Artificial Cytoplasm: Crowding Components Causing Anomalous Diffusion

Elizabeth M. Seider

December 11, 2023
Updated, July 2024 by Sara Conti, Madeleine Petro, and Maya Kannan

The original thesis is available at
`https://physerver.hamilton.edu/People/vhorowitz/theses/Seider_original.pdf`

Faculty Advisor: Viva R. Horowitz

Hamilton College

Department of Physics

Submitted as a Partial Requirement for the Senior Project

Abstract

The cytoplasm of biological cells is a complex and crowded environment. Biomolecules, such as lipids and proteins, as well as organelles, make the cytoplasm an anisotropic medium. Anomalous diffusion is observed in the context of eukaryotic cells. Consequently, it is desirable to develop a statistical mechanical framework to better understand and describe how crowding affects diffusion in such an environment. Here, we complete three studies varying the concentration, molecular weight, and polydispersity of an artificial crowding agent, polyethylene glycol (PEG), to model how different components of molecular crowding affect diffusivity, as measured by the diffusive exponent α. We found that an increase in concentration and size of the crowding molecule leads to a decrease in diffusivity, and our findings suggest that polydispersity actually leads to increased diffusivity when compared to an approximately monodisperse solution of the same average molecular weight. Subdiffusion, a type of anomalous diffusion, was observed using PEG as an artificial crowding agent. Our results provide a detailed model for how different aspects of molecular crowding affect diffusion, and can be used to better understand how biomolecules affect diffusivity in an intracellular environment.

Contents

1 Acknowledgements

I would like to begin by thanking my thesis advisor, Viva Horowitz, for her expertise, mentorship, and support since I began working in her lab a year and a half ago. Not only would I not have been able to complete this thesis without her guidance, but I would not have the confidence that I now do to pursue a career in research without her help. She has pushed me and challenged me, all while giving me the tools necessary to succeed, and I am endlessly grateful.

I would also like to thank other professors in the Hamilton College physics department who have had a huge impact on my physics and academic journey. Kristin Burson for first giving me the belief that I could study physics. Kate Brown for encouraging me to stick with physics because I find it interesting and to stop worrying about whether or not I am smart enough (and also for giving me the hardest midterm exam that I have ever taken). Seth Major and Brian Collett for challenging me in the classroom and giving me excellent advice while serving as my faculty advisors.

I am also thankful for my classmates and friends in the physics department who have made all of the years taking classes together so enjoyable. Bess, I am particularly grateful to you for your years of friendship, and also for leaving the artificial cells project in such a great place.

Lastly, I would like to thank all of the wonderful people in my life outside of the physics department who have supported me and helped me get to where I am. Mom, Dad, Will, Charlie, and John, thank you for all of your love and support, as well as your willingness to pick up the phone for situations that are not nearly as catastrophic as I think they are. To my friends, thank you for encouraging me, supporting me, laughing with me, and especially Babbitt 32: thank you for listening every time I announced that I was starting my thesis over again.

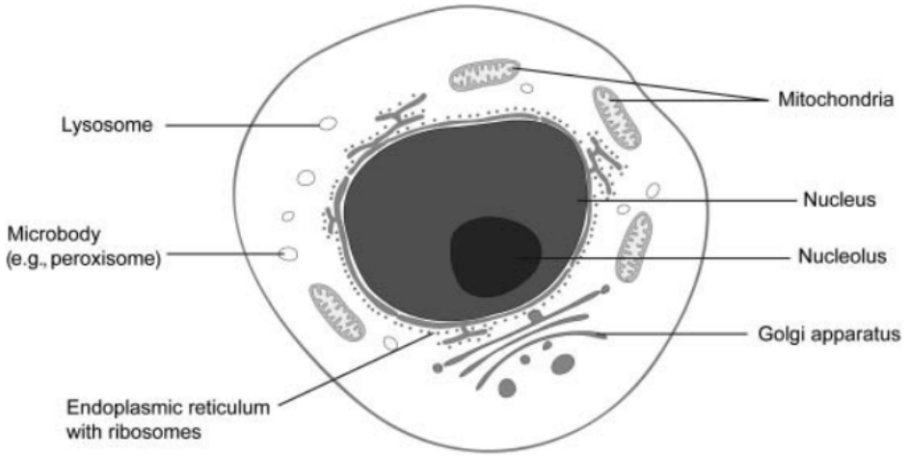

Figure 1: Diagram of organelles in the cytoplasm of a typical eukaryotic cell. Reprinted from [2].

2 Introduction

The cytoplasm of biological cells is a complicated and crowded environment, with well-documented non-Newtonian properties. Biomolecules, such as proteins and ribosomes, create macromolecular crowding within cells, and thus influence the physico-chemical properties of the cells via steric repulsion [1]. This crowding, and cellular structure, exists on many different length scales, with organelles such as the mitochrondria, endosomes, and Golgi apparatus (as seen in Fig. 1) found on the microscale [3]. Consequently, this study will focus on microscale crowding through the use of colloidal suspensions. Colloids have emerged as an important mechanism for studying biomimetic systems in experimental physics because they possess particles that are large enough to be viewed with an optical microscope, while still being small enough to succumb to thermal fluctuations [4]. The study of colloids in what will hereafter be referred to as artificial cells has the potential to provide a statistical mechanical framework for understanding the diffusive phenomena that occur in biological cells and that are consequently an essential property of life.

2.1 Passive Diffusion

Brownian motion (or passive diffusion) occurs when molecules randomly collide, which causes fluctuations from thermal equilibrium and subsequently the random motion of particles [5]. These random walks can be categorized by an exponent of diffusion $\alpha = 1$, as will be explained in the theory section.

2.2 Anomalous Diffusion

Anomalous diffusion is observed when a system has a diffusive exponent not equal to one, as will be explained in the theory section.

2.2.1 Subdiffusion

Subdiffusion due to cytoplasmic crowding is well-documented in biological contexts. Within the cytoplasm, it has been found that diffusional mobility is inversely related to the size of the tracked particle, and particles with a radius greater than $25-30$nm are immobile [3, 6–8]. Additionally, Saxton *et al.* used computer simulations to show that molecular mobility is hindered when the particle diffuses in a "maze-like environment" [9]. Banks *et al.* used flourescence correlation spectroscopy (FCS) to categorize anomalous diffusion of proteins in crowded solutions and they observed subdiffusion of proteins in dextran solutions with concentrations up to 400mg/mL [10]. They note that there is an indirect relationship between α and obstacle concentration and molecular weight [10], with α being the exponent of diffusion as will be explained in the theory section. In biological cells, biomolecules (or obstacles) present in concentrations that range from $80-400$mg/mL [1]. Thus, in order to create a colloidal system in which to study the relationship between cytoplasmic crowding and subdiffusion, it is necessary to introduce a crowding agent. In this study, the long chain polymer polyethylene glycol (PEG), with chemical structure shown in Fig. 2, will be used in varying concentrations, chain lengths, and mixes to better understand the effects of concentration, size, and diversity of biomolecules on diffusion

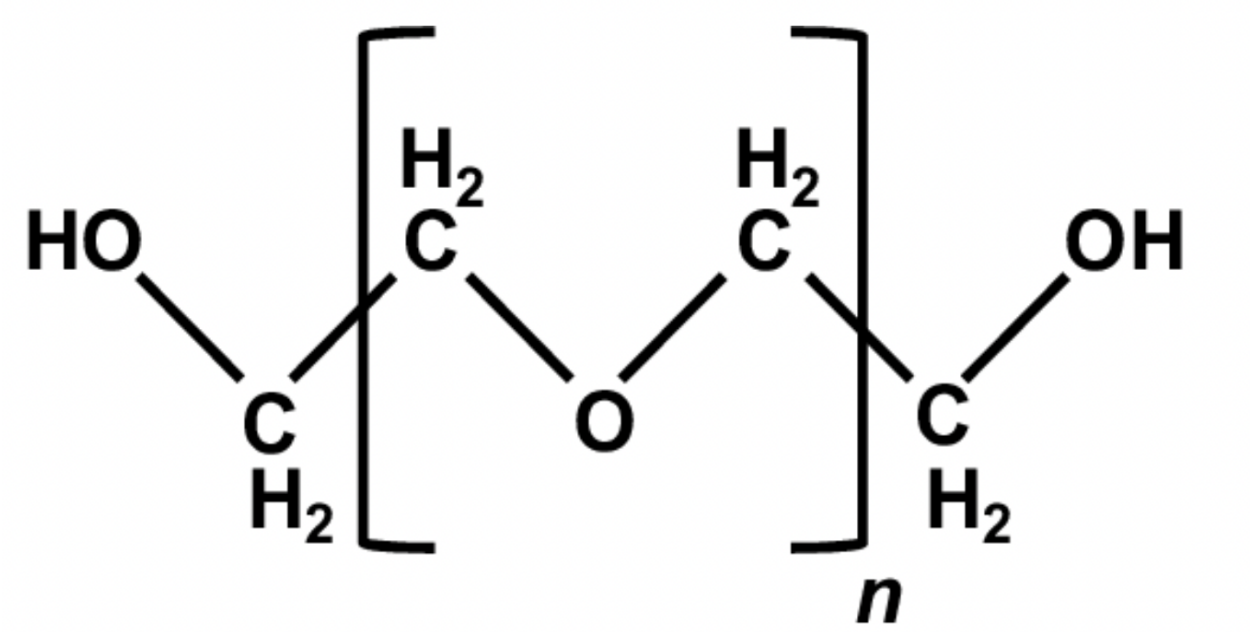

Figure 2: Chemical structure of the organic polymer polyethylene glycol (PEG). Reprinted from [11].

in biological cells.

2.2.2 Superdiffusion

Active colloids or "swimmers" have emerged as a promising way to understand the non-equilibrium properties of biological cells, where diffusion frequently occurs with a diffusive exponent greater than one. Examples of biological microswimmers include bacteria (which move in a characteristic "run and tumble" motion), unicellular protozoa, spermatozoa, and algae [5, 12, 13]. Additionally, for biomedical applications, it is essential to understand the behavior of swimmers in non-Newtonian fluids, as blood, synovial, and cerebrospinal fluids, as well as saliva, are all non-Newtonian [5]. The study of active swimmers not only illuminates the non-equilibrium physics of eukaryotic cells, but also has applications in health care, environmental sustainability, and security due to the ability of such swimmers to carry and deliver micro-and-nanoscale loads [5].

Active colloids can be achieved by a chemical reaction, and a variety of synthetic particles have

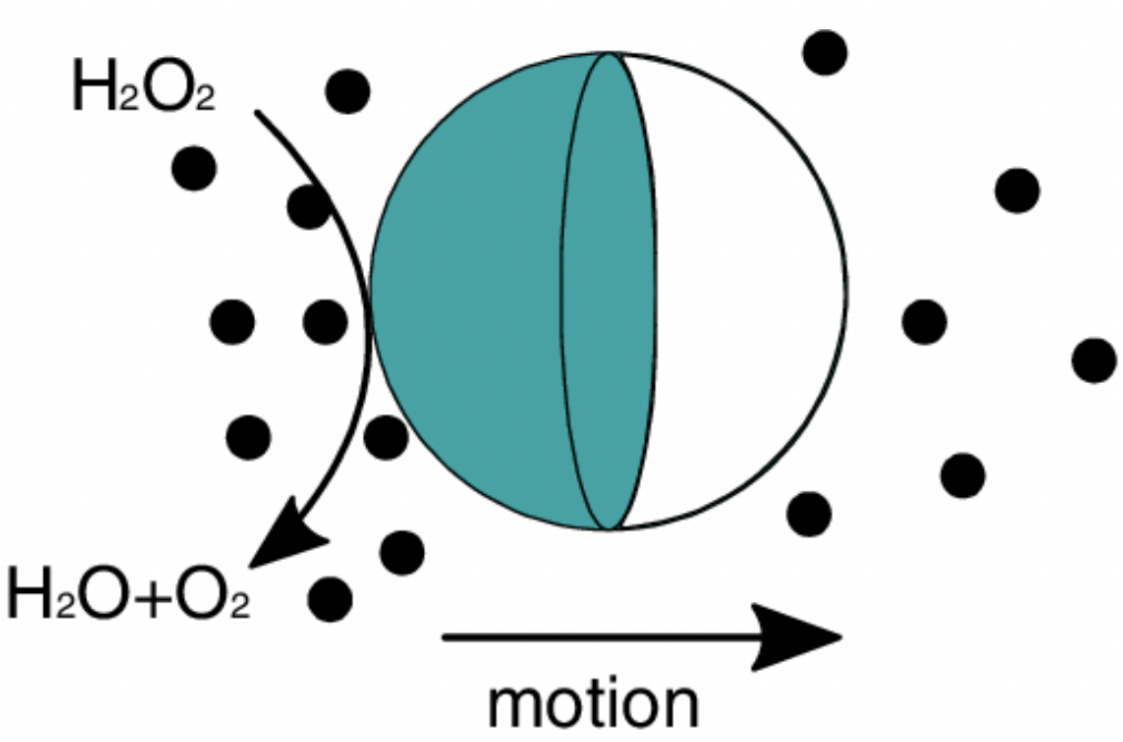

Figure 3: Diagram of the hydrogen peroxide reaction catalyzed on a Janus particle and the consequent self-diffusiophoresis. Reprinted from [18].

been used as "swimmers" in such a reaction. Paxton *et al.* used platinum-gold nanorods as their swimmers, while Fournier-Bidoz *et al.* used gold-nickel nanorods, and Mano and Heller used enzymes to create self-propelled motion via chemical reaction [14–16]. Howse *et al.* used spherical particles with an asymmetric coating of a catalyst to induce a reaction gradient that propels the particles by self diffusiophoresis [17]. In their study, they determined that the spherical particles are more advantageous than the rods, for which a reduction in the velocity by a factor of the aspect ratio occurs [17].

Janus particles, which are named for the two-faced Greek god, are a type of asymmetric, synthetic particle that are frequently used in the study of active colloids [13], and which will be used in this study. The particles consist of polystyrene particles that are half-coated in platinum (hence their "two-faced" name) that catalyze a reaction with hydrogen peroxide and consequently undergo self-diffusiophoresis due to the resulting pressure gradient [13, 17, 19] as

seen in Fig. 3. On short time scales, Howse *et al.* observed a direct correlation between velocity and concentration of fuel particles for their synthetic particles, in accordance with Michaelis-Menten kinetics frequently observed in biochemistry[1], while on longer time scales there is a return to the random walks characteristic of passive diffusion [17]. Additionally, Leptos *et al.* observed that an increase in the concentration of swimmers leads to the increased transport of passive particles as a result of more frequent interactions [21].

In our research, Janus particles can be used in an artificial cell environment consisting of a constant hydrogen peroxide bath in order to observe the effects of swimmers on diffusion for short time scales. Coupled with data on the effects of molecular crowding—simulated by the use of polyethylene glycol of varying molecular weights and in varying concentrations—on diffusion in artificial cells, we aim to use these synthetic means to create a framework for understanding factors affecting diffusion in biological cells through the lens of non-equilibrium physics and statistical mechanics.

2.2.3 Goals

The aim of this research is to provide a better understanding of how macromolecular crowding affects diffusion in the cytoplasm of eukaryotic cells using an artificial environment. Because this research builds a statistical mechanic framework for understanding biological processes, it has important applications in a variety of fields, from ecology to pharmacology. This research will build upon previous research, including that of Lawrence observing subdiffusion in polydisperse solutions [22]. In order to build a more robust understanding of the effects from macromolecular crowding, we will complete three studies that each vary one aspect of macromolecular crowding: the concentration of the crowding agent, the size of the crowding molecule, and the diversity of the crowding molecule(s). In order to vary these factors we used different concentrations and chain lengths of PEG, as well as solutions with different ratios of

[1] In biochemistry, Michaelis-Menten behavior is used to describe reactions where substrate binding and dissociation happen more quickly than product formation [20].

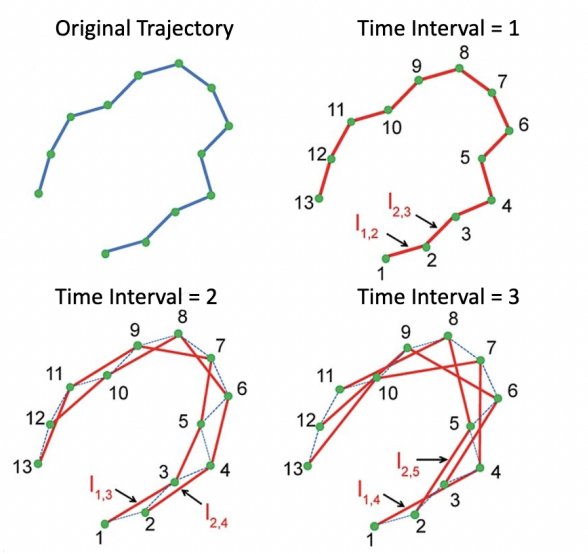

Figure 4: Mean square displacement from a single particle trajectory. Reprinted from [24].

different chain lengths of PEG for the polydispersity study. Our goal through the three studies is to build a better understanding of how the macromolecular crowding aspect studied in each affects diffusion in an artifical cellular environment, and ultimately in a biological one.

3 Theory

To build a statistical mechanical framework for the diffusive phenomena that occur in biological cells, it is necessary to begin with the theoretical relations that describe individual particle motion. In order to do so, we will derive the mean square displacement (MSD) and then the diffusion coefficient, before discussing anomalous diffusion as done by previous thesis students E. Lawrence and R. Smolarsky [22, 23].

3.1 Mean Square Displacement and The Einstein Relation

To develop our framework to describe diffusion, we begin with an essential component of the scaffolding: mean square displacement. As seen in Fig. 4, the mean square displacement can be found from averaging the trajectory of a particle over time intervals. As is the case in Fig.

4, the trajectory of each particle yields the individual mean square displacement (iMSD). For a group of particles, such as those suspended in a colloid, we can find the ensemble mean square displacement (eMSD), which is the average of all of the trajectories of the suspended particles. We now need to develop a relation between diffusion and MSD.

To find the Einstein relation for diffusion, we begin with Fick's second law, which can be derived from Fick's first law. Fick's first law is used to describe the flux, $\vec{J}$, across a surface in a region. It can be assumed, as with the analogous Fourier heat conduction law and equation for the coefficient of viscosity, that $|\vec{J}|$ is proportional to dv/dx, where v is the particle concentration [25]. Fick's first law is then given by [25]

$$J_x = -D\frac{dv}{dx} \tag{1}$$

where the constant of proportionality D is the diffusion coefficient. As seen in Eq. 1, Fick's law can only be used to describe steady state systems, thus creating the need for Fick's second law to describe time variant systems [26]. If we let v equal the time and position dependent function $f(x,t)$ and $J(x)$ be the flux at a position x, then we have [26]

$$\delta f(x) = \frac{[J(x) - J(x+dx)]\delta t}{dx}. \tag{2}$$

and

$$J(x+dx) = J(x) + \frac{\partial J}{\partial x}dx. \tag{3}$$

Substituting Eq. 3 into Eq. 2, we find that

$$\frac{\partial f}{\partial t} = -\frac{\partial J}{\partial x}. \tag{4}$$

We can then substitute for J using Fick's first law, given in Eq. 1, to arrive at Fick's second law:

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x^2} \tag{5}$$

where the constant of proportionality D is the diffusion coefficient. Note the time and position dependence of Eq. 5. Using the initial condition that each particle is at its origin at time $t = 0$, then we know that at some elapsed time $t = t$ particles have undergone some displacement $x + dx$, and the number of particles that have done so is given by $f(x,t)dx$ [27]. Using these conditions, changes in the function f must also satisfy Eq. 5, and so we have for $x < 0$ or $x > 0$ and $t = 0$ that [27]

$$f(x,t) = 0 \tag{6}$$

and

$$\int_{-\infty}^{\infty} f(x,t)dx = n \tag{7}$$

where n is the number of particles. Using the conditions given by Eq. 6 and Eq. 7, Einstein found the solution to the partial differential equation given by Eq. 5 to be [27]

$$n(x,t) = \frac{N}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}}. \tag{8}$$

It is then observed that Einstein's solution has a similar form to the general Gaussian curve function, which is given by

$$f(u) = Ce^{-\frac{u^2}{2\sigma^2}} \tag{9}$$

where C is a constant of proportionality, u is an arbitrary variable, and σ^2 is the expected deviation (variance). Using Eq. 9 as an analog to Eq. 8, we can observe that

$$4Dt = 2\sigma^2 \tag{10}$$

and thus

$$\sigma_x^2 = 2Dt. \tag{11}$$

Since D has units of distance per time, and we are multiplying by time, we can conclude from Eq. 11 that the variance represents the square of a particle's average displacement: the mean square displacement. Therefore, we can write the Einstein relation for mean square displacement in one dimension, x, as [27]

$$\langle x^2 \rangle = \sigma^2 = 2Dt. \tag{12}$$

Addition is a linear operator, so to find the mean square displacement for particles with a trajectory in multiple dimensions, we simply add that dimension. For a particle with motion in both the x and y directions, we can then extrapolate from Eq. 12 that the MSD is given by

$$\langle x^2 \rangle + \langle y^2 \rangle = 2Dt + 2Dt = 4Dt. \tag{13}$$

Following the same process and adding a third dimension, z [23], for the MSD in three dimensions we have

$$\langle r^2 \rangle = \langle x^2 \rangle + \langle y^2 \rangle + \langle z^2 \rangle = 2Dt + 2Dt + 2Dt = 6Dt. \tag{14}$$

Now that we have developed a relation for the MSD, from which the eMSD can be used to measure particle motion in the artificial cytoplasm, we need to find a relation for the coefficient of diffusion, D, in order to fully develop our model.

3.2 The Diffusion Coefficient and the Stokes-Einstein Relation

To derive the Stokes-Einstein relation we begin with Fick's second law, as given by Eq. 5. From the Maxwell-Boltzmann distribution, we know that the equilibrium distribution is also propor-

tional to $e^{-\frac{\alpha x^2}{2KT}}$ [28]. Thus, we arrive at the fluctuation-dissipation relation [28]

$$D = \mu k_B T \tag{15}$$

where μ is the mobility of the particle, T is temperature, and k_B is the Boltzmann constant. The mobility of the particle can be expressed in terms of Stokes' law, which is used to find the drag force on spherical objects in viscous fluids. Importantly, the law is for objects with a small Reynolds number, which is the ratio of inertial forces to viscous forces. Stokes' law is represented by the equation

$$F = 6\pi a \eta v \tag{16}$$

where η is the viscosity of the medium and a and v are the sphere's radius and velocity, respectively. We then find that [28]

$$\mu = \frac{1}{6\pi a \eta} \tag{17}$$

and can combine Eq. 17 with Eq. 15 to yield the Stokes-Einstein relation:

$$D = \frac{K_B T}{6\pi a \eta}. \tag{18}$$

Here we can see that if temperature and particle radius are held constant, the particle's diffusion coefficient is indirectly proportional to, and completely dependent on, the viscosity of the medium. Therefore, we can theoretically determine that in a more crowded, and consequently more viscous, medium, particles will have a lower coefficient of diffusion.

