

Directional change during active diffusion of viral ribonucleoprotein particles through cytoplasm

Kathleen C. Smith,¹ Ryan Oglietti,² Steven J. Moran,³ Jed C. Macosko,⁴ Douglas S. Lyles,^{3,*} and George Holzwarth⁴

¹Department of Chemistry, Wake Forest University, Winston-Salem, North Carolina; ²Department of Biology, Wake Forest University, Winston-Salem, North Carolina; ³Department of Biochemistry, Wake Forest University School of Medicine, Winston-Salem, North Carolina; and ⁴Department of Physics, Wake Forest University, Winston-Salem, North Carolina

ABSTRACT A mesh of cytoskeletal fibers, consisting of microtubules, intermediate filaments, and fibrous actin, prevents the Brownian diffusion of particles with a diameter larger than 0.10 μm , such as vesicular stomatitis virus ribonucleoprotein (RNP) particles, in mammalian cells. Nevertheless, RNP particles do move in random directions but at a lower rate than Brownian diffusion, which is thermally driven. This nonthermal biological transport process is called “active diffusion” because it is driven by ATP. The ATP powers motor proteins such as myosin II. The motor proteins bend and cross-link actin fibers, causing the mesh to jiggle. Until recently, little was known about how RNP particles get through the mesh. It has been customary to analyze the tracks of particles like RNPs by computing the slope of the ensemble-averaged mean-squared displacement of the particles as a signature of mechanism. Although widely used, this approach “loses information” about the timing of the switches between physical mechanisms. It has been recently shown that machine learning composed of variational Bayesian analysis, Gaussian mixture models, and hidden Markov models can use “all the information” in a single track to reveal that the positions of RNP particles are spatially clustered. Machine learning assigns a number, called a state, to each cluster. RNP particles remain in one state for 0.2–1.0 s before switching (hopping) to a different state. This earlier work is here extended to analyze the movements of a particle within a state and to determine particle directionality within and between states.

SIGNIFICANCE The speed with which particles move through the cytoplasm of mammalian cells depends on their size relative to the size of pores in the mesh formed by the cytoskeleton. Particles such as vesicular stomatitis virus ribonucleoprotein, with diameter approximately 0.17 μm , are impeded but not categorically blocked by the mesh. One of the physical mechanisms that allows such particles to move through the cytoplasm is termed “active diffusion” because it requires ATP and causes motion in random directions. Machine learning code designed to detect spatial clustering of particle locations reveals a cage-hop transport mechanism that plays a leading role in the active diffusion of these particles.

INTRODUCTION

The movement of particles through the cytoplasm of live cells depends on their size relative to the size of pores in

the mesh of cytoskeletal fibers. Measurements show that the exclusion pore diameter of this mesh in a typical cell is approximately 0.10 μm (1–4). Small molecules with diameter less than 0.10 μm can move between the fibers by thermally driven diffusion (5). Ribonucleoprotein (RNP) particles of vesicular stomatitis virus (VSV) have a hydrodynamic diameter = 0.17 μm (6). They move through the mesh by two mechanisms: 1) “directed motion” as the cargo of processive motor proteins, in which particles take bursts of steps in the same direction along stiff cytoskeletal fibers, and 2) “active diffusion.” The term arises because this motion depends upon molecular motors that consume ATP (7,8).

Our studies use a recombinant strain of VSV containing enhanced green fluorescent protein (eGFP) inserted into

Submitted September 26, 2023, and accepted for publication April 23, 2024.

*Correspondence: dlyles@wakehealth.edu

Kathleen C. Smith’s present address is National Institute on Aging, Baltimore, Maryland

Ryan Oglietti’s present address is Wake Forest University School of Medicine, Winston-Salem, North Carolina

Steven J. Moran’s present address is Department of Microbiology and Immunology, Cornell University College of Veterinary Medicine, Ithaca, New York

Editor: Mark Alber.

<https://doi.org/10.1016/j.bpj.2024.04.025>

© 2024 Biophysical Society.



the hinge region of the viral P-protein (9). Each RNP (molecular mass 87.1 MDa (10)) contains approximately 400 copies of the P-protein (10). Our previous experiments used a machine learning approach based on variational Bayesian analysis, Gaussian mixture models, and hidden Markov models to demonstrate that most of the RNP transport in the cytoplasm is due to active diffusion (11,12). The positions of the particle are not randomly located in space, but instead, they form clusters. Clusters are objectively detected by the machine learning code. Each cluster is assigned a number, its state S . Machine learning evaluates the most probable number of states in that track and the most probable value of S at each time point in that track. All the particles in a cluster have the same value of S .

There is no net diffusion if S does not change. Active diffusion occurs when S changes with time, because a change in S means a change in cluster center. We call this hopping. For tracks consisting of 400 frames acquired at 100 frames/s (fps), some RNP particles remain in the same state for the entire 4 s. Other particles occupy two, three, four, five, or six clusters during 4 s. Exploratory tests for seven or eight states showed that they occurred rarely and were difficult to interpret because bouts were short. Confined motion within a trap consists of rapid bouncing from one side of the trap to the other. The most probable diameter of a trap, estimated from the lengths of the axes of the two-dimensional Gaussians that defined each state, was 0.24 μm . The most probable distance between the centers of adjacent traps was 0.17 μm . Particles stayed in traps for approximately 0.2–1 s. The dependence on time delay of mean-squared displacement (MSD) of single particles while confined to a trap was shown to be essentially 0.0.

In addition to many observations in live cells, active diffusion has been demonstrated by others in cell-free mixtures composed of actin fibers, microtubules, molecular motors, ATP, and fluorescent beads (13–15). Movement and bending of the cytoskeletal fibers and diffusion of the reporter beads are similar to those observed for endogenous particles of similar size in live cells.

