Differential Dynamic Microscopy

Differential Dynamic Microscopy: a High-Throughput Method for Characterizing the Motility of Microorganism

Vincent A. Martinez [1], Rut Besseling [1], Ottavio A. Croze [2], Julien Tailleur [3], Mathias Reufer [1], Jana Schwarz-Linek [1], Laurence G. Wilson [1,4] Martin A. Bees [2], Wilson C. K. Poon [1]

(1) SUPA, School of Physics and Astronomy, University of Edinburgh Mayfield Road, Edinburgh EH9 3JZ, UK

(2) School of Mathematics and Statistics, University of Glasgow Glasgow G12 8QW, UK

(3) Université Paris Diderot, MSC Sorbonne Paris Cité UMR 7057 CNRS, F75205 Paris, France

(4) Rowland Institute at Harward Cambridge, Massachusetts, 02142, USA

February 9, 2012

[1]Corresponding author. email: vincent.martinez@ed.ac.uk

Abstract

We present a fast, high-throughput method for characterizing the motility of microorganisms in 3D based on standard imaging microscopy. Instead of tracking individual cells, we analyse the spatio-temporal fluctuations of the intensity in the sample from time-lapse images and obtain the intermediate scattering function (ISF) of the system. We demonstrate our method on two different types of microorganisms: bacteria, both smooth swimming (run only) and wild type (run and tumble) *Escherichia coli*, and the bi-flagellate alga *Chlamydomonas reinhardtii*. We validate the methodology using computer simulations and particle tracking. From the ISF, we are able to extract (i) for *E. coli*: the swimming speed distribution, the fraction of motile cells and the diffusivity, and (ii) for *C. reinhardtii*: the swimming speed distribution, the amplitude and frequency of the oscillatory dynamics. In both cases, the motility parameters are averaged over $\sim 10^4$ cells and obtained in a few minutes.

Key words: Particle Tracking; Bacteria; Algae; Microswimmer

Introduction

The motility of micro-organisms, both prokaryotes and eukaryotes, is important in biology and medicine. For example, the virulence of the bacterial pathogen *Helicobacter pylori* depends on its locomotion through host epithelial mucosa (1); the phototaxis of *Chlamydomonas reinhardtii* and similar algae, and therefore their photosynthesis, is predicated on motility. Animal reproduction relies on motile spermatozoa. In all cases, organisms with typical linear dimension R in the range $0.5\ \mu m \lesssim R \lesssim 10\ \mu m$ swim with speeds of $v \sim 10 - 100\ \mu m/s$, so that the Reynolds number, $\mathrm{Re} = \rho v R / \eta$ (ρ and η are the density and viscosity of the liquid medium) are vanishingly small: in aqueous media, $\mathrm{Re} \lesssim 10^{-3}$. In this 'creeping flow' regime, micro-organisms have evolved a variety of strategies for generating the non-reciprocating motion necessary for propulsion, e.g. by rotating or beating one or more flagella. The resulting motility phenotypes are hugely varied. Amongst these, the motility of the enteric bacterium *Escherichia coli* is perhaps best understood (2).

In wild type (WT) *E. coli*, a single cell (roughly a 1 μm $\times$ 2 μm spherocylinder) is equipped with six to ten helical flagella (each $6 - 10\ \mu$m long). When these rotate counterclockwise (CCW, viewed from flagella to cell body), the individual flagella bundle together and propel the cell forward in a straight line (known as a 'run'), with directional deviations brought about by orientational Brownian motion. Every second or so, one or more of the flagella reverse to clockwise (CW) rotation briefly, unbundle, and the cell undergoes large-angle re-orientation known as 'tumble'. When all motors rotate CCW again, the cell begins a new run in an essentially random direction. Such 'run and tumble' gives rise to a random walk, which the cell can bias by decreasing the tumble frequency when running in the direction of a favorable chemical gradient (chemotaxis).

Such detailed information in *E. coli* or other micro-organisms can only be obtained by single-cell tracking. On the other side, tracking is laborious, and seldom averages over more than a few hundred cells, limiting the statistical accuracy. Moreover, since three-dimensional (3D) tracking requires specialized equipment (3–5), the usual practice is to rely on measuring 2D projections in a single imaging plane, which further limits statistical accuracy because of cells moving out of the plane.

We recently proposed (6) that differential dynamic microscopy (DDM) can be used for characterizing the motility of micro-organisms. DDM is fast,

fully 3D and provides statistics from much larger samples than tracking, allowing averaging over $\sim 10^4$ cells in a few minutes, using standard microscopy imaging. We demonstrated DDM for swimming wild-type *E. coli*, and validated the method using simulations.

In this paper, we explain the full details of our method, discuss its limitations, and provide in-depth justification for the approximations made using computer simulation and tracking. We applied the method to a smooth-swimming (run only) mutant of *E. coli*, investigated its use on the WT (run and tumble) in more detail, and extended it for the first-time to study the swimming of bi-flagellate WT alga *C. reinhardtii*, a completely different microorganism in terms of time scale, length scale, and swimming dynamics. Our results should give the basis for generalizing DDM to many other biologically and medically important micro-swimmers, including spermatozoa.

1   Differential Dynamic Microscopy

The key idea of DDM (6–8) is to characterize the motility of a population of particles (colloids or micro-organisms) by studying the temporal fluctuations of the local number density of particles over different length scales via image analysis. It yields the same quantity accessed in dynamic light scattering (DLS), the 'intermediate scattering function' (ISF). The advantage of DDM is that the required range of length scale to study microorganism motility, such as bacteria or algae, is easily accessible in contrast to DLS (9).

In DDM, one takes time-lapse images of particles, described by the intensity $I(\vec{r}, t)$ in the image plane, where $\vec{r}$ is pixel position and t is time. As particles move, $I(\vec{r}, t)$ fluctuates with time. The statistics of these fluctuations contain information about the particle motions. To quantify these fluctuations, DDM measures the Differential Image Correlation Function (DICF), $g(\vec{q}, \tau)$, i.e. the square modulus of the Fourier transform of the difference of two images separated by τ in time

$$g(\vec{q}, \tau) = \left\langle |I(\vec{q}, t + \tau) - I(\vec{q}, t)|^2 \right\rangle_t. \tag{1}$$

Here, $\langle ... \rangle_t$ means average over the initial time t, and $I(\vec{q}, t)$ is the Fourier transform of $I(\vec{r}, t)$, which picks out the component in the image $I(\vec{r}, t)$ that varies sinusoidally with wavelength $2\pi/q$ in the direction $\vec{q}$. In isotropic samples (no preferred direction of motion), the relevant variable is the magnitude

q of $\vec{q}$. It can be shown (6–8) that $g(q, \tau)$ is related to the ISF, $f(q, \tau)$, by

$$g(q, \tau) = A(q)[1 - f(q, \tau)] + B(q), \tag{2}$$

where $A(q)$ depends on the optics, the individual particle shape and the sample's structure, and $B(q)$ represents the camera noise. For independent particles, the ISF is given by (10)

$$f(q, \tau) = \left\langle e^{-i\vec{q}\cdot\Delta\vec{r}(\tau)} \right\rangle, \tag{3}$$

with $\Delta\vec{r}(\tau)$ the single-particle displacement and $\langle ... \rangle$ an average over all particles.

Eq. 3 shows that $f(q, 0) = 1$ and $f(q, \tau \to \infty) = 0$. This decay of the ISF from unity to zero reflects the fact that particle configurations (and therefore images) separated by progressively longer delay time, τ, become more decorrelated due to particle motion. The precise manner in which $f(q, \tau)$ decays contains information on these motions on the length scale $2\pi/q$. The analytic form of the ISF is known for a number of ideal systems (10). As an example, for identical spheres undergoing purely Brownian motion with diffusion coefficient D, $f(q, \tau) = e^{-Dq^2\tau}$. On the other hand, for an isotropic population of straight swimmers in 3D with speed v, $f(q, \tau) = \sin(qv\tau)/qv\tau$.

Figure 1 shows the calculated $f(q, \tau)$ for (i) diffusing spheres with about the same volume as a typical *E. coli* cell; (ii) isotropic swimmers with a speed distribution $P(v)$ typical of *E. coli*; and (iii) for a mixture of these (11) (see later, Eq. 4) at $q = 1$ μm^{-1}. The curve for the mixture illustrates the utility of plotting the ISF against $\log \tau$: it renders obvious that there are two processes, a fast one due to swimming that decorrelates density (or, equivalently, intensity) fluctuations over $\sim 10^{-1}$ s, and a slower process due to diffusion that decorrelates over ~ 1 s (at this q). Their fractional contributions can be visually estimated to be $\approx 7 : 3$.

