Tuning the performance of an artificial protein motor

Nathan J. Kuwada, Martin J. Zuckermann, Elizabeth H. C. Bromley, Richard B. Sessions, Paul M. G. Curmi, Nancy R. Forde, Derek N. Woolfson, and Heiner Linke

Phys. Rev. E **84**, 031922 — Published 21 September 2011

DOI: 10.1103/PhysRevE.84.031922

Tuning the performance of an artificial protein motor

Nathan J. Kuwada

Department of Physics and Materials Science Institute,

1274 University of Oregon, Eugene, OR 97403, USA and

Department of Physics, University of Washington, Box 351560, Seattle, WA 98195[*]

Martin J. Zuckermann

Department of Physics and IRMACS Centre, Simon Fraser University,

Burnaby, British Columbia, V5A 1S6, Canada

Elizabeth H. C. Bromley

Department of Physics, Durham University,

Durham DH1 3LE, United Kingdom

Richard B. Sessions

School of Biochemistry, University of Bristol, Bristol BS8 1TD, United Kingdom

Paul M. G. Curmi

School of Physics, University of New South Wales,

Sydney, New South Wales 2052, Australia

Nancy R. Forde

Department of Physics, Simon Fraser University,

Burnaby, British Columbia, V5A 1S6, Canada

Derek N. Woolfson

School of Chemistry and School of Biochemistry,

University of Bristol, Bristol BS8 1TS, United Kingdom

Heiner Linke

The Nanometer Structure Consortium (nmC@LU) and Division of Solid State Physics,

Lund University, Box 118, 22100 Lund, Sweden[†]

(Dated: August 31, 2011)

Abstract

The Tumbleweed (TW) is a concept for an artificial, tri-pedal, protein-based motor designed to move unidirectionally along a linear track by a diffusive tumbling motion. Artificial motors offer the unique opportunity to explore how motor performance depends on design details in a way that is open to experimental investigation. **Prior studies have shown that TW's ability to complete many successive steps can be critically dependent on the motor's diffusional step time.** Here, we present a simulation study targeted at determining how to minimize the diffusional step time of the TW motor as a function of two particular design choices: non-specific motor-track interactions and molecular flexibility. We determine an optimal non-specific interaction strength and establish a set of criteria for optimal molecular flexibility as a function of the non-specific interaction. **We discuss our results in the context of similarities to biological, linear stepping diffusive molecular motors with the aim of identifying general engineering principles for protein motors.**

* nateman@uw.edu

† heiner.linke@ftf.lth.se

I. INTRODUCTION

Inspired by biological molecular motors and the desire to produce controllable nano-scale machinery, the effort to design and construct artificial molecular-scale motors has received considerable attention. One successful class of artificial molecular motors uses relatively small synthetic molecules whose conformation can be controlled by chemistry or external light sources [1–4]. Examples of this class are motors consisting of ring-like chiral molecules that can be rotated unidirectionally about a fixed axis by external control, allowing for molecular 'shuttles' [5] and 'molecular muscles' [6]. A second class of artificial molecular motor uses oligonucleotides as the motor building blocks [7], which, due to relatively straightforward self-assembly rules, have allowed for the construction of programmable 'DNA walkers' [8], 'molecular spiders' [9], and nanoscale 'assembly lines' [10]. However, most biological molecular motors are based on proteins. Designing an artificial molecular motor using protein components may thus provide insight into subtle structure-function aspects of bio-molecular motors. We have proposed an artificial motor concept, the Tumbleweed, which uses proteins as motor building blocks [11]. Proteins offer more diverse design choices than oligonucleotide structures because of the relatively large set of available amino acid building blocks that can produce large, three-dimensional structures. The Tumbleweed motor, thus, offers a unique opportunity to not only design a functioning motor but also to actively tune the molecular design to *optimize* motor performance. **The overall aim of this modeling study is to determine how motor performance depends on design details that are, at least in principle, open to experimental investigation. In addition, we are interested in discovering advantageous design strategies that also can be identified in biological motors.**

The Tumbleweed (TW) is a tri-pedal construct consisting of two main protein components: a designed coiled-coil, Y-shaped central hub consisting of three 'legs,' and three unique DNA-binding repressor proteins (R_{A}, R_{B}, and R_{C}) attached to each leg (Fig. 1) [11]. Each repressor protein 'foot' binds with high affinity to a unique double-stranded DNA (dsDNA) recognition sequence only when it has bound a specific ligand (*a, b,* and *c*, respectively) whose concentration in solution can be controlled externally. Thus, by using a dsDNA 'track' with cyclic, equally spaced repeats of the three unique repressor motifs, motor stepping can be achieved as follows: by cyclically changing the buffer around the motor,

the ligand concentration is cycled in the order $[a,b]$, $[b,c]$, $[c,a]$, where each concentration is held constant for an **experimentally defined time $\tau_{\mathbf{ligand}}$ [12].** When the ligand concentration is changed, one foot loses its ligand as its concentration drops in solution and the foot releases from the DNA, while the other foot remains tightly bound. The molecule then undergoes unbiased tethered diffusion until the third foot locates its recognition sequence and tightly binds.

In order for the TW motor to complete successive steps, it must coordinate a variety of processes across many time scales. During a single motor step, the motor must diffusively locate its next binding site before the ligand concentration is changed (that is, during τ_{ligand}) otherwise the motor will lose register and fall off the track. In addition, the characteristic lifetime of the tightly bound state of the stationary foot (τ_{bound}) must be longer than τ_{ligand} otherwise the motor is likely to fall off the track. These requirements can be expressed in the inequality

$$\tau_{\mathrm{diff}} \;<\; \tau_{\mathrm{ligand}} \;<\; \tau_{\mathrm{bound}}, \tag{1}$$

where τ_{diff} is the diffusional search (step) time. A prior simulation study of the TW motor using a classical Master equation model has shown that the ability of the motor to perform successive stepping events can be critically sensitive to the diffusional step time even if τ_{diff} is three orders of magnitude smaller than τ_{ligand} and τ_{bound} [13], because the motor is especially sensitive to track detachment during the stepping process when it is bound by only one foot. Thus, a basic molecular and experimental design question is: What are the physical contributions to τ_{diff}, and how can they be adjusted?

The choice of using DNA-binding proteins as the motor feet may have unintended consequences on the diffusional search time of the motor. Although the repressor protein feet bind tightly to a specific DNA sequence, there is also a weaker sequence-independent attraction present. *In vivo*, this non-specific binding is believed to be a crucial component of the site-specific search process of DNA-binding proteins because it can facilitate a relatively fast 1D 'sliding' diffusional search along the DNA, versus a completely 3D diffusional search. But it is not clear *a priori* how non-specific binding may affect the TW motor: although the leading foot may be assisted by non-specific binding, the lagging foot will also be affected and may take longer to release from the track. Another particular design choice that may affect the diffusional search time of the motor is molecular flexibility. The molecule has

two well defined 'joints' whose flexibility is adjustable at least in principle. A completely rigid motor will have less diffusional space to explore and may bind more quickly than a completely flexible motor. But previous modeling results have shown that a rigid motor is very sensitive to the binding site separation (determined by the structure of the dsDNA track), and thus it is also not clear *a priori* what optimal flexibility minimizes the diffusional search time [11].

