Sapienza Università di Roma
Facoltà di Scienze Matematiche, Fisiche e Naturali

Dottorato di ricerca in Genetica e Biologia Molecolare
XXXV ciclo (A.A. 2023/2024)

The effect of RNA on alpha synuclein aggregation

Dottorando:

Jakob Rupert

Relatore:

Prof. Gian Gaetano Tartaglia

Coordinatrice:

Prof.ssa Isabella Saggio

Doktorsko nalogo posvečam spominu na mojega dedka, Prof. dr. Zorana Železnika (9.11.1930 – 13.3.2020)

This work is dedicated to the memory of my grandfather, Prof. dr. Zoran Železnik (9.11.1930 – 13.3.2020)

Ma io penso diverso, cancello di nuovo
Cerco quello che non trovo
Non ubbidisco, esco dal coro
Distruggo e ricostruisco chi sono
(Colle der Fomento, Penso Diverso)

Anlatmam derdimi dertsiz insana
Dert çekmeyen dert kıymetin bilemez
Derdim bana derman imiş bilmedim
Hiçbir zaman gül dikensiz olamaz, aman aman
(Aşık Veysel Şatıroğlu)

Acknowledgements

First and foremost, I would like to thank my mentor, Prof. Gian Gaetano Tartaglia, for the contagious enthusiasm which makes every result a story worth telling, and every discussion a chance to delve a little bit deeper into all the mysteries of the universe. Thank you for the freedom and consideration you gave me and my ideas, for the willingness to discuss and chat about anything and nothing. *Cum grano salis*, as always. I am eagerly awaiting your interview on the Lex Fridman podcast, to play it on repeat while I run towards new adventures.

Endless thanks go to Elsa, who was the single person most responsible for me even starting this journey, still in London. Thank you for always treating me as an equal, even when I had no business to be even close, for the incredible support you gave me, the objective and rational approach that finally channelled my chaotic mind into something vaguely resembling a system and by always leading by example both in kindness, steadfastness, and approach to life. Every bit of my success is just as much yours.

I would also like to thank Prof. Alfonso De Simone and Prof. Stefano Ricagno for agreeing to act as external revisors for the thesis. They themselves have made important contributions to the understanding of the topics covered in the dissertation and it is both a pleasure and an honour to have them reviewing the work presented.

A big thank you goes to all the lab members in Genova, Núria, Laura, Guy and Martina (in chronological order). Thanks for all the help, support and company throughout the years, for tolerating my endless shaking of aggregation samples and all of my music choices.

Another huge thank you for all the other members of the Tartaglia group. Giorgio, Magda, Michele, Alessio, Jonny, Gabriele, Andrea, Alex and Elias, you have all contributed in bigger or smaller steps to everything that is written on the following pages.

Tantissimi ringraziamenti a tutti i bravi fisici di Roma, Valeria, Antonia, Eleonora, Raffaella e Michele, che mi aveste accolto nel vostro lab, e che vi siate lasciati convincere la storia con la sinucleina, infatti, potesse portarci a qualche parte. Non dimenticherò mai né la skyrella, né le spiegazioni del quantum cascade laser.

Mille grazie anche ad Eleonora e Sabrina qui a Genova, per tutto il vostro aiuto con la microscopia, per l'opportunità di condividere gli impegni, i momenti bassi quando le bolle non si formavano, e gli sguardi confusi quando vedevamo le ciambelle sullo schermo.

Vorrei ringraziare a tutti gli altri membri dei laboratori dove ho lavorato durante gli ultimi anni. Ancora nella fase nomadica a Roma, grazie a Monica Ballarino per avermi accolto all'inizio nel lab 2 per fare i clonaggi, a Giacomo Parisi per avermi accompagnato e aiutato organizzare tutto al CLNS e Dipartimento di Biochimica. Finalmente, tantissimi ringraziamenti a tutti nei laboratori di CHT a Genova per tutto il tempo passato insieme, gli aperitivi e la compagnia anche nei momenti difficili. Soprattutto vorrei ringraziare Edo e Yeraldin a Genomica, per tutto l'aiuto, disponibiltà e pazienza con, in retrospettiva, assolutamente troppi campioni in cosa qualcuno avrebbe potuto descrivere come relazione abusiva :)

Un grandissimo ringrazio va anche a Paolo, per avermi supportato tutto questo tempo anche alla distanza. Da Anzio, Parco degli Acquedotti, la tua bellissima Roma, ma poi anche a Genova, nelle Marche, le altezze del Gran Sasso, l'ambiente gangsta del nostro hotel a Queens, quello aristocratico del Cold Spring Harbor a quello alternativo del Brooklyn Park, in aereo, macchina, o sul motorino, non dimenticherò mai tutte le cose che abbiamo fatto insieme!

Seveda ne gre tudi brez slovenščine in zahvale Juretu in Vidu za ves napor, ki gre v ohranjanje stikov. Od smučanja na Rogli, večernih iger, zagovorov doktorskih nalog in na koncu še poroke, v teh štirih letih smo imel res celo koalicijo izzivov in hvala res za vse!

Na koncu gre neizmerna zahvala tudi mojim, za neizčrpno potrpežljivost in ljubezen in ne samo v zadnjih štirih letih, ko sem vam pobegnil čez mejo. Vse, kar sem se naučil, vse kar me zanima in povsod, kjer sem bil, je samo posledica vzgoje in podpore ljubeče družine. Hvala, ker ste vedno verjeli vame, mi dali vso širino znanja, zanimanja in radovednosti in vso podporo, da sem to lahko dosegel. Vas mam nns radajc xxx

E alla fine, il ringrazio più grande e più profondo va a Rebecca. Il nostro percorso era per caso iniziato nello stesso tempo del dottorato, e coincidetemene si è evoluto in qualcosa ancora più grande, qualcosa per sempre. Era per te che sono passato tutti gli ostacoli, era per te che sono diventato chi sono oggi e per la tua pazienza ed enorme capacità di fidarsi che ci siamo ancora, alla distanza, ma più vicini che mai prima. E questo mi significa e dà più motivazione che qualsiasi altra cosa.

Table of contents

Acknowledgements IV
Table of contents VI
List of abbreviations VIII
Summary X
1. Introduction 1
1.1 Protein aggregation and misfolding 1
1.1.1 Molecular mechanisms of protein aggregation 3
1.1.2 Structural characteristics of protein aggregates 5
1.1.3 The kinetics of protein aggregation 7
1.1.4 Aggregation in physiology and functional amyloids 10
1.1.5 Aggregation in disease 11
1.1.6 The role of RNA in protein aggregation 12
1.2 Alpha synuclein 13
1.2.1 The biochemical characteristics of aS 14
1.2.2 The aggregation of aS 15
1.2.3 aS and nucleic acids 16
2. Aims 18
3. Results 19
3.1 RNA accelerates the aggregation of aS 19
3.1.1 The N-terminal of aS is predicted to interact with RNA 19
3.1.2 Optimising the aggregation protocol for aS 20
3.1.3 RNA accelerates the aggregation of both aS103 and aS140 22
3.2 RNA affects the structural properties of aS aggregates 26
3.2.1 Introduction and overview of infrared spectroscopy and its applications 26
3.2.2 Characterisation of the influence of RNA on aS aggregation with micro-FTIR 28
3.2.3 Characterisation of fibril mechanical properties with AFM-IR 33
3.2.4 RNA speeds up the aggregation and changes the fibril conformation of aS 35
3.3 aS sequesters RNA upon aggregation 36
3.3.1 aS103 aggregates sequester larger quantities of RNA than aS140 36
3.3.2 RNA is sequestered to a larger extent than DNA in aS aggregates 37

3.3.3 Monomeric aS does not bind RNA 40
3.4 Aggregation increases the RNA-binding propensity of amyloidogenic proteins 43
3.4.1 Short introduction 43
3.4.2 Calculations recapitulate experimental results on the ability of tau fibrils to attract RNA 43
3.4.3 Physicochemical properties of aS aggregates 47
3.4.4 Summary and conclusions 49
3.4.5 Closing remarks 49
3.5 RNA can modulate the enzymatic activity of aS amyloid assemblies 51
3.5.1 Identification of the putative catalytic site in aS fibrils 51
3.5.2 aS fibrils have comparable catalytic activity to reported values 52
3.5.3 RNA inhibits the catalytic activity of aS fibrils 55
3.5.4 RNA changes the morphology of aS fibrils to render them catalytically inactive upon coaggregation 57
4. Discussion 58
5. Conclusions and future perspectives 65
6. Materials and Methods 68
6.1 Preparation of bacterial expression vectors for aS 68
6.2 Bacterial expression of recombinant aS 69
6.3 Isolation and purification of aS constructs 70
6.4 Protein aggregation assays 71
6.5 Analysis of aggregation data 72
6.6 Nucleic acid quantification and extraction from aS aggregates 73
6.7 Biolayer interferometry 74
6.8 Microscale thermophoresis 75
6.9 Fourier-transform infrared microspectroscopy (micro-FTIR) 76
6.10 Atomic force microscopy – infrared spectroscopy 76
6.11 Fibril enzymatic activity assays 77
6.12 Analysis of enzymatic assays data 78
7. List of publications 80
8. Bibliography 81

List of abbreviations

aS	alpha synuclein
RNA	ribonucleic acid
DNA	deoxyribonucleic acid
RBP	RNA-binding protein
AFM	atomic force microscopy
Aβ	amyloid beta
AD	Alzheimer's disease
IDP	intrinsically disordered protein
APR	aggregation-prone region
ThT	thioflavin T
cryo-EM	cryogenic electron microscopy
ssNMR	solid-state nuclear magnetic resonance
mRNA	messenger RNA
hnRNPDL-2	heterogenous nuclear ribonucleoprotein D-like
TDP-43	TAR-DNA binding protein with a molecular weight of 43 kDa
CPEB	cytoplasmic polyadenylation element-binding protein
ALS	amyotrophic lateral sclerosis
FTLD	frontotemporal lobar degeneration
FUS	protein fused in sarcoma
TMEM106b	transmembrane protein with a molecular mass of 106 kDa
MSA	multiple systems atrophy
SOD1	superoxide dismutase 1
rIGSRNA	ribosomal intergenic spacer noncoding RNA
aS140	alpha synuclein 1-140
aS103	alpha synuclein 1-103
ANOVA	analysis of variance
IR	infrared
micro-FTIR	Fourier-transform infrared microspectroscopy
AFM-IR	atomic forced microscopy coupled with infrared spectroscopy
BSA	bovine serum albumin
rEMSA	RNA electromobility shift assay
BLI	biolayer interferometry
MST	microscale thermophoresis
SERF1	small EDRK-rich factor 1
PDB	Protein Data Bank
AUC	area-under-curve
ROC	receiver operating characteristic
TPR	true positive rate
FPR	false positive rate
pNPA	para-nitrophenyl acetate
pNPP	para-nitrophenyl phosphate

pNP	para-nitrophenyl
LBD	Lewy-body dementia
LCD	low complexity domain
FCS	fluorescence correlation spectroscopy
PCR	polymerase chain reaction
LB	Luria broth
dsDNA	double-stranded DNA
IPTG	isopropyl thiogalactoside
CV	column volume
SDS PAGE	polyacrylamide gel electrophoresis in the presence of sodium dodecylsulphate
DTT	dithiotreitol
FWHM	full-width at half-maximum

Summary

Protein aggregation is a widespread phenomenon, closely linked to protein folding and function. A tightly regulated process, it can lead to a wide variety of polymorphic protein assemblies with incredible thermodynamics stability and mechanical properties. Protein aggregates appear in functional roles in all trees of life, from performing catalytic reactions to regulating gene expression, in processes that are thought to go back to the prebiotic stage of life. Aggregation has also attracted significant attention over the last decades due to the involvement of aggregated proteins in debilitating neurodegenerative diseases. One of these proteins is alpha synuclein, shown to form insoluble aggregated assemblies in a plethora of pathological conditions such as Parkinson's disease, Lewy-body dementia, multiple systems atrophy etc.

Nucleic acids can act as potent modulators of protein aggregation, and RNA has the ability to either hinder or facilitate protein assembly and regulate protein phase transitions. The aim of the research, describe in the following dissertation, was to assess the potential modulating effect RNA has on the aggregation of aS, with the long-term ambition of developing RNA-based therapeutic and diagnostic agents.

We have shown that RNA accelerates the amyloid assembly of aS with fluorescence-based aggregation assays and infrared spectroscopy by increasing the rate of aggregation and decreasing the aggregation lag time. We have also determined that RNA influences the structural characteristics of aS aggregates and alters their conformation and mechanical properties. Finally, we have determined that aS aggregated species both directly bind RNA, as well as sequester it during coaggregation in a protein-dependent manner. We have determined the higher affinity of aS fibrils for RNA compared to DNA by means of quantifying the nucleic acids extracted from aggregates after aggregation. We have also shown directly that RNA binds to the fibrils by inhibiting their enzymatic activity and we have determined the approximate binding affinity. In addition, we have computationally determined that aggregation leads to an increase in the aggregation propensity of proteins by partitioning of physicochemical properties in the inner and outer parts of the amyloid core. We have also shown that RNA- and DNA-binding domains tend to remain outside the amyloid core and can thus also contribute to specific RNA sequestration by canonic, aggregation-prone RBPs. Thus, we can conclude that RNA plays an important role and has a high potential as a modulator not only of aS aggregation, but of other aggregation-prone proteins in general.

1. Introduction

Since the early days of molecular biology, proteins have been identified as the "functional" entities of molecular biology, primarily due to the relative simplicity of measuring and characterising the catalytic function of enzymes. The view of proteins as the principal actors by which a cell can perform and maintain all the molecular processes that enable life was further enshrined in the Central Dogma of Molecular Biology, proposed by Francis Crick in 1958[1]. The view that information can flow unidirectionally from DNA to proteins via RNA, has since remained as a principle that all early-stage students encounter at the beginning of lessons explaining the basics of molecular biology. Another all-time classic example that usually follows in a lesson soon thereafter, is the Anfinsen experiment with an RNase, showing the correlation between a protein's amino acid sequence, its three-dimensional fold and its function[2]. While exceptions to these two principles were of course added over years and years of research, the basic message conveyed to the students remained: proteins are the cell's effectors and require to be folded into a stable, three-dimensional structure in order to perform their functions. In this context, the phenomenon of protein aggregation upon misfolding is of course presented as a disruptive process, associated with loss-of-function, disease and cell death. While this is by no means an inaccurate description, as is the case with the previously presented "dogmas", numerous exceptions and open questions were added to this line of thought. Aggregation has turned out to be a general concept of protein dynamics, basing itself on the exact same thermodynamic principles as folding itself[3]. The aggregated proteins themselves have also been shown to be functional entities that perform important functions in a variety of cellular mechanisms[4]. And, last-but-not-least, aggregation has been shown to be a strong, ambivalent factor in protein evolution[5]. The research, presented in this thesis, aims to tackle both sides of these ever-changing "dogmas". It examines the aggregation of alpha-synuclein (aS), a protein associated with multiple neurodegenerative disorders, with the aim of evaluating the feasibility of developing a novel platform for drug development. It also shows aggregation can also be a function-generating process, leading to proteins acquiring novel functions upon aggregation. And finally, it demonstrates the role of RNA in the process, highlighting the interconnectedness and interdependency of molecular mechanisms that govern such a dynamic and potentially deadly process as aggregation.

1.1 Protein aggregation and misfolding

The term aggregation can be used as the generic term that describes the clustering of proteins into large, insoluble macromolecular assemblies. These then tend to be subdivided into either amyloids or amorphous aggregates, which differ according to their structural characteristics[6]. While amyloids are generally easily identifiable due to their highly characteristic cross-β structure, amorphous aggregates tend to simply be classified as "non-amyloid" aggregates[7]. However, even this classic perception of

"β-aggregation" has been challenged lately by the discovery of an all-α amyloid by Tayeb-Fligelman and colleagues[8]. Therefore, the terms "aggregation", "aggregate" and "amyloid" have been used rather intermittently in the following text, with aggregation and aggregates used generally to describe macromolecular assemblies of aS, and amyloid and fibrils alluding specifically to the amyloid fibril structure as described below and, in the context of discussing results, referencing the mature, fibrillar assemblies of aS as confirmed by experimental analysis with atomic force microscopy (AFM).

The term amyloid was first used by the German medical doctor Rudolph Vierchow in 1854 to describe the filamentous inclusions he observed in the post-mortem brain of patients, known at the time as *corpora amylacea* [9]. After staining the fibrillar inclusions with iodine, a known method for detecting starch, he initially concluded the observed species are likely cellulose and gave them the name derived from the Latin (*amylum*) or Greek (*ἄμυλο*) word for starch[9]). The subsequent research showed these inclusions as having a very high nitrogen content and are thus likely proteic in origin, as well as later showing that not all of them are formed by the same peptides and likely have different origins[10]. Further breakthroughs came through the advances in methods of structural biology and the identification of various amyloidogenic proteins, with the most notable being the discovery of prions and their by Prusiner in 1982 [11]. In fact, it was prions that first contributed to the greater public awareness of aggregation through the mad cow disease scare in the late 1980s and 1990s. In parallel, the discovery of amyloid beta (Aβ) as the major component of amyloid plaques found in the brain of Alzheimer disease (AD) patients[12], first showed the involvement of aggregation in neurodegeneration, an association that continues to this day as the primary area of application for aggregation-associated studies. Onwards, the work of Sir Christopher Dobson and colleagues in Cambridge and Oxford greatly expanded the theoretical framework of both the molecular basis of protein misfolding and aggregation, the exact mathematical description of its associated processes, as well as contributed to the most important notion: misfolding and amyloid formation is not a "special case", but a general process, rooted in the thermodynamics of protein folding, that can be applied to any protein under specific conditions[13–15]. In recent years, the increasing awareness of the social burden of neurodegeneration and aggregation as one its primary hallmarks, have led to an explosion of new results, both by identifying new molecular mechanisms and cofactors that influence aggregation[3], the experimental methodology to study[16] and analytical methods to analyse it[17], the growth of prediction methodology[18] and the advances of population genetics leading to the identification of new risk factors and mutations[19,20]. Especially the advancement of the study of phase transitions, a field similar to aggregation in its booming transition from a couple of interesting "special cases" to a general cellular mechanism, has provided an important cellular framework for the molecular basis aggregation *in vivo* [21,22].

1.1.1 Molecular mechanisms of protein aggregation

Since Anfinsen's experiment, the general principle stands that the primary structure of proteins determines its three-dimensional fold. In terms of globular proteins, the primary structure determines the formation of the secondary structure elements, which then fold into a three-dimensional structure around the hydrophobic core to form a stable, globular structure[23]. Similarly, the discovery of intrinsically disordered proteins (IDPs) has also brought the understanding that primary structure elements of IDPs allow them to remain dynamically unstructured[24]. These structural transitions between thermodynamic states with corresponding changes in free energy are usually depicted through the so-called free energy landscapes[23,25]. The vertical axis of the free energy landscape represents the free energy of the corresponding folded state, while the horizontal axis shows the multitude of possible structural conformations (Fig. 1.1).

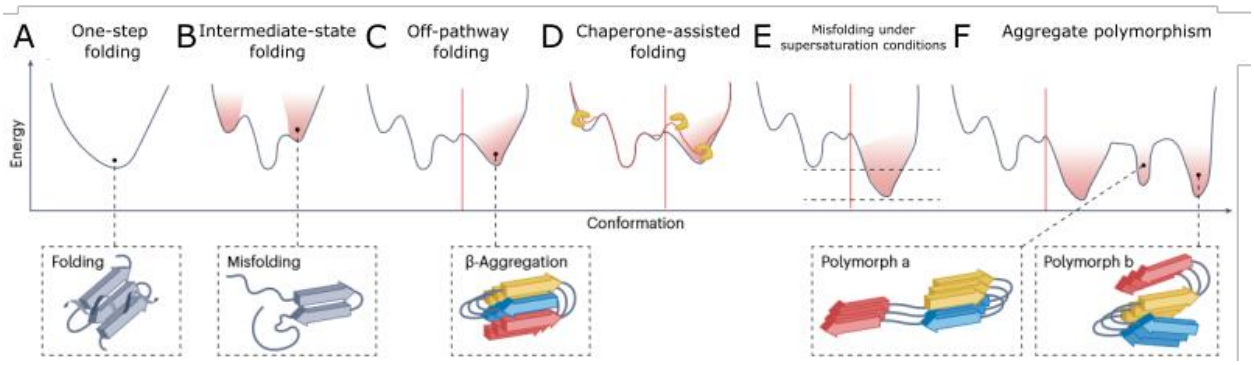


Figure 1.1. The view of protein folding and misfolding through a free energy landscape. (**A**) In its simplest form, it forms a smooth funnel where a multitude of unfolded states with high free energy converge into a single minimum, the native fold. (**B**) The folding of proteins undergoes various steps, with intermediate conformations sometimes encountering small energetic barriers, represented as local minima, that need to be overcome in order to reach the native state. (**C**) As the aggregated state can be as energetically favourable as the native fold, if not more, sometimes proteins become trapped in the minima associated with aggregation. (**D**) Chaperones and other cofactors are often what is preventing the misfolding of proteins *in vivo*, or helping them overcome the local energetic barriers to rescue them from misfolded states. (**E**) In conditions of protein supersaturation, where intermolecular contacts are much more favourable compared to protein-solvent contacts, the aggregated state has a much lower free energy compared to the native state. (**F**) As misfolding is also a process with many intermediate states due to kinetic and enthalpic contributions, many energetic minima usually exist that correspond to various structural polymorphs. This is reflected in the broad structural ensemble of amyloid structures, seen for most proteins. Adapted from[3].

These illustrate well the fact that protein folding involves a large number of intermediate states and conformations, represented by the rugged shape of the landscape with many local free energy minima. These intermediate states can range from local secondary structure rearrangements to partial unfolding and can often be similar to the native fold, as well as functional, thus leading to terms such as metastable states or partially

unfolded intermediates[26,27]. These states then lead progressively to the lowest free energy states, corresponding to either the native fold, dominated by intramolecular interactions, or the aggregated state, dominated by intermolecular interactions[28]. As is the case with native folding, the determinants of aggregation can also be traced to the primary structure level. Termed aggregation-prone regions (APRs), these are often short segments of hydrophobic amino acid residues (6 - 10 residues) that favour the formation of β-sheets[14]. Due to their hydrophobic nature, APRs are most often found in the hydrophobic core of globular proteins, contributing to its stability. APRs are also common in the membrane-spanning segment of transmembrane proteins[29], rendering that part highly aggregation-prone as seen in the example of Aβ[30]. IDPs also contain APRs, albeit to a lesser extent, and have been shown to help facilitate the formation of transient structure, required for the interaction and functionality of certain IDPs[31]. The dynamic structure of IDPs also means that APRs tend to be more exposed than APRs in globular proteins, which can make certain IDPs particularly aggregation prone[32]. In order to counterbalance the misfolding propensity of APRs, amino acid residues on either flanks of these hydrophobic patches have been evolutionary selected to act as a counteracting force by means of unfavourable electrostatic repulsion of charged amino acids, as well as inducing unfavourable structural constraints with proline residues[33].

Aggregation or aggregate formation can thus be considered an alternative folding pathway, a term often used when referring to functional amyloids where amyloidogenesis is required for the protein's activity[34], while misfolding is associated with disease conditions where proteins lose or alter their functionality upon amyloid formation[35]. Proteins achieve the structural transition to the aggregated state either through off-pathway intermediary states upon folding, or by destabilisation of the native state and again the stabilisation of aggregation-prone, off-pathway intermediary states[3]. The destabilisation of the native, globular state can be achieved by many intrinsic or extrinsic mechanisms. Mutations of the protein primary structure are examples of intrinsic factors that can decrease the conformational stability in the native fold and alter the aggregation propensity of proteins[36]. Many studies, mostly conducted *in vitro*, have shown the effect of extrinsic physicochemical factors such as changes in pH, pressure, solvent and co-factors, that promote the transition from a globular to an aggregated state[37]. Also, the presence of existing aggregated forms can both destabilise the existing monomeric, globular protein, as well as promote its aggregation in a process known as "seeding" and best known for as the method of propagation of the prion protein[38]. Furthermore, a process known as "cross-seeding" can also cause the propagation of aggregation-prone proteins by amyloids of another species[39]. All these processes depend greatly on the formation of favourable intermolecular contacts between proteins, meaning the local concentration of the protein greatly affects its aggregation propensity[35]. At a critical concentration, the contacts between proteins are much more energetically favourable than contacts between the protein and the solvent, leading to a phenomenon known as phase separation[40]. The local protein concentration

in a separate phase can rise to mM levels[41], which can in turn lead to another phenomenon, known as supersaturation. Supersaturation occurs when a protein remains in solution at a concentration exceeding its maximum solubility for a prolonged period of time, which further reduces the free energy of the aggregated state and renders aggregation more thermodynamically favourable than the native state[21,42]. This poses a challenge in itself, as most proteins in the cell are expressed at their solubility limit[43], an observation termed "Life on the Edge"[44]. Phase separation and supersaturation have also been shown to promote partial unfolding and conformational changes in protein structure[45]. These findings help explain why aggregation can occur spontaneously under physiological conditions under the condition of supersaturation, when the aggregated state is more energetically favourable than the native state.

In order to prevent aggregation, cells have developed a system of protein homeostasis, which both assists the proteins in the folding process, as well as clears any misfolded or aggregated species by targeted degradation[46]. The role of chaperones is particularly important, both in assisting the proteins in achieving the correct fold upon translation, as well as interacting with already formed aggregates in an attempt to disaggregate and refold them[47]. Aggregated proteins are also degraded by proteasomes and intracellular autophagy[48]. The decline of these mechanisms with ageing is thought to be one of the major factors contributing to an increase in risk of neurodegenerative disease, linked to protein aggregation[49].

1.1.2 Structural characteristics of protein aggregates

As already mentioned, the term aggregate and amyloid are often used interchangeably, likely stemming from the fact that amyloids are by far the most studied and well-characterised aggregate architecture[7]. Amyloids are by no means the only architecture of protein aggregates. Amorphous aggregates are still formed by protein misfolding and exhibit a substantial percentage of β-structure[50], yet are much more difficult to structurally characterise due to their mostly macroscopic differences[7]. Amorphous aggregates also exhibit slightly different kinetics of formation compared to amyloids[51] and are thought to be the favoured aggregate morphology formed by proteins with very similar amino acid sequences[52]. Aggregates can also be formed by proteins under minimal unfolding, exhibiting a native-like structure with certain dynamic elements allowing the establishment of new intermolecular contacts[53]. The amyloid fold, however, is still by far the best characterised aggregate architecture, mostly due to the large number of atomic-level structures solved by various methods and its prevalence in disease conditions. Its signature structural hallmark is the cross-β fold, with its name deriving from the perpendicular orientation of the β-sheets in respect to the fibril orientation[54]. These cross-β folds are stabilised by a number of intermolecular interactions through sidechain-sidechain and sidechain-mainchain H-bonds, buried salt bridges and polar contacts that favour the conformational stabilisation of what Gallardo and colleagues have termed "polar steric zippers"[54]. They differentiate these from "dry

steric zippers", which are formed by sidechain-sidechain contacts between two identical sequences, which are mostly found on the protofibril interface and contribute to the packing of these structures into mature amyloid fibrils[54]. As mentioned before already, these numerous intermolecular contacts confer great stability to the amyloid fold, to the extent that the hypothesis has arisen that postulates amyloids as the first effectors in the prebiotic world[55] and the amyloid fold as the precursor to the development of other, globular proteins folds[56]. Unlike the folding of native proteins, where generally the native fold only refers to a single conformation with the lowest energy, the thermodynamic stability and the many kinetics barriers in the process of amyloid folding can lead to an incredible polymorphism of structures[28]. This is further reflected in the great differences between structures of *in vitro* or *in vivo* fibrils of the same protein. In the case of aS, the two structures of fibrils extracted from post-mortem brain tissue[57,58] are widely different from the ones produced *in vitro* and harbouring disease-associated mutations[59]. Recent work has revealed that despite the prevalent association with the cross-β fold, amyloids can also fold in a cross-α architecture[8]. These motifs seem to be prevalent in cytotoxic amyloids of microbes[60] and in the case of uperin 3.5 can fold with into a cross-α or a cross-β architecture based on the cellular environment and its interactors[61], further highlighting the structural polymorphism and adaptability of the amyloid fold.

The other characteristic of amyloids is that often it is not the full-length protein that gets folded into the characteristic amyloid fold[62]. Especially for longer sequences, the unfavourable conformations that come with the enforced torsion angles can lead to local frustration that prevents these fragments from amyloid formation. Sometimes, these external residues are removed by proteolysis, which has been shown to further stabilise the amyloid fold[63]. More often, they remain and constitute what has been termed the "fuzzy coat" to distinguish it from the "amyloid core"[64]. These residues and their conformational flexibility have both functional and structural roles in many stages of aggregate formation[65]. They provide an interaction surface with its own physicochemical properties that enables the amyloid to facilitate heterogeneous interactions with many binding partners (Fig. 1.2).

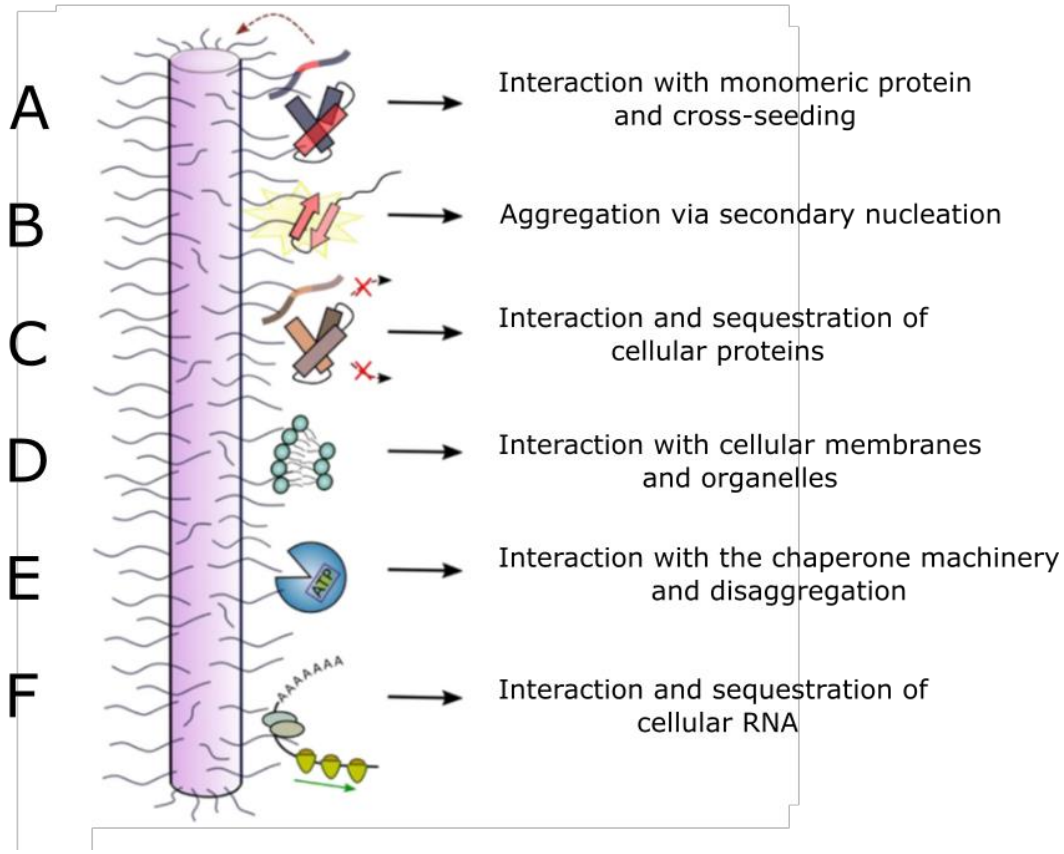


Figure 1.2. The residues outside the amyloid core that constitute the fuzzy coat have many functions ranging from fibril growth to interaction with other molecules. Adapted from [64].

These include interaction with monomeric proteins, either of the same or other species to seed or cross-seed the fibril growth, as has been shown in the case of aS and tau[66]. These interactions affect the growth kinetics themselves, either through interaction with monomeric proteins that promote fibril elongation[67] or through a process termed "secondary nucleation", which will be discussed in more detail below[68]. Interactions with other cellular components such soluble proteins[69], chaperones[70], membranes[71] and nucleic acids[72,73] can also lead to both functional and aberrant mechanisms in the cell. The case of hnRNPDL-2 fibrils is especially relevant and fascinating, as Garcia-Pardo and colleagues show that the RNA-binding domains stay partially folded in the fuzzy coat and still exhibit specificity in RNA-binding[73], implying that not all of the components of the fuzzy coat are actually fuzzy or disordered, but can preserve a level of folding even upon aggregation.

1.1.3 The kinetics of protein aggregation

Macroscopically, the process of aggregation is a very complex process, with sequential conformational changes and species interconversion ranging from a misfolded monomeric protein to a mature amyloid fibril consisting of more than one structurally diverse filaments. It can be divided into three distinct phases, based on the characteristic

plot of time-resolved fluorescence of an aggregation assay (Fig 1.3). During the lag phase, the initial nuclei of unfolded or misfolded proteins are formed in the shape of small, oligomeric species. These then act both as elongation, as well as seeding templates that promote the rapid acceleration of protofibril growth during the exponential phase. The seeded aggregation assay, where pre-formed fibrillar or oligomeric species are added to the reaction mix at the beginning, therefore shows a distinct lack of the lag phase (Fig 1.3). The last is the plateau phase, where the growth stops either due to a lack of available monomeric protein or lack of elongation-compatible ends, and where the protofibrils and filaments assemble into the mature fibrils[3,74].

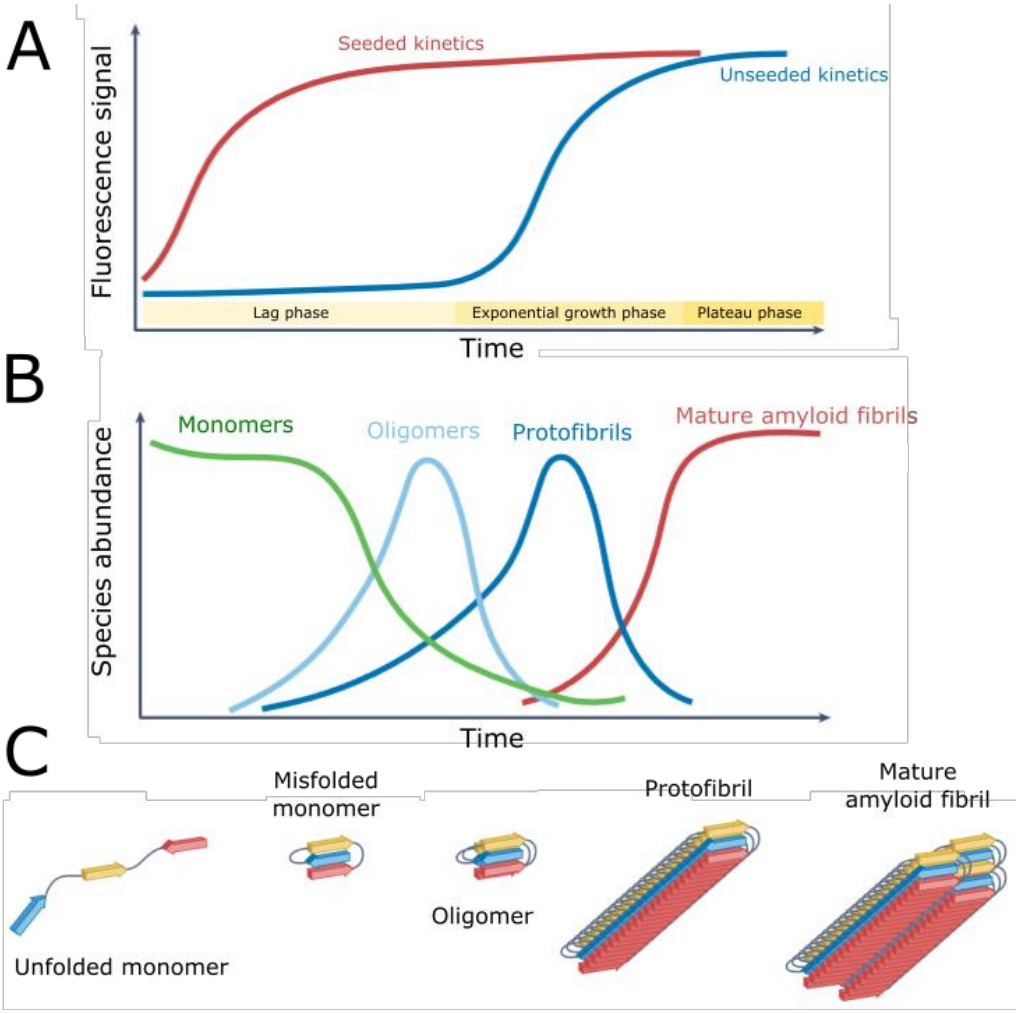


Figure 1.3. The kinetics of protein aggregation with its characteristic phases (**A**). (**B**) Different species are formed and then interconverted during the aggregation process, leading to the formation of the mature, amyloid fibrils. (**C**) A graphical representation of the major species formed during the aggregation process of proteins. Adapted from [3].