3.3 Anomalous Diffusion

The relations derived above prove accurate for isotropic fluids, which have direction-independent properties [10]. However, our goal is to build a model for diffusion in the cytoplasm, for which we cannot assume isotropy. In such cases, a power law can be implemented to fit empirical

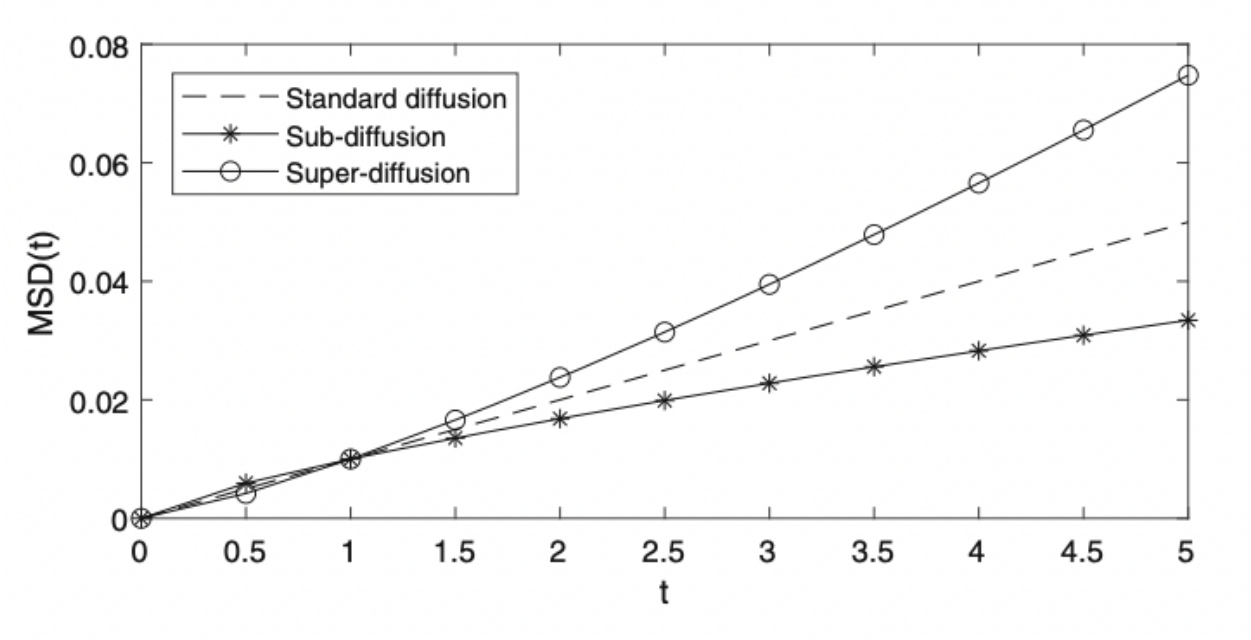

Figure 5: A log-log plot of MSD versus lag time for subdiffusion, normal diffusion, and superdiffusion. Reprinted from [29].

data. Combining this notion of a power law with Eq. 14, we find that

$$\langle r^2 \rangle = 6Dt^{\alpha} \tag{19}$$

where α is the exponent of diffusion, or diffusive exponent. This exponent of diffusion is essential in the development of our framework, as it can be used to characterize a system as one undergoing subdiffusion, normal diffusion, or superdiffusion. It then follows that a value of $\alpha = 1$ represents normal diffusion, where the system obeys the Stokes-Einstein relation, and any value $\alpha \neq 1$ is anomalous. For $\alpha < 1$ the system can be characterized as subdiffusive and for $1 < \alpha < 2$ the system can be characterized as superdiffusive [10]. As seen in Fig. 5, the log-log plot of MSD versus lag time is linear in the case of normal diffusion, and diverges at longer lag times for sub-and-superdiffusion. In the classical, constant velocity scenario where $x = vt$, $\alpha = 2$ represents ballistic motion given that $\langle r^2 \rangle \; \alpha \; t^2$ [30]. This study will primarily explore the conditions that cause subdiffusion in an attempt to build a model for eukaryotic biological cells.

Table 1: Tracer Particle and PEG Concentration for first test varying chain length.

Concentration (mg/ml)	Volume of Tracer Particle Solution (μL)	Volume of PEG 50mg/ml (μL)
40	2	8
30	4	6
20	6	4
10	8	2

4 General Methods

As previously mentioned, polyethylene glycol (PEG) was used to mimic biomolecular crowding withing the cytoplasm. It is desirable to know the effects of the molecular weight, the concentration of PEG, and polydispersity on diffusion, so the study was done in three parts.

4.1 Optimizing Tracer Particle Concentration

It was first necessary to determine the appropriate concentration of 1.1μm in diameter FluoroMax red fluorescent microspheres (1% solids; Thermo Scientific, USA), hereafter referred to as tracer particles. A rough estimate of the lower and upper bound of tracer concentration were determined to find an appropriate concentration for the studies. A more detailed review was later completed in order to fully understand the effects of varied tracer concentration. A first attempt used a 1 : 3 tracer particle and RO water solution to dilute 50mg/ml PEG solutions to the desired concentrations, as shown in Table 1 for PEG 200, 2000, 8000, and 20000. This approach proved problematic because the concentration of tracer particles was too high, which led to an image with indistinguishable particles and the possibility of bombardment occurring from the tracer particles themselves. Following this trial, it was determined that a lower concentration of tracer particles should be used.

Subsequently, PEG solutions were remade according to Table 1, but with a diluted tracer par-

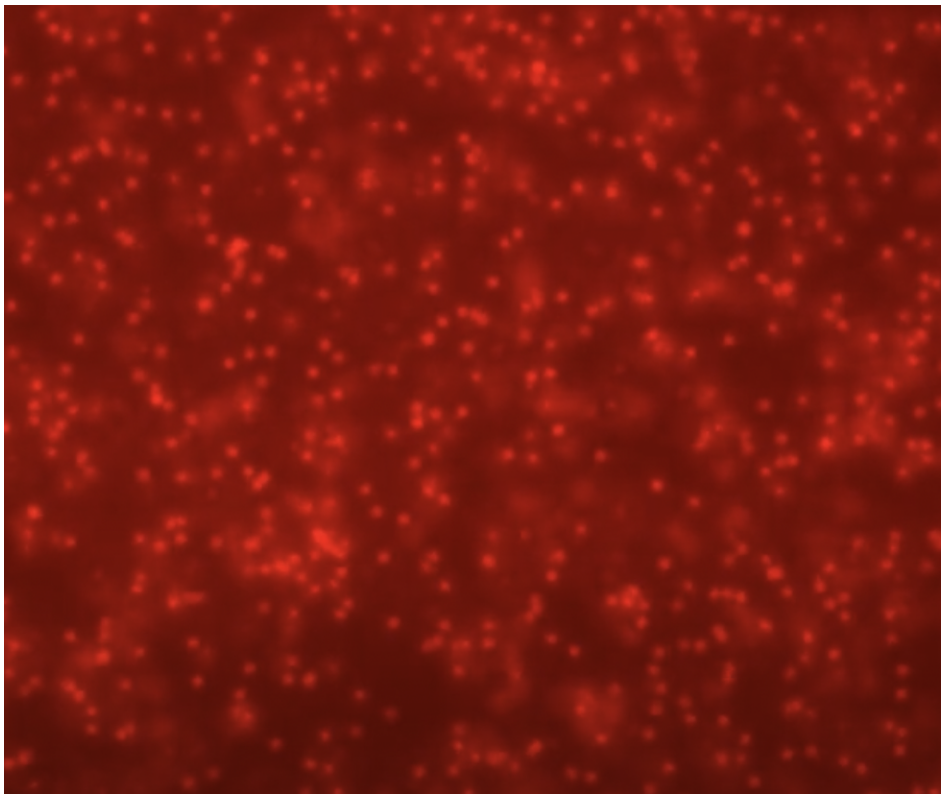

Figure 6: RO water with 0.04% tracer particles. Particles are 1.1μm.

ticle solution that was 1 : 99 of the original tracer solution (1%) and RO water. Upon imaging, we determined that this new concentration of tracer particles was too dilute; too few particles leads to low number statistics and consequently noise in the data. Additionally, we determined that the same concentration of tracer particles should be used, regardless of the concentration of PEG. Varying the concentration of tracer particles by using one tracer and RO water solution of constant concentration to dilute the PEG introduces an new, undesirable independent variable. Due to the length of time needed to take a large data run of the four PEG variants at four concentrations each, we began preparing and imaging a series of slides in order to determine the appropriate tracer particle concentration. An overall particle concentration of 0.04% (4% of the 1% Fluoro-Max solution) was found to have appropriate crowding of tracer particles, as seen in Fig. 6.

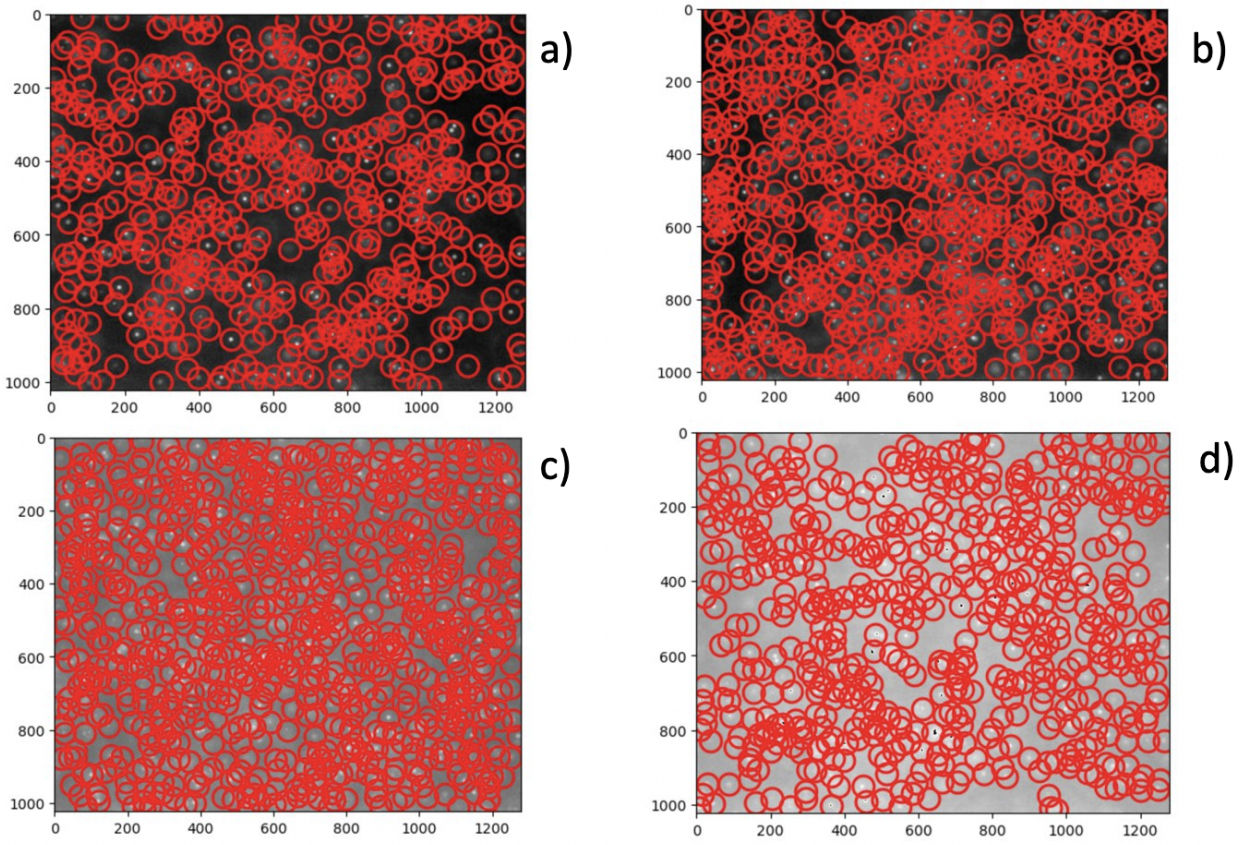

Figure 7: Tracer particle concentration of **a)** 0.02%, **b)** 0.03%, **c)** 0.04%, and **d)** 0.05%. Red circles are Trackpy's identification of particle centers. Particles are 1.1 μm.

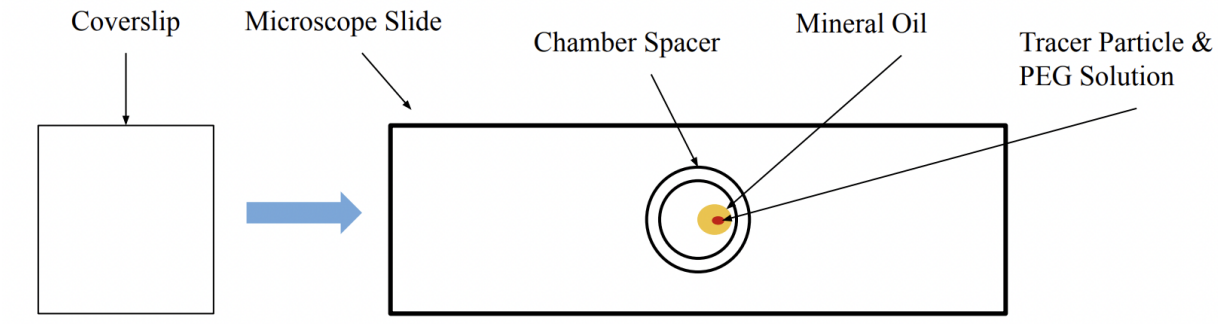

Figure 8: A top-down view of slide construction with mineral oil.

A later study was completed to further hone in on a desirable tracer particle concentration. Tracer particle concentration was varied in a sample with just RO water in order to find a concentration that was high enough to yield quality videos for analysis (without being oversaturated), but in which the tracer particles themselves were not functioning as a crowding agent. This study, with diffusive exponent data included in Appendix 1, helped us determine that tracer concentration is optimized for imaging between 0.02% and 0.04% tracer particles as seen in Fig. 7. Below 0.02% tracer particle concentration, there would not be enough particles to track and we would once again run into the error of noise in the data. At 0.05% tracer particle concentration, the videos were oversaturated and Trackpy was not able to appropriately determine the center of each particle, as seen in Fig. 7d. Additionally, the study was completed using both cold solutions (chilled in a refrigerator) and room temperature ones. No trend was observed between the cold and room temperature data sets. For consistency, the three crowding component studies were completed using room temperature samples.

4.2 Slide Preparation

Solutions were prepared as prescribed in the methods section of each study. Slide preparation was the same for all studies. Solutions that had been refrigerated were allowed to come to room temperature and slides were then constructed as seen in Fig. 8. Four microliters of mineral oil were pipetted onto the slide prepared with a plastic chamber spacer, and 2μl of the tracer par-

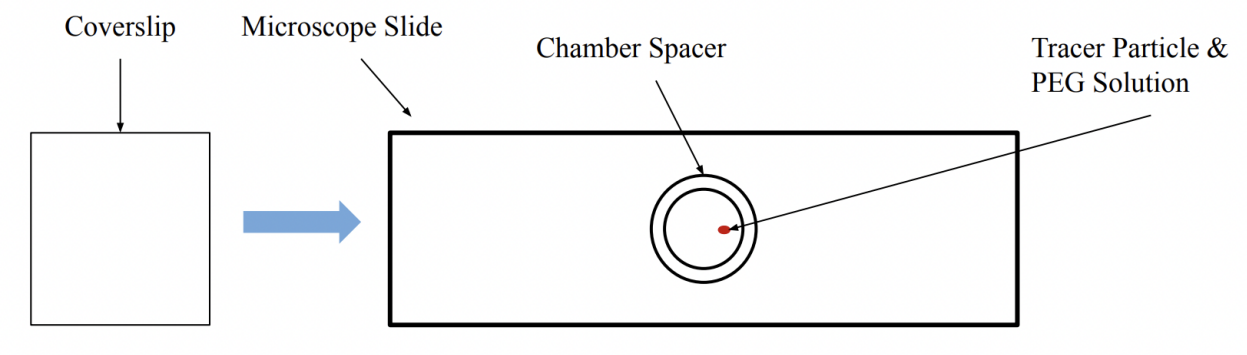

Figure 9: A top-down view of revised slide construction, without mineral oil.

ticle and PEG solution were then pipetted into the mineral oil. It was later observed that slides constructed in this manner yielded a likely systematic error. Thus a new approach was used, as seen in Fig. 9, where the slides were constructed in the same manner, but without the mineral oil and with 4μL of the tracer particle and PEG solution. This second approach is appropriate because there is the potential that the mineral oil creates a diffusion gradient with the PEG at high concentrations, leading to superdiffusion. Additionally, it is unnecessary to use mineral oil to prevent evaporation when slides are imaged immediately after construction, rather than being stored and then imaged.

4.3 Imaging

Slides were imaged using a Nikon Ti-E Inverted microscope. A quad filter was used for fluorescence microscopy. Videos were taken using a 40x objective at a frame rate of 20FPS with a Thorlabs camera. In order to avoid systematic error introduced by the adhesion of particles to the slide glass or the coverslip, videos were taken in the middle of the slide as calculated by finding the working distance to the top and bottom layers of particles on the coverslip and slide glass, respectively. Three videos, each 1000 frames long and at different positions in the same z-plane of the cell, were taken for all of the samples in order to provide an uncertainty in the

Table 2: Solutions for Concentration Study

PEG Volume Percent	Volume of PEG 200 (μL)	Volume of Fluoro-Max Solution (μL)	Volume of RO Water (μL)
96	480	20	0
86	430	20	50
76	380	20	100
66	330	20	150
50	250	20	230
26	130	20	350
0	0	20	480

measurements and help quantify random error. The python library Trackpy was used to locate the centers of the red fluoromax particles and track the trajectories. Each particle's trajectory yields the iMSD. Translational and rotational drift in three dimensions were modeled using a singular value decomposition (SVD) method based on the particles' trajectories in two dimensions [30]. Rotational drift was then subtracted to produce drift-subtracted values which are then used to calculate the eMSD.

5 Concentration Study

5.1 Methods

PEG 200 was selected for the concentration study because it is liquid at room temperature, allowing for high concentrations to be reached due to the lack of a solubility barrier. Solutions were prepared using volume fractions, with the most concentrated being 96% PEG 200. This value was chosen to still allow for the 4% of the 1% Fluoro-Max solution that was previously deemed necessary for imaging. Solutions were made as shown in Table 2. We chose to focus on the higher concentrations, as this is where theory and literature dictates subdiffusion is more likely to occur. The solution without PEG was used as a control, as RO water should undergo passive diffusion.

5.2 Results and Discussion

The results of our concentration experiment are not surprising when an analogy is drawn between the artificial cytoplasm and a room full of people, say at a cocktail party. The more people there are in the room, the harder it would be to move from one side of it to the other. We found the same when varying the concentration of PEG 200.

As previously mentioned, PEG 200 was selected for the concentration study because of its liquid phase at room temperature. This solubility allowed us to test high concentrations, with the results shown in Table 7.

The clear relationship between eMSD and concentration of the crowding agent is shown in Fig. 10, which is a comparative plot of the drift-subtracted eMSD versus lag time for each sample. The median value among the three videos for each sample was chosen to visualize the difference in eMSD between samples, as shown in Fig. 10. This was done in order to reduce crowding on the plot, on the premise that there is very little variability between the three videos for each sample. As seen in Fig. 10, the eMSD decreases as the concentration increases. Like in the trivial cocktail party example, more crowding makings it harder to move, be it people or tracer particles. The correlation is also non-linear, as seen by the fact that the eMSD for 50% PEG by volume is closer to that of the 96% by volume sample than the 26% sample, particularly at longer lag times. As shown in Eq. 19 of the theory section, the diffusive exponent can be then be extrapolated from the MSD.

As seen in Table 7, for the trial with 96% PEG 200 by volume we determined the diffusive exponent, a unitless quantity, to be 0.644 ± 0.009. As expected, this result is highly subdiffusive. The diffusive exponent was found to be 0.674 ± 0.015 for 86% by volume, 0.75 ± 0.02 for 76%,

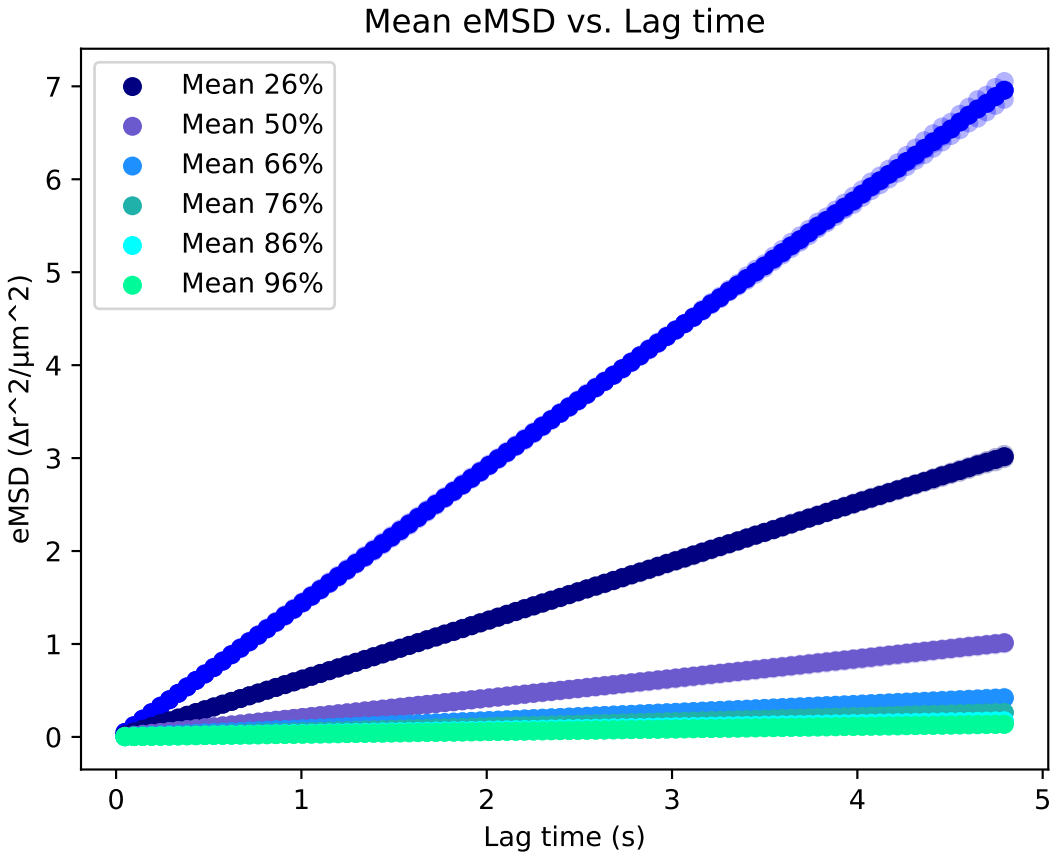

Figure 10: The drift-subtracted eMSD for different concentrations of PEG 200 (given as a percent volume) as a function of lag time. The opaque markers represent the mean eMSDs, while the translucent markers show the eMSDs for each video. (Updated by Sara Conti)

Table 3: Average diffusive exponent for different concentrations of PEG 200. (NOT UPDATED)

Percent PEG by Volume	Diffusive Exponent (α)	Uncertainty
96	0.644*	0.009
86	0.674*	0.015
76	0.75*	0.02
66	0.792*	0.004
50	0.898*	0.011
26	0.982	0.006
0	1.013	0.003

*Indicates samples where the power law breaks down due to extremely anomalous diffusion.