2   Methods

2.1   Samples

E. coli AB1157 (WT and $\Delta cheY$ strains (12)) were grown in Luria-Bertani broth (LB) at 30°C and shaken at 200 rpm, harvested in the exponential phase, washed three times by careful filtration (0.45 μm filter) with and resuspended in motility buffer (6.2 mM K_2HPO_4, 3.8 mM KH_2PO_4, 67 mM

NaCl, 0.1 mM EDTA, pH=7.0) to optical density 0.3 (at 600 nm), corresponding to $\approx 5 \times 10^8$ cells/ml, and $\approx 0.06\%$ by cell volume. Care was taken throughout to minimize damage to flagella. A $\approx 400~\mu$m deep flat glass cell was filled with $\approx 150~\mu$l of cell suspension, sealed, and observed at $22\pm1°$C. Swimming behavior was constant over a 15 minute period.

Batch cultures of WT *C. reinhardtii* (CCAP 11/32B) were grown on 3N Bold's medium (15), and concentrated in cotton using gravitaxis (16). Concentrated cell stock was diluted in growth media to optical density 0.175 (at 590 nm), corresponding to 1.4×10^6 cells/ml, and $\approx 0.002\%$ by volume of cells. Cells were observed at $22\pm1°$C in the same glass cells used for *E. coli* under a 600 nm long pass filter (Cokin) to avoid stimulating the phototaxis (17). The sample dimensions are sufficiently large to avoid boundary effects and small enough to avoid bioconvection or thermal convection (5). The algae motility was constant for 20 minutes. In all cases, we waited at least one minute before capturing images to avoid drift due to mixing flows.

2.2 Differential dynamic microscopy

We collected movies using a Nikon Eclipse Ti inverted microscope and a high-speed camera (Mikrotron MC 1362) connected to a PC with a frame grabber card with 1GB onboard memory. The CMOS pixel size (14 μm $\times$ 14 μm) and magnification determine the inverse pixel size k (in pixel/ μm) in the image plane, which, together with the image size L (in pixels), define the spatial sampling frequency ($q_{\min} = 2\pi k/L$). For bacteria, $10\times$ phase-contrast movies were acquired at $L = 500$, which gives $k = 0.712~\mu$m^{-1} and a q range of $0.01 \lesssim q \lesssim 2.2~\mu$m^{-1}. This allows the imaging of $\sim 10^4$ bacteria cells at a bulk cell density of 5×10^8 cells/ml in a 0.49 mm^2 field of view with a depth of field $\delta \approx 40~\mu$m, over 38 s at a frame rate of 100 fps. For algae, $4\times$ bright-field movies were acquired at $L = 500$, giving $k = 0.285~\mu$m^{-1} and a q range of $0.004 \lesssim q \lesssim 0.9~\mu$m^{-1}, and allowing the imaging of $\sim 10^4$ algae cells at a bulk cell density of 1.4×10^6 cells/ml in a 3.2 mm^2 field of view with $\delta \approx 200~\mu$m, over 3.8 s at a frame rate of 1000 fps. We imaged at $\approx 200~\mu$m from the bottom of a 400 μm thick glass capillary cell to minimize wall effects.

2.3 Data reduction and fitting

The image processing and fitting analysis are easily automated and details are given below (18). Figure 1b illustrates how we obtain the DICF, $g(q, \tau)$, from the movies. For a given delay time τ, the difference images $D_i(\vec{r}, \tau) = I(\vec{r}, t_i + \tau) - I(\vec{r}, t_i)$ are calculated for a set of N different initial times t_i (typically $i = 1, 4, 7, \ldots, 313$). After computing the fast Fourier transform, $F_{Di}(\vec{q}, \tau)$, of each $D_i(\vec{r}, \tau)$ and calculating the squared modulus, $|F_{Di}(\vec{q}, \tau)|^2$, we average over initial times t_i, giving $g(\vec{q}, \tau) = \langle |F_{Di}(\vec{q}, \tau)|^2 \rangle_i$, to improve the signal-to-noise ratio (averaged image appears less grainy, Fig. 1b).

For isotropic swimmers, $g(\vec{q}, \tau)$ is azimuthally symmetric and can be azimuthally averaged to give $g(q, \tau) = \langle g(\vec{q}, \tau) \rangle_{\vec{q}}$. We linearly interpolate between four adjacent points in discrete $\vec{q}$-space to find values for $g(\vec{q}, \tau)$ along a circle of equidistant points with radius q. The finite image size causes numerical artefacts (8) mainly along the horizontal and vertical center lines of the image $g(\vec{q}, \tau)$; these are reduced by omitting the values for $q_x = 0$ and $q_y = 0$ during the azimuthal averaging. The procedure is repeated for a set of delay times τ to obtain the full time-evolution of $g(q, \tau)$. Calculations were done in LabView (National Instruments) on an 4-core PC (3 GHz Quad core, 3 GB RAM) . Processing $\lesssim 4000$ frames with $L = 500$ and averaging over ≈ 100 initial times t_i take ≈ 5 min.

We fitted independently each $g(q, \tau)$ to Eq. 2 using the appropriate parametrized model for $f(q, \tau)$. At each q, non-linear least-squares fitting based on χ^2 minimization using the Levenberg-Marquardt algorithm and the 'all-at-once-fitting' procedure in IGOR Pro (WaveMetrics) returns $A(q)$, $B(q)$ and motility parameters.

2.4 Simulation

We carried out Brownian dynamics simulations in 3D of non-interacting point particles ('bacteria') at a concentration and in a sample chamber geometry directly comparable to our experiments, using periodic boundary conditions to keep the bulk density of swimmers constant. Each particle has a drift velocity whose direction and magnitude were chosen from uniform and Schulz distributions, respectively. During a tumble event, a wild-type (run and tumble) swimmer undergoes standard Brownian diffusion and a new swimming direction is chosen uniformly at random after each 'tumble'. The swimming speed is constant for each bacterium.

From these simulations, we constructed 2D pixelated 'images'. Particles in a slice of thickness d, centered at $z = 0$, contribute to the image. A particle at (x, y, z) is 'smeared' into an 'image' covering the pixel containing (x, y) and the 8 neighboring pixels. To define the image contrast of a bacterium, which depends on z, we used the experimentally measured z-contrast function $c(z)$ (Section 3.4). This mimics the finite depth of field in a microscope.

2.5 Tracking

Both experimental and simulated data were analyzed using standard particle tracking (19) giving the 2D tracks $r_{2D}(t)$. We used inverted $20\times$ videos of *E. coli* with bright cells of ≈ 3 pixels on a dark background and a running average of 3 frames to improve the features' signal over noise ratio. In simulations, equivalent $10\times$ videos could be tracked due to the absence of noise. In all cases, > 400 features were identified per frame, using only high brightness features near the focal plane. Tracking of simulated movies of non-motile (NM) or motile cells reproduced the input diffusion coefficient D and swimming speed distribution $P(v)$(20). Tracking of experimental data of purely NM *E. coli* yields the same D as from DDM.

The analysis of (simulated or experimental) mixed populations of motile and NM cells is more challenging. We generalized a recently-proposed method (21) to analyze such data. Each trajectory is split into short elementary segments of duration Δt over which an average swimmer moves ≈ 1 pixel. First, the mean angle $\langle |\theta| \rangle$ between successive segments is calculated; $\langle |\theta| \rangle = \pi/2$ for a random walk and $\langle |\theta| \rangle = 0$ for a straight swimmer. Then, using the trajectory's start-to-end distance L, duration T, and the mean elementary segment length $\Delta r_{2D}(\Delta t)$, we calculate the parameter $N_c = \frac{L/\Delta r_{2D}}{T/\Delta t}$. Thus $N_c = 0$ for a random walk with $T \to \infty$ and $N_c = 1$ for a straight swimmer. Previous tracking of mixed swimmers and diffusing particles in 2D (at a wall) (21) returned two well-separated clusters in the $(N_c, \langle |\theta| \rangle)$ plane, from which motile and NM populations could be separated and the respective $P(v)$ and D - via fitting of the mean-squared displacement (MSD), $\langle \Delta r_{2D,NM}^2(\tau) \rangle = 4D\tau$ - could be extracted.

However, our bulk data (Section. 3.3) show a much less well defined separation in contrast to the clear distinction in (21) between motile and NM populations (specific for near wall dynamics). Therefore, we studied the dependence of motility parameters with the population selection criteria $(N_c, \langle |\theta| \rangle)$. In addition, we use another estimate for the diffusion coefficient,

D_g, obtained by fitting the distribution of 1D displacements, $P(\Delta x_{NM}(\tau))$, to a Gaussian and using the linear increase of the variance of the fitted distribution with τ to obtain D_g.

Finally we tracked the *C. reinhardtii* videos, identifying ≈ 300 algae per frame with ≈ 5 pixels per cell, and applied the above track diagnostic method (using Δt such that $\Delta r(\Delta t) \approx 1$ pixel on average) to separate 'straight' tracks in the imaging plane from other tracks. Further details are given in Section 5.3.

