Diffusion behaviour of proteins and polyelectrolytes interrogated via a microfluidic device

Tiago Picanço Castanheira da Silva

Thesis to obtain the Master of Science Degree in

Biological Engineering

Supervisors: Doctor Marcel Ottens

Doctor Ana Margarida Nunes da Mata Pires de Azevedo

Examination Comittee

Chairperson: Professor Helena Maria Rodrigues Vasconcelos Pinheiro

Supervisor: Doctor Ana Margarida Nunes da Mata Pires de Azevedo

Member of the Committee: Doctor Pedro Carlos De Barros Fernandes

“Sometimes it's the very people who no one imagines anything of who do the things no one can imagine.”

Christopher Morcom

Acknowledgements

First of all, I would like to thank Ana Azevedo for putting me in contact with Marcel Ottens and for all the support displayed throughout these months and more specifically in the last moments of the thesis, when stress levels are high.

I would also like to thank Marcel Ottens and TU Delft for the great opportunity to develop my work and ideas together with all the group, for all the availability regarding materials and equipment and the great conditions offered for the best work I could develop. I would also like to thank Kawieta for all the help with paperwork during my time there.

I would like to thank both Marcel Ottens and Geert-Jan Witkamp for all the guidance in the work I developed and all the advices along my time in the TU Delft.

I would like to thank all the BPE group for all the great lunch moments, coffee breaks and delicious cakes whenever it was someone's birthday or just because we needed some sugar to keep our mind focused and for the feedback during the presentations and good moments of work, like the great win we achieved during Sports Day. Special thanks to Marcel's group as all of you, Silvia, Deborah, Monica, Bianca, Victor, Shima, Marcelo, Miao, Carmen, Tessa, Suk for the feedback and brainstorming during all the meetings, as I am sure I learned greatly from all of you and all your projects and you allowed me to develop my scientific thinking.

I would like to thank Marcelo, my "twin", for all the great moments in Delft and in the TU Delft. It was very enjoyable to discuss everything from science to sports and to have an "older version of me" that I could talk with and rely on and someone I can call a friend. You taught me many things about chromatography that I wasn't aware but most importantly how to think, structure my ideas and pursue my instincts. Who said that a full day in the lab needs to be boring?

Now I address my words to Miao. You were more than a guide throughout all my time in Delft. I learned a lot from you, specially concerning the efforts I had to do and pick "which fights I needed to fight", as time urged, and data was needed. You have a great scientific mind and a great will to discover the unknown and I believe that is one of your greatest assets so please keep it forever! I think that more than guidance, you gave me your friendship.

I would like to thank all the lab technicians, Stef, Max and Song for the support during my time in the university. I would like to specially thank Song for all the help with the AKTA and all the chromatography I needed to run when time was running out. You put a lot of extra hours to stand beside me and help me when I felt that nothing was working and when I thought I could not finish it. All of our long conversations about work and life in general were really productive and meant a lot to me and I am sure I learned a lot from you. Keep working as hard as you have been, but don't forget your "fotografie" my friend, we need to have pleasures in life and I know that that is yours!

I would like to thank my housemates for all the great times in The Netherlands. It is never easy to “fly away from the nest”, especially to a much colder nest! These housemates were more than that, they were true friends. Thank you, Fernando, for needing a room, because if it weren’t for you, none of us would have had a house to stay in the first few days. Thanks for sticking beside me and trying every single type of beer possible, and for the “paio” you brought from Portugal, as it was great to have a taste from home halfway through our time there. Thanks to you, we have every single trip we had documented! I would also like to thank Lenny and Ticha for all the times together in Delft. I won’t forget all those kebabs and certainly won’t forget how much I learned from you about bikes and all the times you hooked me up when I had a flat tire. Playing by your side was also great, but you need to improve your shooting to score more! It is amazing how much you can grow and how much a friendship can grow in so little time, but with you three it certainly has grown a lot!

Thank you for all my friends that helped me get through what being abroad meant. Thanks to all my friends that despite being away were always close. Thank you all for the great moments throughout university and for the great journey it was to have studied in Técnico for these past five years. I am certain that we all grew with this experience and will continue to do so, together and for fifty more years!

For the ones that were close to me physically, I would like to thank all of you or the great moments. I will not forget all the great moments we had all around The Netherlands and moments like King’s Night, King’s Day and Ascension Day will be forever engraved in my memory. Thank you Grilo, Raquel and Rita for putting up with me during these months.

I would now like to thank my family and dearest friends. Diva (should I say Catarina?) and Rodrigo, you know the importance you’ve had along these five years and I am certain you know that I hold you in the highest standards. I have some friends I can call family, but to you both, I can call brother and sister. I love you from the bottom of my heart and I know, despite all those disagreements, you know that I mean it.

To my grandparents, whom I love and always showed support in this journey, I would like to thank not only for these last five years but for the last twenty-two and much more to come. Let’s continue to eat “lampreia” every year!

To my father, who once said that “parents are the bone in which children sharpen their teeth”. I couldn’t agree more and that is an essential part of growing and I believe that you helped me to shape the person that I become, with great values. From you I learned two of my biggest values that are based on the word “brio”, to always be proud of what you have done and to thrive to do more because it is always possible to do more and “respect”, to always respect the next and face the consequences of my decisions. Thank you for allowing us to have Gullit and if the beginning was rough, I am pretty certain that nowadays it would be easier for you to put me outside the house than what you would do with him!

Finally, to my sister and Gonçalo. With both of you I can certainly say I improved my emotional intelligence and that showed to be of great importance so far and I am sure it will continue to be for many years to come. Nez, you were the greatest sibling one could ask for. You got me excited about science when I was so little, and helped me in my studies (but only the biology, as the maths I leaned myself). You showed me the importance of a great work ethics and that it is important to stand by what you believe, knowing the limitations of when you are being stubborn. The moments we shared have been amazing, but I have a feeling that there are many more to come. We grew a lot by each other's' side and I know that we will continue to do so. You have been much more than a sister, and I know you know that.

Thank you all

Abstract

The diffusion coefficient (*D*) of different proteins and dextran sulphate sodium salt (DSS) was determined by using a microfluidics H-cell. The mass transfer was determined with an inline measurement by a spectrophotometer (proteins) or a conductivity meter (polyelectrolyte) and *D* was calculated using COMSOL and Matlab, making use of the mass transfer between two streams flowing in laminar flow that had different solute concentration and considering the channel's dimensions and residence time in the microchannel. The *D* value of lysozyme and cytochrome c was determined for different experimental conditions, having compared the values and trends with previous literature and models. The *D* value was also determined for six different proteins. It was found that the increase of the ionic strength was inversely proportional to the *D* value of lysozyme, varying from $D = 4.40 \pm 0.17 \times 10^{-10}\, m^2/s$ to a value between 1.2×10^{-10} and $1.5 \times 10^{-10}\, m^2/s$ for higher ionic strength values. The variation with the solvent's viscosity and with the increase of the proteins' molecular weight followed the trend predicted by the Stokes-Einstein correlation suggesting that it is valuable for predicting a trend but not for precise value for the studied molecules. Lysozyme and cytochrome c showed an increase of *D* value with increase of protein concentration suggesting that there are predominant repulsive interactions between the proteins. The *D* value of DSS showed no significant trend or variation with the increase of solute concentration. It is demonstrated that the H-cell is a high-throughput method, being accurate and precise, with flexibility in the detection method.

Keywords: diffusion coefficient, H-cell, microfluidics, protein, polyelectrolyte

Resumo

O coeficiente de difusão (*D*) de diferentes proteínas e de sulfato de dextrano sal sódio (DSS) foi determinado usando uma H-cell de microfluídica. A transferência de massa determinou-se com medição em linha por um espectrofotómetro (proteínas) ou um condutivímetro (polielectrólito) e *D* foi calculando usando os *softwares* COMSOL e MatLab, aproveitando a transferência de massa entre as duas correntes em fluido laminar, com diferentes concentrações de soluto e considerando as dimensões e o tempo de residência no microcanal. Determinou-se o *D* de lisozima e citocromo c sob diferentes condições experimentais e compararam-se os resultados e tendências com literatura e modelos antigos determinando-se ainda *D* de seis diferentes proteínas. Verificou-se que um aumento da força iónica é inversamente proporcional ao *D* da lisozima, variando de $D = 4.40 \pm 0.17 \times 10^{-10}\, m^2/s$ para um valor entre 1.2×10^{-10} e $1.5 \times 10^{-10}\, m^2/s$ para uma força iónica maior. A variação de *D* com a viscosidade do solvente e o peso molecular das proteínas seguiu a tendência prevista pela correlação de Stokes-Einstein, provando ser uma boa correlação para prever tendências, mas não para valores precisos. Verificou-se um aumento de *D* com o aumento da concentração de lisozima e citocromo c sugerindo que as forças repulsivas predominavam nas soluções testadas. O *D* de DSS não variou de maneira significativa nem apresentou nenhuma tendência clara com o aumento da concentração deste nas soluções testadas. Demonstrou-se que a H-cell é um método confiável e preciso, com flexibilidade no método de deteção.

Palavras-chave: coeficiente de difusão, H-cell, microfluídica, proteína, polielectrólito

Index

List of Abbreviations

DSS – Dextran Sulphate Sodium Salt

J – Mass flux

D – Diffusion Coefficient

C – Concentration

t – Time

k_B – Boltzmann constant

T – Temperature

f – Friction Coefficient

R_H – Hydrodynamic radius

M_W – Molecular Weight

R_G – Radius of Gyration

pI – Isoelectric point

pKa – Acid Dissociation Constant

$D_{w,20}$ – Diffusion coefficient under standard conditions (water, 20ºC)

D_0 – Diffusion coefficient at infinite dilution

BSA – Bovine Serum Albumin

PEC – Polyelectrolyte Complexes

SIC – Self-Interaction Chromatography

AB – Acetate Buffer

PB – Phosphate Buffer

CV – Column Volumes

I – Ionic Strength

List of tables

List of figures

1. Introduction

1.1 Microfluidics

Microfluidics has been labelled as the fluid manipulation within channels, with dimensions that can be in the order of tens or hundreds micrometres (Whitesides, 2006). Such as these microchannels, many other components have been scaled down, bringing many advantages to its mass use. Looking back, it is possible to see big differences to nowadays' common day objects, such as a smartphone (that has the processing ability of a computer) when compared to the first computer, which needed a whole room for itself and, even then, had much weaker processing capacity.

Over the years there has been an evolution towards the miniaturization of several processes. Smaller devices have many advantages, as they can be more easily transported and require less sample which for the case of valuable samples allows researchers to save money. The miniaturization is conjugated with an increasingly higher computational power, provided by the development of technology in general.

The development of the several sciences led to the manufacturing of smaller and more capable devices. It is of high importance to have a high throughput method which is at the same time reliable, fast and with small sample consumption. This is the gap that microfluidics fills, as it provides several devices with different purposes, but all rely in the same rational: to have a fluid system at a microscale, allowing the full control of the flows and the process, all inside a single chip, the so-called Lab-on-a-Chip.

1.1.1 First applications and development

Initially, microfluidics was used in molecular analysis (Whitesides, 2006), (Stone, Stroock, & Ajdari, 2004). The miniaturization of processes and the need to lower analysis costs were one of the driving forces for the development of microfluidics in this field. It is also important to keep in mind that, as time passed, better fabrication technologies emerged and therefore better microfluidic devices emerged. The analysis using microfluidics showed great compromise between sensitivity, resolution and amount of sample used and the microscale became more widespread (Whitesides, 2006).

The increasing knowledge on microfluidics helped in the development of the field and the range of materials used. If initially the already well-known materials were used, such as glass and silicon, later materials were developed in order to fulfil the needs of the field of studies (Whitesides, 2006), (Whitesides, 2002). Poly(dimethylsiloxane) (PDMS) is used in microfluidics for a wide range of studies (e.g. (Silva, et al., 2017), (Ebara, Hoffman, Hoffman, & Stayton, 2006)) and the manufacturing techniques are very precise and easy to perform if the right equipment is used. The only drawback with this material, as with other materials, is the need to have a suitable material for a certain purpose. PDMS may perform better when the aim is to have a functionalized membrane than when the objective is to have a stiff material or a thermal conductor such as steel.

Microfluidics has really evolved from its initial applications. With the development of microfabrication techniques there was also a development of the new layouts and structures of the channels. It evolved from having a channel structure to having valves, mixers, chambers and all sorts of different channels or a combination of all above mentioned in one single device (Whitesides, 2006).

The microfluidics chip is meant to serve a specific purpose. Therefore most of these chips are connected to other devices, either to make the fluid input and output onto the chip, like fluid pumps, or to make the detection of the phenomena that the chip is used to study, through the use of detection devices (Whitesides, 2006). This versatility and ability to interchange between detection methods is one key aspect of what the development of the microfluidics field has become and can achieve in the near future. Nowadays it is very important to have reliable and versatile devices as well as to have an adaptable device to the unique needs.

With all these developments, microfluidics presents itself as a technique with several advantages when compared to others because it has high throughput, it is time saving, requires less sample and presents lower costs (Nguyen & Wu, 2004).

1.1.2 Laminar Flow in Microfluidics

The advantages of microfluidics do not rely solely on data collection using a small amount of resources and time. The control of the fluid dynamics is a very important aspect of this area, as the use of laminar flows allows to have a bigger control of molecules' concentration in space and time (Whitesides, 2006).

This is possible due to the use of a laminar flow. Laminar flow is characteristic of a fluid with a Reynolds number, Re, lower than 1 800. Re is a dimensionless number used to study fluid mechanics and different fluid flows (Bergman & Incropera, 2011). The use of laminar flow in microfluidics helps to avoid the convective mixing of the inlet fluids and, therefore, reassures that the molecular transport from one fluid to the other doesn't occur by convective lateral mixing but by molecular diffusion through the fluids' interface (Kamholz & Yager, 2001), which is one of the main objectives of the present work.

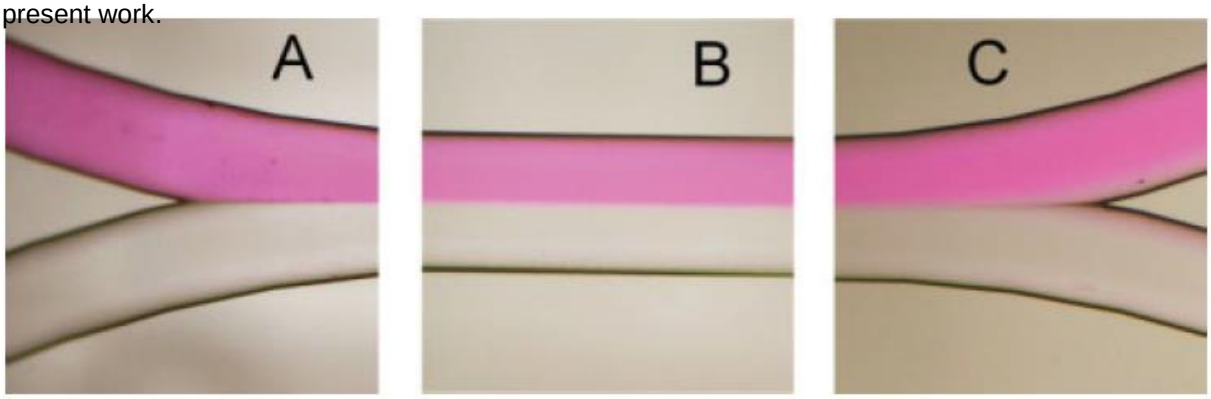

Figure 1 – Two different aqueous streams in a microfluidics chip. (A) represents the point of the joining of the two streams in the inlet, (B) a middle point along the channel and (C) the separation of the two streams, close to the outlet. In this figure we can see that the laminar flow avoids the convective transport of the colouring compound rhodamine B as we can see that in a middle point of the channel we can see a clear difference between the colour of both streams. However, when looking at (C) one can see that some very tenuous diffusion happened, as is possible to see in the transparent stream a small amount of pink ink close to where the two streams separate. Image adapted from *(van Leeuwen & et al, 2009)*.

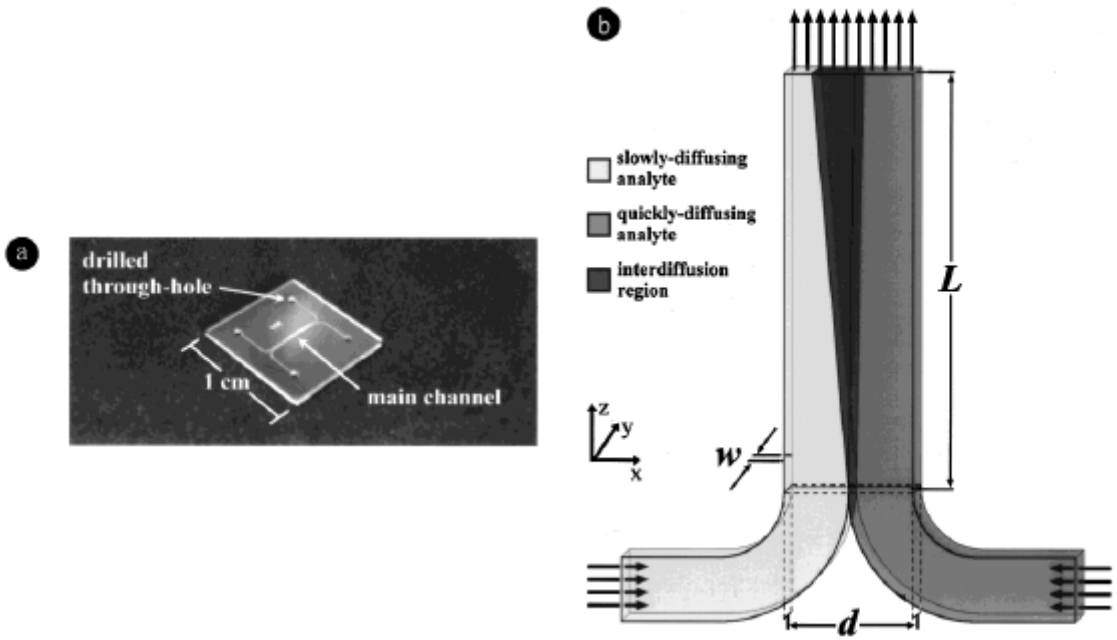

Figure 2– Flow of two different aqueous streams in a microfluidics device. In this figure is showed a scenario where the inlet molecules have different diffusivities. In (b), the light grey inlet shows a slow diffusing analyte and the dark grey shows a quickly diffusing analyte. As diffusion happens through the interface, the fast diffusing analyte will diffuse onto the stream that has the slow diffusing analyte. This is shown by the black region, where one can see a slight diffusion onto the dark grey stream but a very high diffusion onto the light grey stream. Adapted from *(Kamholz, Weigl, Finlayson, & Yager, 1999)*.

Figure 1 and Figure 2 show good examples of what is the laminar flow in microfluidics. As previously stated, the laminar flow keeps both streams flowing independently and avoids convective movement from the molecules in each stream, as seen in Figure 1. In the case of no diffusing or very low diffusing molecules in the inlets, it is expected to have a scenario similar to the one on Figure 1 whereas if the molecules are diffusing faster than it is expected to have a scenario like Figure 2.

1.2 Mass Diffusion

Diffusion of molecules is a random molecular motion of molecules and as acknowledged by Fick, “diffusion is a dynamic molecular process” (Cussler, 2009). Translational diffusion is, at the same time, a random and dynamic process. Fick stated that the flux of molecules and, therefore, the diffusion of these molecules is directly proportional to the concentration difference of two fluids (Mehrer & Stolwijk, 2009), in the case where two different solutions are put together.

Scottish botanist Robert Brown also presented a theory for the diffusion of molecules. According to Brown, molecules move in a random way along spatial coordinates even when there’s no influence of streams in the fluid they’re dissolved in. If the fluid is still, at a macro and microscopic scale, there is still a random movement of the molecules, inherent to their own properties. This phenomenon was later baptized as Brownian Motion (Mehrer & Stolwijk, 2009).

The first equations for mass diffusivity were derived from previous equations on heat diffusivity. Fick adapted Fourier's equations for heat diffusion to the mass diffusion phenomena, reaching the Fick's first law of diffusion:

$$J = -D\frac{\delta C}{\delta x} \tag{1.1}$$

As stated before, this first equation considers that diffusion occurs due to a difference in concentration between two regions, where J represents the mass flux, *D* represents the diffusion coefficient, C represents the molecule's concentration and x the space coordinate where diffusion is occurring. The negative sign in the right side of the equation is due to the fact that the direction of diffusion is opposite to the one of increasing concentration (Sun, 2004), (Cussler, 2009).

This equation then evolved into understanding diffusion phenomena with time. This lead to what is called second Fick's Law and is given by equation (1.2): (Cussler, 2009), (Sun, 2004)

$$\frac{\delta C}{\delta t} = D\frac{\delta^2 C}{\delta x^2} \tag{1.2}$$

Later, Einstein developed a relation for diffusion, based on the principle of Brownian motion. This equation was then adapted to what is known as the Stokes-Einstein correlation for the diffusion of particles in a solvent with low Reynolds number. This correlation is used for the estimation of the diffusion coefficient, D, of large molecules when these molecules are much bigger than the solvent. It takes into account the Stoke's law for the drag coefficient.

$$D = \frac{k_B T}{f} = \frac{k_B T}{6\pi\eta R_0} \tag{1.3}$$

In equation (1.3), *f* represents the friction coefficient of the solute and R_0 the solute radius whereas k_B is the Boltzmann constant and η is the solvent's viscosity. Despite this being a good correlation, research showed that only for about twenty percent of the cases it can accurately estimate the exact value of the diffusion coefficient. Nonetheless, it can still contribute to understand which parameters may influence the most the diffusion coefficient (Cussler, 2009).