0.792 ± 0.004 for 66%, 0.90 ± 0.01 for 50%, 0.982 ± 0.006 for 26%, and 1.013 ± 0.003 for the RO water control (0% PEG). All samples other than the control were found to be subdifussive, with a low amount of uncertainty. The sample with the highest percent uncertainty was the 76% PEG 200 by volume sample, which has an uncertainty of 2.7%. The overall low amount of fractional uncertainty allows us to be confident about the negligible effects of random error in the concentration study. The diffusive exponent of the control sample is slightly superdiffusive which indicates a possible systematic error, since we know from the Stokes-Einstein relation that water should undergo passive diffusion with $\alpha = 1$. This potential superdiffusivity will be discussed further in a later section. There is also additional uncertainty, not reflected in the error bars, for the points circled in green in Fig. 19 due to the breakdown of the power law in diffusive environments that are highly anomalous. [Viva's note: This is not correct.]

Figures 11-17 show a log-log plot of the eMSD as a function of lag time. From these figures, we can observe that the power law holds for the control sample and for the 26% by volume sample. However, from Figures 13-17 we see that as concentration increases the power law fit breaks down more and more. Figure 18 shows a comparative log-log plot of the eMSD as a function of lag time for the different concentrations, which further highlights that diffusion becomes increasingly anomalous as concentration increases. In these figures we actually see multiple behaviors on different time scales, and thus one fit, be it linear or others, would not suffice. The break down of the power law signifies that we have achieved significantly anomalous diffusion at high concentrations.

As seen in Fig. 19, where we plot the diffusive exponent versus the percent PEG 200 by volume, there is a clear relationship between how crowded our artificial cell is and how diffusion occurs in the medium. We fit this trend with a second degree polynomial curve that has a coefficient of determination, R^2, of approximately 0.98, indicating very little variance between our data and the trendline. This curve can be used to predict the subdiffusive nature of a solution upon

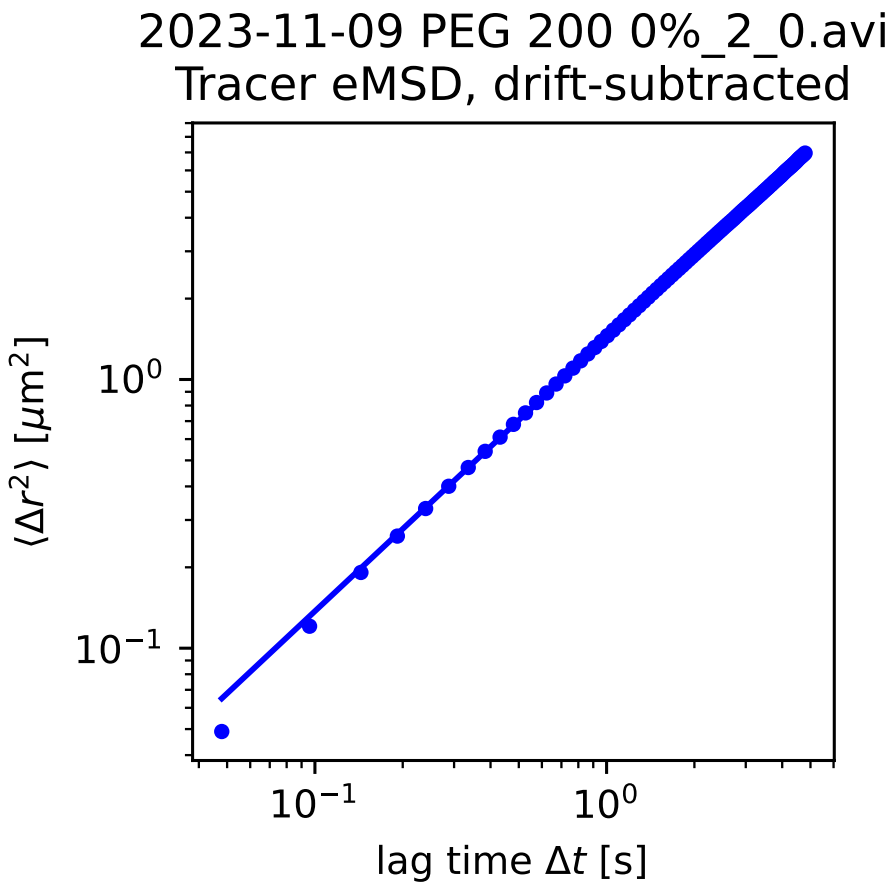

Figure 11: A log-log plot of the eMSD as a function of lag time for the second video of the RO control. The power law fit is shown by the solid blue line. (Updated by Sara Conti)

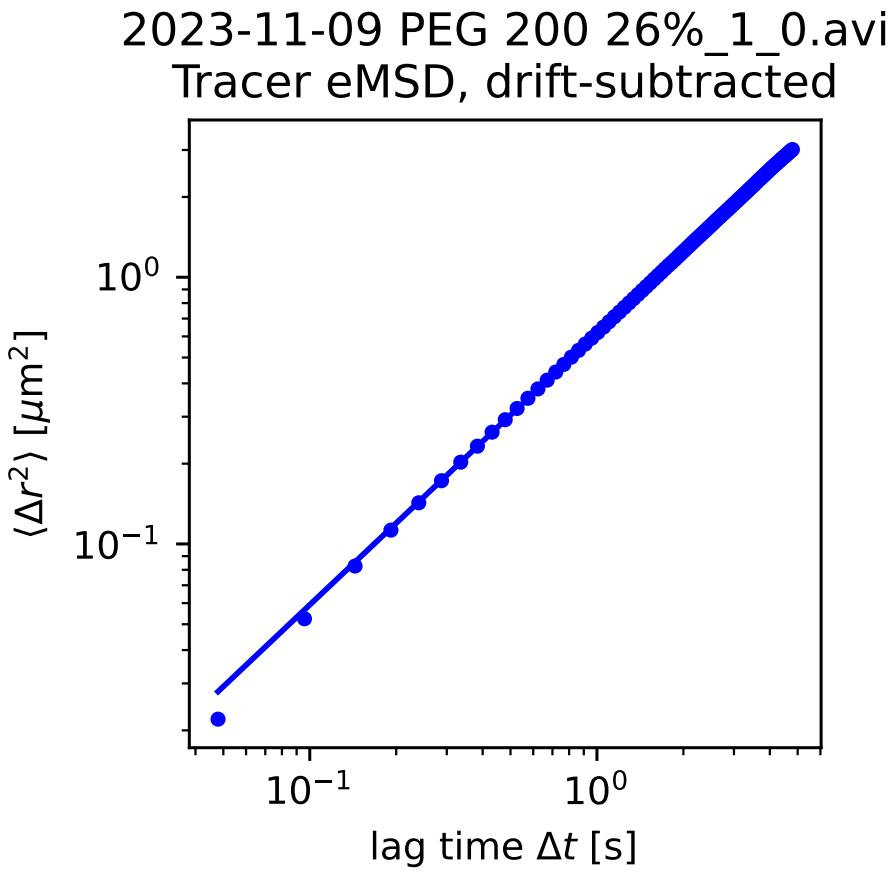

Figure 12: A log-log plot of the eMSD as a function of lag time for the first video of PEG 200 at 26% by volume. The power law fit is shown by the solid blue line. (Updated by Sara Conti)

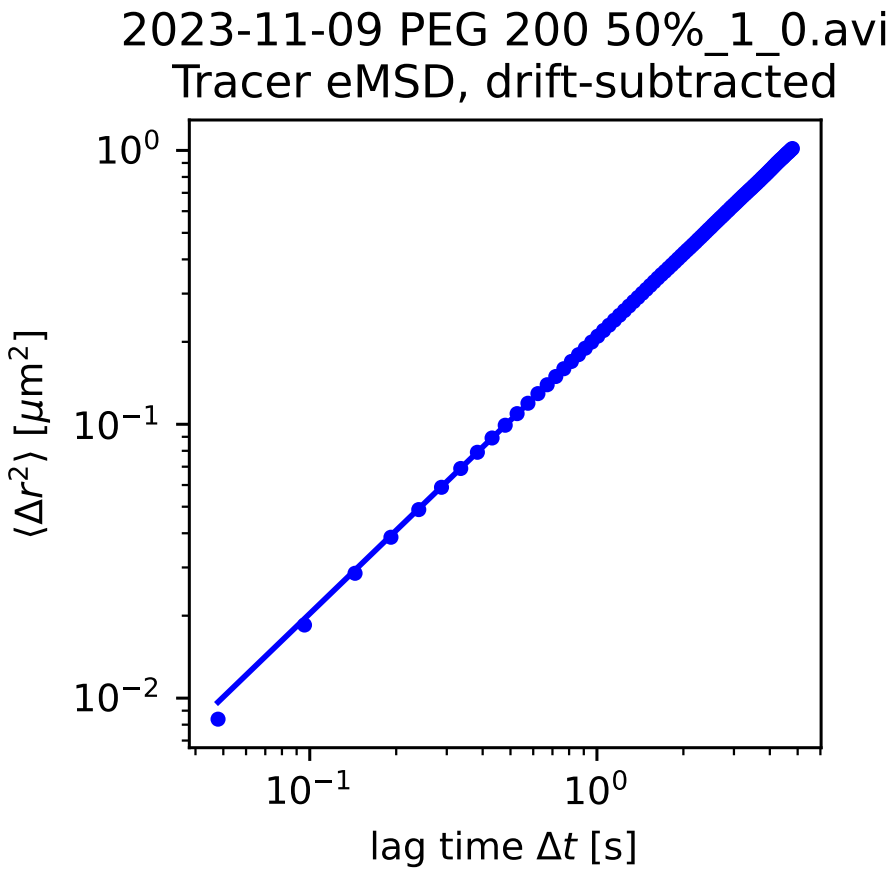

Figure 13: A log-log plot of the eMSD as a function of lag time for the first video of PEG 200 at 50% by volume. The power law fit is shown by the solid blue line. (Updated by Sara Conti)

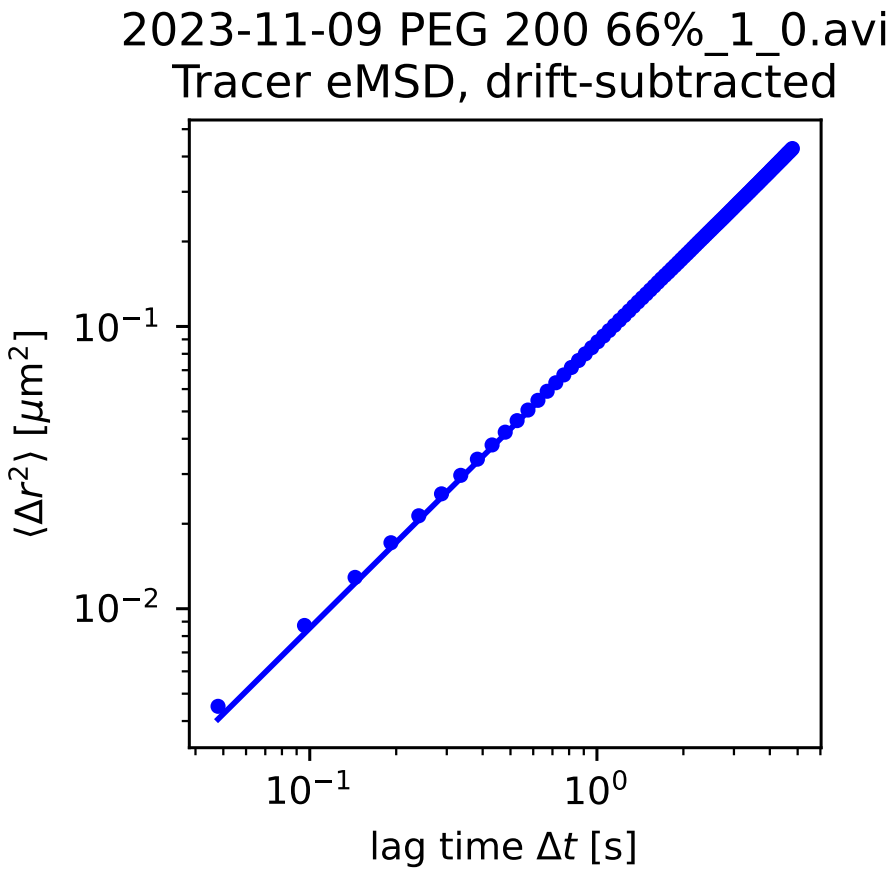

Figure 14: A log-log plot of the eMSD as a function of lag time for the first video of PEG 200 at 66% by volume. The power law fit is shown by the solid blue line. (Updated by Sara Conti)

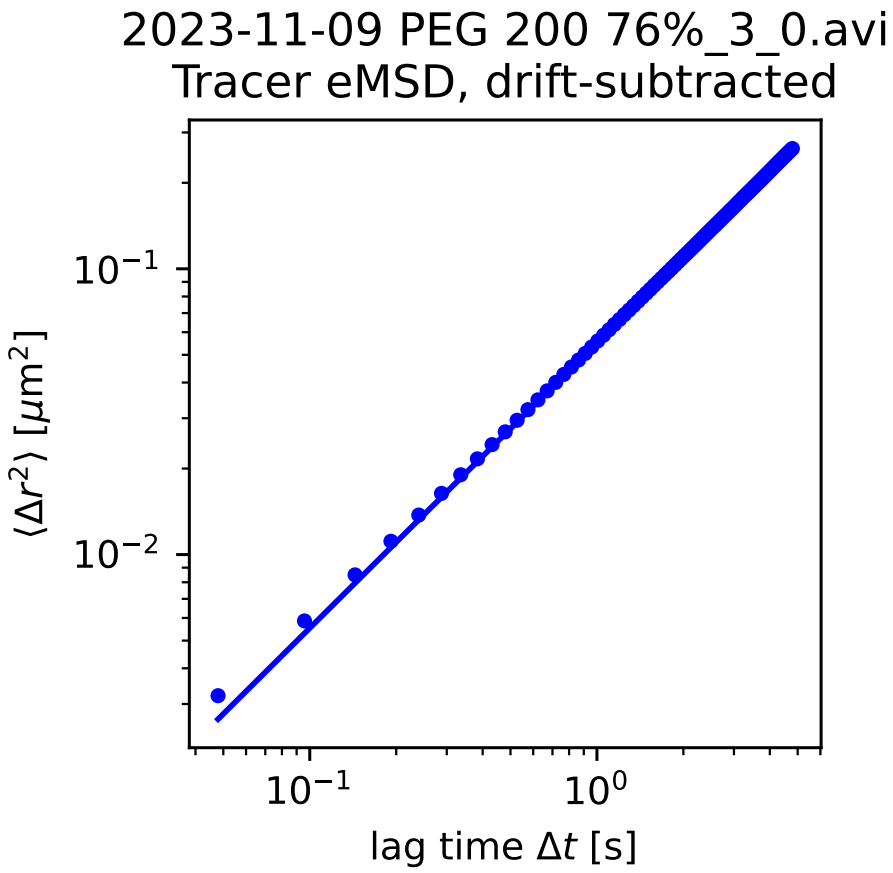

Figure 15: A log-log plot of the eMSD as a function of lag time for the first video of PEG 200 at 76% by volume. The power law fit is shown by the solid blue line. (Updated by Sara Conti)

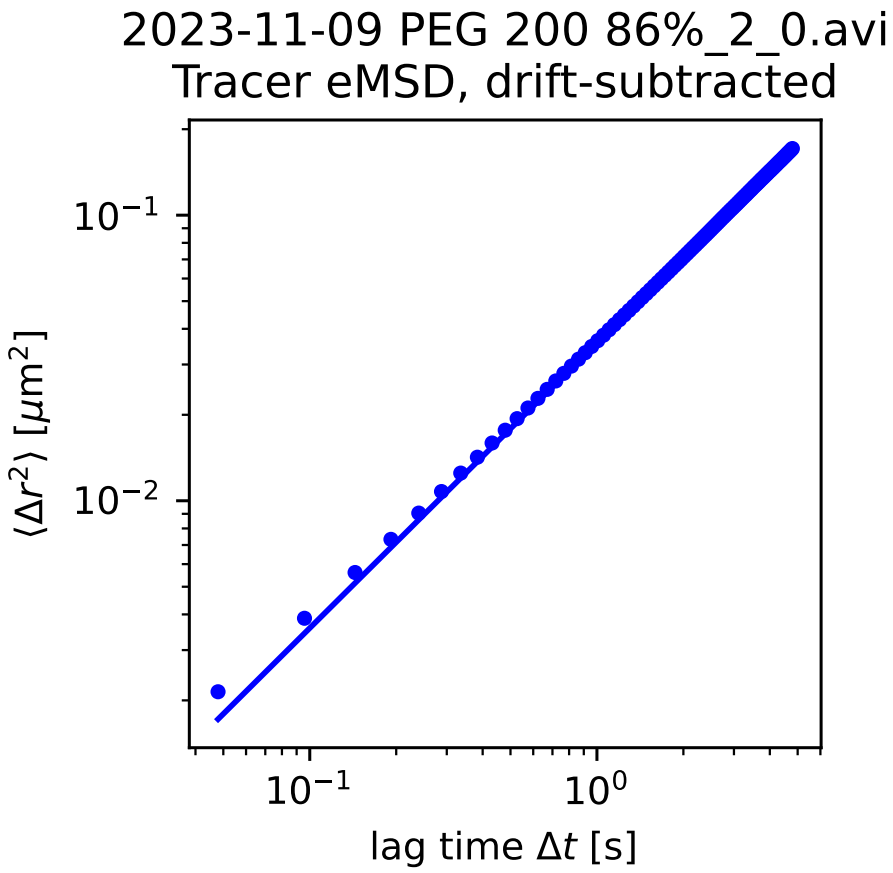

Figure 16: A log-log plot of the eMSD as a function of lag time for the second video of PEG 200 at 86% by volume. The power law fit is shown by the solid blue line. (Updated by Sara Conti)

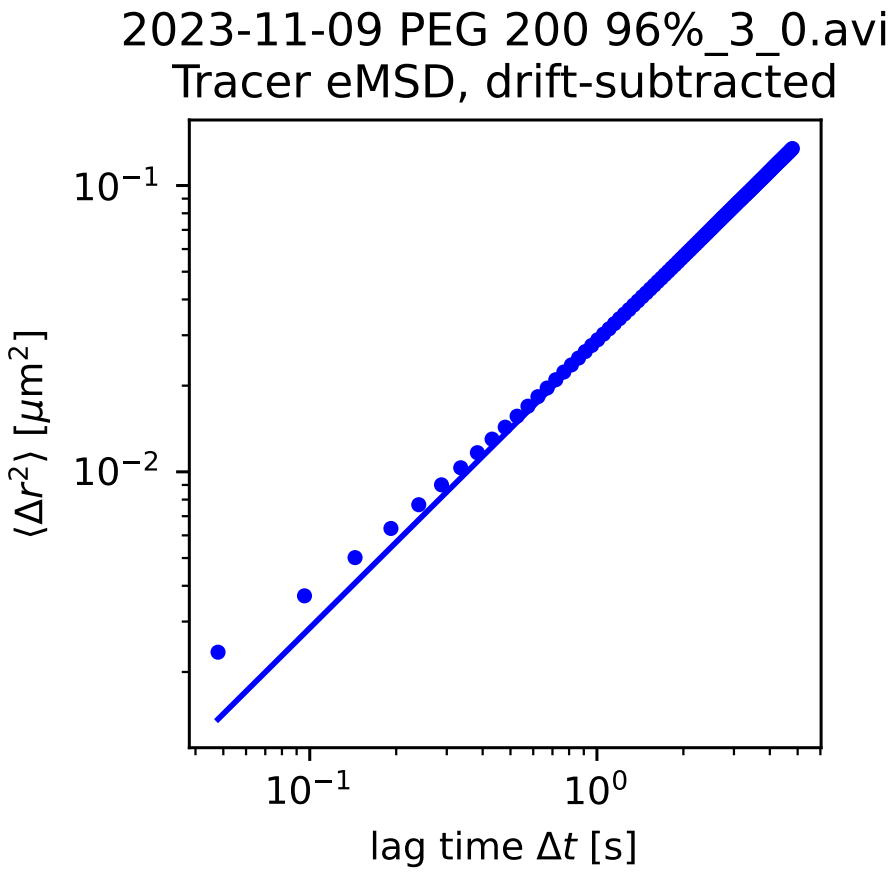

Figure 17: A log-log plot of the eMSD as a function of lag time for the third video of PEG 200 at 96% by volume. The power law fit is shown by the solid blue line. (Updated by Sara Conti)

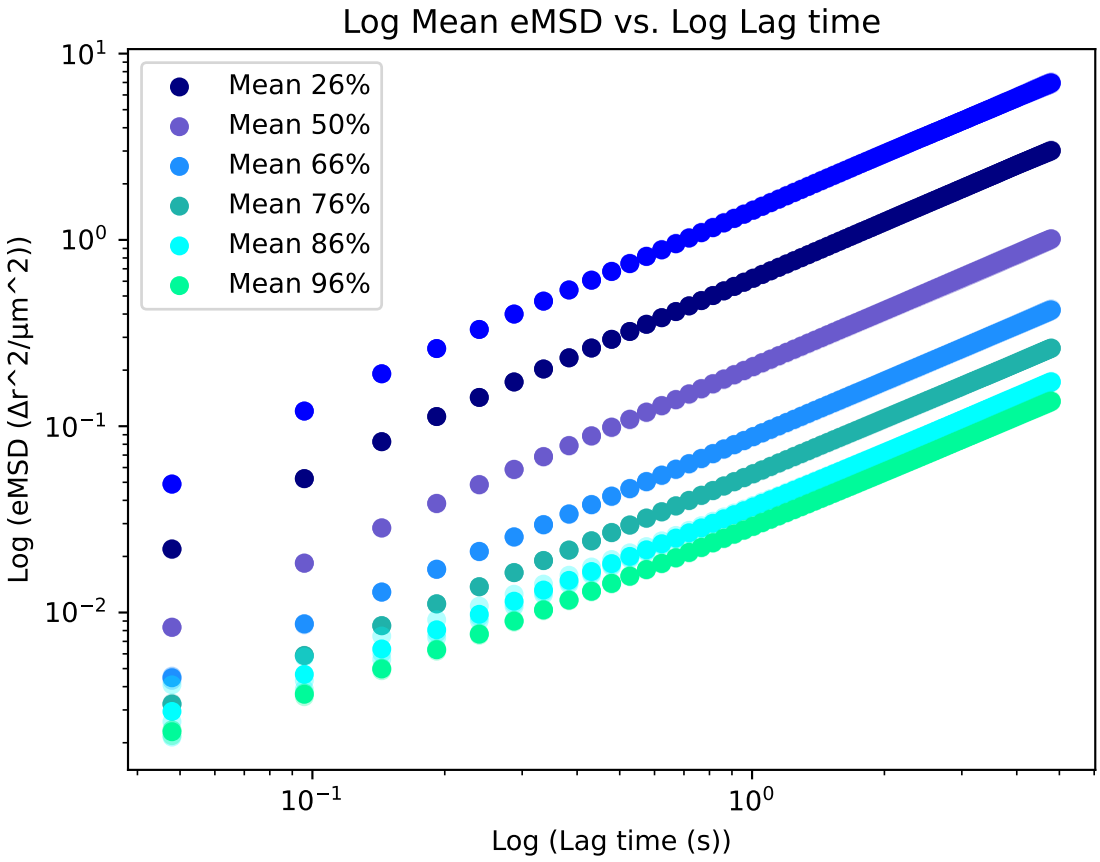

Figure 18: A comparative log-log plot of drift-subtracted eMSD for different concentrations of PEG 200 (given as a percent volume) as a function of lag time. The opaque markers represent the mean eMSDs, while the translucent markers show the eMSDs for each video. (Updated by Sara Conti)

