

# **How sequence specific changes in the N-terminus of synucleins control their aggregation**

**Sabine Marianne Ulamec**

University of Leeds

Astbury Centre for Structural and Molecular Biology

Submitted in accordance with the requirements for the degree of

*Doctor of Philosophy*

March 2022



# Declaration

The candidate confirms that the submitted work is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated overleaf. The candidate confirms the appropriate credit has been given within the thesis where reference has been made to the work of others. This copy has been supplied on the understanding that it is copyright material and that no quotation from this thesis may be published without proper acknowledgement.

© 2022 The University of Leeds and Sabine Marianne Ulamec



# Jointly authored publications

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others:

**Chapter 1** includes work from the following publication:

Ulamec, S. M., Brockwell D. J., and Radford, S. E., "Looking beyond the core: The role of flanking regions in the aggregation of amyloidogenic peptides and proteins." *Frontiers in Neuroscience* 1216, (2020).

For this publication, the author contribution statement reads:

SMU, DJB, and SER wrote the manuscript. All authors contributed to the article and approved the submitted version.

**Chapter 3** includes work from the following publication:

Doherty, C. P. A.\*, Ulamec, S. M\*., Maya-Martinez, R., Good, S. C., Makepeace, J., Khan, G. N., van Oosten Hawle, P., Radford, S. E., and Brockwell, D. J., "A short motif in the N-terminal region of  $\alpha$ -synuclein is critical for both aggregation and function". *Nature structural & molecular biology*, 27, 249-259, (2020).

\* These authors contributed equally to this work.

For this publication, the author contribution statement reads:

C.P.A.D. and S.M.U. prepared samples and designed and performed fluorescence, NMR, EM and other biochemical studies. J.M., S.C.G. and P.v.O.-H. performed the experiments with *C. elegans*. C.P.A.D., S.M.U. and G.N.K. performed CD experiments. R.M.-M. performed NMR assignments and assisted with NMR data analysis and interpretation. S.E.R. and D.J.B. developed the ideas and supervised the work. All authors contributed to the preparation of the manuscript.



# Acknowledgments

Most importantly, I would like to thank my supervisors, Prof David Brockwell and Prof Sheena Radford, for their continued guidance and support throughout my PhD. Thank you for your enthusiasm and excitement that kept me going. I am more than grateful for all your time and effort that has gone into this project.

I have been extremely lucky, to be part of a great PhD program. I would like to acknowledge the Wellcome Trust for its funding and I would like to thank Prof Alan Berry for offering me this PhD position and running this amazing program. I am very grateful for all the wonderful people that I have met through this program.

The past four years in the Brockwell/Radford lab have been an amazing time because of every single person. Thank you for all the support in the lab, fun lunch and coffee breaks, great times in the pub and the best trips to Dublin. Nothing would have been the same without Dr Ciaran Doherty, who got me started in the lab and patiently explained everything about alpha-synuclein. Thank you Ciaran for being this great person that has taught me so much. I would also like to thank Dr Roberto Maya-Martinez for his fantastic support for all the NMR studies and providing tequila whenever needed. This project would not have been possible without you. Further thanks goes to Nasir Khan, the heart and soul of the lab, for always keeping me going by offering biscuits, pep talks or a helping hand in the lab. I would also like to acknowledge the fantastic work of Helen McAllister, who keeps us all organised.

Huge thank you to Emily Byrd and Katy Dewison for many useful discussions about synuclein as well as the MBiol students Emily Prescott and Keri Moore for contributing to this project. Seeing the synuclein team grow and being a part of it has been a great pleasure. Additional thanks goes to all the people who have helped, supported, and offered me advice during my project including Dr Yong Xu, Dr Leon Willis, Dr Martin Walko, Dr Bob Schiffrin, Dr Atenas Posada-Borbon, Dr Madeline Brown, and Dr Chalmers Chau.

I would like to acknowledge all collaboration partners on my project. Thank you Patricija van Oosten-Hawle (University of Leeds), Prof Vladimir Buchman (University of Cardiff), George Heath (University of Leeds), and Prof Tiago Outeiro (University of Göttingen). I have learned a lot through your input and different points of view on this work.