3   Smooth swimming *E. coli*

In (6), we demonstrated DDM using a WT strain of *E. coli*, in which cells 'run and tumble'. We suggested that while salient features of bacterial motility could be explained ignoring the effect of tumbling, some details of the data, such as a small q-dependence in the fitted swimming velocity, could only be understood by taking tumbling into account. Here, we present measurements for a *smooth swimming* mutant of *E. coli*. The simplicity of the motion compared to the WT makes this mutant the ideal organism for presenting the details of DDM. We return to the WT in Section 4.

3.1   Model of $f(q, \tau)$

In a smooth swimming (SS) mutant, each cell possesses a flagellar bundle that rotates exclusively CCW (at ~ 100 Hz); this propels the cell in a straight line, but angular deviations accumulate from orientational Brownian motion. Since the whole 'cell+flagella' complex must be torque free, the cell body rotates CW (at ~ 10 Hz). Moreover, the flagella bundle in general propels the cell off-centered, therefore the cell body 'wobbles'.

To extract motility parameters from the ISF, $f(q, \tau)$, it is important to measure this function in the appropriate q range. An upper bound for q exists because at $q \gtrsim 2\pi/R \sim 6 \ \mu\text{m}^{-1}$, where $R \sim 1 \ \mu\text{m}$ is a typical cell size, both swimming and body wobble contribute to the decay of $f(q, \tau)$, so that it is impractical to extract swimming parameters cleanly in this regime. We thus need to access lower q, or, equivalently, larger length scales. A lower bound for q is set by deviations from straight-line swimming due to Brownian orientational fluctuations and/or tumbling. For *E. coli*, cells run for $\sim 20 \ \mu\text{m}$ between tumbles; this is also the 'persistence length' of the trajectory of SS

cells due to orientational fluctuations. Thus, we do not want to probe much below $q \sim 0.5 \ \mu\mathrm{m}^{-1}$.

Within the range $0.5 \ \mu\mathrm{m}^{-1} \lesssim q \lesssim 6 \ \mu\mathrm{m}^{-1}$, it is possible to model a population of swimming *E. coli* as straight swimming particles with a speed distribution $P(v)$ and uniformly distributed directions. Each particle also undergoes Brownian motion, with diffusivity D. To model a natural population, which inevitably contain non-motile cells, we specify that only a fraction α of the particles are swimming. The resulting ISF has been derived before (11):

$$f(q,\tau) = (1 - \alpha)e^{-q^2 D\tau} + \alpha e^{-q^2 D\tau} \int_0^\infty P(v) \frac{\sin(qvt)}{qvt} dv. \tag{4}$$

To use Eq. 4 to fit experimental data, we need a parametrized form for $P(v)$. Limited previous data (11, 22) suggest that $P(v)$ is peaked. Using a Schulz (or generalised exponential) distribution

$$P(v) = \frac{v^Z}{Z!} \left(\frac{Z+1}{\bar{v}}\right)^{Z+1} \exp\left[-\frac{v}{\bar{v}}(Z+1)\right], \tag{5}$$

where Z is related to the variance σ^2 of $P(v)$ via $\sigma = \bar{v}(Z+1)^{-1/2}$, leads to the following analytical solution of the integration in Eq. 4 (24)

$$\int_0^\infty P(v) \frac{\sin(qv\tau)}{qv\tau} dv = \left(\frac{Z+1}{Zq\bar{v}\tau}\right) \frac{\sin(Z\tan^{-1}\Lambda)}{(1+\Lambda^2)^{Z/2}}, \tag{6}$$

where $\Lambda = (q\bar{v}\tau)/(Z+1)$.

Figure 1 (green curve) shows an example of an ISF calculated at $q = 1 \ \mu\mathrm{m}^{-1}$ using typical *E. coli* motility parameters in Eqs. 4-6. The ISF shows a characteristic two-stage decay. The integral in Eq. 4 due to the straight-line motion of swimmers dominates the first, faster, process, while the purely diffusive first term due to the Brownian motion of non-swimmers dominates the second, slower, process.

Much can be learnt from visual inspection of this $f(q,\tau)$. The relative amplitudes of the fast and slow processes can easily be estimated to be $\approx$ 7 : 3, which gives an estimated $\alpha \approx 0.7$. The length scale probed at this q is $\ell \sim 2\pi/q \sim 6 \ \mu\mathrm{m}$. Either by extrapolating the green curve or by reference to the red curve for pure swimmers, it can be estimated that the fast process decays completely in $\tau_{\mathrm{swim}} \sim 0.5$ s. An order of magnitude estimate of the

swimming speed is therefore $v \sim \ell/\tau_{\mathrm{swim}} \sim 12\ \mu\mathrm{m/s}$. The slower, diffusive, process decays completely in $\tau_{\mathrm{diff}} \sim 20$ s, so that an estimate of the diffusion coefficient of the non-swimmers can be obtained from $6D\tau_{\mathrm{diff}} \sim \ell^2$, giving $D \sim 0.35\ \mu\mathrm{m}^2/\mathrm{s}$. These are credible estimates of the parameters used to generate this ISF: $v = 15\ \mu\mathrm{m/s}$, $D = 0.3\ \mu\mathrm{m}^2/\mathrm{s}$, and $\alpha = 0.7$.

3.2 DDM Results

Fig. 2a shows typical DICFs, $g(q,\tau)$, measured using DDM in the range $0.45 \leq q \leq 2.22\ \mu\mathrm{m}^{-1}$ for a suspension of SS *E.coli* mutant $\Delta cheY$. The measured $g(q,\tau)$ have a characteristic shape reminiscent of the calculated $f(q,\tau)$ shown in Fig. 1 (Note the log-scale for the y-axis in Fig. 2a); indeed, Eq. 2 shows that $g(q,\tau)$ should take the shape of an (un-normalized) 'up-side-down' $f(q,\tau)$. Eq. 2 also shows that the value of $g(q,\tau)$ at small τ gives a measure of the camera noise parameter $B(q)$, which is therefore seen to be more or less q-independent. The total amplitude of $g(q,\tau)$ measures the pre-factor $A(q)$, which evidently increases rapidly as q decreases. This reflects the strong q dependence of both the form factor of a single bacterium and the contrast function of the microscope objective.

The above qualitative remarks can be quantified by fitting the measured $g(q,\tau)$ using Eqs. 2, 4 & 6. From the fit, we extract the six parameters $\bar{v}$, σ, D, α, A and B. The fitted functions $A(q)$ and $B(q)$ allow us to calculate $f(q,\tau)$ from the measured $g(q,\tau)$ via Eq. 2 (25), Fig. 2b.

The ISFs calculated from experimental data (especially those for $q \approx 1\ \mu\mathrm{m}^{-1}$) show the characteristic shape already encountered in the theoretical ISF shown in Fig. 1: a fast decay due to swimming, followed by a slow decay due to diffusion. The identity of these two processes is confirmed by the different scaling of the time axis required to collapse the data at different q values: the slow (diffusive) decay scales as $q^2\tau$, Fig. 3a, and the fast (swimming, or ballistic) decay scales as $q\tau$, Fig. 3b.

A clear separation of the swimming and diffusive decays is important for robust fitting of the ISF using Eq. 4. Such separation of time scales will be achieved if a cell takes much less time to swim the characteristic distance probed, $\ell = 2\pi/q$, than to diffuse over the same distance (in the image plane), i.e. $\tau_{\mathrm{swim}} \sim \ell/v \ll \tau_{\mathrm{diff}} \sim \ell^2/4D$, which requires $q \ll q_c \sim v/D \sim 20 - 50\ \mu\mathrm{m}^{-1}$ for typical *E. coli* values of v and D. All the data shown in Fig. 2b fit comfortably into this regime (23).

Fig. 4 shows the fit parameters $(\bar{v}, \sigma, \alpha, D, A, B)$ from Eq. 4-6 as functions

of q. A common feature, particularly evident in $D(q)$, is the enhanced noise at low q. This is because at low q, the long-time, diffusive part of $f(q, \tau)$ has not fully decayed yet in our time window, rendering it harder to determine D accurately. This can be improved by probing $g(q, \tau)$ over long times. To within experimental uncertainties, the motility parameters $(\bar{v}, \sigma, \alpha, D)$ are all q independent for $q \gtrsim 1$ μm^{-1}, which suggests that our model, Eq. 4, is indeed able to capture essential aspects of the dynamics of a dilute mixture of non-motile and smooth swimming *E. coli*. Note that a fit using fixed D, over the full q-range, results in q-independent motility parameters only when the value used for D is within 10% of the value found in free fitting (data not shown). Averaging over q, in the range $0.5 \lesssim q \lesssim 2.2$, yields $\bar{v} = 10.9 \pm 0.3$ μm/s, $\sigma = 6.43 \pm 0.04$ μm/s, $\alpha = 0.585 \pm 0.002$ and $D = 0.348 \pm 0.003$ μm^2/s with error bars being the standard deviation of the mean in all cases except for $\bar{v}$ where estimated error bars reflects the residual q dependence.