As mentioned above, these individual phases refer to the specific parts of the fluorescence-based aggregation assays. These rely on the fluorescent detection of

aggregated species, specifically through the use of amyloid-specific dyes such as thioflavin T[75]. The increase in fluorescence therefore directly correlates with the presence of amyloid species, yet does not necessarily reflect the presence of other, intermediary species. The key to these are the oligomeric species, formed during the lag phase through primary nucleation. They can undergo significant structural changes during the lag phase[76,77] that can also lead to the formation of structures, known as kinetic traps, that prevent further fibrillar assembly[78]. Due to these dynamic structural rearrangements and relative conformation flexibility, the hydrophobic patches in oligomers tend be more exposed, which can lead to their increased cytotoxicity in respect to mature amyloids[79,80]. Oligomers have been shown to interact with and disrupt cellular membranes[71], inhibit cellular processes associated with cognition[80], sequester other cellular components[81] and seed monomer aggregation[82] in ways similar to mature amyloids. Oligomers can then both self-assemble, grow through elongation, or promote the seeding of other nuclei, which leads to the exponential growth of protofibril and fibrils[82]. The main difference between protofibrils and mature amyloid fibrils lies in their structure and formation process. Protofibrils are considered to be a mechanistic precursor of fibrils, and they possess some similarities to mature fibrils, but they also have distinct structural features[83]. Protofibrils tend to be monofilament, contain similar β-elements to mature fibrils, albeit to a lesser extent, with certain internal segments of the amyloid fold undergoing structural rearrangements upon mature fibril formation[83]. Mature fibrils can also be composed of more than one filament and often of filaments with different conformations, adding even more elements to the architectural complexity of amyloids. In their excellent review, Gallardo and colleagues have classified mature fibrils into four different classes, based on their architecture[54]:

- Type 1: Different packing arrangements of the same protofibril, as was observed for amyloidogenic peptides of transthyretin[84] and immunoglobulin LC λ-1[85]
- Type 2: An identical region of both filaments has an identical fold, while the rest have a different one. This was observed in the 'Rod' and 'Twister' polymorphs of α-synuclein fibrils[86]
- Type 3: Combines Types 1 (different packing of protofilaments) and 2 (partial common fold) and has been observed for $A\beta_{42}$ in structures elucidated by cryo-EM[87] and ssNMR[88]
- Type 4: The most complex, with both the protofilament structure and packing interactions varying. Observed in polymorphs of Tau fibrils formed *in vitro* in the presence of heparin[89] and TDP-43 C-terminal domain[90].

The underlying kinetic processes of the formation of these species have been well-characterised with clear mathematical definitions of individual contributions[91]. The key processes involved in the growth of fibrils are both primary and secondary nucleation, fragmentation, and elongation, with monomer dissociation and subsequent oligomer disaggregation counteracting it (Fig 1.4).

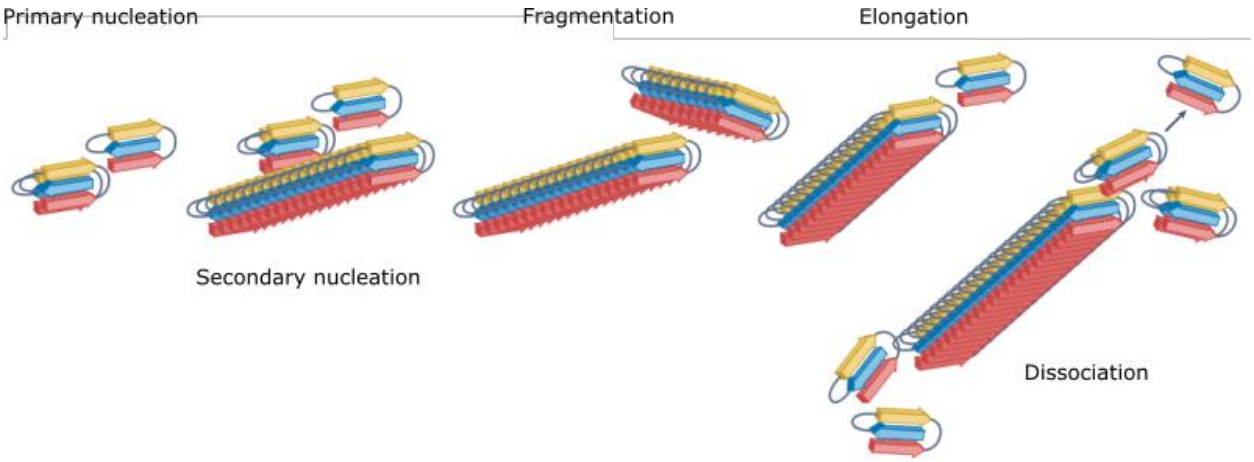


Figure 1.4. The kinetic processes of protein aggregation and their graphical depictions. Adapted from [3].

Primary nucleation represents the misfolding and formation of aggregation-prone nuclei. Due to the many conformational changes associated with misfolding, primary nucleation is entropically unfavourable, which reflects in the slow kinetics[91]. As such, primary nucleation is the process most affected by physicochemical factors, as well as interaction partners[3,91]. Secondary nucleation represents the misfolding on the lateral sides of already aggregated species[68]. After the initial formation of seeding-competent nuclei with primary nucleation, secondary nucleation increases in prominence with the growing abundance and length of aggregates, thus significantly accelerating the process[68,91]. The growth of aggregates is further promoted by the process of elongation or the incorporation of monomers at growth-competent fibril ends. Elongation is closely associated with fragmentation, which can create more growth-competent end, especially since elongation is usually an asymmetric process with one end of the fibril exhibiting increased rates of monomer incorporation compared to the other[92]. Determining the microscopic constants of the aggregation subprocesses is a laborious task and highly dependent on the protein and solution conditions[93], however can be critical in the assessment of the impact of disease-associated mutations, as well as the effect of potential drugs that act as modulators of protein aggregation[94].

1.1.4 Aggregation in physiology and functional amyloids

While disease, especially neurodegeneration, is the most prevalent association with aggregation, protein aggregates are prominently featured in a variety of functional roles in all trees of life[4], which was predominantly what led to the hypothesis of the "amyloid world", the prebiotic world where thermodynamically stable amyloid assemblies of the first peptides are perceived as the precursor species of modern proteins[55]. Short segments of amyloid-forming peptides have indeed been shown to harbour the ability to perform enzymatic reactions[95]. These so-called catalytic amyloids are usually shorter peptide segments that can assemble into stable scaffolds that catalyse a variety of chemical reaction, either in the presence of a cofactor or without[34]. While the rational, guided design of these peptides can produce catalytic amyloids with reaction kinetics

close to those of globular enzymes, several naturally occurring amyloids have been characterised as catalytic and capable of performing basic chemical reactions, mostly as hydrolases with broad specificity[96–98]. This has opened up new perspectives of potential amyloid activity in the cell, particularly cytotoxicity by (un)specific degradation of cellular components by catalytically active amyloid assemblies, although these have yet to be demonstrated *in vivo*.

Amyloid structures with well-characterised cellular functions have been predominantly explored in yeast and bacteria. One of the best studied cases is an example of curli fibrils that form the cell wall of *Escherichia coli* and other bacteria[99]. Formed by two structural and five regulatory proteins, the assembly of curli is a highly regulated process, essential for biofilm formation, cell attachment and infection[100]. The assembly of amyloids is also a key functional aspect of yeast and fungal reproduction. The soluble HET-s protein converts into the highly aggregation-prone HET-S, inducing cell death by forming pores in the membrane of genetically incompatible *Podospoda anserina* fungi[101]. Whi3, instead, regulates the cell cycle by binding yeast cyclin mRNA and preventing the onset of the S phase[102]. This is the key step for the sexual reproduction of yeast cells, however if the compatible cells do not unite in 4 h, Whi3 starts aggregating and releases the cyclin mRNA to induce the onset of the S-phase and vegetative reproduction[102]. Whi3 aggregates then template further soluble Whi3 aggregation, rendering the cell permanently sterile. They are however retained in the mother cell, meaning any daughter cell are again perfectly capable of sexual reproduction[103]. The Whi3 system of amyloid-mediated epigenetic regulation of gene expression is viewed by some researchers as a primitive memory system and the term "mnemons" has been coined for such amyloids[104]. A similar case is the conserved yeast Vts1 protein (Smaug in *D. melanogaster*), which unlike Whi3 is propagated amongst daughter cells, and again regulates the mitotic/meiotic cycle based on available nutrients in the immediate environment[105]. Orb2, the *Drosophila* homologue of mammalian CPEB, is involved in the process of long-term potentiation in the brain by regulating the local translation of certain mRNA[106]. Upon self-assembly, the RNA interactome of Orb2 changes, meaning the aggregated protein still retains the ability to bind RNA and thus regulate local protein translation in the synapse[107].

1.1.5 Aggregation in disease

The misfolding and aggregation of proteins however still mostly results in major disruption of cellular processes and is thus primarily associated with disease conditions. This was fuelled by the discovery of prion protein aggregates as primary causes for the highly infectious prion diseases such as bovine spongiform encephalopathy in cattle and Creutzfeldt-Jakob disease in humans[108]. Prion diseases and protein misfolding disease share common background and motifs and are characterised by the accumulation of insoluble protein deposits in various tissues[14,109]. A significant number of these deposits are found in the brain and are closely associated with neurodegenerative diseases[110].

There are many excellent reviews on the topic[22,35,111] and I will only briefly touch on some of these cases for context.

Aβ has been shown to be primarily linked to plaques found in the patients with AD, as has been the protein tau, with aggregates affecting the cell's microtubule integrity, synaptic functions etc[112]. Amyotrophic lateral sclerosis (ALS) is a complex disease characterised by aberrant phase transitions of various proteins and often with an overlapping clinical spectrum with frontotemporal lobar degeneration (FTLD)[113]. ALS-FTLD represent a very complex molecular background with many networks involved, however the principal proteins that are also found aggregated in both disease conditions are TDP-43 and FUS[113]. Both of these proteins are canonical RBPs, with a characteristic modular structure of two nucleic acid-binding domains and a low-complexity, prion-like domain responsible for phase transitions[114]. The aggregation of these two abundant proteins can have dramatic alterations in a variety of cellular processes, ranging from RNA processing and transport[114,115]. More recently, TMEM106b was identified as the fibril component of amyloid plaques found in patients with Alzheimer's disease as well as FTLD and multiple systems atrophy (MSA)[116,117]. The fibrils of TMEM106b, extracted from patients' brain, have been shown to be very polymorphic across the disease spectra, as has been the case of TDP-43 and Aβ, indicating the existence of potential disease-specific amyloid folds[116]. There have also been several other proteins identified in the aggregates across the ALS-FTLD spectrum, from annexin A11[118], SOD1[119] to dipeptide repeats from the *c9orf72* gene[120].

Aggregation is not linked only to neurodegeneration. In certain cases, amyloid deposits in a variety of tissues such as heart, liver and kidneys can lead to a class of diseases known as systemic amyloidosis[121]. In systemic amyloidosis, extracellular amyloid assemblies form because of aggregation of proteins such as β2-microglobullin, transthyretin, the light chain of immunoglobulin etc[122]. Mutations in p53 affect its aggregation and have been linked to cancer[123].

All these cases exhibit a significant genetic background, as is common with aggregation-prone proteins, with mutations significantly affecting the misfolding and aggregation in disease[20,124]. These examples show that protein aggregation is directly involved in several disease conditions in a variety of cellular contexts.

1.1.6 The role of RNA in protein aggregation

One of these cellular contexts that has only recently started emerging is the role RNA and play in the formation of protein aggregates. RNA has long been associated with aggregation both directly and indirectly. A high-throughput study of the yeast proteome for potential prion proteins has identified several RBPs as candidates with a high aggregation propensity[125]. The interplay of RNA-binding and prion-like domains has been shown to play an important role in protein phase transitions[126] and amyloids have

also been shown to have an increased propensity for interaction with their own transcripts[127].

As mentioned in the sections above, many aggregation-prone protein are RBPs that harbour domains that specifically interact with RNA. Other aggregation-prone proteins have been shown to interact with RNA, albeit in a less specific way. RNA has been shown to modulate the aggregation of the human prion protein by directly interaction[128], as well as promote the aggregation of tau into an-AD-like topology[129]. These tau aggregates have also been shown to sequester RNA upon aggregation[130] and show a certain extent of preference and specificity in terms of RNA binding[131].

These interactions between RNA and proteins are one of the cornerstones of cellular biology, however not all is clear as to the evolutionary process that led to the emergence of specific RNA-binding domains. Reports of peptide-RNA coaggregation into specific assemblies with distinct molecular properties, based on hydrophobic[132] and electrostatic interactions[133] has added another dimensions to the amyloid world hypothesis[134]. RNA-amino acid interactions in these assemblies have dramatically increased their stability, suggesting that these contributed to their coevolution in the pre-biotic world[135,136].

RNA has been shown to directly induce the aggregation of several cellular proteins in the nucleolus[137]. These long non-coding RNA termed rIGSRNA induce the assembly of the VHP protein upon heat shock into amyloid structures termed A-bodies, that are disassembled by the cellular chaperone machinery after shock termination[138]. The induction of aggregation with RNA is critical in the process, however unlike many phase transitions, the formation of A-bodies was not shown to undergo liquid-liquid phase separation[139]. RNA-amyloid interactions also play a critical role in the formation and maintenance of the Balbiani body[140].

Importantly, RNA has also been shown to stabilise the aggregation-prone protein and abrogate its aggregation almost completely as shown in the case of TDP-43[141]. This important observation has directly led to the development of synthetic RNA aptamers with a wide spectrum of applications[142] and most importantly highlighted that while RNA can play a significant role in protein aggregation, it can also significantly contribute to its modulation and potential therapies.

1.2 Alpha synuclein

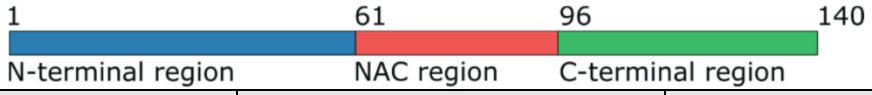
Alpha synuclein (aS) is a 140 amino acid long intrinsically disordered protein, found abundantly expressed in mammalian cells[143]. While its primary has yet to be clearly established, it has been shown to be involved in a variety of cellular processes associated with lipid-binding[144], such as synaptic vesicle trafficking, exocytosis[145,146], mitochondria homeostasis[147] and DNA repair[148]. However, it is likely best known as the

primary component of Lewy bodies[149,150] its involvement in a spectrum of neurological disease known collectively as synucleinopathies[151].

1.2.1 The biochemical characteristics of aS

aS is a distinctly modular protein with three distinct domains or segments which differ a lot according to their primary structure and composition. The N-terminal is predominantly made of polar, amphipathic and basic residues and represents the binding surface for interaction with several cellular components (see Table 1)[144,152]. The NAC region is nominally part of the N-terminal region and represents part of the protein with the highest hydrophobicity and forms the central core of the amyloid assembly upon aggregation[59,153]. The C-terminal region is instead mostly negatively charged and has been shown to remain mostly unstructured in most cellular contexts[154,155].

Table 1. An overview of aS structure and the functions associated with different regions. Adapted from [64].



aS residues	Function	Reference
1	N-terminal acetylation	Bartels *et al.*, 2011[156]
1 – 14	Membrane insertion	Cholak *et al.*, 2020[157]
1 – 25	Membrane binding	Fusco *et al.*, 2014[144]
N-terminus, especially Tyr39	Hsc70 chaperone binding	Burmann *et al.*, 2020[158]
36 - 42 and 45 - 57	Involved in liposome clusterin	Doherty *et al.*, 2020[152]
36 - 42 and 45 - 57	Involved in inter- and intramolecular contacts	Doherty *et al.*, 2020[152]
37 - 54	Forms a transient β-sheet in the partially folded intermediate state, critical for aggregation propensity	Mirecka *et al.*, 2014[159]
His50	Likely critical for enzymatic activity	Horvath *et al.*, 2023[98]
C-terminus (91-140)	Protects against aggregation by intramolecular contacts, shielding the NAC region	Hoyer *et al.*, 2004[154] Stephens *et al.*, 2019[153]
Asn103	Cleavage site for asparagine endopeptidase	Zhang *et al.*, 2017[160]
C-terminus (110-140)	Interaction with SERF1	Falsone *et al.*, 2012[161] Meyer *et al.*, 2020[162]
C-terminus (107-137)	Ca^{2+}-binding and lipid binding	Lautenschläger *et al.*, 2018[163]
C-terminus (125-129)	Dopamine-binding modulates aggregation	Herrera *et al.*, 2008[164]
C-terminus (119-140)	Critical for oligomerisation	Farzadfard *et al.*, 2022[155]

As mentioned, aS is an IDP in solution, however has been shown to form a transient α-helix in the N-terminus and part of the NAC region upon the binding of lipids[165]. aS has also undergoes several post-translational modifications. It has been shown to be N-terminally acetylated *in vivo*, which affects the α-helix folding propensity and the affected functionalities[166]. The phosphorylation of Ser129 has been shown to affect the aggregation propensity and amyloid structure formation and has been used as a biomarker of PD[167]. aS can also be extensively glycated by methyl-glyoxal, ribose and other reactive sugars, affecting its aggregation kinetics[168]. Most importantly, aS has been shown to be truncated both N-terminally and C-terminally[169]. These truncations can occur *in vivo* as results of protease activity[160] and are associated with a more aggressive form of PD[170]. The removal of the negative charge of the C-terminus also drastically changes the aggregation pathways, both accelerating the kinetics and promoting direct fibrillation[155,171].

1.2.2 The aggregation of aS

Due to its involvement in several neurodegenerative disorders, the aggregation of aS is well-studied both from the kinetic and structural, as well as functional perspectives. aS aggregates have been shown to grow via monomer addition and an intermediate step that involves the interconversion of oligomeric species[172]. The initial, slow step of primary nucleation is very much dependent on solution conditions, such as pH, salt concentration etc[93]. The key to the process is perturbing the intramolecular interactions between the opposite charges on the N- and C-termini, thus exposing the NAC region and promoting the folding via the transient β-sheet formation in the partially folded form[153,173,174]. Charged polymers are have consequently been also show to greatly affect the aggregation kinetics of alpha synuclein[175,176]. The addition of lipids also induces aggregation via primary nucleation[177]. C-terminal truncations also diminish the protective charge shielding of the NAC region and have been shown to significantly speed up the aggregation kinetics[178], however it has been determined the more significant contribution to the process is the secondary nucleation[171]. In general, while primary nucleation has been shown to the primary process contributing to aS aggregation at physiological pH values[93], secondary nucleation dominates at pH values lower than 6[179]. As the isolelectric point of aS is 4.76, the lowering of the pH diminishes the overall charge of the protein and weakens the protective intramolecular interaction[179].

As already mentioned before, the aggregation process has also been shown to involve several intermediates. Cremades and colleagues have shown the early stages of aggregation involve the fast formation of oligomeric species[77], which then interconvert into more rigid, proteinase K-resistant oligomeric species[78]. This process is characterized by the transition from a predominantly parallel to an antiparallel β-sheet orientation[180]. The C-terminus play a critical role in the process of self-assembly of oligomeric species[155], and the truncated species tend to form fibrils directly without the

intermediate conformations[171]. Mature fibrils are characterized by various polymorphs, with little resemblance between the structures of fibrils grown *in vitro* and the two folds from *post mortem* tissue. The N- and especially the C-terminal form the fuzzy coat around the amyloid core, formed by the NAC region, as determined by ssNMR[181]. The two structures of *in vivo* amyloid fibrils derive from patients with MSA[58] and LBD[57]. Both show unknown electronic densities in the structure, solved by cryo-EM, indicating the fibrils likely contain other heterogenous components. Lipids are a likely candidate for them, since aS interacts with fibrils in its monomeric form and aS fibril, grown *in vitro* in the presence of lipids, also show lipid-derived electronic densities in the amyloid core and around[182]. Apart from lipids, heparin has also been shown to affect the fibril structure when coaggregated with aS *in vitro*[183]. The recently solved cryo-electron tomography image of *in vivo* aS aggregates by Trinkaus and colleagues has also shown that fibrils sequester cellular components such as membrane organelles and proteasomes[184]. Furthermore, they showed that different seed polymorphs seed different fibrillar architectures *in vivo*, and that aS fibrils likewise exhibit polar growth also in cellular context[184]. They are also targeted by the cellular chaperone machinery, primarily via Hsp70[185]. Hsp70 is recruited to aS fibrils by DnaJB1 co-chaperone and together with the third factor, Apg2, disaggregates the fibril by removing monomeric units from the fibril ends[186]. Another chaperone, Hsc70, has been shown to interact with the N-terminus of aS, with the binding abolished upon Tyr39 phosphorylation[158]. The disaggregation of aS fibrils has also been shown to be detrimental in certain cases, as the species removed can also have cytotoxic effects themselves[187].

1.2.3 aS and nucleic acids

aS has not been characterized as a canonical nucleic acid-binding protein as it lacks a canonical nucleic acid-binding motif, such as an RRM motif common in canonical RBPs. However, aS has nonetheless been associated with nucleic acids in a number of different studies, both directly and indirectly. Cherny and colleagues first reported that double-stranded DNA interacts with monomeric aS and accelerates its aggregation[188]. While the morphology of the fibrils did not show any significant changes, electron microscopy images revealed that plasmid DNA bounds directly to the fibrils of both wild-type aS as well as disease-associated mutants[188]. A further study then reported DNA induces partial folding in aS, with both the aggregation-prone partial folding observed, as well as an increase in alpha helicicity[189]. These different DNA species correspondingly accelerated (via the partially-folded intermediate) or decelerated (via the α-helix formation) the aggregation kinetics[189]. Similarly, aS was shown to interact with DNA i-motifs and induce certain conformational changes[190], with the C-terminal truncations affecting DNA conformations in an opposite manner compared to full-length aS[191]. DNA aptamers have also been developed against aS by using SELEX, however they have been shown to also bind other amyloids such as Aβ[192]. Other DNA aptamers have been shown to significantly slow done aS aggregation kinetics, however no structural or binding mechanisms have been proposed so far[193].

Recently, Schaser and colleagues demonstrated that aS both binds dsDNA and facilitates non-homologous end-joining reaction in cells, thus implicating aS in DNA repair machinery[148]. Similarly, Hallacli and colleagues have reported the association of aS with processing bodies and affecting mRNA metabolism via direct interaction with the decapping mechanisms[194]. No direct interaction with RNA has been demonstrated so far, with the study of polyanionic compounds published by Cohlberg and colleagues showing no significant difference in aggregation kinetics in the presence of both total cell as well as isolated tRNA[176].

We can conclude that while there is plenty of evidence linking aS to nucleic acids, these are mostly limited to DNA. The effect of RNA has yet to be properly investigated, however with the ability of DNA to modulate aS aggregation and the general ability of RNA to heavily affect protein phase transitions, one can make an argument that these studies harbour plenty of potential.

2. Aims

The aim of this thesis was to evaluate the potential of RNA as a modulator of aS aggregation. Given the past experience and results in the study and design of RNA molecules that can influence the aggregation of RBPs such as TDP-43[141,142], and the computational study by Zanzoni and colleagues indicating aS might interact with its mRNA[127], the goal was to determine the feasibility of a similar approach on aS. The development of DNA aptamers by Tsukakoshi and colleagues showed that aS can be successfully targeted by nucleic acids[192]. However, since aS has not been characterised or classified as an RNA-binding protein, a strategy of first determining the general role RNA might have on its aggregation was prioritised, with the eventual progress determining the feasibility of a more specific approach.

With this purpose in mind, the concrete aims of the project were:

- Perform and optimise an aggregation assay protocol of aS co-incubation with RNA to determine the effect of RNA on the kinetics of the process, e.g. whether RNA increases, decreases or does not affect the rate of aggregation, time-of-onset (lag time) etc.
- Evaluate the structural characteristics of aS aggregates in our system and determine the influence RNA might have on the morphology of aS aggregated species. This was especially important given the general view that aS oligomers might be the actual toxic species that contribute to the pathology. Determining whether RNA can influence conformation of aS to the extent of altering the characteristics of its aggregates would be of great interest in this perspective.
- Determine and characterise the interaction between aS and RNA and, if possible, quantify the binding parameters. This would provide critical information on the actual feasibility of targeting aS and ensure the affinity and specificity of these aptamers within a variety of applications.

The success in these aims would lay the foundation for the design and identification of specific RNA aptamers that could be used to target and modulate the aggregation of aS. Furthermore, the use for these specific RNA aptamers can range from applications in imaging via super-resolution microscopy, improved diagnostics via specific recognition of aggregated aS in patient samples and eventual tests as a potential clinical therapy.

3. Results

3.1 RNA accelerates the aggregation of aS

3.1.1 The N-terminal of aS is predicted to interact with RNA

Alpha synuclein is not considered to be a classic or a canonical RBP. As already mentioned in the introduction, it does not harbour any known nucleic acid binding motifs or domain and its binding preferences are currently unknown. A software used to predict RNA-protein interaction propensities, such as the well-established catRAPID[195], gives the overall binding of the protein as very low, especially compared to canonical RBPs such as TDP-43 (Fig. 3.1). However, the individual domains of aS have by themselves very different propensities towards binding RNA, mainly due to the electrostatic charge distribution. The C-terminal domain has the lower predicted binding propensity, likely due to the negative charge at physiological conditions, followed by the hydrophobic NAC domain[196] (Fig. 3.1).

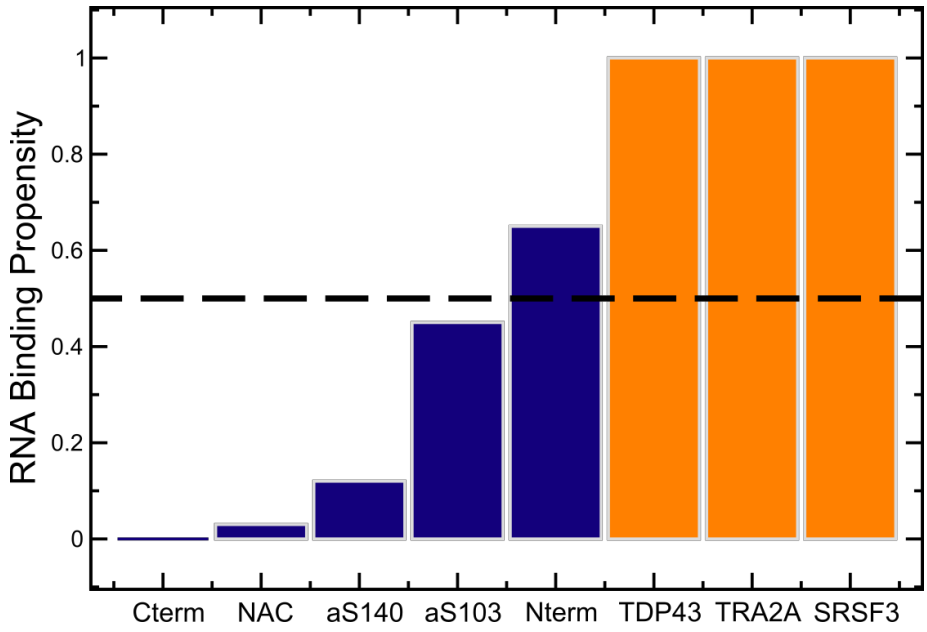


Figure 3.1. RNA-binding predictions of different regions of aS140 and aS103 using the catRAPID tool[197]. The regions as labelled in the figure correspond to aS residues 1-60 ('Nterm'), 61-95 ('NAC') and 96-140 ('Cterm'). Compared with canonical RBP a threshold of 0.5 was chosen to indicate non-negligible interaction propensities. Figure used from [196] with the authors' permission.

The N-terminus on the other hand has a much higher overall binding propensity towards binding the RNA, again presumably due to the positive overall charge that also drives the association with the negatively charged lipid membranes in the cell[144]. This led us to believe that any effect RNA might have on the aggregation of aS might be amplified in the case of a C-terminal truncation. These species can be produced in disease by various endopeptidases and have been shown to exacerbate pathological symptoms[169,170].

We chose the truncated form cleaved by asparagine endopeptidase behind Asn103[160] due to the good overall characterisation of its aggregation characteristics[171,198] and termed it aS103.

3.1.2 Optimising the aggregation protocol for aS

The first step towards characterising the exact effect of RNA on the aggregation of aS was to develop a robust *in vitro* model that enables us to effectively monitor the aggregation kinetics of the two protein constructs, the wild-type aS140 and truncated aS103. This entailed the determination of the temperature and length of incubation, the level of agitation and modality of fluorescence measurements (see also Materials and Methods). aS has been shown to take days to initiate the aggregation process, even when incubated to 37 °C [199]. Its aggregation can be sped up significantly with agitation and even further so using teflon or inert glass beads added to the wells[93]. Given that RNA is a molecule highly sensitive to degradation and hydrolysis, especially at prolonged incubation times at higher temperatures, it was desirable to reduce the aggregation time as much as possible while still retaining the reproducibility between the individual replicates to maintain RNA integrity during the experiment. Adding a glass borosilicate bead to each well does indeed dramatically speed up the aggregation as well as improve the reproducibility of the results for both protein constructs (Fig. 3.2 A,B, Table 2). A detailed description of the analysis of time resolved fluorescence aggregation assays can be found in Materials and Methods.

The aggregation of aS140 under constant agitation at 37 °C only started after on average 60 h, with the results showing great inconsistency both within and between individual experiments (σ = 146 % of the mean, p = 0.033). This improved consistently with the addition of the bead (σ = 27 % of the mean, p = 0.299). aS103 on the other hand aggregated much faster, as already indicated in previous studies[171,198] and adding the bead significantly decreased the aggregation delay time further ($p < 1x10^{-5}$). The calculated aggregation rates did not significantly differ both within replicates ($p >> 0.05$) and between experimental conditions (p = 0.32), however we chose to continue with the experiments using the borosilicate beads for aS103 as well. This was due to comparability with aS140, especially considering the consistent effect of the conditions on the RNA. Adding the bead also improved the variance between the aggregation rates of experimental replicates (σ = 20 % of the mean compared to 50 % of the mean without the bead). No significant difference was also observed between measurements of fluorescence at the top and bottom of the well (Fig 3.2 C,D). In the end, using bottom measurement was preferred due to higher sensitivity as well as the potential of small evaporation droplets forming on the cover of the well and interfering with the measurements at longer incubation times (data not shown).

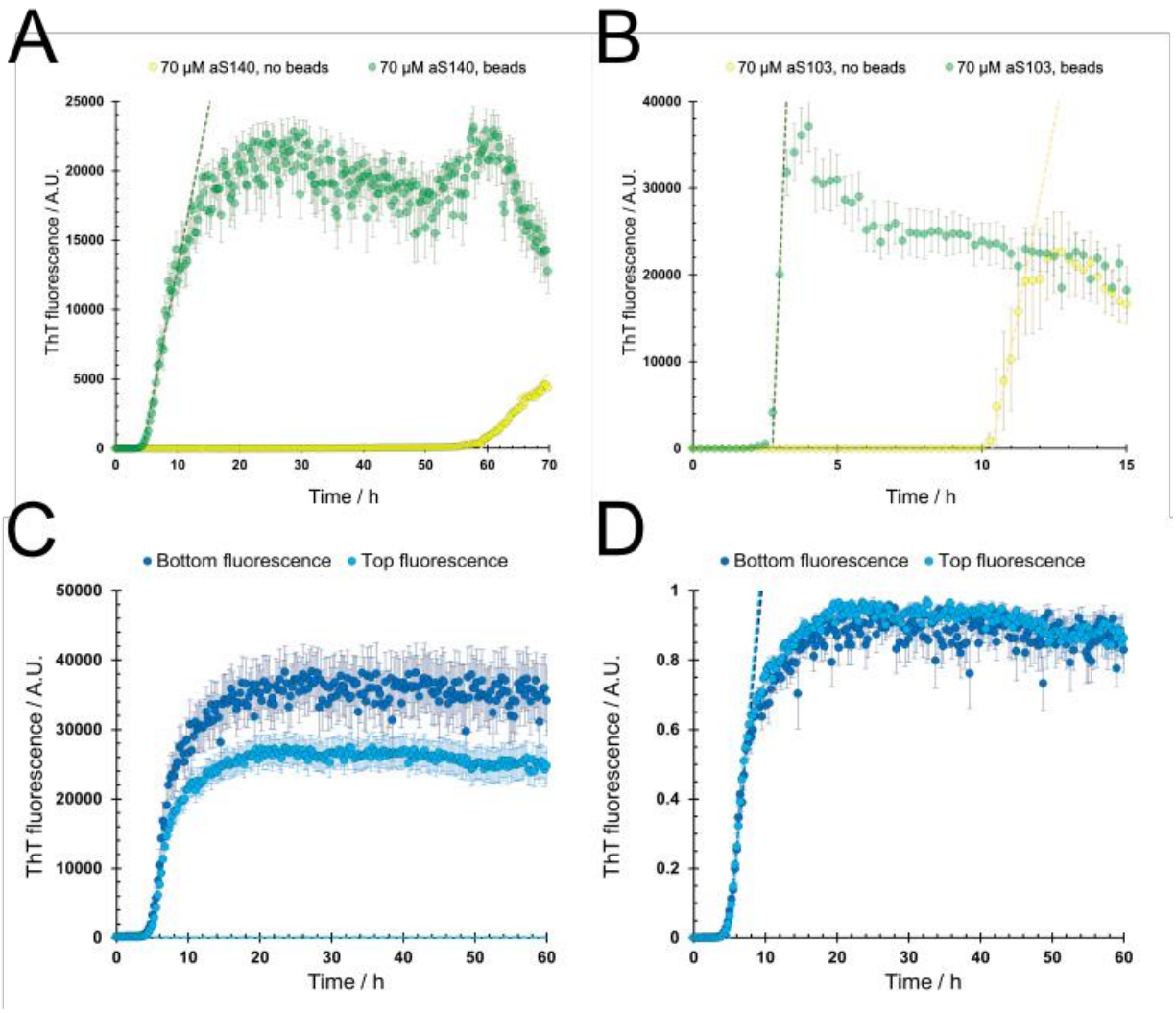


Figure 3.2. Optimisation aggregation assays performed to determine the optimal parameters. **(A)** aS140 starts aggregating after around 60 h with only agitation, while adding a glass bead in the same conditions means the fluorescence reaches a plateau already after less than 30 h (mean of 6 replicates with standard deviations shown). Adding the glass bead also significantly changes the rate of aggregation (Table 2). **(B)** Similarly, glass beads speed up the aggregation of aS103 and improve the reproducibility of the kinetics traces within replicates (Table 2). **(C)** Apart from reaching a higher maximum value of fluorescence, there is no significant difference between the measurements of fluorescence from the top or the bottom of the wells. **(D)** Normalised fluorescence traces with the fit of the exponential part of the sigmoid further shows the repeatability (p-value 0.18, average variance < 10 %, two-way paired Student t-test).