The Bayesian machine learning approach we have taken to analyze RNP tracks is similar to other quantitative methods for analyzing particle tracking data. For example, Kalafut and Visscher (16) describe a method for the detection of nonuniform steps in particle tracks, which like ours uses a stepwise test of increasing numbers of change points and a statistical test of the goodness of fit. Instead of using Bayes's theorem and the Kulbeck-Liebler divergence (Eq. 9.72 in Bishop (17)), they use the Schwartz information criterion. Their method is one dimensional, suitable for directed motion, and ours is two dimensional (2D), appropriate for active diffusion. Recently published nonparametric methods of particle track analysis (18,19) are similar to ours in using Bayesian inference to identify the most likely number of states and hidden Markov models to account for time

dependence. However, the distinction between states in these methods depends on differences in diffusivities reflecting different mechanisms of transport, whereas ours depends on differences in the particle's spatial location. Deep learning neural network methods have been developed to distinguish between diffusion types such as Brownian motion, fractional Brownian motion, continuous time random walk, or Lévy walk in single-particle tracks (18,20). Our studies of RNP particles in live cells are different. They focus on the size and location of clusters, almost all of which are created by one physical mechanism, trapping. We point out that the spatial features of interest to us, such as the size and location of cluster ellipses, are defined by 10–20 pixels. This reduces the quantitative consequences of subpixel localization errors.

It has become customary in the field of single-particle tracking to use the slope of a plot of the MSD against time interval τ as a signal for identifying different classes of particle motion. For an ensemble of N_T tracks, each with N frames acquired at discrete time intervals separated by time Δ , the 2D ensemble-averaged MSD is a double sum given in Eq. 1.

$$\langle \text{msd}(n\Delta) \rangle = \frac{\Delta}{N_T(N-n)} \sum_j^{N_T} \sum_{i=1}^{N-n} |\mathbf{X}_j((i+n)\Delta) - \mathbf{X}_j(i\Delta)|^2 \quad (\text{Equation 1})$$

Here $\langle \text{msd}(n\Delta) \rangle$ is the n th MSD value at time interval $n\Delta$, the angle brackets $\langle \rangle$ denote ensemble average, and $\mathbf{X}_j(i\Delta)$ is the position vector at frame i within track j (21). In the log-log plot of this ensemble-averaged MSD versus τ , slope = 0 is interpreted as confined motion, slope = 1 as random motion (e.g., diffusion), and slope = 2 as directed motion. Intermediate slopes are often described as “anomalous diffusion,” where “anomalous” means “not Brownian”. Ensemble averaging “loses information” about the timing of the switches between physical mechanisms (21). As shown in the current paper, this loss of information is an unnecessary price to pay because machine learning applied to single-particle tracks can use “all the information” to assign a state to each data point in the track if data are acquired at 100 or 600 frames/s. Our model is that clusters exist because the cytoplasm contains a mesh of actin fibers, microtubules, and intermediate filaments that trap the RNPs. The fibers that make up the trap are dynamic. This model is supported by experiments with three drugs known to affect actin fibers, which also affect the size and dynamics of RNP clusters (11).

As noted above, machine learning evaluates the most probable state of the particle at each time point. This lets us snip the track into one to six shorter tracks (bouts) of 40–200 frames during which the particle is in a single state

S. This allows the evaluation of the MSD of each bout using Eq. 2:

$$msd_S(n\Delta) = \frac{\Delta}{(N_S - n)} \sum_{i=1}^{N_S-n} |\mathbf{X}_j((i+n)\Delta) - \mathbf{X}_j(i\Delta)|^2 \quad (\text{Equation 2})$$

These MSDs often have markedly different slopes for different states S (12). Some segments of a track may show driven motion along microtubules or actin filaments. These have a slope of 1.2–1.8. Other segments “in the same track” may have a slope near 0.0. In this case, we say that the particle is caged or trapped during this time segment. We will show below that the particles are not stationary while caged; they move back and forth within the cage but do not get anywhere. If we had used Eq. 1 and averaged the MSDs of 10 or so particles, there would be no objective basis for snipping the tracks into intervals corresponding to the different states. The MSD would be less noisy, but information about the timing of jumps from state to state would be lost.

Burov and colleagues (21) suggest that directional change from frame to frame in single-particle tracks is a better metric than the slope of the ensemble MSD. Because our MSD (Eq. 2) is for a single state, we can take their concept one step further: directional change can be used to understand the motion of single particles within a single state and to determine whether RNP particles show directional preferences while in that state. Faster frame rates (600 fps) were used in the experiments described here to better evaluate directionality within a cluster.

