same sequence as Substrate 1 (brown) while the top of the Signal strand is fully complementary to the leg of Spider 2 (peach). The 3' end is labeled with a fluorescent probe (red).

orthogonal substrate sequences that are partially complementary to Leg 1 and Leg 2, respectively. The tracks for both spiders are of the same length. At the end of the track are noncleavable strands, Goal 1 and Goal 2, to ensure that the spider does not leave the end of the track once it has won the race. The Signal strand is a fluorescently-labeled DNA-RNA strand. The 5' end of the Signal strand (the portion closer to the origami) contains the same sequence as Substrate 1 while the 3' end (the portion farther from the origami) is fully complementary to Leg 2 such that Leg 2 is unable to cleave the Signal strand. Therefore, if Spider 1 wins the race, the fluorescently labeled Signal is cleaved and diffuses into solution. This is detected on a single-molecule fluorescence level by the loss of fluorescent signal. If Spider 2 wins the race, the Signal remains tethered to the surface by Spider 2's leg even if Spider 1 arrives at a later time. The winner of the race is biased by the buffer condition used. If Buffer 1 is used, Leg 1 should have a higher cleavage rate constant, resulting in it being the victor, and similarly for Buffer 2 and Leg 2. To bring the spider race from concept to existence, a second spider with orthogonal reaction conditions was fabricated. It was also verified through simulations that a higher rate constant would theoretically bias the race outcome in the way proposed.

In designing a new spider, two different leg designs were tested for cleavage rate: Leg 2A based on a previously determined deoxyribozyme[3] and Leg 2B which was a modified version of Leg 2A in which the leg sequence was modified to replicate the binding energy for Leg 1. Leg 1 and Leg 2A had similar cleavage rate constants of 5.06 min^{-1} and 4.06 min^{-1} (Figure A.2A). Leg 2B was found to have a lower cleavage rate constant of 3.28 min^{-1} and the extent of cleavage was less than 50% after 32 min. Given this striking difference in activity between deoxyribozymes with two very similar sequences (Leg 2A and Leg 2B), Leg 2A was chosen for the leg of Spider 2 and will henceforth be called Leg 2.

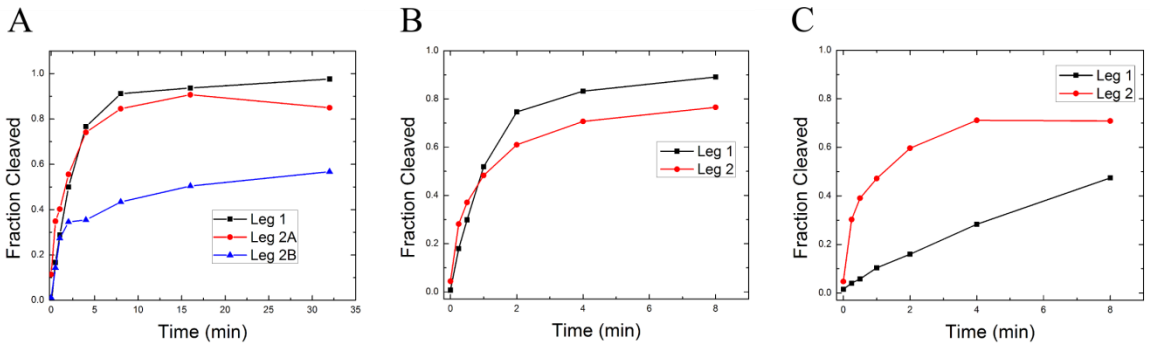


Figure A.2 Cleavage rates of spider legs. (A) Fraction of substrates cleaved in the presence of an excess of its substrate under each leg's optimized buffer conditions. (B) Fraction of substrates cleaved under multiple-turnover conditions in Buffer 1. (C) Fraction of substrates cleaved under multiple-turnover conditions in Buffer 2.

The cleavage activity for Leg 1 was selected for in the presence of 1 mM Zn^{2+} and is thus optimized for that divalent metal ion[6]. Leg 2 was optimized for 7.5 mM Mg^{2+} and 7.5 mM Mn^{2+}; thus the two legs are optimized under different buffer conditions[3]. To determine how effective each leg was at cleaving its respective substrate under the other's buffer conditions, we performed a cleavage assay for each spider under the other's buffer conditions (Figure A.2B). In 1 mM Zn^{2+}, Leg 1 and Leg 2 had comparable cleavage rate constants of 10 min^{-1} and 8 min^{-1} respectively. In 7.5 mM Mg^{2+} and 7.5 mM Mn^{2+}, Leg 1 had a cleavage rate of 1.74 min^{-1}, while Leg 2 had a cleavage rate of 8 min^{-1}. The activity of Leg 1 was greatly impaired in Buffer 2. While the opposite is not true, the probabilities of outcomes can still be taken into account: in Buffer 1, the probability of the two outcomes will be more equal to one another while in Buffer 2, Spider 2 will be favored.

To obtain a better understanding of the proposal that a higher cleavage rate would increase a spider's speed and to gain a sense of winner probabilities, Monte Carlo simulations were implemented. A decrease by a factor of 2 in the cleavage rate demonstrated greatly impaired motion of the spider (Figure A.3A). By doubling the cleavage rate, Spider 2 can arrive at the end position a few minutes before Spider 1 in a ~30 min race, but increasing it further did not greatly benefit influence the outcome. This is confirmed by plotting the probability of the race outcome as a function of the cleavage rate (Figure A.3B). Given the cleavage rate for the spiders under Buffer 2, the significant decrease in the cleavage rate of Spider 1 should be sufficient to greatly bias the race outcome.

A.4 Discussion

With the details of the racetrack oligonucleotides designed and Leg 2 chosen, the next step of this design is to move from schematics to assembly. The origami were fabricated by Dr.

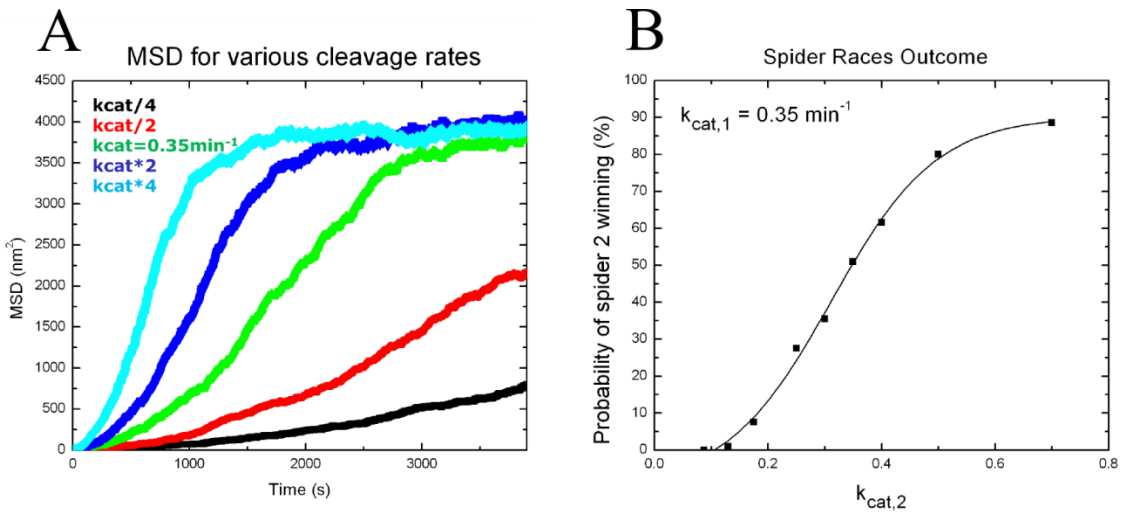


Figure A.3 Monte Carlo spider race simulations. (A) Ensemble mean-squared-displacement of the spider movement along the race track for various cleavage rate constants. (B) Winner outcome probability dependence on the cleavage rate constant of Spider 2.

Jeanette Nangreave from Dr. Hao Yan's lab at Arizona State University and the spiders were fabricated by Dr. Steven Taylor from Dr. Milan Stojanovic's lab at Columbia University. We encountered several challenges upon receiving the samples. The initial problem was that the origami would not bind to the surface. After several trials, an additional thymine spacer was added to give the biotinylated strand more flexibility to find the NeutrAvidin-coated surface[7]. The next hurdle was getting the spider to bind to the origami at high yield. In previous work[1], the spider would occasionally bind to the origami, but a high yield was never achieved. This remains true. Both the spider and origami are large and cumbersome. Previously published work accomplished this binding by incorporating additional Start strands into the origami design[8]. Even in adding these strands and using similar conditions, we obtained a very low binding yield.

For a statistically significant conclusion that is dependent on probabilities like the outcome of the spider race, high yield is important. Because we only intend to detect two different types of fluorophores for a given experiment, we would only be able to fluorescently label two different components of the race at a time and would need to rely on those high yields to determine that all the components needed for the race are in place, since we would not be able to detect all the components at once. If very few origami have Spider 1 on them, there is a low probability of each origami having two spiders immobilized on the surface let alone a Signal.

A.5 Acknowledgements

I would like to thank Jeanette Nangreave from Hao Yan's lab at Arizona State University for her help in designing and preparing the origami. I would also like to thank Steven Taylor from Milan Stojanovich's lab at Columbia University for the fabrication of the spiders. And

finally, I would like to thank all the rotation students who have worked with me on this project including Devon Triplett in the Signal strand design, Berhane Assefa in the cleavage assays, and Sherry Xie for the single-molecule cleavage assays (data not shown).

A.6 References

1 Lund, K. *et al.* Molecular robots guided by prescriptive landscapes. *Nature* **465**, 206-210, doi:Doi 10.1038/Nature09012 (2010).

2 Michelotti, N., de Silva, C., Johnson-Buck, A. E., Manzo, A. J. & Walter, N. G. A Bird's Eye View: Tracking Slow Nanometer-Scale Movements of Single Molecular Nano-Assemblies. *Method Enzymol* **475**, 121-148, doi:Doi 10.1016/S0076-6879(10)75006-0 (2010).

3 Schlosser, K., Gu, J., Sule, L. & Li, Y. F. Sequence-function relationships provide new insight into the cleavage site selectivity of the 817 RNA-cleaving deoxyribozyme. *Nucleic Acids Res* **36**, 1472-1481, doi:Doi 10.1093/Nar/Gkm1175 (2008).

4 Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302 (2006).

5 Taylor, S. Unpublished. (2005).

6 Li, J., Zheng, W., Kwon, A. H. & Lu, Y. In vitro selection and characterization of a highly efficient Zn(II)-dependent RNA-cleaving deoxyribozyme. *Nucleic Acids Res* **28**, 481-488, doi:gkd147 [pii] (2000).

7 Jungmann, R. *et al.* Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami. *Nano Lett* **10**, 4756-4761, doi:10.1021/nl103427w (2010).

8 Fu, J., Liu, M., Liu, Y., Woodbury, N. W. & Yan, H. Interenzyme substrate diffusion for an enzyme cascade organized on spatially addressable DNA nanostructures. *J Am Chem Soc* **134**, 5516-5519, doi:10.1021/ja300897h (2012).

APPENDIX B:
MACGYVER NANOPATTERNER

B.1 Introduction

Current lithographic top-down fabrication methods are able to serially generate structures with features typically on the order of microns[1]. These micropatterning techniques are expensive, requiring the use of large equipment and clean room facilities. Even top-down DNA-based approaches to nanopatterning involve the use of customized equipment and is a serial process[2].

A DNAzyme-based molecular walker called a "spider" has previously been developed that modifies its own environment by cleaving its track composed of DNA-RNA substrates[3,4]. In this sense, the spider is a nanopatterner, patterning the surface as it walks along it. Thus far, this property of the spiders has not been fully exploited. One reason is that if the spider is placed on a 2D surface saturated with substrates, the trajectory of the spider is random, therefore severely limiting the usefulness of this property. A means for user-controlled biasing of the spider motion needs to be implemented. Also, the pattern is of little use if it can't be detected. A means for acquiring the data processed is needed. To this end, an electric field has been applied across the fluidic channel to behave as a remote control, and two postprocessing imaging techniques have been designed based on various biosensors.

This nanopatterning technique requires only materials commonly found around a wet lab: a low voltage power supply, platinum wire from a broken gel apparatus, insulated copper wires, NeutrAvidin-coated slides, coverslips, double-sided sticky tape, epoxy, biotinylated DNA-RNA,

streptavidin, DNA, and buffer. Hence, this nanopatterner is named after the well-loved master of piecing household items together to engineer anything, MacGyver. The MacGyver nanopatterner makes this parallel processing fabrication technique far more economical than previous methods, but has some optimization to go through before it can be widely utilized.

B.2 Materials and Methods

NeutrAvidin Slide Preparation. This technique has been previously described in detail[4]. Two 1-mm holes were drilled in a quartz slide (Finkenbeiner, fused silica, 1"x3"x1mm) to be used for buffer exchange. The slides were submerged in a boiling "piranha" solution (5% (v/v) ammonium hydroxide (Aldrich, 231-704-8) and 14% (v/v) hydrogen peroxide (Acros, 202465000) for 20 min. The slides were rinsed with deionized water, and flamed with a propane torch for 1 min each. The slides were sonicated in 1 M potassium hydroxide (KOH; Aldrich, 215813) for 30 min. After rinsing with deionized water, and acetone, they were aminosilanized by being immersed in a 5% (v/v) 3-aminopropyltriethoxysilane (APTES; Sigma-Aldrich, A3648-100ML) acetone solution for 1 h. They were rinsed with acetone and dried at 80°C for 1 h. Para-diisothiocyanatobenzene (PDTIC; Acros, 417510050) was covalently coupled to the aminosilanized surface[5] by incubating the slides in 0.2% (w/v) PDTIC and 10% (v/v) pyridine (Fisher, P368-500) in dimethylformamide (Acros, spectroscopic grade, 40832-5000) for 2 h. The slides were rinsed with acetone and methanol (Fisher, A452-4), and allowed to dry. The slides were then coated with 0.5 mg/mL NeutrAvidin (Sigma-Aldrich, A9275-2MG) for 2 h. The slides were rinsed with deionized water. To quench the PDTIC not bound to avidin, 1 M NaCl and 40 mM NaOH was applied to the slide surface for a few minutes. Slides were rinsed with deionized water and dried with nitrogen. A flow channel between the drilled holes on each

slide was constructed using two pieces of double-sided tape (Scotch, permanent, ½ inch) spaced 2-3 mm apart. A coverslip was placed over the tape to cover the channel. Epoxy glue (Hardman Adhesives 04001) was added along the edges of the coverslip to seal the channel. These NeutrAvidin-coated slides were stored at 4°C in a desiccator for a maximum of four weeks.

Applying a Voltage Across Fluidic Channel. Fluidic wells that were cut from the thicker end of two pipette tips (Eppendorf, 05-403-33) were attached to the slide using epoxy (Hardman Adhesives, 04001) so as to surround the holes on the slides. Platinum wire was taken from a broken gel apparatus. Due to the limited supply of platinum wire and to insulate the extended wire, the wire was attached to an insulated copper wire (Mouser, 8523 013100) in order to connect the buffer in the wells to the low voltage power supply (Elenco Precision, XP-660).

Electroosmotic flow control. The fluidic channel was filled with 1XHBS (10 mM HEPES, 150 mM NaCl, pH 7.4) along with the two wells, 80 µL of 1XHBS being placed in each. 1 µL of 1:100 diluted blue food coloring was added to one well while 1 µL of 1:100 diluted yellow food coloring was added to the opposite well. The positive lead was placed in the well containing the blue food coloring while the negative lead was placed in the well containing the yellow food coloring. 20 V was applied across the channel. See Figure B.1B.

Imaging beads or spiders in the presence of an electric field. 1XHBS was flowed through the NeutrAvidin-coated channel. 10 µM biotinylated substrate was flowed through the channel and incubated for at least 30min. Excess substrate was flushed out with 1XHBS. A 1:2,000 dilution of TetraSpeck Microspheres (Invitrogen, T-7279) was introduced to the channel and settled for 5

min. These microspheres will henceforth be called fluorescent beads. Excess beads were flushed out. For experiments when spider movement was being observed, the beads were bleached for one hour so the fluorescent signal was comparable to the Cy3-labeled spider. 50 pM spiders were added and incubated for 15min and the excess flushed out with 1XHBS. Reaction buffer (10 mM HEPES, 150 mM NaCl, 1 mM $ZnSO_4$, pH 7.4) was flowed through the channel and added to each reservoir. The platinum end of the platinum-copper wires was placed in the wells. The experiment was imaged using the Prism-type TIRF microscope described in chapter 3. When the spiders were observed, the illumination was shuttered so the slide was exposed to the laser for 0.5-s increments every 30 s. When the beads trajectories were observed, they were illuminated every 15 min. for 6 hrs. For the first two hours, the power supply was off. For hours 2-4, the power supply was turned on. For hours 4-6, the electrical polarity of the leads was switched.

Hairpin Imaging Strand. Oligonucleotides were designed such that there was enough sequence in the loop region that was complementary to the leg, enabling the leg to open the hairpin and cleave the sequence. This made a region that was complementary to the Probe strand accessible to permit the Probe to bind. The sequence complementary to the Probe would need to remain completely hidden until after the reaction had occurred. The sequences for this technique include the following: Leg [5' – CTC TTC TCC GAG CCG GTC GAA ATA GTG AAA A], Hairpin [5' – TGA CTC TTC TTT TCA CTA TrAG GAA GAG TCA AG/3Cy3Sp], Probe [5' -/5Cy5/CTT GAC TCT TCC]. All strands were ordered from IDT.

Binding Detection Assay. 20 μM Hairpin was incubated with either 1 μM Leg (for the multiple-turnover reaction) or 20 μM Leg (for the single-turnover reaction) for at least 5 min. Before

adding the reaction buffer, a small volume of the Hairpin-Leg was placed in 75 mM EDTA and 2 μM Probe for the zero min time-point. A small volume of Reaction buffer was then added. The Reaction buffer was added to initiate cleavage. For the reaction, the concentrations of the Hairpin and Leg were diluted to 10 μM Hairpin, 0.5 μM Leg or 10 μM Leg for the multi-turnover or single-turnover reactions respectively. The reaction was stopped after 2, 4, 8, 16, 32, 60, and 1290 min (21.5 h) by being placed in 75 mM EDTA and 2 μM Probe.

The results were analyzed by loading the contents on a 20% nondenaturing PAGE gel at 4°C. The gel was scanned for fluorescence using a Typhoon fluorescence scanner with 532 nm excitation and Cy5 emission settings (GE Healthcare Life Sciences).

Sandwich Imaging Strand. Oligonucleotides were designed to increase binding specificity with the aid of VisualOMP (DNA Software). Ideal binding energies for each strand was taken from previously designed biosensors[6-8]. The strands required for the technique include the following: The modified substrate [5' - staple probe or biotin – GCT GGT CGA TTT TCA CTA TrAG GAA GAG], the Target [5' - GTT CCA GCC AGT GAA GCC TAT ATA GTG AAA ATC GAC CAG C], the Target Decoy [5' – CTC TTC CTA TAG TGA AAA TCG ACC AGC], and the Reporter [5' – AGG CTT CAC TGG CTG GAA CC - Cy5]. All strands were ordered from IDT.

Single-molecule Probe Detection Assay on origami. 1XHBS was flushed through the NeutrAvidin-coated channel. 100 pM of biotinylated rectangular origami[9] with 48 overhangs or 50pM of biotinylated rectangular origami with 187 overhangs was added to the channel and allowed to incubate for at least 10 min. Excess origami was flushed out with 1XHBS. 50 nM of

either product (5' – (GAT GAA TGG TGG GTG AGA GG)G CTG GTC GAT TTT CAC TAT rA) or substrate (5' – (GAT GAA TGG TGG GTG AGA GG)G CTG GTC GAT TTT CAC TAT rAGG AAG AG) strands with a section complementary to the overhangs on the origami were flushed through the channel and incubated for 30min. Excess strands were flushed out with 1XHBS. 50 nM of Target, Target Decoy, and Reporter were added to the channel and incubated for 30min. The imaging strands were flushed out with 1XHBS. An oxygen scavenging system[10] (OSS) was added to prolong fluorophore lifetime for imaging (5 nM protocatechuate (PCA), 50 nM protocatechuate dioxygenase (PCD), and 2 mM Trolox). The Prism-type TIRF microscope was used for imaging. 0.5-s exposure time was used to collect photons. The 638 nm laser used for imaging was set to an output power of 3 mW.

For detecting fluorescent origami islands in a vast sea of substrates, 1XHBS buffer was flushed through the channel. 133 pM 187 overhang origami was flushed through and allowed to incubate for 10min. Excess origami was flushed out with 1XHBS. 50nM of the product strand was flushed through and allowed to incubate for 10min. Another 50nM product strand was then flushed through for another 10min. 500nM biotinylated substrate was added for 30min. After the excess was flushed out using 1XHBS, 750nM Target was added for 45 min. Excess was flushed out and 750nM Target Decoy was added for 1 hour to allow ample time for the Target Decoy to displace any Target hybridized to the product. Then, 50nM Reporter was added for 15 min. 50nM Reporter was then again added for 15min. The channel was flushed out with 1XHBS followed by OSS for imaging.

B.3 Results

In order to obtain customized patterns, remote control ability over the direction of the spider movement must be made possible. Directional control could in principle be achieved using an electric field to bias the direction of motion (Fig B1.A). Applying a voltage across a nanofluidic channel is known to generate electroosmotic flow, which could be used to bias the movement of particles at the surface. To determine the effect that applying an electric field across our slides would have, food coloring was placed in each well. When attached to a low voltage power supply, the dye migrated from the well with the negative lead (yellow) to the well with the positive lead (blue; Figure B.1B). Dye colors were reversed to ensure the color change was not an artifact of pH change, and the blue dye migrated to the yellow (data not shown). The change in volumes between the two wells was negligible. This implies that the biased movement of the dye particles was driven by electrophoretic mobility as opposed to electroosmotic flow. Because the spider's legs are composed of DNA strands that are negatively charged, electrophoretic mobility is an adequate means of biasing the movement of the spiders.

Spiders were then placed on a slide surface that was coated with DNA-RNA cleavable substrates spaced ~5 nm apart[11] for the spider to walk on. Fluorescent beads were nonspecifically bound to the surface to be used as fiduciary markers. When a potential was applied across the channel, the spiders were unstably attached to the substrates and completely detached from the surface in the presence of the electric field. The beads, however, were more stably nonspecifically bound to the slide surface. Looking at the trajectories of the beads seemed to suggest that the direction of the bead movement might be influenced by the electric field (data not shown). Perhaps the beads could be used as the body of a nanowalker instead of the streptavidin.

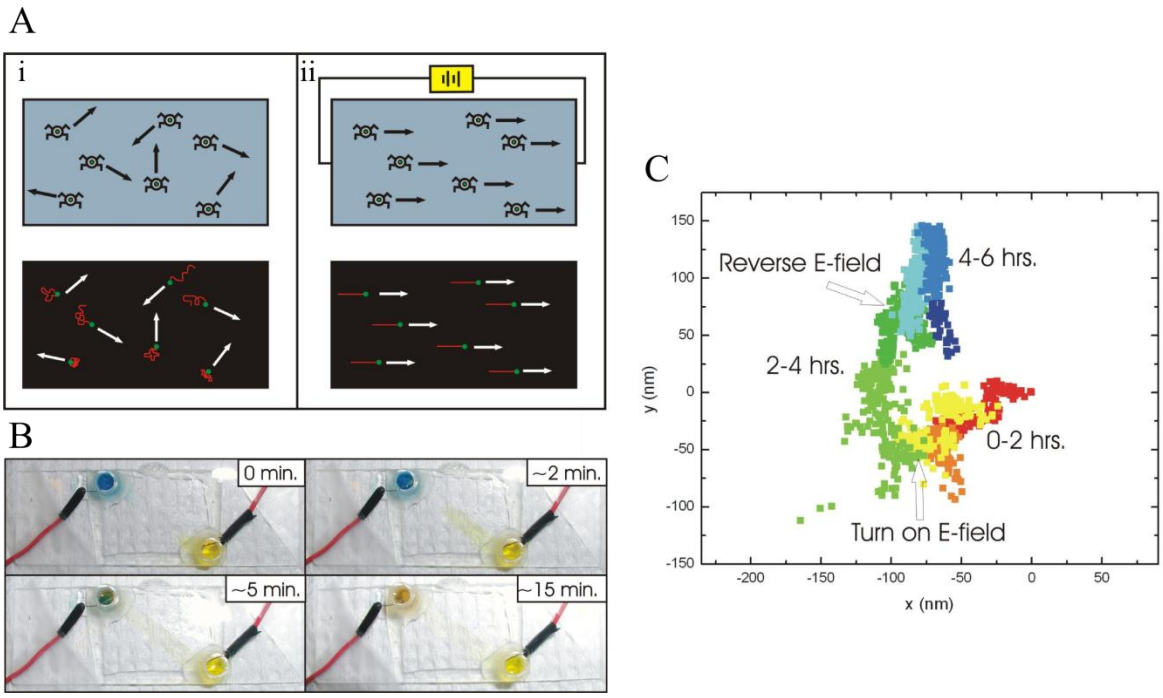


Figure B.1 Remote-controlled spider. (A) Schematic for controlling the direction of spider movement by applying an electric field. (i) When presented with a vast array of substrate, the spider will meander in an unbiased random walk fashion. Therefore, in detecting the pattern generated using single-molecule TIRF microscopy, the Cy5-labeled Probes (red) that blaze the spider's trail will be detected along with the Cy3-labeled spider (green). (ii) When a potential is applied across the fluidic channel, the spiders' movement should become a biased random walk, leading to the formation of a more uniform pattern across all spiders (red). (B) The electrophoretic mobility of food coloring with 20V applied across the channel for 15 min. (C) Fluorescent microsphere in the presence of an electric field with 15-s time resolution. Four beads were Gaussian fitted to form their average trajectory. Hours 0 to 2 (red to yellow) are in the absence of an electric field. Hours 2-4 (light green to dark green) are in the presence of an electric field. Hours 4-6 (light blue to dark blue) are in the presence of an electric field in the opposite direction.

In previous nanowalker applications, streptavidin-coated beads that could bind to the biotinylated legs were dismissed as a potential body. There were multiple reasons for this: Since there would be multiple streptavidins on each bead, each body would contain an indeterminate number of legs; the increase in streptavidins would also increase the number of legs, and it was previously reported that having too many legs per spider body would result in a decrease in nanowalker movement and a higher risk of getting stuck[12]; because the beads are able to nonspecifically bind to the slide surface, it was likely that nanowalkers with bead bodies would never be able to walk along the substrate, but would remain stuck on the surface. However, the presence of an electric field might be enough to aid a bead-based nanopatterner such that the previous concerns would no longer apply.

To help determine the effectiveness of using the fluorescent bead as a body, three different potential settings were applied over 6 h to the substrate-coated channel upon which the fluorescent beads sat: 0 V, 20 V, and -20 V (Figure B.1C). In the absence of an electric field and a fiduciary marker, the supposed trajectory imaged for the beads is only that of stage drift. Once a potential was applied, the beads immediately shifted in direction and slightly increased in speed (increasing from a mean velocity of ~45 nm/h to a mean velocity of ~65 nm/h). When the positive and negative leads were flipped, it took almost an hour for the direction to change. When the direction did change, it changed by 180°, suggesting that the change in movement was caused by the change in the direction of the electric field. No fiduciary marker was using in tracking the movement of the fluorescent beads, so the data is difficult to interpret. A fiduciary could be supplied by implementing the use of a biotinylated bead that would covalently bind to the NeutrAvidin-coated surface in future experiments. Regardless, the current data suggest that the bead could be a promising alternative to the current streptavidin spider-body.

The other vital component of the nanopatterner design are the trail blazes, the markers left by the spiders so that others may detect where they have tread. Here, two designs were proposed and tested to fluorescently detect product (i.e., fluorescently detect where the spider had previously been). The first was based on hairpin biosensors[13-15] (Figure B.2A). In a hairpin biosensor, a strand hybridizes to the loop region of the hairpin, opening the hairpin to make the aptamer-binding region accessible for a probe to hybridize. In the case of the spider, the substrates for the spider are redesigned to be hairpins. The spider leg binds so as to open the loop region and cleave the Hairpin at the cleavage site within the loop region. This permanently disrupts the Hairpin, causing a portion of it to diffuse into solution while the other portion remains tethered to the surface. The product still tethered to the surface is now exposed so that a fluorescent Probe strand may come and hybridize to the remaining strand, which will allow detection of the path traversed by the spider.

Ensemble cleavage assays of the Hairpin design were used to test the specificity of the binding of the Probe to its target sequence, meaning that the Probe should only bind to the product and not the full Hairpin. Under single-turnover conditions in which the leg and Hairpin concentrations are equal, the Hairpin is able to be cleaved and it is found favorable for the Probe to bind to this cleaved strand (Figure B2.B). Under multiple-turnover conditions, however, very few Hairpin strands are cleaved even after 22 h (Figure B2.C). This suggests that the rate at which the leg is able to dissociate from a previously cleaved Hairpin and find a new Hairpin to cleave is extremely slow, even though there is a great excess of Hairpin strands. It is important for this rate to be optimized in order to have spiders walk along a surface in a reasonable amount of time that competes with other nanopatterning techniques. Also, both the single-turnover and multiple-turnover conditions contain a band that suggests excessive hybridization of Probe to

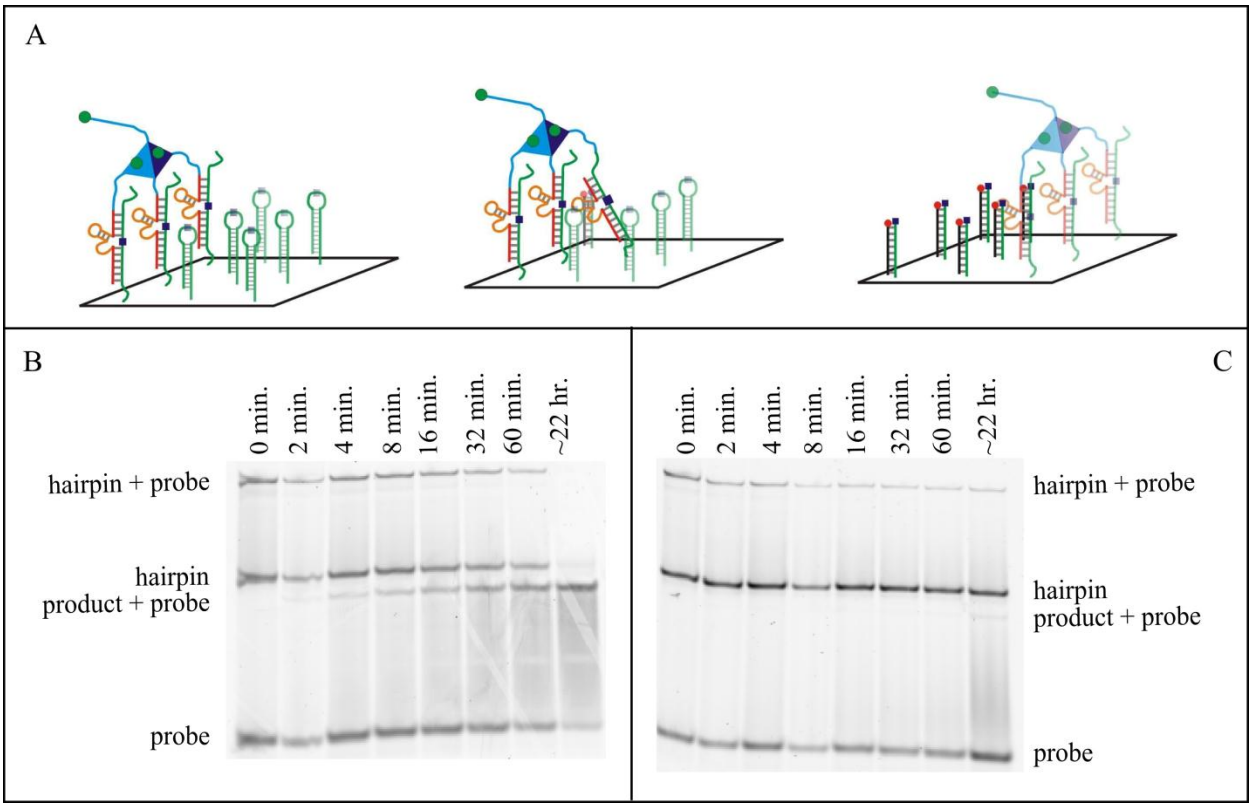


Figure B.2 Hairpin design for nanopatterning. (A) The spider (blue) opens each Hairpin substrate (green) as its legs (red/orange) bind to the loop region. The legs cleave the Hairpin at the cleavage site (purple) before proceeding to the next strand. Once the spider has cleaved the Hairpins, a Cy5-labeled Probe strand (black/red) is added to the slide surface to hybridize to the exposed binding region in the product. (B) Single-turnover and (C) multiple-turnover cleavage and binding specificity assays. As the Hairpin is cleaved, the Probe is able to bind to its product as demonstrated by the decrease in Hairpin and increase in product + Probe over time. While there seems to be significant cleavage activity under single-turnover conditions, a faint product + Probe only appears after 22 hours under multiple-turnover conditions. The significant volume of Hairpin + Probe suggests a high level of premature binding.

uncleaved Hairpin strands, which would lead to far too many false positive fluorescent signals in a single-molecule experiment.

The second design was inspired by sandwich biosensors[6-8] (Figure B.3A). In these biosensors, the Probe binds to a strand, here called the Target, that is complementary to the sequence of interest and contains an overhang to which the Probe strand can bind. In this application, the 3' region of the Target contains a sequence that is fully complementary to the product while the 5' region is fully complementary to the Probe. Because the substrate and product only differ in sequence by 7 basepairs, there is a high probability of false positive signals. To overcome this difficulty, another strand was implemented, dubbed the "Target Decoy." This strand is fully complementary to the substrate. The binding affinity for the Target Decoy to substrate will thus be higher than for the Target to the substrate, resulting in fewer false positive signals. The Target Decoy can also be effectively used in strand displacement to remove the Target from substrate to minimize false signals.

First, we needed to determine whether there was a difference between the fluorescent signals obtained from substrate strands or product strands. Instead of covering the entire slide surface with either product or substrate strands, little islands of strands were made by using origami tiles with 48 overhangs to which either product or substrate could bind. The slide containing only substrate-decorated origami had a significantly lower fluorescent signal than the product-decorated origami, suggesting that the imaging strands are working effectively (Figure B.3B).

