

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Single-Molecule and Super-Resolution Microscopy of Fast Intracellular Diffusion

Permalink

<https://escholarship.org/uc/item/673781pp>

Author

Choi, Alexander Akira

Publication Date

2024

Peer reviewed|Thesis/dissertation

Single-Molecule and Super-Resolution Microscopy
of Fast Intracellular Diffusion

By

Alexander Akira Choi

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ke Xu, Chair
Professor Evan R. Williams
Professor Seung-Wuk Lee

Spring 2024

Single-Molecule and Super-Resolution Microscopy
of Fast Intracellular Diffusion

Copyright © 2024

Alexander Akira Choi

Abstract

Single-Molecule and Super-Resolution Microscopy of Fast Intracellular Diffusion

by

Alexander Akira Choi

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Ke Xu, Chair

Intracellular diffusion defines the kinetics and timescales of biological processes. However, due to the fast motion of biomolecules and the sub-micron scale of intracellular structures, elucidating how unbound molecules diffuse through the cell is challenging. Here, I introduce a recently developed techniques, single-molecule displacement/diffusivity mapping (SMdM), which capitalizes on super-resolution capabilities of single-molecule localization microscopy (SMLM), while uniquely maps out the motion and diffusivity of freely diffusing molecules at the nanoscale.

In this dissertation, I present my recent efforts to transform SMdM into a powerful quantification tool by leveraging high-throughput single-molecule statistics. I first investigated enzymatic diffusion during catalysis challenging existing theories by utilizing high statistics of sub-millisecond displacements of unhindered single enzymes freely diffusing in common buffers to quantify D for four enzymes under catalytic turnovers to show no significant changes in D during catalysis, and additionally formulated how $\sim \pm 1\%$ D precisions can be achieved.

By similarly leveraging high-throughput single-molecule statistics to determine D , in the subsequent study, I showed that SMdM can characterize D up to $350 \mu\text{m}^2/\text{s}$ for water-soluble dyes and small-molecule metabolites. I unveil that small-molecule intracellular diffusion is dominated by vast regions of high diffusivity $\sim 60\text{-}70\%$ of that *in vitro*, lifting a paradoxical speed limit of intracellular diffusion as suggested by previous studies.

Next, I uncover strong *in vitro* charge-driven protein-protein interactions replicating the sign-asymmetric charge effects on intracellular diffusivity discovered in mammalian cells. By converting between D to molecular weight, we were provide novel insight into how positively charged protein at the limit of low diffuser concentration avoid aggregation and coacervation

inside of cytoplasm, by modulating the concentration of negatively charged protein. Here we showed the formation *in situ* that positively charged diffusers drag a single monolayer of negatively charged macromolecules forming a net negatively charged complex, shedding new insight into mechanisms that drive charged-based diffusion interactions inside of the cell.

In collaboration with Rebecca Heald's group, I work with *Xenopus laevis* extracts to study how both charge and molecular size impacts diffusivity of macromolecules in membrane-less cytoplasm for a wide range of protein and demonstrate that negatively charged macromolecules diffuse approximately 45% slower in extract, but positively charged macromolecules diffuse substantially slower. We then demonstrate how nonspecific RNA interactions drive the solubilization of positively charged macromolecules, where depletion of RNA resulted in release of positively charged ribosomes that resulted in cytoplasmic aggregation. Lastly, we show that the filamentous actin cytoskeleton induces macromolecular crowding demonstrating nanoscale intracellular structures shape diffusion processes inside of the cell.

Collectively, by uniquely mapping out fast diffusion at the single-molecule and super-resolution levels, my research sheds new insight on the fundamental laws governing molecular transport within living cells.

Acknowledgements

To Ke, who has been an amazing mentor to me throughout my career and has provided critical support and guidance which has been instrumental in much of my success. I am inspired by his creativity, enthusiasm, and approach to research. His openness to pursue new directions has made my PhD unpredictable, but exciting.

To all past and previous group members for mentoring, motivating, and being good friends. I have been fortunate to have worked with many of you on various projects and have learned much but more importantly, it has made these projects more fun. I have also been responsible for mentoring a few of you to practice my pedagogy. I hope my teaching was mildly beneficial, and I apologize if I was too terse at times.

To collaborators for enabling me to learn new systems such as *Xenopus laevis* egg extracts, HP1a phase separation in *Drosophila melanogaster*, and perovskite nanocrystals. I found that each collaboration enriched my viewpoint on biology or material science and has made me a more well-rounded researcher.

To past mentors at Georgia Tech who inspired and encouraged me to pursue physical chemistry and research. It was motivating to meet so many people excited to talk and work so hard on research.

To my family for providing support throughout this journey. I am fortunate to have parents who are both scientists that can engage me in discussion on research such as on toucans and provide career guidance. Back in high school, my dad advised me against going into research because it is a difficult and competitive career, which I largely ignored. I think he has since come around on his opinion. It is unfortunate that my sister ruptured her Achilles tendon right before my graduation, but her independence and confidence towards her pursuit to become a pilot is proof that you can achieve anything if you put your mind to it.

Dedication

To my parents

Table of Contents

Acknowledgements	iii
Dedication	iv
List of Figures	vii
List of Tables	ix
Chapter 1: Introduction	1
1.1 Thesis Introduction	1
1.2. Single-Molecule Imaging and Microscopy	1
1.2.1. Why Fluorescent Microscopy for Biological Imaging	1
1.2.2. Quantum Mechanics of Light, Absorption and Fluorescence	1
1.2.3 Fluorescence Microscopy and the Diffraction Limit	2
1.3. Single-Molecule Localization Microscopy	3
1.3.1. Single-Molecule Localization Microscopy (SMLM) and Stochastic Optical Reconstruction Microscopy (STORM)	3
1.3.2. Functional Super-Resolution Microscopy	4
1.4. Mapping Intracellular Diffusion	5
1.4.1. Motion and Diffusion	5
1.4.2 Super-Molecule displacement Mapping (SMdM)	6
1.5. Thesis Overview	7
Chapter 2: Displacement statistics of unhindered single molecules show no enhanced diffusion in enzymatic reactions	10
2.1. Introduction	10
2.2 Results and Discussion	11
2.3 Conclusion	15
2.4 Supplemental Materials and Methods	16
Chapter 3: Single-molecule displacement mapping indicates unhindered intracellular diffusion of small (<~1 kDa) solutes	27
3.1. Introduction	27
3.2 Results and Discussion	28
3.3 Conclusion	33
3.4 Materials and Methods	34
Chapter 4: Single-molecule diffusivity quantification unveils ubiquitous net charge-driven protein-protein interaction	40
4.1 Introduction	40
4.2 Results and Discussion	41
4.3 Conclusion	46

4.4 Supplemental Material and Methods	46
Chapter 5: Single-molecule diffusivity quantification in <i>Xenopus laevis</i> egg extract elucidates cytoplasmic molecular environments	52
5.1. Introduction.....	52
5.2 Results and Discussion	52
5.3 Conclusion	53
5.4 Material and Methods	62
Chapter 6: Conclusion and Outlooks	69
6.1. Concluding Remarks.....	69
6.2 Outlooks.....	69
6.2.1 Live-Cell Single-Molecule Diffusion Mapping of Ribonucleoprotein	69
6.2.2 Single-Molecule Microscope to Measure Net Charge.	73
References	75

List of Figures

Figure 1.1. Single-Molecule Localization Microscopy.....	3
Figure 1.2. Single-Molecule displacement Mapping.....	6
Figure 2.1. High-throughput statistics of unhindered single-molecule displacements.....	11
Figure 2.2. Precise determination of diffusivity <i>via</i> single-molecule statistics.....	12
Figure 2.3. No substantial change in single-molecule displacements observed for catalase under turnover.....	13
Figure 2.4. Results on urease, alkaline phosphatase, and aldolase also show no changes in single-molecule displacements under catalytic turnover.....	14
Figure 2.5. Typical timing waveforms in this work.....	22
Figure 2.6. Additional characterizations of single-molecule displacements from SMdM, for Cy3B-labeled catalase diffusing in PBS.....	23
Figure 2.7. SMdM results for Cy3B-labeled catalase at a longer pulse separation of $\Delta t = 1$ ms and an increased single-molecule density.....	24
Figure 2.8. Activity assays for Cy3B-labeled and unlabeled enzymes.....	25
Figure 2.9. Simulated substrate concentrations, reaction rates, and Gibbs energy dissipation rates for the different enzymes in our SMdM experiments.....	26
Figure 3.1. SMdM quantifies the fast diffusion of small solutes <i>in vitro</i>	29
Figure 3.2. SMdM of the water-soluble dye SR101 in primary rat hippocampal astrocytes.....	30
Figure 3.3. SMdM of Cy3B and AF647 dyes delivered into the A549 epithelial cell line via graphene-based electroporation.....	31
Figure 3.4. SMdM D maps of graphene-electroporation-delivered ATP-Cy3B (a), ATP-AF647 (b), and cAMP-CF640R (c) in A549 cells.....	32
Figure 3.5. Oscilloscope-measured typical timing waveforms in this study.....	36
Figure 3.6. SMdM results on Cy3B diffusing in PBS containing different amounts of glycerol...37	
Figure 3.7. SMdM under the same 561-nm excitation laser for an astrocyte not loaded with SR101 but expressing the mEos3.2 FP.....	38
Figure 3.8. Two-dimensional vectorial single-molecule displacement analysis.....	39
Figure 4.1. Single-molecule diffusivity quantification unveils strong HEWL-BSA interactions in the low-concentration limit.....	41
Figure 4.2. The HEWL-BSA interaction is driven by electrostatic attractions based on the protein net charge.....	43
Figure 4.3. Generalizing charge-driven interactions to HEWL-ovalbumin and avidin-BSA systems.....	44
Figure 4.4. SMdM survey of the diffusivity of different proteins in BSA solutions shows consistent slowdowns for positively charged species.....	45
Figure 4.5. SMdM-measured D values of the Cy3B-labeled proteins in Fig. 4 in the starting buffer, plotted against their molecular weights.....	50
Figure 4.6. SMdM of the Cy3B dye with BSA addition.....	51

Figure 5.1. SMdM-based protein diffusivity quantification in the <i>Xenopus</i> egg extract versus in PBS.....	54
Figure 5.2. Net-charge effects in protein diffusion in the <i>Xenopus</i> egg extract.....	55
Figure 5.3. SMdM diffusivity mapping of RNase-treated egg extracts further underscores charge-sign asymmetry of the cytoplasmic environment.....	57
Figure 5.4. Importance of the actin cytoskeletal network in molecular size-dependent diffusion suppression.....	59
Figure 5.5 SMdM diffusivity mapping of Cy3B-labeled HEWL in untreated and BSA-added <i>Xenopus</i> egg extracts.....	68
Figure 6.1 <i>In vitro</i> SMdM of RNA-protein complexes.....	69
Figure 6.2. Live Cell SMdM of Cas9-mEos4B in the nucleus.....	71
Figure 6.3. Live Cell SMdM of Cas13-mEos4B in the nucleus.....	71
Figure 6.4. Live Cell SMdM of Telomerase-mEos4B.....	72
Figure 6.5. Single-molecule <i>electrophoresis</i> microscopy.....	73

List of Tables

Table 2.1. SMdM-determined D values for different protein samples in PBS.....	20
Table 4.1. Accounting for molecular charges in the mammalian cytoplasm.....	48
Table 4.2. Proteins used in this work, their Uniprot IDs, molecular weights (MWs), estimated net charges at pH 7.3, and sources.....	50
Table 5.1. Diffuser proteins used in this work, their Uniprot IDs, molecular weights (MWs), estimated net charges, and sources.....	64
Table 5.2. List of top cytoplasmic proteins in the <i>Xenopus</i> egg cytoplasm extract and their estimated net charge q	66
Table 5.3. List of top RNA-binding and ribosomal proteins in the <i>Xenopus</i> egg cytoplasm extract as detected by mass spectrometry ¹ , and their estimated net charges, generated similarly as Table 5.2.....	67

Chapter 1: Introduction

1.1 Thesis Introduction

The proverbs ‘seeing is believing’ and ‘a picture is worth a thousand words’ highlight the pivotal role imaging approaches play across every scientific discipline, from biology to astronomy. In transforming highly complex and intricate datasets into accessible visible representations, imaging not only elucidates the spatial arrangements of objects in an image, but also unveils rich multidimensional information through color, brightness/contrast, and motion through videos. In this thesis, I will present how modern imaging advancements namely in super-resolution microscopy through the development of a new microscopy-approach called single-molecule *displacement* mapping (SMdM)² enables imaging of sub-microsecond dynamic and motion at nanoscale resolutions. Using this approach, I will highlight how motion at the nanoscale sheds new insight into fundamental principles such as molecular charge and size, which govern critical cellular phenomena and structure.

1.2. Single-Molecule Imaging and Microscopy

1.2.1. Why Fluorescent Microscopy for Biological Imaging

Fluorescence microscopy is a powerful tool in biological imaging, offering distinct advantages for imaging biomolecules in the complex cellular environment. Imaging contrast using absorbed light to identify distinct and sometimes rare species within a cell is challenging given the high concentrations of proteins and other biomolecules, all of which are typically similar in their absorption properties. Visible light emitted by fluorescence is a powerful means of identifying analytes inside of cells due to its low absorption by water, the most ubiquitous molecule in biology, and biomolecules generally low autofluorescence. As a result, fluorescence microscopy using bright, fluorescent light emitted from specific targets such as proteins, DNA, RNA, and small-molecule metabolites provides insight into the activities of individual analytes without the distraction from the plethora of other unlabeled species.

1.2.2. Quantum Mechanics of Light, Absorption and Fluorescence

Light consists of photons which act as both particles and waves, fundamental to quantum mechanics. The energy of a photon is defined by its wavelength, where shorter wavelength light is higher in energy. Absorption is the result of a light-matter interaction whereupon photons that match the energy difference of typically electrons in the matter. Absorption is characterized by excitation from a ground state electronic transition (S_0) to an excited electronic state (S_1). As a result of the ~femtosecond timescales with which absorption happens at as described by the Franck-Condon principle, excited states evolve by electronic excitations that result in minimal change in nuclear position. Following absorption into S_1 , excited state electrons have various modes with which de-excitation back to S_0 can occur, defined by the timescales and therefore the probabilities of these processes. Because of the shorter lifetime of internal conversions (~fs-ps)³

when compared to fluorescence (~ns), vibrational de-excitation and nuclear relaxations result in a ground level vibrational excited state, $S_{1,0}$. While de-excitation can continue along internal conversions, due to the typically large energy difference between ground state vibrational $S_{1,0}$ and S_0 , it becomes more probabilistic that fluorescence occurs. This energy difference between $S_{1,0}$ and S_0 defines the fluorescence wavelength.

1.2.3 Fluorescence Microscopy and the Diffraction Limit

As previously described, fluorescence microscopy is a powerful tool enabling strong contrast or high signal to noise (S/N) to enable biological imaging in complex environments, coupled with its high specific enabled by modern labeling approaches or use of fluorescent proteins. Samples containing fluorophores are illuminated with light necessary for absorption of light by the fluorophore to promote fluorescence. Given that fluorescence is redshifted per Kasha's rule from the absorbed frequencies of light, fluorescence can be collected by filtering the excitation light from the fluorescence to give low background and high S/N for the fluorescence signal.

While there are a variety of fluorescence microscopy approaches such as epifluorescence microscopy or confocal microscopy, I will focus on total internal reflection fluorescence (TIRF) microscopy used to illuminate the interface of samples. TIRF works based on Snell's law which states that the critical angle or 'TIRF' angle with which the light is reflected is described by equation 2. This is significant as molecules located at this interface are excited by the light as described by evanescent waves yielding low background from molecules above the interface. By moderately changing the angle from the TIRF angle, fluorophores deeper into the sample can be illuminated while still conferring the S/N gain from being close to the TIRF angle, sometimes referred to as 'dirty' TIRF.

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \text{ (eqn. 1)}$$

$$n_1/n_2 = \sin \theta_2 \text{ (eqn. 2)}$$

TIRF microscopy is a particularly attractive approach for imaging cellular samples because it reduces background fluorescence from regions that are out-of-focus, thus improving S/N, and results in higher spatial resolution imaging. In addition, cells are quite thin and flat, such that many features can be viewed as they are near the interface. Indeed, for thicker samples such as tissues, alternative microscopy approaches should be used to image such samples.

One question that arises is the spatial resolution of fluorescence microscopy. A consequence of Young's double-slit experiment was the observation of light diffraction and the ability to differentiate between two distinct entities. Ernst Abbe defined the resolution of an Airy disk and therefore a microscope as shown in equation 3. Where $R_{x,y}$ represents the angle required X,Y resolution, λ is the wavelength of light, and NA is numerical aperture.

$$R_{x,y} = \lambda/2NA \text{ (eqn. 3)}$$

For visible light which spans 400-700 nm, a maximum of resolutions $\sim 200\text{-}350$ nm in x,y can be achieved due to this diffraction limit of light. While such a resolution enables observation of many structures, it limits imaging of particularly cell biology where the typical structure inside of cells is far below <200 nm.

1.3. Single-Molecule Localization Microscopy

1.3.1. Single-Molecule Localization Microscopy (SMLM) and Stochastic Optical Reconstruction Microscopy (STORM)

Over the past few decades, various approaches have emerged to bypass the diffraction limit imposed by the wavelength of light. One such class of techniques is called single-molecule localization microscopy (SMLM), which enables precise localization of individual molecules with resolutions below the diffraction limit.

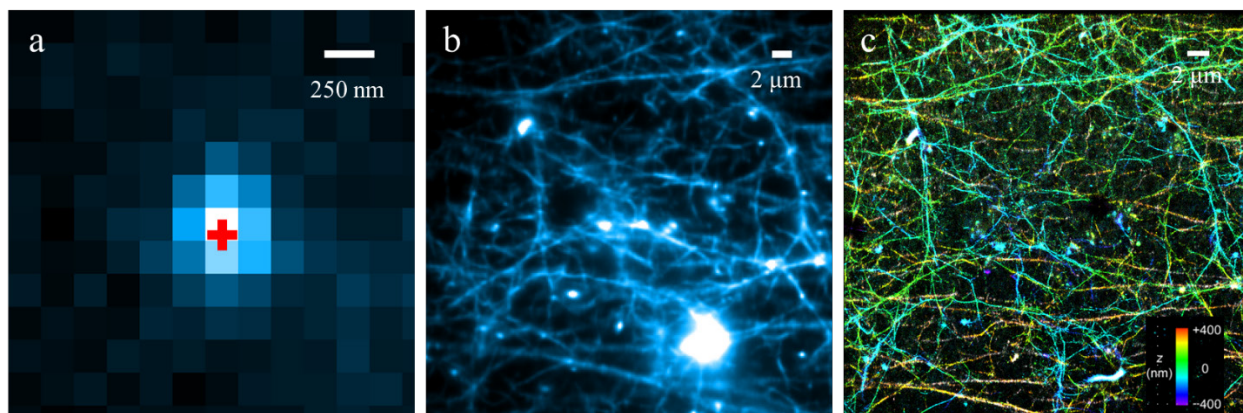


Figure 1.1: Single-Molecule Localization Microscopy. (a) Single molecules of phalloidin-Alexa Fluor 647 imaged on a camera with their position marked by red crosses. (b) Conventional fluorescence image of actin filaments stained by phalloidin-Alexa Fluor 647 in *Xenopus laevis* extract. (c) STORM image of the actin filaments (b) to show higher image resolution by image reconstruction.

When single-molecules are imaged on a camera (Figure 1.1a), their images are diffraction-limited as shown by their blurred pattern referred to as the point spread function (PSF). However, despite this limitation, their positions can be determined with higher precision than what is defined by the diffraction limit. This can be achieved by identifying the center of this PSF using gaussian fitting to define a new localization precision.

$$\text{Localization precision: } \delta \approx \Delta / N^{1/2} \text{ (eqn. 4)}$$

Given that a single photon imaged on a camera would give a diffraction pattern due the wave-like nature of light, the localization precision can be defined by the shot noise of the system. Here, the localization precision is inversely proportional to the square root of the number of photons collected from a single molecule. For fluorescent light that is 600 nm, to achieve 10 nm resolution, ~ 900 photons need to be collected. Camera background limits the lower bound of the localization precision that can be achieved as described by Thompson et al.⁴, where sub-10 nm

resolution cannot be achieved even when collecting 1000s of photons from the same single molecule. For the popular single-molecule dye, Alexa Fluor 647 hundreds of thousands of photons can be collected within seconds of imaging⁵, quickly yielding highly precise localization precisions for single-molecule localization, demonstrating that by utilizing bright single-molecules quickly yield high localization precision enabling high resolution imaging and is important later when discussing single-molecule tracking of quickly diffusing species.

STORM⁶ and Photoactivatable Localization Microscopy (PALM)⁷ are a class of SMLM that utilizes the blinking of single-molecules to transiently turn on and off single-molecules enabling sparsely dense fluorescent samples, critical for observing single-molecules. Here, in this example (Figure 1.1b), actin filaments inside of *Xenopus laevis* extract are stained by Alexa Fluor 647, the filaments appear blurred because it is diffraction limited. By controlling the ‘blinking’ buffer concentration that controls dark and bright state of fluorophores and UV light intensity, single-molecule concentrations can be precisely controlled to enable SMLM. Single-molecule images using the positions obtained through SMLM can be reconstructed to form a ‘super-resolution’ image as shown in Fig 1.1c, where the structure of filaments is much less blurred and more well-resolved as a result of SMLM and the blinking properties of fluorophores. Overall, applying principles of SMLM to reconstruct super-resolution images provides a powerful tool to observe the ultrastructure within cell biology.

1.3.2. Functional Super-Resolution Microscopy

While super-resolution microscopy is a power tool to resolve structures at nanometer resolutions, due to the complexity of cell biology, understanding structures may be insufficient in understanding how nanoscale structures impact biological function. Just like nanoscale intracellular ultrastructure, due to the nonequilibrium nature of the intracellular environment, physicochemical parameters such as chemical polarity, diffusivity and molecular concentrations may be differentially organized to dictate cellular function through the formation of ‘microenvironments’ or ‘nanodomains’ necessitating super-resolution chemical or physical imaging to resolve.

Functional super-resolution microscopy^{8, 9} provides a new possibility to encode physicochemical parameters into details of single-molecules to enable multidimensional super-resolution where nanoscale structures can be related to variations in chemical or physical properties. Encoding physicochemical properties requires utilization of properties of light that are sensitive to the parameter with which is being encoded. In previous work, spectrally resolved super-resolution microscopy of the solvachromatic dye Nile Red was used to determine the hydrophobicity of cellular membranes in live cells by encoding the emission spectra of the single-molecules by hydrophobicity¹⁰. However, other properties of single-molecule properties can be used such as polarization, lifetime, or motion to encode for physicochemical properties for super-resolution imaging. While there are many ways with which single-molecules can be encoded for

physicochemical parameters, this thesis will focus on how the motion of single-molecules can be used to understand how specifically its charge and size influence intracellular dynamics.

1.4. Mapping Intracellular Diffusion

1.4.1. Motion and Diffusion

Diffusion in biology defines the spread of molecules as a function of time. Different from speed or velocity, diffusion provides a better description of random or Brownian motion that molecules take as they traverse intracellular space.

Diffusion is a critical parameter to consider in cellular biology as it defines the timescale and kinetics that processes occur at. One feature of the intracellular environment is that it is a highly crowded space containing high concentrations of macromolecules including protein and RNA that interact and significantly slowdown diffusion inside of intracellular spaces, and therefore the speed with which things are transported. While the cell has evolved active transport-based mechanisms that defy random-walk Brownian motion, because these processes involve transport by transporter proteins, these processes are only faster than free diffusion inside of the cell over long distance scales or for large diffusers. For the typical protein, free diffusion remains both a fast and ‘cheap’ means of delivery as it requires no energy input by ATP. As a result, characterizing free diffusion and processes that shape free diffusion is critical when understanding how intracellular structures shape intracellular transport and kinetics.

There exist various tools to measure intracellular diffusivity classified into either bulk-based approaches or single-molecule approaches. Bulk based approaches such as fluorescence recovery after photobleaching (FRAP) or fluorescence correlation spectroscopy are powerful tools to measure diffusivity of a wide range of fluorescent molecules from large proteins to small-molecules, as their name suggests, because of how they use typically large areas to determine diffusion of molecules in and out of the region, are limited in their spatial resolution, sometimes insufficient to resolve how diffraction-limited structures shape diffusion within their region. Single-molecule approaches such as single-molecule tracking leverage high resolution diffusivity mapping as it uses SMLM to precisely identify positions of single-molecules, it is limited in the scope of molecules with which it can look at, as molecules that diffuse too quickly cannot be easily tracked, and therefore their diffusivity cannot be measured. As a result, single-molecule tracking has been primarily applied to membrane or chromosomal diffusion which is often quite slow and often confined in 2D such as in membrane diffusion. However, it is incapable of measuring free diffusion of proteins inside of the cell, critical when addressing how molecules diffuse in the cytoplasm. As a result, it is critical to develop a technique capable of measuring diffusion of quickly moving species and map at super-resolution scales to understand how intracellular structures and parameter shape diffusivity at the nanoscale.

1.4.2 Super-Molecule displacement Mapping (SMdM)

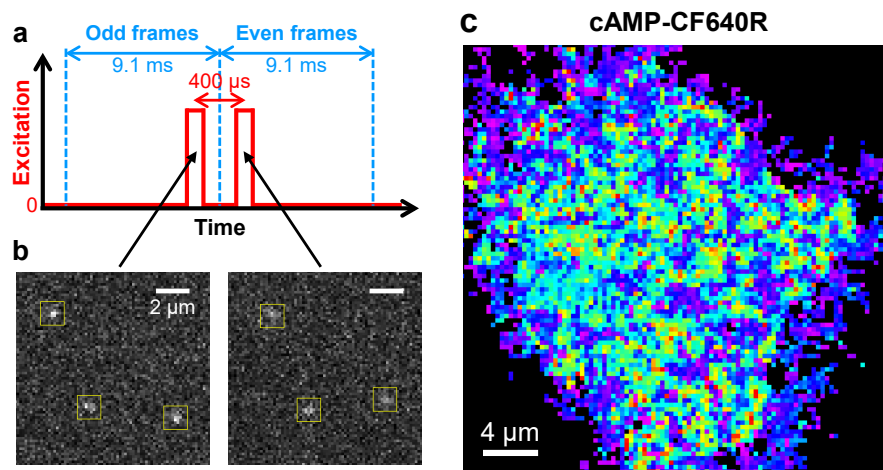


Figure 1.2. Single-Molecule *displacement* Mapping. (a) Schematics: To detect the fast motion, tandem stroboscopic excitation pulses of 200 μs duration are applied across paired camera frames at a $\Delta t = 400 \mu\text{s}$ center-to-center separation. (b) SMdM diffusion spatial map of cAMP-CF640R resolves nanoscale features that drive diffusion heterogeneity in live A549 cells. Figure reproduced from A.A. Choi et. al.¹¹, with permission.

Although single-molecule tracking has previously enabled motion and diffusion measurements, it is often limited to the slow diffusion of sparse molecules, *e.g.*, diffusion in membranes or of motor protein, both of which are orders of magnitude slower than the diffusion of most biomolecules inside the cell such as proteins, mRNA or small molecules. Meanwhile, fluorescence correlation spectroscopy works with faster diffusion, but provides limited spatial resolutions and mapping capabilities. Here, my work has discussed the development of a novel microscope approach capable of quantifying the diffusion of fast single molecules at sub-diffraction limit resolutions to uncover the physical principles that define motion inside the cell.

In 2020, our lab published a novel microscopy method, single-molecule displacement/diffusivity mapping (SMdM)¹. In a typical single-molecule tracking experiment, single-molecules will be excited for the duration of a camera frame, which for a standard EM-CCD camera running at 110 Hz ($\sim 10 \text{ ms}$), this is insufficiently slow to track quickly diffusing species as they will have moved out-of-focus before they can be observed in the subsequent frame. By applying stroboscopic excitation pulses evenly, but closely spaced between camera frames, single-molecules can be tracked with much lower time separations than what is defined by the frame rate of a camera. In this example, by evenly spacing illumination pulses 0.4 ms apart, quickly diffusing protein can be tracked such that it can be identified and localized using SMLM. Based on the coordinates of these single-molecules, the displacement can be determined for each paired single-molecule. Repeating this tandem-pulse illumination scheme can enable quick accumulation of displacement statistics to fit to determine D coefficient.

$$P(r) = \frac{2r}{a} \exp\left(-\frac{r^2}{a}\right) + br \quad (\text{eqn. 4})$$

Here, r is the single-molecule displacement at fixed time interval Δt , $a = 4D\Delta t$, and b is a background term accounting for mismatched molecules.

By locally spatially binning displacements, the D coefficient can be locally fit to render a diffusion spatial map. By selecting bin sizes that are below the diffraction limit such as 200 nm, super-resolution diffusion mapping can be used to resolve sub-diffraction features within the cell to understand how such structural properties impact diffusivity at this size scale.

1.5. Thesis Overview

This thesis will highlight four projects that describe the development and application of SMdM. I will then highlight in the final chapter opportunities in single-molecule diffusivity mapping to answer a variety of questions ranging from cell biology to development of new functional super-resolution techniques.

In Chapter 2, I discuss the use of SMdM to provide insight into whether catalytic reactions enhance the diffusion coefficients D of enzymes. Through high statistics of the transient (600 μs) displacements of unhindered single molecules freely diffusing in common buffers, we here quantify D for four enzymes under catalytic turnovers. We thus formulate how $\sim \pm 1\%$ precisions may be achieved for D , and show no changes in diffusivity for catalase, urease, aldolase, and alkaline phosphatase under the application of wide concentration ranges of substrates. Our single-molecule approach thus overcomes potential limitations and artifacts underscored by recent studies to show no enhanced diffusion in enzymatic reactions. The work is reproduced with permission from the following reference¹².

In Chapter 3, I discuss a fundamentally important question regarding the intracellular diffusion of small ($\lesssim 1$ kDa) solutes, which has been challenging to study due to challenges in labeling and measurement, due to its fast diffusivity. Here we quantify and spatially map the translational diffusion patterns of small solutes in mammalian cells by integrating several recent advances. In particular, by executing tandem stroboscopic illumination pulses down to 400 μs separation, we extend single-molecule displacement/diffusivity mapping (SMdM), a super-resolution diffusion quantification tool, to small solutes with high diffusion coefficients D of $>300 \mu\text{m}^2/\text{s}$. We thus show that for multiple water-soluble dyes and dye-tagged nucleotides, intracellular diffusion is dominated by vast regions of high diffusivity $\sim 60\text{--}70\%$ of that *in vitro*, up to $\sim 250 \mu\text{m}^2/\text{s}$ in the fastest cases. Meanwhile, we also visualize sub-micrometer foci of substantial slowdowns in diffusion, thus underscoring the importance of spatially resolving the local diffusion behavior. Together, these results suggest that the intracellular diffusion of small solutes is only modestly scaled down by the slightly higher viscosity of the cytosol over water but otherwise not further hindered by macromolecular crowding. We thus lift a paradoxically low speed limit for

intracellular diffusion suggested by previous experiments. The work is reproduced with permission from the following reference¹¹.

In Chapter 4, I discuss the formation of protein nanoclusters driven by strong, electrostatic interactions to understand macromolecular structures inside the cell. Based on recent microscopy and NMR studies, it has been noticed that positively charged proteins are substantially slowed in intracellular environment, suggesting an overlooked role of nonspecific electrostatic interactions based on the predominantly negatively charged intracellular environment. Utilizing single-molecule detection and statistics, here we quantify in aqueous solutions how protein diffusion, in the limit of low diffuser concentration to avoid aggregate/coacervate formation, is modulated by differently charged interactor proteins over wide concentration ranges. We thus report substantially suppressed diffusion when oppositely charged interactors are added at ppm levels, yet unvaried diffusivities when same-charge interactors are added beyond 1%. The electrostatic attraction-driven suppression of diffusion is sensitive to the protein net charge states, as probed by varying the solution pH and ionic strength or chemically modifying the proteins and is robust across different diffuser-interactor pairs. By converting the measured diffusivities to diffuser diameters, we further show that in the limit of excess interactors, a positively charged diffuser molecule effectively drags along just one monolayer of negatively charged interactors, where further interactions stop. We thus unveil ubiquitous, net charge-driven protein-protein interactions, and shed new light on the mechanism of charge-based diffusion suppression in the living cell. This work is reproduced with permission from the following reference.¹³

In Chapter 5, I will uncover how molecular size and charge influence D properties of proteins in model cytoplasm using *Xenopus laevis* egg extracts. While we have shown both *in vivo* and *in vitro* charge asymmetric D slowdown due to strong, oppositely charged electrostatic interactions, they are presently limited by being either difficult to control regarding both size and charge of diffusers (*in vivo*) or are too artificial (*in vitro*). *Xenopus laevis* extract bridges both limitations by being a homogenized, undiluted, plasma membrane-less cytoplasm, enabling controlled and facile study of macromolecular size and charge role in shaping D inside of the cell. Leveraging high-throughput single-molecule statistics, we quantify D of protein over a wide range of charges and sizes with $\sim 1\%$ precision. We thus report that negatively charged proteins diffused $\sim 45\%$ of their diffusivity in buffer, showing D scaling based on Young-Carroad-Bell modeling relating D to the molecular weight of the protein, and consistent with the $\sim 2\text{x}$ -fold higher viscosity in extract. Notably, D did not show macromolecular crowding-based size-dependence, likely the result of extracts inhibiting actin filament formation using cytochalasins. Additionally, we find that positively charged protein ($>+5$) diffused at 10-20% of their PBS diffusivities, recapitulating *in vivo* diffusivity behaviors where positively charged proteins were more significantly slowed when compared to negatively charged compartments. Next, by treating extract with RNase to deplete cytoplasmic RNA, we show that the extract aggregates akin to adding excessive positively charged macromolecules, thus uncovering that the release of the predominately positively charged ribosomes drive aggregation, which only moderately increases D of positively charged diffusers

~25%, suggesting that the predominately negatively charged protein environment still suppresses D for positively charged diffusers. Lastly, by using that do not contain cytochalasins, dense intact, filamentous actin networks form akin to cellular actin. We thus show in that diffusion in these dense actin networks that size-dependent D or macromolecular crowding is induced demonstrating that nanoscale intracellular architectures shape D properties based on molecular size. In summary, we present a quantitative understanding of how macromolecular size and charge define D inside of the cell and shed new insight into intracellular mechanisms that regulate these intrinsic properties.