MATERIALS AND METHODS

Cell line, virus, and fluorescence microscopy

Human lung adenocarcinoma cells A549 (CVCL_0023; CCL-185; American Type Culture Collection (ATCC)) were cultured in Dulbecco's modified Eagle medium (DMEM; 11,995-065; Thermo Fisher Scientific) containing 7% heat-inactivated fetal calf serum (FCS; 26,140-079; Thermo Fisher Scientific) at 37°C and 5% CO₂ (11). The A549 cells were seeded in 35-mm glass-bottom dishes (70,670-02; Fisher Scientific) at 2.5×10^5 cells/dish and incubated at 37°C and 5% CO₂. After 24 h, the cells were infected with VSV-PeGFP (9) at a multiplicity of infection of 3. After a 1-h adsorption period at 37°C, the inoculum was removed, and cells were refed with FluoroBrite DMEM (A1896701; Thermo Fisher Scientific) containing 2% heat-inactivated FCS. Between 3 and 4 h post infection, cells were imaged using a Nikon Ti inverted epifluorescence microscope fitted with a 60× NA 1.4 oil-immersion objective. The microscope was set up on a modally damped optical breadboard (Newport) supported by four pneumatic vibration isolators (Thorlabs PWA075 or PWA090). The stage environment was set to 37°C and 5% CO₂. Excitation of eGFP in the RNP particles (400 copies of eGFP/particle) was accomplished by 495-nm light from the cyan light-emitting diode of a Spectra III fluorescence source (Lumencor, Beaverton, OR). The epi-illumination cube used to separate excitation from emission consisted of a 470/40-nm excitation filter, a 495-nm dichroic filter, and a 525/50-nm emission filter (Chroma 49,002, Chroma Inc, Bellows Falls, VT). A high-speed complementary metal oxide semiconductor camera (PCO Edge 5.5, PCO Inc., Romulus, MI, or Hamamatsu ORCA Fusion C14440-20UP; pixel size 6.5 μm × 6.5 μm for both cameras)

captured 16-bit tiff images. Magnification of the microscope/camera system was 9.23 pixels/μm. For maximum speed, images were transferred from the camera to the computer using a CameraLink framegrabber. Image acquisition and excitation light were coordinated by the camera software to ensure that fluorescent particles were exposed to excitation light only during image acquisition. The software used to achieve this coordination was PCO Camware 4 or Hamamatsu HCImageLive for the respective cameras. These software packages use time-to-live (TTL) communications between the camera, the light engine, and the computer for fastest communication.

The eGFP fluorophores in RNP particles bleach after a few seconds of irradiation. To maintain a constant signal/noise ratio at higher frame rates, source intensity must be increased, so track durations must be shortened. In practice, for 400 frames, this limits the duration of a track to 0.67 s at 600 fps and 4 s at 100 fps.

Tracking

Images of individual cells were cropped using ImageJ and saved as 16-bit tiff stacks containing 400 images. These files were analyzed by Video Spot Tracker (VST) software (version 8.01; Computer Integrated Systems for Microscopy and Manipulation, University of North Carolina-Chapel Hill (CISSM-UNC-CH)). RNPs were tracked at subpixel resolution using the VST cone function, precision 0.001, sample spacing 0.2, and radius 2 or 5. RNPs located between the cell nucleus (visible as a shadow in the fluorescence image) and the edge of the cell were chosen for analysis based on the following properties: 1) fluorescence intensity consistent with a single RNP particle, 2) did not encounter another particle within the detection radius, and 3) remained in the focal plane for the entire 400 frames. The vrpn files generated by VST were converted to MATLAB files using the VRPNLogToMatlab utility (version 05.00; CISSM-UNCCH).

Definition of directional angles θ and θ'

Each single-particle track, designated by $X(t)$, comprises values of x and y position recorded at 400 values of time $t(n)$; n is frame number. A diagram defining the change of direction θ for the particle at time $t(n)$ is shown in Fig. 1 A. This definition is the same as that used by Burov (21). The directional angle θ is based on three values of X : $X(n-1)$, $X(n)$, and $X(n+1)$ for offset $\tau = 1$. Two vectors $v(n)$ and $v(n+1)$ connect the three positions $X(n-1)$, $X(n)$, and $X(n+1)$. The angle $\theta(t, \tau)$ between the two vectors is shown in Fig. 1 A. Evaluation of $\theta(t, \tau)$ is then done by inverting the equation $\tan(\theta) = \sin(\theta)/\cos(\theta)$ using the four-quadrant MATLAB function `atan2`. Fig. 1 B is a diagram defining the three positions $X(n-1)$, $X(n)$, and X_{nuc} . X_{nuc} is the location of the center of the nucleus. It is manually selected from an image of the cell. The three positions are used to define two vectors, $v(n) = X(n) - X(n-1)$ and $v_{rnp,nuc} = X_{nuc} - X(n)$. The angle $\theta'(t, \tau)$ between the two vectors is shown in Fig. 2 B. Unlike the angle θ , the angle θ' between $v(n)$ and $v_{rnp,nuc}$ is a measure of a particular direction in the cell, in this case the direction toward the center of the nucleus.

Machine learning

An overview of the workflow of this code for a single-particle track is as follows:

For $K = 1:K_{max}$.

- 1) Construct a candidate model with K 2D Gaussians to create a Gaussian mixture model (gmm). Add hidden variables z_n (the probability of each point belonging to each state), $n = 1:N$, $N = \text{numFrames}$, to account for the initially unknown state of the system at each time point. Randomly initialize Θ , the vector of five parameters for each Gaussian (amplitude, center_x, center_y, variance_x, variance_y).
- 2) Compute an initial value for L , the lower bound of the probability that the model fits the data according to Eq. 10.70 in Bishop (17) derived

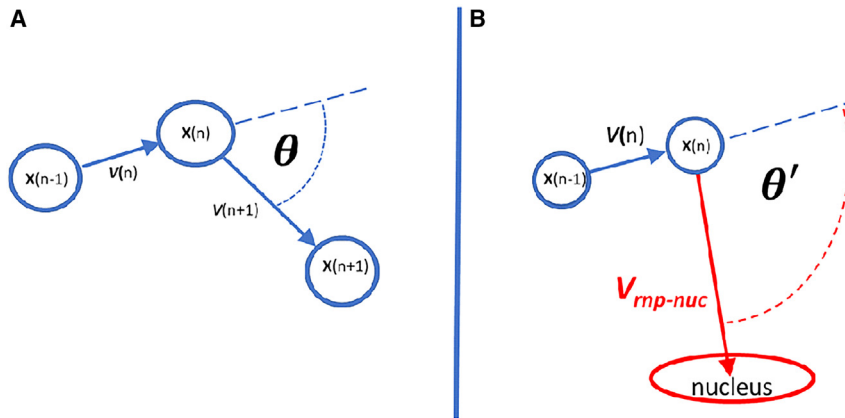


FIGURE 1 Definition of the angles θ and θ' . (A) definition of $\theta(t, \tau)$ based on particle locations only. (B) definition of $\theta'(t, \tau)$. For this panel, the second vector defines a direction in the cell.

from Bayes's theorem applied to 2D Gaussian mixture models. Adjust Θ using two steps of K-means to improve on the initial estimates of model parameters. Data points are assumed to be independent.