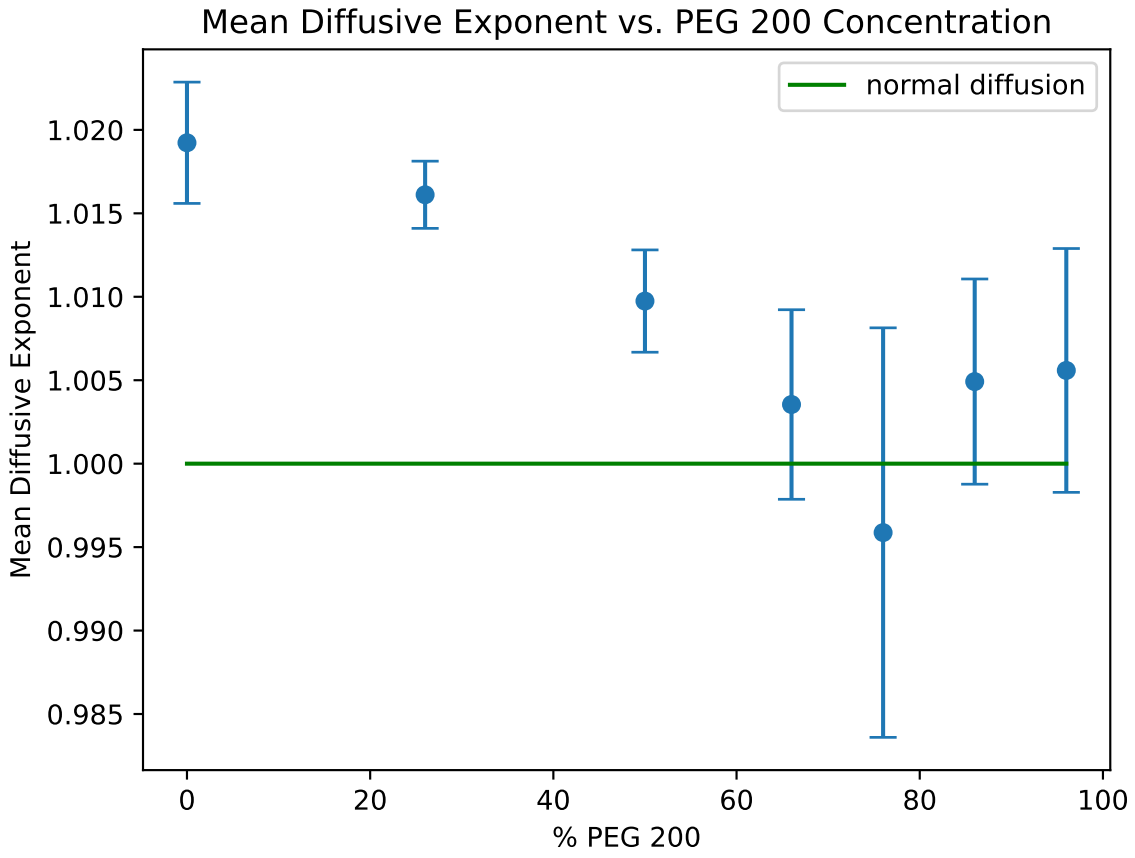

Figure 19: Diffusive exponent α as a function of concentration of PEG 200 (given as a percent volume). The green line represents $\alpha = 1$ (passive diffusion). Error bars are from the range method for the three videos. (Updated by Sara Conti) [Viva update: the superdiffusion is not reproducible, and may depend on the choice of max lag time for the MSD analysis.]

Table 4: Average excess kurtosis for different concentrations of PEG 200. (Not yet updated)

Percent PEG by Volume	Excess Kurtosis	Uncertainty
96	28.12	0.86
86	25.19	0.93
76	21.8	1.5
66	19.3	0.7
50	13.2	0.7
26	7.10	0.65
0	2.04	0.14

introduction of a crowding agent.

Our model uses a power law to fit the eMSD, as derived in the theory section. However, as seen for Fig. 17, the data diverges strongly from the power law fit, particularly at shorter lag times. A log-log plot of the drift-subtracted eMSD versus lag time with a power law fit for every sample in this study can be found in Appendix 2. The Van Hove plot (seen in Fig. 20 for the same sample as Fig. 17), which gives a histogram of particle displacements for a certain lag time, contains "fat tails." These tails are indicative of a propensity to produce outliers, which can be measured by the kurtosis, κ [31]. It is commonly believed that kurtosis is a measure of "peakedness," but it is actually a quantification of the tail behavior of a distribution [31]. If a system has an excess kurtosis value greater than three, it is said to be leptokurtic [31].

A plot of the excess kurtosis versus concentration is shown in Fig. 21, and the values can be found in Table 4. The plot is shown with a linear fit that has an R^2 valued of 0.99, indicating that the excess kurtosis has a strong linear dependence on the concentration of the crowding agent. This linear dependence highlights that as concentration increases, the diffusion becomes increasingly anomalous.

The relationship that we have observed is fundamental in developing a framework for understanding how molecular crowding affects diffusion in eukaryotic cells. In biological cells,

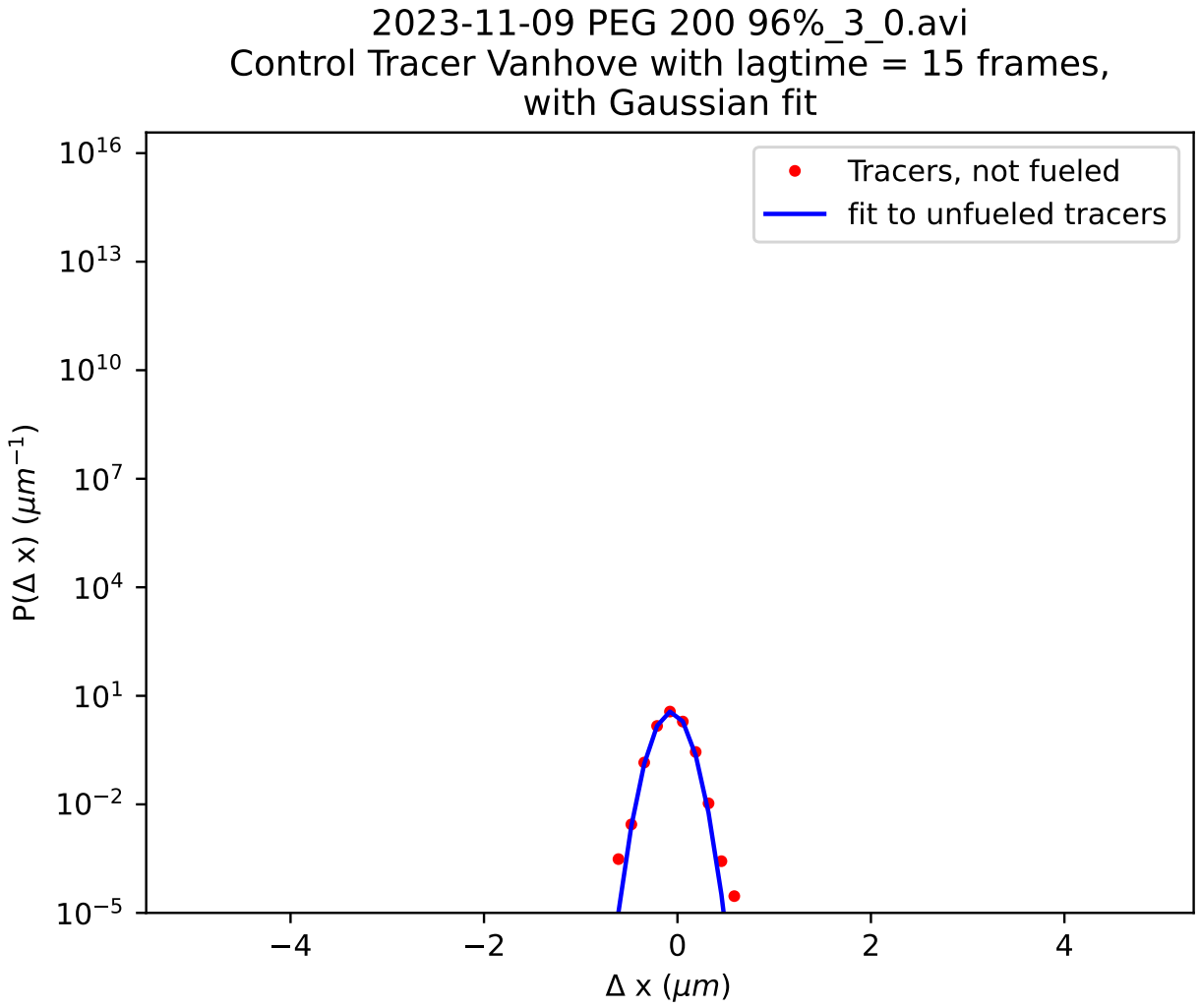

Figure 20: Van Hove plot at a lag time of 0.72 seconds for the third video of PEG 200 at 96% by volume. Shown with a Gaussian fit in blue. (Updated by Sara Conti)

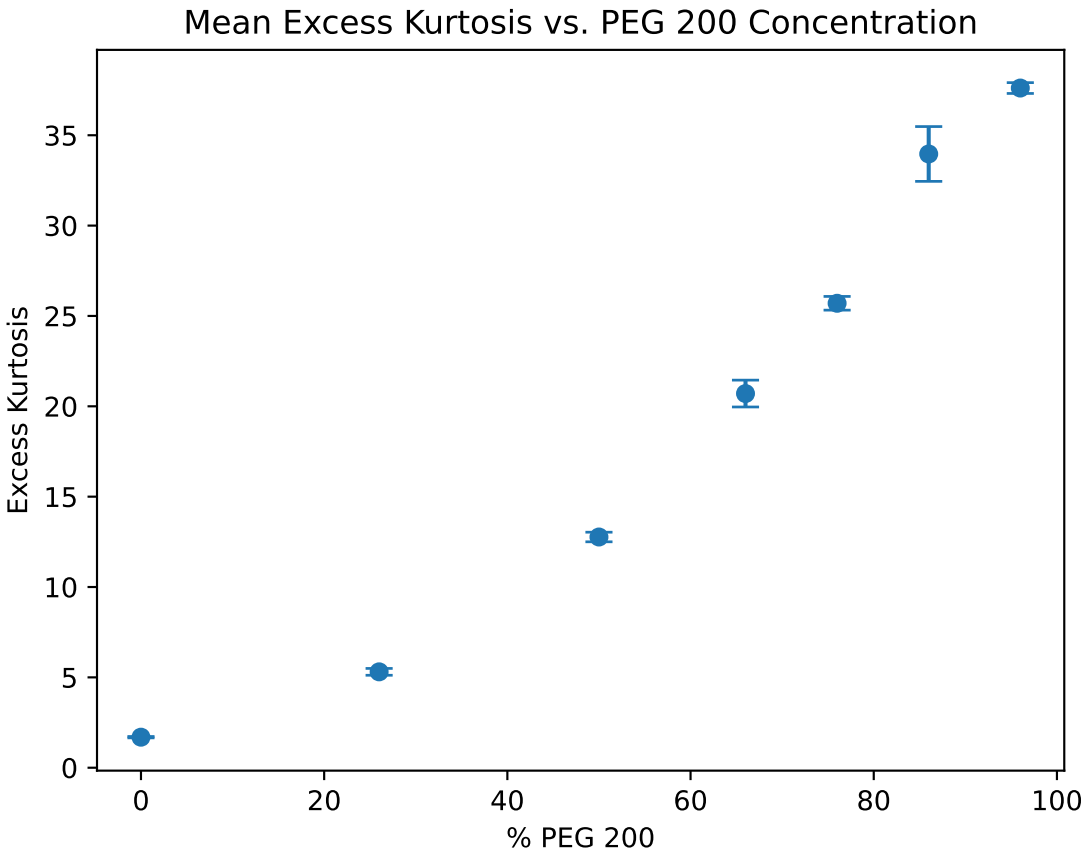

Figure 21: Excess kurtosis as a function of concentration for different volume percents of PEG 200. All samples containing PEG are shown to be leptokurtic ($\kappa > 3$). (Updated by Sara Conti)

proteins, lipids, and sugars occupy up to 40% of cytoplasmic volume [3, 32]. While our study primarily focuses on higher concentrations, we also have a sample at a lower volume fraction (26%). We were able to find a clear relationship between both diffusive exponent and PEG 200 concentration, as well as between excess kurtosis and PEG 200 concentration, exhibiting a successful model for how concentration of biomolecules affects diffusion in biological cells.

6 Molecular Weight Study

6.1 Methods

In order to identify the effect of the size of biomolecules on diffusion within the cytoplasm, for the first part of our study we held the concentration of PEG constant at 40mg/ml while varying the molecular weight of PEG (PEG 200, 2000, 8000, and 20000). Solutions of each molecular weight were prepared by vortexing a 50mg/ml PEG solution, RO water, and the 1% Fluoro-Max tracer particle solution in a $20:1:4$ ratio, respectively. Slides were then prepared as shown in Fig. 9.

6.2 Results and Discussion

Our second study examined the affects of the size of the crowding agent. Since the length of the PEG molecule is directly related to its mass, the higher the molecular weight the longer the chain. To begin our analysis of the affects of chain length on diffusion, we first look at the eMSD. A comparative plot of the drift-subtracted median eMSD for each molecular weight is shown in Fig. 22. Once again the median was comfortably selected for visualization due to a low variability between videos for the same sample. From this plot we can observe a clear relationship between mobility and chain length, particularly at longer lag times. As chain length increases, the eMSD over a given lag time decreases. From the log-log plot of the drift-subtracted eMSD versus lag time given in Fig. 23, we can see that the lower average molecular weights obey the

Table 5: Average diffusive exponent for different molecular weights of PEG. (Not yet updated)

PEG Molecular Weight (g/mol)	Diffusive Exponent (α)	Uncertainty
200	1.01	0.01
2000	0.99	0.03
8000	0.987	0.007
20000	0.953	0.013

power law more closely and are consequently not anomalous. The drift subtracted eMSD as a function of lag time with a power law fit for every sample in this study can be found in Appendix 3.

Due to the relationship between eMSD and diffusive exponent, it is unsurprising that a lower diffusive exponent is observed for longer chain lengths. As seen in Table 24, the diffusive exponent for PEG 200 was found to be 1.01 ± 0.01, right around the passive diffusive exponent $\alpha = 1$. The diffusive exponent was found to be 0.99 ± 0.03 for PEG 2000, 0.987 ± 0.007 for PEG 8000, and 0.953 ± 0.013 for PEG 20000. The fractional uncertainty is highest for the PEG 2000, but still relatively low with a value of 3.0%.

The diffusive exponent is plotted versus molecular weight in Fig. 24. From this figure, we can deduce that there is a negative linear relationship between diffusive exponent and molecular weight. The data is fit with the line $\alpha = -(2.78 \times 10^{-6})x + 1.01$ where x is the molecular weight in g/mol, and has an R^2 value of 0.96. As we increase the molecular weight, the diffusive exponent decreases. While the values for the diffusive exponent are close for all of the samples, we are confident that there is in fact a decreasing linear relationship because there is no overlap in the error bars between the PEG 200 and PEG 20000 samples. Additionally, those error bars give two standard deviations , which is a 95% confidence interval.

Our work determining a relationship between size of the crowding agent and diffusive exponent is a further step towards understanding how different biomolecules affect cytoplasmic diffusion. Organelles, and other crowding features, vary greatly in size and length-scale. By

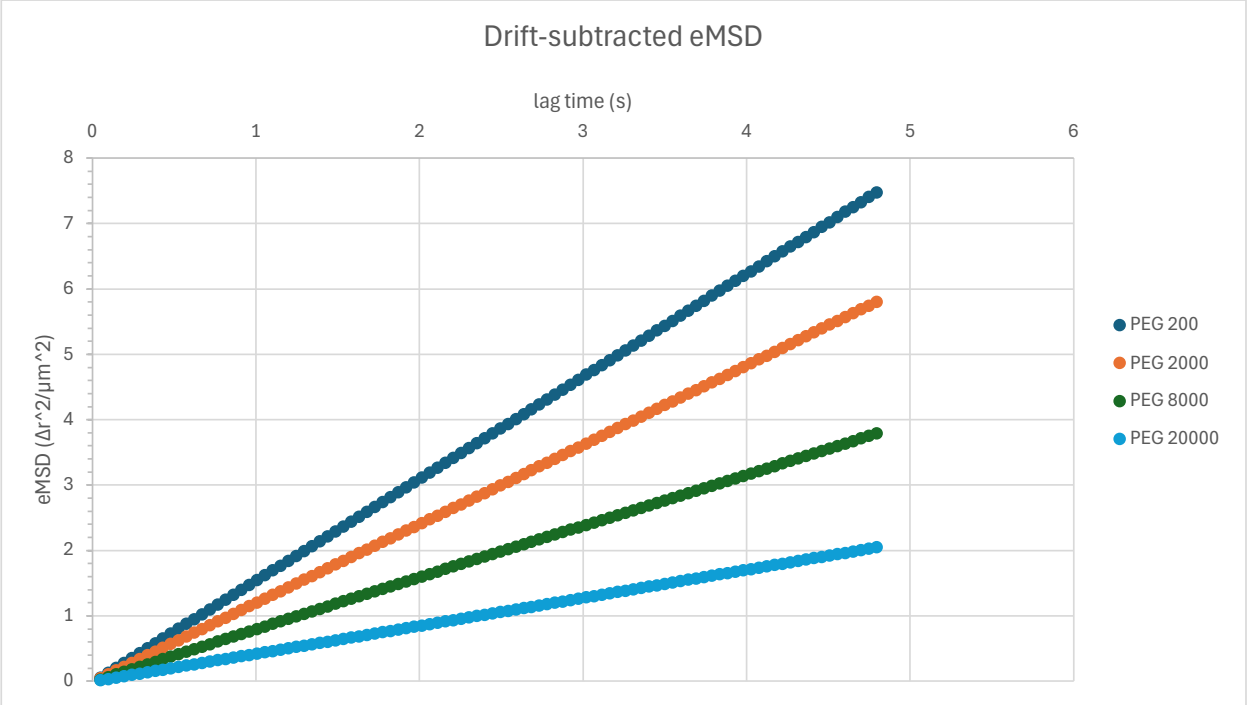

Figure 22: The drift-subtracted eMSD for different molecular weights of PEG as a function of lag time. Concentration is 40mg/ml

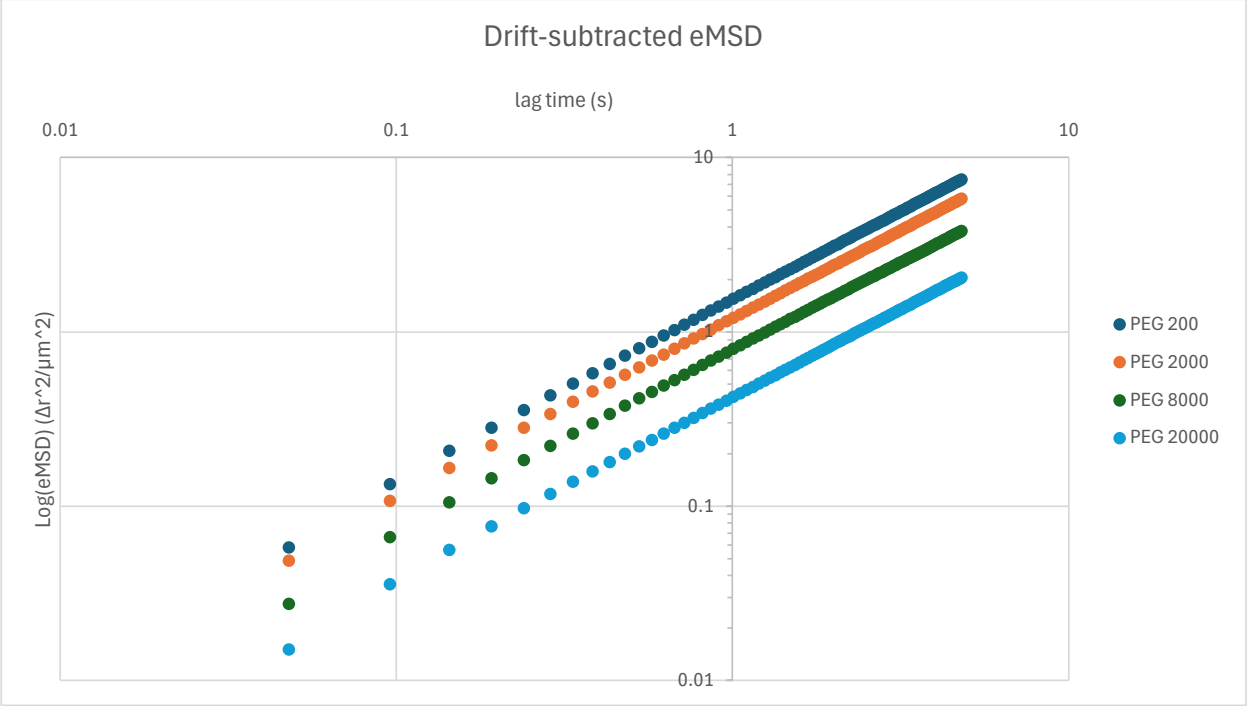

Figure 23: A log-log plot of the drift-subtracted eMSD for different molecular weights of PEG as a function of lag time. Concentration is 40mg/ml

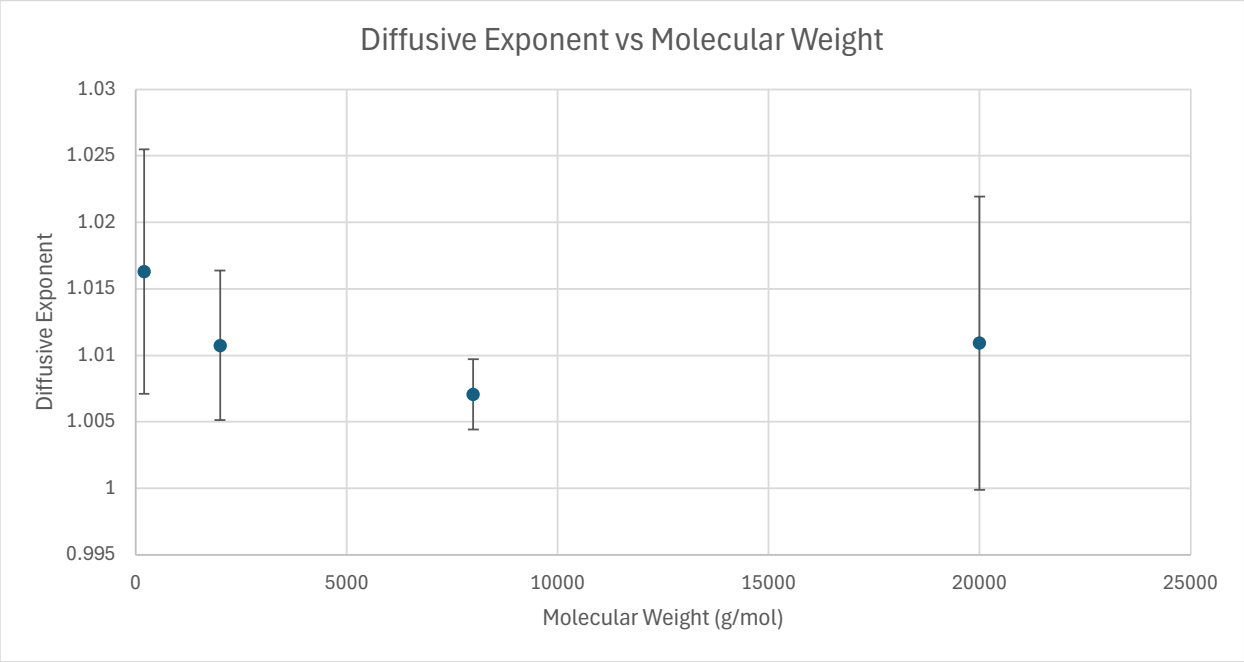

Figure 24: Diffusive exponent α as a function of molecular weight for different chain-lengths of PEG at a concentration of 40mg/ml. Shown with a linear fit with an R^2 value of 0.96. The yellow line represents $\alpha = 1$ (passive diffusion). Error bars are two standard deviations, representing a 95% confidence interval. [Viva update: the superdiffusion is not reproducible, and may depend on the choice of max lag time for the MSD analysis.]

understanding how this length affects diffusion, we are able to take further steps towards developing an accurate model.

