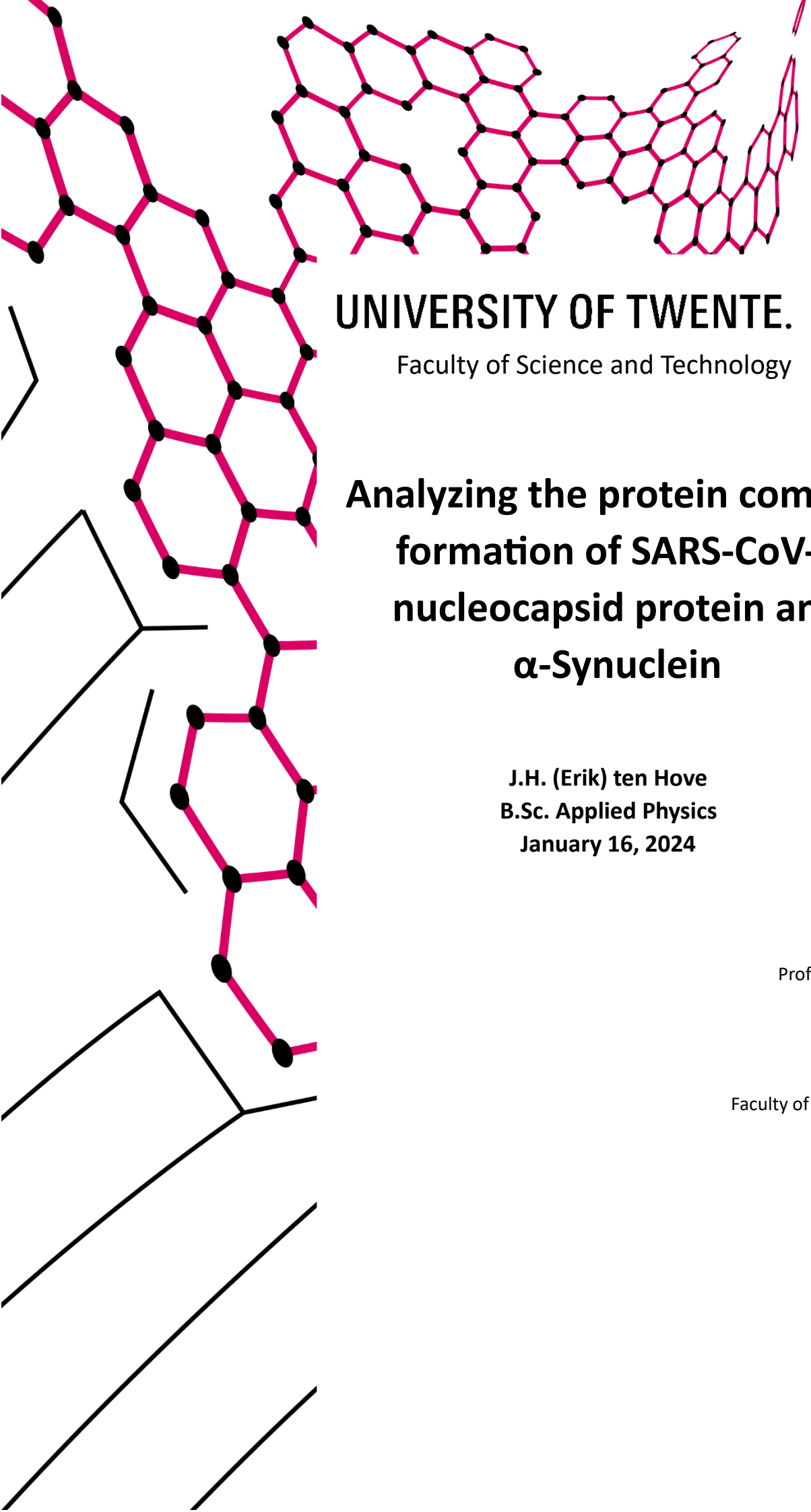


UNIVERSITY OF TWENTE.

Faculty of Science and Technology

Analyzing the protein complex formation of SARS-CoV-2 nucleocapsid protein and α-Synuclein

J.H. (Erik) ten Hove
B.Sc. Applied Physics
January 16, 2024

Supervisors:
Prof. Dr. M.M.A.E. Claessens
Dr. Ir. S.A. Semerdzhiev
Dr. M.H.G. Duits

Nanobiophysics Group
Faculty of Science and Technology
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

Abstract

Parkinson's disease is a neurodenerative diseasse caused by aggregation of α-Synuclein which can be triggered by a viral infection. Research suggest that SARS-CoV-2 could be such a viral infection and initiate amyloid formation causing Parkinson's disease. To gain insight in the relation between these two diseases on molecular niveau, this thesis aims to investigate the interactions between the nucleocapsis protein of SARS-CoV-2 and α-Synuclein further.

The described study attempts to determine the change in protein complex size for the nucleocapsid protein and α-Synuclein using fluorescence correlation spectroscopy. The proteins are labeled and in a low concentration prepared for a fluorescence measurements, where the proteins are detected when they move to the focal volume. This yields a fluorescent intensity trace that can be analyzed with an autocorrelation function to extract the diffusion coefficient of the protein complexes in the sample. The diffusion coefficient of α-Synuclein is found to be $85.22 \pm 56.47\ \mu m^2/s$, where the diffusion coefficient for the nucleocapsid protein is found to be $23.0 \pm 2.9\ \mu m^2/s$.

The binding events and affinity between the two proteins for different concentrations of α-Synuclein are also investigated using microscale thermophoresis. This measures a fixed concentration labeled nucleocapsid protein with different concentrations of an unlabeled ligand protein. This thesis assesses the binding events for both unlabeled nucleocapsid protein and α-Synuclein. The equilibrium dissociation constant found for the interaction between the labeled and unlabeled nucleocapsid protein is $K_D = 0.49 - 0.62\ \mu M$, where the dissociation constant for the interaction between the SARS-CoV-2 nucleocapsid protein and α-Synuclein turned out to be $K_D = 3.3122 \pm 1.1948\ \mu M$.

Acknowledgements

First of all, I would like to thank Slav Semerdzhiev for his help during the time that this project spanned, which was a lot longer than initially intended. You taught me a lot about professionally working in a laboratorium and how I had to critically evaluate my own findings. I felt I could ask you anything and this helped me tremendously with understanding proteins better, since this area of study was quite new to me. Thank you for being patient with me!

I also want to thank Mireille Claesens and Christian Blum who took the time to show me a lot of projects and helped me pick one that I found interesting. Also thanks for the meetings, results discussions and recommendations what to do next during my project.

The same goes for Gobert Heesink who helped me get familiar with the FCS setup and answered my questions regarding the Picoquant and the accompanying software.

Furthermore, I would like to thank Robert Molenaar and Kirsten Leijenhorst-Groener for familiarizing me with respectively the optical laboratory and the biochemical laboratory.

I also want to thank Michel Duits for being my external committee member and taking the time to read and evaluate my report.

Lastly, I want to thank everybody close to me for pushing me to finally finish this project and get my bachelors degree. I am very grateful for the support I got that got me through quite a rough time and finally got me to the point to finish my bachelor.

Contents

1 Introduction

Parkinsons disease (PD) is a neurodegenerative disease which is caused by protein aggregation in the brain of the protein α-Synuclein (αS). There is still a lot unknown about the factors that can trigger PD, but we know that approximately 15% of the patients has a first-degree relative with the disease. αS assembles into amyloid fibrils, which spread from cell to cell and reduce the brain function.[1, 2]

A viral infection can initiate the protein aggregation, which then causes a chain reaction of the formation of protein complexes, and thus triggers a neurodegenerative disease as PD. The relation between viral infections and neurodegenerative diseases such as PD has already been established by several different researchers.[3, 4, 5]

It would not be the first time that a pandemic viral infection coincides with a spike in neurodegenerative diseases. The Spanish flu was a very lethal influenza pandemic in 1918 that caused a lot of deaths not only because of the symptoms of influenza, but also since this pandemic occurred simultaneously with the encephalitis lethargica pandemic which caused a lot of post-encephalitic Parkisonism.[6, 7].

The recent Covid-19 pandemic as a consequence of Sars-CoV-2 infections also caused neurlogical complications, such as loss of smell. This is one of the premotor symptoms of PD and taken in combination with cases of Parkisonism discovered in patients after a Sars-CoV-2 infection suggested a connection between the viral infection and the neurodegenerative disease as had been observed in the past. In fact, research indeed already seems to suggest a relation between a Sars-CoV-2 infection and the development of Parkinson.[8, 9, 10]

To learn more about PD and its accelerated development after a Sars-CoV-2 infection it is important to understand the relation on a molecular level. It could also improve our understanding of the relation between viral diseases and neurodegenerative diseases in general and pave the way for preventive measures in the future as well as improving therapeutic methods to treat neurodegenerative diseases. Research shows that the Sars-CoV-2 nucleocapsid protein (N-protein) seems to interact with αS and consequently speeding up the aggregation process of αS.[11]

That is why this thesis will focus on the protein complex formation of αS under influence of Sars-CoV-2. The focus of this thesis is determining the size of the formed protein complexes of αS in the presence of N-protein for different concentrations of αS as well as analyzing the binding events and affinity between both proteins. To obtain this goal two different optical experimental methods are used the following goals for this thesis are formulated:

1. ***Determine the size of the protein complexes for both the nucleocapsis protein of Sars-CoV-2 and α-Synuclein.***

2. ***Analyze the binding events and affinity between the nucleocapsis protein of Sars-CoV-2 and α-Synuclein and express it as a function for different concentrations α-Synuclein.***

This thesis will attempt to obtain the first goal by using the optical technique Fluorescence Correlation Spectroscopy (FCS), which is a mathematical analysis of fluorescence measurements in a small volume with a low concentration of biomolecules. For this purpose N-protein is labelled with Alexa 488 and αS with Alexa 568 to determine each of the proteins individiual diffusion coefficients. The labeled N-protein is then mixed with different concentrations of wild type or "natural" αS (WT-αS) to find the diffusion coefficients of the formed protein complexes.