Second, we wanted to measure the sensitivity of this technique by determining whether the fluorescent signal would increase given a higher density of product strands. To this end, we used origami containing 187 overhangs to which substrate or product could bind. There was a

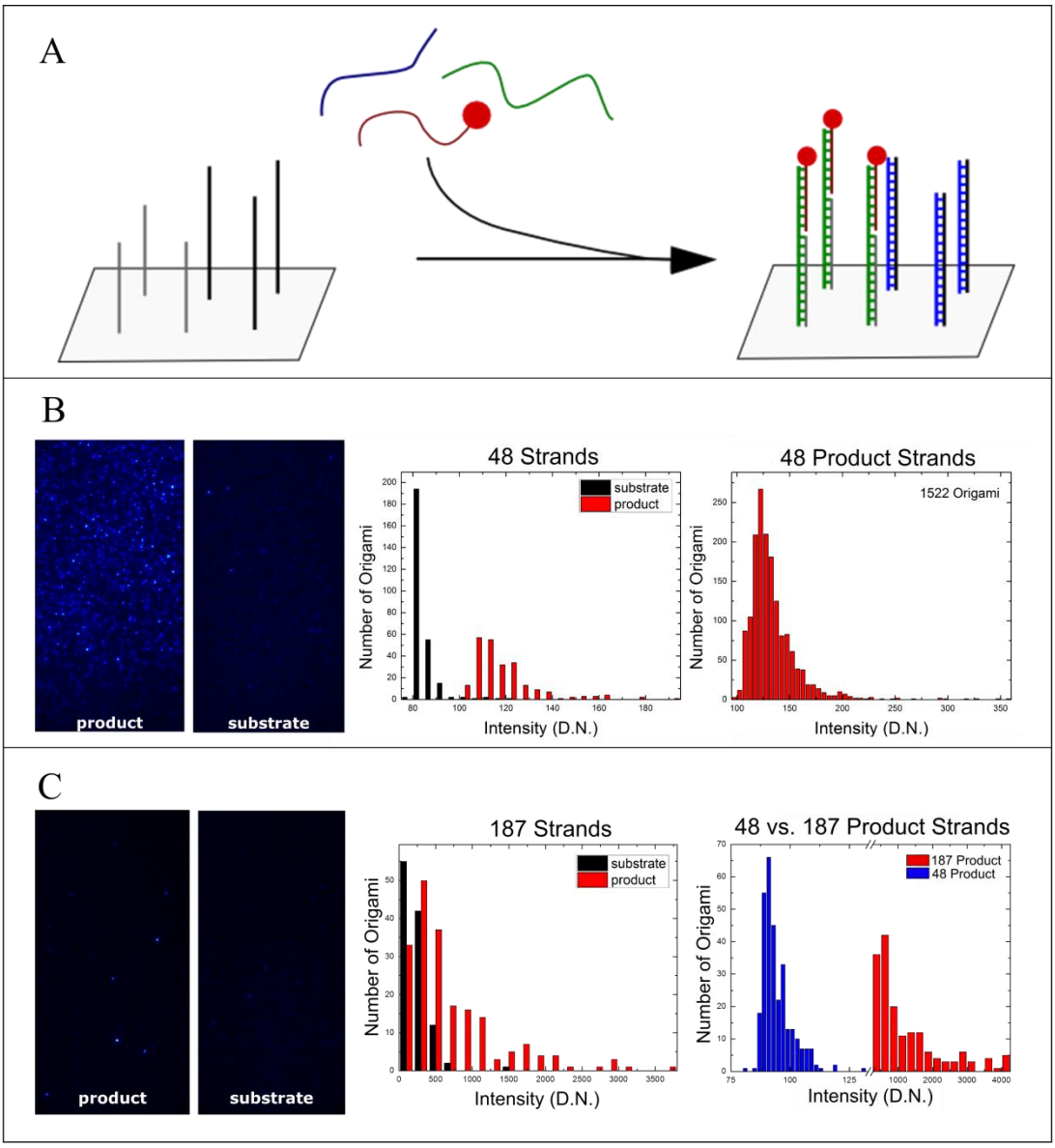


Figure B.3 Sandwich imaging strands. (A) The spiders walk along their substrate (black) that they cleave to form product (gray). After the spider has patterned its surface by cleaving its substrates, three imaging strands are added: the Target (green), the Target Decoy (blue), and the Probe (red). The Target and the Target Decoy can bind to either the substrate or the product, but the Target Decoy has a higher binding affinity for the substrate. The 5' end of the Target is complementary to the Probe, which is fluorescently-labeled, thus signifying the presence of the product in a detectable manner. This technique was tested using origami with (B) 48 overhangs and origami with (C) 187 overhangs. Histograms were made of the fluorescence intensity of the probes on each origami when they were decorated with either product or substrate strands.

significant 10-fold shift in intensity given a 4-fold increase in strands, suggesting that the fluorescent signal is sensitive to changes in strand density (Figure B.3C).

This technique needs to also work when the islands of product are surrounded by a sea of substrate. To test the technique's ability to cope with this large excess of substrates, origami with 187 overhangs decorated with product were placed on the NeutrAvidin-coated slide surface, followed by biotinylated substrate that saturated the remainder of the surface. Initial trials proved promising, but the signal-to-noise is still too low for these strands, let alone for the far smaller number of products in the midst of substrate that would be generated by our remote-controlled nanopatterners.

B.4 Discussion

While some initial progress has been made toward spider nanopatterners, there is still a long way to go. Beads show some promise as a potential material to for a body, but the speeds at which they traverse the surface are not at a point where they can compete with current nanopatterning methods. Also, there are challenges when implementing an electric field through a fluidic channel, such as how to provide a constant supply of metal ions necessary for cleavage on the slide surface when the metal ions will be affected by the electric field. Buffer exchange could be implemented regularly, but this will only slow down the process further.

The detection technique will also have its own challenges. Even if islands of 187 strands may be seen with improved signal-to-noise, the spiders will most likely cleave far fewer strands given their low speeds. The false positive signal would have to be vastly reduced. However, these are relatively manageable challenges. One method of doing this may be to use PEGylated slides which have been found to reduce nonspecific binding of strands. Also, locked nucleic acids (LNA) could be used to improve specificity. If spiders have an applicable future ahead of

them, it is my opinion that MacGyver nanopatterners may be one “biased direction” the field of DNAzyme nanowalkers will take.

B.5 References

1 Michelotti, N., Johnson-Buck, A., Manzo, A. J. & Walter, N. G. Beyond DNA origami: the unfolding prospects of nucleic acid nanotechnology. *Wiley interdisciplinary reviews. Nanomedicine and nanobiotechnology* **4**, 139-152, doi:10.1002/wnan.170 (2012).

2 Kufer, S. K., Puchner, E. M., Gumpp, H., Liedl, T. & Gaub, H. E. Single-molecule cut-and-paste surface assembly. *Science* **319**, 594-596, doi:10.1126/science.1151424 (2008).

3 Lund, K. *et al.* Molecular robots guided by prescriptive landscapes. *Nature* **465**, 206-210, doi:Doi 10.1038/Nature09012 (2010).

4 Michelotti, N., de Silva, C., Johnson-Buck, A. E., Manzo, A. J. & Walter, N. G. A Bird's Eye View: Tracking Slow Nanometer-Scale Movements of Single Molecular Nano-Assemblies. *Method Enzymol* **475**, 121-148, doi:Doi 10.1016/S0076-6879(10)75006-0 (2010).

5 Guo, Z., Guilfoyle, R. A., Thiel, A. J., Wang, R. F. & Smith, L. M. Direct Fluorescence Analysis of Genetic Polymorphisms by Hybridization with Oligonucleotide Arrays on Glass Supports. *Nucleic Acids Research* **22**, 5456-5465 (1994).

6 Pavlov, V. *et al.* Amplified chemiluminescence surface detection of DNA and telomerase activity using catalytic nucleic acid labels. *Analytical chemistry* **76**, 2152-2156, doi:10.1021/ac035219l (2004).

7 Niazov, T., Pavlov, V., Xiao, Y., Gill, R. & Willner, I. DNAzyme-Functionalized Au Nanoparticles for the Amplified Detection of DNA or Telomerase Activity. *Nano Letters* **4**, 1683-1687, doi:10.1021/nl0491428 (2004).

8 Yao, X. *et al.* Sub-attomole oligonucleotide and p53 cDNA determinations via a high-resolution surface plasmon resonance combined with oligonucleotide-capped gold nanoparticle signal amplification. *Analytical biochemistry* **354**, 220-228, doi:10.1016/j.ab.2006.04.011 (2006).

9 Rothemund, P. W. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302, doi:10.1038/nature04586 (2006).

10 Aitken, C. E., Marshall, R. A. & Puglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophysical journal* **94**, 1826-1835, doi:10.1529/biophysj.107.117689 (2008).

11 Manzo, A. J., Taylor, S.K., Pei, R., Stojanovic, M.N. & Walter, N.G. Direct super-resolution video imaging reveals heterogeneous diffusive walks among swarms of single nanorobots. *Unpublished* (2012).

12 Pei, R. *et al.* Behavior of polycatalytic assemblies in a substrate-displaying matrix. *Journal of the American Chemical Society* **128**, 12693-12699, doi:10.1021/ja058394n (2006).

13 Zhang, J., Chen, J., Zhang, X., Zeng, Z., Chen, M. & Wang, S. An electrochemical biosensor based on hairpin-DNA aptamer probe and restriction endonuclease for ochratoxin A detection. *Electrochemistry Communications* **25**, 5-7 (2012).

14 Du, H., Disney, M. D., Miller, B. L. & Krauss, T. D. Hybridization-based unquenching of DNA hairpins on au surfaces: prototypical "molecular beacon" biosensors. *Journal of the American Chemical Society* **125**, 4012-4013, doi:10.1021/ja0290781 (2003).

15 Du, H., Strohsahl, C. M., Camera, J., Miller, B. L. & Krauss, T. D. Sensitivity and specificity of metal surface-immobilized "molecular beacon" biosensors. *Journal of the American Chemical Society* **127**, 7932-7940, doi:10.1021/ja042482a (2005).

APPENDIX C:
SUPPLEMENTARY INFORMATION FOR CHAPTER 2

Monte Carlo Simulation Results.

To approximate the results from experiment, the width and height of the cytoplasmic region were set to 100 nm. The length was set to 160 nm. The nucleoplasmic region had the same dimensions. All dimensions of the central region were set to 40 nm. 1,750 particles were simulated for the mRNA complex results and 675 particles were simulated for the Tap-p15 and GFP results. The first conclusion provided by numerical simulations was that the number of particles tracked during the experiment is sufficient for convergence of the final distribution of particles, as was determined by running the simulation ten times and determining the standard deviation for each region. For the distribution of the mRNA-complexes, the outer sub-regions contained 787 ± 35 particles (45 ± 2% of the total number of particles) and the central region contained 54 ± 8 particles (8.8 ± 0.4%). For the distribution of Tap-p15, the outer sub-regions contained 304 ± 14 particles (45 ± 2%) and the central region contained 54 ± 7 particles (8.3 ± 1.0%). For the distribution of GFP, the outer sub-regions contained 293 ± 14 particles (43 ± 2%) and central region contained 93 ± 7 particles (13.8 ± 1.0%). Incorporating the experimental uncertainty in position determination of 8 nm into the simulation did not influence the numerical results, but did broaden the Gaussians, consistent with the experimental results.

To obtain a distribution of particles resembling the distribution observed experimentally across each sub-region, the maximum step-size through the central region was decreased 2.8-fold for the mRNA-complex and GFP, and 3.6-fold for Tap-p15, implying that Tap-p15 has greater

interaction with the FG Nups, an interaction that has previously been reported[1]. This result was independent of step-size and the number of steps, under the converging conditions tested numerically. This fractional decrease in maximum step-size of 2.8 (3.6) is equivalent to a decrease in the diffusion coefficient by a factor of 7.8 ± 0.4 (12.9 ± 0.4), as was determined by fitting ensemble mean-square-displacement plots for the particles with the same maximum step-size conditions without boundaries to the equation $\langle x^2 \rangle = 4Dt$. This result strongly argues against the proposal that an increase in the diffusion coefficient is necessary to obtain the binomial distribution found experimentally[2]. On the contrary, because of the decrease in volume through the central region, a *decrease* in the diffusion coefficient is necessary to obtain the observed binomial distribution.

Cell type	RNA Type (3.3 kb)	Label	Detecting Single mRNAs on or in	Detection Time per Frame (ms)	Dwell Time (ms)	Efficiency (%)
MEF	β-actin	YFP-MCP	NE	20	21±1	32±5
HeLa	Firefly luciferase	GFP-MCP	NE	20	16±2	31±5
MEF	β-actin	YFP-MCP	NE	2	12±1	36±5
HeLa	Firefly luciferase	mCherry-MCP	Single NPC	2	12±2	34±5
HeLa	Firefly luciferase	mCherry-MCP	Nucleus, cytoplasm and NPC-deficit NE	200	t_1=620±130 t_2=65,000±21,000	n/a

Table C.1 Nucleocytoplasmic transport efficiencies for transport receptors, facilitated translocation, and passively diffusing molecules in permeabilized and living cells.

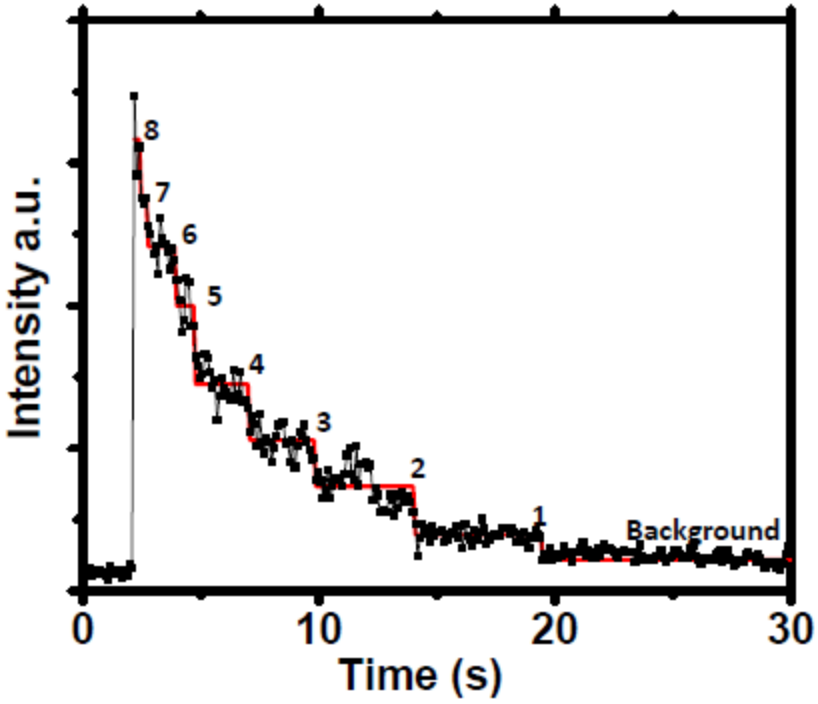


Figure C.1 Selected NPCs contained ~8 copies of GFP-POM121. A typical photobleaching curve of eight copies of GFP-POM121 in a single NPC of a living HeLa cell. The change points between steps were resolved by the published maximum likelihood ratio method[3,4].

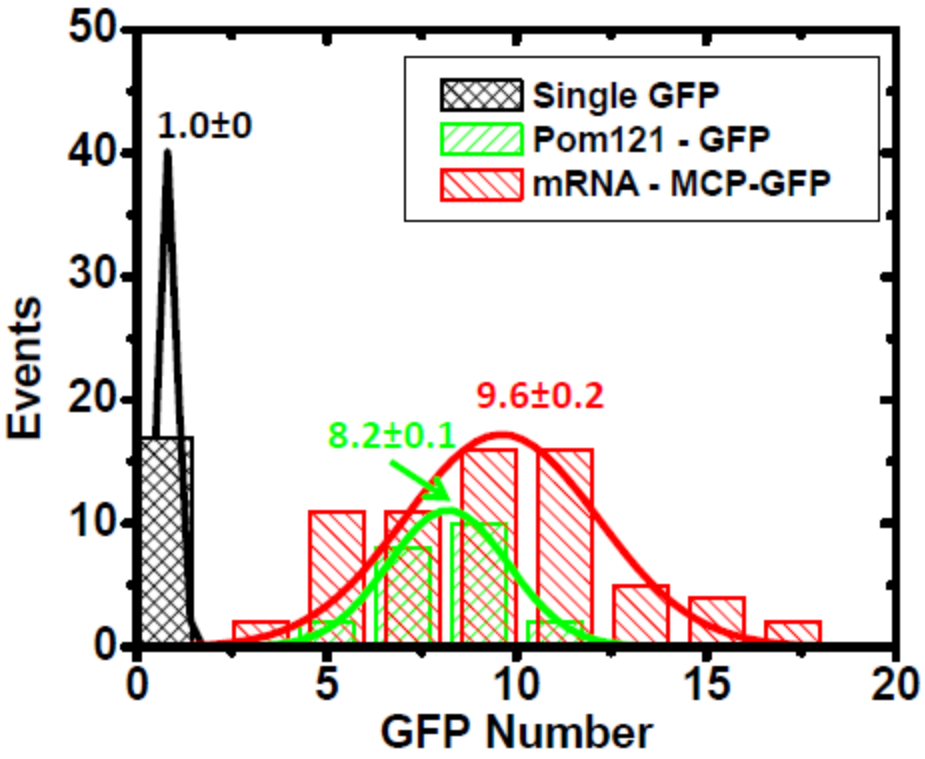


Figure C.2 Each mRNP contained ~10 copies of MCP-FP. Comparisons between the intensities of single GFP, GFP-NPC and GFP-tagged mRNP fitted with Gaussian functions revealed that there were ~8 copies of GFP in each NPC and ~10 copies of MCP-GFP per mRNP. MCP-mCherry-mRNA was constructed with the exact same procedure as MCP-GFP-mRNA, and thus ~10 copies of MCP-mCherry per mRNA were concluded.

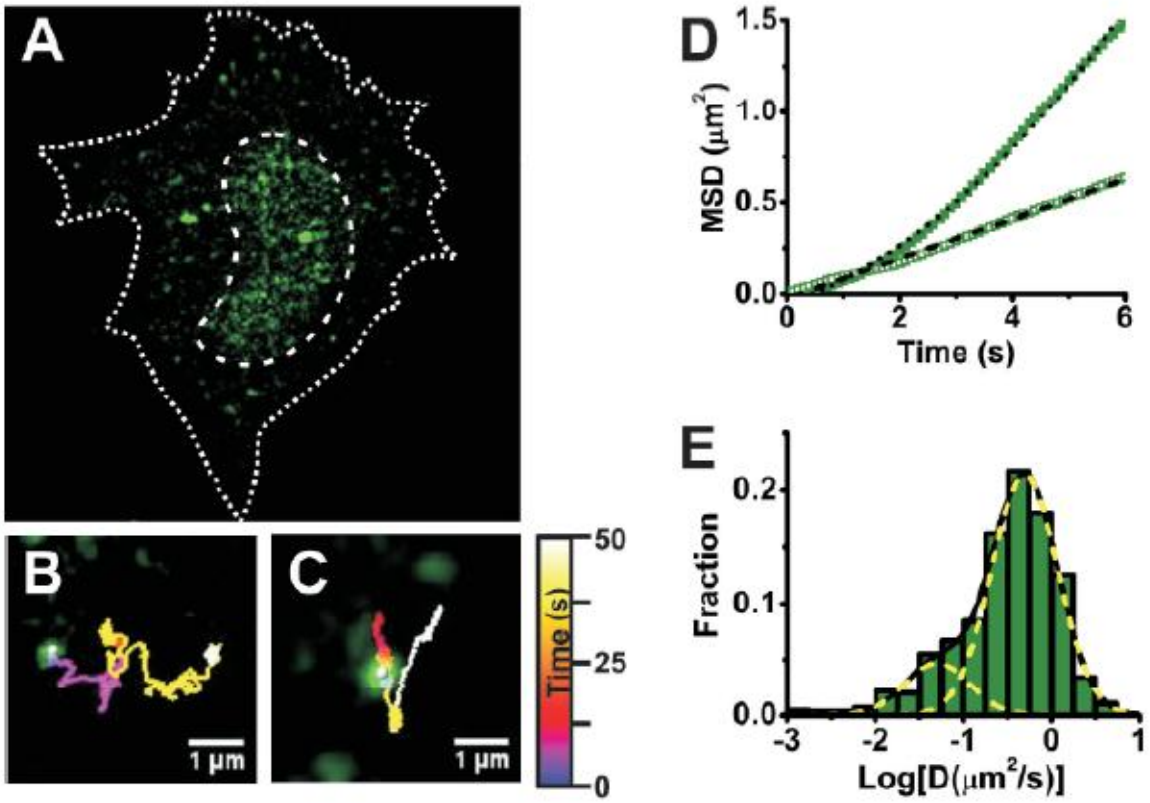


Figure C.3 Cytoplasmic mRNP diffusion. A) A deconvolved (AutoQuant X3) image of a cell transiently expressing the mRNA labeling system. Individual green spots in the cytoplasm represents GFP labeled mRNPs. Low expression levels of both these constructs ensured visualization of spatially distinct single mRNPs. **B** and **C**) Individual mRNPs undergoing brownian and biased diffusion respectively. **D**) Mean-squared displacement as a function of time for mRNPs in B and C. **E**) Distribution of cytoplasmic mRNP diffusion coefficients in HeLa cells. Diffusion coefficients were calculated from MSD vs time plot of individual particles (962 particles, 4 cells) that were visible for at least 9 consecutive frames. mRNPs distributed predominantly into two Gaussian distributions of diffusion coefficients, with average diffusion coefficients of ~0.49 um^2/s and ~0.05 um^2/s for the major (81%) and minor (19%) populations respectively.

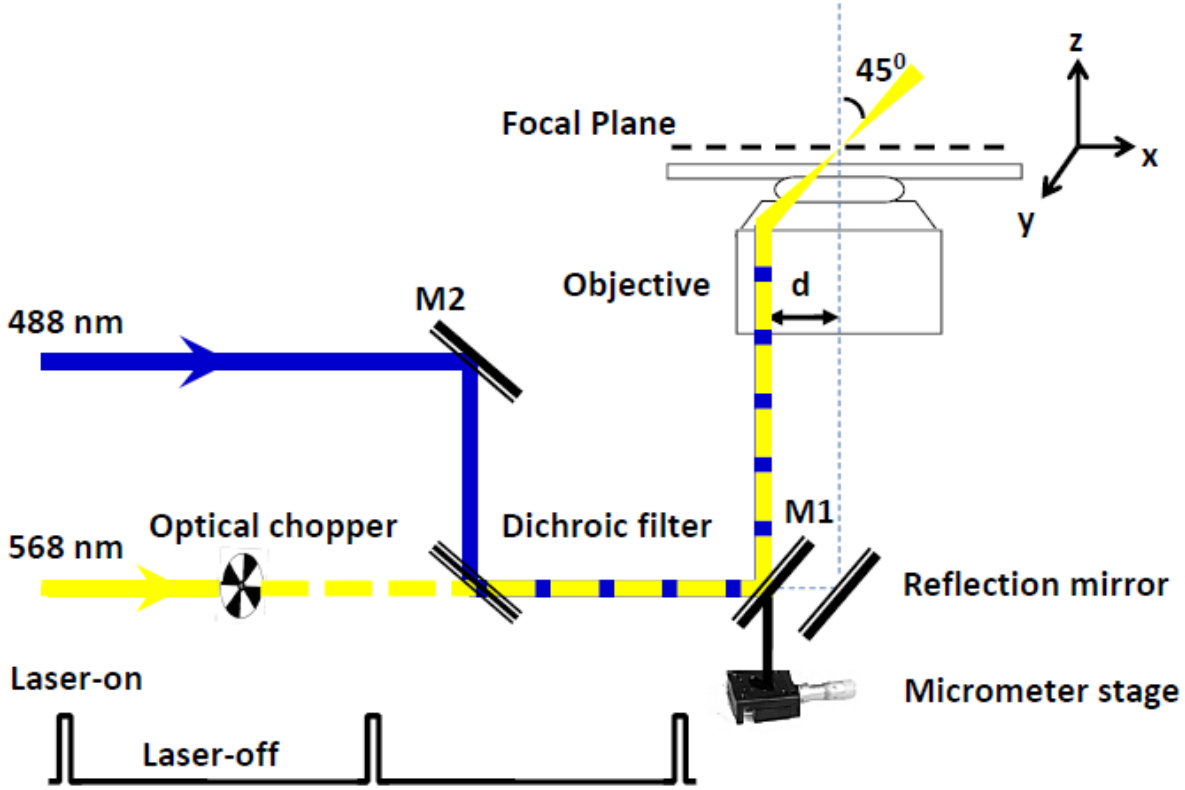


Figure C.4 Optical schematic of the SPEED microscope. After combining the 488 and 568 nm lasers, their beams were shifted using a micrometer stage by ~237 μm (d) from the optical axis of the objective to generate an inclined illumination volume at an angle of 45° to the perpendicular direction. The 568 nm laser was chopped by an optical chopper to achieve an on-off laser mode with a laser-on time of 60 ms and a laser-off time of 140 ms. The longer laser-off time gives particles transiting the NPC sufficient time to escape from the illumination volume and for fresh fluorescent cargo to diffuse from the cytoplasm or the nucleus into the NPC. To ensure that complete transport events through the NPC were captured, we chose the illumination (and photobleaching) time of single fluorescent particles (~60 ms) to be considerably longer than their nuclear transport time (~12 ms).

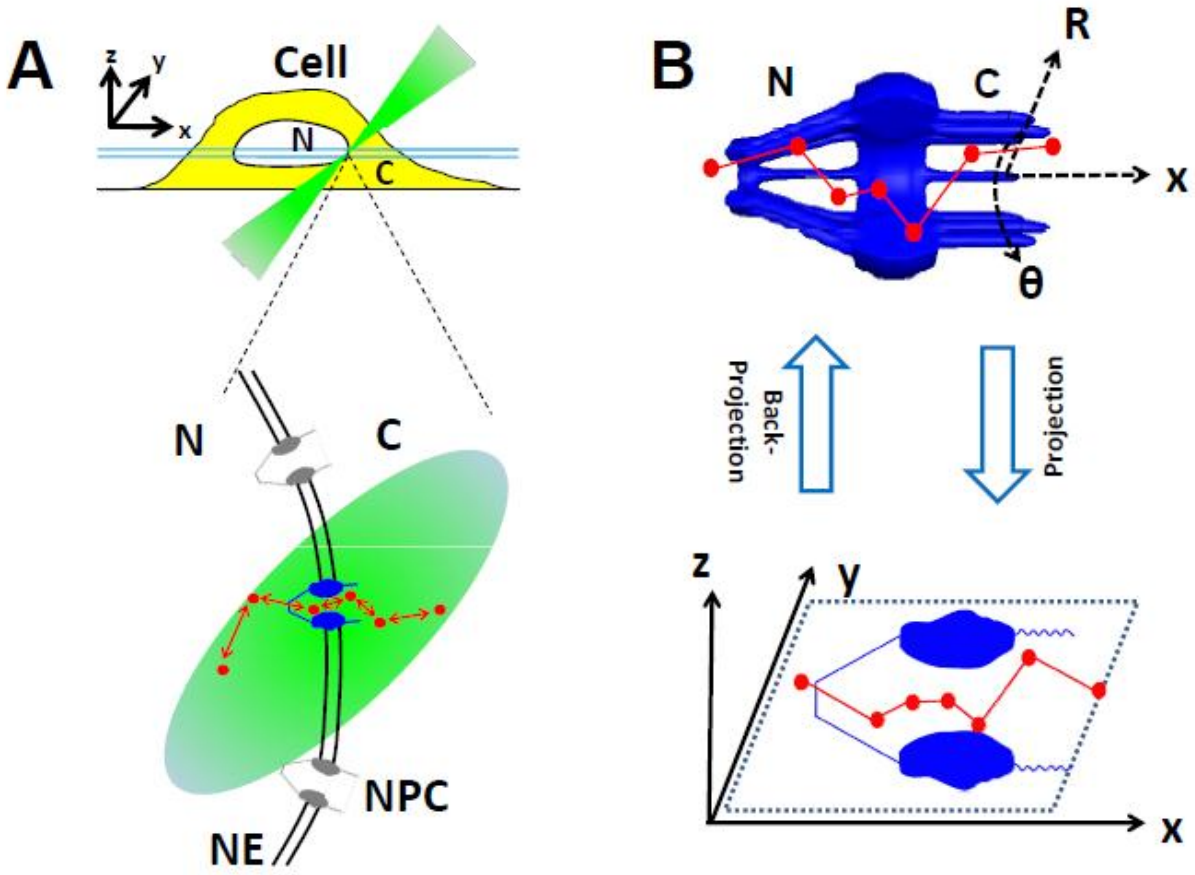


Figure C.5 Schematic of the analysis of a single NPC in a living HeLa cell by SPEED microscopy and 2D-to-3D deconvolution. (**A**) Only a single labeled NPC (green) on the NE was illuminated by an inclined diffraction-limit illumination point spread function (which forms an angle of 45^{o} to the z direction) of SPEED microscopy at the equatorial plane of a HeLa cell nucleus in the focal plane (between the double light blue lines). Single transiting particles (red dots) were tracked as they diffused through the illuminated NPC. N, nucleoplasm; C, cytoplasm. (**B**) The 2D-to-3D deconvolution process. Using established deconvolution algorithms[5,6], the 3D spatial locations of transiting molecules in the NPC in a cylindrical coordinate system (R, θ, x) were recovered from the obtained 2D spatial locations in the x,y plane.

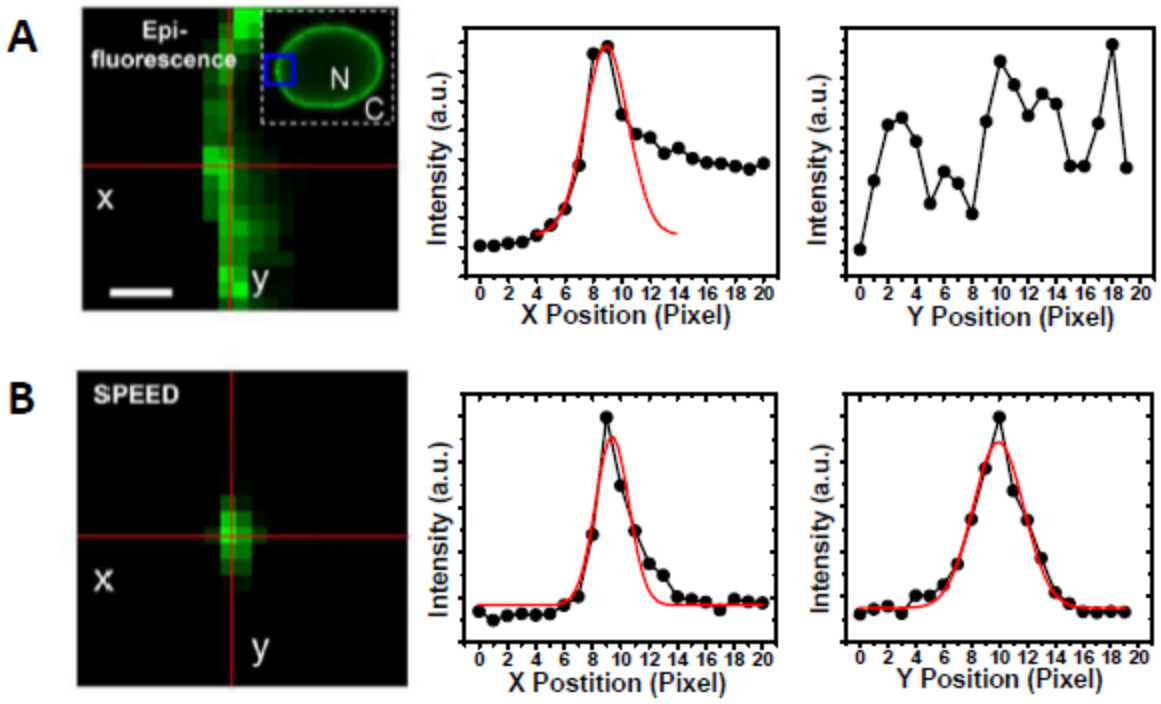


Figure C.6 Localization of a single NPC on the NE. (**A**) Multiple GFP-NPCs were excited using wide-field epi-fluorescence microscopy. The selected area is enclosed by a blue box in the image of the entire fluorescent NE (inset). The heavily overlapped fluorescence of multiple GFP-NPCs prevented Gaussian fitting of individual NPCs on the NE using epi-illumination. N, nucleoplasm; C, cytoplasm; scale bar, 1 μm. (**B**) In contrast, only a single GFP-NPC was excited in the illumination volume of the SPEED microscope. The fluorescent spot was fitted well by Gaussian function in both x and y directions.

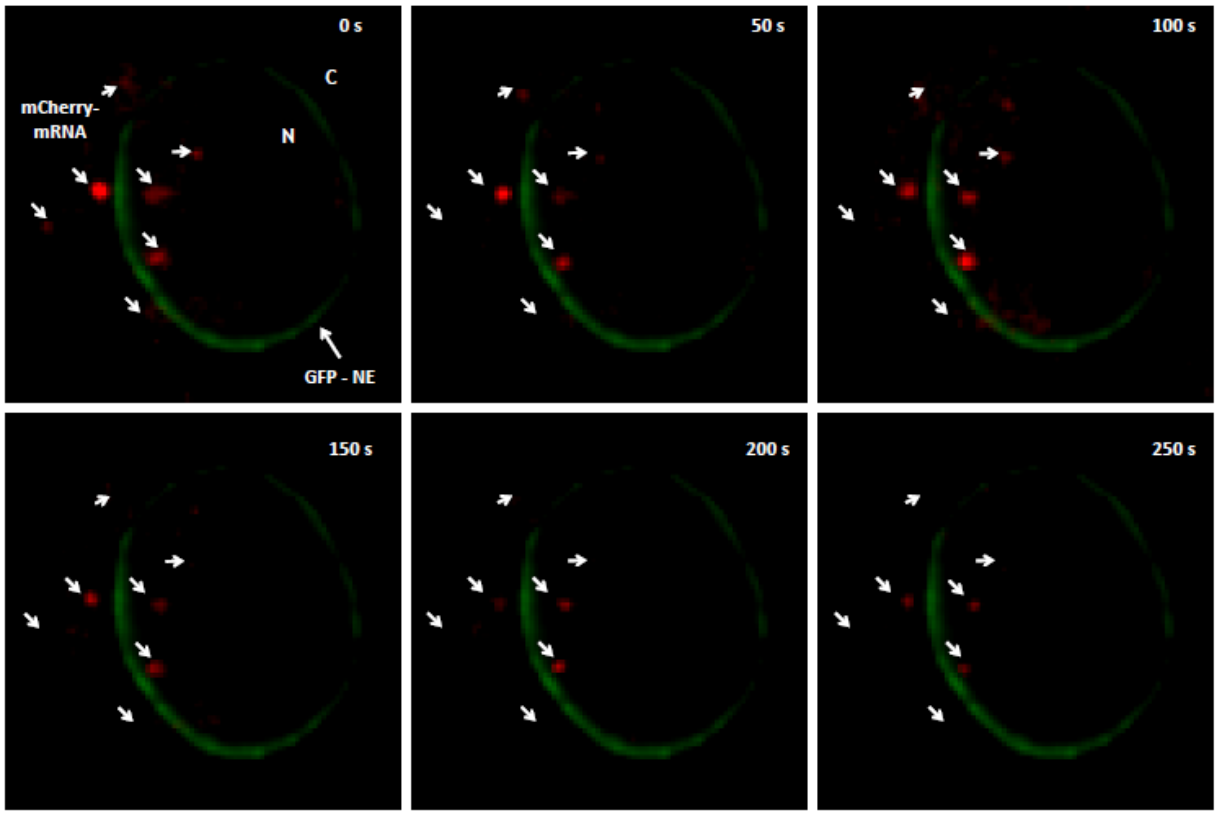


Figure C.7 Almost immobile or slowly moving mRNPs in nucleus and cytoplasm. Multiple mCherry-tagged mRNPs (red spots marked by arrows) stayed in the periphery of the NE (green ring) without obvious movements over 250 s until they photobleached.

REFERENCES

1 Carmody, S. R. & Wente, S. R. mRNA nuclear export at a glance. *J Cell Sci* **122**, 1933-1937 (2009).

2 Grunwald, D. & Singer, R. H. In vivo imaging of labelled endogenous beta-actin mRNA during nucleocytoplasmic transport. *Nature* **467**, 604-607, doi:10.1038/nature09438 (2010).

3 Watkins, L. P. & Yang, H. Detection of intensity change points in time-resolved single-molecule measurements. *The journal of physical chemistry. B* **109**, 617-628, doi:10.1021/jp0467548 (2005).

4 Boudjellaba, H., MacGibbon, B., Sawyer, P. On exact inference for change in a Poisson sequence. *Commun. Statist.-Theory Meth.* **30**, 407-434 (2001).

5 Ma, J., Goryaynov, A., Sarma, A. & Yang, W. Self-regulated viscous channel in the nuclear pore complex. *Proc Natl Acad Sci U S A* **109**, 7326-7331, doi:10.1073/pnas.1201724109 (2012).

6 Ma, J. & Yang, W. Three-dimensional distribution of transient interactions in the nuclear pore complex obtained from single-molecule snapshots. *Proc Natl Acad Sci U S A* **107**, 7305-7310, doi:10.1073/pnas.0908269107 (2010).

