The Nanocaterpillar's Random Walk: Diffusion With Ligand-Receptor Contacts

Journal:	*Soft Matter*
Manuscript ID	SM-ART-10-2021-001544.R1
Article Type:	Paper
Date Submitted by the Author:	21-Feb-2022
Complete List of Authors:	Marbach, Sophie; New York University Courant Institute of Mathematical Sciences, Math; CNRS, Sorbonne Universite, Physicochimie des Electrolytes et Nanosystemes Interfaciaux Zheng, Jeana; New York University, Department of Physics Holmes-Cerfon, Miranda; New York University Courant Institute of Mathematical Sciences, Mathematics

Soft Matter

ARTICLE TYPE

The Nanocaterpillar's Random Walk: Diffusion With Ligand-Receptor Contacts[†]

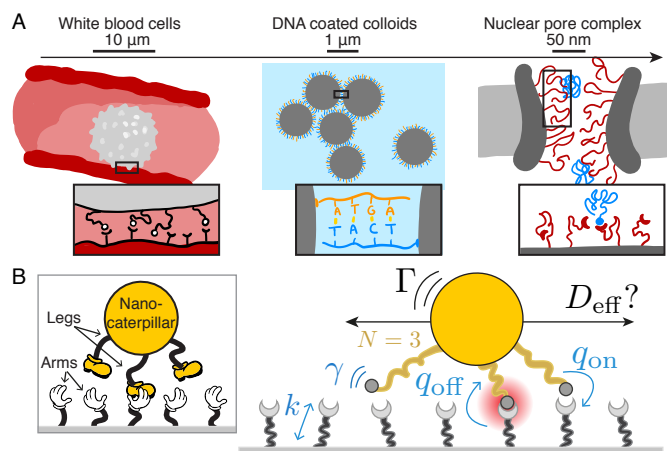Sophie Marbach,[*a,b] Jeana Aojie Zheng,[c] and Miranda Holmes-Cerfon[a]

Particles with ligand-receptor contacts bind and unbind fluctuating "legs" to surfaces, whose fluctuations cause the particle to diffuse. Quantifying the diffusion of such "nanoscale caterpillars" is a challenge, since binding events often occur on very short time and length scales. Here we derive an analytical formula, validated by simulations, for the long time translational diffusion coefficient of an overdamped nanocaterpillar, under a range of modeling assumptions. We demonstrate that the effective diffusion coefficient, which depends on the microscopic parameters governing the legs, can be orders of magnitude smaller than the background diffusion coefficient. Furthermore it varies rapidly with temperature, and reproduces the striking variations seen in existing data and our own measurements of the diffusion of DNA-coated colloids. Our model gives insight into the mechanism of motion, and allows us to ask: when does a nanocaterpillar prefer to move by *sliding*, where one leg is always linked to the surface, and when does it prefer to move by *hopping*, which requires all legs to unbind simultaneously? We compare a range of systems (viruses, molecular motors, white blood cells, protein cargos in the nuclear pore complex, bacteria such as Escherichia coli, and DNA-coated colloids) and present guidelines to control the mode of motion for materials design.

Particles with ligand-receptor contacts – or *nanocaterpillars* – harvest binding and unbinding dynamics of their fluctuating *legs* at the nanoscale to move, target, stick, or assemble into large structures[1–4]. Nanocaterpillars are found across multiple scales, spanning a great variety of systems in biology and biomimetic assays – see Fig. 1-A. To name but a few, microscale white blood cells with protein linkers stick and roll on blood vessel walls until they reach a healing target[5–7]. Microscale droplets with protein linkers are used to study cellular-like adhesion[8–10]. Microscale to nanoscale colloids coated with complementary deoxyribonucleic acid (DNA) strands self-assemble into macroscopic crystals[4,11,12] with novel optical or selectivity properties[13–16]. Nanoscale viruses transiently adhere with spike proteins to the respiratory mucus to find vulnerable host cells[1,17–19]. At even smaller scales, protein cargos bind to receptors in the nuclear pore complex for selective transport to a cell's nucleus[20,21].

For all these systems to function, a nanocaterpillar must *move* relative to the surface to which its legs are attracted. An important question therefore is to characterize *how* it moves, over scales

* E-mail: sophie@marbach.fr   [a] Courant Institute of Mathematical Sciences, New York University, NY, 10012, U.S.A.

[b] CNRS, Sorbonne Université, Physicochimie des Electrolytes et Nanosystèmes Interfaciaux, F-75005 Paris, France.

[c] Department of Physics, New York University, NY, 10012, U.S.A.

Fig. 1 Overview of nanocaterpillars. (A) Multivalent ligand-receptor systems span the micro to nanoscales. White blood cells stick to vessel walls through selectin mediated bonds (inspired from Ref. 7); DNA-coated colloids self-assemble through hybridization of complementary DNA strands; Protein cargos translocate through the polymer mesh of the nuclear pore complex (inspired from Ref. 22). (B) Ligand-receptor systems are modeled here with an arbitrary number of legs N (ligands) and/or arms (receptors). The stochastic model includes binding and unbinding rates q_{on} and q_{off}, spring constant k, and leg friction γ (all fast, in blue); and the bare friction coefficient Γ of the nanocaterpillar (slow, in black). We seek the long-time effective longitudinal diffusion coefficient D_{eff}.

much larger than individual legs. Since legs constantly bind and unbind to the surface, imparting force each time they do so, the particle's macroscopic mobility depends on the microscopic details of its legs. For example, leg flexibility and bond lifetimes control the average mobility of the particle [19,23,24], and differences in both parameters can be harvested to detect infected cells [25–27] or prevent viral infections [28]. As another example, leg density affects how DNA-coated colloids nucleate and grow into crystals [29,30] and governs the long-range alignment of crystals [31–33]. Overall, microscopic details underlie a variety of large-scale modes of motion, such as hopping [3,17,34,35], cohesive motion including rolling and crawling [17,36], and also transient or firm arrest [3,5,37], resulting in large differences in macroscopic mobility.

Investigating how microscopic binding details lead to macroscopic mobility is challenging, as it requires probing time and length scales that can often be quite different [19,38] – legs can be much smaller than the nanocaterpillar they are attached to, while leg dynamics can be orders of magnitude faster than the timescales of macroscopic motion. Furthermore, many systems have a valency of thousands of leg contacts [31,38,39], too many degrees of freedom to resolve experimentally or computationally [22,40]. To make progress, numerical and analytical models often rely on simplified assumptions, *e.g.* excluding stochastic relaxation of the legs [41,42], limiting the analysis to a small number of legs [41,43,44], or assuming small perturbations [22]. Such models have given insight into a variety of phenomena, such as how specific parameters could favor rolling over sliding [7,41,43,45,46] or how specific mechanisms could increase overall mobility (with coupling effects such as binding dynamics depending on bond number [47–49] or when numerous adhesive sites are available for a single ligand [22,50,51]). Nevertheless, such modeling assumptions are not always justified; for example stochasticity plays a critical role for mobility, facilitating rolling [37], targeted arrest [40], or other walking modes [52]. Furthermore, such models can also not reproduce the order of magnitude decrease of diffusion of DNA-coated colloids [31,39]. Hence, a systematic derivation of macroscopic mobility from microscopic details that is valid under a broad range of parameters is needed.

In this paper we derive an analytical expression for the effective mobility of a nanocaterpillar in an overdamped system, by systematically coarse-graining over the microscopic details of its legs. Starting from a model that includes the detailed spatial fluctuations of the legs, we use homogenization techniques [22,53,54] to average over these fluctuations. We obtain an analytical expression for the effective long-time translational diffusion coefficient of the particle, $D_{\text{eff}}(N, \Gamma, \gamma, k, q_{\text{off}}, q_{\text{on}})$, as a function of the microscopic parameters governing the legs (Eq. (15); see also Fig. 1-B and Sec. 1.) The expression depends in a non-trivial way on the friction coefficients of the individual components of the system (legs and particle), with the frictions either adding up arithmetically (like springs in parallel) or harmonically (like springs in series) according to the mechanistic details. We validate our analytical calculations with numerical simulations, which show the expression is accurate over a wide range of parameter values.

Our model gives insight into the mechanism of nanocaterpillar motion, as it allows us to distinguish between two long term

modes of motion: *sliding*, where at least one bond is always attached to the surface, and *hopping*, where the particle detaches completely, moves in free space and reattaches. These regimes are controlled by physical properties of the legs, such as stiffness and adhesive strength, allowing us to investigate existing biological and biomimetic systems in a so-called Ashby chart for nanocaterpillars (Sec. 2). We identify how critical design parameters (such as the coating density for DNA-coated colloids) controls the preferential mode of motion and reconcile disparate experimental observations on similar systems [31,39].

Importantly, the effective diffusion can sometimes be orders of magnitude smaller than the background diffusion coefficient, showing the critical effect of the legs on the particle's mobility. This analytical prediction of a dramatically decreased diffusivity is borne out with experimental measurements of the diffusion of DNA-coated colloids, both from existing data [31,39] and additionally measured in this study. Our model agrees with the data within experimental accuracy over a range of temperatures and for different DNA coating densities on the colloids (Sec. 2).

Finally, we derive the effective diffusion coefficient for several variations of the model with varying assumptions, and show that our model incorporates these assumptions as special limits [22,54], but is accurate over a broader range of parameters and system designs (Sec. 3). In particular, previous approaches can not describe the observed orders of magnitude decrease in diffusion [22]. Overall, our results lay the ground to tune mobility features in artificial designs, and provide methodological tools to study more complex motion mediated through ligand-receptors, including rolling or self-avoiding walks due to active cutting of bonds.

1 Deriving an analytical formula for the effective diffusion coefficient

In Sections 1.1-1.3 we illustrate our homogenization technique pedagogically by considering a 1-legged caterpillar. Our main result for the effective diffusion coefficient of an N-legged caterpillar, Eq. (15), is presented in Section 1.4.

1.1   1-legged caterpillar: constitutive equations

We begin with the simplest possible model: a nanocaterpillar with a single leg (Fig. 2). The leg is permanently fixed to the caterpillar while its other end is mobile, and can attach anywhere on the binding surface. We consider for now a one-dimensional model, where leg fluctuations and particle motion occur on a line, longitudinal to the surface.

The dynamics of the particle position $x(t)$ and leg length $l(t)$ occur over nano to microscales, mostly in dense fluids such as water. In this context, dynamics are well captured by overdamped Langevin equations [55], where inertia plays a negligible role. This is in contrast to previous modeling efforts which used the Langevin equation (with inertia) [54], a point we return to in Sec. 3, where we show that the two approaches can give predictions that are orders of magnitude different in certain parameter regimes.
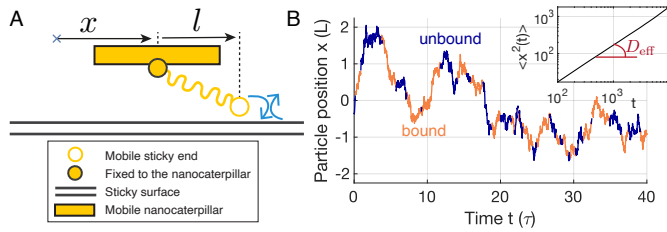
Fig. 2 1-legged nanocaterpillar model. (A) The longitudinal extension of the single leg (l) is monitored and feeds back into the longitudinal position (x) of the particle. (B) Simulation trace of the position of a 1-legged particle with time. (inset) The effective long time diffusion D_{eff} is half the slope of the mean squared displacement over long times.

When the legs are unbound they evolve as

$$\frac{dl}{dt} = -\frac{k}{\gamma}(l(t) - l_0) + \sqrt{\frac{2k_BT}{\gamma}}\eta_l(t).\qquad (1)$$

Here k is a spring constant describing the recoil force of the leg material, γ is its friction coefficient, l_0 its rest length, k_B is Boltzmann's constant, T is temperature and η_l is a Gaussian white noise satisfying $\overline{\eta_l(t)} = 0$ and $\overline{\eta_l(t)\eta_l(t')} = \delta(t-t')$ where $\overline{\cdot}$ is the average over realizations of the noise. In most systems we consider, legs are made of polymers or proteins, where small leg deformations around equilibrium are well captured by a constant spring constant k[56–58].

The particle's position x when the leg is unbound obeys

$$\frac{dx}{dt} = \sqrt{\frac{2k_BT}{\Gamma}}\eta_x(t)\qquad (2)$$

where Γ is the friction coefficient of the particle and $\eta_x(t)$ is a Gaussian white noise uncorrelated with $\eta_l(t)$. The diffusion coefficient for the unbound particle is $D_0 = \frac{k_BT}{\Gamma}$.

We consider for now that the surface is uniformly coated with receptors. The leg can thus bind at any location on the surface with a constant binding rate q_{on} and constant unbinding rate q_{off}. Detailed balance requires $\frac{q_{\text{on}}}{q_{\text{off}}} = \frac{\pi_b}{\pi_u}$ where $\pi_{b/u}$ is the equilibrium probability of the system to be bound or unbound. Typically $\frac{\pi_b}{\pi_u} = e^{-\beta\Delta G}$, where $\beta^{-1} = k_BT$ and $\Delta G < 0$ is the free energy change when the leg binds to the surface[38,59].

