Biophysical approaches to characterising protein-protein interactions and intermediate species in protein aggregation

Author:
Catherine Kejia Xu

Supervisors:
Prof. Christopher Dobson
Prof. Tuomas Knowles

September 2020
Trinity College
Department of Chemistry, University of Cambridge

This thesis is submitted for the degree of Doctor of Philosophy
Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.
Biophysical approaches to characterising protein-protein interactions and intermediate species in protein aggregation
Catherine Kejia Xu
Department of Chemistry, University of Cambridge

The phenomenon of protein misfolding and aggregation has been associated with over 50 human diseases, including Parkinson’s disease (PD). While the hallmark deposits in aggregation-associated diseases are primarily composed of fibrillar species, intermediate oligomeric species that form during the aggregation process are believed to be a major cause of toxicity. Such species are relatively poorly characterised due to several challenges that render them inaccessible to most conventional techniques; oligomers are only present at extremely low concentrations in the aggregation reaction, and are additionally highly heterogeneous and often transient in nature. In this thesis, I present complementary approaches to address these difficulties. Firstly, an ensemble of stable, kinetically trapped oligomers with varying biophysical characteristics was established, enabling detailed structure-toxicity relationships to be investigated. Secondly, a method for the simultaneous single-molecule level characterisation and fractionation of oligomeric species under native conditions was established and applied to studying intermediate species in Parkinson’s disease-associated protein aggregation. Finally, I present a general approach to optimising experimental design, which maximises both the efficiency and information gain of experiments, and demonstrate its validation and application to several experimental systems.
Acknowledgements

This thesis is dedicated to the late Professor Sir Christopher Dobson, without whom none of this work would have been possible. Chris could not have been a kinder or more supportive mentor, and I will cherish the time I spent under his supervision.

I would like to thank Professor Tuomas Knowles for his support and for providing me with the opportunity to work in his group, and for always providing insightful ideas and advice. Special thanks go to Dr Janet Kumita, for her unwavering patience and incredible efficiency, and who pushed me towards being independent. Thank you to Professor Michele Vendruscolo for his additional input and advice, and for welcoming me into his group. Thanks in particular to the following people within the group: Dr Georg Meisl who was there for and put up with everything; Dr Aviad Levin for endless hours of round the clock therapy; Marta Castellana-Cruz who always understood and kept me sane; Dr Georg Krainer for his invaluable advice and support; Dr William Arter as an amazing collaborator; Dr Serene Chen who knew the answer to everything; and Dr Itzel Condado Morales who introduced me to microfluidics.

Thanks also to everyone in the group, especially those who contributed to the work presented herein: Zhen Du the cell expert and confidante; Manuela Zimmermann for her immense coding knowledge and never judging me; Dr Magda Czekalska, Akhila Jayaram, and Greta Musteikyte for their work on lipids; Dr Tom Scheidt, Matthias Schneider, Ashley Priddey, Marc Emmenegger, and Mengsha Hu for their work relating to diffusional sizing; Jordan McKittrick for his help with information theory; Dr Benedetta Mannini for her expertise on cells and oligomers; Dr Quentin Peter and Raphaël Jacquat for their contributions to the single molecule data analysis; Dr Alexander Dear for oligomer expertise; Arran Collis for his work on lipid disruption and for all-round being a great student; Dr Sean Chia who knows everything about everything; Katarina Pisani and Dr Rodrigo Lessa Cataldi for their Aβ expertise and emotional support; Dr Patrick Flagmeier for his help and guidance; Dr Kadi Liis Saar and Dr Therese Herling for their input on microfluidic electrophoresis; and Ewa Klimont and Swapan Preet for their help in protein purification and all things lab-related. Thank you to the following people outside the group: Dr Nunilo Cremades for her continued support and guidance; Mr Vasilis Kosmoliaptis, Professor Sara Linse, Professor Adriano Aguzzi, and Professor Laura Itzhaki for their collaborations.
Thank you to Dr Karin Müller and all at the Cambridge Advanced Imaging Centre for their invaluable help with transmission electron microscopy, and to Dr Katherine Stott and the Biophysics Facility at the Department of Biochemistry for the training and use of the analytical ultracentrifuge. Additionally, thank you to Professor Johan Paulsson for allowing me to visit his group, and to Jacob Shenker and Dr Noah Olsman for teaching me about and involving me in their work in the Paulsson group.

Outside of the lab, I thank everyone who has put up with me, in particular Dr Raj Pandya, as a reminder that there’s always someone more insane and chaotic than you. Last but not least, I thank the Herchel Smith Fund and Paulina Pieprzyca for providing an incredible network of support and science, along with amazing opportunities, and the Cambridge Trust and Trinity College for their support.
Contents

List of figures ......................................................... 16
List of tables ......................................................... 17

1 Introduction ......................................................... 19
  1.1 Protein misfolding and aggregation .......................... 19
  1.2 Molecular mechanisms of protein aggregation ............... 24
  1.3 Oligomer toxicity .............................................. 30
  1.4 Effects of extrinsic factors on protein aggregation .......... 32

2 Structural characterisation of kinetically trapped α-synuclein oligomers 35
  2.1 Summary ....................................................... 35
  2.2 Background and motivation ................................... 36
  2.3 Approach ....................................................... 39
  2.4 Results and discussion ....................................... 46
    2.4.1 All α-synuclein variants form oligomers with similar size and morphology ........................................... 46
    2.4.2 Tyrosine environments indicate structural differences between variants ........................................... 50
    2.4.3 G51D oligomers display varying degrees of α-helical structure ........................................... 52
    2.4.4 G51D oligomers display high stability towards urea-induced dissociation ........................................... 58
    2.4.5 G51D oligomers display decreased surface hydrophobicity ........................................... 64
    2.4.6 α-Synuclein oligomer cytotoxicity is correlated with α-helical content ........................................... 69
  2.5 Conclusions and future directions ............................. 75
### 3 Microfluidic methods for studying oligomers in protein aggregation reactions

- **3.1 Summary** .................................................. 77
- **3.2 Background and motivation** .......................... 78
- **3.3 Approach** .................................................. 80
- **3.4 Results and discussion** .............................. 83
  - **3.4.1 Bulk fractionation of aggregation reactions** ...... 83
  - **3.4.2 Simultaneous fractionation and single molecule characterisation of kinetic oligomers** .......... 92
  - **3.4.3 Downstream functional characterisation of oligomers** .......... 114
- **3.5 Conclusions and future directions** ................... 123

### 4 Optimisation of experimental design

- **4.1 Summary** .................................................. 126
- **4.2 Background and motivation** .......................... 127
- **4.3 Approach** .................................................. 128
- **4.4 Implementation** .......................................... 131
  - **4.4.1 Quantification of binary molecular interactions** .......... 131
  - **4.4.2 In silico comparison of traditional and prediction methods** .......... 136
  - **4.4.3 In vitro validation and application: Characterising therapeutic antibodies against Aβ isoforms** .......... 148
  - **4.4.4 Application: Antibody quantification for the determination of transplant compatibility** .......... 154
  - **4.4.5 Application: Quantification of severe acute respiratory syndrome coronavirus 2-reactive antibodies in patient sera** .......... 157
- **4.5 Conclusions and future directions** ................... 160

### 5 Materials and methods

- **5.1 Methods: Structural characterisation of kinetically trapped α-synuclein oligomers** .............................. 163
  - **5.1.1 Sample preparation** ................................ 163
  - **5.1.2 Biophysical characterisation** ........................ 165
  - **5.1.3 Cellular toxicity** .................................... 169
- **5.2 Methods: Microfluidic methods for studying oligomers in protein aggregation reactions** .............................. 171
5.2.1 Preparation of samples ........................................ 171
5.2.2 Biophysical characterisation of samples ......................... 172
5.2.3 Preparation of microfluidic devices .............................. 173
5.2.4 μFFE device operation ........................................... 175
5.2.5 Widefield epifluorescence microscopy for μFFE ................. 175
5.2.6 Confocal microscopy for μFFE .................................... 176
5.2.7 Analysis of single molecule data ................................. 178
5.2.8 Lipid disruption device operation ................................. 178

5.3 Methods: Optimisation of experimental design ...................... 180
5.3.1 Characterising therapeutic antibodies against Aβ isoforms ... 180
5.3.2 Antibody quantification for the determination of transplant
compatibility .............................................................. 181
5.3.3 Quantification of severe acute respiratory syndrome coronavirus
   2-reactive antibodies in patient sera ................................. 182
5.3.4 Analysis of microfluidic diffusional sizing binding data ........ 183
5.3.5 Bayesian inference details ......................................... 184

References ................................................................. 187
List of Figures

1.1 Protein folding landscape. ........................................... 20
1.2 Protein fates in the cell. ........................................... 21
1.3 Amyloid fibril structure. ........................................... 22
1.4 Structures of fibril polymorphs of α-synuclein. ............... 23
1.5 Mechanisms of protein aggregation. ................................ 25
1.6 Determination of aggregation mechanisms through fitting kinetic data. 26
1.7 Determination of oligomer reaction pathways through fitting kinetic data. ........................................... 28
1.8 Diversity of oligomers in aggregation reactions. ............... 29
1.9 Mechanisms by which oligomers may exert toxicity. .......... 31

2.1 α-Helical structure correlates with cytotoxicity. ............... 36
2.2 Primary structure of α-synuclein, showing the locations of the familial Parkinson’s disease-associated mutations. ............... 37
2.3 Structure of monomeric α-synuclein when membrane-bound. 39
2.4 Conservation of structural features between oligomers formed through aggregation and lyophilisation. ......................... 41
2.5 Structures of α-synuclein oligomers. ................................ 43
2.6 Schematic representations of toxic and non-toxic α-synuclein oligomers interacting with membranes. ......................... 44
2.7 Structural determinants of HypF-N oligomer toxicity. .......... 45
2.8 TEM images of variant α-synuclein oligomers. .................. 46
2.9 TEM images of variant α-synuclein fibrils. ....................... 47
2.10 DLS analysis of variant α-synuclein oligomers. ................. 48
2.11 Native-PAGE analysis of variant α-synuclein oligomers. .... 48
2.12 Sedimentation velocity analysis of variant α-synuclein oligomers. 49
2.13 Determination of the molar extinction coefficients of variant α-synuclein oligomers. 51
2.14 Intrinsic fluorescence properties of α-synuclein variants. 51
2.15 Characterisation of Y39F α-synuclein oligomers. 52
2.16 CD spectra of variant α-synuclein species. 53
2.17 Oligomer structure is not propagated through successive preparations. 54
2.18 Vacuum pump pressure during lyophilisation does not determine oligomer structure. 55
2.19 Oligomer structure is determined during lyophilisation. 56
2.20 Secondary structure does not alter G51D oligomer size. 56
2.21 FTIR spectra of variant α-synuclein oligomers. 57
2.22 FTIR and CD spectra of G51D oligomers with varying secondary structure. 58
2.23 Stability of oligomers over time. 59
2.24 Quantification of monomer concentrations in urea-induced dissociation reactions. 60
2.25 Native-PAGE analysis of oligomers incubated with urea. 61
2.26 Urea-induced dissociation of variant α-synuclein oligomers. 62
2.27 Denaturation of α-synuclein fibrils. 63
2.28 ANS fluorescence spectra upon binding oligomers. 65
2.29 CD spectra of monomeric α-synuclein binding to SUVs. 66
2.30 CD spectra of monomeric α-synuclein variants upon incubation with trifluoroethanol. 67
2.31 CD spectra of oligomeric α-synuclein binding to SUVs. 67
2.32 Secondary structure changes of α-synuclein variants upon binding to SUVs. 68
2.33 MTT reduction upon incubation with variant α-synuclein oligomers. 70
2.34 Fitted CD spectra of G51D oligomers. 71
2.35 G51D oligomer α-helical content does not influence surface hydrophobicity. 72
2.36 Dot blots of WT and G51D oligomer structural variants. 72
2.37 α-Helical content correlates with oligomer cytotoxicity. 73
2.38 Analysis of correlation between secondary structure and MTT reduction. 74
3.1 Microfluidic electrophoresis methods. 82
3.2 Cysteine location and conjugated dye selection. 84
3.3 Biophysical characterisation of Alexa-488-labelled N122C oligomers. 85
3.4 Analysis of kinetically trapped α-synuclein oligomers by µFFE. 86
3.5 Comparison of µFFE with sedimentation velocity analysis. 89
3.6 Analysis of α-synuclein aggregation mixtures by µFFE. 91
3.7 smFFE experimental overview and confocal microscope setup. 94
3.8 smFFE analysis of monomeric α-synuclein. 96
3.9 smFFE analysis of kinetically trapped α-synuclein oligomers. 98
3.10 Monitoring α-synuclein aggregation kinetics by fluorescence quenching. 100
3.11 smFFE analysis of 23 h kinetic α-synuclein oligomers. 102
3.12 α-Synuclein aggregation timepoints investigated by smFFE. 103
3.13 Continuous scans of aggregation supernatant samples at 300 V. 104
3.14 Proportions of total photon intensities detected in kinetic oligomer events. 105
3.15 Histograms of count rates of kinetic oligomers with different electrophoretic mobilities. 107
3.16 Histograms of count rates of kinetic oligomers with different apparent monomer unit numbers. 108
3.17 Oligomer populations during aggregation. 110
3.18 Schematic of the cross-sectional profile of the µFFE device in the confocal setup. 111
3.19 Oligomer and fibril concentrations of oligomers and fibrils during α-synuclein aggregation. 113
3.20 Schematic of lipid disruption detection by calcein release. 115
3.21 High-sensitivity methods for monitoring lipid disruption. 117
3.22 Lipid disruption device design. 118
3.23 Encapsulation of calcein in vesicles. 119
3.24 Fluorescence microscopy images of on-chip lipid disruption. 120
3.25 Quantification of on-chip lipid disruption. 121
3.26 Redesigned lipid disruption device to monitor disruption kinetics. 122
4.1 Schematic of the prediction approach to experimental design. 127
4.2 Schematic diagrams of commonly used surface-based methods for studying binding interactions. 134
4.3 Schematic diagram of the microfluidic diffusional sizing method. 135
4.4 Binding curves are not always sufficient even in cases when $K_d$ is the only parameter to be determined. ........................................... 138
4.5 Demonstration of the prediction approach when $K_d$ is the only parameter to be determined. ........................................... 141
4.6 Binding curves are not always sufficient when both $K_d$ and concentration are parameters to be determined. ......................... 142
4.7 Validation of the approximations in the weak and strong binding regimes.144
4.8 Binding surface, showing the weak and strong binding regimes. .... 145
4.9 Demonstration of the prediction approach when $K_d$ and concentration are both parameters to be determined. ......................... 147
4.10 Aβ42 fibril structure, determined by ssNMR. ............................... 149
4.11 Comparison of the binding curve approach with the prediction method for gantenerumab/fibril binding. ................................. 151
4.12 Aβ/antibody binding parameters determined through prediction. ... 152
4.13 Binding modes of antibodies to fibrils. ...................................... 153
4.14 Validation of the method in the quantification of HLA-specific antibody binding in serum. ....................................................... 155
4.15 Quantification of alloantibodies in patient serum. ....................... 156
4.16 Quantification of RBD-specific antibodies in SARS-CoV-2 patient samples. ................................................................. 159
4.17 Schematic of experimental automation with prediction. ............... 161

5.1 Fabrication of microfluidic masters and devices. .......................... 174
List of Tables

2.1 Binding parameters for monomeric α-synuclein with SUVs composed of DOPC:DOPE:DOPS (2:5:3 w/w) . . . . . . . . . . . . . . . . . . . . . . . . . . 69

4.1 Affinities and stoichiometries of Aβ42 monomers and fibrils binding to therapeutic antibodies. . . . . . . . . . . . . . . . . . . . . . . . . . 150
Chapter 1

Introduction

1.1 Protein misfolding and aggregation

Proteins play a key role in almost all biological processes, and in order to carry out their proper native function, proteins must fold into their correct three-dimensional structure, determined by their amino acid sequences [1, 2]. However, proteins are also able to form structures other than their native state, which may also be minima in the free energy landscape of protein folding (Figure 1.1) [3, 4]. In order to prevent the formation of these aberrant structures, the complex protein folding process is aided and accelerated in vivo through the constitutive expression of molecular chaperones [5]. Chaperone expression is upregulated in response to an increase in cellular levels of misfolded proteins, through the unfolded protein response (UPR) in the endoplasmic reticulum (ER) and the heat shock response (HSR) in the nucleus and cytosol [6, 7]. If misfolded proteins cannot be refolded, protein degradation systems including the proteasome, ER-associated degradation (ERAD) and autophagy are employed to restore protein homeostasis (Figure 1.2) [5, 2].

However, these systems are not always sufficient to clear misfolded and aggregating proteins, especially as their efficacy may decline as organisms age [8]. The formation of highly ordered, β-sheet rich structures known as amyloid fibrils is believed to be ubiquitous to all proteins, and may represent a global minimum in the free energy of protein structures (Figure 1.1) [9, 3]. This process of protein aggregation is associated with over 50 human disorders, many of which are age-related such as Alzheimer’s and Parkinson’s diseases [4]. Deposits composed of protein aggregates and other cellular components are characteristic of such diseases [10]. However, amyloid fibrils are not
Chapter 1. Introduction

Figure 1.1: Protein folding landscape; protein folding conformations are proposed to lie on a funnel shaped surface corresponding to their free energies. Chaperones aid the transition from partially folded states to the native state local minimum. However, proteins may also misfold and form aggregated species, which may comprise further local minima and possibly the global maximum in the energy landscape. Figure reproduced from [2].
1.1. Protein misfolding and aggregation

Figure 1.2: Protein fates in the cell. Protein aggregates can arise through the unfolding of folded proteins, or through the aggregation of newly synthesised, partially folded proteins. The proteostasis network contributes to the prevention of aggregation, with approximately 180 different chaperone components and regulators (numbers of constituent system components are highlighted in yellow). The actions of such components include the disaggregation of aggregates and assistance in the correct folding of proteins. Figure reproduced from [2].
Chapter 1. Introduction

Figure 1.3: The general structure of amyloid fibrils, illustrated by fibrils of Aβ42. (a) Transmission electron microscopy images of fibrils, showing their periodicity in an image averaged over three fibril segments. (b) Three-dimensional reconstructed structure of the fibril determined by cryo-electron microscopy (cryo-EM), showing the canonical parallel cross-β structure, whereby the subunits form β-sheet structure in a direction perpendicular to the fibril axis. (c) Cross-sectional structure of the fibril, showing the interface between two protofilaments. (d) Residue-level detailed view of the cross-sectional structure. Figure adapted from [13].

only found in disease contexts, but also exist in functional roles [11]. The mechanical properties of the fibril structure have found a role in bacterial biofilms, with FapC and CsgA proteins being the main components of protein fibrils in *Escherichia coli* and *Pseudomonas*, respectively [12]. While functional amyloid seems to be less prevalent in animals, they may be involved in processes such as hormone storage and reproduction [11, 4].

The general structure of amyloid fibrils is conserved between those formed of any constituent protein subunit. Monomeric proteins align in β-sheet formation, perpendicular to the fibril axis, forming protofilaments, which twist together in the mature fibril structure (Figure 1.3) [13]. Strikingly, unlike globular proteins whose structure is generally determined solely by the primary amino acid sequence, the same protein or peptide can form a range of fibril structures (Figure 1.4) [14, 15]. The structures of *in vitro*-generated fibrils can be modulated by environmental conditions such as salt concentration [16]. Moreover, *in vivo*, such polymorphism can give rise to strains, with the parent fibril structure being propagated through fibril amplification.
1.1. Protein misfolding and aggregation

Figure 1.4: Structures of fibril polymorphs of α-synuclein, determined by solid-state NMR (a, f: [19]) and cryo-EM (b: [20], c, d: [21], e: [22]). The N-terminal region is shown in grey, while the NAC region is shown in red, with residues E46 (green) and A53T (blue) highlighted by their space fill representation, for orientation. Figure adapted from [14].

steps [17]. However, although recent work has identified fibril structures at the atomic level that are characteristic of certain diseases, the mechanisms that relate the fibril structure to the disease pathology are not clear [18].

Moreover, the mechanisms that relate protein aggregation to disease are not well understood in the majority of cases [4]. In the few diseases where this relationship is known, the predominant disease mechanism is the build-up of fibrillar protein, such as in cataract formation, where high molecular weight aggregates of lens crystallin proteins causes the opacity of the lens [23]. Another example finds itself in systemic amyloidoses, where toxicity can be, at least partially, attributed to the mechanical load of the large masses of fibrils on internal organs [9].

For the vast majority of protein aggregation-associated disorders however, the pathophysicsology has not yet been elucidated [4]. A key area of study is therefore the determination of the molecular mechanisms of protein aggregation, and their relationships to toxicity. This thesis focuses primarily on α-synuclein, the major protein component of the hallmark Lewy bodies and Lewy neurites of Parkinson’s
disease (PD) [10].

1.2 Molecular mechanisms of protein aggregation

The kinetics of amyloid fibril formation can be followed over time by the use of aggregate-specific fluorescent dyes, most commonly thioflavin-T (ThT) [24]. Analysis of the resulting aggregation curves by the application of chemical kinetics has thus identified the main molecular mechanisms behind protein aggregation [25]. In the general scheme, primary nucleation entails the formation of a growth-competent seed from monomers, which can then grow by the further addition of monomeric or oligomeric subunits, in elongation steps (Figure 1.5) [26]. Fibril mass can amplify exponentially through the secondary processes of fragmentation and secondary nucleation. Secondary nucleation describes the process by which the surfaces of formed fibrils catalyse the further formation of aggregates, while fragmentation of fibrils creates additional sites for elongation through the production of more fibril ends [27, 28]. By considering the differences in dependence of the rates of these processes on monomer and fibril concentrations, analysis of aggregation kinetics can be used to determine the dominant aggregation mechanisms under given conditions (Figure 1.6) [29].

However, while these general reaction pathways are known, the details of the intermediate processes and species are not understood [31, 32]. For example, while it is known that autocatalytic secondary nucleation processes take place, the detailed mechanism/s have not been determined [33]. The fibril surface could contain discretised binding sites at which secondary nucleation can take place, or it could provide a homogeneous catalytic surface. In addition, it is not known whether monomers assemble on the surface and thereby form aggregation-competent nuclei, or whether these secondary nuclei are generated by pre-formed oligomeric species undergoing a structural conversion step, catalysed by the fibril surface (Figure 1.5) [33].

In order to understand exactly how and why proteins aggregate, and the relationship of protein aggregation to disease phenotypes, we must therefore discern the properties, both structural and kinetic, of intermediate species [30, 34]. These species, termed collectively as oligomers, are relatively poorly characterised, due to several challenges in their analysis that render them invisible to most conventional biophysical techniques [35]. One major obstacle is that oligomers only comprise an
1.2. Molecular mechanisms of protein aggregation

**Figure 1.5:** Mechanisms of protein aggregation. (a) General scheme of aggregation, with monomers (m, light blue), progressing through intermediate oligomer species (S, green) before forming fibrils (P, with mass concentration M, dark blue). (b) Microscopic mechanisms of protein aggregation, showing the rate constants used in kinetic models. Monomers (light blue) first undergo a primary nucleation step to form oligomers (green). Oligomers undergo a conversion step to form fibrillar species (dark blue), which can elongate by the addition of further monomers, and also participate in secondary pathways. Secondary pathways may include fragmentation (bottom right) of fibrils into smaller fibril fragments, and secondary nucleation (bottom left), whereby fibril surfaces are able to catalyse the conversion of monomers to oligomers. Figure adapted from [30].
Figure 1.6: Determination of aggregation mechanisms through fitting kinetic data. By fitting aggregation data to kinetic models involving different mechanistic steps (Figure 1.5), the dominant mechanism of aggregation can be determined. For $\alpha\beta_{42}$ aggregation, the fit does not describe the data well when fragmentation is the only secondary pathway included (left), whereas when the secondary mechanism included is secondary nucleation, the fit now overlays with the data (right). The dominant secondary pathway is therefore secondary nucleation. Figure adapted from [29].
1.2. Molecular mechanisms of protein aggregation

extremely low mass fraction of the total aggregation reaction mixture, often less than 1%, corresponding to nanomolar concentrations within reaction mixtures on the order of micromolar total concentrations [36, 30]. Furthermore, even if measurement methods can characterise species at those concentrations, oligomers are not a homogeneous population, and it has been observed that the composition of these species in terms of number of monomers can span orders of magnitude [36, 14]. Finally, many oligomeric states are expected to be unstable and likely to convert rapidly [36, 37, 35, 30, 34].

The role of oligomers in aggregation kinetics are best characterised in the case of the A\(\beta\) peptide, whose aggregation is associated with Alzheimer’s disease [3]. Both its 40- (A\(\beta\)40) and 42-residue (A\(\beta\)42) isoforms have been associated with the disease; A\(\beta\)42 is less abundant in vivo but aggregates more rapidly than A\(\beta\)40 [38]. The total oligomer mass concentration during aggregation was monitored by isolating oligomeric species by size exclusion chromatography (SEC), with subsequent fitting to kinetic models revealing that most oligomers dissociate to monomers, rather than converting to fibrils (Figure 1.7) [31].

Similarly, fitting oligomer populations determined through single molecule methods has yielded insights into the diversity of oligomer species formed both in and between different protein aggregation systems (Figure 1.8) [30, 34]. The kinetic stability of oligomers (indicated by their half-life) and propensity to convert to fibrillar species instead of dissociating to monomers (kinetic productivity) vary widely between different oligomer populations. Moreover, the mechanistic roles of individual subpopulations can be determined; in the case of \(\alpha\)-synuclein, two subpopulations differentiated by their intermolecular Förster resonance energy transfer (FRET) efficiencies, Type A and Type B, have been observed [36, 39, 40]. Kinetic analysis has revealed that all non-toxic Type A oligomers convert to toxic Type B oligomers before conversion to fibrils [34]. The level of detail to which such mechanisms can be determined is dependent on the resolution of the experimental method; further methodological advances in oligomer characterisation and quantification will therefore enable elucidation of detailed molecular mechanisms of aggregation [30, 34, 31]. Chapter Three, "Microfluidic methods for studying oligomers in aggregation reactions" presents the development of methods for the simultaneous characterisation and fractionation of oligomers.
Figure 1.7: Determination of oligomer reaction pathways through fitting kinetic data. Analagous to fitting ThT curves representing the fibril mass concentration of the reaction mixture to determine aggregation mechanisms, as in Figure 1.6, when oligomer population data are included, oligomer reaction pathways can also be determined. In all cases, the kinetic model can reproduce the ThT aggregation data (left plots), and fibril mass data is therefore insufficient to determine oligomer reaction mechanisms. When an oligomer-to-monomer dissociation step is not included in the kinetic model, the oligomer concentrations recorded experimentally do not match those predicted from theory (a, b). Once this step is included, the experimental and theoretical concentrations correspond exactly (c), indicating that oligomer dissociation is a crucial process in the aggregation process. Figure reproduced from [31].
Figure 1.8: Diversity of oligomers in aggregation reactions. (a) The kinetic behaviour of oligomers in aggregation reactions varies widely both between proteins and subpopulations formed from the same protein. Oligomers can be parametrised by their kinetic productivity (their propensity to convert into fibrils rather than dissociate to monomers) and their half-life with respect to both conversion to fibrils and dissociation to monomers, a measure of their stability. These parameters were determined through fitting of smFRET data. Panel reproduced from [30]. (b) Kinetic analysis of Type A and Type B oligomers observed by smFRET in α-synuclein aggregation. Measured concentrations of the two species during the aggregation reaction (b, c), and the classification of oligomers observed during aggregation of monomers (d) and disaggregation of fibrils (e) into Type A and Type B, based on their intermolecular FRET efficiencies. (f) Through such analysis, it could be determined that all oligomers need to convert from the initially formed Type A structure, to Type B, before converting to fibrils. Figure reproduced from [34].
1.3 Oligomer toxicity

In addition to being crucial for the determination of molecular mechanisms of protein aggregation, oligomers have attracted attention due to their potential role in cellular dysfunction [35, 37]. While deposits of fibrillar protein are hallmarks of protein aggregation diseases, oligomeric intermediates are proposed to be the major source of toxicity [10, 4]. Many efforts have therefore striven to determine the structure of oligomers, the features that confer toxicity, and the mechanisms by which oligomers exert such detrimental effects [41, 42]. Despite the multitude of oligomeric structures that may form, a general structural property that oligomers across different systems seem to possess is a highly hydrophobic surface, which may be required for toxicity [43, 44]. Moreover, although the mechanisms by which oligomers exert toxicity seem to be conserved between different proteins, the structural features that dictate the level of toxicity induced by oligomers are relatively poorly understood [43, 44, 42, 35].

