Macromolecular Crowding and Intrinsically Disordered Proteins

by

©Yanitza Trosel

A Dissertation submitted to the School of Graduate Studies in partial fulfillment ofthe requirements for the degree of

Doctor of Philosophy

in the

Department of Physics and Physical Oceanography

Faculty of Science

Memorial University of Newfoundland

June 2023

St. John's, Newfoundland

Abstract

The intracellular milieu is crowded and heterogeneous, and this can have profound consequences for biomolecule motions and biochemical kinetics. Macromolecular crowding has been traditionally studied in artificial crowders like Ficoll and dextran or globular proteins such as bovine serum albumin. It is, however, not clear if the effects of artificial crowders on such phenomena are the same as the crowding that is experienced in a heterogeneous biological environment. Bacterial cells, for example, are composed of heterogeneous biomolecules with different sizes, shapes and charges. Using crowders composed of one of three different pretreatments of bacterial cell lysate (unmanipulated, ultracentrifuged, and anion exchanged), we examine the effects of crowding on the diffusivity of a model polymer and an intrinsically disordered protein (IDP). We measure the translational diffusivity, via diffusion NMR, of the test polymer polyethylene glycol (PEG) and the IDP α-synuclein in these bacterial cell lysates. We show that the small ($R_g \sim 5$ nm) test polymer shows a modest decrease in self-diffusivity with increasing crowder concentration for all lysate treatments. The corresponding self-diffusivity decrease in the artificial Ficoll crowder is much more pronounced. For the IDP, we show that α-synuclein has a significant decrease in the diffusivity in the presence of lysate crowders at low concentrations. Moreover, a comparison of the rheological response of biological and artificial crowders shows that while the artificial crowder Ficoll exhibits a Newtonian response even at high concentrations, the bacterial cell lysate is markedly non-Newtonian; it behaves like a shear-thinning fluid with a yield stress. While at any concentration the rheological properties are sensitive to both lysate pretreatment and batch-to-batch variations, the PEG and α-synuclein diffusivity is nearly unaffected by the type of lysate pretreatment.

COVID-19 Statement

As everyone in the world COVID-19 (2020-2021) had a high impact for me. However, this should be a general events statement in my case, since I was an unusual case. Right at the start of the first news about COVID-19 (end 2019), I was diagnosed with breast cancer. I received the cancer news some weeks before I was about to present my comprehensive examination, which had to be postponed. During my chemotherapy treatment (winter 2020), the pandemic lockdown started, and suddenly it felt like the whole world was going through treatment with me.

Even thought I had chemo-brain during most of my treatment, I managed to successfully present my comprehensive in August 2020. During part of this period, research labs at Memorial University were closed.

Naturally, my thesis plans needed to be modified. I had envisioned more diffusion experiments for chapter 7. My goal was to measure the IDP with a bigger range of crowder concentrations, as well as compare it with Ficoll and BSA as the crowders. However, the protein purification process required several weeks of preparation with the risks of protein aggregation, which did happen to me. Meaning that the sample was damage.

Beside my health issues and COVID-19, my experimental work was also affected by the closure of biochemistry laboratories because the department was moving to the new science building. During that period, I had to stop in the middle the protein purification experiments.

Co-authorship Statement

Some of the work presented in this thesis has been published in the journal:

- Trosel, Y., Gregory, L. P., Booth, V. K., & Yethiraj, A. (2023). Diffusion NMR and Rheology of a Model Polymer in Bacterial Cell Lysate Crowders. *Biomacromolecules*. [1]

My contribution to the paper was writing the original draft preparation, editing, conceptualization, NMR experimentation and analysis, DLS and zeta potential experiments, lysate crowder preparation and sample preparations.

The parts of this thesis not included in the paper and authored solely by me are:

- Partially the introduction.
- Chapter 2, all sections with the exception of two paragraph in 2.3.
- Chapter 3, all sections.
- Chapter 4, sections 4.2, parts of 4.3, 4.4 and 4.6.
- Chapter 5, Figure 5.3 and section 5.2.
- Introductory paragraph of chapter 6.
- Chapter 7, all sections.
- Chapter 8, section 8.3.
- Appendix A and B.

I clearly attributed the parts of the thesis where I have consulted the published work of others.

I have acknowledged all sources of help used in this thesis.

The parts of the thesis (some results in chapter 7) where the work was done jointly with others is clearly acknowledged.

Acknowledgements

Words cannot express my gratitude to Dr. Anand Yethiraj and Dr. Valerie Booth, from both who I learned so much. I am extremely grateful they accepted me in their labs, considering I am from a country (Venezuela) with very low to no opportunities to grow as an aspirant scientist, which made me a not strong competitor compared to aspirants from more develop countries.

Thank you, Dr. Yethiraj, for your excellent guidance, I admire your way to approach problems. I appreciate that even if you were not sure about how to solve something in our research, you dedicated the hours to find a way to solve it with me. That taught me on how to approach to scientific problems, which for me it is more valuable than the answer itself. I learned so much from your thought process.

Thank you, Dr. Booth, for being an excellent supervisor, I admire, and I am deeply thankful for your guidance during my PhD. I learned so much from your organization and detail-oriented way to approach any scientific problem, particularly with breaking down into smaller pieces overwhelming scientific question. Your orientation and motivation always kept me on track with the next goal and set of experiments to do. I am deeply thankful with your patience toward me, especially with the biochemistry questions.

I want to also thank to my committee member Dr. Lourdes Peña-Castillo and Dr. Ivan Saika-Voivod. Special thanks to Dr. Mike Morrow for his valuable teaching in NMR. Thank you to the previous and current members of Dr. Booth and Dr. Yethiraj labs, with special mention to Dr. Nury Paula Santisteban, Dr. Sarika Kumari, Sheyla Montero Vega, Liam Gregory and Sina Abdolrahim Poor Heravi.

Thank you to the faculty and staff in the departments of Physics and Biochemistry at Memorial University of Newfoundland.

Thank you to Dr. Céline Schneider for the training and assistance with the NMR experiments.

Special thanks to my funding agencies, Memorial University of Newfoundland, and the School of Graduate Studies, as well as the National Science and Engineering Research Council.

I could not have undertaken this journey without the support of my family, especially my parents Elipsa Arroyo and Freddy Trosel. Thank you for encouraging little me to follow my dream of being a scientist, decision I made when I was still a child. My parents were my first supervisors, and I appreciate they always motivate me to keep curious about

studying and learning. Thank you to my siblings Yeniree Trosel and Freddy Jesus Trosel for their unconditional support.

Lastly, I want to thank to the friends I met in Newfoundland for being the greatest support I could have asked during all my PhD challenging times: Viral Galaiya, Diana Sankar, Dr. Elena Salogni, Thomas Hardy, and Dr. Alfred Dyck.

Contents

List of Figures

Abbreviations

αS	α-**s**ynuclein
Arg	**Arg**inine
CI2	**C**hymotrypsin **i**nhibitor **2**
CD	**C**ircular **D**ichroism
DLS	**D**ynamic **L**ight **S**cattering
DSS	So**D**ium trimethyl**S**ilylpropane**S**ulfonateoly
DOSY	**D**iffusion **O**rdered **S**pectroscopy
DEAE 52	**D**i**e**thyl**a**mino**e**thyl-cellulose resin
E. coli	***E**scherichia **Coli***
FCS	**F**luorescence **C**orrelation **S**pectroscopy
FID	**F**ree **I**nduction **D**ecay
Glu	**Glu**tamic acid
HSQC	**H**eteronuclear **S**ingle-**Q**uantum **C**orrelation
IPTG	**I**so**p**ropyl β- d-1-**t**hio**g**alactopyranoside
IDPs	**I**ntrinsically **D**isordered **P**roteins
IDRs	**I**ntrinsically **D**isordered **R**egions
INEPT	**I**nsensitive **N**uclei **E**nhanced by **P**olarization **T**ransfer
LB	**L**ysogeny **B**roth
Lys	**Lys**ine
NMR	**N**uclear **M**agnetic **R**esonance
OD	**O**ptical **D**ensity
PMSF	**P**henyl**m**ethyl**s**ulfonyl **f**luoride
PGSTE	**P**ulse **G**radient **ST**imulated **E**cho
PFG	**P**ulse **F**ield **G**radient
PGSE	**P**ulse **G**radient **S**pin **E**cho

PEG	**P**oly**E**thylene **G**lycol
PVP	**P**oly**v**inyl**p**yrrolidone
RF	**R**adio**f**requency
SAXS	**S**mall-**A**ngle **X**-ray **S**cattering
SANS	**S**mall-**A**ngle **N**eutron **S**cattering
TXI	**T**riple resonance **X**1+ X2 Nucleus Decoupling **I**nverse 1H Probe
TBS	**T**ris-**b**uffered **s**aline
WATERGATE	**WATER** suppression by **G**r**A**dient **T**ailored **E**xcitation

Physical Constants

Boltzmann constant	k_B	=	$1.380 \times 10^{-23}\ m^2 kg.s^{-2} K^{-1}$
Planck constant	h	=	$6.626 \times 10^{-34}\ J \cdot s$

Symbols

c^*	overlap concentration	mg/ml
D	diffusivity or diffusion coefficient	$m^2 s^{-1}$
T	Temperature	K
h	Height	m
r_h	Hydrodynamic radius	m
A	Area	m^2
F	Force	N
v	Velocity	m/s
η	Viscosity	$Pa.s$
σ	Shear Stress	Pa
$\dot{\gamma}$	Shear Rate	$1/s$
G^*	complex shear modulus	Pa
G'	Elastic modulus	Pa
G''	Loss modulus	Pa
$\tan\delta$	Damping factor	
γ	Gyromagnetic ratio	$rad\ s^{-1}T^{-1}$

Chapter 1

Introduction

How do macromolecules behave in a real intracellular environment when compared with a dilute solution in a test tube? To begin answering this question, let's first focus on the different biomolecules necessary for the function of biological cells. The biomolecules in the intracellular milieu are composed of different proteins as well as nucleic acids such as DNA and RNA.

Proteins are macromolecules with the most diverse function in living organisms. It is estimated that a bacteria of 1 μm^3 of volume has approximately three millions of proteins [5]. The role of proteins ranges from DNA replication to the transport of molecules. Traditionally, proteins have been studied in dilute solutions.

The total concentration of macromolecules (proteins and nucleic acids) within biological cells is between 200 and 400 mg/ml [6–8], depending on the cell type and compartment. Macromolecules make the environment crowded and heterogeneous [9–12]. Due to the high concentration of macromolecules, around 20 and 40% of the physical cell volume occupied is unavailable for other biomolecules [10–15].

The biological effects arising from the high concentration of intracellular and/or extracellular macromolecules that are not functionally related are known as macromolecular crowding [7, 16]. Microscopically, the causes of macromolecular crowding effects are 'hard' nonspecific interactions, also known as excluded volume effects, and 'soft' nonspecific interactions, e.g. electrostatic and hydrophobic interactions [17, 18]. Of course, in real living systems, the crowders could also be substrate molecules that are involved

in enzymatic reactions, so we are considering the simplest cases. Most efforts in understanding macromolecular crowding have considered excluded volume effects, however, such effects are not the whole picture in the crowding world. More recently, Minton, Rivas [19] and Pielak [9, 18] have pointed out the role of soft nonspecific interactions in macromolecular crowding. Since such interactions can be either attractive or repulsive, they either stabilize or destabilize a molecule in crowding. Macroscopically, the total concentration of macromolecules in the intracellular milieu makes the solution more viscous as the concentration increases [20, 21].

Two important parameters affected by crowding are protein size (protein folding) and diffusion. Conversely, most protein diffusion experiments are performed in dilute buffer. Diffusion is one of the principal mechanisms of displacement of proteins within the cell; two examples are enzymes finding substrates, and signalling proteins finding each other. It is important to understand how protein diffusion is impacted by the interactions within the intracellular milieu [22, 23]. A fundamental question that we chose to address is how a realistic crowding environment – characterized by macromolecules with and without charge as well as naturally occurring polydispersity in size – affects test molecule diffusivity [16, 20, 24]. We separated the main question into three research questions: how do biological and artificial crowders compare macroscopically? how do biological crowders like cell lysate affect the translational diffusion of a test molecule, e.g., polyethylene glycol polymer? how do biological crowders like cell lysate affect the translational diffusion of a test molecule, e.g., α-synuclein IDP?.

The diffusion of a test molecule in the presence of crowders has been studied using a bottom-up approach with several artificial crowding agents, such as polyvinylpyrrolidone (PVP) [17, 25], PEG [26], and charged and uncharged versions of the polysaccharides Ficoll [4, 27], and dextran [20, 28, 29]. Other recent studies have used a top-down approach with biological crowder agents such as the globular protein lysozyme [20, 30], the serpin-like protein ovalbumin [20], *E. coli* lysate [20], mammalian cell cytoplasm [31], the negatively charged protein serum albumin [21, 32], and ribonuclease A [32]. A study by Wang *et al* [20] explored the effects of artificial synthetic polymer crowders, biological protein crowders, and lysate on the NMR-measured rotational and translational diffusion of the protein Chymotrypsin Inhibitor-2 (CI2); their study suggests that artificial crowders such as PVP and Ficoll might not be suitable to mimic cellular environments, possibly as a consequence of soft nonspecific interactions in realistic cellular

environments.

Most crowding studies have been done with proteins as the probe molecules. In this thesis, we aim to bridge model systems with real biological cells. In order to increase complexity in a controlled way, we report the diffusion NMR measurements of a probe molecule PEG, and the intrinsically disordered protein α-synuclein (αS), in the presence of biological crowders. PEG is a flexible and uncharged but polar polymer. PEG is also used to precipitate proteins and has been found to phase separate in solution[33]. According to diffusion NMR experiments performed by Palit *et al* [4], the PEG overlap concentration c^* is of order $\sim 10\ mg/ml$; other measurements put c^* higher. The overlap concentration in a polymer indicates (roughly) the point of crossover between the dilute and semidilute regime. Our diffusion NMR experiments in PEG were performed in the semi-dilute regime. We compared the results with artificial polysaccharide crowder (Ficoll). The overlap concentration c^* for Ficoll70 is reported to be $\sim 250\ mg/ml$ by Acosta *et al* [34].

To understand the role of crowder composition, we used bacterial cell lysate, i.e. sheared open *E. coli* bacteria. We employ a top-down approach and use three different treatments of inhomogeneous *E. coli* cell lysate. The main components of *E. coli* lysate are the remnants of the inner and outer membrane, peptidoglycan, ribosomes, proteins, metabolites, inorganic ions, DNA, and RNA [35]. Diffusion of fluorescent proteins in the cytoplasm of the gram positive *Lactococcus lactis* bacteria and the *Haloferax volcanii* archaeon have been previously studied by Schavemaker and coworkers [36]. In their research, they found that the diffusivity of the protein in each cytoplasm depends greatly on the net charge of the protein surface. Since we are using initially uncharged PEG as the probe molecule, we performed our experiments only in *E. coli* cell lysate.

In addition, we tested the rheological properties of both biological lysate crowders and Ficoll to compare their flow behaviours. We wish to correlate the rheology of lysed *E. coli* bacteria, used as crowders, with the diffusivity of probe polymer. The question we sought to answer is how complex crowders like cell lysate with many macromolecular components (either unmanipulated or manipulated by ultracentrifuge and anion exchange) affect both the rheology of the environment and the translational diffusion of two types of probe molecules, PEG or αS. As a remark, the rheological results presented

in this thesis are the first in the literature where the role of the non-Newtonian character of the crowding environment is examined, which is an important consideration in understanding the crowded environment present in real cells. Experiments on the diffusivity of macromolecules in biological crowded environments might help to understand the processes in biological cells. They could also enable us, in the future, to design new and more realistic environments for *in vitro* testing of drug molecules.

This thesis is divided into eight chapters. In chapter 2, we reviewed the theories of diffusion, polymers, macromolecular crowding, and intrinsically disordered proteins. Additionally, we discuss the viscous and elastic response of complex materials. Chapter 3 addresses the experimental techniques used in this thesis. The materials and methods used in the thesis are explained in chapter 4.

Chapter 5 is a rheological characterization of the lysate crowders. In chapter 6, we present the diffusivity of PEG, as the probe molecule, in different lysate crowders. Chapters 5 and 6 results were accepted for publication in *Biomacromolecules*. Chapter 7 shows the diffusivity of the intrinsically disordered protein αS, as the probe molecule, in different lysate crowders. Finally, the conclusions and future directions are presented in chapter 8.

Chapter 2

Background and Theory

2.1 Diffusion

Microscopic molecules, ions, and organelles move in biological cells as a vital mechanism of transport in cells. This phenomenon is known as *diffusion* [12]. Diffusion is the random motion of molecules in a solution due to internal thermal energy, also known as Brownian motion. Brownian motion was first observed in 1827 by botanist Robert Brown while observing in a microscope the jostling of fragment of pollen grain in water [37, 38]. It was not until 1905 that Albert Einstein proposed in his paper *Investigation on the theory of the Brownian movement* that the fragments of pollen grain were big enough to be visible under a microscope but small enough to be affected by the thermal motion of water molecules [39]. Brownian motion is observed in molecules in the nanometer-to- micrometer size range, which is exactly the range of length scales for molecules in cells. In 1926, Jean Perrin won the Nobel prize in physics for his experimental studies on the structure of matter, which proved Einstein's theory of Brownian motion.

Diffusion is the result of a net force experienced by the particle due to the imbalance of collisions by the molecules of the solution, which causes the random or Brownian motion of the macromolecules. The rate of diffusion is described by the diffusivity or diffusion coefficient, which has units of area per time or in SI unit system $m^2 s^{-1}$ [40].

Before Einstein's contribution, one of the first mathematical models to describe the diffusion process was presented in 1855 by Adolf Fick. Fick established the movement of particles from higher concentration to lower concentration, as shown in figure 2.1 [41].

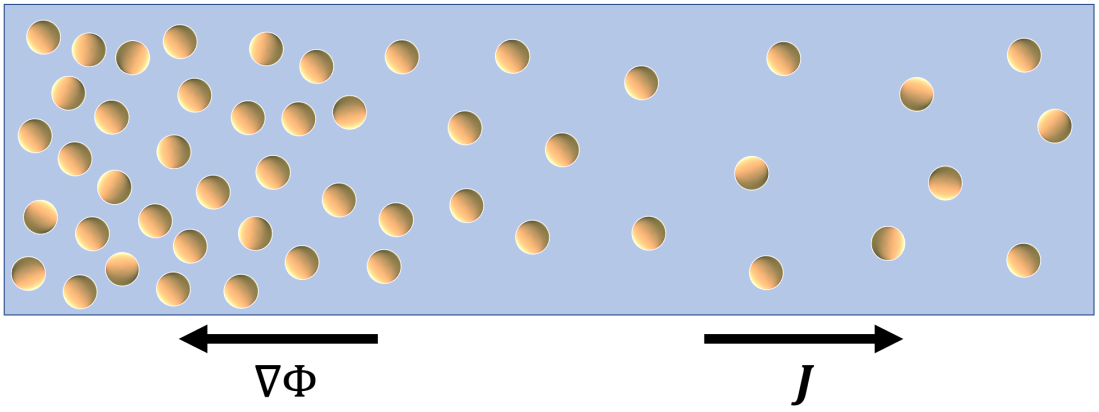


FIGURE 2.1: Representation of the diffusion flux J and the gradient concentration $\nabla\Phi$.

In Fick's first law, the diffusion flux J is proportional to the gradient of the concentration of particles Φ, or

$$\mathbf{J} = -D\nabla\Phi(\mathbf{r}, t), \tag{2.1}$$

where D is the diffusivity or diffusion coefficient. Moreover, Fick relates the rate of change of concentration to the divergence of the flux

$$\frac{\delta\Phi}{\delta t} = -\nabla \cdot \mathbf{J}, \tag{2.2}$$

leading to

$$\frac{\delta\Phi}{\delta t} = D\nabla^2\Phi. \tag{2.3}$$

Equations 2.1 and 2.3 are also known as Fick's first and second law respectively. Fick's laws describe particles' movement in solution from higher to lower concentration, also known as *mutual diffusion* [40].

Although Fick's equations describe the diffusion process, it does not explain the movement of individual diffusing particles. At the beginning of 1900, Albert Einstein (1905) [39], Marian Smoluchowski (1906) [42], and Paul Langevin (1908) [43] were able to independently relate the diffusion coefficient with the trajectory of the particles. They

found a relationship between the mean square displacement $< |\Delta r(t)|^2 >$ as a physical quantity to describe the displacement of a particle related to time t [44], or

$$\langle |\Delta r(t)|^2 \rangle = 2nD\Delta t \tag{2.4}$$

n is the dimension number, and Δt is the time travelled by the particle. Smoluchowski also calculated the mean square displacement, but his calculation was off by a factor of $\sqrt{27/64} = 0.65$ [42].

Einstein's contribution to the understanding of Brownian motion is important because it links a macroscopic measurement like diffusion with the thermal energy of molecules, where no macroscopic concentration is needed in the calculation. Einstein based his derivation on the osmotic pressure of suspended particles in a semipermeable membrane. He found that the diffusion coefficient D is inversely proportional to the Stokes drag $(6\pi\eta r_h)$ of a spherical particle [45], or

$$D = \frac{k_B T}{6\pi\eta r_h}. \tag{2.5}$$

where $k_B = 1.38 \times 10^{-23}\ m^2 kg s^{-2} K^{-1}$ is the Boltzmann constant, T is the temperature, η is the surrounding solution viscosity and r_h is the hydrodynamic radius of the particle. Equation 2.5 is also known as the Stokes-Einstein equation, although William Sutherland arrived at the same result in 1904. The Einstein diffusion derivation refers to a translational *self-diffusion coefficient*. Equation 2.5 is only valid in dilute systems.

2.2 Polymers

Some of the components in biological cells are made up of long chains, which can be classified as organic polymers. Polymers are long chains built up of many repeat molecular units linked by covalent bonds, also known as macromolecules [46]. Polymers can be natural or synthetic. Some examples of natural polymers are DNA, proteins, or polysaccharides like starch. Examples of synthetic polymers are the plastics polystyrene and polyethylene.

Polymers chains may be *linear* or *branched* [47]. Linear polymers can be characterized by the degree of polymerization N, which is the number of repeated units. N is proportional to the relative molecular mass of the chain M. Examples of linear polymers are polyethylene, polystyrene, and most proteins [46].

Branched polymers have secondary chains attached to the primary chain. An example of a branched polymer is the synthetic hydrophilic polysaccharide, Ficoll, which is used in this thesis.

The heterogeneity of a polymer can be measured by its polydispersity index (PDI). The polydispersity index measures how broad the molecular weight distribution is, and is defined as

$$PDI = \frac{\bar{M}_W}{\bar{M}_n}, \tag{2.6}$$

where $\bar{M}_W$ is the weighted average of the molecular weight, and $\bar{M}_n$ is the number-averaged molecular weight. PDI is equal to 1 for monodisperse polymers, which means the chains have the same length [48]. Solutions of polymers in liquids are usually viscous and sometimes show viscoelastic properties; this will be discussed in section 2.5.

PEG is one of the most simple synthetic linear polymers, with chemical structure $H{-}(O{-}CH_2{-}CH_2)_N{-}OH$. PEG is water soluble, and it is highly used in the medical and industrial fields. In this work, we used PEG at a molecular weight 22,000 g/mol and PDI=1.10.

2.3 Macromolecular Crowding

Macromolecular crowding is the biological effect caused by the high intracellular and/or extracellular concentrations of different functionally unrelated biomolecules rather than a high concentration of a single macromolecular species [7, 9, 16]. A schematic illustration in figure 2.2 shows a representation of the intracellular environment in the *Escherichia coli* (*E. coli*) bacteria.

Traditionally, experiments on macromolecular crowding have used artificial crowders like PVP [17, 25], uncharged PEG [26], the polysaccharides charged and uncharged Ficoll

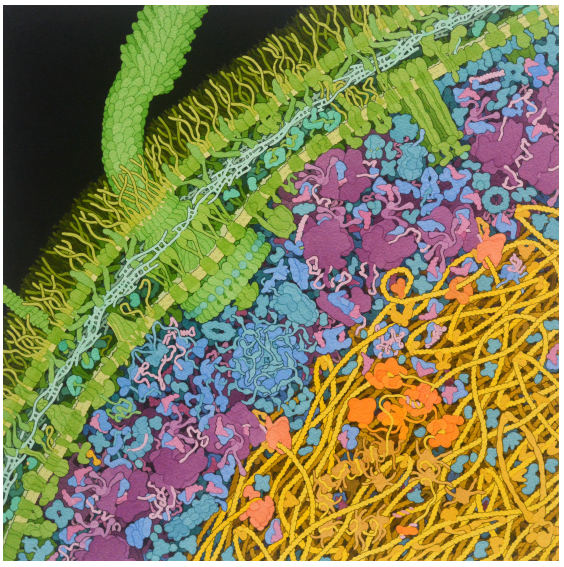


FIGURE 2.2: ***E. coli* bacteria**. Artistic illustration representing the volume occupied by macromolecules inside the bacterium *E. coli*. Illustration by David S. Goodsell [2].

[4, 27], and dextran as the crowders [20, 28, 29]. This approach is known as 'bottom-up' , where the idea is to understand simple models and gradually increase the level of complexity [49]. In contrast, experiments can be explored breaking down the whole complexity of the system within the intact cell, this approach is known as 'top-down'.

Some studies used a top-down approach to understand macromolecular crowding with biological crowders such as globular protein lysozyme [20, 30], the serpin-like protein ovalbumin [20], *E. coli* lysate [20], mammalian cell cytoplasm [31], the negatively charged protein serum albumin [21, 32], and ribonuclease A [32].

The physiological consequences of macromolecular crowding on proteins can be separated into two major effects; changes in the hydrodynamic radius and changes in the associated state. The changes in the hydrodynamic radius are observed in the chemical equilibria such as conformational behavior and stability, protein structure, protein folding, protein shape, binding of small molecules, enzymatic activity and compaction. The changes in the associated state are related to phase separation, aggregation, protein-protein interaction, protein-nucleoid acid interaction and oligomeric proteins [8, 22, 50, 51].

The causes of macromolecular crowding effects are "soft" nonspecific interactions, such as electrostatic and hydrophobic interactions, and "hard" nonspecific interactions, known as excluded volume (steric repulsion) [7, 11, 16, 52–57]. Nonspecific interactions depend upon the overall properties of macromolecules, such as the polarity of surface residues,

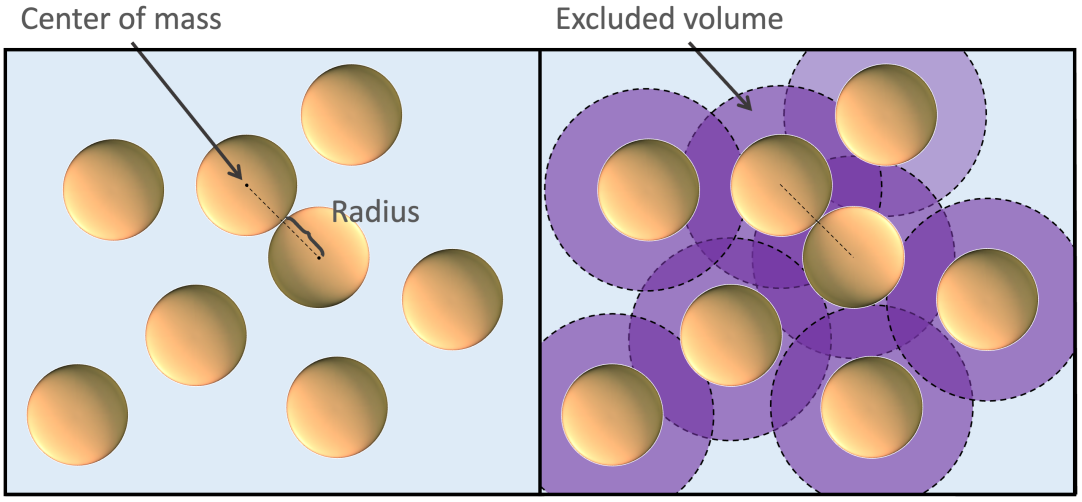


FIGURE 2.3: **Representation of spherical macromolecules in a solution**. (Left) The closest the center of mass of two idealistic spherical macromolecules can touch is the total sum of their radii. (Right) The area in purple represents the excluded volume for other centers of mass of other macromolecules, meanwhile, the green area is accessible for other macromolecules of the same size or smaller.

net charge, dipole or multiple moments and the shape of macromolecules. Nonspecific interactions can be repulsive (steric, electrostatic), or attractive (electrostatic, hydrophobic, depletion) [18, 58].

The excluded volume effects arise from the fact that macromolecules have a physical space occupied that is not accessible for other macromolecules [51]. As a result of steric repulsion, two macromolecules cannot overlap. To understand a simplified example of excluded volume, several spherical macromolecules of the same radius are in a solution (figure 2.3 left). The position of the macromolecules is defined by its center of mass. The closest two centers of mass can get together is the total sum of the two macromolecular radii. Thus, around each macromolecule is an excluded volume for the other center of mass (figure 2.3 right) [14, 16, 54, 58–60].

The presence of excluded volume in the crowded milieu decreases the randomness in the distribution of the crowder macromolecules and the probe molecules, which means that the entropy of the system decrease in comparison with an ideal solution where the crowder molecules occupy no volume [22].

In seeming opposition to steric repulsion, there is an effect known as the depletion attraction. The depletion attraction is also entropic in origin, and occurs when several small molecules and a few bigger macromolecules are in solution. The small molecules

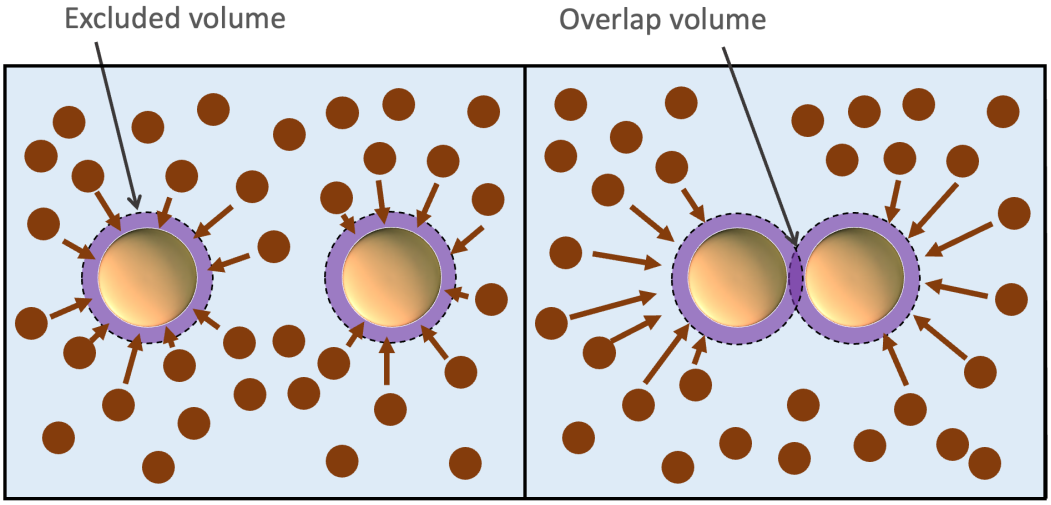


FIGURE 2.4: **Representation of depletion attraction with small crowder molecules and bigger probe macromolecules in a solution**. (Left) The smallest molecules are constantly bombarding the big macromolecules, represented by arrows. (Right) When two big macromolecules get close, the small molecules exert an unopposed force on the opposite side of the big macromolecules, causing an attraction with an overlap volume.

are constantly bombarding the big macromolecules due to Brownian motion (figure 2.4 left). When two big macromolecules get close, the small molecules exert an unopposed force on the opposite side of the big macromolecules (figure 2.4 right). The exerted force derives from the osmotic pressure of the small molecules. In terms of thermodynamics, the solution has more accessible space for the smaller molecules, which means the entropy of the global system is maximized and the excluded volume is minimized [61, 62].

The presence of crowder macromolecules not only causes a change in the microscopic interaction between molecules, it also affects the viscosity of the solution on the macroscopic level. With the increase of the concentration of the total macromolecules in the intracellular milieu, there is an increase in the macroscopic viscosity of the solution [20, 21]. An increase in the concentration within the intracellular milieu should profoundly affect the probe protein environment [22, 23]. In dilute solutions, the viscosity is related to the diffusivity of macromolecules due to Brownian motion by the Stokes-Einstein equation (equation 2.5). In crowded solutions, the Stokes-Einstein equation is technically not valid. Moreover, there can be conformational changes that alter the probe protein's hydrodynamic radius.

One important parameter affected by crowding is protein diffusion. Conversely, most protein diffusion experiments are performed in dilute buffer. Diffusion is one of the principal mechanisms of displacement of proteins within the cell; two examples are enzymes finding substrates, and signalling proteins finding each other. It is important to understand how protein diffusion is impacted by the interactions within the intracellular milieu [22, 23]. Li and Pielak used diffusion NMR to study Chymotrypsin Inhibitor 2 (CI2, 7.4 kDa) and differentiate the effects of soft nonspecific interactions from the increased viscosity of the bovine serum albumin (BSA) crowder [21]. Sarkar *et al* [63] studied CI2 stability in total protein lysate and anion protein crowders. They found that anion protein crowder destabilizes CI2 in a manner similar to that of total protein lysate, despite differences in the crowders' net charges. However, crowder composition does matter as the same probe molecules were slowed down substantially more in *E. coli* cytoplasm (11-fold slower) compared to eukaryotic cytoplasm (3-fold slower) [7], although the eukaryotic cytoplasm is more dilute than the *E. coli* cytoplasm [13]. The nature of the probe molecule also makes a difference. For example in dilute solution, CI2, a globular protein, diffuses faster than the intrinsically disordered protein alpha-synuclein. However, under crowded conditions α-synuclein diffuses faster than CI2 [64]. What causes these differences is unknown [7], but could be related to differences in diffusivity in crowded environments between compact and chain-like probe molecules [4].