We now seek to describe motion of the system when the leg is bound. In this case, variables are constrained as $x(t) + l(t) - x_r = 0$ where x_r is the location of the receptor where the leg tip is attached, which is constant until the leg detaches and reattaches to another location. The stochastic dynamics Eqns. (1) and (2) must be projected[32,60] onto the constraint surface, see Appendix A. We obtain

$$\frac{dx}{dt} = -\frac{dl}{dt} = \frac{k}{\Gamma+\gamma}(l(t)-l_0) + \sqrt{\frac{2k_BT}{\Gamma+\gamma}}\eta(t)\qquad (3)$$

where $\eta(t)$ is a Gaussian white noise. Here we see that the projected dynamics have a natural expression where the effective friction in the bound state is the arithmetic sum of the friction coefficients in the unbound states, $\Gamma+\gamma$. Note that this projection is a crucial step that is often ignored in such derivations[22,32,54], and modifies the dynamics in non trivial ways especially with a

large number of legs.

The dynamics are now specified through the set of Eqns. (1)-(3), together with the binding and unbinding dynamics. To see what happens over long times, we simulate trajectories for 1 leg – see Fig. 2-B (and simulation details in Appendix B). Over long times, the particle's mean-squared displacement grows linearly with time, and we may extract an effective long time diffusion coefficient D_{eff} – see inset of Fig. 2-B.

1.2 Homogenization to coarse-grain the fast dynamics

The computational cost of simulating Eqns. (1)-(3) is high, since small time steps are required to resolve the fast relaxation and binding events. We therefore seek an analytical method to coarse-grain over these fast timescales. To apply this method we identify a non-dimensional separation of scales, which is novel compared to other approaches[22,51,54] and will allow us to find a result valid over a broad range of parameters. We use homogenization theory to average over the fast scales, eventually obtaining an effective diffusion equation, Eq. (10), with effective diffusivity (Eq. (11)) and related effective friction (Eq. (12)), which is one of the main results of this paper for the special case of a 1-legged caterpillar. A reader interested in the results and physical implications may skip to Section 1.3.

1.2.1 Set up: partial differential equations to be coarse-grained

The set of stochastic Eqns. (1)-(3) defines a Markov process that is conveniently studied via the Fokker-Planck equation and its adjoint, the Kolmogorov backward equation[53,61]. Let $p(x,l,t) = (p_u(x,l,t), p_b(x,l,t))^T$ be the probability density function of finding the system at time t and positions x, l in the unbound or bound states. We obtain from Eqns. (1)-(3) the Fokker-Plank equation

$$\partial_t p = \mathscr{L}^{\star} p,\qquad (4)$$

with $\mathscr{L}^{\star} = \mathscr{V}^{\star} + \mathscr{Q}^{\star}$ where

$$\mathscr{V}^{\star} = \text{diag}\begin{pmatrix} \partial_l\left(\frac{k}{\gamma}(l-l_0) + \frac{k_BT}{\gamma}\partial_l\right) + \frac{k_BT}{\Gamma}\partial_{xx} \\ (\partial_l - \partial_x)\left(\frac{k}{\Gamma+\gamma}(l-l_0) + \frac{k_BT}{\Gamma+\gamma}(\partial_l - \partial_x)\right) \end{pmatrix},$$

$$\mathscr{Q}^{\star} = \begin{pmatrix} -q_{\text{on}} & q_{\text{off}} \\ q_{\text{on}} & -q_{\text{off}} \end{pmatrix},$$

with an appropriate initial condition. Additionally we require the flux in either state to vanish at infinity, to conserve total probability. The stationary solution of Eq. (4) is $\pi = \frac{e^{-\beta k(l-l_0)^2/2}}{Z}(q_{\text{off}}, q_{\text{on}})^T$ where Z is a normalization constant. This is therefore the equilibrium probability density of the system; it satisfies detailed balance.

While probability densities have an intuitive physical meaning, in the following it will be easier – and mathematically better posed – to consider the adjoint of the Fokker-Planck equation and the corresponding dual functions. These are functions $f(x,l,t) = \int p(x',l',t|x,l)g(x',l')dl'dx'$ that give the expectation of any scalar function $g(x(t),l(t))$, given an initial condition $x(0) = x, l(0) = l$. Once we know how such functions f evolve,

we may calculate any statistic g of our stochastic process. Writing $f(x,l,t) = (f_u(x,l,t), f_b(x,l,t))^T$, we have that f satisfies the Kolmogorov backward equation[61]

$$\partial_t f = \mathscr{L} f, \quad f(x,l,0) = g(x,l). \tag{5}$$

Here $\mathscr{L}$ is the adjoint operator of $\mathscr{L}^\star$, defined by the operator that satisfies $\langle f, \mathscr{L}^\star p \rangle = \langle \mathscr{L} f, p \rangle$ for any probability density p and statistic f, where $\langle f, p \rangle = \iint (f_u p_u + f_b p_b) dl dx$ is the inner product.

1.2.2 Non-dimensionalization and assumptions on scales.

We now seek to coarse-grain the fast dynamics, by applying homogenization techniques to the backward equation, Eq. (5). To start, we non-dimensionalize the equation using

$$x \to L_x \tilde{x}, \ l - l_0 \to L \tilde{l}, \ t \to \tau \tilde{t},$$

where $L = \sqrt{k_B T / k}$ is the reference length of the leg fluctuations, L_x is the scale for the long-time average motion of x, and τ is the timescale associated with this average motion. The latter two scales are not determined *a priori* by any intrinsic scales in the system, but rather are chosen large enough that averaging will be appropriate over such scales; hence we choose $L_x = L/\varepsilon$ where $\varepsilon \ll 1$ is a small non-dimensional number. We are interested in long time scales corresponding to the diffusion of the particle, hence we expect $\tau = L_x^2/D_0$, which corresponds to $\tau = \frac{1}{\varepsilon^2} \frac{\Gamma}{k}$. Importantly, and in contrast with other works[22,51], here ε does not measure the value of physical parameters, but rather, it measures the large observation time scale over which the coarse-grained model is valid. Such long observation times are quite likely in experiments, as typical binding rates and leg dynamics occur at most over 1 ms $-$ 1 s while observation (or other biophysical processes such as internalisation for viruses[17]) happens over the course of 10 min at least[38]. This non-dimensionalization step is crucial as it will allow us to find order of magnitude changes in the diffusion coefficient according to the physical parameters, something that was not captured by previous perturbative approaches[22,51].

We now assume that the observation time scale is long enough, such that binding and unbinding events, as well as relaxation dynamics, will both occur on comparably short time scales. We can therefore write $\tilde{q}_i = q_i \Gamma/k = O_\varepsilon(1)$ and $\gamma/\Gamma = O_\varepsilon(1)$. In Sec. 3 we will see that taking different limits for these physical parameters (such as $\gamma/\Gamma \ll 1$) yields the same result as applying these limits to the final result. Our choices of scalings are therefore quite general and can be easily adapted to more detailed systems.

Using non-dimensional variables (and dropping the $\tilde{}$ for simplicity) we obtain from the backward equation Eq. (5) a separation in orders of ε as

$$\partial_t f = \mathscr{L} f = \left(\frac{1}{\varepsilon^2} \mathscr{L}_0 + \frac{1}{\varepsilon} \mathscr{L}_1 + \mathscr{L}_2 \right) f \tag{6}$$

where

$$\mathscr{L}_0 = \begin{pmatrix} -q_{on} + \frac{\Gamma}{\gamma}(-l\partial_l + \partial_{ll}) & q_{on} \\ q_{off} & -q_{off} + \frac{\Gamma}{\Gamma+\gamma}(-l\partial_l + \partial_{ll}) \end{pmatrix},$$

$$\mathscr{L}_1 = \text{diag}\left(0, \frac{\Gamma}{\Gamma+\gamma}(l\partial_x - 2\partial_{lx}) \right),$$

$$\mathscr{L}_2 = \text{diag}\left(\partial_{xx}, \frac{\Gamma}{\Gamma+\gamma}\partial_{xx} \right).$$

1.2.3 Homogenization method.

We seek a solution to Eq. (6) of the form $f = f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \dots$. We obtain a hierarchy of equations at different orders in ε:

$$O_\varepsilon\left(\frac{1}{\varepsilon^2} \right): \qquad \mathscr{L}_0 f_0 = 0, \tag{7}$$

$$O_\varepsilon\left(\frac{1}{\varepsilon} \right): \qquad \mathscr{L}_0 f_1 = -\mathscr{L}_1 f_0, \tag{8}$$

$$O_\varepsilon\left(1 \right): \qquad \mathscr{L}_0 f_2 = \partial_t f_0 - \mathscr{L}_1 f_1 - \mathscr{L}_2 f_0, \tag{9}$$

$$\vdots \qquad\qquad \vdots$$

and we solve these iteratively for f at each order in ε. At lowest order we obtain from Eq. (7) and the vanishing flux at boundaries, $f_0 = a(x,t)\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, where $a(x,t)$ is an unknown function of the slow variable x, whose dynamics we seek to determine. The associated equilibrium distribution at lowest order, $\mathscr{L}_0^\star \pi_0 = 0$ is simply the full one $\pi_0 = \pi$.

At the next order, one can check that

$$f_1 = \begin{pmatrix} \gamma q_{on} \\ \Gamma + \gamma q_{on} \end{pmatrix} \frac{l\partial_x a}{\Gamma(1 + q_{off}) + \gamma(q_{on} + q_{off})}$$

is a particular integral of Eq. (8), and is the unique solution since we impose that f_1 does not contain terms in the nullspace of $\mathscr{L}_0$.

Finally Eq. (9) possesses a solution if and only if it satisfies the Fredholm alternative[53]

$$\langle (\partial_t f_0 - \mathscr{L}_1 f_1 - \mathscr{L}_2 f_0), \pi_0 \rangle = 0.$$

Standard algebra yields an effective long time diffusion equation for a (in dimensional variables)

$$\partial_t a = D_{eff} \partial_{xx} a, \tag{10}$$

where

$$D_{eff} = \frac{k_B T}{\Gamma_{eff}}, \tag{11}$$

with

$$\frac{1}{\Gamma_{eff}} = \frac{p_0}{\Gamma_0} + \frac{p_1}{\Gamma_1}, \quad \text{with } \Gamma_0 = \Gamma, \ \Gamma_1 = \Gamma + \gamma_{eff}$$

$$\text{and } \gamma_{eff} = \gamma + k\left(\frac{1}{q_{off}} + \frac{\gamma}{k}\frac{q_{on}}{q_{off}} \right). \tag{12}$$

In the above expressions, $p_0 = \frac{q_{off}}{q_{off}+q_{on}}$ is the equilibrium probability to have no bond, and $p_1 = 1 - p_0$ the equilibrium probability

to have one bond. $\Gamma_0 = \Gamma$ is the friction in the unbound state and Γ_1 is the effective friction contributing to the bound state.

Eq. (10), which is the backward equation for the particle+leg over long times, is one of the main results of this paper, in the case of a 1-legged caterpillar. It is the backward equation for a particle that evolves as

$$\frac{dx}{dt} = \sqrt{2D_{\text{eff}}}\,\eta_x(t). \qquad (13)$$

That is, the particle diffuses, with effective diffusion coefficient D_{eff} and effective friction Γ_{eff}. The effective diffusivity and friction have the usual interpretation. In particular, if a potential $\mathscr{U}(x)$ were added to the particle Eqns. (2) and (3), one would recover in Eq. (13), following the same coarse-graining procedure, a term $-\frac{1}{\Gamma_{\text{eff}}}\partial_x\mathscr{U}$.

In Fig. 3 we compare the analytical result obtained in Eq. (12) (gray line) to numerical simulations of the full stochastic Eqns. (1)-(3) (gray dots). We show the results for a number of system parameters and find perfect agreement over several orders of magnitude of physical parameters. We also predict order of magnitude changes in the diffusion coefficient as the microscopic parameters change.

1.3 Microscopic parameters determine long term diffusion

How shall we interpret the expressions for the effective diffusivity Eq. (11) and the effective friction Eq. (12)? The effective diffusivity is a weighted sum of the diffusivity in each state, $D_{\text{eff}} = p_0 D_0 + p_1 D_1$ where the weights correspond to the probability to be in either state, and $D_i = k_B T / \Gamma_i$. The effective friction, on the other hand, is a harmonic weighted sum of the friction coefficients. That the diffusivity averages arithmetically is to be expected, since the mean squared displacement is an extensive quantity in a system with multiple states. Over a time t we can write

$$\overline{x^2(t)} = 2D_{\text{eff}}t = 2D_0 p_0 t + 2D_1 p_1 t$$

$$= 2D_0 t_0 + 2D_1 t_1 = \overline{x^2(t)}|_0 + \overline{x^2(t)}|_1,$$

where t_0 and t_1 refer to the time spent in either state. The novelty here is that the diffusivity in the bound state,

$$D_1 = k_B T (\Gamma + \gamma_{\text{eff}})^{-1} \neq k_B T (\Gamma + \gamma)^{-1},$$

is obtained not just from the friction in the bound state, see Eq. (12), but is modified by spring resistance during binding events by an additional term $\gamma_{\text{eff}} - \gamma$.