In Chapter 6, I will summarize the work conducted and presented in this thesis. I will then discuss opportunities in single-molecule diffusivity mapping to answer a variety of questions ranging from cell biology to development of new functional super-resolution techniques.

Publications reproduced in this dissertation are listed below. Permission has been granted by all critical co-authors for the reproduction of the work in the dissertation.

*A.A. Choi, H.H. Park, K. Chen, R. Yan, W. Li, K. Xu, "Displacement statistics of unhindered single molecules show no enhanced diffusion in enzymatic reactions," J. Am. Chem. Soc., **144**, 4839-4844, 2022.*

*A.A. Choi, L. Xiang, W. Li, K. Xu, "Single-molecule displacement mapping indicates unhindered intracellular diffusion of small ($\lesssim 1$ kDa) solutes," J. Am. Chem. Soc., **145**, 8510-8516, 2023.*

*A.A. Choi, K. Xu, "Single-molecule diffusivity quantification unveils ubiquitous net charge-driven protein-protein interaction" J. Am. Chem. Soc., **146**, 10973-10978, 2024.*

*A.A. Choi, C.Y. Zhou, A. Tabo, R. Heald, K. Xu, "Single-molecule diffusivity quantification in *Xenopus laevis* egg extract elucidates cytoplasmic molecular environments."*

Chapter 2: Displacement statistics of unhindered single molecules show no enhanced diffusion in enzymatic reactions

The work is reproduced with permission from the following publication:

A.A. Choi, H.H. Park, K. Chen, R. Yan, W. Li, K. Xu, “Displacement statistics of unhindered single molecules show no enhanced diffusion in enzymatic reactions,” *J. Am. Chem. Soc.*, **144**, 4839-4844, 2022. Copyright © 2022 American Chemical Society.

2.1. Introduction

The recent reporting of accelerated diffusion of enzymes in catalytic reactions has generated keen research interest and debate.¹⁴⁻²¹ Measurements based on fluorescence correlation spectroscopy (FCS) have documented substantial ($\sim 30\%$) increases in the diffusion coefficient D for multiple enzymes under catalytic turnovers.^{14, 15, 22-29} Several recent studies, however, emphasize FCS artifacts^{16, 30, 31} and report contradicting results. Dynamic light scattering (DLS) experiments, with low sensitivity for the very small sizes of proteins, also yield inconsistent results either supporting^{27, 29} or refuting³² substrate-enhanced diffusion. Likewise, different nuclear magnetic resonance (NMR) experiments have separately reported the presence or absence of reaction-enhanced D for enzymes and small molecules,³³⁻³⁶ with unsettled debate on data interpretation.^{21, 37, 38}

Single-molecule experiments³⁹⁻⁴¹ provide potential means to achieve precise D measurements by probing one molecule at a time. However, with expected $D \sim 30\text{-}50 \mu\text{m}^2/\text{s}$,^{19, 42} the fast motions of typical enzymes make it difficult to detect freely diffusing single molecules. Using reducing-oxidizing chemicals that suppress fluorophore blinking, a recent study achieved long-time anti-Brownian electrokinetic (ABEL) single-molecule trapping of dye-labeled alkaline phosphatase (ALP) to show no change in D with the addition of its substrate.³¹ Meanwhile, single-molecule tracking has been reported for dye-labeled urease and ALP that were substantially slowed down by using viscous reagents as methylcellulose^{43, 44} and glycerol^{31, 44} or by tethering to supported lipid bilayers,⁴⁴ which yielded contrasting results of no change versus up to 3-fold increases in D upon substrate addition. Thus, it remains difficult to quantify D through single-molecule measurements for molecules freely diffusing in regular buffers, whereas the addition of blinking-suppressing chemicals and viscous reagents complicates data interpretation.

We recently developed single-molecule displacement/diffusivity mapping (SMdM),^{2, 45} a single-molecule imaging-based super-resolution method^{9, 46, 47} capable of determining molecular diffusivity. In SMdM, stroboscopic excitation pulses of $\tau < 1$ ms duration reduce motion-blur to capture snapshots of fast-moving molecules in the wide field. The excitation pulses are applied as pairs across tandem camera frames with a fixed $\Delta t \sim 1$ ms interval (Figure 2.1a), so that images captured in the two frames (Figure 2.1b) can be compared to determine the transient displacements

d of the observed molecules within the Δt time window. Repeating this tandem excitation scheme then enables the accumulation of d over time for D -value determination.

Whereas we previously focused on the high spatial resolution of SMdM to study intracellular diffusion heterogeneities,^{2, 45} here we harness its capability for high single-molecule statistics to enable the precise quantification of D for enzymes freely diffusing in regular buffers. We thus formulate how $\sim\pm 1\%$ precisions may be achieved for D , and then utilize this capability to show no changes in diffusivity for four enzymes under wide concentration ranges of substrates.

2.2 Results and Discussion

To optimize SMdM for the $\sim 60\%$ higher D versus the $\sim 20\text{--}30\ \mu\text{m}^2/\text{s}$ values in our previous work on intracellular diffusion,² we accordingly reduced the excitation pulse duration τ and center-to-center separation Δt to 300 and 600 μs , respectively (Figure 2.1a and Figure 2.5). The use of Cy3B,⁴⁸ a bright synthetic dye comparable in size to those used in previous enzyme-diffusion studies, to label the enzymes enabled single-molecule detection with good signals (Figure 1b). By running the recording EM-CCD camera at 110 frames per second, $\sim 10^4$ pairs of tandem frames were executed in minutes, from which $>5 \times 10^4$ single-molecule displacements were typically accumulated (Figure 2.1c). Maximum likelihood estimation (MLE) with a predefined model based on normal diffusion plus background² (Methods) yielded good fits, from which D values were obtained with their 95% confidence intervals, *e.g.*, $D = 47.3 \pm 0.4\ \mu\text{m}^2/\text{s}$ for Cy3B-labeled catalase in phosphate-buffered saline (PBS) (Figure 2.1c). We also analyzed the single-molecule displacements in two dimensions,⁴⁵ and found the diffusion isotropic, with identical D values calculated along different directions (Figure 2.6). As we varied Δt between 400 μs and 1 ms, we further observed a linear dependence for the MLE-extracted mean squared displacements (Figure 2.6). Thus, SMdM characterized normal (Gaussian) diffusion in our experiments.

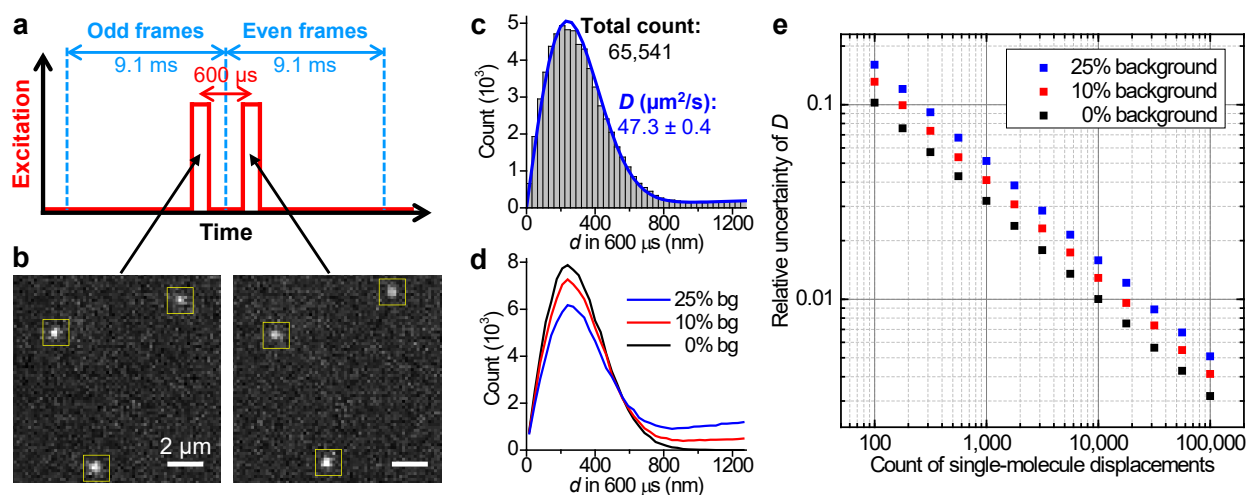


Figure 2.1. High-throughput statistics of unhindered single-molecule displacements. (a) Schematics: tandem excitation pulses of 300 μs duration are applied at $\Delta t = 600\ \mu\text{s}$ center-to-center separations across paired frames of the recording EM-CCD camera. (b) Example single-molecule images captured in the paired frames, for Cy3B-labeled

catalase freely diffusing in the PBS buffer, which enabled the determination of the transient displacements d of the 3 molecules marked by the yellow boxes in the $\Delta t = 600 \mu\text{s}$ time window. (c) Histogram: distribution of the 65,541 transient single-molecule displacements collected in 9 min by repeating the tandem excitation scheme 30,000 times. Blue line: MLE yielding a diffusion coefficient D of $47.3 \pm 0.4 \mu\text{m}^2/\text{s}$ (95% confidence interval). (d) Example distributions of 100,000 simulated displacements, of which 0%, 10%, and 25% are backgrounds from molecules randomly entered the vicinity while the rest are 600- μs displacements with $D = 50 \mu\text{m}^2/\text{s}$. (e) Relative uncertainty of D (σ_D/D) from MLE results of the simulated data at different levels of backgrounds, as a function of the counts of displacements. The standard deviations σ_D were calculated from the differences between 5,000 rounds of simulations under each condition.

To define how the achievable precision in D depends on the count of single-molecule displacements N , we simulated 600- μs displacements d for molecules diffusing at $D = 50 \mu\text{m}^2/\text{s}$. Different levels of backgrounds were further included, *i.e.*, assuming 0%, 10%, or 25% of the total counts were due to extraneous molecules that randomly entered the vicinity during Δt (Figure 2.1d). MLE of the simulated data showed that with zero background, the relative uncertainty of D (σ_D/D) was $\sim 1/\sqrt{N}$, so that 10% and 1% uncertainties were achieved with 100 and 10,000 counts of d , respectively (Figure 2.1e). This result is expected: each displacement provides an independent reporting of diffusivity, so the signal-to-noise ratio improves as $\sim \sqrt{N}$. At background levels of 10% and 25%, 1% uncertainty in D was achieved with $\sim 20,000$ and $\sim 30,000$ counts of d , respectively. As our experimental data contained $<10\%$ background (compare Figure 2.1c and Figure 2.1d), $\sim 0.5\%$ relative uncertainty (hence an $\sim \pm 1\%$ bracket for the 95% confidence interval) was thus obtained (Figure 2.1e) with the typical $>50,000$ counts of d in our experiments, in agreement with our above MLE results on the experimental data (Figure 2.1c). We further compared experimental results from different days and observed D values in the range of 46.3 – $48.1 \mu\text{m}^2/\text{s}$ for Cy3B-labeled catalase in PBS (Figure 2.6). The $\sim \pm 2\%$ variations between runs are larger than the above MLE-deduced $\sim \pm 1\%$ bracket from the single-molecule statistics alone and may be attributed to uncontrolled experimental factors. In particular, an $\sim 2\%$ increase in D is expected per 1 K temperature increase, mainly attributable to changes in water viscosity.¹⁹

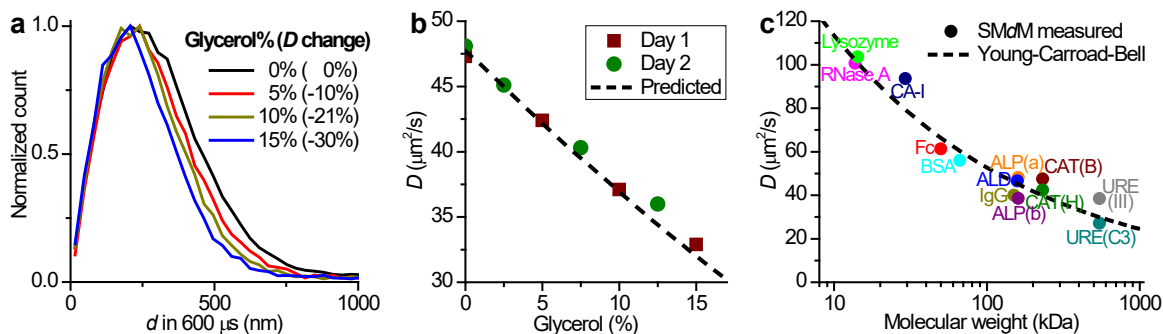


Figure 2.2. Precise determination of diffusivity *via* single-molecule statistics. (a) Normalized count distributions of the measured 600- μs single-molecule displacements for Cy3B-labeled catalase diffusing in PBS buffers containing 0%, 5%, 10%, and 15% glycerol. (b) The measured D as a function of the glycerol percentage, for two batches of samples measured a month apart (squares and circles; 95% confidence intervals are comparable to the symbol sizes), compared to that is predicted by scaling the 0% glycerol value with the known glycerol concentration-dependent

viscosity (dash line). (c) SMdM-measured D values of 13 different protein samples of varying sizes in PBS and Tris buffers. Dash line: expected values at 20 °C according to the Young-Carroad-Bell model. See protein sample details in Table S1.

To verify whether the above high precisions would enable the experimental detection of small differences in D , we examined the diffusion of Cy3B-labeled catalase in PBS buffers containing varying amounts of glycerol. The addition of 5%, 10%, and 15% weight percentages of glycerol led to substantial shifts in the measured d distribution (Figure 2.2a), with MLE yielding ~10%, ~21%, and ~30% decreases in D , respectively (Figure 2.2b). Two batches of samples measured a month apart yielded consistent D values, as well as good matches to the trend predicted by the glycerol concentration-dependent viscosity⁴⁹ (Figure 2.2b). Among these results, the 2.5% glycerol sample showed ~5% reduction in D and was well differentiated from the glycerol-free sample (Figure 2.2b). Whereas it is difficult to produce substantially higher D with the same enzyme, we employed SMdM to measure the diffusion coefficients of 13 different protein samples in a wide 14-545 kDa size range in PBS and Tris buffers (Figure 2.2c and Table 2.5). We thus obtained results in agreement with that is predicted by the Young-Carroad-Bell model,^{19, 42} including for the fast-diffusing small proteins as RNase A and lysozyme at $>100 \mu\text{m}^2/\text{s}$. Thus, SMdM is well poised for detecting D differences over a broad range.

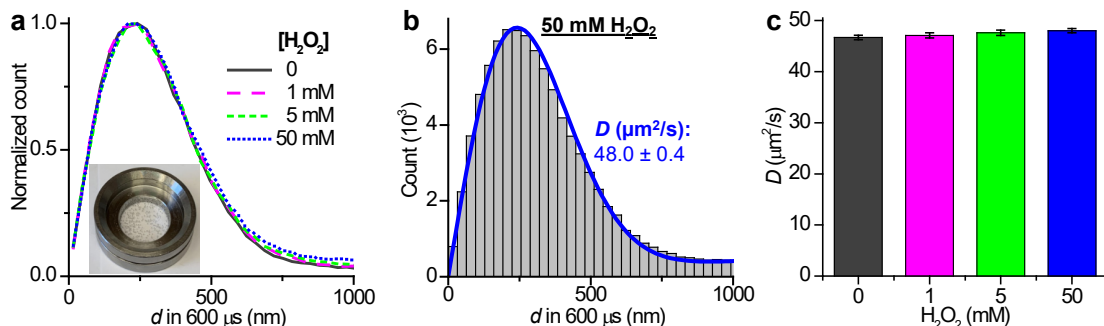


Figure 2.3. No substantial change in single-molecule displacements observed for catalase under turnover. (a) Normalized count distributions of the measured 600- μs single-molecule displacements for Cy3B-labeled catalase diffusing in PBS buffers with H_2O_2 added at 0, 1, 5, and 50 mM. Inset: photo of a sample with 50 mM H_2O_2 , showing the generation of bubbles from the catalytic reaction. (b) Histogram: the count distribution with the 50 mM H_2O_2 solution. Blue line: MLE yielding $D = 48.0 \pm 0.4 \mu\text{m}^2/\text{s}$ (95% confidence interval). (c) MLE-determined D at the different substrate concentrations. Error bars: 95% confidence intervals of MLE.

For catalytic reactions, we started with catalase in the presence of its substrate H_2O_2 , for which system FCS has reported ~30-40% increase in D with ~25 mM H_2O_2 .^{15, 22, 24} SMdM statistics showed that the addition of 1, 5, and 50 mM H_2O_2 did not noticeably shift the single-molecule displacement distribution (Figure 2.3a). The catalytic reaction generated oxygen bubbles, which were more notable for samples with higher H_2O_2 concentrations (Figure 2.3a inset). The bubbles scattered light and increased the fluorescence background to slightly reduce our capability to track molecules across the tandem frames, and so led to a moderately increased tail in the d distribution (Figure 2.3a). Nonetheless, MLE yielded good fits (Figure 2.3b), and the resultant D values

showed minimal variations in the range of 46.5-48.0 $\mu\text{m}^2/\text{s}$ between the different substrate concentrations (Figure 2.3c). We also conducted SMdM experiments at a longer pulse separation of $\Delta t = 1$ ms, as well as at increased single-molecule densities for higher backgrounds in the d distribution, and similarly observed unvaried D independent of the presence of the H_2O_2 substrate (Figure 2.7). These results contrast with the ~30-40% substrate-induced D enhancements reported in previous FCS studies.^{15, 22, 24}

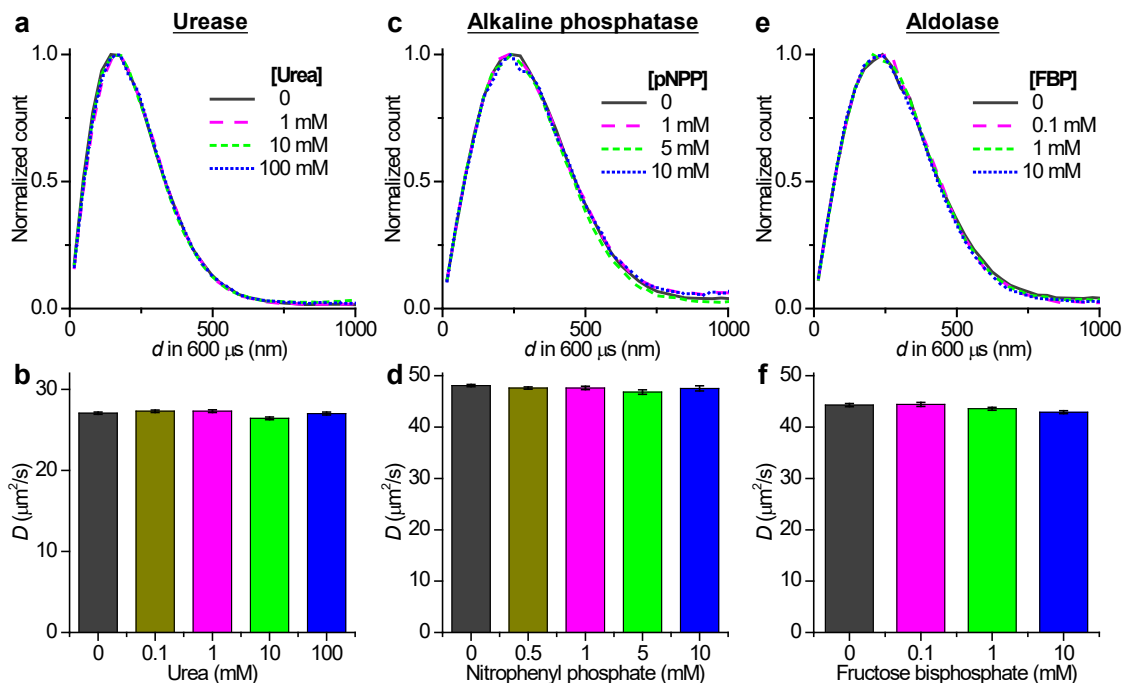


Figure 2.4. Results on urease, alkaline phosphatase, and aldolase also show no changes in single-molecule displacements under catalytic turnover. (a) Normalized distributions of the measured 600- μs single-molecule displacements for Cy3B-labeled urease in PBS, with and without the addition of urea at different concentrations. (b) MLE-determined D values at different urea concentrations. Error bars: 95% confidence intervals of MLE. (c,d) Similar to that in (a,b), but for Cy3B-labeled alkaline phosphatase in 0.1 M Tris-HCl (pH = 9.8) with the addition of 1 mM MgCl_2 and 20 μM ZnCl_2 , with p -nitrophenyl phosphate (pNPP) as the substrate. (e,f) Similar to that in (a,b), but for Cy3B-labeled aldolase in 0.1 M HEPES (pH = 7.4), with fructose 1,6-bisphosphate (FBP) as the substrate.

We next examined urease, for which previous FCS experiments have reported up to ~40% D enhancement in the presence of 10-100 mM of its substrate urea.^{14, 15, 25, 27-29} Moreover, ~3-fold substrate-induced D enhancement is recently reported for single urease molecules diffusing in the highly crowded or membrane-tethered systems at orders of magnitudes lower D values.^{43, 44} SMdM showed virtually identical distributions for the 600- μs single-molecule displacements of urease in PBS with and without the addition of 0.1-100 mM of urea (Figure 2.4a). MLE hence yielded comparable D values of 26.4-27.3 $\mu\text{m}^2/\text{s}$ for the five conditions with no noticeable trend on the substrate concentration (Figure 2.4b).

For ALP, FCS has initially reported up to an 80% increase in D with the addition of 2.6 mM of its substrate *p*-nitrophenyl phosphate (pNPP).¹⁵ However, recent FCS experiments,^{16, 29, 31} as well as single-molecule trapping in blinking-suppressed buffers and tracking in glycerol-slowed solutions³¹ indicate no substrate-induced D enhancement. For ALP in its working buffer of 100 mM Tris-HCl (pH = 9.8) containing 1 mM MgCl₂ and 20 μ M ZnCl₂, SMdM showed no substantial changes in single-molecule displacements with the addition of 0.5-10 mM pNPP (Figure 2.4c). Accordingly, the resultant D values of 46.8-48.1 μ m²/s indicated no enhanced diffusion (Figure 2.4d).

We further examined aldolase, for which previous FCS experiments have either reported ~30% increase²³ or no change^{27, 29} of D in the presence of ~1 mM of its substrate, fructose 1,6-bisphosphate (FBP), whereas DLS^{27, 32} and NMR³⁴ experiments have indicated no D enhancement over broad substrate ranges. SMdM of aldolase in a HEPES buffer (pH = 7.4) showed no substantial changes in single-molecule displacements with and without the addition of 0.1-10 mM of FBP (Figure 2.4e), and hence no enhanced diffusion of D values in the range of 42.9-44.4 μ m²/s (Figure 2.4f).

To verify that the enzymes were active throughout our experiments, we performed standard enzyme activity assays under identical conditions as in SMdM. We thus showed that the enzymes were active, comparable to that reported in previous studies (Figure 2.8). Using these results, we further calculated the expected substrate concentrations and reaction rates as a function of time (Figure 2.9). Changes in both parameters were small over 10 min. Indeed, as SMdM is based on single-molecule imaging, its ~200 pM working enzyme concentration is an order of magnitude lower than that is typically used in FCS,^{15, 29} and so substrate consumption during measurement is less concerning. Following recent discussions that D may increase linearly with the free energy released in the enzyme-catalyzed reaction,²⁹ we also calculated the expected Gibbs energy dissipation rates for the different systems (Figure 2.9). The resultant values for urease and aldolase were comparable to the study,²⁹ which would predict >10% reaction-induced diffusion enhancement for the former. For catalase, the calculated Gibbs energy dissipation rates were one order of magnitude higher than the highest values discussed in the study.²⁹ Our observation of no enhanced diffusion for both systems thus does not support the exergonic argument.

2.3 Conclusion

Together, with the high precision achieved through the high single-molecule statistics of SMdM, our results consistently showed unvarying D for four enzymes that have been reported as exhibiting enhanced diffusion under catalytic turnovers. These results, together with recent experiments that reexamine artifacts due to photophysical processes as fluorescence quenching and blinking,^{16, 30, 31} underscore the difficulties with quantifying D through FCS. Although blinking-suppressing buffers partly alleviate these issues while also enabling single-molecule trapping for D measurements,³¹ the addition of reducing and oxidizing chemicals complicates data interpretation. Meanwhile, the high D values of enzymes have largely limited single-molecule

tracking to systems in which viscous reagents as methylcellulose^{43, 44} and glycerol^{31, 44} are added to slow down molecular motions.

In contrast, with SMdM, here we were able to work with the diffusion of enzymes in regular buffers without the addition of extraneous components that may potentially affect the system: As SMdM captured and accumulated the transient (600 μ s) displacements of single molecules, it removed the needs to establish long-term fluorophore photostability or to impede diffusion to permit multi-frame tracking. Whereas in this study we have employed MLE to extract D values based on normal (Gaussian) diffusion, the highly invariant distributions of single-molecule displacements we observed under different substrate concentrations suggest no changes in diffusivity independent of the specific model. The sub-millisecond motion-detection time window of SMdM is comparable to that of FCS, which has been the primary method used in previous studies to advocate enhanced diffusion in the presence of substrates. For the four enzymes examined, this time window should accommodate multiple turnovers for catalase, urease, and ALP, but not aldolase (based on reaction rates in Figure 2.9). Our results thus support recent analysis that reaction-facilitated diffusion vanishes at the molecular scale,²¹ and call for the reexamination of factors that may have inadvertently affected previous measurements, including both photophysical processes^{16, 30, 31} and the dissociation of enzyme subunits.^{16, 27} The application of the tools developed in this work to other systems in which precise D values are important, potentially including the even faster diffusion of smaller molecules, awaits future efforts.

2.4 Supplemental Materials and Methods

Fluorescent labeling of enzymes. Sources of enzymes and other proteins examined are listed in Table S1. The proteins were reacted with Cy3B NHS (*N*-hydroxysuccinimidyl) ester (Cytiva, PA63101) in 0.1 M NaHCO₃ at a ~10:1 dye-to-protein ratio at room temperature for 4 h. Excess dye was removed using Amicon 3k MWCO Centrifugal Filters (Millipore, UFC500396). Absorbances at 280 and 560 nm, as measured by a NanoDrop 2000c spectrometer (ThermoFisher), showed that the labeled product had 0.5-2 dyes per protein with a protein concentration of 0.1-1.0 mg/mL. The labeled catalase, and urease were kept in Dulbecco's Phosphate-Buffered Saline (PBS) on ice. The labeled alkaline phosphatase (ALP) was kept in 100 mM Tris pH = 8.0, 2 mM MgCl₂, 50 μ M ZnCl₂ on ice. The labeled aldolase was kept in 100 mM HEPES (pH = 7.4) on ice. Single-molecule experiments and enzyme activity assays were performed on the same day of labeling.

Enzyme activity assays. Enzyme activity assays for both Cy3B-labeled and unlabeled enzymes were performed in standard 10-mm path quartz cuvettes using the cuvette mode of a NanoDrop 2000c spectrometer (ThermoFisher). Catalase activity assays were performed in PBS by monitoring the hydrogen peroxide (VWR, BDH7742-1) concentration *via* absorbance at 260 nm ($\epsilon_{260} = 13 \text{ M}^{-1}\text{cm}^{-1}$). Urease activity assays were performed in PBS using the Urease Activity Assay Kit (MAK120, Sigma-Aldrich), following the manufacturer's instructions. Alkaline phosphatase assays on the p-nitrophenyl phosphate (pNPP; Sigma-Aldrich; 4876) substrate were performed in

0.1 M Tris-HCl (pH = 9.8) with 1 mM MgCl₂ and 20 μ M ZnCl₂, and the reaction velocity was determined from the 405 nm absorbance of the product 4-nitrophenol ($\epsilon_{405\text{nm}} = 18,000 \text{ M}^{-1}\text{cm}^{-1}$). Aldolase activity assays on the fructose 1,6-bisphosphate (FBP; Sigma-Aldrich, F6803) substrate were performed in 0.1 M HEPES, pH 7.4, using the α -glycerophosphate dehydrogenase-triosephosphate isomerase (GDI-TPI; G1881, Sigma-Aldrich) assay with 0.1 mM NADH (N8129, Sigma-Aldrich). The reaction velocity was determined using the absorbance of NADH at 340 nm ($\epsilon_{340\text{nm}} = 6,220 \text{ M}^{-1}\text{cm}^{-1}$) with a stoichiometry of 2 NADH:1 FBP.

SMdM sample preparation. 25-mm diameter, #1.5 coverslips were treated in a heated 3:1 H₂SO₄ (98%) and H₂O₂ (30%) mixture for 45 minutes, rinsed until neutral using Milli-Q water, and then dried using N₂ gas. Coverslips were functionalized with 2 mM methoxy PEG silane (5 kDa, PG1-SL-5k, Nanocs) in 95% ethanol/water for 30 minutes in an ethanol vapor. Coverslips were rinsed and sonicated for 5 minutes in Milli-Q water before being mounted into an imaging chamber (ThermoFisher, A7816). Cy3B-labeled enzymes were diluted to $\sim 200 \text{ pM}$ into the same buffer as the above buffers for enzyme activity assays with or without the addition of substrates (same as above used in the enzyme activity assays) at different concentrations, and SMdM was immediately performed.

SMdM of enzymes. SMdM was performed on a Nikon Ti-E inverted fluorescence microscope, as described previously,² but with reduced excitation pulse duration τ and pulse-to-pulse separation Δt for the fast free diffusion in buffers. Briefly, a 561 nm laser (OBIS 561 LS, Coherent, 165 mW) was focused to the back focal plane of an oil-immersion objective lens (CFI Plan Apochromat Lambda 100X Oil, NA = 1.45) to enter the sample slightly below the critical angle of the glass-solution interface, so that it illuminated a few micrometers into the solution. The focal plane was set at $\sim 2 \text{ }\mu\text{m}$ away from the glass surface into the solution, which was maintained by the native Nikon Perfect Focus System (PFS) on the microscope during the experiment. A multifunction I/O board (PCI-6733, National Instruments) detected the exposure timing signal of the recording EM-CCD camera (iXon Ultra 897, Andor), and accordingly modulated the emitting power of the 561 nm excitation laser. Paired pulses of $\tau = 300 \text{ }\mu\text{s}$ duration and $\Delta t = 600 \text{ }\mu\text{s}$ center-to-center separation were thus repeatedly applied across tandem camera frames, with the timing waveforms confirmed by a GW Instek GDS-1054B oscilloscope (Figure 2.5). For Figure 2.6e-h and Figure 2.7, we further compared results obtained with $\Delta t = 400 \text{ }\mu\text{s}$, $800 \text{ }\mu\text{s}$, and 1 ms center-to-center separations, for which the pulse durations τ were accordingly varied as $\Delta t/2$. Typical runs executed $1.7\text{-}3.0 \times 10^4$ pairs of tandem frames at 110 frames per second (fps), thus data-collection times of 5-9 min. Fluorescence emission was filtered by a longpass (ET575lp, Chroma) and a bandpass filter (ET605/70m, Chroma).

Data analysis. SMdM data were analyzed as described previously.² Briefly, single molecules were first super-localized in each recorded frame (e.g., Figure 2.1b), and then their two-dimensional displacements d between the paired frames were calculated with a search radius R of $1.3 \text{ }\mu\text{m}$. The accumulated $>5 \times 10^4$ single-molecule displacements in each run were pooled together for plotting

of distribution (e.g., Figure 2.1c) and MLE with the below probability model based on two-dimensional random walk:²

$$P(r) = \frac{2r}{a} \exp\left(-\frac{r^2}{a}\right) + br \quad (\text{eqn. 1})$$

Here r is the magnitude of single-molecule displacement in the fixed time interval Δt , and $a = 4D\Delta t$ is the mean squared displacement (MSD) in two dimensions. b accounts for a uniform background due to extraneous molecules that randomly diffuse into the search radius during Δt , which we rationalized and validated previously with experiments performed at varied single-molecule densities.² In the current study, we kept the enzyme samples at low concentrations to minimize the b term. In Figure 2.7, we further compared results in which we intentionally increased the sample concentration to induce a higher background in the d distribution, and MLE still yielded consistent D values. MLE with MATLAB, in which a and b were independent free parameters, returned 95% confidence intervals for both parameters. The fitted models matched well with the experimental results, with slight differences attributable to small amounts of impurity and/or oligomerization forms in the sample, which likely distorted the single-molecule displacement distribution. The a results (MSDs) were directly converted to D values through $D = a/(4\Delta t)$ using the fixed $\Delta t = 600 \mu\text{s}$ (or $\Delta t = 1 \text{ ms}$ in Figure 2.7). In Figure 2.6e-h, the MLE-extracted a values (MSDs) from SMdM at different Δt time separations were further plotted as a function of Δt . In Figure 2.6a-d, we further examined the directional single-molecule displacements in two dimensions, as detailed in Ref⁴⁵. The principal direction of diffusion θ was evaluated together with an anisotropy α (0 and 1 for fully isotropic and anisotropic diffusions, respectively). The single-molecule displacements are then projected along and perpendicular to the calculated θ direction, and each fit to a modified one-dimensional random-walk model with the probability distribution:

$$P(x) = \frac{1}{\sqrt{a\pi}} \exp\left(-\frac{x^2}{a}\right) + b' \quad (\text{eqn. 2})$$

Here x is the projected displacement in one dimension, $a = 4D\Delta t$, and b' is the background.