Macromolecular crowding typically results in a reduction in the diffusivity of a test probe macromolecule in the intracellular milieu, compared to its diffusion in dilute solution [20, 26, 54, 65–69]. Palit and coworkers (2017) explored the effect of artificial charged and uncharged Ficoll crowders on PEG as the test probe polymer, showing that the crowder charge has a small effect on the diffusivity of the uncharged PEG, reflecting the possible importance of 'soft' interactions [4, 70]. Although artificial and biological crowders have been examined for their effects on the diffusion of macromolecules such as proteins, there is still much to investigate about the role of biological crowders, such as how a realistic crowding environment – characterized by macromolecules with and without charge as well as naturally occurring polydispersity in size – affects test molecule diffusivity [16, 20, 24].

2.4 Intrinsically Disordered Proteins

Cells contain thousands of different protein machines in charge of structure, transport, receiving and sending signals, movement, antibodies and catalysis of biochemical reactions. Proteins are polymers formed by a sequence of amino acids that make up the polypeptide chains.

The function of a protein is determined by its amino acid sequence [12, 35]. Typically, proteins have a well-defined three dimensional structure, which is known as an ordered protein: examples are globular proteins. In *E. coli* bacteria, there are around 3,000 different proteins [35]. Some of those protein's specific function depends on its 3D structure, which is described in four different levels, primary, seconday, tertiary, and quaternary structures.

The primary structure is the amino acid sequence held together by the peptide bonds. The secondary structure is the local structure of the amino acid sequences and is most commonly α-helix and β-sheets stabilized mainly by hydrogen bonds. The tertiary structure is the three-dimensional form resulting from how the secondary structures of the protein pack together to form the overall 3D structure. The tertiary structure is determined by the interactions of the amino acid sidechains. The proteins fold maximizing water entropy by hiding the hydrophobic amino acids inside. Depending on the amino acids involved there can be hydrophobic, and van der Waals interactions, as well as disulfide bonds, ionic bonds, and hydrogen bonds. Quaternary structure is present only in proteins with multiple polypeptide chains and describes how the chains come together. Changing conditions, such as temperature or pH, can cause the protein to unfold, i.e. denature it, and disrupt its function [35].

However not all proteins are organized hierarchically into four different levels of ordered, approximately between 35 and 51% of proteins in eukaryotes are disordered. Meanwhile, in bacteria, the estimate for the fraction of disordered proteins is between 6 and 33% [64, 71]. Disorder in a protein means the absence of a well-defined folded structure. Such proteins are known as Intrinsically Disordered Proteins (IDPs) and disordered regions in a protein are known as intrinsically disordered regions (IDRs). Intrinsically disordered proteins lack entirely (IDPs) or partially (IDRs) of a tertiary structure. A feature of IDPs is their low hydrophobicity and high net charge [35, 72, 73], in opposite to a more

structured protein with typically higher hydrophobicity and lower net charge. This feature is a result of the primary sequence of the IDPs, with the influence of charged amino acid residues such as lysine (Lys), arginine (Arg), and glutamic acid (Glu) [35].

The forces that define a protein's structure are finely balanced and not all proteins have a well defined tertiary structure. IDPs lack a stable three-dimensional structure, making them flexible [74–78]. Their flexibility often allows them to bind multiple partners and thus act like a network hub that synthesize functions, such as recognition, signaling, and regulation [72, 76, 79].

An example of an IDP is α-synuclein (αS), one of the most explored IDPs found in the human brain, known for its role in Parkinson's disease and others neurodegenerative disorders [80]. An exploration of IDPs in the presence of more realistic crowders can help to understand more about their behavior and develop efficient treatments. The importance of IDPs relates to their dynamic ability to bind with different partners and ligands, such as nucleic acids. A failure in an intrinsically disordered protein is more likely to be lethal than for an ordered protein [81, 82].

What does disorder in a protein mean? An ordered protein has a well-defined folding conformation; an example is a globular protein. A disordered protein has no folding conformation on its own. However, IDPs can reach a globular-like form upon binding [83, 84]. In the "protein trinity"model, proteins can exist in one of three states: ordered, collapsed-disordered, or extended disordered [81, 85].

Experimentally, an IDP's structure is mostly characterized by techniques such as nuclear magnetic resonance (NMR) spectroscopy, and circular dichroism (CD) spectroscopy [85, 86]. Hydrodynamic radius can be explored using dynamic light scattering (DLS). Meanwhile, gyromagnetic radius can be measured by scattering techniques such as small-angle x-ray scattery (SAXS) [87] and small-angle neutron scattering (SANS) [77].

Experimental characterization of the structural characteristics of IDPs are important to understand misfolding related diseases. An abnormal aggregation of misfolded and insoluble proteins inside nerve cells gives rise to some neurodegenerative disease; such is the case of Lewy body dementia, commonly known as Parkinson's disease. The major component of the Lewy body is the protein αS [80, 88]. αS is an IDP of 140 amino acids and 14 kDa of size (figure 2.5), mostly abundant in the human brain [89]. αS's

residues have three regions; the residues 1 to 60 have a net positive charge N-terminus, residues 61 to 95 show non-amyloid component hydrophobicity, and from 96 to 140 have a highly net negative charge C-terminus [80, 90]. Interestingly, unlike other common IDPs, αS does not exhibit low hydrophobicity and high net charge [91]. Nevertheless, analysis using CD spectroscopy of α-synuclein shows its lack of a well defined secondary structure [90, 92]. The functions associated with αS disordered regions include targeting, and molecular recognition.

```
  1: MDVFMKGLSK AKEGVVAAAE KTKQGVAEAA GKTKEGVLYV GSKTKEGVVH
 51: GVATVAEKTK EQVTNVGGAV VTGVTAVAQK TVEGAGSIAA ATGFVKKDQL
101: GKNEEGAPQE GILEDMPVDP DNEAYEMPSE EGYQDYEPEA
```

FIGURE 2.5: **Amino acid sequence**. αS amino acid sequence [3].

A small change in the IDP binding partner interaction can cause a big impact in the cell function. On the other hand, the crowded nature of cells can affect the protein's structural preference and protein- protein binding. Moreover, it could affect the IDP function.

What is the nature of the conformation and the structure of IDPs within the crowded cell environment? The answer is under study, but some exploration of IDPs in the presence of artificial crowders suggests that they still have some level of disorder [93].

Furthermore, a diffusion ^{19}F-NMR experiment for a globular protein (CI2) and an IDP (αS) in the presence of artificial and protein crowders suggest that globular proteins diffuse faster in dilute solution compared to the IDP, but, in the presence of crowders the IDP diffuses faster than the globular protein [64].

Other studies using fluorescence correlation spectroscopy (FCS) show that the translational diffusion coefficient of rhodamine green, dextran, DNA, and albumin is reduced in the presence of Ficoll as the crowder [94]. Increasing research on macromolecular crowding suggests that the nature of the crowders, such as size, charge and their composition might affect the diffusing macromolecules [18, 95, 96].

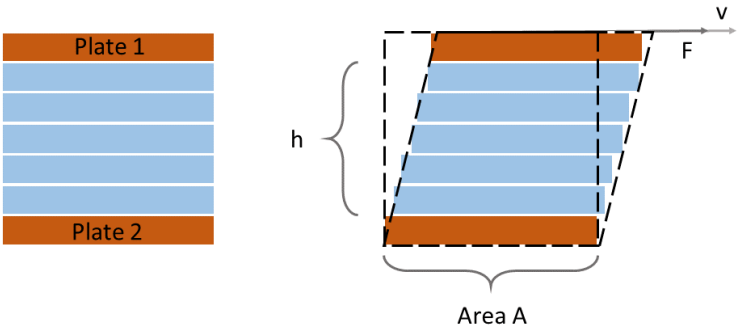


FIGURE 2.6: **Two-plates model**. Two-plates model to calculate the shear stress and shear rate of viscous material.

2.5 Rheology: viscoelastic behaviour

Soft condensed materials such as soup, glue, or pastes cannot be considered as either solids or liquids. The way to classify these materials is by studying their macroscopic shear stress response when a shear rate or velocity gradient is applied [47]. To understand the meaning of shear stress and shear rate, we consider a two-plate model (figure 2.6). This model consists of two plates separated at a distance h with fluid in between them. The upper and lower plates have an area of A. The upper plate is in motion due to a shear force F, with a velocity v. Meanwhile, the lower plate is stationary, with a velocity of $v = 0$. We assume that the fluid is in good contact between the plate, there is no glide, and the flow is laminar (flow by layers) [97].

The shear stress (σ) is defined by the force F applied to the upper plate divided by its area A, with units of pressure Pa. Shear stress is tangential to the surface,

$$\sigma = \frac{F}{A}. \tag{2.7}$$

Meanwhile, the shear rate ($\dot{\gamma}$) is defined by the velocity v divided by the distance h between the plates, with units of $\dot{\gamma}$ is $1/s$, known as *reciprocal seconds*. Shear rate is a measure of the rate at which the fluid layers move in reference to each other,

$$\dot{\gamma} = \frac{v}{h}. \tag{2.8}$$

In a rheology experiment, depending on the response of shear stress and shear rate, we can classify the material according to its flow. In a Newtonian fluid, the shear stress is

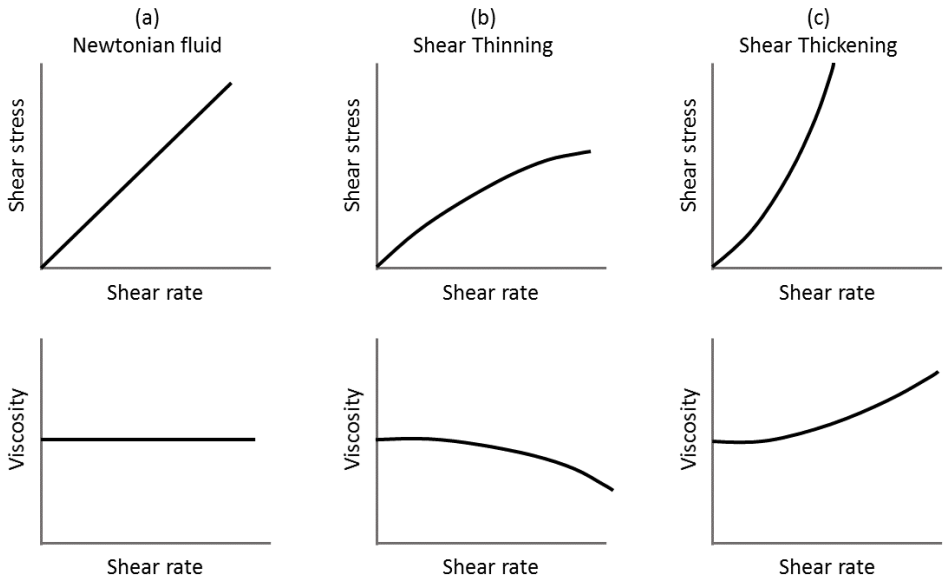


FIGURE 2.7: **Shear stress and apparent viscosity**. Shear stress and apparent viscosity as a function of shear rate for a Newtonian fluid (left), a shear-thinning fluid (middle), and a shear-thickening fluid (right).

proportional to the shear rate (figure 2.7 a), and the constant of proportionality is the viscosity η,

$$\sigma = \eta\dot{\gamma}. \tag{2.9}$$

However, viscoelastic materials have a different response compared to Newtonian fluids. Viscoelastic materials have viscous and elastic behaviors. Some typical flow response for viscoelastic materials are shear-thinning and shear-thickening (figure 2.7 (b) and (c)). The ratio of shear stress σ to shear rate $\dot{\gamma}$ is known as the apparent viscosity. In a Newtonian fluid, the apparent viscosity is constant, and equals the true fluid viscosity. In a non-Newtonian fluid, the apparent viscosity is a function of the shear rate $\dot{\gamma}$ (bottom panels of figure 2.7 (b) and (c))

Shear-thinning flow behavior refers to material where the apparent viscosity decreases with shear rate. Some examples are polymer solutions, shampoos, and glues. In polymer solutions, the shear-thinning behavior depends on the concentration of the polymer. If the concentration is smaller than the critical concentration ($c < c_{crit}$) the polymer solution behaves as an ideal-viscous fluid. If the concentration is greater than the critical concentration ($c > c_{crit}$) the polymer solution behaves as a shear-thinning fluid.

This response is because the polymer macromolecules at rest are organized in a low energy configuration, where the macromolecules are entangled with neighboring macromolecules. When shear is applied, the entangled macromolecules start to open with their length in the direction of the applied shear. As a consequence, as the flow resistance decreases, so the viscosity decreases (figure 2.7 b).

The degree of chain entanglement depends on the concentration of the polymer and the shear load. The apparent viscosity is a function of the shear rate ($\eta = \eta(\dot{\gamma})$). When the shear rate is increased, the number of macromolecules with entanglements decreases: therefore, the viscosity decreases.

In a shear-thickening flow behavior, the apparent viscosity increases with the shear rate. When the applied shear rate increases, the flow resistance increases, so the apparent viscosity increases (figure 2.7 c). Some examples of material with shear-thickening responses are ceramic suspensions, starch dispersions, and dispersions with a high concentration of solid or gel-like particles. One of the reasons for shear-thickening behavior is a hydrodynamic flow instability [97], which is related to turbulence in the fluid.

In flow processes, the complex shear modulus ($G^* = G' + \imath G''$) parameter can quantify the viscous and elastic contribution of the material. G' represents the elastic behaviour, while G'' represents the viscous behavior. The complex shear modulus G^* is related to the shear stress and shear rate by the equation

$$G^* = \frac{\sigma(t)}{\dot{\gamma}(t)}. \tag{2.10}$$

The storage, or elastic, modulus (G') measures the deformation energy stored by the material when a shear is applied. After the load process is removed, the energy stored is used up in the reformation process. The loss modulus (G'') measure the deformation energy used up by the material when a shear is applied. The energy is lost for the sample after the load process is removed [97].

The ratio of the storage and loss modulus of a viscoelastic material is known as the damping factor (or loss factor) $\tan\delta$. The damping factor is calculated using the equation

$$\tan\delta = \frac{G''}{G'}. \tag{2.11}$$

Biological cells are composed of several different sized macromolecules in a solution. When the membrane is sheared open, all the macromolecules are contained in the fluids of the cell. This process is called *lysis* of the cell. The resulting sheared open cells are called *cell lysate*. The methods to lysis are chemical (detergents), acoustic (sonicators), and mechanical (the French press). We used the French press, standard tool to disrupt cells, to lyse our bacteria because it is the least disruptive technique for the macromolecules in the cell.

A study of *E. coli* cell lysate treated with an alkali detergent shows shear-thickening flow behavior at a low shear stress [98]. Meanwhile, for an alkaline *E. coli* lysate treated for plasmid DNA recovery shows a non-Newtonian response at a low shear rate, but a Newtonian flow at a higher shear rate [99]. Furthermore, non-Newtonian behavior has also been observed in other biological environments, such is the case of salmon ovarian fluid. Salmon ovarian fluid was found to have shear-thinning flow, which facilitates the selection of the faster sperm [100].

We tested the rheological properties of both biological crowders and Ficoll to understand if both flow behaviour are comparable and if it is valid to use artificial crowder to mimic cellular environment. Previous rheological studies of intact bacterial *E. coli* in suspensions indicated a non-Newtonian shear thinning flow behaviour [101, 102], both in motile and non-motile intact *E. coli*. In spite this indication, no previous macromolecular crowding studies in the last years have examined the role of the non-Newtonian character of the environment, even though the Stokes-Einstein equation uses a Newtonian assumption.

Chapter 3

Experimental Techniques

3.1 Nuclear Magnetic Resonance

In 1952, Felix Bloch and Edward Mills Purcell were awarded the Nobel prize in physics "for their development of new methods for nuclear magnetic precision measurements and discoveries in connection therewith" [103], a technique that is known nowadays as *Nuclear Magnetic Resonance (NMR)*. NMR is a spectroscopy technique highly used in the fields of medicine, physics, biology, and chemistry. NMR spectroscopy shows the intensity of absorption or emission of energy as a function of frequency, represented as a spectrum. Magnetic resonance results from the interaction between the magnetic moment of the atomic nuclei and an external magnetic field.

The popularity of NMR lies in its ability to provide identification of molecules, sample structure, dynamic in liquids, and others atomic resolution information without being invasive to the sample [104]. There are different approaches to NMR spectroscopy, such as liquid-state and solid-state NMR, both technique can be used for protein NMR. Each approach provides advantages, like information on J-coupling interactions and protein structure. Nevertheless, as does every experimental technique, NMR has some disadvantages. Some of the disadvantages of NMR are that the size of the macromolecule cannot exceed approximately 35 kDa, in high resolution NMR [105], a strong magnetic field is needed to get a large spectrum signal especially in protein studies [104], and in my personal opinion the instrument is economically inaccessible to many scientists

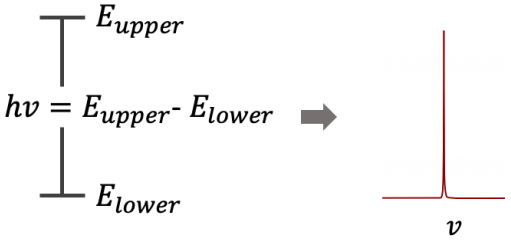


FIGURE 3.1: **Basic description of NMR spectrum using energy levels**. The difference between two energy levels results in absorption line (right side) at a frequency υ.

living in developing countries, because of the need for cryogenics for a superconducting magnet.

To understand a simple NMR spectrum, let's consider the transition between two energy levels of an atomic nucleus in the presence of an external magnetic field B [106]. The difference between two energy levels, with energies E_{upper} and E_{lower}, is

$$\Delta E = E_{upper} - E_{lower} \tag{3.1}$$

A molecule can absorb a photon if its energy is equal to the difference between two energy levels, or moreover

$$h\upsilon = \Delta E = E_{upper} - E_{lower} \tag{3.2}$$

where $h = 6.626 \times 10^{-34} J \cdot s$ is the Planck constant. A NMR spectrum contains lines where the frequencies depend on the separation between energy levels, as shown in Figure 3.1. For an isolated spin-half nucleus, the energy difference between two levels are

$$E_m = -m\hbar\gamma B_0, \quad \text{with} \quad m = +\frac{1}{2} \quad \text{or} \quad -\frac{1}{2} \tag{3.3}$$

where m is the quantum number associated with the spin, $\hbar = h/2\pi$ is the reduced Planck constant in $J \cdot s/rad$, γ is the gyromagnetic ratio in $rad\ s^{-1}T^{-1}$, and B_0 is the magnetic field strength in Tesla (T).

TABLE 3.1: Gyromagnetic ratio γ values for common isotope used in protein NMR.

Nucleus	γ (10^6 $rad\ s^{-1}T^{-1}$)
^{1}H	267.522
^{13}C	67.283
^{15}N	-27.116

In NMR, the value of the gyromagnetic ratio γ depends on the nucleus under study. γ values for nuclei commonly used in protein NMR, the isotopes of hydrogen (^{1}H), nitrogen (^{15}N), and carbon (^{13}C) are listed in table 3.1. The sign of γ is an indication of the sense of nuclear spin precession.

The role of γ in NMR is its relationship with the *Larmor frequency*, ω_0. The Larmor frequency is the product of γ and the magnetic field strength

$$\omega_0 = -\gamma B_0 \quad \text{in} \quad rad\, s^{-1} \tag{3.4}$$

or

$$\nu_0 = -\gamma \frac{B_0}{2\pi} \quad \text{in} \quad Hz. \tag{3.5}$$

From equations 3.3 and 3.1, the transition between two states with a spin value for the lower level $m = +1/2$, and for the upper energy level equal to $-1/2$, and with a positive γ is

$$\begin{aligned} \Delta E &= \frac{1}{2}\hbar\gamma B_0 - \left(-\frac{1}{2}\hbar\gamma B_0\right) \\ &= \hbar\gamma B_0 \\ &= h\gamma \frac{B_0}{2\pi} \end{aligned} \tag{3.6}$$

Comparing equations 3.5 and 3.6, the transition between two energy levels is

$$\Delta E = -h\nu_0 \tag{3.7}$$

As mentioned at the beginning of the chapter, NMR spectra are reported as signal intensity versus a frequency. Since the frequency values are typically in the range of 10 and 800 MHz (radiofrequency wavelength: 30-40 m), there is a practice in the NMR community to present the NMR spectrum intensities as a function of a *chemical shift* (δ) scale, referred in 'parts per million' (ppm) [106]. Doing so makes the frequency axis of the spectra independent of the external magnetic field strength B_0, allowing the spectra from different NMR instrument magnets to be compared easily. The conversion for the chemical shift (δ) is

$$\delta(ppm) = 10^6 \times \frac{\nu - \nu_{ref}}{\nu_{ref}} \tag{3.8}$$

where ν is the frequency of the spectrum line, and ν_{ref} is the frequency for the compound of reference. Traditionally, the compound of reference for ^{1}H, ^{13}C, and ^{15}N is tetramethylsilane (TMS), $Si(CH_3)_4$.

TMS is traditionally used as the reference compound in ^{1}H NMR because has only one hydrogen environment meaning its peak is far away from other peaks. The nucleus is highly shielded by silicon atoms. Biological samples do not usually have silicon, and TMS is soluble in most organic solvents. However, TMS has low solubility in water. Therefore for aqueous samples, like those in protein NMR, sodium trimethylsilylpropanesulfonate (DSS) is used due to its higher water solubility.

3.1.1 Basic NMR experiment and pulse sequence

To understand the basic NMR experiment let's consider the vector model. The vector model establishes the key ideas to understand uncoupled spin experiments without using quantum mechanics [106].

Each nucleus has a nuclear spin magnetic moment, which means the nucleus behaves as a bar magnet that generates a small magnetic field. The small magnetic fields that each nucleus generates are cancelled out by one another in the absence of an external field; as a consequence, there is no net magnetization as shown in Figure 3.2 (a).

When an external magnetic field is applied to the sample (e.g., along the z axis), there is an energetic preference for the individual magnetic moments to be aligned with the

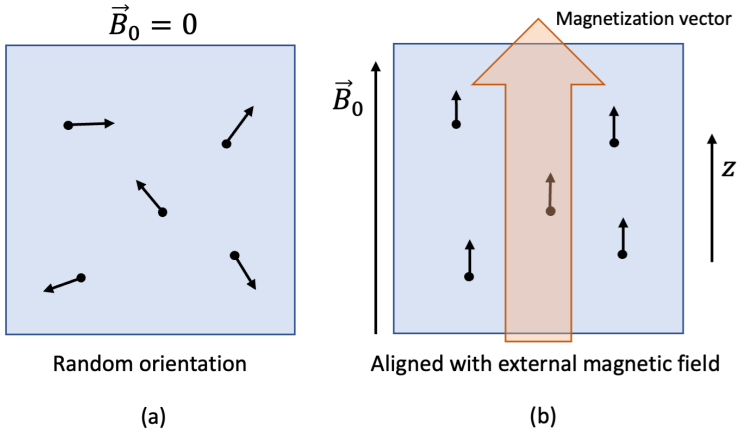


FIGURE 3.2: **Illustration of the nuclei**. (a) Representation of the individual magnetic moments. (b) Individual magnetic moments aligned with the external field.

external magnetic field. The total sum of the individual magnetic moments is known as the *bulk magnetization* of the sample. Represented by a *magnetization vector* (shown exaggeratedly in Figure 3.2 (b)). The net magnetization reaches its equilibrium value along the z axis after several seconds after the external magnetization is applied [106].

What is detected in a NMR experiment is the precession of the magnetization vector after applying a pulse that tips the net magnetization off the z axis. This is done by applying a radiofrequency field using a coil around the sample, aligned in the xy plane. The precession of the magnetization vector, also known as Larmor precession, induces a current in the coil that is amplified and recorded using the same coil. The resulting recorded signal is known as *free induction signal* or *free induction decay* (FID).

Two magnetic fields are involved in a NMR experiment. The first field is the external magnetic field B_0, which aligns with the magnetization vector with the z axis. The second magnetic field B_1 is generated by applying an amplified radiofrequency signal by the x axis coil, same coil as the radiofrequency field and free induction signal detection. The second magnetic field will position the magnetization vector in the xy plane. From a stationary laboratory frame of reference, the magnetization motion is a blur, since the vector is precessing and rotating in the xy plane.

It is useful to consider the behaviour of the bulk magnetization, not in the laboratory frame, but in a frame rotating at the resonance frequency. An analogy to understand the importance of a rotation frame is illustrated in Figure 3.3. In Figure 3.3 (a), a hobbit is trying to read the ring of power inscription from The Lord of the Rings, while Gandalf is spinning it over its center on a plane. From the laboratory frame, the inscription

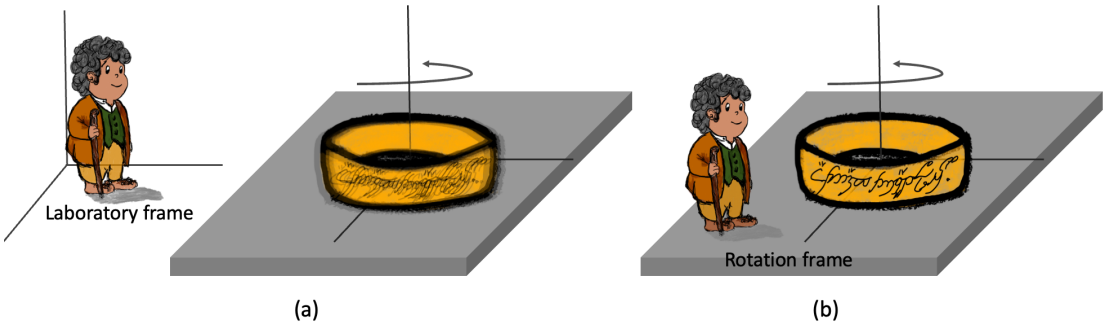


FIGURE 3.3: **Analogy for rotation frame**. (a) The observer (Hobbit) is looking at the rotating ring from the laboratory frame, therefore the ring inscription is a blur. (b) The observer 'jumps' into the rotation frame, and since the ring is rotating at the same speed the ring inscription is easy to read.

is a blur. However, if the hobbit 'jumps' onto the ring's rotating reference frame, the inscription now appears stationary and is therefore easy to read (Figure 3.3 (b)).

In the rotational frame, the effect of B_0 disappears. The frame rotates at the Larmor frequency, which means the magnetization appears to be stationary (apparent $\omega_0 = 0$). In reality, the vector is still precessing, so it is necessary transform the apparent Larmor frequency in the rotating frame. If the frequency in the rotating frame is $\omega_{rot,frame}$, the difference between the Larmor frequency ω_0 and $\omega_{rot,frame}$ is known as the *offset* Ω,

$$\Omega = \omega_0 - \omega_{\text{rot,frame}}. \tag{3.9}$$

The basic NMR pulse sequence (Figure 3.4 (a)) can be divided in three sections. In section 1, the net vector reaches the equilibrium magnetization along the z axis. In section 2, the coil in the xy plane produces a small magnetic field oscillating in time at or near the Larmor frequency along the x axis. This x pulse will rotate the magnetization towards -y, and the magnetization is in the yz plane. This field is known as the radiofrequency or RF field, because the oscillation frequency is typically of order 500 MHz. The 90° pulse takes the system from equilibrium magnetization along the z axis to a precession of the vector in the xy plane. In NMR, pulses are represented as bars as shown in Figure 3.4 (a). Pulses can be applied along the x or y axis. In section 3, the pulse is off and the magnetization keeps precessing until it once again reaches the equilibrium magnetization along the z axis.

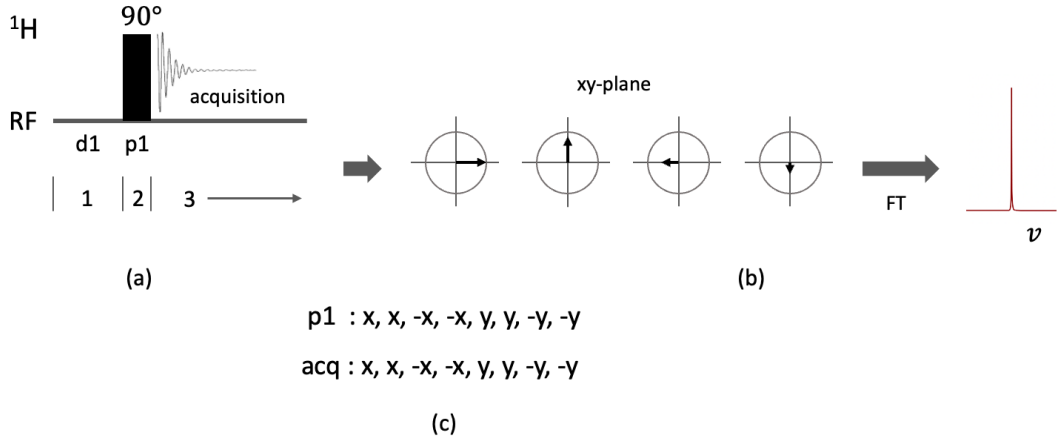


FIGURE 3.4: **Basic ^{1}H pulse sequence and phase cycle**. (a) Basic pulse sequence. (b) View from the xy plane of the magnetization vector precession until equilibrium. (c) phase cycle.

Phase cycling is commonly used to remove artifacts from the spectrum and a simple phase cycling scheme is shown in Figure 3.4. Figure 3.4 (b) shows a schematic representation of the precession vector sequence until equilibrium from the xy plane view during section 2. The phase cycling for a basic pulse sequence is shown in Figure 3.4 (c), the phase cycling is simply averaging over any inhomogeneities since x is arbitrary, x, -x, y and -y are equivalent. The other idea of the phase cycle is to select a particular quantum coherence pathway. In the transmitter phase, the NMR instrument generates the pulses (p1). In the receiver phase, the signal is amplified and acquired (acq). Experimentally, the required duration for a 90° pulse of a sample is calculated by first determining the duration for a 180° pulse when the magnetization vector is along the z axis, therefore resulting in a minimum in the signal. This is the 180° pulse. The 90° pulse is simply half that duration.

3.2 Spin echo

The main principle for diffusion NMR pulse sequences is to employ, in addition to RF pulses, time-dependent magnetic field gradients, referred to as gradient pulses, in a *spin echo* or *Hahn spin echo* sequence. Before explaining more about gradient pulses, let's consider a spin echo pulse sequence with only RF pulses.

So far, we have consider a homogeneous static magnetic field, but in reality, there are some local magnetic field inhomogeneities which result in spatially varying Larmor frequencies. The field inhomogeneities causes the magnetization vectors at different positions to precess at different frequencies. Because of different precession frequencies in the sample in the rotation frame, the vectors spread out. Since the total magnetization vector is the sum of individual vectors, the acquisition signal decays [106].

The main purpose of a spin echo sequence is refocusing the signal. The spin echo sequence is illustrated in Figure 3.5(a). Figure 3.5(b) is a series of graphics that corresponds to the first step in the phase cycling in Figure 3.5(a). The variation of p1 (x, $-x$, y, $-y$) is just geometrical, as discussed for phase cycles. For p2, y, $-y$ applies the 180° pulse in two equivalent ways, clockwise and counterclockwise. Initially, the magnetization vector is in equilibrium and is parallel to the external magnetic field B_0 (the first panel of Figure 3.5(b)). In the x direction, a 90° pulse is applied such that the magnetization vector is along -y, and therefore in the xy plane, as shown in the second graphic of Figure 3.5(b). After removing the RF pulse, the coherence of the spins starts to be lost (the third graphic of Figure 3.5(b), viewed from the rotation frame). These variations in phase are due to inhomogeneities in the external field, spin-spin interactions, and chemical shift.

In order to reverse these variations, a 180° pulse is applied along the y direction at a time τ. The spins will remain in the xy plane but with inverted relative positions as illustrated in the fourth graph of Figure 3.5(b). Since the spins keep precessing, they will momentarily regain coherence at a time 2τ, forming an echo.

3.3 Diffusion/DOSY NMR

Translational self-diffusion measurements can be done by a pulse field gradient (PFG) sequence, sometimes referred to as a DOSY (**D**iffusion **O**rdered **S**pectroscopy) experiment [106]. First, let's introduce the idea of the magnetic field gradient pulse. A magnetic field gradient is applied for a short amount of time δ. The purpose of this gradient pulse is to enhance the dephasing process, and make the signal depend on space. We know that the Larmor frequency is related to the external magnetic field by $\omega_0 = -\gamma B_0$.

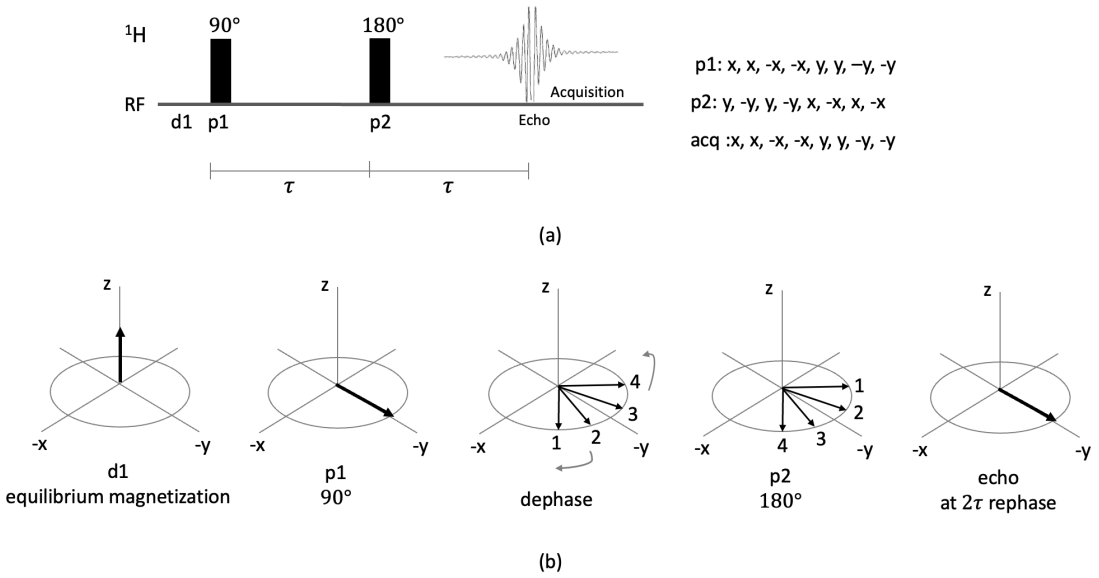


FIGURE 3.5: **Representation of a spin echo pulse sequence and phase cycle**. Two pulses are applied at 90° and 180°.