In order to achieve the second goal the optical technique Microscale Thermophoresis (MST) is used to analyze dilution series of labelled N-protein-488 with different concentrations of WT-αS to determine a binding curve between the two proteins. Since this process can also be influenced by the interaction of N-protein with itself, a MST experiment is also done for N-protein-488 with wild type N-protein as a ligand.

2 Theory

To obtain the formulated research goals it is crucial to first explain some underlying theoretical principles about the optical experimental techniques that are used and fluorescence in general. These principles explain how the proteins will behave, which will be crucial in modelling the behaviour of these proteins during experiments.

2.1 Fluorescence

The concept of fluorescence is important to this research, considering the biophysical techniques that are used conducting these experiments utilize fluorescence to locate proteins in a solution. When an electron is excited by an incoming photon, it is transitioned to a higher energy state. If this excited energy state is a singlet state, the electron is coupled to an electron in the ground state with opposite spin. The opposite spin allows the excited electron to fall back to the ground state, while emitting energy in the form of a photon. This emitted photon is the fluorescent light we observe. The energy transitions during the process of fluorescence can be seen in Figure 1. The fluorescent lifetime is the time between the excitation and the emission and is normally around 1 - 10 ns. Because of the short fluorescent lifetime, experiments require very precise and sophisticated instruments.

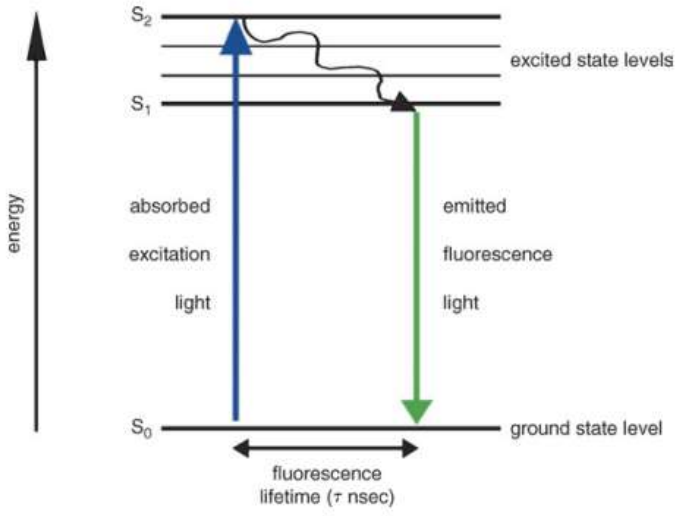


Figure 1: A Jablonski diagram of a fluorescent energy transition[12]

The figure above shows a so called Jablonski diagram. It describes the energy transition during an example of fluorescence. At first the electron is excited by blue light and absorbs its energy, thus exciting to a higher energy level S_1 or

higher. The system then relaxes causing the electron to fall back to the lowest excited state S_1. The electron then falls back to the ground state by emitting a photon, which can be seen as green light in this example. The energy of the photon and the color of the light are related by the Planck-Einstein relation:

$$E = \frac{hc}{\lambda} \tag{1}$$

where h is known as the Planck's constant. The color of the light is determined by the wavelength λ[13].

2.1.1 Protein labeling

The wild type version of the proteins used for this thesis, N-protein and αS, are not fluorescent, so they are chemically labeled in order to become fluorescent. N-protein is labeled with Alexa 488. This label is suited for the 488 nm laser, since it emits fluorescent light when excited with light of a wavelight around 490 nm. The αS is labeled using Alexa 568, which emits fluorescent light when excitated by the 568 nm laser.

2.2 Fluorescence Correlation Spectroscopy

Fluorescence Correlation Spectroscopy (FCS) is an optical technique used to analyze biomolecules in low concentrations. FCS can be used to analyze almost all physical phenomena that cause variations in fluorescence intensity. The focus of the technique is not primarily on the intensity of the fluorescence peaks but more on the fluctuations in intensity caused by minuscule deviations from the thermal equilibrium in the system.[14, 15] A confocal fluorescence microscope is used to focus on a small excitation volume in the protein solution. When a labeled protein moves through the excited volume it emits fluorescent light, as shown in Figure 2. This creates a fluorescence signal over time and FCS uses a quite straight forward mathematical analysis of the fluorescence signal that correlates the signal with itself over time to gather data on the movements of the biomolecules. This research will focus mainly on finding the diffusion constant of the protein complex, since this will help analyze the size and speed of the protein complexes from which conclusions about the formation of the protein complexes can be drawn.

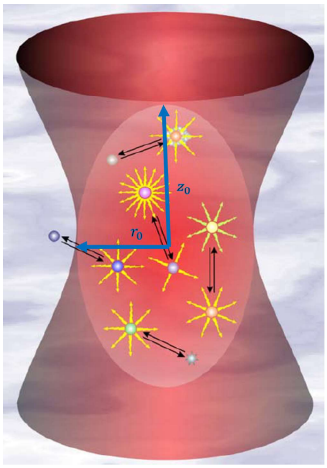


Figure 2: An excitation volume containing diffusing fluorescent particles[14]

The mathematical method used to analyze the fluorescence signal for FCS is called temporal autocorrelation. This means that a signal is analyzed with respect to itself after a certain lag-time τ. Usually the autocorrelation function is normalized so it is defined as[14]:

$$G(\tau) = \frac{\langle \delta F(t) \cdot \delta F(t+\tau) \rangle}{\langle F(t) \rangle^2} \tag{2}$$

Here the average fluorescence intensity is defined as $\langle F(t) \rangle$ and the fluctuation in the fluorescence at a time t as $\delta F(t) = F(t) - \langle F(t) \rangle$.

The measured fluorescence depends on a lot of individual parameters, including the detection efficiency and the optical transfer function of the microscope. It is almost impossible to determine all these parameters, so we use a Gaussian function in three dimensions to model the spatial distribution of light $W(r)$.

$$W(r) = e^{-2\frac{x^2+y^2}{r_0^2}} \cdot e^{-2\frac{z^2}{z_0^2}}$$

where r_0 is the lateral distance and z_0 is the axial distance of the focal volume, as is shown in Figure 2, and where $W(r)$ is decayed by $\frac{1}{e^2}$. The fluctuation in the fluorescence signal can then be written as

$$\delta F(t) = \int_V W(r) \delta(\eta \cdot C(r,t)) \cdot dV \tag{3}$$

Here η is the photon count rate per detected molecule per second and $C(r,t)$ is the fluorescent particle concentration at point r and time t.
Equation 2 can then be written as

$$G(\tau) = \frac{\iint W(r) W(r') \langle \delta(\eta \cdot C(r,t)) \delta(\eta \cdot C(r', t+\tau)) \rangle \, dV dV'}{\left(\int W(r) \langle \delta(\eta \cdot C(r,t)) \rangle dV \right)^2} \tag{4}$$

η is a constant which gives $\delta\eta = 0$. This gives

$$\delta(\eta \cdot C(r,t)) = C\delta\eta + \eta\delta C = \eta$$

Equation 4 can then be written as

$$G(\tau) = \frac{\iint W(r)W\left(r'\right)\left\langle \delta C(r,0)\delta C\left(r',\tau\right)\right\rangle dVdV'}{\left(\langle C\rangle \int W(r)dV\right)^2} \tag{5}$$

The term $\langle \delta C(r,0)\delta C\left(r',\tau\right)\rangle$ is called the density autocorrelation term and is calculated by:

$$\langle \delta C(r,0)\delta C\left(r',\tau\right)\rangle = \langle C\rangle \frac{1}{(4\pi D\tau)^{\frac{3}{2}}} \cdot e^{-\frac{(r-r')^2}{4D\tau}}$$

D is called the diffusion coefficient and it is a physical constant depending on the size of a molecule. The diffusion coefficient is an important tool in analyzing protein complexes, since it indicates the size of a measured particle.

Combining this with the definitions for the lateral diffusion time τ_D, which is the time that a particle stays in the focal volume and the effictive volume

$$\tau_D = \frac{r_0^2}{4 \cdot D} \tag{6}$$

$$V_{eff} = \frac{\left(\int W(r)dV\right)^2}{\int W^2(r)dV} = \pi^{\frac{3}{2}} \cdot r_0^2 \cdot z_0$$

gives the following autocorrelation function for freely diffusion fluorescent particles

$$G(\tau) = \frac{1}{V_{eff}\langle C\rangle} \cdot \frac{1}{\left(1+\frac{\tau}{\tau_D}\right)} \cdot \frac{1}{\sqrt{1+\left(\frac{r_0}{z_z}\right)^2 \cdot \frac{\tau}{\tau_D}}} \tag{7}$$

Equation 7 is the final version of the autocorrelation function used to model FCS measurements and determine the diffusion constant of the proteins.

This diffusion constant can then be used to find the size of the biomolecule using the Stokes-Einstein equation for diffusion.

$$D = \frac{k_B T}{6\pi\eta r} \tag{8}$$

Where k_B is the Boltzmann constant, T is the temperature in K, η is the dynamic viscosity in $kg\,m^{-1}s^{-1}$ and r is the radius of the spherical particle in m.

2.3 Microscale Thermophoresis

The second optical technique used in this thesis to analyze the protein complexes is Microscale Thermophoresis (MST). MST is a technique which uses the phenomenon called thermophoresis to analyze biomolecules[16, 17, 18, 19]. It changes the temperature in a very small volume of the sample which causes the molecules to start moving along this temperature gradient, meaning particles start to move from the warmer to the colder area. By measuring the fluorescence in the heated area, it is possible to observe a change in fluorescent particles in the focal volume which gives information about binding events, molecule size and other physical properties[20].