Further thanks goes to the amazing biostructure facilities in Leeds. This project would not have been the same without the help of the NMR, EM, bioimaging and Mass Spec facility.

Being away from my home and family has sometimes been hard, but I couldn't have found better friends here in Leeds to support me and always be there for me. Thank you Dr Anna Higgins and Dr Jess Ebo for being among the first people to welcome me here. I know I can always count on you and I have enjoyed every second we spent together. Thanks to Dr Nicolas Guthertz (and Mélanie) for being my allies in this country and for inviting me over for dinner many times. Thank you James Whitehouse for always making me laugh and Dr Paul White for the best after-work drinks together. Last but not least, thank you Jane Thorn for being the best 'non-Uni' friend one can ask for. I hope to stay in contact with all of you and see you soon in the UK, Germany, Austria or anywhere else.

I would further like to thank Nina L'Hoest. Thank you for always staying in touch, for remembering every important presentation and meeting, wishing me luck beforehand and for telling me that I can do it.

Lastly, a huge thank you to the people who have supported me most. To my family, Mama, Papa and Thomas, thank you for always believing in me, I could not have done this without your love and support. Thank you for coming to visit me in Leeds so many times, spending hours on the phone with me and always being there for me. I would also like to thank you, David, for pushing me to be the best version of myself. Thank you for convincing me to go to Leeds with you and for the great time we spent here. I am excited for what is coming next and grateful that I know we will do it together.



# Abstract

The intrinsically disordered protein (IDP) alpha-synuclein ( $\alpha$ Syn) is known to be involved in neurodegenerative disorders such as Parkinson's disease (PD) which affects more than 10 million people worldwide. Patients show multiple motor (e.g. tremor, lack of coordination) and non-motor (e.g. depression, anxiety) symptoms, as well as pathological indicators, characterised by the loss of dopaminergic neurons and the formation of aggregated  $\alpha$ Syn-containing Lewy bodies in the brain. Understanding the process of amyloid formation from the highly dynamic  $\alpha$ Syn monomer to highly ordered fibrils, formed by  $\beta$ -sheets, is therefore crucial.

The central part of the  $\alpha$ Syn sequence contains the non-amyloid  $\beta$ -component (NAC) region crucial for aggregation. In this thesis, the importance of regions N-terminally flanking NAC is demonstrated and characterised in detail. Using bioinformatics, aggregation assays, *in vivo* systems, mutations, and NMR led to the identification of a 7-residue sequence, named P1 (residues <sup>36</sup>GVLVGS<sup>42</sup>) that modulates aggregation and function (in synergy with P2 (residues <sup>45</sup>KEGVVHGVATVAE<sup>57</sup>)) of full-length  $\alpha$ Syn. Deletion or substitution (with a Gly-Ser-linker) of these motif(s) drastically slow down fibril formation and interrupt the ability of remodelling membranes. It could be shown that the P1 (and P2) region acts as a 'master controller' via transient intra- and inter-molecular interactions throughout the protein.

An in-depth analysis of the P1 region narrowed down the key residues to control self-assembly into amyloid structures to be L38, Y39 and S42. The results demonstrated a remarkable sequence specificity with different substitutions at the same position resulting in enhanced or retarded fibril formation. Interestingly, the two amino acid substitutions in P1 between  $\alpha$ Syn and its significantly less aggregation-prone paralogue gamma-synuclein ( $\gamma$ Syn), L38M and S42A, were found to inhibit fibrillation of  $\alpha$ Syn. Changing the P1 sequence in  $\gamma$ Syn to that of  $\alpha$ Syn does not switch on aggregation, highlighting a complex interplay of interactions required for amyloid formation.

Strikingly, an Amyotrophic Lateral Sclerosis (ALS) linked substitution located within P1 (M38I) was found in  $\gamma$ Syn, and *in vitro* and in cell studies confirmed that this variant has an increased potential to form fibrils. Further studies on different substitutions at this site in  $\gamma$ Syn suggest that  $\beta$ -sheet forming amino acids at residue 38 drive fibril formation, with Ile and Val enhancing aggregation and Met, Ala, Leu to slowing down fibrillation.