- 3) Construct a first-order hidden Markov model (HMM) to describe transitions between the hidden states z from frame to frame to account for the time dependence of the data. Apply a variational Bayes algorithm to this gmm-HMM model to improve Θ , and assign posterior probabilities $p(z_i = k | x_i, \Theta)$ to each z . The variational update equations are coupled, so updating must be done iteratively over alternating E and M steps.
- 4) If L has increased, continue the iteration over the E and M steps alternately.
- 5) If L has decreased, terminate the iteration for that value of K .
- 6) Keeping Θ fixed, adjust the responsibilities of each hidden variable z_n to improve L . Use the Viterbi and Beal algorithms to find the hidden state "sequence" with the highest global probability. This reassigns the hidden variable z of each frame from a set of K responsibilities to a single responsibility (state), which may not be the state with the highest responsibility. Very short bouts lasting for five frames or less between long bouts were reassigned to their adjacent bouts. This affected about 30% of tracks when $K = 5$ or 6.
- 7) Repeat steps 1–8 for the next value of K until $K = K_{max}$.
- 8) The result is an optimized set of Gaussian parameters, L , and responsibilities for all values of K . Compare the maximum value of L for all values of K . The value of K having the largest value of L is the most probable value of K for this track. Note that the most probable value of K was in most cases not the largest value of K tested because Bayes' theorem penalizes overfitting.
- 9) An "ellipse plot" of the $\pm 2\sigma$ boundaries of each state in xy space is generated. This is an invaluable visual tool to understand the location and size of each cluster discovered by steps 1–8.

Our code is based on a pattern recognition program called vbFRET (22) written to analyze Förster resonance energy transfer data, which is one dimensional. The core routines are from Ian Nabney's NETLAB library (23). We modified vbFRET so that it detects clustering in two dimensions. Our machine learning code, called vbTRACK_2D, is available on GitHub (GitHub: https://github.com/hoadleysilas/vbTRACK_2D). The notation is identical to that used in Bishop's book (17).

Fig. 2 shows highlights of the machine learning workflow for one particle acquired at 600 fps. A particularly simple track with two states and a hop in the approximate middle of the track was selected to illustrate the machine learning procedure. Fig. 2 A shows the raw data, the position of the particle during 400 frames. Fig. 2 B shows values of the probability lower bound L that each of the six models tested fits the data in Fig. 2 A. The model with two states has the best fit. Fig. 2 C shows the time dependence of the states occupied by this RNP particle, which occupies state 1 for the first 0.43 s and then occupies state 2 for the next 0.24 s. Fig. 2 D shows an "ellipse plot" for the optimized model with two states. Each state is represented by an ellipse

that is ± 2 standard deviations for the 2D distributions of data points in that state. Fig. 2 E shows log-log plots of the MSD for state 1 (red) and state 2 (green). The slope is essentially 0.0 for both states, which shows that this particle is trapped in one state or the other; it does not hop back and forth between the neighboring states more than once. The instrumental localization error, which can arise from stage vibrations, and the subpixel tracking error generated by Video Spot Tracker, were determined from the value of $\log(\text{MSD})$ for 1- μm latex beads stuck to the surface of a coverslip. This value is approximately -16 for $\log \tau = -2.7$ (G.H., unpublished data), so system noise can be safely neglected.

Number of tracks analyzed

A previous study (12) of 232 tracks at 100 fps found $n = 7, 14, 33, 28, 58$, and 76 for $K = 1:6$. Only seven out of the 232 tracks showed directionality in at least one bout. These values of n are small, so the precision is poor. In a later study on the effects of latrunculin and blebbistatin on RNP movement, many more tracks were obtained at 100 fps (11) with similar results. For the latrunculin controls, $n = 324$ tracks, and for the blebbistatin controls, $n = 358$.

The current paper on directionality uses new data acquired at a frame rate of 600 fps. The number of tracks analyzed was 101; the number of cells was 10. At this high frame rate, the number of states for 400 frames was one, two, or three. Tracks with more than three states were not observed because of the short duration of the data sets.