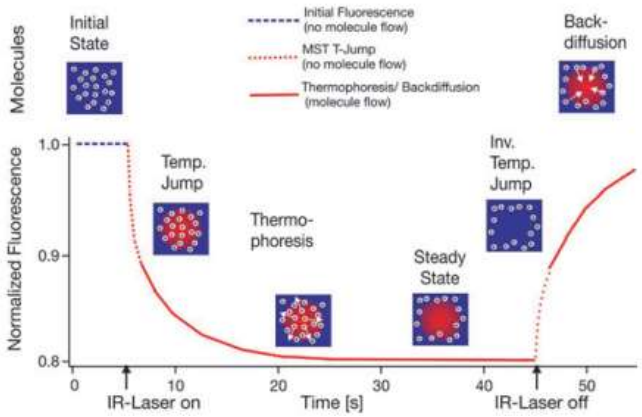


Figure 3: An example of a normalized MST fluorescence trace over time. The points in time where the heating laser is turned on and off are clearly detectable.[21].

Figure 3 shows the change in fluorescence due to thermophoresis. Before the infrared (IR) laser is turned on and there is no temperature gradient in the sample the particles are distributed equally over the sample volume. When the IR-laser is turned on, the temperature increases for a small part of the sample which causes the molecules in the sample to move along the temperature gradient and out of the heated volume. This is observed as a drop of fluorescence in the focal volume due to less fluorescent molecules being present. It is important that the focal volume of the fluorescence microscope is overlapping with the volume that is heated by the IR-laser, since we want to observe the change in a sample when the temperature changes.

2.3.1 The equilibrium dissociation constant K_D

We can use MST to quantify the binding between proteins. For a MST experiment we usually have two binding partners. One of them is labeled and kept at a fixed known concentration and the concentration of the other binding partner

varies per measurement.

The fluorescence trace as shown in Figure 3 is used to calculate the normalized fluorescence as follows:

$$F_{norm} = \frac{F_{hot}}{F_{cold}} \tag{9}$$

So the value of the fluorescence when the IR-laser is turned on is divided by the initial fluorescence to obtain the relative fluorescence which is used for quantifying the affinity between the proteins by analyzing the change in F_{norm} as a function of the concentration of the variable binding partner[17].

The normalized fluorescence consists of a linear superposition of particles in an unbound state and particles which are bound to the ligand. We can write F_{norm} then as

$$F_{\text{norm}} = (1 - FB)F_{\text{norm, unbound}} + (FB)F_{\text{norm, bound}} \tag{10}$$

where FB is the fraction bound, $F_{\text{norm, unbound}}$ is the normalized fluorescence of the proteins in unbound state and $F_{\text{norm, bound}}$ the normalized fluorescence of the proteins in bound state.

We can describe the process between the two binding partners to the formation of a protein complex as follows:

$$A + B \leftrightarrows AB$$

where A and B are the two binding partners and AB is a protein complex of the two proteins.

The ratio between the concentrations of the proteins in unbound state and the concentration of the protein complex is called the equilibrium dissociation constant K_D.

$$K_D = \frac{[A]_{free}[B]_{free}}{[AB]} \tag{11}$$

Since the free concentrations are unknown, they have to be derived from the known total concentrations

$$[A] = [A]_{free} + [AB]$$
$$[B] = [B]_{free} + [AB]$$

Equation 11 then becomes:

$$K_D = \frac{[A]_{free}[B]_{free}}{[AB]} = \frac{([A] - [AB])\,([B] - [AB])}{[AB]} \tag{12}$$

K_D is then modelled by using the fraction bound, since

$$FB = \frac{[AB]}{[B]} = \frac{[A] + [B] + K_D - \sqrt{([A] + [B] + K_D)^2 - 4[AB]}}{2[B]} \tag{13}$$

The fraction bound is now expressed in the known concentrations $[A]$, $[B]$ and $[AB]$, so the only free model parameter is K_D. Since FB scales linearly with F_{norm}, it can be directly fitted to equation 10, which we then use to model the data for different ligand concentrations, so we can extract K_D.

3 Materials and Methods

In this chapter there will be a more in depth description of the materials and methods used in the experiments performed for this thesis. As explained before, there are basically two experimental methods used to observe the formed protein complexes between the N-protein of Sars-CoV-2 and α-synuclein. Both of these proteins will be further specified below. Furthermore the methods and limitations of both of the experimental techniques, Fluorescence Correlation Spectroscopy (FCS) and Microscale Thermophoresis (MST), will be discussed.

3.1 Protein samples and buffers

This thesis revolves around the interactions between two different proteins. One being the nucleocapsid protein (N-protein) of Sars-CoV-2 and one being α-synuclein (αS) which aggregation is involved in the pathology of PD. The N-protein is labelled with Alexa Fluor 488 and the αS is labelled with Alexa Fluor 568. All proteins were diluted using a buffer of 10 mM Tris.

3.1.1 Sample preparation

For the FCS measurements, the protein samples were droplets of approximately $50\mu L$. The sample was deposited on a 22×22 mm high precision cover glass with a thickness of 170 ± 5 μm Since the microscope was focussed on a small volume inside the droplet it is not necessary to subject the cover slip to an extensive cleaning process.

To obtain a measurement suitable for FCS analysis it is important that there are not too many fluorescent particles in the excitation volume at the same time. Hence, the protein concentration should not be to high and ideally range between 10 and 50 nM. For the experiments determining the individual diffusion coefficients of N-protein and αS, the labelled protein samples were diluted down to the desired concentration using Tris as a buffer. N-protein was labelled with Alexa 488 and αS with Alexa 568. For the experiments on the protein complexes consisting of a combination of both proteins a dilution series was made of WT-αS, with concentrations ranging from 7.81 nM to 1 μM. Then a constant concentration of N-protein-488, which concentration should be in the range suitable for FCS, was added to the dilution series.

Then for the MST measurements a dilution series was made consisting of 16 different concentrations of the ligand protein. The wild type N-protein started from 10 μM and then diluted 15 times, each time halving the proteins concentration, thus the lowest concentration was 0.31 nM. The dilution series, consisting of 16 samples of approximately 30 μL with WT-αS started from 125 μM, so diluting 15 times gave a dilution series with a concentration of WT-αS ranging from 3.81 nM to 125 μM. Finally, the dilution series was finished by adding a constant concentration of N-protein-488. This fixed concenctration can be determined by testing some samples with N-protein-488 with concentrations between 50 nM and 500 nM and observing which give the best results. The

fluorescence intensity peak should be clear and between 200-1000 units. Select the concentration that gives that result to be the fixed concentration of the labelled N-protein. Then all the different dilution are added to the high precision capillaries and inserted in the Nanotemper Monolith.

3.2 Fluorescence Correlation Spectroscopy

3.2.1 Picoquant Microtime 200

As mentioned before, the microscope used for the FCS measurements is the Picoquant Microtime 200. This is a time-resolved confocal fluorescence microscope with has an extremely high temporal resolution and unique single molecule sensitivity.[22]

The specific setup used for the experiments of this thesis is shown in Figure 4. The experiments make use of two different fluorescent labels and thus uses the 488 nm and 560 nm lasers of the system to excitate the samples depending on the fluorescent label of the sample.

The emission light of the fluorescent proteins is then focussed on a confocal pinhole to dispose of the out-of-focus light, following a 488 nm long-pass filter to filter the excitation light out. When the 560 nm a notch filter of 560 nm is added before the confocal pinhole. The focused light is thereafter split by a 560 dichroic mirror, which ensures the emission light of the 560 nm fluorescent proteins ends up at the αS detector and the emission light of the 488 nm fluorescent proteins ends up at the N-protein detector. The N-protein detector is combined with a 520/35 bandpass filter, that allows a range of $502 - 538$ nm to pass and the αS detector is combined with a 620/60 bandpass filter that allows a range of $590 - 650$ nm to pass.

Figure 4: The Picoquant Microtime 200. 1=Notch filter. 2=488LP. 3=Confocal pinhole. 4=560 dichroic mirror. 5=620/60. 6=520/35. 7=αS detector. 8=N-protein detector.

3.2.2 Calibration

To succesfully perform a FCS analysis, it is important to first calibrate the setup before each series of measurements, in order to determine the size of the excitation volume. The 488 nm laser is calibrated using the dye Alexa 488, which has a known diffusion coefficient of 435 $\mu m^2 s^{-1}$. First a 50 μL droplet of Alexa 488 is measured to determine the measurements of the excitation volume from the FCS curve, using the known diffusion coefficient. The size of the excitation volume can then be used to model the FCS curves of the protein samples. Given that αS is labelled with Alexa 568 and measured with the 560 nm laser, a different calibration sample is needed. For this the dye Atto Rho was used wich has a fluorescence wavelength of 560 nm. The calibration method is the same as for Alexa 488.

Lastly, it is important that the excitation volume is located completely inside the sample droplet. For this purpose, live imaging of the microscope was used to find the droplet-coverslip border and focus the microscope above this border, making sure the excitation volume contains only the sample solution.

3.2.3 Analysis

The FCS experiments are performed using the SymPhoTime 64 software as grouped experiments. Every 30 seconds the FCS measurement is restarted to obtain 10 different results all stemming from the same focal volume. The size of this focal volume can then be determined using the result of the calibration measurement performed on the calibration dye. The software then models the 10 FCS curves using equation (7) to extract the average diffusion time of the diffusing species. Inserting this diffusion time in equation (6) then yields the diffusion coefficient of the protein complex inside the sample.

The model also takes into account the fluorescent blinking due to the triplet state. The fluorescent intensity trace was also filtered by imposing a upper threshold to filter the biggest intensity peaks, since this will most likely be protein aggregates. The value of this threshold was determined by inspecting the intensity trace and setting it so the biggest peaks were filtered out.