Table 2. Statistical analysis of the optimisation of aS aggregation, shown for both constructs. Non-normalised rates are shown as seen for the data in figure 3.2. Number of experiments refers to individual experiments conducted separately (technical replicates), while number of replicates is the number of replicates within an experiment. The p-value within the experiment was obtained with a one-factor ANOVA test for aS140 experiments with beads (N. of experiments > 2) and with a two-way paired Student t-test for the rest. The p-values comparing the experiments with and with a bead added were obtained using a one-factor ANOVA test.

aS103

		N^{exp}	$N^{replicates}$	*Mean*	*Std*	*p within experiment*	*p between experiments*
Aggregation rate	bead	2	6	88915.91	17652.23	0.52381	0.318423
	no bead	2	6	70233.27	35410.07	0.235978	0.318423
Aggregation delay time	bead	2	6	2.763941	0.089223	0.051798	$1.86x10^{-6}$
	no bead	2	6	11.19832	0.815115	0.848225	$1.86x10^{-6}$

aS140

		N^{exp}	$N^{replicates}$	*Mean*	*Std*	*p within experiment*	*p between experiments*
Aggregation rate	bead	3	6	5287.376	1410.092	0.299995	$2x10^{-5}$
	no bead	2	6	1763.962	2517.055	0.033656	$2x10^{-5}$
Aggregation delay time	bead	3	6	4.318563	0.677011	0.603143	0.004247
	no bead	2	6	31.2903	30.36945	0.158986	0.004247

3.1.3 RNA accelerates the aggregation of both aS103 and aS140

Using these optimised measurement parameters, aggregation assays in the presence of total yeast RNA were performed to determine the effect on aggregation kinetics of aS140 and aS103. Due to the much faster kinetics of aS103 aggregation, we had to increase the sampling to one read every 10 min instead of 15 min as for aS140 measurements. The proteins were incubated in the presence of an increasing concentration of RNA up to 500 ng/µL. Unlike in the optimisation runs, Proteostat™ was used as a reporter dye instead of ThT due to its much lower propensity to bind nucleic acids and thus lower background as shown in previous studies[141].

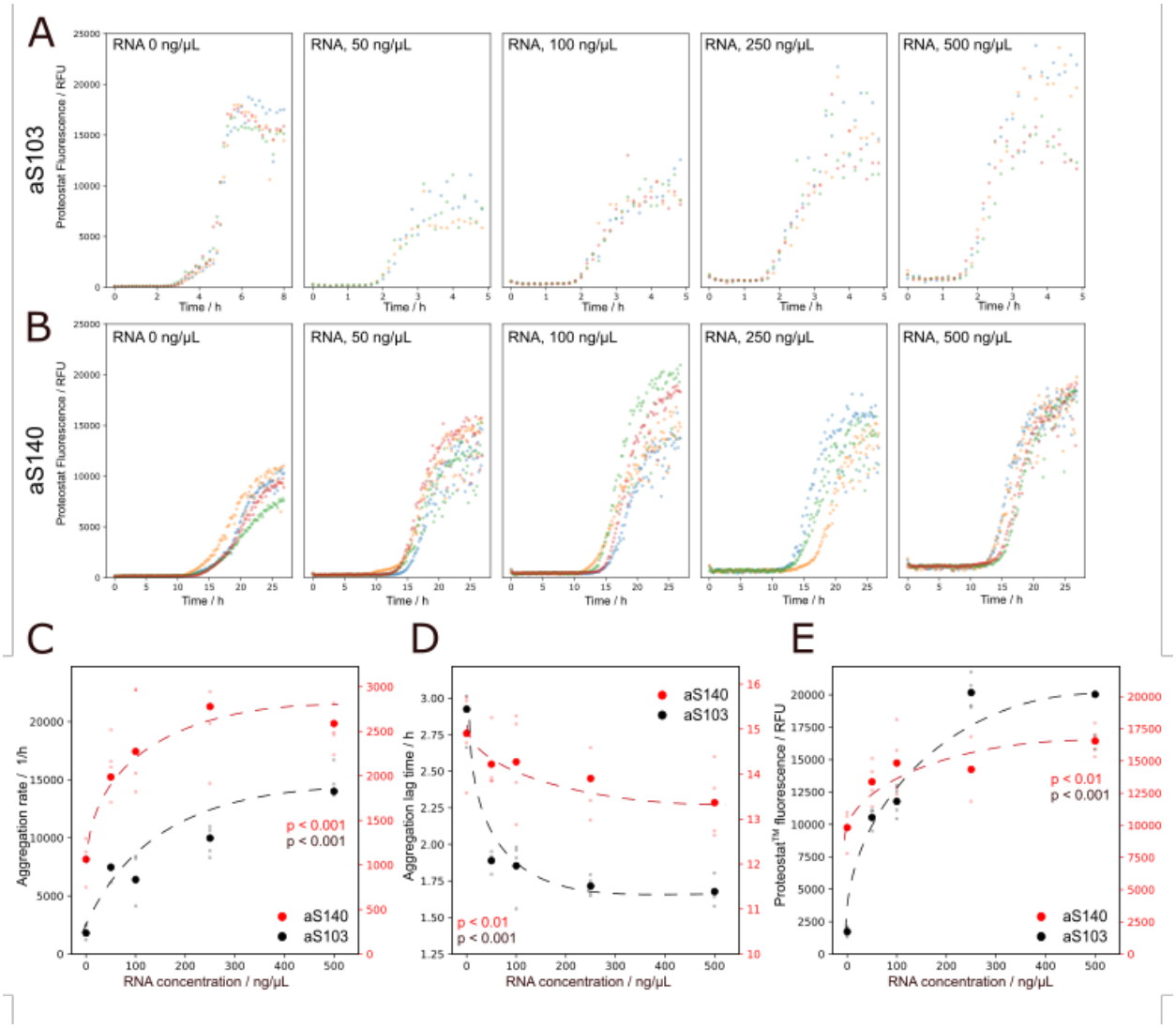


Figure 3.3. Aggregation of aS140 and aS103 in the presence of RNA. Data points in **(C-E)** corresponding to each protein are plotted against their own respective y-axis (aS103 to the left and aS140 to the right axis). **(A)** aS103 shows a much faster onset of the exponential phase of growth 2 h after the start of the assay which progressively shortens with increasing RNA concentration. (**B**) An identical trend is observed for aS140. The protein without RNA has a higher fluorescence value at plateau compared to aS103, while values at higher RNA concentrations are comparable. **(C)** Calculated aggregation rates from time-resolved Proteostat™ fluorescence. Data shown are calculated from parameters derived from the fitted curve to 6 technical replicates of 3 separate experiments and statistical significance was calculated for rates in the presence of RNA as compared to the protein-only samples (see Materials and Methods). The presence of RNA significantly increases the rate of aggregation for both aS103 (Fisher exact test, $p < 0.001$) and aS140 (Fisher exact test, $p < 0.001$), with the aS140 rate reaching a plateau with concentrations higher than 250 ng/µL. **(D)** aS103 has a significantly shorter aggregation lag time than aS140. The addition of RNA further decreases the lag phase of both proteins (aS103, $p < 0.001$ and aS140, $p < 0.01$, Fisher exact test), however increasing concentrations do not seem to drastically affect aS103 further, while the opposite is true for aS140. **(E)** Increasing concentration of RNA also elevates the plateau of fluorescence for both proteins, especially significantly for aS103 ($p < 0.001$ compared to $p < 0.01$ for aS140, Fisher exact test).

As already established, aS103 (Fig. 3.3 A) aggregates much faster both in terms of the onset of the elongation phase as well as the aggregation rate itself compared to aS140 (Fig. 3.3 B). This is true in all the conditions, irrespective of the concentration of RNA. Analysing the experimental data shows that RNA significantly increases the aggregation rate for both aS140 as well as aS103 ($p < 0.001$, Fisher exact test) when taking into account 3 separate experiments with 6 replicates each. The calculated aggregation rate of aS103 is approximately 2-fold higher than aS140 in the absence of RNA ($p < 0.0001$, Fisher exact test). An initial look at the raw data might counter that as fluorescent signal for the elongation stage in the absence of RNA looks especially steep (Fig. 3.3 A). However, a closer look identified a distinct two-step aggregation mechanism with a first, slower step occurring between roughly 2,5 to 5 h. This is consistent with the data from the work of Van der Wateren and colleagues[171] (see Fig. 6 B and E), who identified the aggregation of aS103 as a multi-step process characterised by the formation of earlier proto-fibrils that later assemble into the mature fibril in a faster secondary step. As the earlier step is distinctively slower, we decided to take into account its specific kinetics when calculating and comparing the aggregation rates.

Further addition of RNA significantly increases the aggregation rate for both proteins (Fig. 3.3 C). The rate for aS140 increases up to approximately 5-fold at 250 ng/µL and then reaches a plateau ($p > 0.3$ comparing the rates at 250 ng/µL and 500 ng/µL, one-factor ANOVA). The rate for aS103 increases more than 10-fold at 500 ng/µL and does not seem to stabilise ($p < 0.0001$ comparing 250 ng/µL and 500 ng/µL, one-factor ANOVA). Curiously, additions of RNA also eliminate the rate-limiting first step of aggregation for aS103 (Fig. 3.3 A), indicating a change in the mechanism of aggregation to a faster, single-step process.

RNA also decreases the aggregation lag time of both protein constructs, however barely significantly for aS140 ($p < 0.01$, Fisher exact test) until higher concentrations (Fig. 3.3 D). In contrast, the lag time for aS103 decreases dramatically after the addition of RNA ($p <<< 0.0001$, comparing 0 ng/µL and 50 ng/µL, one-factor ANOVA), with increasing concentrations decreasing it to a lesser extent ($p \approx 0.01$ for 50, 100, 250 and 500 ng/µL, one-factor ANOVA). This fact further suggests a role of RNA in the early aggregation of aS103 by eliminating the rate-limiting first step, with higher concentrations affecting mostly the elongation rate and not the initiation.

The last parameter analysed was the maximum value of fluorescence upon reaching the plateau at the end of the elongation phase (Fig. 3.3 E). Here it is important to mention that while Proteostat™ does interact with nucleic acids, leading to a higher background fluorescence in the presence of RNA, the data analysed have the background value subtracted from control wells containing only RNA and the dye. The maximum fluorescence value increases slightly for aS140 ($p < 0.01$, Fisher exact test) until 100 ng/µL and then plateaus in a manner similar to the aggregation rate itself ($p > 0.05$, between

100, 250 and 500 ng/μL). The results appear more distinct for aS103 ($p < 0.001$, Fisher exact test), however as referenced before, the data for the aggregation of aS103 in the absence of RNA only refer to its first step process. Taking the whole dataset into account shows the RNA actually significantly decreases the fluorescence plateau ($p <<< 0.0001$, one-factor ANOVA), with higher concentrations increasing it back to its original value at 500 ng/μL. Since the overall fluorescence depends on the binding of Proteostat™ dye to the aggregated protein, any changes in its value reflect either an overall decrease of aggregated protein, or the change of aggregate conformation altering the number of available binding sites. It has been shown the various protein conformations within aggregated species greatly affect ThT binding[200], however Proteostat™ is much less characterised in terms of aggregate binding and the overall interpretation is limited. It is however clear that RNA affects either the conformation and consequently the structure of aS aggregates, or the overall level of insoluble protein upon the completion of the elongation phase.

Overall, the results of the aggregation assays clearly indicate that RNA significantly affects the mechanism of self-assembly of both aS140 and aS103. The different levels of effect suggest the effect is protein-specific and in both cases exacerbates the aggregation by both accelerating the elongation phase as well as decreasing the time of its onset.

3.2 RNA affects the structural properties of aS aggregates

Note: the data presented in the following chapter represent a summary of the results from the upcoming manuscript with the title "Investigating the effect of RNA on α-synuclein aggregation by infrared microscopy". The work was done in collaboration with the group of Professor Michele Ortolani from the Department of Physics "Enrico Fermi" at Sapienza University of Rome. The chapter briefly introduces infrared spectroscopy and atomic force microscopy correlated with infrared nanospectroscopy as experimental techniques, then presents the results of the characterisation of aS aggregation. The precise description of the experimental procedures can be found in Materials and Methods.

3.2.1 Introduction and overview of infrared spectroscopy and its applications

In order to further investigate the structural impact of RNA on the aggregation of aS, we employed Fourier-transform infrared micro-spectroscopy (micro-FTIR) and atomic force microscopy coupled with infrared spectroscopy (AFM-IR). Both techniques are applications of infrared (IR) spectroscopy that enable the direct observation of the conformation of chemical bonds and consequently the structural conformations of biological macromolecules[201]. IR spectroscopy is based on the differential absorbance of infrared light by chemical bonds between the atoms. The exact position and strength of the absorption band in the spectrum are determined by several intra- and intermolecular factors that affect the vibrational frequencies of the bonds. The position depends on the type of bond (single, double, tripe), the atomic composition, the electronic couplings as well as the influence of other vibrating masses in its immediate environment. The intensity of the absorption band also increases with the increased polarity of the vibrating bonds[201]. Thus, the vibrational frequencies of certain characteristic bonds of biological macromolecules change depending on their molecular environment, such as the absorbance of the peptide bond in different secondary structure conformations[201]. All of the bonds in a molecule contribute to the overall absorption spectrum, meaning IR spectroscopy can be a powerful tool to study biological macromolecule conformational transitions without the need of a specific label. On the other hand, this fact means that IR spectra of molecules containing a high number of bonds, such as most biological molecules, contain a lot of overlapping absorption bands that need to be identified and evaluated. Specifically for protein conformation characterisation, the most robust method is based on the deconvolution of the amide I band[202]. The amide I band, usually positioned at roughly 1650 cm^{-1} (Table 3) arises mainly from the stretching vibration of the C=O bond in the peptide bond, and depends on the protein secondary structure conformation whilst not being affected by the side chains of the amino acids involved[202]. Analysing the line shape and intensity of the amide I band therefore enables direct tracking of the protein secondary structure and its changes associated with folding, catalysis, misfolding and aggregation[203–206] (Fig. 3.4).

Table 3. The approximate positions of amide I band absorption peaks at different wavenumbers in protein infrared spectra[201].

Secondary structure	Band position in $^1H_2O/cm^{-1}$		Band position in $^2H_2O/cm^{-1}$	
	Average	Extremes	Average	Extremes
α-helix	1654	1648–1657	1652	1642–1660
β-sheet	1633, 1684	1623–1641, 1674–1695	1630, 1679	1615–1638, 1672–1694
Turns	1672	1662–1686	1671	1653–1691
Disordered	1654	1642–1657	1645	1639–1654

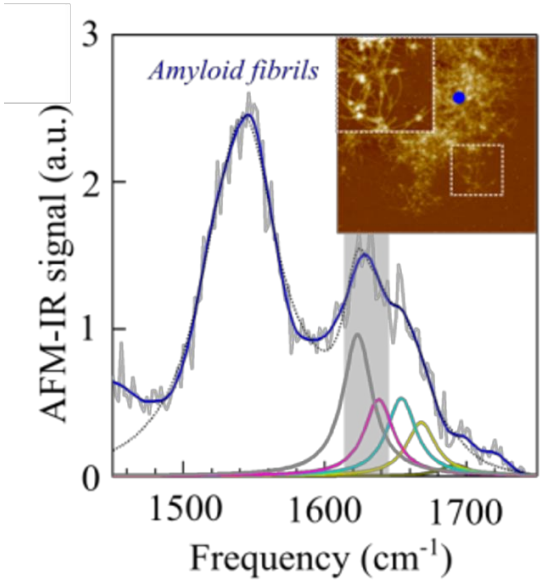


Figure 3.4. Characteristic amide I and amide II absorption peaks in a protein IR spectrum, obtained with AFM-IR. The inset shows the corresponding AFM image with the exact point of the measurement as a blue dot. Shown below the spectrum is the deconvolution of the main absorption band using the five Lorentzian curves representing the different structural elements contributing to the spectrum: antiparallel and parallel β-sheet (grey), disordered (magenta), α-helix (light blue), turn (green) and antiparallel β-sheet (olive).

Micro-FTIR is an application of the technique, based on the coupling of an infrared microscope to a FTIR spectrometer[205]. The smaller field of vision of the microscope allows to record spectra of samples in a much smaller window down to 50 x 50 μm^2, suitable for samples like protein deposits or membrane proteins[207,208]. Despite the much-enhanced resolution, it still suffers from the "bulk" syndrome as the spectrum recorded is the averaged spectrum of all the deposits in the sample. In order to gain an

insight into the structural characteristics of single particles, IR spectroscopy was coupled with AFM to detect the thermal expansion of the molecules upon irradiation with a quantum cascade laser[209,210]. This method does not rely on spectroscopic, but a mechanic detection and can be used at room temperature with a variety of samples ranging from protein complexes[211], membranes[212] to bacteria[213]. Since protein aggregates can range from nanoscale oligomeric structure to microscale fibrillar assemblies, AFM-IR is very suitable for characterisation of protein aggregation, as it enables to both visually distinguish morphologically different assemblies as well as record their individual IR spectra for nanoscale analysis[214,215]. Its sensitivity has been shown to enable the detection of a ligand bound to amyloid fibrils[216]. All of these facts together make a combined approach from micro-FTIR and AFM-IR uniquely suitable for the characterisation of aS aggregation in the presence of RNA.

3.2.2 Characterisation of the influence of RNA on aS aggregation with micro-FTIR

Aggregation is generally characterised by the misfolding and assembly of monomeric proteins into an amyloid fold, enriched with β-sheet structural components[217]. These can be either parallel or antiparallel, with each having a distinct fingerprint in the IR spectrum that is different from that of regular β-sheet conformations in folded proteins[203]. In general, the structure of the amyloid cross-β structure features a highly compact set of extended β-sheets with strong inter-sheet bonding and an orientation perpendicular to the axis of the fibril (hence the name cross-β)[218]. Tracking the evolution of the amide I peak at various time points and performing the deconvolutions of the spectrum allowed a direct tracking of the aggregate-associated spectral markers (Fig. 3.5 and 3.6).

Analysis and comparison of the amide I spectra obtained with micro-FTIR reveal the faster aggregation of aS140 in the presence of RNA (Fig. 3.5). Specifically, the differences between the antiparallel and parallel β-sheet component ($(\beta\text{-sheet})_{P,AP}$) at 1625 cm^{-1} and the antiparallel β-sheet component ($(\beta\text{-sheet})_{AP}$) at 1690 cm^{-1} reveal a significant increase in $(\beta\text{-sheet})_{P,AP}$ already 48 h after the start in presence of RNA while there is no significant change between 24 h and 48 h in protein-alone (Fig. 3.5 A,B). There are no major differences in the other components of the spectrum between the samples, thus indicating that while RNA speeds up the aggregation of aS140, it likely does not drastically change the conformation of the amyloid fibrils. This is confirmed by the similar shape of overlaid spectra, indicating an equal contribution of all the structural components at the end of aggregation (Fig. 3.5 C).

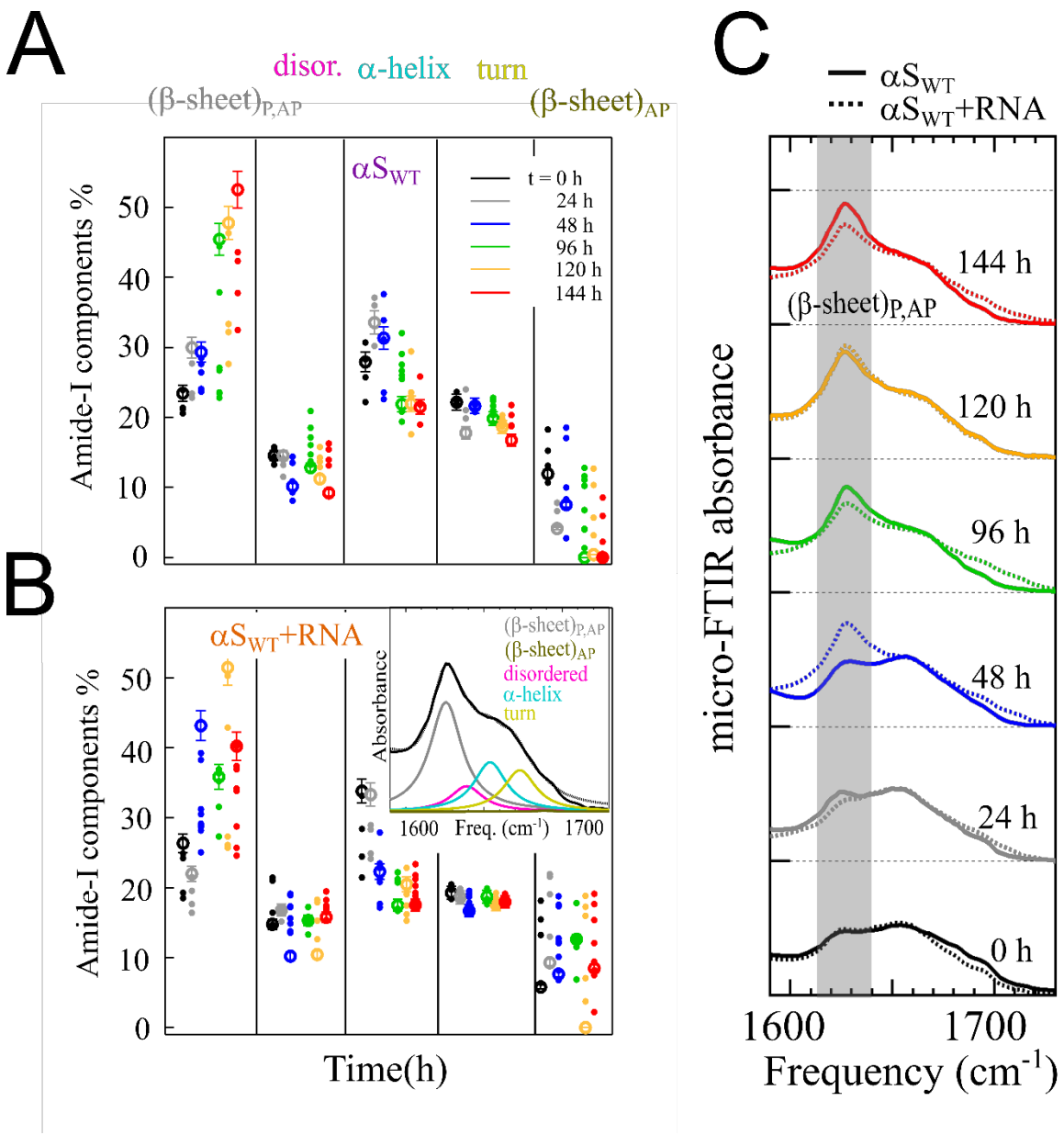


Figure 3.5. Analysis of aS140 aggregation in the presence and absence of RNA with micro-FTIR. **(A)** Percentage of amide I components after the deconvolution of the amide I band into five Lorentzian components, centred at 1625, 1640, 1652, 1670, and 1690 cm^{-1}. The percentages were computed as the ratio of each component's intensity to the sum of the intensities of the five components used to fit the amide I band. The gradual increase in the $(\beta\text{-sheet})_{P,AP}$ clearly indicates the protein aggregation into a characteristic cross-β amyloid structure. **(B)** The coaggregation of aS140 with RNA increases the $(\beta\text{-sheet})_{P,AP}$ percentage starting at t = 48 h, indicating RNA speeds up the formation of the amyloid structure. **(C)** An example of the micro-FTIR spectra with the spectra in the presence (full line) and absence (dashed line) of RNA. Spectra were normalised at 1660 cm^{-1}. The shape of the spectrum at t = 48 h shows the much higher contribution for the $(\beta\text{-sheet})_{P,AP}$ component at roughly 1625 cm^{-1} in the presence of RNA.

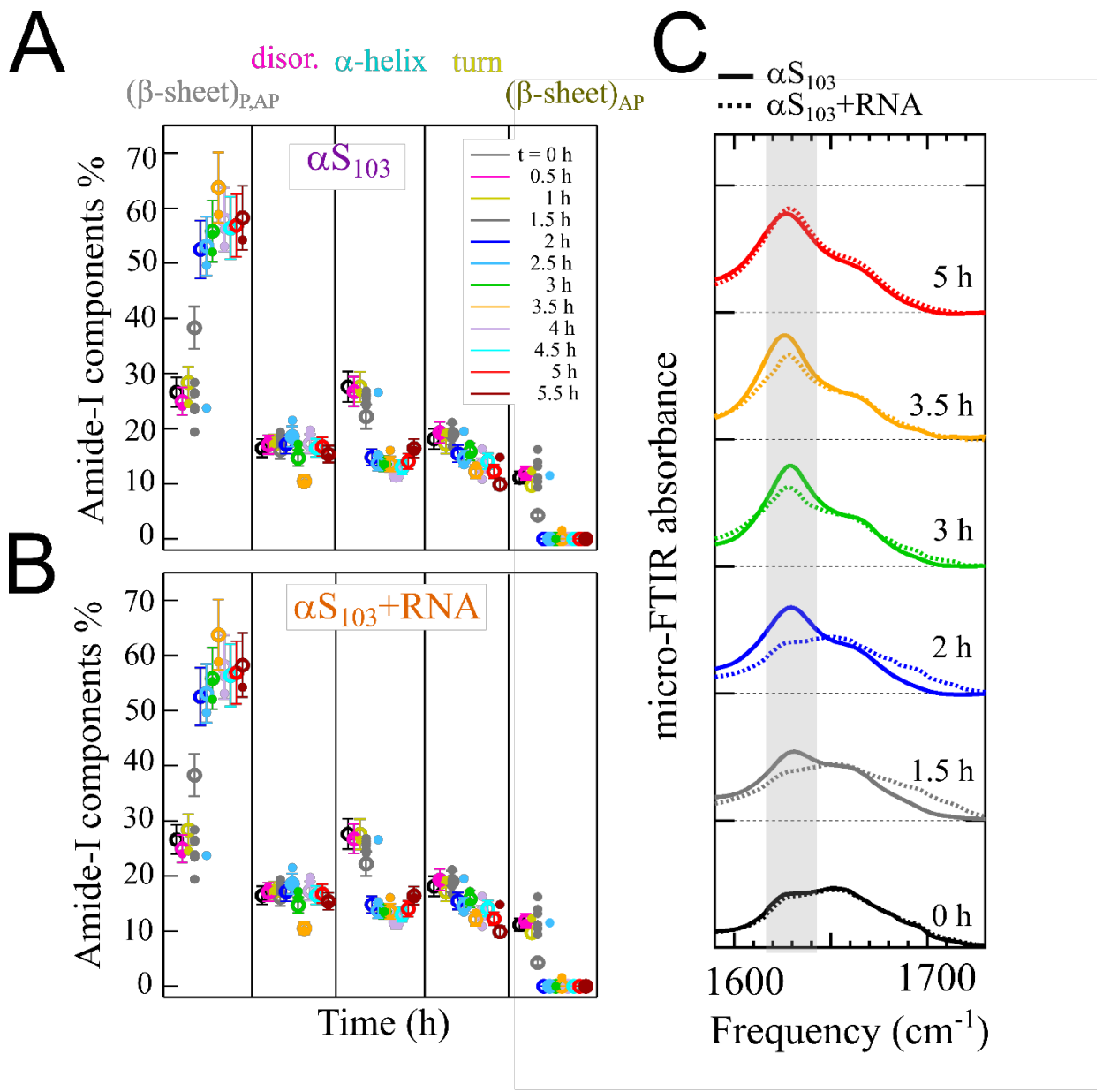


Figure 3.6. Analysis of aS103 aggregation in the presence and absence of RNA with micro-FTIR. **(A)** The measurements in the absence of RNA shows an increase of the $(\beta\text{-sheet})_{P,AP}$ component until the percentage reaches a plateau at roughly t = 2 h. In contrast, the $(\beta\text{-sheet})_{AP}$ component percentage starts decreasing at the same time point. Apart from the slight decrease in the percentage of α-helical structures, there are no major changes in the overall distribution of other components. **(B)** On the contrary, aS103 in the presence of RNA shows a later onset of the $(\beta\text{-sheet})_{P,AP}$ increase, yet the $(\beta\text{-sheet})_{AP}$ component percentage increases slightly at the beginning. Again, apart from the decrease in the α-helix percentage little change is observed. **(C)** The overlap of the spectra in the presence (full line) and absence (dashed line) of RNA confirms the observation. The major differences are observed at intermediate timepoints while the initial and final spectrum appear identical.

The analysis of aS103 aggregation confirms the faster aggregation kinetics compared to aS140, as already shown in Section 3.1 and known from literature[171,198].

Micro-FTIR analysis however points out a major difference between the presence and absence of RNA in the aggregation process. Time-resolved fluorescence experiments show RNA significantly speeds up the aggregation of aS103 (Fig. 3.3), yet the contribution of
(β-sheet)$_{P,AP}$ increases faster in the absence of RNA (Fig. 3.6 A). The (β-sheet)$_{P,AP}$ starts increasing after 1 h and reaches a plateau after 3 h, while in the presence of RNA, the (β-sheet)$_{P,AP}$ starts increasing after 2 h and reaches a plateau only after 5 h. This seemingly contradicts the kinetic data, however it has to be noted that the measurements simply acquire the total fluorescence value in the well without specifically differentiating between aggregate populations with different structural conformations. Analysing the different structural components with micro-FTIR instead reveals that while (β-sheet)$_{P,AP}$ increases faster in the absence of RNA, the presence of RNA favours an rapid increase in the (β-sheet)$_{AP}$ component in the early stages of aggregation (Fig. 3.6 B). The percentage of (β-sheet)$_{AP}$ then drastically decreases towards 0 in perfect coordination with the start of the exponential growth of the (β-sheet)$_{P,AP}$ in both conditions (Fig. 3). The addition of RNA therefore appears to favour the formation of an early-stage intermediate with a higher percentage of antiparallel β-sheets, which then rapidly converts into a more mature amyloid structure with a much lower percentage of the antiparallel β-sheet. Again, as in the case of aS103, it is interesting to note that in terms of amide I components, the spectra in the presence and absence of RNA look almost identical at the beginning and the end of aggregation, with the major differences occurring in the intermediate steps (Fig. 3.6 C). Similarly to aS140, there is also a decrease in the α-helix component during aggregation, while the percentages of disordered and turn remain unchanged.

In order to further understand potential differences between the structural conformations of aggregates grown in the presence and absence of RNA, we shifted our focus from the intensities of the (β-sheet)$_{P,AP}$ to its central frequency (ω_β). As mentioned, the exact position of the (β-sheet)$_{P,AP}$ in the IR spectrum depends on its molecular environment or the conformation of the protein backbone[201]. Different conformations not only cause an increase in the component intensity but can also lead to a shift of its central frequency. In the case of ω_β, factors such as the number of β-strands in a sheet, their length, the number of inter-strand H-bonds etc.[203] The time-resolved shift of ω_β aS140 and aS103 with and without RNA reveals a distinct trend between the two conditions (Fig. 3.7). The ω_β for aS140 increases at t = 48 h in agreement with the increase of the (β-sheet)$_{P,AP}$ component, decrease of (β-sheet)$_{AP}$ and thus indicating the formation of a parallel amyloid structure. The presence of RNA slightly shifts the ω_β at t = 0 h upwards, then shifts towards lower frequencies again at t = 96 h, only to end up more or less at a similar level at 144 h (Fig. 3.7 B,D).

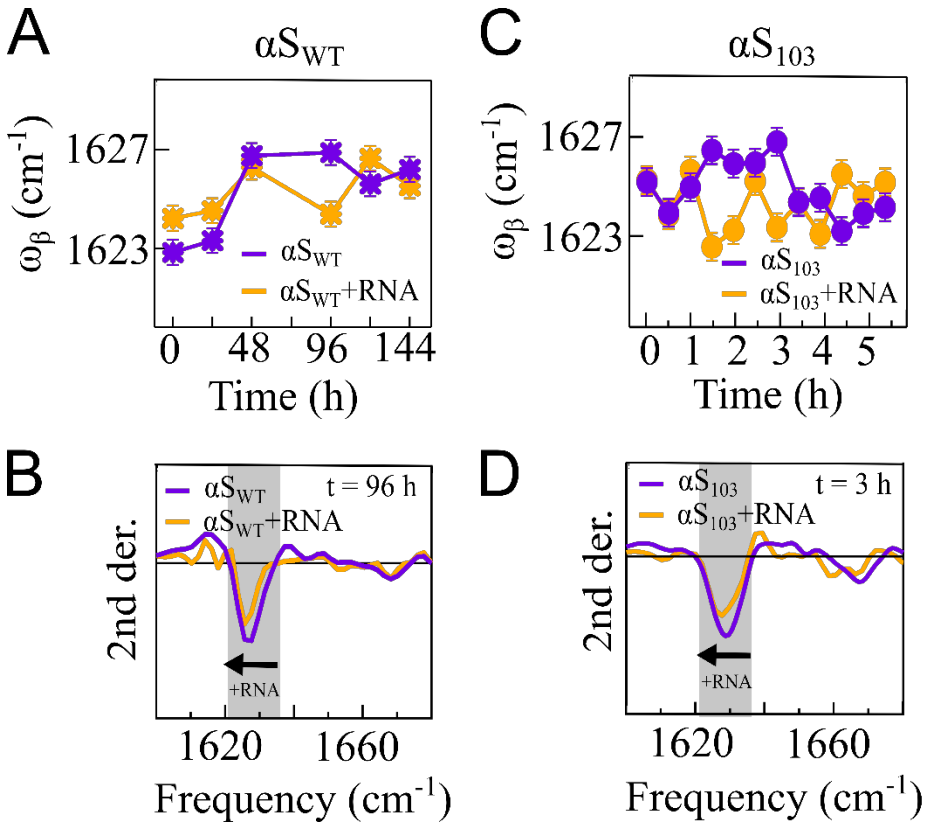


Figure 3.7. The spectral analysis of the central frequency (ω_β) shift of the micro-FTIR spectra $(\beta\text{-sheet})_{P,AP}$ component. **(A)** Central frequency of the $(\beta\text{-sheet})_{P,AP}$ component, with constant error bars ± 0.5 cm^{-1}, resulting from the best fitting curves of the micro-FTIR spectra acquired for the aS140 in the presence and absence of RNA. The frequency of ω_β in the presence of RNA shifts from higher frequencies at t = 0 h towards a slightly lower one at the end compared to the protein-only sample. **(B)** The second order derivative of the spectra at t = 96 h clearly shows the shift towards a lower frequency in the presence of RNA. **(C)** The analysis shows several shifts of ω_β towards lower frequencies during the aggregation process with RNA, however the endpoint is found at a slightly higher frequency. **(D)** The same analysis shown for the second derivative of the spectrum at t = 96 h.

The second derivative of the spectrum at t = 96 h confirms the visible shift towards lower frequencies (Fig. 3.7 B). A similar shift towards lower frequencies occurs during the aggregation of aS103 with RNA (Fig. 3.7 C,D). This further indicates potential structural rearrangements during fibril growth and maturation. In general, the more rigid an amyloid fibril is, the lower the shift of the ω_β in the spectrum[219], indicating the presence of RNA could potentially lead to a more compact and rigid intermediate structure. Variations of ω_β over time could also instead point to the formation of various polymorphs during the aggregation process with their unique spectral contributions. Mature fibrils can also be composed of several filaments and the shifts could be attributed to different filament assembly pathways[220].

3.2.3 Characterisation of fibril mechanical properties with AFM-IR

To solve these questions, AFM-IR was used to give the IR measurements the additional dimension of visualising the aggregates and analysing their mechanical and spectroscopic properties. Representative AFM images of aS samples at various times reveal the first fibrillar growth occurring already at 24 h in the presence of RNA (Fig. 3.8 A). This confirms the observed faster kinetics compared to protein-only conditions, as the first fibrillar assemblies are observed only at 96 h. AFM images also offer an explanation as to the observed non-zero contribution of $(\beta\text{-sheet})_{P,AP}$ in micro-FTIR measurements already at t = 0 h, since non-fibrillar structures are present already at very early time. This is consistent with the reported observation of fast-forming aS non-fibrillar oligomeric species, especially in a buffer containing higher salt concentrations such as 100 mM KCl used in the experiment[221].

Analysing the dimensions of the fibrils shows a wide distribution of the diameters ranging from 40 to 130 Å (Fig. 3.8 B). This is consistent with reported analysis from previous AFM studies[222]. The fibrils grown in the presence of RNA exhibit a smaller diameter with an average of 45 Å and a much more uniform distribution. These fibrils also exhibit a higher persistence length, which is a parameter used to indicate the stiffness or the rigidity of the fibril[223] (Fig. 3.8 C). Finally, they also exhibit a much more uniform distribution of the ω_β for the $(\beta\text{-sheet})_{P,AP}$, together with the overall ω_β towards lower frequencies, compared to the bimodal distribution of high- and low-frequency ω_β for aggregation without RNA (Fig. 3.8 D). As mentioned previously, a lower ω_β corresponds to a higher strand content, length and inter-strand chemical bonds, all hallmarks of a much more compact fibril architecture. Together with the fibril diameter and persistence length measurements, these data show in agreement that RNA changes the structural conformation and consequently the mechanical properties of aS aggregates, as previously indicated from micro-FTIR data.