Several mechanisms have emerged by which oligomers may cause cellular dysfunction (Figure 1.9) [41]. Oligomers may interact with cell surface receptors, leading to their aberrant activation; in the case of Aβ-associated toxicity, oligomers may activate tumour necrosis factor receptors, leading to cell death [45, 46, 47]. Receptor binding or direct interactions with membranes may also lead to the oligomers gaining entry to the cell, upon which they may bind non-specifically to and sequester cellular components, potentially leading to effects such as mitochondrial dysfunction and proteasome impairment [48, 49, 50, 51, 52, 53]. However, the mechanism that has received the most attention in biophysical studies is the direct interaction with and disruption of lipid membranes [54, 55, 56, 57, 58].

This permeabilisation of lipid membranes may occur by several mechanisms, with two of the major proposed classes being the formation of a pore and a general disruption to the membrane integrity [59]. The amyloid pore hypothesis is based on the observation of pore-like oligomer structures for a range of protein aggregation systems, both in the presence and absence of lipid membranes [54, 55, 56, 37]. In addition to structural evidence indicating a central cavity in the oligomer structure, electrophysiology measurements have found stepwise changes in electrical conductivity across the membrane, suggesting that oligomers may form "open" and "closed" states [60, 61, 62, 63]. Alternatively, the insertion of oligomers into lipid membranes may already be sufficient to disturb the membrane integrity, even without forming a
1.3. Oligomer toxicity

Figure 1.9: Mechanisms by which oligomers may exert toxicity. (Left) Oligomers may interact directly with lipid membranes, leading to membrane permeabilisation, and subsequent disruption of ion homeostasis. (Centre) Oligomers have been found to interact with cell surface receptors, leading to their aberrant activation, which may result in the activation of apoptotic mechanisms. (Right) Oligomers may gain entry to the cell or be formed intracellularly, where they can sequester cellular components such as chaperones and thereby impair proper cellular function.
physical channel through the oligomer [64, 59]. In both cases, the exact structural
determinants that confer toxicity to oligomers are not clear, and are discussed in more
detail in the following chapter, "Structural characterisation of kinetically trapped
α-synuclein oligomers".

1.4 Effects of extrinsic factors on protein aggregation

Given the potential causative role of protein aggregation in disease, the aggregation
process has attracted attention as a key target for therapeutic intervention [65, 4, 66].
A vast body of work is therefore focused on understanding the effects of external
factors on the aggregation reaction, both in the development of therapeutic strategies
and in understanding how cells prevent aggregation. Chaperone proteins may be
involved in curbing aggregation before disease onset; the decline of cellular anti-
aggregation pathways such as chaperone-mediated refolding and degradation is
believed to be a key contributor to the development of age-related aggregation
diseases such as PD and AD [2, 4]. Understanding how chaperone proteins clear
misfolded and aggregated proteins may both inform on how disease pathophysiology
begins, and how to develop successful therapies based on chaperone action.

A central theme in the study of extrinsic factors in protein aggregation is how they
influence the microscopic steps in the reaction. Determining quantitatively which
steps in the aggregation mechanism (Figure 1.5) are affected is not trivial, due to the
complexity of the reaction network [27]. Recently, strategies to quantitatively assess
the effects of potential inhibitors have been developed, requiring the measurement of
affected aggregation kinetics in the presence of varying concentrations of seed fibrils
[67, 66]. Additionally, such approaches are often associated with large errors on rate
constants, due to the large number of free parameters [67, 68].

Moreover, simply inhibiting the aggregation process is unlikely to be sufficient in
alleviating toxicity, since intermediate oligomeric species, rather than mature fibrils,
are believed to be the major toxic entities [4, 42]. Accordingly, the small molecule
trodusquemine has been found to accelerate aggregation, but decrease the associated
toxicity by reducing the concentration of toxic oligomers [69]. In order to be able to
most effectively design therapeutic strategies based around aggregation-associated
toxicity, we therefore need to both determine the mechanisms that produce toxic
species, and be able to isolate these mechanisms in the analysis. Chapter Four, "Optimisation of experimental design" presents key steps towards this goal.
Chapter 2

Structural characterisation of kinetically trapped $\alpha$-synuclein oligomers

The mammalian cell-based experiments presented in this chapter were performed by Marta Castellana-Cruz.

2.1 Summary

A wide variety of oligomeric structures is formed during the aggregation of proteins associated with neurodegenerative diseases. Such soluble oligomers elicit a range of biological responses and some are thus believed to be key toxic species in the related disorders. The identification of the structural determinants of toxicity is therefore of central importance. We have analysed the structure and cytotoxicity of a variety of oligomers formed by $\alpha$-synuclein and its mutational variants, whose aggregation is associated with sporadic and familial Parkinson’s disease, respectively, in order to identify biophysical characteristics that confer toxicity. By exploiting the remarkable structural polymorphism of G51D oligomers, we correlate the level of induced cellular dysfunction induced with the changes in secondary structure elements, thus identifying $\alpha$-helical content as a determinant of toxicity. Given the observation of $\alpha$-helical intermediates during the aggregation of a number of other proteins, our findings may have broader implications for amyloid toxicity.
2.2 Background and motivation

In Parkinson’s disease (PD), aggregates of the 14 kDa protein α-synuclein are the major component of Lewy bodies and neurites, which emerge as the pathological hallmarks of the disease. In solution, α-synuclein is intrinsically disordered, however, upon insertion into membranes, α-synuclein has been observed to adopt an α-helical structure [70]. The association with membranes appears to be a key trigger of aggregation [71, 72, 73], and α-synuclein is likely to be associated with membranes in an in vivo functional context [74]. During the aggregation process to form the canonical β-sheet-rich fibrillar structures, α-synuclein forms oligomers with varying toxicity. Recently through a combination of solution and solid-state NMR studies, two structural features were identified to be relevant for α-synuclein oligomer toxicity: an N-terminal region that anchors the oligomer to the lipid membrane and a hydrophobic cross-β structure that inserts into the lipid bilayer, causing strong membrane perturbation and disruption [75, 64].

In addition to its abundance in the characteristic amyloid deposits in PD, α-synuclein is further implicated in PD disease development by the finding that
2.2. Background and motivation

Figure 2.2: Primary structure of α-synuclein, showing the locations of the familial Parkinson’s disease-associated mutations. The N-terminal, NAC, and C-terminal regions are highlighted in pink, blue, and yellow, respectively. The locations of the familial PD-associated mutations are highlighted in dark blue.

 duplications and triplications of the WT α-synuclein gene, as well as a number of single-point mutations, are associated with familial cases of PD [76, 77] (Figure 2.2). The causative relationship between an increased load of the protein, which results from the duplication or triplication of the gene, and earlier onset of the disease is likely to be simply a result of the increased aggregation propensity of α-synuclein due to its increased concentration. By contrast, the aetiology of the familial cases associated with the pathological variants remains unknown and a variety of mechanisms and reasons for the connection of these mutations to disease have been proposed [77, 78, 79, 80, 81, 82, 83, 54, 84].

Each mutation is associated with different signatures of disease symptoms, and varying ages of onset; while disease onset has been observed as early as age 19 in individuals with the G51D mutation, cases have only been found starting at age 56 for the H50Q mutation [85]. The effects of the mutations in cellular contexts, most prominently in yeast models, have been extensively studied, finding effects of the mutations on the localisation and accumulation of α-synuclein. While the WT and A53T GFP-fused proteins formed cytoplasmic aggregates, the A30P construct remained cytoplasmically diffuse, and the E46K construct localised at the plasma membrane with little apparent aggregation behaviour [86, 87, 88]. Similarly, the
Chapter 2. Structural characterisation of kinetically trapped \(\alpha\)-synuclein oligomers

G51D has been reported to behave similarly to the A30P strain, forming few inclusions and exhibiting impaired membrane association ability when expressed in yeast [89]. However, the behaviour seems to be somewhat construct and strain-dependent, with other work finding that both the E46K and G51D mutations increased the frequency of inclusion formation [90]. Moreover, such studies in vivo all require the fusion of \(\alpha\)-synuclein with a fluorescent protein, to enable visualisation. The presence of the fluorescent protein is likely to induce artefacts in the observed behaviour of \(\alpha\)-synuclein; biophysical approaches to understanding the label-free behaviour of these variants have therefore been employed in tandem [91].

All familial PD-associated mutations identified thus far are located in the N-terminal region of \(\alpha\)-synuclein. This region of the protein is required for its membrane binding properties, which is hypothesised to be required for its native function [92, 93, 74, 70]. The familial PD-associated mutations have been found to alter the membrane binding properties of \(\alpha\)-synuclein [94, 95, 96, 97, 98, 99, 100, 89, 80, 101, 102, 103]. However, while the A30P and G51D mutations have been reported to abrogate \(\alpha\)-synuclein/membrane interactions, the E46K variant may enhance membrane binding, so the potential role of these mutations’ influence on membrane binding properties is not clear [100, 94, 98, 95, 99, 104].

Furthermore, studies on the in vitro aggregation kinetics of these \(\alpha\)-synuclein variants have yielded conflicting results [105, 78, 79, 106, 80, 89, 81, 82, 83, 107, 54, 108, 77]. While the E46K, H50Q, and A53T variants have been found to aggregate more rapidly than the WT, the G51D exhibits a lower propensity for aggregation, and the A30P appears to be the most variable in behaviour. Nevertheless, it is clear that aggregation kinetics are insufficient to explain the link between these mutations and disease.

Another aspect of protein aggregation that may be involved in toxicity is the nature of the intermediate oligomeric species [3, 43, 109]. Such intermediates have been found to exert greater cellular toxicity than fibrils, due to their highly hydrophobic surface [110, 111]. Consequently, several studies have sought to characterise the effects of familial PD-associated mutations on \(\alpha\)-synuclein oligomers. Paslawski et al. used hydrogen/deuterium exchange mass spectrometry to study oligomers of WT \(\alpha\)-synuclein and the A30P, E46K, and A53T variants, finding very subtle differences between deuterium exchange profiles [112]. Furthermore, through single molecule FRET (smFRET) experiments on the same variants, the concentrations
2.3 Approach

Amyloid oligomers are extremely challenging to study, due to their intrinsic heterogeneity, transiency and presence at low concentrations in aggregation reactions.
These properties therefore do not permit their detailed biophysical characterisation by most conventional techniques. In order to avoid these difficulties, many efforts have focused on establishing stable, kinetically trapped oligomers that enable in-depth characterisation, while mimicking oligomers that are produced during protein aggregation [37, 113]. Such models have been developed for several protein aggregation systems, including the $\alpha$-synuclein, the $\alpha$-synuclein ([37, 31]). In order to avoid these difficulties, many efforts have focused on establishing stable, kinetically trapped oligomers that enable in-depth characterisation, while mimicking oligomers that are produced during protein aggregation [37, 113]. Such models have been developed for several protein aggregation systems, including the $\alpha$-synuclein, the $\alpha$-synuclein. Preparation of kinetically trapped oligomers generally involves the use of stabilising agents such as metal ions and small molecules, or lyophilisation [113, 114, 14]. In the case of $\alpha$-synuclein, the most highly characterised oligomers are formed through lyophilisation, although the mechanism by which freeze-drying induces oligomer formation is unclear. While variations on the intermediate steps have been used, the general preparation method involves lyophilisation of the protein, followed by resuspension and incubation (with or without shaking) at high (hundreds of $\mu$M) concentrations. Following removal of monomeric and large fibrillar species, the resulting oligomers are highly stable, being resistant to dissociation even on the order of weeks, and comprise a fairly homogeneous population [115, 37]. Moreover, the FRET efficiencies and cellular toxicities of species formed in this way are almost identical to those of the toxic Type B oligomers identified through smFRET, and have thus been named Type B* (Figure 2.4) [36, 37]. In addition, DNA aptamers raised against lyophilised oligomers are able to bind oligomers generated in aggregation reactions and are now heavily used in super-resolution imaging of $\alpha$-synuclein aggregates, suggesting conserved structural features (Figure 2.4) [116, 117, 118]. Taken together, these findings indicate that kinetically trapped oligomers produced by lyophilisation provide an appropriate model system for oligomers that are formed during aggregation.

Accordingly, lyophilised oligomers have been used to determine the potential structures that $\alpha$-synuclein oligomers can form. Through hydrogen-deuterium exchange mass spectrometry (HDX-MS), two major subpopulations were identified [112]. Two subpopulations were also identified in oligomers formed by a similar preparation method, by analytical ultracentrifugation (AUC) and cryo-electron microscopy (cryo-EM) measurements, which may correspond to those identified by HDX-MS (Figure 2.5) [37]. Small angle X-ray scattering (SAXS) has also been employed to obtain morphological information on stable $\alpha$-synuclein oligomers, revealing elongated rigid core structures with 18.8 nm height and 9.4 nm width, surrounded by a $\sim$5 nm
2.3. Approach

**Figure 2.4:** Conservation of structural features between oligomers formed through aggregation and lyophilisation. (a-c) DNA aptamers raised against α-synuclein oligomers formed through lyophilisation also recognise oligomers formed during aggregation reactions. Panels reproduced from [117]. (a) Schematic of multiple DNA aptamers binding to one α-synuclein aggregate. The fluorescently labelled imaging strand binds transiently to the docking strand on the aptamer, facilitating super-resolution imaging. (b) Time series (0.5 s between images) of imaging the same aggregate, with time progressing left to right, and then top to bottom. Scale bar = 1 µm. (c) Example images of an α-synuclein oligomer (top) and fibril (bottom) obtained through diffraction limited microscopy using ThT as the fluorescent probe (left) and by aptamer-based super-resolution microscopy (right). Scale bar = 500 nm. (d-e) Biophysical characterisation of oligomers prepared by lyophilisation and type B oligomers from aggregation reactions. Panels reproduced from [37]. (d) Frequency distributions of FRET efficiencies for lyophilised oligomers (grey bars) compared to those obtained for type A and B oligomers in aggregation mixtures, determined by smFRET. (e) Estimated mass distributions of lyophilised oligomers (blue) and oligomers in aggregation reactions (grey), based on smFRET intensities.
stable oligomers of α-synuclein have also been formed by the addition of small molecules. While lyophilisation seems to produce toxic oligomers, the properties of small molecule-stabilised oligomers are more variable. Oligomers produced by incubation of monomeric α-synuclein with 3,4-dihydroxyphenyl-acetaldehyde (DOPAL), a metabolite of dopamine, were able to induce mitochondrial dysfunction [121]. In contrast, oligomers stabilised by epigallocatechin gallate (EGCG) do not affect cell viability and have thus been named Type A*, for their similarity to the Type A oligomers observed by smFRET [36, 64, 122].

In addition to determining structural properties of α-synuclein oligomers, such kinetically trapped species have been used to understand how oligomers induce toxicity through membrane permeabilisation [58, 64]. The interactions of non-toxic EGCG-stabilised Type A* oligomers and toxic Type B* oligomers with membranes was investigated by solid-state nuclear magnetic resonance (ssNMR). Through dye leakage assays, Type A* oligomers were found not to induce membrane disruption, while Type B* oligomers readily permeabilised lipid membranes. ssNMR measurements revealed that, while both oligomer types interacted with membranes mimicking the composition of synaptic vesicles, only the Type B* oligomers were able to insert into the hydrophobic core. This differential membrane disruption ability seems to be due to the availability of the N-terminus to act as a membrane anchor, and the presence of a rudimentary β-sheet core structure that can insert into the membrane, thereby providing a molecular level rationalisation of the differential membrane permeabilisation ability of the two oligomer types (Figure 2.6) [64, 123].

Furthermore, kinetically trapped models have been invaluable in determining structure-toxicity relationships, particularly with the use of the HypF-N system. The N terminal domain of the *E. coli* Hyp-F protein is frequently used as a model protein aggregation system, most notably in oligomer studies [43, 124, 125]. Through modulating the environmental conditions for oligomer formation, HypF-N reliably forms two different oligomer structures at high yields. These oligomer subtypes are denoted Type A and Type B, with the former exhibiting toxicity towards cells, and Type B being inert in this regard [44, 126, 127]. Through introducing mutations, the surface hydrophobicity and size of the oligomers, probed by 1-anilino-8-naphthalene sulphate (ANS) dye fluorescence and static light scattering (SLS), respectively, could be modulated. By relating these properties to cell viability measurements, it was...
Figure 2.5: Structures of α-synuclein oligomers. (a) Structure of α-synuclein oligomers formed during aggregation, determined by small angle X-ray scattering (SAXS). Figure reproduced from [119]. (b) Dimensions of purified stable oligomers obtained through lyophilisation, determined by SAXS. The structured core is approximately ellipsoidal, with axes of length 4.7 and 9.4 nm, surrounded by an unstructured region which is roughly 5 nm thick. Figure reproduced from [120]. (c) Cryo-EM derived structures of two subpopulations in a kinetically trapped oligomer sample of α-synuclein, generated through lyophilisation. Both populations are approximately cylindrical, with a central cavity with a diameter of 25 Å. Figure adapted from [37].
Figure 2.6: Schematic representations of toxic and non-toxic α-synuclein oligomers interacting with membranes. The non-toxic Type-A* (left) oligomer remains fully disordered in structure upon membrane binding, and is unable to insert in the membrane. In contrast, the N-terminal regions in the toxic Type-B* oligomer (right) fold into α-helices upon membrane binding, allowing the rigid β-sheet oligomer structure to insert into the membrane and thereby disrupt its integrity. Figure reproduced from [64].

determined that small size and high surface hydrophobicity are key determinants of cellular toxicity (Figure 2.7) [42].

Such detailed structure-toxicity studies have not been carried out on other oligomer systems. In particular, the role of secondary structure in toxicity has only received limited attention. However, secondary structure polymorphism in fibrils has been studied in detail in the context of strains and toxicity [15, 128, 129]. Although the vast majority of such structures consist predominantly of the canonical amyloid β-sheet structure, recent work has identified α-helical content as a key structural element in fibrils of the most toxic member of the phenol soluble modulin (PSM) family, with toxicity being clearly linked to the secondary structure [130, 131]. In light of the profound effect of small structural changes on the properties of fibrils, the significant intrinsic heterogeneity of oligomer structures may be a key determinant of their aggregation propensity and toxicity [132]. However, it has recently been reported that β-sheet structure plays a minimal role in toxicity [133]. This chapter therefore presents the investigation of the role of oligomer secondary structure as a determinant of toxicity and its connection to disease-related mutations of α-synuclein.
2.3. Approach

Figure 2.7: Structural determinants of HypF-N oligomer toxicity. (a) Schematic showing the effects of size and surface hydrophobicity on oligomer toxicity. (b) Relationship of cell viability determined by MTT assay, with the maximum wavelength emission of ANS, an indicator of surface hydrophobicity, and Rayleigh ratio, which correlates with oligomer size. (c) Zoomed in view of the surface in b. Figure adapted from [42].
2.4 Results and discussion

2.4.1 All α-synuclein variants form oligomers with similar size and morphology.

Oligomers from the familial PD associated α-synuclein variants were successfully generated using our previously described protocols [37]. Transmission electron microscopy (TEM) images showed that these variant oligomers have similar size and morphology to the WT oligomers, being approximately spherical with a diameter of around 5-15 nm (Fig. 2.8). The variants were additionally confirmed to form fibrillar species, with very different morphology from the oligomers (Figure 2.9).

\[\text{Figure 2.8: All } \alpha\text{-synuclein variants form oligomers with similar size and morphology: TEM images of variant oligomers, confirming that relatively homogeneous oligomer populations are produced in all cases, with roughly spherical shape and 5-15 nm diameter (scale bar = 100 nm).}\]
2.4. Results and discussion

Figure 2.9: TEM images of variant α-synuclein fibrils, all showing similar morphology. Scale bar = 500 nm

The sizes of the oligomers were measured using dynamic light scattering (DLS), with oligomers having a major peak around 20 nm diameter, in contrast to the monomeric peak at 4 nm (Figure 2.10). The higher size estimate obtained from DLS compared to TEM is likely due to the effects of solvation and bias towards larger species in DLS [134]. Previous characterisation of the WT oligomers showed the presence of subpopulations with different sizes, which would also increase the apparent size according to DLS, given the scaling of scattered intensity with the radius to the power of six [37].

Further investigation of the sizes of the oligomers demonstrated clear differences in the size distributions of the variants, demonstrated by native polyacrylamide gel electrophoresis (native-PAGE) (Figure 2.11). Sedimentation velocity analysis by analytical ultracentrifugation (AUC) confirmed the rough distributions afforded by native-PAGE, and again demonstrated conserved stable sizes for such oligomers (Figure 2.12). All oligomers, except for G51D, contained oligomer populations at both 10S and 15S, while A53T contained an additional populations at 19S, and a lowly populate species at 24S. In all cases, the size distributions obtained are very well-defined, indicating that only specific sizes are stabilised in each variant.
Chapter 2. Structural characterisation of kinetically trapped α-synuclein oligomers

Figure 2.10: All variant oligomers exhibit the same size distribution profile by DLS. DLS profiles of all variant oligomers (solid line) are shown alongside the profiles for their corresponding monomeric species (dotted line).

Figure 2.11: Native-PAGE analysis of variant α-synuclein oligomers. All variants form oligomers in the same size range, with all variants except for the G51D forming multiple subpopulations.
Figure 2.12: Sedimentation velocity analysis of variant α-synuclein oligomers. The size distribution profiles are consistent with the Native-PAGE analysis. In addition to the monomer peak around 1 S, all variants except for the G51D display major oligomer populations at 10 S and 15 S. Representative profiles are shown (n ≥ 3).
2.4.2 Tyrosine environments indicate structural differences between variants

From the native-PAGE analysis, we also observed that the G51D band was less dense than the other variants, indicating that the protein mass loaded was lower than for the WT. Since all the oligomer concentrations were determined by assuming a molar extinction of 5600 M$^{-1}$ cm$^{-1}$, that of the monomer, this result suggested that the molar extinction coefficient, a reporter of the light absorbance properties, is not constant between the variant oligomers. Using amino acid analysis coupled with BCA and UV-vis absorbance analyses, we found that A30P and G51D oligomers had a higher molar extinction coefficient (approximately 12,444 M$^{-1}$ cm$^{-1}$) than the WT and other variant oligomers (approximately 7000 M$^{-1}$ cm$^{-1}$) (Figure 2.13). These values were additionally confirmed by dissociating WT and G51D oligomer samples by the addition of guanidinium thiocyanate, as representative variants for the two oligomer extinction coefficient classes. The dissociated samples were then analysed by SDS-PAGE, finding that the monomer band densities were the same, thereby validating the determined molar extinction coefficients (Figure 2.13).

We further analysed the spectral properties of the species, finding marked differences between their intrinsic fluorescence properties. The monomeric proteins display a maximum fluorescence emission at 305 nm, typical of tyrosine (Figure 2.14). However, all oligomers display a maximum emission peak around 345 nm, which may be due to the formation of tyrosinate in the excited state [135, 136]. The relative intensity of this 345 nm peak to the tyrosine emission peak at 305 nm is much greater in A30P and G51D oligomers than in the WT oligomers, suggesting a stronger stabilization of the tyrosinate form in the excited state in the A30P and G51D oligomers, which explains the increased extinction coefficient observed for A30P and G51D oligomers [135, 136] (Figure 2.14).

α-Synuclein contains four tyrosine residues, located at positions 39, 125, 133, and 136. The latter three are unlikely to be located in the oligomer core; previous evidence suggests that the oligomer core is primarily composed of the N-terminal and NAC regions [37, 64, 123]. Y39 is thus the only tyrosine residue that is likely to be in a different local environment in the monomeric and oligomeric species, so we formed oligomers from the Y39F variant to investigate the role of the Y39 residue in the spectral properties of the oligomers. Since phenylalanine lacks the
2.4. Results and discussion

**Figure 2.13:** Determination of the molar extinction coefficients of variant α-synuclein oligomers. (a) Estimation of molar extinction coefficients at 275 nm by comparison of UV-vis absorbance spectra with protein concentration by BCA assay. (b) Estimation of molar extinction coefficients of the WT and G51D oligomers by amino acid analysis, in combination with UV-vis spectroscopy. In both a and b, error bars are shown for standard deviation, with individual repeats indicated by points. (c) SDS-PAGE analysis of WT and G51D oligomers, following incubation with guanidinium thiocyanate, at oligomer concentrations determined using the molar extinction coefficient values determined in a and b.

**Figure 2.14:** Intrinsic fluorescence properties of α-synuclein variants. Intrinsic fluorescence emission (276 nm excitation: solid lines) and excitation (305 nm emission: dotted lines) spectra of variant α-synuclein monomers (left) and oligomers (right).
Chapter 2. Structural characterisation of kinetically trapped α-synuclein oligomers

**Figure 2.15:** Characterisation of Y39F α-synuclein oligomers. (a) α-Synuclein sequence, with the N-terminal region highlighted in pink, NAC region in blue, and C-terminal region in yellow. The tyrosine residues are shown in bold, and the Y39F mutation highlighted in dark blue. (b) Chemical structures of tyrosine and phenylalanine; the only difference in structure is the presence/absence of the hydroxyl group. (c) Intrinsic fluorescence emission (276 nm excitation: solid lines) and excitation (305 nm emission: dotted lines) spectra of WT and Y39F monomers. (d) Intrinsic fluorescence emission (276 nm excitation: solid lines) and excitation (305 nm emission: dotted lines) spectra of WT and Y39F oligomers. (e) CD spectra of Y39F (pink) and WT (black) oligomers, showing similar degrees of the expected β-sheet structure of oligomers [37].

Hydroxyl group that differentiates tyrosine, the intrinsic fluorescence of the Y39F oligomer should be unchanged from that of the monomer, if Y39 is indeed the source of the differences observed for the familial variants. The Y39F oligomer secondary structure was determined by CD spectroscopy, to confirm that no major structural changes were induced by the mutation. The intrinsic fluorescence emission spectra of the Y39F monomers and oligomers display very few differences, supporting our hypothesis that Y39 is indeed the residue that is involved in the spectral changes displayed by the oligomers (Figure 2.15).

### 2.4.3 G51D oligomers display varying degrees of α-helical structure

Analysis of the oligomers by circular dichroism (CD) spectroscopy revealed that all the oligomers contain β-sheet structure, intermediate between that of their respective
disordered monomeric and β-sheet rich fibrillar states (Fig. 2.16). Remarkably, for the G51D oligomer preparations, we observed differing amounts of additional α-helical content.

Figure 2.16: Representative CD spectra of variant α-synuclein monomers (dotted) (10 µM), oligomers (solid) (8 µM WT, E46K, H50Q, and A53T, 4.5 µM A30P and G51D), and fibrils (dashed) (10 µM).

However, these G51D oligomers with a wide range of α-helical structural content were all prepared under seemingly identical conditions. Since the monomers from previous oligomer preparations were reused to produce more oligomers, we first explored whether the structure of the previous preparation dictated the structure of the following preparation. Tracking oligomer ‘lineages’ in this way, no link was found between successive oligomer preparation rounds, indicating that environmental conditions during the preparation process are a more likely source of structural determination (Figure 2.17). Moreover, these results indicate that the variation in oligomer structures is not due to the modification of the monomer, for example by oxidation or strong interactions with contaminants such as metal ions or detergents, since this would result in propagation of the structure through successive preparations.

The oligomers are formed during the lyophilisation step of the preparation procedure, so this was the step to which we turned our attention. Protein solutions were always lyophilised after flash-freezing in liquid N$_2$ in a narrow range of concentrations (140-250 µM) and volumes (2-3 mL), in the same tubes, making lyophilisation
Chapter 2. Structural characterisation of kinetically trapped α-synuclein oligomers

Figure 2.17: Oligomer structure is not propagated through successive preparations. Two representative experiments are shown (left and right panels). CD spectra of oligomers (dashed) produced from the remaining monomeric flow-throughs do not necessarily overlap with those from the previous oligomer preparation (solid).

concentration, volume, and surface area unlikely sources of the structural variations. Although almost all oligomers were prepared using the same lyophiliser, the lyophiliser vacuum pump strength was not consistent, so protein aliquots from the same preparation were divided between three different lyophilisers, with vacuum pump pressures ranging between 0.03 and 0.7 mBar. No clear relationship between the lyophilisation pressure and oligomer structure could be determined (Figure 2.18). Moreover, with the exception of these samples prepared at 0.7 mBar, all oligomers were formed using pump pressures $\geq \ 0.3$ mBar, and formed a wide range of structures, demonstrating that pump pressure is not sufficient to explain the structural variations.
Figure 2.18: Vacuum pump pressure during lyophilisation does not determine oligomer structure. CD spectra of G51D oligomers produced using three different lyophilisers with different vacuum pump pressures, finding no clear correlation between secondary structure and pressure. All oligomers were subsequently prepared using pump pressures $\geq 0.3$ mBar.

