# Appendix A

# Partial Supplementary Material for Molecular Robots Guided by Prescriptive Landscapes[0]

## A.1 Supporting Discussion

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.

[0]This work was coauthored by Kyle Lund, Anthony J. Manzo, Nadine Dabby, Nicole Michelotti, Alexander Johnson-Buck, Jeanette Nangreave, Steven Taylor, Renjun Pei, Milan N. Stojanovic*, Nils G. Walter*, Erik Winfree*, & Hao Yan* and published in 2010 [Lund et al., 2010] with the following contributions: AFM experiments were performed by K.L. (majority), J.N., and N. D.; analysis was performed by N. D., K.L., J.N., S.T., and supervised by E.W., and H.Y.; fluorescence microscopy and particle tracking analysis were performed by A.J.M., N.M., A.J.B, and supervised by N. G. W.; spiders were synthesized, purified, and their integrity confirmed and monitored by S.T.; SPR experiments were performed by R. P.; research coordination by M.N.S., material transfer coordination by S.T., J.N., and K.L. Experimental design and manuscript was done with input from all authors.

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 [Braitenberg, 1984, Brooks, 1991, Simon, 1996, Siegwart, 2004] 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 [Stojanovic et al., 2002]; interactions between multiple molecular robots could lead to collective behavior [Kube and Zhang, 1993, Rus et al., 2002, Dorigo and Stützle, 2004]; 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 [Turing, 1936, Von Neumann and Burks, 1966, Bennett, 1982, Gajardo et al., 2002].

Research in programmable DNA walkers [Bath and Turberfield, 2007] started with non-autonomous remote-controlled systems [Sherman and Seeman, 2004, Shin and Pierce, 2004], progressed to autonomous walkers that modify visited sites to achieve directed (but brief) motion on linear tracks [Bath et al., 2005, Tian et al., 2005, Yin et al., 2008, Omabegho et al., 2009], or to achieve continuous processive (but undirected) motion in two or three dimensions [Pei et al., 2006], and shows promise for processive and directed walking on undisturbed tracks [Green et al., 2008]. While synthesizing suitably well-defined tracks has been an important technical challenge (no previous walker has been demonstrated to take more than three steps on a linear track), our interest here 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 [Adleman, 1994, Stojanovic et al., 2003, Seelig et al., 2006, Yin et al., 2008], complex structures [Yan et al., 2003, Rothemund et al., 2004, Seeman, 2005, Rothemund, 2006, Aldaye et al., 2008, He et al., 2008, Jungmann et al., 2008] and actuators [Pei et al., 2006, Ding

and Seeman, 2006]. 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 [Santoro and Joyce, 1997]. In the context of substrates that are immobilized at sites on a surface, spider behavior can be modeled using local rules [Antal, 2007]: 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) to 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 [Sherman and Seeman, 2004, Shin and Pierce, 2004, Bath et al., 2005, Tian et al., 2005, Omabegho et al., 2009] and observed in nature [Saffarian et al., 2004], 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 [Rothemund, 2006], the previously introduced processive but random walker [Pei et al., 2006] 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.

## A.2 Materials and Methods

### A.2.1 Abbreviations

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

### A.2.2 Preparation of Spiders

**Materials and Instrumentation for the Preparation and Characterization of NICK$_{3.4A+1}$ and NICK$_{3.4A+1}$· (Cy3)$_3$.** Synthesis and purification of the modified DNA strands used to construct NICK$_{3.4A+1}$ and NICK$_{3.4A+1}$· (Cy3)$_3$ were carried out by Integrated DNA Technologies (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, $4.6 \times 50$ mm (IDxL) (Tosoh Biosciences). Concentrations of oligonucleotides were determined on an Amersham Biosciences Ultrospec 3300 pro UV/visible spectrophotometer. **Assembly of NICK$_{3.4A+1}$.** 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 $K_3PO_4$, pH 6.5). The desired one-to-one conjugate product (STV-(C)1) was purified by ion exchange (IE) HPLC. Part B; deoxyribozyme leg [5′ - /5BioTEG//iSp18//iSp18/TCT CTT CTC CGA GCC GGT CGA AAT AGT GAA AA - 3′] (L) (100$\mu$M, 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. NICK$_{3.4A+1}$. 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. The

assembly was purified by IE-HPLC and the volume of the elutent reduced (by centrifugation) to give a final concentration of $2.3\mu$M, as determined by absorbance at $260$ nm. Characterization of the assembly was carried out by IE-HPLC and PAGE. 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 NICK$_{3.4A+1}$ 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 ($\epsilon_{max}150,000$ M$^{-1}$cm$^{-1}$) and $260$ nm (Extinction coefficient max $1,220,000$ M$^{-1}$cm$^{-1}$) respectively.

One should note that 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)$_4$.