Here we address the above questions using a coarse-grained Langevin dynamics model. We find that there is an optimal strength of the non-specific binding interaction to reduce the diffusional search time of the motor and that molecular flexibility appears generally preferable for low non-specific binding strengths. However, some molecular rigidity can mitigate problems associated with high non-specific interactions. We discuss how these results relate to biological molecular motors.

II. COARSE-GRAINED MODEL

To simulate TW diffusional stepping, we use a **three dimensional coarse-grained Langevin dynamics model.** The construct, shown in Fig. 1(a), is represented as four spheres, where the size of the spheres (therefore, their drag coefficient) is set to match the approximate size of the corresponding protein components (feet and legs). Unless stated otherwise, the hub-foot separation is assumed constant, matching the structural rigidity of the coiled-coil 'leg' proteins, such that the sphere configuration matches the geometry and flexibility of the original molecule, as shown in Fig. 1(b). The equation of motion for each sphere is an overdamped Langevin equation [15],

$$v_i(t) \;=\; F_i(t)/\gamma_i \;+\; \xi_i(t), \tag{2}$$

where the instantaneous velocity $v_i(t)$ of sphere i is determined by the sum of conservative forces on the sphere, $F_i(t)$, by the viscous drag coefficient of the sphere (calculated using Stokes' Law), γ_i, and by a Gaussian white noise term simulating Brownian motion, $\xi_i(t)$, which is a random number pulled from a Gaussian distribution with zero mean and variance $< \xi_j \; \xi_k > \;\; = \; \delta_{jk} \; 2kT/\gamma_i \Delta t$, where Δt is the time step of the simulation. The conservative forces on the spheres are calculated as the sum of gradients of potentials, $F_i \;=\; \sum_n (-\nabla_{r_i} V_n)$. The position of each foot sphere relative to the hub is maintained by

a harmonic tethering potential,

$$V_1(r_{ik}) \;=\; V_{\rm harm}\;\left(r_{ik} - r_{\rm leg}\right)^2 \tag{3}$$

where r_{ik} is the instantaneous distance between foot sphere i (i=1,2,3) and the hub sphere (k) while $r_{\rm leg}$ is the equilibrium sphere separation, matching the length of a leg in the original molecule. **All simulations presented here use $V_{\rm harm}$ = 100 kT.** The excluded volume of the molecule is modeled by a repulsive Lennard-Jones potential between all spheres,

$$V_2(r_{ij}) = \begin{cases} \left(\frac{\zeta}{r_{ij}}\right)^{12} - \left(\frac{\zeta}{r_{ij}}\right)^{6}) & r_{ij} < 2^{1/6}\zeta \\ -1/4 & r_{ij} > 2^{1/6}\zeta \end{cases} \tag{4}$$

where ζ is the **steric diameter** of each sphere and r_{ij} is the center-to-center distance between spheres i and j.

The original molecule has two well-defined regions of flexibility: the central hub where the three coiled-coils meet and the links between each coiled-coil and its repressor, referred to as the 'hub joint' and 'ankle joint,' respectively. Each of these joints consists of a relatively short polypeptide chain, or linker, whose contour length is not *a priori* determined by any structural constraints. The flexibility of these joints is expected to be determined by two factors: First, the length of the linker (number of amino acids), e. g. if the ankle joint linker is very short, collisions between the coiled-coil and the repressor protein are likely, and will lead to a constrained configuration space for the ankle joint; Second, flexibility depends on whether the designed flexible linker adopts a non-intended well-defined rigid structure to increase the rigidity.

In order to explore the general form of the ankle joint potential we performed an all-atom, three-dimensional molecular dynamics simulation of a foot bound to DNA with its associated coiled coil in explicitly modeled water. The ankle joint flexed through some 80° during this 70 ns unconstrained simulation (Fig. 2(a)). Structures sampled during the trajectory were used to guide the construction of an arc of conformations defining a particular pathway for the ankle-bending process (Fig. 2(b)). We used umbrella sampling to estimate the free-energy cost of movement along this pathway (further details are provided in Supplemental Material [14]). The profile shown in Fig. 2(c) comprises a relatively flat region across the centre, bounded by steep rises on either side. We interpret this

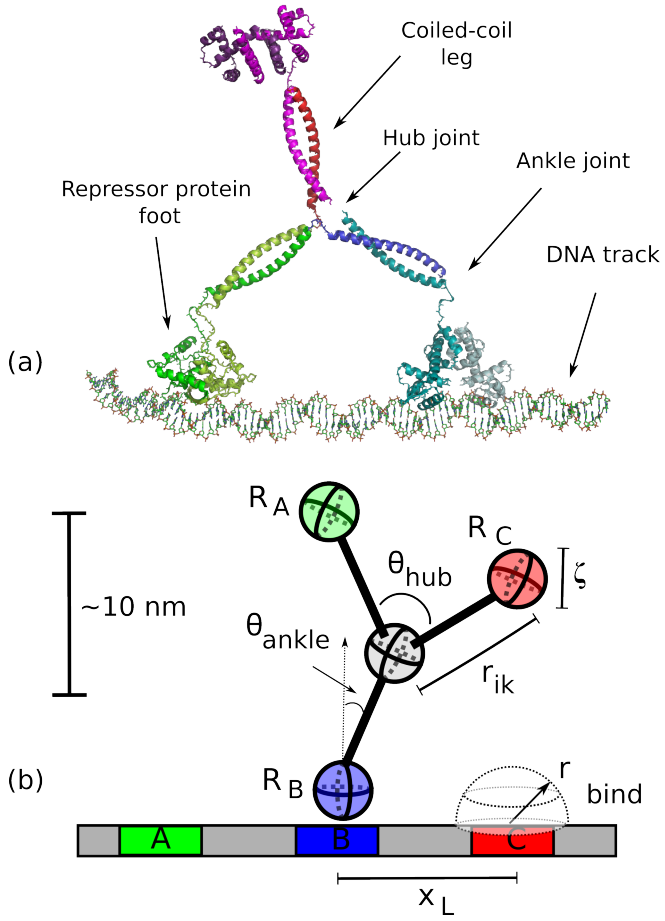

FIG. 1. (Color online) (a) Protein structure of one possible realization of the TW motor [11].Three repressor protein 'feet' are attached by 'ankle' joints to a hub made up of three rigid, coiled-coil 'legs.' (b) Three-dimensional, coarse-grained Tumbleweed model used in Langevin simulations. The molecule is represented as four connected spheres. The separation between spheres, $r_{\rm ik}$ is maintained by a harmonic potential, and the excluded volume (ζ) of the molecule is maintained by a repulsive Lennard-Jones potential. A molecule with a 'rigid' ankle minimizes the angle $\theta_{\rm ankle}$ between the bound leg and a vector normal to the track, and a molecule with a rigid central hub joint has $\theta_{\rm hub} = 120^\circ$ **(where $\theta_{\rm hub}$ is the angular separation of two legs in the plane defined by the two legs)**. Site-specific foot binding is assumed if the foot is a distance less than $r_{\rm bind}$ from its specific binding site.