3.3 Microscale Thermophoresis

3.3.1 Nanotemper Monolith

The MST experiments are carried out using the Nanotemper Monolith. This instrument has two fluorescent channels, one for blue and one for green light and can measure the affinity for ligand concentrations ranging from 1 nM to 1 mM. The concentration of fluorescent proteins should also be in a range of nM to mM and the laser power can be adjusted to make sure the fluorescent intensity is not too high, leading to overexposure and not too low, which can not be detected sufficiently. The instrument can contain 16 samples per measurement and the capillaries containing these fluorescent protein samples have a volume up to 4 μL.[23]

These 16 capillaries are filled with protein samples containing fluorescently labeled N-protein at a fixed concentration and 16 different concentrations of WT-αS as a ligand. The Nanotemper Monolith then performs the same experiment on every capillary. First measure the fluorescence of the sample for 5 seconds, followed by 30 seconds of heating the sample using a LED, while still measuring the fluorescence. The instrument keeps measuring the fluorescence for 5 seconds after the heating LED is turned off.

3.3.2 Analysis

The goal of the MST analysis is determining the equilibrium dissociation constant K_D from 16 fluorescent traces of different ligand concentrations. The first 5 seconds of these traces are used to normalize the trace and called F_{norm} in equation (9). F_{hot} is chosen from the 30 seconds where the heating LED is turned on, on an interval that has a smooth fluorescence trace. The software then plots F_{norm} against the ligand concentration $[B]$.

The free model parameter K_D is then numerically determined by modeling the $F_{norm} - [B]$ plot, using equations (10) and (13), where $[B]$ is the ligand

concentration, using the known concentrations and calculated F_{norm}. This is already explained in more depth in chapter 2.3.1. The K_D-model can be used, since there is no cooperative binding, but merely a 1:1 interaction. Otherwise, the binding curve should be modeled by the Hill model.

The equilibrium dissocation constant measures the ratio between the free protein molecules and the protein molecules that are associated with each other, as is shown in equation (11). The lower this constant, the easier the binding between the present biomolecules.

4 Results

4.1 Fluorescence Correlation Spectroscopy experiments

The first part of this thesis consists of using FCS to analyze the protein complexes of N-protein with itself, αS with itself and a combination of the two proteins. The goals of these FCS experiments is trying to determine the diffusion coefficient of the formed protein complexes, since this relates to the size of the protein complex via Stokes-Einstein equation, and trying to determine how these diffusion coefficients depend on the concentrations of the proteins.

4.1.1 α-Synuclein

The first FCS experiment is focussing on αS protein complexes to determine the size of αS protein complexes in the absence of N-protein. For this purpose a series of 10 FCS measurements were conducted on 10 nM αS and the results used to model the FCS curve using equation (7), taking in account the blinking due to the triplet state, to get a value for the diffusion coefficient of αS. An example of the results of one of the measurements, is shown in Figure 5.

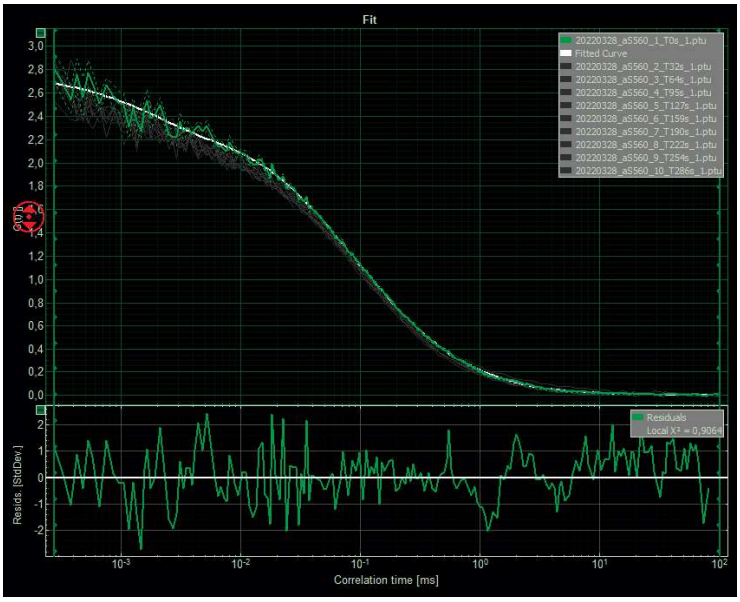


Figure 5: Results of series of FCS measurements on 10 nM αS. The first measurement is highlighted in green. The modeled FCS curve is shown in white.

The diffusion coefficient that was found modeling the series of FCS measurements using equations (6) and (7) is shown in Table 1.

Protein	D ($\mu m^2/s$)
α-Synuclein	85.22 ± 56.47

Table 1: Table containing the model result for the diffusion constant of the FCS measurements for 10 nM of αS.

This diffusion coefficient seems close to the value that was expected from literature[11] and thus a realistic value, only the uncertainty is quite large. The fitting of the FCS curves for 1 diffusing species seem to be quite accurate, as can be seen in Figure 5. None of the FCS curves shows strange behaviour and there seems to be no outlier.

4.1.2 N-protein

The experiment described in chapter 4.1.1 is then repeated for N-protein to also try to determine the diffusion coefficient of the formed N-protein complexes to determine the size of the N-protein in absence of αS. This can be used to verify that, when put together, the αS does in fact attach to the N-protein. 10 FCS measurements were conducted on 25 nM N-protein and thereafter modeled using equation (7) to find the diffusion coefficient. An example of the results is shown in Figure 6

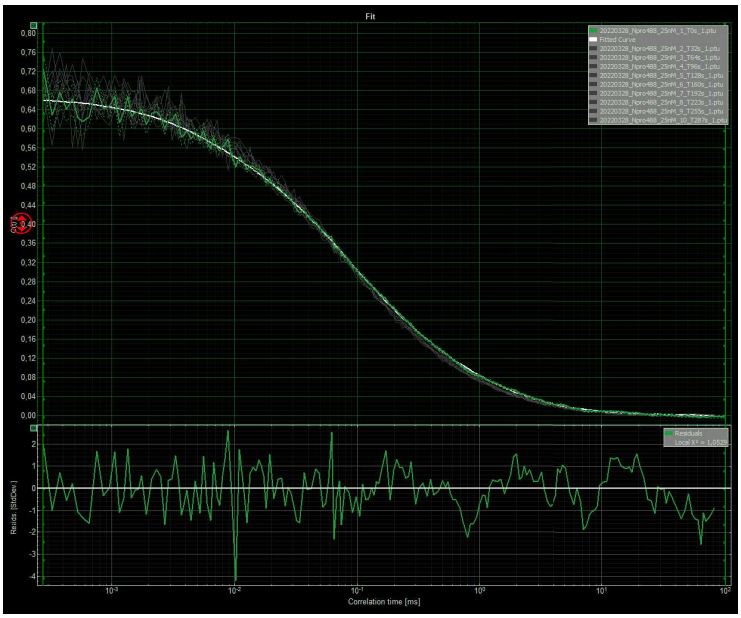


Figure 6: Results of series of FCS measurements on 25 nM N-protein. The first measurement is highlighted in green. The modeled FCS curve is shown in white.

The results are modeled using equation (7). One of the modeling parameters is n_{Diff}, which is the number of independently diffusing species. The model gave the best results for $n_{Diff} = 2$, where we get one high diffusion coefficient, in the range 10^2, and a diffusion coefficient in the range 10^1. These diffusion coefficients are shown in Table 2.

Protein	$\mathbf{D_1}$ **(µm²/s)**	$\mathbf{D_2}$ **(µm²/s)**
N-protein	23.0 ± 2.9	264 ± 32

Table 2: Table containing the model result for the diffusion constant of the FCS measurements for 25 nM of N-protein.

4.1.3 N-protein - α-Synuclein

The last experiment regarding the FCS part of this thesis focuses on the protein complex formation of N-protein with αS. This is done to investigate if the size of the protein complex changes for different concentration of αS. For this purpose a series of experiments were done on a fixed concentration of 25 nM of fluorescently labelled N-protein and different concentrations of wild type αS (WT-αS) ranging from 7.81 nM to 1 μM. For each concentration of αS, a series of FCS measurements was done to determine the diffusion coefficient of the protein complexes. The measured FCS curves are modeled using equation (7) and this model is then used to extract the diffusion coefficient using equation (6). Appendix A contains the FCS curves, which appear to be fitted quite accurately. The FCS curves are fitted for 2 diffusing species, since the sample contains bound and unbound αS. Table 6 shows the found diffusion coefficients that are plotted in Figure 7.

The computated diffusion coefficients of the protein complexes are then plotted against the concentration αS, which is displayed in Figure 7.

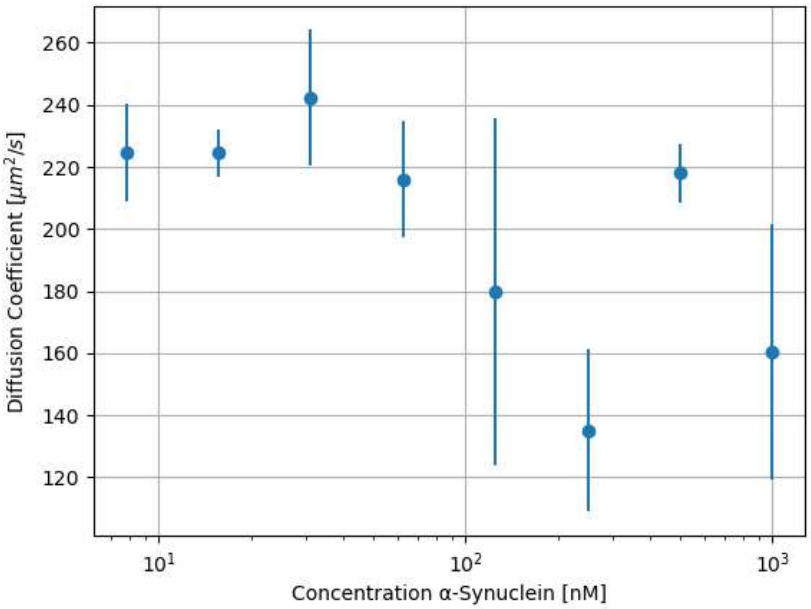


Figure 7: Diffusion coefficients of protein complex between 25 nM N-protein and WT-αS with a concentration ranging from 7.81 nM to 1 μM

The diffusion coefficients appear to be on the high side, since the protein complexes diffuse faster than αS aggregates. Some of the uncertainty margins are also quite high relative to the measured diffusion coefficient.