To be able to use this correlation one needs to assume that proteins are hard spheres. While this assumption may not be completely wrong to some protein (spherical proteins in solution may act like hard spheres), for others it may not be adequate. An adaptation to the equation (1.3) was suggested, and the solute radius was taken as hydrodynamic radius, R_H, for the case of proteins. Besides this small adaptation, more equations were proposed in order to have prediction models for the diffusion based on different criteria.

1.2.1 Diffusion Coefficient Prediction

Along the years, many have tried to reach an equation or correlation to accurately predict the diffusion coefficient of molecules, for example proteins were very studied for this specific matter. In many cases, it is important to have models for the estimation of parameters in order to save the

determination of these given parameters using slow methods. In industry this becomes more important as different models for different unit operations rely on these parameters and therefore raises the need to have prediction methods. However, these methods are only relevant if they can accurately predict the parameter's value, as one doesn't want to have a parameter with a wrong value.

Table 1 - Different correlations for the estimation of the Diffusion coefficient of proteins. Equations adapted from *(He, 2003)*.

Based On		Equation	Parameters
Molecular Weight		$D = 8.34 \times 10^{-12} \cdot \left(\frac{T}{\eta \cdot M_w^{\frac{1}{3}}} \right)$ (1.4)	D – Diffusion Coefficient ($m^2 s^{-1}$) T – Temperature (K) η – solvent viscosity (cP) M_w – Molecular Weight (g mol^{-1})
Radius of Gyration		$D = 5.78 \times 10^{-12} \cdot \left(\frac{T}{\eta \cdot R_G} \right)$ (1.5)	D – Diffusion Coefficient ($m^2 s^{-1}$) T – Temperature (K) η – solvent viscosity (cP) R_G – Radius of Gyration (Å)
Geometry	Cylindrical	$D = \frac{1.69 \cdot 10^{-9}}{0.1 \cdot L}$ (1.6)	D – Diffusion Coefficient ($m^2 s^{-1}$) L – Length of rod-like proteins (Å)
	All shapes	$D = \frac{6.85 \cdot 10^{-15} \cdot T}{\eta \cdot \sqrt{M_w^{\frac{1}{3}} \cdot R_G}}$ (1.7)	D – Diffusion Coefficient ($m^2 s^{-1}$) T – Temperature (K) M_w – Molecular Weight (kg $kmol^{-1}$) R_G – Radius of Gyration (Å) η – solvent viscosity (cP)

When the shape of the protein draws away from the simplest shapes (spherical shape), then it becomes hard to have an accurate model for the prediction of the Diffusion using simpler correlations like (1.3) or even (1.4). Although equation (1.6) already has some lights on different shapes, like the cylindrical shape, this means that for one to be able to use this equation, the protein shape must be known beforehand, and we need to keep in mind that not all non-spherical proteins are cylindrical, in fact many have all kinds of shape, that can go from cylindrical or rod-like shape to ellipsoidal. Niemeyer and He showed that the shape of the proteins has a big influence on the hydrodynamic radius, as one could expect. So, it is understandable why this correlation, although it is an improvement from previous ones, is not good enough (He, 2003).

Trying to improve a correlation that could fit a protein's shape, Niemeyer and He tried to develop a model that took into account almost all available parameters for proteins. The authors state that the ratio between R_H and R_0 (where R_0 is defined as the radius of a sphere of volume identical to the volume of the studied protein) is constant for proteins and changes depending on their shape. Further, it is explained that the use of the M_w is handy, as this parameter is well known for the proteins and that

$M_w^{1/3}$ is proportional to R_0 and so they were able to use both of these parameters in the same correlation. A more detailed explanation is given in (He, 2003).

The drawback of these equations is the empirical parameters all use, based on experimental trials. Although it is necessary to have empirical parameters, when talking about diffusion coefficient and proteins, there can be many combinations in which these correlations don't apply. Diffusion coefficient depends on several parameters, being one of the most important the solute itself, but other parameters also have a very big influence and the empirical values may not have been taken under the conditions we want to know the *D* value.

1.3 Diffusion Coefficient Determination Methods

Diffusion coefficients can be estimated using a variety of different methods. Contrary to the early scenario, when the first *D* values were estimated experimentally, current technology allows for accurate and reliable methods, with data being generated in a short amount of time when compared to the very first methods (Cussler, 2009).

Estimation of *D* values is very important because it can help to avoid experimental work on the determination of it. However, and taking into account what was previously discussed, *D* values can change greatly with different conditions. This is why experimental data is so important for both research and industry (Cussler, 2009).

In Table 2, it is possible to compare different methods used for the determination of the diffusion coefficients. There are other methods besides the ones listed on Table 2 that were left out for different reasons. Some were left out because they are not so commonly used and others, like the steady-state methods because it includes the microfluidics device.

Cussler refers to the steady state methods as a "nightmare" that have complicated requirements and consume a huge amount of sample (Cussler, 2009). While that could be true for first generation steady state methods, an extensive explanation on the advantages of microfluidics was given and the problems raised by Cussler can easily be overcome (Nguyen & Wu, 2004).

Table 2 - Summary of the most used methods for the determination of Diffusion coefficients. Several aspects are tabled for the mainly used methods, both in the past and present, like the nature (underlaying mechanism) of the diffusion, the cost and easiness to assemble the required equipment, if it is necessary or not to have a concentration difference, how long does it take to collect the data and the method's accuracy. Data adapted from *(Cussler, 2009)* except for (*) which was taken from *(Pecora, 2013)*.

	Nature of Diffusion	**Equipment (Cost; Assembly)**	**Concentration Difference Required**	**Duration of Measurements**	**Accuracy (%)**
Diaphragm Cell	Pseudosteady state	Small; Easy	Big	Days	98
Taylor Dispersion	Decay of a Pulse	Moderate; Moderate	Moderate	Days	99
NMR	Decay of a Pulse	Big; Hard	Not Required	Hours	95
DLS	Decay of a Pulse	Big; Hard	Not Required	Hours	90(*)
Gouy Interferometry	Unsteady in an infinite cell	Big; Moderate	Small	Hours	99,9
Capillary	Unsteady out of finite cell	Small; Easy	Moderate	Days	99

1.4 Macromolecules Diffusion

Diffusion is often a rate limiting step in several industries and industrial processes, as it is a very slow process. Diffusion in gases is very fast when compared to liquids and both gas and liquid diffusion are several orders of magnitude bigger than solid diffusion. The main interest in this parameter is in liquid diffusion, as almost every industrial process for the chemical and biochemical industries are done in liquids and, according to Cussler, “the sloth characteristic liquid diffusion” is what defines it as a rate limiting step (Cussler, 2009).

Diffusion coefficients are a property of a given molecule as each macromolecule (whether it is a protein, polyelectrolyte, among others) has its own diffusion coefficient value. However, it is important to highlight that despite being an inherent property of these molecules it may vary depending on the conditions and environment in which the protein is.

1.4.1 Protein Diffusion

As stated above, the knowledge behind protein diffusion is very important. With the rising of the biotechnology field, biotechnology industries are growing and producing more and more bioproducts for the everyday use.

Niemeyer and He briefly summarized why is the data on this subject so important for the current industries is mainly in “process design and analysis” (He, 2003). The ability to design a process and create a model that describes an industrial process is of the utmost importance for the industries and is often necessary to have data on different parameters. Diffusion coefficient is a parameter often used in mass transfer models, that can be implemented in a model for a certain chromatography or mass transport models.

Proteins are complex molecules that have subunits called amino acids, connected to each other through covalent bonds, that in their hand also have subunits. Proteins have all different kind of shapes, sizes and properties (Sun, 2004). All these aspects have an influence in several parameters, one of them being the *D* value for each protein.

The complexity of these molecules makes it not only hard to predict what happens when they’re put in different media but also to predict their properties in different media. As stated in 1.4, diffusion coefficient is not independent of the conditions in which the macromolecules are put into and, despite not having a significant change in terms of increasing/decreasing in several orders of magnitude, the diffusion coefficient can be highly influenced by external influence. This raises the need to have a fast and reliable way to determine how *D* is affected by different parameters.

It is not always easy to study the influence of different parameters on diffusion coefficient because when changing a given parameter, more than one condition can be altered. This is the case when, for instance, one wants to study the influence of the pH but wants to change the pH in many pH values. Most of the times it is not possible to have the same buffer with very different pH values, as phosphate buffer, for example, ranges between 5.8 and 8 pH values. If one wants to study the proteins on a pH

lower than 5.8, one would have to change the species in solution, preparing another buffer. In this new buffer, one can not only have the influence of the pH but also the influence of different ions in solution.

1.4.1.1 Effect of the pH

Proteins, like many other molecules, have a characteristic pH value in solution at which their net charge is neutral. This parameter is characteristic from each protein and represents the solution pH at which the residues that are positively charged are in the same amount as the negatively charged residues, making the net charge of the protein zero. This is the isoelectric point of a protein, pI. The net charge of a protein can be determined by calculating the pK values of the ionizable groups at a given pH (Tanford, 1963).

When pH has the same value the as pI, the net charge of the protein is neutral whereas as if the pH is higher or lower than the pI value, the net charge of the protein is negative or positive, respectively. Working at a pH close to the pI means that the protein is processed in such conditions where it has a net charge close or equal to zero (depending if the pH is close or equal to the pI value). It is important to keep in mind that, despite the net charge of the protein being zero, the protein still has charged residues, but in such a way that they cancel each other out (Tanford, 1963).

Having a neutral net charge doesn't necessarily mean higher stability or solubility in solution. In fact, different authors have stated that a protein is less soluble closer to the pI value (Kakalis & Regenstein, 1986), (De Wit, 1989), (Wong, Camirand, Pavlath, Parris, & Friedman, 1996). The pH will, therefore, affect the surface charge of the proteins. This will imply that the protein will be more or less charged (either negative or positively charged, depending on the pI value) according to how far the solution's pH is from the pI, increasing the electrostatic interactions between proteins. Pelegrine and Gasparetto state that the closer proteins are in solution to the pI value, the less electrostatic interactions will occur (Pelegrine & Gasparetto, 2005). So, the furthest to this value, more electrostatic interactions will occur.

1.4.1.2 Effect of the protein concentration

In the case of a very dilute solution, or even infinite dilution, the solute only "sees" the solvent around itself. This means that the solute will only interact with the solvent and the *D* will only be influenced by a solute-solvent interaction (Cussler, 2009). However, when the concentration of the solute increases, it starts "seeing" other solute molecules and interacting with them. It goes from an exclusive solute-solvent interaction to a mix of interactions between solute-solute and solute-solvent.

It is therefore expected that an increase in concentration has an effect in the effective diffusion coefficient value of a molecule. The increase of concentration will also increase the amount of solute-solute interactions occurring, that can be hydrophobic or electrostatic interactions (depending on the molecules, one can have simply electrostatic interactions, if we are in the presence of a purely electrolyte material like salts, or a combination of hydrophobic, van der Waals and electrostatic, if we are in the presence of a protein).

If a given solute has attractive interactions with itself, this will mean that it will diffuse less on the solvent because it will tend to keep close to itself in solution. Aggregation of molecules is a strong

attractive interaction that can become irreversible. Precipitation is an aggregation that turns soluble solutes into solids in suspension. It occurs due to a decrease in solubility of the solute triggered by adding a precipitating agent, by changing buffer conditions or the temperature of the solution or by adding more solute/evaporating the solvent.

On the other hand, when a solute has repulsive interactions with itself, this will mean that it will diffuse more in solution because when it interacts with itself, the tendency will be for both molecules to "push" each other away.

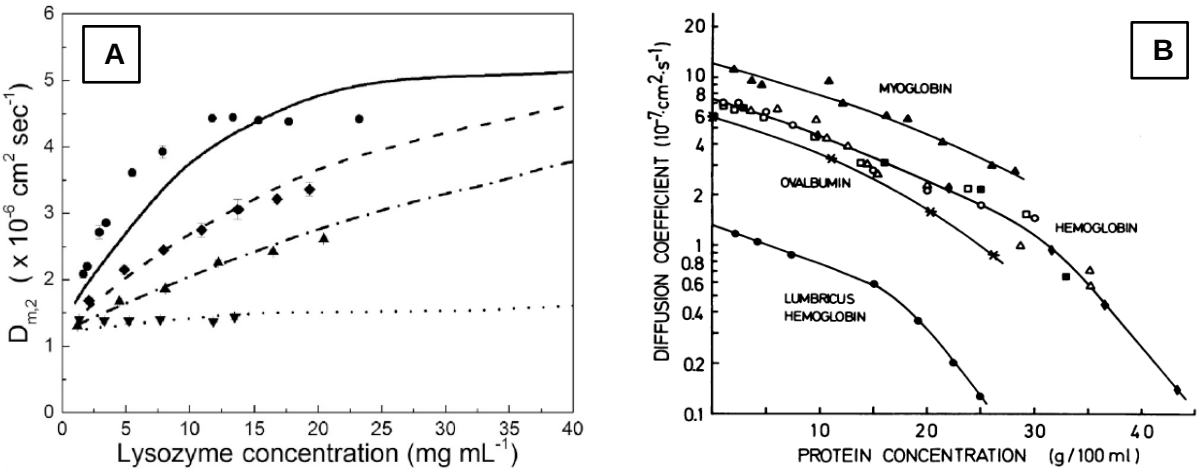

Figure 3 - Diffusion Coefficient with increasing concentration of protein. (A) shows *D* values at different solute (lysozyme) concentration in different experimental conditions. The circles data points were obtained in a 2 mM KCl solution, the diamonds in a 5 mM KCl solution, the triangles in a 10 mM KCl solution and the inverted triangles in a 90 mM KCl solution, all at pH 6.0. (B) shows *D* values for different proteins under different conditions with increasing protein concentration. (A) and (B) were adapted from *(Sorret, DeWinter, Schwartz, & Randolph, 2016)* and *(Gros, 1978)*, respectively.

In Figure 3 it is possible to observe two distinct types of behaviour. In the graph on the left it is possible to see that the diffusion coefficient of lysozyme increases with a concentration increase whereas the opposite happens for the other proteins. This shows that different molecules show different behaviours and therefore the diffusion coefficient varies differently. To better understand why *D* increases or decreases it is important to understand under which conditions the experiments were performed and understand what kind of interactions are predominant in each situation. We may have a predominance of electrostatic repulsive interactions when *D* increases with concentration, making it clear that an increase in concentration will promote more repulsion between the molecules and consequently more diffusion.

1.4.1.3 Effect of the solvent

In section 1.4.1 it was stated that *D* values can vary with different experimental conditions in which the measurements are performed. This is something to keep in mind, as it is important to raise awareness on this topic.

Some authors, for instance, have performed diffusion trials to determine the diffusion coefficient of a given protein under a specific set of conditions varying solely its concentration. In these trials, the

proteins were solubilized in a specific buffer and then the desired concentration was achieved. Some of the authors then used this knowledge to estimate the *D* value at infinite dilution under standard conditions ($D_{w,20}$), only correcting for the difference in the buffer's viscosity (Dubin, Clark, & Benedek, 1971). Although this argument is not fully wrong, as one can try to extrapolate the diffusion coefficient at infinite dilution (D_0) from several concentration trials, as long as the data is in good accordance, it is hard to make the "jump" from infinite dilution in buffer and in water.

Guy *et al.* showed that a change from 0 M to 0.01 M KCl buffers can have a significant impact on the diffusion coefficient of lysozyme. It is hypothesized that the ions in solution may have a shielding effect on the surface charges of the protein in study. These ions will prevent the protein surface charges from interacting with each other causing a decrease of D, as can be observed in Figure 4 (Cadman, Fleming, & Guy, 1982).

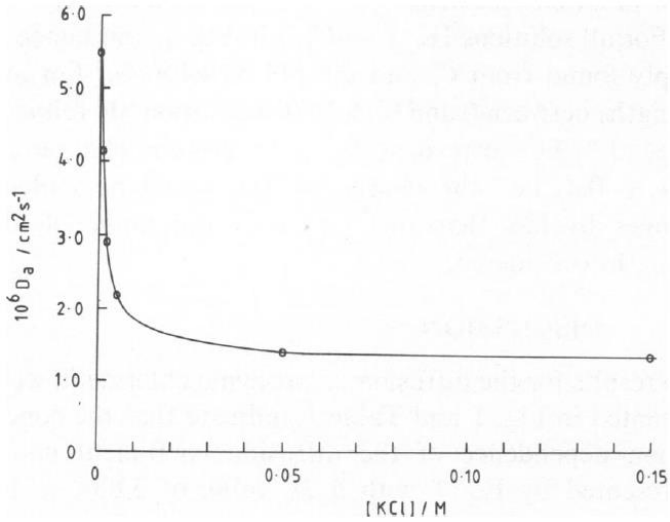

Figure 4 - Effect of KCl concentration in the diffusion coefficient of lysozyme. Adapted from *(Cadman, Fleming, & Guy, 1982)*.

Not only there is a difference in having or not a salt in solution, but there is also a difference in having different types of salts. Different salts have a different electrostatic interaction with the protein in solution and will affect the diffusion coefficient, as shown in Figure 5 (Medda, Monduzzi, & Salis, 2015).

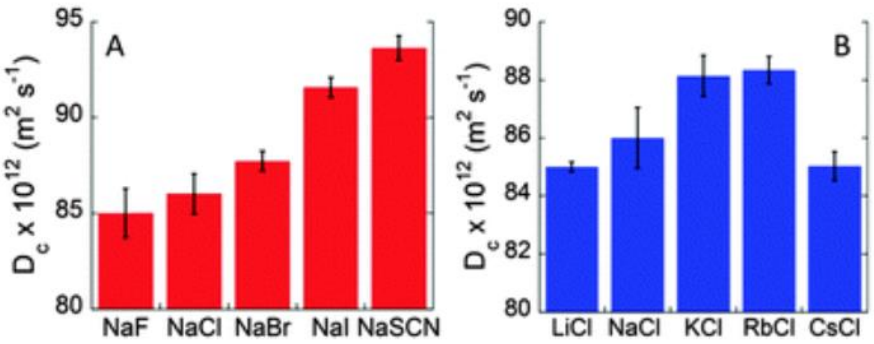

Figure 5 - Variation of the diffusion coefficient of BSA in a 40 mg/mL solution. The measurements were performed in a 10 mM phosphate buffer solution at pH 7, with different salts at 100 mM concentration. (A) shows the effect of differing anions, as sodium is used in all solutions and (B) shows the effect of different cations, as chloride is used in all solutions. Adapted from *(Medda, Monduzzi, & Salis, 2015)*.

1.4.1.4 Effect of solvent viscosity

Looking at equation (1.3), it is possible to see that, for hard sphere molecules diffusing through a liquid, the higher the viscosity, the lower the *D* will be, as also exemplified by Figure 6. This happens because there is an increase in the drag coefficient and, therefore, it will be harder for the molecules to move through the liquid. Mathematically speaking one can see that *D* varies in an inversely proportional way to the viscosity, η (Cussler, 2009).

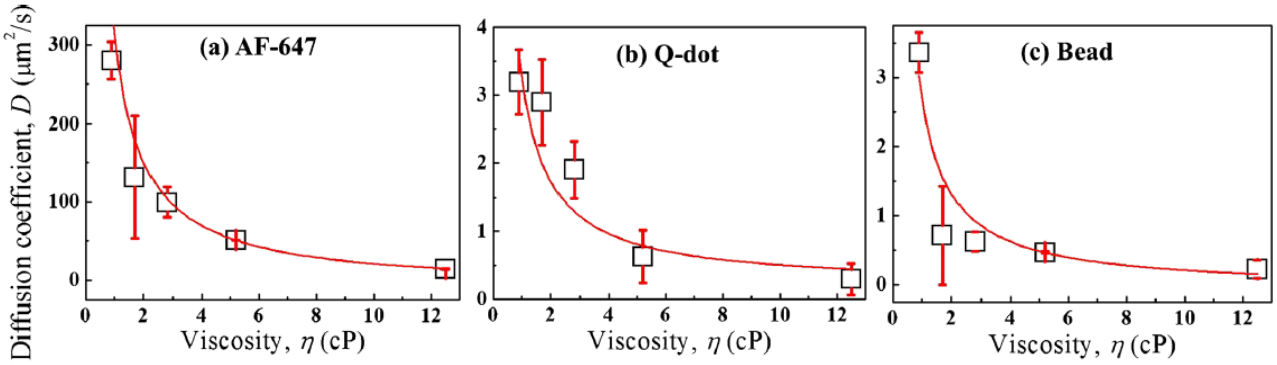

Figure 6 - Diffusion coefficient of three different fluorescent proteins with increasing medium viscosity. The three substances are (a) AF-647, (b) quantum dot, (c) fluorescent beads. It is visible that *D* decreases inversly proportional to the increase in medium viscosity. Adapted from *(Jung, Lee, Kang, & Kim, 2014)*.

This observation is typical in presence of stable solutes like proteins and electrolytes. In the presence of unstable solutes, like sodium dodecyl sulphate (SDS), a different phenomenon can be seen. Weinheimer et al. measured the diffusion coefficient of SDS solutions at increasing solute concentration and concomitant increasing viscosities. It is possible to see in Figure 7 that the concentration increase of the solute (and medium's viscosity) did not cause a decrease but an increase in *D* (Weinheimer, Evans, & Cussler, 1981). The authors suggest that this unstable molecule creates micelles in solution and, therefore, different interactions occur in solution. We need to consider that in the given study, the factor to consider is not only the medium's viscosity change, but also the concentration of the solute. The increase of *D* in this specific situation is proposed by the authors to be the result of aggregation and electrostatic interaction.