rise in energy as being due to expected steric clashes between the coiled-coil and the repressor. We note, however, that the simulations explore only one possible pathway from one side to the other, and the observed barrier value

is therefore an upper limit. Furthermore, other pathways would be expected to report different details about the landscape at the bottom of the potential. Consequentially, we do not model this surface in detail here (for example, many more pathways would have to be explored in order to determine whether the double minimum observed is a general property of the energy surface or merely special to this particular pathway). Nevertheless, the simulations supply us with important information about the range of angles that can be explored by the ankle. In our model we will use Eq. 5 to describe the ankle motion, and, based on our simulations, we establish the range 10 kT to 50 kT as meaningful for $V_{\mathbf{ankle}}$, with 30 kT being the best value (Fig. 2).

In the coarse-grained model, the general shape of the potential can be described to a first approximation as

$$V_3(\theta) \;=\; V_{\text{ankle}} \;\left(1 - cos\left(\theta_{\text{ankle}}\right)\right), \tag{5}$$

where θ_{ankle} is the angle between a specifically-bound motor leg and a vector normal to the track, and V_{ankle} parameterizes ankle joint flexibility (see Figure 1). Samples of this potential are shown in Fig. 2(c) for a variety of V_{ankle} values. We assume the potential for the hub joint to be of the same form,

$$V_4\left(\theta_{\text{hub}}\right) \;=\; V_{\text{hub}} \;\left(1 - cos\left(\theta_{\text{hub}} \;-\; \theta_0\right)\right), \tag{6}$$

where θ_{hub} is the angle between two legs of the molecule, θ_0 is the equilibrium angle between legs of the molecule, and V_{hub} parameterizes the flexibility of the hub joint. **The angular separation of each pair of legs is determined by an independent potential function of the form of Eq. 6. Initially, we will choose a completely flexible hub joint ($V_{\mathbf{hub}} \;=\; 0$ kT).**

Because we assume the specific binding process to be much faster than the diffusional time scale, site-specific binding occurs instantly in the simulation when the distance between a motor foot and its specific binding site is less than 1 nm, similar to the Debye length in solution. The model also includes a hard floor in the plane of the track, modeled using specular reflection, as a computational convenience.

To characterize the diffusional stepping time of the molecule we build a distribution of

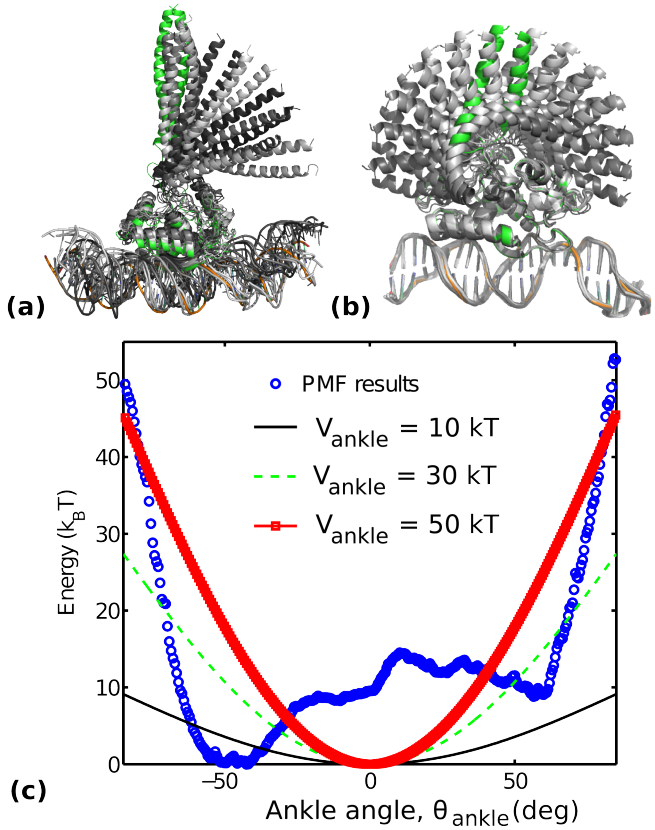

FIG. 2. (Color online) Molecular dynamics study of ankle flexibility. (a) Configurations of the coiled coil with respect to the repressor bound to DNA found during an all-atom simulation of the ankle. The starting configuration is shown in green. (b) Starting configurations for the calculation, the original point taken from the previous simulation is shown in green. (c) Graph of energy as a function of coiled-coil angle generated from umbrella sampling data overlayed with different choices of the ankle potential strength $V_{\rm ankle}$ (Eq. 5) used in the Langevin simulations.

first-passage times. The first-passage time of the motor is defined as the time from lagging foot specific-binding release to leading foot specific-binding. **To get a characteristic first-passage time from a distribution of independent step times, we fit the distribution with a single decaying exponential function, where error bars are taken as the 5% confidence interval of the fit parameter $\tau_{\rm diff}$ [11].**

Parameter	Model Value
Sphere diameter, ζ	4 nm
Leg length, r_{leg}	6.35 nm
Drag coefficient, γ_i	3.8 x 10^{-9} kg/s
Binding site separation, x_L	11 nm
Binding length (capture radius), r_{bind}	1 nm

TABLE I. Langevin Dynamics simulation input parameters.

III. NON-SPECIFIC PROTEIN-DNA BINDING

When modeling repressor proteins, we need to include the possibility of a weaker, sequence-independent DNA-protein binding behavior [17]. *In vivo*, DNA-binding proteins are thought to take advantage of non-specific binding to reduce the time it takes to diffusively find their specific binding site. To locate their specific binding site, binding proteins combine standard 3D diffusion with relatively fast 1D diffusive slides along the DNA [18]. In the cell, DNA is usually tightly packed such that physically adjacent sections of DNA may sequentially be many bases apart. Proteins make 3D diffusive 'hops' between DNA sections, then undergo a facilitated 1D diffusive search along the DNA until they find their specific sequence or dissociate from the track.

In terms of the Tumbleweed motor we may therefore expect that the non-specific interaction reduces the time for the leading foot to diffusively locate its specific binding site. However, if the strength of non-specific binding is too high, the lagging foot may not release from the track on the timescale of ligand exchange. It is therefore *a priori* not clear whether non-specific binding is advantageous or disadvantageous for TW stepping, and whether an optimal value of non-specific binding strength exists.