RESULTS

The distribution of directional change θ for particles in a trap

Fig. 3 A shows the observed track of a particle that stays within a single cluster. The compact "pin cushion" pattern of the cluster is strikingly different from simulated tracks of Brownian motion obtained by solving the Langevin equation for a particle with the mass and radius of an RNP suspended in homogeneous medium with the viscosity of cytosol (12). Fig. 3 B is the averaged histogram of the directional change of 12 RNP particles selected because machine learning showed that they stayed in one cluster for 400 frames. The histogram shows pronounced probability peaks at $\theta = -\pi$ and $+\pi$. However, in the center of the histogram ($\theta = 0$), the probability falls to 2.5%, which is below the 5% value expected for a random distribution with 20 bins. This shows that as a particle approaches

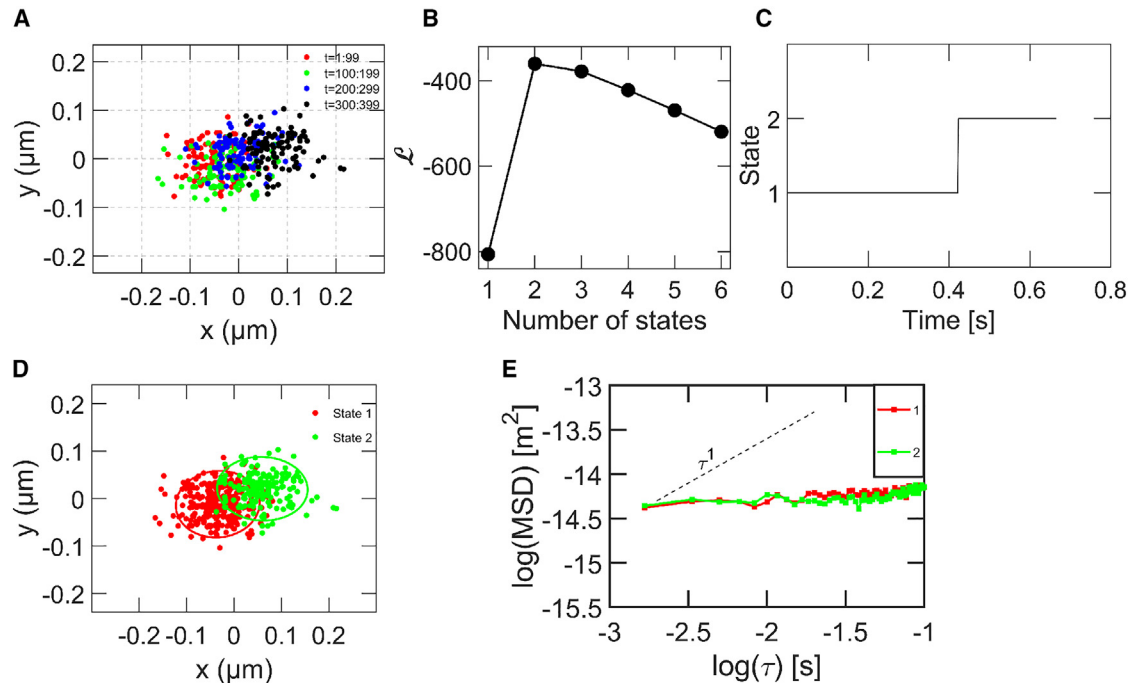


FIGURE 2 Highlights of the machine learning workflow for one particle acquired at 600 fps. A particularly simple track with two states and a hop in the approximate middle of the track was selected to illustrate the machine learning procedure. (A) Raw data, position of the particle during 400 frames. (B) Values of the probability lower bound L that each of the six models tested fits the data in (A). The model with two states has the best fit. (C) Time dependence of the states occupied by this RNP particle. The particle occupies state 1 for the first 0.43 s and then occupies state 2 for the next 0.24 s. (D) An “ellipse plot” for the optimized model with two states. Each state is represented by an ellipse that is ± 2 standard deviations for the 2D distributions of data points in that state. (E) Log-log plot of the MSD for state 1 (red) and state 2 (green).

the cage, it experiences a repulsive force that pushes the particle back toward the center of the cage. The cytoskeletal fibers themselves, which make up the cage, are not imaged in our experiments, and no convincing model of the interaction between a probe particle and the cytoskeletal fibers of its trap has been published yet.

It is also apparent in Fig. 3 B that the distribution of θ is largely independent of τ over the range of τ analyzed. At sufficiently short τ , it is expected that the distribution of θ for a confined particle would become uniform as in the case for a random walk (21). The data in Fig. 3 B indicate that even at the fastest frame rates, the motions of the RNPs within individual states are too rapid to accurately determine their diffusivities. This result is consistent with the independence of MSD with respect to τ over this range of τ (Fig. 2 E).

To shed additional light on the origin of the spikes seen in Fig. 3 A and the histogram in Fig. 3 B, we divided the cluster of 400 xy locations in Fig. 3 A into three annular rings. After calculating the mean x and mean y to define the center of the clump, the smallest ring includes all points with $0 < r < 0.08 \mu\text{m}$. The middle ring includes all points with $0.08 \leq r < 0.14 \mu\text{m}$. Finally, the largest ring includes all points with $r \geq 0.14 \mu\text{m}$; the edges were chosen to give each ring approximately the same number of particles. A histogram of directional change was then calculated for the data points in each ring. Finally, we averaged each of

the three histograms with corresponding histograms for 11 other trapped particles to reduce noise. The resultant three histograms are plotted in Fig. 3 C.

The histogram for the largest ring shows strong directionality because these particles are close to cytoskeletal fibers on one side of the cage. In addition to the thermal forces, which are nondirectional, the particles in this ring experience a strong force pushing them back toward the center of the cage. This strong force causes the particles to reverse direction, seen as spikes on the outer edges of the track in Fig. 3 A. The histogram for the smallest ring shows much weaker directionality because these particles are furthest from all the fibers, and the forces from fibers in different directions largely cancel one another. The histogram for the middle ring shows intermediate directionality. A mathematical model of rigid spherical beads trapped in mucus (24) provides some guidance through the thicket of issues that would be involved in modeling the physics of our flexible RNP particle trapped in a mesh of cytoskeletal fibers.

