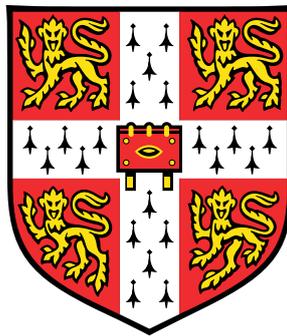


Development of Original and Notable Techniques for Protein Analysis, Native Identification, and Characterisation



Raphaël Philippe Bernard Jacquat

Department of Physics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

St Edmund's College

February 2022

I would like to dedicate this thesis not only to my family who supported me but to all the people I met and who through discussions made me evolve and shape my way of thinking. A special dedication to my fiancée Itzel Condado Morales who supports me even through the distance in the hard moments of doubts, but most importantly for the happy moments we share together.

Declaration

I hereby declare that, except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted, in whole or in part, for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgements. This dissertation contains fewer than 60,000 words including appendices, bibliography, footnotes, tables and equations.

Raphaël Philippe Bernard Jacquat
February 2022

Acknowledgements

I would like to thank Professor Tuomas Knowles for his precious time that he dedicated to me, his insightful discussions on my work, and for giving me this opportunity to participate in such exciting projects. I would also like to thank Dr Thomas Müller, who has dedicated a lot of his time in my project, and for mentoring me. And I would like to thank Dr Georg Krainer for supervising and guiding different projects. The PhD would not have been the same without Dr Quentin Peter guidance and support during all these PhD years we shared together, and with whom a lot of the projects would have not been possible. Thanks for all the collaborator in the project with the university of Zürich. Matthias Schneider, Dr Magdalena Czekalska, Ewa Andrzejewska, Dr Kevin Baumann , Dr Zenon Toprakcioglu, Dr Tuuli Hakala, Timothy Welsh, Dr Tomas Sneideris, Greta Sneideriene, Jenny Fan, Dr Catherine Xu, Dr Will Arter and Dr Tadas Kartanas for helping in the lab and for the long discussions as funny as helpful which has given to me a different point of view. Lastly, I would also like to thank all the lab members, especially the Knowles group and CMD members, for making the lab a nice place to work in. A special thanks for my sensor CDT cohort, whose people give me advice from field outside my expertise.

I would like to thank the rigorous proofreading and feedback of Dr Itzel Condado Morales, Dr Georg Krainer, Jan Heck, Matthias Schneider, Dr Quentin Peter, Johanna Kölbl, Thank you all.

I could not have completed this work without the support of my friends in and outside Cambridge. Thank you Loïc Reymond, Luca Mascheroni , Francesca van Tartwijk, Lorena Gordillo Dagallier, Johanna Kölbl, Sebastian Horstmann, Charles Christensen, Paul Martens, Quentin Peter and many others who were nice to me in difficult times. To Itzel, for all the helps, the share and support during moment of doubts as well as the support during the hard time of distance. An important part of this work would not have been possible without you. Thank you for being you and making me happy.

Table of contents

List of figures	xiii
1 Introduction	1
1.1 Motivation: Protein characterisation in heterogeneous solutions	1
1.2 Amyloids	4
1.3 Microfluidic advantage and characterisation	4
1.3.1 Physics in microfluidics	5
1.3.2 Fabrication	6
1.4 Development of a separation capillary electrophoresis device combined with H-filter detection	7
1.5 Microscopy	8
1.5.1 Label free detection	8
1.5.2 Single molecule detection	9
1.6 Structure of this thesis	9
2 Coupled Capillary Electrophoresis and Diffusional Sizing	11
2.1 Introduction	12
2.2 Theory	13
2.2.1 Capillary Electrophoresis	13
2.2.2 Diffusional sizing with a H-filter	17
2.3 First generation device	18
2.3.1 Injection and separation	18
2.3.2 Multiple injection	20
2.3.3 Analysis and Results	21
2.3.4 The diffusional sizing method	21
2.3.5 Separation	21
2.3.6 H-filter, or size characterisation	23

2.4	Second generation CE/H-filter device	25
2.4.1	Design process	25
2.4.2	Results	26
2.4.3	Methods	33
2.5	Conclusion	33
3	Single molecule detection and limitations of microfluidic experiments	35
3.1	Single molecule confocal setup	36
3.1.1	Detection and representation	37
3.1.2	Limit of detection and heterogeneous measurement	38
3.2	Single Molecule Microfluidic Diffusional Sizing (smMDS)	39
3.2.1	Results	41
3.3	Micro Free Flow Electrophoresis	52
3.3.1	Results	53
3.4	Conclusion	56
4	Nano-Cavity diffusional sizing (NDS)	59
4.1	Introduction	59
4.1.1	Theory and Simulation	61
4.1.2	Experimental demonstration of the NDS approach	65
4.1.3	Conclusion	66
5	Interferometric scattering correlation microscopy - iSCORR	69
5.1	Introduction	69
5.2	Theory	71
5.2.1	iSCAT	71
5.2.2	Correlation function and iSCORR	73
5.2.3	Dynamic light scattering	75
5.2.4	Contribution of polarisation	76
5.3	Methods	76
5.3.1	The iSCORR work flow	76
5.3.2	Setup	78
5.3.3	Limit of detection	78
5.3.4	Data analysis	79
5.4	Results and discussion	80
5.4.1	iSCAT	80

5.4.2	iSCORR	80
5.5	Conclusion	83
6	Application of protein detection in seroprevalence studies of SARS-COV2	85
6.1	Introduction	86
6.2	Methodology	87
6.2.1	Cohort selection	88
6.2.2	Measure	90
6.3	Results and discussion	92
6.3.1	Cohort characterisation, and biases	92
6.3.2	TRABI	93
6.3.3	Control and comparison of TRABI	95
6.3.4	Temporal evolution of seropositivity prevalence in greater Zürich	96
6.4	Conclusions	100
6.5	Acknowledgement	103
7	Conclusions	105
7.1	Future Applications	106
	References	109

List of figures

1.1	Illustration of the amyloid beta precursor protein (image taken from the protein data bank http://pdb101.rcsb.org).	2
1.2	Mechanism of amyloid formation from the primary and secondary pathways.	4
2.1	illustration of electro osmosis flow in capillary without pressure driven flow	14
2.2	Comparison of peak capacity and time detection in a CE without a pressure driven flow	16
2.3	Schematic of H-filter	17
2.4	Custom-made device used for CE	19
2.5	Microfluidics chip combining CE separation and diffusional sizing.	19
2.6	Images of the injection process	20
2.7	Data analysis of the CE part results	22
2.8	Data analysis of the H-filter part results	23
2.9	Fluorescein sizing using the first design (CE combined with H-filter)	24
2.10	Second device work flow	27
2.11	Model used for second device. Reproducibility and dilution of electropherogram	28
2.12	Explanation of diffused/undiffused electropherogram curve, and relation to the size	30
2.13	Electropherograms of N122C α syn mutant monomer and large unilamellar vesicle of DOPS	31
2.14	Electropherogram of α syn oligomer	32
3.1	Single molecule setup description with example traces	38
3.2	Particles detection probabilities	40
3.3	Principle and implementation of single molecule Molecular Diffusional Sizing	43
3.4	Influence on sizing due to single molecule parameter	45

3.5	Detection limit of widefield microscopy	46
3.6	Single molecule diffusional sizing of nanoscale analytes and protein complexes	47
3.7	Binding curve at high affinity using smMDS technique	49
3.8	Heterogeneous mixture sizing using single molecule microfluidic diffusional sizing of self-assembly α -syn and HSA	51
3.9	uFFE chips and flow of the process to analysis	52
3.10	Example of applications for single molecule analysis with uFFE chips . . .	54
4.1	Principle of nanocavity diffusional sizing (NDS)	61
4.2	Theory and simulation of diffusion under nano-confinement.	63
4.3	Nano-fluidic diffusional sizing of single particles in solution.	67
5.1	Path of the light to focal plane	72
5.2	iSCAT & iSCORR microscope and Work flow of the iSCORR method . . .	77
5.3	Results of iSCAT images	81
5.4	Size analysis results of iSCORR	82
6.1	Pipeline of the study	88
6.2	Cohort characterisation.	94
6.3	Sample provenance	94
6.4	ROC curves and comparison with exiting assays	95
6.5	Evolution of CoV2 prevalence in a cohort of Zürich University Hospital (USZ) patients and donors from the blood donation service (BDS)	97
6.6	Post stratification and antibody waning	98
6.7	Prevalence estimation in two large cohorts.	99
6.8	Seroprevalence maps for municipalities in the canton of Zürich.	101

Chapter 1

Introduction

1.1 Motivation: Protein characterisation in heterogeneous solutions

Proteins are the building blocks of life. They regulate almost every function of living cells from their structure to catalytic activities through molecular motor or messenger transmission [125, 154]. As an illustration, actin filaments, which compose the cytoskeleton, serve as a support to myosine, a motor protein involved in muscle contraction. Another example of protein function is DNA polymerase, an enzyme engaged in the replication and reparation of DNA [43, 54]. Every protein is formed by a chain composed of amino acids. The chain folds into a tridimensional structure depending on the thermodynamic stability of the system, and the resulting shape will define the protein's function. To compose the chain, human cells have a pool of 20 different amino acids [121], and any living cell so far discovered has approximately as many as a human's. That is why, even though there exists a huge diversity of proteins, they share similar characteristics due to the composition of the score of amino acids [154]. In order to function properly, the proteins need to fold into the correct tertiary structure, and incorrect folding of the proteins may in some cases lead to aggregation which is associated with disease, such as neurodegenerative disorders and systemic amyloidoses [34]. In order to extend our knowledge about life, it is fundamental to study how proteins work and interact. There exist two main research directions: static studies focusing on composition and/or structure, and dynamic studies focusing on interactions. The former consist in determining whether a particular protein exists in a subsystem and, if so, to find the structure or determine the principal properties characterising the protein. It is thus essential to be able to distinguish one protein from another. Static studies are, for instance, used to detect

diseases through the specific proteins they produce, named biomarkers. On the other hand, dynamic studies focus on the evolution of a system with proteins: the kinetics of protein reactions. One example is the kinetics of protein aggregation associated with misfolding diseases. Globular proteins fold into a configuration which minimises their potential, termed the native state [33]. The native shapes lie in a local minimum, which corresponds to the state with the lowest free energy. Under certain circumstances, an unfolded or partially unfolded conformation can corrupt other “well-folded” proteins, which could then promote aggregation. These types of misfolded proteins appear in diseases such as Alzheimer’s, Parkinson’s and Creutzfeldt-Jakob’s, which are respectively associated with the proteins amyloid beta ($A\beta$), α -synuclein and the prion protein (PrP). In each of these diseases, autopsies have shown the presence of aggregates or plaques of proteins in the brain, but fall short of implying a cause/consequence relation [47, 88, 100, 102, 191, 193, 213]. An illustration of the amyloid beta precursor protein is shown in Figure 1.1. The amyloid hypothesis states that the formation and deposition of the amyloid beta peptide in the brain is the direct cause of Alzheimer’s disease [78]. The hypothesis is supported by genetic evidence in the case of Parkinson’s disease, where some mutations in the α -synuclein gene are directly associated with an accelerated aggregation of the protein [61]. However, recent evidence shows that the most toxic species are not the amyloid fibrils but rather the intermediate oligomers formed during the aggregation process [19], though in some diseases like systemic amyloidoses, the fibrils are the direct cause of organ dysfunction. Thus the aggregation process should be prevented in any case. Cells have their own protective mechanisms against the pathogenic self-assembly of proteins, and they can consume energy in order to reconfigure proteins into their native functional shape, as in the case of molecular chaperones [74] or via degradation with the proteasome system [85].

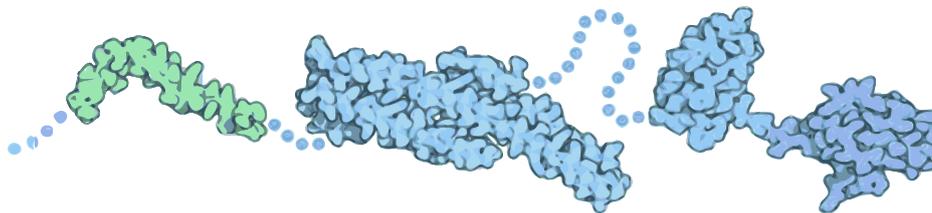


Fig. 1.1 Illustration of the amyloid beta precursor protein (image taken from the protein data bank <http://pdb101.rcsb.org>).

To distinguish between the different protein entities, several physical properties can be probed such as the entity structure, charge, size, mass, hydrophobicity and affinity. Currently, many techniques are exploited to determine these properties:

Size and structure can be directly measured by depositing the sample on a surface, and probing with an atomic force microscope (AFM) [166], low-energy electron holography techniques [128] or X-ray diffraction. The last technique usually requires protein crystal formation [178], and this is a major limitation in the study of proteins. These direct probing approaches are strong against artefacts and allow a good overview of the protein. However, they are often not cost effective, and involve adsorption on a surface where proteins are not in solution, which could lead to a change in their configuration that does not correspond to their biologically relevant state. Otherwise, NMR or cryo-electron microscopy (CryoEM) allow the resolution of protein structure in solution to a certain extent [161, 217] but require a purified sample. NMR has the main disadvantage that resolving the spectra of large proteins becomes quite challenging, as well as the need for a high concentration of pure protein which in some cases is difficult to obtain [64]. CryoEM, on the other hand, has a lower limit in protein size. The composition of a protein can also be determined with a mass-spectrometer, which provides information on the mass/charge ratio of sub-units of the protein. However, in addition to a high price tag, the main disadvantage is a non-native view of the specimens, which are sputtered in gaseous form in order to be analysed.

The size analysis of proteins can be probed indirectly in bulk by dynamic light scattering (DLS) [188], with the main disadvantage that the results may be biased towards the larger colloids present in the system that scatter more light. Another alternative is the use of microfluidic diffusional sizing [10, 211].

Separation techniques can qualitatively compare species. Capillary electrophoresis, for example, separates colloids in terms of their mobility. Different chromatographic techniques enable separation according to several parameters including charge (ion-exchange chromatography), size (with gel filtration or size exclusion chromatography), hydrophobicity (hydrophobic-interaction or reverse phase chromatography) or any affinity with a specific tag (tag affinity chromatography) [116, 129].

All these techniques are not necessarily independent and can be combined in order to create a multidimensional separation/characterisation of protein samples. One of the purposes of my work is to combine the separation of capillary electrophoresis and size analysis of microfluidic diffusional sizing, explained in detail below. Another purpose is the development of new microfluidic techniques in order to expand the limits of existing tools.

1.2 Amyloids

The polymerisation process of proteins was first described by Oosawa [148] and can be applied to functional protein aggregates. It is described as a two-step model that includes the formation of nuclei above a critical concentration of protein, followed by growth of the nuclei by addition of monomers. The main difference between functional and aberrant aggregation of proteins is perhaps the presence of secondary processes in the aberrant amyloid formation [6] which lead to the uncontrolled formation of aggregates. The kinetics of the aberrant formation of proteins involves both primary (nucleation and elongation) as well as secondary processes (secondary nucleation and fragmentation) [37, 103] as depicted in Figure 1.2. The secondary processes accelerate the formation of fibrils by the formation of nuclei via secondary nucleation in the surface of existing fibrils, or via fragmentation of the fibrils which increases the number of fibril ends. One of the most widely used techniques for studying amyloid formation is the use of the dye Thioflavin-T (ThT) which is only fluorescent when bound to the amyloid beta-sheet structures of the fibrils, and thus allows tracking of the fibril mass over time [21]. However, the formation of primary and secondary nuclei or small oligomers remains challenging to pinpoint with the use of ThT and has been proven with the use of radiolabelling techniques [122]. Another challenge in the study of oligomer formation during the aggregation process is the heterogeneity as well as the low concentration and transient nature of such species, which are difficult to track with the sensitivity and time-frames of traditional techniques.

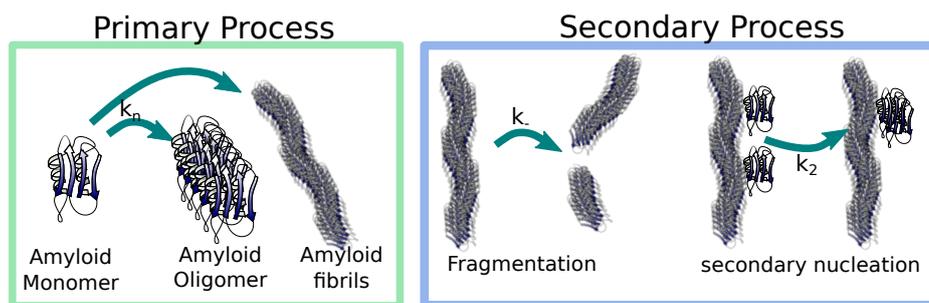


Fig. 1.2 Mechanism of amyloid formation from the primary and secondary pathways.

1.3 Microfluidic advantage and characterisation

Microfluidics is a set of techniques for manipulating liquid at submillimetre scale. At this scale, viscous forces dominate the inertial ones, which acts to reduce the turbulent

flows. With laminar flow, microfluidics enables better transport control and prediction than in the bulk situation [23]. These techniques work at a size closer to living cells, and at reduced volume sample per experiment. They reduce the cost of prototypes and experiments, as well as allowing the possibility of high throughput design. Multiple and varied experiments in the field of biotechnology use microfluidics, like DNA sequencing, reactor, cells culture and separation, immunoassay, and single molecule detection at low concentration [75, 111, 133, 150].

1.3.1 Physics in microfluidics

The physics of incompressible fluids can be described by the reduced Navier-Stokes equation:

$$\rho[\partial_t \vec{v} + (\vec{v} \cdot \vec{\nabla}) \vec{v}] = -\vec{\nabla} p + \eta \nabla^2 \vec{v} + \rho \vec{g} \quad (1.1)$$

where ρ , η represent respectively the density and the viscosity of the fluid, and $\vec{\nabla} p$ refers to the change of pressure within the fluid. Note that for a conservative external force, the force term can be absorbed into the pressure term: $-\vec{\nabla} p + \rho \vec{g} = -\vec{\nabla}(p + V) = -\vec{\nabla} \bar{p}$. A unitless parameter called the Reynolds number can be defined to highlight when part of this equation can be neglected.

$$Re \equiv \frac{\rho v_0 L_0}{\eta} \quad (1.2)$$

where v_0 and L_0 are characteristic scales of the system. A unitless version of equation (1.1) can be written as:

$$Re[\tilde{\partial}_t \tilde{\vec{v}} + (\tilde{\vec{v}} \cdot \tilde{\nabla}) \tilde{\vec{v}}] = -\tilde{\nabla} \tilde{p} + \tilde{\nabla}^2 \tilde{\vec{v}} \quad (1.3)$$

where tildes represent dimensionless quantities, for example $\vec{v} = v_0 \tilde{\vec{v}}$. For low Reynolds numbers, the inertial term as well as the time dependence can be neglected.

In a microfluidic system, small L_0 pushes the Reynolds number to low values and therefore reduces turbulence to something negligible; the modelling is more predictable. In addition to the property of laminar flow, micro/nano-fluidics has a high surface volume ratio compared to larger scale fluidics. Wall effects are non-negligible, such as wall charges. These charges tend to be screened by counter-ions once in solution. This phenomena affects not only walls but also particles present in the liquid, with particles being any molecules, proteins, or colloids in this example. Two different layers can be distinguished in the screening area, called a double layer. The first, called the Stern layer, is composed of solution ions fixed to the surface. The second, called the diffuse layer, screens the potential smoothly. The concentration of counter-ion/co-ion increases/decreases from the solution's concentration

toward the wall or particle. The expression for the ion concentration can be approached by the Poisson-Boltzmann mean-field approximation [208]:

$$c_{\pm} = c_0 \left(\exp \left(\frac{\mp e\phi}{k_B T} \right) - 1 \right) \quad (1.4)$$

where c_{\pm} are the concentrations of the ion/co-ion, c_0 the concentration of the ion in solution, k_B the Boltzmann constant, T the temperature, e the electrical charge and ϕ the electrostatic potential. The range where the double layer exerts an effect on the solution is called the Debye length, written λ_D or κ^{-1} . The strength of electrical interaction forces or phenomena such as diffusiophoresis or electrophoresis are often expressed in terms of the unitless parameter κa , where a is the size of the particles. The double layer is dependent on the pH value or the solution's salt concentration, for example the size decreases for higher salt concentrations. The solution can be tuned in order to increase or decrease the effect.

1.3.2 Fabrication

Microfluidics chips are mostly fabricated using soft lithography techniques [50]. Fabrication in our lab consists of two main steps, the fabrication of the master, and the replication of the device from the master. The former is made of a thin uniform layer of photoresist material SU-8 on a silicon wafer. The uniformity of the layer is ensured by spinning SU-8. Different rotational speeds, times, and viscosities of SU-8 ensure different channel heights. After heating a first time (soft-bake) on a hotplate to fix the photoresist, the SU-8 is exposed to UV-light with a mask lying on top of it. A longer heating time (hard-bake) on a hotplate is performed to ensure fixation of the illuminated pattern. Non-exposed SU-8 is removed during the development time by washing it with PGMEA. The second step, which is the device fabrication, can be repeated on the master. It consists of pouring a mixture of PDMS and the curing agent on the master of the desired thickness, and after accelerating the curing process in an oven (at 60 °C), the PDMS is peeled off. Microfluidic channels are now trenches, in which holes are punched at the inlets and outlets before it is covered by bounding a glass/PDMS slide using a plasma oven to activate the surface.

1.4 Development of a separation capillary electrophoresis device combined with H-filter detection

The development of a separation and characterisation technique for protein samples usually comes from a set of sub-techniques. It is called a multi-dimensional separation system, as each technique isolates a certain physical property (dimension) of the protein [95, 210]. It is used mainly in proteomics [36], in order to determine the protein composition of a biological sample. The 2D-microchip separation technique, used in one of my PhD projects, combines the capillary zone electrophoresis technique for separating the protein sample in different species according to their mobility with the H-filter, for performing diffusional sizing. The capillary zone electrophoresis applied in this microchip has been widely used in microchips since the year 2000 [48]. The separation of proteins by electrophoresis was first described in 1930 [26, 192]. In 1967, Hjerten used capillaries and reduced the thermal broadening [84], with the explanation described later [200]. The emergence of capillary electrophoresis (CE) arrived after the description by Jorgenson and Lukacs in the 1980s with a smaller diameter capillary [96, 98]. The technique consists of separating different proteins when exposed to an electric field in terms of a parameter called mobility. The mobility is a result of the protein charge, geometry, mass and size. It is dependent on the solution in which the protein is immersed. As a rule of thumb, proteins of similar charge will move faster if they are smaller. Proteins of similar size/mass will move faster if they carry higher charge. The H-filter consists of two streams flowing in parallel in the same channel, one with the desired sample to analyse and the other with a similar buffer. As larger species diffuse slower than smaller ones, the diffusion profile at the end of the H-filter depends on the particle size. The configuration originally was used to separate smaller species from the sample, and the microfluidic design kept the name [15, 79, 204]. The technique in microfluidics consisting of recovering the diffusion coefficient of species looking at the evolution of the profile in time and space is called diffusional sizing. Different geometries exist with different names, from T-junction to H-filter, as well as a centred flow surrounded by the auxiliary buffer streams [10, 30, 211]. Interestingly, a second configuration consisting of separating the two fractions of diffused/undiffused parts within an H-filter before using the CE gives some advantages discussed in the chapter 2.

1.5 Microscopy

With the aim of characterising proteins in solution, the detection of the protein plays an important role in the system, and therefore a microscopy technique specifically for the microfluidic chips has to be developed in parallel. Two main ways of detection have been investigated and can be combined in future work. The first is label free detection which allows the study of proteins without the addition of a dye which will influence even minimally the system. The second is the detection of single molecule events, which allows a reduction in the concentration needed and allows the study of sample interactions at close to physiological conditions [117]. If this microscopy technique is used with the combination of microfluidic, not only the sample volume to analyse is reduced but the sample can be at low concentration, and thus without techniques of pre-concentration. It therefore opens up the door for new diagnostic techniques.

1.5.1 Label free detection

Label free detection is often referred to as wide-field microscopy, where the scattering signal of the light is used to detect the particles. This technique with no addition method is limited to a spatial precision of the order of the wavelength we are looking at [198]. And if we want to detect a single molecule event regardless of the spatial shape of the particle, it is limited by the detection "power" of our system. Indeed the Rayleigh scattering intensity decreases as a power of 6 of the size of the object [17]. An interesting way to reduce the Rayleigh limit is to combine this scattering with a carry wave. This combination decreases as a power of 4 instead of 6. This type of microscope based on the interference between carry wave and scattering has been developed by Philipp Kukura and Vahid Sandoghdar, and is now known as iSCAT [108]. A second label free detection method comes from the intrinsic fluorescence property of proteins containing aromatic amino acids. These specific proteins have an autofluorescence response from an excitation in the UV. The limit is then no longer based on the scattering power of the particle, and proteins can be detected regardless of their size if they contain enough aromatic ring. The use of a 280 nm wavelength laser allows maximising the response. The technique can be combined with single molecule technique like a confocal system [29, 41, 144].

1.5.2 Single molecule detection

Single molecule detection has attracted increased interest during the last 20 years [45, 169, 190, 198]. It is physically constrained by the so-called diffraction limit, which is the impossibility to detect two single emitters separated by a distance d smaller than the Abbe limit:

$$d_{Abbelimit} = \frac{\lambda}{2n \sin(\theta)}$$

where λ is the wavelength of the emitter, n is the medium's refractive index and θ determines the angle of aperture. A large number of publications try to beat this physical constraint. With the exception of structure illumination microscopy [163], the main approaches consist in forcing two close emitters to not emit at the same time. Therefore the individual position of the emitter can be tracked back, and is often referred to as super-resolution microscopy techniques. It is the case for stimulated emission depletion microscopy (STED), which creates a depletion emission around the excitation light [199], or the PALM/STORM which relies on a blinking emitter and a large collection of images in order to have in mean only one emitter excited in a circle of a radius d_{Abbe} [157, 179]. Despite an increased interest in techniques for imaging below the diffraction limit. Single molecule detection does not require imaging, and can be achieved without these techniques. In this case the density or concentration of the particle of interest used should be such that the probability to encounter two or more particles closer than the diffraction limit is quasi null. In the case of the confocal technique, the excitation light is reduced to the Abbe limit size using a pinhole [206]. The excitation volume is approximatively of the order of a femtolitre. When the particle of interest crosses the spot, it either scatters or emits some photons. The microscope can be enhanced by using a single photon detector, like photomultiplier tubes (PMTs) and an avalanche photo-diode detector (APD).

1.6 Structure of this thesis

The main challenge is measuring heterogeneous solutions of biological samples with similar characteristics, and developing a model given the minimal size detection that can be obtained through these techniques. This has been developed through four different projects. The first aimed to measure the mobility and size in a heterogeneous solution by combining a capillary electrophoresis separation with a diffusion device. The second project looked at combining existing microfluidic designs using the potential of a single molecule detection device, in order to increase the sensitivity in measuring mobility or radius of proteins. The third project

used nanofluidics in order to measure the size of proteins. It uses a confocal spot which illuminates a nanocavity. Using a population of the residence time of the particle within the cavity, it is possible to size the molecules. The last project used a novel approach to size label free particles in an interferometric scattering microscope (iSCAT), which is similar to the way a DLS measures particles in a wide-field microscope. Finally, due to a lab closure in 2020, I have been able to use my time in offering new analytic approaches of existing ELISA assay in order to increase the sensitivity of the assay and record the seroprevalence of a sub-population in the great area of Zürich.

Chapter 2

Coupled Capillary Electrophoresis and Diffusional Sizing

This chapter is based on a patent, and a manuscript in preparation:

Raphael P.B. Jacquat, Quentin A. E. Peter*, Magdalena Czekalska, Greta Musteikyte, Tuomas Müller, & Tuomas P. J. Knowles. "Couple Capillary electrophoresis signal to diffusional sizing for heterogeneous sample".*

*(*contributed equally).*

My contributions to this paper are the development of the instruments, the data acquisition, data analysis, and simulations. Therapeutic drug monitoring (TDM), which consists of monitoring the concentrations of drugs in patients' biofluids in order to adapt their dosage, has become increasingly important in recent years [3, 68, 132, 134]. It requires fast techniques allowing detection of specific biomarkers, drugs, or proteins at low concentration in a heterogeneous solution like blood or serum. In addition to TDM, such detection can be useful in medical diagnostics. The diffusional sizing technique can size particles in solution, but the technique only provides an average size of the mixture. A separation technique is required in order to separate and size individual components. Such separation can be achieved using capillary electrophoresis on microfluidic chips. In this chapter, a proof-of-concept experiment is presented using a two-dimensional microfluidic analysis. Two systems have been tested: one requires an extrinsic fluorescent label for the sizing, while in the second the label is added after the sizing step. Among various biomedical applications, this technique is especially useful for detection and characterisation of early oligomeric species formed during the aggregation of proteins involved in the occurrence of neurodegenerative diseases. The α Syn protein is therefore used as proof-of-concept of this tool.

2.1 Introduction

Although sizing homogeneous solutions is useful to characterise a sample or a chemical reaction [168, 172, 173], sizing heterogeneous solutions can open many doors such as in medical diagnostics, where the challenge is to separate the target protein from a complex mixture. Purification can be tedious and time consuming, risking the loss of sample and effectively inaccurate measurement. Gel electrophoresis is commonly used in chemistry labs but suffers from a lack of precision. Capillary electrophoresis (CE) with a strong voltage seems to be a good way to separate species within solution based on their electrophoretic mobility. The challenge posed by this method is often the interpretation of each peak. The peaks can have a position which varies from run to run and is influenced by any tag attached to the target species. While adding a tag can be an advantage in selecting a given species, it also adds new peaks within the detection, leading to a more complex interpretation as to which molecule is highlighted. A way to reduce the difficulty of interpreting a CE curve is to combine the technique with another that allows characterisation of each peak individually. Diffusional sizing seems to be a good candidate. If each species is separated, each peak can have an associated size, enabling a direct interpretation of the peaks.

In this chapter, two designs are developed and presented.

The first design consists of a separation step, followed by a sizing step of the separated species. The separation is performed using CE, and the sizing uses a diffusional sizing H-filter design. The second design has a less obvious workflow. First, the separation into two fractions of the mixture within an H-filter is performed, creating a diffused fraction and non-diffused fraction. From this separation, it is only possible to probe an average size of the mixture. Each of the two fractions will then pass through a CE device, and give electropherogram peaks. The two electropherograms (diffused and non-diffused) can be overlapped in time and their ratio extracted. For each peak an average ratio is extracted and can be directly related to the hydrodynamic radius of the separated components.

While the first design is no longer developed and used in our lab, it is discussed here in the hope of being helpful in some future project. The second design is more promising and is still in use. I hope it will help future research, especially in protein self-association studies.

2.2 Theory

2.2.1 Capillary Electrophoresis

Electro-osmotic flow and electrophoretic mobility μ

If a potential is applied in a conductive solution confined between charged walls, the liquid will move due to the unbalanced screening charge inside the wall double layer. These ions are attracted by the opposite charge potential and will induce an electro-osmotic flow, abbreviated EOF. An illustration of the phenomena is shown in Figure 2.1. The flow has the advantage of having a flat distribution compared to the pressure driven flow, which has a parabolic distribution. It results in a much easier analysis of the data. One disadvantage is that the flow depends on the interaction between the solution and the surface charge on the channel walls, and therefore depends on the material and the surfactant used. It also introduces the constraint to use similar materials for the whole channel. Microfluidic chips commonly used in research labs are made of three walls of PDMS and a wall of glass. However, a thin layer of PDMS can be added on the glass side. One exception where a thin layer of PDMS cannot be used is for UV-spectroscopy, where the bottom quartz glass cannot be coated. This will increase the difficulty of post-processing the data due to flow shear. Applying a potential within a microfluidic device will, in addition to the EOF, move the charged proteins as a function of their mobility, μ , defined as the ratio of the particle velocity under an applied electric field to the magnitude of the field:

$$\mu = \frac{v}{E} \quad (2.1)$$

This quantity depends on the protein's intrinsic parameters such as shape, charge distribution, and mass. A simplified model commonly used is the one of Henry [39, 82], which assumes spherical particles with uniform low charge distribution:

$$\mu = \frac{Q}{4\pi\eta R_H(1 + \kappa R_H)} f(\kappa R_H) \quad (2.2)$$

where Q is the net charge, R_H is the hydrodynamic radius, κ the inverse Debye layer, η the solution viscosity, and $f()$ is Henry's function which is bounded and monotonic. Further developments exist taking the shape into consideration as well as quadrupole effects and other parameters tweaking the mobility model prediction [101, 147]. This will not be discussed further here, as the main interest is the dependence of the mobility on the double layer. As described in the introduction chapter, the double layer depends on salt concentration and

pH. We can define the apparent mobility $\mu^* = \vec{v}^*/\vec{E}$, with \vec{v}^* the particle velocity measured inside the microfluidic chip, as the velocity of the particle measured minus the velocity of the flow, combining EOF \vec{v}_{EOF} , and pressure driven flow \vec{v}_p .

$$\mu_{\text{prot}} = \frac{\vec{v}_{\text{prot}}^* - \vec{v}_{\text{EOF}} - \vec{v}_p}{\vec{E}} \quad (2.3)$$

A simplified approach to calculate the velocity of EOF is described by the Helmholtz-

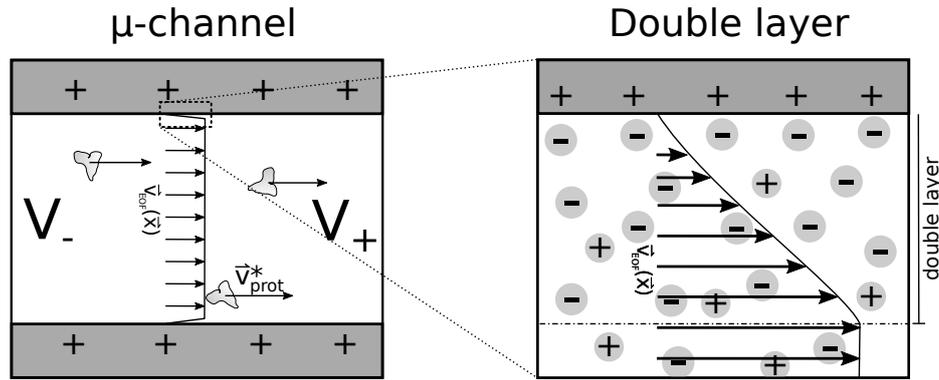


Fig. 2.1 Illustration of electro-osmosis flow in a capillary without pressure driven flow. On the left-hand side, flow profile $\vec{v}_{\text{EOF}}(\vec{x})$ in a section of a microfluidic channel exposed to an electrical potential between V_- and V_+ , with walls positively charged. Protein velocity $\vec{v}_{\text{prot}}^* = \vec{v}_{\text{EOF}}(\vec{x}) + \mu_{\text{prot}}\vec{E}$ is illustrated. On the right-hand side, a zoom on the electric double layer phenomena is shown. The negative ions are more abundant in the screening area (see equation 1.4) of the positively charged wall inducing flow velocity.

Smoluchowski expression [118, 184, 201]:

$$\frac{v_{\text{EOF}}}{E} = \frac{\varepsilon\zeta}{4\pi\eta} \quad (2.4)$$

where ε is the buffer solution dielectric constant, ζ the zeta-potential across the diffuse layer, η the viscosity of the solution, and E the electric field.

Separation

As μ_{prot} is an intrinsic parameter of the protein, capillary electrophoresis can separate heterogeneous mixtures of proteins. For a given microfluidic device, the separation channel length and cross section, solution/buffer used, voltage applied and wall materials are known. Therefore, the electro-osmotic flow should be well characterised and reproducible. The separation power of a device is measured or calculated via the following quantities: peak capacity, plate number, or resolution [176]. The definition of the number of plates is similar

Molecule	D [m ² s ⁻¹]	R_H [nm]	μ_{part} [m ² V ⁻¹ s ⁻¹]	q [e]	M_W [kDa]
Fluorescein	$4.27 \cdot 10^{-10}$	0.502	$-3.7 \cdot 10^{-8}$	-1.02	0.33
ERK1	$7.47 \cdot 10^{-11}$	2.87	$-1.5 \cdot 10^{-8}$	-5	43
Rabbit anti MAP Kinase (ERK1/2)	$4.97 \cdot 10^{-11}$	4.33	$9.8 \cdot 10^{-9}$	5	151.5

Table 2.1 Values of different molecules used for comparison and simulation [212].

for all standard chromatography techniques [71] and is defined as $N = L^2/\sigma_x^2$, where L is the separation channel length and σ_x the standard deviation of the peak broadening. Compared to pressure driven flow, the plug-shape flow from CE does not create dispersion peak broadening from the parabolic-shape and the peak broadening is limited by species diffusion and other effects such as thermal peak capacity [72]. Assuming no pressure driven flow as described in [97], the number of plates no longer depends on the capillary length L but only on the applied potential V :

$$N = \frac{\mu V}{2D} \quad (2.5)$$

The peak capacity is given by the ratio between the separation distance L and 2 times the peak broadening: $4 \cdot \sigma_x$. It gives approximately the number of peaks that are able to be stacked and distinguished in the capillary for a given mobility μ :

$$\text{Peak capacity} = \frac{1}{4} \frac{L}{\sigma_x} = \frac{N^{1/2}}{4} \quad (2.6)$$

The peak capacity is not representative of the maximum resolution achievable, as it depends only on one mobility coefficient of the solution, but this unit can be used to compare the performance of different systems, buffers or coatings [53]. As for the number of plates, the length of the channel and the electric field are not the main contribution, only the voltage is relevant. The stronger the voltage, the better the separation. But the time to arrive at the detection point can vary. For 20 kV in a 20 cm long channel, it is possible to tune the detection point in order to find a good ratio between time spent and resolution of the peak capacity needed (see Figure 2.2, which uses Table 2.1 as reference). The final quantity used to characterise the separation power of a system is the resolution. It is a quantity specific to two different analytes in a system with given mobilities μ_1 and μ_2 . It shows the ability to distinguish the peak, $R = \frac{\Delta x}{4\bar{\sigma}_x}$, where Δx is the distance between peak 1 and 2, and $\bar{\sigma}_x$ is the

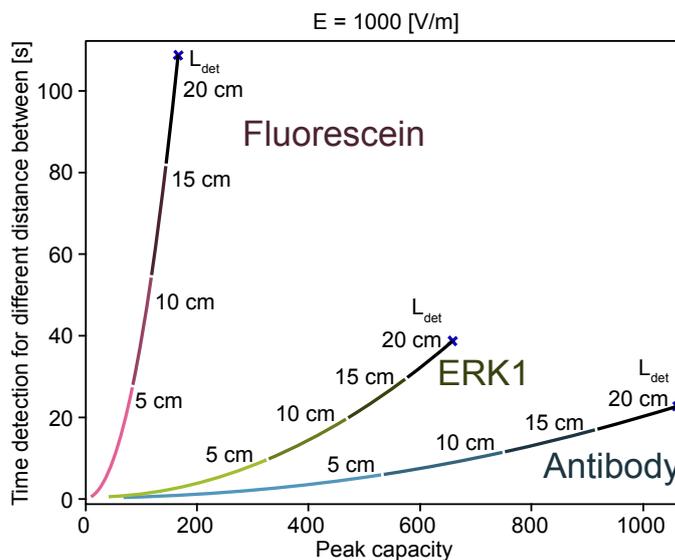


Fig. 2.2 Peak capacity and time before detection, for three different biomolecules, at different detection points (L_{det}). The electric field is fixed at 20 kV applied over a channel of total length (L) of 20 cm.

mean of both standard deviations. It is described in [71, 97] and derived as:

$$R = \frac{N^{1/2} \Delta v}{4 \bar{v}} = \frac{1}{4} \left(\frac{\bar{\mu} + \mu_{EOF}}{2D} V \right)^{1/4} \left(\frac{\mu_1 - \mu_2}{\bar{\mu} + \mu_{EOF}} \right) \quad (2.7)$$

where \bar{v} is the mean velocity, Δv is the difference velocity, and $\bar{\mu}$ is the mean mobility of the two species. Tuning μ_{EOF} , with different coatings in order to have $\bar{\mu} + \mu_{EOF} \rightarrow 0$ for two given specific proteins, can enhance the resolution. However, this condition implies a very low protein speed. The separation resolution increases with the time passed into the separation channel, within the limit of elution time in the whole capillary given by $t_{max} = \frac{L^2}{2D}$, where L is the capillary length and D the particles' diffusion coefficient. In addition to the diffusion, adsorption on the wall can lead to peak broadening and reduce the resolution [175]. The broadening of a sharp peak due to a heat gradient introduced into the system should be taken into account. The heat is dependent on the current and voltage applied. Heat not only increases the diffusion coefficient; it also enhances the mobility by 2% °C⁻¹[97]. Heat also denatures proteins. A smaller channel cross section area reduces the temperature difference between the centre and the wall by the square of the radius (in case of a spherical channel), and therefore reduces broadening due to speed difference. Heating depends on the current, therefore a low conductance buffer is used in order to reduce the heating power [14]. Finally,

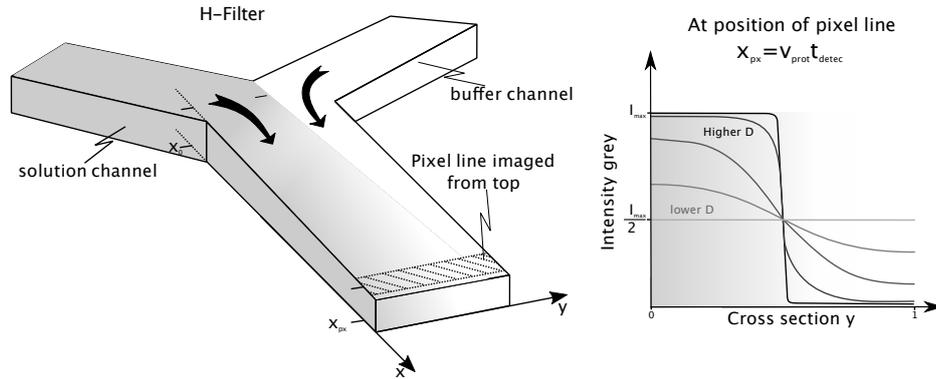


Fig. 2.3 Schematic of the H-filter On the left: Schematic of the set-up simulated with the convention parameter used in this report. The thick arrows indicate the flow direction. On the right: Different diffusion profiles expected depending of the diffusion coefficient and thus the hydrodynamic radius.

as this broadening is theoretically from an infinitesimally small plug of protein, an additional broadening due to the physical size of the injection should be added.