However, since oligomer formation is induced during the lyophilisation step, it seems reasonable that the oligomer structure is also determined simultaneously. Preparing oligomers on different occasions from aliquots that were lyophilised together (and stored at -20 °C for different lengths of time) resulted in oligomers with identical structures, thus confirming that oligomer structure is determined during lyophilisation (Figure 2.19). However, the exact relationship between lyophilisation conditions and structure could not be determined. The structure of every subsequent G51D oligomer preparation was therefore recorded, in order to be able to relate secondary structure with the other biophysical aspects under investigation. Despite the large-scale structural differences, the highly $\alpha$-helix-enriched (>20%) oligomers did not display any differences in size from those with minimal (<15%) $\alpha$-helical content (Figure 2.20).
Chapter 2. Structural characterisation of kinetically trapped α-synuclein oligomers

Figure 2.19: Oligomer structure is determined during lyophilisation. CD spectra of G51D oligomers formed from three batches (red, blue, and black) of lyophilised protein, following different storage times at -20 °C (solid and dashed).

![Figure 2.19](image)

Figure 2.20: Secondary structure does not alter G51D oligomer size. (a) Native-PAGE analysis of G51D oligomers with differing degrees of α-helical structure. (b) TEM image of G51D oligomers with >20% α-helical content (scale bar = 100 nm).

![Figure 2.20](image)

We next used FTIR spectroscopy to probe the β-sheet content of all the variant oligomers in more detail (Figure 2.21). All variant oligomers, including structural variants of G51D oligomers, contained antiparallel β-sheet structure, evidenced by an absorbance peak around 1695 cm⁻¹, in contrast to the predominantly parallel β-sheet structure of the fibrils. This is in line with previous characterisation of such...
2.4. Results and discussion

Figure 2.21: FTIR spectra of variant α-synuclein monomers (dotted), oligomers (solid), and fibrils (dashed). The monomeric spectra indicate the sole presence of random coil structure in all variants, while the oligomer and fibril spectra show the presence of β-sheet structure (1630 cm⁻¹). In all cases, the oligomers display a peak around 1695 cm⁻¹, indicative of antiparallel β-sheet structure, which notably is not present in the fibril spectra.

Kinetically trapped oligomers, as the large degree of structural conversion required to form parallel β-sheet structure from antiparallel likely contributes to the high stability of these oligomers [37].

Further structural analysis of the G51D oligomers using FTIR spectroscopy confirmed that all G51D preparations possessed significant β-sheet content, despite differences in the extent of α-helical content [37]. These results demonstrate that the β-sheet structural element, including the anti-parallel structure, is largely conserved in G51D oligomers, indicating that the α-helical structure is likely to arise from regions that remain disordered in the WT and other variant oligomers (Fig 2.22).
Chapter 2. Structural characterisation of kinetically trapped α-synuclein oligomers

Figure 2.22: FTIR (left) and CD (right) spectra of G51D oligomers with varying secondary structure (solid and dashed). Despite differences in the α-helical content, the β-sheet structure is still preserved.

2.4.4 G51D oligomers display high stability towards urea-induced dissociation

α-Synuclein oligomers produced through lyophilisation have previously been found to be extremely stable, both with respect to time and denaturant-induced dissociation [37, 115]. We thus set out to determine how the mutations affect oligomer stability. The stabilities of the variant oligomers were monitored by CD spectroscopy, making use of the clear structural differences between monomeric and oligomeric α-synuclein (Figure 2.16). Given the apparent inability of these oligomers to aggregate, the samples were incubated at room temperature with no aggregation into fibrils observed, and their spectra recorded every day [37]. All oligomers displayed extremely high stability towards dissociation, and in most cases, were still largely intact weeks after purification, reaching an apparent equilibrium of around 20% monomer in most cases (Figure 2.23). Nevertheless, in order to maintain accurate concentrations and minimise the risk of structural rearrangements, all experiments were performed within 24 hours of oligomer preparation.
2.4. Results and discussion

![Graphs showing percentage monomer over time for different mutations.

Figure 2.23: Stability of oligomers over time, monitored by CD spectroscopy. The percentage monomer content was estimated by the CD signal at 220 nm, assuming that only oligomers and monomers were present, giving rise to a linear combination of their respective signals.

This high resistance towards dissociation has been reported previously for oligomers prepared through lyophilisation [115, 37]. Furthermore, such oligomers have also been found to be highly resistant towards urea-induced dissociation. In order to reduce sample volumes required, the relative concentrations of oligomer and monomer were instead determined by native-PAGE analysis. The ability of native-PAGE to yield quantitative concentrations based on band density was first confirmed by loading varying concentrations of monomeric α-synuclein, yielding a linear relationship between the quantified density and monomer concentration (Figure 2.24). Oligomers were incubated overnight at room temperature with varying concentrations of urea, and the percentage monomer content determined by gel densitometry analysis (ImageJ) by comparison to a lane loaded with a known mass of monomer (Figure 2.25).

Previous work on the WT oligomers revealed, by AUC, that the 15S oligomer population was more susceptible to urea-induced dissociation than the 10S population [37]. With the exception of the G51D variant, which only exhibits one major subpopulation by native-PAGE, for all variants the smallest species (around 10S,
**Figure 2.24:** Quantification of monomer concentrations in urea-induced dissociation reactions. (a) Native gel band density is linearly related to protein concentration, with the relationship being strongest at low monomer concentrations. Gel densitometry analysis (ImageJ) of native-PAGE bands of varying α-synuclein monomer concentrations, alongside the line of best fit (dotted), demonstrating the linear relationship between density and [monomer]. (b) Centrifugation conditions (21,130 rcf, 15 min) are sufficient to separate fibril and monomer. Percentage monomer concentrations of WT fibrils (10 µM) incubated with urea (ON, RT) determined by BCA assay analysis of the supernatant following centrifugation, and native-PAGE analysis of the total reaction mixture. Error bars in both cases are shown for standard deviation.
Figure 2.25: Representative native gels of variant oligomers (10 μM) following incubation (ON, RT) with urea. Monomeric bands were quantified by gel densitometry analysis (ImageJ) and the absolute concentration determined by comparison to reference lanes with known monomer concentration (far right of each gel).
as determined by AUC) was the most stable towards urea. However, the oligomer bands were too diffuse to analyse quantitatively (Figure 2.25). Through extracting the percentage monomer content from the native gels, the relative stabilities of the variant oligomers could be compared (Figure 2.26). While the A30P, E46K, and H50Q oligomers exhibited identical stabilities to the WT oligomers, the A53T oligomers were slightly less stable, and the G51D oligomers displayed markedly higher resistance towards urea-induced dissociation. Given that the 10S oligomer population seems to be the more stable of those observed by sedimentation velocity analysis, the increased sensitivity of the A53T oligomers towards urea could be due to its relatively high percentage of large oligomers (Figure 2.12). Similarly, the altered size distribution and other structural aspects of the G51D oligomers could be behind their increased stability.

However, generation of urea-induced dissociation curves for the variant fibrils revealed a different pattern of stabilities. Fibrils were incubated overnight with urea and centrifuged (21,130 rcf, 15 min) to pellet down fibrils, while monomeric species
2.4. Results and discussion

Figure 2.27: Denaturation of α-synuclein fibrils. (Left) Urea-induced dissociation curves for variant fibrils (colours), with WT data shown in black. (Right) Guanidinium chloride-induced fibril dissociation curves, with data from variants (colours) shown alongside WT data (black). In both cases, fibrils were incubated with urea or GdmCl (ON, RT), and centrifuged (21,130 rcf, 15 min) to remove fibrils. The monomer concentration in the supernatant was quantified by BCA assay. Error bars shown for standard error (n = 3).
remained in the supernatant; the percentage monomer content was thus determined by a bicinchoninic acid (BCA) protein quantification assay. In order to confirm that the centrifugation was sufficient to separate monomers from fibrils, the supernatant concentrations determined by BCA assay were compared to native-PAGE analysis of the whole fibril/monomer mixture, yielding identical results (Figure 2.24). The BCA assay was therefore used for all the variant fibrils, given its higher throughput due to the plate-based format. Unlike their highly stable oligomeric counterparts, the G51D fibrils were of a similar stability to the WT and most other variants (Figure 2.27). In the case of the fibrils, the H50Q fibrils were the most stable towards urea; this difference in trend is perhaps not surprising given the FTIR data showing the large structural rearrangements required to interconvert these oligomers and fibrils (Figure 2.21). Equivalent experiments using guanidinium hydrochloride (GdmCl) instead of urea, to disrupt ionic interactions, yielded the same trends (Figure 2.27). The identical fibril dissociation curves obtained for urea and GdmCl suggest that the fibrils are stabilised by polar interactions and hydrogen bonding, with very little stabilisation from ionic interactions [137].

2.4.5 G51D oligomers display decreased surface hydrophobicity

Oligomer toxicity has been shown to be largely dependent on high surface hydrophobicity, proposed to increase oligomer toxicity by increasing the oligomers’ affinity for the membrane interior via non-specific interactions, thus facilitating membrane disruption and cellular dysfunction [44, 138, 139]. Surface hydrophobicity was therefore investigated by the use of 8-anilinonaphthalene-1-sulphonic acid (ANS) dye, whose fluorescence emission is increased upon binding to hydrophobic regions [42]. Interestingly, whereas A30P, E46K, H50Q and A53T oligomers showed similar solvent-accessible hydrophobicity as the WT oligomers, the oligomers generated by the G51D variant exhibited a significantly reduced hydrophobic surface (Figure 2.28). The marked reduction in surface hydrophobicity of the G51D oligomers relative to the WT oligomers may account for a previous report that G51D oligomers induced less membrane disruption than other PD variants [102].

Given that the surface hydrophobicity of oligomers is believed to be linked with their membrane interaction properties, the membrane interactions of the variants
2.4. Results and discussion

Figure 2.28: Representative ANS fluorescence spectra upon binding oligomers (solid lines), compared to fibrils (dashed line) and ANS in PBS alone (dotted). G51D oligomers display markedly decreased ANS binding. Protein samples (5 µM) were incubated with ANS (250 µM) for 30 minutes, before acquisition of fluorescence emission spectra (n ≥ 3).

were investigated, using small unilamellar vesicles (SUVs) with a vesicle composition of 2:5:3 DOPC:DOPE:DOPS (w/w) [103, 140]. This vesicle size and chemical composition was chosen for their similarities to synaptic vesicles, interactions with which are hypothesised to be crucial for the native function of α-synuclein [74]. Upon binding to membranes, monomeric α-synuclein adopts a high degree of α-helical structure, which can be detected and quantified by CD spectroscopy [70]. Titrating α-synuclein monomers with SUVs, the clear increase in α-helical content was monitored by CD spectroscopy (Figure 2.29) [77]. While the WT, E46K, H50Q, and A53T monomers reached the same maximum α-helical content in the fully bound state, the A30P and G51D monomers reached a plateau state with diminished α-helical content (Figure 2.29). This difference in monomer binding has previously been observed with other membrane compositions, indicating a difference in membrane binding mode induced by these mutations [77, 103]. Furthermore, all mutations lie within the N-terminal region of α-synuclein, the region which forms the membrane-bound α-helices. In order to confirm that this difference in α-helical content is specific to membrane binding, and not a diminished structural propensity intrinsic to the variants, the monomeric variants were all incubated with trifluoroethanol (TFE), a solvent which induces α-helical structure in peptides and proteins [141, 142, 143]. Upon incubation with 60% TFE,
all variants adopted the same degree of α-helical structure, confirming that the structural differences of the membrane-bound A30P and G51D monomers are specific to the membrane-binding, and not inherent to the protein’s propensity for forming α-helical structures (Figure 2.30).

Having established that all variant monomers bind the SUVs, the membrane binding abilities of the variant oligomers was studied. While the monomeric protein undergoes a dramatic structural change on binding membranes, the structural change in the oligomers is much less pronounced [64]. As in the monomers, the oligomers adopt some α-helical structure, but to a much lesser extent. The changes in oligomer structure were therefore monitored by CD spectroscopy (Figure 2.31). All variants gave rise to similar changes in α-helical content. Previous work reported that G51D oligomers have a much reduced membrane disruption ability, relative to the other variants [102]. However, membrane disruption and membrane binding appear to be distinct processes, with a high binding affinity not necessarily being correlated to a strong ability to induce membrane disruption [59].

The degree of α-helical content induced upon membrane binding can be estimated by the decrease in MRE signal at 220 nm. Extracting the changes in MRE\textsubscript{220} values, we can obtain binding curves for the α-synuclein/membrane interactions, where we assume that the change in MRE\textsubscript{220} values is linearly correlated to the concentration.
2.4. Results and discussion

Figure 2.30: CD spectra of monomeric α-synuclein variants (10 µM) in the absence (dotted) and presence (solid) of 60% trifluoroethanol in PBS.

Figure 2.31: CD spectra of oligomeric α-synuclein (4.5 µM A30P and G51D, 8 µM all other variants) binding to SUVs.
of bound α-synuclein (Figure 2.32). The monomeric binding can be fitted well, yielding dissociation constants \( K_d \) for all the variants that lie within each other’s error ranges, and binding stoichiometries in the same order of magnitude (Table 2.1). However, the oligomeric data are too noisy to yield meaningful fits, and can only be considered qualitatively; all oligomers undergo very similar structural changes on binding SUVs, with apparently similar affinities.
2.4. Results and discussion

<table>
<thead>
<tr>
<th>Variant</th>
<th>$K_D$ /µM</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>1.1 ± 0.5</td>
<td>212 ± 15</td>
</tr>
<tr>
<td>A30P</td>
<td>0.59 ± 0.31</td>
<td>278 ± 25</td>
</tr>
<tr>
<td>E46K</td>
<td>0.32 ± 0.11</td>
<td>223 ± 10</td>
</tr>
<tr>
<td>H50Q</td>
<td>0.50 ± 0.17</td>
<td>242 ± 14</td>
</tr>
<tr>
<td>G51D</td>
<td>0.54 ± 0.21</td>
<td>340 ± 21</td>
</tr>
<tr>
<td>A53T</td>
<td>0.30 ± 0.19</td>
<td>239 ± 15</td>
</tr>
</tbody>
</table>

Table 2.1: Binding parameters for monomeric $\alpha$-synuclein with SUVs composed of DOPC:DOPE:DOPS (2:5:3 w/w). L corresponds to the average number of lipid molecules interacting with each $\alpha$-synuclein molecule.

2.4.6 $\alpha$-Synuclein oligomer cytotoxicity is correlated with $\alpha$-helical content

With this set of oligomeric species, we sought to investigate their toxicity towards cells using the MTT test, an indicator of cellular stress [144, 145]. The MTT test provides a readout of cellular metabolic activity, by using the dye 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) which, upon metabolism by mitochondrial reductase, undergoes a colour change from yellow to purple with the formation of formazan. We have previously investigated the toxicity of WT oligomers to SH-SY5Y cells, and the MTT reduction to 83% (standard deviation = 8%) reported here is in good agreement with previous studies [146]. Surprisingly, despite A30P, E46K, H50Q, and A53T variant oligomers displaying apparently identical structural properties to the WT oligomers, none of these four variants caused a significant increase in cellular dysfunction under the experimental conditions used here. In contrast, G51D oligomers had a remarkably increased reduction in MTT activity, indicative of significantly higher toxicity than the WT oligomers (Figure 2.33).

In order to confirm that the high toxicity of the G51D oligomers was not due to increased lipopolysaccharide (LPS) levels in the G51D oligomer samples, the LPS content of the protein stocks was quantified, showing no correlation between MTT reduction and LPS content (Figure 2.33) [147, 148]. The high toxicity of the G51D oligomers is surprising given their diminished surface hydrophobicity; previous studies correlating surface hydrophobicity with cytotoxicity would predict the G51D oligomers to be less toxic.
Chapter 2. Structural characterisation of kinetically trapped α-synuclein oligomers

Figure 2.33: (a) MTT reduction upon incubation (24 h) with 0.3 µM variant α-synuclein oligomers. Due to the low yield of G51D oligomers, high volumes of samples were required for cell treatment, necessitating a separate control with the appropriate volume of PBS (PBS (G51D)). Error bars show the standard error, with individual repeats indicated by points. (b) Quantification of LPS concentrations in α-synuclein stocks.

oligomers to be the least toxic (Figure 2.28) [44, 138, 139, 42]. Furthermore, given that all variant oligomers exhibited the same size range, the previously identified toxicity determinants of small size and high hydrophobicity were clearly not sufficient to explain the dramatically higher cytotoxicity of the G51D oligomers [42].

We additionally observed that the variation in cellular dysfunction between experiments for the G51D oligomer treated cells was higher than any of the other variants. Combined with the observation that the G51D oligomers exhibited different degree of α-helical content depending on the preparation, these data therefore suggest that the increased variance in the measured cell toxicity may be due to this observed structural polymorphism.

In order to explore whether the variation in the α-helical content correlates with changes in cellular dysfunction, we characterized structurally distinct G51D oligomers. By deconvoluting the CD spectra, we were able to estimate the relative secondary structural content of oligomer preparations [149, 150]. These fits reproduced our experimental data with extremely low residuals, indicating that this is a robust method for comparatively analysing our spectra. Deconvolution of CD spectra of
2.4. Results and discussion

Figure 2.34: Fitted CD spectra (blue dotted) of several preparations of G51D oligomers with varying secondary structure, shown alongside experimentally acquired spectra (black solid).

WT oligomers suggests that they contain around 11% \( \alpha \)-helical structure, which was not detected by solid-state NMR analysis [64], indicating that the percentage \( \alpha \)-helical content we report here should only be used as a relative quantification between oligomer samples.

Comparing preparations containing similar \( \beta \)-sheet content but enriched in \( \alpha \)-helical structure (considered to be oligomers with >20% \( \alpha \)-helical structure, determined by deconvolution of CD spectra), we determined that their hydrophobicities by ANS were almost identical (Fig 2.35). Furthermore, the WT and G51D oligomers were detected by the A11 antibody (proposed to bind toxic oligomers) with similar affinities (Fig 2.36) [151]. Combined with the oligomer size analysis (Fig 2.20), we thus found that the secondary structure content is the only significant structural and morphological difference between these G51D oligomer preparations. Upon testing the toxicities of these distinct G51D oligomers, we identified a clear correlation between increased \( \alpha \)-helical content and increased cellular dysfunction (Fig. 2.37). However, no correlation was observed between cell toxicity and level of \( \beta \)-sheet or random coil structures, suggesting that the variations in cellular dysfunction can be solely attributed to the \( \alpha \)-helical content (Fig 2.38). In \( \alpha \)-synuclein oligomers, \( \alpha \)-helical content therefore presents itself as an additional determinant of cytotoxicity.
Chapter 2. Structural characterisation of kinetically trapped α-synuclein oligomers

Figure 2.35: G51D oligomer α-helical content does not influence surface hydrophobicity. ANS emission spectra of WT oligomers (black), G51D oligomers (blue) with varying degrees of α-helical content, and ANS in PBS alone (black dotted).

Figure 2.36: Dot blots of WT and G51D oligomer structural variants. All oligomers gave rise to identical signals when probed with the 211 antibody, which binds to the C-terminus. All oligomers showed the same reactivity towards the A11 antibody, which binds to amyloid oligomers.
2.4. Results and discussion

Figure 2.37: Oligomer toxicity is correlated with their degree of α-helical content. MTT reduction as a function of α-helical content of oligomers as estimated by far-UV CD spectra deconvolution [149, 150]. The data correlation obtained was (linear relationship visualized as a grey dashed line): (fraction viability) = -0.016 \times (percentage \ α-helical content) + 1.06, with a Pearson correlation coefficient of -0.54, an R^2 value of 0.29 and p value of 0.0026.
Figure 2.38: (a) MTT reduction of G51D oligomers, compared to their structural content according to fitted CD spectra. Correlation is only observed with respect to α-helical content. (b) Omitting the possible outlier at 29% α-helical content and almost no MTT reduction in the analysis still results in a significant correlation (dotted line) with a Pearson correlation coefficient of -0.44, $R^2$ value of 0.19, and $p$ value of 0.019.
2.5 Conclusions and future directions

Through a detailed biophysical characterisation of α-synuclein oligomers, we identified a link between α-helical content and cytotoxicity. α-Helical content may be particularly significant in the context of the aggregation and toxicity of α-synuclein, given its propensity to form highly α-helix rich structures upon binding to lipids, which is believed to be the first requirement for triggering oligomer-mediated cell damage [58]. Indeed, α-helical content has previously been observed during the aggregation of several α-synuclein variants [152, 153]. Detailed work on the WT oligomers generated through our methods has identified the mechanistic features of α-synuclein oligomer-induced membrane disruption: first the disordered N-terminal regions of the oligomers were found to act as anchors to the membrane by folding into α-helices, allowing the structured hydrophobic β-sheet core of the oligomer to insert into the interior of the lipid bilayer [64]. The membrane anchoring step, therefore, seems to be critical for the induction of toxicity of α-synuclein oligomers. In our study, despite the lower hydrophobic nature of the G51D oligomers, we observe an enhanced cellular toxicity, which is correlated with an increased extent of α-helical structure, relative to the WT oligomers. This suggests that the pre-formed helical structure in the G51D oligomers facilitates the anchoring of the oligomers to the membranes, allowing for a more efficient insertion of the hydrophobic β-sheet core into the lipid bilayer, thus causing membrane disruption and cellular dysfunction.

Moreover, given the observation of α-helical intermediates in multiple systems such as IAPP (islet amyloid precursor protein), the Aβ peptide associated with Alzheimer’s disease, and recent work on the PSMα3 peptide, one of a family of peptides implicated in Staphylococcus aureus pathogenicity, α-helical content within amyloid structures may have broader implications [130, 131, 154, 155, 156, 157, 158]. Indeed, α-helical structure may play a general role in amyloid associated toxicity: structural polymorphs of amyloid fibrils derived from the PSMα3 peptide have recently been found to show variation in the levels of cellular toxicity, with the α-helical form showing more toxicity towards eukaryotic cells [131, 130]. Furthermore, studies on the PSMα3 peptide indicate that these differences in toxicity may be due to their interactions with lipid membranes, supporting our suggestion that the α-helical content within α-synuclein oligomers promotes toxicity by facilitating membrane binding and thus insertion [154].
By varying α-helical content in otherwise unchanged oligomers, we were able to clearly delineate α-helical content as a determinant of cytotoxicity. Given the observed cytotoxicity of β-sheet protein assemblies with partial α-helical structure in unrelated systems, our findings may hold broad significance for systems that exert their toxicity via interactions of oligomeric protein species with cellular membranes.

In conclusion, our finding that G51D oligomer cytotoxicity is associated with increasing α-helical structure demonstrates that, in this system, the biological activity cannot be explained on the basis of size and hydrophobicity, two structural parameters previously identified as determinants of oligomer toxicity [138, 44, 139, 159, 111]. Rather, we find here that the cytotoxicity of G51D oligomers opposes that predicted by surface hydrophobicity. Our results show that instead, α-helical content is a good predictor of cytotoxicity and thus constitutes an additional structural attribute which correlates with cellular dysfunction. More generally, α-helical structure has been identified in multiple protein aggregation systems, indicating that this structural element may play a key role in amyloid oligomer toxicity; future work could therefore probe the generality of this structural link to toxicity.
Chapter 3

Microfluidic methods for studying oligomers in protein aggregation reactions

The microfluidic free-flow electrophoresis work has been done in collaboration with Dr William Arter and Dr Georg Krainer. Dr William Arter designed the desalting µFFE device and acquired the widefield data presented in the section "Bulk fractionation of aggregation reactions". Dr Georg Krainer and Dr Kadi Liis Saar built the confocal setup, the burst extraction code was written and developed by Dr Georg Krainer, Raphaël Jacquat, and Dr Quentin Peter, and the analysis code was written by Dr William Arter and myself. The second version of the lipid disruption device was designed by Arran Collis under my supervision. This chapter is based, in part, on the following preprint:

doi: 10.1101/2020.03.10.985804

3.1 Summary

Oligomers produced during aggregation reactions pose many challenges to their study and characterisation (cf. Introduction: section 1.2). Such species, henceforth referred to as kinetic oligomers, exist as a highly heterogeneous mixture at extremely low
concentrations, and are additionally often metastable and highly sensitive towards environmental changes. In the previous chapter, we probed oligomer structure-toxicity relationships through the use of stable oligomers, however due to their kinetically trapped state and unknown relationship to aggregation reactions, such model systems cannot yield any insights into oligomer kinetics during aggregation. In this chapter, I present the development of microfluidic methods for characterising oligomers produced during aggregation reactions. Initially, a microfluidic free-flow electrophoresis (µFFE) method was developed for the simultaneous fractionation and characterisation by single molecule spectroscopy of oligomers in aggregation mixtures, under native conditions. Then, a method for monitoring lipid vesicle disruption on-chip was designed, to be used downstream of the initial fractionation by µFFE, to provide functional characterisation of oligomer subpopulations.

3.2 Background and motivation

Due to their presence in extremely low quantities and short-lived nature in aggregation reactions, characterisation of kinetic oligomers is generally challenging to achieve by conventional biophysical techniques, including DLS and AUC [31]. Despite the technical difficulties, some studies have employed bulk spectroscopy techniques including CD and Raman spectroscopy, to monitor structural changes during aggregation, which can be attributed to the formation of oligomeric species [152, 153].

However, while bulk techniques may be able to capture some of the structural changes that occur during oligomerisation, oligomers exist as a highly heterogeneous population [36, 160, 30]. In order to characterise the range of oligomeric species, most studies utilise single molecule level techniques [36, 39, 24]. One such approach is atomic force microscopy (AFM), which has been employed in the characterisation of oligomers formed during the aggregation of a wide range of proteins. AFM can yield images at extremely high resolution, and reveal minute alterations in the morphology and mechanical properties of assembled species [161, 61, 162, 163]. However, obtaining such high resolution data requires the drying of samples, which may introduce artefacts from this process, such as structural and morphological changes in the species under study [164]. More recently, AFM images of aggregating proteins have been obtained in solution, avoiding this potential source of structural
perturbation, but the resolution of this technique in the application to small oligomeric species is still limited [165].

Super-resolution microscopy has also been employed for the study of the biophysical properties of kinetic oligomers. Spectrally-resolved points accumulation for imaging in nanoscale topography (sPAINT) [166, 167, 168] has thus been applied in mapping and quantifying the surface hydrophobicities of α-synuclein and Aβ oligomers. Applying sPAINT to kinetic oligomers at timepoints during the aggregation process revealed a wide range of oligomer hydrophobicities. The degree of variation in surface hydrophobicities was found to decrease at later timepoints, which may have implications for toxicity, given the association of surface hydrophobicity with toxicity [44, 42, 167]. Similarly, total internal reflection fluorescence (TIRF) microscopy has been used to monitor fluorescence anisotropy of ThT bound to aggregates including small oligomeric species, providing an estimation of the degree of ordered β-sheet structure [169].

Another successfully employed fluorescence-based technique is single molecule Förster resonance energy transfer (smFRET), which allows the monitoring of the FRET efficiency between donor and acceptor dyes conjugated to the protein. This intermolecular FRET efficiency provides an estimate of how compact the oligomer structure is; more compact oligomer structures may bring dye molecules from different monomers into closer proximity with each other, thereby leading to a higher FRET efficiency. smFRET studies on α-synuclein have revealed two major populations with different FRET efficiencies, denoted Type A and Type B, with low and high FRET efficiencies, respectively. While Type A oligomers predominate in the early stages of aggregation, Type B oligomers are more prevalent in later timepoints [36, 39, 40]. smFRET has also been used in the study of tau aggregation, allowing the identification of differences in the kinetic oligomer populations induced by mutations linked with familial frontotemporal dementia (fFTD) [170, 171].

Two colour coincidence detection (TCCD) is a related method to smFRET, which also relies on the presence of orthogonally labelled monomers in the same oligomer complex. While smFRET detects interactions between the two dyes, TCCD instead identifies oligomers by the simultaneous detection of photons in both fluorescence channels. TCCD has been applied for the study of the aggregation of the SH3 domain of PI3 kinase, obtaining size distributions for oligomers observed during the aggregation reaction [172, 173]. However, these single molecule fluorescence-based
methods require extreme dilution of samples (up to $10^5$-fold), which may result in dissociation of weakly stable oligomers, and under-sampling of rare species [36, 172].

While such methods enable the characterisation of kinetic oligomers, in order to assess their advanced functional properties, such as toxicity and aggregation seeding capability, a separation step is required. Without a further separation step, previous studies have relied on studying the entire heterogeneous aggregation mixture at timepoints where one population is enriched over another [36]. However, this dependence on large differences in relative subpopulation levels renders rare populations almost impossible to study, and also relies on different species possessing distinct functional properties. Therefore, while this approach enables some coarse-grained causality relationships to be determined, this is not akin to directly relating structure to properties.