4.2 Microscale Thermophoresis experiments

For the second part of this thesis the protein complexes of N-protein and α-Synuclein were investigated using MST measurements. The goal is to analyze the binding events between N-protein and αS as well as taking into account the affinity N-protein has to bind to itself. For this purpose two experiments were performed, one using labeled N-protein with wild type N-protein as ligand and one with labeld N-protein with WT-αS as ligand.

4.2.1 N-protein - N-protein

First the binding events between N-proteins are analyzed using MST to determine the equilibrium dissociation constant that describes its affinity to bind with itself. The experiment uses a fixed N-protein at a concentration of 200 nM, since the peaks were optimized as explained in chapter 3.1.1. The ligand wild type N-protein concentration ranges from 0.31 nM to 10 μM. The MST power was set to 60% and 80%.

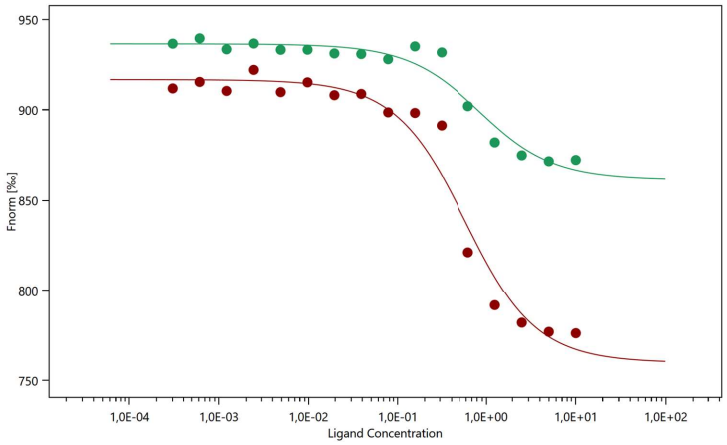


Figure 8: MST binding curves for N-protein - N-protein interactions with a MST power of 60% (green curve) and a MST power of 80% (red curve) and the ligand concentration of wild type N-protein in nM on the horizontal axis.

In Figure 8 the binding curves are shown for the N-protein interactions for two different MST power outputs. The F_{norm} for every ligand concentration is calculated as shown in equation (9). The K_D-model, as explained in chapter 3.3.2 is then used to model the binding curves and find the dissociation constant. In Table 3 the K_D values for both binding curves are shown to model the experimental results.

MST Power (%)	**Kd (µM)**
60	0.74081 ± 0.2537
80	0.48065 ± 0.1348

Table 3: Table containing the dissociation constant for modeling the binding curves as shown in Figure 8 for different MST power outputs.

4.2.2 N-protein - α-Synuclein

Chapter 4.2.1 shows the experimental results of the N-protein - N-protein interactions, but the main focus of this thesis is on the interactions and complex formation of N-protein with α-Synuclein. To investigate the binding relation between the two proteins, two experiments were done with WT-αS as a ligand with concentrations ranging between 3.81 nM and 125 μM. The N-protein - N-protein interactions are also taken into consideration, by performing two different experiments. In Figure 8 the binding curve clearly has two different plateaus. A high plateau for the lower ligand concentrations and a low plateau for the higher ligand concentrations. The MST experiments on N-protein interactions with αS are performed with a fixed concentration of N-protein. The first one with a low fixed concentration of N-protein chosen from the concentration

range spanned by the high plateau in Figure 8 and one with a high fixed concentration of N-protein picked from the range spanned by the low plateau. This way it is possible to investigate the interaction of αS with small N-protein molecules and also with bigger N-protein aggregates. The concentrations N-protein are based on the binding curve shown in Figure 8. The lower concentration is chosen from the upper level of the binding curve and set at 100 nM and the higher concentration is chosen from the lower level and set at 10 μM.

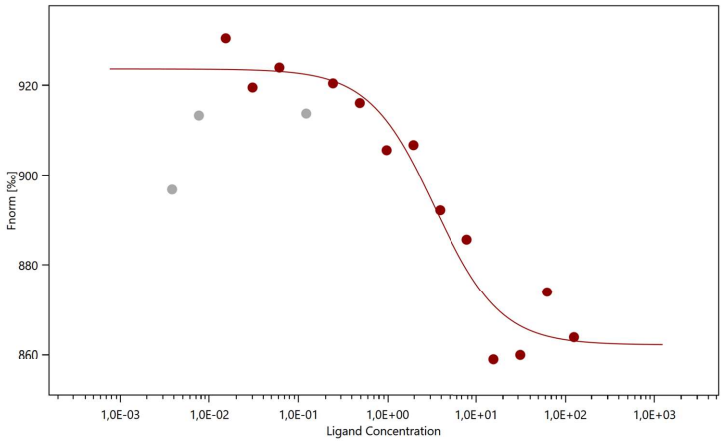


Figure 9: MST bindingcurve for N-protein - α-Synuclein interactions with fixed concentration fluorescently labelled N-protein of 100 nM at a MST power of 80 % and the ligand concentration αS in nM on the horizontal axis. The greyed out points are measured data points that are left out while modeling the K_D values, since they showed anomalies in their fluorescence trace.

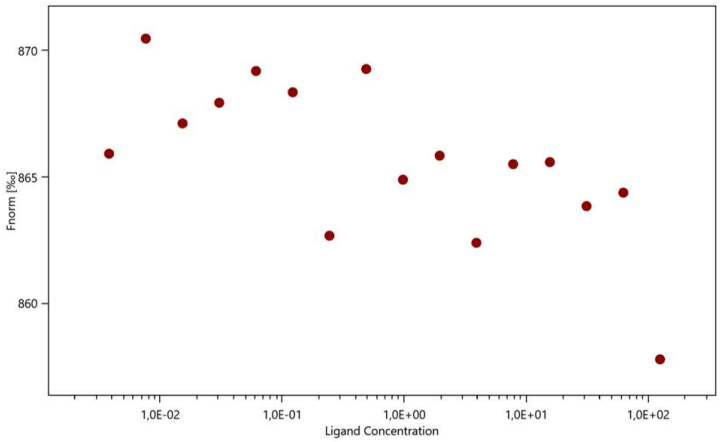


Figure 10: MST dose response plot for N-protein - α-Synuclein interactions with fixed concentration fluorescently labelled N-protein of 10 μM at a MST power of 80 % and the ligand concentration αS in nM on the horizontal axis.

Figures 9 and 10 show the binding curves for the interactions between N-protein and αS. The datapoints are the calculated F_{norm} for every ligand concentration using equation (9). These points are modeled using the Kd-model to determine the dissociation constant for the interaction between N-protein and αS. Table 4 shows the results of the modeling for both experiments.

[N-protein] (µM)	**Kd (µM)**
0.1	3.3122 ± 1.1948
10	None

Table 4: Table containing the dissociation constant K_D for both the fixed concentrations of N-protein used for modeling the binding curves as shown in Figures 9 and 10.

5 Discussion

5.1 Size of protein complexes

The first part of this thesis uses FCS to investigate the size of protein complexes for N-protein - N-protein interactions and N-protein - αS interactions. These protein sizes are derived from the diffusion coefficients that follow from modeling the FCS autocorrelation curves to a single diffusing species, for αS and the combination complexes of N-protein and αS, but modeling to two different diffusing species They are shown in Tables 1 and 2 for the single proteins and in Figure 7 for the combined N-protein and αS samples.

5.1.1 Size of α-Synuclein

Table 5 shows the comparison of the diffusion coefficient of αS found in the literature and according to the experimental results from this thesis. The experimental found diffusion coefficient is very close to the value found in literature[11], but there is a large error margin which follows from the modeling software SymPhoTime 64.

The diffusion coefficient is derived from modeling 10 different FCS measurements on the same focal volume. Modeling such a FCS measurement depends on a lot of different parameters, which are mostly unknown. This can increase the inaccuracy of estimating these parameters and can be an explanation for the proportion of the error margin.

Since the diffusion coefficient is a weighted average over 10 different FCS measurements, it could also be possible that one of the measurements is an outlier and throws the average off, but observing the raw data and FCS curves, as shown in Appendix A, it seems that there is no such outlier and the FCS curves are quite similar in form, so this explanation is not too plausible.

Protein	**D (literature)** $(\mu m^2/s)$	**D (experimental)** $(\mu m^2/s)$
α-Synuclein	86	85.22 ± 56.47

Table 5: Comparison diffusion coefficient αS between literature and experimental values

5.1.2 Multiple diffusing species in N-protein sample (n=2)

The diffusion coefficients experimentally found for the experiments on N-protein are shown in Table 2. Literature research did not yet yield a diffusion coefficient of SARS-CoV-2 N-protein to compare to, so it is difficult to comment on the reliability of the result via literature, so this has to be done purely through discussing the modeling process and result of SymPhoTime 64. As already mentioned discussing the size of αS, the software models several different parameters at once, using the size of the focal volume.

It is interesing to note that that modeling the FCS curve for only 1 diffusing species did not fit the experimental FCS curves as good as modeling for 2

different diffusing species, which rendered two different diffusion coefficients. This would implicate that there is another fluorescent protein diffusing through the excitation volume. Alexa Fluor 488 has a diffusion coefficient of 435 $\mu m^2/s$, so it is not probable that the second diffusing species is free dye. A possible explanation is that the fast diffusing species consist of N-protein monomers, so single protein molecules and the slower diffusing species are actually oligomers formed of multiple N-protein molecules linked together.

To analyze the FCS data, a fluorescence threshold was implemented to filter out bigger clumps of fluorescent proteins, but it is possible that smaller oligomers were still detected and not filtered out by this threshold.