2.2.2 Diffusional sizing with a H-filter

In microfluidics, under low Reynolds number conditions, i.e. where inertial force are negligible, the dominant mechanisms are advection and diffusion [143]. At the junction of two streams flowing in parallel, the only relevant mixing is caused by diffusion. This is the configuration of the main channel in an H-filter, or after a T-junction. Looking at the density profile across the two streams at different positions along the channel and knowing the flow velocity inside the channel is sufficient to recover the diffusion coefficient of a homogeneous solution. Figure 2.3 shows the principle of the microfluidic device, as well as different spatial distributions expected as a function of the diffusion coefficient, D , of the species. Each position x along the channel represents a timepoint t , which is the time spent to diffuse. The link between the diffusion coefficient and the hydrodynamic radius is given by:

$$D = \frac{k_B T}{6\pi R_H \eta} \quad (2.8)$$

where k_B is the Boltzmann constant, T is the temperature, R_H the hydrodynamic radius, and η the viscosity of the liquid. The hydrodynamic radius is defined as the radius of a sphere which diffuses at the same rate as the given object. This diffusional sizing method has been developed in different configurations and is described in several articles, for example in [10, 30, 211]. The main limitation comes from the inability to fit polydisperse data, such as heterogeneous mixtures with similar hydrodynamic radii.

2.3 First generation device

Capillary electrophoresis (CE) depends mainly on the electrical potential applied. Common CE techniques use a range of ~ 10 kV. A 20 kV power supply is used to create the separation technique, but the electrical workshop custom-made device is certified for 10 kV. A second source of 500 V is used to have a voltage driven injection into the separation channel. Both power supplies apply electrical potential at 4 points in the microchip, but never simultaneously; a system switch ensures there are no links between the two sources of potential. A simplified model of the electrical circuit as well as an image of the custom-made device are shown in Figure 2.4. The red and green lines represent the two independent electrical circuits, the high voltage and 500V circuit, respectively. The red resistances enforce a certain voltage. The blue area represents the microfluidic chips. The equivalent of the last circuit design can be seen in Figure 2.5. The capillary electrophoresis part and the H-filter part represent the resistance R_4 . R_1 , R_2 and R_3 are the different channels from buffer inlet, sample inlet, and sample waste, respectively, until they join each other. The H-filter buffer inlet (HBi) is at the same potential as the buffer inlet (Bi), and represents the extremity of the blue resistance R_1 . The H-filter buffer waste (HBw) and H-filter sample waste (HSw) are also connected electrically and represent the extremity of the resistance R_4 in Figure 2.4. In practice it was more convenient to experiment with a simpler design and many capillary electrophoresis experiments have been performed with no inlet/outlet at HBi, HBw, HSw, and only at the end of the CE part. Further discussions on the CE and H-filter part are given in different sections below.

2.3.1 Injection and separation process

A limited quantity of sample needs to be injected into the capillary part before starting the separation process. The injection is driven by EOF. A voltage difference between the sample inlet (Si) and the sample waste outlet (Sw) is generated. With no coating on the PDMS channel, the EOF is directed through the negative electrode. Ideally, the voltage at the buffer inlet (Bi) and at the outlet of the H-filter (HBw, HSw) should be tuned in order to avoid flow direction in the separation channel, meaning having the same value as the cross intersection voltage V_x . Once the sample is loaded into the separation channel, the injection driven voltage is turned off, and the switch changes the circuit from the 500 V power supply to the high voltage source. High tension is applied between buffer inlet and buffer waste. The voltage at the sample inlet and sample waste are chosen to move the sample away from the separation channel, as shown in Figure 2.6.

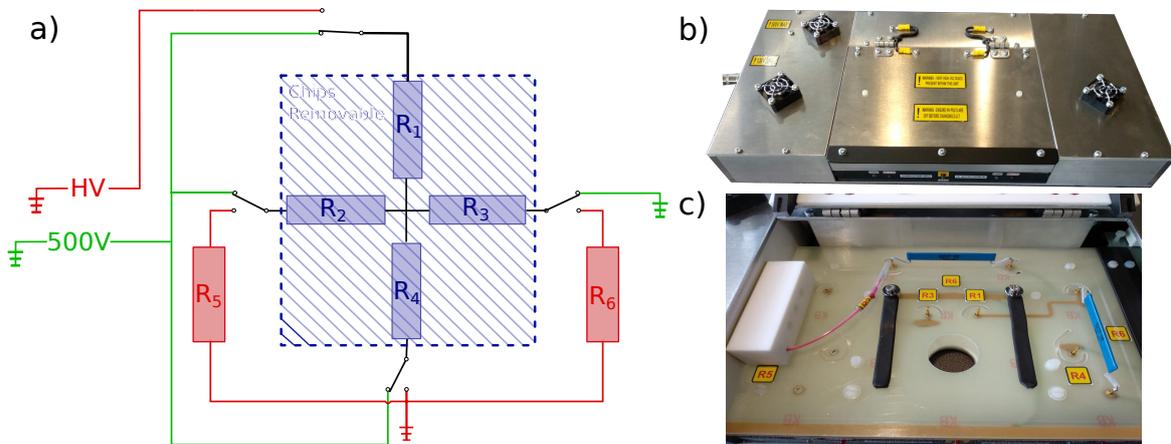


Fig. 2.4 Custom-made device used for CE. **a.** Schematic of the electrical connection: two circuits exist in parallel and can be switched from one to another. In green the injection 500 V circuit, in red the separation driven by high voltage (HV). The red resistances are used to reduce the voltage difference between the centre of the chips and the extremity of R_2 and R_3 , and they can be exchanged and modified easily. The blue area represents the consumable PDMS chips in which the channels are represented by resistances, depending on the design and buffer used the value may differ. In general R_1 , R_2 and R_3 have a similar value while R_4 is the highest. **b.** Picture of the over-secured encapsulated box. **c.** Photo of the inside view: the hole allows the microscope objective to image the PDMS slide which will stand on top of it. Platinum wires will be connected at the positions labelled R_3 , R_1 and R_4 and on the red wire, and they will connect with a small well into the PDMS chip.

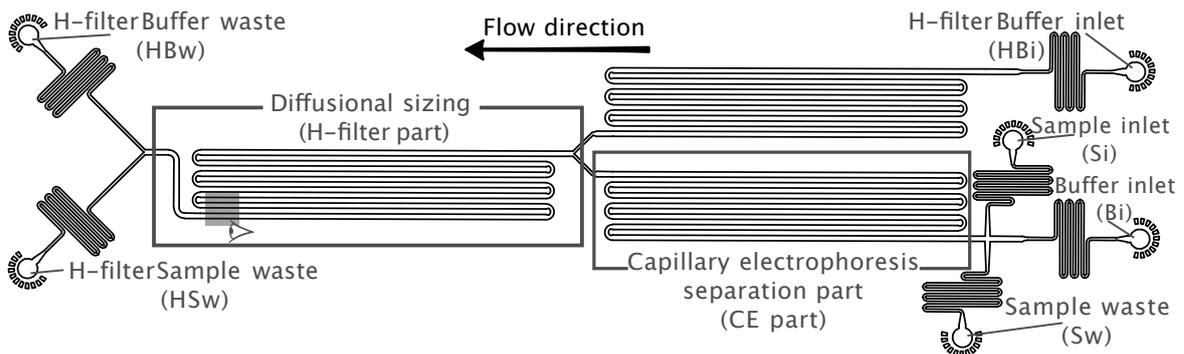


Fig. 2.5 Microfluidics chip combining CE separation and diffusional sizing. The proteins are injected through sample inlet (Si) to the sample waste (Sw) during the loading time. Once the concentration is reasonably high at the cross intersection, injection mode/time starts. An electric field is applied between HBi, Bi and HSw, HBw and it injects loaded proteins inside the CE separation part. The separation depends on the mobility of the molecules. At the end of the CE part, the H-filter is used to diffuse proteins inside the channel. The diffusion profile is detected near the end of the H-filter part. Each diffusion profile depends on the protein size.

2.3.2 Multiple injection

The advantage of performing on-chip capillary electrophoresis is the ability to repeat the experiments with high frequency compared to traditional techniques. Indeed, there is no need to wait for the solution plug to reach the end of the separation channel before introducing a second. Automation of the potential setting for injections and separation phases make programming possible. Multiple injections have been recorded successfully and are displayed in Figure 2.6. The intensity and flow rate are quite repeatable. The difference in intensities in the red area is mainly due to the difference in frequencies between the frame rate of the camera and the rate of the injection. Several measurements at the end of the channel reduce the error in the size and mobility. On the other hand, the time elapsed during the injection phase increases the diffusion of the species along the channel which will weaken the separation power, like the resolution or peak capacity.

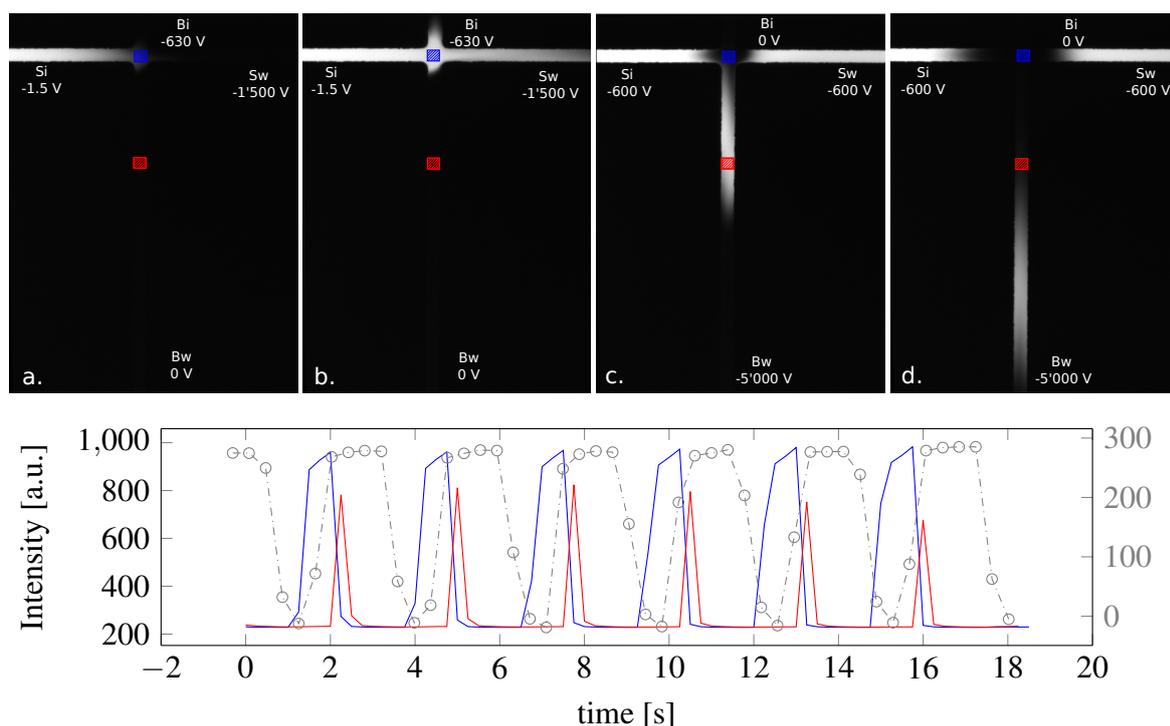


Fig. 2.6 Images of the injection process. a. & b. Images taken during the sample loading process. **c. & d.** Images taken during the sample separation time. The graph represents the mean intensity in the picture corresponding to the coloured area as a function of time. The grey dashed dotted line represents the measurement of the flow sensor.

2.3.3 Analysis and Results

Image analysis

The volume injected at the beginning of the channel spreads out over several frames by the time it reaches the end of the channel. Representative frames are displayed in Figure 2.7. Looking at the mean intensity of an area in the channel as a function of time can provide information on the separation of different samples. For example, the bottom left graph in Figure 2.7 displays how two different substances separate. More information can be extracted by superposing several images. The algorithm used to analyse the data consists of three steps. The first reduces artefacts due to non-uniform illumination; this spatial deformation of the image is flattened by subtracting a polynomial fitting which does not take the channel into account; for having uniformity in time images, each image is divided by a reconstructed image where each pixel is the median value over time. The second step consists of finding the speed of each peak. Each image is reduced to a pixel array representing the mean over the cross section of the channel, and stacked in time (see Figure 2.7). At each position, the peak is reconstructed and the position of the maximum recorded. Extracting the slope of the position of the maximum in time then gives the speed of each peak. The final step consists of finding the size of the peak. It is recovered by stacking the different profiles placed according to the speed velocity. Such a stack of profiles with Gaussian fitting is shown at the bottom right of Figure 2.7, with each colour representing one image frame.

2.3.4 The diffusional sizing method

2.3.5 Separation

A solution composed of fluorescein and FITC BSA separates weakly (Fig. 2.7). BSA has a lower peak capacity compared with fluorescein, which behaves contrary to expectations as BSA is larger than fluorescein and therefore should diffuse less. In reality, the protein/wall interaction cannot be neglected as in the theoretical calculation. The BSA peak widens due to absorption and desorption on the wall of the channel, and coats the channel slightly. The coating of BSA on the surface changes the EOF as well, and reduces the injection reproducibility during the experiment, especially when comparing the beginning and the end.

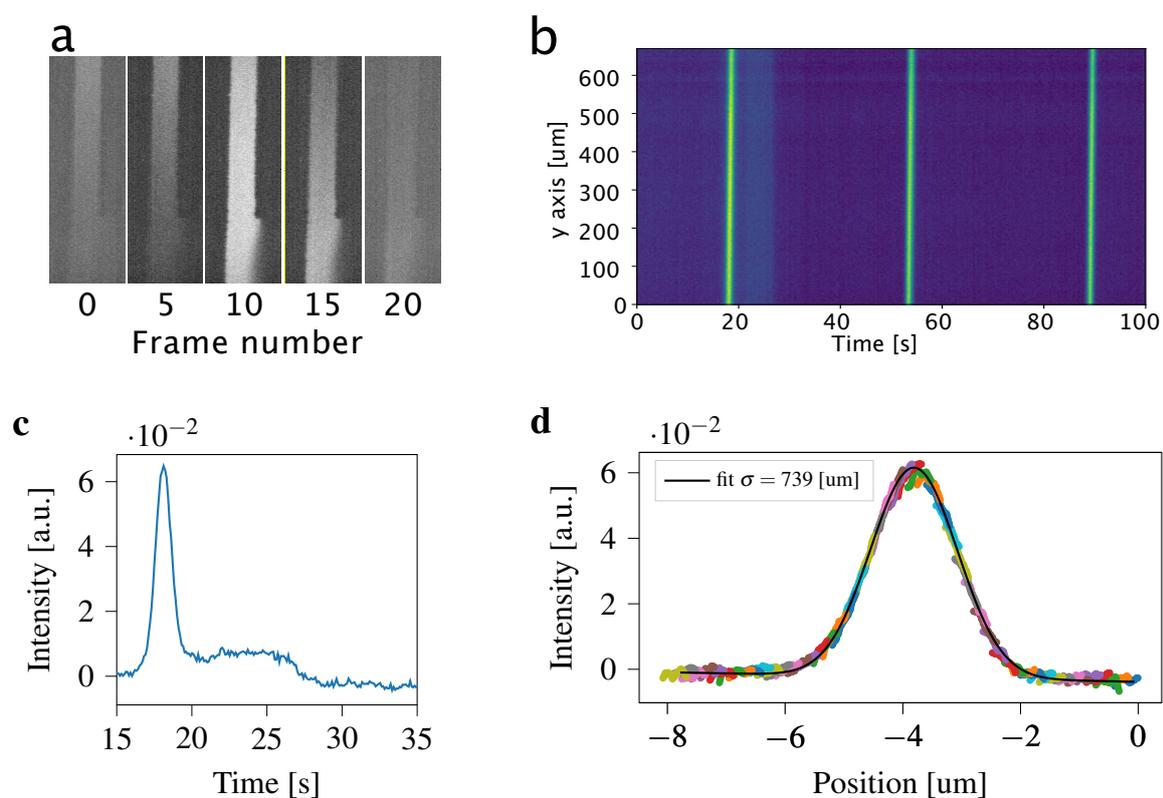


Fig. 2.7 Data analysis of the CE part results. **a.** Different frames of the channel, when the solution flows through. **b.** The projection of the inside of the channel in time. The line represents 3 injected solutions. The slope represents the velocity of the analyte. **c.** The temporal profile of a line across the channel, the second bump is BSA **d.** Profile intensity, with a Gaussian fit from the stack of images before the H-filter, the field of view of each frame represented in different colours.

2.3.6 H-filter, or size characterisation

A result of the H-filter sizing part is shown in Figure 2.8. It was taken from a device without the CE part. The sample flowed continuously, driven by EOF. It explains the low voltage used compared to the CE, as the length of the channel is much smaller. Fluorescein is sized with a 1D simulation of the profile. The known parameters are the initial profile, the flow rate, and the channel size. The flow rate and the dimensions of the channel give the time-points for different positions x for which the profile should be minimised. The orange crosses in the bottom left inset in Figure 2.8 illustrate the last pixel row, which represents different time spent in the channel by the fluorescein depending on the flow. The blue crosses represent the initial profile and the colormap profiles represent different time steps of the simulation. Higher voltage represents a higher velocity profile. The quality of the flow sensor used affects the result, which is reflected in the large error bar on the right side of Figure 2.8. Indeed, the flow sensor used for this experiment was at its detection limit. Also, the flow sensor was calibrated for water solution and could not measure the analyte velocity correctly. Other

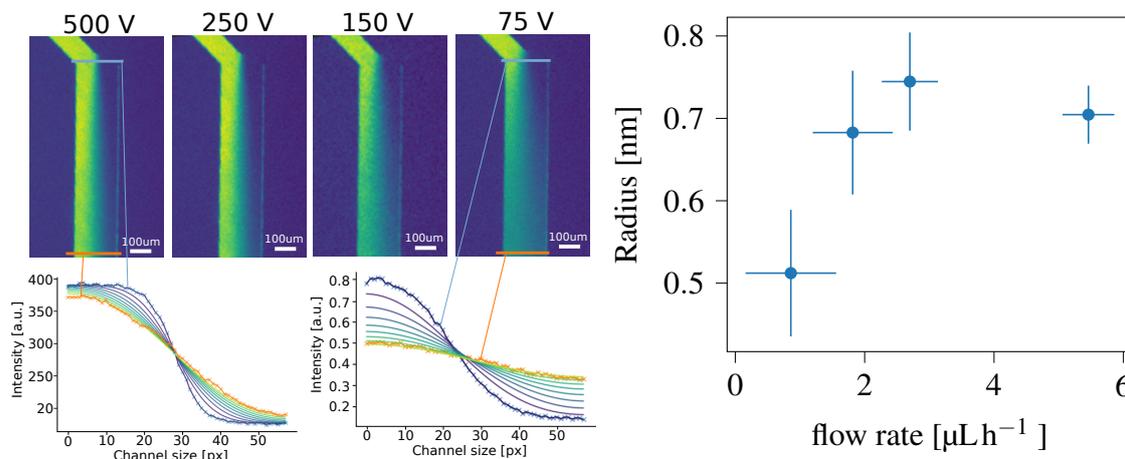


Fig. 2.8 Data analysis of the H-filter part results. Left: Images of the H-filter experiment at different potentials and thus different flow rates. Below are the results of the simulation from the grey blue line to the orange line. Right: Results of the calculated hydrodynamic radius. The value from the literature is 0.576 ± 0.001 nm [99]. The Error bars represent the standard deviation on 10 consecutive measurements.

results of experiments using an H-filter are presented in Figure 2.9, which shows the sizing calculated during a separation run at different time points. When no analyte is inside the H-filter, the fitted size is random, as it measures the convergence of the algorithm between 2 random noise profiles. Between frames 530 and 545, the solution is spread across the entire H-filter, the size thus should be approximatively the same for each frame, but this is not the case. The size seems to increase until it reaches a plateau at a correct value for

fluorescein. Sizing in this dynamic solution adds difficulty compared to the previous figures within a steady state experiment (the intensity is similar in time). The baseline is dependent on the intensity of the plug at a given position. A solution to overcome this problem is to run the simulation on each profile until it reaches a uniform distribution. Each baseline is then corrected in order to have the same profile at the end. The algorithm still cannot retrieve the correct value, as can be seen from the size measurement spread of 2 orders of magnitude between frames 530 and 545. This disadvantage of varying intensity is compensated by a more precise velocity for the protein, v_p^* , calculated without the flow sensor. Coupling H-filter and CE separation is then better than a single H-filter driven by EOF, despite the steady state images which can reduce the noise by increasing the exposure time or number of frames taken.

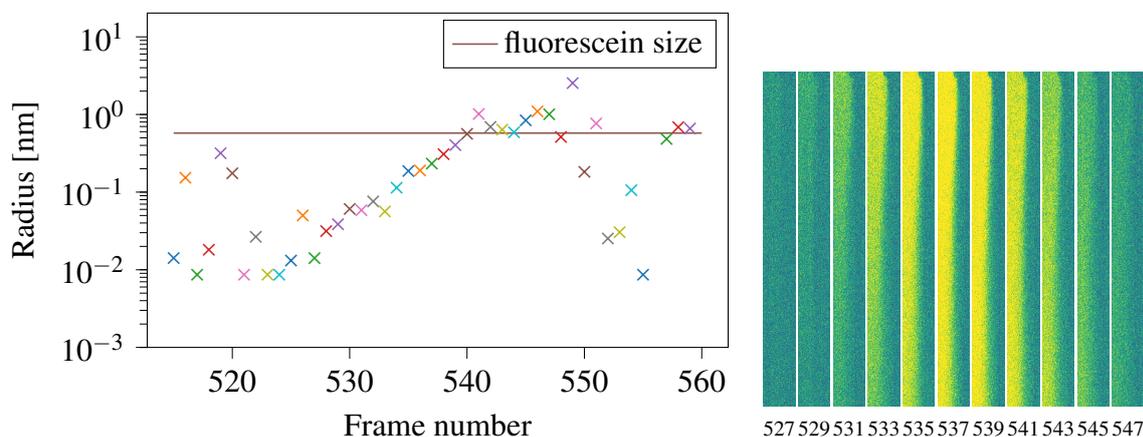


Fig. 2.9 Fluorescein sizing using the first design (CE combined with H-filter). Left: The different values of radii found for different frames; the solution arrived in the H-filter at frame 527 and left at frame 547. Outside the value of the radii which are noisy represent the convergence to the noise of the algorithm. The first frames and last frames with solution inside cannot be completely correct as the first and last profiles are not completely filled. Note the exponential line (linear in semi log y) aspect of the radius until it reaches the theoretical size of the fluorescein. Right: The corresponding image frames of the H-filter main channel for one run; it shows that in frame 533 the size should be reconstructable as well as in frame 543, which is not the case.

Flow control

One of the main limitations is the calculation of the flow velocity. The flow of the liquid caused by EOF and pressure driven flow can be measured by using a flow sensor at the end of the channel, or by measuring the velocity using the camera, but it will fail to detect a difference of flow between the two separation channels before the H-filter. A change in the flow rate in one of these channels, due to an impurity or an air bubble, can lead to a

miscalculation of the size. A possible solution is to introduce a standard protein in each of the mixtures. From the known size and the different profiles recorded, the flow rate of the two channels can be recovered. This solution nevertheless provides the flow rate at a time t in the experiment, and impurities can always appear and change the flow rate at a later stage.

Conclusion

The design described above has many parameters that are necessary to control, and each of them can potentially fail without an easy way to notice. The main problem is the impossibility to record the flow rate at the inlets of the H filter. The device nevertheless exhibits interesting features which are useful to keep in mind for anyone pursuing this project as a direct imaging of the different separated components, and the ability to size each of them. An improvement would be to use a more flexible high-voltage power supply with different channels, and better control on the pressure driven force by introducing a pressure driven pump. For me, the largest limitation is that the diffusion time cannot be tuned without affecting the separation as well.

2.4 Second generation CE/H-filter device

One way to reduce the complexity of the device described above is to separate the system into two microchips. It is however not straightforward to first separate and collect the different species from the capillary electrophoresis separation before running them in a diffusional sizing chip. By contrast, collecting fractions from an H-filter diffusional sizing experiment can be easily achieved without valves or time critical steps. This is due to the steady state nature of diffusional sizing.

2.4.1 Design process

A workflow of how to recover sizes from heterogeneous solutions within this configuration is described in Figure 2.10.a-d. The first step is the separation of the sample into two fractions, the diffused fraction and the non-diffused fraction. It is called the sizing step (**a**). After this step both fractions contain information on the size of the different species. This enables us to include a second step which is the labelling step (**b**). Indeed, now that the species are fractionated in terms of their size, it is possible to add a label in order to recover the native size without influencing the end result. If needed, the label can be added before the fraction. The measured size would then be the complex protein-label. With the two fractions labelled,

the last measurement step is to separate each fraction. The separation is made using on-chip capillary electrophoresis (c). The same chip is used twice, in order to have similar peak position for each fraction. The recorded intensity is proportional to the concentration. The concentration ratio of each separated peak gives information on the size of the corresponding species (d). The excess labels, when added after the sizing step (b-step), cannot be sized. The non-diffused fraction would contain less free label than the diffused fraction. Several designs of H-filter are used and a slight overflow of the buffer inlet is needed to achieve a better result. The optimisation of the H-filter was performed by Itzel Condado Morales with the help of Thomas Muller and can be found in her PhD thesis [40]. The simulated expected flow profile of a given design is displayed in Figure 2.10.e for different particle sizes. The ratio of the integrated non-diffused concentration (-50 nm to 0) divided by the integrated diffused concentration (0 to 50 nm) gives a unique fraction for a given size (Fig. 2.10.f). The ratio error has been calculated using an error of 5% of concentration in each fraction. Due to the sigmoidal shape of the relation between radius and ratio, the error on the radius can be significant, especially for the smallest and largest species. For a given design of H-filter, the flow rate can be tuned in order to decrease the error for a given size range.

2.4.2 Results

The example used for the proof of concept of the techniques is a study of interaction between lipid vesicles made of DOPS and α -synuclein monomer and oligomers (see Figure 2.11.a). The α -synuclein of cysteine mutant N122C is the one used for precise labelling at desired positions with Alexa 647 maleimide dye. Studies exist on α -synuclein monomer interaction with vesicles made of DOPS [66]. This particular mutant associated with the lipid vesicle has already been characterised within fibril form on vesicles composed partially of DOPS [113], but further work on the early stage of aggregation would be interesting. The adsorption on the lipid layer increases the probability that two species encounter. Indeed, the system passes from a 3D random walk for species in solution to a 2D random walk for species on the surface. Further knowledge can lead to a better understanding of the early stages of self-association close to the surface. These results are preliminary data intended to show how the technique works, and further research is required for an explanation of interactions of early aggregation species with lipid vesicles.

Repeatability & Dilution. Two factors are determinant within an electropherogram: the repeatability of the intensity of the peak and the time it appears. Figure 2.11.b shows the repeatability of different runs of the capillary electrophoresis experiment. The repeatability

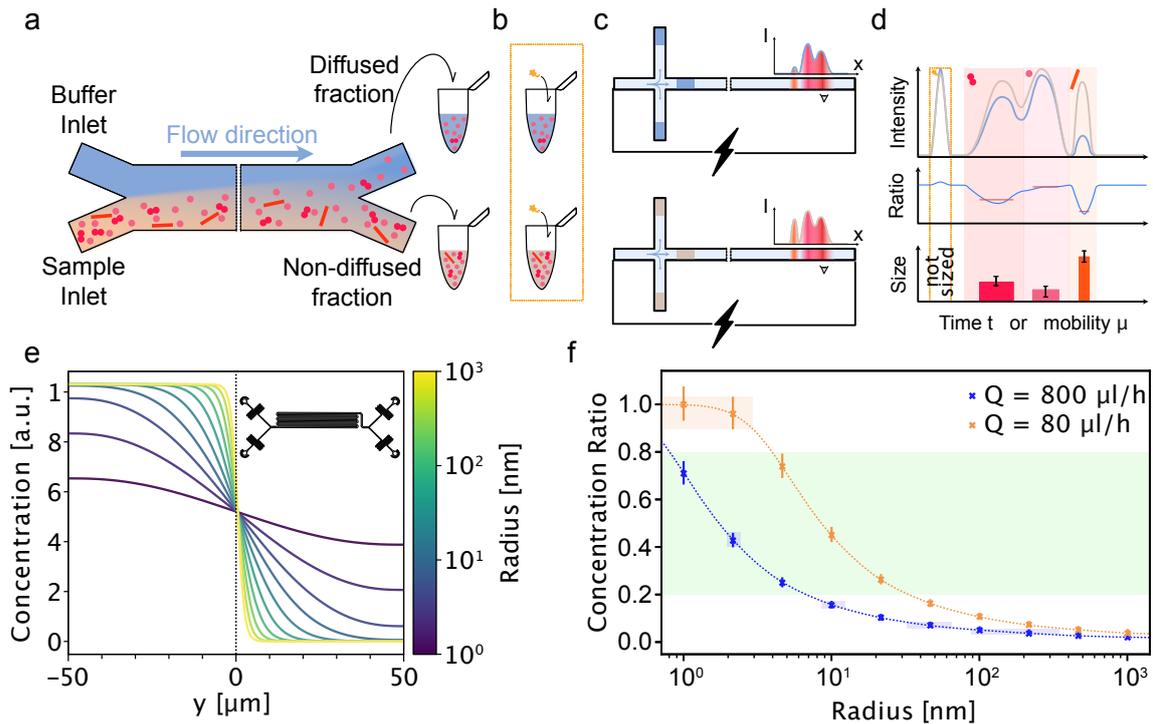


Fig. 2.10 Second device work flow. **a-e.** Pipeline of the measurement process. **a.** Sizing step. Separation and collection of diffused fraction and non-diffused fraction using H-filter. **b.** Labelling step. Possibility to label after the sizing step. **c.** Separation step. Separation of species in each fraction using capillary electrophoresis. **d.** Measurement. Superposing the two fractions and sizing. **e.** Calculation of the concentration profile for different particle sizes. This should be carried out for different devices. **f.** Size of particles as a function of the fraction for a given device at a given total flow rate. It is important to tune the flow rate in terms of the expected size. The shaded areas show the error in size for a given point, the green shaded area gives the ideal fraction value for minimising the error. The error bars represent the maximal and minimal ratio errors assuming 5% error in the concentration measurements.

of the position works better without the introduction of the lipid vesicles; indeed the vesicles tend to burst and coat the surface. This not only modifies the electro-osmotic flow, but also creates interaction between the species and the newly coated surfaces. This change increases with the time passed within the capillary. Such change can be seen in the last peak of the blue curve, which is shifted. Interestingly, the next two bumps seem not to be affected. The influence of the modification of electro-osmotic flow is negligible and most of the impact comes from the interaction of species with the coated layers. Figure 2.11.c shows a control on the behaviour of intensity with concentration dilution. It is important to make sure that the intensity I is proportional to the concentration c of the species. The result coincides with the expected trend in dashed blue. The shaded value corresponds to the standard deviation computed over several runs. The dilutions 1 to 2 and 4 change in intensity from run to run, which may be due to the quantity of sample injected. When the mean is considered the value is consistent. Furthermore, the chip and focus can give variation on the intensity. It is therefore important to use and load the different samples within the same chip and without changing the detection position.

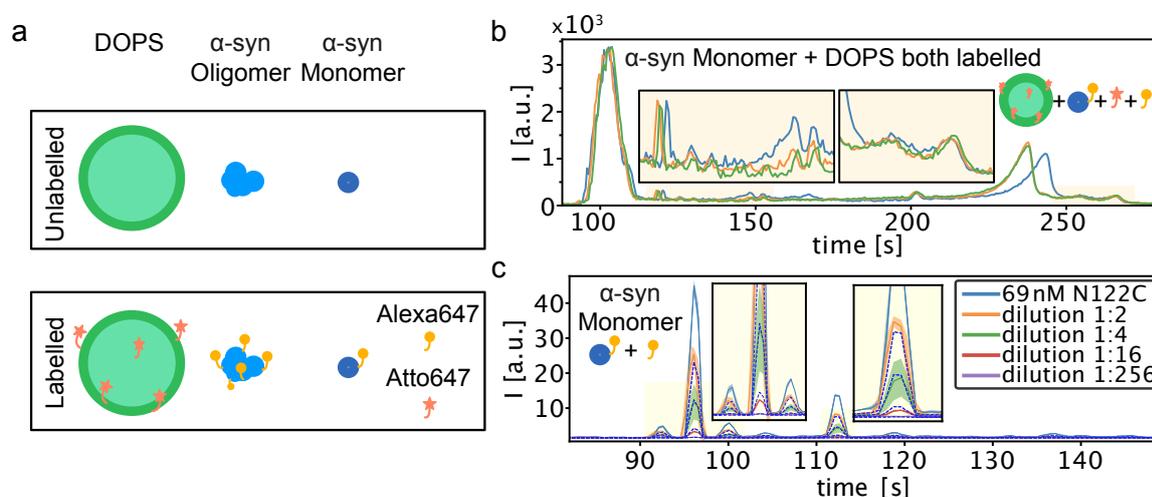


Fig. 2.11 Model used for second device. Reproducibility and dilution of electropherogram. **a.** Description of the model used. Three different components are analysed. These components interact with each other: Unilamellar vesicles made of DOPS, α -syn oligomer and α -syn monomer. Two different dyes are used to stain the component, Atto647 and Alexa647. **b.** Several runs are displayed to show the reproducibility. The solution is a mixture of 50 μ M DOPS with 1% atto dye and N122C monomer 1 μ M previously labelled with Alexa647. When DOPS is used within the chips some peaks are shifted, reducing the reproducibility. **c.** Controlled dilution. Curve with dilution made to see if the intensity is proportional to the concentration. The dash blue lines correspond to expected values given the 69 nM N122C curve.

CE run and sizing. As explained above, the size can be extracted from the concentration ratio between the diffused and undiffused fraction. Taking at least three different runs in order to get the average intensity of each of the runs, it is possible to obtain the ratio and therefore the size (see Figure 2.12). The example of this figure shows a mixture of unlabelled vesicles with labelled α -synuclein monomers. The first peak size appears to be above 15.37 nm, which corresponds to a ratio of less than 0.2. The size error starts to be too large to give a quantitative result. It is the same for smaller radii. The shaded red area highlights the size which corresponds to a ratio above 0.8. The size of the vesicle is expected to be much larger than 15 nm. The difference in size between the protein and the vesicle is not suitable for an accurate simultaneous sizing of each of the species. The mixing with DOPS creates a strange behaviour on the tail of the trace. The mixture of species arrives way after it was expected, probably due to absorption on the wall of the label. In the next figures, the focus will be on earlier times. The shaded area corresponds to the standard deviation of triplicate measurements. Figure 2.13 shows the value of α -synuclein monomers with dye in (a), and within a mixture with DOPS vesicles in (b). The first panel (a) highlights the dye values using shaded yellow areas. As expected, the fitted size is very small, and some points have a ratio above zero due to the error. It seems that the Alexa647 mixture has several mobility peaks. The monomer N122C peak is highlighted in blue and is sized with a value of $2.71 \text{ nm} \pm 0.21 \text{ nm}$. Literature reports a value closer to $2.9 \text{ nm} \pm 0.1 \text{ nm}$ [56]. Panel (b) shows different variants of the mixture: first with the label on the monomer species, then on the vesicle only and lastly on both species. The size of the labelled monomer is similar to the previous figure. It contains several plateaus visible after the green area of the vesicle. It is difficult to find a correct interpretation for each of them. Interestingly, the last one with a size of $4.22 \text{ nm} \pm 0.31 \text{ nm}$ is close to the position where the labelled monomer should stand. It is larger than expected. The difference in size can be due to the overlapping of the peaks. The second value of the panel (b) could not show any single monomer as none of them are labelled - no strong difference appears in the electropherogram of labelled vesicles alone. The size of the direct measurement of the vesicles gives a value up to 100 nm and saturated the simulation of the ratio. This value is closer to the expected size even if this sizing is purely qualitative. The last graph of the panel (b) corresponds to both labelled vesicle and labelled monomer. Two peaks can be seen. The second - most likely related to the monomer part of the electropherogram - has a calculated size of $5.3 \text{ nm} \pm 1.2 \text{ nm}$, which is two times higher than expected. The reason for conducting the three labelling cases of all species in solution is to ensure that the labelling does not strongly affect the interaction between monomer and vesicle. A final experiment of the sizing of the stabilised

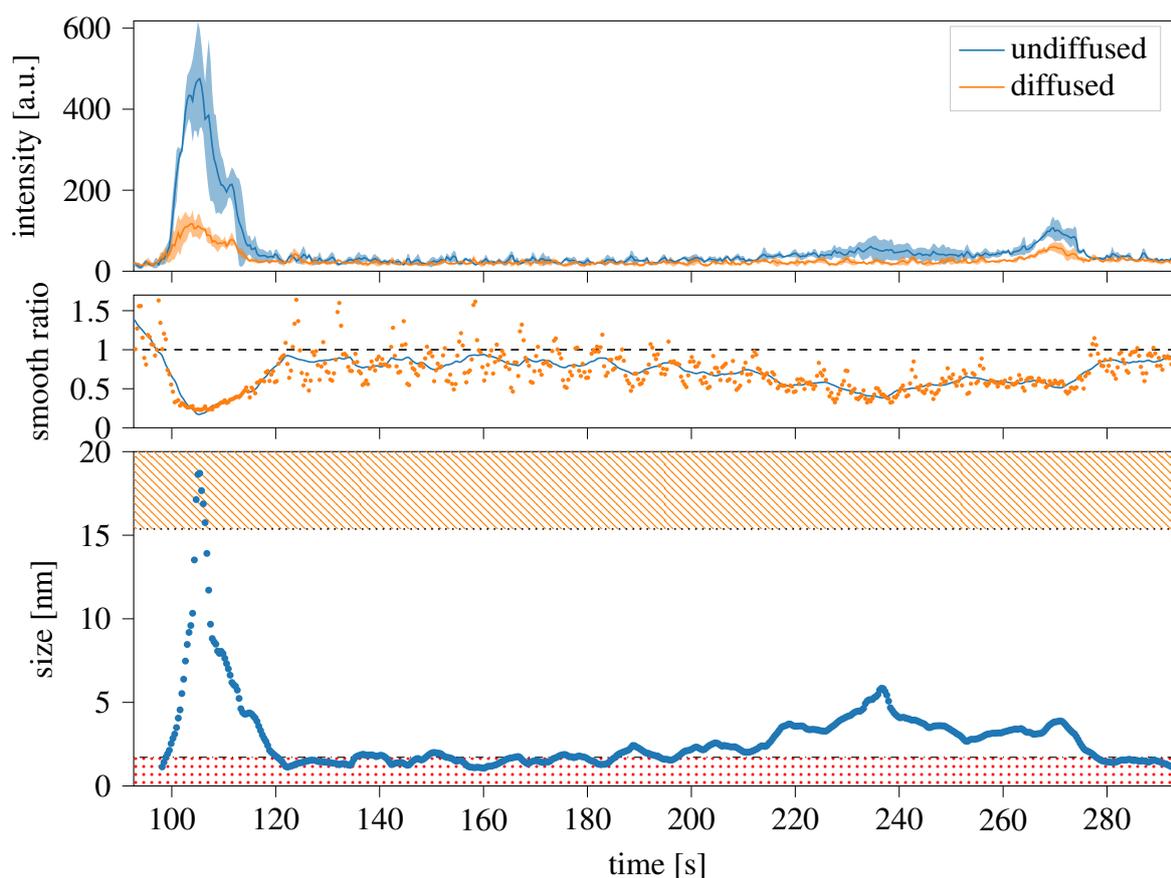


Fig. 2.12 Explanation of diffused/undiffused electropherogram curve, and relation to the size. Sizing of DOPS mixture with labelled α syn N122C mutant. The top panel shows the diffused and undiffused run, the shaded area corresponds to the standard deviation of three different runs. The panel in the middle shows with the orange dots the calculated ratio between diffused and non-diffused traces. The blue curve corresponds to a smooth trace extracted from the dots. The bottom panel shows the size found as a function of the smooth ratio. The orange-lined shaded area corresponds to an error on the size too big to be quantitative, the value can be much higher. The red dotted area corresponds to a ratio too small to be properly sized, and the calculated radius in this area can be zero.