Note that the value of $D = 0.35$ μm^2/s is higher than the value (6) of $D = 0.30$ μm^2/s expected for a suspension of purely non-motile *E. coli* with "paralyzed" flagella ($\Delta motA$) with similar geometry as the wild type. This is due to an enhancement of the diffusion of non-motile cells (and other colloidal-sized objects) in a suspension of motile organisms (6, 21). Moreover, the fitting of D is dominated by the diffusion of non-motile organisms: changing our model from Eq. 4 to one in which the motile cells do not diffuse does not change the results (data not shown).

We used a Schulz distribution for modelling $P(v)$ for analytic convenience in the integration of Eq. 4. In Fig. 5a we show the average speed obtained by fitting with three different probability distributions. The results for the Schulz and log-normal distributions agree closely, but using a Gaussian form produced noisier data and a significantly lower $\bar{v}$. The latter is because $P(v = 0) \neq 0$ for the Gaussian distribution, strongly overestimating the number of slow swimmers. The presence of these spurious slow swimmers in turns renders it difficult to fit D, causing noisier data for all motility parameters.

We were able to fit the data satisfactorily irrespective of whether bright field, phase contrast or fluorescence imaging was used. However, phase-contrast imaging shows a better signal to noise ratio $(A(q)/B(q))$. In particular, changing $A(q)$ and $B(q)$ by using a 20$\times$ phase-contrast objective (which is suboptimal for our experiment) produced the same results in the relevant q range (data not shown).

3.3 Tracking results

Figure 6(a) show the probability density of the track diagnostics $(N_c, \langle|\theta|\rangle)$ (Section 2.5). Recall that $(N_c, \langle|\theta|\rangle) = (1,0)$ for straight swimming and $(0, \pi/2)$ for Brownian diffusion. While two clear maxima corresponding to diffusion and (nearly-straight) swimming are observed, there is a substantial statistical weight of tracks with intermediate $(N_c, \langle|\theta|\rangle)$ values. The actual distribution obtained depends on Δt, the elementary time interval into which we segment trajectories. Our optimal choice, $\Delta t = 0.1$s, over which the average swimming distance is ≈ 1 pixel, gave the most sharply separated peaks. However, the motile and NM populations are still not cleanly separated in our $(N_c, \langle|\theta|\rangle)$ data, Fig. 6a. We therefore select various populations of motile and NM cells by including tracks with $(N_c, \langle|\theta|\rangle)$ values within progressively larger circles centered on their respective peaks in the $(N_c, \langle|\theta|\rangle)$ space. The radius of the circle (ϵ) is measured in units such that the $(0 \leq N_c \leq 1, 0 \leq \langle|\theta|\rangle \leq \pi/2)$ space in Fig. 6a is a 10×10 rectangle.

For motile cells, $P(v)$ was determined, at each ϵ, by calculating the speed, $v = \langle \Delta r_{2D}(\tau)/\tau \rangle_T$, for each trajectory, averaged over the trajectory duration T, for various τ. The limit $\tau \to 0$ gives the instantaneous linear velocity. In practice, the lowest reasonable τ is set by $\Delta t = 0.1$s. Figure 6b shows $P(v)$ at $\epsilon = 3$ for $\tau = 0.1$ s and 0.4 s, while Fig. 6c shows $\bar{v}$ and σ of $P(v)$ for $\tau = 0.1$ s and $\tau = 1$ s. Unsurprisingly, $\bar{v}$ decreases with ϵ as progressively more 'non-ideal' swimming tracks are included: first more curved trajectories and then (at larger ϵ), some diffusive ones. Thus, there are ambiguities involved in motility characterisation using tracking. It was also not possible to extract reliably a value for α due to strong dependence on ϵ.

However, the results for the other motility parameters show reasonable agreement with DDM (Fig. 4). In particular, using $(\Delta t, \tau) = (0.1\text{s}, 0.1\text{s})$ (•, Fig. 6c), and averaging over all ε, $\bar{v} = 10.7 \pm 0.3$ μm/s and $\sigma = 5.1 \pm 0.1$ μm/s. The mean speed $\bar{v}_\varepsilon$ for each ε is also consistent with the MSD of the swimmers, Fig. 6e. The measured $P(v)$ depends on τ, e.g. some fast swimmers will not be tracked for large τ unless perfectly aligned with the image plane, while for very short τ the 2D-projection contributes to $P(v)$ at small v (20). Yet, for $\tau \sim 0.1 - 0.2$s, our measured $P(v)$ agrees with the Schulz distribution inferred from DDM, Fig. 6c.

For NM cells, we determined D by fitting the MSDs for selected tracks at several ϵ. We again found a dependence on $(\Delta t, \tau)$ and ϵ. The MSD for $\epsilon > 1$ showed deviations from purely diffusive behaviour and/or the resulting values

of D depend significantly on ε (Fig. 6d, □). Both effects are due to (likely artificial) non-Gaussian tails in $P(\Delta x_{NM}(\tau))$ (not shown). Another estimate of the diffusion coefficient is D_g (based on Gaussian fits to $P(\Delta x_{NM}(\tau))$, Section 2.5), shown as a function of ε in Fig. 6d. The average value $D_g = 0.36$ $\mu\text{m}^2/\text{s}$ agrees with the DDM value of 0.35 $\mu\text{m}^2/\text{s}$ and to the one from the MSD for $\varepsilon = 1$. Figure. 6e shows both the incorrect MSD of the diffusers obtained for $\varepsilon = 3$ and the appropriate MSD based on D_g.

3.4 Vertical motion and depth of field

Our derivation of Eq. 2 assumes that the image contrast of a bacterium does not vary with its position along the vertical (optical) z-axis, i.e. assumes an infinite depth of field (δ). The validity of this assumption depends on how fast cells move relatively to the finite δ in reality. Giavazzi et al. presented a complex theoretical model, based on the coherence theory, to take into account this effect (8). Here, we suggest a simple model and use simulations to investigate this effect and its importance over the accessible q-range. Our simple model captures the essential features and reproduces qualitatively and quantitatively the experimental results.

Experimentally (6), the intensity profile of a bacterium along the z-axis can be described by the contrast function

$$C(z) = C_B - C_0 \left(1 - \frac{4z^2}{\delta^2} \right) \tag{7}$$

where C_B and C_0 are the background and the amplitude of an object in the focal plane ($z = 0$) respectively. We determined C_B and C_0 experimentally, and then used this function to 'smear' the simulated data previously presented (6) to give simulated 'images' at a range of δ. At each δ except the lowest, the input values $\{\bar{v}, \sigma, \alpha, D\}$ are recovered from DDM analysis of these images at $q \gtrsim 2\pi/\delta$; the case of $\bar{v}$ is shown in Fig. 7. However, for $q \lesssim 2\pi/\delta$, the analysis returns $\bar{v}$ and D values that are too high: disappearance of cells along the z axis due to the rapid fading of $C(z)$ is mistaken as swimming and diffusion. Comparison between simulated data (Fig. 7) and experimental data (Fig. 4 & 5b) shows that the effect of finite depth of field, δ, is negligible for $q > 0.5$ μm^{-1} using 10x phase-contrast imaging and that our experimental depth of field $\delta \gtrsim 20$ μm.

4   Wild-Type E.coli

The motility pattern of WT *E. coli* in the absence of chemical gradients is well known (4). A cell alternates between running for $t_{\mathrm{run}} \approx$ s and tumbling for $t_{\mathrm{tum}} \approx$ s. During the latter they change direction abruptly. After many tumbling events, the bacterium effectively performs a 3D random-walk.

Modelling the ISF using Eq. 4 assumes that swimmers follow straight trajectories and neglects the effect of tumbling. We have previously applied our method to the WT *E. coli* (6). Here we study the effect of tumbling by comparing systematically the q-dependence of the average speed obtained from DDM for WT (run and tumble) and SS (run only) swimmers. Since several experimental data sets were obtained from different batches of cells, we report the speed normalised to $\langle \bar{v} \rangle_{\mathrm{high\ q}}$, the average in the range ($2.0 < q < 2.2\ \mu\mathrm{m}^{-1}$).

Simulations and experiments show, Fig. 5, a qualitative difference in $\bar{v}(q)$ of SS and WT cells. All data for WT cells show a slight *decrease* in $\bar{v}(q)$ towards low q, while data for SS show the opposite trend. The increase towards low q in the $\bar{v}(q)$ of the SS is presumably largely due to depth of field effects (Section 3.4). The opposite trend in the behaviour of $\bar{v}(q)$ for the WT can be understood as follows. The mean speed, $\bar{v}(q)$, measured by DDM at a certain q is estimated by $\bar{v}(q) \sim (2\pi/q)/\tau_q$, i.e. the time ($\tau_q$) taken to advect density between two points spatially separated by distance $2\pi/q$. For a straight swimmer, the tracklength s will be equal to the distance between the two points, i.e. $s = 2\pi/q$, so that $\bar{v}(q) = \bar{v}$. Any deviation from a straight track, e.g. due to changes in direction from tumbling, renders $s > 2\pi/q$. Since $\tau_q = s/\bar{v}$, we now have $\bar{v}(q) < \bar{v}$. This effect becomes progressively more pronounced at low q, as observed.