Simulation. Monte Carlo simulation was performed with MATLAB, borrowing components from our recent simulation of more complicated systems.⁵⁰ Single-molecule displacements in $\Delta t = 600 \mu\text{s}$ were simulated for different total counts based on normal diffusion (two-dimensional random walk)⁵¹ with $D = 50 \mu\text{m}^2/\text{s}$, with the addition of Gaussian blurs of $\sigma = 20 \text{ nm}$ for both the starting and ending positions to simulate the uncertainties in localizing the single molecules. The $\sim 250 \text{ nm}$ single-molecule motion-blur during the $300 \mu\text{s}$ pulse window is smaller than the diffraction-limited single-molecule image size and is not further simulated in this work. We next allowed extraneous molecules to randomly enter the simulated area with a uniform spatial density to contaminate 10% or 25% of the total counts. The resultant simulated single-molecule displacement data (see

example distributions in Figure 2.1d) were fed as input for the MLE determination of D based on eqn. 1 as described above. The simulation under each condition was repeated 5,000 times, so that the differences in the extracted D values between the 5,000 runs were used to calculate the standard deviation σ_D . The relative uncertainty σ_D/D was then plotted as a function of the count of simulated displacements (Figure 2.1e).

Supplementary Table

Table 2.1. SMdM-determined *D* values for different protein samples in PBS (except for ALPs in Tris buffers, as noted)

Label in Figure 2.2c	MW (kDa)	Measured <i>D</i> ($\mu\text{m}^2/\text{s}$)	Protein	Provider	Cat #
<u>Used in Figures 2.3 & 2.4</u>					
ALD	158	46.5*	Aldolase from rabbit muscle	Sigma	A8811
ALP(a)	160	48.1 [#]	Alkaline phosphatase from bovine intestinal mucosa	Sigma	P7923
CAT(B)	232	47.5	Catalase from bovine liver	Sigma	C30
URE(C3)	545	27.1	Urease from jack bean, Type C-3	Sigma	U0251
<u>Not in Figures 2.3 & 2.4</u>					
RNase A	13.7	100.6	Ribonuclease A from bovine pancreas	Sigma	R5500
Lysozyme	14.3	103.5	Lysozyme from chicken egg white	Sigma	L6876
CA-I	29.2	93.6	Carbonic anhydrase I from human erythrocytes	Sigma	C6653
Fc	50	61.3	Fc fragment of human immunoglobulin G	Athens	16-16-090707-FC-IMM
BSA	66.5	56.0	Bovine serum albumin	Sigma	A3059
IgG	150	40.0	Mouse immunoglobulin G (monoclonal anti-IRE1a)	Santa Cruz	sc-100772
ALP(b)	160	38.5 [^]	Alkaline phosphatase from bovine intestinal mucosa	Sigma	P7640
CAT(H)	232	42.4	Catalase from human erythrocytes	Sigma	C3556
URE(III)	545	38.5	Urease from jack bean, Type III	Sigma	U1500

* The aldolase values in Figure 2.4f are slightly lower than the value here separately measured in the PBS, attributable to the ~5% higher viscosity of the 0.1 M HEPES buffer.⁵²

Measured in a 0.1 M Tris-HCl (pH = 9.8) buffer containing 1 mM MgCl₂ and 20 μM ZnCl₂.

^ Measured in a 5 mM Tris-HCl (pH = 8.0) buffer containing 2 mM MgCl₂ and 50 μM ZnCl₂.

Supplementary Figures

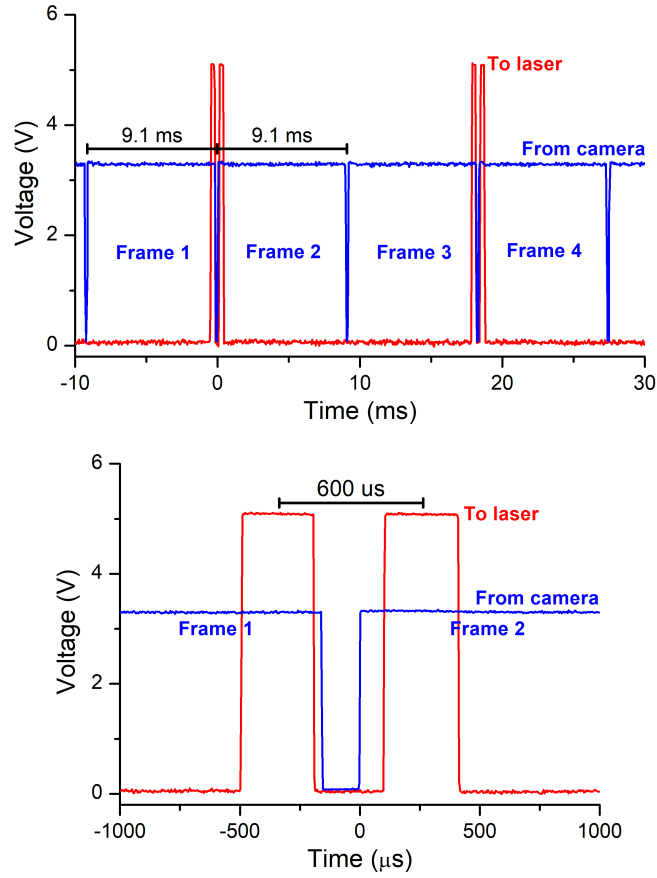


Figure 2.5. Typical timing waveforms in this work. Left panel: zoom-out for four consecutive camera frames. Right panel: zoom-in for the tandem pulses across two frames. Blue traces: the “fire” TTL signal from the EM-CCD camera (Andor iXon Ultra 897), which ran at 110 fps (9.1 ms/frame). 3.3 V corresponds to when exposure occurred, whereas 0 V corresponds to the dead time between frames. Red traces: output from the PCI-6733 card, which drives the 561 nm excitation laser between the off (0 V) and on (5 V) states *via* digital modulation. A pair of pulses each of $\tau = 300 \mu$ s duration are applied across the two frames at a $\Delta t = 600 \mu$ s center-to-center separation, avoiding the dead time of the camera.

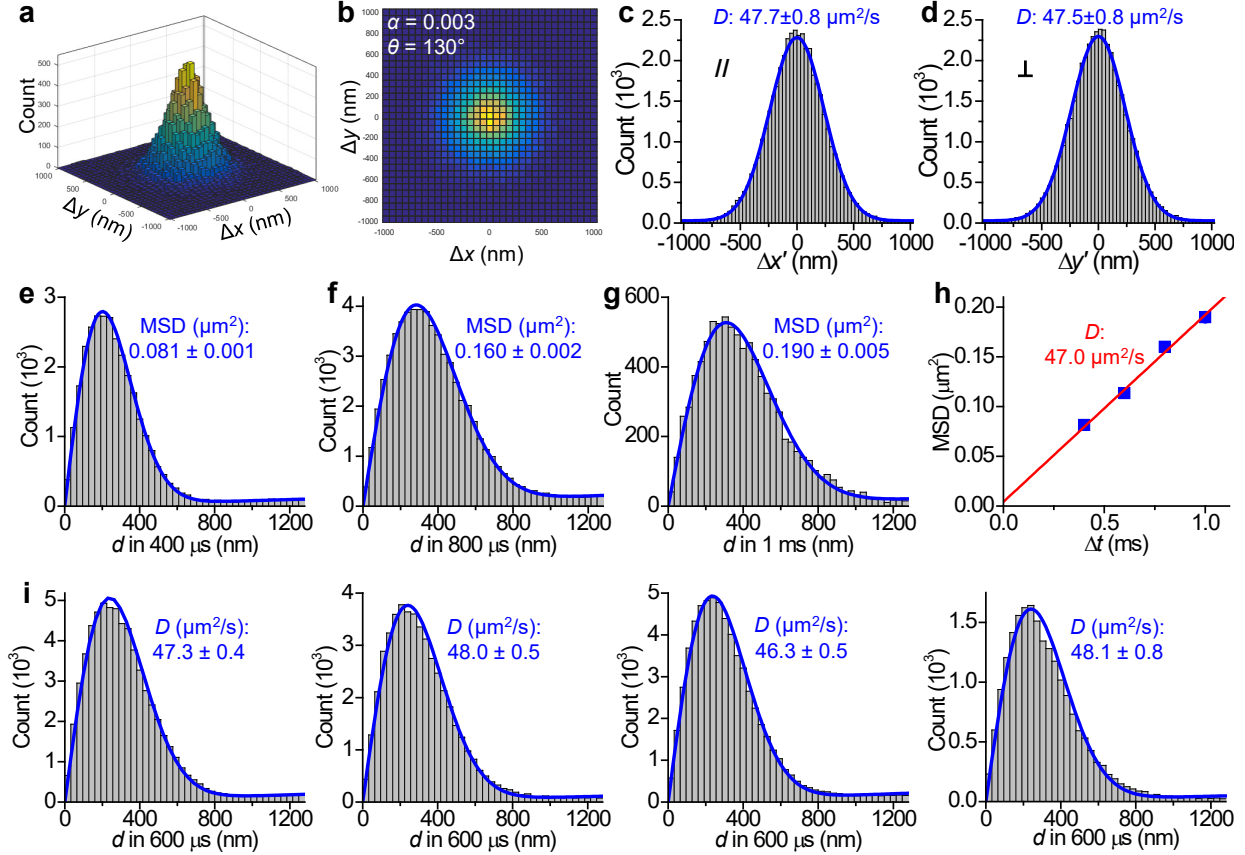


Figure 2.6. Additional characterizations of single-molecule displacements from SMdM, for Cy3B-labeled catalase diffusing in PBS. (a,b) Histogram (a) and heatmap (b) of two-dimensional single-molecule displacements in $\Delta t = 600 \mu\text{s}$. A near-zero anisotropy α of 0.003 is calculated with a principal direction θ of 130° . (c,d) One-dimensional distributions of the displacements in (a,b) projected along (c) and perpendicular (d) to θ . Blue curves: MLE with our one-dimensional diffusion model (eqn. 2), yielding $D = 47.7$ and $47.5 \mu\text{m}^2/\text{s}$. (e-g) Distributions of the magnitude of single-molecule displacements for SMdM experiments performed at $\Delta t = 400 \mu\text{s}$ (e), $800 \mu\text{s}$ (f), and 1 ms (g). Blue curves: MLE with our two-dimensional diffusion model (eqn. 1), with resultant mean squared displacements (MSDs; a in eqn. 1) and 95% confidence intervals labeled in each graph. (h) Blue squares: the MLE-extracted MSDs for the different Δt experiments. Error bars are smaller than the symbol sizes. Red line: linear fit, yielding slope $4D = 187.9 \mu\text{m}^2/\text{s}$; thus $D = 47.0 \mu\text{m}^2/\text{s}$. (i) Comparison of the $\Delta t = 600 \mu\text{s}$ results obtained on different days. Blue lines: MLE fits with resultant D values and 95% confidence intervals labeled in each graph.

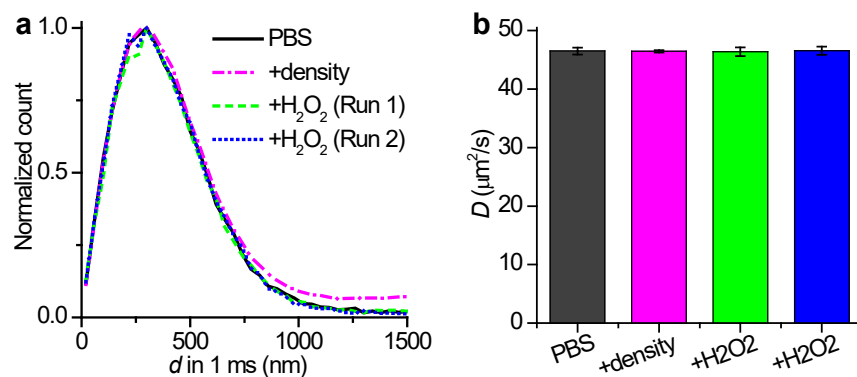


Figure 2.7. SMdM results for Cy3B-labeled catalase at a longer pulse separation of $\Delta t = 1$ ms and an increased single-molecule density. (a) Normalized count distributions of the measured 1-ms single-molecule displacements, for a typical control sample (black), a sample of substantially higher catalase concentration (and thus single-molecule density) to generate a higher background in the d distribution (magenta), and two samples in which 50 mM H_2O_2 were added (green and blue). (b) MLE-determined D for the four samples in (a). Error bars: 95% confidence intervals of MLE. Consistent D values are obtained for the regular- and high-density data despite the increased tail distribution of the latter. This result is expected, as in eqn. 1, the a parameter is sensitive to the peak position of the distribution whereas b fits to a background that increases linearly with d . Meanwhile, the addition of 50 mM H_2O_2 did not change the diffusion behavior of catalase here detected under the $\Delta t = 1$ ms condition.

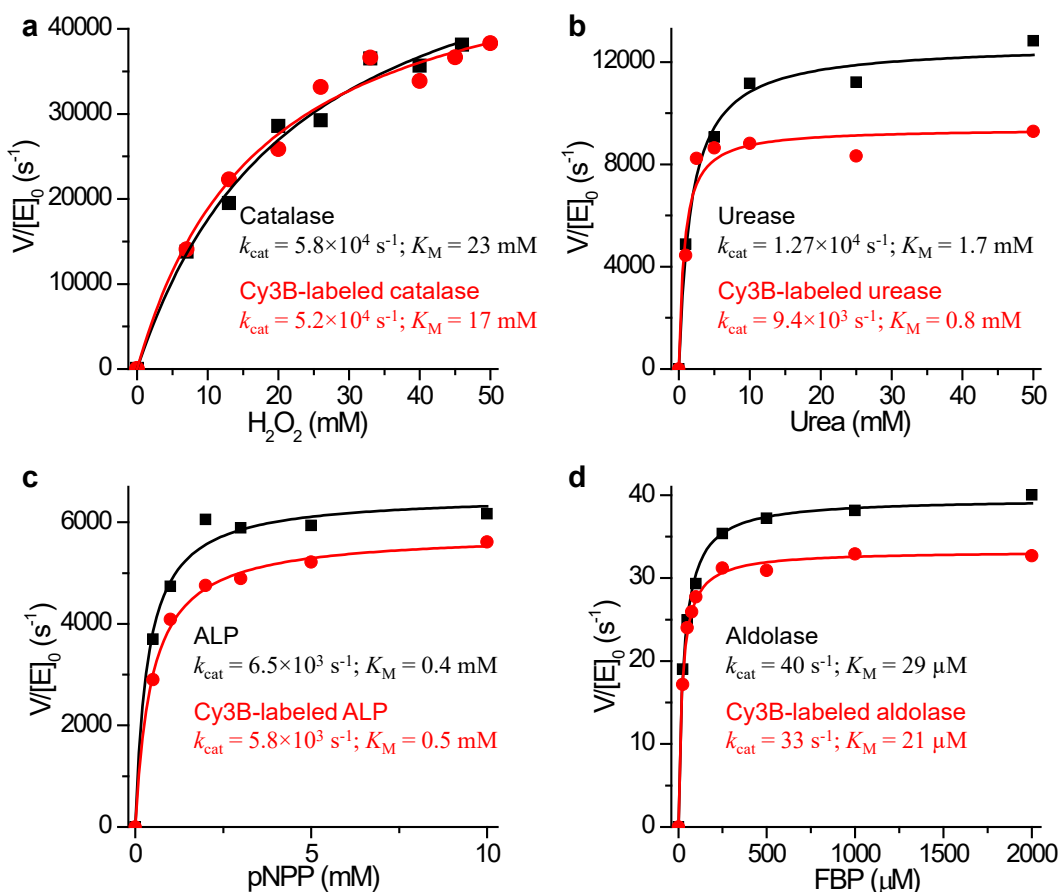


Figure 2.8. Activity assays for Cy3B-labeled and unlabeled enzymes. Data points: Reaction rate per enzyme $V/[E]_0$ at varied substrate concentrations $[S]$, where V is the measured reaction rate and $[E]_0$ is the concentration of enzyme used in the assay. Curves: fits to the Michaelis-Menten kinetics $V/[E]_0 = k_{\text{cat}} \frac{[S]}{K_M + [S]}$. The resultant turnover number k_{cat} and Michaelis constant K_M are

labeled in each graph. (a) Catalase activity assays were performed for ~ 1 nM enzyme in PBS. Enzyme activities are comparable to previous work.^{15, 22, 53} (b) Urease activity assays were performed for ~ 1 nM enzyme in PBS, yielding results comparable to previous work.^{15, 29} (c) ALP activity assays were performed for ~ 50 pM enzyme in 0.1 M Tris-HCl (pH = 9.8) with the addition of 1 mM MgCl_2 and 20 μM ZnCl_2 . Results are comparable to previous work.³¹ (d) Aldolase activity assays were performed for ~ 10 nM enzyme in a 0.1 M HEPES (pH = 7.4) buffer. Results are comparable to previous work.^{32, 34}

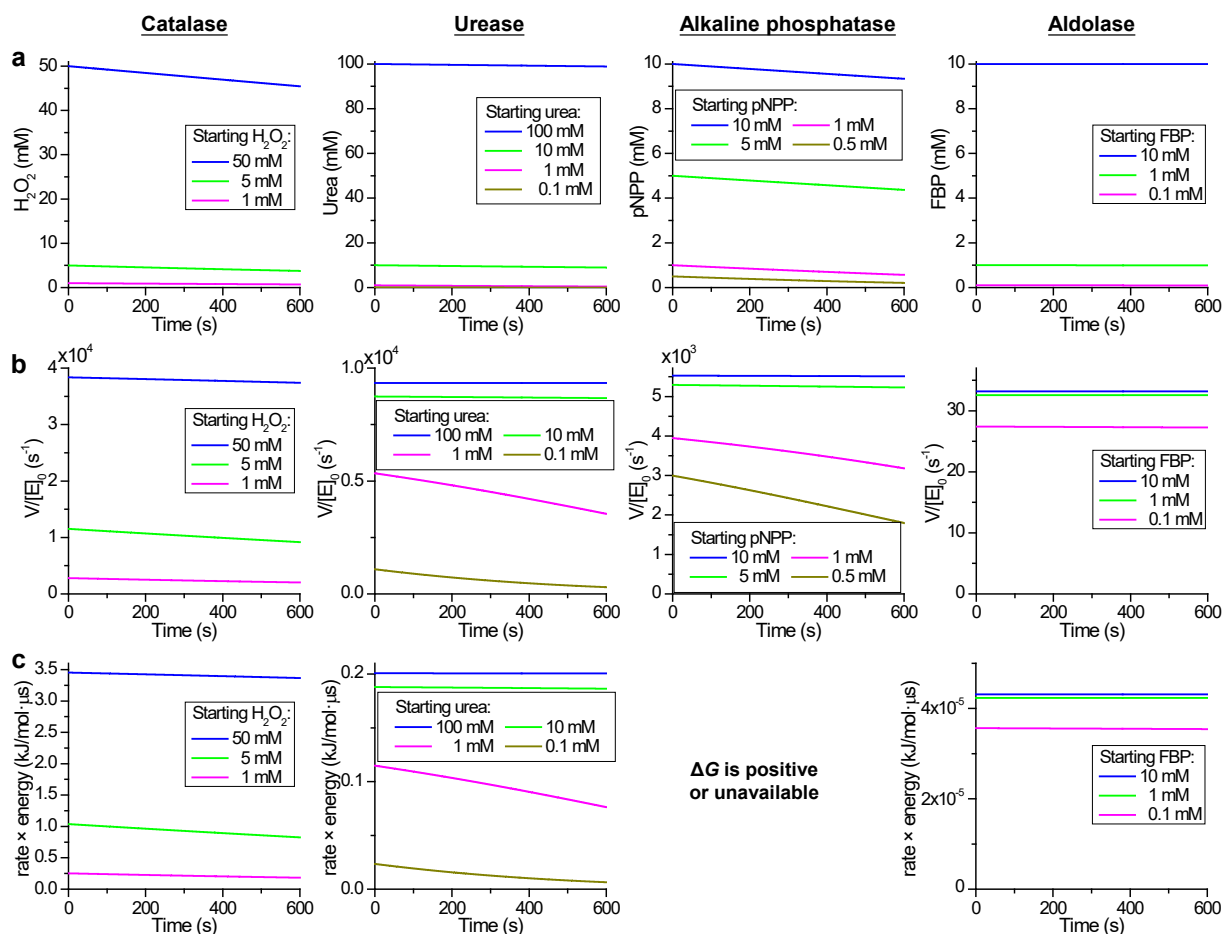


Figure 2.9. Simulated substrate concentrations, reaction rates, and Gibbs energy dissipation rates for the different enzymes in our SMdM experiments. (a, b) Simulated substrate concentrations (a) and reaction rates per enzyme (b) as a function of time, for the different starting substrate concentrations in Figures 2.3 and 2.4. Simulation was based on Michaelis-Menten kinetics for 200 pM enzyme using the k_{cat} and K_{M} values determined for the Cy3B-labeled enzymes in Figure 2.8. (c) Simulated Gibbs energy dissipation rates, as calculated by multiplying the reaction rates in (b) with the free energy release in the reaction, $-\Delta G$, as discussed in ref²⁹. The ΔG values of the reactions of urease and aldolase are from Table 1 of ref²⁹. The resultant Gibbs energy dissipation rates are comparable to that in ref²⁹; the urease values of $\sim 0.2 \text{ kJ/mol} \cdot \mu\text{s}$ would thus predict $>10\%$ enhancement in D with the master curve in ref²⁹. For catalase, which is not included in ref²⁹, ΔG of the reaction is from eQuilibrator.⁵⁴ The resultant Gibbs energy dissipation rates of $\sim 3.5 \text{ kJ/mol} \cdot \mu\text{s}$ are one order of magnitude higher than the highest values discussed in ref²⁹ for other enzymes. For the ALP reaction, ref²⁹ lists a positive ΔG value whereas eQuilibrator⁵⁴ indicates that ΔG cannot be estimated “because the uncertainty is too high”.

Chapter 3: Single-molecule displacement mapping indicates unhindered intracellular diffusion of small (~ 1 kDa) solutes

The work in this chapter is reproduced with permission from the following publication:

A.A. Choi, L. Xiang, W. Li, K. Xu, “Single-molecule displacement mapping indicates unhindered intracellular diffusion of small ($\lesssim 1$ kDa) solutes,” *J. Am. Chem. Soc.*, **145**, 8510-8516, 2023. Copyright © 2023 American Chemical Society.

3.1. Introduction

Molecular diffusion is a fundamental process of the cell that underlies key functions.⁵⁵⁻⁶² Whereas the translational diffusion of macromolecules as proteins is routinely probed in the mammalian cell with modern live-cell tagging and diffusivity measurement approaches, the intracellular diffusion of unbound small (~ 1 kDa) solutes has been more difficult to elucidate.

On the labeling side, as small fluorescence probes cannot be expressed in the cell [typical fluorescence proteins (FPs) are ~ 30 kDa], they often need to be delivered exogenously. While lipophilic fluorescent dyes pass the plasma membrane, their strong lipid interactions make them mainly report on the slow (diffusion coefficient $D < \sim 3 \mu\text{m}^2/\text{s}$) diffusion of intracellular membranes.^{63, 64} Conversely, hydrophilic dyes are typically membrane impermeable. Capping the polar functional groups, *e.g.*, with acetoxymethyl ester or acetate ester, enables intracellular delivery with subsequent de-capping, yet dyes delivered this way may be sequestered in vesicles.⁶⁵ More generally, independent of the loading method, a fraction of the delivered probes may be bound to different intracellular structures,⁶⁶⁻⁶⁸ leading to reduced apparent diffusivity.

On the measurement side, the very fast diffusion ($D > \sim 200 \mu\text{m}^2/\text{s}$; see below) of unbound small molecules is difficult to quantify: the range is orders of magnitude higher than what is workable with typical single-molecule tracking,^{41, 69} while also causing large uncertainties for other common microscopy-based approaches as fluorescence recovery after photobleaching (FRAP) and fluorescence correlation spectroscopy (FCS).^{55, 59, 61, 70} It is even more difficult to map out spatial heterogeneities in intracellular diffusion, and the global readout may be biased by the locally bound molecules. Separately, whereas in-cell NMR provides alternative means to probe molecular diffusion,⁷¹ ensemble results are collected from many cells; cell confinement effects and signal from the extracellular space, together with the limited signal-to-noise ratio, complicate data interpretation.

With the above challenges, our current model of how small solutes diffuse in the cell^{56, 60} has often focused on an early FRAP study⁷² in which the acetoxymethyl ester-loaded 520 Da dye

BCECF is found to diffuse four-time slower in the mammalian cytoplasm versus *in vitro* ($D_{\text{cyto}}/D_{\text{water}} \sim 0.25$). While this large reduction in D sets an alarming speed limit for the intracellular diffusion of small molecules, it contrasts with the modest (~ 20 - 40%) drop in intracellular rotational diffusivity of the same dye, which suggests only a slight increase in the cytoplasm viscosity over water.⁷³ Although this paradox may be reconciled by considering hindered translational diffusion *via* collision with cell solids,^{56, 72} similar measurements on intracellularly expressed GFP show slightly less suppressed D ($D_{\text{cyto}}/D_{\text{water}} \sim 0.31$),⁷⁴ even as stronger macromolecular crowding-induced hindrances may be expected for the much larger GFP. Later experiments have reported scattered intracellular diffusivities for different small solutes,⁷⁵⁻⁸⁰ and with the above-noted limitations in cell loading and measurement, have not resolved the $D_{\text{cyto}}/D_{\text{water}} \sim 0.25$ speed-limit paradox.

Here we overcome previous challenges to quantify and map out the intracellular diffusivity of small solutes by integrating several recent advances. For intracellular delivery, we exploit the discovery that astrocytes actively uptake the hydrophilic dye sulforhodamine 101 (SR101),^{81, 82} and then move to other small solutes *via* our recently devised graphene-based *in situ* electroporation.⁸³ For D measurement, we harness our recent development of single-molecule displacement/diffusivity mapping (SMdM),² which has previously enabled the super-resolution D mapping of ~ 30 kDa FPs in live cells at $D \sim 20 \mu\text{m}^2/\text{s}$. Extending SMdM to unbound small molecules, we here quantify D values of $>300 \mu\text{m}^2/\text{s}$ *in vitro* and $>200 \mu\text{m}^2/\text{s}$ in mammalian cells. We thus show that for multiple water-soluble dyes and dye-tagged nucleotides, the intracellular diffusion is dominated by large regions of unhindered diffusion with $D \sim 60$ - 70% of the diffusion *in vitro*.

3.2 Results and Discussion

In SMdM, wide-field images are recorded for diffusing single molecules with a typical EM-CCD camera. Stroboscopic illumination is employed to mitigate single-molecule motion-blur,⁸⁴ and the stroboscopic pulses are applied in tandem across paired camera frames, so that single-molecule displacements are recorded for time separations substantially shorter than the camera frame time (Figure 3.1a and Figure 3.5).² The paired excitation pulses across tandem frames further leave ample time between the anti-paired pulses for molecules outside the illuminated area to diffuse in, thus creating a mechanism to dynamically exchange and replenish molecules during imaging and permitting the use of conventional, non-photoswitchable fluorophores in SMdM.^{2, 9}

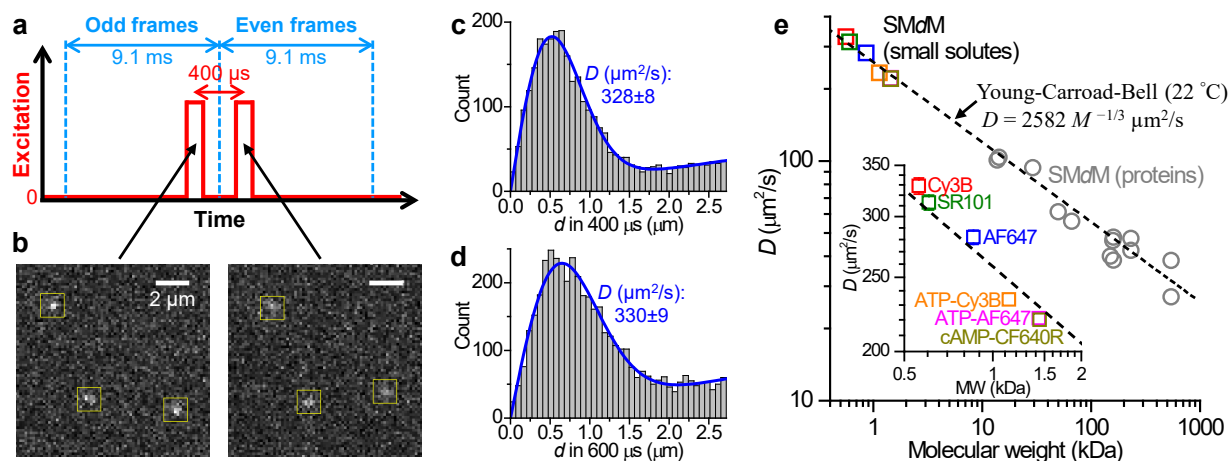


Figure 3.1. SMdM quantifies the fast diffusion of small solutes *in vitro*. (a) Schematics: To detect the fast motion of small solutes, tandem stroboscopic excitation pulses of 200 μs duration are applied across paired camera frames at a $\Delta t = 400 \mu\text{s}$ center-to-center separation. (b) Example single-molecule images captured in a tandem frame pair, for Cy3B freely diffusing in the PBS buffer at room temperature. (c,d) Typical distributions of the recorded single-molecule displacements for two experiments carried out at $\Delta t = 400 \mu\text{s}$ (c) and $600 \mu\text{s}$ (d), respectively, for Cy3B in PBS. Blue lines: maximum likelihood estimation (MLE) fits to our model, yielding diffusion coefficients D of 328 ± 8 and $330 \pm 9 \mu\text{m}^2/\text{s}$, respectively. (e) Summary of the SMdM-measured D values of 6 different small solutes in PBS (color squares), as well as 13 protein samples we previously reported (gray circles),¹² plotted as a function of molecular weight on a logarithmic scale. Dash line: the Young-Carroad-Bell model at 22 °C: $D = 2582 M^{-1/3} \mu\text{m}^2/\text{s}$, M being the molecular weight in Da. Inset: zoom-in of the small-molecule data. Error bars are shown for MLE uncertainties when larger than the symbol size.

We previously worked with a typical center-to-center pulse separation of $\Delta t = 1 \text{ ms}$ when probing the intracellular diffusion of FPs.² Here for the fast-diffusing small solutes, we dropped Δt down to $400 \mu\text{s}$ with the duration of each pulse τ accordingly reduced to $200 \mu\text{s}$ (Figure 3.5), and we focused on bright fluorescent dyes.^{5, 85} Good single-molecule images were obtained, as shown in Figure 3.1b for the free diffusion of the smallest (and thus fastest) probe examined in this work, Cy3B (560 Da), in the phosphate-buffered saline (PBS). Accumulating the resultant 400- μs displacements calculated from the super-localized positions from many paired frames yielded distributions (Figure 3.1c) well fitted by our model based on normal diffusion plus a background term for mismatched molecules (Methods).² Relaxing Δt to $600 \mu\text{s}$ led to a broader distribution of single-molecule displacements and a slight increase in the background (Figure 3.1d). Nonetheless, similar D values of $\sim 330 \mu\text{m}^2/\text{s}$ were obtained for both experiments, comparable to that previously measured for fluorescent probes of similar sizes.^{86, 87} Note that as PBS and other common buffers have viscosities near-identical to water, we collectively discuss diffusivity in them as typical “*in vitro*” values. To assess whether SMdM could correctly quantify impeded diffusion, we further examined Cy3B diffusion in PBS containing varying amounts of glycerol. We thus found that as glycerol addition increased the solution viscosity, the SMdM-measured diffusivity decreased accordingly following the expected trend over a wide range (Figure 3.6), down to $\sim 20\%$ of its initial value at 50% glycerol.

Figure 3.1e and inset summarize the SMdM-determined D in PBS for the 6 different small solutes we examined in this work. A notable drop in D was observed for the larger molecules. Moreover, as we plotted the D values against the molecular weight (M) on a logarithmic scale and further added in our previous *in vitro* SMdM results of 13 protein samples,¹² they all agreed reasonably well with the Young-Carroad-Bell Model⁴² without any fitting or scaling (Figure 3.1e). This agreement was not expected *a priori* (as the Model was summarized from >3 kDa proteins), yet reasonable in hindsight: few-heteroatom organic compounds may have densities comparable to that of protein. Their molar volumes, and hence D , may thus follow similar trends versus M as proteins.

After validating SMdM's applicability to the fast diffusion of small solutes, we started cell experiments with the special case of the 607 Da, yellow-excited dye SR101 (Figure 3.2b). While being water-soluble and generally cell-impermeable, SR101 is actively uptaken by astrocytes.^{81, 82} Thus, by incubating cultured primary rat hippocampal astrocytes with an SR101-containing medium, the dye was readily loaded into the cells. SMdM registered good single-molecule images (Figure 3.2a). As the camera recorded at ~ 110 frames per second, time separations between the anti-paired pulses were >35 times longer than between the paired pulses (Figure 3.5). Thus, whereas single-molecule displacements were tracked between the paired frames (e.g., Frames 1-2 and Frames 3-4 in Figure 3.2a), between the anti-paired pulses, molecules in the view were exchanged via diffusion (e.g., Frames 2-3 in Figure 3.2a). As we illuminated and imaged an ~ 1 μm layer at the bottom of the cell (Methods), diffusional exchanges enabled sustained single-molecule imaging with only a modest drop in the molecular count over $>7 \times 10^4$ frames (Figure 3.2c). Spatially binning the accumulated single-molecule displacements onto 320 nm-sized grids enabled local fitting^{2, 9} to extract D values and construct D maps (Figure 3.2d).