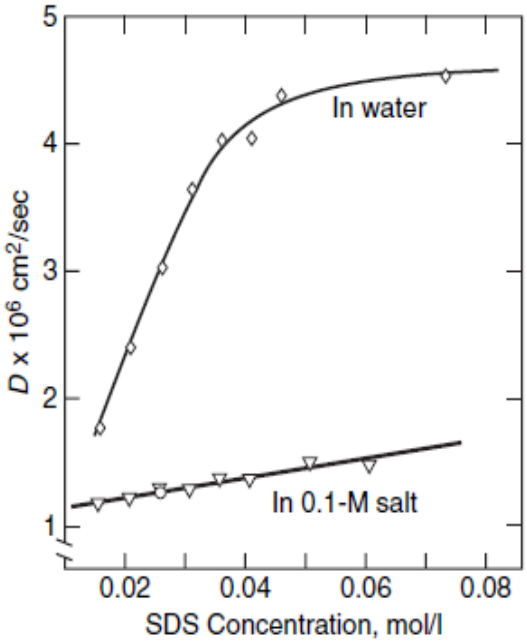

Figure 7 – Diffusion coefficient of SDS with the increase of concentration. It is possible to observe that *D* increase with the increase of solute's concentration. This increase in concentration is also responsible for the increase of the solution's viscosity. Adapted from *(Cussler, 2009)*, *(Weinheimer, Evans, & Cussler, 1981)*.

1.4.1.5 Effect of the Temperature

Once again, by analysing equation (1.3), it is possible to see that *D* will increase proportionally to the increase of temperature. In the case of proteins, it is well known that they are only stable at a given temperature range, as they denature above the maximum temperature for stability or below the minimum temperature for stability.

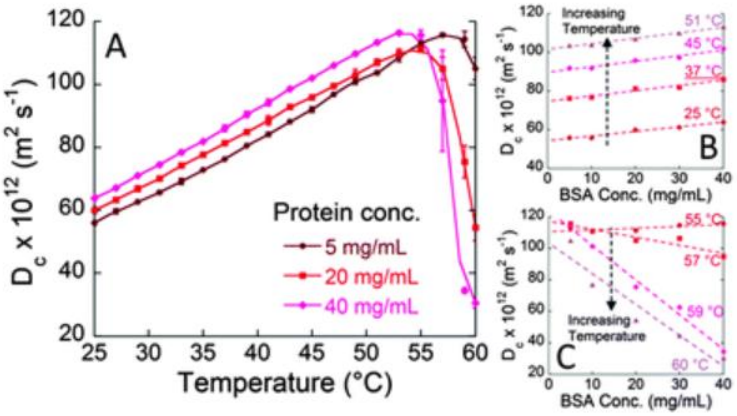

Figure 8 - Diffusion coefficient determination of BSA with the increase of temperature, in 10 mM Phophate Buffer at pH 7 and 100 mM NaCl, at different protein concentrations. Adapted from *(Medda, Monduzzi, & Salis, 2015)*.

In Figure 8 it is possible to see that there is an increase of *D* with an increase of the temperature. However, when temperature reaches around 50ºC, there is a sharp decrease of *D*. This happens because proteins do have a stability range. Above a given maximum temperature the decrease on the *D* value is explained due to the unfolding of the protein and the consequent aggregation (Medda, Monduzzi, & Salis, 2015). As mentioned in section 1.4.1.2, protein aggregation will lead to larger bulks of BSA molecules all together, increasing the size of this bulks making it harder for the protein to diffuse.

1.4.1.6 Model protein

For the majority of the studies present in this work, egg white lysozyme was used. Lysozyme has been studied for a long time now, and is really very well characterized in literature. Chicken egg lysozyme became very popular and still is nowadays due to in-depth knowledge on it, the result of an investigation of many years, performed by different researchers. This protein has a simple structure, as it is composed by a single polypeptide chain, with a 14.3 kDa molecular weight and an isoelectric point of 11.3. This protein is readily soluble in a variety of aqueous solutions (Wetter & Deutsch, 1951).

It also has a somewhat spherical shape, which is not the case of other proteins, for instance immunoglobulins. This is useful because it presents itself as the simplest case for diffusion of proteins, as the protein can be considered as a hard sphere. All of these characteristics conjugated with the fact that egg white lysozyme is an inexpensive material, when compared to other proteins, and a very well-studied protein make it a good model protein.

1.4.2 Polyelectrolyte Diffusion

Like stated in section 1.4.1, polyelectrolyte diffusion coefficients are also important for a variety of process design. Proteins can, themselves, be considered polyelectrolytes and are indeed just that. However, we are considering polyelectrolytes to be a polymer with *n* identical repeating units that have an electrolyte group and are much less complex than proteins (Du, Dai, Liu, & Dankovich, 2006).

Polyelectrolytes have been emerging as important molecules in the biomedical field, as some present themselves as biocompatible molecules, being non-toxic and well-tolerated by the human body. Polyelectrolytes can form complexes called Polyelectrolyte complexes (PEC). These complexes are made by mixing two polyelectrolytes with opposite charges, causing strong but reversible electrostatic interactions (Il'ina & Varlamov, 2005). This is one of the advantages of using polyelectrolytes, when compared to other synthetic polymers. Furthermore, as synthetic polymers rely on the cross-linking of molecules, if unreacted functional groups are not completely blocked, they could cause side-effects and consequently several verification steps are required for the use of these compounds (Berger, et al., 2004).

To use polyelectrolytes and proteins coupled in the same system there is the need to better understand polyelectrolytes themselves. Many polyelectrolytes have been widely studied throughout

the times, however, there is a lack of studies on some of the properties of these polyelectrolytes, and many times diffusion data lacks (Cussler, 2009). O-SO_3Na

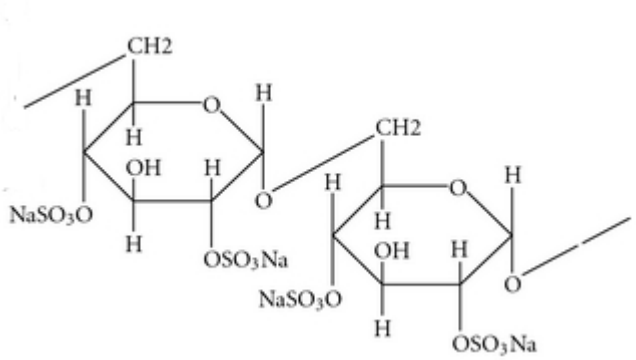

Figure 9 - Dextran Sulphate Sodium Salt.

Looking at Figure 9 it is possible to see a polyelectrolyte composed by dextran with sulphate groups covalently bound to the carbon where the -OH groups would be in the glucose molecules. Dextran Sulphate Sodium Salt (DSS) was the polyelectrolyte used for the studies in the present work.

1.4.2.1 Effect of the different parameters in polyelectrolyte diffusion

It is expected that several parameters can have an effect on polyelectrolyte diffusion, similar to the effect on proteins.

It is important to keep in mind that polyelectrolytes can also be very different from each other. For instance, the same polyelectrolyte can have a larger or smaller polymeric chain, depending on the number of repeating units, *n*. DSS can go from small molecules (in the order of magnitude of 5 kDa) to very high molecular weight molecules (500 kDa).

Like for proteins, size and shape may have an influence in the *D* values. As the major interactions in play with polyelectrolytes are electrostatic interactions, the influence of the pH and the buffer on the diffusion of such molecules is expected to be similar to the influence of these parameters on proteins.

The same rational applies for both temperature and solvent viscosity. With the temperature increase there should be an increase in the *D* value of the polyelectrolytes, like for proteins. The increase of the medium's viscosity will increase the frictional force of the solvent and, therefore, make it harder for the solute to diffuse.

1.4.3 Protein-Polyelectrolyte Complexes – Future of Drug Delivery?

Proteins and enzymes have been used for the treatment of diseases or disorders in the modern world for a long time now. Despite the long aproval time and rates by drug regulatory agencies, dozens have been approved or were going through approval and by 2015 many were available to the consumer (Struck, 1994), (Kinch, 2015), with more than 130 making it through the FDA approval and reaching the consumer (Muheem, et al., 2016).

The route of administration used for the therapy using proteins is an important aspect of the drug, as it influences the efficacy and also the kinetics of the drug. While research is undergoing to discover new routes of administration, the mainly used nowadays are the parenteral routes, which comprises injections or administration through a catheter (Muheem, et al., 2016). As proteins are not very stable in the human body, they have a short half-life in the serum, lasting only a maximum of few hours. Besides this, the "filters of the human body" also contribute to rapidly degrading proteins in circulation, making it mandatory to have several injections of the given proteins in order to maintain a threshold level of the drug in the system (Burke, 2000).

The frequent injections, to maintain this threshold level, leads to the oscillation of the drug's concentration in the patient's blood. Although this route is the most commonly used, there is still the need to overcome the low acceptability of patients to these drugs (Muheem, et al., 2016). Taking these criteria into consideration, there is the need to create a delivery approach where is important to avoid degradation of the drug whilst maintaining all of its biological activity (Gupta, et al., 2013).

This is where PECs may have a key role. As polyelectrolytes are biocompatible and proteins also behave like polyelectrolytes, a PEC can be therefore formulated using both of these molecules. For instance, the pKa of DSS is lower than 2, which means that above pH 2 this molecule will be negatively charged (Sacco & Dellacherie, 1986). Theoretically, to create a PEC molecule using DSS, all one needs is to have a protein that has a positive net charge and then they will interact, forming a PEC.

There is already a study on PECs with lysozyme and DSS (among other polyelectrolytes). In this study, the authors tested the release rate of protein from the PEC using strong or weak polyacids. To fully understand the underlying mechanisms of the release of proteins from the PEC, it is important to know what drives this release (Kamiya & Klibanov, 2003).

Although the first data available on the subject was not exciting, years of research in the field helped to improve the yields and possibilities of this technology. Shu *et al.* have developed a method using nanoparticles of dextran sulphate and quaternized chitosan and the data revealed that there was fast release of the studied protein (BSA) in phosphate buffer at pH 7.4 (Shu, Zhang, Wu, Wang, & Li, 2011).

1.5 Self-Interaction Chromatography (SIC)

A better understanding on how proteins interact with each other is important, with emphasis on the weak interactions between proteins, as they are important for different reasons that can be related to the industry, namely, in the form of how the stabilization of a given formulation for therapy behaves (Cleland, Powell, & Shire, 1993) or in the case of the present study, to understand protein diffusion.

Self-Interaction Chromatography (SIC) is a relatively recent technique that was first described in 1996 by (Patro & Przybycien, 1996). SIC is a type of chromatography that is based on weak affinity chromatography principles because the target protein of the study is, at the same time, the ligand and the ligate. As the name suggests, the protein of interest will interact with itself and the outcome of the interaction is the aim of the SIC. For the interaction to be possible, the protein needs to be irreversibly bound to the solid state chromatographic phase (Patro & Przybycien, 1996).

After the protein subjected to the study is immobilized to the column, a concentration pulse of the same protein, solubilized in a suitable buffer, is loaded into the column. After injection of this pulse, the chromatography is ran normally, in the same buffer the protein was solubilized. The elution of the protein is done by simply running the buffer in isocratic conditions, meaning the buffer composition for the elution is kept constant. The eluate is monitored with a suitable method for the detection of proteins. Most times, a UV-Vis detector is used and the breakthrough point is determined (Tessier, Lenhoff, & Sandler, 2002). It is assumed that the immobilized protein keeps its native three-dimensional as well as its secondary structure and that it was immobilized in different orientations.

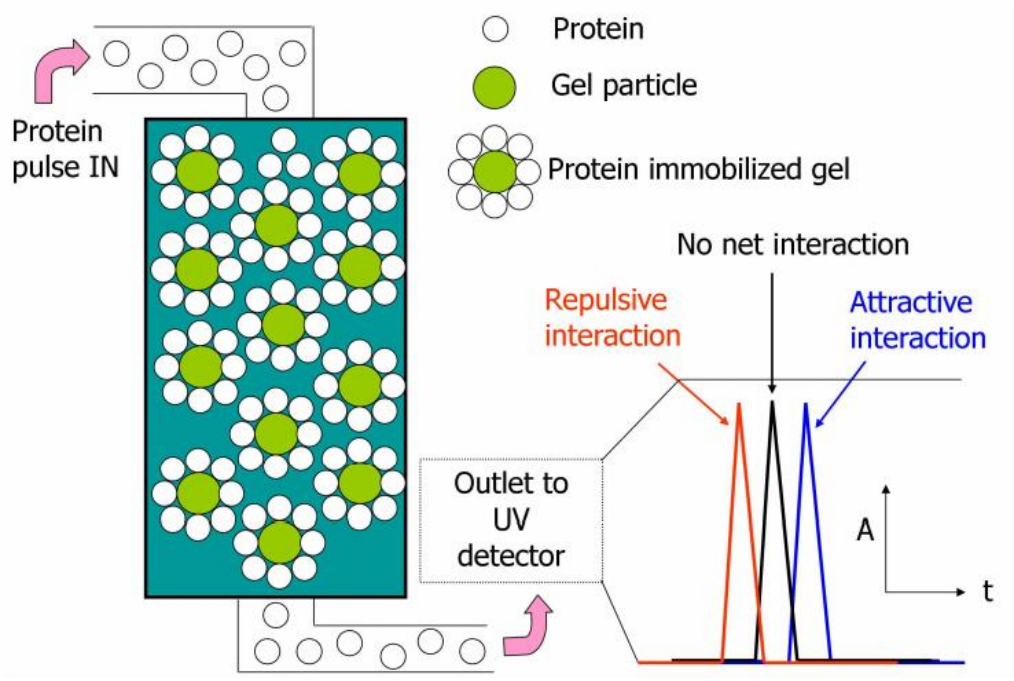

Figure 10 - Schematic of the Self-Interaction Chromatography process. It is possible to see that the ligand and the ligate are the same, represented by the white balls. There is an outlet UV detector for the concentration of the protein. Adapted from *(Deshpande, et al., 2009)*.

2. Aim of studies

The diffusion coefficients of proteins and polyelectrolytes are of great interest for the design of experimental processes. Throughout the past decades, several equations and correlations have been described for the determination of the diffusion coefficient of different molecules, being the Stokes-Einstein correlation the most widely known and accepted, and having suffered several adaptations along the years, to better fit the prediction of the diffusion coefficient.

Although the prediction is important, it would be better to have a method that could provide reliable experimental data, in which the diffusion is measured under very specific experimental conditions. Several methods have been used along the years and, most recently, microfluidics arose as a way to have fast and reliable methods without having to rely on very expensive equipment and allowing for time and sample saving. Furthermore, there has been an ongoing research on the complexes that can be formed between proteins and polyelectrolytes and all the potential behind this technology, as they can be used for therapeutic delivery systems, as it avoids the fast degradation of proteins.

The diffusion coefficient is a reflection of molecular stability and more knowledge on protein/polyelectrolyte behaviour under different conditions can be obtained by studying the diffusion coefficients of these molecules. One of the targets of this study was to determine the diffusion coefficient of different proteins under different experimental conditions, varying conditions that could affect the diffusion, such as ionic strength, solvent's viscosity, solute concentration and protein's molecular weight. The other target was to develop a fast and reliable method, making use of the microfluidics device and a suitable detection method.

The present work took place at the Technische Universiteit Delft in Delft, The Netherlands. It was conducted under the supervision of Professor Marcel Ottens, Professor Geert-Jan Witkamp, Miao Yu and Professor Ana Azevedo.

3. Materials and Methods

3.1 Materials

3.1.1 Chemicals

Different chemicals were used to prepare the different buffer solutions used in the diffusion coefficient measurements. For the preparation of the acetate and phosphate buffers, CH_3COOH glacial (MERCK, ≥ 99,7% assay), CH_3COONa (MERCK, ≥ 99% assay) and Na_2HPO_4 (Sigma-Aldrich, ≥ 99% assay), NaH_2PO_4 (Sigma-Aldrich, ≥ 99% assay) were used, respectively. The viscous solutions were prepared using glycerol (MERCK, ≥ 99% assay) whereas the solutions with ionic strength were prepared using NaCl (J.T. Baker, 100%). A different buffer was used for the Enbrel ® IgG solution using L-arginine monohydrochloride (Sigma-Aldrich, ≥ 99,5%), HCl (MERCK, 37-38%), Na_2HPO_4 (Sigma-Aldrich, ≥ 99%), NaCl and sucrose (MERCK). Milli Q water was used as the solvent for all buffers. The cleaning solution was prepared by solubilization of NaOH pellets (Sigma-Aldrich, ≥ 98% assay). For the SIC trials, $NaHCO_3$ (JT Baker, ≥ 99% assay) and ethanolamine (Sigma-Aldrich, ≥ 98% assay) were used, besides some of the components previously described. Blue Dextran (DxBlue) (Sigma-Aldrich) and acetone (Sigma-Aldrich, ≥ 99.9% assay) were used to assess the SIC columns' integrity.

Lysozyme from chicken hen egg white (Sigma-Aldrich, 14.3 kDa, 70 000 U/mg), cytochrome c from bovine heart (Sigma-Aldrich, 12.327 kDa, ≥ 95% assay), myoglobin from horse heart (Sigma-Aldrich, 17 kDa, ≥ 90% assay), ovalbumin from chicken egg white (Sigma-Aldrich, 44.3 kDa, ≥ 98% assay) human Enbrel® IgG (Pfizer, 150 kDa), bovine serum albumin (BSA) (Sigma-Aldrich, 66.5 kDa, ≥ 95% assay) and dextran sulphate sodium salt (Sigma-Aldrich, 5 kDa) were used. All protein/polyelectrolyte solutions were used using acetate buffer, phosphate buffer or milli-Q water, and will be referred accordingly.

3.1.2 Solutions Preparation

The solutions were prepared by weighing (AG 204 Delta Range scale, Mettler Toledo, Switzerland) every component needed and then adding Milli-Q water. Three different buffers were prepared: i) acetate buffer (AB) at pH 4.2; ii) phosphate buffer (PB) at pH 7.2 and iii) IgG buffer at pH 6.3. After the preparation, every buffer solution was filtered using a 0.2 µm filter (Whatman ®, GE Healthcare, United Kingdom).

After solubilization of the proteins/polyelectrolyte the final solution was also filtered using a 0.2 µm filter for every protein, except for IgG. As the IgG solution was prepared via dilution of a supplier's solution, it is assumed that there would be no insolubilized particles of the protein.

The solutions were used within one week of its preparation, to guarantee that there was no aggregation, precipitation or degradation of the samples.

3.1.3 Set-up used for Diffusion Coefficient measurements

The setup that was used for the measurements of the diffusion coefficient of proteins and dextran sulphate is shown in Figure 11.

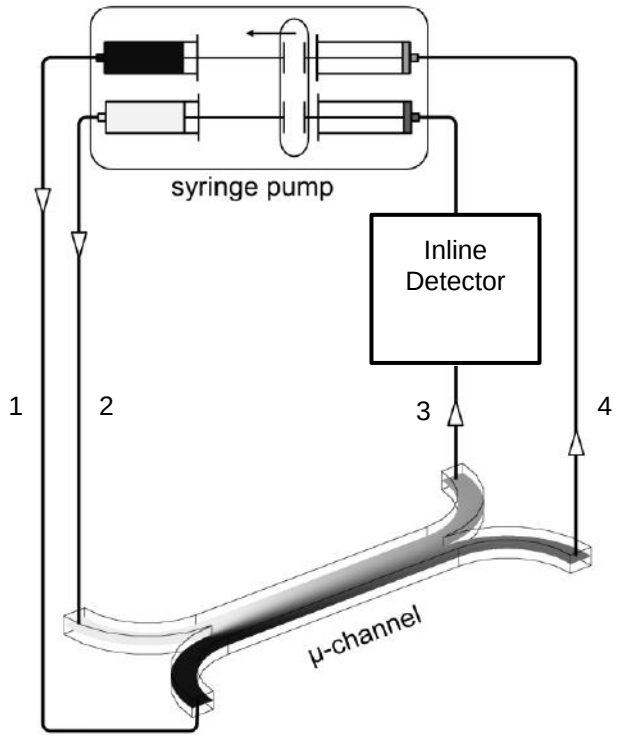

Figure 11 - Experimental setup used for the determination of the diffusion coefficient of the studied molecules. The system comprises a syringe pump module, a microfluidics channel (H-cell) and a detector (Spectrophotometer/Conductivity meter). The syringe module contains two syringes for the inlet of the giving fluid (stream 1) and receiving fluid (stream 2) and two syringes for the outlet of the giving fluid (stream 4) and receiving fluid (stream 3). Image adapted from *(Häusler, Domagalski, Ottens, & Bardow, 2012)*.

The system consists of a syringe pump (Model 200, KD Scientific, Massachusetts) that has four 1 mL syringes (Omnifix®-F, B.Braun, Germany) operating simultaneously. Two syringes are pumping the fluid into the H-cell (Micronit®, Netherlands) (streams 1 and 2) while two syringes are withdrawing the fluid out (streams 3 and 4) at the same rate. This helps to ensure that there is an equal inflow and outflow of fluid in the microchannel system, ensuring that there is a clear interface between the two fluids right in the centre of the channel. The syringes are connected to the microfluidics H-cell (Micronit®, Netherlands) with PEEK^{TM} tubes (Upchurch Scientific®, VWR, Pennsylvania).

Between the H-cell (Micronit®, Netherlands) and the outlet syringe for the receiving fluid, in its path, is an in-line detector, that can either be a spectrophotometer (SPD-20AV, Shimadzu, Japan), to detect the absorbance of the proteins, or a conductivity meter (C3010 Multi-Parameter Analyser, Consort, Belgium), to detect the conductivity of the polyelectrolyte.

The microfluidics chip has a total of five channels. The channels vary in length and width but only one was used. The chosen channel for the study was both the longest and widest channel as seen in Figure 12 (blue outer channel).