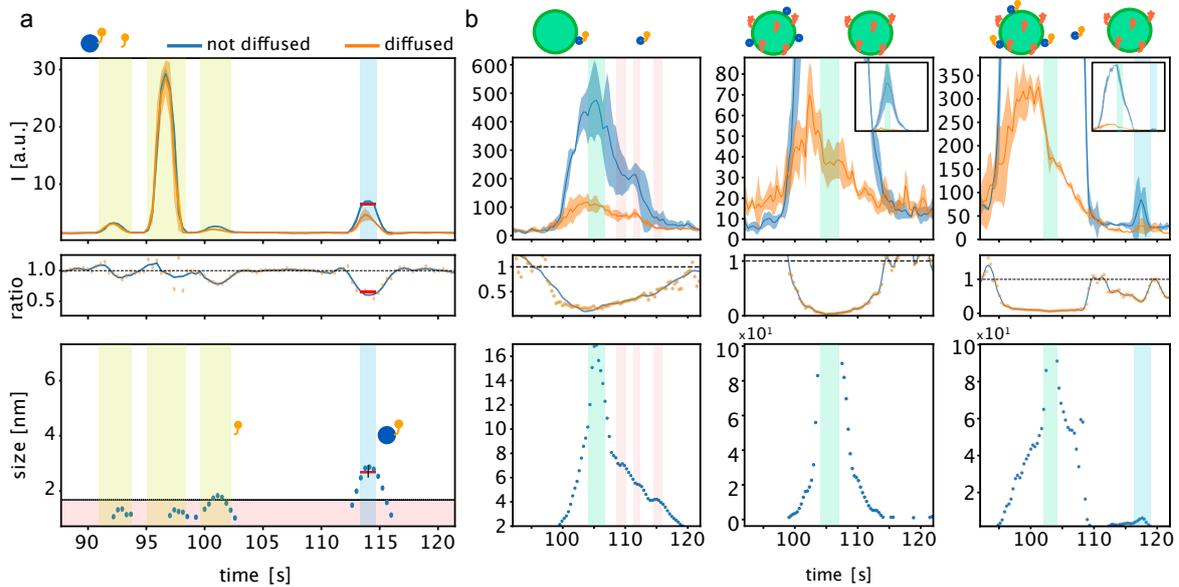


Fig. 2.13 Electropherograms of α syn monomer, and large unilamellar vesicle of DOPS. **a.** Sizing of N122C Alexa647-labelled α syn mutant monomer. The shaded area highlights free Alexa dye in yellow and the complex monomer/Alexa in blue. **b.** Sizing of $1\ \mu\text{M}$ of N122C mixed with $50\ \mu\text{M}$ of Large Unilamellar Vesicle of DOPS, from left to right it represents labelled monomer only, labelled vesicle only, and both labelled. The shaded areas around the lines on the top panels correspond to the standard deviation of triplicate measurements.

α -synuclein oligomer has been conducted. Figure 2.14 represents the result. Two more peaks than with the monomer only curve appeared in the experiment. Interestingly, it appears with a lower mobility than the expected monomer peak. This is unexpected as the mobility from self-associated particles should always be higher than its monomeric form [13]. One explanation could be that oligomer species interact much more with the surface and therefore arrive later in the detection area. The sizing part seems to suffer from a miss-alignment of peaks. The alignment should be performed carefully as the intensity depends of the speed of the species at the detection area, but as the length of the channel is constant the speed is inversely proportional to the time until reaching the detection area. The integration of the third and fourth peaks give a value of $4.16\ \text{nm} \pm 0.22\ \text{nm}$ and $6.8\ \text{nm} \pm 0.7\ \text{nm}$, respectively. The size of monomeric species from these runs is overestimated and the oligomer run is underestimated compared to the traditional diffusional sizing experiments, which size the oligomer closer to $7.5\ \text{nm}$.

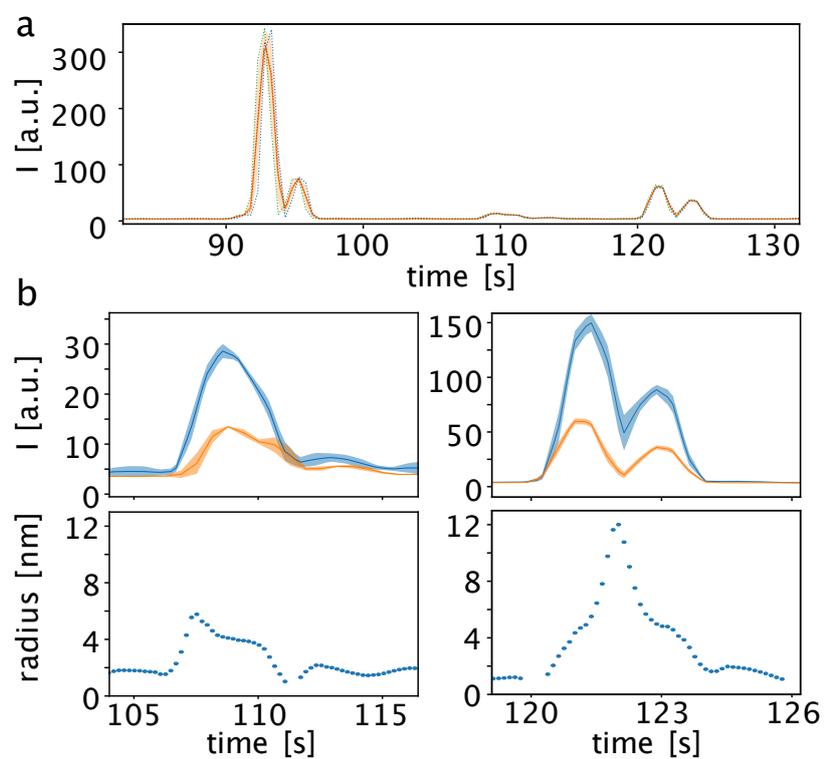


Fig. 2.14 Electropherogram of α syn oligomer. **a.** Oligomer trace within capillary electrophoresis. **b.** Sizing of the third and fourth peaks, the small change within the curve showing a great change in size. The integral over a peak allows an average size of the species within this peak, and is more robust. The shaded area corresponds to the standard deviation of three different runs.

2.4.3 Methods

H-filter The flow in the H-filter is controlled by syringe pumps in pushing mode. The sample is loaded within one syringe, pushing in the sample inlet and a second push the buffer through the buffer inlet. The flow rate ratio is 0.55/1 for the buffer inlet. For a total flow rate of $80 \mu\text{Lh}^{-1}$, the buffer is injected at $44 \mu\text{Lh}^{-1}$ and sample at $36 \mu\text{Lh}^{-1}$.

CE The custom-made device used above has been replaced by a more versatile unit. The potential is created by four different positive high voltage power supplies controlled by 0-10V USB Analog Output Module. One power supply can reach 0-25kV, and the three others 0-5kV. Each of the independent wires is connected to the microfluidic device. This allows a higher flexibility in the separation phase, allowing the current to be adjusted such that most of the sample plug goes within the separation channel. The injections made in the experiment are made using a pressure driven injection. Four pressure driven controllers from -1 to +1 bar are connected as well within the chips and allowed to fine tune the injection process.

Large Unilamellar Vesicles (LUVs) preparation A dry and thin lipid (DOPS from Avanti Polar Lipids Inc. +1 mol % DOPE-Atto647 from ATTO-TEC GmbH) film was prepared by gently evaporating the chloroform from the stock solution of lipids in chloroform using a nitrogen stream. A 100mM HEPES-BisTris buffer was added to hydrate the lipid film to the concentration of 5 mM and the solution was stirred at room temperature. The lipid solution was extruded 31 times through polycarbonate membranes (1 μm and 100 nm from Avanti Polar Lipids Inc.) using a manual extruder (avanti polar lipids Inc.). The DLS was used to control the uniformity of produced vesicles. The liposome solution was stored at 4°C and used within 2 days.

2.5 Conclusion

The first generation device I have shown couples CE & H-filter in the same chip. It gave promising results and helped me to understand how the general principle works, but it was not the most suitable to perform reproducible experiments. The number of inlets and outlets make the device hardly usable with the 4-pressure controller that we had available. It makes it more complicated to control the flow within the chip and creates a source of error. Splitting the system into two successive experiments allowed better control of different sources of error. The second generation device offers greater freedom for conducting experiments, and gives more reliable results. Capillary electrophoresis allowed separation of different

species, while the diffusional sizing of each of the components through the H-filter gives an interesting way to get the size distribution of a mixture. The non-obviousness of being able to size the different components after passing in the H-filter makes this invention more versatile, as changing the order of the steps allows adding the dye after the sizing step of individual species, even if the information is not yet available. The proof of concept worked, and future experiments are already planned, including the sizing of oligomer species at different stages of the aggregation process, or sizing within human serum of specific antibody. Considering that the resolution of the microscope will improve the precision of the ratio, the described method combined with single molecule detection microscopy opens up the possibility of experiments at a lower detection limits, but also improved precision of obtained values of sizes. Better objectives nevertheless have short working distances, which are often too short for reaching the solution with the added layer of PDMS on top of the coverslip. Plastic injection moulding microfluidic chips have therefore been ordered and will be tested on a confocal microscope.

Chapter 3

Single molecule detection and limitations of microfluidic experiments

The content of this chapter is based on three manuscripts in preparation.

Jieyuan Fan, Georg Krainer*, Raphael P. B. Jacquat*, Matthias M. Schneider*, Timothy J. Welsh*, Quentin A. E. Peter, Ewa A. Andrzejewska, Lin Chai, William E. Arter, Kadi L. Saar, Therese W. Herling, Vasilis Kosmoliaptis, and Tuomas P.J. Knowles*
"Single-molecule microfluidic diffusional sizing (smDS)"

Georg Krainer, Kadi L. Saar, William E. Arter, Raphael P. B. Jacquat, Quentin Peter, Pavankumar Challa, Christopher G. Taylor, David Klenerman, Tuomas P.J. Knowles
"Direct digital sensing of proteins in solution through single-molecule optofluidics"

Georg Krainer, Catherine K. Xu*, William E. Arter*, Raphael P. B. Jacquat, Quentin Peter, Timothy J. Welsh, Marta Castellana-Cruz, Therese W. Herling, Kadi L. Saar, Janet R. Kumita, Christopher M. Dobson, Tuomas P. J. Knowles*
"In-solution structural characterization and detection of α -synuclein oligomers by single-molecule free-flow electrophoresis"

(contributed equally)*

A preprint version of the second paper can be found in the reference [105]. I contributed by developing the analysis tools, the data analysis, designing the different experiments in the first paper, and performing some measurements. This chapter describes the adaptation of two existing microfluidic techniques to single molecule measurements, both with the aim of achieving the lowest concentration detection limit using a confocal setup. The first method is based on diffusional sizing (DS), the second on micro free-flow electrophoresis (uFFE). The

first part of this chapter focuses on the application of microfluidic diffusional sizing at low concentration. Scanning a confocal volume across a microfluidic chip improves the limit of detection from nanomolar down to femtomolar concentration compared to epifluorescence detection. An increase in sensitivity enables us to detect and to quantify high-affinity reaction events where the sigmoid curve of the reaction is switched at low concentrations [87]. The device combined with the confocal spot gives as well the possibility to distinguish the brightest species in a heterogeneous solution, or even several species within the same reaction. The second part of this chapter explores briefly the application of uFFE. An electric field in a large channel perpendicular to the flow separates heterogeneous solutions into different species based on their mobility. The confocal volume scans across the main chamber and allows us to distinguish the different subpopulations of heterogeneous analyte mixtures. Applications include separation of monomer from oligomer or aggregate, or identification binding and quantification of a reaction rate.

3.1 Single molecule confocal setup

Single molecule detection allows a more sensitive approach in the detection of biological phenomena like the understanding of protein folding [92, 110], or interaction between two particles [44, 87, 112, 164]. The experiments described in this chapter were performed with a confocal microscope setup using a 488 nm laser which allows us to detect a specific dye attached to a single particle within a spot about four femtoliter sized. A representation of the instrument is shown in Figure 3.1. The dynamic range of confocal microscopy is limited at high concentrations in distinguishing two particles arriving at the same time, and at low concentration limited by the time for a particle to reach the detection limit. In order to decrease the lower limit, a flow can be applied decreasing the time for the confocal to detect the desired molecules. This comes at the cost of reducing the total brightness emitted by the molecules, as they spend less time inside the laser spot. The maximal flow rate depends then on the noise level of the confocal. An interesting solution has been developed and put on the market using Quanterix [42]. Its principle is based on capturing molecules of interest on microbeads and uses a concentration such that the Poisson probability to find one or zero molecule is much higher than the probability to find two or more molecules bound, then compartmentalise the microbeads and make a chain reaction in order to detect any fluorescence. This should work with a compartmentalised droplet, in a similar way to the work of Pfammatter [155]. The combination of this more sensitive microscope with well-established microfluidic techniques allowed us to push the probing to a lower concentration

than traditional epifluorescence and without requiring single molecule analysis. The confocal has a better sensitivity and better signal to noise ratio than a CCD camera. Furthermore, microfluidics allowed a flow control and thus a better discrimination of false positive triggers. The single molecule detection of the trace seems to reduce the noise and therefore gives a better result. However, the noise level can be considered as constant, and single molecule analysis has a probability of false negative and false positive counts that reduces the quality of the reading. Nevertheless, single molecule analysis allows us to extract extra information, such as the number of photons emitted by a single event (also called burst), if the sample is diluted enough to be in the single molecule regime. The number of photons emitted per event informs about a number of free dyes exposed to the confocal by the particle. Therefore it is possible up to a certain extent to distinguish species in a heterogeneous sample. This is limited, as the response to an excitation depends on how the protein is labelled and it does not scale necessarily linearly with the number of dyes attached to them, due to quenching effects. The labelling efficiency is an additional factor which reduces the ability to distinguish between different sizes of self-associated proteins. As explained by *Dr Seven F. Lee*, even if the labelling efficiency reaches 90%, only 35% of protein complexes composed of 10 proteins will be fully labelled [115].

3.1.1 Detection and representation

The trace representation from a photo-detector is often in number of photons over time (see Figure 3.1 panel b) even if this representation loses some information compared to a graph of interphoton time versus index of photon pairs. Indeed the interphoton time representation (Figure 3.1 panel c) does not request to bin a number of photons per an arbitrary time. The detection part is composed of two different thresholds based on the interphoton time representation. An event is considered as detected if enough photons arrive within a given time frame. The probability of having two photons coming from the noise in a short time difference depends on a Poisson distribution. The probability of a false positive decreases if several photons arrive in a short time respective to each other. This is why two thresholds are used. The first determines the interphoton under which the event occurred. The second threshold selects a minimum number of photons to distinguish the potential event from the first threshold. Figure 3.1.c shows in dotted black line the interphoton time threshold, from all potential events occurring below the line. Only the parts highlighted in red show the selected events which have a high enough photon count. The determination of the threshold can be achieved using a background profile, but it would not be accurate because the sample

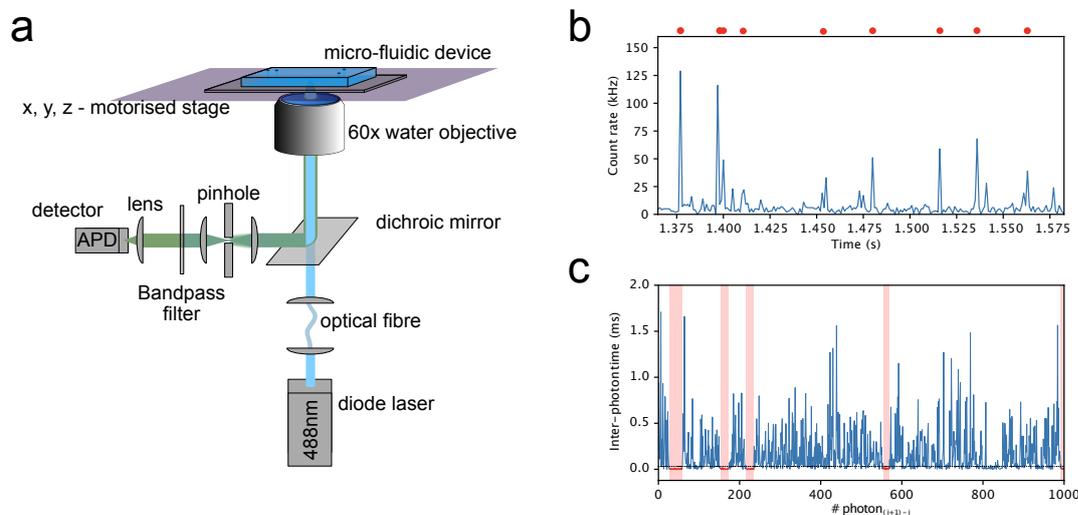


Fig. 3.1 (a) Schematic of the setup used. (b) Example of a trace representing the time versus number of photons per binning of time. The sample is 10 pM of HSA in $100 \mu\text{L h}^{-1}$ flow rate, the red dots represent the single molecule detected. (c) Index of interphoton versus the interphoton time. An event (in red) is considered as detected when a certain number of photons arrived below a certain number of interphoton time (dotted black line).

itself would modify the noise level; indeed, analyte samples often contain different slightly fluorescent molecules. Another technique to automatise the threshold level is by selecting a percentile of the total interphoton time distribution. Nevertheless this solution requires us to know the particle concentration level. Often a visual check by an operator still needs to occur to ensure a correct determination of the parameter.

3.1.2 Limit of detection and heterogeneous measurement

The confocal technique by itself decreases the detection limit, which means it can achieve detection of molecules at lower concentrations than epifluorescence microscopes. It is not necessarily dependent on any single molecule detection regime. The limit within the single molecule detection depends not only on the sensitivity of the techniques, but also on the distinction between two molecules. If several particles stand within the confocal volume, it cannot be resolved and only one event or burst will be counted. The upper concentration limit of single molecule detection from a confocal depends therefore on the laser spot volume present within the solution. The confocal volume is 4 fl , its upper concentration limit is 50 pM , and above this concentration the probability to have two or more molecules detected at the same time becomes larger than 1% (see Figure 3.2 .c). If the volume of liquid exposed is reduced due to application of nanofluidic devices, this upper limit can be expanded. In

theory, if the dye used is bright enough, there is no lower limit for the concentration. However, the time required to detect at least one molecule based only on diffusion can be very long. It can be calculated using Poisson statistics and looking at the probability to encounter one molecule at a given time t within the confocal volume (see Figure 3.2.a). An example is shown in Figure 3.2.b for BSA protein with a diffusion coefficient of $5.9 \times 10^{-11} \text{ m}^2 \text{ s}^{-1}$, which requires more than 5 minutes to have a 99% chance to detect at least one particle at 1 pM concentration, and this is without taking into account false negative probability. In the case of 1 fM, the time required to detect one molecule is closer to 10h. The time to detect particles is reduced under continuous flow. It affects the sensitivity as the dye spends less time within the detection volume. There is therefore a trade-off between the integration time required to have enough molecules detected, and the ability to detect the particle. It depends on the system, and the brightness of the dye. The fluorescence can be used to distinguish species by their brightness. In the case of monomer versus oligomer, if the dye is attached to the monomer, the brightness should in theory scale with the number of monomers composing the oligomer. But the path through the confocal spot is not necessarily centred at the middle of the laser (see Figure 3.2.d). This problem could be resolved by a normalisation of the number of photons detected with the time spent within the confocal detection volume. The number of photons per unit of time is a more accurate unit to compare, but not only the time spent plays a role, the illumination within the detection volume is non-uniform. Hence, the quicker events have a different normalised intensity than the one passing through the centre. It is still possible in continuous flow microfluidic experiments to create a sub-selection of the longest events, and reduce the pool of detected particles. The influence of the flow rate on the particle velocity should be dominant compared to the Brownian motion. The microfluidic conditions ensure no turbulent flow which may modify the path in the confocal spot. As the flow rate is controlled within the microfluidic, bursts that are longer than expected can be rejected.

3.2 Single Molecule Microfluidic Diffusional Sizing (smMDS)

Microfluidic diffusional sizing (MDS) has been widely used in research labs for ten years [8, 31, 168, 173, 207]. It allows us to measure the size of a component/components in solution flowing on-chip. It uses the physics behaviour at low Reynolds numbers from microfluidics to remove most of the convective mixing and allows only diffusive mixing. Two streams are injected and intersect in a single channel (diffusion channel) where they co-flow in parallel and mix only by diffusion. The time spent in the channel at a given flow rate depends on the

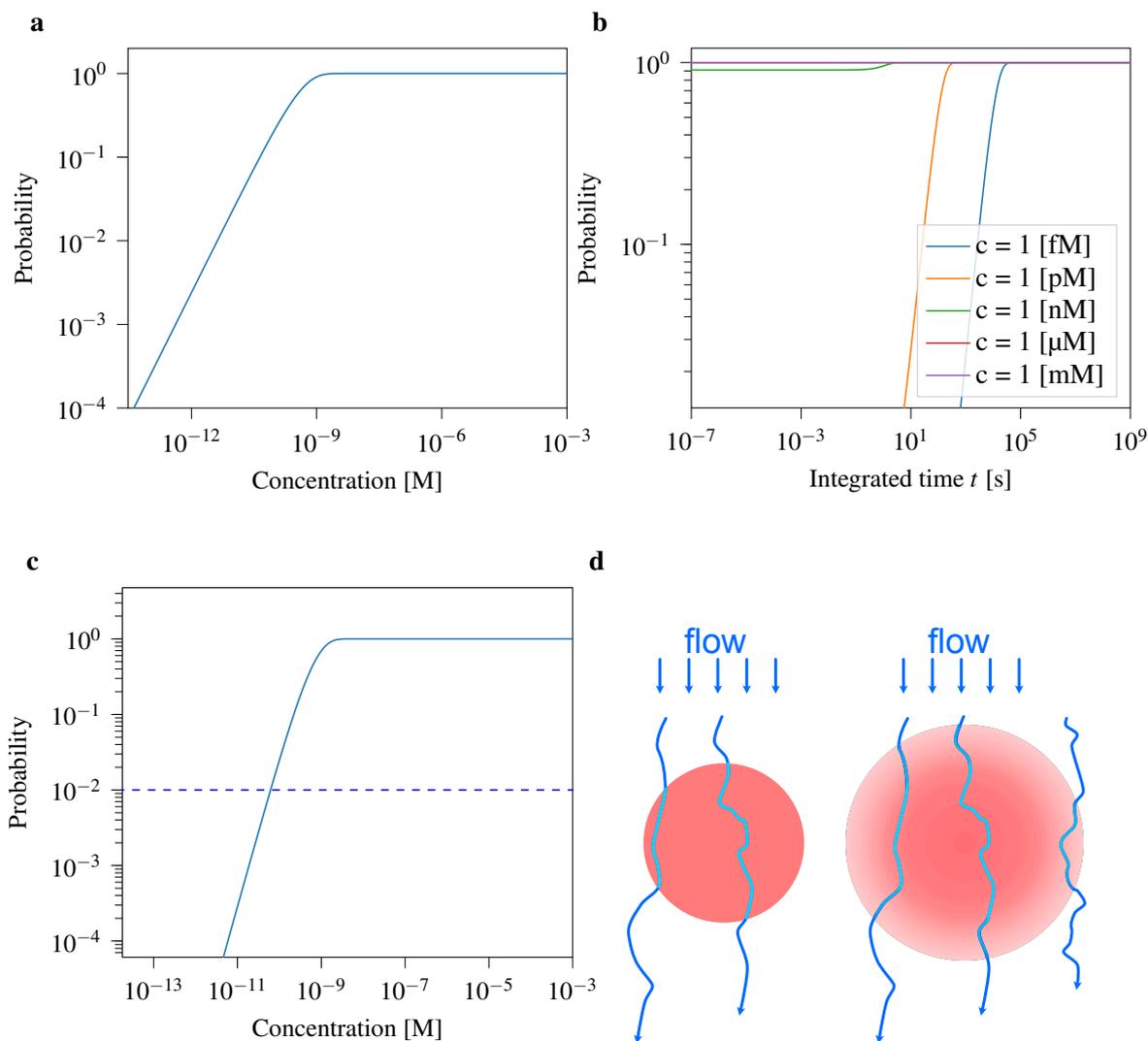


Fig. 3.2 a. Probability of encountering at least one molecule inside the confocal detection volume of 4 fL as a function of the concentration. **b.** Probability of encountering at least one molecule within a given time frame at different concentrations. The molecule is diffusion limited, driven by Brownian motion. It has the diffusion coefficient of BSA. **c.** Probability of encountering 2 or more molecules simultaneously for a given concentration. The 1% probability is highlighted in dashline. **d.** The image on the left represents the ideal case of a uniform excitation light confocal spot. The normalisation of the total photons emitted by the time passed within the confocal spot allows comparison between particles. The image on the right side represents an illustration of non-uniformity of the laser spot. The normalisation by the particle travelled time within the confocal cannot compare events passing on the side and in the middle of the spot.

distance travelled within the main channel. Recording the evolution of the diffusion profile at different parts of this channel allows us to deconvolute the diffusion coefficient of the components. The diffusion coefficient, D , is inversely proportional to the hydrodynamic radius R_H of the analyte. This technique is reliable, and requires only a small volume of sample. One downside is the inability to dissociate different components of similar radius, e.g. monomer & dimer. The fusion of these techniques with confocal microscopy opens up several possibilities discussed below, with the ability to solve heterogeneous mixtures and the determination of high affinity binding events thanks to the high sensitivity of the microscope.

3.2.1 Results

Principles and experimental demonstration of the smMDS platform The working principle and experimental implementation of the smMDS platform is depicted in Figure 3.3.a. A sample containing the fluorescent analyte is injected into the microfluidic chip. The analyte stream at the centre of the channel is co-flowed with an auxiliary buffer at the analyte entry nozzle, leading to diffusion of the analyte to either side of the channel when travelling downstream. To detect fluorescently labelled molecules in the channel and monitor the micron-scale diffusive mass transport, laser confocal fluorescence microscopy is used to scan the confocal volume across the microfluidic chip at the mid-height of the channel perpendicular to the flow direction. Two fiducial marks are used to keep a record of the position within the chip. The scan trajectory is chosen such that various positions along the channel are probed, including positions that are close to the nozzle where the sample meets the carrier medium, and others further away. The four innermost channels of the device are scanned, as these cover a wide range of distances and time points along the channel, such that biomolecular analytes (e.g. proteins) with hydrodynamic radii (R_h) in the range of 1–10 nm can be analysed. To demonstrate the principle of smMDS, the molecular diffusivity and size of human serum albumin (HSA) is measured as a control. Firstly, the sensitivity of the method is examined by measuring a dilution series of HSA from bulk concentrations (10 μM) down to single-molecule conditions (100 fM). To this end, HSA is pre-labelled with a fluorescent dye (Alexa Fluor 488), the sizing experiments are performed by moving the observation volume of the confocal microscope through the chip while maintaining a constant total flow rate of 100 μLh^{-1} . The buffer was phosphate buffered saline (PBS) supplemented with 0.01% Tween20 to prevent adhesion of molecules on chip surfaces. The recorded diffusion profiles are shown in Figure 3.3.b-c. In the range from 1 μM down to tens of pM HSA (Figure 3.3.b), sufficient molecular flux of protein molecules allowed

for the recording of diffusion profiles from continuous scanning experiments. As a result, the confocal spot is moved through the device at a constant scan speed of $20 \mu\text{m s}^{-1}$ and the fluorescence intensity from molecules flowing through the confocal volume is recorded. The diffusion profiles show the broadening due to diffusion of molecules along the channels. Narrow peaks are observed at channel positions close to the nozzle where the sample meets the auxiliary buffer and the peaks broaden further in the downstream channel positions. As the concentration approaches the single molecule regime (i.e., at and below 20 pM) (Figure 3.3.c), instead of continuously scanning through the device, the confocal volume is moved along the same trajectory in a stepwise manner through the channels (henceforth denoted as step scan experiments). Thereby, bursts of fluorescence corresponding to the passage of single analyte molecules passing through the confocal volume are recorded at each channel position for a specific period of time. In total, the confocal microscope performed 200 to 400 scan steps across the chip and recorded 2 to 4 second long fluorescence traces at each position. Using my burst-analysis algorithm, single molecule events were extracted for each position during the entire time. The resulting numbers of counted molecules as a function of chip position created a diffusion profile. In this way, size is obtained from diffusion profiles for HSA from 20 pM down to 100 fM. Notably, the obtained curves from both continuous-scan and single-molecule step-scan experiments exhibited similar shapes and showed the characteristic broadening due to diffusion along the channels. To extract the hydrodynamic radii R_H of the protein species in solution, the diffusion profiles of HSA are analysed using an advection-diffusion model. The algorithm used is available at doi:10.5281/zenodo.3881940; it was developed by Quentin A. E. Peter, based on an early code of Thomas MÅ¼ller. The first step of the algorithm is based on deconvolving the experimental profiles into a linear combination of profiles expected for particles with known diffusion coefficients (basis functions). A least-squares error algorithm was then used to find the linear combination yielding the lowest residuals, allowing the average hydrodynamic radii of the analyte to be determined. Applying this procedure to the HSA profiles acquired under all conditions yielded excellent fits of the experimental data (Figure 3.3.b-c). The extracted hydrodynamic radii were, within 10% error, in excellent agreement among all concentrations with previously reported values for HSA under similar experimental conditions [2, 7, 94] (Figure 3.3.d). The burst-analysis algorithm fitting parameters change as a function of the chip angle position, the slide cleanliness and purity of the buffer and solution. The validity and influence of the parameter selection for burst analysis of sizing experiments in the single-molecule regime has been studied (Figure 3.4). A wide range of burst selection parameters provides the expected size information, and this demonstrates the robustness of the technique. Hence,

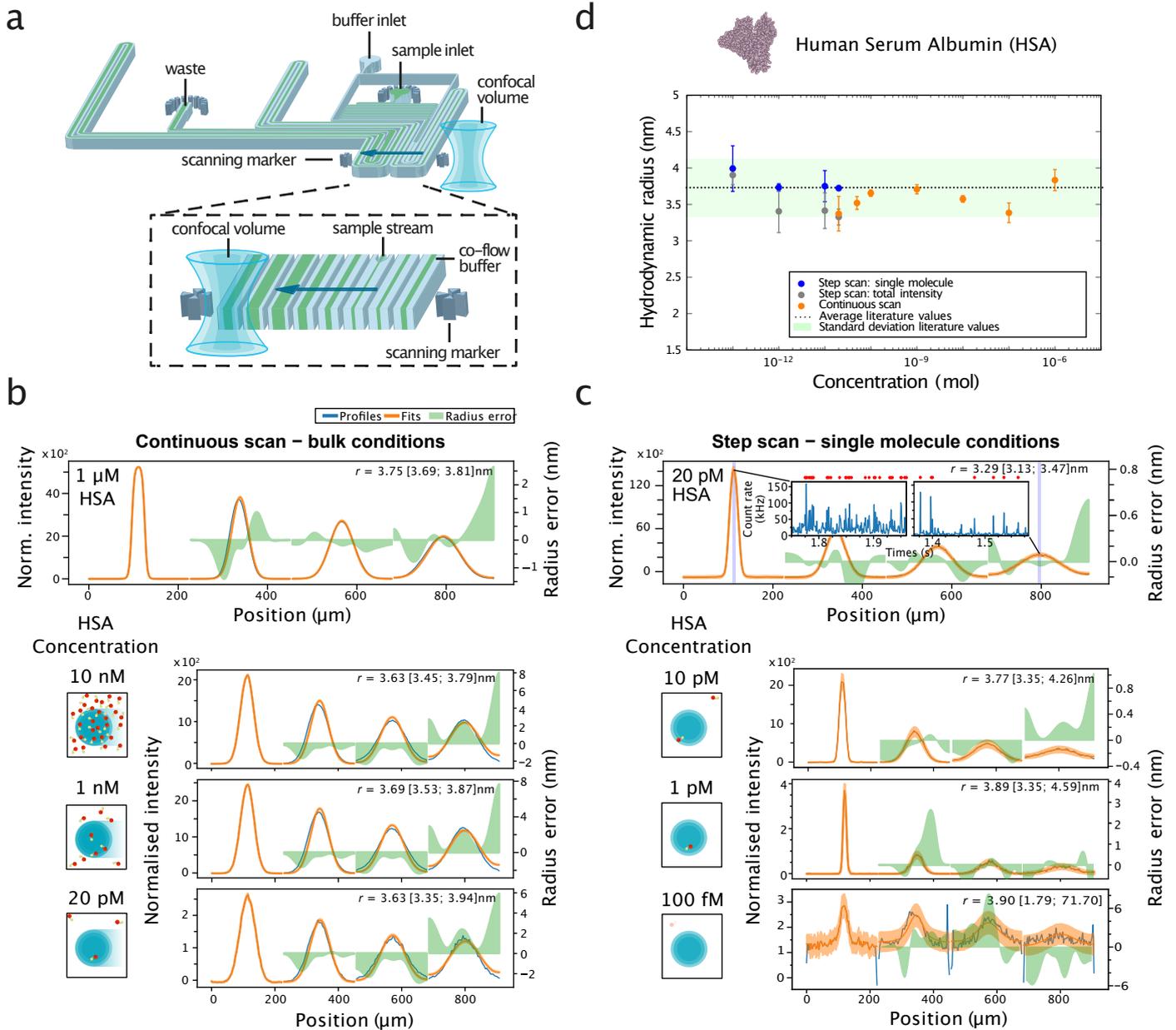


Fig. 3.3 Principle and experimental implementation of Single Molecule Microfluidic Diffusional Sizing (smMDS)

a. Schematic of the microfluidic chip design. The sample is injected into the sample inlet and arrives in the main channel at the sample stream, where it encounters the auxiliary buffer. Two scanning markers allow better reproducibility of the experiment, the confocal volume scan in between both markers. Eight sections of the channel can be scanned, and the focus is mainly on the four centred channel. **b.** Intensity profiles at different concentration of HSA, with corresponding size. All profiles are acquired using the continuous scan mode; single molecule detection is not possible. **c.** Intensity profiles at different concentrations of HSA, with corresponding size. All profiles are acquired using the step scan mode, working at the concentration range in which only one protein passes through the confocal volume, allowing single molecule counting. An example of single step trace is shown within the 20 pM HSA for the top of the first profile and last profile. **d.** Values of different sizes of has measured at different concentrations. The error bars represent the standard deviations of triplicate measurements. The literature values are taken from [2, 7, 94]. All data points have been acquired by Matthias M. Schneider and myself, and the analysis of the curve has been performed by myself.

the results demonstrate that smMDS provides accurate sizing information down to the fM concentration range. To compare the sensitivity of the smMDS technique to conventional MDS experiments, fluorescence widefield imaging experiments have been performed (Figure 3.5). Three different concentrations have been recorded: 1 μM , 100 nM, and 50 nM of labelled HSA. Each of the images have their uneven illumination corrected by dividing the result of a two-dimensional least squares polynomial fit of degree 2. While at the lowest concentration the sizing algorithm cannot even be run (50 nM), the algorithm converges to a size for the 100 nM has, after processing the image. The brightest 0.5% of pixels with an extreme difference to their neighbours ($>100\%$) have been rendered blank, and have not been taken into account for calculating the mean of the intensity profile. The measured sizes are $\sim 25\%$ higher than expected; the limit of conventional fluorescence widefield imaging is of the order of magnitude of 100 nM, similar to the autofluorescent device [30]. This shows that the smMDS technique presented here extends the sensitivity range of diffusional sizing experiments by more than 6 orders of magnitude.

Sizing of nanoscale analytes and protein complexes by smMDS Next, validation of the smMDS technique is performed by comparing the hydrodynamic radii of a series of nanoscale analytes obtained by smMDS with values obtained from conventional techniques. To this end, a set of molecules was selected, including a small organic fluorophore (Alexa 488), proteins (RNase A, Lysozyme, α -Synuclein monomer, Human Leukocyte Antigen (HLA), HSA, Thyroglobulin), and protein assemblies (α -Synuclein oligomers) covering a wide range of sizes in the 1 to 10 nm regime (Figure 3.6.a). The protein analytes were fluorescently labelled and purified before analysis. Every smMDS measurement was performed under single-molecule conditions at an analyte concentration of 10 pM in buffer as described above. The confocal spot scans in a stepwise manner through the channels and extracted single-molecule events for each analyte at each channel position and fitted the obtained diffusion profiles with the advection-diffusion model. The estimated hydrodynamic radii were then plotted against previously reported values (Figure 3.6.a). The values obtained by smMDS followed the expected trend within the margins of error. This demonstrates the excellent agreement between the values obtained from smMDS and literature values, revealing the reliability of the single molecule microfluidic assay. The average deviation in R_H for different repeats of the same sample is less than 5%. We further plotted the experimentally obtained hydrodynamic radii against the expected molecular weight. For both folded (RNase A, Lysozyme, Human Leukocyte Antigen (HLA), HSA, Thyroglobulin) and unfolded protein species (α -Synuclein monomer and oligomers), the data points followed the trend as expected

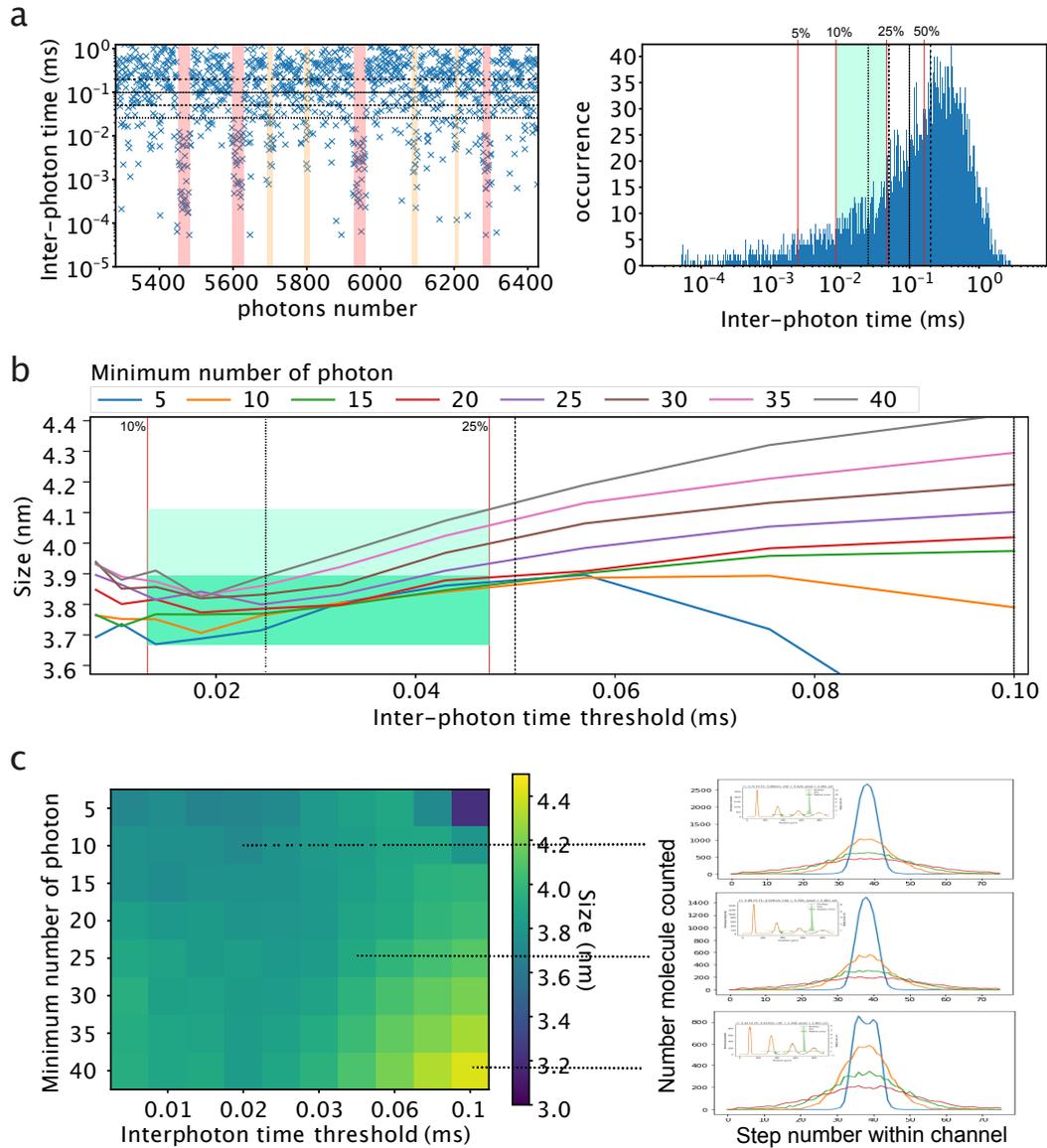


Fig. 3.4 Influence on sizing due to single molecule parameter on 20 pM HSA. **a.** On the left, interphoton time versus the indexed photons pairs, with the different black lines representing several thresholds. The red areas correspond to certain events of particle detection, the orange areas correspond to uncertain situations. On the right, the histogram of interphoton time (binning following a log space). The highlighted area in green shows the percentile range between 10% and 25% of event. **b.** Size found as a function of interphoton time in the x-axis and minimum number of photons in colour. The same green area as the histogram above is highlighted. **c.** A 2D representation of part b in a colour-map grid which represents the estimated size. On the right, three profiles from three different threshold conditions are displayed. A visual inspection of the main profiles extracted allowed us to discard the extreme values. The blue profile of the bottom graph shows that the minimum threshold is too high, merging different events together.