5.1.3 Size of protein complex N-protein - α-Synuclein

The last series of FCS experiments were carried out to determine the influence of N-protein presence on the diffusion of αS. The result of these series of experiments with varying concentrations of ligand αS and a fixed concentration of labeled N-protein-488 are shown in Figure 7. The FCS curve is fitted for only 1 diffusing species.

Figure 7 shows the diffusion coefficient of the protein complex logarithmically plotted against the concentration αS, including the error margins calculated by the modeling software. The values of the diffusion are of the same order of magnitude as the high diffusion coefficient found in the FCS experiments on N-protein. That seems to be on the high side, since the proteins are expected to form protein complexes and thus should be diffusing slower than they do as monomers.

Another reason for strange behaviour of the proteins can be the incubation time of the samples. The dilution series with different varying ligand concentrations were all prepared at the same time, but could not be tested at the same time, creating a difference in incubation time for the different solutions.

The data set also seems to show a binding curve with an upper plateau for the lower ligand concentrations and a lower plateau for the higher ligand concentrations, so the protein complexes in the sample seem to slow down for a high enough ligand concentrations. The K_D value appears to be $\sim 120\ \mu m^2/s$ as is shown fitting the data using $K_D = 124\ nM$ in Figure 11.

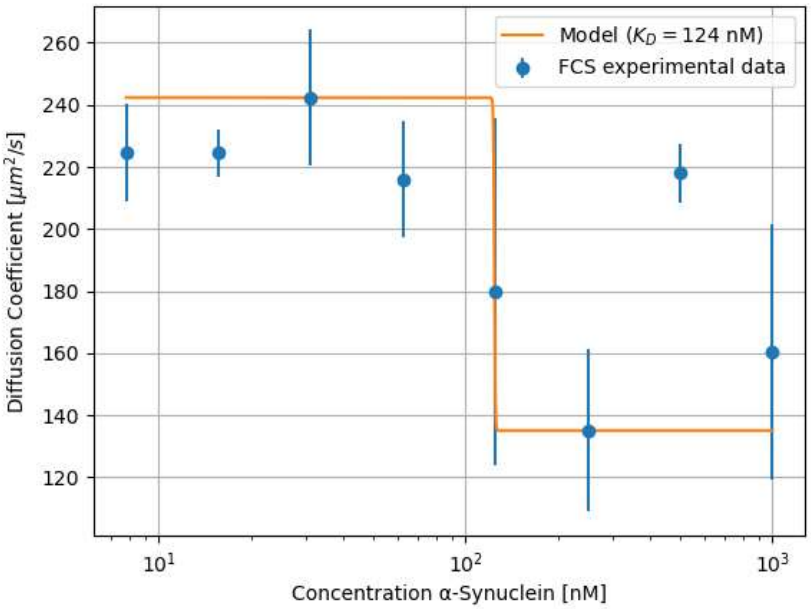


Figure 11: Experimentally determined diffusion coefficients (blue) for a fixed concentration of fixed N-protein-488 and a variable ligand concentration of αS fitted using $K_D = 124\,nM$ (orange).

5.2 Binding curves

Where the first part of this thesis focuses mostly on the size of N-protein and αS and protein complexes formed by these two proteins, there already seemed to be indications for affinity between N-protein and αS, as is shown by the binding curve in Figure 11. The MST experiments attempt to further investigate the interactions between these proteins. The results for the experiments regarding the interactions between labeled N-protein-488 and wild type N-protein are shown in Figure 8 and Table 3 and the results for the experiments on labeled N-protein-488 with WT-αS as a ligand are shown in Figures 9 and 10 as well as in Table 4.

From the experiments it seems that N-protein has quite a strong affinity to bind to itself, with a K_D value that lies in a order of magnitude of $10^2\,nM$. The binding curve in Figure 8 shows two plateaus. The lower ligand concentrations yield a higher relative fluorescence intensity and vice versa for the higher ligand concentrations. To test the affinity to αS, a MST experiment is performed with WT-αS as ligand and a fixed concentration N-protein from the lower ligand concentration plateau, where the N-protein will interact little with itself and the same is done with a fixed concentration N-protein from the higher ligand

concentration plateau, where the N-protein will definitely also interact with itself.

The smaller concentration of N-protein in combination with αS had a dissociation constant $K_D = 3.3122 \pm 1.1948\ \mu M$. This indicates that N-protein forms protein complexes with αS, although the interaction is almost a factor of 10 times weaker than the interaction N-protein has with itself. This also follows from the result of the MST experiment with the high fixed concentration of N-protein. Figure 10 clearly shows that the data can not be properly fitted, and the relative fluorescence intensity for all ligand concentrations spread around 845‰, so the protein complex is in all likelihood dominated by N-protein complexes. It is possible that the αS still forms complexes attached to N-protein, but this experiment can not clarify this, since the diffusion coefficients for N-protein and αS already showed that the N-protein oligomers are quite a lot larger than αS.

6 Conclusion

This thesis was built on two research goals. One was analyzing the size of N-protein and α-Synuclein by evaluating the diffusion coefficients of the biomolecules after fitting the measured FCS curves. In addition the size of protein complexes formed by N-protein in combination with αS as a ligand was to be analyzed using the same technique by fitting the FCS curves. The second part focused on analyzing the affinity and binding events of N-protein with itself and with αS as a ligand using MST and determining the equilibrium dissociation constant K_D of the interactions.

The diffusion coefficient of αS was determinted to be $85.22 \pm 56.47\ \mu m^2/s$, but had quite a large uncertainty. Following the FCS experiments on N-protein, the diffusion coefficient of this protein was quantified to be $23.0 \pm 2.9\ \mu m^2/s$, but the FCS curves had to be fit using a model for two diffusing species, so it could be that this coefficient applies to oligomers of N-protein instead of monomers. The FCS experiments on the protein complexed formed by a combination of N-protein and αS yielded a binding curve, where the diffusion of the protein complexes was shown as a funcion of the concentration of the ligand αS. The dissociation constant was established to be $\sim 124\ nM$.

The binding curves for N-protein interacting with itself are shown in Figure 8. The equilibrium dissociation consant of this interaction was found to be $K_D \approx 0.49 - 0.62\ \mu M$, so the interaction is quite strong. Especially compared to the interaction between N-protein and αS. The dissociation constant of this interaction for a low concentration of N-protein was found to be $K_D = 3.3122 \pm 1.1948\ \mu M$, which indicates a significantly weaker interaction strength than was the case for N-protein with itself. For a high concentration of N-protein, it was not possible to fit a binding curve through the data, since all data point seemed to flock around the same relative fluorescence intensity. This can be a consequence from the greater affinity to bind with itself of N-protein.

For further research it could be interesting to also perform MST experiments on labels αS with itself as ligend and with N-protein as a ligand as well. This could complete the picture of the interactions between N-proteins and αS that now has been sketched partly by this thesis. This way, it would also be possible to analyze the comparison between the aggregation of αS on its own and under influence of N-protein. Additionaly, it could also be interesting to look into the time it takes for αS to form aggregates and try to find if N-protein also influences the speed of aggregate formation.

It could be worthwile as well to use FCS more to better establish the affinity between N-protein and αS. This thesis showed already that measuring the diffusion of the protein complexes for different ligand concentration produces a binding curve, but a lot of improvement can be done to analyzing the FCS data and fitting the curves, such as determining more physical parameters beforehand, so the model has less free parameters and also attempt to filter out the oligomers as much as possible.

Bibliography

References

[1] Ali Samii, John G Nutt, and Bruce R Ransom. "Parkinson's disease". In: *The Lancet* 363.9423 (2004), pp. 1783–1793. DOI: 10.1016/s0140-6736(04)16305-8.

[2] Patrik Brundin, Ronald Melki, and Ron Kopito. "Prion-like transmission of protein aggregates in neurodegenerative diseases". In: *Nature Reviews Molecular Cell Biology* 11.4 (Apr. 2010), pp. 301–307. DOI: 10.1038/nrm2873.

[3] Collin M. Bantle et al. "Infection with mosquito-borne alphavirus induces selective loss of dopaminergic neurons, neuroinflammation and widespread protein aggregation". In: *npj Parkinson's Disease* 5.1 (2019). DOI: 10.1038/s41531-019-0090-8.

[4] A. J. Espay and K. K. Henderson. "Postencephalitic parkinsonism and basal ganglia necrosis due to Epstein-Barr virus infection". In: *Neurology* 76.17 (2011), pp. 1529–1530. DOI: 10.1212/wnl.0b013e318217e7dd.

[5] Hristina Vlajinac et al. "Infections as a risk factor for parkinson's disease: A case–control study". In: *International Journal of Neuroscience* 123.5 (2013), pp. 329–332. DOI: 10.3109/00207454.2012.760560.

[6] Charles P. Maurizi. "Why was the 1918 influenza pandemic so lethal? the possible role of a neurovirulent neuraminidase". In: *Medical Hypotheses* 16.1 (1985), pp. 1–5. DOI: 10.1016/0306-9877(85)90034-9.

[7] R. T. Ravenholt and William H. Foege. "1918 influenza, encephalitis lethargica, parkinsonism". In: *The Lancet* 320.8303 (Oct. 1982), pp. 860–864. DOI: 10.1016/s0140-6736(82)90820-0.

[8] Ingrid Faber et al. "Coronavirus disease 2019 and parkinsonism: A non-post-encephalitic case". In: *Movement Disorders* 35.10 (Aug. 2020), pp. 1721–1722. DOI: 10.1002/mds.28277.

[9] Mikhal E Cohen et al. "A case of probable parkinson's disease after SARS-COV-2 infection". In: *The Lancet Neurology* 19.10 (Oct. 2020), pp. 804–805. DOI: 10.1016/s1474-4422(20)30305-7.

[10] Marcelo Merello, Kailash P Bhatia, and Jose A Obeso. "SARS-COV-2 and the risk of parkinson's disease: Facts and fantasy". In: *The Lancet Neurology* 20.2 (Nov. 2020), pp. 94–95. DOI: 10.1016/s1474-4422(20)30442-7.