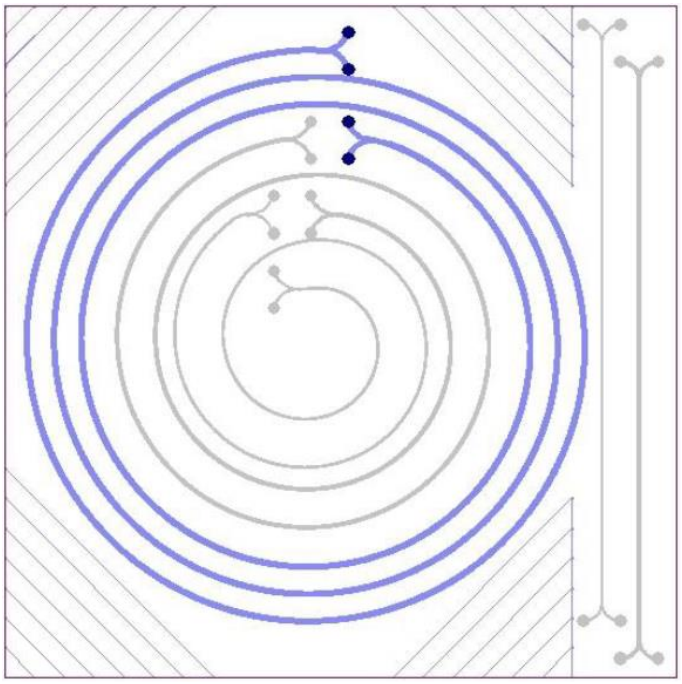

Figure 12 - H-cell profile. There are five available channels in this microfluidics device with varying lenght and width. There are two straight channels and three spiral channels. The channel highlighted with the blue colour is the one used for the current study.

The channels dimensions were provided by the supplier but were measured with image analysis a cross section of the chip outer channel (Annexes) and are summarized in Table 3. The cross-section schematics can be seen in Figure 13.

Table 3 - Summary of chip channel dimensions. The supplier doesn't distinguish between the top and bottom width, therefore the value of 800 µm is presented for both widths. Measurements performed with Olympus (Tokyo, Japan) scale-implemented microscope.

	Supplier	**Measurements**
Depth (µm)	70	69 ± 2
Top width (µm)	800	809 ± 16
Bottom width (µm)		697 ± 18
Average Width (µm)		753 ± 14
Lenght (cm)	62,5	-

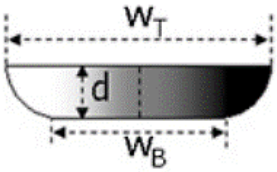

Figure 13 - Cross-section profile of the microfluidics channel. w_T is the top width, d is the depth and w_B is the bottom width. Image adapted from *(Häusler, Domagalski, Ottens, & Bardow, 2012)*.

The H-cell was held using a chip holder, which consisted of a bottom part, in plastic, and top part in metal. Four plastic connectors were used to connect the tubes with the channel, using ferules with rubber rings, to avoid leakage.

3.2 Methods

3.2.1 Determination of protein initial concentration

The protein concentration determination for each sample after solubilization in a suitable solution and filtration (either buffers or water) was done by UV-Vis spectrophotometry. After filtration, it is assumed that some protein could have been adsorbed to the filter or precipitated and, therefore, filtered out. The determination of the initial protein concentration allows to take these events into consideration.

The measurements were performed in a spectrophotometer (UV-1800, Shimadzu, Japan). From the absorbance outlet signal from the equipment and with the quartz cell length (1 cm path length) it was possible to determine the concentration of the solutions using Beer-Lambert law.

3.2.2 Viscosity Measurements

The viscosity measurements were done using and Ostwald viscometer (ROWEEL Electronic, China). The viscosity measurements values were compared to the distilled water assuming a value of 0.001 Pa•s for the viscosity of water.

3.2.3 Mass Balance

The diffusion coefficient is calculated through a mass balance to the chip channels. The two inlet concentrations are known, being always needed to have a receiving fluid and a giving fluid (as previously stated in section 3.1.3). After the diffusion takes place, inside the H-cell, the outlet concentration of the receiving fluid is measured. Knowing this concentration, a mass balance can be done to estimate the total amount of mass that diffused from the giving channel to the receiving channel. Cussler explains this mass balance (Cussler, 2009), that can be applied to the interface of the channel, where both fluids contact with each other but do not mix.

This mass balance should be applied when steady state is reached. In this study, there is a steady-state diffusion across a thin, moving liquid film. Three assumptions were made in order to obtain the mass balance, which are: (Cussler, 2009)

- The solution is diluted;
- The only resistance to mass transfer is the liquid;
- Mass transport has two main components: diffusion along the x-axis and convection along the y-axis (Figure 14). Any other phenomenon for mass transfer is negligible.

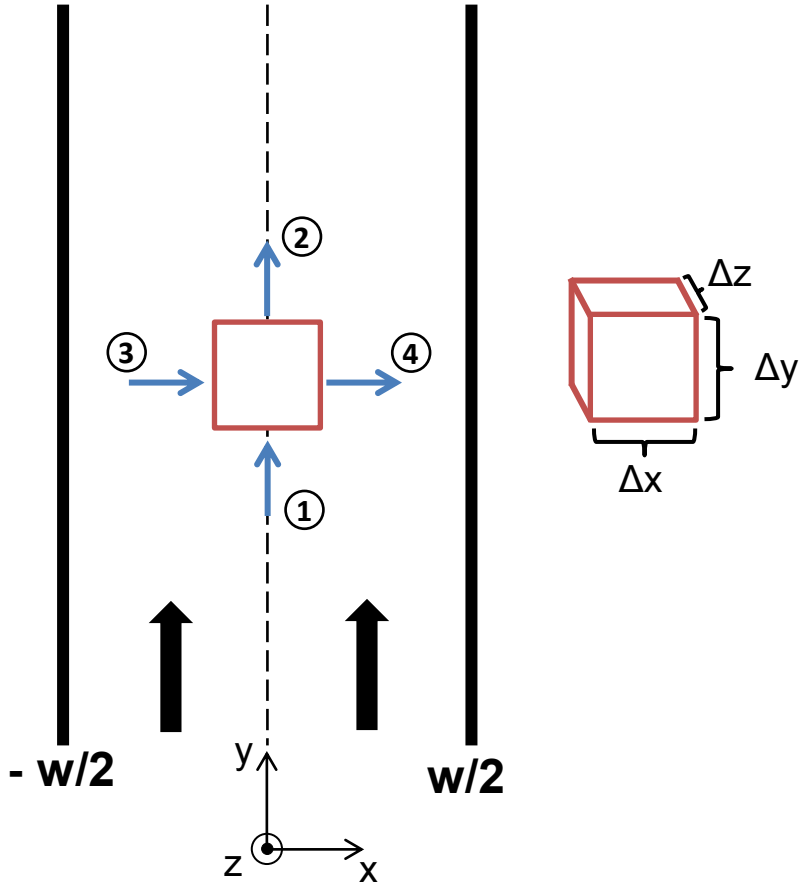

Figure 14 - Schematic of the steady-state diffusion mass balance. The two thick black arrows represent the fluid flow, whereas the blue arrows represent the streams used for the mass balance. Arrows 3 and 4 refer to mass transport by diffusion and arrows 1 and 2 represent mass transport by convection. The red shape, in this case represented by a square, has its dimensions defined as Δx and Δy. The three-dimensional axis is represented in the bottom right position. The half width of the channel is represented by w/2.

The last assumption is the most important of them all, as it states that the mass transfer by convection is negligible in the x-direction and mass transfer by diffusion in the y-direction is assumed much slower than the one by convection and therefore is of much less importance (Cussler, 2009).

It is important to keep in mind that microfluidics channels are not perfect rectangles as the manufacturing is not perfect, generating different geometries than the desired ones, such as round-edged trapezoids as in this case. Previous studies (Häusler, Domagalski, Ottens, & Bardow, 2012) showed that round-edged or trapezoidal geometries differ less than 1% from simplified rectangular geometries using average-width rectangles. The velocity profile for the fluids resembles an “egg shell” (or “bullet”) shape both in the x-axis (or x-y plane) and z-axis (or z-y plane). Erik *et al.* showed that using a x-z-averaged velocity or using a x-averaged velocity showed a difference of less than 0,1% in the calculated outlet concentrations. Therefore, the velocity profile for this study only takes into account the averaged x-velocity (Häusler, Domagalski, Ottens, & Bardow, 2012).

The mass balance, based in Fick’s law of diffusion for steady-state and in Figure 14, is described in equation 3.1.

$$\bar{v}(x) \cdot C \cdot \Delta x \Delta z|_y - \bar{v}(x) \cdot C \cdot \Delta x \Delta z|_{y+\Delta y} - D \cdot \Delta y \Delta z \cdot \frac{\partial C}{\partial x}\Big|_x + D \cdot \Delta y \Delta z \cdot \frac{\partial C}{\partial x}\Big|_{x+\Delta x} = 0 \qquad (3.1)$$

Dividing all of the equation's terms for the characteristic dimensions Δx Δy Δz, equation 3.2 is obtained.

$$\bar{v}(x) \cdot \frac{\partial C}{\partial y} = D \frac{\partial^2 C}{\partial x^2} \qquad (3.2)$$

3.2.4 H-cell measurements

To correctly operate the H-cell and have reliable results, there are some aspects to take into account. The most important one is to ensure that there are no gas bubbles inside the H-cell, meaning that the H-cell must be filled with liquid for the experiment. This makes it possible to assume that the diffusion only takes place in a moving liquid fluid and not due to some turbulence caused by air bubbles. Besides this, the outlet signal measured is also affected when there is air in the system.

3.2.4.1 Determination using an in-line detector

The determination of the outlet concentration was done using an in-line detector. Depending on the nature of the solute, the detector was either a spectrophotometer (SPD-20AV, Shimadzu, Japan) or a conductivity meter (C3010 Multi-Parameter Analyser, Consort, Belgium). The first one was used for protein measurements while the second was used for polyelectrolyte measurements.

The stream 3 in Figure 11 has the solution of the receiving fluid and passes through the detector. The molecules of interest, when passing the detector, will generate a signal, that can be in Absorbance Units (AU) or in µS/cm, whether it is proteins or the polyelectrolyte that are being measured.

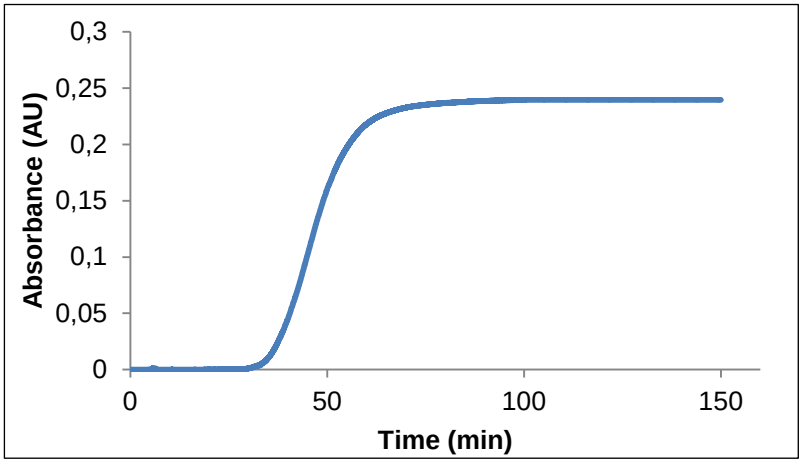

Figure 15 - Signal profile over time. Representation of a measurement using the spectrophotometer (SPD-20AV, Shimadzu, Japan) with a wavelength of 550 nm, to detect the presence of cytochrome c in the receiving fluid stream, in 10 mM acetate buffer, at pH 4.2. This profile was obtained in an experiment where the receiving fluid has 0 mg/mL and the giving fluid has 5 mg/mL of the protein.

With the information of the profile of the signals with time, such as the one represented in Figure 15, one can monitor if the steady state was reached or not and then use the signal response at the steady-state to calculate the outlet concentration which is done using a calibration line.

3.2.4.1.1 Calibration lines

For each different molecule of interest, it was done a calibration line with a minimum of 4 data points. The solutions of known concentration are pumped through the system, using the same solution for both giving and receiving fluids. When passing through the detector, each stock solution originated a different signal and, after plotting the data points a calibration line was obtained, as depicted in Figure 16.

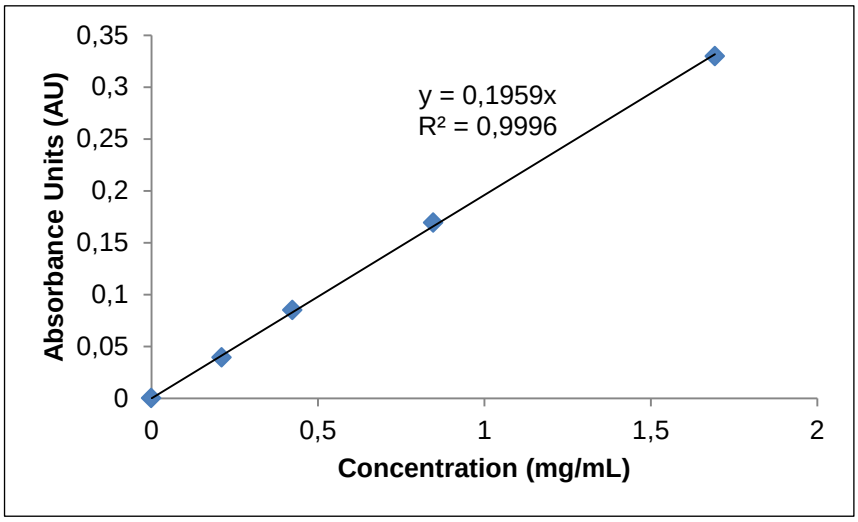

Figure 16 - Calibration line for ovalbumin in phosphate buffer, at pH 7.2. Representation of a calibration using the spectrophotometer (SPD-20AV, Shimadzu, Japan) with a wavelength of 280 nm. The equation of the calibration line is $y = 0.1969x$ and the correlation coefficient is R^2=0.9996

3.2.4.2 Protein adsorption to the chip surface

To quantify the amount of protein adsorbed to the microchip surface walls, *attachment trials* were performed. The experimental procedure was similar to the one described for the outlet concentration of a *normal trial*, with the difference that both outlet streams were connected to an in-line detector. Using the data collected by these devices and the calibration curves, it was possible to make a mass balance in steady-state and assess the amount of protein attached to the chip surface.

3.2.4.3 Determination using an Off-line detector

For trials with higher concentration of protein the detector's signal was saturated. This made it necessary to have an off-line measurement method, in order to know the outlet concentration of the receiving fluid stream. Therefore, a 100 µL loop was installed between the chip and the detector, to enable the collection of the sample only when steady state was reached (as the detector is after the loop, as soon as it was possible to clearly see steady-state in the measurement device it would mean that the solution inside the loop was already in steady-state). After collecting the sample in a microtube, the solution was diluted using the same solvent as the measurement solution has and then measured off-line in a spectrophotometer (UV-1800, Shimadzu, Japan).

3.2.5 Diffusion Coefficient determination

To determine the diffusion coefficient two different software were used, with a combination of Matlab (Mathworks, USA) and COMSOL Multyphysics, v5.2 (COMSOL, Sweden) (Häusler E. B., 2011).

By looking at equation (3.2) it is possible to see that there is the need to know with certainty some variables, to calculate the diffusion coefficient. The initial and boundary conditions were previously set by Häusler *et al* (Häusler, Domagalski, Ottens, & Bardow, 2012) and include: i) the initial concentration of both receiving and giving fluids, ii) no concentration difference on the walls of the chip (meaning that $-\frac{\partial C}{\partial x} = 0$ when $x = -\frac{w}{2} \vee x = \frac{w}{2}$), and iii) the velocity dependence with x. With these parameters and also the volumetric flow rate, the known values are inputted in Matlab, that serves as a user interface. COMSOL was used to build the model of the channels and to solve second order differential equations (Häusler E. B., 2011).

Although it is important to understand how the model works, this model was previously developed and validated in previous work (Häusler E. B., 2011) and is considered out of the scope of the project.

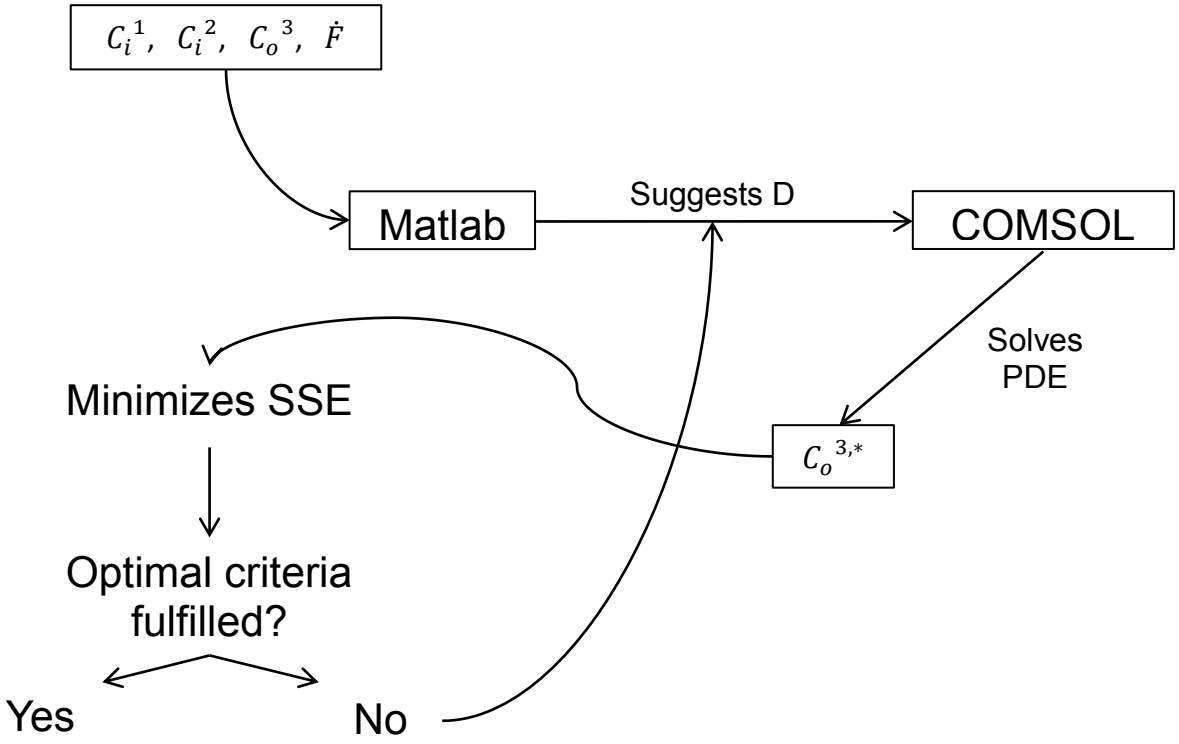

Figure 17 – Model mode of operation. The user inputs inlet concentrations ($C_i^1$and C_i^2) as well as the outlet concentration calculated by the detector (C_o^3) and the volumetric flow rate ($\dot{F}$) to Matlab. Then, Matlab generates a diffusion coefficient (*D*) and inputs it to COMSOL. COMSOL then solves the partial deferential equation (PDE) using both initial concentration, the volumetric flow rate and the *D* value from Matlab. Then calculates the outlet concentration ($C_o^{3,*}$) using the model for mass transfer. This value is then inputted to Matlab, that will compare it with the outlet concentration given by the user, by minimizing the sum of the squared errors (SSE). Then it checks if the optimal criteria were fulfilled and if this is not the case, Matlab will suggest other value for D, until the criterium is fulfilled.

3.2.6 Optimization criteria

As said previously, the Fourier number is a representation of the contact time between the two streams, as low Fourier numbers represent low contact times between both fluids and vice versa. Previous work showed the usage of low Fourier numbers (Costin & Synovec, 2002), (Hotta, Nii, Yajima, & Kawaizumi, 2007). However, Erik *et al.* showed that an optimization regarding this criterion

should be done, in order to have more certainty in the diffusion coefficient determination (Häusler, Domagalski, Ottens, & Bardow, 2012).

The optimal Fourier number can be estimated, when the chip dimensions are known, using the equation (3.3) (Häusler, Domagalski, Ottens, & Bardow, 2012):

$$Fo_{OPT} = 0.299 \left(\frac{b}{a}\right)^{-0.0983} \tag{3.3}$$

Where 2a represents the channel width and 2b represents the channel height.

Despite knowing that having a diffusion coefficient that has an optimal Fourier would be better, this is not always possible, as the optimal Fourier is 0.378 for the channel used (using dimensions in Table 3) and the slightest change in the value of *D* would influence this value. The optimization criteria adapted was a Fourier number between 0.3 and 0.39 (Häusler, Domagalski, Ottens, & Bardow, 2012), where it is possible to have the value of *D* with a large degree of confidence without having the need to always have the optimal one.

3.2.7 Volumetric flow-rates

The volumetric flow-rate used depended on the solute used and on the diffusion coefficient obtained. Ideally, Häusler *et al.* developed a method to estimate the needed flow rate for an experiment depending on the diffusion coefficient. The flow rates were then adjusted as experiments followed to try and achieve a Fourier according to equation (3.3).

Table 4 - Range o volumetric flow-rate (µl/min) used for the several experiments of protein and polyelectrolyte. The flow rate was initially estimated depending on the expected diffusion and according to the Fourier number desired and, after the first estimantion of the D value, it was adjusted experimentally

	Trials	**Range of volumetric flow-rate used (µl/min)**
Lysozyme	*I*	3 - 8.5
	Concentration	3 - 8
	Viscosity	1.5 - 3.5
	MW	3.75
Cytochrome C	*I*	2.75 - 5
	Concentration	3.75 - 7.5
	Viscosity	1.5 - 3.5
	MW	3.5
Myoglobin	MW	2.75
Ovalbumin	MW	1.5
BSA	MW	1.25
Enbrel ® IgG	MW	0.75
DSS	Concentration	9 - 12

3.2.8 Self-Interaction Chromatography (SIC)

For the self-interaction chromatography (SIC) trials, a 1 ml commercial column (Hi-Trap™ NHS-activated, GE Healthcare) connected to an AKTA Avant 25 (GE Healthcare) under the control of the UNICORN software was used.