When an additional magnetic field B_z in the z axis is included, the effective frequency is

$$\omega_{\text{eff}} = \omega_0 + \gamma B_z. \tag{3.10}$$

Assuming that the gradient is constant, $g = \frac{\partial B_z}{\partial z}$, ω_{eff} is

$$\omega_{\text{eff}} = \omega_0 + \gamma g z. \tag{3.11}$$

This means that the effective Larmor frequency depends on the spatial location of the spin along z.

The **P**ulse **G**radient **S**pin **E**cho (PGSE) sequence was first proposed in 1965 by E.O. Stejskal and J.E. Tanner [106]. In PGSE, there are two (nearly square) trapezoidal magnetic-field gradients (g) after each pulse (90° and 180°), as shown in Figure 3.6 (a). Each gradient pulse has a duration of δ, and a separation time between pulses Δ.

If there is no diffusion, the magnetization vector will be rephased during the second gradient. The acquisition signal should not loose amplitude (Figure 3.6 (b)). For diffusing molecules, the total magnetization after rephasing will be reduced, meaning the

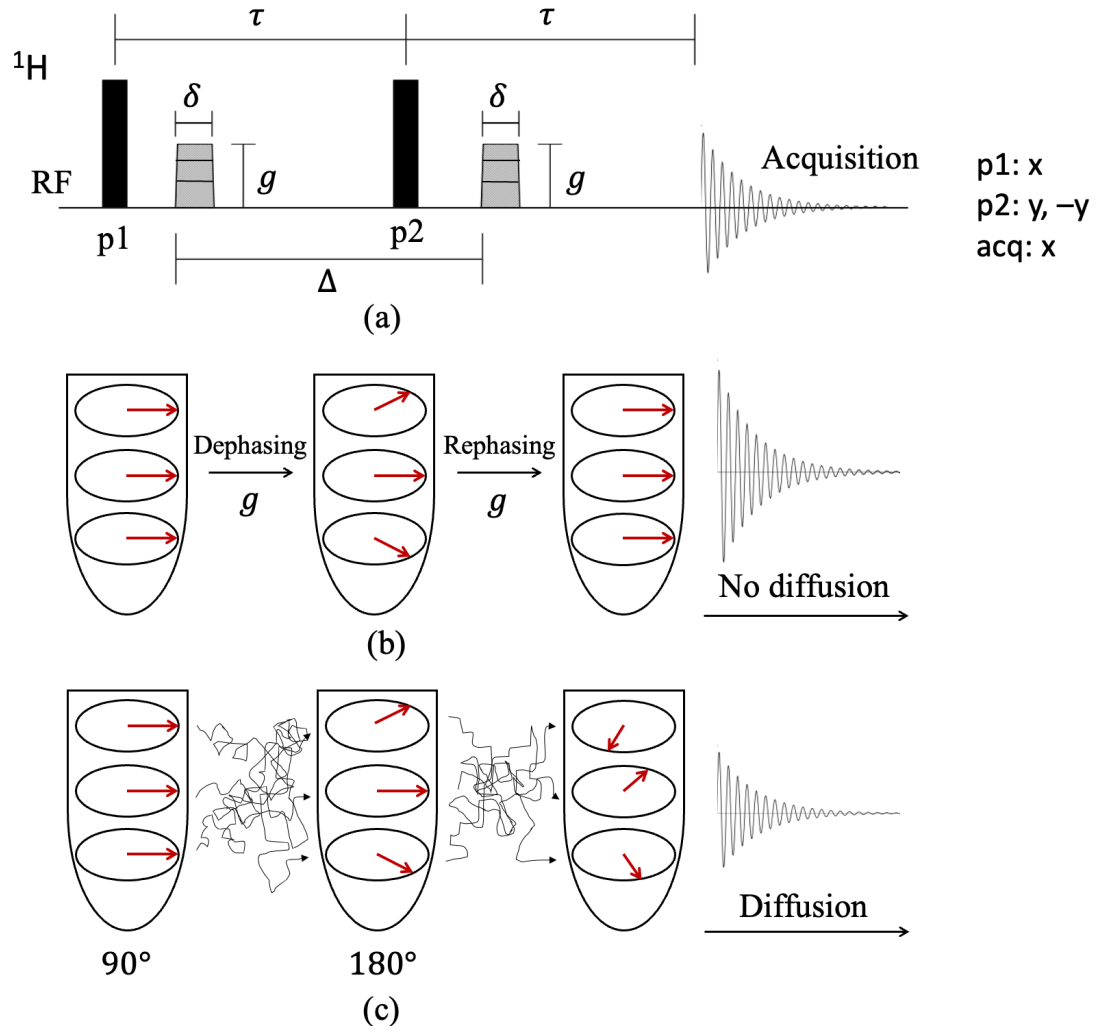


FIGURE 3.6: **Representation of a Pulse Gradient Spin Echo (PGSE) sequence**. Two pulses are applied at 90° and 180°, followed by a gradient. The gradients incremented during the experiment.

amplitude of the total signal will be smaller (Figure 3.6 (c)). The signal decay, processed from the Fourier transform and then selected from a particular spectral region, is fitted by a signal attenuation equation known as the Stejskal-Tanner equation,

$$S(k) = S(0)\exp^{-kD} \tag{3.12}$$

where $k = (\gamma\delta g)^2(\Delta - \delta/3)$, and D is the translational diffusion coefficient.

In this thesis, we used a **P**ulse **G**radient **ST**imulated **E**cho (PGSTE) (Figure 3.7) [107]. In this sequence, there are three pulses (90°, 90°, and 90°). Three gradient pulses are applied after the RF pulses, respectively, as shown in Figure 3.7.

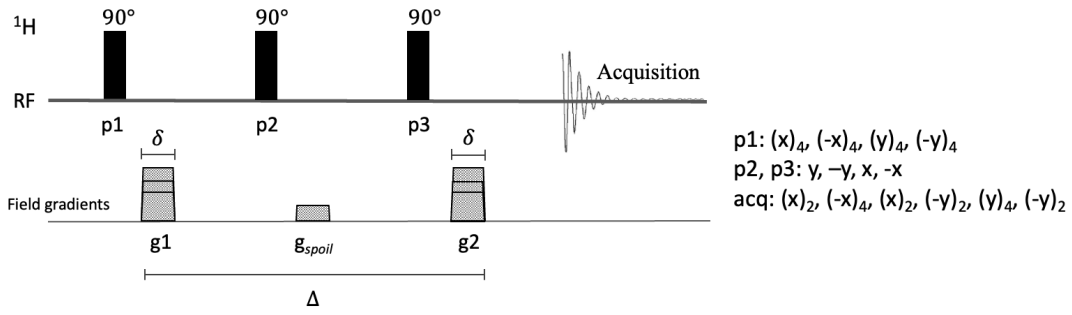


FIGURE 3.7: **Representation of a Pulse Gradient Stimulated Echo (PGSTE) sequence**. Three 90° pulses are applied, followed by a gradient. The second gradient corresponds to a gradient spoil pulse. The gradients incremented during the experiment.

However, the second gradient pulse is a gradient spoil pulse (g_{spoil}). The idea of g_{spoil} is to remove in-plane coherence during the evolution period. Thus all the magnetization is stored along the z axis. The third RF pulse returns the magnetization to the xy plane, and the third gradient pulse rephases the magnetization, either completely when there is no diffusion, or only partially when there is. The signal attenuation equation is unchanged, but the diffusion time Δ is now the time between the first and third pulses.

For DOSY experiments, we used the pulse sequence shown in Figure 3.8 [106, 107]. The difference is that bipolar pulses (g1) are applied. The idea of the bipolar gradient pulses is to apply two pulses that are equal in magnitude but with different directions (i.e., a gradient with magnetic field increasing along $+z$ or increasing along -z). Water resonance peaks are suppressed by the WATERGATE (**WATER** suppression by **G**r**A**dient **T**ailored **E**xcitation) (p5-p6) sequence (discussed in the next section) in all my diffusion experiments. The gradient pulses in DOSY are sinusoidal rather than trapezoidal as in diffusion NMR sequences. When a sinusoidal gradient is used, k in equation 3.12 changes from $k = (\gamma\delta g)^2(\Delta - \delta/3)$ to $k = (\gamma\delta g)^2(\Delta - \delta/4)$.

3.4 ^{1}H NMR with water suppression by WATERGATE

Most organic samples contains water molecules, which means that it is necessary to suppress the water peak (4.7 *ppm*) in the spectra, such that other component peaks are visible. If there is too much solvent signal, the peaks of interest will not be detected. One way to do it is using the WATERGATE technique [107]. This technique is based on the gradient spin echo used in DOSY and diffusion NMR.

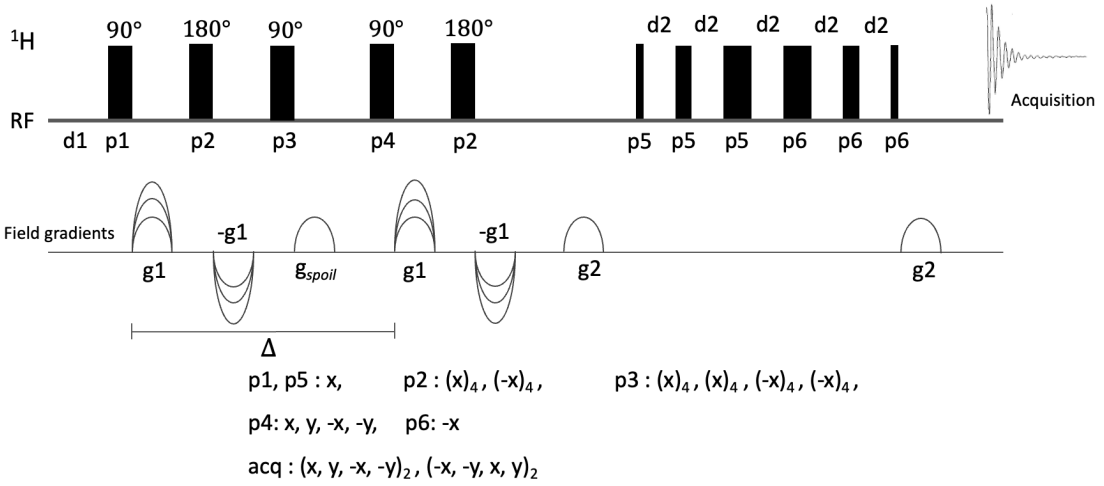


FIGURE 3.8: **Representation of a DOSY sequence and its phase cycle**. Bruker pulse program name is stebpgp1s19, which has a stimulated echo, bipolar gradients, and water suppression. g1 are bipolar pulses. The gradients incremented during the experiment.

The main advantage of this technique is that does not suppress signals from nuclei of interest exchanging protons with the water solvent. This is in contrast to, other water suppression techniques like pre-saturation that irradiate the water resonance, but also suppress nuclei exchanging protons with the water nuclei. The disadvantage of the WATERGATE technique is that also suppress the signal that is very close to the water resonance. Our NMR peaks of interest are sufficiently far away from the water resonance at 4.7 *ppm*.

The pulse sequence (Figure 3.9 (a)) for water suppression by WATERGATE consists of 1 *s* relaxation delay (d1) to reach the equilibrium magnetization. This is followed by a 90° pulse (p1). Then, a sinusoidal-shaped field gradient (g1) is applied for 2 *ms*. The purpose of the pulse is to dephase all the resonance peaks. Following this, a composite pulse (p2, p3, p4, p5, p6, p7) is applied, acting as a 180° over all peaks except on the water, which is on resonance. The composite pulse is a cluster of contiguous pulses with different phases, eliminating errors in the magnetization inversion. p2, and p7: 0.231·p1. p3, and p6: 0.692·p1. p4, and p5: 1.462·p1. Between each pulse (p2 to p7), there is a delay (d2) of 300 μs. A second gradient pulse is applied (g2=g1), dephasing the water resonance peak but rephasing all the other peaks inverted by the 180° composite pulse. The FID is acquired with the water peak suppressed, dephased by the two gradients. The phase cycling for WATERGATE technique is in Figure 3.9 (b).

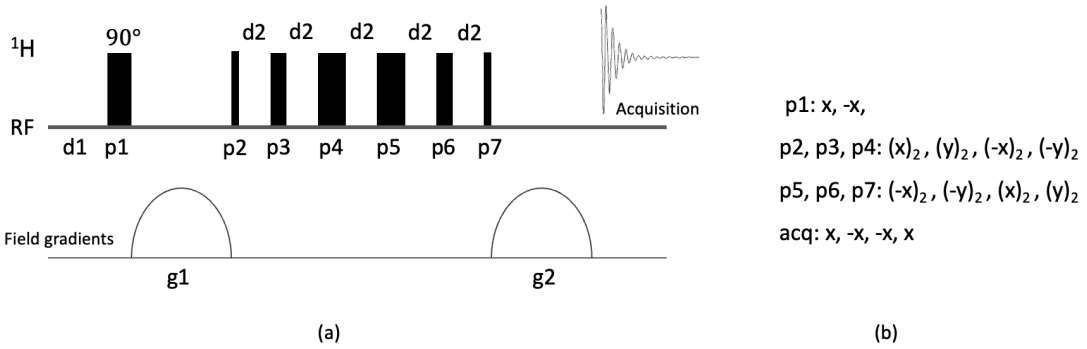


FIGURE 3.9: **Representation of a WATERGATE sequence and its phase cycle**. (a) WATERGATE pulse sequence. (b) Phase cycling.

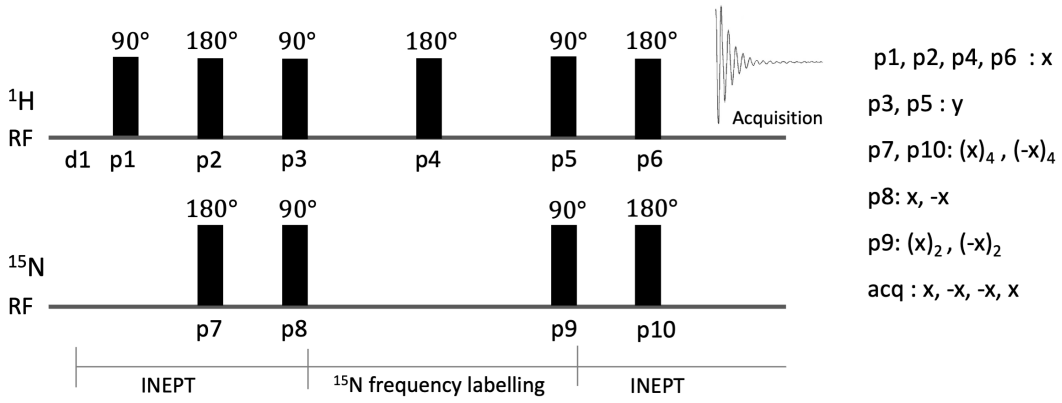


FIGURE 3.10: **A 2-dimensional pulse sequence with RF pulses in the ^{1}H-^{15}N channels**. (a) The HSQC pulse sequence. (b) Phase cycling for the HSQC.

3.5 HSQC

One-dimensional NMR spectra are plotted as intensity versus frequency (or chemical shift). Two-dimensional spectra are plotted as a frequency, F_1, versus another frequency, F_2, with contour lines to represent signal intensity in the third dimension. The **H**eteronuclear **S**ingle-**Q**uantum **C**orrelation or HSQC sequence reveals the direct bonds between proton and nitrogen or carbon [106]. All HSQC experiments presented in this manuscript were measured in the proton ^{1}H and nitrogen ^{15}N channels. Two-dimensional NMR experiments consists of a preparation, evolution, and mixing, followed by the acquisition. The HSQC spectrum of a protein can be used to monitor structural changes.

The HSQC pulse sequence is based on the **I**nsensitive **N**uclei **E**nhanced by **P**olarization **T**ransfer sequence, also known as INEPT. The first section in Figure 3.10 corresponds to the INEPT sequence. The idea of an INEPT experiment is to transfer magnetization from a nucleus with a high gyromagnetic ratio γ, like proton, to a nucleus with a lower

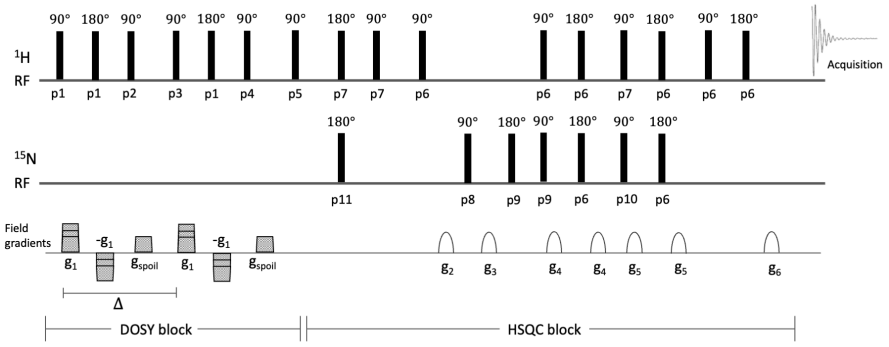


FIGURE 3.11: **Pulse sequence for a HSQC-DOSY experiment**.

γ, like nitrogen or carbon. This is achieved after applying a series of pulses that excite the high γ nucleus, and then trasferring its magnetization to the low γ nucleus. The purpose is to enhance the signal from the heteronucleus.

In the preparation period of a ^{1}H-^{15}N HSQC (Figure 3.10), an INEPT pulse sequence is applied to transfer the magnetization from the proton nucleus to nitrogen. Then, during the middle section, a ^{1}H : 180° pulse is applied to rephase the magnetization. In the second INEPT, the ^{15}N magnetization is transferred back to the proton for acquisition.

3.6 HSQC-DOSY

To measure the translational diffusion of the intrinsically disordered protein αS, a 1D diffusion stimulated echo with bipolar gradients are used with a ^{1}H-^{15}N HSQC pulse sequence. The resulting acquisition is a series of 1D ^{1}H-^{15}N HSQC with increasing gradients [108]. The gradients used in the experiment were 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% of the probe gradient strength [109].

Figure 3.11 illustrates the pulse sequence for the HSQC-DOSY experiment. The first block corresponds to the DOSY pulse sequence, in which the two variable bipolar pulses dephase and rephase the magnetization along + and - z. The second block corresponds to the HSQC pulse sequence, on which the magnetization is transfer from the ^{1}H to the ^{15}N nuclei [110, 111].

3.7 Rheology

Rheological experiments were performed to measure shear stress as a function of shear rate. All experiments were done using a cone-and-plate measuring system as illustrated in Figure 3.12 from the side view (a) and the top view (b). The geometry of the cone-and-plate system consists of a stationary flat circular plate on the bottom, and a circular cone of radius R on the top. The upper circular cone has a cone angle α. For rheological measurements, it is recommended that the cone angle α is between 1° and 4°. In our experiments we used $\alpha = 1°$ [97], with a plate diameter of 50 mm. This is the CP50-1 geometry.

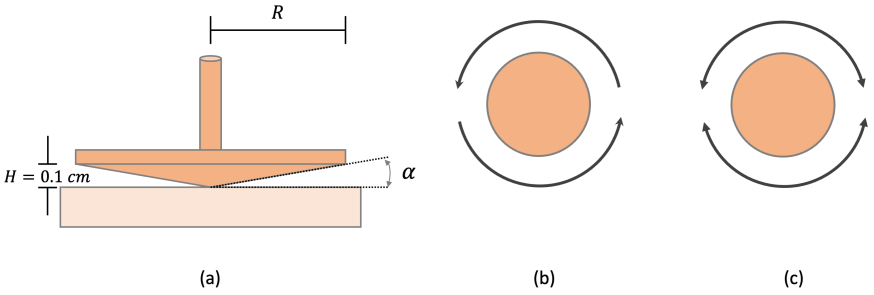


FIGURE 3.12: **Measuring system for rheology experiments**. (a) Cone-and-plate geometry for rotational viscometer. (b) Upper view of continuous rotation for steady shear measurements. (c) Bidirectional rotation for oscillatory shear.

In this experiment, what is actually measured is the torque (mNm), deflection angle (μrad), and the speed at the edge of the cone (min^{-1}). These parameters are related to shear stress σ, and shear rate $\dot{\gamma}$.

Shear stress σ is defined as the fraction between the shear force F and the area A. For a cone-and-plate rheological experiment, the force can be related to the torque ($torque = \mathbf{R} \times \mathbf{F}$), so that

$$\sigma = \frac{F}{A} = \frac{\frac{torque}{R\sin(\theta_{TR})}}{A_{cone}} = C_{ss} \cdot torque. \tag{3.13}$$

C_{ss} is a constant that only depends on the dimensions of the cone, and θ_{TR} is the angle between the force and the direction of R. The units of C_{ss} are $[1/m^3]$ [97]. For shear rate, we know it is defined as the speed divided by distance between both plates. In a cone-and-plate measuring system, the speed and maximum gap are measured at the edge of the cone. From figure 3.12 (a), we know that $\tan\alpha = H_{max}/R$ or $H_{max} = R\tan\alpha$.

For the speed v_{max} is equal to rotational speed ω multiplied by the radius R. Therefore the shear rate $\dot{\gamma}$ is

$$\dot{\gamma} = \frac{v_{max}}{H_{max}} = \frac{\omega R}{R \tan \alpha} = \frac{\omega}{\tan \alpha} \approx \frac{\omega}{\alpha} \approx C_{sr}\omega. \tag{3.14}$$

C_{sr} is a constant that depends on the cone angle with units of $[1/rad]$ [97]. Notice that a small angle approximation was used in equation 3.14, which is the reason why small angle cones are recommended for the measurements using a cone-and-plate system. Figure 3.12 (b) illustrated the upper view of the cone-and-plate measuring system, which rotates in one direction.

For oscillatory measurements, Figure 3.12 (c) shows an upper view representation of the rotations performed by the cone-and-plate measuring system. In this case, the rotation goes back and forward. From this experiments, we get the storage (G') and the loss (G") modulus by the equations:

$$G' = \left(\frac{\tau_A}{\gamma_A} \cos(\delta_A) \right), \tag{3.15}$$

and

$$G'' = \left(\frac{\tau_A}{\gamma_A} \sin(\delta_A) \right), \tag{3.16}$$

where τ_A represents the shear stress amplitude measured in Pa, γ_A is the amplitude sweep in %, and δ_A is the phase shift angle between the preset and the resulting curve.

Chapter 4

Materials and Methods

4.1 Biological crowder preparation

Unmanipulated bacterial cell lysate. In a sterile environment, 75 ml of lysogeny broth (LB) media in a 500 ml Erlenmeyer flask was inoculated with 1 ml of frozen cell stocks of *E. coli* strain JM109 bacteria. The overnight culture was incubated at 30 $^{\circ}C$, 150 rpm between 15 and 20 hours. Using a disposable 10 ml serological pipette, 10 ml of the overnight culture was transferred into $1L$ of LB media in a $4L$ Erlenmeyer flask at 37 $^{\circ}C$. After reaching an optical density with an absorbance at 600 nm (OD_{600nm}) of 0.6 against an LB media blank, the cells were harvested by centrifugation (Thermo Sorvall RC6+ centrifuge) for 15 minutes at $4690g$ and 4 $^{\circ}C$. With the spatula, I transfer the pellet without adding water into the French press.The pellet was lysed using three rounds of French press at $10,000$ psi. Note that the lysate was not subjected to sonication. The lysed cells were flash frozen in liquid nitrogen and then lyophilized by freeze-drying (Labconco Freezone 12 freeze dryer) for three days.

Ultracentrifuged bacterial cell lysate. The starting point of this treatment was after the unmanipulated cell lysate was lysed. High molecular weight molecules and complexes were removed from the lysate crowder. The large nucleic acids, lipid structures and other cell debris in the lysed cells were removed by ultracentrifugation (Beckman L90K centrifuge) at $100,000g$ at 4 $^{\circ}C$ for 40 minutes. Considering a 5 cm centrifuge tube and 1.41 g/cm^3 [112] the average density of proteins in *E. coli*, we estimated that ultracentrifugation should get rid of macromolecules with an approximate hydrodynamic

radius of 10 nm and larger (calculation in appendix B). The supernatant was carefully separated from the pellet using a sterile pipette. The supernatant was flash frozen in liquid nitrogen and then lyophilized for three days.

Anion exchange chromatography bacterial cell lysate. The starting point of this treatment was after ultracentrifugation in the ultracentrifuged cell lysate. To remove some of the negatively charged crowder molecules, 1 g of diethylaminoethyl-cellulose resin (Biophoretics, referred as DEAE 52 in the text) was mixed with 15 ml of TBS (Tris-Buffered Saline: 50 mM Tris-Cl, 150 mM $NaCl$, pH 7.5) was vortexed at room temperature to form a slurry. The DEAE 52-TBS mixture was added to the ultracentrifuged bacterial cell lysate in equal proportion by volume. The solution was stirred for five minutes at room temperature to allow the resin to bind the negatively charged molecules. The stirred solution was centrifuged at $7850g$ at 4 $^{\circ}C$ for 15 minutes and the supernatant carefully separated from the pellet using a sterile pipette. The pellet from the centrifuged was poured onto a PD-10 column (GE Heathcare). The pellet mixed with the resin was discarded. The pH of the collected flowthrough from the column and the supernatant were adjusted to 7.0, which afterwards was flashed frozen in liquid nitrogen and then lyophilized for three days.

Crowder concentration. Initially, we intended to estimate the concentration from fresh (not lyophilized) cell lysate. To do so, 6 flask of *E. coli* strain JM109 bacteria were grown following the beginning protocol of unmanipulated lysate. For each flask, the growth was stopped at different OD_{600nm}, flask 1=0.2, flask 2=0.4, flask 3=0.6, flask 4=0.7, flask 5=0.8 and flask 6=1.0 (Figure 4.1(a)). After reaching the desired optical density, the cells were harvested by centrifugation for 15 minutes at $4690g$ and 4 $^{\circ}C$. The harvested bacteria (pellet) were flash frozen in liquid nitrogen and then lyophilized by freeze-drying for three days. The mass of the dry weight was calculated as $dry.weight = mass_{centrifuge.tube+bacteria} - mass_{centrifuge.tube}$. The dry weight concentration ($dry.weight/volume$) was estimated by dividing the mass of the dry weight by the initial centrifuge tube volume before lyophilization. The dry weight concentration of the growth as a function of the optical density OD_{600nm} is shown in Figure 4.1(b). The curve in Figure 4.1(b) was fitted as $log(dry.weight) = A \times OD_{600nm} + B$, where A and B are constant from the fit. The experiment was repeated a second time to guarantee consistency in the resulting fit. However, the fitted curves were not consistent. Therefore, resuspension from lyophilization were used in all experiments for accuracy in the

final crowder concentration value.

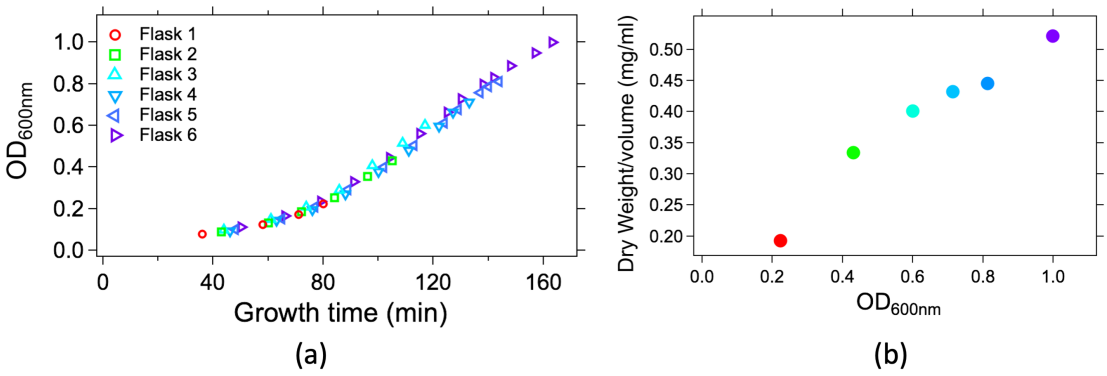


FIGURE 4.1: **Growth curve of *E. coli* bacteria**. (a) Optical density OD_{600nm} of the bacteria *E. coli* as a function growth time. The growth for each flask was stopped at different OD's, starting with flask 1: 0.2, flask 2: 0.4, and lastly flask 6: 1.0. (b) Logarithmic dry weight of cell lysate as a function of the OD_{600nm}.

4.2 α-synuclein protein purification

Non-labelled αS protein purification.

BL21 (DE3) pLysS (Promega) cells had previously been transformed with a pET21 DNA plasmid (Genscript) that codes for the α-synuclein expression construct (see Figure 4.2).

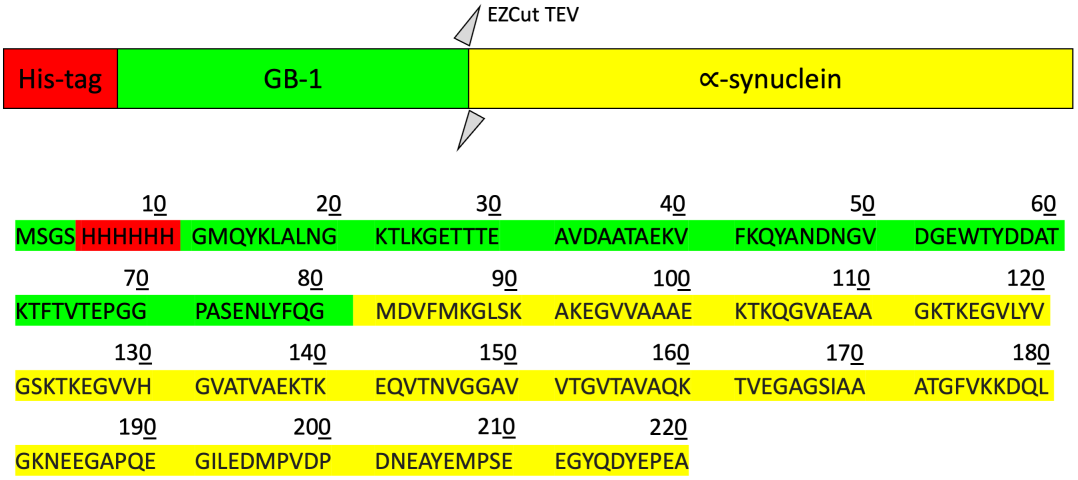


FIGURE 4.2: **Fusion protein sequence**. The red bar corresponds to the 6-histidine tag. The green bar remarks the GB-1 protein attached to the disordered protein αS in yellow bar with the cleavage site marked by grey triangles.

1 ml of transformed frozen cell stock were inoculated into 75 ml sterilized lysogeny broth (LB) media in a 500 ml Erlenmeyer flask with 100 $\mu g/ml$ ampicillin (Fisher, filtered with Sartorius 0.22 μm filter) and 35 $\mu g/ml$ chloramphenicol (Sigma Aldrich, filtered with

Sartorius 0.22 μm filter). The bacteria were incubated (Geneq Inc SI incubator) at 30 $°C$ and 150 rpm overnight for 17 hours. 1 L of LB media with 100 $\mu g/ml$ ampicillin and 35 $\mu g/ml$ chloramphenicol was warmed in a 4 L flask at 35 $°C$ overnight.

Each 1 L LB media was inoculated with 10 ml of the overnight culture. The flasks were incubated at 37 $°C$ and 175 rpm. One of the flasks was measured periodically against a blank of cold LB media in a spectrophotometer (Genesys 10S UV-vis) in order to obtain the OD_{600nm}. Once the bacteria reached an OD_{600nm} of 0.6, 1 ml of isopropyl β- d-1-thiogalactopyranoside (IPTG, Gold Biotechnology) was added to each flask to a final concentration of 1M. The bacteria were incubated until $OD_{600nm} = 1.0$, typically around 30 minutes. The purpose of the IPTG is to trigger the expression of αS. The cells were harvested in a centrifuge (Sorvall RC6+) at 4690g and 4 $°C$. The pellet was collected and resuspended in 20 ml TBS (Tris-Buffered Saline: 50 mM Tris-Cl, 150 mM $NaCl$, pH 7.5) with 1 mM PMSF (phenylmethylsulfonyl fluoride). PMSF is a protease inhibitor. The resuspension was stored at -20 $°C$ until the next day.

Using a french press (Amico), the thawed pellet mixed with TBS/PMSF was lysed at 10 000 psi for three cycles at 4 $°C$. After french press, the lysate was sonicated (Branson sonifier cell disruptor) on setting 7, repeating for three cycles of 30 seconds each. To avoid overheating, the sample was kept on ice. The lysate was centrifuged at 100 000g and 4 $°C$ for 40 minutes using the Beckman L-90K ultracentrifuge and the rotor 70Ti. Using a disposable serological pipette, the supernatant, which contains αS, was removed and placed on ice.

Immobilized metal affinity chromatography (IMAC) beads was used to purify αS from the other bacterial proteins. The expressed αS contains a 6-histidine tag which allows it to bind to the IMAC resin, while the rest of the bacterial proteins should not. The column preparation steps were carried out as follows. To charge 5 ml of sepharose beads (IMAC sepharose 6 fast flow, GE Healthcare) with metal ions, a 0.2 M solution of nickel (II) sulfate hexahydrate (Sigma) ions was dissolved in distilled water. The sepharose resin was placed in a PD-10 column (GE Healthcare) with two column volumes (CV) of distilled water. 1 CV=5 ml. Then, 0.2 CV (1 ml) of the 0.2 M nickel solution was applied to the column. The column was washed with 5 CV (25 ml) of distilled water to remove the excess metal ions. Then, it was washed with 5 CV of loading buffer

(5 mM imidazole, 50 mM Trizma base, 150 mM sodium chloride ($NaCl$), pH 7.5) to pre-equilibrate the column.