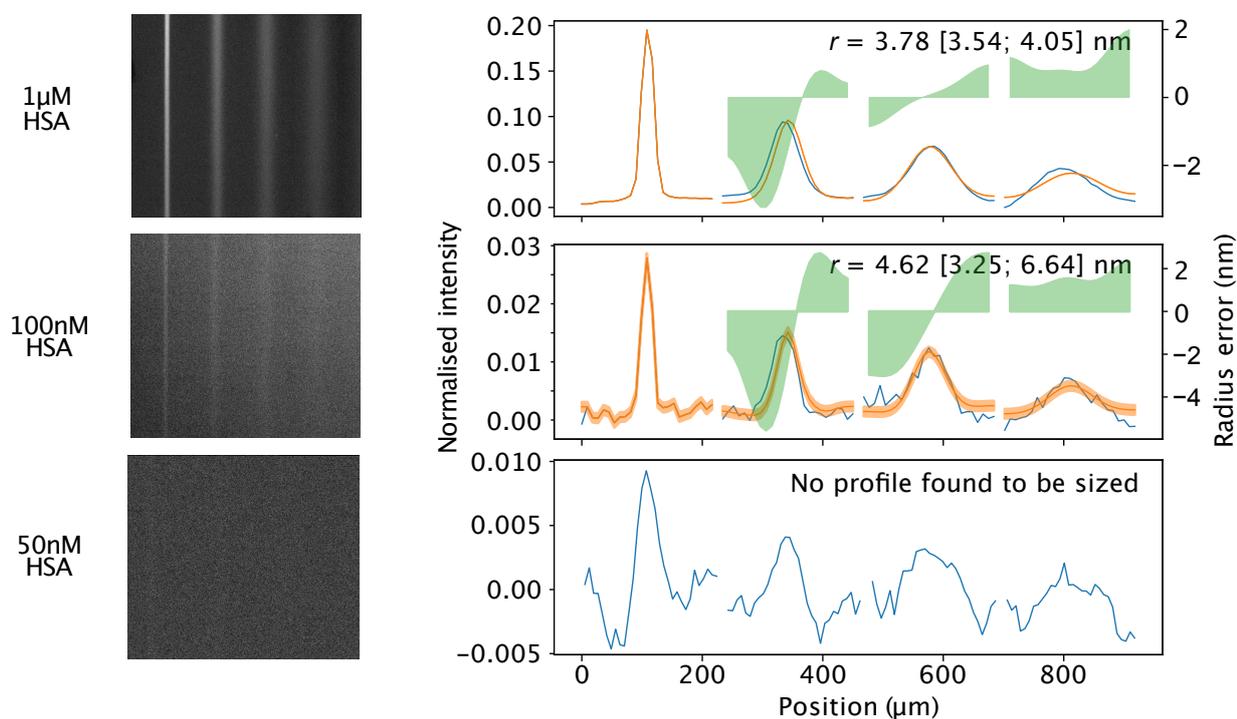


Fig. 3.5 Detection limit of widefield microscopy. Three images taken with widefield imaging, showing three different concentration profiles within microfluidic diffusional sizing chips. The diffusional sizing algorithm script returns a value of 3.78 nm for the concentration of 1 μM without prior data processing within the image. It returns a value of 4.62 nm for 100 nM with image processing prior to running the script, and stopped returning values at 50 nM.

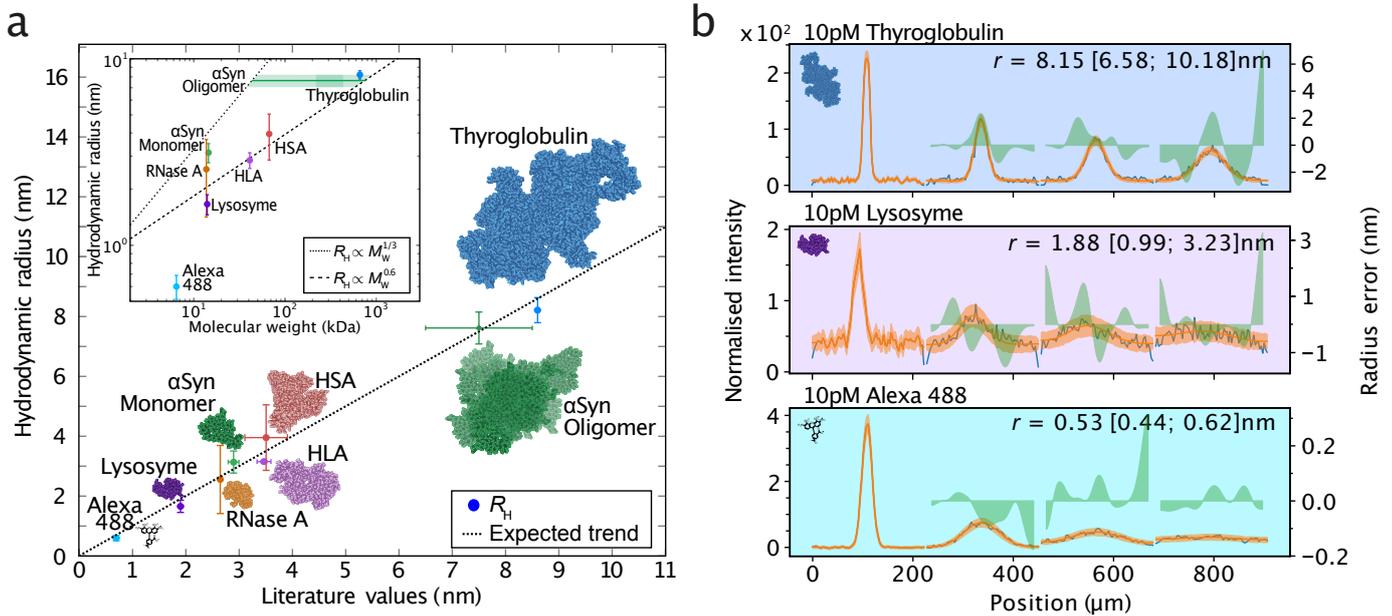


Fig. 3.6 Single molecule diffusional sizing of nanoscale analytes and protein complexes. a. Plot of experimentally determined hydrodynamic radii of biomolecular analytes by smMDS against literature values. The dashed line depicts the expected trend. Literature values come from [173] for HLA, [216] for lysozyme, [49] for RNaseA [174], [32, 56] for α Syn, [140] for Alexa488 and [2, 7, 94] for HSA. Inset figure compares the hydrodynamic radius R_H to the molecular weight M_W . Proteins should stand between two trends, the globular shape $R_H \propto M_W^{0.6}$ and unfolded shape $R_H \propto M_W^{1/3}$. The dye is completely off this trend. α Syn oligomers do not have a clear M_W , the shaded area corresponds to potential values, the darker one corresponds to the most probable, based on the assumption of the number of monomers per oligomer expected (~ 11). The error bars correspond to the standard deviation of triplicate measurements. **b.** Example of three different traces: Thyroglobulin, HLA, and Alexa488. *With the exception of HLA data points acquired by Jenny Fan, all other data points have been acquired by Matthias M. Schneider and myself, and the analysis of the curve has been performed by myself.*

for well-folded globular proteins or proteins that exhibit disordered protein regions (Figure 3.6.a, inset). The scaling exponent between the hydrodynamic radius of a protein and its molecular weight used for the trend is approximately 0.6, and 0.33 for unfolded or globular proteins [86, 159]. Note that these trends do not apply for small molecules such as Alexa 488, which has a hydrodynamic radius twice smaller than the model predicts.

Quantifying biomolecular interactions by smMDS A second application demonstrates the capability of smMDS to determine the binding affinity of biomolecular interactions. Interactions of proteins with secondary molecules, in particular with other proteins, are of great importance across the life sciences, and quantitative measurements of affinity constants

(i.e., dissociation constants K_D) have become vital in biomedical research and clinical diagnostics for affinity profiling [31, 59, 195]. Diffusional sizing allows for the detection of biomolecular interactions by monitoring the increase in size associated with binding and complex formation [8, 31, 123, 168, 173, 207]. By acquiring binding isotherms, affinity constants of the interaction can be determined under native conditions, without the need for purification or immobilisation on a surface. Furthermore, reaching lower concentration limits allowed us to determine higher affinity systems, as the binding curve occurs at lower K_D [112]. To demonstrate direct detection of protein binding and quantification of dissociation constants by smMDS at the single molecule level, we probed the binding of a clinically relevant antibody-antigen interaction. Specifically, we focus on the binding interaction between HLA A*03:01, an isoform of the major histocompatibility complex type I (MHC) and a key factor in the human immune system [22], and the antibody W6/32, a generic antibody that binds to all class I HLA molecules [91] (Figure 3.7). A series of smMDS experiments were performed by keeping the HLA concentration constant at 400 pM and adding increasing amounts of the W6/32 antibody. The different smMDS diffusion profiles were acquired and fitted to obtain effective radii across the concentration series. An increase in radius is observed from $3.18 \text{ nm} \pm 0.04 \text{ nm}$ for pure HLA, corresponding to a molecular weight of $\sim 50 \text{ kDa}$, as expected for pure HLA, to $5.08 \text{ nm} \pm 0.01 \text{ nm}$, corresponding to a molecular weight of 215 kDa, consistent with the binding of a 150 kDa antibody to HLA. Fitting the binding isotherm with a Bayesian inference combined with a mass balance model (see the methods in [173]), the dissociation constant was determined ($K_d = 400.5 \text{ pM} \pm 39.6 \text{ pM}$), in good agreement with previous results [173]. Remarkably, HLA is an extensively used clinical biomarker to assess, for example, the risk of allograft rejection. These results therefore outline a path towards detection and affinity profiling of antibody responses at the single molecule level with minimal sample consumption, and low concentration detection needed for high affinity measurements [87].

Differentiation within heterogeneous mixtures A key benefit of the smMDS technique, due to its single molecule sensitivity, is the capability to detect specific molecules within heterogeneous mixtures. That is to say, by rapidly detecting single photons, the single molecule extraction algorithm is able to easily resolve the bursts of intensity that can come from larger assemblies of molecules, oligomers or higher-order clusters, within a lower background signal of monomeric molecules. To this end, the capability for detecting two different types of assemblies has been displayed, all of which have distinct functions in functional biology and disease. The differentiation can be performed between single molecule

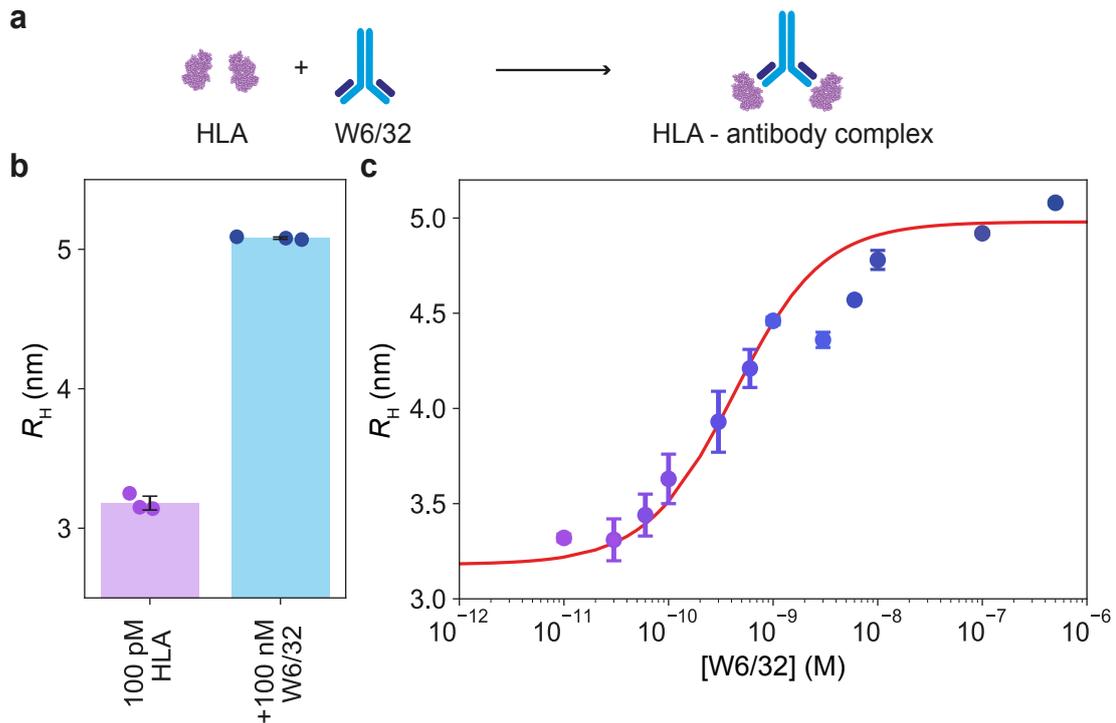


Fig. 3.7 a. Reaction between HLA and the antibody W6/32 exposing two binding sites resulting in a HLA-antibody complex. **b.** End result of the two extreme cases measured between 100pM HLA solution without W6/32, and 100pM HLA with 100nM of W6/32. **c.** Titration curve of antibody W6/32 within a solution of 100pM HLA, giving the binding curve of antibody W6/32 with HLA. *The error bars (in b & c) correspond to the standard deviations of triplicate measurement. All data points were obtained by Jenny Fan, and the analysis of the curve was performed by Matthias M. Schneider and myself.*

components, or between bright single molecules species within a bulk background. In the latter case, the bulk can be dissociated from bright single molecules and both components can be sized. In the first case a solution of α -syn oligomer is analysed; it is created out of 50 pM labelled monomeric concentration. The discrimination factor to remove the smaller species is the total number of photons. Indeed oligomeric species composed of a larger monomer unit will appear brighter. The measured sizes span from 3.6 nm for the average size by taking even the smallest species, to 16.5 nm by selecting only the brightest. The result is depicted in Figure 3.8.a. These values are in accordance with *Serene W. Chen, 2015* [32], which describes a variability in size within oligomer mixtures from 3 to 16 nm. The change in the minimum number of photons threshold is inadequate to differentiate the species directly. The less bright one would always be an average of every species, because the total number of emitted photons does not depend only on the number of emitters, but also on the path within the confocal volume. A very bright molecule can pass quickly on the side, and therefore emit less photons. Furthermore, the excitation intensity is not uniform, and it is therefore difficult to compare a normalised intensity of photon per time. As the particle time passed within the laser is determined by the flow rate ($\sim 100 \mu\text{L h}^{-1}$) rather than the diffusion coefficient, a sub-selection of the longest burst ensures that the particle path gets closer to the centre of the spot and the normalised illumination should be comparable. Naturally, bursts that last longer than the time to cross the confocal spot are discarded as they are most probably two subsequent events. An example of detected bursts normalised intensity distribution is displayed in the histogram of Figure 3.8.b. HSA is known to have a mixture of monomers, dimers and trimers which are roughly 89%, 10% and 1%, respectively [214], and previous studies using iSCAT microscopy can dissociate the different molecular weights [214]. In the histogram in Figure 3.8.b, no populations are distinguishable due to the broadening of this technique. However, knowing the different populations and their respective proportions, three ranges of normalised intensity have been selected and only the particles corresponding to this range are sized. In the right panel of Figure 3.8.b, the different sizes found in a selected range are shown. This is an average size of the different species composing the range. The error becomes important in the timer part as the number of molecules detected is very low. The sizes of 3.57 nm, 4.61 nm, and 5.13 nm correspond to a molecular mass of globular protein of 72 kDa, 157 kDa and 218 kDa, which seems to confirm the presence of monomer, dimer and trimer.

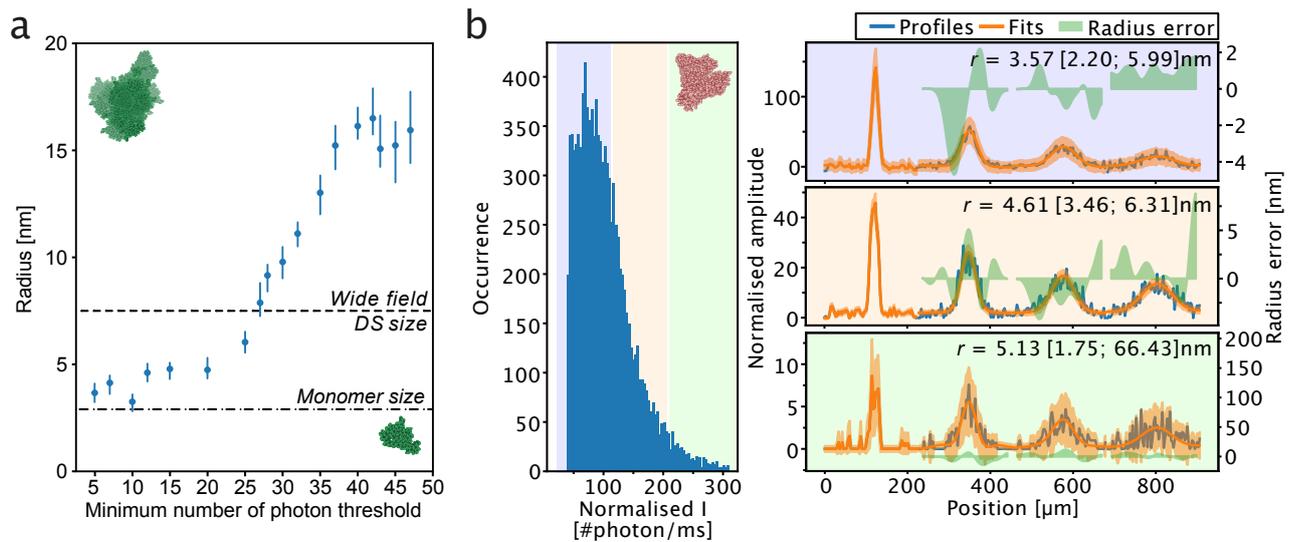


Fig. 3.8 Heterogeneous mixture sizing using single molecule microfluidic diffusional sizing of self-assembly α -syn and HSA. **a.** Size evolution of oligomer sizing, depending of the minimum number of photons threshold parameter. It allows the distinction between oligomer and a mix of oligomer/monomer with a lower threshold. Made on a solution of 50pM monomeric concentration. The results are in accordance with the range between 3 and 16 nm in *PNAS, Serene W. Chen, 2015* [32]. The error bars correspond to the standard deviations of triplicate measurements. **b.** Size of different forms of HSA (monomer, dimer, tri & quadrimer), measured using a 100 pM solution. Separation is created using different ranges of normalised intensity. The histogram of normalised intensity for all events shows a peak at 70 photons per millisecond, corresponding with the monomer form value. From this peak three different areas have been selected. Result of selected events size 3.57, 4.61 and 5.13 nm is in concordance with, respectively, monomer, dimer, and trimer expected size. The number of events is quite low for properly sized dimer; for the trimer the events have been collected over a larger range in order to have enough events to run the diffusional sizing script. Note that the error increases as well as the size. *All data points were obtained by Matthias M. Schneider and myself, and the analysis of the curve was done by myself.*

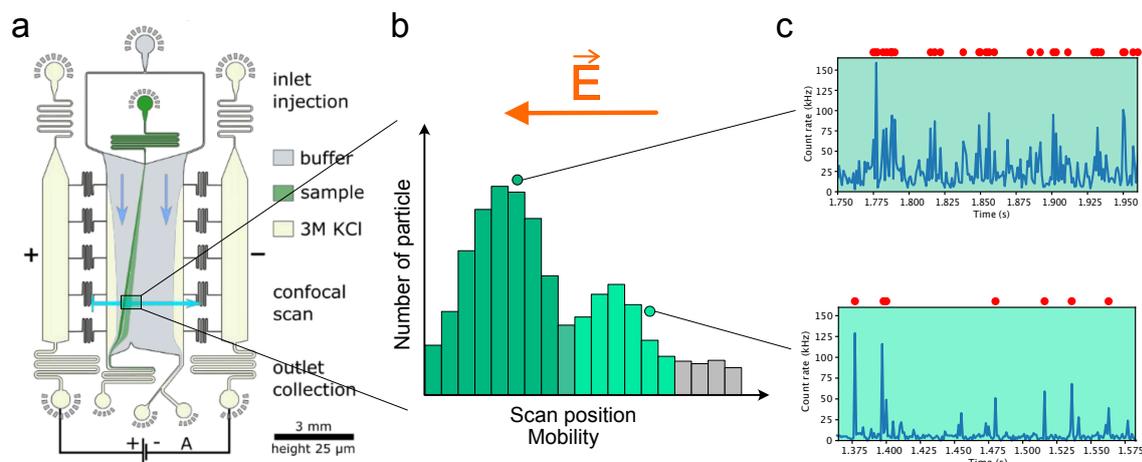


Fig. 3.9 Micro free flow electrophoresis (uFFE) chip design and flow of the process to analyses. **a.** Microfluidic design. The sample is injected in the middle of the main chamber. An electric field deviates the analytes in terms of their mobilities (green flow). The electric field is generated by a potential applied around the two electrodes composed of 3 M KCl. The confocal spot (light blue) crosses the main chamber and records the number of particles per position. **b.** Example of data acquired. The number of particles detected is plotted as a function of the position, which can be related to the mobility. **c.** Example of possible traces. The graph on top depicts a higher particle concentration than the one below. For each position a trace is recorded for a given time (typically 2-4 s). The single molecule analysis algorithm extracts the number of molecules per position.

3.3 Micro Free Flow Electrophoresis

The microfluidic free flow electrophoresis device separates the particles in terms of their mobility. A schematic of the device is displayed in Figure 3.9.a. An electric field is applied perpendicularly to the flow direction. The particles exposed to the electric field deviate their trajectory towards the electrode. Particles with a stronger mobility have a stronger deviation. The mobility is a metric depending of the ratio between the size and the charge of the particles. The electric field within the device is created by liquid electrodes [167]. It consists of two streams of 3M KCl solution flowing on both sides of the main microfluidic chamber. A potential is applied between each solution in order to create an electric field, \vec{E} . The confocal spot scans perpendicularly to the main chamber between two fiducial marks on the microfluidic device. The single molecule detection algorithm can be applied and results in a number of events versus the position. The analytes' positions across the chamber depend of their respective mobilities, see figure 3.9.b. In the case of a heterogeneous solution with different mobility values between the species, several populations can be separated. Figure 3.9.c illustrates the difference in particle density between two subpopulations.

3.3.1 Results

Two experiments are briefly explained as potential applications of this technique. The first focuses on the separation at single molecule level of protein binding reaction, rather than on the quantification of the mobility. The separation allows us to quantify different metrics like the end concentration of reactant and products, or the affinity of a reaction K_d . The second application tries to characterise the structural shape of protein self-assembly by looking at the change of apparent mobility as a function of the brightness of the aggregate.

Digital sensing of proteins in solution Two examples taken from the preprint [105] are presented in Figure 3.10 a-b, showing respectively the electropherogram of the complex biotinylated probe DNA with monovalent streptavidin, and the electropherogram of the complex aptamer with IgE. Both are binding experiments at single molecule level. The single molecule level allows us to directly count the number of complex versus individual particles. The number of molecules associated with the flow rate applied and the geometry of the chip enable us to calculate the concentration of molecules within the sample. Indeed the flux of particles detected J_p can be approximated by the total number of particles N detected over the time scan with the factor of correction of the total area scan over the total cross section of the chamber ($h \cdot L$):

$$J_p = \frac{N}{t} \frac{h \cdot L}{A_{laser} \cdot n_{step}} = \frac{N h \cdot d_{step}}{t \pi \cdot \frac{z}{2} \frac{w}{2}} \quad (3.1)$$

where A_{laser} is the projected area of the laser on the channel cross section, approximated by an ellipse shape of major and minor axes z and w , and where n_{step} is the number of steps scan within the chamber and d_{scan} the distance between two consecutive steps. From the flux J_p , the concentration of particle c_p can be estimated using the flow rate of the sample applied Q :

$$c_p = \frac{J_p}{N_A \cdot Q} = \frac{4N \cdot h \cdot d_{step}}{N_A \cdot Q \cdot \pi \cdot z \cdot w} \quad (3.2)$$

The concentration of complexes has been calculated from the total number of molecules counted in the grey region shown in Figure 3.10 a-b. The concentration is corrected by subtracting the concentration of the control experiment found in the same area (background subtraction). The concentration of the complex found is of 25.5 ± 0.8 pM from the mixture of 5 pM monovalent streptavidin with 50 pM biotinylated probe DNA, which indicates that most of the streptavidin binds to the labelled DNA. The measured complex concentration is equal to 21.7 ± 0.7 pM for the mixture of 40 nM IgE and 50 pM aptamer probe. By assuming a one to one binding, it is possible to recover the dissociation constant K_D , from the initial

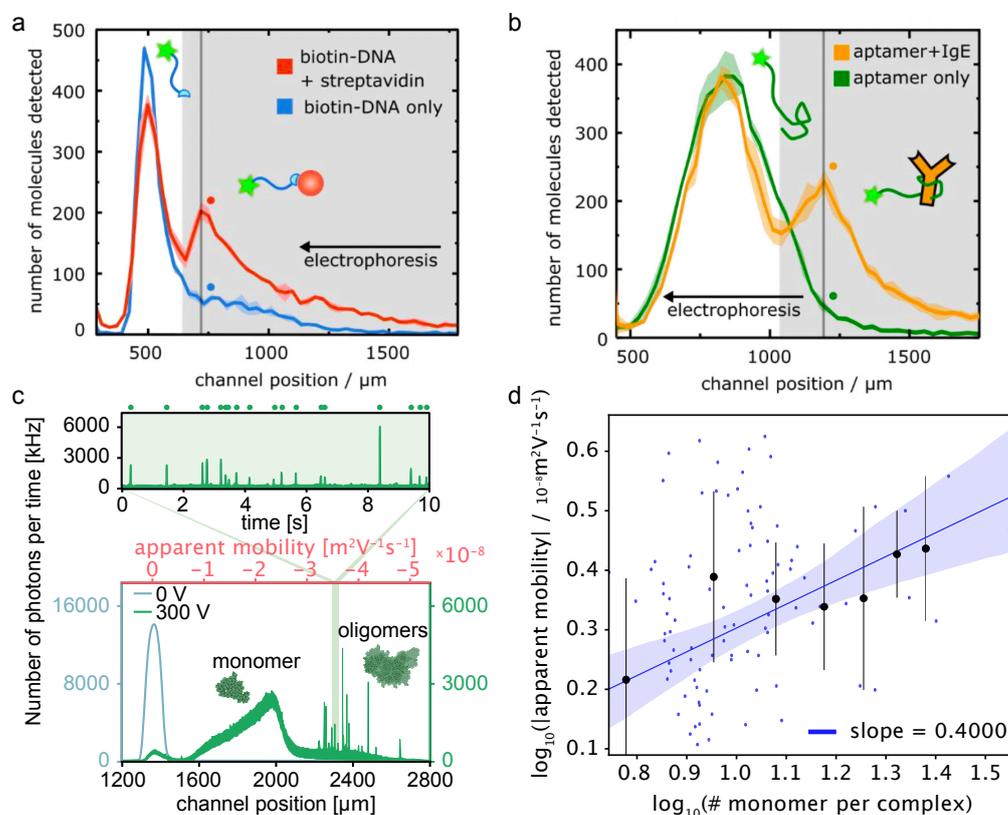


Fig. 3.10 Example of applications for single molecule analysis with uFFE chips. **a.** Electropherogram at 150 V of 50 pM biotinylated probe DNA (in blue) and a mixture of 25 pM monovalent streptavidin with 50 pM biotinylated probe DNA (in red). The region shaded in grey was used to count the number of complexes. **b.** Electropherogram at 150 V of 50 pM aptamer probe (in green), and a mixture of 40 nM IgE and 50 pM aptamer probe. The region shaded in grey was used to count the number of complexes. **c.** uFFE experiment using a solution on α -syn oligomer. In the bottom graph, the continuous scan at 0V and 300V. With the electric field a peak of α -syn spreads, at lower mobility -2.5 to -5 different oligomers at the single molecule regime can be probed. In the top graph, an example of step scan trace at the position of 2350 μm is depicted. **d.** The result of the mean size of oligomers found per apparent mobility. A slope of the loglog graph suggests a power law between the mobility of the oligomers and the number of monomers ($\mu_o \propto n_m^\nu$, $\nu = 0.4$). The error bars correspond to the standard deviations of the different points within range. Range size is kept constant. *The shaded areas around the curves (in a & b) correspond to the standard deviation of triplicate measurements. All data points were obtained by Catherine K. Xu, Georg Krainer, and William E. Arter. The analysis of the curve was performed by William E. Arter, Catherine K. Xu and myself. Panels a and b are figures reproduced from the BioRxiv paper [105].*

concentration c_{IgG} and $c_{Aptamer}$:

$$K_D = \frac{c_{IgG} \cdot c_{Aptamer}}{c_{complex}} - c_{IgG} - c_{Aptamer} + c_{complex} = 52.1 \pm 3.0 \text{ nM} \quad (3.3)$$

The resulting K_d is consistent with values from the literature of the same complex IgE/aptamer, with K_d estimated at 48 nM and 64 nM [70, 194].

Structural characterisation and detection of α -synuclein oligomers In addition to calculating the concentration of species, uFFE chips associated with confocal microscopy can probe the structure of aggregation. Looking at how mobilities change with the size of the self-assembly system gives information on the repartitions of the size and the shape of the assembly. Indeed, the mobility of the complex can be seen as a function of the charge and size of monomer unit. Depending of the shape and charge repartitions of the oligomer, the function follows different behaviour. Let us first assume the Debye layer to be much smaller than the hydrodynamic radius $R_H \gg \lambda_D$. In this approximation, the Henry equation described in equation 2.2 is simplified as [89]:

$$\mu = \frac{q}{6\pi R_H \eta} \propto \frac{q}{R_H} \quad (3.4)$$

The hydrodynamic radius of the oligomer r_c can be written as a function of the monomer radius r_m and the number of monomers n_m whether it has a globular shape ($r_o = r_m \cdot n_m^{1/3}$) or a rod-like shape ($r_o = r_m \cdot \frac{n_m}{\ln(n_m)}$) [57]. Similarly, the charge of the oligomer q_o can be written as a function of q_m and n_m , whether the charges cumulate ($q_o = n_m \cdot q_m$) or with surface charge only on a globular shape ($q_o \propto q_m n_m^{2/3}$), or with surface charge only on the rod-like shape ($q_o \propto q_m \forall n_m$). The mobility of the oligomer species as a function of the monomers in the polymer chain leads to different scaling behaviour:

$$\mu_o \propto \begin{cases} \mu_m n_m^{2/3}, & \text{for globular, without internal shielding} \\ \mu_m n_m^{1/3}, & \text{for globular, surface charge only} \\ \mu_m \ln(n_m), & \text{for rod-like, without internal shielding} \\ \mu_m \frac{\ln(n_m)}{n_m}, & \text{for rod-like, surface charge only and } n_m \gg 1 \end{cases} \quad (3.5)$$

Within the uFFE these studies have been performed with α -syn oligomer. Figure 3.10.c depicts the experiment. A solution of α -syn oligomer is injected within the uFFE chip. In Figure 3.10.c bottom graph, a continuous scan shows no deviation and no separation when the

voltage is null (light blue/green). At a voltage of 300V, a separation of different components can be seen in the green curve, with a big contribution from the monomer species down to an apparent mobility of $-2.5 \times 10^{-8} \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$. Further down in mobilities, oligomeric species become apparent as individual bursts. In this area a step scan is performed (see the graph at the top of Figure 3.10.c). The brightness gives an idea of the number of monomer units. The number of monomers per oligomer is found by dividing the average brightness of oligomer by the average brightness of the monomer. For each apparent mobility (equivalent to position), the average oligomer brightness is extracted. As explained above, this might vary significantly with the labelling efficiency of the dye/protein complex. Figure 3.10.d indicates the points of mobility versus number of monomer per complex in a log-log plot. A trend has been drawn, suggesting a power law of slope $\nu = 0.4$, but the noise is too strong to conclude a real trend. Even if the quality of the data does not help to provide a conclusion about the structure of the oligomer form, it shows an interesting application of the information which can be extracted from the combination between uFFE and the confocal microscope.

3.4 Conclusion

The combination of well-known microfluidic platforms and confocal microscopy can not only decrease the limit of detection of the system, allowing us to perform experiments at lower concentrations, but provides information which could not be achieved without the confocal. From the diffusional sizing experiment, I have shown the limit of detection of the system, the new application for drawing a binding curve with a higher affinity reaction, but I have also shown the ability to size different species within an heterogeneous sample looking at α -syn oligomer and HSA, not available with a traditional epifluorescence. From the micro free flow electrophoresis experiment, I have emphasised the possibility to digitally count a complex from a binding reaction, allowing us to determine high affinity K_d , plus the possibility to have an idea of the structure of the oligomer complex by looking at the scaling behaviour of the mixture compared to the monomer. This work is mainly a proof of concept study, which opens up different applications. Future directions include the addition of a second laser with orthogonal excitation/emission wavelengths, which allows both binding partners in a bimolecular binding event to be labelled and detected independently, thereby reducing the false positive rate. With the additional wavelength, new experiments can be developed like affinities characterisations in a competitive assay solution. The second wavelength can help to train an algorithm to better define an event on a single wavelength experiment, using a better training data set due to the improvement of event selection. Further improvements to

the system would be the creation of a UV confocal microscope, allowing us to probe label free proteins containing aromatic amino acids. Previous work has reached a sizing detection limit down to ~ 100 nM with the diffusional sizing microfluidic device [30]. By analogy with the detection limit from epifluorescence microscopy, the UV-confocal microscope could reach a precision of the order of pico molar concentrations.

Chapter 4

Nano-Cavity diffusional sizing (NDS)

This chapter is based on a manuscript in preparation:

Raphael P. B. Jacquat, Georg Krainer*, Quentin A. E. Peter*, & Tuomas P. J. Knowles.
"Single-molecule sizing through nano-fluidic confinement".*

*(*contributed equally).*

My contributions to this paper are the development of the theory, the data analysis and simulation. An approach relying on nano-cavity confinement is developed in this chapter for the sizing of nanoscale particles and single biomolecules in solution. The approach, termed nano-cavity diffusional sizing (NDS), measures particle residence times within fluidic nano-cavities to determine their hydrodynamic radii. Using theoretical modelling and simulation, it has been shown that the residence time of particles within nano-cavities above a critical timescale depends on the diffusion coefficient of the particle, which allows estimation of the particle's size. This approach experimentally has been demonstrated through measurement of particle residence times within nano-fluidic cavities using single-molecule confocal microscopy. These results show that the residence times scale linearly with the sizes of nanoscale colloids, protein aggregates and single DNA oligonucleotides. NDS thus constitutes a new single molecule optofluidic approach that allows rapid and quantitative sizing of nanoscale objects for potential application in nanobiotechnology, biophysics, and clinical diagnostics.

4.1 Introduction

Many important biomolecules, including proteins and protein assemblies as well as natural and synthetic biopolymers and colloids, have sizes in the nanometre range [77, 181]. Achieving rapid, accurate, and reliable measurement of their sizes under native solution

conditions has therefore become a key objective in many areas including nanobiotechnology, biophysics, and clinical diagnostics [51, 137, 142]. For example, the sizing of proteinaceous particles at nanometre scales is critical in studies that further our understanding of protein misfolding and aggregation processes which lie at the heart of a wide range of human diseases [9, 83, 102]. Moreover, characterising the assembly state of biomacromolecules is important when assessing, for example, biopharmaceutical product stability and efficacy of proteins or biocolloids in drug delivery systems and formulations [27, 104, 180]. Sizing techniques are therefore considered work horses in many areas of fundamental and applied science [73, 137]. Hence, the development of experimental approaches for high sensitivity detection and characterisation of nanoscale entities in the fluid phase remains an area of great current interest. Several techniques are available in order to measure the nanoscale size of proteins and nano-colloids in solution [51, 137, 142]. Most are based on determining the particle's diffusion coefficient D in solution, which is related via the Stokes-Einstein equation to the hydrodynamic radius R_H of the particle. One of the most widely used techniques is dynamic light scattering (DLS) [188]. Other widely used methods include nuclear magnetic resonance (NMR)-based techniques (e.g., pulse-field gradient NMR) [205], chromatographic techniques [25], and surface deposition microscopy, like atomic force microscopy (AFM) or scanning/transmission electron microscopy (SEM/TEM) [63, 170]. These techniques suffer from relatively high sample consumption and long acquisitions times or surface immobilisation, and often require sophisticated instrumentation. In recent years, a number of techniques have been established that operate with minimal sample requirements and sensitivities down to the single molecule regime, and operate directly in solution. These include microfluidic techniques [12, 83] such as microfluidic diffusional sizing (MDS) [9], Taylor dispersion analysis (TDA) [141], nanoparticle tracking analysis (NTA) [60] and fluorescence correlation spectroscopy (FCS) [151], or a combination of interferometric scattering (iSCAT) microscopy [215] with electrostatic trapping [139]. Such fluidic and single molecule-based approaches offer great potential for the sizing of nanoparticles and nano-colloids. However, they are often limited in the size range that can be detected and require complex models to analyse the size distributions. In this chapter, I introduce a nanofluidic approach, termed nano-cavity diffusional sizing (NDS), which allows for the robust determination of the sizes of single particles directly in solution through measurement of particle residence times within fluidic nano-cavities. The approach extracts sizes of nanoparticles through nano-confinement and allows sizing of particles in the range from a few nanometres up to hundreds of nanometres. The implementation and experimental realisation of the NDS approach is outlined in Figure 4.1. The observation volume of a confocal microscope is placed within one of the

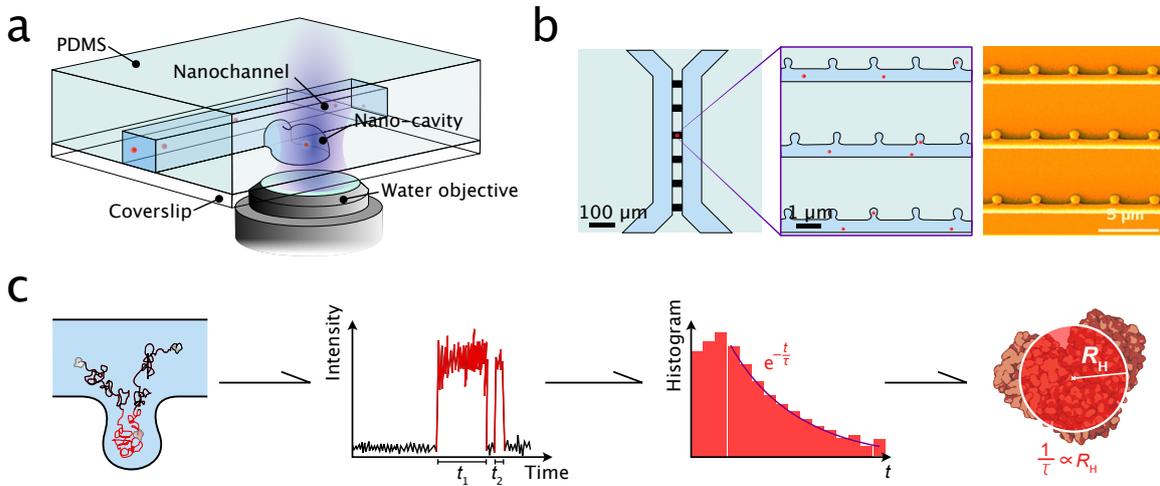


Fig. 4.1 Principle of nano-cavity diffusional sizing (NDS). **a.** Three-dimensional illustration of the experimental implementation of NDS showing the positioning of the confocal detection volume inside the nano-cavity of the nanofluidic chip. A high numerical aperture objective is used for single molecule detection of molecules within the nano-cavity of the chip, which is fabricated by hybrid lithography and moulding in PDMS. Biomolecules are depicted in red. **b** Schematic of the nanofluidic chip used for NDS measurements. The nano-cavities are located adjacent to nanofluidic channels on a microfluidic chip. A SEM image of the chip is depicted in the right panel. **c.** Workflow of the sizing experiment. First, the particles are detected by confocal microscopy as they diffuse into and out of the nano-cavity. Then, the residence times t are extracted from the recorded time trace and binned in a residence time histogram according to the occurrence of residence time. This histogram is then fit with an exponential function of the type ($e^{-t/\tau}$), from which the hydrodynamic radius (R_H) can be extracted. The coefficient τ is inversely proportional to the hydrodynamic radius.

trapping cavities of a nanofluidic device, which itself is filled with an aqueous solution that contains the biomolecule of interest (Figure 4.1.a). The chip design is shown in Figure 4.1.b. To extract the sizes of single molecules, time trajectories of particles diffusing into and out of the observation volume are recorded, and residence time distributions are extracted (Figure 4.1.c). Fitting of the obtained residence time histograms with a quantitative model provides R_H and, thus, the size of the nanoparticle of interest (Figure 4.1.c).