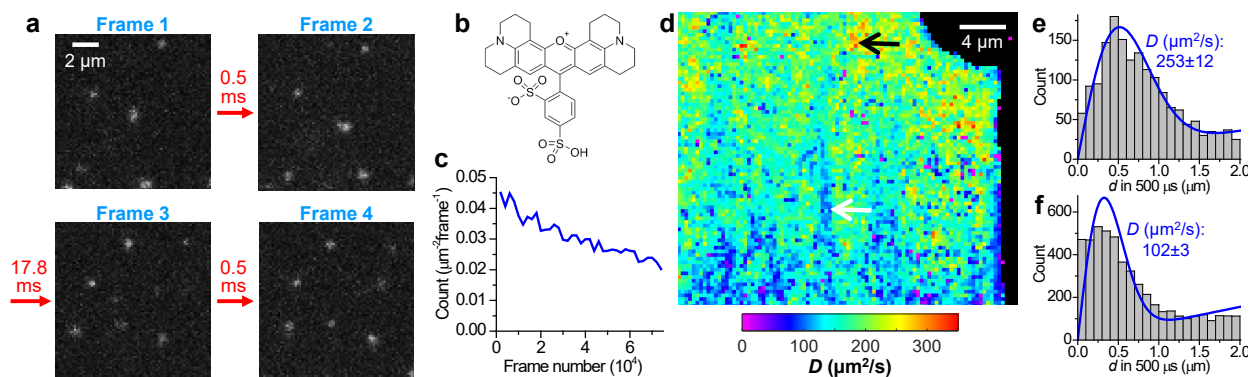


Figure 3.2. SMdM of the water-soluble dye SR101 in primary rat hippocampal astrocytes. (a) Four consecutive frames of the SMdM-recorded single-molecule images of SR101 diffusing inside an astrocyte, with excitation pulse length $\tau = 200$ μs and center-to-center separation $\Delta t = 500$ μs . (b) Chemical structure of SR101. (c) Count of detected single molecules per μm^2 per frame over the recording of 7.5×10^4 frames. (d) Color-coded SMdM D map of SR101 in the cell. (e,f) Local distributions of the recorded 500- μs single-molecule displacements for ~ 1 μm^2 regions of fast (e) and slow (f) diffusion in the cell, as pointed to by the black and white arrows in (d), respectively. Blue lines: MLE fits to our model, yielding D of 253 ± 12 and 102 ± 3 $\mu\text{m}^2/\text{s}$, respectively.

We thus found that for vast regions of the astrocyte cytoplasm, SR101 diffused at $D > \sim 200 \mu\text{m}^2/\text{s}$ (Figure 3.2d), $> \sim 60\%$ of that in PBS (Figure 3.1e). As a control, we compared SMdM under the same excitation laser for astrocytes not loaded with SR101 but expressing the mEos3.2 FP, and observed $D \sim 20 \mu\text{m}^2/\text{s}$ (Figure 3.7), comparable to our previous results in mammalian cells.² For the intracellular diffusion of SR101, the fastest regions reached $D \sim 250 \mu\text{m}^2/\text{s}$ (Figure 3.2de), $\sim 80\%$ of that in PBS. Meanwhile, micrometer-sized structures of lower D were also noted (Figure 3.2d), where the apparent D dropped to $\sim 100 \mu\text{m}^2/\text{s}$ and the distribution was not well fit by our model based on single-component Gaussian diffusion (Figure 3.2f), indicative of population heterogeneity, *e.g.*, a fraction interacting with intracellular organellar membranes. As such interactions may be dynamic and complex, we did not attempt to fit the single-molecule displacements with multiple components; the overall smaller displacements were captured by the imperfect single-component fit as low nominal D values.

Although SR101 is easily loaded into astrocytes, this strategy cannot be generalized to other molecular species or cell types. To deliver additional small solutes into mammalian cells, we harnessed our recent development of graphene-based *in situ* electroporation.⁸³ By coating the glass coverslip with graphene, a conductive monolayer of carbon atoms, the application of voltage pulses across the graphene-coated substrate and a counter electrode (Figure 3.3b) enables electroporation and fluorescent-probe delivery for adherent cells while maintaining full compatibility with single-molecule imaging.

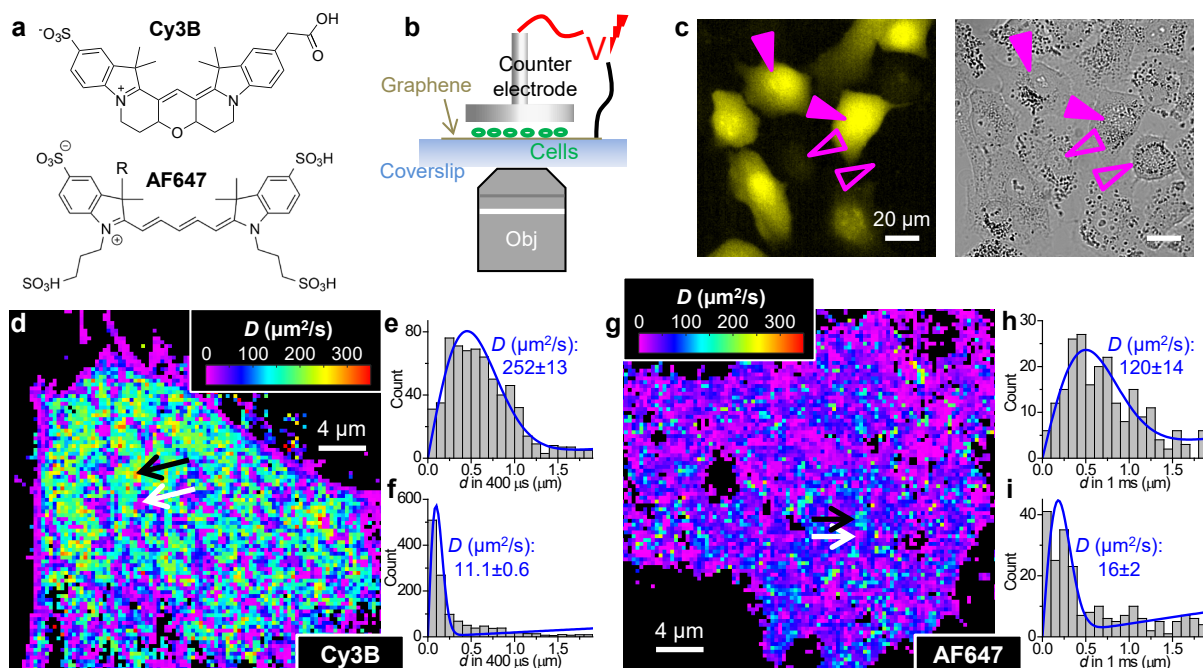


Figure 3.3. SMdM of Cy3B and AF647 dyes delivered into the A549 epithelial cell line via graphene-based electroporation. (a) Molecular structures of the two dyes. (b) Schematics: electroporation by applying a voltage pulse between the graphene-coated coverslip substrate and a counter electrode. (c) Example fluorescence image of Cy3B after electroporation loading into A549 cells (left), vs. bright-field image of the same sample (right). Filled and hollow

arrowheads point to loaded and unloaded cells in the same sample, respectively. (d) SMdM D map of the delivered Cy3B in an A549 cell. (e,f) Local distributions of single-molecule displacements for $\sim 1 \mu\text{m}^2$ regions of fast (e) and slow (f) diffusion in the cell, as pointed to by the black and white arrows in (d), respectively. (g) SMdM D map of the delivered AF647 in an A549 cell. (h,i) Local distributions of single-molecule displacements for $\sim 1 \mu\text{m}^2$ regions of relatively high (h) and low (i) diffusivities. Blue lines in (e,f,h,i): MLE fits to our model, with resultant D values and uncertainties marked in each plot.

We start with two water-soluble fluorescent dyes, the green/yellow-excited Cy3B and the red-excited Alexa Fluor 647 (AF647) (Figure 3.3a). Both dyes are bright and widely used in single-molecule microscopy,^{5, 85} and exhibit no appreciable interactions with lipid bilayers⁶⁶ so may be less trapped by intracellular membranes. Graphene-based electroporation successfully delivered the dyes into A549 human epithelial cells (Figure 3.3c).

SMdM unveiled contrasting intracellular diffusion behaviors for the two dyes. The 560 Da Cy3B exhibited vast regions of $D > \sim 200 \mu\text{m}^2/\text{s}$ (Figure 3.3d), $> \sim 60\%$ of that in the PBS, and the faster regions reached $D \sim 250 \mu\text{m}^2/\text{s}$ (Figure 3.3de), $\sim 75\%$ of that in the PBS (Figure 3.1c-e). Together with our above SR101 results in astrocytes (Figure 3.2), these observations indicate that small solutes may diffuse unhindered in the cytoplasm, only being mildly slowed by the $\sim 20\text{--}40\%$ higher viscosity of the cytosol over water. Meanwhile, we also observed (sub)micrometer-sized features with low nominal D of $\sim 10 \mu\text{m}^2/\text{s}$ and displacement distributions not well fitted by our single-component model (Figure 3.3df), indicative of substantial binding to intracellular structures, *e.g.*, organellar membranes. These results underscore the importance of spatially resolving the local diffusion behavior, so that unhindered diffusion can be separated from the locally bound molecules, as opposed to a mixed global readout that may bias the result to an artificially low D . In contrast, the 959 Da AF647 showed generally slow diffusion in the cell, with vast regions exhibiting nominal $D < \sim 50 \mu\text{m}^2/\text{s}$ (Figure 3.3gi), $< \sim 18\%$ of that in the PBS (Figure 3.1e). Micrometer-sized domains of faster diffusion were also observed, where D reached up to $\sim 120 \mu\text{m}^2/\text{s}$ (Figure 3.3gh), $\sim 40\%$ of that in PBS (Figure 3.1e).

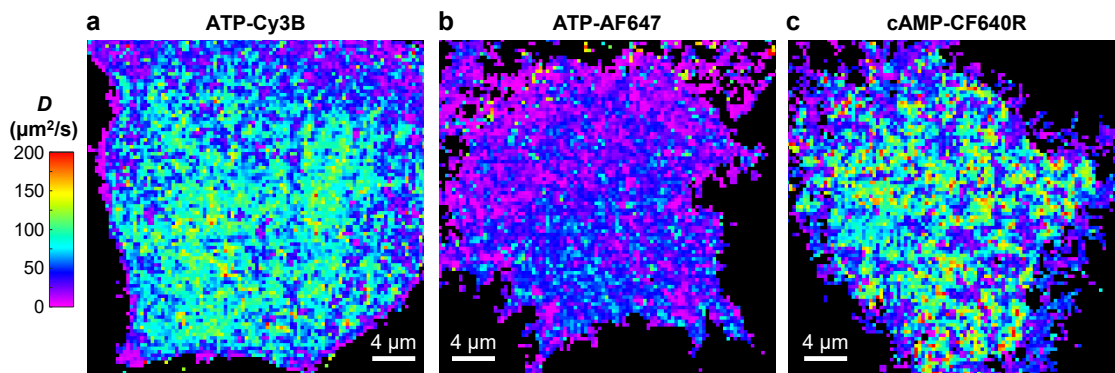


Figure 3.4. SMdM D maps of graphene-electroporation-delivered ATP-Cy3B (a), ATP-AF647 (b), and cAMP-CF640R (c) in A549 cells. Note the different D color scales versus Figures 2 and 3 for the slower diffusion of the larger molecules.

To clarify whether the substantially more hindered intracellular diffusion of AF647 could be due to its larger size over Cy3B, we next examined the intracellular diffusion of Cy3B- and AF647-tagged adenosine triphosphate (ATP). SMdM showed that the 1136 Da ATP-Cy3B, after being loaded into A549 cells *via* graphene-based electroporation, diffused at $D \sim 120 \mu\text{m}^2/\text{s}$ over vast regions (Figure 3.4a), $\sim 50\%$ of that in PBS (Figure 3.1e). Thus, the intracellular diffusion of ATP-Cy3B is less hindered than that of AF647 despite its larger size. In contrast, the intracellular diffusion of the 1434 Da ATP-AF647 (Figure 3.4b) was substantially slower than in the PBS, with vast regions exhibiting $D < \sim 30 \mu\text{m}^2/\text{s}$ ($\sim 15\%$ of in PBS) and sporadic domains reaching $\sim 60 \mu\text{m}^2/\text{s}$ ($\sim 30\%$ of in PBS), analogous to what we observed for the AF647 dye itself (Figure 3.3g). Previous work has shown that AF647 tends to cluster up in mammalian cells,^{67, 68, 88} which may account for its significant slowdown in diffusion.

To further substantiate AF647 as the outlier, we examined CF640R, a different red-excited dye that has shown minimal aggregations in live cells.^{67, 68} With a 1445 Da cyclic adenosine monophosphate (cAMP) conjugate, near-identical in size to ATP-AF647, we observed fast diffusion of $D \sim 130 \mu\text{m}^2/\text{s}$ over vast regions in the A549 cytoplasm (Figure 3.4c), $\sim 60\%$ of that in PBS.

3.3 Conclusion

Together, by utilizing two new strategies to deliver small solutes into mammalian cells and pushing the working D range of SMdM to $>300 \mu\text{m}^2/\text{s}$, we have unveiled the intracellular diffusion patterns for several hydrophilic dyes and their nucleotide conjugates. We thus showed that for these 560-1450 Da solutes, D values $\sim 60\text{-}70\%$ of that in the PBS could be reached in vast regions of the cell, while also noting substantial local slowdowns at the sub-micrometer scales that emphasize the importance of D spatial mapping. The outlying results of AF647 with substantially reduced intracellular D further underscore potential artifacts that may be caused by the label. As the dyes were larger in size than the tagged nucleotides examined in this work, diffusion properties were likely dominated by the former.

The $D_{\text{cyto}}/D_{\text{water}} \sim 60\text{-}70\%$ result suggests that the intracellular diffusion of small solutes is only modestly reduced by the $\sim 20\text{-}40\%$ higher viscosity of the cytosol over water, but otherwise not further hindered by macromolecular crowding as seen with FPs.^{55, 74} While this result contrasts with the general notion that small solutes diffuse at $D_{\text{cyto}}/D_{\text{water}} \sim 25\%$ based on early FRAP measurements on acetoxymethyl ester-loaded BCECF,^{56, 60, 72} NMR has indicated unhindered diffusion of the 507 Da ATP in the rat skeletal muscle,⁷⁵ and fluorescent reporters have implied fast diffusion of the 329 Da cAMP and 345 Da cyclic guanosine monophosphate in mammalian cells consistent with unhindered diffusion.⁸⁹ Unhindered diffusion is plausible for small solutes: proteins and other intracellular macromolecular crowders are often substantially larger than 1 kDa, and it is likely that they do not act as effective diffusion barriers for small solutes. As the cytoplasmic transport of small solutes mainly relies on passive diffusion, lifting the previously suggested, paradoxical speed limit could carry significant implications.

3.4 Materials and Methods

Fluorescent probes. SR101 was from Sigma-Aldrich (S7635). Cy3B and AF647 free dyes were prepared by hydrolyzing the corresponding NHS (*N*-hydroxysuccinimidyl) esters (PA63101, Cytiva, and A200006, Invitrogen) in 0.1 M sodium bicarbonate overnight. ATP-AF647 was from Invitrogen (A22362). cAMP-CF640R was from Biotium (#00037). ATP-Cy3B was conjugated by reacting EDA-ATP (A 072, BIOLOG) with Cy3B-NHS ester at a high excess ratio (1:80). Thin-layer chromatography confirmed negligible unconjugated Cy3B in the product.

Optical setup for SMdM. SMdM was performed on a setup² based on a Nikon Ti-E inverted fluorescence microscope, but with the addition of a 642-nm laser with pulse control. Briefly, 561 nm (OBIS 561 LS, Coherent, 165 mW) and 642 nm (Stradus 642, Vortran, 110 mW) lasers were focused at the back focal plane of an oil-immersion objective lens (CFI Plan Apochromat Lambda 100 $\times$, NA = 1.45) to enter the coverslip-sample interface slightly below the critical angle, thus illuminating a few micrometers into the sample. The illuminated area was ~ 40 μ m in diameter, yielding ~ 10 kW/cm² peak power density. Single-molecule images were recorded in the wide field with an EM-CCD camera (iXon Ultra 897, Andor) that ran continuously at 110 frames per second. A multifunction I/O board (PCI-6733, National Instruments) detected the exposure timing signal of the camera and accordingly modulated the 561 nm (for excitation of SR101 and Cy3B) and 642 nm (for excitation of AF647 and CF640R) lasers. Paired pulses were thus applied across tandem camera frames (Figure 1a). Typical center-to-center separations between the paired pulses were $\Delta t = 400$ -1000 μ s, with the pulse durations $\tau = \Delta t/2$. The timing waveforms were confirmed by a GW Instek GDS-1054B oscilloscope (Figure 3.5). The tandem excitation scheme was repeated $\sim 10^4$ times for each SMdM run. As we illuminated a few micrometers into the sample, we imaged and photobleached a thin layer at the bottom. Diffusional exchange with deeper parts of the sample enabled the stochastic replenishment of single molecules in the view for sustained imaging over long recording times, a process facilitated by the ample time (~ 18 ms) between the anti-paired pulses in our excitation sequences.

SMdM in PBS. 25-mm diameter, #1.5 coverslips were treated in a heated 3:1 H₂SO₄:H₂O₂ (30%) mixture for 45 minutes, rinsed with Milli-Q water, and then dried using N₂ gas. To reduce out-of-focus backgrounds due to fluorescent probes at the coverslip surface, the coverslips were functionalized with 10 mg/mL methoxy PEG silane (5 kDa, M-SLN-5000, JenKem Technology) in 95% ethanol/water for 30 min, and rinsed and sonicated for 5 min in Milli-Q water. Fluorescent probes were diluted to ~ 100 pM in PBS (14190144, Gibco), and SMdM was performed ~ 2 μ m away from the glass surface in the solution.

SMdM of SR101 in primary rat hippocampal astrocytes. Primary rat hippocampal astrocytes were from BrainBits and cultured on polylysine-coated #1.5 glass coverslips (GG-18-1.5-PDL, Neuviro) in the NbAstro medium following the recommended protocols. For SR101 loading, the sample was incubated in the NbAstro medium with 100 nM SR101 for 20 min, followed by rinsing with PBS for three times. Previous studies^{81, 82} have applied SR101 at a wide range of concentrations (~ 0.1 -500 μ M), under which conditions astrocytes stay healthy, and we found 100

nM optimal for our SMdM experiments. SMdM of the labeled cells was carried out in Leibovitz's L-15 medium (21083027, Gibco) with 20 mM HEPES (15630106, Gibco).

Graphene electroporation-based intracellular delivery and SMdM in A549 cells. Graphene electroporation was performed as described previously.⁸³ Briefly, monolayer graphene (Graphene Supermarket) was transferred onto #1.5 glass coverslips. A549 cells (University of California Berkeley Cell Culture Facility) were cultured following standard protocols in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with GlutaMAX-I (10566-016, Gibco) and 10% fetal bovine serum, and plated onto the graphene-coated coverslip. After ~24 h, the culture medium was changed to an electroporation buffer (1652676, Bio-Rad) containing the fluorescent probe to be delivered at ~100 nM. Electroporation was achieved by applying a 15 V voltage pulse (~5-10 ms halftime) across the graphene-coated substrate and a counter electrode ~500 μm above. After ~10 min incubation, the sample was rinsed with the imaging buffer (L-15 medium + 20 mM HEPES) 3 times, and SMdM was performed as described above.

Data analysis. SMdM data were analyzed as described previously.² Briefly, single molecules were super-localized in each recorded frame, and their two-dimensional displacements between paired frames were calculated. The accumulated displacement values were spatially binned onto a $320 \times 320 \text{ nm}^2$ grid. For each spatial bin, the distribution of local displacements was fitted through maximum likelihood estimation (MLE) to a probability model based on two-dimensional random walk:²

$$P(r) = \frac{2r}{a} \exp\left(-\frac{r^2}{a}\right) + br \quad (\text{eqn. 1})$$

where r is the single-molecule displacement in the fixed time interval Δt , $a = 4D\Delta t$, and b is a background term accounting for mismatched molecules that randomly diffuse into the field of view during Δt , as discussed previously.² Performing MLE for all spatial bins generated a map of local D for color rendering. Analysis of two-dimensional vectorial single-molecule displacements showed no noticeable anisotropy, and so fitting to a one-dimensional random-walk model⁶⁴ yielded comparable D values (Figure 3.8).

Supplementary Figures

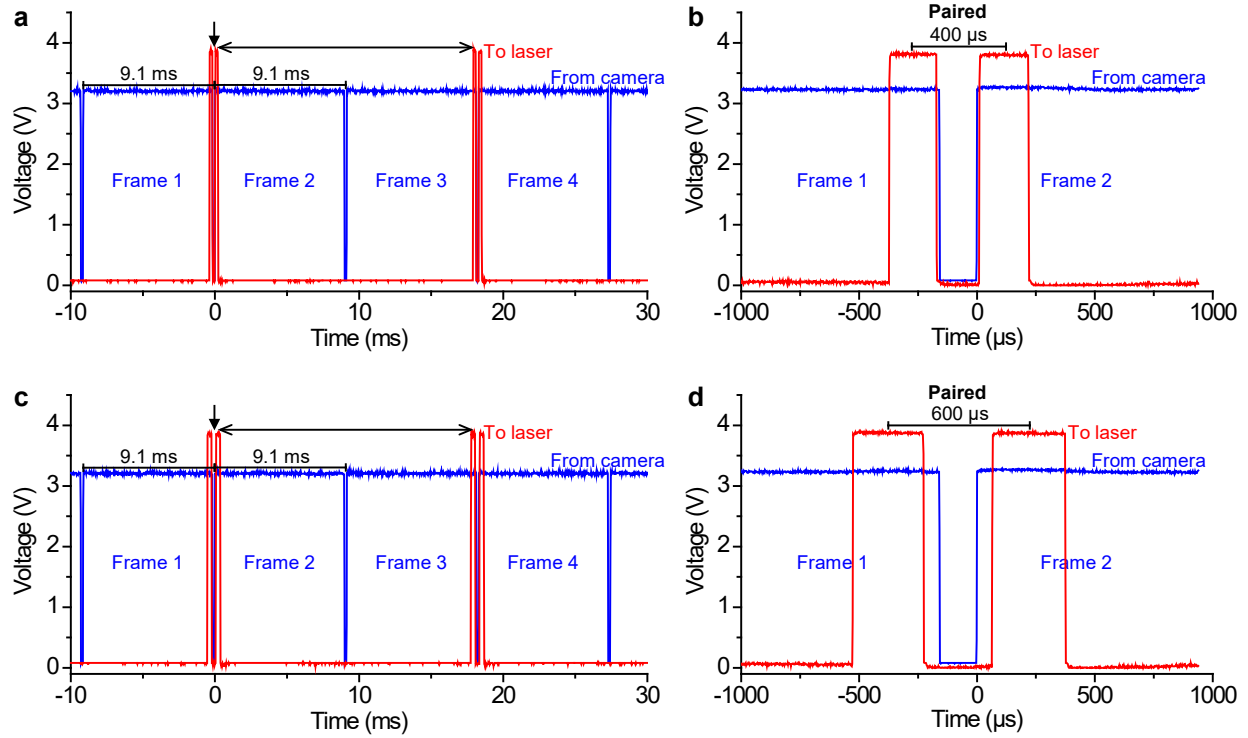


Figure 3.5. Oscilloscope-measured typical timing waveforms in this study. Blue traces: the “fire” TTL signal from the EM-CCD camera, which ran at 110 fps (9.1 ms/frame). 3.3 V: exposure; 0 V: dead time between frames ($\sim 160 \mu\text{s}$). Red traces: output from the PCI-6733 card, which modulates the 561- or 642-nm excitation lasers between off (low voltage) and on (high voltage) states. (a,b) Pulse center-to-center separation $\Delta t = 400 \mu\text{s}$ and pulse durations $\tau = 200 \mu\text{s}$, shown as zoom-out for four consecutive camera frames (a) and zoom-in for the tandem pulses across the two paired frames (b). (c,d) Pulse center-to-center separation $\Delta t = 600 \mu\text{s}$ and pulse durations $\tau = 300 \mu\text{s}$. Note the very long ($\sim 18 \text{ ms}$) time separations between the anti-paired pulses.

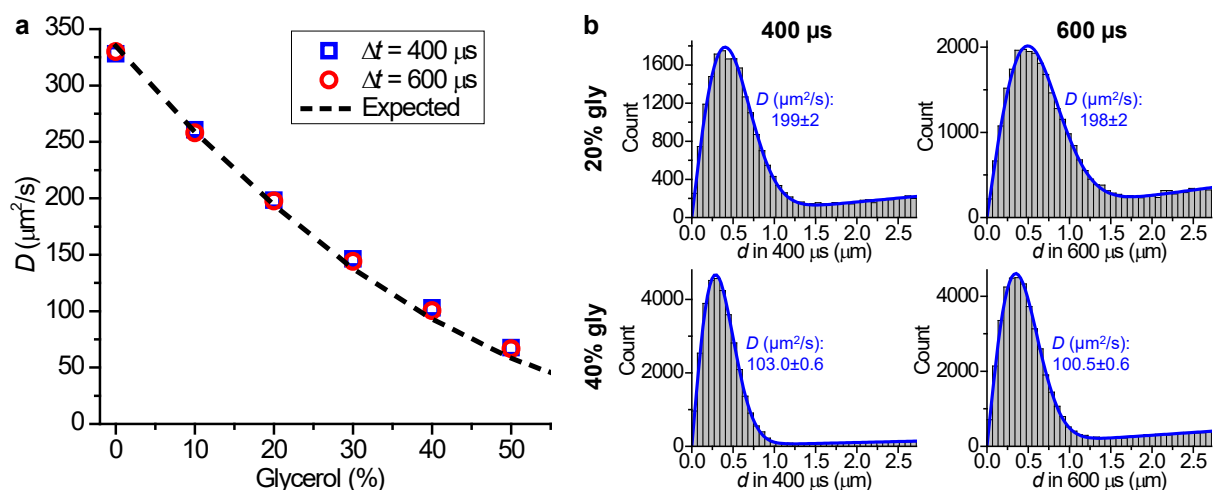


Figure 3.6. SMdM results on Cy3B diffusing in PBS containing different amounts of glycerol. (a) SMdM-determined diffusion coefficient D of Cy3B as a function of the glycerol weight percentage, for experiments carried out at center-to-center pulse separations Δt of $400 \mu\text{s}$ (blue squared) and $600 \mu\text{s}$ (red circles), compared to that is predicted by scaling the 0% glycerol value with the known glycerol concentration-dependent viscosity (dash line).⁹⁰ (b) Distributions of the recorded single-molecule displacements for results in 20% and 40% glycerol at $\Delta t = 400$ and $600 \mu\text{s}$. Blue lines: maximum likelihood estimation (MLE) fits to our model, with resultant D values and uncertainties marked in each plot.

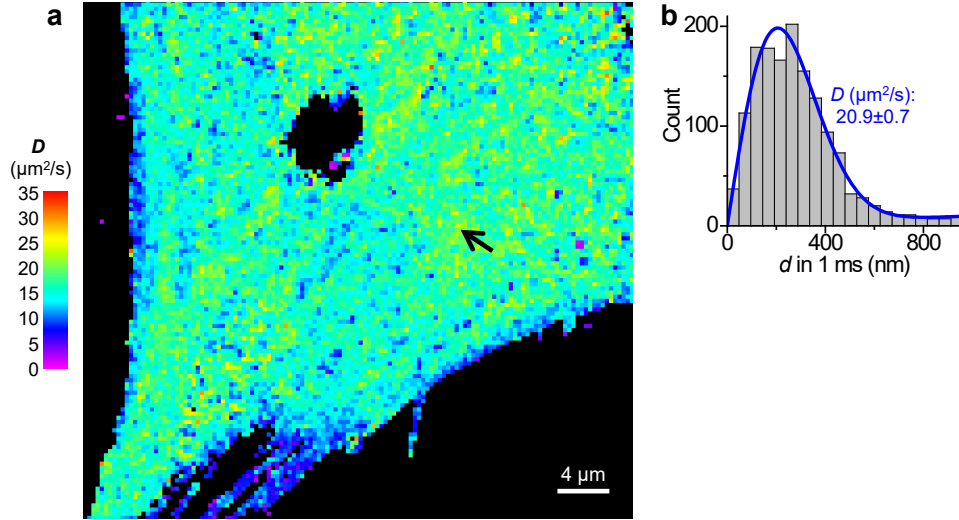


Figure 3.7. SMdM under the same 561-nm excitation laser for an astrocyte not loaded with SR101 but expressing the mEos3.2 FP. (a) Color-coded SMdM D map. Note the $10\times$ lower D values for the color scale compared to Figure 2d. (b) Local distribution of the recorded 1-ms single-molecule displacements for an $\sim 1 \mu\text{m}^2$ region pointed to by the black arrow in (a). Blue line: MLE fit to our model, yielding $D = 20.9 \pm 0.7 \mu\text{m}^2/\text{s}$.

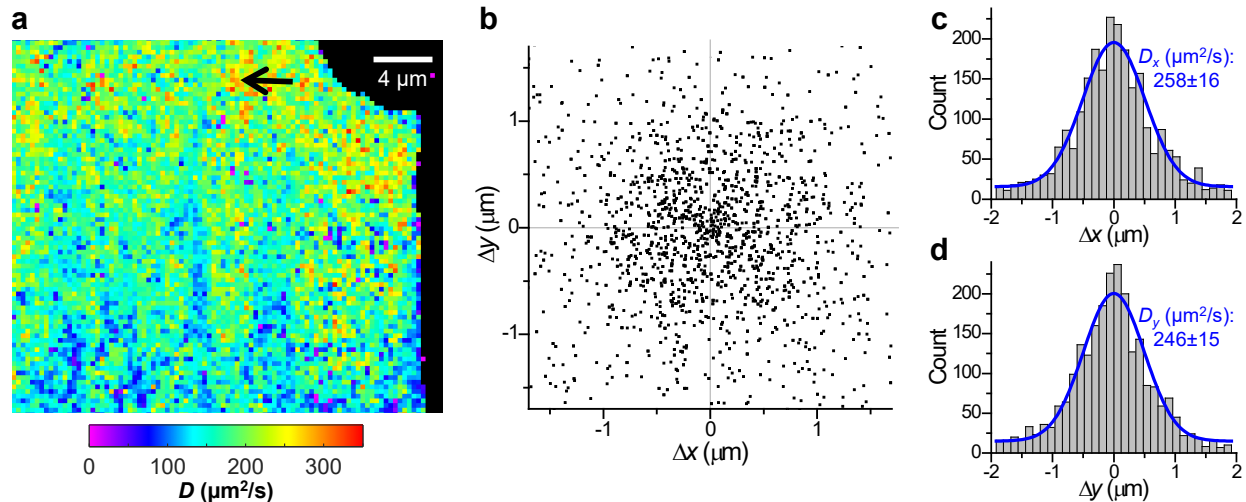


Figure 3.8. Two-dimensional vectorial single-molecule displacement analysis. (a) Color-coded principal-direction SMdM (pSMdM) diffusivity map for the same data shown in Figure 2d, obtained by first calculating the preferred diffusion direction in each spatial bin, and then fitting the distribution of single-molecule displacements along that direction to a modified one-dimensional random-walk model with the probability distribution of $P(x) = \frac{1}{\sqrt{a\pi}} \exp(-\frac{x^2}{a}) + b$, where $a = 4D\Delta t$ ($\Delta t = 500 \mu\text{s}$) and b accounts for a uniform background.⁶⁴ (b) Two-dimensional plots of the vectorial 500- μs single-molecule displacements accumulated for a region pointed by the black arrow in (a), showing no noticeable anisotropy. (c,d) Distributions of the displacements in (b) along the horizontal (c) and vertical (d) directions. Blue curves: MLE fits to the modified one-dimensional random-walk model above, yielding $D_x = 258 \pm 16$ and $D_y = 246 \pm 15 \mu\text{m}^2/\text{s}$, comparable to D obtained from fitting the scalar magnitude of single-molecule displacements to the two-dimensional random-walk model (Figure 3.2e).

Chapter 4: Single-molecule diffusivity quantification unveils ubiquitous net charge-driven protein-protein interaction

The work in this chapter was reproduced with permission from the following publication:

A.A. Choi, K. Xu, “Single-molecule diffusivity quantification unveils ubiquitous net charge-driven protein-protein interaction” *J. Am. Chem. Soc.*, **146**, 10973-10978, 2024. Copyright © 2024 American Chemical Society.

4.1 Introduction

Molecular diffusion stands as a fundamental process of the cell while also providing a valuable window into intermolecular interactions.^{55, 56, 59, 61} We recently developed single-molecule displacement/diffusivity mapping (SMdM),² in which the mass statistics of transient ($\sim <1$ ms) single-molecule displacements enables the quantification of molecular translational diffusion *in vitro*^{11, 12, 91} and in the living cell.^{2, 11, 64, 92} For fluorescent proteins (FPs) expressed in the mammalian cell, SMdM unexpectedly unveils the protein net charge as a key determinant of their diffusivity. Intriguingly, across diverse cytoplasmic, nuclear, and intraorganellar environments, the possession of positive, but not negative, net charges substantially impedes diffusion.^{2, 92} These results echo other recent microscopy and NMR observations in bacterial and mammalian cells generally showing slower diffusion for positively charged proteins.⁹³⁻⁹⁷

Our analysis of the net charges of major proteins in the mammalian cell (per mass spectrometry-detected abundances^{98, 99}) shows an excessive preponderance of negative values.^{2, 92} Combined with the known high abundance of intracellular small cations (~ 150 mM) but low availability of small anions (~ 20 mM),¹⁰⁰⁻¹⁰² we thus reason that in the mammalian cytoplasm, positive charges are predominantly carried by small ions, whereas negative charges are mainly carried by RNA and proteins (Table 4.1). Consequently, it is plausible that in the mammalian cell, positively charged proteins will be dragged by the predominantly negatively charged intracellular macromolecules, thus explaining the above experimental observations of charge-sign asymmetric impediment of diffusion.