Current preparative separation steps are mostly limited to size exclusion chromatography (SEC) and sucrose density gradients [174, 31]. SEC has been used to isolate and quantify oligomers in Aβ42 aggregation, thus confirming kinetic models of oligomer populations [31]. However, it is not clear whether SEC has sufficient resolution to allow differentiation of oligomer subpopulations. Moreover, the technique relies on interactions with an external matrix over the course of minutes, which may perturb oligomer structures, while also masking the presence of short-lived species [36]. Similarly, while sucrose density gradients allow for the separation of oligomer subpopulations, the technique requires samples to be taken out of their native buffer conditions, for significant periods of time, on the order of minutes at minimum [175, 159, 174]. We therefore aimed to develop a method to achieve rapid simultaneous fractionation and quantitative structural characterisation of oligomers under native conditions [176]. We applied these new approaches to investigate the aggregation of α-synuclein, which was chosen based on our ability to form well-defined oligomeric species, and for their ease of comparison to previous detailed studies on kinetic α-synuclein oligomers.

### 3.3 Approach

Our approach combines microfluidic free-flow electrophoresis (μFFE) with single molecule spectroscopy, to simultaneously fractionate and characterise oligomeric species formed during protein aggregation, which can then be directly coupled to
3.3. Approach

downstream functional analyses. In this chapter, I present the development and validation of this method, and its application to studying the processes involved in α-synuclein aggregation.

Microfluidic techniques have been employed in a wide range of applications, particularly in the analysis of biological samples [177, 178]. The low sample quantities make microfluidics an attractive option in cases where sample availability is low [179]. Furthermore, microfluidics allows precise spatial and temporal control over sample components and their mixing or separation. This precise control is in large part due to the behaviour of fluids at such small length-scales [180]. In contrast to the turbulent flow observed under bulk conditions, microfluidic techniques exhibit laminar flow, characterised by the Reynolds number, defined as:

\[ Re = \frac{\rho u L}{\mu} \]  
(3.1)

where \( \rho \) is the density of the fluid, \( u \) is the flow speed, \( L \) is the characteristic length scale of the fluid (corresponding to the hydraulic diameter of the microfluidic channel), and \( \mu \) is the dynamic viscosity of the fluid. At Reynolds numbers below 2000, flow enters the laminar regime, whereby inertial forces are low relative to viscous forces, leading to highly predictable flow patterns [181]. This property of laminar flow has enabled the use of microfluidic devices in applications such as diffusional sizing (discussed in more detail in Chapter 4, "Optimisation of experimental design") and separation techniques [182, 183, 184, 28].

Microfluidic separation techniques rely on the slow diffusion of molecules relative to the degree of separation induced. A range of methods inducing differential movement of molecules according to their properties, thereby leading to separation, have been developed, including electrophoresis, magnetophoresis, and acoustophoresis [185]. This chapter focuses primarily on the electrophoretic separation of protein aggregation mixtures. Several microfluidic solution-based separation methods employing electrophoresis have been developed, in two major modes, namely capillary electrophoresis (CE) and free-flow electrophoresis [186]. Such methods allow for the separation of species based on their electrophoretic mobility (\( \mu \)), which is determined by their hydrodynamic radius and charge, according to the following equation:

\[ \mu = \frac{q}{6\pi \eta R_H} \]  
(3.2)

81
where \( q \) and \( R_H \) are the charge and hydrodynamic radius of the species, respectively, and \( \eta \) is the viscosity of the medium [184, 181]. In the case of CE, the electric field is applied along a microchannel, such that the distance moved by the species is dependent on their electrophoretic mobility [187]. CE can achieve very high separation resolution, but the volume of sample that can be loaded is highly restricted [188, 189]. However, in the case of FFE, samples are flowed continuously, and the electric field is applied in a direction perpendicular to that of the flow [190, 191]. Given the low abundance of oligomers in aggregation reactions, the throughput of FFE in terms of sample volume is much more appropriate for this application, since a higher sample volume and thus number of oligomers can be studied [31]. FFE thus allows us to acquire fluorescence microscopy images or single molecule data for arbitrarily long times, and thus obtain more information on rare, lowly populated oligomeric species.

However, one of the limitations of the FFE method lies in its incompatibility with samples containing high salt concentrations (>10 mM), as the presence of ions diminishes the electric field imposed on the sample, thus resulting in little deflection and separation. Aggregation reactions have been found to be highly dependent on buffer conditions, with salt content additionally affecting oligomer structure [16, 192]. In order to overcome this challenge and study aggregation reactions under physiologically relevant conditions, a desalting module prior to the electrophoresis chamber was introduced to the device. This feature exploits the

Figure 3.1: Microfluidic electrophoresis methods: general principles of capillary electrophoresis (CE, a) and free-flow electrophoresis (FFE, b) devices. In CE, the electric field is applied in the same direction as sample flow, whereas in FFE, the field is applied perpendicular to the flow direction. Figure adapted from [181].
difference in sizes, and thus rates of diffusion, between the protein species of interest and salt ions, in order to selectively allow salt ions to diffuse out into a waste stream while larger species, including oligomers, are retained for analysis. This module is able to reduce the salt content by up to 20× in 1 second, thus making buffers that mimic physiological conditions accessible. Furthermore, unlike manual dilution into low salt buffer or water, this microfluidic approach means that samples are only taken out of their native conditions for one second prior to analysis, a timescale which should not significantly affect oligomer structure.

3.4 Results and discussion

3.4.1 Bulk fractionation of aggregation reactions

µFFE is able to resolve monomeric and oligomeric α-synuclein species

In order to detect the deflection of α-synuclein species, α-synuclein monomers were labelled with a fluorescent dye in a 1:1 molar ratio, by use of a thiol-maleimide conjugation, so that fluorescence intensity could be used to quantify protein concentration. The N122 residue was chosen to be the site of the cysteine mutation for the dye conjugation, due to its location outside the WT fibril core. It was therefore hypothesised to perturb aggregation kinetics and aggregate structures less severely than the A90C variant (Figure 3.2), which has been previously characterised in smFRET studies [36, 19, 21, 20, 22]. Furthermore, its distance from the C-terminus ensured that any truncation or degradation events would not separate the dye from the aggregation-competent NAC region (Figure 3.2).
Chapter 3. Microfluidic methods for studying oligomers in protein aggregation reactions

Figure 3.2: Cysteine location and conjugated dye selection. (a) Primary sequence of WT α-synuclein, with the N-terminal region shown in pink, amyloidogenic NAC region in blue, and disordered C-terminal region in orange. Locations of the A90 and N122 residues in the NAC and C-terminal regions, respectively, are highlighted. (b and c) Structures of the AlexaFluor-488 and AlexaFluor-546 dyes, showing similar charge distributions. (d) Cross-section of the structured core of α-synuclein fibrils, determined by cryo-EM, with the A90 residue location indicated by arrows. The N122 residue is located outside the fibril core, and its mutation and dye conjugation are thus expected to affect aggregation to a lesser extent than the A90. Panel adapted from [21].

Following the preparation of fluorescently labelled monomers, we sought to test whether the FFE device allows us to resolve monomeric and oligomeric α-synuclein, by forming kinetically trapped oligomers of the N122C variant labelled with AlexaFluor-488 (Alexa-488) in PBS. The mutation and presence of the dye, with the dye-conjugated protein denoted N122C488, did not significantly perturb oligomer structure (Figure 3.3) [37, 193].
3.4. Results and discussion

**Figure 3.3:** Biophysical characterisation of Alexa-488-labelled N122C (N122C488) oligomers (top row), in comparison to unlabelled WT oligomers (bottom row). CD (a) and FTIR (b) spectra of N122C488 oligomers, showing no difference in secondary structure from WT oligomers; N122C488 oligomers display the same extent of antiparallel $\beta$-sheet structure. TEM (c) and Native-PAGE (d) analysis of N122C488 oligomers, exhibiting the same size ranges as WT oligomers [37]. Scale bars = 100 nm.

The N122C488 oligomers were injected onto the desalting electrophoresis device and the electric field applied. In order to ascertain the electrophoretic mobilities of the samples, fluorescence images were taken in the analysis region (Figure 3.4a), and the fluorescence intensity of the Alexa-488-labelled protein across the image quantified (ImageJ). Two major populations were resolved at high field strengths, corresponding to the oligomers and monomer component, which arise through the dissociation of oligomers (Figure 3.4c). The high field strengths necessary to resolve oligomeric species were accessed by the use of liquid electrolyte electrodes [184]. The specific µFFE device design used in this study contains liquid electrolyte electrodes (3 M KCl + 1 µM fluorescein for visualisation), to connect solid electrodes to the electrophoresis sample chamber [184, 194]. Removing the solid electrode-solution interface prevents the formation of bubbles of electrolysis products which perturb flows, at high voltages, thus allowing the application of high voltages.

In order to enable facile assignment of signals without requiring perfect baseline separation between monomer and oligomer, monomers labelled with a complementary fluorescent dye were added to the desalting water inlet ("reference" inlet in Figure 3.4a). By adding the reference monomers to the desalting inlet rather than to the
Chapter 3. Microfluidic methods for studying oligomers in protein aggregation reactions

Figure 3.4: Analysis of kinetically trapped α-synuclein oligomers by µFFE. (a) Desalting µFFE device design, showing the locations of the sample inlets and outlets. The electric field is applied by the contact of the KCl electrolyte with metal electrodes, connected to an external circuit. (b) Schematic of the experiment, and principle of the Alexa-546 signal subtraction. The N122C546 monomer signal is normalised to the N122C488 monomer signal, and then subtracted to yield electropherograms of only the N122C488 oligomeric species. (c) Fluorescence images of the analysis region of kinetically trapped N122C488 oligomers spiked with Alexa-546 monomer, under 0 V and 300 V conditions. (Figure prepared by Dr William Arter.)

sample in the syringe, we minimise its contact time with the Alexa-488 sample under study, thereby minimising the chance of cross-reactivity between the orthogonally labelled samples. AlexaFluor-546 (Alexa-546) was chosen for this purpose, due to its similarity in structure and charge state to Alexa-488, which would not lead to differences in the electrophoretic mobilities of the Alexa-488-labelled and Alexa-546-labelled α-synuclein monomers (Figure 3.2b and c). Importantly, the excitation and emission wavelengths of the Alexa-488 and Alexa-546 do not overlap, so the two dyes can be detected independently from each other. The normalised Alexa-546 signal was subtracted from the Alexa-488 signal, in order to obtain electropherograms containing fluorescence signals arising solely from oligomeric species.
3.4. Results and discussion

The oligomeric species were thus confirmed to have a higher electrophoretic mobility than the monomers, which can be rationalised by considering how radius and charge scale with the number of monomer subunits. An oligomer has mobility \( \mu_o \), proportional to the ratio of its charge to radius \( \frac{q_o}{r_o} \). The oligomer charge can be assumed to scale with its number of monomeric subunits, so that:

\[
\mu_o \propto \frac{q_o}{r_o} \propto \frac{n_m}{r_o} \tag{3.3}
\]

where \( n_m \) is the number of monomers that form the oligomer, and \( \nu \) is the scaling coefficient relating monomer and oligomer net charges, which is expected to be between 0 and 1. Making the approximation that monomers and oligomers have the same mass density, we can express \( n_m \) as:

\[
n_m = \frac{V_o}{V_m} = \frac{r_o^3}{r_m^3} \tag{3.4}
\]

We can thus express \( r_o \) as \( r_m n_m^{\frac{1}{3}} \), and obtain the following:

\[
\mu_o \propto \frac{n_m^{\nu - \frac{1}{3}}}{r_m} \tag{3.5}
\]

The higher mobility of oligomers relative to monomers is thus consistent with physical considerations and, furthermore, oligomer mobility is predicted to increase with the number of monomeric subunits. We next sought to validate this general scaling principle by comparing the electropherograms to sedimentation velocity analyses. The sedimentation coefficient \( (s) \) of a particle is given by:

\[
s = \frac{m}{6\pi \eta r} \tag{3.6}
\]

where \( m \) is the mass of the particle and \( r \) its radius. Since the mass of an oligomer scales linearly with the number of monomers, we have the following relationship:

\[
s \propto \frac{m_o}{r_o} \propto \frac{n_m}{r_o} \tag{3.7}
\]

We previously made the approximation that \( r_o \) can be expressed as \( r_m n_m^{\frac{1}{3}} \), yielding:

\[
s_o \propto \frac{n_m}{r_m n_m^{\frac{1}{3}}} \tag{3.8}
\]
Chapter 3. Microfluidic methods for studying oligomers in protein aggregation reactions

\[ s_o \propto \frac{n_m^2}{r_m} \]  

(3.9)

implying that the AUC sedimentation velocity profiles are expected to scale very similarly to the electropherograms, and should yield similar depictions of oligomer heterogeneity. Through comparing profiles obtained from the two techniques, we observed a strong similarity, confirming that the broad spread of the oligomer signal in the electropherogram is not a result of hydrodynamic broadening, but rather a true indicator of the structural heterogeneity of the sample (Figure 3.5).

Having validated that μFFE facilitates the resolution of oligomeric species, in order to confirm the correct performance of the desalting module, oligomers were prepared in 10 mM sodium phosphate buffer (pH 7.4) and electrophoresed in the absence of this module. Comparing these results with those obtained from oligomers prepared in PBS and injected into the desalting-μFFE device, the degree of separation between monomers and oligomers was almost identical, confirming that the desalting-μFFE device is compatible with physiological buffers such as PBS (Figure 3.5).
3.4. Results and discussion

Figure 3.5: Comparison of µFFE with sedimentation velocity analysis. Electropherograms (a, b) and sedimentation velocity analyses (c, d) of kinetically trapped N122C488 oligomers in PBS (left) and 10 mM sodium phosphate (right). (Panels a and b) prepared by Dr William Arter.)

Evolution of oligomer populations during aggregation

Following our validation that the µFFE approach is suitable in resolving monomers and kinetically trapped oligomers on a timescale of seconds at physiological ionic strengths, we next investigated its applicability towards kinetic oligomers. N122C488 monomer was incubated (100 µM, 37 °C, 200 rpm), and aliquots periodically withdrawn for analysis. Following centrifugation (21,130 rcf, 10 min) to remove large fibrillar species that would block the device, the supernatant was carefully removed. A small portion of the supernatant was retained for concentration determination by UV-vis spectroscopy, to ascertain the degree of aggregation, and the remainder injected onto the device for electrophoretic analysis.
Chapter 3. Microfluidic methods for studying oligomers in protein aggregation reactions

At 0 h of aggregation, no oligomeric species were observed, confirming the purity of the N122C488 monomer (Figure 3.6). Following aggregation initiation, before large aggregated species could be detected by a decrease in the supernatant concentration, oligomeric species could be observed with higher electrophoretic mobilities than the monomers, demonstrating the sensitivity of the method. The total oligomer concentration (in terms of monomer subunits) at timepoints could be estimated by integration of the fluorescence signal, following subtraction of the Alexa-546 monomer signal, to remove the monomer contribution, yielding concentrations in the low micromolar range, consistent with previous work [36].

As the aggregation reaction proceeded, two major populations of oligomers with different mobilities were observed. Initially, only the low mobility population was detected, before the emergence and dominance of the high mobility species. Notably, the range of electrophoretic mobilities observed for the kinetic oligomers was similar to that obtained for the stable oligomers, confirming their suitability as a model oligomer system.

The emergence of two oligomer populations has been previously observed in single molecule FRET (smFRET) experiments, where Type A oligomers with low FRET efficiency being the major species at early aggregation timepoints, while Type B oligomers with high FRET efficiency dominating at late stages of aggregation [36]. These findings could be consistent with our electrophoresis results, since an increase in FRET efficiency is likely to correspond to an increase in oligomer structure compaction. Such an increase in structural density would result in a decrease in hydrodynamic radius, but will not affect charge, thus increasing the electrophoretic mobility. The low and high mobility species identified through µFFE could therefore correspond to the Type A and Type B oligomers, respectively. In order to determine the structural characteristics of the oligomer subpopulations according to electrophoretic mobility, we next combined the µFFE method with single molecule spectroscopy.
3.4. Results and discussion

Figure 3.6: Analysis of α-synuclein aggregation mixtures by µFFE. (a) Fluorescence images of the aggregation mixture supernatant after 24 hours of aggregation, under the application of 400 V. (Scalebar = 300 µm) (b) Quantification of the fluorescence profiles, with a zoomed-in view of the oligomer region shown in the inset. (c) Apparent oligomer concentrations (green, in monomer equivalents) obtained through integration of the electropherograms, shown alongside the fibril concentration (blue), during the aggregation timecourse. The fibril concentration was estimated by quantification of the supernatant concentration by UV-visible spectroscopy, assuming that all fibrillar species remained in the insoluble pellet. (d) Electropherograms of the monomer-subtracted oligomer region, showing evolution of the oligomer populations during the course of the aggregation reaction. (Figure prepared by Dr William Arter.)
Simultaneous fractionation and single molecule characterisation of kinetic oligomers

Integration of µFFE and single molecule spectroscopy

smFRET studies have previously identified a range of oligomer sizes within the Type A and Type B subpopulations [36, 39, 40]. Due to their low concentrations and the inherent sensitivity limit of the fluorescence microscope, the oligomer signal we obtained is very weak (Figure 3.6), which may obscure further subpopulations. Therefore, in order to further resolve heterogeneity within the oligomeric populations, we combined the µFFE method with single molecule spectroscopy.

The implementation of µFFE has several advantages over the original smFRET experiments; in order to access the single molecule regime of all protein species for smFRET, the aggregation mixture must be diluted by up to $10^5$-fold [172, 36]. This high dilution factor may cause dissociation of weakly stable oligomers, and may also result in under-sampling of rare species. Our µFFE approach minimises the need for dilution, by separating the highly abundant monomeric species from the rare oligomer populations, so that instead every protein species in the mixtures, only the oligomers need to be in the single molecule detection regime. In addition, the physical separation enables the collection of oligomer subtypes for further downstream analysis, such as toxicity, which can then be correlated with their electrophoretic mobility.

The µFFE was therefore carried out using confocal microscopy. In contrast to the widefield epifluorescence microscopy used in the initial experiments where a large area of the device is illuminated and imaged, in confocal microscopy only a small volume (the confocal volume) is monitored. Photons are detected as fluorescent particles pass through the confocal volume. In addition to the increased detection sensitivity, confocal microscopy also provides an additional advantage in improving the apparent resolution of electrophoretic separation. By focusing the confocal volume in the centre of the device height, only the central section of the total chamber height is monitored. This thereby reduces the effects of hydrodynamic broadening of deflected sample streams, which results from inconsistent flow rates across the channel height due to interactions of the flowing solution with the device surfaces [195].

The confocal volume was scanned across the electrophoresis chamber by controlled movement of the stage, in a direction perpendicular to that of the sample flow (Figure
3.4. Results and discussion

3.7). Two scanning modes were used: continuous and stepping; in the continuous scan mode, the stage holding the device was moved continuously, to provide a quick overview of the species present in the sample. In stepping mode, the movement was stopped for seconds to minutes at defined points along this scanning path. This longer acquisition time at different position allowed us to obtain more statistically robust populations of oligomers at different electrophoretic mobilities.
Chapter 3. Microfluidic methods for studying oligomers in protein aggregation reactions

Figure 3.7: smFFE experimental overview and confocal microscope setup. (a) Experimental overview: samples containing N122C488 oligomers and monomers are injected into the device, and electrophoresed. The confocal volume is moved across the electrophoresis chamber from the cathode to the anode, in order to obtain electropherograms. (b) Design of the confocal microscope. A laser beam at 488 wavelength is directed into a single-mode optical fibre, and following collimation, reflected by a dichroic mirror and focused onto a spot in the device through a 60x objective. Photons arising from fluorescence were focused onto a detector after passing through a pinhole to remove out-of-focus light and a bandpass filter. (Original figure prepared by Dr Georg Krainer and Dr William Arter.)
smFFE is able to reproduce the results from epifluorescence microscopy

We first applied the single molecule FFE (smFFE) method to Alexa-488-labelled N122C monomers, in order to confirm that the on-chip behaviour observed using the bulk epifluorescence microscopy could be reproduced. Injecting 100 nM monomer into the device so that burst events from individual monomers could be accessed, we achieved very similar mobilities to the widefield experiments (Figure 3.5). The monomer peak shape obtained in the widefield data was also reproduced in the confocal setup, by running 5 μM monomer, confirming that the device behaviour was preserved between the two microscopy methods (Figure 3.8f). The increased apparent signal broadness obtained by smFFE is thus likely to be due to the increased sensitivity of the detection method, and not to inherently different behaviour of the protein between the two setups.

By acquiring data in stepping mode, distributions of photon counts from individual monomers could be determined. The distribution of burst intensities is due to the different paths through the confocal volume that particles can take, since the illumination in the confocal volume is not uniform. Particles that pass through the centre of the volume are excited to a higher degree than those that pass through the edge of the volume, thus resulting in a range of fluorescence emission intensities. The histograms of burst intensities at the lowest and highest mobility positions for the monomeric sample were identical (Figure 3.8), confirming that the large spread of apparent monomer mobilities is due to diffusion of the sample stream, and not caused by the presence of dimers or higher order species.
Chapter 3. Microfluidic methods for studying oligomers in protein aggregation reactions

Figure 3.8: smFFE analysis of monomeric α-synuclein. (a) Continuous acquisition of 100 nM N122C488 monomers in the presence and absence of the applied electric field. (b) Burst event rates (light blue) obtained at different mobilities, acquired in stepping mode (10 s acquisition time per position). Median burst intensities (dark blue line) shown alongside the interquartile ranges of detected burst intensities (dark blue shaded region). No difference in burst intensity is observed between different mobilities. (c, d) Time traces of photons detected at low (c) and high (d) mobility regions. (e) Histogram of burst intensities of all monomeric species at all mobilities. (f) Continuous scan of 5 μM N122C488 monomers, yielding the same peak shape as obtained in the widefield experiments.
Having confirmed that the smFFE data are consistent with the bulk epifluorescence data, we next applied the smFFE approach to kinetically trapped oligomers. While we required N122C546 monomer in the bulk experiments to reliably differentiate between monomeric and oligomeric signals, the single molecule detection approach allows us to differentiate between monomeric and oligomeric signals based on their intensity. Using this classification, we again observed that the oligomers possessed higher electrophoretic mobilities than the monomer. Since each monomer is labelled with one dye molecule, the number of monomers comprising each oligomer can be estimated by comparison of the oligomer burst intensity with the monomer burst intensity, obtained through experiments on the pure monomeric sample (Figure 3.8).

We previously obtained the following relationship for electrophoretic mobility and monomer composition:

\[ \mu_o \propto \frac{n^{\nu-1}}{r_m} \]  

(3.10)

With the caveat that we have made the assumption of constant mass density of all particles, this relationship predicts that the electrophoretic mobility of oligomers increases with mass (number of monomeric subunits). We therefore conducted a step scan to test the validity of this proposed relationship, recording high intensity burst events for periods of 30 seconds at different positions across the electrophoresis chamber. In line with equation 3.10, we observed an increase in the median burst intensities with increasing mobility (Figure 3.9).
Chapter 3. Microfluidic methods for studying oligomers in protein aggregation reactions

Figure 3.9: smFFE analysis of kinetically trapped α-synuclein oligomers. (a) Continuous scans of kinetically trapped N122C488 oligomers in the presence (300 V, dark blue) and absence (light blue) of an applied voltage. (b) Median photons per complex (PPC) (dark blue line) shown alongside the interquartile ranges (dark blue shaded region) of the PPC for each mobility position. The light blue line represents the detected complexes per second. (c) Photon detection time traces for different mobility regions, showing a general increase in PPC at increasing mobility values.
3.4. Results and discussion

Establishment of reproducible α-synuclein aggregation kinetics

Having confirmed that smFFE is able to simultaneously fractionate and characterise stable, kinetically trapped oligomers, we next turned to kinetic oligomers, produced during aggregation reactions. In order to aid the acquisition of oligomer burst events in statistically sound numbers, rather than the qualitative analysis of the epifluorescence microscopy, we first determined a method for aggregating α-synuclein with highly reproducible kinetics. Unlike the aggregation of the Aβ peptide, α-synuclein aggregation has been found to be far more stochastic in nature, and requires shaking, stirring, or fibrillar seeds in order to undergo aggregation initiation [38, 71, 77, 196]. Under conditions of neutral pH, secondary nucleation rates are minimal in α-synuclein; nucleation is thus dominated by primary processes, and a significant proportion of oligomers in the early stages of aggregation are likely to arise through primary nucleation [16, 196]. This means that the oligomer populations are likely to be highly dependent on the aggregation kinetics of the individual aggregation reaction mixture, so in order to meaningfully compare oligomer populations observed in different reactions, the reactions should have the same aggregation kinetics.

Incubating α-synuclein in eppendorf tubes does not allow facile monitoring of aggregation kinetics, as timepoints have to be collected manually and centrifuged to estimate the fibril mass content. Therefore, a plate-based format was preferable, to allow continuous monitoring of multiple aggregation mixtures simultaneously, without manual involvement or perturbation of the aggregation reactions. The standard method of recording aggregation kinetics in plate reader experiments is to monitor the fluorescence emission intensity of thioflavin-T (ThT), a dye whose fluorescence is increased upon binding to β-sheet-rich fibrils [24]. However, the fluorescence of the Alexa-488 dye is not compatible with ThT, so a different method was required, namely fluorescence quenching [36]. The fluorescence emission intensity of the Alexa-488 dye is quenched upon fibril formation by N122C488, which provides a convenient readout of fibril concentration (Figure 3.10).

Aggregating α-synuclein under shaking conditions (100 µM, 37 °C) in the presence of a glass bead yielded highly reproducible fluorescence emission intensities, both at the start and end points of aggregation (Figure 3.10). The exact emission intensity values at any point during the reaction can therefore be compared, without the requirement for normalisation to internal reference fluorescence values. Moreover, the reproducibility between wells and protein preparations was reasonably high,
with most reactions giving rise to half times of around 22-25 h. These fluorescence quenching profiles were normalised to the ranges of intensities, to yield aggregation kinetic curves expressed as the degree of fibril formation.

In order to confirm that the degree of fluorescence quenching is linearly related to the fibril concentration, samples at various timepoints throughout the reaction were withdrawn and centrifuged (21,130 rcf, 10 min) and the supernatant concentration quantified by UV-vis spectroscopy. These supernatant concentrations, consisting of unaggregated species, were in good agreement with those estimated by the fluorescence quenching, thereby validating this method of monitoring aggregation kinetics. In order to maximise reproducibility, only aggregation curves whose fluorescence intensities lay within ~5% of the values obtained from the standard curve in figure 3.10b were used in the smFFE experiments.

**Figure 3.10:** Monitoring α-synuclein aggregation kinetics by fluorescence quenching. (a) Aggregation curves from three different protein preparations (red, blue, and black), monitored by fluorescence quenching of the Alexa-488 dye. (b) Standard normalised aggregation curve, to which aggregation kinetics were compared to determine suitability for use in the smFFE experiments. The fluorescence has been normalised to the minimum and maximum intensities observed during the reaction. (c) Comparison of fibril quantification by fluorescence quenching and supernatant concentration determination by UV-vis spectroscopy, following centrifugation (21,130 rcf, 10 min), with equal concentration readouts shown by the black line.

**Evolution of oligomers in aggregation reactions at single molecule resolution**

The first timepoint we investigated was at 23 h (60% aggregated), which was chosen due to previous studies which found maximal oligomer concentrations to be around
the aggregation half time [36, 39, 40]. This timepoint was posited as a good test case for determining the applicability of smFFE to kinetic oligomers. An aggregation reaction displaying similar fibril formation kinetics to the standard curve was selected and centrifuged (21,130 rcf, 10 min) to remove large fibrillar species. The supernatant was diluted tenfold to a total protein concentration of 4 µM (monomer equivalents), as testing of different dilution factors showed this to be the minimum dilution required to access oligomers in the single molecule regime. The lowest dilution factor possible was required, as this both minimises the perturbations to the aggregation mixture and maximises the number of oligomer burst events detected.

Injecting the aggregation mixture into the device, we observed an extremely high intensity peak, due to the high (4 µM) concentration of protein present. Upon application of the electric field, the separation of monomer and oligomer was achieved (Figure 3.11a). In the continuous scanning mode, oligomers were clearly distinguished from the large monomer peak, due to their higher electrophoretic mobility and high burst intensity. Analysing the photon burst events obtained in stepping mode for these 23 h kinetic oligomers, we again found that median burst intensity increased with mobility, similar to the stable oligomers (Figure 3.11).