The supernatant containing αS was combined with the previously prepared IMAC resin in a cold room at 4 $°C$. The mixture was gently agitated by hand for 5 minutes. The resin with bound αS was isolated by centrifugation (Sorvall RC6+) at 1000g for 15 minutes at 4 $°C$. Using a serological pipette, the supernatant was decanted into a centrifuge tube. The pellet containing the resin with αS bound to it was transferred into the PD-10 column. The remaining flowthrough liquid coming from the column, which should not have αS bound to it, was collected in the supernatant centrifuge tube and stored at 4 $°C$. The resin (pellet) was washed with 5 CV of loading buffer. The flowthrough was collected as 1.5 ml fractions in Eppendorf tubes. Then, the resin was washed with 5 CV of washing buffer (20 mM imidazole, 50 mM Trizma base, 150 mM $NaCl$, pH 7.4). The flowthrough was collected in aliquots of 1.5 ml into Eppendorf tubes. Then, the resin was washed with 5 CV of elution buffer (300 mM imidazole, 50 mM Trizma base, 150 mM $NaCl$, pH 7.4) with the purpose of releasing the αS from the resin. The flowthrough was collected in fractions of 1.5 ml into Eppendorf tubes.

The optical density at a wavelength of 280 nm of each flowthrough fraction was measured using the corresponding buffer (loading, washing or elute) as the blank. The reason 280 nm is used is because that is the maximum absorbance for proteins with aromatic sidechains. The optical densities were measured using a quartz 10.00 mm cuvette in a Genesys 10S UV-vis spectrophotometer. Between each measurement, the cuvette was cleaned with distilled water. Each fraction was stored back in its eppendorf tube. The fractions with higher optical density were pooled and run in a sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) experiment. The fractions with higher optical density are more likely to contain a higher concentration of the αS protein.

SDS-PAGE experiments were performed using a 12% Mini-PROTEAN® TGX^{TM} pre-cast 50 μl 10-well protein gel (Bio-rad). 50 μl of sample buffer (100 mM Tris (Sigma), 0.05% (w/v) bromophenol blue (Biorad), 20% (w/v) glycerol (Sigma), 1 mM dithiothreitol (DTT, Sigma), 4% w/v sodium dodecyl sulfate (SDS, Sigma, SDS was added after adjusting the pH), pH 6.8) were mixed with 50 μl of high OD fraction. The mixture was submerged in a water bath at 100 $°C$ for 5 min. Each well of the cell was filled with 20 μl of the previously boiled mixture, leaving one well for 5 μl of protein marker

as a reference. The loaded gel was set up into a cell with 1X diluted from 10X running buffer (10X stock: 250 mM Tris base, 1.92 M glycine, 1% SDS, 1000 ml distilled water, pH 8.3) at 80 V for 10 minutes, intermediately followed by an increase to 155 V for at least 40 minutes or until the blue dye front reached the black line at the bottom. After electrophoresis, the gel was submerged in Coomassie Brilliant Blue R-250 staining solution (Bio-Rad) and agitated at 100 rpm (Orbit Shaker-lab-line) for 1 hour. Then, the gel was transferred into a destaining solution (5% acetic acid, 10% ethanol) and agitated at 100 rpm overnight. Gel images were captured using ImageQuant LAS 4000.

The expressed protein construct is a fusion protein with αS (yellow) plus guanine nucleotide-binding protein subunit beta (GB-1) (green), as shown in Figure 4.2. The reason why the expressed protein is a fusion of GB-1 and αS is to avoid aggregation, since aS by itself is prone to aggregation. EZCut tobacco etch virus (TEV) protease (Bio-vision) was used to cleave GB-1 from αS. 19 μl of 1 $\mu g/ml$ EZCut TEV (triangles) protease was added per mg of fusion protein. The volume of the sample was doubled with TBS buffer and incubated at 34 °C for 90 minutes. The concentration of the fusion protein was calculated from the absorbance at 280 nm using the Lambert-Beer Law, $A = \epsilon cl$, where A is the measured optical density, $\epsilon = 17420\ 1/Mcm$ is the molar extinction coefficient for the amino acid sequence of the fusion protein (web.expasy.org/protparam), c is the concentration of the absorbing species, and l the path length.

At this point the αS should be cleaved from the GB-1, but a further step is needed to separate the two proteins. This was done with a second IMAC column. GB-1, which retains the 6-histidine tag (his-tag), should stick to the resin, whereas the αS, which now has no histidine tag, should flow through. The his-tag (red) are the six consecutive histidine residues located in the n-terminal (Figure 4.2).

If re-used, the resin was regenerated by adding 5 CV of IMAC stripping buffer (20 $mM\ Na_3PO_4$, 500 mM sodium chloride, 50 mM EDTA, pH 7.4), followed by 5 CV of deionized water. The column is charged with 1.2 ml of 0.2 $NiSO_4$, and then washed with 5 column volumes of deionized water. Finally, the column is pre-equilibrated with 5 column volume loading buffer.

The GB-1 plus αS mixture was applied to the IMAC column, and the flowthrough collected in 1.5 ml fractions. The resin was washed with 5 CV of TEV washing buffer (50 mM Trizma base, 100 $mM\ NaCl$, pH=8.0) and collected in 1.5 ml fractions. Using

quartz cuvette and the spectrophotometer, the optical density of each fraction was measured at a wavelength of 280 nm. Using the highest optical density values, a SDS-PAGE was run and fractions containing pure αS were pooled together.

The pooled sample was dialyzed with a membrane of molecular weight cut-off (MWCO): 6- 8,000 Da (Spectrumlabs) at 4 $°C$ and for 24 hours against sodium phosphate buffer (10 mM sodium phosphate dibasic, anhydrous, 0.05 % sodium azide, pH 7.0). αS has a molecular mass of 14,460 Da, which means that the MWCO of the membrane allows the buffer components to cross but not the protein. 12 % SDS-PAGE was performed with the pooled sample after dialysis (Figure 4.3). Lane 1 (after dialysis) was compared with the flowthrough before dialysis (lane 2), standard proteins (lane 3) and with the fusion protein (lane 4).

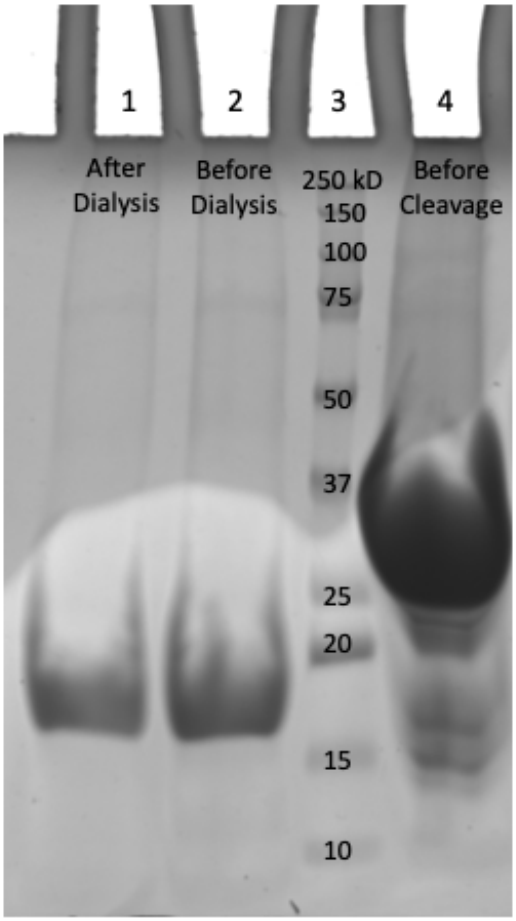


FIGURE 4.3: 12 % SDS-PAGE of αS at different stages of protein purification. Lane 1: αS with an approximate size of 14 kDa. Lane 2: flowthrough before dialysis. Lane 3: protein standard of sizes 250, 150, 100, 75, 50, 37, 25, 20, 15 and 10 kDa. Lane 4: flowthrough before cleavage of the fusion protein.

TABLE 4.1: 400X trace elements minerals for M9 media. Final solution must be filter-sterilized.

Ingredient	Concentration	Source
$FeCl_3 \cdot 6H_2O$ iron (III) chloride hexahydrate	$2.7g/100ml$	Sigma Aldrich
$ZnCl_2 \cdot 4H_2O$ manganese (II) chloride tetrahydrate	$0.2g/100ml$	Sigma Aldrich
$CoCl_2 \cdot 6H_2O$ cobalt (II) chloride hexahydrate	$0.2g/100ml$	Fisher
$Na_2MoO_4 \cdot 2H_2O$ sodium molybdate dihydrate	$0.2g/100ml$	Sigma Aldrich
$CaCl_2 \cdot 2H_2O$ calcium chloride dihydrate	$0.1g/100ml$	Sigma Aldrich
$CuCl_2 \cdot 6H_2O$ copper chloride hexahydrate	$0.13g/100ml$	Sigma Aldrich
H_3BO_3 boric acid	$0.05g/100ml$	Sigma Aldrich
thiamine hydro-chloride	10 % v/v	Sigma Aldrich

^{15}N labelled αS protein preparation. To prepare αS with ^{15}N labels, protein is prepared as above, except that the bacteria that produce the protein are given a growth media that only contains one nitrogen source- and that nitrogen source contains the ^{15}N isotope. 1 L of M9 media was prepared as follows, 6 g/L Na_2HPO_4 sodium phosphate dibasic (Sigma Aldrich), 3 g/L KH_2PO_4 potassium phosphate monobasic (Sigma Aldrich), 0.5 g/L $NaCl$ (Fisher), and pH 7.4 was autoclaved and then incubated at 37 °C overnight before adding the other M9 media ingredients.

The other M9 media ingredients are 1 ml of 100 mg/ml ampicillin (filtered), 1 ml of 35 mg/ml chloramphenicol (filtered with Sartorius 0.22 filter), 2 ml of 1 M $MgSO_4 \cdot 7H_2O$ magnesium sulfate heptahydrate (Sigma Aldrich, filtered), 100 μl of 1 M $CaCl_2 \cdot 2H_2O$ calcium chloride dihydrate (Sigma Aldrich, filtered), 1 ml of 10 mM $FeSO_4 \cdot 7H_2O$ iron (II) sulfate heptahydrate (Fisher, filtered), 1 ml of 5 mg/ml thiamine hydro-chloride (Sigma Aldrich, filtered), 2.5 ml of 400x trace elements minerals (filtered, table 4.1), 1 ml of 1000x vitamin supplements (filtered, table 4.2), 1 g of ^{15}N-labelled ammonium chloride (Cambridge Isotopes, $^{15}NH_4Cl$), and 4 g of glucose (Sigma Aldrich). αS expressing bacteria were grown as in the section 4.1 and then switched into M9 media at the point the overnight culture was inoculated into the large-scale cultures. At OD_{600nm} of 0.6, 1 ml of 1M IPTG (prepared in stock) was added to each flask. The protein was then purified as for the unlabelled protein (section 4.2). It should be noted that cells grew faster in LB media than M9 media.

TABLE 4.2: 1000X vitamin supplement for M9 media. Final solution must be filter-sterilized.

Ingredient	Concentration	Source
D-Biotin	$0.1g/100ml$	Novabiochem
Choline chloride	$0.05g/100ml$	Sigma Aldrich
Folic acid	$0.05g/100ml$	Sigma Aldrich
Myo-Inositol	$0.1g/100ml$	Sigma Aldrich
Nicotinamide	$0.05g/100ml$	Sigma Aldrich
D-Pantothenic acid	$0.05g/100ml$	Sigma Aldrich
Pyridoxal hydrochloride	$0.05g/100ml$	Sigma Aldrich
Riboflavin	$0.005g/100ml$	Sigma Aldrich
Thiamine hydrogen chloride	$0.005g/100ml$	Sigma Aldrich

4.3 NMR sample preparation

PEG in lysate crowder. Polyethylene glycol, with an average molecular weight 22 000 g/mol, polydispersity index 1.10 and neutral charge, was purchased from Polymer Source. Used as the probe macromolecule, PEG in distilled water was added to the lyophilized lysate to a final PEG concentration of 30 mg/ml. The lyophilized cytosol was mixed with distilled water and stirred using a vortex. The clumps were pressed with a spatula and then stirred with the vortex. The process was repeated many times until no clumps were visible. The mix was not filtered. The pH for the resuspended lysate is 7.0 for all experiments. Three kinds of lysate crowders were used in the experiments– unmanipulated, ultracentrifuged, and anion exchanged– at a crowder concentration of 50, 100, 150, 200, 250, and 300 mg/ml. The concentration of the crowders was chosen to include the overall concentration of a cell (200- 400 mg/ml). For each diffusion NMR measurement, the solution was prepared the same day and discarded after the experiment. PEG in distilled water was used as a reference sample.

αS in lysate crowder. ^{15}N-labelled αS (pH 7.4) with a molecular weight 14 kDa at a fixed concentration of (2.08$\pm$0.31) mg/ml or 148.6 μM was added into lyophilized lysate in 90:10 (by volume) H_2O:D_2O and 0.4 mM sodium trimethylsilylpropanesulfonate (DSS). Three kinds of lysate crowders were used in the experiments– unmanipulated, ultracentrifuged, and anion exchanged– at a concentration of 25, 50, 75, and 100 mg/ml. The concentrations were chosen to test αS diffusivity at low crowder concentrations. ^{15}N-labelled αS served as the probe macromolecule. ^{15}N-labelled αS in 90:10 H_2O:D_2O was used as a reference sample. For each HSQC-DOSY NMR experiment, the solution was prepared the same day and discarded after experimentation.

4.4 NMR

^{1}H 1D NMR. ^{1}H 1D spectra were measured in a Bruker Avance II 500 MHz spectrometer. The experiments were performed with a TXI probe with a B_0 magnetic field magnitude of 11.75 T. 800 μl 90:10 H_2O:D_2O of sample was transferred to a 7" long and 5 mm diameter pyrex glass NMR tube (ACP Chemicals Inc.). ^{1}H 1D NMR spectra were measured using **WATER** suppression by **G**r**A**dient **T**ailored **E**xcitation (a pulse sequence known as WATERGATE) [113, 114]. All experiments were performed at a temperature of 25 $°C$. Spectra processing was carried out using TopSpin and MestreNOVA. The phase was corrected by automatic processing and the baseline with an auto-correct 3rd degree polynomial.

^{1}H PFG NMR. Translational self-diffusion coefficients were measured on a Bruker Avance II 600 MHz spectrometer. The diffusion experiments in chapter 6 were performed with a diffusion Diff30 probe with a maximum gradient strength of 1800 G/cm (18 T/m) and a B_0 magnet field magnitude of 14.08 T. The maximum gradient used in our experiments was 1000 G/cm. 1 ml of the sample was transferred to a 7" long and 5 mm diameter pyrex glass NMR tube. Diffusion coefficients were measured using a nearly square trapezoidal gradient pulse in a **P**ulse **G**radient **S**timulated **E**cho (PGSTE) sequence experiment [115]. All experiments were performed at a temperature of 25 $°C$. Analysis of measured diffusion coefficients was carried out using previously described methods [27].

The diffusion coefficients were obtained from the attenuation of the signal according to the bi-exponential equation

$$S(k)/S(0) = f\exp(-D_{\text{fast}}k) + (1-f)\exp(-D_{\text{slow}}k) \tag{4.1}$$

where $S(k)$ and $S(0)$ are the intensities of the signals with and without the field gradient pulses, respectively. D_{fast} represents the diffusion coefficients for the fast component and D_{slow} is the diffusion coefficient for the slow component. The fraction of the fast component is represented by f. The field gradient pulse is given by $k = (\gamma\delta g)^2(\Delta - \delta/3)$, where the gyromagnetic ratio for proton is $\gamma = \gamma^H = 2.657 \times 10^8 T^{-1}s^{-1}$, the duration of the field gradient pulse is $\delta = 2\ ms$, the time between two gradient pulsed is $\Delta = 100$

ms, and the amplitude of the field gradient strength g was from 50 to 1000 G/cm, with 64 gradient steps, and 16 scans at each step. All self diffusion values were calibrated using the diffusion coefficient of trace HDO in D_2O, which is 1.902×10^{-9} m^2/s [116].

DOSY. Translational self-diffusion coefficients were measured in a Bruker Avance II 500 MHz spectrometer. The protein diffusion experiments in chapter 7 were performed with a TXI probe with a maximum gradient strength of 33.71 G/cm and a B_0 magnetic field magnitude of 11.75 T. We used a sinusoidal shape for the gradient. The maximum value for the gradient strength at 100% is 33.71 G/cm. For a smoothed square gradient, the maximum value at 100% is 48.15 G/cm. 800 μl of sample in 90:10 H_2O:D_2O was transferred to a 7" long and 5 mm diameter pyrex glass NMR tube. Experiments were carry out using standard Bruker pulse program stebpgp1s19 to obtain diffusion coefficients by DOSY. The pulse program applies WATERGATE for water suppression and stimulated echoes with bipolar gradient pulse pairs. All experiments were performed at a temperature of 25 $°C$. Analysis of measured diffusion coefficients was carried out using a mono-exponential equation (3.12) of the signal attenuation decay. $\gamma = \gamma^H = 2.657 \times 10^8 T^{-1} s^{-1}$, $\delta = 3$ ms, and $\Delta = 100$ ms.

HSQC-DOSY with ^{15}N-labelled αS. For ^{15}N-labelled αS, diffusion coefficients were measured with a ^{15}N edited technique to avoid the protein signal being overwhelmed by background signals from the lysate. ^{15}N-labelled protein translational self-diffusion coefficients were measured in a Bruker Avance II 500 MHz spectrometer. The experiments were performed with a TXI probe with a maximum gradient strength of 48.15 G/cm and a B_0 magnetic field magnitude of 11.75 T. We used a smooth square gradient. 800 μl of sample in 90:10 H_2O:D_2O was transferred to a 7" long and 5 mm diameter pyrex glass NMR tube. DOSY experiments were carried out using a series of ^{1}H-^{15}N HSQC experiments at 1, 10, 20, 30, 40, 50, 60, 70, 80, and 90 % of the gradient strength [108]. Each 2-dimensional HSQC gradient spectrum was processed in TopSpin as a 1D experiment. The diffusivity values were obtained from the attenuation of the signals peak area under the peaks from each gradient, as shown in Figure 4.4. The stacked spectra in Figure 4.4 shows the peaks from the regions between 10 and 6 ppm. Each spectrum was processed with TopSpin, using automatic phasing and auto-correct polynomial baseline with degree 3.

The first gradient 0.482 G/cm corresponds to 1 %, 4.815 G/cm to 10 %, and so on until

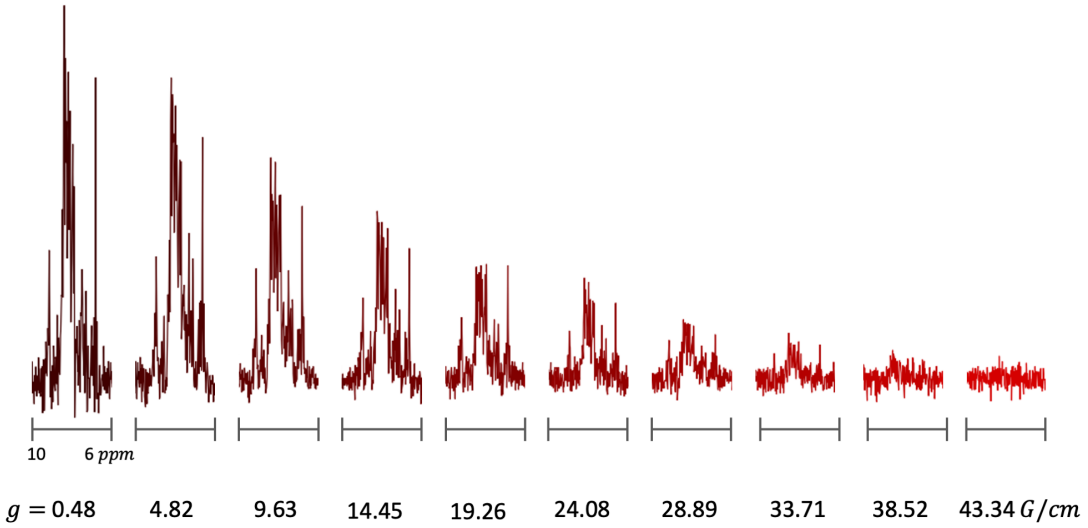


FIGURE 4.4: **Signal attenuation of labelled αS**. ^{1}H-^{15}N HSQC-DOSY of 0.1 mM ^{15}N labelled αS in no crowder. The attenuated signal is shown as a function of the field gradient pulse g. Each spectrum is shown between 10 and 6 ppm.

10 %. The area under the region between 8.7 and 7 ppm was integrated in TopSpin. The first gradient was normalized to 1, and following gradients were adjusted to the previous integration. The translational diffusion coefficient associated with αS was obtained from the signal attenuation of an exponential fitting, also known as Stejskal-Tanner equation:

$$S(k) = S(0)exp(-Dk). \tag{4.2}$$

$S(k)$ and $S(0)$ are the signals with and without the field gradient pulses, respectively. $k = (\gamma g \delta \Delta)^2(\Delta - \delta/3)$, where the proton gyromagnetic ratio value is $2.657 \times 10^8\ T^{-1}s^{-1}$, the duration of the field gradient pulse is $\delta = 3\ ms$, the time period between two pulses is $\Delta = 100\ ms$, and the amplitude of the field gradient pulse g.

The translational diffusion experiments with PEG polymer could be carried out at high gradients ($\approx 1000\ G/cm$, corresponding to $k_{max} \sim 10^{12}\ s/m^2$). For the protein samples in crowded enviroments, we were limited to the TXI probe (in order to use the ^{15}N channel). Thus maximum gradients were $g < 50\ G/cm$, and $k_{max} \sim 10^{10}\ s/m^2$. Thus, the αS diffusion measurements in chapter 7 are not able to detect the slow component of the diffusivity, if it exists.

4.5 DLS and zeta potential

Dynamic light scattering and zeta potential. Dilutions were performed with re-suspended lyophilized lysate in distilled water prior to the measurements. The lysate concentration was first adjusted to 50 mg/ml, and mixed using a vortex mixer. Next, serial dilutions were performed to produce samples at a final concentration of 1 mg/ml. 1 ml of sample was pipetted into a disposable polystyrene cuvette for size measurements. A Malvern Panalytical Zetasizer Nano ZS was used to measure the hydrodynamic size and zeta potential values at 25 $°C$. The viscosity η of the solvent was set in software to 0.89 cP, the refractive index to 1.450, and the equilibration time was 120 seconds. The sample cell used for zeta potential measurements was the ZEN1002 (Malvern). The measurement time was set to a minimum of 50 repetitions. The analysis model used for data processing was "general purpose," since there is a variety of particle sizes in the sample.

4.6 Rheology

Shear stress experiments. Shear stress σ versus shear rate $\dot{\gamma}$ flow curves were the primary rheometry method employed. All rheology experiments were performed using an Anton Paar MCR 301 rheometer with a CP50-1 cone-and-plate tool that is 50 mm in diameter and has a 1° cone angle. All samples were pre-sheared for 60 s before the data was collected. The shear stress σ was measured as a function of the shear rate $\dot{\gamma}$ which was varied from 2 to 200 s^{-1}. The apparent viscosity η at a particular shear rate $\dot{\gamma}$ can be obtained from the equation $\eta(\dot{\gamma}) = \frac{\sigma}{\dot{\gamma}}$.

Steady small shear and oscillatory shear experiments. Oscillatory shear experiments were also performed with the CP50-1 cone-and-plate tool. All samples were pre-sheared for 60 s before the data was collected. Oscillatory measurements as a function of angular frequency were done at 5 different amplitude sweeps: 1, 5, 10, 30, and 50%. The damping factor G"/G' can be obtained by dividing the viscous modulus G" and the elastic modulus G'. The viscoelastic moduli were measured as a function of the angular frequency which was varied from 2 to 200 s^{-1}.

Chapter 5

Rheology of cell lysate

Here we compare the rheological response between *E. coli* bacterial cell lysate and the artificial crowder Ficoll. We compared to Ficoll since it is a commonly used in crowding studies. We performed shear stress versus shear rate experiments to obtain the type of flow and apparent viscosity of our samples. Furthermore, we explore the viscoelastic properties of cell lysate.

In this chapter, we used different batches of bacterial cell lysate. As a guide for the reader, table 5.1 lists the different bacterial cell lysate growth (batch) with its corresponding Figure.

TABLE 5.1: Different batches used in the thesis related to their corresponding figures.

	Batches		
Figure	Unmanipulated	Ultracentrifuged	Anion exchanged
5.1	1 & 2	3	4
5.2	2	-	-
5.3	2	-	-
5.4	2	3	4
5.5	1	-	-

5.1 Rheological characteristics of bacterial cell lysate

We studied the bulk rheological properties of the different treatments of lysate. Figure 5.1 shows the dependence of the shear stress σ (Figure 5.1(a)) and apparent viscosity $\eta = \sigma/\dot{\gamma}$ (Figure 5.1(b)) as a function of the shear rate for two batches of unmanipulated

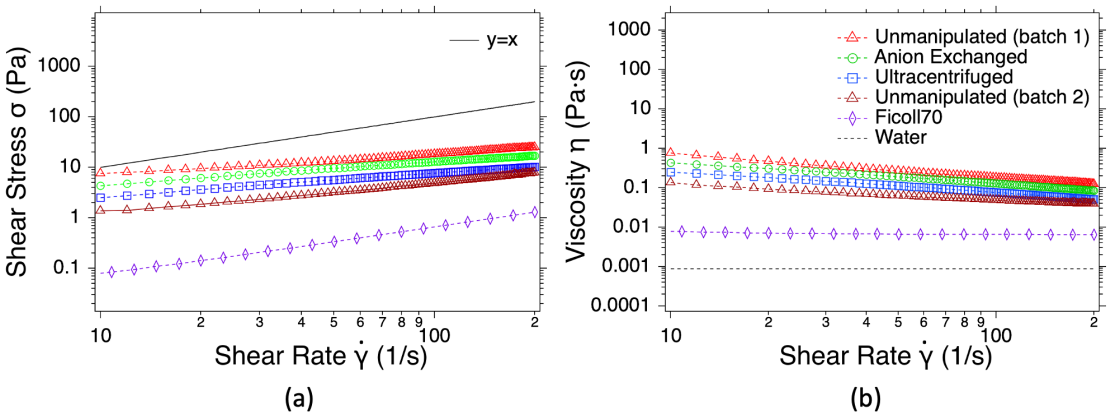


FIGURE 5.1: **Rheology of crowders at 200 mg/ml**. (a) Shear stress versus shear rate of 200 mg/ml cell lysate and Ficoll at 25 $°C$. The line $y = x$ of slope 1 is shown as a guide to the eye. (b) Viscosity versus shear rate of 200 mg/ml cell lysate and Ficoll. The three lysate treatments show shear-thinning flow. Viscosity of water is shown as a guide to the eye.

cell lysate as well as ultracentrifuged and anion exchanged lysates. We compare our cell lysate results to Ficoll. The σ vs $\dot{\gamma}$ plot for Ficoll (purple diamonds in Figure 5.1(a)) is linear and has a slope of unity on a log-log scale, i.e. Ficoll solutions have a Newtonian rheological response, $\eta = \sigma/\dot{\gamma}$, that is a constant, independent of the shear rate [4]. For the unmanipulated, ultracentrifuged and anion exchanged cell lysates, the flow curves are not linear but exhibit power law behavior, i.e., they are linear on a log-log scale, but with a slope less than 1. In addition, they do not approach zero stress at zero shear rate, i.e., they exhibit a "yield stress". There are also variations from batch to batch for the lysate, as shown in Figure 5.1 (a,b) for unmanipulated lysate. Nevertheless, in spite of these variations, the cell lysates are distinct from Ficoll.

To capture both the non-zero yield stress and the power-law dependence on shear rate, the shear stress curve was fitted using the Herschel-Bulkley (HB) model [97],

$$\sigma = \sigma_0 + k_0 \dot{\gamma}^n \tag{5.1}$$

where σ is the shear stress, σ_0 is the yield stress, k_0 is the consistency index, and n is the flow index. A fluid is shear-thinning, e.g. like ketchup, when $n < 1$, and thickening, like cornstarch in water, when $n > 1$. The HB model is a generalized rheological model for a non-Newtonian fluid. For a Newtonian fluid, $\sigma_0 = 0$, $n = 1$ and $k_0 = \eta$ is the bulk viscosity. Another way of visualizing this non-Newtonian behaviour is that the apparent viscosity $\eta = \sigma/\dot{\gamma}$ decreases with shear rate (Figure 5.1(b)). Since this apparent viscosity

decreases with lysate concentration, they are shear-thinning, while the Ficoll solution is not.

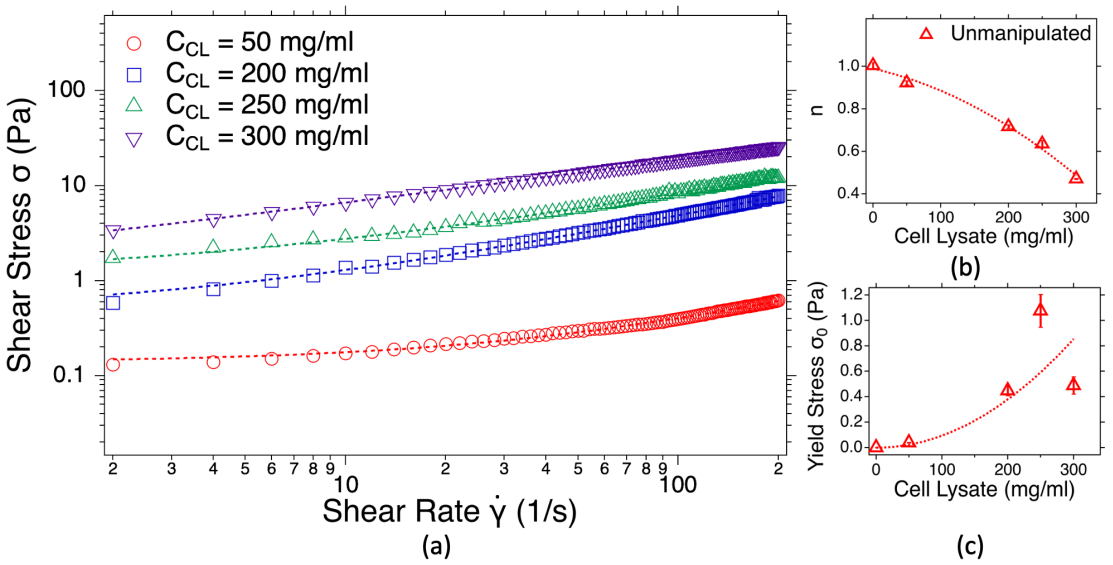


FIGURE 5.2: **Rheological characterization of unmanipulated cell lysate at 25** $°C$. (a) Shear stress versus shear rate. (b) Herschel-Bulkley exponent n for lysate concentrations of 50, 200, 250 and 300 mg/ml decreases from $n = 1$ at 0 concentration to $n = 0.471 \pm 0.002$ at 300 mg/ml: the more concentrated cell lysate shows more pronounced shear-thinning. (c) Yield stress σ_0 versus lysate concentration was obtained from the Herschel-Bulkley fit, and increases from 0 to (0.8 ± 0.2) Pa. Error bars represent standard deviation in the fit. The dotted lines in Figures (b) and (c) are a guide to the eye.

Focusing on a single batch of unmanipulated cell lysate, we examine the concentration dependence of the rheological properties. In Figure 5.2(a), we see the shear stress versus shear rate of unmanipulated cell lysate at concentrations of 50, 200, 250, and 300 mg/ml. Each concentration was successfully fitted using the HB model, obtaining a HB exponent n. In Figure 5.2(b), the HB exponent n decreases with increasing unmanipulated lysate concentration, from 1 (i.e., Newtonian) at 0 mg/ml to ~ 0.5 at 300 mg/ml. The values of the HB exponent n decrease steadily from 1 to 0 with increasing concentration. Thus, the unmanipulated lysate exhibits non-Newtonian, shear-thinning behaviour that is more pronounced at higher concentration.

From the fits to the HB model, the yield stress as a function of the concentration of the unmanipulated lysate is shown in Figure 5.2(c). The yield stress is a representation of how much stress must be applied in the unmanipulated lysate in order to start it flowing. As observed from Figure 5.2(c), the yield stress σ_0 increases from zero to a value of 0.8 ± 0.2 Pa as the concentration of lysate increases from 0 to 300 mg/ml. The

elasticity ($\sim 1\ Pa$) is very weak; in comparison jello would have an elastic modulus on the order of kilopascals [117].

A possible explanation for shear-thinning behaviour in bacterial cell lysate might be due to macromolecular entanglement. In the presence of shear, cell lysate macromolecules are oriented and disentangled in the direction of the shearing. Meanwhile, Ficoll is roughly an artificial compact macromolecule, which cannot align in the direction of the shear load.

In order to test the non Newtonian behaviour at different temperatures, we measured the stress response from 0 to 40 $°C$, as shown in Figure 5.3(a). As a first observation, 200 mg/ml unmanipulated cell lysate shows a shear thinning behaviour when increasing the temperature from 0 to 40 $°C$. Before changing the temperature during the experiment, the initial measurement was done at 25 $°C$. Later, the temperature was changed to 0 $°C$ and gradually increasing until 40 $°C$, repeating the measurement at 25 $°C$, for each stress-rate curve. Both curves, the first measurement stress-rate at 25 $°C$ and the second measurement during the temperature gradually increasing, did not show any flow behaviour or shear stress magnitude differences.

As for Figure 5.2, we fitted each stress-rate curve to the Herschel-Bulkley model (equation 5.1) using a global fit in Igor Pro 8, where HB exponent was fixed to result in the same value for all curves. From the yield stress σ_0 as a function of temperature (Figure 5.3(b)), we observe σ_0 value increases at 30 $°C$, with the exception of 0 $°C$. In addition, the consistency index k_0 shows an abrupt change at 30 $°C$. From k_0 values at 0 $°C$, the lysate shows more movement resistant compared to the same-batch lysate at 25 $°C$. However the curve direction changes at 30 $°C$, where k_0 increases again after initially decrease. This result might suggest that maybe diffusion experiments would be better to study at the body temperature at 37 $°C$. We disregard the abrupt change in yield stress at 0 $°C$ could be an artefact due to freezing. However, the abrupt increase in yield stress at 30 $°C$ is likely a real effect that should be examined further. It should be stated that in spite of this difference, the essential result that the cell lysate shows a non-Newtonian HB rheological response is valid at all temperatures.