[11] Slav A. Semerdzhiev et al. "Interactions between SARS-COV-2 N-protein and α-synuclein accelerate amyloid formation". In: *ACS Chemical Neuroscience* 13.1 (2021), pp. 143–150. DOI: 10.1021/acschemneuro.1c00666.

[12] David Llères, Samuel Swift, and Angus I. Lamond. "Detecting Protein-Protein Interactions In Vivo with FRET using Multiphoton Fluorescence Lifetime Imaging Microscopy (FLIM)". In: *Current protocols in cytometry* (Nov. 2007). DOI: 10.1002/0471142956.cy1210s42.

[13] Joseph R. Lakowicz. *Principles of Fluorescence Spectroscopy*. Vol. 3. New York: Springer, 1983.

[14] Petra Schwille and Elke Haustein. *Fluorescence Correlation Spectroscopy*. Göttingen, Germany: Max-Planck-Institute for Biophysical Chemistry.

[15] Antonie J.W.G. Visser and Mark A. Hink. "New Perspectives of Fluorescence Correlation Spectroscopy". In: 9 (1 Oct. 1998). DOI: 10.1023/A:1020595926133.

[16] Stefan Duhr and Dieter Braun. "Why molecules move along a temperature gradient". In: *The Proceedings of the National Academy of Sciences (PNAS)* (Oct. 2006). DOI: 10.1073/pnas.0603873103.

[17] Susanne A.I. Seidel et al. "Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions". In: *Elsevier* (Dec. 2012). DOI: 10.1016/j.ymeth.2012.12.005.

[18] Thomas H. Scheuerman et al. "On the acquisition and analysis of microscale thermophoresis data". In: *Elsevier* (Mar. 2016). DOI: 10.1016/j.ab.2015.12.013.

[19] Peter Geelhoed et al. "Thermophoresis". In: *Encyclopedia of Microfluidics and Nanofluidics*. Ed. by Dongqing Li. Boston, MA: Springer US, 2013, pp. 1–6. ISBN: 978-3-642-27758-0. DOI: 10.1007/978-3-642-27758-0_1582-2. URL: https://doi.org/10.1007/978-3-642-27758-0_1582-2.

[20] Moran Jerabek-Willemsen et al. "MicroScale Thermophoresis: Interaction analysis and beyond". In: *Elsevier* (Mar. 2014). DOI: 10.1016/j.molstruc.2014.03.009.

[21] Moran Jerabek-Willemsen et al. "Molecular Interaction Studies Using Microscale Thermophoresis". In: *ASSAY and Drug Development Technologies* 9 (4 Aug. 2011). DOI: 10.1089/adt.2011.0380.

[22] PicoQuant. *MicroTime 200*. 2023.

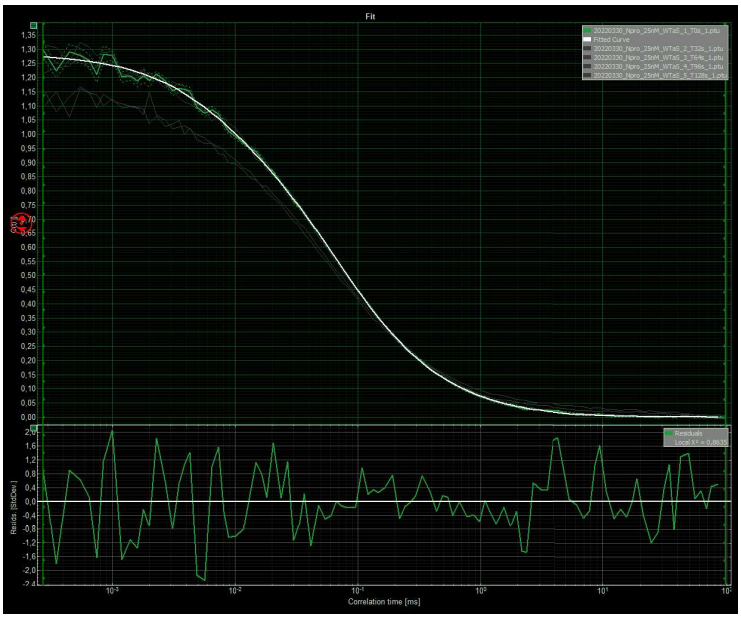
[23] NanoTemper. *Monolith*. 2023.

A FCS experimental results

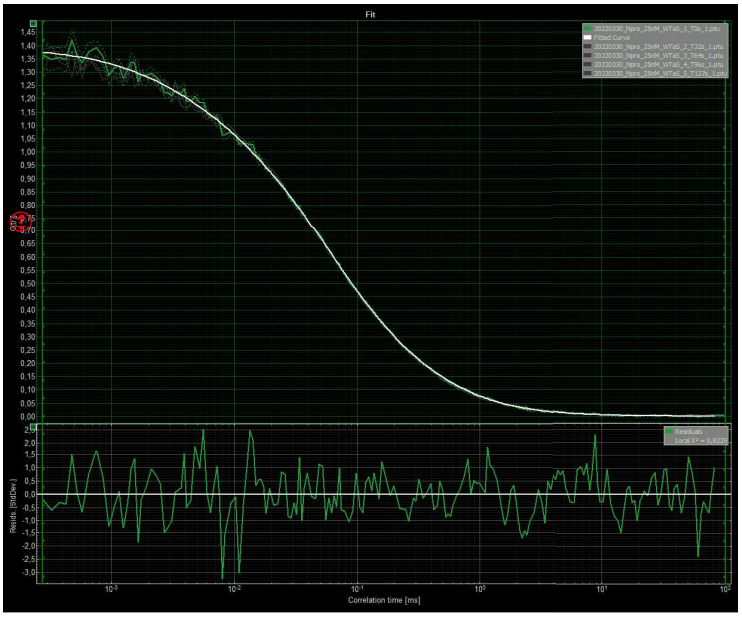
Table 6 shows the diffusion coefficients found for the different ligand concentrations of αS for the FCS experiment.

[**αS**] (nM)	**D** $(\mu m^2/s)$
1000.0	160.3474 ± 41.0951
500.0	217.9677 ± 9.3433
250.0	135.0787 ± 26.1693
125.0	179.8206 ± 55.7858
62.5	215.9386 ± 18.8028
31.25	242.3435 ± 21.8371
15.625	224.5734 ± 7.6196
7.8125	224.5884 ± 15.517

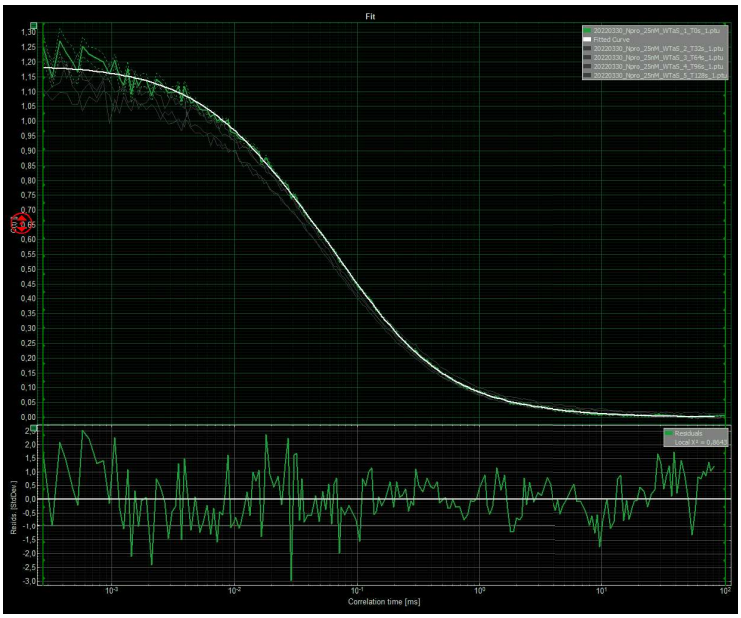
Table 6: Diffusion coefficient for protein complexes between N-protein and αS as a function of concentration of ligand αS.