However, the above model, which would imply strong ubiquitous net charge-driven protein-protein interactions in intracellular environments, has yet to be validated experimentally or recapitulated *in vitro*. As intracellular proteins are predominantly negatively charged, protein-induced macromolecular crowding¹⁰³⁻¹⁰⁵ is often examined between proteins of the same, negative

charge sign. Meanwhile, oppositely charged biomacromolecules have been mixed at near equimolar ratios at high (>0.1 wt% or >10 μM) concentrations to generate polyelectrolyte complexes/coacervates as phase-separated, micrometer-sized droplets,¹⁰⁶⁻¹¹⁰ with recent work extended to cells.^{111, 112}

Uniquely enabled by the single-molecule detection mechanism of SMdM, here we examine in aqueous solutions, in the low-concentration (~ 200 pM) limit of diffuser proteins, how interactor proteins of opposite and same charge signs, over wide concentration ranges, differently interact with the diffuser and thus modulate its translational diffusivity. This regime avoids the potential interferences of aggregate/coacervate formation, and thus better emulates how individual diffuser molecules interact with the crowded, predominantly negatively charged intracellular environment. We thus document ubiquitous, net charge-driven nonspecific protein-protein interactions, wherein electrostatic attractions between species of opposite net charges substantially suppress diffusion. By converting the measured diffusivity to diffuser diameters, we further show that in the limit of excess interactors, a positively charged diffuser molecule effectively drags along a monolayer of negatively charged interactors, thus a model to elucidate the experimentally observed diffusion suppression of positively charged proteins in the living cell.

4.2 Results and Discussion

We started by examining the diffusion of hen-egg-white lysozyme (HEWL) in a solution of bovine serum albumin (BSA). With an isoelectric point of ~ 10.7 , the diffuser HEWL has an estimated net charge of $+7$ at the cytoplasmic pH of 7.3 (Table 4.2). Meanwhile, the interactor, the common protein standard BSA has an isoelectric point of ~ 5.0 and is -17 charged at pH 7.3 (Table 4.2).

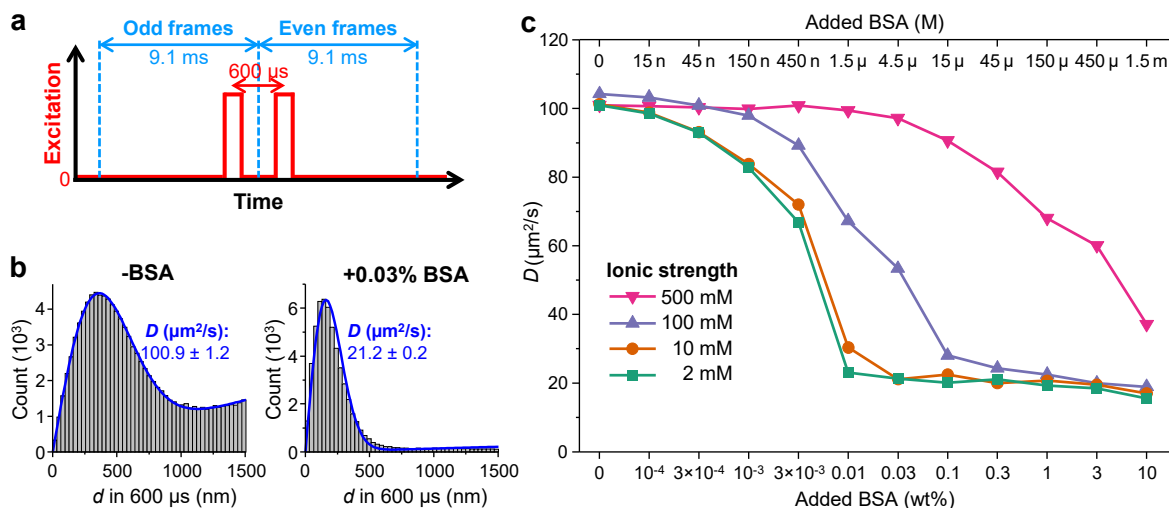


Figure 4.1. Single-molecule diffusivity quantification unveils strong HEWL-BSA interactions in the low-concentration limit. (a) Schematics of the excitation sequence in SMdM, in which paired pulses were repeatedly applied across tandem camera frames to capture transient single-molecule displacements in the wide field over the time window defined by the pulse separation. (b) Example histograms of single-molecule displacements from SMdM, for 200 pM Cy3B-labeled HEWL diffusing in a pH = 7.3 buffer of 2 mM ionic strength, with and without the addition

of 0.03% BSA. Blue lines: fits to the SMdM diffusion model, yielding diffusion coefficients D of 100.9 ± 1.2 and $21.2 \pm 0.2 \mu\text{m}^2/\text{s}$ (95% confidence intervals), respectively. (c) SMdM-determined D values for 200 pM Cy3B-labeled HEWL in pH = 7.3 buffers at ionic strengths of 2, 10, 100, and 500 mM, as a function of added BSA (bottom scale: weight percentage wt%; top scale: molar concentration).

Fluorescently labeled HEWL was first diluted to ~ 200 pM in a pH = 7.3 buffer. At this low concentration, the average distance between individual HEWL molecules in the solution was $\sim 2 \mu\text{m}$, thus permitting robust single-molecule detection and minimizing the likelihood of interactions between different HEWL molecules. For SMdM, an excitation laser repeatedly illuminated the sample as paired stroboscopic pulses across tandem camera frames at a fixed center-to-center separation of $600 \mu\text{s}$ (Figure 4.1a), thus enabling the rapid wide-field detection of $\sim 10^5$ transient single-molecule displacements for the $600 \mu\text{s}$ time window in ~ 3 min (Methods).^{2, 12} Fitting the accumulated single-molecule displacements to our diffusion model (Figure 4.1b and Methods) yielded diffusion coefficient D with $\sim 1\%$ precisions.¹² While the sub-millisecond time scale we probe does not resolve the individual transient interactions between biomolecules, it characterizes the averaged effects relevant to intracellular transports at sub-micrometer length scales.

For the starting 1 mM phosphate pH = 7.3 buffer (ionic strength ~ 2 mM), as well as buffers with KCl added to ionic strengths of 10, 100, and 500 mM, SMdM determined comparable D values of 101 - $104 \mu\text{m}^2/\text{s}$ for Cy3B-labeled HEWL (Figure 4.1b). These values agree with our previous SMdM results, expected based on the 15 kDa molecular weight.^{12, 42}

Notably, as we added BSA to the sample, SMdM detected substantial drops in D in a concentration-dependent manner (Figure 4.1c). In buffers of low ionic strengths of 2 and 10 mM, $\sim 10\%$ suppression in D was noted at 3×10^{-4} wt% (weight percentage), or 3 ppm (parts per million), addition of BSA. Note that 3 ppm BSA (45 nM) was still >200 -fold over the ~ 200 pM concentration of the HEWL diffuser used in the measurement, and so interactions were not limited by the available amount of the BSA interactor in the solution. Increasing the BSA concentration to 0.01% (100 ppm; $1.5 \mu\text{M}$) suppressed the HEWL D values to $<30 \mu\text{m}^2/\text{s}$, and this value stabilized to $\sim 20 \mu\text{m}^2/\text{s}$ for BSA concentrations of $>0.03\%$ ($4.5 \mu\text{M}$).

With the higher ionic strength of 100 mM, $\sim 10\%$ suppression in D was observed at 30 ppm (450 nM) BSA addition, and 300 ppm and 0.3% BSA additions suppressed the HEWL D values to ~ 50 and $\sim 20 \mu\text{m}^2/\text{s}$, respectively. For the very high ionic strength of 500 mM, diffusion suppression occurred at $>\sim 0.1\%$ (15 μM) BSA addition, with 1% and 10% BSA additions suppressing D to 68 and 38 $\mu\text{m}^2/\text{s}$, respectively.

To further elucidate the effects of ionic strength in this system, in another experiment, we initiated BSA-induced diffusion slowdown for HEWL in a 1 mM phosphate pH = 7.3 buffer, and then gradually added KCl to raise the ionic strength. SMdM detected HEWL diffusivity recovery in this process (Figure 4.2a). Together, the observed ionic strength-dependent BSA suppression of

HEWL diffusion suggests that the HEWL-BSA interaction is driven by electrostatic attractions and thus reversible when the protein charges are screened by high salts.

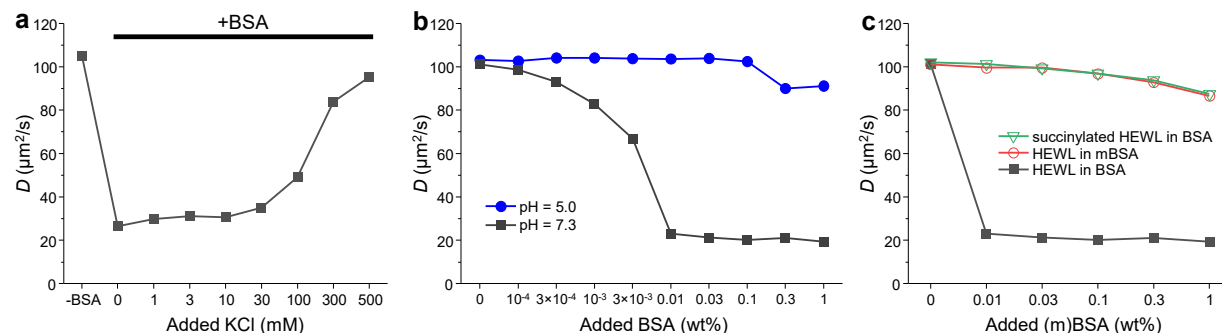


Figure 4.2. The HEWL-BSA interaction is driven by electrostatic attractions based on the protein net charge. (a) SMdM-determined diffusion coefficient D for 200 pM Cy3B-labeled HEWL initially diffusing in a pH = 7.3 buffer of 2 mM ionic strength, to which 0.03% BSA and then KCl of different concentrations were sequentially added. (b) SMdM-determined D values for 200 pM Cy3B-labeled HEWL diffusing in 1 mM phosphate (pH = 7.3) or acetate (pH = 5.0) buffers, at different wt% of added BSA. (c) SMdM-determined D values in pH = 7.3 buffers of 2 mM ionic strength, for 200 pM Cy3B-labeled HEWL in the presence of varied wt% of methylated BSA (mBSA) (red) and 200 pM Cy3B-labeled succinylated HEWL (green) and HEWL (black) in the presence of varied wt% of BSA.

To directly probe the charge effects, we compared SMdM results in a 1 mM acetate buffer of pH = 5.0, the isoelectric point of BSA. We thus observed no changes in the HEWL diffusivity for BSA additions up to 0.1%, and only ~10% drops at higher BSA additions (Figure 4.2b). These results corroborate our mechanism based on charge interactions, which vanished as BSA became neutral at its isoelectric point.

With the pH = 7.3 buffer, we further examined the diffusions of HEWL in methylated BSA (mBSA) and succinylated HEWL in BSA. These chemical modifications flip the net charges of BSA and HEWL to positive and negative, respectively.^{113, 114} We thus found that for both systems, no appreciable suppression of diffusion occurred for up to 0.1% mBSA/BSA additions (Figure 4.2c). ~10% D depressions were noted at 1% mBSA/BSA additions, attributable to slight increases in the solution viscosity. These results suggest negligible interactions between diffusers and interactors of the same charge signs. Together, SMdM showed that the BSA suppression of HEWL diffusion is driven by electrostatic attractions owing to the opposite net charges carried by the two proteins.

To generalize the above findings of electrostatic attraction-driven protein interactions, we next probed the diffusion of HEWL in the presence of the 45 kDa ovalbumin (isoelectric point: 4.5; estimated net charge at pH 7.3: -12), another common protein standard readily obtainable at bulk quantities. SMdM showed substantial, ionic strength-sensitive suppressions of diffusion as ovalbumin was added (Figure 4.3a), similar to our results above with BSA addition.

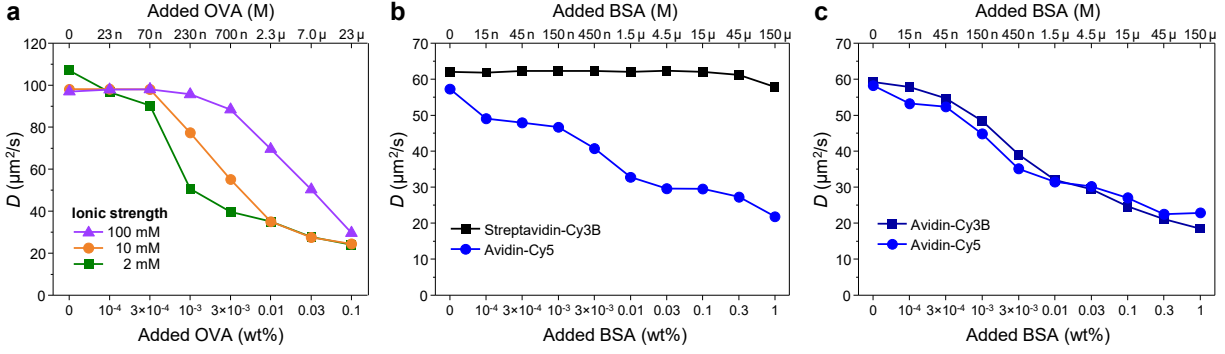


Figure 4.3. Generalizing charge-driven interactions to HEWL-ovalbumin and avidin-BSA systems. (a) SMdM-determined diffusion coefficient D for 200 pM Cy3B-labeled HEWL in pH = 7.3 buffers at ionic strengths of 2, 10, and 100 mM, as a function of added ovalbumin (bottom scale: weight percentage wt%; top scale: molar concentration). (b) Two-color SMdM D values for 200 pM Cy5-labeled avidin and 200 pM Cy3B-labeled streptavidin concurrently added to a pH = 7.3 buffer of 2 mM ionic strength, as a function of added BSA. (c) Two-color SMdM D values in another sample under similar conditions as (b), but with Cy3B-labeled avidin replacing Cy3B-labeled streptavidin.

In another assay, we compared the diffusion behaviors of avidin and streptavidin in BSA solutions. While similar in structure and both tightly bind biotin for bioconjugation applications, the ~68 kDa avidin tetramer has an isoelectric point of 10.5 and an ~+23 net charge at pH 7.3, whereas the ~60 kDa streptavidin tetramer has an isoelectric point of 6.4 and an estimated net charge of -4 at pH 7.3. Notably, as we separately labeled avidin and streptavidin with Cy5 and Cy3B and added both into the same sample to perform SMdM in two color channels, the addition of BSA suppressed diffusion of the former but not the latter (Fig. 4.3b). In comparison, in another sample in which Cy5- and Cy3B-labeled avidin were added into the same solution for two-color SMdM, similar diffusion suppressions by BSA were observed for both (Fig. 4.3c).

To further generalize the above net-charge-based interaction/noninteraction rules, we compared the diffusion of 8 different proteins (Table 4.2), each at ~200 pM concentration, in pH 7.3 buffers with varied BSA additions (Figure 4.4a). Before adding BSA, the SMdM-determined D values of the different proteins (Figure 4.5) showed molecule-weight dependences in excellent agreement with the Young-Carroad-Bell model,⁴² consistent with our previous results.¹² With BSA added, the 3 negatively charged proteins, Cy3B-labeled BSA, α -lactalbumin, and streptavidin, exhibited unvaried diffusivities up to 0.1% BSA addition and <10% suppressions at 1% BSA addition attributed to the aforementioned increases in solution viscosity. In contrast, the 4 substantially positively charged (>+5) proteins (Cy3B-labeled HEWL, CXCL9, avidin, and Histone H1.0) all exhibited markedly suppressed diffusion at ppm-level BSA additions, and converged to similar D values of ~20 $\mu\text{m}^2/\text{s}$ at >0.1% BSA levels. The mildly positively charged Cy3B-labeled RNase A (~+3 net charge) showed an intermediate behavior in which modest diffusion suppressions were observed for >0.01% BSA additions, with an ~40% suppression achieved at 1% BSA. Together, SMdM has demonstrated ubiquitous, nonspecific protein-protein interactions owing to electrostatic attractions between opposite net charges.

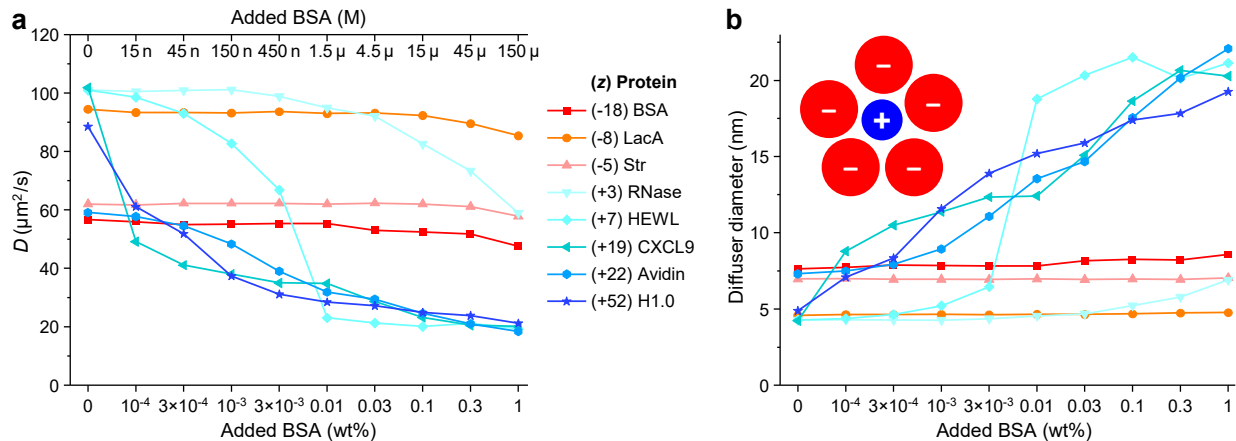


Figure 4.4. SMdM survey of the diffusivity of different proteins in BSA solutions shows consistent slowdowns for positively charged species. (a) SMdM-determined diffusion coefficients D for 8 different Cy3B-labeled diffuser proteins separately added at 200 pM to pH = 7.3 buffers of 2 mM ionic strength, as a function of added BSA (bottom scale: weight percentage wt%; top scale: molar concentration). The proteins are listed in the legend with their estimated net charge z . LacA, α -lactalbumin; Str, streptavidin; H1.0, Histone H1.0. See also Table S2. (b) Effective diffuser diameters ($2 \times$ Stokes radii) of the different diffusers at different BSA concentrations, as converted from the D values in (a) through the Stokes-Einstein equation (methods). Inset model: A positively charged diffuser drags along a dynamic shell of negatively charged BSA interactors.

To rationalize the above curious convergence of D to $\sim 20 \mu\text{m}^2/\text{s}$ at high BSA concentrations for the positively charged proteins, we converted the D values in Fig. 4.4a to effective diameters ($2 \times$ Stokes radii) of the diffusers using the Stokes-Einstein equation (Methods). We thus found that as the differently sized diffuser proteins started with dissimilar diameters in the buffer, the negatively charged species maintained their effective diameters with BSA addition, whereas the positively charged species had their effective diameters quickly grown and stabilized to ~ 17 – 22 nm for $>0.1\%$ BSA additions (Fig. 4.4b). This size is $\sim 3 \times$ the diameter of the interactor BSA (~ 7.5 nm based on our D values and previous studies), thus pointing to a model in which the positively charged diffuser, in the limit of low abundance, drags along a potentially dynamic shell of negatively charged BSA (Fig. 4.4b inset) to manifest a diffusion coefficient commensurate with that size.

As the positive net charge on a diffuser molecule is readily overcompensated by just a few BSA molecules, after the positive diffuser attracts a first layer of BSA interactors, the resultant dynamic complex will appear to the outside as being negatively charged. It is thus likely that electrostatic attraction stops right with the first BSA layer, which defines the apparent D of the diffuser regardless of further increases in BSA concentration. Before that limit is reached, how the D value evolves as a function of the interactor concentration may depend on the sizes and charge distributions of the diffuser and the interactor, awaiting future theoretical and experimental investigations. The low diffuser concentration allowed in SMdM measurements was instrumental in elucidating the above model, so that the amounts of the interactor remained in high excess over the diffuser throughout the measurement. It is thus unlikely for further complexing to occur

through the interaction of multiple diffuser molecules with the same interactor molecule, a condition previously shown to generate micrometer-sized aggregates or coacervates when oppositely charged biomacromolecules are mixed at near equimolar ratios.^{106-109, 111, 112}

4.3 Conclusion

In conclusion, through SMdM, we have quantified how protein diffusion, in the limit of low diffuser concentrations, is modulated by differently charged interactor proteins over wide concentration ranges. We thus unveiled ubiquitous, net charge-driven protein-protein interactions and showed that in the limit of excess interactors, a positively charged diffuser molecule effectively drags along one monolayer of negatively charged interactors.

As the mammalian intracellular environment is dominated by negatively charged proteins^{2, 92} and RNA counterbalanced by small cations (Table 4.1), the rarely encountered positively charged proteins would similarly drag along a dynamic shell of negatively charged macromolecules without further inducing aggregates or coacervates. The resultant complex thus provides an intuitive, refined model for the mechanism underlying the experimentally observed diffusion slowdown of positively charged proteins in the living cell. Intracellular macromolecular crowding, which progressively obstructs the diffusion of larger particles,^{56, 103} conceivably further amplifies diffusion suppression, and such effects may be examined in future studies. Interactions with cellular membranes, membrane proteins, and membrane-associated cytoskeletons,¹¹⁵ as well as the varied environments inside diverse organelles,⁹² further add to the equation. Meanwhile, the demonstrated SMdM quantification of effective diffuser diameters opens new paths to elucidating molecular interactions through single-molecule detection.

4.4 Supplemental Material and Methods

Fluorescent labeling of proteins. Sources of proteins used are shown in **Table S2**. The proteins were labeled with Cy3B NHS (*N*-hydroxysuccinimidyl) ester (Cytiva PA63101) or Cy5 NHS ester (Lumiprobe 13020) in 0.1 M NaHCO₃ at an ~1:1 initial dye-to-protein ratio for 1 h at room temperature. Unconjugated dye was removed by filtering through Amicon 10k or 30k MWCO centrifugal filters (Millipore, UFC501096 and UFC503096) six times. Absorbances at 280 and 560 nm, as measured by a NanoDrop 2000c spectrometer (ThermoFisher), indicated that the final product had ~0.5 dyes per protein on average. Consequently, the fluorescently detected molecules in our SMdM experiments typically had only one dye on the protein, while the unlabeled fraction of the protein was undetected. Succinylation of Cy3B-labeled HEWL was done by adding excessive solid succinic anhydride (Fisher Scientific AC158760050) into the sample¹¹⁶ and then cleaning up with Amicon centrifugal filters. SMdM experiments were performed on the day of labeling.

Sample preparation. #1.5 Glass coverslips were treated in a heated 3:1 H₂SO₄/H₂O₂ (30%) mixture for 45 min, rinsed with Milli-Q water, and then dried with N₂ gas. The coverslips were

passivated with 10 mg/mL methoxy-PEG silane (MW 5000, PG1-SL-5k, Nanocs) in 95% ethanol/water for 30 min, and then rinsed and sonicated for 5 min in Milli-Q water.¹² The coverslips were then each mounted with a plastic tube (cut from a 1.6 mL microcentrifuge tube, inner diameter ~9 mm) to form an imaging chamber, as described previously.⁸³ The 1 mM phosphate buffer (pH = 7.3) was prepared by diluting a 0.1 M phosphate buffer (Sigma P5244) with Milli-Q water, yielding an initial ionic strength of ~2 mM. The 1 mM acetate buffer (pH = 5.0) was prepared by diluting a 0.1 M potassium acetate buffer (Sigma-Aldrich, SAE0157). The ionic strengths of the buffers were then adjusted by adding KCl. The above dye-labeled proteins were diluted to ~200 pM into the buffers, and then added into the above assembled imaging chambers for SMdM. Interactor proteins were first prepared as 1% or 10% stock solutions in the same buffers, adjusted back to the original pH of the buffers as necessary, and then gradually added to the sample with appropriate serial dilutions as necessary.

SMdM. SMdM was performed on a Nikon Ti-E inverted fluorescence microscope, as described previously.^{2, 12} Briefly, a 561 nm laser (OBIS 561 LS, Coherent, 220 mW) and a 642 nm laser (Stradus 642, Vortran, 110 mW) were focused at the edge of the back focal plane of an oil-immersion objective lens (CFI Plan Apochromat Lambda 100×, NA = 1.45) to enter the sample slightly below the critical angle of the glass-solution interface to illuminate a few micrometers into the solution. The focal plane was set at ~3 μm away from the glass surface into the solution, which was maintained during the experiment by the Nikon Perfect Focus System. An I/O board (PCI-6733, National Instruments) detected the exposure timing signal of the recording EM-CCD camera (iXon Ultra 897, Andor), which operated continuously at 110 frames per second (fps), and accordingly modulated the emitting power of the excitation lasers. Paired pulses of $\tau = 300 \mu\text{s}$ duration and $\Delta t = 600 \mu\text{s}$ center-to-center separation were thus repeatedly applied across tandem camera frames (Figure 1a), so that wide-field images were captured for the diffusing single molecules to evaluate their displacements over the Δt time window. Typical runs executed 1.0×10^4 pairs of tandem frames, thus data-collection times of ~3 min for each condition.

Data analysis. SMdM data were analyzed as described previously.^{2, 12} Briefly, single molecules were super-localized in each recorded frame, and then their two-dimensional displacements d were calculated between the paired frames. The accumulated $0.5\text{-}5 \times 10^5$ single-molecule displacements in each run were pooled together for plotting of distribution (*e.g.*, Fig. 4.1b) and fitting to the model:²

$$P(r) = \frac{2r}{a} \exp\left(-\frac{r^2}{a}\right) + br \quad (\text{eqn. 1})$$

Here $a = 4D\Delta t$ with D being the diffusion coefficient, and b accounts for a uniform background due to extraneous molecules that randomly diffuse into the search radius during Δt .² MLE with MATLAB returned 95% confidence intervals for a , which were then converted to D using the fixed Δt (600 μs) in the experiment.

Estimation of effective diffuser diameters. Effective diffuser diameters were calculated from the SMdM-determined D values as $2\times$ the Stokes radius through the Stokes-Einstein equation:

$$Diameter = 2r = 2 \frac{k_B T}{6\pi\eta D} = \frac{k_B T}{3\pi\eta D} \quad (\text{eqn. 2})$$

Here, k_B is the Boltzmann constant, $T = 295$ K is the temperature, and η is the solution viscosity, which was taken as 1.00 mPa·s for solutions with <0.1 wt% BSA addition, and 1.02 and 1.06 mPa·s for the 0.3% and 1% BSA solutions, respectively.

Table 4.1. Accounting for molecular charges in the mammalian cytoplasm

Small ions:^{100, 101}

Positive charges

K^+ : 140 mM

Na^+ : 12 mM

Negative charges

Cl^- : ~7 mM

HCO_3^- : ~10 mM

$H_2PO_4^-/HPO_4^{2-}$: ~2 mM

Total: ~150 mM positive charges and ~20 mM negative charges.

RNA:

Every nucleotide gives one negative charge to the RNA backbone.

Every 1 mg/mL of RNA ~3 mM nucleotides → Net charge ~ -3 mM

Assuming cytoplasmic RNA is ~20 mg/mL → **Total net charge:** = -60 mM

Proteins:

Assuming typical proteins carry net charges similar to that of BSA. Table S2 of Ref² analyzes the net charges of the most abundant proteins in the mammalian cytoplasm, and shows most of them negatively charged.

Every BSA molecule carries -17 net charges at the cytoplasmic pH of 7.3.

For every 1 wt% BSA, $c = 0.15$ mM → Total net charge: -17×0.15 mM = -2.55 mM

Assuming cytoplasmic proteins is ~25 wt% → **Total net charge:** 25×-2.55 mM = -64 mM

Tally of above:

Positive charges: ~152 mM from small ions, mostly K^+ ; negligible from macromolecules.

Negative charges: ~20 mM from small ions, ~60 mM from RNA, and ~64 mM from proteins.

Significantly, the above counted total negative charges of ~144 mM largely balance the total positive charges, suggesting that the above analysis has accounted for the majority of charged species in the cytoplasm.

Table 4.2. Proteins used in this work, their Uniprot IDs, molecular weights (MWs), estimated net charges at pH 7.3, and sources. Net charges are estimated with Protein Calculator (<http://protcalc.sourceforge.net>), using the canonical “Chain” sequences (signal peptides removed) according to Uniprot. A shift of -1 net charge is assessed for samples labeled by Cy3B NHS ester. Whereas Cy3B NHS ester is charge-neutral, its conjugation to lysine removes one positive charge on the labeled protein and so shifts the net charge by -1 . Meanwhile, Cy5 NHS ester is $+1$ charged, and its conjugation to lysine does not alter the net charge of the labeled protein.

Protein	Uniprot ID	MW (kDa)	Charge @ pH 7.3	+Cy3B MW (kDa)	+Cy3B Charge	Provider	Cat#
Bovine serum albumin (BSA)	P02769	66.4	-17	67	-18	Sigma-Aldrich	A3059
Ovalbumin	P01012	42.7	-11.6	-	-	Sigma-Aldrich	A5503
α -Lactalbumin (LacA)	P00711	14.2	-7.2	14.8	-8.2	Sigma-Aldrich	L5385
Streptavidin (tetramer)	P22629	60	-3.6	61	-4.6	Sigma-Aldrich	S4762
RNase A	P61823	13.7	+3.9	14.3	+2.9	Sigma-Aldrich	R5500
Lysozyme (HEWL)	P00698	14.3	+7.5	14.9	+6.5	Sigma-Aldrich	L4919
Methylated BSA		~67	~+15 (ref ¹¹³)	-	-	Sigma-Aldrich	A1009
CXCL9	Q07325	11.7	+20	12.3	+19	PeptoTech	300-26
Avidin (tetramer)	P02701	68	+23	69	+22	Sigma-Aldrich	A9275
Histone H1.0	P07305	20.7	+53	21.3	+52	New England Biolabs	M2501S

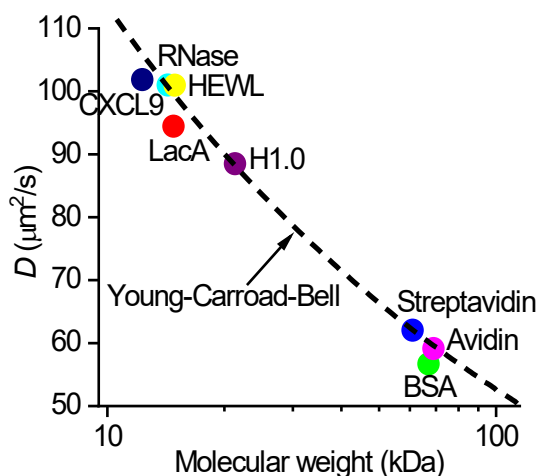


Figure 4.5. SMdM-measured D values of the Cy3B-labeled proteins in Fig. 4.4 in the starting buffer, plotted against their molecular weights. Dashed line: expected values according to the Young–Carroad–Bell model.^{12, 42}

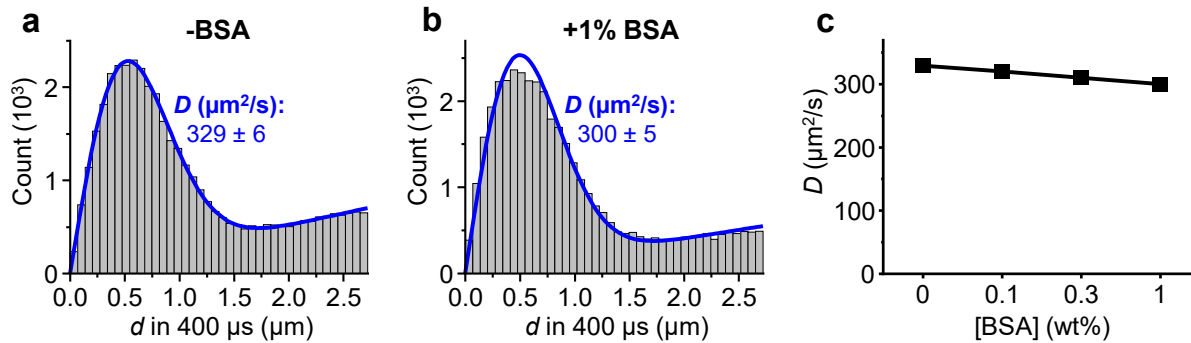


Figure 4.6. SMdM of the Cy3B dye with BSA addition. (a,b) Distributions of single-molecule displacements of Cy3B in a pH = 7.3 buffer of 2 mM ionic strength, with and without the addition of 1% BSA, from SMdM performed with a $\Delta t = 400 \mu\text{s}$ center-to-center separation between paired pulses. Blue lines: fits to the SMdM diffusion model, yielding diffusion coefficients D of 329 ± 6 and $300 \pm 5 \mu\text{m}^2/\text{s}$ (95% confidence intervals), respectively. (c) SMdM-measured D values at different wt% of added BSA.

Chapter 5: Single-molecule diffusivity quantification in *Xenopus laevis* egg extracts elucidates cytoplasmic molecular environments

The work in this chapter was conducted by A.A. Choi in collaboration with Coral Y. Zhou, Ayana Tabo, Rebecca Heald* and Ke Xu*. Permission was received from all co-authors.