5   Swimming algae: wild-type C. *reinhardtii*

As a final example we apply DDM to the bi-flagellate freshwater alga *Chlamydomonas reinhardtii*, a model for eukaryotic flagellar motility (15). *C. reinhardtii* has a prolate spheroidal cell body about 10 μm across with two flagella roughly 10-12 μm long (15). Beating the latter at $\simeq$ 50Hz in an alternation of effective (forward moving) and recovery (backward moving) strokes propels the cell body forward, oscillating as it advances. The flagellar beat is not perfectly planar, so cells precess around their long axis at $\simeq$ 2Hz; this

rotation, critical to phototaxis (17), results in helical swimming trajectories. For length scales $\sim 100\ \mu m$ the direction of the axis of the helical tracks is approximately straight, but on larger scales the stochastic nature of the flagellar beat causes directional changes resulting in random walk (5, 26, 27).

Racey *et al.* (28) carried out the first high-speed microscopic tracking study of *C. reinhardtii*, obtaining the cells' swimming speed distribution, with mean speed $\bar{v} = 84\ \mu m/s$, as well as the average amplitude $\bar{A} = 1.53\ \mu m$ and frequency $\bar{f} = 49$ Hz of the beat. A more recent tracking study obtained a 2D swimming speed distribution with $\bar{v} \approx 100\ \mu m/s$ (29). The swimming of *C. reinhardtii* has also been studied with dynamic light scattering (DLS) (28, 30). However, studies at appropriately small q are difficult, severely limiting useful data (the limitation is even greater than for bacteria: algae swim on larger length scales, requiring smaller values of q). We present here the first characterisation of the swimming motility of *C. reinhardtii* using DDM. As for *E. coli*, the technique allows to study the 3D motion of much larger numbers of cells ($\sim 10^4$) than tracking, and allows easy access to larger length scale than DLS.

5.1 Model of $f(q, \tau)$

The swimming dynamics of *C. reinhardtii* are on larger length scales, shorter time scale (algae swim faster) and of a different nature than *E. coli*, so models and experimental conditions used for DDM need to be adjusted for this organism. In particular, the decay of the ISF, $f(q, \tau)$, will reflect these algae's peculiar dynamics. Cells oscillate at lengthcales $< 10\ \mu m$, translate in the range $10\ \mu m < L < 30\ \mu m$, spiral over $30\ \mu m < L < 100\ \mu m$, and diffuse for $L > 100\ \mu m$. A schematic representation of a helical trajectory, highlighting the small scale oscillatory motion is shown in inset of Fig. 8 (diffusive length scales are not shown).

At length scales $L \lesssim 30\ \mu m$, the swimming of *C. reinhardtii* can be approximated as a sinusoidal oscillation superimposed on a linear progression, so that the particle displacement $\Delta r(\tau)$ of a cell after a time interval τ is given by (30)

$$\Delta r(\tau) = v\tau + A_0[\sin(2\pi f_0 \tau + \phi) - \sin(\phi)], \tag{8}$$

where A_0 and f_0 are the amplitude and frequency of the swimming oscillation and ϕ is a random phase to ensure the swimming beats of different cells

are not synchronised. Substituting this into Eq. 3, averaging over ϕ, and assuming a Schulz distribution for the swimming speed (Eq. 5), we obtain the ISF,

$$f_{\text{algae}}(q,\tau) = \frac{1}{2} \int_{-1}^{1} \frac{\cos\left[(Z+1)\tan^{-1}(\Lambda\chi)\right]}{\left[1+(\Lambda\chi)^2\right]^{(Z+1)/2}} J_0[2qA_0\chi\sin(\pi f_0\tau)]\,d\chi, \quad (9)$$

where $\Lambda \equiv \frac{q\bar{v}\tau}{Z+1}$, $\chi \equiv \cos\psi$, where ψ is the angle between $\vec{q}$ and $\vec{r}$, J_0 is the zeroth order Bessel function. All other terms are as previously defined. The first and second term describe the contribution from straight swimming and oscillatory beat, respectively. In the limit $qA_0 \ll 1$ ($J_0 \to 1$), Eq. 9 analytically integrates to Eq. 6, the same expression as for the progressive model used for *E. coli*.

The derivation of Eq. 9 assumes the distributions $P(A)$ and $P(f)$ for swimming amplitude and frequency, respectively, are narrowly centred around the values A_0 and f_0 and ignores (i) the negligible diffusion of non-motile algae; (ii) any bias in the swimming direction caused by gravitaxis (31); and (iii) the helical nature of the swimming.

5.2 DDM results

Fig. 8a shows a typical DICF, $g(q,\tau)$, at $q = 0.52$ μm^{-1} ($l \approx 12$ μm), for a suspension of WT alga *C. reinhardtii* measured using DDM. The reconstructed ISFs, $f(q,\tau)$, are shown in Fig. 8b in the range $0.2 \lesssim q \lesssim 0.9$ μm^{-1}, corresponding to a length scale range of $7 \lesssim l \lesssim 30$ μm^{-1}. $f(q,\tau)$ shows a characteristic shape for all q's: a fast decay at $\tau \leq 0.02$ s due to the oscillatory beat and a slower decay at $\tau \geq 0.02$ s due to swimming. The identity of these two processes is confirmed by their difference in $\tau-$ and $q-$ dependencies. The characteristic time of the fast process is independent of q, while its amplitude decreases with q. Both observations fully agree with the term (J_0) due to the oscillatory contribution in Eq. 9. Moreover, 0.02 s corresponds to the period of a 50 Hz oscillatory beat. Finally, the slow process scales perfectly with $q\tau$ (data not shown) confirming the ballistic nature (swimming) of this process.

Fig. 9 shows the fitting parameters ($\bar{v}, \sigma, A_0, f_0$) from Eq. 2 using the oscillatory model (Eq. 9) as a function of q. All parameters display a small q-dependence. This is likely due to effects not captured by the simple oscillatory model (e.g. body precession and helical swimming) and will be discussed

elsewhere. Averaging over q yields $\bar{v} = 89.6 \pm 2.8$ μm/s, $\sigma = 24.9 \pm 4.6$ μm/s, $A_0 = 0.98 \pm 0.06$ μm and $f_0 = 48.6 \pm 0.55$ Hz, with estimated error bars reflecting the residual q dependence. Fitting the experimental data using Eq. 9 requires numerical integration. Using instead the linear model (Eq. 6), thus ignoring the oscillatory beat, yields similar results for $(\bar{v}, \sigma)$ as shown in Fig. 9. This is because the fast process is mainly ignored when performing such fit as shown in Fig. 8a. Moreover, using the linear model and a movie, for which the lowest $\tau \gtrsim 1/f_0$ (for example 100 fps) and therefore the oscillatory beat is not contributing to the ISF, yields the same $(\bar{v}, \sigma)$, thus allowing high-throughput economical measurements of the mean speed of biflagellate algae.

5.3 Tracking results

Tracking of *C. reinhardtii* resulted in two well separated groups of $(N_c, \langle |\theta| \rangle)$ values (Section 2.5) independent of Δt, provided $\Delta t > 1/f_0$, Fig. 10a. We used tracks with $(N_c \geq 0.7, \langle |\theta| \rangle < 0.5)$, reflecting nearly straight swimmers aligned with the image plane, to obtain $P(v)$. Misaligned tracks are excluded this way: motion perpendicular to the helical axis enhances the circular contribution in the 2D projection, thus reducing N_c and increasing $\langle |\theta| \rangle$ (insets to Fig. 10a).

We measured $P(v)$ for several τ, Fig 10b, and found a slight τ- dependency, e.g. due to undetected fast swimmers for large τ. Note that our $P(v)$ is smaller at small v than in (29), where all *projected* trajectories were considered. Our small v data are likely closer to the true distribution, due to our exclusion of misaligned tracks. Moreover, our $P(v)$ are in reasonable agreement with the result inferred from DDM (Fig 10b). We find $\bar{v} = 81 \pm 1$ μm/s and $\sigma = 22 \pm 3$ μm/s (averaged over different τ), in excellent agreement with the values from DDM (Fig. 9). Extending the selected trajectories to $(N_c \geq 0.55, \langle |\theta| \rangle < 0.7)$, changes $P(v)$ and $\bar{v}$ by less than 5%. We analysed the oscillating component of the displacement, $r_{os}(t)$, for 'straight' tracks (32). From Fourier analysis of $r_{os}(t)$, we obtained $f_0 = 49.3 \pm 0.5$ Hz. We identified an additional modulation frequency of ~ 10 Hz (i.e. an extra peak at $f - f_0 \simeq 10$ Hz in the power spectrum of $r_{os}(t)$, to be discussed elsewhere). From the rms value of $r_{os}(t)$, we determined the average oscillation amplitude $A_0 = \sqrt{2\langle x_{os}^2 + y_{os}^2 \rangle} = 0.93 \pm 0.22$ μm. These values are in agreement with previous (28) and DDM results.