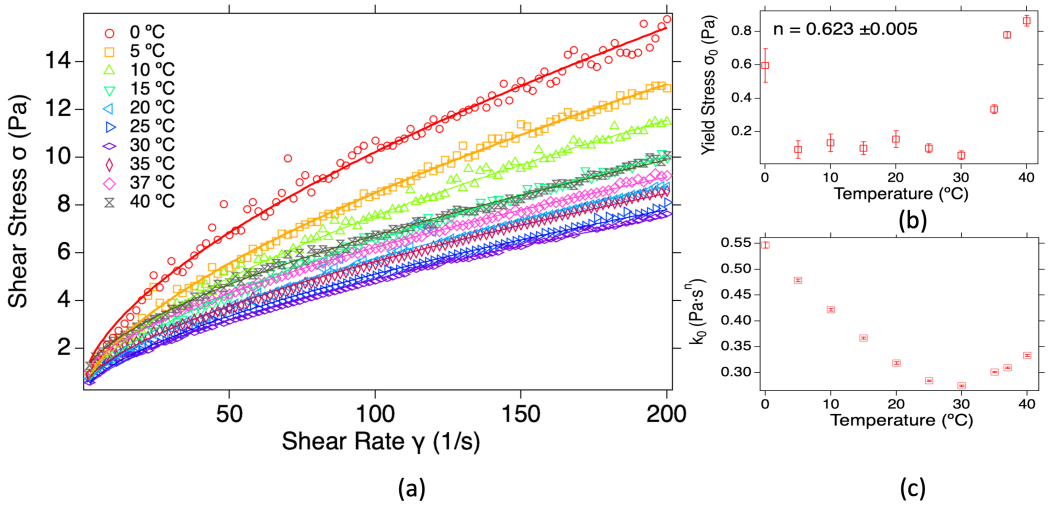


FIGURE 5.3: **Temperature dependent rheology of unmanipulated cell lysate at 200 mg/ml**. (a) shear stress versus shear rate measured at different temperatures for same sample of lysate. (b) Herschel-Bulkley exponent n for the unmanipulated cell lysate as a function of the temperature. (c) Yield stress σ_0 versus temperature. The yield stress was obtained from the Herschel-Bulkley fit. Error bars represent standard deviation in the fit.

5.2 Steady small shear and oscillatory shear experiments

We tested the viscoelastic response of cell lysate. Samples of unmanipulated, ultracentrifuged and anion exchanged lysate were investigated at the amplitude $\dot{\gamma}_A$ of 5%. In Figure 5.4, qualitatively, the storage G' and loss G'' moduli increases with increasing the angular frequency ω. The angular frequency dependence is sublinear for both G' and G''. In the case of G'', a sublinear ω dependence with $G'' \sim \omega^{\alpha}$ ($\alpha < 1$) is consistent with the shear thinning behaviour observed in Figure 5.1. Comparing the magnitudes of G' and G'', we see that $G'' < G'$ for the unmanipulated cell lysate which means the elastic-like behaviour dominates over the viscous-like, but G' and G'' are very similar in magnitude for the ultracentrifuged and anion exchanged lysates which means both behaviour are present at 5%.

The damping factor (G''/G') is calculated from the division of the viscous modulus (G'') by the elastic modulus component (G'). For values < 1, elastic-like behaviour dominates in the lysate. While for values of the damping factor > 1, viscous-like behaviour dominated in the lysate. For $G''/G' = 0$, it means that the elastic modulus completely dominates over the viscous modulus, i.e. ideal-elastic behavior. For $G''/G' = \infty$, the viscous modulus completely dominates over the elastic modulus, i.e. ideal-viscous behavior. For $G''/G' = 1$, the viscous and elastic behavior are exactly equal.

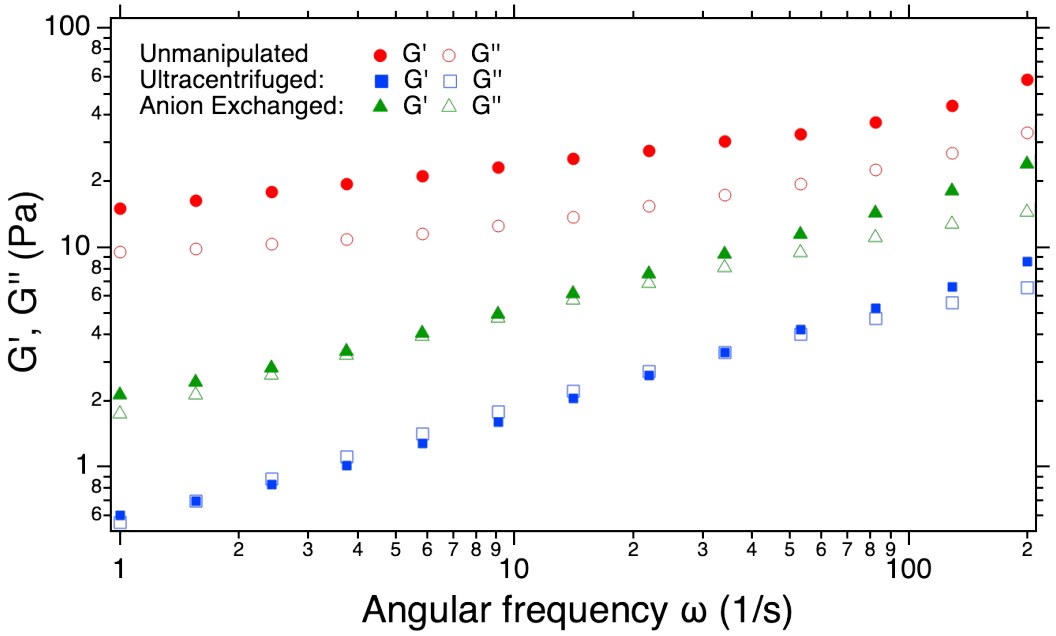


FIGURE 5.4: **Storage G' and loss G'' moduli for the different treatments of bacterial cell lysate**. G' and G'' are presented as a function of the angular frequency ω. The amplitude $\dot{\gamma}_A$ used is 5%.

Figure 5.5 shows the damping factor G''/G' at different lysate concentrations (50, 200, 250, 300 mg/ml) and different amplitudes (1, 5, 10, 30, 50%). From Figure 5.5, from top to bottom, the damping factor increases as the amplitude $\dot{\gamma}_A$ increases, i.e., it becomes more apparent liquid-like. Meanwhile, from left to right, the damping factor becomes smaller as the concentration increases, i.e., it becomes more apparent solid-like. This is true for all amplitudes and cell lysate concentrations.

From our rheological results, it is clear that bacterial cell lysate has a non-Newtonian behaviour, specially at the typical concentrations of cells (200-400 mg/ml). As a remark, from the oscillatory tests, it seems that at least unmanipulated cell lysate has a solid-like behaviour as the concentration increases. Consequently, Ficoll's Newtonian behaviour might not be the most suitable crowder to mimic real cellular environment.

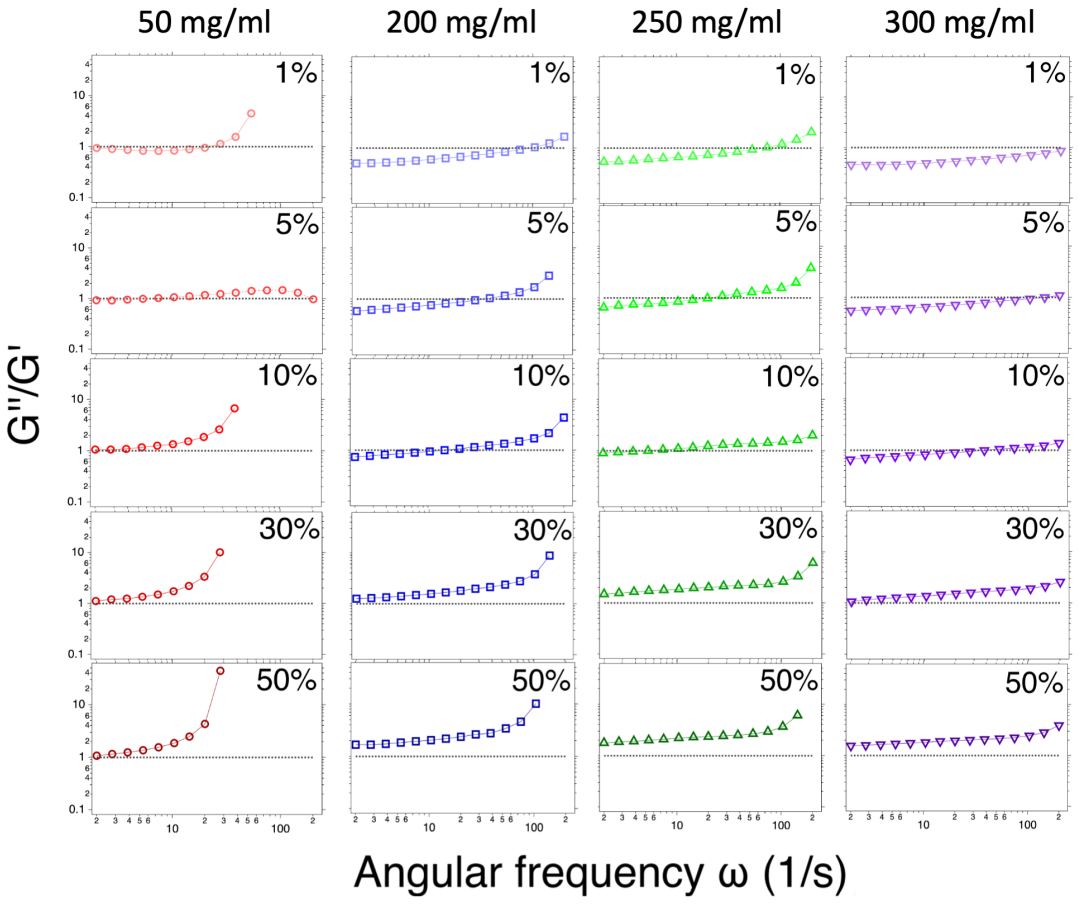


FIGURE 5.5: **Oscillatory test of unmanipulated cell lysate at different concentrations**. The horizontal black like at 1 is a reference to the eye for < 1 elastic-like behaviour, and > 1 viscous-like behaviour.

Chapter 6

Diffusion NMR study of a model polymer in bacterial cell lysate crowders

We measured the diffusivity of PEG in three different lysate treatments: unmanipulated, ultracentrifuged, and anion exchanged. We used the linear polymer PEG as a model intrinsically disordered protein. The idea of the three different lysate treatments was to explore into the non-specific interactions We used diffusion NMR to study the diffusivity of PEG in the lysate crowders. We compared our PEG diffusivity results with PEG in Ficoll as the crowder. Note: the PDI reported in the paper [1] was removed from table 6.1 in this manuscript because they were incorrect. They have no impact in the rest of the results.

6.1 Cell lysate characterization

We first present the ^{1}H (proton) NMR spectra of PEG in lysate crowders. The NMR signal intensity of the PEG peak with lysate in Figure 6.1(a) shows a much larger PEG peak at 3.6 *ppm* compared to lysate peaks found at similar frequencies. The high intensity of the PEG signal compared to the lysate signals was an advantage when measuring PEG translational diffusion using NMR. The difference between PEG and

the lysate signal intensities remained even at higher concentrations of lysate (spectra not shown).

We distinguish between three different lysate treatments. Unmanipulated lysate is sheared-open *E. coli* bacteria, but otherwise not manipulated. Ultracentrifuged lysate corresponds to the bacterial lysate after removing the biggest structures such as DNA and lipids. The third treatment, anion exchange, partially depleted the lysate of negatively charged macromolecules by anion exchange chromatography. Figures 6.1(b,c,d) correspond respectively to 1D proton NMR spectra of unmanipulated, ultracentrifuged and anion exchanged lysate without the probe molecule PEG. The spectrum of unmanipulated lysate (Figures 6.1(b)) exhibits significant broadening, while there is no substantial difference between the ultracentrifuged and anion exchanged treatments (Figures 6.1(c,d)), where the peaks are significantly narrower in comparison.

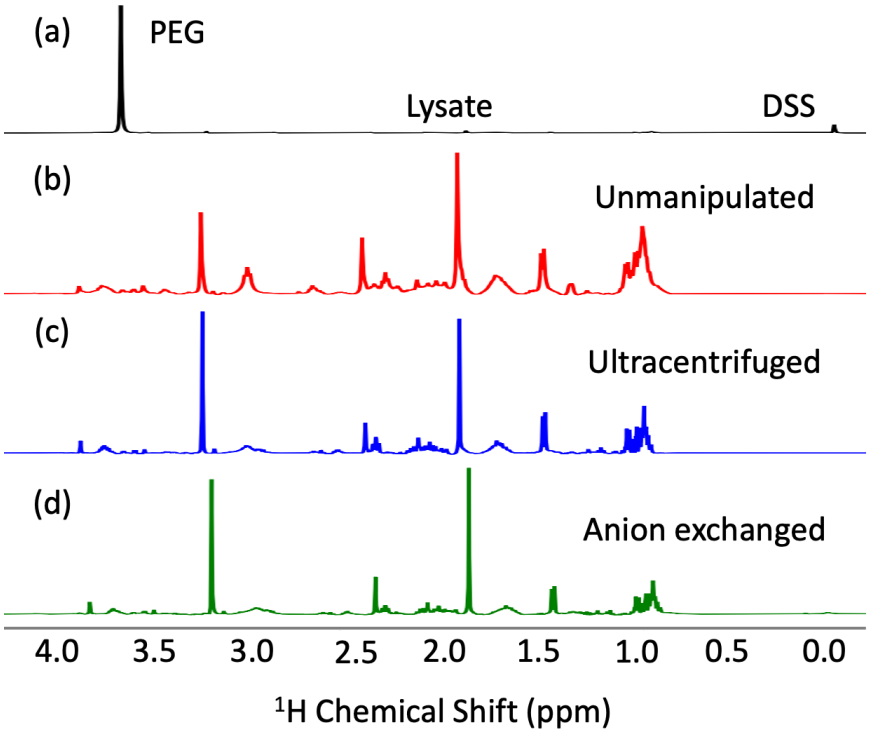


FIGURE 6.1: **^1H NMR spectra for different treatments of cell lysate at 100** mg/ml. NMR spectra measured with WATERGATE water suppression. The PEG peak is at 3.6 ppm and the main lysate peaks are between 4 and 0.5 ppm. (a) 30 mg/ml PEG mixed with unmanipulated lysate. The intensity of the PEG peak signal is distinguishably greater than all the peaks corresponding to the different treatments of lysate. The intensity signal decay to calculate the translational diffusion of PEG was obtained from its peak at 3.6 ppm. (b) Unmanipulated lysate without PEG. (c) Ultracentrifuged lysate without PEG. (d) Anion exchanged lysate without PEG. The intensities in the spectra b,c, and d are not scaled with the intensity in a. Sodium trimethylsilylpropanesulfonate (DSS) is used as chemical shift reference.

TABLE 6.1: (Top row) The mean hydrodynamic radius $\overline{R_{H.Int}}$ (weighted by intensity) at 1 mg/ml for each cell lysate treatment. (Middle row) The number weighted hydrodynamic radius $\overline{R_{H.num}}$. (Bottom row) The average ζ potential and the width of the ζ potential distribution DW, at 1 mg/ml for each cell lysate treatment. DW is equal to the full width half maximum of the distribution. The number-weighted values are significantly smaller than the intensity-weighted values, as seen in other particulate suspensions [118].

Unmanipulated		Ultracentrifuged		Anion Exchanged	
$\overline{R_{H.Int}}$ (nm)		$\overline{R_{H.Int}}$ (nm)		$\overline{R_{H.Int}}$ (nm)	
244 ± 14		34 ± 1		26 ± 1	
$\overline{R_{H.num}}$ (nm)		$\overline{R_{H.num}}$ (nm)		$\overline{R_{H.num}}$ (nm)	
36 ± 8		4 ± 1		4 ± 1	
ζ (mV)	DW (mV)	ζ (mV)	DW (mV)	ζ (mV)	DW (mV)
-21	31	-37	6	-39	6

To characterize the crowder, we measured diffusion coefficients of the lysate macromolecules in dilute solutions using dynamic light scattering (DLS) for each lysate treatment. The diffusion coefficient D in dilute solution is converted to a hydrodynamic radius, R_H, using the Stokes-Einstein equation

$$D = \frac{k_B T}{6\pi\eta R_H}, \tag{6.1}$$

where η is the solvent viscosity. These average hydrodynamic radii $\overline{R_H}$ are shown in Table 6.1 (top row). For unmanipulated lysate, the average hydrodynamic radius is found to be 244 ± 14 nm. $\overline{R_H}$ drastically decreases for ultracentrifuged (34 ± 1) nm and anion exchanged lysate (26 ± 1) nm when compared with the unmanipulated lysate. As a reference, the estimated radius for *E. coli* bacteria is 0.5 μm. This dramatic decrease is expected for both cell treatments given that the large components like lipid structures and DNA are removed from the lysate. Meanwhile, there is only a small difference between the $\overline{R_H}$ of the ultracentrifuged and the anion exchanged lysates. Anion exchange removed $\sim$ 27 % of the dried mass from the ultracentrifuged lysate, which is consistent with depletion of negatively charged species, although it should be noted that for the work presented below, all treatments of the lysate were adjusted to the same concentrations.

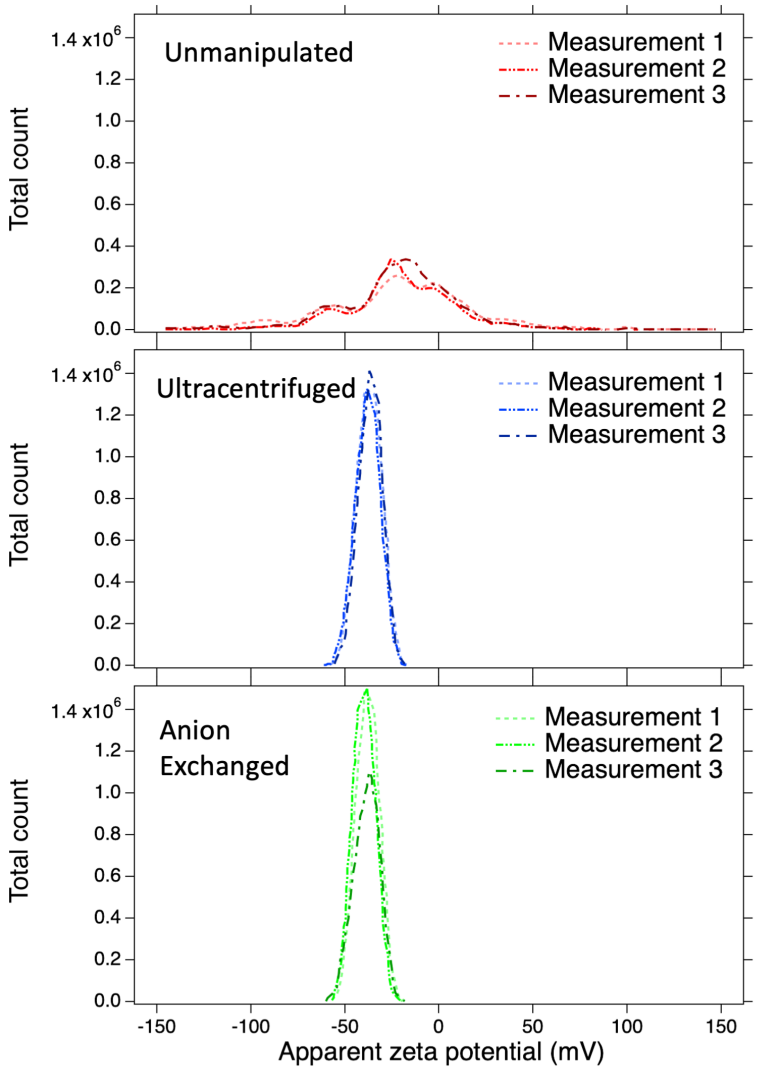


FIGURE 6.2: **Zeta potential distributions for all treatments of cell lysate at 1** mg/ml. Each value was measured three times. While the distributions for anion exchanged (bottom) and ultracentrifuged lysate (middle) are similar, and have a single peak near -40 mV, the distribution for unmanipulated lysate (top) is multimodal and broad.

We also performed the zeta potential experiments using a dip cell for all treatments of cell lysate at 1 mg/ml. It is seen in Figure 6.2, and summarized in Table 6.1 (bottom), that the width of the distribution (DW) for unmanipulated cell lysate is large, 31 mV, indeed exceeding the magnitude of the average value of ~ -20 mV. For both ultracentrifuged and anion exchanged lysates, the average $\zeta \sim -40$ mV. The distribution is unimodal in both cases, and the width is much smaller, 6 mV. The average charge for *E. coli* bacteria proteins is negative [119].Note that although the uncertainty in the average value of the zeta potential is small, the distribution of zeta potential values is much larger, hence this is more useful to report. Based on the zeta potential distributions in

Figure 6.2, ultracentrifuging the lysate seems to make the charge distribution narrower, but its value is not significantly altered by anion exchange.

6.2 Diffusion coefficient of PEG in cell lysate

The decay of the NMR signal intensity of PEG was used to measure the translational diffusion coefficient, which indicates how quickly PEG diffuses. To do so, the gradient strength g was increased from 50 G/cm to 1000 G/cm in a PFG NMR double stimulated echo sequence. The signal decay of PEG is presented as a function of the gradient strength parameter k, as shown in Figure 6.3. The gradient strength parameter $k = (\gamma\delta g)^2(\Delta - \delta/3)$ depends on the gyromagnetic ratio $\gamma = \gamma^H = 2.657 \times 10^8\ T^{-1}s^{-1}$, the duration of the field gradient pulse $\delta = 2\ ms$, and the time between two gradient pulses $\Delta = 100\ ms$. For a single macromolecule species one expects

$$S(g) = S(0)\exp(-kD) \tag{6.2}$$

and so a plot of $\ln(S(g)/S(0))$ vs k is linear with the slope yielding the diffusion coefficient [27]. S_0 is the initial signal without a gradient pulse. The concentration of PEG was fixed at 30 mg/ml for all experiments. Meanwhile, the concentrations for each of the three lysate crowder treatments were 50, 100, 150, 200, 250 and 300 mg/ml. This choice of concentration for the crowder was guided, in part, by the fact that the estimated total protein concentration in *E. coli* bacteria is $\sim 200 - 400\ mg/ml$ [5, 13, 120, 121]. The translational diffusion of PEG was also measured in the absence of crowders, i.e. PEG in water, as a reference.

In Figure 6.3, we see that, for PEG in water (green diamond symbols), $\ln(Signal)$ decreases linearly with k and can be fit with a single exponential. For PEG in lysate, the signal attenuation is not mono-exponential, and was thus fit with a bi-exponential. In this case, the bi-exponential equation represents two components for the diffusion of PEG. The first component represents the fraction that diffuses faster, and the second component is related to a slower diffusing component, as shown in the equation,

$$S(g)/S(0) = f\exp(-D_{\text{fast}}k) + (1-f)\exp(-D_{\text{slow}}k) \tag{6.3}$$

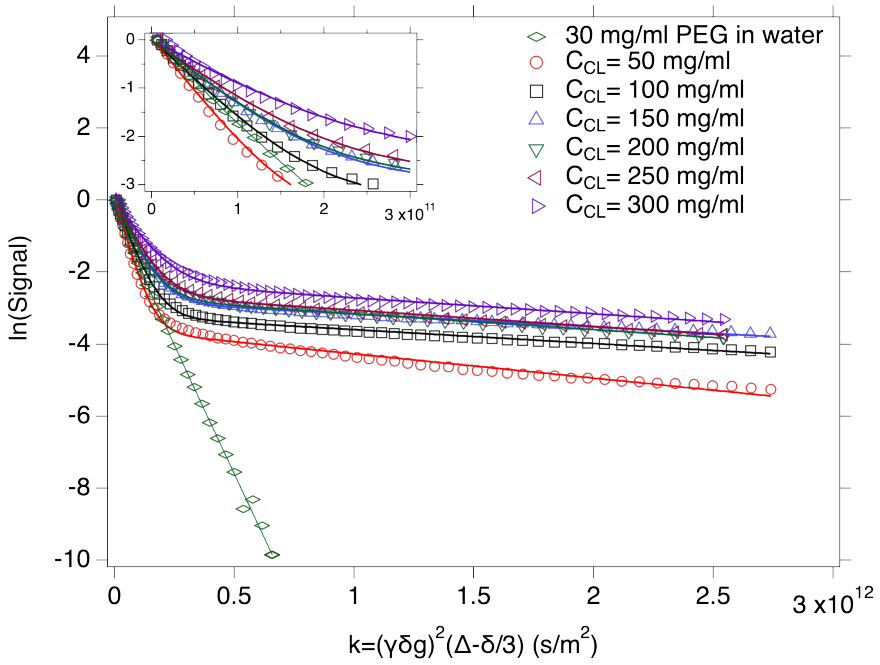


FIGURE 6.3: **PEG in unmanipulated cell lysate crowder has two diffusion coefficients**. The logarithm of the attenuated signal decay versus the gradient strength parameter (k) for 30 mg/ml PEG with different concentrations of unmanipulated cell lysate C_{CL}. Each signal intensity decay was fitted using a bi-exponential equation with a fast and slow component diffusion coefficient. The inset shows an expanded view of the curves at small k. The mono-exponential decay (linear dependence of $\ln(Signal)$ as a function of k) of PEG in water is shown for comparison.

where f is the fraction representing the fast component, D_{fast} is the diffusion coefficient of the fast component, and D_{slow} is the diffusion coefficient of the slow component. The fast component is expected to be PEG diffusing as an isolated chain. On the other hand, the slow component might be due to the aggregation of either PEG with itself or PEG with components of the lysate crowder. To support our interpretation, we added Figure C.3 and C.4 in our Appendix C. Figure C.3 corresponds to a smaller PEG concentration than the used for all our experiments. From Figure C.3, we can notice that 3 mg/ml PEG in distilled water diffusivity can be calculated from a mono-exponential. However, this is not true for 30 mg/ml PEG in distilled water, where there are (minimally) two diffusivity components. Meaning, that the slow component could exclusively correspond to PEG aggregating with itself.

Figure 6.4 shows a summary of all the fast component diffusion coefficients for PEG at different concentrations of lysate. All our PEG diffusivity values reported in Figure 6.4 are the result from an average over at least 4 repeated experiments with freshly prepared samples. These 4 measurements were indeed taken at sequential time points, and no

systematic variation in diffusivity was observed. The error bars represent the standard error. In order to address reproducibility and variability of the diffusion coefficient results in different batches of cell lysate, all results in Figure 6.4 are presented as averages of measurements from 4 freshly prepared samples and the error bars represent the standard error.

For unmanipulated cell lysate concentrations of 200 mg/ml and higher, there was some sample separation (the top part of the sample was more transparent compared to the bottom) observed when the NMR tube was removed at the end of the NMR experiment, which could contribute to the large error bars on some of the data. The non-mono-exponential signal attenuation of PEG, both in water and in lysates indicates PEG-containing aggregates that are larger than the monomeric PEG. Thus, it is quite feasible that the observed sample separation might also be due to the interaction of PEG with the component of the lysate, resulting in protein precipitation [33].

At the lowest lysate concentration (25 mg/ml), we compare the unmanipulated cell lysate (red bar) with Ficoll crowders (purple bars). At this concentration, the PEG signal attenuation is mono-exponential in both cases, with D_{PEG} being faster in the cell lysate compared to in Ficoll. At a cell lysate concentration of 50 mg/ml, the average fast diffusion value of PEG in unmanipulated lysate (red bars) seems to be larger than in the presence of anion exchange treated lysate. However, for the concentrations of 100, 150, 200, 250 and 300 mg/ml, it seems that the removal of the larger macromolecules in the crowder does not affect the diffusion coefficient of PEG significantly.

For PEG in unmanipulated cell lysate the diffusion coefficient (fast component) decreases from 24×10^{-12} to 17×10^{-12} m^2/s between 25 and 150 mg/ml. The fast component of the diffusion coefficient of PEG decreases roughly 4-fold at the maximum lysate concentration. On the other hand, the diffusion coefficient for PEG in the artificial spherical crowder Ficoll decreases nearly twice as rapidly as the diffusion coefficient in cell lysate to 6×10^{-12} m^2/s as the concentration approaches 150 mg/ml.

The purple asterisks in Figure 6.4 represent the Ficoll concentrations where it was not possible to measure a diffusion coefficient due to phase separation. Solutions with 30 mg/ml PEG were stable up to 150 mg/ml Ficoll, while in the unmanipulated cell lysate, the solution stability began to be affected at 200 mg/ml and above. These are indicated by red asterisks. Since the slowest-diffusing aggregates of molecules are likely not seen

in our NMR spectra, the observed diffusivity in these cases may be considered to be an upper bound.

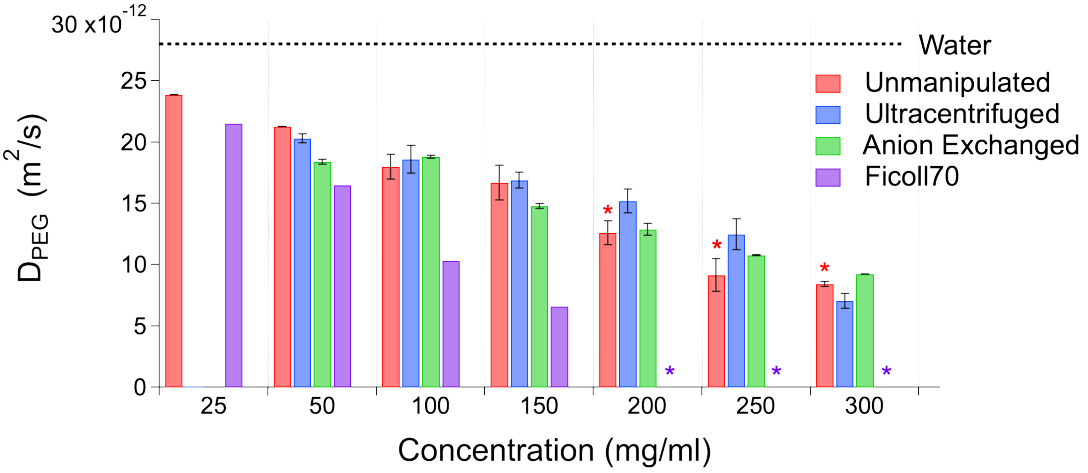


FIGURE 6.4: **Self-diffusion of PEG in cell lysate: the fast component**. The average over 4 replicates for 30 mg/ml PEG in different concentrations and treatments of cell lysate (unmanipulated, ultracentrifuged, and anion exchanged). PEG in water and Ficoll70 data was obtained from [4]. The error bars come from the standard error of multiple measurements. Asterisks indicate samples for which phase separation was observed.

Figure 6.5 provides information about the slow component obtained from the same diffusion measurements. Figure 6.5(a) shows the average fractions f of the sample that is diffusing with the faster diffusion coefficient; $(1 - f)$ is of course the fraction of the slow component. The fraction f associated with the fast component of PEG diffusion coefficient is ≈ 1 at low concentrations and decreases as a function of increasing lysate crowder concentration for unmanipulated cell lysate (red triangles in Figure 6.5(a)). This suggests that as concentration increases, PEG aggregates more, either with itself or with lysate macromolecules. The same conclusion cannot be drawn from the ultracentrifuged or anion exchanged lysate as f does not change with lysate concentration. For the ultracentrifuged and anion exchanged lysates, we see in Figure 6.5(a) that there is effectively only one component. Figure 6.5(b) shows the diffusion coefficient, D_{slow}, of the slow component for the unmanipulated cell lysate. Signal attenuations and D_{slow} for ultracentrifuged and anion exchanged are shown in Appendix C. The mean value for PEG diffusivity in lysate at 100 mg/ml and above is 0.4×10^{-12} m^2/s, about 30-50 times lower than $D_{PEG,fast}$. If we associate the slow component with clusters of PEG, this implies rather large clusters that are 30-50 times larger than the PEG molecule (which has a radius of gyration $R_g \sim 5$ nm). This estimate assumes that the Stokes-Einstein

relation is valid, which is only expected for dilute solutions. However, the lack of a significant concentration dependence of D_{slow} suggests this is a reasonable estimate.

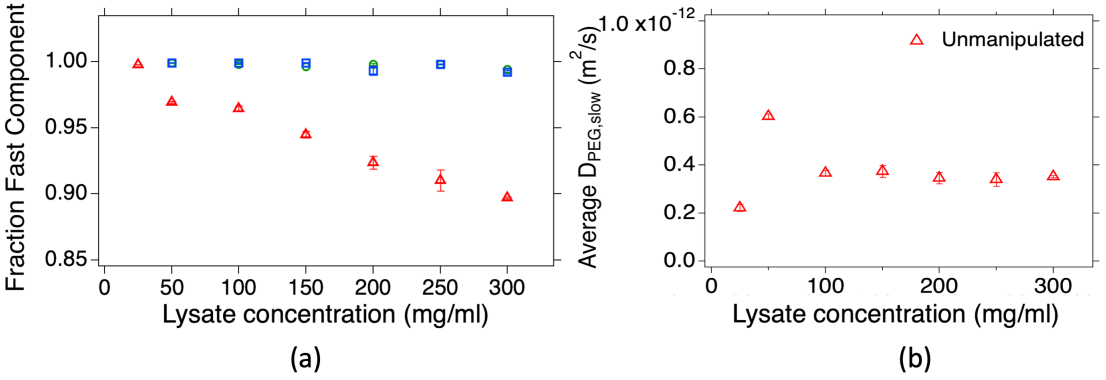


FIGURE 6.5: **Self-diffusion of PEG in cell lysate: fractions and the slow component**. (a) Average fraction corresponding to the fast diffusion coefficient component as a function of the lysate concentration. The fraction for the unmanipulated lysate shows that increasing the concentration of the crowder decreases the fraction of PEG that diffuses quickly. (b) Average diffusion coefficient of the slow component of PEG. (Blue squares) Fraction fast component for ultracentrifuged lysate. (Green circles) Fraction fast component for anion exchanged lysate.

