INTERNATIONAL DOCTORAL SCHOOL OF THE USC

Sara
Illodo Brea

PhD Thesis

# Study of biomolecular self-assembly and host-guest association using spectroscopic and microscopic techniques

Santiago de Compostela, 2024

**Doctoral Programme in Chemical Science and Technology**

ESCOLA DE DOUTORAMENTO
INTERNACIONAL DA USC

DOCTORAL THESIS

# Study of biomolecular self-assembly and host-guest association using spectroscopic and microscopic techniques

Author

Sara Illodo Brea

Supervisor/s: Mª de la Flor Rodríguez

Mª de la Merced Novo Rodríguez

Tutor: Mª de la Merced Novo Rodríguez

**PHD PROGRAMME IN CHEMICAL SCIENCE AND TECHNOLOGY**

SANTIAGO DE COMPOSTELA

# Acknowledgments

This PhD thesis was funded by the Xunta de Galicia predoctoral contracts program and by the Spanish Government I+D+i fundings.

I would like to express my gratitude to all those people who made this thesis possible and who helped me along the way.

First, to my supervisors Prof. Flor Rodríguez Prieto and Prof. Mercedes Novo Rodríguez for guiding me and providing me of all the teaching and advice that I needed during this thesis, and even before the thesis started. I would also like to thank all the other professors from our research group who have helped me along the way: Prof. Carmen Ríos Rodríguez, Prof. Wajih Al-Soufi and Prof. Manuel Mosquera.

I want to express my most sincere words of gratitude to Prof. Sang-Hak Lee from Pusan National University, who gave me the opportunity to work under his supervision during my research stay. Thank you for all the guidance you gave, not only academically but also personally. I will always be indebted to you.

To my lab colleagues, from the PhD students to all the undergraduate students that have worked in our lab, I want to say thank you for making the long experiments easier and for the warm work environment I was surrounded with. I also want to express my gratitude to Cibrán and Ángel, for all the insights and advice they have given me. I have learnt a lot from all of you and I sincerely hope all the best to each and one of you. I cannot forget about the Bionanoscopy Lab researchers at PNU that welcomed into their lab with open arms and helped me along the way. My stay at Busan was unforgettable because of all the help I received from you.

On a more personal note, thank you to all my friends. Carmen, Elena, Laura and Pablo, who have known me for a lifetime and had always been there for me. They have heard me talk about this thesis so many times, even if they didn´t understand a word, that it is also theirs. I also want to thank Andrés, Stella, Jesús Fernando, Samira, Jose and Víctor, whom I met thanks to chemistry and have made these years more beautiful. To Jiwon and Hyeryeong, who have become a great support despite the distance, I want to thank them for their unconditional support.

Lastly, but most importantly, I want to thank my family who has always been supportive of my decisions and believed in me blindly. Especially, my mom, Saky and siblings - Mia, Olaya and Iago - who gave me the strength I needed to keep on going. This thesis was possible because of you.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

## Index



USC DE COMPOSTELA

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

# List of abbreviations

| | |
|---|---|
| AD | Alzheimer's Disease |
| Aβ | Amyloid-beta |
| Aβ40 | Amyloid-beta(1-40) |
| Aβ42 | Amyloid-beta(1-42) |
| ALS | Amyotrophic Lateral Sclerosis |
| APP | Amyloid Precursor Protein |
| AICP | APP Intracellular Domain |
| BSA | Bovine Serum Albumin |
| CBTOH | 2-cyano-6-hydroxybenzothiazole |
| CD | Circular Dichroism |
| CSBII | Conserved block region II |
| FCS | Fluorescence Correlation Spectroscopy |
| FRET | Förster Resonance Energy Transfer |
| G4 | Guanine quadruplexes |
| HSA | Human Serum Albumin |
| IDP | Intrinsically Disordered Protein |
| Aβ40* | Labelled HiLyte-488-Amyloid-beta(1-40) |
| Aβ42* | Labelled HiLyte-488-Amyloid-beta(1-42) |
| LLPS | Liquid-liquid Phase Separation |
| Phe, F | Phenylalanine |
| PE | Plasma Exchange |
| R123 | Rhodamine 123 |
| Spm | Spermine |
| Spd | Spermidine |
| ThT | Thioflavin T |
| Trp, W | Triptophane |
| Tyr, Y | Tyrosine |
| TICT | Twisted Intramolecular Charge Transfer |

More specific abbreviations will be presented and explained in their respective sections.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# Summary

In this PhD thesis, different biomolecules linked to neurodegenerative diseases, especially Alzheimer's disease (AD), were studied. These types of conditions entail a public health risk globally, as their development is linked to aging. Thus, the prevalence is high among the population over 65 years old, and especially in individuals over 80 years old. The inherent conditions of these pathologies, along with the global aging population, make these diseases into a societal problem.

Biomolecules are essential components in biological processes due to their ability to form complex structures through self-assembly processes or specific intermolecular interactions. These processes are key to the correct functioning of organisms, but their incorrect regulation can lead to the development of numerous diseases. An example of this is neurodegenerative diseases such as Alzheimer's or Parkinson's, whose appearance is highly linked to the incorrect regulation of proteins that undergo aggregation or misfolding processes. However, there are cases where association processes can be used for therapeutic purposes by inhibiting pathological processes of protein aggregation or interaction. Therefore, understanding the mechanism underlying self-assembly and aggregation processes is essential for the development of new therapeutic strategies. In this way, during this thesis, the study of nucleic acids and proteins involved in Alzheimer's disease was carried out.

Spectroscopic techniques are widely used tools in diverse research fields. From a molecular perspective, these techniques allow the detection and quantification of molecules and molecular processes. The application of these techniques to study the dynamics of association and self-assembly of biomolecular processes is of particular interest, as they allow both quantitative thermodynamic and kinetic studies. For this reason, this thesis was conducted using different spectroscopic techniques, from the most conventional ones (such as UV-Vis absorption or steady-state fluorescence) to more innovative techniques that allow reaching the scale of individual molecules (such as FCS or fluorescence microscopy).

**Section 5.1: Spectroscopic Characterisation of Mitochondrial DNA G-Quadruplexes.**

In the first part of this work, which corresponds to the first section of **Results and discussion**, guanine quadruplexes (G4) formation was studied both in DNA, RNA and

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

DNA:RNA hybrids sequences. The self-assembly that DNA and RNA sequences present in our organism can undergo leads to the formation of structures known as G4. The role these structures play is still under study, but many studies suggest that they are strongly linked to transcription, translation and gene expression regulation. This implies that they could be potential key factors in the development of numerous diseases, among which is also AD. In this PhD thesis the results obtained using circular dichroism (CD) as the main technique are presented. CD is a widely used technique in the biochemistry field due to its ability to distinguish between different secondary structures biomolecules can adopt. Complementarily, fluorescence and gel electrophoresis studies were performed for CD data corroboration. Among the used fluorescence techniques are both steady-state and time-resolved fluorescence emission and anisotropy. Using CD allowed us to carry out a quantitative study of the G-quadruplexes formation dynamics. This way, we elucidated how sequence length, its nature and even the presence of an initial guanine 1 bp upstream could impact both the formation and the adopted topology. We also investigated the role of monovalent cations ($Na^+$ e $K^+$) in G4's stabilisation and obtained the equilibrium constants for G4 formation in both cases.

The presented results demonstrate G4 formation in the studied DNA sequences, adopting a parallel topology. They also corroborate that a higher number of guanines yields more G4 formation, and also that the presence of the upstream initial guanine plays a crucial role in the self-assembly impacting significantly the degree to which formation takes place. Our investigations also evidence the role sodium and potassium cations have in the self-assembly process. Both cations promote a conformational change from a parallel topology to a highly ordered parallel topology. This was corroborated by fluorescence emission and anisotropy. These experiments showed that the binding of the fluorophore to the second topology led to a restriction of the rotation of the molecule, forcing it to adopt a twisted structure, while when binding with the first topology the fluorophore adopted a planar conformation. The comparison between both cations evidenced that potassium had a higher influence than sodium, promoting the rigid G4 formation to a higher degree. Studies of RNA and DNA:RNA hybrid sequences also presented evidences of mainly parallel G4 formation. However, for RNA sequences we observed a second structure in CD spectra, which could be attributed to anti-parallel or hybrid topologies.

**Section 5.2: Critical aggregation concentration and reversibility of Aβ(1-40) oligomers.**

In the second part of this PhD thesis, which includes **section 5.2** to **5.4**, the investigation focused the aggregation of amyloid peptides. Amyloid peptides are considered one of the main precursors of AD. The formation of amyloid aggregates, especially of early oligomers, was suggested as a crucial in neurodegeneration development. In our organism exit two

USC UNIVERSIDADE

main types of amyloids, Aβ40 and Aβ42. Even though Aβ40 is the most abundant type of amyloid, Aβ42 is the most studied due to its high aggregation tendency. In this part of the thesis, the study of Aβ40 early aggregates using FCS allowed us to characterise the aggregation process, obtaining both the critical aggregation concentration and the size distribution of the formed aggregation at different amyloid concentrations. In order to do this, FCS correlation curves were analysed to obtain the amyloid's diffusion times at different concentrations, which allowed for differentiation between monomers and aggregates. From the obtained diffusion times and the contribution of each one to the correlation curve, we applied a complex data analysis and obtained the degree of aggregation and the mean size of the formed aggregates for each sample. We also studied the reversibility and stability of the aggregation process. Throughout **section 5.2** a comparison between the aggregation mechanism obtained for Aβ40 and the mechanism previously reported by our group for Aβ42.

Our studies evidenced the differences between aggregation of Aβ40 and Aβ42. First, our results showed a two-step aggregation process for Aβ40, contrary to the single-step process obtained for Aβ42, obtaining two critical aggregation concentrations (cac). In Aβ40, the first step leads to aggregates similar in size and shape to the ones obtained in Aβ42. However, this takes places at 10 times higher concentrations for Aβ40. Then, aggregation can grow as the concentration increases, so that after the second phase has finished, the size of the aggregates is significantly bigger. Our reversibility studies demonstrated that the formed aggregates can disaggregated upon dilution, but this disaggregation does not happen completely for those samples in which the formed aggregates are the smallest one, as they need longer times after dilution to form monomers. On the other hand, stability studies, in which we studied the same samples at different times after preparation, evidenced that the aggregates do not undergo significant variations in size nor composition even after one week. This stability is also different that the one seen for Aβ42.

**Section 5.3: Amyloid autofluorescence as method for amyloid's early aggregation detection.**

As early aggregates are linked to AD, it is necessary to find tools that help to selectively detect amyloid peptides and allow for differentiation between monomers and aggregates. To this day, the most widespread fluorescent probe for these studies is Thioflavin T (ThT), as it presents important spectral variations upon binding to amyloid peptides, especially when bound to fibrils. However, this fluorophore has a complex mechanism that complicates data analysis. Moreover, ThT can also bind to other proteins with relatively small association constant similar to the ones of association with amyloids. Therefore, finding fluorescence probes with improved characteristic is still of great interest. Amyloid peptides have within its sequence one tyrosine residue, which confers them

autofluorescence. The intrinsic fluorescence of the tyrosine residue could be a potential tool for amyloid detection, being as well a simple methodology that does not need additional molecules for emission studies. For this reason, we carried out a study on amyloid's autofluorescence as potential toll for amyloid detection and distinction between monomers and aggregates. This study was done using fluorescence emission as main technique and FCS as a complementary technique.

The autofluorescence study evidence the potential use of the intrinsic emission of the Aβ40's tyrosine residue as a detection method. Amyloid mission and excitation spectra of mainly monomeric and aggregated samples showed sufficient intensity variations and spectral shifts in the emission maxima to be used for discerning whether amyloid has undergone aggregation or not. When amyloid is in its monomeric form, the emission spectrum has a singles emission band around 350 nm. This band shifts towards lower wavelengths (303 nm) upon aggregation. Moreover, an emission band at higher wavelengths (500 nm) was also observed, which can be attributed to an aggregation induced emission as its contribution is higher in more aggregated samples. Sample's aggregation was corroborated with FCS measurements. Taking into consideration the obtained results, we proposed a model for the photophysical mechanism of tyrosine, in which it undergoes a deprotonation on the excited state when the peptide has not yet aggregated. This deprotonation is hindered when amyloid aggregates are formed. This could be due to variations on the local environment of tyrosine, which could be less exposed to the aqueous solution after aggregation hampering deprotonation. However, the obtained data are not of sufficient quality for quantitative study of the process and, therefore, performed investigation is only semi-quantitative and needs more experiments to better clarify the behaviour of the emission band at higher wavelengths. Thus, the proposed model must be considered as a hypothesis that can successfully explain our data but must be corroborated with further experiments.

**Section 5.4: Electrostatic interactions impact on amyloid-beta aggregation.**

Amyloids, when in physiological conditions, are negatively charged proteins. In our organism exist several small, charged biomolecules with positive net charged in physiological pH, such as polyamines. Non-covalent interactions are of great importance in numerous intermolecular association processes. For this reason, the presence of molecules like polyamines could impact the aggregation of amyloids by electrostatic interactions that happen because of the difference in charge of both molecules. These interactions are also studied as one of the causes of liquid-liquid phase separation (LLPS). Several studies evidence that in proteins that undergo aggregation like amyloids, the aggregation process happens through LLPS. This way, condensates where a local environment with high protein concentration would be formed, promoting aggregation inside them. Considering this, in this PhD thesis amyloid aggregation was studied through

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

a LLPS perspective. In this word we studied by fluorescence microscopy if amyloid aggregation also occurs by LLPS. Then, we also investigated the influence of electrostatic interaction between amyloid and charged biomolecules (spermine, spermidine and magnesium salt) could have in the aggregation process. This study was done at different aggregation phases: before and after fibril formation. Microscopy images were treated with SRRF, which allows for obtention of super-resolution images without super-resolution equipment.

By using the described technique, amyloid samples at different incubation times were studied, which evidenced how Aβ42 aggregates until fibril formation through LLPS. That is, through circular condensates formation, inside which aggregates were formed, that disappeared as fibril formation took place. On the other hand, our results also demonstrated the influence spermine, that has a net charge of +4, has on fibril formation. Addition of this polyamine to samples with already formed fibrils resulted in the appearance of condensates and progressive disappearance of fibrils, with both incubation time and spermine concentration dependency. When the added molecules were spermidine (with a net charge of +3) or magnesium chloride (with a +2 net charge), this process was not observed as clearly. In the case of spermidine no variation was observed under the studied conditions, neither with concentration nor time, which would suggest the need of higher charges for these charge-charge interactions to impact aggregation. However, in the studies done with magnesium condensates formation was seen at the highest concentration conditions and longest incubation times, which indicates that for smaller charge is not as promoted. These studies also suggest that charge might not be the only factor for LLPS to happen. The investigation of the size of the condensates revealed no variation in size and form, neither with concentration nor incubation time. This hints towards reversibility of fibrillation not being complete. At last, we also carried out studies before fibril formation that showed small variation depending on concentration and incubation time. These variations were both of fibril and condensates presence, as well as of the size of the condensates. However, as the observed tendencies were not clear enough more studies at different incubation times and concentration would be needed for more solid conclusions to be drawn.

**Section 5.5: 2-cyano-6-hydroxybenzothiazole as fluorescence probe for human serum albumin detection.**

The last part of this thesis focused on the study of blood plasma transport proteins, as they play a crucial role in in amyloid transport from the blood stream to the cerebrospinal fluid. This specific function human serum albumin has is the basis for an AD treatment through blood plasma exchange (PE) with human serum albumin reposition (known as AMBAR). This treatment is in clinical trials. Moreover, recent studies suggest that the association between HSA and amyloid peptides favors amyloid aggregation inhibition,

UAB UNIVERSITAT AUTÒNOMA DE BARCELONA

which highlights the protein role in potential therapeutic methodologies for AD. Due to HSA involvement with AD the necessity of finding fluorescence probes that allow for selective detection of HSA and amyloid, or of the binding between both molecules, becomes evident. As explained before, the use of Thioflavin T although very common, has important drawbacks. 2-cyano-6-hydroxybenzothiazole (CBTOH) is a fluorescence molecule that has already been characterized by our group in aqueous solution and other solvents. This study revealed the photoacid behaviour of the fluorophore, which undergoes proton transfer in the excited state. In addition, it also demonstrated the spectral variation that CBTOH can present depending on the polarity of the media it is in, which makes the molecule a candidate for biomolecule's detection. On the other hand, bovine serum albumin (BSA) is commonly used in biological studies as a model of HSA. The reason this is done this way is due to the difficulties HSA obtention and purification presents, which makes it also less economically accessible than BSA. This way, we studied CBTOH as a potential fluorescence probe for serum albumin detection by both UV-vis and fluorescence emission and anisotropy spectroscopic techniques. Both steady-state and time-resolved fluorescence emission experiments were carried out. The association studies between the fluorophore and both serum albumins allowed us to propose the mechanism underlying the host-guest association and to obtain the equilibrium constant with both proteins.

As the first step of the association between CBTOH and blood plasma proteins, the fluorophore's acid-base equilibrium was studied in physiological conditions of ionic strength. This revealed a slight variation towards lower $pK_a$ values than the ones described for aqueous solutions without salt addition. Then, the association studies were carried out and the results evidenced the effective binding between the dye and both proteins. The association can be observed as both absorption and emission spectra present intensity variation and spectral shifts upon protein's concentration increase, allowing for monitoring of the association process by absorption and emission spectroscopy. In both cases, the binding process can be explained with a 1:1 host-guest association model, which can reproduce our data accurately. In this regard, differences between proteins were observed. At physiological pH, CBTOH can exist in solution in anionic and neutral forms, and both species can bind to serum albumins. Nonetheless, fluorescence emission of the neutral form cannot be observed as it undergoes a rapid proton transfer in the excited state, being the anionic form the only one presenting emission. Our studies revealed that the proton transfer does not happen, or happens to a small degree, when CBTOH is bound to BSA, but it does take places when bound to HSA. This is evidence that, although used as analogous proteins, the differences in their sequences confer them different properties. Thus, they present different binding affinities and slightly different binding sites. The variations can also be observed in the obtained association constant value, which are nonetheless of the same order of magnitude. Overall, the present results

USC UNIVERSIDADE

demonstrate CBTOH characteristics to be used as probe for blood plasma protein detection. In addition, as there exist spectral differences between the formed complexes with each protein, it can also be used for identifying the protein to which it is bound. This also suggest that it could present spectral variation when bound to other proteins, such as amyloid. Moreover, as the differences can be seen both in absorption and emission, a combination of both techniques with a ratiometric analysis of the spectra could be of use in detection quantification.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# Resumo

Nesta tese doutoral estudáronse diferentes biomoléculas ligadas a enfermidades neurodexenerativas, especialmente á enfermidade de Alzhéimer (EA). Este tipo de doenzas supoñen un risco para a saúde pública a nivel global, xa que a súa aparición está ligada ao envellecemento, polo que a súa frecuencia é alta na poboación maior de 65 anos e, especialmente, en persoas maiores de 80 anos. Os propios condicionamentos destas patoloxías, xunto co proceso de envellecemento poboacional que se está a vivir a escala global, converten este tipo de enfermidades nun problema para a sociedade.

As biomoléculas son pezas esenciais nos procesos biolóxicos grazas á habilidade que presentan para formar estruturas complexas mediante procesos de autoensamblaxe ou interaccións especificas intermoleculares. Estes procesos son chave para o correcto funcionamento dos organismos, mais a súa incorrecta regulación pode provocar o desenvolvemento de numerosas doenzas. Un exemplo disto son as enfermidade neurodexenerativas como o Alzhéimer ou o Parkinson, cuxa aparición está altamente vinculado a un regulamento incorrecto de proteínas que sofren procesos de agregación ou pregamentos inadecuados. Porén, existen casos nos que procesos de asociación poden ser empregados con fins terapéuticos inhibindo procesos patolóxicos de agregación ou de interacción de proteínas. Por iso, entender o mecanismo polo cal ocorren os procesos de autoensamblaxe e agregación é imprescindible para o desenvolvemento de novas estratexias terapéuticas. Deste xeito, durante esta tese realizouse o estudo de ácidos nucleicos e proteínas involucradas na enfermidade de Alzhéimer.

As técnicas espectroscópicas son ferramentas moi empregadas en diversos campos de investigación. Dende o punto de vista molecular, son técnicas que permiten a detección e cuantificación de moléculas e procesos moleculares. É de especial interese a aplicación destas técnicas para o estudo de dinámicas de asociación e auto-ensamblaxe de procesos biomoleculares, xa que permiten estudos cuantitativos tanto termodinámicos como cinéticos. Por iso, esta tese foi realizada empregando diferentes técnicas espectroscópicas, dende as máis comúns (como a absorción UV-Vis ou a fluorescencia de estado estacionario) ata técnicas máis novidosas que permiten chegar á escala de moléculas individuais (como FCS ou microscopia de fluorescencia).

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

**Sección 5.1: Caracterización espectroscópica de G-cuadruplexes de ADN mitocondrial.**

Na primeira parte deste traballo, que corresponde co primeiro capítulo de resultados e discusións, levouse a cabo un estudo sobre a formación de cuadruplexes de guanina (G4) tanto en secuencias de ADN como de ARN e de híbridos de ADN:ARN. A autoensamblaxe de secuencias de ADN e ARN do noso organismo pode dar lugar á formación de estruturas coñecidas como G4. O papel destas estruturas aínda é obxecto de estudo, pero numerosos estudos defenden que está fortemente ligado á regulación da transcrición, tradución e expresión xenética. Isto implica que poden ser pezas chaves na aparición de diversas doenzas, entra as que tamén se atopa á EA. Nesta tese preséntanse os resultados obtidos empregando como técnica principal o dicroísmo circular (CD), cuxo uso está moi estendido no campo da bioquímica pola súa capacidade para discernir entre diferentes estruturas secundarias de biomoléculas. De forma complementaria, realizáronse estudos de fluorescencia e electroforese para a corroboración dos datos de CD. As técnicas de fluorescencia empregadas inclúen a emisión e a anisotropía de fluorescencia, ámbalas dúas tanto en estado estacionario como resoltas no tempo. O uso do CD permitiunos estudar cuantitativamente a dinámica da formación dos cuadruplexes. Deste xeito, elucidamos como a lonxitude da secuencia, a natureza e inclusive a presenza dunha guanina extra no comezo da secuencia podía influenciar a formación e a topoloxía adoptada pola molécula. Tamén estudamos o papel dos catións monovalentes ($Na^+$ e $K^+$) na estabilización dos G4 e obtivemos as constantes do proceso de formación para ambos casos.

Os resultados presentados demostran a formación de cuadruplexes nas secuencias de ADN estudadas, adoptando en todas elas unha topoloxía paralela. Tamén corroboran que unha maior presenza de guaninas resulta nunha formación maior de G4, e que a existencia dunha guanina inicial na secuencia xoga un papel importante nesta autoensamblaxe, influenciando significativamente o grao de formación. Os nosos estudos tamén evidenciaron o papel do sodio e do potasio no proceso de autoensamblaxe. Ambos catións provocan un cambio conformacional dunha topoloxía paralela a outra paralela altamente ordenada no G4 formado. Isto foi corroborado por emisión e anisotropía de fluorescencia, onde se aprecia que a unión da sonda á segunda conformación provoca unha restrición no movemento de rotación da molécula obrigándoa a adoptar unha estrutura cunha certa torsión, mentres que na primeira topoloxía adopta unha conformación plana. A comparación entre os dous tipos de catión demostrou que o potasio ten unha influencia maior que o sodio, fomentando en maior medida a topoloxía ríxida dos G4. Os estudos das secuencias de ARN e de híbridos de ADN:ARN tamén mostraron evidencias da formación de G4 principalmente paralelos.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Porén, no caso dos ARN observamos unha segunda estrutura nos resultados de CD, que podería corresponderse a topoloxías anti-paralelas ou híbridas.

**Sección 5.2: Concentración crítica de agregación e reversibilidade dos oligómeros de Aβ(1-40).**

Na segunda parte da tese, que abrangue as **seccións 5.2** a **5.4**, o traballo centrouse no estudo da agregación dos péptidos de amiloide. Os péptidos de amiloide están considerados como uns dos principais responsables da aparición da EA. A formación de agregados deste péptidos, en especial dos agregados temperás, foi suxerida como peza chave para o desenvolvemento da neurodexeneración. Existen principalmente dous tipos de amiloides no noso organismo, o Aβ40 e o Aβ42. Malia seren o Aβ40 o amiloide máis abundante no organismo, é o Aβ42 o máis estudado pola súa gran tendencia á agregación. Nesta tese, o estudo dos agregados temperás do Aβ40 mediante a técnica de FCS permitiunos caracterizar o proceso de agregación, obtendo tanto as concentración críticas de agregación como a distribución de tamaños dos agregados formados a diferentes concentracións de amiloide. Para isto, as curvas de correlación de FCS foron analizadas para conseguir os tempos de difusión do amiloide en mostras con diferentes concentracións, podendo diferenciar os monómeros do resto de agregados. A partir dos tempos e da contribución de cada un á curva de correlación, aplicando un complexo análise de datos, conseguimos o grao de agregación e o tamaño medio dos agregados para cada mostra. Tamén se investigou a reversibilidade e a estabilidade temporal do proceso de agregación. Ao longo do capítulo fíxose unha comparativa entre o mecanismo de agregación obtido para o Aβ40 e o mecanismo xa publicado anteriormente polo grupo para o Aβ42.

As investigacións realizadas evidenciaron as diferencias existentes entre a agregación dos amiloides Aβ40 e Aβ42. En primeiro lugar, os nosos resultados mostraron un proceso de dúas etapas na agregación do Aβ40 en oposición á unha única etapa para o Aβ42, obtendo dúas concentracións críticas de agregación (cac). No Aβ40, a primeira fase dá lugar a agregados de tamaño e forma similares aos observados anteriormente para o Aβ42, mais isto ten lugar a concentracións case 10 veces maiores para o Aβ40. Tras isto, os agregados poden seguir medrando ao aumentar a concentración, de forma que, despois de que a segunda fase teña rematado, os tamaños dos agregados son significativamente maiores. Os estudos da reversibilidade demostraron que os agregados formados poden ser desagregados tras a dilución da mostra, mais esta desagregación non é completa para aquelas mostras nas que os agregados formados son os máis pequenos, que precisan de maiores tempos dende a dilución para monomerizarse. Por outra banda, os estudos da estabilidade temporal, nos que se estudaron as mesmas mostras a diferentes tempos dende a súa preparación, demostraron que os agregados formados non sofren variacións importantes no seu

tamaño nin na súa composición aínda despois dunha semana. Esta estabilidade tamén é diferente que para o caso do Aβ42.

**Sección 5.3: Autofluorescencia de amiloides como metodoloxía para a detección de agregados temperás.**

Ao terse relacionado os agregados temperás dos amiloides coa EA, é necesario buscar ferramentas que axuden á detección de amiloides de forma selectiva e que permitan a diferenciación entre amiloides monoméricos e agregados. Ata a data, a sonda fluorescente máis estendida para realizar estes estudos é a Tioflavina T, xa que presenta cambios importantes na súa emisión cando se une aos péptidos, especialmente ás fibras de amiloide. Porén, esta sonda presenta un mecanismo complexo que dificulta o tratamento dos datos. Ademais, tamén é capaz de unirse a outras proteínas con constantes de asociación relativamente baixas e moi semellantes. Polo tanto, a busca de sondas con mellores características é aínda de gran interese. Os amiloides contan con un residuo de tirosina na súa secuencia peptídica, o que lles confire certa autofluorescencia. A fluorescencia intrínseca dos amiloides podería ser unha vía para a detección destes, sendo ademais un método máis sinxelo que non necesitaría de moléculas adicionais para o estudo da emisión. É por iso que realizamos un estudo da autofluorescencia dos amiloides como posible ferramenta para a detección destes péptidos e a distinción entre péptidos agregados e monoméricos. Este estudo foi levado a cabo empregando a emisión de fluorescencia como técnica principal, e a espectroscopia FCS como técnica complementaria.

O estudo realizado sobre a autofluorescencia dos amiloides evidenciou o potencial uso da emisión intrínseca da tirosina do Aβ42 como método de detección. Os espectros de emisión e de excitación do amiloide de mostras principalmente monómericas e mostras agregadas mostraron variacións en intensidade e na posición dos máximos de emisión suficientes para seren usadas para discernir se o proceso de agregación do péptido tivo lugar ou non. Cando o amiloide se atopa en forma de monómero, o espectro de emisión ten unha única banda a lonxitudes de onda arredor dos 350 nm. Esta banda desprázase a menores lonxitudes de onda (303 nm) a medida que a agregación vai avanzando. Ademais, tamén se observou unha banda a maiores lonxitudes de onda (500 nm) que atribuímos á unha emisión inducida pola agregación, posto que a súa contribución é maior en mostras de maior grao de agregación. A agregación das mostras foi corroborada con medidas de FCS. En vista dos resultados, propuxemos un modelo para o mecanismo fotofísico da tirosina, no cal esta sofre unha desprotonación no estado excitado cando o péptido non está agregado, que se ve impedida cando o péptido se agrega. Isto pode ser debido ao cambio no entorno próximo da tirosina, que tras a agregación pode estar menos exposta á disolución acuosa impedindo a perda do protón. Malia todo, os datos obtidos non teñen a calidade suficiente para facer un estudo cuantitativo do proceso,

polo que o estudo realizado é tan só semi-cuantitativo, e precisa de máis experimentos para aclarar de forma máis precisa a natureza e comportamento da banda a maiores lonxitudes de onda. Por iso, o modelo proposto debe ser considerado unha hipótese que explica debidamente os nosos datos, pero debe ser corroborada con máis experimentos.

**Sección 5.4: Influenza das interaccións electrostáticas na agregación de amiloides.**

Os amiloides, cando se atopan en condicións fisiolóxicas, contan cunha carga neta negativa. No noso organismo existen variadas biomoléculas pequenas que presentan cargas netas positivas cando se atopan en pH fisiolóxicos, como é o caso das poliaminas. As interaccións non covalentes son de gran importancia en numerosos procesos de asociacións intermoleculares. Por iso, a presencia no medio de moléculas como as poliaminas podería influír no proceso de agregación dos amiloides mediante forzas electrostáticas debidas á diferencia de carga entre ditas moléculas. Estas interaccións tamén son obxecto de estudo como un dos motivos que provocan a separación de fases líquido-líquido (LLPS das súas siglas en inglés). Existen diversos estudos nos que se evidencia que nas proteínas que sofren procesos de agregación como os dos amiloides, este proceso ocorre mediante unha LLPS. Desta maneira, formaríanse uns condensados onde se crearía un medio local cunha concentración de proteínas elevada que provocaría a agregación. Tendo en conta todo isto, nesta tese estudouse a agregación dos amiloides tamén dende o punto de vista da LLPS. Neste traballo estudamos, empregando a microscopia de fluorescencia, en primeiro lugar se a agregación do amiloide se producía por LLPS. Tras isto, estudamos tamén a influenza que as interaccións electrostáticas entre amiloides e outras biomoléculas cargadas (espermina, espermidina e sal de magnesio) podería ter no proceso de agregación. Este estudo foi realizado en diferentes etapas da agregación: antes e despois da formación de fibras. Ás imaxes de microscopia aplicóuselle un tratamento coñecido como SRRF, que permite a obtención de imaxes de alta resolución sen a necesidade empregar microscopias de súper-resolución.

Mediante a técnica descrita, estudáronse mostras de amiloide a diferentes tempos de incubación, que permitiron evidenciar como o Aβ42 se agrega ata formar fibras mediante a separación de fases, é dicir, formándose condensados circulares nos que se formaban os agregados e que desaparecían a medida que avanzaba a formación de fibras. Por outra banda, os nosos resultados tamén demostraron a influencia da espemina, que conta cunha carga neta de +4, na formación de fibras. A adición desta poliamina a mostras con fibras xa formadas resultou na aparición de condensados e desaparición progresiva de fibras, tanto dependendo do tempo de incubación como da concentración de poliamina engadida. Cando as moléculas engadidas eran a espermidina (con carga +3) ou o cloruro de magnesio (con carga +2), este proceso non foi observado tan claramente. No caso da espermidina non se observaron variacións en ningunha das condicións observadas, nin de concentración nin de tempo, o que indicaría a necesidade de cargas maiores para que

USC UNIVERSIDADE

as interaccións carga-carga teñan un impacto. Porén, os estudos feitos con magnesio si mostraron formación de condensados nas concentracións máis altas e nos tempos de incubación máis elevados, o que indica que para cargas menores o proceso non se ve tan favorecido. Estes estudos mostraron tamén que a carga non é o único factor que inflúe na LLPS. A investigación do tamaño dos agregados revelou que non presentaban variación en tamaño nin forma, nin coa concentración nin co tempo de incubación. Isto suxire que a reversibilidade da formación de fibras non é completa. Por último, tamén se fixeron estudos antes da formación de fibras, que revelaron pequenas variacións en función da concentración e do tempo de incubación. Estas variacións eran tanto da presencia de fibras e condensados como do tamaño destes. Malia todo, as tendencias non foron o suficientemente claras polo que se precisarían estudos a diferentes tempos de incubación e a diferentes concentracións para poder establecer conclusións máis sólidas.

**Sección 5.5: 2-ciano-6-hidroxibenzotialzol como sonda fluorescente para a detección de seroalbumina humana.**

Na última parte desta tese estudáronse as proteínas transportadoras do plasma sanguíneo, posto que estas xogan un papel importante no traslado dos amiloides dende o torrente sanguíneo e o líquido cefalorraquídeo. Este papel que presenta a seroalbumina humana é a base do tratamento da enfermidade de Alzhéimer por intercambio de plasma con reposición de seroalbúmina humana (coñecido como AMBAR), que se atopa en ensaios clínicos. Ademais, algúns estudos recentes suxeriron que a unión da HSA cos amiloides favorece a inhibición da agregación destes, evidenciando aínda máis o seu papel en posibles metodoloxías terapéuticas para o tratamento da EA. Debido á importancia da HSA nesta enfermidade, faise necesaria a busca de sondas fluorescentes que permitan a detección selectiva da proteína e dos amiloides, ou da unión entre ambas moléculas. Como xa se mencionou anteriormente, o uso da Tioflavina T, aínda que moi estendido, conta con desvantaxes importantes. A molécula 2-ciano-6-hidroxibenzotiazol (CBTOH) é unha especie que xa foi caracterizada polo noso grupo en disolucións acuosas e outros disolventes. Ese estudo revelou o carácter fotoácido desta sonda, que sofre unha transferencia protónica no estado excitado. Ademais, tamén revelou as variacións espectrais que sofre en función da polaridade do medio no que se atopa, o cal a fai unha candidata para a detección de biomoléculas. Por outra banda, a seroalbúmina bovina (BSA) adoita ser empregada en estudos biolóxicos como modelo da HSA. Isto é debido á dificultade que supón a obtención e purificación da HSA comparada coa BSA, o que tamén fai que esta última teña un custo económico menor. Deste xeito, estudamos o CBTOH como posible sonda fluorescente para a detección de seroalbuminas con técnicas espectroscópicas tanto de absorción UV-vis como de emisión e anisotropía de fluorescencia. A emisión de fluorescencia estudouse en estado estacionario e resolta no

USC UNIVERSIDADE DE COMPOSTELA

tempo. O estudo da unión do colorante tanto coa BSA como coa HSA permitiunos propoñer o mecanismo de asociación e obter as constantes de unión da sonda con ambas proteínas.

Como primeira etapa do estudo da asociación do CBTOH ás albuminas, investigouse o equilibrio ácido-base en condicións fisiolóxicas de forza iónica . Isto revelou unha pequena variación do $pK_a$ cara a valores lixeiramente inferiores respecto aos descritos anteriormente en disolución acuosa sen sales engadidas. Despois, fíxose o estudo da asociación e os resultados obtidos demostraron unha unión efectiva entre a sonda e ámbalas dúas proteínas. A asociación pode observarse posto que os espectros, tanto de absorción como de emisión, presentan cambios de intensidade e desprazamentos espectrais a medida que se aumenta a concentración de proteína, podendo ser monitorizado o proceso de asociación por absorción e emisión. En ambos casos, a asociación pode explicarse cun modelo de equilibrio de asociación hóspede-anfitrión (*host-guest*) 1:1, que permitiu reproducir adecuadamente os datos. A este respecto, observáronse diferencias entre as proteínas. A pH fisiolóxico, o CBTOH pode existir en forma aniónica e neutra, e ambas formas poden unirse á proteína. Non obstante, a fluorescencia da forma neutra non é observable xa que sofre unha transferencia protónica moi rápida no estado excitado, observándose soamente a emisión da forma aniónica. Os estudos revelaron que esta transferencia protónica non ocorre, ou ocorre en moi pouca medida, cando a sonda está unida á BSA, pero si ten lugar cando se atopa unida á HSA. Isto evidencia que, malia ser proteínas usadas como análogas, as diferencias na secuencia confírenlle propiedades diferentes, tendo distintas afinidades de asociación e sitios de unión lixeiramente diferentes. As diferenzas tamén se presentaron nos valores das constantes de asociación obtidos, que non obstante son da mesma orde de magnitude. De forma xeral, os resultados demostran que o CBTOH conta coas características para poder ser empregado como sonda para a detección das proteínas do plasma sanguíneo. Ademais, como existen diferenzas espectrais entre os complexos formados por cada proteína, tamén permite a identificación da proteína á que esta unido. Isto tamén suxire que a sonda podería mostrar diferenzas espectrais de unirse a outras proteínas, como os amiloides. Ademais, ao aparecer as diferenzas tanto en absorción como en emisión, unha combinación de técnicas cun análise ratiométrico dos espectros podería ser de utilidade na cuantificación da detección.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# 1. Introduction

Living organisms are remarkably complex and rely on a meticulous network of molecular interactions to function. Numerous interconnected processes are taking place simultaneously to sustain life. Cellular activity, from energy production to gene regulation, depends on conserving the balance of all biochemical interactions. Within all involved factors are biomolecules. These molecules, that include lipids, proteins, nucleic acids, and carbohydrates, are fundamental to the structure and function of cells. They play crucial roles in the biochemical processes essential for life.[1]

Proteins serve as the primary drivers of the cell, regulating cellular processes, catalysing biochemical reactions, and providing structural support. Nucleic acids, such as DNA and RNA, store and transmit genetic information, ensuring that each cell functions according to its genetic blueprint. Lipids, which include fats and oils, form the basis of cellular membranes, creating barriers that protect the cell and enable compartmentalization of different cellular functions. Carbohydrates, on the other hand, provide energy and serve as signalling molecules, facilitating communication between cells.[1]

One of the reasons why biomolecules are fundamental pieces of all biological processes is due to their ability to self-assembly into complex structures and participate in specific interactions. Self-assembly is involved in cellular organization, enabling the formation of structures such as membranes, ribosomes, and the cytoskeleton. Their specific binding interactions are key for the regulation of cellular activities. For example, enzymes selectively bind to substrates to catalyse reactions, while receptors on the cell surface bind to signalling molecules to trigger responses within the cell.

Self-assembly refers to the process by which molecules spontaneously organize into ordered structures without external guidance. This phenomenon is driven by a combination of non-covalent interactions such as hydrogen bonding, van der Waals forces, hydrophobic effects, and electrostatic interactions.[2,3] This is evidenced in numerous biological events such as the formation of cellular membranes, the folding of proteins into functional three-dimensional shapes, and the construction of multi-molecular complexes. On the other hand, binding interactions involve the specific and often reversible association between biomolecules, leading to the formation of complexes. These interactions are critical for numerous biological functions, including enzyme-substrate recognition, signal transduction, immune response, and gene

USC UNIVERSIDADE DE SANTIAGO

regulation. The specificity and affinity of these interactions are governed by the precise fit between the interacting molecules, often described by the "lock-and-key" or "induced fit" models.

While the self-assembly, host-guest, and other binding interactions of biomolecules are indispensable for the normal function of cells, aberrations in these processes can lead to a wide variety of pathological conditions. These issues may appear from the absence of these processes, their presence, or their failure to occur properly. For example, in neurodegenerative diseases like Parkinson's, Alzheimer's, or amyotrophic lateral sclerosis (ALS), the development of the disorder is thought to be caused by misfolding and aggregation of key biomolecules involved.[4–6] These self-assembly processes would disrupt the normal cell function leading to the pathogenesis of the aforementioned diseases. Another example would be cancer, where misregulated protein-protein interactions along with self-assembly of signalling complexes can promote uncontrolled cell proliferation and metastases.[7–9]

In contrast to the above, the binding processes can be exploited as well for therapeutic purposes. The design of small molecules that selectively bind to proteins inhibiting their aggregation or pathological protein-protein interactions show promising results as treatments for neurogenerative diseases. The use of other existing biomolecules is another alternative that is also being studied. For example, in the case of Alzheimer's disease (AD), there is a promising treatment using blood plasma exchange with clean human serum albumin (HSA) which is already on clinical trials.[10]

The duality of biomolecular self-assembly and binding between being essential for life and potential precursors of numerous diseases highlights the importance of this field of study. The interest in the aggregation and binding processes of biomolecules related to neurodegenerative diseases has been increasing over the past decades because of the need to understand the mechanism underlying these conditions.

Neurodegenerative diseases symbolise a growing challenge to public health across the globe as the number of patients developing these conditions is continuously increasing worldwide. These illnesses are highly linked to the ageing process. Hence, the global increase in life expectancy and the growing proportion of elderly individuals in the overall population are key factors contributing to their prevalence. With approximately 55 million people affected globally currently[11], these diseases which are detrimental both for the patients and their families, become also a risk to society. In 2019 the Alzheimer Europe organization published a report about the prevalence of AD within European countries.[12] The results showed that in 2018, AD's estimated prevalence rate among the European total population from ages 65 to 69 was about 1.3 % (1.1 % for men and 1.5 % for women, Figure 1.1), but it increased rapidly to a 12.1 % for population aging above 80. In Spain, the

USC DE SANTIAGO DE COMPOSTELA

number of individuals affected by AD in 2018 was over 800 000, and it is expected to increase by more than 100 % by 2050 (representing a 4 % prevalence of the total population).[12]

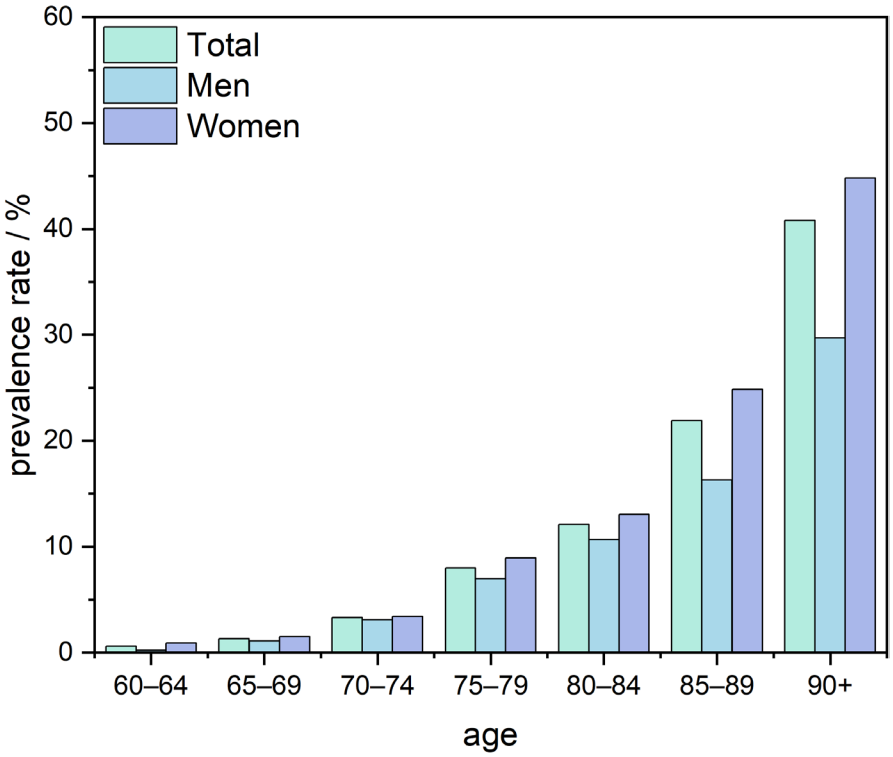



*Figure 1.1. Comparison of Alzheimer's prevalence rates in 2018 between Europe's total population, men population and women population within different age ranges. This figure was self-made using data from reference 12.*

To this day, the reason of AD development is still unclear. However, several hypotheses have been suggested where different biomolecules were pointed out to be involved in AD. The two main accepted molecular players are beta-amyloid peptides and tau proteins. Amyloid-beta peptides aggregation outside neurons along with the accumulation of misfolded tau proteins inside neurons are thought to be responsible of the neurodegeneration caused in AD.[13–18] Even though these two proteins have been widely studied, there are other biomolecules that have also been linked to the disease like apolipoprotein E, Guanine quadruplexes (G-quadruplexes or G4s) or even one of the blood plasma transport proteins HSA.[16,18–26]

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

## G-quadruplexes

G-quadruplexes are four-stranded non-canonical structures that DNA and RNA sequences can adopt. They appear in guanine-rich nucleic acids sequences and its presence can be predicted in sequences that follow a $G_{\geq 3}N_{1\text{-}7}G_{\geq 3}N_{1\text{-}7}G_{\geq 3}$, where N is one of the four possible nucleobases. G4s are formed by the union or two or more G-quartets, its primary unit. G-quartets consist of the assembly of four guanines that bind together by Hoogsteen hydrogen bonds forming a square planar structure (Figure 1.2). G4 stability depends not only on the π-π interactions between the G-quartets that conform it but also on stabilising cations that coordinate with the guanine's oxygen atoms. The most reported stability trends are $K^+ > Rb^+ > Na^+ > Li^+$, $Cs^+$ and $Sr^{2+} > Ba^{2+} > Ca^{2+} > Mg^{2+}$,[5-8] although they may slightly vary depending on the nucleic acid sequence. Moreover, despite the large number of cations analysed, the most relevant at physiological conditions, and thus the most studied, are $Na^+$ and $K^+$. The union between G-quartets can be inter or intramolecular, and depending on the combination of the strand direction the resulting G4 can adopt different topologies. These topologies are parallel (all strands have the same orientation), antiparallel (all strands have opposite directions) and hybrids (Figure 1.2).

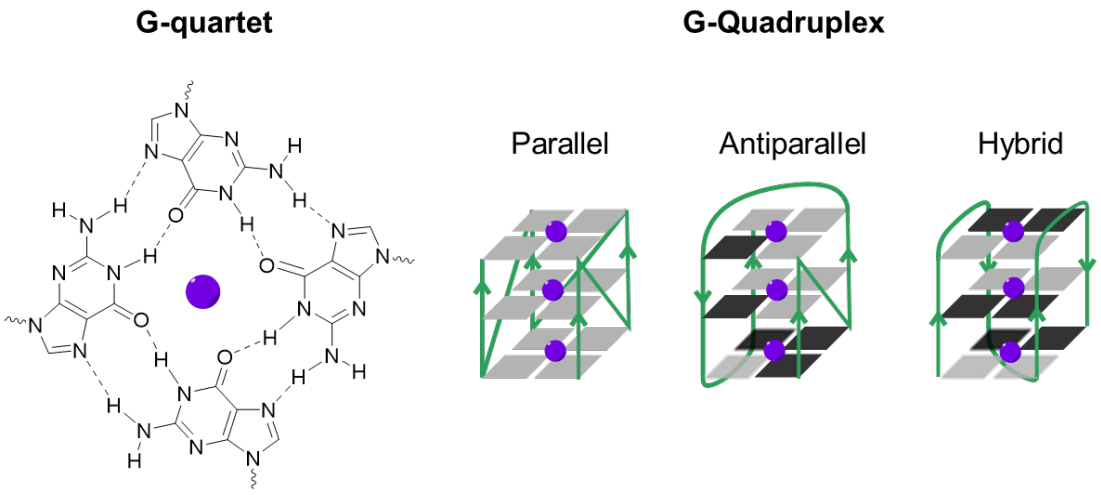



*Figure 1.2. Chemical structure of a G-quartet stabilized by a cation (purple circle in the middle), and schematic representation of the three possible topologies G4 can adopt: parallel, antiparallel and hybrid.*[27]

The first study reporting G4 was published in 1962[28] and since then numerous investigations have already described its existence. A recent bioinformatics analysis of the human genome estimated over 350,000 motifs with the ability to form a G4.[29] Although the formation of G4 structures within the genome has been controversial,[30] there is clear evidence of their formation in vivo, and it has been reported their presence at the telomeres[31,32], nuclei[33], nucleoli[34,35], cytoplasm[36–38] and mitochondria.[39,40] In fact, more recent computational and sequencing studies show an even greater presence of potential G4-forming sequences, identifying of up to 716,310 distinct G4 structures.[20,21]

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Interestingly, these assays detected a high density of G4s in functional regions such as 5' untranslated regions, splicing sites, telomeres and gene promoters.

The importance of guanine quadruplexes lies in their biological functions, as they are structures found in the DNA of many species.[41] In nuclear DNA, G-quadruplexes are associated with genome instability and the regulation of transcription, with more than 40 % of promoter regions in human DNA containing sequences capable of folding into these quadruplexes.[42–45] Additionally, they are involved in gene expression, including the regulation of messenger RNA (mRNA), the translation process into proteins, and polyadenylation.[39,45,46] G4s can also be found in telomeres (linear structures formed at the ends of chromosomes crucial for their stability), where they seem to be capable of inhibiting the enzyme telomerase, which is responsible for the elongation of telomeres.[39,43,45,46] Within the mitochondrial DNA, there is a guanine-rich conserved sequence block (CSB II) in which three different G4 can be formed (Figure 1.3): 1) an unimolecular DNA G4 formed at the non-template DNA strand, 2) an unimolecular RNA G4 formed at the transcript strand and 3) a DNA:RNA hybrid between both strands. It has been pointed out that these mitochondrial G4 could play crucial roles in replication, transcription, and translation.[39,43,45,47]

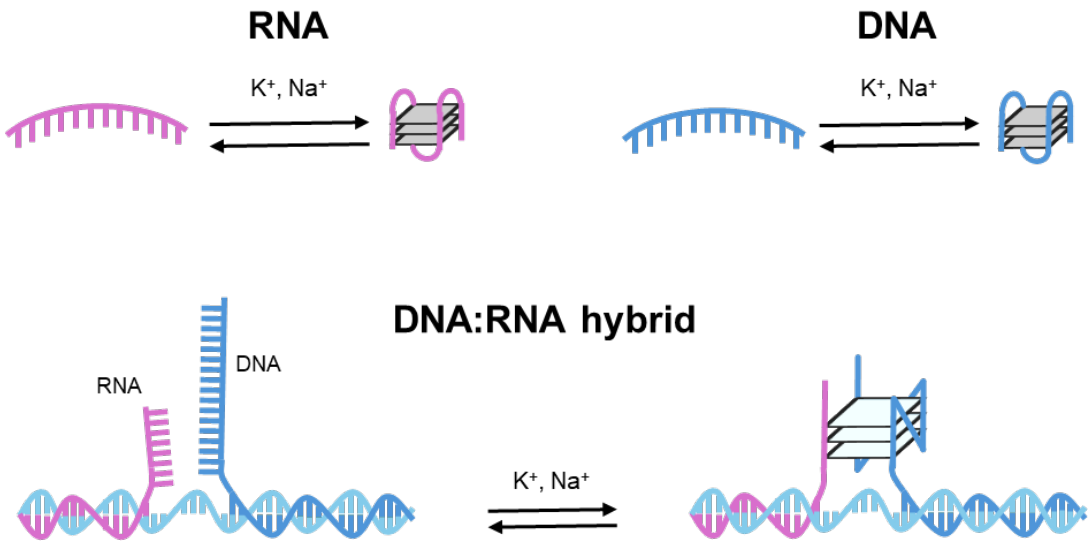



*Figure 1.3. Schematic representation of the three distinct G4s that can be formed within mitochondrial CSB II: unimolecular RNA G4 assembly adopted within the RNA transcript, unimolecular DNA G4 formed within the non-template strand and bimolecular DNA:RNA hybrid formed between the non-template DNA and the RNA transcript.*[27]

Because of their implication in translation, G-quadruplexes are also gaining attention for their potential roles in neurodegenerative diseases.[20,21,48–51] It has been reported that G4s regulate the expression of genes associated to the pathways of AD.[20,50] For example, the elimination of amyloid-beta (Aβ) plaques in AD depends on the enzyme ADAM10, which processes amyloid precursor protein (APP) to release a protective fragment called sAPP-α. An RNA G4 structure in ADAM10 mRNA inhibits its translation. When this G4 structure is mutated, ADAM10 translation increases, leading less Aβ production. A

methylquinolinium derivate chemical entity, known as compound 24, binds to the G4 structure, promoting ADAM10 translation and reducing Aβ production. These researches show that the G4 structure in ADAM10 mRNA plays a crucial role in regulating ADAM10 levels and suggests that targeting such RNA structures could help treat AD.[52,53]

## Amyloid-beta

As mentioned above, amyloid-beta is one of the different biomolecules involved in the pathogenesis in AD. Amyloid peptide derives from a protein known as the amyloid precursor protein (APP), which is an integral transmembrane protein expressed in many tissues, including the brain.[54] Through the action of specific enzymes known as secretases, APP is sequentially fragmented to produce various species, with Aβ being one of the most significant in the context of Alzheimer's disease. The fragmentation of APP can happen following two different pathways depending on the secretase involved (Figure 1.4). On one hand, α- and γ-secretase follow the non-amyloidogenic pathway, where their combined action processes APP releasing shorter amyloids known as P3 peptides. On the other hand, first β-secretase and γ-secretase follow the amyloidogenic processing of APP leading to the formation of Amyloid-β(1-40) (Aβ40) mainly and to lesser extent Amyloid-β(1-42) (Aβ42). The mechanisms of both pathways, even though similar, result in different species because of the regions where the cleavage takes places.[18,54,55]

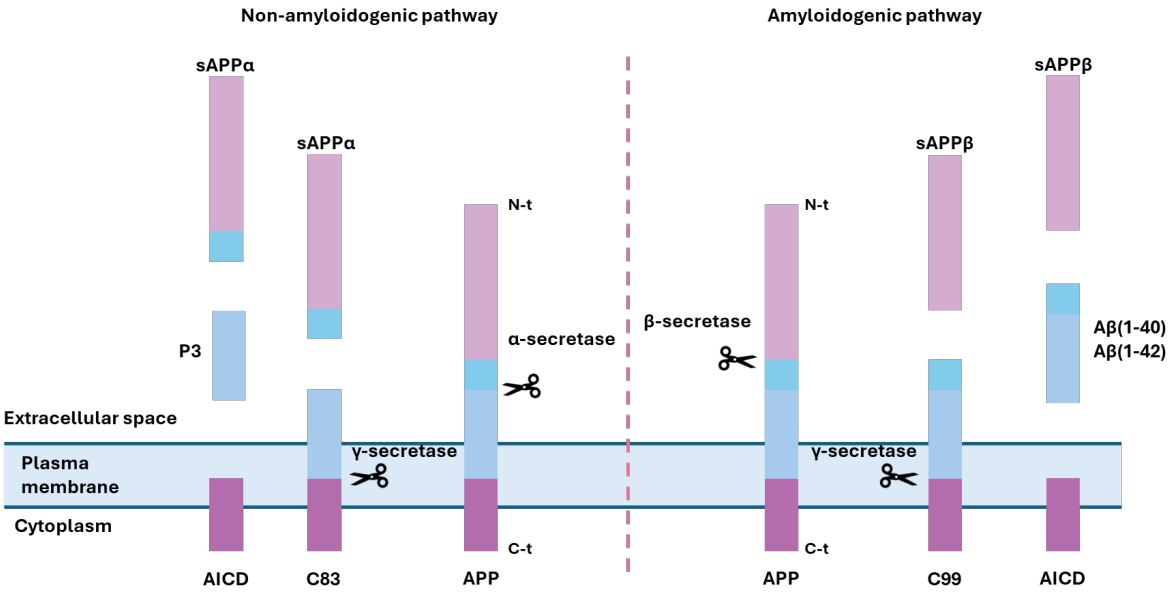



*Figure 1.4. APP cleavage pathways. This figure is self-made adapting figures from references 18 and 54.*

In the non-amyloidogenic pathway (Figure 1.4, left), first the α-secretase cuts the APP by the AB domain within Leu688 and Lys687 creating a soluble fragment known as sAPPα and a membrane-bound fragment called C38. C38 is then processed by γ-secretase resulting in two knew fragments: the P3 peptides and the intracellular APP domain (AICP).

P3 are shorter amyloid peptides that are not involved in AD. These peptides correspond to the Aβ(17-40) and Aβ(17-42) sequence. Contrary, in the amyloidogenic pathway (Figure 1.4, right), instead of the α-secretase, the β-secretase (also known as BACE1) is responsible of cutting the APP and this takes place outside the Aβ region, at the Aβ N-terminus around the Met671 and Asp672, generating the soluble fragment sAPPβ and the membrane-bound C99 fragment. Then, as in the previous pathway, γ-secretase cleaves the C99 fragment forming the AICD and the Aβ peptides.[18,54,55]

The Aβ peptides formed in the amyloidogenic pathway are, even though smaller than most proteins, a type of intrinsically disordered proteins (IDP). IDP are those proteins that lack a stable three-dimensional conformation and so can adopt different conformations. Another characteristic of IDP is their high tendency towards assemblage. During the aggregation process, diverse species like oligomers, protofibrils and fibrils are formed. In the brain of AD patients, accumulations and deposits of amyloid fibrils can be found, which is why they were first associated with neurodegeneration. However, numerous studies reported that the actual neurotoxic species are the early oligomers and protofibrils formed at the beginning of the aggregation process. Even though the exact mechanism by which oligomers promote neurotoxicity is still unclear, different ways in which they could affect neurotoxicity were reported: i) extracellular amyloid can interact with cellular membranes and cellular receptors affecting them in multiple ways, ii) those amyloids that were able to enter the cell can influence the system causing cellular damage and leading to apoptosis; iii) both extracellular and intracellular amyloids present cell-to-cell transmission.[56]

As mentioned above, different forms of Aβ have been identified presenting different lengths and conformations. Among the various amyloid peptides, Aβ40 and Aβ42 are recognized as primary contributors to Alzheimer's. Despite differing by only two amino acids at their C-termini (Figure 1.5), these peptides exhibit significantly distinct properties. Aβ40, the most prevalent form (comprising 80-90 % of amyloid peptides),[57] is more soluble and less prone to aggregation compared to Aβ42. Due to its higher propensity for aggregation, Aβ42 has been identified as a key player in AD and has been the focus of extensive research. However, studies indicate that the aggregation mechanisms may vary depending on peptide length and that there are interactions between different amyloid-beta peptides.[58–67]

For Aβ42, our group has found in previous studies that it only aggregates when the concentration of amyloid monomers exceeds a given concentration, known as critical aggregation concentration (*cac*).[68] This behaviour is typical of spontaneous aggregation process and indicates that below the *cac* only monomers would be present in solution, whereas after exciding this concentration aggregation would take place steeply. This spontaneous, cooperative process in Aβ42 yields a *cac* of 91 ± 14 nmol/L and stable

USC UNIVERSIDADE

micelle-like oligomers. The aggregates size (≈50 monomers, $R_h$ ≈ 7.3-10.7 nm) and elongated shape are independent of incubation time and peptide concentration. These results presented by our group for the Aβ42 oligomers are in very good agreement with the reported cylindrical micellar assemblies of Aβ40.[69]

On the other hand, the literature on Aβ40 aggregation presents notable discrepancies, as some authors suggest that Aβ40 aggregation leads directly to fibrils without forming oligomers nor protofibrils, while others report results that hint at oligomer formation.[62–66,69–74] Among those who reported oligomers formation, inconsistencies in the reported sizes of these oligomers can be found as well. [62,72,74,75] Because of this, and given that Aβ40 is the dominant species in the human body, understanding its aggregation mechanisms is crucial for comprehending the overall process of amyloid aggregation in Alzheimer's disease.

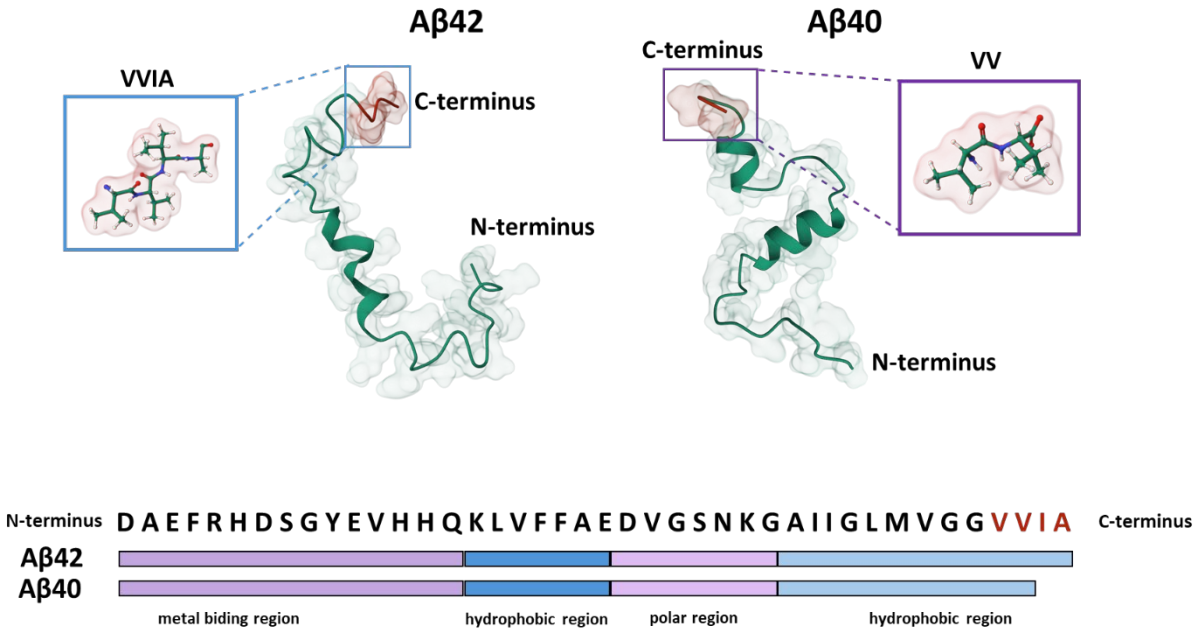



*Figure 1.5. Amyloid-beta 1-42 and 1-40 3D structures (upper panel) and their respective amino acid sequences (lower panel).*

Amyloid aggregation mechanisms have been widely studied and different mechanisms have been proposed. Traditionally, most mechanisms suggested a multi-step aggregation with different pathways, but all followed a transformation from monomers to fibrils through the formation of different species. The most accepted mechanism hints towards a two-step aggregation with a first nucleation phase taking place followed by an elongation phase (Figure 1.6). The nucleation phase would lead to the formation of the first oligomers which would act as the nucleus for the elongation phase. These oligomers would consist only of the accumulation and aggregation of monomers producing a more spherical conformation. Once oligomers are formed in solution, they growth into larger

USC DE SANTIAGO DE COMPOSTELA

aggregates and their elongation into fibrils would take place. Some authors suggest an additional step, a secondary nucleation, that would occur due to already formed fibrils interacting with monomers in solution. These interactions could act as catalysers of the aggregation process. It has also been reported that even though fibrils are quite stable structures they could potentially break into small oligomers. The overall kinetic profile of the aggregation process from this mechanism would correspond to a sigmoidal curve, were the nucleation would be the slow phase and the decisive step in the aggregation rate. If already formed oligomers are added to the solution they would act as seeds for nucleation and speed up the aggregation process.[56,76]

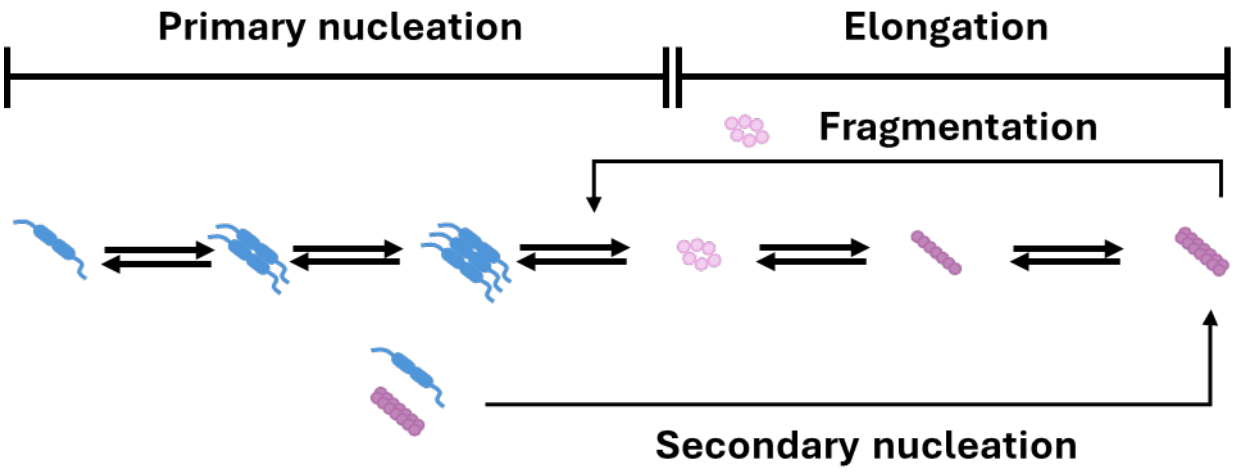



*Figure 1.6. Schematic representation of the amyloid aggregation process through a nucleation-elongation step process, according to the literature. This figure is self-made adapting figures from references 56 and 76.*

Recently, multiple studies suggested that intrinsically disordered proteins like α-synuclein and amyloid, which tend to form fibrils, may undergo a liquid-liquid phase separation (LLPS).[77–81] LLPS refers to the process by which certain proteins separate from the bulk solution to form distinct, dynamic, liquid-like condensates usually referred to as droplets. These droplets act as microenvironments with significantly higher local concentrations of proteins, in which the aggregation would take place. The local increase on protein concentration promotes aggregation and, therefore, fibril formation (Figure 1.7).

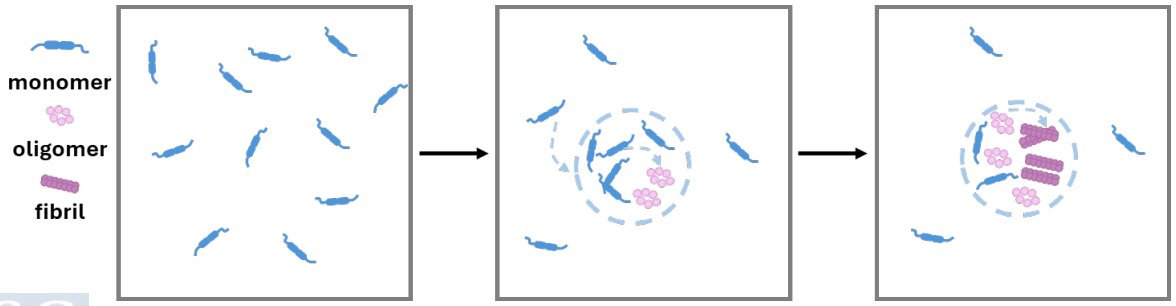



*Figure 1.7. Schematic representation of the liquid-liquid phase separation process.*

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Because the formation of droplets creates an environment in which the critical concentration exists for amyloid aggregation to start, fibril growth is easily achieved. Thus, LLPS may be a key step in amyloid pathological aggregation as condensates locally accelerate the kinetics of the aggregation process. Moreover, the environment that droplets provide to amyloids may influence its biochemical properties, potentially changing their tendency to aggregate. Still, the exact reason why this LLPS occurs during protein aggregation is still unknown, but weak, multivalent interactions, including hydrophobic interactions, hydrogen bonding, and particularly charge-charge interactions among the protein molecules have been proposed as key for induction of LLPS.[80,82,83] Charge-charge interactions could be crucial as proteins are highly charged molecules and its charge distribution could need stabilisation by electrostatic interactions with other charged biomolecules inherent to our organism like polyamines or salt cations.[32,34,3]

## Human serum albumin

Human serum albumin (HSA) is a protein that presents three homologous domains which at the same time are split into two subdomains each. It is the most abundant protein in blood plasma, and it is synthesised in the liver composed of 585 amino acids.[84] HSA plays a crucial role in maintaining oncotic pressure and serving as transport for both exogenous and endogenous substances. As it functions as a transport protein, HSA presents great ability to bind to a wide variety of substances like metals, metabolites and numerous proteins and peptides, such as Aβ.[24,84,85] Interestingly, more than 90 % of all amyloid in human blood exists bound to HSA, only a small proportion of it being free.[86] This high binding ability HSA has with Aβ has been reported in the literature highlighting the protein's potential implications in AD.[86–88]

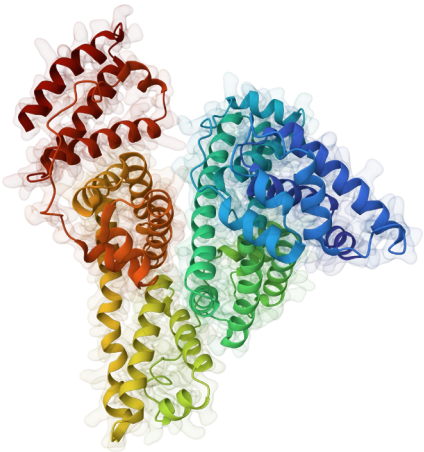


*Figure 1.8. Schematic representation of the liquid-liquid phase separation process.*

As mentioned above, numerous potential treatments for AD are still on ongoing research. Considering the implications that HSA has with Aβ, a plasma exchange (PE) treatment has been proposed by the Spanish pharmaceutical company Grifols S.A.[10,24] The PE with human serum albumin reposition (known as AMBAR) has already shown promising results and is currently in clinical trials (2b/3 phase).[24,89] AMBAR takes advantage of the binding ability HSA has and replaces the already existing bound HSA in blood plasma with clean HAS that would potentially bind with Aβ. This way, the new serum albumin could bind to both Aβ in the blood plasma and the brain. There is a dynamic equilibrium of Aβ between blood plasma and the cerebrospinal fluid, which would help HSA trap amyloids leading them to exit into the blood stream and avoid accumulation and aggregation of Aβ in the brain (Figure 1.9).[24,90]

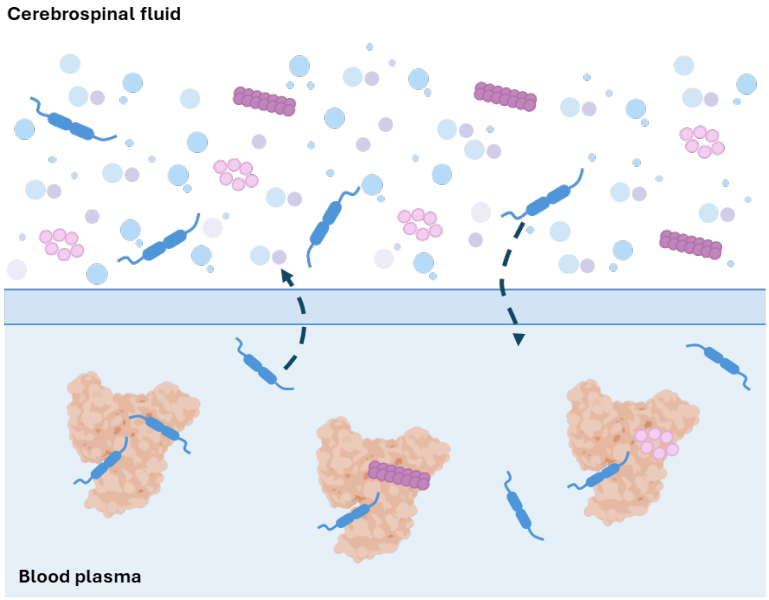



*Figure 1.9. Schematic representation of the Aβ exchange at the cerebrospinal fluid and blood plasma barrier and the binding of the amyloids to human serum albumin.*

Additionally, several authors have reported that the association between HSA and Aβ can also inhibit amyloid aggregation.[91,87,88] The binding process is thought to suppress Aβ self-assembly by interfering with the aggregation mechanism. This could be a result of HSA binding preferably to amyloid oligomers rather than monomers, which would restrain the growth toward bigger aggregates. However, the presence of fatty acids seems to have an impact on the inhibition as it was reported by Guo and Zhou.[92]

Even though human serum albumin therapeutical relevance has been made apparent, the exact mechanisms by which these processes take place are still unclear. For this reason, further studies are needed.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

## Fluorescence sensors

Fluorescence-based techniques are indispensable tools widely used in various fields, including chemistry, biological science, and materials science. These techniques usually present high sensitivity, specificity, and versatility, and some are non-destructive and non-invasive. These characteristics, intrinsic to fluorescence emission, allow for studies of diverse systems with real-time monitoring. However, for samples to be able to be studied by fluorescence emission, the presence of an emitting molecule is needed. For this reason, fluorescence techniques rely on the use of fluorescent sensors.