APPENDIX D:
SUPPLEMENTARY INFORMATION FOR CHAPTER 3

SUPPORTING DISCUSSION: MOLECULAR ROBOTICS

Robots are often defined by their ability to sense their environment, perform computations, and take actions; as such, they have revolutionized our ability to automate factories, send autonomous vehicles to remote or dangerous locations, and improve our daily lives. The potential for autonomous sensing and acting at the molecular scale is illustrated by the sophisticated machinery within biological cells, where molecular motors and biochemical circuitry coordinate the cell's active responses to its environment. From a chemist's perspective, the potential for molecular robotics goes far beyond what is observed in biology, but the challenges of realizing that potential are daunting due to the need to synthesize behavior.

As with protein motors, an isolated molecular robot by itself serves no purpose; to be useful, it must interact with its environment of other molecules and molecular machines; it must behave. Despite vast differences in size, classical robotics[1-4] can provide a framework for designing interacting molecular machines with complex behaviors within their environments.

A simple example of a molecular robot would be a "walking" DNA molecule that can recognize and follow an arbitrary trail ("bread crumbs"). If such a simple molecular robot could be demonstrated, its capabilities then could be expanded by incorporating additional layers of control mechanisms from DNA nanotechnology and concepts from computer science. For example, integration of logic and memory into the robot's body would enhance the robot's ability to respond to its environment intelligently[5]; interactions between multiple molecular robots could lead to collective behavior[6-8], and the ability to read and transform the landscape (e.g., pick up and deposit loads) would in theory provide the essential mechanism for Turing-universal algorithmic behavior[9-12].

Research in programmable DNA walkers[13] started with non-autonomous remote-controlled systems[14,15], progressed to autonomous walkers that modify visited sites to achieve directed (but brief) motion on linear tracks[16-19], or to achieve continuous processive (but undirected) motion in two or three dimensions[20], and shows promise for processive and directed walking on undisturbed tracks[21]. Theoretical work has envisioned how certain types of DNA walkers could be augmented with additional control mechanisms to act as finite state machines, universal Turing machines, and programmable molecular robots [22,23], but these schemes have not been experimentally demonstrated. Other work related to molecular robotics has demonstrated the controlled movement of dendrimers[24], nanowire motors[25], and nanocars[26]. While synthesizing suitably well-defined tracks has been an important technical challenge for DNA-based walkers (no previous walker has been demonstrated to take more than three steps on a linear track), our

interest herein is in how robotic behavior can be obtained from the interaction between a simple random walker and its environment.

In this work, we present an implementation of molecular robots that integrates aspects of DNA-based computing devices[16,27-29],complex structures[30-36] and actuators[13,20,37]. The DNA walkers chosen for this work, called "molecular spiders", comprise an inert body and multiple catalytic "legs". Specifically, here we use three-legged spiders with a streptavidin body. Spider legs are adapted from DNA enzyme 8-17 that binds and cleaves single-stranded oligodeoxynucleotide substrates with a single ribose moiety into two shorter products that have a lower affinity for the enzyme[38]. In the context of substrates that are immobilized at sites on a surface, spider behavior can be modeled using local rules[39]: a leg bound to substrate will cleave it at a low rate; a leg bound to product will detach at an intermediate rate; and a free leg will quickly bind (with little or no bias) a nearby substrate or product. For a multipedal spider positioned at the interface between regions of product and substrate, these rules predict that after a given leg cleaves and then lifts, it will by trial-and-error search out a nearby substrate to bind, thus moving the spider's body toward the substrate region while enlarging the product region behind it. A Monte Carlo simulation using these rules is presented further below. On 2D surfaces or in a 3D matrix, such spider movement results in a random walk with memory of visited sites, while on a 1D linear track it results in directed motion as the substrate is consumed. Crucially, unlike related "burnt bridge" Brownian ratchet mechanisms used in DNA walkers[14,15,17-19] and observed in nature[40], these local rules predict that multipedal spiders will not readily dissociate even from tracks consisting exclusively of product strands, and indeed will perform a rapid unbiased random walk there until they again encounter substrate.

Considering spider legs to be simultaneously sensors that detect nearby oligonucleotides and actuators that modify their environment to inhibit reverse motion, we exploit this sensor-actuator feedback to design prescriptive landscapes that direct the spiders' motion along a predefined path (Figs 1c and d). A spider traversing this landscape of oligonucleotide substrates can sense the set of available cues within its reach and take action accordingly. Here, we show that in the context of a precisely-defined track laid out on two-dimensional (2D) DNA origami[33], the previously introduced processive but random walker[20] becomes a processive and directed walker capable of path-following behavior. The importance of these results lies not in the walkers reaching stable thermodynamic endpoints, but in reaching those points through autonomously guided dissipative processes that can be programmed. Such processes could, in the future, be used to couple the behavior of multiple walkers through their interactions with a common landscape.

Figure D.1 IE HPLC trace showing: 'Part A' mixture for **NICK3.4A+1** from which "STV-(**C**)$_1$" was isolated (lowest trace); Other traces show the 'Part A' mixture with increasing amounts of **C** added (Note: "equivalent" amounts were based on the reported lyophilized amount of product supplied and not determined by absorption at 260 nm, which likely accounts for the introduction of a systematic error in the actual number of equivalents as observed by excess oligonucleotide present in the top trace). The 260nm/280nm ratios for the peaks of the middle trace are (left-to-right) 1.06, 1.28, 1.39, and 1.44, consistent with each consecutive peak containing a higher ratio of DNA-to-streptavidin than the peak preceding it. The 260nm/280nm ratio for peak "(**C**)" is 1.94, consistent with the absorption characteristics of pure DNA. See right y-axis for buffer B gradient (dotted line) as a percentage of buffers A plus B. Buffer A was composed of 20 mM TRIS, and buffer B, 20 mM TRIS/1 M NaCl, both adjusted to pH 7.4. The total flow rate of buffer A and B was 1 min^{-1}.

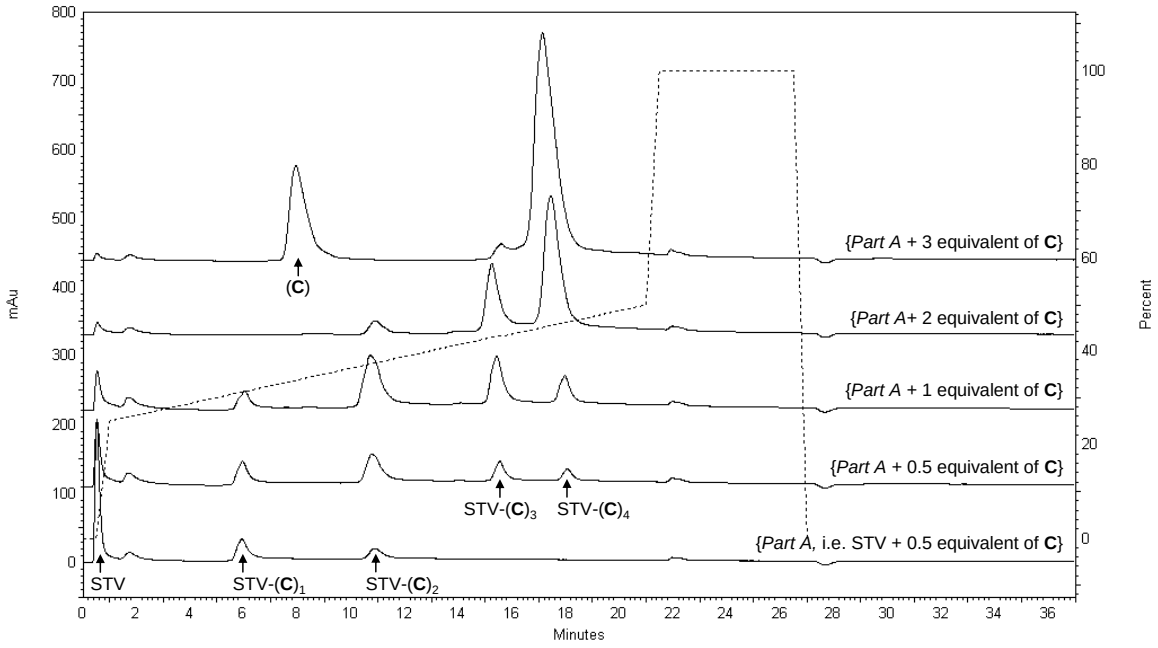


Figure D.2 IE HPLC trace showing titration of STV-(C)1 with increasing equivalents of L. 260nm/280nm ratios are STV-(**C**)$_1$ 1.09; STV-(**C**)$_1$(**L**)$_1$ 1.38; STV-(**C**)$_1$(**L**)$_2$ 1.53 and 1.49 (taken at the two apparent maxima respectively for STV-(**C**)$_1$(**L**)$_2$); **NICK$_{3.4A+1}$** 1.59; and **L** 2.05 (see caption for Supplementary Fig. 1 for explanation of absorption wavelength ratio 260/280). See right y-axis for buffer B gradient (dotted line) as a percentage of buffers A plus B. Buffer A was composed of 20 mM TRIS, and buffer B, 20 mM TRIS/1 M NaCl, both adjusted to pH 7.4. The total flow rate of buffer A and B was 1 min^{-1}.

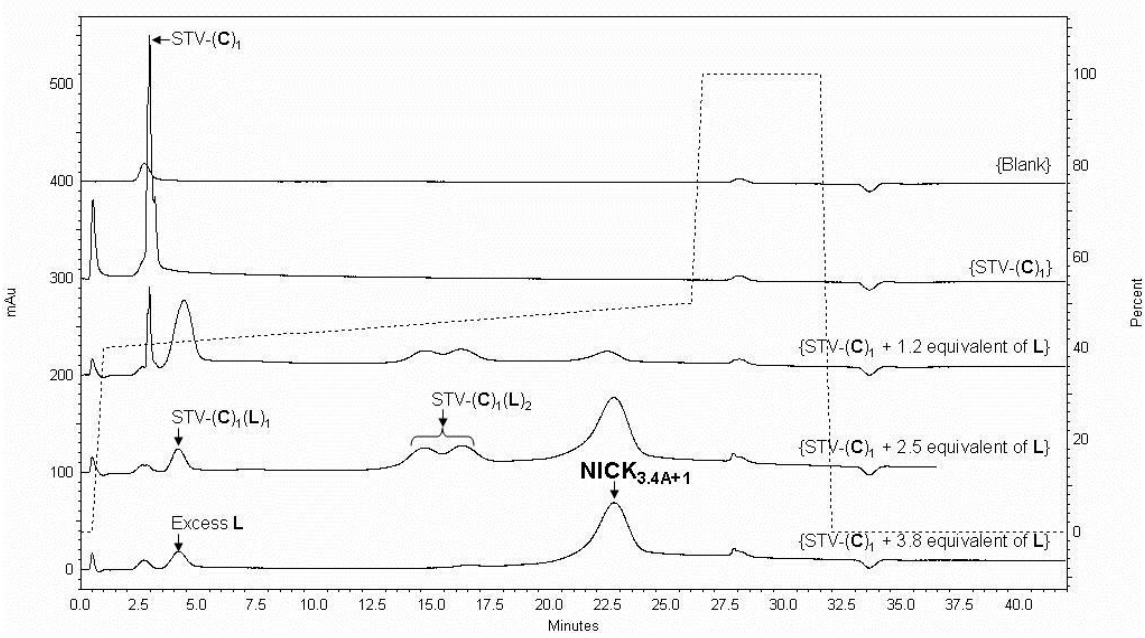


Figure D.3 Gel characterization of spider assembly. PAGE characterization of **NICK$_{3.4A+1}$** showing that isolated **NICK$_{3.4A+1}$ (**lane 1**)** contains the strand C, i.e. the capture strand 5′ - GCC GAG AAC CTG ACG CAA GT/iSp18//iSp18//3Bio/ - 3′, and strand L, i.e. the deoxyribozyme or "leg" strand 5′ - /5BioTEG//iSp18//iSp18/TCT CTT CTC CGA GCC GGT CGA AAT AGT GAA AA - 3′ in a ratio of 1:3. Native stacking gel with a 12% acrylamide separation layer and a 4% acrylamide stacking layer; running buffer is TRIS-glycine. Bands were stained with SYBR Gold (Invitrogen). STV is streptavidin. 1XC is an assembly consisting of one streptavidin conjugated to one capture strand C, 2XC is an assembly consisting of one streptavidin conjugated to two capture strands, etc (assignments of bands 1XC, 2XC, 3XC, and 4XC are made based on results shown in Supplementary Fig. 1). Lane 1 is the isolated **NICK$_{3.4A+1}$** assembly; Lane 2 is the isolated streptavidin-(mono)capture strand conjugate (STV-(C)$_1$) used to form **NICK$_{3.4A+1}$** by adding the "leg" strand L to the three remaining biotin binding sites; Lane 3 is the unpurified result on adding 3.5 equivalents of "leg" strand, L, to STV-(C)$_1$; Lane 4 is the titration of a half equivalent of C with STV-(C)$_1$ showing migration distances of STV-(C)$_n$ (where n = 1-3); Lane 5 is the titration of a half equivalent of C with STV; Lane 6 the titration of a excess C with STV; Lane 7 is the titration of a half equivalent of L with STV, where 1XL is an assembly consisting of one streptavidin conjugated to one "leg" strand L, 2XL is an assembly consisting of one streptavidin conjugated to two "leg" strands L etc. (assignments of bands 1XL, 2XL, 3XL, and 4XL are made based on results shown in Supplementary Fig. 4 and in reference 20); Lane 8 is the titration of excess L with STV.

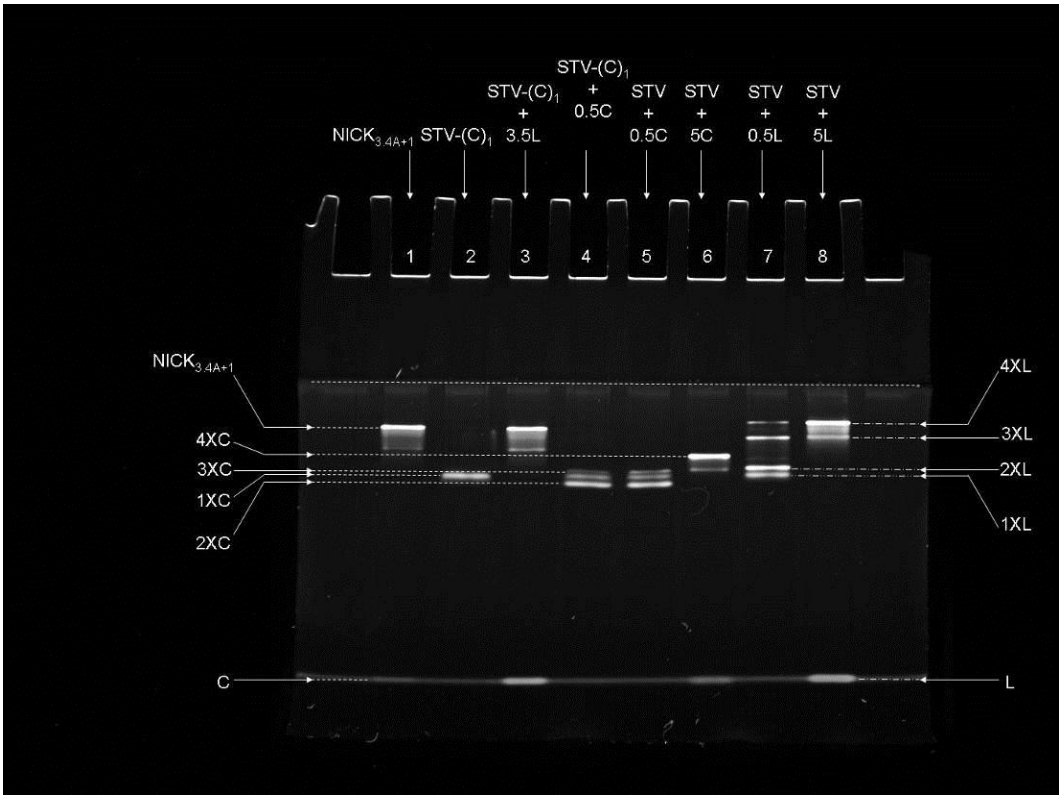


Figure D.4 Gel characterization of assembled spiders. PAGE characterization of **NICK**$_{\mathbf{n.4A}}$ (where n = 1-to-4) supporting assignments of lane 8 in Supplementary Fig. 3. Native stacking gel with a 10% acrylamide separation layer and a 4% acrylamide stacking layer; running buffer is TRIS-glycine. Bands were stained with SYBR Gold (Invitrogen). STV-(**L**)$_n$ conjugates used in this gel were isolated and characterized as previously described above (Supplementary Fig. 1). 1XL is an assembly consisting of one streptavidin conjugated to one "leg" (i.e. deoxyribozyme strand L), 2XL is an assembly consisting of one streptavidin conjugated to two "legs," etc.

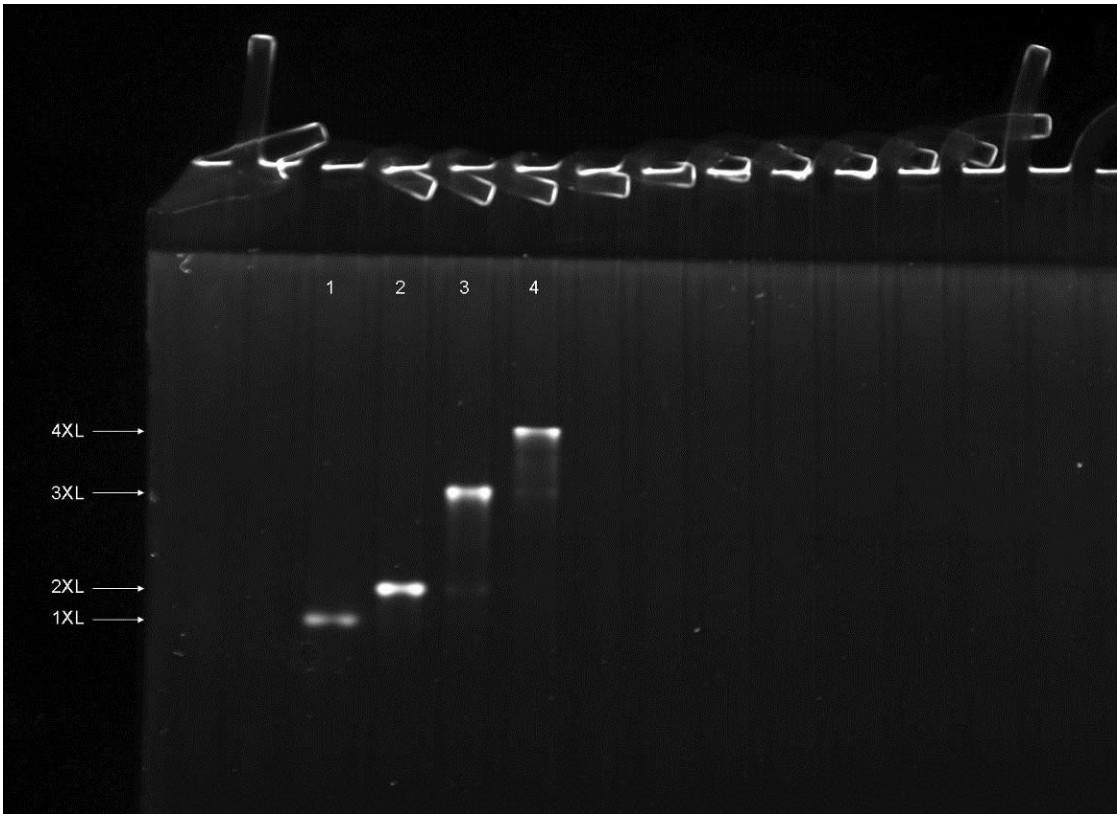


Figure D.5 Spider dissociation. Dissociation curves for **NICK$_{3.4A+1}$** spider from non-cleavable substrate (black trace, 1:89 ratio of spider to substrate) and product (green trace, 1:97 ratio of spider to product) on the 2D monolayer surfaces.

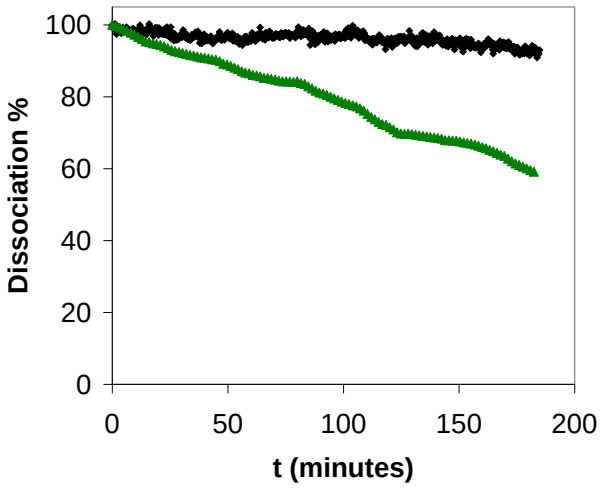
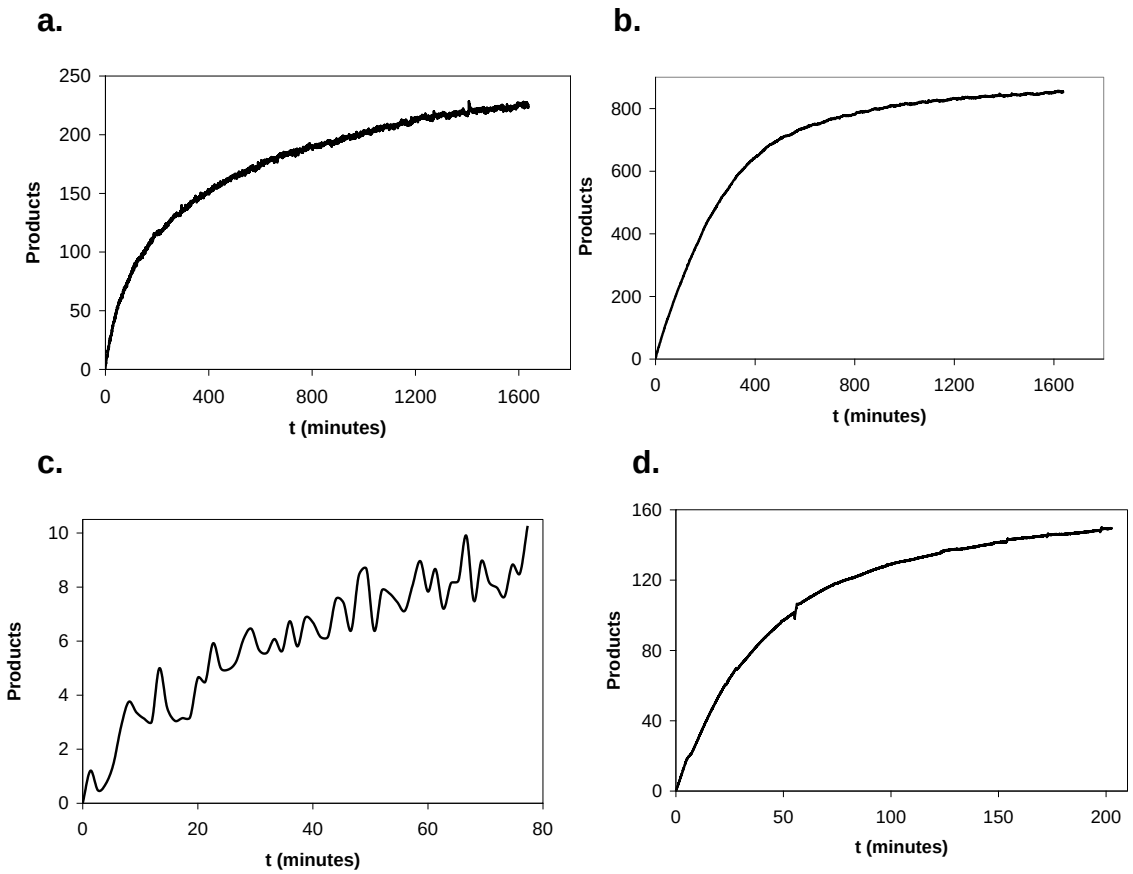


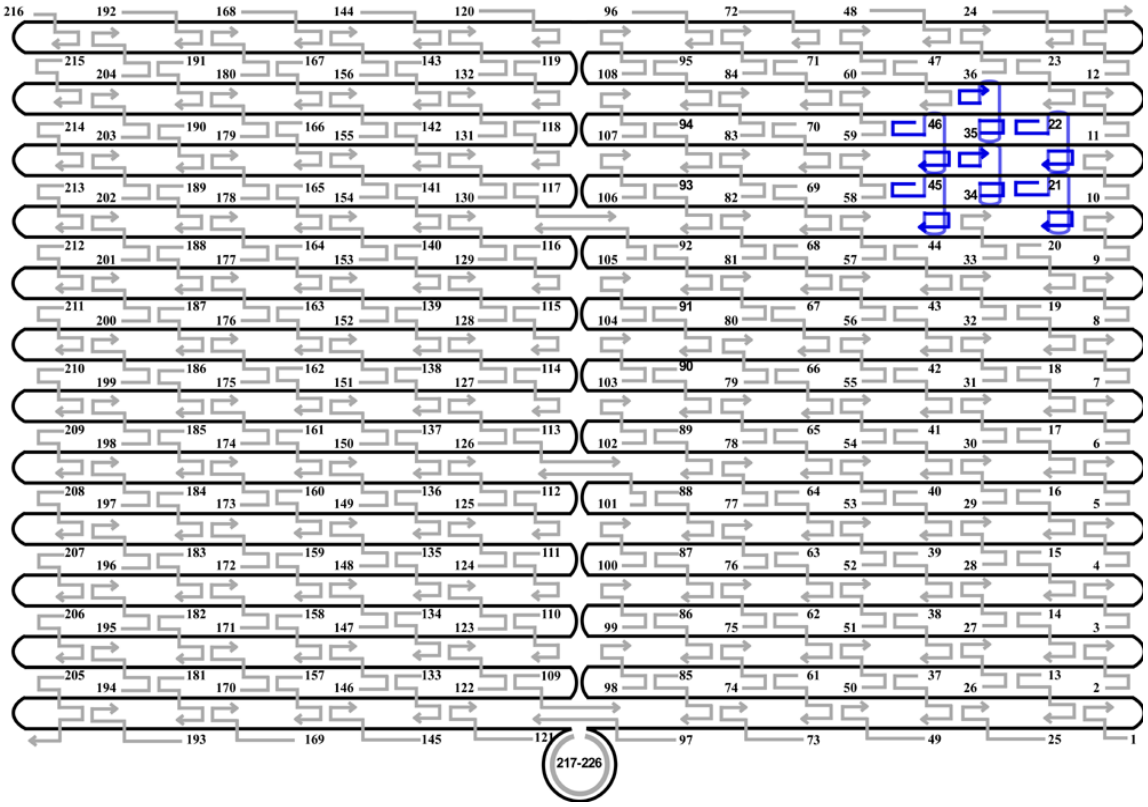
Figure D.6 Spider cleavage sensorgrams. a. Sensorgram (y-axis is products released per spider, the number of products released was obtained by conversion of SPR response unit (RU) to mass using the standard formula 1,000 RU= 1ng·mm^{-2}) of **NICK$_{3.4A+1}$** spider on the 2D monolayer surface showing the real-time substrate cleavage at a 1:291 ratio of spider to substrate with a cleavage rate of 1.42 min^{-1} per spider in 1× TA-Mg buffer with 1mM $ZnCl_2$. **b**. Sensorgram of **NICK$_{3.4A+1}$** spider on the pseudo-2D matrix surface showing the real-time substrate cleavage at a 1:990 ratio of spider to substrate with a cleavage rate of 2.81 min^{-1} per spider in 1× TA-Mg buffer with 1mM $ZnCl_2$. **c**. Sensorgram of **NICK$_{3.4A+1}$·(Cy3)$_4$** spider on the pseudo-2D matrix surface showing the real-time substrate cleavage at a 1:50 ratio of spider to substrate with a cleavage rate of 0.18 min^{-1} per spider in 1× SSC with 2 mM $ZnSO_4$. **d**. Sensorgram of **NICK$_{3.4A+1}$·(Cy3)$_4$** spider on the pseudo-2D matrix surface showing the real-time substrate cleavage at a 1:180 ratio of spider to substrate with a cleavage rate of 2.72 min^{-1} per spider in HBS buffer with 1mM $ZnSO_4$. All cleavage reactions were monitored with a flow rate of 20 μL/min.



DNA Sequences

M13mp18 sequence can be found at the following web-address http://www.neb.com/nebecomm/tech_reference/restriction_enzymes/sequences/m13mp18.txt.

Figure D.7 Schematic of the rectangular shaped DNA origami structure with the staple strand location and numbering marked. This is a representation of a plain origami structure with the marker included. In this drawing, the continuous black colored strand represents circular M13 viral genome and all the staple strands are shown in grey with arrows pointing the 3′- ends of the sequences. Numbers denote the sequence of the strands below. The blue strands denote the dumbbell hairpins used as a marker to aid in identification of origami by AFM.



Name	Sequence
1	TTTTCGATGGCCCACTACGTAAACCGTC
2	TATCAGGGTTTTCGGTTTGCGTATTGGGAACGCGCG
3	GGGAGAGGTTTTTGTAAAACGACGGCCATTCCCAGT
3A	GGGAGAGGTTTTTGTAAAAC
3B	Biotin GACGGCCATTCCCAGT
4	CACGACGTTTTTGTAATGGGATAGGTCAAAACGGCG
5	GATTGACCTTTTGATGAACGGTAATCGTAGCAAACA
6	AGAGAATCTTTTGGTTGTACCAAAAACAAGCATAAA
7	GCTAAATCTTTTCTGTAGCTCAACATGTATTGCTGA
8	ATATAATGTTTTCATTGAATCCCCCTCAAATCGTCA
9	TAAATATTTTTTGGAAGAAAAATCTACGACCAGTCA
10	GGACGTTGTTTTTCATAAGGGAACCGAAAGGCGCAG

11 ACGGTCAATTTTGACAGCATCGGAACGAACCCTCAG
11A ACGGTCAATTTTGACAGCAT
11B Biotin CGGAACGAACCCTCAG
12 CAGCGAAAATTTTACTTTCAACAGTTTCTGGGATTTTGCTAAACTTTT
13 TGGTTTTTAACGTCAAAGGGCGAAGAACCATC
14 CTTGCATGCATTAATGAATCGGCCCGCCAGGG
15 TAGATGGGGGGTAACGCCAGGGTTGTGCCAAG
16 CATGTCAAGATTCTCCGTGGGAACCGTTGGTG
17 CTGTAATATTGCCTGAGAGTCTGGAAAACTAG
18 TGCAACTAAGCAATAAAGCCTCAGTTATGACC
19 AAACAGTTGATGGCTTAGAGCTTATTTAAATA
20 ACGAACTAGCGTCCAATACTGCGGAATGCTTT
21 CTTTGAAAAGAACTGGTCCTCTTTTGAGGAACAAGTTTTCTTGT CTCATTATTTAATAAA
22 ACGGCTACTTACTTAGTCCTCTTTTGAGGAACAAGTTTTCTTGT CCGGAACGCTGACCAA
23 GAGAATAGCTTTTGCGGGATCGTCGGGTAGCA
24 ACGTTAGTAAATGAATTTTCTGTAAGCGGAGT
25 ACCCAAATCAAGTTTTTTGGGGTCAAAGAACG
26 TGGACTCCCTTTTCACCAGTGAGACCTGTCGT
27 GCCAGCTGCCTGCAGGTCGACTCTGCAAGGCG
28 ATTAAGTTCGCATCGTAACCGTGCGAGTAACA
29 ACCCGTCGTCATATGTACCCCGGTAAAGGCTA
30 TCAGGTCACTTTTGCGGGAGAAGCAGAATTAG
31 CAAAATTAAAGTACGGTGTCTGGAAGAGGTCA
32 TTTTTGCGCAGAAAACGAGAATGAATGTTTAG
33 ACTGGATAACGGAACAACATTATTACCTTATG
34 CGATTTTAGAGGACAGTCCTCTTTTGAGGAACAAGTTTTCTTGT ATGAACGGCGCGACCT
35 GCTCCATGAGAGGCTTTCCTCTTTTGAGGAACAAGTTTTCTTGT TGAGGACTAGGGAGTT
36 AAAGGCCGAAAGGAACAACTAAAGCTTTCCAG
37 AGCTGATTACAAGAGTCCACTATTGAGGTGCC
38 CCCGGGTACTTTCCAGTCGGGAAACGGGCAAC
39 GTTTGAGGGAAAGGGGGATGTGCTAGAGGATC
40 AGAAAAGCAACATTAAATGTGAGCATCTGCCA
41 CAACGCAATTTTTGAGAGATCTACTGATAATC
42 TCCATATACATACAGGCAAGGCAACTTTATTT
43 CAAAAATCATTGCTCCTTTTGATAAGTTTCAT
44 AAAGATTCAGGGGGTAATAGTAAACCATAAAT
45 CCAGGCGCTTAATCATTCCTCTTTTGAGGAACAAGTTTTCTTGT TGTGAATTACAGGTAG
46 TTTCATGAAAATTGTGTCCTCTTTTGAGGAACAAGTTTTCTTGT TCGAAATCTGTACAGA
47 AATAATAAGGTCGCTGAGGCTTGCAAAGACTT
48 CGTAACGATCTAAAGTTTTGTCGTGAATTGCG

49	GTAAAGCACTAAATCGGAACCCTAGTTGTTCC
50	AGTTTGGAGCCCTTCACCGCCTGGTTGCGCTC
51	ACTGCCCGCCGAGCTCGAATTCGTTATTACGC
52	CAGCTGGCGGACGACGACAGTATCGTAGCCAG
53	CTTTCATCCCCAAAAACAGGAAGACCGGAGAG
53A	CTTTCATCCCCAAAAA
53B	Biotin CAGGAAGACCGGAGAG
54	GGTAGCTAGGATAAAAATTTTTAGTTAACATC
55	CAATAAATACAGTTGATTCCCAATTTAGAGAG
56	TACCTTTAAGGTCTTTACCCTGACAAAGAAGT
57	TTTGCCAGATCAGTTGAGATTTAGTGGTTTAA
57A	TTTGCCAGATCAGTTG
57B	Biotin AGATTTAGTGGTTTAA
58	TTTCAACTATAGGCTGGCTGACCTTGTATCAT
59	CGCCTGATGGAAGTTTCCATTAAACATAACCG
60	ATATATTCTTTTTTCACGTTGAAAATAGTTAG
61	GAGTTGCACGAGATAGGGTTGAGTAAGGGAGC
62	TCATAGCTACTCACATTAATTGCGCCCTGAGA
63	GAAGATCGGTGCGGGCCTCTTCGCAATCATGG
64	GCAAATATCGCGTCTGGCCTTCCTGGCCTCAG
65	TATATTTTAGCTGATAAATTAATGTTGTATAA
66	CGAGTAGAACTAATAGTAGTAGCAAACCCTCA
67	TCAGAAGCCTCCAACAGGTCAGGATCTGCGAA
68	CATTCAACGCGAGAGGCTTTTGCATATTATAG
69	AGTAATCTTAAATTGGGCTTGAGAGAATACCA
70	ATACGTAAAAGTACAACGGAGATTTCATCAAG
71	AAAAAAGGACAACCATCGCCCACGCGGGTAAA
72	TGTAGCATTCCACAGACAGCCCTCATCTCCAA
73	CCCCGATTTAGAGCTTGACGGGGAAATCAAAA
74	GAATAGCCGCAAGCGGTCCACGCTCCTAATGA
75	GTGAGCTAGTTTCCTGTGTGAAATTTGGGAAG
76	GGCGATCGCACTCCAGCCAGCTTTGCCATCAA
77	AAATAATTTTAAATTGTAAACGTTGATATTCA
78	ACCGTTCTAAATGCAATGCCTGAGAGGTGGCA
79	TCAATTCTTTTAGTTTGACCATTACCAGACCG
80	GAAGCAAAAAAGCGGATTGCATCAGATAAAAA
81	CCAAAATATAATGCAGATACATAAACACCAGA
82	ACGAGTAGTGACAAGAACCGGATATACCAAGC
83	GCGAAACATGCCACTACGAAGGCATGCGCCGA
84	CAATGACACTCCAAAAGGAGCCTTACAACGCC
85	CCAGCAGGGGCAAAATCCCTTATAAAGCCGGC
86	GCTCACAATGTAAAGCCTGGGGTGGGTTTGCC
87	GCTTCTGGTCAGGCTGCGCAACTGTGTTATCC
88	GTTAAAATTTTAACCAATAGGAACCCGGCACC
89	AGGTAAAGAAATCACCATCAATATAATATTTT
90	TCGCAAATGGGGCGCGAGCTGAAATAATGTGT