Chapter 7

Diffusion NMR study of an intrinsically disordered protein in bacterial cell lysate crowders

We measured the diffusivity of the intrinsically disordered protein αS in different cell lysate crowder treatments at different concentrations. We used and HSQC-DOSY NMR experiment to calculate the diffusivity of ^{15}N labelled αS in cell lysate. The idea of the experiments presented in this chapter is to have an insight on the non-specific interactions (changes in size and charge of the crowder) effects on the IDP αS.

7.1 Spectra of α-synuclein

In the Materials and Methods chapter (section 4.2), we described the methods for the purification of ^{15}N-labelled αS. Next, we needed to make sure we have the right protein. To do so, we performed a SDS-PAGE (Figure 7.1) experiment to confirm we have an approximately 14 kDa protein. For our experiments, we used the pool corresponding to protein after dialysis (rightmost vertical band in gel). We compared the protein after dialysis band with the protein marker (leftmost vertical band in gel), on which we observed that effectively we have a protein around 15 kDa, considering that the molecular weight accuracy in a gel is around 10%. The middle bands (fusion protein +EZCUT and protein before dialysis) are to compare with previous steps in the process

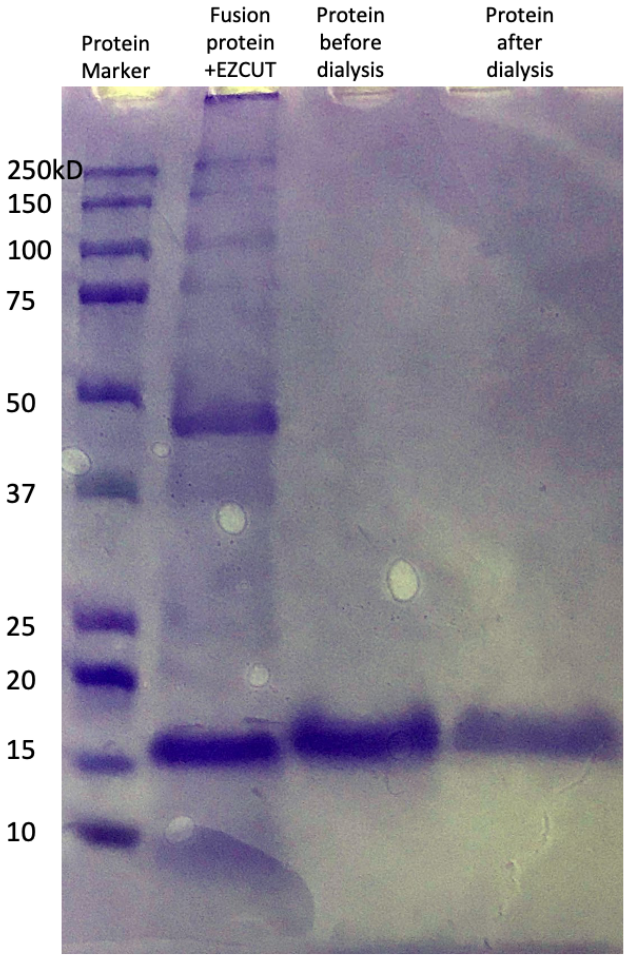


FIGURE 7.1: **SDS-PAGE gel for ^{15}N-labelled αS**. On the leftmost vertical band are represented the molecular weight for the protein marker. The rightmost vertical band corresponds to αS after dialysis. The two middle columns are a reference during the steps of the protein purification process.

of protein purification. Since we confirmed our protein size, the next step was to use NMR to measure the spectrum.

Our probe molecule is the intrinsically disordered protein, αS, whose concentration is (2.08 ± 0.31) mg/ml (0.1 mM). The final concentration of αS was calculated from the average between two measurements. First, the concentration was estimated using the Lambert-Beer equation ($A = \epsilon cl$), where ϵ is molar extinction coefficient and a $l = 1$ cm is the path length. For αS, $\epsilon = 5960$ $M^{-1}cm^{-1}$. The absorbance (A) at 280 nm was measured using a quartz 10.00 mm cuvette in a Genesys 10S UV-vis spectrophotometer. The concentration was also estimated with a second measurement and the Lambert-Beer equation, with the same values for ϵ and l. In the second measurement, the absorbance

(A) at 280 nm was measured with the ThermoScientific Nanodrop 2000. The reported concentration for αS used in this manuscript is the average from these two measurements.

We performed all experiments in the NMR 500 MHz at 25 $^{\circ}C$ and pH 7.0, unless stated otherwise. We distinguish between three different lysate treatments: unmanipulated, ultracentrifuged, and anion exchanged.

We first present the ^{1}H (proton) NMR spectra of ^{15}N-labelled intrinsically disordered protein αS in pretreated lysate *E. coli* bacteria (Figure 7.2). The peak at 4.7 *ppm* corresponds to water after water suppression by WATERGATE. The spectra in Figure 7.2 is for αS in unmanipulated cell lysate (top), ultracentrifuged (middle), and anion exchanged (bottom). αS in the lysate treatments was measured right after mixing both the protein and the lysate crowder, as well as 20 hours after mixing. From the spectra we can noticed that in general the peaks do not shift or appear new peaks. We do not show the spectra for ^{15}N αS without crowder because it is easier to do comparison with the protein in crowder using HSQC spectra.

Our region of interest is between 10 *ppm* and 6 *ppm*, since it is used for ^{15}N diffusion experiments. This region is shown in Figure 7.3. From the spectra, we can notice that there are no significant changes between αS in ultracentrifuged and anion exchanged cell lysate. However, the peaks seems broad when the protein is mixed with unmanipulated lysate. From our lysate characterization in chapter 6, we learned that there are no significant differences between ultracentrifuged and anion exchanged lysate, so this is consistent with our previous observations.

^{15}N-^{1}H HSQC spectra of ^{15}N αS are shown in Figure 7.4. Figure 7.4(a) corresponds to ^{15}N αS in aqueous solution (dialysed TEV washing buffer described in chapter 4 with 10% D_2O). We compared our HSQC spectrum with the one reported by Alderson and Markley [122] for αS. The clustered peaks around 8 in the ^{1}H channel correspond to a disordered protein, a folded protein has spread-out peaks. The two peaks between 6.8 *ppm* and 7.8 *ppm* (^{1}H channel) and 112 (^{15}N channel) correspond to the residue side chain [122]. Figures 7.4(b,c,d) show ^{15}N αS in 100 mg/ml lysate treatments (unmnanipulated (b), ultracentrifuged (c), and anion exchanged lysate (d)). We notice there are no changes in big peaks, new peaks do not appear or disappear, and there is no *ppm* shift in any of the major peaks. The experiments were not repeated with another newly prepared protein, since the process of protein purification took a long time. Only the

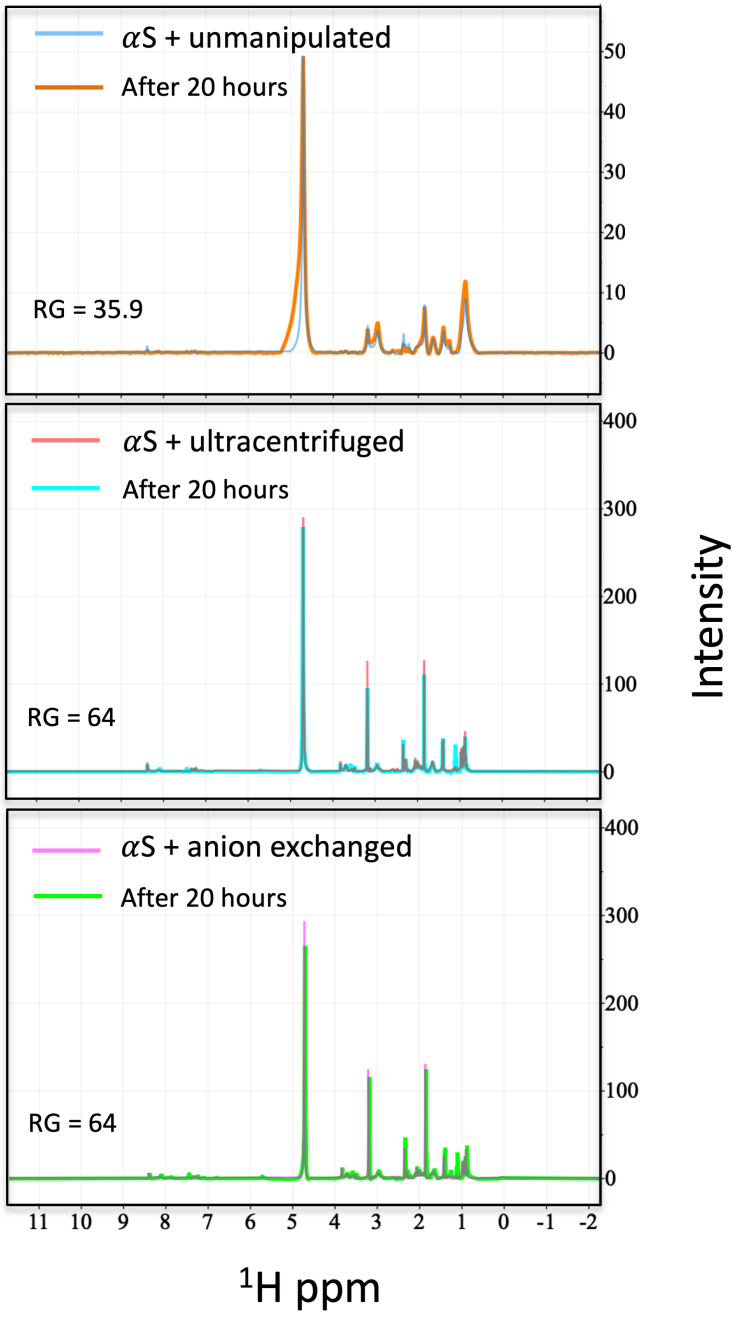


FIGURE 7.2: ^{1}H **NMR spectra of** ^{15}**N** α**S in cell lysate freshly mixed and 20 hours after mixing at 25** $^{\circ}C$. NMR spectra measured with WATERGATE water suppression with a TXI probe and 500 MHz. The concentration of cell lysate crowder is 100 mg/ml. αS concentration is (2.08 ± 0.31) mg/ml (0.1 mM) and pH 7.0. RG is the receiver gain used in the experiment.

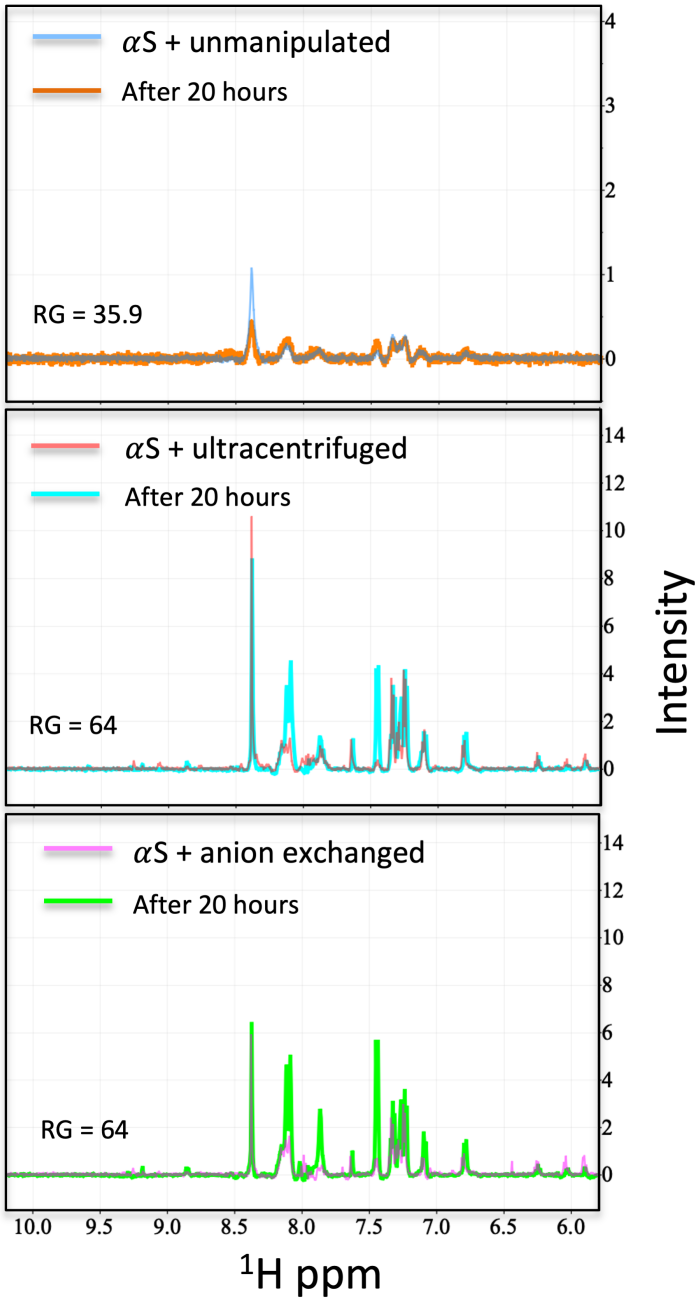


FIGURE 7.3: **Region 10-6** *ppm* **in** ^{1}H **NMR spectra of** ^{15}**N** α**S at 25** $^{\circ}C$. αS is presented in different treatments of cell lysate freshly mixed and 20 hours after mixing. Region view is maximized from the spectra in Figure 7.2

peaks with ^{15}N-^{1}H coupling are visible in the spectra. An interesting observation for αS in **E. coli** lysate crowder is that the protein seems to preserve the disordered structure. This result is in agreement with McNulty and coworkers [123], where they found αS keeps the unfolded structure in the presence of **E. coli** crowder. McNulty'experiments were performed at 35 $^{\circ}C$ with ^{15}N-labelled αS within *in vivo* **E. coli**.

Since our diffusion experiments have a duration of approximately 10 hours, we decided to look for possible changes in the structure due to age. To do so, we compared HSQC spectra for our protein αS in lysate crowders before and after 20 hours from preparing the sample. Figure 7.5 compares the HSQC spectra for freshly mixed protein-crowder (red) and the spectra obtained after 20 hours of mixing (green) lysate. In figure 7.5, we compared αS spectra mixed with unmanipulated (a), ultracentrifuged (b), and ultracentrifuged with anion exchanged (c) lysate at 100 mg/ml. In contrast with the 1D spectra from Figure 7.3, where one mostly observes the lysate peaks, there are no significant changes on the HSQC spectra after 20 hours of experiments.

From the HSQC experiments, it seems there are not significant spectral changes or major *ppm* shifts in the protein peaks with freshly mixed αS with lysate and after 20 hours from mixing. This observation is useful considering that the HSQC-DOSY experimentation have an approximate duration of 10 hours per sample.

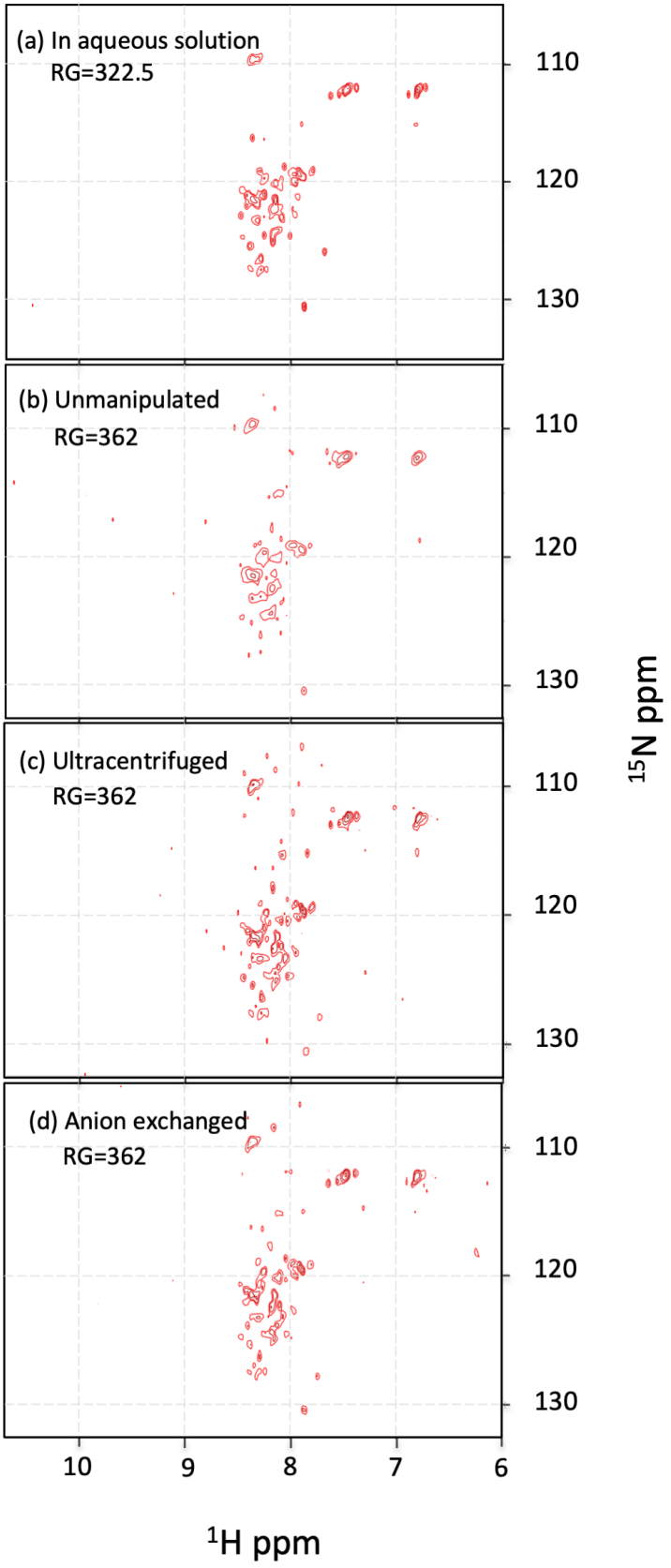


FIGURE 7.4: **^{15}N-^{1}H HSQC spectra of αS** at 25 $°C$ in (a) aqueous solution (no crowder, this data was measured in 500 MHz and the experiment was prepared and run by Sina Heravi), (b) unmanipulated, (c) ultracentrifuged, and (d) anion exchanged cell lysate. All lysate concentration is 100 mg/ml. The samples with crowders were measured in the 500 MHz NMR around 30 minutes after it was mixed.

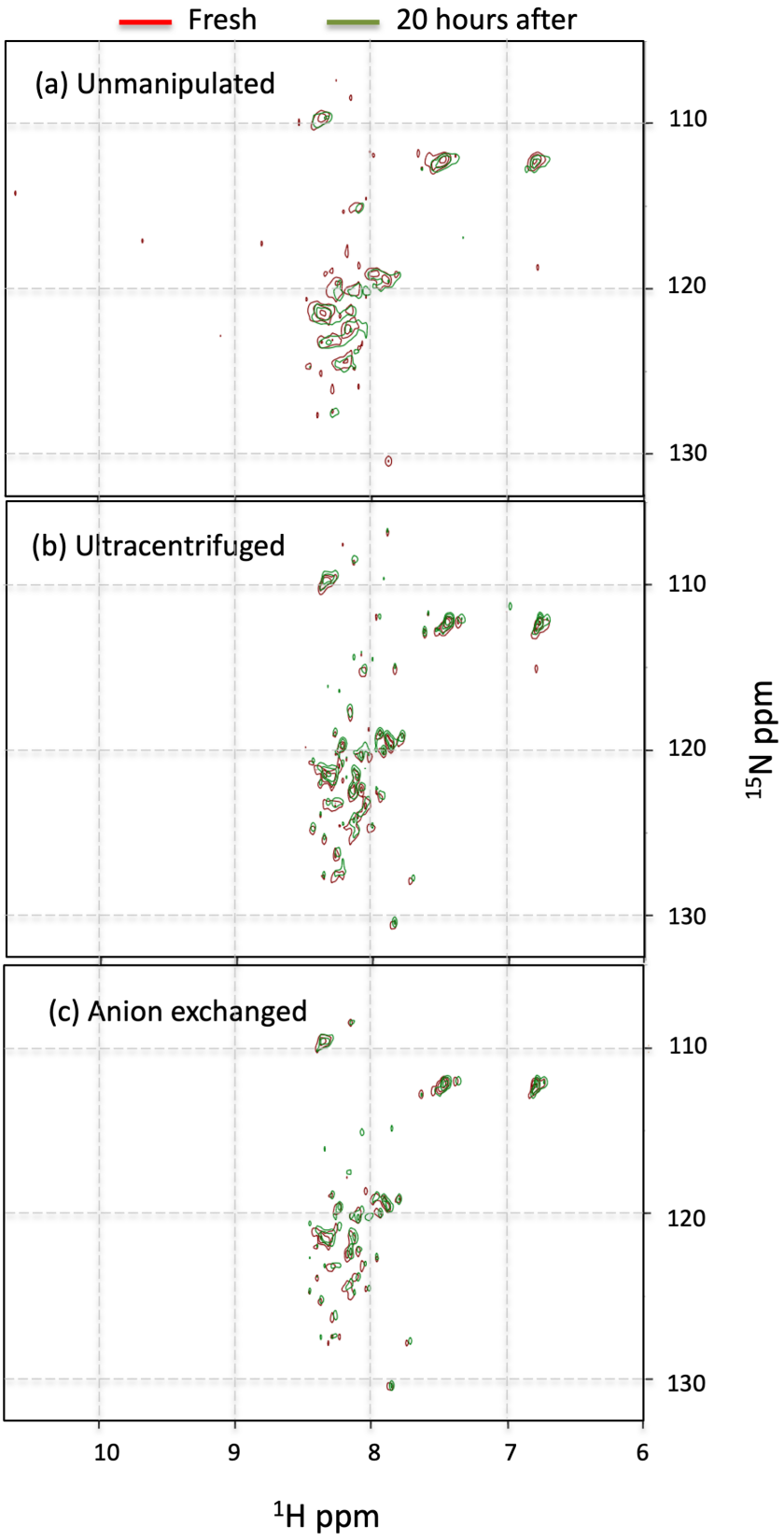


FIGURE 7.5: **^{15}N-^{1}H HSQC spectra of αS in cell lysate freshly mixed and 20 hours after**. (a) unmanipulated, (b) ultracentrifuged, and (c) anion exchanged cell lysate. All lysate concentration is 100 mg/ml.

7.2 Diffusion of α-synuclein in cell lysate

Using SDS-PAGE and NMR, we were able to check and compared with the literature that we effectively have the intrinsically disordered protein αS. Our next step was to determine a region of integration withing 10-6 *ppm*. All self-diffusivity measurements of ^{15}N αS were done in NMR 500 MHz using a HSQC-DOSY experiment (Figure 3.11 described in chapter 3).

Figure 7.6 shows 10 1D spectra obtained from a quasi-2D HSQC-DOSY experiment, where each spectrum is obtained at different pulsed field gradient strengths (increasing from bottom 1% to top 90%). Figure 7.6(a) corresponds to αS in aqueous solution, and Figure 7.6(b) is αS in 100 mg/ml unmanipulared cell lysate. From Figure 7.6, we notice that the decay is faster for αS in aqueous solution (a) compared to αS in 100 mg/ml unmanipulated lysate (b).

The peaks between 9.5 and 8.9 *ppm* correspond to ^{15}N GB-1 protein from the fusion protein. Since the GB-1 region is not shown in our HSQC spectra (Figure 7.4(a)), we do not consider any peaks beyond 8.9 *ppm*. In the presence of crowders, the GB-1 peak is not seen, which might be because it binds with the lysate crowder and therefore is invisible to the instrument. The peaks between 8.9 and 7.1 *ppm* most likely correspond to the aromatic amino acids (phenylalanine, tryptophan, and tyrosine), for both samples with and without crowders. Considering our HSQC spectra for ^{15}N αS, our region of integration should be between 8.7 and 7 *ppm*. However, the region of integration changes slightly when the lysate is mixed with our protein. From Figure 7.4(b), the region of integration should be between 6.5 *ppm* to 9 *ppm*. The reason is since we are integrating by area, we need to choose two extremes with well defined baseline.

With the region of integration determined, we needed to measure the diffusivity of a well known molecule, formamide, to calibrate our measurements. ^{15}N formamide was suspended in dimethyl sulfoxide (DMSO-d6). To compare the effectiveness of the HSQC-DOSY pulse program, we measured as well the diffusivity of our molecule using DOSY experiment described in chapter 3.1 Figure 3.8. An important remark is that the shape of the gradient for the DOSY experiment is sinusoidal (SINE.100), which means that the maximum gradient 100% corresponds to 33.71 G/cm. However, the gradient in the HSQC-DOSY pulse programs has a smooth square shape, which makes

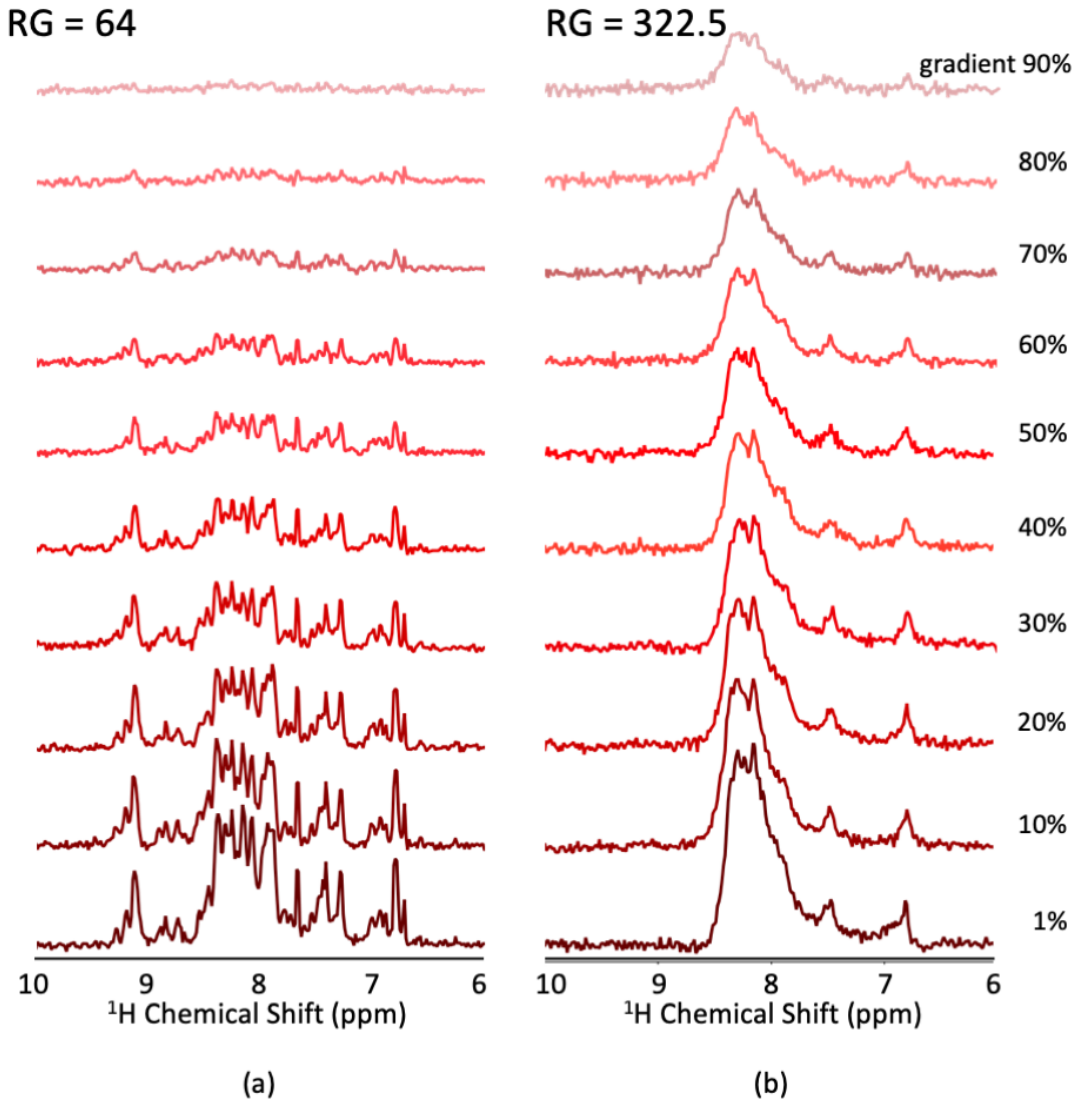


FIGURE 7.6: **Signal decay for ^{15}N αS**, (a) in aqueous solution, and (b) at 100 mg/ml of unmanipulated lysate. αS in aqueous solution data was measured in 500 MHz and the experiment was prepared and run by Sina Heravi.

the maximum gradient 100% equal to 48.15 G/cm. The shape of the pulses must be taken into consideration when calculating k values.

Table 7.1 shows the results from our measurements for formamide. We measured the diffusivity with both, 500 and 600 MHz NMR. Our results were mostly consistent between all measurements.

Our next step was to measure diffusivity of αS without crowder and with lysate crowders. Figure 7.7 shows the signal attenuation of ^{15}N αS between the region 9 ppm and 6.5 ppm in unmanipulated (a), ultracentrifuged (b), and anion exchanged lysate (c) as a function of the gradient strength parameter k. In each graph, the signal attenuation

TABLE 7.1: **Summarized diffusion values using a mono-exponential fit for the signal decays of formamide**. Experiment were prepared and run in collaboration with Sina Heravi. Error bars represent standard deviation in the fit.

	Region $[ppm]$/ NMR Field	DOSY	HSQC-DOSY
Formamide	7.7-6.8/ 500 MHz	$(5.99 \pm 0.04)10^{-10}$	$(6.15 \pm 0.06)10^{-10}$
	600 MHz	$(5.13 \pm 0.01)10^{-10}$	$(4.91 \pm 0.28)10^{-10}$

TABLE 7.2: **Summarized diffusion values of ^{15}N αS in cell lysate crowders**. Each D value was obtained from the Stejskal-Tanner equation. The experiments were performed in 500 MHz NMR. The row on the left corresponds to the lysate crowder concentrations C_{CL}. Error bars represent standard deviation in the fit.

C_{CL} mg/ml	Unmanipulated	Ultracentrifuged	Anion exchanged
	$D_{\alpha S.unm} \times 10^{-10}$ m^2/s	$D_{\alpha S.ult} \times 10^{-10}$ m^2/s	$D_{\alpha S.ani} \times 10^{-10}$ m^2/s
0	1.67 ± 0.04	1.67 ± 0.04	1.67 ± 0.04
50	1.06 ± 0.03	1.49 ± 0.03	1.13 ± 0.04
75	0.89 ± 0.05	1.12 ± 0.04	1.06 ± 0.06
100	0.82 ± 0.03	0.82 ± 0.03	0.89 ± 0.05

for αS in aqueous solution is shown for comparison. The decays were fitted using the Stejskal-Tanner equation $\ln(S(k)) = \ln(S(0)) - Dk$ (Equation 3.12, mono-exponential), where the slope corresponds to the self-diffusivity D. From the fitting for ^{15}N αS in no crowder, we obtained a diffusivity value $(1.67 \pm 0.04) \times 10^{-10}$ m^2/s. Using Stokes-Einstein equation 2.5, with $\eta = 0.89$ $Pa \cdot s$ for water at 25 $°C$, we can estimate that the hydrodynamic radius of αS is $(1.47 \pm 0.06)nm$. The hydrodynamic radius of αS has been reported to be approximetly 4 nm by Li, Uversky and Fink [124].

The resulting slopes from the mono-exponential fitting are represented in the bar graphs in Figure 7.8. As a first observation, there is a steady decrease in diffusivity when the concentration of the lysate crowder increases. Table 7.2 lists all the diffusion values for ^{15}N αS in lysate crowders. In table 7.2, the % uncertainties are roughly 3 or 6%. However, only a single replicate has been done, and it is quite possible that sample-to-sample variations are much bigger than the intrinsic fitting errors reported in the table.