A fluorescence sensor can be defined as an emissive molecule that presents a high fluorescence quantum yield. They usually are sensitive to changes in their environment and respond to them by varying their emission intensity or the wavelength at which their emission takes place. Their spectral variations allow investigators to obtain valuable information about their molecular structure and interactions with other molecules. Molecular interactions, whether they are intra- or intermolecular, provoke changes in the molecular environment. These changes can be presented as polarity, pH, or viscosity variations. As a result, a change in the photophysical properties of the sensor is triggered. If applied to biological systems, they provide a way of studying association dynamics, self-assembly and folding processes, as well as cell localization.

There are two groups into which fluorescence sensors can be categorized: intrinsic and extrinsic. Intrinsic fluorescence sensors are those who rely on the natural fluorescence properties of the sample of study, without the need for any additional molecule. Some biomolecules like aromatic amino acids, flavins or chlorophyll present emission which makes them intrinsic fluorescence sensors. On the other hand, extrinsic fluorescence sensors are those emissive molecules that are added to the sample of study to modify its fluorescence properties.

Within this thesis, both intrinsic and extrinsic fluorescence sensors were used. As intrinsic probes, tyrosine present in amyloid peptides was studied, whereas rhodamine 123, Thioflavin T, HiLyte Fluor 488 and 2-cyano-6-hydroxibenzothiazole were used as extrinsic sensors for several applications. We present these probes below.

### Fluorescent amino acids

Many biomolecules (nucleic acids, lipids, carbohydrates) do not show fluorescence, but some proteins present great fluorescence properties. The reason why these proteins are somewhat fluorescent is because the aromatic amino acids that may be present within their sequences. The fluorescent amino acids are tryptophan (Trp or W), phenylalanine (Phe or F), and tyrosine (Tyr or Y). However, proteins which present these amino acids are

USC UNIVERSIDADE

not as common as one might expect. All three amino acids present aromatic groups and residues that confer each of them different fluorescence profiles Figure 1.10). However, fluorescent proteins can have more than one aromatic amino acid in their sequences, usually making identification of their intrinsic fluorescence harder. Additionally, proteins are supramolecular systems with complex environments in which numerous interactions may take place. As other fluorescence sensor, the amino acids fluorescence is also sensitive to their surroundings. Thus, the emission spectra of the aromatic amino acids can vary depending on the protein microenvironment.

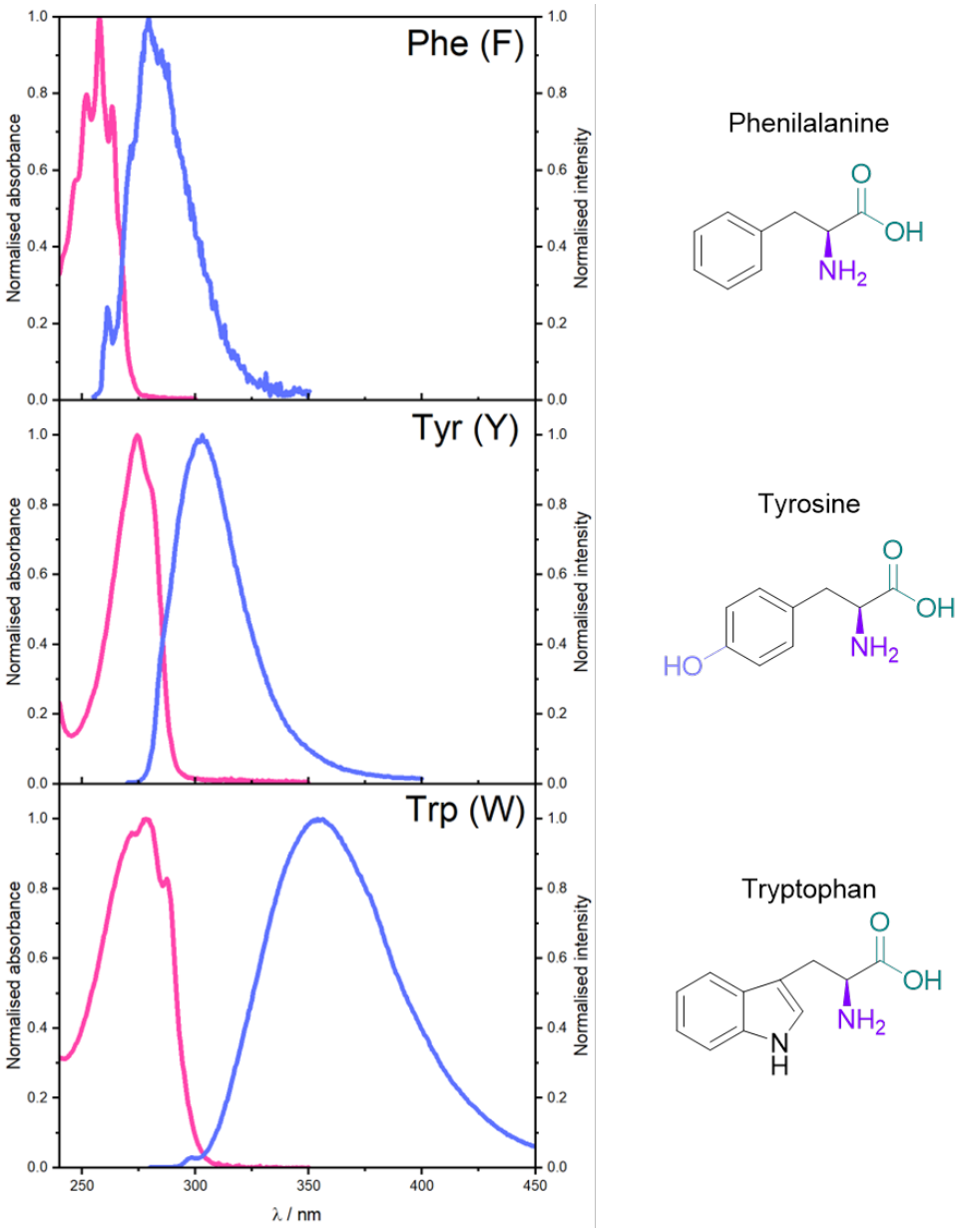



*Figure 1.10. Absorption (pink) and emission (blue) spectra of phenylalanine, tyrosine and tryptophane at pH 7 aqueous solution, and their respective molecular structures. Spectra were obtained from PhotochemCad repository.*[93]

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Tryptophane is the fluorescent amino acid that presents the highest molar absorption coefficient (Table 1.1) and highest fluorescence quantum yield, which make it the more emissive amino acid among the three. In addition, it is also the most sensible to changes in its surroundings as its 350 nm emission band is highly dependent on the polarity of the media. Its fluorescence emission can present quenching from other amino acids residues' side chains. When the three amino acids are present in the same protein, tryptophane benefits for the Förster resonance energy transfer (FRET) that takes places between them. Usually, phenylalanine acts as donor for tyrosine and tyrosine, at the same time, acts as donor for tryptophane. This way, tryptophane is the main contributor to emission from proteins. These properties are the reason why tryptophane is the most appealing intrinsic fluorescence sensor among the three aromatic amino acids. However, tyrosine can also showcase great properties as a fluorescence sensor.

*Table 1.1. Spectroscopic properties of the three aromatic amino acids.*[93,94]

| | Tryptophane | Tyrosine | Phenylalanine |
|---|---|---|---|
| **Absorption maximum / nm** | 280 | 275 | 257 |
| **Molar absorption coefficient / $M^{-1}$ $cm^{-1}$** | 5579 | 1405 | 195 |
| **Emission maximum / nm** | 348 | 303 | 282 |
| **Fluorescence quantum yield** | 0.2 | 0.14 | 0.03 |
| **Lifetime / ns** | 0.4-2.8 | ≈3 | ≈2 |

Even though tyrosine is commonly thought to present a simple photophysical behaviour, it can also present complex spectral variations. As shown in Figure 1.10 and summarised in Table 1.1, tyrosine presents an absorption maxima at 275 nm and an emission maximum around 303 nm. Contrary to tryptophane, tyrosine fluorescence is not sensitive to the polarity of the media. However, in the excited state it can suffer a deprotonation of the aromatic hydroxyl group forming a new species known as tyrosinate. This deprotonation appears as a result of ionisation which lowers the hydroxyl $pK_a$ from 10 on the ground state to 4 on the firs-excited singlet state. The formation of tyrosinate also leads to a significant spectral shift, changing the emission maximum from 303 nm to 350 nm. It has been reported that tyrosinate formation can happen even at neutral pH and that it may be promoted in solvents where proton acceptors are present. In the case of proteins, tyrosine excited-state ionisation also depends on the degree of tyrosine exposition to the aqueous phase. If the adopted secondary structure of the protein inhibits tyrosine of being in an aqueous environment, then the ionisation would be as well hindered.[93]

Three phenylalanine residues and one tyrosine residue are present in amyloid-beta 40 and 42, whose sequences are shown in Figure 1.5. This means that amyloid peptides should present fluorescence emission that could be used for their detection. Amyloid peptides detection, especially in the early stages of the aggregation, is key in the

prevention of AD. This early detection must be achievable with non-expensive and non-invasive methods that could be applied to a large segment of the population. Given that amyloid peptides have their own intrinsic fluorescence sensors, taking advantage of this could help to develop a method that would allow not only amyloid detection, but also the characterisation of the species formed in the studied sample. Thus, this would be a potential non-invasive and cheap tool for amyloid detection, as no additional fluorophore would be needed, and conventional, more broadly available fluorescence techniques could be used.

**Rhodamine 123**

Rhodamine is a family of xanthene's derivate fluorescence sensors. Its core structure is represented in black in Figure 1.11. There exist numerous variations of rhodamine depending on the functionalisation added to the core structure. The added functional groups confer the molecules different fluorescence properties. Rhodamine sensors are widely use in biochemistry and chemistry fields in techniques like flow cytometry, fluorescence microscopy, fluorescence correlation spectroscopy and conventional fluorescence. One of the reasons why their use is common is because they allow for cellular localisation and colocalization.[95] Among the most used rhodamines in biochemistry is rhodamine 123 (Figure 1.11), which presents its emission maxima at 525 nm.

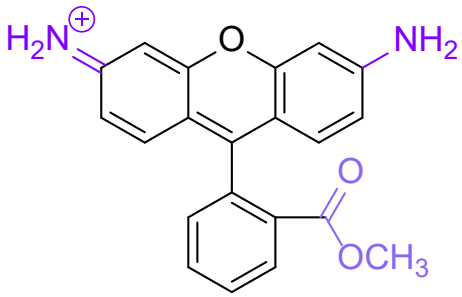



*Figure 1.11. Rhodamine 123 molecular structure. Rhodamine's family core structure is represented in black while the functional groups that characterise R123 are represented in different shades of purple.*

Rhodamine 123 (R123, Figure 1.11) is a rhodamine functionalised with two amino groups and one methoxycarbonyl. As mentioned above, R123 presents an emission maximum around 525 nm in aqueous solution, which shifts depending on the solvent. In the case of ethanol, the maximum presents a spectral shift of 5 nm with respect to water, having its maximum at 530 nm. Even though it presents spectral variations depending on the media, its properties do not vary with pH except for extreme conditions. Some intensity decrease can be seen at very acidic and basic pH (below 1 and above 12). The key point of R123 is

its high fluorescence quantum yield, of around 0.8 in water,[95–97] Its photophysical characteristics make R123 a reference fluorescence sensor in several fields. For example, in biological fields is commonly used for mitochondrial labelling as it can accumulate selectively inside them.[98,99] It can also be used as an anticancer agent as it presents photosensitiser properties.[100] Its emission properties also make R123 one of the reference dyes used in some microscopy techniques. In the case of fluorescence correlation spectroscopy (FCS), it works as the reference dye for laser alignment and sample volume calculation (see **Experimental methods** for further information).

In the context of this thesis, rhodamine 123 will be used exclusively as the reference dye for all fluorescence correlation spectroscopy studies. For these studies, we derived the diffusion coefficient of the dye to be $D_{R123}$ = (4.7 ± 0.3) × $10^{10}$ $m^2$ $s^{-1}$ at 25ºC based on the reported values in the literature.[101–103]

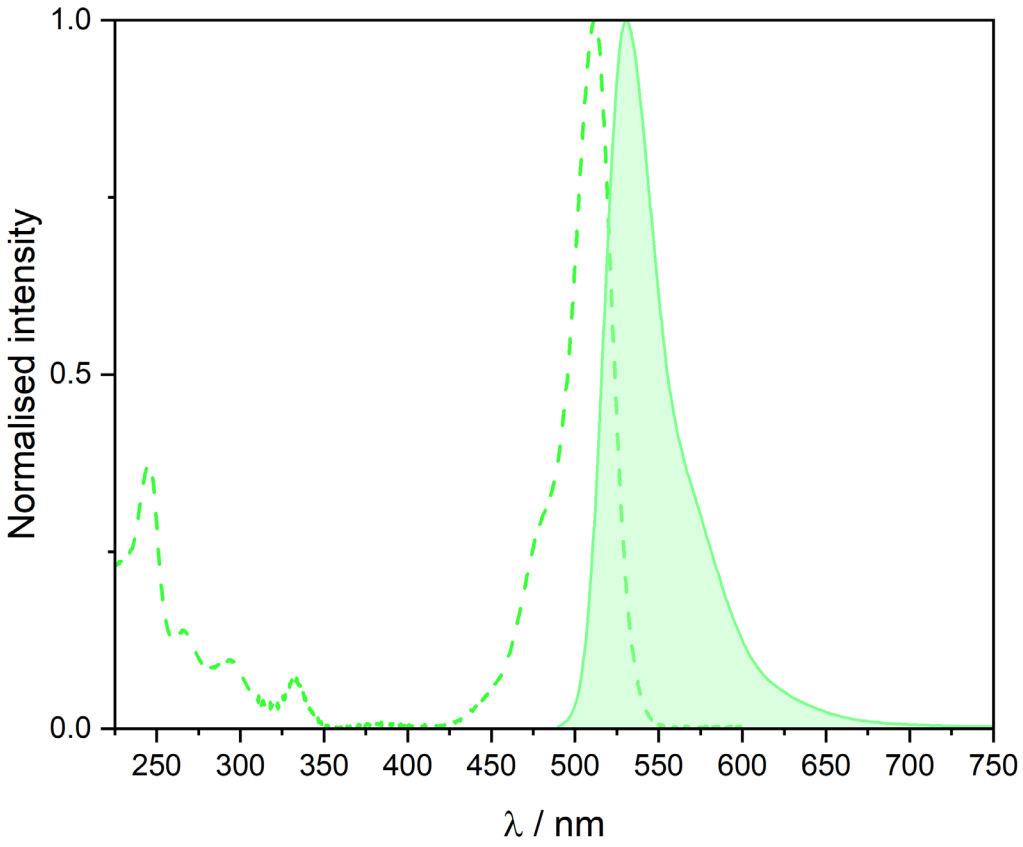



*Figure 1.12. Absorption (dashed line) and emission (shaded area) spectra of rhodamine 123 in aqueous solution. Spectra were obtained from PhotochemCad repository.[104]*

**HiLyte Fluor 488**

Hilyte Fluor 488 is a commercial dye from the family of the xanthene fluorophores. As seen in Figure 1.13, its core structure is the same as the rhodamine dyes, but it is a highly

functionalized derivate. It exhibits high fluorescence intensity with an emission maximum at 525 nm. This dye is quite photostable and is not greatly affected by the pH of the environment when working within the range of 4 to 10.[105–107] Moreover, it can also be easily conjugated to biomolecules such as proteins or antibodies. This binding is also specific, which allows for tailored labelling of the biomolecular systems of study. [92,108]

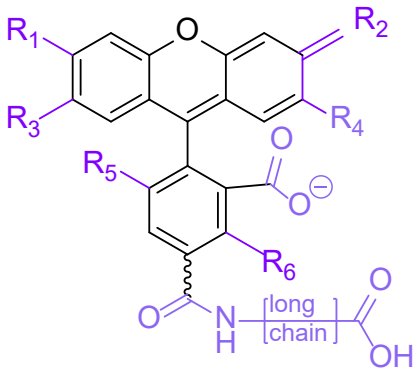



*Figure 1.13. Hilyte Fluor 488 fluorescence sensor molecular structure.*

Hilyte Fluor 488 is a widely used fluorophore in amyloid studies, as it presents a sensitive binding to amino residues in its acidic form. Amyloid labelling happens at the N-terminus of the peptide. Its binding to amyloid peptides allows for visualisation and characterisation of the different species that exist during the amyloid aggregation. Due to its emission properties, amyloid peptides labelled with Hilyte fluor 488 can be used in a great variety of fluorescence techniques, allowing even for real-time monitoring in living cells.[105,109–111]

In this thesis, Hilyte Fluor 488 is used used as amyloid peptide label to allow amyloid aggregation studies to be carried out with fluorescence correlation spectroscopy and super-resolution microscopy.

**Thioflavin T**

Thioflavin T (ThT) is a fluorophore from the benzothiazole family composed of benzothiazole moiety bind to a phenyl ring functionalised with a dimethylamine group at its para position (Figure 1.14). ThT is also a very common extrinsic fluorophore in the study of biomolecular systems with fluorescence-based techniques owing to its photophysical properties. It is thought to be a highly selective fluorophore that presents important spectral variations upon binding to different biomolecules.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

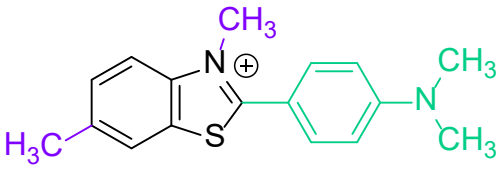



*Figure 1.14. Molecular structure of Thioflavin T.*

Thiofalvin T emission in water presents a maximum at around 454 nm with a quite low fluorescence quantum yield. However, its emission spectrum can change significantly both in intensity and in position when the environment the fluorophore is in contact with is modified. ThT is sensitive to variations in the polarity and viscosity of the media. For example, when changing the solvent from water to chloroform a spectral shift to 500 nm and a increase in fluorescence emission occurs.[112] Additionally, when binding to supramolecular systems like amyloids and nucleic acids a similar behaviour is detected where, even if the spectral shift is not large, the increase in its fluorescence quantum yield is of great importance. Numerous studies have been published about ThT photophysical behaviour, stablishing the reason of its useful properties on the dye being a molecular rotor. This sensor is able to undergo a twisted intramolecular change transfer (TICT) in the excited-state due to the ability of the phenyl ring to rotate, thereby changing the molecular structure from a completely planar to perpendicular (Figure 1.15).[112–115]

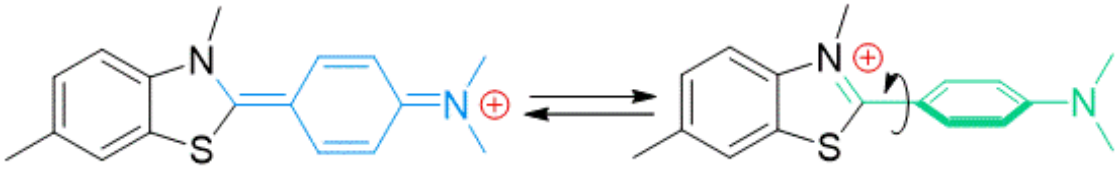



*Figure 1.15. Schematic representation of the molecular structures associated with the two limit conformations that ThT can adopt.*

In the context of AD, ThT has long been stablished as the main dye for amyloid detection and characterisation, especially for fibrils.[113,114,116–119] However, as it has a quite complex photophysical behaviour that makes it difficult to work with. ThT presents a strong solvent dependency as it can acquire a wide distribution of conformations with different behaviours in the ground state. Additionally, ThT can bind to other proteins like serum albumins with similar binding constants to that with amyloids, which makes it not specific enough.[112] Although ThT is indeed still a great proposal for Aβ detection, researchers are still working on new fluorescent probes that could give better results.[118,120–123] Recent studies also suggest that ThT can selectively bind as well to nucleic acids allowing the identification of G4 and other DNA conformations.[124–128]

In this thesis, ThT was used during the study of G-quadruplexes in the fluorescence assays for detection and characterisation of the different G-quadruplexes topologies that our system presented.

USC
DE SANTIAGO
DE COMPOSTELA

**2-cyano-6-hydroxybenzothiazole**

2-cyano-6-hydroxybenzothiazole (CBTOH, Figure 1.16) is a fluorescent molecule from the benzothiazole family. Benzothiazoles are commonly used as fluorescence sensors as they present high photostability and photophysical properties that vary depending on the functional groups present at the second position of the heterocycle as well as at the benzene. They usually present high absorption coefficients and fluorescence quantum yields, which make them appealing for spectroscopic studies. They are used in a variety of fields including agriculture (for insecticides), new materials (as semiconductors) and in pharmacology (as antimicrobial agents).[129–132]

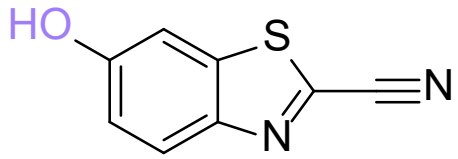



*Figure 1.16. 2-cyano-6-hydroxybenzothiazole molecular structure.*

CBTOH is an intermediate in the biosynthesis of D-luciferin. In the cycle D-luciferin is first oxidised by an enzymatic reaction resulting in the formation of oxyluciferin. Then, oxyluciferin is once again oxidized yielding CBTOH and thioglycolic acid. The formed CBTOH reacts with cysteine by a condensation process regenerating into luciferin. This biological process that takes place in fireflies has been studied in the biological fields leading to new applications in in vivo imaging and protein labelling. Previous studies done by our research group reported that CBTOH emissive properties depend on the solvent nature and on the pH of the environment.[133] CBTOH is both a photoacid and a photobase, as the acid-basic properties of both the hydroxyl group and the thiazole nitrogen change in the first singlet excited state. In the ground state, CBTOH exhibits a p$K_a$ of 8,01 ± 0.02, but upon transition into the excited state, the p$K_a$ drastically decreases resulting in a value of −1.8. This variation is due to the intramolecular charge transfer that takes place upon excitation to the first-excited singlet state, and triggers proton-transfer processes if the solvent can support them. In water, a proton-transfer from the hydroxyl group to the solvent takes places in the excited-state, causing a spectral change which shifts the emission maximum from 430 nm to 504 nm. Additionally, the nitrogen atom from the heterocycle can undergo protonation in the excited state, leading to the formation of a zwitterionic species. This protonation occurs after the deprotonation of the hydroxyl and also provokes a spectral shift from 504 nm to 625 nm.[133]

Considering the photophysical behaviour CBTOH presents, it could also be a great potential fluorescence sensor for biomolecular systems. The molecule binding to another one could promote spectral variations as the fluorophore environment changes. In this

thesis, it will be studied as a candidate for human serum albumin detection. As mentioned before, it is crucial to find a fluorophore with better properties than ThT for AD detection. We believe that finding the correct fluorescence probe that would selectively bind to HSA in the presence of amyloid or that can present a significant spectral change between free HSA and bound HSA is of high importance. The use of this fluorescence molecule could quantify the association between both molecules.

## This thesis

In this thesis, the self-assembly and supramolecular association processes of the mentioned biomolecules related to Alzheimer's and other neurodegenerative diseases were studied. To obtain the equilibrium constants and the aggregation dynamics, we used spectroscopic and microscopic techniques. Apart from the conventional absorption and fluorescence spectroscopies, single-molecule fluorescence techniques to obtain structural and dynamic information of the systems under study were also used. Both Fluoresce Correlation Spectroscopy (FCS) and super-resolution microscopy were key techniques during this thesis. In addition, some studies were made with circular dichroism (CD) spectroscopy. The combination of all these techniques allow to understand the complex mechanisms of the processes studied.

In **section 5.1**, we present the results of an already published study on the formation of G4s in mitochondrial nucleic acids sequences from the CSB II.[27] The tendency of DNA, RNA and a DNA:RNA hybrid between the DNA and RNA sequences to form G4s was investigated. A study of the structure of G4s and the influence of monovalent cations ($K^+$ and $Na^+$) on the stability of G4s was also carried out. Here, the main technique was circular dichroism as it allowed us the study of all sequences without the need of an additional fluorescent molecule. In addition, gel electrophoresis was used as a complementary technique. To get a further insight into the topologies appearing in the systems of study, fluorescence studies were needed. Thus, steady-state and time-resolved fluorescence emission and fluorescence anisotropy techniques were also used.[27]

In **section 5.2**, the results obtained on the aggregation process of amyloid-beta (1-40) peptide will be discussed and compared to the previously studied process of amyloid (1-42). For this study, FCS was the only technique used. With FCS measurements, we were able to characterize in size and shape both the monomeric and the early oligomeric species of Aβ40. In addition to the characterization, the critical aggregation concentration of the peptide was calculated. Whether or not the aggregation process for Aβ40 is reversible will also be discussed within this chapter.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

The results shown in **section 5.3** focus on amyloid's autofluorescence and its potential use as a detection method for amyloid aggregation. Using steady-state fluorescence emission in combination with FCS, we studied the aggregation process of Aβ40 and if it led to any spectral changes in its emission spectra. The obtained results will serve to propose a model based on amyloid's autofluorescence for amyloids aggregation.

**Section 5.4** reports the results on the influence on amyloid aggregation and liquid-liquid phase separation of electrostatic interactions between amyloids and small charged biomolecules (polyamines like spermine and spermidine, and $Mg^{2+}$), using super-resolution microscopy. We first studied if amyloid undergoes LLPS during its aggregation at different incubation times. Then, we investigated the influence of the mentioned cations on the LLPS before and after fibril formation. This work was carried out under the supervision of Professor Sang-Hak Lee during a six-month research stay at Pusan National University in South Korea.

As the last part of the thesis, **Section 5.5** will present 2-cyano-6-hydroxibenzothiazole (CBTOH) as a potential fluorescent probe for human serum albumin detection. First, we studied the acid-base properties of the molecule in physiological conditions. Then, we investigated if CBTOH could bind with bovine and human serum albumin and quantified the binding process using spectroscopic techniques like UV-vis absorption, steady-state fluorescence emission and anisotropy and time-resolved fluorescence emission.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# 2. Objectives

The general objective of this thesis is the study of different supramolecular processes of self-assembly and host-guest interaction in biomolecules related to different diseases, especially to neurodegenerative diseases like Alzheimer's. Diseases are usually studied from different fields to comprehend their development and find both a cure and a preventive method. However, in most cases the studies are done primarily from a medical or pharmacological perspective. For this reason, this thesis aims to bridge the knowledge gap between the fields by studying the structural changes mechanisms that key molecules in Alzheimer's disease undergo.

For this goal, in this thesis we made use of conventional and groundbreaking spectroscopic and microscopic techniques. By a thorough data analysis, we seek to obtain information about the structure and the dynamics of association of the different biomolecules under study. Various factors that can impact association and interactions will be studied: concentration, charges, and time stability.

Then, specific objectives were set up for each study and are as follows:

- **Spectroscopic Characterisation of Mitochondrial DNA G-Quadruplexes (Section 5.1).**
  G-quadruplexes formation in different mitochondrial DNA, RNA and hybrid DNA:RNA sequences will be studied. All studied sequences are sequences that can be found in the CSBII region of the mitochondria. The formation of non-canonical DNA conformations will be studied as well as other possible conformational changes that may take place under the studied conditions. Circular Dichroism will be the main technique of this study, as it allows to follow conformational changes without the need for an external fluorescence molecule. Complementary techniques will also be used to corroborate the obtained results: gel electrophoresis, and fluorescence emission and anisotropy (both steady-state and time-resolved) experiments will be performed. Characterisation of G-QP will be carried out under the studied conditions, obtaining the equilibrium constant of the formation. Cation power of stabilisation will be studied as well, seeing which cation presents a greater impact. Length sequence and sequence's nature influence on G-QP formation will also be studied.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

- **Critical aggregation concentration and reversibility of Aβ(1-40) oligomers (Section 5.2).**
  Following the group's experience with amyloid (1-42) study, we will study the aggregation of amyloid (1-40). Using FCS as main technique of the study, aggregated amyloid fraction curves based on the solution's total concentration of amyloid will be obtained. This will allow us to obtain the critical concentration at which AB40 aggregates are formed and to compare the aggregation process with the AB42's process. Shape and size of the aggregates will be also obtained and compared to the ones obtained for AB42. Additionally, a reversibility study will also be carried out as well as the time evolution of the formed aggregates.
- **Amyloid autofluorescence as method for amyloid's aggregation detection (Section 5.3).**
  Using emission fluorescence and FCS simultaneously AB40's aggregation will be studied. Using the fluorescence that the tyrosine present in AB40 sequence yield, we will follow the aggregation process and see what changes it may cause in fluorescence emission. We will study if the aggregation process can be followed using a widely available conventional technique, like fluorescence emission is, and propose a model that links the spectral changes with the degree of aggregation. FCS experiments will help corroborate that the obtained results from autofluorescence are accurate.
- **Electrostatic interactions impact on amyloid-beta aggregation (Section 5.4).**
  Prof. Lee's group from Pusan National University (South Korea) main research interest is the role of small-charged biomolecules have in neurodegenerative diseases. Given its expertise, AB42 aggregation in presence of small-charged biomolecules will be studied using super-resolution microscopy. The obtained images will then be analysed with Super-Resolution Radial Fluctuations for the obtention of super-resolution images. First, the studied of how charge biomolecules like polyamines can affect fibril formation by electrostatic interactions will be studied. We will study if previously formed fibrils undergo any change when exposed to molecules that can bind by electrostatic interactions. We will also study how the concentration of said charged molecules affects fibrillation. The same study will be carried out for different charged molecules (spermine, spermidine and $Mg^{2+}$) found in our organism, each of which has a different net charge of +4, +3 and +2 respectively. This way, we will study how charge affects fibrils formation and if it can reverse the process.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

- **2-cyano-6-hydroxybenzothiazole as fluorescence probe human serum albumin detection (Section 5.5).**
  2-cyano-6-hydroxybenzothiazole will be studied as a probe for detection of blood plasma proteins related to Alzheimer's disease (serum albumins). First, we will characterise the probe under physiological conditions, as previous studied from our group were performed only in water. Then, the host-guest association of the probe to different serum albumins (bovine and human) will be studied by different spectroscopic techniques. Both steady-state and time-resolved techniques will be used to obtain the association constant and characterise the host-guest interaction between said proteins and the fluorescence probe. This way, we will study if the fluorescence probe is adequate for protein detection and, if it is, it will present a first step towards amyloid detection as human serum albumin is the protein in charge of amyloid transportation from the bloodstream to the cerebrospinal fluid.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# 3. Experimental methods

## 3.1. Experimental Techniques

### 3.1.1. Conventional Spectroscopic Techniques

#### 3.1.1.1. UV-Vis Absorption

All UV-visible absorption spectra were obtained using a double beam Cary 100 spectrophotometer from *Varian* (Figure 3.1). It has two different light sources that allow for a wide spectral range from 190 nm to 900 nm. The light source for the visible region of the spectra (350-900 nm) is a tungsten lamp, while the one for the UV region (190-350 nm) is a deuterium lamp. As shown on Figure 3.1, it has one monochromator and one pre-monochromator for wavelength selection. Excitation beam direction is controlled by a chopper, which rapidly alternates the path between the sample and the reference to continuously measure both signals. Then a second chopper alternatively leads the beam of each path into the detector, which is a PMT (photomultiplier tube). Having a double beam allows for spectral correction with a reference. Moreover, the continuous comparison of the sample and reference signals increases the stability of the measurement, efficiently correcting the signal drift over time due to lamp or detector instabilities.

For temperature control, the spectrophotometer was connected to a thermostat from VWR. The sample holder was connected directly to the thermostat. Temperature was kept constant at 25.0 ± 0.1 $^{o}$C.

For absorption measurements, 10 mm x 2 mm quartz absorption cuvettes from *Hellma* were used. This type of cuvette is meant to be used with lower volumes than conventional 10 mm x 10 mm cuvettes, which makes them appropriate for use with relatively expensive samples. The optical path length of these cuvettes is10 mm. For the reference solution, 10 x 10 mm quartz cuvettes were used, as they have the same optical path length, and they are easier to clean.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

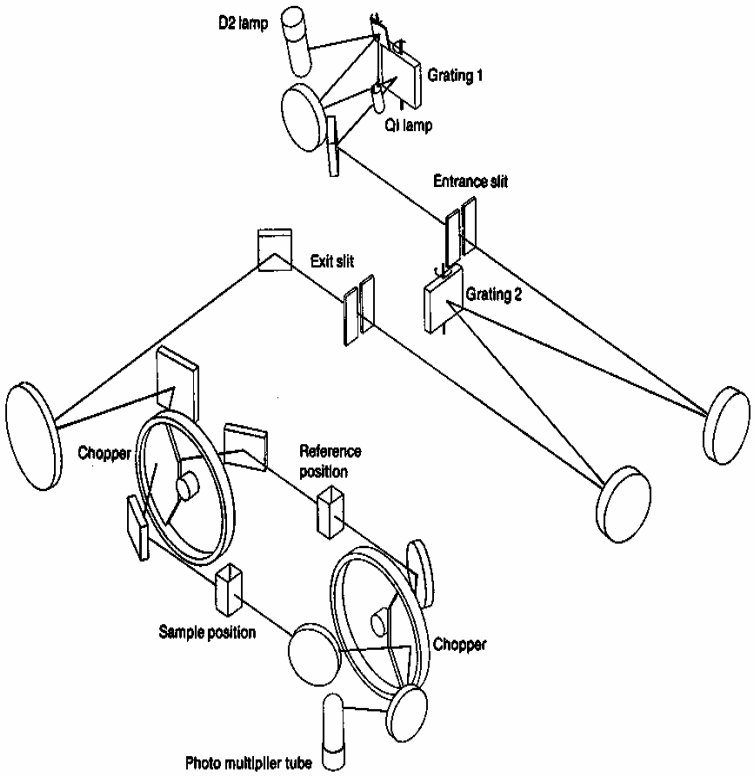



*Figure 3.1. Optical scheme of the UV-Vis absorption Cary 100 spectrophotometer, as published by the manufacturer.*

Calibration of the equipment was performed before carrying out the experiments. For a better baseline calibration that could be used even with high absorbing samples, a two-step calibration was done. First, a conventional baseline measurement was obtained, which consisted of a 100% transmittance solution. Then, the light path was blocked with a black piece that allowed to measure 0% transmittance.

### 3.1.1.2. Circular Dichroism

Circular dichroism (CD) is a widely used spectroscopic technique in different disciplines that include chemistry, structural biology, material science and others. It can be used to analyse the secondary structure, conformational changes, and interactions between molecules if there is a certain degree of asymmetry. Thus, it is very useful for chiral molecules such as proteins and nucleic acids.

The foundation of this technique is very similar to the one used in UV-Visible absorption spectroscopy. However, CD uses circularly polarised light. The sample is irradiated alternatively by right-handed circularly polarised light and left-handed circularly polarised light, obtaining as a result the difference in absorption between both sides (Eq. (3.1)). After passing through the sample, the light polarization can undergo a transformation resulting in the elliptical polarization of the light. This means that the polarization vector traces an ellipse in its plane of propagation. This transformation takes

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

place due to the differential absorption that chiral molecules show between the left-handed and right-handed circularly polarized light. Because of this, usually CD spectra are represented in terms of ellipticity ( $\theta_{exp}$ , Eq (3.2)) or molar ellipticity ( $\theta$ ). Molar ellipticity can be calculated following Eq. (3.3), where C is the concentration in mol $dm^{-3}$, l is the light path of the cuvette in cm and x is the ellipticity in deg.

$$\theta_{exp} = \Delta A = A_{left} - A_{right} \tag{3.1}$$

$$\Delta A = \frac{4\pi x}{180 \ln(10)} \tag{3.2}$$

$$\theta = \frac{100\theta_{exp}}{Cl} \tag{3.3}$$

As the light of the excitation source is circularly polarised and the obtained data is the difference between left-handed and right-handed light, samples for CD need to present a certain degree of chirality. Chirality of the samples may be intrinsic to them, or it may be induced by the environment. However, if no chirality exists no signal will be detected in CD. This happens because when any sort of asymmetry is present in the sample, its optical properties will also vary depending on the polarisation direction of the light. The obtained spectra may look similar to the absorption spectra in shape, but they can also show negative ellipticity values depending on which direction absorbs more light. Figure 3.2 shows different examples of hypothetical CD spectra compared to absorption spectra on samples that show chirality and a sample that does not.

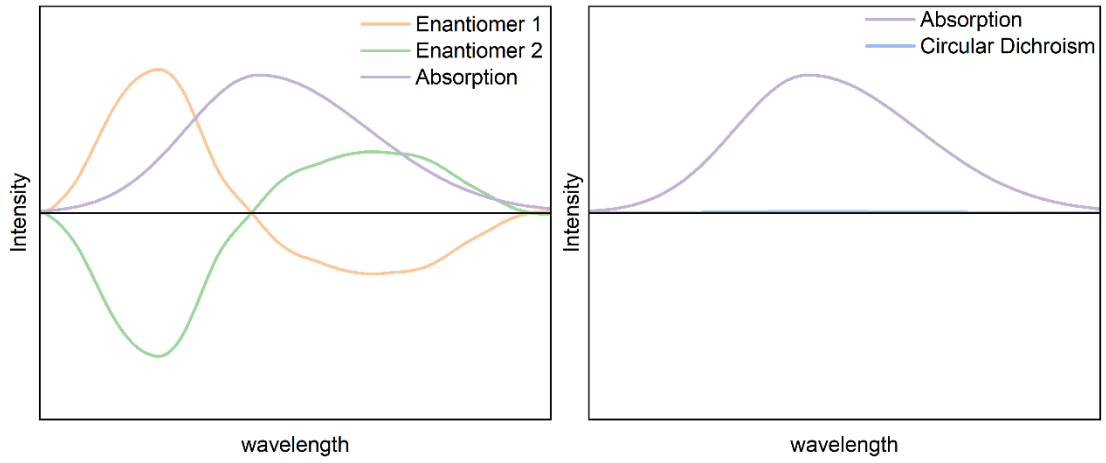



*Figure 3.2. Simulated circular dichroism and absorption spectra. Left: circular dichroism spectra for two enantiomers and its absorption spectrum. Right: CD and absorption spectra of a non-chiral molecule.*

For CD spectra acquisition, specific equipment is needed (Figure 3.3 and Figure 3.4). CD equipment use a xenon lamp as light source that allow for a wide work range at low wavelengths. Usually, the sensitive range goes from 180 nm to 800 nm. As the spectral range includes the distant UV, which is highly energetic, the equipment needs to be under

USC
DE SANTIAGO
DE COMPOSTELA

nitrogen purge before and during use to prevent oxygen from reacting with UV radiation to form ozone. To obtain circularly polarized light in both directions, a photo-elastic modulator (PEM) is used. PEM transforms linearly polarised light that comes from the polariser to circularly polarised light.

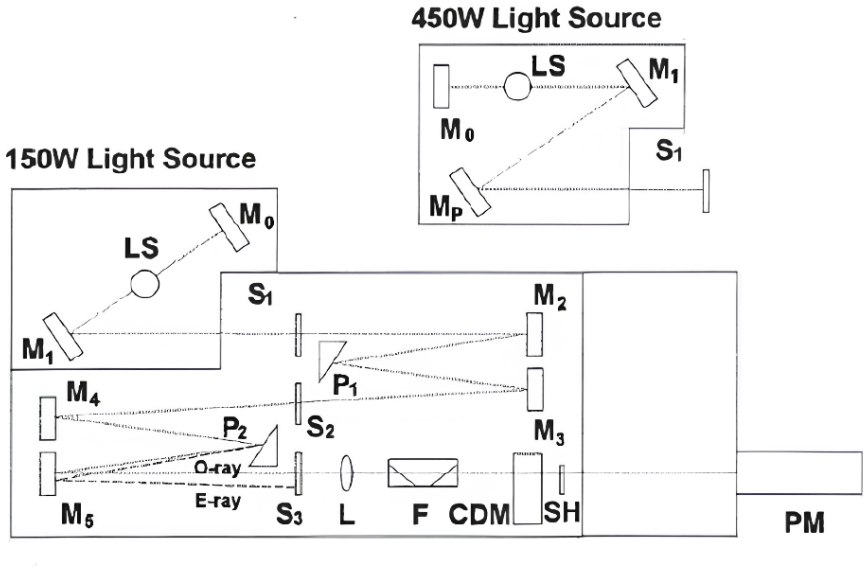



*Figure 3.3. Jasco-715 spectropolarimeter optical scheme, as published by the manufacturer.*

For this thesis, all regular circular dichroism spectra were acquired on a Jasco-715 spectropolarimeter (Hachioji, Japan; Figure 3.3) at 20 ºC in a 2 mm quartz cuvette (Hellma, Germany). Melting experiments were performed under phosphate buffer conditions with a 2 mm quartz cuvette (Hellma, Germany) and a temperature range from 20 to 95 ºC on a Jasco-1100 spectropolarimeter (Hachioji, Japan; Figure 3.4). The absorption spectra of the samples were also recorded to determine their concentrations for the correction of the CD spectra.

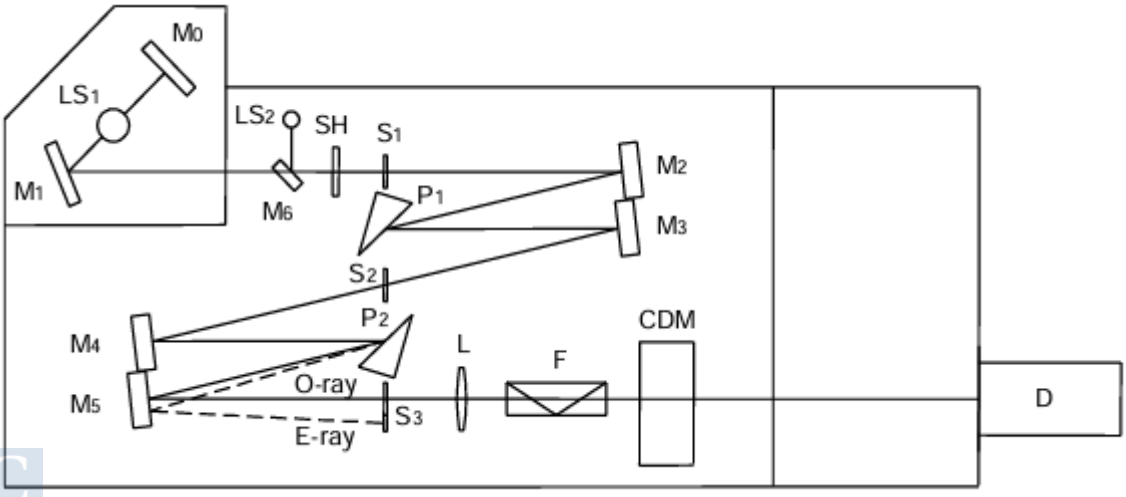



*Figure 3.4. Optical scheme of the spectropolarimeter Jasco-1100, obtained from the manufacturer manual.*

USC DE COMPOSTELA

### 3.1.1.3. Steady-state fluorescence

Steady-state emission spectra are obtained by fixing the excitation monochromator on a wavelength were molecule of study presents absorption and changing the emission monochromator wavelength within a desired range on which the molecule presents emission. On the contrary, steady-state excitation spectra can be obtained by fixing the emission monochromator to a wavelength were the studied molecule presents emission and changing the excitation monochromator within the wavelength were the molecule presents absorption. The excitation spectrum should coincide with the absorption spectrum for a given fluorescent species if no other species exist in the ground state. On the contrary, the spectra differ if there are several different species in the ground state, which can also be formed by a single molecular species (i.e., aggregates, tautomeric forms, etc.).

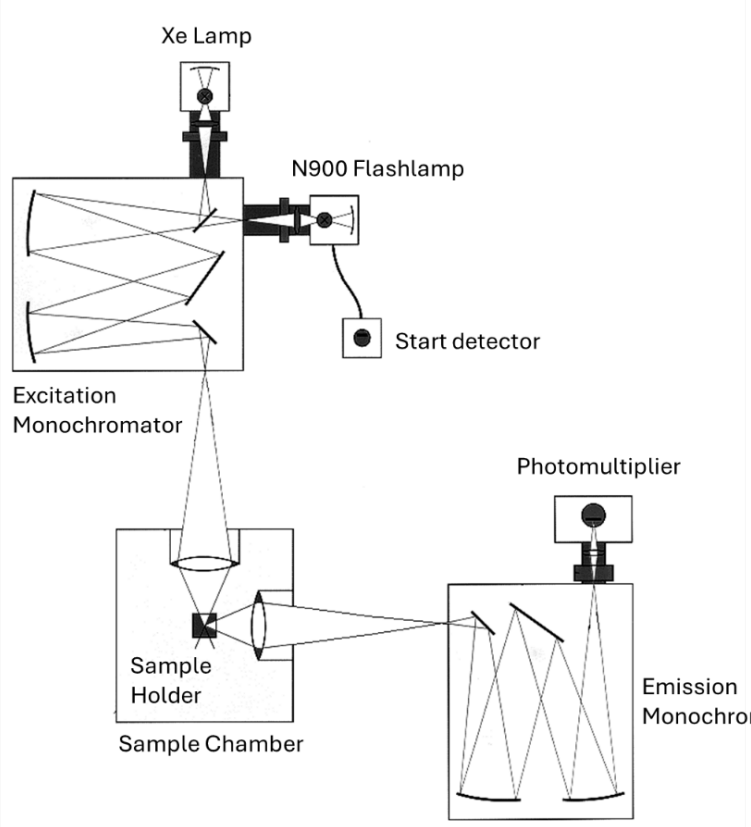



*Figure 3.5. Edinburgh Instruments F900 spectrofluorometer optical scheme, as published by the manufacturer.*

Steady-state fluorescence emission and excitation measurements as well as steady-state anisotropy measurements were carried out using F900 from Edinburgh Instruments (Livingstone, UK) for DNA studies. As shown in Figure 3.5, F900 spectrofluorometer uses as excitation source a 450 W Xenon Lamp which allows the spectral range to be from 200 to 2600 nm. However, the actual spectral range of the equipment if 200-900 nm because of the monochromators limit. F900 has two monochromators, one for excitation located

USC DE COMPOSTELA

before the sample chamber, and the other for emission located after the sample chamber. Both monochromators use a Czerny-Turner diffraction grating (1800 g/mm) and have manually variable slits ranging from 10 µm to 10 mm. The sample chamber has three different sample holders that can be moved controlled by the software. Sample's temperature is controlled by a thermostat keeping the temperature constant at 25 ± 1ºC. As a detector, F900 has a photomultiplier tube. Additionally, right after the sample chamber, F900 is equipped with manual polarisers allowing for fluorescence anisotropy measurements to be performed as well.

All amyloid studies were carried out using a FS5 spectrofluorometer from *Edinburgh Instruments* for steady-state emission, excitation, and anisotropy measurements. Figure 3.6 shows the equipment scheme. We used quartz cuvettes from Hellma with a pathway of 1.5 x 1.5 mm, 3 x 3 mm or 10 x 10 mm, depending on the experiment. The microcuvettes were used to avoid using high quantities of amyloid peptide.

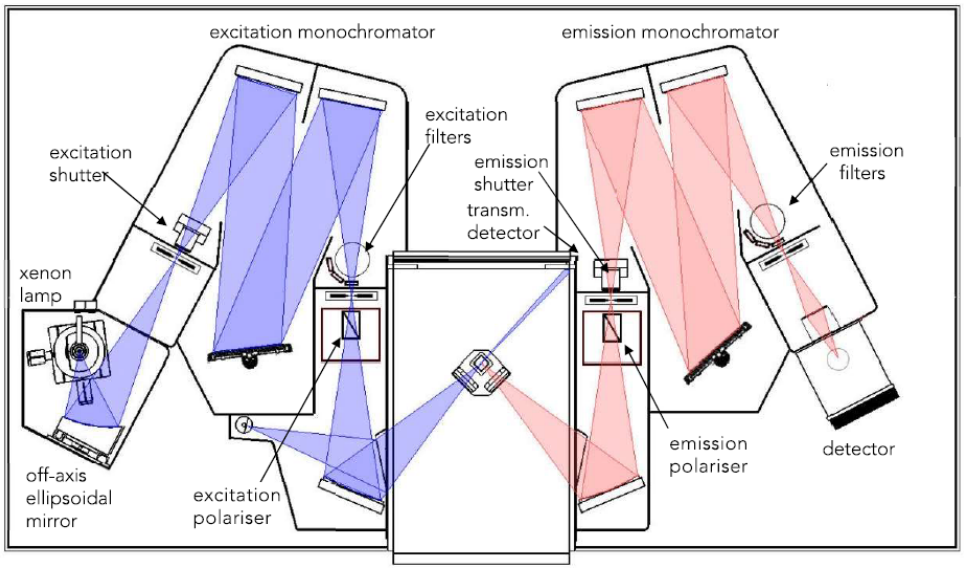



*Figure 3.6. Edinburgh Instruments FS5 spectrofluorometer optical scheme, as published by the manufacturer.*

The FS5 spectrofluorometer uses a Xenon Arc 150W lamp as excitation source. It has an excitation and emission monochromator that allows to choose wavelength from 200-1000 nm. An excitation shutter is controlled by the software, so it only opens when a measurement is running or when checking the signal rate. This is for prevention of the photodegradation of the sample. The emission shutter is connected to the sample chamber's lid. Thus, it will only open when the lid is closed. This is to prevent damage of the detectors by over-exposure to bright light. It also has excitation and emission filters that prevent the appearance of second and third order diffraction bands. It is equipped with three different detectors. First, the reference detector, an UV enhanced silicon photodiode, which is used to correct for undesirable source-light-intensity shifts. The

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

transmission detector and the signal detectors are both R928P type temperature-stabilised photomultiplier detectors.

Additionally, for steady-state fluorescence anisotropy measurements, this spectrofluorometer is equipped with two motorized polarisers, one for excitation and one for emission. Both polarisers have αBBO polarising prims which are controlled by the software with an accuracy of 0.5º.

All steady-state fluorescence experiments in **section 5.3** were performed with two identical FS5 spectrofluorometers, one in the laboratory of Lugo (FS5 A) and the other in Santiago (FS5 B). The reason for specifying the apparatus used is the distortions observed in the Lugo equipment, which appear to be due to a poor adjustment of the optical path as. For all experiments of this chapter, both emission and excitation slits were set to 5 nm.

In the case of human serum albumin detection (**section 5.5**), for steady-state fluorescence studies 2.0 nm excitation and 1.5 nm emission slits were employed, while for anisotropy studies 2 nm and 3 nm slits were used for excitation and emission respectively.

### 3.1.1.4. Time-resolved fluorescence

***Time-Correlated Single Photon Counting***

Time-resolved fluorescence measurements can be carried out using two different methodologies: pulsed light methods and frequency-domain methods. Among the pulsed light methodology, different techniques have been developed, like Time-Correlated Single Photon Counting (TCSPC) and Multichannel Scaling (MSC). All of them offer both advantages and disadvantages and are usually chosen based on several factors like sensitivity, temporal resolution, and compatibility.

In our group, the available technique is TCSPC, which we used for all fluorescence lifetimes measurements. TCSPC is widely used in different fields from chemistry, biology, and physics. While it was first introduced as a tool for fluorescence lifetime acquisition, it is also used in other fluorescence techniques like Fluorescence Correlation Spectroscopy (FCS, which will be explained later in this chapter) and Single-Molecule Spectroscopy as well as other areas like Quantum Cryptography or Time-Domain Diffuse Optical Tomography.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

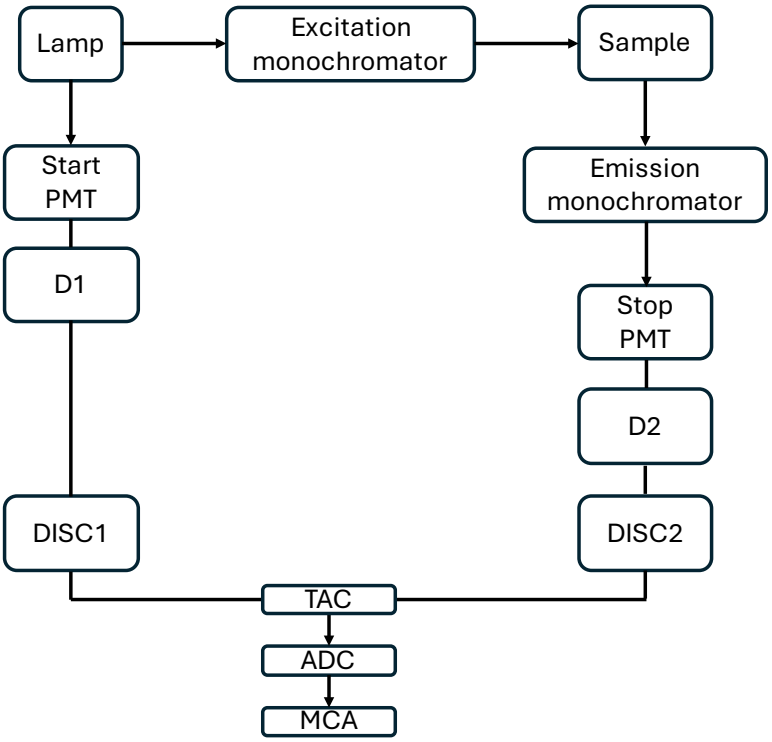



*Figure 3.7. Diagram of a conventional TCSPC equipment. Lamp is the excitation source, Start and stop PMT are the start and stop photomultiplier respectively. D1 and D2 are the retardation lines. DISC1 and DISC2 are the discriminator for the trigger and the fluorescence signal respectively. TAC is the time-amplitude transformer, ADC is the analogical-digital converter and MCA is the multichannel analyser. The lamp is nowadays usually replaced by a monochromatic laser or LED source, in which case no excitation monochromator is needed.*

TCSPC operates on the principle that the probability distribution of emitting a particular photon following excitation aligns with the intensity-time distribution of all emitted photons resulting from excitation. Therefore, by determining the first photon that arrives at the detector after excitation and conducting a large number of excitations, it is possible to know the temporal dependence curve of the fluorescence intensity of the sample under study for a given emission wavelength. The TCSPC technique has the advantage of its high sensitivity, allowing the measurement of short times and multi-exponential decays with exceptional accuracy.

The technique consists of sample excitation with a light pulse, usually generated by a laser. This pulse excites the sample while it is detected by the start photomultiplier tube (start PMT, Figure 3.7), triggering an electric signal that will activate the time to amplitude converter (TAC). After activation, TAC generates a voltage linearly increasing with time. When the stop photomultiplier (stop PMT in Figure 3.7) detects the first emitted photon, a new pulse that stops the voltage increase in the TAC. This way, the accumulated voltage will be proportional to the time interval between both start and stop pulses. Then, the multichannel analyser (MCA) transforms the accumulated voltage value into a digital value and stores one count in the corresponding channel. This process is repeated for a high number of excitations, usually until 10000 counts are registered in the maxima's

channel for better precision. However, data can be acquired by a set time instead of desired maxima counts. If all data is obtained until 10000 counts the time it takes each measurement to finish will vary depending on its intensity and that information is stored for analysis purposes. Instead, if an appropriate time is set for all samples to obtain reasonable maxima counts, then the information on difference in intensity is also obtained. When the measurement is finished, a fluorescence decay curve is obtained, which consist of a histogram with very narrow intervals.

Typically, emission decay will follow a first order kinetic ( $I = I_0 e^{-kt}$ ). So, if we represent fluorescence intensity in logarithmic scale versus time, we shall obtain a straight line if there is only one emitting species in our sample and the lifetime is long enough in comparison to the excitation pulse duration. Nevertheless, if this last condition is not met, a deconvolution analysis of the instrument response function (IRF) and the exponential decay of the sample is needed. Moreover, more complex decay patterns can happen, for example if more than one species is emitting, so in these cases a multiexponential fitting is needed (see section 4.1.4 in **Data Analysis** for further explanation).

Using TCSPC offers several advantages. First, because it is based on low-intensity light measurements, the sensitivity it offers is quite high. Also, photon counting obeys a perfectly defined Poisson distribution, meaning that the standard deviation will be the square root of the number of acquired counts. However, one should pay attention to its disadvantages as well. There are other effects that can give a distorted curve: the pile-up effect. If more than one photon reaches the stop PMT the decay will be shifted towards lower times than it should because the second photon cannot be considered as TAC is not activated when it reaches the stop PMT. To prevent pile-up from happening fluorescence intensity must always be reduced so the sample photon frequency is no higher than 2% of the excitation frequency.

***Time-resolved fluorescence equipment***

Time-resolved emission spectra were obtained using a LifeSpec-ps equipment from Edinburgh Instruments (Livingstone, UK). The equipment allows for two lasers to be connected at the same time. Thus, a 375 nm diode laser from PicoQuant (PicoQuant LDH-P-C-375) was always kept connected and the other laser was changed depending on the requirements of the experiment (Table 3.1). The equipment also allows to measure time-resolved anisotropy, as it has a Glan Thompson polarizer. The detector is a Hamamatsu multichannel plate photomultiplier (model R3809U-50) with a response time of 50 ps and a wavelength detection range of 200 nm – 850 nm. The equipment uses the technique Time Correlated Single Photon Counting (TCSPC). The instrument response function (IRF) was obtained before each measurement using a scattering solution in a 10 x 10 mm quartz cuvette. The emission wavelength was selected to match the wavelength of the excitation source and the decay was recorded to obtain a peak maximum of 10000 counts.

USC UNIVERSIDADE DE SANTIAGO

A value of IRF ≈ 100 ps full width at half maximum (FWHM) was obtained with a 60 ps FWHM laser as excitation source.

*Table 3.1. Excitation wavelength, type and brand of the different lasers and LEDs used as excitation source in time-resolved fluorescence measurements.*

| Wavelength/ nm | FWHM / ps | Type | Brand |
|---|---|---|---|
| 340 | 1050 | Pulsed LED, EPLED | Edinburgh Instruments |
| 445 | 85 | Pulsed laser, EPL | Edinburgh Instruments |

Time-resolved fluorescence measurements in **section 5.1** and **section 5.5** were performed with 2 mm slits for emission. For excitation, in QP studies (**section 5.1**) a 445 nm picosecond diode laser for excitation, while for HSA studies (**section 5.5**) a 375 nm diode laser and a 340 nm LED were used. For both of these studies, 3 x 3 mm quartz cuvettes were used.

## 3.1.2. Fluorescence microscopy techniques

### 3.1.2.1. Fluorescence Correlation Spectroscopy

***What is Fluorescence Correlation Spectroscopy***

Fluorescence Correlation Spectroscopy (FCS) was first introduced in 1972 by Madge and collaborators.[134] Despite its long history, FCS remains an evolving and expanding technique. It has become a widespread method for the study of biomolecular processes, offering researchers a non-invasive means to investigate phenomena at the single-molecule level.

The technique analyses the fluorescence intensity fluctuations of a freely diffusing sample coming in and out of the focus volume of a confocal microscope. From the correlation of the observed fluorescence intensity fluctuations over time, FCS is able to obtain both chemical and physical parameters like molecular diffusion, flow rates, triplet state lifetimes, binding kinetics, and conformational changes with exceptional sensitivity and temporal resolution. Each process will have an individual correlation time in their corresponding timescale, which can range from milliseconds (molecular diffusion) to picoseconds (triplet state relaxation or rotational diffusion).[135,136]

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

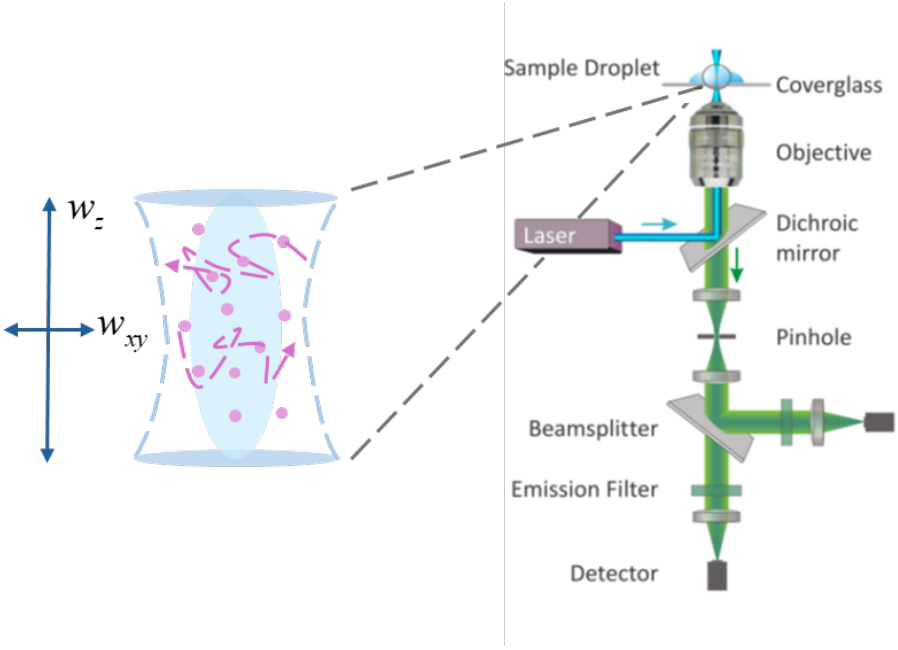



*Figure 3.8. Fluorescence Correlation Spectroscopy principles. Right panel: schematic representation of a confocal microscope used for FCS curves acquisition. Left panel: representation of dye molecules (pink dots) diffusing freely in and out of the microscope's focus volume (in blue).*

As samples used in FCS are in solution, the molecules are continuously diffusing freely within the sample volume, coming in and out of the focus of the confocal microscope (Figure 3.8). The volume of the microscope focus is defined by an ellipsoid as shown in Figure 3.8 (left). The fluorophore concentration within this volume changes as the molecules move through it in a random Brownian motion. Because the number of fluorescent molecules in the focus changes with time, the emission intensity is also affected, generating intensity fluctuations. Additionally, to the diffusion of the molecules, other processes can take place that also alter the emission intensity of the sample. The influence of each process in a given correlation curve depends on the system and the conditions it was measured in. They usually have distinct time scales and can be distinguished easily because of this. In the following, we briefly discuss some of these processes.[136–139]

- **Antibunching:** When a molecule is excited, it remains in the excited state for a certain time before it deactivates by emitting a photon. During this time, the molecule cannot absorb again. This means that the time elapsed between consecutive emitted photons depends on the excited-state lifetime. This response time interval between the emission of two consecutive photons by the same molecule is intrinsic to the fluorophore and is determined by the lifetime of the dye in the excited state. This is known as antibunching. The typical time scale of antibunching processes is of nanoseconds or under. Usually, it appears as a sharp decrease in the correlation curve at very low correlation times.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

- **Rotational Diffusion:** The probability of exciting a dye with light depends on the angle between its transition dipole moment and the polarization vector of the exciting light. In a freely rotating fluorescent molecule, its excitation and emission rates fluctuate. This leads to a decay in the correlation function, with a characteristic rotational correlation time. Generally, this term is difficult to detect, and, when it does appear, it has correlation times in the nanoseconds scale.
- **Triplet state:** Typically, the excitation of an organic molecule leads to a singlet excited state whose deactivation process can occur directly through fluorescence emission or it can transition from the singlet state into the triplet state. The transition of the dye to a triplet state produces a fluctuation in fluorescence intensity, as the molecules in said state do not emit fluoresce. The number of molecules than can transition into the triplet state depend on the light source irradiance. At high irradiance levels, more molecules transition to the triplet state, reducing fluorescence emission. These molecules return to the ground state more slowly, contributing to the triplet relaxation time included in the correlation curve. The amplitude of the triplet contribution to the correlation curve accounts for the portion of molecules that exist in this state. Triplet relaxation times often appear in the microsecond scale, even up to a few milliseconds. To reduce the impact of the triplet relaxation time, that can overlap with the diffusion time if its slow enough, it is advisable to use low irradiances.
- **Translational diffusion:** The mobility of the fluorescent molecules in the sample cause variations in the number of detected photons per unit time. The faster the molecules diffuse in solution; the more fluctuations will be detected. The velocity at which a molecule can diffuse depends on its size. Thus, bigger molecules will exhibit slower diffusion and small molecules will present longer diffusion times. For this reason, diffusion times can range from microseconds up to seconds.
- **Reaction:** Interactions between the dye and other molecules, such as through reaction or association processes, can cause changes in its fluorescence intensity. The reaction correlation time that results from these processes taking place is directly related to the direct and reverse reaction rates of the reaction. Reaction times occur in a very wide time range, and thus only fast reactions can be seen in FCS without overlap with the diffusion time.

All the mentioned processes correspond to a region in the correlation function defined by their characteristic time scales. Figure 3.9a shows how the occurrence of a binding process between two fluorophores, combined with the translational diffusion of the fluorophore, result in different time scales fluctuations. These fluctuations then lead to a correlation curve similar to the one shown in Figure 3.9b.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

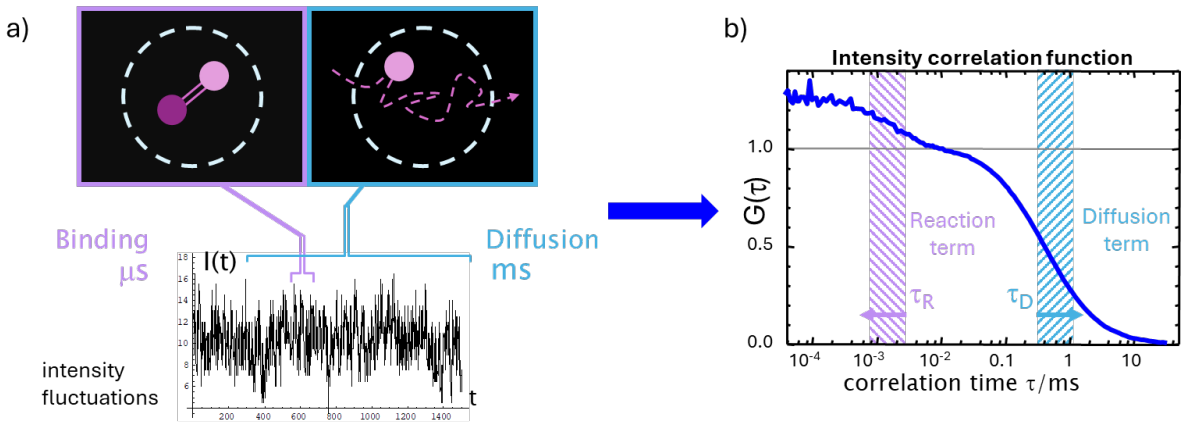



*Figure 3.9. Scheme of the correlation curve obtention. a) Left-up and right-up panels represent dye molecules (pink dots) undergoing association and diffusion (respectively) within the focus volume (blue dashed circle). The lower graph represents the fluorescence intensity fluctuations variation in time. b) typical correlation curve obtained from the intensity fluctuations.*

### *FCS set-up*

A house-made equipment was used for the FCS measurements (see Figure 3.10). It has been already described by our group in the literature.[140,141] Samples were excited by a diode laser at 485 nm (Becker&Hickl, BDL-485-SMC). The laser beam was first collimated by a fiber-collimator (60FC-4-A3.1-01-DI, Schäfter+Kirchhoff GmbH) before entering the microscope. Then the excitation light was directed towards the microscope objective (Olympus, UPLSAPO 60xW/1.20, water immersion) for sample excitation by a dichroic mirror (DI02-R532, AHF analysentechnik). Fluorescence emission was collected by the same objective and focused on a pinhole (Thorlabs, Ø=50 µm, US). Then it was filtered by a band-pass filter (Semrock, Brightline HC 525/45, US) and split into two beams by a nonpolarizing beamsplitter cube (Newport, 05BC17MB.1, US). Each beam was then focused onto an avalanche photodiode (MPD50CTC APD, Ø=50 µm, MPD, Italy).

Amyloid aggregation, reversibility and time stability studies presented in **section 5.2** were done with two TCSPC-modules (SPC 132, Becker & Hickl GmbH, Berlin, Germany) for processing and storing the detector signals. However, after this the equipment was changed and for the autofluorescence studies described in **section 5.3** a MultiHarp 150 high-throughput multichannel event timer & TCSPC unit (PicoQuant, Germany) was used instead.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

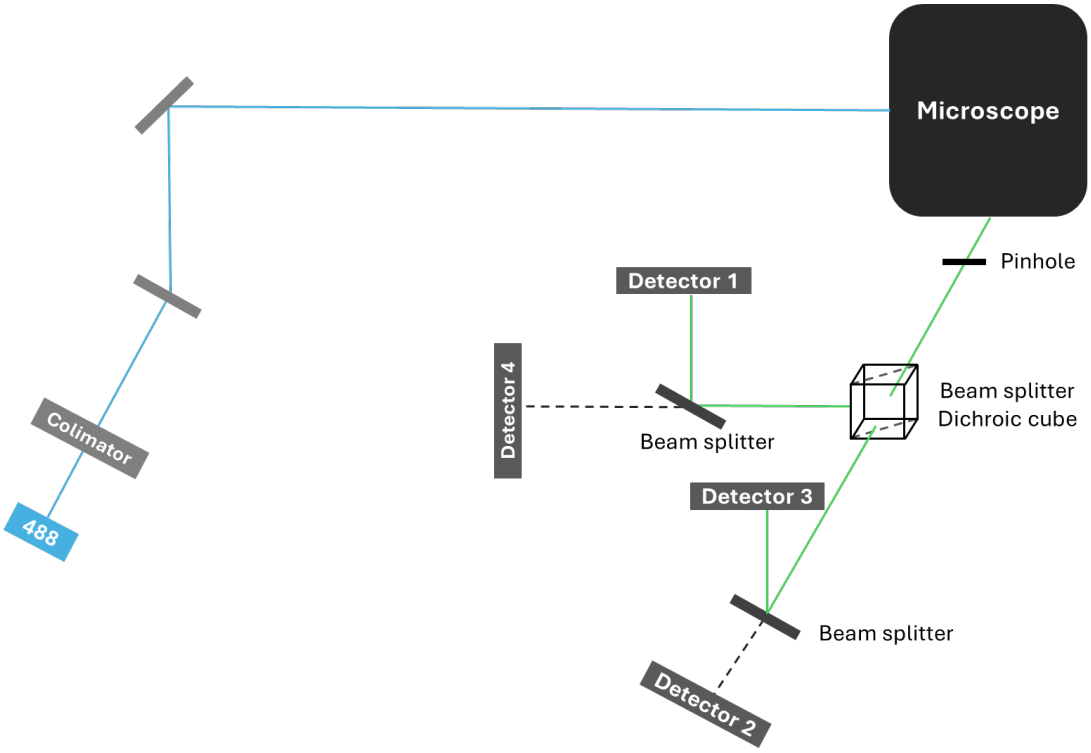



*Figure 3.10. FCS optical set-up scheme. Blue line is laser excitation path, green line is the emission path. Dotted grey line is the possible emission path after colour split if detectors 2 and 4 are activated. Grey rectangles are conventional optical mirrors.*

All measurements were carried out using 96 well plates (Whatman) which were kept at constant temperature of 21.5 ± 0.5 ºC and 22.0 ± 0.5 ºC for the results shown in **section 5.2** and **section 5.3**, respectively.

For fluorescence correlation spectroscopy, one needs to know the sample volume of each experiment as accurately as possible. For this reason, one must make a microscope alignment each day before recording any data. The alignment is carried out measuring the excitation laser power before entering the microscope, after passing through the objective and aligning the mirrors and the collimator until the maximum possible power is reached in both positions. Once the collimator and mirrors of the excitation pathway are properly aligned, the emission pathway is aligned as well. For this, a reference dye is needed. As we used a 485 nm diode laser as excitation source, our reference dye was Rhodamine 123 (R123). The emission intensity of R123 is checked in the data acquisition software for both used detectors. Detectors are aligned until the maximum intensity is achieved in both. After this, a correlation curve of the R123 sample is recorded and analysed as explained in the **Data Analysis** chapter (section 4.3). The diffusion coefficient of the R123 is known to be $D = 4.4 \times 10^{10}\ m^2 s^{-1}$ at 22 ºC.[101–103] After checking that the obtained parameters for the R123 sample are correct, the data acquisition can begin.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

It is important to note that FCS is very susceptible to photobleaching and optical saturation as it would lead to diffusion times lower that the actual values. The photobleaching phenomenon occurs when samples are irradiated with high irradiances, which provokes degradation of the fluorescent molecules during their transit through the focus. It affects the diffusional times associated with the fluorophore, because molecules which are still in focus can degrade and suddenly stop emitting, which would represent a high fluctuation in emission intensity. Optical saturation also appears at high irradiances, but it is due to the concentration of molecules transitioning into the triplet state. The transition into the triplet state also leads to a decrease in fluorescence emission intensity, affecting the resulting diffusional time. The irradiance at which optical saturation and photobleaching happens depends on the system of study. For photobleaching, the dye stability and the diffusional time of the fluorophore are important factors. For optical saturation, the key factors are the photophysical properties of the dye and the exposure time. For this reason, the laser irradiance should be optimised for each fluorescent system.