Thus, our results simultaneously validate the DDM methodology and the

simple model (Eq. 8 & 9) for the swimming of *C. reinhardtii*. Our method can therefore be used for the fast and accurate characterisation of the motility of large ensembles of this organism, and, potentially, of other algae.

Conclusions

We have shown that DDM is a powerful, high-throughput technique to characterise the 3D swimming dynamics of microorganisms over a range of time scales and length scales (~ 3 and ~ 1 order of magnitude respectively) simultaneously in a few minutes, based on standard imaging microscopy. The time scales and length scales of interest depend on the swimming dynamics of the microorganism, and are easily tuned by changing the frame rate or optical magnification respectively.

We have studied in considerable detail how to use DDM for characterising the motility of smooth-swimming (run only) and wild-type (run and tumble) *E. coli*, as well as the wild-type alga *C. reinhardtii*. We validated the methodology using tracking and simulations. The latter was also used to investigate the effect of a finite depth of field and tumbling. Using DDM, we were able to extract (i) for *E. coli*: the swimming speed distribution, the fraction of motile cells and the diffusivity; and (ii) for *C. reinhardtii*: the swimming speed distribution, the amplitude and frequency of the oscillatory dynamics. In both cases, these parameters were obtained by averaging over many thousands of cells in a few minutes without the need for specialised equipment.

Further developments are possible. For *E. coli*, analytic expressions for $\bar{v}(q)$ taking into account either trajectory curvature due to rotational Brownian motion (smooth swimmers) or directional changes due to tumbling (wild type), can be derived. Fitting these expressions to data should then yield quantitative information on the respective features. For *C. reinhardtii*, the helical motion, the asymmetric nature of the swimming stroke, and the higher harmonics in the body oscillations observed by tracking could be explored theoretically and using DDM experiments. This will allow us to test simulations that use the method of regularised stokeslets to reproduce the fine details of the swimming of bi-flagellate algae (33).

DDM is based on the measurement of the spatio-temporal fluctuations in intensity, and therefore does not require good optical resolution of the motile objects. Thus, DDM can probe a large field of view, yielding good

statistics even under relatively poor imaging conditions. Moreover, DDM could also be used to probe anisotropic or asymmetric dynamics (34) of microorganisms. Finally, the analysis is not restricted to dilute suspensions and can be used to investigate the swimming dynamics at different time scales and length scales of the collective behaviour of populations. However, quantitative interpretation of the resulting data will require new models of the ISF.

With the availability of DDM, quantitative characterisation of motility can become a routine laboratory method, provided suitable theoretical models are available for fitting of the ISF. Even without such models, however, qualitative features of the measured ISF can still allow conclusions to be drawn and trends to be studied (e.g. the speeding up of the decay of the ISF almost invariably correspond to faster motion). DDM should therefore be a powerful tool in future biophysical studies of microorganismic locomotion.

Acknowledgment

References

1. Yoshiyama, H., and T. Nakazawa. 2000. Unique mechanism of *Helicobacter pylori* for colonizing the gastric mucus. *Microbes and Infection*, 2, 55-60.

2. Berg, H.C. 2004. *E. coli in motion.* Springer.

3. Berg, H.C. 1971. How to track bacteria. *Rev. Sci. Instr.*, 42, 868-871.

4. Berg, H.C. and D.A. Brown. 1972. Chemotaxis in *Escherichia coli* analysed by Three-dimensional Tracking. *Nature*, 239, 500-504.

5. Drescher, K., K.C. Leptos, and R.E. Goldstein. 2009. How to track protists in three dimensions. *Rev. Sci. Instrum.* 80:014301.

6. Wilson, L.G., V.A. Martinez, J. Schwarz-Linek, J. Tailleur, G. Bryant, P.N. Pusey and W.C.K. Poon. 2011. Differential Dynamic Microscopy of Bacterial Motility. *Phys. Rev. Lett.*, 106:018101.

7. Cerbino, R., and V. Trappe. 2008. Differential Dynamic Microscopy: Probing Wave Vector Dependent Dynamics with a Microscope. *Phys. Rev. Lett.*, 100:188102.

8. Giavazzi, F., D. Brogioli, V. Trappe, T. Bellini and R. Cerbino. 2009. Scattering information obtained by optical microscopy: Differential dynamic microscopy and beyond *Phys. Rev. E*, 80:031403.

9. Boon, J.P., R. Nossal and S.H. Chen. 1974. Light-scattering spectrum due to wiggling motions of bacteria. *Biophys. J.*, 14:847–864.

10. Berne, B.J. and R. Pecora. 2000. *Dynamic Light Scattering*, Dover.

11. Stock, G.B. 1978. The measurement of bacterial translation by photon correlation spectroscopy. *Biophys. J.*, 22:79–96.

12. P1 phage transduction (13) was used to create a smooth swimming (AB1157 $\Delta cheY$) using the appropriate E. coli K-12 single knockout mutant from the KEIO collection (14). Kanamycin (final concentration 30 mg/l) was added to all growth media for AB1157 $\Delta cheY$.

13. Miller, J.H. 1972. *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory Press.

14. Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K.A. Datsenko, M. Tomita, B.L. Wanner and H. Mori. 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, 2, 111.

15. Harris, E.H. 2009. The *Chlamydomonas* sourcebook. Academic Press.

16. Croze, O.A., E.E. Ashraf, and M.A. Bees. 2010. Sheared bioconvection in a horizontal tube. *Phys. Biol.* 7:046001.

17. Foster, K.W. and R.D. Smyth. 1980. Light antennas in phototactic algae. *Microbiol. Rev.* 44:572–630.

18. All relevant softwares are available on request.

19. Crocker, J.C. and D.G Grier. 1996. Methods of digital video microscopy for colloidal studies *J. Coll. Int. Sci.*, 179:298–310.

20. 2D tracking measures $P(v\cos(\beta))$ (β is the angle with the image plane) rather than $P(v)$, but this has only a small effect on the results. By tracking only the brightest features within z of the focal plane, only tracks with $\tau v\sin(\beta) < z$ contribute, which supresses projection effects. Moreover, the diagnostic method to select 'straight' swimmers, presented later, further excludes large β tracks from the motile population, since these tracks exhibit a stronger diffusive (*E. coli*) or circular component (algae), a result of the projection.

21. Miño, G., T.E. Mallouk, T.Darnige, M. Hoyos, J. Dauchet, J. Dunstan, R. Soto, Y. Wang , A. Rousselet and E. Clement. 2011. Enhanced diffusion due to active swimmers at a solid surface *Phys. Rev. Lett.*, 106:048102.

22. Nossal, R., S.H. Chen and C.C. Lai. 1971. Use of laser scattering for quantitative determinations of bacterial motility. *Opt. Commun.*, 4:35–39.

23. In the regime of $q \gg q_c$, the ISF separates into a fast diffusive process followed by a slower swimming process.

24. Pusey, P.N. and W. van Megen. 1984. Detection of small polydispersities by photon correlation spectroscopy J. Chem. Phys., 80:3513–3520.

25. The determination of $A(q)$ and $B(q)$ does not necessarily requires fitting of $g(q,\tau)$ and therefore a model for $f(q,\tau)$. They could simply be determined from the short and long time limit of $g(q,\tau)$ such as: $B(q) = f(q,\tau \to 0)$ and $A(q) = f(q,\tau \to \infty) - B(q)$. However, the latter requires well-defined "plateau" at both short and long time regimes, which are not always observed depending on the value of q, and the experimental time window restricted: at short-time by the frame rate and at long-time by the duration of the movie.

26. Hill, N.A. and T.J. Pedley. 2005. Bioconvection. *Fluid. Dyn. Res.*, 37:1–20.

27. Polin, M., I. Tuval, K. Drescher, J.P. Gollub and R.E. Goldstein. 2009. Chlamydomonas swims with two "gears" in a eukaryotic version of run-and-tumble locomotion. *Science*, 325:487–490.

28. Racey, T.J., R. Hallett and B. Nickel. 1981. A quasi-elastic light scattering and cinematographical investigation of motile *Chlamydomonas reinhardtii. Biophys. J.*, 35:557–571.

29. K.C. Leptos, J.S. Guasto, J.P. Gollub,A.I. Pesci and R.E. Goldstein. 2009. Dynamics of Enhanced Tracer Diffusion in Suspensions of Swimming Eukaryotic Microorganisms *Phys. Rev. Lett.*, 103:198103.

30. T.J. Racey and F.R. Hallett, 1983. A low angle quasi-elastic light scattering investigation of *Chlamydomonas reinhardtii J. Musc. Res. Cell. Motil.*, 4:321–331.