The physical details of the non-specific interaction are not entirely understood. Recent experimental results [19] support existing theory [20, 21] that some non-specifically bound proteins diffuse in a helical path along the groove of the DNA, but this behavior has not been confirmed for a wide variety of DNA-binding proteins [17]. Because the non-specific interaction is not characterized for Tumbleweed's specific DNA-binding proteins, we instead choose to model the non-specific interaction as isotropic along the DNA, as calculated by

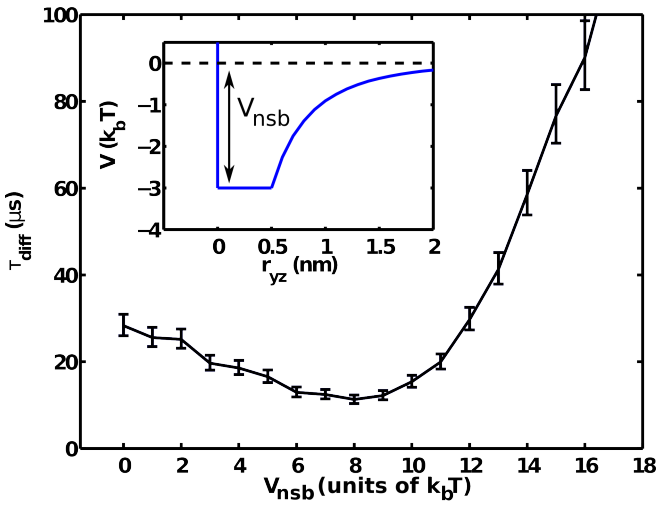

FIG. 3. (Color online) Tumbleweed motor diffusional step time ($\tau_{\rm diff}$) as a function of non-specific binding strength, **with $V_{\rm ankle} = V_{\rm hub} = 0\ kT$.** For an isotropic non-specific binding potential (inset, shown for $V_{\rm nsb} = 3\ kT$), $\tau_{\rm diff}$ initially decreases as a function of $V_{\rm nsb}$, reaches a minimum value at $V_{\rm nsb} \sim 8\ kT$, then dramatically increases for higher values of $V_{\rm nsb}$.

Dahriel, et. al. for a generically shaped DNA-binding protein [22]. According to this work, the energy landscape of the non-specific interaction has a 3 kT minimum 0.5 nm from the DNA surface and a shallow tail that extends $\sim$ 2 nm into solution. **We use the following potential function to model this behavior (shown in the inset of Fig. 3):**

$$V_5(r_{yz}) = \begin{cases} -V_{\rm nsb}\ \frac{e^{-r_{yz}/\xi}}{(r_{yz}/\xi)} & \text{for } r_{yz} > r_{cut} \\ -V_{\rm nsb} & \text{for } r_{yz} < r_{cut} \end{cases} \tag{7}$$

where r_{yz} is the distance away from the DNA track **(taking the x-axis parallel to the track)** and ξ is the characteristic interaction length. The strength of the non-specific interaction is parameterized by $V_{\rm nsb}$, which is the depth of the potential function. **The specific potential function Eq. 7 was chosen to satisfy the following desired properties: Firstly, it is flat from $r_{cut} = 0.5$ nm to the hard-walled track surface in order to avoid a singularity of the potential at $r_{yz} = 0$ nm and to maintain a potential well (please refer to the inset of Fig. 3). Secondly, it provides a gradual rise at larger distances, parameterized by $\xi = 1$ nm, combarable to the Debye length in solution.** Note that the non-specific binding potential is modeled to be ligand-independent, i. e. a repressor with or without its associated ligand is treated in the same manner.

Figure 3 shows the diffusional step time, τ_{diff}, as a function of V_{nsb}. Without any non-specific interaction, we see the step time of the motor is $\sim$ 30 μs. This is in good accordance with previous analytic estimations of the step time of the motor modeled as a cylinder rotationally diffusing 120° and translationally diffusing 11 nm [11]. As V_{nsb} is increased, we see a decrease in τ_{diff}, which reaches a minimum value of $\sim$ 15 μs for $V_{\text{nsb}} \sim 8\ kT$. As V_{nsb} is increased further, though, the diffusional step time increases dramatically, likely due to the lagging foot taking longer to release from the track. One possible way to promote lagging foot release, and potentially decrease τ_{diff}, may be adding some rigidity to the ankle joint of the tightly bound stationary foot, which may act as a lever arm and tear a non-specifically bound lagging foot off of the track. We address this possibility in the next section.

IV. ANKLE JOINT FLEXIBILITY AND NON-SPECIFIC BINDING

Figure 4 shows the diffusional step time of the Tumbleweed motor as a function of ankle rigidity, V_{ankle}, for different fixed values of V_{nsb}. For relatively weak non-specific binding ($V_{\text{nsb}} < 10\ kT$), we see that the diffusional step time of the motor always increases for increasing ankle rigidity. But for high non-specific binding strength, where we expect rigidity to help the motor release its non-specifically bound lagging foot, we do in fact see a slight reduction in τ_{diff} for increasing rigidity.

If the strength of non-specific binding is fixed, we now have an understanding of how to adjust the molecular design to retain the best performing motor: for low non-specific interactions, the ankle joint should be as flexible as possible to reduce τ_{diff}, while for high non-specific interactions, a rigid ankle can somewhat mitigate the ill effects of non-specific binding. But what if the flexibility of the ankle joint is constrained by design requirements? Can we similarly adjust the non-specific binding strength to mitigate the ill effects of a rigid ankle?

Figure 5 shows τ_{diff} as a function of V_{nsb} for different fixed values of V_{ankle}. We see that the general shape of Fig. 3 is maintained for all values of V_{ankle}, although the maximum at $V_{\text{nsb}} = 0$ appears much more pronounced for higher values of V_{ankle}. Compared to Fig. 4, where the reduction in τ_{diff} was at most $\sim$ 30 %, we see that τ_{diff} can be reduced nearly an order of magnitude for $V_{\text{ankle}} = 50\ k_B T$ by tuning V_{nsb}.

Tuning V_{nsb} appears much more effective at reducing τ_{diff} for a motor with rigid ankles

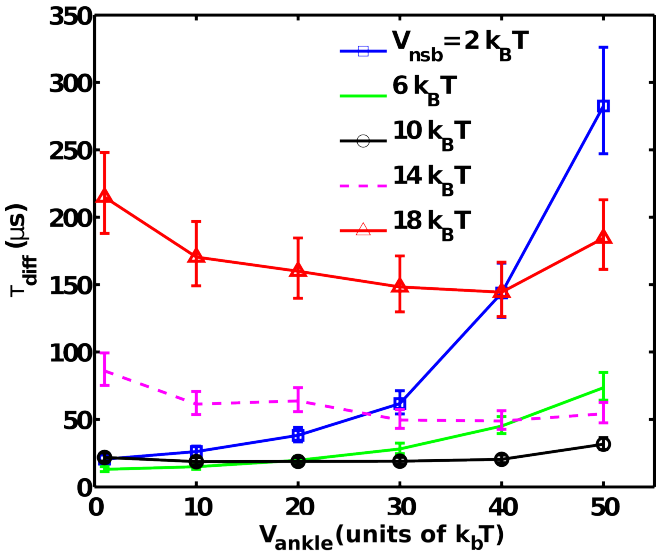

FIG. 4. (Color online) $\tau_{\rm diff}$ as a function of $V_{\rm ankle}$, each curve representing different fixed values of $V_{\rm nsb}$(**with** $V_{\rm hub} = 0\ kT$). For $V_{\rm nsb} < 10\ kT$, increasing ankle rigidity always increases the diffusional step time. For $V_{\rm nsb} > 10\ kT$, $\tau_{\rm diff}$ can be slightly reduced by tuning the rigidity of the ankle joint, with a ~ 30 % reduction for $V_{\rm nsb} = 18\ kT$ and $V_{\rm ankle} = 40\ kT$.