The correct irradiance to choose for a given system would be the one that allows higher emission intensities without photobleaching or optical saturation taking place. To optimise this experimental parameter, laser power series studying the variation of the fluorescence counts per second (ADC), the fluorescence intensity per molecule (cpm) and the diffusion time with laser irradiance must be performed. ADC corresponds to the fluorescence emission of the fluorophores in the sample at each moment, and it is directly recorded by the software when obtaining a correlation curve. The cpm can be calculated as a function of the ADC divided by the average number of molecules in the focus volume. All three parameters present a linear dependency with irradiance at low irradiance values. A loss in linearity usually means that either photobleaching, optical saturation or both are taking place. Also, the value at which linearity is lost will depend on when photobleaching and optical saturation start. It can change with the focus volume bur also when the dye molecules bind to other molecules. Thus, the optimisation must be carried out at a fixed focus volume and for both free and complexed fluorophore. The highest irradiance at which linearity still applies should be used for the studied system.

**Data acquisition conditions**

For FCS experiments presented in **section 5.2**, for each sample, several correlation curves, each consisting of 5 million photons, were recorded to ensure high-quality data. The FCS focus volume was of $V = 0.52 \pm 0.04\ \mu m^3$ with a radial width of $\omega_{xy} = 0.27 \pm 0.01\ \mu m$. From power series of monomeric and aggregated samples of the peptide (Figure 6.1) an excitation power of 30μW, corresponding to an irradiance of 24 kW $cm^{-2}$ was selected to prevent photobleaching.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

In **section 5.3** the data acquisition software was modified allowing for longer acquisition times. Therefore, for each sample, one correlation curve consisting of 30 million photons was recorded to ensure high-quality data. Binary data was recorded with MultiHarp 150 software from PicoQuant and correlation curves were extracted from said data using a home-built routine that runs under LABVIEW (National Instruments). After calibration, the focus volume for these studies was of $V$=0.54 ± 0.04 μm$^3$ with a radial width of $\omega_{xy}$ = 0.271 ± 0.006 μm. Excitation power was the same as before, 30 μW, which corresponds to an irradiance of around 26 kW cm$^{-2}$.

### 3.1.2.2. Super-Resolution Fluorescence Microscopy

Fluorescence microscopy was first introduced in the early 20$^{th}$ century, and it rapidly gained the attention of scientists because it allowed visualization of fluorescent molecules with high specificity and sensitivity. Its impact was such that by the mid-20$^{th}$ century it was already a widely used technique in biological research. Although it brought great advances with its appearance, resolution could not go lower than the diffraction limit (200-300 nm). Therefore, shortly after, higher resolution microscopy techniques like confocal microscopy were developed looking to improve the resolution, especially in biological samples. These new techniques allowed to bring the resolution down to 100-150 nm. Lowering the detection limit allowed researchers the visualization of subcellular structures as well as capability for three-dimensional imaging. By the end of the 20$^{th}$ century, the first super-resolution microscopy (SRM) was introduced (STED - Stimulated Emission Depletion). Since then, numerous SRM techniques have been developed including SIM (Structured Illumination Microscopy), PALM (Photoactivated Localisation Microscopy), and STORM (Stochastic Optical Resolution Microscopy). These techniques allowed researchers to achieve resolutions below 100 nm. By surpassing the diffraction limit, scientists could now study molecular processes with high resolution at the nanometre-scale, enabling detailed imaging of molecular processes within cells. This was such a revolutionary advance in fluorescence microscopy that the main contributors (Stefan W. Hell, William E. Moerner, and Eric Betzig) were awarded the Nobel Prize in Chemistry in 2014. Although all these techniques are still very revolutionary, their application was strained because they require the use of highly specialised and complex microscopes which are as well quite expensive. For this reason, single-molecule localisation microscopy (SMLM) techniques and fluorescence-fluctuation based microscopy technologies (FF-SRM) were introduced as alternative techniques.[142–145]

***Fluorescence-fluctuation SRM: Super-resolution radial fluctuations theory***

Super-resolution techniques looking to optimise image resolution have been increasing for the last couple of years. However, as exposed above, most of them are camera-based technologies that require expensive additional optical components. Another approach that had widespread were the SMLM techniques. These techniques are based on the

USC UNIVERSIDADE DE SANTIAGO

localisation of blinking fluorophores. They can construct super-resolution images by collecting large sequences of frames containing small populations of non-overlapping transiently emitting fluorophores. Then, by analytical detection and localization of the fluorophores, an accurate map of the positions can be obtained. However, these approaches can only be applied when the distance between emitting fluorophores is no lower than 1.5 µm. This means that to apply this approach, one needs low density populations of fluorophores emitting in each frame. This can be tough to obtain when working with dynamic processes in cells that often present high heterogeneity in fluorophore densities. For these reasons, numerous FF-SRM methodologies have been developed (SOFI, 3B, deconSTORM, SRRF and others).[142,145,146]

At its core, these new algorithms operate on a fundamentally different principle compared to traditional super-resolution methods. They take a different route by using the intrinsic properties of image sequences acquired from standard widefield microscopes or total internal reflection setups. Their main feature lies in their ability to extract high-resolution information directly from raw image data without the requirement for individual fluorophore detection or localization. This novel approach eliminates the need for costly additional hardware, making super-resolution imaging more accessible to researchers across diverse fields.

SOFI (Super-resolution Optical Fluctuation Imaging) was the first algorithm of this type to be developed and was first introduced in 2009 and uses higher-order statistical analysis of temporal fluorophore intensity fluctuations. Since then, many updates have come out. 3B, on the other hand, uses a completely different approach consisting of Bayesian analysis to estimate the position of the fluorophore molecules. deconSTORM reconstructs images by averaging extensively deconvolved images of sub-populations of fluorophores. However, all three of these approaches still show significant drawbacks as they require big amounts of frames and present a medium to high dependency on fluorophore blinking.

Super-resolution radial fluctuations (SRRF) was first introduced by Gustafsson et al. in 2016[146] and follows similar principles as the other mentioned algorithms. However, contrary to previous algorithms, it operates without a detection step on a single analytical framework, allowing for high-fidelity reconstructions for a wide range of different data sets. This gives SRRF great adaptability to different imaging conditions and sample types, including images containing overlapping fluorophores.

Instead of pinpointing the precise locations of fluorophores, SRRF analyses the temporal and spatial fluctuations present in the image sequence, generating super-resolved images with only 100 frames with a maximum resolution of 60 nm. This transformation preserves the temporal information within the data while enhancing spatial resolution,

USC
DE SANTIAGO
DE COMPOSTELA

setting SRRF apart from traditional single-molecule localization microscopy (SMLM) approaches. To do this, SRRF first conducts a spatial analysis and then a temporal analysis of the resulting data from the first analysis (Figure 3.11).

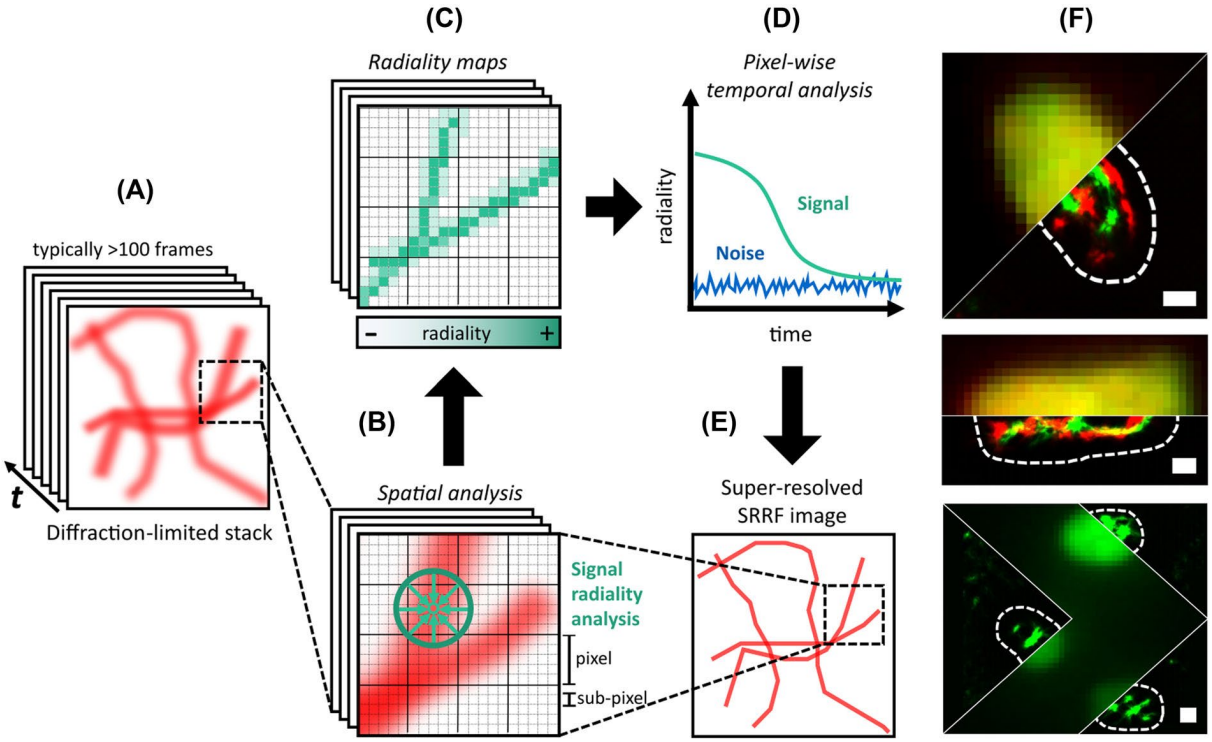



*Figure 3.11. Schematic representation of super-resolution radial fluctuation algorithm's analysis procedure. A) Raw data. B) Obtained subpixels after interpolation of the raw data. C) Generated radiality map from measuring local radial symmetries. D) Radiality variation in time. E) Final reconstructed image. F) Super-resolved images of Escherichia coli imaged on HILO microscopy. Reconstructions were generated with SRRF: Ring radius = 0.5, magnification = 10, axes in ring = 8. Scale bars: 350 nm .*[145]

First, the spatial analysis starts by applying a bicubic interpolation to the raw image that effectively magnifies each pixel into subpixels (Figure 3.11B). Then, to each of the generated subpixels a non-binary value is assigned based on the probability of it containing a fluorophore. This assignation is done by measuring the local radial symmetries in the image, which generates a radiality map like the one shown in Figure 3.11C. Radiality measures the degree of local gradient convergence across the entire frame on a sub-pixel basis. If a subpixel is located close to the centre of a fluorophore molecule, its degree of convergence will be high. While, if a subpixel is located on the image background region, the surrounding intensity gradients vectors will present low convergence values. This way, SRRF is capable of discerning two overlapping fluorescent molecules with only one radiality map. However, on images presenting significant noise it can be misinterpreted as a fluorophore signal and generate artefacts. To prevent these artefacts from appearing a temporal analysis must be performed (Figure 3.11D).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

The temporal analysis is affected by the drift that samples can present when acquiring the images. For this reason, before implementing this analysis, the drift correction is applied to the data. After this, the radiality fluctuations over time are analysed with higher-order temporal statical methods like TRM (temporal radiality maximum), TRA (temporal radiality average), TRAC (temporal radiality auto cumulant) or TRPPM (temporal radiality pairwise product mean). Temporal analysis plays a key role in enhancing the resolution and fidelity of SRRF reconstructions. Because of this, the resolution of the final image will depend on the degree of fluctuations exhibited in the raw data.[142,145,146]

Despite its strengths, SRRF is not without limitations. In noisy images or high-density fluorophore samples, the algorithm may generate artifacts or overly narrow structures, leading to the need of careful parameter optimization and error correction strategies. To address these challenges, other algorithms continue to be developed up to this day.[145] An example of this is the enhanced version of SRRF, known as eSRRF. This new version incorporates Fourier transform interpolation, radial gradient convergence (RGC), and sensitivity control, offering improved performance and user-friendliness.[147]

The implementation of the SRRF algorithm, as well as many of the other existing algorithms, is a freely available open-source plugin compatible with popular image analysis software like ImageJ or Fiji. In the case of SRRF, it comes as the NanoJ-SRRF and NanoJ-eSRRF software packages.[146,147]

***Microscope set-up***

Super-resolution fluorescence microscopy images were obtained with a Nikon, Eclipse TE2000-E wide-field inverted microscope. The microscope set up was as shown in Figure 3.12. For sample irradiation a Cobolt, 06-MLD 488 nm CW-Diode laser was used. The laser first passes through a beam expander (Thorlabs GBE15-A) to increase the laser beam diameter. Then the beam enters the microscope and passes through a Semrock BrightLine, Di03-R405/488/561/635-t1-25x36 dichroic mirror that directs the excitation light to a Nikon, Apo TIRF 100xW/ 1.49, oil immersion objective. After passing through the objective the sample is irradiated and the emitting light goes through the same objective and the dichroic mirror to be directed towards the detector. As a detector an Andor, EMCCD Monochrome camera (IXON-L-897 512 x 512) was used. Before reaching the detector, the emission light is filtered with a HQ512/35M emission filter.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

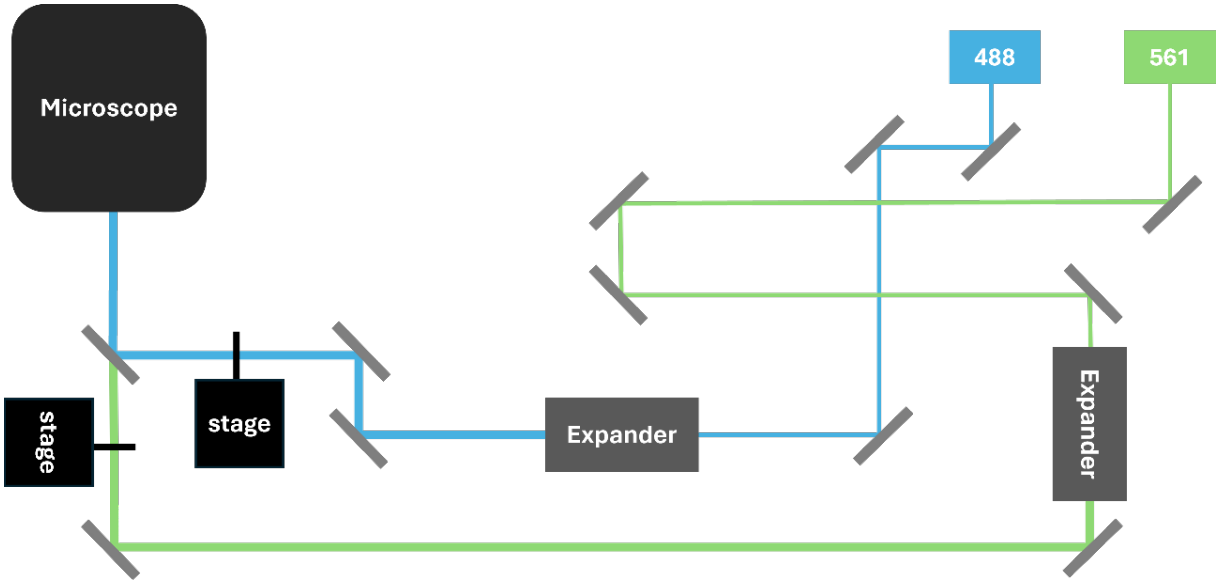



*Figure 3.12. Wide-field microscope optical Set-up scheme. Grey rectangles are optical mirrors, blue line is the 488 nm laser excitation pathway and green line is the 561 nm excitation pathway.*

All imaged were acquired at room temperature and the room was kept at a constant temperature of 18ºC. Images were acquired with Andor Solis software at 0.2 second exposure time and one accumulation per image. Raw image sequences acquired in the wide-field microscope were then transformed into super-resolution images using the SRRF algorithm. The above mentioned NanoJ-SRRFF software package was installed as a plugin in ImageJ.

***Slide preparation for imaging experiments***

For super-resolution imaging acquisition, sample preparation does not consist only of solution preparation but also on microscope slides preparation. For our studies, all experiments were carried out with the same slide cleaning and sample channel preparation protocol.

First, both slides and coverslips were cleaned by rinsing them with acetone, water and ethanol in that order. Immediately after rinsing, they were dried using high-pressure air and stored on a clean surface. Next, the sample channels were constructed using double-sided tape. Strips of tape were cut to the desired width and length to create three sample channels per slide. The tape was carefully laid on top of the slide, as shown in Figure 3.13, and air bubbles were removed by gently applying pressure. Once all the tape was in place, a previously cleaned coverslip was placed on top. To ensure the coverslip adhered to the tape, pressure was applied to each tape section until no air bubbles were left.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

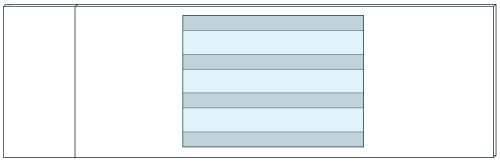


*Figure 3.13. Microscope slide.*

To load the samples into de sample channels, we first cleaned each channel several times with PBS. After passing PBS through the channels at least two times, samples were loaded into their respective channels. To ensure that the solutions stay in place, right after all samples were loaded glue was added to all sides of the coverslip sealing the samples.

## 3.2. Materials

All the studies that appear throughout this thesis were performed under buffering conditions to ensure the desired pH were obtained for all samples. Table 3.2 shows the different reagents used for buffer preparation alone with the manufacturer that they were purchased from and the purity.

*Table 3.2. Used reagents for buffer solutions preparations, with commercial specifications (brand and purity).*

| Reagent | Brand | Purity / % |
|---|---|---|
| Sodium phosphate | Sigma Aldrich | ≥99 |
| Potassium phosphate | Sigma Aldrich | ≥99 |
| Sodium perchloride | Panreac | ACS grade |
| Potassium perchloride | Acros Organics | ACS grade |
| Sodium chloride | Sigma Aldrich | ≥99 |
| Potassium chloride | Merck | ≥98 |
| Trizma hydrochloride | Sigma Aldrich | ≥99 |
| Chlorohydric acid 37% | Merck | ACS grade |
| Sodium hydroxide | Sigma | ≥98 |
| Sodium carbonate | Sigma Aldrich | ACS grade |
| Sodium acetate | Sigma Aldrich | ACS grade |
| Acetate | AnalaR Normapur | ACS grade |
| PBS tablets | Sigma Aldrich | - |

Additionally, to match the ionic strength of the buffers to physiological conditions (0.18 M), in buffers that originally do not require it like Tris, carbonate and acetate, NaCl was added to obtain an ionic strength of 0.18 M. If necessary, pH adjustments were made by HCl or NaOH addition.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

- **G-quadruplexes studies**

DNA and RNA sequences were purchased from Integrated DNA Technologies and Biomers, and dissolved in water, purified using a Millipore Milli-Q system, to obtain stock solutions at concentrations of the order of 100 μM. All studied sequences and its given abbreviations are shown in Table 3.3. Sodium and potassium phosphate salts were used to prepare the Phosphate buffer (20 mM, pH 7.0) and Trizma hydroclorhide for the Tris buffer (20 mM, pH 7.5). Thioflavin T (ThT, Sigma) was used without further purification both in fluorescence and electrophoresis experiments. For fluorescence measurements, sodium and potassium perchlorates were used instead of chlorides to avoid quenching.

*Table 3.3. Abbreviation and descriptions of the used nucleic acid sequences.*[27]

| Abbreviation | Type | Sequence (5' to 3') |
|---|---|---|
| G5AG7 | DNA | GGGGGAGGGGGGG |
| GCG5AG7 | DNA | GCGGGGGAGGGGGGG |
| rGCG5AG5 | RNA | GCGGGGGAGGGGGGG |
| G6AG7 | DNA | GGGGGGAGGGGGGG |
| GCG6AG7 | DNA | GCGGGGGGAGGGGGGG |
| rGCG6AG7 | RNA | GCGGGGGGAGGGGGGG |
| G6AG8 | DNA | GGGGGGAGGGGGGGG |
| GCG6AG8 | DNA | GCGGGGGGAGGGGGGGG |
| rGCG6AG8 | RNA | GCGGGGGGAGGGGGGGG |
| GCG6AG7 MUT | DNA | GCGGAGGGAGGAGGGG |
| hGCG6AG7 | DNA | GCGGGGGGAGGGGGGGAAGACAGCACTCAGCAAGCA |
| hGCG6AG8 | DNA | GCGGGGGGAGGGGGGGGAAGACAGCACTCAGCAAGCA |
| DNAh | DNA | TGCTTGCTGAGTGCTGTCAAGCGTCTCGTCCTGATGA |
| RNAh | RNA | UCAUCAGGACGAGACGCAAGGGAA |

- **Amyloid aggregation studies**

For amyloid aggregation studies both labelled (Aβ40* and Aβ42*) und unlabelled (Aβ40 and Aβ42) commercial amyloid peptides were used. Unlabelled amyloids were purchased from Anaspec Inc. and Genscript USA Inc. with a purity of ≥95 %. Labelled

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

amyloids Hilyte Fluor 488- β-amyloid(1-40) and Hilyte Fluor 488- β-amyloid(1-42) were obtained from Anaspec Inc. with a ≥95 % purity as well.

Studies described in **section 5.2** and **5.3** utilized the HFIP protocol for reaggregation of the amyloid (see 3.3.2.1 section for further details). To prevent contamination, aliquoted 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) ≥99% from Sigma Aldrich was used. Aqueous solutions under physiological conditions were prepared using PBS (7.5, ionic strength 0.18 M, phosphate concentration 13 mM and salt concentration of 142 mM sodium and 2.7 mM potassium) as the buffer. Tris buffer (pH 7.5, ionic strength 0.18 M) was prepared with Trizma hydrochloride and NaCl. For pH adjustment HCl (Merck, 37% ACS grade) or NaOH (Sigma Aldrich, ≥98 % purity) were used.

For the experiments presented in **section 5.4** amyloids were stored at -20ºC after stock preparation following the $NH_4OH$ protocol explained in the 3.3.2.2 section. Ammonium hydroxide solution (2 % NH3 in water, ≥ 99 %) was purchased from Sigma Aldrich. Samples were prepared in PBS pH 7.5, which was prepared using PBS tablets from Sigma Aldrich (pH 7.2-7.6). For the cation, spermine (≥ 97 % GC, Merck), spermidine (≥99 %, Alfa Aesar), and magnesium chloride (anhydrous, ≥ 98 %, Sigma Aldrich) were used.

- **Human serum albumin detection studies**

All studies were done using solutions of different concentrations of CBTOH (Sigma Aldrich, ≥ 97 %), which was first dissolved in ethanol (Scharlau, HPLC purity). Samples were prepared using PBS (pH 7.4-7.5), acetate buffer (pH 5.0) or carbonate buffer (pH 10.0) depending on the experiment. Globulin-free bovine serum albumin (BSA) was obtained from Sigma Aldrich with a purity of ≥ 98 %, and high purity (≥ 95 %, Fraction V) human serum albumin (HSA) was purchased from Merck.

# 3.3. Sample preparation

Sample preparation depends highly on which technique is going to be used to carry out the experiment, as the required concentrations and volumes are not the same. Additionally, protein and DNA or RNA samples require different and specific treatments. Thus, this chapter will explain very briefly the followed methodologies for sample preparation.

## 3.3.1. Sample preparation for G-Quadruplexes studies

All DNA and RNA sequences were commercial samples that were stored as obtained from the manufacturer at -20ºC in its lyophilised state until needed for preparation. For DNA stock preparation, samples were diluted using Phosphate buffer (20 mM, pH 7.0) to obtain concentrations of around 200 µM. Then, each stock concentration was calculated

USC DE SANTIAGO DE COMPOSTELA

using its absorption spectra. Stocks were then stored again at -20ºC. RNA samples were prepared and used in the same day, without storing them for later studied because of their lower stability.

First, sample solutions were prepared in the phosphate (20 mM, pH 7.0) or Tris (20 mM, pH7.5) buffer with a concentration about 6 μM of the nucleic acid and 0-200 mM of the desired cation. The samples were then heated in an Eppendorf ThermoMixer (Hamburg, Germany) at 95 ºC for 15 min and then cooled with ice for another 15 min to achieve a fast-cooling process and avoid the formation of intermolecular structures. It must be noted that, when using the phosphate buffer, the cation concentration was already 30 mM without the further addition of salt.
Although samples in which the formation of a hybrid G-quadruplexes between DNA and RNA sequences was required were prepared in the same way, it must be noted that to obtain a final concentration of 6 μM a mixture of three sequences was added to the solution. By adding this mixture and promoting hybridization at 95 ºC, two DNA:RNA G4 hybrids are formed between the non-template DNA (DNAh, light blue strand in Figure 1.3), the RNA transcript (RNAh, pink strand in Figure 1.3) and the G4 sequence DNA (hGCG6AG7 or hGCG6AG8, light blue strand in Figure 1.3):

- DNA:RNA-hGCG6AG7: DNAh hybridised with hGCG6AG7 and RNAh
- DNA:RNA-hGCG6AG8: DNAh hybridised with hGCG6AG8 and RNAh

For the fluorescence experiments samples were prepared following the explained protocol with same nucleic acid concentration and then ThT was added into the cuvette to achieve a ThT concentration of 12 μM.

In electrophoresis experiments samples were prepared the same way, but for a final volume of 30 μL and 300 pmol of DNA or RNA. Before loading the samples, 10 μL of 40% saccharose were added to them. These experiments were carried out with 14 % polyacrylamide gels on 0.5X TB buffer (0.05M Tris borate). Gels were run for 1-2 hours at 80-90 V with 0.5X TB buffer and developed using ThT, as it was previously reported to be a selectively fluorescence probe that binds to G4 DNA.[128,148–150] Gel evolution was followed by addition of bromophenol blue to one of the wells of the gel.

## 3.3.2. Amyloid-β monomeric stock preparation

When working with amyloid peptides, pre-storage treatment is needed to prevent undesired amyloid aggregation. Several protocols for amyloid treatment can be found on the literature. We describe below the methods used in this thesis.

### 3.3.2.1. Amyloid sample preparation for storage with HFIP

USC UNIVERSIDADE DE SANTIAGO

The protocol for monomeric amyloid preparation with HFIP has already been described in the literature by Stine and colleagues.[151,152] The addition of this solvent to the commercial

peptide will disaggregate any potentially formed oligomer, allowing us to make sure that our stored amyloid is indeed all in monomeric form.

- Amyloid peptide is dissolved directly on the commercial tube by adding HFIP for a final concentration of 1 mg/mL.
- The sample is incubated for 1 hour at room temperature with occasional mixing – done by vortex. Amyloid solution should be a clear and uncoloured solution.
- The sample is mixed for 20 minutes on a vortex.
- The solution is separated into aliquots of the desired volume that will allow later to obtain the correct amyloid concentration for each experiment. Before pipetting the solution from the vial to an Eppendorf tube, as HFIP is highly volatile, chloroform is pipetted several times.
- HFIP is evaporated using nitrogen.
- All aliquots are stored in a desiccator for at least 4 hours under vacuum conditions to make sure no HFIP remains on the samples.
- The samples are stored at -20ºC.

Stock amyloid samples for studies present in **section 5.2** and **5.3**were prepared using this procotocol.

#### 3.3.2.2. Amyloid sample preparation for storage with $NH_4OH$

The protocol with $NH_4OH$ has been described as well[153], and it is advised by the manufacturer for amyloid solution preparation. This protocol is easier and faster than the above described. However, amyloid oligomers may form as it is stored in aqueous solution.

- The amyloid peptide is dissolved with 80 μL of 1% $NH_4OH$ for each mg of peptide.
- PBS pH 7.4 buffer is added for a final amyloid concentration of 1mg/mL. Higher concentrations of amyloid will lead to aggregates formation. 1% $NH_4OH$ without buffer cannot be used for long-term storage.
- Mix the solution for several minutes.
- Separate the solution into desired volume aliquots.
- Store aliquots at -20ºC. Refreezing will lead to aggregate formation.

Stock amyloid samples for studies present in **section 5.4** were prepared using this procotocol.

### 3.3.3. Amyloid samples for FCS studies

All samples were prepared for lyophilised amyloid aliquots that were previously prepared following the HFIP protocol reported by Stine et al[151,152], which is described in the **Experimental methods** chapter (3.3.2.1 section).

In **section 5.2** all samples were prepared maintaining the concentration of labelled amyloid around 20 nM while changing the concentration of the unlabelled peptide. Different procedures were applied depending on the specific objectives of each experiment.

- **Independent samples for aggregation studies**

These samples were prepared individually to investigate the potential dependence of the aggregation process on the incubation time and to obtain independent measurements for each concentration, as reported in previous studies.[68,154]

First, an aliquot of dry monomeric labelled amyloid was dissolved in PBS to achieve the desired concentration of about 20 nM. After verifying the absence of aggregates by measuring a FCS curve, this solution was used to directly dissolve an aliquot of the unlabelled peptide. The final solution was then mixed thoroughly at room temperature for 5 minutes and used for FCS measurements. The concentration of unlabelled amyloid was varied from 0 to 32μM.

- **Samples for reversibility studies**

Reversibility of the aggregation process was studied by dilution of concentrated amyloid samples covering peptide concentrations ranging from 150μM to 0.5μM, divided into three series with nominal concentration ranges of 0.5-25μM, 5-100μM, and 50-150 μM.

For each series, the sample of the highest concentration was prepared as described above for the individual samples. Diluted samples were then prepared directly in the microplate well during the FCS experiment using an extraction-dilution method. A small volume was extracted from the total sample volume of 100 μL and replaced with the same volume of a solution containing the same concentration of labelled amyloid but no unlabelled peptide. This method ensured the concentration of labelled amyloid and the sample volume remained constant while the unlabelled amyloid was diluted. The sample was mixed in the well by repeated pipetting. FCS curves were measured for the solutions obtained at each step within few minutes after dilution. These extraction-dilution-measurement cycles were repeated to cover the desired concentration range.

- **Samples for time stability studies**

For the stability studies, the independent samples from the aggregation studies were stored at room temperature. For short incubation times, they were kept in the well plate covered with Parafilm sealing film and a silicone sealing mat to avoid evaporation and contamination. FCS curves were measured every 10 to 20

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

minutes for 2 hours. For longer incubation times the same samples were stored in microcentrifuge tubes for the desired incubation period and then transferred to the well plate for FCS measurements.

In **section 5.3** samples were used both for steady-state fluorescence and for FCS measurements. First, samples were prepared for its use in steady-state emission with two different methods: dilution-extraction and independent sample preparation. For independent sample preparation, an unlabelled aliquot was dissolved in a PBS or Tris buffering solution, obtaining the desired Aβ40 concentration. In the dilution-extraction method, a first sample was prepared with the independent sample preparation method at the highest Aβ40 concentration, then a certain volume of PBS or Tris was added to the Aβ40 solution, and the new solution was homogenized. After the solution was homogenized, a volume of the solution was removed (the same volume that was previously added), keeping sample volume constant. This step was repeated systematically until the desired range of concentrations was covered. This method was used because it allows us to use smaller Aβ40 amounts, but has the downside of an error addition as the iterations go on.[155] The needed steady-state measurements were performed for each sample and then, a small fraction of the sample was transferred to a new Eppendorf were labelled Aβ40 was added to a final nominal concentration of labelled amyloid of around 20 nM. After adding the labelled amyloid, samples were used for FCS data acquisition. Both steady-state fluorescence and FCS measurements were taken right after the solution preparation and 5-10 minutes of mixing.

### 3.3.4. Amyloid samples for microscopy studies

Amyloid stocks for storage were prepared using the $NH_4OH$ protocol, as explained in this chapter section 3.3.2.2. Then, samples were prepared as follows: First, an amyloid solution containing 10 % of labelled Aβ42 and 90 % of unlabelled Aβ42 were prepared to a total concentration of 10-12.5 μM. This solution was incubated at 37ºC from periods ranging from 24 hours up to 1 week for fibril formation. After fibrils were form samples with cations concentrations ranging from 0 to 1 mM were prepared by adding the desired cation (spermine, spermidine or $Mg^{2+}$) leading to a final concentration of Aβ42 of 10 μM. Cation concentration was adjusted to follow a cation:amyloid ratio range of 0, 0.1, 1, 10 and 100. These samples were incubated again at 37 ºC for 24 hours up to 1 week.

### 3.3.5. Sample preparation for human serum albumin detection studies

Samples were prepared in buffered solutions of different pH depending on the desired experiment. For studies under physiological conditions samples were prepared with PBS

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

pH 7.4 - 7.5, while for association studies at pH 5.0 and 10.1 acetate buffer and carbonate buffer were used respectively. All buffers were first prepared for a higher concentration than needed (0.02 or 0.03 M) and dissolved to the desired final concentration of 0.01 M when preparing the samples. A NaCl 1 M was prepared in MilliQ water and used to bring the ionic strength of all samples to a physiological level, resulting on a final NaCl concentration of 0.16 M and an ionic strength of 0.18 M in all samples. A stock sample of CBTOH at 2.5 mM was prepared in ethanol, as CBTOH does not present great solubility in water and needs to be first dissolved in organic solvents. Then, the needed volume of this stock solution was transferred to a Falcon tube, and ethanol was evaporated under a nitrogen stream until CBTOH was dry. Once all ethanol was evaporated, buffer (PBS, acetate or carbonate) was added to dilute the dye to a final concentration ranging from 2 to 25 μM depending on the experiment as detailed in Table 3.4.

*Table 3.4. CBTOH and buffer conditions used for each technique.*

| Technique | CBTOH / μM | Buffer |
|---|---|---|
| UV-Vis absorption | 25 | PBS, acetate and carbonate |
| Fluorescence emission | 2 | PBS |
| Fluorescence anisotropy | 2 | PBS |
| Fluorescence lifetimes | 5 | PBS |

For pH titrations, samples were prepared by maintaining the final volume, CBTOH and buffer concentration, and ionic strength constant in all samples. The procedure used involved preparing a new CBTOH stock in phosphate buffer, which was then added for a final concentration of 25 μM, then NaCl was added to match the ionic strength and the pH was modified by adding HCl or NaOH. MilliQ water was added as needed to reach a certain total volume.

For association studies, two solutions of CBTOH were prepared, one without protein and the other with a high concentration of protein. To prepare the first solution, CBTOH was diluted with the needed buffer (PBS, acetate or carbonate) for the needed final concentration of dye depending on the experiment (Table 3.4). The second solution consisted of CBTOH with added protein and was prepared to match the CBTOH concentration of the first solution, with a protein concentration of around 250-300 μM. These two solutions were there systematically mixed following a dilution-extraction methodology to prepare samples with a constant volume and concentration of CBTOH and varying concentrations of protein. The final range of protein concentration studied went from 0 μM to 250-300 μM.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# 4. Data Analysis

## 4.1. Absorption and Emission Spectra Analysis

### 4.1.1. UV-vis absorption spectra

To correctly measure the UV_vis absorption spectra, the baseline must be previously measured, and the data acquisition software directly corrects the absorbance differences. Nevertheless, when doing measurements for long periods of time, the signal can show some instability. This can be easily detected in the no absorptive range. This means that when the sample should present an absorbance equal to zero, the measured absorption might be slightly higher or lower than it should be. This can happen because the measured baseline was not correctly stabilized. Usually, the absorption change happens to the whole spectrum and not only to the low absorption range. Therefore, additional spectral correction must be applied. First, the region of the spectrum where the absorbance should not change must be selected. The wider this region is, the better. Then, an average of the absorbance values of that region is calculated and this average is subtracted from the spectrum values.

Absorption spectra were also corrected by the concentration of the dye. This is done because when doing titrations where the dye concentration must be kept constant, a slight change in the concentration might lead to errors in the obtained absorption. Thus, because all samples are prepared weighting the added volumes, the real concentrations of the dye on each sample can be calculated. The concentration correction is done according to Eq (4.1) where $C_i$ is the dye concentration of the sample whose spectrum is going to be corrected and C is the average of the dye concentrations of all the samples measured in the same experiment.

$$A_{corr} = \frac{AC_i}{\overline{C}} \tag{4.1}$$

### 4.1.2. Circular Dichroism spectra

As in the case of UV-vis absorption spectra, circular dichroism spectra also need to be corrected by the signal the solvent might have. In this case, there is no automatic baseline correction done by the data acquisition program, so we manually subtract the circular dichroism spectrum of the blank sample (a sample were just the used solvent is measured) from all our sample's CD spectra. After this, spectra are corrected by the baseline using the same procedure as explained for absorption spectra. Then, the resulting corrected spectra are divided by the nucleic acid concentration to obtain the

USC DE COMPOSTELA

molar ellipticity (θ) in deg $cm^2$ $dmol^{-1}$ as shown in Eq. (4.2), where $\theta_{exp}$ is the corrected experimental ellipticity in deg, l is the pathway in cm and C is the concentration in mol $dm^3$. As the obtained spectra presented a lot of noise, a moderate smoothing was applied to reduce said noise. Because no further smoothing was applied to prevent any spectral distortion, CD spectra still show higher noise than usual.

$$\theta = \frac{100\theta_{exp}}{Cl} \tag{4.2}$$

## 4.1.3. Steady-state emission and excitation spectra

Emission and excitation spectra are highly influenced by the characteristics of the equipment. Usually, light sources are not perfectly stabile in time, meaning they present ups and downs in its intensity which inevitable affects the obtained fluorescence intensity. Additionally, not all monochromators have the same spectral properties, which can also lead to distortions in the spectra. The spectral sensitivity of the detectors is also a key factor in determining the fluorescence intensity, because their sensitivity depends on wavelength. All these factors can lead to distortions in the spectra as well as lower or higher fluorescence intensity than the actual intensity and, therefore, need to be corrected before data analysis can be performed. It must be noted that these corrections can be done automatically by the software of the equipment in almost all commercially available spectrofluorometers. However, one must pay attention to whether or not the obtained data are already corrected.

- **Lamp intensity fluctuations**

Continuous light sources typically used for steady-state fluorescence emission and excitation spectra acquisition, though being constantly exciting the sample, show instabilities or intensity drifts over time. A change in the light source intensity will also translate in a change in fluorescence intensity. However, this can be corrected easily if the light source profile is also measured alongside the desired emission or excitation spectrum. If the lamp intensity is also measured at the same time, then its intensity fluctuations can be corrected using Eq (4.3).

$$F_{C2}(\lambda) = \frac{F_{C1}(\lambda)F_{ref}^{m}(\lambda)}{F_{ref}(\lambda)} \tag{4.3}$$

Corrected fluorescence intensity values of the sample ( $F_{C2}(\lambda)$ ) will be obtained by normalizing the measured lamp intensity ( $F_{ref}(\lambda)$ ) with the average intensity of all measured lamp spectra ( $F_{ref}^{m}(\lambda)$ ).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

- **Spectral sensitivity**

As mentioned above, monochromators, detectors, and other optical elements do not show the same sensitivity at all wavelengths. This means that, at wavelengths were the sensitivity drastically decreases, the fluorescence intensity will also be significantly lower than the actual intensity of the sample. This effect depends on the specific elements used and can be corrected by obtaining its sensitivity spectrum. Typically, the correction spectrum is already measured by the manufacturer for each specific device and supplied within the software, so this correction can be automatically done by the software.

$$F'(\lambda) = F(\lambda) \times C(\lambda) \tag{4.4}$$

For this correction the experimental spectra ( $F(\lambda)$ ) has to be multiplied by the wavelength-dependent device sensitivity factor ( $C(\lambda)$ ), as shown in Eq (4.4).

## 4.1.4. Fluorescence lifetime decay curves

As explained in the **Experimental methods** chapter, time-resolved fluorescence emission measurements result in fluorescence emission decay curves. Fluorescence intensity in time-resolved fluorescence curves is a temporal distribution given by the convolution of the excitation pulse L(t) and the time-dependent fluorescence response D(t) obtained after sample excitation with a short laser pulse (Eq (4.5)). As samples are excited with a short light pulse, the laser profile represents the evolution of the intensity of the excitation source with time. Therefore, the number of molecules that can transition into the excited state will be proportional to the intensity of the excitation light and the final detected intensity of the samples will be the sum of all the excited molecules at the different times. Equation (4.5) can be solved directly by the data acquisition program from Edinburgh Instruments using the iterative deconvolution by least squares method.

$$F(t) = \int_0^t L(t')D(t-t')dt \tag{4.5}$$

Usually, when a single fluorophore is contributing to the fluorescence emission, the fluorescence emission decay (D(t)) follows a first-order kinetics and therefore can be represented by a mono-exponential function (Eq (4.6)). When more than one fluorophore is present in solution, or a single species has more than one emitting conformer or tautomer, then these curves are the sum of various exponential functions as shown in Eq. (4.7). Fitting the decay curves with Eq (4.6) or (4.7) will allow for the determination of fluorophores lifetimes ( $\tau_i$ ) and their respective contributions to the total fluorescence intensity ($B_i$). Additionally, a third parameter known as shift is obtained. This parameter is the temporal shift between the fluorescence decay and the lamp profile.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

$$D(t) = D_0 e^{-kt} \tag{4.6}$$

$$D(t) = \sum_{i=1}^{n} B_i e^{-t/\tau_i} \tag{4.7}$$

When fitting fluorescence emission decay curves, usually, each exponential term in the fit corresponds to a distinct lifetime associated with a single emissive species. If a single exponential fits the decay curve accurately, it suggests that only one emissive species is contributing to the fluorescence emission. But, if multiple exponentials are needed for the fit, then it indicates the existence of, at least, as many species or processes as needed exponentials with different lifetimes contributing to fluorescence emission.

This fit can be done individually for each decay measurement with the data acquisition software. For this, the Marquardt-Levenberg algorithm is used, which allows to fit a given curve with up to 4 exponentials. However, this was used only as a preliminary analysis. In this thesis, fluorescence decays were studied as a function of two different variables: concentration and wavelength. To do this, a Global Analysis developed by our group and implemented in Matlab was used.[156] This type of analysis will be explained later in this chapter, in the section about Principal Component Global Analysis.

### 4.1.5. Fluorescence anisotropy spectra

Fluorescence anisotropy is a powerful technique that gives information about the rotation that a given molecule undergoes during its excited-state lifetime. Anisotropy depends on the fluorophore rotational mobility in the excited state. Thus, when the molecule interacts with its surroundings, a change in its rotation speed may happen, leading to different anisotropy values. This way, we can obtain information to characterize the mobility, shape and size of a molecule and the rigidity of the medium. This information is especially useful to characterize binding equilibria, as the free and bound fluorophore may have different rotational characteristics.

If a sample is excited with polarized light, then its emission will become polarized as well to a certain degree, meaning that its electric field will oscillate in a certain direction instead of in every direction. Anisotropy (r) measures the degree of polarization exhibited by the emitted light. To determine this degree, it is necessary to measure the fluorescence intensity with different polarizations. As shown in Figure 4.1, there is one excitation polarizer and one emission polarizer, and both can select the light with a given polarization direction.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

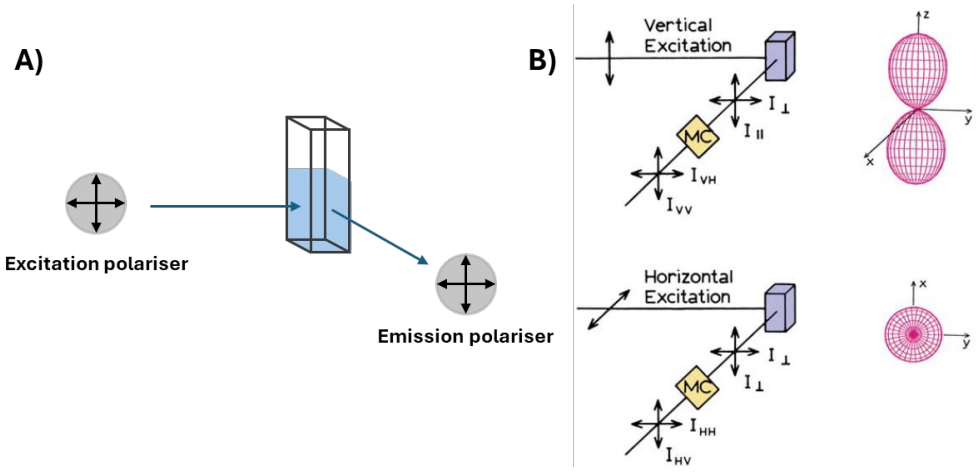



*Figure 4.1. A) Schematic representation of how anisotropy polarizers are positioned. B) Schematic representation of fluorescence anisotropy measurements on different planes (MC: monochromator); the figures on the right represent the distribution of excited-state molecules, photoselected by the polarized excitation.*[93]

Fluorescence anisotropy is defined by Eq (4.8), where $I_{\parallel}$ and $I_{\perp}$ are the intensities of the emitted light measured with the linear polarizer for emission parallel and perpendicular, respectively, to the direction of polarization of the linearly polarized excitation light. The difference between parallel and perpendicular polarization intensities is normalized by the total intensity ($I_T = I_{\parallel} + 2I_{\perp}$). Thus, anisotropy is dimensionless.

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}} \tag{4.8}$$

Because the anisotropy measurements involve different orientations of the emission monochromator. one must keep in mind that the optical components, mainly the monochromator, show polarization-dependent efficiencies. Then, to determine the anisotropy, it is necessary to measure the emission intensity of the sample with the excitation polarizer and the emission polarizer in vertical positions (VV), both polarizers in horizontal positions (HH) and both polarizers in positions perpendicular to each other (VH and HV). The G factor, defined by Eq (4.9), is an instrument-dependent parameter that measures the sensitivity ratio of the optical system for vertically and horizontally polarized light. Then, the experimental anisotropy has to be calculated using a slight correction from Eq (4.8) that takes into consideration the correction factor (Eq (4.10)).

$$G = \frac{I_{HV}}{I_{HH}} \tag{4.9}$$

$$r = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}} \tag{4.10}$$

In a freely rotating molecule, the emitted light will be depolarized over time because the molecule reorients itself randomly in solution. However, if the molecule rotation is restricted due to interactions with other molecules or structural constraints, the emitted

light will retain some degree of polarization. Samples where molecules can rotate freely (fast rotation) will present little to no anisotropy, while samples where this is restricted or slowed down will present higher anisotropy values (Figure 4.2). Usually, rigid systems where no depolarization can take place present maximum anisotropy values of $r = 0.4$. Additionally, if no spectral artefacts are present, anisotropy will be independent of fluorophore concentration.

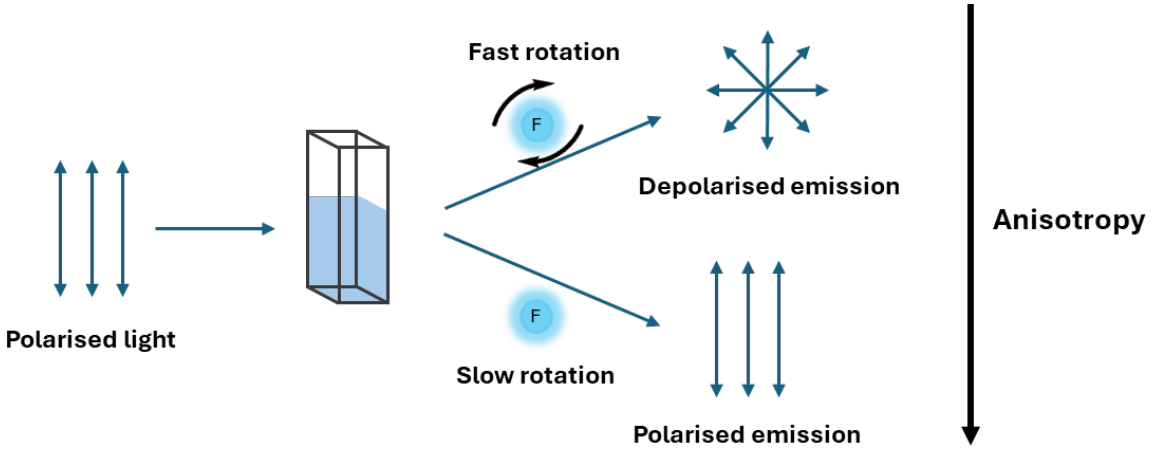



*Figure 4.2. Schematic diagram showing how polarized excitation leads to different degrees of polarized emission depending on the sample nature and how this translates into higher or lower anisotropy values.*

## 4.2. Principal Components Global Analysis

Traditionally, when analysing spectral series, two different methodologies were followed. One of then consisted of proposing first a hypothetical mechanism for the studied system and then, selecting a few wavelengths to be fitted with the proposed model. If the resulting fit was accurate, the model was accepted. The second approach was making assumptions about the shape and overlap of the spectra and then determine the individual spectra by normalization and subtraction. With this method, the concentration profiles were obtained and could be fitted with a model. The major drawback of this procedure is that to know something as basic as the number of species, highly restrictive assumptions must be made. Moreover, even after those assumptions were made, both approaches could only use a few selected wavelengths for model fitting instead of the whole range of measured wavelengths. This is because such large amounts of data are difficult to work with. These methodologies are the ones that can be carried out by using almost any of the commercially available data analysis programs.

However, obtaining spectra at different concentrations or pH can give more information when using all the measured wavelengths instead of just a few of them. To avoid using the above-exposed methodologies, a statical method known as Principal Components Analysis (PCA) can be used. And, if instead of using a few selected wavelengths, one

USC UNIVERSIDADE DE COMPOSTELA

would like to use the whole range then, Global Analysis (GA) can be carried out. If used together, then it can be called Principal Components Global Analysis (PCGA).

For a more detailed analysis of data obtained from both absorption and fluorescence conventional techniques a different PCGA programmed by our research group were used. For absorption, circular dichroism and fluorescence emission and excitation spectra a PCGA programmed in Wolfram Mathematica was used.[157] For fluorescence lifetimes, PCGA was implemented in Matlab.[156]

## 4.2.1. Principal Components Analysis

In the studies done with spectroscopic techniques one usually has spectral series of the chosen technique where an external variable changes systematically. Whether it is absorption, CD or fluorescence emission, spectra are obtained changing the concentration or the pH. Spectral data can then be organized as a data matrix (Y) where $n_s$ represents the number of different spectra, and $n_w$ represents the number of wavelengths at which the spectra were measured.

$$\underset{n_s \times n_w}{\mathbf{Y}} = \begin{pmatrix} y_{1,1} & \cdots & y_{1,n_w} \\ \cdots & \cdots & \cdots \\ y_{n_s,1} & \cdots & y_{n_s,n_w} \end{pmatrix} \tag{4.11}$$

To apply PCA to a data set, the only necessary condition is that the variation with the external variable of the spectra must lead to a linear combination of the spectra of each contributing species. The spectra of the component species are known as pure spectra and are linearly independent. Knowing this, we can then express the theorical spectra ($\hat{Y}$) as the product of the component spectra matrix (E) and the coefficients matrix (D).

$$\underset{n_s \times n_w}{\hat{\mathbf{Y}}} = \underset{n_s \times n_c}{\mathbf{D}} \cdot \underset{n_c \times n_w}{\mathbf{E}} = \begin{pmatrix} d_{1,1} & \vdots & d_{1,n_c} \\ \vdots & \vdots & \vdots \\ d_{n_s,1} & \vdots & d_{n_s,n_c} \end{pmatrix} \cdot \begin{pmatrix} e_{1,1} & \cdots & e_{1,n_w} \\ \cdots & \cdots & \cdots \\ e_{n_c,1} & \cdots & e_{n_c,n_w} \end{pmatrix} \tag{4.12}$$

Neither E nor D or the number of contributing species ($n_c$) are known beforehand, so estimates need to be found for all of them. As the $\hat{Y}$ spectra are linear combinations of $n_c$ component spectra ($e_j$), the spectra $e_j$ can be interpreted as basis vectors covering a dimension $n_c$. Now, each of the $\hat{Y}$ spectra are defined by the contributions $d_{ij}$, which are the space defined by the basis vectors. To estimate $n_c$, we look for the minimal number of linearly independent spectra (pure spectra) that allows to reproduce the variations present in spectra $\hat{Y}$.

$$\hat{\mathbf{y}}_i = d_{i1} \cdot \mathbf{e}_1 + d_{i2}\mathbf{e}_2 + \cdots + d_{in_c}\mathbf{e}_{n_c} \tag{4.13}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

To obtain the component spectra E from Eq (4.12), we need to express the experimental spectra by an apparent spectra V with coefficients Z. Z and E are related by a rotation matrix P. This way, the minimal number of needed vectors (V) are the number of observable species. The vectors will reproduce all the variation in the experimental data.

$$\underset{n_s \times n_w}{\hat{\mathbf{Y}}} = \underset{n_s \times n_c}{\tilde{\mathbf{Z}}} \cdot \underset{n_c \times n_w}{\tilde{\mathbf{V}}} = \begin{pmatrix} z_{1,1} & \vdots & z_{1,n_c} \\ \vdots & \vdots & \vdots \\ z_{n_s,1} & \vdots & z_{n_s,n_c} \end{pmatrix} \cdot \begin{pmatrix} v_{1,1} & \cdots & v_{1,n_w} \\ \cdots & \cdots & \cdots \\ v_{n_c,1} & \cdots & v_{n_c,n_w} \end{pmatrix} \tag{4.14}$$

$$\hat{\mathbf{y}}_i = z_{i1}\,\mathbf{v}_1 + z_{i2}\,\mathbf{v}_2 + \cdots + z_{i\,n_c}\,\mathbf{v}_{n_c} \tag{4.15}$$

$$\underset{n_c \times n_w}{\mathbf{E}} = \underset{n_c \times n_c}{\mathbf{P}} \cdot \underset{n_c \times n_w}{\tilde{\mathbf{V}}} \;;\; \underset{n_s \times n_c}{\tilde{\mathbf{Z}}} = \underset{n_s \times n_c}{\mathbf{D}} \cdot \underset{n_c \times n_c}{\mathbf{P}} \tag{4.16}$$

After this, we need to determine the number of components by analizing the results of the analysis. To do this, we use a combination of statistical and visual criteria of the obtained vectors (V) and the resulting residuals of the analysis. This is done by looking at the eigenvector diagram, the logarithmic eigenvalue diagram, the mean residual error diagram, the resulting value of the Durbin Watson Test and the visual analysis of the residuals.

- **Eigenvector Diagram:** the first eigenvectors are plotted against wavelength, concentration, or pH. We can see the number of eigenvectors that present a systematic variation, which indicates the minimum number of components present in system.
- **Eigenvalues logarithmic diagram:** Eigenvalues indicate the variation of the corresponding principal components. Their relative value indicates the number of principal components with a significantly different variation from experimental noise. Ideally, data will show no noise leading to eigenvalues to take a zero value. However, data always shows some sort of noise that needs to be explained by eigenvalues. The plot of eigenvalues against the eigenvalue number shows that the last eigenvalues have a practically constant value corresponding to residual principal components that reproduce spectral random noise (see Figure 4.3). The number of principal components with an eigenvalue above this variance level corresponds to the upper limit of the number of structural components.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

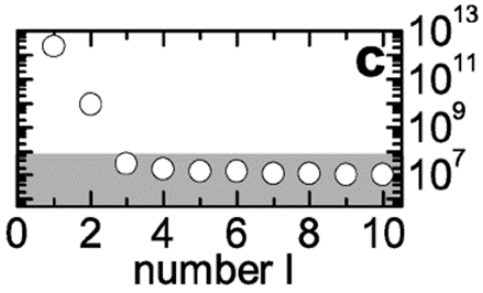



*Figure 4.3. Example of an eigenvalues logarithmic diagram for a system that needs two principal components.*

- **Mean residual error diagram**: This type of logarithmic diagram reflects the mean residual error, which is an estimate of the standard deviation between experimental and theoretical spectra including a certain number of principal components. The mean residual error depends on the number of principal components included in the theoretical spectra and should reach a constant value once systematic variation has been explained.
- **Durbin Watson Test**: It represents the value of the Durbin-Watson parameter for an increasing number of included components. Once the systematic error of the experimental spectra is reproduced by the first principal components, the residuals between the theoretical and experimental spectra should show only noise, i.e., random variations without correlation.
- **Visual analysis of the residuals:** plotting the residuals of the analysis for an increasing number of needed components allows to judge the goodness of the fit of each spectra applying different numbers of components. Contrary to the rest of test were only the minimum number of components needed is obtained, the analysis of the residuals gives us the exact number of components present in the sample.

Once the number of principal components is determined by PCA, we then apply the GA to our data.

## 4.2.2. Global Analysis

The second step in our PCGA analysis is the determination of the pure spectra and the model parameters of our system. This step is the Global Analysis. As exposed above, GA is an analysis methodology that allows the use of the whole data set of a given spectral series in terms of an external parameter such as concentration or pH. It can be implemented already in commercially available data analysis programs, but in those case one must still prepare the data manually to link each value of a specific wavelength to the given independent variable. However, using the one programmed by our group that work is not needed. The program can use all wavelengths from a spectral series and fit it to a specific model. For an equilibrium constant of a system to be obtained, the model that

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

applies to said system must be selected or introduced in the code (all used models in this thesis are explained later in this chapter in section 4.5). Using more wavelengths for a given parameter increases the total number of values, which, statistically leads to a more precise data analysis. This way, the uncertainty calculated by GA can be lowed up to approximately $\sqrt{n}$, where n is the number of analysed wavelengths for a given factor. However, we have come to the conclusion that this program underestimates the real uncertainty of the obtained parameters and, in order to correct this underestimation, one must multiply the obtained uncertainty value by $\sqrt{n}$.[158]

GA is carried out in three steps: model proposal, matrix D and E (Eq (4.12)) calculation and corroboration of the validity of the results.

- **Model proposal:** We postulate a model that describes the spectral variations in matrix D as a function of the external variable ( $x_i^1, x_i^2, \ldots$ ) and the needed parameters (a,b...). The pure spectra $\hat{E}$ are still unknown but can be replaced by the product of the calculated eigenvector with the rotation matrix $\hat{P}$ (Eq (4.18)). For all spectra except CD one condition will be carried out: pure spectra cannot have negative values.

$$\hat{d}_{il} = f_l(a,b,\ldots;x_i^1,x_i^2,\ldots) \tag{4.17}$$

$$\hat{E} = \hat{P}V \tag{4.18}$$

- **Matrix D and E obtention:** The proposed model is then solved to obtain matrix D and E using the least squares method. This means that a nonlinear fit procedure that varies iteratively the parameters of the equilibrium constant and the elements of the rotation matrix minimising the sum of the squared different is carried out.
- **Corroboration of the validity of the results:** As with any other data analysis, the goodness of the fit shall be corroborated. For this, we visually analyse the obtained residuals for the curve fits as well as the similarity between the experimental spectra and the calculated spectra. An appropriate fit would lead to random residual curves with no systematic variations.

# 4.3. Fluorescence Correlation Spectroscopy Data Analysis

## 4.3.1. Fluorescence Correlation Curves

As explained in the previous chapter (section 3.1.2.1), in FCS measurements the detectors record the fluorescence emission intensity as a function of the time. The fluctuations on the intensity registered by each detector can then be transformed into a

correlation curve by correlating each detector with each other (cross-correlation) or with itself (autocorrelation). This transformation is shown in Figure 4.4. On the left side, the fluorescence emission intensity is represented as a function of time, on the right side the autocorrelation values obtained from the data analysis are represented as a function of time in a logarithmic scale.

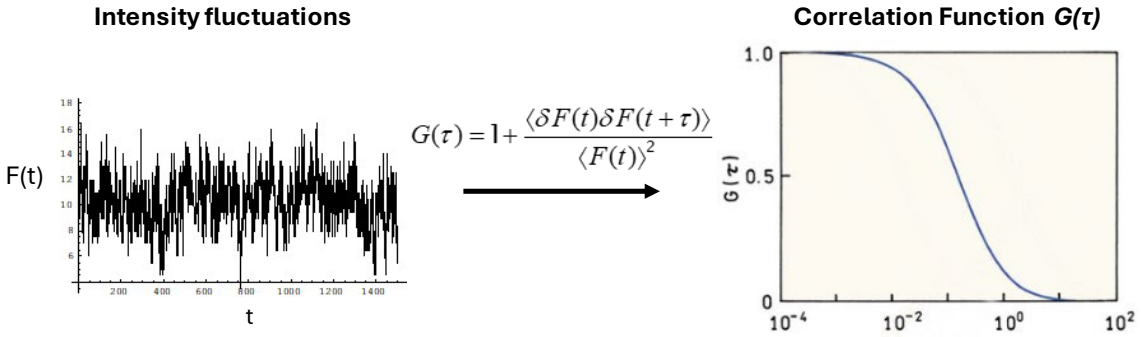



*Figure 4.4. Schematic representation of the transformation FCS undergoes when obtaining the correlation curve from the intensity fluctuations.*

To do this, the detected intensity is analysed by correlating the fluctuations around the average measured intensity in a period of time. The autocorrelation equation gives the probability of detecting a photon with a temporal delay of τ after the detection of a first photon at a time t. The delay is known as correlation time. This probability can be expressed as:

$$G(\tau) = 1 + \frac{\langle \delta F(t) \delta F(t+\tau) \rangle}{\langle F(t) \rangle^2} \tag{4.19}$$

Where <> denotes the temporal average, and δF represents the variation of the fluorescence signal, defined as the difference between the current fluorescence intensity and its mean: $\delta F = F(t) - \langle F(t) \rangle$. The autocorrelation function decays over the correlation time τ because fluctuations become less correlated with increasing temporal separation. Essentially, the longer the time since a molecule emitted a photon, the lower the likelihood of it emitting another photon due to factors like diffusion, triplet states and others that were explained in section 3.1.2.1.

Correlation curves are obtained directly from the data acquisition program (SPC-150 from Becker&Hickl GmbH and MultiHarp 150 from PicoQuant) and need further analysis to obtain the desired molecular parameters. First, the correlation curves, were fitted with correlation functions for the translational diffusion of one (Eq (4.20)) or two (Eq (4.21)) independent species across the sample volume and an additional term due to the transitions of the dye to the dark triplet state:[93,136,159]

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

$$G_D = \frac{1}{N}\left(1+\frac{\tau}{\tau_D}\right)^{-1}\left(1+\frac{\tau}{\omega^2\tau_D}\right)^{-\frac{1}{2}}\left(1+A_T e^{-\tau/\tau_T}\right) \tag{4.20}$$

$$G_D = \frac{1}{N}\left(R\left(1+\frac{\tau}{\tau_{D1}}\right)^{-1}\left(1+\frac{\tau}{\omega^2\tau_{D1}}\right)^{\frac{1}{2}} + (1-R)\left(1+\frac{\tau}{\tau_{D2}}\right)^{-1}\left(1+\frac{\tau}{\omega^2\tau_{D2}}\right)^{\frac{1}{2}}\right)\left(1+A_T e^{-\frac{\tau}{\tau_T}}\right) \tag{4.21}$$

For the expressions shown above, first we must assume that the measured fluorescence intensity follows a Gaussian-like distribution in all three dimensions. Here, N is the mean number of fluorescent molecule in the sample volume, $\tau_{Di}$ is the diffusion correlation time of species *i* across the sample volume and $\omega$ is the aspect ratio of the ellipsoid of the sample volume. The aspect ratio is defined by the radius of the ellipsoid in the xy and z directions (Eq (4.22)).

$$\omega = \frac{w_z}{w_{xy}} \tag{4.22}$$

The sample volume is equal to the volume of an ellipsoid with radius $w_{xy}$ and height $w_z$, related by the aforementioned aspect ratio $\omega$. Thus, it can be calculated following Eq (4.23).

$$V = \frac{4}{3}\pi w_{xy}^3 \omega \tag{4.23}$$

The apparent number of particles $N$ depends on the average number of each type of particles in the sampling volume ($N_i = V\,C_i$), as well as on the brightness of the particles $Q_i$. The brightness of a particle *i* is defined as the product of its molar absorption coefficient, its fluorescence quantum yield and the detection efficiency, $Q_i = \varepsilon_i\,\phi_{(F)i}\,g_i$.

$$\frac{1}{N} = \frac{Q_1^2 N_1 + Q_2^2 N_2}{\left(Q_1 N_1 + Q_2 N_2\right)^2} \tag{4.24}$$

Factors $R$ and $1-R$ are the relative contributions of each type of particle to the FCS curve.

$$R = \frac{Q_1^2 N_1}{Q_1^2 N_1 + Q_2^2 N_2} \tag{4.25}$$

In correlation curves other type of process can appear like the triplet state of the fluorophore, which is the last term shown in Eq (4.21). The triplet state has a characteristic lifetime ($\tau_T$) and amplitude ($A_T$). The amplitude of the triple state term is related to the fraction of molecules of fluorophore that transition into the triplet state (Eq (4.26)).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

$$A_T = \frac{T}{1-T} \tag{4.26}$$

Another process that may appear is the antibunching. The velocity with which two successive photons can be detected is determined by the excitation and emission velocities. The minimum time needed for detection is the antibunching term in the correlation curve ( $G_F(\tau)$ ) and has a characteristic time ( $\tau_F$ ) and amplitude ( $A_F$ ). The antibunching time is the inverse of the sum of the excitation rate constant and the decay rate constant of the dye ( $\tau_F = 1/(k_{exc} + k_{decay})$ ). If the excitation rate constant is much smaller than the emission rate constant then, the antibunching is defined by the fluorescence lifetime of the fluorophore.

$$G_F(\tau) = (1 - A_F e^{-\tau/\tau_F}) \tag{4.27}$$

The complete correlation function will be the product of the three individual functions if the condition of $\tau_D \gg \tau_T \gg \tau_F$ is maintained.

## 4.3.2. Diffusion coefficient and hydrodynamic radius obtention

The diffusion time obtained from the correlation curves is related to the diffusion time of the molecule and the radius of the sample volume. Applying Eq (4.28) we can transform the obtained diffusion times into diffusion coefficients, $D$.

$$D = \frac{w_{xy}^2}{4\tau_D} \tag{4.28}$$

We need the diffusion coefficient because it is a parameter related to the size of the molecule. The bigger the molecule is, the lower its diffusion coefficient will be, as their diffusion will be hindered by the size. For molecular sizes significantly bigger than the size of the solvent molecules, the diffusion coefficient can be is related to the inverse of the hydrodynamic radius ( $R_h$ ) by the Stokes-Einstein Equation (Eq (4.29)). In this equation, $k$ is the Boltzmann constant, T is the temperature at which the experiment was carried out, and $\eta$ is the viscosity.[137]

$$R_h = \frac{kT}{6\pi\eta D} \tag{4.29}$$

The hydrodynamic radius of the fluorophore depends on its molecular mass, as the diffusion coefficient is inversely proportional to the molecular mass (M) to the power of the shape factor ( $\upsilon$ ) (see Eq (4.30), where $a$ is a proportionality factor). The shape factor is a coefficient that accounts for the molecule geometry. The more compact a molecular conformation is, the lower the shape factor would be. For example, a dense sphere, which is the most compact conformation a molecule can acquire, would have a shape factor of $\upsilon \approx 0.3$, while for a random coil is $\upsilon \approx 0.5$ and for a rigid rod is $\upsilon \approx 1$ (Figure 4.5).

UFG UNIVERSIDADE

$$D = aM^{-\upsilon} \tag{4.30}$$

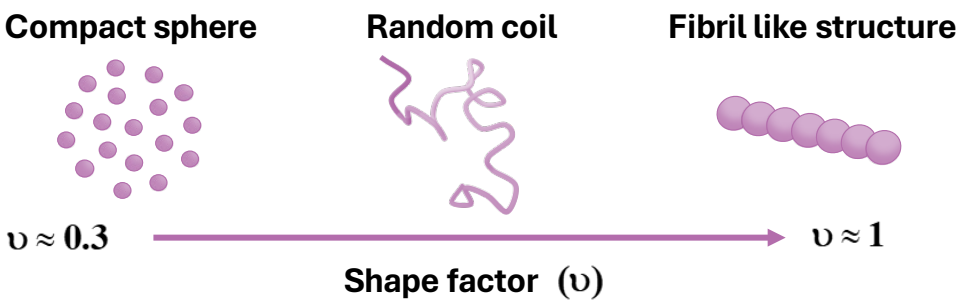



*Figure 4.5. Schematic representation of the shape factor variation for different molecular structures.*

As explained in the previous chapter (section 3.1.2.1.2), FCS needs to be aligned with a reference dye. From this alignment, we obtain the volume of the sample as well as the diffusion time of the reference dye ($D_{ref}$). The molecular mass of the reference dye is also known ($M_{ref}$). Thus, we can divide Eq (4.30) applied to the reference dye and the molecule under study and obtain an expression were the proportionality factor disappears (Eq (4.31)).

$$\frac{D_{ref}}{D_x} = \frac{M_{ref}^{-\upsilon_{ref}}}{M_x^{-\upsilon_x}} \tag{4.31}$$

# 4.4. Image Analysis

Images acquired using the wide-field microscope were treated with the SRRF algorithm explained previously in section 3.1.2.2. After applying this algorithm, videos were transformed to high-resolution images. These files were treated with two different analysis programs to obtain both the number ratio between fibrils and droplets forming at different ion concentrations as well as the size distribution of the formed droplets. For the first analysis, we developed a process in ImageJ, whereas for the second one we used NanoPArt, a program developed by PhD Víctor Leborán from Prof. Rivadulla's group at the University of Santiago de Compostela. Data obtained from both programs was then further analysed using Origin Pro.

## 4.4.1. Fibrils and droplets ratio

Images obtained from high-resolution fluorescent imaging experiments turned out to be quite diverse in terms of shapes that could be seen on them (Figure 4.6). Some conditions allowed only for fibril-like shape formation, while others allowed for only round-like shapes to be seen. However, most conditions allowed for both shapes to appear. For image analysis, the mixture of two significantly different shapes is a challenge without

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

previous treatment. Thus, we developed the following procedure in ImageJ that allowed us to easily determine the percentage of each shape appearing in the image.

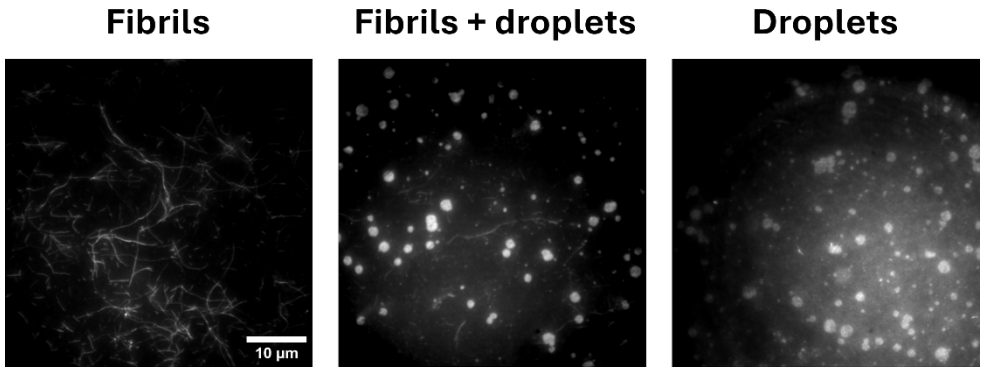



*Figure 4.6. Example images of the types of shapes that can be found in amyloid fluorescence microscopy data.*

First, the contrast was modified manually to achieve the highest contrast possible between the white particles and the background without losing information. When an adequate contrast was obtained, we applied a Hessian filter, which is available to download at ImageJ as port of FeatureJ plugin. The Hessian filter discriminates between plate-like, blob-like and line-like structures within a given image by computing for each pixel the eigenvalues of the Hessian matrix.[160–162] We applied this filter for the smallest eigenvalues of the Hessian tensor and a smoothing scale of 1.2 (Figure 4.7b shows an example of a resulting image after applying this filter). Then, we apply the Otsu auto-threshold to obtain again a black and white high contrast image (Figure 4.7c). If the studied image presents only fibril-like structures, then no further treatment is needed. However, if it does present droplet-like structures, then additional treatment for structure differentiation must be carried out. Thus, we open the original file (Figure 4.7a) again and directly apply the auto-threshold feature, but this time using Yen's threshold as it only detects the droplet-like structures as they tend to be brighter (Figure 4.7d). After this, we subtract this droplet-only image from the image obtained from applying the Otsu auto-threshold. This way we have two different images (Figure 4.7d and Figure 4.7e), each containing only one type of structure that can be directly analysed with ImageJ Black to white ratio calculator. It is important to note that, because some images may present data only in one region of the image due the microscope focus, instead of using the whole image to obtain the black-to-white ratio, only a significant section of the original image was selected for the analysis.