We can interpret this additional term by writing it as

$$\gamma_{\text{eff}} - \gamma = k\tau_{\text{eff}}, \quad \text{where} \quad \tau_{\text{eff}} = \tau_b + \tau_u^{\text{relax}}$$

is the typical time over which the leg's spring resistance acts, with $\tau_b = 1/q_{\text{off}}$ representing the average bound time, and $\tau_u^{\text{relax}} = \frac{\gamma}{k}\frac{q_{\text{on}}}{q_{\text{off}}} = \frac{\gamma}{k}\frac{\tau_b}{\tau_u}$ representing the bare relaxation time γ/k increased by the ratio of average bound time to average unbound time. This is coherent as the leg fluctuations may only relax in the unbound state. The interpretation of τ_{eff} is comparable to that in Ref. 54

although the results of Ref. 54 were obtained from underdamped dynamics.
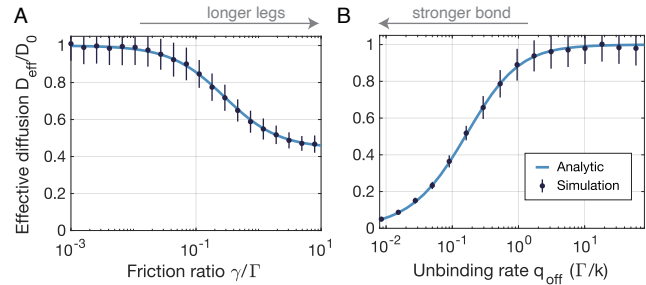
Fig. 3 Effective diffusion D_{eff} of a 1-legged particle. Simulation and analytical result Eq. (12) for a 1D system with 1 leg, with respect to (A) friction ratio γ/Γ and (B) unbinding rate q_{off}. (A) and (B) share the same y-axis. The other numerical parameters are $q_{\text{on}}\Gamma/k = 1.0$, and for (A) $q_{\text{off}}\Gamma/k = 0.8$ while for (B) $\gamma/\Gamma = 0.1$. Error bars represent one standard deviation for 100 independent runs.

Fig. 3 shows how the effective diffusion coefficient depends on microscopic parameters such as the leg friction and binding rates. As the leg friction γ increases, the effective diffusion of the particle decreases (Fig. 3-A). When the leg friction γ is large compared to all other contributions to friction, diffusion in the bound state is frozen $D_1 = 0$, and the effective diffusion corresponds only to mobility in the unbound state $D_{\text{eff}} = p_0 D_0$ ($p_0 = 0.8/1.8 \simeq 0.44$ in Fig. 3-A). As leg friction is typically proportional to the size of the legs, it is thus expected that the bigger the legs, the slower the particle. As the unbinding rate q_{off} decreases, D_{eff} decreases to arbitrarily small values (Fig. 3-B). This slow down is due to spring recoil forces acting over longer times, eventually freezing the particle in a given location. Note that similar qualitative dependencies of the diffusion coefficient on the unbinding rate ($D_{\text{eff}} \sim k_B T q_{\text{off}}/k$) were noted in a numerical model of multivalent transport on discrete sites[44], in a scaling law investigation of sticky reptation in polymers[62], and experimentally in Influenza A viruses[19].

As a test of modeling choice, the analytical expression may also be plotted against numerical simulations of the non-dimensional equations with any value of ε. We find perfect agreement up to $\varepsilon \lesssim 10$ (Supplementary Fig. S1), regardless of the choice of physical parameters. This highlights that the natural choice $\varepsilon = L/L_x$ for coarse-graining purposes, corresponding to bound leg length scales versus unbound particle long range motion, is especially well suited for these types of problems. In the following ε is not incorporated in numerical simulations.

1.4 Diffusion of N-legged caterpillar spans orders of magnitude

We extend our framework to probe nanocaterpillar dynamics with an arbitrary number of legs N (see Fig. 4-A). Eq. (1) is repeated for each unbound leg, and each leg binds to the surface with rates $q_{\text{on}}, q_{\text{off}}$ independently. Eq. (2) gives the particle dynamics when no legs are bound. When n legs are bound, indexed by $i = 1, \ldots, n$, the dynamics of the particle and bound legs are constrained as

(Supplementary 1.2)

$$\frac{dx}{dt} = -\frac{dl_i}{dt} = \frac{k}{\Gamma + n\gamma}\sum_{i=1}^{n}(l_i - l_0) + \sqrt{\frac{2k_BT}{\Gamma + n\gamma}}\eta. \quad (14)$$

Note here that the projection step yields a friction coefficient scaling linearly with the number of bonds n, and hence is not a perturbative effect [22]. The set of stochastic equations is now fully determined and can be simulated for any N, see Fig. 4-B.
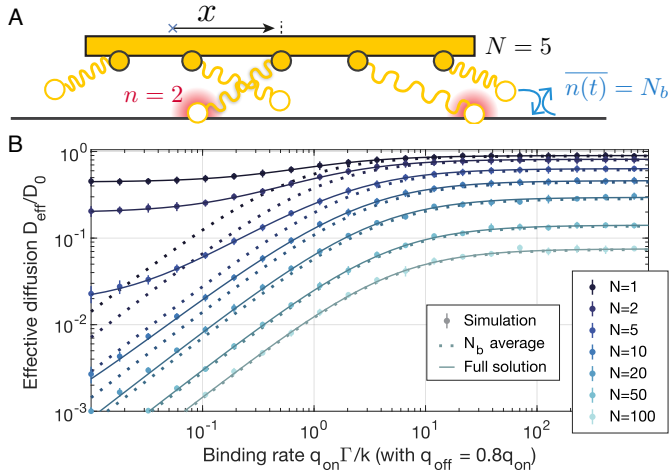
Fig. 4 N-legged nanocaterpillar model. (A) The longitudinal extension of N legs are monitored (here $N = 5$) with binding and unbinding. The number of bonds $n(t)$ changes in time, here $n(t) = 2$. The average number of bonds $\overline{n(t)} = N_b$ depends on the binding and unbinding rates. (B) Simulations and analytical results of the effective diffusion coefficient for N-legs according to the binding rate $q_{on}\Gamma/k$. "N_b average" corresponds to Eq. (16) and "full solution" to Eq. (15). The other numerical parameters are $\gamma/\Gamma = 0.1$ and $q_{off} = 0.8q_{on}$.

Similarly as in Sec. 1.2, coarse-graining predicts a long time effective diffusion with N legs as (Supplementary 1.2)

$$D_{eff}^{N\,legs} = \frac{k_BT}{\Gamma_{eff}^{N\,legs}} = k_BT\sum_{n=0}^{N}\frac{p_n}{\Gamma_n} \quad (15)$$

where $p_n = \binom{N}{n}\frac{q_{off}^{N-n}q_{on}^n}{(q_{off}+q_{on})^N}$ is the equilibrium probability to have n bonds and Γ_n is the friction coefficient in a state with n bonds. The frictions $\{\Gamma_n\}$ solve a linear system of equations that does not have a simple analytical solution (see Eqns. (S1.20-22)), but can be solved using numerical linear algebra for given parameters as reported in Supplementary 1.2.

Eq. (15) is one of the main results of this paper. It predicts the long-term diffusion coefficient of a nanocaterpillar, as a nontrivial function of the microscopic parameters of the legs. We compare the numerically solved Eq. (15) (full lines) to numerical stochastic simulations with N legs (dots) in Fig. 4-B and find excellent agreement.

The coefficients Γ_n contributing to each bound state can be further investigated to yield an analytical approximation for $\Gamma_{eff}^{N\,legs}$. When a large number of legs N is involved in the process, the dominant term in the sum of Eq. (15) corresponds to the average number of bonds $N_b = \sum_{n=0}^{N}np_n = \frac{q_{on}}{q_{off}+q_{on}}N$. Furthermore, one expects that the coefficients vary weakly around $n = N_b$, simplifying

the linear system for the $\{\Gamma_n\}$, yielding

$$\frac{1}{\Gamma_{eff}^{N\,legs}} \underset{N \gg 1}{\simeq} \frac{1}{\Gamma_{N_b}} = \frac{1}{\Gamma + N_b\gamma_{eff}}. \quad (16)$$

The right hand side of Eq. (16) is valid regardless of parameter values (Fig. S3) and provides a good approximation for $\Gamma_{eff}^{N\,legs}$ for large values of N (Fig. S2). For example, close agreement with Eq. (15) is obtained as early as $N = 20$, while good qualitative agreement is obtained for $N = 5$ (see Fig. 4-B, dotted line). Eq. (16) shows that the effective friction with N legs decays linearly with the *average* number of bonds N_b. For systems with a large number of legs (and hence potentially a large average number of bonds) [31,38,39], we therefore expect a strong diffusion decrease, covering potentially several orders of magnitude, due to enhanced friction with the surface.

2   Do nanocaterpillars hop or slide?

Our model and analytical formula Eq. (15) are useful not only for quantitatively predicting the diffusion coefficients of existing nanocaterpillar systems, but also to obtain insight into the *mechanism* by which particles diffuse. Different experiments with DNA-coated colloids made puzzling and seemingly contradictory observations, whereby similar systems appear to diffuse in different ways. For example, some DNA-coated colloids appear to diffuse through a succession of uncohesive moves, namely hops above the surface [39], while others move cohesively along the surface [31]. The difference between cohesive and uncohesive modes of motion has been noted in a variety of other systems, ranging from virus mobility on surfaces [17,19] to sticky polymer reptation [62]. Yet the parameters that characterize and quantify these different modes of motion remain to be elucidated. Our model gives insight into this question – do nanocaterpillars prefer to diffuse by "sliding" along the surface, or by "hopping" along it (see Fig. 5-A)?

2.1   What are hopping and sliding?

We start by quantifying the diffusion associated with either hopping or sliding. The mean squared displacement of a particle whose diffusion coefficient is determined from Eq. (15) can be split into two contributions, as

$$\langle x^2 \rangle = 2D_{eff}t = 2p_0\frac{k_BT}{\Gamma_0}t + 2\sum_{n=1}^{N}p_n\frac{k_BT}{\Gamma_n}t$$

$$\equiv 2D_{hop}t + 2D_{slide}t.$$

We identify (a) a *hopping* mode (in accordance with Refs. 34 and 39) where the particle detaches *all bonds* with the surface and moves in free space (see Fig. 5-A), until it forms another bond. In this hopping mode

$$D_{hop} = p_0\frac{k_BT}{\Gamma} = \left(\frac{q_{off}}{q_{off}+q_{on}}\right)^N\frac{k_BT}{\Gamma}. \quad (17)$$

We also isolate (b) a *sliding* mode (see Fig. 5-A) where the particle keeps at least one bond with the surface, a form of walking with

no preferred direction,

$$D_{\text{slide}} = \sum_{n=1}^{N} \frac{p_n}{\Gamma_n} \simeq \frac{k_B T}{\Gamma_{N_b}} = \frac{k_B T}{\Gamma + N \frac{q_{\text{on}}}{q_{\text{off}} + q_{\text{on}}} \gamma_{\text{eff}}}. \tag{18}$$

The total mean-squared displacement can be broken up into the sum of the mean-squared displacement when hopping, and the mean-squared displacement when sliding, as $\langle x^2 \rangle = 2D_{\text{hop}}t + 2D_{\text{slide}}t = \langle x^2 \rangle_{\text{hop}} + \langle x^2 \rangle_{\text{slide}}$.
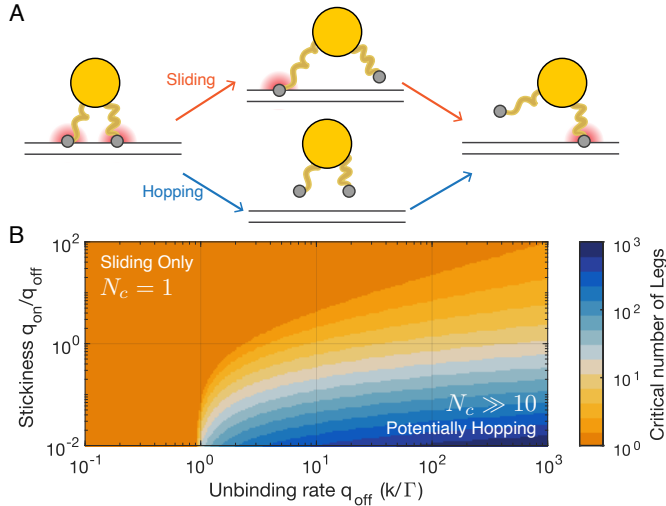
Fig. 5 Nanocaterpillar diffusion modes with N legs. (A) Typical modes of motion with N bonds: the nanocaterpillar may either slide (at least one bond remains attached to the surface) or hop (all bonds detach for the particle to move). (B) Critical number of legs N_c required for sliding to be more effective than hopping as a function of stickiness $q_{\text{on}}/q_{\text{off}}$ and unbinding rate.

An important observation is that D_{slide} decays with the number of legs roughly as $1/N$, while D_{hop} decays exponentially with N, *i.e.* much faster. As soon as a few legs are involved, we may therefore expect that *sliding* dominates *hopping*. This interpretation is natural, since when a system has just a few legs ($N \simeq 1-2$), the odds that the legs all *detach at once* are quite high, therefore favoring hopping. In contrast, in a system with a large number of legs, the odds that all legs *simultaneously* detach are simply too small, and the system walks randomly, remaining close to the surface. In a sense, nanocaterpillars truly are caterpillars walking with nanoscale legs. The scaling quantifying both modes of motion is another essential analytical result of our work.