7 Polydispersity Study

7.1 Methods

For the polydispersity study, PEG 200 and PEG 20000, both at a concentration of 25mg/ml, were mixed in different ratios to characterize the effects of polydispersity on diffusion. PEG 200 and 20000 were chosen because they have the greatest difference in molecular size among the polymers included in our study. The ratios chosen were: 95% PEG 200 and 5% PEG 20000, 75% PEG 200 and 50% PEG 20000, 50% PEG 200 and 50% PEG 20000, 25% PEG 200 and 75% PEG 20000, and 5% PEG 200 and 95% PEG 20000. The polydispersity was varied with different ratios of PEG 200 to PEG 20000 (i.e. 95% : 5% versus 5% : 95%) in order to not only examine the effect of polydispersity, but whether molecular weight is actually the dominant factor.

7.2 Results and Discussion

This study seeks to answer how the diversity of biomolecules in the cytoplasm may affect diffusion. As with the concentration and molecular weight studies, we begin our analysis by looking at a comparative plot of the eMSDs for the different samples.

Figure 25 shows the median eMSD versus lag time for PEG 200 and PEG 20000 mixed in five different ratios. Polydisperse solutions are one in which the polymer chains are not equal in length. We see however that the eMSD is not most dependent on the polydispersity, but rather the average molecular weight. We additionally observe this trend in Fig. 26, where we see that particles diffuse more in solutions of lower average molecular weight as evidenced by position on the y-axis. The shape of the log-log plot is similar, and almost linear for all four average

Table 6: Ratio of PEG 200: PEG 20000 and the average molecular weight for the crowding agent.

PEG 200 : PEG 20000	Average Molecular Weight (g/mol)
95 : 5	1190
75 : 25	5150
50 : 50	10100
25 : 75	15050
5 : 95	19010

molecular weights, which indicates that diffusion is not highly anomalous as it was for some of the samples as in the concentration study. A log-log plot of the drift-subtracted eMSD versus lag time, shown with a power law fit, can be found in Appendix 4 for every sample in this study.

The average molecular weight for each ratio, calculated using a weighted average, is given in Table 6. We see therefore that tracer particles are most mobile in solutions containing the lowest average molecular weight of PEG, and the least mobile in solutions containing the highest average molecular weight of PEG. The disparity between the eMSD for samples with a lower and higher average molecular mass is increased at longer lag times.

The diffusive exponent follows from the eMSD, and we similarly find a strong correlation between average molecular weight and diffusive exponent. We find a diffusive exponent of 1.016 ± 0.006 for the 95 : 5 ratio of PEG 200 to PEG 20000, 1.013 ± 0.009 for 75 : 25, 1.003 ± 0.007 for 50 : 50, 0.9951 ± 0.0007 for 25 : 75, and 0.990 ± 0.007 for 5 : 95. For the samples containing 50 or more percent PEG 200, the determined diffusive exponent is slightly superdiffusive. This potential superdiffusivity will be explored in a later section. Here, the drop in diffusive exponent as the concentration of PEG 20000 is increased and the concentration of PEG 200 is decreased (consequently the average molecular weight is increased) is of note.

In Fig. 27 we show the diffusive exponent as a function of average molecular weight. From this plot, it is evident that there is once again a strong, negative linear relationship between the dif-

Table 7: Average diffusive exponent for polydisperse solutions of PEG 200 and PEG 20000 at a concentration of 25mg/ml. (Not yet updated)

PEG 200 : PEG 20000	Diffusive Exponent (α)	Uncertainty
95 : 5	1.016	0.006
75 : 25	1.013	0.009
50 : 50	1.003	0.007
25 : 75	0.9951	0.0007
5 : 95	0.990	0.007

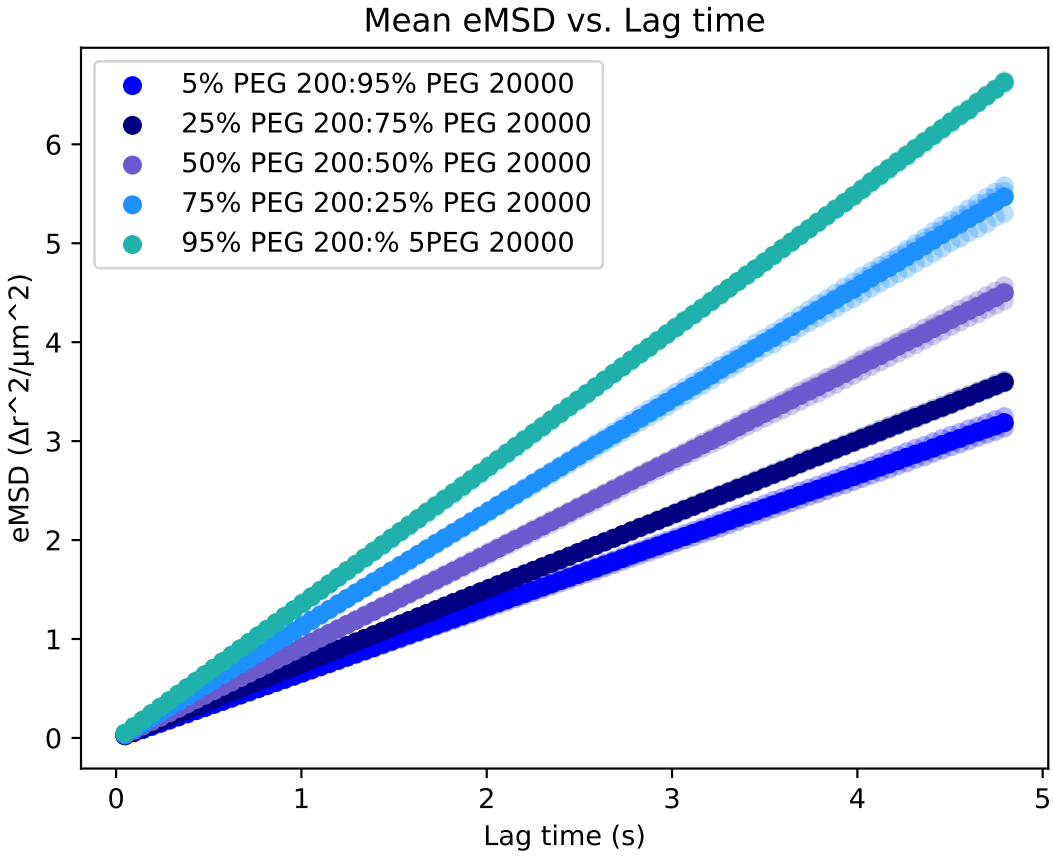

Figure 25: The drift-subtracted eMSD for different ratios of PEG 200 to PEG 20000 as a function of lag-time. PEG concentration is 25mg/ml for all. The opaque markers represent the mean eMSDs, while the translucent markers show the eMSDs for each video. (Updated by Sara Conti from Madeleine Petro's Analysis)

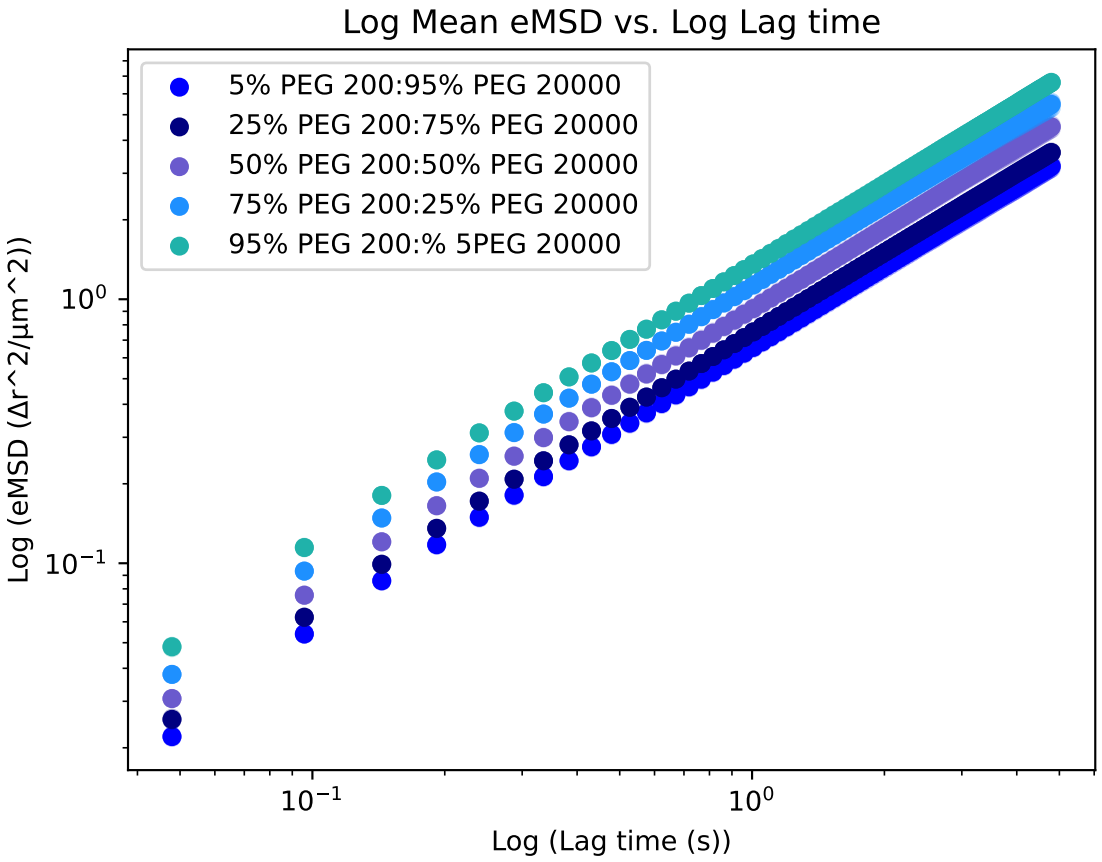

Figure 26: A log-log plot of drift-subtracted eMSD for different ratios of PEG 200 to PEG 20000 as a function of lag-time. PEG concentration is 25 mg/ml for all. The opaque markers represent the mean eMSDs, while the translucent markers show the eMSDs for each video.(Updated by Sara Conti from Madeleine Petro's Analysis)

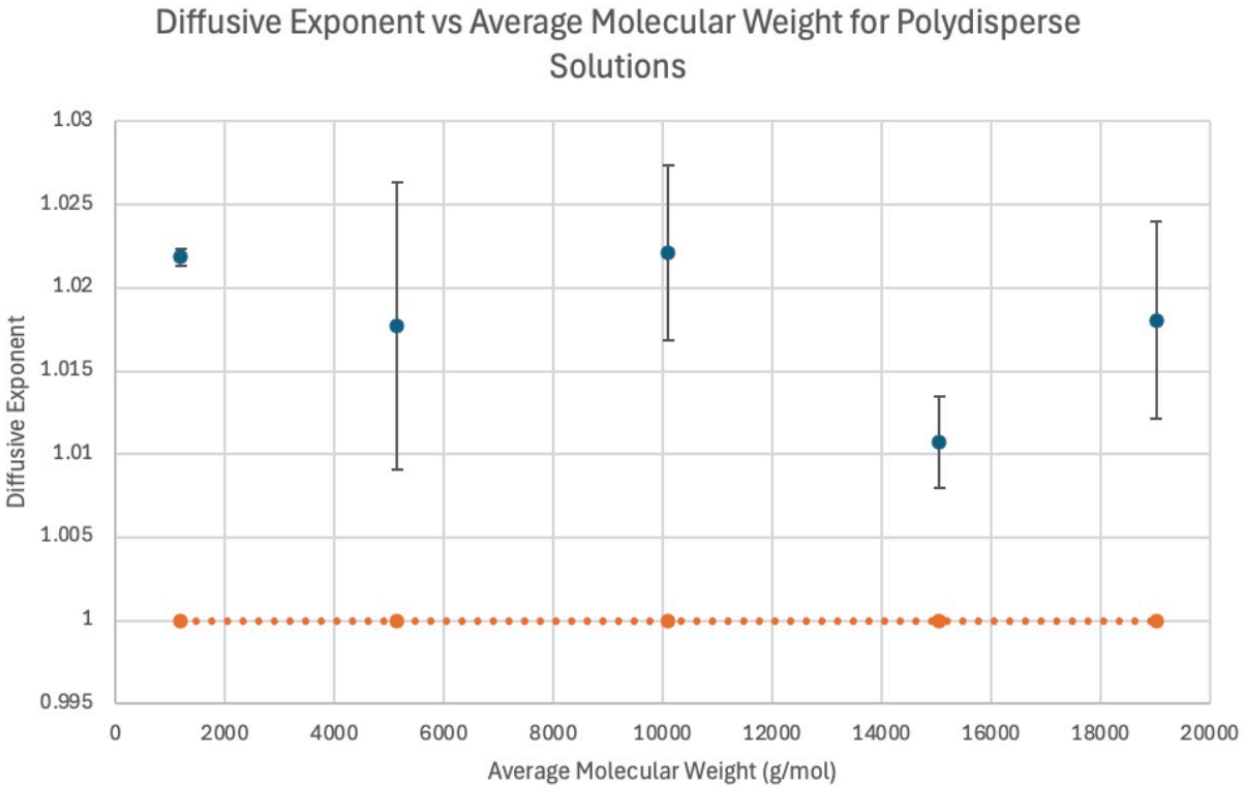

Figure 27: Diffusive exponent α as a function of average molecular weight in g/mol. The yellow line represents $\alpha = 1$ (passive diffusion). Error bars are two standard deviations , representing a 95% confidence interval. (Updated by Madeleine Petro)

fusive exponent and average molecular weight. The data is fit with error bars of two standard deviations , representing a 95% confidence interval. The data is also fit with a linear trend line given by the equation $\alpha = -(1.935 \times 10^{-6})x + 1.014$. This trend line has an R^2 value of 0.99, indicating of a strong fit.

This study suggests that the diffusive exponent is affected more by the size of the crowding molecule(s) than the diversity of molecules present. Furthermore, it is reasonable to conclude that longer chains have more of a caging effect than the same concentration of crowding agent but with shorter chain length (i.e. PEG 200). Further work is needed to explore this effect, and the next section will compare the diffusive exponent and coefficient of diffusion for approximately monodisperse solutions and polydisperse solutions of the same concentration and average molecular mass.

8 Polydispersity Versus Approximate Monodispersity

8.1 Methods

For this study, polydisperse (i.e. bidisperse) and approximately monodisperse solutions of the same concentration and average molecular weight were prepared in order to determine if polydispersity affects diffusivity. Solutions containing only one chain length of PEG are said to be *approximately* monodisperse because there is some uncertainty in the degree of polymerization.

For the approximately monodisperse sample, we made a solution of PEG 8000 at a concentration of 25mg/ml using 250μL of a 50mg/mL PEG 8000 solution mixed with 230μL of RO water and 20μL of the 1% Fluoro-max solution (for an overall tracer particle concentration of 0.04%). For the polydisperse sample, we first determined the ratio that we needed to mix PEG 2000 and PEG 20000 in to achieve an overall average molecular weight of 8000g/mol and concentration

of 25mg/ml. Polyethylene glycol 2000 and 20000 were used because they have a large difference in chain length. The 2000g/mol polymer was chosen over PEG 200, because for PEG 200 we observed a bimodality in the data as explained in the next section. In order to determine the amount of PEG 2000 and 20000 needed to make a solution of our desired concentration and average molecular weight, we solved the system of equations

$$2000x + 20000y = 8000 \tag{20}$$

and

$$x + y = 1 \tag{21}$$

where x (y) is the necessary concentration of PEG 2000 (20000) in mg/ml. Using our solution to the given system, we mixed 333 μL of PEG 2000 with 167 μL of PEG 20000, as well as 540 μL of RO water and 40 μL of the 1% Fluoro-max solution to make our desired polydisperse solution with an average molecular weight of 8000g/mol and concentration of 25mg/ml. Slides were constructed in the same manner as the other studies.

8.2 Results and Discussion

As shown in the plot of eMSD versus lag time given in Fig. 28 and the log-log plot of eMSD versus lag time given in Fig. 29, we can observe a clear difference in tracer particle diffusion in polydisperse versus approximately monodisperse solutions of the same concentration and average molecular weight. For the approximately monodisperse solution, we find $\alpha = 1.002 \pm 0.007$. For the polydisperse solution, $\alpha = 1.008 \pm 0.003$. In both, the uncertainty was determined using two standard deviations, providing a 95% confidence interval. There is actually overlap in the uncertainty for the diffusive exponent between the two samples, and from Fig. 29 we can observe that there is not much of a difference in the shape of the data, indicating that there is very little difference in the diffusive exponent of the two samples. However, when we look at A, which is directly proportional to the coefficient of diffusion, we observe a difference between polydisperse and

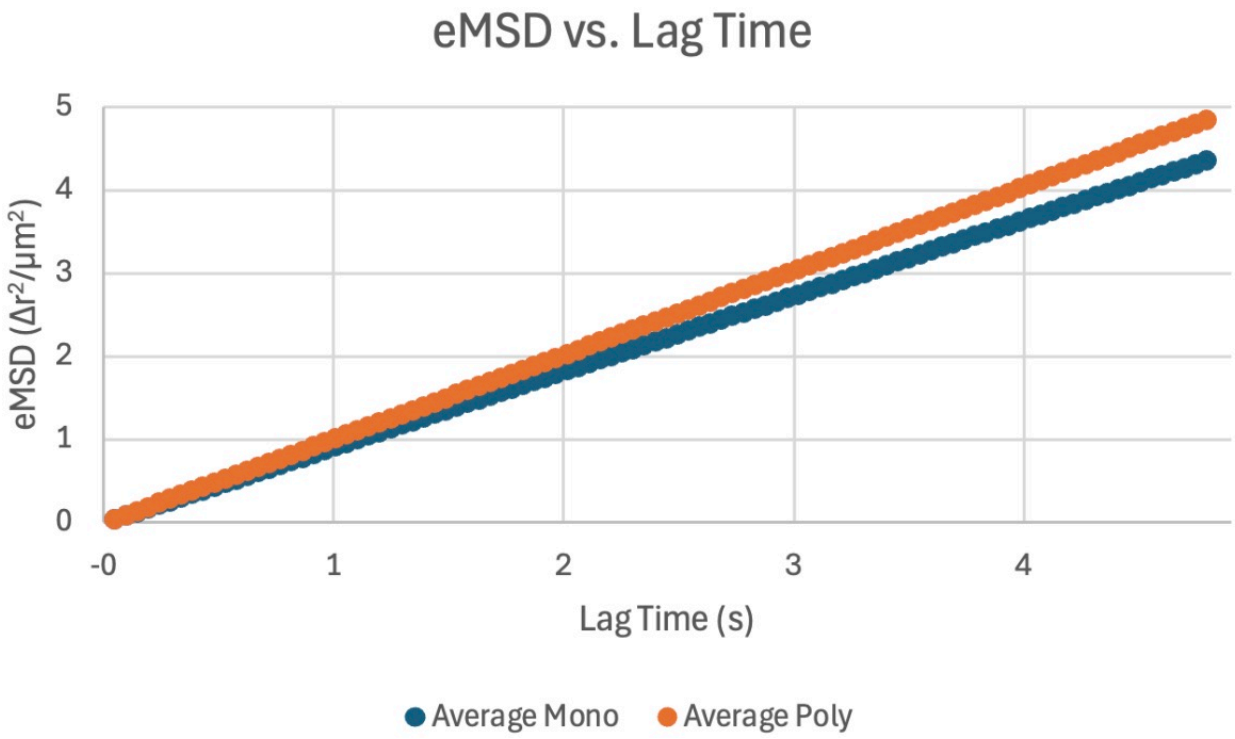

Figure 28: eMSD versus lag time for approximately monodisperse PEG 8000 and a polydisperse solution of PEG 2000 and PEG 20000 with an average molecular weight of 8000g/mol. Green is data from three videos of the polydisperse sample and blue is data from three videos of the approximately monodisperse sample. Concentration for both is 25mg/ml. (Updated by Madeleine Petro)