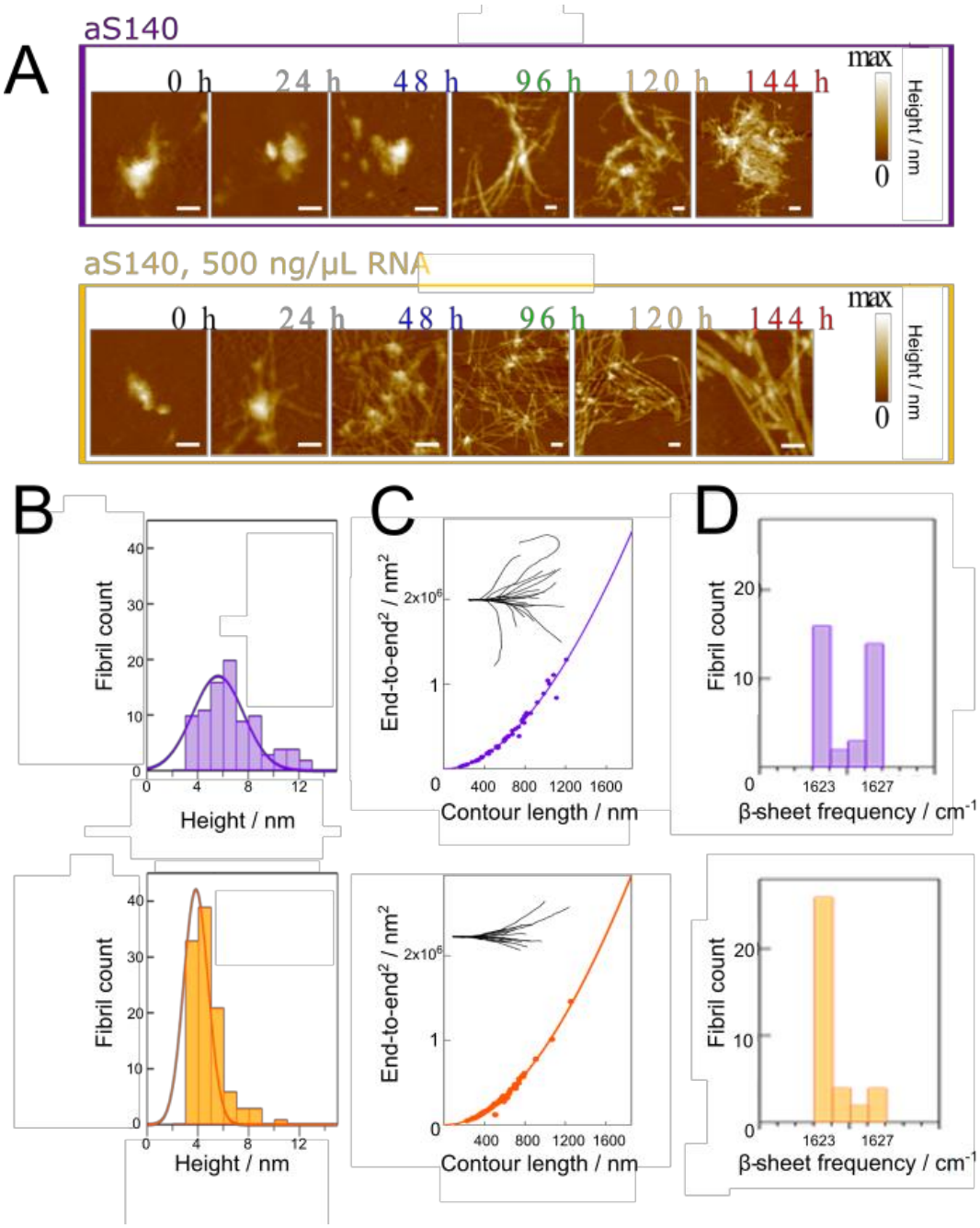


Figure 3.8. Analysis of aS aggregation in presence and absence of RNA with AFM-IR. **(A)** AFM micrograms show the gradual evolution of fibrillar structures at different timepoints. These start appearing at t > 96 h in the absence of RNA, however some protofibrils are already visible at t = 24 h in the presence of RNA. **(B)** Analyses of the fibril thickness and the persistence length of the fibrils show further differences. Coaggregation with RNA leads to fibrils at t = 144 h with a smaller diameter ($d_{no\ RNA}$ = 6.06 ± 2.90 nm (N = 89) compared to d_{RNA} = 4.45 ± 1.31 nm (N = 106)) and a higher persistence length ($L_{no\ RNA}$ = 1435 ± 185 nm compared to L_{RNA} = 1938 ± 217 nm). Both parameters indicate a much more compact structure in the presence of RNA. **(C)** Example of amide I band deconvolution from the spectra obtained with AFM-IR. The point of data acquisition is shown in the inserted AFM microgram, indicated with a blue dot. **(D)** Deconvolution of amide I band in the presence and absence of RNA shows a more significant shift of ω_β towards lower frequencies in the presence of RNA.

3.2.4 RNA speeds up the aggregation and changes the fibril conformation of aS

Combining the data from the experiments with both micro-FTIR and AFM-IR shows clearly that RNA does in fact speed up the aggregation of aS. This last part of the chapter is intended to briefly summarise the contributions of these techniques on individual experimental conditions. Micro-FTIR measurements of aS140, combined with AFM data, show that a certain population of non-fibrillar oligomers is formed in the very early stages of the experiment. Their formation does not seem to be influenced by the presence of RNA since the spectra at t = 0 seem to overlap well, as do the percentages of individual components. These oligomeric structures appear to be mostly composed of antiparallel β-sheets, which is in good agreement with previously reported data[180]. Coincubation of aS140 with RNA then causes a faster increase in the percentage of the $(\beta\text{-sheet})_{P,AP}$ component, which coupled with AFM images points to a faster fibrillation. Characterisation of the fibrils at the endpoint shows that there is a mixed population, likely composed of two major species in the absence of RNA as indicated by the spectral markers in Fig. 3.8 D. Coaggregation with RNA in turn appears to favour the formation of the more compact, thinner population with a smaller diameter and increased persistence length. As previously mentioned, unfortunately we do not have the measurements for aS103 available at the moment of writing. However, micro-FTIR results indicate that similarly to aS140, there is a population of oligomeric species formed very early in the aggregation process as indicated by the non-zero initial values of β-sheet components. Similarly to aS140, these species appear to have a structure composed mostly of antiparallel β-sheets. However, the addition of RNA increases the formation of these antiparallel β-sheets species temporarily, until they get converted into the mostly parallel β-sheet rich fibrillar species at a higher rate compared to the protein without RNA. The end structures appear to be spectrally similar, however macroscopically it is difficult to conclude about their overlap without AFM image analysis. During the analysis we unfortunately could not establish reliably whether RNA directly interacts with the protein or contributes to the spectrum. It has been previously shown that AFM-IR can be successfully used in determining the binding between a small molecule ligand and an amyloid assembly[216], however in our case the lack of visible changes might have been due to the overall lack of clear, RNA-specific spectral markers due to their overlap with the amide I region of proteins[224]. Overall, we can still conclude that spectroscopic analysis of aS aggregation clearly indicates that RNA influences not only the kinetics, but also the structural and mechanical properties of the aggregates formed. In the next chapters, the ability of aS aggregates to bind RNA will be demonstrated and put into a wider context, further expanding on the already presented data in the previous two chapters and highlighting its role as modulator of protein aggregation.

3.3 aS sequesters RNA upon aggregation

Having observed the effect RNA has both on the aggregation kinetics of aS as well as the actual structural mechanism and architecture, the next question that posed itself was an obvious one: does RNA interact directly with the protein? What are the characteristics of this interaction? Recent works has shown that aS is directly present in cellular mechanisms the involve nucleic acids[148,194] and several DNA aptamers[192,193] as well as double-stranded DNA[189] are known to bind aS.

3.3.1 aS103 aggregates sequester larger quantities of RNA than aS140

As an estimate of the RNA degradation during the aggregation assays, its concentration was always measured before and after the experiment. These experiments were always performed with control wells containing only RNA and after quantification, an interesting observation was made that the protein wells consistently had much less soluble RNA present than the control wells. In order to determine whether this was due to degradation or some other process, we collected the samples in the wells after the aggregation, centrifuged them and quantified the RNA found in both soluble and insoluble fractions (see Materials and Methods). The concentration was measured individually for each well and normalised according to the initial RNA concentration in the same well, in order to account for individual rates of RNA degradation. Also, to avoid any interference by remaining non-aggregated protein and smaller aggregates, we measured the RNA concentration of using an RNA-specific fluorescent dye. Importantly, to bypass any potential differences between the volumes of the soluble and insoluble fractions, concentrations were recalculated as RNA mass and presented as the percentage of total RNA mass after aggregation in different fractions.

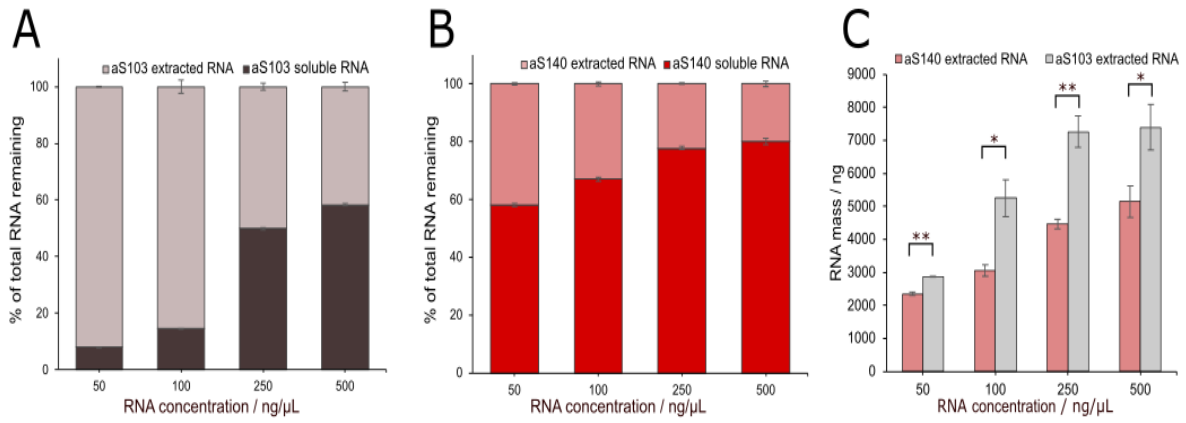


Figure 3.9. Quantification of RNA found in soluble and insoluble fractions after the aggregation assays. The data represent average values from 3 different experiments with 3 replicates each. **(A)** At lower initial RNA concentrations, the vast majority of RNA is found in the insoluble fraction with aS103, then the percentage gradually decreases as more RNA is left in solution afterwards. **(B)** The trend is similar with aS140, however the percentage of RNA found in the insoluble fraction is smaller. **(C)** The average amount

of RNA extracted from aS103 and aS140 aggregates at different initial starting concentrations show again that aS103 aggregates have a higher propensity of sequestering soluble RNA upon aggregation (Two-way paired Student t-test, N = 9, * $p < 0.05$, ** < p 0.01).

These percentages were individually adjusted to 100% to again take into account the different rates of RNA degradation during experiment *Fig. 3.9 A,B). These results clearly indicate that RNA is found sequestered in the protein aggregates of both aS103 and aS140. Furthermore, there is a clear difference between the amount of RNA sequestered by individual protein constructs, with significantly higher amounts of RNA found in aS103 aggregates in all conditions (Fig. 3.9 C). As the starting concentration of the protein is always 50 μM, it is safe to assume the difference is due to the higher sequestration capacity of aS103 aggregates. Higher initial concentration of RNA increases the amount sequestered until it appears to reach saturation at 250 ng/μL for aS103 ($p > 0.8$ for comparison between 250 and 500 ng/μL, two-way paired Student t-test), indicating all the binding spots on the aggregates appear saturated.

In order to confirm that aggregation is indeed the underlying mechanism of the RNA sequestration, we repeated the experiments in the same conditions with bovine serum albumin, a non-aggregating, non-RNA-binding protein. The quantification of RNA at the end of the assay showed comparable values to the control values of total RNA without any protein added (Fig. 3.10 A). Also, no visible aggregates were observed compared to wells containing aS (Fig. 3.10 B).

3.3.2 RNA is sequestered to a larger extent than DNA in aS aggregates

We repeated the aggregation of aS103 and aS140 also in the presence of a total cell DNA extract, since aS aggregates have been previously shown to bind both plasmid and linearised DNA[188]. aS103 has a higher propensity to sequester DNA as well, however to a lesser extent compared to RNA (Fig. 3.10 D). aS140 still shows the ability to sequester DNA at lower initial DNA concentrations, however higher concentrations apparently limit the binding ability. aS140 barely sequesters any DNA in the aggregates, with less than 10% of the initial DNA extracted from the insoluble part. Comparing the actual quantities of nucleic acids extracted from the aggregates show that a higher amount of RNA was extracted from aS140 aggregates compared to DNA (Fig. 3.11 A). The quantities are significantly higher in all conditions ($p < 0.01$, two-way paired Student t-test) apart from 500 ng/μL ($p > 0.1$, two-way paired Student t-test). On the other hand, the is little difference between RNA and DNA for aS103 until 500 ng/μL, where the amount of DNA continues to rise while an apparent saturation is reached for (Fig. 3.11 B).

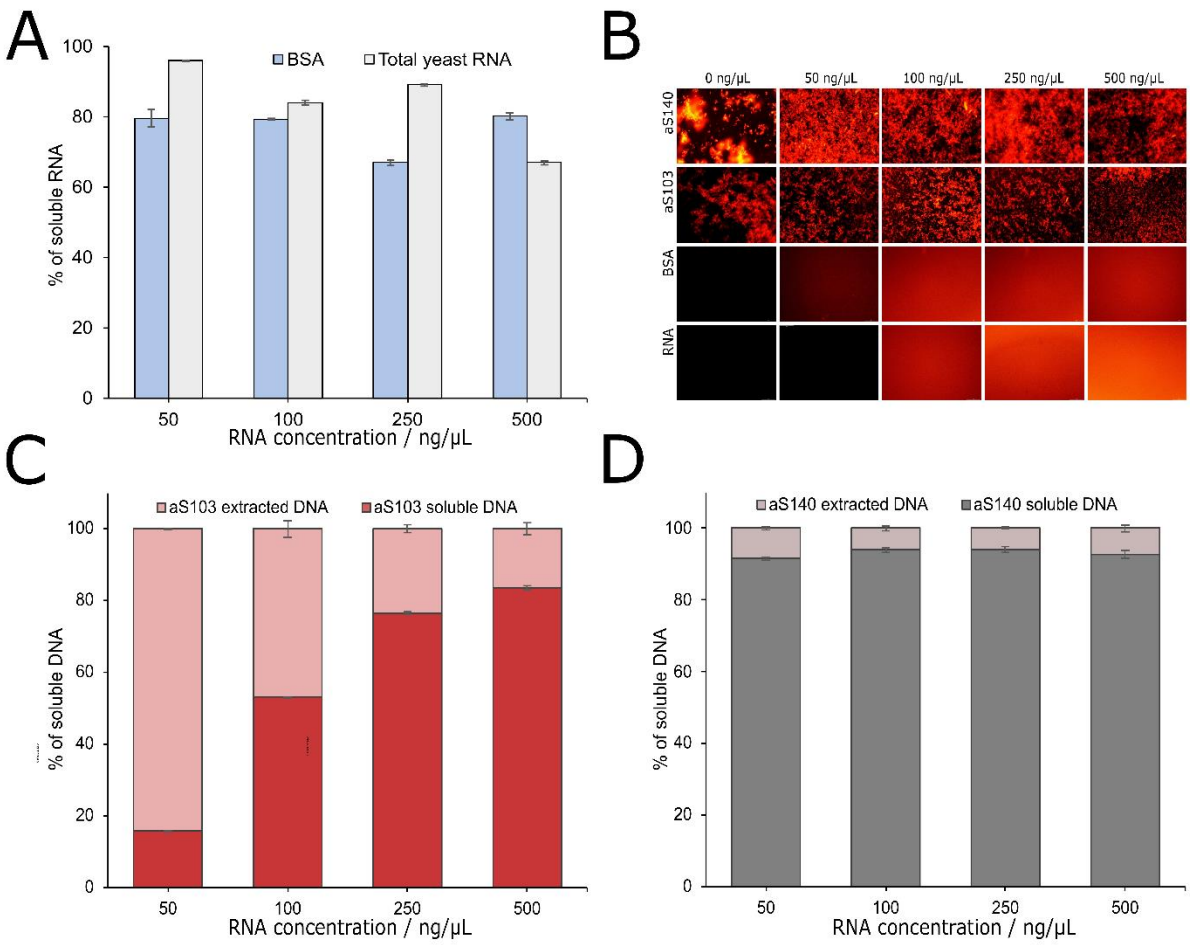


Figure 3.10. Additional control experiments demonstrating that aS aggregation is indeed the mechanism of nucleic acid sequestration. **(A)** Quantification of soluble RNA after incubation with BSA shows comparable levels to RNA alone. **(B)** Fluorescent microscope images of the wells at 20x magnification. There was no aggregation observed in conditions with BSA and RNA alone compared to aS103 and aS140 and no pellet was observed after centrifugation (data not shown). **(C,D)** Quantification of total cell DNA found in both the soluble and insoluble fractions of aS103 and aS140. The trends are similar to the trends in the presence of RNA, however the percentages of DNA found in the insoluble fraction is much lower compared to RNA, especially for aS140.

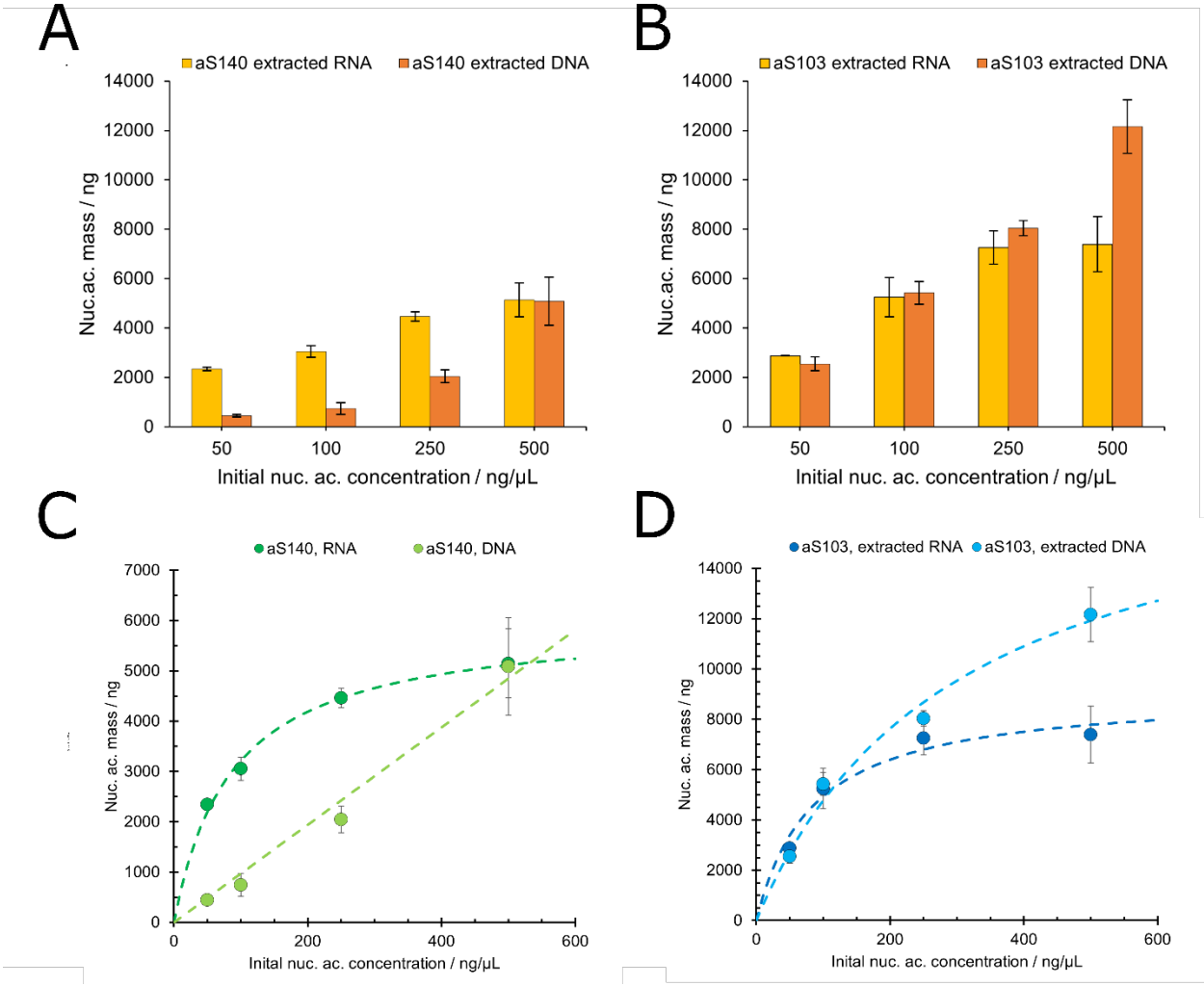


Figure 3.11. Comparing the amounts of RNA and DNA isolated from the aggregates of aS. (**A**) aS140 aggregates sequester larger amounts of RNA compared to DNA until 500 ng/µL when the difference becomes non-significant ($p > 0.9$, N = 9 for RNA and N = 6 for DNA, two-way paired Student t-test). (**B**) aS103 aggregates sequester nucleic acids to a much larger extent compared to aS140. There is no significant difference between the amount of RNA and DNA quantified ($p > 0.05$, N = 9 for RNA and N = 6 for DNA, two-way paired Student t-test) until 500 ng/µL, when the amount of DNA significantly increases compared to RNA ($p < 0.01$, two-way paired Student t-test). (**C**) Fitting the aS140 data with the basic 1:1 interaction model at presumed [RNA/DNA] >> [aggregate] reveals the approximate binding constant at 96 ± 61 ng/µL for total yeast RNA, while the data for DNA quantification could not be fitted. (**D**) The same comparison for aS103 gives the approximate K_D for RNA at 93± 61 ng/µL and for DNA at 309 ± 89 ng/µL.

In an attempt to quantify and compare the binding affinities of the fibrils for RNA and DNA, the extracted quantities were plotted against initial concentration and fitted to the simplistic model of reversible 1:1 binding to one binding spot (Fig. 3.11 C,D):

$$F + NA \rightarrow FNA$$

with F standing for fibrils/aggregates and NA for nucleic acid, the constant of binding is:

$$K_{D,app} = \frac{[F][NA]}{[FNA]}$$

The data can therefore be fitted to the basic equation of:

$$[FNA] = \frac{[FNA]_{max} \cdot [FNA]}{K_{D,app} + [FNA]}$$

Where $[FNA]_{max}$ refers to the concentration of the complex at saturation of binding. The constant of binding is just an estimate in this case, better termed as $K_{apparent}$ or $K_{D,app}$, since the mechanism of binding as well as the number of binding sites is unknown. Even more importantly, the nucleic acids in this case are total cell extracts with molecules of different molecular masses, therefore the exact number of molecules bound is unknown and the constant cannot be determined correctly. This enabled us to determine the $K_{D,app}$ for the binding of RNA for both aS140 (96 ± 61 ng/µL) and aS130 (93± 61 ng/µL) as well as the $K_{D,app}$ between DNA and aS103 aggregates (309 ± 89 ng/µL). As simplistic as the model is, these values can still serve as a rough estimate of the binding affinities and show the higher preference of both aS103 and aS140 aggregated forms towards RNA compared to DNA.

3.3.3 Monomeric aS does not bind RNA

To further contextualise the affinity of the aggregated species towards RNA, we attempted to determine and quantify the binding between the monomer and the RNA as well. Monomeric aS has been shown to interact and bind DNA[148,194], however no definite mechanism of binding as well as affinity constants have been determined as of yet. The initial binding screening was performed with an RNA electromobility shift assay (rEMSA). Incubating increasing concentrations of aS in the presence of RNA did not give any significant shifts of the protein (data not shown). Furthermore, due to the necessity of performing it in native conditions, the intrinsically disordered structure of the protein caused it to migrate in the gel in the form of a long smear, rendering any analysis of the gel shift impossible. Therefore, we have tried to quantify the interaction with two further techniques, commonly used to quantify RNA-protein interactions: microscale thermophoresis (MST) and biolayer interferometry (BLI) (see Materials and Methods).

Both techniques have their own advantages and disadvantages, with the chief difference being the necessity to immobilise the protein in the case of BLI[225]. We performed the MST experiments with a specific RNA at a 100-molar excess, yet observed only a slight shift in the fluorescence curve that was well within the signal-to-noise ratio, indicating no significant binding (Fig. 3.12 A,B). A similar thing occurred with BLI, with both protein constructs failing to induce any significant shift in the biosensor response (Fig. 3.12 C).

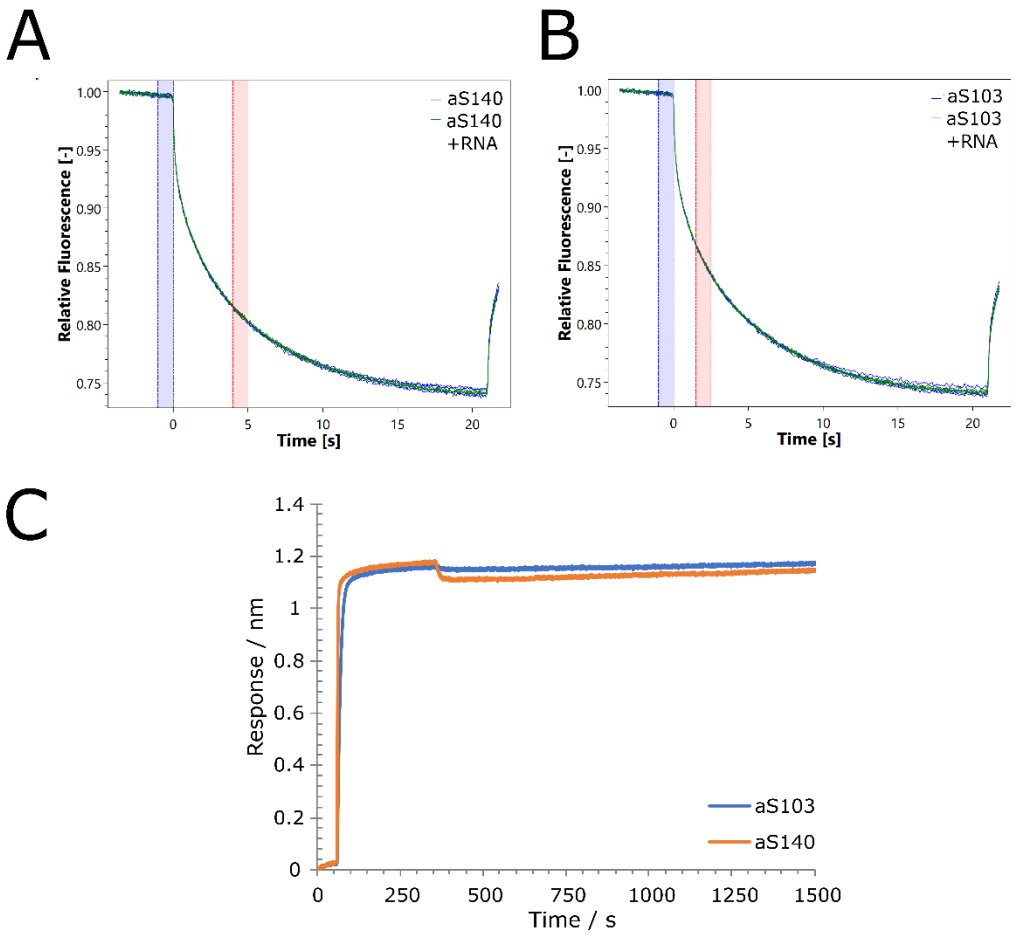


Figure 3.12. Attempts to quantify the binding of monomeric aS to RNA. **(A,B)** MST thermograms show no difference between the protein-only (blue) and protein-RNA (green), indicating no significant interaction occurs in the selected conditions. **(C)** Response traces for BLI measurements of aS103 (blue) and aS140 (orange) in the presence of RNA. The measurements were performed with 500 ng/µL total yeast RNA and the proteins were immobilised on the streptavidin biosensor as described in Materials and methods. The protein loads onto the sensor well enough during the loading phase, yet no shifts are observed during association and dissociation phases, indicating no detected binding.

The failure to obtain a positive result further indicates that any interaction between monomeric aS and RNA are likely transient and occurring in the context of a disordered conformational ensemble, or "fuzzy" according to a term that has been gaining in

popularity recently[226]. A common example of this kind of interaction mode between RNA and proteins is SERF1[162], which coincidentally is also an interacting partner of aS. These highly transient interactions with low or no binding specificities are difficult to quantify with classic techniques and most often rely on NMR analysis of structural perturbations upon binding, something that unfortunately was not realistic for the timeline of our study. Nonetheless, this gives further emphasis on the fact that the propensity of aS to bind and sequester nucleic acids significantly increases upon aggregation. As the next chapter illustrates, this is part of a broader trend of amongst aggregation-prone proteins, where the structural changes upon misfolding promote RNA binding and can lead to proteins acquiring novel interaction propensities upon aggregation.

3.4 Aggregation increases the RNA-binding propensity of amyloidogenic proteins

Note: This chapter is intended as a summary of the contents of the recently published article titled "RNA sequestration driven by amyloid formation: the alpha synuclein case" in the journal Nucleic Acids Research[196]. The chapter recaps the computational aspects of the work which aim to give context to the experimental results discussed in the first three chapters of the thesis. The introduction and experimental results have been omitted from the text since these have already been presented and expanded upon in the previous chapters.

3.4.1 Short introduction

This article builds on the previous work and experience within our group when predicting the structural and physicochemical properties of amyloids, especially on computational approaches the enable the successful prediction of protein aggregation propensities and identification of their amyloid core regions[227–229]. To briefly outline the workflow, the Zyggregator[227] method was applied to a recently published curated list of aggregation-prone proteins[230]. Since Zyggregator can effectively identify the amyloid core, we used the analysis to effectively discriminate between the residues that form the core region and the ones that are left outside. We then analysed the physicochemical properties of these "inner" and "outer" regions based on parameters known to influence RNA binding, such as hydrophobicity, disorder, charge, as well as use catRAPID, to assess potential differences in the RNA-binding propensity. This method was applied first to microtubule-associated protein tau, a known factor involved in various neurodegenerative disorders[231]. After validation, the method was used to analyse the aforementioned dataset of proteins to identify potential broader trends. Finally, both aS140 and aS103 were assessed in the same way in order to give context to the experimentally obtained results. For a more in-detail overview of the computational methodology, the reader is invited to refer to the Materials and Methods section of the publication.

3.4.2 Calculations recapitulate experimental results on the ability of tau fibrils to attract RNA

As already mentioned above, we started by assessing the robustness of our calculations of aggregation propensities on the known case of tau. A cryo-EM structure of tau fibrils (PDB 7SP1), grown in the presence of total RNA, shows an electron density in a pocket on the external region of the aggregate where an RNA molecule was later additionally docked into[232]. The 36-amino acid region of the C-terminal disordered domain (residues 391–426) is identified in the cryo-EM study as the fibril core[232] and another work indicates that the third microtubule-binding repeat (residues 306–335)[233,234] contributes to tau aggregation. Using the Zyggregator algorithm, which predicts aggregation-prone elements in protein sequences[227], we are able to successfully identify these two regions as the ones with highest aggregation propensity (peaks at residues 320 and 409,

Fig. 3.13 A). This analysis indicates that Zyggregator is able to identify the inner regions of the tau aggregate in detail.

At the experimental level, the binding of docked RNA is stabilised through hydrogen bonds between the phosphate backbone, the 2-OH group of the ribose, and additional contacts between the RNA bases and the aggregate core[232]. More in detail, the cryo-EM structure of aggregated tau reveals that residues R406 and H407 on the tau fibril surface directly interact with sequestered RNA and that residues E391-S396, V398, S400, D402, S404, D421, P423, I425, A426 are exposed to the solvent (Fig. 3.13 B). An examination of residues 391–426 reveals that the calculated physicochemical properties are consistent with cryo-EM results. This confirms that the experimentally identified surface regions exhibit a higher propensity for RNA-binding. Additionally, these regions are predicted to possess less hydrophobicity and more disorder, as detailed in Fig. 3.13 B.

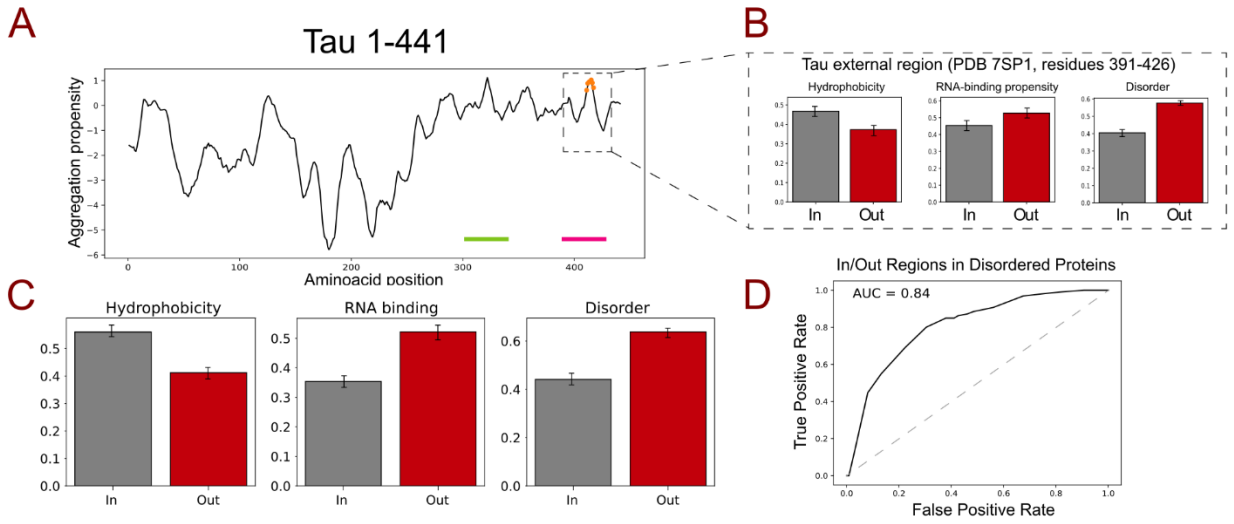


Figure 3.13. Analysis of tau aggregation and prediction of physicochemical properties. (**A**) Aggregation propensity profile for tau, with orange dots representing the predicted internal part of the fibril, and green and pink bars indicating regions experimentally identified as part of the fibril core. (**B**) Comparison of hydrophobicity, RNA-binding, and disorder propensities for residues 391-426, with features of experimentally identified internal (In) and external (Out) regions marked in black and red, respectively (mean and standard deviation shown for these regions). (**C**) Analysis of hydrophobicity, RNA-binding propensity, and disorder across the entire tau sequence, using the Zyggregator method to distinguish between internal and external regions (with a threshold of 0.8; error bars are calculated by varying the Zyggregator threshold within the 0.76-0.84 range). (**D**) Performance evaluation of the Zyggregator method in differentiating between the internal and external regions of experimentally identified protein aggregates, with an optimal threshold of 0.8, as indicated by the Youden index of the Receiver Operating Characteristic (ROC) curve. With regard to panel C, we note that in the range 0.76-0.84 the Zyggregator scores are associated with a True Positive Rate (TPR) of 0.75 and a False Positive Rate (FPR) of 0.25.

We broadened our analysis to encompass the entirety of the tau sequence, thereby incorporating regions that lie outside of the core aggregate at residues 391–426 (Fig. 3.13 C). Upon examining the regions associated with low Zyggregator scores, it is notable that peaks within residues 175–220 and 230–260, which correspond to the proline-rich region and first microtubule-binding repeat, have been highlighted in previous research for their direct involvement in interactions with nucleic acids[235]. This finding suggests a high likelihood of their localization within the external regions of the amyloid structure (Fig. 3.13 C). We observe that these regions, characterised by inherent disorder, may evade detection through traditional experimental techniques such as cryo-EM. Consequently, their potential contribution to the sequestration of RNA during aggregation cannot be dismissed. This possibility has also been acknowledged by the authors of the cryo-EM study[232].