Taken together, the work presented in this thesis highlights the crucial role of the P1 region to modulate fibril formation in synucleins. The identification of this 'master controller' region offers a new target for therapeutics preventing aggregation in disease.



# Table of Contents

<b>Declaration</b> .....	<b>iii</b>
<b>Jointly authored publications</b> .....	<b>v</b>
<b>Acknowledgements</b> .....	<b>vii</b>
<b>Abstract</b> .....	<b>ix</b>
<b>Table of Contents</b> .....	<b>xi</b>
<b>List of Figures</b> .....	<b>xv</b>
<b>List of Tables</b> .....	<b>xxi</b>
<b>List of abbreviations</b> .....	<b>xxiii</b>
<b>List of amino acids</b> .....	<b>xxvii</b>
<b>Chapter 1</b> .....	<b>1</b>
1.1 Protein Folding and Misfolding .....	1
1.2 Amyloid .....	5
1.2.1 Historical perspective .....	6
1.2.2 Structure of amyloids .....	7
1.2.3 Mechanism of amyloid formation.....	13
1.2.4 Toxicity in amyloid proteins .....	15
1.2.5 The role of flanking regions .....	16
1.3 Intrinsically disordered proteins .....	21
1.4 Alpha-synuclein .....	24
1.4.1 Physiological function.....	24
1.4.2 Synucleopathies.....	26
1.4.3 Structural properties.....	30
1.4.4 Lipid interactions .....	37
1.5 Synuclein paralogues .....	39
1.5.1 Gamma-synuclein .....	40
1.6 Thesis aim.....	42
<b>Chapter 2</b> .....	<b>45</b>
2.1 Chemicals .....	45
2.2 Bacterial strains.....	45
2.3 Vectors .....	45
2.4 Peptides .....	46
2.5 Molecular Biology .....	46
2.5.1 Site directed mutagenesis .....	46
2.5.2 Subcloning .....	49

2.5.3 Transformation .....	49
2.5.4 Sequencing .....	49
2.6 Protein expression and purification .....	50
2.6.1 $\alpha$ Syn .....	50
2.6.2 $\gamma$ Syn .....	51
2.6.3 Chimeric proteins .....	51
2.7 General biochemical techniques .....	52
2.7.1 SDS-PAGE .....	52
2.7.2 Agarose gel electrophoresis .....	52
2.7.3 Immunoblotting analysis .....	53
2.7.4 Preparation of disulfide locked dimeric $\alpha$ -synuclein species .....	53
2.7.5 SEC MALS .....	53
2.7.6 Dynamic Light Scattering .....	54
2.7.7 MTSL-labelling .....	54
2.7.8 Liposome Preparation .....	55
2.7.9 Peptide synthesis .....	55
2.7.10 Limited proteolysis .....	55
2.8 Aggregation assays .....	56
2.8.1 Seed preparation .....	56
2.8.2 Thioflavin T assay .....	56
2.8.3 ThT assay data analysis .....	57
2.8.4 Pelleting assay .....	57
2.8.5 Characterisation of supernatant and pellet from pelleting assay .....	58
2.9 NMR experiments .....	59
2.9.1 NMR backbone assignment .....	59
2.9.2 Paramagnetic relaxation enhancement experiments .....	59
2.9.3 NMR experiments on liposomes .....	61
2.9.4 NMR experiments in the presence of peptides .....	61
2.10 Circular dichroism .....	62
2.10.1 Binding to liposomes .....	62
2.10.2 Binding to peptides .....	64
2.11 Mass Spectrometry analysis .....	64
2.12 Transmission Electron Microscopy .....	64
2.13 Atomic force microscopy (AFM) .....	64
2.14 Computational techniques .....	65