Initially, the isopropanol was removed from inside the columns with ice-cold HCl (wash-out buffer) with 6 column volumes (CV) at a flow rate of 1 ml/min.

Table 5 - Buffers used for the preparation of the columns for SIC and for the experiments with SIC. *-not measured

Buffer	Concentration (M)						
	$NaHCO_3$	NaCl	HCl	Na_2HPO_4	Ethanolamine	Sodium Acetate	pH
Wash-out	-	-	0.001	-	-	-	*
A	-	0.5	-	-	0.5	-	8.3
B	-	0.5	-	-	-	0.1	4
Standard Coupling	0.2	0.5	-	-	-	-	8.5
Storage	-	-	-	0.01	-	-	8.5
Elution	-	2	-	-	-	-	*

3.2.8.1 Coupling of the ligand

After washing out the isopropanol, 10 ml of a solution of 3 mg/ml of the ligand dissolved in standard coupling buffer was left to recirculate in an ice bath during 4h, with a flow rate of 1 ml/min. Then, the coupling solution is washed out with 3 CV of standard coupling buffer, to further measure the coupling efficiency. Then, 6 CV of buffer A were flowed through the column, followed by 6 CV of buffer B and then 6 more CV of buffer A, at a flow rate of 1 ml/min. The column was then left for 15 to 30 minutes at room temperature. Additionally, 6 CV of buffer B were flowed through the column, followed by 6 CV of buffer A and then 6 CV of buffer B. The column was then stored in storage buffer.

The same protocol is used for the blocked column with the exception that, for this column, no ligand solution is flowed through the column, meaning that the step used after the washing out of the isopropanol is the 6 CV of buffer A, and so on.

3.2.8.1.1 Coupling efficiency

To estimate the coupling efficiency, a Size Exclusion Chromatography (SEC) was run to desalt the solution of 3 CV collected from the SIC column for further analysis in a spectrophotometer. The SEC experiment was performed in a commercial column (PD-10™ desalting column, GE Healthcare).

To do this, the column was first equilibrated using phosphate buffer 10 mM. After that, 0.5 ml of the coupling solution collected in the coupling step of the immobilization of lysozyme in the SIC column was loaded onto the column and the eluent discarded. After that, 2 ml of the equilibration buffer

(phosphate buffer 10 mM) was loaded onto the column and let run, discarding the eluent. Then, 1.5 ml of equilibration buffer was loaded onto the column and the eluent was collected to a microtube.

The concentration of lysozyme was then determined after analysing the resulting solution on a microtube by UV spectrophotometry.

It is necessary to have a coverage area of the column between 20 to 30%, for the results to be considered. To estimate this, the area of one lysozyme molecule is calculated using its hydrodynamic radius (Gull, Ishtikhar, Alam, Andrabi, & Khan, 2017) and the equation of the area of a circle (1.13×10^{-17} m^2/molecule). Knowing the lysozyme molecular weight and how many grams a molecule of lysozyme has (using Avogadro's constant), it is possible to estimate that lysozyme has a total of 4.21×10^{16} molecules/mg. With this, it is possible to estimate the area occupied per mg of lysozyme, that is 0.477 m^2/mg.

The accessible surface area per volume of packed column was estimated by Ahamed *et al.* for an NHS-Sepharose column using lysozyme and its value is 42.5 m^2/ml of column (Ahamed, Ottens, van Dedem, & van der Wielen, 2005). So, to have a coverage area of 20 to 30% it is needed to have the immobilization of 17.8 to 27.7 mg of lysozyme. From the SEC experiment, it was estimated that there was a concentration of 0.86 mg/ml in the microtube collected in the final step of the SEC experiment. This means that the loaded sample had a concentration of 2.58 mg/ml (3 times more concentrated). As the concentration of the loaded sample is the same as the sample collected from the SIC experiment, the amount of lysozyme that did not couple was 7.7 mg. This means that 22.3 mg coupled to the column which translates in a coupled area of 25%. This value is within the desired range and the column can be used for the experiments.

3.2.8.2 Determination of the retention volume

After having both columns ready, the measurements can take place.

An initial blank run is made at a flow rate of 0.75 ml/min, until a stable UV signal is reached. Upon doing so, the column is equilibrated with 10 CV of the experimental buffer. Then, 50 µl of a 1.5 mg/ml solution of lysozyme is injected onto the column. Lysozyme is further washed out from the column with 10 CV of the experimental buffer at a flow rate of 0.5 ml/min. It is expected that the lysozyme exits the column around the 1 ml mark. After this, an elution step with 3 CV of elution buffer is performed, to make sure every lysozyme molecule that could still be in the column is washed out. After the elution step, the column is re-equilibrated if another run is to be performed or cleaned with a storage buffer. The described methodology is applied for both the ligand and the blocked column.

3.2.8.3 Column integrity

To analyse the columns' integrity, a pulse of a 1% (v/v) of acetone in water solution and a 2 mg/ml dextran blue (DxBlue) in water solution were passed through the column. It was expected that acetone showed a sharp and symmetric peak, as it is a very small molecule and flows through the inter and intramolecular space in the column. DxBlue is a molecule with a much higher molecular weight that will only flow through the intermolecular space of the column, eluting earlier. The DxBlue peak showed a tailing due to the poor mass transfer through the pores, resulting in an asymmetric peak. The elution of the DxBlue needs to be done using very high ionic strength, and 5 CV of 1 M NaCl were used for an effective elution.

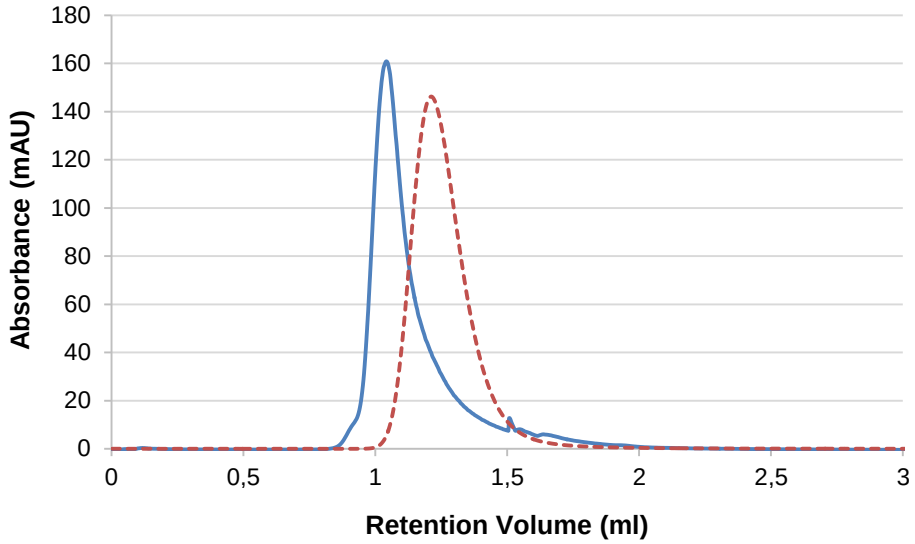

Figure 18 - Retention volume of DxBlue (—) and Acetone (- - -) in a HiTrap™ NHS-activated (GE Healthcare) column. The assessment depicted is for the blocked column. The experiments were run at a controlled temperature of 20ºC.

4. Results and Discussion

4.1 Chip Validation Measurements

To validate the chip and the chosen channel for diffusion, different studies were made. To better understand if the chip was reliable in the generated data, it was important to know if it could provide data in accordance to previous literature and to know what was the chip surface influence in more complex molecules, like proteins.

4.1.1 Sodium Chloride Diffusion

Previous work was done using microfluidics to predict diffusion coefficients. However, this method is still under development stage and more data needs to be gathered using this methodology, to further validate it. One thing that can be done, and considering that the experimental data on molecules like proteins is less than the one for electrolytes, is to use electrolytes for the study as they are more stable molecules than proteins.

To measure the diffusion coefficient, it was necessary to know the outlet concentration of the receiving fluid, in order to assess the molecules that diffused from one fluid to the other. To do this the conductivity meter mentioned on section 3.2.4.1 was used, with the respective calibration line. After the determination of the outlet concentration, the calculation of the *D* value was performed based on equation (3.2) and making use of the Matlab (Mathworks, USA) and COMSOL Multyphysics, v5.2 (COMSOL, Sweden) combination, as mentioned in section 3.2.5.

For this experiment, and for all that will be further mentioned, there were a minimum of three trials that respected the optimization criteria (Fourier number between 0.3 and 0.39) and the average value was considered, with the calculation of the standard deviation.

Erik *et al.* (Häusler, Domagalski, Ottens, & Bardow, 2012) also estimated the diffusion coefficient of different electrolytes, using a microchannel even though with different dimensions. Data gathered along this thesis will be compared with this literature as well as with other literature values.

Table 6 - Measurement of the diffusion coefficient of sodium chloride in aqueous solution using the microfluidics' chip outer channel. The measurements were performed at a controlled room temperature of 20ºC. * - *(Häusler, Domagalski, Ottens, & Bardow, 2012)*, ** - *(Cussler, 2009)*

	Measured *D* (m^2/s)	Literature *D* (m^2/s)	
		Häusler et al.*	Cussler**
Sodium Chloride	$1.55 \pm 0.04 \times 10^{-9}$	$1.60 \pm 0.01 \times 10^{-9}$	1.61×10^{-9}

The experiments for the measurement of the *D* value of sodium chloride show a deviation of about 3.1%, which is considered to be a reasonable deviation. The values from Häusler et al. (Häusler, Domagalski, Ottens, & Bardow, 2012) were measured by having a fluid with no sodium chloride and

the other with 10 mol/m^3 salt. The current experimental data was collected using 0 and 50 mol/m^3 for the receiving and giving fluid, respectively.

4.1.2 Protein attachment to the glass surface

Proteins are complex molecules and have different properties than electrolytes. Therefore, there was the need to understand what was the influence of the chip glass surface on these molecules. As the *D* values are estimated through a mass balance in steady-state, if there is protein adsorption on the chip surface, the outlet concentration of the receiving fluid could be misrepresented. To study this phenomena lysozyme and other proteins were used under different experimental conditions and taking the procedure as explained in section 3.2.4.2.

Table 7 - Assessment of the attachment of different proteins to the glass surface. Different proteins under different conditions were studied. The acetate buffer (AB) and phosphate buffer (PB) had a pH of 4.2 and 7.2, respectively. The lysozyme solutions in water (H_2O) had a pH around 3.5. The measurements were performed at a controlled room temperature of 20ºC.

Protein	Inlet Fluid Concentration (mg/ml)	Solution	Recovery (%)
Lysozyme	0 vs 2	H_2O	98.1 ± 3.0
Lysozyme	0 vs 5	H_2O	$108.9 \pm .4$
Lysozyme	0 vs 10	H_2O	99.4 ± 6.1
Lysozyme	0 vs 2	H_2O + 10 mM NaCl	103.8 ± 3.3
Lysozyme	0 vs 2	H_2O + 500 mM NaCl	103.7 ± 1.3
Lysozyme	0 vs 2	10 mM PB	97.9 ± 2.3
Lysozyme	0 vs 2	10 mM PB	103.1 ± 9.0
Lysozyme	0 vs 2	10 mM AB	105.9 ± 2.1
Cytochrome C	0 vs 2	10 mM PB	96.5 ± 5.9
Myoglobin	0 vs 2	10 mM PB	107.5 ± 0.8
Ovalbumin	0 vs 2	10 mM PB	99.5 ± 1.7
BSA	0 vs 2	10 mM PB	103.1 ± 0.4
Enbrel® IgG	0 vs 2	10 mM PB	101.9 ± 2.8

Looking at the results of Table 7 it is possible to see that the attachment of protein to the glass is neglectable. Overall, the average recovery was 102.3 ± 3.8 %.

4.2 Protein Diffusion

The measurement of the diffusion coefficients of proteins was carried out using a high throughput device, by the conjugation of a microfluidics H cell and a suitable method, which in the case of proteins was a spectrophotometer. The experiments on protein diffusion and its influencing factors can be divided into 4 major studies: i) ionic strength of the solution; ii) effect of protein concentration; iii) effect of medium's viscosity; and iv) effect of protein molecular weight. A comparison between measured and predicted values was also performed using models available in the literature for some conditions used.

4.2.1 Ionic Strength

Initially the goal of the studies was to study how the diffusion of proteins was affected by the ionic strength of the solution. To do this, two different proteins, lysozyme (lys) and cytochrome c (cyt. c), were used using different conditions. All the trials were performed using a water solution (for ionic stregth of 0 M), 10 mM and 100 mM phosphate and acetate buffers, at pH 7.2 and 4.2, respectively, with different amount of NaCl added to increase the ionic strength of the solutions.

Phosphate/acetate groups are not strong electrolytes like sodium chloride, so there was the need to calculate the ionic strength (*I*) of the buffer solution, using the equation (4.1) as described in (Prausnitz, Lichtenthaler, & de Azevedo, 1998), where c_i represents the concentration of "i" ion in the solution and z_i is the charge of this ion.

$$I = \frac{1}{2}\sum_{i=1}^{n} c_i \cdot {z_i}^2 \qquad (4.1)$$

Strong monovalent electrolytes fully dissociate in solution and, therefore, its concentration corresponds to its *I*, in molarity, in solution. However, weak electrolytes as phosphate/acetate salts don't fully dissociate in solution and remain in equilibrium in solution. This chemical equilibrium is dependent also on the solution's pH and this raises the need to calculate the *I* of each solution.

Table 8 - Summary of all the different solutions used in the study. PB represents phosphate buffer, AB represents acetate buffer and *I* represents ionic strength. The ionic strength of each solution was estimated using equation (4.1). * - The solution's pH is due to the self-buffering capacity of lysozyme.

Protein	Solution	NaCl Concentration (mM)	pH	I (mM)
Lysozyme	Water	0	3.5*	0
	PB 10 mM	0	7.2	20
		200		220
		500		520
	PB 100 mM	0		200
	AB 10 mM	0	4.2	3
	AB 100 mM	0	4.2	33
Cyt. C	PB 10 mM	0	7.2	20
		200		220
		500		520
	AB 10 mM	0	4.2	3

The measurements were done using a concentration of protein of 2 mg/mL in the giving fluid and a concentration of 0 mg/mL in the receiving fluid.

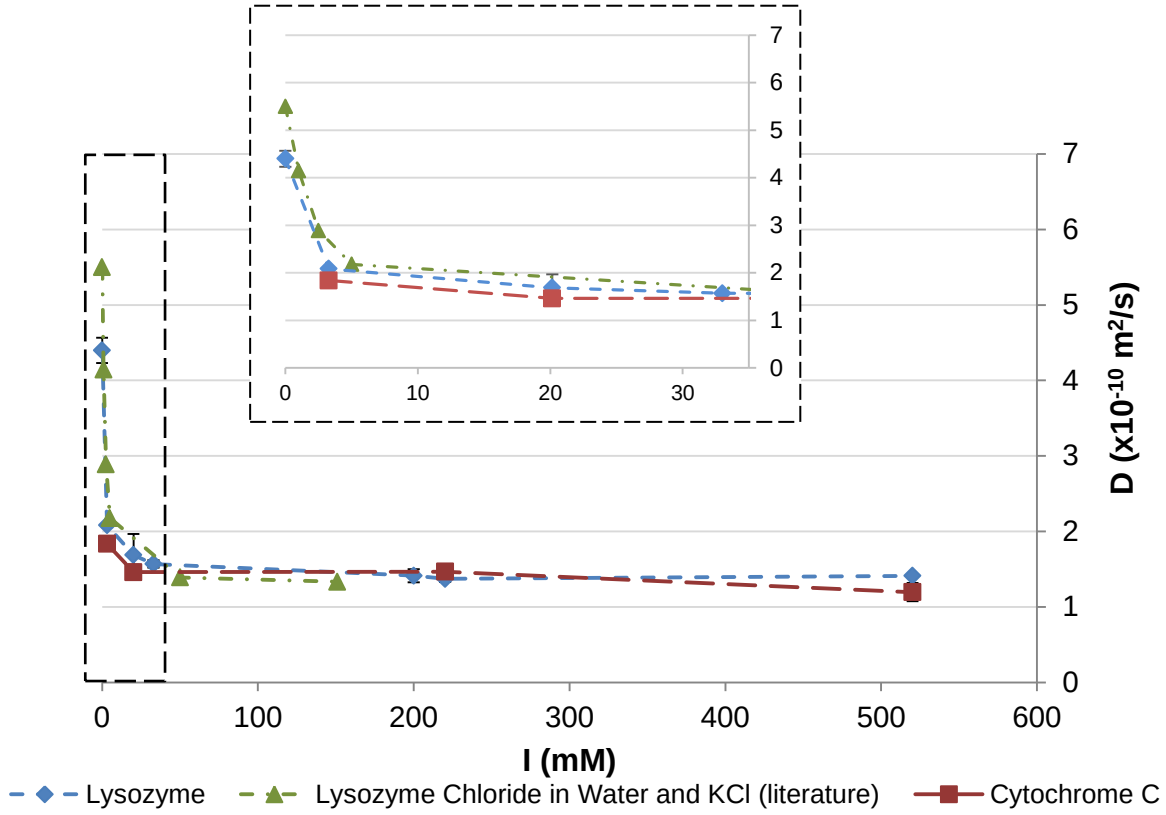

Figure 19 - Diffusion of lysozyme and cytochrome c in different ionic strength solutions and at pH 7.2 (for acetate buffer), ph 4.2 (for acetate buffer) and pH 3.5 (for water solution). The solutions used for the experiments are summarized in Table 8. The green triangles (▲) represent literature values for lysozyme in water, with varying concentration of KCl, for the variation of the Ionic Strength (Cadman, Fleming, & Guy, 1982). The blue diamonds (♦) represent lysozyme experiments in different buffers and the red squares (■) represent the cytochrome C experiments in different buffers with varying concentration of NaCl. The dashed lines in data points are used to guide the viewers eye. Measurements performed at a controlled room temperature of 20ºC. The black dashed line is used to highlight the lower ionic strength region.

Looking at Figure 19 it is possible to see that there is a sharp decrease in the *D* value of lysozyme (triangles and diamonds) in the region of very low ionic strength. On the other hand, experimental data on lysozyme and cytochrome c in a region from low to relatively high ionic strength showed that there is not a significant change in the diffusion of these proteins. It was also observed that there is not a considerable difference in having ionic strength using a higher concentration of phosphate buffer (for example 100 mM) or using a lower concentration buffer and adding salt (for example 10 mM and 200 mM NaCl), as it is possible to see that the two data points for lysozyme at an *I* around 200 mM show virtually the same value for *D* for both cases. This is also observed when comparing similar ionic strengths (PB 10 mM and AB 100 mM, with an *I* value of 20 and 33) from different ions, as the acetate buffer value and the phosphate buffer value seem to have a similar value, that follows the trend of the experimental and literature values.

Although the estimated values are slightly different from the literature values [for instance, for very similar *I* of 3 (2.5 mM for literature, and 3.2 for experimental) the literature value is $2.887 \times 10^{-10}\ m^2/s$ where the experimental value is $2.08 \pm 0.03 \times 10^{-10}\ m^2/s$ and also one can see the same for water only measurements, as the literature value is $5.838 \times 10^{-10}\ m^2/s$ and the experimental value is $4.40 \pm 0.17 \times 10^{-10}\ m^2/s$], both follow the same trend.

The literature data present in Figure 19 is from Cadman and his colleagues work (Cadman, Fleming, & Guy, 1982) and for higher ionic strength the experimental data is very comparable to the data presented by the authors. Kuehner *et al.* also studied the diffusion of lysozyme with varying ionic strengths (Kuehner, et al., 1997). In their work, ammonium sulphate is used to regulate the desired ionic strength, which varied from 50 mM to 5 M, which is very high. Comparing these authors' data with the previously obtained data gathered from literature, the values for the diffusion coefficient seem to have good agreement with both experimental data and Cadman and colleagues' data. Indeed, the *D* value obtained experimentally in this thesis for an ionic strength between 25 and 500 mM is quite similar to the data obtained by Kuehner *et al.* between 50 and 300 mM and by Cadman *et al.* between 50 and 150 mM.

The values of *D* obtained for very low ionic strength are much higher than the one for high ionic strengths. Both lysozyme and cytochrome c have a very high pI value (11.3 and 10, respectively), which means that for all the tested conditions the net charge of the proteins was positive. Cytochrome c was not tested for 0 M ionic strength and the lowest *I* value assessed for this protein was around 3 mM, using 10 mM AB.

Being both proteins spherical and positively charged, the main interactions are repulsive electrostatic, which explains the decrease of the diffusion coefficient for both proteins, although this phenomenon is easier to see for the case of lysozyme. Both for Cadman data and for experimental data, the *D* value of lysozyme decreases around 77% and 69%, respectively, for an ionic strength increase from 0 to 150 mM for Cadman experiments and from 0 to 220 mM for the experimental data of the present work. This shows that the diffusion coefficient varies greatly with the ionic strength of the solution.

An increase of the ionic strength of the solutions will lead to an increase of the free ions in solution that can establish electrostatic interactions with the proteins. Increasing I, the free electrostatic energy of the proteins will decrease and that will make it less likely for them to repel one another, reducing the diffusion coefficient. Ions in solution have a shielding effect over the charged residues of the proteins and residues that were previously charged and able to interact with other charge residues are not available anymore, decreasing the repulsive interactions between the molecules (Cadman, Fleming, & Guy, 1982).