91 AAGAGGAACGAGCTTCAAAGCGAAGATACATT
92 GGAATTACTCGTTTACCAGACGACAAAAGATT
93 CCAAATCACTTGCCCTGACGAGAACGCCAAAA
94 AAACGAAATGACCCCCAGCGATTATTCATTAC
95 TCGGTTTAGCTTGATACCGATAGTCCAACCTA
96 TGAGTTTCGTCACCAGTACAAACTTAATTGTA
97 GAACGTGGCGAGAAAGGAAGGGAACAAACTAT
98 CCGAAATCCGAAAATCCTGTTTGAAGCCGGAA
99 GCATAAAGTTCCACACAACATACGAAGCGCCA
100 TTCGCCATTGCCGGAAACCAGGCATTAAATCA
101 GCTCATTTTCGCATTAAATTTTTGAGCTTAGA
102 AGACAGTCATTCAAAAGGGTGAGAAGCTATAT
103 TTTCATTTGGTCAATAACCTGTTTATATCGCG
103A TTTCATTTGGTCAATA
103B Biotin ACCTGTTTATATCGCG
104 TTTTAATTGCCCGAAAGACTTCAAAACACTAT
105 CATAACCCGAGGCATAGTAAGAGCTTTTTAAG
106 GAATAAGGACGTAACAAAGCTGCTCTAAAACA
107 CTCATCTTGAGGCAAAAGAATACAGTGAATTT
108 CTTAAACATCAGCTTGCTTTCGAGCGTAACAC
109 ACGAACCAAAACATCGCCATTAAATGGTGGTT
110 CGACAACTAAGTATTAGACTTTACAATACCGA
111 CTTTTACACAGATGAATATACAGTAAACAATT
112 TTAAGACGTTGAAAACATAGCGATAACAGTAC
113 GCGTTATAGAAAAAGCCTGTTTAGAAGGCCGG
114 ATCGGCTGCGAGCATGTAGAAACCTATCATAT
115 CCTAATTTACGCTAACGAGCGTCTAATCAATA
116 AAAAGTAATATCTTACCGAAGCCCTTCCAGAG
117 TTATTCATAGGGAAGGTAAATATTCATTCAGT
118 GAGCCGCCCCACCACCGGAACCGCGACGGAAA
119 AATGCCCCGTAACAGTGCCCGTATCTCCCTCA
120 CAAGCCCAATAGGAACCCATGTACAAACAGTT
121 CGGCCTTGCTGGTAATATCCAGAACGAACTGA
122 TAGCCCTACCAGCAGAAGATAAAAACATTTGA
123 GGATTTAGCGTATTAAATCCTTTGTTTTCAGG
124 TTTAACGTTCGGGAGAAACAATAATTTTCCCT
125 TAGAATCCCTGAGAAGAGTCAATAGGAATCAT
126 AATTACTACAAATTCTTACCAGTAATCCCATC
127 CTAATTTATCTTTCCTTATCATTCATCCTGAA
128 TCTTACCAGCCAGTTACAAAATAAATGAAATA
129 GCAATAGCGCAGATAGCCGAACAATTCAACCG
130 ATTGAGGGTAAAGGTGAATTATCAATCACCGG
128 AACCAGAGACCCTCAGAACCGCCAGGGGTCAG
132 TGCCTTGACTGCCTATTTCGGAACAGGGATAG
133 AGGCGGTCATTAGTCTTTAATGCGCAATATTA
134 TTATTAATGCCGTCAATAGATAATCAGAGGTG

135 CCTGATTGAAAGAAATTGCGTAGACCCGAACG
136 ATCAAAATCGTCGCTATTAATTAACGGATTCG
137 ACGCTCAAAATAAGAATAAACACCGTGAATTT
138 GGTATTAAGAACAAGAAAAATAATTAAAGCCA
139 ATTATTTAACCCAGCTACAATTTTCAAGAACG
140 GAAGGAAAATAAGAGCAAGAAACAACAGCCAT
141 GACTTGAGAGACAAAAGGGCGACAAGTTACCA
142 GCCACCACTCTTTTCATAATCAAACCGTCACC
143 CTGAAACAGGTAATAAGTTTTAACCCCTCAGA
144 CTCAGAGCCACCACCCTCATTTTCCTATTATT
145 CCGCCAGCCATTGCAACAGGAAAAATATTTTT
146 GAATGGCTAGTATTAACACCGCCTCAACTAAT
147 AGATTAGATTTAAAAGTTTGAGTACACGTAAA
148 ACAGAAATCTTTGAATACCAAGTTCCTTGCTT
149 CTGTAAATCATAGGTCTGAGAGACGATAAATA
150 AGGCGTTACAGTAGGGCTTAATTGACAATAGA
151 TAAGTCCTACCAAGTACCGCACTCTTAGTTGC
152 TATTTTGCTCCCAATCCAAATAAGTGAGTTAA
153 GCCCAATACCGAGGAAACGCAATAGGTTTACC
154 AGCGCCAACCATTTGGGAATTAGATTATTAGC
155 GTTTGCCACCTCAGAGCCGCCACCGATACAGG
156 AGTGTACTTGAAAGTATTAAGAGGCCGCCACC
157 GCCACGCTATACGTGGCACAGACAACGCTCAT
158 ATTTTGCGTCTTTAGGAGCACTAAGCAACAGT
159 GCGCAGAGATATCAAAATTATTTGACATTATC
160 TAACCTCCATATGTGAGTGAATAAACAAAATC
160A TAACCTCCATATGTGA
160B Biotin GTGAATAAACAAAATC
161 CATATTTAGAAATACCGACCGTGTTACCTTTT
162 CAAGCAAGACGCGCCTGTTTATCAAGAATCGC
163 TTTTGTTTAAGCCTTAAATCAAGAATCGAGAA
164 ATACCCAAGATAACCCACAAGAATAAACGATT
164A ATACCCAAGATAACCC
164B Biotin ACAAGAATAAACGATT
165 AATCACCAAATAGAAAATTCATATATAACGGA
166 CACCAGAGTTCGGTCATAGCCCCCGCCAGCAA
167 CCTCAAGAATACATGGCTTTTGATAGAACCAC
168 CCCTCAGAACCGCCACCCTCAGAACTGAGACT
169 GGAAATACCTACATTTTGACGCTCACCTGAAA
170 GCGTAAGAGAGAGCCAGCAGCAAAAAGGTTAT
171 CTAAAATAGAACAAAGAAACCACCAGGGTTAG
172 AACCTACCGCGAATTATTCATTTCCAGTACAT
173 AAATCAATGGCTTAGGTTGGGTTACTAAATTT
174 AATGGTTTACAACGCCAACATGTAGTTCAGCT
175 AATGCAGACCGTTTTTATTTTCATCTTGCGGG
176 AGGTTTTGAACGTCAAAAATGAAAGCGCTAAT

177 ATCAGAGAAAGAACTGGCATGATTTTATTTTG
178 TCACAATCGTAGCACCATTACCATCGTTTTCA
179 TCGGCATTCCGCCGCCAGCATTGACGTTCCAG
180 TAAGCGTCGAAGGATTAGGATTAGTACCGCCA
181 CTAAAGCAAGATAGAACCCTTCTGAATCGTCT
182 CGGAATTATTGAAAGGAATTGAGGTGAAAAAT
183 GAGCAAAAACTTCTGAATAATGGAAGAAGGAG
184 TATGTAAACCTTTTTTAATGGAAAAATTACCT
185 AGAGGCATAATTTCATCTTCTGACTATAACTA
186 TCATTACCCGACAATAAACAACATATTTAGGC
187 CTTTACAGTTAGCGAACCTCCCGACGTAGGAA
188 TTATTACGGTCAGAGGGTAATTGAATAGCAGC
189 CCGGAAACACACCACGGAATAAGTAAGACTCC
190 TGAGGCAGGCGTCAGACTGTAGCGTAGCAAGG
191 TGCTCAGTCAGTCTCTGAATTTACCAGGAGGT
192 TATCACCGTACTCAGGAGGTTTAGCGGGGTTT
193 GAAATGGATTATTTACATTGGCAGACATTCTG
194 GCCAACAGTCACCTTGCTGAACCTGTTGGCAA
195 ATCAACAGTCATCATATTCCTGATTGATTGTT
196 TGGATTATGAAGATGATGAAACAAAATTTCAT
197 TTGAATTATGCTGATGCAAATCCACAAATATA
198 TTTTAGTTTTTCGAGCCAGTAATAAATTCTGT
199 CCAGACGAGCGCCCAATAGCAAGCAAGAACGC
200 GAGGCGTTAGAGAATAACATAAAAGAACACCC
201 TGAACAAACAGTATGTTAGCAAACTAAAAGAA
202 ACGCAAAGGTCACCAATGAAACCAATCAAGTT
203 TGCCTTTAGTCAGACGATTGGCCTGCCAGAAT
204 GGAAAGCGACCAGGCGGATAAGTGAATAGGTG
205 AAACCCTCTTTTACCAGTAATAAAAGGGATTCACCAGTCACACGTTTT
206 GATGGCAATTTTAATCAATATCTGGTCACAAATATC
206A GATGGCAATTTTAATCAATA
206B Biotin TCTGGTCACAAATATC
207 AAAACAAATTTTTTCATCAATATAATCCTATCAGAT
208 ACAAAGAATTTTATTAATTACATTTAACACATCAAG
209 TAAAGTACTTTTCGCGAGAAAACTTTTTATCGCAAG
210 TATAGAAGTTTTCGACAAAAGGTAAAGTAGAGAATA
211 GCGCATTATTTTGCTTATCCGGTATTCTAAATCAGA
212 TACATACATTTTGACGGGAGAATTAACTACAGGGAA
213 AGCACCGTTTTTTAAAGGTGGCAACATAGTAGAAAA
214 ACAAACAATTTTAATCAGTAGCGACAGATCGATAGC
214A ACAAACAATTTTAATCAGTA
214B Biotin GCGACAGATCGATAGC
215 AGGGTTGATTTTATAAATCCTCATTAAATGATATTC
216 TTTTTATAAGTATAGCCCGGCCGTCGAG
217 AACATCACTTGCCTGAGTAGAAGAACT
218 TGTAGCAATACTTCTTTGATTAGTAAT

219 AGTCTGTCCATCACGCAAATTAACCGT
220 ATAATCAGTGAGGCCACCGAGTAAAAG
221 ACGCCAGAATCCTGAGAAGTGTTTTT
222 TTAAAGGGATTTTAGACAGGAACGGT
223 AGAGCGGGAGCTAAACAGGAGGCCGA
224 TATAACGTGCTTTCCTCGTTAGAATC
225 GTACTATGGTTGCTTTGACGAGCACG
226 GCGCTTAATGCGCCGCTACAGGGCGC

The following three sequences are attached to the 5′ end of the staple sequences, as a probe, for the START position, binding of the cleavable substrate, and binding of the non-cleavable substrate. For fluorescence microscopy, strands 3A, 3B, 11A, 11B, 206A, 206B, 214A, 214B were incorporated into the origami and CONTROL staples were replaced with staples lacking the non-cleave-able substrate probes.

Spider START (green)
5′- GATGTCTACTTGCGTCAGGTTCTCGGC[staple]

Spider Cleavable Substrate Probes (brown)
5′- CCTCTCACCCACCATTCATC[staple]

Spider Non-Cleavable Substrate Probes (for STOP and CONTROL; red)
5′- GGTTCAGTTCGTTGAGCCAG[staple]

Spider Cleavable Substrate
5′- GATGAATGGTGGGTGAGAGGTTTTTCACTATrAGGAAGAG

Spider Non-Cleavable Substrate (STOP and CONTROL)
5′- CTGGCTCAACGAACTGAACC TTTTTCACTATAGGAAGAG

Spider Non-Cleavable Substrate (STOP) for fluorescence microscopy
5′- CTGGCTCAACGAACTGAACC TTTTTCACTATAGGAAGAG-Cy5

Spider TRIGGER Strand
5′- GCCGAGAACCTGACGCAAGTAGACATC

Figure D.8 Schematic of the ABD origami design. Green represents the START position, brown the probes for the substrate, and red the probes for the STOP and CONTROL.

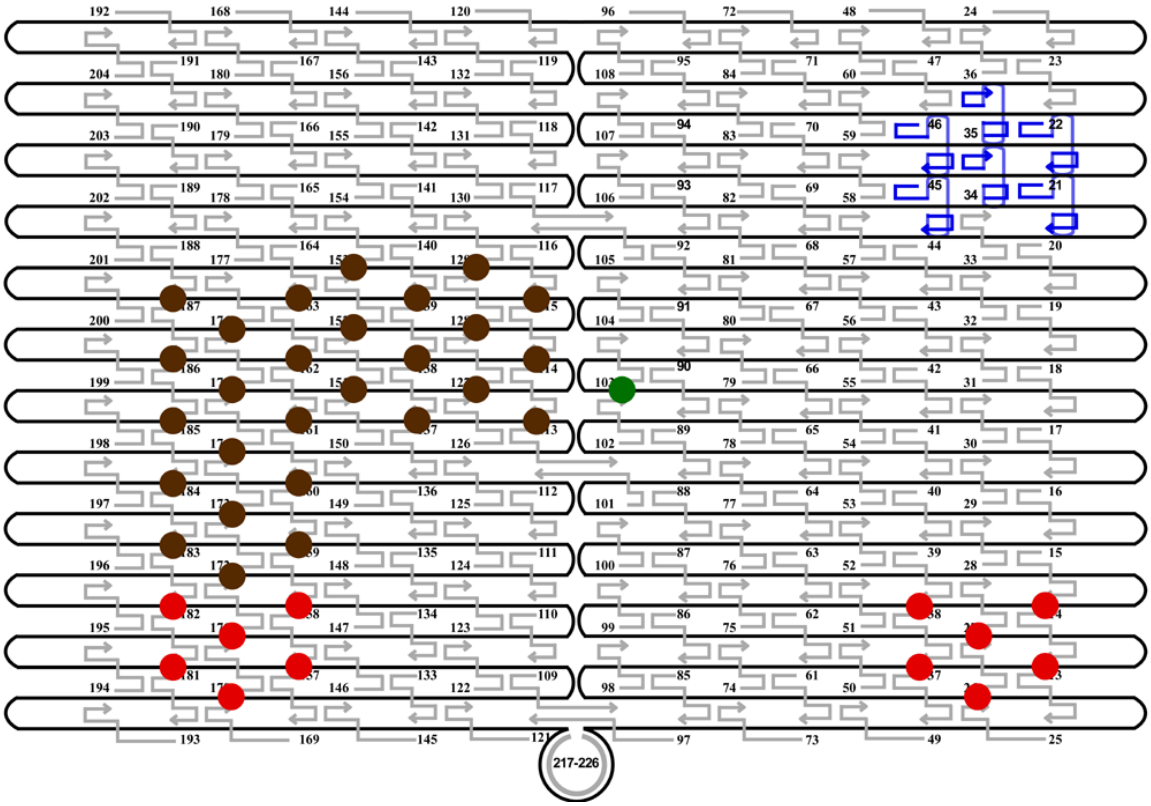
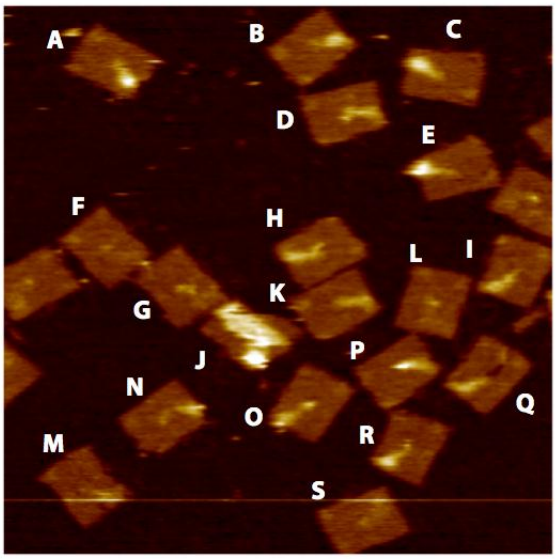
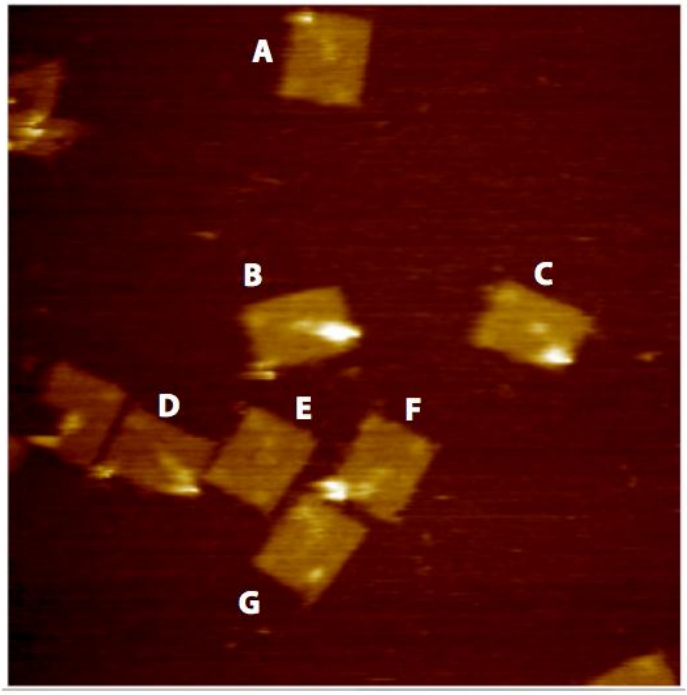


Figure D.9 Wide Field AFM images and classifications used for statistical analysis of ABD design. AFM images of the spider before release.



	up	Down	#Spiders	START	TRACK	STOP	CONTROL	inscrutable
a	0	1	1	0	0	1	0	0
b								1
c								1
d	0	1	0	0	0	0	0	0
e	0	1	1	0	0	1	0	0
f								1
g	1	0	1	1	0	0	0	0
h								1
i	0	1	0	0	0	0	0	0
j								1
k	0	1	0	0	0	0	0	0
l	1	0	1	1	0	0	0	0
m								1
n	0	1	1	0	0	1	0	0
o	0	1	1	0	0	1	0	0
p	0	1	1	1	0	0	0	0
q								1
r	0	1	1	0	0	1	0	0
s	1	0	0	0	0	0	0	0

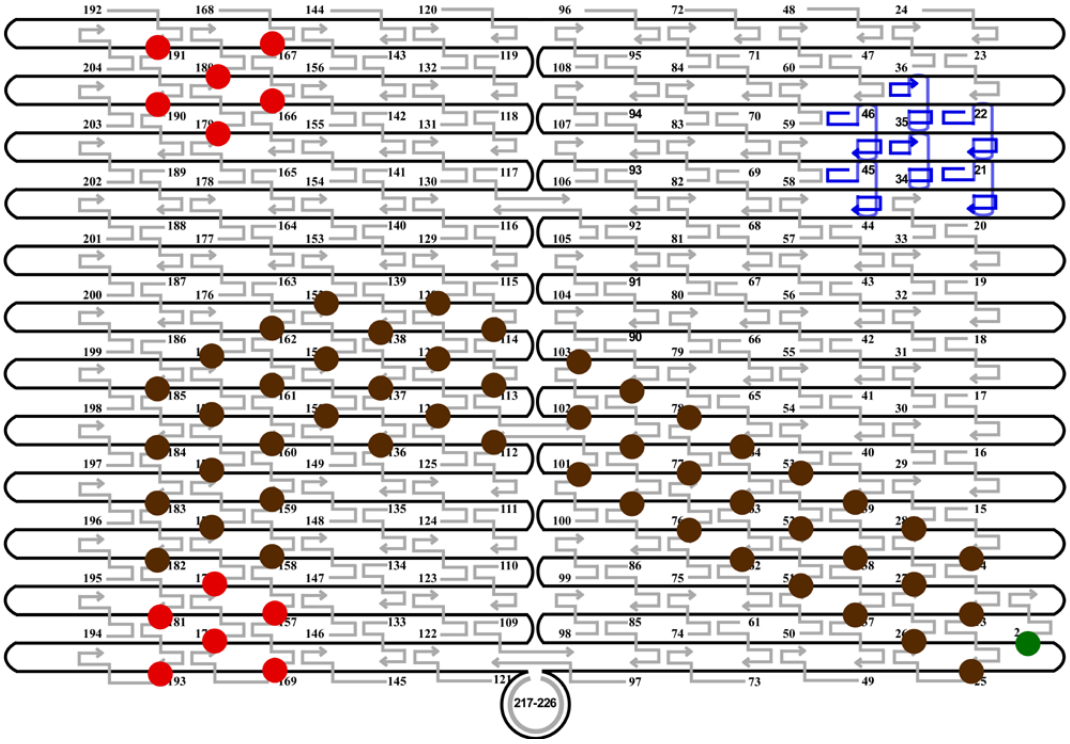
Figure D.10 Wide Field AFM images and classifications used for statistical analysis of ABD design. AFM images of the spider after release.



	up	down	#Spiders	START	TRACK	STOP	CONTROL	inscrutable
a	1	0	1	0	0	1	0	0
b	0	1	1	0	0	1	0	0
c	0	1	1	0	0	1	0	0
d	0	1	0	0	0	0	0	0
e								1
f	0	1	1	0	0	1	0	0
g	1	0	0	0	0	0	0	0

Figure D.11 Schematic of the EABD origami design. Green represents the START position, brown the probes for the substrate, and red the probes for the STOP and CONTROL. **a,** AFM design and **b,** fluorescence microscopy design.

a



b

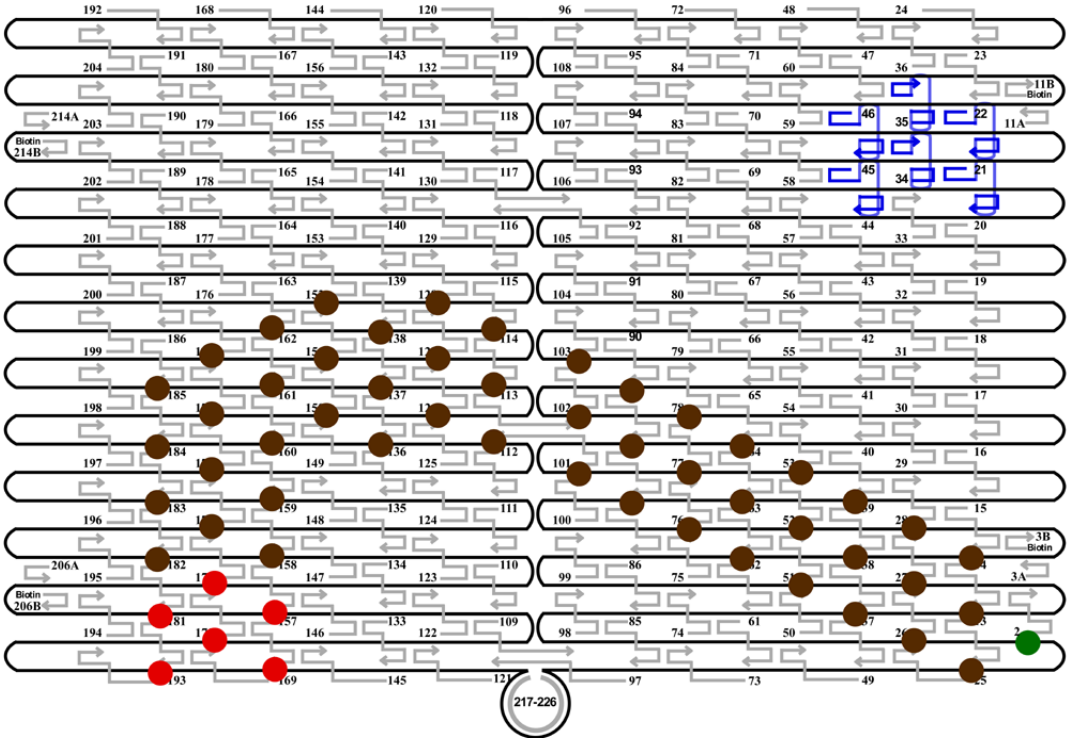
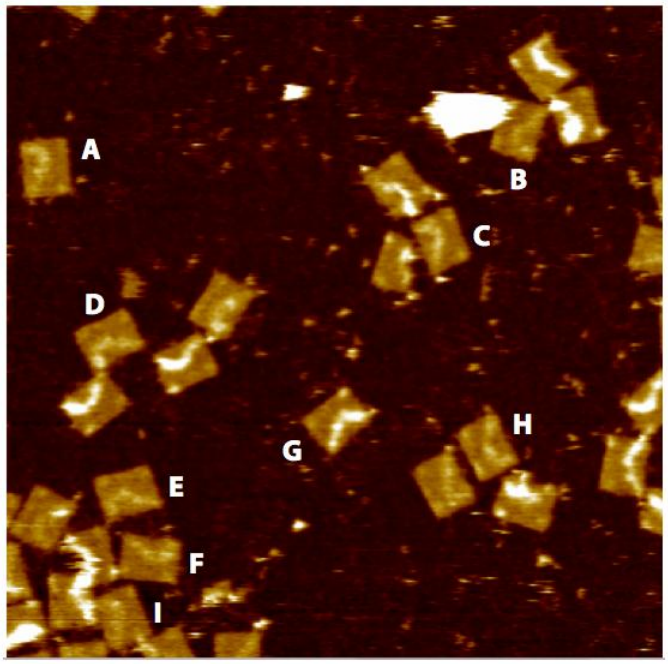
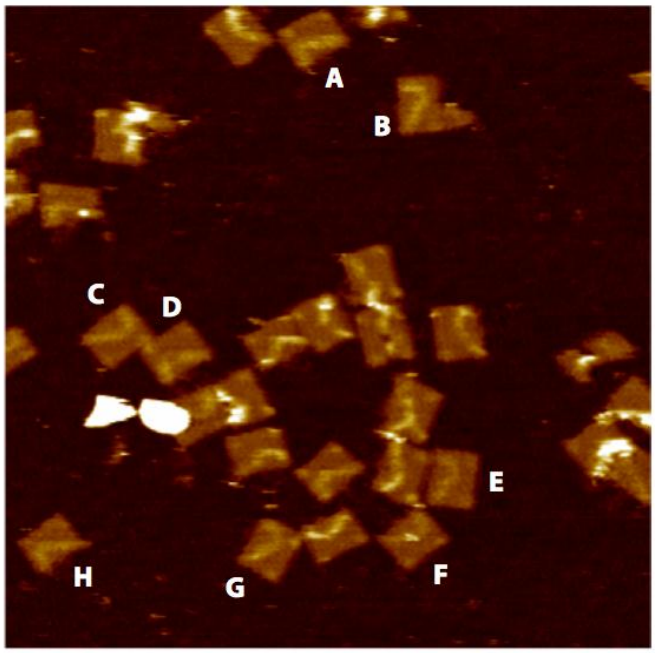


Figure D.12 Wide Field AFM images and classifications used for statistical analysis of EABD design. AFM images of the spider before release.



	up	Down	#Spiders	START	TRACK	STOP	CONTROL	inscrutable
a	1	0	0	0	0	0	0	0
b								1
c	1	0	1	1	0	0	0	0
d	1	0	1	0	0	1	0	0
e	1	0	0	0	0	0	0	0
f	1	0	0	0	0	0	0	0
g	1	0	1	1	0	0	0	0
h	1	0	1	1	0	0	0	0
i								1

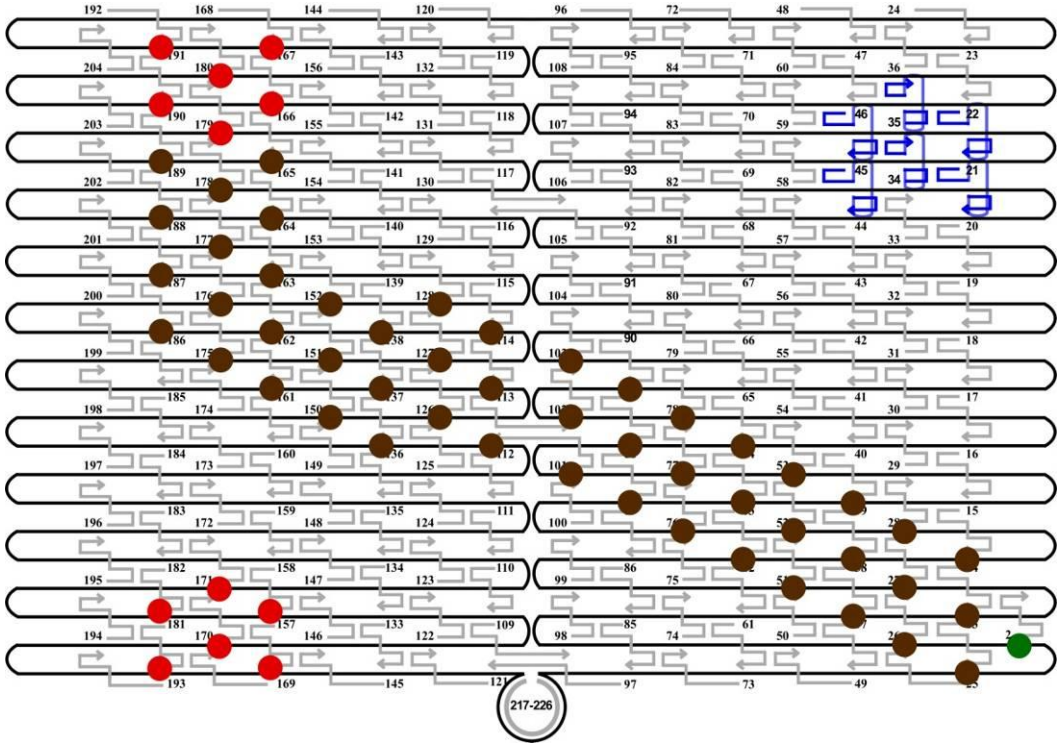
Figure D.13 Wide Field AFM images and classifications used for statistical analysis of EABD design. AFM images of the spider after release.



	up	down	#Spiders	START	TRACK	STOP	CONTROL	inscrutable
a	1	0	0	0	0	0	0	0
b								1
c	1	0	0	0	0	0	0	0
d	1	0	0	0	0	0	0	0
e	1	0	0	0	0	0	0	0
f	1	0	1	0	1	0	0	0
g	1	0	0	0	0	0	0	0
h	1	0	0	0	0	0	0	0

Figure D.14 Schematic of the EABC origami design. Green represents the START position, brown the probes for the substrate, and red the probes for the STOP and CONTROL. **a,** AFM design and **b,** fluorescence microscopy design.

a



b

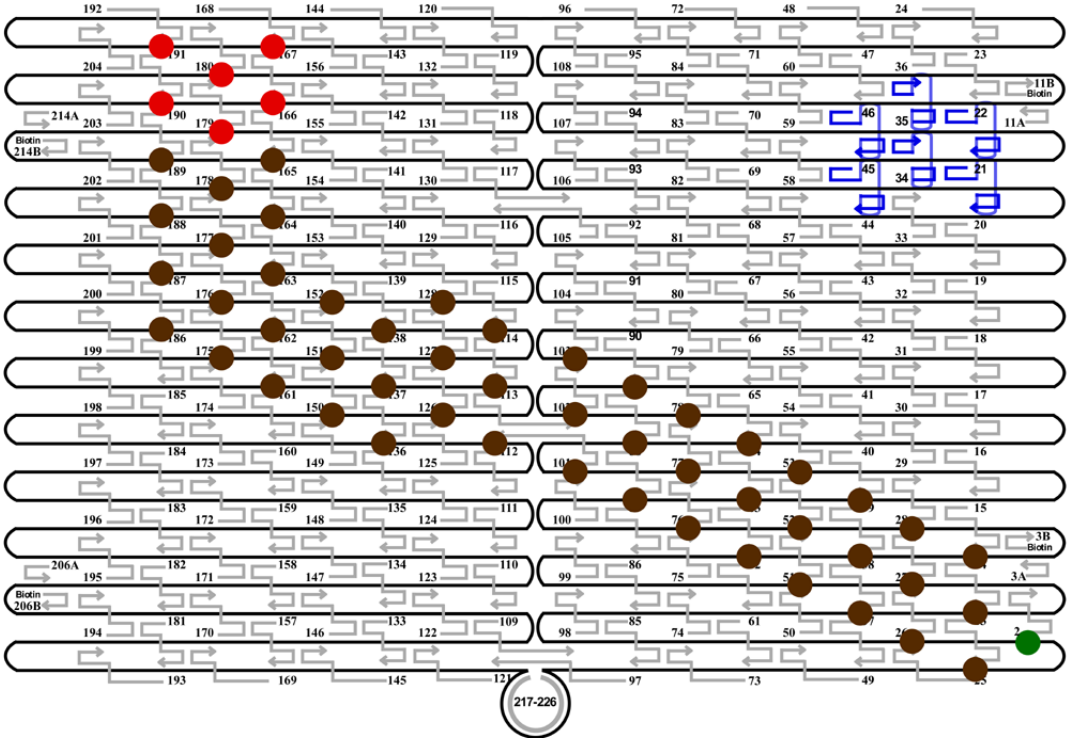
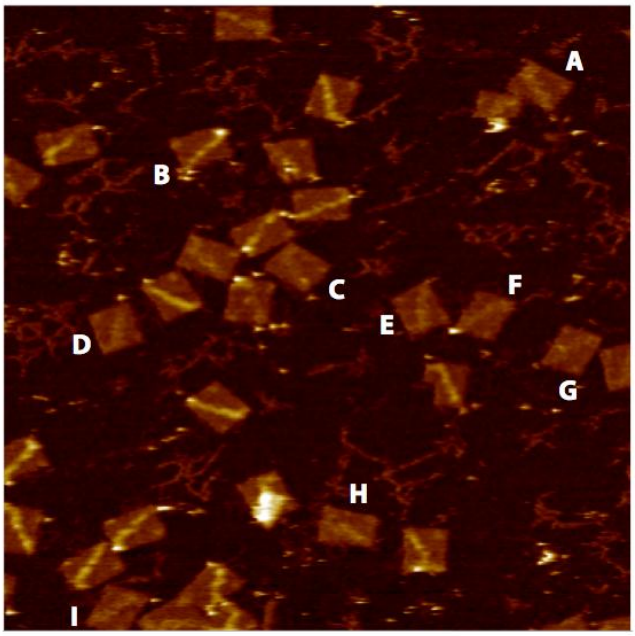
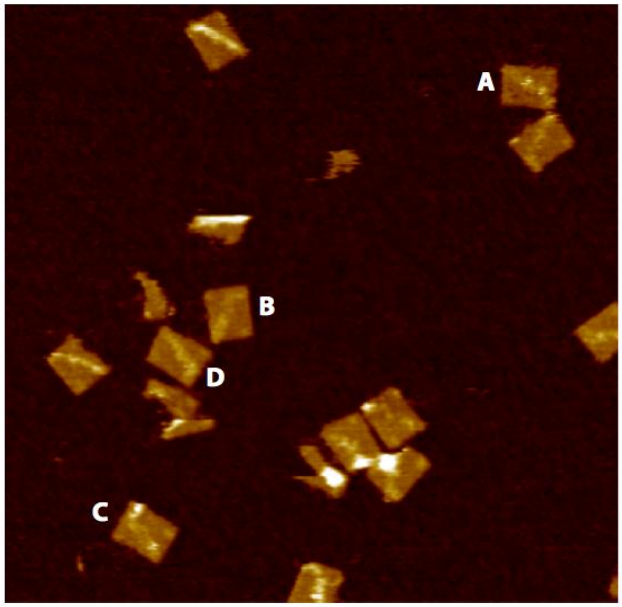