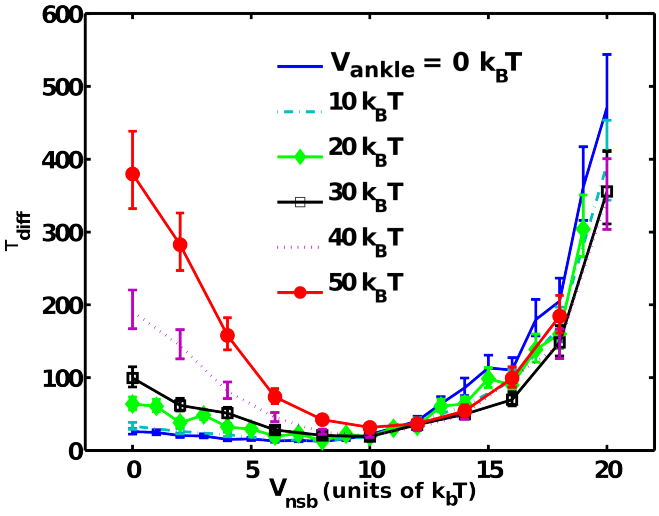

FIG. 5. (Color online) $\tau_{\rm diff}$ as a function of $V_{\rm nsb}$, different curves corresponding to different fixed values of $V_{\rm ankle}$(**with** $V_{\rm hub} = 0\ kT$). Tuning the strength of non-specific binding can greatly mitigate the ill effects of a rigid ankle, reducing $\tau_{\rm diff}$ nearly an order of magnitude for $V_{\rm ankle} = 50\ kT$.

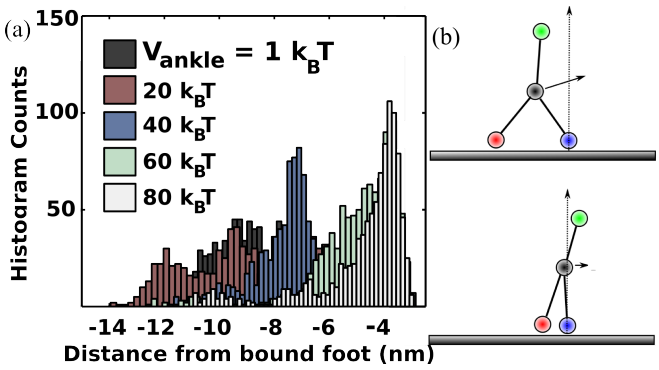

FIG. 6. (Color online) (a) Histograms of a non-specifically bound lagging foot for increasing ankle rigidity. Because the central hub joint of the molecule is completely flexible, the non-specifically bound lagging foot slides right next to the bound stationary foot as the ankle joint straightens (shown schematically in (b)). **The effect becomes increasingly strong for values of $V_{\mathbf{ankle}} > 20\ kT$.**

than tuning V_{ankle} for a motor with high non-specific interactions. What are the physical reasons for this behavior? Recall that the central joint of the motor, representing the Y-shaped coiled-coil hub joint, is assumed to be a completely flexible joint. This flexibility allows the motor feet to slide along the track while non-specifically bound. Consider a motor with a rigid ankle and high non-specific binding interactions immediately after the lagging foot releases from its binding site (shown in Fig. 6(b)). Because there is no barrier to sliding along the track (besides viscous drag on the repressor protein), the rigid ankle joint can simply slide the non-specifically bound lagging foot right next to the specifically bound stationary foot (Fig. 6(b)). In this configuration, the ankle joint is nearly vertical, and thus does not exert much vertical force on the non-specifically bound lagging foot to help release it from the track as expected. To confirm this behavior, we can plot histograms of the position along the track of a non-specifically bound lagging foot as a function of V_{ankle}, shown in Fig. 6(a). We see as V_{ankle} is increased, the lagging foot does indeed spend most of its time near the specifically bound stationary foot.

V. RIGID CENTRAL HUB JOINT

Based on our discussions in the previous sections, we hypothesize that a rigid ankle may be more effective at reducing τ_{diff} in a system with high V_{nsb} if the central hub joint of the molecule is also made rigid. In this situation, the non-specifically bound feet are no longer able to slide along the DNA track. Thus, for the ankle joint to straighten, it must tear the non-specifically bound lagging foot from the track. But a rigid hub may unfortunately also remove the positive effects of non-specific binding. The reduction of τ_{diff} as a function of non-specific binding was due to the ability of the motor feet to slide along the DNA by facilitated 1D diffusion. With a hub that is too rigid, this motion is no longer allowed.

Figure 7 shows the first passage time of a motor with $V_{\text{hub}} = 100\ kT$ (see Eq. 6) for (a) fixed values of V_{ankle} and (b) fixed values of V_{nsb}. As expected, we no longer see a reduction in τ_{diff} as a function of V_{nsb}, but we still see an increase in τ_{diff} at high V_{nsb}. Although the feet cannot slide on the track, they can still be non-specifically bound to their binding site in the absence of their associated binding ligand. Figure 7(b) confirms that a motor with a rigid hub can more effectively mitigate high V_{nsb}: for $V_{\text{nsb}} = 16\ kT$, we see nearly a 50% reduction in τ_{diff} by adjusting V_{ankle} from 0 kT to 40 kT. But also notice that the overall scale of τ_{diff} has increased nearly an order of magnitude from the free-hub motor step time. Thus, although the rigid-hub motor is more effective at tearing the lagging foot from the track, the loss in facilitated 1D sliding diffusion of the motor feet ultimately makes the rigid-hub motor a poor choice for reducing the diffusional stepping time of the motor.

VI. LOAD FORCE

We would like to briefly explore the diffusional behavior of the motor under load force as a function of V_{nsb} and V_{ankle}. In prior studies of motor performance, the force behavior of molecular motors is characterized by the stall force. A useful definition of stall force is the amount of force applied to the motor such that the motor takes forward and backward steps at equal rates [23]. Due to the design of the TW system, though, backward motor steps are not realistically possible (except in a very contrived manner). We therefore focus on characterizing τ_{diff} as a function of applied load force. For a given τ_{diff}, the overall performance of the motor is then determined primarily by whether or not Eq. 1 is fulfilled.