Another phenomenon noticeable is that, from a given *I* value onwards, there is virtually no variation of the diffusion coefficient. This happens because only a certain ionic strength is needed to shield all protein charged residues and the proteins will present a typical hard sphere diffusion, behaving like there are no more proteins in solution. This was visible until an ionic strength of 500 mM. However, if *I* is increased to very high values (*e.g.* 3M - a very high value of ionic strength) then the diffusion coefficient can be, again, affected by the ionic strength. In this scenario, the influence of the *I* is not the prevention of the electrostatic interactions between proteins but rather the possible occurrence of agglomeration or precipitation (salting-out). Larger molecules diffuse slower than smaller molecules due to distinct reasons, one being the larger drag coefficient that bigger molecules have when compared to smaller molecules.

4.2.2 Concentration

From the ionic strength results, it is possible to understand how protein molecules can influence the diffusion by interacting with each other's. Another parameter worth investigating is the concentration of the target protein. To perform these concentration trials, a different experimental approach from what had been previously used was performed.

In previous experiments the receiving fluid and a giving fluid had an initial concentration of protein of 0 and 2 mg/ml, respectively. Häusler *et al.* stated that the diffusion coefficient measured in a H-cell is the average of all diffusion coefficients in the cell. This means that for previous experiments, it was considered to be a diffusion coefficient of a protein at a concentration of 1 mg/ml (the average concentration of protein within the H-cell). For the experimental setup to work, it is necessary to have a concentration difference between the receiving and the giving fluid and the minimum concentration difference accepted was considered to be 2 mg/ml. So, to determine the *D* value of a certain protein at different protein concentrations, it was necessary to adjust the concentration of the giving and receiving fluid according to the objective. To determine *D* at 1 mg/ml the receiving and giving fluid had a concentration in protein of 0 and 2 mg/ml, respectively; to determine at 3 mg/ml the receiving and giving fluid had a concentration of 2 and 4 mg/ml, respectively; to determine at 5 mg/ml the receiving and giving fluid had a concentration of 4 and 6 mg/ml, respectively.

Studies already published in the literature have tried to correlate the diffusion coefficient with the concentration of the solute, and an equation that expresses the variation of the diffusion coefficient with the concentration of the solute was determined (Han & Akcasu, 1981), (Anderson, Rauh, & Morales, 1978).

$$D = D_0 \times (c \cdot k_c + 1) \qquad (4.2)$$

This equation relies in experimental data and empirical values. k_c is a concentration coefficient that needs to be estimated, namely estimated through mathematical equations, like Han and Akcasu or Anderson *et al.* showed. The units of the coefficient are the inverse units of concentration.

To understand how the two studied proteins behaved as their concentration increases, different concentrations were tested in different buffers. Lysozyme and cytochrome c trials were performed using solutions 10 mM AB or PB and also water for lysozyme and the *D* values obtained are represented in Figure 20 and Figure 21, respectively for lysozyme and cytochrome c.

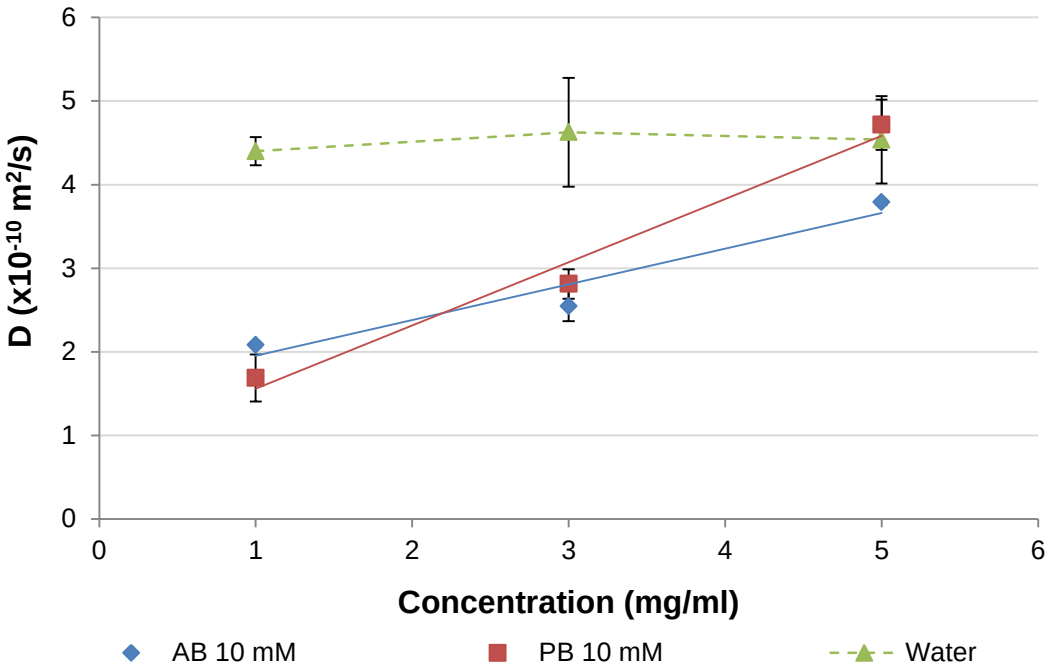

Figure 20 - Lysozyme diffusion with varying concentration, in different solutions. The green triangles (▲) represent the water trials, the blue diamonds (♦) represent the trials in 10 mM acetate buffer and the red squares (■) represent the trials in 10 mM phosphate buffer. Measurements performed at a controlled room temperature of 20ºC. The dashed green line in data points is used to guide the viewers eye. A linear regression for the trials was done except for water. For 10 mM AB, the equation of the regression was $y = 0.427x + 1.523$ with a correlation coefficient of R^2=0.935 (full blue line). For 10 mM PB, the equation of the regression was $y = 0.757x + 0.801$ with a correlation coefficient of R^2=0.978 (full red line).

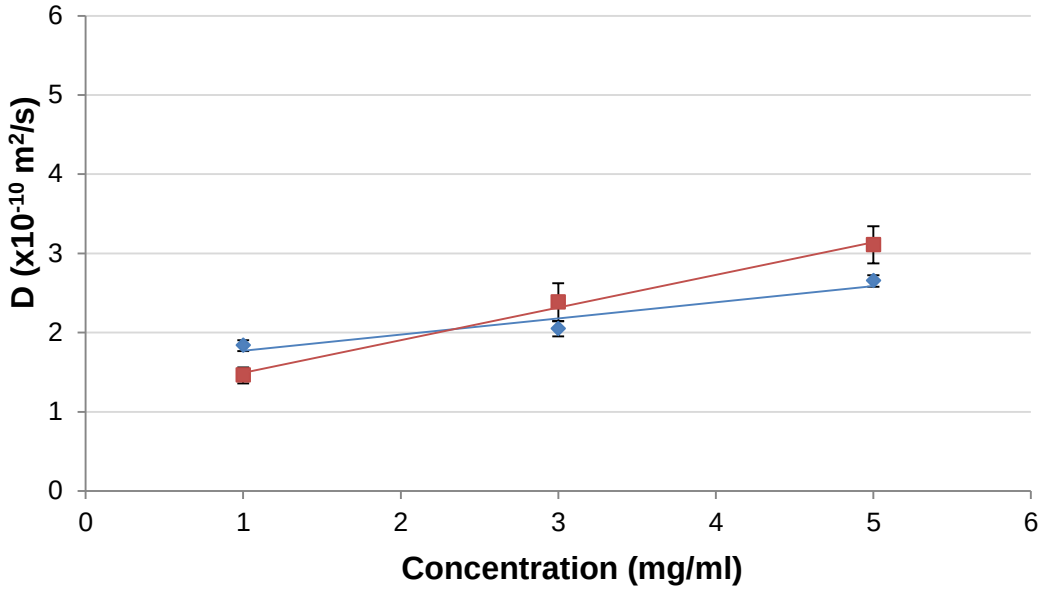

Figure 21 – Cytochrome C diffusion with varying concentration, in different solutions. The blue diamonds (♦) represent the trials in 10 mM acetate buffer and the red squares (■) represent the trials in 10 mM phosphate buffer. Measurements performed at a controlled room temperature of 20ºC. A linear regression for the buffers was done. For 10 mM AB, the equation of the regression was $y = 0.204x + 1.568$ with a correlation coefficient of R^2=0.928 (full blue line). For 10 mM PB, the equation of the regression was $y = 0.412x + 1.081$ with a correlation coefficient of R^2=0.995 (full red line).

According to Figure 20 and Figure 21, it is possible to see that both lysozyme and cytochrome c diffusion coefficient increases with an increase of the concentration of protein for the tested buffers, with low ionic strength (3 mM for the acetate buffer and 20 mM for the phosphate buffer). For the lowest ionic strength tested (water solution, with 0 mM of ionic strength) the diffusion coefficient was constant for the tested concentrations.

The increase of the diffusion coefficient of the proteins with the increase in concentration is explained with the repulsive interactions that these proteins have with themselves in the tested conditions as this is a similar phenomenon to a previously reported one for a different protein (α-chymotripsinogen) (Blanco, Perevozchikova, Martorana, Manno, & Roberts, 2014).

Once again, it is challenging to fully separate the inherent mechanisms that affect the diffusion of both lysozyme and cytochrome c. The observed data can be explained by the low ionic strength of the solutions used. For lysozyme the solutions studied include water, acetate buffer and phosphate buffer with an ionic strength of 0, 3 and 20 mM, respectively, and for cytochrome c the solutions studied were acetate and phosphate buffer with an ionic strength of 3 and 20 mM, respectively. Considering the available ions to shield the surface charge of the proteins, for the same ionic strength, it will be easier for the ions in solution to shield the number of lysozyme molecules at a concentration of 1 mg/ml than at a concentration of 5 mg/ml, as the number of molecules, in the same volume, is five times higher. So, for lower concentrations of protein, these can behave as hard-spheres because it has its surface charges shielded by the ions in solution. Theoretically, it is expected that there will be no change (either increase or decrease) when the available ions in solution are enough to shield all protein net charges at the tested concentrations (in this scenario the molecules also behave like hard-spheres).

It may not be correct to say that *D* is influenced by the concentration of proteins, even though the data shows exactly that. It may be more accurate to say that the *D* of these proteins is influenced by the ratio of ions to protein molecules, as reported by Sorret *et al.* (Sorret, DeWinter, Schwartz, & Randolph, 2016). Simplifying the phenomena, when this ratio is enough to shield all the charges of the protein, it acts as if the protein doesn't "see" other proteins in solution and for this reason, doesn't interact with other protein molecules via electrostatic interactions. Decreasing this ratio will make less ions available per protein, and the molecules will interact more easily with each other.

This can be one of the reasons why the *D* value virtually doesn't change with the concentration of the protein in water (the values are virtually the same, taking into account the error bars). Despite Cadman *et al.* showing that the diffusion coefficient decreases with the increase of the concentration of lysozyme, the changes presented by these authors are not very significant and it is usually very hard to collect such accurate data for the diffusion coefficient. Other authors, however, reported that the diffusion coefficient of lysozyme didn't change in a distilled water solution (Kim & Myerson, 1996). It is challenging to compare the experimental data with any of the data from these two references as the values predicted by Cadman *et al.* are a little bit higher than the values estimated and show a decrease of *D* with the increase of lysozyme concentration, while the values from Kim and Myerson follow the same trend (constant value of *D* in the tested concentration region) but differ considerably

in value when compared with experimental data obtained (literature value around $2.2 \times 10^{-10} m^2/s$ compared to experimental value of $4.5 \times 10^{-10} m^2/s$).

4.2.2.1 Self-Interaction Chromatography (SIC)

Self-interaction chromatography (SIC) was used to study the interactions that occurred in the diffusion trials, namely if a protein, in a given condition, has with itself a repulsive or an attractive interaction or if it doesn't interact with itself in any way. These interactions can be quantified with further estimation of the second virial coefficient, B_{22}, that is positive for repulsive interactions and negative for attractive interactions. The experimental procedure used is described in the sub-section 3.2.8.

After testing to the columns' integrity and the estimation of the void volume, the trials followed (the void volume was found, and is an important parameter for the estimation of the B_{22}). As the equilibrium and washing are done using the buffer in which we want to test the interactions, and considering that the protein doesn't bind to the ligand (which is the protein itself), it is expected that the protein present in the sample elutes during the washing steps. The elution step (with high salt concentration) is used to make sure that no protein remained inside the columns. The SIC experiments were only performed with lysozyme.

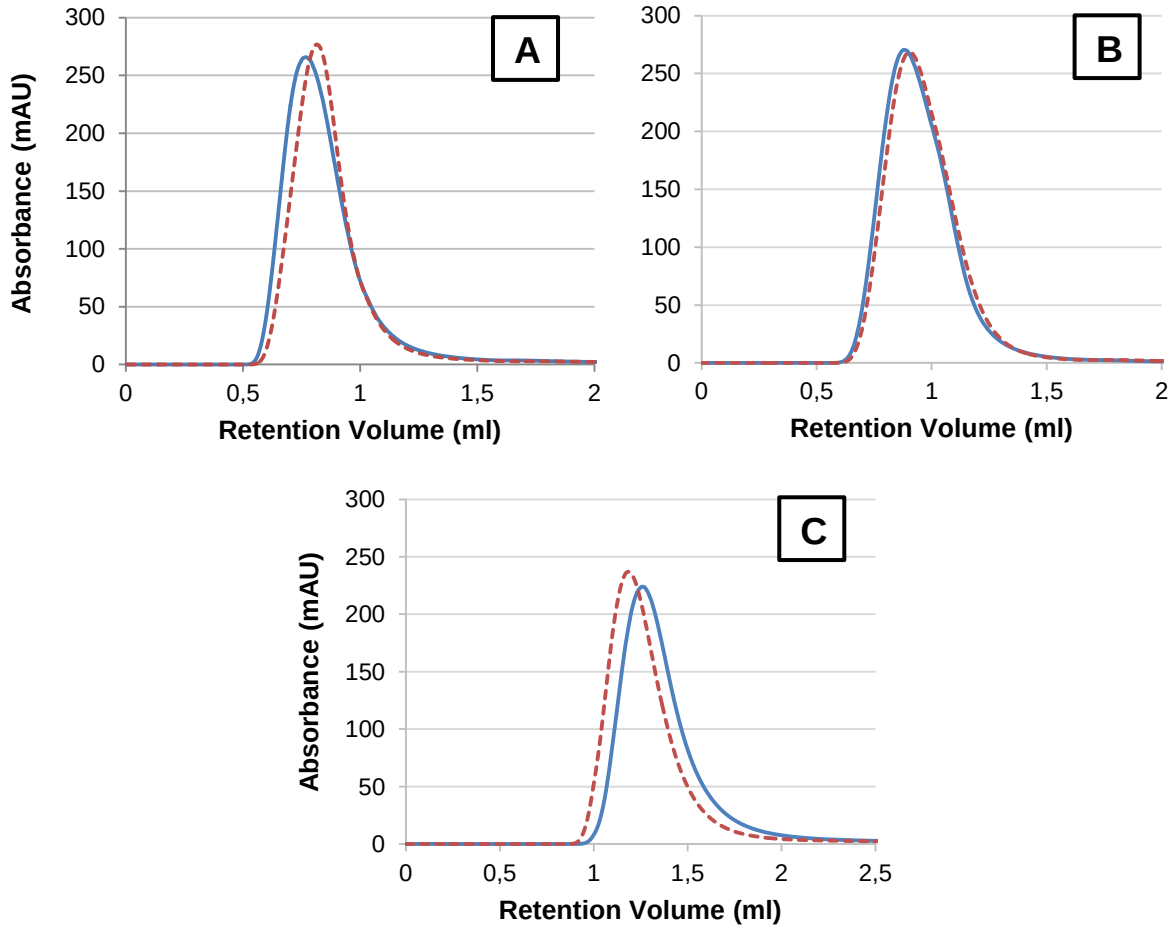

Figure 22 - Determination of the retention volumes of ligand column (represented by the blue line) and blocked column (represented by the red dashed line.). A represents the trials in water solution, B represents the trial in 10 mM acetate buffer solution and C represents the trials in 10 mM phosphate buffer solution. The trials were performed as mentioned in sub-section 3.2.8 at a controlled room temperature of 20ºC.

It is accepted that lysozyme molecules' interactions are mainly based on two phenomena, repulsive electrostatic interactions and attractive van der Waals interactions (Haynes, Tamura, Korfer, Blanch, & Prausnitz, 1992). Therefore, when there is a low ion to protein ratio, the predominant interactions between these molecules are the electrostatic interactions, making the molecules repel each other. However, when this ratio is increased, it is expected that the electrostatic interactions lose their predominance over the attractive van der Waals interactions.

The SIC data shows that the predominant interactions in water and acetate buffer are repulsive, as lysozyme exhibits a lower retention time in the ligand column; while in phosphate are attractive, as lysozyme elutes with a higher retention time in the ligand column. This observation goes against what was hypothesized before, as it was expected to have repulsive interactions between lysozyme in all the solutions studied, since the *D* value in water is considerably higher than the values obtained in AB and PB for a concentration of 1 mg/ml, with the diffusion coefficient increasing with an increase of the protein concentration in both buffers.

The SIC experiments were performed using an inject concentration of lysozyme around 1.5 mg/ml. It is important to keep in mind that the interactions between molecules, with varying concentration, depend on the ion/protein ratio. This can mean that at a low concentration of protein in a buffer solution, the ionic strength might be enough to shield all protein charges, decreasing the electrostatic interactions (Quigley, Heng, Liddell, & Williams, 2013) and, therefore, a larger influence of other type of interactions will rule the diffusion.

This was not the case for the measurements performed in water, as there were no ions in solution that could shield the charges of the proteins. For AB and PB conditions are different, as both buffers contribute to the ionic strength which is higher in PB than in AB. When testing the buffer samples in the SIC experiments, a low concentration of protein was used. At this concentration the presented data shows that lysozyme has repulsive interactions for AB and attractive interactions for PB.

This ca be explained considering the different ion/protein ratios between buffers and comparing the concentration trials ratio and the SIC ratio. Therefore, it could be possible that an experiment suggests that lysozyme has attractive interactions with itself when the opposite was expected. The lysozyme has, in fact, attractive interactions with itself, but in the tested SIC conditions, as the concentration of protein increases, the ion/protein ratio decreases, leading to a higher influence of the electrostatic repulsive interactions. Furthermore, in the assay performed in PB, at pH 7.2, lysozyme is less positively charged, and consequently the repulsive electrostatic interactions will be weaker that at pH 4.2, for AB. Additionally, the ionic strength in PB is almost 10 times higher than in AB, and thus charge shielding is more effective. Besides, at pH 7.2, half of the phosphate ions are divalent (charge -2) and can promote interaction by ion-pairing, explaining the attractive behaviour observed during the chromatography runs.

The aim of the SIC experiments was to have a qualitative determination of the protein interactions rather than a quantitative determination. SIC is a precise method and the smallest deviation could influence the final value for the B_{22} value. Nonetheless, it is important to have the possibility to do the

quantitative determination of the lysozyme interactions via calculation of B_{22}. Both AB and PB showed reproducible results with very low deviation between each trial. However, and although the water trials show the same trend, that is the repulsive interactions between lysozyme molecules in water solution, the difference in the differences of the retention volumes can change up to 0.1-0.2 ml. Although this may not seem much, for the determination of the B_{22} value it can have a significant influence (Quigley, Heng, Liddell, & Williams, 2013). The difference between the reproducibility of the trials can rely on the changes of the conditions in which the protein is solubilized and in which the experiments are performed. The lysozyme water solution has a pH value of around 3.5, due to the self-buffering capacity of lysozyme.

The SIC water trials were ran using distilled water at room temperature, that has a neutral pH. The inject volume was very small (50 µl) that will flow through a 1 ml column. At the moment of the injection, the full amount of injection volume is placed on top of the column, promoting a dilution of the analysed solution and also a change of the pH of the solution in which lysozyme was solubilized (it was previously at pH 3.5 and then is injected in a neutral pH). This phenomenon, that does not help on the protein stability, is possible to be the cause of the non-reproducibility of the water trials for the SIC experiments, even though a repulsive trend is always verified.

4.2.3 Viscosity

After understanding the effect of the ionic strength on *D* and how diffusion can change depending on protein concentration, the influence of the solvent's viscosity was subsequently studied. To understand how the diffusion is affected by the viscosity of the solvent a wide variety of buffers was tested, ranging from low viscosity (water viscosity, aqueous solution) to very high viscosity. The solutions were prepared in 10 mM phosphate buffer pH 7.2, by adding glycerol at 10, 20 and 40% (v/v). As the glycerol was very viscous and despite several washes of the glass beaker to make sure that all glycerol was in solution, it was decided that it would be better to accurately know the viscosity of each solution, using an Ostwald viscometer, as explained in section 3.2.2. The solutions used for the trials are shown in Table 9.

Table 9 - Solutions used in the measurement of the diffusion coefficient with the variation of solvent's viscosity. The measured viscosity is an average of three trials but the errors were so minimal that were not represented.

Added glycerol (% v/v)	Measured Viscosity (Pa·s)
0	1×10^{-3}
10	1.36×10^{-3}
20	1.84×10^{-3}
40	4.57×10^{-3}

The measurements were made using a concentration of 2 mg/ml of protein in the giving fluid and 0 mg/ml in the receiving fluid.