In each of the resulting droplet and fibre only images, the number of white pixels is proportional to the total area of droplets or fibres. To compare images of different size we indicate the percentage of white pixel in the total pixel count.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

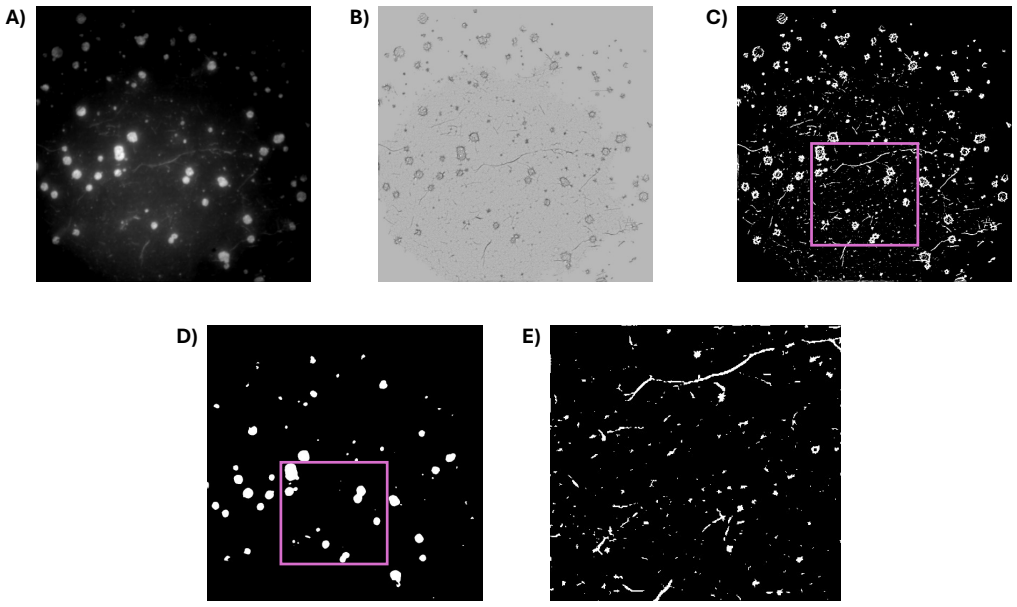



*Figure 4.7. Transformations an image suffers in each step of the image analysis. A) original image, B) image after Hessian filter application, C) image after Otsu auto-threshold application, D) image after Yen auto-threshold application to the original image, E) resulting image from subtracting the selected regions inside the pink squares of image D from the same region in image C.*

### 4.4.2. Size distribution

Images where a significant population of droplets appeared in the previous analysis were selected to obtain its size distribution. Each selected image was uploaded to NanoPArt. This program, which is implemented in MatLab and is free to download, analyses a given image after selecting the scale and shape that one would like to model. For our images, the "Ext circle" shape was selected. After this, the program allows to corroborate if the found structures are correct. If any structure is missing or any of the selected ones is not actually correct, then it allows to manually delete or add as necessary. After selecting the structures that we want to study, the program automatically prints out a histogram of the size of the structures. The calculated data includes the area (pixel$^2$) and the radius (both in pixels and in nm). Radius size distribution was then plotted in OriginPro and analysed to fit a LogNormal distribution.

## 4.5. Model deductions

### 4.5.1. Hill equation for G4 intramolecular association constant

Data obtained for the study of G4 formation from circular dichroism experiments were fitted with a semi-empirical equation commonly used in biochemistry known as Hill's Equation (Eq (4.32)). This equation was first introduced by Hill to describe the equilibrium between the saturation of haemoglobin and oxygen tension. Since then, it has been widely used in the biochemistry and the pharmaceutic fields as a set model to study the

USC UNIVERSIDADE

binding behaviour between a ligand and a macromolecule. It describes the relationship between the concentration of a ligand and the saturation of the binding sites on a macromolecule. It considers the cooperativity a biological system can present, where the binding of one ligand affects the affinity of the rest of the ligands near the binding site.

$$Y = \frac{[L]^n}{K_d^n + [L]^n} \tag{4.32}$$

In Hill's Equation, [L] is the concentration of the ligand, *n* is the degree of cooperativity and is knowns as the Hill coefficient, and $K_\mathrm{d}$ is the dissociation constant. If the Hill coefficient is equal to 1, it means that the studied system is a non-cooperative system. On the other hand, if it is different than 1, then there is cooperativity in the system. This cooperative binding can be positive (n > 1) if one ligand binding to the macromolecule increases the affinity of the binding process, or negative (n < 1), if the affinity of the binding process decreases after one ligand binds to the macromolecule.

Hill's equation is based on one important assumption, and that is that the binding process can be descried by equations that phenomenological simulate an infinite cooperative system. For an infinite cooperativity (all-or-none system) the equation can be obtained as follows.[163]

Taking a binding equilibrium between a given molecule (M) and a ligand (L) with a stoichiometry of 1:n (Eq (4.33)):

$$M_0 + nL \rightleftarrows M_n \tag{4.33}$$

The dissociation constant can be written as shown in Eq (4.34).

$$K_d = \frac{[M_0][L]^n}{[M_n]} = K^n \tag{4.34}$$

Now, the total concentration of the molecule is:

$$[M]_{total} = [M_0] + [M_n] \tag{4.35}$$

As a result of combining Eq (4.34) and Eq (4.35), we obtain Eq (4.36), where F is the fraction of molecules bound to the ligand. For our system, F would be the fraction of G4 that were stabilized by the cation.

$$F = \frac{[M]_n}{[M]_{total}} = \frac{[L]^n}{K^n + [L]^n} \tag{4.36}$$

Now, to apply this equation to circular dichroism spectral series, we have to consider that the total ellipticity at a given wavelength is the combination of the ellipticity of all the molecules contributing to it at that wavelength (Eq (4.37)).

USC DE SANTIAGO DE COMPOSTELA

$$\theta^{\lambda} = \theta_0^{\lambda} l[M_0] + \theta_n^{\lambda} l[M_n] \tag{4.37}$$

Combining Eq (4.36) and Eq (4.37) together, we obtain the following expression for CD spectra.

$$\theta^{\lambda} = \theta_0^{\lambda} l[M]_{total} \frac{K^n}{K^n + [L]^n} + \theta_n^{\lambda} l[M]_{total} \frac{[L]^n}{K^n + [L]^n} = A \frac{K^n}{K^n + [L]^n} + B \frac{[L]^n}{K^n + [L]^n} \tag{4.38}$$

In our case, the sodium or potassium cations are the ligands, and the nucleic acid experiments a change in topology due to the binding to the ligand, which can be described by the following equilibrium equation:

$$G4_i + cation \rightleftarrows G4_f \tag{4.39}$$

where $G4_i$ is the initially adopted topology and $G4_f$ the final topology.

Applying Hill's equation to this equilibrium, the following equation is obtained for the dependency of the molar ellipticity $\theta$ with cation concentration:

$$\theta^{\lambda} = A \frac{K^n}{K^n + [cation]^n} + B \frac{[cation]^n}{K^n + [cation]^n} \tag{4.40}$$

where $K$ is the equilibrium constant of the backward process of topology reorganization described by equation (4.39) and $n$ is the Hill's coefficient. $A$ and $B$ account for the molar ellipticity of the two G4 topologies.

## 4.5.2. Boltzmann equation for nucleic acids melting temperature

CD melting experiments' data were analysed using an empirical function, the modified Boltzmann sigmoidal function:

$$\theta = \theta_2 + \frac{\theta_1 - \theta_2}{1 + e^{\frac{T - T_m}{c}}} \tag{4.41}$$

where $\theta_1$ is the ellipticity value of the plateau at low temperatures and $\theta_2$ the one at high temperatures, and $c$ is a coefficient which indicates the steepness of the curve. The melting temperature ($T_m$) of the nucleic acid corresponds to the inflection point of the curve.

## 4.5.3. Amyloid-beta(1-40) early aggregates analysis procedure

To extract quantitative information about the early amyloid aggregation we make use of a framework of equations describing the species and their properties as observed in the experimental data. These equations have been derived and described in detail in the supplementary information of Novo et al.[68] for the analysis of the aggregation of Aβ42. Here we repeat only briefly the relevant parts and adapt them to the case of the Aβ40

USC DE COMPOSTELA

aggregation. We apply a data analysis procedure based on these equations that is similar to that described in previous works.[68,164]

### 4.5.3.1. Abbreviations

| | |
|---|---|
| A | Amyloid (in any form, unlabelled or labelled, free or aggregated) |
| $A_f$, $A_g$ | Free monomeric amyloid, aggregated amyloid |
| $A^\circ$, $A^*$ | unlabelled and labelled amyloid |
| $\overline{n}$ , $\overline{n}_u$ , $\overline{n}_c$ | Mean number of amyloid monomers per aggregate, uncorrected or corrected, (*mean aggregation number*) |
| $\overline{i}$ | Mean number of labelled amyloid per aggregate (*mean occupation number*) |
| $\overline{i}^*$ | Mean number of labelled amyloid per *labelled* aggregate |
| $\gamma$ | Amyloid aggregation fraction (*degree of aggregation*) |
| $\alpha$ | fraction of labelled amyloid of the total amyloid |
| *cac* | critical aggregation concentration |
| *r* | relative transition width around the *cac* (micelle model) |
| $\tau_{D1}$ | diffusion time of monomeric amyloid |
| $\tau_{D2}$ | diffusion time of aggregated amyloid |
| $N_1$ | number of monomeric labelled amyloid monomers per FCS sample volume |
| $N_2$ | number of labelled amyloid aggregates per FCS sample volume |
| $Q_1$ | brightness of monomeric labelled amyloid |
| $Q_2$ | mean brightness of labelled aggregated amyloid |
| *q* | brightness ratio between a monomeric and an aggregated labelled amyloid |
| $Q_1'$ , *q'* | reference values of $Q_1$ and *q* determined from samples free of aggregates. |
| *I* | registered total fluorescence intensity (photon counts per second) |
| *R*, 1-*R* | relative contributions of the monomeric amyloid and the aggregates to the full amplitude of the diffusion term |
| *cpm* | photon counts per second and molecule |
| *NA* | total number of amyloid molecules per FCS sample volume |
| *NG* | number of aggregates per FCS sample volume |
| $NA_f$, $NA_g$ | number of aggregated amyloid molecules per FCS sample volume |

### 4.5.3.2. Assumptions for the data analysis

The data analysis rests on the following assumptions: [68]

A1. The occupation numbers $\overline{i}$ and $\overline{i}^*$ (mean number of labelled amyloids per (labelled) aggregate) follow a Poisson distribution.

A2. The labelling has neither effect on the aggregation properties of the amyloid monomer nor on its adsorption affinity. This implies that the fraction α of labelled

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

amyloid is the same for free and for aggregated amyloid, as well as for adsorbed and dissolved amyloid.

A3. The brightness $Q_g$ of a labelled amyloid molecule in an aggregate is independent of the size $\bar{n}$ or the occupation $i$ of the aggregate.

### 4.5.3.3. General equations

FCS yields absolute numbers of particles (dye-labelled monomers or aggregates) observed in the sample volume. Therefore, we express the following equations in "numbers" (per volume). These can be converted to concentrations $[X] = NX / V$ with the known sample volume *V*.

The **total number of amyloid** molecules (*NA*) is the sum of the numbers of free unlabelled $(NA_f^\circ)$, free labelled $(NA_f^*)$, aggregated unlabelled $(NA_g^\circ)$, and aggregated labelled amyloid $(NA_g^*)$:

$$NA = NA_f^\circ + NA_f^* + NA_g^\circ + NA_g^* = NA_f + NA_g \tag{4.42}$$

Note that *NA* refers to the available amyloid in the solution as determined from the FCS measurements.

Following assumption A2, the **fraction of labelled amyloid *α*** does not depend on the aggregation, and is the same for free and aggregated amyloid:

$$\alpha = \frac{NA^*}{NA} = \frac{NA_f^*}{NA_f} = \frac{NA_g^*}{NA_g} \qquad = \frac{\textit{labelled A}\beta}{\textit{total A}\beta} \tag{4.43}$$

The fraction of aggregated amyloid (*degree of aggregation*) *γ* can be determined from the labelled amyloid alone:

$$\gamma \equiv \frac{NA_g}{NA} = \frac{NA_g^*}{NA^*} \qquad = \frac{\textit{aggregated A}\beta}{\textit{total A}\beta} \tag{4.44}$$

The ***mean number of amyloid molecules per aggregate* (*size*, or *aggregation number*)** $\bar{n}$ is given by the number of aggregated amyloid and the number of aggregates:

$$\bar{n} = \frac{NA_g}{NG} \qquad = \frac{\textit{aggregated A}\beta}{\textit{number of aggregates}} \tag{4.45}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

The labelled amyloid is randomly distributed among the aggregates, leading to aggregates that bear none, one, or multiple labels. The mean number of labelled amyloid per aggregate, ***mean occupation number*** $\bar{i}$ , is given by the product between the labelling-fraction *α* and the aggregation number $\bar{n}$ :

$$\bar{i} = \frac{NA_g^*}{NG} = \alpha \cdot \bar{n} \qquad \frac{\textit{labelled aggregated A}\beta}{\textit{number of aggregates}} \quad (4.46)$$

The **probability** that an aggregate contains $i$ labelled amyloids is given by a Poisson distribution (assumption A2):

$$P(i) = \frac{NG^i}{NG} = \frac{\bar{i}^{\,i} e^{-\bar{i}}}{i!} \qquad \frac{\textit{aggregates with i A}\beta^*}{\textit{number of aggregates}} \quad (4.47)$$

The **number of labelled aggregates** $NG^*$ is then given by

$$NG^* = NG \cdot (1 - e^{-\bar{i}}) = \frac{NA_g^*}{\bar{i}} \cdot (1 - e^{-\bar{i}}) = \frac{NA_g^*}{\alpha \cdot \bar{n}} \cdot (1 - e^{-\alpha \cdot \bar{n}}) \quad (4.48)$$

The *brightness of labelled amyloid* may change upon aggregation. According to the assumption A3 the **ratio *q* between the brightness of a free ($Q_f$) and an aggregated ($Q_g$) labelled amyloid molecule** is independent on $\bar{n}$ or $i$ (note that $q$ refers to a single labelled amyloid molecule):

$$q = \frac{Q_g}{Q_f} \quad (4.49)$$

The **mean brightness** $Q_2 \equiv Q_{G^*}$ **of labelled aggregates** can be estimated from the brightness of aggregates with a *single* labelled amyloid, $Q_g$, multiplied by the ***mean number of labelled amyloid per* labelled *aggregate*** $\bar{i}^*$ :

$$Q_2 \equiv Q_{G^*} = \bar{i}^* Q_g \quad (4.50)$$

This occupation number $\bar{i}^*$ is different from $\bar{i}$ , which counts also unlabelled aggregates. Using Eq (4.51) we get:

$$\bar{i}^* = \frac{NA_g^*}{NG^*} = \bar{i}\left(1 - e^{-\bar{i}}\right)^{-1} = \alpha\bar{n}\left(1 - e^{-\alpha\bar{n}}\right)^{-1} \quad \frac{\textit{labelled aggregated A}\beta}{\textit{number of labelled aggregates}} \quad (4.51)$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

The ratio between the mean brightness of labelled aggregates and the brightness of free amyloid is then:

$$\frac{Q_2}{Q_1} \equiv \frac{Q_{G^*}}{Q_f} = q\bar{i}^* = q\frac{\bar{i}}{1-e^{-\bar{i}}} = q\frac{\alpha\bar{n}}{1-e^{-\alpha\bar{n}}} \tag{4.52}$$

Equation (4.52) thus relates the mean aggregation number $\bar{n}$ to the brightness of labelled free amyloid and aggregates.

With Eq (4.53) the degree of aggregation $\gamma$ can then be estimated from $N_1$ and $N_2$ determined in FCS:

$$\gamma = \frac{N_2 \cdot \bar{i}^*}{N_1 + N_2 \cdot \bar{i}^*} \tag{4.53}$$

### 4.5.3.4. Estimation of the mean aggregation number from the diffusion times

As explained before, the diffusion coefficient of homogeneous particles across the FCS sample volume change with their molar mass following a power law, where the exponent $\upsilon$ is related to the geometry of the particle[165]. Because the diffusion coefficient is related to the diffusion time, we can rearrange Eq (4.30) and Eq (4.28) to obtain the diffusion time as a function of the molar mass.

$$\tau_D = \frac{w_{xy}^2}{4 \cdot D} = \frac{w_{xy}^2}{4 \cdot a \cdot M^{-\upsilon}} = b \cdot M^{\upsilon} \tag{4.54}$$

The mean molar mass of the aggregates $M_2$ is given by the mean aggregation number and the molar mass of the amyloid monomers $M_1$ (disregarding the mass of the labels):

$$M_2 = \bar{n} \cdot M_1 \tag{4.55}$$

The ratio of the diffusion times of aggregates and monomers is then,

$$\frac{\tau_{D,2}}{\tau_{D,1}} = \frac{b_2 M_2^{\upsilon_2}}{b_1 M_1^{\upsilon_1}} = \frac{b_2 M_1^{\upsilon_2} \bar{n}^{\upsilon_2}}{b_1 M_1^{\upsilon_1}} = \frac{b_2}{b_1} M_1^{(\upsilon_2 - \upsilon_1)} \bar{n}^{\upsilon_2} \tag{4.56}$$

$$\log \frac{\tau_{D,2}}{\tau_{D,1}} = \log \left[ \frac{b_2}{b_1} M_1^{(\upsilon_2 - \upsilon_1)} \right] + \upsilon_2 \log \bar{n} \tag{4.57}$$

The slope of the linear relationship given in Eq (4.57), $\upsilon_2$ gives thus information about the geometry of the aggregates.

### 4.5.3.5. Fraction of aggregated amyloid

In FCS only labelled aggregates contribute to the signal. From equations (4.21)-(4.25), it follows:

$$N_1 = \frac{Q_2^2 N R}{\left(Q_1(1-R)+Q_2 R\right)^2},\ N_2 = \frac{Q_1^2 N(1-R)}{\left(Q_1(1-R)+Q_2 R\right)^2} \tag{4.58}$$

The registered fluorescence intensity *I* is given by

$$I = I_1 + I_2 = Q_1 N_1 + Q_2 N_2 \tag{4.59}$$

Substituting in Eq (4.58) the number of free labelled amyloid can be determined as

$$N_1 = \frac{R}{N}\frac{I^2}{Q_1^2} \tag{4.60}$$

and the number and brightness of labelled aggregates are given by

$$N_2 = \frac{I - Q_1 N_1}{Q_2} \tag{4.61}$$

$$Q_2 = \frac{1-R}{R}\frac{Q_1^2 N_1}{I - Q_1 N_1} \tag{4.62}$$

Substituting in Eq (4.53) we obtain for the fraction of aggregated amyloid *γ*:

$$\gamma = \frac{1-R}{1+R(\bar{i}^* q^2 - 1)} \tag{4.63}$$

The fraction of aggregated amyloid *γ* can therefore be estimated from the FCS-parameter *R*, avoiding the use of the absolute intensities detected in FCS which suffer from great uncertainties.

### 4.5.3.6. Total Amyloid Concentration in Solution

With Eq (4.43) and (4.53) the total number of amyloid in the sample volume (FCS focus) is obtained as

$$NA = \frac{NA^*}{\alpha} = \frac{N_1 + N_2 \cdot \bar{i}^*}{\alpha} \tag{4.64}$$

The true total amyloid concentration in solution is then obtained with the sample volume *V* determined from measurements with the reference dye with known diffusion coefficient as described above ($N_A$: Avogadro number):

$$[A] = \frac{NA}{N_A V} \tag{4.65}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

### 4.5.3.7. Data analysis procedure

The followed procedure for amyloid data analysis was already reported by our group in the Supplementary Information of Novo et al.[68] and in the method that was recently published.[164] The following procedure was applied to the raw FCS curves measured with different amyloid concentrations to obtain the results presented in this thesis in **section 5.2**.

**Step 1: Fit of the FCS curves.**

The reference value of **the monomer brightness** $Q_1'$ of free labelled amyloid was determined from FCS experiments with monomeric labelled amyloid at the lowest Aβ concentration, fitting the one species model, Eq (4.20).

All other correlation curves were fitted with the correlation function for the translational diffusion of two independent species, Eq (4.21), obtaining the **diffusion times** of the monomer ($\tau_{D1}$) and the oligomer ($\tau_{D2}$), as well as their **relative contributions** (***R***) and the **total number of observed molecules** in the sample volume (***N***). (Step 1.[68] and 3.3.1.[164])

For each FCS curve, we calculated from these parameters and the total detected fluorescence intensity *I*, the **brightness $Q_2$ of the oligomers** and the **numbers of particles $N_1$, $N_2$**, using Eq (4.60)-(4.62).

The **labelling fraction *α*** of each sample was calculated from the corresponding concentrations of labelled and unlabelled Aβ.

**Step 2: Brightness ratio *q*.**

Labelled amyloid presents a slightly different brightness ($Q_g$) in aggregates than as free monomer ($Q_f$). We determined the **brightness ratio *q*** = $Q_g/Q_f = Q_2/Q_1$ from samples with a very low labelled amyloid fraction ($\alpha < 0.1$) that assures that aggregates contain no more than one labelled amyloid (see Eq (4.52)). The mean of $Q_2/Q_1'$ ratios of samples with low *α* yields the reference value of *q'*, which will be used as well as $Q_1'$ for all subsequent analyses. (Step 2.[68] and 3.3.2.[164])

**Step 3: Individual label fractions and uncorrected mean aggregation numbers.**

Using the reference values $Q_1'$ and *q'* we then calculated for each experiment the **uncorrected mean number of labelled monomers per aggregate** from the individual values of $Q_2$ (see (4.50)):

$$\overline{i_u}^* = \frac{Q_2}{q' Q_1'} \tag{4.66}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Then we estimated for each experiment a first **uncorrected aggregation number** $\bar{n}_u$ solving numerically (Eq (4.51)):

$$\bar{i}_u^* = \frac{\alpha \bar{n}_u}{1 - e^{-\alpha \bar{n}_u}} \tag{4.67}$$

These values of the aggregation number based on the brightness $Q_2$ are strongly affected by the fluctuations in the total intensity. They will be used in the next step but will then be substituted by better estimates derived from the diffusion times. (Step 3.[68] and 3.3.3.[164])

**Step 4: Molar mass dependence of the diffusion times. Shape factor.**

Because the diffusion time of the aggregates is dependent on their molar mass $M_2$, Eq (4.54), which in turn can be estimated from the mean aggregation number $\bar{n}$ and the monomers' molar mass, $M_2 = \bar{n} M_1$, the ratio between the diffusion times of aggregates, $\tau_{D2}$, and monomers, $\tau_{D1}$, is related to $\bar{n}$ by Eqs (4.56) and (4.57):

$$\log \frac{\tau_{D2}}{\tau_{D1}} = intercept + \upsilon_2 \log \bar{n}_u \tag{4.68}$$

The slope of this linear relationship is an estimate of the exponent in the mass dependence of the diffusion coefficient $D \sim M_2^{-\upsilon_2}$. This "shape factor", $\nu_2$, gives information about the geometry of the aggregates. A dense sphere would correspond to $\nu_2 \approx 0.3$, a random coil to $\nu_2 \approx 0.5$ and a rigid rod to $\nu_2 \approx 1$. (Step 4.[68] and 3.3.4.[164])

**Step 5: Corrected mean aggregation numbers.**

Using the slope and the mean intercept from the previous step, we calculate for each experiment a new corrected aggregation number $\bar{n}_c$ solving the equation of the diffusion time ratios for $\bar{n}$ (Eq (4.69)). (Step 5.[68] and Note 10.[164])

$$\bar{n}_c = \left( \frac{\tau_{D,2}}{\tau_{D,1}} e^{-intercept} \right)^{1/\upsilon_2} \tag{4.69}$$

**Step 6: Concentration dependence of the degree of aggregation γ.**

Next, with the values of the corrected mean aggregation number, $\bar{n}_c$, we calculate corrected values of the labelled occupation number, $\bar{i}_c^*$ (Eq (4.70)).

$$\bar{i}_c^* = \frac{\alpha \bar{n}_c}{1 - e^{-\alpha \bar{n}_c}} \tag{4.70}$$

With the values of $\bar{i}_c^*$, $q'$ and the individual values of $R$, we determine the degree of aggregation γ for each sample applying Eq . (Step 6.[68] and 3.3.5.[164])

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

To estimate the true amyloid concentration in each sample, we first calculate corrected values of the brightness of the oligomers:

$$Q_{2c} = \overline{i}_c^{*} q' Q_1' \tag{4.71}$$

Then we calculate corrected particle numbers $N_{1c}$ and $N_{2c}$ using Eq (4.72) and the corrected parameters obtained from the diffusion times.

$$N_{1c} = \frac{{Q_{2c}}^2 N R}{\left(Q_1'(1-R) + Q_{2c} R\right)^2}, \qquad N_{2c} = \frac{Q_1'^2 N(1-R)}{\left(Q_1'(1-R) + Q_{2c} R\right)^2} \tag{4.72}$$

The total number, *NA*, of amyloid molecules (both monomeric and aggregated) in the FCS sample volume is obtained from Eq (4.73):

$$NA = \frac{N_{1c} + N_{2c}\overline{i}_c^{*}}{\alpha} \tag{4.73}$$

The true total amyloid concentration is then obtained with the sample volume *V* determined from measurements with the reference dye with known diffusion coefficient as described above and the Avogadro constant $N_A$:

$$[A\beta 40]_{true} = \frac{NA}{N_A V} \tag{4.74}$$

**Step 7: Critical aggregation concentrations.**

The critical aggregation concentrations of Aβ40 are determined from fits of the aggregation-concentration-model with two critical aggregation concentrations, Eq (4.79) -(4.82), to the dependence of the degree of aggregation *γ* with the amyloid concentration of the complete data set.

We obtained *γ*-values from 23 samples, each measured at 5 to 12 different incubation times, in total about 160 values. We represent each sample by the mean and the standard deviation of the *γ*-values of each repetition and then fit these mean values weighted by their standard deviations.[68] This fit gives those samples with a tighter standard deviation a higher weight and thus reduces the influence of samples with stronger fluctuations. This weighted fit of the mean values yields in our opinion the best estimate for the *cac* using this fit model and is the one used in the article.

**Step 8: Distribution of the aggregate size.**

To construct the size distribution histogram, we calculated for each FCS curve the number of aggregates (labelled or not) in the sample volume from $N_{2c} \equiv NG^{*} = NG\cdot(1-e^{-\overline{i}})$, Eq (4.75):

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

$$NG = \frac{N_{2c}}{1 - e^{-\alpha \bar{n}_c}} \tag{4.75}$$

using the corrected number of labelled aggregates $\bar{n}_c$ and the corrected mean aggregation number $\bar{n}_c$ determined before. (Step 7.[68] and 3.3.6.[164])

Then, using Eq (4.76), we normalize the relative number of aggregates to the total number of Aβ molecules *NA*, considering that each sample had different amyloid concentrations.

$$NG_{norm} = \frac{NG}{NA} \tag{4.76}$$

Then, the size distribution histogram was constructed summing up the values of $NG_{norm}$ of all experiments in discrete intervals of $\bar{n}_c$ . This histogram was fitted to a double Lognormal distribution.

$$f(x) = \frac{1}{x\sqrt{2\pi}}\left[\frac{1}{\sigma_1}\exp\left(-\frac{1}{2\sigma_1^2}\ln(\frac{x}{\mu_1^*})^2\right) + \frac{1}{\sigma_2}\exp\left(-\frac{1}{2\sigma_2^2}\ln(\frac{x}{\mu_2^*})^2\right)\right] \tag{4.77}$$

The mean aggregation number of each distribution is indicated as $\bar{n}_{mean,2} = \mu_2^*$ and $\bar{n}_{mean,2} = \mu_2^*$, with widths $\sigma_1$, and $\sigma_2$. From these values the 68.3% intervals [$\mu^*/e^{\sigma}$, $\mu^* \times e^{\sigma}$] were calculated.[166]

## 4.5.4. Double micelle model for amyloid aggregation

As in the case of Aβ42, the concentration of monomeric and aggregated amyloid around the *cac* can be estimated with the aggregation concentration model we reported previously for the micelle formation.[167] However, this model was developed for a single critical aggregation process. To consider two different critical aggregation concentrations as observed in Aβ40, we extended the model as follows:

We assume two types of aggregates, A and B, that coexist in solution and are formed from Aβ40 monomers above their respective critical concentrations, $cac_A$ and $cac_B$. The concentration of amyloid forming aggregates A and B is [$A_{g,A}$] and [$A_{g,B}$], respectively.

The degree of aggregation of total Aβ40 is the sum of the contributions of aggregates A and B:

$$\gamma = \frac{[A_g]}{[A]} = \frac{[A_{g,A}] + [A_{g,B}]}{[A]} = \gamma_A + \gamma_B \qquad \gamma_A = \frac{[A_{g,A}]}{[A]},\ \gamma_B = \frac{[A_{g,B}]}{[A]} \tag{4.78}$$

Then we define a maximum contribution of each type of aggregate to the degree of aggregation at very high total amyloid concentration:

$$\gamma_{A,max} = \lim_{[A]\to\infty} \gamma_A \qquad \gamma_{B,max} = \lim_{[A]\to\infty} \gamma_B \qquad \gamma_{max} = \gamma_{A,max} + \gamma_{B,max} \tag{4.79}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

To calculate the concentrations of amyloid forming aggregates A and B, $[A_{g,A}]$ and $[A_{g,B}]$, we apply our reported concentration model[167]. This model estimates the concentration of free amyloid $[A_f]$ as a function of the total amyloid concentration [A], the critical aggregations concentrations $cac_A$ and $cac_B$, the relative widths of the transition regions, $r_A$ and $r_B$.

$$[A_f]([A],cac,r) = cac\left\{1-\frac{a}{2}\left[\sqrt{\frac{2}{\pi}}\, r\, e^{-\frac{1}{2r^2}\left(\frac{[A]}{cac}-1\right)^2}+\left(\frac{[A]}{cac}-1\right)\left(\mathrm{erf}\left[\frac{1}{\sqrt{2}r}\left(\frac{[A]}{cac}-1\right)\right]-1\right)\right]\right\}$$
$$a = 2\left(1+\sqrt{\frac{2}{\pi}}\, r\, e^{-\frac{1}{2r^2}}+\mathrm{erf}\left[\frac{1}{\sqrt{2}r}\right]\right)^{-1} \quad (4.80)$$

With the limiting degrees of aggregation, $\gamma_{A,max}$ and $\gamma_{B,max}$ the concentrations of amyloid forming aggregates A and B, $[A_{g,A}]$ and $[A_{g,B}]$ can be calculated:

$$[A_{g,A}] = \gamma_{A,\max}\left([A]-[A_f]([A],cac_A,r_A)\right)$$
$$[A_{g,B}] = \gamma_{B,\max}\left([A]-[A_f]([A],cac_B,r_B)\right) \quad (4.81)$$

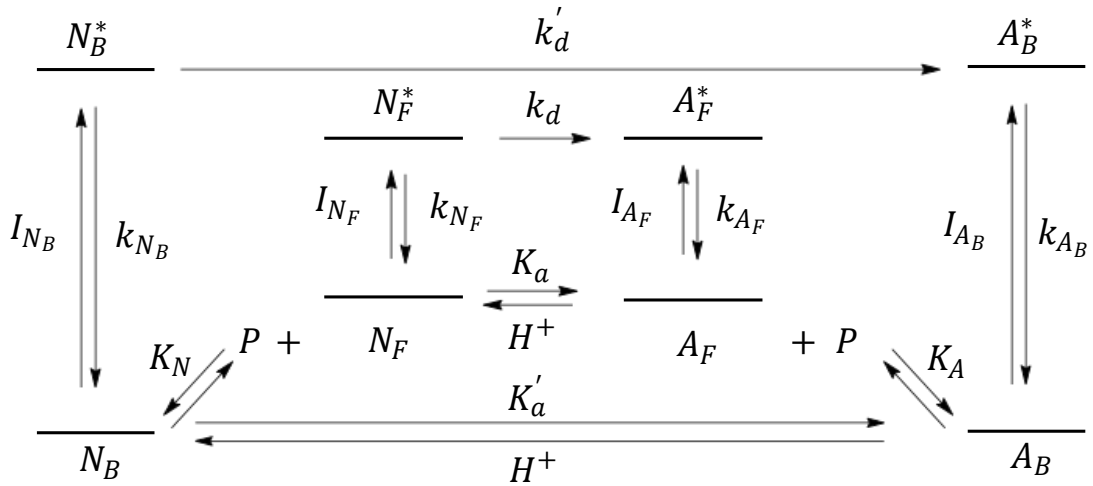
With Eq (4.78) we obtain a fit function for the degree of aggregation $\gamma$ that includes a baseline $\gamma_0$:

$$\gamma = \gamma_{A,\max}\frac{[A]-[A_f]([A],cac_A,r_A)}{[A]}+\gamma_{B,\max}\frac{[A]-[A_f]([A],cac_B,r_B)}{[A]}+\gamma_0 \quad (4.82)$$

## 4.5.5. 1:1 host-guest association for CBTOH:SA system

The model applied to the association between 2-cyano-6-hydroxybenzothiazole (CBTOH) and serum albumins blood proteins (SA) was a 1:1 host-guest interaction. To obtain this model, we have first to consider the whole mechanism that takes place both in the fundamental state and the excited state. As shown in Scheme 4.1, we proposed a mechanism were two host-guest 1:1 binding process can take place in the fundamental state as the CBTOH molecules continue to be irradiated and, thus, transition in and out of the excited state. As there is an acid-base equilibrium between the free neutral CBTOH ($\mathbf{N_F}$) and the free anionic $CBTO^-$ ($\mathbf{A_F}$), and both coexist in the studied conditions, the binding process can occur for both species forming CBTOH:SA (bound $\mathbf{N_B}$) and CBTOH:SA (bound $\mathbf{A_B}$).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA




*Scheme 4.1. Proposed binding mechanism between CBTOH and blood plasma protein (P is either BSA or HSA), together with the excitation and decay processes.*

Following the proposed mechanism, the total concentration of the dye in solution would be the sum of all its species (Eq (4.83)).

$$a_0 = [N_F] + [A_F] + [N_B] + [A_B] \tag{4.83}$$

For the ground-state acid-base process, de dissociation constant is as shown in Eq (4.84), and that for a 1:1 host-guest binding process the equilibrium constant is given by Eq (4.85), where G is the guest (or protein in our system), H is the host (or dye in our system) and G:H is the formed complex.

$$Ka = \frac{[A^-][H^+]}{[AH]} \tag{4.84}$$

$$K_D = \frac{[G:H]}{[G][H]} \tag{4.85}$$

Now, if we apply this equations to our system we obtain Eq (4.86) and Eq (4.87) for the acid-base processes and Eq (4.88) and Eq (4.89) for the host-guest process.

$$K_a = \frac{[A_F][H^+]}{[N_F]} \tag{4.86}$$

$$K_a' = \frac{[A_B][H^+]}{[N_B]} \tag{4.87}$$

$$K_N = \frac{[N_B]}{[N_F][P]} \tag{4.88}$$

$$K_A = \frac{[A_B]}{[A_F][P]} \tag{4.89}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

If we solve for $A_F$, $N_B$ and $A_B$ from equations (4.86), (4.88) and (4.89), and then substitute their solutions in Eq (4.83), we get Eq (4.90), where $R_H$ is the ratio between the acidic constant for the free dye ($K_a$) and the proton concentration, $R_H = (K_a / [H^+])$.

$$[N_F] = \frac{a_0}{1 + R_H + (K_N + K_A R_H)[P]} \tag{4.90}$$

And for the rest of species concentrations, we obtain:

$$[N_B] = \frac{a_0 K_N [P]}{1 + R_H + (K_N + K_A R_H)[P]} \tag{4.91}$$

$$[A_F] = \frac{a_0 R_H}{1 + R_H + (K_N + K_A R_H)[P]} \tag{4.92}$$

$$[A_B] = \frac{a_0 R_H K_A}{1 + R_H + (K_N + K_A R_H)[P]} \tag{4.93}$$

To apply this model to fluorescence spectra, we must assume a steady-state condition and take into consideration that the total fluorescence intensity at a given wavelength ($F^\lambda$) will be a linear combination of the fluorescence of each of the contributing species. The fluorescence intensity of each species will be given by its radiative deactivation constant ($k_r^i$), its excited-state concentration in steady-state conditions ($[i^*]_{ss}$), and a parameter ($\varpi_i^\lambda$) that depends on instrument characteristics and on the species emission spectra. Applying this to our system, we get:

$$F^\lambda = \varpi_{A_F}^\lambda k_r^{A_F} [A_F^*]_{ss} + \varpi_{N_B}^\lambda k_r^{N_B} [N_B^*]_{ss} + \varpi_{A_B}^\lambda k_r^{A_B} [A_B^*]_{ss} \tag{4.94}$$

From Scheme 4.1, we can deduce the following expressions for the steady-state concentrations of each of the involved species, where $I_{N_F}$, $I_{N_B}$, $I_{A_F}$, $I_{A_B}$ represent the (spectral) photon flux density (colloquially intensity) absorbed by each species, $k_{N_F}$, $k_{N_B}$, $k_{A_F}$, $k_{A_B}$ are their deactivation rate constants and $k_d$ and $k_d^{'}$ are the acid dissociation constant of $\mathbf{N_F^*}$ and $\mathbf{N_B^*}$.

$$[N_F^*]_{ss} = \frac{I_{N_F}}{k_d + k_{N_F}} \tag{4.95}$$

$$[N_B^*]_{ss} = \frac{I_{N_B}}{k_{N_B} + k_d^{'}} \tag{4.96}$$

$$[A_F^*]_{ss} = \frac{I_{A_F} + k_d [N_F^*]_{ss}}{k_{A_F}} \tag{4.97}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

$$[A_B^*]_{ss} = \frac{I_{A_B} + k_d'[N_B^*]_{ss}}{k_{A_B}} \tag{4.98}$$

The light intensity ($I_i$) absorbed by a given emissive species ($i$) can be calculated from its absorbance $A_i$ using the Lambert-Beer equation (Eq (4.99)), which relates its absorbance to the ratio between the incident light intensity ($I_0$) and the transmitted light intensity, ($I_{t(i)}$). The intensity absorbed by species $i$ is the difference between the incident and the transmitted light intensity (Eq (4.99)). If we rewrite the equation as the product of the incident light intensity multiplied by a factor *B*, which contains an exponential term of the absorbance. Then, we can apply Taylor's series to the exponential term to simplify this factor, obtaining Eq (4.100). For low values of the absorbance, as is usually the case in fluorescence measurements, the Taylor series can be truncated at the linear term, which results in the absorbed light intensity being proportional to the absorbance. However, it is important to keep in mind that linearity is only maintained in low absorption samples, and therefore, this can only be applied to samples with absorbance values approximately below 0.1.

$$\log_{10}\left(\frac{I_0}{I_{t(i)}}\right) = A_i \Rightarrow I_{t(i)} = I_0 10^{-A_i} \Rightarrow I_i = I_0 - I_{t(i)} = I_0(1-10^{-A_i}) = I_0 B \tag{4.99}$$

$$B = 1 - \left[1 - 2.303A_i + \left(\frac{2.303A_i}{2}\right)^2 + \ldots\right] \approx 2.303A_i \tag{4.100}$$

If we replace the *B* factor in Eq (4.99) for the value obtained in Eq (4.100)then we get the following expression for the light intensity absorbed by each species at a given wavelength:

$$I_i^\lambda = 2.303 I_0^\lambda A_i^\lambda = 2.303 I_0^\lambda \varepsilon_i^\lambda l[i] \tag{4.101}$$

Replacing Eq (4.90)-(4.93) in Eq (4.101), we obtain the absorbed light intensity expression for each species in our system.

$$I_{N_F}^\lambda = 2.303 I_0^\lambda \varepsilon_{N_F}^\lambda l[N_F] \tag{4.102}$$

$$I_{A_F}^\lambda = 2.303 I_0^\lambda \varepsilon_{A_F}^\lambda l[A_F] \tag{4.103}$$

$$I_{N_B}^\lambda = 2.303 I_0^\lambda \varepsilon_{N_B}^\lambda l[N_B] \tag{4.104}$$

$$I_{A_B}^\lambda = 2.303 I_0^\lambda \varepsilon_{A_B}^\lambda l[A_B] \tag{4.105}$$

If we now replace these expressions (Eq (4.102)-(4.105)) in Eq (4.94) we obtain the following expression (Eq (4.106)) where $F_{A_F}^\lambda$, $F_{N_B}^\lambda$ and $F_{A_B}^\lambda$ are the fluorescence intensities at a given wavelength for $A_F^*$, $N_B^*$ and $A_B^*$ if the dye only existed as $A_F$, $N_B$, and $A_B$, respectively, in solution.

USC UNIVERSIDADE

$$F^{\lambda} = \frac{F^{\lambda}_{A_F}\left[R_H + \frac{\varepsilon^{\lambda}_{N_F}}{\varepsilon^{\lambda}_{A_F}}\frac{k_d}{k_d + k_{N_F}}\right] + \left[F^{\lambda}_{N_B}K_N + F^{\lambda}_{A_B}R_H K_A + F^{\lambda}_{A_B}\frac{\varepsilon^{\lambda}_{N_B}}{\varepsilon^{\lambda}_{A_B}}\frac{k_d^{'}}{k_d^{'} + k_{N_B}}\right][P]}{1 + R_H + (R_H K_A + K_N)[P]} \quad (4.106)$$

If we divide both the numerator and the denominator by $1+R_H$ we can now simplify Eq (4.106) as follows:

$$F^{\lambda} = \frac{a^{\lambda} + b^{\lambda}K_{ap}[P]}{1 + K_{ap}[P]} \quad (4.107)$$

$$a^{\lambda} = \frac{F^{\lambda}_{A_F}\left[R_H + \frac{\varepsilon^{\lambda}_{N_F}}{\varepsilon^{\lambda}_{A_F}}\frac{k_d}{k_d + k_{N_F}}\right]}{1 + R_H} \quad (4.108)$$

$$b^{\lambda} = \frac{F^{\lambda}_{N_B}K_N + F^{\lambda}_{A_B}R_H K_A + F^{\lambda}_{A_B}\frac{\varepsilon^{\lambda}_{N_B}}{\varepsilon^{\lambda}_{A_B}}\frac{k_d^{'}}{k_d^{'} + k_{N_B}}K_N}{K_N + R_H K_A} \quad (4.109)$$

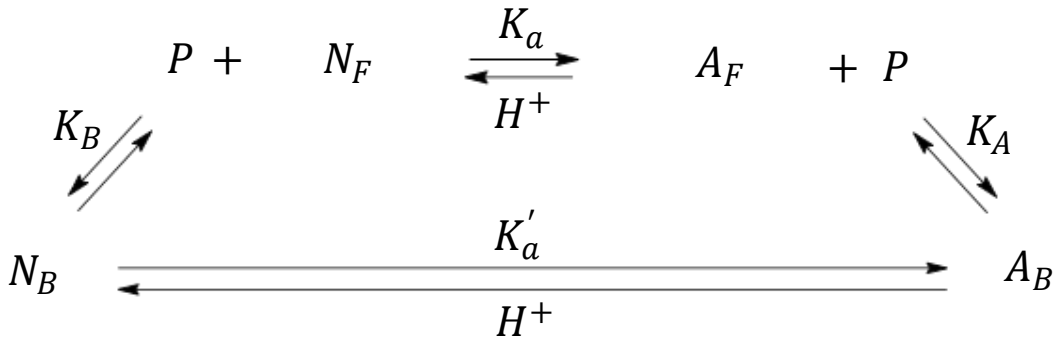
$$K_{ap} = \frac{K_N + R_H K_A}{1 + R_H} \quad (4.110)$$

Now, in the conditions of our experiments, at a physiological pH of around 7.4, the fluorophore will present a $R_H$ of around 0.5, which will give us that the apparent constant expression can be simplified as:

$$K_{ap} = \frac{1}{3}K_A + \frac{2}{3}K_N \quad (4.111)$$

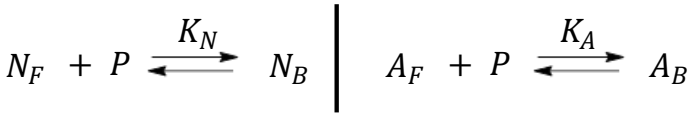
For the electronic ground state, the process is highly simplified, as no excited state transitions are taking place (Scheme 4.2). In addition, changing the pH media to highly acid or highly basic pH allowed us to choose which binding process to study. This way, the acid-basic equilibrium will be shifted towards the anionic species or the neutral species allowing for only one of them to exist in solution.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA




*Scheme 4.2. Proposed binding mechanism between CBTOH and blood plasma protein (P is either BSA or HAS) for the electronic ground state.*

Thus, the process in Scheme 4.2 can be split into two independent processes as shown in Scheme 4.3.




*Scheme 4.3. Binding mechanism between CBTOH and blood plasma protein in different pH conditions. On the left: mechanism in acidic media. On the right: mechanism in basic media.*

In these limiting conditions, as only one host-guest process is taking place, the equilibrium constants $K_N$ and $K_A$ for the neutral and anionic species can be easily calculated if we consider that the total absorbance of a sample at a given wavelength is a linear combination of the absorbances of the different species contributing to the spectra. For our system, the expressions for the neutral and anionic process would be as shown in Eq (4.112) and Eq (4.113) respectively.

$$A^{\lambda} = \varepsilon^{\lambda}_{N_F} l[N_F] + \varepsilon^{\lambda}_{N_B} l[N_B] \tag{4.112}$$

$$A^{\lambda} = \varepsilon^{\lambda}_{A_F} l[A_F] + \varepsilon^{\lambda}_{A_B} l[A_B] \tag{4.113}$$

Now, the total concentration in solution is given by Eq (4.114) for the neutral and Eq (4.115) for the anionic form.

$$a^{N}_{0} = N_F + N_B \tag{4.114}$$

$$a^{A}_{0} = A_F + A_B \tag{4.115}$$

If we replace the concentration of the complex in Eq (4.88) and (4.89) and then in Eq (4.112) and Eq (4.113), we obtain an expression that allows for equilibrium constants calculation as a function of the absorbance (Eq (4.116) and Eq (4.117) for the neutral and anionic process, respectively).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

$$A_N^{\lambda} = \frac{A_{N_F}^{\lambda} + A_{N_B}^{\lambda} K_N P}{1 + K_N P} \tag{4.116}$$

$$A_A^{\lambda} = \frac{A_{A_F}^{\lambda} + A_{A_B}^{\lambda} K_A P}{1 + K_A P} \tag{4.117}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# 5. Results and discussion

## 5.1. Spectroscopic Characterisation of Mitochondrial DNA G-Quadruplexes

**This section is a reproduction of an already published research.**

The results presented on this section have already been published as:

Illodo, S.[1,2]; Pérez-González, C.[1,2]; Barcia, R.[3]; Rodríguez-Prieto, F.[2]; Al-Soufi, W.[1]; Novo, M[1]. Spectroscopic Characterization of Mitochondrial G-Quadruplexes. Int. J. Mol. Sci. **2022**, 23, 925. https://doi.org/10.3390/ijms23020925

[1]Departamento de Química Física, Facultade de Ciencias, Universidade de Santiago de Compostela, 27002 Lugo, Spain

[2]Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Física, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain

[3]Departamento de Bioquímica e Bioloxía Molecular, Facultade de Veterinaria, Universidade de Santiago de Compostela, 27002 Lugo, Spain

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

In this section of Results and Discussion the obtained results of the study of formation of G4 in mitochondrial DNA, RNA and hybrid DNA:RNA sequences of the CSB II region will be discussed. The specific experimental parameters and conditions were previously presented in **Experimental methods** (sections 3.1.1.2, 3.1.1.3 and 3.1.1.4) and the sample preparation protocol was explained in section 3.3.1.

## 5.1.1. Circular Dichroism

### 5.1.1.1. DNA quadruplexes: influence of the sequence

To study the influence of the length sequence on GP formation, we carried out CD measurements. For this purpose, CD spectra were acquired for the DNA sequences with the guanine 1 bp upstream (GCGXAGY) and without it (GXAGY) and all three lengths of the chains (see Table 3.3 in section 3.2 of **Experimental methods**).These spectra are shown in Figure 5.1 for the X-Y = 6-7 and in Figure 5.2 for all three lengths (X-Y = 5-7, 6-7 and 6-8). For all sequences, two bands are observed, one negative band with minimum at about 245 nm and another intense, positive band at around 265 nm. The positions of these two bands coincide with those reported for parallel G4 conformations.[42] Moreover, the CD spectra do not show significant bands at wavelengths higher than 290 nm, which would be characteristic of both the antiparallel and the hybrid topologies. This is evidence that all the studied sequences (GCGXAGY and GXAGY) form parallel G4s.

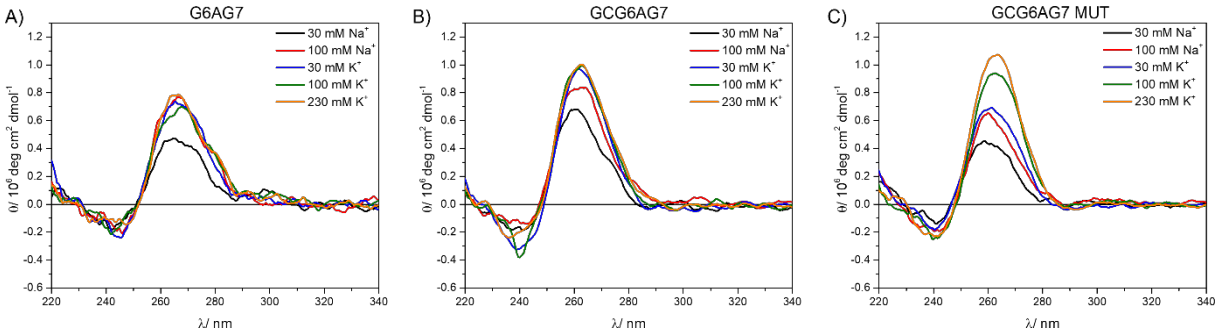



*Figure 5.1. Circular Dichroism spectra for DNA sequences G6AG7, GCG6AG7 and GCG6AG7 MUT in Phosphate buffer with cations $Na^+$ and $K^+$ in the concentration range from 30 mM to 230 mM.*

Even though all studied sequences show CD bands that are coincident with the reported CD bands for parallel G4 conformations, if we compere the spectra of GXAGY sequences and GCGXAGY, subtle differences can be observed. GCGXAGY spectra present higher intensities than their analogous GXAGY sequences in all the studied conditions (compare Figure 5.1A with Figure 5.1B and the upper row with the lower row in Figure 5.2). In addition, GCGXAGY sequences have better defined CD bands, and their maxima show a slight shift to the red with respect to GXAGY (Figure 5.1B and Figure 5.2D-F). These significant differences observed between GXAGY and GCGXAGY CD spectra indicate that the additional upstream guanine actually plays an important role in the G4 formation. This has been further corroborated by the study of a mutant sequence, where some

guanines have been replaced by adenines (GCG6AG7 MUT, see Table 3.3 in section 3.2 of **Experimental methods**). As shown in Figure 5.1C, this mutant sequence presents very similar CD spectra to those of GCG6AG7, with identical positions of the bands and similar intensities. This not only evidences the role the 1 bp upstream guanine has, but also that its presence is more important for G4 formation than the lower number of guanines in the chain. Thus, these observations suggest that the additional guanine is part of the G4 structure. However, the stronger dependency with cation concentration of GCG6AG7 MUT indicates a lower stability of G4s in this mutant sequence in comparison to GCG6AG, which could be caused by the lower presence of guanines. Applying the same argument to the other sequences, we can conclude that all studied sequences of the CSB II form parallel G4s which are favoured when a guanine is present 1 bp upstream.

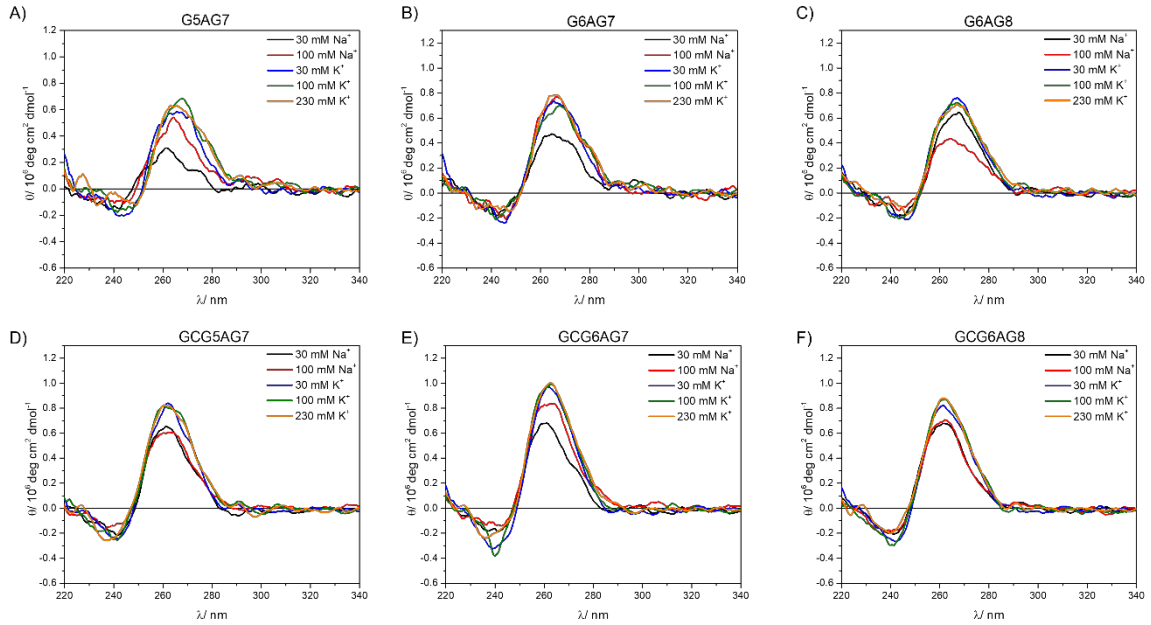



*Figure 5.2. CD spectra for DNA sequences G5AG7, G6AG7, G6AG8, GCG5AG7, GCG6AG7 and GCG6AG8 in Phosphate buffer with cations $Na^+$ and $K^+$ in the concentration range 30-230 mM.*

Additionally, Figure 5.2 shows the influence of the number of guanines: as compared to the longer sequences, the CD bands of the shorter ones (G5AG7 in Figure 5.2A, and GCG5AG7, Figure 5.2D), have less intensity and show negligible spectral shift with increasing cation concentration, suggesting a less effective G4 formation, even when the guanine 1 bp upstream is present. The results of CD melting experiments confirm this interpretation. Figure 5.3 shows the melting data for the three different GCGXAGY sequences fitted using Boltzmann's equation to obtain the melting temperature ($T_m$). The small variation with temperature together with a poorly defined $T_m$ of the shorter sequence (black dots in Figure 5.3 and Table 5.1) confirm its exiguous G4 stability. Instead, the longer sequences show typical melting curves, with identical melting temperatures around 59 °C. This value, in very good agreement with those reported in the

USC DE COMPOSTELA

literature for different G4s of parallel topology[168], confirms the stability of the G4s for the longer DNA sequences. Moreover, melting experiments at higher cation concentrations of these longer sequences GCG6AG7 and GCG6AG8 show no variations of the CD spectrum (data not shown), proving the high stability of the quadruplexes formed by these sequences.

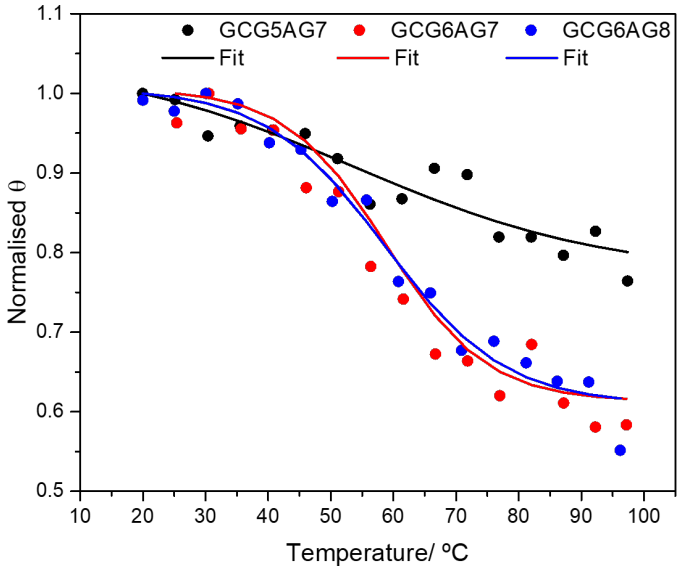



*Figure 5.3. CD melting data: normalised value of the ellipticity Θ at the maximum of the CD spectra as a function of temperature and Boltzmann fits to the experimental data for DNA sequences GCG5AG7 (black dots), GCG6AG7 (red triangles) and GCG6AG8 (blue squares) in Phosphate buffer solution with 30 mM of $K^+$. The corresponding fitted melting temperatures are given in Table 5.1.*

Thus, all these findings lead to the conclusion that the longer sequences, GCG6AG7 and GCG6AG8, which are also more abundant in the CSB II, have a higher tendency to form quadruplexes. In consequence, from now on, we will focus our study on those longer sequences.

*Table 5.1. Melting temperatures obtained in the Boltzmann fits of the CD melting data for the studied DNA sequences.*

| **Sequence** | $T_m$ **/ºC** |
|---|---|
| GCG5AG7 | 54 ± 3 |
| GCG6AG7 | 58.9 ± 0.7 |
| GCG6AG8 | 58.7 ± 0.4 |

### 5.1.1.2. DNA quadruplexes: influence of the cation

CD spectra in Figure 5.1 and Figure 5.2 show that the nature of the cation ($Na^+$ or $K^+$) and its concentration have a significant effect both in the intensity and in the position of the

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

bands. Thus, G4 formation seems to be favoured by higher cation concentrations, and more by the cation $K^+$ than by $Na^+$.

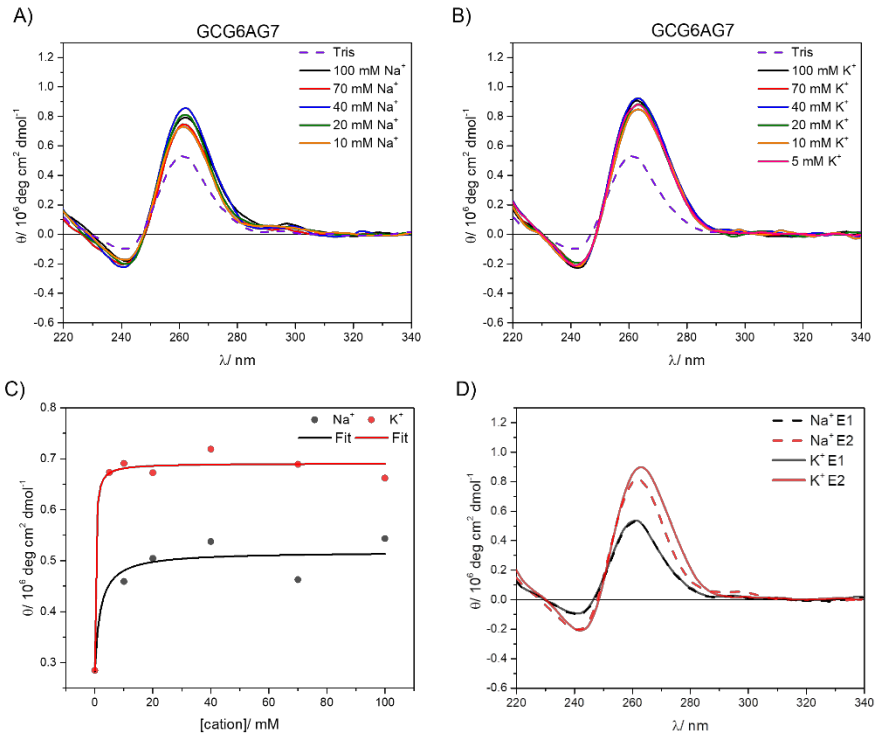



*Figure 5.4. A) CD spectra of GCG6AG7 in Tris buffer at different added sodium concentrations. B) CD spectra of GCG6AG7 in Tris buffer at different added potassium concentrations. C) Experimental data and fitted curves using Hill's model of the intensities at the maxima for the CD spectra shown in A and B. D) Pure CD spectra of GCG6AG7 obtained from the non-linear global fit of Hill's model to the full CD spectra of potassium (continous line) and sodium (dashed line) titrations.*

To gain deeper insight into the influence of the cation concentration on G4 formation, we changed the buffer to one that does not contain sodium nor potassium cations itself, so that samples without these cations and lower cation concentrations could be obtained. Hence, we measured additional CD titrations for the GCG6AG7 sequence in Tris buffer varying the concentrations of sodium or potassium (Figure 5.4A-B). In the absence of sodium or potassium, the CD spectrum of GCG6AG7 shows the two previously observed bands. The intensities of these bands increase significantly with the addition of very low concentrations of the cations, especially with $K^+$. This increase is associated with a slight but noticeable shift of the positive band towards longer wavelengths suggesting a change of topology correlated with the presence of the cations.

In order to extract as much information as possible from the CD titrations, PCA combined with GA analysis were applied. PCA reveals that the experimental CD spectra are contributed by two different species. Using Hill's model (Eq (4.43) in section 4.5.1) to fit the variation of the experimental data (Figure 5.4C) by GA, the "pure" spectra related to those two species are obtained. These spectra are shown in Figure 5.4D with their real

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

intensities and in Figure 5.5A normalized for a better comparison, being E1 (black lines) the CD spectrum of the GCG6AG7 sequence in the absence of cation and E2 (red lines) the CD spectra of GCG6AG7 with sodium (dashed red line) and potassium (solid red line).

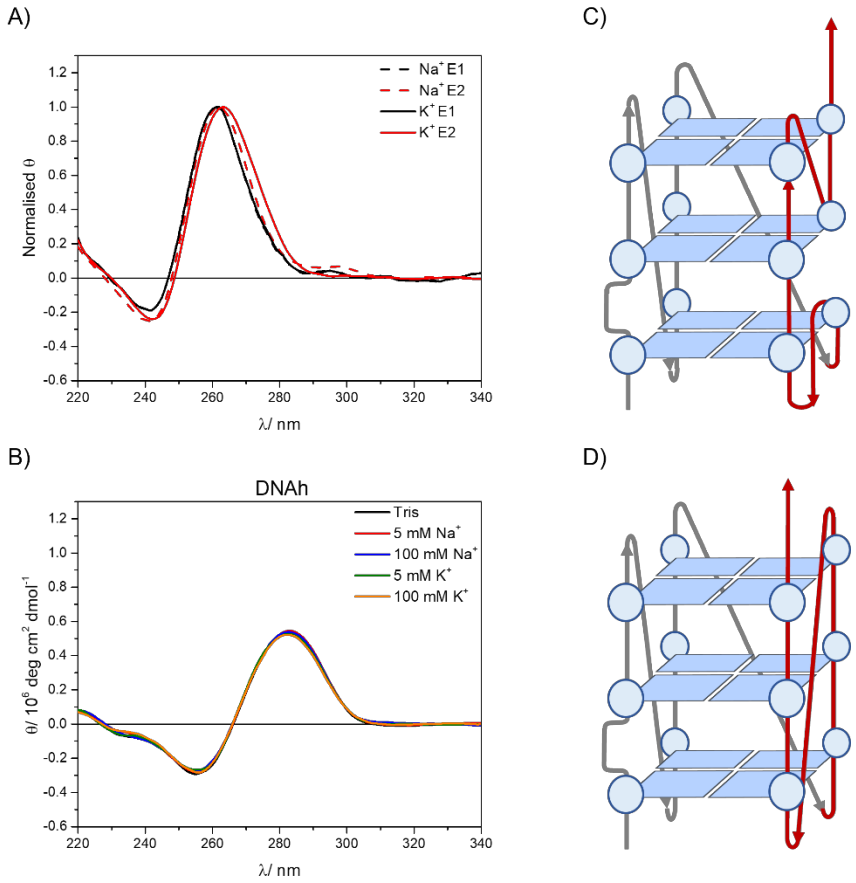



*Figure 5.5. A) Normalised pure spectra obtained from the non-linear global analysis of the CD titration data for the GCG6AG7 sequence with potassium (black lines) and sodium (dashed lines) using Hill's equation. B) CD spectra of DNAh in Tris buffer and different added sodium and potassium concentrations. C) Proposed structure for the less ordered, (-p-p-p+p) parallel G4. D) Proposed structure for the highly ordered, -(ppp) parallel G4.*

The pure spectra resulting from this analysis show that, even in the absence of sodium or potassium cations, the GCG6AG7 sequence forms parallel G4s. Nevertheless, a slight shift of the two CD bands to the red is associated to the addition of the cations, more significantly in the case of potassium, which coincides with that reported for parallel G4s of two topologies with different types of loops.[42] Following the folding topologies described by Karsisiotis and collaborators, we propose that, in the absence of cations, GCG6AG7 forms a less ordered, parallel G4 with looping sequence (-p-p-p+p) (Figure 5.5C), which reorganizes in the presence of the cations, yielding a highly ordered, more stable G4 of topology –(ppp) (Figure 5.5D). The given looping sequences follow the convention of a positive sign (+) for tetrad loops progressing clockwise and negative sign (-) for those progressing counter-clockwise. The letter *p* stands for the type of the loop being a propeller.[42]

The titrations showed that only small concentrations of the cations are needed to bring about the reported G4 rearrangement, especially in the case of potassium. The fits of the Hill equation to the titration data also yield an estimation of the relative stability of the two G4s, based on the value of the rearrangement constant (K), which corresponds to the equilibrium shown in Eq (4.33) of section 4.5.1 backwards, that is, the change from the ordered $G4_f$ to the less ordered $G4_i$. The values obtained for K are 0.34 ± 0.06 mM with potassium and 2.8 ± 0.4 mM with sodium, with a Hill's coefficient of about 1 which shows no cooperativity of the binding process. Hence, the ordered $G4_f$ is better stabilized by potassium than by sodium and needs only a very small cation concentration to form. This behaviour is in accordance with that reported for other G4s.[169,170]

To rule out the potential contribution of other DNA conformations in the sequences under study, CD spectra were measured for a DNA without any guanine-rich sequence (DNAh, Table 1.1) that presumably cannot form G4. Comparing these spectra (Figure 5.5B) with those of guanine-rich sequences (Figure 5.1 and Figure 5.2), the differences between non-G4 and G4 CD bands are clear, being the positions of the bands significantly red shifted for the non-G4 DNA (negative band at about 255 nm and positive band at about 285 nm). Moreover, no change of the CD spectra is observed with cation type and concentration, confirming that the features observed in the CD spectra of guanine-rich sequences are due to the formation of G4 conformations.

#### 5.1.1.3. DNA, RNA and DNA:RNA hybrids quadruplexes

When comparing same length RNA (rGCGXAGY) and DNA (GCGXAGY) sequences under the same cation concentration conditions, we observe similar patterns in the CD spectra (Figure 5.6). This confirms that these RNA sequences also form quadruplexes and with a parallel topology. The positions of the CD bands correspond to the highly ordered G4 which is formed in Phosphate buffer even without addition of salt. However, RNA sequences show lower intensities for all sequences and cation concentrations than DNA, suggesting a lower stability of the RNA-G4s (Figure 5.6). Also, a very low but persistent positive band is observed around 290 nm for all three rGCGXAGY sequences, which is not present in the corresponding DNA sequences and might be due to the formation in a small extent of antiparallel and the hybrid topologies.[42]

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

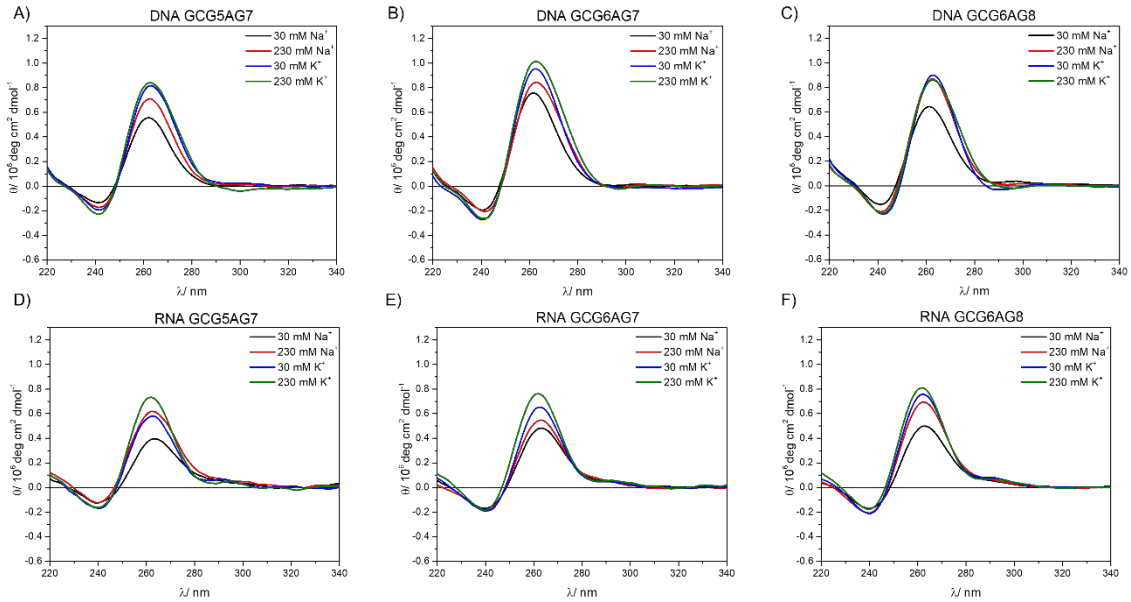



*Figure 5.6. Comparison of the CD spectra of DNA (A-C) and RNA (B-F) sequences in the presence of 30 mM and 230 mM of sodium (black and red) and potassium (blue and green).*

Finally, CD spectra of the DNA:RNA hybrids with the longer sequences, DNA:RNA-hGCG6AG7 and DNA:RNA-hGCG6AG8 (Table 3.3 and section 3.2 of **Experimental methods**), were measured. Raw data showed very broad bands (Figure 5.7A and B) which were obviously contributed by two species: the G4 and the non-G4 DNA that appears due to an incomplete hybridization between DNA and RNA chains. Since the CD spectrum of the non-G4 DNA (DNAh) was known (Figure 5.5B), it was possible to remove this contribution and to restore the CD spectra of the hybrid alone (Figure 5.8). As DNA:RNA hybrid sequences are the result of hybridisation between three different sequences, and some unhybridized molecules could exist in the media, we subtracted from the DNA:RNA hybrid data the contribution of the non-G4 DNA band (DNAh). As shown in Figure 5.7C, the correction of the experimental CD spectra of the hybrid sequences (blue shaded area) by directly deducting the contribution of the DNAh spectra (orange shaded area) results in a parallel G4-like band (green shaded area).

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

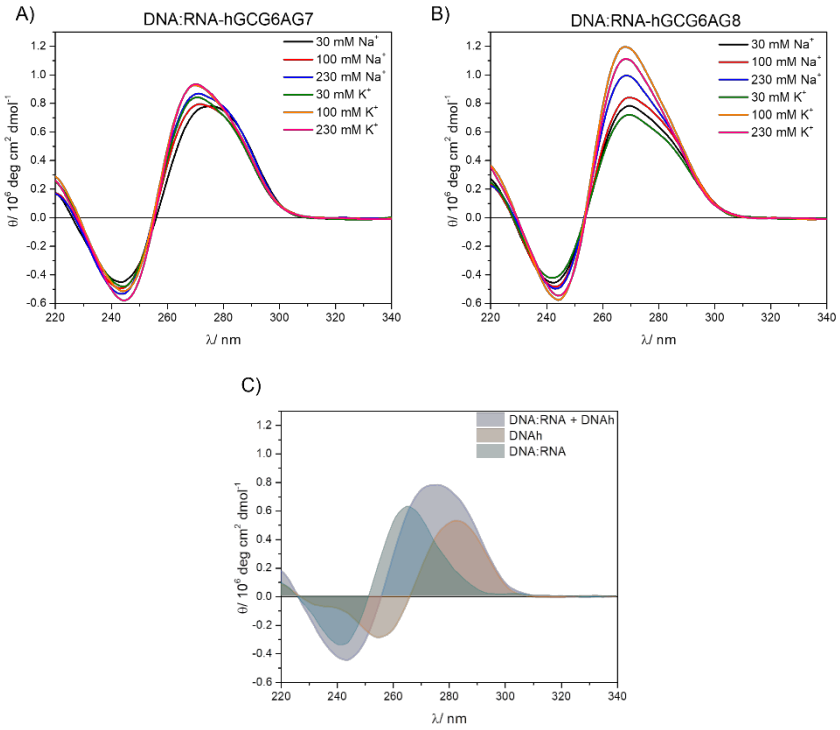



*Figure 5.7. CD experimental spectra of DNA:RNA hybrids at different cation concentrations: A) hGCG6AG7; B) hGCG6AG8; C) Schematic representation of CD hybrid spectra analysis with ssDNA band substraction.*

These CD spectra show similar patterns as the DNA and RNA sequences, confirming the formation of parallel quadruplexes in these hybrid forms with similar stability as for DNA sequences. No further bands are observed at higher wavelengths suggesting the formation of pure parallel G4s. DNA:RNA hybrid G4s formed by telomeric DNA and RNA have been recently reported with parallel conformation and similar melting temperatures both in vitro and in the environmental conditions of HeLa cells.[171]

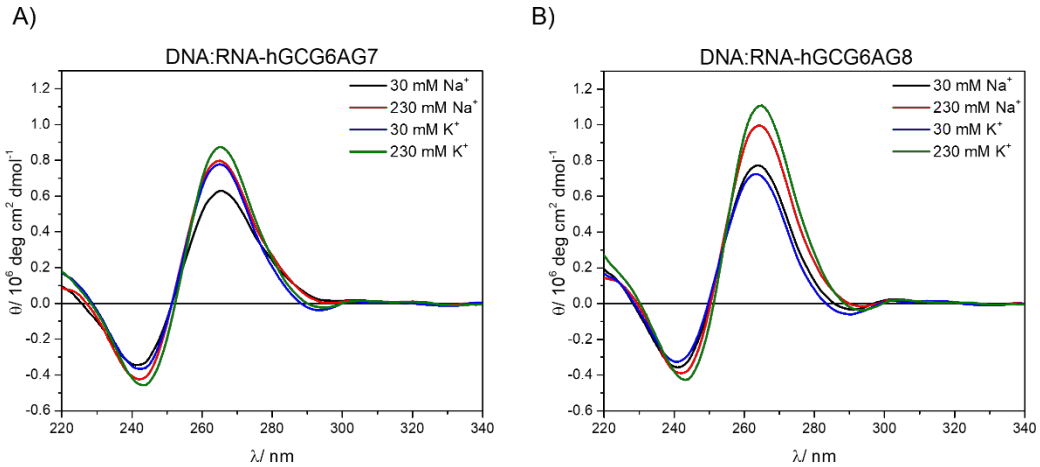



*Figure 5.8. CD spectra of the hybrids DNA:RNA-hGCG6AG7 (A) and DNA:RNA-hGCG6AG8 (B) in Phosphate buffer at different concentrations of sodium and potassium after removing the contribution of non-G4 DNA (see Supplementary Figure SI4).*

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Figure 5.9 summarizes the results obtained from CD measurements, allowing the comparison among the three forms of quadruplexes under study (DNA, RNA and DNA:RNA hybrid) and the three lengths of sequences. The CD spectra show very similar bands in all types and sequences, typical for parallel G4. In the case of DNA:RNA hybrids, the observed slight shift to the red might be due to a somewhat different topology, but it can also be explained by a residual contribution of the non-G4 DNA. Comparison among the three different types of G4s (Figure 5.9D and Figure 5.9E) show that only in the case of RNA there might be present small proportions of antiparallel or hybrid topologies, characterized by a positive band at about 290 nm.