In general, the critical number of legs $N_c(q_{\text{on}}, q_{\text{off}}, k, \gamma, \Gamma)$ required to favor sliding ($N \geq N_c$) over hopping ($N \leq N_c$) satisfies

$$\frac{\langle x^2 \rangle_{\text{hop}}}{\langle x^2 \rangle_{\text{slide}}} = \frac{D_{\text{hop}}}{D_{\text{slide}}} = \left(\frac{q_{\text{off}}}{q_{\text{off}} + q_{\text{on}}} \right)^{N_c} \left(1 + N_c \frac{q_{\text{on}}}{q_{\text{off}} + q_{\text{on}}} \frac{\gamma_{\text{eff}}}{\Gamma} \right) = 1. \tag{19}$$

The critical number of legs is controlled by the ratio $q_{\text{on}}/q_{\text{off}}$, termed henceforth *stickiness*, and by the magnitude of the effective friction in the bound states γ_{eff}, itself dominated in most systems by the unbinding rate q_{off}. We can therefore investigate N_c as a function of stickiness $q_{\text{on}}/q_{\text{off}}$ and unbinding rate q_{off} (Fig. 5-B). Overall, a system with say $N = 10$ legs is typically dominated by sliding motion. Yet hopping may still occur *e.g.* with large

unbinding rate q_{off}. In fact q_{off} increases the friction γ_{eff} in the bound states and reduces D_{slide}. The number of legs is thus a critical parameter for nanocaterpillar diffusion: controlling both the magnitude of the diffusion decrease and the mode of motion.

2.2 Distinguishing the diversity of biophysical nanocaterpillars

Whether a nanocaterpillar slides or hops, as predicted by Eq. (19), depends on numerous system parameters. Existing biological and biomimetic systems cover a broad range of parameters that we now explore, to ask which systems prefer to move by sliding and which by hopping, within the framework of our model.

Our model relies on 6 physical parameters $k, \gamma, q_{\text{off}}, q_{\text{on}}, \Gamma, N$ that can be estimated from the literature for many systems : viruses, molecular motors, white blood cells, protein cargos in the nuclear pore complex, bacteria such as Escherichia coli, and DNA-coated colloids (Supplementary 3). Typically, stickiness values are similar across systems with $q_{\text{on}}/q_{\text{off}} \sim 0.05 - 0.8 \geq 1$ – when the system is not thermally manipulated as will be explored in Sec. 2.3. Therefore we consider $q_{\text{on}}/q_{\text{off}} \simeq 0.1$. Additionally, as legs are generally small compared to particles, $\gamma/\Gamma \simeq 10^{-3} - 10^{-1}$ and therefore the dominant factor in $\gamma_{\text{eff}}/\Gamma$ is usually controlled by spring recoil force and unbinding times, as $k/\Gamma q_{\text{off}}$. We find $k/\Gamma q_{\text{off}} \simeq 10^{-2} - 10^{8}$ in the range of systems studied, confirming that this is a critical factor to discriminate nanocaterpillars. Additionally, as systems have a varied number of legs N, we define an effective relaxation rate

$$\frac{k^{(N)}}{\Gamma} = \frac{k}{\Gamma} N \frac{q_{\text{on}}}{q_{\text{off}} + q_{\text{on}}} \left[\left(\frac{q_{\text{off}} + q_{\text{on}}}{q_{\text{off}}} \right)^N - 1 \right]^{-1}$$

that will allow us to predict either sliding or hopping.

We sort systems in a so-to-speak *Ashby* chart, according to the effective relaxation rate $k^{(N)}/\Gamma$ and unbinding rate q_{off} (Fig. 6). This chart summarizes parameter ranges for different systems, and predicts which systems move by sliding and which move by hopping, within the assumptions of our model. If $k^{(N)}/\Gamma q_{\text{off}} \leq 1$, according to Eq. (19), sliding (orange region) is favored over hopping (blue region). While other modes of motion could occur for such complex systems, our aim here is to observe these systems in the "projected" sub-space where only sliding and hopping is considered. Interestingly, we find that different groups of systems emerge according to this classification, that we review below.

2.2.1 Sticky hoppers

We predict that viruses, white blood cells, and molecular motors cannot slide. These systems show very long bond lifetimes, with $\tau_{\text{off}} = q_{\text{off}}^{-1} \simeq 1 - 100$ s. This is characteristic of strong bonds, for which the interaction energy $|\Delta G| \gg k_B T$. Since for the protein ligands in these systems, $k \simeq 10^{-4}$ N/m and $\Gamma \simeq 10^{-9}$ N.s/m for 1μm particles, we expect $k/\Gamma \simeq 10^5 \gg q_{\text{off}}$ and $\gamma_{\text{eff}} \gg \Gamma$. Therefore such systems simply can not slide. Sliding is even more disfavored for coronaviruses (Sars CoV 1 and 2), since the legs are made of very rigid proteins, with $k \simeq 0.5$ N/m [63,64]. Hopping is therefore a probable mode of motion for these systems.

These predictions are qualitatively consistent with experimen-
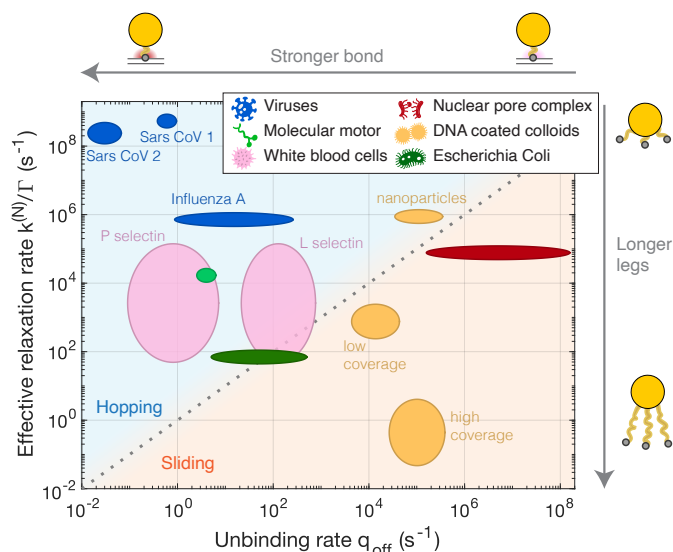
Fig. 6 Sorting biophysical systems. Expected regimes of sliding or hopping according to the effective relaxation rate $k^{(N)}/\Gamma$ and unbinding rate q_{off}. The gray line corresponds to $k^{(N)}/\Gamma = q_{\text{off}}$ and separates the sliding and the hopping regions. Circles represent the range of values found in the literature for parameters of each system. Systems are color coded according to their category in the legend. When multiple systems belong to a category, details are indicated next to the circles. Low and high coverage DNA-coated colloids refer to 1 μm size colloids and nanoparticles to 15 nm size.

tal measurements. The diffusion coefficient of an influenza A virus on protein-coated surfaces was measured as $D_0/D_{\text{eff}} \simeq 4 - 190$[17,19]. Estimating the typical number of available legs $N \simeq 10$[65,66] and the bound probability $q_{\text{on}}/(q_{\text{on}} + q_{\text{off}}) = 20\%$[66] yields $D_0/D_{\text{hop}} = [q_{\text{off}}/(q_{\text{on}} + q_{\text{off}})]^N \simeq 10$, in the range of measured values. Our model predicts that hopping is therefore more probable than sliding for influenza A, at least when considering its translational motion under passive binding and unbinding. This is consistent with Ref. 17, which observed infrequent yet very long spatial steps, termed *gliding* moves. We note that the influenza A virus has also been observed to move via cohesive short spatial steps, that have been attributed to rolling motion[5,17,19,41], which may be due in this context to active bond cleaving[17,19,41] that is beyond the scope of passive binding as presented here.

Turning to DNA-coated colloids, while the binding kinetics are roughly independent of colloid size, the effective relaxation rate can vary strongly. Nanometre-sized DNA-coated colloids (yellow nanoparticles) have fast relaxation rates as they are small (and therefore Γ is smaller), and are thus sticky hoppers. In contrast, microscale colloids have slower relaxation rates $k^{(N)}/\Gamma$, all the more as usually a great number of bonds $N \simeq 100$ are involved in the binding process, and thus are prone to slide. We will turn in more detail to DNA-coated colloids in Sec. 2.3.

2.2.2 Slippery sliders

Reciprocally, we predict that systems with weak adhesion (equivalent to short bond lifetimes, *i.e.* large q_{off}) may move by sliding. Such systems include proteins translocating through the nuclear pore complex, or white blood cells adhering through L-selectin

linkers, which are notably weaker than P-selectin[23]. Sliding may also be accessible to systems with short effective relaxation rate, for which the sticky friction mediated by k/Γ is low. This corresponds to large particles with long legs, as is the case for Escherichia Coli[57] (dark green). DNA-coated colloids with high DNA coverage are prone to slide due to their large number of legs.

2.3 DNA-coated colloids hop *and* slide, with order of magnitude decrease in their diffusion coefficient

We now turn to probe in more detail the predicted modes of motion and strong decrease in diffusion of DNA-coated colloids by comparing our model's predictions with experimental measurements of DNA-coated colloids. DNA-coated colloids provide a well-controlled model system for testing our analytical results, especially their dependence on N, since the number of DNA legs involved in the sticking process may be easily tuned by changing the temperature[38]. Our aim here is not to build a detailed model to describe all the possible modes of motion of DNA-coated colloids. Rather, we seek potential key parameters that control the magnitude of the diffusion and the mode of motion. To do so, we test whether the predicted strong decrease is coherent with experimental observations over a range of temperatures and for three different experimental designs.

2.3.1 Model parameters can be directly established from experimental data.

We predict the diffusion coefficients D_{eff} (and D_{slide} and D_{hop}) for three different experimental systems, by determining the parameters involved in Eq. (15) from the literature or from independent measurements, with no fitting parameters (apart from calibrating to the melting temperature, as discussed below). The diffusion coefficients for DNA-coated colloids on flat DNA-coated surfaces have been measured in two different experimental systems reported in the literature[31,39]. These studies report only very few data points around the melting temperature where motion is diffusive, since in these experimental systems diffusive motion is only observed in a narrow range of temperatures, so the studies focused mainly on the low temperature regime where motion is subdiffusive. We complemented the scarce existing data by performing our own experiments, using recently-developed fabrication[38] and acquisition techniques[31,39], and we observe diffusive motion over a wider range of temperatures (Supplementary 2). For each of the three experimental datasets, we map reported experimental parameters to the parameters of the model, and detail our process below.

Some parameters are easily estimated using standard results, see Table 2. The friction coefficient Γ is taken as the hindered lateral hydrodynamic friction near a wall[67]; γ and k correspond to hydrodynamic friction and spring resistance of the polymer linker (that links the surface and the complementary DNA strand) and are directly established from polymer dynamics[56]. The binding rate q_{on} depends on the exact – known – DNA sequence used for the complementary stickers and the density of coated DNA strands on surfaces[70].

Other parameters, such as N and N_b (or equivalently N and the

Table 1 Method used to calculate model parameters for the DNA-coated colloids studied experimentally in this work. Parameter values are reported only at the melting temperature T_m. Their dependence on temperature is indicated in the "Comments and References" column.

Parameter	Formula used	Value at T_m	Comments and References
Γ	$\Gamma = 2 \times 6\pi\eta(T)R$	1.6×10^{-8} N.s/m	hydrodynamic friction near a surface[67]; colloid radius $R = 500$ nm; $\eta(T)$ water viscosity with temperature.
γ	$\gamma = 6\pi\eta(T)h$	1.8×10^{-10} N.s/m	with brush height $h \simeq 22$ nm, calculated with Milner-Witten-Cates theory[68], and accounting for increased brush density due to Pluronic F127 (see Ref. 38).
k	$k = 3k_BT/2L\ell$	0.16 mN/m	spring constant for polymers[56]; extended brush length $L \simeq 84$ nm (6500 g/mol PEO + 20 single stranded DNA (ssDNA) bases); persistence length $\ell = 0.5$ nm (average of PEO + ssDNA at 140 mM salt concentration[69])
q_{on}	$q_{\text{on}} = k_{\text{on}}\bar{\sigma}/h\mathcal{N}_A$	4 kHz	where $k_{\text{on}} = 1.6 \times 10^6$ M^{-1}.s^{-1} from Ref. 70, using the exact sequence as in our experiments; $\bar{\sigma} = \sqrt{\sigma\sigma_g}$ where $\sigma = 1/(3.27 \text{ nm})^2$ is the particle coating density and $\sigma_g = 1/(10.8 \text{ nm})^2$ is the glass substrate coating density; Avogadro's number $\mathcal{N}_A$; Independent of T.
q_{off}	$q_{\text{off}} = q_{\text{on}}\frac{N(T)-N_b(T)}{N_b(T)}$	18 kHz	N_b average number of bound legs and N total number of legs available for binding in the interaction region; Dependent on T.

ratio $q_{\text{on}}/q_{\text{off}}$) require more extensive modeling of the detailed leg-arm interactions to be evaluated. Recently Refs. 38 and 59 have shown how to establish N and N_b with no fitting parameters, taking as input parameters the DNA sequence used, the coating densities, and the properties of the DNA linker (see Fig. S5), and we employ the method we have developed in Ref. 38.

Finally, since measurements include colloid vertical motion beyond the binding range*, we further include vertical motion and hence particle buoyancy through a 2×1D model. Such vertical motion is generally slow and only affects the effective probabilities p_n, not the friction coefficients Γ_n. Motion in two lateral dimensions can be straightforwardly extended from our 1D model (see Supplementary 2 for more details).