This initial result indicates that regions outside the core of tau amyloid fibrils have the potential to interact with RNA and we are able to identify them in great detail using our computational approaches. This finding also suggests that other aggregating proteins might behave as tau.

In an effort to explore the prevailing trend that regions capable of binding RNA are located on the exterior of amyloids, we sourced a dataset of amyloids from a previous publication[230]. We then analysed their physicochemical properties employing the Zyggregator algorithm (Supplementary Table S1). This allowed us to pinpoint the regions constituting the amyloid core. We focused on proteins that partially exhibit an unfolded state in their monomeric form. We especially prioritised proteins that not only can be found in their amyloid form, but that have been also empirically determined and stringently validated. Our focus has been on the following proteins: calcitonin[236,237], amylin[238], glucagon[239], Aβ42[240,241], alpha synuclein[242–244], hnRNPDL-2[73] and tau[232]. The model achieves an Area Under the Curve (AUC) of 0.84 on the Receiver Operating Curve (ROC), thereby effectively identifying the amyloid core of the aggregates. The optimal Zyggregator threshold is noted to be 0.8, as indicated by the Youden Index of the ROC curve, which corresponds to a True Positive Rate (TPR) of 0.75 and a False Positive Rate (FPR) of 0.25 (Fig. 3.13 D).

As expected, our calculations show that the inner region of an amyloid is more hydrophobic than the outer one (Fig. 3.14 A). There is a significant shift of the RNA-binding propensity and structural disorder in favor of the external region (Fig. 3.14 B,C) while the electrostatic charge distribution does not differ substantially (Supplementary Figure S1A and B). Notably, the trend is observed with different physicochemical propensity predictors (Supplementary Figure S2). Given the overall negative charge of the phosphate backbone of RNA and no significant difference in the aggregate charge distribution, we find strong RNA-binding propensity exclusively for positively charged protein regions.

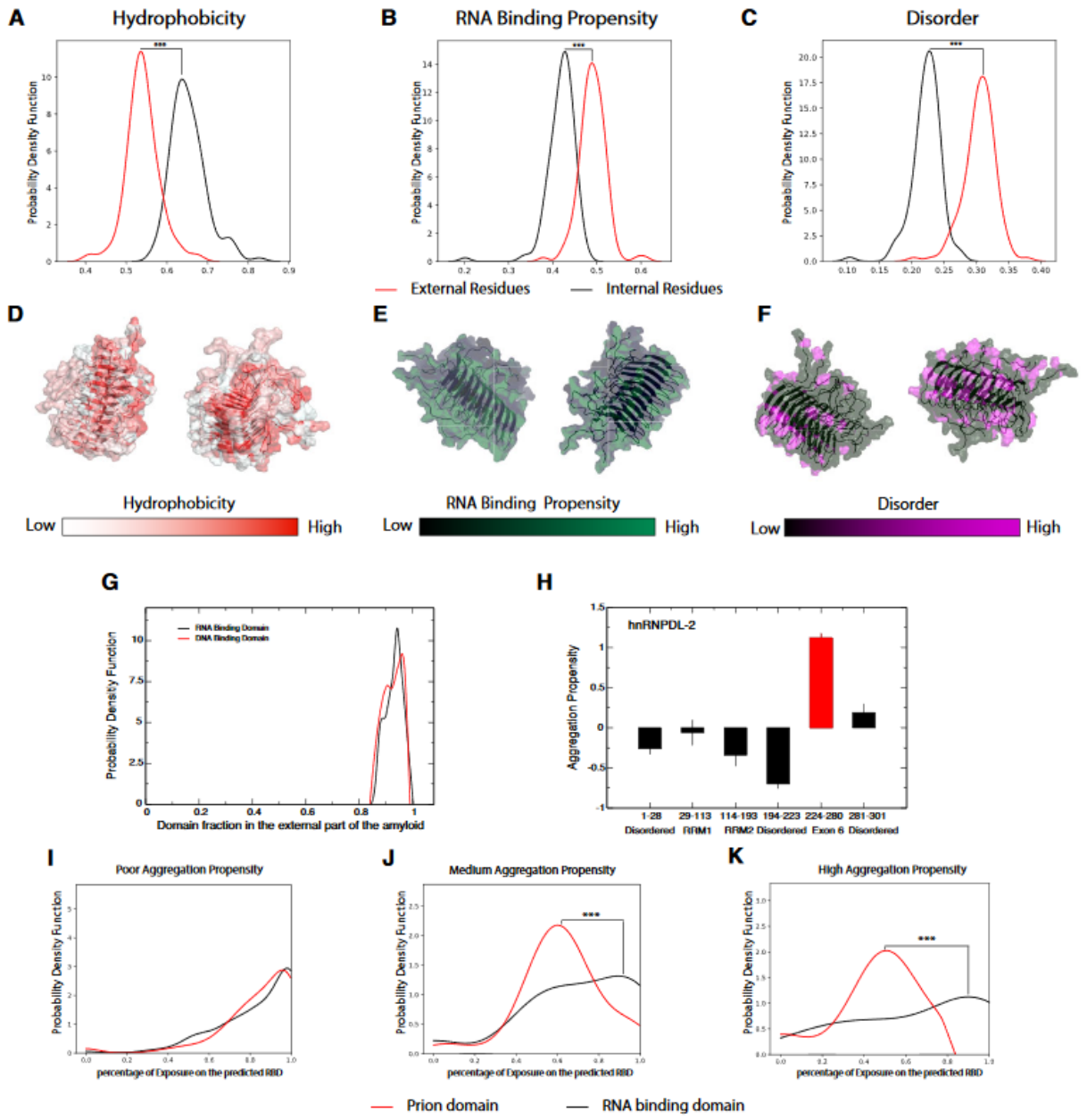


Figure 3.14. Physicochemical properties of internal and external regions of amyloids. **(A-C)** Differences in hydrophobicity, RNA-binding, and disorder propensities were calculated for the internal and external portions of amyloids, which were identified using the Zyggregator method (Supplementary Table 2; Supplementary Figure 2). **(D-F)** Molecular structure of the HET-s prion protein (PDB 2RNM) solved by solid-state NMR, depicted by different colour scales for hydrophobicity, RNA-binding propensity, and disorder, respectively. **(G)** Fraction of the RNA and DNA binding domains predicted to belong to the outer region of amyloid fibrils. **(H)** Zyggregator scores for hnRNPDL-2 domains, showing the average values along with the corresponding standard error of the mean. The red colour highlights Exon 6, which has experimentally demonstrated to form the core of the aggregate. **(I-K)** Fraction of exposed residues in the RNA-binding and prion domains of the human proteome at different Zyggregator scores. *** indicate p-values < 0.0001 (Kolmogorov-Smirnov test).

To better visualise this trend, we chose the structure of the amyloid formed by *Podospora anserina* protein HET-s (PDB 2RNM) as an example (Fig. 3.14 D-F). The structure was determined using ssNMR and included the flexible outer region of the amyloid. Using different colour scales, the distribution trends are clearly visible, especially in the hydrophobic amyloid core (Fig. 3.14 D-F). The RNA-binding propensity is considerably higher for the unstructured external portion of the fibrils, similarly to the structural disorder (Fig. 3.14 D-F, Supplementary Figure S2B). Overall, the analysis of the HET-s amyloid structure shows excellent agreement with the predicted results from the computational analysis.

Within the set of amyloids, we analysed proteins that have RNA and DNA binding domains annotated in Uniprot (Supplementary Tables S2 and S3). Consistent with our calculations of RNA-binding abilities, the nucleic acid-binding domains are predicted to be displayed on the surface of the aggregate (Fig. 3.14 G). This is well exemplified by hnRNPDL-2, a protein in which residues 224–280 were found by cryo-EM to constitute the core of the fibril (PDB 7ZIR), while two RNA-binding domains are external[73]. The aggregation propensity predicted by Zyggregator captures the inner part of the aggregate very well, and the RNA-binding domains are indeed predicted to be more exposed (Fig. 3.14 H). Further investigating this phenomenon at the human proteome level, we carried out predictions of prion and RNA-binding domains, respectively using PLAAC[245] and *cat*RAPID signature[246]. Proportionally to the overall aggregation propensity computed with Zyggregator, we estimated that prion domains are in the core of the fibril while RNA-binding domains tend to be on the outer regions (Fig. 3.14 I-K, Supplementary Table S4). The literature is in agreement with our analysis of hnRNPDL-2[73].

3.4.3 Physicochemical properties of aS aggregates

We next analysed the aggregation propensity of alpha synuclein (aS or aS140), a protein with a well-characterised amyloid core[228]. We computed the aggregation propensity of the wild-type sequence (Fig. 3.15 A) and performed a mutational analysis with 10000 random single point variations[5]. As supported by data reported in the literature[247], the central region of the protein emerges as a key determinant in protein aggregation, exhibiting a prominent Zyggregator peak that aligns with the NAC region (residues 61–95; Fig. 3.15 A). Furthermore, the work by Doherty and colleagues[152] sheds light on additional smaller regions, specifically residues 36–42 and 45–57, which coincide with minor Zyggregator peaks and play a crucial role in modulating aS aggregation (Fig. 3.15 A).

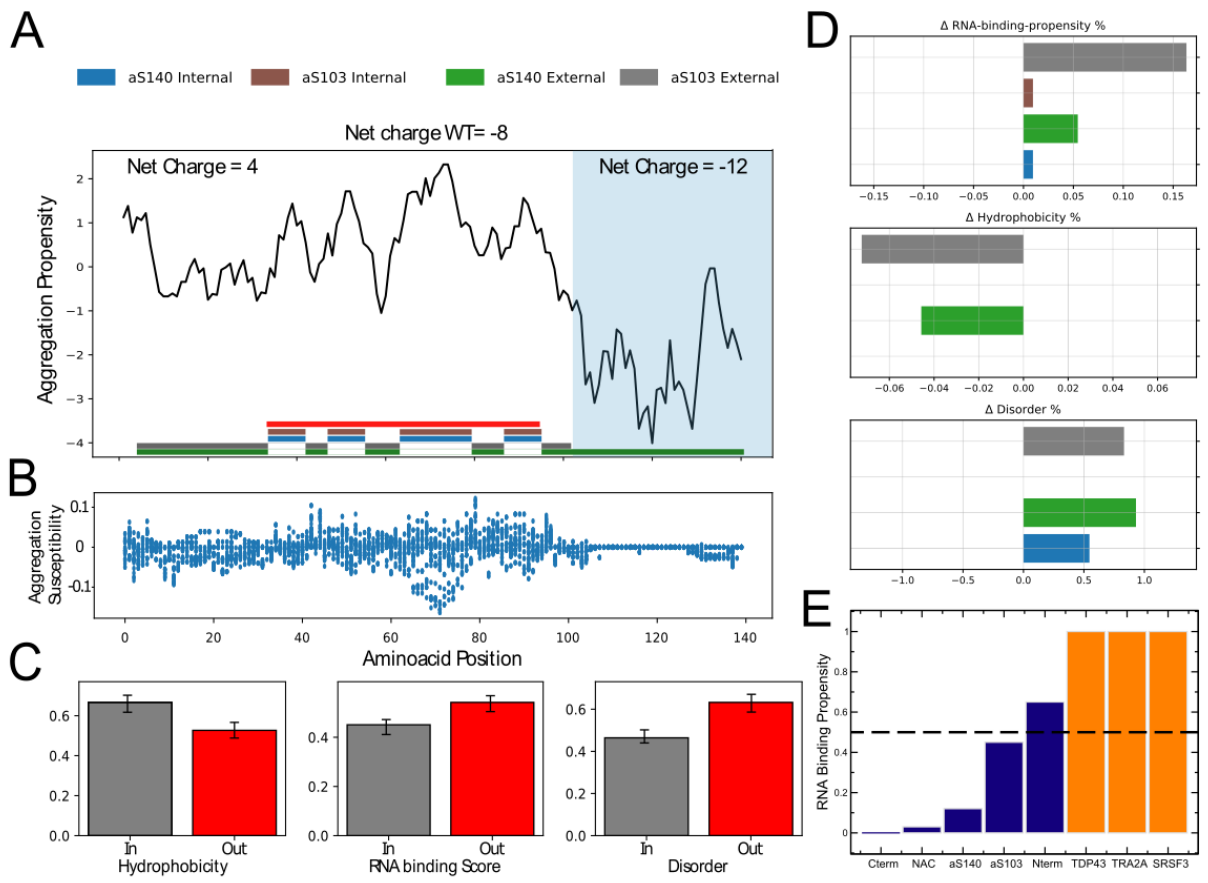


Figure 3.15. Aggregation propensity of aS140 and aS103. **(A)** Zyggregator profile of aS140 and aS103, with coloured bars indicating predicted inner and outer amyloid regions. The experimentally determined amyloid core of aS is marked with a red bar. Orange indicates minor Zyggregator peaks playing a role in modulating aS aggregation. **(B)** Random mutagenesis analysis showing that the NAC region and N-terminus of aS are susceptible to significant changes in aggregation propensity. **(C)** Physicochemical properties of inner (In, grey) and outer (Out, red) regions of aS140 and aS103. **(D)** RNA-binding propensity, hydrophobicity, and disorder of aS relative to its internal and external parts. Differences with respect to the mean of the amyloid histograms (**Fig. 3.14 A-C**) are shown. (**E)** RNA-binding predictions of different regions of aS140 and aS103, compared with canonical RBP (scores > 0.5 indicate non-negligible interaction propensities).

Mutations in the NAC region display a strong impact on aS140 aggregation (Fig. 3.15 B) and the outer regions of aS140 aggregates are predicted to be disordered, to display low overall hydrophobicity, and to be able to interact with RNA molecules (Fig. 3.15 C). Amino acids 67–82 appear to be particularly sensitive to change and dramatically affect the propensity of aS to aggregate, in agreement with previous experimental work[248]. Random single mutations in the C-terminus have instead negligible effects. Yet, 14 out of the 24 negatively charged residues in aS140 (18 Glu and 6 Asp) are found in this region, indicating that the truncated form of aS140 -namely aS103-, lacking all amino acids in positions 104–140, could behave differently compared to the wild-type protein in terms of aggregation propensity as well as RNA-binding ability.

Computing the scores of the physicochemical properties for both aS140 and aS103 reveals a strong difference between the two protein variants (Fig. 3.15 D). The removal of the 12 negatively charged residues in the C-terminal portion results in a net charge increment. Accordingly, the external regions of aS103 and aS140 show a lower degree of hydrophobicity compared to the average of the dataset and an increase in RNA-binding propensity (Fig. 3.15 D) due to the abundance of charged and polar residues outside the NAC region. We note that removing the C-terminus slightly affects the predicted structural disorder, which is consistent with the overall unfolded nature of the whole aS (Fig. 3.15 D). We next used the *cat*RAPID approach[195,197] to estimate the ability of aS140 and aS103 to interact with RNA. While the C-terminus and NAC show low interaction propensity, the N-terminal domain is predicted to be exposed to the solvent in the aggregate and available to interact with RNA (Fig. 3.15 E).

3.4.4 Summary and conclusions

In the study above a computational approach has been used to first discriminate between the amyloid core and the regions outside in aggregation-prone proteins. Then, these parts were analysed according to parameters that determine the RNA-binding propensity of the protein primary structure. After validation on examples of proteins with known amyloid structures, the analysis of a broad range of aggregation-prone proteins shows a clear trend of physicochemical properties discrimination. The amyloid core is tendentially hydrophobic and highly ordered while external regions, which are more disordered, display a distinctly higher propensity to interact with nucleic acids. Protein domains that function as nucleic acids-binding regions also tend to predominantly remain outside the amyloid core upon aggregation. aS is a great example of the trend, with the hydrophilic N- and C-terminal domains shown to remain outside the amyloid core and contributing to the separation of the physicochemical properties upon aggregation. The N-terminus also has a higher overall RNA-binding propensity that by itself could be considered significant compared to other RBPs. This is offset by the much lower propensity of the C-terminus in the monomeric protein, however upon aggregation and the partition of the physicochemical determinants, the propensity to interact with RNA increases in aS aggregates. This is further potentiated in the case of aS103, where the hydrophobic core structure is retained, and the lack of the C-terminal negative charge confer the aggregates a much higher overall RNA-binding affinity compared to the monomeric protein.

3.4.5 Closing remarks

The study presented above systematically addresses the role aggregation plays in determining the RNA-binding ability of a protein. The partitioning of the molecular RNA-binding determinants upon amyloid folding also explains very well the different trends in nucleic acid sequestration observed in aS140 and aS103 aggregation and presented in Section 3.3. It also gives further context as to why monomeric aS does not

exhibit a clear RNA-binding preference, yet RNA still plays a significant role in its aggregation. This point will be further addressed in the Discussion part of the thesis. Furthermore, the major implications of the study can be twofold. Firstly, the increase upon RNA-binding propensity upon aggregation could potentially lead to proteins that do not bind RNA as monomers to acquire a potentially toxic gain-of-function upon aggregation by means of RNA sequestration. Secondly, the trend that nucleic acid-binding protein domains tend to remain outside the amyloid core also implies that aggregation prone RBPs can preserve their RNA-binding function and even specificity upon aggregation, as already experimentally demonstrated in the work of Garcia-Pardo and colleagues[73]. This could lead to toxic RNA sequestration by the RBP aggregates formed in disease and contribute to the pathology.

The experimental results, presented in Chapters 1, 2 and 3, together with the background context of the computational work in Chapter 4, paint a clear image of the impact of RNA on the process of aS aggregation. RNA affects both the kinetics, the structural characteristics, as well as directly interacts with aggregates. In the last chapter, I will outline how all these observations can be linked together and how RNA can also directly influence the function of the aS aggregates.

3.5 RNA can modulate the enzymatic activity of aS amyloid assemblies

3.5.1 Identification of the putative catalytic site in aS fibrils

As outlined in the Introduction, amyloid fibrils can act as catalysts for various chemical reactions[34,249]. The unique arrangement of longitudinal repeats of amino acid sidechains with or without cofactors can in turn protonate or deprotonate amide or hydroxyl groups to act as the active site[95]. The mechanism is similar to the catalytic triads of nucleases or proteases, without a globular active site or pocket to achieve substrate specificity[250,251]. As mentioned, one of the most prominent catalytic markers is the histidine residue due to the unique chemistry of the imidazole ring that allows it to act both as a proton donor and acceptor at physiological pH[252]. aS contains only one histidine residue at position 50, which is found on the surface of most of aS fibril polymorphs[59]. Especially curious is the arrangement of the amino acid residues in the structure termed the rod 1a polymorph[86] (Fig. 3.16 A).

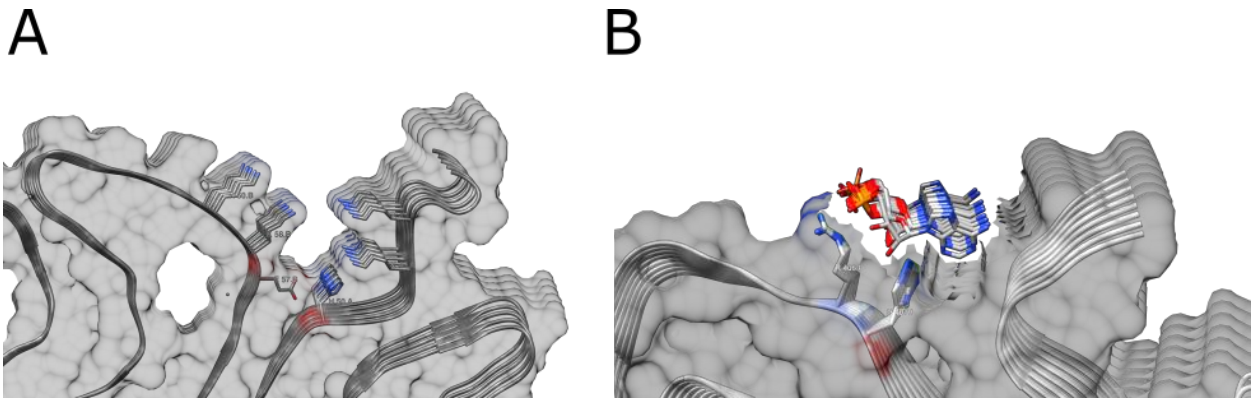


Figure 3.16. (A) Molecular and structural basis of potential catalytic activity of aS fibrils (depicted is the putative active site of rod 1a polymorph, PDB 6CU7). The arrangement of aminoacid residues around the longitudinal repeats of His50 shows a negatively (E57) and positively (K43) charged residue capable of either protonating or deprotonating the histidine residues, thus creating a reactive group. **(B)** The site of tau fibrils, grown in the presence of RNA, where a poly-G molecule was docked later during analysis (PDB 7SP1). The close spatial arrangement of histidine and arginine residues shows a degree of similarity to the putative aS fibril catalytic site.

This arrangement raised suspicion that aS fibrils could act as catalytic amyloids when (mis)folded in a favourable conformation that allowed for a chemical activation of the sidechains of certain residues, such as His50. Indeed, during the course of research thesis, aS fibrils were shown to be able to catalyse the hydrolysis of both esterase and phosphatase substrates, with the mutation His50Ala almost completely abolishing the enzymatic activity[98]. An intriguing aspect of this fact is, however, that apart from His50, the putative active site in the rod1a polymorph is formed mostly by positively charged residues of 4 lysine residues (Fig. 3.16 A). The overall structural conformation also reminded us of the putative binding site of RNA in the cryo-EM structure of tau fibrils, grown in the presence of RNA[232] (Fig. 3.16 B). This immediately posed a question,

whether RNA could bind to aS fibrils in a similar way and modulate their catalytic properties. Furthermore, the characterisation of the potential modulator mechanism of RNA would enable the independent estimate of the affinity of RNA towards the fibrils, verifying the estimates from the quantification of sequestered RNA.

3.5.2 aS fibrils have comparable catalytic activity to reported values

We decided to follow the protocol for the measurement of catalytic activity of aS fibrils established by Horvath and colleagues[98] and based on measuring the esterase and phosphatase activity on standardised, colorimetric substrates (see Materials and Methods). The reaction is based on the conversion of 4-nitrophenol (or para-nitrophenol, pNP) derivatives such as 4-nitrophenyl acetate (pNPA) and 4-nitrophenyl phosphate (pNPP) into pNP, a coloured compound with a peak absorbance at 410 nm. The reaction can thus be tracked by measuring the absorbance of the reaction mix at 410 nm (A_{410}), convert the A_{410} into pNP molar concentration and thus calculate the initial reaction rates as already outlined[97,98].

Firstly, we estimated the overall kinetics of the reaction, and the influence RNA might have on the kinetic parameters of pNP derivate hydrolysis. The fibrils grown exhibited both an esterase activity towards pNPA as well as phosphatase activity towards pNPP (Fig. 3.17 A-D). In order to determine the basic kinetic parameters of their enzymatic activity, the reaction was followed at different substrate concentrations. The initial velocities (v_0) at different concentrations of substrate were plotted against the substrate concentration in the Michaelis-Menten plot in order to determine the kinetics parameters of the reaction according to the analysis reported in literature[98]. While monomeric aS has not been shown to contribute to the enzymatic activity[98], we performed a control assay to determine that is indeed the case. As expected, there is no indication of substrate degradation in the presence of monomeric aS (Fig. 3.17 E,F).

The parameters obtained in the tests with pNPA and pNPP both show that the fibrils behave similarly to the values reported in literature[98] (Table 4).

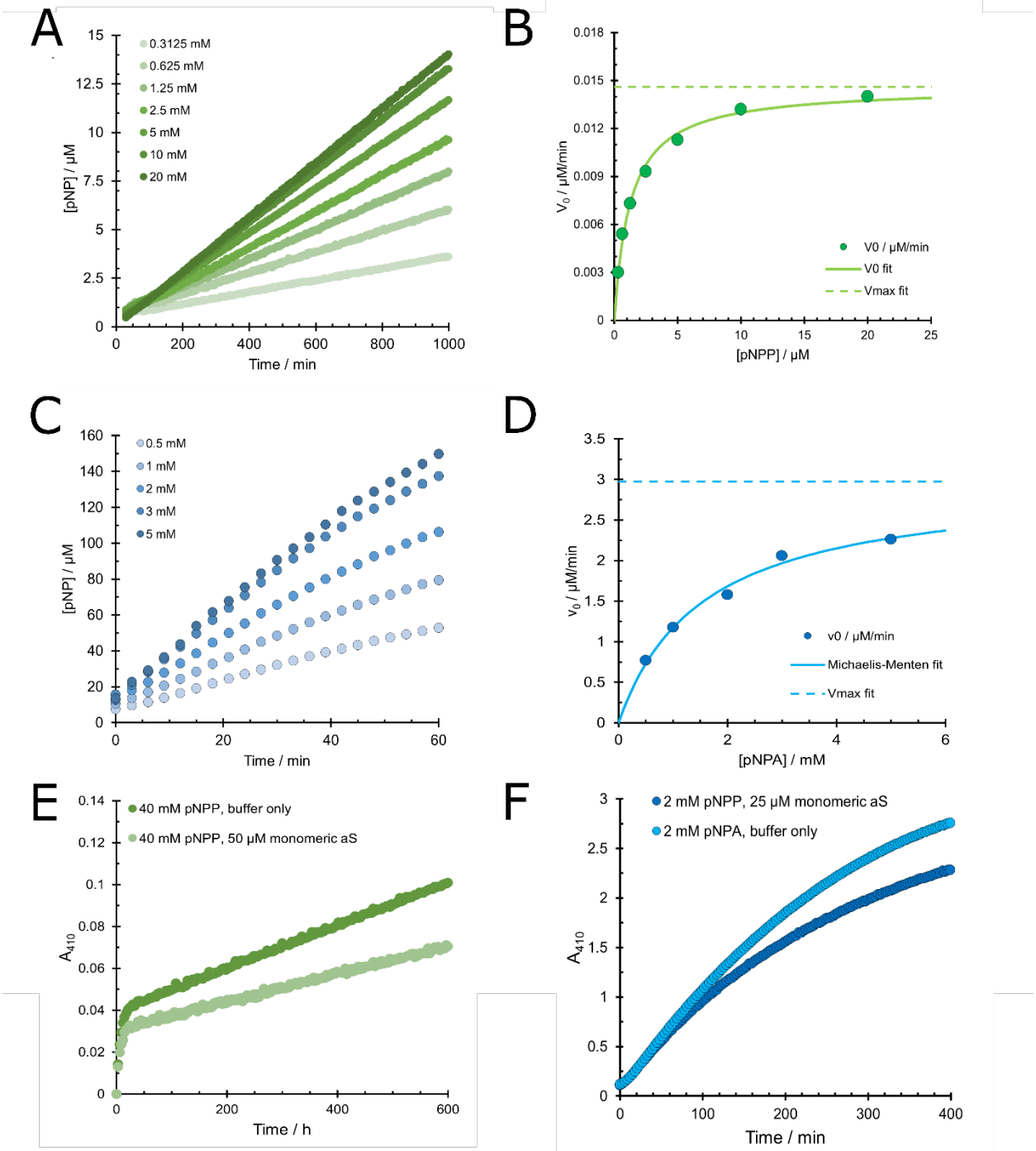


Figure 3.17. Enzymatic activity of aS fibrils. **(A)** 40 µM aS fibrils are able to degrade pNPP, a standard substrate used to test compounds for phosphatase activity. The activity is increased with increasing concentrations of pNPP. **(B)** Michaelis-Menten plot for aS fibrils and pNPP, together with the fitted curve and the kinetic parameters highlighted (K_M and V_{max}). **(C)** Esterase activity of 25 µM aS fibrils measured by pNPA degradation. Again, the enzymatic activity increases with a higher initial substrate concentration. **(D)** Michaelis-Menten plot for aS fibrils and pNPA. **(E,F)** Monomeric aS does not exhibit any enzymatic activity towards both pNPA and pNPP. The A_{410} values are seemingly negative due to the subtraction of control samples containing only substrate and buffer.

Table 4. Kinetic parameters of aS aggregate enzymatic activity compared to values reported in literature[98]. The values reported have been calculated from parameters, obtained in 7 (pNPP) and 3 independent experiments (pNPA). The fibrils were prepared freshly for each individual experiment (see Materials and Methods).

	Phosphatase (pNPP)		**Esterase (pNPA)**	
	Present work	Horvath *et al*	Present work	Horvath *et al*
K_M / *mM*	1.07 ± 0.72	0.5 ± 0.13	1.54 ± 0.44	4.3 ± 2.5
k_{cat} / s^{-1}	$0.9 \cdot 10^{-5} \pm 0.5 \cdot 10^{-5}$	$30 \cdot 10^{-5} \pm 10 \cdot 10^{-5}$	0.005 ± 0.002	0.012 ± 0.003
ε / $M^{-1}s^{-1}$	0.01 ± 0.006	0.6	3.2 ± 0.6	2.9

The parameters obtained in the tests with pNPA and pNPP both show that the fibrils behave similarly to the values reported by Horvath and colleagues (Table 4), especially considering enzyme kinetics notoriously differ from one laboratory to another[253]. Out of the parameters listed, the Michaelis-Menten constant (K_M) can be considered the most reliable comparison. K_M describes the behaviour of the enzyme at half-saturation with the substrate according to basic steady-state kinetics, which can be interpreted as the affinity of the enzyme towards the specific substrate[254]. Very importantly, K_M is not dependent on enzyme concentration[254]. This is especially significant in the case of aggregate enzymatics, where the aggregate population is a mixture of different molecular weight species and thus an exact molarity of the enzyme cannot be determined. The other two values reported, k_{cat} and ε, are used to describe enzyme behaviour at different substrate concentrations. The catalytic constant k_{cat}, also called the turnover constant, represents the number of substrate molecules that are converted to product per active site per unit of time when the enzyme is fully saturated with substrate. This reflects the behaviour of the enzyme when fully saturated with substrate at high concentrations[254]. The enzyme specificity constant ε, calculated by dividing k_{cat} with K_M, on the other hand describes the behaviour of an enzyme at low substrate concentrations. The k_{cat} values in our experiments are an order of magnitude lower compared to literature, while ε for pNPA is very similar and an order of magnitude lower for pNPP (Table 4). The discrepancy between these values likely comes from a different fibril preparation and quantification, since these values are calculated from the concentration of the enzyme[254].

3.5.3 RNA inhibits the catalytic activity of aS fibrils

The next step was determining how RNA can influence the catalytic properties of aS aggregates. Co-incubation with Increasing concentrations of RNA at constant substrate and fibril concentration drastically decreases the enzymatic activity towards pNPA (Fig. 3.18 A) as well as pNPP (Fig. 3.18 D). The initial velocity of the reaction decreases in what appears to be a hyperbolic trend, reaching saturation at very high concentration in the case of pNPA, while continually decreasing for pNPP. While the trend appears to be hyperbolic, it is very difficult to reliably establish a mechanism of inhibition for just the shape of the curve in this case.

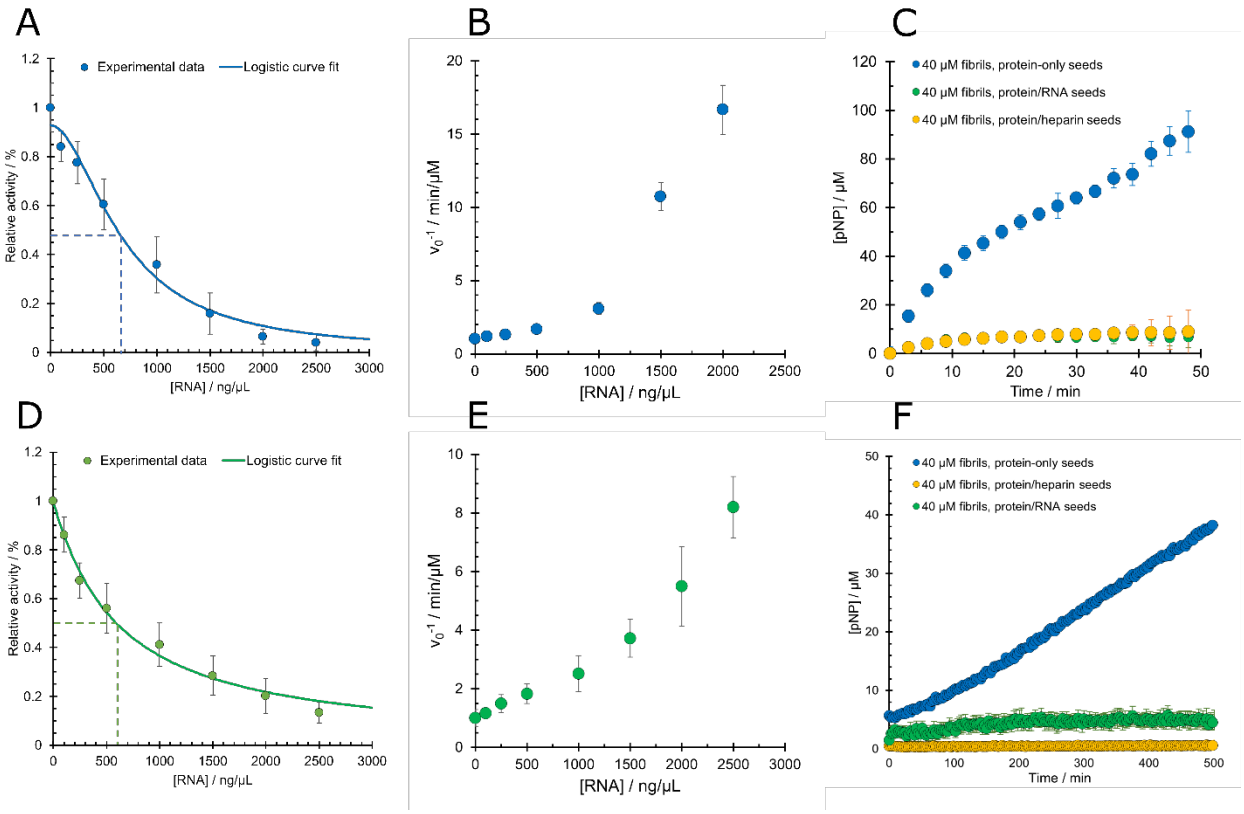


Figure 3.18. RNA inhibits the enzymatic activity of aS fibrils. **(A)** RNA acts as an inhibitor of αS esterase activity towards pNPA. The fit of initial reaction velocities with the logistic equation yields an apparent affinity constant $K_{0,5}$ of 680 ng/µL. **(B)** The Dixon plot for RNA reveals a parabolic inhibition mechanism, which occurs when a modulator binds the enzyme at both the active site and at additional sites with different affinities. (**C**) Fibrils, seeded by seeds grown in the presence of either RNA or heparin, show almost negligent esterase activity towards pNPA. (**D**) Similar to pNPA, RNA also inhibits the catalytic activity of αS fibrils towards pNPP. The fit with the logistic equation indicates a $K_{0,5}$ of 602 ng/µL. **(E)** Again, the Dixon plot reveals a hyperbolic/parabolic mechanism of inhibition, indicating both that the modulator binds at multiple binding sites, as well as the enzyme-modulator complex being catalytically active. **(F)** As with pNPA, coaggregation with RNA produces fibrils with very little catalytic activity towards phosphatase substrates.

Since the exact mechanism is more complicated than a simple linear inhibition, the curve was fitted with a logistical equation with four parameters as reported for the interactions of human elastase-2 with its modulators[255].This equation allows for an estimation of the apparent affinity constant ($K_{0,5}$) when the mechanism of the enzyme modifier is too complicated to be characterised precisely due to missing kinetic parameters, unknown number of binding spots, unknown allosteric mechanisms etc. Using this equation, the fit of the data revealed the $K_{0,5}$ for total RNA to be 684 ng/µL (pNPA) and 602 ng/µL (pNPP). As was the case with the RNA extracted from the aggregates in Chapter 3, this is an apparent constant due to unknown molarities of both the fibrils as well as RNA. The value is much higher compared to the aforementioned values which was very crudely established at 96 ng/µL, however due to the completely stochastic method it can be considered a better estimate of the actual binding of RNA to the fibrils. It also better reflects the reversible binding of the RNA on the fibril surface, since in the case of extracted RNA there is no way of knowing if the RNA is simply bound to the aggregates or actively sequestered. The latter would also imply an irreversible process and not a dynamic equilibrium. The fit also reveals a significant cooperative effect (h = 1,96 for pNPA), which reflects the complexion of the system observed. As does an attempt to elucidate the mechanism of modulation using the Dixon plot. The Dixon plot is used as a quick and reliable way of approximately determining the mechanism of the enzyme modulator from the shape of the curve $v_0^{-1}([I])$[254]. The plots for pNPA and pNPP confirms the highly complicated nature of the modulation, since the curve appears to suggest a parabolic mechanism (Fig. 3.18 B,E)[256]. Parabolic inhibition is a very rare mechanism in nature, reported for example in the case of modulation of human dipeptidyl peptidase by peptides derived from HIV Tat protein[257]. It occurs when a modulator binds the enzyme at both the active site and at additional sites with different affinities, with the binding completely excluding the substrate. Thus, while appearing linear at lower inhibitor concentrations, higher inhibitor concentrations exponentially decrease the catalytic activity[256]. This interpretation, however imprecise, suits the system observed. The total cell RNA likely contains a variety of fragments with different binding propensities that can individually bind the fibril with many different motifs at numerous sites on the actual fibril. Therefore, this approach of measuring the binding affinity of RNA cannot be considered a reliable method, yet it could still present a viable approach if a specific RNA or small molecule is used. This way, the exact mechanism of modulation could be determined by performing the experiments at different concentrations of the substrate and plotting the curves in order to properly estimate both the binding affinity as well as the exact mechanism of modulation.