Figure D.15 Wide Field AFM images and classifications used for statistical analysis of EABC design. AFM images of the spider before release.



	up	down	#Spiders	START	TRACK	STOP	CONTROL	inscrutable
a	1	0	0	0	0	0	0	0
b	1	0	0	0	0	0	0	0
c	1	0	0	0	0	0	0	0
d	1	0	1	1	0	0	0	0
e	1	0	0	0	0	0	0	0
f	1	0	1	1	0	0	0	0
g								1
h	1	0	0	0	0	0	0	0
i	1	0	0	0	0	0	0	0

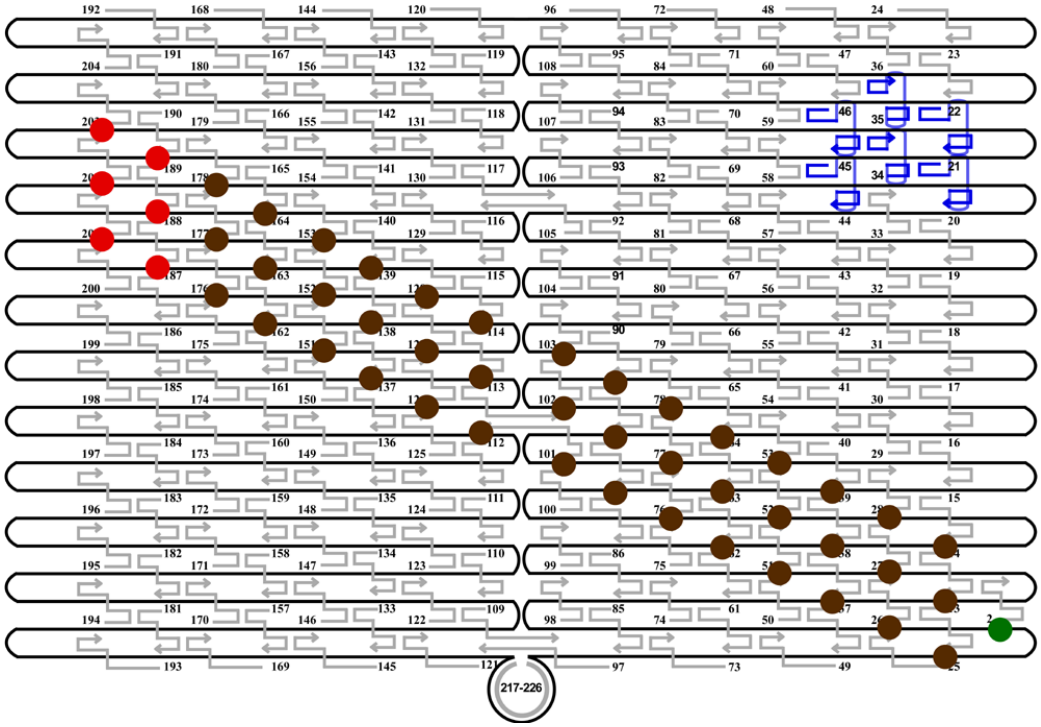
Figure D.16 Wide Field AFM images and classifications used for statistical analysis of EABC design. AFM images of the spider after release.



	up	down	#Spiders	START	TRACK	STOP	CONTROL	inscrutable
a	1	0	2	0	2	0	0	0
b	1	0	0	0	0	0	0	0
c	1	0	2	0	1	1	0	0
d	1	0	0	0	0	0	0	0

Figure D.17 Schematic of the EAC origami design. Green represents the START position, brown the probes for the substrate, and red the probes for the STOP and CONTROL. **a,** AFM design and **b,** fluorescence microscopy design.

a



b

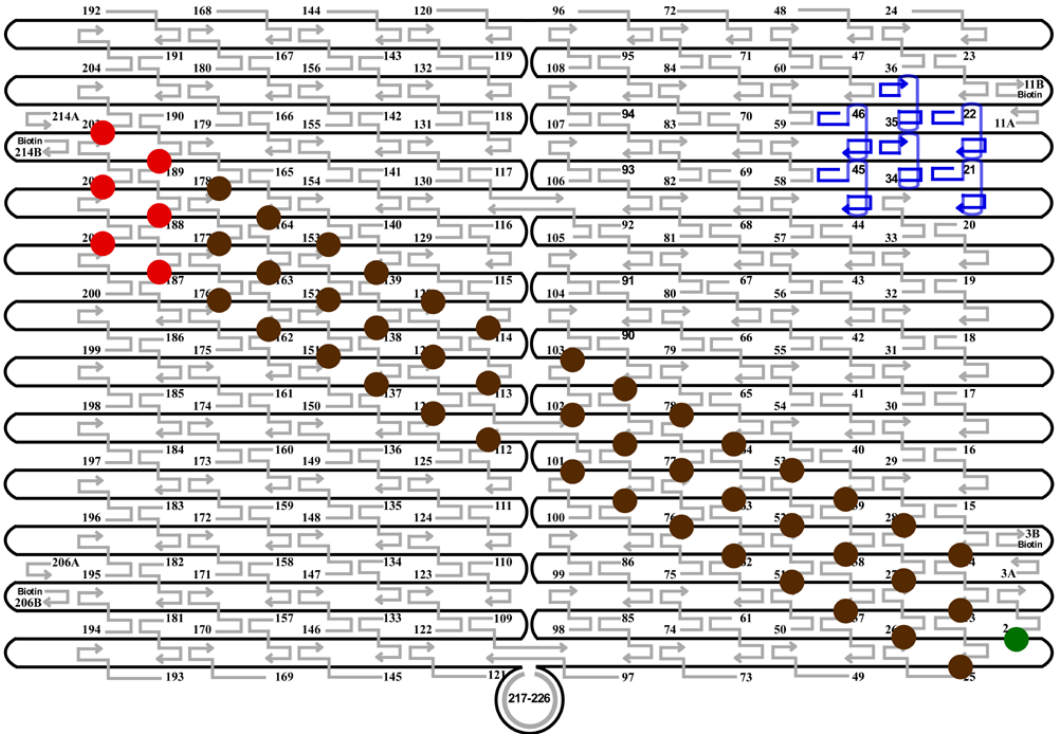
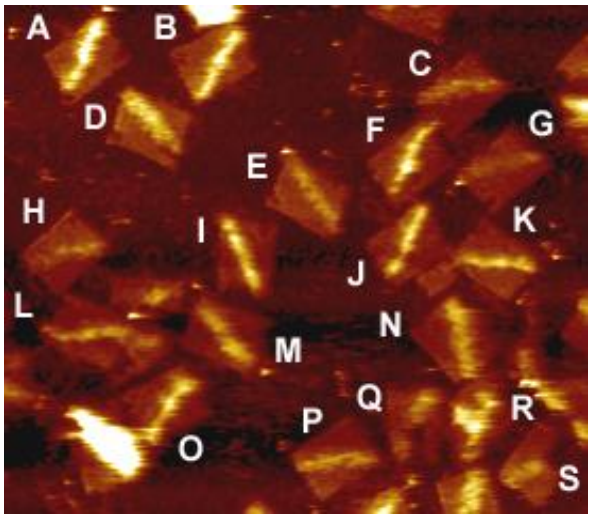
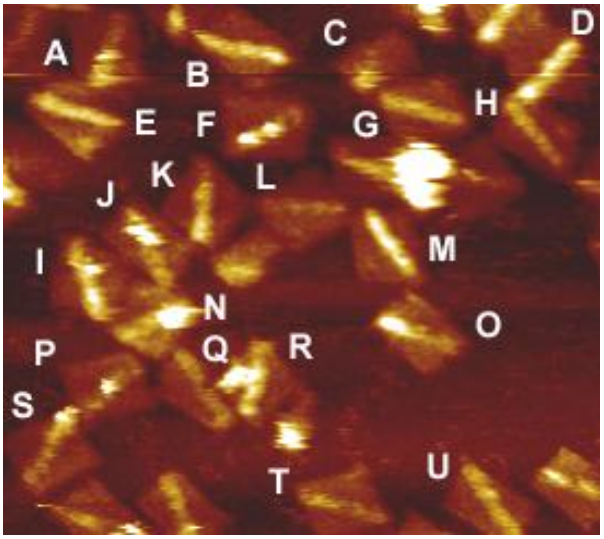


Figure D.18 Wide Field AFM images and classifications used for statistical analysis of EAC design. AFM images of the spider before spider is released.



	up	down	# spiders	START	TRACK	STOP	CONTROL	inscrutable
a		1						0
b		1						0
c	1	0	0	0	0	0	0	0
d	1	0	0	0	0	0	0	0
e	1	0	0	0	0	0	0	0
f		1				0	0	0
g	1	0	1	1	0	0	0	0
h	1	0	0	0	0	0	0	0
i		1						0
j		1						0
k		1						0
l		1						0
m								1
n	1	0	0	0	0	0	0	0
o		1						0
p	1	0	0	0	0	0	0	0
q								1
r								1
s								1

Figure D.19 Wide Field AFM images and classifications used for statistical analysis of EAC design. AFM images of the spider after release.



	up	down	# spiders	START	TRACK	STOP	CONTROL	inscrutable
a								1
b								1
c								1
d		1						0
e		1						0
f		1						0
g	1		0	0	0	0	0	0
h	1		0	0	0	0	0	0
i								1
j	1		1		1			0
k		1						0
l	1	0	0	0	0	0	0	0
m		1						0
n								1
o		1						0
p		1						0
q	1	0	0	0	0	0	0	0
r								1
s		1						0
t	1	0	0	0	0	0	0	0
u	1	0	0	0	0	0	0	0

Figure D.20 Schematic and AFM images of "face-up" vs. "face-down" arrays. **a**, ABD, **b**, EABD **c**, EABC **d**, EAC. Images are 300 x 300 nm.

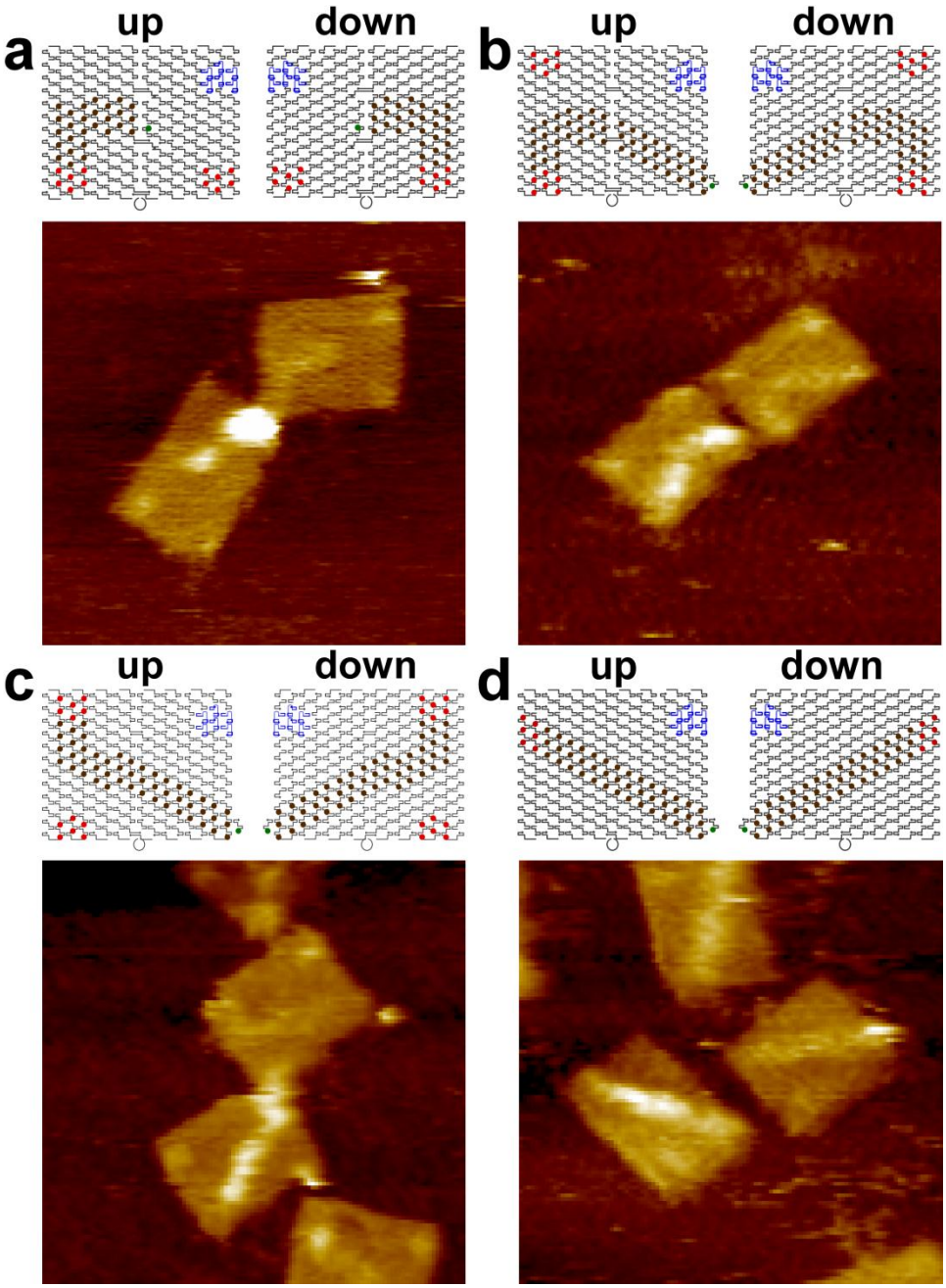


Figure D.21 Schematic and AFM images of "face-up" arrays with no spiders. a, ABD **b**, EABD **c**, EABC **d**, EAC. Images are 300 x 300 nm.

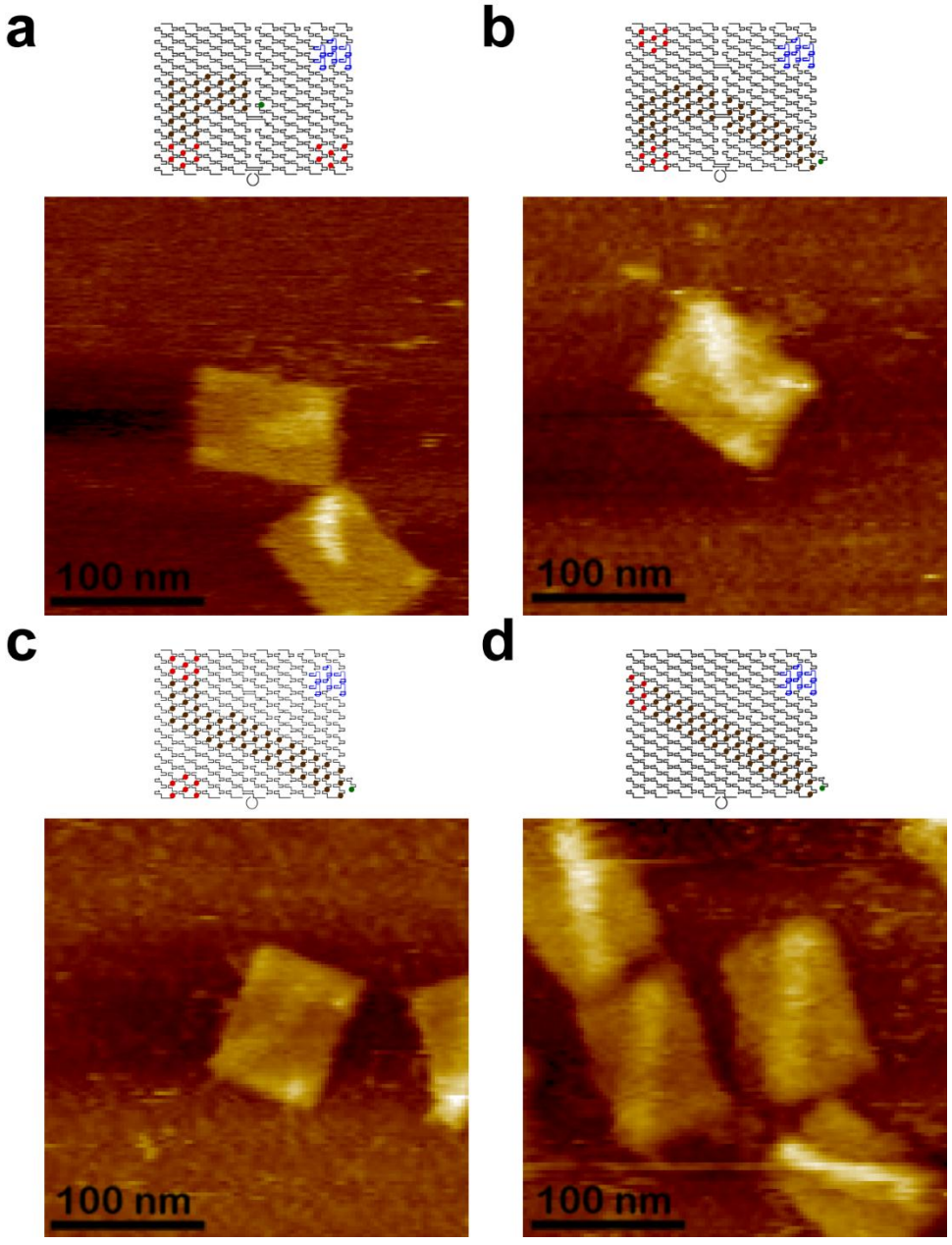


Figure D.22 Schematic and AFM images of false positive (no spider) and multiples. a, Image of "face-up" EABD origami without any spider and **b**, a "face down" EABD from same sample that appears to have a spider near the STOP (even though there was no spider in this sample) **c**, **d**, EABD (**c**) and EAC (**d**) image of two spiders on the array**.** Images are 300 x 300 nm.

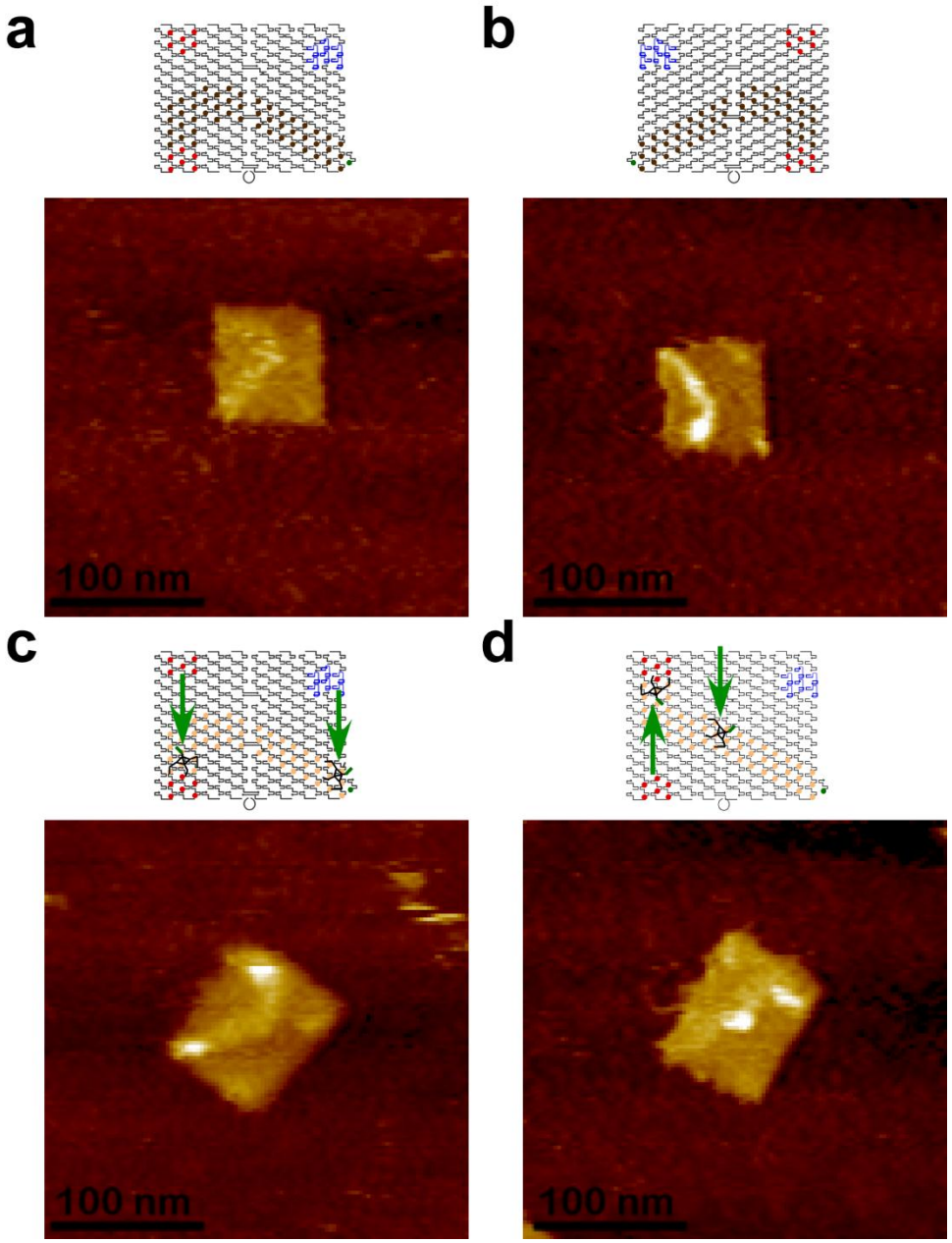


Figure D.23 Schematics, AFM images and Graph of EAC before vs. after. a, Schematics and AFM images of the EAC walk before addition of TRIGGER and 30 min after addition of TRIGGER. **b**, Statistical graph of EAC before vs. after.

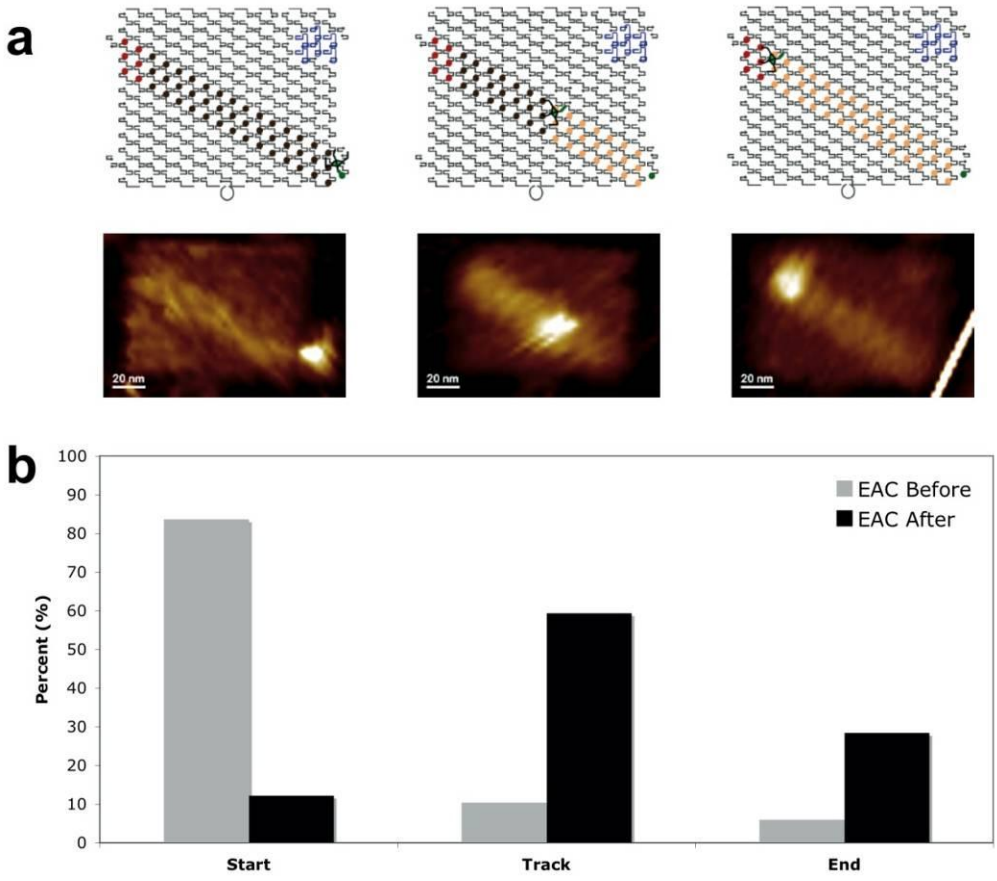
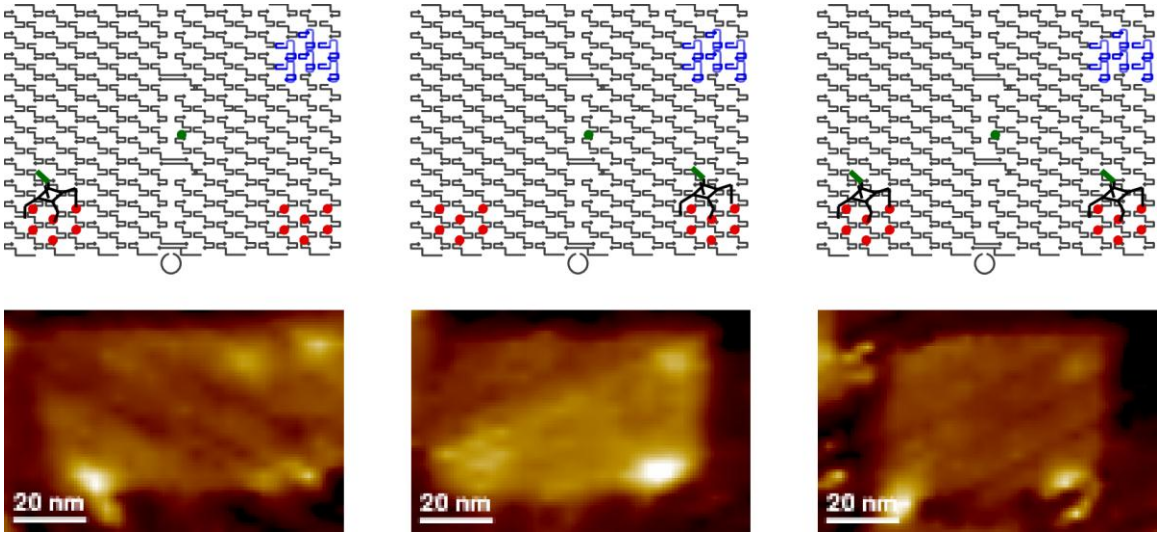


Table D.1 Data and statistics of "face-up" origami arrays. The number of spiders is the total number of spiders found at START, TRACK, STOP and CONTROL sites on singly-occupied origami. Since the 'before' state is the same for the EABD 15 min, 30 min, and 60 min samples, only one set of 'before' statistics was collected. We assume that the 'before' state for the EABD product 60 min sample was also similar.

Track	Before							After						
	Total spiders	START (% of total)	TRACK (% of total)	STOP (% of total)	CONTROL (% of total)	Multiples	None	Total spiders	START (% of total)	TRACK (% of total)	STOP (% of total)	CONTROL (% of total)	Multiples	None
ABD 30 min	31	28 (90.3%)	1 (3.2%)	2 (6.5%)	0 (0%)	0	67	22	1 (4.5%)	4 (18.2%)	16 (72.7%)	1 (4.5%)	0	42
EABD 15 min	113	98 (86.7%)	8 (7.1%)	4 (3.5%)	3 (2.7%)	18	159	106	19 (17.9%)	68 (64.2%)	15 (14.1%)	4 (3.8%)	15	165
EABD 30 min								76	11 (14.5%)	33 (43.4%)	29 (38.2%)	3 (3.9%)	13	135
EABD 60 min								97	11 (11.3%)	23 (23.7%)	60 (61.9%)	3 (3.1%)	10	179
EABD product 60 min								26	4 (15.4%)	9 (34.6%)	12 (46.2%)	1 (3.9%)	2	37
EABC 30 min	98	78 (79.6%)	8 (8.2%)	12 (12.4%)	0 (0%)	14	149	42	2 (4.8%)	20 (47.6%)	20 (47.6%)	0 (0%)	8	70
EAC 30 min	67	56 (83.6%)	7 (10.4%)	4 (6%)	N/A	7	233	74	9 (12.2%)	44 (59.4%)	21 (28.4%)	N/A	8	197

Figure D.24 Schematics and AFM images of spider release control. The spider was released without the TRACK present and allowed to traverse the array for 30 minutes in solution. The images below show the spider at the STOP and CONTROL of this array and an instance where two spiders were seen occupying both positions on one array.

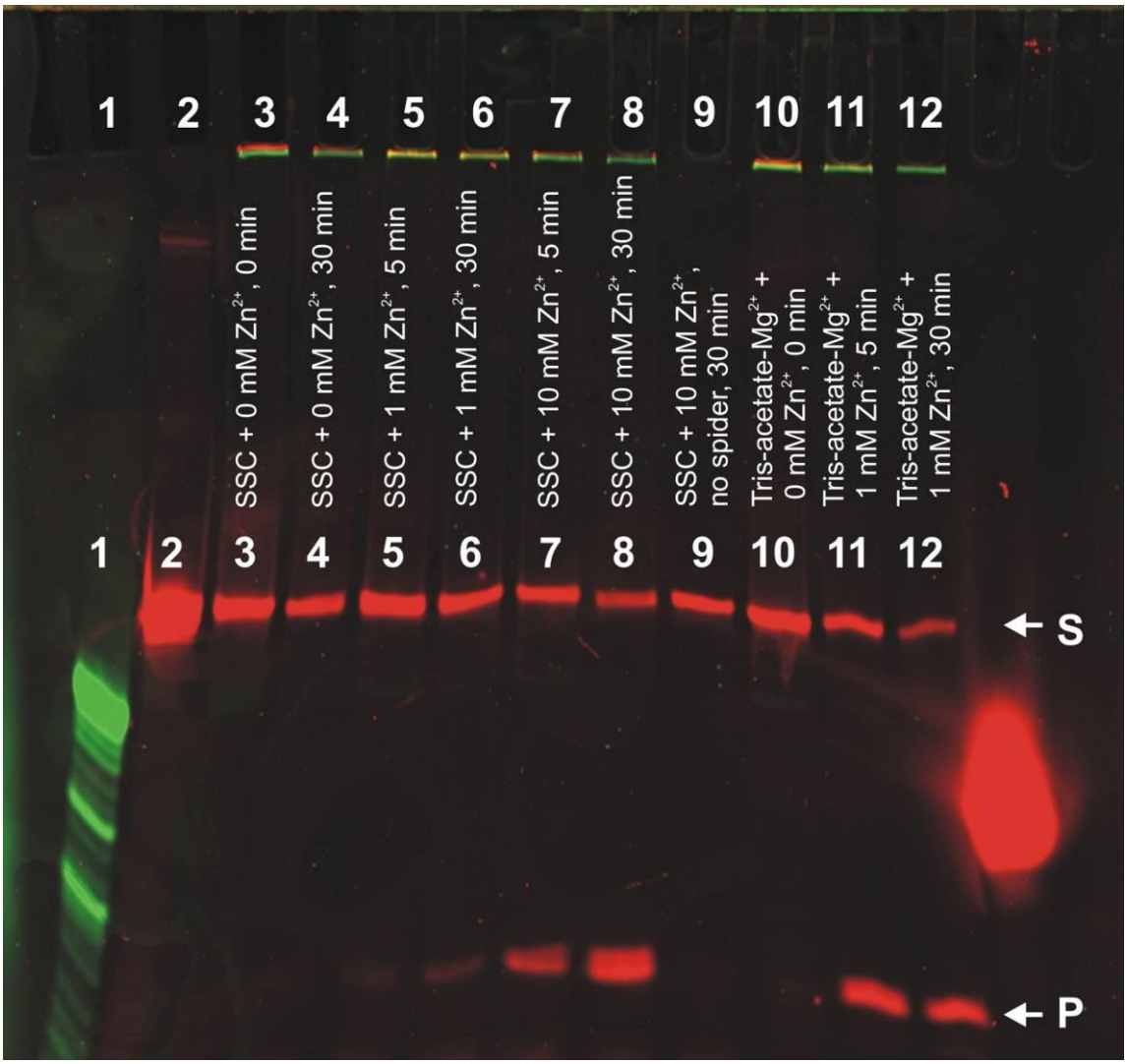


Supplementary Table 2. Data of Spider release without the TRACK

Track	START	STOP	CONTROL
ABD	7	48	45

Figure D.25 PAGE Characterization of Spider Activity in Solution. a, Fluorescence scan of Cy3 and Cy5 in polyacrylamide gel (24% acrylamide). Lane 1 contains an alkali hydrolysis RNA ladder (sequence: 5'-pUGCGUUAGUAGGUUGUAUAGUU-Cy3). Lane 2 contains Cy5-substrate incubated at pH 12 for 5 min at 70°C. Lanes 3-12 contain the products of reactions between spider and Cy5-substrate (S) to form product (P) under the conditions shown in the respective lanes. No cleavage was detected after 30 minutes in absence of either $ZnSO_4$ (lane 4) or spider (lane 9). **b,** Fraction of substrate cleaved versus incubation time in TA-Mg + 1 mM Zn^{2+} (red triangles). SSC + 10 mM Zn^{2+} (blue circles), SSC + 2 mM Zn^{2+} (green diamonds) or SSC + 1 mM Zn^{2+} (black squares). The cleavage assay in SSC + 2 mM Zn^{2+} is not shown in (**a**) but was performed in an identical manner to the other assays in a separate experiment. Each curve is fit to a single exponential decay function.

a



b

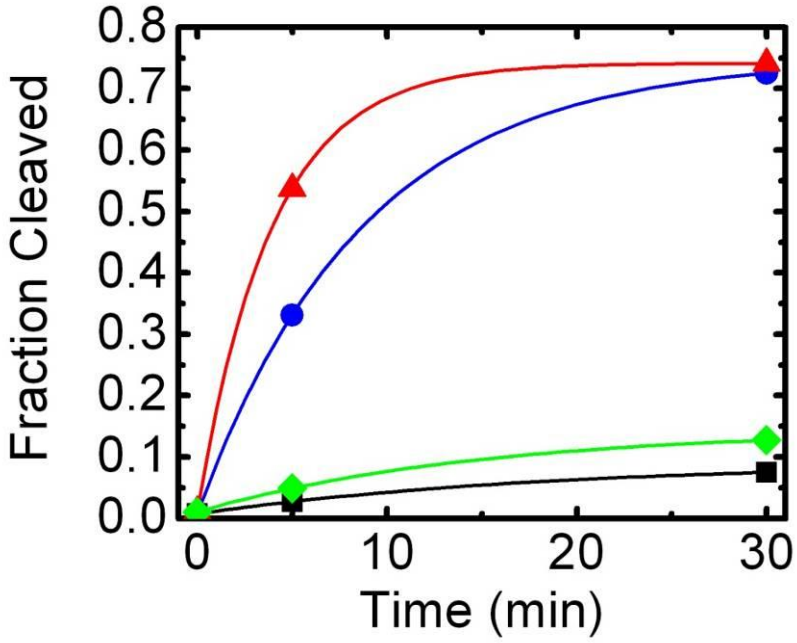


Figure D.26 Preparation of microscope slides. Surface coating of the microscope slide, showing the aminosilane (blue), PDITC (yellow), and covalently bound avidin layers (red).