31. N.A. Hill and D.P. Häder. 1997. A Biased Random Walk Model for the Trajectories of Swimming Micro-organisms *J. Theor. Biol.*, 186:503–526.

32. We define $r_{os}(t) = (x_{os}, y_{os})(t) = r_{2D}(t) - \langle r_{2D}(t) \rangle_{dt}$ with $dt = 2/f_0$. $r_{os}(t)$ is well resolved due to subpixel accuracy ~ 0.2 μm of the coordinates.

33. O'Malley, S. and M.A. Bees, 2011. The Orientation of Swimming Biagellates in Shear Flows *Bull. Math. Bio.*, 74:232–255.

34. Reufer, M., V.A. Martinez, P. Schurtenberger and W.C.K. Poon. 2012. Differential dynamic microscopy for anisotropic colloidal dynamics. *Langmuir*. In press.
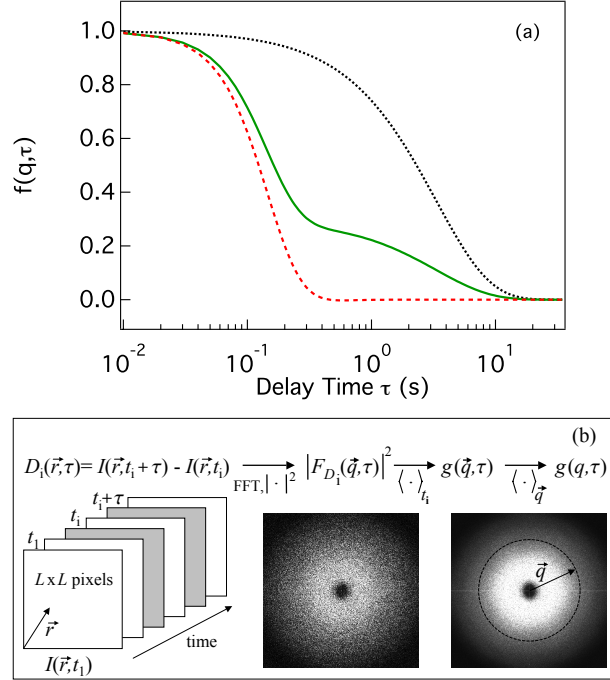
Figure 1: (a) Theoretical ISF, $f(q, \tau)$, vs τ at $q = 1$ μm^{-1}, for (black dotted line) a population of diffusing spheres with $D = 0.3$ $\mu m^2/s$, (red dashed line) a population of equivalent-size spheres swimming isotropically in 3D with a Schulz speed distribution $P(v)$ with average speed $\bar{v} = 15$ $\mu m/s$ and width $\sigma = 7.5$ $\mu m/s$, and (green line) a 30:70 mixture of diffusers and swimmers. (b) Schematic of the image processing to obtain the DICFs, $g(q, \tau)$, from the videos (left) collected in an experiment. (Middle) Non-averaged image, $|F_{Di}(\vec{q}, \tau)|^2$ and (right) averaged image, $g(\vec{q}, \tau)$, over initial times t_i at $\tau = 0.52$ s.
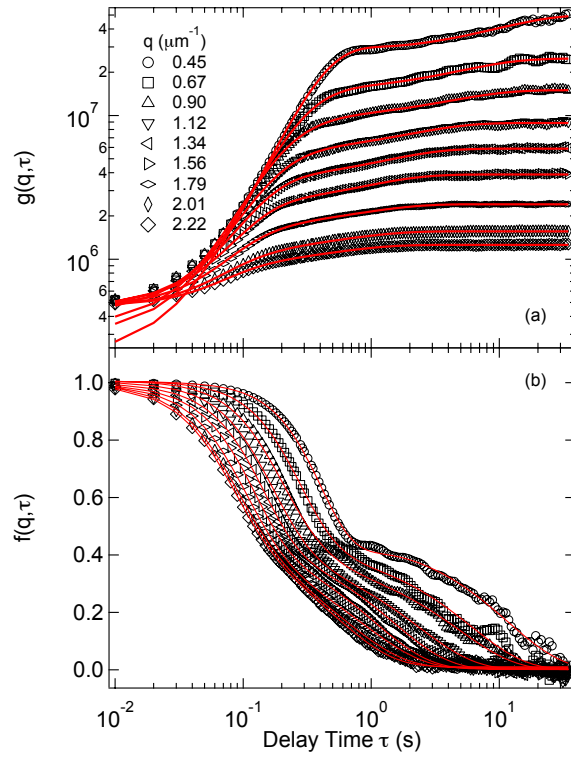
Figure 2: DDM for smooth swimming *E. coli*. (a) Measured (symbols) and fitted (lines) DICFs, $g(q,\tau)$. (b) ISFs, $f(q,\tau)$, reconstructed from $g(q,\tau)$ using Eqs. 2,4 & 6.
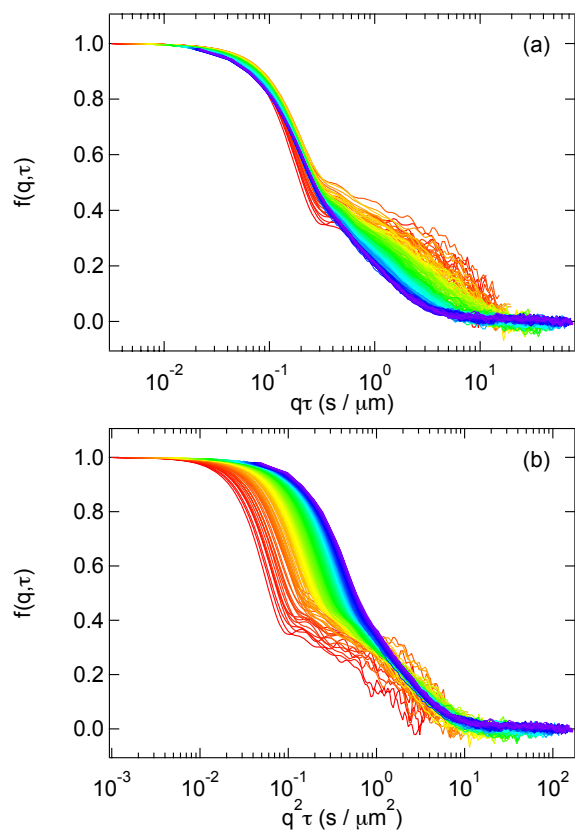
Figure 3: The reconstructed ISFs, $f(q, \tau)$, shown in Fig. 2 plotted against (a) $q\tau$ and (b) $q^2\tau$. Data collapse for the fast process in (a) and for the slow process in (b). q value increases from red to blue end of the spectrum colour in the range $0.3 \leq q \leq 2.2$ μm^{-1}.
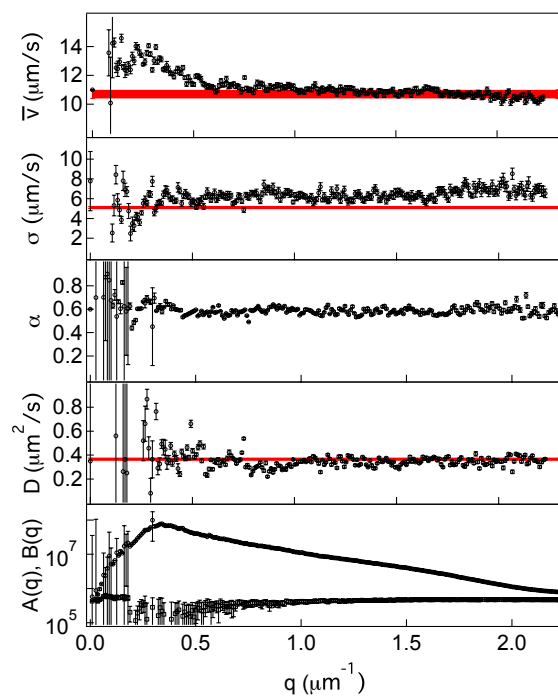
Figure 4: SS *E. coli*. Fitting parameters vs q using Eqs. 2, 4 & 6. From top to bottom: $\bar{v}$ and σ of the Schulz distribution, motile fraction α, diffusivity D, and $A(q)$ ($\circ$) and $B(q)$ ($\square$). Red lines are results from tracking, with thickness corresponding to the error bars. No reliable value for α could be obtained from tracking.
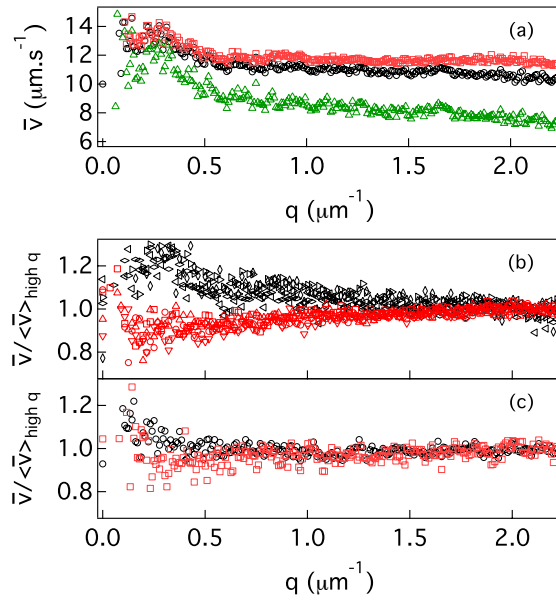
Figure 5: Swimming speed vs. q from DDM. (a) Effect of using different forms of the speed distribution: lognormal (□), Schulz (○), and Gaussian (△). (b) Effect of tumbling (experiments): four data sets from the SS (black symbols) and four data sets from the WT (red symbols). (c) Effect of tumbling (simulations): SS (○) and WT (□). Note that for (b) and (c) panels, the swimming speed has been normalised to $\langle v \rangle_{\text{high q}}$ to enable comparison and highlight the difference in q-dependence.