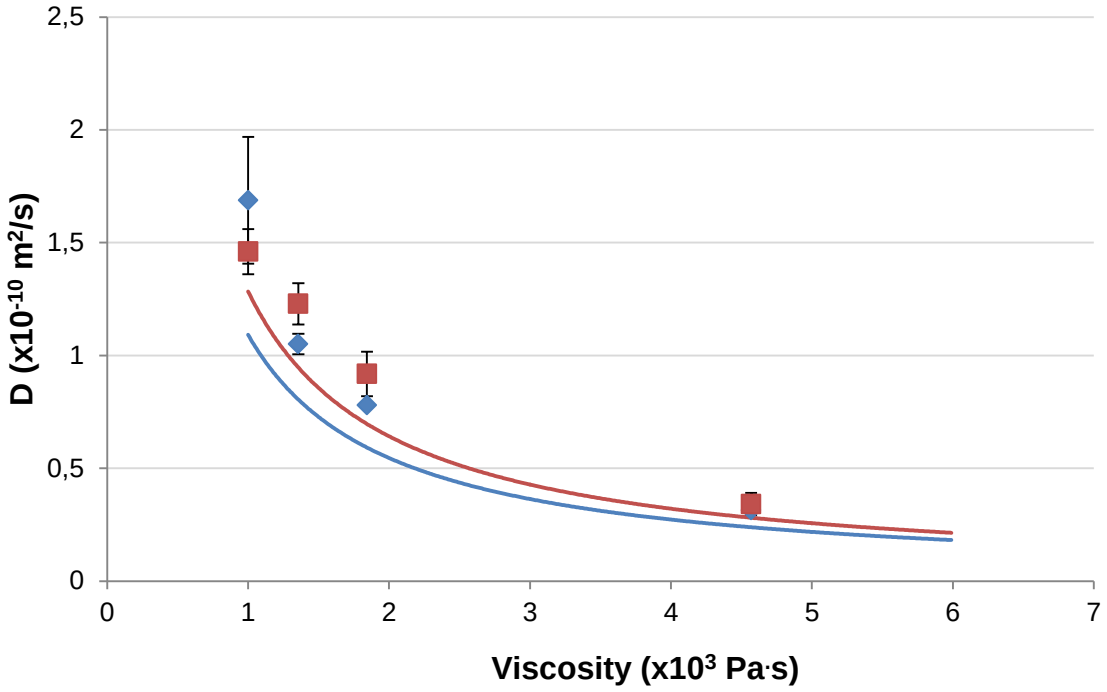

Figure 23 – Protein diffusion with varying viscosity, in PB 10 mM. The blue diamonds (♦) represent the lysozyme trials and the red squares (■) represent the cytochrome C trials. The blue and red lines represent the variation of the diffusion coefficient with the variation of the solvent's viscosity and are based on Stokes-Einstein correlation (equation (1.3)) for lysozyme and cytochrome C, respectively. The hydrodynamic radius used for the use of the correlation were 19 Å for lysozyme and 17 Å for cytochrome C *(Gull, Ishtikhar, Alam, Andrabi, & Khan, 2017)*. The concentration of protein in the giving fluid was 2 mg/ml and 0 mg/ml in the receiving fluid. The measurements were performed at a controlled room temperature of 20ºC.

It is possible to see from Figure 23 that there is a decrease in the diffusion coefficient of both proteins with the increase of the solvent's viscosity. For the case of molecules diffusing in a solvent that is much smaller, the Stokes-Einstein correlation is good for the prediction of the variation of *D*. The increase of the viscosity will lead to a larger drag force of the solvent applied in the solute, as the friction of the aqueous solution increases. Stokes-Einstein correlation also predicts an inverse proportionality between *D* and the medium's viscosity, and that is what is observed.

It is important to highlight that Stokes-Einstein is a very simplistic model, that doesn't take into account inter-molecular interactions. However, and as discussed in sections 4.2.2 and 4.2.2.1, for the used experimental conditions the electrostatic interactions influence in the diffusion phenomena is shielded due to the ions in solution, and only the van der Waals interactions have a role, having both proteins a hard-sphere diffusion behaviour.

The observation of Figure 23 shows that the prediction of the Stokes-Einstein correlation could be rather accurate for this specific case. It appears that what keeps the equation's values from being closer to the experimental ones is the hydrodynamic radius chosen for each of the proteins. In fact, for

lysozyme several hydrodynamic radii have been suggested for similar experimental conditions to the ones used (at the lowest viscosity), and range from 17 to 20 Å and the chosen to the Stokes-Einstein correlation was 20 Å. For bovine cytochrome c there is limited data on this subject and a R_H of 17 Å was used. Using a different value for the hydrodynamic radius of both proteins could show a very precise fitting of the Stokes-Einstein correlation to the experimental data, as a hydrodynamic radius of 16 Å for lysozyme and 15 Å for cytochrome C would largely improve the fitting of the correlation to the experimental data.

4.2.4 Molecular Weight

The next step was to understand how diffusion changed depending on the solute, more specifically how the diffusion varied with an increase of the size of the diffusing molecules. To do so, the diffusion coefficient of different proteins was determined, using the H-cell. The proteins used in this study were the lysozyme and cytochrome C, as before, and myoglobin, ovalbumin, BSA and Enbrel ® IgG. All the trials performed in the present study were performed using 10 mM PB pH 7.2 except for IgG, for which a different buffer was used, recommended by the manufacturer, and a concentration of 2 mg/ml in the giving fluid of the proteins and a concentration of 0 mg/ml in the receiving fluid.

The proteins differ not only in size, but also in shape and net charge. Lysozyme and cytochrome c stand out for their very high pI value (11.3 and 10, respectively), being positively charged for all the tested conditions in the present study. However, myoglobin presents a pI value of 6.8-7.2, meaning that the net charge of the protein will either be negative or neutral. The fact that the pI of this protein is so close to experimental pH was noticeable during the preparation of the solution of myoglobin, as it clearly took a long time to fully dissolve in the used buffer. Ovalbumin, BSA and Enbrel ® IgG have pI values of 4.5, 4.7 and 4.4-5.5, so all of these proteins will have a negative net charge in solution.

Although the different proteins have a different net charge in solution, it is considered that this will not greatly influence the diffusion, as the ionic strength of the solution (24 mM) is enough to shield the charged protein residues, preventing electrostatic interactions, meaning the prevalent protein-protein interactions will be due to van der Waals interactions.

Table 10 - Molecular Weight of the different used proteins.

Protein	Molecular weight (kDa)
Cytochrome C	12.3
Lysozyme	14.3
Myoglobin	17.0
Ovalbumin	44.3
BSA	66.5
Enbrel ® IgG	150

The different proteins also have different shapes. Lysozyme, cytochrome c and myoglobin have a somewhat spherical shape while ovalbumin has an ellipsoidal shape and BSA has a “bulkier” shape that is heart-shaped and Enbrel ® IgG has a Y shape. The fact that most of these proteins does not have a spherical or even close to spherical shape makes it harder to use the Stokes-Einstein correlation. The gathered data is presented in Figure 24.

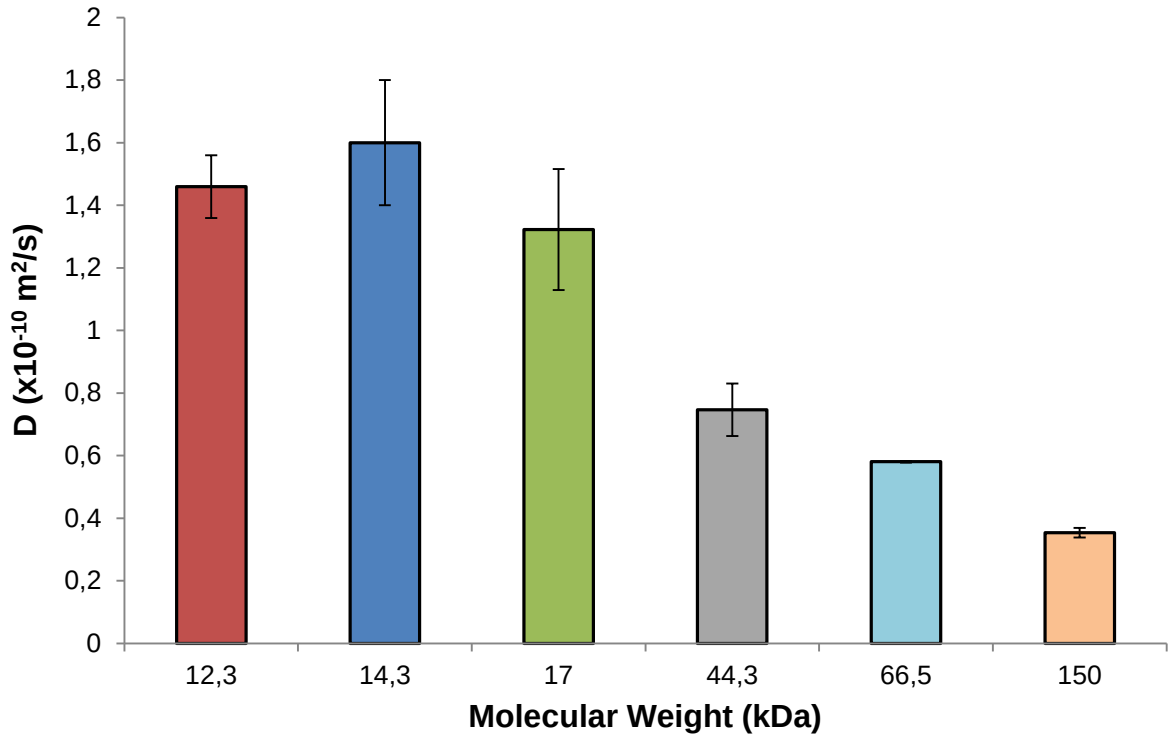

Figure 24 – Protein diffusion of different proteins with different molecular weights. From left to right, the proteins are Cytochrome C, Lysozyme, Myoglobin, Ovalbumin, BSA and Enbrel ® IgG. The concentration of protein in the giving fluid was 2 mg/ml and 0 mg/ml in the receiving fluid. The trials for all proteins were done in phosphate buffer 10 mM except for IgG, that had its own buffer. The measurements were performed at a controlled room temperature of 20ºC.

The data shows that an increase in the size (in this case showed by the molecular weight) will lead to a decrease in the diffusion coefficient. The first three proteins (cytochrome c, lysozyme and myoglobin) have similar molecular weights and shapes and the only difference (besides the MW difference) is the pI value of each. This is in accordance with what the Stokes-Einstein correlation (equation (1.3)) predicts, as an increase in the molecular weight is most of the times synonymous of an increase in the hydration radius. The influence of the follows the same trend of the influence of the solvent’s viscosity, as an increase in this value will lead to an inversely proportional decrease of *D*, as a smaller molecule will be subjected to a lower friction from the solvent, when moving freely in solution, when compared with a larger molecule.

A comparison with equation (1.4) was done, where the values were calculated using the tested proteins’ molecular weight. For the three proteins with the lowest molecular weight, this model predicts lower values than the ones obtained experimentally (deviations of 25-35 %). However, for ovalbumin and BSA, the deviation between H-cell values and model values is much smaller (around 5% deviation). For Enbrel ® IgG the same thing that happened with lysozyme was observed, as there was

a deviation of around 32% between the model and the experimental value. These deviations can be explained by the way the model was obtained. Looking at the model, it is possible to see that the equation has a constant value, that was determined based on experiments, and is empirical. This model may be suitable for a certain range of molecular weights, but proves to be limited for a broader range.

4.3 Polyelectrolyte Diffusion

4.3.1 Concentration

The protein concentration trials were performed using all three solutions previously used for the model protein: water, 10 mM AB and 10 mM PB. It is important to highlight that the solutions' pH values are 6.9, 4.2 and 7.2, respectively. The concentration trials were performed using the same methodology used for proteins, meaning that the concentration under analysis was the average of the concentration between the two inlet fluids, i.e. if the desired concentration is 2 mg/ml, than the giving and receiving fluids should have a concentration of DSS of 1 and 3 mg/ml, respectively.

DSS solubility is much higher than for proteins. The solubility of this compound is around 100 mg/ml, according to the manufacturer, so a higher concentration could be tested for the determination of the *D* value for this compound. The non-saturation of the conductivity signal using higher concentrations of the polyelectrolyte also allowed for the higher concentrations to be tested. DSS has a pKa value lower than 2, which means that for every tested condition it will have a negative net charge. The resulting trials of the concentration experiments for DSS are presented in Figure 25.

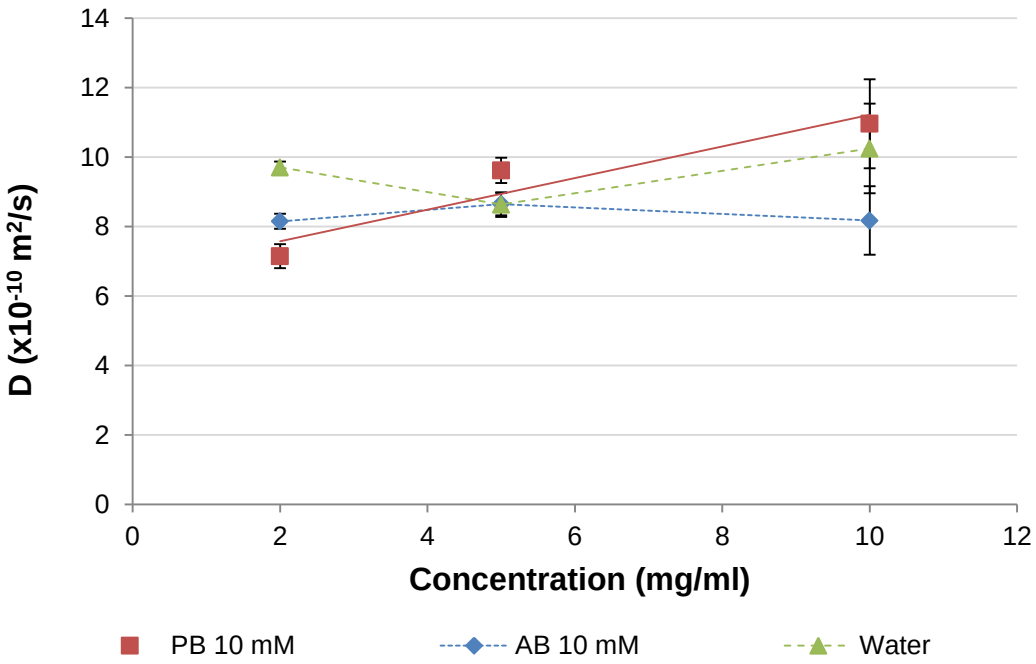

Figure 25 - DSS diffusion with varying concentration, in different solutions. The green triangles (▲) represent the water trials, the blue diamonds (♦) represent the trials in 10 mM acetate buffer and the red squares (■) represent the trials in 10 mM phosphate buffer. Measurements were performed at a controlled room temperature of 20ºC. For 10 mM PB, the equation of the regression was $y = 0.456x + 6.663$ and a correlation coefficient of R^2=0.905 (full red line). The dashed lines are used to guide the viewer's eye.

From Figure 25 it is possible to see that there is no apparent change in the diffusion coefficient of DSS for water and AB solutions. On the other hand, it is noticeable that there is an increase of the diffusion coefficient of DSS with the increase of the solute's concentration for the PB experiments.

Dextran sulphate is in the form of sodium salt. This means that for each single unit of glucose sulphate there are three units of sodium. When put in solution, the sodium salt will dissociate in solution, leading to an increase of the solution's ionic strength. To understand how the sodium present in DSS affects the solution's ionic strength, there was the need to estimate the solution's ionic strength with the concentration of DSS. To do this estimation, a single unit of the polyelectrolyte was considered (in Figure 9 there are two units and only one was considered) and the number of these units per mol was estimated (considering that a mol has a molecular weight of 5 000 g/mol and that each of the singular units has a MW of 420.18 g/mol). With these considerations, it was estimated that a mol of DSS has a total of 11,9 units (that we will considered to be 12 units, for simplification of further calculations). So, the concentration of sodium in solution will be 36 times the concentration of DSS. So, using equation (4.1) to estimate the ionic strength, for a concentration of DSS of 2 mg/ml, the contribution of DSS for the ionic strength will be of 7.2 mM. For a concentration of DSS of 5 and 10 mg/ml, the contribution for the ionic strength will be 18 and 36 mM, respectively.

For the lowest concentration, the highest value for the diffusion coefficient was obtained in water followed by the acetate buffer. This could suggest that DSS has a similar behaviour as lysozyme, meaning that the lower the ionic strength, the higher the diffusion coefficient value. However, upon further analysis of Figure 25, it is possible to see that for PB the trend follows as lysozyme (*D* increases with concentration) but for acetate buffer the *D* value is virtually constant, considering the error bars.

Using the same analysis used for proteins, the increase of the concentration of the polyelectrolyte could decrease the ion/polyelectrolyte ratio and this will lead to an enhanced importance of the electrostatic interactions between polyelectrolyte molecules. However, as explained above, the ion/polyelectrolyte ratio is kept constant in water since DSS has sodium ions in its composition. It is true, however, that increasing the concentration of DSS will decrease the ratio mentioned in the buffer solutions because for a smaller concentration, this ratio is influenced by the sodium in DSS molecules plus the ions in solution. For small concentrations of DSS the ratio of ion/polyelectrolyte is higher than when the concentration of DSS in solution is higher, as the influence of the sodium ions in DSS will have a bigger importance as the sodium in the DSS molecules will be more than the ones present in solution.

It could be hypothesized that the decrease in this ratio, with the increase in concentration, would have an effect in the diffusion coefficient with the increasing concentration of DSS. This could be an explanation for the increase observed for PB, however a similar phenomenon was expected to happen for AB and was not verified. With the presented data it is hard to postulate on which interactions are predominant for DSS diffusion, since the van der Waals interactions are highly unlikely to be something other than attractive (Israelachvili, 2011). It is, therefore, unknown what is happening in

solution and why there is an increase in the diffusion coefficient of DSS for PB and the same trend is not verified for AB.

It is suggested that this study is conducted under higher ionic strength conditions, to make sure that the solubilization of DSS doesn't influence the ionic strength and, therefore, try to understand how the concentration increase can influence the diffusion phenomena. This can, however, prevent the electrostatic interactions from happening due to the very high shielding that will occur. Another study that can be done is the use of dextran sulphate without being in a salt form. This would lead to a simplification of what happens in solution and a better understanding of the influence that each parameter (ionic strength and solute concentration) can have on the diffusion behaviour of dextran sulphate.

5. Conclusions and Future Work

The main objective of the present work was to determine the diffusion coefficient of different molecules, such as different proteins and a polyelectrolyte, using a microfluidics chip (H-cell). The determination of such parameters aims to understand how the diffusion of these molecules is affected by the conditions of the solution in which they're solubilized.

These determinations were done through a mass balance between two co-current fluid streams in the microfluidics channel. To have the mass balance, the parameters used were the initial concentration of solute in both inlet streams as well as the outlet concentration, the flowrates and the dimensions of the microfluidics channel. The first trials (sodium chloride diffusion measurement and protein attachment to the chip surface) allowed to conclude that the microfluidics channel predicted a diffusion coefficient with a 3.1% deviation from previously reported data, which allowed to conclude that the method was good for the prediction of this parameter. It is also possible to conclude that there is no significant loss of protein by attachment to the glass surface of the H-cell, as a recovery of virtually 100% was obtained.

From the different protein trials, with the variation of ionic strength, protein concentration, solvent's viscosity and different size of the proteins it is possible to conclude that the diffusion coefficient varies with the different solvent's conditions or solute's conditions (when varying the molecular weight and using different proteins, there's a variation not only in the size of the proteins but also in the intrinsic properties of these proteins, such as pI values). It was noticeable that the diffusion of proteins is affected by the ion/protein ratio (in this study it was studied by means of increasing concentration of protein), by the solution's ionic strength and viscosity and by the size of the diffusing proteins, as the larger proteins diffused slower than the smaller proteins. For the MW trials, it could be better to use the same molecule (same protein or even use a polyelectrolyte with different MW's). Using the same protein, the trials could be conducted with a protein that has monomers, dimers and trimers, after making sure that in solution these dimers and trimers kept existing avoiding the differences in intrinsic properties, as previously mentioned.

The diffusion coefficient of the polyelectrolyte (dextran sulphate sodium salt) is higher than the one for proteins as expected, due to the lower size of the molecule. The studies on the variation of diffusion of DSS with its concentration in solution were inconclusive and further studies need to be done in order to understand how the medium influences its diffusion coefficient.

We can conclude that there was a decrease in the diffusion coefficient of lysozyme with the increase of the ionic strength in a very small range of tested ionic strength. With this, we can conclude that lysozyme has electrostatic repulsive interactions with itself that are progressively reduced with the increase of ionic strength, due to the shielding of the surface charges of the proteins. It is important to have a study on the ionic strength where the ionic strength is solely controlled by the concentration of a strong electrolyte (sodium chloride, for example). With this, there is very high certainty in the ionic strength of the solution and a very broad spectrum of ionic strength values can be tested (the buffers

have an ionic strength according to their interacting species that are in a ratio good enough to keep a pH value constant). With the use of a strong electrolyte, the slope presented in the study can be better characterized and a comparison between the ionic strength from buffers and the strong electrolyte can be done, to fully understand if the salt species has an influence on the diffusion of proteins, besides the ionic strength.

The SIC trials were used to better understand what type of interactions were predominant between lysozyme molecules in the tested buffers, as a consequence of the concentration trials. The data suggests that there are repulsive interactions between lysozyme molecules in water and AB solutions and attractive interactions between lysozyme molecules in PB solution. Future work needs to be done to mimic the experimental conditions from the diffusion measurements in the SIC columns. To do it, the water solution during the SIC trials should have its pH adjusted to the value that the lysozyme solutions present when the diffusion measurements are taking place and the lysozyme solutions in water, AB and PB need to be tested in the different concentration conditions to better proof that there is a variation in the interactions between the molecules. By applying these changes, a quantitative estimation of these interactions can also be done (calculation of the B_{22} value).