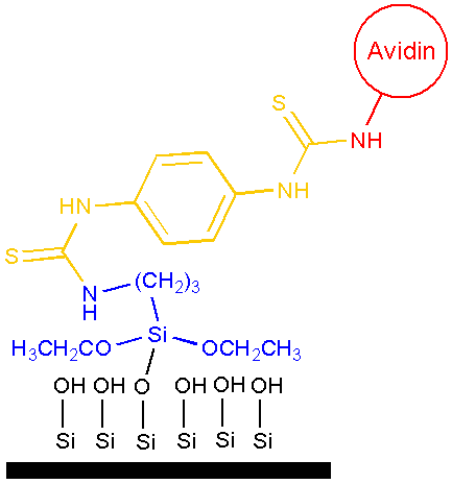


Figure D.27 Schematic representation of the EAC (**a**), EABD (**b**), and EABC (**c**) constructs for fluorescence microscopy. The spider is labeled with 2-3 Cy3 molecules (green) and the STOP with up to 6 Cy5 molecules (red). PSFs from spider-origami pairs are imaged over time and fit, frame-by-frame, to a 2-D Gaussian function; the fit has low residuals (**d**). The coordinates of each PSF exhibit significant drift through time (**e, g**) which is corrected by subtracting the coordinates of Cy5 from its proximal Cy3. The resulting coordinate plots (**f, h**) track the motion of each spider relative to its STOP position. In absence of Zn^{2+} but in presence of TRIGGER in SSC buffer, primarily stationary spiders are observed (**e, f**); the standard deviations σ_x and σ_y give an estimate of precision in position measurements. In contrast, a spider incubated with TRIGGER and zinc in SSC (**g, h**) shows a distinctly biased pattern of motion when subtracted (**h**). The trajectory in (**e**, **f**) corresponds to trajectory 4 in Supplementary Figure D.28d, and the trajectory in (**g**, **h**) corresponds to EAC 2 (Supplementary Figure D.29).

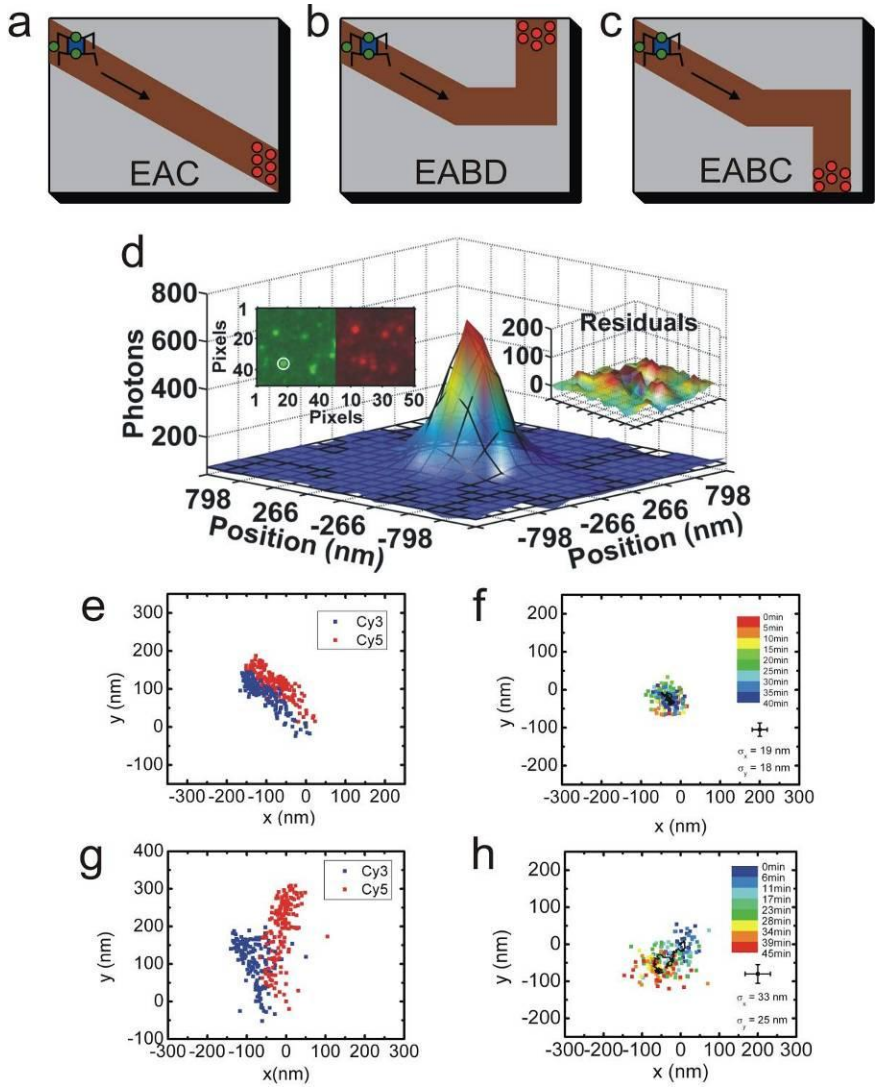
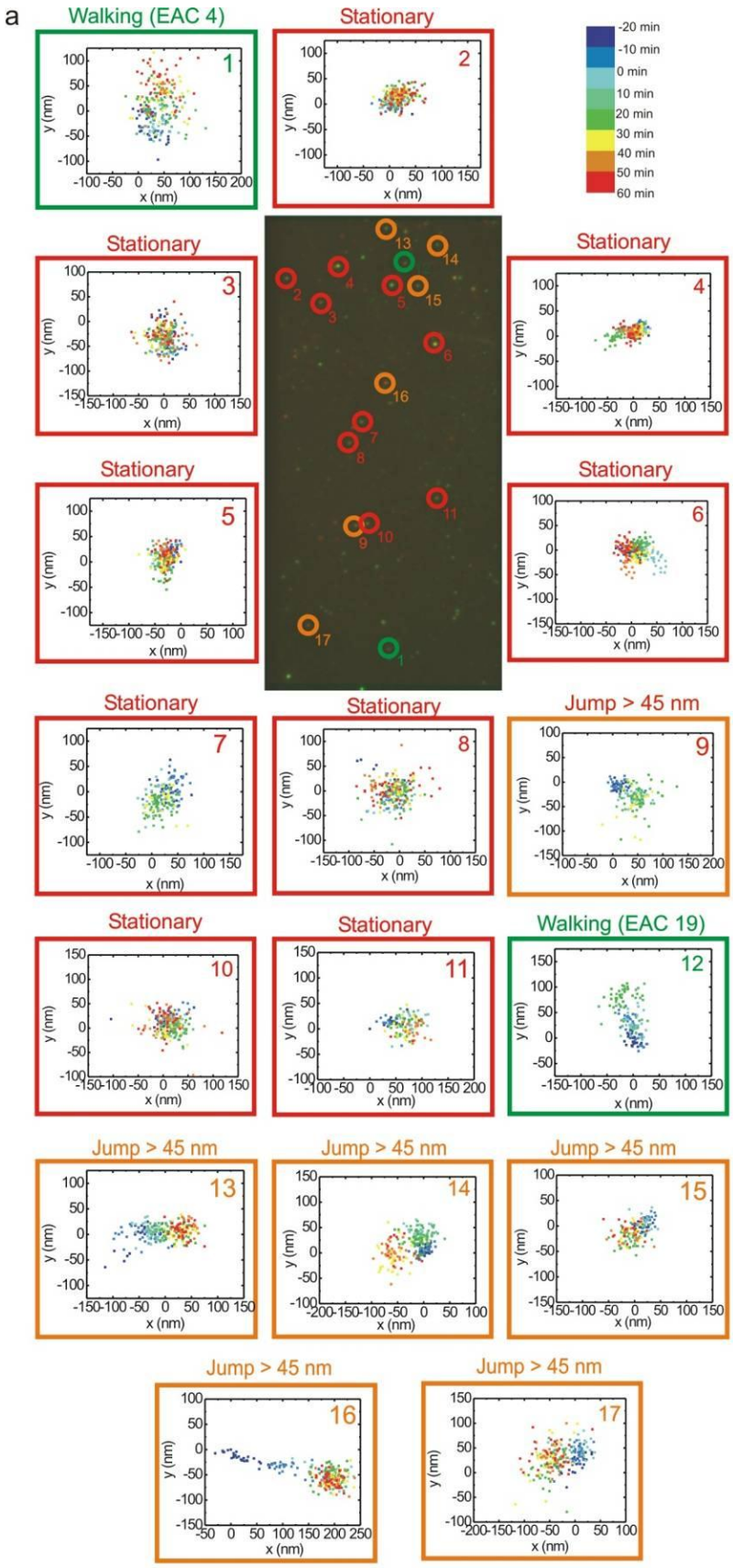
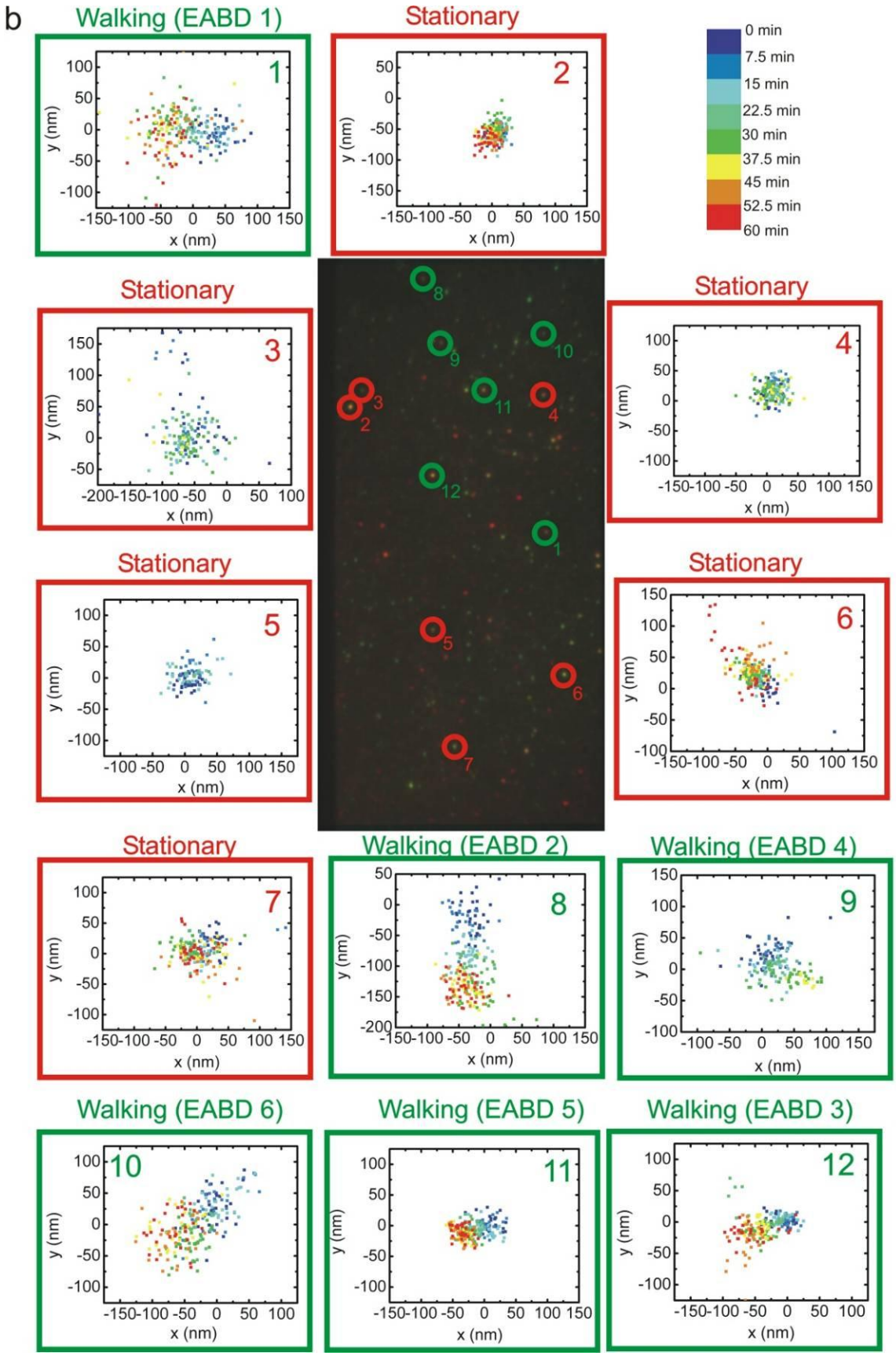
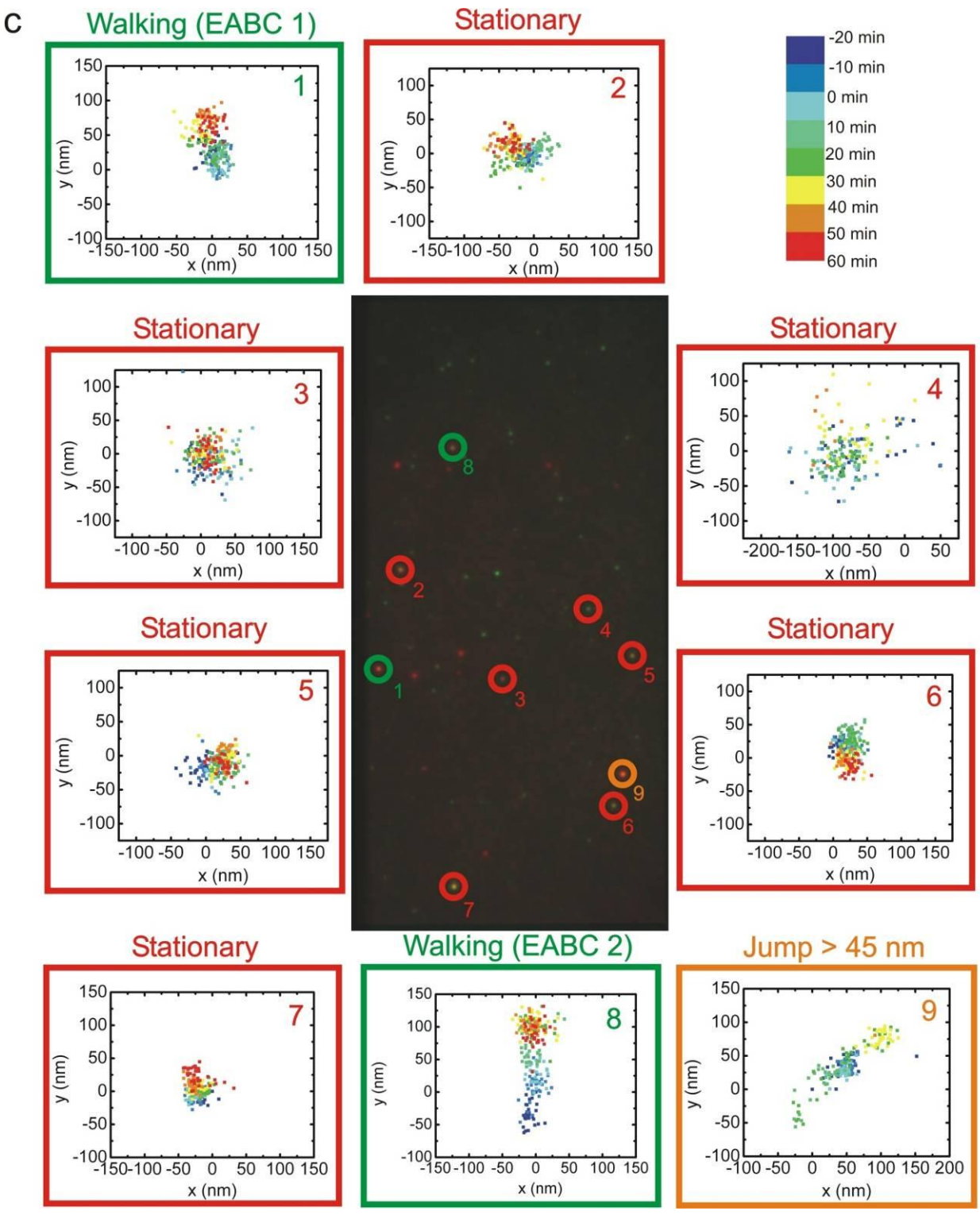


Figure D.28 example CCD camera images from one of each type of experiment: EAC (**a**), EABD (**b**), EABC (**c**), and EAC in the absence of zinc (**d**). Both EAC experiments are performed in SSC buffer, and the EABC and EABD experiments in TA-Mg buffer as described in the text. The numbered circles mark the coordinates of the PSFs whose positions over time were determined using the Gaussian fitting technique described in Supplementary Fig. 27 as displayed in their corresponding numbered trajectory graphs. The trajectory graphs include spiders that walked continuously with a net displacement > 45 nm (green) as determined from the criteria in the text; PSFs that exhibited discontinuous displacement(s) > 45 nm and were thus determined to not be analyzable spiders (orange); and spiders or PSFs that remained stationary or displayed movement ≤ 45 nm (or 2-3 standard deviations, red). The lack of movement in the (-) zinc control (**d**) is consistent with the fact that cleavage activity is dependent on zinc, and supports the notion that movements seen in the experiments with zinc addition are not optical artifacts. Additionally, the presence of many apparently stationary spiders in the (+) zinc experiments (**a-c**) strongly suggests that the motion of adjacent spiders does not result from systematic instrument drift.


a
Walking (EAC 4)
Stationary
-20 min
-10 min
0 min
10 min
20 min
30 min
40 min
50 min
60 min
Stationary
Stationary
Stationary
Stationary
Stationary
Jump > 45 nm
Stationary
Stationary
Walking (EAC 19)
Jump > 45 nm
Jump > 45 nm
Jump > 45 nm
Jump > 45 nm
Jump > 45 nm
y (nm)
x (nm)


b
Walking (EABD 1)
Stationary
0 min
7.5 min
15 min
22.5 min
30 min
37.5 min
45 min
52.5 min
60 min
Stationary
Stationary
Stationary
Stationary
Stationary
Walking (EABD 2)
Walking (EABD 4)
Walking (EABD 6)
Walking (EABD 5)
Walking (EABD 3)
y (nm)
x (nm)


C
Walking (EABC 1)
Stationary
-20 min
-10 min
0 min
10 min
20 min
30 min
40 min
50 min
60 min
Stationary
Stationary
Stationary
Stationary
Stationary
Walking (EABC 2)
Jump > 45 nm
y (nm)
x (nm)

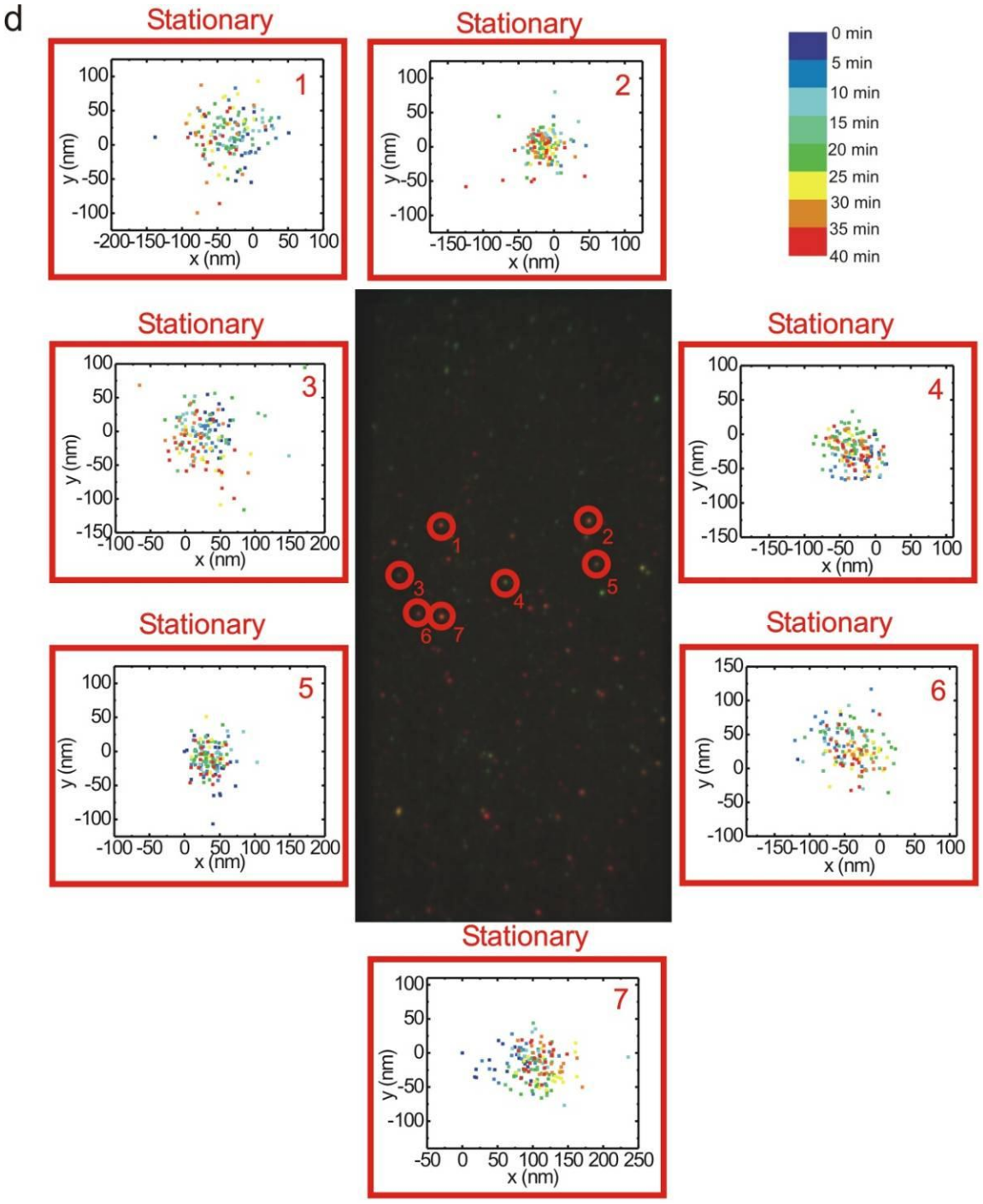
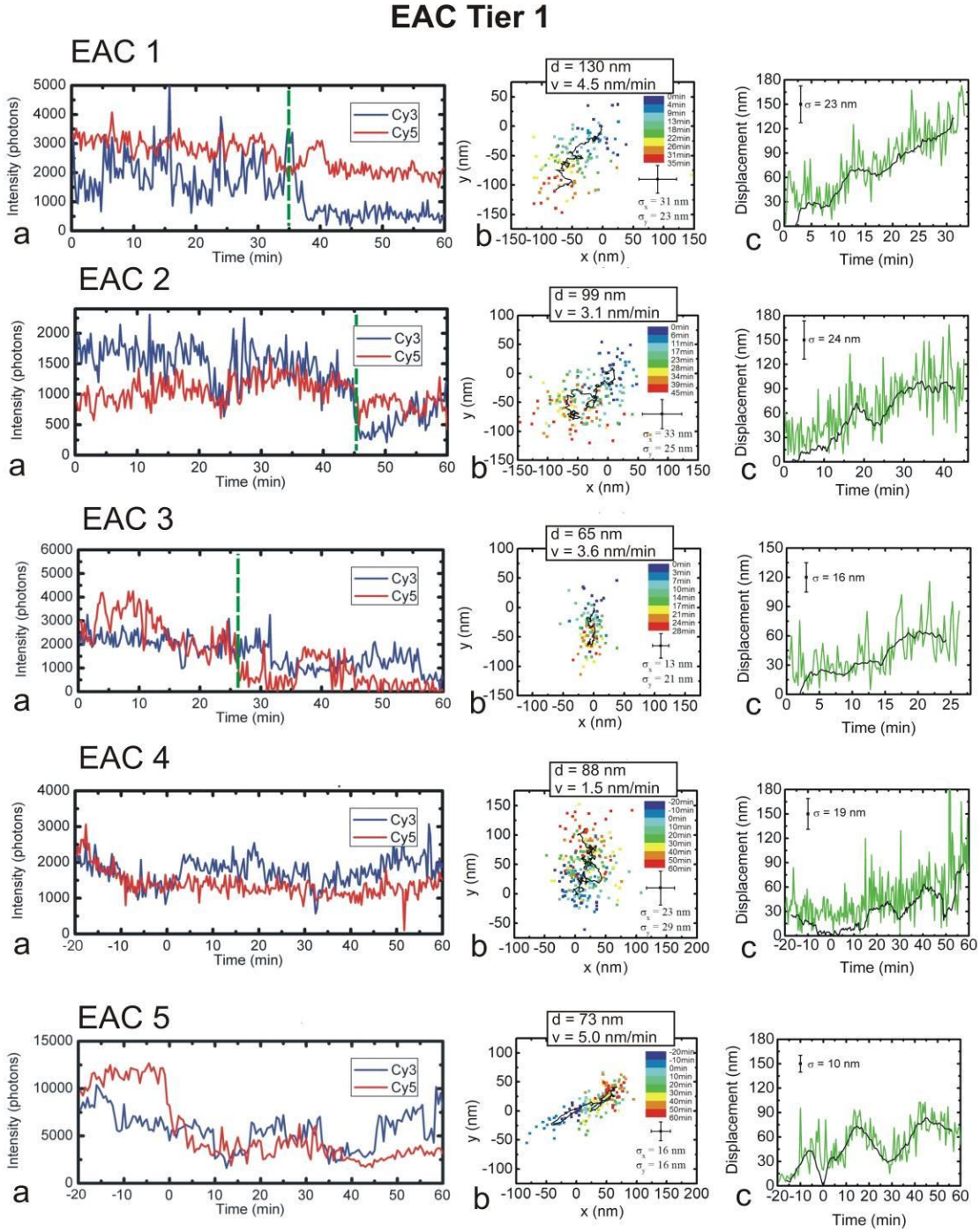
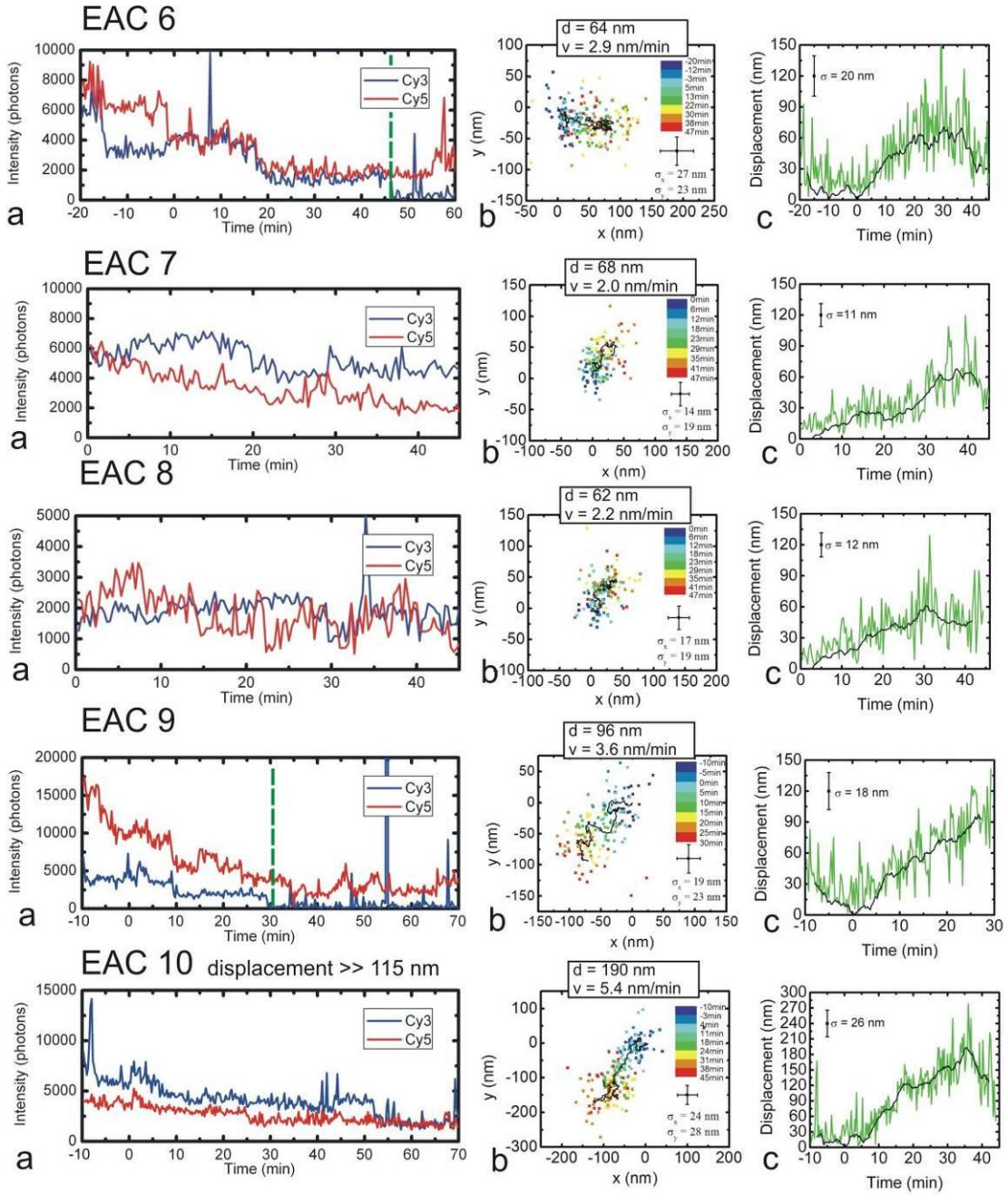

d
Stationary
1
Stationary
2
0 min
5 min
10 min
15 min
20 min
25 min
30 min
35 min
40 min
Stationary
3
Stationary
4
Stationary
5
Stationary
6
Stationary
7
y (nm)
x (nm)

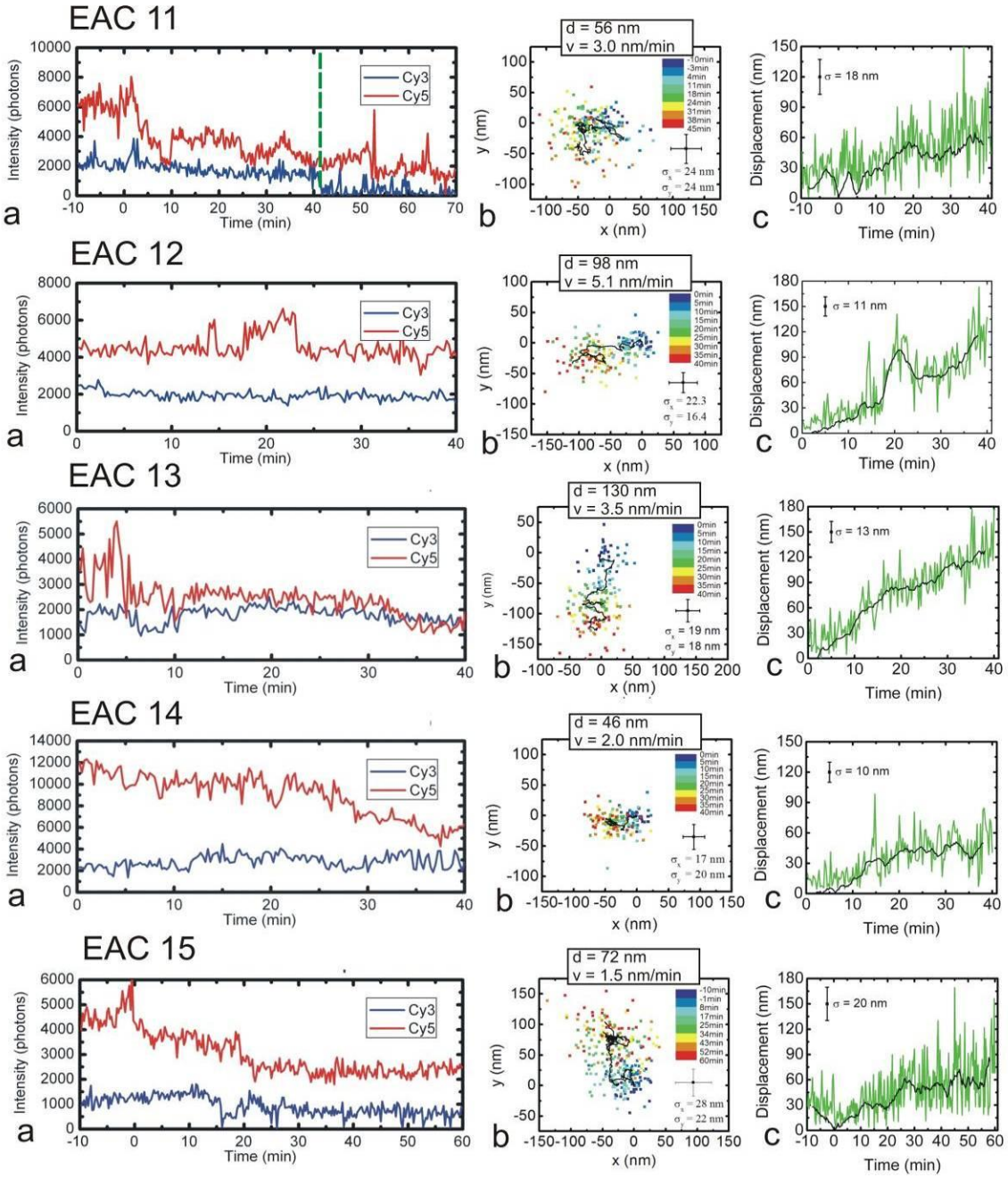
Table D.3 Trajectory filtering statistics for spiders imaged by fluorescence microscopy on substrate tracks. These statistics reflect the filtering of raw fluorescence microscopy data to yield the spider trajectories shown in Supplementary Figs 29 and 30. Total PSF pair candidates (column 1) include PSFs identified as possible signal by an automated image analysis routine. Most of these contain detectible signal from only Cy3, only Cy5, or neither. This is highly variable from experiment to experiment and may depend on how free a particular slide is of fluorescent contaminants. PSF pairs were fit to Gaussians (column 2) if both Cy3 and Cy5 were present for at least 25 minutes with at least 1,000 photon counts per movie frame. PSF pairs with satisfactory fitting (column 3) lacked excessive blinking or interference from other nearby PSFs. Finally, putative moving spiders (column 4) satisfy the selection criteria listed in the supplement text and are also shown in Supplementary Figs 29 and 30.