5.1. Introduction

The living cell creates a unique molecular environment within its micrometer-sized bound volume^{104, 105, 117, 118}. This milieu is notably difficult to recapitulate *in vitro*: Simple buffers often do not maintain the native protein interactions or achieve the optimal conditions for activities seen inside the cell. While macromolecular crowding plays a role, adding inert crowders like dextrans does not recover cellular performance. A precise mix of proteins, nucleic acids, peptides, as well as small molecules and ions at the correct concentrations appears necessary to attain the magic of life. How this myriad of molecular species interacts with each other remains a mystery.

Diffusion measurements provide a valuable window into molecular environments and interactions^{55, 56, 59}. With the development of single-molecule displacement/diffusivity mapping (SMdM), a high-precision diffusion quantification method based on high-throughput single-molecule microscopy and statistics^{2, 119}, we have recently reported that for fluorescent proteins (FPs) expressed in mammalian cells, the diffusion of positively, but not negatively, charged species is substantially impeded in diverse cytoplasmic and intraorganellar environments^{2, 92}. Together with other recent microscopy and NMR studies noting reduced diffusivities of positively charged proteins in bacterial and mammalian cells⁹³⁻⁹⁷, these results point to an intriguing, charge sign-asymmetric intracellular environment.

Expression of FPs in the cell, however, offers limited control over the types and amounts of diffusers that can be examined, and nuances in diffusivity may be masked by intracellular structures and cell-to-cell variations. We have thus also examined protein interactions in solution, and demonstrated that simple negatively charged proteins like bovine serum albumin (BSA)

effectively impede the diffusion of positively charged proteins in an ionic strength-dependent fashion ¹²⁰. While easy to manipulate, in-solution experiments are limited by their simplicity and artificial nature.

Xenopus laevis egg extracts present an opportunity to bridge the above two limits. Prepared as homogenized, undiluted cytoplasm, they stand as an excellent model that retains the native intracellular environment ¹²¹⁻¹²³. Many complex cellular processes have been recapitulated and/or studied in this system, including mitotic spindle formation, chromosome condensation, cell-like self-organization, and apoptosis ^{121, 123-128}. Several recent studies have highlighted the use of *Xenopus* egg extracts to interrogate the biophysical properties of the cytoplasm ¹²⁹⁻¹³¹, in which molecular diffusion plays key roles.

Capitalizing on the versatility and manipulability of *Xenopus laevis* egg extracts and the high-precision diffusivity quantification of SMdM, here we examine how dye-tagged proteins, in well-defined low-concentration limits, diffuse in the complex cytoplasmic molecular environment. The freedom to introduce different proteins and control the ionic strength of the sample allows us to first establish size trends and elucidate charge effects. The facile introduction of ribonucleases then enables us to disintegrate RNA and ribosomes to both isolate their contributions to charge interactions and manifest unexpected aggregation mechanisms due to the liberation of ribosomal proteins. Finally, by inhibiting or enhancing actin polymerization, we underscore its key role in the size-dependent suppression of diffusion in the cytoplasm.

5.2 Results and Discussion

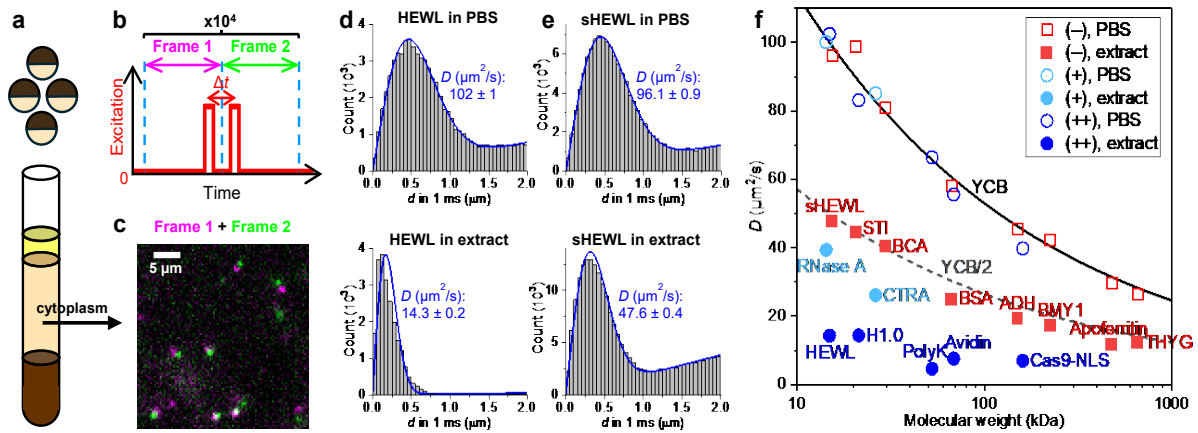


Figure 5.1. SMdM-based protein diffusivity quantification in the *Xenopus* egg extract versus in PBS. (a) Schematic: Cytoplasmic extracts from *Xenopus* eggs. (b) Schematic: In SMdM, paired excitation pulses are repeatedly applied across tandem camera frames, so that transient single-molecule displacements are captured in the wide field for the time window defined by the separation between the paired pulses, Δt . (c) Example single-molecule images of Cy3B-labeled bovine carbonic anhydrase diffusing in the extract, shown as a magenta-green overlaid image for a tandem pair of frames at $\Delta t = 1$ ms. (d) Example distributions of SMdM-recorded 1-ms single-molecule displacements for ~ 200 pM Cy3B-labeled HEWL diffusing in the PBS (top) versus in the extract (bottom). Blue curves: fits to our single-mode diffusion model, with resultant apparent diffusion coefficients D and 95% confidence intervals marked in each plot. (e) Similar to (d), but for succinylated HEWL (sHEWL). (f) SMdM-determined D values for 15 proteins of varied sizes and charges (**Table S1**), in PBS (hollow symbols) and in the extract (filled symbols). Red squares: Negatively charged proteins. Blue circles: Positively charged proteins with $>+5$ net charges. Light-blue circles: Weakly positively charged proteins with $\sim +2$ net charges. Solid curve: Expected values in PBS at room temperature according to the Young–Carrood–Bell (YCB) model. Dashed curve: The YCB values divided by 2.

Xenopus laevis egg cytoplasm extracts (hereafter “extracts”) were prepared following standard protocols (**Fig. 5.1a** and Methods)¹²². Dye-labeled proteins (Methods and **Table 5.1**) were added into the extract sample at <1 nM, optimal for SMdM single-molecule detection and diffusion quantification while minimizing the possible interactions between the added molecules¹²⁰. For SMdM (Methods)^{2, 12}, an excitation laser illuminated ~ 3 μm into the sample for the wide-field recording of single-molecule images with an EM-CCD camera. The laser was repeatedly modulated as paired stroboscopic pulses across tandem camera frames at a fixed center-to-center separation of $\Delta t = 1$ ms (**Fig. 5.1b**), so that transient displacements in the Δt time window were detected for molecules that diffused into the field of view (**Fig. 5.1c**). The execution of $\sim 10^4$ frame pairs over ~ 3 min thus recorded $\sim 10^5$ transient single-molecule displacements. The accumulated displacements were either spatially binned for local statistics^{2, 9} to generate a diffusivity map (**Fig. 5.1**), or pooled for global fitting to yield a diffusion coefficient D with an $\sim \pm 1\%$ bracket at 95% confidence (**Fig. 5.1d,e**)¹².

Figure 5.1d compares SMdM results of Cy3B-tagged hen egg white lysozyme (HEWL) in the phosphate-buffered saline (PBS) versus in the extract. With a net charge of $\sim +7$ at physiological pH, the ~ 15 kDa protein serves as a model to examine charge interactions¹²⁰. In PBS, the SMdM-recorded single-molecule displacements were well-fitted to our single-mode diffusion model to yield $D = 102 \mu\text{m}^2/\text{s}$ (**Fig. 5.1d**), as expected^{12, 120}. In contrast, markedly suppressed displacements were observed in the extract, which conformed less well to the single-mode model and yielded a substantially lower apparent D value of $14 \mu\text{m}^2/\text{s}$ (**Fig. 5.1d**). Spatial mapping of D showed no noticeable features (**Fig. 5.5**), expected for the homogenized extract, and suggests the above-

observed deviation from the single-mode diffusion model was due not to spatial variations but likely to heterogeneous transient interactions with extract components. To understand if the positive net charge of HEWL drove the above behavior, we succinylated HEWL to inverse its net charge to negative (Methods and **Table 5.1**)^{114, 120}. SMdM thus yielded a notably higher D of 48 $\mu\text{m}^2/\text{s}$ for succinylated HEWL (sHEWL) in the extract with a good fit to the single-mode diffusion model (**Fig. 1e**).

To establish a trend, we quantified the diffusivity of 15 proteins of diverse sizes and origins (**Table 5.1**) in PBS and in the extract. As we plotted all D values in PBS (hollow symbols in **Fig. 5.1f**) against the protein molecular weight, we noted a monotonic decrease in agreement with the Young-Carroad-Bell (YCB) model (solid curve in **Fig. 5.1f**) that relates the expected protein diffusivity to the molecule size and medium viscosity⁴², consistent with our previous results^{11, 12}.

For diffusion in the extract, the SMdM-determined D values of the 8 negatively charged proteins (filled red squares in **Fig. 5.1f**) showed a good size dependence that roughly followed (or slightly lower than) 50% of the YCB values (dashed curve in **Fig. 5.1f**). An early NMR study has examined the apparent diffusion coefficients of small molecules in *Xenopus* oocytes and shown a uniform 45% scaling versus that in water¹³². Thus, our results on negatively charged proteins are consistent with a universal diffusion slowdown due to an ~ 2 -fold higher viscosity of the *Xenopus* oocyte cytoplasm over water.

In comparison, the 5 substantially positively charged ($>+5$ net charge; **Table 5.1**) proteins consistently exhibited drastically reduced diffusivities (filled blue circles in **Fig. 5.1f**) far below the 50% YCB values (hence the diffusivities of negatively charged proteins). Meanwhile, the two weakly positively charged ($\sim +2$ net charge; **Table 5.1**) proteins displayed smaller, yet still noted deviations from the 50% YCB values (filled light-blue circles in **Fig. 5.1f**).

When plotted as a function of the protein net charge, the SMdM-determined D values in the extract relative to that in PBS showed a good trend (**Fig. 5.2a**): Whereas the negatively charged proteins all diffused at 40-50% of their diffusivities in PBS, proteins carrying $>+5$ net charges diffused at 10-20% of their PBS diffusivities, and the weakly positively charged proteins exhibited intermediate results.

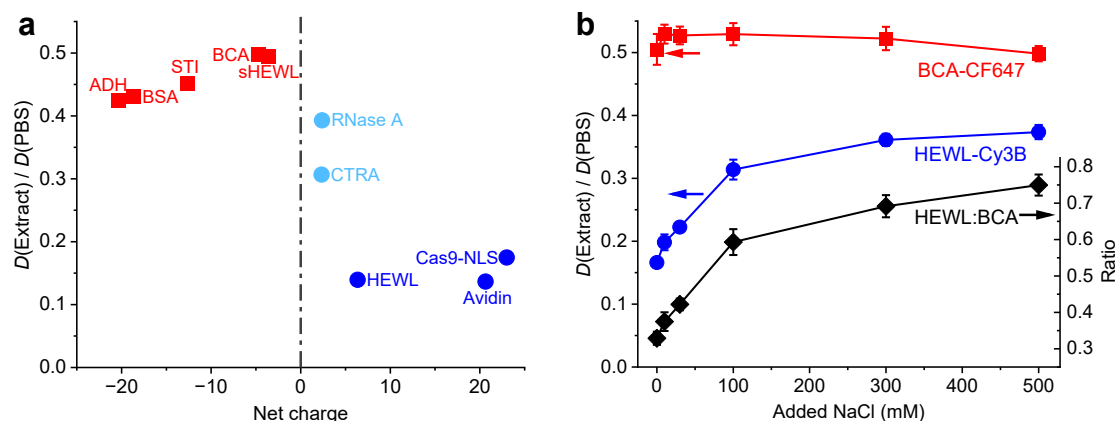


Figure 5.2. Net-charge effects in protein diffusion in the *Xenopus* egg extract. **(a)** SMdM-determined diffusion coefficients D in the extract relative to those in PBS for different proteins, as a function of their net charges in the range of -25 to +25. **(b)** Blue and red symbols: Relative D values in the extract over in PBS as a function of added NaCl, for the positively charged HEWL-Cy3B (blue) and the negatively charged BCA-CF647 (red), measured in the same sample through sequential SMdM in two color channels. Error bars: Sample standard deviations between results from two or three SMdM measurements at each data point. Black diamonds (y-axis on the right): Ratio between the PBS-normalized D values of the two proteins.

To further substantiate charge interactions as the drive behind this behavior, we examined ionic-screening effects by adding salts. To directly compare positively and negatively charged proteins in the same extract samples, we separately labeled HEWL and bovine carbonic anhydrase (BCA) with Cy3B and CF647 and performed SMdM in two color channels.

We thus found that for samples added with increasing amounts of NaCl, the positively charged HEWL (blue circles in **Fig. 5.2b**) progressively recovered its diffusivity from 16% to 37% of the PBS value. In comparison, the negatively charged BCA (red squares in **Fig. 5.2b**) stayed 50-53% of its PBS value. Assuming the remaining minor variations in the BCA data reflected non-charge effects, *e.g.*, mechanical disturbances from NaCl addition, the ratio between the PBS-normalized D values of HEWL and BCA (black diamonds in **Fig. 5.2b**) showed consistent increases with NaCl addition down to the initial 10 mM addition, and continued to rise at 500 mM added NaCl to 75%. Thus, ionic screening efficiently reduces the likely dynamic interactions between the positively charged HEWL and the negatively charged species in the extract, yet the interactions are still not fully eliminated at 500 mM ionic strength, reminiscent of our recent observations on HEWL-BSA interactions in solution¹²⁰.

Together, our SMdM results in the extract are consistent with a predominately negatively charged macromolecule environment that biasedly impedes positively charged diffusers. A

previous study attributes suppressed diffusion of positively charged proteins in bacteria to interaction with the negatively charged ribosomes⁹³. With our SMdM results in mammalian cells, we have evaluated the net charges of major proteins in the mammalian cytoplasm and organelles and noted the dominance of negatively charged species^{2, 92}. The negatively charged protein environment may drive the diffusion slowdown of positively charged proteins, which appears necessary to explain results in compartments devoid of ribosomes, *e.g.*, the endoplasmic-reticulum lumen⁹². With in-solution experiments, we have further shown that negatively charged proteins alone substantially impede the diffusion of positively charged proteins¹²⁰.

Here, we assessed the net charges of major proteins in the *Xenopus* egg cytoplasm based on mass spectrometry-detected abundances¹, and similarly found most to be negatively charged (**Table 5.2**). Although many ribosomal proteins possess high positive charges on their own (**Table 5.3**), their assembly with the highly negatively charged ribosomal RNA (rRNA) leads to negative net charges of the resultant ribosomes, hence the above model of charge-driven ribosome interactions⁹³. These findings echo our analysis of human cell lines^{2, 92}. Contrasting with cell lines, however, the extract provides a manipulatable system in which ribosomes, together with other RNA forms, can be readily disintegrated by incubating with ribonucleases (RNases)^{131, 133, 134}, hence a potential path to isolate their contributions.

Interestingly, we found that while RNase addition did not immediately alter the diffusion behavior of the positively charged HEWL in the extract (**Fig. 5.3a-b** vs. **Fig. 5.5**), micrometer-scale domains emerged in the SMdM diffusivity map after 3 h, where the local apparent D values dropped to $\sim 2 \mu\text{m}^2/\text{s}$ (magenta in **Fig. 5.3c**), suggesting the formation of insoluble aggregates. Moving away from microscopy, we further noted that RNase-treated extracts became visibly cloudy (**Fig. 5.3e**).

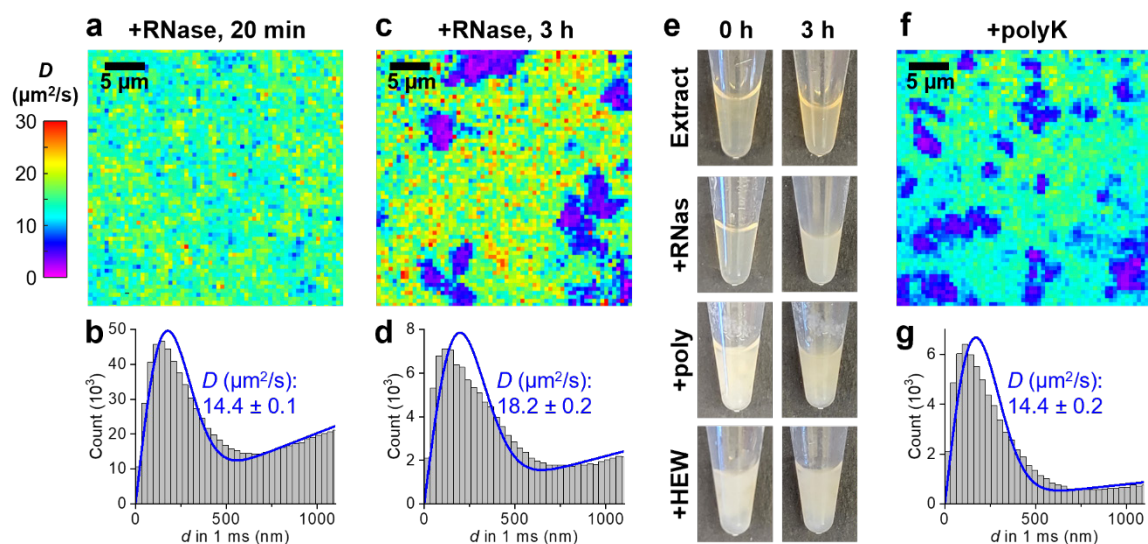


Figure 3. SMdM diffusivity mapping of RNase-treated egg extracts further underscores the charge-sign asymmetry of the cytoplasmic environment. (a) Color-coded SMdM diffusivity map of Cy3B-labeled HEWL in an extract sample 20 min after RNase addition. (b) Distributions of single-molecule displacements for the data in (a). Blue line: fit to the SMdM diffusion model, with resultant apparent diffusion coefficient D and 95% confidence intervals marked. (c) SMdM diffusivity map of the same sample after 3 h. (d) Distributions of single-molecule displacements for regions outside the low-diffusivity domains in (c). (e) Photos for extract samples taken immediately (left) and after 3 h (right), with or without the addition of RNase or 1 mg/mL of polylysine or HEWL. (f) Color-coded SMdM D map of Cy3B-labeled HEWL in an extract sample added with 1 mg/mL polylysine. (g) Distribution of single-molecule displacements for regions outside the low-diffusivity domains in (f).

A review of past studies indicated that RNase treatments of mammalian and bacterial lysates cause widespread protein aggregation^{135, 136}. Earlier work has separately shown that RNase treatments of *Xenopus* egg extracts abolish cellular processes^{133, 134}. While these observations have been attributed to translation-independent RNA functions, our starting point of ribosome interactions inspires alternative interpretations.

The different ribosomal protein forms in the *Xenopus* egg extract amount to $\sim 250 \mu\text{M}$ ¹, translating to $\sim 4 \text{ mg/mL}$. As another estimation, the $\sim 0.7 \mu\text{L}$ *Xenopus* egg cytoplasm contains $\sim 5 \mu\text{g}$ rRNA^{137, 138}, suggesting $\sim 5 \text{ mg/mL}$ ribosomal proteins given their typical ratios versus rRNA. The RNase degradation of rRNA thus releases up to 5 mg/mL ribosomal proteins, many highly positively charged (**Table 5.3**), into the negatively charged cytoplasmic background. This substantial mix of oppositely charged proteins may induce aggregation¹⁰⁷⁻¹⁰⁹.

To examine this hypothesis, we added positively charged proteins into untreated extracts. Notably, the addition of 1 mg/mL of polylysine and HEWL both induced immediate clouding of

the extract (**Fig. 5.3e**) and micrometer-sized low-diffusivity domains in the SMdM diffusivity map (**Fig. 5.3f**). In comparison, adding negatively charged BSA to the extract did not induce aggregation (**Fig. 5.5**). A recent study shows that introducing cationic “killer peptides” into cell lysates induces aggregation¹³⁹, echoing earlier findings in bacteria¹¹¹. Our results suggest that in the cytoplasmic environment, substantial addition or liberation of positively charged proteins, regardless of specific forms, both prompt aggregation. The former scenario is fast due to immediate charge interactions whereas the latter is slow, pending the graduate release of positively charged proteins.

Beyond the difference in aggregation speed, the RNase-treated extract further showed a mild increase in D of HEWL from ~ 14 to $\sim 18 \mu\text{m}^2/\text{s}$ for regions outside the low-diffusivity aggregates (compare **Fig. 5.3d** to **Fig. 5.3b,g**). This reduced suppression of diffusion may be attributed to diminished impediments from ribosomes and RNA. Yet, the recovery is far from complete given the above $D = 48 \mu\text{m}^2/\text{s}$ of sHEWL in the extract, suggesting that the predominately negatively charged protein environment still suppressed the diffusion of the positively charged HEWL¹²⁰.

We next revisit the above-noted relatively invariant 40-50% scaling of D in the extract versus in PBS for the different negatively charged proteins (**Fig. 5.2a**). While this observation simplified our comparison with positively charged diffusers, it conflicts with the notion that intracellular diffusion is generally more hindered for larger molecules^{56, 60, 140, 141}. To reconcile this issue, we noted that in the standard extract preparation protocol, cytochalasin is added to inhibit actin polymerization¹²². It has been shown that the size-dependent DNA mobility in mammalian cells depends on the actin cytoskeleton, and that in solutions, the addition of 8 mg/mL polymerized actin recapitulate size-dependent diffusion slowdown whereas soluble crowding agents, including cytosol extracts, do not¹⁴⁰. Our recent SMdM results in expandable hydrogels have also shown that obstruction due to immobile meshworks is vital in the size-dependent suppression of molecular diffusivity¹⁴².

To probe the likely role of the actin cytoskeleton in size-dependent diffusion suppression, we prepared actin-intact *Xenopus* egg extracts in which cytochalasin was omitted, as well as actin-supplemented samples to which extra actin was added at 5 mg/mL. Chemically fixing the samples

for phalloidin labeling¹⁴³ and three-dimensional stochastic optical reconstruction microscopy (3D-STORM)^{6, 144} super-resolution imaging resolved dense actin networks (**Fig. 5.4a**) analogous to those seen in animal cells¹⁴³. In comparison, no phalloidin staining was observed in the standard actin-inhibited samples (not shown).

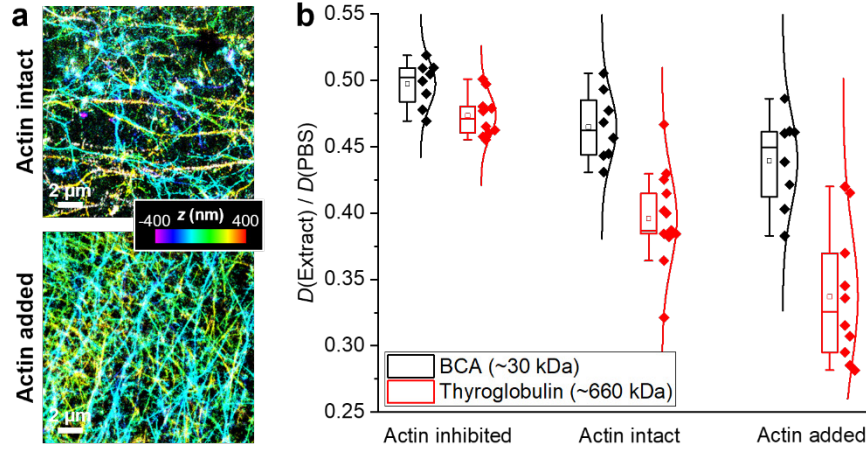


Figure 4. Importance of the actin cytoskeletal network in molecular size-dependent diffusion suppression. **(a)** 3D-STORM super-resolution images of phalloidin-labeled actin filaments in actin-preserved (top) and actin-supplemented (bottom) *Xenopus* egg extracts. Color presents axial (depth) information. **(b)** SMdM-determined D values in the extract relative to in PBS, for the ~30 kDa BCA (black) and the ~660 kDa thyroglobulin (red), in actin-inhibited, actin-intact, and actin-supplemented samples. Each data point corresponds to one independent SMdM measurement for a different sample region, from ~3 samples under each condition.

SMdM thus showed that whereas in the actin-inhibited extract, the 660 kDa thyroglobulin exhibited similar scaling of its D value over in PBS when compared to the 30 kDa BCA (47% vs. 50%), it experienced progressively stronger suppression in diffusivity in the actin-intact and actin-supplemented extracts, reaching 39% and 32% of the PBS value, respectively (**Fig. 5.4b**). In comparison, the 30 kDa BCA displayed notably smaller drops in D to reach 46% and 43% of the PBS value in the actin-intact and actin-supplemented extracts, respectively (**Fig. 5.4b**). Increased variations in the measured D values were also noted in the presence of actin (**Fig. 5.4b**), likely related to the observed spatial inhomogeneity in local actin density (**Fig. 5.4a**). These results underscore the key role of actin cytoskeleton in the size-dependent suppression of diffusion in the cytoplasm.

5.3 Conclusion

The complex intracellular environment has been difficult to elucidate, with nuances in molecular behaviors readily masked by subcellular and cell-to-cell heterogeneities. *Xenopus* egg extracts offer a homogenized, near-native cytoplasm system that can be closely examined and readily manipulated. Meanwhile, through high-throughput single-molecule statistics, SMdM quantifies diffusivity with $\sim 1\%$ precision at sub-nM diffuser concentrations. Integration of SMdM with the extract system thus provides a powerful path toward understanding molecular interactions in the cytoplasm environment.

Starting with 15 proteins of diverse properties and origins, we first found that whereas the in-extract diffusivities of the negatively charged species displayed a clear size trend $\sim 50\%$ of their in-PBS values, the positively charged species diffused substantially slower. While the former observation implies that the diffusion of negatively charged proteins is simply scaled by an $\sim 2\times$ cytoplasm viscosity over water, the latter evidences further diffusion suppression owing to charge-based interactions. Notably, increasing the sample ionic strength effectively alleviated, but did not fully eliminate, this charge-based suppression. Examination of the mass spectrometry-ranked major proteins showed most as negatively charged; while ribosomal proteins tend to be positively charged, their assembly with rRNA results in negatively charged ribosomes. A predominately negatively charged macromolecule environment thus likely drove the observed biased diffusion suppression for positively charged proteins.

To isolate the contributions from ribosomes and RNA, we next incubated the extract with RNase. Interestingly, while we observed a moderate reduction in diffusion suppression, the more striking results were the gradual formation of aggregates. Calculations indicated that the breakdown of ribosomes releases >1 mg/mL positively charged proteins in the extract; correspondingly, we showed that adding 1 mg/mL positively charged proteins into the extract induced immediate aggregation. Besides further manifesting the importance of maintaining a negatively charged macromolecular environment in the cytoplasm, these findings call for reexaminations of previous results involving RNase treatments.

To answer why our results above did not yield size-dependent suppression of diffusion, we noted that actin polymerization was inhibited in the typical extract preparation protocols. With actin-intact and actin-supplemented extract samples, we thus showed that with enhanced actin

filament levels, the diffusion of a large protein was progressively more suppressed than a small protein, recapitulating the typical trend in cells.

Together, our results shed new light on how differently charged and sized proteins interact with the complex cytoplasm environment. Besides fundamental insights into the inner workings of the cell, such understandings further carry implications for how the extracts may be best utilized as model systems. More generally, the findings that the cytoplasm provides both a dominantly negatively charged macromolecular environment and crosslinked meshworks may prove important in designing *in vitro* systems to recapitulate cellular performances. The application of the methodologies developed in this work to resolve diffusion patterns in diverse cytoplasmic processes that can be initiated in the extract system, as well as in extracts from *Xenopus* embryos at different development stages, represents exciting future challenges.

5.4 Material and Methods

***Xenopus* egg cytoplasm extracts.** *Xenopus* egg cytoplasm extracts were prepared following standard protocols¹²². Briefly, eggs from *Xenopus laevis* were centrifuged at 10,000 rpm for 15 min to be crushed. The cytoplasm was retrieved and supplemented with 10 µg/mL LPC protease inhibitors, 20 µg/mL cytochalasin B, and 1x energy mix (4 mM creatine phosphate, 0.5 mM ATP, 0.05 mM EGTA, and 0.5 mM MgCl₂). Actin-intact extracts were similarly prepared but excluded cytochalasin B. Actin-supplemented extracts were prepared by adding actin from rabbit skeletal muscle (Cytoskeleton, Inc. AKL95-B) to the actin-intact extract to 5 mg/mL. For RNase treatment, Ribonuclease A from bovine pancreas (Sigma R5500) was added to the extract to 1 mg/mL. For the addition of different proteins, poly(D,L-lysine hydrobromide)₂₅₀ (Alamanda Polymers 000-RKB250), HEWL (Sigma L4919), or BSA (Sigma A3059) were separately added to the extract to 1 mg/mL.

Dye-labeling of proteins. Sources of proteins used are shown in **Table 5.1**. The proteins were labeled with Cy3B NHS (*N*-hydroxysuccinimidyl) ester (Cytiva PA63101) or CF647 NHS ester (Biotium 92135) in 0.1 M NaHCO₃ at an ~1:1 initial dye:protein ratio for 1 h. Unconjugated dye was removed by filtering through Amicon centrifugal filters six times. Absorbances at 280 and 560 nm, as measured by a NanoDrop 2000c spectrometer (ThermoFisher), indicated that the final product had ~0.5 dyes per protein on average. Thus, the fluorescently detected molecules in SMdM typically had only one dye on the protein, while the unlabeled fraction of the protein was undetected. Succinylation of Cy3B-labeled HEWL was done by adding excessive solid succinic anhydride (Fisher Scientific AC158760050) into the sample¹¹⁶ and then cleaning up with Amicon centrifugal filters.

Imaging devices. #1.5 Glass coverslips were acid-treated and passivated with 10 mg/mL methoxy-PEG silane (MW 5000, PG1-SL-5k, Nanocs) in 95% ethanol/water for 30 min, and then rinsed and sonicated for 5 min in Milli-Q water¹². The coverslips were then each mounted with a plastic tube (cut from a 0.65 mL microcentrifuge tube) to form an imaging chamber⁸³. For SMdM, ~100 μ L of extract or PBS, with dye-labeled proteins added at 0.2-1.0 nM, were added into the imaging chamber. We separately compared imaging chambers with Bioinert surfaces (ibidi 80800), which have been used in recent studies¹²⁹, and obtained indistinguishable results.

SMdM. SMdM was performed on a Nikon Ti-E inverted fluorescence microscope, as described previously^{2, 12}. Briefly, a 561 nm and a 642 nm laser were focused at the edge of the back focal plane of an oil-immersion objective lens to illuminate a few micrometers into the sample. The focal plane was maintained at ~3 μ m into the sample, and single-molecule images were recorded in the wide field with an EM-CCD camera (iXon Ultra 897, Andor) that operated continuously at 110 frames per second. The laser was repeatedly modulated as paired stroboscopic pulses across tandem camera frames at a fixed center-to-center separation of $\Delta t = 1$ ms (**Fig. 1b**). The recorded images across the pair frames thus captured the transient single-molecule displacements in the $\Delta t = 1$ ms time window. Typical runs executed 1.0×10^4 frame pairs over ~3 min, accumulating $0.5\text{-}5 \times 10^5$ single-molecule displacements. These displacements were pooled for global fitting or spatially binned with a grid size of 500 nm for local fitting to the probability model:

$$P(r) = \frac{2r}{a} \exp\left(-\frac{r^2}{a}\right) + br \quad (\text{Eqn. 1})$$

Here $a = 4D\Delta t$ with D being the diffusion coefficient, and b accounts for a uniform background due to extraneous molecules that randomly diffuse into the search radius during Δt ². For spatially binned data, the fitted D value in each spatial bin was assigned a color on the continuous D scale for rendering into a color-coded D map.

3D-STORM imaging of actin filaments. For 3D-STORM imaging of actin filaments¹⁴³ in the actin-intact and actin-added extracts, 100 μ L of the extract was added into an 8-well glass coverslip chamber (ibidi 80807) and incubated for 45 min at room temperature. The sample was fixed with 2% glutaraldehyde in the cytoskeleton buffer (10 mM MES, pH = 6.1, 150 mM NaCl, 5 mM EGTA, 5 mM glucose, and 5 mM MgCl₂) for 30 min, and then reduced using freshly made 0.1% (w/v) NaBH₄ in PBS. Actin filaments were stained by 400 nM phalloidin Alexa Fluor 647 (Cell Signaling Technology 8940S) overnight at 4 C. The sample was washed with PBS and mounted in a STORM imaging buffer containing 100 mM Tris-HCl, pH = 7.5, 100 mM cysteamine, 5% (w/v) glucose, 0.8 mg/mL glucose oxidase (Sigma-Aldrich, G2133), and 40 μ g/mL catalase (Sigma-Aldrich, C30). 3D-STORM was performed as described previously^{6, 144, 145}.

Table 5.1. Diffuser proteins used in this work, their Uniprot IDs, molecular weights (MWs), estimated net charges, and sources. Net charges are estimated using the canonical “Chain” sequences (signal peptides removed) according to Uniprot. Charge estimation is performed with Prot pi (<https://www.protpi.ch/>) for pH 7.7 (typical pH of the *Xenopus* egg cytoplasm and extract ¹²²) using the ExPASy isoelectric point values, with disulfide bonds applied globally. A shift of -1 net charge is assessed for samples labeled by Cy3B NHS ester. Whereas Cy3B NHS ester is charge-neutral, its conjugation to lysine removes one positive charge on the labeled protein and so shifts the net charge by -1 .