Stokes-Einstein correlation still presents itself as a good prediction equation. Although it has some flaws regarding the accuracy of the prediction, it is very good to understand the trend of the variation of the diffusion coefficient with parameters like solvent's viscosity, temperature and solute's size (in the form of the hydrodynamic radius).

The device showed to be a reliable and high throughput method for the determination of the diffusion coefficient, whilst allowing to have a smaller sample and time consumption when compared to other determination methods. There is still some work to do to add value to this method. An automation of the sampling (with an autosampler, for example) could make an effective use of the presented technology to avoid human manipulation, reducing the risk of errors and saving human power. The miniaturization of the detection methods and the system pumps could elevate this technology to an even higher standard. If the detection methods could be small enough, we could have a lab-on-a-chip, with the conjugation of the H-cell and a suitable detection method.

6. References

Ahamed, T., Ottens, M., van Dedem, G. W., & van der Wielen, L. A. (2005). Design of self-interaction chromatography as an analytical tool for predicting protein phase behavior. *Journal of Chromatography A, 1089*(1), 111-124.

Anderson, J. L., Rauh, F., & Morales, A. (1978). Particle diffusion as a function of concentration and ionic strength. *The Journal of Physical Chemistry, 82*(5), 608-616.

Berger, J., Reist, M., Mayer, J. M., Felt, O., Peppas, N. A., & Gurny, R. (2004). Structure and interactions in covalently and ionically crosslinked chitosan hydrogels for biomedical applications. *European Journal of Pharmaceutics and Biopharmaceutics, 57*(1), 19-34.

Bergman, T. L., & Incropera, F. P. (2011). *Fundamentals of heat and mass transfer.* John Wiley & Sons.

Blanco, M. A., Perevozchikova, T., Martorana, V., Manno, M., & Roberts, C. J. (2014). Protein–protein interactions in dilute to concentrated solutions: α-chymotrypsinogen in acidic conditions. *The journal of physical chemistry. B, 118*(22), 5817.

Burke, P. (2000). Controlled release protein therapeutics: Effects of process and formulation on stability. Em D. Wise, *Handbook of pharmaceutical controlled release technology* (pp. 661-692). New York: Wise DL.

Cadman, A. D., Fleming, R., & Guy, R. H. (1982). Diffusion of lysozyme chloride in water and aqueous potassium chloride solutions. *Biophysical journal, 37*(3), 569-574.

Cleland, J. L., Powell, M. F., & Shire, S. J. (1993). The development of stable protein formulations: a close look at protein aggregation, deamidation, and oxidation. *Critical reviews in therapeutic drug carrier systems, 10*(4), 307-377.

Costin, C. D., & Synovec, R. E. (2002). Measuring the transverse concentration gradient between adjacent laminar flows in a microfluidic device by a laser-based refractive index gradient detector. *Talanta, 58*(3), 551-560.

Cussler, E. L. (2009). *Diffusion: mass transfer in fluid systems.* Cambridge university press.

De Wit, J. N. (1989). Functional properties of whey proteins. *Developments in dairy chemistry, 4*, 285-321.

Deshpande, K. S., Ahamed, T., Ter Horst, J. H., Jansens, P. J., Van Der Wielen, L. A., & Ottens, M. (2009). The use of self-interaction chromatography in stable formulation and crystallization of proteins. *Biotechnology journal, 4*(9), 1266-1277.

Du, J., Dai, J., Liu, J. L., & Dankovich, T. (2006). Novel pH-sensitive polyelectrolyte carboxymethyl Konjac glucomannan-chitosan beads as drug carriers. *Reactive and Functional Polymers, 66*(10), 1055-1061.

Dubin, S. B., Clark, N. A., & Benedek, G. B. (1971). Measurement of the rotational diffusion coefficient of lysozyme by depolarized light scattering: configuration of lysozyme in solution. *The Journal of Chemical Physics, 54*(12), 5158-5164.

Ebara, M., Hoffman, J. M., Hoffman, A. S., & Stayton, P. S. (2006). Switchable surface traps for injectable bead-based chromatography in PDMS microfluidic channels. *Lab on a Chip, 6*(7), 843-848.

Gros, G. (1978). Concentration dependence of the self-diffusion of human and Lumbricus terrestris hemoglobin. *Biophysical journal, 22*(3), 453-468.

Gull, N., Ishtikhar, M., Alam, M. S., Andrabi, S. N., & Khan, R. H. (2017). Spectroscopic studies on the comparative refolding of guanidinium hydrochloride denatured hen egg-white lysozyme and Rhizopus niveus lipase assisted by cationic single-chain/gemini surfactants via artificial chaperone protocol. *RSC Advances, 7*(45), 28452-28460.

Gupta, S., Jain, A., Chakraborty, M., Sahni, J. K., Ali, J., & Dang, S. (2013). Oral delivery of therapeutic proteins and peptides: a review on recent developments. *Drug delivery, 20*(6), 237-246.

Han, C. C., & Akcasu, A. Z. (1981). Concentration dependence of diffusion coefficient at various molecular weights and temperatures. *Polymer, 22*(9), 1165-1168.

Häusler, E. B. (2011). *Master's Thesis: Rapid diffusivity measurements using microfluidics.* Rijswijk: Technische Universiteit Delft.

Häusler, E., Domagalski, P., Ottens, M., & Bardow, A. (2012). Microfluidic diffusion measurements: The optimal H-cell. *Chemical engineering science, 72*, 45-50.

Haynes, C. A., Tamura, K., Korfer, H. R., Blanch, H. W., & Prausnitz, J. M. (1992). Thermodynamic properties of aqueous. alpha.-chymotrypsin solution from membrane osmometry measurements. *The Journal of Physical Chemistry, 96*(2), 905-912.

He, L. &. (2003). A novel correlation for protein diffusion coefficients based on molecular weight and radius of gyration. *Biotechnology progress, 19*(2), 544-548.

Hedberg, S. H., Heng, J. Y., Williams, D. R., & Liddell, J. M. (2015). Self-Interaction Chromatography of mAbs: Accurate Measurement of Dead Volumes. *Pharmaceutical research, 32*(12), 3975-3985.

Hotta, T., Nii, S., Yajima, T., & Kawaizumi, F. (2007). Mass Tansfer Characteristics of a Microchannel Device of Split-Flow Type. *Chemical engineering & technology, 30*(2), 208-213.

Il'ina, A. V., & Varlamov, V. P. (2005). Chitosan-based polyelectrolyte complexes: a review. *Applied Biochemistry and Microbiology, 41*(1), 5-11.

Israelachvili, J. N. (2011). *Intermolecular and surface forces* (3rd ed.). Academic press.

Jung, C., Lee, J., Kang, M., & Kim, S. W. (2014). Viscosity-dependent diffusion of fluorescent particles using fluorescence correlation spectroscopy. *Journal of fluorescence, 24*(6), 1785-1790.

Kakalis, L. T., & Regenstein, J. M. (1986). Effect of pH and salts on the solubility of egg white protein. *Journal of Food Science, 51*(6), 1445-1447.

Kamholz, A. E., & Yager, P. (2001). Theoretical Analysis of Molecular Diffusion in Pressure-Driven Laminar Flow in Microfluidic Channels. *Biophysics Journal, 80*(1), 155-160.

Kamholz, A. E., Weigl, B. H., Finlayson, B. A., & Yager, P. (1999). Quantitative analysis of molecular interaction in a microfluidic channel: the T-sensor. *Analytical chemistry, 71*(23), 5340-5347.

Kamiya, N., & Klibanov, A. M. (2003). Controling the rate of protein release from polyelectrolyte complexes. *Biotechnology and bioengineering, 82*(5), 590-594.

Kim, Y. C., & Myerson, A. S. (1996). Diffusivity of protein in aqueous solutions. *Korean Journal of Chemical Engineering, 13*(3), 288-293.

Kinch, M. S. (2015). An overview of FDA-approved biologics medicines. *Drug discovery today, 20*(4), 393-398.

Kuehner, D. E., Heyer, C., Rämsch, C., Fornefeld, U. M., Blanch, H. W., & Prausnitz, J. M. (1997). Interactions of lysozyme in concentrated electrolyte solutions from dynamic light-scattering measurements. *Biophysical journal, 73*(6), 3211-3224.

Medda, L., Monduzzi, M., & Salis, A. (2015). The molecular motion of bovine serum albumin under physiological conditions is ion specific. *Chemical Communications, 51*(30), 6663-6666.

Mehrer, H., & Stolwijk, N. A. (2009). Heroes and highlights in the history of diffusion. *Diffusion Fundamentals, 11*(1), 1-32.

Muheem, A., Shakeel, F., Jahangir, M. A., Anwar, M., Mallick, N., Jain, G. K., & Ahmad, F. J. (2016). A review on the strategies for oral delivery of proteins and peptides and their clinical perspectives. *Saudi Pharmaceutical Journal, 24*(4), 413-428.

Nguyen, N. T., & Wu, Z. (2004). Micromixers—a review. *Journal of micromechanics and microengineering, 15*(2), R1.

Patro, S. Y., & Przybycien, T. M. (1996). Self-interaction chromatography: a tool for the study of protein–protein interactions in bioprocessing environments. *Biotechnology and bioengineering, 52*(2), 193-203.

Pecora, R. (2013). *Dynamic light scattering: applications of photon correlation spectroscopy.* Springer Science & Business Media.

Pelegrine, D. H., & Gasparetto, C. A. (2005). Whey proteins solubility as function of temperature and pH. *LWT-Food Science and Technology, 38*(1), 77-80.

Prausnitz, J. M., Lichtenthaler, R. N., & de Azevedo, E. G. (1998). *Molecular thermodynamics of fluid-phase equilibria.* (3rd ed.). Pearson Education.

Quigley, A., Heng, J. Y., Liddell, J. M., & Williams, D. R. (2013). The accurate measurement of second virial coefficients using self-interaction chromatography: Experimental considerations. *European Journal of Pharmaceutics and Biopharmaceutics, 85*(3), 1103-1111.

Sacco, D., & Dellacherie, E. (1986). Interaction of a macromolecular polyanion, dextran sulfate, with human hemoglobin. *FEBS letters, 199*(2), 254-258.

Saluja, A., Badkar, A. V., Zeng, D. L., Nema, S., & & Kalonia, D. S. (2007). Ultrasonic storage modulus as a novel parameter for analyzing protein-protein interactions in high protein concentration solutions: correlation with static and dynamic light scattering measurements. *Biophysical journal, 92*(1), 234-244.

Shu, S., Zhang, X., Wu, Z., Wang, Z., & Li, C. (2011). Delivery of protein drugs using nanoparticles self-assembled from dextran sulfate and quaternized chitosan. *Journal of controlled release, 152*, 170-172.

Silva, D. F., Azevedo, A. M., Fernandes, P., Chu, V., Conde, J. P., & Aires-Barros, M. R. (2017). Determination of partition coefficients of biomolecules in a microfluidic aqueous two phase system platform using fluorescence microscopy. *Journal of Chromatography A, 1487*, 242-247.

Sorret, L. L., DeWinter, M. A., Schwartz, D. K., & Randolph, T. W. (2016). Challenges in predicting protein-protein interactions from measurements of molecular diffusivity. *Biophysical journal, 111*(9), 1831-1842.

Stone, H. A., Stroock, A. D., & Ajdari, A. (2004). Engineering Flows in Small Devices: Microfluidics Toward a Lab-on-a-Chip. *Annual Review of Fluid Mechanics, 36*, 381-411.

Struck, M. M. (1994). Biopharmaceutical R&D success rates and development times. *Nature Biotechnology, 12*(7), 674-677.

Sun, S. F. (2004). *Physical chemistry of macromolecules: basic principles and issues.* John Wiley & Sons.

Tanford, C. (1963). The interpretation of hydrogen ion titration curves of proteins. *Advances in protein chemistry, 17*, 69-165.

Tessier, P. M., Lenhoff, A. M., & Sandler, S. I. (2002). Rapid measurement of protein osmotic second virial coefficients by self-interaction chromatography. *Biophysical journal, 82*(3), 1620-1631.

van Leeuwen, M., & et al. (2009). Quantitative determination of glucose transfer between cocurrent laminar water streams in a H-shaped microchannel. *Biotechnology progress, 25*(6), 1826-1832.

Weinheimer, R. M., Evans, D. F., & Cussler, E. L. (1981). Diffusion in surfactant solutions. *Journal of Colloid and Interface Science, 80*(2), 357-368.

Wetter, L. R., & Deutsch, H. (1951). Immunological studies on egg white proteins IV. Immunochemical and physical studies of lysozyme. *The Journal of Biological Chemistry, 192*, 237-242.

Whitesides, G. M. (2002). Poly(dimethylsiloxane) as a Material for Fabricating Microfluidic Devices. *Accounts of chemical research, 35*(7), 491-499.

Whitesides, G. M. (2006). The origins and the future of microfluidics. *Nature, 442*(7101), 368-373.

Wong, D. W., Camirand, W. M., Pavlath, A. E., Parris, N., & Friedman, M. (1996). Structures and functionalities of milk proteins. *Critical Reviews in Food Science & Nutrition, 36*(8), 807-844.

7. Annexes

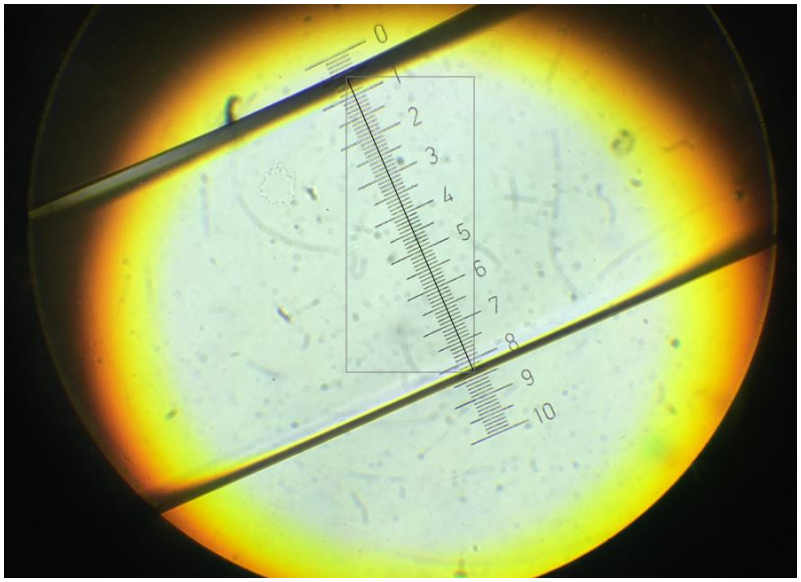

Figure 26 – Top view of the microfluidics H-cell. The channel was filled with air.

Figure 27 – Cross-sectional view of the microfluidics H-cell. The channel was filled with air. It is possible to see that the cross-section is not a perfect rectangle and that the bottom and top width do not have the same dimension.

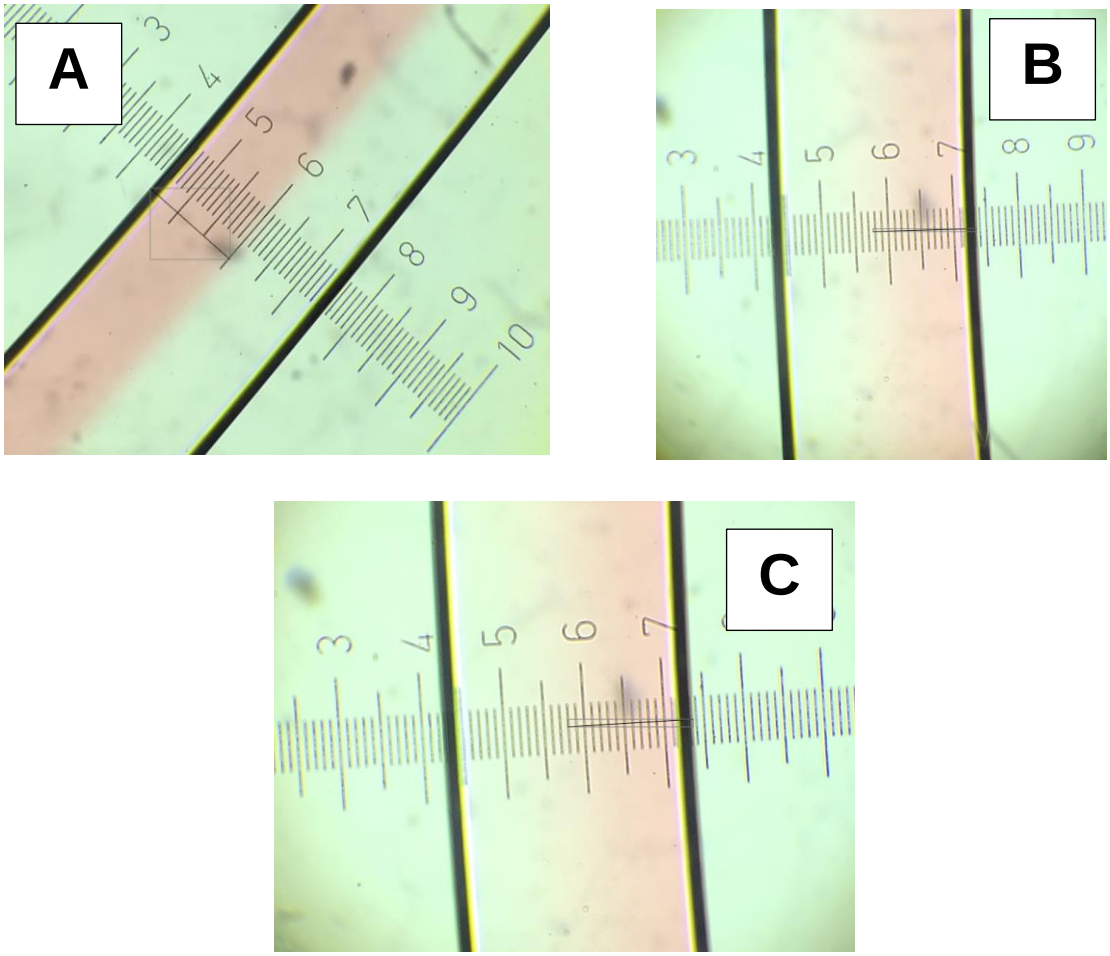

Figure 28 – Top view of the microfluidics H-cell. The input streams of the channel had distilled water (transparent) and red food colouring (red horseradish). In A we have the closest point of the inlet, in B we have a middle point and in C we have the closest point to the outlet possible. The flow rate used was 6 µl/min. It is possible to see that in B and C that the interface is no longer very clear, as some diffusion occurred during the residence time of both streams in the channel.

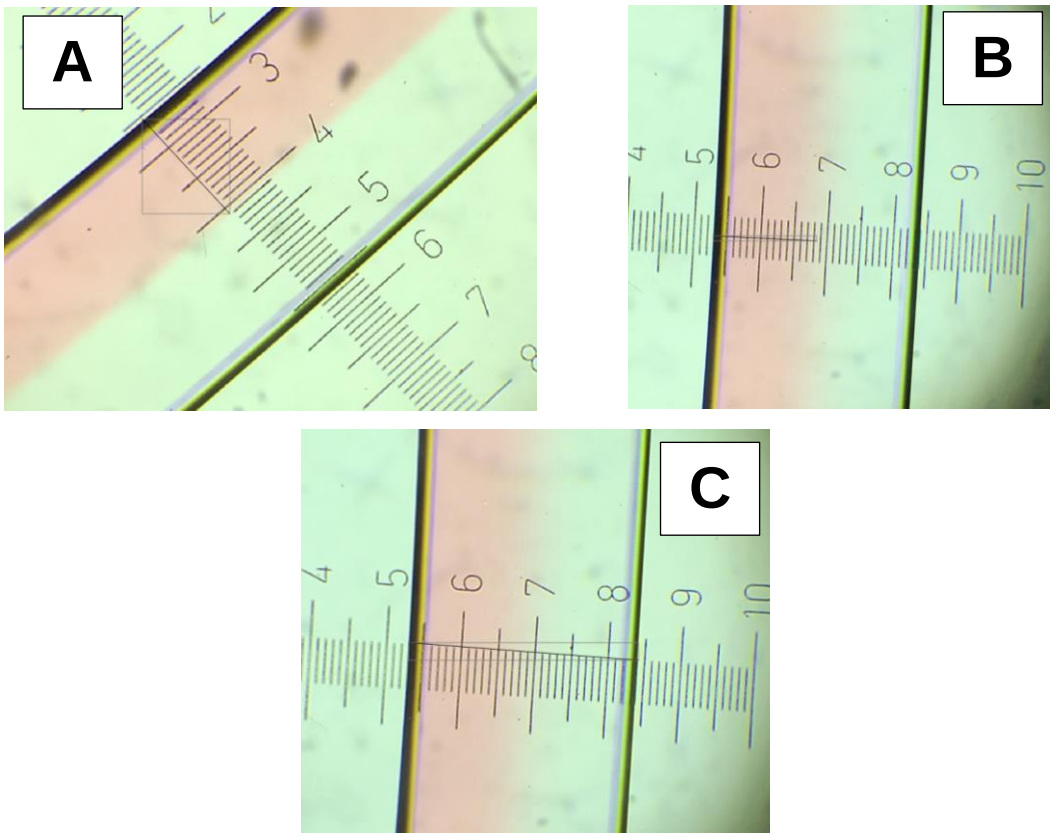

Figure 29 – Top view of the microfluidics H-cell. The input streams of the channel had distilled water (transparent) and red food colouring (red horseradish). In A we have the closest point of the inlet, in B we have a middle point and in C we have the closest point to the outlet possible. The flow rate used was 50 µl/min. It is possible to see that in B that there is still a very clear interface. In C, this interface is no longer very clear, but still clearer than the one in Figure 28, and this is because the residence time of the streams in the microchannel is much lower and, therefore, the horseradish molecules are not able to reach a steady-