4.1.1 Theory and Simulation

An analytical model is first developed to examine the diffusive behaviour of particles within nano-cavities and quantitatively describe how particle size relates to residence time. This allows us to assess the scaling behaviour of particle residence times and provides a theoretical framework for the analysis of the experimental results. The diffusion of particles is modelled as Brownian motion with reflective boundary conditions at the walls. No other potential was

considered. Under the solution and geometric conditions, the direct electrostatic interactions between analyte particle and the cavity walls are negligible as the particle is typically separated by more than the screening length from the walls. The nano-cavity, as shown in Figure 4.1.a, was modelled as a cavity of rectangular shape (Figure 4.2.a), which was perpendicularly connected to adjacent nanochannels. The particle, in the model, can therefore only enter and exit the cavity by diffusing perpendicular to the nanochannel axis. Assuming that diffusion is isotropic, the model, for the case of a rectangular cavity, can therefore be reduced to a one-dimensional (1D) diffusion problem. The residence time of a particle is determined by the probability of the particle exiting the cavity over a given period of time. Representative time trajectories of a particle entering and exiting the cavity are shown in Figure 4.2.a. The particle enters the cavity at time $t = 0$ and the time points to describe particle entry and exit are denoted as t_{in} and t_{out} , which yields the residence time $t = t_{out} - t_{in}$. The length of the cavity is d_E , where d_E is the effective depth of the cavity and is given by the depth of the cavity d_C minus the particle radius R :

$$d_E = d_C - R \quad (4.1)$$

As depicted in Figure 4.2.a-b, two regimes can be observed in the residence time distribution: a short and a long timescale regime. These two regimes are separated by a critical time t_c , which corresponds to the mean time for a particle with diffusion coefficient D to reach the bottom wall of the nano-cavity according to:

$$t_c = d_E^2/2D \quad (4.2)$$

For short timescales, the particle resides within the cavity only for a short period of time such that it typically does not diffuse to the far end of the cavity (Figure 4.2.a, upper panels). The probability distribution of the particle position inside the cavity is therefore concentrated near the original position. The distribution in this regime can be described by an unconstrained random walk model and is given by the first passage time density:

$$p(t,y) = \frac{|y - y_0|}{\sqrt{(4\pi Dt^3)}} \exp(-(y - y_0)^2/4Dt) \quad (4.3)$$

where y is the coordinate position between the inside/outside of the cavity (position where $t = 0$) and y_0 is the distance that the particle diffuses inside the cavity in the y -direction. For

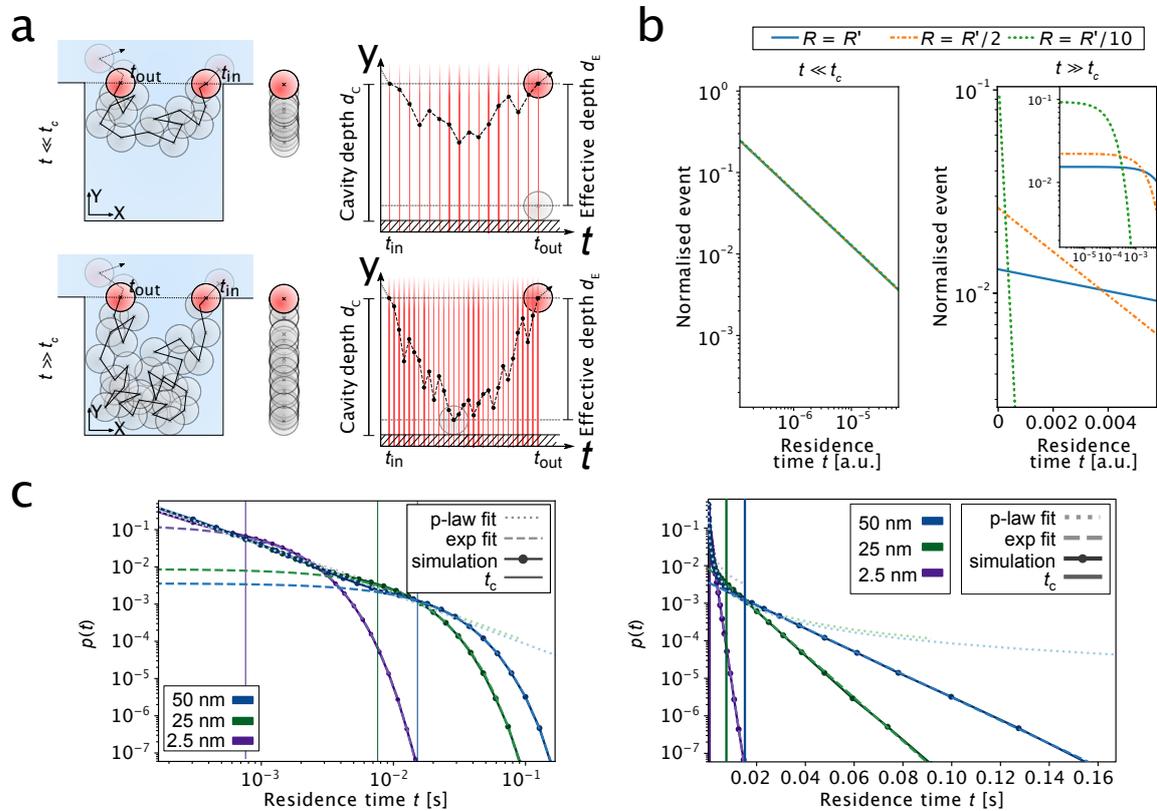


Fig. 4.2 Theory and simulation of diffusion under nano-confinement. **a.** For a particle diffusing within a nano-cavity, two diffusive scenarios can be distinguished: (i) the particle enters the cavity and exits it without reaching the bottom wall (top panels) or (ii) the particle enters the cavity and reaches the bottom of the well before exiting it (bottom panels). The right panels show displacement in time along the y axis of the diffusion processes. The particle enters at time t_{in} and exits at time t_{out} . The depth of the cavity is d_c , and d_E denotes the effective cavity depth ($d_E = d_c - R$), with R being the radius of the particle. **b.** Analytical modelling of particle diffusion at short time-scales ($t \ll t_c$, left panel) and long time-scales ($t \gg t_c$, right panel). t_c denotes the critical time that separates the two regimes and corresponds to the mean time for a particle with diffusion coefficient D to reach the end of the nano-trap (see main text). Shown are residence time probability plots. At short time-scales ($t \ll t_c$), the residence time is scale invariant and does not depend on the size of the particle. Conversely, at long time-scales ($t \gg t_c$), the residence time follows an exponential decay which is dependent on the size of the particle. Modelled were three particles with different radii at R' fixed: $R = R'$ (blue), $R = R'/2$ (orange), and $R = R'/10$ (green). Diffusion was modelled as a 1D random walk. **c.** Simulation results for the diffusion of particles within a nano-cavity. Shown are residence time probability plots (log–log plot, left panel; linear–log plot, right panel) for particles of different sizes (50 nm, blue; 25 nm, green; 2.5 nm, purple). At short residence time, particles are scale invariant, as evident by a linear behaviour in the log–log plot (i.e., power law behaviour). Conversely, at long timescales, particle residence times exhibit an exponential decay, which is dependent on the size of the particle, due to the linear behaviour in the linear–log plot. Data points represent simulation results. Long and short dashed lines depict fits of the simulation data by power law and exponential functions, respectively.

very small Δy , such that $(\Delta y^2)/4D \ll t \ll t_c$, the density can be expressed as:

$$p(t) = \frac{\Delta y}{\sqrt{(4\pi Dt^3)}} \propto t^{-\frac{3}{2}} \quad (4.4)$$

Accordingly, for the system at $\frac{(\Delta y^2)}{4D} \ll t \ll t_c$, the residence time follows a power law which is independent of D , because Δy scales with \sqrt{D} . This behaviour is displayed in Figure 4.2.b, left panel for particles with different diffusion coefficients. For long timescales, the cavity starts playing a role in the diffusion process, as the particle has enough time to explore the confined volume through diffusion. The free random walk model can therefore no longer be applied. As described in an analogous situation in [16] (Equation 5.47 therein), the residence time distribution in this regime has an exponential dependence:

$$P(t) \propto \frac{D}{d_E} \exp - \frac{t}{\tau} \quad (4.5)$$

with the decay time τ being:

$$\tau = \frac{2d_E^2}{\pi^2 D} \quad (4.6)$$

Accordingly, τ is inversely proportional to D . Hence, for the system at $t \gg t_c$, the size of the particle is linked to the decay time. This scaling behaviour is shown in Figure 4.2.b, right panel for particles with different diffusion coefficients. Residence time measurements of particles at large timescales ($t \gg t_c$) allow estimations of the size of particles. This model can therefore size particle without the requirement to have an energetic contribution to any kind of trapping free energy potential. To corroborate the results from analytical modelling, numerical simulations of a particle 1D random walk within a nano-cavity are developed to extract residence time probability distributions for differently sized particles. The simulations were performed using a reflective boundary condition for the wall on the bottom of the cavity. A Gaussian random number generator was used to simulate diffusive steps. Details of the simulations are given in the Methods section. Obtained residence time probability distributions are shown in Figure 4.2.c. The results are consistent with the theory above, in that at short timescales, the particle's residence times follow a power law behaviour ($t^{(-3/2)}$), as evident in a linear decay in the log-log plot, whereas at long timescales, the residence time decays exponentially ($e^{(-t/\tau)}$), as evident in the linear decay in the linear-log plot. Fitting of the simulation results recovered the initial input values with a relative error of 0.5%, demonstrating the robustness of this analysis approach.

4.1.2 Experimental demonstration of the NDS approach

After having explored the possibility of sizing particles in nano-cavities on a theoretical basis, I next set out to demonstrate the NDS approach experimentally. Conceptually, the experimental implementation involves the following steps. First, the duration of trapping events within nano-cavities needs to be recorded, from which probability distributions of residence times $p(t)$ are generated. Then, by fitting this distribution with an exponential function, Eqs. 4.5 and 4.6 can be used to compute the diffusion coefficient from the decay time of this exponential. Based on these considerations, I set out to experimentally demonstrate and realise the NDS approach for the sizing of single particles in solution. A nano-fluidic device, previously developed in the laboratory, helps to measure particle residence times within nano-cavities (Vanderpoorten et al., under review). A schematic of the fluidic platform is shown in Figure 4.3.a, upper panel. The device consists of arrays of nano-cavities which are connected to nanofluidic channels. These nanofluidic functionalities lie in between two microfluidic reservoirs with inlets and outlets that serve as fill ports for the sample solution. The nano-cavities are of cylindrical shape with a radius of 350 nm and a height of 650 nm. The connecting nanofluidic channels are 650 nm wide and 750 nm high. SEM images of the channel and nano-cavity geometries are shown in Figure 4.3.a, lower panels. For the detection of single particles, a confocal fluorescence microscope is used. Samples were excited with a continuous wave diode laser and their fluorescence collected using avalanche photodiodes, which allowed us to readout the fluorescent signal of molecules with high sensitivity and monitor their residence times within nano-cavities with high temporal resolution. Using this optofluidic platform, the residence times of nanoscale particles are measured for size determination by NDS. Measurements are performed on fluorescent nanoscale colloids (50 and 20 nm radius), fluorescently labelled oligomeric aggregates of the protein α -synuclein with average sizes of around 11.5 nm, and fluorescently labelled DNA oligonucleotides (45 bp, 9.5 nm). Sample solutions were injected into the fluidic device and the confocal observation volume parked in the middle of one of the nano-cavities of the device. After a short equilibration period to ensure hydrostatic balance, single molecule fluorescence of particles diffusing into and out of the well were recorded. Examples of time traces are shown in Figure 4.3.b. As anticipated from the theoretical considerations, larger particles/molecules resided longer within the nano-wells as compared to smaller ones. For each detected event, the associated residence time is extracted. Individual residence times were pooled in a histogram to obtain residence time histograms. Due to the nature of the measurement, short residence events are under-sampled, which would create an artefact in the distribution. A threshold was therefore applied in order to represent residence times only at longer timescales.

Moreover, residence times at short timescales are scale invariant; hence for size determination this regime can be omitted (see Theory and Simulations above). The experimentally obtained residence time distributions for the four tested species are shown in Figure 4.3.c. The residence time decays follow a linear behaviour in the linear–log plot, as expected for an exponential behaviour due to the biased random walk of the particles within the nano-cavity, as predicted from the theoretical modelling and simulations (see above). Accordingly, the data are fitted with an exponential function of the form $e^{(-t/\tau)}$. This allows the extraction of a decay time τ , which is proportional to the size of the particle according to the theory derived above, considering also the effective well depth. A plot of the extracted decay times versus the size (i.e., R_H) of the particles yields a linear relation (Figure 4.3.d), as anticipated from the theory. Using such calibration, it is thus possible to size particles in solution using NDS. Finally, NDS is robust against measurement noise. Single molecule experiments usually provide data at low signal-to-noise ratios. However, because size information in this approach is extracted from long timescale events, false positive events are exponentially unlikely for longer residence times, as required in this approach. In other words, the likelihood for false particle detection, which mainly happens for the detection of events on short timescales, is minimised as NDS extracts information from long timescale events. This feature of obtaining data in the high signal-to-noise regime makes NDS robust against measurement noise and thus ensures accurate and reliable measurement of a molecule's size.

4.1.3 Conclusion

In this chapter, we have established an approach for the sizing of particles using single molecule detection and nanofluidics. The NDS approach harnesses the size-dependent diffusional escape of particles under nano-confinement to obtain size information from the particle's diffusive properties. Using the theoretical modelling and simulations, it has been shown that above a critical timescale, the scaling of the particle's residence time changes from a power law, which is size independent, to an exponential, size-dependent behaviour. This realisation forms the basis of our approach and yields a linear behaviour of the size of a particle versus its residence time within a nano-cavity. Using a nanofluidic chip combined with confocal spectroscopy, I have experimentally validated both the exponential scaling behaviour for nanoscale particles and biomolecules, and shown that the decay rate follows a linear behaviour with respect to the diffusion coefficient. Using such calibration, this approach can yield rapid, accurate, and reliable sizing of particles and biomolecules. The NDS approach lines up with other techniques such as FCS and NPT analysis in terms of

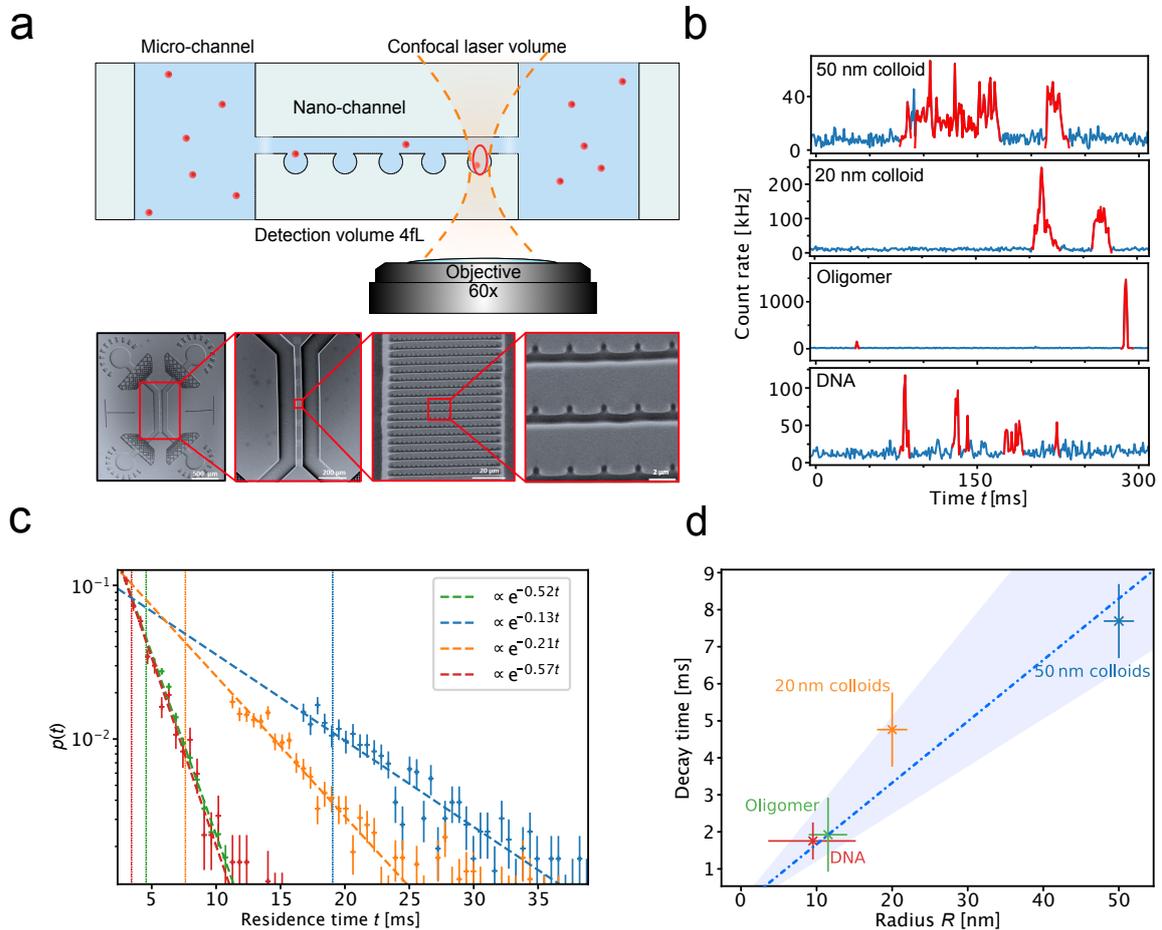


Fig. 4.3 Nanofluidic diffusional sizing of single particles in solution. **a.** Experimental setup of the NDS experiment. The observation volume of the confocal microscope is placed within a nano-cavity. Fluorescence of particles or biomolecules of interest are observed as they diffuse in and out of the confocal volume. Lower panel: SEM micrographs of the nanofluidic device with trapping functionalities used in NDS experiments. PDMS nanofluidic device imprints were fabricated via hybrid UV mask lithography and 2-photon lithography. **b.** Examples of time traces from fluorescence detection of nano-colloids, α -synuclein protein oligomers, and DNA oligonucleotides. Highlighted in red are the times when a particle was present within the confocal detection volume. **c.** Normalised residence time probability histograms of particles within the nano-cavity. The data were fit with an exponential function of the form $e^{-t/\tau}$. The slope of the curves give the decay time. The error bars correspond to the square root of occurrences found divided by the normalisation factor. **d.** Extracted decay times versus hydrodynamic radii; the dotted line corresponds to the fit on the data. The y-error bars represent the error of the slope on **c**. The x-error bars correspond to the solution accuracy.

measurement times, yet no correlation analysis is needed for size determination, and particle sizes from a few nanometres up to tens of nanometres can be determined, which is hardly achievable with other techniques. For example, NTA tracks particles only down to ca. 30 nm, while FCS is most sensitive to molecules in the low nanometre regime. By nature, NDS is a single particle counting analysis technique. Such analysis offers the advantage to size heterogeneous mixtures with components of different sizes and brightness. The implementation, as demonstrated here, uses fluorescence single molecule detection. However, other readout modalities including total internal reflection microscopy or scattering-based techniques (e.g. iSCAT) can be envisaged as well. In summary, with NDS I have presented a new single molecule optofluidic approach that allows for a rapid and quantitative sizing of nanoscale objects, which opens up potential applications in areas including nanobiotechnology, biophysics, and clinical diagnostics.

Chapter 5

Interferometric scattering correlation microscopy - iSCORR

The content of this chapter is based on a manuscript in preparation: Quentin A. E. Peter, Raphael P.B. Jacquat*, Georg Krainer* & Tuomas P. J. Knowles.*

"Interferometric scattering correlation (iSCORR) microscopy".

*(*contributed equally).*

I contributed to this manuscript by developing the theory with Quentin Peter, building the microscope with Georg Krainer, writing a compatible sdk and gui for the use of the high frame rate camera, conducting experiments, assisting with coding the analysis tools, and I performed the analysis of the iSCAT and iSCORR data. Biomolecules in their native state have been studied during the last twenty years with several different approaches, from optofluidic to nano-patterning [55, 138, 218]. The development of the interferometric scattering microscope (iSCAT) allowed investigation of a protein solution without the introduction of a label [149, 215]. In this chapter, a novel technique to characterise a protein solution is developed by investigating the time correlation of iSCAT images. It allows the determination of the diffusion coefficient of an analyte in a wide concentration range, as it can obtain results with only a single particle in the camera field of view as well as several thousands.

5.1 Introduction

The need to study detection and identification of biomolecules is ubiquitous when trying to understand life and its mechanisms, as well as for improving medicine development or sample analysis. One widely used method to detect proteins is through fluorescent labels. But even if the detection is simplified with such a tag, it needs to bind to or to engineer a

biomolecule in order to detect it, thereby modifying the particle which may affect its native and natural behaviour.

Furthermore, fluorescence emitters by nature lack a temporal and spatial resolution limit due to saturation of emission photons [182], as well as being limited in a total exposition period, resulting in photo-bleaching [46].

None of the above drawbacks exist in elastic scattering. However, the signal strength drops quickly with respect to the size because Rayleigh scattering depends on the sixth power of size [17]. The interferometric scattering (iSCAT) technique allows us to decrease the dependency on the size to a power of three due to the use of a carrier wave [149]. This additional carrier wave leads to possible detection of small biomolecules on surfaces like DNA down to 100 bp [119]. In most of the cases, the carrier wave is a reflected wave on the interface with the solution. In this case, the fraction of the light reflected is fixed by the refractive index. The excitation wave must be strong in order to have a sufficient scattering signal. However, the reflected signal can be too strong and quickly saturate the camera. To solve the problem the carrier wave must be attenuated without affecting the scattering signal. Several papers describe techniques to reduce this saturation by adding a mask at the Fourier plane where the scattering signal is separated from reflected light, or by using a high speed camera allowing us to increase the effective well depth over the same time [38, 120, 187]. All of these techniques, however, reduce the number of photons and do not increase the signal-to-noise ratio.

With the exception of a highly scattering tag, such as a gold nanoparticle, the detection of proteins with iSCAT is often limited to particles which can be immobilised on the surface [119, 131, 185, 215]. Indeed, the phase varies rapidly when a particle diffuses in solution, a change of a quarter wavelength in height resulting in transforming the pattern from in-phase to out-of-phase. In contrast, an event where particles bind to the surface can be integrated for a long time before and after the event in order to increase the signal [156]. In this chapter, interferometric scattering correlation (iSCORR) microscopy is presented. This method, developed by Quentin Peter and myself, combines image correlation analysis of single particles with the improved signal of interferometric scattering microscopy (iSCAT) to enable label-free sizing of particles in solution. It utilises a high frame rate camera in order to capture the change in the interference pattern. Indeed, in order to capture the interference change of phase the image rate should be faster than the time for a particle to diffuse a quarter of a wavelength. The signal received can be enhanced by a combination of several particles in the field of view, allowing the detection of particles which scatter less. This chapter develops

the theory, then an overview of usual iSCAT data are presented before the first measurement made with the iSCORR method is shown.

5.2 Theory

5.2.1 iSCAT

The scattering from particles smaller than the light wavelength (λ) is described by Rayleigh scattering. The intensity at a distance R and angle θ of a particle of diameter d is proportional to the incident intensity I_0 :

$$I_s = I_0 \frac{1 + \cos^2 \theta}{2R^2} \left(\frac{2n_m}{\lambda} \right)^4 \left(\frac{n_p^2 - n_m^2}{n_p^2 + 2n_m^2} \right)^2 \left(\frac{d}{2} \right)^6 \quad (5.1)$$

where n_m is the refractive index of the surrounding medium and n_p is the refractive index of the particle. Because of the d^6 dependency on the particle size, the signal is very hard to separate from background light. The weak signal can be better detected by adding a coherent carrier wave and by measuring the interferometric signal of the combination of the weak scatter wave and carrier wave. This is the principle of interferometric scattering (iSCAT). In the most common approach, the carrier wave is created by reflection of the light at a glass-water interface, where the reflection coefficient is denoted by r . The light which passes within the medium can be scattered by particles inside, creating a scattering contribution s (see Figure 5.2.a). The total intensity measured is affected by both waves interfering:

$$I = I_0(|r|^2 + 2|r||s| \cos \Phi + |s|^2) \quad (5.2)$$

where I_0 is the excitation intensity and Φ the interference phase. For small d , the contribution of the scattering only ($|s|^2$) can be assumed to be negligible ($|s|^2 \ll |r|^2$). The shot-noise variance of the measurement (σ_I^2) is defined as proportional to the total intensity I :

$$\sigma_I^2 = I\sigma^2 \quad (5.3)$$

where σ^2 is the normalised variance. The variance σ_I^2 can be approximated as:

$$\sigma_I^2 \approx I_0|r|^2\sigma^2 \quad (5.4)$$

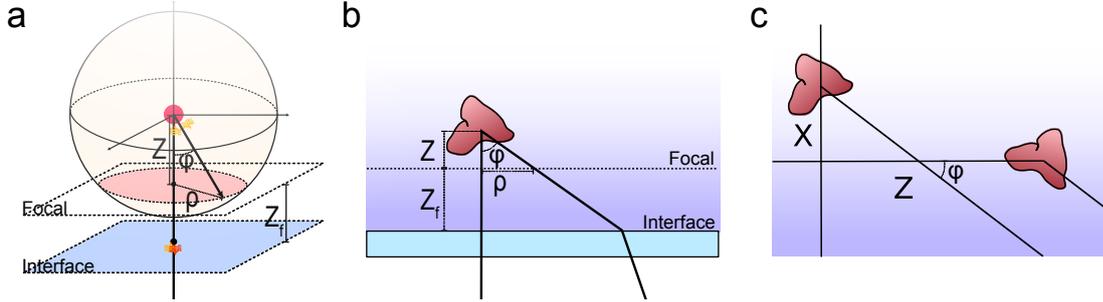


Fig. 5.1 Path of the light to focal plane. **a.** 3D view of the optical path from the change of glass-water interface to the focal plane. **b.** Same view in 2D. **c.** Optical path in diffusional light scattering (DLS) system.

The path difference between the reflected light and the scattered light to the camera provides the phase (Φ). The image of the camera being the image at the focal plane, the phase is hence determined by the path difference up to the focal plane. It is possible to calculate each path knowing three different parameters (see Figure 5.1): the particle distance to the focal plane (Z), the distance of the focal plane from the interface (Z_f), and the horizontal distance of the particle (ρ). The phase is hence:

$$\Phi = \frac{2\pi}{\lambda_m} \left(\sqrt{\rho^2 + Z^2} + Z + 2Z_f \right) \quad (5.5)$$

where λ_m is the laser wavelength in the medium ($\lambda_m = \lambda / n_m$). Note the small dependence on Z for particles below the focal length ($Z < 0$). In the case of no deviation of the light from the optical axis after scattering ($\varphi = 0$), the path difference is zero. For particles above the focal plane ($Z > 0$), the change in Z strongly influences the pattern intensity, for example by passing from constructive to destructive interference for a quarter wavelength displacement in Z (for $\varphi = 0$).

Tracking a dim pattern in solution is difficult, especially when it changes from in-phase to out-of-phase. Integration of frames can reduce the total information given per frame. One way to be able to combine this information over frames constructively is to prevent the particle from moving. In *Young et al.* [215], the authors employed this approach for detecting BSA using adsorption on a glass slide. Another way to conserve information is by correlation analysis, i.e. by combining frames, each of them containing a small amount of information.

5.2.2 Correlation function and iSCORR

The fluctuation of the pattern intensity can be characterised by a temporal autocorrelation function (g_I^2):

$$g_I^2(\Delta t) \equiv \frac{\langle (I(t) - I(t + \Delta t))^2 \rangle_T}{\langle I(t) \rangle_T} \quad (5.6)$$

Under the condition $|s|^2 \ll |r|^2$, which corresponds to a small object with low scattering, $|s|^2$ can be neglected and the total intensity can be approximated as:

$$I(t) = I_0(t) (|r(t)|^2 + 2|r(t)||s(t)|\cos\Phi(t)) \quad (5.7)$$

When assuming constant laser intensity ($I_0(t) = I_0$) and constant reflection on the cover slip ($|r(t)| = |r|$), as well as assuming that the phase variation has a stronger contribution than the variation in the scattering ($|\cos(\Phi(t)) - \cos(\Phi(t + \Delta t))| \gg ||s(t)| - |s(t + \Delta t)||$ for small Δt), the normalised temporal autocorrelation function defined in equation 5.6 can be written as:

$$\begin{aligned} g_I^2 &= \frac{4I_0^2|r|^2|s|^2\langle (\cos(\Phi(t)) - \cos(\Phi(t + \Delta t)))^2 \rangle_T + 2\sigma_I^2}{I_0|r|^2} \\ &= 4I_0|s|^2\langle (\cos(\Phi(t)) - \cos(\Phi(t + \Delta t)))^2 \rangle_T + 2\sigma^2 \end{aligned} \quad (5.8)$$

where the square of the noise averages to the variance σ_I^2 and not to zero, and the factor two comes from the fact that two time points are used. The standard deviation of the correlation function is given by the sampling error when averaging over n frames [1]:

$$\sigma_{g_I^2} = 2\sqrt{\frac{2}{n}}\sigma^2 \quad (5.9)$$

Up to this point no assumption has been made about the movement of the particles within the field of view and the phase can be calculated for different scenarios. First, a stagnant system (i.e. no flow) is considered. In the case of a random walk, the mean displacement of a particle in any direction x after a time Δt will be $\langle (\Delta x)^2 \rangle_T = 2D\Delta t$, with D the diffusion coefficient of the particle. The phase difference $\Delta\Phi$ can be written as a contribution of the

two directions ρ and Z , see Figure 5.1.a-b:

$$\begin{aligned}
\Delta\Phi &= \frac{2\pi}{\lambda_m} \left(\Delta Z + \sqrt{(\rho + \Delta\rho)^2 + (Z + \Delta Z)^2} - \sqrt{\rho^2 + Z^2} \right) \\
&\approx \frac{2\pi}{\lambda_m} \left(\Delta Z \left(1 + \frac{Z}{\sqrt{\rho^2 + Z^2}} \right) + \Delta\rho \frac{\rho}{\sqrt{\rho^2 + Z^2}} \right) \\
&\approx \frac{2\pi}{\lambda_m} (\Delta Z a + \Delta\rho b)
\end{aligned} \tag{5.10}$$

where a and b are variables introduced to simplify the notation. Note that $\Delta\Phi$ is a Gaussian random variable as it is the sum of the Gaussian random variables ΔZ and $\Delta\rho$. Furthermore, ΔZ and $\Delta\rho$ are uncorrelated to each other, which allows reduction of the cross correlation term:

$$\begin{aligned}
\sigma_\Phi^2 &\equiv \langle (\Delta\Phi)^2 \rangle_T \approx \frac{4\pi^2}{\lambda_m^2} \langle (\Delta Z)^2 a^2 + 2\Delta Z \Delta\rho ab + \Delta\rho^2 b^2 \rangle_T \\
&\approx \frac{4\pi^2}{\lambda_m^2} (\langle (\Delta Z)^2 \rangle_T a^2 + \langle \Delta\rho^2 \rangle_T b^2) \\
&\approx \frac{4\pi^2}{\lambda_m^2} 2D\Delta t \left(1 + \frac{2Z}{\sqrt{\rho^2 + Z^2}} + \frac{\rho^2 + Z^2}{\rho^2 + Z^2} \right) \\
&\approx \frac{4\pi^2}{\lambda_m^2} 4D\Delta t \left(1 + \frac{Z}{\sqrt{\rho^2 + Z^2}} \right)
\end{aligned} \tag{5.11}$$

From equation 5.8, the temporal correlation of the square $\Delta\cos\Phi$ is obtained using the identities $\langle \sin^2 \Phi(t) \rangle_T = \frac{1}{2}$, $\langle \cos \Phi(t) \rangle_T = 0$ and $2\cos a \cos b = \cos(a-b)\cos(a+b)$:

$$\langle (\cos \Phi(t) - \cos(\Phi(t + \Delta t)))^2 \rangle_T = 1 - \langle \cos \Delta\Phi \rangle_T \tag{5.12}$$

It is then possible to solve the cosine contribution using a power series:

$$\begin{aligned}
\langle \cos \Delta\Phi \rangle_T &= \sum_{n=0}^{\infty} \frac{(-1)^n \langle \Delta\Phi^{2n} \rangle_T}{(2n)!} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n \sigma_\Phi^{2n} (2n)!}{(2n)! (2^n n!)} \\
&= \exp\left(\frac{-\sigma_\Phi^2}{2}\right)
\end{aligned} \tag{5.13}$$

where in the second line the variance from a power of a Gaussian random variable is utilised. Combining equations 5.8 and 5.13, the correlation function can be written as:

$$g_I^2 = 4I_0|s|^2 \left(1 - \exp\left(-\frac{t}{\tau}\right)\right) + 2\sigma^2 \quad (5.14)$$

where the correlation time τ is defined as:

$$\tau = \frac{1}{D} \left(\frac{\lambda_m}{2\pi}\right)^2 \frac{1}{4 \left(1 + \frac{Z}{\sqrt{\rho^2 + Z^2}}\right)} \quad (5.15)$$

Flow measurement In addition to the measurement of pure diffusion, the iSCORR method allows use on a sample under constant flow. Indeed, using the property of the Fourier space that a translation is equivalent to a phase difference [162], it is possible to pass into the Fourier space in order to detect and to remove translation from the diffusion coefficient. *The decorrelation of the flow has been developed by Quentin A. E. Peter only and is further detailed in his thesis [153]. I would nevertheless like to mention this piece of work as it creates many more applications for iSCORR.*

5.2.3 Dynamic light scattering

Dynamic light scattering (DLS) is a method used to size particles by measuring the correlation time of scattering interference. While similar to iSCORR, the end result is not the same. In this subsection, the correlation function of DLS is described to show the originality of iSCORR. Assuming two diffusing particles with similar scattering cross sections ($|s_1| = |s_2| \equiv |s|$), the scattered signal is given by:

$$I = I_0(|s_1|^2 + 2|s_1||s_2|\cos\Phi + |s_2|^2) = I_0|s|^2(1 + \cos\Phi) \quad (5.16)$$

The change of phase Φ is extracted from the path difference as illustrated in Figure 5.1, and is given by:

$$\Delta\Phi = \frac{2\pi}{\lambda_m} (\Delta X \sin\varphi + \Delta Z(\cos\varphi - 1)) \quad (5.17)$$

And the variance of the phase in the case of a pure random walk is calculated as:

$$\sigma_\Phi \equiv \langle \Delta\Phi^2 \rangle_T = \frac{2\pi}{\lambda_m} 8D\Delta t \sin^2\left(\frac{\varphi}{2}\right) \quad (5.18)$$

Equations 5.12 and 5.13 remain true for DLS, and the correlation function is similar:

$$\begin{aligned}
 g_{DLS}^2 &\equiv \frac{\langle \Delta I^2 \rangle_T}{\langle I \rangle_T^2} = \langle (\Delta \cos \Phi)^2 \rangle_T + \frac{\sigma^2}{2I_0 |s|^2} \\
 &= \left(1 - \exp\left(-\frac{\sigma_\Phi^2}{2}\right) \right) + \frac{\sigma^2}{2I_0 |s|^2} \\
 &= \left(1 - \exp\left(-\frac{t}{\tau}\right) \right) + \frac{\sigma^2}{2I_0 |s|^2}
 \end{aligned} \tag{5.19}$$

with the correlation time defined as

$$\tau = \frac{1}{D} \left(\frac{\lambda_m}{2\pi} \right)^2 \frac{1}{8 \sin^2\left(\frac{\varphi}{2}\right)} \tag{5.20}$$

5.2.4 Contribution of polarisation

Light arriving at an angle to an interface is transmitted differently depending on whether it is polarised perpendicularly or parallel to the plane of incidence (s-polarised or p-polarised, respectively) [146]. The transmission of p-polarised light is decreased due to the angle φ with the interface ($\cos(\varphi)^2$) and all light, independent of polarisation, arrives at an angle at the detection device adding another $\cos(\varphi)$ contribution. Furthermore, the light passing through the interface follows the Fresnel equation due to the change of medium.

$$|s|^2 \propto (|t_s(\varphi)|^2 + (|t_p(\varphi)| \cos(\varphi))^2) \frac{\cos(\varphi)}{R^2(\varphi)} \tag{5.21}$$

5.3 Methods

5.3.1 The iSCORR work flow

Figure 5.2.d shows the work flow used to size a solution with the iSCORR method. After collecting a range of images, an optional filter is applied by passing through the Fourier space. It allows us to calculate and remove the flow contribution but also reduces the noise outside the frequency range of interest. The next step is the fit of the correlation function. The fit not only gives information about the noise level $\frac{\sigma^2}{2I_0}$ (labelled N in the figure), it also yields the size of the scattering object due to the correlation time τ , which is proportional to the hydrodynamic radius R_H .

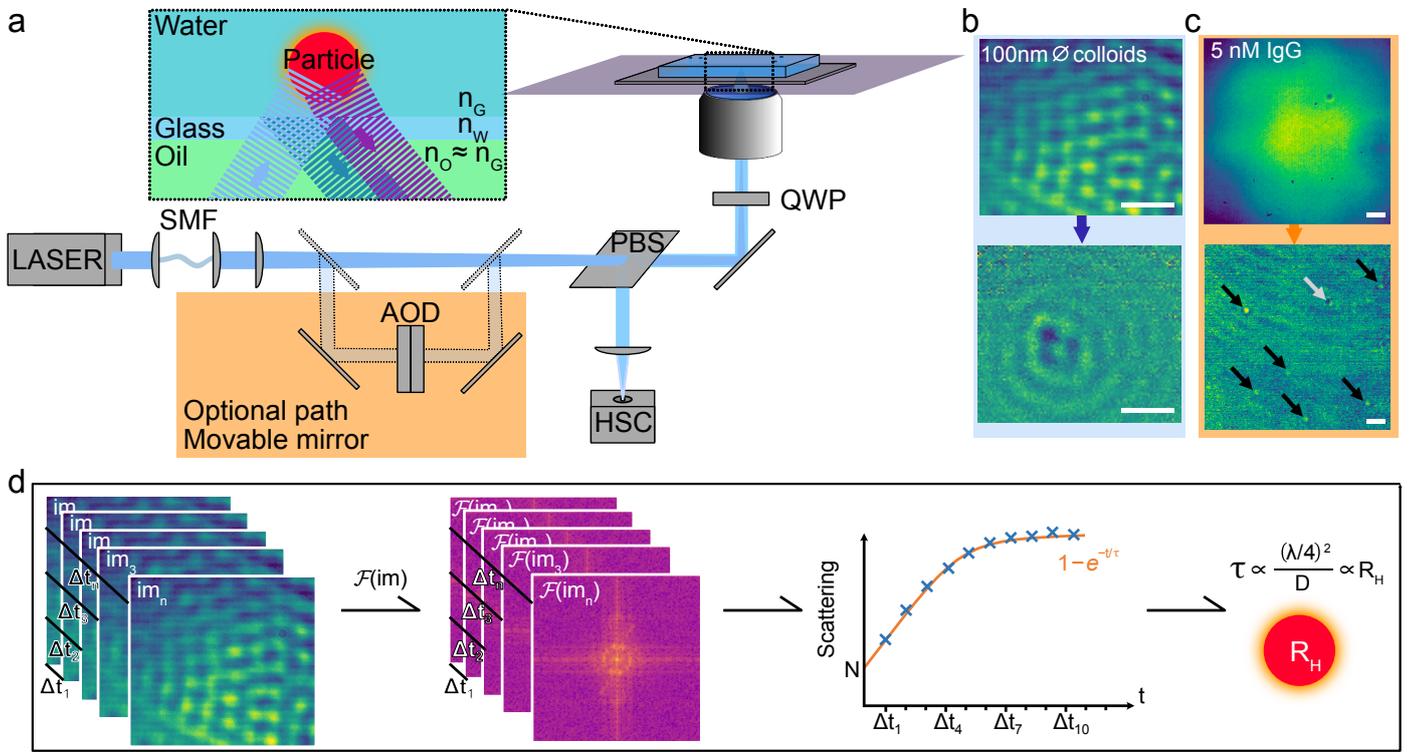


Fig. 5.2 iSCAT and iSCORR microscope and work flow **a.** Microscope representation of the optical path. The laser beam passes first through a single-mode fibre (SMF) and is collimated. The laser is then focused by a lens to the back focal plane of the objective. The light passes in between through a polarising beam splitter (PBS) and a quarter wave plate (QWP) that transforms the linearly polarised light into circularly polarised light, allowing modulation of the intensity of the reflective wave front that the camera receives. The light arrives focused on the back focal plane of the objective and leaves collimated toward the sample. The only change of refractive index is between the glass and the medium, and the medium with a particle in the light path. Due to the interface, both reflected waves return to the objective and are reflected to the PBS before being focused on the high speed camera (HSC). An optional path can be added by adding two mirrors, and the light path then includes an acousto-optic deflector (AOD). The AOD creates a scan over the field of view in x and y, resulting in a flat illumination. Both light paths, the optional and normal one, are designed to have the same length, which is achieved with additional mirrors not included in the drawing for simplicity. **b.** Image of 50 nm radius polystyrene colloids which scatter the light, image prior to processing on top and after processing below. Processing is performed by taking the difference of the two integrated frames. **c.** Image of landing events of IgG from the optional path, image prior to processing on top and after processing below. Landing events are highlighted with dark arrows, an unknown particle/dust is moving out (highlighted with the light arrow). **d.** Work flow of the iSCORR process: images are acquired at a fixed frame rate. An option is to compute the 2D Fourier transform for each frame allowing different subsequent analysis (like removing flow rate contributions). A correlation fit is created for the scattering of each difference image using basis functions. The hydrodynamic radius (R_H) can be calculated from the fit parameter τ , which is proportional to R_H .

5.3.2 Setup

The setup is based on previous work on iSCAT firstly made by Sandoghdar, and with some additions of Kukura [11, 69]. It has been developed in order to be able to work in two different modes (see Figure 5.2.a), one which has a restricted field of view due to non-uniform illumination, and a second mode (optional path) which has a slower image acquisition time but includes better illumination using acousto-optic deflectors (AODs). A laser beam with a wavelength of 445 nm is focused on the back focal plane of an oil immersion objective, in order to illuminate a solution with a collimated beam. At the cover slip interface with the water, the change of refractive index reflects part of the beam. The beam passing through the solution scatters on the suspended particle, a part of the scattered light returns to the objective and interferes with the reflected light, and the combined wave is captured by a fast camera. In order to have a stronger scattering signal, the illumination intensity should be maximal. Nevertheless the camera is often saturated. Two systems are used to capture as much light as possible without saturating the camera. The first uses a high frame rate camera (FASTCAM Mini UX100 type 800K-M-16G), allowing us to double the well depth over a time t by doubling the frame rate. The down side is often a reduction of the region of interest (RoI) with the increase of the image acquisition frequency. A second method to reduce the saturation is the addition of a quarter wave plate positioned below the objective. It allows us to tune the signal by rotating the plate, such that the camera is closest to its maximal capacity for a given frame rate. In addition to the intensity, the coherence of the light should be larger than the path difference between reflected surface and scattered particle. A too large coherence will increase the risk of interference between the different optical parts. A laser diode with limited coherence length is therefore used in the system. Although scattering intensity increases with lower wavelength to a power of 4 (see the Rayleigh scattering equation 5.1), decreasing the wavelength does not necessarily improve the signal due to the efficiency of the different optical components. Even if UV transparent components are available, the camera efficiency is decreased; besides, UV illumination may damage biological samples. These conditions motivated the choice of a light source at a wavelength of 445 nm.