3.5.4 RNA changes the morphology of aS fibrils to render them catalytically inactive upon coaggregation

While RNA undoubtedly binds the fibrils and thus inhibits their catalytic activity, the experiments presented above were always performed with fibrils seeded from protein left to aggregate by itself. Since the results presented in the preceding chapters clearly indicated that coaggregation of aS and RNA changes the conformation of the aggregates, the question remained whether these were in fact catalytically active. In order to determine that, the seeds were prepared by coaggregation with both RNA and heparin as a control (Materials and Methods). Afterwards, the fibrils were washed several times with buffer in an attempt to remove as much of the RNA and heparin bound to the aggregates as possible. The aggregates were then sonicated, washed again and finally added to fresh monomeric protein without any further RNA or heparin. These fibrils exhibit significantly less catalytic activity compared to the control condition, prepared in the standard way (Fig. 3.18 C,F). This indicates that coaggregation with RNA changes the structure of the fibrils drastically enough to almost completely abrogate their catalytic properties.

4. Discussion

The results presented in this work clearly indicate that RNA has a significant and multi-faceted effect on aS aggregation. To briefly recap, the results shown in Section 3.1 show that RNA can significantly alter the aggregation kinetics of aS. Section 3.2 shows that RNA can influence the structural conformations during the process of fibrillation and the mechanical properties of the final fibril population. Section 3.3 demonstrates that aS aggregates can sequester RNA in protein-sequence dependent manner and that aggregation is the driver of the process. Chapter 4 gives a computational and wider background on how aggregation can influence the misfolding of proteins and thus lead to them acquiring novel functional properties such as RNA binding. Finally, Chapter 5 shows unequivocally that RNA does bind the fibrils by directly inhibiting their enzymatic activity, as well as changes the structural conformation of the mature fibrils by changing the conformation that confers the fibrils their catalytic properties.

What these results do not provide, however, is a mechanism or a molecular-level mode of action. Why and how does RNA act to influence the trends in the way that it does, and could that be significant?

As outlined in the introduction, the current stand on the actual trigger mechanism of aS aggregation lies in the interruption of the intramolecular contacts between the N-terminal and the C-terminal domain[153,155,258], that can lead to the formation of the aggregation-prone partially-folded intermediate[173]. The explanation for the role of RNA is therefore rather simple and straightforward. By interaction, albeit transiently, with the N-terminal region, RNA likely displaces the C-terminal tail and increases the conformational flexibility of the NAC region and the potential formation of the partially folded intermediate. This resembles the one-step nucleation mechanism triggered by anionic lipids[177]. In this interpretation, RNA is acting almost like a molecular mimetic of the C-terminal domain and increasing its concentration should make the aggregation of the wild-type aS140 behave in a way that resembles the C-terminally truncated aS103. Our results confirm this, as the increasing concentration of RNA does drastically increase the aggregation rate, with the peak values reaching a plateau strikingly similar to the rate of aS103 without RNA (Fig. 3.3). We can therefore conclude that increasing the starting concentration of RNA leads to an increased perturbation of the contacts between the N-terminus and the C-terminus, until at around 250 ng/μL, the concentration of RNA reaches equimolar conditions compared to the C-terminus and the aggregation rate increase effectively plateaus. The C-terminus of aS has also been shown to inhibit fibrillation and promote oligomerisation[155,171]. This is reflected in the AFM data, showing a much faster appearance of fibril-like structures in the presence of RNA (Fig. 3.8). These fibrils also have different mechanical properties, especially a higher rigidity and a lower diameter (Fig. 3.8), which is precisely what has been observed in electron microscopy analysis of fibrils produced by C-terminally truncated aS

constructs[178]. However, the fact that fibrils are more compact and rigid does not mean that they're mechanically less prone to breaking[259]. This could in turn mean the increase in the aggregation rate observed for aS140 in the presence of RNA is partially caused by the much higher rate of fragmentation in the experimental setup used for aggregation assay. The unfortunate downside of using a system with both continuous shaking as well as a glass bead in the well is the likely fragmentation of aggregates, induced by the shearing forces[260]. The method of fitting the data acquired in this way is also not as precise and informative as other approaches which allow for the direct identification of various individual rates[93,179]. These models require a detailed knowledge of the parameters and additional measurements that were unavailable to us such as consistent AFM or EM imaging of aggregates to determine the amount of monomeric units in a fibrils[17,93]. Another widely used option is the Finke-Watzky equation[261], which describes a two-step model with individual rates for nucleation and elongation. Importantly, however, it does not take into account the secondary nucleation, one of the critical steps in the aggregation of aS[93,179]. Overall, further experiments are needed to clarify exactly which is the process that RNA affects the most, likely with a different experimental approach that eliminates the variable of fragmentation and the resulting altering of kinetics.

The postulation that RNA effectively displaces the C-terminal domain and increases the aggregation rate by promoting the partially folded intermediate formation holds well in view of the results obtained with the full-length aS140. However, it does not provide an explanation as to why RNA has a similar scaling effect on aS103, which lacks almost the entirety of the C-terminus. The IR results show a slower increase in $(\beta\text{-sheet})_{P,AP}$ in the presence of RNA compared to aS103 (Fig. 3.6). Also, the fluorescence data indicate that RNA overall decreases the aggregation rate at 50 ng/μL if the slower first step is not considered, as well as decreases the fluorescence plateau (Fig. 3.3). Since the N-terminus does not change chemically, RNA still likely interacts transiently with it, just like in the case of aS140, and this interaction could similarly favour the primary nucleation step via the partially-folded intermediate. However, Van der Wateren and colleagues found out that in the case of aS103, secondary nucleation is the dominating process which reflects in the much higher aggregation rate compared to aS140[171]. This could mean that while RNA speeds up the primary nucleation step to, it blocks or slows down the secondary nucleation step. This hypothesis neatly explains the very different shape of the fluorescence curve in Fig. 3.3B when comparing 0 and 50 ng/μL. Increasing additions of RNA then further stimulate the primary nucleation as seen by the increasing aggregation rate, until it reaches the values more or less similar to the secondary nucleation-dominated aS103 in the absence of RNA. A possible explanation for this could be the increased propensity of interaction with RNA by the aggregated form of aS103, whereas the RNA bound or sequestered in the aggregated form prevents the interaction between the monomers and the fibrils and thus blocks secondary nucleation. If so, this

could present a very important discovery in the development of RNA-based therapeutics that limit the extent of protein aggregation in cells.

The mechanism of fibril-induced secondary nucleation in the case of aS has been explored by Kumari and colleagues[262]. They showed that upon aggregation, the ordered C-termini forming the fuzzy coat of aS fibrils transiently interact with the N-termini of the monomers in solution. This then enables the nucleation of the monomers in the same way as negatively charged polymers can stimulate primary nucleation, by destabilising the intramolecular contacts in the monomeric aS. The addition of RNA in this model can thus block the monomer-fibril contact in two ways. The first is by competing with the fibril for interacting with the monomer N-terminus, and the second is by actively interacting with the fibril and limiting the available binding spots for the monomer. Our hypothesis works by assuming that a.) the interaction propensity between the N-terminus of the monomer and the RNA is stronger than the interaction between the N-terminus of the monomer and the C-termini of the fuzzy coat, and b.) The interaction between the aggregated aS and the RNA is stronger than between the aggregated and monomeric aS. The latter assumption can be quickly confirmed by the estimated binding propensities, obtained by enzymatic inhibition assays. The $K_{apparent}$ for fibril-RNA interaction can be roughly estimated at around 600 ng/µL, meaning at the working RNA concentration of 500 ng/µL, roughly 40 % of the available binding spots on the fibrils are occupied by the RNA. Kumari and colleagues estimate the K_D for fibril-monomer interaction to be roughly 1 mM[262], meaning at the working monomer concentration of 50 µM and the fibril concentration much lower, the affinity for the monomer is much lower than the affinity for the RNA. The former assumption can, however, only be evaluated indirectly from the calculated aggregation rates. As mentioned above, at the RNA concentration of 250 ng/µL, the aggregation rate of aS140 is comparable to the aggregation rate of aS103 alone. This likely means that around the same concentration, RNA is likely the preferred interaction partner of the N-terminus. In any case, the fulfilment of just one of these requirements can still effectively abrogate the secondary nucleation by preferentially interacting with the fibril compared to the monomer.

This fact further highlights the importance of the fibril-RNA interaction, which have so far remained very much understudied. One reason for this could be the simple fact that so far, no atomic-level structures of protein fibrils have been solved with the RNA directly bound. There is the structure of tau, grown in the presence of RNA, recently solved by cryo-EM by Abskharon and colleagues[232], yet the RNA has been docked into the structure in the post-processing phase. The closest equivalent in the case of aS is the structure of heparin-remodelled aS fibrils, recently published by Tao and colleagues[183], where a 6-mer heparin polysaccharide has also been docked into the structure post-processing. Still, these are all structures of fibrils originating *in vitro*, where isolated conditions may facilitate the interaction more easily. However, the overview of all the 86 solved structures of aS fibrils from AmyloidAtlas[263,264] reveals an interesting fact. Out of all the

structures deposited in the curated database, only a couple originate from the post-mortem tissues. One is the structure of fibrils from a patient with Lewy Body Dementia[57] and the other from a patient with Multiple System Atrophy [58] (Fig. 4.1).

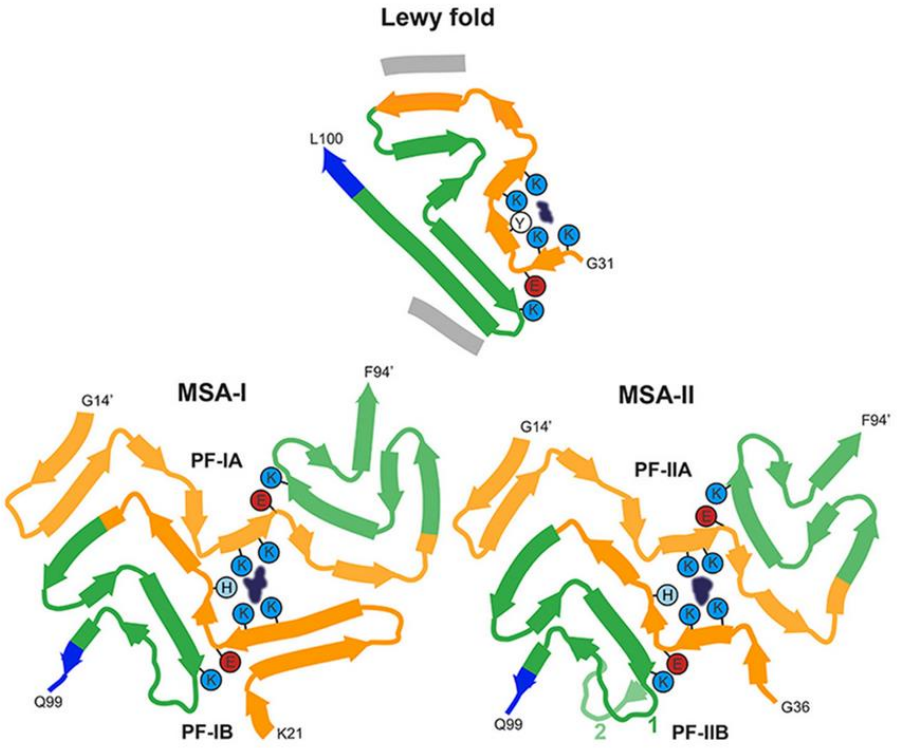


Figure 4.1. Comparison of the structures of amyloid fibrils extracted from a patient with Lewy Body Dementia[57] (PDB 8A9L) and Multiple System Atrophy[58] (PDB 6XYP and 6XYO). Both structures show unknown electronic densities that interact with positively charged pockets both on the fibril surface (Lewy fold) and in the fibril core (MSA fold).

Curiously, both structures clearly contain unassigned electron densities whose origins the authors could not determine. These densities appear bound to the fibril surface of the LBD fold, with the interaction stabilised by four lysine residues and a tyrosine, as well trapped in the core of the fibril and again bound by the four lysine residues and the one histidine. This immediately indicates the unknown ligand is likely negatively charged and hydrophilic, fitting the description of RNA. Furthermore, the ligand binding site on the fibrils is located precisely where monomeric aS was predicted to have the highest propensity of interacting with its own transcript[127]. Moreover, these fibrils have a unique fold amongst all the solved structures, further implying that RNA has the potential to alter the amyloid structure in a new and unique way. While determining the structure of aS fibrils as grown in the presence of RNA would be absolutely the best way to address all of these questions, it was unfortunately out of reach for the scope of this project.

Due to the diversity of fibrils structures solved so far, also in the presence of various cofactors, it would be reasonable to assume that total cell RNA such as the ones used in our experiments could induce the formation of a population of structurally diverse fibrils. In fact, that could be a fascinating experiment, studying the effect of structurally and compositionally diverse RNA on the amyloid fold of aS. This also further highlights the question that was left unaddressed so far: what happens with the RNA? The RNA used in the experiments was a total yeast extracted RNA, which is largely fragmented and represents a good average population of different motifs and structures. This is precisely why it was chosen for the experiments, to avoid any sort of bias by a certain overrepresented RNA species. However, it makes the RNA unusable for any sort of sequencing applications due to the low quality and also the lack of suitable kit for the preparation of sequencing libraries. There is also the possibility that fibrils themselves can affect the RNA bound to them. Enzymatic tests conducted with the fibrils showed that they are catalytically active both towards esterase as well as phosphatase substrates. RNA itself as a molecule contains both types of these chemical bonds, which could potentially be cleaved by the catalytically active fibrils. In this case, what appears to be competitive inhibition by means of binding to the catalytic site could actually be an apparent inhibition by means of RNA representing a competitive substrate. Still, competitive inhibitor or competitive substrate, this does not change the fact that RNA is directly bound by aS fibrils. In fact, it could add another potential mechanism of cytotoxicity to the aS fibrils by degrading RNA and potentially altering the transcriptome in affected cells. For now, there is no linking catalytic activity to *in vivo* determined structures of aS fibrils. The two structures of aS fibrils from the post-mortem brain for example do not exhibit the catalytic motif mentioned in Chapter 5 of the Results. The key residue, His50, is actually seen directly interacting with the unknown moiety in the core of the fibril in the case of the MSA fibril and is not solvent exposed (Fig. 4.1). Furthermore, the fibrils grown from the seeds aggregated in the presence of RNA do not exhibit the same catalytic activity, indicating that while the subject of enzymatically active fibrils remains fascinating, it currently lacks any evidence it occurs *in vivo* as well.

In general, triggering aS aggregation in a cellular model has remained a challenge[265] and there is no available structure of aS fibrils isolated from cells. This field, however, changed rapidly and started evolving right during the time of the research for this thesis with the discovery made by Ray and colleagues that aS undergoes both liquid-to-liquid and liquid-to-solid phase separation[266]. The chapter of Discussion was generally meant not to directly comment on the experimental results, but to give them a wider context, and the last part will be dedicated precisely to this new direction the research is currently evolving. Biological phase separation has come to the forefront of life science research in the last 15 years, both as an essential functional and organisational process in cells[267] as well as potential precursor state leading to conformational and morphological states, associated with disease and protein aggregation[268]. aS has been shown to undergo phase separation both *in vitro* as well as *in vivo*[266,269]. Furthermore, it has been shown

that these condensate can mature into amyloid assemblies[270] and the kinetics of these processes have been determined[41]. Especially of interest is the work by Dada and colleagues, which show that the actual kinetic model of aS aggregation inside the condensate is not drastically different from what occurs in non demixed conditions[41]. Still, the initial nucleation step is required to trigger the condensation, yet unlike the solution-like conditions, the aggregation is then dominated by the secondary nucleation. Importantly, aS fibrils have been shown to be uptaken into aS liquid-like condensates and can thus initiate the liquid-to-solid phase transition via secondary nucleation. The major difference between the model used in our research and the condensate one is the huge difference in local concentrations. Dada and colleagues have estimated the approximate concentration of aS to reach around 30 to 40 mM inside the condensate. This is approximately an order of magnitude higher than the concentration used in our experiments. This fact, together with the crowded conditions, might strengthen interactions found transient and insignificant in our system, such as the one between monomeric aS and RNA. A very intriguing result indirectly confirming this has been the recent discovery by Dhakal and colleagues that aS can modulate the internal organisation of condensates composed by TDP-43 low-complexity domain, aS and RNA[271]. aS has been shown to form clusters on the outer rim of the condensate and sequestering the RNA on the surface of the condensate, thereby promoting TDP-43 LCD aggregation. An N-terminally truncated aS constructed did not show this sequestration ability, indirectly confirming the partitioning of RNA inside these condensates is mediated by the N-terminal domain of aS. These results confirm that aS-RNA interactions can be completely re-contextualised in the condensate system, where the very high local concentrations might render otherwise insignificant interaction very important. Unfortunately, Dhakal and colleagues do not comment on the extent of aS aggregation in this system. It would be incredibly interesting to see the effect of aS aggregation of this partitioning of RNA inside the condensate. In accordance with our computational model, the increase in RNA-binding propensity by aS aggregation might further sequester RNA and prevent it to interact with TDP-43 LCD, thus abolishing the protective effect of RNA and increasing its aggregation. On the other hand, the sequestration of RNA by the initial aggregates of aS might prevent any further exacerbation by blocking the secondary nucleation process as in the case of aS103.

While this is all speculation at the current moment, the work of Dhakal and colleagues can potentially be viewed as indirectly confirming and validating the results of the research work presented in this thesis. The increase in RNA-binding propensity by amyloid formation can be viewed as a process that leads to aggregation-prone proteins acquiring novel functional properties. This sequestration of RNA can drastically influence the aggregation kinetics by modulating the primary nucleation and potentially blocking secondary nucleation by way of RNA binding to the surface of the fibrils. This increase in RNA-binding propensity can also drastically alter the internal dynamics of a condensate by sequestering RNA, often one of the key components contributing to the

liquid-like properties, especially in the case of multi-component systems with more than one aggregation-prone protein. This could further give context to one of the less prominent results of the computational analysis, presented in Section 3.4. Namely, the fact that nucleic acid-binding domains are shown to predominantly be located outside the amyloid core upon aggregation (Fig. 3.14 I-K). This could present an evolutionary "safety valve", assuring that a protein domain with a higher propensity for RNA binding is in close proximity to the amyloid core to prevent it from (un)specifically sequestering RNA upon aggregation. Or, from an inverse perspective, ensure that there is always an RNA molecule present to the highly aggregation-prone prion regions to modulate and prevent its misfolding. This interpretation can be thought of as complementary to the previous work of Gotor and colleagues, showing the interplay between RNA-binding domains and prion domains that leads to the modulation of phase separation[126], with the "soluble" prion domain replaced by the amyloid core.

Overall, the observations based on the results of the research work presented in this dissertation provide new insights into the modulation of aS aggregation by RNA, a field almost completely unexplored until now. However, some interpretations based on the framework presented here can reach far beyond aS. The dynamic structure-function relationship has so far been limited to the study of globular proteins, yet our results show that any amyloid assembly can be treated as a functional fold, with its own binding or catalytic properties. This can in turn have potentially far-reaching implications for the comprehension of protein behaviour, emphasising the intricate interplay between aggregation, interactional and functional diversity across the biological landscape.

5. Conclusions and future perspectives

Based on the results presented and with the original aims of the project in mind, the following conclusions can be drawn.

We have successfully optimised the protocol for the *in vitro* characterisation of aS aggregation in the presence and absence of RNA. We have determined by way of time-resolved fluorescence with Proteostat™ and micro-FTIR measurements that total yeast RNA accelerates the aggregation of both aS140 and aS103 by increasing the rate of aggregation and decreasing the aggregation lag time. We interpreted the results of aS140 by means of perturbation of aS intramolecular contacts by RNA interacting with the N-terminus, promoting the partial folding and increasing the primary nucleation rate. This was especially visible in the case of aS103, where the accelerated primary nucleation overtook secondary nucleation as the primary contribution towards the aggregation process.

We have also determined that RNA influences the structural characteristics of aS aggregates and alters their conformation. AFM-IR and micro-FTIR results have shown RNA promotes the growth of thinner, more compact, and more rigid fibrils. We have also determined that RNA alters the intermediate aggregation pathway by promoting early protofibril growth in place of the slow interconversion of oligomeric species. Enzymatic activity tests have shown that fibrils amplified from seeds grown in the presence of RNA do not exhibit any catalytic activity compared to fibrils amplified from protein-only seeds. Since the catalytic activity of aS fibrils is highly conformation-dependent, this directly shows an altering of the fibrils structure by coaggregation with RNA.

Finally, we have determined that aS aggregated species both directly bind RNA, as well as sequester it during coaggregation. We have determined the higher affinity of aS fibrils for RNA compared to DNA by means of quantifying the nucleic acids extracted from aggregates after aggregation. We have also shown directly that RNA binds to the fibrils by inhibiting their enzymatic activity and we have determined the approximate binding affinity.

In addition, we have computationally determined that aggregation leads to an increase in the aggregation propensity of proteins by partitioning of physicochemical properties in the inner and outer parts of the amyloid core. We have also shown that RNA- and DNA-binding domains tend to remain outside the amyloid core and can thus also contribute to specific RNA sequestration by canonic, aggregation-prone RBPs. We have also verified the computational analysis with the experiments conducted with aS.

Overall, the idea of the project was to evaluate aS as a potential target for the development of RNA-based molecules for specific targeting and modulation of its

pathological processes. Based on the results, we can conclude that RNA does play an important role and has a high potential as a modulator not only of aS aggregation, but of other aggregation-prone proteins in general.

The results of the study can of course be considered preliminary, and a lot more work needs to be done in order to further clarify all the aspects of the RNA-based modulation. We are still conducting the AFM-IR experiments on the aS103 to determine whether RNA has a similar effect on the mechanical properties of its fibrils. Also, the whole set of enzymatic assays needs to be finalised for aS103 to determine whether the inhibition in the presence of RNA is in fact stronger and a consequence of the predicted higher affinity of aS103 fibrils towards RNA. This will complete the foundations of the project, outlined in this thesis.

An obvious next step would be to further investigate the kinetics of the influence of RNA on aS aggregation, specifically the effect of RNA on the secondary nucleation. This could be easily performed by conducting aggregation assays in the presence and absence of fibrillar seeds. Also, a potential alternative to the experimental setup with glass beads and constant agitation could be considered in order to evaluate the effect that fragmentation has on the kinetics of aS aggregation in the presence and absence of RNA. However, to fully evaluate that, a more complex kinetics model needs to be taken into account to properly quantify these specific processes. Additionally, with the information that we possess about the importance of fibril-RNA interaction, a completely new kinetic model could be developed that incorporates these interactions and the way they modulate the individual processes. An example would be the introduction of a fibril-RNA binding constant and a monomer-RNA binding constant into the secondary nucleation equation to account for the interference of RNA into this process. An earlier attempt at such a mechanism can be traced at the older versions of the manuscript on bioRxiv[272], however with all the new information it should be expanded and polished.

The key question after demonstrating that aS fibrils sequester and bind RNA is which kind of RNA get sequestered and whether a certain preference exists in terms of motifs, structures etc. This could be crucial also for the identification of potential natural aptamers, e.g. sequences that have an especially high interaction propensity with the fibrils. As already mentioned, the use of total yeast RNA prevented any attempts for sequencing of the original samples due to a lack of material available for sequencing libraries preparation. To address this, we have performed a similar experiment with total RNA, extracted from human-derived cells such as HEK293T and SH-SY5Y. We have extracted the RNA sequestered by the aggregates during coaggregation, as well as the RNA bound to the preformed fibrils in order to determine the potential differences in the types of RNA bound. We are currently awaiting the results of the sequencing, hence the omission of these experiments from the thesis, however we have high expectations for

the outcome of these experiments. The identification of enriched motifs would enable us to study how these perturb the aggregation of aS on the model already established. Also, the use of a specific RNA with a defined molecular mass could also allow the precise calculation of all the enzymatic parameters and determine an actual molar affinity constant, as well as a correct inhibition mechanism.

As already hinted at in the Discussion, another promising future goal would be the characterisation of the role of RNA, aggregation and subsequent processes in the liquid-liquid and liquid-solid phase transitions of aS. We have also already initiated work in this way with our collaborators, using the technique of fluorescence correlation spectroscopy (FCS)[273]. These experiments will enable us to determine the molecular dynamics and mobility of fluorescently-tagged aS both in the diluted, as well as the dense phase over time. This should allow us to determine the role RNA plays in the formation, dynamics and maturation of aS condensates and compare the results to the ones already obtained with the methodology described in this work. An even further goal, in combination with the sequencing experiments, would be to further extract any RNA entrapped in these solid-like condensates, perform RNA sequencing, and then compare them to the results obtained via "classic" amyloid aggregation. The comparison could further explore the roles different types of RNA have when it comes to different phase transitions and amyloid structures.

Other, more distant and less concrete goals and future perspectives include the characterisation of aS aggregation and the potential role RNA plays in it in a cellular model, either by extracting the cellular RNA bound to the aggregated species of aS or determine the transcriptomic changes in the cells upon aggregation. Another one, mentioned at several points during the Discussion, would also be to attempt to determine the atomic structure of aS aggregates, formed in the presence of RNA, by cryo-EM. A high-resolution structure would provide invaluable insight into the effect RNA has in perturbing the conformations of aS during aggregation, as well as comparing the structure to the existing ones and perhaps even recognising RNA as one of the unknown electronic densities.

A final, more conceptual approach would also be to follow upon the computational predictions and attempt to quantify the changes in RNA-binding propensity not only for aS, but also for other amyloidogenic, non-canonic RBPs. Or assess the changes upon RNA binding patterns before and after aggregation for canonic RBPs, similar to the experiment performed for hnRNPDL-2[73]. We have been proceeding ourselves in this direction as well, by evaluating the RNA-binding propensity changes for Aβ1-40, with potential for many other proteins tested with the methodology outlined by this thesis.

6. Materials and Methods

6.1 Preparation of bacterial expression vectors for aS

The plasmid pET21a containing the sequence of wild-type human aS was obtained from Professor Annalisa Pastore (Fig. 6.1 A). The plasmid allows for the expression of recombinant aS without any tags or labels via the T7 expression system in bacterial cells. The expression vector was used as such.

The plasmid for the expression of aS103 (Fig. 6.1 B) was prepared by deleting the 111 bp of the original vector coding for amino-acid residues 104-140 via polymerase chain reaction (PCR) using the Q5 polymerase (New England Biolabs, USA). Table 5 shows the oligonucleotides used and the reaction conditions. The resulting product was treated with DpnI (New England Biolabs, USA) for 30 min at 37 °C to degrade the matrix DNA and subsequently purified with the Monarch PCR & DNA Cleanup kit (New England Biolabs, USA). The clean amplicons were then ligated with QuickLigase (New England Biolabs, USA) at room temperature for 15 min and immediately used to transform competent *Escherichia coli* DH5α cells (New England Biolabs, USA). Positive colonies were tested for the correct insert via PCR and Sanger sequencing, using the T7_for and aS103_rev oligonucleotides (Table 6).

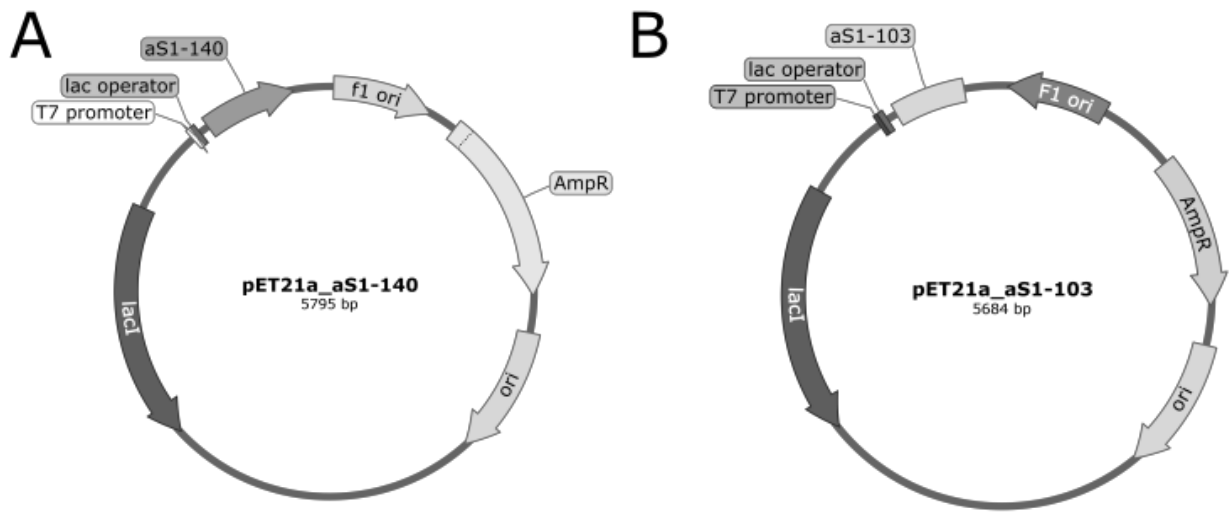


Figure 6.1. Plasmid maps for the pET21a_aS1-140 (**A**) and pET21a_aS1-103 expression vectors (**B**).

Table 5. List of oligonucleotides used in the preparation of aS103 and validation of the correctness of the insert.

Oligo label	Oligo sequence (5'→ 3')
aS103_for	TAAGCGGCCGCACTCGAGCAC
aS103_rev	ATTCTTGCCCAACTGGTCCTTTTTGACAAAGC
T7_for	TAATACGACTCACTATAGGG

Table 6. Conditions of the PCR reaction used to shorten the aS cDNA for the preparation of aS103. The thermocycler was set to 32 cycles.

Temperature	Time	
98 °C	10 s	
98 °C	10 s	x 32
68 °C	20 s	
72 °C	30 s	
72 °C	120 s	

6.2 Bacterial expression of recombinant aS

Competent *E. coli* BL21 [DE3] cells (Thermo Fisher Scientific, USA) were transformed with both constructs according to standard protocols. The pre-inoculum was grown by adding 500 µL of cryo-preserved bacteria in 50 % glycerol/LBA (Merck KGaA, Germany) into 50 mL of autoclaved LB medium (Merck KGaA, Germany) supplemented with 50 µg/mL ampicillin (Merck KGaA, Germany). The culture was grown overnight at 37 °C at 160 rpm. For full-scale expression of aS, 2 L of autoclaved 2xYT broth (Thermo Fisher Scientific, USA) supplemented with 50 µg/ml ampicillin in a 5 L erlenmeyer flask were inoculated with overnight cultures in a 100:1 ratio. Cultures were grown at 37°C up to an optical

density of ca. 0.8 at 600 nm. Protein expression was induced with 0.5 mM IPTG (Thermo Fisher Scientific, USA) and the cultures were shaken for 4 h at 37 °C at 180 rpm. Cells were harvested by centrifugation at 4500 g at 4 °C for 30 min, afterwards the pellet was resuspended on ice in 20 mM potassium phosphate buffer pH 7.2 (Merck KGaA, Germany), frozen immediately and kept at - 20 °C until use. The expression conditions were identical for both protein constructs.

6.3 Isolation and purification of aS constructs

For protein isolation the frozen cell suspensions were thawed in cold water and lysed by boiling at 95 °C for 30 min. The boiled suspension was transferred on ice for 20 min, then centrifuged at 18 000 g at 4 °C for 60 min. The supernatant was transferred into a glass beaker and powdered streptomycin sulfate (Merck KGaA, Germany) was added with constant stirring at room temperature up to a final concentration of 10 mg/mL in order to precipitate soluble nucleic acids. After the addition of all the powder, the beaker was transferred to a magnetic mixer at 4 °C for 20 min to further stimulate nucleic acid precipitation. The solution was then centrifuged at 18 000 g at 4 °C for 30 min and the pellet containing precipitated impurities was discarded. The supernatant was again transferred to a fresh glass beaker and the protein was slowly precipitated by gradual addition of crystalline ammonium sulfate at constant stirring up to a final concentration of 360 mg/mL at room temperature. The beaker was then moved to the 4 °C and stirred for a further 20 min. The suspension was subsequently centrifuged at 18 000 g at 4 °C for 30 min. The supernatant was discarded, and the pellet was resuspended in 20 mM Tris–HCl pH 7.4 (Merck KGaA, Germany) with 400 mM KCl (Merck KGaA, Germany) and dialysed against 20 mM Tris–HCl pH 7.4 overnight at 4 °C. The isolation protocol used was identical for both protein constructs.

aS140 was purified using ion exchange chromatography using a HiTrap™ Q FastFlow column (Cytiva, USA) connected to the ÄKTA pure™ chromatography system (Cytiva, USA). The dialysed protein was loaded directly onto the column equilibrated in 20 mM Tris–HCl pH 7.4 at the flow of 5 mL/min, washed until absorbances of UV light at 280 (A_{280}), 260 (A_{260}) and 230 nm (A_{230}) returned to the baseline values. The flow was reduced to 1 mL/min and the protein was eluted by gradual addition of 20 mM potassium phosphate pH 7.2, 800 mM NaCl (Merck KGaA, Germany). The gradient used was 16 mM NaCl/min at the flow 1 mL/min with the aim of reaching 800 mM NaCl within 5 column volumes (CVs) or 50 mL. The eluted protein was pooled and immediately loaded onto a HiLoad™ 16/600 Superdex™ 75 gel filtration column (Cytiva, USA) equilibrated in 20 mM potassium phosphate pH 7.2, 100 mM KCl (Merck KGaA, Germany). The separation was followed by measuring A_{280} and A_{230} and the elution peak was pooled, aliquoted and frozen immediately. An aliquot was kept for protein concentration and purity determination via BCA assay, spectropolarimetry and SDS PAGE.

aS103 was purified via heparin affinity chromatography on a HiTrap™ Heparin column (Cytiva, USA) connected to the ÄKTA pure™ chromatography system. The dialysed protein was again loaded directly onto the column equilibrated in 20 mM Tris–HCl pH 7.4 at the flow of 5 mL/min and washed until the UV absorbance values returned to baseline values. The protein was eluted by gradual addition of 20 mM potassium phosphate pH 7.2, 1 M NaCl (gradient of 20 mM/min at the flow of 1 mL/min). The eluted protein was again pooled and immediately loaded onto the HiLoad™ 16/600 Superdex™ 75 gel filtration column (Cytiva, USA) equilibrated in 20 mM potassium phosphate pH 7.2, 100 mM KCl. The subsequent steps were identical to those described for aS140.

6.4 Protein aggregation assays

aS aggregation assays were performed by tracking the fluorescence of a marker dye that binds to aggregated species. As time progresses and the protein starts forming characteristically misfolded species, the overall fluorescence level increases, until the aggregation reaches an equilibrium between the various aggregated species and the monomeric protein and the fluorescence level stabilises at a plateau (Fig. 1.3). The fluorescence trace can then be used to determine kinetics parameters of aggregation and compare the effect of various molecules as modulators of the process. The fluorescence marker that is most often used in protein aggregation assays is thioflavin T (ThT)[200]. However, since ThT has also been shown to bind nucleic acids and thus shows a very high background when performing aggregation assays in their presence[141], we performed the aggregation assays in the presence of nucleic assays with the dye Proteostat™. The aggregation assays were performed by quickly thawing a frozen aliquot of the protein, either aS140 or aS103, and immediately filtering them with a 0.22 µm syringe filter (Merck KGaA, Germany). Total yeast RNA (totRNA, Roche, Switzerland) was resuspended in aggregation assay buffer (20 mM potassium phosphate pH 7.2, 100 mM KCl, 5 mM $MgCl_2$) to form the stock solution. For aggregation assays in the presence of DNA, total DNA from herring sperm (totDNA, Merck KGaA, Germany) was resuspended in sterile, nuclease-free water for the stock solution and sonicated with a tip sonicator for 5 min at room temperature. Protein, totRNA or totDNA were diluted to the final concentration in the aggregation assay buffer with a 1000-time diluted Proteostat™ detection dye.