All parameters are thus readily expressed from detailed experimental system design. The diffusion coefficient D_{eff} is decreased by orders of magnitude at low temperatures. It progressively increases to its "bare" value – corresponding to non-sticky DNA – at high temperatures, with a sharp transition. This sharp transition from the bound to unbound state occurs at a melting temperature T_m specific to each experimental design. The predicted T_m is always close to the experimentally measured T_m (less than 1°C difference) with no fitting parameters.

Nonetheless, intrinsic variations remain in experimental parameters. In particular, different e.g. humidity conditions can affect the coating process and exact coating density obtained, and hence the experimental T_m, over about 2°C. To investigate data over the relevant short temperature range where diffusion can be measured, one option could be to fit e.g. the value of the coating density on colloids, to obtain the exact experimental T_m – effectively fitting the location of the sharp transition. Instead, we choose to align all data (theoretical or experimental) with respect to its own melting point T_m (predicted or measured). This has the advantage of avoiding fitting and allowing us to easily compare similar experimental systems with slightly different T_m (Supplementary 2).
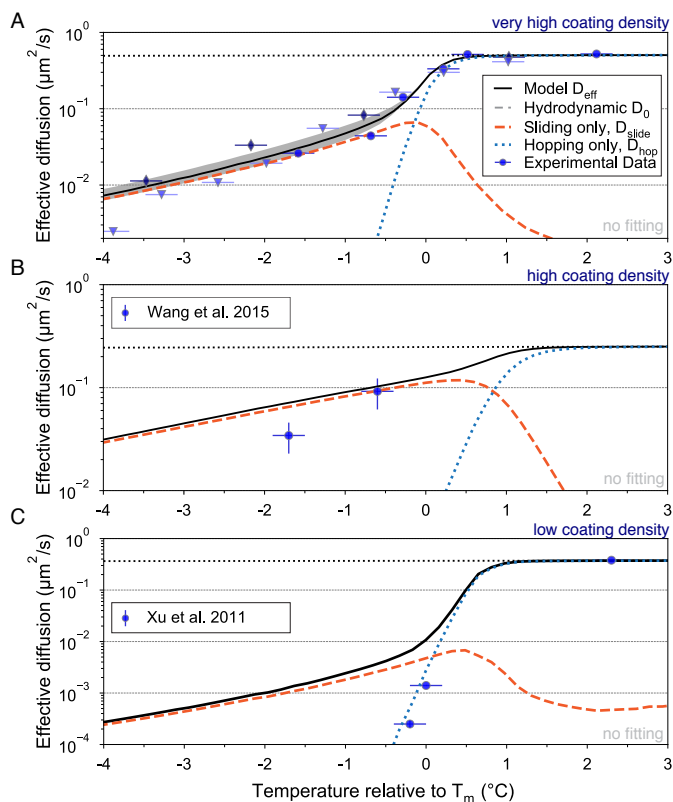
Fig. 7 Diffusion coefficients of DNA-coated colloids. Comparison between experimentally measured diffusion coefficients of DNA-coated colloids on DNA-coated surfaces and analytical predictions of D_{eff}, D_{slide}, and D_{hop} (Eqns. (15), (18) and (17)). The DNA-coated colloids have (A) highly dense coatings (1 DNA per 10 nm^2, Supplementary 2) (B) dense coatings (1 DNA per 27 nm^2) from Ref. 31 and (C) sparse coatings (1 DNA per 144 nm^2) from Ref. 39. In (A) the gray region corresponds to uncertainties on the coating density of the substrate, and the different symbols correspond to repeated experiments repeated. The hydrodynamic diffusion $D_0 = k_BT/12\pi\eta R$ corresponds to lateral diffusion near a flat rigid wall, where R is the radius of the colloid and η the solution viscosity. Horizontal error bars correspond to uncertainties on imposed temperature and vertical error bars correspond to uncertainties in determining the diffusion coefficient from data (Supplementary 2).

---

* The binding range is about 20 nm, but this is not optically removable as the vertical resolution is about 200 nm

2.3.2 The coating density controls the mode of motion and the magnitude of the diffusion coefficient decrease.

The number of legs implied in the sticking process N changes significantly with temperature. At low temperatures $N \gtrsim 100$; the colloids are strongly bound. With increasing temperatures N decreases until the particles are completely unbound and $N = 0$ (see Fig. S5), with a sharp transition at the melting temperature T_m. Importantly, the number of legs is the parameter that changes the most with temperature and controls therefore the magnitude of the long time diffusion D_{eff}.

The three experimental systems differ mainly in the DNA coating density, which implicitly controls the number of legs N involved in the binding process. For densely coated colloids (Fig. 7, A and B), we find excellent agreement between our model calculation for D_{eff} and experimental data, predicting a fast diffusion decrease over 2 orders of magnitude in barely a few temperature degrees. Further, we predict that sliding, or some form of cohesive motion with the surface, is the dominant mode of motion below the melting temperature T_m. In fact the high number of available legs, $N \simeq 100$, due to high coverage, prevents hopping below the melting temperature and colloids primarily slide, consistent with the observed cohesive motion[31]. Hopping emerges as a favorable mode above the melting point, where the average number of available and bound legs significantly decreases due to particle lift-off from the surface. This prediction is consistent with our qualitative observations above the melting point: particles perform long moves over short time intervals, accompanied by more frequent and longer excursions far from the surface. The transition between motion modes occurs for about $N = 40$ legs in contact (Fig. S5).

For DNA-coated colloids with low coverage densities, as in Ref. 39 (Fig. 7 C), our model predicts a diffusion coefficient that is far too large. Yet, D_{hop} is in remarkable agreement with experimental data. In fact, D_{eff} contains sliding motion yet the spacing between legs in Ref. 39 is too large and geometrically prevents sliding. Hence only hopping, or uncohesive motion with the surface, is possible. In fact, for such systems only hopping is observed, resulting in a much stronger slow down of diffusion with decreasing temperature[39]. The DNA coating density therefore appears to be a significant factor in determining how DNA-coated colloids move, allowing it to vary from sliding to hopping.

2.3.3 Other possible modes of motion.

There are other ways that DNA-coated colloids could move in specific experimental regimes, that could be probed with the analytical tools set forth here, yet that we have not yet explored. At lower temperatures, particles don't diffuse, they rather subdiffuse[31,39], potentially due to inhomogeneities in the coated surfaces[31,39,42]. Such spatial dependencies are not accounted for in our model but could be studied through spatially dependent attachment rates $q_{\mathrm{on}}(x)$ or leg number $N(x)$.

Particles may also move by rolling instead of by sliding[31], a motion that could also be investigated with homogenization techniques. Rolling may have a higher mobility at some temperatures[33,54], since the strands closest to the contact point on the surface do not resist rolling, for geometrical reasons. Yet when a large number of bonds are implied in the binding process, numerous bonds are actually far from the contact point and hence resist rolling. It is possible that rolling is thus favorable only over a small range of temperatures.

Although our model lacks these more complex ingredients and geometries, it is in surprisingly good agreement with our experimental measurements. This suggests we have identified some critical parameters controlling the observed effective diffusion, precisely the coating density and working temperature as they set the number of legs N. Even in a more complex model, containing *e.g.* inhomogeneous coating density, or rotational degrees of freedom, we therefore expect these parameters to play an important role in mobility.

2.4 Design rules for sliding versus hopping

Herewith we can draw simple design rules for sliding or hopping. Numerous, long wobbly legs with weak adhesive bonds are well adapted for sliding. Short and stiff legs with strong adhesive bonds facilitate hopping. DNA-coated colloids offer various design features to control their mobility: for example, larger particle size, higher DNA coverage, and lower temperature all favor sliding. Further control can be achieved by tuning the microscopic features of the legs, such as their spring constants k, for example by choosing the length of the ligand leg[38]. However, such control is especially hard to achieve experimentally *without* changing other experimental features at the same time. For example, current coating processes generally result in less dense coatings for longer legs[38].

Overall, these design rules allow one to tune artificial systems to control their mobility. This could have consequences in particular in the field of self-assembly of artificial structures, where facilitated cohesive motion is believed to be essential for long-range alignment[31–33].

3 Coarse-graining under different models and assumptions

In the physical and biological systems we explored, the range of physical parameters was quite broad, suggesting that other scaling ansätze might be appropriate to study long term dynamics. We review alternative approximations and modeling assumptions and compare them to the predictions of the model presented in Section 1. We find that our model is the most general, encapsulating perturbative results obtained with other approximations, and that it is naturally modified to account for additional features (such as arms as well as legs). To make the argument simpler, we mainly focus on a 1-legged caterpillar; the comparisons should be similar for a multi-legged caterpillar. Detailed coarse-graining steps are reported in Supplementary 4. All results are summarized in Table 2 (displayed in the Appendix) and compared in Fig. 8.

3.1 Dynamics with inertia

One may include particle inertia with a small yet finite mass $m \neq 0$, by starting with the underdamped Langevin equations for the particle (rather than the overdamped as we have done) – see

Ref. 54. To understand the scales associated with mass, one can compare the correlation time of the particle's velocity when spring recoil forces are at play, $\tau_v \simeq \frac{m(L_x/\tau)}{Lk}$, to the time scale of observation τ [54]. Coarse-grained dynamics require $\frac{\tau_v}{\tau} = \frac{mL_x}{Lk\tau^2} = O(\varepsilon)$, which is apparently coherent with a small mass.

Coarse-graining steps (Supplementary 4.1) lead to an effective friction

$$\Gamma_{\text{eff}}^m = p_0\Gamma_0 + p_1\Gamma_1. \tag{20}$$

Notice that the effective friction is the *arithmetic* sum of the frictions in each state – not the *harmonic* sum obtained in Eq. (12)[†]. Eq. (20) is equivalent to Eq. (12) in the limit where the friction correction is small, $\gamma_{\text{eff}} \ll \Gamma$ – see Fig. 8-B (yellow).

However, differences arise beyond this regime. For stiff legs ($\gamma/\Gamma \gg 1$, $k/q_{\text{off}}\Gamma \gg 1$) one finds $\Gamma_{\text{eff}}^m \sim 0$ while $\Gamma_{\text{eff}} \sim \Gamma$. This stark difference has an intuitive explanation: the particle may not move when it is attached with the stiff leg, but it can still move when it is unbound, and therefore effective friction should remain finite. This is true *unless* the particle has significant inertia and therefore does not have the time to accelerate within the unbound periods. In fact, in the non-dimensionalization we implicitly assumed that $m/\Gamma = \varepsilon Lk\tau^2/\Gamma L_x = \Gamma/k\varepsilon^2$, such that the inertial relaxation time was in fact assumed to be large compared to the time scale of velocity fluctuations.

This drives the general question of how to account for inertia in such systems, and whether inertia plays a role in the macroscopic diffusion of nanocaterpillars. We will address this question thoroughly in another paper [71], in which we reconcile Eq. (20) and Eq. (12).
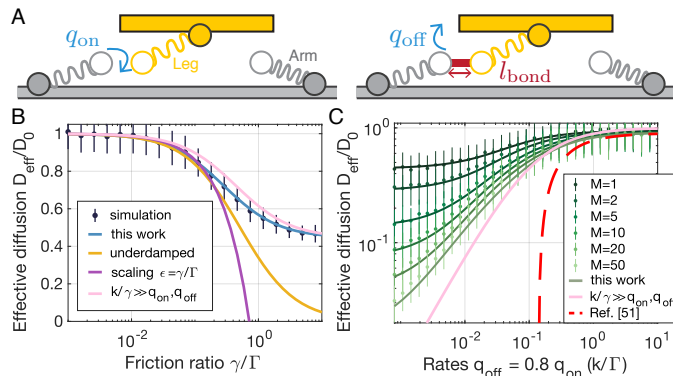
Fig. 8 Comparing with other coarse-grained models and assumptions. (A) Schematic for arm and leg dynamics considered in this work. (B) Effective diffusion with respect to friction ratio γ/Γ: calculated with Eq. (12) ("This work"), Eq. (20) ("underdamped"), Eq. (21) ("scaling $\varepsilon = \gamma/\Gamma$") and Eq. (22) ("$k/\gamma \gg q_{\text{on}}, q_{\text{off}}$"). (C) Effective diffusion with respect to binding and unbinding rates (keeping $q_{\text{on}}/q_{\text{off}}$ constant), for a particle with 1 leg facing $M = 1 - 50$ arms: calculated with Eq. (25) ("This work") and Eq. (22) ("$k/\gamma \gg q_{\text{on}}, q_{\text{off}}$"), taking $p_0 = 0$ and $p_1 = 1$ to match the limits in $M \to \infty$. Ref. 51 corresponds both to $k/\gamma \gg q_{\text{on}}, q_{\text{off}}$ and $\gamma/\Gamma = \varepsilon$ and was plotted for consistency. For (A) and (B), shared numerical parameters are $q_{\text{on}}\Gamma/k = 1.0$, $q_{\text{off}}\Gamma/k = 0.8$ and $\gamma/\Gamma = 0.1$.

[†] Eq. (20) corresponds to the result derived in Ref. 54, with in addition projected dynamics for the bound state, and base friction of the particle ($\Gamma \neq 0$)

3.2 Choice of time-scale hierarchy

There are other choices for the ordering of time scales. We review these below: we describe their experimental relevance, then briefly examine the effective friction under these different approximations and compare it to our main result Eq. (12).