We now compare our results for the intrinsically disordered protein αS with the non-ionic polymer (PEG) studied in chapter 6. Table 7.3 has summarized all the relative

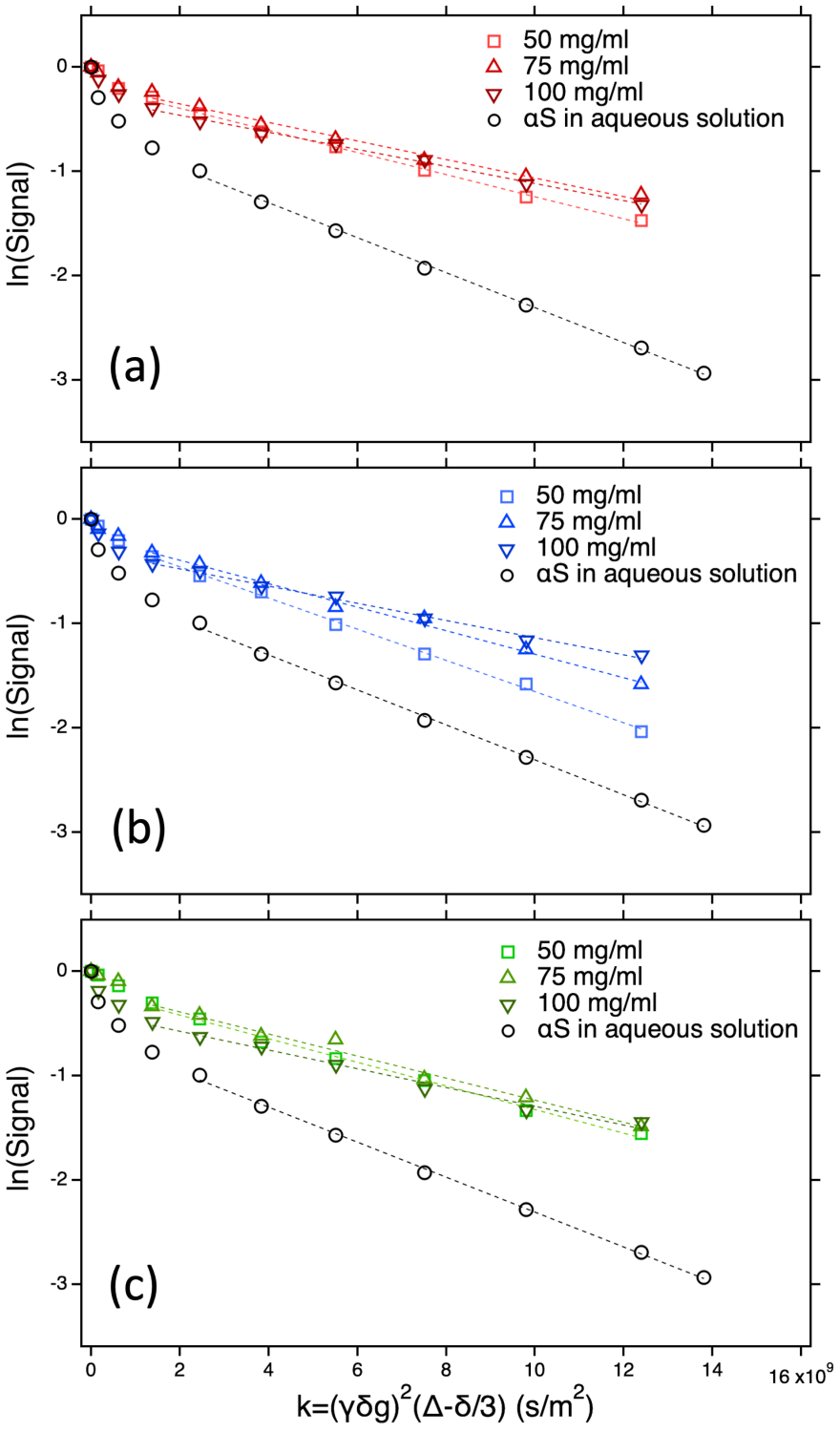


FIGURE 7.7: **Signal attenuations of ^{15}N αS in lysate crowder using 500 MHz at 25 °C.** Logarithmic scale of the attenuated signal decay versus the gradient strength parameter (k) for alphaS with different concentrations of (a) unmanipulated, (b) ultracentrifuged, and (c) anion exchanged cell lysate. The curve fits were done using the Stejskal-Tanner equation 3.12. αS in aqueous solution data was measured in 500 MHz and the experiment was prepared and run by Sina Heravi.

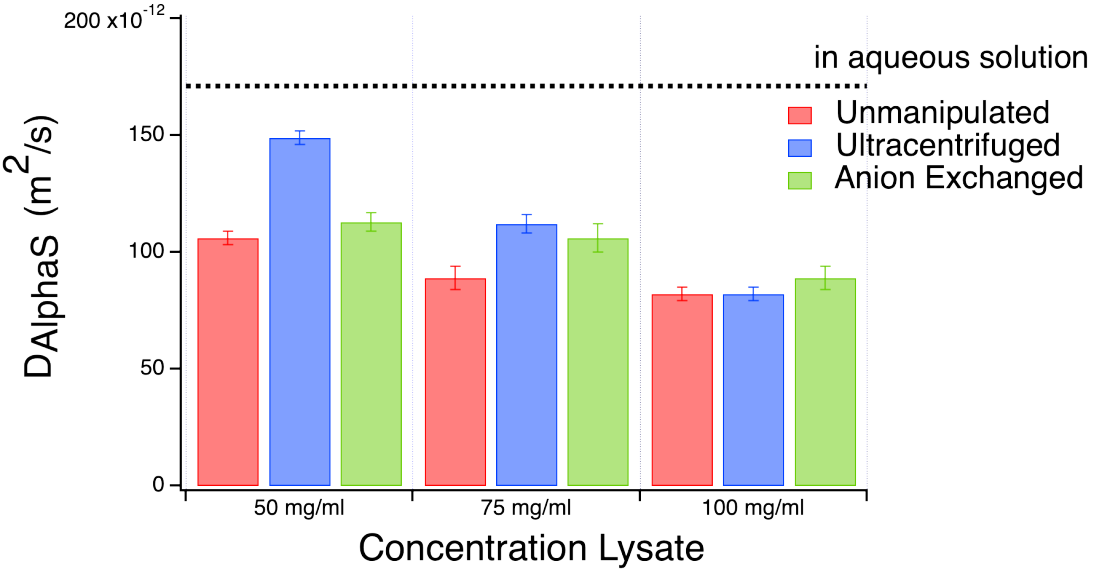


FIGURE 7.8: **Self-diffusion coefficient of αS in cell lysate at 25 °C.** For each ^{15}N αS, we used 2560 scans (NS in NMR) to improve the resolution of the signal. This data is listed in table 7.2. Error bars represent standard deviation in the fit.

TABLE 7.3: **Summarized relative diffusion values of ^{15}N αS and PEG in cell lysate crowders.**

Relative diffusion	50 mg/ml	100 mg/ml
Unmanipulated (αS)	0.635	0.491
Unmanipulated (PEG)	0.853	0.386
Ultracentrifuged (αS)	0.892	0.491
Ultracentrifuged (PEG)	0.726	0.664
Anion exchanged (αS)	0.677	0.533
Anion exchanged (PEG)	0.657	0.671

diffusion values for our protein and PEG at the concentrations of 50 and 100 mg/ml.

Figure 7.9 shows the relative diffusivity for αS and PEG. The PEG diffusivity are taken from chapter 6. For each test molecule (αS and PEG) in the different lysate treatments, there is a steady decrease. However, whereas in PEG the diffusion in ultracentrifuged and anion exchanged lysates looks very similar, this is not the case in αS. As a remark, we only have a single replicate for our disordered protein studies, so this unusual result indicates that multiple replicates might be required.

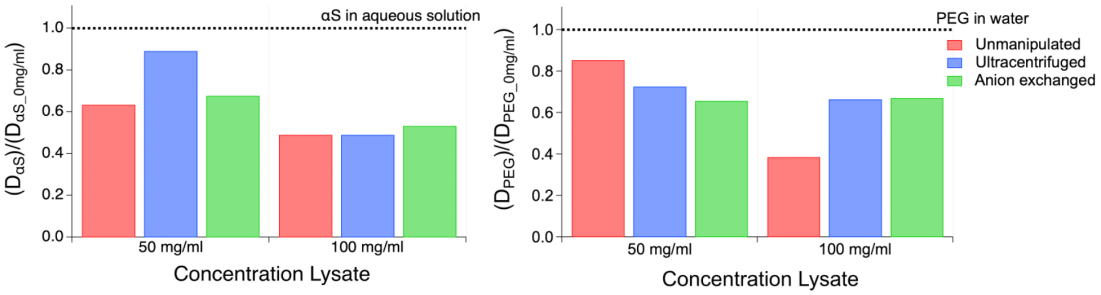


FIGURE 7.9: **Relative diffusion coefficient of αS and PEG in cell lysate at 25 °C at two concentrations**. Left graph corresponds to the relative diffusion of αS in lysate crowders. Right graph corresponds to the relative diffusion of PEG in lysate crowders.

Chapter 8

Conclusions and future directions

A controllable way to mimic a realistic cellular environment is using bacteria cell lysate. In this thesis, we have examined the nature of the realistic cellular environment (with rheology) and molecules (polymers and proteins) being transported by diffusion in cell lysate. Insights into macromolecular motions in realistic cell-like environments is very relevant for the effective delivery of drug molecules inside the body.

We measured, via NMR, the translational self-diffusion coefficient of PEG in different treatments of cell lysate: unmanipulated; ultracentrifuged; and anion exchanged (after ultracentrifugation). We also measured the rheological properties of the cell lysate, focusing on the concentration dependence of the unmanipulated cell lysate. First, it should be noted that there was no noticeable quantitative difference between the surface charge characteristics of ultracentrifuged and ultracentrifuged-plus-anion-exchanged pretreated cell lysates. Electrophoretic mobility measurements yield the same value of average (negative) charge with zeta potential ζ close to $-40\ mV$, and a rather similar (roughly unimodal) distribution of ζ. The unmanipulated cell lysate, on the other hand has an average negative charge with $\zeta = -20\ mV$, but with a much wider distribution of ζ values. Surprisingly, we found that the zeta potential results were more negative for ultracentrifuged compared to unmanipulated lysate. The reason for this is unclear.

As a contribution to the field of macromolecular crowding, this is the first research in literature to examine the non-Newtonian characteristics of cell lysate as a crowder at different concentrations and compare it to the artificial Newtonian crowder Ficoll. Furthermore, the diffusivity of PEG in the different treatments of cell lysate gives an

insight into the contribution of the non-specific interactions when compared with PEG's diffusivity in Ficoll. Moreover, the different treatments of cell lysate seem to have no effect in the diffusivity of αS, however, repetitions on the experiments and proper comparison with αS in Ficoll must be done.

8.1 Rheological characteristics of bacterial cell lysate

Unmanipulated bacterial cell lysate crowder is distinctly shear-thinning in the crowding limit. We used the Herschel-Bulkley model to quantify the degree of shear-thinning for the unmanipulated lysate as we increase the lysate concentration. Unmanipulated cell lysate showed a decrease of the HB exponent n from 1 to 0, as the concentration of lysate increases. The HB exponent is 1 for a Newtonian solution like Ficoll. A possible explanation for this result would relate to the nature of cell lysate and Ficoll, i.e. macromolecules in cell lysate are prone to entanglement as opposed to the artificial compact Ficoll molecules, whose structure has been recently studied in depth by Ranganathan *et al* [125]. Non-Newtonian cytoplasm of intact cells has been previously inferred via intracellular measurements (Puchkov *et al.* and Śmigiel *et al*) where diffusion has been found to be location dependent [126, 127]. In addition, our rheological data shows that a higher "yield" stress is required to induce flow as the concentration of lysate increases.

So far, we have examined the non-Newtonian rheology of cell lysate, and the quiescent diffusivity of a test macromolecule in cell lysate. However, this study suggests an intriguing possibility: the shear-thinning nature of cell lysate implies that active swimmers experiencing a high local shear rate would experience a 'thinner' medium than Brownian diffusers [128–130]. This will be a subject of future study. It is, however, interesting that the non-associating fraction (the "fast" component) of the PEG is not affected much by the concentration-dependent increase in this non-Newtonian character. As a note, PEG and αS diffusivity values were not affected by the batch-to-batch variation of lysate treatments.

As a point, PEG has been used in many studies [52, 131, 132] as an inert polymeric crowder. However, the strong association of PEG with components of cell lysate indicates that PEG is not an appropriate excluded-volume crowder and should be strictly avoided in use as a crowding agent. This is also in strong agreement with the finding of Crowley *et*

al [133]. But more importantly, most artificial excluded-volume crowders miss multiple important characteristics that biologically relevant crowders embody: a wide distribution of macromolecular size *and* charge, and a non-Newtonian rheological response that is consistent with some form of network-like structure much larger than the size of proteins.

As a remark, the rheological results presented in this thesis are the first in literature where the role of the non-Newtonian character of the crowding environment is examined. This is important because it is likely relevant to crowding living.

8.2 Diffusion coefficient of PEG in cell lysate

Our NMR data shows two diffusion components for PEG in unmanipulated cell lysate. The fast diffusion component corresponds to PEG as a single polymer chain. Meanwhile, the slow diffusion component might be the result of PEG binding either with itself or with macromolecules in the cell lysate. In light of the indications from rheology that the unmanipulated cell lysate may contain network-like structures that give rise to an elastic component to the response, the low value of $D_{PEG,slow}$ ($D_{PEG,fast}/D_{PEG,slow} \geq 30$) suggests that PEG might be interacting with these network-like structures in the lysate, or with large biomolecules.

In addition, there is a dependence of the diffusion coefficient of PEG on the concentration of the crowder. When the concentration of lysate increases, PEG diffuses slower; this is true for unmanipulated, ultracentrifuged and anion exchanged bacterial cell lysate. As we approach 150 mg/ml, PEG diffuses almost twice as fast in cell lysate as in Ficoll. Even though Ficoll has a significantly lower apparent viscosity than lysate, it seems it has a larger effect on PEG diffusivity than the lysate crowders. A possible explanation for this might be due to excluded volume effects. PEG and Ficoll have an estimated hydrodynamic radius of 5 nm. Meanwhile, cell lysate has different macromolecules at different sizes. This means that the PEG probe molecule encounters more obstacles around its own size when Ficoll is used as the crowder. On the other hand, for a given concentration of lysate, there are not big differences between the fast component of diffusion coefficients of PEG at the individual concentrations for the pretreatment lysates. However, in the unmanipulated lysate, up to 10% of the PEG is associated in a larger aggregate.

Overall, our measurements with *E. coli* lysate crowders show a decrease for the fast diffusion coefficient component of roughly 4-fold when the concentration is increased from 0 to 300 mg/ml. This result is consistent with the diffusivity of probe molecules in eukaryotic cytoplasm, 4-fold slower compared to no crowder. However, it is a weaker dependence than the 11-fold slower diffusivity in *E. coli* cytoplasm reported by Ellis [7]. Palit and coworkers [4, 70] showed that the diffusivity of a test molecule slows down more when the test molecules are larger than the crowder, which indicates that the relative size of the probe molecule to that of the crowder molecules should be relevant. Simulations have examined the effect of mixed macromolecular crowders on protein folding and misfolding [52, 134], but similar studies on the effects on diffusion would be welcome.

The modest dependence of the diffusion coefficient on the lysate treatment suggests that the smaller macromolecules in the lysate ($R_H \approx 4\ nm$), which are not removed by centrifugation, contribute more strongly to the slowing down of PEG. Since the R_H of PEG is $\approx 5\ nm$, this might suggest the slowing down is more pronounced when the crowder size and probe size are comparable. Another remark from our diffusion data is that PEG at 30 mg/ml in unmanipulated cell lysate is stable up to 200 mg/ml, and the ultracentrifuged and anion exchanged lysates are stable to 300 mg/ml, i.e. it is not phase separated right after mixing. Meanwhile, the Ficoll-PEG solution is phase separated beyond 150 mg/ml.

8.3 Diffusion coefficient of αS in cell lysate

As a first study for the intrinsically disordered protein αS (our test molecule), we measured HSQC spectra for αS freshly mixed with the lysate crowder and after 20 hours from mixing. Our HSQC data for αS shows no significant changes in the αS peaks.

Next, we used a HSQC-DOSY NMR experiment to measured the diffusivity of ^{15}N αS. However, we needed to check carefully that the diffusion coefficient were reasonable. To do so, we calibrated our experiments with a well known molecule: ^{15}N-labelled formamide. We compared the diffusivity values of formamide using two kind of experiments: DOSY and HSQC-DOSY. From both experiments, the diffusivity values of formamide were consistent.

After the calibration process, we carried out diffusivity studies of ^{15}N αS in three treatments of lysate crowders: unmanipulated, ultracentrifuged, and anion exchanged. The experiments were performed at a constant protein concentration but changing the lysate crowder concentration: 50, 75, and 100 mg/ml.

We compared our diffusivity results for αS with the non-ionic polymer (PEG). For each diffusivity value, there is a steady decrease as lysate concentration increases. For PEG diffusivity in ultracentrifuged and anion exchanged lysates, the behaviour looks very similar. However, that is not the case in αS. We want to remark that αS diffusivity experiments were done with a single replicate, so this unusual result indicates that multiple replicates might be required.

From the PEG study, it appeared that there was not much difference between ultracentrifuged and anion exchanged lysates. If that is indeed the case also in αS, one could treat the difference between ultracentrifuged and anion exchanged to reflect the sample-to-sample variation. This would also indicate that one would need, in future work, multiple replicate experiments of each lysate preparation.

8.4 Future directions

In this thesis, we have examined diffusion in a realistic crowding environment in comparison with the artificial crowder Ficoll. We have also examined diffusion of an intrinsically disordered protein, in comparison with a linear chain polymer, PEG, in realistic crowding environments.

However, there are still questions to explore regarding diffusivity in realistic crowders. As future directions, αS diffusivity could be measured at different concentrations, as well as different temperature. We could test αS hydrodynamic radius with realistic crowders. The diffusivity of αS should be studied and compared with artifical crowders, charged and uncharged, to evaluate the soft interaction effects using a bottom-up approach.

The study done in this thesis uses only different treatments of bacteria *E. coli* as the crowder. According to Schavemaker and Poolman [36], the diffusion of fluorescent proteins in the cytoplasm of the gram positive *Lactococcus lactis* bacteria and the *Haloferax volcanii* archaeon depends greatly on the net charge of the protein surface. Therefore,

the study should be expanded with different kinds of cells like the gram-positive bacteria *bacillus subtilis* and mammalian cells.

For the field as a whole, if artificial crowders are to be used, one should take into consideration both the effect of soft non-specific interactions and the non-Newtonian character of realistic crowding environments.

Appendix A

Appendix A

A.1 Unmanipulated cell lysate diffusion values

Table A.1 shows the diffusion values for 0.03 g/ml PEG, using as a crowder unmanipulated cell lysate at different concentrations (C_{CL}). Each diffusion coefficient was calculated with a bi-exponential fit.

The average diffusion coefficient for PEG is shown in table A.2. While there is not a clear trend with the slow component, the faster self diffusion values for PEG decreases with increasing the concentration of the crowders.

TABLE A.1: Fast (D_{fast} m^2/s) and slow (D_{slow} m^2/s) component self diffusion values for 0.03 g/ml PEG in unmanipulated cell lysate at different concentrations (C_{CL}). Data was obtained using diffusion NMR.

C_{CL}	f	$D_{fast} \times 10^{-11}$	$D_{slow} \times 10^{-13}$	Gradient (G/cm)	Folder
50 mg/ml	0.9702	2.14	6.13	50-1000	50LysTreatms
	0.9685	2.13	5.92	50-1000	(Sep 4,2019)
100 mg/ml	0.9639	2.10	11.1	50-400	LysFeb25,2019
	0.9702	1.70	3.45	50-1000	LysFeb27,2019
	0.9610	1.71	3.70	50-1000	LysMar2119
	0.9618	1.70	3.87	50-1000	LysMar2119
150 mg/ml	0.9431	1.96	12.1	50-400	LysFeb25,2019
	0.9516	1.86	4.24	50-1000	LysFeb27,2019
	0.9436	1.40	3.50	50-1000	LysMar2119
	0.9407	1.45	3.44	50-1000	LysMar2119
200 mg/ml	0.9336	1.47	15.2	50-300	LypOct09 (2018)
	0.9076	1.45	11.9	50-400	LysFeb26,2019
	0.9337	1.32	3.91	50-1000	LysMar01,2019
	0.9219	1.01	3.26	50-1000	LysMar2519
	0.9212	1.04	3.18	50-1000	LysMar2519
250 mg/ml	0.9259	1.18	3.95	50-1000	LysMar01,2019
	0.9013	0.77	3.16	50-1000	LysMar2519
	0.9032	0.79	3.09	50-1000	LysMar2519
300 mg/ml	0.8959	0.81	3.47	50-1000	LysMar2519
	0.8983	0.85	3.50	50-1000	LysMar2519
	0.8965	0.88	3.60	50-1000	LysMar2519

TABLE A.2: Average fast and slow component self diffusion values for 0.03 g/ml of PEG in unmanipulated cell lysate at different concentrations (C_{CL}).

C_{CL}	$< D_{fast} > \times 10^{-11}$ m^2/s	$< D_{slow} > \times 10^{-13}$ m^2/s
50 mg/ml	2.13	6.02
100 mg/ml	1.80	3.67
150 mg/ml	1.67	3.73
200 mg/ml	1.26	3.45
250 mg/ml	0.92	3.40
300 mg/ml	0.85	3.52

A.2 Ultracentrifuged cell lysate diffusion values

All experiments performed in diffusion NMR for PEG using lysate crowder treated by ultracentrifuged are displayed in table A.3. Each diffusion value was obtained from the fitting of bi-exponential equation. The fraction (f) represents the contribution of the

fast component in the bi-exponential fitting and are close to 1, which is an indicator that the graph can be fitted using the Stejskal-Tanner equation.

TABLE A.3: Fast (D_{fast} m^2/s) and slow (D_{slow} m^2/s) component self diffusion values for 0.03 g/ml PEG in ultracentrifuged cell lysate at different concentrations (C_{CL}). Data was obtained using diff NMR.

C_{CL}	f	$D_{fast} \times 10^{-11}$	$D_{slow} \times 10^{-13}$	Gradient (G/cm)	Folder
50 mg/ml	0.9997	2.07	5.64	50-1000	50LysTreatms
	0.9995	2.00	10.8	50-1000	50LysTreatms
100 mg/ml	0.9989	1.66	8.72	50-1000	LysMar2819
	0.9986	1.67	8.66	50-1000	LysMar2819
	0.9998	2.06	0.62	50-1000	UltraLyPEG
	0.9996	2.05	8.33	50-1000	UltraLyPEG
150 mg/ml	0.9995	1.60	4.24	50-1000	LysMar2819
	0.9994	1.55	2.80	50-1000	LysMar2819
	0.9993	1.78	3.04	50-1000	UltraLyPEG
	0.9985	1.82	4.96	50-1000	UltraLyPEG
200 mg/ml	0.9979	1.33	8.18	50-1000	LysMar2819
	0.9863	1.71	40.1	50-1000	LysMar2819
	0.9880	1.67	44.4	50-1000	LysMar2819
	0.9992	1.39	6.08	50-1000	UltraLyPEG
250 mg/ml	0.9994	1.21	2.40	50-1000	UltraLyPEG
	0.9999	1.49	8.07	50-1000	UltraLys2
	0.9985	1.06	6.54	50-1000	UltraLys2
300 mg/ml	0.9969	0.82	8.23	50-1000	UltraLyPEG
	0.9933	0.64	2.46	50-1000	UltraLys2
	0.9918	0.65	2.92	50-1000	UltraLys2

Table A.4 displays the average self diffusion values for the fast and slow components of PEG in ultracentrifuged cell lysate.

TABLE A.4: Average fast and slow component self diffusion values for 0.03 g/ml of PEG in ultracentrifuge cell lysate at different concentrations (C_{CL}).

C_{CL}	$< D_{fast} > \times 10^{-11}$ m^2/s	$< D_{slow} > \times 10^{-13}$ m^2/s
50 mg/ml	2.03	8.21
100 mg/ml	1.86	8.57
150 mg/ml	1.69	3.76
200 mg/ml	1.52	0.25
250 mg/ml	1.25	5.67
300 mg/ml	0.70	0.29

A.3 Anion exchanged cell lysate diffusion values

All the fast and slow components of the self diffusion coefficient for PEG using as a crowder anion exchange cell lysate are shown in the table A.3. All the diffusion values were obtained from a bi-exponential fitting. The average values for the diffusion of PEG are in table A.6.

TABLE A.5: Fast (D_{fast} m^2/s) and slow (D_{slow} m^2/s) component self diffusion values for 0.03 g/ml of PEG in anion exchange cell lysate at different concentrations (C_{CL}). Data was obtained using diff NMR.

C_{CL}	f	$D_{fast} \times 10^{-11}$	$D_{slow} \times 10^{-13}$	Gradient (G/cm)	Folder
50 mg/ml	0.9988	1.82	8.87	50-1000	50LysTreatms
	0.9983	1.86	1.06	50-1000	50LysTreatms
100 mg/ml	0.9983	1.87	9.18	50-1000	AnioExcpH7
	0.9986	1.89	9.04	50-1000	AnioExcpH7
150 mg/ml	0.9959	1.46	9.63	50-1000	AnioExcpH7
	0.9959	1.50	9.65	50-1000	AnioExcpH7
200 mg/ml	0.9987	1.24	6.69	50-1000	AnioExcpH7
	0.9971	1.34	7.02	50-1000	AnioExcpH7
250 mg/ml	0.9982	1.07	3.99	50-1000	AnioExcpH7
	0.9979	1.08	2.22	50-1000	AnioExcpH7
300 mg/ml	0.9949	0.96	7.24	50-1000	AnioExcpH7
	0.9935	0.89	6.96	50-1000	AnioExcpH7

TABLE A.6: Average fast and slow component self diffusion values for 0.03 g/ml of PEG in anion exchange cell lysate at different concentrations (C_{CL}).

C_{CL}	$< D_{fast} > \times 10^{-11}$ m^2/s	$< D_{slow} > \times 10^{-13}$ m^2/s
50 mg/ml	1.84	9.72
100 mg/ml	1.88	9.11
150 mg/ml	1.48	9.64
200 mg/ml	1.29	6.86
250 mg/ml	1.08	3.11
300 mg/ml	0.92	7.10

Appendix B

Appendix B

B.1 Calculation of macromolecule radii removed in ultracentrifugation

From Stokes' Law, we know that the drag force F is

$$F = 6\pi\eta R_H v_t, \tag{B.1}$$

where R_H is the hydrodynamic radius, v_t is the velocity of the macromolecule, and viscosity $\eta = 0.001\ Pa \cdot s$. Combining Stoke's law with Newton's law, we can obtain for v_t

$$v_t = \frac{ma}{6\pi\eta}\frac{1}{R_H}. \tag{B.2}$$

The mass can be obtained assuming an spherical macromolecule with volume $V = \frac{4}{3}\pi\Delta\rho R_H^3$, where $\Delta\rho$ is its density (approximately 1.4 g/m^3 for proteins [112]). In this case, the acceleration corresponds to the ultracentrifuge acceleration equal to 100,000 g. So, the velocity of the macromolecule is

$$v_t = \frac{\frac{4}{3}\pi\Delta\rho 10^5 g}{6\pi\eta}R_H^2. \tag{B.3}$$

Considering a 5 cm long sample size centrifuged tube and 2400 seconds the duration of ultracentrifugation, we can estimate the velocity of the macromolecules by $v_t = length/time = 2.08 \times 10^{-5}\ m/s$. Replacing all this parameters, we can calculate from equation B.3

$$R_H = \sqrt{\frac{v_t 6\pi\eta}{\frac{4}{3}\pi\Delta\rho 10^5 g}}. \tag{B.4}$$

From this calculation, it is estimated that macromolecules with a hydrodynamic radius around 10 nm and larger should be removed by ultracentrifugation.

Appendix C

Appendix C

C.1 Supporting material for Chapter 6

We want to clarify that we do not separate the cytoplasm from the cell membrane for our unmanipulated lysate crowders. In our lysate preparation, we use the pellet that contains most of the cell material, however some of the smaller macromolecules, like some proteins, are discarded with the supernatant.

In order to estimate the protein concentration in our lysate crowders we performed a Bradford assay experiment of unmanipulated, ultracentrifuged and anion exchanged lysate. As a standard curve, we use the protein BSA. From our results we estimated that the weight fraction of protein in unmanipulated lysate is 39 ± 6 %, in ultracentrifuged is 33 ± 5 %, and in anion exchanged is 59 ± 9 %. This fraction is small; however, the Bradford assay is not the most accurate technique for our kind of sample, due to the variety of proteins on it.

The Bradford assay Coomassie Brilliant Blue G-250 dye binds mainly to the basic amino acids residues, like arginine (specially arginine), lysine, and histidine, as well as, binds to the aromatic amino acids residues, like phenylalanine, tyrosine, tryptophan and histidine [135]. Considering that Shemesh and coworkers reported that the most abundant amino acids in *E. coli* are methionine followed by lysine [136], the fraction reported from the Bradford assay should be treated as a lower bound.

A better technique to estimate protein concentration in our lysate sample would be a modified Lowry Protein Assay (as carried out by Sarkar 2013 [137]), since it depends on the interaction of the proteins with Cooper and Folin reagent. This will be done in the future.

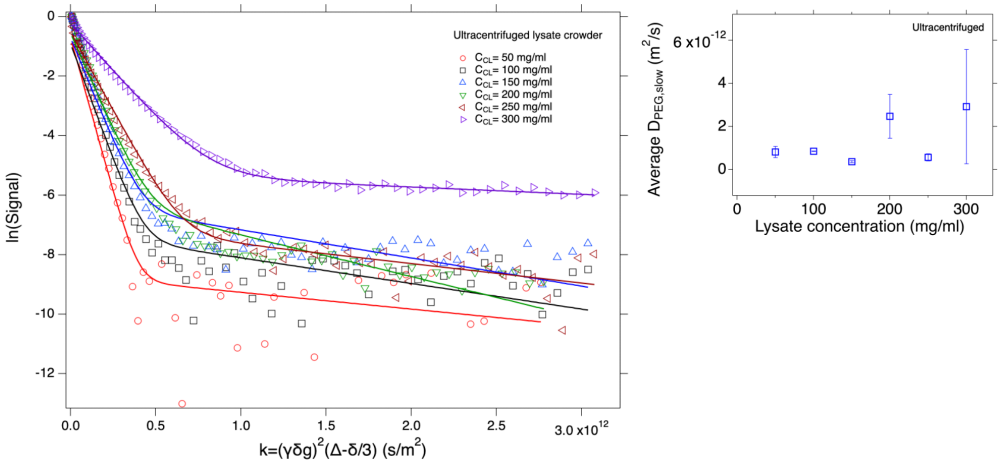


FIGURE C.1: **PEG in ultracentrifuged cell lysate crowder has two diffusion coefficients**. The logarithm of the attenuated signal decay versus the gradient strength parameter (k) for 30 mg/ml PEG with different concentrations of ultracentrifuged cell lysate C_{CL}. Each signal intensity decay was fitted using a bi-exponential equation with a fast and slow diffusion coefficient components (Left Figure). Average slow diffusion component (Right Figure)

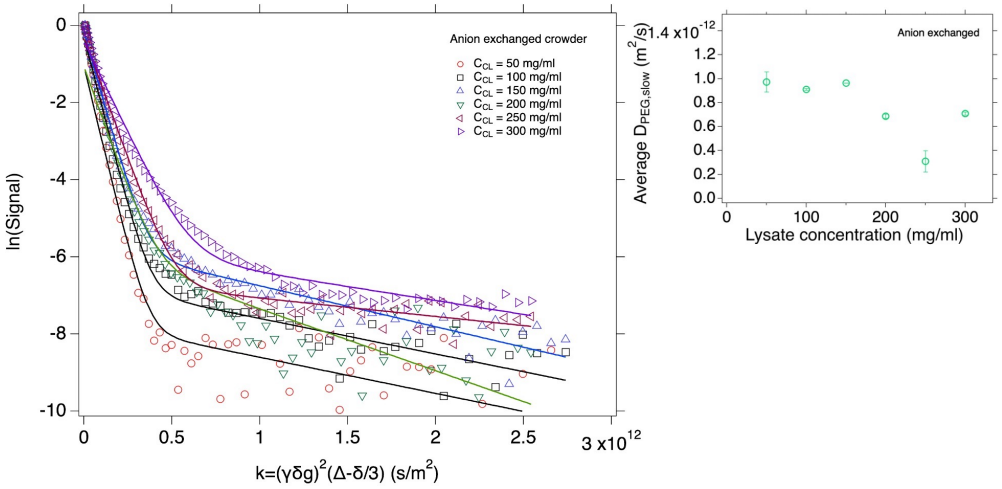


FIGURE C.2: **PEG in anion exchanged cell lysate crowder has two diffusion coefficients**. The logarithm of the attenuated signal decay versus the gradient strength parameter (k) for 30 mg/ml PEG with different concentrations of anion exchanged cell lysate C_{CL}. Each signal intensity decay was fitted using a bi-exponential equation with a fast and slow component diffusion coefficient (Left Figure). Average slow diffusion component (Right Figure).

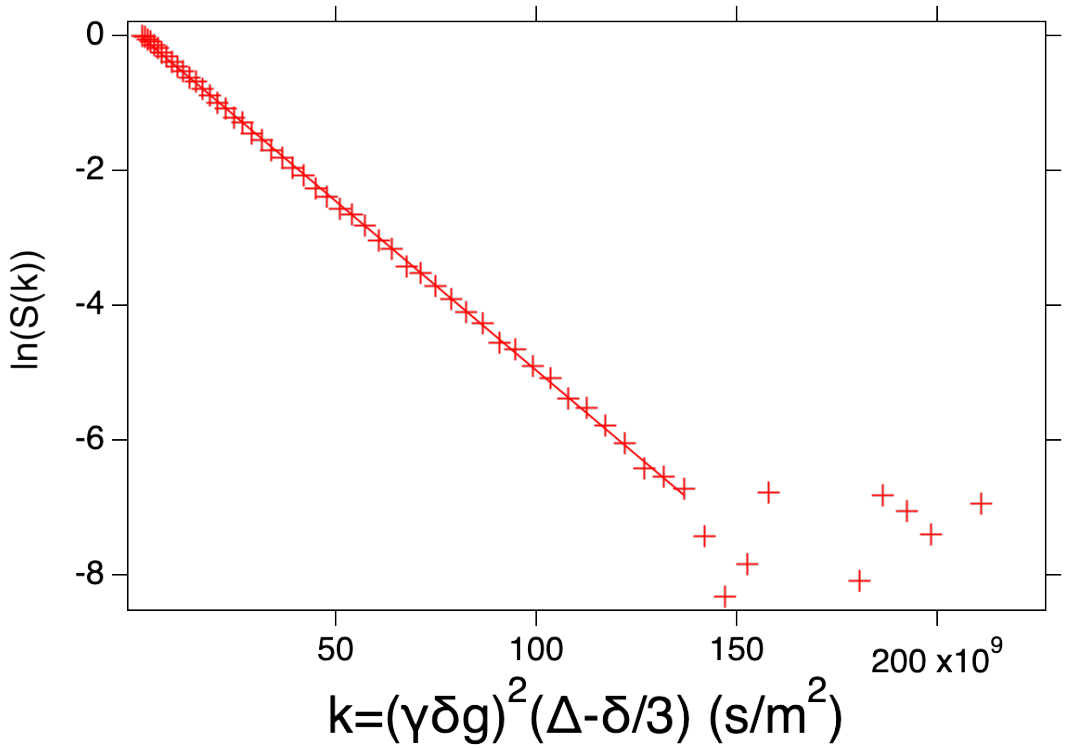


FIGURE C.3: **3** *mg/ml* **PEG in distilled water has one diffusion coefficient**. The logarithm of the attenuated signal decay versus the gradient strength parameter (k). The signal intensity decay was fitted using a mono-exponential equation.