5.3.3 Limit of detection

Using the highest possible frame rate, except for limiting the saturation, does not necessarily add information for iSCORR. As shown in equation 5.9, the noise of the signal depends on the square root of the number of frames, and linearly on the exposition time. Therefore,

the noise can be decreased by combining the frame prior to performing the correlation. Combining frames over a time longer than the correlation time τ will delete information about the system. Hence, the signal-to-noise ratio reaches a maximum for a single frame exposure time at half the correlation time. With this information in mind, the minimum frame rate needed to size a particle can be calculated. For example, sizing BSA with a hydrodynamic radius of approximately 3.5 nm which corresponds to a diffusion coefficient $D = 7 \cdot 10^{-11} \text{ m}^2 \text{ s}^{-1}$, the integration time should be of the order of 20 μs at most, which corresponds to a frame rate of 50 kHz. In addition to this, some high speed cameras have two systems to read out the frames in alternation, which means that the even and odd numbered frames are slightly different. These frames need to be combined, which transforms the frame rate needed to 100 kHz, which is above the 80 kHz maximal frame rate of the camera. Even if the theoretical maximum signal-over-noise ratio cannot be achieved, the measurement and sizing can still be accomplished with a lower frame rate.

5.3.4 Data analysis

The fit of the correlation data from an iSCAT measurement depends of the two parameters ρ and Z . This information can be found only if the position of the object is known. For one object which scatters in the field of view at a given time t , ρ lies between zero and the length of field of view, however this is not the case for Z . On the other hand, as the particle moves and is recorded on several frames, it can be assumed that the particle covers the entire Z possible above the focal plane. The experimental curve should hence correspond to an integral over all Z positions and ρ of the correlation function.

$$G^2 = 2\sigma^2 + 4I_0 \int_{Z_f}^{\infty} dZ \int_0^{\rho_{max}} \rho d\rho |s(Z, \rho)|^2 \left(1 - \exp\left(-\frac{\Delta t}{\tau(Z, \rho)}\right) \right) \quad (5.22)$$

If the particle is located below the focal plane, it only makes a weak contribution to the light path difference and therefore does not have to be taken into account. Furthermore, in practice the operator usually focuses slightly above the interfaces in order to reduce the number of particles below the focal plane, as well as reducing the contribution from particles which are adsorbed on the surface. Similar to the diffusion profile fit algorithm used in the chapter above, the algorithm to fit the size in iSCORR is based on deconvolving the correlation curve into a linear combination of correlation functions calculated beforehand for particles with known diffusion coefficients (basis functions). A least-squares error algorithm is then used to find the linear combination yielding the lowest residuals, allowing determination

of the diffusion coefficient of the particles within the field of view. From the diffusion coefficient the apparent hydrodynamic radius (R_H) of the particles can be calculated using the Stokes-Einstein equation.

5.4 Results and discussion

5.4.1 iSCAT

In order to ensure the quality of the custom-made microscope, state-of-the-art experiments have been replicated. The landing events of several particles have been recorded by focusing directly on the surface. *Young et al.* [215] have shown that the landing event contrast is directly linked to the molecular mass of the particle with such precision that the different oligomeric states of BSA (monomer, dimer, trimer) could be measured. Examples of measured landing events from high to low contrast are shown in Figure 5.3.a. The automatised detection of landing events yields probably a high number of false negative events at low contrast. Indeed, the replication of the observation of BSA landing with similar concentration (10 nM) and record time compared to [215] failed to properly resolve the different values of monomer, dimer, and trimer (see histogram in Figure 5.3.b). The contrast for different numbers of events is reported in the graph below the histogram. Nevertheless, the detection of BSA protein landing confirms that the scope is working close to the optimal condition. In addition to the BSA experiment, the contrast of three other proteins has been extracted (Enolase, α -syn oligomer and IgG), and is depicted in Figure 5.3.c (with respective mean contrast values of $3.1 \cdot 10^{-3}$, $6.7 \cdot 10^{-3}$ and $8.1 \cdot 10^{-3}$). Even if the molecular weight of α -syn is not well defined, it should be larger than that of the IgG (≈ 145 nM). The contrast of α -syn spreads more than the other proteins, probably due to the heterogeneity of the complex within the solution.

5.4.2 iSCORR

Recording of the landing events has the disadvantage of localising the adsorption in a first step and only subsequently obtaining information of protein adsorbed on a surface. In the case of the iSCORR analysis technique, the information on the localisation of the particle is not considered, thus reducing the filtering of the data. Furthermore, the sample is characterised in solution, which means the environment is closer to the native state. In order to reduce the influence of the glass/water interface on the correlation measurement, the light is focused slightly above the glass on the order of 10 μm .

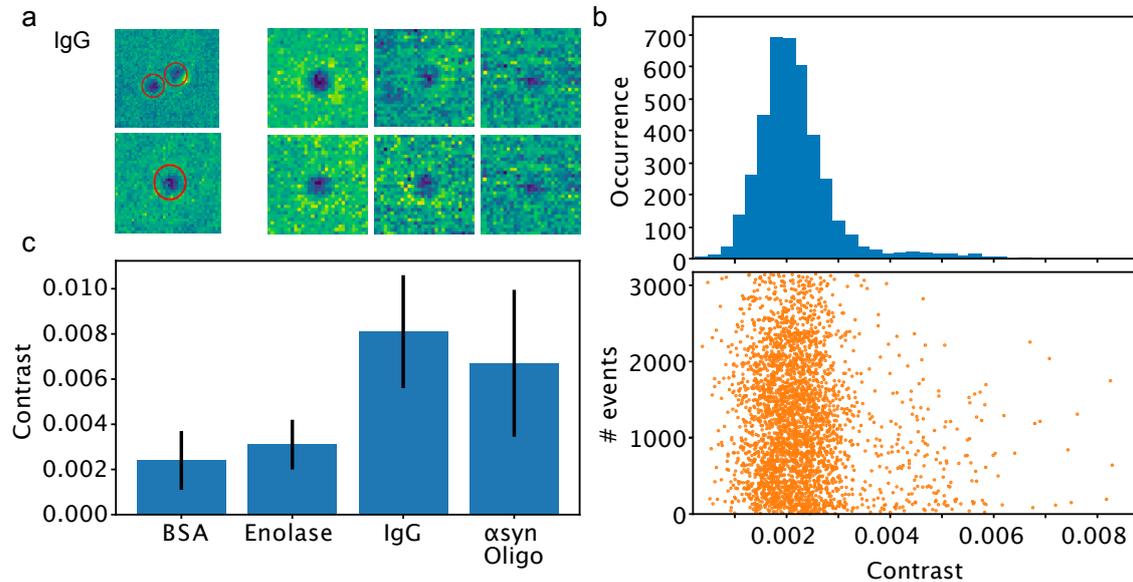


Fig. 5.3 Results of iSCAT images. **a.** Picture of landing events detected from an IgG solution, using the iSCAT microscope. The first column shows two frames with the red circle highlighting automatic event detection. The second to fourth columns show a zoom on events from higher contrast to lower. **b.** Histogram of the contrast found for each of the BSA landing events detected. The report of the contrast of each individual event is shown below. **c.** Contrast found for different proteins and protein complexes. A heavier protein means a stronger contrast.

An example of different polystyrene colloid sizes (10, 20, 50 nm radius) and BSA at two different concentrations is shown in Figure 5.4.a. The iSCAT image and the image from the difference of two frames for 50 nm colloid and BSA are shown, respectively, in Figure 5.4.b and in Figure 5.4.c. The given scattering signal and its associated correlation time τ and size r can be seen in the graphs below the images. The error shown in panel a is obtained through the standard deviation of repeated experiments, and not directly from the fitting error. Proteins in solution are sized using the diffusion coefficients found, and assuming a viscosity of water at 20 $\hat{\text{A}}$ C. The sizing of both concentrations of BSA (50 mgmL $^{-1}$ and 133 mgmL $^{-1}$) then lead to a larger apparent size as the viscosity at this concentration influences the diffusion of particles. Similar work by *Arosio et al.* [9] allows us to compare the apparent size of BSA at different concentrations with a diffusional sizing technique. Both BSA solutions nevertheless appear larger than the sizes reported in the literature. This might be due to a stronger signal from the heavier protein complex, and therefore biases the sizing towards the heavier species. Even if in theory iSCORR is able to size a single particle in the field of view, the signal is often too weak and several particles need to be present in the image to be able to estimate their size. Nevertheless, it measures the correlation of the interference between particle and the surface rather than between different particles as in

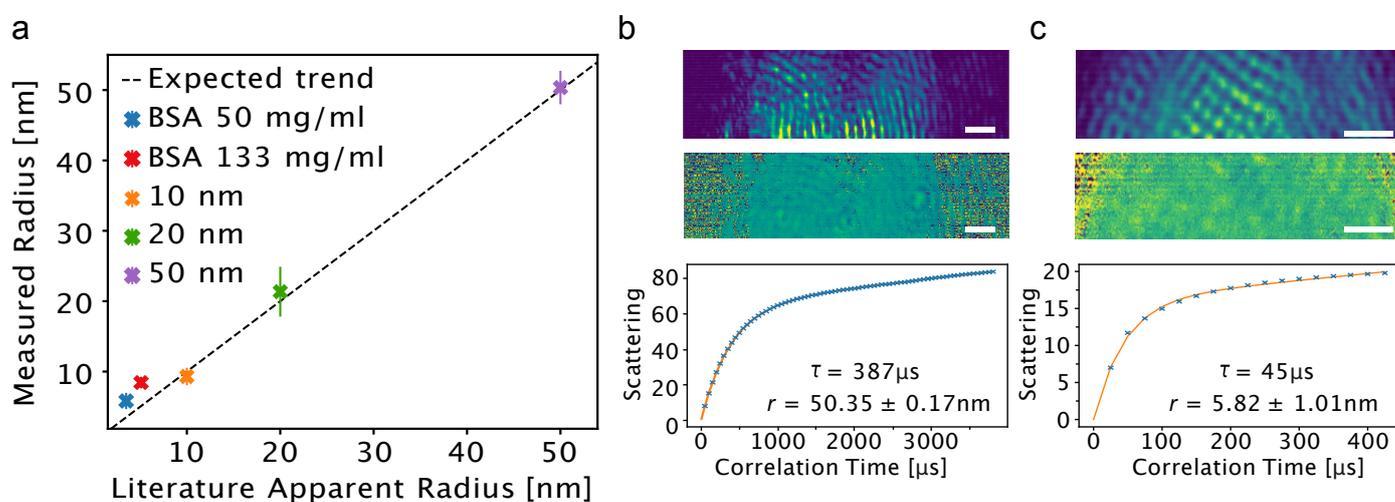


Fig. 5.4 Size analysis results of iSCORR. a. Measured hydrodynamic radius versus expected hydrodynamic radius for BSA and different colloidal particles. The literature radius represents the apparent size of BSA [9] if the viscosity would have been that of water. The error bars correspond to the standard deviations of triplicate measurements. **b.** Raw image and processed image used for calculating the scattering signal of 50 nm polystyrene colloidal particles, three of which are in field of view. At the bottom is the scattering signal, with its associated fit. The correlation time τ equal 387 μs , and corresponding size r is 50.35 nm **c.** Raw image and processed image for calculating the scattering signal of 50 mg ml^{-1} of BSA protein. At the bottom is the scattering signal, with its associated fit. The correlation time τ equal 45 μs , and the corresponding size r is 50.35 nm.

the case of DLS. Figure 5.4.b illustrates the scattering of three particles with the surface; the scattering between each other is negligible. The fact that the noise is decoupled from the scattering signal from the correlation function g_I^2 allows us to estimate how cleanly the signal was recorded. It also allows us to compare the total scattering signal between different setups. The total scattering signal should be able to predict the concentration of the protein, once it is associated with the corresponding size. However, this assumes that the scattering of different proteins is similar for proteins of similar mass, and could not be used with different types of materials. Indeed, gold nanoparticles at similar concentration would scatter more light than organic components such as proteins.

5.5 Conclusion

The iSCORR technique is a label free method that can be implemented with any existing iSCAT microscope. It enables us to detect and to size particles in solution. While iSCAT can record small proteins only by detecting landing events, iSCORR manages to detect them in solution. When iSCAT investigates a discontinuity event between the states "bound to the surface" and "unbound", iSCORR is continuously separating the scattering signal from the background while the protein is diffusing. Therefore the iSCORR signal is weaker and harder to measure. Nevertheless, the data presented in this chapter demonstrates its ability to detect proteins as small as 3.5 nm. Further work needs to be carried out to push the detection limit down to lower concentrations, and to allow more accurate sizing. The downside of this tool is the inability to distinguish different particles within the field of view, and thus a bias of the average size in heterogeneous mixtures toward the heaviest species is observed. But contrary to the NDS technique described in the previous chapter, it does not require a collection of single molecule events to size particles. A possibility to resolve heterogeneous solutions with iSCORR is to enforce only one particle at a time in the field of view, and this can be achieved by working in a low concentration regime. Development of the theory can be enhanced by adding a rotational diffusion coefficient for anisotropic objects in which the variation of the scattering may have a stronger contribution than phase variation [20]. Furthermore, the ability mentioned in this chapter to not only measure the flow, but to remove its contribution to size the proteins, presents a promising combination with microfluidic experiments, and thus can be coupled with separation tools to size different components of a heterogeneous solution under continuous flow. The light wavelength and power used make iSCORR an appealing method to use with living cells, allowing measurement of the mean

diffusion coefficient of a region at the surface or in the cells up to 20 μm , without the use of any tags. This opens up new possibilities for cell biology studies.

Chapter 6

Application of protein detection in seroprevalence studies of SARS-COV2

The content of this chapter is based on a manuscript in preparation:

Marc Emmenegger, Elena De Cecco, David Lamparter[†], Raphaël P. B. Jacquat[†], Julien Riou, Daniel Ebner, Mathias M. Schneider, Itzel Condado Morales, Berre Doğançay, Jingjing Guo, Anne Wiedmer Julie Domange, Marigona Imeri, Rita Moos, Chryssa Zografou, Leyla Batkitar, Lidia Madrigal, Chiara Trevisan, Andres Gonzalez-Guerra, Alessandra Carrella Irina L. Dubach, Catherine K. Xu, Georg Meisl, Vasilis Kosmoliaptsis, Tomas Malinauskas, Nicola Burgess-Brown, Ray Owens, Stephanie Hatch, Juthathip Mongkolsapaya, Gavin R. Screaton, Katharina Schubert, John D. Huck, Feimei Liu, Florence Pojer, Kelvin Lau, David Hacker, Elsbeth Probst-Müller, Carlo Cervia, Jakob Nilsson, Onur Boyman, Lanja Saleh, Katharina Spanaus, Arnold von Eckardstein, Dominik J. Schaer, Nenad Ban, Ching-Ju Tsai, Jacopo Marino, Gebhard F. X. Schertler, Nadine Ebert, Volker Thiel, Jochen Gottschalk, Beat M. Frey, Regina Reimann, Simone Hornemann, Aaron M. Ring, Tuomas P. J. Knowles, Christian L. Althaus, Ioannis Xenarios, David I. Stuart, and Adriano Aguzzi

"Continuous population-level monitoring of SARS-CoV-2 seroprevalence in a large metropolitan region".

A previous preprint version is available in medrxiv:

Marc Emmenegger, Elena De Cecco, David Lamparter[†], Raphaël P. B. Jacquat[†], ... Tuomas P. J. Knowles, Christian L. Althaus, Ioannis Xenarios, David I. Stuart, Adriano Aguzzi
"Early peak and rapid decline of SARS-CoV-2 seroprevalence in a Swiss metropolitan region"[52].

I contributed to this manuscript by developing the analysis, which consists in the analysis of the different ELISA, the elaboration of the new metric which combines the different assays, the prevalence results, as well as writing the paper. The seroprevalence of a given antibody within a population allows us to predict the percentage of the population that have started to develop a defence against a pathogen. In the case of the emergence of a new virus, it helps to predict the percentage that have been exposed to it, when no vaccination is available, and therefore helps decision-makers to base their policies on factual numbers. The study in this chapter is based on two different subsets of the Zürich population, and has a representativeness bias. It focuses on immunoglobulin G (IgG) detection against three different parts of a protein expressed by the SARS-CoV-2 virus. The three IgG results have been combined into one metric in order to have a more robust detection of seroprevalence at a given time of the year. Furthermore, the study develops a model assuming a raising and waning of the antibody IgG in order to get a measure of the people exposed. This chapter provides a direct example of how protein detection can be used for research. Even if for this specific study, instead of the previews method described, an ELISA has been used to detect the antibodies.

6.1 Introduction

In early 2020, a few months after the onset of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic, some hospital systems around the world were on their knees and hundreds of thousands of fatalities from coronavirus disease 2019 (COVID-19) have been registered. Research labs all over the world shut down for non-determined periods, and researchers were looking for ways to be useful to their societies. It is in this context that it became necessary to quickly collect quantifiable and reliable data on SARS-CoV-2 prevalence within the population, and thus for best helping governmental authorities to plan future policies to fight the virus. In order for the data collected to be usable and as minimally distorted as possible, it is necessary to develop a rapid and scalable test which makes it achievable to measure a large representative quantity of civil society. In addition, due to the nature of the low prevalence in the population, the test must meet high criteria for false positive and false negative [18]. The intuitive approach of directly measuring the infection by looking at PCR-based diagnostics has been discarded. This was not only because the nasal swab probing of the population requires more medical personnel and puts them in a potentially hazardous position, but also because, at the time, a study reported false-negative rates of 25% even under the best conditions [107]. The serological approach

has been adopted because of the direct capacity of the lab to perform ELISA at a high rate, and because of the accessibility to a serum bank program within the University Hospital of Zürich (USZ). Furthermore, serological assays picture the immune response of hosts. The antibodies measured can be detected for months in the blood after an infection, contrary to a viral load which decreases within weeks [81]. The serological study therefore gives at a given time the percentage of the population which has encountered the virus, is protected against it, and has a lower chance to spread it. Thanks to an existing structure, a low cost and high rate fully automatised measurement has been possible for thousands of blood samples. With this structure, we aimed to investigate the evolution of the CoV2 seroprevalence in the canton of Zürich using an in-house developed tripartite automated blood immunoassay (TRABI). This canton, despite having a high international population, represents a particularly low prevalence region. Although large serological surveys have been carried out in several countries [5, 28, 158, 165, 202], there is a lack of continuous seroprevalence data. As the waning of CoV2 antibodies has been reported in multiple instances [24, 35, 90, 127, 177, 203], single timepoint serology estimates may yield misleading insights into the true extent of CoV2 spread. Continuous immunosurveys were conducted in a large cohort from the University Hospital of Zürich (n=55814 samples) and blood donors from the Blood Donation Services of the canton of Zürich (n=16291), over a period from December 2019 to December 2020. The vaccination programme started in early 2021 in Zürich has therefore not been considered. Lastly, in 2021, the monitoring of hospital patients and blood donors was pursued in the Aguzzi lab with the same analysis pipeline, but my contribution stopped at the beginning of the year.

6.2 Methodology

The seroepidemiological survey of CoV2 infection in the greater area of Zürich is a population-based study to investigate the temporal evolution of seropositivity for CoV2 in two large independent cohorts. The cohorts come from two independent sources: patients at the University Hospital of Zürich (n=55814 samples) and blood donors in the canton of Zürich (n=16291). Each cohort is separate in two main categories: repandemic samples, or negative subcohort from sampling prior to December 2019, and copandemic samples for sampling from December 2019 to December 2020. A third category from the copandemic is the COVID sample or known positive subcohort, which is necessary to evaluate the efficacy of this method (see Figure 6.1, Cohort). Each of the blood samples has been measured using ELISA toward the presence within the blood of three specific IgG against nucleocapsid

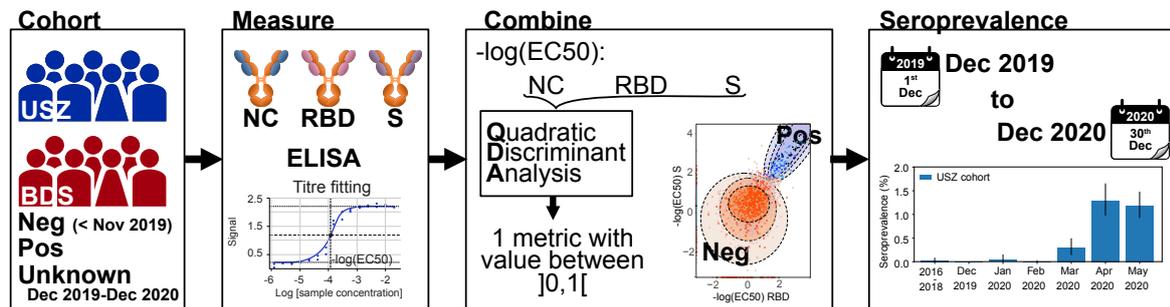


Fig. 6.1 Pipeline of how the study has been conducted. Blood plasma samples from two different cohorts have been collected: the UniversitätsSpital Zürich (USZ) and the Blood Donation Services (BDS). Each of these cohort have a negative sample from blood collected before the pandemic (before 2019), and known positive control from PCR tested patients. The blood is analysed against three proteins/protein parts (NC, RBD, S). The results of the ELISA titration values $-\log(\text{EC}_{50})$ are combined in one metric using a QDA algorithm. Finally, the results of each blood sample are bin per month to plot a seroprevalence of the population.

protein (NC), the receptor binding domain of the spike protein (RBD) and the spike protein (S) (see Figure 6.1, Measure). The three results of the ELISA assays are combined using a quadratic discriminant analysis in order to get only one new metric, which is the posterior probability to be seropositive toward SARS-CoV2 knowing the seroprevalence at a time t of sampling (see Figure 6.1, Combine). Once the collection of points have individual probability, the prevalence can be computed and displayed by month for each cohort (see Figure 6.1, Seroprevalence). All experiments and analyses involving samples from human donors were conducted with the approval of the local ethics committee (KEK-ZH-Nr. 2015-0561, BASEC-Nr. 2018-01042, and BASEC-Nr. 2020-01731), in accordance with the provisions of the Declaration of Helsinki and the Good Clinical Practice guidelines of the International Conference on Harmonisation.

6.2.1 Cohort selection

USZ. The sample used comes from the surplus plasma samples from inpatients and outpatients admitted to the USZ, collected daily (Monday-Friday) and used for population-wide interrogations of the antibody repertoire. For the CoV2 seroprevalence study, we included 4379 samples prior to December 2019 (prepandemic samples) and 51435 samples from December 2019 to December 2020 (copandemic samples). The criteria for this study to include a sample into the analysis were: (1) The patients' blood was sent to the Institute of Clinical Chemistry (at USZ), (2) there was enough residual heparin plasma (at least 150 μl) for the automated generation of a research aliquot, (3) no aliquot from the same patient was

already provided within the same month, and (4) additional information (age, sex, clinical ward to which the patient was admitted) was available. Point (3) led to the exclusion of 415 samples and point (4) to the exclusion of 30 samples for the calculation of the seroestimates. While not being representative of the population of the canton of Zürich in a strict sense, this patient cohort has been selected due to the depth of available medical data that will allow us to trace long-term effects of CoV2 infections from a clinical stance. At the same time, many hospital patients are among the most susceptible within a population and are thus in need of substantial monitoring. The annotation as condition positives was performed post-hoc using USZ databases in the absence of serological data. First, we identified all samples with known positive CoV2 RT-qPCR results ($n=320$). Condition-positive samples ($n=78$) were defined as those with: (1) clinically manifest COVID-19 pneumonia, (2) positive RT-qPCR for CoV2, and (3) venipuncture occurring *geq* 14 days after the positive qPCR to account for seroconversion.

BDS. Similar to others [4, 24, 183, 197, 209], we have investigated CoV2 IgG seroprevalence in a healthy adult population, complementing the hospital patients, in blood donors from the Blood Donation Service of the canton of Zürich. Overall, 16291 samples (thereof 1096 prior to December 2019) from blood donors who consented to further use of their samples for research were randomly selected every month (1170 samples per month on average from December 2019 to December 2020) and analysed. The positives samples ($n=76$) were collected from convalescent individuals with PCR-confirmed CoV2 infection recruited through the blood donation service for a plasmapheresis study. The criteria to be admitted for blood donation are in line with international standards of blood donation services. Blood donors with a confirmed CoV2 infection are excluded from donating blood for four weeks, following the full remission of symptoms. Blood donors have to be at least 18 years of age, weigh at least 50 kg, and feel healthy. In order to be included for blood donation, donors must have not undergone a substantial surgery or pregnancy/birth in the past 12 months, not been subjected to dental treatments in the past 72 hours, and not received foreign blood since January 1, 1980. Moreover, the inclusion mandates that blood donors have not been to an area at risk of malaria or another region with a high prevalence of infectious diseases. Blood donors are only admitted if they have not been tattooed or acquired permanent make-up in the past four months. A positive test for HIV, syphilis, hepatitis C or B leads to a definite exclusion. Additionally, blood donors are excluded if they have had new sexual partners within the last four months and if they display sexually risky behaviour. Lastly, donors must not have been to England, Wales, Scotland, Northern Ireland, Isle of Man, Channel Islands,

Gibraltar or to the Falkland Islands for more than six months between 1980 and 1986. Blood donors over age 65, until maximally age 75, can continue donating blood if they have donated blood earlier (the last, complication-free donation has to date back no longer than two years) and the health survey does not indicate any particular health risk.

6.2.2 Measure

A tripartite automated blood immunoassay (TRABI) has been developed utilising contactless acoustic dispensing [106, 109] to transfer diluted plasma droplets (2.5 nl) into 1536-well plates (total volume 3 μ l) and measuring the IgG response against viral proteins by ELISA. TRABI targets three IgG antigens against the three following protein and protein subdomains: the CoV2 spike protein (S) [186], its receptor binding domain (RBD, amino acids 330-532 of the S protein), and the nucleocapsid protein (NC, amino acids 1-419). Each sample was tested at eight consecutive two-fold dilution points (1:50 to 1:6000), and the resulting data fitted to a sigmoidal curve by logistic regression. The inflection point (or $-\log(\text{EC}_{50})$) of each sigmoid was defined as the respective antibody titer. As reference samples for assay establishment, we utilised a collective of 55 venous plasma samples drawn at various days post-onset of symptoms (dpo) from 27 RT-qPCR confirmed patients suffering from COVID-19 and hospitalised at the University Hospital of Zürich (USZ, true positives, see Table S1), as well as 90 anonymised USZ samples from the prepandemic era (true negatives). We then constructed receiver-operating-characteristic (ROC) curves to assess the assay quality for each antigen individually. Finally, we created a composite metric that integrates S/RBD/NC measurements using quadratic discriminant analysis (QDA). While each single antigen showed excellent discrimination of negatives and positives on samples drawn at ≥ 14 dpo, the compound models outperformed the individual antigen measurements at 7-13 dpo, where the emergence of an IgG response is expected to be variable (Fig. 6.4a, upper panel). We therefore used the QDA modelling assumptions to infer the prevalence in large cohorts based on the distributional information of true negatives and true positives using information gained from all three antigens.

ELISA measurement

Eight-dilution points equally spaced on a logarithmic scale are fitted with an equation derived from a simple binding equilibrium. The inflection point ($\log\text{EC}_{50}$) is extracted from the fit. Baseline and plateau values are fixed by the respective positive and negative controls in a

plate-wise fashion, and the signal is fitted following these equations:

$$c_{bound} = 1 - \frac{1}{2}c_a c + k_d + 1 - \sqrt{(c_a c + k_d)^2 + 2k_d - c_a c + 1}, \quad (6.1)$$

where c_{bound} , c_a and c are antigen-antibody, antigen, and blood concentrations, respectively.

$$OD_{signal} = c_{bound}(baseline - plateau) + plateau \quad (6.2)$$

Combining different metrics in one and seroprevalence analysis

The combined data were first subjected to a preprocessing part. All samples that yielded a $-\log(EC_{50})$ of below -3 on any antigen were labelled as non-fittable and non-detectable. Their dilution curves cannot be differentiated from baseline and therefore only an upper bound for $-\log(EC_{50})$ can be determined. These samples were therefore excluded from data fitting but were of course included in ROC analysis and prevalence estimation as a negative result. The second step was the calculation of TRABI using QDA.

QDA, LDA, and Prevalence estimation. Assume that we have data for m samples with known serostatus and antibody measurements; that is, we have $X_i, Y_i, i = 1, \dots, m$, where X_i is the vector of size p (in our case our antigen measurements) and Y_i is a Boolean variable defining group membership (in our case, whether the individual is seropositive or not). The QDA model assumes multivariate normally distributed X_i given Y_i :

$$X|Y = j \sim \mathcal{N}_p(\mu_j, \Sigma_j). \quad (6.3)$$

Further, the model assumes that the prior, that is, the distribution of Y_i , is known such that $P(Y = j) = \pi_j$. The quadratic discriminant classifier simply assigns each sample to the group which has the larger posterior $P(Y = j | X)$, which is proportional to the joint probability $P(Y = j, X)$.

Therefore, we assign sample i to group 1 if

$$\log(f_{x|y=1}(x_i)) + \log(\pi_1) > \log(f_{x|y=0}(x_i)) + \log(\pi_0), \quad (6.4)$$

and to group 0 otherwise. To set the prior, one option is to take just the mean of the group sizes. However, this is not an ideal option in our case, where we have an additional n samples with unknown serostatus to classify: the prevalence in the m samples with known serostatus might deviate substantially from the prevalence in the population with unknown serostatus. We therefore estimate π_1 directly from the data of unknown serostatus using a

simple expectation maximisation scheme. Proceeding in an iterative fashion, from a given estimate π_1^k , we define the posterior (E step):

$$t_1^k x_i = \frac{\pi_1^k f_{x|y=1} x_i}{\pi_1^k f_{x|y=1} x_i + (1 - \pi_1^k) f_{x|y=0} x_i}. \quad (6.5)$$

Then we update our estimate of π_1 (M step):

$$\pi_1^{k+1} = \sum_i \frac{t_1^k x_i}{n}. \quad (6.6)$$

After convergence, this yields our estimate of the positive serostatus prevalence in the samples. Note that the sample ordering according to this classifier is independent of the prior and therefore has no impact on an analysis via ROC curves. Further, note that evaluating QDA via ROC analysis, an out of sample scheme should be employed to avoid biased estimates of performance; we chose 10-fold cross-validation throughout. Lastly, note that the strategy does not critically depend on the normality assumption but just requires an estimate for the density functions, $f_{x|y=j}(x_i)$. Even nonparametric estimates could be an option. For the LDA approach, we first collapse the antigen measurements per sample according to the linear discriminant classifier:

$$z_i = x_i^T \Sigma_0 \mu_1 - \mu_0, \quad (6.7)$$

where Σ_0 is the covariance estimated from the known negatives only and μ_1 , μ_0 are the means of the known positives and negatives, respectively. The above algorithm is then applied to the resulting one-dimensional variable z_i . 95% confidence intervals were derived by bootstrap drawing 1000 bootstrap samples, where the number of samples drawn from each annotation group (known positives, known negatives and unannotated) was kept constant.

6.3 Results and discussion

6.3.1 Cohort characterisation, and biases

CoV2 antibodies were measured with TRABI in 66630 copandemic samples (collected between December 2019 and December 2020), 51435 belonging to patients of the USZ and 15195 to blood donors. The median age of the USZ patients was 55 (40-68) years (Fig. 6.2.a) and 42 (28-45) years of the blood donors (Fig. 6.2.a), which was stable over the time span of our measurements for the USZ patients (Fig. 6.2.b) but showed deviations for the blood donors, with a decrease in overall age between April and August

2020, followed by an increase in age henceforth (Fig. 6.2.b). The sex distribution in the USZ sample was stable over time, with a female/male ratio close to parity (Fig. 6.2.c). The BDS sample contained slightly more men than expected (Fig. 6.2.c). Most of the hospital patients included in this study were adult residents of the canton of Zürich (Fig. 6.2.d) and were treated in one of the many clinical departments (Fig. 6.3), the highest number in Medical Oncology and Hematology, followed by Cardiology, Infectious Diseases and Hospital Hygiene, Rheumatology, and Gastroenterology and Hepatology. The distribution of samples originating from these hospital wards was relatively stable over time (Fig. 6.3). The USZ and BDS cohorts' age and sex representatively are not entirely congruent with the population's (Fig. 6.2.a.). This bias can be reduced using a post-stratification on both the sex and age knowing the distributional information from the canton of Zürich. Additionally, we aimed to assess the extent of a bias posed by patients with severe COVID-19, hospitalised at the USZ for this reason. Therefore, comparison of seroprevalence should be conducted with the removal of patients (1) admitted to the Infectious Disease and Hospital Hygiene or the Internal Medicine wards, or (2) with annotations like "acute respiratory failure" or "special procedures for testing for CoV2" from the dataset and re-evaluated the course of seroprevalence for the cohort of hospital patients.

6.3.2 TRABI

The question of the advantage of combining three metrics in one is a legitimate query. It is important to emphasise the difference between repeating three of the same measurements, and the three different proteins. If we assume no measurement error, repeating three times will lead to the same value. The system is perfectly correlated for negative and positive values. Furthermore, depending of the intrinsic error of measurement, the way the metric is combined will lead to discarding positive results when a measurement failed, and the system would be less strong. But if we assume no error in measurement and three different proteins are measured, the positive case would be certainly correlated, while the negative will not be. The fact that the body can fight the disease by creating several antibodies leads some targets to be less strong rather than others. If an antibody is too weak (either in affinity or low concentration) to be properly detected, TRABI can compensate by detecting a stronger second antigen.

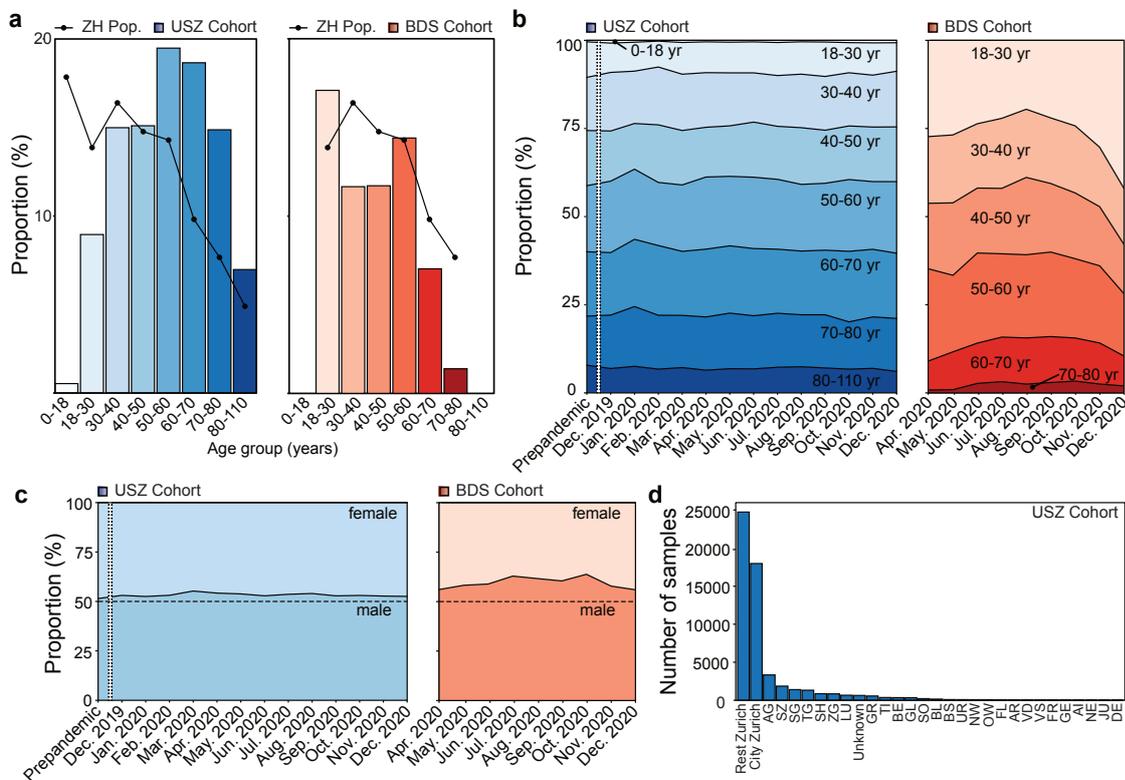


Fig. 6.2 Cohort characterisation. a. Sample distribution according to age groups for USZ and for BDS. b. Proportion of age groups over time for USZ and for BDS. c. Sex distribution over time for USZ and for BDS. d. Repartition geographically of the USZ cohort (BDS cohort data unavailable). Cantonal abbreviations are used, as well as Germany as DE.

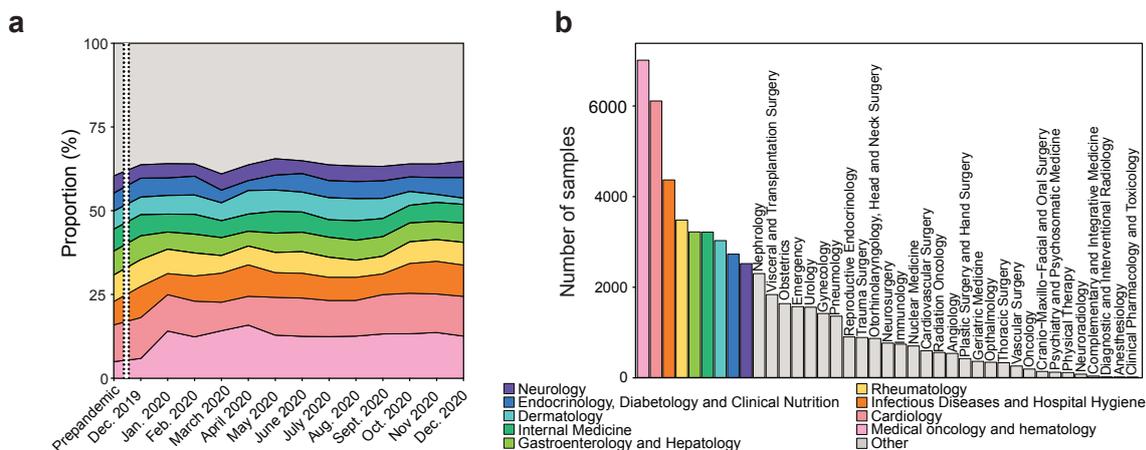


Fig. 6.3 Sample provenance and clinical wards of the USZ patients. a. Distribution of clinical departments over time of study. b. Clinical departments at which patients were treated.

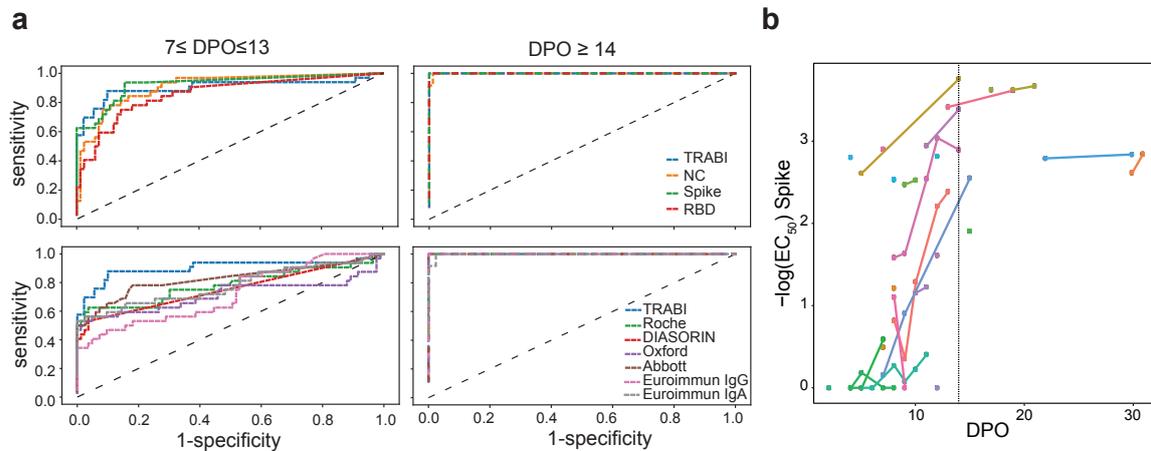


Fig. 6.4 ROC curve and comparison with existing assays, and evolution of Spike pEC₅₀ **a.** Upper panel: Using 53 samples from confirmed CoV2 patients and 83 prepandemic samples, we assessed the specificity-sensitivity relationship for all antigens individually and combining all results into a single score (TRABI) using QDA-based posterior probability. Between 7 and 13 dpo, approximately 60% of samples were positive (i.e. posterior probability > 0.5) at 100% specificity cutoff, whereas 100% sensitivity was reached at ≥14 dpo. Lower panel: COVID and prepandemic samples were used to assess the performance of TRABI, commercial tests (Roche, DiaSorin, Abbott, Euroimmun) and an assay developed at the Target Discovery Institute (Oxford). While all tests scored equally at ≥14 dpo, TRABI outperformed all other assays at ≤13 dpo. **b.** Time course of IgG response in 55 samples from 27 COVID patients. IgG antibodies were reliably detectable at ≥13 dpo. Colours represent individual patients, the dotted line emphasises 14dpo.