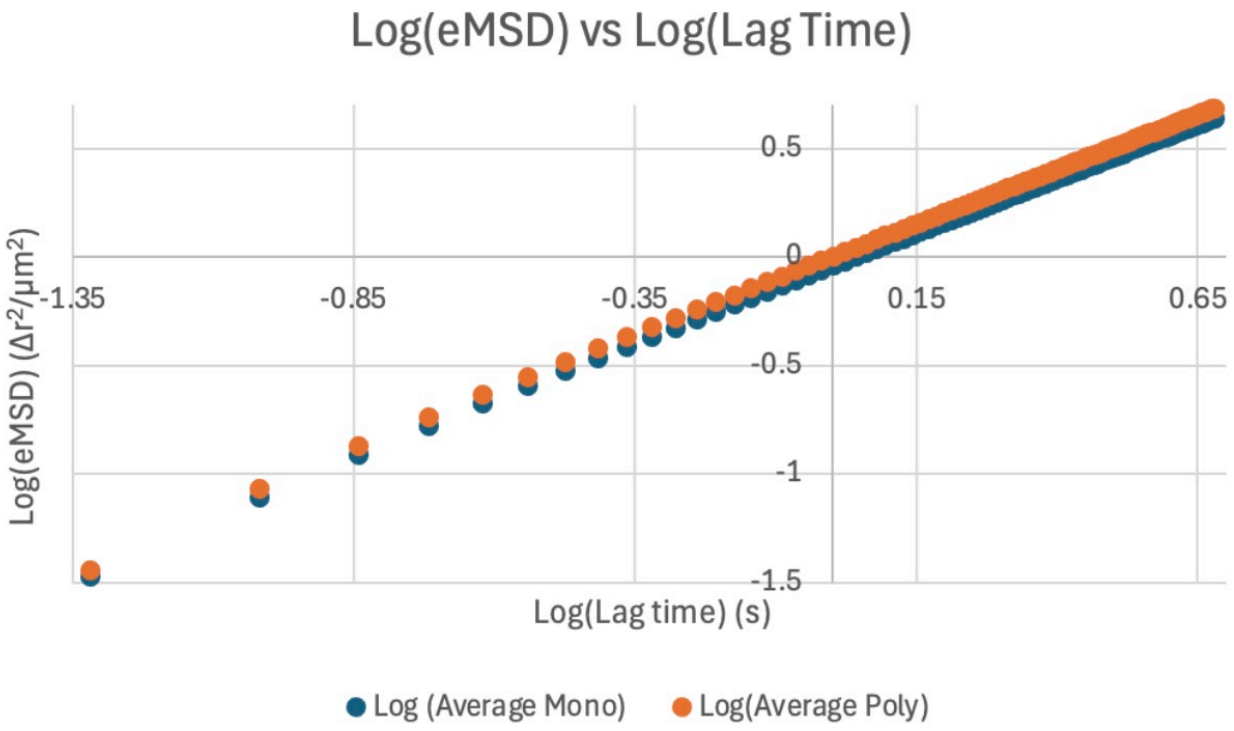

Figure 29: Log-log plot of eMSD versus lag time for approximately monodisperse PEG 8000 and a polydisperse solution of PEG 2000 and PEG 20000 with an average molecular weight of 8000g/mol. Green is data from three videos of the polydisperse sample and blue is data from three videos of the approximately monodisperse sample. Concentration for both is 25mg/ml. (Updated by Madeleine Petro)

approximately monodisperse solutions. For the approximately monodisperse solution we find that A=0.902 ± 0.014, while for the polydisperse solution we find that A=0.999 ± 0.014. Uncertainty was once again calculated using two standard deviations for a 95% confidence interval. There is no overlap in the uncertainty for A, and we can conclude that tracer particles in a polydisperse solution diffuse more than those in a monodisperse solution. This can also be seen in Figures 28 and 29, where the points corresponding to the polydisperse solution are shifted up on the y-axis from those corresponding to the approximately monodisperse solution.

In a theoretical study, Gonzalez *et al.* found that many-body hydrodynamic couplings are weakened in polydisperse solutions, which leads to increased diffusivity for small particles [33]. Additionally, the polydisperse solution will have a higher entropy of mixing, which could lead to enhanced diffusion. More work is needed to examine polydispersity versus approximate monodispersity, as will be expanded upon in the future research section.

9 Interesting Phenomena

9.1 Bimodality

[Note from Viva: The bimodality is an artifact of tracking things that are not particles. It goes away entirely when we add a minimum threshold to the brightness of the brightest pixel of what we will consider a particle.]

Bimodality was repeatedly seen in the iMSD of samples with low concentrations of PEG 200 and in the RO water control samples. As seen in Figures 30-34, which show the iMSD versus lag time, there are two distinct populations. The iMSDs before drift-subtraction were selected because drift-subtraction smoothes out the bimodality, perhaps creating a false comfort. One population follows the expected displacement pattern while the particles in the other population appear to be relatively stuck (their displacement does not change, even over longer lag times). Having two distinct populations is a potential source of error, as our analysis relies on

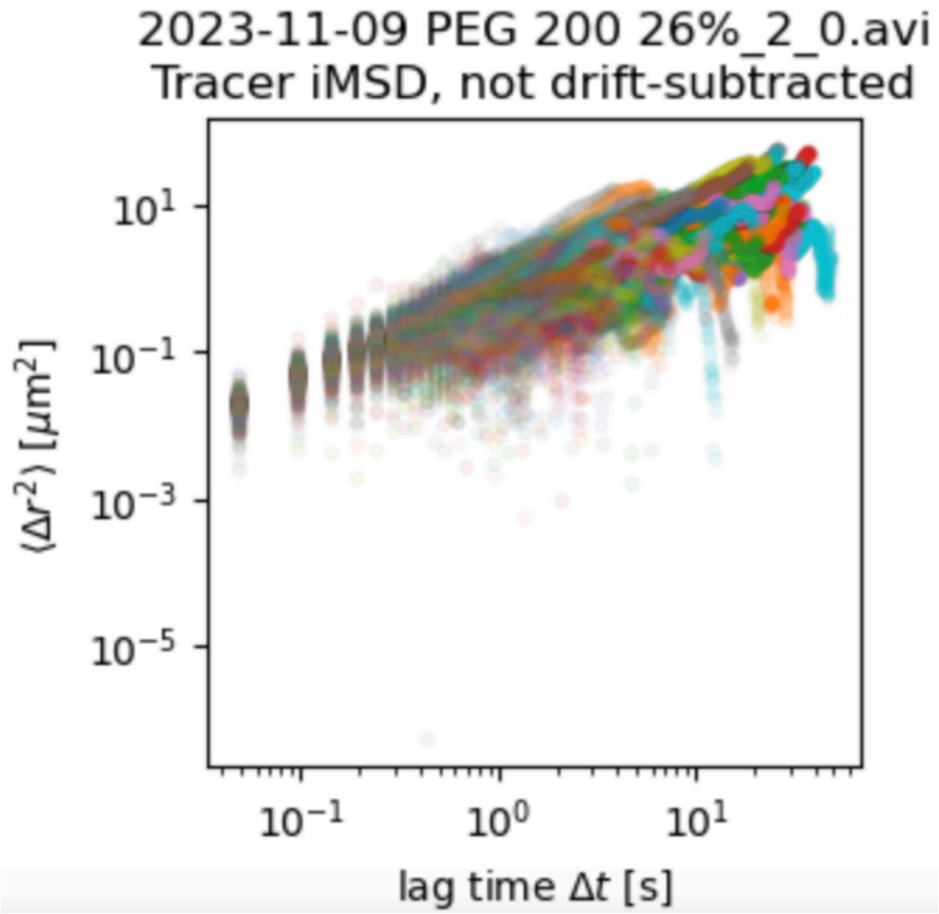

Figure 30: The iMSD without drift subtraction for PEG 200 at 26% volume (video 2). (Updated by Madeleine Petro)

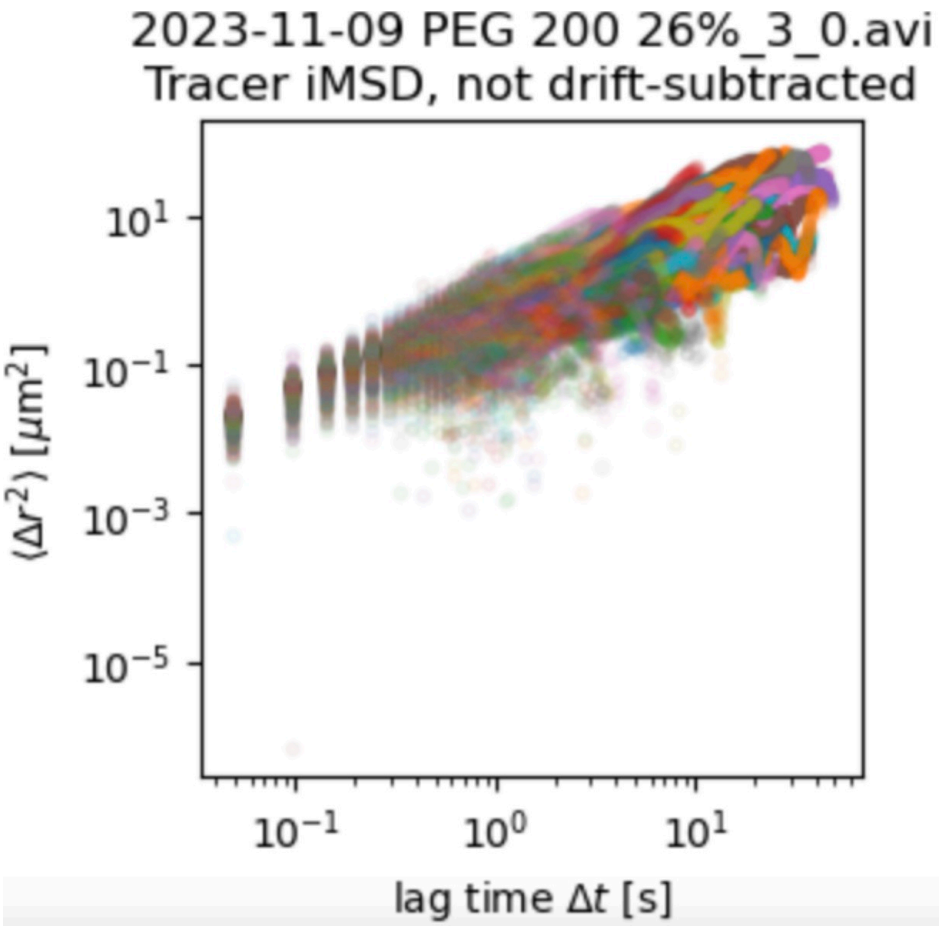

Figure 31: The iMSD without drift subtraction for PEG 200 at 26% volume (video 3). (Updated by Madeleine Petro)

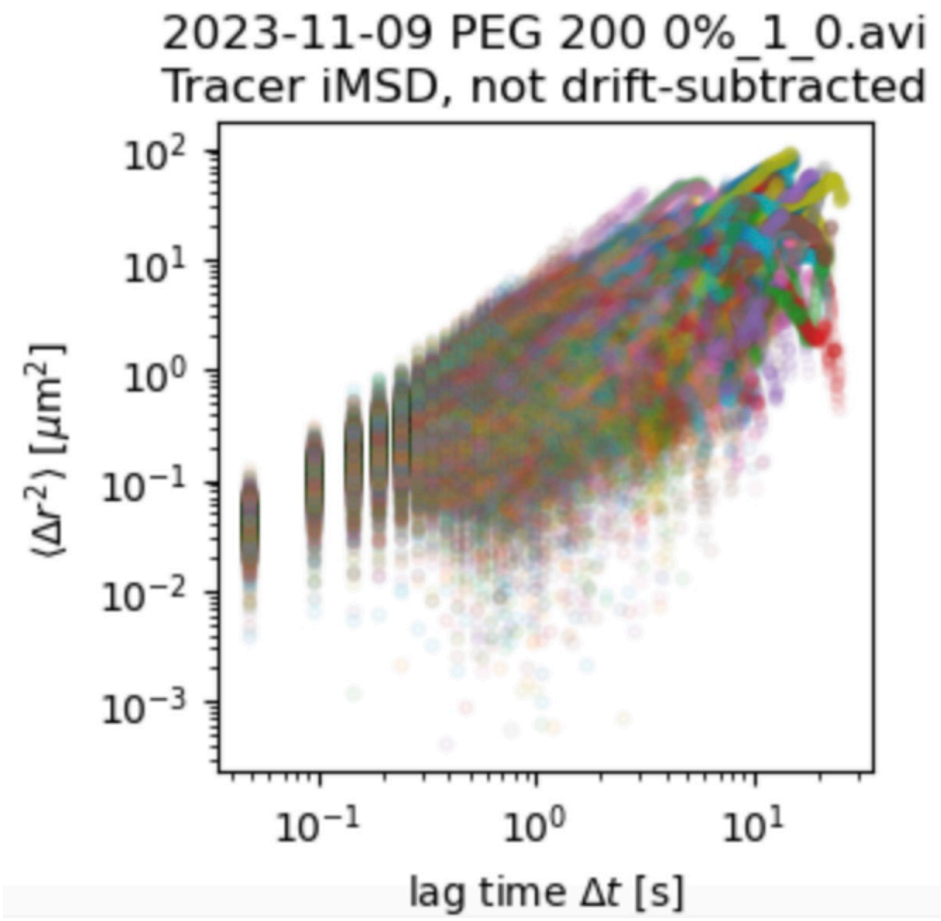

Figure 32: The iMSD without drift subtraction for the RO water control from the concentration study (video 1). (Updated by Madeline Petro)

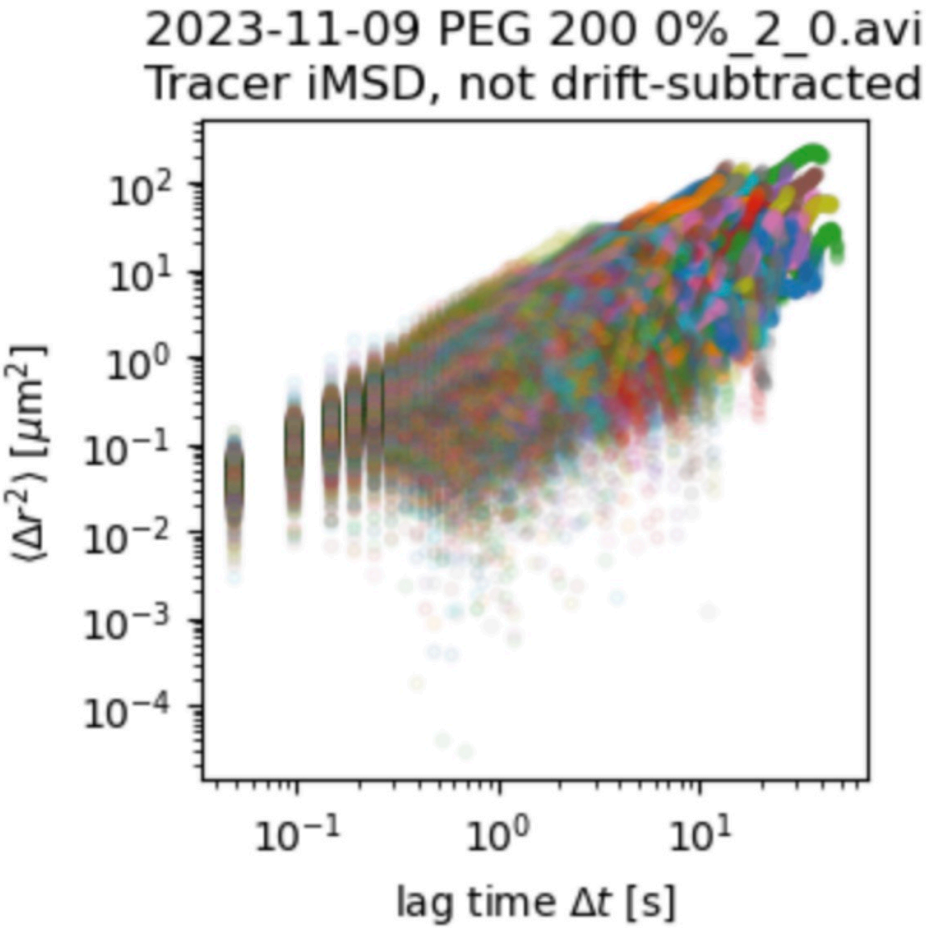

Figure 33: The iMSD without drift subtraction for the RO water control from the concentration study (video 2).(Updated by Madeleine Petro)

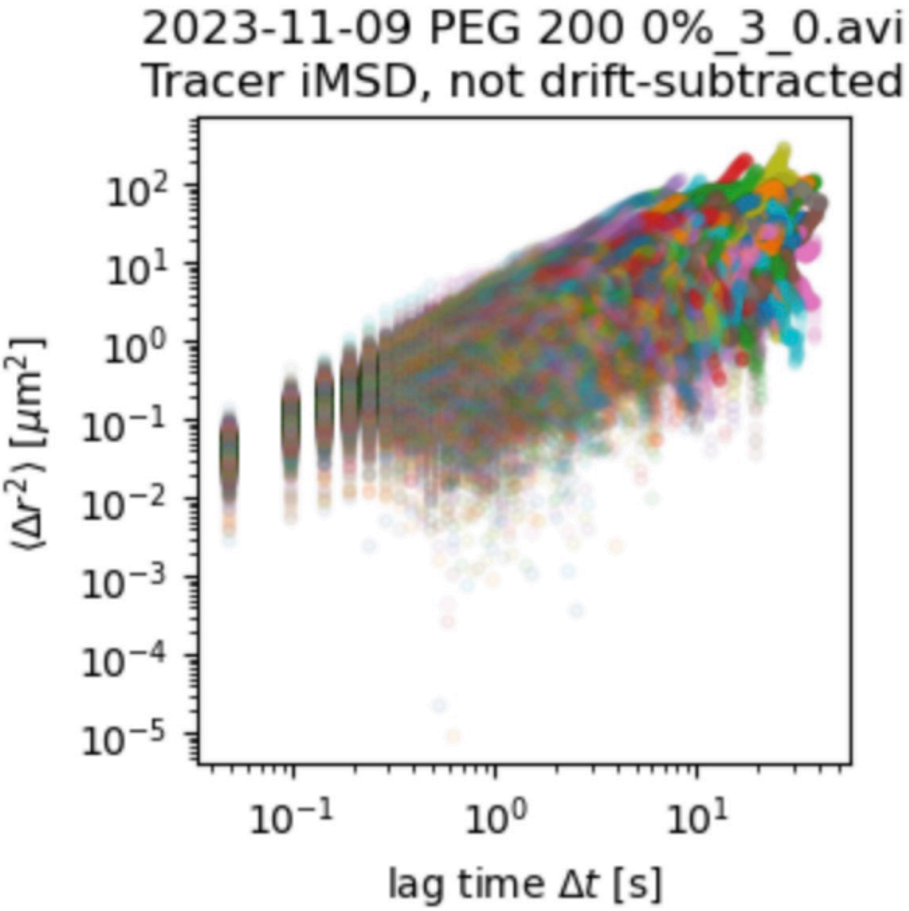

Figure 34: The iMSD without drift subtraction for the RO water control from the concentration study (video 3). (Updated by Madeleine Petro)

taking averages. Averaging two distinct populations as one is statistically inaccurate. It is notable that this bimodality primarily occurs with lower concentration PEG solutions and RO and tracer particle solutions.

More work is needed to determine the cause of such a bimodality, but we can turn to existing literature for possible explanations. Polyethylene glycol 200 is the only polymer that we used with a liquid phase at room temperature. Raman spectroscopy, which measures scattered photons, can be used to examine the vibrational modes of molecules. Kuzmin *et al.* compared the experimental Raman spectra of PEG at molecular masses that are liquid, semi-solid, and powders at room temperature [34]. They determined that the liquid phase samples have a significantly different Raman spectrum, particularly around the 1100 cm^{-1} peak [34]. The shift in this peak position can be used to determine the molecular length, and it's noted that the O-C-C-O torsion angles change significantly between the monomers at the center of a molecule versus at the ends [34]. A difference in molecular structure between PEG 200 and its solid phase counterparts could potentially produce a different caging effect.

For the RO water control, the presence of the "stuck" particles is equally puzzling due to the lack of a crowding agent. As with the bimodality of the PEG 200 sample, further work is needed. In the meantime, we can turn to existing literature for possible explanations. Our tracer particles are made of polystyrene, which is a non-polar, and consequently hydrophobic, polymer. Hydrophobic particles in water can become entrapped, forming clathrate hydrate at low temperature and reasonably high pressure [35]. In ambient conditions, certain biomolecules can form clathrate-like water, as seen in Fig. 35. Parui *et al.* note that clathrate formation is dependent on a confined, hydrophobic environment, both of which our artificial cell model with polystyrene particles possess. Future work can use tracer particles made from a hydrophilic material to test this hypothesis of the formation of clathrate-like water in the RO control.

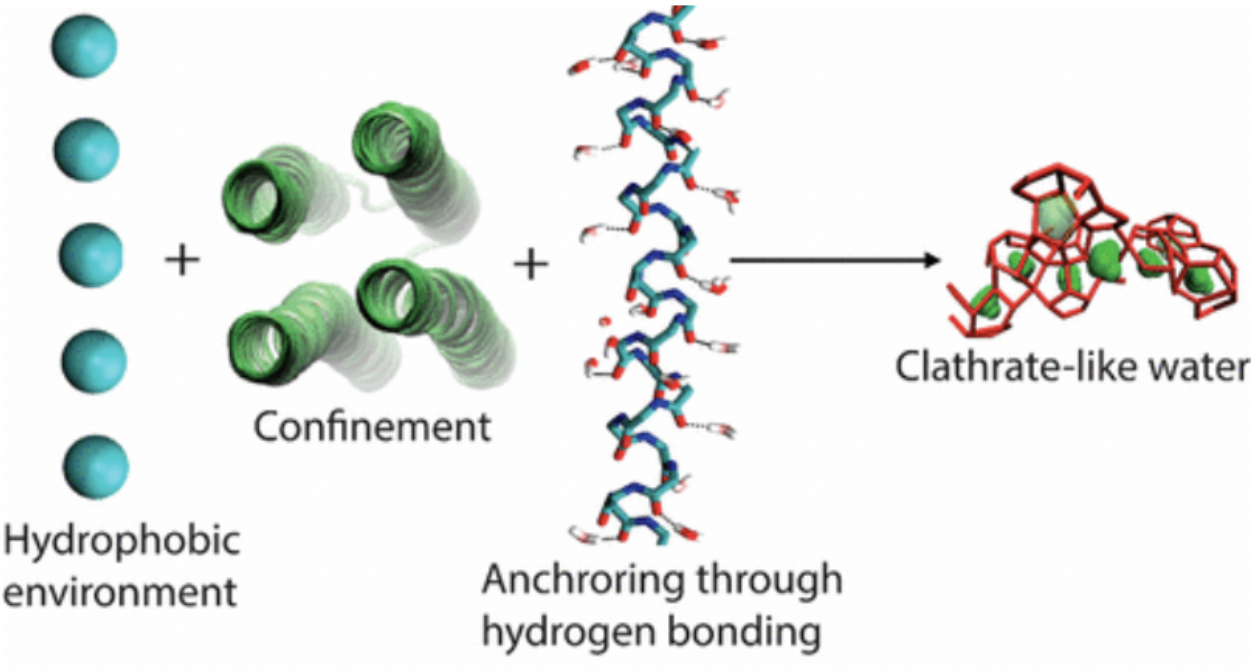

Figure 35: Diagram showing the formation of clathrate-like water. Reprinted from [35].