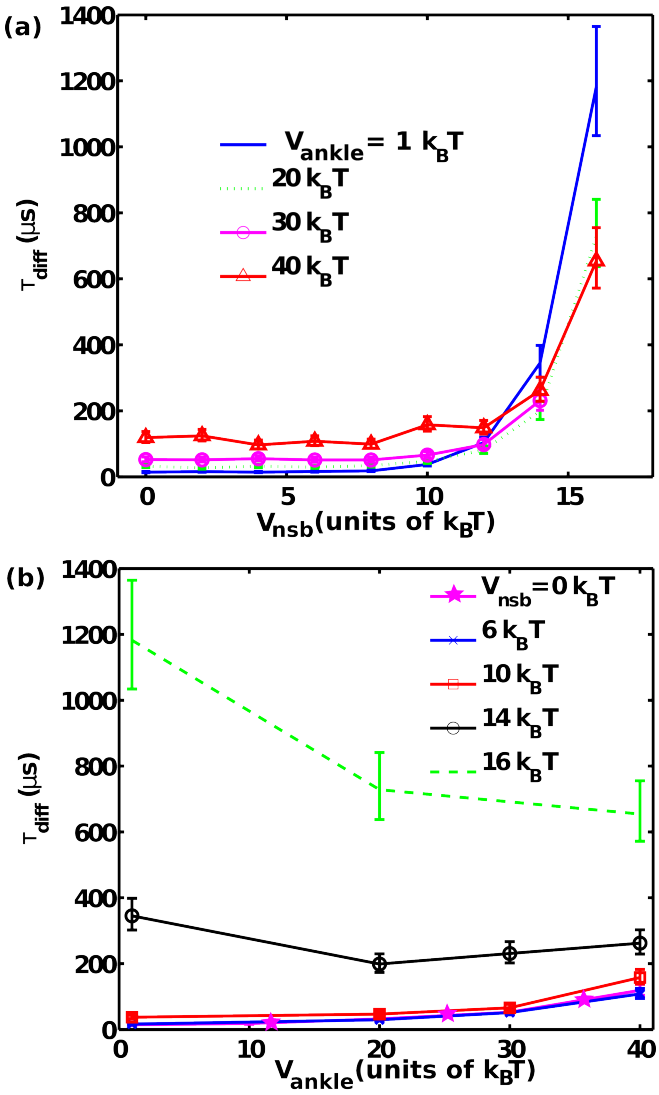

FIG. 7. (Color online) $\tau_{\rm diff}$ for a motor with a rigid central hub joint ($V_{\mathbf{hub}} = 100\ kT$) (a) as a function of $V_{\rm nsb}$, where different curves represent different fixed values of $V_{\rm ankle}$, and (b) as a function of $V_{\rm ankle}$, where different curves represent different fixed values of $V_{\rm nsb}$. As expected, tuning $V_{\rm nsb}$ no longer decreases $\tau_{\rm diff}$ because the motor feet cannot slide along the DNA track. Tuning $V_{\rm ankle}$ better mitigates the effects of high $V_{\rm nsb}$, but the overall scale of $\tau_{\rm diff}$ has also increased nearly an order of magnitude.

To obtain a first estimate of the reasonable range of load that the motor will be able to overcome, we consider that 1 kT = 4.14 pN nm at room temperature. Given that the motor steps purely diffusively, and assuming that it can overcome an energy barrier of several kT over its 11 nm step, the motor should be able to overcome up to about 1 pN

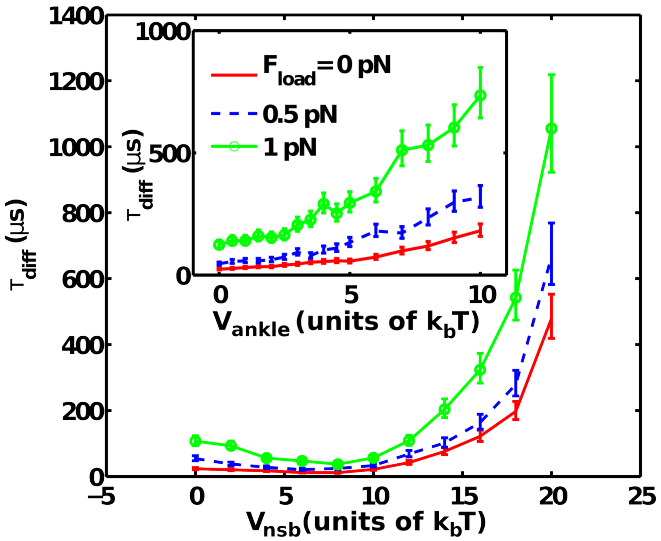

FIG. 8. (Color online) Diffusional step time as a function of $V_{\rm nsb}$ (with $V_{\rm ankle} = 0$ kT) and $V_{\rm ankle}$ (inset, with $V_{\rm nsb} = 0$ kT) for different values of load force from 0 - 1 pN. For both interactions, $\tau_{\rm diff}$ increases for all values of $V_{\rm ankle}$ and $V_{\rm nsb}$ with increasing load while the generic shape of each function remains the same.

($\sim$ 2.5 $kT/11$ nm) of load.

Figure 8 shows $\tau_{\rm diff}$ as a function of $V_{\rm nsb}$ and $V_{\rm ankle}$ for a constant horizontal load force on the central joint of the molecule between 0 and 1 pN. As one may expect, the diffusional stepping time of the motor appears to always increase with increasing load at fixed values of $V_{\rm nsb}$ and $V_{\rm ankle}$. However, the qualitative shape of the curves for fixed values of load force remain similar to the non-load curve, such that we can assume that the design strategies found so far apply also under load force. In particular, Fig. 8 clearly shows that optimizing the non-specific binding strength is highly beneficial for minimizing $\tau_{\rm diff}$ under load.

VII. LEG FLEXIBILITY

The original Tumbleweed design shown in Fig. 1 uses stiff (coiled-coil) legs. This design was introduced based on the intuitive expectation that a well-defined step size is important for TW performance, specifically to avoid backstepping (binding to a rearward rather than a forward binding site). However, our results so far indicate that, overall, high molecular flexibility is of advantage for motor performance. We therefore ask: would a Tumbleweed

with flexible legs formed by polypeptide chains perform better than the original design with rigid, coiled-coil legs?

In order to answer this question, we explore TW motor stepping in terms of flexible rather than rigid legs. We adjust the model of Section II by replacing the y-shaped motor hub with three **self-avoiding chains** (SAC; see inset of Fig. 9). The SAC model is a basic representation of a polymer, where the length of the chain segments ($r_{\rm chain}$ in Fig. 9) is twice the persistence length. Each SAC has $textrmN_p$ segments of equal length, $r_{\rm chain}$, plus an end segment of length $d_{\rm rep} = 2.5$ nm, representing the repressor **(the extra 0.5 nm difference from ζ (Sec. II) is included to correctly model the excluded volume of the chain segments)**. The segments are modeled by a harmonic tethering potential (Eq. 3) and the interactions between them are given by a Lennard-Jones potential (Eq. 4).