6.3.3 Control and comparison of TRABI

Comparison To benchmark TRABI, we compared the results with a high-throughput assay under development at the time at the University of Oxford, as well as assays commercialised by Roche (Elecsys), DiaSorin, EuroImmune, and Abbott (Fig. 6.4a, lower panel). This comparative assessment was based on 136 of 146 samples (10 samples were removed from the analysis because of insufficient sample volume to perform all tests). While all assays displayed 100% specificity/sensitivity at late time points, TRABI scored best at early time points. When these results were plotted as a function of dpo, a temporal pattern emerged consistent with the gradual emergence of IgG antibodies within 14 dpo (Fig. 6.4).

Reproducibility To assess the technical reproducibility of TRABI, we repeated the assay on 200 and 112 randomly selected positive and negative samples, respectively. This repeat screen was found to reproduce the original TRABI results ($R^2 = 0.85$, Fig. 6.5.d).

SARS-CoV vs SARS-CoV2 Antibodies against the RBD of SARS-CoV can bind to the CoV2 RBD 21. We therefore tested whether samples with high anti-CoV2-RBD titers display cross-reactivity with SARS-CoV RBD. For visualisation, we binned samples into groups of absent, moderate and high CoV2 RBD titers ($-\log(\text{EC}_{50}) < 1.5$, $1.5-2$, and > 2.5 , respectively) and computed their respective QDA-derived posterior probability. For individuals with CoV2 RBD titers < 2 , a small fraction showed binding to SARS-CoV RBD at $-\log(\text{EC}_{50}) > 2$ (Fig. 6.5.e). However, those with strong binding properties to CoV2 RBD (> 2.5) clustered at high values for SARS-CoV RBD, indicating that some anti-CoV2 RBD antibodies were cross-reactive to SARS-CoV RBD.

6.3.4 Temporal evolution of seropositivity prevalence in greater Zürich

We applied the QDA-based probability model to estimate the monthly prevalence, from December 2019 to December 2020, using the USZ and the BDS cohorts. No substantial shift above baseline was inferred for samples screened until February 2020 (Fig. 6.5.c). In March 2020, the USZ-based prevalence increased to 0.5% (95% confidence intervals: 0.3%-0.7%) and to 1.6% (CI95%: 1.2%-2.0%) in April 2020, with blood donors displaying a comparable course of seroconversion, with the prevalence approximating 1.3% in April (CI95%: 1.0%-2.0%). The blood donors then reached a first peak in May 2020, with a prevalence of 1.8% (CI95%: 1.3%-2.5%), while the USZ patients plateaued. Following an initial decline in June (USZ: 1.0% (CI95%: 0.8%-1.2%), BDS: 1.4% (CI95%: 0.6%-2.3%)), the seroprevalence fluctuated at around 0.8% over the course of the summer. These summer months were generally characterised by a low reported incidence (4106 new PCR-confirmed cases and 16 COVID-19-associated deaths from July 1 to September 30 in the canton of Zürich [for Open Government Data Canton of Zurich]), until a second wave surged in October. A sharp rise in seroprevalence was observed for November (USZ: 4.0% (CI95%: 3.4%-4.5%), BDS: 2.4% (CI95%: 1.5%-3.2%)) and beginning/mid-December 2020 (USZ: 6.3 (CI95%: 5.5%-7.2%), BDS: 5.1% (CI95%: 4.2%-6.4%)). In addition to the QDA-based model that assumes that both the condition-positive and negative data follow distinct multivariate Gaussian distributions with unequal covariances (Fig. 6.5.a,b), we tested a model based on Gaussian distributions with equal covariances: linear discriminant analysis (LDA) (Fig. 6.7.a). LDA allows us to verify the distributional assumptions more readily (Fig. 6.7.b). Using the distributions of the condition negatives and the condition positives, we computed the posterior probability (i.e. the probability of an individual to be seropositive as modelled via the distribution of the known condition-negatives and known condition-positives) for all

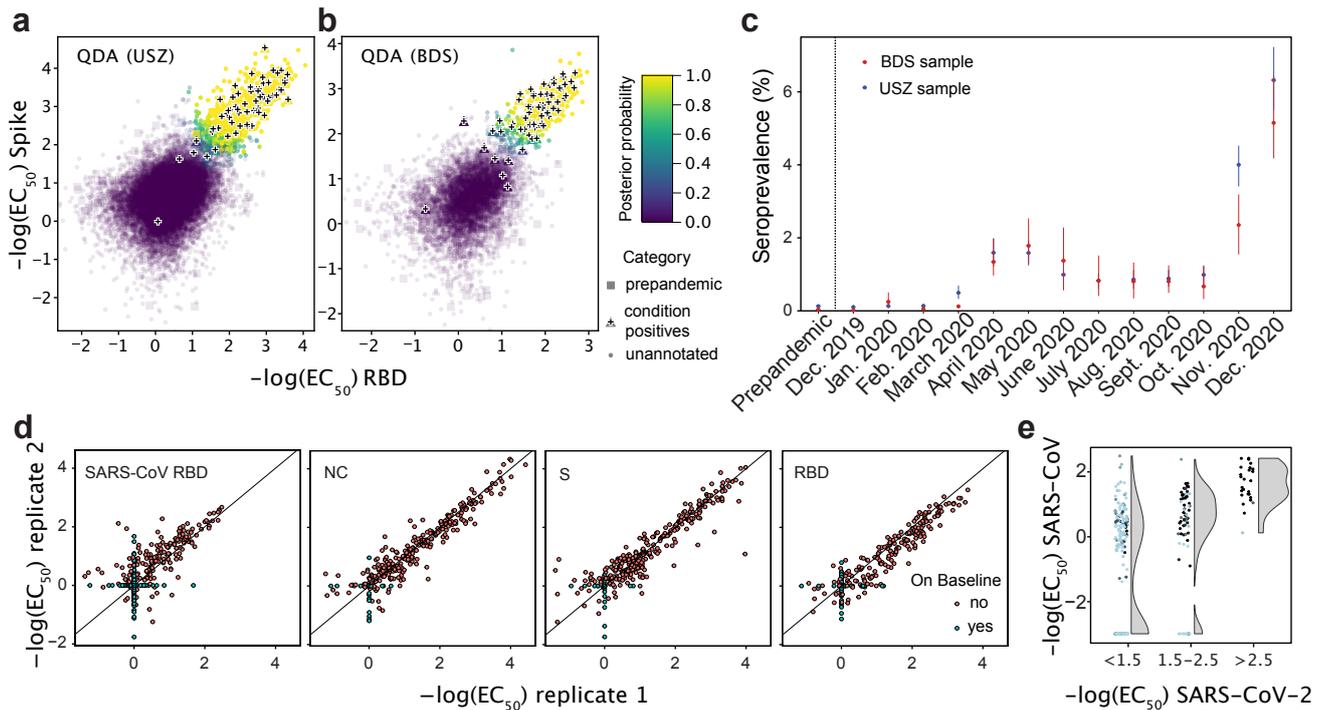


Fig. 6.5 Evolution of CoV2 prevalence in a cohort of Zürich University Hospital (USZ) patients and donors from the blood donation service (BDS). **a-b.** Inflection points of dilution curves, denoted $\log(\text{EC}_{50})$, of plasma titrated against S and RBD in the USZ and BDS cohorts. Posterior probabilities were calculated using QDA assuming a multivariate Gaussian distribution. **c.** Prevalence of CoV2 seropositivity in prepandemic (before December 2019) and co-pandemic samples (from December 2019 to December 2020) estimated using the posterior probabilities from the multivariate Gaussian distribution (QDA). Bar: 95% confidence intervals (CI). **d.** TRABI reproducibility was assessed using duplicates run in pairs of independent assay plates. **e.** To assess potential cross-reactivity of CoV2 seropositive individuals, we tested 200 high-scoring samples and 112 random samples for binding to the RBD of SARS-CoV. CoV2 RBD binders with a high posterior probability (same colour maps as in B) segregated within the higher anti-SARS-CoV-RBD titers.

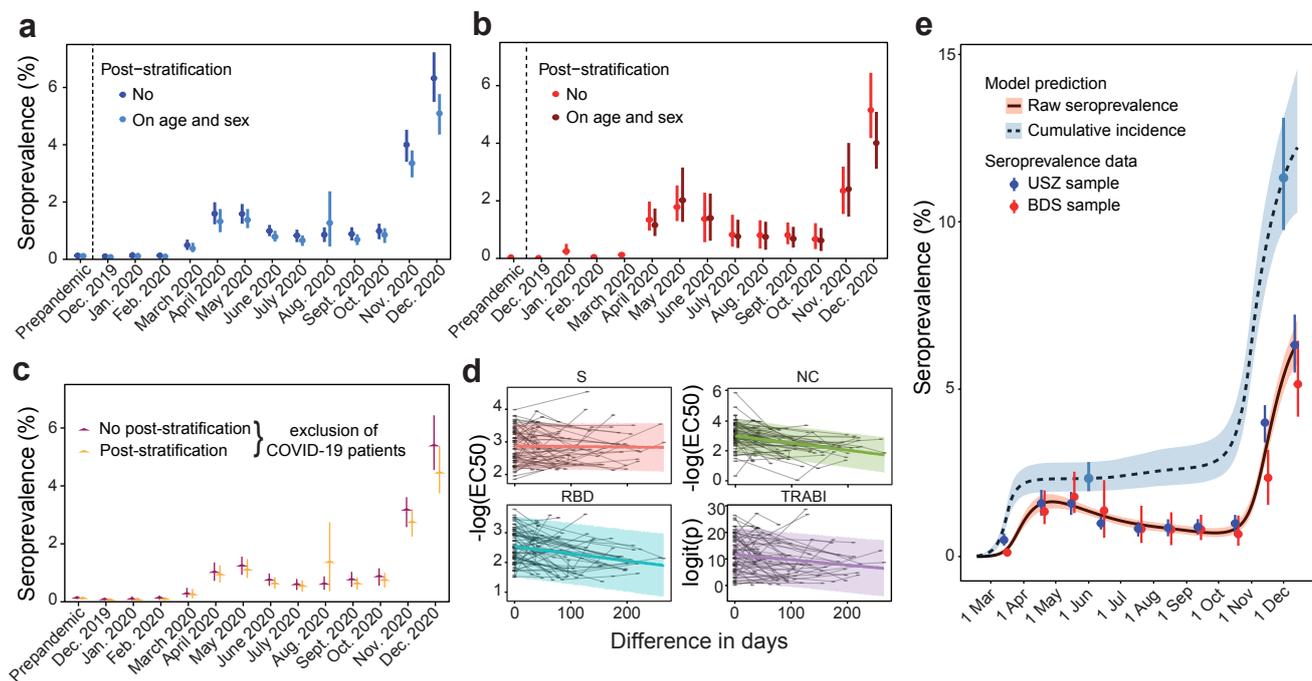


Fig. 6.6 Post stratification and antibody waning. **a.** Seroprevalence in USZ patient cohort after post-stratification on age and sex using the age and sex distributions of the canton of Zürich. **b.** Seroprevalence in BDS cohort after post-stratification on age and sex using the age and sex distributions of the canton of Zürich. **c.** Seroprevalence in USZ patient cohort after removal of patients hospitalised because of COVID-19, for both raw seroprevalence and seroprevalence data after post-stratification on age and sex. **d.** Antibody waning observed with longitudinal sampling. **e.** Cumulative incidence or infection attack rate calculated from seroprevalence corrected for waning. *All the post-stratification data were processed by Julien Riou.*

data points. The respective ROC curves were then plotted (Fig. 6.7.c). At 100% specificity, we identified 78% of the annotated true positives for the USZ (Fig. 6.7.c) and 67% annotated true positives for the BDS cohort (Fig. 6.7.c). For both the USZ and the BDS cohorts, the sensitivity increased rapidly with a slight decrease in specificity (at a false-positive-rate of 0.001, we identified 82% condition positives for USZ and 89% for BDS). After correcting the prevalence using post-stratification on both the sex and age in the USZ and BDS cohorts (Fig. 6.6.a, b), we conclude that this correction led to only minor changes. The maximal effect observed is during December 2020 where the prevalence reaches 5.1% (CI95%: 4.2%-6.4%) unadjusted versus 4.0% (CI95%: 3.1%-5.1%) adjusted. This small difference in prevalence suggests that the two cohorts appropriately reflect the seroprevalence of the adult population. With the removal of the patients with severe COVID-19 hospitalisation, we found that COVID-19 patients contribute to the prevalence observed during both the first as well as

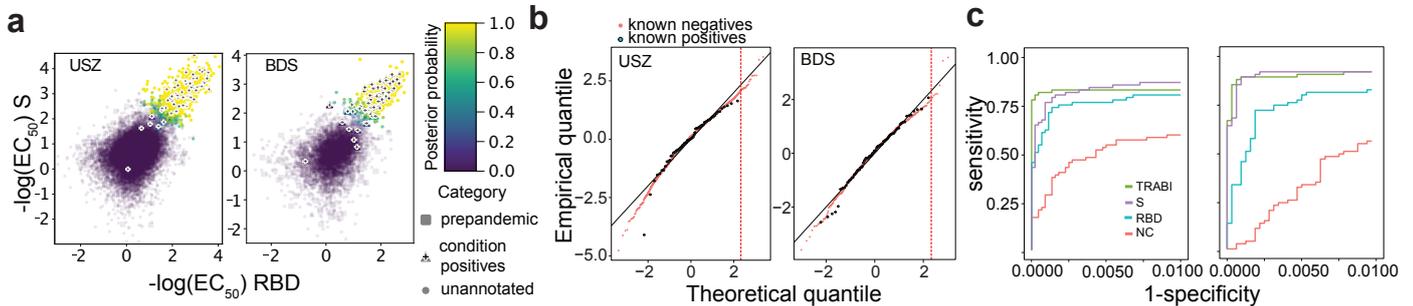


Fig. 6.7 Prevalence estimation in two large cohorts. **a.** Depicted are all the $-\log(\text{EC}_{50})$ values calculated for S and the RBD for the USZ and the BDS cohorts. Posterior probabilities were calculated using LDA. **b.** Q-Q plots to check the Gaussian distributional assumption of the LDA model. For the USZ and BDS cohorts, respectively, we collapsed the 3 measures per sample according to the linear discriminant classifier. We then scaled known positives and negatives to mean zero and unit variance and compared their distributions to the univariate Gaussian distribution via Q-Q plot. We saw that for the relevant upper tail, the distribution of known negatives followed the normal distribution with only a mild deviation, such that, for instance, the upper one percent quantile is reached at 2.07 and 1.99 rather than at their theoretical value of 2.33. **c.** ROC curves for the USZ and BDS cohorts using the pre-pandemic samples (including condition negatives from December 2019 and January 2020 for BDS) as condition negatives and selected condition positives from both cohorts.

the second wave (Fig. 6.6c). Yet, the application of post-stratification on age and sex and the removal of COVID-19 patients did not change the overall dynamics of seroprevalence.

Antibody waning and cumulative incidence The decrease in seroprevalence observed after the peak of the first wave is suggestive of the waning of antibodies at the population level. The availability of repeated samples from the hospital patients allowed us to explore the titers individually. Using data from 65 individuals with a posterior probability ≥ 0.5 and at least two seroestimates, we observe a decrease of all measurements, except for the S protein, over time, including the compound metric (Fig. 6.6), in line with a previous report [58]. We then estimated the half-life of the decrease of the antibody titer directly from the seroprevalence data, using an extension of the classic Susceptible-Exposed-Infectious-Removed (SEIR) model. Assuming an average time to seroconversion of 1/14 days [76, 93, 126], an average generation interval of 5.2 days [67] and an average time from disease onset to death of 20.2 days [124], the overall waning rate observed at the level of the population is 75 (CI95% 55-103) days (unadjusted) or 88 (CI95%: 61-128) days (post-stratification for age and sex). We then computed the cumulative incidence of CoV2, i.e. the seroprevalence corrected for antibody waning, for the population of the canton of Zürich (Fig. 6.6e). The cumulative incidence first raised in March and slowly but gradually increased over the summer period, cumulating to 2.3% (CrI95%: 2.0%-2.8%) in June 2020. A sharp escalation was detectable at the begin-

ning of November, mounting in an cumulative incidence of 12.2% (CrI95%: 10.3%-14.6%) in mid-December 2020. This suggests that over 180000 people had contracted CoV2 until mid-December 2020 in the canton of Zürich. Thus, the cumulative number of cases detected by PCR (55375 until December 13, 2020 [for Open Government Data Canton of Zurich]) is likely to underestimate the true prevalence by approximately factor 3 on average. However, the hidden epidemic ratio (i.e. the number of unobserved cases for each reported case) has changed over time, with a drastic underestimation of cases at the time of the first wave (over 50 cases to 1 detected), a clearly improved precision just before summer 2020 (less than 2 cases for one detected), and a significant underestimation during the second wave (about 5 cases for one detected).

Spatial coverage of the canton of Zürich We aimed to further depict the evolution of seroprevalence in the canton of Zürich. As we avail ourselves of the postcodes, we first mapped the total number of hospital patients per postcode for the months March-July (first wave) and September-December (second wave), 2020 (Fig. 6.8.a,b), only considering the fraction of patients from the canton of Zürich (Fig. 6.2.d). We then investigated the fraction of seropositive hospital patients over the total number of hospital patients per postcode, for above time periods but restricting the analysis to municipalities with at least 50 patients in total, to avoid statistical variability. In line with the overall increased seroprevalence, we observed more than double the number of municipalities (97) showing a prevalence higher than 2% during the second wave, compared to 45 in the first wave (Fig. 6.8.c,d). This result is indicative that the epidemic outbreak in Zürich is not local but extends throughout the canton, with similar rates of increase. The decrease of the fold-change of positive/total cases in the city of Zürich compared to the rest of the canton from the first to the second wave substantiates the observation that after a slightly more localised first outbreak and a remission phase, the second wave is characterised by a non-local spread.

6.4 Conclusions

Using a high-throughput CoV2 serology pipeline, we drew a detailed picture of the evolution of CoV2 seroprevalence in a large central-European metropolitan area. If antibody titers were stable after infection, the seroprevalence would reflect the entirety of the population infected since inception of the pandemic. However, anti-CoV2 titers were found to decay in multiple studies [24, 35, 90, 127, 177, 203], with a half-life of approximately 106 (CI95% 89 to 13) days [24], and others suggesting an even shorter half-life of 26 to 60 days [90].

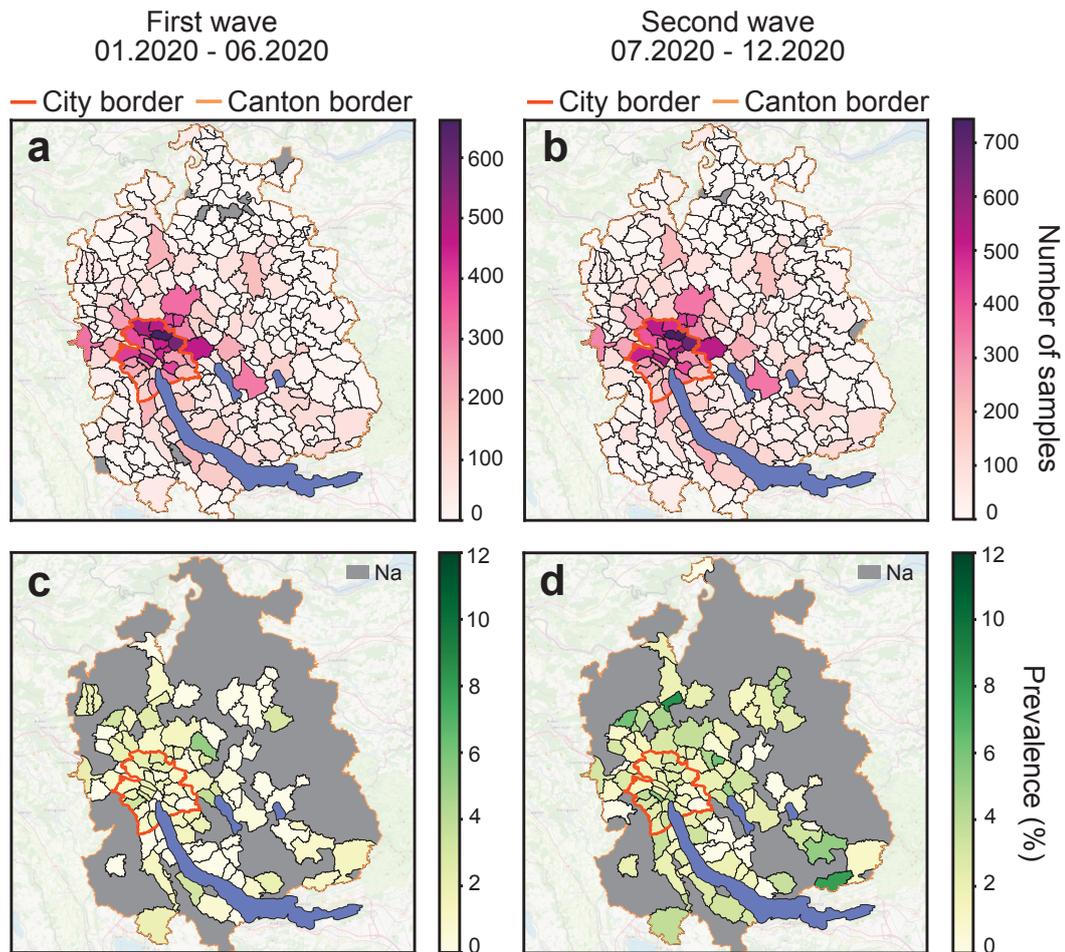


Fig. 6.8 Seroprevalence maps for municipalities in the canton of Zürich. **a.** Samples of hospital patients residing in Zürich sorted according to postcodes. Data from January 2020 to June 2020, including first wave. **b.** Samples of hospital patients residing in Zürich sorted according to postcodes. Data from July 2020 to December 2020, including second wave. **c.** Seropositive samples of hospital patients residing in Zürich sorted according to postcodes. Data from January 2020 to June 2020, including first wave. **d.** Seropositive samples of hospital patients residing in Zürich sorted according to postcodes. Data from July 2020 to December 2020, including second wave. **c** and **d**: Only municipalities with at least 50 samples per postcode are displayed.

This decrease in titers over time was confirmed in neutralisation assays, shown in various studies [35, 127, 177]. Indeed, between April and July 2020 the prevalence of seropositivity fell by $\approx 60\%$ in our cohorts, which confirms the waning of humoral immunity at the population level. Using an extended SEIR model, we estimated that the population-wide half-life of seropositivity is 75 (CI95% 55-103) days (unadjusted seroprevalence data) or 88 (CI95%: 61-128) days (after post-stratification for age and sex). If our sampling methodology suffers from systematic errors, the cohorts sampled here may not be representative of the population studied. In order to minimise such issues, we surveyed two non-overlapping cohorts: hospital in- and outpatients and healthy blood donors. Neither cohort can be assumed to represent a representative random sample of the population. However, post-stratification by age and sex led to only minor changes in seroestimates, indicating that our cohorts are largely representative of the population of the canton of Zürich. However, we have not investigated the extent of CoV2 spread in children in the canton, which was recently done by others [196]. The dynamics of the seroepidemiology confirms that the outbreak followed three distinct phases. The cumulative incidence rose during the first wave in spring 2020, with 2.3% (CrI95%: 2.0%-2.8%) having contracted CoV2 by June 2020. There was a modest increase over the summer months, followed by a rapid rise in late 2020. We estimate that 10.3-14.6% had undergone an infection with CoV2 by mid-December 2020. Thereby, we could delineate the precise serological status in the population of the canton of Zürich in a continuous manner, rather than at single points in time. These estimates of CoV2 antibodies were performed on a highly sensitive immunoassay (TRABI) that combines antibody measurements against three CoV2 proteins in a QDA-based compound metric, a system developed in-house. In view of the critique levelled at past serological studies [18, 189], we have gone to great lengths to assess and validate our technology, using several orthogonal techniques. A recent publication [145] has shown pre-existing anti-CoV2 antibodies in unexposed humans. Antibody sizing [9, 171] and immunoblots, however, point to fundamental differences between pre-pandemic seropositivity and the immune responses of CoV2-infected individuals. While the latter consistently showed high-affinity responses that were clearly visible in Western blotting, the few seropositive pre-pandemic sera were unanimously negative in Western blotting, and equilibrium displacement ELISA of one pre-pandemic plasma sample suggested a much lower affinity despite similar antibody EC50 titers. We conclude that any immune response in uninfected individuals, whether it represents cross-reactivity with common-cold coronaviruses or something else, is of inferior quality and may less likely be protective. A blinded comparison with commercial test kits showed that our approach was suitable for large-scale epidemiologic studies and that the compound

metrics did indeed lead to a power gain, as shown by the enrichment of samples with high posterior probabilities in excess of the single assays during the epidemic. The comparably low seroprevalence of CoV2 in the canton of Zürich, in particular during the first wave, is compatible with other more affected regions, based on the reported IFR, in Switzerland [152] and in European areas with similar medical infrastructure [80]. While some large-scale serological surveys performed throughout the globe revealed CoV2 spread slightly exceeding the values we observed in Zürich [114, 158, 202], other studies identified regions with seroprevalence surpassing 50%, e.g. in some areas in the Amazonas state in Brazil [24] or in slums in Mumbai, India [130]. Yet, since antibody waning has been reported in multiple instances [24, 35, 90, 127, 177, 203], discrete seroestimates may reflect snapshots of the immunity status of a population at a certain time, rather than the true cumulative case incidence. Conversely, we have accounted for antibody waning using a model fit developed by data obtained through continuous CoV2 seromonitoring. Thereby, we were able to derive the cumulative incidence rate for both the first and the second wave of the epidemic in the canton of Zürich and have shown that the nationwide antigen testing underestimated the true number of CoV2 infections by approximately a factor of 3, similar to what was found in France [160].

In addition to the underlying cumulative incidence, from the USZ cohort, an identification of co-morbidities has been conducted within the positive patient blood. This is not developed in this thesis, because the low number of positive cases has not allowed a strong correlation with disease entities beyond those already reported. For future work, as TRABI takes into account the NC IgG level within the detection of the immune response due to SARS-CoV2, it can technically separate the population having antibodies from the mRNA vaccine based on the S protein to a real infection. This is possible only until October 2021, as the only vaccine accepted in Switzerland until this date was vaccine with mRNA technology, assuming as well low the percentage of the population coming from abroad and a more traditional vaccine.

6.5 Acknowledgement

This study is the result of effort from many people, whose help was crucial to complete it. From the synthesis of the S/RBD/NC proteins to the antibody production for the control, as well as the early access collection of test assay from different institutes, and to the blood sample collection from the blood donation centre to the hospital. Part of the studies involving epidemiology/biostatistic tools like the stratification and cumulative waning of the antibody needed the help of J. Riou & Althaus. I would like to thanks Prof. A. Aguzzi for his trust

and allowing me to join this large collaboration. I would like to thank, more specifically, D. Lamparter & I. Xenarios for supporting me and providing knowledge in statistical population study.

Chapter 7

Conclusions

In this thesis I present several novel techniques for protein detection and characterisation. These methods enable us to resolve a variety of biophysical properties such as a protein's size, charge, molecular weight, or the structure of biological complexes. Furthermore, these tools make it possible to study protein and protein aggregates close to biologically relevant conditions, such as at low concentrations and in heterogeneous solutions. Overall, my contribution focuses on the development of the diverse techniques and data analysis, rather than on data acquisitions itself. I would not have been able to cover so many different projects without the contribution of my colleagues and collaborators in Cambridge and in Zürich. The tools described in chapters 2 to 4 explore microfluidics engineering combined with clever optics to extract information about proteins in solution. The capillary electrophoresis (CE) device in combination with the H-filter described in chapter 2 has shown proof-of-concept in extracting sizes within a heterogeneous solution. Two different device designs have been tested. While the first design performing the electrophoretic separation prior to diffusional fractionation was challenging to operate, the second configuration with the inverted work flow performed better separation and sizing. In chapter 3, the aim was to apply confocal microscopy to two well-established microfluidic devices, the diffusional sizing device and the micro free flow electrophoresis (uFFE). While the diffusional sizing device is widely used in the lab, the combination with confocal microscopy enabled us not only to push the detection limit down to picomolar scales, but also to size heterogeneous samples and characterise binding affinity at the low picomolar range. The uFFE device with the improved sensitivity could probe reactions digitally by counting the number of complexes directly. Furthermore, this device can help to resolve the structure of oligomers, helping the research into the toxicity of such aggregates [65]. The fourth chapter investigates a new method to size nanoscopic objects using the residence time inside a nano-cavity. After recording thousands of events,

the residence time distribution shows an exponential decay which is size dependent. A simple mathematical model explains the phenomena without implementing any sort of trapping free energy potential. The last technique developed in the fifth chapter is purely optic based, and uses the principle of interferometric scattering (iSCAT). After presenting some state-of-the-art results based on this technique, I demonstrate the principle of interferometric correlation spectroscopy, which enables the sizing of colloids and small particles in solution. The results of this technique are still preliminary but they promise new applications in combination with microfluidic devices. Finally, the last chapter is a direct application using a well-established analytical tool (ELISA) for studying the SARS-COV2 antibody concentration in human plasma samples. I developed with my colleagues a new metric allowing us to use triparty immunoassay, based on three antigens (Spike/Nucleocapsid/RBD). I also performed the statistical analysis of the COVID seroprevalence within the great area of Zürich.

7.1 Future Applications

The techniques described in this thesis are all still at an early stage of development within the lab. I would like to see the 2D separation CE/H-filter applied to study protein aggregation kinetics. It could then confirm or invalidate some hypotheses on oligomer formation during the aggregation time [136]. Since the oligomers are the intermediate species of the aggregation reactions, they are usually transient and formed at low concentrations, which make them difficult to study with traditional biophysical tools. Confocal microscopy combined with microfluidic devices allows the study of very high affinity reactions with a very low sample volume. An improvement of the scope using two different wavelengths would give more reliable results in single molecule detection, and thus would help to study complex solutions such as serum or body fluids. This technique could be used in diagnostics or treatment monitoring. Finally, the work on iSCORR is in my opinion the most promising to be further developed in the lab. It allows label-free single molecule sizing, without requiring a surface-immobilisation step. In the context of misfolding diseases, the tool can be used to record the aggregation rate over days in a label-free manner. Its downside of not being able to distinguish different protein components within the field of view can be overcome with the aid of microfluidics for separating components of a heterogeneous solution. Thanks to its non-ionised wavelength, the method can be used to study the kinetics of diffusion through tissue or cells without labelling contrary to FRAP [135]. In addition, of interest for all the techniques mentioned in this thesis is, firstly, their development in a high-throughput system. The main reason that the majority of research labs in academia and pharmacology are using

surface-based techniques such as ELISA, or SPR for affinity measurements, is due to their availability in automatised systems. Furthermore, a final aim would be to perform all these assays into a more user-friendly device, as the case of the translation of one version of the diffusion device [9] into a commercial instrument, Fluidity-One W.

References

- [1] Ahn, S. and Fessler, J. A. (2003). Standard errors of mean, variance, and standard deviation estimators. *EECS Department, The University of Michigan*, pages 1–2.
- [2] Akbarzadehlaleh, P., Mirzaei, M., Mashahdi-Keshtiban, M., Shamsasenjan, K., and Heydari, H. (2016). Pegylated human serum albumin: review of pegylation, purification and characterization methods. *Advanced pharmaceutical bulletin*, 6(3):309.
- [3] Alffenaar, J.-W. C., Heysell, S. K., and Mpagama, S. G. (2019). Therapeutic drug monitoring: the need for practical guidance. *Clinical Infectious Diseases*, 68(6):1065–1066.
- [4] Amorim Filho, L., Szwarcwald, C. L., Mateos, S. d. O. G., Leon, A. C. M. P. d., Medronho, R. d. A., Veloso, V. G., Lopes, J. I. F., Porto, L. C. d. M. S., Chieppe, A., and Werneck, G. L. (2020). Seroprevalence of anti-sars-cov-2 among blood donors in rio de janeiro, brazil. *Revista de saude publica*, 54:69.
- [5] Anand, S., Montez-Rath, M., Han, J., Bozeman, J., Kerschmann, R., Beyer, P., Parsonnet, J., and Chertow, G. M. (2020). Prevalence of sars-cov-2 antibodies in a large nationwide sample of patients on dialysis in the usa: a cross-sectional study. *The Lancet*, 396(10259):1335–1344.
- [6] Andreasen, M., Meisl, G., Taylor, J. D., Michaels, T. C., Levin, A., Otzen, D. E., Chapman, M. R., Dobson, C. M., Matthews, S. J., and Knowles, T. P. (2019). Physical determinants of amyloid assembly in biofilm formation. *Mbio*, 10(1):e02279–18.
- [7] Armstrong, J. (2004). Wenby rb, meiselman hj, fisher tc. *Biophys. J*, 87:4259–4270.
- [8] Arosio, P., Michaels, T. C., Linse, S., Månsson, C., Emanuelsson, C., Presto, J., Johansson, J., Vendruscolo, M., Dobson, C. M., and Knowles, T. P. (2016a). Kinetic analysis reveals the diversity of microscopic mechanisms through which molecular chaperones suppress amyloid formation. *Nature communications*, 7(1):1–9.
- [9] Arosio, P., Mueller, T., Rajah, L., Yates, E. V., Aprile, F. A., Zhang, Y., Cohen, S. I., White, D. A., Herling, T. W., De Genst, E. J., et al. (2016b). Microfluidic diffusion analysis of the sizes and interactions of proteins under native solution conditions. *ACS nano*, 10(1):333–341.
- [10] Arosio, P., Müller, T., Rajah, L., Yates, E. V., Aprile, F. A., Zhang, Y., Cohen, S. I., White, D. A., Herling, T. W., De Genst, E. J., et al. (2015). Microfluidic diffusion analysis of the sizes and interactions of proteins under native solution conditions. *ACS nano*, 10(1):333–341.

- [11] Arroyo, J. O., Cole, D., and Kukura, P. (2016). Interferometric scattering microscopy and its combination with single-molecule fluorescence imaging. *Nature protocols*, 11(4):617–633.
- [12] Arter, W. E., Levin, A., Krainer, G., and Knowles, T. P. (2020a). Microfluidic approaches for the analysis of protein–protein interactions in solution. *Biophysical reviews*, 12(2):575–585.
- [13] Arter, W. E., Xu, C. K., Castellana-Cruz, M., Herling, T. W., Krainer, G., Saar, K. L., Kumita, J. R., Dobson, C. M., and Knowles, T. P. (2020b). Rapid structural, kinetic, and immunochemical analysis of alpha-synuclein oligomers in solution. *Nano letters*, 20(11):8163–8169.
- [14] Atamna, I. Z., Metral, C. J., Muschik, G. M., and Issaq, H. J. (1990). Factors that influence mobility, resolution, and selectivity in capillary zone electrophoresis. iii. the role of the buffers' anion. *Journal of liquid chromatography*, 13(16):3201–3210.
- [15] Atencia, J. and Beebe, D. J. (2004). Controlled microfluidic interfaces. *Nature*, 437(7059):648.
- [16] Balluffi, R. W., Allen, S. M., and Carter, W. C. (2005). *Kinetics of materials*. John Wiley & Sons.
- [17] Barnett, C. (1942). Some applications of wave-length turbidimetry in the infrared. *The Journal of Physical Chemistry*, 46(1):69–75.
- [18] Bendavid, E., Mulaney, B., Sood, N., Shah, S., Bromley-Dulfano, R., Lai, C., Weissberg, Z., Saavedra-Walker, R., Tedrow, J., Bogan, A., et al. (2021). Covid-19 antibody seroprevalence in santa clara county, california. *International journal of epidemiology*, 50(2):410–419.
- [19] Benilova, I., Karran, E., and De Strooper, B. (2012). The toxic $\alpha\beta$ oligomer and alzheimer's disease: an emperor in need of clothes. *Nature neuroscience*, 15(3):349.
- [20] Berger Bioucas, F. E., Damm, C., Peukert, W., Rausch, M. H., Koller, T. M., Giraudet, C., and Froeba, A. P. (2019). Translational and rotational diffusion coefficients of gold nanorods dispersed in mixtures of water and glycerol by polarized dynamic light scattering. *The Journal of Physical Chemistry B*, 123(44):9491–9502.
- [21] Biancalana, M. and Koide, S. (2010). Molecular mechanism of thioflavin-t binding to amyloid fibrils. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1804(7):1405–1412.
- [22] Biddison, W. E., Anderson, R. W., Cowan, E. P., Turner, R. V., Coligan, J. E., Hannestad, K., Hansen, T., and Maloy, W. L. (1989). Structural studies of an hla-a03 alloantigenic epitope defined by a human hybridoma antibody. *Immunogenetics*, 30(1):54–57.
- [23] Bruus, H. (2015). Chapter 1 governing equations in microfluidics. pages 1–28.
- [24] Buss, L. F., Prete, C. A., Abraham, C. M., Mendrone, A., Salomon, T., de Almeida-Neto, C., França, R. F., Belotti, M. C., Carvalho, M. P., Costa, A. G., et al. (2021). Three-quarters attack rate of sars-cov-2 in the brazilian amazon during a largely unmitigated epidemic. *Science*, 371(6526):288–292.

- [25] Cabré, F., Canela, E. I., and Canela, M. A. (1989). Accuracy and precision in the determination of Stokes radii and molecular masses of proteins by gel filtration chromatography. *Journal of Chromatography A*, 472:347–356.
- [26] Camilleri, P. (1997). *Capillary electrophoresis: theory and practice*. CRC press.
- [27] Caputo, F., Clogston, J., Calzolari, L., Rösslein, M., and Prina-Mello, A. (2019). Measuring particle size distribution of nanoparticle enabled medicinal products, the joint view of euclid and nci-ncl. a step by step approach combining orthogonal measurements with increasing complexity. *Journal of Controlled Release*, 299:31–43.
- [28] Carrat, F., de Lamballerie, X., Rahib, D., Blanche, H., Lapidus, N., Artaud, F., Kab, S., Renuy, A., de Edelenyi, F. S., Meyer, L., et al. (2020). Seroprevalence of sars-cov-2 among adults in three regions of France following the lockdown and associated risk factors: a multicohort study.
- [29] Challa, P. K., Kartanas, T., Charmet, J., and Knowles, T. P. (2017). Microfluidic devices fabricated using fast wafer-scale led-lithography patterning. *Biomicrofluidics*, 11(1):014113.
- [30] Challa, P. K., Peter, Q., Wright, M. A., Zhang, Y., Saar, K. L., Carozza, J. A., Benesch, J. L., and Knowles, T. P. (2018). Real-time intrinsic fluorescence visualization and sizing of proteins and protein complexes in microfluidic devices. *Analytical chemistry*, 90(6):3849–3855.
- [31] Charmet, J., Arosio, P., and Knowles, T. P. (2018). Microfluidics for protein biophysics. *Journal of molecular biology*, 430(5):565–580.
- [32] Chen, S. W., Drakulic, S., Deas, E., Ouberai, M., Aprile, F. A., Arranz, R., Ness, S., Roodveldt, C., Guilliams, T., De-Genst, E. J., et al. (2015). Structural characterization of toxic oligomers that are kinetically trapped during α -synuclein fibril formation. *Proceedings of the National Academy of Sciences*, 112(16):E1994–E2003.
- [33] Chiti, F. and Dobson, C. M. (2009). Amyloid formation by globular proteins under native conditions. *Nature chemical biology*, 5(1):15.
- [34] Chiti, F. and Dobson, C. M. (2017). Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. *Annual review of biochemistry*, 86:27–68.
- [35] Choe, P. G., Kang, C. K., Suh, H. J., Jung, J., Song, K.-H., Bang, J. H., Kim, E. S., Kim, H. B., Park, S. W., Kim, N. J., et al. (2021). Waning antibody responses in asymptomatic and symptomatic sars-cov-2 infection. *Emerging infectious diseases*, 27(1):327.
- [36] Choudhary, J. and Grant, S. G. (2004). Proteomics in postgenomic neuroscience: the end of the beginning. *Nature neuroscience*, 7(5):440.
- [37] Cohen, S. I., Vendruscolo, M., Dobson, C. M., and Knowles, T. P. (2012). From macroscopic measurements to microscopic mechanisms of protein aggregation. *Journal of molecular biology*, 421(2-3):160–171.

- [38] Cole, D., Young, G., Weigel, A., Sebesta, A., and Kukura, P. (2017). Label-free single-molecule imaging with numerical-aperture-shaped interferometric scattering microscopy. *ACS photonics*, 4(2):211–216.
- [39] Compton, B. J. (1991). Electrophoretic mobility modeling of proteins in free zone capillary electrophoresis and its application to monoclonal antibody microheterogeneity analysis. *Journal of Chromatography A*, 559(1-2):357–366.
- [40] Condado Morales, I. (2019). *Microfluidic Approaches for Investigating Aggregated Forms of Disease-Related Proteins*. PhD thesis, University of Cambridge.
- [41] Croce, A. and Bottiroli, G. (2014). Autofluorescence spectroscopy and imaging: a tool for biomedical research and diagnosis. *European journal of histochemistry: EJH*, 58(4).
- [42] Davies, C. (2005). Introduction to immunoassay principles. *The immunoassay handbook*, 3:40.
- [43] Davies, R. (1963). A molecular theory of muscle contraction: calcium-dependent contractions with hydrogen bond formation plus atp-dependent extensions of part of the myosin-actin cross-bridges. *Nature*, 199(4898):1068.
- [44] Deniz, A. (2008). a; mukhopadhyay, s.; lemke. *E. a. Single-molecule biophysics: at the interface of biology, physics and chemistry. JR Soc. Interface*, 5:15–45.
- [45] Deniz, A. A., Mukhopadhyay, S., and Lemke, E. A. (2008). Single-molecule biophysics: at the interface of biology, physics and chemistry. *Journal of the Royal Society Interface*, 5(18):15–45.
- [46] Dickson, R. M., Cubitt, A. B., Tsien, R. Y., and Moerner, W. E. (1997). On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature*, 388(6640):355–358.
- [47] Dobson, C. M. (2001). The structural basis of protein folding and its links with human disease. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 356(1406):133–145.
- [48] Dolník, V., Liu, S., and Jovanovich, S. (2000). Capillary electrophoresis on microchip. *ELECTROPHORESIS: An International Journal*, 21(1):41–54.
- [49] Douglas Mankiewicz, Bruce B. Weiner, Y. Z. and Brousseau, J.-L. (2019). *Application note Brook Haven Instruments*.
- [50] Duffy, D. C., McDonald, J. C., Schueller, O. J., and Whitesides, G. M. (1998). Rapid prototyping of microfluidic systems in poly (dimethylsiloxane). *Analytical chemistry*, 70(23):4974–4984.
- [51] Eaton, P., Quaresma, P., Soares, C., Neves, C., De Almeida, M., Pereira, E., and West, P. (2017). A direct comparison of experimental methods to measure dimensions of synthetic nanoparticles. *Ultramicroscopy*, 182:179–190.