	Total PSF Pair Candidates	PSF Pairs Fit to Gaussians	PSF Pairs with Satisfactory Fitting	Putative Moving Spiders
EAC SSC + Zn^{2+}	3,821	139	83	21
EAC SSC - Zn^{2+}	303	15	7	0
EAC HBS + Zn^{2+}	384	22	15	11
EAC HBS - Zn^{2+}	127	22	21	3
EABD TA-Mg + Zn^{2+}	477	28	11	6
EABC TA-Mg + Zn^{2+}	227	9	7	2

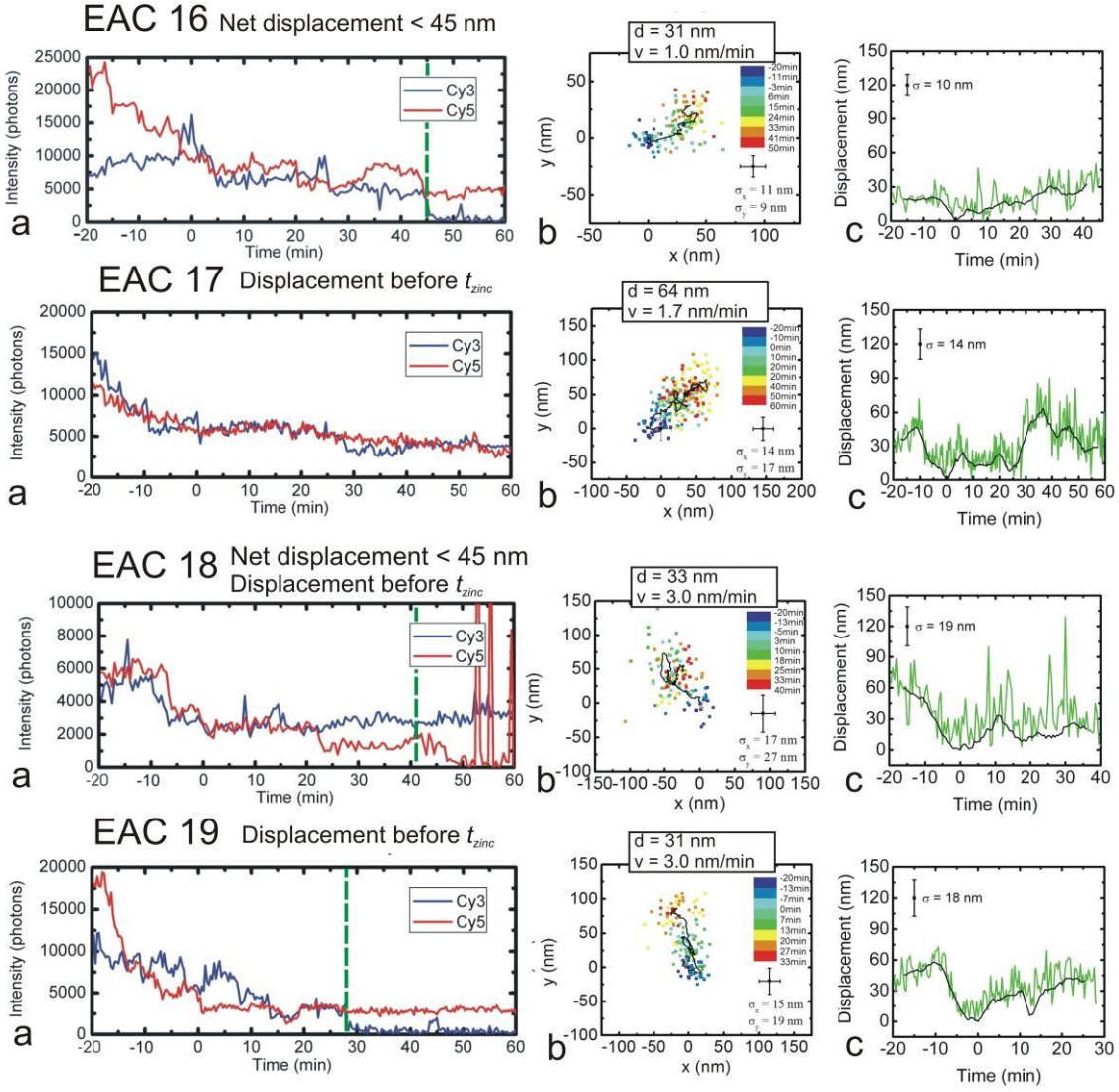
Figure D.29 Fluorophore emission intensity over time (**a**) and (non-averaged) 2-dimensional trajectories of the motion of individual spiders relative to the Cy5 PSF (**b**) for the EAC (EAC 1-19), EABD (EABD 1-6), and EABC (EABC 1-2) substrate tracks as viewed by fluorescence microscopy. The EAC traces in this figure were imaged in SSC buffer, and the EABD and EABC traces in TA-Mg buffer as described in the text. EAC traces are divided into "Tier 1" (EAC 1-15) and "Tier 2" (EAC 16-19), traces in the latter group having a lower probability of representing single walking spiders than the former due to reasons state above each trace and discussed in the text. In (**a**), the dashed green line represents the point after which the trace is no longer analyzed due to photobleaching. The black line in (**b**) represents the smoothed trajectory obtained by applying a 16-frame rolling average as described in the text. The position of the origin is arbitrarily chosen as the start of the trajectory. Plots of displacement versus time for the raw trajectory (**c,** green line) and smoothed trajectory (**c,** black line) are also shown for each trace. The addition of 1 mM $ZnSO_4$ (1-10 mM $ZnSO_4$ for EAC traces) occurred at t = 0 min. The intensity traces show the number of photons collected from each Cy3 or Cy5 point spread function over time, with sharp drops in photon count upon photobleaching of individual fluorophores. Only those time intervals with adequate tracking precision for both fluorophores – generally with more than 1,000 photon counts per frame – are shown in the 2-D trajectories (**b**), and it is these intervals which were analyzed to produce Fig. 4. In the 2-D trajectories, the axes represent spatial dimensions in the fluorescence microscopy image after drift correction. Also shown in panel b are values of net displacement (d) and mean velocity (v) for the EAC track, and mean velocity for the long leg (prior to the 90-degree turn) of the EABC and EABD tracks.


EAC Tier 1
EAC 1
Intensity (photons)
Time (min)
Cy3
Cy5
a
d = 130 nm
v = 4.5 nm/min
y (nm)
x (nm)
b
Displacement (nm)
σ = 23 nm
c
EAC 2
d = 99 nm
v = 3.1 nm/min
σ = 24 nm
EAC 3
d = 65 nm
v = 3.6 nm/min
σ = 16 nm
EAC 4
d = 88 nm
v = 1.5 nm/min
σ = 19 nm
EAC 5
d = 73 nm
v = 5.0 nm/min
σ = 10 nm

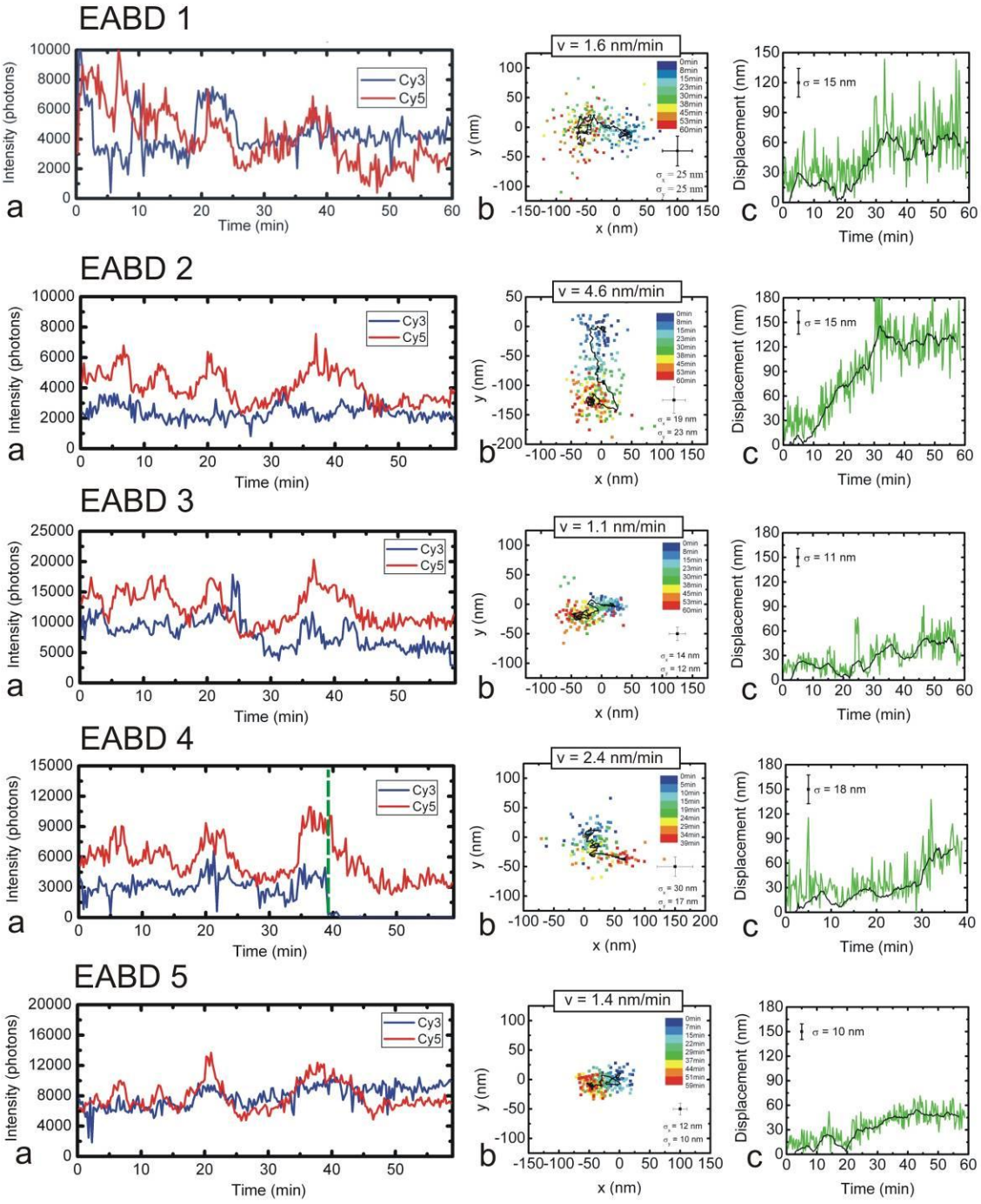

EAC 6
Intensity (photons)
Time (min)
Cy3
Cy5
a
d = 64 nm
v = 2.9 nm/min
y (nm)
x (nm)
b
Displacement (nm)
σ = 20 nm
c
EAC 7
d = 68 nm
v = 2.0 nm/min
σ =11 nm
EAC 8
d = 62 nm
v = 2.2 nm/min
σ = 12 nm
EAC 9
d = 96 nm
v = 3.6 nm/min
σ = 18 nm
EAC 10 displacement >> 115 nm
d = 190 nm
v = 5.4 nm/min
σ = 26 nm


EAC 11
Intensity (photons)
Time (min)
Cy3
Cy5
a
d = 56 nm
v = 3.0 nm/min
y (nm)
x (nm)
b
Displacement (nm)
σ = 18 nm
c
EAC 12
d = 98 nm
v = 5.1 nm/min
σ = 11 nm
EAC 13
d = 130 nm
v = 3.5 nm/min
σ = 13 nm
EAC 14
d = 46 nm
v = 2.0 nm/min
σ = 10 nm
EAC 15
d = 72 nm
v = 1.5 nm/min
σ = 20 nm

EAC Tier 2


EAC 16 Net displacement < 45 nm
d = 31 nm
v = 1.0 nm/min
σ = 10 nm
EAC 17 Displacement before t_{zinc}
d = 64 nm
v = 1.7 nm/min
σ = 14 nm
EAC 18 Net displacement < 45 nm
Displacement before t_{zinc}
d = 33 nm
v = 3.0 nm/min
σ = 19 nm
EAC 19 Displacement before t_{zinc}
d = 31 nm
v = 3.0 nm/min
σ = 18 nm
Cy3
Cy5
Intensity (photons)
Time (min)
x (nm)
y (nm)
Displacement (nm)
a
b
c

EABD


EABD 1
Intensity (photons)
Cy3
Cy5
a
Time (min)
v = 1.6 nm/min
y (nm)
x (nm)
b
Displacement (nm)
σ = 15 nm
c
Time (min)
EABD 2
Intensity (photons)
Cy3
Cy5
a
Time (min)
v = 4.6 nm/min
y (nm)
x (nm)
b
Displacement (nm)
σ = 15 nm
c
Time (min)
EABD 3
Intensity (photons)
Cy3
Cy5
a
Time (min)
v = 1.1 nm/min
y (nm)
x (nm)
b
Displacement (nm)
σ = 11 nm
c
Time (min)
EABD 4
Intensity (photons)
Cy3
Cy5
a
Time (min)
v = 2.4 nm/min
y (nm)
x (nm)
b
Displacement (nm)
σ = 18 nm
c
Time (min)
EABD 5
Intensity (photons)
Cy3
Cy5
a
Time (min)
v = 1.4 nm/min
y (nm)
x (nm)
b
Displacement (nm)
σ = 10 nm
c
Time (min)

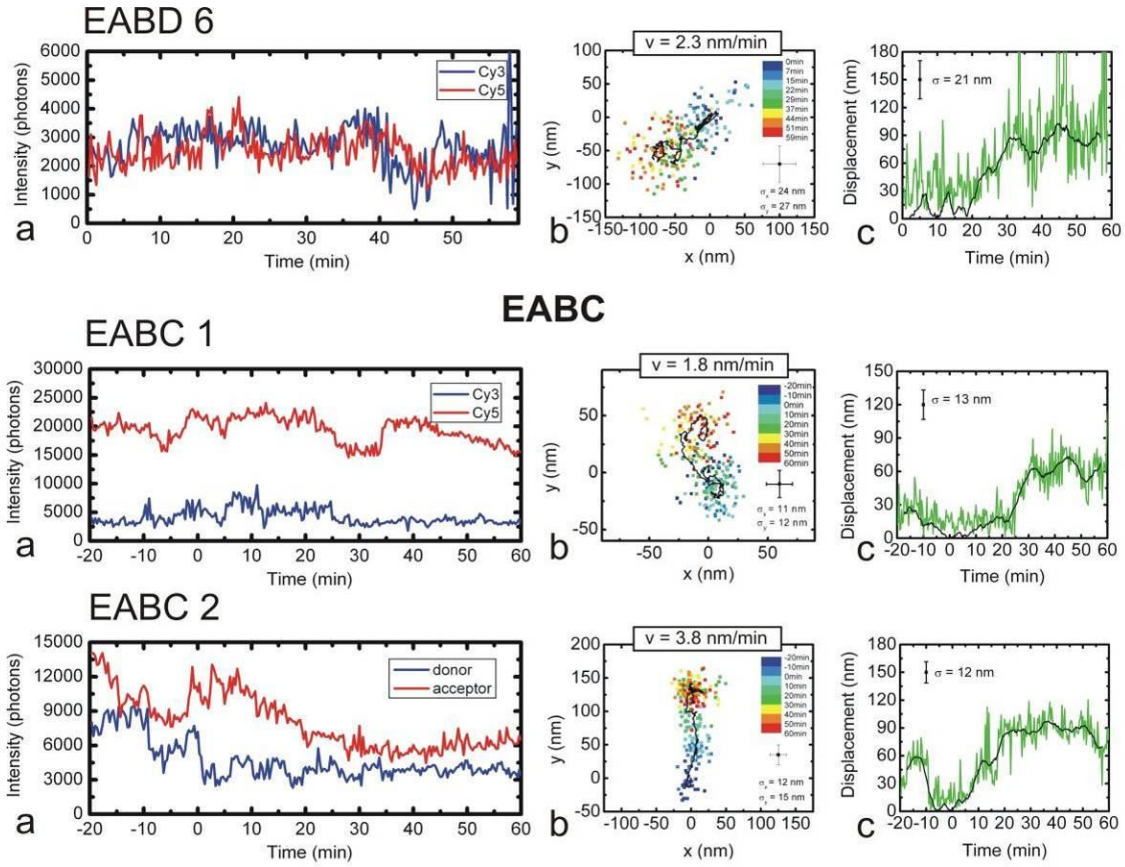
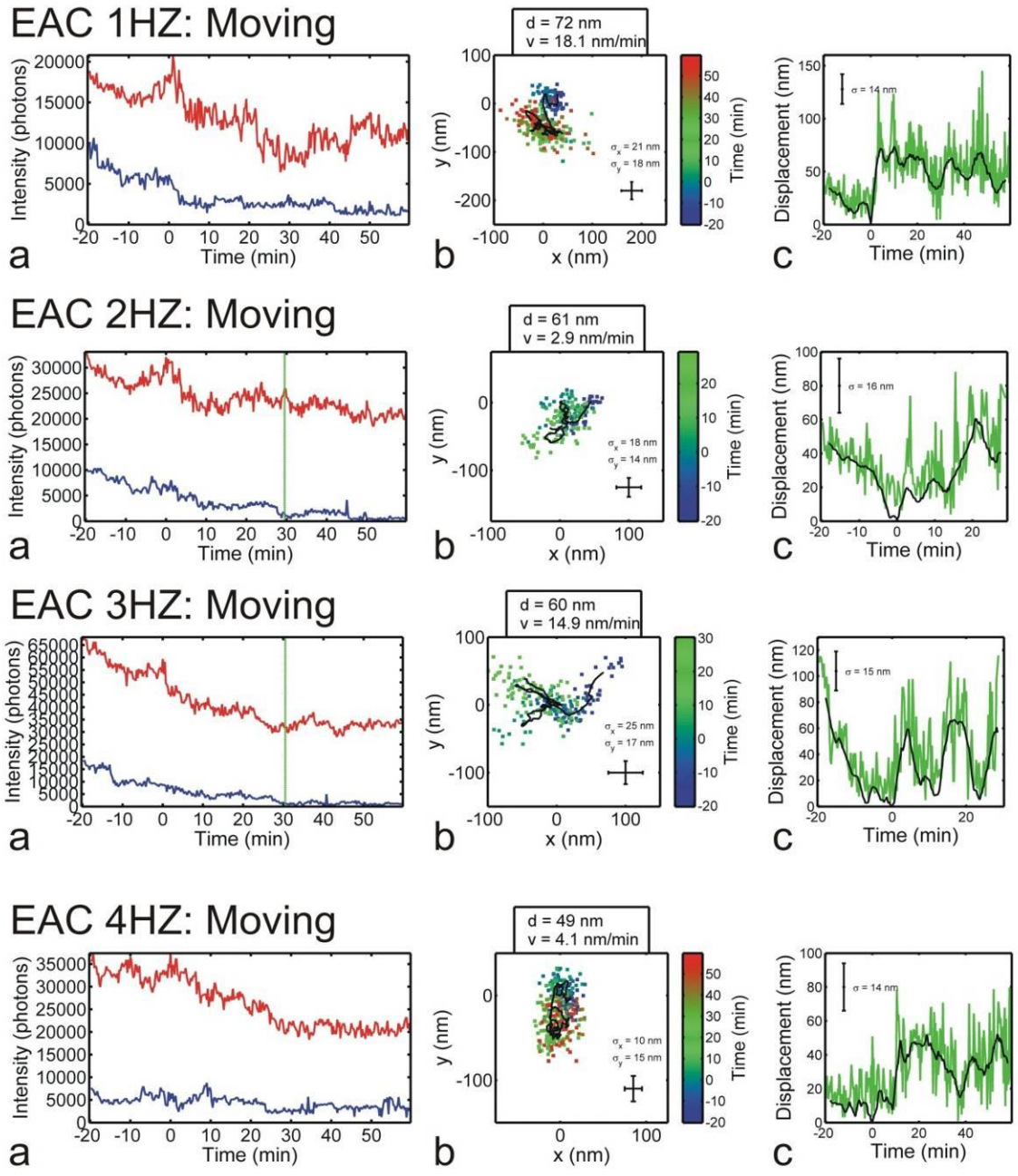
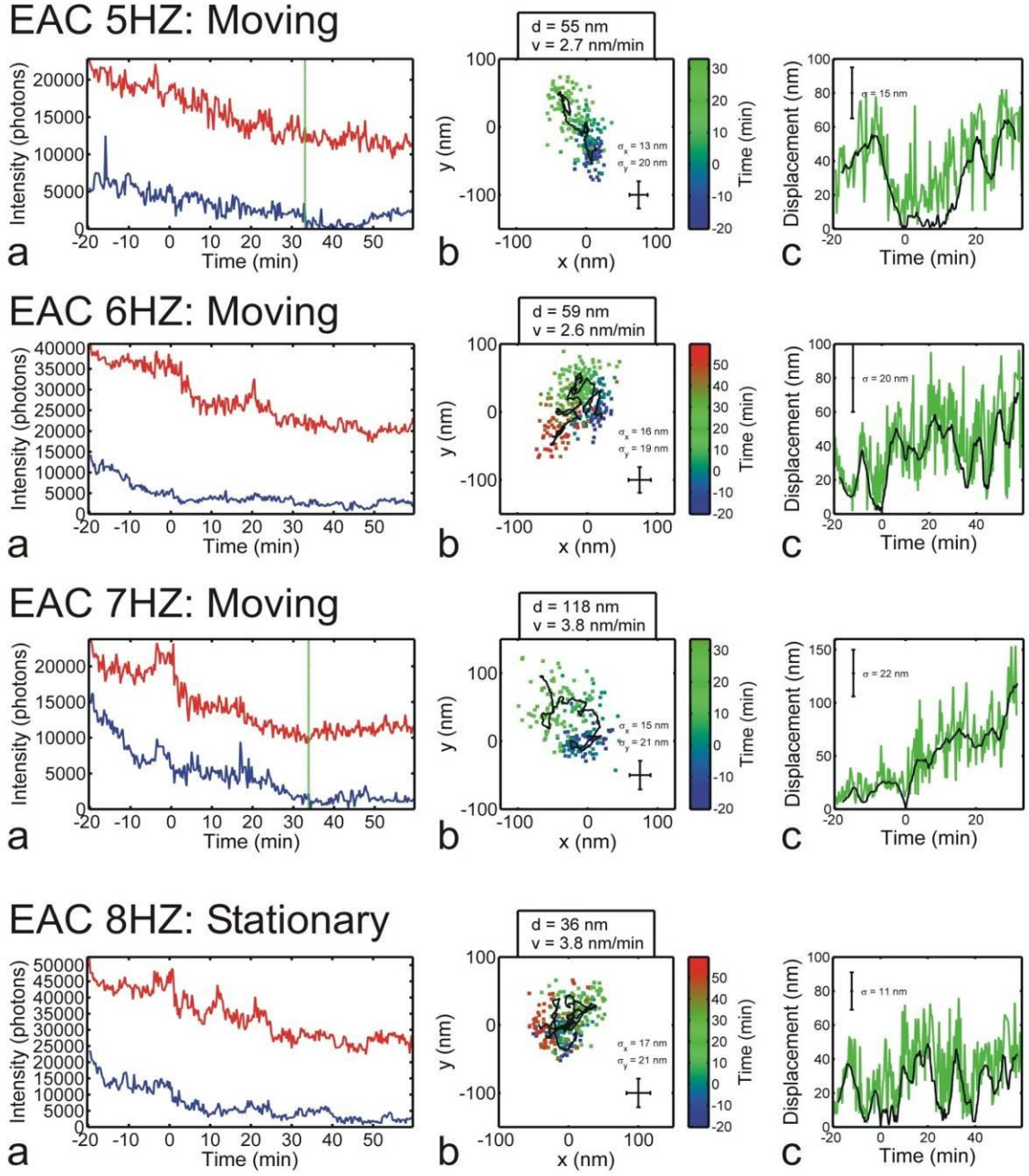
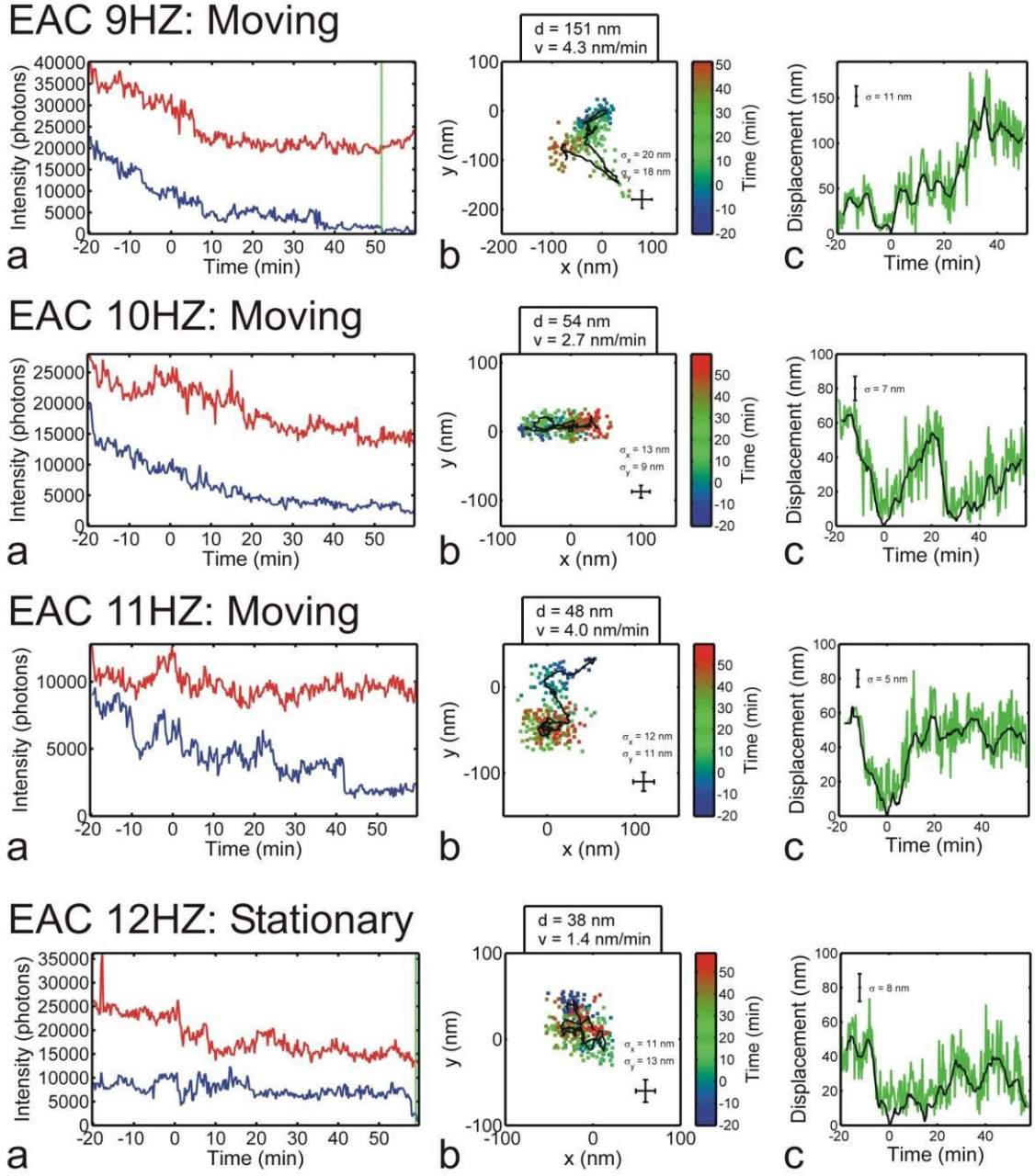

EABD 6
Intensity (photons)
Cy3
Cy5
Time (min)
a
v = 2.3 nm/min
y (nm)
x (nm)
b
Displacement (nm)
σ = 21 nm
Time (min)
c
EABC
EABC 1
Intensity (photons)
Cy3
Cy5
Time (min)
a
v = 1.8 nm/min
y (nm)
x (nm)
b
Displacement (nm)
σ = 13 nm
Time (min)
c
EABC 2
Intensity (photons)
donor
acceptor
Time (min)
a
v = 3.8 nm/min
y (nm)
x (nm)
b
Displacement (nm)
σ = 12 nm
Time (min)
c

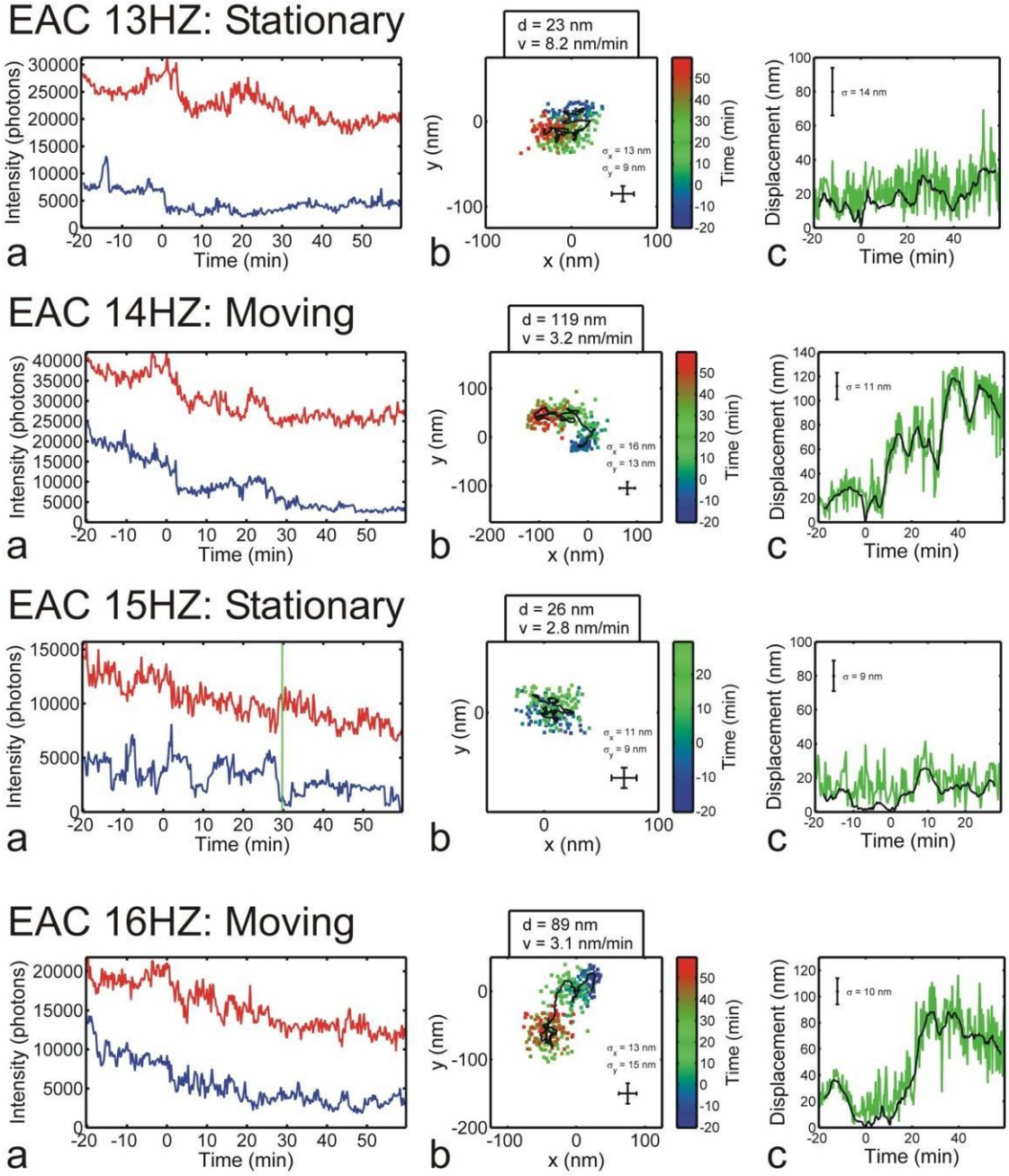
Figure D.30 Fluorophore emission intensity over time (**a**), (non-averaged) 2-dimensional trajectories of the motion of individual spiders relative to the Cy5 PSF (**b**), and displacement versus time plots of individual spiders (**c**) on the EAC substrate track as imaged by fluorescence microscopy in 1× HBS buffer with 1× HBS buffer containing 5 mM (EAC HZ 1-16) or 0 mM $ZnSO_4$ (EAC H 1-21) added after 20 minutes of imaging. In (**a**), the vertical green line represents the point after which the trace is no longer analyzed due to photobleaching. The black line in (**b**) represents the smoothed trajectory obtained by applying a 16-frame rolling average as described in the text. The origin is chosen to coincide with each spider's coordinates at the time of adding 1× HBS buffer containing 0 or 5 mM $ZnSO_4$. Plots of displacement versus time for the raw trajectory (**c**, green line) and smoothed trajectory (**c,** black line) are also shown for each trace. The addition of 5 mM $ZnSO_4$ occurred at t = 0 min. Also shown in panel b are values of net displacement (d) and mean velocity (v) calculated as described in the fluorescence microscopy analysis section. For comparison between experiments performed in 0 and 5 mM $ZnSO_4$, all stationary spiders (those with net displacements less than or equal to 45 nm after zinc addition) observed under each set of conditions are also shown. In the presence of 5 mM $ZnSO_4$, 12 of 16 trajectories move > 45 nm, while only 3 of 21 trajectories collected in absence of Zn^{2+} ions appear to move > 45 nm.