3.2.1 Fast leg dynamics compared to particle dynamics

One common approximation is to assume rapid leg dynamics compared to particle dynamics, with $\varepsilon = \gamma/\Gamma$ [51]. Such an approximation is consistent with numerous experiments, as legs are typically short, hence fast because of Stokes relation, compared to the large particles investigated (such as white blood cells [7] or DNA-coated colloids [72]).

With this assumption one typically relaxes the restriction on lengthscales, as $L \sim L_x$. The observation time-scale is $\tau = L^2/D_0 = \Gamma/k$ and binding and unbinding are taken to be fast compared to this time scale, $q_{\text{on}} \sim q_{\text{off}} \sim 1/\tau\varepsilon$. One obtains (Supplementary 4.2.1)

$$\frac{1}{\Gamma_{\text{eff}}^{\varepsilon=\gamma/\Gamma}} = \frac{p_0}{\Gamma} + \frac{p_1}{\Gamma}\left(1 - \frac{\gamma_{\text{eff}}}{\Gamma}\right). \tag{21}$$

Eq. (21) results in a small correction to the effective friction, of order ε. It is equivalent to Eq. (12) in the limit where $\gamma_{\text{eff}} \ll \Gamma$ is small. The assumption $\varepsilon = \gamma/\Gamma$ appears thus quite restrictive as it implicitly also requires to observe the system at long time scales compared to the other time scales in the system. Furthermore, contrary to Eq. (12) where the small parameter ε disappears, here $1/\Gamma_{\text{eff}}^{\varepsilon=\gamma/\Gamma}$ is a first order expansion in $\varepsilon \sim \gamma_{\text{eff}}/\Gamma$. We present Eq. (21) against Eq. (12) in Fig. 8-B (purple vs black) and find that Eq. (21) is indeed only valid for small values of γ/Γ. Our choice of scaling $\varepsilon = L/L_x$ can thus account for a broad range of bare friction values. Additionally, such an approach can only account for small perturbations to the background mobility, while we find perturbations over several orders of magnitude.

3.2.2 Fast leg dynamics compared to binding dynamics

Another approximation assumes fast leg relaxation dynamics compared to binding dynamics, $k/\gamma \gg q_{\text{on}}, q_{\text{off}}$ (and both are fast compared to particle dynamics). In this case leg lengths are sampled from their equilibrium distribution when they bind, corresponding to a "pre-averaging" approximation. Leg lengths are not tracked when they are unbound, allowing to speed up simulations [22,33,51,73]. This limit is relevant to describe stiff legs, *e.g.* rigid polymers such as double stranded DNA – see Table S1.

Coarse-graining gives (Supplementary 4.2.2)

$$\frac{1}{\Gamma_{\text{eff}}^{k/\gamma \gg q}} = \frac{p_0}{\Gamma} + \frac{p_1}{\Gamma + \gamma + \frac{k}{q_{\text{off}}}}. \tag{22}$$

The pre-averaged result Eq. (22) is comparable to Eq. (12), yet misses the relaxation term involving $\tau_{\text{u}}^{\text{relax}}$ in γ_{eff}. This confirms that $\tau_{\text{u}}^{\text{relax}}$ originates from unbound relaxation dynamics. This difference results in some differences in D_{eff}, depending on the microscopic parameters (Fig. 8-B). Additionally, the pre-averaged limit may be viewed as the limit regime for a nanocaterpillar with a large number of legs, say $N \gg 1$, where on average 1 or 0 leg is bound to the surface, $N_b \lesssim 1$. This typically requires

$q_{\text{on}} \ll q_{\text{off}} \ll k/\gamma$, and indeed Eq. (15) converges to the pre-averaged result in that limit (Supplementary Fig S4).

The validity of pre-averaging is limited to the domain $q_{\text{on}}, q_{\text{off}} \ll k/\gamma$. In systems such as DNA-coated colloids, binding rates q_{on} and q_{off} may be manipulated over orders of magnitude[74], by choosing the DNA sequence or by adjusting temperature, potentially accessing $q_{\text{on}} \gg q_{\text{off}} \gg k/\gamma$ at low temperatures. In this regime, Eq. (12) predicts that the nanocaterpillar is frozen in the bound state, while pre-averaged dynamics still predict a non zero mobility. In these situations pre-averaged dynamics are therefore not suitable. We show later however that introducing numerous arms – more generally a lot of binding partners – can extend the validity range of pre-averaging.

3.2.3 Fast binding dynamics compared to leg dynamics

Finally, one can consider fast binding dynamics compared to leg dynamics, $q_{\text{on}}, q_{\text{off}} \gg k/\gamma$. Although this limit is not often considered in simulations, it is relevant for dense arrangements of receptor sites[72]. In fact as the binding rate q_{on} scales linearly with the concentration of receptors, it can increase by orders of magnitude for a leg potentially in contact with a dense array of arms – see Table S1.

Coarse-graining yields (Supplementary 4.2.3)

$$\frac{1}{\Gamma_{\text{eff}}^{q\,\text{fast}}} = \frac{p_0}{\Gamma} + \frac{p_1}{\Gamma + \gamma + k\left(\frac{\gamma}{k}\frac{q_{\text{on}}}{q_{\text{off}}}\right)} \tag{23}$$

which is exactly what is expected in the limit $q_{\text{on}}, q_{\text{off}} \gg k/\gamma$ in Eq. (12). Again, this highlights the physical mechanisms yielding the different contributions in γ_{eff}. Here the average bound time of the leg is small, $\tau_{\text{b}} \ll \gamma/k$, and therefore does not contribute to γ_{eff}.

3.3 Arms and/or legs

The diversity of nanocaterpillars resides also in their geometry: some particles have legs that attach to a surface[75], some have no legs (or infinitesimally small legs), with binding sites directly on the particle that attach to outstretched receptors on the surface that we refer to as arms[22,51] (1 arm case in Table 2) and some have both outstretched legs connecting to outstretched arms[33] (arms and legs in Table 2).

3.3.1 Arms or legs

A particle with a leg or a bare particle attaching to an arm (1-legged and 1-armed respectively, see Table 2) have nearly equivalent effective dynamics. The only difference resides in the interpretation of Γ in the unbound leg dynamics Eq. (2) – see Supplementary 4.3.1. For the 1-legged case, if the leg's center of mass corresponds to the point grafted to the particle, the unbound friction coefficient is simply increased by the leg as $\Gamma \to \Gamma + \gamma$, where Γ is the bare particle friction coefficient and γ the leg's. If the leg's center of mass is offset from the grafting point on the surface, minor modifications have to be made to Eq. (2) yet lead to very similar dynamics overall. For the 1-armed case, we simply have the unbound friction coefficient Γ to be the bare friction coefficient of the particle. This justifies our approach in Sec. I, where

we ignore the details of the leg or arm location and simply treat them as mathematically equivalent.

3.3.2 Arm and leg

A 1-legged particle attaching to 1 arm has slightly more interesting dynamics. To investigate this case, we simplify the problem and consider that the leg can bind to the arm regardless of their relative location, with a rigid rod of length l_{bond} that bridges the gap between the sticky points (see Fig. 8-A). In the bound state the constraint is thus $x + l_{\text{leg}} - l_{\text{arm}} = l_{\text{bond}}$. The relative distance l_{bond} is unimportant and can be assumed to be zero, and therefore this model effectively creates an arm with the correct length each time the leg binds.

Although the model is simplistic, it is realistic in the presence of a dense periodic array of arms and allows us to compare the mechanical properties of this geometry compared to a single leg or arm. We find using similar coarse-graining techniques (Supplementary 4.3.2)

$$\frac{1}{\Gamma_{\text{eff}}^{leg+arm}} = \frac{p_0}{\Gamma} + \frac{p_1}{\Gamma + \gamma_{\text{eff},1}(1,1)} \text{ where } \gamma_{\text{eff},1}(1,1) = \frac{\gamma_{\text{eff}}}{2}. \tag{24}$$

The added friction in the bound state is only half that with a single leg or a single arm: friction is distributed harmonically, like the effective spring constant of two springs in series[‡]. Slightly improved mobility is therefore achieved with both an arm and a leg, while the qualitative behavior of the original model is preserved.

3.3.3 Leg facing numerous arms

We now consider a leg that can bind to *multiple* arms at the same time. As in the previous section, the M arms do not have particular locations but rather appear with the correct lengths when needed. In that case, the binding rate depends on the number of bound legs. For a given leg, the effective binding rate is $(M-n)q_{\text{on}}$, where n is the current number of bound legs, such that $M-n$ corresponds to the number of available binding sites. The effective unbinding rate of each leg remains q_{off}. Following the formalism of arm and leg dynamics detailed above (Supplementary 4.3.3) one finds that with M arms,

$$\frac{1}{\Gamma_{\text{eff}}^{leg+M\,arms}} = \frac{p_{M,0}}{\Gamma} + \frac{p_{M,1}}{\Gamma + \gamma_{\text{eff},1}(M,1)} \tag{25}$$

where $p_{M,0} = q_{\text{off}}/(q_{\text{off}} + Mq_{\text{on}})$ and $p_{M,1} = 1 - p_{M,0}$ are the probabilities to have 0 or 1 bond. The added friction $\gamma_{\text{eff},1}$ is a harmonic average when M is large

$$\frac{1}{\gamma_{\text{eff},1}(M,1)} \underset{M \gg 1}{\simeq} \frac{1}{\gamma_{\text{eff},M,1}} + \frac{1}{\gamma_{\text{eff},1,1}}, \tag{26}$$

with $\gamma_{\text{eff},M,1} = k\left(\frac{1}{q_{\text{off}}} + \frac{\gamma}{k}\frac{(M-1)q_{\text{on}}+q_{\text{off}}}{q_{\text{off}}}\right)$ the effective friction due to the leg $\gamma_{\text{eff},1,1} = k\left(\frac{1}{q_{\text{off}}} + \frac{\gamma}{k}\right)$ due to arms. We see that the factors implying the unbound relaxation time $\tau_{\text{u}}^{\text{relax}}$ are modified in each

---

[‡] Note however that attaching springs with different spring constants would not lead to a similar harmonic sum of effective frictions, as the effective friction contains more contributions than those originating from the spring recoil force (analytical results not shown here).

case. We give the following interpretation: the average unbound time for the leg is $\tau_u = 1/(M-1)q_{on}$, due to $M-1$ other available arms to bind to. For the arms, $\tau_u = \infty$ as there are no other legs to bind to once the only leg is bound. The harmonic average in Eq. (26) highlights again that the leg-arm configuration is mathematically similar to the effective force of springs in series.

In the limit of a large number of arms M, the leg is always bound to the surface ($p_1 = 1$) and the correction to the bound state friction converges to

$$\gamma_{eff,1}(M,1) \xrightarrow[M\to\infty]{} \gamma_{eff,1}(1,1) = \gamma + \frac{k}{q_{off}}, \qquad (27)$$

which is the correction to the effective friction for the pre-averaged result, Eq (22).

This limit is surprising. Sec 1, Eq. (12) showed that for a leg binding to a uniformly sticky surface, in the limit where the leg is always bound ($p_1 = 1$), the nanocaterpillar is frozen and $D_{eff} = 0$. When the leg is bound to a great many arms this is no longer the case: we recover the diffusion coefficient associated with pre-averaging. We interpret this discrepancy as follows. With many arms binding to a leg, the particle may still move, even in a parameter regime where the leg is always bound. In fact, the leg rapidly swaps between different arms, which have different random lengths and hence apply different random forces, causing the particle's position to fluctuate. Indeed, in Eq. (27) it is apparent that the remaining friction is due to arms and not to the leg. Swapping the particle upside down, this is equivalent to a particle with a large number M of legs binding to a uniformly sticky surface, but where on average only 0 or 1 leg is bound to the surface at a time. Therefore, this limit is equivalent to the pre-averaged result: each time a new arm is bound it is sampled from its equilibrium distribution – as so many arms are within reach.

Simulations with M arms are presented in Fig. 8-C with analytical solutions Eq. (25) (green colors). They indeed converge to the pre-averaged result (pink). For consistency, we also record the result of Ref. 51 (Eq. (2.48)) that corresponds to pre-averaging *and* assumes $\varepsilon = \gamma/\Gamma$. It is plotted in Fig. 8-C (red) and agrees with our result only over a limited range of parameters, corresponding to the validity range of Ref. 51.

3.3.4 Numerous legs facing numerous arms

N legs binding to M arms induce a long time effective friction that encapsulates the previous result for M arms and that for N legs in Sec. 1.4 (Supplementary 4.3.4). Eq. (15) still holds with adapted bond probabilities p_n, and γ_{eff} in Eq. (16) is the harmonic average between arm and leg contributions, $(\gamma_{eff,n}(M,N))^{-1} = \gamma_{eff,M,n}^{-1} + \gamma_{eff,N,n}^{-1}$.

Overall, spanning different limits shows that our methodology to investigate long time dynamics is robust, as it accounts for a broad range of physical parameters and a variety of geometries. It also justifies the use of "pre-averaging" approximations (sampling leg lengths from equilibrium distributions upon binding) to accelerate simulations in specific situations. It also highlights that taking limits of various parameters is subtle, and care must be taken when doing so as the limits do not commute in general.

Conclusion

When a particle is coated with ligands that bind and unbind stochastically to receptors on a surface, the ligands impart a random force to the particle each time they bind, causing the particle to undergo a random walk on long timescales. We constructed a model for the coupled dynamics of such a nanocaterpillar and its leg-like ligands, and derived an analytical expression for the nanocaterpillar's long-term effective diffusion coefficient as a function of the microscopic leg parameters. Our simulations showed this expression is valid over a broad range of parameters. Our expression predicts a dramatic decrease in the diffusion coefficient, by several orders of magnitude, as temperature decreases by a few degrees, a prediction that is borne out in our experimental measurements.