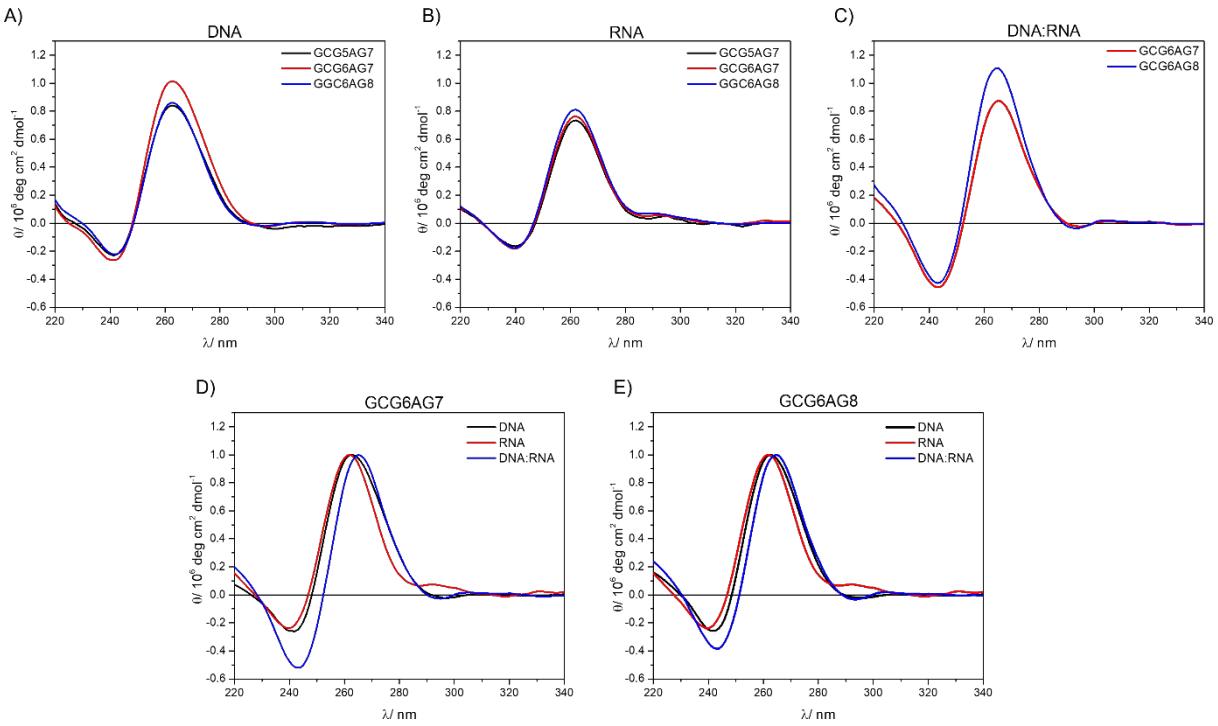



*Figure 5.9. A) CD spectra for different length DNA sequences at 230 mM K⁺; B) CD spectra for different length RNA sequences at 230 mM K⁺; C) CD spectra for different length DNA:RNA hybrids at 230 mM K⁺; D) normalised CD spectra of GCG6AG7, r GCG6AG7 and DNA:RNA-hGCG6AG7; E) normalised CD spectra of GCG6AG8, r GCG6AG8 and DNA:RNA-hGCG6AG8.*

On the basis of these results, we can conclude that all three possible forms of quadruplexes in the CSB II, DNA, RNA and DNA:RNA hybrid, fold into parallel G4s with any sequence's length but more efficiently in GCG6AG7 and GCG6AG8, which are more abundant in the CSB II. The stability of the G4s increases with addition of small cation concentrations, especially potassium.

## 5.1.2. Gel Electrophoresis

Gel electrophoresis experiments were carried out in order to confirm G4 formation. Gels with all DNA sequences (GXAGY, GCGXAGY and GCGXAGY MUT) were prepared using sodium and potassium phosphates (cation concentration of 30 mM) and were developed

USC DE SANTIAGO DE COMPOSTELA

with ThT, a fluorescence probe that is presumed to bind selectively to G4 DNA (Figure 5.10A).[128,148–150] These gels show fluorescent bands for all sequences at similar heights that can be attributed to the G4s. However, the GXAGY sequences, without a guanine 1 bp upstream, present much lower intensities, in line with their lower efficiency of quadruplex formation. In the case of GCGXAGY mutants additional upper bands are observed with very short migration, which can be attributed to the formation of intermolecular G4 structures. Thus, the fact that all studied sequences can fold into G4 structures is reinforced as well as the important influence of the first guanine in the G4 formation.

We have also carried out gel electrophoresis for RNA sequences (rGCGXAGY) and DNA:RNA hybrids (DNA:RNA-hGCGXAGY), as well as for the non-G4 sequence DNAh. Figure 8B shows these gels together with those of the corresponding DNA sequences for comparison. It is observed that RNA sequences do not show clear bands, what is attributed to the lack of RNAase free conditions. Nevertheless, the wide bands of higher intensity present similar migration as the corresponding DNA sequences (Figure 5.10B, left).

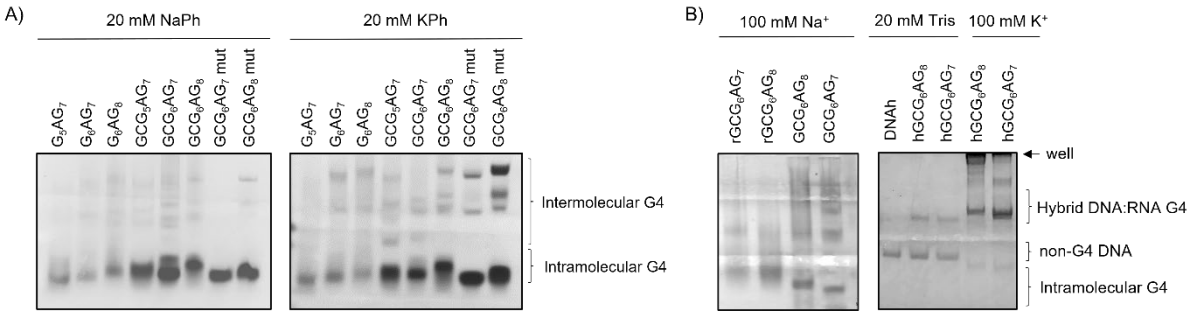



*Figure 5.10. A) Gel electrophoresis developed with ThT for different DNA sequences in 20 mM sodium phosphate and 20 mM potassium phosphate (cation concentrations are 30 mM); B) Gel electrophoresis developed with ThT for RNA (rGCGXAGY) and DNA (GCGXAGY) sequences in 100 mM $Na^+$ and for DNAh and DNA:RNA (DNA:RNA-hGCGXAGY) sequences in 20 mM Tris (no cation) and 100 mM $K^+$.*

DNA:RNA hybrid sequences show two different bands whose intensity change with the addition of potassium cations, being the upper bands much more intense in presence of potassium. Comparing the positions of the DNA:RNA-hGCGXAGY bands and their retention factor (Rf, Table 5.2) with those of the corresponding DNA sequences, we can conclude that the upper bands that increase intensity when adding salt correspond to the G4s formed after the hybridisation of the three strands, which show much lower Rf than the corresponding DNA sequences due to their higher mass. The lower band, which is much more intense in the absence of salt, presents identical migration as the non-G4 DNAh and is therefore attributed to the residual DNAh strands from the incomplete hybridisation.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

In Figure 5.10B we can see that all sequences present bands even though we are staining with ThT, which has been reported as an optimum probe for G4 recognition.[128,148–150] However, our results show that ThT binds also to some extent to non-G4 sequences, although the emission intensity is significantly lower than with G4 sequences. These results are also supported by other previous fluorescence studies of the fluorophore.[115,126,127,172,173]

*Table 5.2. Obtained retention factors (Rf), in percentage, for each sequence used in gel electrophoresis (Figure 5.10).*

| Sequence | Rf/ % | Sequence | Rf/ % |
|---|---|---|---|
| **GCG6AG7** | 79 | **DNA:RNA-hGCG6AG7 Tris** | 37 / 64 |
| **GCG6AG8** | 77 | **DNA:RNA-hGCG6AG8 Tris** | 37 / 64 |
| **DNAh** | 58 | **DNA:RNA-hGCG6AG7 $K^+$** | 33 |
| | | **DNA:RNA-hGCG6AG8 $K^+$** | 33 |

## 5.1.3. Fluorescence measurements

In order to obtain further insight into the influence of the cation concentration in the G4 formation as well as to prove the binding of ThT to non-G4 structures, fluorescence measurements were performed. The advantage of ThT is its negligible emission in aqueous solution in contrast to its high brightness when bound to proteins or other macromolecules. However, its complex photophysical behaviour may complicate the interpretation of the observed fluorescence properties.[112]

Steady-state and time-resolved fluorescence emission and anisotropy of ThT in presence of DNA GCG6AG7 were measured as a function of potassium concentration and compared with those of ThT in presence of DNAh with and without 100 mM of potassium. For all these experiments, perchlorate salts were used instead of chlorides to prevent any possible quenching of the fluorescence. As pointed out above, we can assume that the observed fluorescence is only due to ThT bound to the DNA. We will denominate ThT bound to GCG6AG7 also as ThT:GCG6AG7 complex and ThT bound to DNAh as ThT:DNAh complex.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

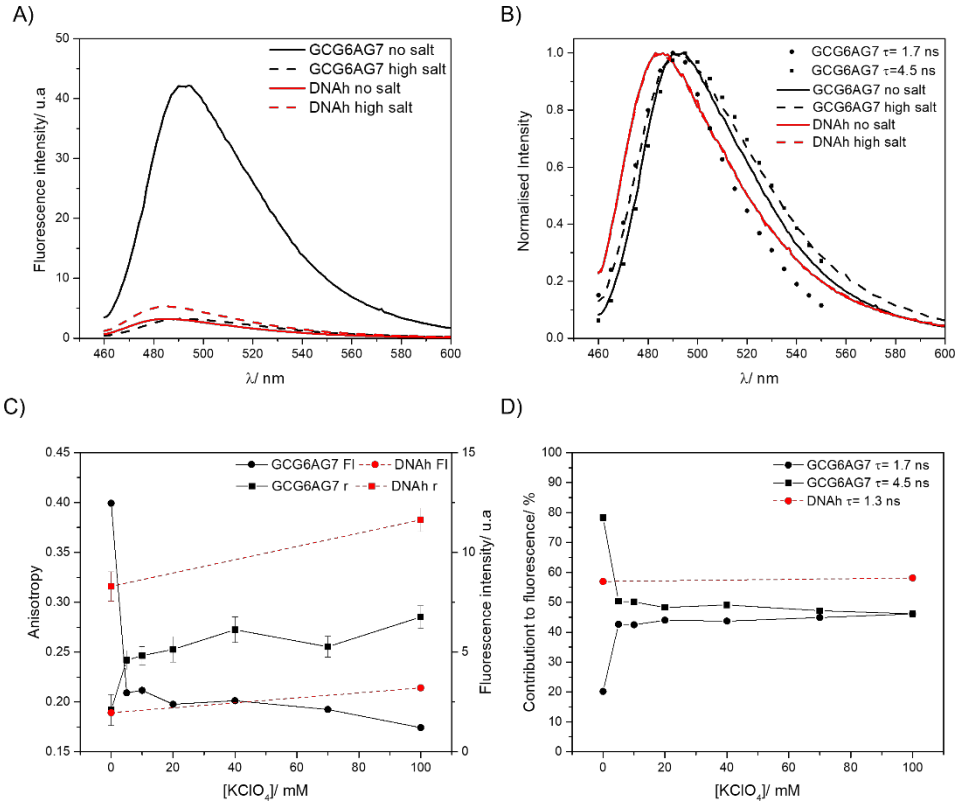



*Figure 5.11. A) Fluorescence emission spectra of ThT with GCG6AG7 and DNAh sequences in Tris buffer in the absence of cations (solid lines) and with 100 mM potassium (dashed lines); B) Normalized fluorescence emission spectra of graph A compared with the time-resolved emission spectra of the ThT species with τ = 1.7 ns and τ = 4.5 ns ; C) Steady-state fluorescence anisotropy (squares, left scale) and fluorescence intensity (circles, right scale) of ThT with GCG6AG7 (black symbols) and with DNAh (red symbols) sequences in Tris buffer with different concentrations of potassium. D) Percentage contributions to fluorescence of the two ThT species observed with GCG6AG7 (black symbols) and the one with DNAh (red symbols) in Tris buffer with different concentrations of potassium.*

Fluorescence emission spectra are shown as absolute intensities in Figure 5.11A and normalized in Figure 5.11B. In the absence of the cation, ThT bound to GCG6AG7 has an about 10-fold higher intensity than ThT bound to the non-G4 DNAh. The addition of potassium causes a sharp decrease of the fluorescence intensity of ThT bound to GCG6AG7, whereas in the case of ThT bound to DNAh it increases slightly (Figure 5.11A and Figure 5.11C). The normalized spectra of the emission spectra of ThT bound to the non-G4 DNAh (Figure 5.11B) show a considerable blue shift as compared to those of ThT bound to the G4. These findings confirm that ThT is not specific for G4s but unveil the distinct fluorescence properties of ThT when bound to GCG6AG7 or to DNAh.

Time-resolved fluorescence measurements allow us to identify the number and type of ThT species responsible for the observed emission. When bound to GCG6AG7, ThT shows two main lifetimes which correspond to two different species of the ThT:GCG6AG7 complex (Table 5.3). The one with the longer lifetime of about 4.5 ns is the main species in the absence of cation, explaining 80% of the total fluorescence (black squares in Figure 5.11D), and its contribution decreases sharply to about 50% with the addition of potassium. On the contrary, the ThT:GCG6AG7 complex with the shorter lifetime of

1.7 ns contributes only 20% without potassium and around 45% in the presence of the cation (black circles in Figure 5.11D).

*Table 5.3. Fluorescence lifetimes (τ) and rotational correlation times (ρ) of ThT bound to GCG6AG7 and ThT bound to DNAh obtained by global analysis of the time-resolved fluorescence and anisotropy decays, respectively. The lifetime $\tau_1$ is a very short lifetime with low contribution to the observed fluorescence which is attributed to a small proportion of ThT molecules undergoing TICT.*[112]

| ThT | $\tau_1$/ns | $\tau_2$/ns | $\tau_3$/ns | $\rho$ (no $K^+$)/ns | $\rho$ (with $K^+$)/ns |
|---|---|---|---|---|---|
| GCG6AG7 | 0.40 ± 0.05 | 1.7 ± 0.2 | 4.5 ± 0.3 | 3.20 ± 0.07 | 5.6 ± 0.2 |
| DNAh | 0.39 ± 0.09 | 1.3 ± 0.2 | - | 3.6 ± 0.2 | 3.6 ± 0.2 |

ThT presents a wide distribution of ground-state conformations with different fluorescence properties, going from those with nearly perpendicular benzothiazole and aminobenzene moieties (Figure 5.12, left), characterized by blue-shifted fluorescence, low quantum yield and a short lifetime, to configurations which are almost planar (Figure 5.12, right), with red-shifted fluorescence, high quantum yield and long lifetime.[112] On this basis, the ThT bound to GCG6AG7 with the longer lifetime, predominant in the absence of cation, corresponds to a virtually planar conformation, explaining the high fluorescence intensity under such conditions (Figure 5.11A and Figure 5.11C). Instead, the ThT bound to GCG6AG7 with the shorter lifetime presents some degree of torsion and is responsible for the lower, blue-shifted fluorescence observed in the presence of potassium (Figure 5.11A-C). The time-resolved fluorescence spectra obtained for these two types of bound ThT (black symbols in Figure 5.11C) are in agreement with this explanation.

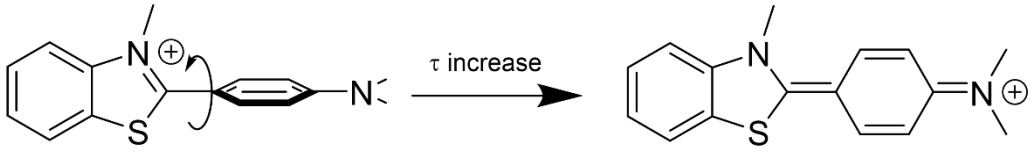



*Figure 5.12. Schematic representation of the two limit conformations adopted by ThT.*[112]

Further insight into the conformations of ThT bound to GCG6AG7 is provided by fluorescence anisotropy measurements, which give information about the mobility of the fluorescent probe within its microenvironment. The steady-state fluorescence anisotropy value (Figure 5.11C) of ThT bound to GCG6AG7 increases significantly as potassium concentration is increased, indicating a decrease in the mobility of ThT when the cation is added. This can be explained on the basis of the rotational correlation times obtained in the time-resolved anisotropy measurements (Table 5.3). In the absence of potassium, most ThT molecules live enough in the excited state (4.5 ns) to be able to rotate to some extent with the corresponding rotational time of 3.20 ns. Instead, in the

USC
DE SANTIAGO
DE COMPOSTELA

presence of potassium the rotational correlation time is higher (5.6 ns) and ThT molecules do not stay enough time in the excited state to rotate and depolarize.

Nevertheless, the most important result of these fluorescence data is the confirmation of the rearrangement of the DNA GCG6AG7 quadruplex in the presence of low amounts of cation, as observed in CD measurements. The sharp variations of the fluorescence intensity, steady-state anisotropy, and lifetimes' contributions (Figure 5.11C and Figure 5.11D) with the addition of potassium support this hypothesis, showing that ThT must adapt to the topology of the G4.

Finally, fluorescence measurements show that ThT bound to DNAh presents very different fluorescence properties: much lower fluorescence intensity which increases slightly when potassium is added (Figure 5.11A and Figure 5.11C), significant blue-shift of the emission spectrum (Figure 5.11B), shorter lifetime (Table 5.3) and high anisotropy due to the ineffective rotation during the lifetime (Figure 5.11C and Table 5.3). These results prove a higher degree of torsion between the moieties of ThT when bound to a non-G4 DNA (Figure 5.12) due to the different microenvironment of the probe.

Overall, the detailed analysis shown in this study evidences the formation of G4 structures with parallel topology for the three G-rich main sequences within mitochondrial CSB II, those with a higher number of guanines and a guanine located 1 bp upstream being more stable. We observed a rearrangement of the G4 topology to a highly ordered structure with the addition of small amounts of cation, especially potassium. All three potential G4s forms within CSB II: RNA G4 at the RNA transcript, DNA G4 within the non-transcribed strand and DNA:RNA hybrid between the RNA transcript and the non-transcribed strand. These findings contribute to a better knowledge of the G4 conformations found in mitochondria, which are thought to have a relevant role in transcription termination and R-loop stabilization.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# 5.2. Critical aggregation concentration and reversibility of Aβ(1-40) oligomers

**This section is a reproduction of a submitted research.**

Here we present the study of amyloid (1-40) aggregation in physiological conditions, which was carried out using FCS (see section 3.1.2.1). We also present a comparison between our recent studies and the previously reported work for amyloid (1-42).

## 5.2.1. FCS Power series

As explained in the **Experimental methods** section 3.1.2.1, photobleaching can interfere significantly in the measured diffusion times, so optimization of the measurement conditions is needed. For this reason, FCS power series of both monomeric and aggregated samples of Aβ40 were measured, as shown in Figure 5.13, panel A.

The registered fluorescence intensity *I* and, consequently, the brightness (fluorescent counts per molecule, *cpm* = *I*/*N*) of the monomeric Aβ40* samples increase linearly with the laser power up to an irradiance of about 50 kW $cm^{-2}$ (laser power of 60 μW) but deviate at higher power due to the effects of photobleaching and optical saturation (Figure 5.13A and C). Also, at high irradiances the FCS curves shift towards lower correlation times (Figure 5.13A), with the corresponding decrease of the diffusion time due to photobleaching. Thus, monomeric samples could be excited up to 50 kW $cm^{-2}$. However, addition of unlabelled Aβ40 leads to the formation of oligomers that stay longer in the focus because of their bigger sizes, leading to increased photobleaching. For oligomeric samples (Figure 5.13C, Aβ40* + Aβ40), linearity is already lost above 30 kW $cm^{-2}$. Therefore, to avoid photobleaching and to obtain correct values of the diffusion times, all samples were measured with an irradiance of 24 kW $cm^{-2}$ (laser power of 30μW).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

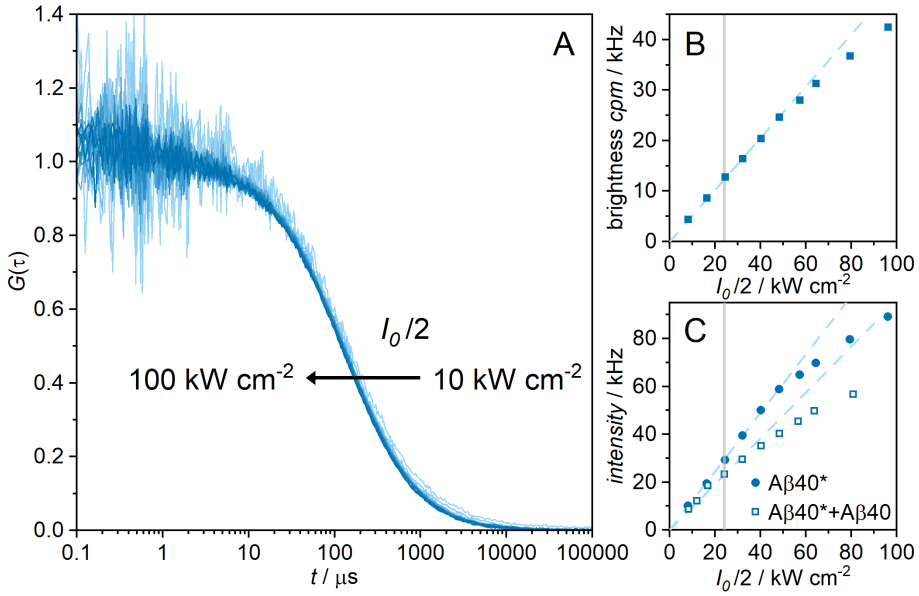



*Figure 5.13. FCS power series. Panel A: normalized fluorescence correlation curves of an Aβ40* sample acquired at different laser irradiances. Panel B: fluorescence brightness expressed as photon counts per molecule and second, cpm, determined from the curves in panel a. Panel C: fluorescence intensity determined from the curves in panel A (filled circles) and from measurements with added unlabelled Aβ40 (open squares). The grey vertical lines indicate the excitation irradiance of 24 kW cm$^{-2}$ chosen for all further measurements.*

## 5.2.2. Monomeric Aβ40 characterization

First, we characterized the monomeric Aβ40 to confirm that the starting peptide samples were free of aggregates. FCS measurements of samples of low concentrations of labelled amyloid (Aβ40*) fit well with a single diffusion term (Eq (20)), yielding a diffusion time of 151 ± 2 µs, and a value of the diffusion coefficient of monomeric Aβ40 of (1.30 ± 0.14) × 10$^{-10}$ m$^2$ s$^{-1}$ at 25°C. This value is consistent with that obtained for monomeric Aβ42[68], (1.27 ± 0.09) × 10$^{-10}$ m$^2$ s$^{-1}$ at 25°C converted to the R123 reference value used in this work. Both values follow the relative molar mass dependency $D \sim M^{-v}$ with a value of $v$ = 0.44 as determined by Danielsson *et al.* [174] for a series of amyloid peptides using PFG-NMR.

Comparing the absolute diffusion coefficients, our values of labelled monomeric amyloids are approximately 10% lower than those reported by Danielsson *et al.*[174], a difference within the uncertainty range primarily due to the uncertainty of the reference value. This discrepancy may also arise from a potential change in the scaling factor $v$ in the labelled amyloids due to the fluorophore or to the presence of small amounts of dimers in the samples. A dimer would have a diffusion coefficient at least $2^{0.33}$ = 1.26 times lower than that of the monomer, assuming a compact sphere – other shapes would lead to even higher factors.[174] Thus, Aβ40 dimers or higher assemblies would correspond to significantly smaller diffusion coefficients than those observed. We therefore conclude that our starting amyloids are indeed predominantly monomeric.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Using the Stokes-Einstein equation, we obtained a hydrodynamic radius for the Aβ40 monomer of $R_h$ = 1.88 ± 0.20 nm, which coincides within its error with that reported by our group for Aβ42 (1.93 ± 0.13 nm)[68] and with literature values found for Aβ40 monomers (ranging from 1 nm to 2 nm).[62,175,176,66]

## 5.2.3. Aggregation of Aβ – Degree of Aggregation

To determine the degree of aggregation, we followed the data analysis procedure described in the Section 4.5.3.7 of **Data Analysis**.

The first step involved fitting the FCS curves (Section 4.5.3.7of **Data Analysis**, Step 1) obtained for solutions with varying concentrations of unlabelled amyloid and a constant single-molecule concentration of labelled amyloid. For this Eq (4.21) was globally fitted with a shared monomer diffusion time ($\tau_{D1}$), and free diffusion times of the oligomers ($\tau_{D2}$), relative contributions (R) and the total number of observed molecules (N).

### 5.2.3.1. Time dependency: short incubation times

To study the potential influence of the incubation time, FCS curves were recorded up to 180 minutes after sample preparation. For each sample, series of FCS curves were measured covering the entire observation time (Figure 5.14a) and fitted with the two species model (Eq (4.21)) with a monomer diffusion time shared among all curves. We observed no systematic variation in the diffusion times of the aggregates (Figure 5.14b) or in the relative contributions of the species (Figure 5.14c), indicating that the aggregates do not grow or change during the studied time. This finding is consistent with previous observations for Aβ42.[68]

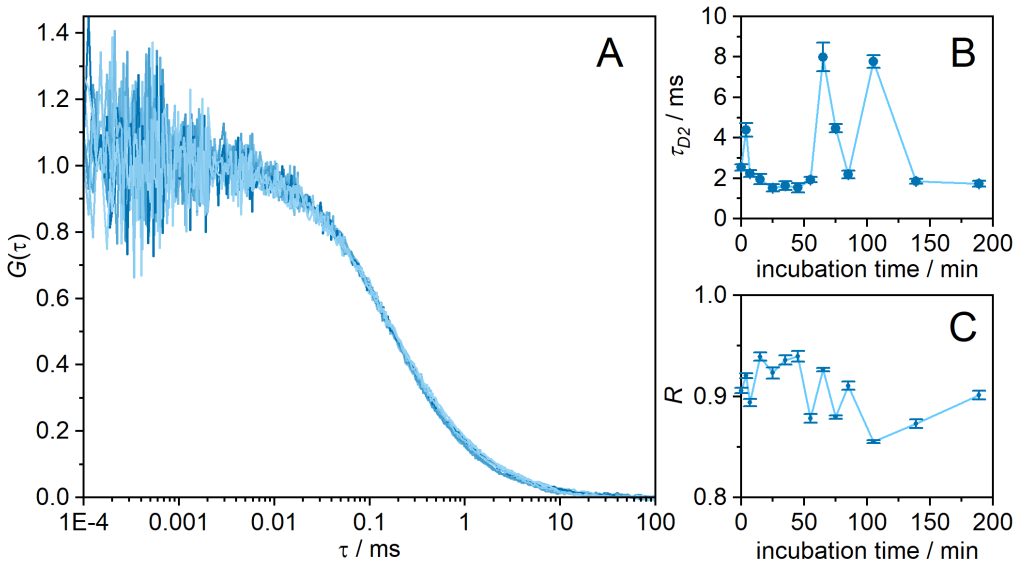



*Figure 5.14. Panel A: Normalised FCS curves of a given amyloid sample (nominal total amyloid concentration 4 μM) recorded during 180 minutes at the incubation times indicated in panels B and C. Panel B: Diffusion times of the aggregates, $\tau_{D2}$, obtained from global fits of eq. SI2 to the FCS curves shown in Panel A as a function of incubation time.*

*The monomer diffusion time, $\tau_{D1}$, was fixed to the value obtained for monomeric Aβ40. Panel C: Factor R representing the relative contribution of monomeric amyloids to the FCS curves obtained from each of the curves shown in Panel A over the incubation time. The FCS curves in panel A were normalized using the value of N obtained from the fits.*

### 5.2.3.2. Brightness ratio

The fit results show that the brightness of labelled amyloid is lower in aggregates ($Q_2$) than in monomers ($Q_1$). For Aβ40* we determined a brightness ratio $q' = Q_2/Q_1 = 0.53 \pm 0.09$ (Section 4.5.3.7of **Data Analysis**, Step 2), which coincides with the value of 0.6 ± 0.1 obtained for Aβ42.[68]

### 5.2.3.3. Aggregation number of the oligomers

From the brightness ratio $q'$ and the monomer brightness $Q_1$ we obtained the mean number of labelled Aβ per aggregate $\bar{i}_u^*$ and a preliminary uncorrected mean aggregation number $\bar{n}_u$ of each experiment (Section 4.5.3.7 of **Data Analysis**, Step 3). The double-logarithmic plot of the diffusion time ratios $\tau_{D2}/\tau_{D1}$ versus $\bar{n}_u$ (Figure 5.15) yields a scaling factor of the aggregates $v_2 = 0.71 \pm 0.07$ (Section 4.5.3.7 of **Data Analysis**, Step 4). This relatively high value indicates a shift from the random coil conformation of the monomer to a more elongated, fibril-like conformation of the early aggregates. Such a conformation was already observed in the case of Aβ42 ($v = 0.75$)[68] and is in good agreement with the reported spherocylindrical geometry of Aβ40 micellar assemblies determined from small angle neutron scattering.[69]

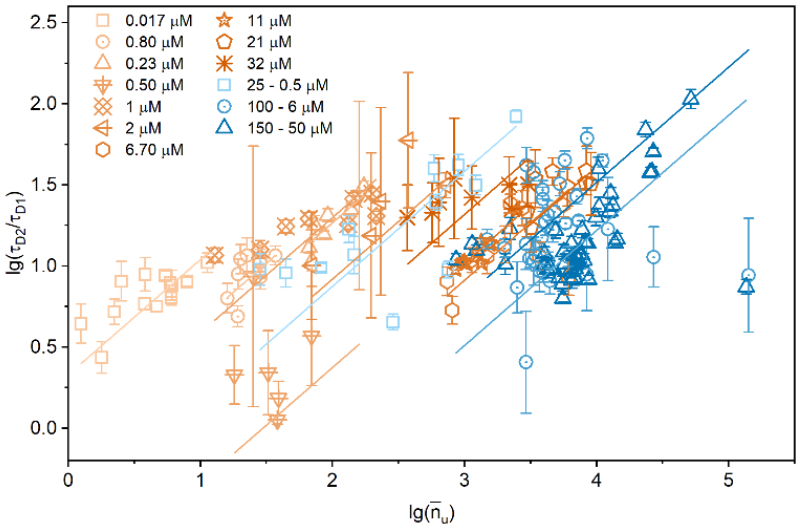



*Figure 5.15. Double logarithmic plot of the diffusion times ratio ($\tau_{D2}/\tau_{D1}$) of Aβ40 versus the uncorrected aggregation numbers from repetitive measurements of samples of different amyloid concentrations (nominal concentrations are indicated). Straight lines are the results of the global linear fit with free intercepts but a shared slope of $v_2 = 0.71 \pm 0.07$.*

Using the results of the previous step, we calculate a new corrected aggregation number $\bar{n}_c$ for each measurement (Section 4.5.3.7of **Data Analysis**, Step 5). Figure 5.16 shows the corrected mean aggregation number as a function of the nominal total Aβ40 concentration. No systematic variation of the aggregation numbers with the incubation

time or the total amyloid concentration is observed, consistent with the findings for Aβ42.[68] Mean aggregation numbers range between approximately 10 and 200, although some values fall outside this range with high uncertainties, especially in the dilution series. We further analyse the size of the aggregates below.

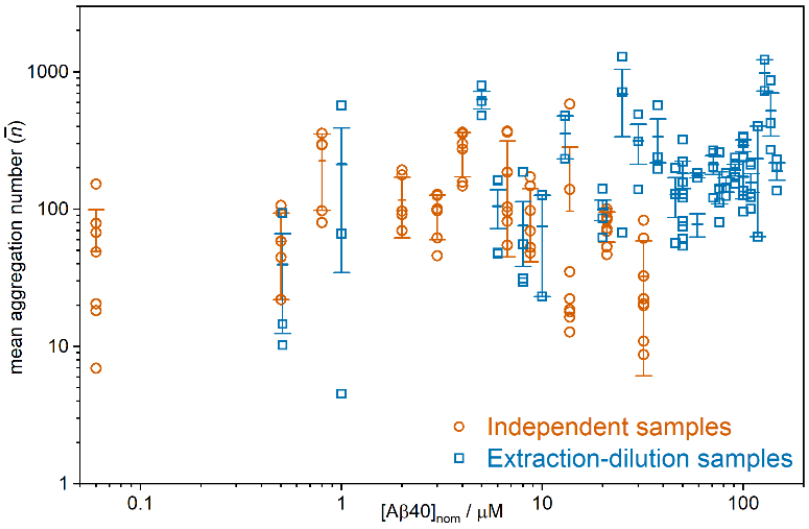



*Figure 5.16. Double logarithmic plot of the corrected mean aggregation numbers of the aggregates as a function of the nominal total amyloid concentration obtained from independent samples (orange circles) and extraction-dilution samples used for reversibility studies (blue squares). The coloured horizontal bars represent the mean value of each sample. The error bars correspond to the standard error of the mean value of each sample.*

### 5.2.3.4. Degree of aggregation

Next, we analyse the fraction of the total amyloid that is aggregated (degree of aggregation *γ*) as a function of the total true Aβ40 concentration. The procedure is described in Steps 5 and 6 of the Section 4.5.3.7 of **Data Analysis**. Due to the high tendency of amyloids to adsorb to surfaces, determining the true amyloid concentration is crucial in this analysis, although for Aβ40 a much better correlation between nominal and true concentrations than for Aβ42 is observed (Figure 5.17).[68]

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

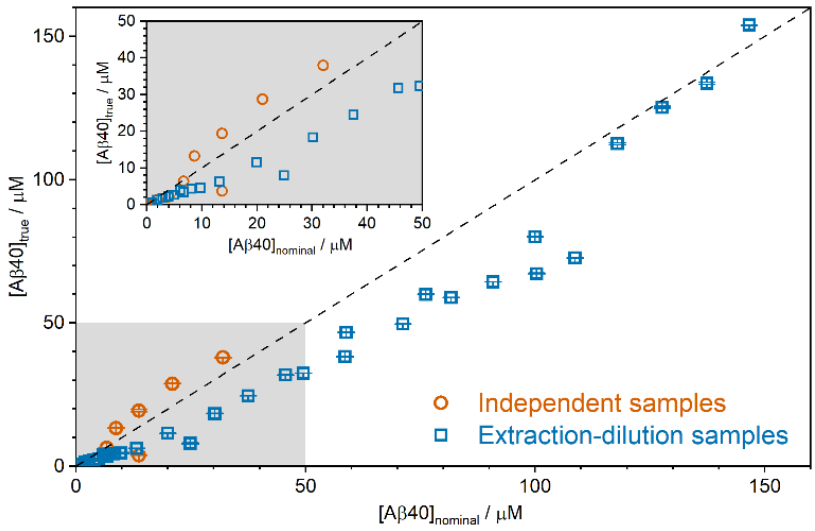



*Figure 5.17. True amyloid concentration calculated from FCS curves versus the nominal amyloid concentration for independent samples (orange circles) and extraction-dilution samples (blue squares). Inset: lower sample concentration range. The dashed lines indicate a 1:1 correspondence between nominal and true concentrations.*

Figure 5.18 shows the degree of aggregation determined from individually prepared samples as a function of the total true Aβ40 concentrations (depicted by orange circles in Figure 5.18, nominal concentrations ranging from 0.005 µM to 43 µM). The fraction of aggregated amyloid depends strongly on the total Aβ40 concentration, revealing three distinct regions. Below 0.1 µM minimal aggregation is observed. Between 0.1 µM and 10 µM, the degree of aggregation rises, reaching up to 10% of the total Aβ40. Above 10 µM, the degree of aggregation increases markedly.

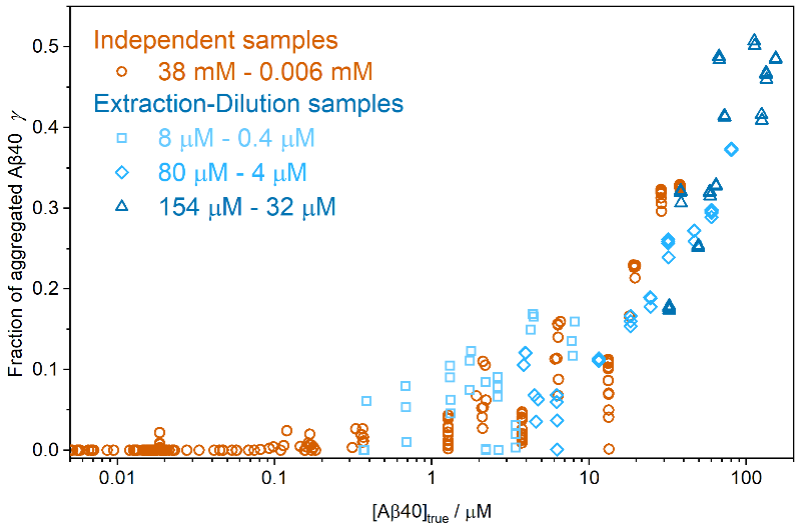



*Figure 5.18. Degree of aggregation γ of Aβ40 samples, determined using the procedure outlined in the section 4.5.3.7 of the Data Analysis, plotted against the true total amyloid concentration in solution. Independently prepared samples are compared with series of progressively diluted samples covering the indicated nominal concentration intervals.*

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

### 5.2.3.5. Reversibility and long-term stability

To analyse the reversibility of Aβ40 aggregation, dilution series covering the full concentration range were prepared. Figure 5.18 shows the results of the degree of aggregation of these samples as a function of the true total Aβ40 concentration, together with those obtained for the independently prepared samples for comparison. It can be seen that the degree of aggregation decreases with the dilution of Aβ40, with values that overlap among the dilution series. Moreover, the degree of aggregation of samples diluted from higher concentrations decrease to the same value obtained for samples directly prepared from lower Aβ40 concentrations, hinting at an effective disaggregation of the Aβ40 oligomers. This reduction of the degree of aggregation upon dilution is accompanied by the expected increase of the relative contribution *R* of the monomers to the amplitude of the FCS curves (see Figure 5.19). These findings suggest that the Aβ40 aggregation is reversible within the studied concentration interval.

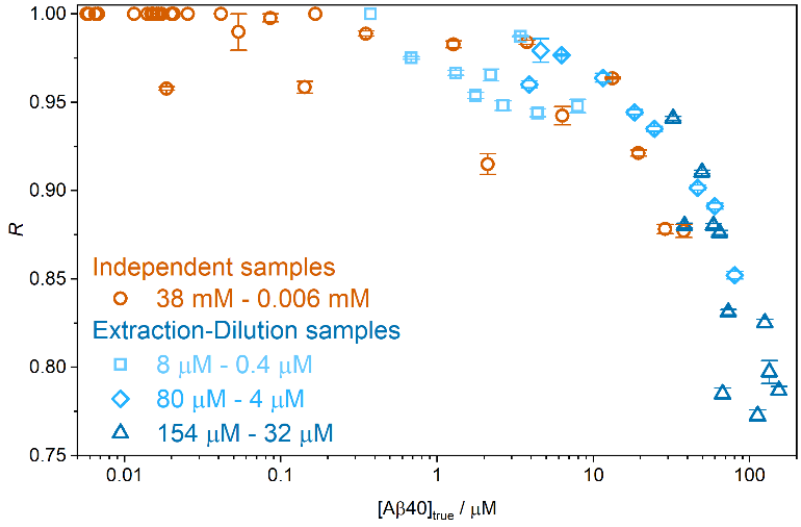



*Figure 5.19. Relative contribution R of monomeric amyloid to the full amplitude of the diffusion term as determined from fits to FCS curves of independent samples and series of successively diluted samples of Aβ40.*

The dilution series also provide data for the degree of aggregation at high Aβ40 concentrations, not reached in the individually prepared samples. As shown in Figure 5.18, at concentrations higher than 10 μM, the degree of aggregation increases strongly from values of $\gamma \approx 0.05$ - 0.1 to around 0.5, corresponding to roughly 50% of aggregated Aβ40 above 100 μM. A very good agreement is observed between the different types of samples.

Further studies were performed to check the stability of the samples over extended incubation periods. Figure 5.20 compares data from independently prepared samples measured during 2-3 hours after preparation and again after one week. No significant changes were observed in the diffusion times of the aggregates (Figure 5.20A and B) or in the relative contribution *R* of the monomers to the FCS curves (Figure 5.20C and D),

except that for low concentrations, *R* values were closer to 1 after one week. This finding suggests that there is a kinetic barrier for the disaggregation of the oligomers, indicating that reversibility would require longer times to be completed. This could also explain the highly dispersed data in the dilution data in the range 8 – 0.4 µM (Figure 5.18, blue squares).

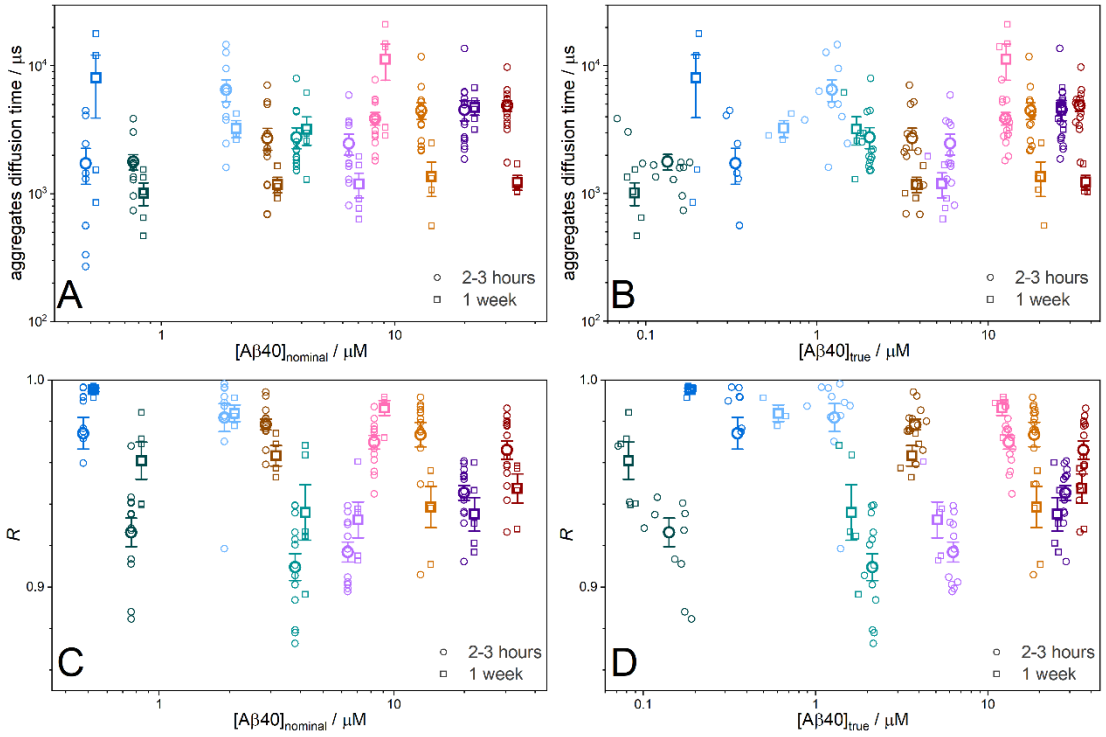



*Figure 5.20. Comparison of data of independently prepared samples measured during 2-3 hours after preparation (circles) and after one week (squares) versus the nominal (left, A, C) and the true (right, B, D) total amyloid concentration. Panels A and B show the aggregates' diffusion times and panels C and D the relative contribution of monomers to the FCS curves. The bigger symbols represent the mean value of each sample. The error bars correspond to the standard deviations of the individual values. In panels A and C the values of the same sample were slightly shifted below and above the corresponding nominal concentration to improve their visibility in the plot.*

The true total amyloid concentration remained constant after one week, showing no precipitation of the aggregates, which would be expected for assemblies larger than the soluble oligomers. However, a much higher dispersion of the data at long incubation times is observed, as indicated by the large uncertainties in the relative contribution *R* against the true total amyloid concentration (Figure 5.20D), suggesting an increase in polydispersity.

### 5.2.3.6. Critical aggregation concentrations

Combining all *γ*-values from independent samples and dilution series (Figure 5.18), and calculating the mean values and uncertainties with all measurements of each sample, we obtain the data shown in Figure 5.21 which will allow us to obtain the critical aggregation concentration (Section 4.5.3.7 of **Data Analysis**, Step 7). The observed

USC DE SANTIAGO DE COMPOSTELA

variation in the degree of aggregation with the true total amyloid concentration suggests two cooperative aggregation processes at two critical aggregation concentrations. To address this, we extended the aggregation concentration model originally developed for the self-assembly of surfactants[167] to include two independent aggregations with their corresponding *cac* values, as described in the Section 4.5.3.7of **Data Analysis**.

The continuous line in Figure 5.21 represents the result of the weighted nonlinear fit of Eq (4.82) to the experimental data. The double aggregation model fits the data well, indicating a first aggregation at $cac_1$ = 0.5 ± 0.3 μM and a second at $cac_2$ = 18.9 ± 2.2 μM. The corresponding transition widths ($\sigma_i = r_i \cdot cac_i$) of $\sigma_1$ = 0.2 μM and $\sigma_2$ = 5.7 μM, respectively, indicate that under the studied conditions the transition from monomeric amyloid to the first type of aggregate takes place between 0.4 - 0.7 μM, while the transition from this first oligomeric amyloid to the next one occurs in a concentration range from 13 μM to 25 μM. The fit curve also indicates that the first aggregation leads to less of 10% of aggregated amyloid and to around 50% well over the second *cac*, indicating that a significant fraction of Aβ40 remains monomeric even at very high amyloid concentrations.

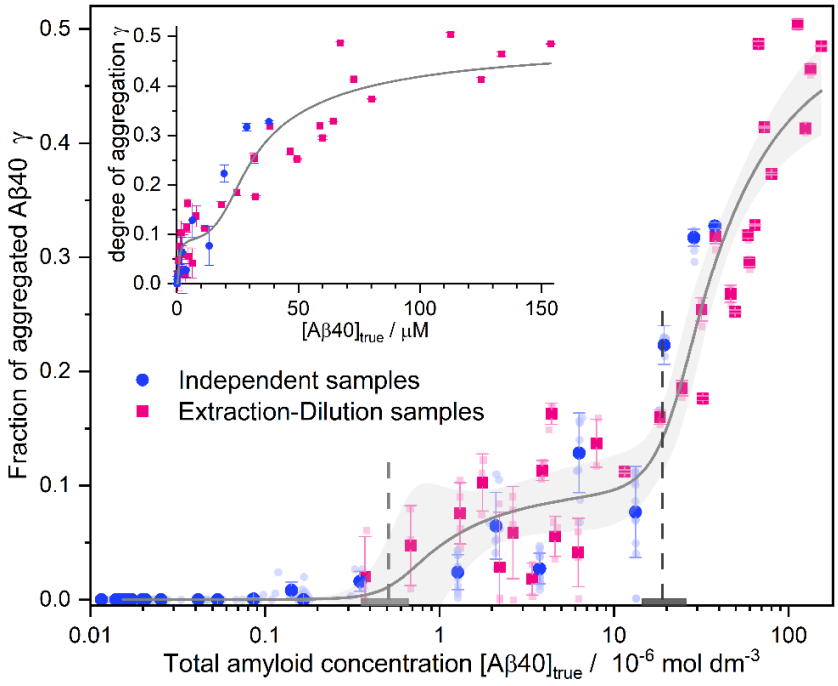



*Figure 5.21. Degree of aggregation of Aβ40, γ, as a function of the true total amyloid concentration in solution. The filled symbols are the mean of the γ-values shown in Figure 6.6. The error bars correspond to the standard deviations. The solid line is the result of a weighted nonlinear fit of the aggregation model with a first cac of (0.5 ± 0.3) μM (vertical grey dotted line) and a second cac of (18.9 ± 2.2) μM (vertical grey dashed line) and widths of the transition regions σ = r·cac = ±0.2 μM and ±5.7 μM, respectively (shaded bars around the cacs). The shaded area around the fit curve indicates 95% confidence bands. Main Panel: logarithmic concentration scale. Inset: linear concentration scale.*

For Aβ42, we previously found a single critical aggregation process with a *cac* around 0.09 μM.[68] However, the concentration ranges are significantly different in both studies. Due to the strong adsorption of Aβ42, in our previous study the true concentration of Aβ42

did not exceed 1.5 µM, about 17 times its *cac*. In contrast, the much lower tendency to adsorb of Aβ40 allowed us in this study to reach much higher concentrations (300 times the $cac_1$) and thus to observe different aggregation processes not visible in the previous study.

The $cac_1$ of Aβ40 is approximately 5 times higher than that of Aβ42. However, when comparing a concentration range up to 17 × $cac_1$, the behaviour of Aβ40 is very similar to that of Aβ42. Below the *cac,* only small or no aggregates are found. Above it, stable aggregates with a relatively narrow size distribution are formed. For Aβ42, no further aggregation process was found due to the limited concentration range. In contrast, for Aβ40, a second aggregation process was observed above the $cac_2$ (≈ 25 × $cac_1$), resulting in much bigger aggregates. This second aggregation can be explained by the further growth of the initial oligomers to form protofibrils, as described in the literature.[66]

The higher $cac_1$ of Aβ40 and its much lower degree of aggregation compared to Aβ42 align with its significantly lower tendency to aggregate. Moreover, this significant difference in the aggregation behaviour could be underestimated since, in contrast to the Aβ40 samples used in this work, the Aβ42 samples for oligomerization studies were treated with DMSO prior to solubilization with PBS[68], possibly leading to a somewhat higher cac of Aβ42.

Compared to literature data, the $cac_2$ of Aβ40 determined in this work is close to the value of 35 µM reported by Yong *et al*.[69] The low degree of aggregation could also explain the mechanism proposed for Aβ40 of direct fibril formation.[66]

## 5.2.4. Size distribution of the aggregates

We estimate the size distribution of the aggregates from the histogram of the mean aggregation number of the oligomers shown in Figure 5.22 (Section 4.5.3.7 of **Data Analysis**, Step 8). The histogram reveals a high number of very small aggregates with $\bar{n} <$ 30, and a second population with higher aggregation numbers between 40 and 80. Furthermore, a small number of much bigger aggregates were observed, with sizes up to several hundreds of monomers. The number of these big aggregates is clearly underestimated due to limitations of the FCS experiment.

Fitting the histogram with a triple log-normal distribution, we obtain a first group of small aggregates of 5 - 25 monomers ( $\bar{n}_{mean,1} = 14$) and a second wider group with 40 to 80 monomers ( $\bar{n}_{mean,2} = 62$) per aggregate. The third distribution represents much bigger aggregates.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

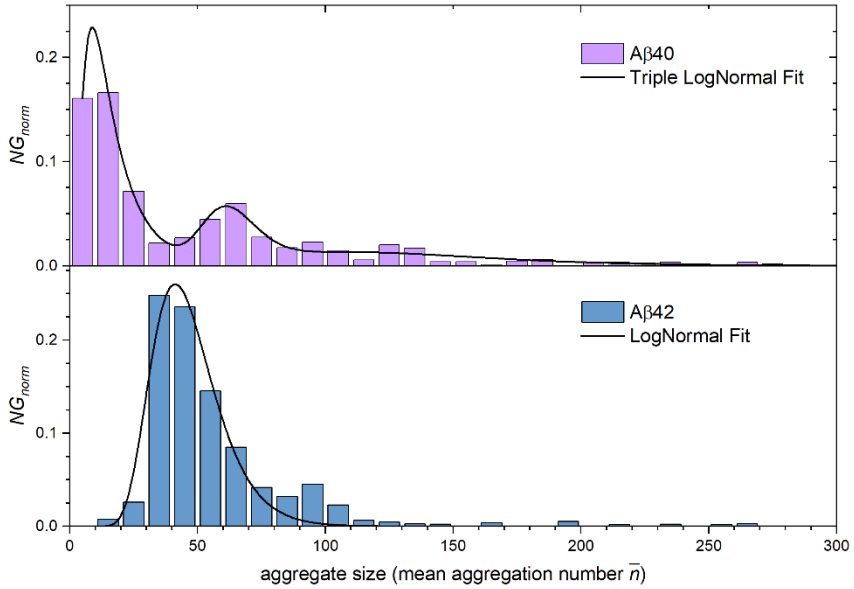



*Figure 5.22. Size distribution of Aβ aggregates. Upper panel: Histogram of the mean aggregation number obtained from all Aβ40 samples at different amyloid concentrations. The curve corresponds to the fit of a triple Lognormal distribution with* $\bar{n}_{mean,1}$ *= 14,* $\bar{n}_{mean,2}$ *= 62, and* $\bar{n}_{mean,3}$ *= 128. Lower panel: Histogram of Aβ42 aggregates recalculated with data from reference 68. The curve corresponds to the fit of a Lognormal distribution with* $\bar{n}_{mean}$ *= 45.*

To relate this size distribution to the double aggregation process deduced from the degree of aggregation (Figure 5.21) we constructed separate size histograms from samples with very low and very high true Aβ40 concentrations (Figure 5.23). Below the first critical aggregation concentration ($cac_1$), only small aggregates were found (Figure 5.23A). Once the $cac_1$ is reached a relatively homogeneous population of aggregates with $\bar{n} = 62$ forms. Well above the second aggregation concentration mostly much bigger aggregates are observed with a very wide size distribution (Figure 5.23B).

A possible interpretation of these findings is that during the first aggregation at $cac_1$ a small fraction of the monomeric Aβ40 forms stable oligomers with predominantly $\bar{n} = 62$ and a much lower amount of metastable oligomers with $\bar{n} = 14$, which convert into the former as concentration increases. At higher concentrations, above $cac_2$, these oligomers coalesce and/or seed much bigger aggregates involving a much higher fraction of the amyloid. In fact, at the studied higher concentrations the critical aggregation model used for the fitting may underestimate the degree of aggregation, which may increase to much higher values than the model indicates. Experiments with much higher amyloid concentration would be needed to confirm one or the other prediction.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

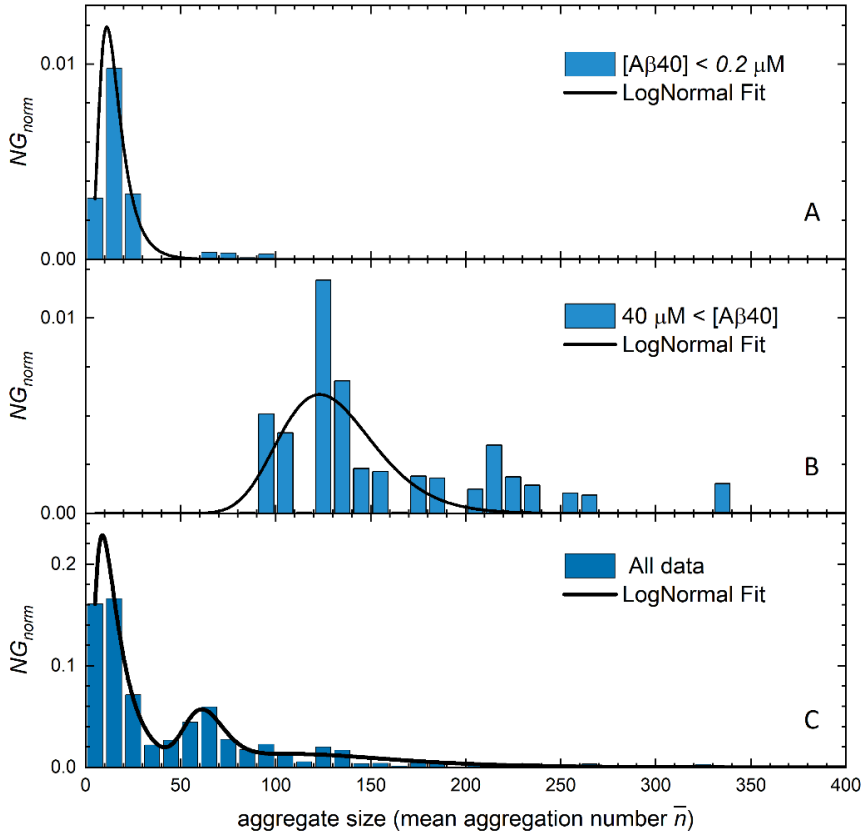



*Figure 5.23. Size distribution of Aβ40 aggregates. Histograms of the mean aggregation number obtained from samples in the indicated amyloid concentration intervals. The closed curves in panels A - C correspond to the fits of lognormal distributions with $n_{mean,1}$ = 14, $n_{mean,2}$ = 62, and $n_{mean,3}$ = 128, respectively.*

Focussing on the smaller oligomers, their aggregation numbers correspond to values of the hydrodynamic radius around 4 nm for the metastable aggregates and 12 nm for the larger ones. These results are in very good agreement with the reported micelle-like oligomers of Aβ40 with hydrodynamic radii around 5 nm at 120 min (metastable) and 10-20 nm at 720 min.[73] Additionally, two populations of low (0.9 nm <$R_h$ <5 nm) and higher (5 nm <$R_h$ <50 nm) molar mass were observed by Deleanu *et al.*[66], with a very small percentage of the total amyloid, which is consistent with the low fraction of aggregated amyloid that we observed in the first aggregation step.

In contrast, in our study of the early aggregation of Aβ42[68], we only observed the oligomer species with larger aggregation number around 50 and hydrodynamic radius between 7.3 nm and 10.7 nm. This difference was also reported by Morel et al.[73], who found only stable oligomers around 15 nm.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

### 5.2.5. Aβ40 and Aβ42 aggregation comparison

In this study, we have systematically investigated the aggregation behaviour of Aβ40 using fluorescence correlation spectroscopy (FCS). Our results reveal a two-step cooperative aggregation process for Aβ40, which contrasts with the single-step aggregation observed for Aβ42.

The first critical aggregation concentration ($cac_1$) for Aβ40 is approximately 0.5 ± 0.3 μM, about five times higher than that of Aβ42. Below this concentration, no significant aggregation is observed. Surpassing $cac_1$, small metastable oligomers of about 14 monomers and stable oligomers of 62 monomers form, together not exceeding 10% of the total Aβ40 concentration. These stable aggregates are similar in size and shape to those observed in Aβ42, but the higher *cac* and the much lower fraction of aggregated amyloid highlight Aβ40's much lower aggregation tendency.

A second critical aggregation concentration ($cac_2$) at 18.9 ± 2.2 μM leads to the formation of much larger aggregates, possibly protofibrils, with around 50% aggregated amyloid. This demonstrates the potential for further growth of initial oligomers at higher concentrations.

The transition widths for these aggregation steps indicate relatively narrow concentration ranges in which these processes occur, underscoring the distinct stages of Aβ40 aggregation.

Our results suggest that early Aβ40 oligomers are reversible upon dilution, although a kinetic barrier exists, indicating that longer times are necessary for complete reversibility.

Our comparative analysis shows that while Aβ40 and Aβ42 share some similarities in their aggregation behaviour, Aβ40 has a markedly lower aggregation tendency. This is evidenced by its higher $cac_1$ and lower degree of aggregation. Finally, the stability studies indicate that Aβ40 oligomers remain stable over time, with no significant changes in aggregate size or composition after one week, although an increase in polydispersity is noted. This stability, combined with the higher aggregation threshold, further underscores the distinct aggregation dynamics of Aβ40 compared to Aβ42.

Our findings provide a deeper understanding of the aggregation mechanisms of Aβ40 and highlight the importance of concentration and incubation time in studying amyloid aggregation. These insights could inform therapeutic strategies targeting amyloid aggregation in Alzheimer's disease.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# 5.3. Amyloid autofluorescence as method for amyloid's early aggregation detection

The tyrosine residue present at the amyloid's sequence gives the peptide an intrinsic fluorescence that could be a promising tool for amyloid detection. Here, we present a study on amyloid autofluorescence and a semi-quantitative characterisation of tyrosine's photophysical behaviour. The experimental procedures were previously explained in **Experimental methods** and sample preparation was carried out as explain in section 3.3.3.

## 5.3.1. Oligomer's autofluorescence: spectral variations

We performed fluorescence emission and excitation measurements for samples with only unlabelled Aβ40 at different amyloid concentrations. The concentrations were chosen based on the results obtained from the previous work (**section 5.2**) to cover all three steps of the aggregation process: around the first $cac_1$ = 0.5 μM, between $cac_1$ and $cac_2$, and well above the second $cac_2$ = 19 μM. Figure 7.1 shows the emission and excitation spectra obtained for Aβ40 solutions 86 μM, 6 μM and 0.47 μM using different excitation and emission wavelengths.

The emission spectra obtained with both 245 nm excitation and 290 nm excitation (Figure 5.24a-b) for the lowest amyloid concentration present one emission band at around 350 nm with a small shoulder at 500 nm. When amyloid concentration increases above the first critical aggregation concentration, which was determined in the previous study (**section 5.2**) to be of 0.5 ± 0.3 μM, a spectral shift of the band at 350 nm towards lower wavelengths can be seen. This shift is not as evident in the emission spectra monitored at 290 nm excitation (Figure 5.24b) due to the Raman band which is present at approximately 310 nm. Nonetheless, the disappearance of the band at 350 nm after amyloid concentration is increased is observed. Moreover, the contribution at 500 nm increases and can be distinguished as a second emission band. The observed spectral shift becomes even clearer at the highest amyloid concentration studied, 86 M μM, where the first emission band appears at around 300 nm and the second emission band appears at around 550 nm.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

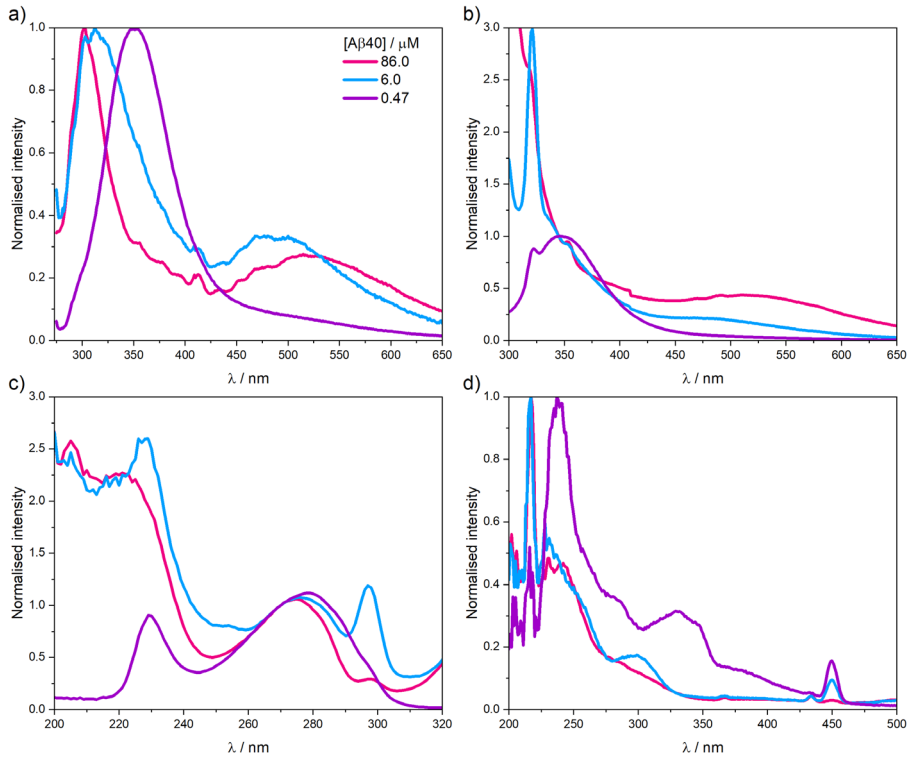



*Figure 5.24. Autofluorescence of Aβ40 in PBS at different concentrations: a) emission spectra normalised at their maxima with $\lambda_{ex}$=245 nm, b) emission spectra normalised at 350 nm with $\lambda_{ex}$=290 nm, c) excitation spectra normalised at 270 nm with $\lambda_{em}$=330 nm and, d) excitation spectra normalised at their maxima with $\lambda_{em}$=530 nm. 86 μM spectra were acquired with a 1.5x1.5 mm microcuvette while 6.0 μM and 0.47 μM spectra were obtained with a 3x3 mm one.*

Excitation spectra obtained at the emission wavelength of 330 nm (Figure 5.24c) show a band at low wavelengths around 260-280 nm, which is the expected absorption range for the aromatic amino acids. The band seems to shift to higher wavelengths when amyloid concentration decreases. A narrow band at 300 nm can be seen as well in the excitation spectra, which must be attributed to the water Raman band. Excitation spectra at 530 nm emission (Figure 5.24d), on the other hand, present a spectral shift towards lower wavelengths with amyloid concentration, and its main band appears at around 330 nm for the lowest concentration and 300 nm for the other two concentrations. However, it must be noted that excitation spectra for the 530 emissions are affected by the second order Rayleigh band, which should appear at around 260 nm.

The short-wavelength band of the Aβ40 emission spectra can be attributed to tyrosine emission, which emits in the form of tyrosinate (345 nm) at low amyloid concentrations.[93] When increasing the amyloid concentration, this band shifts towards shorter wavelengths showing the typical emission of protonated tyrosine (303 nm) at high amyloid concentrations. This behaviour indicates that aggregation of amyloids promotes a structural change in its tyrosine residue. Tyrosine deprotonates in the excited state when the peptide is monomeric, because it can be reached by water and weak bases present in the medium. This deprotonation can happen because, as mentioned in the

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

*Introduction*, the p$K_a$ of the phenolic proton of tyrosine varies from a value of 10.3 in the ground-state to about 4 in the excited state, and the ionization is catalyzed by bases.[93] In the aggregated form, the deprotonation of the tyrosine would be hindered as the conformational change of the peptide would protect it from the aqueous environment. Thus, it would be less accessible for basic species to reach it.

The band at 530 nm, which appears at the higher amyloid concentrations (Figure 5.24a-b), seems to be related to the presence of aggregates and it resembles the visible bands described in the literature for a long series of proteins and peptides.[177] However, this band seems to be contributed by an intrinsic emission or scattered light coming from the quartz microcuvettes when the excitation wavelength is set in the UV spectral range. As shown in Figure 5.25a, Tris buffer spectra present a band at around 550 nm when the used cuvette is 3x3 mm or 1.5x1.5 mm. This contribution disappears when the used cuvette is a conventional 10x10 mm cuvette. This effect could be due to the thick quartz walls of the microcuvettes, which is also black tinted. Although for most fluorescence studies they do not present any problem, when working in the UV range they seem to cause artefacts. This behaviour can be also seen in Figure 5.25b, where the sample amyloid sample was measured in a 3x3 mm and 10x10 mm cuvette in the two different FS5 fluorometers available in our group. Additionally, we observed that the contribution is higher when the used spectrofluorometer is FS5 A than FS5 B, which might be explained by an incorrect alienation of the equipment. It must be noted that both are brand new equipment and, therefore, the differences in calibration come from the manufacturer.

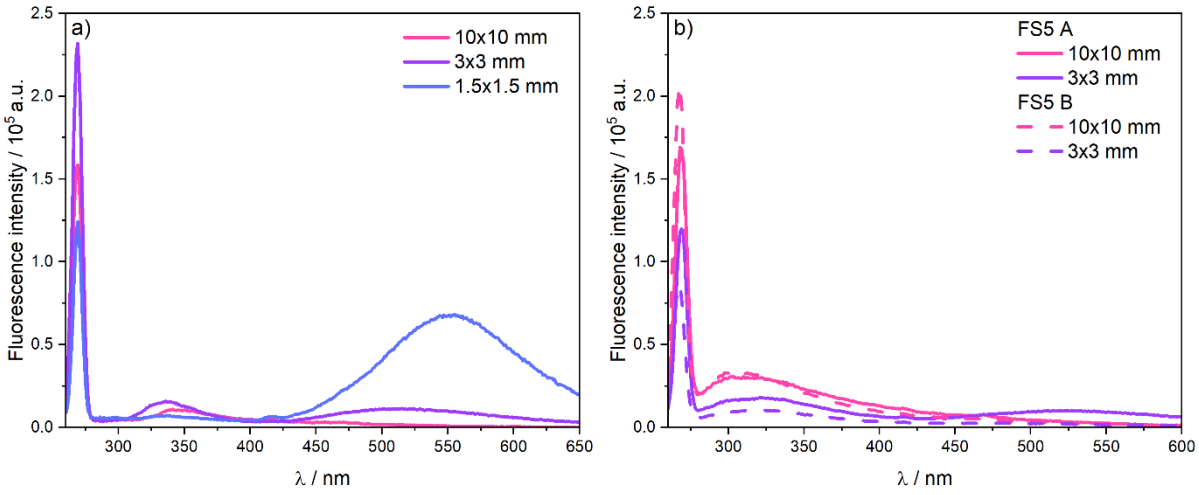



*Figure 5.25. a) Tris spectra obtained in the FS5 A with different cuvettes. b) Same amyloid sample in different cuvettes and equipments. 10x10 mm cuvette spectra were obtained adding a home-made mask to reduce the measured volume.*

Since the emission band does not depend on excitation wavelength, as it appears at around 530 nm for both studied excitation conditions (Figure 5.24a-b), scattering alone cannot explain the emission band appearing at higher wavelengths. Nor can the intrinsic quartz emission account for the long-wavelength emission. Therefore, we think that the

observed band is a real emission from the sample, but it is contributed by instrumental distortions caused by the microcuvettes.

### 5.3.2. Oligomer's autofluorescence: intensity variations and aggregation reversibility

To get a further insight into the autofluorescence properties of amyloid peptides when they undergo early aggregation, we carried out two series of emission experiments: one with a low concentration range (0.47–6.0 μM, Figure 5.26 left) and another at a high concentration range (9.9 – 86 μM, Figure 5.26 right). For these experiments samples were systematically diluted following the dilution-extraction protocol. As shown in Figure 5.26, amyloid's aggregation promotes both intensity and spectral variations. In both cases the band at higher wavelengths can also be seen, even though it appears at somewhat lower wavelengths (500 nm) for the low concentration range than for the high concentration range (520 nm). Additionally, we record the used cuvette's emission spectra for each series, being the 3x3 mm cuvette for the low concentrations and the 1.5x1.5 mm cuvette for the high concentrations. It is evident that, although the contribution is much higher for the smaller cuvette, both present intensities that should be considered.