9.2 Superdiffusion In Unfueled Systems

For all three studies in this research, a superdiffusive exponent was observed for samples at the low range of their crowding factor (be it molecular weight or concentration), as well as for the RO water control samples. We know from our theoretical power law that water should be perfectly linear, with diffusive exponent $\alpha = 1$. Theory dictates that superdiffusion should not be observed without the addition of fuel, or some external energy source, to the system.

Normally, a superdiffusive exponent would be attributed to drift caused by air bubbles in the system. However, drift is a more random source of error and we observe repeated superdiffusion. Another potential cause of this drift is the flourescent lamp we use as a light source. It is possible that the lamp is adding thermal energy to the system.

We can also observe from our log-log eMSD plots that there are multiple distinct behaviors over different lag times that cannot all be treated with the same fit. This observation of multiple behaviors for eMSD over different lag times has been observed before in a different system, as seen in Fig. 36 which shows the eMSD versus lag time for Janus swimmers in hydrogen peroxide solution. Discussion of how to analyze these behaviors will be included in the following section.

10 Future Research

As discussed in the previous section, future research should seek to provide a more robust explanation for the bimodality observed for low concentrations of PEG 200 and RO water with polystyrene tracer particles, as well as the observed supperdiffusivity for tracer particles in unfueled systems. To examine if the bimodality is due to to the formation of clathrate-like water, we will use silica microspheres which are hydrophilic rather than hydrophobic like our current polystyrene tracer particles. With regards to the superdiffusivity, we will cut out the short lag times from our fit, where there are fewer points, in order to ensure that we are only fitting one

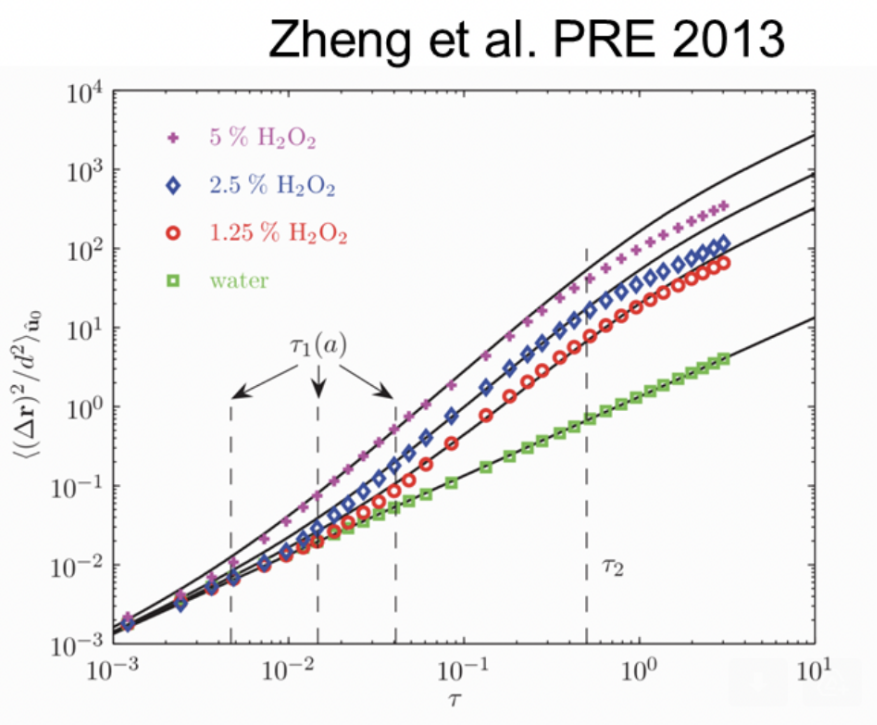

Figure 36: eMSD versus some lag time τ for Janus particles in hydrogen peroxide solution. Reprinted from [36].

diffusion behavior at a time. Additionally, we will use a faster camera to determine if we are actually observing different diffusive behaviors.

Our observed increase in diffusion for tracer particles in polydisperse solutions as compared to approximately monodisperse solutions of the same concentration and average molecular weight is a promising result from a clean data set. We will build upon this result by making polydisperse solutions from PEG 2000 and 20000 of the same average molecular weight as different approximately monodisperse solutions of PEG (i.e. PEG 12000).

Using the results from each of the three studies (concentration, molecular weight, and polydispersity) we have established a more accurate and complete understanding of how macromolecular crowding affects diffusivity in the cytoplasm. Now, to expand upon our model, it is necessary to incorporate Janus swimmers to mimic the work done by motor proteins in eukary-

otic cells. As mentioned in the introduction, the platinum-coated side of the Janus swimmers catalyzes a reaction with hydrogen peroxide that releases molecules of oxygen as a byproduct. This molecular oxygen is an issue, as it causes drift in the system. Therefore, in order to appropriately introduce Janus swimmers into our model, we need to build a drift reduction chamber that contains particles while allowing molecules of oxygen to diffuse out. Future research will build upon the work of Palacci [19], Smolarsky [23], and Lawrence [22] to construct such a chamber using an agarose hydrogel, and consequently complete our model through the introduction of Janus swimmers.

11 Conclusion

This research included three primary studies varying the concentration, average molecular weight, and polydispersity of PEG, which was used to mimic biomolecular crowding in artificial cells. Through these studies, we were able to observe clear relations between how different aspects of macromolecular crowding affect diffusivity.

From the concentration study, it was determined that an increase in concentration of the crowding agent leads to a decrease in diffusivity of fluorescent tracer particles. The molecular weight study similarly shows that an increase in the size of the crowding agent leads to a decrease in diffusivity. Finally, the polydispersity study confirmed that diffusivity decreases as average molecular weight increases. The data collected comparing polydisperse and approximately monodisperse solutions of the same concentration and average molecular weight suggests that polydispersity actually increases diffusivity. The results of this study provide a more detailed explanation, complete with control data, of the previous findings of Smolarsky and Lawrence. We have developed a clean framework for replicating and observing subdiffusion varying different crowding conditions, which sets the stage for the introduction of Janus swimmers to continue building the artificial cell model.

References

1. Shahid, S., Ahmad, F., Hassan, M. I. & Islam, A. Relationship between protein stability and functional activity in the presence of macromolecular crowding agents alone and in mixture: An insight into stability-activity trade-off. *Archives of Biochemistry and Biophysics* **584,** 42–50. ISSN: 0003-9861. https://www.sciencedirect.com/science/article/pii/S0003986115300412 (2023) (Oct. 2015).

2. Van Der Giezen, M. Mitochondria and the Rise of Eukaryotes. en. *BioScience* **61,** 594–601. ISSN: 1525-3244, 0006-3568. https://academic.oup.com/bioscience/article-lookup/doi/10.1525/bio.2011.61.8.5 (2023) (Aug. 2011).

3. Weiss, M., Elsner, M., Kartberg, F. & Nilsson, T. Anomalous Subdiffusion Is a Measure for Cytoplasmic Crowding in Living Cells. *Biophysical Journal* **87,** 3518–3524. ISSN: 0006-3495. https://www.sciencedirect.com/science/article/pii/S0006349504738163 (2023) (Nov. 2004).

4. Zeravcic, Z., Manoharan, V. N. & Brenner, M. P. *Colloquium* : Toward living matter with colloidal particles. en. *Reviews of Modern Physics* **89,** 031001. ISSN: 0034-6861, 1539-0756. https://link.aps.org/doi/10.1103/RevModPhys.89.031001 (2023) (Sept. 2017).

5. Bechinger, C. *et al.* Active particles in complex and crowded environments. *Reviews of Modern Physics* **88,** 045006. https://link.aps.org/doi/10.1103/RevModPhys.88.045006 (2023) (Nov. 2016).

6. Luby-Phelps, K., Taylor, D. L. & Lanni, F. Probing the structure of cytoplasm. en. *The Journal of cell biology* **102,** 2015–2022. ISSN: 0021-9525, 1540-8140. https://rupress.org/jcb/article/102/6/2015/58506/Probing-the-structure-of-cytoplasm (2023) (June 1986).

7. Seksek, O., Biwersi, J. & Verkman, A. Translational Diffusion of Macromolecule-sized Solutes in Cytoplasm and Nucleus. en. *The Journal of Cell Biology* **138,** 131–142. ISSN: 0021-

9525, 1540-8140. https://rupress.org/jcb/article/138/1/131/15536/Translational-Diffusion-of-Macromolecule-sized (2023) (July 1997).

8. Arrio-Dupont, M., Foucault, G., Vacher, M., Devaux, P. F. & Cribier, S. Translational Diffusion of Globular Proteins in the Cytoplasm of Cultured Muscle Cells. en. *Biophysical Journal* **78,** 901–907. ISSN: 00063495. https://linkinghub.elsevier.com/retrieve/pii/S0006349500766471 (2023) (Feb. 2000).

9. Saxton, M. J. Lateral diffusion in an archipelago. Dependence on tracer size. en-US. *Biophysical Journal* **64,** 1053–1062. ISSN: 0006-3495. https://www.sciencedirect.com/science/article/pii/S0006349593814711 (2023) (Apr. 1993).

10. Banks, D. S. & Fradin, C. Anomalous Diffusion of Proteins Due to Molecular Crowding. *Biophysical Journal* **89,** 2960–2971. ISSN: 0006-3495. https://www.sciencedirect.com/science/article/pii/S0006349505729404 (2023) (Nov. 2005).

11. Ray Foster, L. J. en. in *Biopolymers* (ed Elnashar, M.) (Sciyo, Sept. 2010). ISBN: 9789533071091. http://www.intechopen.com/books/biopolymers/pegylation-and-biopegylation-of-polyhydroxyalkanoates-synthesis-characterisation-and-applications (2023).

12. Cates, M. E. Diffusive transport without detailed balance in motile bacteria: does microbiology need statistical physics? *Reports on Progress in Physics* **75,** 042601. ISSN: 0034-4885, 1361-6633. https://iopscience.iop.org/article/10.1088/0034-4885/75/4/042601 (2023) (Apr. 2012).

13. Wang, W., Lv, X., Moran, J. L., Duan, S. & Zhou, C. A practical guide to active colloids: choosing synthetic model systems for soft matter physics research. en. *Soft Matter* **16,** 3846–3868. ISSN: 1744-683X, 1744-6848. http://xlink.rsc.org/?DOI=D0SM00222D (2023) (2020).

14. Paxton, W. F. *et al.* Catalytic Nanomotors: Autonomous Movement of Striped Nanorods. en. *Journal of the American Chemical Society* **126,** 13424–13431. ISSN: 0002-7863, 1520-5126. https://pubs.acs.org/doi/10.1021/ja047697z (2023) (Oct. 2004).

15. Fournier-Bidoz, S., Arsenault, A. C., Manners, I. & Ozin, G. A. Synthetic self-propelled nanorotors. en. *Chemical Communications,* 441–443. ISSN: 1364-548X. https://pubs.rsc.org/en/content/articlelanding/2005/cc/b414896g (2023) (Jan. 2005).

16. Mano, N. & Heller, A. Bioelectrochemical Propulsion. en. *Journal of the American Chemical Society* **127,** 11574–11575. ISSN: 0002-7863, 1520-5126. https://pubs.acs.org/doi/10.1021/ja053937e (2023) (Aug. 2005).

17. Howse, J. R. *et al.* Self-Motile Colloidal Particles: From Directed Propulsion to Random Walk. en. *Physical Review Letters* **99,** 048102. ISSN: 0031-9007, 1079-7114. https://link.aps.org/doi/10.1103/PhysRevLett.99.048102 (2023) (July 2007).

18. Chen, L., Mo, C., Wang, L. & Cui, H. Direct numerical simulation of the self-propelled Janus particle: use of grid-refined fluctuating lattice Boltzmann method. en. *Microfluidics and Nanofluidics* **23,** 73. ISSN: 1613-4982, 1613-4990. http://link.springer.com/10.1007/s10404-019-2230-1 (2023) (May 2019).

19. Palacci, J., Cottin-Bizonne, C., Ybert, C. & Bocquet, L. Sedimentation and Effective Temperature of Active Colloidal Suspensions. en. *Physical Review Letters* **105,** 088304. ISSN: 0031-9007, 1079-7114. https://link.aps.org/doi/10.1103/PhysRevLett.105.088304 (2023) (Aug. 2010).

20. *Michaelis-Menten Kinetics and Briggs-Haldane Kinetics* https://depts.washington.edu/wmatkins/kinetics/michaelis-menten.html#:~:text=Michaelis%20and%20Menten%20assumed%20that,KD%20and%20KM. (2023).

21. Leptos, K. C., Guasto, J. S., Gollub, J. P., Pesci, A. I. & Goldstein, R. E. Dynamics of Enhanced Tracer Diffusion in Suspensions of Swimming Eukaryotic Microorganisms. *Physical Re-*

view Letters **103,** 198103. https://link.aps.org/doi/10.1103/PhysRevLett.103.198103 (2023) (Nov. 2009).

22. Lawrence, E. *Artificial Cytoplasm: Observing Anomalous Diffusion & Progress Towards Drift-Reduction* May 2023.

23. Smolarsky, R. *Diffusion and Drift Reduction in Artificial Cells* Dec. 2022.

24. Wang, Y., Jeong, Y., Jhiang, S., Yu, L. & Menq, C.-H. *Calculation of mean square displacement (MSD) and determination of motion type.* en. figure. Dec. 2015. https://figshare.com/articles/figure/_Calculation_of_mean_square_displacement_MSD_and_determination_of_motion_type_/1051270/1 (2023).

25. Schroeder, D. V. *An introduction to thermal physics* ISBN: 9780192895547 9780192895554 (Oxford University Press, New York, NY, 2021).

26. Shuman, D. *Lecture 4: Diffusion: Fick's second law* 2004. https://www-eng.lbl.gov/~shuman/NEXT/MATERIALS&COMPONENTS/Xe_damage/ficks2ndlaw.pdf.

27. Einstein, A. *Investigations on the Theory of, the Brownian Movement* (ed Fürth, R.) trans. by Cowper, A. original-date: 1926 (Dover Publications, 1956).

28. Neher, R. *Brownian motion and Diffusion* en-US. Oct. 2018. https://neherlab.org/20181004_theoretical_biophysics.html (2023).

29. Lindwall, G. & Gerlee, P. Fast and precise inference on diffusivity in interacting particle systems. en. *Journal of Mathematical Biology* **86,** 64. ISSN: 0303-6812, 1432-1416. https://link.springer.com/10.1007/s00285-023-01902-y (2023) (May 2023).

30. Horowitz, V. R., Chambers, Z. C., Gözen, İ., Dimiduk, T. G. & Manoharan, V. N. Active colloidal particles in emulsion droplets: a model system for the cytoplasm. en. *The European Physical Journal Special Topics* **227,** 2413–2424. ISSN: 1951-6401. https://doi.org/10.1140/epjst/e2019-800026-y (2023) (Mar. 2019).

31. Westfall, P. H. Kurtosis as Peakedness, 1905 – 2014. R.I.P. *The American statistician* **68,** 191–195. ISSN: 0003-1305. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4321753/ (2023) (2014).

32. Fulton, A. B. How crowded is the cytoplasm? en. *Cell* **30,** 345–347. ISSN: 00928674. https://linkinghub.elsevier.com/retrieve/pii/0092867482902318 (2023) (Sept. 1982).

33. Gonzalez, E., Aponte-Rivera, C. & Zia, R. N. Impact of polydispersity and confinement on diffusion in hydrodynamically interacting colloidal suspensions. en. *Journal of Fluid Mechanics* **925,** A35. ISSN: 0022-1120, 1469-7645. https://www.cambridge.org/core/journals/journal-of-fluid-mechanics/article/abs/impact-of-polydispersity-and-confinement-on-diffusion-in-hydrodynamically-interacting-colloidal-suspensions/17D33C3C04C20CFEBA5B6522368B367B (2023) (Oct. 2021).

34. Kuzmin, V. *et al.* Raman spectra of polyethylene glycols: Comparative experimental and DFT study. en. *Journal of Molecular Structure* **1217,** 128331. ISSN: 00222860. https://linkinghub.elsevier.com/retrieve/pii/S0022286020306566 (2023) (Oct. 2020).

35. Parui, S. & Jana, B. Factors Promoting the Formation of Clathrate-Like Ordering of Water in Biomolecular Structure at Ambient Temperature and Pressure. en. *The Journal of Physical Chemistry B* **123,** 811–824. ISSN: 1520-6106, 1520-5207. https://pubs.acs.org/doi/10.1021/acs.jpcb.8b11172 (2023) (Jan. 2019).

36. Zheng, X. *et al.* Non-Gaussian statistics for the motion of self-propelled Janus particles: Experiment versus theory. *Physical Review E* **88,** 032304. https://link.aps.org/doi/10.1103/PhysRevE.88.032304 (2023) (Sept. 2013).

A Appendix

A.1 Tracer Particle Concentration Data

Table 8: Average diffusive exponent for different concentrations of tracer particles. (Not yet updated.)

Tracer Particle Concentration (Percent Volume)	Diffusive Exponent (α)	Uncertainty
0.02%	1.019	0.009
0.03%	1.011	0.010
0.04%	1.018	0.002
0.05%	1.001**	0.003

**Cannot accurately determine diffusive exponent because Trackpy fails to identify centers at this concentration.

A.2 Log-log eMSD Versus Lag Time With Power Law Fit For Concentration Study

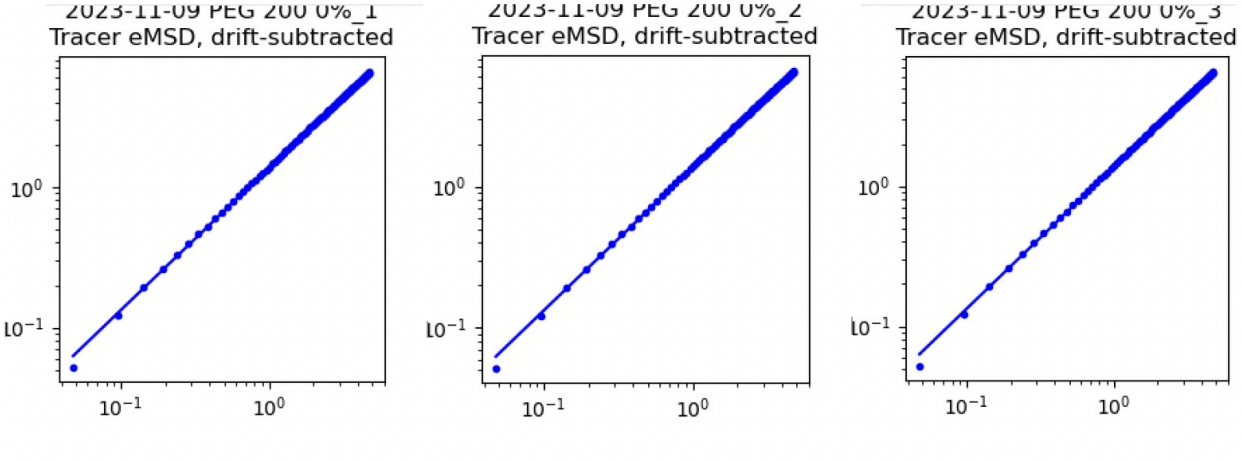

Figure 37: RO Water Control.

.

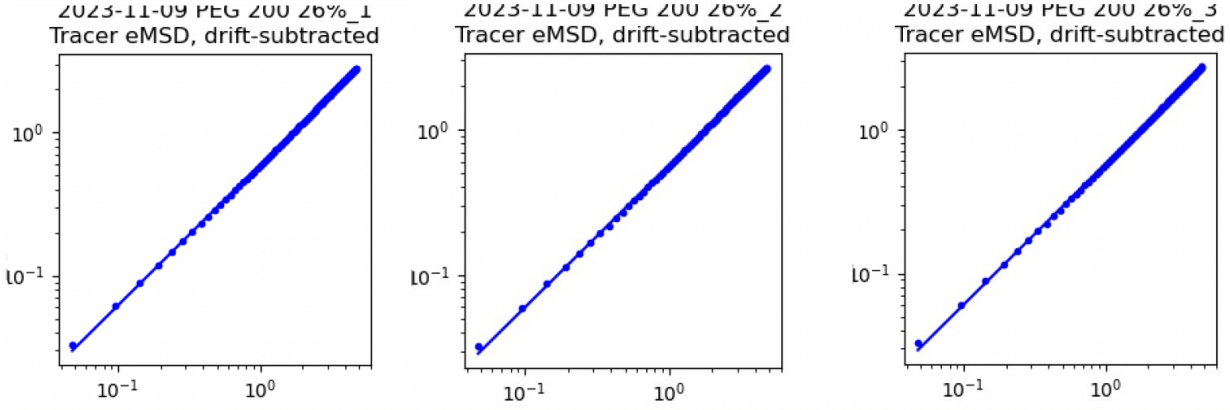

Figure 38: PEG 200 at 26% by volume.

.

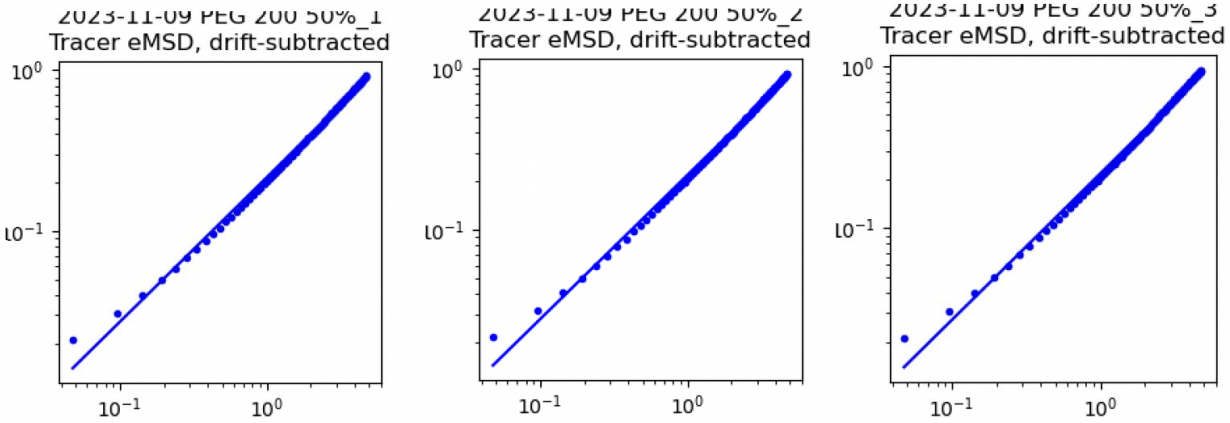

Figure 39: PEG 200 at 50% by volume.