Having confirmed the applicability of smFFE for the study of kinetic oligomers, we next characterised several timepoints of α-synuclein aggregation reactions, in order to determine how oligomer populations evolve during the process. Based on the kinetic curves obtained through fluorescence quenching, we selected timepoints that covered the whole range of aggregate concentrations (Figure 3.12). Following extraction of the supernatant and dilution, the total monomer concentration in the kinetic oligomer samples injected into the device was up to ∼5 µM. Since this is much higher than the 100 nM monomer used in the initial studies, we confirmed that the 5 µM concentration did not induce oligomer formation, by the observation of no oligomer bursts in the 5 µM monomer sample (Figure 3.13).
Figure 3.11: smFFE analysis of 23h kinetic α-synuclein oligomers. (a) Continuous scans of the tenfold diluted supernatant after 23 h of aggregation, at 0 V (light blue) and 300 V (dark blue). (b) Events detected per second (light blue) and median photons per event (dark blue), shown alongside the interquartile range of photons per event (dark blue shaded), at different electrophoretic mobilities. (c) Time traces of photons detected at various positions in the channel, showing variations in the oligomer sizes observed at different mobilities. (d) Histograms of oligomer burst intensities detected at the positions shown in c.
3.4. Results and discussion

Figure 3.12: α-Synuclein aggregation timepoints investigated by smFFE, shown alongside the standard aggregation curve.

Continuous scans of the diluted supernatants from the selected timepoints revealed the presence of oligomeric species in all samples, with the exception of the pure monomeric sample (Figure 3.13). These scans suggest that oligomer concentration increases during the early stages of aggregation, reaching a maximum around 23 hours (60% aggregated). Interestingly, soluble oligomeric species are still present after the aggregation "endpoint" plateau has been reached.

We next sought to characterise the kinetic oligomer populations in more detail, by acquiring burst intensity data in stepping mode. Data were acquired for periods of seconds to minutes at various points across the electrophoresis chamber, in order to determine the oligomeric species present with different electrophoretic mobilities. Oligomeric species with higher mobilities than the monomer could be detected by a burst extraction algorithm based on the interphoton detection times and intensity thresholds. However, oligomers with mobilities similar to that of the monomer would not be detectable above the high background signal of the monomer. Nonetheless, given the scaling relationships determined previously (Equation 3.10), oligomers would not be expected to lie in this region. Furthermore, no oligomers were detected at mobilities lower than that of the monomer, providing further evidence of the robustness of this relationship, despite the approximation of constant mass density (Figures 3.13 and 3.14).
Figure 3.13: Continuous scans of aggregation supernatant samples at 300 V. The timepoint and dilution factor are given in the plot labels. In all cases except for the monomeric 0 h sample, high intensity events arising from oligomers are detected at mobilities higher than that of the monomer. No oligomers were detected at lower mobilities than the monomer.
3.4. Results and discussion

**Figure 3.14:** Proportions of total photon intensities detected in kinetic oligomer events. Total photon detection rate (lines) shown against the detection rate of photons arising from oligomer events at each channel position (shaded).
In order to ensure the accurate calibration of the device field efficiency and thereby mobilities for each dataset, an internal reference was used, rather comparing the currents of experiments and the device when filled with KCl, the method used for the widefield epifluorescence experiments. Given the highly reproducible behaviour and high concentrations of monomer present in the kinetic oligomer samples, the monomer deflection distance was used as the reference to calibrate the field strength. In addition, in order to relate the observed burst intensities with oligomer mass, the burst intensities were divided by the mean monomer burst intensity, obtained through scans of pure monomeric samples (Figure 3.8). The monomer units obtained for the oligomers are therefore approximate, and we selected the mean monomer photon count value to eliminate systematic biases about the exact path of particles through the confocal volume.

Integrating data acquired from multiple step scans, totalling several minutes of acquisition time per channel position, to minimise random and time-dependent variations between experiments, we obtained oligomer count rates as a function of electrophoretic mobility (Figure 3.15). Similar to the widefield epifluorescence data, oligomers first formed with a fairly uniform range of mobilities, before evolving into a peaked population centred around a mobility of $-2 \cdot 10^{-8} \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ at 18 h. As the aggregation reaction proceeded, oligomer populations shifted towards higher mobilities. In contrast to their mobilities, the size distributions did not show any clear changes over course of the aggregation reaction, with the exception of the endpoint plateau sample (46 h), in which small oligomers were much more prevalent than larger oligomers (Figure 3.16). Combining the size distribution and electrophoretic mobilities information, we can also visualise the oligomer populations by heatmaps (Figure 3.17). Through these plots, the peak in oligomer concentration at the aggregation half-time is evident.
3.4. Results and discussion

Figure 3.15: Histograms of count rates of kinetic oligomers with different electrophoretic mobilities. The count rates have been normalised according to the dilution factor, so the values represent those of the total, undiluted aggregation mixture. Oligomers initially span a wide range of mobilities at 17 h. At 18 h, oligomers all have fairly low mobilities, and evolve towards higher mobility species.
Figure 3.16: Histograms of count rates of kinetic oligomers with different apparent monomer unit numbers. Interestingly, no clear time dependence is evident for the size distributions.
3.4. Results and discussion

17h 12%

18h 18%

20h 34%

22h 55%

23h 60%

25h 80%

46h 100%
Figure 3.17: Oligomer populations during aggregation. Heatmaps representing histograms of apparent oligomer sizes at different electrophoretic mobilities (main plots), with histograms of frequencies with respect to mobility only (top plots) and to apparent size only (right plots). The aggregation time and percentage of fibril formation by mass are given in each plot label. Initially, a highly heterogeneous population of oligomers is observed at 17 h, whereas a less heterogeneous population in terms of mobility is found at 18 h and later timepoints. Oligomer populations evolve towards higher mobilities during the aggregation timecourse, without a clear change in size distribution by mass.
3.4. Results and discussion

By counting individual complexes that pass through the confocal volume, we were able to obtain an absolute number concentration of oligomers. The burst detection rate \( b_c \) in Hz at each mobility was first corrected by the dilution factor to obtain the total oligomer concentration in the supernatant. Since not all oligomers at a given channel position and mobility value will pass through the confocal volume and be detected, we next calculate the total burst rate \( b_t \) in Hz at each mobility step in the scan by considering the relative areas of the cross sections:

\[
b_t = b_c \cdot \frac{h \cdot d}{\frac{\pi}{4} \cdot z \cdot w} \quad (3.11)
\]

where \( h \) is the height (27 µm) of the device, \( d \) is the distance between scanned step positions, and \( z \) and \( w \) are the semi-major (3 µm) and semi-minor (0.4 µm) axes, respectively, of the confocal volume cross section, assuming ellipsoidal geometry (Figure 3.18) [197].

![Figure 3.18: Schematic of the cross-sectional profile of the µFFE device in the confocal setup, showing the dimensions used to estimate total oligomer concentration based on burst detection rate.](image)

Given that the sample is injected into the device at a flow rate of 10 µL/h, we can obtain the total events per litre of sample \( b_l \):

\[
b_l = \frac{3600 \cdot 10^6 b_t}{10} \quad (3.12)
\]

Converting to molar units in terms of oligomer events \( b_o \):
Chapter 3. Microfluidic methods for studying oligomers in protein aggregation reactions

\[ b_o = \frac{b_t}{6.02 \cdot 10^{23}} \]  

(3.13)

This oligomer concentration was then adjusted to account for the loss of approximately 30% of oligomers in the desalting step, assuming that all oligomeric species have a hydrodynamic radius of at least 10 nm [183, 198].

\[ b_o = \frac{b_o}{0.7} \]  

(3.14)

In order to compare this more easily with previous data, we can convert this to an approximate concentration in terms of monomer equivalents \( (b_m) \):

\[ b_m = b_o \cdot M \]  

(3.15)

Finally, the ratio of the total intensity of protein detected to the known total protein concentration (Figure 3.10) was used to calibrate the calculated oligomer concentrations to account for any protein loss that may have occurred between the aggregation reaction and the chip. The main source of protein loss is likely to be through sticking to surfaces such as the syringe and connecting tubing. We thus obtain estimates of the oligomer concentrations in the timepoints studied, finding a peak in oligomer concentration of 2.5 µM at the aggregation half-time, in a total reaction concentration of 100 µM (2.5%) (Figure 3.19). This concentration is similar to those obtained in the widefield microscopy experiments (peak of 1.8 µM in 100 µM total aggregation mixture: 1.8%). The slight difference may be due to the low signal:noise ratio under the conditions used, and imperfect background subtraction of the Alexa-546 reference signal. These detected oligomer concentrations are also very similar to those obtained through smFRET (peak of 750 nM in 70 µM total aggregation mixture: 1.1%) and SEC (peak of 2.2 µM in 70 µM total aggregation mixture: 3.1%) experiments [36]. While the experimental conditions varied between these measurements (aggregation was carried out in eppendorf tubes for all data except for the smFFE, and the smFRET and SEC data are for the A90C variant with a mixture of Alexa-488 and Alexa-647 dyes), the higher concentrations detectable by smFFE compared to smFRET may be due to the lower dilution factors required, such that oligomers are retained under conditions closer to their native state [192]. smFFE thus enables the simultaneous fractionation and characterisation of kinetic oligomers under native conditions.
Figure 3.19: Oligomer and fibril concentrations (monomer equivalents) during α-synuclein aggregation. Fibril mass fraction was determined by fluorescence quenching, while oligomer mass concentration was estimated by counting oligomeric events observed by smFFE.
3.4.3 Downstream functional characterisation of oligomers

One of the advantages of the µFFE method is its ability to fractionate samples and we have thus shown that it is capable of resolving oligomeric subpopulations within both stable and kinetic oligomer samples. This therefore opens up the possibility of performing downstream analyses on fractionated oligomer samples, in order to relate biophysical properties (mass, electrophoretic mobility) with functional aspects, such as toxicity and seeding capability. These properties may provide insights into disease pathology and spreading, in relation to aggregation mechanisms, and thereby inform on which microscopic processes should be targeted by therapeutic strategies [199, 200, 201]. In order to be compatible with the low volumes processed by µFFE, and to enable efficient coupling of methods, microfluidic assays are thus preferable for such downstream characterisation. Here, I present the development of an on-chip lipid disruption assay.

Methods to study lipid vesicle disruption

As discussed in detail in the previous chapter ("Structural characterisation of kinetically trapped α-synuclein oligomers"), a major feature of oligomer toxicity is believed to be their ability to interact with and disrupt the integrity of lipid membranes [42, 64]. A key factor in understanding aggregation-associated toxicity is elucidating the structural properties of oligomers that confer toxicity, and monitoring their formation during the reaction.

A wide range of methods have been developed for quantifying lipid disruption, most of which rely on the differential fluorescence of a dye, based on the integrity status of the lipid vesicles [202]. The standard method for monitoring disruption of lipid vesicles utilises calcein, a fluorescent dye which self-quenches at concentrations higher than 70 mM [203, 204, 205]. By encapsulating calcein at high self-quenching concentrations within vesicles, lipid disruption can be monitored by the fluorescence increase caused by calcein leakage (Figure 3.20) [203, 206]. The fluorescence signals observed can then be normalised to by the addition of a detergent to lyse all vesicles. Commonly, this assay is carried out on a fluorimeter or platereader, and requires relatively high volumes and concentrations of samples, which are not compatible with the µFFE separation [37, 207].

One recently developed ultrasensitive method has been demonstrated to monitor
3.4. Results and discussion

Figure 3.20: Schematic of lipid disruption detection by calcein release. Calcein is encapsulated inside vesicles at high concentrations, such that its fluorescence is quenched. Upon disruption of the vesicle by oligomers, calcein is released from the vesicles and fluorescence is detected.

calcium ion influx into vesicles by the fluorescence of a calcium-sensitive fluorescent dye, Cal-520 [208]. Individual vesicles are monitored by TIRF microscopy, and their fluorescence normalised to their potential maximum fluorescence intensity by the subsequent addition of the ionophore molecule, ionomycin. However, this method requires the addition of calcium ions to the sample buffer, a change in environment which may alter the protein species present. Several other methods also involve the tethering of vesicles to glass slides, however, a format that is challenging to couple in-line with the microfluidic FFE device [209, 210, 211, 206].

Several microfluidic assays for lipid disruption have been developed to trap and thereby immobilise vesicles, whose fluorescence can thus be monitored [212, 213]. However, due to the low protein concentrations obtained through fractionation by µFFE, the characterisation should ideally occur under continuous flow conditions of both protein and vesicles, in order to acquire data on sufficient numbers of oligomers [212, 213]. Several continuous flow methods for monitoring lipid permeation by small molecules have been developed, but are not easily adaptable for protein samples [214, 215, 216, 217, 218]. One such method relies on intravesicle pH changes induced by the entry of basic compounds, which activate the fluorescence of an encapsulated
pH-sensitive dye [214]. A similar microfluidic continuous flow method relies on the autofluorescence of small drug-like molecules to monitor their ability to enter lipid vesicles (Figure 3.21) [215, 216, 217, 218].
3.4. Results and discussion

Figure 3.21: High-sensitivity methods for monitoring lipid disruption. (a) Ultrasensitive detection of Ca^{2+} influx by the tethering of vesicles to glass slides, through biotin-neutravidin interactions. The fluorescence of the encapsulated calcium-sensitive dye (Cal-520) is quantified by TIRF microscopy. Figure reproduced from [208]. (b) Microfluidic continuous flow method for monitoring the entry of autofluorescent small molecules into vesicles. The fluorescence of vesicles is monitored throughout their retention time on-chip, in order to determine permeation kinetics. Figure reproduced from [215]. (c) Vesicle trapping device: vesicles are trapped based on their size, which does not allow them to pass through constrictions, so that thousands of vesicles can be retained in a single device. Figure reproduced from [213]. (d) Microfluidic valve-based trapping and immobilisation of single giant unilamellar vesicles. Individual vesicles can be easily monitored, but at the cost of throughput; only tens of vesicles can be trapped per device. Figure reproduced from [212].
Figure 3.22: Lipid disruption device design. Vesicles and sample are mixed on-chip, and flowed down a long incubation channel (width 500 µm). Triton-X is added to fully lyse vesicles, to provide a constant normalisation marker. Images are taken at the intersection of the pre- and post-triton-X channels (shown in red).

Development of an on-chip lipid vesicle disruption assay

Since calcein release does not require extrinsic factors such as calcium ions, and does not impose restrictions on the compatibility of samples, such as sample autofluorescence or charge state, calcein was selected for this assay (Figure 3.20). Previous studies on α-synuclein-induced vesicle disruption have identified the presence of negatively charged lipid headgroups as a pre-requisite for oligomer/membrane interactions [219, 220, 221]. In order to develop the assay, 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (POPS) was therefore selected as the model membrane for the vesicles, given its high susceptibility towards α-synuclein oligomer-induced membrane disruption [37]. Incubation times for lipid disruption are usually on the order of minutes to hours, timescales that are much longer than the usual on-chip retention time; the device was therefore designed with a long incubation channel (83 cm length × 400 µm width × 50 µm height) to maximise the protein/vesicle incubation time [220, 37, 208]. In order to provide a normalisation fluorescence intensity, corresponding to the maximum possible fluorescence observed upon complete vesicle disruption, an additional inlet was added for the delivery of the detergent triton-X (Figure 3.22).

Large unilamellar vesicles (LUVs) formed from POPS containing calcein at high concentrations were generated by freeze-thaw cycles and extrusion, followed by desalting columns to remove unencapsulated dye [220, 37]. The integrity and
size of the LUVs was confirmed by dynamic light scattering (DLS), showing a narrowly distributed diameter of 100 nm. The successful encapsulation of calcein was demonstrated by fluorescence spectroscopy, while vesicles were dissociated to micelles upon the addition of 1% triton-X detergent (Figure 3.23). Intact vesicles gave rise to very little fluorescence emission intensity in the calcein fluorescence emission wavelength range. Following addition of 1% triton-X, a high fluorescence signal was detected, indicating the release of calcein from the vesicles (Figure 3.23).

**Figure 3.23:** Encapsulation of calcein in vesicles. The integrity of calcein-filled vesicles was confirmed by DLS (a), and the ability of 1% triton-X to fully dissociate vesicles into micelles. (b) The successful encapsulation of calcein was confirmed by fluorescence emission measurements. Intact vesicles (blue) gave rise to very little fluorescence emission intensity, whereas a large signal was detected following the addition of 1% triton-X detergent (black).

Having confirmed the ability of the calcein-filled LUVs to report on lipid disruption, the device was tested with samples of PBS, kinetically trapped WT α-synuclein oligomers, and triton-X. The device was imaged at the intersection of the pre- and post-triton-X inlet, in order to simultaneously record the degree of disruption induced by the sample and the triton-X (Figure 3.22). When PBS was run as the sample, very little fluorescence was detected before the addition of triton-X, indicating that no disruption was induced (Figure 3.24). The low levels of fluorescence are likely to be free dye that persist after vesicle purification (Figure 3.23). Upon the addition of 5 µM α-synuclein oligomers, the fluorescence of the incubation channel increased relative to that of the triton-X incubation channel. When 5% triton-X was injected into the sample inlet, a further increase was observed, demonstrating that 5 µM oligomer was not sufficient to saturate the vesicles (Figure 3.24).
Chapter 3. Microfluidic methods for studying oligomers in protein aggregation reactions

Figure 3.24: Fluorescence microscopy images of on-chip lipid disruption. When only PBS and LUVs are incubated, very little fluorescence is detected (left). The fluorescence of the sample/LUVs incubation channel relative to that of the triton-X channel increases upon addition of oligomers (centre). When 5% triton-X detergent is run as the sample, the vesicles are fully disrupted, leading to a high signal in the sample incubation channel (right). This signal is higher than obtained by the 5 µM oligomer sample, demonstrating that 5 µM oligomer is not sufficient to reach saturation of the disruption signal.

The dependence of the device on oligomer and LUV concentration was next tested. In order to quantify the degree of disruption, the fluorescence intensity of the sample incubation channel was normalised by the intensity of the triton-X incubation channel. LUV concentrations were tested in a range of 25-100 µM, finding no difference in assay sensitivity between these concentrations (Figure 3.25). In contrast, an approximately linear relationship was found between the relative fluorescences of the incubation channel and the oligomer concentration, with the minimum detectable oligomer concentration being around 1 µM (Figure 3.25). Furthermore, by acquiring images of the incubation channel at multiple positions, corresponding to different incubation times, the kinetics of lipid disruption could be monitored. The disruption kinetics thus provide an additional parameter by which different oligomer populations may be parametrised. However, the error in this quantification is high in the original device design, due to the number of images that need to be acquired and normalised to each other. In order to facilitate the determination of lipid disruption kinetics, the device was redesigned to facilitate the monitoring of multiple incubation timepoints simultaneously (Figure 3.26). In order to increase the sensitivity of the assay, the single molecule confocal microscope will be used instead of the epifluorescence microscope. This may additionally negate the need for the triton-X normalisation incubation, given the digital nature of detection.
3.4. Results and discussion

Figure 3.25: Quantification of on-chip lipid disruption. (a) Relative fluorescence intensities before and after the addition of triton-X for varying concentrations of both oligomer and LUVs. The fluorescence intensities resulting from oligomer-induced vesicle disruption are not affected by the total vesicle concentration used, but do seem to increase approximately linearly with oligomer concentration, with a detection limit of around 1 µM. (b) Vesicle disruption kinetics, acquired by imaging the device at multiple points along the sample/LUVs incubation channel, converted to incubation time by considering the flow rates and device dimensions. (c) Device design showing the incubation channel positions imaged (red stars) to determine vesicle disruption kinetics in b.
Figure 3.26: Redesigned lipid disruption device to monitor kinetics of disruption using confocal microscopy. A linear scan path encompassing several widely-spaced positions along the sample/LUVs incubation channel, in addition to the triton-X incubation channel, now replaces the imaging region in the original device design.
3.5 Conclusions and future directions

In conclusion, this chapter presents the development of microfluidic methods to study intermediate species formed during protein aggregation reactions. The first method discussed combined microfluidic free flow electrophoretic separation with single molecule spectroscopy, to simultaneously fractionate and characterise oligomeric species, while the second method works towards the functional characterisation of oligomeric subpopulations.

By applying smFFE to aggregation mixtures of α-synuclein, we have been able to obtain biophysical characteristics of transient, heterogeneous oligomer populations, and monitor how these evolve during the course of the aggregation reaction. These obtained subpopulations are now being used to generate refined mechanistic models of α-synuclein aggregation. Moreover, by further developing the burst extraction method used here, through the incorporation of Bayesian nonparametric models, the diffusion coefficients and thus hydrodynamic radii of oligomers can be determined simultaneously from the interphoton times. Combined with the burst intensities, this would thereby provide a readout of oligomer mass density, a parameter that is likely to be correlated with the degree of ordered structure in the oligomer.

Although we have so far focused on studying α-synuclein by µFFE, this is not a limitation of the smFFE method. The Aβ peptide in its 42-residue isoform aggregates by the most well-understood mechanism, with highly reproducible aggregation kinetics [27, 38]. In this system, detailed studies on oligomer concentrations during aggregation have revealed that most oligomers dissociate, rather than proceeding to form fibrils [31, 30]. By obtaining detailed characterisation of oligomer populations, this view of Aβ oligomers could be refined further. Additionally, under the neutral pH conditions used in our α-synuclein aggregation reactions, the rate of secondary nucleation is negligible, so all oligomers formed arise through primary nucleation or fibril dissociation [16, 30]. In contrast, Aβ, when aggregated under its standard conditions of low micromolar concentrations in pH 8 sodium phosphate buffer, undergoes both primary and secondary nucleation at significant rates [27, 38]. A key avenue of investigation would then be to determine differences in the properties of oligomers that form via primary and secondary nucleation.

In addition to elucidating the aggregation mechanisms of pure protein, the effects of extrinsic factors, such as chaperones, small molecules, or antibodies, could be
probed [222, 223, 224, 67, 65, 225, 68, 66, 69, 30]. Given that oligomers are believed to be a major toxic species in protein aggregation-associated diseases, understanding how such factors can influence oligomer populations could be extremely useful in determining the relationship of aggregation with disease pathology, and thus in developing potential strategies to alleviate toxicity [223, 31].

Finally, the development of downstream characterisation methods to be applied to the fractionated oligomer populations will provide valuable insights into the functional properties of the oligomer subpopulations and determinants thereof. Such characterisation methods include the lipid disruption assay presented here, once higher sensitivity has been achieved. Additional parameters include the seeding capability of oligomers; self-replication and amplification of aggregates is a key feature of protein aggregation, and may be involved in the spreading of aggregates between cells during disease progression [199, 200, 201]. Further structural characterisation could be achieved through harnessing other imaging methods, such as sPAINT, to determine the surface hydrophobicities of oligomers [166]. The combination of separation and characterisation methods will enable the detailed determination of mechanisms and rates of oligomer formation, conversion, and dissociation in aggregation reactions, and their contributions to toxicity.
Chapter 4

Optimisation of experimental design

This work has been carried out in collaboration with Dr Georg Meisl. Experimental data were obtained by Dr Tom Scheidt (Aβ/antibodies), Ashley Priddey (HLA), Matthias Schneider (HLA and SARS-CoV-2), and Dr Itzel Condado Morales (SARS-CoV-2) in collaboration with Professor Sara Linse (Lund University, Sweden), Dr Vasilis Kosmoliaptsis (University of Cambridge, UK), Professor Adriano Aguzzi (University Hospital Zürich, Switzerland), and Fluidic Analytics (Cambridge, UK).

This chapter is based, in part, on the following publication and preprints:

doi: 10.1038/s41594-020-0505-6


4.1 Summary

In order to maximise the efficiency of experiments, we should only perform those which will be the most informative. Here, I present the development of a method to optimise experimental design through the iterative prediction of which future experiments will be the most informative (Figure 4.1). This approach entails two key advantages, one qualitative and one quantitative. Firstly, obtaining the most information-rich data possible reduces the number of experiments required, thereby reducing both time and sample quantities needed. In the specific experimental cases tested, this reduction is usually by a factor of 3-4. Secondly, the choice of experiments commonly used in the absence of a rigorous mathematical approach is often insufficient to fully constrain the fitted parameters, in particular if multiple parameters are unknown. Choosing the experiments that are able to constrain the data quickly becomes a non-trivial process, and therefore some information is simply not readily accessible without a robust experimental design theory. The measurement points that our approach proposes completely negate that problem and thus enable one to obtain information not otherwise accessible. I demonstrate the validation and use of the prediction approach in three experimental applications: characterising therapeutic antibodies for Alzheimer’s disease, predicting transplant compatibility, and quantifying antibody responses to SARS-CoV-2 infection.
4.2 Background and motivation

In the work described thus far, the experiments have focused on the aggregation of pure α-synuclein, in order to gain insights into the fundamental mechanisms of aggregation. However, there is considerable interest in understanding perturbations to the aggregation process by extrinsic factors, in particular for developing therapeutics that target the aggregation process [65, 68, 66]. Determining the effects of such factors, which include chaperones, antibodies, and small molecules, on kinetic rate constants can be challenging, due to the complexity of the system [67]. Experiments to disentangle these processes generally require acquisition of aggregation data at varying concentrations of monomer, fibril seeds, and inhibitor [65]. These experiments would be much more efficient, with a rigorous approach to optimising experimental design. Furthermore, the degree of certainty to which quantitative effects on rate constants are determined can be low; the approach presented here ensures that the necessary information can be extracted from the data and maximises and quantifies
Chapter 4. Optimisation of experimental design

the degree of certainty [67].

This chapter presents the development of a general method for optimising experimental design, demonstrated through the application to equilibrium binding measurements, as a key biologically relevant process, that can be formulated in a simple mathematical description. Determining the biophysical parameters governing antibody binding, namely the dissociation constant ($K_d$) and stoichiometry of binding, is of particular interest due to their potential in both diagnostic and therapeutic applications [226, 227]. Such measurements are frequently performed on patient samples, where efficiency in both time and sample usage are often paramount [228, 229, 230, 231].

The theory of optimal experimental design has been explored in detail, and has been applied in several contexts in computational work [232, 233, 234]. However, such concepts have not been implemented in large-scale experimental biological work; methods to optimise experiments in the discrimination of biochemical network models have been proposed, but have not been widely adopted in lab-based settings [235, 236, 237]. This is likely to be due to both the high complexity of theory, the significant computational demands of these optimisation methods, and the level of know-how required to apply the methods in practice. In this chapter, I describe the development and implementation of a robust framework for experimental optimisation which works through an iterative predictive approach, based on Bayesian inference and information theory, and that requires minimal user know-how [238]. The efficacy of this method is demonstrated through its application to the characterisation of binary protein interactions, in several model systems.

4.3 Approach

The mathematical framework for the prediction method holds Bayesian inference at its core, which allows us to generate probability distributions in the parameters to be determined (henceforth referred to as unknown parameters), through the use of Bayes’ theorem:

$$P(A|B)P(B) = P(B|A)P(A)$$  \hspace{1cm} (4.1)

In the context of Bayesian inference, the variables A and B are considered to be the vectors of unknown parameters and measured data.
4.3. Approach

\[ P(\text{parameters}|\text{data}) = \frac{P(\text{data}|\text{parameters})P(\text{parameters})}{P(\text{data})} \]  \hspace{1cm} (4.2)

\( P(\text{data}) \), the prior predictive, is not accessible and thus we usually remove it in favour of a normalisation factor. We thus obtain the generic Bayesian inference statement:

\[ P(\text{parameters}|\text{data}) \propto P(\text{data}|\text{parameters})P(\text{parameters}) \]  \hspace{1cm} (4.3)

where \( P(\text{parameters}) \) describes the information we have about the unknown parameters before we take the measurement, known as the prior, since this contains our prior knowledge of the system. \( P(\text{data}|\text{parameters}) \), the probability of the data given the unknown parameters, is the likelihood function, which describes what measurements we would expect based on our chosen model, for a given choice of parameters and level of experimental noise in the measurements. \( P(\text{parameters}|\text{data}) \), the probability of the parameter values given the measured data, is the posterior probability distribution, which describes our knowledge about the unknown parameter values, following the measurements. This is the main quantity of interest and, under particular choices for the likelihood and prior, the maximum of the posterior probability distribution corresponds to the best fit in a least squares approach.

We can thus obtain probability distributions in our unknown parameters through Bayesian inference. The information contained in these distributions can be quantified through the application of information theory. Shannon introduced the information entropy, \( H(\text{A}) \) for a random variable \( \text{A} \) with discrete states \( a_i \), given by equation 4.4.

\[ H(\text{A}) = -\sum_i P(a_i) \ln(P(a_i)) \]  \hspace{1cm} (4.4)

In the context of our measurements, \( \text{A} \) would be the unknown parameter, or vector of unknown parameters. \( H(\text{A}) \) is minimised when \( P(a_i) = 1 \), when we have absolute certainty of the outcome, and is maximised when \( P(a_i) \) are equal for all \( i \), a uniform distribution. The information therefore refers to the ‘surprise’ in the outcome, the amount of new information obtained about the distribution. Here, information therefore denotes the degree of certainty about the fitted values of the unknown parameters. Therefore, the more peaked a distribution, the higher its information content, and the lower its entropy. We therefore seek to minimise the entropy-based information, \( H(\text{A}) \). In order to avoid confusion, we will henceforth
use the term "information" with its information theoretical definition, and the term "certainty" to refer to the degree of certainty about the unknown parameter values.