2.14.1 <i>In silico</i> identification of aggregation-prone regions .....	65
2.14.2 ConSurf analysis .....	65
2.14.3 Genomic analysis of $\gamma$ Syn .....	65
2.15 Biological techniques .....	66
2.15.1 Cellular viability assays .....	66
2.15.2 Inclusion formation in cells .....	67
2.15.3 Experiments with <i>C. elegans</i> .....	67
<b>Chapter 3</b> .....	<b>71</b>
3.1 Introduction .....	71
3.2 Identification of a N-terminal aggregation-prone motif .....	72
3.3 Engineering recombinant $\alpha$ Syn .....	74
3.3.1 Molecular biology .....	74
3.3.2 Expression and purification .....	74
3.4 N-terminal motif controls $\alpha$ Syn aggregation <i>in vitro</i> .....	76
3.5 The N-terminal motifs control aggregation <i>in vivo</i> .....	84
3.6 Assignment and Evaluation of NMR experiment conditions for analysis of $\alpha$ Syn variants .....	87
3.6.1 The effect of solution conditions on NMR spectra .....	88
3.7 Intra- and inter-molecular interactions involving N-terminal motifs promote aggregation .....	90
3.7.1 Intramolecular interactions with P1 and P2 determine aggregation ....	90
3.7.2 Intermolecular interactions with P1 determine aggregation .....	98
3.8 P1 and P2 are conserved in synucleins .....	100
3.9 Importance of P1 and P2 for membrane remodelling .....	100
3.10 Discussion .....	105
<b>Chapter 4</b> .....	<b>111</b>
4.1 Introduction .....	111
4.2 The P1 sequence drives aggregation .....	114
4.3 Binding mode between $\alpha$ Syn and P1-peptide .....	120
4.4 Specific residues within P1 are responsible for controlling aggregation in $\alpha$ Syn 126	
4.4.1 Alanine scan to identify key residues in P1 .....	126
4.4.2 The variants Y39A and S42A form toxic oligomers .....	128
4.4.3 Capacity to seed $\alpha$ Syn WT fibrils for Ala-scan variants .....	133
4.5 Residue 38 in P1 is a key regulator of $\alpha$ Syn aggregation .....	136
4.6 The role of the P1 region in $\gamma$ Syn aggregation .....	139

4.7 Single residues in $\alpha$ Syn fine-tune aggregation <i>in vivo</i> .....	140
4.8 Interactions between N- and C-terminal regions control synuclein aggregation .....	143
4.9 Membrane binding of L38M and $\gamma$ Syn WT .....	144
4.10 Discussion .....	146
<b>Chapter 5</b> .....	<b>151</b>
5.1 Introduction.....	151
5.2 Reduced aggregation propensity for $\gamma$ Syn might be driven by sequence changes in NAC .....	153
5.3 Identification of an ALS-associated substitution outside NAC .....	157
5.4 Protein expression and purification .....	158
5.5 Presence of isoleucine at position 38 dramatically increase $\gamma$ Syn propensity to form amyloid structures <i>in vitro</i> and in cell culture .....	159
5.6 Met38Ile substitution does not change toxicity in <i>C. elegans</i> .....	162
5.7 Ile38 aggregates under all biologically relevant conditions.....	164
5.8 High sequence specificity is required to accelerate $\gamma$ Syn fibril formation..	166
5.9 $\gamma$ I38 shows a distinct fibril architecture compared with $\alpha$ Syn WT .....	167
5.10 PRE NMR to investigate conformational changes between $\gamma$ WT and $\gamma$ I38 .....	171
5.11 Discussion.....	172
<b>Chapter 6</b> .....	<b>175</b>