.

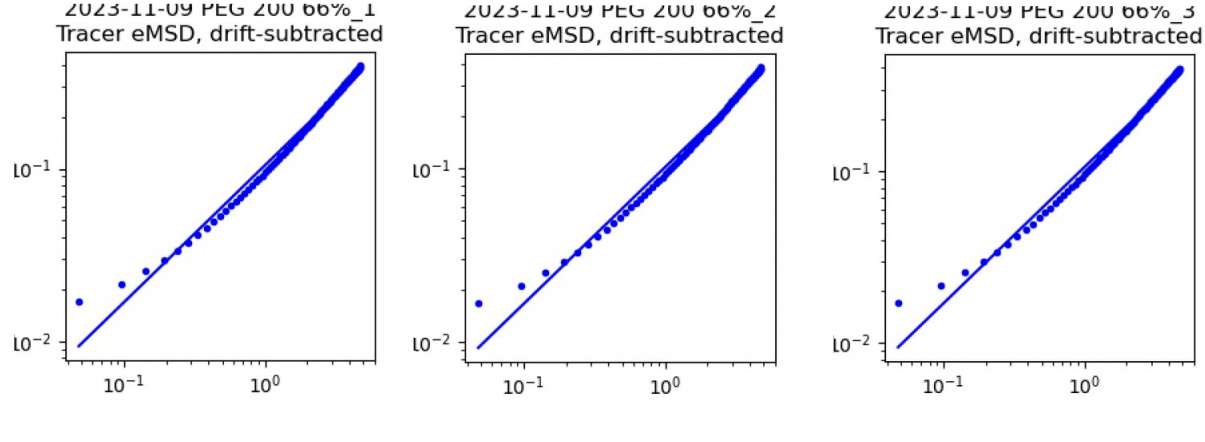

Figure 40: PEG 200 at 66% by volume.

.

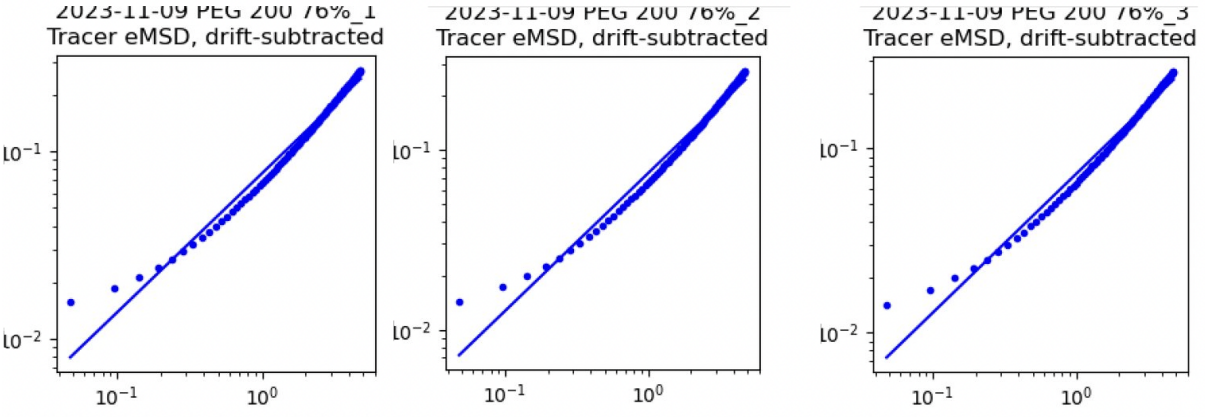

Figure 41: PEG 200 at 76% by volume.

.

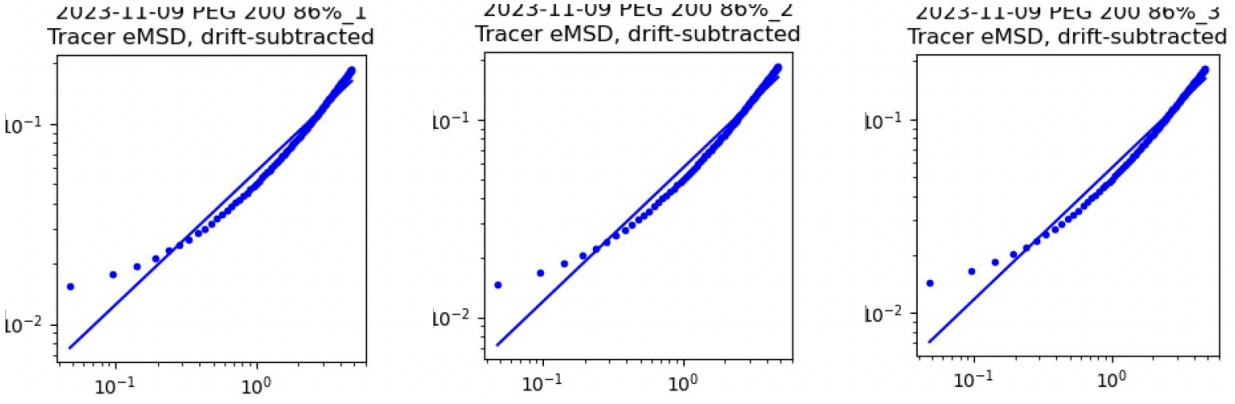

Figure 42: PEG 200 at 86% by volume.

.

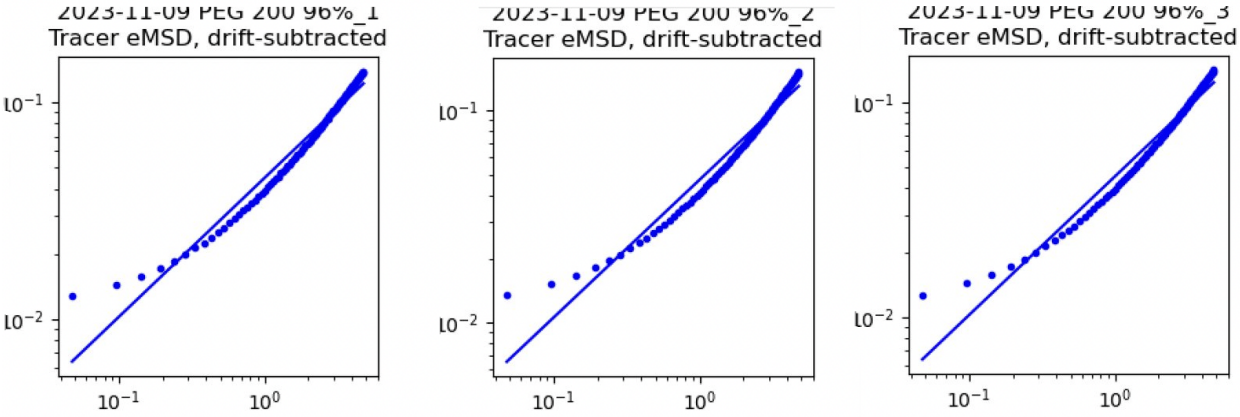

Figure 43: PEG 200 at 96% by volume.

.

A.3 Log-log eMSD Versus Lag Time With Power Law Fit For Molecular Weight Study

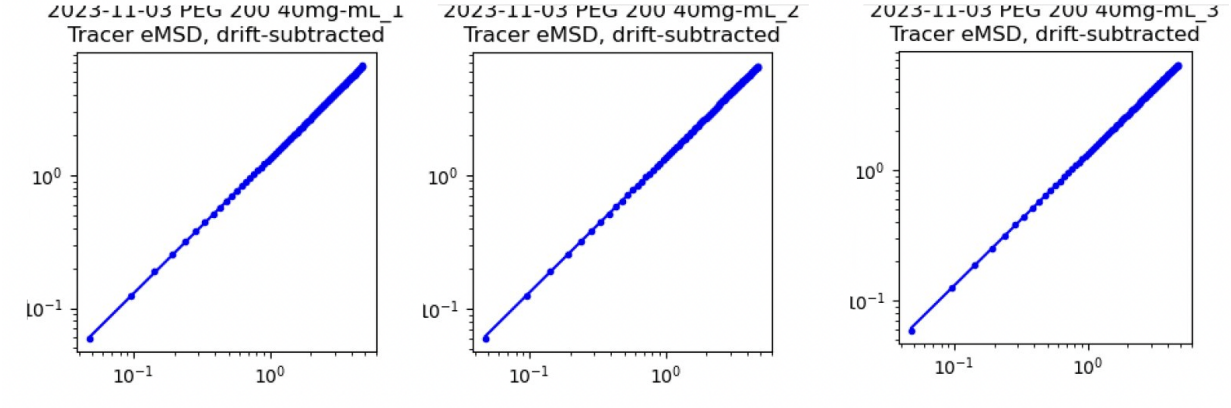

Figure 44: PEG 200 at 40mg/ml.

.

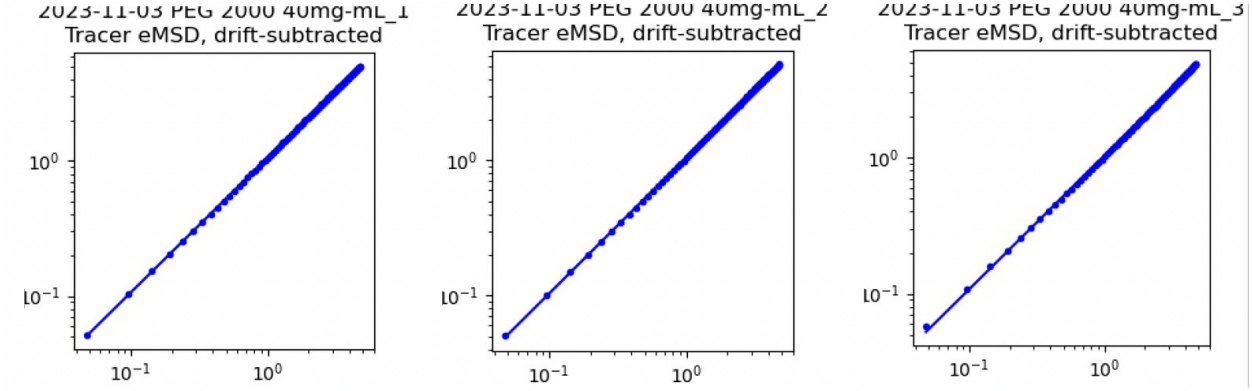

Figure 45: PEG 2000 at 40mg/ml.

.

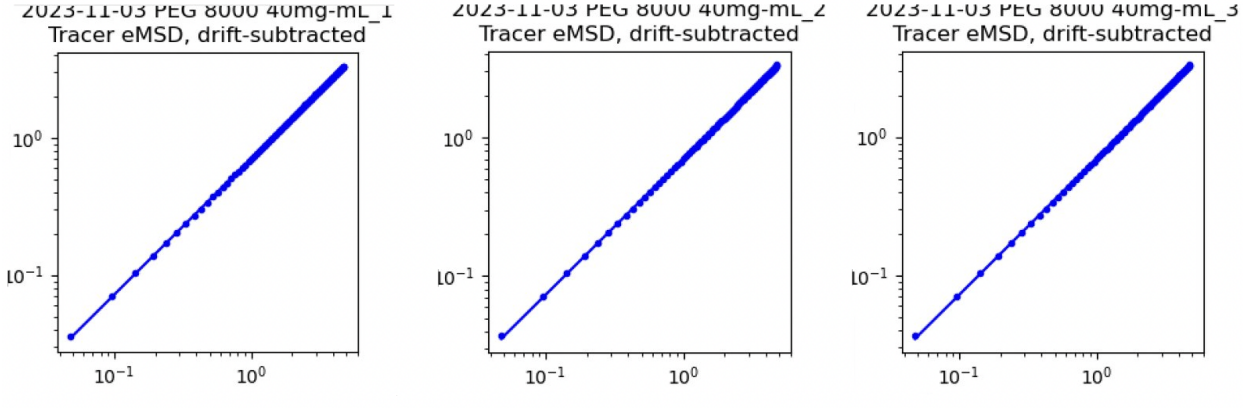

Figure 46: PEG 8000 at 40mg/ml.

.

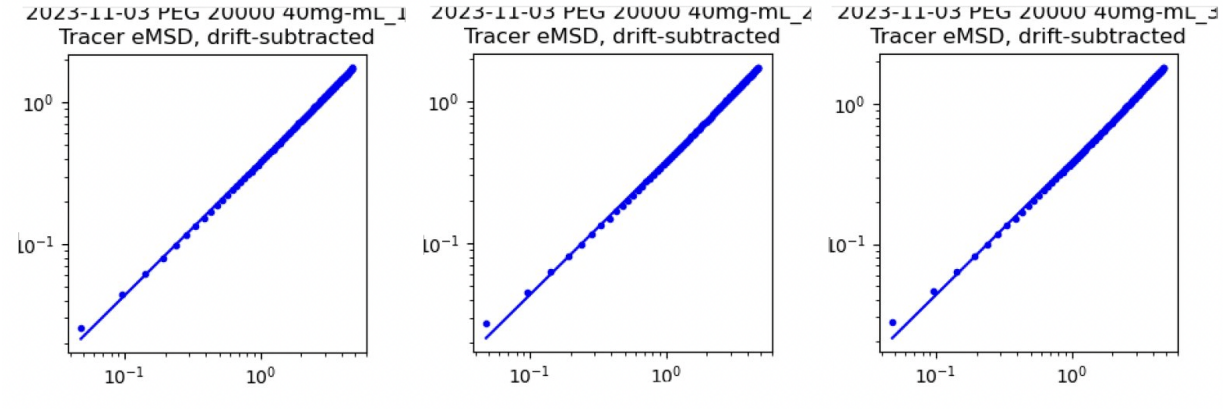

Figure 47: PEG 20000 at 40mg/ml.

.

A.4 Log-log eMSD Versus Lag Time With Power Law Fit For Polydispersity Study

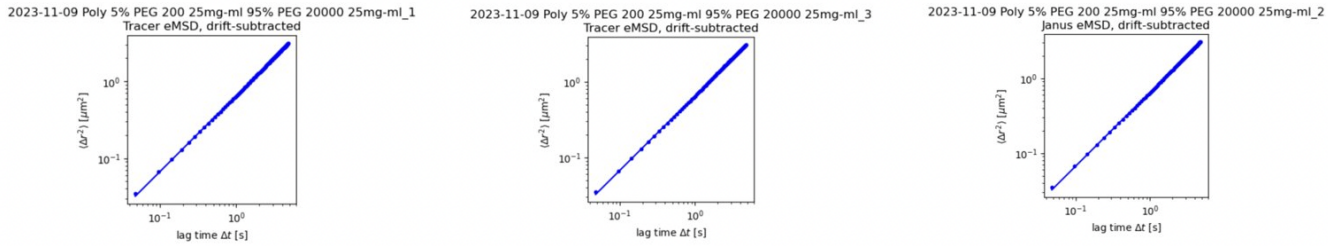

Figure 48: 5% PEG 200: 95% PEG 20000.

.

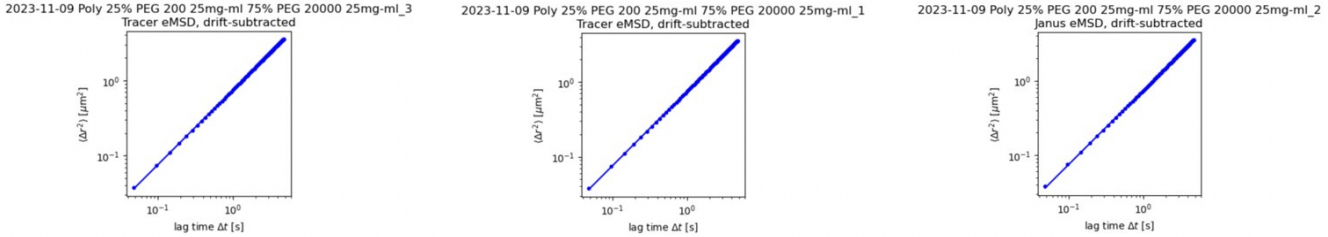

Figure 49: 25% PEG 200: 75% PEG 20000.

.

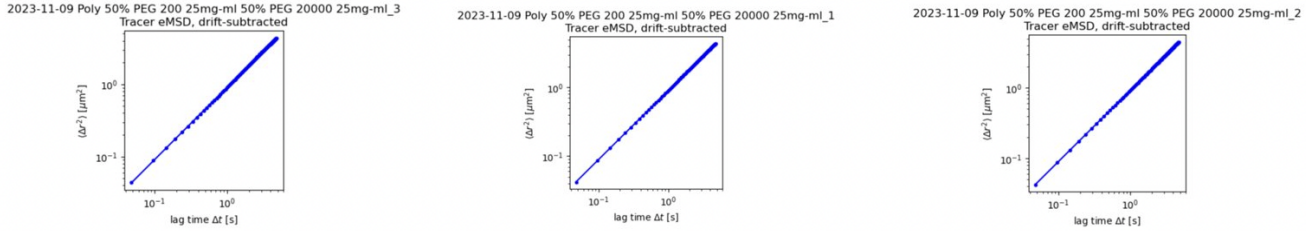

Figure 50: 50% PEG 200: 50% PEG 20000.

.

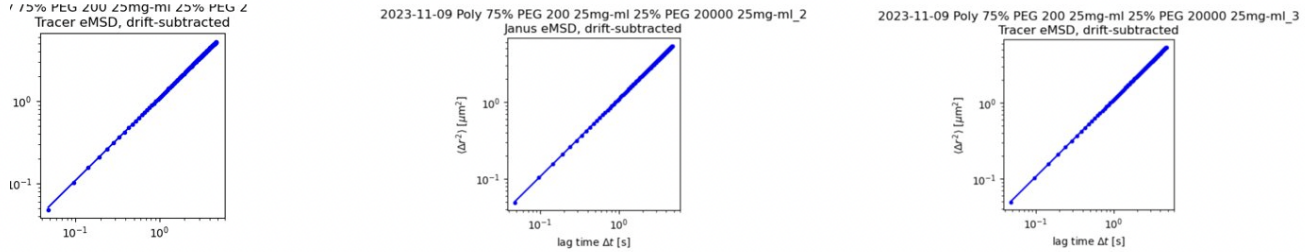

Figure 51: 75% PEG 200: 25% PEG 20000.

.

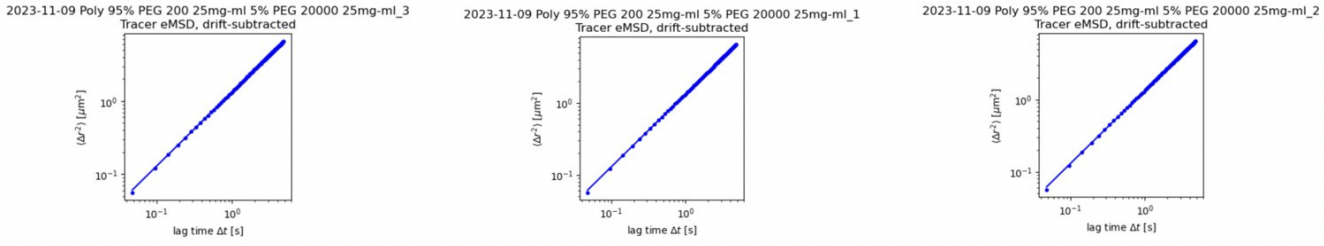

Figure 52: 95% PEG 200: 5% PEG 20000.

.

A.5 Log-log eMSD Versus Lag Time With Power Law Fit For Approximate Monodispersity Versus Polydispersity

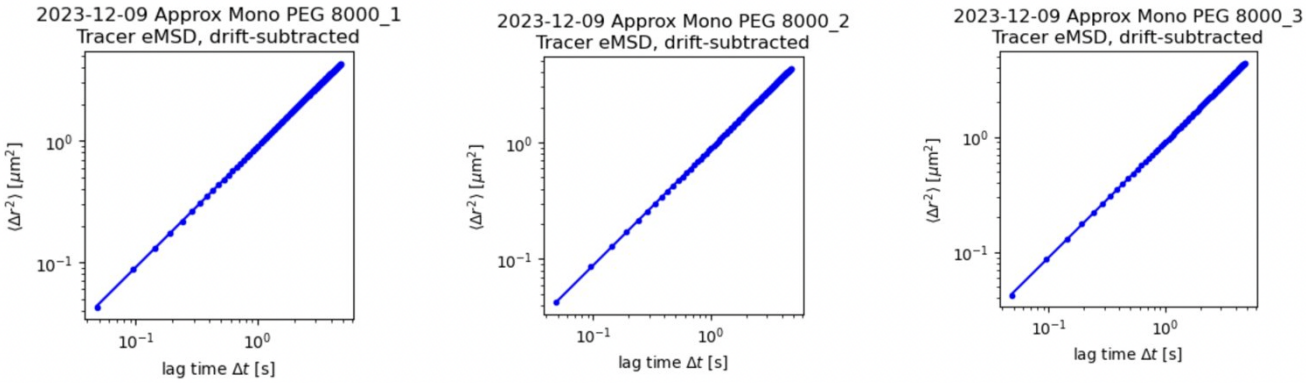

Figure 53: Approximately monodisperse PEG 8000 at 25 mg/ml.

.

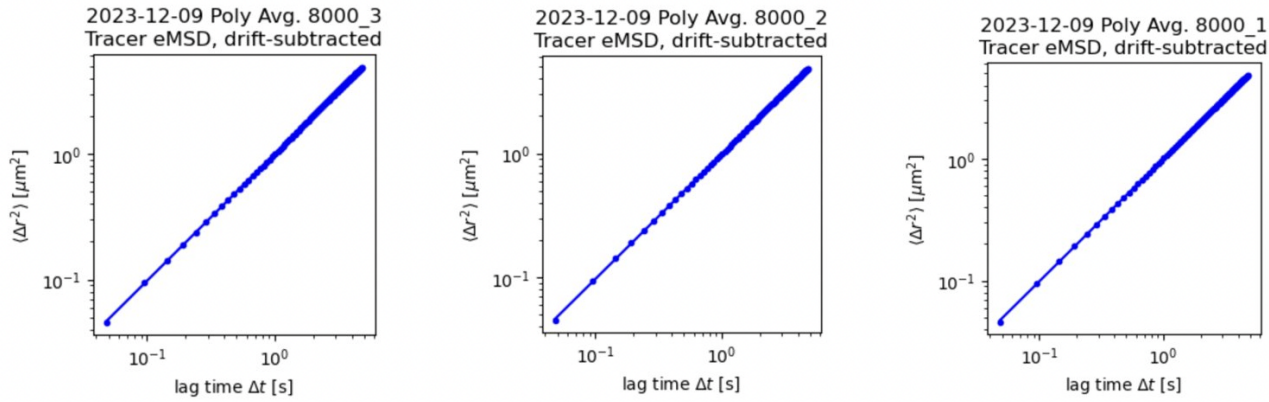

Figure 54: Polydisperse solution of PEG 2000 and PEG 20000 with concentration of 25 mg/ml and average molecular weight of 8000g/mol.

.