$H(A)$ from equation 4.4 is only applicable for discrete variables. In the vast majority of biological experiments, the unknown parameters are continuous, such as the dissociation constant of a binding equilibrium, so equation 4.4 must be adapted for continuous variables $A$. Equation 4.5, the differential entropy, has been proposed as the extrapolation of equation 4.4 to continuous variables $A$.

$$h(A) = - \int P(a) \ln(P(a)) \, dx$$

While this extension to continuous distributions does not retain all the properties of the discrete entropy (equation 4.4), it is sufficient for our purposes. The differential entropy is no longer invariant to unit changes, and may also take negative values, since probability density functions can have values greater than one. The limiting density of discrete points, formulated by Jaynes, addresses these losses of functionality. However, for our purposes, where we wish to compare the differential entropies of distributions of parameters under the same supports and units, this correction is not necessary and the differential entropy suffices.

We can thus quantify the information content of our posterior probability distribution by using the differential entropy. Having taken our $n^{th}$ measurement, yielding data point $y_n$, the information $I$ contained in our posterior probability distribution, $P(K|\{y\}_n)$, is therefore given by equation 4.6.

$$h(P(K|\{y\}_n)) = I = - \int P(K|\{y\}_n) \ln(P(K|\{y\}_n)) \, dK$$

Next, we wish to predict which measurement point will yield the maximum certainty posterior. We therefore need to determine the expected information of a given measurement result $y_n$. The expectation value of a continuous probability distribution, $P(Y)$, is given by $E[Y] = \int Y P(Y) \, dY$. This substitution yields equation 4.7.

$$\langle I|X \rangle = - \int \int P(K|\{y\}_n) \ln(P(K|\{y\}_n)) \, dK P(y_n|X, K) \, dy_n$$

where $x$ and $K$ are vectors of the values of the independent variables and unknown parameters, respectively. However, equation 4.7 relies on us knowing the values for $K$. In an experiment to determine $K$, we of course do not have this knowledge.
already. We therefore express \( \langle I|\mathbf{x} \rangle \) using the expectation values of \( \mathbf{K} \), using the posterior obtained from the first \( n-1 \) measurements, \( P(\mathbf{K}|\{y\}_{n-1}) \) as the probability distribution over \( \mathbf{K} \), generating equation 4.8.

\[
\langle I|\mathbf{x} \rangle = -\int \int P(\mathbf{K}|\{y\}_n) \ln(P(\mathbf{K}|\{y\}_n)) d\mathbf{K} \int P(\mathbf{K}|\{y\}_{n-1}) P(y_n|\mathbf{x}, \mathbf{K}) d\mathbf{K} dy_n
\]

(4.8)

Equation 4.8 now only contains distributions that we have access to before we measure the \( n^{th} \) point, and describes the expected information of the posterior probability distribution that results from taking a measurement with the independent variables \( \mathbf{x} \). We note that \( P(y_n|\mathbf{x}, \mathbf{K}) \) is simply the likelihood function. Our aim is to determine \( \mathbf{x} \) such that \( \langle I|\mathbf{x} \rangle \) is minimised.

We must therefore calculate \( \langle I|\mathbf{x} \rangle \) for all \( \mathbf{x} \), and the \( \mathbf{x} \) that minimises \( \langle I|\mathbf{x} \rangle \) is the measurement that should produce the posterior with the highest certainty. For the first measurement, \( P(\mathbf{K}|\{y\}_{n-1}) \), is simply the prior. After each measurement, \( P(\mathbf{K}|\{y\}_{n-1}) \) is updated to be the posterior from all existing measurements, and \( \langle I|\mathbf{x} \rangle \) is computed again to determine the next measurement to be taken. The approach therefore works through the iterative calculation of 'best' measurements.

It is not generally possible to obtain a closed form expression for equation 4.8. \( \langle I|\mathbf{x} \rangle \) was therefore computed numerically using Python. In this chapter, I showcase this software through its application to several experimental systems, and demonstrate the theory behind the \( \mathbf{x} \) values predicted.

4.4 Implementation

4.4.1 Quantification of binary molecular interactions

The determination of the biophysical parameters that govern binary molecular interactions is of intense interest, particularly in biomedical studies [239]. Indeed, the proper functioning of cells and organisms is dependent on the occurrence of appropriate molecular interactions, for example in the control of cytoskeletal protein polymerisation, or in enzyme activity [240, 241, 242]. Such biophysical parameters, which include the binding affinity, quantified by the dissociation constant \( K_d \), and stoichiometry of binding, have attracted particular interest in the context of antibody binding [243, 244]. The determination of the affinity and concentration of antibodies
in patient sera can function as a readout of the immune response to disease, and may therefore contribute to our understanding of disease progression [228, 229, 230, 231]. Moreover, antibodies are also receiving increased attention for their ability to themselves be therapeutic agents [243, 244]. Characterising their binding to therapeutic targets is therefore of extreme importance.

Equilibrium binding measurements are performed in order to determine binding affinities and stoichiometries. Some of the most widely used methods for detecting protein binding interactions involve the immobilisation of one binding partner on a surface, to which varying concentrations of a potential binder are added in solution. The degree of binding can be monitored by several methods, including enzyme-linked immunosorbent assays (ELISA), surface plasmon resonance (SPR), and biolayer interferometry (BLI) (Figure 4.2) [245, 246, 247, 248, 249, 250, 251]. These assays can be extremely sensitive and easily carried out in high volume by parallelisation in well-plate formats, and are therefore employed in situations where throughput is crucial. However, the surface immobilisation of one binding partner presents several major problems for the determination of binding affinity. Firstly, this surface-bound substrate concentration is difficult to determine; its equivalent in-solution concentration is not clear, and binding affinities are thus given as EC50 values, corresponding to the concentration of solution component that gives rise to 50% of the maximum, fully-bound signal. The relation of EC50 values to the $K_d$, is therefore not known, and not necessarily correlated in a trivial manner.

Surface-based measurements that are able to measure association ($k_{on}$) and dissociation ($k_{off}$) rates, such as SPR and BLI, are able to circumvent this challenge, due to the relationship between the quantities:

$$K_d = \frac{k_{off}}{k_{on}}$$ (4.9)

Secondly, the concentration of surface-immobilised reagent cannot easily be modulated. As will be explored in this chapter, in order to achieve both maximal experimental efficiency and certainty, the concentrations of both binding components need to be varied. Both of these problems can potentially be overcome by experimental development and calibration work, but the surface-bound nature of these assays presents a fundamental drawback; the immobilisation of substrate on the surface may introduce avidity effects due to the high local concentration at the surface, and the binding affinities determined are therefore not always representative of the
in-solution interactions that would occur \textit{in vivo} \cite{252}. Additionally, there may be non-specific binding to the surface, or the surface immobilisation may block epitopes \cite{253, 254}.

In order to reliably determine $K_d$ values, we therefore turn our attention to solution-state methods. Furthermore, solution-state methods are generally much lower in throughput than surface-based assays and therefore present a greater need for optimisation methods. Existing solution-state measurements include isothermal titration calorimetry (ITC), fluorescence correlation spectroscopy (FCS), and microscale thermophoresis (MST) \cite{255, 256, 257, 258, 259, 260}. ITC determines heat capacities by monitoring temperature changes induced by binding events, thus enabling determination of the enthalpy and entropy of binding, while FCS and MST monitor changes in the behaviour of a fluorescently labelled protein. More recently, microfluidic diffusional sizing (MDS) has been developed and applied to determine binding affinities, and is used in this chapter to demonstrate the power of the optimisation approach by prediction \cite{28}.

In the MDS method, the rates of diffusion of species, and thus their diffusion coefficients and hydrodynamic radii are determined on a microfluidic chip \cite{183, 182}. One species is labelled with a fluorescent dye, and its apparent diffusion coefficient is decreased upon binding to another species (Figure 4.3). By monitoring the diffusional behaviour of the mixture, we can infer the concentrations of bound and free labelled species. Due to the low sample volumes required, and its flexibility in sample buffer conditions, we have centred the development of this prediction method around MDS. All validation and applications of the method presented in this chapter were carried out using MDS.
Figure 4.2: Schematic diagrams of commonly used surface-based methods for studying binding interactions. (a) Variants of the ELISA method. In direct ELISA assays, the antigen (red) is immobilised on the surface, and detected by an antibody (green) directly conjugated to the detection molecule (blue). In the indirect ELISA assay, the antigen is immobilised to the surface, a primary antibody binds to the surface-bound antigen, and a secondary antibody (dark green) is used to determine the amount of primary antibody bound. The sandwich ELISA assay works in the same way, except that in this case the antigen is immobilised on the surface by binding to a surface-conjugated antibody (light green) with a different epitope from the primary antibody of interest. (b) In the bio-layer interferometry (BLI) method, the interference pattern of light reflected from the antigen-coated biosensor tip (right) in the presence of antibody is compared to that of a coated reference tip in the absence of antibody (left). Changes in the interference pattern can be recorded in real time, such that binding kinetics can be measured. (c) In the surface plasmon resonance (SPR) assay, changes in the binding status of surface-immobilised antigens lead to changes in the refractive index at the surface. This difference in refractive index can be used to quantify the degree of binding, and can be measured in real time, in order to record binding kinetics.
Figure 4.3: Schematic diagram of the microfluidic diffusional sizing method. Sample and buffer are flowed along a diffusion channel under conditions of laminar flow, in which diffusion dictates lateral movement of particles. The rate of particle diffusion is determined by their hydrodynamic radius; smaller particles diffuse more quickly than larger particles. When no antibody is present (top), the antigen diffuses quickly. When antibody is present, the binding of antibody to antigen increases its apparent size, decreasing the rate of diffusion (bottom). This apparent size increase caused by binding can be quantified by measuring the concentration of antigen (by the fluorescence intensity of a conjugated dye molecule) in the proximal and distal outlets (right).
4.4.2 *In silico* comparison of traditional and prediction methods

The traditional approach to determining binding parameters is to measure a binding curve, in which all measurement concentrations are decided on before any measurements are taken. One species is held at a concentration, and the concentrations of the binding partner are evenly spaced in logarithmic space [261]. However, even in situations where the binding stoichiometry is known and the $K_d$ is therefore the only unknown parameter, this is not the most efficient way, nor is it always sufficient, to determine the $K_d$.

To demonstrate this, we simulated binding data for an interaction in which $K_d$ is the only unknown parameter. Rather than directing us to take a traditional binding curve, with the labelled species concentration held at a constant value, our method instead determines that the best strategy is to vary the concentrations of both labelled (monitored) and unlabelled species (Figure 4.4). In reference to the MDS method, the labelled species refers to the component whose free/bound state is monitored, and unlabelled to the species which is invisible to the measurement technique. In order to be able to visualise the goodness of the fit, we therefore plot the "binding curves" against the concentration of free unlabelled species, rather than the total concentration of unlabelled species. This allows us to project the three-dimensional binding surface onto a single curve, demonstrated by the following equations:

Given the general binding interaction of species $A$ with $B$:

$$A + B \rightleftharpoons AB \quad (4.10)$$

The $K_d$ for this interaction is defined as:

$$K_d = \frac{[A][B]}{[AB]} \quad (4.11)$$

where $[A]$ and $[B]$ are the concentrations of free $A$ and $B$ in the equilibrium mixture, respectively, and $[AB]$ is the concentration of bound $AB$ complex. Rearranging, and applying conservation of mass laws:

$$[AB] = \frac{([A]_0 - [AB])[B]}{K_d} \quad (4.12)$$
4.4. Implementation

<table>
<thead>
<tr>
<th>Standard Binding Curve</th>
<th>Our Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N=30</strong></td>
<td><strong>N=9</strong></td>
</tr>
<tr>
<td>Distribution of</td>
<td></td>
</tr>
<tr>
<td>measurement points</td>
<td></td>
</tr>
<tr>
<td>Correct $K_D$</td>
<td></td>
</tr>
<tr>
<td>$K_D^{(fit)}$</td>
<td>$K_D^{(fit)}$</td>
</tr>
<tr>
<td>$K_D^{(real)} = 1$</td>
<td>$K_D^{(real)} = 1$</td>
</tr>
<tr>
<td>Incorrect $K_D$</td>
<td></td>
</tr>
<tr>
<td>$K_D^{(fit)}$</td>
<td>$K_D^{(fit)}$</td>
</tr>
<tr>
<td>$K_D^{(real)} = 0.01$</td>
<td>$K_D^{(real)} = 0.01$</td>
</tr>
<tr>
<td>Probability of $K_D$</td>
<td></td>
</tr>
<tr>
<td>Real $K_D$</td>
<td>Real $K_D$</td>
</tr>
</tbody>
</table>

Correct $K_D$ Incorrect $K_D$ Distribution of measurement points $K_D^{(fit)}$ $K_D^{(real)} = 1$ $K_D^{(fit)}$ $K_D^{(real)} = 0.01$ Probability of $K_D$ real $K_D$
Figure 4.4: Binding curves are not always sufficient even in cases when $K_d$ is the only parameter to be determined. (Top row) Concentrations of labelled and unlabelled species for which binding measurements were simulated, with the colour of the point corresponding to the mean simulated measured radius. In the standard binding curve case, three replicate "experiments" were simulated for each concentration pair, whereas measurements were simulated in duplicate in the case of our prediction method. (Second row) Simulated measurement data shown alongside the theoretical curve calculated using the true $K_d$ value. (Third row) Simulated measurement data shown alongside the theoretical curve calculated using a $K_d$ two orders of magnitude below the true one. (Bottom row) Posterior probability distributions obtained from the two approaches.

where $[A]_0$ is the total concentration of $A$ in the reaction. This yields the following expression for the fraction of $A$ that is bound to $B$ in the mixture, $f_A$:

$$\frac{[AB]}{[A]_0} = f_A = \frac{[B]}{K_d + [B]}$$

(4.13)

$f_A$ can thus be expressed as a function of $[B]$ only. If $A$ is defined to be the labelled species, and its total concentration ($[A]_0$) is therefore known, $f_A$ can be directly converted into an effective measured radius. Plotting the effective radius against $[B]$ therefore results in a one-dimensional plot even when both $[A]_0$ and $[B]_0$ are varied. It should be noted, however, that this transformation requires knowledge of $K_d$, and so determining $f_A$ for different values of $K_d$ will cause the same data to appear at different positions in this plot. Therefore, this projection is useful in judging the goodness of a fit, but offers little additional insight.

In the case where the concentration of labelled species is larger than the $K_d$, we are only able to obtain an upper bound on the $K_d$, as demonstrated by the lack of a single maximum in the posterior probability distribution (Figure 4.4). This is exemplified by plotting the calculated binding curves alongside the data, using the true $K_d$ value that was used to simulate the data, and a $K_d$ value one hundred-fold lower. In both cases, the fit describes the data well, illustrating our inability to determine the correct $K_d$ value from these data.

We next applied the prediction method to this situation. In our method, after each simulated measurement, the data are fed back into the prediction software, in order
to generate the next measurement point which will best increase our certainty about the value of $K_d$, illustrated in figure 4.5. Accordingly, measurements taken in regions predicted to not be informative have very little effect on the posterior probability distribution over $K_d$, while those predicted to have a large effect significantly increase our certainty on $K_d$. Applying this to the same simulated experiment as the binding curve, we obtain a very narrow posterior probability distribution for $K_d$. In this case, only the calculated binding curve using the true $K_d$ value, and not the $K_d$ two orders of magnitude smaller, is able to reproduce the data (Figure 4.4). The binding curve is therefore not sufficient in this case to determine $K_d$, whereas our prediction method ensures that $K_d$ can be determined.

We next considered the case where, in addition to the $K_d$, the binding stoichiometry is also a parameter to be determined. In this case, the traditional binding curve is rarely sufficient for obtaining constraints on both unknown parameters (Figure 4.6). In cases where the binding curve at a constant concentration of labelled protein is recorded at a concentration lower than the $K_d$, then only the ratio of $K_d$ to stoichiometry can accurately be determined. On the other hand, if the labelled concentration is held constant at a concentration significantly higher than that of the $K_d$, then the stoichiometry can be determined, but only an upper bound on the $K_d$ can be extracted, analogous to the situation when only $K_d$ was unknown. Both $K_d$ and stoichiometry could potentially be determined accurately if the binding curve happens to be recorded at a concentration of labelled protein similar to that of the $K_d$. Of course, when determining binding affinities, we do not have prior knowledge of the $K_d$, and can therefore only rely on chance to choose a concentration equal to the $K_d$ for the experiment.

In order to understand the different successes in determining $K_d$ and stoichiometry from these three simulation scenarios, the equations that describe the binding interaction were considered. Rearranging the binding equilibrium equation (Equation 4.11) and applying conservation of mass laws, we find that the concentration of bound complex ($[AB]$), is given by:

$$[AB] = \frac{[A][B]}{K_d} = \frac{([A]_0 - [AB])([B]_0 - [AB])}{K_d}$$

(4.14)

where $[A]_0$ is the total concentrations of $A$ in the mixture. Solving for $[AB]$, we find:

$$[AB] = \frac{[A]_0 + [B]_0 + K_d - \sqrt{([A]_0 + [B]_0 + K_d)^2 - 4[A]_0[B]_0}}{2}$$

(4.15)
Chapter 4. Optimisation of experimental design

![Expected information](image1)

![Posterior probability distribution](image2)

![Best fit](image3)

![Expected information](image4)

![Posterior probability distribution](image5)

![Best fit](image6)

![Expected information](image7)

![Posterior probability distribution](image8)

![Best fit](image9)

![Expected information](image10)

![Posterior probability distribution](image11)

![Best fit](image12)

![Expected information](image13)

![Posterior probability distribution](image14)

![Best fit](image15)

![Expected information](image16)

![Posterior probability distribution](image17)

![Best fit](image18)
4.4. Implementation

Figure 4.5: Demonstration of the prediction approach when \( K_d \) is the only parameter to be determined. Each row demonstrates the simulation of successive acquisition of measurements. The expected information is shown as heatmaps (left) over all the possible combinations of concentrations of unlabelled and labelled species, with regions of high expected information content in yellow, and regions of low expected information content in blue. The point of maximum information is indicated by the red point, and measurements that have been taken are shown by black points with white outlines. The evolution of the posterior probability distribution over \( K_d \) is shown in the central column, with the \( K_d \) value used for simulation shown by the black dotted line, while the best fit alongside the data is displayed on the right. After the first data point ([Unlabelled] = 1e-7 M, [Labelled] = 5e-7 M), very little is known about \( K_d \). For illustration, we then simulate the measurement of a point at [Unlabelled] = 5e-10 M and [Labelled] = 4e-7 M, where the expected information content is low, resulting in very little change in the posterior. Successive measurements were simulated at the most informative values indicated by the red point.

If species \( A \) is labelled, we thus monitor the concentrations \([A]\) and \([AB]\), those of free and bound \( A \), respectively. When \([A]_0 \gg K_d\), and we are in the strong-binding regime, we can take the limit and obtain:

\[
[AB] = \frac{[A]_0 + [B]_0 - \sqrt{([A]_0 + [B]_0)^2 - 4[A]_0[B]_0}}{2}
\]  
(4.16)

so that

\[
\frac{[AB]}{[A]_0} = \frac{1}{2[A]_0}([A]_0 + [B]_0 - ||[A]_0 - [B]_0||)
\]  
(4.17)

Therefore, if \([A]_0 > [B]_0\):

\[
\frac{[AB]}{[A]_0} = \frac{[B]_0}{[A]_0}
\]  
(4.18)

Or if \([A]_0 < [B]_0\):

\[
\frac{[AB]}{[A]_0} = 1
\]  
(4.19)

In both cases, the expression for fraction of \( A \) bound is independent of the \( K_d \), and we therefore only obtain information on the total concentration of \( B \).
Chapter 4. Optimisation of experimental design

Figure 4.6: Binding curves are not always sufficient when both $K_d$ and concentration are parameters to be determined. Joint posterior probability distributions (top: yellow = high probability density, blue = low probability density) over $K_d$ and [Total unlabelled], and the data shown alongside the best fit (bottom). (Left) When the concentration of labelled protein ([Labelled] = $10^{-9}$ M) is lower than the $K_d = 10^{-8}$ M (weak binding regime), we are only able to determine the ratio of concentration to $K_d$, demonstrated by the straight line of high probability in the posterior. (Centre) When [Labelled] is exactly equal to the $K_d$ of interaction (both $10^{-8}$ M), we are able to determine approximate constraints on both $K_d$ and concentration, but the correlation between the parameters is still evident. (Right) When [Labelled] exceeds the $K_d$ (here: [Labelled] = $10^{-7}$ M and $K_d = 10^{-8}$ M), the concentration can be determined to a high degree of certainty, but only an upper bound on the $K_d$ can be extracted. In all cases, the data are well described by the fit, and it we can determine that the parameters are not both constrained in all cases.
When \([A]_0 \ll K_d\), in the weak-binding regime, we know that only a very small fraction of species \(B\) is bound, and we can therefore approximate the equilibrium binding equation as:

\[
K_d = \frac{([A]_0 - [AB])B_0}{[AB]} \quad (4.20)
\]

Rearranging, we obtain:

\[
\frac{[AB]}{[A]_0} = \frac{[B]_0}{K_d + [B]_0} \quad (4.21)
\]

If \([B]_0 \ll K_d\), we have:

\[
\frac{[AB]}{[A]_0} = \frac{[B]_0}{K_d} \quad (4.22)
\]

whereas if \([B]_0 \gg K_d\), we have:

\[
\frac{[AB]}{[A]_0} = 1 \quad (4.23)
\]

In the weak binding regime, we therefore obtain information on the ratio of \([B]_0\) to \(K_d\). In order to verify that the approximations made in these limits are valid, we compare them with the exact analytical solutions (Figure 4.7), and find that they are in good agreement. These equations therefore explain the results of the Bayesian inference, where only the ratio of stoichiometry to \(K_d\) could be determined in the \([A]_0 \ll K_d\), and only stoichiometry could be determined in the \([A]_0 \gg K_d\) situation. In order to visualise these two regimes, the binding surface plot was computed for given \(K_d\) and stoichiometry values (Figure 4.8). A clear shift in the shape of the surface can be seen when \([A]_0 \sim K_d\); when \([A]_0 \gg K_d\), the contour lines are diagonals, whereas when \([A]_0 \gg K_d\) the contour lines are vertical.

We next applied the prediction method to this situation, where both \(K_d\) and stoichiometry are unknown (Figure 4.9). Note that this is analogous to not knowing the total \([B]\) concentration. Mathematically, this can be expressed as follows:

\[
[AB] = [A]_0 + \beta[B] + K_d - \sqrt{([A]_0 + \beta[B] + K_d)^2 - 4[A]_0\beta[B]} \quad (4.24)
\]

where \(\beta\) is the fraction of the total stock of \(B\) in the mixture, and \([B]\) is the total stock concentration of \(B\). We therefore infer \([B]\) (total [unlabelled]) and \(K_d\) from
Figure 4.7: Validation of the approximations in the weak and strong binding regimes (equations 4.21 and 4.17, respectively). The calculations were performed using $K_d = 10^{-8}$ M, and $[A]_0 = 10^{-7}$ M (strong binding), or $[A]_0 = 10^{-9}$ M (weak binding).

the analysis. In this case, the deviation of the predicted measurement points from a traditional binding curve is particularly striking. The software predicts that the optimal experimental strategy is to acquire data in both the weak and the strong binding regimes, and based on the current data at each step updates where the switch between these regimes is expected to be (Figure 4.9). This is consistent with our finding that data in the weak binding regime only allow us to determine the ratio of $K_d$ and concentration, whereas the strong binding regime provides absolute constraints on concentration (Figure 4.6). Combining data from the two regimes therefore allows us to determine both parameters, in order to determine the value of [Labelled] at which the shift in binding surface shape occurs (Figure 4.8), which corresponds to the $K_d$. 
**Figure 4.8**: Binding surface, showing the weak and strong binding regimes, with colours representing the fraction of unlabelled species that is bound to labelled species. The boundary lines at which [Labelled] and [Unlabelled] are equal to the $K_d$ are indicated in dotted lines. There is a clear shift in the topology of the binding surface when [Labelled] = $K_d$. When fitting binding data to determine $K_d$, we are therefore looking for the point at which this shift occurs. When [Labelled] $\gg K_d$, the fraction bound is solely determined by the concentration of unlabelled protein. When [Labelled] $\ll K_d$, the fraction bound is now dependent on both the concentration of unlabelled protein and the $K_d$. 
Figure 4.9: Demonstration of the prediction approach when $K_d$ and concentration are both parameters to be determined. Each row demonstrates the simulation of successive acquisition of measurements. The expected information is shown as heatmaps (first column) over all the possible combinations of concentrations of unlabelled and labelled species, with regions of high expected information content in yellow, and regions of low expected information content in blue. The point of maximum information is indicated by the red point, and measurements that have been taken are shown by black points. Successive measurements were simulated at the values indicated by the red point. The evolution of the posterior probability distribution over $K_d$ and concentration is shown in the second column, while the best fit alongside data is displayed in the third and fourth columns. A $K_d$ of 1e-8 M and total unlabelled concentration of 5e-7 M were used to simulate the data.
4.4.3 *In vitro* validation and application: Characterising therapeutic antibodies against Aβ isoforms

Having tested the prediction algorithm using simulated data, we next sought to validate the applicability of the method in real wet-lab experiments. Since our theory work demonstrated that the most significant improvement over standard binding curve approach could be achieved when both binding stoichiometry and $K_d$ are unknown. In this case, a single binding curve is rarely sufficient to determine both parameters, and the prediction method is necessary (Figure 4.6).

As discussed previously, protein aggregation is under heavy investigation for its potential as a therapeutic target in its associated diseases [65, 66]. Multiple clinical trials have focused on antibody-based strategies to target the Aβ peptide, one of the two proteins which forms the characteristic aggregate deposits in Alzheimer’s disease [10, 262, 263, 264, 265, 266]. Some of the most recently studied antibodies are aducanumab, solanezumab, bapineuzumab, and gantenerumab [262, 263, 264, 265, 266]. These four antibodies were all produced to bind epitopes on Aβ. Aducanumab has shown promise in phase 3 trials, through effects on both clinical symptoms and disease-related biomarkers, and binds to the N-terminal region containing residues 3-7 [265, 267, 268]. Gantenerumab is the subject of several phase 3 studies, and recognises epitopes at residues 3-11 and 18-27 [263, 269]. Clinical trials have been discontinued for bapineuzumab, which recognises residues 1-5 [264]. The epitope of solanezumab is located at residues 16-26 and is under investigation in reducing amyloid levels in cognitively normal patients [266, 270]. Considering the fibril structure of Aβ42, the binding specificities of the antibodies can be rationalised; of the four antibodies, only the epitope of solanezumab is located in the fibril core (Figure 4.10), thus making solanezumab the only monomer-specific antibody of the four [271, 272].

In order to determine the biophysical parameters that describe binding of these antibodies to Aβ fibrils, we employed microfluidic diffusional sizing. Fluorescently labelled antibodies were incubated with unlabelled Aβ42 fibrils, and the increase in antibody size due to binding interactions was monitored by diffusional sizing [28]. Although the total fibril mass concentration is known, the concentration of antibody binding sites on the fibrils is unknown. For example, antibodies could bind only to fibril ends, or additionally to sites present at an unknown density along the fibril surface. We therefore have two unknown parameters to be determined, the $K_d$.
Figure 4.10: Aβ42 fibril structure, determined by ssNMR. Figure adapted from [271]. (a) Cross sectional view of the fibril, showing the positions of residues. Residues 15-42 are located within the fibril core; antibodies which recognise epitopes in this region are therefore unable to bind fibrils. (b) Fibril structure, showing the parallel cross-β structure. (c) Cross sectional structure, showing the disordered N-terminal regions, which are therefore exposed and available for binding to antibodies.
and the binding site concentrations, which we express here as the binding sites per monomer unit in the fibrils. We conducted binding experiments for gantenerumab with fibrils, both using the traditional binding curve approach and our prediction method (Figure 4.11). Data from both experiments were well described by the fit, but while the binding curve yielded a fully correlated $K_d$ and binding site concentration in the posterior probability distribution, our prediction approach was able to provide absolute constraints on both parameters independently (Figure 4.11).

Having validated the performance and effectiveness of the prediction method in the case of antibody/fibril binding, we were now able to quantitatively compare the binding of the four therapeutic antibodies to both monomers and fibrils. In the monomer binding experiments, the stoichiometry of 1:2 antibody-to-monomer was known, based on the size increase observed and structural considerations, and so binding curves were sufficient to determine $K_d$ values for each antibody (Figure 4.12a). Applying our prediction method to the fibril binding experiments, we were able to determine both stoichiometries and affinities of binding for all antibodies except for solanezumab, which did not bind to fibrils, as expected based on its epitope location (Figures 4.10 and 4.12b-d, Table 4.1).