Histogram for θ'

Directional preference with respect to the cell coordinates, rather than local coordinates, is of great biological interest. Do RNPs move in the best direction to achieve final assembly and extrusion from the cell quickly? We can determine

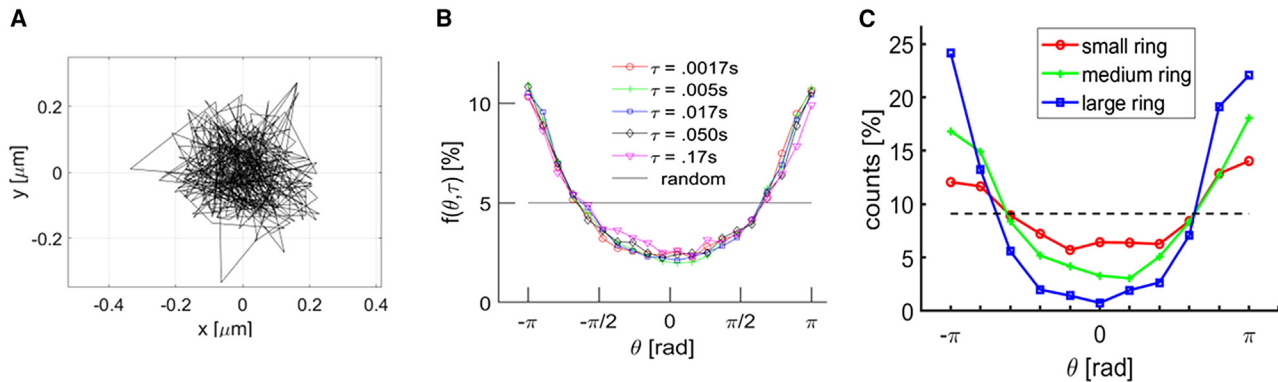


FIGURE 3 Particle in a single trap. (A) High magnification image of a single particle in a trap: 400 frames, 600 fps. (B) Averaged histogram of the directional angle θ for the trapped RNP in (A). To reduce noise, the histogram is an average for the particle in (A) and 11 other trapped particles. (C) Histograms of the directional angle θ after the 400 data points in (A) are divided into three groups defined by angular rings around the center of the cluster. The smallest ring is defined by $0 < r < 0.08 \mu\text{m}$. The middle ring is defined by $0.08 \leq r < 0.14 \mu\text{m}$. The largest ring is defined by $r \geq 0.14 \mu\text{m}$. The value of τ is .0017 s (1 frame at 600 fps).

directionality with respect to a vector from the RNP to the center of the nucleus, as described in [materials and methods](#). Fig. 4 shows a histogram of θ' for the data shown in Fig. 3 B. The angle θ' is defined in [materials and methods](#). Unlike the histogram for θ , the histogram for θ' is flat, within the observed noise. This shows that RNPs in living A549 cells probably do not sense the shortest route to the edge of the cell in terms of motion within clusters.

Radial distribution of the particles in a cluster

Another way to think about the forces causing clustering is to evaluate the particle distribution for a single cluster in polar coordinates, with radial coordinate r and azimuthal coordinate Φ . Assuming that the distribution is independent of Φ , the distribution in r can be evaluated. For data like that in Fig. 3 B, we determined the number of particles $N(r)$ within 20 concentric rings of identical width Δr . Dividing $N(r)$ by the area of each ring gives N/divA . Fig. 5 is a bar chart of N/divA against r . A single exponential, $y = A \cdot \exp(-r/r_{\text{fit}})$, fits the data well. The value of r_{fit} is $.061 \mu\text{m}$, with 95% confidence limits equal to $[\.059 \mu\text{m}, .063 \mu\text{m}]$.

DISCUSSION

Previous research from our group demonstrated that VSV RNPs move by active diffusion in mammalian cells. Because these RNPs are a little larger than the pores in the cytoskeletal mesh, they get trapped by the mesh. The mesh is composed of microtubules, intermediate filaments, and actin fibers. It is cross-linked by active (ATP-driven) myosin II motors, which make the trap dynamic. When the trap opens, the RNP can escape from its original pore to an adjacent pore; we call this hopping. Hopping events occurred, on average, every 0.2–1.0 s in the cells we studied. We wrote machine learning code that uses a variational Bayes, Gaussian mixture, hidden Markov algorithm to fit

single-particle tracks and to pinpoint objectively the time and place at which hopping occurs.

The present manuscript adds the determination of directionality of single states to the analysis of single-track data. This was proposed by Burov et al. (21) as a better way to detect the physical mechanisms underlying particle movement than that provided by the slope of the ensemble-averaged MSD. Burov proposed a mathematical definition of directionality and a useful way to display it, as a histogram over the range of possible directional angles, $-\pi$ to $+\pi$. Burov applied his procedure to single-particle tracks, which are often the result of several physical mechanisms at different times. Machine learning is a powerful tool to divide single-particle tracks into shorter tracks in which all frames are in the same state, so that we can apply Burov's procedure to single-state tracks. We applied this approach to particles in a trap. Histograms of directionality θ for a particle in a trap are not random at all; they show a strong preference for $-\pi$ and $+\pi$. This is evidence that as the particle in a single trap moves away from the center of the trap, it experiences a strong force pushing it back to the center of the trap. This ties studies of the mechanics of filaments in cytoplasm (7,8,13) to our studies of the motions of particles in a mesh of filaments. None of this is revealed in the slopes of ensemble-averaged MSD curves for similar experimental data because the MSD data for different states are combined.