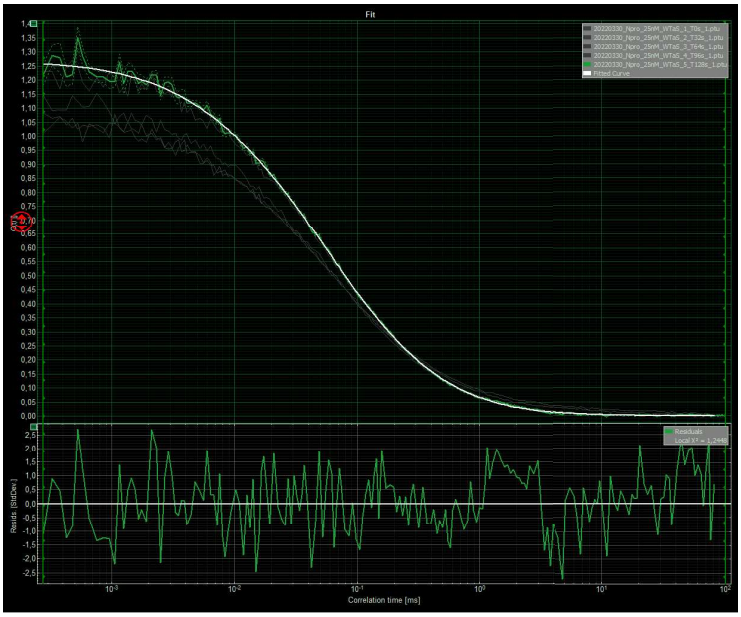
(a) $[\alpha S] = 1000.0\ nM$



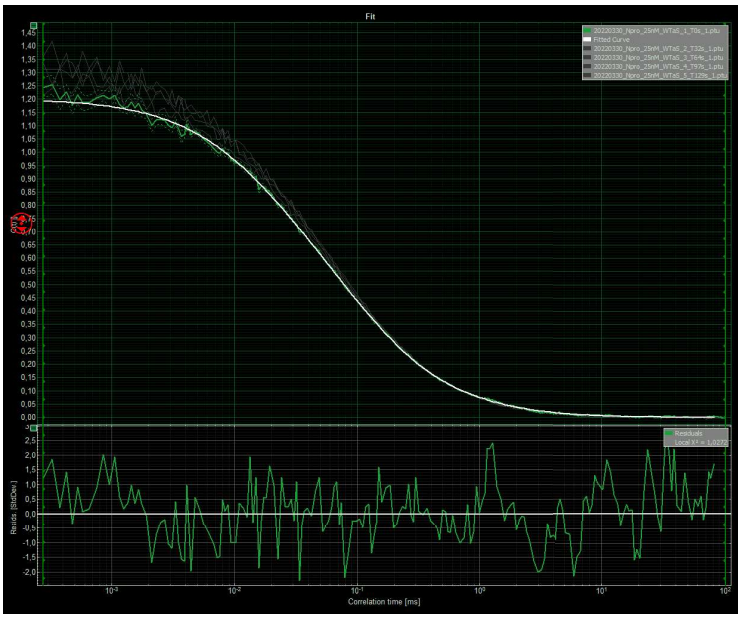
(b) $[\alpha S] = 500.0\ nM$



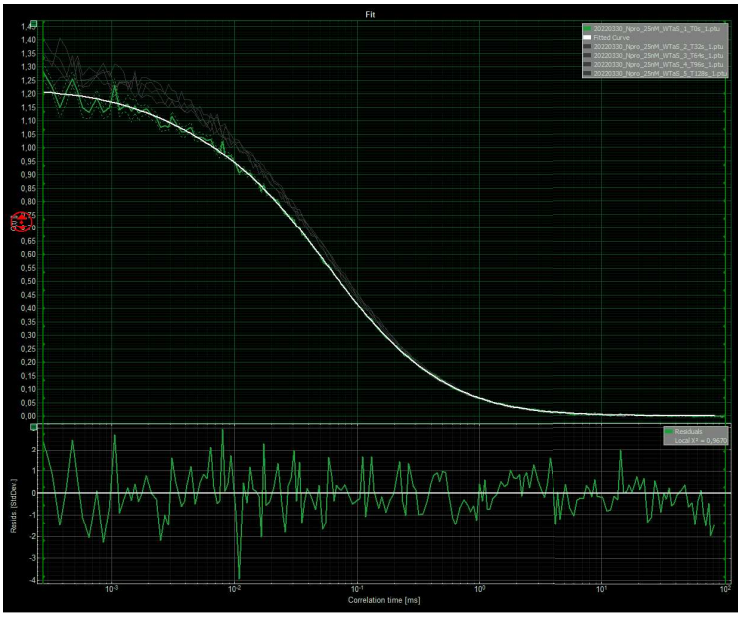
(c) $[\alpha S] = 250.0\ nM$



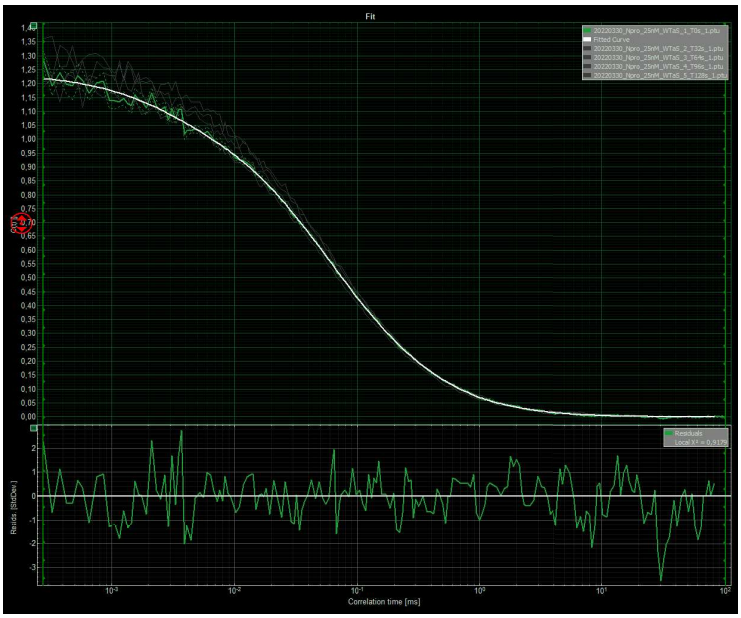
(d) $[\alpha S] = 125.0\ nM$



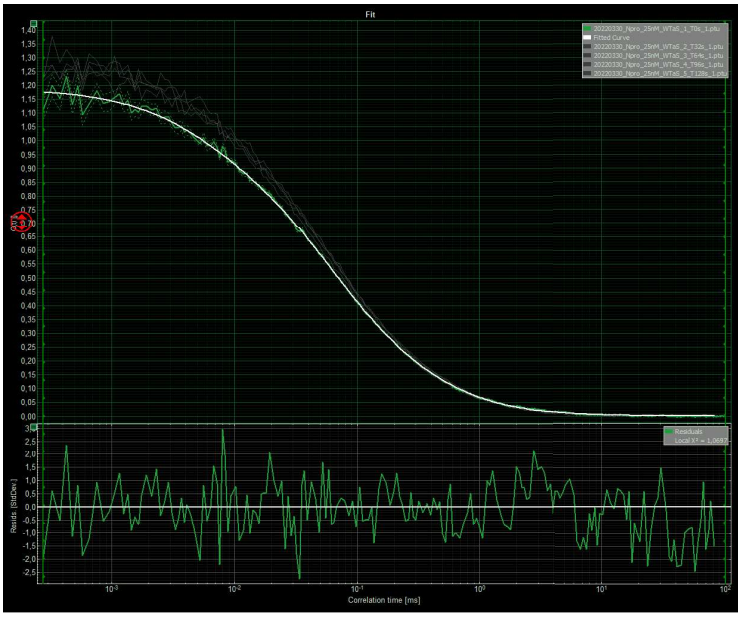
(e) $[\alpha S] = 62.5\ nM$



(f) $[\alpha S] = 31.25\ nM$



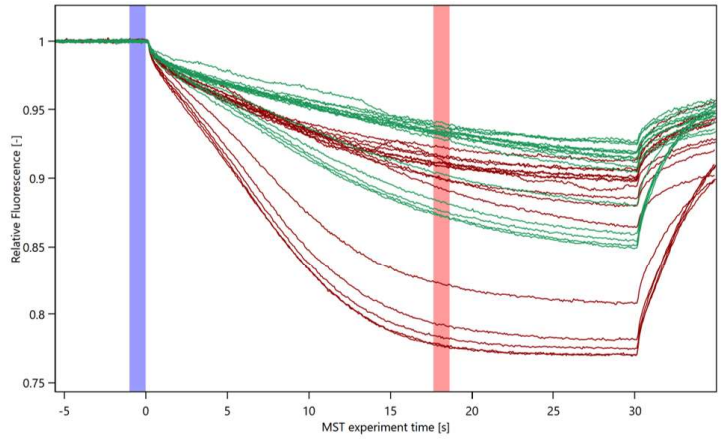
(g) $[\alpha S] = 15.625\ nM$



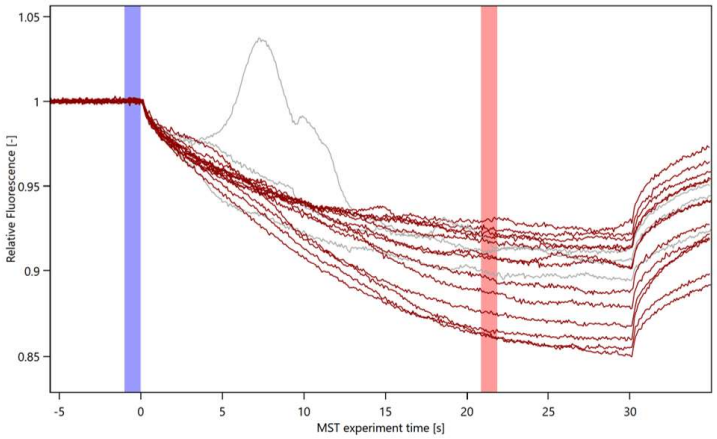
(h) $[\alpha S] = 7.8125\ nM$

Figure 12: Examples for fitting FCS curves for different ligand concentrations.

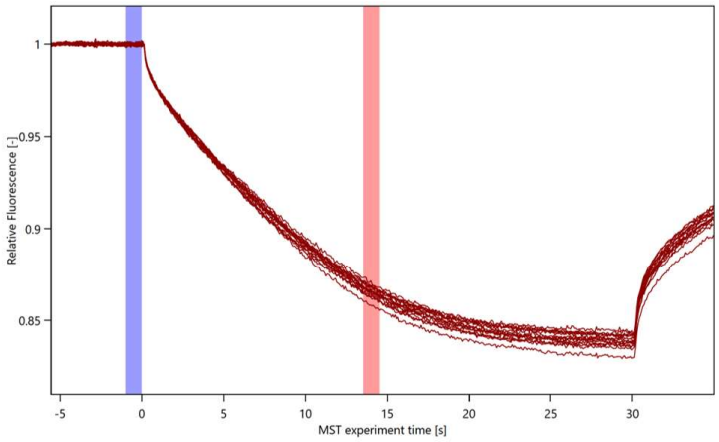
B MST experimental results



(a) Fluorescence intensity trace over time for N-protein - wild type N-protein interactions with a fixed concentration of $[N-protein-488] = 200\ nM$ for a MST power of 60% (green) and a MST power of 80% (red



(b) Fluorescence intensity trace over time for N-protein - αS interactions with a fixed concentration of $[N-protein-488] = 100\ nM$



(c) Fluorescence intensity trace over time for N-protein - αS interactions with a fixed concentration of $[N - protein - 488] = 10\ \mu M$

Figure 13: MST fluorescence intensity traces for all MST experiments.