As a first step, we keep the polymer's contour length constant and equal to the length of the rigid legs in the model of Section II. The length of each equal segment is given by $r_{\rm chain} = (r_{\rm leg}$ - $d_{\rm rep})/{\rm N_p}$ and the value of $\rm N_p$ is increased from $\rm N_p = 1$ to $\rm N_p = 4$. As $\rm N_p$ increases, the corresponding segment length decreases and the inset in Fig. 9 shows that this increasing flexibility results in an increase in $\tau_{\rm diff}$ by more than an order of magnitude, because the legs can fold onto themselves. This is due to an increase in the entropy of the legs which makes it more difficult for them to reach the next binding site, given that the contour length is independent of $\rm N_p$.

As a second step, we again explore motor stepping but now in terms of a variable contour length of the SAC. For compatibility with the case of fixed contour length in the previous paragraph, we choose the value of $r_{\rm chain}$ to be equal to that for N = 4 above, i.e. $r_{\rm chain}=$ 0.9625 nm (comparable to twice the persistence length of polypeptide chains [24, 25]) and we use a steric length scale $\zeta_{\rm chain} = r_{\rm chain}$ for each monomer. The contour length of the resulting SAC, including the repressor segment, is then given by $({\rm N}r_{\rm chain} + d_{\rm rep})$. Fig. 9 shows $\tau_{\rm diff}$ as a function of N.

As shown in Fig. 9, added contour length can bring the value of $\tau_{\rm diff}$ back down to about 30 μs (for N = 8), that is, to values comparable to those observed with a rigid coiled-coil leg. For even higher N (and larger contour length), $\tau_{\rm diff}$ slowly increases. The initial decrease of $\tau_{\rm diff}$ is due to the increased contour length compensating for the fact that flexible legs fold onto themselves, allowing the flexible legs to bind more easily. However, as N increases even further, the configuration space that is explored diffusively becomes larger, and $\tau_{\rm diff}$

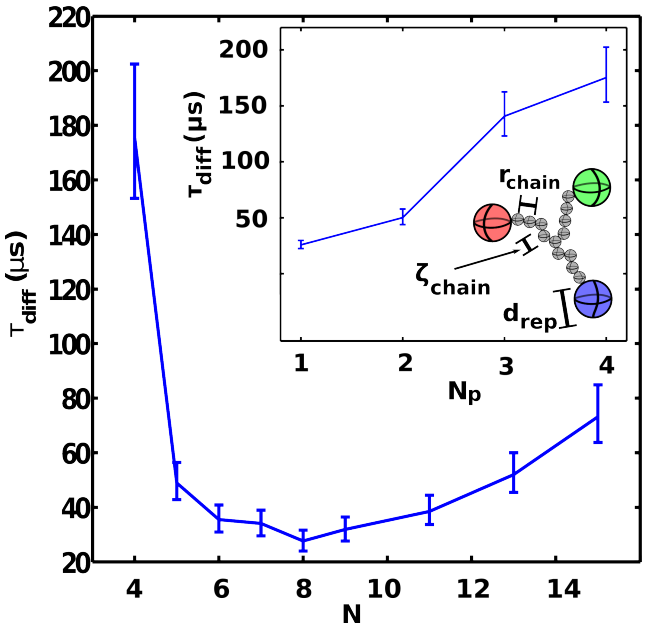

FIG. 9. (Color online) A possible motor design replaces the three rigid coiled-coil hub arms with flexible polypeptide chains (schematically shown in inset). We adjust our coarse-grained model by replacing the central monomer with freely-jointed chains built from spherical monomers of size $\zeta_{\text{chain}} = 1$ nm. (Inset) One approach is to keep the contour length of the TW legs constant and increase the number of polymer segments (N_p). For this approach, τ_{diff} increases monotonically with increasing N_p. Another approach is to keep sphere separation $r_{\text{chain}} \sim 1$ nm constant and increase the number of monomers (N). This approach recovers the stiff-leg τ_{diff} for $N \sim 8$.

increases. It is interesting to note that backstepping (binding to a rearward binding site) which becomes sterically possible for N = 11, is not actually observed in simulations until N = 15.

Whereas simulating the relevant physical details of an actual peptide chain is beyond the capabilities of the coarse-grained model presented here, we can nevertheless conclude that added leg flexibility does not give any clear advantages in stepping behavior over a rigid, coiled-coil leg. Given that a long, flexible polypeptide chain may be very difficult to design, and may form unintended, stable secondary structures on its own together with the repressor feet, a coiled-coil design remains a preferable design choice.

VIII. DISCUSSION

The starting point for this study was the finding of an earlier Master equation study [13] that TW's diffusive stepping time $\tau_{\rm diff}$ should be kept at least three orders of magnitude below the other characteristic times in Eq. 1, namely $\tau_{\rm ligand}$ (to avoid misstepping) and $\tau_{\rm bound}$ (to avoid unbinding from the track) [12]. Given that, for perfect stepping, the motor speed is given by $\mathrm{x}_L/\tau_{\rm ligand}$, and it may be possible to operate TW with $\tau_{\rm ligand}$ as short as 0.1 s, it is desirable to keep $\tau_{\rm diff}$ as short as possible, and below about 100 μs.

The main findings from the present study are then as follows: (i) $\tau_{\rm diff}$ is minimized by keeping TW's hub and ankle joint flexible ($V_{\rm hub}$ and $V_{\rm ankle}$ should be chosen as small as possible); (ii) Non-specific repressor-DNA binding reduces $\tau_{\rm diff}$, in particular in the presence of finite $V_{\rm ankle}$, up to an optimal value of $V_{\rm nsb} \sim\ 8\ kT$; and (iii) additional flexibility of the legs does not appear to carry any advantages in the present design, such that coiled-coil legs continue to be a good choice.

Given the practical challenges involved in constructing a novel multi-subunit protein, it is likely that molecular flexibility will be difficult to accurately predetermine. Therefore, the role of non-specific binding that emerges from this study is a very important tuning parameter. The precise strength of NSB of TW's repressor protein feet is not known, but there is evidence that the non-specific protein-DNA binding is highly dependent on ionic strength. Revzin and von Hippel determined association constants for the *Escherichia coli* lac repressor protein to non-operator sections of DNA as a function of ionic strength, and reported more than an order of magnitude increase in the association constant for only a 25% reduction in ionic concentration [26]. Therefore, it may be realistic to experimentally tune $V_{\rm nsb}$ over at least part of the optimal range of $V_{\rm nsb} = \ 5 - 12\ kT$ (Figs. 3 and 5).

In this context, the details of the NSB model used here deserve further discussion. In our model we assumed that a non-specifically bound protein can linearly diffuse along DNA, and that the track is designed such that all of the repressor-binding sites face in the same direction. However, recent studies suggest that certain non-specifically bound DNA-binding proteins diffuse along the groove of the DNA in helical manner [19]. If this should be true also for TW's repressor protein feet, we do not expect the behavior shown in Fig. 5 to remain. However, based on our simulation results, we can speculate how helical diffusion might change TW's behavior. Due to the structural rigidity of the coiled-coil protein hub, a

foot would likely not be able to slide very far along the DNA with another foot specifically bound. The resulting behavior would be qualitatively similar to the results shown in Fig. 7, where the central hub joint is rigid, and where non-specific binding increases τ_{diff}. Based on these results, in the presence of helical diffusion of repressor feet along the DNA, it may be beneficial to give the ankle joint some rigidity and to reduce the non-specific interaction as much as possible. However, it is likely that leg flexibility may help to restore some of the positive effects of NSB in the presence of helical diffusion, and the role of leg flexibility should be revisited in the context of a future TW model that incorporates different modes of 3D diffusion along DNA.