Although our approach effectively models active 2D diffusion for particles in a certain size range, it does have limitations. Firstly, the experimental image data are limited to two dimensions, x and y . In flat cells, the structure of the cytoskeleton in the z direction may be different from the xy projection of the mesh. Another consequence of 2D image acquisition is that particles that move rapidly in the z direction will go out of focus and thus disappear. We have not analyzed such particles. This biases the results toward slower-moving particles moving in the xy plane. Secondly,

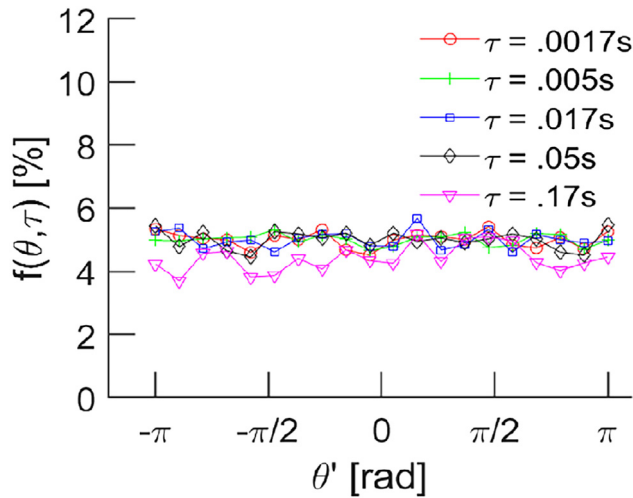


FIGURE 4 Histogram of the directional angle θ' for the same experimental data shown in Fig. 3 B.

our experimental data reveal only the fluorescently labeled particles; the cytoskeletal fibers are not seen because they are not labeled. Third, because of bleaching of the fluorophores, especially GFP, data are acquired for a short time only. The mechanisms of active diffusion in living cells over longer times are still uncertain.

Two software packages, VST and vbTRACK_2D, have been used in this paper. vbTRACK_2D uses machine learning to find the five parameters needed to describe each Gaussian. Future work on this problem might consider analyzing the hopping motions of RNP particles over longer times with nonparametric code packages (18–20).

The tracks of RNP particles are similar in appearance to the tracks of colloidal spheres in silica during the glass tran-

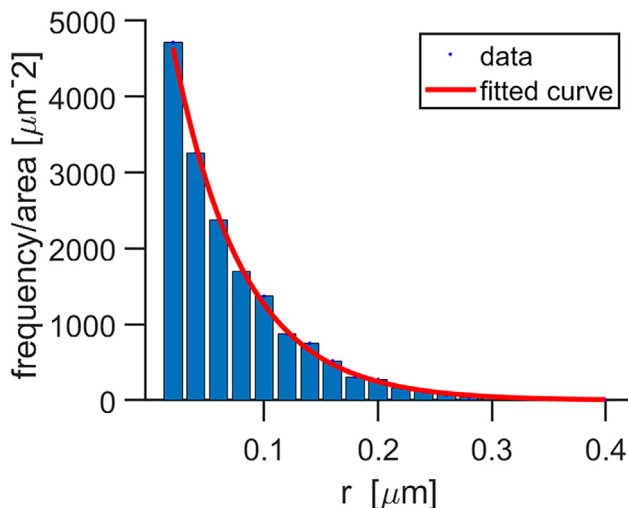


FIGURE 5 The radial distribution of an RNP within a trap. It is assumed that the distribution does not depend on azimuth. The red line shows the fit of a single exponential, $y = A \cdot \exp(-r/r_{\text{fit}})$ computed by MATLAB's fit function (The MathWorks, Natick MA). The characteristic length r_{fit} of the fitted exponential is 0.059 μm .

sition of this inorganic material (Fig. 3 in Chaudhuri, Berthier, and Kob (25)). Theoretical analysis shows that the silica particles carry out Gaussian displacements after a distribution of waiting times. Glassy behavior is sometimes proposed to describe cell mechanics. Two recent papers from the Mizuno lab (26, 27) use active microrheology data to propose that the cytosol of live HeLa cells shows glassy behavior at frequencies below 1 Hz because of crowding. The lower bound of the spatial resolution of this active microrheology study is unfortunately 1 μm , the diameter of the probe beads. The most probable size of the clusters revealed in our analysis of RNP transport is only 0.25 μm (11). The theoretical description of the glassy behavior of cytoplasm is further complicated by the multicomponent nature of the cytoskeletal mesh and cytosol of live cells. This spatial heterogeneity in physical properties is well documented (28).

AUTHOR CONTRIBUTIONS

Conceptualization and experimental design: G.H. and D.S.L. Cell and virus growth and sample preparation: S.J.M. and D.S.L. Image acquisition: R.O., K.C.S., D.S.L., and G.H. Coding: K.C.S., R.O., and G.H. Image processing: K.C.S., R.O., D.S.L., and G.H. Initial draft of the manuscript: K.C.S. Critical discussions and revisions of the manuscript: K.C.S., R.O., S.J.M., J.M., D.S.L., and G.H.

ACKNOWLEDGMENTS

The authors deeply mourn the death of coauthor Dr. George Holzwarth on March 13, 2024. Dr. Holzwarth was the senior (corresponding) author on this manuscript. This study was supported by a grant from the U.S. National Institute of Allergy and Infectious Diseases RO1 AI20623 (D.S.L., PI). K.C.S. and R.O. were supported by summer fellowships from Wake Forest University. S.J.M. was supported in part by National Institutes of Health Training Grant T32 AI007401. We also acknowledge use of the Cell Engineering Shared Resource of the Comprehensive Cancer Center of Wake Forest University, which is supported by National Cancer Institute Grant P30 CA012197.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Osborn, M., W. W. Franke, and K. Weber. 1977. Visualization of a system of filaments 7–10 nm thick in cultured cells of an epithelioid line (Pt K2) by immunofluorescence microscopy. *Proc. Natl. Acad. Sci. USA.* 74:2490–2494.
- Janson, L. W., K. Ragsdale, and K. Luby-Phelps. 1996. Mechanism and size cutoff for steric exclusion from actin-rich cytoplasmic domains. *Biophys. J.* 71:1228–1234.
- Chang, L., K. Barlan, ..., R. D. Goldman. 2009. The dynamic properties of intermediate filaments during organelle transport. *J. Cell Sci.* 122:2914–2923.
- Svitkina, T. 2022. Imaging Cytoskeleton Components by Electron Microscopy. *In Cytoskeleton.* R. H. Gavin, ed Springer US, New York, NY, pp. 25–52.