Label in figures	Protein	Uniprot ID	MW (kDa)	Charge @pH7.7	+Cy3B MW(kDa)	+Cy3B Charge	Provider	Cat#
Apo ferritin	Apo ferritin, equine (24-mer)	P02791 Q8MIP0	480	-235	480	-236	Sigma	A3660
THYG	Thyroglobulin, bovine (dimer)	P01267	660	-83	660	-84	Sigma	T1001
BM Y1	Beta amylase, sweet potato (tetramer)	P10537	224	-59	224	-60	Sigma	A8781
ADH	Alcohol dehydrogenase, yeast (tetramer)	P00330	150	-19.4	151	-20.4	Sigma	A8656
BSA	Bovine serum albumin	P02769	66.4	-17.7	67	-18.7	Sigma	A3059
STI	Soybean trypsin inhibitor	P01070	20.0	-11.6	20.6	-12.6	Sigma	T2327
BCA	Bovine carbonic anhydrase	P00921	29.1	-3.8	29.7	-4.8	Sigma	C2624
sHEWL*	Succinylated HEWL	-	14.8	-2.6	15.4	-3.6	From HEWL	From HEWL
CTRA	Chymotrypsinogen A, bovine	P00766	25.7	3.3	26.3	2.3	Sigma	C4879
RNase A	RNase A, bovine	P61823	13.7	3.4	14.3	2.4	Sigma	R5500
HEWL	Lysozyme from	P00698	14.3	7.4	14.9	6.4	Sigma	L4919

	chicken egg white							
Avidin	Avidin from chicken egg white (tetramer)	P02701	68	21.6	68.6	20.6	Sigma	A9275
Cas9-NLS[#]	Cas9, recombinant, with 1X NLS	Q99ZW2	160	24.0	161	23.0	Sigma	CAS9PROT
H1.0	Histone H1.0, human, recombinant	P07305	20.7	52	21.3	51	NEB	M2501S
PolyK	poly(D,L-lysine hydrobromide) ₂₅₀	-	52	250	52.6	250	Alamanda Polymers	000-RKB250

*sHEWL: Assuming Cy3B NHS ester conjugated to one of the 6 lysine residues, while succinylation of the remaining 5 lysine residues inversed five +1 charges to -1 ¹¹⁴. This still leaves one N-terminus amine that may be succinylated.

[#]Cas9-NLS: Based on the sequence of Cas9 plus the +5 charged typical nuclear localization sequence (NLS) PKKKRKV.

Table 5.2. List of top cytoplasmic proteins in the *Xenopus* egg cytoplasm extract and their estimated net charge q . Cytoplasm proteins (not including RNA-binding and ribosomal proteins in Table S3) are ordered by their estimated concentrations c in the extract according to the mass-spectrometry data of ¹, based on updated annotations by Wühr Lab (<https://wuehr.scholar.princeton.edu/protein-concentrations-xenopus-egg>). Net charges at pH 7.7 are estimated for the annotated sequences similarly as described in Table S1. We thus note that most proteins in this top list are negatively charged or neutral. A few weakly positively charged (+1 to +3) proteins are found for those involved in glycolysis (GAPDH and AldOa). This behavior is similar to what we previously noticed in the human cytoplasm ²: The weak positive net charge may be intended for their binding to the negatively charged, phosphorylated glucose metabolites, which thereby neutralizes the total charge.

<i>Laevis</i> gene	Human gene	Protein description	c (μ M)	Net q
serpina6.L	SERPINA1	Alpha-1-antitrypsin	8.7	0.2
actc1.L	ACTC1	Actin, alpha cardiac muscle 1	8.0	-13.8
eno1.L	ENO1	Alpha-enolase	7.9	-6.9
cgl.2.L	ITLN2	Intelectin-2	6.3	-7.7
hbz.L	HBZ	Hemoglobin subunit zeta	5.9	-4.8
actg1.S	ACTG1	Actin, cytoplasmic 2	5.8	-12.8
gapdh.L	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	5.8	2.1
tpi1.S	TPI1	Triosephosphate isomerase	5.8	-0.9
ldhb.S	LDHB	L-lactate dehydrogenase B chain	5.5	-6.6
nme2.S	NME2	Nucleoside diphosphate kinase B	5.3	-1.9
fabp4	FABP4	Fatty acid-binding protein, adipocyte	4.6	-1.9
rps27a.S*	RPS27A	Ubiquitin	4.4	-0.9
pkm.L	PKM	Pyruvate kinase PKM	4.2	-8.8
plin2.L	PLIN2	Perilipin-2	4.2	-16.9
grhpr.2.L	GRHPR	Glyoxylate reductase/hydroxypyruvate reductase	4.1	-4.9
nme2.L	NME2	Nucleoside diphosphate kinase B	4.1	-1.9
mdh1.S	MDH1	Malate dehydrogenase, cytoplasmic	4.1	-1.9
hba-l5.L	HBZ	Hemoglobin subunit zeta	4.0	-4.8
ckb.S	CKB	Creatine kinase B-type	3.9	-7.8
pgk1.L	PGK1	Phosphoglycerate kinase 1	3.9	-1.9
gapdh.S	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	3.9	1.1
aldoa.S	ALDOA	Fructose-bisphosphate aldolase A	3.7	3.2
gstml.S	GSTM1	Glutathione S-transferase Mu 1	3.7	0.1
hpgds.S	HPGDS	Hematopoietic prostaglandin D synthase	3.6	-2.0
ahcy.S	AHCY	Adenosylhomocysteinase	3.6	-6.9
ftmt.S	FTH1	Ferritin heavy chain	3.5	-11.7
gpi.L	GPI	Glucose-6-phosphate isomerase	3.4	-2.7
tpi1.L	TPI1	Triosephosphate isomerase	3.3	-1.9

* rps27a.S (RPS27A) encodes both ubiquitin and the 40S ribosomal protein S27a. The charge here is calculated for ubiquitin.

Table 5.3. List of top RNA-binding and ribosomal proteins in the *Xenopus* egg cytoplasm extract as detected by mass spectrometry ¹, and their estimated net charges, generated similarly as Table S2. Note that when assembled with rRNA into ribosomes, the heavy negative charges on rRNA (−1 net charge per nucleotide to the RNA backbone) overcompensate the protein charges to render the assembled ribosomes highly negatively charged ^{93, 146}.

<i>Laewis</i> gene	Human gene	Protein description	<i>c</i> (μM)	Net <i>q</i>
rpl32.L	RPL32	60S ribosomal protein L32	5.0	25.1
rpl23a.S	RPL23A	60S ribosomal protein L23a (Fragment)	4.7	25.1
rps27a.S*	RPS27A	40S ribosomal protein S27a	4.4	15.4
cirbp.S	CIRBP	Cold-inducible RNA-binding protein	4.1	2.1
rps14.L	RPS14	40S ribosomal protein S14	4.1	12.2
rps9	RPS9	40S ribosomal protein S9	4.0	21.1
rpl27.S	RPL27	60S ribosomal protein L27	3.9	24.0
rpl7a.S	RPL7A	60S ribosomal protein L7a	3.8	40.0
rps18	RPS18	40S ribosomal protein S18	3.8	33.1
rps16.S	RPS16	40S ribosomal protein S16	3.8	16.1
rps3.L	RPS3	40S ribosomal protein S3	3.7	11.1
rps8.L	RPS8	40S ribosomal protein S8	3.7	33.1
rps28p9.L	RPS28	40S ribosomal protein S28	3.6	4.2
LOC398653	RPLP2	60S acidic ribosomal protein P2	3.5	-8.9
rplp0.S	RPLP0	60S acidic ribosomal protein P0	3.5	-6.9
rps4x.L	RPS4X	40S ribosomal protein S4, X isoform	3.4	23.2
rps12.L	RPS12	40S ribosomal protein S12	3.4	-0.9
rpl27a.S	RPL27A	60S ribosomal protein L27a	3.4	22.3
rpl26.L	RPL26	60S ribosomal protein L26	3.3	24.1
rpl22.L	RPL22	60S ribosomal protein L22	3.3	6.1
rpl8.L	RPL8	60S ribosomal protein L8	3.3	36.2
rps19.S	RPS19	40S ribosomal protein S19	3.3	16.2
rps17.S	RPS17	40S ribosomal protein S17	3.3	8.1
rps25.L	RPS25	40S ribosomal protein S25	3.3	20.0
rps11	RPS11	40S ribosomal protein S11	3.1	22.1
rpl19	RPL19	60S ribosomal protein L19	3.1	43.1
rpl7.L	RPL7	60S ribosomal protein L7	3.1	41.0
rps6.L	RPS6	40S ribosomal protein S6	3.0	40.0
rpl4.S	RPL4	60S ribosomal protein L4	3.0	63.0

* rps27a.S (RPS27A) encodes both ubiquitin and the 40S ribosomal protein S27a. The charge here is calculated for the 40S ribosomal protein S27a.

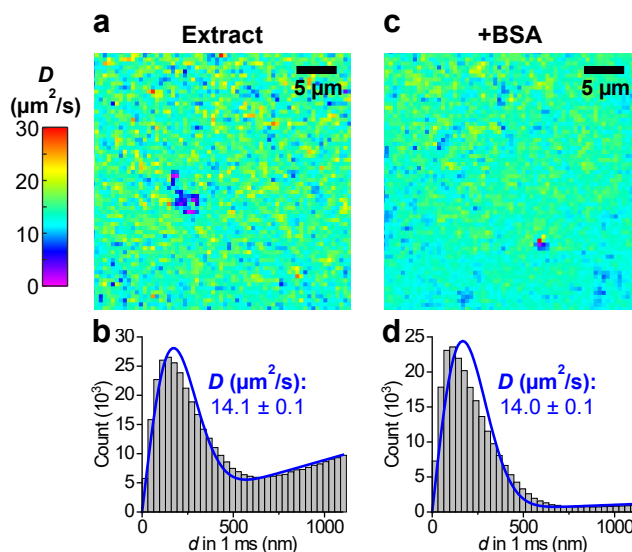


Figure 5.5. SMdM diffusivity mapping of Cy3B-labeled HEWL in untreated and BSA-added *Xenopus* egg extracts. (a) Color-coded SMdM diffusivity map for an untreated sample, presented on the same color scale as **Fig. 5.3**. (b) Distribution of the 1-ms single-molecule displacements from (a). Blue curve: fit to our single-mode SMdM diffusion model, with resultant apparent diffusion coefficients D and 95% confidence intervals marked. Note that in this sample, a higher concentration of Cy3B-labeled HEWL (~ 1 nM) was used to increase the detected single-molecule density to facilitate spatial mapping. While this led to a higher background in the displacement distribution (when compared to **Fig. 5.3D**), the D value obtained with our fitting model (Eqn. 1) stays consistent, in agreement with what we have analyzed and demonstrated previously². (c) Color-coded SMdM diffusivity map for another sample to which BSA was added at 1 mg/mL. (d) Distribution of the 1-ms single-molecule displacements from (c) and fit to our model.

Chapter 6: Conclusion and Outlooks

6.1. Concluding Remarks

In previous chapters, I described the evolution of SMdM into a quantitative single-molecule diffusion measurement tool capable of measuring a variety of analytes ranging from proteins to small-molecules in living cells, providing new insight into how physical parameters such as molecular size and charge are shaped by the intracellular space. Here, I leveraged the sensitivity of SMdM to observe formation of protein clusters formed *in situ* due to strong electrostatic interactions, elucidating the charge landscape of the intracellular cytoplasm, and how it is organized and structured by electrostatic interactions. In collaboration with Professor Rebecca Heald's group, I worked with the *Xenopus laevis* extract system to uncover biophysical laws pertaining to size and charge in a plasma membrane-less cytoplasm where we observed the importance of the intracellular actin cytoskeleton in inducing macromolecular crowding as well as the importance of RNAs in solubilizing positively charged proteins.

6.2 Outlooks

While this thesis primarily focuses on method development in SMdM, I will discuss promising directions in single-molecule microscopy. While SMdM has primarily been applied to characterize diffusivity of proteins. First, I will highlight some work exploring diffusion of the ribonucleoprotein (RNA-protein complexes) such as used in the CRISPR-Cas9 system to understand how the RNA in ribonucleoproteins shape the diffusive properties within the nucleus. Then, I will describe efforts made towards developing a novel functional super-resolution microscopy tool through the development of single-molecule *electrophoresis* microscope capable of identifying the net charge of single-molecules.

6.2.1 Live-Cell Single-Molecule Diffusion Mapping of Ribonucleoprotein

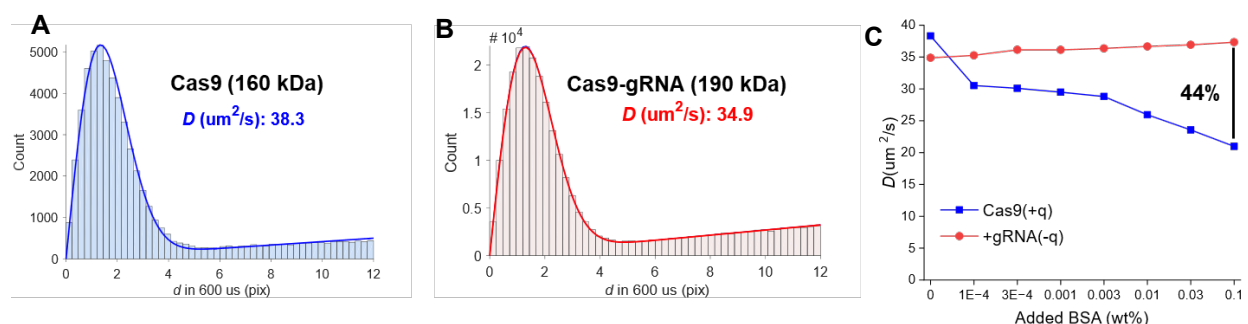


Figure 6.1 Single-molecule diffusivity quantification shows strong electrostatic attractions between Cas9 and BSA, but not for Cas9-gRNA and BSA. Example histograms of single-molecule displacements for 200 pM (A) Cas9-Cy3B and (B) Cas9-Cy3B-gRNA in pH = 7.3 buffer of ionic strength of 2 mM. (C) SMdM-determined diffusion coefficients D for Cy3B-labeled Cas9 with and without gRNA added at 200 pM to pH = 7.3 buffers of 2 mM ionic strength, as a function of wt% added BSA.

Many RNA- and DNA-binding proteins are positively charged, likely to facilitate binding to the highly negatively charged nucleic acids. However, this has implications on their intracellular diffusivity where positively charged, but not negatively charged protein slowdown substantially due to strong electrostatic interactions with the predominantly negatively charged macromolecules^{2, 92} present in intracellular environment. It came to our attention that Cas9, Cas13, and telomerase, while highly positively charged on their own, would become net negatively charged when forming ribonucleoprotein complexes, which may impact its diffusive properties. Based on our previous work on how net charge influenced diffusion dynamics in live cells, we hypothesized that this charge inversion would increase diffusivity, and would have important implications in organelles such as the nucleus where for proteins such as Cas9 utilizes 3D diffusion to identify genomic sequences in DNA that are complementary to its gRNA¹⁴⁷.

In collaboration with Chun Ying Wu, we first started by examining the diffusion of Cas9 in a pH = 7.3 of ionic strength of 2 mM to mimic the low ionic strength of the intracellular environment. By labeling Cas9 with the bright fluorophore, Cy3B, we were able to track freely diffusing single-Cas9-Cy3B protein using $\Delta t = 0.6$ ms spaced center-to-center from the end of the camera frame. Accumulation of $>10^4$ displacements yielded a displacement distribution that was well fit by our diffusion equation² and resulted in a $D = 38.3 \mu\text{m}^2/\text{s}$ (Fig 6.1A). In addition, the D coefficient agreed well with the Young-Carroad-Bell model for a protein of ~ 160 kDa, which relates the diffusion coefficient to the molecular weight of protein⁴². After validating that we could accurately measure the D coefficient of Cas9, we measured the D coefficient of the Cas9-gRNA complex by labeling Cas9 with Cy3B, and then mixing with nontargeting gRNA in excess. Similarly, using the same buffer composition as Cas9, we measured a $D = 34.9 \mu\text{m}^2/\text{s}$ (Fig 6.1B) for Cas9-gRNA-Cy3B, resulting in a moderately reduced D coefficient when compared to Cas9, but plausible given the ~ 30 kDa increase in molecular weight from the addition of gRNA to the Cas9-gRNA complex.

We then examined the influence of electrostatic interactions of the positively Cas9 and negatively charged Cas9-gRNA complex in a solution containing varying concentrations of the negatively charged bovine serum albumin (BSA). Using the same buffer as described above, we titrated in increasingly high concentrations of BSA and measured the D for both Cas9 and Cas9-gRNA labeled by Cy3B. While Cas9-gRNA maintained a $D \sim 35 \mu\text{m}^2/\text{s}$ even at 1% wt BSA concentrations, Cas9 had $\sim 44\%$ reduction in its overall D coefficient at 1% BSA with $D = 21.4 \mu\text{m}^2/\text{s}$, showing that while Cas9-gRNA was unaffected by even high concentrations of BSA, Cas9 alone was strongly suppressed in D as a result of additions of BSA. This substantiated the formation of ribonucleoprotein complex *in vivo* may reduce unspecific electrostatic interactions that slow diffusion, especially prevalent in the nucleus where there exist high high concentrations of negatively charged DNA and RNA.

After validating SMdM to measure the diffusivity of single-Cas9 protein in buffer conditions, we started cell experiments to study its diffusivity in the nuclei of live cells. We did

this by expressing dCas9-mEos4b using plasmid transfection containing a nuclear localization sequence (NLS) to localize dCas9-mEos4b to the nucleus and used a catalytically ‘dead’ Cas9 to prevent endonuclease activity when bound to gRNA. To spatially map the diffusivity of dCas9 in the nucleus using SMdM, we used 2 ms center-to-center separation excitation pulses, and spatially mapped single-molecule displacements by locally binning and fitting the diffusivity using 320 nm sized spatial bins to resolve nanometer sized spatial heterogeneities in diffusivity. Here, we showed that for dCas9-mEos4b, that vast regions of the nucleus had $D \sim 2 \mu\text{m}^2/\text{s}$ (Fig. 6.2A) and was generally spatially homogeneous in diffusivity except for regions near the nucleolus.

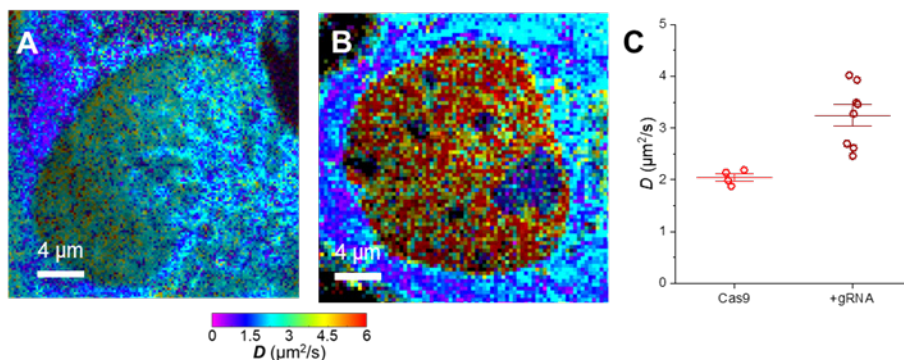


Figure 6.2. SMdM quantification of diffusion in the nucleus of Cas9-mEos4b of live COS-7 cells. Representative D spatial maps of (A) dCas9-mEos4b, (B) dCas9-mEos4b and nontargeting gRNA, in the nucleus of live COS-7 cells using 2 ms single-molecule displacements. (C) SMdM-determined D values for dCas9-mEos4b with and without gRNA expression. In (C), each circle represents the holistic D values from the nucleus of an individual cell.

Next by co-expressing Cas9-mEos4b with a nontargeting gRNA (Fig. 6.2B), we measured a D in the nucleus $\sim 3 \mu\text{m}^2/\text{s}$ consistent with previous results using dCas9-HaloTag using a nonsense gRNA¹⁴⁸. By measuring the holistic D coefficient for the whole nuclei for each cell but excluding the nucleolus, (Fig. 6.2C), cells expressing gRNA had $\sim 50\%$ increase in diffusivity suggesting that the formation of dCas9-gRNA complex had higher intranuclear diffusivity than dCas9 alone, suggesting that negatively charged gRNA in the Cas9-gRNA complex may increase its intranuclear diffusivity.

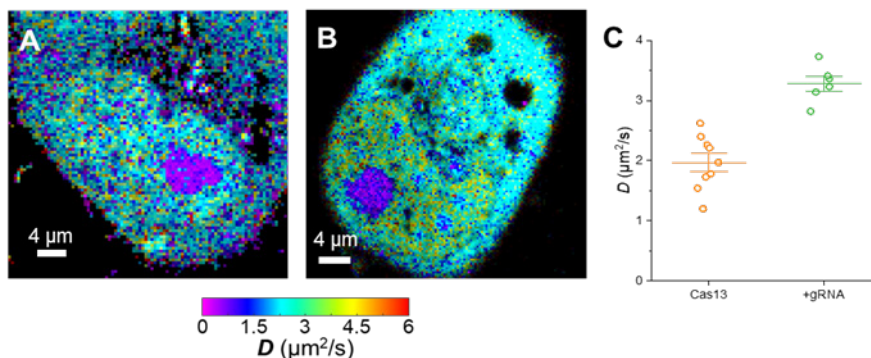


Figure 6.3. SMdM quantification of diffusion in the nucleus of dCas13b-mEos4b of live COS-7 cells. Representative D spatial maps of (A) dCas13b-mEos4b, (B) dCas13b-mEos4b and nontargeting gRNA, in the nucleus

of live COS-7 cells using 2 ms single-molecule displacements. (C) SMdM-determined D values for dCas13b-mEos4b with and without gRNA expression. In (C), each circle represents the holistic D values from the nucleus of an individual cell.

To generalize these results, we selected the positively charged dCas13b which similarly forms ribonucleoprotein by binding to gRNA to become negatively charged, but unlike Cas9 which binds to DNA when complexed with gRNA, dCas13b binds RNA. By using plasmid transfection to express the catalytically dead dCas13b-mEos4b, using SMdM, we were able to track freely diffusing dCas13b-mEos4b in the nucleus using $\Delta t = 2$ ms center-to-center separation, and spatially mapped diffusivity using 320 nm sized spatial bins (Fig. 6.3A). Similarly to dCas9-mEos4b in the nucleus, the D coefficient of dCas13b-mEos4b was $\sim 2 \mu\text{m}^2/\text{s}$, which is reasonable given that both proteins are ~ 160 kDa.

Next to study how the formation of the ribonucleoprotein complex of gRNA-dCas13 influences intranuclear diffusion, we co-expressed nontargeting gRNA for Cas13b with dCas13b-mEos4b (Fig. 6.3B). Similarly to dCas9-mEos4b-gRNA, dCas13b-mEos4b-gRNA had $D \sim 3 \mu\text{m}^2/\text{s}$ or 50% increase in D (Fig. 6.3C), further supporting that net negatively charged ribonucleoproteins facilitates intracellular diffusion if the initial protein is net positively charged.

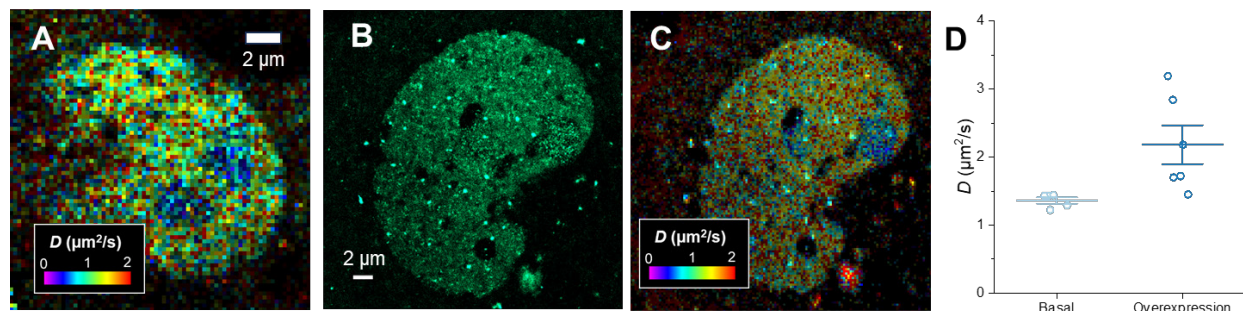


Figure 6.4. SMdM quantification of diffusion in the nucleus of human telomerase-mEos4b of live COS-7 cells. Representative D spatial maps of (A) telomerase-mEos4b, (B) telomerase-mEos4b and telomerase RNA, in the nucleus of live COS-7 cells using 2 ms single-molecule displacements. (C) SMdM-determined D values for telomerase-mEos4b at basal telomerase RNA expression or at overexpression of telomerase RNA. In (D), each circle represents the holistic D values from the nucleus of an individual cell.

Given that both Cas9 and Cas13b are not endogenous to mammalian cells, we chose the endogenous protein, human telomerase, which forms the ribonucleoprotein by binding to telomerase RNA (teloRNA), a type of long noncoding RNA^{149, 150} to form the human telomerase reverse transcriptase (hTERT) complex. By plasmid transfection, we expressed human-telomerase-mEos4b, and spatially mapped its diffusivity using $\Delta t = 2$ ms center-to-center separation, and spatially mapped diffusivity using 320 nm sized bins, measuring $D \sim 1.4 \mu\text{m}^2/\text{s}$ (Figure 6.4A) consistent with previous measurements of human telomerase¹⁵¹. Unlike Cas9 and Cas13 which are not endogenous to mammalian cells, telomerase and teloRNA are basally expressed in mammalian cells suggesting some population of hTERT-teloRNA-mEos4b complex is likely present and tracked by SMdM. However, due to the low levels of expression of both,

spatial maps likely represent mostly the non-complexed telomerase-mEos4b, which is overexpressed by plasmid transfection.

Upon co-expression of teloRNA with telomerase-mEos4b, we observed spatial heterogeneity in the distribution of telomerase-mEos4b where it localized to nanometer sized regions resembling telomeres (Figure 6.4B). Spatially mapping its diffusivity, we observed heterogenous diffusivity near these telomeric regions of $\sim 1 \mu\text{m}^2/\text{s}$, but was higher at $\sim 2 \mu\text{m}^2/\text{s}$ in more spatially homogeneous regions of the nucleus (Fig. 6.4C). By measuring the holistic D value for both cells overexpressing teloRNA and basal expression, we showed that teloRNA overexpression resulted in a $\sim 40\%$ increase in D (Fig. 6.4D).

In conclusion, we have demonstrated for three ribonucleoproteins, the formation of the ribonucleoprotein in nuclei of live COS-7 cells results in increased D for dCas9, dCas13 and telomerase. While we still have not directly linked charge as the main determinant for the diffusion slowdown of those protein, future work should focus on measuring the diffusivity of net negatively charged protein that form ribonucleoprotein complexes or by engineering the net charge state of Cas9, Cas13b or telomerase to be negatively charged without their corresponding RNA to see if the same diffusion change comparable to its ribonucleoprotein diffusivity in the nucleus. Overall, I believe such a study will impact protein engineering where charge optimization is crucial for achieving the fastest intracellular diffusion kinetics.

6.2.2 Single-Molecule Microscope to Measure Net Charge.

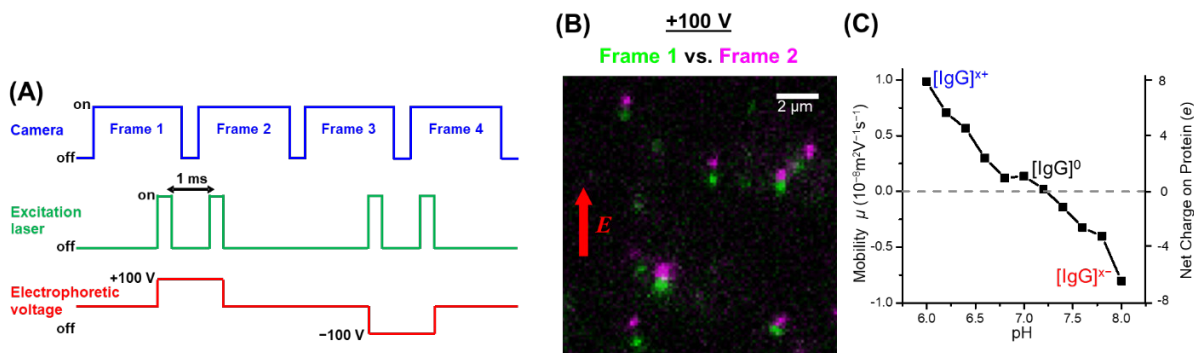


Figure 6.5. Single-molecule *electrophoresis* microscopy. (A) Representative schematic of single-molecule *electrophoresis* microscopy showing electrophoretic. (B) Overlaid single-molecule fluorescence images of dye tagged IgG molecules recorded between paired frames (green vs. magenta) (C) IgG at different pHs show different electrophoretic mobilities based on the protonation state of protein sidechain groups.

One main emphasis in my research is to develop novel microscopes capable of providing physicochemical imaging to SRM, coined ‘functional’ SRM. While we have developed multiple functional SRM methods to study cellular hydrophobicity¹⁰ and diffusion², we have yet to develop one to quantify the net charge of a single molecule.

In Figure 6.5.A, I present a method developed in collaboration with my coworkers, Dr. Kun Chen and Dr. Wan Li, which is a microscope capable of characterizing the net charge of single molecules. In this approach, we implement the same optical scheme as SMdM where we apply short excitation pulses between paired frames to track displacements of quickly moving single molecules. However, we now additionally apply a voltage pulse during the excitation to displace charged molecules in or away from the direction of the applied electric field. In Figure 6.5B., a Cy3B-tagged IgG protein is systematically displaced in the direction of the electric field, suggesting that it was positively charged. Based on their displacements, we can decouple the displacement due to random diffusion and displacement due to electrophoresis, and directly determine the net charge of single molecules. As one example, we showed that by varying the pH, we could identify small charge changes in the IgG that were sensitive to 1 electronic charge change (Figure 6.5C).

We have successfully characterized the net charge of a variety of protein *in vitro* using this approach that match well with theoretical predictions of net charge, and since moved into the *Xenopus laevis* extract systems to understand the charge environment of a cell. As discussed in Chapter 5, for positively charged proteins, we found that although their diffusion in the *Xenopus laevis* extract is substantially suppressed possibly from the attraction of a layer of negatively charged macromolecules such as RNA or protein, the detected single-molecule electrophoresis results nonetheless indicated positive net charges matching their native states. This suggested that the negatively charged shell structures formed *in situ* that impede diffusion are highly transient.

References

1. Wühr, M.; Freeman, Robert M.; Presler, M.; Horb, Marko E.; Peshkin, L.; Gygi, Steven P.; Kirschner, Marc W., Deep proteomics of the *Xenopus laevis* egg using an mRNA-derived reference database. *Curr. Biol.* **2014**, *24* (13), 1467-1475.
2. Xiang, L.; Chen, K.; Yan, R.; Li, W.; Xu, K., Single-molecule displacement mapping unveils nanoscale heterogeneities in intracellular diffusivity. *Nature Methods* **2020**, *17* (5), 524-530.
3. Walla, P. J., Modern Biophysical Chemistry Detection and Analysis of Biomolecules Foreword. *Modern Biophysical Chemistry: Detection and Analysis of Biomolecules, 2nd Edition* **2014**, Xiii-Xiv.
4. Thompson, R. E.; Larson, D. R.; Webb, W. W., Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.* **2002**, *82* (5), 2775-2783.
5. Dempsey, G. T.; Vaughan, J. C.; Chen, K. H.; Bates, M.; Zhuang, X., Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nature Methods* **2011**, *8* (12), 1027-1036.
6. Rust, M. J.; Bates, M.; Zhuang, X. W., Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods* **2006**, *3* (10), 793-795.
7. Betzig, E.; Patterson, G. H.; Sougrat, R.; Lindwasser, O. W.; Olenych, S.; Bonifacino, J. S.; Davidson, M. W.; Lippincott-Schwartz, J.; Hess, H. F., Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **2006**, *313* (5793), 1642-1645.
8. Yan, R.; Wang, B. W.; Xu, K., super-resolution microscopy of the cell. *Curr Opin Chem Biol* **2019**, *51*, 92-97.
9. Xiang, L.; Chen, K.; Xu, K., Single molecules are your quanta: A bottom-up approach toward multidimensional super-resolution microscopy. *ACS Nano* **2021**, *15* (8), 12483-12496.
10. Moon, S.; Yan, R.; Kenny, S. J.; Shyu, Y.; Xiang, L. M.; Li, W.; Xu, K., Spectrally Resolved, Functional Super-Resolution Microscopy Reveals Nanoscale Compositional Heterogeneity in Live-Cell Membranes. *J. Am. Chem. Soc.* **2017**, *139* (32), 10944-10947.
11. Choi, A. A.; Xiang, L.; Li, W.; Xu, K., Single-molecule displacement mapping indicates unhindered intracellular diffusion of small (≤ 1 kDa) solutes. *J. Am. Chem. Soc.* **2023**, *145* (15), 8510-8516.
12. Choi, A. A.; Park, H. H.; Chen, K.; Yan, R.; Li, W.; Xu, K., Displacement statistics of unhindered single molecules show no enhanced diffusion in enzymatic reactions. *J. Am. Chem. Soc.* **2022**, *144* (11), 4839-4844.
13. Choi, A. A.; Xu, K., Single-molecule diffusivity quantification unveils ubiquitous net charge-driven protein-protein interaction. *Journal of the American Chemical Society* **2024**, DOI: 10.1021/jacs.4c02475.
14. Muddana, H. S.; Sengupta, S.; Mallouk, T. E.; Sen, A.; Butler, P. J., Substrate catalysis enhances single-enzyme diffusion. *J. Am. Chem. Soc.* **2010**, *132* (7), 2110-2111.
15. Riedel, C.; Gabizon, R.; Wilson, C. A. M.; Hamadani, K.; Tsekouras, K.; Marqusee, S.; Presse, S.; Bustamante, C., The heat released during catalytic turnover enhances the diffusion of an enzyme. *Nature* **2015**, *517* (7533), 227-230.
16. Gunther, J. P.; Borsch, M.; Fischer, P., Diffusion Measurements of Swimming Enzymes with Fluorescence Correlation Spectroscopy. *Accounts Chem. Res.* **2018**, *51* (9), 1911-1920.
17. Agudo-Canalejo, J.; Adeleke-Larodo, T.; Illien, P.; Golestanian, R., Enhanced Diffusion and Chemotaxis at the Nanoscale. *Accounts Chem. Res.* **2018**, *51* (10), 2365-2372.