For the control aggregation assays with bovine serum albumin (BSA), ultra-pure, DNA- and RNA-free BSA was used (Merck KGaA, Germany) and diluted to the final concentration in the sample buffer as described above for aS.

The samples at a constant concentration of protein at 50 µM and at various concentrations of nucleic acids were prepared in nucleic-acid low-binding Eppendorf tubes (Merck KGaA, Germany) and immediately distributed in wells on a 96-well black plate with transparent bottom (Corning, USA). As described in Results, chapter 1, a single

borosilicate glass bead (Merck KGaA, Germany) was added to each well to ensure sample homogeneity and reproducibility.

The experiment was performed in the Tecan Spark multiplate reader (Tecan, Switzerland). The excitation wavelength was set to 435 nm for ThT and 505 nm for Proteostat™, while the emission wavelength was set to 480 nm for ThT and 590 nm for Proteostat™ respectively at a gain setting of 60. Plates were incubated at 37 °C with constant double-orbital shaking at 200 rpm. Reads were taken every 10 min for aS140 and every 5 min for aS103 for 24 h, with each read represented as an average of 10 scans over an area of the whole well. All protein samples were aggregated in parallel to total RNA alone at the same starting concentrations and conditions for RNA degradation control.

6.5 Analysis of aggregation data

The fluorescence traces acquired during the aggregation assays were first corrected for background fluorescence by subtracting the values from wells containing only buffer and the aggregation dye. The data were then fitted to a sigmoid curve using the Hill function (1):

$$h(x) = A \frac{x^h}{x^h + K^h}$$

by means of non-linear least-square regression analysis using the curve_fit function of Scientific Python (SciPy), where *A*, *h* and *K* are the fitting parameters.

To compute the aggregation rate, the gradient of the fitted curve in the point of inflection before plateau was determined. With the previously set parameters and a Taylor expansion of $h(x)$ around K , the aggregation rate α was calculated as (2):

$$x \equiv h\,(x = K) = h(K) + \frac{\partial h(x)}{\partial x}\,;\; x = K$$

$$\alpha = A\frac{h}{2\,K}$$

The aggregation lag time was defined as the intersection of the fitting line with the slope α and centered in $(x = K)$ and $(y = \frac{A}{2})$. The aggregation lag time τ was defined resolving the linear system (3):

$$y(x) = \alpha x + q$$

$$y\left(\frac{K}{2}\right) = \alpha\frac{K}{2} + q = \frac{A}{2} \rightarrow q = \frac{A - \alpha K}{2}$$

$$\alpha\tau + \frac{A - \alpha K}{2} = 0 \rightarrow \tau = \frac{\alpha K - A}{2\alpha}$$

$$\tau = K\frac{(h - 1)}{h}$$

The fitted parameters of individual replicates at the same concentration of RNA/DNA were averaged and the standard deviation was calculated. The parameters of experiments at identical RNA concentrations were analysed for significance using a two-way Student t-test, while the one-factor ANOVA test was used for comparing parameters of experiments at different RNA concentrations or with a different number of replicates. The Fisher exact test was used to compare the parameters at different concentrations between aS140 and aS103.

6.6 Nucleic acid quantification and extraction from aS aggregates

After the completion of the aggregation assay, the samples from individual wells on the 96-well plate were transferred to 1,5 mL protein low-binding Eppendorf tubes and centrifuged at 4 °C, 18 000 g for 30 min. The soluble fraction was transferred to another 1,5 mL Eppendorf tube and the RNA found in solution was quantified using Qubit RNA BR fluorescent dye on the Qubit Flex instrument (Thermo Fisher Scientific, USA). We have used the RNA-specific Qubit dye to avoid the potential interference from the remaining monomeric protein in solution and any potential lower molecular weight protein aggregated species left in solution after the centrifugation. Soluble DNA after aggregation was likewise quantified with Qubit DNA BR dye using the Qubit Flex instrument (Thermo Fisher Scientific, USA).

For the quantification of nucleic extracts sequestered in the insoluble fraction, the pellet obtained after the initial centrifugation step was resuspended in 25 µL of digestion buffer (2 M urea, 100 µg/ml proteinase K, 3 mM DTT) and incubated at room temperature for 15 min. 125 µL of 6 M guanidine hydrochloride was added to a final concentration of 5 M and the sample was incubated at room temperature for further 30 min. 1 ml of Trizol™ (Thermo Fisher Scietific, USA) was added to extract the RNA. 200 µL of ice-cold chloroform (Merck KGoA, Germany) was added to initiate the separation of the organic and inorganic phases. After a centrifugation step at 12.000 g, 4 °C for 15 min, the RNA was extracted from the inorganic phase using the RNA Clean & Concentrator-5 kit (Zymo Research, USA) according to the standard protocol specified by the manufacturer. The isolated RNA was eluted using nuclease-free water into 1,5 mL nucleic acid low binding Eppendorf tubes and again quantified using the Qubit RNA BR fluorescent dye. For the extraction of DNA, the addition of Trizol™ was omitted and the DNA was instead

extracted directly from the digested pellet with the QIAamp DNA Mini Kit (Qiagen, Germany) according to the manufacturer protocol. DNA was eluted with 50 µL of nuclease-free buffer AE. The extracted DNA was again quantified using the Qubit DNA BR dye.

A similar protocol for soluble RNA quantification was adopted for the control experiments using the nuclease-free BSA and RNA alone. Due to the use of amyloid-specific fluorescent dye Proteostat™, the insoluble aggregates after centrifugation were thus stained and visible with the naked eye even though the overall amount was low. However, in these cases no pellet could be observed after the centrifugation. This was confirmed by a visual inspection of the wells with a fluorescence microscope. As a result, while the soluble RNA was quantified as described above, no analysis of the insoluble RNA was performed due to a lack of pellet.

Due to different rates of degradation and protocols used from the quantification of sequestered nucleic acids, several steps were included in the analysis to achieve the most comparable results. The concentration of RNA/DNA before and after aggregation was measured in each individual well. The resulting mass concentration before and after aggregation were then recalculated into RNA mass according to the volume of the individual sample. This was done in order to account for the slightly different RNA/DNA degradation rate in each individual well, as well as potential minor errors in pipetting. The mass of the RNA/DNA left in the solution and the one extracted from the aggregates were normalised against the starting nucleic acid mass and expressed as percentages. The final result was thus presented as an average of the individual percentages of all the replicates from one starting RNA concentration. The individual percentages at different starting nucleic acids concentration were then compared between each other for significance using a two-way Student t-test.

6.7 Biolayer interferometry

Biolayer interferometry (BLI) is a spectroscopic method used for the study of macromolecular interactions in solution, based on the reflection of white light on optical biolayers. The technique relies on the immobilisation of one of the interactors on an optical biosensor, which is then dipped in the solution containing the free ligand. As the ligand binds to the surface of the sensor, the thickness of the layer of molecules increases, which in turn increases the white light reflection. The increase in interference due to the increasingly disparate phases of the individual reflections produces a spectral shift which represents the measured signal. This signal increases over time with increased binding until an equilibrium is reached or until the saturation of the binding spots occurs. The sensor with the ligand still bound is then transferred to a well containing only buffer, which causes it to dissociate, decreasing the spectral shift, until a dynamic equilibrium is reached again. The possibility of measuring both the

association and dissociation step separately allows the estimation of real-time kinetic constants for both processes. This provides a very reliable way of estimating the affinity constant K_D as the ratio between the association and dissociation constants (k_{on} and k_{off}).

The BLI experiments of the binding characteristics between monomeric aS and RNA were performed on the Octet Red instrument (Sartorius, Germany). Due to the difficulty of labelling and immobilising the total RNA, the protein was instead labelled using the biotin-maleimide reagent (Thermo Fisher Scientific, USA) according to the manufacturer's instructions. The binding assays were performed in the aggregation assay buffer (20 mM potassium phosphate pH 7.2, 100 mM KCl, 5 mM $MgCl_2$) with the addition of 0.05 % Tween 20 (Merck KGoA, Germany) at 25 °C. Streptavidin-coated biosensors (Sartorius, Germany) were loaded with 10 μM biotinylated aS140 or aS103 and assayed against the range of RNA concentrations used in the aggregation assays (500, 250, 100 and 50 ng/μL). The lack of signal during the association and dissociation phases (Fig. 3.12 C) reflects no quantifiable interaction and thus the data could not have been fitted for further analysis.

6.8 Microscale thermophoresis

Microscale Thermophoresis (MST) is fluorescence-based technique, based on measuring the diffusion times of molecules in a laser-induced thermal gradient. The movement along the temperature gradient is dependent on the hydrodynamic radius of the molecules, which changes upon structural conformations or interaction with other molecules. MST can thus be used for tracking the folding dynamics of biological macromolecules or their interaction kinetics. The measurements are performed in a glass capillary, requiring only small volumes and no immobilisations. Measurements can also be performed on label-free molecules if their intrinsic fluorescence is high enough, otherwise the molecules need to be tagged with a fluorescent dye.

Due to the very low base intrinsic fluorescence of aS, the protein was tagged with the dye Cy5 according to the instructions provided by the manufacturer. The labelled protein was then diluted to 50 μM and RNA added up to 500 ng/μL in aggregation buffer (20 mM potassium phosphate pH 7.2, 100 mM KCl, 5 mM $MgCl_2$). 10 μL of sample was pipetted into the glass capillary and the measurements were performed with the Monolith NT.115 MST instrument (NanoTemper, Germany) at 37 °C with the excitation wavelength set to 625 nm and the emission to 680 nm. Measurements of aS alone and aS in the presence of RNA did not show any difference (Fig. 3.12 A,B), hence no kinetic analysis of the results was performed.

6.9 Fourier-transform infrared microspectroscopy (micro-FTIR)

As for fluorescence aggregation assays, the protein samples were quickly defrosted and immediately filtered with a 0.22 μm syringe filter. Total yeast RNA (Roche) was resuspended in the aggregation buffer (20 mM potassium phosphate pH 7.2, 100 mM KCl, 5 mM $MgCl_2$). Protein and RNA were mixed immediately before the start of the assay to the final concentrations of 50 μM and 0.5 mg/mL, respectively, with the addition of NaN_3 to a final concentration of 0.05 % . The samples were prepared in 1.5 mL protein low-binding Eppendorf tubes and incubated at 37 °C with constant agitation at 800 rpm for 144 h for aS140 and 6 h for aS103. At designated time points, 20 μL of sample was removed from the Eppendorf tubes and cast onto CaF_2 windows due to its optical properties and lack of absorbance in the IR spectrum. Four droplets of 5 μL were deposited at the centre of CaF_2 windows and dried completely under low-humidity conditions.

Micro-FTIR measurements were performed with an IFS 66v/S spectrometer combined with a Hyperion IR-microscope (Bruker GmbH, Ettlingen, Germany). The Hyperion microscope was equipped with a 15x Cassegrain objective lens with a numerical aperture of 0.4 and a liquid nitrogen cooled MCT detector. Spectra were recorded in transmission mode within the spectral range between 500 and 6000 cm^{-1} at a resolution of 2 cm^{-1}. A sample area with the dimensions of 200x200 μm^2 was scanned and the final spectrum represents the average of 512 scans.

The obtained spectra were baseline-corrected, smoothed with a spline algorithm, and normalized by the Wavemetrics software. As described in Chapter 2 of Results, the amide I was deconvolved using five Lorentzian components corresponding to characteristic absorbances of various secondary structure elements (Fig. 3.4). These fitted curves were centred at 1625, 1640, 1652, 1670, and 1690 $cm^{-1} \pm 2$ cm^{-1} with a fixed full width at half maximum (FWHM) equal to 25 cm^{-1} and intensity as the free parameter. The secondary structure assignment was carried as follows: 1623-1627 cm^{-1} was assigned to parallel and antiparallel β-sheet, 1638-1642 cm^{-1} to disordered, 1650-1654 cm^{-1} to α-helix, 1668-1672 cm^{-1} to turn and 1688-1692 cm^{-1} to antiparallel β-sheet. The second-order derivatives of the IR spectra determined the number and peak positions of the Lorentzian line shapes. For the frequency shift analysis of the $(\beta\text{-sheet})_{P,AP}$ component (Fig. 3.4), the spectra were fitted using the same parameters mentioned above with the only exception of giving a higher frequency freedom to the $(\beta\text{-sheet})_{P,AP}$ component, e.g. 1623-1633 cm^{-1} instead of 1623-1627 cm^{-1}.

6.10 Atomic force microscopy – infrared spectroscopy

The aggregation of protein samples in the presence and absence of RNA was set up as described in Section 3.9. For the preparation of samples used for the AFM-IR experiments, the aliquots were taken from the samples at the same time points and

prepared by cast-dropping 2 µL three times onto a 1 cm x 1 cm template-stripped gold surface with a nominal roughness of 0.3 nm (Platypus Technologies, WI, USA). The samples were left to deposit for 5 min at room temperature, then rinsed two times with 40 µL of Milli-Q water to achieve a sparse area of deposited protein aggregates. The samples were then let to dry for 1 h in a controlled low-humidity environment.

AFM images were acquired at a line rate of 1 - 1.5 Hz and a resolution of xxx pixels per line using the nanoIR2 instrument platform (Anasys Instruments Inc., CA, USA). The probes were PR-EX-nIR2-10 gold-coated silicon cantilevers with an elastic constant of 0.07 - 0.4 Nm^{-1} and a nominal radius of 25 nm. All measurements were done in contact mode and at room temperature. The AFM images were flattened using the manufacturer-provided Analysis Studio software (Anasys Instruments Inc., Santa Barbara, CA, USA).

6.11 Fibril enzymatic activity assays

The protocol for fibril enzymatic activity was used as described by Horvath and colleagues[98]. The main concern, as the fibril enzymatic activity is highly determined on the aggregate structure, was to ensure a homogenous population of fibrils. This was done through seeded aggregation of monomeric aS under quiescent conditions to minimise potential changes in conformation. Seeds for fibril growth were prepared as described[98]. Briefly, 150 µM aS was aggregated on a 96-well plate without Proteostat™ added as described in Section 3.4. After a roughly 16 h incubation, the aggregate suspension was transferred from the wells to 1.5 mL protein low-binding Eppendorf tubes. The samples were centrifuged at 18.000 g at room temperature for 30 min, upon which the supernatant was removed, and the pellet was resuspended in equal volume of sterile 1xPBS. The pellet was again centrifuged at 18.000 g at room temperature for 10 min, after which the supernatant was again removed and the pellet resuspended in sterile, 1xPBS. This washing step was done in order to avoid any potential monomeric or oligomeric aggregates in the final fibril stock. The resuspended fibrils were sonicated in a water bath at room temperature for 15 min at the setting of 8. Afterwards, the suspension was aliquoted in protein low-binding Eppendorf tubes and stored at – 80 °C until use. The seeds grown in the presence of RNA and heparin were prepared similarly, with the addition of another washing step after sonication to further facilitate the removal of RNA/heparin. Purified recombinant aS was quickly thawed and immediately filtered into protein low-binding Eppendorf tubes, the concentration adjusted to 250 µM and seeds added to 5 % of final volume. The tubes were then incubated at 37 °C for 5 days under quiescent conditions.

The fibrils were always prepared fresh on the day of the enzymatic assay. The samples incubated at 37 °C were centrifuged at 18.000 g for 1 h at room temperature and the protein concentration of the supernatant determined with a BCA assay. The fibril concentration was then determined by subtracting the supernatant concentration from

the initial monomer concentration (250 µM). The pellet after centrifugation was resuspended with sterile 1xPBS to a monomer-equivalent concentration of 250 µM.

Reaction mixes were prepared by addition of sterile 1xPBS, 0.5 M $MgCl_2$ and freshly resuspended RNA. Fibrils were sonicated in the water bath for 15 min at room temperature and the sonicator setting of 8, then added into the reaction mixes. To stimulate fibril-RNA interaction, the reaction mixes were incubated at room temperature with gentle mixing for 30 min. The substrate was then added to a final concentration of 10 mM (pNPP) or 5 mM (pNPA). The reaction mixture was briefly mixed with a pipette and aliquoted into Costar transparent 96-well plates (Corning Inc, NY, USA). The plates were sealed and incubated at 37 °C for up to 16 h in quiescent condition. The reactions were followed by measuring the absorbance at 410 nm (A_{410}) in a Tecan Spark multiplate reader (Tecan, Switzerland) every 3 min.

6.12 Analysis of enzymatic assays data

The time-resolved A_{410} data from the enzymatic assays were first blanked by subtracting the A_{410} values from wells containing only substrate and buffer. This way, any catalytic activity due to potential enzymatic contaminants in the buffer, as well as auto-hydrolysis of the substrate, was eliminated. The A_{410} values were then converted into the molar concentration of the product 4-nitrophenol (pNP) according to the compound's extinction coefficient of 7500 $M^{-1}cm^{-1}$ measured in 50 mM phosphate buffer at pH 7.

For the determination of basic kinetic parameters, the initial velocities of reaction (v_0) were calculated as the gradient of the linear fit of the plot of pNP concentration against time. The fitted v_0 were then plotted against initial substrate concentration to obtain the characteristic Michaelis-Menten plot of $v_0([S])$. The experimental values were fitted by least-square linear regression onto the Michaelis-Menten curve (1) to obtain the Michaelis-Menten constant K_M and maximum reaction velocity V_{max}:

(1) $$v_0 = \frac{v_{max} \cdot [S]}{K_M + [S]}$$

The parameters obtained from the fitting of the curve were then used to calculate the remaining two kinetic parameters. The turnover number k_{cat} was obtained by dividing the V_{max} value with the molar concentration of the enzyme (2), and the enzyme specificity ε by dividing the k_{cat} by K_M (3).

(2) $$k_{cat} = \frac{v_{max}}{[E]}$$

(3) $$\varepsilon = \frac{k_{cat}}{K_M}$$

For the determination of the binding affinity of RNA, the fitted v_0 at different concentrations of RNA were plotted against the initial RNA concentration. The resulting values were fitted by least-square linear regression with a logistic curve with four parameters (4):

(4) $$v_i = v_0 - \frac{(v_0 - v_\infty)\,[S]^h}{K_{0.5}^h + [S]^h}$$

with v_0, v_∞, $K_{0.5}$ and h as the fitting parameters.

The estimation of the inhibition mechanism was done using the shape of the curve in the Dixon plot. In the Dixon plot, the inverse of the initial velocity is plotted against the inhibitor concentration, unlike the classic inhibition plot characterised by the Eq. 4. The mechanism of modulation can subsequently be identified from the shape of the curve, and its competitive character from the overlap of curves at different substrate concentrations[254,256].

7. List of publications

1. Colantoni A., **Rupert J.**, Vandelli A., Tartaglia G.G., Zacco E. Zooming in on protein–RNA interactions: a multilevel workflow to identify interaction partners *Biochemical Society Transactions* (2020) 48 (4): 1529–1543. DOI: 10.1042/BST20191059

2. Vandelli A., Monti M., Milanetti E., Armaos A., **Rupert J.**, Zacco E., Bechara E., Delli Ponti R., Tartaglia G. G. Structural analysis of SARS-CoV-2 genome and predictions of the human interactome. *Nucleic Acids Research* (2020), 48 (20): 11270–11283. DOI: 10.1093/nar/gkaa864

3. Armaos A., Colantoni A., Proietti G., **Rupert J.**, Tartaglia G. G. catRAPID omics v2.0: going deeper and wider in the prediction of protein-RNA interactions. *Nucleic Acids Research* (2021), 49 (W1): W72–W79. DOI: 10.1093/nar/gkab393

4. Louka A., Bagnoli S., **Rupert J.**, Esapa B., Tartaglia G. G., Cellerino A., Pastore A., Terzibasi Tozzini E. New lessons on TDP-43 from old N. furzeri killifish. *Aging Cell* (2022), 21 (1): e13517. DOI: 10.1111/acel.13517

5. Giambruno R., **Rupert J.**, Zacco E. Chapter 10 - Experimental methods to study protein–nucleic acid interactions. Advances in Protein Molecular and Structural Biology Methods. Academic Press (2022), p. 137-161, ISBN 9780323902649. DOI: 10.1016/B978-0-323-90264-9.00010-6

6. **Rupert J.**, Monti M., Zacco E., Tartaglia G.G. RNA sequestration driven by amyloid formation: the alpha synuclein case. *Nucleic Acids Research* (2023), 51 (21): 11466–11478. DOI: 10.1093/nar/gkad857

7. Perego E., Zappone S., Castagnetti F., Mariani D., Vitiello E., **Rupert J.**, Zacco E., Tartaglia, G. G., Bozzoni I., Slenders E., Vicidomini G. Single-photon microscopy to study biomolecular condensates. *Nature Communications* (2023), 14 (8224). DOI: 10.1038/s41467-023-43969-7

8. Parisi G., Piacentini R., Incocciati A., Bonamore A., Macone A., **Rupert J.**, Zacco E., Miotto M., Milanetti E., Tartaglia G. G., Ruocco G., Boffi A., Di Rienzo, L. Design of protein-binding peptides with controlled binding affinity: the case of SARS-CoV-2 receptor binding domain and angiotensin-converting enzyme 2 derived peptides. *Frontiers in Molecular Biosciences* (2024), 10, DOI: 10.3389/fmolb.2023.1332359

9. Intze A., **Rupert J.**, Polito R., Temperini M. E., Zacco E., Mura F., Veber A., Puskar L., Schade U., Ortolani M., Tartaglia G. G., Giliberti V. Investigating the effect of RNA on α-synuclein aggregation by infrared microspectroscopy. *Biophysical Journal*, submitted

8. Bibliography

1. Crick, F. H. On protein synthesis. *Symp. Soc. Exp. Biol.* **12**, 138–163 (1958).

2. Anfinsen, C. B. Principles that Govern the Folding of Protein Chains. *Science* **181**, 223–230 (1973).

3. Louros, N., Schymkowitz, J. & Rousseau, F. Mechanisms and pathology of protein misfolding and aggregation. *Nat. Rev. Mol. Cell Biol.* **24**, 912–933 (2023).

4. Otzen, D. & Riek, R. Functional Amyloids. *Cold Spring Harb. Perspect. Biol.* **11**, a033860 (2019).

5. Monti, M., Armaos, A., Fantini, M., Pastore, A. & Tartaglia, G. G. Aggregation is a Context-Dependent Constraint on Protein Evolution. *Front. Mol. Biosci.* **8**, 678115 (2021).

6. Dobson, C. M. Principles of protein folding, misfolding and aggregation. *Semin. Cell Dev. Biol.* **15**, 3–16 (2004).

7. Rousseau, F., Schymkowitz, J. & Serrano, L. Protein aggregation and amyloidosis: confusion of the kinds? *Curr. Opin. Struct. Biol.* **16**, 118–126 (2006).

8. Tayeb-Fligelman, E. *et al.* The cytotoxic Staphylococcus aureus PSMα3 reveals a cross-α amyloid-like fibril. *Science* **355**, 831–833 (2017).

9. Sipe, J. D. & Cohen, A. S. Review: History of the Amyloid Fibril. *J. Struct. Biol.* **130**, 88–98 (2000).

10. Rambaran, R. N. & Serpell, L. C. Amyloid fibrils. *Prion* **2**, 112–117 (2008).

11. Prusiner, S. B. Novel proteinaceous infectious particles cause scrapie. *Science* **216**, 136–144 (1982).

12. Glenner, G. G. & Wong, C. W. Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem. Biophys. Res. Commun.* **120**, 885–890 (1984).

13. Dobson, C. M. Protein folding and misfolding. *Nature* **426**, 884–890 (2003).

14. 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).

15. Baum, J. *et al.* Homage to Chris Dobson. *Front. Mol. Biosci.* **6**, (2019).

16. Housmans, J. A. J., Wu, G., Schymkowitz, J. & Rousseau, F. A guide to studying protein aggregation. *FEBS J.* **290**, 554–583 (2023).

17. Meisl, G. *et al.* Molecular mechanisms of protein aggregation from global fitting of kinetic models. *Nat. Protoc.* **11**, 252–272 (2016).

18. Navarro, S. & Ventura, S. Computational methods to predict protein aggregation. *Curr. Opin. Struct. Biol.* **73**, 102343 (2022).

19. Gan, L., Cookson, M. R., Petrucelli, L. & La Spada, A. R. Converging pathways in neurodegeneration, from genetics to mechanisms. *Nat. Neurosci.* **21**, 1300–1309 (2018).

20. van Rheenen, W. *et al.* Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nat. Genet.* **48**, 1043–1048 (2016).

21. Michaels, T. C. T. *et al.* Amyloid formation as a protein phase transition. *Nat. Rev. Phys.* **5**, 379–397 (2023).

22. Mathieu, C., Pappu, R. V. & Taylor, J. P. Beyond aggregation: Pathological phase transitions in neurodegenerative disease. *Science* **370**, 56–60 (2020).

23. Gruebele, M., Dave, K. & Sukenik, S. Globular Protein Folding In Vitro and In Vivo. *Annu. Rev. Biophys.* **45**, 233–251 (2016).

24. Tompa, P. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* **37**, 509–516 (2012).

25. Chong, S.-H. & Ham, S. Folding Free Energy Landscape of Ordered and Intrinsically Disordered Proteins. *Sci. Rep.* **9**, 14927 (2019).

26. Englander, S. W. Protein Folding Intermediates and Pathways Studied by Hydrogen Exchange. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 213–238 (2000).

27. Ghosh, D. K. & Ranjan, A. The metastable states of proteins. *Protein Sci. Publ. Protein Soc.* **29**, 1559–1568 (2020).

28. Adamcik, J. & Mezzenga, R. Amyloid Polymorphism in the Protein Folding and Aggregation Energy Landscape. *Angew. Chem. Int. Ed Engl.* **57**, 8370–8382 (2018).

29. Pawar, A. P. *et al.* Prediction of 'aggregation-prone' and 'aggregation-susceptible' regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol.* **350**, 379–392 (2005).

30. Sisodia, S. S. Beta-amyloid precursor protein cleavage by a membrane-bound protease. *Proc. Natl. Acad. Sci.* **89**, 6075–6079 (1992).

31. Oldfield, C. J. & Dunker, A. K. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* **83**, 553–584 (2014).

32. Uversky, V. N. Intrinsically disordered proteins and their (disordered) proteomes in neurodegenerative disorders. *Front. Aging Neurosci.* **7**, 18 (2015).

33. De Baets, G., Van Durme, J., Rousseau, F. & Schymkowitz, J. A genome-wide sequence-structure analysis suggests aggregation gatekeepers constitute an evolutionary constrained functional class. *J. Mol. Biol.* **426**, 2405–2412 (2014).

34. Marshall, L. R. & Korendovych, I. V. Catalytic amyloids: Is misfolding folding? *Curr. Opin. Chem. Biol.* **64**, 145–153 (2021).

35. Knowles, T. P. J., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384–396 (2014).

36. Bolognesi, B. *et al.* The mutational landscape of a prion-like domain. *Nat. Commun.* **10**, 4162 (2019).

37. Chiti, F. *et al.* Designing conditions for in vitro formation of amyloid protofilaments and fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 3590–3594 (1999).

38. Saborio, G. P., Permanne, B. & Soto, C. Sensitive detection of pathological prion protein by cyclic amplification of protein misfolding. *Nature* **411**, 810–813 (2001).

39. Ivanova, M. I., Lin, Y., Lee, Y.-H., Zheng, J. & Ramamoorthy, A. Biophysical processes underlying cross-seeding in amyloid aggregation and implications in amyloid pathology. *Biophys. Chem.* **269**, 106507 (2021).

40. Boeynaems, S. *et al.* Protein Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol.* **28**, 420–435 (2018).

41. Dada, S. T. *et al.* Spontaneous nucleation and fast aggregate-dependent proliferation of α-synuclein aggregates within liquid condensates at neutral pH. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2208792120 (2023).

42. Babinchak, W. M. & Surewicz, W. K. Liquid-Liquid Phase Separation and Its Mechanistic Role in Pathological Protein Aggregation. *J. Mol. Biol.* **432**, 1910–1925 (2020).

43. Vecchi, G. *et al.* Proteome-wide observation of the phenomenon of life on the edge of solubility. *Proc. Natl. Acad. Sci.* **117**, 1015–1020 (2020).

44. Tartaglia, G. G., Pechmann, S., Dobson, C. M. & Vendruscolo, M. Life on the edge: a link between gene expression levels and aggregation rates of human proteins. *Trends Biochem. Sci.* **32**, 204–206 (2007).

45. Ruff, K. M. *et al.* Sequence grammar underlying the unfolding and phase separation of globular proteins. *Mol. Cell* **82**, 3193-3208.e8 (2022).

46. Kundra, R., Ciryam, P., Morimoto, R. I., Dobson, C. M. & Vendruscolo, M. Protein homeostasis of a metastable subproteome associated with Alzheimer's disease. *Proc. Natl. Acad. Sci.* **114**, E5703–E5711 (2017).

47. Tyedmers, J., Mogk, A. & Bukau, B. Cellular strategies for controlling protein aggregation. *Nat. Rev. Mol. Cell Biol.* **11**, 777–788 (2010).

48. Lamark, T. & Johansen, T. Aggrephagy: selective disposal of protein aggregates by macroautophagy. *Int. J. Cell Biol.* **2012**, 736905 (2012).

49. Hipp, M. S., Kasturi, P. & Hartl, F. U. The proteostasis network and its decline in ageing. *Nat. Rev. Mol. Cell Biol.* **20**, 421–435 (2019).

50. Yoshimura, Y. *et al.* Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation. *Proc. Natl. Acad. Sci.* **109**, 14446–14451 (2012).

51. Wilcken, R., Wang, G., Boeckler, F. M. & Fersht, A. R. Kinetic mechanism of p53 oncogenic mutant aggregation and its inhibition. *Proc. Natl. Acad. Sci.* **109**, 13584–13589 (2012).

52. Borgia, M. B. *et al.* Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins. *Nature* **474**, 662–665 (2011).

53. Bemporad, F., De Simone, A., Chiti, F. & Dobson, C. M. Characterizing intermolecular interactions that initiate native-like protein aggregation. *Biophys. J.* **102**, 2595–2604 (2012).

54. Gallardo, R., Ranson, N. A. & Radford, S. E. Amyloid structures: much more than just a cross-β fold. *Curr. Opin. Struct. Biol.* **60**, 7–16 (2020).

55. Maury, C. P. J. Amyloid and the origin of life: self-replicating catalytic amyloids as prebiotic informational and protometabolic entities. *Cell. Mol. Life Sci.* **75**, 1499–1507 (2018).

56. Greenwald, J. & Riek, R. On the possible amyloid origin of protein folds. *J. Mol. Biol.* **421**, 417–426 (2012).

57. Yang, Y. *et al.* Structures of α-synuclein filaments from human brains with Lewy pathology. *Nature* **610**, 791–795 (2022).

58. Schweighauser, M. *et al.* Structures of α-synuclein filaments from multiple system atrophy. *Nature* **585**, 464–469 (2020).

59. Guerrero-Ferreira, R., Kovacik, L., Ni, D. & Stahlberg, H. New insights on the structure of alpha-synuclein fibrils using cryo-electron microscopy. *Curr. Opin. Neurobiol.* **61**, 89–95 (2020).

60. Landau, M. Mimicking cross-α amyloids. *Nat. Chem. Biol.* **14**, 833–834 (2018).

61. Salinas, N. *et al.* The amphibian antimicrobial peptide uperin 3.5 is a cross-α/cross-β chameleon functional amyloid. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2014442118 (2021).

62. Eisenberg, D. S. & Sawaya, M. R. Structural Studies of Amyloid Proteins at the Molecular Level. *Annu. Rev. Biochem.* **86**, 69–95 (2017).

63. Ye, X., Hedenqvist, M. S., Langton, M. & Lendel, C. On the role of peptide hydrolysis for fibrillation kinetics and amyloid fibril morphology. *RSC Adv.* **8**, 6915–6924 (2018).

64. Ulamec, S. M., Brockwell, D. J. & Radford, S. E. Looking Beyond the Core: The Role of Flanking Regions in the Aggregation of Amyloidogenic Peptides and Proteins. *Front. Neurosci.* **14**, 611285 (2020).

65. Khare, S. D., Chinchilla, P. & Baum, J. Multifaceted interactions mediated by intrinsically disordered regions play key roles in alpha synuclein aggregation. *Curr. Opin. Struct. Biol.* **80**, 102579 (2023).

66. Nonaka, T., Watanabe, S. T., Iwatsubo, T. & Hasegawa, M. Seeded Aggregation and Toxicity of α-Synuclein and Tau: CELLULAR MODELS OF NEURODEGENERATIVE DISEASES *. *J. Biol. Chem.* **285**, 34885–34898 (2010).

67. Tompa, P. Structural disorder in amyloid fibrils: its implication in dynamic interactions of proteins. *FEBS J.* **276**, 5406–5415 (2009).

68. Törnquist, M. *et al.* Secondary nucleation in amyloid formation. *Chem. Commun. Camb. Engl.* **54**, 8667–8684 (2018).

69. Oikawa, T. *et al.* α-Synuclein Fibrils Exhibit Gain of Toxic Function, Promoting Tau Aggregation and Inhibiting Microtubule Assembly. *J. Biol. Chem.* **291**, 15046–15056 (2016).

70. Wentink, A. S. *et al.* Molecular dissection of amyloid disaggregation by human HSP70. *Nature* **587**, 483–488 (2020).

71. Fusco, G. *et al.* Structural basis of membrane disruption and cellular toxicity by α-synuclein oligomers. *Science* **358**, 1440–1443 (2017).

72. Khan, M. R. *et al.* Amyloidogenic Oligomerization Transforms Drosophila Orb2 from a Translation Repressor to an Activator. *Cell* **163**, 1468–1483 (2015).

73. Garcia-Pardo, J. *et al.* Cryo-EM structure of hnRNPDL-2 fibrils, a functional amyloid associated with limb-girdle muscular dystrophy D3. *Nat. Commun.* **14**, 239 (2023).

74. Cohen, S. I. A., Vendruscolo, M., Dobson, C. M. & Knowles, T. P. J. From macroscopic measurements to microscopic mechanisms of protein aggregation. *J. Mol. Biol.* **421**, 160–171 (2012).

75. Aliyan, A., Cook, N. P. & Martí, A. A. Interrogating Amyloid Aggregates using Fluorescent Probes. *Chem. Rev.* **119**, 11819–11856 (2019).

76. Breydo, L. *et al.* Structural differences between amyloid beta oligomers. *Biochem. Biophys. Res. Commun.* **477**, 700–705 (2016).

77. Cremades, N. *et al.* Direct observation of the interconversion of normal and toxic forms of α-synuclein. *Cell* **149**, 1048–1059 (2012).

78. Chen, S. W. *et al.* Structural characterization of toxic oligomers that are kinetically trapped during α-synuclein fibril formation. *Proc. Natl. Acad. Sci.* **112**, E1994–E2003 (2015).

79. Campioni, S. *et al.* A causative link between the structure of aberrant protein oligomers and their toxicity. *Nat. Chem. Biol.* **6**, 140–147 (2010).

80. Lasagna-Reeves, C. A. *et al.* Alzheimer brain-derived tau oligomers propagate pathology from endogenous tau. *Sci. Rep.* **2**, 700 (2012).

81. Fang, Y.-S. *et al.* Full-length TDP-43 forms toxic amyloid oligomers that are present in frontotemporal lobar dementia-TDP patients. *Nat. Commun.* **5**, 4824 (2014).

82. Katzmarski, N. *et al.* Aβ oligomers trigger and accelerate Aβ seeding. *Brain Pathol. Zurich Switz.* **30**, 36–45 (2020).

83. Kheterpal, I., Chen, M., Cook, K. D. & Wetzel, R. Structural differences in Abeta amyloid protofibrils and fibrils mapped by hydrogen exchange--mass spectrometry with on-line proteolytic fragmentation. *J. Mol. Biol.* **361**, 785–795 (2006).

84. Fitzpatrick, A. W. P. *et al.* Atomic structure and hierarchical assembly of a cross-β amyloid fibril. *Proc. Natl. Acad. Sci.* **110**, 5468–5473 (2013).