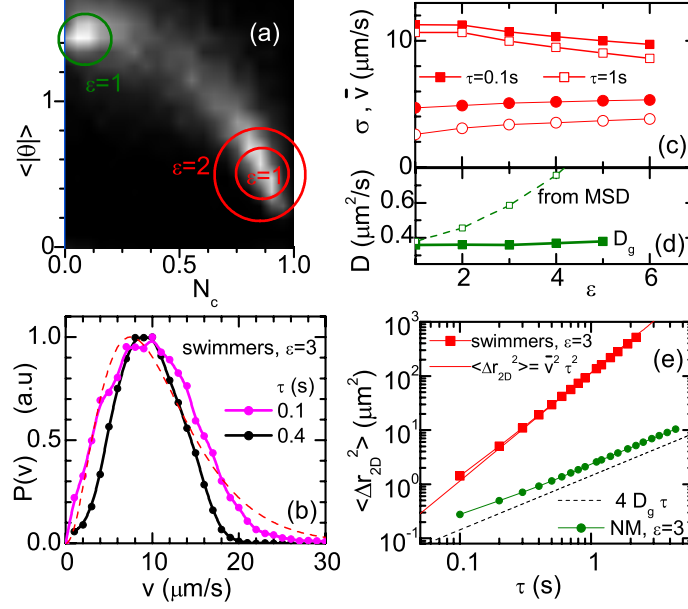
Figure 6: Tracking of SS *E. coli*. (a) Probability $P(N_c, \langle|\theta|\rangle)$ ($\Delta t = 0.1$ s) for all tracks. White denotes large values of P. Circles (radius ε, see text) are selection criteria for motile (red, lower right) or NM (green, upper left) cells. (b) $P(v)$ for $\varepsilon = 3$ ($\Delta t = 0.1$ s) for two time lags τ. Dashed line: Schulz distribution from DDM. (c) $\bar{v}$ (squares) and σ (circles) of $P(v)$ vs ε for $\tau = 0.1$ s (filled) and $\tau = 1$ s (open). (d) Diffusion coefficient of NM cells vs. ε, from Gaussian fits to $P(\Delta x_{NM}(\tau))$ (■) and from linear fits to the MSD (□). (e) MSD vs. τ for motile (■) and NM cells (•) for $\varepsilon = 3$, $\Delta t = 0.1$ s. Line: motile MSD calculated using $\bar{v}_{\varepsilon=3}$ in (c); dashed line: NM MSD calculated using $D_{g,\varepsilon=3}$ from (d).
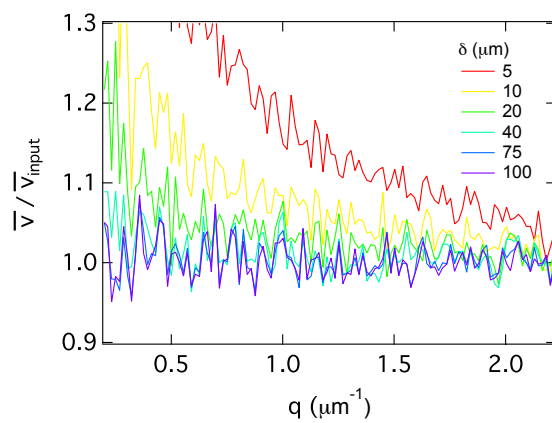
Figure 7: Effect of the depth of field on DDM analysis of simulated data. Normalised mean swimming speed $\bar{v}/\bar{v}_{input}$ versus q for smooth swimmers and several values of depth of field δ. $\bar{v}_{input} = 15\mu m/s$ is the input mean swimming speed used to generate the simulated data.
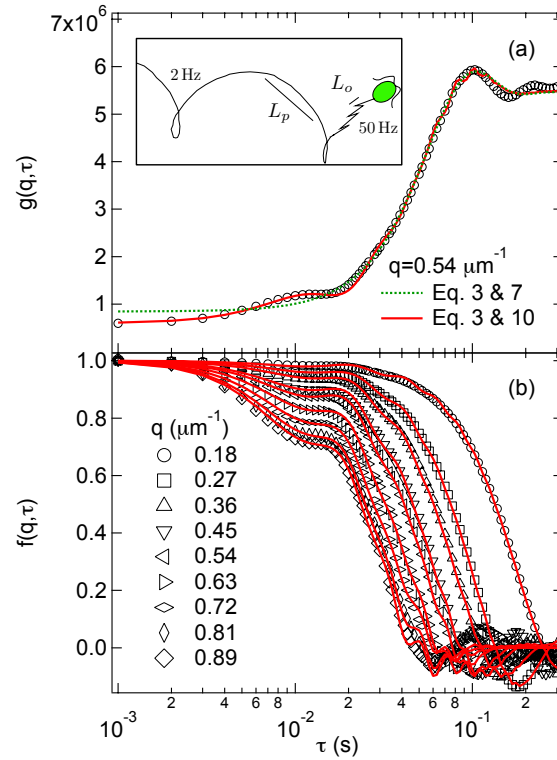
Figure 8: DDM for WT C. *reinhardtii.* (a) Measured (circles) $g(q,\tau)$. Line and dashed line are fits using the oscillatory model (Eq. 9) and the linear model (Eq. 6), respectively. Inset: Portion of a helical *C. reinhardtii* trajectory. The progressive, L_p, and (zoomed-in) oscillatory, L_0, length scales probed by DDM are shown, with the frequencies of the helical precession (2 Hz) and oscillatory swimming (50 Hz). (b) ISFs, $f(q,\tau)$, using Eqs. 2 & 9.
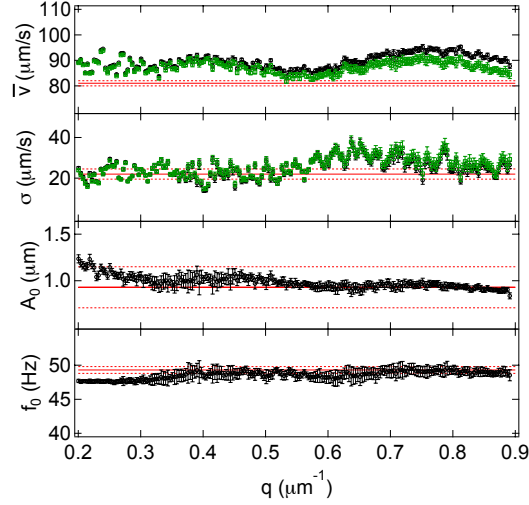
Figure 9: Fitting parameters using the oscillatory model Eq. 9 (circles) or linear model Eq. 6 (squares) as function of q for *C. reinhardtii*. From top to bottom: $\bar{v}$ and σ of the Schulz distribution, amplitude A_0, and frequency f_0. Red lines are results from tracking with dashed lines corresponding to error bars.
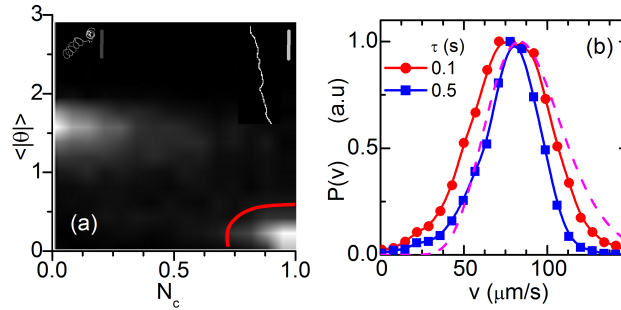

Figure 10: Tracking of *C. reinhardtii*. (a) Probability $P(N_c, \langle|\theta|\rangle)$ of all tracks ($\Delta t = 0.05$ s). Tracks within the red bordered region (example in top right inset, 7 s, scalebar 150 μm) are used to measure $P(v)$; Top left inset: an excluded track with $N_c < 0.4$ (30 s, scalebar 30 μm). (b) Normalised $P(v)$ from tracks selected in (a), for two values of τ. Dashed line: $P(v)$ from DDM analysis.