18. Zhao, X.; Gentile, K.; Mohajerani, F.; Sen, A., Powering motion with enzymes. *Accounts Chem. Res.* **2018**, *51* (10), 2373-2381.
19. Zhang, Y. F.; Hess, H., Enhanced diffusion of catalytically active enzymes. *ACS Central Sci.* **2019**, *5* (6), 939-948.
20. Feng, M.; Gilson, M. K., Enhanced Diffusion and Chemotaxis of Enzymes. *Annual Review of Biophysics* **2020**, *49* (1), 87-105.
21. Zhang, Y.; Hess, H., Chemically-powered swimming and diffusion in the microscopic world. *Nature Reviews Chemistry* **2021**, *5*, 500-510.
22. Sengupta, S.; Dey, K. K.; Muddana, H. S.; Tabouillot, T.; Ibele, M. E.; Butler, P. J.; Sen, A., Enzyme Molecules as Nanomotors. *J. Am. Chem. Soc.* **2013**, *135* (4), 1406-1414.
23. Illien, P.; Zhao, X.; Dey, K. K.; Butler, P. J.; Sen, A.; Golestanian, R., Exothermicity Is Not a Necessary Condition for Enhanced Diffusion of Enzymes. *Nano Lett* **2017**, *17* (7), 4415-4420.
24. Jiang, L.; Santiago, I.; Foord, J., Observation of nanoimpact events of catalase on diamond ultramicroelectrodes by direct electron transfer. *Chemical Communications* **2017**, *53* (59), 8332-8335.
25. Jee, A. Y.; Dutta, S.; Cho, Y. K.; Tlusty, T.; Granick, S., Enzyme leaps fuel antichemotaxis. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (1), 14-18.
26. Jee, A. Y.; Cho, Y. K.; Granick, S.; Tlusty, T., Catalytic enzymes are active matter. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (46), E10812-E10821.
27. Jee, A. Y.; Chen, K.; Tlusty, T.; Zhao, J.; Granick, S., Enhanced diffusion and oligomeric enzyme dissociation. *J. Am. Chem. Soc.* **2019**, *141* (51), 20062-20068.
28. Ghosh, S.; Mohajerani, F.; Son, S.; Velegol, D.; Butler, P. J.; Sen, A., Motility of enzyme-powered vesicles. *Nano Lett* **2019**, *19* (9), 6019-6026.
29. Jee, A. Y.; Tlusty, T.; Granick, S., Master curve of boosted diffusion for 10 catalytic enzymes. *Proc Natl Acad Sci U S A* **2020**, *117* (47), 29435-29441.
30. Kandula, H. N.; Jee, A. Y.; Granick, S., Robustness of FCS (Fluorescence Correlation Spectroscopy) with Quenchers Present. *J Phys Chem A* **2019**, *123* (46), 10184-10189.
31. Chen, Z.; Shaw, A.; Wilson, H.; Woring, M.; Darzacq, X.; Marqusee, S.; Wang, Q.; Bustamante, C., Single-molecule diffusometry reveals no catalysis-induced diffusion enhancement of alkaline phosphatase as proposed by FCS experiments. *Proceedings of the National Academy of Sciences* **2020**, *117* (35), 21328-21335.
32. Zhang, Y. F.; Armstrong, M. J.; Kazeruni, N. M. B.; Hess, H., Aldolase Does Not Show Enhanced Diffusion in Dynamic Light Scattering Experiments. *Nano Lett* **2018**, *18* (12), 8025-8029.
33. Pavlick, R. A.; Dey, K. K.; Sirjoosingh, A.; Benesi, A.; Sen, A., A catalytically driven organometallic molecular motor. *Nanoscale* **2013**, *5* (4), 1301-1304.
34. Gunther, J. P.; Majer, G.; Fischer, P., Absolute diffusion measurements of active enzyme solutions by NMR. *J. Chem. Phys.* **2019**, *150* (12), 124201.
35. MacDonald, T. S. C.; Price, W. S.; Astumian, R. D.; Beves, J. E., Enhanced Diffusion of Molecular Catalysts is Due to Convection. *Angewandte Chemie International Edition* **2019**, *58* (52), 18864-18867.
36. Wang, H.; Park, M.; Dong, R.; Kim, J.; Cho, Y.-K.; Tlusty, T.; Granick, S., Boosted molecular mobility during common chemical reactions. *Science* **2020**, *369* (6503), 537-541.

37. Günther, J.-P.; Fillbrook, L. L.; MacDonald, T. S. C.; Majer, G.; Price, W. S.; Fischer, P.; Beves, J. E., Comment on “Boosted molecular mobility during common chemical reactions”. *Science* **2021**, *371* (6526), eabe8322.
38. Wang, H.; Park, M.; Dong, R.; Kim, J.; Cho, Y.-K.; Tlustý, T.; Granick, S., Response to Comment on “Boosted molecular mobility during common chemical reactions”. *Science* **2021**, *371* (6526), eabe8678.
39. Ha, T., Single-molecule methods leap ahead. *Nature Methods* **2014**, *11* (10), 1015-1018.
40. Moerner, W. E.; Shechtman, Y.; Wang, Q., Single-molecule spectroscopy and imaging over the decades. *Faraday Discuss* **2015**, *184*, 9-36.
41. Shen, H.; Tauzin, L. J.; Baiyasi, R.; Wang, W. X.; Moringo, N.; Shuang, B.; Landes, C. F., Single particle tracking: from theory to biophysical applications. *Chem. Rev.* **2017**, *117* (11), 7331-7376.
42. Young, M. E.; Carroad, P. A.; Bell, R. L., Estimation of diffusion coefficients of proteins. *Biotechnology and Bioengineering* **1980**, *22* (5), 947-955.
43. Xu, M. Q.; Ross, J. L.; Valdez, L.; Sen, A., Direct Single Molecule Imaging of Enhanced Enzyme Diffusion. *Phys. Rev. Lett.* **2019**, *123* (12), 128101.
44. Xu, M.; Rogers, W. B.; Ahmed, W. W.; Ross, J. L., Comparison of different approaches to single-molecule imaging of enhanced enzyme diffusion. **2021**, arXiv:2012.15424.
45. Yan, R.; Chen, K.; Xu, K., Probing nanoscale diffusional heterogeneities in cellular membranes through multidimensional single-molecule and super-resolution microscopy. *J. Am. Chem. Soc.* **2020**, *142*, 18866-18873.
46. Xu, K.; Shim, S.-H.; Zhuang, X., Super-resolution imaging through stochastic switching and localization of single molecules: an overview. In *Far-Field Optical Nanoscopy*, Tinnefeld, P.; Eggeling, C.; Hell, S. W., Eds. Springer: Berlin, 2015; pp 27-64.
47. Möckl, L.; Moerner, W. E., Super-resolution microscopy with single molecules in biology and beyond-essentials, current trends, and future challenges. *J. Am. Chem. Soc.* **2020**, *142*, 17828-17844.
48. Cooper, M.; Ebner, A.; Briggs, M.; Burrows, M.; Gardner, N.; Richardson, R.; West, R., Cy3B™: Improving the Performance of Cyanine Dyes. *Journal of Fluorescence* **2004**, *14* (2), 145-150.
49. Sheely, M. L., Glycerol viscosity tables. *Ind. Eng. Chem.* **1932**, *24* (9), 1060-1064.
50. Truong, M. E.; Bilekova, S.; Choksi, S. P.; Li, W.; Bugaj, L. J.; Xu, K.; Reiter, J. F., Vertebrate cells differentially interpret ciliary and extraciliary cAMP. *Cell* **2021**, *184* (11), 2911-2926.
51. Berg, H. C., *Random walks in biology*. Princeton University Press: Princeton, N.J., 1993; p vii, 152 p.
52. Kurniawan, N. A.; van Kempen, T. H. S.; Sonneveld, S.; Rosalina, T. T.; Vos, B. E.; Jansen, K. A.; Peters, G. W. M.; van de Vosse, F. N.; Koenderink, G. H., Buffers Strongly Modulate Fibrin Self-Assembly into Fibrous Networks. *Langmuir* **2017**, *33* (25), 6342-6352.
53. Islamovic, S.; Galic, B.; Milos, M., A study of the inhibition of catalase by dipotassium trioxohydroxytetrafluorotriborate K₂[B₃O₃F₄OH]. *Journal of Enzyme Inhibition and Medicinal Chemistry* **2014**, *29* (5), 744-748.
54. Beber, M. E.; Gollub, M. G.; Mozaffari, D.; Shebek, K. M.; Flamholz, A. I.; Milo, R.; Noor, E., eQuilibrator 3.0: a database solution for thermodynamic constant estimation. *Nucleic Acids Research* **2022**, *50* (D1), D603-D609.

55. Lippincott-Schwartz, J.; Snapp, E.; Kenworthy, A., Studying protein dynamics in living cells. *Nat. Rev. Mol. Cell Biol.* **2001**, *2* (6), 444-456.
56. Verkman, A. S., Solute and macromolecule diffusion in cellular aqueous compartments. *Trends Biochem.Sci.* **2002**, *27* (1), 27-33.
57. Soh, S.; Byrska, M.; Kandere-Grzybowska, K.; Grzybowski, B. A., Reaction-Diffusion Systems in Intracellular Molecular Transport and Control. *Angew. Chem.-Int. Edit.* **2010**, *49* (25), 4170-4198.
58. Bressloff, P. C.; Newby, J. M., Stochastic models of intracellular transport. *Reviews of Modern Physics* **2013**, *85* (1), 135-196.
59. Machan, R.; Wohland, T., Recent applications of fluorescence correlation spectroscopy in live systems. *FEBS Lett.* **2014**, *588* (19), 3571-3584.
60. Milo, R.; Phillips, R., *Cell Biology by the Numbers*. Garland Science: New York, NY, 2016; p xlii, 356 pages.
61. Lippincott-Schwartz, J.; Snapp, E. L.; Phair, R. D., The development and enhancement of FRAP as a key tool for investigating protein dynamics. *Biophys J* **2018**, *115* (7), 1146-1155.
62. Mogre, S. S.; Brown, A. I.; Koslover, E. F., Getting around the cell: physical transport in the intracellular world. *Phys Biol* **2020**, *17* (6), 061003.
63. Shim, S. H.; Xia, C.; Zhong, G.; Babcock, H. P.; Vaughan, J. C.; Huang, B.; Wang, X.; Xu, C.; Bi, G. Q.; Zhuang, X., Super-resolution fluorescence imaging of organelles in live cells with photoswitchable membrane probes. *Proc Natl Acad Sci U S A* **2012**, *109* (35), 13978-13983.
64. Yan, R.; Chen, K.; Xu, K., Probing nanoscale diffusional heterogeneities in cellular membranes through multidimensional single-molecule and super-resolution microscopy. *J. Am. Chem. Soc.* **2020**, *142* (44), 18866-18873.
65. Slayman, C. L.; Moussatos, V. V.; Webb, W. W., Endosomal accumulation of pH indicator dyes delivered as acetoxymethyl esters. *Journal of Experimental Biology* **1994**, *196* (1), 419-438.
66. Hughes, L. D.; Rawle, R. J.; Boxer, S. G., Choose your label wisely: water-soluble fluorophores often interact with lipid bilayers. *PLoS One* **2014**, *9* (2), e87649.
67. Hayashi-Takanaka, Y.; Stasevich, T. J.; Kurumizaka, H.; Nozaki, N.; Kimura, H., Evaluation of Chemical Fluorescent Dyes as a Protein Conjugation Partner for Live Cell Imaging. *PLOS ONE* **2014**, *9* (9), e106271.
68. Bosch, P. J.; Correa, I. R., Jr.; Sonntag, M. H.; Ibach, J.; Brunsfeld, L.; Kanger, J. S.; Subramaniam, V., Evaluation of fluorophores to label SNAP-tag fused proteins for multicolor single-molecule tracking microscopy in live cells. *Biophys J* **2014**, *107* (4), 803-814.
69. Manzo, C.; Garcia-Parajo, M. F., A review of progress in single particle tracking: from methods to biophysical insights. *Rep. Prog. Phys.* **2015**, *78* (12), 124601.
70. Ries, J.; Schuille, P., Fluorescence correlation spectroscopy. *Bioessays* **2012**, *34* (5), 361-368.
71. Theillet, F.-X., In-Cell Structural Biology by NMR: The Benefits of the Atomic Scale. *Chem. Rev.* **2022**, *122* (10), 9497-9570.
72. Kao, H. P.; Abney, J. R.; Verkman, A. S., Determinants of the translational mobility of a small solute in cell cytoplasm. *J Cell Biol* **1993**, *120* (1), 175-184.
73. Fushimi, K.; Verkman, A. S., Low Viscosity in the Aqueous Domain of Cell Cytoplasm Measured by Picosecond Polarization Microfluorimetry. *J Cell Biol* **1991**, *112* (4), 719-725.

74. Swaminathan, R.; Hoang, C. P.; Verkman, A. S., Photobleaching recovery and anisotropy decay of green fluorescent protein GFP-S65T in solution and cells: Cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion. *Biophys. J.* **1997**, 72 (4), 1900-1907.
75. de Graaf, R. A.; van Kranenburg, A.; Nicolay, K., In vivo ^{31}P -NMR diffusion spectroscopy of ATP and phosphocreatine in rat skeletal muscle. *Biophys. J.* **2000**, 78 (4), 1657-1664.
76. Vendelin, M.; Birkedal, R., Anisotropic diffusion of fluorescently labeled ATP in rat cardiomyocytes determined by raster image correlation spectroscopy. *Am J Physiol Cell Physiol* **2008**, 295 (5), C1302-C1315.
77. Mika, J. T.; Van Den Bogaart, G.; Veenhoff, L.; Krasnikov, V.; Poolman, B., Molecular sieving properties of the cytoplasm of *Escherichia coli* and consequences of osmotic stress. *Molecular Microbiology* **2010**, 77 (1), 200-207.
78. Agarwal, S. R.; Clancy, C. E.; Harvey, R. D., Mechanisms restricting diffusion of intracellular cAMP. *Sci Rep* **2016**, 6, 19577.
79. Bock, A.; Annibale, P.; Konrad, C.; Hannawacker, A.; Anton, S. E.; Maiellaro, I.; Zabel, U.; Sivaramakrishnan, S.; Falcke, M.; Lohse, M. J., Optical Mapping of cAMP Signaling at the Nanometer Scale. *Cell* **2020**, 182 (6), 1519-1530.
80. Kwapiszewska, K.; Szczepański, K.; Kalwarczyk, T.; Michalska, B.; Patalas-Krawczyk, P.; Szymański, J.; Andryszewski, T.; Iwan, M.; Duszyński, J.; Hołyst, R., Nanoscale viscosity of cytoplasm is conserved in human cell lines. *The Journal of Physical Chemistry Letters* **2020**, 11 (16), 6914-6920.
81. Nimmerjahn, A.; Kirchhoff, F.; Kerr, J. N. D.; Helmchen, F., Sulforhodamine 101 as a specific marker of astroglia in the neocortex in vivo. *Nature Methods* **2004**, 1 (1), 31-37.
82. Schnell, C.; Hagos, Y.; Hülsmann, S., Active sulforhodamine 101 uptake into hippocampal astrocytes. *PLOS ONE* **2012**, 7 (11), e49398.
83. Moon, S.; Li, W.; Hauser, M.; Xu, K., Graphene-enabled, spatially controlled electroporation of adherent cells for live-cell super-resolution microscopy. *ACS Nano* **2020**, 14 (5), 5609-5617.
84. Xie, X. S.; Choi, P. J.; Li, G. W.; Lee, N. K.; Lia, G., Single-molecule approach to molecular biology in living bacterial cells. *Annual Review of Biophysics* **2008**, 37, 417-444.
85. Ha, T.; Tinnefeld, P., Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. *Annual Review of Physical Chemistry* **2012**, 63, 595-617.
86. Culbertson, C. T.; Jacobson, S. C.; Michael Ramsey, J., Diffusion coefficient measurements in microfluidic devices. *Talanta* **2002**, 56 (2), 365-373.
87. Petrášek, Z.; Schwille, P., Precise Measurement of Diffusion Coefficients using Scanning Fluorescence Correlation Spectroscopy. *Biophys. J.* **2008**, 94 (4), 1437-1448.
88. Dubielecka, P. M.; Cui, P.; Xiong, X.; Hossain, S.; Heck, S.; Angelov, L.; Kotula, L., Differential regulation of macropinocytosis by Abi1/Hssh3bp1 isoforms. *PLOS ONE* **2010**, 5 (5), e10430.
89. Nikolaev, V. O.; Bünnemann, M.; Hein, L.; Hannawacker, A.; Lohse, M. J., Novel Single Chain cAMP Sensors for Receptor-induced Signal Propagation. *Journal of Biological Chemistry* **2004**, 279 (36), 37215-37218.
90. Segur, J. B.; Oberstar, H. E., Viscosity of Glycerol and Its Aqueous Solutions. *Ind. Eng. Chem.* **1951**, 43 (9), 2117-2120.

91. He, C.; Wu, C. Y.; Li, W.; Xu, K., Multidimensional super-resolution microscopy unveils nanoscale surface aggregates in the aging of FUS condensates. *J. Am. Chem. Soc.* **2023**, *145* (44), 24240-24248.
92. Xiang, L.; Yan, R.; Chen, K.; Li, W.; Xu, K., Single-molecule displacement mapping unveils sign-asymmetric protein charge effects on intraorganellar diffusion. *Nano Lett* **2023**, *23* (5), 1711-1716.
93. Schavemaker, P. E.; Smigiel, W. M.; Poolman, B., Ribosome surface properties may impose limits on the nature of the cytoplasmic proteome. *eLife* **2017**, *6*, e30084.
94. Ye, Y.; Wu, Q.; Zheng, W.; Jiang, B.; Pielak, G. J.; Liu, M.; Li, C., Positively charged tags impede protein mobility in cells as quantified by ^{19}F NMR. *The Journal of Physical Chemistry B* **2019**, *123* (21), 4527-4533.
95. Leeb, S.; Sörensen, T.; Yang, F.; Mu, X.; Oliveberg, M.; Danielsson, J., Diffusive protein interactions in human versus bacterial cells. *Current Research in Structural Biology* **2020**, *2*, 68-78.
96. Leeb, S.; Yang, F.; Oliveberg, M.; Danielsson, J., Connecting longitudinal and transverse relaxation rates in live-cell NMR. *The Journal of Physical Chemistry B* **2020**, *124* (47), 10698-10707.
97. Vallina Estrada, E.; Zhang, N.; Wennerström, H.; Danielsson, J.; Oliveberg, M., Diffusive intracellular interactions: On the role of protein net charge and functional adaptation. *Current Opinion in Structural Biology* **2023**, *81*, 102625.
98. Beck, M.; Schmidt, A.; Malmstroem, J.; Claassen, M.; Ori, A.; Szymborska, A.; Herzog, F.; Rinner, O.; Ellenberg, J.; Aebersold, R., The quantitative proteome of a human cell line. *Mol. Syst. Biol.* **2011**, *7*, 549.
99. Itzhak, D. N.; Tyanova, S.; Cox, J.; Borner, G. H. H., Global, quantitative and dynamic mapping of protein subcellular localization. *eLife* **2016**, *5*, e16950.
100. Theillet, F. X.; Binolfi, A.; Frembgen-Kesner, T.; Hingorani, K.; Sarkar, M.; Kyne, C.; Li, C. G.; Crowley, P. B.; Gierasch, L.; Pielak, G. J.; Elcock, A. H.; Gershenson, A.; Selenko, P., Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs). *Chem. Rev.* **2014**, *114* (13), 6661-6714.
101. Lodish, H. F.; Berk, A.; Kaiser, C.; Krieger, M.; Bretscher, A.; Ploegh, H. L.; Amon, A.; Martin, K. C., *Molecular cell biology*. Eighth edition. ed.; W.H. Freeman-Macmillan Learning: New York, 2016; p xl, 1170, G-28, I-36 pages.
102. Wennerström, H.; Estrada, E. V.; Danielsson, J.; Oliveberg, M., Colloidal stability of the living cell. *Proceedings of the National Academy of Sciences* **2020**, *117* (19), 10113-10121.
103. Dix, J. A.; Verkman, A. S., Crowding effects on diffusion in solutions and cells. *Annual Review of Biophysics* **2008**, *37*, 247-263.
104. Rivas, G.; Minton, A. P., Macromolecular crowding *in vitro*, *in vivo*, and in between. *Trends Biochem.Sci.* **2016**, *41* (11), 970-981.
105. Guin, D.; Gruebele, M., Weak Chemical Interactions That Drive Protein Evolution: Crowding, Sticking, and Quinary Structure in Folding and Function. *Chem Rev* **2019**, *119* (18), 10691-10717.
106. Srivastava, S.; Tirrell, M. V., Polyelectrolyte complexation. *Adv. Chem. Phys.* **2016**, *161*, 499-544.
107. Croguennec, T.; Tavares, G. M.; Bouhallab, S., Heteroprotein complex coacervation: A generic process. *Advances in Colloid and Interface Science* **2017**, *239*, 115-126.

108. Zheng, J.; Tang, C.-h.; Sun, W., Heteroprotein complex coacervation: Focus on experimental strategies to investigate structure formation as a function of intrinsic and external physicochemical parameters for food applications. *Advances in Colloid and Interface Science* **2020**, *284*, 102268.
109. Kapelner, R. A.; Yeong, V.; Obermeyer, A. C., Molecular determinants of protein-based coacervates. *Curr Opin Colloid In* **2021**, *52*, 101407.
110. Galvanetto, N.; Ivanović, M. T.; Chowdhury, A.; Sottini, A.; Nüesch, M. F.; Nettels, D.; Best, R. B.; Schuler, B., Extreme dynamics in a biomolecular condensate. *Nature* **2023**, *619* (7971), 876-883.
111. Zhu, Y.; Mohapatra, S.; Weisshaar, J. C., Rigidification of the *Escherichia coli* cytoplasm by the human antimicrobial peptide LL-37 revealed by superresolution fluorescence microscopy. *Proceedings of the National Academy of Sciences* **2019**, *116* (3), 1017-1026.
112. Yeong, V.; Werth, E. G.; Brown, L. M.; Obermeyer, A. C., Formation of Biomolecular Condensates in Bacteria by Tuning Protein Electrostatics. *ACS Central Sci.* **2020**, *6* (12), 2301-2310.
113. Saroff, H. A.; Rosenthal, N. R.; Adamik, E. R.; Hages, N.; Scheraga, H. A., The methylation of bovine serum albumin. *J Biol Chem* **1953**, *205* (1), 255-270.
114. Gitlin, I.; Carbeck, J. D.; Whitesides, G. M., Why are proteins charged? Networks of charge-charge interactions in proteins measured by charge ladders and capillary electrophoresis. *Angew. Chem.-Int. Edit.* **2006**, *45* (19), 3022-3060.
115. Rentsch, J.; Bandstra, S.; Sezen, B.; Sigrist, P.; Bottanelli, F.; Schmerl, B.; Shoichet, S.; Noé, F.; Sadeghi, M.; Ewers, H., Sub-membrane actin rings compartmentalize the plasma membrane. *J Cell Biol* **2024**, *223* (4), e202310138.
116. Geoghegan, K. F., Modification of amino groups. *Curr. Protoc. Protein Sci.* **2016**, *86* (1), 15.2.1-15.2.20.
117. Stadtmiller, S. S.; Pielak, G. J., Protein-complex stability in cells and *in vitro* under crowded conditions. *Curr. Opin. Struct. Biol.* **2021**, *66*, 183-192.
118. Bonucci, M.; Shu, T.; Holt, L. J., How it feels in a cell. *Trends Cell Biol.* **2023**, *33* (11), 924-938.
119. Steves, M. A.; He, C.; Xu, K., Single-molecule spectroscopy and super-resolution mapping of physicochemical parameters in living cells. *Annu. Rev. Phys. Chem.* **2024**, DOI: 10.1146/annurev-physchem-070623-034225.
120. Choi, A. A.; Xu, K., Single-molecule diffusivity quantification unveils ubiquitous net charge-driven protein-protein interaction. *J. Am. Chem. Soc.* **2024**, *146* (15), 10973-10978.
121. Field, C. M.; Pelletier, J. F.; Mitchison, T. J., *Xenopus* extract approaches to studying microtubule organization and signaling in cytokinesis. *Methods Cell Biol* **2017**, *137*, 395-435.
122. Good, M. C.; Heald, R., Preparation of cellular extracts from *Xenopus* eggs and embryos. *Cold Spring Harb. Protoc.* **2018**, (6), doi:10.1101/pdb.prot097055.
123. Gibeaux, R.; Heald, R., The use of cell-free *Xenopus* extracts to investigate cytoplasmic events. *Cold Spring Harb. Protoc.* **2019**, (6), doi:10.1101/pdb.top097048.
124. Heald, R.; Tournebise, R.; Blank, T.; Sandaltzopoulos, R.; Becker, P.; Hyman, A.; Karsenti, E., Self-organization of microtubules into bipolar spindles around artificial chromosomes in *Xenopus* egg extracts. *Nature* **1996**, *382* (6590), 420-425.
125. Good, M. C.; Vahey, M. D.; Skandarajah, A.; Fletcher, D. A.; Heald, R., Cytoplasmic Volume Modulates Spindle Size During Embryogenesis. *Science* **2013**, *342* (6160), 856-860.

126. Zhou, C. Y.; Dekker, B.; Liu, Z.; Cabrera, H.; Ryan, J.; Dekker, J.; Heald, R., Mitotic chromosomes scale to nuclear-cytoplasmic ratio and cell size in *Xenopus*. *eLife* **2023**, *12*, e84360.
127. Cheng, X. R.; Ferrell, J. E., Spontaneous emergence of cell-like organization in *Xenopus* egg extracts. *Science* **2019**, *366* (6465), 631-637.
128. Newmeyer, D. D.; Farschon, D. M.; Reed, J. C., Cell-free apoptosis in *Xenopus* egg extracts: Inhibition by Bcl-2 and requirement for an organelle fraction enriched in mitochondria. *Cell* **1994**, *79* (2), 353-364.
129. Huang, W. Y. C.; Cheng, X. R.; Ferrell, J. E., Cytoplasmic organization promotes protein diffusion in *Xenopus* extracts. *Nat. Commun.* **2022**, *13* (1), 5599.
130. Chen, Y.; Huang, J.-H.; Phong, C.; Ferrell, J. E., Viscosity-dependent control of protein synthesis and degradation. *Nat. Commun.* **2024**, *15* (1), 2149.
131. Keber, F. C.; Nguyen, T.; Mariossi, A.; Brangwynne, C. P.; Wuhr, M., Evidence for widespread cytoplasmic structuring into mesoscale condensates. *Nat. Cell Biol.* **2024**, *26* (3), 346-352.
132. Sehy, J. V.; Ackerman, J. J.; Neil, J. J., Apparent diffusion of water, ions, and small molecules in the *Xenopus* oocyte is consistent with Brownian displacement. *Magn. Reson. Med.* **2002**, *48* (1), 42-51.
133. Blower, M. D.; Nachury, M.; Heald, R.; Weis, K., A Rae1-Containing Ribonucleoprotein Complex Is Required for Mitotic Spindle Assembly. *Cell* **2005**, *121* (2), 223-234.
134. Aze, A.; Fragkos, M.; Bocquet, S.; Cau, J.; Méchali, M., RNAs coordinate nuclear envelope assembly and DNA replication through ELYS recruitment to chromatin. *Nat. Commun.* **2017**, *8* (1), 2130.
135. Aarum, J.; Cabrera, C. P.; Jones, T. A.; Rajendran, S.; Adiutori, R.; Giovannoni, G.; Barnes, M. R.; Malaspina, A.; Sheer, D., Enzymatic degradation of RNA causes widespread protein aggregation in cell and tissue lysates. *EMBO Rep.* **2020**, *21* (10), e49585.
136. Park, C.; Han, B.; Choi, Y.; Jin, Y.; Kim, K. P.; Choi, S. I.; Seong, B. L., RNA-dependent proteome solubility maintenance in *Escherichia coli* lysates analysed by quantitative mass spectrometry: Proteomic characterization in terms of isoelectric point, structural disorder, functional hub, and chaperone network. *RNA Biol.* **2024**, *21* (1), 1-18.
137. Okuyama, K.; Sagata, N.; Shiokawa, K.; Yamana, K., Localization of poly(A) sequences that are lost during oocyte maturation of *Xenopus laevis*. *Cell Struct. Funct.* **1981**, *6* (3), 279-282.
138. Tokmakov, A. A.; Hashimoto, T.; Hasegawa, Y.; Iguchi, S.; Iwasaki, T.; Fukami, Y., Monitoring gene expression in a single *Xenopus* oocyte using multiple cytoplasmic collections and quantitative RT-PCR. *FEBS J.* **2014**, *281* (1), 104-114.
139. Boeynaems, S.; Ma, X. R.; Yeong, V.; Ginell, G. M.; Chen, J.-H.; Blum, J. A.; Nakayama, L.; Sanyal, A.; Briner, A.; Haver, D. V.; Pauwels, J.; Ekman, A.; Schmidt, H. B.; Sundararajan, K.; Porta, L.; Lasker, K.; Larabell, C.; Hayashi, M. A. F.; Kundaje, A.; Impens, F.; Obermeyer, A.; Holehouse, A. S.; Gitler, A. D., Aberrant phase separation is a common killing strategy of positively charged peptides in biology and human disease. *bioRxiv* **2023**, DOI: 10.1101/2023.03.09.531820.
140. Dauty, E.; Verkman, A. S., Actin cytoskeleton as the principal determinant of size-dependent DNA mobility in cytoplasm. *J. Biol. Chem.* **2005**, *280* (9), 7823-7828.

141. Kwapiszewska, K.; Szczepanski, K.; Kalwarczyk, T.; Michalska, B.; Patalas-Krawczyk, P.; Szymanski, J.; Andryszewski, T.; Iwan, M.; Duszynski, J.; Holyst, R., Nanoscale viscosity of cytoplasm is conserved in human cell lines. *J. Phys. Chem. Lett.* **2020**, *11* (16), 6914-6920.
142. Park, H. H.; Choi, A. A.; Xu, K., Size-dependent suppression of molecular diffusivity in expandable hydrogels: A single-molecule study. *J. Phys. Chem. B* **2023**, *127* (14), 3333-3339.
143. Xu, K.; Babcock, H. P.; Zhuang, X., Dual-objective STORM reveals three-dimensional filament organization in the actin cytoskeleton. *Nat. Methods* **2012**, *9* (2), 185-188.
144. Huang, B.; Wang, W.; Bates, M.; Zhuang, X., Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* **2008**, *319* (5864), 810-813.
145. Wojcik, M.; Hauser, M.; Li, W.; Moon, S.; Xu, K., Graphene-enabled electron microscopy and correlated super-resolution microscopy of wet cells. *Nat. Commun.* **2015**, *6*, 7384.
146. Knight, A. M.; Culviner, P. H.; Kurt-Yilmaz, N.; Zou, T. S.; Ozkan, S. B.; Cavagnero, S., Electrostatic effect of the ribosomal surface on nascent polypeptide dynamics. *ACS Chem. Biol.* **2013**, *8* (6), 1195-1204.
147. Knight, S. C.; Tjian, R.; Doudna, J. A., Genomes in Focus: Development and Applications of CRISPR-Cas9 Imaging Technologies. *Angew. Chem.-Int. Edit.* **2018**, *57* (16), 4329-4337.
148. Knight, S. C.; Xie, L. Q.; Deng, W. L.; Guglielmi, B.; Witkowsky, L. B.; Bosanac, L.; Zhang, E. T.; El Beheiry, M.; Masson, J. B.; Dahan, M.; Liu, Z.; Doudna, J. A.; Tjian, R., Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science* **2015**, *350* (6262), 823-826.
149. Mattick, J. S.; Amaral, P. P.; Carninci, P.; Carpenter, S.; Chang, H. Y.; Chen, L. L.; Chen, R. S.; Dean, C.; Dinger, M. E.; Fitzgerald, K. A.; Gingeras, T. R.; Guttman, M.; Hirose, T.; Huarte, M.; Johnson, R.; Kanduri, C.; Kapranov, P.; Lawrence, J. B.; Lee, J. T.; Mendell, J. T.; Mercer, T. R.; Moore, K. J.; Nakagawa, S.; Rinn, J. L.; Spector, D. L.; Ulitsky, I.; Wan, Y.; Wilusz, J. E.; Wu, M., Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.* **2023**, *24* (6), 430-447.
150. Cong, Y. S.; Wright, W. E.; Shay, J. W., Human telomerase and its regulation. *Microbiol Mol Biol R* **2002**, *66* (3), 407-+.
151. Schmidt, J. C.; Zaug, A. J.; Cech, T. R., Live Cell Imaging Reveals the Dynamics of Telomerase Recruitment to Telomeres. *Cell* **2016**, *166* (5), 1188-+.