Our model allows us to distinguish between two modes of motion, sliding and hopping, and to identify parameters that govern which mode of motion is preferred, across a wide range of biophysical systems. Typically, systems with a large number of legs will slide, since the mean-squared displacement due to hopping decreases exponentially with the number of bound legs. Hopping is favored for systems with short, stiff legs, and/or strong bonds. Regardless of the mode of motion, the fast binding and relaxation dynamics at the microscale result in an overall slow diffusion of the nanocaterpillar, sometimes many times smaller than the background hydrodynamic diffusion.

We derived the effective diffusivity for a range of other models and scaling assumptions, which allowed us to tease out *e.g.* the effect of having arms (flexible receptors) as well as legs, having significantly more arms than legs or *vice versa*, having significant inertia, *etc*. In particular, we explored the validity range of specific approximations used to accelerate simulations, such as that upon binding, leg lengths are sampled from their equilibrium distributions[22,33,51]. We showed this approximation is valid for fast leg dynamics $\gamma/k \ll q_{on}, q_{off}$ in 1D, or when binding to a great number of binding partners, such as many arms, $M \gg 1$, yet its validity should be reassessed in more complex geometries.

There are numerous ways to build upon our model to address additional complexities within the same coarse-graining framework. An important step would be to incorporate particle rotational degrees of freedom, and to ask how rolling compares to hopping and sliding. Rolling has been predicted to lead to a low effective friction in systems with stiff legs, because it doesn't require stretching legs at the contact point[33,54]. While rolling has been modeled in special situations, none of these account for the full stochastic nature of the motion, nor do they systematically derive a coarse-grained equation from microscopic parameters[41]. A systematic derivation of a rolling diffusion coefficient would involve a few additional mathematical subtleties beyond those that occur here, such as including binding rates with spatial dependencies to account for the variable separation between surfaces[76,77], but we may nevertheless expect similar parameters (such as spring relaxation times and unbinding rates) to discriminate between rolling and other modes of motion.

Going further, other effects that could be studied include the details of binding kinetics, *e.g.* non-exponential kinetics in DNA

hybridization[78–80], which could also impact the long time response[42]; mobility of the leg roots, corresponding to fluidity of the bilayer[10,81]; and out-of-equilibrium effects, such as white blood cells streaming in blood flow[5,77], active stepping of molecular motors[49,82,83], or proteins that actively cleave bonds on influenza A[17,36]. Accounting for such effects would require adapting bond dynamics to include increased bond rigidity or bond lifetime in flow[3,84–88]; binding kinetics coupled to the number of bonds[47,49]; or memory effects associated with dead zones created by cleaved bonds[26,36,52]. Importantly, such improvements require carefully adapting binding rates to preserve detailed balance and physical constraints[32,76].

Furthermore, detailed hydrodynamic effects may be important to describe certain kinds of nanocaterpillar dynamics. We have accounted for hydrodynamics via the bare friction coefficients (Γ, γ), but these coefficients themselves are coarse-grained, and in reality depend on the distance of a nanocaterpillar to a surface[67] and are coupled to the details of the polymer leg mesh. Indeed, elasticity of the polymer mesh could modify the particle's mobility near the interface, as was predicted for elastic membranes[89,90]. A more detailed description of the hydrodynamic flow near a nanocaterpillar could help shed light on other systems where mobility through fluid is mediated by slender legs, such as for the Vampire amoeba[91].

Beyond its biophysical details, nanocaterpillar motion resonates with other fields where mobility is determined through adhesive contacts. For example, solid state sliding friction is created by bonds breaking between atoms. Close neighbor interactions between bonds, originating from mechanical interactions, can result in dramatic avalanches of bond breaking that change the sliding motion[92,93]. Similar correlations between nearby bonds could be at play in some nanocaterpillars. For example, in white blood cells, membrane tension mediates bond-bond interactions[47,48]. It is therefore interesting to speculate whether avalanches of bond unbinding could also occur for nanocaterpillar systems. Overall, the mathematical framework of coarse-graining is well suited to explore how microscopic features determine macroscopic modes of motion for nanocaterpillars and could facilitate predictive capacity for materials design and biophysical systems.

Author Contributions

S.M. derived the mathematical framework, and solved it in all cases; acquired biological data for the Ashby chart; designed and analyzed the numerical simulations; found predictions for the diffusion of DNA-coated colloids. J.A.Z. synthesized DNA-coated colloids, conducted the experiments, and analyzed the experimental data to find diffusion coefficients. M.H.C. supervised the project. S.M. and M.H.C. wrote the paper.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Appendix

Appendix A: Projection of the dynamics in the bound state

To project the stochastic dynamics Eqns. (1) and (2) in the bound case we use a formalism (and notations) similar to Ref. 32; see also[60,94]. This projection consists in using stiff springs to impose each constraint, and considering the limit where the spring constants go to infinity. The resulting projected equations can be obtained by directly pursuing the steps below (without redoing the reasoning with stiff springs).

We start from stochastic equations in the (x,l) space and seek to project them on the constraint manifold, defined by the constraint $q(x,l) = x + l - x_{\mathrm{r}} = 0$. The constraint matrix is therefore

$$C = (\nabla q)^T = \begin{pmatrix} 1 & 1 \end{pmatrix}. \tag{28}$$

We obtain the projector

$$P = I - C^T(CC^T)^{-1}C = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{29}$$

Initially the dynamics of $X = (x,l)^T$ may be written as

$$\frac{dX}{dt} = -\tilde{\Gamma}^{-1}\nabla\mathscr{U}(X) + \sqrt{2k_BT\tilde{\Gamma}^{-1}}\eta_{xl}(t) \tag{30}$$

where the potential $\mathscr{U}(X) = kl^2/2$, the noise $\eta_{xl} = (\eta_x, \eta_l)^T$ and the friction matrix is

$$\tilde{\Gamma} = \begin{pmatrix} \Gamma & 0 \\ 0 & \gamma \end{pmatrix}. \tag{31}$$

The projected friction and its Moore-Penrose pseudo-inverse are

$$\Gamma_P = P\tilde{\Gamma}P \quad = \frac{\Gamma+\gamma}{4}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \tag{32}$$

$$\Gamma_P^\dagger \quad = \frac{1}{\Gamma+\gamma}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \tag{33}$$

with a square root

$$\sigma_P = \sqrt{\Gamma_P^\dagger} = \frac{1}{\sqrt{\Gamma+\gamma}}\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}. \tag{34}$$

We obtain the projected dynamics

$$\frac{dX}{dt} = \quad -\Gamma_P^\dagger\nabla\mathscr{U}(X) + \sqrt{2k_BT\Gamma_P^\dagger}\eta_{xl}(t) \tag{35}$$

where additional terms are needed if C is not constant over the constraint manifold[60,94]. One can check that this exactly yields

Table 2 Summary of different models and their effective long time friction. The 1-leg case corresponds to a system where the leg's center of mass is fixed on the particle. Apart from the 1-leg case, we ignore differences between Γ and $\tilde{\Gamma}$ to simplify notations.

Model	Sketch	Result
Main geometries		
1-arm		$\frac{1}{\Gamma_{\text{eff}}} = \frac{p_0}{\Gamma_0} + \frac{p_1}{\Gamma_1}$, $\Gamma_0 = \Gamma$, $\Gamma_1 = \Gamma + \gamma_{\text{eff}}$, $\gamma_{\text{eff}} = k\left(\frac{1}{q_{\text{off}}} + \frac{\gamma}{k}\frac{q_{\text{on}}}{q_{\text{off}}}\right)$
1-leg		$\frac{1}{\Gamma_{\text{eff}}} = \frac{p_0}{\Gamma_0} + \frac{p_1}{\Gamma_1}$, $\Gamma_0 = \tilde{\Gamma}$, $\Gamma_1 = \tilde{\Gamma} + \gamma_{\text{eff}}$, $\tilde{\Gamma} = \Gamma + \gamma$
N-legs		$\frac{1}{\Gamma_{\text{eff}}} = \sum_{n=0}^{N} \frac{p_n}{\Gamma_n}$, $p_n = \binom{N}{n}\frac{q_{\text{off}}^{N-n}q_{\text{on}}^{n}}{(q_{\text{off}}+q_{\text{on}})^N}$, $\Gamma_n \underset{N\gg 1}{\simeq} \Gamma + n\gamma_{\text{eff}}$
Inertial dynamics		
1-leg, inertia		$\Gamma_{\text{eff}} = p_0\Gamma_0 + p_1\Gamma_1$, $\Gamma_0 = \Gamma$, $\Gamma_1 = \Gamma + \gamma_{\text{eff}}$
Limit regimes		
Small legs	$\gamma/\Gamma \ll 1$	$\frac{1}{\Gamma_{\text{eff}}} = \frac{p_0}{\Gamma} + \frac{p_1}{\Gamma}\left(1 - \frac{\gamma_{\text{eff}}}{\Gamma}\right)$
Fast legs	$k/\gamma \gg q$	$\frac{1}{\Gamma_{\text{eff}}} = \frac{p_0}{\Gamma_0} + \frac{p_1}{\Gamma_1}$, $\Gamma_0 = \Gamma$, $\Gamma_1 = \gamma + \frac{k}{q_{\text{off}}}$
Fast binding	$k/\gamma \ll q$	$\frac{1}{\Gamma_{\text{eff}}} = \frac{p_0}{\Gamma_0} + \frac{p_1}{\Gamma_1}$, $\Gamma_0 = \Gamma$, $\Gamma_1 = \gamma + k\left(\frac{\gamma}{k}\frac{q_{\text{on}}}{q_{\text{off}}}\right)$
Extended geometries		
1-arm, 1-leg		$\frac{1}{\Gamma_{\text{eff}}} = \frac{p_0}{\Gamma_0} + \frac{p_1}{\Gamma_1}$, $\Gamma_0 = \Gamma$, $\Gamma_1 = \Gamma + \frac{1}{2}\gamma_{\text{eff}}$
M-arms, N-legs		$\begin{cases} \frac{1}{\Gamma_{\text{eff}}} = \sum_{n=0}^{N} \frac{p_n}{\Gamma_n}, \Gamma_n = \Gamma + n\gamma_{\text{eff,n}}(M,N), \\ (\gamma_{\text{eff,n}}(M,N))^{-1} \simeq (\gamma_{\text{eff,M,n}})^{-1} + (\gamma_{\text{eff,N,n}})^{-1}, \gamma_{\text{eff,P,n}} = \gamma + k\left(\frac{1}{q_{\text{off}}} + \frac{\gamma}{k}\frac{(P-n)q_{\text{on}}}{q_{\text{off}}}\right) \end{cases}$

the bound dynamics Eq. (3), with $\eta = \eta_x$ (this decomposition of the noise is not unique but this does not impact the dynamics in a weak sense).

Appendix B: Numerical simulations

Stochastic simulations of particle and leg dynamics are conducted using a custom made Fortran 90 routine. Fast random number generation is performed according to a Mersenne twister algorithm. Normally distributed random numbers are used for particle displacement while uniformly distributed random numbers are used to determine binding events. Equations are simulated in their non-dimensional form. The step dt was chosen to be much small than all other time scales of the system. Typically $dt = \frac{1}{100}\min\left(\frac{q_{\text{on}}\Gamma}{k}, \frac{q_{\text{on}}\Gamma}{k}, \frac{\gamma}{\Gamma}\right)$. The system is simulated for $N_T = 10^8$ time steps, and the simulation is repeated over $N_{\text{runs}} = 100$ independent runs (with renewed random number seed).

To simulate binding and unbinding events, for each leg, at each time step, we choose a random number R uniformly distributed between 0 and 1 and then:

- if the leg is bound, and if $R > q_{\text{off}}dt$ then the leg becomes unbound. Otherwise it remains bound.

- if the leg is unbound, and if $R > q_{\text{on}}dt$ then the leg becomes bound. Otherwise it remains unbound.

This simulation routine approximates well the exponential binding dynamics expected from the continuous equations since $dt \ll q_{\text{off}}^{-1}, q_{\text{on}}^{-1}$. To simulate all other stochastic equations we use a standard Euler-Maruyama discretization.

The particle position x is saved every 10^4 time steps, and the mean squared displacement $< (x(t+t_0) - x(t_0))^2 >_{t_0}$ (averaged over initial times t_0) is computed up to $N_T/100 = 10^6$ time steps. The effective diffusion coefficient for each run $D_{\text{eff,i}}$ is obtained from the analytical least square regression of $< (x(t+t_0) - x(t_0))^2 >_{t_0}$ with time. The average value over the runs $D_{\text{eff}} = \frac{1}{N_{\text{runs}}}\sum_i D_{\text{eff,i}}$ is retained as the effective long time diffusion coefficient. The standard deviation of $D_{\text{eff,i}}$ allows to draw error bars in all simulation plots.