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Monomer $K_d$/nM</th>
<th>Monomer $K_d$/nM</th>
<th>Binding sites per monomer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aducanumab</td>
<td>9200 ± 1300</td>
<td>1.5 ± 1.2</td>
<td>0.22 ± 0.03</td>
</tr>
<tr>
<td>Gantenerumab</td>
<td>490 ± 85</td>
<td>29 ± 14</td>
<td>0.023 ± 0.003</td>
</tr>
<tr>
<td>Bapineuzumab</td>
<td>36 ± 9</td>
<td>0.32 ± 0.17</td>
<td>0.025 ± 0.002</td>
</tr>
<tr>
<td>Solanezumab</td>
<td>4.4 ± 3.1</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

**Table 4.1**: Affinities and stoichiometries of Aβ42 monomers and fibrils binding to therapeutic antibodies. The errors given are the standard deviations of the one-dimensional marginal distributions.

Strikingly, aducanumab exhibited a very different binding stoichiometry from gantenerumab and bapineuzumab, with an order of magnitude more binding sites on the fibril surface. This suggests that the antibodies have different binding modes, with the low stoichiometry of bapineuzumab and gantenerumab indicating that
4.4. Implementation

**Figure 4.11:** Comparison of the binding curve approach with the prediction method for gantenerumab/fibril binding. (Top) Schematic of antibodies binding to fibrils. While the concentration of fibrils is known, the concentration of antibody binding sites on the fibril surface is unknown. (Middle) Binding data shown alongside the best fit. Insets show the range of possible experiments, in terms of antibody and fibril concentrations, with the colour of points corresponding to the mean measured radius. (Bottom) Joint posterior probability distributions of $K_d$ and stoichiometry, expressed as the number of binding sites per monomer unit in the $A_{\beta}$ fibrils.
Figure 4.12: Aβ/antibody binding parameters determined through prediction. (a) Posterior probability distributions over $K_d$ obtained for monomeric Aβ42 binding to therapeutic antibodies. (b) Joint posterior probability distributions over $K_d$ and stoichiometry (binding sites per monomer in Aβ42 fibrils) for the binding of Aβ42 fibrils to antibodies. Points denote the maximum posterior values of the parameters. No binding was detected for solanezumab, and it is therefore not shown here. The one-dimensional marginals of these distributions are shown in c and d.
they can only bind to the ends of fibrils (Figure 4.13b). In contrast, the high stoichiometry of binding for aducanumab implies that it can bind to sites along the fibril surface (Figure 4.13a). These binding modes are reflected in the effects of the antibodies on aggregation kinetics. While the dominant effect of gantenerumab and bapineuzumab is to inhibit elongation, consistent with their binding to fibril ends, aducanumab primarily affects secondary nucleation, a behaviour which can be attributed to its binding along the fibril surface. Solanezumab, which only binds to monomeric A\textsubscript{\beta}42, inhibits primary nucleation of monomers but does not affect elongation or secondary nucleation [268]. Furthermore, the differential binding of these antibodies to monomers and fibrils show different preferences; while the affinities of gantenerumab and bapineuzumab are selective for fibrils by only 1-2 orders of magnitude, the affinity of aducanumab for fibrils is roughly 4 orders of magnitude higher, which again may contribute to its greater success in clinical trials, due to its ability to inhibit secondary nucleation and thereby the spread of aggregates [265, 201].

**Figure 4.13:** Binding modes of antibodies to fibrils. Schematics of antibody binding sites being distributed along the fibril surface (a), as in the case of aducanumab, or located only at fibril ends (b), as in the cases of gantenerumab and bapineuzumab.
4.4.4 Application: Antibody quantification for the determination of transplant compatibility

With the prediction method now verified in a real experimental setting, we next moved to an analogous application; the determination of antibody concentrations and binding affinities in patient serum samples. This is mathematically equivalent to the fibril binding, where both the total binding site concentration and $K_d$ are unknown. The first system to which we applied this is the case of Human Leukocyte Antigen (HLA) reactive antibodies, termed alloantibodies. The reactivity of these alloantibodies to specific HLA variants have been found to be strong biomarkers of transplant graft rejection, and their characterisation is therefore an essential factor in pre-transplant assessment and post-transplant immune monitoring [273, 274, 275, 276, 277]. By determining the HLA and alloantibody expression profiles of transplant recipients and donors, we aim to predict compatibility and thereby reduce the chances of rejection.

First, we verified that our method of combining MDS with prediction can yield accurate antibody concentrations and affinities in human serum. The A*02:01 variant of HLA was conjugated to a fluorescent dye, AlexaFluor-647, and incubated with a non-HLA reactive serum sample, spiked with known concentrations of a purified antibody known to bind HLA A*02:01. The antibody concentrations determined through the prediction and analysis correspond very well to the known values, assuming a binding stoichiometry of 1:2 (antibody-to-HLA) based on the radius of the bound complex (Figure 4.14).
Figure 4.14: Validation of the method in the quantification of HLA-specific antibody binding in serum. Known concentrations of HLA-specific antibody were spiked into unreactive human serum, and their binding to HLA quantified to determine [antibody] and the \( K_d \) of interaction. (Top) Measured radii shown alongside the best fit, with error bars for standard deviation. In all cases, the data are well-described by the fit. (Middle) Posterior probability distributions over \( K_d \) obtained for each dataset. The grey region corresponds to the 95% confidence intervals obtained through fitting all datasets combined, using the known antibody concentrations, and is thus an indicator of the true \( K_d \) of interaction. In all cases, this region overlaps significantly with the region of high probability determined in the analysis of each individual dataset, although the 3 nM dataset is approaching the lower limit of detection as reflected by the spread out posteriors (Bottom) Posterior probability distributions of antibody concentrations, assuming 1:2 antibody-to-HLA molar binding stoichiometry. The shaded region corresponds to the range of antibody concentrations within the estimated error. In all cases, the expected concentration is approximately equal to the best fit value determined by the MDS analysis.

Having confirmed that the combination of MDS and prediction yields accurate \( K_d \) and antibody concentration values, we next aimed to characterise the alloantibody content of a patient sample. The patient had become sensitised after receiving a kidney transplant expressing the HLA variant A*24:02. The patient serum was tested for antibodies specific to the A*24:02 and A*02:01 variants of HLA, which share
Figure 4.15: Quantification of alloantibodies in patient serum. Posterior probability distributions obtained for antibodies in the same patient serum sample against the A*02:01 (left) and A*24:02 (right) variants of HLA.

Some sequence similarity, finding comparable $K_d$ values for antibodies against both variants in preliminary experiments (Figure 4.15) [278]. However, the concentration of antibodies reactive against the A*24:02 variant was higher than that determined for the A*02:01 variant, suggesting that there is at least one population of alloantibodies which recognise the A*24:02 variant, but not the A*02:01. Current work is now focused on expanding these data to be able to correlate these antibody profiles with clinical data.
Having confirmed that our method enables the accurate determination of both antibody $K_d$ and concentration in human serum samples, we next used the method to quantify antibody responses to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. Coronavirus disease 2019 (COVID-2019) is caused by SARS-CoV-2, and was first reported in China in December 2019 [279]. Since then, its rapid spreading throughout the world has had massive economic and social consequences, and effective therapeutic measures are therefore under heavy investigation [280]. SARS-CoV-2 infections elicit a wide range of responses in patients; while a significant fraction of patients exhibit no symptoms, up to 10% of cases are fatal [281, 282]. While pre-existing health conditions and advanced age have been found to cause increased susceptibility to the virus, the source of most of this range in disease severity is still unknown [283]. Much effort has been directed at understanding the antibody responses of patients, in order to determine if and how antibody production is related to disease progression and severity [284, 285, 286, 287, 288]. As discussed earlier, the ELISA methods for assessing antibody responses are surface-based, and the quantities thus determined may contain contributions from both the $K_d$ and concentration of antibody present. Using our prediction method, we therefore investigated the extra information contained in determining these two parameters independently.

The SARS-CoV-2 viral particles are believed to infect patient host cells by the interaction of the viral spike receptor protein with angiotensin converting enzyme 2 (ACE2) receptors on the host cell surface [289]. Structural studies of this interaction have identified the receptor binding domain (RBD) of the spike protein to be the relevant site of interaction on the virus particle surface [290, 291]. In order to maximise the size increase upon binding to antibodies while still maintaining the presence of physiologically relevant epitopes, the RBD was therefore selected for these binding studies.

39 patient serum samples were tested, after testing positive for RBD-reactive antibodies by ELISA [292]. Of these 39 patients, 3 required hospitalisation in the intensive care unit, and the remaining 36 were asymptomatic or only reported mild symptoms. Due to difficulties in objectively distinguishing between asymptomatic
and mildly symptomatic based on patient reports, these 36 patients were combined into one class. Antibodies with quantifiable binding were detected by MDS in all patient samples, except for 7 of the non-hospitalised patients. The exact cause of the apparent absence of RBD-reactive antibodies in these 7 patients could be due to a false positive ELISA result or to the quantities being too low to be detected by MDS.

Comparing the $K_d$ and antibody concentrations (assuming a 1:2 antibody-to-RBD molar binding stoichiometry) obtained for the 32 samples with quantifiable binding, we found generally similar values for the antibody concentration, with most patients being within an order of magnitude of each other, whereas the $K_d$ values varied more widely between patients, with values for some samples being sub-nanomolar, and others being tens of nanomolar (Figure 4.16a). While only three patients who required hospitalisation were studied, all three had higher antibody concentrations than the average of the overall population. By considering the physical requirements for significant binding to occur, the total antibody concentration must exceed the $K_d$. Accordingly, in all 32 samples where binding could be quantified, the determined antibody concentration is greater than the $K_d$. In many cases, however, this is not by a large factor, less than an order of magnitude in most cases.

In order to better understand the dynamics of the antibody response, we obtained serum samples from the three hospitalised patients at different timepoints during the infection. All three patients suffered from diabetes, with patients 2 and 3 presenting with additional cardiovascular conditions and requiring hospitalisation due to pneumonia. Analyses were performed for patient 1 (days post onset (DPO) 9 -13), patient 2 (DPO 8 -14), and patient 3 (DPO 7 -15). In all cases, no binding was detected by MDS until 12 DPO, consistent with the ELISA data [292]. Analysis of plasma samples taken from patients 1 and 2, taken one-two days apart, indicate that antibody concentration increases with no change in binding affinity (DPO 12 and 13 for patient 1, and DPO 12 and 14 for patient 2) (Figure 4.16b). These data suggest that after producing an antibody with sufficiently high affinity, the affinity is not increased further, in favour of instead increasing antibody production. While our data are limited to just two timepoints in two patients, this effect is striking, and in contrast to previous work on Ebola, where antibody affinity has been found to increase as a function of time [293]. Our data therefore indicate that determining both the concentration and affinity independently from each other may be important in understanding antibody responses in disease.
Figure 4.16: Quantification of RBD-specific antibodies in SARS-CoV-2 patient samples. Posterior probability plots, shown alongside the maximum likelihood values of $K_d$ and antibody concentration (assuming a 1:2 antibody-to-RBD molar binding stoichiometry) for all patient samples studied (a) and longitudinal data of hospitalised patients (b). In both plots, the grey shaded region indicates the area where the concentration of antibody binding sites is less than $K_d$, and binding therefore cannot be quantified, while the dashed line at 4 nM [Ab] represents the lowest antibody concentration we can quantify by MDS, due to the sensitivity of the measurement. In a, the dotted line shows where the concentration of binding sites exceeds $K_d$ by a factor of 10. Timepoint 1 represents 12 DPO for both patients 1 and 2, while Timepoint 2 represents 13 DPO for patient 1 and 14 DPO for patient 2.
4.5 Conclusions and future directions

In this chapter, I have presented the development of a method to optimise the efficiency and information gain of experiments, and its application in several experimental settings. Using our prediction method, it is now possible to determine both the concentration and affinity of specific antibodies in patient serum samples, a situation where the standard binding curve approach does not suffice. We have applied this in determining the biophysical binding parameters of anti-Alzheimer’s antibodies against Aβ species, providing insights into the factors that govern therapeutic success. We have then extended the approach to studying patient serum samples, in the contexts of transplant compatibility and antibody responses to SARS-CoV-2 infection.

The mathematical framework behind the prediction method is fully general, and its applications are therefore not limited to binary molecular interactions. This method will be integrated into the analysis and design of aggregation kinetic reactions, particularly in the presence of external factors, in order to quantify their effects on individual rate constants of the microscopic mechanisms involved (e.g. elongation and primary nucleation). This will therefore enable the determination of specificities of such factors, and thereby aid in the elucidation of molecular mechanisms of aggregation.

At present, the method optimises experimental design based solely on the tightness of the posterior probability distribution. However, this will be extended to also be able to optimise experimental efficiency in other contexts, such as model differentiation. While in the examples presented here, equilibrium binding was known to be the appropriate mathematical model, this is not always the case. Binding experiments may be performed to determine whether the binding is cooperative, or to determine whether inhibitors are competitive or allosteric. In these situations, simply determining the parameters of a given model is not necessarily the best thing to do, but measurements should instead be chosen such that they maximise our confidence in assigning the correct model.

In order to make the prediction method accessible, we are in the process of creating an online server (free for academic use), onto which users can upload data and perform analysis and prediction. Furthermore, the next step to maximise experimental efficiency further, particularly in standardised methods, would be to
4.5. Conclusions and future directions

Figure 4.17: Schematic of experimental automation with prediction. In the traditional plate layout (top), complete dilution series (different shades) of samples (different colours) are measured in each plate. In order to maximise experimental efficiency and information gain (bottom), this could be altered so that each plate contains many different samples, and additional dilutions of each sample are then run on future plates, based on the measurements dictated by the in-built prediction algorithm.

remove manual input and integrate the prediction component into lab machines. For example, when scanning multiple samples, instead of running multiple dilutions of a given sample all on the same plate, an automated setup could run just one dilution of each sample on one plate, and measure additional dilutions of samples on subsequent plates, by using the data of previous plates as input for the prediction method (Figure 4.17).
Chapter 5

Materials and methods

5.1 Methods: Structural characterisation of kinetically trapped α-synuclein oligomers

5.1.1 Sample preparation

Protein purification

α-Synuclein was purified from Escherichia coli BL21 cells, transformed with the pT7-7 plasmid encoding the protein. Cells were grown in 2xYT with ampicillin (100 µg/ml), induced with IPTG (1 mM), incubated overnight at 28 °C, and harvested by centrifugation in a Beckman Avanti J25 centrifuge with a JLA-8.1000 rotor at 6238 x g (Beckman Coulter UK Ltd., High Wycombe, UK). The cell pellet was resuspended in 10 mM Tris-HCl (pH 7.7), 1 mM EDTA, cOmplete protease inhibitor cocktail (1 tablet/40 mL; Roche, West Sussex, UK), and lysed by sonication. Following centrifugation (JA-25.50 rotor, 39191 x g), the supernatant was boiled for 20 minutes and centrifuged (39,191 x g). Streptomycin sulphate was added to the supernatant to a final concentration of 10 mg/mL and centrifuged (39,191 x g). Ammonium sulphate was added to the supernatant to 361 mg/mL and stirred at 4 °C for 30 minutes. The mixture was centrifuged (39191 x g), and the pellet resuspended in 25 mM Tris-HCl (pH 7.7) and loaded onto a HiLoad™ 26/10 Q sepharose high performance column (GE Healthcare Ltd., Little Chalfont, UK) and eluted at 350 mM NaCl with a salt gradient from 0 M to 1.5 M NaCl. Selected fractions were subsequently loaded onto a Superdex 75 26/60 (GE Healthcare Ltd.) and eluted
in PBS. Protein concentration was determined by absorbance at 275 nm, using an extinction coefficient of 5600 M$^{-1}$ cm$^{-1}$.

**Preparation of kinetically trapped oligomers**

Oligomers were prepared as previously described [37]. Briefly, $\alpha$-synuclein was purified into PBS [294], and subsequently dialysed against water (4 L) (ON, 4 °C). 6 mg aliquots were lyophilised (48 h), followed by resuspension in buffer (500 µL PBS). The resuspended protein was passed through 0.22 µm filters and incubated (20-24 h, 37 °C). The samples were ultracentrifuged (1 h, 288,000 rcf, 20 °C) in a TLA 120.2 rotor, using an Optima TLX Ultracentrifuge (both Beckman Coulter, High Wycombe, UK) to remove aggregates and large oligomers. Remaining monomer was removed using a 100 kDa centrifugation filter (4 × 2 min, 9300 rcf). The flow-through containing predominantly monomer from the first three passes was kept and reused up to five times. Oligomer concentration was determined using UV-vis spectroscopy, using molar extinction coefficients of 7000 M$^{-1}$ cm$^{-1}$ for the WT, E46K, H50Q, and A53T, and 12,444 M$^{-1}$ cm$^{-1}$ for A30P and G51D, with molar extinction coefficients determined using amino acid analysis and BCA assays.

Due to the clear differences between the CD spectra of the monomers and oligomers, we were able to use CD spectroscopy to monitor oligomer dissociation to monomer over time (Fig. 2.23), finding that all oligomers reached an equilibrium with monomer after several days of incubation at room temperature. Nevertheless all experiments were carried out within 24 h of oligomer purification, in order to keep concentrations accurate.

**Preparation of fibrils**

Fibrils were prepared as previously described [37]. $\alpha$-Synuclein monomer in PBS was incubated (3 days, 37 °C, 70 µM, 300 µL) under shaking conditions (200 rpm). Fibrils, henceforth referred to as f0 fibrils, were next centrifuged (16,100-21,130 rcf, 15 minutes, 20 °C) and resuspended in PBS three times. F0 fibril concentration was determined by UV-Vis spectroscopy, following incubation (30 minutes, room temperature) of an aliquot with GdmCl (4 M). F0 fibrils were sonicated (30% cycles, 10% power, 1 minute), before incubation of 10 µM f0 fibrils with 100 µM monomers under quiescent conditions (overnight, 37 °C), to yield f1 fibrils. F1 fibrils were collected and their concentration determined, as described for the f0 fibrils, and
finally sonicated (30% cycles, 10% power, 20 seconds) before use. In order to prevent bacterial growth, 0.02% (w/v) sodium azide was added to both incubation steps. Amyloid fibril morphology was confirmed by TEM (Fig 2.9).

**Preparation of small unilamellar vesicles**

Lipid mixtures of DOPC:DOPE:DOPS (2:5:3 w/w/w) suspended in chloroform were purchased from Avanti Polar Lipids. Known amounts of lipids were dried under nitrogen to a thin film and lyophilised overnight, before resuspension in PBS. Resuspended lipids were subjected to 3-5 freeze-thaw cycles in liquid nitrogen, and subsequently sonicated (50% cycles, 10% power, 10 \times 2 minute bursts) on ice, yielding small unilamellar vesicles (SUVs) with a diameter of 30-50 nm, confirmed by dynamic light scattering (DLS).

**5.1.2 Biophysical characterisation**

**Analytical ultracentrifugation**

Sedimentation velocity measurements were performed at 20 °C, 38,000-43,000 rpm (106,750-136,680 \times g) by using a Beckman-Coulter Optima XL-I analytical ultracentrifuge equipped with UV-visible absorbance optics and an An50Ti rotor. The sedimentation coefficient distributions, corrected to standard conditions by using the SEDNTERP program, were calculated via least-squares boundary modelling of sedimentation velocity data using the c(s) and ls-g*(s) methods, as implemented in the SEDFIT program (www.analyticalultracentrifugation.com/default.htm) [295].

**ANS binding**

8-Anilino-1-sulfonic acid (ANS) was added to samples (5 µM protein) to a final concentration of 250 µM and subsequently incubated (30 min, 20 °C). Fluorescence emission spectra were recorded between 400-650 nm with an excitation wavelength of 350 nm, using a Cary Eclipse Fluorescence spectrophotometer (Agilent, Santa Clara, CA).
Bicinchoninic acid assay

Bicinchoninic acid (BCA) assays were performed using a kit and bovine serum albumin (BSA) (both Thermo Scientific, Rockford, IL), and carried out in Corning 96 well plates (3635). 200 µL working reagent was added to 25 µL of sample, and incubated at 37 °C for 45 min. Following incubation, absorbance at 562 nm of each sample was recorded on a FLUOstar Optima plate reader (BMG Labtech, Aylesbury, UK). A standard curve was generated using concentrations of stock BSA between 0 and 250 µg mL⁻¹, which was used to determine protein concentrations in the sample. In order to account for potential differences in behaviour of BSA and α-synuclein in the assay, samples were normalised to a known α-synuclein standard sample.

Binding of α-synuclein to small unilamellar vesicles

α-Synuclein samples (4.5-10 µM) were incubated with SUVs (30 min, RT) prior to measurements. The secondary structures of samples were determined by CD spectroscopy; far-UV spectra were recorded in a quartz cuvette (1 mm path length), on a JASCO J-810 equipped with a Peltier thermally controlled cuvette holder (Jasco (UK) Ltd, Dunmow, UK) at 20 °C. 15 spectra were averaged and recorded between 250 and 200 nm, with a data pitch of 0.5 nm, bandwidth of 1 nm, scanning speed of 50 nm/min, and response time of 4 sec. In order to quantify the binding interactions, the CD signal was converted to the mean residue ellipticity (MRE):

\[
MRE_\lambda = \frac{\theta_\lambda M_0}{10cl}
\]  

(5.1)

where \(MRE_\lambda\) is the MRE at wavelength \(\lambda\), \(\theta_\lambda\) is the observed ellipticity in mdeg at \(\lambda\), \(M_0\) is the mean residue weight of the protein \(\frac{\text{Molecular weight}}{\text{Number of amino acids}}\), \(c\) is the protein concentration in mg ml⁻¹, and \(l\) is the path length in cm.

The MRE was used to determine what fraction of α-synuclein was membrane-bound. The CD signal can be broken down into two additive components. As MRE scales linearly with CD signal, we can represent this as:

\[
MRE_{\text{obs}} = \chi_B MRE_B + \chi_F MRE_F
\]  

(5.2)

where \(\chi_B\) and \(\chi_F\) are the fractions of α-synuclein bound to membrane and free in solution, respectively, and \(MRE_B\) and \(MRE_F\) the CD signals of fully bound and free α-synuclein, respectively. We can obtain \(\chi_B\) from our titrations.
5.1. Methods: Structural characterisation of kinetically trapped α-synuclein oligomers

\[ \chi_B = \frac{MRE_{\text{obs}} - MRE_F}{MRE_B = MRE_F} \]  \hspace{1cm} (5.3)

We can represent the binding of α-synuclein to SUVs as an equilibrium reaction:

\[ K_d = \frac{[F][SUVs_L]}{[B.SUVs_L]} \]  \hspace{1cm} (5.4)

\[ [\alpha - \text{synuclein}] = [F] + [B.SUVs_L] \]  \hspace{1cm} (5.5)

\[ [SUVs] = L([SUVs_L] + [B.SUVs_L]) \]  \hspace{1cm} (5.6)

\[ \chi_B = \frac{([\alpha - \text{syn}] + \frac{[SUVs_L]}{L} + K_d) - \sqrt{([\alpha - \text{syn}] + \frac{[SUVs_L]}{L} + K_d)^2 - \frac{4[SUVs][\alpha - \text{syn}]}{L}}}{2[\alpha - \text{syn}]} \]  \hspace{1cm} (5.7)

Matlab (Version R2016b, The MathWorks Inc., Natick, MA) was used to fit the data from the CD-monitored titrations to this model.

Circular dichroism spectroscopy

α-Synuclein far-UV spectra were recorded in a quartz cuvette (1 mm path length), on a JASCO J-810 equipped with a Peltier thermally controlled cuvette holder (Jasco (UK) Ltd, Dunmow, UK) at 20 °C. 15-30 spectra were averaged and recorded between 250 and 200 nm, with a data pitch of 0.5 nm, bandwidth of 1 nm, scanning speed of 50 nm/min, and response time of 4 sec. Spectra were deconvoluted using the BestSel web server [149, 150].

Dot blot

1 µg monomer or oligomer was deposited onto a 0.2 µm PVDF membrane (Millipore) and left to dry at room temperature. The membranes were then blocked (5% BSA in PBS, 1 h, room temperature) and subsequently incubated with A11 or 211 antibody in 5% BSA in PBS (overnight, 4 °C) at concentrations of 1:5000 and 1:2000, respectively. Membranes were washed in PBS + 0.01% Tween-20 (PBST) (3 × 10 min, room temperature) then incubated with secondary antibody (AlexaFluor-488 goat anti-mouse for A11, and AlexaFluor-488 goat anti-rabbit for 211, both 1:5000
(Invitrogen)) in PBST (1 h, room temperature). Following washing in PBST (3 × 10 min, room temperature), membranes were imaged on a Typhoon Trio scanner and the images analysed using ImageQuant TL v2005 (both Amersham Bioscience).

**Dynamic light scattering**

Dynamic light scattering (DLS) measurements were performed on a Zetasizer Nano ZSP (Malvern Instruments, Malvern, UK) at 25 °C. with the following settings: refractive index of 1.45, respectively, absorption of 0.001, dispersant of ICN PBS tablets, and 173° backscatter measurement angle.

**Fourier transform infrared spectroscopy**

FTIR measurements were performed on a Vertex 70 (Bruker, Billerica, MA), fitted with a Platinum ATR (Diamond F). 2 μL 2-15 μM oligomer sample was deposited onto the detector and dried, followed by washing with milliQ water. 5 spectra each averaged over 128 scans, with atmospheric compensation and background correction, were recorded per sample. Recorded spectra were baseline corrected in the 1720-1580 cm⁻¹ (amide I) region, and normalised.

**Intrinsic fluorescence spectroscopy**

Intrinsic fluorescence spectra were recorded on a Cary Eclipse fluorescence spectrophotometer (Agilent, Santa Clara, CA). For emission spectra, samples were excited at 276 nm, while for excitation spectra, emission was monitored at 305 nm.

**Native polyacrylamide gel electrophoresis**

Samples were mixed with 4× loading buffer and run on NativePAGE™ 4-16% Bis-Tris gels (Thermo Scientific, Rockford, IL), alongside NativeMark™ Unstained Protein Standard (Invitrogen by Thermo Fisher Scientific, Carlsbad, CA), for 105 min at 125 V using NativePage™ Cathode Buffer Additive and NativePage™ Running Buffer (both Novex, Carlsbad, CA). Gels were destained with a mixture of water, ethanol, and acetic acid (5:4:1 volume ratio).
5.1. Methods: Structural characterisation of kinetically trapped α-synuclein oligomers

SDS polyacrylamide gel electrophoresis

Samples were mixed with 4 × loading buffer and run on 4-12% Bis-Tris gels (Thermo Scientific, Rockford, IL), alongside Tri-Colour Marker (Expedeon, UK) for 35 minutes at 200 V in NuPAGE® MES SDS Running Buffer (Novex, Carlsbad, CA), and stained with InstantBlue™. Quantification of bands was achieved by densitometry analysis with ImageJ 1.x.

Transmission electron microscopy

Each sample (10 µM, 10 µL) was adsorbed onto carbon-coated copper grids. Once dry, grids were washed (2 × 10 µL water), followed by staining with 2% (w/v) uranyl acetate, and further washes (2 × 5 µL water). The samples were imaged on a Phillips CM-100 transmission electron microscope operating at 80 kV (Cambridge Advanced Imaging Centre, University of Cambridge, UK).

5.1.3 Cellular toxicity

Lipopolysaccharide concentration determination

α-Synuclein monomer stock samples in PBS were diluted (50 µM) into sterile water (Thermo Scientific, Waltham, MA) were tested for lipopolysaccharide (LPS) contamination using the Pierce™ LAL Chromogenic Endotoxin Quantitation Kit. 50 µL sample or LPS standard was incubated with 50 µL Limulus Amebocyte Lysate (LAL) reagent (37 °C, 10 min) in a 3635 plate (Corning). 100 µL prewarmed Chromogenic Substrate solution was added and incubated (30 °C, 6 min). The absorbance at 405 nm was measured on a plate reader (Clariostar, BMG Labtech, Offenburg, Germany), and LPS concentrations calculated by reference to the standard calibration curve.

MTT cell viability assay

MTT cell viability experiments were performed by Marta Castellana-Cruz. Human SH-SY5Y neuroblastoma cells (A.T.C.C., Manassas, VA) were cultured in 1:1 Dulbecco’s modified Eagle’s medium (DMEM)-F12+GlutaMax supplement (Thermo Fisher Scientific, Waltham, MA) supplemented with 10% fetal bovine serum. The cells were maintained in a 5.0% CO₂ humidified atmosphere at 37 °C and grown until 80% confluence for a maximum of 20 passages. SH-SY5Y cells were plated in
Chapter 5. Materials and methods

a 96-well plate at a concentration of 10,000 cells/well and treated for 24 h at 37 °C with the different α-synuclein species. After this, the cells were incubated with 0.5 mg/mL MTT 23 (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) in RPMI (Thermo Fisher Scientific, Waltham, MA) solution at 37 °C for 4 h and subsequently lysed with a solution of 20% SDS, 50% N,N-dimethylformamide, pH 4.7 at 37 °C for 2 h. Absorbance values of blue formazan were determined at 590 nm using a Clariostar plate reader (BMG Labtech, Aylesbury, UK).