- [52] Emmenegger, M., De Cecco, E., Lamparter, D., Jacquat, R. P., Ebner, D., Schneider, M. M., Morales, I. C., Schneider, D., Doğançay, B., Guo, J., et al. (2020). Early peak and rapid decline of sars-cov-2 seroprevalence in a swiss metropolitan region. *MedRxiv*, (20118554).
- [53] Engelhardt, H. and Cunat-Walter, M. (1995). Use of plate numbers achieved in capillary electrophoretic protein separations for characterization of capillary coatings. *Journal of Chromatography A*, 717(1-2):15–23.
- [54] Eun, H.-M. (1996). *Enzymology primer for recombinant DNA technology*. Elsevier.
- [55] Evanko, D. (2010). Label-free microscopy. *Nature Methods*, 7(1):36–36.
- [56] Falke, M., Victor, J., Wördehoff, M. M., Peduzzo, A., Zhang, T., Schröder, G. F., Buell, A. K., Hoyer, W., and Etzkorn, M. (2019). α -synuclein-derived lipoparticles in the study of α -synuclein amyloid fibril formation. *Chemistry and physics of lipids*, 220:57–65.
- [57] Favro, L. D. (1960). Theory of the rotational brownian motion of a free rigid body. *Physical Review*, 119(1):53.
- [58] Fenwick, C., Croxatto, A., Coste, A. T., Pojer, F., André, C., Pellaton, C., Farina, A., Campos, J., Hacker, D., Lau, K., et al. (2021). Changes in sars-cov-2 spike versus nucleoprotein antibody responses impact the estimates of infections in population-based seroprevalence studies. *Journal of virology*, 95(3):e01828–20.
- [59] Fiedler, S., Piziorska, M. A., Denninger, V., Morgunov, A. S., Ilsley, A., Malik, A. Y., Schneider, M. M., Devenish, S. R., Meisl, G., Kosmoliaptsis, V., et al. (2021). Antibody affinity governs the inhibition of sars-cov-2 spike/ace2 binding in patient serum. *ACS infectious diseases*.
- [60] Filipe, V., Hawe, A., and Jiskoot, W. (2010). Critical evaluation of nanoparticle tracking analysis (nta) by nanosight for the measurement of nanoparticles and protein aggregates. *Pharmaceutical research*, 27(5):796–810.
- [61] Flagmeier, P., Meisl, G., Vendruscolo, M., Knowles, T. P., Dobson, C. M., Buell, A. K., and Galvagnion, C. (2016). Mutations associated with familial parkinson’s disease alter the initiation and amplification steps of α -synuclein aggregation. *Proceedings of the National Academy of Sciences*, 113(37):10328–10333.
- [for Open Government Data Canton of Zurich] for Open Government Data Canton of Zurich, S. U. Sars-cov-2 open government data reported by the swiss cantons and the principality of liechtenstein.
- [63] Franken, L. E., Boekema, E. J., and Stuart, M. C. (2017). Transmission electron microscopy as a tool for the characterization of soft materials: application and interpretation. *Advanced Science*, 4(5):1600476.
- [64] Frueh, D. P. (2014). Practical aspects of nmr signal assignment in larger and challenging proteins. *Progress in nuclear magnetic resonance spectroscopy*, 78:47–75.

- [65] Fusco, G., Chen, S. W., Williamson, P. T., Cascella, R., Pemi, M., Jarvis, J. A., Cecchi, C., Vendruscolo, M., Chiti, F., Cremades, N., et al. (2017). Structural basis of membrane disruption and cellular toxicity by α -synuclein oligomers. *Science*, 358(6369):1440–1443.
- [66] Gang, H., Galvagnion, C., Meisl, G., Muller, T., Pfammatter, M., Buell, A. K., Levin, A., Dobson, C. M., Mu, B., and Knowles, T. P. (2018). Microfluidic diffusion platform for characterizing the sizes of lipid vesicles and the thermodynamics of protein–lipid interactions. *Analytical chemistry*, 90(5):3284–3290.
- [67] Ganyani, T., Kremer, C., Chen, D., Torneri, A., Faes, C., Wallinga, J., and Hens, N. (2020). Estimating the generation interval for coronavirus disease (covid-19) based on symptom onset data, march 2020. *Eurosurveillance*, 25(17):2000257.
- [68] Garzón, V., Pinacho, D. G., Bustos, R.-H., Garzón, G., and Bustamante, S. (2019). Optical biosensors for therapeutic drug monitoring. *Biosensors*, 9(4):132.
- [69] Gemeinhardt, A., McDonald, M. P., König, K., Aigner, M., Mackensen, A., and Sandoghdar, V. (2018). Label-free imaging of single proteins secreted from living cells via iscat microscopy. *JoVE (Journal of Visualized Experiments)*, (141):e58486.
- [70] German, I., Buchanan, D. D., and Kennedy, R. T. (1998). Aptamers as ligands in affinity probe capillary electrophoresis. *Analytical chemistry*, 70(21):4540–4545.
- [71] Giddings, J. C. (1969). Generation of variance, “theoretical plates,” resolution, and peak capacity in electrophoresis and sedimentation. *Separation Science and Technology*, 4(3):181–189.
- [72] Gobie, W. A. and Ivory, C. F. (1990). Thermal model of capillary electrophoresis and a method for counteracting thermal band broadening. *Journal of Chromatography A*, 516(1):191–210.
- [73] Gollwitzer, C., Bartczak, D., Goenaga-Infante, H., Kestens, V., Krumrey, M., Minelli, C., Pálmai, M., Ramaye, Y., Roebben, G., Sikora, A., et al. (2016). A comparison of techniques for size measurement of nanoparticles in cell culture medium. *Analytical Methods*, 8(26):5272–5282.
- [74] Goloubinoff, P., Sassi, A. S., Fauvet, B., Barducci, A., and Los Rios, P. (2018). Chaperones convert the energy from atp into the nonequilibrium stabilization of native proteins. *Nature chemical biology*, 14(4):388.
- [75] Greener, J. (2014). 14. microfluidics for synthesis and biological functional materials: from device fabrication to applications. In *Functional Materials*, pages 375–414. De Gruyter.
- [76] Guo, L., Ren, L., Yang, S., Xiao, M., Chang, D., Yang, F., Dela Cruz, C. S., Wang, Y., Wu, C., Xiao, Y., et al. (2020). Profiling early humoral response to diagnose novel coronavirus disease (covid-19). *Clinical infectious diseases*, 71(15):778–785.
- [77] Hamley, I. W. (2019). Protein assemblies: nature-inspired and designed nanostructures. *Biomacromolecules*, 20(5):1829–1848.

- [78] Hardy, J. and Selkoe, D. J. (2002). The amyloid hypothesis of alzheimer's disease: progress and problems on the road to therapeutics. *science*, 297(5580):353–356.
- [79] Hatch, A., Garcia, E., and Yager, P. (2004). Diffusion-based analysis of molecular interactions in microfluidic devices. *Proceedings of the IEEE*, 92(1):126–139.
- [80] Hauser, A., Counotte, M. J., Margossian, C. C., Konstantinoudis, G., Low, N., Althaus, C. L., and Riou, J. (2020). Estimation of sars-cov-2 mortality during the early stages of an epidemic: A modeling study in hubei, china, and six regions in europe. *PLoS medicine*, 17(7):e1003189.
- [81] He, X., Lau, E. H., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y. C., Wong, J. Y., Guan, Y., Tan, X., et al. (2020). Temporal dynamics in viral shedding and transmissibility of covid-19. *Nature medicine*, 26(5):672–675.
- [82] Henry, D. (1931). The cataphoresis of suspended particles. part i.—the equation of cataphoresis. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 133(821):106–129.
- [83] Herling, T. W., Levin, A., Saar, K. L., Dobson, C. M., and Knowles, T. P. (2018). Microfluidic approaches for probing amyloid assembly and behaviour. *Lab on a Chip*, 18(7):999–1016.
- [84] Hjertén, S. (1967). Free zone electrophoresis. *Chromatographic reviews*, 9(2):122–219.
- [85] Hoeller, D. and Dikic, I. (2016). How the proteasome is degraded. *Proceedings of the National Academy of Sciences*, 113(47):13266–13268.
- [86] Hofmann, H., Soranno, A., Borgia, A., Gast, K., Nettels, D., and Schuler, B. (2012). Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proceedings of the National Academy of Sciences*, 109(40):16155–16160.
- [87] Holzmeister, P., Acuna, G. P., Grohmann, D., and Tinnefeld, P. (2014). Breaking the concentration limit of optical single-molecule detection. *Chemical Society Reviews*, 43(4):1014–1028.
- [88] Horwich, A. (2002). Protein aggregation in disease: a role for folding intermediates forming specific multimeric interactions. *The Journal of clinical investigation*, 110(9):1221–1232.
- [89] Hosse, M. and Wilkinson, K. J. (2001). Determination of electrophoretic mobilities and hydrodynamic radii of three humic substances as a function of ph and ionic strength. *Environmental science & technology*, 35(21):4301–4306.
- [90] Ibarondo, F. J., Fulcher, J. A., Goodman-Meza, D., Elliott, J., Hofmann, C., Hausner, M. A., Ferbas, K. G., Tobin, N. H., Aldrovandi, G. M., and Yang, O. O. (2020). Rapid decay of anti-sars-cov-2 antibodies in persons with mild covid-19. *New England Journal of Medicine*, 383(11):1085–1087.

- [91] Ivanyi, D. and Van de Meughevel, W. (1984). A monomorphic hla-specific monoclonal antibody, w6/32, reacts with the h-2d b molecule of normal mouse lymphocytes. *Immunogenetics*, 20(6):699–703.
- [92] Jia, Y., Talaga, D. S., Lau, W. L., Lu, H. S., DeGrado, W. F., and Hochstrasser, R. M. (1999). Folding dynamics of single gcn-4 peptides by fluorescence resonant energy transfer confocal microscopy. *Chemical physics*, 247(1):69–83.
- [93] Jiang, C., Wang, Y., Hu, M., Wen, L., Wen, C., Wang, Y., Zhu, W., Tai, S., Jiang, Z., Xiao, K., et al. (2020). Antibody seroconversion in asymptomatic and symptomatic patients infected with severe acute respiratory syndrome coronavirus 2 (sars-cov-2). *Clinical & translational immunology*, 9(9):e1182.
- [94] Jiang, X., Weise, S., Hafner, M., et al. (2010). Jr soc. *Interface*, 7:5.
- [95] Johnson, A. C. and Bowser, M. T. (2017). High-speed, comprehensive, two dimensional separations of peptides and small molecule biological amines using capillary electrophoresis coupled with micro free flow electrophoresis. *Analytical chemistry*, 89(3):1665–1673.
- [96] Jorgenson, J. W. and Lukacs, K. (1981a). Free-zone electrophoresis in glass capillaries. *Clinical chemistry*, 27(9):1551–1553.
- [97] Jorgenson, J. W. and Lukacs, K. D. (1981b). Zone electrophoresis in open-tubular glass capillaries. *Analytical Chemistry*, 53(8):1298–1302.
- [98] Jorgenson, J. W. and Lukacs, K. D. (1983). Capillary zone electrophoresis. *Science*, 222:266–274.
- [99] Kapusta, P. (2010). Absolute diffusion coefficients: Compilation of reference data for fcs calibration.
- [100] Kelly, J. W. (1998). The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Current opinion in structural biology*, 8(1):101–106.
- [101] Kim, J. Y., Ahn, S. H., Kang, S. T., and Yoon, B. J. (2006). Electrophoretic mobility equation for protein with molecular shape and charge multipole effects. *Journal of colloid and interface science*, 299(1):486–492.
- [102] Knowles, T. P., Vendruscolo, M., and Dobson, C. M. (2014). The amyloid state and its association with protein misfolding diseases. *Nature reviews Molecular cell biology*, 15(6):384–396.
- [103] Knowles, T. P., Waudby, C. A., Devlin, G. L., Cohen, S. I., Aguzzi, A., Vendruscolo, M., Terentjev, E. M., Welland, M. E., and Dobson, C. M. (2009). An analytical solution to the kinetics of breakable filament assembly. *Science*, 326(5959):1533–1537.
- [104] Kopp, M. R., Wolf Pérez, A.-M., Zucca, M. V., Capasso Palmiero, U., Friedrichsen, B., Lorenzen, N., and Arosio, P. (2020). An accelerated surface-mediated stress assay of antibody instability for developability studies. In *Mabs*, volume 12, page 1815995. Taylor & Francis.

- [105] Krainer, G., Saar, K. L., Arter, W. E., Jacquat, R. P., Peter, Q., Challa, P., Taylor, C. G., Klenerman, D., and Knowles, T. P. (2020). Direct digital sensing of proteins in solution through single-molecule optofluidics. *bioRxiv*.
- [106] Kryvtuufek, Robin and Vacha, P. (2019). Increasing the throughput of crystallization condition screens: Challenges and pitfalls of acoustic dispensing systems. *MethodsX*, 6:2230–2236.
- [107] Kucirka, L. M., Lauer, S. A., Laeyendecker, O., Boon, D., and Lessler, J. (2020). Variation in false-negative rate of reverse transcriptase polymerase chain reaction–based sars-cov-2 tests by time since exposure. *Annals of internal medicine*, 173(4):262–267.
- [108] Kukura, P., Ewers, H., Müller, C., Renn, A., Helenius, A., and Sandoghdar, V. (2009). High-speed nanoscopic tracking of the position and orientation of a single virus. *Nature methods*, 6(12):923–927.
- [109] Kuleskiy, E., Saarela, J., Turunen, L., and Wennerberg, K. (2016). Precision cancer medicine in the acoustic dispensing era: ex vivo primary cell drug sensitivity testing. *Journal of laboratory automation*, 21(1):27–36.
- [110] Kuzmenkina, E. V., Heyes, C. D., and Nienhaus, G. U. (2005). Single-molecule Förster resonance energy transfer study of protein dynamics under denaturing conditions. *Proceedings of the National Academy of Sciences*, 102(43):15471–15476.
- [111] Lagally, E., Medintz, I., and Mathies, R. (2001). Single-molecule dna amplification and analysis in an integrated microfluidic device. *Analytical chemistry*, 73(3):565–570.
- [112] Laurence, T. A. and Weiss, S. (2003). How to detect weak pairs. *Science*, 299(5607):667–668.
- [113] Lautenschläger, J., Stephens, A. D., Fusco, G., Ströhl, F., Curry, N., Zacharopoulou, M., Michel, C. H., Laine, R., Nespovitaya, N., Fantham, M., et al. (2018). C-terminal calcium binding of α -synuclein modulates synaptic vesicle interaction. *Nature communications*, 9(1):1–13.
- [114] Le Vu, S., Jones, G., Anna, F., Rose, T., Richard, J.-B., Bernard-Stoecklin, S., Goyard, S., Demeret, C., Helynck, O., Escribe, N., et al. (2021). Prevalence of sars-cov-2 antibodies in france: results from nationwide serological surveillance. *Nature communications*, 12(1):1–7.
- [115] Lee, S. F. and Klenerman, D. (2018). Weighing one protein with light. *Science*, 360(6387):378–379.
- [116] Lee, W.-C. and Lee, K. H. (2004). Applications of affinity chromatography in proteomics. *Analytical biochemistry*, 324(1):1–10.
- [117] Leeman, M., Choi, J., Hansson, S., Storm, M. U., and Nilsson, L. (2018). Proteins and antibodies in serum, plasma, and whole blood—size characterization using asymmetrical flow field-flow fractionation (af4). *Analytical and bioanalytical chemistry*, 410(20):4867–4873.
- [118] Li, S. and Wu, Y. (2000). Capillary electrophoresis. *Electrophoresis*, pages 1176–1187.

- [119] Li, Y., Struwe, W. B., and Kukura, P. (2020). Single molecule mass photometry of nucleic acids. *Nucleic acids research*, 48(17):e97–e97.
- [120] Liebel, M., Hugall, J. T., and van Hulst, N. F. (2017). Ultrasensitive label-free nanosensing and high-speed tracking of single proteins. *Nano letters*, 17(2):1277–1281.
- [121] Liedy, W. (2005). Why does the genetic code include only 20 amino acids? *IUBMB Life*, 57(7):523–524.
- [122] Linse, S. (2017). Monomer-dependent secondary nucleation in amyloid formation. *Biophysical reviews*, 9(4):329–338.
- [123] Linse, S., Scheidt, T., Bernfur, K., Vendruscolo, M., Dobson, C. M., Cohen, S. I., Sileikis, E., Lundqvist, M., Qian, F., O'Malley, T., et al. (2020). Kinetic fingerprints differentiate the mechanisms of action of anti- $\alpha\beta$ antibodies. *Nature Structural & Molecular Biology*, 27(12):1125–1133.
- [124] Linton, N. M., Kobayashi, T., Yang, Y., Hayashi, K., Akhmetzhanov, A. R., Jung, S.-m., Yuan, B., Kinoshita, R., and Nishiura, H. (2020). Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. *Journal of clinical medicine*, 9(2):538.
- [125] Lodish, H., Darnell, J. E., Berk, A., Kaiser, C. A., Krieger, M., Scott, M. P., Bretscher, A., Ploegh, H., Matsudaira, P., et al. (2008). *Molecular cell biology*. Macmillan.
- [126] Long, Q.-X., Liu, B.-Z., Deng, H.-J., Wu, G.-C., Deng, K., Chen, Y.-K., Liao, P., Qiu, J.-F., Lin, Y., Cai, X.-F., et al. (2020a). Antibody responses to sars-cov-2 in patients with covid-19. *Nature medicine*, 26(6):845–848.
- [127] Long, Q.-X., Tang, X.-J., Shi, Q.-L., Li, Q., Deng, H.-J., Yuan, J., Hu, J.-L., Xu, W., Zhang, Y., Lv, F.-J., et al. (2020b). Clinical and immunological assessment of asymptomatic sars-cov-2 infections. *Nature medicine*, 26(8):1200–1204.
- [128] Longchamp, J.-N., Rauschenbach, S., Abb, S., Escher, C., Latychevskaia, T., Kern, K., and Fink, H.-W. (2017). Imaging proteins at the single-molecule level. *Proceedings of the National Academy of Sciences*, 114(7):1474–1479.
- [129] Majors, R. E., Barth, H. G., and Lochmueller, C. H. (1984). Column liquid chromatography. *Analytical Chemistry*, 56(5):300–349.
- [130] Malani, A., Shah, D., Kang, G., Lobo, G. N., Shastri, J., Mohanan, M., Jain, R., Agrawal, S., Juneja, S., Imad, S., et al. (2021). Seroprevalence of sars-cov-2 in slums versus non-slums in mumbai, india. *The Lancet Global Health*, 9(2):e110–e111.
- [131] McDonald, M. P., Gemeinhardt, A., Koenig, K., Piliarik, M., Schaffer, S., Voelkl, S., Aigner, M., Mackensen, A., and Sandoghdar, V. (2018). Visualizing single-cell secretion dynamics with single-protein sensitivity. *Nano letters*, 18(1):513–519.
- [132] McKeating, K. S., Aubé, A., and Masson, J.-F. (2016). Biosensors and nanobiosensors for therapeutic drug and response monitoring. *Analyst*, 141(2):429–449.

- [133] Mehling, M. and Tay, S. (2014). Microfluidic cell culture. *Current opinion in Biotechnology*, 25:95–102.
- [134] Meneghello, A., Tartaglia, S., Alvau, M. D., Polo, F., and Toffoli, G. (2018). Biosensing technologies for therapeutic drug monitoring. *Current medicinal chemistry*, 25(34):4354–4377.
- [135] Meyvis, T. K., De Smedt, S. C., Van Oostveldt, P., and Demeester, J. (1999). Fluorescence recovery after photobleaching: a versatile tool for mobility and interaction measurements in pharmaceutical research. *Pharmaceutical research*, 16(8):1153–1162.
- [136] Michaels, T. C., Šarić, A., Curk, S., Bernfur, K., Arosio, P., Meisl, G., Dear, A. J., Cohen, S. I., Dobson, C. M., Vendruscolo, M., et al. (2020). Dynamics of oligomer populations formed during the aggregation of alzheimer's $\alpha\beta 42$ peptide. *Nature chemistry*, 12(5):445–451.
- [137] Modena, M. M., Rühle, B., Burg, T. P., and Wuttke, S. (2019). Nanoparticle characterization: what to measure? *Advanced Materials*, 31(32):1901556.
- [138] Moerner, W. E. and Orrit, M. (1999). Illuminating single molecules in condensed matter. *Science*, 283(5408):1670–1676.
- [139] Mojarad, N. and Krishnan, M. (2012). Measuring the size and charge of single nanoscale objects in solution using an electrostatic fluidic trap. *Nature nanotechnology*, 7(7):448–452.
- [140] Moparthy, S. B., Thieulin-Pardo, G., Mansuelle, P., Rigneault, H., Gontero, B., and Wenger, J. (2014). Conformational modulation and hydrodynamic radii of cp 12 protein and its complexes probed by fluorescence correlation spectroscopy. *The FEBS journal*, 281(14):3206–3217.
- [141] Moser, M. R. and Baker, C. A. (2021). Taylor dispersion analysis in fused silica capillaries: a tutorial review. *Analytical Methods*.
- [142] Mudalige, T., Qu, H., Van Haute, D., Ansar, S. M., Paredes, A., and Ingle, T. (2019). Characterization of nanomaterials: Tools and challenges. *Nanomaterials for food applications*, pages 313–353.
- [143] Müller, T., Arosio, P., Rajah, L., Cohen, S. I., Yates, E. V., Vendruscolo, M., Dobson, C. M., and Knowles, T. P. (2016). Particle-based monte-carlo simulations of steady-state mass transport at intermediate pécelet numbers. *International Journal of Nonlinear Sciences and Numerical Simulation*, 17(3-4):175–183.
- [144] Neves-Petersen, M. T., Petersen, S., and Gajula, G. P. (2012). Uv light effects on proteins: from photochemistry to nanomedicine. In *Molecular Photochemistry-Variou Aspects*. IntechOpen.
- [145] Ng, K. W., Faulkner, N., Cornish, G. H., Rosa, A., Harvey, R., Hussain, S., Ulferts, R., Earl, C., Wrobel, A. G., Benton, D. J., et al. (2020). Preexisting and de novo humoral immunity to sars-cov-2 in humans. *Science*, 370(6522):1339–1343.

- [146] Novotny, L. and Hecht, B. (2012). *Principles of nano-optics*. Cambridge university press.
- [147] Ohshima, H. (2017). A simple algorithm for the calculation of an approximate electrophoretic mobility of a spherical colloidal particle based on the modified poisson-boltzmann equation. *Colloid and Polymer Science*, 295(4):543–548.
- [148] Oosawa, F. and Kasai, M. (1962). A theory of linear and helical aggregations of macromolecules. *Journal of molecular biology*, 4(1):10–21.
- [149] Ortega-Arroyo, J. and Kukura, P. (2012). Interferometric scattering microscopy (iscat): new frontiers in ultrafast and ultrasensitive optical microscopy. *Physical Chemistry Chemical Physics*, 14(45):15625–15636.
- [150] Paegel, B. M., Blazej, R. G., and Mathies, R. A. (2003). Microfluidic devices for dna sequencing: sample preparation and electrophoretic analysis. *Current opinion in biotechnology*, 14(1):42–50.
- [151] Pal, N., Verma, S. D., Singh, M. K., and Sen, S. (2011). Fluorescence correlation spectroscopy: an efficient tool for measuring size, size-distribution and polydispersity of microemulsion droplets in solution. *Analytical chemistry*, 83(20):7736–7744.
- [152] Perez-Saez, J., Lauer, S. A., Kaiser, L., Regard, S., Delaporte, E., Guessous, I., Stringhini, S., Azman, A. S., Alioucha, D., Arm-Vernez, I., et al. (2021). Serology-informed estimates of sars-cov-2 infection fatality risk in geneva, switzerland. *The Lancet Infectious Diseases*, 21(4):e69–e70.
- [153] Peter, Q. (2021). *Advances in biophysical methods for protein detection and characterisation*. PhD thesis, University of Cambridge.
- [154] Petsko, G. A., Ringe, D., and Charnot, M. D. (2008). *Structure et fonction des protéines*. De Boeck Supérieur.
- [155] Pfammatter, M., Andreasen, M., Meisl, G., Taylor, C. G., Adamcik, J., Bolisetty, S., Sánchez-Ferrer, A., Klenerman, D., Dobson, C. M., Mezzenga, R., et al. (2017). Absolute quantification of amyloid propagons by digital microfluidics. *Analytical chemistry*, 89(22):12306–12313.
- [156] Piliarik, M. and Sandoghdar, V. (2014). Direct optical sensing of single unlabelled proteins and super-resolution imaging of their binding sites. *Nature communications*, 5(1):1–8.
- [157] Pinotsi, D., Buell, A. K., Galvagnion, C., Dobson, C. M., Kaminski Schierle, G. S., and Kaminski, C. F. (2014). Direct observation of heterogeneous amyloid fibril growth kinetics via two-color super-resolution microscopy. *Nano letters*, 14(1):339–345.
- [158] Pollán, M., Pérez-Gómez, B., Pastor-Barriuso, R., Oteo, J., Hernán, M. A., Pérez-Olmeda, M., Sanmartín, J. L., Fernández-García, A., Cruz, I., de Larrea, N. F., et al. (2020). Prevalence of sars-cov-2 in spain (ene-covid): a nationwide, population-based seroepidemiological study. *The Lancet*, 396(10250):535–544.

- [159] Polson, A. (1950). The some aspects of diffusion in solution and a definition of a colloidal particle. *The Journal of Physical Chemistry*, 54(5):649–652.
- [160] Pullano, G., Di Domenico, L., Sabbatini, C. E., Valdano, E., Turbelin, C., Debin, M., Guerrisi, C., Kengne-Kuetche, C., Souty, C., Hanslik, T., et al. (2021). Underdetection of cases of covid-19 in france threatens epidemic control. *Nature*, 590(7844):134–139.
- [161] Radamaker, L., Lin, Y.-H., Annamalai, K., Huhn, S., Hegenbart, U., Schönland, S. O., Fritz, G., Schmidt, M., and Fändrich, M. (2019). Cryo-em structure of a light chain-derived amyloid fibril from a patient with systemic al amyloidosis. *Nature communications*, 10(1):1103.
- [162] Reddy, B. S. and Chatterji, B. N. (1996). An fft-based technique for translation, rotation, and scale-invariant image registration. *IEEE transactions on image processing*, 5(8):1266–1271.
- [163] Reymond, L., Huser, T., Ruprecht, V., and Wieser, S. (2020). Modulation-enhanced localization microscopy. *Journal of Physics: Photonics*, 2(4):041001.
- [164] Rissin, D. M., Kan, C. W., Campbell, T. G., Howes, S. C., Fournier, D. R., Song, L., Piech, T., Patel, P. P., Chang, L., Rivnak, A. J., et al. (2010). Single-molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations. *Nature biotechnology*, 28(6):595–599.
- [165] Rudberg, A.-S., Havervall, S., Månberg, A., Falk, A. J., Aguilera, K., Ng, H., Gabrielson, L., Salomonsson, A.-C., Hanke, L., Murrell, B., et al. (2020). Sars-cov-2 exposure, symptoms and seroprevalence in healthcare workers in sweden. *Nature communications*, 11(1):1–8.
- [166] Ruggeri, F. S., Charmet, J., Kartanas, T., Peter, Q., Chia, S., Habchi, J., Dobson, C. M., Vendruscolo, M., and Knowles, T. P. (2018). Microfluidic deposition for resolving single-molecule protein architecture and heterogeneity. *Nature communications*, 9(1):3890.
- [167] Saar, K. L., Zhang, Y., Müller, T., Kumar, C. P., Devenish, S., Lynn, A., Łapińska, U., Yang, X., Linse, S., and Knowles, T. P. (2018). On-chip label-free protein analysis with downstream electrodes for direct removal of electrolysis products. *Lab on a Chip*, 18(1):162–170.
- [168] Scheidt, T., Łapińska, U., Kumita, J. R., Whiten, D. R., Klenerman, D., Wilson, M. R., Cohen, S. I., Linse, S., Vendruscolo, M., Dobson, C. M., et al. (2019). Secondary nucleation and elongation occur at different sites on alzheimer’s amyloid- β aggregates. *Science advances*, 5(4):eaau3112.
- [169] Schmidt, T., Schütz, G., Baumgartner, W., Gruber, H., and Schindler, H. (1996). Imaging of single molecule diffusion. *Proceedings of the National Academy of Sciences*, 93(7):2926–2929.
- [170] Schneider, M. C., Telschow, R., Mercier, G., López-Martinez, M., Scherzer, O., and Schütz, G. J. (2021a). A workflow for sizing oligomeric biomolecules based on cryo single molecule localization microscopy. *Plos one*, 16(1):e0245693.

- [171] Schneider, M. M., Emmenegger, M., Xu, C. K., Morales, I. C., Meisl, G., Turelli, P., Zografou, C., Zimmermann, M. R., Frey, B. M., Fiedler, S., et al. (2021b). Microfluidic affinity profiling reveals a broad range of target affinities for anti-sars-cov-2 antibodies in plasma of covid-19 survivors. *MedRxiv*, pages 2020–09.
- [172] Schneider, M. M., Gautam, S., Herling, T. W., Andrzejewska, E., Krainer, G., Miller, A. M., Trinkaus, V. A., Peter, Q. A., Ruggeri, F. S., Vendruscolo, M., et al. (2021c). The hsc70 disaggregation machinery removes monomer units directly from α -synuclein fibril ends. *Nature communications*, 12(1):1–11.
- [173] Schneider, M. M., Scheidt, T., Priddey, A. J., Xu, C. K., Hu, M., Devenish, S. R., Meisl, G., Dobson, C. M., Kosmoliaptsis, V., and Knowles, T. P. (2020). Microfluidic antibody affinity profiling for in-solution characterisation of alloantibody-hla interactions in human serum. *bioRxiv*.
- [174] Schneider, S. (2014). Determination of protein molecular weight and size using the agilent 1260 infinity multi-detector bio-sec solution with advanced light scattering detection. *Application note Agilent*.
- [175] Schure, M. R. and Lenhoff, A. M. (1993). Consequences of wall adsorption in capillary electrophoresis: theory and simulation. *Analytical Chemistry*, 65(21):3024–3037.
- [176] Schwer, C. and Kenndler, E. (1992). Peak broadening in capillary zone electrophoresis with electroosmotic flow: Dependence of plate number and resolution on charge number. *Chromatographia*, 33(7-8):331–335.
- [177] Seow, J., Graham, C., Merrick, B., Acors, S., Pickering, S., Steel, K. J., Hemmings, O., O’Byrne, A., Kouphou, N., Galao, R. P., et al. (2020). Longitudinal evaluation and decline of antibody responses in sars-cov-2 infection. *Nature microbiology*, 5(12):1598.
- [178] Seuring, C., Ayer, K., Filippaki, E., Barthelmess, M., Longchamp, J.-N., Ringler, P., Pardini, T., Wojtas, D. H., Coleman, M. A., Dörner, K., et al. (2018). Femtosecond x-ray coherent diffraction of aligned amyloid fibrils on low background graphene. *Nature communications*, 9(1):1836.
- [179] Shashkova, S. and Leake, M. C. (2017). Single-molecule fluorescence microscopy review: shedding new light on old problems. *Bioscience reports*, 37(4):BSR20170031.
- [180] Shekunov, B. Y., Chattopadhyay, P., Tong, H. H., and Chow, A. H. (2007). Particle size analysis in pharmaceuticals: principles, methods and applications. *Pharmaceutical research*, 24(2):203–227.
- [181] Shen, Y., Levin, A., Kamada, A., Toprakcioglu, Z., Rodriguez-Garcia, M., Xu, Y., and Knowles, T. P. (2021). From protein building blocks to functional materials. *ACS nano*, 15(4):5819–5837.
- [182] Shi, Z., Huang, J., Huang, X., Huang, Y., Wu, L., and Li, Q. (2020). Resonant scattering enhanced interferometric scattering microscopy. *Nanoscale*, 12(14):7969–7975.

- [183] Slot, E., Hogema, B. M., Reusken, C. B., Reimerink, J. H., Molier, M., Karregat, J. H., IJlst, J., Novotný, V. M., van Lier, R. A., and Zaaijer, H. L. (2020). Low sars-cov-2 seroprevalence in blood donors in the early covid-19 epidemic in the netherlands. *Nature communications*, 11(1):1–7.
- [184] Smoluchowski, M. (1924). Contribution à la théorie de l'endosmose électrique et de quelques phénomènes corrélatifs. *Pisma Mariana Smoluchowskiego*, 1(1):403–420.
- [185] Soltermann, F., Foley, E. D., Pagnoni, V., Galpin, M., Benesch, J. L., Kukura, P., and Struwe, W. B. (2020). Quantifying protein–protein interactions by molecular counting with mass photometry. *Angewandte Chemie*, 132(27):10866–10871.
- [186] Song, W., Gui, M., Wang, X., and Xiang, Y. (2018). Cryo-em structure of the sars coronavirus spike glycoprotein in complex with its host cell receptor ace2. *PLoS pathogens*, 14(8):e1007236.
- [187] Spindler, S., Ehrig, J., König, K., Nowak, T., Piliarik, M., Stein, H. E., Taylor, R. W., Garanger, E., Lecommandoux, S., Alves, I. D., et al. (2016). Visualization of lipids and proteins at high spatial and temporal resolution via interferometric scattering (iscat) microscopy. *Journal of Physics D: Applied Physics*, 49(27):274002.
- [188] Stetefeld, J., McKenna, S. A., and Patel, T. R. (2016). Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophysical reviews*, 8(4):409–427.
- [189] Streeck, H., Schulte, B., Kuemmerer, B., Richter, E., Höller, T., Fuhrmann, C., Bartok, E., Dolscheid, R., Berger, M., Wessendorf, L., et al. (2020). Infection fatality rate of sars-cov-2 infection in a german community with a super-spreading event. *medrxiv*.
- [190] Stroh, C., Wang, H., Bash, R., Ashcroft, B., Nelson, J., Gruber, H., Lohr, D., Lindsay, S., and Hinterdorfer, P. (2004). Single-molecule recognition imaging microscopy. *Proceedings of the National Academy of Sciences*, 101(34):12503–12507.
- [191] Tan, S. and Pepys, M. (1994). Histopathology. *Amyloidosis*, 25:403–414.
- [192] Tiselius, A. (1930). The moving boundary method of studying the electrophoresis of proteins. *Nova Acta Regiae Soc. Sci. Uppsala, Ser. 4*, 7(4):1–107.
- [193] Trieschmann, L., Santos, A. N., Kaschig, K., Torkler, S., Maas, E., Schätzl, H., and Böhm, G. (2005). Ultra-sensitive detection of prion protein fibrils by flow cytometry in blood from cattle affected with bovine spongiform encephalopathy. *BMC biotechnology*, 5(1):26.
- [194] Turgeon, R. T., Fonslow, B. R., Jing, M., and Bowser, M. T. (2010). Measuring aptamer equilibria using gradient micro free flow electrophoresis. *Analytical chemistry*, 82(9):3636–3641.
- [195] Uhlén, M. (2008). Affinity as a tool in life science. *Biotechniques*, 44(5):649–654.
- [196] Ulyte, A., Radtke, T., Abela, I. A., Haile, S. R., Berger, C., Huber, M., Schanz, M., Schwarzmüller, M., Trkola, A., Fehr, J., et al. (2020). Clustering and longitudinal change in sars-cov-2 seroprevalence in school-children: prospective cohort study of 55 schools in switzerland. *medRxiv*.

- [197] Uyoga, S., Adetifa, I. M., Karanja, H. K., Nyagwange, J., Tuju, J., Wanjiku, P., Aman, R., Mwangangi, M., Amoth, P., Kasera, K., et al. (2021). Seroprevalence of anti-sars-cov-2 igg antibodies in kenyan blood donors. *Science*, 371(6524):79–82.
- [198] Verdaasdonk, J. S., Stephens, A. D., Haase, J., and Bloom, K. (2014). B ending the rules: Widefield microscopy and the abbe limit of resolution. *Journal of cellular physiology*, 229(2):132–138.
- [199] Vicidomini, G., Bianchini, P., and Diaspro, A. (2018). Sted super-resolved microscopy. *Nature methods*, 15(3):173–182.
- [200] Virtanen, R. (1974). Zone electrophoresis in a narrow-bore tube employing potentiometric detection-theoretical and experimental study. *Acta Polytechnica Scandinavica-Chemical Technology Series*, (123):1–67.
- [201] Wall, S. (2010). The history of electrokinetic phenomena. *Current Opinion in Colloid & Interface Science*, 15(3):119–124.
- [202] Ward, H., Atchison, C. J., Whitaker, M., Ainslie, K. E., Elliot, J., Okell, L. C., Redd, R., Ashby, D., Donnelly, C. A., Barclay, W., et al. (2020a). Antibody prevalence for sars-cov-2 in england following first peak of the pandemic: React2 study in 100,000 adults. *MedRxiv*.
- [203] Ward, H., Cooke, G., Atchison, C. J., Whitaker, M., Elliott, J., Moshe, M., Brown, J. C., Flower, B., Daunt, A., Ainslie, K. E., et al. (2020b). Declining prevalence of antibody positivity to sars-cov-2: a community study of 365,000 adults. *MedRxiv*.
- [204] Weigl, B. H. and Yager, P. (1999). Microfluidic diffusion-based separation and detection. *Science*, 283(5400):346–347.
- [205] Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999). Hydrodynamic radii of native and denatured proteins measured by pulse field gradient nmr techniques. *Biochemistry*, 38(50):16424–16431.
- [206] Wilson, T. (2011). Resolution and optical sectioning in the confocal microscope. *Journal of microscopy*, 244(2):113–121.
- [207] Wright, M. A., Aprile, F. A., Bellaiche, M. M., Michaels, T. C., Mueller, T., Arosio, P., Vendruscolo, M., Dobson, C. M., and Knowles, T. P. (2018). Cooperative assembly of hsp70 subdomain clusters. *Biochemistry*, 57(26):3641–3649.
- [208] Würger, A. (2010). Thermal non-equilibrium transport in colloids. *Reports on Progress in Physics*, 73(12):126601.
- [209] Xu, R., Huang, J., Duan, C., Liao, Q., Shan, Z., Wang, M., Rong, X., Li, C., Fu, Y., and Wang, H. (2021). Low prevalence of antibodies against sars-cov-2 among voluntary blood donors in guangzhou, china. *Journal of Medical Virology*, 93(3):1743–1747.
- [210] Xu, X., Liu, K., and Fan, Z. H. (2012). Microscale 2d separation systems for proteomic analysis. *Expert review of proteomics*, 9(2):135–147.

- [211] Yates, E. V., Müller, T., Rajah, L., De Genst, E. J., Arosio, P., Linse, S., Vendruscolo, M., Dobson, C. M., and Knowles, T. P. (2015). Latent analysis of unmodified biomolecules and their complexes in solution with attomole detection sensitivity. *Nature chemistry*, 7(10):802.
- [212] Yin, L., Wang, W., Wang, S., Zhang, F., Zhang, S., and Tao, N. (2015). How does fluorescent labeling affect the binding kinetics of proteins with intact cells? *Biosensors and Bioelectronics*, 66:412–416.
- [213] Yokota, O., Terada, S., Ishizu, H., Ujike, H., Ishihara, T., Nakashima, H., Yasuda, M., Kitamura, Y., Ueda, K., Checler, F., et al. (2002). Nacp/ α -synuclein, nac, and β -amyloid pathology of familial alzheimer's disease with the e184d presenilin-1 mutation: a clinicopathological study of two autopsy cases. *Acta neuropathologica*, 104(6):637–648.
- [214] Young, G., Hundt, N., Cole, D., Fineberg, A., Andrecka, J., Tyler, A., Olerinyova, A., Ansari, A., Marklund, E. G., Collier, M. P., et al. (2017). Quantitative mass imaging of single molecules in solution. *bioRxiv*, page 229740.
- [215] Young, G., Hundt, N., Cole, D., Fineberg, A., Andrecka, J., Tyler, A., Olerinyova, A., Ansari, A., Marklund, E. G., Collier, M. P., et al. (2018). Quantitative mass imaging of single biological macromolecules. *Science*, 360(6387):423–427.
- [216] Yuanmung Zhang, Eric Farrell, D. M. and Weiner, Z. (2019). *Application note Brook Haven Instruments*.
- [217] Zahn, R., Liu, A., Lührs, T., Riek, R., von Schroetter, C., García, F. L., Billeter, M., Calzolari, L., Wider, G., and Wüthrich, K. (2000). Nmr solution structure of the human prion protein. *Proceedings of the National Academy of Sciences*, 97(1):145–150.
- [218] Zanchetta, G., Lanfranco, R., Giavazzi, F., Bellini, T., and Buscaglia, M. (2017). Emerging applications of label-free optical biosensors. *Nanophotonics*, 6(4):627–645.