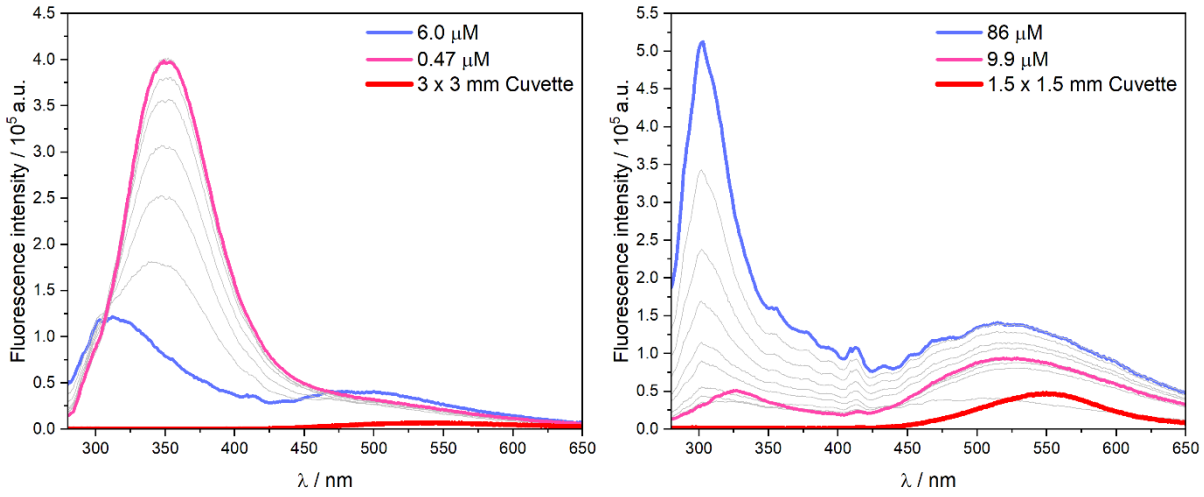



*Figure 5.26. Experimental emission spectra of Aβ40 in PBS obtained with $\lambda_{ex}$=245 nm at concentration ranging from 0.47 μM to 6 μM (left) and 9.9 μM to 86 μM (right). The spectra of the used cuvettes are also presented (red curves).*

Thus, we corrected the emission spectra by directly subtracting the recorded cuvette's emission spectra to the emission spectra of each sample (Figure 5.27). If we study the spectral variations, we see that, in the low concentration range, the intensity of the band at 350 nm decreases upon aggregation (Figure 5.27 left). This diminution in intensity is also accompanied by the spectral blue shift we had already seen in Figure 5.24. This shift towards lower wavelengths continues in the high concentration series (Figure 5.27 right), were a clear rise of the low wavelength band intensity occurs, but in this case the band moves from approximately 330 nm to 303 nm at the highest concentrations of this series

USC DE COMPOSTELA

(Figure 5.27 right) and presents an increasing intensity with concentration instead of a decreasing tendency. This 350 nm to 303 nm shift with increasing concentration is especially interesting because, as it was explained before, it is the expected shift for the tyrosine emission going from the excited tyrosinate in monomers to the excited tyrosine in the aggregates. Hence, data corroborate that in the monomeric Aβ40 the tyrosine residue is exposed to the aqueous medium and undergoes excited-state ionization.

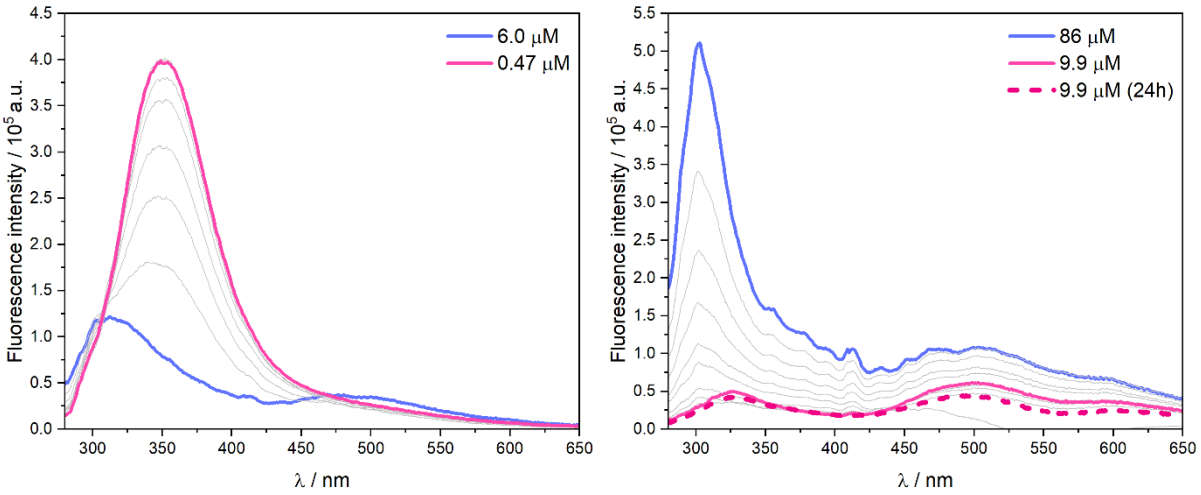



*Figure 5.27. Cuvette corrected emission spectra of Aβ40 in PBS obtained with $\lambda_{ex}$=245 nm at concentration ranging from 0.47 to 6 μM (left) and 9.9 to 86 μM (right). Low concentration series data were acquired with a 3x3 mm cuvette and high concentration series with a 1.5x1.5 mm one.*

Although with less intensity, the high wavelength band is still present after cuvette's contribution subtraction and shows a maximum at approximately 500 nm independently of amyloid concentration (Figure 5.27). This band can be seen in all samples of the higher concentration series (Figure 5.27, right), but it shows less intensity than the 303 nm one. On the other hand, in the low concentration series (Figure 5.27, left), only the 6.0 μM Aβ40 sample shows this band clearly. These findings indicate that the band at higher wavelengths is related to the presence of Aβ40 aggregates in solution. In the low concentration series, aggregates are present at the starting 6.0 μM solution but they gradually disappear as the solution is diluted showing a complete reversibility as observed in the previous aggregation study. However, when the initial concentration is high, well above the second *cac*, disaggregation seems to be less effective and the band at higher wavelengths drop more slowly than that attributed to tyrosine (Figure 5.27, right). This behaviour may be due to the different aggregates formed in the two concentration ranges. As explained in **section 5.2**, at concentrations between the first and second *cac* both small metastable and medium-size, stable aggregates are formed with good reversibility, while, above the second *cac,* the formed oligomers are bigger in size and present less reversibility, suggesting a kinetic barrier for disaggregation. This is also corroborated by the fact that, as seen in Figure 5.27 right, the 9.9 μM of the high concentration series does not decrease in intensity as expected right upon dilution, but

after 24 hours it presents a lower intensity which is coherent with a further disaggregation, as also observed in the previous aggregation study.

The observed intrinsic visible band of Aβ40 could be explained by an aggregation induced emission (*aie*), as reported in the literature for many peptides and solid materials.[177,178] Moreover, structure-specific intrinsic fluorescence at long wavelengths has been observed for tyrosine in amyloid peptides.[179] Recent studies in our group of the free tyrosine amino acid have shown excimer emission at around 410 nm that supports the hypothesis of an aggregation induced emission, which would be due to similar excited-state interactions between tyrosine residues in the aggregates. However, further photophysical characterization of the aggregates must be carried out to confirm this hypothesis.

Now, if we correct the spectra by amyloid concentration (Figure 5.28), we observe that the band at 350 nm becomes even more intense while the band at 303 nm reduces its intensity. This happens at both concentration series and, in addition, in the higher concentrations the fluorescence intensity is significantly reduced being almost 2 orders of magnitude lower that for the lowest concentration series. The band at 350 nm can be attributed to the monomeric form of Aβ40, and it exhibits a strong quenching when amyloid concentration is increased. This can be interpreted as an aggregation-induced quenching by the formation of dimers or oligomers. Concentration corrected spectra also present the 500 nm band, which gains relative intensity after the correction in the higher concentrations (Figure 5.28 right) and is negligible in the low concentration series (Figure 5.28 left). For the higher concentration range an increase in intensity is observed when samples are diluted down to the 9.9 μM (Figure 5.28, right), whereas is hardly observed in the low concentration range due to the strong emission of the monomer. The 500 nm band seems to become especially important at concentrations around 10 μM, between the two *cac* values, suggesting that it is related to the metastable small aggregates and the stable medium size that were seen in the middle concentration ranges in section 5.2.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

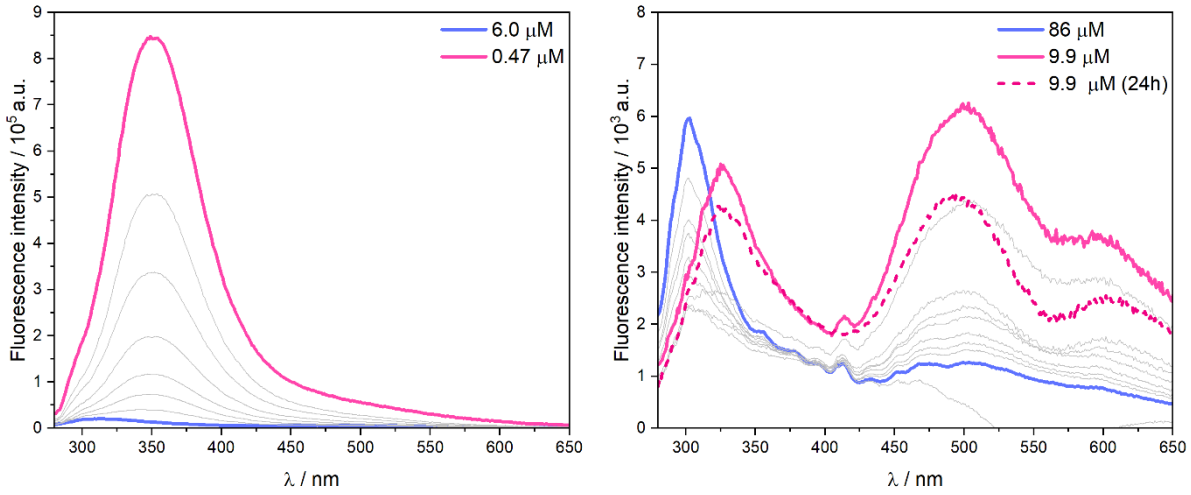



*Figure 5.28. Concentration and cuvette corrected emission spectra of Aβ40 in PBS obtained with $\lambda_{ex}$=245 nm at concentration ranging from 0.47 to 6 μM (left) and 9.9 to 86 μM (right). Low concentration series data were acquired with a 3x3 mm cuvette and high concentration series with a 1.5x1.5 mm one.*

## 5.3.3. Oligomer's autofluorescence: FCS corroboration

We also performed measurements to correlate the Aβ40 autofluorescence with its degree of aggregation determined by FCS, where samples were prepared following both the independent sample protocol and the dilution-extraction protocol. These samples were first measured in the spectrofluorometer and then, after addition of the labelled amyloid, measured FCS to obtain not only their true concentration but also the degree of aggregation. Because of the contributions to fluorescence the microcuvettes had, the cuvette was changed to a 10x10 mm, and a home-made mask was used to allow for lower sample volumes. Additionally, to prevent any potential quenching the presence of phosphate could promote in amyloid's autofluorescence,[93] the buffer solution was changed to Tris buffer for these studies. Figure 5.29 shows the experimental emission spectra obtained with excitation wavelengths of 245 nm. These measurements corroborate the behaviours seen previously, as the band at 303 nm shifts towards higher wavelengths as samples are diluted and, in consequence, amyloid concentration is reduced. However, the spectral shift in these experiments seems to only reach 330 nm instead of 350 nm (Figure 5.29). Additionally, the obtained emission intensities do not increase as expected and stay in much lower values than in the previous studies, which suggests that the disaggregation of Aβ40 is not happening at the same rate.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

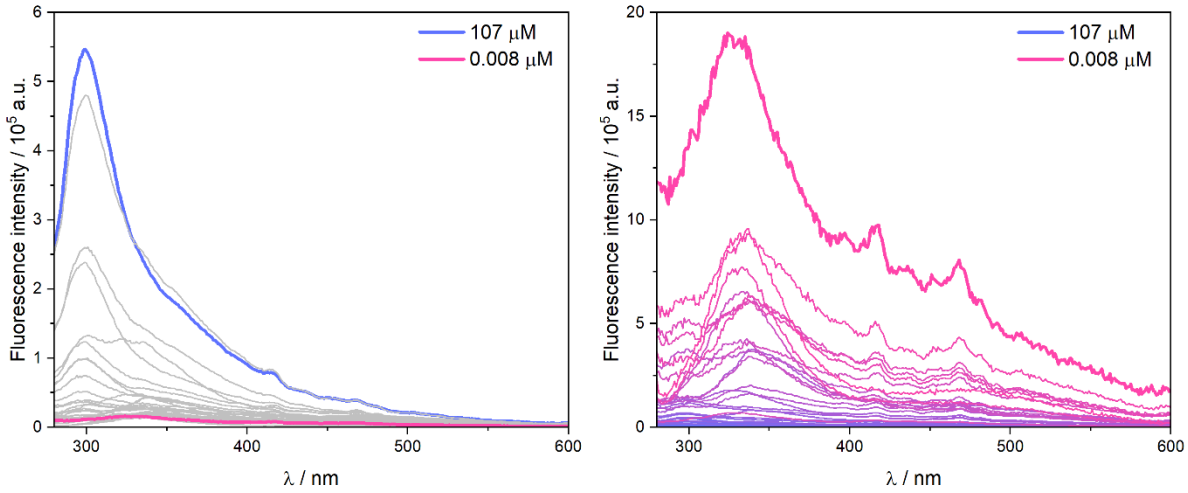



*Figure 5.29. Corrected emission spectra of Aβ40 at different concentrations obtained with a 245 nm excitation wavelength. Left: correction is the subtraction of the Tris buffer emission, and the real concentrations are indicated. Right: further correction dividing by the true concentration indicated for the extreme values.*

As samples were also measured in FCS, we could correct their emission spectra by the true concentration calculated for each sample with FCS data (Figure 5.29, right). The true concentration was determined using the same data analysis protocol as in section 5.2, which is described in the Data Analysis (section 4.5.3.7, Step 6). After the experimental data were corrected by the true concentrations, the obtained spectra for 245 nm (Figure 5.29, right) present the expected intensity variations. That is, the intensity of the low wavelength band decreases as amyloid aggregation takes place. However, the obtained intensities are still not in agreement with the previous studies (Figure 5.28). Additionally, as the raw data intensities were very low, spectra also present great noise. For this reason, the band at 500 nm cannot be separated from the 350 band in the spectra obtained with 245 nm excitation although its contribution is evident.

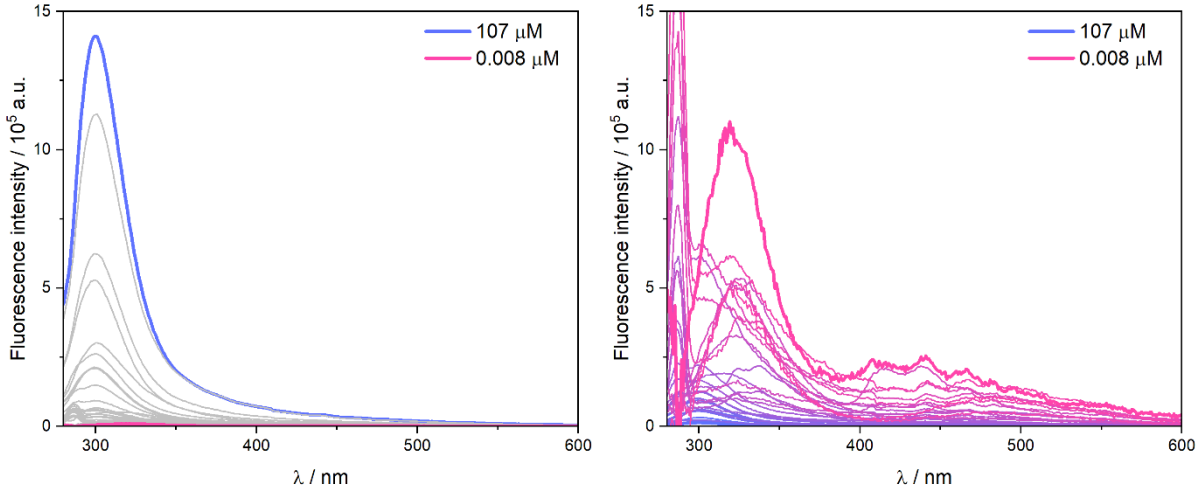



*Figure 5.30. Corrected emission spectra of Aβ40 at different concentrations obtained with a 260 nm excitation wavelength. Left: correction is the subtraction of the Tris buffer emission, and the nominal concentrations are indicated. Right: corrected by the true concentration indicated for the extreme values.*

Figure 5.30 left shows the emission spectra obtained with an excitation wavelength of 260 nm, after the solvent emission had been subtracted. In these measurements we can observe again that even though the band at 303 nm shifts towards higher wavelengths when amyloid concentration decreases, it does not shift to 350 nm but stays at 330 nm. As in the case of the 245 nm excitation, the obtained emission intensities for the 350 nm band do not increase as much as before, being almost imperceptible. The concentration-corrected spectra (Figure 5.30, right) do show the increase in intensity with amyloid disaggregation as the 245 nm data. In the uncorrected data, the band at 500 nm is not distinguishable the spectra show a long tail. However, after the correction of the spectra by the true FCS concentrations, the existence of a high wavelength band is confirmed by the 260 nm excitation data (Figure 5.30, right). It can also be observed that the intensity of this band increases at lower amyloid concentrations.

It must be noted that these series of spectra do not have sufficient quality to extract quantitative results, but support the qualitative explanation given before. It is important to consider that this data is also highly affected by the Raman peak which appears at around 290 nm, distorting specially the spectra of the high concentration range.

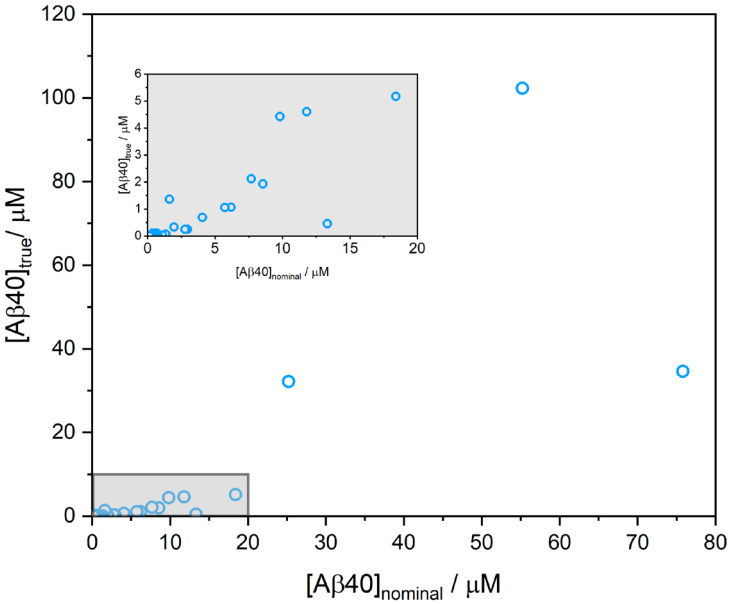
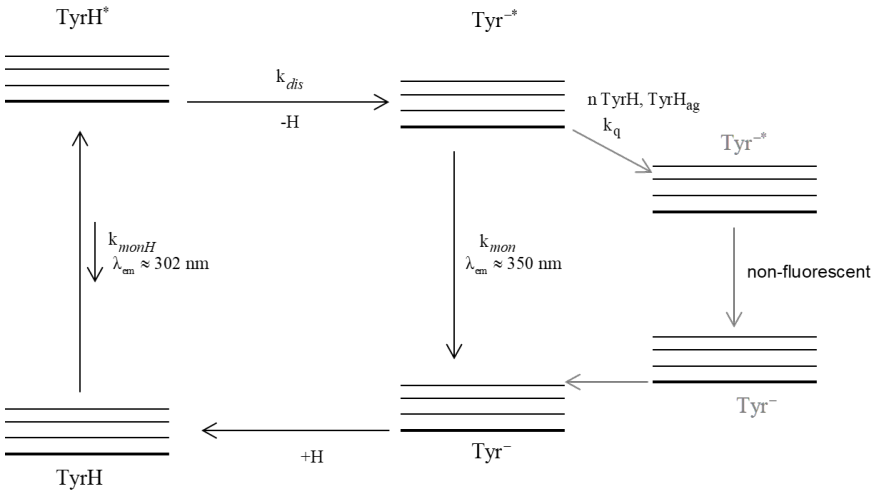



*Figure 5.31. True total Aβ40 concentration calculated from FCS data versus the total Aβ40 nominal concentration.*

Figure 5.31 shows the calculated true concentrations from the FCS measurements versus the nominal concentrations for each sample. This evidences the discrepancy between the two concentrations, which we did not see for the samples in section 5.2. We do not have a sure explanation for this, but we are aware that several variables were changed in these series of data. First, new commercial samples of Aβ40 and the HFIP solvent used for the disaggregation of the amyloid had to be bought, with the corresponding different production batches. In the case of the commercial amyloid,

some batches showed fluorescent impurities which interfered in the autofluorescence studies. Additionally, as explained before, the buffer in these studies was changed from PBS to Tris to prevent the potential quenching that has been described in the literature.[93] Another potential reason for these unexpected results is that the treated amyloid were not properly disaggregated.

Because of these, we believe that even though these data do present the expected variations with amyloid concentration and qualitatively agree with the previous studies, better quality data should be acquired to do a quantitative study. Despite the poor quality of the data, they allow for postulation of a hypothetical photophysical mechanism for the tyrosine residue in both the monomeric amyloid (Scheme 5.1) and the aggregated amyloid (Scheme 5.2).



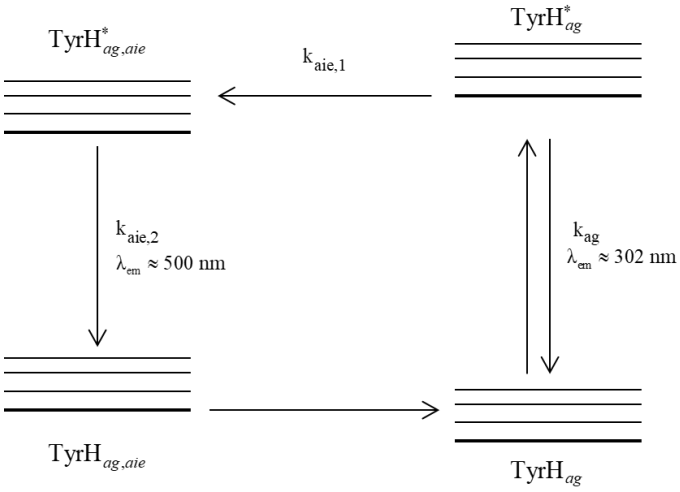

*Scheme 5.1. Proposed photophysical mechanism of the tyrosine present in monomeric Aβ40.*

As shown in Scheme 5.1, if a tyrosine residue inside of a non-aggregated amyloid is excited (**TyrH**), the probability of being deactivated from the tyrosine excited state (**TyrH**$^*$) directly to the tyrosine ground state, that would cause the emission around 303 nm, is very low. Instead, it would more probably be deactivated by proton transfer,[93] transferring a proton to the medium, and becoming excited tyrosinate (**Tyr**$^{-*}$). Then, the excited tyrosinate will deactivate to its ground state (**Tyr**$^-$), with a 350 nm emission. However, when **Tyr**$^{-*}$ collides with other amyloids or amyloid aggregates, it suffers an emission quenching that will deactivate **Tyr**$^{-*}$ through a non-emissive pathway.

On the other hand, when tyrosine is part of Aβ40 aggregates (**TyrH$_{ag}$**, Scheme 5.2), the excited tyrosine is not exposed to the media, preventing its excited state from deprotonation. As the excited state tyrosinate cannot be formed, tyrosine being part of amyloids that are aggregated (**TyrH$^*_{ag}$**), will emit falling directly to its ground state, showing emission at 302 nm. But, as we saw in the previously shown data, an emission around 500 nm happens which is related to the aggregates. We propose that this emission is due

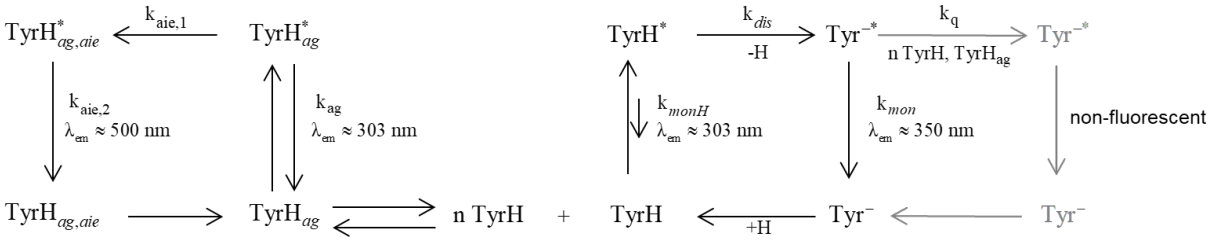
USC DE COMPOSTELA

to the excited-state interaction between the aromatic amino acids in Aβ40 which leads to a red shifted emission similar to the excimer bands called aggregate induced emission (*aie*).




*Scheme 5.2. Proposed photophysical mechanism of the tyrosine residue of an aggregated Aβ40; ag subscript stands for aggregate and aie subscript for aggregated induced emission.*

Combining the two mechanisms, the photophysical mechanism for both non aggregated (Scheme 5.2) and aggregated Aβ40 amyloids (Scheme 5.2), we propose a full model that can explain the experimental data obtained, shown in Scheme 5.3. This model can explain the band at 350 nm at lower Aβ40 concentrations (Figure 5.28, left), and its intensity decrease and shift towards 303 nm as the Aβ40 concentration increases, and also explains the band observed at 500 nm at high concentrations.



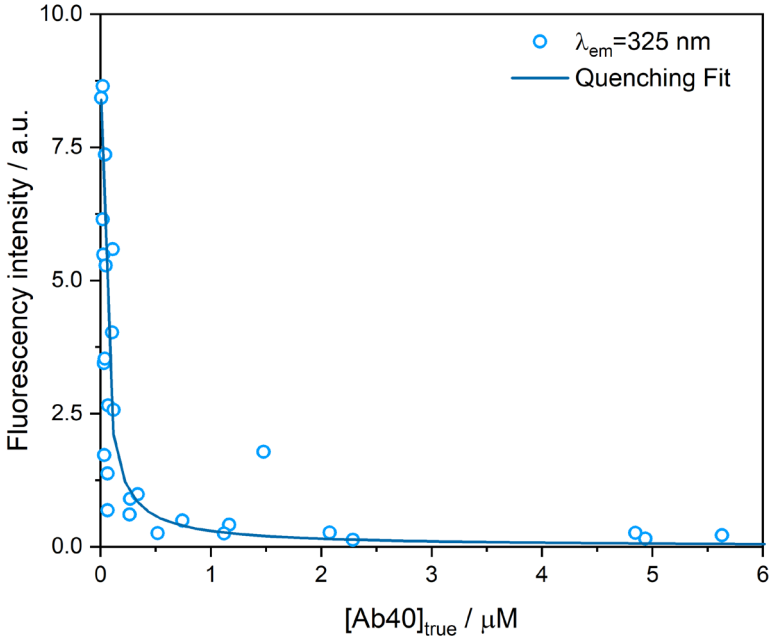

*Scheme 5.3. Diagram of the full proposed mechanism of the Tyrosine present in Aβ40 amyloids, both aggregated and non-aggregated.*

If we apply the proposed photophysical mechanism to the experimental data obtained in Tris that covers the true concentration range from 0.008 to 107 μM, we can see in Figure 5.31 how it is possible to obtain a satisfactory fit of the experimental data with an empirical model using a Stern-Volmer quenching function (Eq (5.1)). Where $F_0$ is the fluorescence intensity in absence of quencher, and $K_q$ is the quenching constant.

$$F^{\lambda} = \frac{F_0}{1 + K_q[Tyr]} \tag{5.1}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

The quenching constant $K_q$ of (3.6 ± 1.7) × $10^4$ $M^{-1}$ is high, in accordance with the strong quenching observed. However, to obtain quantitative results for all the kinetic constants that of the proposed photophysical mechanism, more studies are needed.




*Scheme 5.4. Concentration corrected emission intensity at 325 nm versus true amyloid concentration (blue circles) and the resulted fit curve using a Stern-Volmer equation (blue curve). The data was obtained with a $\lambda_{ex}$= 245 nm.*

Even if the amyloid fluorescent mechanism we propose can explain the observed data it must be considered that it is still a hypothesis that needs to be corroborated with higher quality data, as there are some points that need a deeper study to achieve a good understanding of the process, like the nature of the proposed aggregation induced emission, or the tyrosinate quenching process. The visible fluorescence of peptide and protein aggregates has been detected by other authors,[177] and its mechanism is still under debate, as the usual electronic models cannot fully explain this fluorescence.[180–182]

## 5.3.4. Amyloid's autofluorescence: a promising tool

Considering all the exposed results, we can suggest that autofluorescence of the amyloid peptide is a promising tool for amyloid aggregation detection. Our studies allowed us to semi-quantitatively follow the aggregation process through fluorescence emission, as the emission spectra present changes with aggregation and aggregated induced emission is observed. We have evidenced that the tyrosine residue undergoes proton transfer at the excited state when the amyloids are in monomeric form, shifting its emission band from 303 nm to 350 nm. We have proposed a preliminary photophysical mechanism for the Aβ40 autofluorescence. However, we believe that further studies need to be performed to verify the proposed mechanism and to obtain the deactivation constants of the process.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

## 5.4. Electrostatic interactions impact on amyloid-beta aggregation

Liquid-liquid phase separation has been proposed as pathway different intrinsically disordered proteins undergo during their aggregation process as explain in the **Introduction**. The exact reason for LLPS to take place is still unknown, but electrostatic interactions are thought to play a role. Thus, here we present a fluorescence microscopy study of amyloid-β(1-42) aggregation and the influence that the electrostatic interactions between the peptide and other charged biomolecules can have in the aggregation process. The set-up and experimental conditions have been explained in section 3.1.2.2, and sample preparation was described in section 3.3.4.

### 5.4.1. Fibril formation

As the objective of this project was to study the influence that small, charged biomolecules can have on amyloid aggregation, we first studied the amyloid aggregation in physiological conditions without addition of any cation concentration. It has been reported that Aβ42 critical aggregation concentration is 0.09 μM[68], meaning that at any concentration above that samples will present oligomer formation. However, the reported concentration was the real concentration obtained from FCS measurements and samples were prepared for nominal concentrations ten times higher. Additionally, for fibrillation to occur amyloid concentrations need to be higher, and incubation is also needed. Therefore, for our chosen conditions, which are 10 μM total amyloid concentrations, we needed to study the required incubation time for fibrils to fully form. As shown inFigure 5.32, when incubating our amyloid sample for 24 hours only droplet-like structure can be seen in the microscopy images. Data for longer incubation times, 48 and 72 hours, present both droplets and fibrils, while after 1 week only fibrils can be seen. The longer the incubation times are, the more fibrils are formed, and the less droplets can be seeing. Thus, by the end of fibril formation, which seems to happen after around 1 week incubation, no droplets are left.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

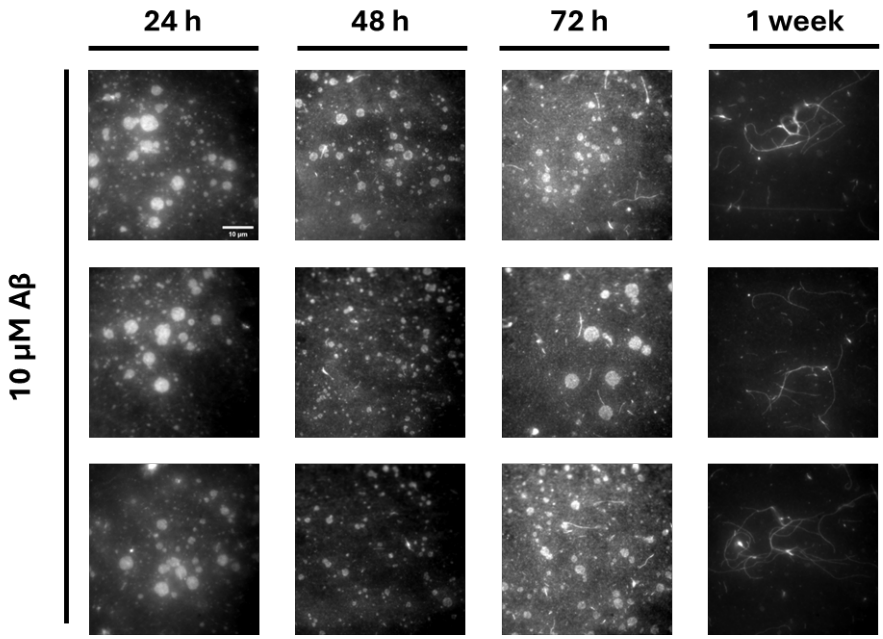



*Figure 5.32. Super-resolution images of 10 μM Aβ42 samples incubated at 37ºC during 24, 48, 72 hours and 1 week.*

The condensates formation that can be seen in the microscopy images corroborates the idea that amyloid peptides indeed aggregate by liquid-liquid phase separation as it has been suggested in the literature.[79,183,184] These condensates, often referred to as droplets, would represent a microenvironment in which local amyloid concentration is much higher promoting Aβ42's aggregation and fibril formation (see **Introduction**; Figure 1.7).

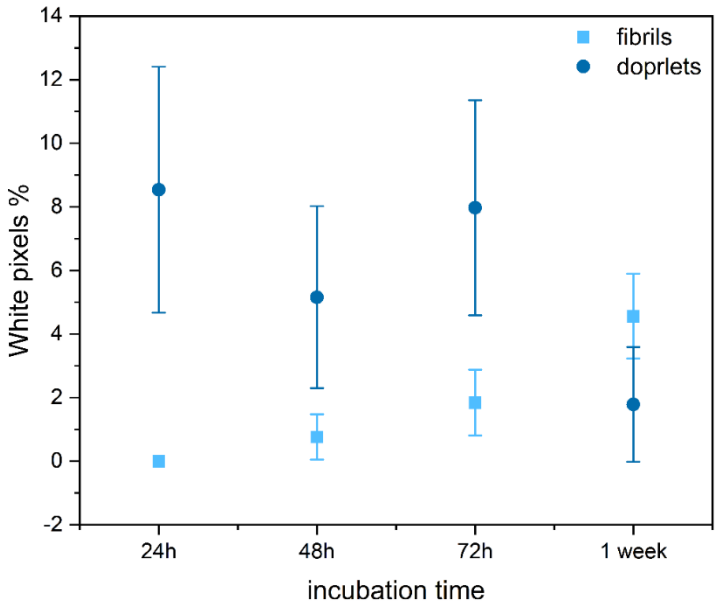



*Figure 5.33. Percentage of white pixels corresponding to fibrils (light blue squares) and droplets (dark blue circles) variation with incubation time. The mean values were calculated using all the values acquired from different images of experiments done under the same conditions. Error bars correspond to the standard deviation obtained from the calculation of the average.*

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

When analysing the percentage of fibrils and droplets variation with time (Figure 5.33) we verified that droplet's presence slightly decreases while fibrils increase. This could suggest that the fibrillation takes place by LLPS. As the condensates form, the local amyloid concentration inside them grows, promoting amyloid aggregation. Aggregates grow in size until fibrils are formed and then, the condensates break letting the fibrils into the solution.

We also studied the size distribution of the formed droplets at the different incubation times. Figure 5.34 shows how the sizes distribution present droplets sizes ranging mainly from 0.5 μm to 5 μm. The shortest incubation period presents a size distribution of the condensates with two size populations (Figure 5.34a). On the other hand, 48 hours and 72 hours incubation times (Figure 5.34b and Figure 5.34c) do not show a distinct second population even though the sizes between which their distributions range are similar to the ones from 24 hours data. The longest incubation time presents a narrower size distribution as it only presents droplets up to 3.5 μm (Figure 5.34d). If we fit the data to a Lognormal distribution we obtain similar mean sizes for all incubation times: 1.50 ± 0.02 μm for the first population at 24 hours, 1.48 ± 0.04 μm for 48 hours, 2.3 ± 0.1 μm for 72 hours and, 1.2 ± 0.04 for the first population at 1 week. These results indicate that there is no significant variation between the droplets size at different incubation times. Thus, we can suggest that under the studies conditions amyloid aggregation does not promote droplet size variation. However, literature suggest that droplets do grow in size as aggregation takes place during the LLPS phase before fibrils are completely formed.[77,185,186] Thus, it must be noted that early oligomers can be seen immediately after dissolving the peptide to concentrations above the critical aggregation concentration and, consequently, under our studied conditions even at the shortest incubation time aggregation should already be advanced. If the aggregation has already reached the elongation phase, droplet growth might not be perceived.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

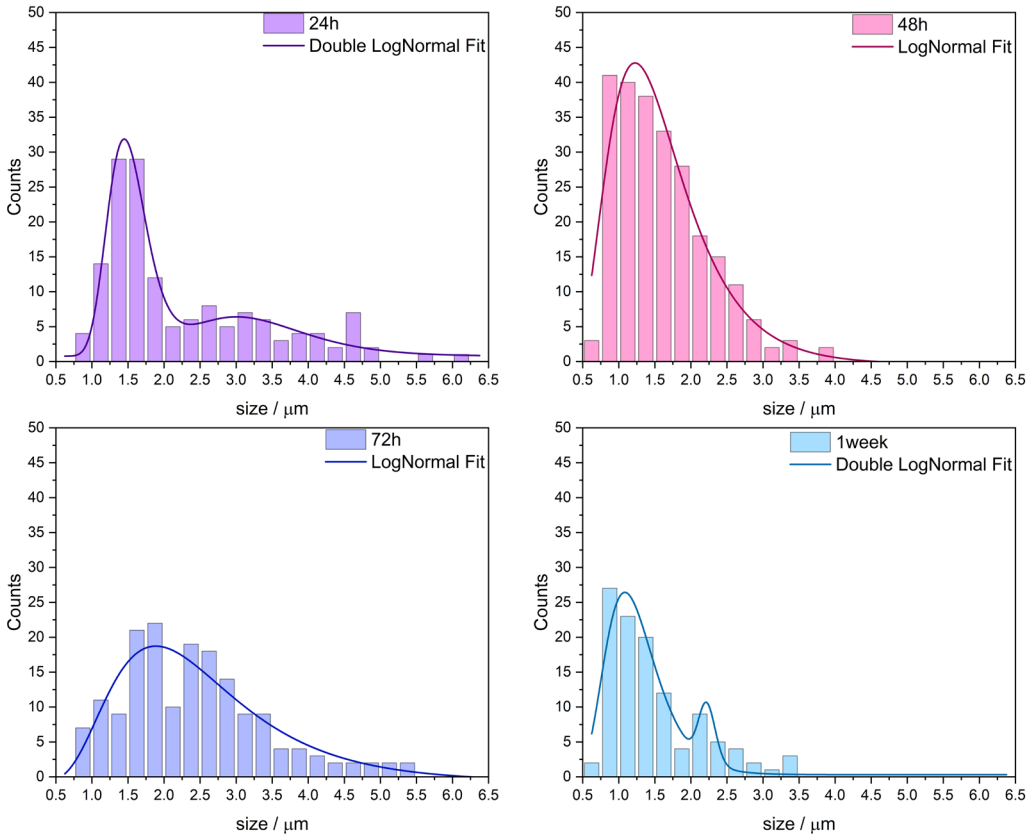



*Figure 5.34. Droplets size distribution acquired at different incubation times without cation presence (bars) and their respective LogNormal fits (continuous lines).*

## 5.4.2. Cation impact on fibrillation

Once the aggregation of amyloids undergoing LLPS was confirmed by our microscopy experiments, we also carried out experiments to help understand the implications electrostatic interactions between amyloid peptides, which present a negative charge at physiological conditions, and other positively charged small biomolecules may have in the aggregation process of Aβ.

### 5.4.2.1. Influence after fibril formation

***Spermine***

Spermine is a polyamine with a positive net charge at physiological conditions, as all its amines are protonated pH above 10. Thus, at pH 7.4 its net charge value is +4. We acquired images of samples with spermine concentrations ranging from 0-1 mM, increasing the spermine to amyloid ratio 10 times between each concentration conditions. After addition of spermidine to the already formed fibrils samples they were incubated for 24 hours at 37 ºC.Figure 5.35 shows the super-resolution images at all the studied concentrations. With a first visual analysis we can see a variation on the shape of the formed species. In the control sample only fibrils can be seen, but after spermine concentration is increased sufficiently droplet like shapes slowly start to appear. At the

highest spermine concentration (1 mM) almost all fibrils have disappeared and mainly only droplets can be seen.

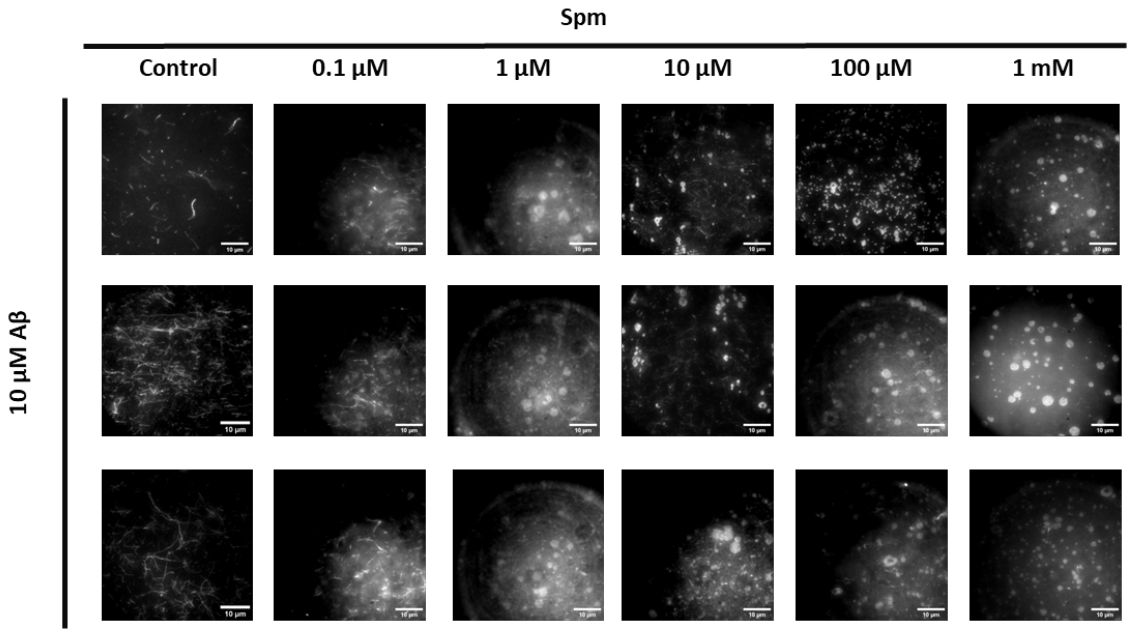



*Figure 5.35. Fluorescence microscopy images of Aβ42 samples incubated during 24 hours at 37 ºC with different spermine concentrations that were added after fibrillation had occurred. Scales for all images are equal to 10 µm.*

To see the degree at which this transformation takes place, we analysed all images obtaining the white pixel percentage of the fibril-like shapes and the droplet-like shapes separately. As shown in Figure 5.36, the percentage of fibrils present in the samples presents a decreasing tendency with spermine concentration, while the percentage of droplets grows with concentration. These results suggest that spermine would be reversing the formation of amyloid fibrils and that the reverse process takes place again by LLPS.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

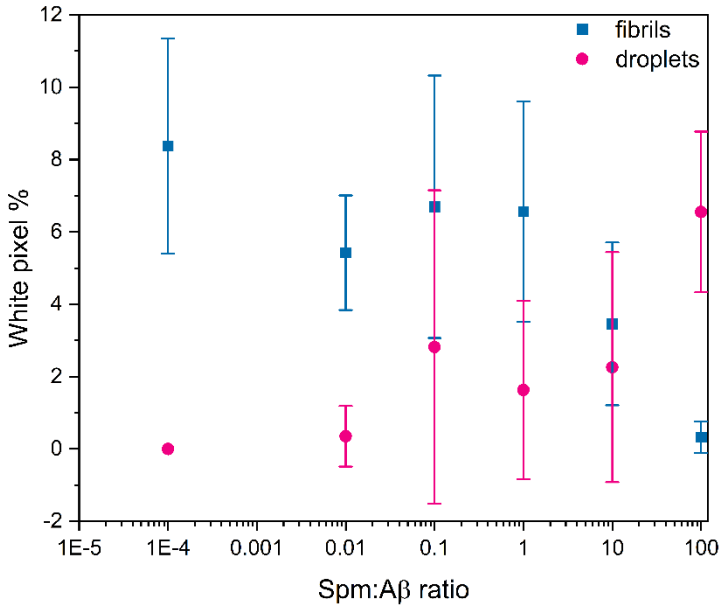



*Figure 5.36. Percentage of white pixels corresponding to fibrils (blue squares) and droplets (pink dots) variation with spermidine concentration in 24 hours incubation samples. Concentration is represented as the cation amyloid ratio, and amyloid concentration is always 10 μM. The control samples, which have no spermine, are represented as a 1x10$^{-4}$ ratio so it could be seen in the logarithmic scale. The mean values were calculated using all the values acquired from different images of experiments done under the same conditions. Error bars correspond to the standard deviation obtained from the calculation of the average.*

We also carried out an analysis on the droplets size, obtaining the distribution histograms shown in Figure 5.37. The obtained sizes are coincident with the above reported droplet sizes (Figure 5.34). The droplet's size at 1 μm (Figure 5.37a) presents a very wide distribution with significant frequencies of bigger droplets (sizes above 3 μm). The bigger condensates seem to disappear upon increasing spermine concentration as the distribution become narrower at 10 μM and 100 μM (Figure 5.37b and c). When concentration reaches 1 mM (Figure 5.37d) aggregates above 3 μm have almost disappeared. This indicates that droplets size decreases with spermine concentration, which could be explained by the degree of reversibility fibril formation undergoes. However, when the LogNormal fit of the data at all concentrations results in similar mean sizes: 2.3 ± 0.3 for 1 μM, 1.34 ± 0.04 for 10 μM, 1.00 ± 0.01 for 100 μM, and 1.43 ± 0.04 for 1 mM. The obtained mean sizes only decrease when increasing spermine concentration from 1 μM to 10 μM and then, do not present any tendency. Additionally, in the aggregation in absence of cation studies the droplets did not grow as aggregation was taking place and the bigger droplets that could be seen did not show a tendency with incubation time. The lack of size variation suggests that LLPS is only happening until the elongation phase, as different aggregation stages would result in different droplet sizes.[185,186]

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

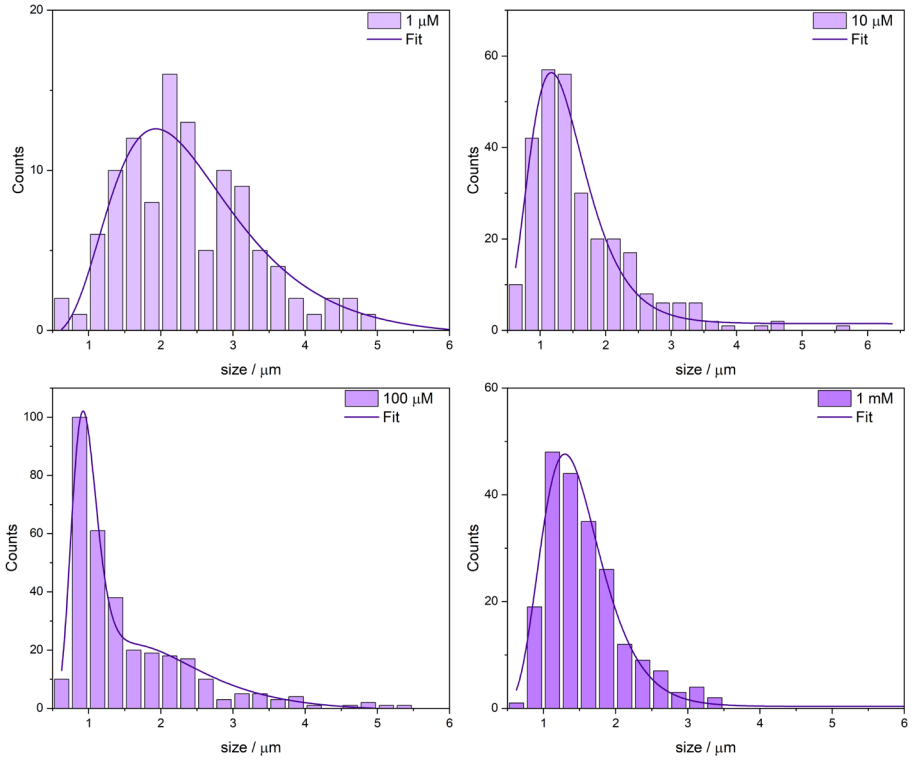



*Figure 5.37. Droplets size distribution acquired at 24 hours times for different spermine concentrations after fibril formation (bars) and their respective LogNormal fits (lines).*

We also studied if by incubating samples at different times the effect spermine presence had on fibrillation was stable. As shown in Figure 5.38, it does not appear to exist a significant variation between the shorter incubation times and the longer ones, as the decrease on fibril formation can be seen in all studied conditions. Condensates appear both at 24 hours and 1 week incubation after at around 1 µM of spermine, which suggests that the spermine can reverse fibril formation and promote LLPS even at longer periods of time. However, even though fibrils do become less when spermine concentration increases at 72 hours incubation, no droplets can be seen on the obtained images. As the quantity of formed fibrils presents the same tendency at all studied conditions, it should be expected that droplets also have the same behavior in all experiments. Thus, the lack of droplets at 72 hours could be due to a bad focus when acquiring the images.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

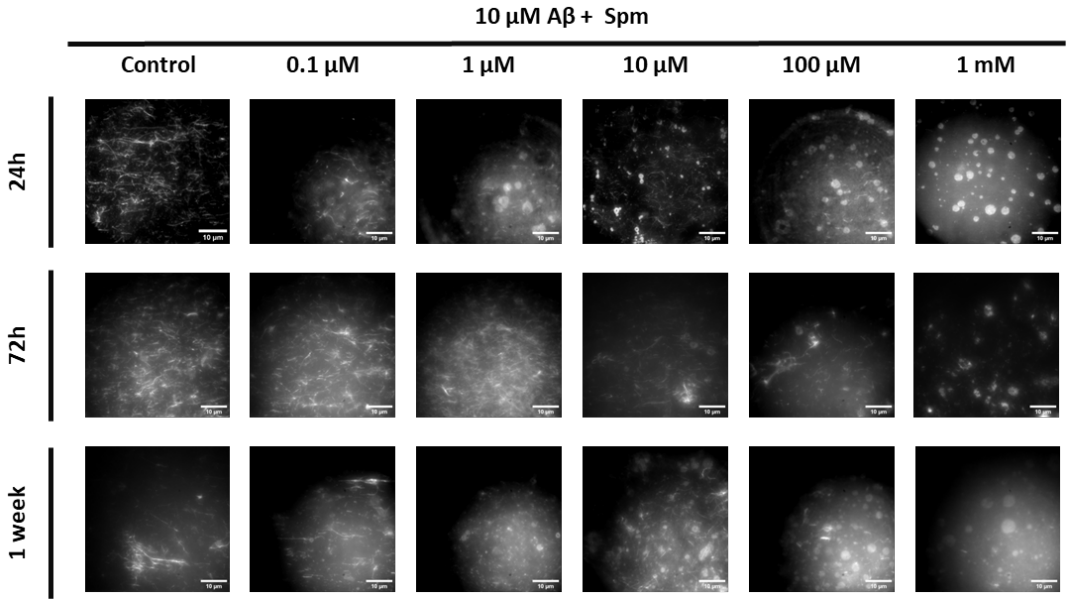



*Figure 5.38. Fluorescence microscopy images of Aβ42 samples with different spermine concentration obtained after being incubated for 24 hours, 72 hours and 1 week. Scales for all images are equal to 10 µm.*

The tendencies that were suggested by visual analysis of the images are corroborated by the analysis of the percentage of formed fibrils and droplets present in each of the samples, which is shown in Figure 5.39. As expected, fibril's percentage decreases with the more spermine is added to the sample with similar tendencies for all incubation times. For 24 hours and 1-week samples (Figure 5.39a and c) this tendency is accompanied by an increase in the percentage of droplets with spermine concentration. However, that is not the case for the 72 hours experiment (Figure 5.39b). Even though the obtained images did not show a significant increase in droplets at 72 hours incubation, it does seem that the droplet formation is somehow happening as the higher spermine concentration do show a slight increase in droplet's percentage. As mentioned before, the fact that fibrils present the expected tendency hints towards an incorrect data acquisition that did not allow for droplets to be seen, and therefore this experiment must be repeated.

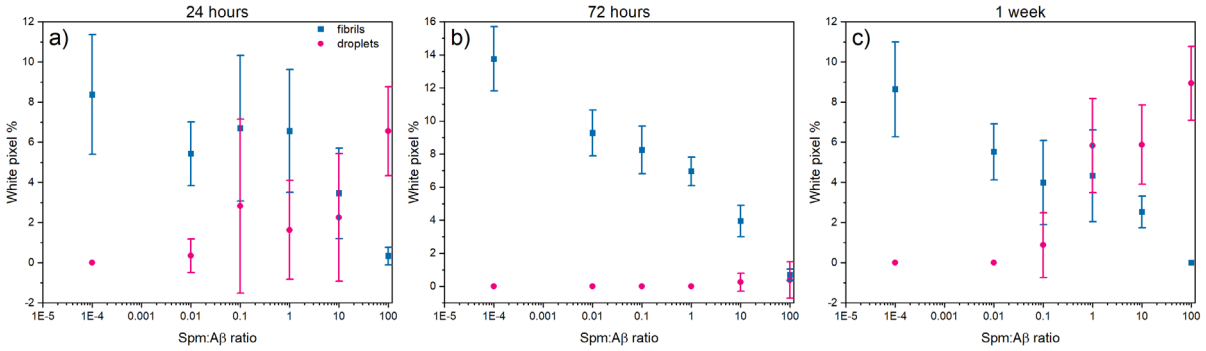



*Figure 5.39. Percentage of white pixels corresponding to fibrils (blue squares) and droplets (pink dots) variation with spermine concentration at different incubation times. a) 24 hours incubation, b) 72 hours incubation, c) 1 week incubation.*

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

We also obtained the droplet's size distribution for the samples with 1 week of incubation and constructed its histogram at different concentrations of cation (Figure 5.40). At 1 µM of spermine the formed droplets show a narrow distribution with no condensates above 3 µm (Figure 5.40a). The distribution broadens and shifts towards higher sizes reaching even 5 µm when spermine concentration is 10 µM (Figure 5.40b). A similar distribution to the one at 10 µM was obtained for 100 µM and 1 mM (Figure 5.40c and d). Droplet's size after 1 week incubation seem to grow when adding spermine into the media, contrary to what was seen for the 24 hours incubation (Figure 5.37).

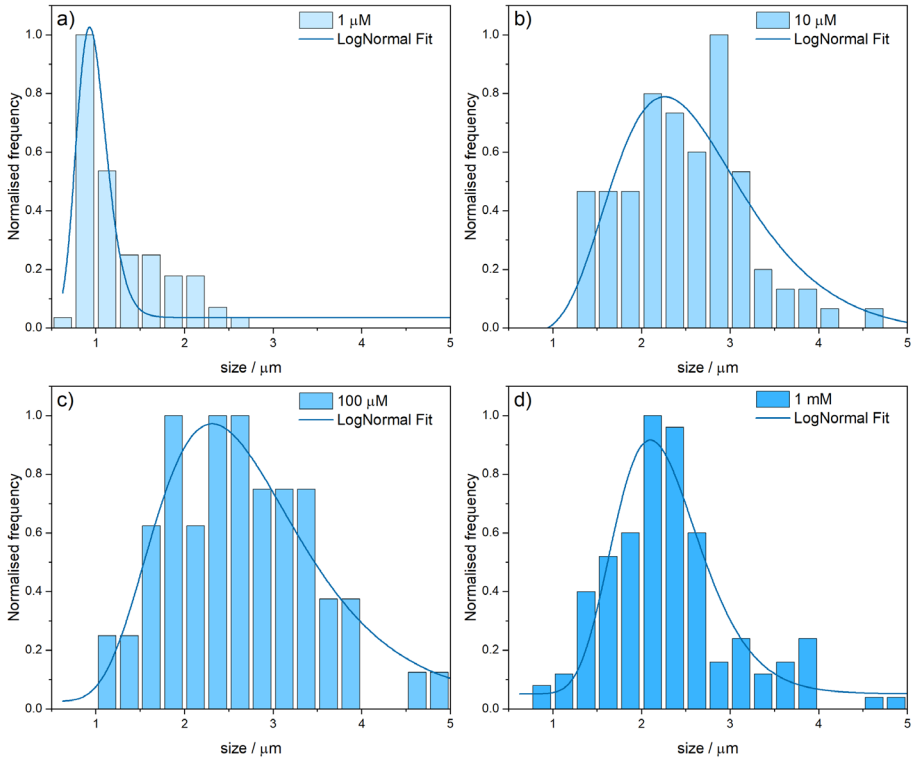



*Figure 5.40. Droplets size distribution acquired at 1 week incubation for different spermine concentrations after fibril formation (bars) and their respective LogNormal fits (lines).*

Additionally, after fitting the data with a LogNormal equation we obtained that the mean size shifts from 0.96 ± 0.03 µm, which is smaller than expect, at 1 µM of spermidine to 2.21 ± 0.06 µm when spermine concentration is 1 mM. The intermediate conditions also resulted in bigger droplets sizes: 2.51 ± 0.09 µm for 10 µM, and 2.60 ± 0.09 µm for 100 µM. Droplet's growth has been reported in the literature when studying protein aggregation through LLPS on other molecules, and it also was reported by Prof. Lee's group when studying the effects of small charged biomolecules on tau-p protein aggregation.[77,185] Usually, as the formed aggregates become bigger so do the droplets in which aggregation is taking place, as the nucleation phase finish and the elongation begins. However, in our studies aggregation and fibrils formation had already happened when spermine was added to the samples and, consequently, it should not undergo LLPS in the aggregation direction. One possible explanation to this behaviour is that at longer incubation times

the degree at which LLPS is reversed at low spermine concentration is higher than what was seen for 24 hours, which is corroborated by the lower droplet size obtained for 1 μM of spermine at 1 week incubation. If longer incubation needs less spermine concentration to reverse fibrillation, adding more positive charge than needed to the solution could have the reverse effect and promote again amyloid aggregation.

***Spermidine***

We then carried out the same procedure as for spermine changing the polyamine to spermidine. Spermidine presents at physiological pH a positive net charge of +3 (Figure 5.41). From the previously obtained result we expected to see again a variation on fibril formation upon increase of spermidine concentration. However, as shown in Figure 5.41, super-resolution images showed only the presence of fibrils at all the studied concentrations. SRRF images also shows that no variation on fibrils concentration seems to be happening at any spermidine concentration.

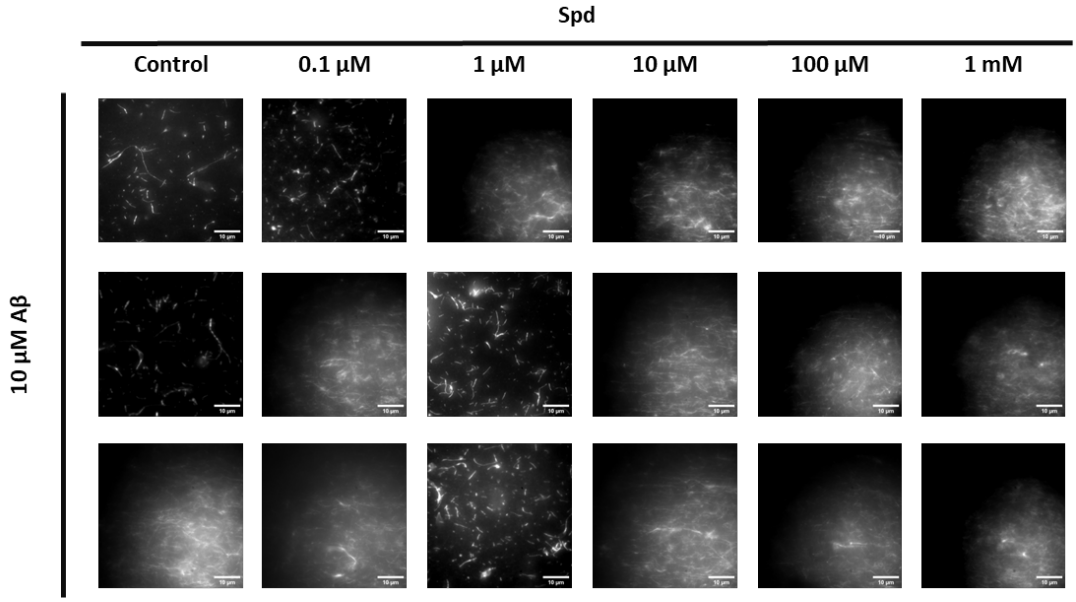



*Figure 5.41. Fluorescence microscopy images of Aβ42 samples incubated during 24 hours at 37 ºC with different spermidine concentrations that were added after fibrillation had occurred. Scales for all images are equal to 10 μm.*

Because spermidine presents a weaker positive net charge, one could suggest that its impact could also be lower. After confirming that even at concentrations of 1 mM there is no influence on fibril formation, we decided to very if the inhibition could take place with longer incubation times. Figure 5.42 shows the comparison between obtained images for all spermidine concentrations at incubation times ranging from 24 hours to 1 week. Again, a visual analysis of said images indicates that droplet formation is not taking place.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

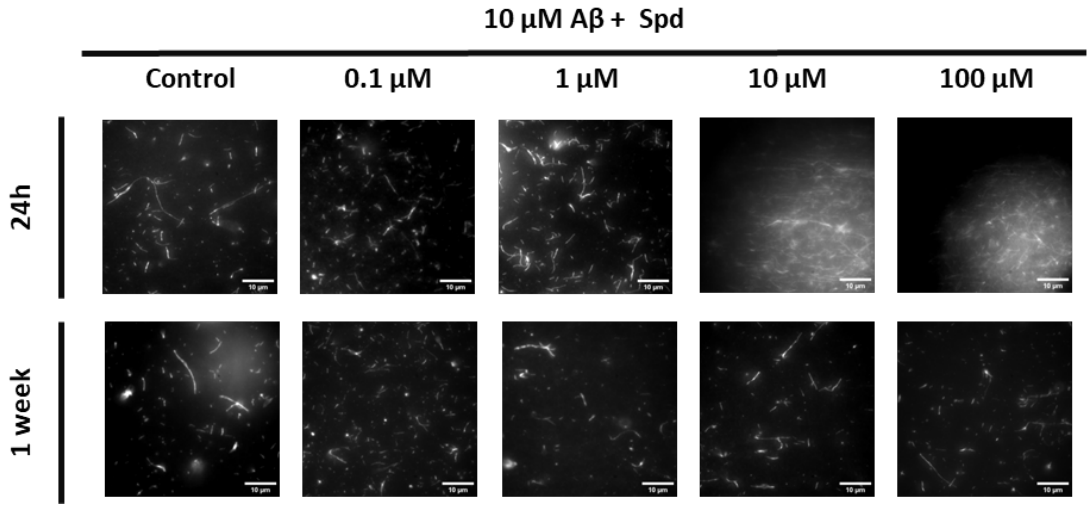



*Figure 5.42. Fluorescence microscopy images of Aβ42 samples with different spermine concentration obtained after being incubated for 24 hours, and 1 week. Scales for all images are equal to 10 µm*

Moreover, when analysing the percentage of white pixels that corresponded to fibrils in the images (Figure 5.43), we saw that there is little to no variation in the number of formed fibrils in the samples. Even though there is a slight fluctuation in the mean values, considering their standard deviation the variation is insignificant. This corroborates that no change takes place when the used cation is spermidine. It is important to remark that experiments were repeated to corroborate the results, as a poor focus of the microscope when acquiring images could lead to no droplets being seen.

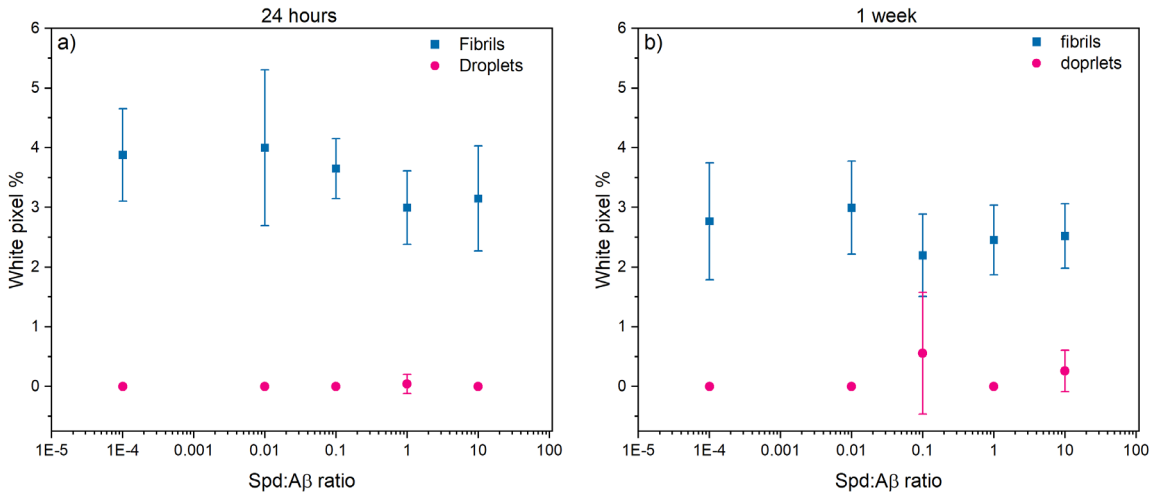



*Figure 5.43. Percentage of white pixels corresponding to fibrils (blue squares) and droplets (pink dots) variation with spermidine concentration at different incubation times. a) 24 hours incubation and, b) 1 week incubation.*

***Magnesium chloride***

We also performed the same studies with magnesium chloride as our cation. Magnesium is also an important ion in biological systems. In this case, the net charge is +2. Figure 5.44 shows the acquired images for different magnesium concentrations after incubating samples with already formed fibrils for 24 hours. As per the case of spermidine, there seems to be no significant droplet formation when the added cation is magnesium. However, super-resolution images at concentrations above 100 µM do seems to present less fibrils than at lower concentrations. In addition, at the higher studied concentrations, contrary to the case of spermine when fibril formation decrease was followed by the formation of condensates that could easily be distinguished with a visual analysis, that does not seem to be the case for magnesium.

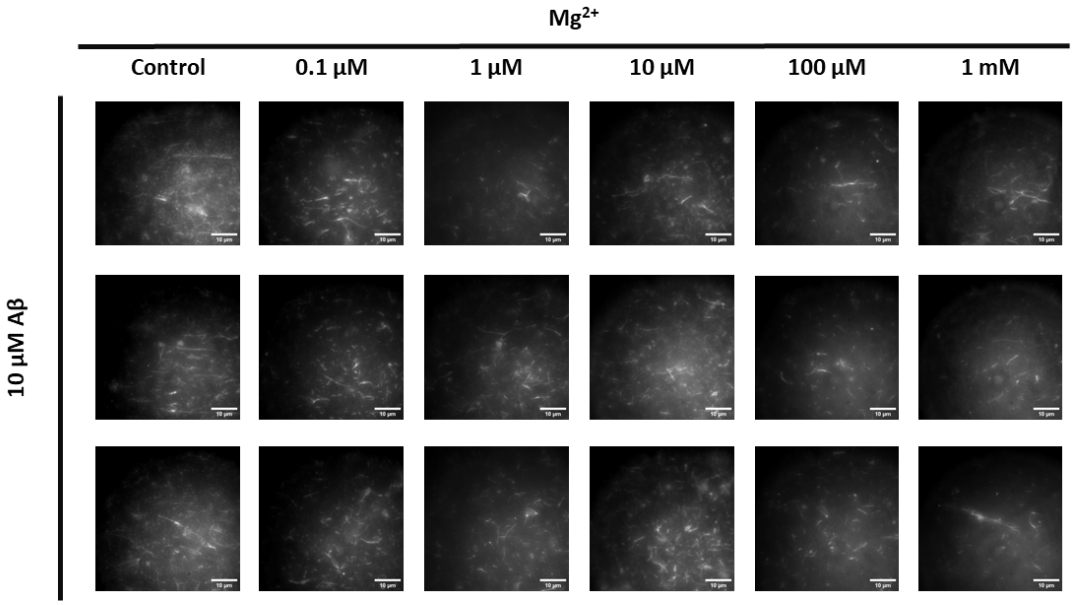



*Figure 5.44. Microscopy images of Aβ42 samples incubated during 24 hours at 37 ºC with different magnesium concentrations that were added after fibrillation had occurred. Scales for all images are equal to 10 µm.*

If we increase the incubation rate to 72 hours for the same $MgCl_2$ concentrations (Figure 5.45), droplet still do not appear at any of the studied concentrations. However, there does seem to be a decrease in fibril formation at the higher magnesium levels, which was also seen for the 24 hours incubation period. 1 week incubation data, on the other hand, does show the formation of droplets at concentrations equal or higher than 10 µM. The appearance of droplets occurs in association with lower fibrillation.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

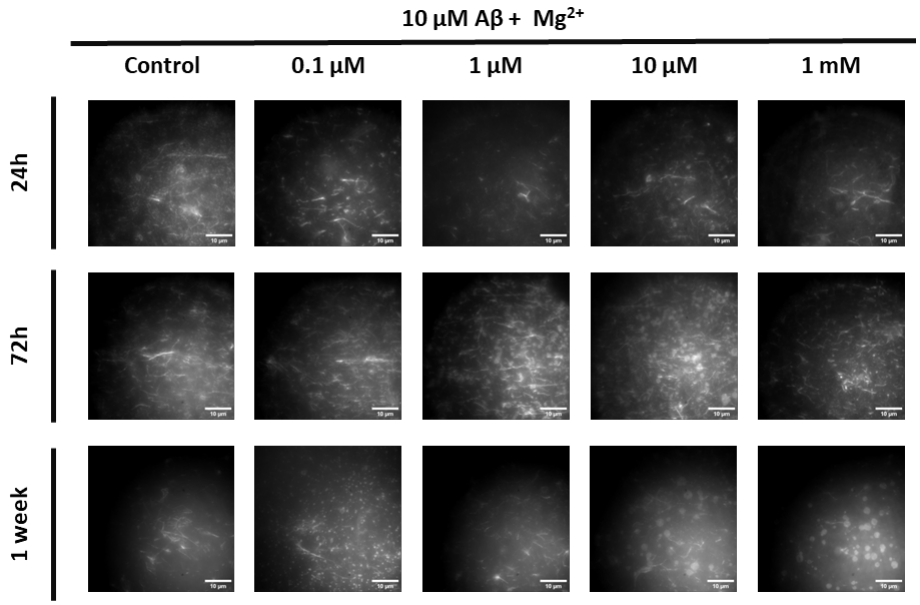



*Figure 5.45. Microscopy images of Aβ42 samples with different magnesium concentrations incubated at 37 ºC for 24 hours, 72 hours and 1 week. Scales for all images are equal to 10 µm.*

The variation of fibrils and droplets percentage (Figure 5.46) for magnesium helps corroborate the initial results from the visual analysis of the microscopy images. At 24 hours incubation (Figure 5.46a), there is a first decrease in the presence of fibrils between having no magnesium in the samples and adding the lowest concentration. However, between the lowest $Mg^{2+}$:Aβ ratio fibril's percentage stays somehow stable until a $Mg^{2+}$:Aβ ratio of 1 (10 µM $Mg^{2+}$). When even more magnesium is added to the sample (cation:amyloid ratio of 10, 100 µM $Mg^{2+}$) another sharp drop of the fibril's percentage appears, which is also accompanied by a very low increase of droplets percentage. These droplets cannot easily be perceived in the images shown in Figure 5.44. In the case of 72 hours incubation (Figure 5.46b), the decrease in fibril's percentage happens between having no magnesium and the first added concentration. Then, it stabilizes and shows no variation with concentration. Alongside fibrils decrease, an increase in droplet formation from 1 µM to 100 µM. Droplets seem to be forming until around 10 µM, where its percentage stabilizes at approximately 2 %. After 1 week of incubation (Figure 5.46c), droplet formation happens at 0.1 µM and increases with concentration. Additionally, the obtained percentages at 1 week incubation are higher than for 24 and 72 hours, which could also explain why droplets could not be distinguish in a first visual analysis at this incubation times. However, the degree at which droplets are formed seems to be lower than for the case of spermine.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

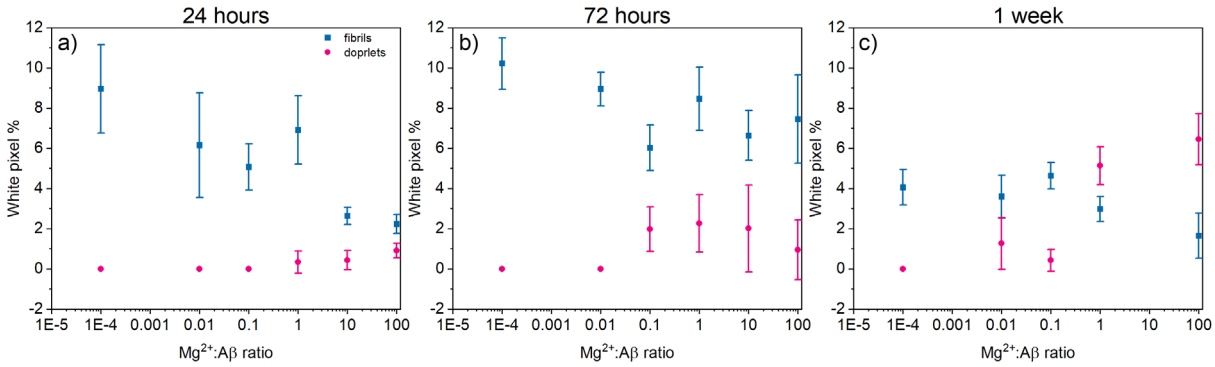



*Figure 5.46. Percentage of white pixels corresponding to fibrils (blue squares) and droplets (pink dots) variation with magnesium chloride concentration at different incubation times. a) Results for 24 hours incubation, b) results for 72 hours incubation, c) results for 1 week incubation.*

Even though spermidine has a higher net charge than magnesium chloride does, spermidine results showed that its presence has less impact on fibrillation than magnesium did even at the longer incubation times. Surprisingly, spermidine had no influence whatsoever on fibril formation and it did not promote LLPS neither, but magnesium did promote it, especially after 1 week of incubation. This suggests that charge is not the only factor involved in LLPS taking place after fibril formation. We must consider that polyamines are bigger molecules than metal ions and, therefore, its interactions could be also hindered by its size. Spermine and spermidine are similar in size, which allows us to conclude that higher positive charges are need for LLPS to be promoted after fibrillation. However, for the case of magnesium even though the net charge is lower so is its size, which could lead to more molecules being able to interact with amyloids reversing fibrillation.