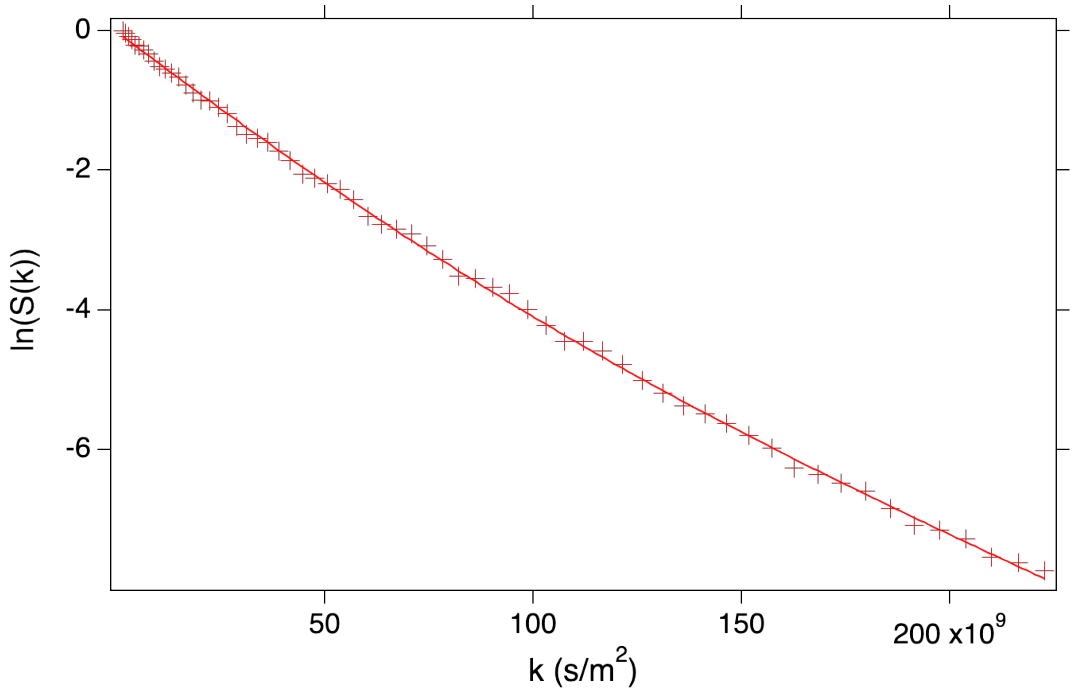
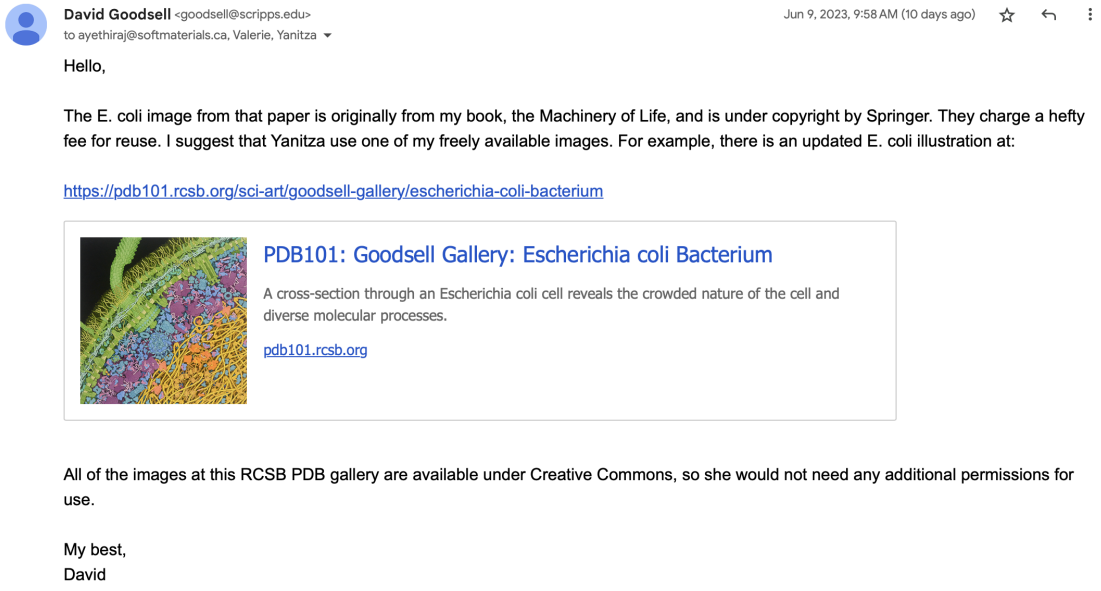


FIGURE C.4: **30** *mg/ml* **PEG in distilled water has two diffusion coefficients**. The logarithm of the attenuated signal decay versus the gradient strength parameter (k). The signal intensity decay was fitted using a bi-exponential equation with a fast and slow component diffusion coefficient.

Appendix D

Permissions

Figure 2.2 is available under Creative Commons.



David Goodsell <goodsell@scripps.edu>
to ayethiraj@softmaterials.ca, Valerie, Yanitza

Jun 9, 2023, 9:58 AM (10 days ago)

Hello,

The E. coli image from that paper is originally from my book, the Machinery of Life, and is under copyright by Springer. They charge a hefty fee for reuse. I suggest that Yanitza use one of my freely available images. For example, there is an updated E. coli illustration at:

https://pdb101.rcsb.org/sci-art/goodsell-gallery/escherichia-coli-bacterium

All of the images at this RCSB PDB gallery are available under Creative Commons, so she would not need any additional permissions for use.

My best,
David

Bibliography

[1] Gregory Liam. P. Booth Valerie. K. Yethiraj Anand Trosel, Yanitza. Diffusion nmr and rheology of a model polymer in bacterial cell lysate crowders. *Biomacromolecules*, 24(6):2469–2478, 2023.

[2] David S. Goodsell. Escherichia coli bacterium. *RCSB Protein Data Bank*, 2021.

[3] Damiano Piovesan, Francesco Tabaro, Ivan Mičetić, Marco Necci, Federica Quaglia, Christopher J Oldfield, Maria Cristina Aspromonte, Norman E Davey, Radoslav Davidović, Zsuzsanna Dosztányi, et al. Disprot 7.0: a major update of the database of disordered proteins. *Nucleic acids research*, 45(D1):D219–D227, 2017.

[4] Swomitra Palit, Lilin He, William A. Hamilton, Arun Yethiraj, and Anand Yethiraj. The effect of crowder charge in a model polymer-colloid system for macromolecular crowding: Polymer structure and dynamics. *Journal of Chemical Physics*, 147:114902, 2017.

[5] Ron Milo. What is the total number of protein molecules per cell volume? a call to rethink some published values. *BioEssays*, 35:1050–1055, 12 2013.

[6] Alice B Fulton. How crowded is the cytoplasm? *Cell*, 30:345–347, 1982.

[7] R. John Ellis. Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Current Opinion in Structural Biology*, 11:114–119, 2001.

[8] Irina M. Kuznetsova, Konstantin K. Turoverov, and Vladimir N. Uversky. What macromolecular crowding can do to a protein. *International Journal of Molecular Sciences*, 15:23090–23140, 2014.

[9] Shannon L Speer, Claire J Stewart, Liel Sapir, Daniel Harries, and Gary J Pielak. Annual review of biophysics macromolecular crowding is more than hard-core repulsions. *Annual Review Biophysics*, 51:267–300, 2022.

[10] R. John Ellis and A. P. Minton. Cell biology- join the crowd. *Nature*, 425:27– 28, 2003.

[11] Sumra Shahid, Md Imtaiyaz Hassan, Asimul Islam, and Faizan Ahmad. Size-dependent studies of macromolecular crowding on the thermodynamic stability, structure and functional activity of proteins: in vitro and in silico approaches. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1861(2):178–197, 2017.

[12] Rob Phillips, Jane Kondev, Julie Theriot, Hernan G. Garcia, and Nigel Orme. *Physical Biology of the Cell*. Garland Science, second edition, 2013.

[13] Francois Xavier Theillet, Andres Binolfi, Tamara Frembgen-Kesner, Karan Hingorani, Mohona Sarkar, Ciara Kyne, Conggang Li, Peter B. Crowley, Lila Gierasch, Gary J. Pielak, Adrian H. Elcock, Anne Gershenson, and Philipp Selenko. Physicochemical properties of cells and their effects on intrinsically disordered proteins (idps). *Chemical Reviews*, 114:6661–6714, 7 2014.

[14] R. John Ellis. Macromolecular crowding: Obvious but underappreciated. *Trends in Biochemical Sciences*, 26:597– 604, 2001.

[15] Steven B. Zimmerman and Stefan O. Trach. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of escherichia coli. *Journal of Molecular Biology*, 222:599–620, 12 1991.

[16] Germán Rivas and Allen P Minton. Macromolecular crowding in vitro, in vivo, and in between. *Trends in biochemical sciences*, 41(11):970–981, 2016.

[17] Andrew C. Miklos, Conggang Li, Naima G. Sharaf, and Gary J. Pielak. Volume exclusion and soft interaction effects on protein stability under crowded conditions. *Biochemistry*, 49:6984–6991, 8 2010.

[18] Mohona Sarkar, Conggang Li, and Gary J. Pielak. Soft interactions and crowding. *Biophysical Reviews*, 5:187–194, 2013.

[19] Germán Rivas and Allen P Minton. Annual review of biochemistry influence of nonspecific interactions on protein associations: Implications for biochemistry in vivo. *Annual Reviews Biochemistry*, 91:321–351, 2022.

[20] Yaqiang Wang, Conggang Li, and Gary J. Pielak. Effects of proteins on protein diffusion. *Journal of the American Chemical Society*, 132:9392–9397, 2010.

[21] Conggang Li and Gary J. Pielak. Using nmr to distinguish viscosity effects from nonspecific protein binding under crowded conditions. *Journal of the American Chemical Society*, 131:1368–1369, 2 2009.

[22] Irina M. Kuznetsova, Boris Y. Zaslavsky, Leonid Breydo, Konstantin K. Turoverov, and Vladimir N. Uversky. Beyond the excluded volume effects: Mechanistic complexity of the crowded milieu. *Molecules*, 20:1377–1409, 2015.

[23] M. Erlkamp, S. Grobelny, and R. Winter. Crowding effects on the temperature and pressure dependent structure, stability and folding kinetics of staphylococcal nuclease. *Physical Chemistry Chemical Physics*, 16:5965–5976, 2014.

[24] Conggang Li, Yaqiang Wang, and Gary J. Pielak. Translational and rotational diffusion of a small globular protein under crowded conditions. *Journal of Physical Chemistry B*, 113:13390–13392, 10 2009.

[25] Lisa M. Charlton, Christopher O. Barnes, Conggang Li, Julian Orans, Gregory B. Young, and Gary J. Pielak. Residue-level interrogation of macromolecular crowding effects on protein stability. *Journal of the American Chemical Society*, 130:6826–6830, 5 2008.

[26] Barak Akabayov, Sabine R. Akabayov, Seung Joo Lee, Gerhard Wagner, and Charles C. Richardson. Impact of macromolecular crowding on dna replication. *Nature Communications*, 4:1615, 2013.

[27] Suliman Barhoum, Swomitra Palit, and Anand Yethiraj. Diffusion nmr studies of macromolecular complex formation, crowding and confinement in soft materials. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 94-95:1–10, 2016.

[28] Antonios Samiotakis, Pernilla Wittung-Stafshede, and Margaret S. Cheung. Folding, stability and shape of proteins in crowded environments: Experimental and computational approaches. *International Journal of Molecular Sciences*, 10:572–588, 2 2009.

[29] Torvard C. Laurent. The interaction between polysaccharides and other macromolecules. *Biochem. J*, 89:253–257, 1963.

[30] Martin Muschol and Franz Rosenberger. Interactions in undersaturated and supersaturated lysozyme solutions: static and dynamic light scattering results. *The Journal of chemical physics*, 103(24):10424–10432, 1995.

[31] Thomas Kühn, Teemu O. Ihalainen, Jari Hyväluoma, Nicolas Dross, Sami F. Willman, Jörg Langowski, Maija Vihinen-Ranta, and Jussi Timonen. Protein diffusion in mammalian cell cytoplasm. *PLoS ONE*, 6:e22962, 2011.

[32] Silvia Zorrilla, Mark A. Hink, Antonie J.W.G. Visser, and M. Pilar Lillo. Translational and rotational motions of proteins in a protein crowded environment. *Biophysical Chemistry*, 125:298–305, 2 2007.

[33] David Knoll and Jan Hermans. Polymer-protein interactions. comparison of experiment and excluded volume theory. *The Journal of Biological Chemistry*, 258:5710–5715, 1983.

[34] Luis C. Acosta, Gerardo M. Perez Goncalves, Gary J. Pielak, and Annelise H. Gorensek-Benitez. Large cosolutes, small cosolutes, and dihydrofolate reductase activity. *Protein Science*, 26:2417–2425, 12 2017.

[35] David Nelson and Michael Cox. *Lehninger Principles of Biochemistry*. W. H. Freeman and Company, sixth edition, 2013.

[36] Paul E Schavemaker, Wojciech M S, and Bert Poolman. Ribosome surface properties may impose limits on the nature of the cytoplasmic proteome. *eLife*, 6:e30084, 2017.

[37] Roland Glaser. *Biophysics*. Springer, second edition, 2012.

[38] Jean Philibert. One and a half century of diffusion: Fick, einstein, before and beyond. *Diffusion Fundamentals*, 2:1–10, 2005.

[39] Albert Einstein. *Investigations on the theory of the Brownian movement, translated by A. D. Cowper*. Dover publications, INC., 1956.

[40] Paul T. Callaghan. Physics of Diffusion. In *Diffusion MRI: Theory, Methods, and Applications*. Oxford University Press, 11 2010.

[41] M. H. Jacobs. *Diffusion Processes*. Springer Berlin Heidelberg, 1967.

[42] Robert M. Mazo. *Brownian Motion: Fluctuations, Dynamics, and Applications*. Oxford University Press, 2009.

[43] Don S Lemons and Anthony Gythiel. Paul langevin's 1908 paper "on the theory of brownian motion"["sur la théorie du mouvement brownien," cr acad. sci.(paris) 146, 530–533 (1908)]. *American Journal of Physics*, 65(11):1079–1081, 1997.

[44] S. Chandrashekhar. Stochastic problems in physics and astronomy. *Reviews of Modern Physics*, 15:1–89, 1943.

[45] Ronald Newburgh, Joseph Peidle, and Wolfgang Rueckner. Einstein, perrin, and the reality of atoms: 1905 revisited. *American Journal of Physics*, 74:478–481, 6 2006.

[46] Gert Strobl. *The physics of polymers: Concepts for understanding their structures and behavior*. Springer Berlin Heidelberg, 2007.

[47] Richard A. L. Jones. *Soft Condensed Matter*. Oxford University Press, first edition, 2002.

[48] Anshuman Shrivastava. *Introduction to Plastics Engineering*. Elsevier Inc., 2018.

[49] Lila M. Gierasch and Anne Gershenson. Post-reductionist protein science, or putting humpty dumpty back together again. *Nature Chemical Biology*, 5:774–777, 2009.

[50] Allen P. Minton. Implications of macromolecular crowding for protein assembly. *Current Opinion in Structural Biology*, 10:34–39, 2 2000.

[51] G. B. Ralston. Effects of "crowding" in protein solutions. *Journal of Chemical Education*, 67:857–860, 1990.

[52] Huan-Xiang Zhou, Germán Rivas, and Allen P. Minton. Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annual Review of Biophysics*, 37:375–397, 2008.

[53] Samuel H Schneider, Schuyler P Lockwood, Dominique I Hargreaves, David J Slade, Micaela A LoConte, Bridget E Logan, Erin E McLaughlin, Michael J Conroy, and Kristin M Slade. Slowed diffusion and excluded volume both contribute to the effects of macromolecular crowding on alcohol dehydrogenase steady-state kinetics. *Biochemistry*, 54(38):5898–5906, 2015.

[54] Allen P Minton. Influence of excluded volume upon macromolecular structure and associations in crowded media. *Current Opinion in Biotechnology*, 8:65–69, 1997.

[55] G Foffi, A Pastore, Federico Piazza, and PA Temussi. Macromolecular crowding: chemistry and physics meet biology (ascona, switzerland, 10–14 june 2012). *Physical biology*, 10(4):040301, 2013.

[56] Pinki Dey and Arnab Bhattacherjee. Role of macromolecular crowding on the intracellular diffusion of dna binding proteins. *Scientific Reports*, 8(1):844, 2018.

[57] Yaqiang Wang, Mohona Sarkar, Austin E Smith, Alexander S Krois, and Gary J Pielak. Macromolecular crowding and protein stability. *Journal of the American Chemical Society*, 134(40):16614–16618, 2012.

[58] Allen P. Minton. The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *Journal of Biological Chemistry*, 276:10577–10580, 4 2001.

[59] Allen P. Minton and Germán Rivas. Biochemical reactions in the crowded and confined physiological environment: Physical chemistry meets synthetic biology, 2011.

[60] Damien Hall and Allen P. Minton. Macromolecular crowding: Qualitative and semiquantitative successes, quantitative challenges. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1649:127–139, 2003.

[61] Davide Marenduzzo, Kieran Finan, and Peter R. Cook. The depletion attraction: An underappreciated force driving cellular organization. *Journal of Cell Biology*, 175:681–686, 2006.

[62] Liel Sapir and Daniel Harries. Is the depletion force entropic? molecular crowding beyond steric interactions. *Current Opinion in Colloid and Interface Science*, 20:3–10, 2015.

[63] Mohona Sarkar, Joe Lu, and Gary J. Pielak. Protein crowder charge and protein stability. *Biochemistry*, 53:1601–1606, 2014.

[64] Yaqiang Wang, Laura A. Benton, Vishavpreet Singh, and Gary J. Pielak. Disordered protein diffusion under crowded conditions. *Journal of Physical Chemistry Letters*, 3:2703–2706, 2012.

[65] Steven B. Zimmerman. Macromolecular crowding effects on macromolecular interactions: Some implications for genome structure and function. *Biochimica et Biophysica Acta*, 1216:175–185, 1993.

[66] Grzegorz Nawrocki, Alp Karaboga, Yuji Sugita, and Michael Feig. Effect of protein-protein interactions and solvent viscosity on the rotational diffusion of proteins in crowded environments. *Physical Chemistry Chemical Physics*, 21:876–883, 2019.

[67] Swomitra Palit and Anand Yethiraj. A new model system for diffusion nmr studies of concentrated monodisperse and bidisperse colloids. *Langmuir*, 24:3747, 2008.

[68] Jessica Balbo, Paolo Mereghetti, Dirk Peter Herten, and Rebecca C. Wade. The shape of protein crowders is a major determinant of protein diffusion. *Biophysical Journal*, 104:1576–1584, 4 2013.

[69] Shruti Mittal, Rimpy Kaur Chowhan, and Laishram Rajendrakumar Singh. Macromolecular crowding: Macromolecules friend or foe. *Biochimica et Biophysica Acta - General Subjects*, 1850:1822–1831, 2015.

[70] Swomitra Palit and Anand Yethiraj. Dynamics and cluster formation in charged and uncharged ficoll70 solutions. *Journal of Chemical Physics*, 147:074901, 2017.

[71] A. K. Dunker, Z. Obradovic, P. Romero, E. C. Garner, and C. J. Brown. Intrinsic protein disorder in complete genomes. *Genome informatics. Workshop on Genome Informatics*, 11:161–171, 2000.

[72] Christopher J. Oldfield, Yugong Cheng, Marc S. Cortese, Celeste J. Brown, Vladimir N. Uversky, and A. Keith Bunker. Comparing and combining predictors of mostly disordered proteins. *Biochemistry*, 44:1989–2000, 2005.

[73] Vladimir N. Uversky, Joel R. Gillespie, and Anthony L. Fink. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins: Structure, Function and Genetics*, 41:415–427, 2000.

[74] A. Keith Dunker, M. Madan Babu, Elisar Barbar, Martin Blackledge, Sarah E. Bondos, Zsuzsanna Dosztányi, H. Jane Dyson, Julie Forman-Kay, Monika Fuxreiter, Jörg Gsponer, Kyou-Hoon Han, David T. Jones, Sonia Longhi, Steven J. Metallo, Ken Nishikawa, Ruth Nussinov, Zoran Obradovic, Rohit V. Pappu, Burkhard

Rost, Philipp Selenko, Vinod Subramaniam, Joel L. Sussman, Peter Tompa, and Vladimir N Uversky. What's in a name? why these proteins are intrinsically disordered. *Intrinsically Disordered Proteins*, 1, 2013.

[75] Peter E. Wright and H. Jane Dyson. Intrinsically unstructured proteins: Reassessing the protein structure-function paradigm. *Journal of Molecular Biology*, 293:321–331, 1999.

[76] Christopher J. Oldfield and A. Keith Dunker. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annual Review of Biochemistry*, 85:553–584, 2014.

[77] V. N. Uversky. Natively unfolded proteins: A point where biology waits for physics. *Protein Science*, 11:739–756, 4 2002.

[78] Vladimir N. Uversky. Intrinsically disordered proteins from a to z. *International Journal of Biochemistry and Cell Biology*, 43:1090–1103, 2011.

[79] Peter Tompa. Intrinsically disordered proteins: A 10-year recap. *Trends in Biochemical Sciences*, 37:509–516, 12 2012.

[80] Klaus Beyer. Mechanistic aspects of parkinson's disease: α-synuclein and the biomembrane. *Cell Biochemistry and Biophysics*, 47:285–299, 2007.

[81] A. Keith Dunker and Zoran Obradovic. The protein trinity - linking function and disorder. *Nature Biotechnology*, 19:805–806, 2001.

[82] H. Jeong, S. P. Mason, A. L. Barabási, and Z. N. Oltvai. Lethality and centrality in protein networks. *Nature*, 411:41–42, 2001.

[83] Andreas Prestel, Katrine Bugge, Lasse Staby, Ruth Hendus-Altenburger, and Birthe B. Kragelund. *Characterization of Dynamic IDP Complexes by NMR Spectroscopy*, volume 611. Elsevier Inc., 1 edition, 2018.

[84] Vladimir N Uversky. *Intrinsically Disordered Proteins*. Springer Briefs in Molecular Science, 2014.

[85] V. N. Uversky. Intrinsically disordered proteins. *Brenner's Encyclopedia of Genetics: Second Edition*, 3263:124–126, 2013.

[86] Vladimir N Uversky. Biophysical methods to investigate intrinsically disordered proteins: avoiding an "elephant and blind men" situation. *Intrinsically Disordered Proteins Studied by NMR Spectroscopy*, pages 215–260, 2015.

[87] Veronique Receveur-Brechot and Dominique Durand. How random are intrinsically disordered proteins? a small angle scattering perspective. *Current Protein & Peptide Science*, 13:55–75, 2012.

[88] Francois Xavier Theillet, Andres Binolfi, Beata Bekei, Andrea Martorana, Honor May Rose, Marchel Stuiver, Silvia Verzini, Dorothea Lorenz, Marleen Van Rossum, Daniella Goldfarb, and Philipp Selenko. Structural disorder of monomeric α-synuclein persists in mammalian cells. *Nature*, 530:45–50, 2016.

[89] Vladimir N. Uversky, Jie Li, Pierre Souillac, Ian S. Millett, Sebastian Doniach, Ross Jakes, Michel Goedert, and Anthony L. Fink. Biophysical properties of the synucleins and their propensities to fibrillate: Inhibition of α-synuclein assembly by β- and γ-synucleins. *Journal of Biological Chemistry*, 277:11970–11978, 2002.

[90] Jia Bai, Maili Liu, Gary J Pielak, and Conggang Li. Macromolecular and small molecular crowding have similar effects on α-synuclein structure. *ChemPhysChem*, 18(1):55–58, 2017.

[91] Jaie Woodard, Kinshuk R Srivastava, Gil Rahamim, Asaf Grupi, Steven Hogan, David J Witalka, Grzegorz Nawrocki, Elisha Haas, Michael Feig, and Lisa J Lapidus. Intramolecular diffusion in alpha-synuclein: It depends on how you measure it. *Biophysj*, 115:1190–1199, 2018.

[92] Vladimir N. Uversky. Neuropathology, biochemistry, and biophysics of α-synuclein aggregation. *Journal of Neurochemistry*, 103:17–37, 2007.

[93] Elio A. Cino, Mikko Karttunen, and Wing Yiu Choy. Effects of molecular crowding on the dynamics of intrinsically disordered proteins. *PLoS ONE*, 7, 2012.

[94] Emmanuel Dauty and A. S. Verkman. Molecular crowding reduces to a similar extent the diffusion of small solutes and macromolecules: Measurement by fluorescence correlation spectroscopy. *Journal of Molecular Recognition*, 17:441–447, 2004.

[95] Daniel S. Banks and Cécile Fradin. Anomalous diffusion of proteins due to molecular crowding. *Biophysical Journal*, 89:2960–2971, 2005.

[96] Aron B. Goins, Hugo Sanabria, and M. Neal Waxham. Macromolecular crowding and size effects on probe microviscosity. *Biophysical Journal*, 95:5362–5373, 2008.

[97] Thomas G. Mezger. *The Rheology Handbook*. Vincentz Network, 2nd edition, 2006.

[98] Dianne Stephenson, Fiona Norman, and R. H. Cumming. Shear thickening of dna in sds lysates. *Bioseparation*, 3:285– 289, 1993.

[99] L. A. S. Ciccolini, P. Ayazi Shamlou, N.J. Titchener-Hooker, J. M. Ward, and P. Dunnill. Rheological properties of chromosomal and plasmid dna during alkaline lysis reaction. *Bioprocess Engineering*, 21:231– 237, 1999.

[100] Marco Graziano, Swomitra Palit, Anand Yethiraj, Simone Immler, Matthew J. G. Gage, and Craig F. Purchase. Frequency-dependent viscosity of salmon ovarian fluid has biophysical implications for sperm–egg interactions. *Journal of Experimental Biology*, 226:1–10, 1 2023.

[101] Jane Y.Y. Chui, Carine Douarche, Harold Auradou, and Ruben Juanes. Rheology of bacterial superfluids in viscous environments. *Soft Matter*, 17:7004–7013, 8 2021.

[102] Jérémie Gachelin, Gastón Miño, Hélène Berthet, Anke Lindner, Annie Rousselet, and Éric Clément. Non-newtonian viscosity of escherichia coli suspensions. *Physical Review Letters*, 110, 6 2013.

[103] Felix Bloch and Edward Mills Purcell. The nobel prize in physics 1952. *Nature*, 170:911–912, 1952.

[104] John C Chatham and Stephen J Blackband. Nuclear magnetic resonance spectroscopy and imaging in animal research. *ILAR Journal*, 42:189–208, 2001.

[105] Hongtao Yu. Extending the size limit of protein nuclear magnetic resonance. *Proceedings of the National Academy of Sciences*, 96(2):332–334, 1999.

[106] James Keeler. *Understanding NMR Spectroscopy*. Wiley, 2005.

[107] Stefan Berger and Siegmar Braun. *200 and more NMR experiments*. Wiley-VCH, 2nd edition, 2004.

[108] Kyungsoo Shin, Muzaddid Sarker, Shuya K. Huang, and Jan K. Rainey. Apelin conformational and binding equilibria upon micelle interaction primarily depend on membrane-mimetic headgroup. *Scientific Reports*, 7, 12 2017.

[109] Donghui Wu, Aidi Chen, and Charles S Johnson Jr. Three-dimensional diffusion-ordered nmr spectroscopy: the homonuclear cosy–dosy experiment. *Journal of Magnetic Resonance, Series A*, 121(1):88–91, 1996.

[110] Fabien Ferrage, Manuela Zoonens, Dror E. Warschawski, Jean Luc Popot, and Geoffrey Bodenhausen. Slow diffusion of macromolecular assemblies by a new pulsed field gradient nmr method. *Journal of the American Chemical Society*, 125:2541–2545, 3 2003.

[111] Marcus L Tillett, Mark A Horsfield, Lu-Yun Lian, and Timothy J Norwood. Protein-ligand interactions measured by 15 n-filtered diffusion experiments, 1999.

[112] Hannes Fischer, Igor Polikarpov, and Aldo F. Craievich. Average protein density is a molecular-weight-dependent function. *Protein Science*, 13:2825–2828, 1 2009.

[113] Martial Piotto, Vladimir Saudek, and Vladimir Sklenar. Gradient-tailored excitation for single-quantum nmr spectroscopy of aqueous solutions. *Journal of Biomolecular NMR*, 2:661–665, 1992.

[114] Vladimir Sklenar, Martial Piotto, Raymond Leppik, and Vladimir Saudek. {Gradient-tailored water suppression for 1H-15N HSQC experiments optimized to retain full sensitivity}. *Journal of Magnetic Resonance, Series A*, 102:241–245, 1993.

[115] William S Price. Pulsed-field gradient nuclear magnetic resonance as a tool for studying translational diffusion: Part 1. basic theory. *Conc. Magn. Reson*, 9:299–336, 1997.

[116] M. Holz and H. Weingartner. Calibration in accurate spin-echo self-diffusion measurements using 1h and less-common nuclei. *Journal of Magnetic Resonance (1969)*, 92:115–125, 1991.

[117] Stephen Pansino and Benoit Taisne. Shear wave measurements of a gelatin's young's modulus. *Frontiers in Earth Science*, 8:1–13, 5 2020.

[118] Hurmiz Shamana and John R. Dutcher. Transition in the glassy dynamics of melts of acid-hydrolyzed phytoglycogen nanoparticles. *Biomacromolecules*, 23:2040–2050, 2022.

[119] Xiao Liang, Chunyu Liao, Michael L. Thompson, Michelle L. Soupir, Laura R. Jarboe, and Philip M. Dixon. E. coli surface properties differ between stream water and sediment environments. *Frontiers in Microbiology*, 7:1732, 11 2016.

[120] Benjamin J. Zeskind, Caroline D. Jordan, Winston Timp, Linda Trapani, Guichy Waller, Victor Horodincu, Daniel J. Ehrlich, and Paul Matsudaira. Nucleic acid and protein mass mapping by live-cell deep-ultraviolet microscopy. *Nature Methods*, 4:567–569, 7 2007.

[121] Kathy R Albe, Margaret H Butler, and Barbara E Wright. Cellular concentrations of enzymes and their substrates. *Journal of theoretical biology*, 143(2):163–195, 1990.

[122] T Reid Alderson and John L Markley. Biophysical characterization of α-synuclein and its controversial structure. *Intrinsically disordered proteins*, 1(1):e26255, 2013.

[123] Brian C McNulty, Gregory B Young, and Gary J Pielak. Macromolecular crowding in the escherichia coli periplasm maintains α-synuclein disorder. *Journal of molecular biology*, 355(5):893–897, 2006.

[124] Jie Li, Vladimir N Uversky, and Anthony L Fink. Effect of familial parkinson's disease point mutations a30p and a53t on the structural properties, aggregation, and fibrillation of human α-synuclein. *Biochemistry*, 40(38):11604–11613, 2001.

[125] Venketesh Thrithamara Ranganathan, Saman Bazmi, Stefan Wallin, Yun Liu, and Anand Yethiraj. Is ficoll a colloid or polymer? a multitechnique study of a prototypical excluded-volume macromolecular crowder. *Macromolecules*, 55:9103–9112, 10 2022.

[126] Wojciech M Śmigiel, Luca Mantovanelli, Dmitrii S Linnik, Michiel Punter, Jakob Silberberg, Limin Xiang, Ke Xu, and Bert Poolman. Protein diffusion in escherichia coli cytoplasm scales with the mass of the complexes and is location dependent. *Science Advances*, 8(31):eabo5387, 2022.

[127] EO Puchkov. Intracellular viscosity: Methods of measurement and role in metabolism. *Biochemistry (Moscow) Supplement Series A: Membrane and Cell Biology*, 7:270–279, 2013.

[128] Christopher Brennen and Howard Winet. Fluid mechanics of propulsion by cilia and flagella. *Annual Review of Fluid Mechanics*, 9(1):339–398, 1977.

[129] Gaojin Li, Eric Lauga, and Arezoo M. Ardekani. Microswimming in viscoelastic fluids. *Journal of Non-Newtonian Fluid Mechanics*, 297:104655, 11 2021.

[130] Eric Lauga. Bacterial hydrodynamics. *Annual Review of Fluid Mechanics*, 48:105–130, 1 2016.

[131] Adrian H. Elcock. Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Current Opinion in Structural Biology*, 20:196–206, 4 2010.

[132] Andrea Soranno, Iwo Koenig, Madeleine B. Borgia, Hagen Hofmann, Franziska Zosel, Daniel Nettels, and Benjamin Schuler. Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proceedings of the National Academy of Sciences of the United States of America*, 111:4874–4879, 4 2014.

[133] Peter B. Crowley, Keith Brett, and Jimmy Muldoon. Nmr spectroscopy revceals cytochrome c-poly(ethylene glycol) interactions. *ChemBioChem*, 9:685–688, 3 2008.

[134] Jeetain Mittal and Robert B. Best. Dependence of protein folding stability and dynamics on the density and composition of macromolecular crowders. *Biophysical Journal*, 98:315–320, 1 2010.

[135] M. Tal, A. Silberstein, and E. Nusser. Why does coomassie brilliant blue r interact differently with different proteins? a partial answer. *Journal of Biological Chemistry*, 260:9976–9980, 1985.

[136] Ronen Shemesh, Amit Novik, and Yossi Cohen. Follow the leader: Preference for specific amino acids directly following the initial methionine in proteins of different organisms. *Genomics, Proteomics and Bioinformatics*, 8:180–189, 9 2010.

[137] Mohona Sarkar, A. E. Smith, and Gary J. Pielak. Impact of reconstituted cytosol on protein stability. *Proceedings of the National Academy of Sciences*, 110:19342–19347, 2013.