Correlation between MTT reduction and α-helical content was assessed by linear regression, yielding an $R^2$ value of 0.29, p value of 0.0026, and 95% confidence intervals for the slope of -0.026, 0.0062. The high apparent toxicity of the G51D oligomer preparation with the highest α-helical content may be an outlier. However, omitting this point in the analysis still results in a significant correlation: $R^2 = 0.19$, p value = 0.019, and 95% confidence intervals for slope = -0.017, -0.0017.
5.2 Methods: Microfluidic methods for studying oligomers in protein aggregation reactions

5.2.1 Preparation of samples

Purification of α-synuclein N122C variant

The N122C variant of α-synuclein was purified as detailed in section 5.1, with the addition of 3 mM dithiothreitol (DTT) to all buffers, in order to prevent dimerisation through disulphide bonds. AlexaFluor dye-labelled oligomers were formed from labelled monomer, as described in section 5.1, and the final oligomer concentration (in monomer equivalents) estimated by absorbance at 495 nm, using a molar extinction coefficient of 72,000 M\(^{-1}\) cm\(^{-1}\).

Labelling of α-synuclein

The N122C variant of α-synuclein was conjugated to AlexaFluor-488 and 546 dyes via a maleimide linker at the cysteine residue. In order to remove DTT from the buffer, the protein was first buffer exchanged into PBS or 10 mM sodium phosphate buffer (pH 7.4) by the use of P10 desalting columns packed with Sephadex G25 matrix (GE Healthcare). The protein was then incubated with a 1.5-fold molar excess of Alexa-488 or Alexa-546 dyes with maleimide moieties (overnight, 4 °C on a rolling system). The labelled α-synuclein was finally isolated from the remaining free dye by size exclusion chromatography on a HiLoad\(^\text{TM}\) 16/600 column packed with pg 200 resin, at room temperature. The concentration of purified labelled monomer was estimated by the absorbance of the fluorophore (Alexa-488: ε of 72,000 M\(^{-1}\) cm\(^{-1}\) at 495 nm; Alexa-546: ε of 104,000 M\(^{-1}\) cm\(^{-1}\) at 546 nm).

Preparation of calcein

Calcein (Sigma-Aldrich) was dissolved in 6 M sodium hydroxide solution to a concentration of 752 mg/mL. The calcein was buffer exchanged into PBS by purification in P10 desalting columns packed with Sephadex LH20 matrix (GE Healthcare). The most concentrated fractions were collected and pooled. Following the adjustment of their pH to 7.4, the concentration of calcein (∼ 100 mM) was determined by UV-vis spectroscopy, using an extinction coefficient of 72,000 M\(^{-1}\) cm\(^{-1}\) at 490 nm.
Preparation of calcein-filled large unilamellar vesicles

POPS suspended in chloroform was purchased from Avanti Polar Lipids. 5 mg lipid was dried under nitrogen to a thin film and lyophilised overnight, and subsequently resuspended in 1 mL calcein solution (~100 mM). Resuspended lipids were subjected to 5 freeze-thaw cycles in liquid nitrogen. The resulting unilamellar vesicles were extruded through membranes with a pore size of 100 nm to yield a homogeneous size population of LUVs (Avanti Polar Lipids). In order to remove free calcein, the vesicle mixture was passed through multiple P10 centrifuge desalting columns packed with Sephadex G50 matrix (GE Healthcare), until the flow through was almost colourless (2-3 ×). The resultant concentration of LUVs was estimated based on the volume obtained following the centrifugation columns, assuming no vesicles were lost during the preparation process (confirmed by Dr Serene Chen in previous work).

Aggregation of labelled α-synuclein

For widefield epifluorescence μFFE experiments, Alexa-488-labelled N122C monomers were aggregated in eppendorf tubes (1 mL, 100 µM, 37 °C) under shaking conditions (200 rpm) in PBS, in the presence of 0.01% (w/v) sodium azide to prevent bacterial growth. 150 µL aliquots were withdrawn and centrifuged (10 min, 21,130 rcf, room temperature) to pellet large fibrillar components of the reaction mixture. The supernatant, containing monomers and soluble oligomers, was carefully removed and immediately injected onto the μFFE device. A small portion of the supernatant was retained for concentration determination, to monitor aggregation reaction progress.

For single molecule confocal experiments, Alexa-488-labelled N122C monomers (100 µL per well, 100 µM, PBS) were aggregated with a glass bead (Sigma-Aldrich) in 3881 plates (Corning) in a FLUOstar Omega plate reader (BMG Labtech) under shaking conditions (200 rpm, 37 °C) in the presence of 0.01% w/v sodium azide. Fluorescence emission was monitored at 520 nm, with an excitation wavelength of 485 nm.

5.2.2 Biophysical characterisation of samples

AUC, CD, DLS, FTIR, Native-PAGE, and TEM analyses were carried out as described in Section 5.1.
Fluorescence of calcein-filled large unilamellar vesicles

Calcein-filled vesicles (100 µM) were incubated with PBS or Triton-X (5 min, RT) before recording their fluorescence emission spectra between 500-600 nm, with an excitation wavelength of 495 nm, on a Cary Eclipse Fluorescence spectrophotometer (Agilent, Santa Clara, CA).

5.2.3 Preparation of microfluidic devices

Fabrication of masters

Device design was carried out using the software AutoCAD (Autodesk). The designs were printed onto acetate transparencies (Micro Lithography Services) to produce photolithographic masks. ∼3 mL SU-8 (3000 series: 3025 for 25 µm height for µFFE devices, or 3050 for 50 µm height for lipid disruption devices) photoresist was deposited onto a silicon wafer (3 inch diameter), following rinsing of the wafer in isopropanol (IPA) and subsequent drying under nitrogen stream. The wafer was spun (500 rpm, 5-10 s followed by 3000 rpm, 30s), coating it in a layer of SU-8 at the desired height. The wafer was then soft-baked on a hotplate (95 °C, 15 min for 25 µm height masters, or 25 min for 50 µm height masters). The photolithographic mask was then fixed in place over the coated wafer using a custom-made clamp. The wafer was exposed to UV light (365 nm, 60 s for 25 µm height masters, or 90 s for 50 µm height masters) using a custom-built LED-based setup, in order to crosslink SU-8 in regions exposed by the mask. The wafer was then hard-baked on a hotplate (95 °C, 5 min for 25 µm masters, or 10 min for 50 µm height masters). The wafer was then submerged in propylene glycol methyl ether acetate (PGMEA) (5 min) to dissolve uncrosslinked SU-8. Following rinsing in IPA and PGMEA and subsequent drying under nitrogen stream, the feature heights on the master were determined using a profilometer (DektakXT, Bruker). Masters for µFFE devices were made to be 25 µm high, while masters for lipid disruption devices were 50 µm in height.

Fabrication of microfluidic devices

Polydimethylsiloxane (PDMS) (Dow Corning, primer and base thoroughly mixed in a 1:10 (w/w) ratio) was applied to the master in a petri dish. In order to fabricate lipid disruption devices, a small amount of black carbon black nanopowder was evenly distributed by mixing into the PDMS, to minimise background fluorescence.
Figure 5.1: Fabrication of microfluidic masters and devices. Silicon wafers were coated in SU-8, soft-baked, and aligned with a photomask. Subsequent exposure to UV-light induced cross-linking of SU-8. Following hard-baking, the uncrosslinked SU-8 was washed away using PGMEA. The master was placed in a petri dish and PDMS poured over the master. PDMS was de-gassed, baked (65 °C, 1 h), and devices cut out from moulds. Following punching of holes using biopsy punches and cleaning, the devices were bonded to glass coverslips or coverslides. Devices were then plasma treated to create hydrophilic surfaces, and filled with water.

Following degassing to remove air bubbles, the devices were baked (65 °C, 1 h). Devices were cut out from the moulds and punched with holes for the connection of tubing (0.75 mm diameter) and electrodes (1.5 mm diameter - µFFE devices only), using biopsy punches. The devices were then cleaned of dust and debris by the application of Scotch tape, sonicated in IPA (5 min), and dried under nitrogen. Devices for use in widefield epifluorescence microscopy were bonded to glass slides, which were first cleaned with IPA and dried under nitrogen. Devices for use in confocal microscopy were bonded to glass coverslips (#1.5), which were first sonicated in 1% Hellmanex (5 min), followed by rinsing with milliQ water and ethanol, and dried under nitrogen. Bonding of PDMS devices to glass slides/coverslips was achieved in an oxygen plasma oven (15 s, 40% power). Prior to use, devices were treated to render the surfaces hydrophilic, by plasma treatment (500 s, 80% power). Following this exposure, the hydrophilicity of the surface was maintained by immediate filling of the device with milliQ water.
5.2.4 µFFE device operation

The µFFE device design used contains liquid electrodes (3 M KCl solution) to connect the electrophoresis chamber to the external electric circuit. These liquid electrodes are connected to the circuit via hollow metal electrodes made from bent syringe tips, which also constitute the outlets for the liquid electrodes. Gaseous products of electrolysis are removed from the device, without disturbing fluid flow, through these outlets [184]. Samples are flowed into the device at controlled flow rates by the use of syringe pumps (Cetoni neMESYS, Korbussen, Germany), connected to polytetrafluoroethylene (PTFE) tubing (0.012" inner diameter × 0.030" outer diameter, Cole-Parmer, St. Neots, UK). The electric field was applied by a benchtop power supply (Elektro-Automatik EA-PS 9500-06, Viersen, Germany) connected to the metal electrode outlets.

5.2.5 Widefield epifluorescence microscopy for µFFE

Widefield epifluorescence µFFE experiments were conducted and analysed by Dr William Arter.

Data acquisition

Experiments in PBS were conducted with the following flow rates: auxiliary buffer, electrolyte, monomer reference (4 µM Alexa-546-labelled N122C monomer in milliQ water) and sample flow rates of 1200, 250, 140 and 10 µL h⁻¹, respectively, for 15× reduction in buffer salt concentration for samples in PBS buffer. Samples in 10 mM sodium phosphate buffer were studied using devices in the absence of the desalting module, the sample flow rate used was 20 µM h⁻¹; auxiliary buffer and electrolyte flow rates were 1200 and 250 µL h⁻¹, respectively. In this case, 1 µM Alexa-546-labelled N122C monomer was mixed with 4 µM Alexa-488-labelled oligomer before injection into the device.

Microfluidic experiments were conducted using an inverted fluorescence microscope (Zeiss AxioObserver D1, Oberkochen, Germany), Alexa-488 and 546-labelled species were observed using appropriate filter sets (49002 and 49004, respectively, both Chroma Technology, Bellows Falls, VT) and camera (Evolve 512 CCD, Photometrics, Tucson, AZ).
Chapter 5. Materials and methods

Voltage efficiency determination

Device voltage efficiency was calibrated by comparison of current-voltage curves of the device operating under assay conditions and when filled with 3M KCl electrolyte. The gradient of the voltage plotted against the current yielded the electrical resistances of the overall device and the electrodes. The voltage efficiency was then calculated by the following relationship: efficiency = 1 - Ω\text{electrode}/Ω\text{device}. Efficiencies were found to be \sim 20%, affording electric fields equivalent to 200-267 V cm\(^{-1}\) for potentials of 300–400 V.

5.2.6 Confocal microscopy for µFFE

Description of the setup

Experiments were performed using a custom-built single molecule confocal fluorescence spectroscopy setup, which was designed and construct by Dr Georg Krainer. The 488 nm wavelength laser beam (Cobolt 06-MLD 488 nm 200 mW diode laser, Cobolt, Stockholm, Sweden) was coupled into a single-mode optical fibre (P3-488PM-FC01, Thorlabs, Newton, NJ) and collimated (60FC-L-4-M100S-26, Schäfter und Kirchhoff, Hamburg, Germany) before being directed into the back aperture of an inverted microscope body (Applied Scientific Instrumentation Imaging, Eugene, OR). The laser beam was then reflected by a dichroic mirror (Di03-R488/561, Semrock, Rochester, NY) and focused to a concentric diffraction-limited spot in the microfluidic channel through a high-numerical-aperture water-immersion objective (CFI Plan Apochromat WT 60\(\times\), NA 1.2, Nikon, Tokyo, Japan). Photons arising through fluorescence emission were detected using the same objective. Fluorescence was then passed through the dichroic mirror and imaged onto a 30 μm pinhole (Thorlabs), removing out of focus light. The signal was then filtered through a bandpass filter (FF01-520/35-25, Semrock), and focused onto a single-photon counting avalanche diode (APD, SPCM-14, PerkinElmer Optoelectronics, Waltham, MA). Photons were recorded using a time-correlated single photon counting (TCSPC) module (TimeHarp 260 PICO, PicoQuant, Berlin, Germany) with 25 ps time resolution. Single-photon counting recordings were obtained using custom-written Python code.
5.2. Methods: Microfluidic methods for studying oligomers in protein aggregation reactions

Data acquisition

µFFE devices were fabricated using thin glass coverslips (#1.5), and experiments performed using flow rates of 1000, 200, 140, 10 µL h⁻¹ for the auxiliary buffer (15× diluted PBS in milliQ water), electrolyte, desalting milliQ water, and sample, respectively. 1 nM Atto-488 dye was added to the 3 M KCl electrolyte, in order to serve as markers for the electrophoresis chamber dimensions. A custom Python script was written by Dr Quentin Peter to control the simultaneous movement of the stage and acquisition of photon data. Two scanning modes were used: continuous and stepping. In continuous mode, the stage was moved at a constant speed (20-200 µm s⁻¹) across the electrophoresis chamber from cathode to anode, perpendicular to the fluid flow direction. In stepping mode, the stage was moved along the same trajectory as in continuous mode, but was stopped at defined positions along this path, and photon count rate data obtained for periods of seconds to minutes at each position.

Voltage efficiency determination

For pure monomer experiments, the voltage efficiency \( E \) of the device was determined by comparison of the measured current under the experimental conditions \( (I_{\text{exp}}) \) and the chip when filled with 3 M KCl \( (I_{\text{KCl}}) \):

\[
E = 1 - \frac{I_{\text{exp}}}{I_{\text{KCl}}} \quad (5.8)
\]

The field \( F \), in units of V cm⁻¹, was then calculated by the following equation:

\[
F = E \cdot V \cdot \frac{1}{W} \quad (5.9)
\]

where \( V \) is the voltage applied (in V), and \( W \) is the width of the electrophoresis chamber (in cm), in this case 0.3.

The mobilities \( M \), in units of m² V⁻¹ s⁻¹ can therefore be calculated:

\[
M = \frac{D}{F \cdot T \cdot 0.01} \quad (5.10)
\]

where \( D \) is the deflection distance (in m) and \( T \) is the time the sample spends in the electrophoresis chamber before detection.

Due to the long acquisition times relative to the widefield epifluorescence experiments to obtain sufficiently high volumes of oligomer data, the device efficiency
may change over the course of experiment. Since a large amount of monomer was present in kinetic oligomer experiments, the calibration was therefore performed by normalisation of mobilities to the monomer, which has a highly reproducible electrophoretic mobility of \(-1.43 \cdot 10^{-8} \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}\).

### 5.2.7 Analysis of single molecule data

A custom-written Python script (written by Raphaël Jacquat) was used to extract single molecule events from the interphoton time data. For extracting single monomer events, cut-offs of 0.005 µs and 17 were used for the maximum interphoton time, and minimum total number of photons. For detecting single kinetically-trapped oligomer events, values used were 0.005 and 100, while 0.0005 and 400 were used for kinetic oligomers. In all cases, a Lee filter of four was used. The intensities of single molecule events were thus determined as a function of channel position and thereby electrophoretic mobility. The apparent sizes of the oligomer in terms of number of monomer subunits were obtained by division of the intensities by the mean burst intensity observed for monomeric α-synuclein.

### 5.2.8 Lipid disruption device operation

#### Data acquisition

Flow rates of 50, 50, and 20 µL h\(^{-1}\) were used for the sample, vesicles, and 5% triton-X, respectively, using syringe pumps (neMESYS, Cetoni) connected to the device by PTFE tubing (0.012" inner diameter × 0.030" outer diameter, Cole-Parmer). Following 15 minutes of incubation, in order to ensure the entire device was re-equilibrated with the new sample, fluorescence images were acquired using an inverted fluorescence microscope (Zeiss AxioObserver D1), fitted with a filter set with the appropriate wavelengths for calcein (excitation/emission maxima at 495 and 505 nm) (49002, Chroma Technology), 4× objective, and camera (Evolve 512 CCD, Photometrics). Image acquisition times of 500-2000 ms were used, depending on the concentration of vesicles.
Data analysis

The total fluorescence intensities across the channels were quantified by densitometry analysis (ImageJ). For analysis of lipid disruption kinetics, the position was converted to incubation time by considering the channel dimensions (400 µm width × 50 µm height). The relative increase in fluorescence between timepoints was determined by comparing fluorescence intensities of adjacent channel positions, normalised to the fluorescence intensity of the previous channel.
5.3 Methods: Optimisation of experimental design

The prediction method was implemented using custom-written Python software, written by myself with input from Dr Georg Meisl. Binding data presented in this thesis were obtained by Dr Tom Scheidt (Aβ/antibodies), Ashley Priddey (HLA), Matthias Schneider (HLA and SARS-CoV-2), and Dr Itzel Condado Morales (SARS-CoV-2) in collaboration with Professor Sara Linse (Lund University, Sweden), Dr Vasilis Kosmoliaptsis (University of Cambridge, UK), Professor Adriano Aguzzi (University Hospital Zürich, Switzerland), and Fluidic Analytics (Cambridge, UK). Protein labelling in the case of the SARS-CoV-2 work was performed by Dr Viola Denninger and Dr Sebastian Fiedler (Fluidic Analytics).

5.3.1 Characterising therapeutic antibodies against Aβ isoforms

Labelling of Aβ

AβMC1-42, a mutant with an extra cysteine reside positioned between the starting methionine and asparagine, was used for fluorophore labelling [296]. Lyophilised purified peptide monomer was dissolved in 6 M GdmCl, 10 mM DTT, pH 8.5, and incubated (1 h, RT). The monomer was isolated by gel filtration in 20 mM sodium phosphate buffer, pH 8.0. Two molar equivalents of Alexa-647 C2 maleimide (Thermo Fisher A20347) were added from a concentrated stock in DMSO, and the mixture incubated in darkness (ON, on ice). Labelled monomer was isolated from free dye by two successive rounds of gel filtration in 20 mM sodium phosphate buffer, pH 8.0.

Labelling of antibodies

Antibodies were labelled by mixing with two molar equivalents of Alexa-647 N-hydroxy succinimidyl ester (Thermo Fisher A20006) after gel filtration of each antibody in PBS. Mixtures were incubated (2 h, 4 °C), followed by two successive rounds of gel filtration to remove free dye. The absence of free dye was confirmed using microfluidic diffusional sizing with fluorescence detection using a Fluidity One-W instrument (Fluidic Analytics).
5.3. Methods: Optimisation of experimental design

Aβ/antibody binding

Binding interactions in solution were assessed through diffusion measurements under laminar flow in a microfluidic device using a Fluidity One-W instrument (Fluidic Analytics). Alexa-647-labelled Aβ42 was used to study interactions between peptide monomers and antibodies, and Alexa-647-labelled antibodies were used to study the interactions between antibodies and Aβ42 fibrils. Alexa-647-Aβ42 monomers were isolated by SEC and flash-frozen in liquid nitrogen or kept on ice until use. Unlabelled Aβ42 fibrils were formed under quiescent conditions at 37 °C in PEGylated plate wells (Corning 3881) from monomer, and then sonicated (20/20 s break cycles) for 6 min on ice, and finally shaken for 30 min at 1800 rpm to provide fibrils of approximately 50 nm length.

The labelled species were combined with the unlabelled binding partner, and incubated (15 min, RT) before the diffusion measurements at 27-28 °C in PBS, pH 7.8. The sample was injected into one half of the channel, and buffer into the other half. Following passage down a microfluidic channel, the fluorescence intensities in the two halves of the channel at the device outlets were analysed separately. Measurements for each condition were repeated 3-7 times. The fluorescence intensities were then analysed by Bayesian inference (see below) to determine the binding parameters.

5.3.2 Antibody quantification for the determination of transplant compatibility

HLA monomers were obtained through the NIH Tetramer Core Facility in Emory, Atlanta, US, in PBS, PBS and AlexaFluor 647 was purchased from Thermo Fisher Scientific Inc., Waltham, US, Human IgG ab205198 was from Abeam, Cambridge, UK, All other chemicals were from Sigma Aldrich, All PBS was supplemented with NaN3 (0.02% (w/v)).

Labelling of HLA

1 molar equivalent of AlexaFluor-647 N-hydroxysuccinimide ester (Thermo Fisher), dissolved in DMSO, was added to HLA in 0.1 M NaHCO3 buffer, pH 8.0, and the mixture was incubated (1h, RT) while protected from light. The labelled HLA was purified from free dye by size exclusion chromatography in PBS, with an approximate yield of 80%, and degree of labelling between 0.33 and 2.25.
Chapter 5. Materials and methods

HLA/antibody binding in human serum

All binding experiments in human serum were carried out using human serum from non-sensitised volunteers, which did not contain HLA-specific antibodies, or from a sensitised kidney transplant patient. Alexa-647-labelled HLA was mixed with non-sensitised serum spiked with specific antibodies or the sensitised patient sample, and incubated (30 min, RT). The effective hydrodynamic radius of the HLA was measured by diffusional sizing, in a Fluidity One-W machine (Fluidic Analytics).

5.3.3 Quantification of severe acute respiratory syndrome coronavirus 2-reactive antibodies in patient sera

Acquisition of patient serum samples

Ethical and biosafety statement: 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.

EDTA plasma from healthy donors and from convalescent individuals was obtained from the Blutspendedienst (blood donation service) Kanton Zürich from donors who gave consent for their samples to be used for conducting research. Samples from hospitalised patients with COVID-19 were collected at the University Hospital Zürich from patients who signed an informed consent form.

Labelling of SARS-CoV-2 receptor binding domain

RBD protein was purchased from SinoBio (Eschborn, Germany) and labelled using classical amine coupling based on NHS chemistry with AlexaFluor-647 dye (Thermo Fisher). RBD protein and 3 molar equivalents of AlexaFluor-647 N-hydroxysuccinimide ester were combined in 0.1 M NaHCO₃, pH 8, and incubated (1 h, RT), while protected from light. Alexa-647-labelled RBD was separated from free dye by size exclusion chromatography in PBS.
RBD/antibody binding in patient samples

Patient serum samples were incubated with Alexa-647-labelled RBD in PBS with 0.05% Tween-20, in a total volume of 20 µL, for 40 min at room temperature, before measurement of the effective hydrodynamic radius of the RBD protein by diffusional sizing on a Fluidity One-W (Fluidic Analytics).

5.3.4 Analysis of microfluidic diffusional sizing binding data

The intensities of the "diffused" (distal: \(I_d\)) and "undiffused" (proximal: \(I_u\)) channels (Figure 4.3) were corrected for the serum autofluorescence in experiments involving serum samples using a custom-written Python script provided by Dr Sean Devenish (Fluidic Analytics). The fraction of labelled protein to diffuse into the diffused channel \(f_d\) can be determined by:

\[
f_d = \frac{I_d}{I_d + I_u}
\]  
(5.11)

For the binding of \(L\) (labelled species) to \(U\) (unlabelled species) to form \(C\) (complex) with dissociation constant \(K_d\), \(I_d\) and \(I_u\) can be expressed as:

\[
I_d = \kappa([C]\rho_b + [L]\rho_f) = \kappa([C][\rho_b + ([L]_0 - [C])\rho_f])
\]  
(5.12)

\[
I_u = \kappa([C](1 - \rho_b) + [L](1 - \rho_f)) = \kappa([C](1 - \rho_b) + ([L]_0 - [C])(1 - \rho_f))
\]  
(5.13)

where \(\kappa\) is a constant that relates the concentration of \(L\) to the fluorescence intensity observed, \(\rho_f\) and \(\rho_b\) are the fractions of free and bound \(L\) that diffuse into the distal channel, respectively, \([L]\) and \([C]\) are the equilibrium concentrations of \(L\) and \(C\), respectively, and \([L]_0\) is the total concentration of \(L\) (i.e. \([L] + [C] = [L]_0\)). We can therefore express \(f_d\) as:

\[
f_d = \frac{[C]\rho_b + ([L]_0 - [C])\rho_f}{[L]_0} = \rho_f + \frac{[C]}{[L]_0}(\rho_b - \rho_f)
\]  
(5.14)

Solving the binding equilibrium equation, we obtain the following expression for \([C]\):

\[
[C] = \frac{[L]_0 + [U]_0 + K_d}{2} - \sqrt{([L]_0 + [U]_0 + K_d)^2 - 4[L]_0[U]_0}
\]  
(5.15)
We therefore have $f_d$, our observable, as an analytical function of $[L]_0$, $[U]_0$, $K_d$, $\rho_f$, and $\rho_b$. In most cases, the binding stoichiometry is not 1:1 $L:U$, and we therefore adjust our expression for $[C]$ accordingly. $U$, $L$, and $C$ now correspond to binding sites, rather than molecules.

$$[C] = \frac{\lambda [L]_0 + \nu [U]_0 + K_d - \sqrt{(\lambda [L]_0 + \nu [U]_0 + K_d)^2 - 4\lambda \nu [L]_0 [U]_0}}{2}$$ (5.16)

where $\lambda$ and $\nu$ are the number of binding sites per $L$ and $U$ molecule, respectively, if the absolute concentrations of $L$ and $U$ are known. For example, if $L$ is a labelled antibody with two binding sites, we would have $\lambda = 2$, and $[L]_0$ = concentration of antibody, and $\lambda[L]$ therefore corresponds to the concentration of binding sites. If the total concentration of $U$ is unknown and to be determined, $[U]_0$ corresponds to the total stock concentration of $U$, and $\nu$ to the fraction of $U$ used in the measurement.

We thus have $f_d(\rho_f, \rho_b, \nu, [U]_0, \lambda, [L]_0, K_d)$.

### 5.3.5 Bayesian inference details

Bayes’ theorem states that:

$$P(\text{parameters}|\text{data}) \propto P(\text{data}|\text{parameters})P(\text{parameters}) \quad (5.17)$$

where $P(\text{parameters}|\text{data})$ is know as the posterior, $P(\text{data}|\text{parameters})$ as the likelihood, and $P(\text{parameters})$ as the prior. The prior probability distribution is an expression of our information about the system before we acquire any measurement data. For $\rho_f$ and $\rho_b$, we assume the prior is flat in linear space, whereas for the $K_d$ and $\nu$, a prior that is flat in logarithmic space is more appropriate, to reflect the scale invariance of the problem.

We assume our experimental data ($f_d$) to be normally distributed about the true value; our likelihood function is therefore a Gaussian centred on the theoretical measurement value:

$$P(\text{data}|\text{parameters}) \propto \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{N} (y_i - f_d(\rho_f, \rho_b, \nu_i, [U]_{0,i}, \lambda, [L]_{0,i}, K_d))^2\right] \quad (5.18)$$

where $y_i$ is the fraction of diffused labelled component in the $i^{th}$ measurement, and $\nu_i$, $[U]_{0,i}$, and $[L]_{0,i}$ are the concentrations or fractions (as applicable) in the
$i^{th}$ measurement. In order to define an appropriate standard deviation, $\sigma$, for each dataset, the standard deviation of repeats of each measurement was calculated, and the maximum of those values was then used as a global standard deviation value for that dataset.
Bibliography


[45] Rabizadeh, S., Bitler, C. M., Butcher, L. L. & Bredesen, D. E. Expression of the low-affinity nerve growth factor receptor enhances beta-amyloid peptide


[97] Jo, E., Fuller, N., Rand, R., St George-Hyslop, P. & Fraser, P. E. Defective membrane interactions of familial parkinson’s disease mutant a30p α-


Bibliography


[263] Bohrmann, B. *et al.* Gantenerumab: A novel human anti-a\(\beta\) antibody demon-
strates sustained cerebral amyloid-\(\beta\) binding and elicits cell-mediated removal

[264] Salloway, S. *et al.* Two phase 3 trials of bapineuzumab in mild-to-moderate

[265] Sevigny, J. *et al.* The antibody aducanumab reduces a\(\beta\) plaques in alzheimer’s

[266] Honig, L. S. *et al.* Trial of solanezumab for mild dementia due to alzheimer’s

[267] Arndt, J. W. *et al.* Structural and kinetic basis for the selectivity of aducanumab

[268] Linse, S. *et al.* Kinetic fingerprints differentiate the mechanisms of action of
Bibliography