Notes and references

1 M. Mammen, S.-K. Choi and G. M. Whitesides, *Angewandte Chemie International Edition*, 1998, **37**, 2754–2794.
2 P. C. Bressloff and J. M. Newby, *Reviews of Modern Physics*, 2013, **85**, 135.
3 D. A. Hammer, *Journal of biomechanical engineering*, 2014, **136**, 021006.
4 W. B. Rogers, W. M. Shih and V. N. Manoharan, *Nature Reviews Materials*, 2016, **1**, 1–14.
5 R. Alon and S. Feigelson, Seminars in immunology, 2002, pp. 93–104.
6 K. Ley, C. Laudanna, M. I. Cybulsky and S. Nourshargh, *Nature Reviews Immunology*, 2007, **7**, 678–689.
7 C. Korn and U. Schwarz, *Physical review E*, 2008, **77**, 041904.
8 Y. Zhang, A. McMullen, L.-L. Pontani, X. He, R. Sha, N. C. Seeman, J. Brujic and P. M. Chaikin, *Nature communications*, 2017, **8**, 1–7.
9 L.-L. Pontani, I. Jorjadze and J. Brujic, *Biophysical journal*, 2016, **110**, 391–399.

10 S. Merminod, J. R. Edison, H. Fang, M. F. Hagan and W. B. Rogers, *Nanoscale*, 2021.

11 R. J. Macfarlane, B. Lee, M. R. Jones, N. Harris, G. C. Schatz and C. A. Mirkin, *science*, 2011, **334**, 204–208.

12 D. J. Lewis, L. Z. Zornberg, D. J. Carter and R. J. Macfarlane, *Nature materials*, 2020, **19**, 719–724.

13 J.-G. Park, S.-H. Kim, S. Magkiriadou, T. M. Choi, Y.-S. Kim and V. N. Manoharan, *Angewandte Chemie International Edition*, 2014, **53**, 2899–2903.

14 M. He, J. P. Gales, É. Ducrot, Z. Gong, G.-R. Yi, S. Sacanna and D. J. Pine, *Nature*, 2020, **585**, 524–529.

15 R. Merindol, N. Martin, T. Beneyton, J.-C. Baret and S. Ravaine, *Advanced Functional Materials*, 2021, **31**, 2010396.

16 C. R. Bilchak, M. Jhalaria, Y. Huang, Z. Abbas, J. Midya, F. M. Benedetti, D. Parisi, W. Egger, M. Dickmann, M. Minelli *et al.*, *ACS nano*, 2020, **14**, 17174–17183.

17 T. Sakai, S. I. Nishimura, T. Naito and M. Saito, *Scientific reports*, 2017, **7**, 1–11.

18 T. Sakai, H. Takagi, Y. Muraki and M. Saito, *Journal of virology*, 2018, **92**, e01522–17.

19 M. Müller, D. Lauster, H. H. Wildenauer, A. Herrmann and S. Block, *Nano letters*, 2019, **19**, 1875–1882.

20 T. D. Allen, J. Cronshaw, S. Bagley, E. Kiseleva and M. W. Goldberg, *Journal of cell science*, 2000, **113**, 1651–1659.

21 I. V. Aramburu and E. A. Lemke, Seminars in cell & developmental biology, 2017, pp. 34–41.

22 B. Fogelson and J. P. Keener, *Biophysical journal*, 2018, **115**, 108–116.

23 R. Alon, S. Chen, K. D. Puri, E. B. Finger and T. A. Springer, *The Journal of cell biology*, 1997, **138**, 1169–1180.

24 R. Shrivastava, A. Rai, M. Salapaka and S. Sivaramakrishnan, *Biochemistry*, 2019, **58**, 4721–4725.

25 A. K. Dasanna and U. S. Schwarz, *Soft matter*, 2018, **14**, 9061–9070.

26 K. Yehl, A. Mugler, S. Vivek, Y. Liu, Y. Zhang, M. Fan, E. R. Weeks and K. Salaita, *Nature nanotechnology*, 2016, **11**, 184–190.

27 R. Karnik, S. Hong, H. Zhang, Y. Mei, D. G. Anderson, J. M. Karp and R. Langer, *Nano letters*, 2008, **8**, 1153–1158.

28 Y.-Y. Wang, A. Kannan, K. L. Nunn, M. A. Murphy, D. B. Subramani, T. Moench, R. Cone and S. K. Lai, *Mucosal immunology*, 2014, **7**, 1036–1044.

29 A. Hensley, W. M. Jacobs and W. B. Rogers, *arXiv preprint arXiv:2105.14631*, 2021.

30 D. J. Lewis, P. A. Gabrys and R. J. Macfarlane, *Langmuir*, 2018, **34**, 14842–14850.

31 Y. Wang, Y. Wang, X. Zheng, É. Ducrot, J. S. Yodh, M. Weck and D. J. Pine, *Nature communications*, 2015, **6**, 1–8.

32 M. Holmes-Cerfon, *Physical Review E*, 2016, **94**, 052112.

33 P. K. Jana and B. M. Mognetti, *Physical Review E*, 2019, **100**, 060601.

34 C. Loverdo, O. Benichou, R. Voituriez, A. Biebricher, I. Bonnet and P. Desbiolles, *Physical review letters*, 2009, **102**, 188101.

35 P. E. Hamming, N. J. Overeem and J. Huskens, *Chemical science*, 2019, **11**, 27–36.

36 M. D. Vahey and D. A. Fletcher, *Elife*, 2019, **8**, e43764.

37 K. Ramesh, R. Thaokar, J. R. Prakash and R. Prabhakar, *Physical Review E*, 2015, **91**, 022302.

38 F. Cui, S. Marbach, J. Zheng, M. Holmes-Cerfon and D. S. Pine, *to appear*, 2021.

39 Q. Xu, L. Feng, R. Sha, N. Seeman and P. Chaikin, *Physical review letters*, 2011, **106**, 228102.

40 C. Etchegaray and N. Meunier, *Journal of mathematical biology*, 2019, **79**, 1665–1697.

41 F. Ziebert and I. M. Kulić, *Physical Review Letters*, 2021, **126**, 218101.

42 N. A. Licata and A. V. Tkachenko, *Physical Review E*, 2007, **76**, 041405.

43 S. Bose, S. K. Das, J. M. Karp and R. Karnik, *Biophysical journal*, 2010, **99**, 3870–3879.

44 A. Kowalewski, N. R. Forde and C. S. Korosec, *The Journal of Physical Chemistry B*, 2021.

45 K. E. Caputo and D. A. Hammer, *Biophysical journal*, 2005, **89**, 187–200.

46 B. Grec, B. Maury, N. Meunier and L. Navoret, *Journal of theoretical biology*, 2018, **452**, 35–46.

47 S. Klumpp and R. Lipowsky, *Proceedings of the National Academy of Sciences*, 2005, **102**, 17284–17289.

48 S. F. Fenz, T. Bihr, D. Schmidt, R. Merkel, U. Seifert, K. Sengupta and A.-S. Smith, *Nature physics*, 2017, **13**, 906–913.

49 C. E. Miles, S. D. Lawley and J. P. Keener, *SIAM Journal on Applied Mathematics*, 2018, **78**, 2511–2532.

50 C. P. Goodrich, M. P. Brenner and K. Ribbeck, *Nature communications*, 2018, **9**, 1–8.

51 B. Fogelson and J. P. Keener, *SIAM Journal on Applied Mathematics*, 2019, **79**, 1405–1422.

52 C. S. Korosec, L. Jindal, M. Schneider, I. C. de la Barca, M. J. Zuckermann, N. R. Forde and E. Emberly, *Soft Matter*, 2021, **17**, 1468–1479.

53 G. Pavliotis and A. Stuart, *Multiscale methods: averaging and homogenization*, Springer Science & Business Media, 2008.

54 J. P. Lee-Thorp and M. Holmes-Cerfon, *Soft matter*, 2018, **14**, 8147–8159.

55 X. Bian, C. Kim and G. E. Karniadakis, *Soft Matter*, 2016, **12**, 6331–6346.

56 M. Rubinstein, R. H. Colby *et al.*, *Polymer physics*, Oxford university press New York, 2003, vol. 23.

57 E. Miller, T. Garcia, S. Hultgren and A. F. Oberhauser, *Biophysical journal*, 2006, **91**, 3848–3856.

58 R. Y. Lim, N.-P. Huang, J. Köser, J. Deng, K. A. Lau, K. Schwarz-Herion, B. Fahrenkrog and U. Aebi, *Proceedings of the National Academy of Sciences*, 2006, **103**, 9512–9517.

59 P. Varilly, S. Angioletti-Uberti, B. M. Mognetti and D. Frenkel, *The Journal of chemical physics*, 2012, **137**, 094108.

60 G. Ciccotti, T. Lelievre and E. Vanden-Eijnden, *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 2008, **61**, 371–

408.

61 C. W. Gardiner *et al.*, *Handbook of stochastic methods*, springer Berlin, 1985, vol. 3.

62 L. Leibler, M. Rubinstein and R. H. Colby, *Macromolecules*, 1991, **24**, 4701–4707.

63 W. Cao, C. Dong, S. Kim, D. Hou, W. Tai, L. Du, W. Im and X. F. Zhang, *Biophysical journal*, 2021, **120**, 1011–1019.

64 M. Ponga, *Scientific reports*, 2020, **10**, 1–7.

65 A. Harris, G. Cardone, D. C. Winkler, J. B. Heymann, M. Brecher, J. M. White and A. C. Steven, *Proceedings of the National Academy of Sciences*, 2006, **103**, 19123–19127.

66 V. Reiter-Scherer, J. L. Cuellar-Camacho, S. Bhatia, R. Haag, A. Herrmann, D. Lauster and J. P. Rabe, *Biophysical journal*, 2019, **116**, 1037–1048.

67 H. Brenner, *Chemical Engineering Science*, 1961, **16**, 242–251.

68 S. T. Milner, T. A. Witten and M. E. Cates, *Macromolecules*, 1988, **21**, 2610–2619.

69 H. Chen, S. P. Meisburger, S. A. Pabit, J. L. Sutton, W. W. Webb and L. Pollack, *Proceedings of the National Academy of Sciences*, 2012, **109**, 799–804.

70 J. X. Zhang, J. Z. Fang, W. Duan, L. R. Wu, A. W. Zhang, N. Dalchau, B. Yordanov, R. Petersen, A. Phillips and D. Y. Zhang, *Nature chemistry*, 2018, **10**, 91–98.

71 S. Marbach and M. Holmes-Cerfon, *arXiv preprint arXiv:2112.05266*, 2021.

72 J. S. Oh, Y. Wang, D. J. Pine and G.-R. Yi, *Chemistry of Materials*, 2015, **27**, 8337–8344.

73 C. Fröhner and F. Noé, *The Journal of Physical Chemistry B*, 2018, **122**, 11240–11250.

74 S. Xu, J. Zhan, B. Man, S. Jiang, W. Yue, S. Gao, C. Guo, H. Liu, Z. Li, J. Wang *et al.*, *Nature communications*, 2017, **8**, 1–10.

75 K.-C. Chang, D. F. Tees and D. A. Hammer, *Proceedings of the National Academy of Sciences*, 2000, **97**, 11262–11267.

76 C. Korn and U. Schwarz, *The Journal of chemical physics*, 2007, **126**, 03B605.

77 U. S. Schwarz and R. Alon, *Proceedings of the National Academy of Sciences*, 2004, **101**, 6940–6945.

78 M. I. Wallace, L. Ying, S. Balasubramanian and D. Klenerman, *Proceedings of the National Academy of Sciences*, 2001, **98**, 5584–5589.

79 W. B. Rogers, T. Sinno and J. C. Crocker, *Soft Matter*, 2013, **9**, 6412–6417.

80 K.-T. Wu, L. Feng, R. Sha, R. Dreyfus, A. Y. Grosberg, N. C. Seeman and P. M. Chaikin, *Physical Review E*, 2013, **88**, 022304.

81 N. Sarpangala and A. Gopinathan, *bioRxiv*, 2021.

82 S. A. McKinley, A. Athreya, J. Fricks and P. R. Kramer, *Journal of theoretical biology*, 2012, **305**, 54–69.

83 C. S. Peskin and T. C. Elston, *SIAM Journal on Applied Mathematics*, 2000, **60**, 842–867.

84 G. I. Bell, *Science*, 1978, **200**, 618–627.

85 P. S. Doyle, B. Ladoux and J.-L. Viovy, *Physical review letters*, 2000, **84**, 4769.

86 S. Chen and T. A. Springer, *Proceedings of the National Academy of Sciences*, 2001, **98**, 950–955.

87 D. E Leckband, *Current Protein and Peptide Science*, 2011, **12**, 743–759.

88 S. Rakshit and S. Sivasankar, *Physical Chemistry Chemical Physics*, 2014, **16**, 2211–2223.

89 A. Daddi-Moussa-Ider, A. Guckenberger and S. Gekle, *Physical Review E*, 2016, **93**, 012612.

90 V. Bertin, Y. Amarouchene, E. Raphael and T. Salez, *arXiv preprint arXiv:2104.00900*, 2021.

91 S. Hess, N. Sausen and M. Melkonian, *PLoS One*, 2012, **7**, e31165.

92 T. W. de Geus, M. Popović, W. Ji, A. Rosso and M. Wyart, *Proceedings of the National Academy of Sciences*, 2019, **116**, 23977–23983.

93 W. Ji, T. W. de Geus, E. Agoritsas and M. Wyart, *arXiv preprint arXiv:2106.13153*, 2021.

94 M. Holmes-Cerfon, S. J. Gortler and M. P. Brenner, *Proceedings of the National Academy of Sciences*, 2013, **110**, E5–E14.