5. Einstein, A. 1956. *Investigations on the Theory of Brownian Motion*, 1905-1908. Dover Publications.
6. Lyles, D. S., M. O. McKenzie, and R. R. Hantgan. 1996. Stopped-flow, classical, and dynamic light scattering analysis of matrix protein binding to nucleocapsids of vesicular stomatitis virus. *Biochemistry*. 35:6508–6518.
7. Brangwynne, C. P., G. H. Koenderink, ..., D. A. Weitz. 2008. Cytoplasmic diffusion: molecular motors mix it up. *J. Cell Biol.* 183:583–587.
8. Brangwynne, C. P., G. H. Koenderink, ..., D. A. Weitz. 2009. Intracellular transport by active diffusion. *Trends Cell Biol.* 19:423–427.
9. Das, S. C., D. Nayak, ..., A. K. Pattnaik. 2006. Visualization of intracellular transport of vesicular stomatitis virus nucleocapsids in living cells. *J. Virol.* 80:6368–6377.
10. Thomas, D., W. W. Newcomb, ..., A. C. Steven. 1985. Mass and molecular composition of vesicular stomatitis virus: a scanning transmission electron microscopy analysis. *J. Virol.* 54:598–607.
11. Moran, S. J., S. Puckett, ..., D. S. Lyles. 2022. Dynamic Actin Filament Traps Mediate Active Diffusion of Vesicular Stomatitis Virus Ribonucleoproteins. *J. Virol.* 96:e0093422–22.
12. Holzwarth, G., A. Bhandari, ..., D. S. Lyles. 2020. Vesicular stomatitis virus nucleocapsids diffuse through cytoplasm by hopping from trap to trap in random directions. *Sci. Rep.* 10, 10643.
13. Mizuno, D., C. Tardin, ..., F. C. MacKintosh. 2007. Nonequilibrium mechanics of active cytoskeletal networks. *Science*. 315:370–373.
14. Anderson, S. J., C. Matsuda, ..., R. McGorty. 2019. Filament Rigidity Vies with Mesh Size in Determining Anomalous Diffusion in Cytoskeleton. *Biomacromolecules*. 20:4380–4388.
15. Anderson, S. J., J. Garamella, ..., R. M. Robertson-Anderson. 2021. Subtle changes in crosslinking drive diverse anomalous transport characteristics in actin–microtubule networks. *Soft Matter*. 17:4375–4385.
16. Kalafut, B., and K. Visscher. 2008. An objective, model-independent method for detection of non-uniform steps in noisy signals. *Comput. Phys. Commun.* 179:716–723.
17. Bishop, C. M. 2006. *Pattern Recognition and Machine Learning*. Springer, New York, NY.
18. Chen, Z., L. Geffroy, and J. S. Biteen. 2021. NOBIAS: Analyzing Anomalous Diffusion in Single-Molecule Tracks With Nonparametric Bayesian Inference. *Front. Bioinform.* 1, 742073.
19. Falcao, R. C., and D. Coombs. 2020. Diffusion analysis of single particle trajectories in a nonparametric framework. *Phys. Biol.* 17, 025001.
20. Granik, N., L. E. Weiss, ..., Y. Shechtman. 2019. Single-Particle Diffusion Characterization by Deep Learning. *Biophys. J.* 117:185–192.
21. Burov, S., S. M. A. Tabei, ..., A. R. Dinner. 2013. Distribution of directional change as a signature of complex dynamics. *Proc. Natl. Acad. Sci. USA*. 110:19689–19694.
22. Bronson, J. E., J. Fei, ..., C. H. Wiggins. 2009. Learning rates and states from biophysical time series: a Bayesian approach to model selection and single-molecule FRET data. *Biophys. J.* 97:3196–3205.
23. Nabney, I. T. 2002. *NETLAB: Algorithms of Pattern Recognition*. Springer, London.
24. Hansing, J., and R. R. Netz. 2018. Particle Trapping Mechanisms Are Different in Spatially Ordered and Disordered Interacting Gels. *Biophys. J.* 114:2653–2664.
25. Chaudhuri, P., L. Berthier, and W. Kob. 2007. Universal Nature of Particle Displacements Close to the Glass and Jamming Transitions. *Phys. Rev. Lett.* 99, 060904.
26. Ebata, H., K. Umeda, ..., D. Mizuno. 2023. Activity-dependent glassy cell mechanics I: Mechanical properties measured with active microrheology. *Biophys. J.* 122:1781–1793.
27. Umeda, K., K. Nishizawa, ..., D. Mizuno. 2023. Activity-dependent glassy cell mechanics II: Nonthermal fluctuations under metabolic activity. *Biophys. J.* 122:4395–4413.
28. Garner, R. M., A. T. Molines, ..., F. Chang. 2023. Vast heterogeneity in cytoplasmic diffusion rates revealed by nanorheology and Doppelgaenger simulations. *Biophys. J.* 122:767–783.