### 5.4.2.2. Influence before fibril formation

As electrostatic interactions between amyloids and the studied cations could also take place before aggregation, we also studied the impact the presence of these charged molecule had on amyloid in samples that were not yet fully aggregated and, therefore, presented no fibrils. Because spermine had the greatest impact on promoting LLPS on already formed fibrils, we decided to carry out this last study using spermine.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

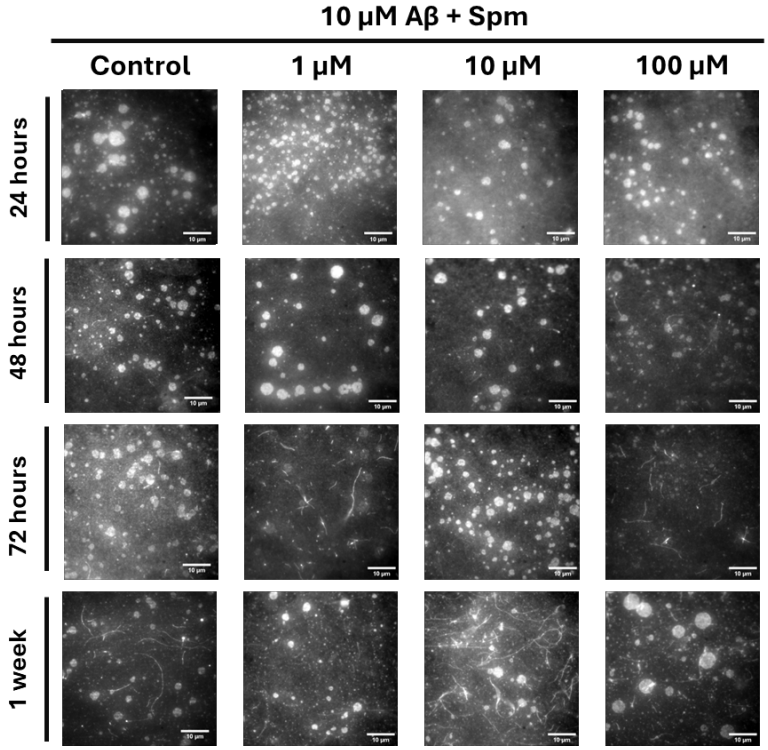



*Figure 5.47. Microscopy images of Aβ42 samples with different spermine concentrations incubated at 37 ºC for 24 hours, 72 hours and 1 week before fibril formation. Scales for all images are equal to 10 μm.*

Figure 5.47 shows the obtained SRRF microscopy images from spermine samples with different concentrations incubated from 24 hours to 1 week. As these studies were done before aggregation was completed, control samples at all incubation conditions have a significant presence of condensate formation. Some fibrils can be seen at the longest incubation period, which is coherent with aggregation taking place over time. 24- and 48-hours incubation samples do not seem to present any variation upon spermine addition to the media, as only droplets are seen at almost concentrations. Fibrils in these conditions appear only when spermine concentration are around 100 μM at 48 hours incubation. If samples are incubated for 72 hours instead, fibril formation starts to appear at lower spermine concentrations. Data from 100 μM of spermine at this incubation period shows mainly the presence of fibril and almost no droplets can be seen. However, the tendency is not clear under the studied conditions. As for the case of 1 week incubation, spermine addition seems to lead to more droplet formation than fibrils. This could happen as under these conditions, some fibrils are already formed and, therefore, the aggregation degree is similar to the one existing in the previous studies, when spermine was added before fibrillation started.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

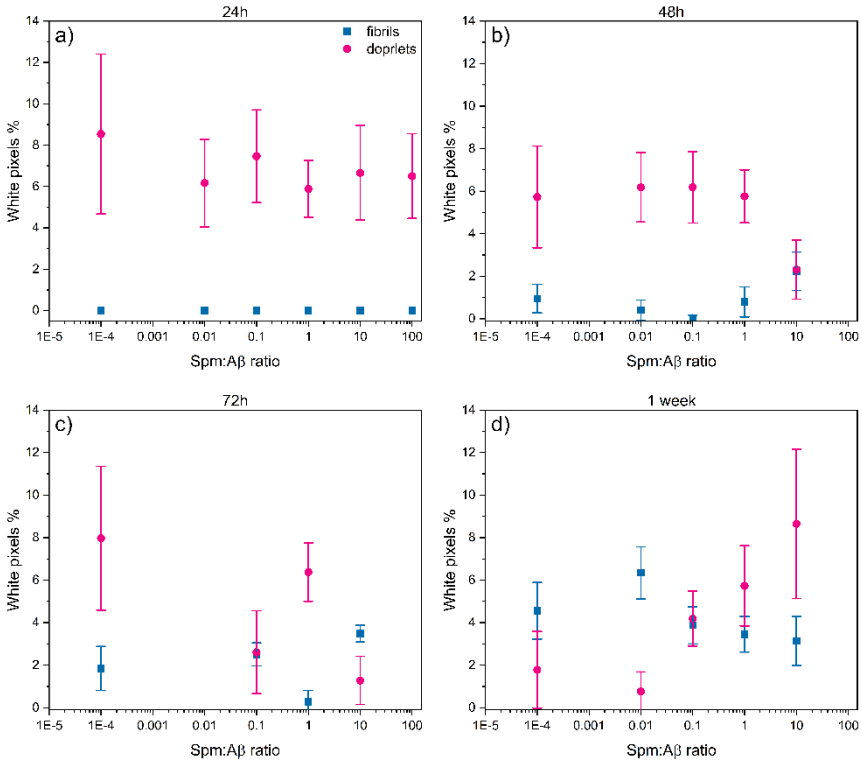



*Figure 5.48. Percentage of white pixels corresponding to fibrils (blue squares) and droplets (pink dots) variation with spermine concentration at different incubation times. a) Results for 24 hours incubation, b) results for 72 hours incubation, c) 48 hours incubation and, d) results for 1 week incubation.*

To better understand the tendencies that appear in the super-resolution images, percentage variation with spermine concentration were calculated and are presented in Figure 5.48 for all incubation times. The obtained values corroborate that spermine addition does not involve any variation in aggregation at 24 hours of incubation, neither promoting it nor reversing it, as both droplets and fibrils percentages are constant at all concentrations (Figure 5.48a). There seems to be an influence when incubation is extended to 48 hours (Figure 5.48b), where droplet's percentage start to decrease at spermine concentrations of around 10 μM while fibril's percentage presents a raise at said conditions. However, the tendency is not so evident in the case of 72 hours incubation (Figure 5.48c) as there is some oscillation in the intermediate concentrations. Still, these results suggest that droplets are being transformed into fibrils upon spermine addition. On the other hand, the tendency at 1 week's incubation is evident as is opposing to the other incubation times (Figure 5.48d), as droplets are percentages rise with spermine concentration and fibrils decrease. These variations support the idea that 1 week incubation samples present an aggregation degree similar to the samples where spermine was added after fibril formation had started.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

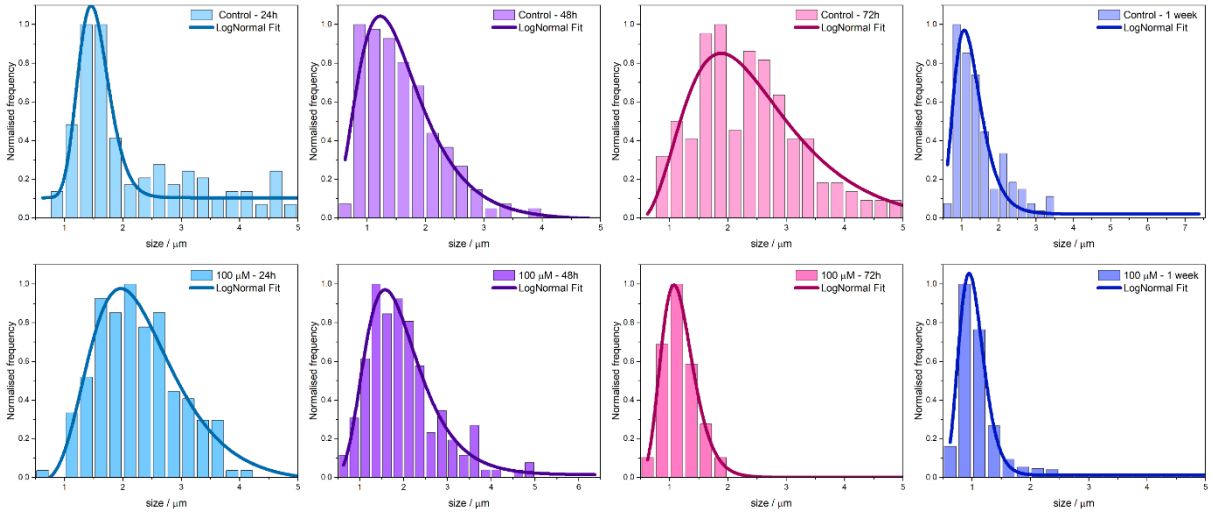



*Figure 5.49. Size distribution histograms at different incubation times and spermine concentrations added before fibrils formation. Up: Control samples at 24h, 48h, 72h and 1 week incubation times respectively. Down: 100 µM samples incubated for 24h, 48h1 72h and 1 week. The bars are the experimental data, and the solid lines are the LogNormal fit results of each histogram.*

Now, if we study the size distribution of the formed droplets at the different incubation times (Figure 5.49), we can see that all distributions range from 1-5 µm. If we compare all control samples at different incubation periods, we can see that the distributions seem to get narrowed when increasing the time of incubation. This behavior is also seen in the case of 100 µM at different times. On the other hand, if we compare the same incubation times at different spermine concentrations we see that at 24 hours and 48 hours incubation the droplets sizes present wider distributions upon increasing spermine concentration, whereas at 72 hours and 1 week the opposite takes place. This could imply that spermine does have an impact in amyloid aggregation at all incubation times, however, to which direction the aggregation is advancing could vary depending on the incubation time. At 24 hours and 48 hours spermine addition seem to be promoting amyloid aggregation as droplets are growing, even though its amount does not show significant variations. On the contrary, 72 hours incubation would be an intermediate step where spermine addition first promotes fibrillation up to a certain concentration. Exceeding this concentration provokes the opposite effect decreasing the number of formed fibrils. This could explain why the percentage did not present a clear tendency and is supported by the droplets size decreasing at the highest studied spermine concentration. 1 week incubation samples had already undergone fibrillation and, therefore, spermine addition seems to be reversing fibrillation. After fitting the data with a LogNormal function, we obtained the mean sizes of each distribution (Table 5.4). The resulting mean droplet sizes are also in agreement with spermine concentration being able to both promote and hinder amyloid concentration, as sizes increase with concentration for 24 hours and 48 hours while they decrease for 72 hours and 1 week incubation experiments.

USC DE COMPOSTELA

*Table 5.4. Mean droplet sizes obtained from the LogNormal fit of the size histograms at different incubation times for control and 100 μM samples.*

| | 24 hours | 48 hours | 72 hours | 1 week |
|---|---|---|---|---|
| **Control** | 1.51 ± 0.02 | 1.48 ± 0.04 | 2.3 ± 0.1 | 1.2 ± 0.03 |
| **100 μM** | 2.23 ± 0.05 | 1.83 ± 0.05 | 1.15 ± 0.03 | 1.00 ± 0.02 |

## 5.4.3. Amyloid fibrillation and charge-charge interactions

Our results suggest that, although charge-charge interactions between charged biomolecules do seem to have an impact on amyloid aggregation, charge is not the only factor controlling the influence of charged biomolecules. Even though the higher the net charge is, the clearer the influence is, the differences between spermidine and magnesium suggest that the number of charged molecules that can bind to amyloids is also important. In addition, spermine results suggest that amyloid aggregation can be both promoted and hindered by electrostatic interactions and that controlling the cation concentration might be key to determining which direction aggregation occurs. However, we believe that further experiments are needed to further clarify the extent to which electrostatic interactions may interfere with amyloid aggregation, with intermediate incubation times and different sized cations.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# 5.5. 2-cyano-6-hydroxybenzothiazole as fluorescence probe for human serum albumin detection

As the last part of this thesis, the study of a fluorophore from the family of the benzothiazoles (CBTOH) as tool for human serum albumin detection was performed. In this section we present the obtained results, while the experimental conditions can be found in section 3.1.1 and the sample preparation in section 3.3.5.

## 5.5.1. CBTOH characterisation in physiological conditions

The characterisation study of CBTOH published by our group studied the acid-base properties of CBTOH obtaining a $pK_a$ for hydroxyl group of the molecule of 8.01 ± 0.02 and estimating a $pK_a$ for the amino group lower than 0.[133] However, it must be noted that the study was done in aqueous solution. Thus, the ionic strength of the samples was not equal to the ones that can be found in physiological conditions. For this reason, we first carried out a study of the CBTOH acid-base properties under physiological conditions of ionic strength. That is, ionic strength of 0.175 M - 0.18 M depending on the pH.

Figure 5.50a shows the variation of the absorption spectra of CBTOH at different pH ranging from 5.4 to 10.1. The spectra show two main absorption maxima at 320 nm and 375 nm, approximately, which are the respective absorption maxima for the two species that can be formed at the studied pH range. Both the peak at 260 nm and the peak at 320 nm decrease in intensity when changing the media from an acidic one into a basic one, whereas the peaks at 280 and 373 nm show the opposite behaviour, increasing their intensity with pH.

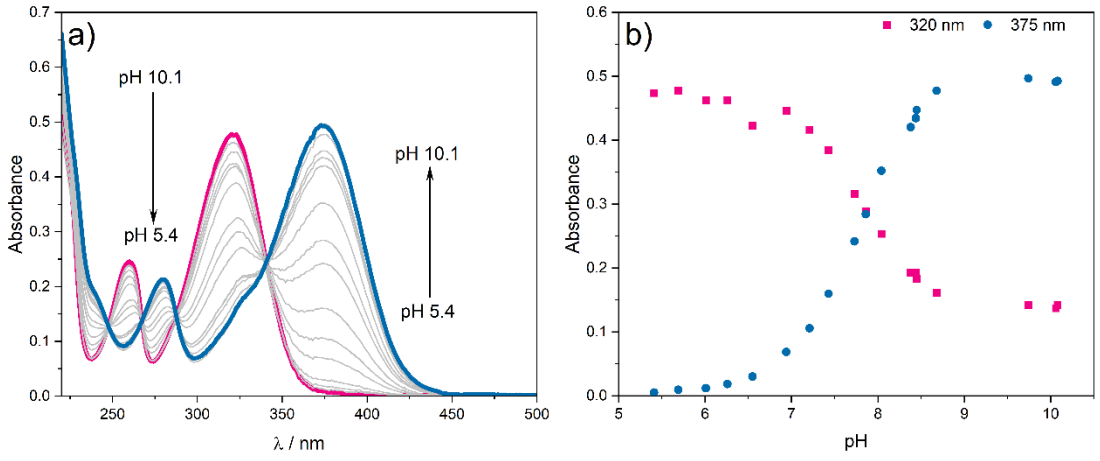
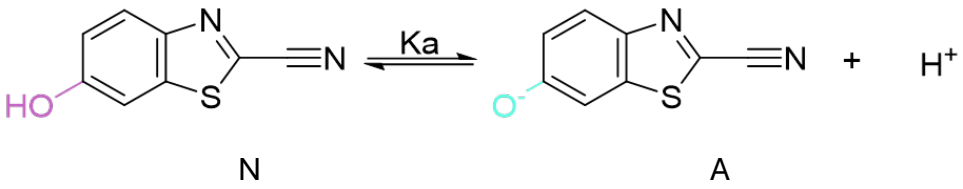



*Figure 5.50. UV-Vis absorption data: a) absorption spectra at different pH for CBTOH at physiological ionic strength conditions (I=0.17M), b) absorbance variation with pH at the absorption maxima of 320 nm (pink squares) and 375 nm (blue dots).*

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

When plotting the absorbance of CBTOH as a function of the pH, a sigmoidal variation can be observed (Figure 5.50b) with clear isosbestic points. This behaviour shows that two species exist in equilibrium under the studied conditions. Thus, we can propose an acid-base equilibrium with two species (Scheme 5.5), a protonated form (**N**) and a hydroxyl- deprotonated form (**A**).




N A

*Scheme 5.5. Schematic representation of the acid-base equilibrium of CBTOH in the ground state.*

Following this mechanism the acidity equilibrium constant $K_a$ for this process would be as shown in Eq (5.2).

$$K_a = \frac{[H^+][A]}{[N]} \tag{5.2}$$

As the total concentration of CBTOH ( $a_0$ ) is constant, the sum of the concentrations of each species is equal to $a_0$ , resulting the concentrations of **N** and **A** as follows:

$$[N] = \frac{a_0[H^+]}{K_a + [H^+]} \tag{5.3}$$

$$[A] = \frac{a_0 K_a}{K_a + [H^+]} \tag{5.4}$$

Considering Lambert-Beer's law and the fact that the absorbance at a given wavelength will be the sum of the absorbance of all species present in solution, for our system we obtain the following expression:

$$A^\lambda = \varepsilon_N^\lambda l[N] + \varepsilon_A^\lambda l[A] \tag{5.5}$$

Where $\varepsilon_i^\lambda$ is extinction molar coefficient of species i at a given wavelength and $l$ is the cuvette's pathway.

If we rewrite Eq (5.5) with Eq (5.3) and Eq (5.4) we obtain the following equation for the absorbance as a function of the proton concentration, where $A_N^{\lambda,0}$ and $A_A^{\lambda,0}$ are the absorbances at the given wavelength when the fluorophore exist only as **N** or **A** in solution, respectively:

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

$$A^{\lambda} = \frac{A_N^{\lambda,0}[N] + A_A^{\lambda,0}K_a}{K_a + [H^+]} \quad (5.6)$$

Using Eq (5.6), we applied PCGA[157,187] (see **Data Analysis**) to analyse the variation of the absorption spectra with the pH and obtained a $pK_a$ value of 7.751 ± 0.002 for CBTOH under physiological conditions. The significantly lower value of $pK_a$ with respect to that obtained in aqueous solution can be explained by the higher ion strength present. We also obtained the pure spectra for the two species that are formed under the studied conditions (Figure 5.51). For the **N** form, the absorption spectrum presents two peaks at 260 nm and 320 nm, while for the anion **A**, the spectrum shows both peaks shifted to higher wavelengths (280 nm and 373 nm, respectively). The obtained pure spectra are in agreement with the spectra reported by our group.[133]

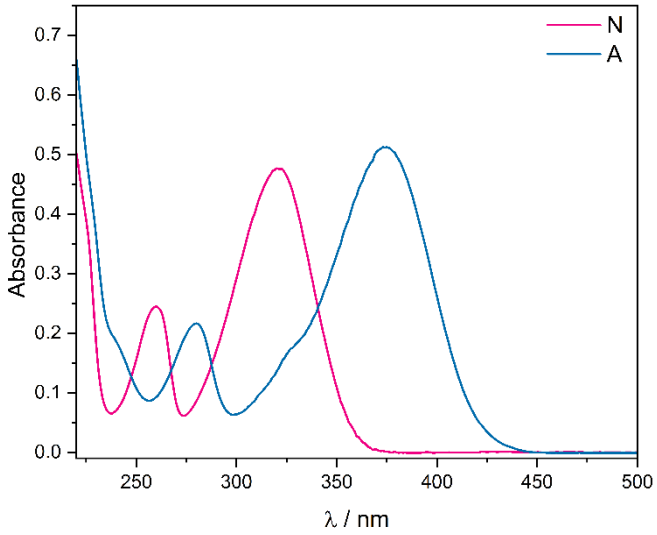



*Figure 5.51. PCGA Results: pure absorption spectra for species N (pink) and A (blue) obtained by PCGA analysis.*

The decrease in the $pK_a$ of the dye upon increasing ionic strength of the solution entails that under physiological conditions of pH and ion concentration, CBTOH will be present in solution under its protonated and deprotonated forms almost in a 1:0.5 **N:A** ratio, as the physiological pH value is of approximately 7.45.

## 5.5.2. Host-guest association studies

### 5.5.2.1. Emission and excitation measurements

To study the binding process between CBTOH and blood plasma serum albumin protein, we first carried out the studies both with BSA and HSA. First, we performed steady-state fluorescence studies and obtained the emission and excitation spectra of CBTOH at different concentrations of protein, ranging from 0 to 300 μM. Figure 5.52 shows the obtained emission spectra at three different excitation wavelengths (375 nm, 320 nm and

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

345 nm). At all studied excitation wavelengths CBTOH presents and emission spectra that varies in intensity and band position with the addition of protein (Figure 5.52).

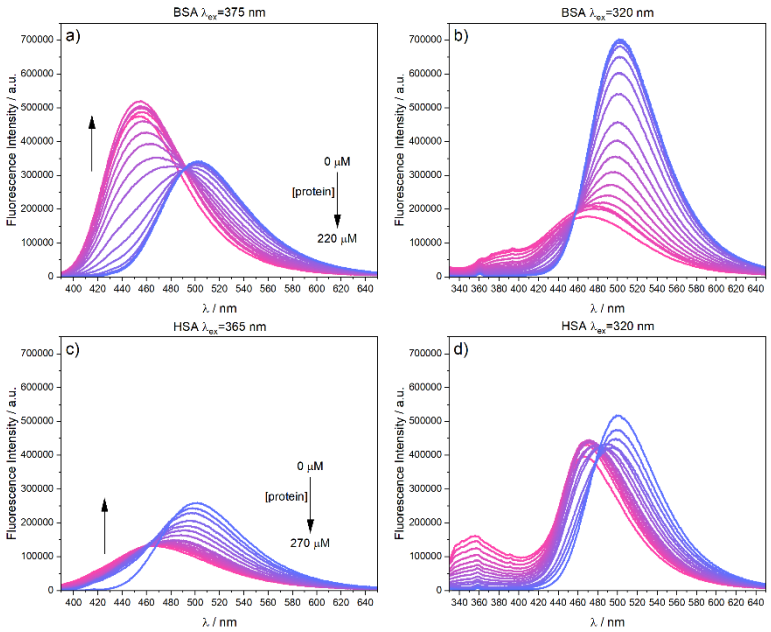



*Figure 5.52. CBTOH emission spectra at different concentrations of protein bovine or human serum albumin. Upper panels: a-b) added protein is BSA and the excitation wavelengths are 375 nm and 320 nm, respectively. Lower panels: c-d) added protein is HSA and excitation wavelengths are 365 nm and 320 nm in said order.*

Figure 5.52a-b shows the emission spectra for the titration of CBTOH with BSA. In the absence of BSA, when all fluorophore is free in solution, the emission band appears at around 500 nm independently of the excitation wavelength. This band must be attributed to the emission of the anionic species **A***, directly excited at 375 nm (Figure 5.51) or formed by photodissociation of the neutral species **N*** when exciting at lower wavelengths, as reported in a previous work.[133] Upon addition of BSA the emission band shifts to lower wavelengths, stabilizing at high protein concentrations. Thus, the new band at 450 nm should correspond to the complex formed between the CBTOH and the BSA (**A:BSA***). If we use the **N** absorption maximum as excitation wavelength (320 nm, Figure 5.51) the obtained spectra (Figure 5.52b) present three overlapping emission bands: the band at 500 nm of the free anion **A***, the complex **A:BSA*** band around 450 nm and a third band at about 380 nm. The fact that the third emission band appears only when exciting **N** and not **A** indicates that it is due to the emission of the complex between **N*** and the protein (**N:BSA***). Thus, the efficient excited-state proton transfer that neutral CBTOH suffers when unbound must be somehow inhibit by the binding to the protein.

For the case of association with HSA, shown in Figure 5.52c-d, similar behaviour as with BSA is observed. When exciting the **A** form ($\lambda_{ex}$=365 nm) only the emission band at 500 nm appears in the absence of HSA due to **A*** emission. This band shifts to 460 nm and decreases its intensity as the HSA concentration is increased. This second band must be

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

attributed to the complex **A:HSA*** and is slightly red-shifted with respect to that of **A:BSA***. When changing the excitation wavelength to 320 nm (Figure 5.52d), where only neutral species **N** absorbs, a new band appears at about 360 nm whose intensity is proportional to protein concentration. As in the case of BSA, this third band must correspond to the complex between the neutral CBTOH and HSA (**N:HSA***), which is slightly blue-shifted with respect to that of the **N:BSA*** complex (Figure 5.52).

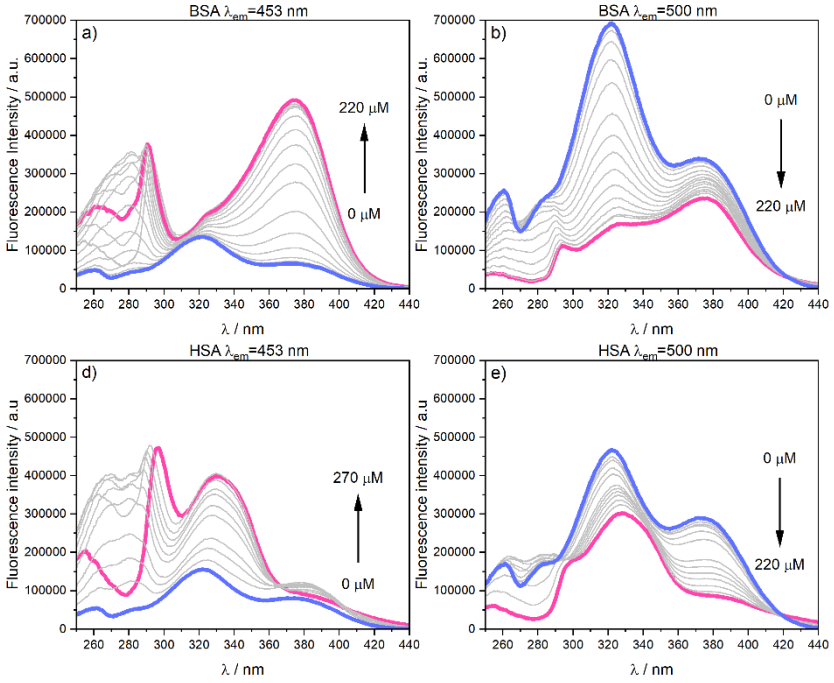
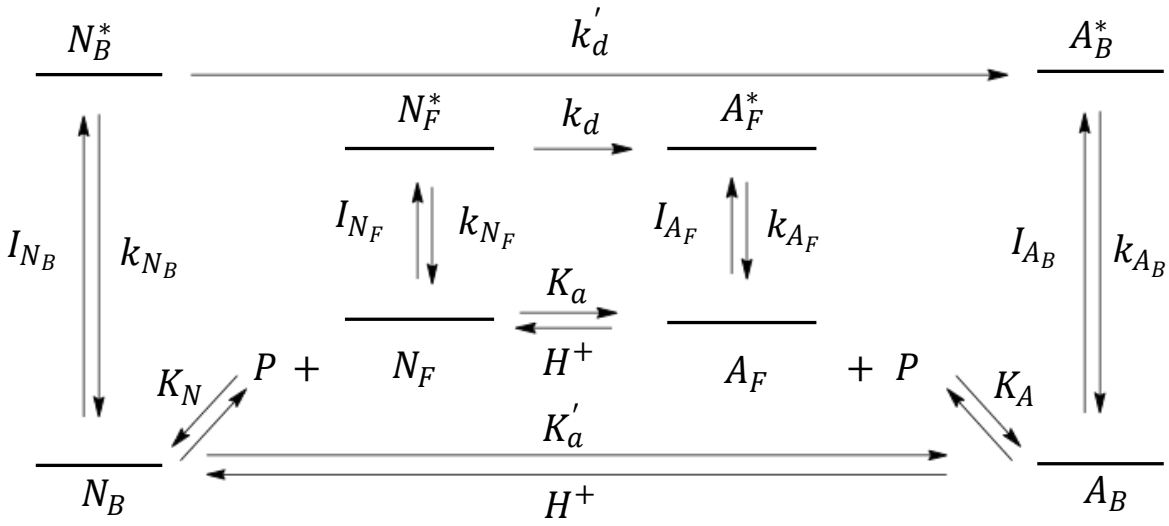



*Figure 5.53. CBTOH excitation spectra at different concentrations of protein bovine or human serum albumin. Upper panels: a-b) added protein is BSA and the emission wavelengths are 453 nm and 500 nm, respectively. Lower panels: c-d) added protein is HSA and excitation wavelengths are 460 nm and 500 nm respectively.*

Excitation spectra of CBTOH at different concentrations of BSA (Figure 5.53a-b) shows two main bands at 320 nm and 375 nm, related to **N** and **A** respectively, that vary in intensity when increasing protein concentration. Additionally, there is a spectral shift in both bands towards higher wavelengths that can be attributed to the ground-state association of the dye with the protein. If the emission wavelength is set to the emission maximum of the complex **A:BSA*** (453 nm, Figure 5.53a), the intensity increases with protein concentration, whereas when setting it to the maximum of **A*** (500 nm, Figure 5.53b) it presents a lowering tendency. This behaviour corroborates that the emission band at 453 is, in fact, caused by the presence of **A:BSA*** in solution. The fact that the absorption band of N decreases in contribution as the protein concentration increases verifies that the proton transfer process of **N*** is restricted when it is bound to BSA, in the complex **N:BSA***. In the case of HSA (Figure 5.53c-d), excitation spectra also showcase a spectral shift and intensity dependency with protein concentration with similar explanation as in BSA. However, both for an emission wavelength of 460 nm (Figure 5.53c) and for 500 nm (Figure 5.53d), the main excitation band at high protein concentrations is

the band at around 325 nm, which is attributed to the complex **N:HSA***. Although **N:HSA*** is expected to have some contribution to 460 nm emission, these results suggest that photodissociation is still happening when CBTOH is bound to HSA.

Considering all the exposed above we can propose the following photophysical mechanism to explain the association of species **N** and **A** with serum albumin (Scheme 5.6). In this mechanism, in the absence of protein, CBTOH can exist both as free species $\boldsymbol{N_F}$ and $\boldsymbol{A_F}$ in the ground state and the corresponding acid-base equilibrium constant is $K_a$. Besides the deactivation of these species in the excited state, photodissociation of $\boldsymbol{N_F^*}$ to yield $\boldsymbol{A_F^*}$ takes place with a rate constant $k_d$. In the presence of serum albumin proteins ($\boldsymbol{P}$), both species of CBTOH can bind to them forming the bound species $\boldsymbol{N_B}$ and $\boldsymbol{A_B}$ (previously denoted by N:BSA or N:HSA and A:BSA or A:HSA) with association equilibrium constants $K_N$ and $K_A$ **,** respectively. The described ground-state equilibria are connected by the acidity constant of the bound species, denoted by $K_a^{'}$. In the excited state the formed complexes are emissive species with deactivation constants $k_{N_B}$ and $k_{A_B}$ and $\boldsymbol{N_B^*}$ is also proposed to undergo photodissociation upon excitation with a rate constant $k_d^{'}$.




*Scheme 5.6. Proposed photophysical mechanism for CBTOH and CBTOH bound to blood plasma proteins.*

From the analysis of the fluorescence data with the derived model (**Data Analysis**, Eq 4.107), we can obtain the apparent association constant of the 1:1 CBTOH-protein binding $K_{ap}$ (Eq (5.7)). The apparent constant is a function of both the association constant when CBTOH binds to **N** ($K_N$) and the association constant when it binds to **A** ($K_A$), considering as well the proportion of each species in solution with the $R_H$, which we defined in section 4.5.5 as $K_a/[H^+]$.

$$K_{ap} = \frac{K_N + R_H K_A}{1 + R_H} \tag{5.7}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

For a $R_H = 0.5$, which corresponds to a pH of 7.45, we obtain the following expression:

$$K_{ap} = \frac{1}{3}K_A + \frac{2}{3}K_N \tag{5.8}$$

We first carried out a global analysis in Origin with a few selected emission wavelengths, Figure 9.5 shows the experimental data and the resulting fit for the association between CBTOH and both proteins. For the CBTOH:BSA complex (Figure 5.54a-b) we obtained a value of $K_{ap}=(3.2 \pm 0.1) \times 10^4$ $M^{-1}$ for the data at $\lambda_{ex}$= 320 nm, and $K_{ap}=(3.8 \pm 0.3) \times 10^4$ $M^{-1}$ at $\lambda_{ex}$= 375 nm. For the CBTOH:HSA complex (Figure 5.54c-d) the global analysis at different excitation wavelengths resulted in an $K_{ap}=(2.6 \pm 0.3) \times 10^4$ $M^{-1}$ at $\lambda_{ex}$= 320 nm, and a $K_{ap}=(2.38 \pm 0.09) \times 10^4$ $M^{-1}$ at $\lambda_{ex}$= 365 nm for the same sample series.

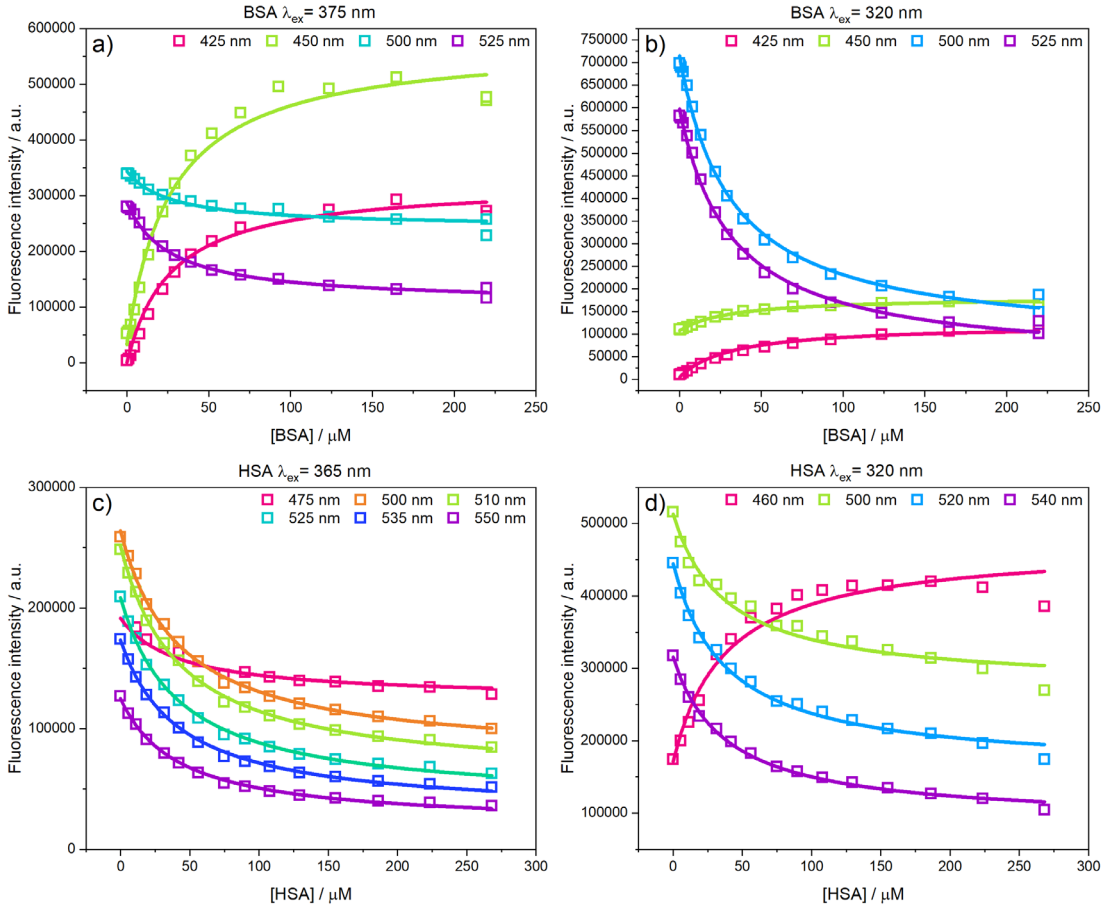



*Figure 5.54. Emission intensity variation with protein concentration at different excitation and emission wavelengths. a) Protein is BSA and excitation wavelength is 375 nm, b) protein is BSA and excitation wavelength is 320 nm, c) protein is HSA and excitation wavelength is 365 nm and, d) protein is HSA and excitation wavelength is 320 nm.*

Giving the discrepancies observed in the first global analysis, we performed a PCA that evidenced the need for an additional third component. After the PCA analysis we noted that the emission spectra at high concentrations of serum albumin was slightly affected by the protein autofluorescence. For this reason, we had to modify the model adding a parameter proportional to protein concentration (C in Eq (5.9)).

USC UNIVERSIDADE DE SANTIAGO

$$F^{\lambda} = \frac{a^{\lambda} + b^{\lambda}[P]}{1 + K_{ap}[P]} + C \tag{5.9}$$

With the modified model we performed the PCGA using the full emission spectra and we obtained the pure spectra for each species shown in Figure 5.55. In both cases, PCGA with the added parameter resulted in a satisfactory fit of the data. For BSA (Figure 5.55a-b) the fit resulted in $K_{ap}=(1.53 \pm 0.01) \times 10^4$ $M^{-1}$ at excitation wavelength of 375 nm and in $K_{ap}=(2.33 \pm 0.03) \times 10^4$ $M^{-1}$ when excitation was set to 320 nm. In the case of HSA the obtained values were of $K_{ap}=(3.65 \pm 0.08) \times 10^4$ $M^{-1}$ and $K_{ap}=(2.93 \pm 0.02) \times 10^4$ $M^{-1}$ at 365 nm and 320 nm excitation wavelengths, respectively (Figure 5.55c-d). The obtained values, although coincident in the order of magnitude, are different between both proteins, having HSA a slightly higher apparent constant than BSA. The variations observed between the different excitation wavelengths in both cases can be explained by the influence of the protein emission which is more significant when excitation wavelength is 320 nm.

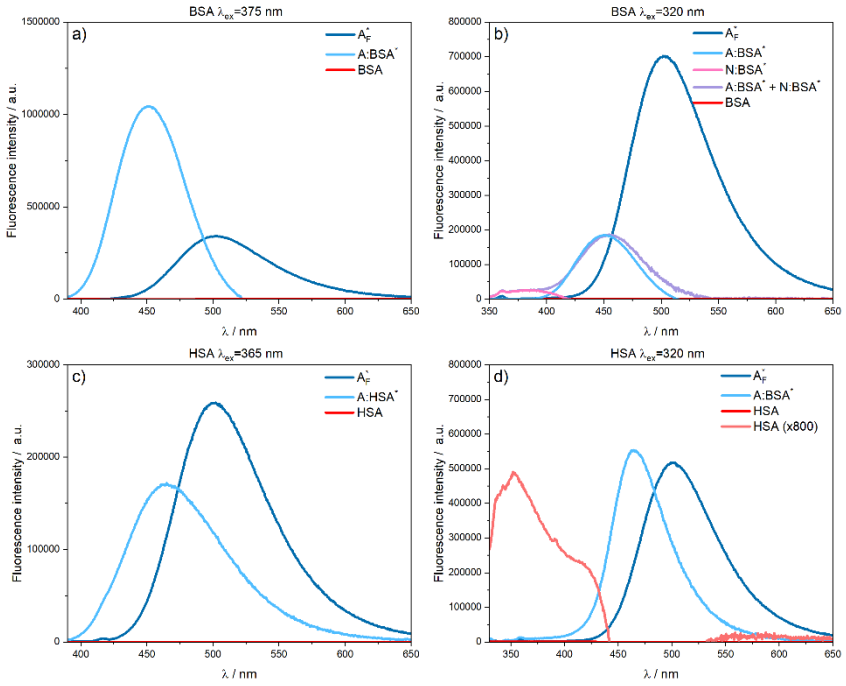



*Figure 5.55. Pure fluorescence emission spectra obtained from the PCGA of titration series obtained. a-b) Results for BSA with excitation wavelengths of 375 nm and 320 nm in that order, c-d) results for HSA with excitation wavelengths of 365 nm and 320 nm, respectively.*

In the case of BSA (Figure 5.55a), when the excitation wavelength is 375 nm, we obtained the two spectra responsible for the emission, corresponding to species **A:BSA*** and **A***. There is a third contribution with very little intensity due to the protein autofluorescence. When exciting at 320 nm (Figure 5.55b), we obtain also two spectra responsible for the emission, one to the free anion **A*** and a second one which is a linear combination of the complex **A:BSA*** and **N:BSA*** emission bands. Due to their identical dependency on

USC DE SANTIAGO DE COMPOSTELA

protein concentration, PCGA analysis cannot separate the emission bands of these complexes. However, we calculated the separated spectra of **A:BSA*** and **N:BSA*** by subtracting the pure spectra of **A:BSA*** obtained exciting at 375 nm from the one exciting at 320 nm. To do this, we first correct the intensity of the pure spectra at $\lambda_{ex}$ = 375 nm to match the intensity at $\lambda_{ex}$ = 320 nm. Then, the new calculated **A:BSA*** was subtracted to combined pure spectra of **A:BSA*** and **N:BSA***, resulting in an estimation of the **N:BSA*** pure spectra (Figure 5.55b). HSA pure spectra present the same behaviour as BSA at 375 nm (Figure 5.55c). However, at 320 nm excitation, the spectrum obtained for the complex corresponds only to the complex **A:BSA***, being the contribution of complex **N:BSA*** attributed completely to the protein. As shown in Figure 5.55d, although the fitted HSA spectrum has very low intensity, the calculated concentration parameter is such that, if applied, it multiplies the intensity almost 1000 times, becoming the main contributor.

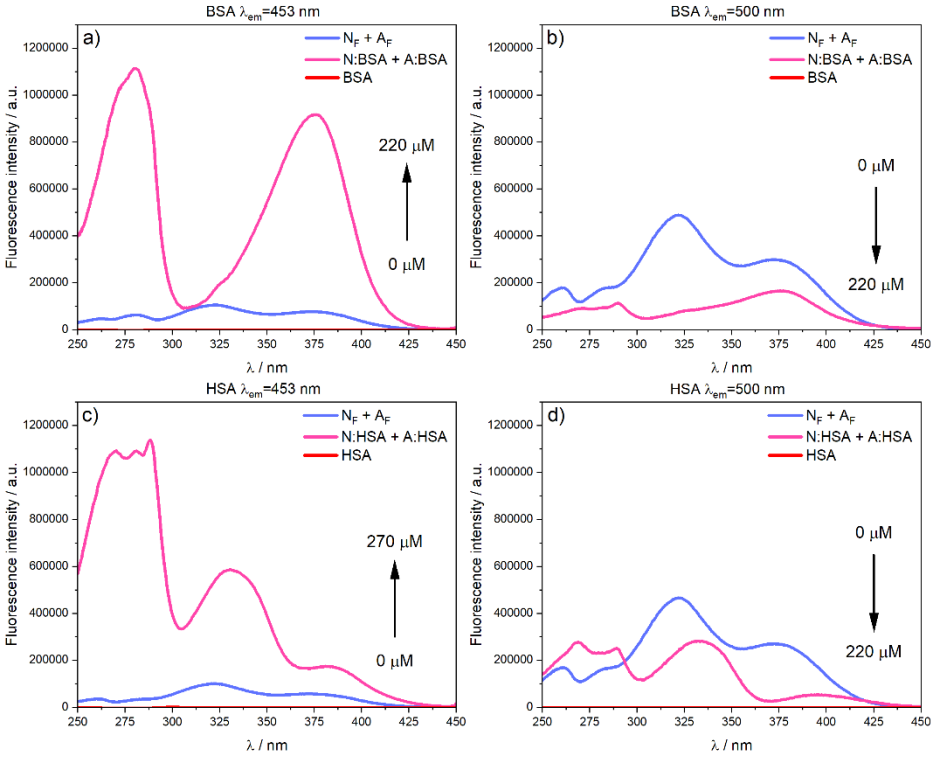



*Figure 5.56. Excitation pure spectra obtained from the PCGA. a) pure spectra of all emissive species formed during the association process with BSA with an emission wavelength of 453 nm, b) pure spectra for CBTOH and BSA binding process with an emission wavelength of 500 nm, c) pure spectra for CBTOH and HSA binding process with an emission wavelength of 453 nm and, d) CBTOH and HSA binding process with an emission wavelength of 500 nm.*

We also carried out the PCGA analysis with the excitation spectra for both proteins using the same procedure as before. Figure 5.56a-b show the obtained pure spectra for BSA at $\lambda_{em}$ = 453 nm and $\lambda_{em}$ = 500 nm, respectively. These pure spectra present a spectral shift towards higher wavelengths for the complex, and only one peak can be seen in the pure spectra of the complex, which corresponds to **A:BSA***. Pure spectra for HSA, shown in Figure 5.56c for $\lambda_{em}$ = 453 nm and Figure 5.56d for $\lambda_{em}$ = 500 nm, present two excitation

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

bands at around 325 nm and 380 nm. The first bands corresponds to **N:HSA*** while the second band can be attributed to **A:HSA***, and appear shifted towards the red. Pure spectra for both BSA and HSA are in agreement with the band observed in the experimental spectra.

### 5.5.2.2. Anisotropy measurements

We also carried out anisotropy measurements that allowed us to see if the environment of the CBTOH upon binding restricted its movement. As shown in Figure 5.57a-b, CBTOH shows an increase in its anisotropy upon adding concentration of BSA. This increment happens at both measured emission wavelengths. This is also seen for the case of HSA in Figure 5.57c-d and it is indicative that the fluorophore molecule's environment is changing from one that allowed the molecule to freely rotate to one that hinders said movement. This change indicates that the binding is indeed taking place.

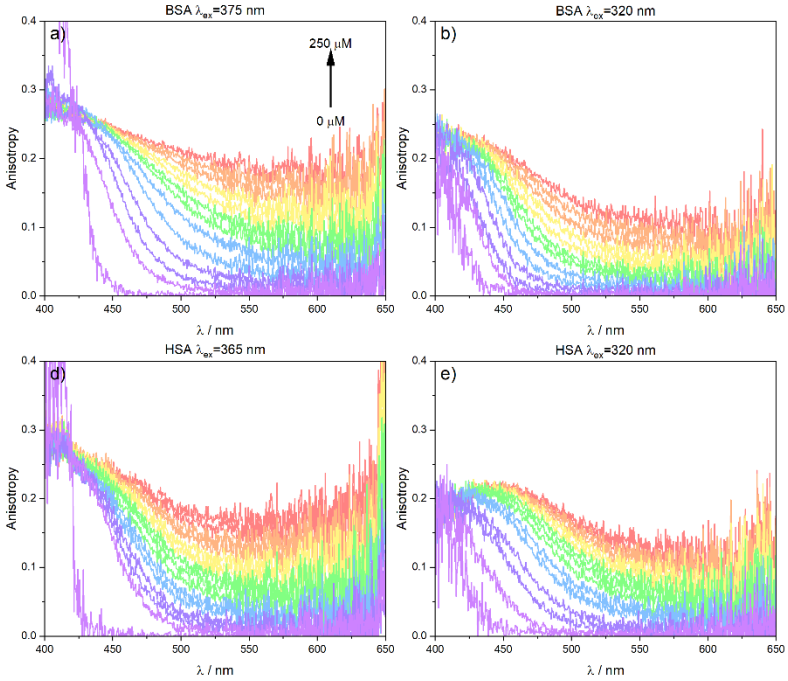



*Figure 5.57. Fluorescence anisotropy spectra of CBTOH:SA complex. a) Spectra obtained with a $\lambda_{ex}$= 375 nm with a BSA concentration range of 0-250 μM, b) spectra obtained with a $\lambda_{ex}$= 320 nm with a BSA concentration range of 0-250 μM, c) spectra obtained with a $\lambda_{ex}$= 375 nm with HSA concentration range of 0-250 μM, and d) spectra obtained with a $\lambda_{ex}$= 365 nm with a HSA concentration range of 0-250 μM.*

If we study the variation of the anisotropy value with protein concentration, we see a clear grow as protein concentration is increased (Figure 5.58). As this variation follows the expected tendency for a 1:1 host-guest association model, the data can be fitted with the same model used in emission and excitation experiments to obtain the apparent binding constant of the system. However, in the case of anisotropy the determined association constant depends on an additional parameter (*q*), which is the emission intensity ratio between the bound CBTOH and the free CBTOH at a given wavelength (Eq (5.10)). This

ratio can be calculated from the experimental spectra of CBTOH, at 0 μM of protein, and CBTOH:SA at 200 μM. For the association with BSA we obtained an apparent constant of (2.8 ± 0.3) × $10^4$ $M^{-1}$ at $\lambda_{ex}$= 320 nm, and of (3.5 ± 0.3) × $10^4$ $M^{-1}$ at $\lambda_{ex}$= 375 nm, which are similar to the obtained values from the emission and excitation analysis. Complexation with HSA resulted in an apparent constant of (2.3 ± 0.2) × $10^4$ $M^{-1}$ at $\lambda_{ex}$= 320 nm, and of (1.4 ± 0.01) × $10^4$ $M^{-1}$ at $\lambda_{ex}$= 365 nm.

$$r^{\lambda} = \frac{r_F^{\lambda} + q^{\lambda} r_B^{\lambda} [P]}{1 + K_{ap} q^{\lambda} [P]} \tag{5.10}$$

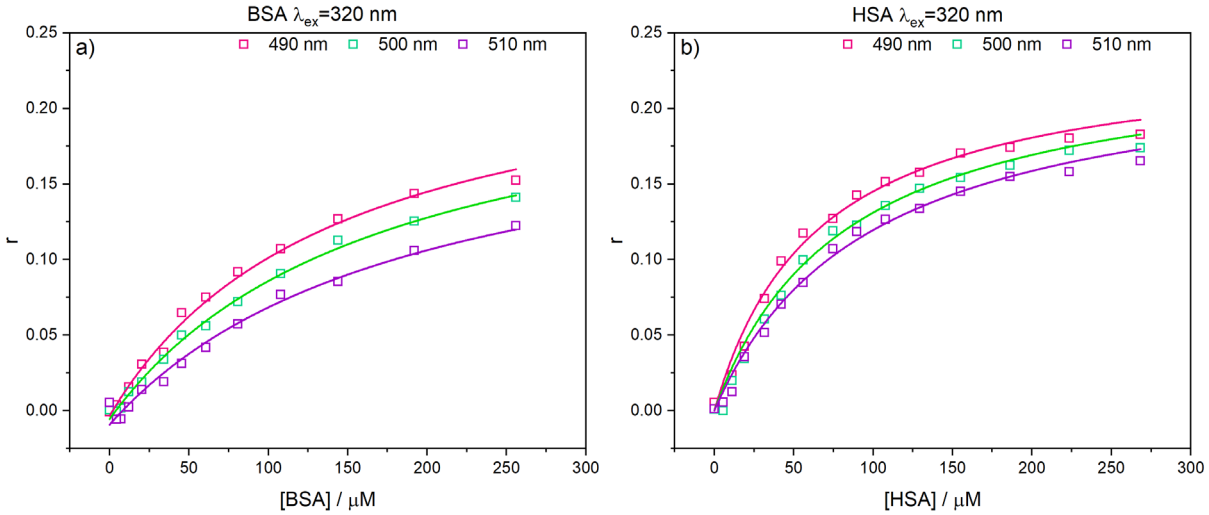



*Figure 5.58. Anistropy (r) variation with protein concentration at different emission wavelengths ($\lambda_{ex}$= 320 nm). a) Variation with BSA concentration, b) variation with HSA concentration.*

### 5.5.2.3. Lifetimes measurements

To get a further insight into the mechanism taking place at the excited state, and in order to characterise better the formed complex between CBTOH and the blood plasma proteins, we also performed time-resolved fluorescence measurements.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

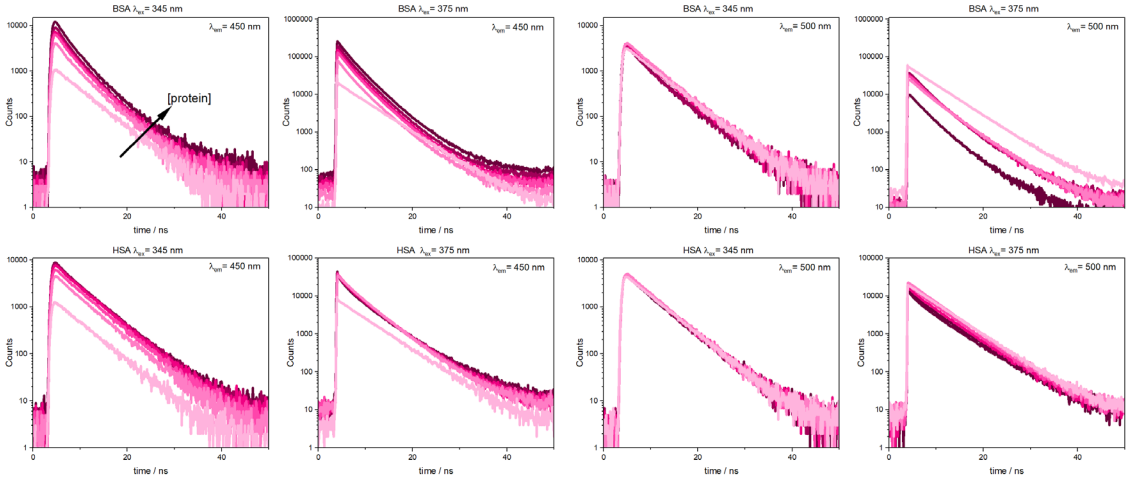



*Figure 5.59. Fluorescence decays of CBTOH variation with serum albumin concentration. Upper panel: decays at different concentrations of BSA and a) $\lambda_{ex}$= 345 nm and $\lambda_{em}$= 450 nm, b) $\lambda_{ex}$= 375 nm and $\lambda_{em}$= 450 nm, c) $\lambda_{ex}$= 345 nm and $\lambda_{em}$= 500 nm, d) $\lambda_{ex}$= 375 nm and $\lambda_{em}$= 500 nm. Lower panel: decays at different concentrations of has and d) $\lambda_{ex}$= 345 nm and $\lambda_{em}$= 450 nm, e) $\lambda_{ex}$= 375 nm and $\lambda_{em}$= 450 nm, f) $\lambda_{ex}$= 345 nm and $\lambda_{em}$= 500 nm, g) $\lambda_{ex}$= 375 nm and $\lambda_{em}$= 500 nm.*

Figure 5.59 shows the experimental fluorescence decay curves variation with protein concentrations at emission wavelengths. Although decays were obtained with two different excitation wavelengths to try to excite both **N** and **A** individually, the available excitation sources at our group only allowed to use 345 nm as excitation wavelength for the **N** instead of 320 nm. Fluorescence decay for BSA experiments present an increase in intensity with protein addition when the emission wavelength is set to 450 nm, which corresponds to the emission band of **A:BSA**$^*$. On the other hand, the intensity lowers with protein concentration when the selected emission wavelength if the one of **A**$^*$ (500 nm). This behaviour can be observed for both studied excitation wavelengths. In the case of HSA, the variation in intensity upon protein addition can also be seen for all the studied conditions with similar tendencies as the ones BSA studies presented.

We performed a PCGA analysis of the series of fluorescence decays at different protein concentrations. Results are summarised in Table 5.5. PCGA resulted in a three-exponential fit for samples with added protein, and a two-exponential fit for samples were only CBTOH was present. This difference is easily explained by the contributions to fluorescence the free and bound species have at said spectral regions. A first very short lifetime was needed to properly fit the data and it was fixed to a value of 0.2 ns. This can be attributed to variations of the laser pulse and other artefacts occurring because of the low intensities the samples present at the studied conditions. Samples with no added protein were studied first to obtain **A*** lifetime, which resulted in a value of 5.30 ± 0.02 ns for $\lambda_{ex}$= 375 nm and a 5.256 ± 0.006 ns for $\lambda_{ex}$= 345 nm. For the association data analysis the second lifetime, corresponding to **A**$^*$ was fixed to 5.3 ns. As shown in Table 5.5, for 200 µM BSA samples, the third lifetime resulted in 2.32 ± 0.02 ns for $\lambda_{ex}$= 375 nm and a 2.45 ± 0.02 ns for $\lambda_{ex}$= 345 nm. However, in the case of HSA samples the obtained lifetime

USC
DE SANTIAGO
DE COMPOSTELA

was of 1.72 ± 0.02 ns for $\lambda_{ex}$= 375 nm and a 1.71 ± 0.02 ns for $\lambda_{ex}$= 345 nm. In both cases, the obtained lifetimes are combinations of the **N:SA*** and the **A:SA*** lifetimes. We were expecting to be able to obtain both values separately when excitation wavelength was set to match, as well as we could, **N** absorption maximum. However, under the studied conditions this could not be achieved for several reasons: both **N** and **A** are being excited at 345 nm, the lifetimes values could be sufficiently similar as to difficult its differentiation and, proton transfer taking place even in the bound species.

*Table 5.5. Fluorescence lifetime values resulting from the PCGA analysis of the fluorescence decays at different protein concentrations for $\lambda_{ex}$ = 375 nm and 345 nm.*

| | [protein] / μM | $\lambda_{ex}$ / nm | $\tau_1$/ ns | $\tau_2$/ ns | $\tau_3$/ ns |
|---|---|---|---|---|---|
| **No protein** | - | 375 | 0.2 | 5.30 ± 0.02 | - |
| | - | 345 | 0.2 | 5.256 ± 0.006 | - |
| **BSA** | 200 | 375 | 0.2 | 5.3 | 2.32 ± 0.02 |
| | 200 | 345 | 0.2 | 5.3 | 2.45 ± 0.02 |
| **HSA** | 200 | 375 | 0.2 | 5.3 | 1.72 ± 0.02 |
| | 200 | 345 | 0.2 | 5.3 | 1.71 ± 0.02 |

If we study the variation of the lifetimes with protein concentration (Figure 5.60) we observe that the third lifetime has no systematic variation with BSA concentration, although it does present some fluctuations. For HSA, on the other hand, there is a clear decreasing tendency of the lifetime value with albumin concentration at both excitation wavelengths. The variation of the lifetime could be caused by different proportions of **N:HSA*** and **A:HSA*** contributing to fluorescence, as the obtained value is a mixture of both species and depends on the degree to which the deprotonation in the excited state has taken place and, therefore, the concentration of each species.

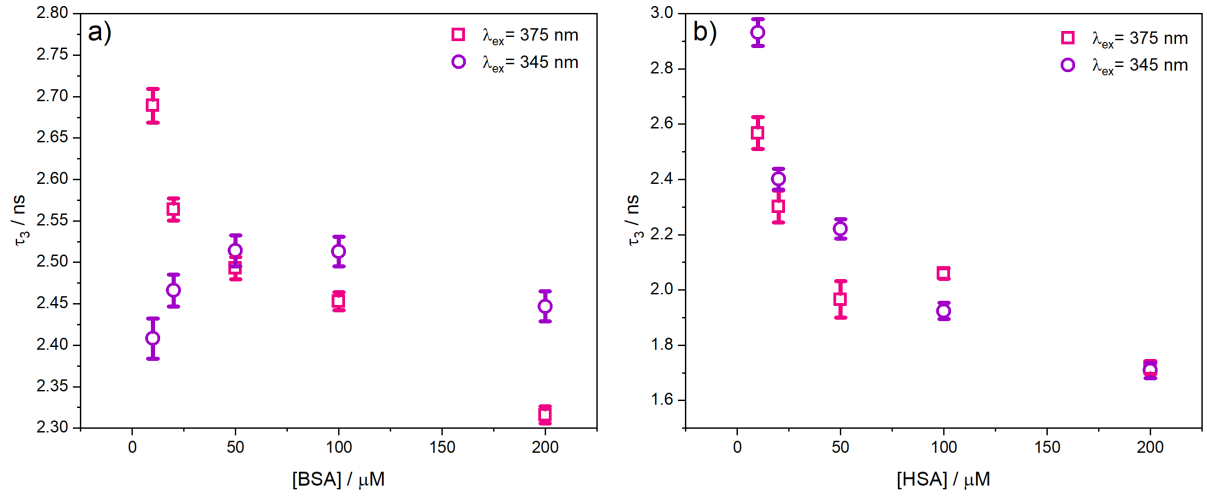



*Figure 5.60. Obtained lifetimes from PCGA variation with protein concentration for $\lambda_{ex}$= 375 nm (pink squares) and $\lambda_{ex}$= 345 nm (purples dots). a) Protein is BSA, b) protein is HSA.*

From PCGA analysis of the decays we also obtained the preexponential factors, or amplitudes of each lifetime. Figure 5.61 shows the amplitudes of the second and third lifetimes (B2 and B3 respectively) dependency with serum albumin concentration. For BSA, the second lifetime contribution decreases for both excitation wavelengths when the emission is set at 500 nm, and it slightly increases when its set at 450 nm (Figure 5.61a-b). The third lifetime, however, increases at all studied emission and excitation wavelengths. For HSA (Figure 5.61c-d) free CBTOH lifetime's amplitude increases with protein concentration when the emission wavelength is 450 nm at both excitation conditions and decreases when emission is set at 500 nm. On the other hand, the complex's lifetime contribution presents an increasing tendency with concentration at all emission and excitation conditions. With both proteins, the free dye lifetime ($\tau_2$) should present a decreasing contribution with concentration as the association process takes places, meaning that less free CBTOH is present in the ground state to be excited and, therefore, less molecules can be deactivated by emission. However, the complex excited-state behaviour of CBTOH together with the different affinity constants of **N** and **A** makes it difficult to predict the variations of the preexponential factors. In any case, even if the lifetime values of the two complexes **N:SA**$^*$ and **A:SA**$^*$ could be determined, it is not possible to extract their individual binding constants from the time-resolved measurements.

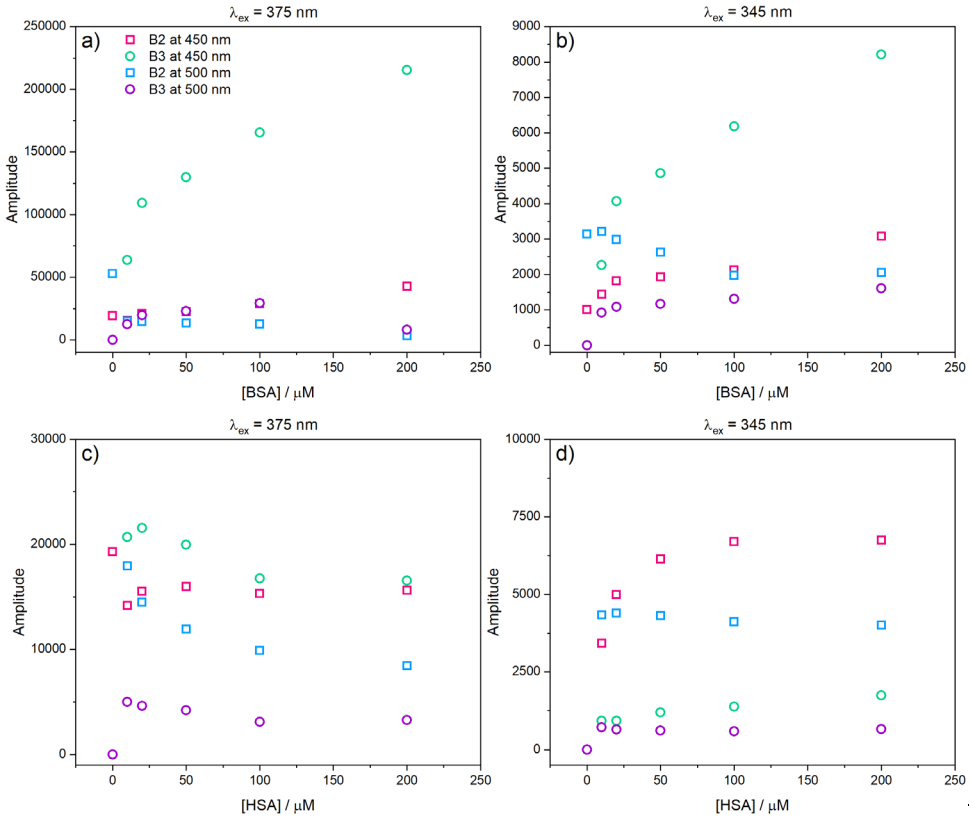



*Figure 5.61. Amplitudes vs protein concentrations. a) BSA for $\lambda_{ex}$= 375 nm, b) BSA for for $\lambda_{ex}$= 345 nm, c) HSA for for $\lambda_{ex}$= 375 nm, d) HSA for for $\lambda_{ex}$= 345 nm.*

### 5.5.2.4. UV-vis absorption measurements

Because CBTOH presents the proton transfer in the excited state, when performing fluorescence base studies only the emission associated with the anionic species of the dye, both unbound and bound, can be seen. Thus, the association constant that we obtained from all the aforementioned studies is actually an apparent constant than includes both the binding constant of **N** ($K_N$) and **A** ($K_A$). In order to try deciphering separately both equilibrium constants we decided to study the association process by UV-vis absorption measurements. To do this, we set conditions that allowed for only one of the species to be present in solution. Consequently, for studies on **N** species carbonate buffer was used, whereas experiments on **A** species were performed at acidic pH having acetate as buffering solution. Figure 5.62 shows the results for the host-guest association of **N** and **A** with both serum albumins.

The absorption spectra for the association between **N** and BSA are shown in Figure 5.62a. Spectra show a shift towards higher wavelengths with protein concentration, moving the maximum from 320 nm to 325 nm. There is also a variation in absorbance, which is lower for the band at lower wavelengths. This means that the absorption spectrum of **N:BSA** has its maximum at 325 nm and that formed complex absorption is more efficient than free **N**. In the case of acetate, absorption spectra of CBTOH in carbonate shift its maximum just by 5 nm, from 320 nm to 325 nm, upon binding to BSA (Figure 5.62b). The absorbance values for **A** are slightly lower than the absorption presented by **A:BSA**. The association process was also studied for HSA and is shown in Figure 5.62c and d for **N** and **A** respectively.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

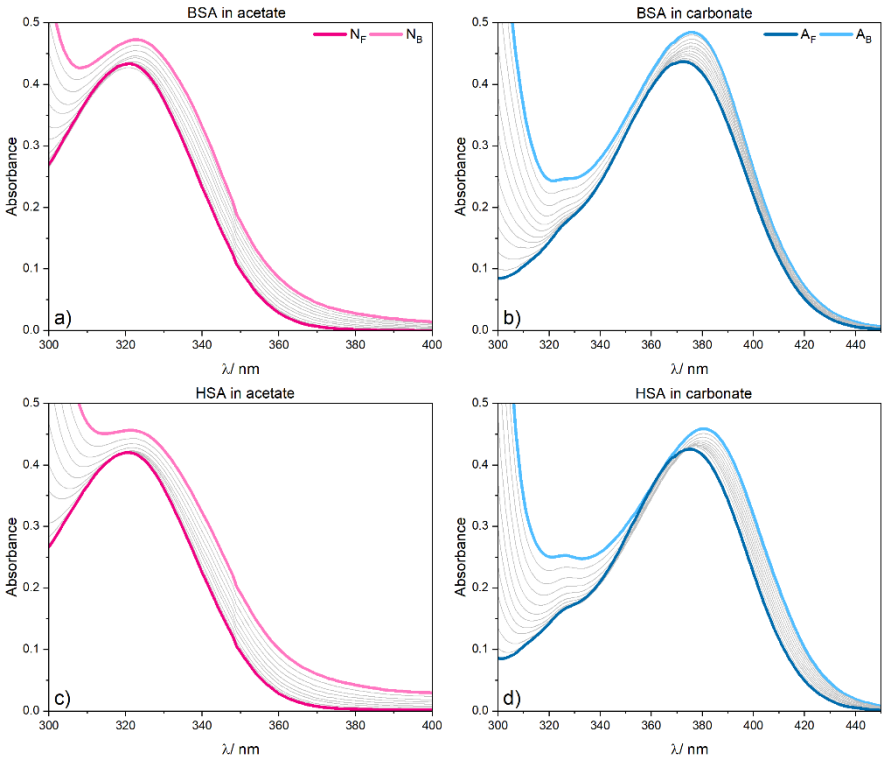



*Figure 5.62. UV-Vis absorption spectra of CBTOH and serum albumin binding studies. a) BSA in acetate buffer, b) BSA in carbonate buffer, c) HSA in acetate buffer, d) HSA in carbonate buffer.*

We performed PCGA analysis without subtraction of the protein contribution to obtain the pure spectra of all the species present in solution (Figure 5.63). To do this we applied the model explain in **Data Analysis** section 4.5.5 (Eq (4.116) and (4.117), and modified to so it could take into consideration an additional parameter C proportional to protein concentration, obtaining Eq (5.11) for the biding process with **N** and Eq (5.12) for the one with **A**. The obtained pure spectra are in agreement with the experimental data both for BSA (Figure 5.63a) and HSA (Figure 5.63b). In addition, the spectral shift towards the red is also evidenced in the pure spectra, where the bands for the complex are at around 325-330 nm (**N:SA**) and 380 nm (**A:SA**).

$$A_N^\lambda = \frac{A_{N_F}^\lambda + A_{N_B}^\lambda K_N P}{1 + K_N P} + C \tag{5.11}$$

$$A_A^\lambda = \frac{A_{A_F}^\lambda + A_{A_B}^\lambda K_A P}{1 + K_A P} + C \tag{5.12}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

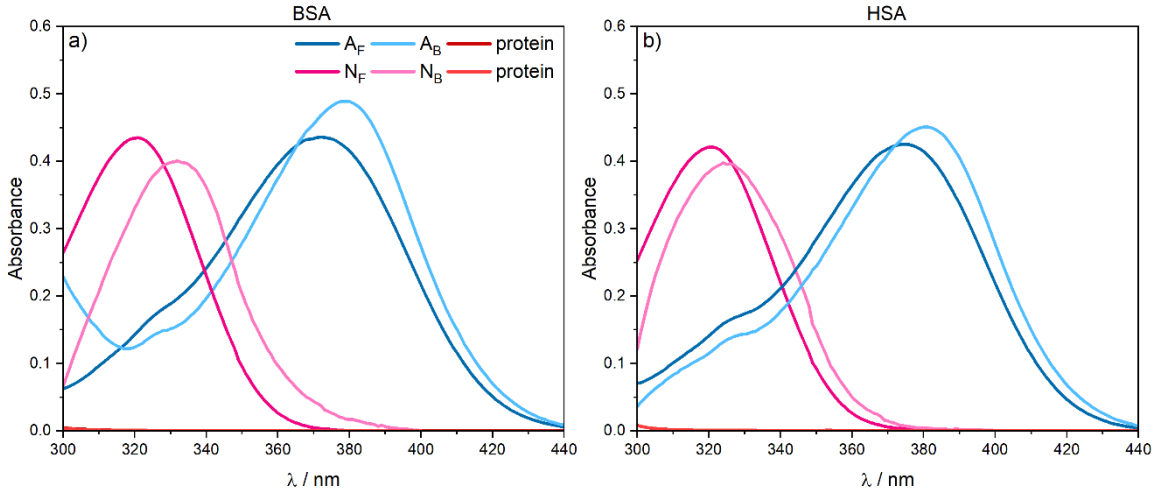



*Figure 5.63. Comparison of the pure spectra of CBTOH free forms (**$N_F$** and **$A_F$**) with pure spectra of the formed complex with serum albumins (**$N_B$** and **$A_B$**). a) Obtained pure spectra from the association studies with BSA, b) pure spectra from the association studies with HSA.*

From the PCGA we also obtained the association constants for **N:SA** and **A:SA** ($K_N$ and $K_A$) for the two serum albumins, which are shown in Table 5.6. If we divide Eq 87 by Eq 86 we obtain the ratio between the acidic constants for the free and bound dye, which only depends on the concentration of each species. This relation can also be obtained by the ratio between the association constants of **N** (Eq 4.88) and **A** (Eq 4.89). Because of this, equating the ratios results in an expression that relates the association constants with the acidity constants of CBTOH ($K_a$) and its complex with serum albumin ($K_a^{'}$). Therefore, from the resulting association constants we can calculate the $pK_a$ of CBTOH in these complexes following Eq (5.13), resulting a value of 7.45 for BSA and 7.52 for HSA. These results indicate that CBTOH is slightly more acidic when complexed with serum albumins than in aqueous physiological medium.

$$K_a^{'} = \frac{K_a K_N}{K_A} \tag{5.13}$$

With $K_N$ and $K_A$ we can also calculate the apparent association constant of the complex, using Eq (5.8). The obtained values (Table 5.6) agree in order of magnitude with those obtained from the emission and excitation studies.