85. Close, W. *et al.* Physical basis of amyloid fibril polymorphism. *Nat. Commun.* **9**, 699 (2018).

86. Li, B. *et al.* Cryo-EM of full-length α-synuclein reveals fibril polymorphs with a common structural kernel. *Nat. Commun.* **9**, 3609 (2018).

87. Gremer, L. *et al.* Fibril structure of amyloid-β(1–42) by cryo–electron microscopy. *Science* **358**, 116–119 (2017).

88. Wälti, M. A. *et al.* Atomic-resolution structure of a disease-relevant Aβ(1–42) amyloid fibril. *Proc. Natl. Acad. Sci.* **113**, E4976–E4984 (2016).

89. Zhang, W. *et al.* Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *eLife* **8**, e43584 (2019).

90. Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P. & Eisenberg, D. S. Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat. Struct. Mol. Biol.* **26**, 619–627 (2019).

91. Michaels, T. C. T. *et al.* Chemical Kinetics for Bridging Molecular Mechanisms and Macroscopic Measurements of Amyloid Fibril Formation. *Annu. Rev. Phys. Chem.* **69**, 273–298 (2018).

92. Heldt, C. L., Zhang, S. & Belfort, G. Asymmetric amyloid fibril elongation: a new perspective on a symmetric world. *Proteins* **79**, 92–98 (2011).

93. Buell, A. K. *et al.* Solution conditions determine the relative importance of nucleation and growth processes in α-synuclein aggregation. *Proc. Natl. Acad. Sci.* **111**, 7671–7676 (2014).

94. Michaels, T. C. T. *et al.* Thermodynamic and kinetic design principles for amyloid-aggregation inhibitors. *Proc. Natl. Acad. Sci.* **117**, 24251–24257 (2020).

95. Rufo, C. M. *et al.* Short peptides self-assemble to produce catalytic amyloids. *Nat. Chem.* **6**, 303–309 (2014).

96. Arad, E., Baruch Leshem, A., Rapaport, H. & Jelinek, R. β-Amyloid fibrils catalyze neurotransmitter degradation. *Chem Catal.* **1**, 908–922 (2021).

97. Arad, E. *et al.* Native Glucagon Amyloids Catalyze Key Metabolic Reactions. *ACS Nano* **16**, 12889–12899 (2022).

98. Horvath, I. & Wittung-Stafshede, P. Amyloid Fibers of α-Synuclein Catalyze Chemical Reactions. *ACS Chem. Neurosci.* **14**, 603–608 (2023).

99. Otzen, D. & Nielsen, P. H. We find them here, we find them there: functional bacterial amyloid. *Cell. Mol. Life Sci. CMLS* **65**, 910–927 (2008).

100. Barnhart, M. M. & Chapman, M. R. Curli biogenesis and function. *Annu. Rev. Microbiol.* **60**, 131–147 (2006).

101. Seuring, C. *et al.* The mechanism of toxicity in HET-S/HET-s prion incompatibility. *PLoS Biol.* **10**, e1001451 (2012).

102. Schlissel, G., Krzyzanowski, M. K., Caudron, F., Barral, Y. & Rine, J. Aggregation of the Whi3 protein, not loss of heterochromatin, causes sterility in old yeast cells. *Science* **355**, 1184–1187 (2017).

103. Saarikangas, J. *et al.* Compartmentalization of ER-Bound Chaperone Confines Protein Deposit Formation to the Aging Yeast Cell. *Curr. Biol.* **27**, 773–783 (2017).

104. Caudron, F. & Barral, Y. A super-assembly of Whi3 encodes memory of deceptive encounters by single cells during yeast courtship. *Cell* **155**, 1244–1257 (2013).

105. Chakravarty, A. K., Smejkal, T., Itakura, A. K., Garcia, D. M. & Jarosz, D. F. A Non-amyloid Prion Particle that Activates a Heritable Gene Expression Program. *Mol. Cell* **77**, 251-265.e9 (2020).

106. Krüttner, S. *et al.* Drosophila CPEB Orb2A mediates memory independent of Its RNA-binding domain. *Neuron* **76**, 383–395 (2012).

107. Hervas, R. *et al.* Cryo-EM structure of a neuronal functional amyloid implicated in memory persistence in Drosophila. *Science* **367**, 1230–1234 (2020).

108. Aguzzi, A. & Falsig, J. Prion propagation, toxicity and degradation. *Nat. Neurosci.* **15**, 936–939 (2012).

109. Scheckel, C. & Aguzzi, A. Prions, prionoids and protein misfolding disorders. *Nat. Rev. Genet.* **19**, 405–418 (2018).

110. Vaquer-Alicea, J. & Diamond, M. I. Propagation of Protein Aggregation in Neurodegenerative Diseases. *Annu. Rev. Biochem.* **88**, 785–810 (2019).

111. Soto, C. & Pritzkow, S. Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases. *Nat. Neurosci.* **21**, 1332–1340 (2018).

112. Bloom, G. S. Amyloid-β and tau: the trigger and bullet in Alzheimer disease pathogenesis. *JAMA Neurol.* **71**, 505–508 (2014).

113. Taylor, J. P., Brown, R. H. & Cleveland, D. W. Decoding ALS: from genes to mechanism. *Nature* **539**, 197–206 (2016).

114. Portz, B., Lee, B. L. & Shorter, J. FUS and TDP-43 Phases in Health and Disease. *Trends Biochem. Sci.* **46**, 550–563 (2021).

115. Hallegger, M. *et al.* TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. *Cell* **184**, 4680-4696.e22 (2021).

116. Chang, A. *et al.* Homotypic fibrillization of TMEM106B across diverse neurodegenerative diseases. *Cell* **185**, 1346-1355.e15 (2022).

117. Schweighauser, M. *et al.* Age-dependent formation of TMEM106B amyloid filaments in human brains. *Nature* **605**, 310–314 (2022).

118. Smith, B. N. *et al.* Mutations in the vesicular trafficking protein annexin A11 are associated with amyotrophic lateral sclerosis. *Sci. Transl. Med.* **9**, eaad9157 (2017).

119. Bruijn, L. I. *et al.* Aggregation and motor neuron toxicity of an ALS-linked SOD1 mutant independent from wild-type SOD1. *Science* **281**, 1851–1854 (1998).

120. Mori, K. *et al.* The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* **339**, 1335–1338 (2013).

121. Wechalekar, A. D., Gillmore, J. D. & Hawkins, P. N. Systemic amyloidosis. *The Lancet* **387**, 2641–2654 (2016).

122. Muchtar, E. *et al.* Systemic amyloidosis from A (AA) to T (ATTR): a review. *J. Intern. Med.* **289**, 268–292 (2021).

123. Silva, J. L., De Moura Gallo, C. V., Costa, D. C. F. & Rangel, L. P. Prion-like aggregation of mutant p53 in cancer. *Trends Biochem. Sci.* **39**, 260–267 (2014).

124. Cirulli, E. T. *et al.* Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science* **347**, 1436–1441 (2015).

125. Newby, G. A. *et al.* A Genetic Tool to Track Protein Aggregates and Control Prion Inheritance. *Cell* **171**, 966-979.e18 (2017).

126. Gotor, N. L. *et al.* RNA-binding and prion domains: the Yin and Yang of phase separation. *Nucleic Acids Res.* **48**, 9491–9504 (2020).

127. Zanzoni, A. *et al.* Principles of self-organization in biological pathways: a hypothesis on the autogenous association of alpha-synuclein. *Nucleic Acids Res.* **41**, 9987–9998 (2013).

128. Kovachev, P. S. *et al.* RNA modulates aggregation of the recombinant mammalian prion protein by direct interaction. *Sci. Rep.* **9**, 12406 (2019).

129. Kampers, T., Friedhoff, P., Biernat, J., Mandelkow, E. M. & Mandelkow, E. RNA stimulates aggregation of microtubule-associated protein tau into Alzheimer-like paired helical filaments. *FEBS Lett.* **399**, 344–349 (1996).

130. Chakraborty, P. *et al.* Co-factor-free aggregation of tau into seeding-competent RNA-sequestering amyloid fibrils. *Nat. Commun.* **12**, 4231 (2021).

131. Lester, E. *et al.* Tau aggregates are RNA-protein assemblies that mislocalize multiple nuclear speckle components. *Neuron* **109**, 1675-1691.e9 (2021).

132. Datta, D., Tiwari, O. & Gupta, M. K. Self-Assembly of Diphenylalanine-Peptide Nucleic Acid Conjugates. *ACS Omega* **4**, 10715–10728 (2019).

133. Rha, A. K. *et al.* Electrostatic Complementarity Drives Amyloid/Nucleic Acid Co-Assembly. *Angew. Chem. Int. Ed Engl.* **59**, 358–363 (2020).

134. Carny, O. & Gazit, E. Creating prebiotic sanctuary: self-assembling supramolecular Peptide structures bind and stabilize RNA. *Orig. Life Evol. Biosphere J. Int. Soc. Study Orig. Life* **41**, 121–132 (2011).

135. Frenkel-Pinter, M. *et al.* Mutually stabilizing interactions between proto-peptides and RNA. *Nat. Commun.* **11**, 3137 (2020).

136. Braun, S. *et al.* Amyloid-associated nucleic acid hybridisation. *PloS One* **6**, e19125 (2011).

137. Audas, T. E., Jacob, M. D. & Lee, S. Immobilization of proteins in the nucleolus by ribosomal intergenic spacer noncoding RNA. *Mol. Cell* **45**, 147–157 (2012).

138. Audas, T. E. *et al.* Adaptation to Stressors by Systemic Protein Amyloidogenesis. *Dev. Cell* **39**, 155–168 (2016).

139. Theodoridis, P. R. *et al.* Local translation in nuclear condensate amyloid bodies. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2014457118 (2021).

140. Boke, E. & Mitchison, T. J. The balbiani body and the concept of physiological amyloids. *Cell Cycle Georget. Tex* **16**, 153–154 (2017).

141. Zacco, E. *et al.* RNA as a key factor in driving or preventing self-assembly of the TAR DNA-binding protein 43. *J. Mol. Biol.* **431**, 1671–1688 (2019).

142. Zacco, E. *et al.* Probing TDP-43 condensation using an in silico designed aptamer. *Nat. Commun.* **13**, 3306 (2022).

143. Lashuel, H. A., Overk, C. R., Oueslati, A. & Masliah, E. The many faces of α-synuclein: from structure and toxicity to therapeutic target. *Nat. Rev. Neurosci.* **14**, 38–48 (2013).

144. Fusco, G. *et al.* Direct observation of the three regions in α-synuclein that determine its membrane-bound behaviour. *Nat. Commun.* **5**, 3827 (2014).

145. Scott, D. & Roy, S. α-Synuclein Inhibits Intersynaptic Vesicle Mobility and Maintains Recycling-Pool Homeostasis. *J. Neurosci.* **32**, 10129–10135 (2012).

146. Burré, J. *et al.* α-Synuclein Promotes SNARE-Complex Assembly in Vivo and in Vitro. *Science* **329**, 1663–1667 (2010).

147. Rocha, E. M., De Miranda, B. & Sanders, L. H. Alpha-synuclein: Pathology, mitochondrial dysfunction and neuroinflammation in Parkinson's disease. *Neurobiol. Dis.* **109**, 249–257 (2018).

148. Schaser, A. J. *et al.* Alpha-synuclein is a DNA binding protein that modulates DNA repair with implications for Lewy body disorders. *Sci. Rep.* **9**, 10919 (2019).

149. Spillantini, M. G., Crowther, R. A., Jakes, R., Hasegawa, M. & Goedert, M. α-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc. Natl. Acad. Sci.* **95**, 6469–6473 (1998).

150. Grazia Spillantini, M. *et al.* Filamentous α-synuclein inclusions link multiple system atrophy with Parkinson's disease and dementia with Lewy bodies. *Neurosci. Lett.* **251**, 205–208 (1998).

151. Goedert, M., Jakes, R. & Spillantini, M. G. The Synucleinopathies: Twenty Years On. *J. Park. Dis.* **7**, S51–S69 (2017).

152. Doherty, C. P. A. *et al.* A short motif in the N-terminal region of α-synuclein is critical for both aggregation and function. *Nat. Struct. Mol. Biol.* **27**, 249–259 (2020).

153. Stephens, A. D. *et al.* Extent of N-terminus exposure of monomeric alpha-synuclein determines its aggregation propensity. *Nat. Commun.* **11**, 2820 (2020).

154. Hoyer, W., Cherny, D., Subramaniam, V. & Jovin, T. M. Impact of the Acidic C-Terminal Region Comprising Amino Acids 109–140 on α-Synuclein Aggregation in Vitro. *Biochemistry* **43**, 16233–16242 (2004).

155. Farzadfard, A. *et al.* The C-terminal tail of α-synuclein protects against aggregate replication but is critical for oligomerization. *Commun. Biol.* **5**, 123 (2022).

156. Bartels, T., Choi, J. G. & Selkoe, D. J. α-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature* **477**, 107–110 (2011).

157. Cholak, E. *et al.* Avidity within the N-terminal anchor drives α-synuclein membrane interaction and insertion. *FASEB J.* **34**, 7462–7482 (2020).

158. Burmann, B. M. *et al.* Regulation of α-synuclein by chaperones in mammalian cells. *Nature* **577**, 127–132 (2020).

159. Mirecka, E. A. *et al.* Sequestration of a β-Hairpin for Control of α-Synuclein Aggregation. *Angew. Chem. Int. Ed.* **53**, 4227–4230 (2014).

160. Zhang, Z. *et al.* Asparagine endopeptidase cleaves α-synuclein and mediates pathologic activities in Parkinson's disease. *Nat. Struct. Mol. Biol.* **24**, 632–642 (2017).

161. Falsone, S. F. *et al.* SERF protein is a direct modifier of amyloid fiber assembly. *Cell Rep.* **2**, 358–371 (2012).

162. Meyer, N. H. *et al.* Structural Fuzziness of the RNA-Organizing Protein SERF Determines a Toxic Gain-of-interaction. *J. Mol. Biol.* **432**, 930–951 (2020).

163. Lautenschläger, J. *et al.* C-terminal calcium binding of α-synuclein modulates synaptic vesicle interaction. *Nat. Commun.* **9**, 712 (2018).

164. Herrera, F. E. *et al.* Inhibition of α-Synuclein Fibrillization by Dopamine Is Mediated by Interactions with Five C-Terminal Residues and with E83 in the NAC Region. *PLOS ONE* **3**, e3394 (2008).

165. Fusco, G. *et al.* Structural basis of synaptic vesicle assembly promoted by α-synuclein. *Nat. Commun.* **7**, 12563 (2016).

166. Iyer, A., Sidhu, A. & Subramaniam, V. How important is the N-terminal acetylation of alpha-synuclein for its function and aggregation into amyloids? *Front. Neurosci.* **16**, (2022).

167. Wang, Y. *et al.* Phosphorylated α-Synuclein in Parkinson's Disease. *Sci. Transl. Med.* **4**, 121ra20-121ra20 (2012).

168. Farzadfard, A. *et al.* Glycation modulates alpha-synuclein fibrillization kinetics: A sweet spot for inhibition. *J. Biol. Chem.* **298**, (2022).

169. Sorrentino, Z. A. & Giasson, B. I. The emerging role of α-synuclein truncation in aggregation and disease. *J. Biol. Chem.* **295**, 10224–10244 (2020).

170. Li, W. *et al.* Aggregation promoting C-terminal truncation of α-synuclein is a normal cellular process and is enhanced by the familial Parkinson's disease-linked mutations. *Proc. Natl. Acad. Sci.* **102**, 2162–2167 (2005).

171. van der Wateren, I. M., Knowles, T. P. J., Buell, A. K., Dobson, C. M. & Galvagnion, C. C-terminal truncation of α-synuclein promotes amyloid fibril amplification at physiological pH. *Chem Sci* **9**, 5506–5516 (2018).

172. Iljina, M. *et al.* Kinetic model of the aggregation of alpha-synuclein provides insights into prion-like spreading. *Proc. Natl. Acad. Sci.* **113**, E1206–E1215 (2016).

173. Uversky, V. N., Li, J. & Fink, A. L. Evidence for a Partially Folded Intermediate in α-Synuclein Fibril Formation*. *J. Biol. Chem.* **276**, 10737–10744 (2001).

174. Yu, H., Han, W., Ma, W. & Schulten, K. Transient β-hairpin formation in α-synuclein monomer revealed by coarse-grained molecular dynamics simulation. *J. Chem. Phys.* **143**, 243142 (2015).

175. Antony, T. *et al.* Cellular Polyamines Promote the Aggregation of α-Synuclein*. *J. Biol. Chem.* **278**, 3235–3240 (2003).

176. Cohlberg, J. A., Li, J., Uversky, V. N. & Fink, A. L. Heparin and Other Glycosaminoglycans Stimulate the Formation of Amyloid Fibrils from α-Synuclein in Vitro. *Biochemistry* **41**, 1502–1511 (2002).

177. Galvagnion, C. *et al.* Lipid vesicles trigger α-synuclein aggregation by stimulating primary nucleation. *Nat. Chem. Biol.* **11**, 229–234 (2015).

178. Sorrentino, Z. A. *et al.* Physiological C-terminal truncation of α-synuclein potentiates the prion-like formation of pathological inclusions. *J. Biol. Chem.* **293**, 18914–18932 (2018).

179. Gaspar, R. *et al.* Secondary nucleation of monomers on fibril surface dominates α-synuclein aggregation and provides autocatalytic amyloid amplification. *Q. Rev. Biophys.* **50**, e6 (2017).

180. Celej, M. S. *et al.* Toxic prefibrillar α-synuclein amyloid oligomers adopt a distinctive antiparallel β-sheet structure. *Biochem. J.* **443**, 719–726 (2012).

181. Tuttle, M. D. *et al.* Solid-state NMR structure of a pathogenic fibril of full-length human α-synuclein. *Nat. Struct. Mol. Biol.* **23**, 409–415 (2016).

182. Frieg, B. *et al.* The 3D structure of lipidic fibrils of α-synuclein. *Nat. Commun.* **13**, 6810 (2022).

183. Tao, Y. *et al.* Heparin induces α-synuclein to form new fibril polymorphs with attenuated neuropathology. *Nat. Commun.* **13**, 4226 (2022).

184. Trinkaus, V. A. *et al.* In situ architecture of neuronal α-Synuclein inclusions. *Nat. Commun.* **12**, 2110 (2021).

185. Gao, X. *et al.* Human Hsp70 Disaggregase Reverses Parkinson's-Linked α-Synuclein Amyloid Fibrils. *Mol. Cell* **59**, 781–793 (2015).

186. Schneider, M. M. *et al.* The Hsc70 disaggregation machinery removes monomer units directly from α-synuclein fibril ends. *Nat. Commun.* **12**, 5999 (2021).

187. Tittelmeier, J. *et al.* The HSP110/HSP70 disaggregation system generates spreading-competent toxic α-synuclein species. *EMBO J.* **39**, e103954 (2020).

188. Cherny, D., Hoyer, W., Subramaniam, V. & Jovin, T. M. Double-stranded DNA stimulates the fibrillation of alpha-synuclein in vitro and is associated with the mature fibrils: an electron microscopy study. *J. Mol. Biol.* **344**, 929–938 (2004).

189. Hegde, M. L. & Rao, K. S. J. DNA induces folding in α-synuclein: Understanding the mechanism using chaperone property of osmolytes. *Arch. Biochem. Biophys.* **464**, 57–69 (2007).

190. K. Mukherjee, S., Knop, J.-M., Oliva, R., Möbitz, S. & Winter, R. Untangling the interaction of α-synuclein with DNA i-motifs and hairpins by volume-sensitive single-molecule FRET spectroscopy. *RSC Chem. Biol.* **2**, 1196–1200 (2021).

191. Jiang, K., Rocha, S., Kumar, R., Westerlund, F. & Wittung-Stafshede, P. C-terminal truncation of α-synuclein alters DNA structure from extension to compaction. *Biochem. Biophys. Res. Commun.* **568**, 43–47 (2021).

192. Tsukakoshi, K., Abe, K., Sode, K. & Ikebukuro, K. Selection of DNA aptamers that recognize α-synuclein oligomers using a competitive screening method. *Anal. Chem.* **84**, 5542–5547 (2012).

193. Tran, C. H., Saha, R., Blanco, C., Bagchi, D. & Chen, I. A. Modulation of α-Synuclein Aggregation In Vitro by a DNA Aptamer. *Biochemistry* **61**, 1757–1765 (2022).

194. Hallacli, E. *et al.* The Parkinson's disease protein alpha-synuclein is a modulator of processing bodies and mRNA stability. *Cell* **185**, 2035-2056.e33 (2022).

195. Armaos, A., Colantoni, A., Proietti, G., Rupert, J. & Tartaglia, G. G. catRAPID omics v2.0: going deeper and wider in the prediction of protein–RNA interactions. *Nucleic Acids Res.* **49**, W72–W79 (2021).

196. Rupert, J., Monti, M., Zacco, E. & Tartaglia, G. G. RNA sequestration driven by amyloid formation: the alpha synuclein case. *Nucleic Acids Res.* **51**, 11466–11478 (2023).

197. Bellucci, M., Agostini, F., Masin, M. & Tartaglia, G. G. Predicting protein associations with long noncoding RNAs. *Nat. Methods* **8**, 444–445 (2011).

198. Ni, X., McGlinchey, R. P., Jiang, J. & Lee, J. C. Structural Insights into α-Synuclein Fibril Polymorphism: Effects of Parkinson's Disease-Related C-Terminal Truncations. *J. Mol. Biol.* **431**, 3913–3919 (2019).

199. Fink, A. L. The Aggregation and Fibrillation of α-Synuclein. *Acc. Chem. Res.* **39**, 628–634 (2006).

200. Biancalana, M. & Koide, S. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1804**, 1405–1412 (2010).

201. Barth, A. Infrared spectroscopy of proteins. *Biochim. Biophys. Acta BBA - Bioenerg.* **1767**, 1073–1101 (2007).

202. Oberg, K. A., Ruysschaert, J.-M. & Goormaghtigh, E. The optimization of protein secondary structure determination with infrared and circular dichroism spectra. *Eur. J. Biochem.* **271**, 2937–2948 (2004).

203. Zandomeneghi, G., Krebs, M. R. H., McCammon, M. G. & Fändrich, M. FTIR reveals structural differences between native β-sheet proteins and amyloid fibrils. *Protein Sci.* **13**, 3314–3321 (2004).

204. Barth, A. & Zscherp, C. Substrate binding and enzyme function investigated by infrared spectroscopy. *FEBS Lett.* **477**, 151–156 (2000).

205. Ami, D. *et al.* Structure, Stability, and Aggregation of β-2 Microglobulin Mutants: Insights from a Fourier Transform Infrared Study in Solution and in the Crystalline State. *Biophys. J.* **102**, 1676–1684 (2012).

206. Dyer, R. B., Gai, F., Woodruff, W. H., Gilmanshin, R. & Callender, R. H. Infrared Studies of Fast Events in Protein Folding. *Acc. Chem. Res.* **31**, 709–716 (1998).

207. Ami, D. *et al.* In situ characterization of protein aggregates in human tissues affected by light chain amyloidosis: a FTIR microspectroscopy study. *Sci. Rep.* **6**, 29096 (2016).

208. Giliberti, V. *et al.* Heterogeneity of the Transmembrane Protein Conformation in Purple Membranes Identified by Infrared Nanospectroscopy. *Small* **13**, 1701181 (2017).

209. Lu, F., Jin, M. & Belkin, M. A. Tip-enhanced infrared nanospectroscopy via molecular expansion force detection. *Nat. Photonics* **8**, 307–312 (2014).

210. Dazzi, A. & Prater, C. B. AFM-IR: Technology and Applications in Nanoscale Infrared Spectroscopy and Chemical Imaging. *Chem. Rev.* **117**, 5146–5173 (2017).

211. Amenabar, I. *et al.* Structural analysis and mapping of individual protein complexes by infrared nanospectroscopy. *Nat. Commun.* **4**, 2890 (2013).

212. Giliberti, V. *et al.* Tip-Enhanced Infrared Difference-Nanospectroscopy of the Proton Pump Activity of Bacteriorhodopsin in Single Purple Membrane Patches. *Nano Lett.* **19**, 3104–3114 (2019).

213. Kochan, K. *et al.* In vivo atomic force microscopy–infrared spectroscopy of bacteria. *J. R. Soc. Interface* **15**, 20180115 (2018).

214. Ruggeri, F. S. *et al.* Infrared nanospectroscopy characterization of oligomeric and fibrillar aggregates during amyloid formation. *Nat. Commun.* **6**, 7831 (2015).

215. Zhou, L. & Kurouski, D. Structural Characterization of Individual α-Synuclein Oligomers Formed at Different Stages of Protein Aggregation by Atomic Force Microscopy-Infrared Spectroscopy. *Anal. Chem.* **92**, 6806–6810 (2020).

216. Ruggeri, F. S. *et al.* Infrared nanospectroscopy reveals the molecular interaction fingerprint of an aggregation inhibitor with single Aβ42 oligomers. *Nat. Commun.* **12**, 688 (2021).

217. Chiti, F. & Dobson, C. M. Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.* **5**, 15–22 (2009).

218. Jahn, T. R. *et al.* The Common Architecture of Cross-β Amyloid. *J. Mol. Biol.* **395**, 717–727 (2010).

219. Moran, S. D. & Zanni, M. T. How to Get Insight into Amyloid Structure and Formation from Infrared Spectroscopy. *J. Phys. Chem. Lett.* **5**, 1984–1993 (2014).

220. Kurouski, D. *et al.* Is Supramolecular Filament Chirality the Underlying Cause of Major Morphology Differences in Amyloid Fibrils? *J. Am. Chem. Soc.* **136**, 2302–2312 (2014).

221. Horrocks, M. H. *et al.* Fast Flow Microfluidics and Single-Molecule Fluorescence for the Rapid Characterization of α-Synuclein Oligomers. *Anal. Chem.* **87**, 8818–8826 (2015).

222. Qin, Z., Hu, D., Han, S., Hong, D.-P. & Fink, A. L. Role of Different Regions of α-Synuclein in the Assembly of Fibrils. *Biochemistry* **46**, 13322–13330 (2007).

223. vandenAkker, C. C., Engel, M. F. M., Velikov, K. P., Bonn, M. & Koenderink, G. H. Morphology and Persistence Length of Amyloid Fibrils Are Correlated to Peptide Molecular Structure. *J. Am. Chem. Soc.* **133**, 18030–18033 (2011).

224. Geinguenaud, F., Militello, V. & Arluison, V. Application of FTIR Spectroscopy to Analyze RNA Structure. in *RNA Spectroscopy: Methods and Protocols* (eds. Arluison, V. & Wien, F.) 119–133 (Springer US, 2020). doi:10.1007/978-1-0716-0278-2_10.

225. Giambruno, R., Rupert, J. & Zacco, E. Chapter 10 - Experimental methods to study protein–nucleic acid interactions. in *Advances in Protein Molecular and Structural Biology Methods* (eds. Tripathi, T. & Dubey, V. K.) 137–161 (Academic Press, 2022). doi:10.1016/B978-0-323-90264-9.00010-6.

226. Fuxreiter, M. Fuzziness in Protein Interactions—A Historical Perspective. *J. Mol. Biol.* **430**, 2278–2287 (2018).

227. Tartaglia, G. G. & Vendruscolo, M. The Zyggregator method for predicting protein aggregation propensities. *Chem Soc Rev* **37**, 1395–1401 (2008).

228. Tartaglia, G. G. *et al.* Prediction of Aggregation-Prone Regions in Structured Proteins. *J. Mol. Biol.* **380**, 425–436 (2008).

229. Agostini, F., Vendruscolo, M. & Tartaglia, G. G. Sequence-Based Prediction of Protein Solubility. *J. Mol. Biol.* **421**, 237–241 (2012).

230. Vendruscolo, M. & Fuxreiter, M. Sequence Determinants of the Aggregation of Proteins Within Condensates Generated by Liquid-liquid Phase Separation. *Protein Interact. Liq.-Liq. Phase Sep.* **434**, 167201 (2022).

231. Iqbal, K., Liu, F. & Gong, C.-X. Tau and neurodegenerative disease: the story so far. *Nat. Rev. Neurol.* **12**, 15–27 (2016).

232. Abskharon, R. *et al.* Cryo-EM structure of RNA-induced tau fibrils reveals a small C-terminal core that may nucleate fibril formation. *Proc. Natl. Acad. Sci.* **119**, e2119952119 (2022).

233. Hernández, F. *et al.* Tau Aggregation. *Neuroscience* **518**, 64–69 (2023).

234. Von Bergen, M. *et al.* Assembly of τ protein into Alzheimer paired helical filaments depends on a local sequence motif (306 VQIVYK 311) forming β structure. *Proc. Natl. Acad. Sci.* **97**, 5129–5134 (2000).

235. Wang, X. *et al.* The proline-rich domain and the microtubule binding domain of protein tau acting as RNA binding domains. *Protein Pept. Lett.* **13**, 679–685 (2006).

236. Moriarty, D. F., Vagts, S. & Raleigh, D. P. A role for the C-terminus of calcitonin in aggregation and gel formation: a comparative study of C-terminal fragments of human and salmon calcitonin. *Biochem. Biophys. Res. Commun.* **245**, 344–348 (1998).

237. Reches, M., Porat, Y. & Gazit, E. Amyloid fibril formation by pentapeptide and tetrapeptide fragments of human calcitonin. *J. Biol. Chem.* **277**, 35475–35480 (2002).

238. Luca, S., Yau, W.-M., Leapman, R. & Tycko, R. Peptide Conformation and Supramolecular Organization in Amylin Fibrils: Constraints from Solid State NMR. *Biochemistry* **46**, 13505–13522 (2007).

239. Pedersen, J. S., Dikov, D. & Otzen, D. E. N- and C-terminal hydrophobic patches are involved in fibrillation of glucagon. *Biochemistry* **45**, 14503–14512 (2006).

240. Török, M. *et al.* Structural and dynamic features of Alzheimer's Abeta peptide in amyloid fibrils studied by site-directed spin labeling. *J. Biol. Chem.* **277**, 40810–40815 (2002).

241. Lührs, T. *et al.* 3D structure of Alzheimer's amyloid-beta(1-42) fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 17342–17347 (2005).

242. Miake, H., Mizusawa, H., Iwatsubo, T. & Hasegawa, M. Biochemical Characterization of the Core Structure of α-Synuclein Filaments*. *J. Biol. Chem.* **277**, 19213–19219 (2002).

243. Der-Sarkissian, A., Jao, C. C., Chen, J. & Langen, R. Structural Organization of α-Synuclein Fibrils Studied by Site-directed Spin Labeling*. *J. Biol. Chem.* **278**, 37530–37535 (2003).

244. Del Mar, C., Greenbaum, E. A., Mayne, L., Englander, S. W. & Woods, V. L. Structure and properties of α-synuclein and other amyloids determined at the amino acid level. *Proc. Natl. Acad. Sci.* **102**, 15477–15482 (2005).

245. Lancaster, A. K., Nutter-Upham, A., Lindquist, S. & King, O. D. PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. *Bioinforma. Oxf. Engl.* **30**, 2501–2502 (2014).

246. Livi, C. M., Klus, P., Delli Ponti, R. & Tartaglia, G. G. catRAPID signature: identification of ribonucleoproteins and RNA-binding regions. *Bioinforma. Oxf. Engl.* **32**, 773–775 (2016).

247. Tartaglia, G. G., Cavalli, A., Pellarin, R. & Caflisch, A. The role of aromaticity, exposed surface, and dipole moment in determining protein aggregation rates. *Protein Sci.* **13**, 1939–1941 (2004).

248. Waxman, E. A., Mazzulli, J. R. & Giasson, B. I. Characterization of hydrophobic residue requirements for alpha-synuclein fibrillization. *Biochemistry* **48**, 9427–9436 (2009).

249. Arad, E. & Jelinek, R. Catalytic amyloids. *Trends Chem.* **4**, 907–917 (2022).

250. Loverix, S., Winqvist, A., Strömberg, R. & Steyaert, J. Mechanism of RNase T1: concerted triester-like phosphoryl transfer via a catalytic three-centered hydrogen bond. *Chem. Biol.* **7**, 651–658 (2000).

251. Hedstrom, L. Serine Protease Mechanism and Specificity. *Chem. Rev.* **102**, 4501–4524 (2002).

252. Schneider, F. Histidine in Enzyme Active Centers. *Angew. Chem. Int. Ed. Engl.* **17**, 583–592 (1978).

253. McDonald, A. G. & Tipton, K. F. Parameter Reliability and Understanding Enzyme Function. *Molecules* **27**, 263 (2022).

254. Athel Cornish-Bowden. *Fundamentals of Enzyme Kinetics, 4th Edition*. (Wiley-Blackwell, 2012).

255. Schenker, P. & Baici, A. Paradoxical interactions between modifiers and elastase-2. *FEBS J.* **277**, 2486–2495 (2010).

256. Vladimir Leskovac. *Comprehensive Enzyme Kinetics*. (Springer New York, NY, 2003).

257. Lorey, S. *et al.* Different modes of dipeptidyl peptidase IV (CD26) inhibition by oligopeptides derived from the N-terminus of HIV-1 Tat indicate at least two inhibitor binding sites. *Eur. J. Biochem.* **270**, 2147–2156 (2003).

258. Stephens, A. D., Zacharopoulou, M. & Schierle, G. S. K. The Cellular Environment Affects Monomeric α-Synuclein Structure. *Trends Biochem. Sci.* **44**, 453–466 (2019).

259. Knowles, T. P. J. & Buehler, M. J. Nanomechanics of functional and pathological amyloid materials. *Nat. Nanotechnol.* **6**, 469–479 (2011).

260. Hill, E. K., Krebs, B., Goodall, D. G., Howlett, G. J. & Dunstan, D. E. Shear flow induces amyloid fibril formation. *Biomacromolecules* **7**, 10–13 (2006).

261. Morris, A. M., Watzky, M. A., Agar, J. N. & Finke, R. G. Fitting neurological protein aggregation kinetic data via a 2-step, minimal/"Ockham's razor" model: the Finke-Watzky mechanism of nucleation followed by autocatalytic surface growth. *Biochemistry* **47**, 2413–2427 (2008).

262. Kumari, P. *et al.* Structural insights into α-synuclein monomer–fibril interactions. *Proc. Natl. Acad. Sci.* **118**, e2012171118 (2021).

263. Sawaya, M. R., Hughes, M. P., Rodriguez, J. A., Riek, R. & Eisenberg, D. S. The expanding amyloid family: Structure, stability, function, and pathogenesis. *Cell* **184**, 4857–4873 (2021).

264. Amyloid Atlas 2023. https://people.mbi.ucla.edu/sawaya/amyloidatlas/.

265. Stojkovska, I. & Mazzulli, J. R. Detection of pathological alpha-synuclein aggregates in human iPSC-derived neurons and tissue. *STAR Protoc.* **2**, 100372 (2021).

266. Ray, S. *et al.* α-Synuclein aggregation nucleates through liquid–liquid phase separation. *Nat. Chem.* **12**, 705–716 (2020).

267. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-Liquid Phase Separation in Biology. *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).

268. Hardenberg, M., Horvath, A., Ambrus, V., Fuxreiter, M. & Vendruscolo, M. Widespread occurrence of the droplet state of proteins in the human proteome. *Proc. Natl. Acad. Sci.* **117**, 33254–33262 (2020).

269. Hardenberg, M. C. *et al.* Observation of an α-synuclein liquid droplet state and its maturation into Lewy body-like assemblies. *J. Mol. Cell Biol.* **13**, 282–294 (2021).

270. Piroska, L. *et al.* α-Synuclein liquid condensates fuel fibrillar α-synuclein growth. *Sci. Adv.* **9**, eadg5663 (2023).

271. Dhakal, S. *et al.* α-Synuclein emulsifies TDP-43 prion-like domain—RNA liquid droplets to promote heterotypic amyloid fibrils. *Commun. Biol.* **6**, 1–16 (2023).

272. Rupert, J., Monti, M., Zacco, E. & Tartaglia, G. G. A Computational Approach Reveals the Ability of Amyloids to Sequester RNA: the Alpha Synuclein Case. 2022.09.20.508776 Preprint at https://doi.org/10.1101/2022.09.20.508776 (2023).

273. Perego, E. *et al.* Single-photon microscopy to study biomolecular condensates. *Nat. Commun.* **14**, 8224 (2023).