# List of Figures

Figure 1.1: Key events in the field of protein structures...	1
Figure 1.2: Schematic of some proposed folding mechanisms discussed in this work..	3
Figure 1.3: Schematic energy landscape of folding and aggregation of proteins. ....	4
Figure 1.4: Schematic presentation of the proteostasis network. ....	5
Figure 1.5: Number of publication in PubMed associated with 'amyloid' from 1900 to date.....	6
Figure 1.6: Summary of fibril diffraction and morphology.....	8
Figure 1.7: Schematic of fibril architecture on the example of $\alpha$ Syn fibrils.....	9
Figure 1.8: Example of wide variety of amyloid fibril structures from different proteins. ....	10
Figure 1.9: Tau folds from fibrils extracted from patients with different diseases. ....	11
Figure 1.10: APRs comprise only a small part of the amyloid core.....	12
Figure 1.11: Schematic of spontaneous/ <i>de novo</i> and seeded amyloid formation. ....	13
Figure 1.12: Schematic illustrating the microscopic processes of fibril formation including secondary aggregation pathways.....	14
Figure 1.13: Factors affecting the conformational properties and interactions with amyloid precursors that retard or accelerate fibril growth. ....	15
Figure 1.14: Mechanisms of toxicity which have been observed for oligomers and fibrils of amyloid proteins.....	16
Figure 1.15: Schematic of interactions between the dynamically disordered regions of fibrils and other molecules. ....	21
Figure 1.16: Energy landscape of a structured proteins vs IDPs. ....	22
Figure 1.17: Charge-hydrophathy phase space .....	22
Figure 1.18: IDPs in disease.....	24
Figure 1.19: Physiological role of $\alpha$ Syn at pre-synaptic termini illustrating the various regulatory functions $\alpha$ Syn exerts over synaptic vesicles.....	26
Figure 1.20: Significance and symptoms of Parkinson's disease. ....	30
Figure 1.21: Morphological and structural characterisation of $\alpha$ Syn monomers, oligomers and fibrils .....	31
Figure 1.22: Sequence identity of $\alpha$ Syn .....	32
Figure 1.23: Oligomer structures and effect on membrane binding .....	36
Figure 1.24: Examples of $\alpha$ Syn fibril structures determined by cryoEM.....	37
Figure 1.25: Summary of $\alpha$ Syn binding to membranes. ....	38
Figure 1.26: Schematic of liposome clustering by $\alpha$ Syn. ....	39

Figure 1.27: Sequence alignment between the synuclein family members $\alpha$ Syn, $\beta$ Syn and $\gamma$ Syn.....	40
Figure 2.1: Molecular weight calibrants for size exclusion chromatography.....	51
Figure 2.2: MTSL labelling reaction.....	54
Figure 2.3: Loading control for fibril yield analysis.....	58
Figure 2.4: Schematic of PRE experiment.....	59
Figure 2.5: Schematic of liposome binding experiment..	61
Figure 2.6: Schematic of tissues and lifecycle of <i>C. elegans</i> .....	68
Figure 3.1: Aggregation and solubility profiles of $\alpha$ Syn.....	73
Figure 3.2: Expression and purification of $\alpha$ Syn WT.....	75
Figure 3.3: The aggregation kinetics of 100 $\mu$ M WT $\alpha$ Syn; $\Delta$ P1; $\Delta$ P2 or $\Delta\Delta$ variants. .....	76
Figure 3.4: The aggregation kinetics of 100 $\mu$ M $\Delta$ N7.....	77
Figure 3.5: Aggregation kinetics of 100 $\mu$ M $\Delta$ C1, P1-GS, and P1P2-GS.....	78
Figure 3.6: The aggregation kinetics of 100 $\mu$ M WT $\alpha$ Syn; $\Delta$ P1; $\Delta$ P2 or $\Delta\Delta$ variants at low salt concentration (20 mM NaCl).....	79
Figure 3.7: Aggregation kinetics and negative-stain TEM images of endpoint (140 h) aggregates of 100 $\mu$ M V40C, V52C or A140C monomers or dimers.....	81
Figure 3.8: Preparation of disulfide locked $\alpha$ Syn variants.....	82
Figure 3.9: Cross-seeding $\alpha$ Syn variants using seeds created from $\alpha$ Syn WT. ....	83
Figure 3.10: Aggregation kinetics of co-incubated $\alpha$ Syn WT and $\Delta$ P1 monomer...	84
Figure 3.11: Deletion of P1 or both P1 and P2 in <i>C. elegans</i> expressing $\alpha$ Syn::YFP suppresses aggregation and proteotoxicity. ....	86
Figure 3.12: Assignment of $\alpha$ Syn WT and $\Delta\Delta$ at pH 4.5.....	88
Figure 3.13: Comparison of chemical shifts for different buffer conditions and synuclein variants.....	89
Figure 3.14: Comparison of separate 1H and 15N chemical shifts for different buffer conditions and synuclein variants shown in Figure 3.13. ....	90
Figure 3.15: Ascorbic acid titration to identify ideal conditions for reducing MTSL in PRE experiments and effect of free MTSL. ....	92
Figure 3.16: Overlaid paramagnetic and diamagnetic spectra for WT $\alpha$ Syn labelled at positions A18C, A90C or A140C. ....	93
Figure 3.17: NMR PRE data of $\alpha$ Syn WT with MTSL labels at position A18C, A90C and A140C to observe intramolecular interactions through the protein.....	94
Figure 3.18: Control intermolecular PRE experiments of WT $\alpha$ Syn in low (20 mM) and high (200 mM) salt (NaCl) at pH 4.5.....	95
Figure 3.19: NMR PRE data of $\alpha$ Syn $\Delta\Delta$ with MTSL labels at position A18C, A90C and A140C to observe intramolecular interactions through the protein.....	96
Figure 3.20: Intramolecular interactions for $\alpha$ Syn P1P2-GS.....	98