HBS, 5 mM Zn^{2+} (EAC HZ)

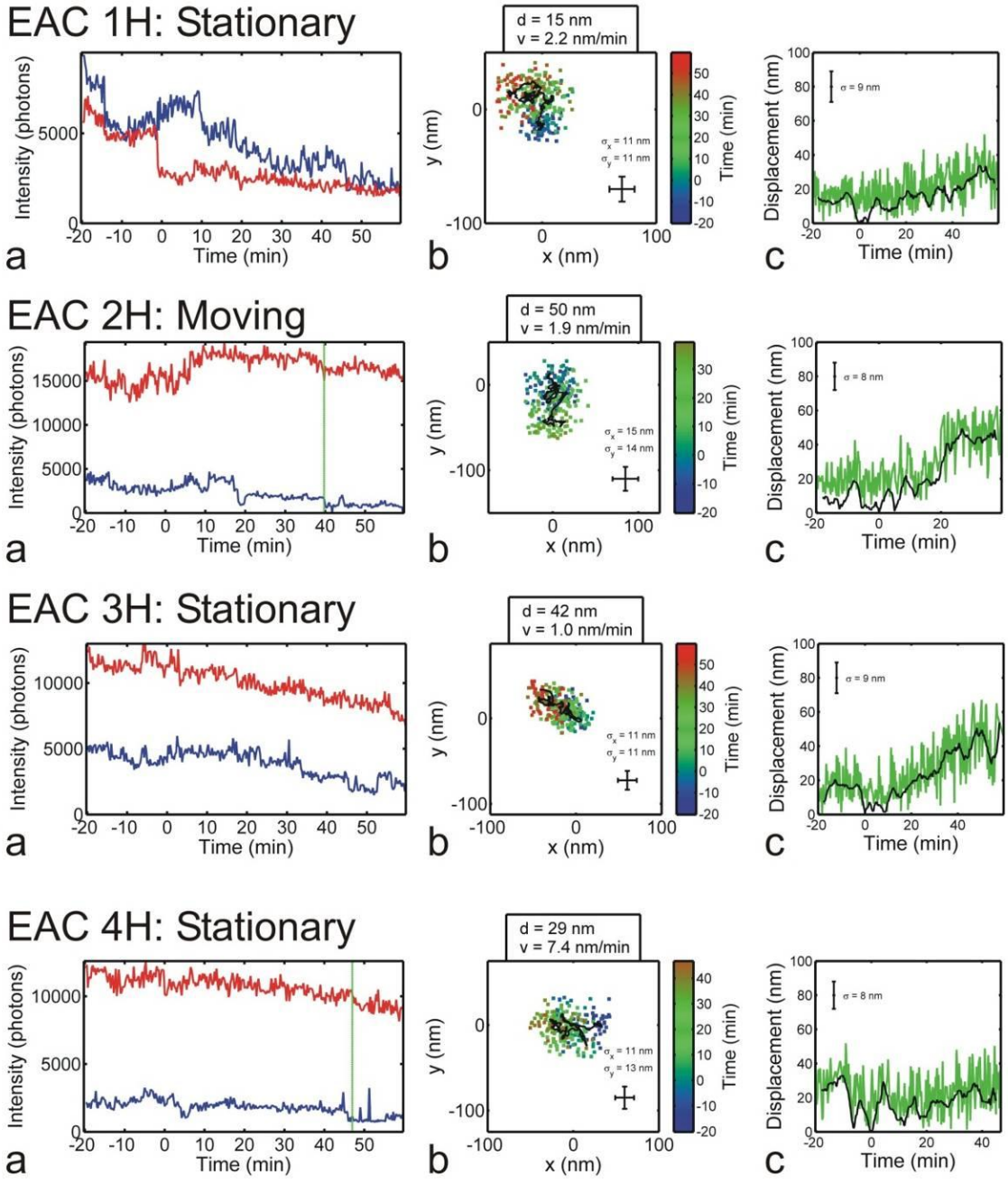


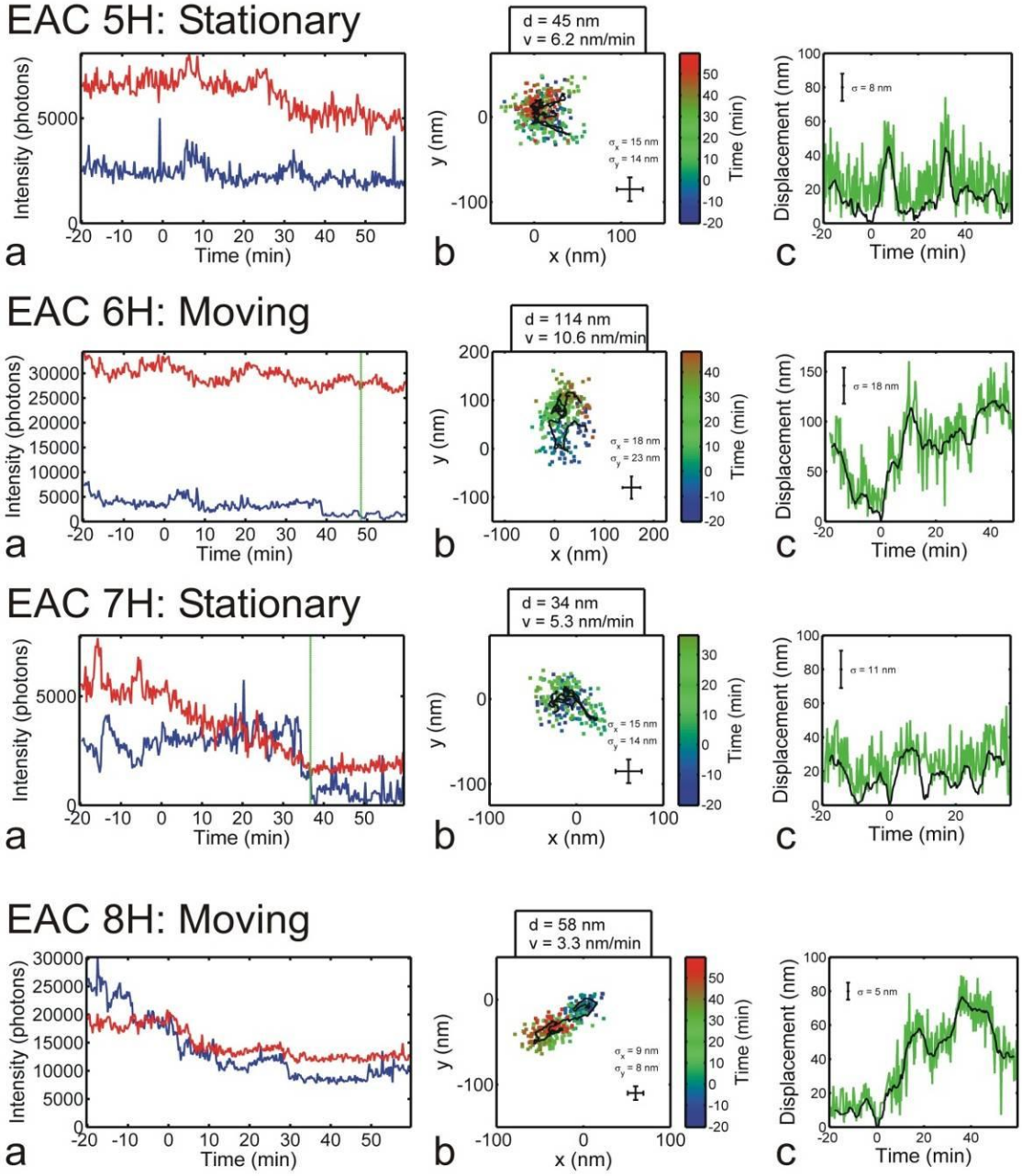

EAC 5HZ: Moving
d = 55 nm
v = 2.7 nm/min
σx = 13 nm
σy = 20 nm
σ = 15 nm
EAC 6HZ: Moving
d = 59 nm
v = 2.6 nm/min
σx = 16 nm
σy = 19 nm
σ = 20 nm
EAC 7HZ: Moving
d = 118 nm
v = 3.8 nm/min
σx = 15 nm
σy = 21 nm
σ = 22 nm
EAC 8HZ: Stationary
d = 36 nm
v = 3.8 nm/min
σx = 17 nm
σy = 21 nm
σ = 11 nm
Intensity (photons)
Time (min)
y (nm)
x (nm)
Time (min)
Displacement (nm)
Time (min)
a
b
c

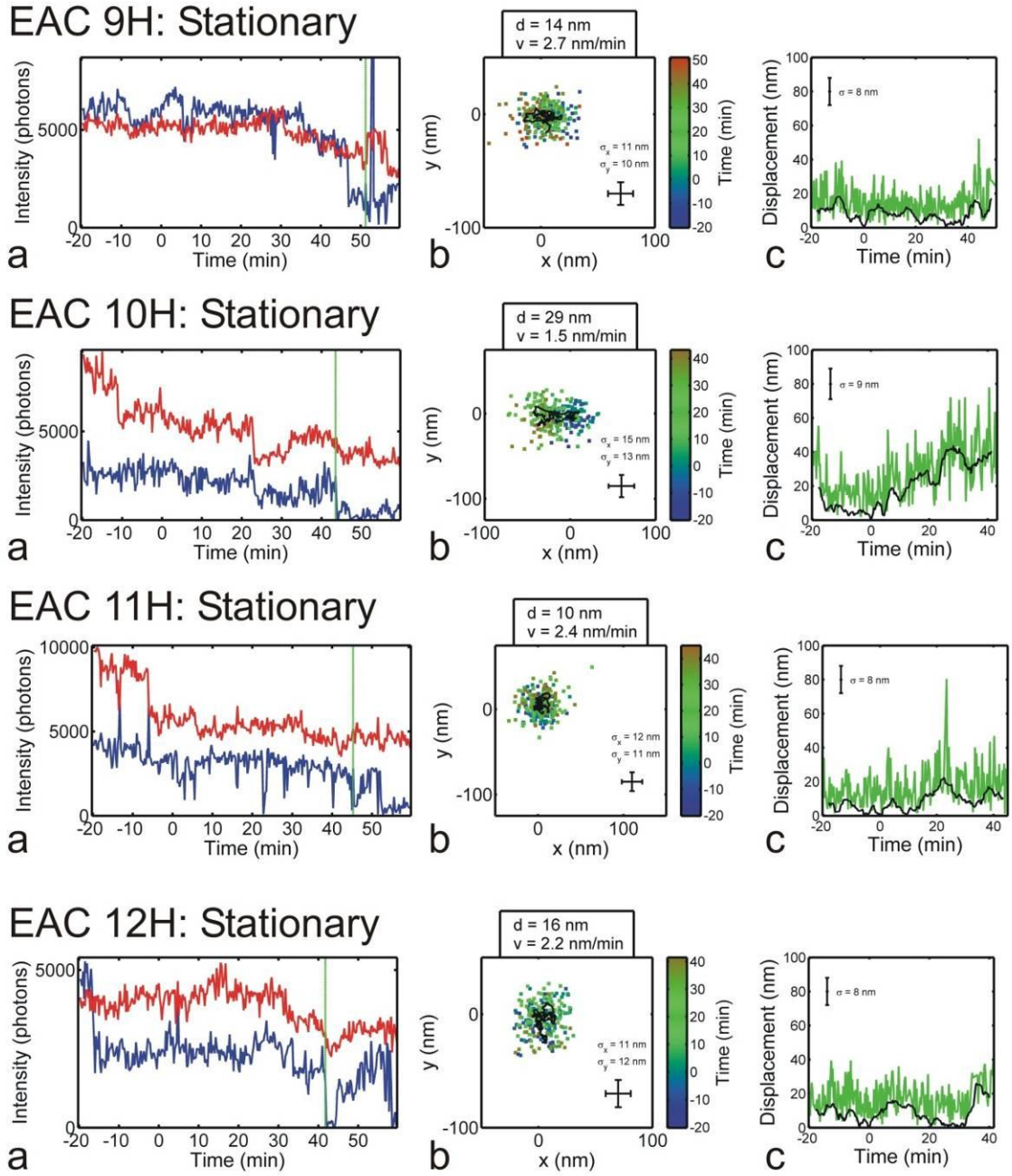

EAC 9HZ: Moving
Intensity (photons)
Time (min)
a
d = 151 nm
v = 4.3 nm/min
σx = 20 nm
σy = 18 nm
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
σ = 11 nm
Time (min)
c
EAC 10HZ: Moving
Intensity (photons)
Time (min)
a
d = 54 nm
v = 2.7 nm/min
σx = 13 nm
σy = 9 nm
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
σ = 7 nm
Time (min)
c
EAC 11HZ: Moving
Intensity (photons)
Time (min)
a
d = 48 nm
v = 4.0 nm/min
σx = 12 nm
σy = 11 nm
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
σ = 5 nm
Time (min)
c
EAC 12HZ: Stationary
Intensity (photons)
Time (min)
a
d = 38 nm
v = 1.4 nm/min
σx = 11 nm
σy = 13 nm
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
σ = 8 nm
Time (min)
c

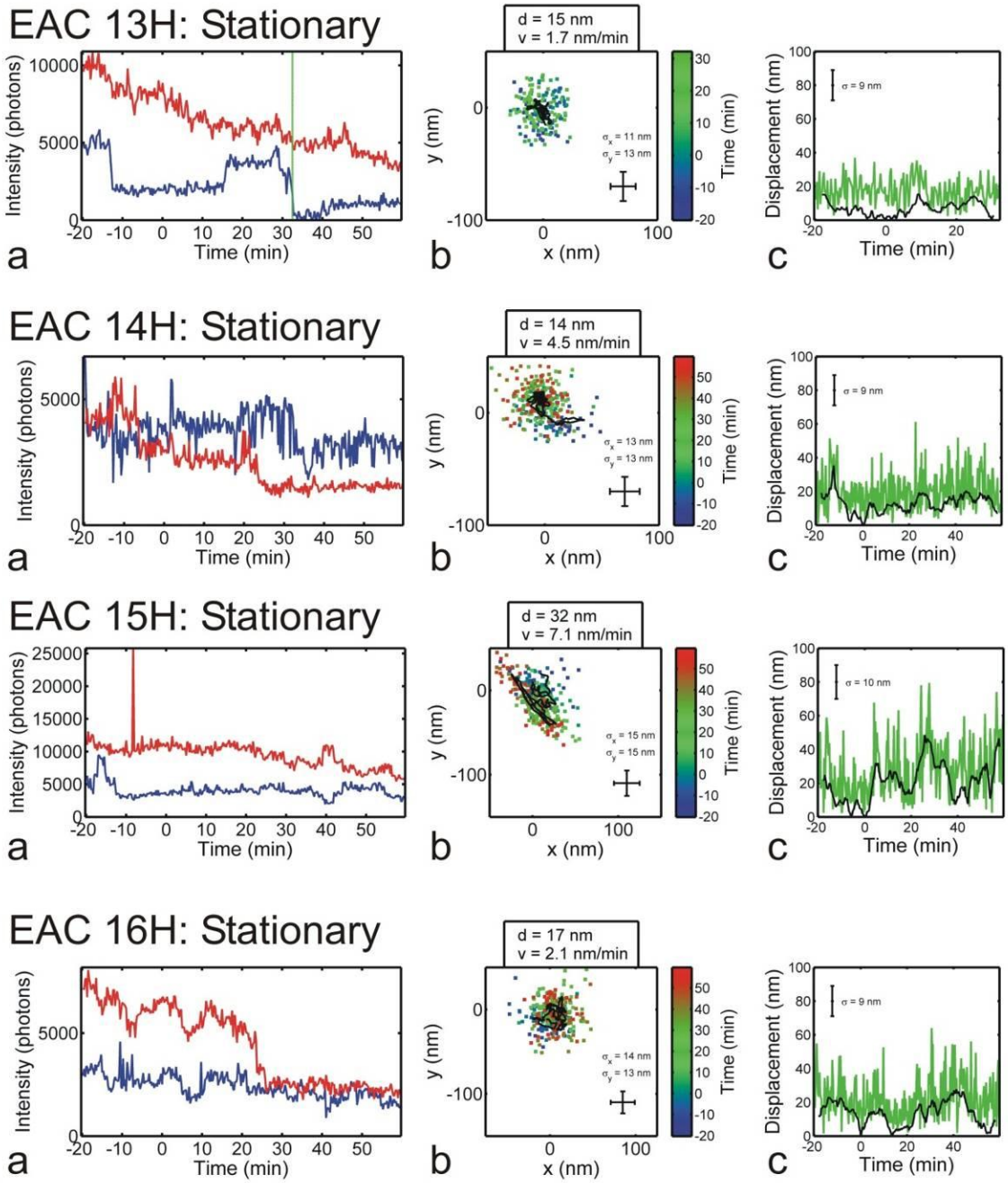

EAC 13HZ: Stationary
Intensity (photons)
Time (min)
a
d = 23 nm
v = 8.2 nm/min
σx = 13 nm
σy = 9 nm
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
σ = 14 nm
Time (min)
c
EAC 14HZ: Moving
Intensity (photons)
Time (min)
a
d = 119 nm
v = 3.2 nm/min
σx = 16 nm
σy = 13 nm
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
σ = 11 nm
Time (min)
c
EAC 15HZ: Stationary
Intensity (photons)
Time (min)
a
d = 26 nm
v = 2.8 nm/min
σx = 11 nm
σy = 9 nm
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
σ = 9 nm
Time (min)
c
EAC 16HZ: Moving
Intensity (photons)
Time (min)
a
d = 89 nm
v = 3.1 nm/min
σx = 13 nm
σy = 15 nm
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
σ = 10 nm
Time (min)
c

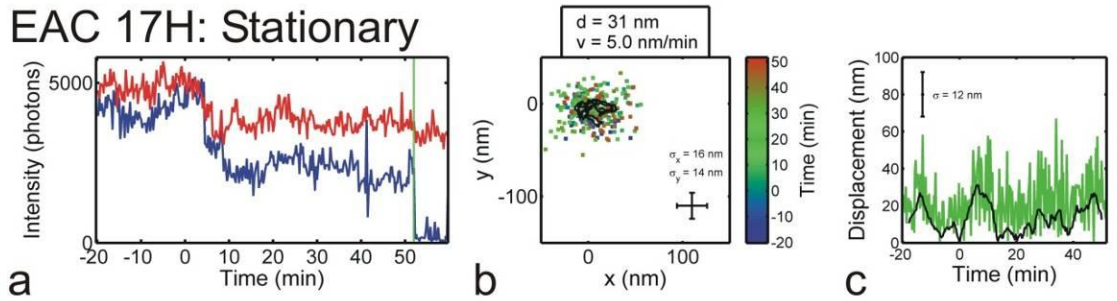
HBS, 0 mM Zn^{2+} Control (EAC H)

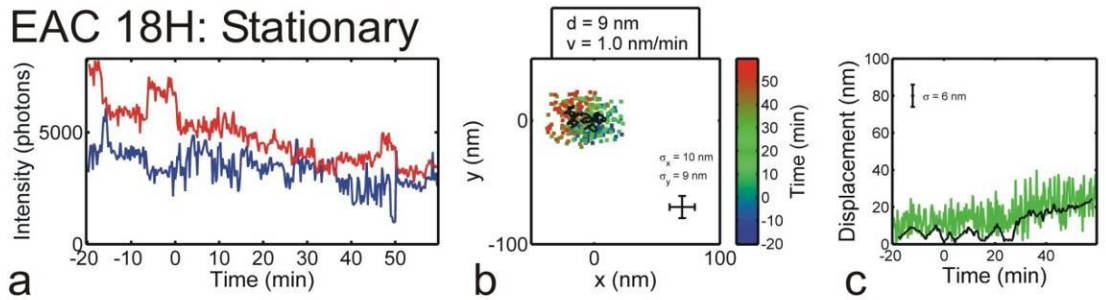


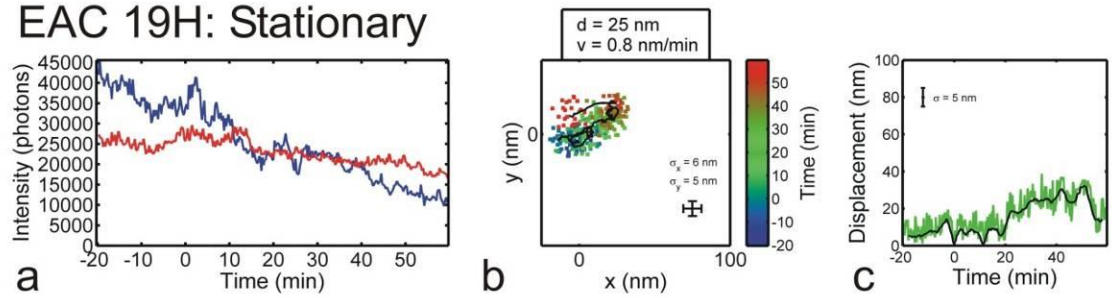

EAC 5H: Stationary
Intensity (photons)
Time (min)
a
d = 45 nm
v = 6.2 nm/min
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
Time (min)
c
EAC 6H: Moving
d = 114 nm
v = 10.6 nm/min
EAC 7H: Stationary
d = 34 nm
v = 5.3 nm/min
EAC 8H: Moving
d = 58 nm
v = 3.3 nm/min

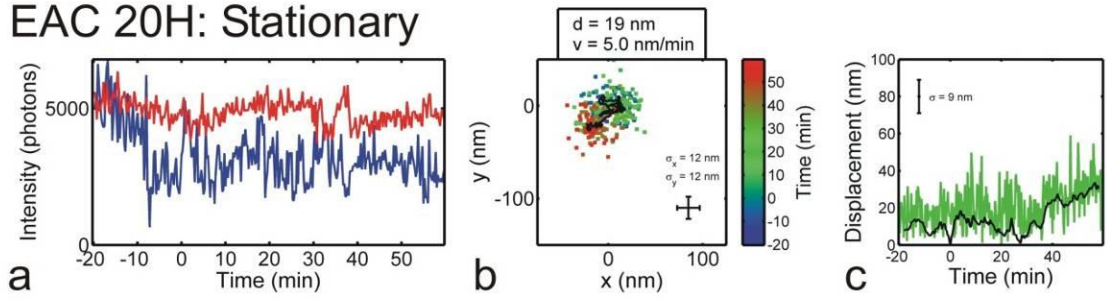

EAC 9H: Stationary
Intensity (photons)
Time (min)
a
d = 14 nm
v = 2.7 nm/min
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
Time (min)
c
EAC 10H: Stationary
d = 29 nm
v = 1.5 nm/min
EAC 11H: Stationary
d = 10 nm
v = 2.4 nm/min
EAC 12H: Stationary
d = 16 nm
v = 2.2 nm/min


EAC 13H: Stationary
Intensity (photons)
Time (min)
a
d = 15 nm
v = 1.7 nm/min
σx = 11 nm
σy = 13 nm
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
σ = 9 nm
Time (min)
c
EAC 14H: Stationary
Intensity (photons)
Time (min)
a
d = 14 nm
v = 4.5 nm/min
σx = 13 nm
σy = 13 nm
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
σ = 9 nm
Time (min)
c
EAC 15H: Stationary
Intensity (photons)
Time (min)
a
d = 32 nm
v = 7.1 nm/min
σx = 15 nm
σy = 15 nm
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
σ = 10 nm
Time (min)
c
EAC 16H: Stationary
Intensity (photons)
Time (min)
a
d = 17 nm
v = 2.1 nm/min
σx = 14 nm
σy = 13 nm
y (nm)
x (nm)
Time (min)
b
Displacement (nm)
σ = 9 nm
Time (min)
c


EAC 17H: Stationary
d = 31 nm
v = 5.0 nm/min
σ = 12 nm
σx = 16 nm
σy = 14 nm
Intensity (photons)
Time (min)
y (nm)
x (nm)
Time (min)
Displacement (nm)
Time (min)
a
b
c


EAC 18H: Stationary
d = 9 nm
v = 1.0 nm/min
σ = 6 nm
σx = 10 nm
σy = 9 nm
Intensity (photons)
Time (min)
y (nm)
x (nm)
Time (min)
Displacement (nm)
Time (min)
a
b
c


EAC 19H: Stationary
d = 25 nm
v = 0.8 nm/min
σ = 5 nm
σx = 6 nm
σy = 5 nm
Intensity (photons)
Time (min)
y (nm)
x (nm)
Time (min)
Displacement (nm)
Time (min)
a
b
c


EAC 20H: Stationary
d = 19 nm
v = 5.0 nm/min
σ = 9 nm
σx = 12 nm
σy = 12 nm
Intensity (photons)
Time (min)
y (nm)
x (nm)
Time (min)
Displacement (nm)
Time (min)
a
b
c

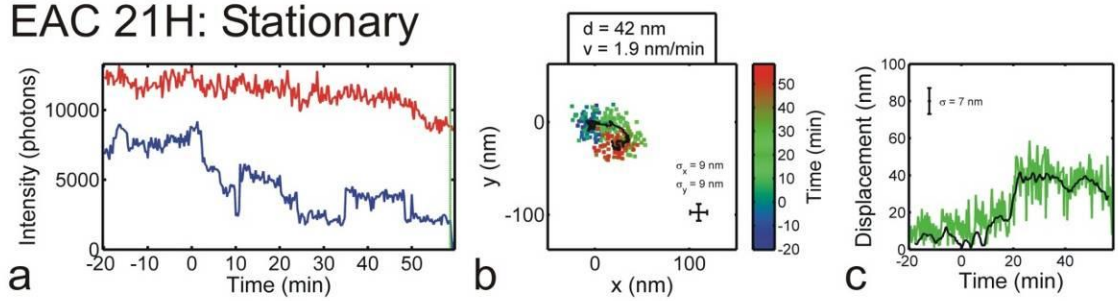

EAC 21H: Stationary
d = 42 nm
v = 1.9 nm/min
σ = 7 nm
σx = 9 nm
σy = 9 nm
Intensity (photons)
Time (min)
y (nm)
x (nm)
Time (min)
Displacement (nm)
Time (min)
a
b
c

Figure D.31 Particularly clean trajectory plots for individual spiders on the EAC (**a**), EABC (**b**), and EABD (**c**) tracks. The EAC trace was collected in 1× HBS + 5 mM $ZnSO_4$ and corresponds to spider EAC 5H (Supplementary Fig. 30), while the EABC and EABD traces were collected in 1× TA-Mg + 1 mM $ZnSO_4$ and correspond to spiders EABC 1 and EABD 1 (Supplementary Fig. 29). The color bars on the left indicate the time in minutes. Zinc was added at time 0. Among the x-y plots for EABC and EABD traces, some were consistent with the prescribed turn (as shown here); however, our resolution was not sufficient to extract features of these landscapes such as turn angles with satisfactory confidence.

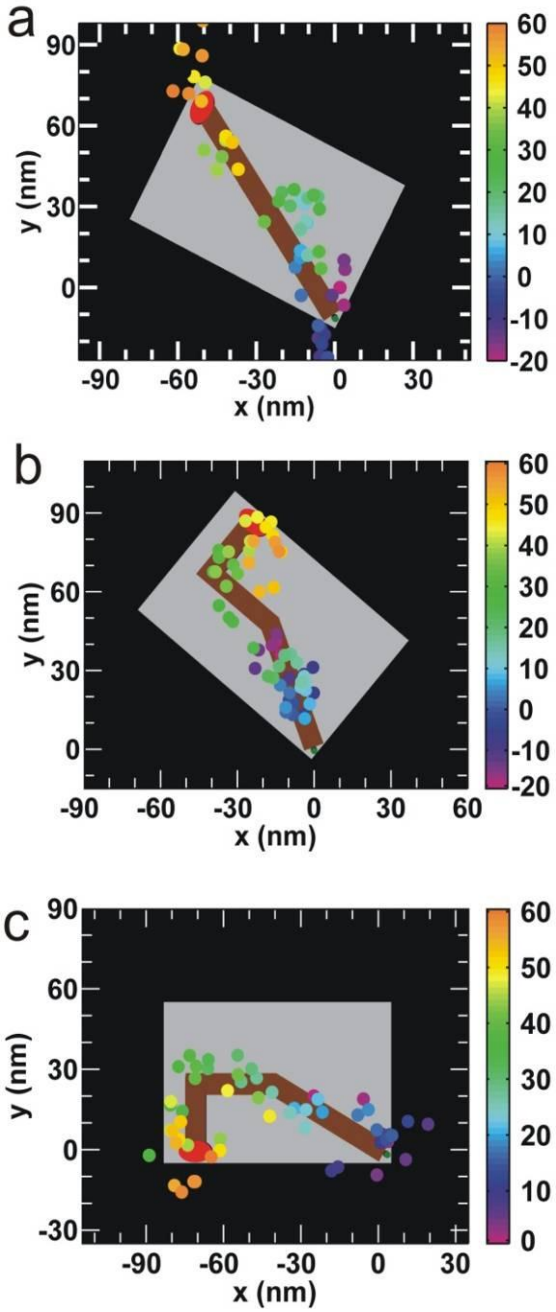
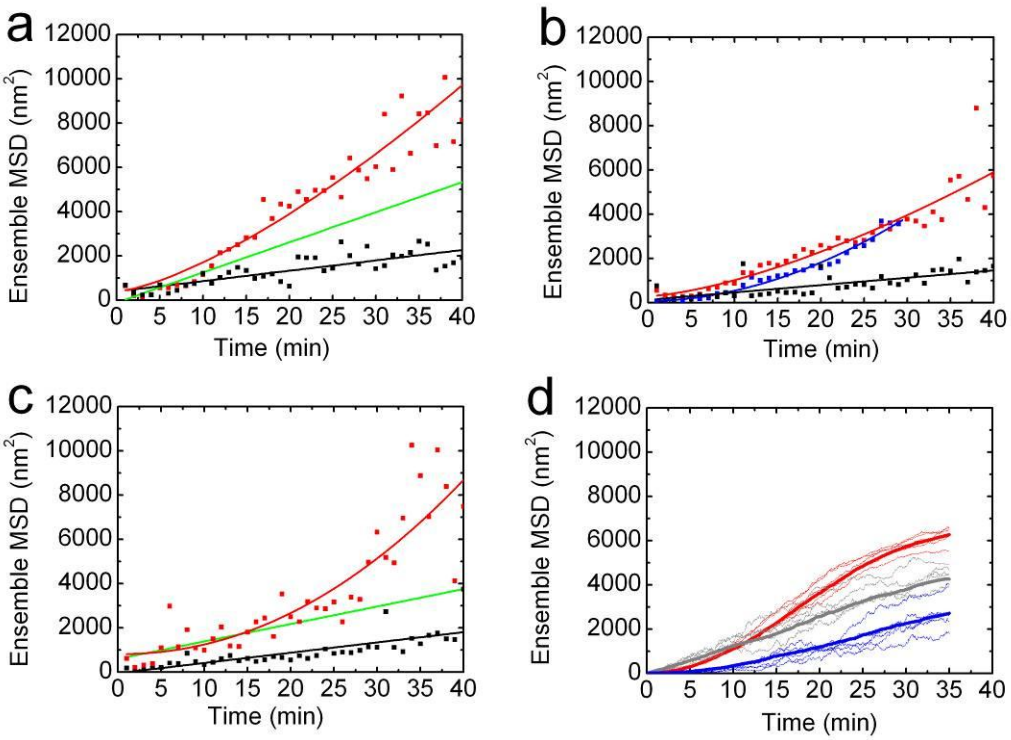


Table D.4 Trajectory filtering statistics for spiders imaged by fluorescence microscopy on the EAC track. These statistics reflect the filtering performed to produce the MSD plot in Supplementary Fig. 32b. The only criterion these spider-origami pairs needed to satisfy is to have detectable Cy3 and Cy5 for at least 1/3 of the experiment's duration (20-30 min). All are imaged in 1× SSC buffer in the presence of Zn^{2+}. The EAC substrate track was in all cases incubated with TRIGGER for 30-60 min prior to imaging, while TRIGGER was added to the EAC product track either 30-60 min (row 2) or 10-15 min (row 3) prior to imaging by fluorescence microscopy.

	Total PSF Pair Candidates	**Spiders Included in S34b**
EAC substrate track	3,821	85
EAC product track, Long TRIGGER pre-incubation	276	29
EAC product track, Short TRIGGER pre-incubation	74	18

Figure D.32 a, Ensemble mean square displacement (MSD) versus time calculated from 15 individual Tier 1 spiders on the EAC substrate track (red squares; EAC 1-15 in Supplementary Fig. 29) in 1× SSC. A power law function (red curve) is fit to the MSD from 1-30 min, and a linear function (green curve) is fit to the first 12 min. For further comparison, an MSD plot is shown that is derived from 7 spiders (black circles; traces found in Supplementary Fig. 28d) from a no-Zn^{2+} control experiment in which spiders are not expected to walk. In both types of experiments, the origami-spider complexes were incubated with TRIGGER for 30-60 min prior to imaging. **b**, Ensemble MSD versus time plots comparing behavior on the substrate-covered (red) and product-covered EAC tracks with TRIGGER added either 30-60 min (long incubation, black) or 10-15 min before (short incubation, blue) imaging by fluorescence microscopy in 1× SSC in the presence of zinc. The MSD values were calculated from 85 individual spiders on the substrate-covered EAC track, 29 spiders on the product-covered EAC track incubated for 30-60 min with TRIGGER, and 18 spiders on the product-covered EAC track incubated for 10-15 min with TRIGGER selected according to intensity, ellipticity, and outlier cutoff criteria stated in the supplement text. All three are fit to power law functions (solid curves). The MSD plot for the substrate track is assembled from the same datasets as the red plot in panel (**a**), but the more relaxed selection criteria result in the inclusion of more stationary or slowly moving spiders, resulting in a shallower curve. **c**, Ensemble MSD versus time plots of spider movement on the EAC substrate track observed in 1× HBS with 1× HBS buffer containing either 0 mM (black curve) or 5 mM (red curve) $ZnSO_4$ added at time t = 0 min. In both types of experiments, the sample was incubated with TRIGGER for 30-60 min prior to the beginning of the experiment. A power law function (red curve) is fit to the MSD with 5 mM zinc from 1-30 min, and a straight line (green curve) is fit to the first 15 min of the MSD with 5 mM zinc. A straight line (black) is also fit to the MSD in 0 mM zinc. **d**, Simulated MSD versus time plots calculated as described in the supplement text from 1,000 spiders (thick lines) or separate trials of fewer spiders (thin lines) for the substrate-covered EAC track (red), and for the product-covered EAC track without (gray) or with a delayed release (blue) from the START region ($t_{1/2release}$ = 0 or 10 min). For the smaller trials, 80 spiders per trial were used for the substrate-covered track, while 20 spiders per trial were used for the product track with and without delayed release (to approximate the numbers of experimental spiders observed in each case).


a
b
c
d
Ensemble MSD (nm^2)
Time (min)
0
2000
4000
6000
8000
10000
12000
0 5 10 15 20 25 30 35 40

References

1 Siegwart, R. & Nourbakhsh, I. R. *Introduction to Autonomous Mobile Robots*. (MIT Press, 2004).
2 Braitenberg, V. *Vehicles: Experiments in Synthetic Psychology*. (MIT Press, 1984).
3 Brooks, R. A. Intelligence without representation. *Artif. Intell.* **47**, 139–159 (1991).
4 Simon, H. A. *The Sciences of the Artificial*. 3rd edn, (MIT Press, 1996).
5 Stojanovic, M. N., Mitchell, T. E. & Stefanovic, D. Deoxyribozyme-based logic gates. *J. Am. Chem. Soc.* **124**, 3555-3561 (2002).
6 Kube, C. R. & Zhang, H. Collective robotics: from social insects to robots. *Adapt. Behav.* **2**, 189-219 (1993).
7 Rus, D., Butler, Z., Kotay, K. & Vona, M. Self-reconfiguring robots. *Commun. ACM* **45**, 39-45 (2002).
8 Dorigo, M. & Stützle, T. *Ant colony optimization*. (MIT Press, 2004).
9 Turing, A. M. On computable numbers, with an application to the Entscheidungsproblem. *Proc. Lond. Math. Soc. Ser.* **2**, 230-265 (1936).
10 Bennett, C. H. The thermodynamics of computation—a review. *Int. J. Theor. Phys.* **21**, 905--940 (1982).
11 Gajardo, A., Moreira, A. & Goles, E. Complexity of Langton's ant. *Discrete Appl. Math.* **117**, 41-50 (2002).
12 Von Neumann, J. & Burks, A. W. *Theory of Self-Reproducing Automata*. (University of Illinois Press, 1966).
13 Bath, J. & Turberfield, A. J. DNA nanomachines. *Nat. Nanotechnol.* **2**, 275-284 (2007).
14 Sherman, W. B. & Seeman, N. C. A precisely controlled DNA biped walking device. *Nano Lett.* **4**, 1203-1208 (2004).
15 Shin, J. S. & Pierce, N. A. A synthetic DNA walker for molecular transport. *J. Am. Chem. Soc.* **126**, 10834-10835 (2004).
16 Yin, P., Choi, H. M. T., Calvert, C. R. & Pierce, N. A. Programming biomolecular self-assembly pathways. *Nature* **451**, 318-U314 (2008).
17 Tian, Y., He, Y., Chen, Y., Yin, P. & Mao, C. D. A DNAzyme that walks processively and autonomously along a one-dimensional track. *Angew. Chem. Int. Ed.* **44**, 4355-4358 (2005).
18 Bath, J., Green, S. J. & Turberfield, A. J. A free-running DNA motor powered by a nicking enzyme. *Angew. Chem. Int. Ed.* **44**, 4358-4361 (2005).
19 Omabegho, T., Sha, R. & Seeman, N. C. A bipedal DNA brownian motor with coordinated legs. *Science* **324**, 67-71 (2009).
20 Pei, R. *et al.* Behavior of polycatalytic assemblies in a substrate-displaying matrix. *J. Am. Chem. Soc.* **128**, 12693-12699 (2006).
21 Green, S. J., Bath, J. & Turberfield, A. J. Coordinated chemomechanical cycles: a mechanism for autonomous molecular motion. *Phys. Rev. Lett.* **101**, 238101 (2008).
22 Reif, J. H. & Sahu, S. Autonomous programmable DNA nanorobotic devices using DNAzymes. *Theoretical Computer Science* **410**, 1428-1439 (2009).
23 Yin, P., Turberfield, A. J., Sahu, S. & Reif, J. H. Design of an autonomous DNA nanomechanical device capable of universal computation and universal translational motion. *Lect. Notes Comput. Sc.* **3384**, 426-444 (2005).

24 Chang, T., Rozkiewicz, D. I., Ravoo, B. J., Meijer, E. W. & Reinhoudt, D. N. Directional movement of dendritic macromolecules on gradient surfaces. *Nano Lett.* **7**, 978-980 (2007).
25 Wang, J. & Manesh, K. M. Motion control at the nanoscale. *Small* **6**, 338-345 (2010).
26 Vives, G. & Tour, J. M. Synthesis of single-molecule nanocars. *Acc. Chem. Res.* **42**, 473-487 (2009).
27 Adleman, L. M. Molecular computation of solutions to combinatorial problems. *Science* **266**, 1021-1024 (1994).
28 Stojanovic, M. N. & Stefanovic, D. A deoxyribozyme-based molecular automaton. *Nat. Biotechnol.* **21**, 1069-1074 (2003).
29 Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585-1588 (2006).
30 Yan, H., Park, S. H., Finkelstein, G., Reif, J. H. & LaBean, T. H. DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science* **301**, 1882-1884 (2003).
31 Rothemund, P. W. K., Papadakis, N. & Winfree, E. Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biol.* **2**, 2041-2053 (2004).
32 Seeman, N. C. DNA enables nanoscale control of the structure of matter. *Q. Rev. Biophys.* **38**, 363-371 (2005).
33 Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 298-302 (2006).
34 Jungmann, R., Liedl, T., Sobey, T. L., Shih, W. & Simmel, F. C. Isothermal assembly of DNA origami structures using denaturing agents. *J. Am. Chem. Soc.* **130**, 10062-10063 (2008).
35 He, Y. *et al.* Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature* **452**, 198-201 (2008).
36 Aldaye, F. A., Palmer, A. L. & Sleiman, H. F. Assembling materials with DNA as the guide. *Science* **321**, 1795-1799 (2008).
37 Ding, B. & Seeman, N. C. Operation of a DNA robot arm inserted into a 2D DNA crystalline substrate. *Science* **314**, 1583-1585 (2006).
38 Santoro, S. W. & Joyce, G. F. A general purpose RNA-cleaving DNA enzyme. *Proc. Natl. Acad. Sci. USA* **94**, 4262-4266 (1997).
39 Antal, T. & Krapivsky, P. L. Molecular spiders with memory *Phys. Rev. E* **76**, 021121 (2007).
40 Saffarian, S., Collier, I. E., Marmer, B. L., Elson, E. L. & Goldberg, G. Interstitial collagenase is a Brownian ratchet driven by proteolysis of collagen. *Science* **306**, 108-111 (2004).