### A.2.3 Surface Plasmon Resonance (SPR)

**Materials and Instrumentation for SPR Experiments.** Immunopure avidin was purchased from Pierce (Rockford, USA). We used a Biacore X system, commercially available Biacore SA sensor chips, and Biacore C1 sensor chips, from GE Healthcare (Piscataway, USA). $1\times$ HBS buffer ($10$ mM HEPES, pH $7.4$ with $150$ mM NaCl) was employed as running buffer.

**Preparation of Substrates on pseudo-2D Hydrogel Matrix Surfaces for SPR**. A $20\mu$M 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 $16$ minutes at $5\mu$L/min, followed by a $60$ second 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 [Pei et al., 2006].

**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 $5\mu$L/min by using a $7$ minute 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 $80\mu$L injection of $1$ mg/mL avidin (in $1\times$ HBS) was flowed over the activated surfaces of both channels for $40$ minutes at $2\mu$L/min. The remaining activated sites on the chip surfaces were blocked with a $35\mu$L injection of an ethanolamine hydrochloride solution ($1$ M, pH $8.5$). Then, a $20\mu$M solution of cleavable substrate was applied to both channels of C1 sensor chip for $20$ minutes at $4\mu$L/min, followed by a $60$ second 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 $P) = Mw(S$ or $P) \times RU(spider)/[Mw(spider) \times RU(S$ or $P)]$. Monitoring the dissociation of the spider was performed in $1\times$ 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 on product. We observed that over the course of 30 minutes more than $92\%$ of spiders remained on a product covered surface and over the course of 60 minutes $86\%$ remained bound. 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 $1\times$ HBS buffer) were loaded only on channel 2 at $5\mu$L/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 $1\times$ TA-Mg buffer with 1 mM $ZnCl_2$ or $1\times$ 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·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.

### A.2.4 Preparation of Spider-Origami Arrays

**Assembly of Spider-Origami Arrays for Atomic Force Microscopy (AFM).** The spider arrays consist of M13mp18 viral DNA (New England Biolabs) and 202 ssDNA staples (Integrated DNA Technologies, see below for DNA sequences). The arrays were annealed in $1\times$ TA-Mg Buffer (40 mM Tris, 20 mM acetic acid, 12.5 mM $Mg^{++}$, 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°C to 25°C using an Eppendorf PCR machine (Eppendorf). The NICK$_{3.4A+1}$ 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 temperature

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°C to 24°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 $Zn^{++}$), 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.

**Modification of Spider-Origami Arrays for Fluorescence Microscopy.** To make the origami arrays compatible with fluorescence microscopy, we returned four 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. 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-cleavable 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.

### A.2.5 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 10 mM $ZnCl_2$ 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 minutes (depending on the experiment) at room temperature before the origami were deposited on mica. Samples ($2\mu$L) were deposited onto a freshly cleaved mica surface (Ted Pella, Inc.) and left to adsorb for 3 minutes. Buffer ($1\times$ TA-Mg, $400\mu$L) 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 $Zn^{++}$ (see AFM Imaging for Movie) and the consistent trends in the time-lapse experiments (Fig. 2.2, main text), 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 ($2\mu$L) was deposited onto a freshly cleaved mica surface and left to bind for 2 minutes. Then $1\mu$L of TRIGGER strand was added to the sample on the surface and after 2 minutes $270\mu$L of buffer and $30\mu$L of 10 mM $ZnCl_2$ was added to the sample cell. The four images were taken over a 26-minute time frame with about 10 minutes 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 Fig. 2.3, 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 Fig. 2.3) 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 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. 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 $1 \times 1\mu$m 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). 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. 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 six 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. 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 (Fig. 2.2).

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 (Fig. 2.2c,g, main text). Our shortest track (ABD, spanning $48$ nm) efficiently delivers spiders to the STOP, with less than $20\%$ of spiders on the TRACK after $30$ minutes (Fig. 2.2c, main text). If the TRACK was omitted on the ABD landscape, spiders were equally distributed between the STOP and CONTROL sites after $30$ minutes, implying that the track is needed for efficient delivery to the STOP site. On longer TRACKS (such as EABD, spanning $90$ nm) $15\%$ of spiders are delivered to the STOP within $15$ minutes after release. Longer incubation times ($30$ and $60$ minutes) increase the efficacy of delivering spiders to the STOP to up to $70\%$, (Fig. 2.2c,g, main text). Even at $60$ minutes, 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) $30$ minutes after release (Fig. 2.2c,g, main text).

### A.2.6 DNA Sequences

| 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 |

| Name | Sequence |
|---|---|
| 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 |

| Name | Sequence |
|---|---|
| 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 |

| Name | Sequence |
|---|---|
| 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 |

| Name | Sequence |
|---|---|
| 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 |

| Name | Sequence |
|---|---|
| 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 |

| Name | Sequence |
|---|---|
| 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 |

| Name | Sequence |
|---|---|
| 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 |

| Name | Sequence |
|---|---|
| 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 |

| Name | Sequence |
|---|---|
| 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 |

| Name | Sequence |
|---|---|
| 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-cleavable 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