One of the goals of our project to design and build an artificial protein motor is to look for common design principles that may also apply to biological motors, for example bi-pedal stepping motors such as kinesin, myosin V, and dynein. In fact, the positive effect of non-specific binding (a short diffusional step time and increased run length) of the TW model used here hints at a possible connection. It has been shown that certain dyneins [27], kinesins [28, 29], and kinesin-related motor proteins [30] exhibit a similar 1D diffusion along their microtubule tracks in the absence of ATP, providing evidence of a weakly-bound state similar to the non-specific protein-DNA interaction. This interaction has been hypothesized to generally help the motors stay attached to the track. Whereas a detailed study of stepping behavior including the weakly-bound state has to our knowledge not yet been performed [23], it is nevertheless interesting to speculate that its beneficial role for reduced stepping time and increased run length may be of universal nature in diffusive stepping motors.

Another curious similarity between an optimized TW design and at least some biological, bi-pedal motors is the interplay of rigid elements and flexible joints. For example, a coarse-grained model for myosin V appears to share several features with the optimized TW, namely a highly flexible neck joint (corresponding to V_{hub})[31–34] and well defined legs with limited flexibility. These features were recently strikingly confirmed by high speed AFM imaging of myosin V stepping [35].

Acknowledgements. We thank Gerhard Blab, Michael Plischke, and the entire HFSP Motor Collaboration for useful discussions. We thank the University of Bristol for the provision of HPC time. Computational resources were also provided by Westgrid. N.R.F. acknowledges support from the Michael Smith Foundation for Health Research. N.R.F. and

M.J.Z. wish to thank NSERC for support. This work was financially supported by the Human Frontier Science Program (RGP0031/2007), by the National Science Foundation under Grant Nos. DGE-0742540 and DGE-0549503, by the Swedish Research Council, and by nmC@LU.

---

[1] B. L. Feringa, Journal of Organic Chemistry, **72**, 6635 (2007).

[2] E. R. Kay, D. A. Leigh, and F. Zerbetto, Angewandte Chemie-International Edition, **46**, 72 (2007).

[3] V. Balzani, A. Credi, S. Silvi, and M. Venturi, Chemical Society Reviews, **35**, 1135 (2006).

[4] G. S. Kottas, L. I. Clarke, D. Horinek, and J. Michl, Chemical Reviews, **105**, 1281 (2005).

[5] J. Sauvage and C. Dietrich-Buchecker, eds., *Molecular catenanes, rotaxanes and knots: a journey through the world of molecular topology* (Wiley-VCH, Weinheim, 1999).

[6] T. J. Huang, B. Brough, C. M. Ho, Y. Liu, A. H. Flood, P. A. Bonvallet, H. R. Tseng, J. F. Stoddart, M. Baller, and S. Magonov, Applied Physics Letters, **85**, 5391 (2004).

[7] J. Bath and A. J. Turberfield, Nature Nanotechnology, **2**, 275 (2007).

[8] R. A. Muscat, J. Bath, and A. J. Turberfield, Nano Letters, **11** (3) (2011).

[9] R. Pei, S. K. Taylor, D. Stefanovic, S. Rudchenko, T. E. Mitchell, and M. N. Stojanovic, Journal of the American Chemical Society, **128**, 12693 (2006).

[10] H. Gu, J. Chao, S. Xiao, and N. Seeman, Nature, **465**, 202 (2010).

[11] E. H. C. Bromley, N. J. Kuwada, M. J. Zuckermann, R. Donadini, L. Samii, G. A. Blab, G. J. Gemmen, B. J. Lopez, P. M. G. Curmi, N. R. Forde, D. N. Woolfson, and H. Linke, HFSP Journal, **3**, 204 (2009).

[12] P. B. Allen, G. Milne, B. R. Doepker, and D. T. Chiu, Lab Chip, **10**, 727-733 (2010).

[13] N. J. Kuwada, G. A. Blab, and H. Linke, Chemical Physics, **375** (2010).

[14] *See Supplemental Material at [URL will be inserted by publisher] for further details and discussion of the all-atom molecular dynamics simulation.*

[15] M. Allen and D. Tildesley, *Computer simulation of liquids* (Oxford University Press, USA, 1990).

[16] S. Kumar, J. Rosenberg, D. Bouzida, R. Swendsen, and P. Kollman, Journal of Computational Chemistry, **13**, 1011 (1992).

[17] J. Gorman and E. Greene, Nature Structural & Molecular Biology, **15**, 768 (2008).

[18] S. E. Halford, Biochemical Society Transactions, **37**, 343 (2009).

[19] P. Blainey, G. Luo, S. Kou, W. Mangel, G. Verdine, B. Bagchi, and X. Xie, Nature Structural & Molecular Biology (2009).

[20] J. Schurr, Biophysical chemistry, **9**, 413 (1979).

[21] B. Bagchi, P. Blainey, and X. Xie, J. Phys. Chem. B, **112**, 6282 (2008).

[22] V. Dahirel, F. Paillusson, M. Jardat, M. Barbi, and J. M. Victor, Physical Review Letters, **102**, 228101 (2009).

[23] N. Carter and R. Cross, Nature, **435**, 308 (2005).

[24] F. Huang, Angewandte Chemie-International Edition, **42** (2004).

[25] V. Singh and L. Lapidus, The Journal of Physical Chemistry B, **112**, 13172 (2008).

[26] A. Revzin and P. Von Hippel, Biochemistry, **16**, 4769 (1977).

[27] R. Vale, D. Soll, and I. Gibbons, Cell, **59**, 915 (1989).

[28] J. Helenius, G. Brouhard, Y. Kalaidzidis, S. Diez, and J. Howard, Nature, **441**, 115 (2006).

[29] J. Cooper, M. Wagenbach, C. Asbury, and L. Wordeman, Nature Structural & Molecular Biology (2009).

[30] R. Chandra, S. Endow, and E. Salmon, Journal of Cell Science, **104**, 899 (1993).

[31] E. M. Craig and H. Linke, Proceedings of the National Academy of Sciences of the United States of America, **106**, 18261 (2009).

[32] A. Dunn and J. Spudich, Nature Structural & Molecular Biology, **14**, 246 (2007).

[33] K. Shiroguchi and K. Kinosita Jr, Science, **316**, 1208 (2007).

[34] S. Burgess, M. Walker, F. Wang, J. Sellers, H. White, P. Knight, and J. Trinick, The Journal of cell biology, **159**, 983 (2002).

[35] N. Kodera, D. Yamamoto, R. Ishikawa, and T. Ando, Nature, **468**, 72 (2010).