*Table 5.6. Obtained association constants from PCGA for BSA and HSA and the calculated apparent constant from said values.*

| | $K_N$ / $10^4$ $M^{-1}$ | $K_A$ / $10^4$ $M^{-1}$ | $K_{app}$ / $10^4$ $M^{-1}$ |
|---|---|---|---|
| **BSA** | 0.96 ± 0.08 | 1.9 ± 0.2 | 1.24 |
| **HSA** | 2.4 ± 0.31 | 4.0 ± 0.3 | 2.91 |

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

As the obtained apparent constant from the absorption measurements resulted in slightly smaller values than the obtained from the PCGA of the emission series, we carried out the same PCGA again, but for this procedure that apparent constant was fixed to a value of $1.24\times10^4$ $M^{-1}$ for BSA association data and of $2.91\times10^4$ $M^{-1}$ for HSA. These values were obtained from the experimental $K_N$ and $K_A$ equilibrium constants applying Eq (5.8). In both cases, PCGA with the fixed apparent constant values resulted in a satisfactory fit of the data (Figure 5.64).

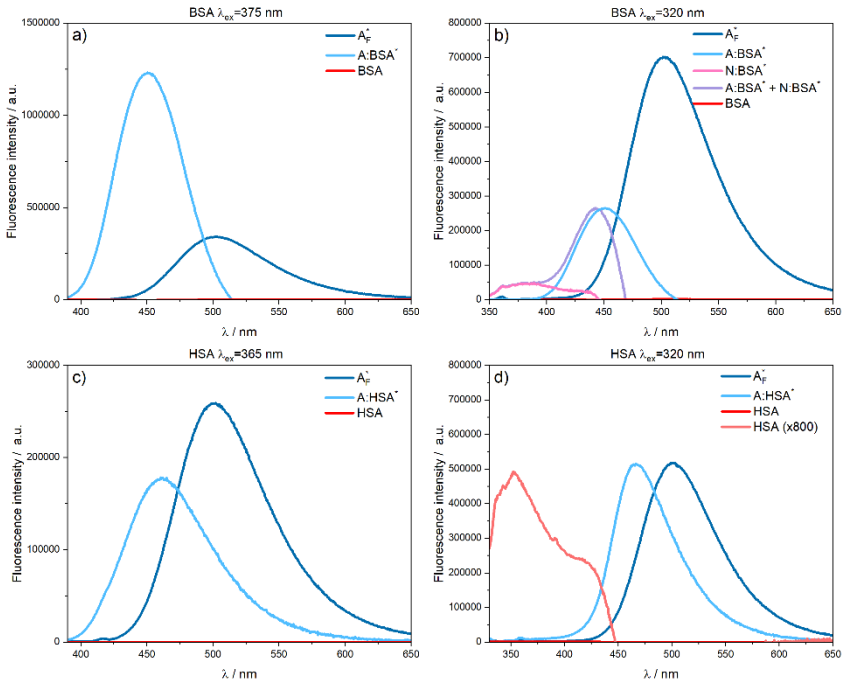



*Figure 5.64. Pure fluorescence emission spectra obtained from the PCGA of titration series. a-b) Results for BSA obtained with excitation wavelengths of 320 nm and 375 nm, c-d) results for HSA obtained with excitation wavelengths of 320 nm and 365 nm.*

As with the pure spectra without fixing the value of the apparent constat, when the excitation wavelength is 375 nm, we observe the two spectra responsible for the emission, which correspond to $A_F^*$ and **A:HSA***, with a third component from the protein emission with almost no intensity . However, at 320 nm the pure spectra of **A:BSA*** and **N:BSA*** could not be separated, so they were calculated the same way as before (Figure 5.64b). The exact same behaviour is seen for HSA (Figure 5.64c and d).

As the association process seems to modify the absorption properties of CBTOH we also studied if the acid-basic properties of the fluorophore were also influenced by the binding to the blood plasma proteins. To do this we carried out some UV-Vis absorption titrations with varying pH. To ensure that no free CBTOH could exist in solution, and based on the results from fluorescence studies, all samples had 200 μM concentration of protein.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

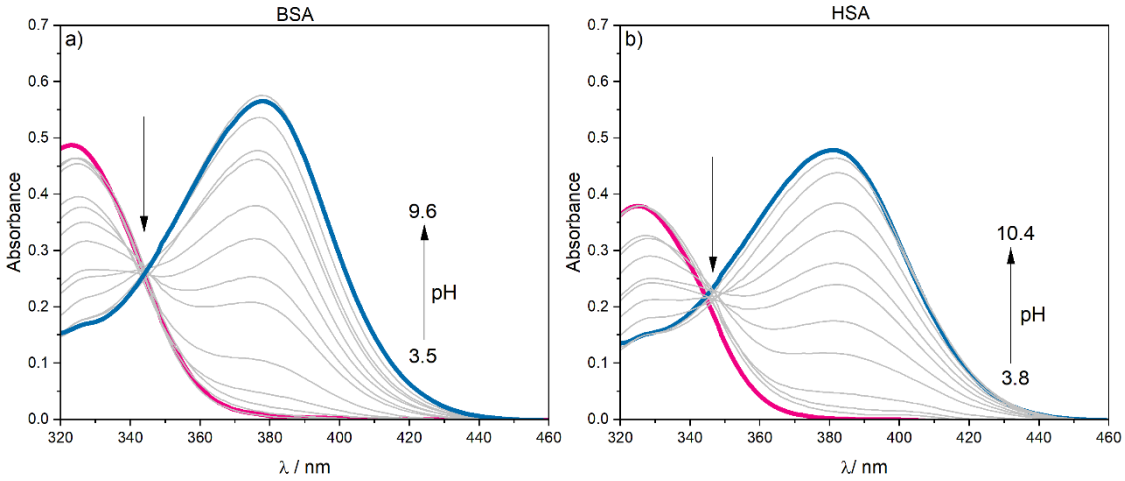



*Figure 5.65. Corrected by subtraction of the protein band UV-Vis absorption spectra at different pH of the complex between CBTOH and BSA (left) and HSA (right).*

As seen in Figure 5.65, the corrected absorption spectra of the complex CBTOH:SA for both proteins seem to follow a similar behaviour as free CBTOH when changing the pH of the media. In both cases two absorption bands can be seen at around 325 nm and 380 nm. As seen for the case of free CBTOH, the band at lower wavelengths decreases in intensity when pH is more basic, whereas the second band presents higher intensity values when increasing the pH of the solution. Thus, as acid-base equilibrium seems to still involve only one transformation, we can use Eq (5.6) as the fitting equation. The band at 325 nm can be attributed to **N:SA** and the second band, at 380 nm corresponds to **A:SA**. The calculated $pK_a$ for the dye:protein complexes is different than the one of the free fluorophore. For the complex with BSA the $pK_a$ resulted in a value of 7.452 ± 0.006, which is in very good agreement with that obtained from the association constants (see above). On the contrary, for the HSA complex we obtained a value of 8.257 ± 0.008, which is higher than the $pK_a$ of the free CBTOH, varying in the opposite direction as determined before.

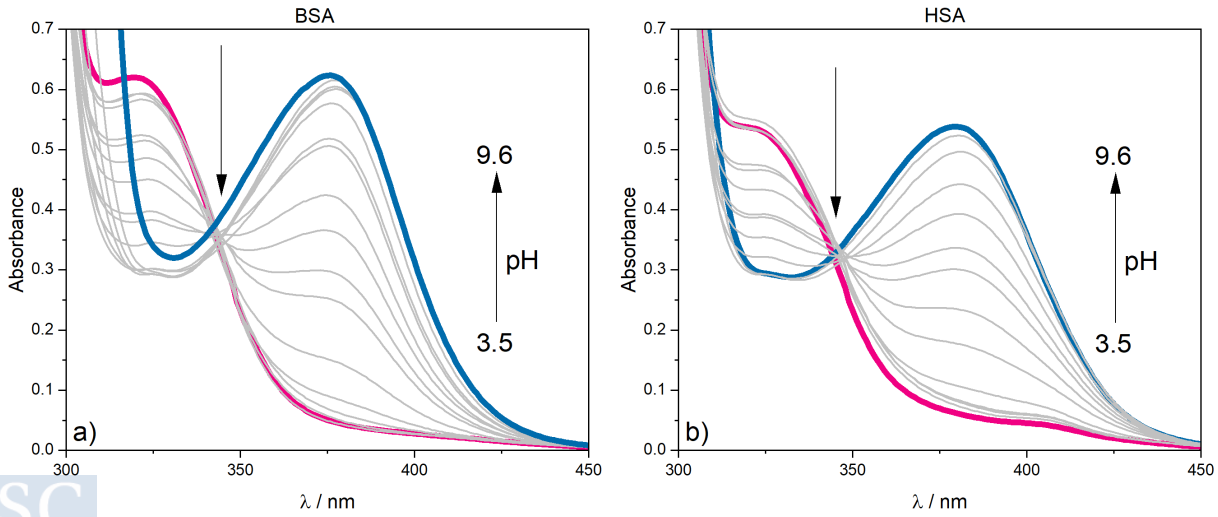



*Figure 5.66. UV-vis absorption spectra before protein band correction at different pH. a) protein is BSA, b) protein is HSA.*

However, it must be noted that the present spectra are a result of a correction of the raw spectra. The correction was needed to try to lower the impact that the protein absorbance has on the dye absorption bands. Figure 5.66 shows the absorption spectra without correction of the protein band, as the used protein concentration was significantly high its absorbance was distorted affecting the spectra even at higher wavelengths were the fluorophore absorption takes place. Due to this distortion as well, the less intense bands that appeared in Figure 5.50 at 260 and 280 nm cannot be seen in the spectra. Moreover, as the protein band is highly distorted due to saturation even after correction of the spectra the information below 310 nm is still lost. To ensure that the protein band subtraction was not distorting the obtained $pK_a$ values, we also carried out a PCGA with the spectra before protein subtraction. Table 5.7 shows the $pK_a$ values for both BSA and HSA analysis before and after protein band correction. The obtained $pK_a$ do not present variations after correction, as the values are equal considering its errors. Thus, this suggests that the protein subtraction does not influence the resulting $pK_a$.

*Table 5.7. Obtained $pK_a$ from the PCGA analysis of corrected and uncorrected absorption spectra in presence of BSA and HSA.*

| | BSA | HSA |
|---|---|---|
| **$pK_a$ corrected spectra** | 7.452 ± 0.006 | 8.092 ± 0.005 |
| **$pK_a$ uncorrected spectra** | 7.44 ± 0.01 | 8.092 ± 0.005 |

## 5.5.3. BSA and HSA behaviour comparison

BSA and HSA are often used interchangeably in a wide range of biological studies due to its similitudes. Commonly, bovine serum albumin is used as model as its obtention, and purification is easier than for HSA.[188] This also translates in different availability of the proteins and in different economic values. Under this premise, we first started these studies using BSA as our initial model before performing HSA studies. However, the presented data throughout this section has evidenced the differences between both serum albumins. Table 5.8 provides a summary of the obtained constant for both the acid-base and association studies with both BSA and HSA as the binding proteins, as well as the calculated $pK_a$ and apparent association constant.

*Table 5.8. Comparison of the experimentally obtained and calculated pKa of complexed CBTOH and association constants of its neutral and the anionic species with BSA and HSA.*

| | $pK_a$ exp | $pK_a$ calculated | $K_N$ / $10^4$ $M^{-1}$ | $K_A$ / $10^4$ $M^{-1}$ | $K_{app}$ / $10^4$ $M^{-1}$ |
|---|---|---|---|---|---|
| **BSA** | 7.452 ± 0.006 | 7.45 | 0.96 ± 0.08 | 1.9 ± 0.2 | 1.2 |
| **HSA** | 8.257 ± 0.008 | 7.52 | 2.4 ± 0.31 | 4.0 ± 0.3 | 2.9 |

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

All the differences in the obtained values between both serum albumins can be explained by their different compositions. Although usually used together, it must be noted that even though their secondary structure is similar, and they are considered almost analogous proteins, their amino acids sequences only present a 76 % of similarity.[188,189] These variations in their sequences give the proteins different properties. For example, as shown in Figure 5.67, the number of aromatic amino acids present in their secondary structures is different. Bovine serum albumin has two tryptophan residues (Trp-134 and Trp-212), while human serum albumin only has one (Trp-214). Moreover, the environment at which the residues are exposed to is slightly different. In the case of BSA, the Trp-134 is in a more hydrophilic environment as is closer to the protein surface, while Trp-212 is located in a closed hydrophobic environment, more protected from the surface. The only tryptophan residue present in HSA is exposed to a similar environment as Trp-212 from BSA.[189,190] As tryptophan is very sensible to the environment and their spectroscopic properties vary depending on the media they are exposed to[93], these differences can explain why BSA and HSA contributions to absorption and fluorescence emission and excitation spectra were different.

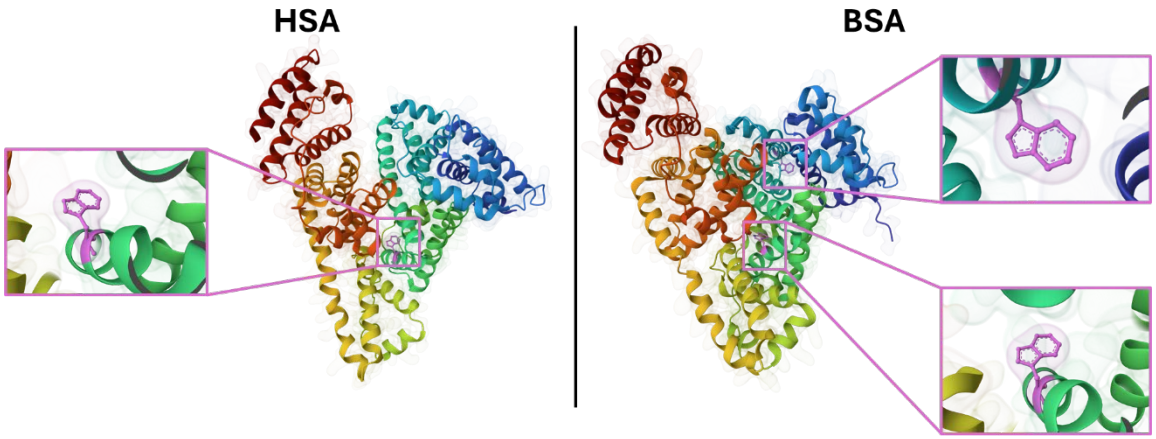



*Figure 5.67. Comparison between human serum albumin (left) and bovine serum albumin (right) 3D secondary structure and the positions tryptophane residues are in each protein.*

Serum albumin sequence variations also translate into structurally different binding sites and affinities. Several studies about their differences can be found in the literature.[191–194] These same variations have also been used to study different approaches for serum albumin detection and differentiation.[189,195–199] On this note, we believe that CBTOH is a promising fluorescence probe for distinction between both albumins as it presents important differences both in absorption and emission spectra (Figure 5.68). With a ratiometric methodology in which absorption and fluorescence emission experiments are combined, the differentiation between both proteins can be achieved. Additionally, as shown in Figure 5.68a, the CBTOH binding to serum albumin allows for determination not

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

only of which protein it is bound to, but also which CBTOH form, **N** or **A**, has bound to the protein. This would allow as well for clarification of the overall acidity of the media.

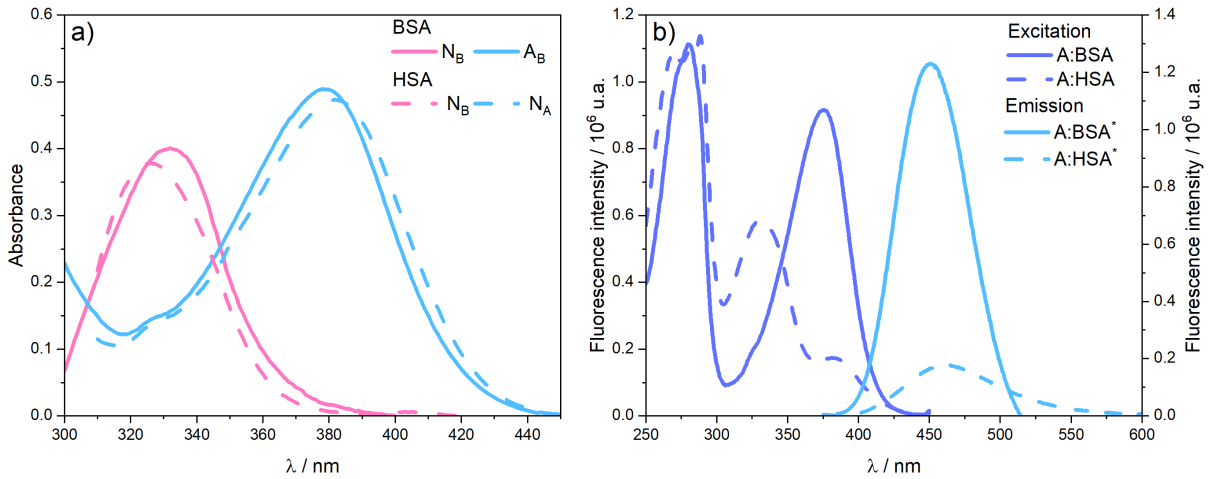



*Figure 5.68. Spectroscopic properties comparison between proteins. a) UV-Vis absorption pure spectra of* ***$N_B$*** *(pink curves) and* ***$A_B$*** *(blue curvas) for BSA (solid curves) and HSA (dashed curves), b) excitation (dark blue) and emission (light blue) pure spectra for BSA (solid curves) and HSA (dashes curves).*

The results presented throughout this work allowed a successful characterisation of the host-guest association process between CBTOH and blood plasma proteins BSA and HSA. We have corroborated the effective binding between the fluorophore using complementary spectroscopic techniques. This allowed us to evidence the potential use of CBTOH as a probe for serum albumin detection, not only being able to bind to BSA and HSA effectively but also allowing for differentiation of the protein it is bound to. Our results also suggest that the use of a radiometric analysis combining absorption and fluorescence emission measurements could potentially discern HSA from BSA. In addition, our results also corroborate CBTOH sensitivity to the pH of the media, which is kept even after binding to the proteins. Even though the acidity constant slightly increases after the association has happened, the presented absorption spectra present spectral variations depending on the CBTOH form that has bound to the protein. These differences could also be used for a qualitative analysis of the acidity of the media. As the overall objective is to find a potential probe for amyloid detection, we believe that this work confirms the potential characteristics of CBTOH as fluorescence probe.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# 6. Conclusions

In this thesis the intra and intermolecular interactions that regulate process of self-assembly, aggregation and host-guest association of different biomolecular systems involved in AD were successfully studied through a spectroscopic point of view. The use and combination of numerous spectroscopic techniques allowed for quantitative and semi-quantitative studies of the dynamics underlying the interactions in these supramolecular biological systems. This way, we studied the mechanisms behind the self-assembly of like G-quadruplexes, the aggregation of amyloid peptides and the host-guest association of blood plasma proteins with fluorescent probes.

In the first part of this thesis (**section 5.1**) we demonstrated that formation of G4 structures with parallel topology occurs in the three G-rich main sequences within mitochondrial CSB II. Our studies also evidence that the number of guanines present in the sequence impacted the stability of the G4, leading to more stable G4 when sequences are richer in guanines. Interestingly, the addition of a guanine located 1 bp upstream also played a key role in the stability of the G-quadruplexes, being more stable when this guanine was part of the sequence. In this work, the effect of cation addition in G4 formation was also studied. The obtained results showed that parallel G4 are formed even in the absence of cations in the media, and that its addition leads to a rearrangement of the G4 topology from a parallel unorder structure to a highly ordered structure. The conformational change happens with the addition of very small amounts of cation, and the constat rate depends on the cation nature. Samples where potassium was added underwent topology rearrangement at lower cation concentrations than those where sodium was the added cation. All three potential G4s forms within CSB II: RNA G4 at the RNA transcript, DNA G4 within the non-transcribed strand and DNA:RNA hybrid between the RNA transcript and the non-transcribed strand were observed both in the CD spectra and the electrophoresis studies. In addition, all three sequence types formed mainly parallel G4 even though the RNA sequences suggested the presence of a slight contribution of different topologies as well. Overall, these findings contribute to a better understanding of the G4 conformations found in mitochondria.

The second part of this thesis, which includes **section 5.2**, **5.3** and **5.4** focused on the study of amyloid peptides aggregation. Within these sections we presented studies on the aggregation process in physiological conditions, how it can be studied with conventional

USC DE COMPOSTELA

spectroscopic techniques for potential diagnosis purposes and how small, charged biomolecules could modify the aggregation process.

In **section 5.2** we investigated the aggregation behaviour of Aβ40 using fluorescence correlation spectroscopy (FCS). Our results revealed a two-step cooperative aggregation process for Aβ40, contrary to the reported single-step aggregation for Aβ42. We also determined the critical aggregation concentrations (*cac*) of both steps. The first critical aggregation concentration ($cac_1$) for Aβ40 is 0.5 ± 0.3 μM, about five times higher than that of Aβ42, while the second critical aggregation concentration ($cac_2$) is 18.9 ± 2.2 μM. No significant aggregation was observed before surpassing $cac_1$, but at concentrations above it, small metastable oligomers of about 14 monomers and stable oligomers of 62 monomers were formed, not exceeding 10% of the total Aβ40 concentration. These stable aggregates are comparable to those reported for Aβ42, but the higher *cac* and the much lower fraction of aggregated amyloid is evidence of Aβ40's much lower aggregation tendency. Exciding the $cac_2$ lead to the formation of much larger aggregates, possibly protofibrils, with around 50% aggregated amyloid. This demonstrates the potential for further growth of initial oligomers at higher concentrations. We also performed reversibility studies that indicate that early Aβ40 oligomers are reversible upon dilution, but for the complete reversibility to occur longer times are necessary. The stability studies suggest that Aβ40 oligomers remain stable over the studied time, with no significant changes in aggregate size even after one week, although an increase in polydispersity is noted. Our results demonstrate the differences between the aggregation mechanisms of Aβ40 and Aβ42 and highlight the importance of concentration and incubation time in studying amyloid aggregation.

The potential use of amyloid's autofluorescence for aggregation monitoring is investigated throughout **section 5.3**, where we evidence that autofluorescence is a promising tool for amyloid aggregation detection. Fluorescence emission spectra of aggregated and monomeric Aβ40 resulted in both spectral and intensity differences. The variations are significant enough to easily know whether an amyloid sample has undergone or not aggregation. Our results indicate that, when aggregation has not taken place, the tyrosine residue undergoes proton transfer on the excited state, forming deprotonated tyrosine, which presents and emission band at 350 nm instead of 303 nm. Aβ40's autofluorescence also presents a 500 nm band that may be attributed to an aggregation induced emission. Although the need for further studies with better quality data, we were able to semi-quantitative the photophysical mechanism that the tyrosine residue undergoes when amyloid is in its monomeric and aggregated state.

The potential impact of electrostatic interactions in amyloid aggregation was presented in **section 5.4**. This was done by addition of small, charged biomolecules like polyamines and metal ions (spermine, spermidine and $Mg^{2+}$) to samples in different aggregation

stages. Fluorescence-imaging followed by the application of SRRF algorithm allowed us to obtain super-resolution images. We first corroborated that amyloid aggregation takes place through liquid-liquid phase separation (LLPS), where droplets with local high amyloid concentrations are formed promoting aggregation inside them. Studies where cation addition was done after fibrils were already formed, evidenced the effect of the highest charged cation, spermine, where LLPS was promoted upon its addition. However, when the polyamine with smaller charge spermidine was added, no changes were seen. This indicates that higher charges could be needed for charge-charge interactions to be significant. Experiments with magnesium as the cation also presented droplet formation at the higher concentrations and incubations times, hinting towards charge not being the only factor controlling the influence of charged biomolecules. The droplet sizes were similar for both spermine and magnesium experiments and did not present variations with time incubation or concentration, meaning that the reversibility is not complete and LLPS only happens up to the last stage before fibril formation. Although it is an undergoing study that needs further data for stronger conclusions to be drawn, our results hint towards charge-charge interactions with cations influencing amyloid aggregation.

Finally, in the last section of the **Results and discussion** chapter of this thesis (**section 5.5**) we have presented the results of the study of host-guest association of blood plasma proteins with CBTOH. With conventional fluorescence techniques, we have demonstrated the effective binding between the fluorophore and both proteins. Our results show that the binding process can be monitored with absorption and emission fluorescence, as both CBTOH absorption and emission undergo intensity and spectral variations. This is evidence of the potential use of CBTOH as a probe for serum albumin detection. We have successfully suggested a binding mechanism following a host-guest 1:1 association equilibrium for both proteins. As the binding process with BSA and HSA yield differences between the two proteins, the use of CBTOH could help to discern between BSA and HSA effectively, which suggest that the dye binding to other proteins like amyloids could also be selective enough to present significant differences. The combination of absorption and emission measurements, using a radiometric analysis, could potentially discern HSA from BSA.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# References


1. Shah, D. Biomolecules: The Elements That Make Up Life. *Int. Res. J. Basic Clin. Stud.* **8**, 1–3 (2023).
2. Subramani, K. & Ahmed, W. Chapter 13 - Self-Assembly of Proteins and Peptides and Their Applications in Bionanotechnology and Dentistry. in *Emerging Nanotechnologies in Dentistry* (eds. Subramani, K. & Ahmed, W.) 209–224 (William Andrew Publishing, Boston, 2012).
3. Fung, S. Y., Hong, Y., Keyes-baig, C. & Chen, P. 12 - Self-assembly of peptides and its potential applications. in *Molecular Interfacial Phenomena of Polymers and Biopolymers* (ed. Chen, P.) 421–474 (Woodhead Publishing, 2005).
4. Lotz, G. P. & Legleiter, J. The role of amyloidogenic protein oligomerization in neurodegenerative disease. *J. Mol. Med.* **91**, 653–664 (2013).
5. Gandhi, J. *et al.* Protein misfolding and aggregation in neurodegenerative diseases: a review of pathogeneses, novel detection strategies, and potential therapeutics. *Rev. Neurosci.* **30**, 339–358 (2019).
6. Kurtishi, A., Rosen, B., Patil, K. S., Alves, G. W. & Møller, S. G. Cellular Proteostasis in Neurodegeneration. *Mol. Neurobiol.* **56**, 3676–3689 (2019).
7. Aranda-Anzaldo, A., Dent, M. A. R., Segura-Anaya, E. & Martínez-Gómez, A. Protein folding, cellular stress and cancer. *Prog. Biophys. Mol. Biol.* **191**, 40–57 (2024).
8. Jiang, S., Fagman, J. B., Chen, C., Alberti, S. & Liu, B. Protein phase separation and its role in tumorigenesis. *eLife* **9**, e60264 (2020).


USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

9. Zhong, Q. *et al.* Protein posttranslational modifications in health and diseases: Functions, regulatory mechanisms, and therapeutic implications. *MedComm* **4**, e261 (2023).
10. Boada, M. *et al.* A randomized, controlled clinical trial of plasma exchange with albumin replacement for Alzheimer's disease: Primary results of the AMBAR Study. *Alzheimers Dement. J. Alzheimers Assoc.* **16**, 1412–1425 (2020).
11. Global Dementia Observatory (GDO). https://www.who.int/data/gho/data/themes/global-dementia-observatory-gdo.
12. *Dementia in Europe Yearbook 2019. Estimating the Prevalence of Dementia in Europe.* (Alzheimer Europe, Luxembourg, 2019).
13. Calabrò, M., Rinaldi, C., Santoro, G. & Crisafulli, C. The biological pathways of Alzheimer disease: a review. *AIMS Neurosci.* **8**, 86–132 (2020).
14. Busche, M. A. & Hyman, B. T. Synergy between amyloid-β and tau in Alzheimer's disease. *Nat. Neurosci.* **23**, 1183–1193 (2020).
15. Muralidar, S., Ambi, S. V., Sekaran, S., Thirumalai, D. & Palaniappan, B. Role of tau protein in Alzheimer's disease: The prime pathological player. *Int. J. Biol. Macromol.* **163**, 1599–1617 (2020).
16. Wegmann, S., Biernat, J. & Mandelkow, E. A current view on Tau protein phosphorylation in Alzheimer's disease. *Curr. Opin. Neurobiol.* **69**, 131–138 (2021).
17. Yin, X. *et al.* The Role of Amyloid-Beta and Tau in the Early Pathogenesis of Alzheimer's Disease. *Med. Sci. Monit. Int. Med. J. Exp. Clin. Res.* **27**, e933084-1-e933084-7 (2021).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

18. Roda, A. R., Serra-Mir, G., Montoliu-Gaya, L., Tiessler, L. & Villegas, S. Amyloid-beta peptide and tau protein crosstalk in Alzheimer's disease. *Neural Regen. Res.* **17**, 1666–1674 (2022).

19. 2023 Alzheimer's disease facts and figures. *Alzheimers Dement.* **19**, 1598–1695 (2023).

20. Wang, E., Thombre, R., Shah, Y., Latanich, R. & Wang, J. G-Quadruplexes as pathogenic drivers in neurodegenerative disorders. *Nucleic Acids Res.* **49**, 4816–4830 (2021).

21. Simone, R., Fratta, P., Neidle, S., Parkinson, G. N. & Isaacs, A. M. G-quadruplexes: Emerging roles in neurodegenerative diseases and the non-coding transcriptome. *FEBS Lett.* **589**, 1653–1668 (2015).

22. Chen, Y., Strickland, M. R., Soranno, A. & Holtzman, D. M. Apolipoprotein E: Structural Insights and Links to Alzheimer Disease Pathogenesis. *Neuron* **109**, 205–221 (2021).

23. Andrade-Guerrero, J. *et al.* Alzheimer's Disease: An Updated Overview of Its Genetics. *Int. J. Mol. Sci.* **24**, 3754 (2023).

24. Boada, M., Martínez-Lage, P., Serrano-Castro, P., Costa, M. & Páez, A. Therapeutic plasma exchange with albumin: a new approach to treat Alzheimer's disease. *Expert Rev. Neurother.* **21**, 843–849 (2021).

25. Rohrer, L., Yunce, M., Montine, T. J. & Shan, H. Plasma Exchange in Alzheimer's Disease. *Transfus. Med. Rev.* **37**, 10–15 (2023).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

26. Picón-Pagès, P. *et al*. Human Albumin Impairs Amyloid β-peptide Fibrillation Through its C-terminus: From docking Modeling to Protection Against Neurotoxicity in Alzheimer's disease. *Comput. Struct. Biotechnol. J.* **17**, 963–971 (2019).

27. Illodo, S. *et al*. Spectroscopic Characterization of Mitochondrial G-Quadruplexes. *Int. J. Mol. Sci.* **23**, 925 (2022).

28. Gellert, M., Lipsett, M. N. & Davies, D. R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci.* **48**, 2013–2018 (1962).

29. Maizels, N. & Gray, L. T. The G4 Genome. *PLoS Genet.* **9**, e1003468 (2013).

30. Lam, E. Y. N., Beraldi, D., Tannahill, D. & Balasubramanian, S. G-quadruplex structures are stable and detectable in human genomic DNA. *Nat. Commun.* **4**, 1796 (2013).

31. Paeschke, K., Simonsson, T., Postberg, J., Rhodes, D. & Lipps, H. J. Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nat. Struct. Mol. Biol.* **12**, 847–854 (2005).

32. Schaffitzel, C. *et al*. In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8572–8577 (2001).

33. Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5**, 182–186 (2013).

34. Laguerre, A., Wong, J. M. Y. & Monchaud, D. Direct visualization of both DNA and RNA quadruplexes in human cells via an uncommon spectroscopic method. *Sci. Rep.* **6**, 1–10 (2016).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

35. Yan, J. W. *et al.* Development of a new colorimetric and red-emitting fluorescent dual probe for G-quadruplex nucleic acids. *Chem. Commun.* **50**, 6927–6930 (2014).

36. Chen, X.-C. *et al.* Tracking the Dynamic Folding and Unfolding of RNA G-Quadruplexes in Live Cells. *Angew. Chem. Int. Ed Engl.* **57**, 4702–4706 (2018).

37. Laguerre, A. *et al.* Visualization of RNA-Quadruplexes in Live Cells. *J. Am. Chem. Soc.* **137**, 8521–8525 (2015).

38. Xu, S. *et al.* Directly lighting up RNA G-quadruplexes from test tubes to living human cells. *Nucleic Acids Res.* **43**, 9575–9586 (2015).

39. Falabella, M. *et al.* G-quadruplex dynamics contribute to regulation of mitochondrial gene expression. *Sci. Rep.* **9**, 5605 (2019).

40. Huang, W. C. *et al.* Direct evidence of mitochondrial G-quadruplex DNA by using fluorescent anti-cancer agents. *Nucleic Acids Res.* **43**, 10102–10113 (2015).

41. Wu, F. *et al.* Genome-wide analysis of DNA G-quadruplex motifs across 37 species provides insights into G4 evolution. *Commun. Biol.* **4**, 1–11 (2021).

42. Karsisiotis, A. I. *et al.* Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angew. Chem. Int. Ed.* **50**, 10645–10648 (2011).

43. Bochman, M. L., Paeschke, K. & Zakian, V. A. DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.* **13**, 770–780 (2012).

44. Wood, B. R. The importance of hydration and DNA conformation in interpreting infrared spectra of cells and tissues. *Chem. Soc. Rev.* **45**, 1980–1998 (2016).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

45. Murat, P. & Balasubramanian, S. Existence and consequences of G-quadruplex structures in DNA. *Curr. Opin. Genet. Dev.* **25**, 22–29 (2014).

46. Harkness, R. W. & Mittermaier, A. K. G-quadruplex dynamics. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1865**, 1544–1554 (2017).

47. Hillen, H. S. *et al.* Mechanism of Transcription Anti-termination in Human Mitochondria. *Cell* **171**, 1082-1093.e13 (2017).

48. Fisette, J.-F., Montagna, D. R., Mihailescu, M.-R. & Wolfe, M. S. A G-rich element forms a G-quadruplex and regulates BACE1 mRNA alternative splicing. *J. Neurochem.* **121**, 763–773 (2012).

49. Crenshaw, E. *et al.* Amyloid Precursor Protein Translation Is Regulated by a 3'UTR Guanine Quadruplex. *PloS One* **10**, e0143160 (2015).

50. Hanna, R., Flamier, A., Barabino, A. & Bernier, G. G-quadruplexes originating from evolutionary conserved L1 elements interfere with neuronal gene expression in Alzheimer's disease. *Nat. Commun.* **12**, 1828 (2021).

51. Vijay Kumar, M. J., Morales, R. & Tsvetkov, A. S. G-quadruplexes and associated proteins in aging and Alzheimer's disease. *Front. Aging* **4**, (2023).

52. Lammich, S. *et al.* Translational Repression of the Disintegrin and Metalloprotease ADAM10 by a Stable G-quadruplex Secondary Structure in Its 5'-Untranslated Region*. *J. Biol. Chem.* **286**, 45063–45072 (2011).

53. Dai, J. *et al.* Discovery of Small Molecules for Up-Regulating the Translation of Antiamyloidogenic Secretase, a Disintegrin and Metalloproteinase 10 (ADAM10), by Binding to the G-Quadruplex-Forming Sequence in the 5' Untranslated Region (UTR) of Its mRNA. *J. Med. Chem.* **58**, 3875–3891 (2015).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

54. Hampel, H. *et al*. The Amyloid-β Pathway in Alzheimer's Disease. *Mol. Psychiatry* **26**, 5481–5503 (2021).

55. Verma, A., Kumar Waiker, D., Bhardwaj, B., Saraf, P. & Shrivastava, S. K. The molecular mechanism, targets, and novel molecules in the treatment of Alzheimer's disease. *Bioorganic Chem.* **119**, 105562 (2022).

56. Lee, S. J. C., Nam, E., Lee, H. J., Savelieff, M. G. & Lim, M. H. Towards an understanding of amyloid-β oligomers: characterization, toxicity mechanisms, and inhibitors. *Chem. Soc. Rev.* **46**, 310–323 (2017).

57. Murphy, M. P. & LeVine, H. Alzheimer's Disease and the β-Amyloid Peptide. *J. Alzheimers Dis. JAD* **19**, 311 (2010).

58. Meisl, G. *et al*. Differences in nucleation behavior underlie the contrasting aggregation kinetics of the Aβ40 and Aβ42 peptides. *Proc. Natl. Acad. Sci.* **111**, 9384–9389 (2014).

59. Chiti, F. & Dobson, C. M. Protein misfolding, amyloid formation, and human disease: A summary of progress over the last decade. *Annu. Rev. Biochem.* **86**, 27–68 (2017).

60. Zheng, W., Tsai, M.-Y. & Wolynes, P. G. Comparing the Aggregation Free Energy Landscapes of Amyloid Beta(1–42) and Amyloid Beta(1–40). *J. Am. Chem. Soc.* **139**, 16666–16676 (2017).

61. Heo, C. E., Choi, T. S. & Kim, H. I. Competitive homo- and hetero- self-assembly of amyloid-β 1–42 and 1–40 in the early stage of fibrillation. *Int. J. Mass Spectrom. Ion Process.* **428**, 15–21 (2018).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

62. Festa, G. *et al.* Aggregation States of Aβ1–40, Aβ1–42 and Aβp3–42 Amyloid Beta Peptides: A SANS Study. *Int. J. Mol. Sci.* **20**, 4126 (2019).

63. Nirmalraj, P. N. *et al.* Complete aggregation pathway of amyloid β (1-40) and (1-42) resolved on an atomically clean interface. *Sci. Adv.* **6**, eaaz6014 (2020).

64. Wang, L., Eom, K. & Kwon, T. Different Aggregation Pathways and Structures for Aβ40 and Aβ42 Peptides. *Biomolecules* **11**, 198 (2021).

65. Deleanu, M. *et al.* Unraveling the Speciation of β-Amyloid Peptides during the Aggregation Process by Taylor Dispersion Analysis. *Anal. Chem.* **93**, 6523–6533 (2021).

66. Deleanu, M. *et al.* Taylor Dispersion Analysis and Atomic Force Microscopy Provide a Quantitative Insight into the Aggregation Kinetics of Aβ (1–40)/Aβ (1–42) Amyloid Peptide Mixtures. *ACS Chem. Neurosci.* **13**, 786–795 (2022).

67. Braun, G. A., Dear, A. J., Sanagavarapu, K., Zetterberg, H. & Linse, S. Amyloid-β peptide 37, 38 and 40 individually and cooperatively inhibit amyloid-β 42 aggregation. *Chem. Sci.* **13**, 2423–2439 (2022).

68. Novo, M., Freire, S. & Al-Soufi, W. Critical aggregation concentration for the formation of early Amyloid-β (1-42) oligomers. *Sci. Rep.* **8**, 1783 (2018).

69. Yong, W. *et al.* Structure determination of micelle-like intermediates in amyloid beta -protein fibril assembly by using small angle neutron scattering. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 150–154.

70. Lin, Y. *et al.* Diverse Structural Conversion and Aggregation Pathways of Alzheimer's Amyloid-β (1–40). *ACS Nano* **13**, 8766–8783 (2019).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

71. Johnson, R. D., Schauerte, J. A., Wisser, K. C., Gafni, A. & Steel, D. G. Direct observation of single amyloid-β(1-40) oligomers on live cells: binding and growth at physiological concentrations. *PloS One* **6**, e23970 (2011).

72. Nag, S. *et al.* Nature of the Amyloid-β Monomer and the Monomer-Oligomer Equilibrium. *J. Biol. Chem.* **286**, 13827–13833 (2011).

73. Morel, B., Carrasco, M. P., Jurado, S., Marco, C. & Conejero-Lara, F. Dynamic micellar oligomers of amyloid beta peptides play a crucial role in their aggregation mechanisms. *Phys. Chem. Chem. Phys.* **20**, 20597–20614 (2018).

74. Sharma, L. G. & Pandey, L. M. Shear-induced aggregation of amyloid β (1–40) in a parallel plate geometry. *J. Biomol. Struct. Dyn.* **39**, 6415–6423 (2021).

75. Das, A., Gupta, A., Hong, Y., Carver, J. A. & Maiti, S. A Spectroscopic Marker for Structural Transitions Associated with Amyloid-β Aggregation. *Biochemistry* **59**, 1813–1822 (2020).

76. Kumar, S. & Walter, J. Phosphorylation of amyloid beta (Aβ) peptides – A trigger for formation of toxic aggregates in Alzheimer's disease. *Aging* **3**, 803–12 (2011).

77. Ray, S. *et al.* α-Synuclein aggregation nucleates through liquid-liquid phase separation. *Nat. Chem.* **12**, 705–716 (2020).

78. Hazra, M. K. & Levy, Y. Biophysics of Phase Separation of Disordered Proteins Is Governed by Balance between Short- And Long-Range Interactions. *J. Phys. Chem. B* **125**, 2202–2211 (2021).

79. Xing, Y. *et al.* Amyloid Aggregation under the Lens of Liquid–Liquid Phase Separation. *J. Phys. Chem. Lett.* **12**, 368–378 (2021).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

80. Dignon, G. L., Best, R. B. & Mittal, J. Biomolecular Phase Separation: From Molecular Driving Forces to Macroscopic Properties. *Annu. Rev. Phys. Chem.* **71**, 53–75 (2020).

81. Zbinden, A., Pérez-Berlanga, M., De Rossi, P. & Polymenidou, M. Phase Separation and Neurodegenerative Diseases: A Disturbance in the Force. *Dev. Cell* **55**, 45–68 (2020).

82. Kang, M. *et al.* Aggregation or phase separation can be induced in highly charged proteins by small charged biomolecules. *Soft Matter* **18**, 3313–3317 (2022).

83. Pak, C. W. *et al.* Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Mol. Cell* **63**, 72–85 (2016).

84. Wang, S. *et al.* Human Serum Albumin (HSA) and Its Applications as a Drug Delivery Vehicle. **14**, 1–8 (2020).

85. Fasano, M. *et al.* The extraordinary ligand binding properties of human serum albumin. *IUBMB Life* **57**, 787–796 (2005).

86. Biere, A. L. *et al.* Amyloid beta-peptide is transported on lipoproteins and albumin in human plasma. *J. Biol. Chem.* **271**, 32916–32922 (1996).

87. Costa, M., Ortiz, A. M. & Jorquera, J. I. Therapeutic albumin binding to remove amyloid-β. *J. Alzheimers Dis. JAD* **29**, 159–170 (2012).

88. Milojevic, J., Costa, M., Ortiz, A. M., Jorquera, J. I. & Melacini, G. In Vitro Amyloid-β Binding and Inhibition of Amyloid-β Self-Association by Therapeutic Albumin. *J. Alzheimers Dis.* **38**, 753–765 (2014).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

89. Feng, C. *et al*. Modulation of inflammatory regulation by plasma exchange with albumin replacement revealed by deep plasma and CSF proteomic profiling of the AMBAR study. *Alzheimers Dement*. **19**, e079113 (2023).

90. Yamamoto, K., Shimada, H., Koh, H., Ataka, S. & Miki, T. Serum levels of albumin-amyloid beta complexes are decreased in Alzheimer's disease. *Geriatr. Gerontol. Int*. **14**, 716–723 (2014).

91. Milojevic, J., Raditsis, A. & Melacini, G. Human Serum Albumin Inhibits Aβ Fibrillization through a "Monomer-Competitor" Mechanism. *Biophys. J.* **97**, 2585–2594 (2009).

92. Guo, C. & Zhou, H. X. Fatty Acids Compete with Aβ in Binding to Serum Albumin by Quenching Its Conformational Flexibility. *Biophys. J.* **116**, 248–257 (2019).

93. Lakowicz, J. R. & Masters, B. R. *Principles of Fluorescence Spectroscopy, Third Edition*. vol. 13 (2008).

94. Michaelian, K. & Simeonov, A. Fundamental molecules of life are pigments which arose and co-evolved as a response to the thermodynamic imperative of dissipating the prevailing solar spectrum. *Biogeosciences* **12**, 4913–4937 (2015).

95. Chow, A., Kennedy, J. C., Pottier, R. & Truscott, T. G. Rhodamine 123 - photophysical and photochemical properties. *Photobiochem. Photobiophys.* **11**, 139–148 (1986).

96. Pal, P. *et al*. Spectroscopic and photophysical properties of some new rhodamine derivatives in cationic, anionic and neutral micelles. *J. Photochem. Photobiol. Chem*. **98**, 65–72 (1996).

97. Kubin, R. F. & Fletcher, A. N. Fluorescence quantum yields of some rhodamine dyes. *J. Lumin*. **27**, 455–462 (1982).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

98. Chen, L. B. Chapter 7 Fluorescent Labeling of Mitochondria. in *Methods in Cell Biology* (eds. Wang, Y.-L., Taylor, D. L. & Jeon, K. W.) vol. 29 103–123 (Academic Press, 1988).

99. Huang, M., Camara, A. K. S., Stowe, D. F., Qi, F. & Beard, D. A. Mitochondrial Inner Membrane Electrophysiology Assessed by Rhodamine-123 Transport and Fluorescence. *Ann. Biomed. Eng.* **35**, 1276–1285 (2007).

100. Biasutto, L., Dong, L.-F., Zoratti, M. & Neuzil, J. Mitochondrially targeted anti-cancer agents. *Mitochondrion* **10**, 670–681 (2010).

101. Novo, M., Felekyan, S., Seidel, C. A. M. & Al-Soufi, W. Dye-Exchange Dynamics in Micellar Solutions Studied by Fluorescence Correlation Spectroscopy. *J. Phys. Chem. B* **111**, 3614–3624 (2007).

102. Petrášek, Z. & Schwille, P. Precise Measurement of Diffusion Coefficients using Scanning Fluorescence Correlation Spectroscopy. *Biophys. J.* **94**, 1437–1448 (2008).

103. Gendron, P.-O., Avaltroni, F. & Wilkinson, K. J. Diffusion Coefficients of Several Rhodamine Derivatives as Determined by Pulsed Field Gradient–Nuclear Magnetic Resonance and Fluorescence Correlation Spectroscopy. *J. Fluoresc.* **18**, 1093–1101 (2008).

104. PhotochemCAD. https://www.photochemcad.com/.

105. Jungbauer, L. M., Yu, C., Laxton, K. J. & LaDu, M. J. Preparation of fluorescently-labeled amyloid-beta peptide assemblies: the effect of fluorophore conjugation on structure and function. *J. Mol. Recognit. JMR* **22**, 403 (2009).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

106. Hnedzko, D., McGee, D. W. & Rozners, E. Synthesis and properties of peptide nucleic acid labeled at the N-terminus with HiLyte Fluor 488 fluorescent dye. *Bioorg. Med. Chem.* **24**, 4199–4205 (2016).

107. Freire Rodríguez, S. Estudio de la asociación de sistemas supramoleculares mediante técnicas de fluorescencia: de ciclodextrinas a beta-amiloides (Study of the association of supramolecular systems through fluorescence techniques: from cyclodextrins to beta-amyloids), PhD-thesis. *PhD Thesis* vol. PhD Thesis (University of Santiago de Compostela, Santiago de Compostela, 2015).

108. Zhang, Y. *et al.* Interaction Between the Glymphatic System and α-Synuclein in Parkinson's Disease. *Mol. Neurobiol.* **60**, 2209–2222 (2023).

109. Esbjörner, E. K. *et al.* Direct Observations of Amyloid b Self-Assembly in Live Cells Provide Insights into Differences in the Kinetics of Ab (1–40) and Ab (1–42) Aggregation. *Chem. Biol.* **21**, 1–11 (2014).

110. Hemion, C. *et al.* Clearance of neurotoxic peptides and proteins by meningothelial cells. *Exp. Cell Res.* **396**, 112322 (2020).

111. Oasa, S. *et al.* Small Molecule Decoys of Aggregation for Elimination of Aβ-Peptide Toxicity. *ACS Chem. Neurosci.* **14**, 1575–1584 (2023).

112. Freire, S., De Araujo, M. H., Al-Soufi, W. & Novo, M. Photophysical study of Thioflavin T as fluorescence marker of amyloid fibrils. *Dyes Pigments* **110**, 97–105 (2014).

113. Stsiapura, V. I. *et al.* Thioflavin T as a Molecular Rotor: Fluorescent Properties of Thioflavin T in Solvents with Different Viscosity. *J. Phys. Chem. B* **112**, 15893–15902 (2008).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

114. Amdursky, N., Erez, Y. & Huppert, D. Molecular rotors: what lies behind the high sensitivity of the thioflavin-T fluorescent marker. *Acc. Chem. Res.* **45**, 1548–1557 (2012).

115. Murudkar, S., Mora, A. K., Singh, P. K. & Nath, S. Ultrafast molecular rotor: an efficient sensor for premelting of natural DNA. *Chem. Commun.* **48**, 5301–5303 (2012).

116. LeVine, H. Thioflavine T interaction with synthetic Alzheimer's disease beta-amyloid peptides: detection of amyloid aggregation in solution. *Protein Sci. Publ. Protein Soc.* **2**, 404–410 (1993).

117. Bourhim, M., Kruzel, M., Srikrishnan, T. & Nicotera, T. Linear quantitation of Abeta aggregation using Thioflavin T: reduction in fibril formation by colostrinin. *J. Neurosci. Methods* **160**, 264–268 (2007).

118. Bertoncini, C., W. & Soledad Celej, M. Small Molecule Fluorescent Probes for the Detection of Amyloid Self-Assembly In Vitro and In Vivo. *Curr. Protein Pept. Sci.* **12**, 206–220 (2011).

119. Aliyan, A., Cook, N. P. & Martí, A. A. Interrogating Amyloid Aggregates using Fluorescent Probes. *Chem. Rev.* **119**, 11819–11856 (2019).

120. Cheng, Y. *et al.* A pyrane based fluorescence probe for noninvasive prediction of cerebral β-amyloid fibrils. *Bioorg. Med. Chem. Lett.* **25**, 4472–4476 (2015).

121. Mora, A. K., Singh, P. K., Patro, B. S. & Nath, S. PicoGreen: a better amyloid probe than Thioflavin-T. *Chem. Commun.* **52**, 12163–12166 (2016).

122. Pandey, S. P. & Singh, P. K. Basic Orange 21: A molecular rotor probe for fluorescence turn-on sensing of amyloid fibrils. *J. Mol. Liq.* **303**, 112618 (2020).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

123. Needham, L. M. *et al*. ThX – a next-generation probe for the early detection of amyloid aggregates. *Chem. Sci.* **11**, 4578–4583 (2020).

124. Verma, S., Ravichandiran, V. & Ranjan, N. Beyond amyloid proteins: Thioflavin T in nucleic acid recognition. *Biochimie* **190**, 111–123 (2021).

125. Yeasmin Khusbu, F., Zhou, X., Chen, H., Ma, C. & Wang, K. Thioflavin T as a fluorescence probe for biosensing applications. *TrAC Trends Anal. Chem.* **109**, 1–18 (2018).

126. Liu, S., Peng, P., Wang, H., Shi, L. & Li, T. Thioflavin T binds dimeric parallel-stranded GA-containing non-G-quadruplex DNAs: a general approach to lighting up double-stranded scaffolds. *Nucleic Acids Res.* **45**, 12080–12089 (2017).

127. Das, S. & Purkayastha, P. Selective Binding of Thioflavin T in Sequence-Exchanged Single Strand DNA Oligomers and Further Interaction with Phospholipid Membranes. *ChemistrySelect* **2**, 5000–5007 (2017).

128. Mohanty, J. *et al*. Thioflavin T as an efficient inducer and selective fluorescent sensor for the human telomeric G-quadruplex DNA. *J. Am. Chem. Soc.* **135**, 367–76 (2013).

129. Keri, R. S., Patil, M. R., Patil, S. A. & Budagumpi, S. A comprehensive review in current developments of benzothiazole-based molecules in medicinal chemistry. *Eur. J. Med. Chem.* **89**, 207–251 (2015).

130. Gill, R. K., Rawal, R. K. & Bariwal, J. Recent advances in the chemistry and biology of benzothiazoles. *Arch. Pharm. (Weinheim)* **348**, 155–178 (2015).

131. Koppolu, V. Synthetic and Biological Applications of Benzothiazole Phosphonates. *Heterocycl. Compd. Biol. Appl.* (2017).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

132. Henary, M., Paranjpe, S. & Owens, E. A. Synthesis and applications of benzothiazole containing cyanine dyes. *Heterocycl. Commun.* **19**, 1–11 (2013).

133. Jadhav, A. S. *et al.* Firefly luciferin precursor 2-cyano-6-hydroxybenzothiazole: Fluorescence à la carte controlled by solvent and acidity. *Dyes Pigments* **177**, 108285 (2020).

134. Magde, D., Elson, E. L. & Webb, W. W. Thermodynamic Fluctuations in a Reacting System - Measurement by Fluorescence Correlation Spectrsocopy. *Phys. Rev. Lett.* **29**, 705 (1972).

135. Yu, L. *et al.* A Comprehensive Review of Fluorescence Correlation Spectroscopy. *Front. Phys.* **9**, (2021).

136. Elson, E. L. Fluorescence Correlation Spectroscopy: Past, Present, Future. *Biophys. J.* **101**, 2855–2870 (12AD).

137. Valeur, B. & Berberan Santos, M. N. *Molecular Fluorescence: Principles and Applications, Second Edition*. (Wiley VCH, Weinheim, 2012).

138. Koenig, M. Multi-parameter-fluorescence-spectroscopic investigations of molecular dynamics on single biomolecules. (Heinrich-Heine Universität, 2006).

139. Rigler, R. & Elson, E. S. *Fluorescence Correlation Spectroscopy: Theory and Applications*. (Springer Verlag, Berlin, 2001).

140. Bordello, J., Novo, M. & Al-Soufi, W. Exchange-dynamics of a neutral hydrophobic dye in micellar solutions studied by Fluorescence Correlation Spectroscopy. *J. Colloid Interface Sci.* **345**, 369–376 (2010).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

141. Granadero, D., Bordello, J., Pérez-Alvite, M. J., Novo, M. & Al-Soufi, W. Host-Guest Complexation Studied by Fluorescence Correlation Spectroscopy: Adamantane–Cyclodextrin Inclusion. *Int. J. Mol. Sci.* **11**, 173–188 (2010).

142. Culley, S., Tosheva, K. L., Matos Pereira, P. & Henriques, R. SRRF: Universal live-cell super-resolution microscopy. *Int. J. Biochem. Cell Biol.* **101**, 74–79 (2018).

143. Shashkova, S. & Leake, M. C. Single-molecule fluorescence microscopy review: shedding new light on old problems. *Biosci. Rep.* **37**, BSR20170031 (2017).

144. Prakash, K., Diederich, B., Heintzmann, R. & Schermelleh, L. Super-resolution microscopy: a brief history and new avenues. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **380**, 20210110 (2022).

145. Alva, A. *et al.* Fluorescence fluctuation-based super-resolution microscopy: Basic concepts for an easy start. *J. Microsc.* **288**, 218–241 (2022).

146. Gustafsson, N. *et al.* Fast live-cell conventional fluorophore nanoscopy with ImageJ through super-resolution radial fluctuations. *Nat. Commun.* **7**, 12471 (2016).

147. Laine, R. F. *et al.* High-fidelity 3D live-cell nanoscopy through data-driven enhanced super-resolution radial fluctuation. *Nat. Methods* **20**, 1949–1956 (2023).

148. Du, Y.-C., Zhu, L.-N. & Kong, D.-M. Label-free thioflavin T/G-quadruplex-based real-time strand displacement amplification for biosensing applications. *Biosens. Bioelectron.* **86**, 811–817 (2016).

149. Xu, S. *et al.* Thioflavin T as an efficient fluorescence sensor for selective recognition of RNA G-quadruplexes. *Sci. Rep.* **6**, 1–9 (2016).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

150. Renaud de la Faverie, A., Guedin, A., Bedrat, A., Yatsunyk, L. A. & Mergny, J.-L. Thioflavin T as a fluorescence light-up probe for G4 formation. *Nucleic Acids Res.* **42**, e65 (2014).

151. Stine, W. B., Dahlgren, K. N., Krafft, G. A. & LaDu, M. J. In vitro characterization of conditions for amyloid-beta peptide oligomerization and fibrillogenesis. *J. Biol. Chem.* **278**, 11612–11622 (2003).

152. Stine, W. B., Jungbauer, L., Yu, C. & Ladu, M. J. Preparing synthetic Aβ in different aggregation states. *Methods Mol. Biol. Clifton NJ* **670**, 13–32 (2011).

153. Ryan, T. M. *et al.* Ammonium hydroxide treatment of Aβ produces an aggregate free solution suitable for biophysical and cell culture characterization. *PeerJ* **1**, e73 (2013).

154. Novo, M., Pérez-González, C., Freire, S. & Al-Soufi, W. Early Aggregation of Amyloid-β(1–42) Studied by Fluorescence Correlation Spectroscopy. in *Protein Aggregation* (ed. Cieplak, A. S.) vol. 2551 1–14 (Springer US, New York, NY, 2023).

155. Jover, A., Meijide, F., Mosquera, V. & Vázquez Tato, J. A Step-by-Step Dilution Extraction Method for Laboratory Experiments. *J. Chem. Educ.* **67**, 530–532 (6AD).

156. Reija, B. Análisis avanzado de espectros de fluorescencia resueltos en el tiempo. (Universidade de Santiago de Compostela, Santiago de Compostela, 2003).

157. Al-Soufi, W., Novo, M. & Mosquera, M. Principal component global analysis of fluorescence and absorption spectra of 2-(2'-hydroxyphenyl)benzimidazole. *Appl. Spectrosc.* **55**, 630–636 (2001).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

158. López Hortas, M. I. Análisis global de series de espectros de fluorescencia. Determinación de intervalos de confianza de parámetros fisicoquímicos mediante simulaciones de Monte Carlo. (Universidade de Santiago de Compostela, 25AD).

159. Al-Soufi, W. *et al*. Fluorescence Correlation Spectroscopy, a Tool to investigate Supramolecular Dynamics: Inclusion Complexes of Pyronines with Cyclodextrin. *J. Am. Chem. Soc.* **127**, 8775–8784 (2005).

160. Frangi, A. F., Niessen, W. J., Hoogeveen, R. M., Van Walsum, T. & Viergever, M. A. Model-based quantitation of 3-D magnetic resonance angiographic images. *IEEE Trans. Med. Imaging* **18**, 946–956 (1999).

161. Sato, Y. *et al*. Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Med. Image Anal.* **2**, 143–168 (1998).

162. FeatureJ. https://imagescience.org/meijering/software/featurej/.

163. Cantor, C. R. & Schimmel, P. R. *Biophysical Chemistry: The Behavior of Biological Macromolecules*. (W.H. Freeman & Company, 1980).

164. Novo, M., Pérez-González, C., Freire, S. & Al-Soufi, W. Early aggregation of Amyloid-β (1-42) studied by Fluorescence Correlation Spectroscopy. in *Methods in Molecular Biology. Protein Aggregation: Methods and Protocols* (ed. Cieplak, A.) (Springer Nature Experiments, 2022).

165. Wilkins, D. K. *et al*. Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques. *Biochemistry* **38**, 16424–16431 (1999).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

166. Limpert, E., Stahel, W. A. & Abbt, M. Log-normal distributions across the sciences: Keys and clues. *Bioscience* **51**, 341–352 (2001).

167. Al-Soufi, W., Piñeiro, L. & Novo, M. A model for monomer and micellar concentrations in surfactant solutions: Application to conductivity, NMR, diffusion, and surface tension data. *J. Colloid Interface Sci.* **370**, 102–110 (2012).

168. Tran, P. L. T., Largy, E., Hamon, F., Teulade-Fichou, M.-P. & Mergny, J.-L. Fluorescence intercalator displacement assay for screening G4 ligands towards a variety of G-quadruplex structures. *Biochimie* **93**, 1288–96 (2011).

169. Sen, D. & Gilbert, W. A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* **344**, 410–414 (1990).

170. Bhattacharyya, D., Arachchilage, G. M. & Basu, S. Metal cations in G-quadruplex folding and stability. *Front. Chem.* **4**, 1–14 (2016).

171. Bao, H.-L. & Xu, Y. Telomeric DNA-RNA-hybrid G-quadruplex exists in environmental conditions of HeLa cells. *Chem. Commun. Camb. Engl.* **56**, 6547–6550 (2020).

172. Liu, L., Shao, Y., Peng, J., Liu, H. & Zhang, L. Selective recognition of ds-DNA cavities by a molecular rotor: switched fluorescence of thioflavin T. *Mol. Biosyst.* **9**, 2512–2519 (2013).

173. Wang, H., Peng, P., Liu, S. & Li, T. Thioflavin T behaves as an efficient fluorescent ligand for label-free ATP aptasensor. *Anal. Bioanal. Chem.* **408**, 7927–7934 (2016).

174. Danielsson, J., Jarvet, J., Damberg, P. & Graslund, A. Translational diffusion measured by PFG-NMR on full length and fragments of the Alzheimer Ab(1-40)

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

peptide. Determination of hydrodynamic radii of random coil peptides of varying length. *Magn. Reson. Chem.* **40**, S89–S97 (2002).

175. Leite, J. P., Gimeno, A., Taboada, P., Jiménez-Barbero, J. J. & Gales, L. Dissection of the key steps of amyloid-β peptide 1–40 fibrillogenesis. *Int. J. Biol. Macromol.* **164**, 2240–2246 (2020).
176. Morel, B., Carrasco-Jiménez, M. P., Jurado, S. & Conejero-Lara, F. Rapid Conversion of Amyloid-Beta 1-40 Oligomers to Mature Fibrils through a Self-Catalytic Bimolecular Process. *Int. J. Mol. Sci.* **22**, 6370 (2021).
177. Balasco, N. *et al.* A Comprehensive Analysis of the Intrinsic Visible Fluorescence Emitted by Peptide/Protein Amyloid-like Assemblies. *Int. J. Mol. Sci.* **24**, 8372 (2023).
178. Xia, Q. *et al.* A historical review of aggregation-induced emission from 2001 to 2020: A bibliometric analysis. *Aggregate* **3**, e152 (2022).
179. Singh, A., Khatun, S. & Nath Gupta, A. Simultaneous Detection of Tyrosine and Structure-Specific Intrinsic Fluorescence in the Fibrillation of Alzheimer's Associated Peptides. *ChemPhysChem* **21**, 2585–2598 (2020).
180. Pinotsi, D. *et al.* Proton Transfer and Structure-Specific Fluorescence in Hydrogen Bond-Rich Protein Structures. *J. Am. Chem. Soc.* **138**, 3046–3057 (2016).
181. Fricano, A., Librizzi, F., Rao, E., Alfano, C. & Vetri, V. Blue autofluorescence in protein aggregates "lighted on" by UV induced oxidation. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1867**, 140258 (2019).
182. Mirón, G. D. *et al.* The carbonyl-lock mechanism underlying non-aromatic fluorescence in biological matter. *Nat. Commun.* **14**, 7325 (2023).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

183. Ahmad, A., Uversky, V. N. & Khan, R. H. Aberrant liquid-liquid phase separation and amyloid aggregation of proteins related to neurodegenerative diseases. *Int. J. Biol. Macromol.* **220**, 703–720 (2022).

184. Gui, X. *et al.* Liquid–liquid phase separation of amyloid-β oligomers modulates amyloid fibrils formation. *J. Biol. Chem.* **299**, (2023).

185. Lee, J. *et al. The Role of Positively Charged Small Biomolecules in the Aggregation of Hyperphosphorylated Tau in Alzheimer's Disease*. http://biorxiv.org/lookup/doi/10.1101/2023.10.19.563032 (2023).

186. Mukherjee, S., Poudyal, M., Dave, K., Kadu, P. & K. Maji, S. Protein misfolding and amyloid nucleation through liquid–liquid phase separation. *Chem. Soc. Rev.* **53**, 4976–5013 (2024).

187. Al-Soufi, W., Novo, M., Mosquera, M. & Rodríguez-Prieto, F. Principal Component Global Analysis of Series of Fluorescence Spectra. in 23–45 (Springer, New York, NY, 2011).

188. Xu, X. *et al.* Applications of human and bovine serum albumins in biomedical engineering: A review. *Int. J. Biol. Macromol.* **253**, 126914 (2023).

189. Wang, P.-Y., Yang, C.-T. & Chu, L.-K. Differentiating the protein dynamics using fluorescence evolution of tryptophan residue(s): A comparative study of bovine and human serum albumins upon temperature jump. *Chem. Phys. Lett.* **781**, 138998 (2021).

190. Tayeh, N., Rungassamy, T. & Albani, J. R. Fluorescence spectral resolution of tryptophan residues in bovine and human serum albumins. *J. Pharm. Biomed. Anal.* **50**, 107–116 (2009).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

191. Nishi, K., Yamasaki, K. & Otagiri, M. Serum Albumin, Lipid and Drug Binding. in *Vertebrate and Invertebrate Respiratory Proteins, Lipoproteins and other Body Fluid Proteins* (eds. Hoeger, U. & Harris, J. R.) vol. 94 383–397 (Springer International Publishing, Cham, 2020).

192. Ketrat, S., Japrung, D. & Pongprayoon, P. Exploring how structural and dynamic properties of bovine and canine serum albumins differ from human serum albumin. *J. Mol. Graph. Model.* **98**, 107601 (2020).

193. Shaghaghi, M. *et al.* Exploring the interactions of a Tb(III)–quercetin complex with serum albumins (HSA and BSA): spectroscopic and molecular docking studies. *Luminescence* **35**, 512–524 (2020).

194. Maier, R., Fries, M. R., Buchholz, C., Zhang, F. & Schreiber, F. Human versus Bovine Serum Albumin: A Subtle Difference in Hydrophobicity Leads to Large Differences in Bulk and Interface Behavior. *Cryst. Growth Des.* **21**, 5451–5459 (2021).

195. Moriyama, Y., Ohta, D., Hachiya, K., Mitsui, Y. & Takeda, K. Fluorescence behavior of tryptophan residues of bovine and human serum albumins in ionic surfactant solutions: A comparative study of the two and one tryptophan(s) of bovine and human albumins. *J. Protein Chem.* **15**, 265–272 (1996).

196. Guria, S. *et al.* Small-Molecule Probe for Sensing Serum Albumin with Consequential Self-Assembly as a Fluorescent Organic Nanoparticle for Bioimaging and Drug-Delivery Applications. *ACS Appl. Bio Mater.* **3**, 3099–3113 (2020).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

197. Starosta, R., Santos, F. C. & de Almeida, R. F. M. Human and bovine serum albumin time-resolved fluorescence: Tryptophan and tyrosine contributions, effect of DMSO and rotational diffusion. *J. Mol. Struct.* **1221**, 128805 (2020).

198. Liu, B. *et al.* Fluorescence discrimination of HSA from BSA: A close look at the albumin-induced restricted intramolecular rotation of flavonoid probe. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **264**, 120306 (2022).

199. Dai, Y., Gong, J., Cao, J., Chen, W. & Fu, N. A deep-red fluorescent probe based on naphthalimide for discrimination of HSA from BSA and tracking HSA by bioimaging. *Dyes Pigments* **222**, 111893 (2024).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# Annex I: List of publications

## 1) Spectroscopic Characterisation of Mitochondrial DNA G-Quadruplexes

International Journal of Molecular Sciences. **2022**, 23, 925. DOI: https://doi.org/10.3390/ijms23020925

### Contributions to the paper:

Data: data acquisition and curation, data formal analysis and data visualization. Writing: original draft preparation, review and editing.

### Journal quality indicators:

The journal where section 5.1 has been published presented the following journal quality indicators in 2022: an impact factor (JIF) of 5.6 (obtained by Clarivate), an SRJ 1.15 (calculated by Scimago), a CiteScore of 7.8 (Scopus, https://www.scopus.com/sources.uri). The journal is placed in the first quartile (Q1) in Biochemistry & Molecular Biology (Clarivate) and Spectroscopy (Scimago).

### Permission to reproduce:

The journal is part of the MDPI editorial which publishes all its articles in open access (https://www.mdpi.com/openaccess). As stated by the journal: "No special permission is required to reuse all or part of the article published by MDPI, including figures and tables. For articles published under an open access Creative Common CC BY license, any part of the article may be reused without permission provided that the original article is clearly cited."

Open Access Article

**Spectroscopic Characterization of Mitochondrial G-Quadruplexes**

by **Sara Illodo** [1,2], **Cibrán Pérez-González** [1,2], **Ramiro Barcia** [3], **Flor Rodríguez-Prieto** [2], **Wajih Al-Soufi** [1] and **Mercedes Novo** [1,*]

[1] Departamento de Química Física, Facultade de Ciencias, Universidade de Santiago de Compostela, 27002 Lugo, Spain

[2] Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Física, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain

[3] Departamento de Bioquímica e Bioloxía Molecular, Facultade de Veterinaria, Universidade de Santiago de Compostela, 27002 Lugo, Spain

[*] Author to whom correspondence should be addressed.

*Int. J. Mol. Sci.* **2022**, *23*(2), 925; **https://doi.org/10.3390/ijms23020925**

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

# Annex II: Rights and permissions

- **Section 5.1** of **Results and discussion**, **Table 3.3**, **section 3.3.1**, **section 4.5.1** and **section 4.5.2** contents are a reproduction of the original work: Spectroscopic Characterization of Mitochondrial G-Quadruplexes. Sara Illodo, Cibrán Pérez-González, Ramiro Barcia, Flor Rodríguez-Prieto, Wajih Al-Soufi and Mercedes Novo. *International Journal of Molecular Sciences* **2022**, *23*(2). This is an open access article distributed under the terms of the Creative Commons CC BY license (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
- **Figure 1.1** was constructed with data obtained from reference 12.
- **Figure 1.2** and **Figure 1.3** are reproductions from the original work: Spectroscopic Characterization of Mitochondrial G-Quadruplexes. Sara Illodo, Cibrán Pérez-González, Ramiro Barcia, Flor Rodríguez-Prieto, Wajih Al-Soufi and Mercedes Novo. *International Journal of Molecular Sciences* **2022**, *23*(2). This is an open access article distributed under the terms of the Creative Commons CC BY license (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
- **Figure 1.4** is an adaptation of Figure 3 from reference 54 and Figure 1 from reference 18. Both works are open access articles distributed under the terms of the Creactive Commos licenses. Reference 54 is published under under the terms of the Creative Commons CC BY license (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Reference 18 is published under Creative Commos Attribution-NonCommercial-ShareAlike License (CC BY-NS-SA), which allows to distribute, remix, adapt, and build upon the material in any medium or format for noncommercial purposes only, and only so long as attribution is given to the creator. The material that was remixed, adapt, or built upon the material, must be licensed the modified material under identical terms.
- **Figure 3.8** and **Figure 3.9** are property of USC-BioPhysChem research group. They are license under the terms of Creative Commos Attribution-NonCommercial-ShareAlike License (CC BY-NS-SA), which allows to distribute, remix, adapt, and build upon the material in any medium or format for noncommercial purposes only, and only so long as attribution is given to the creator.
- **Figure 3.11** is a reproduction of Figure 7 from reference 145. This is an open-access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Neurodegenerative diseases represent a major health problem. Biomolecules involved in these disorders are essential for life, but the misregulation of their interactions can lead to severe damage it. Thus, understanding these biomolecular interactions is vital for developing effective treatments. In this thesis, we study the self-assembly and binding processes of biomolecules like G-quadruplexes, amyloid peptides, and blood plasma proteins, which are thought to be involved in Alzheimer's disease. Using microscopic and spectroscopic techniques, we study the mechanisms behind the self-assembly of G-Quadruplexes, the aggregation of amyloid peptides and the host-guest association of blood plasma proteins with fluorescence probes.