Figure 3.21: Intramolecular interactions between P1, P2 region and C-terminal region control aggregation. ....	99
Figure 3.22: Sequence conservation within the synuclein family calculated using ConSurf. ....	100
Figure 3.23: CD binding assays of $\alpha$ Syn WT, $\Delta\Delta$ and P1P2-GS to DMPS liposomes. ....	101
Figure 3.24: Residues involved in binding to DMPS liposomes investigated by NMR.. ....	103
Figure 3.25: Liposome induced aggregation of $\alpha$ Syn and membrane remodelling. ....	104
Figure 3.26: Summary of processes involving the N-terminal region of $\alpha$ Syn.....	106
Figure 3.27: Schematic of structural features of $\alpha$ Syn including the P1 and P2 regions. ....	108
Figure 4.1: Sequence alignment between the synuclein family members $\alpha$ Syn, $\beta$ Syn and $\gamma$ Syn.....	113
Figure 4.2: <i>In silico</i> analysis of the $\alpha$ -, $\beta$ - and $\gamma$ Syn sequences. Zyggregator, CamSol and ZipperDB profiles for $\alpha$ Syn and $\beta$ Syn and $\gamma$ Syn.....	113
Figure 4.3: pH dependent aggregation of $\alpha$ Syn WT and $\Delta$ P1.....	115
Figure 4.4: Aggregation kinetics of pH titration for $\alpha$ Syn WT and $\Delta$ P1. ....	116
Figure 4.5: TEM images at the endpoint of ThT aggregation assay with $\Delta$ P1. ....	116
Figure 4.6: Aggregation kinetics of WT $\alpha$ Syn, $\Delta$ P1 and $\Delta\Delta$ in the presence of the peptides P1 or P1-SG. ....	118
Figure 4.7: The synthetic peptides P1 and P1-SG do not aggregate in isolation under the conditions employed. ....	119
Figure 4.8: Purification and characterisation of MTSL-labelled Cys-P1 peptide by LC-MS. ....	121
Figure 4.9: The binding of P1-peptide alters the conformational ensemble of WT $\alpha$ Syn. ....	122
Figure 4.10: Interaction between $\alpha$ Syn WT and $\Delta$ P1 with the P1 peptide characterised by native nESI-MS ....	123
Figure 4.11: Cross-peaks of $^1\text{H}$ - $^{15}\text{N}$ HSQC spectra to identify differences between pH titration and P1-peptide addition for $\alpha$ Syn WT. ....	123
Figure 4.12: Binding of peptide P1 alters the conformational ensemble of $\Delta$ P1. ..	124
Figure 4.13: Aggregation kinetics of WT $\alpha$ Syn, $\Delta$ P1 and $\Delta$ N7 in the presence of the N7-peptide. ....	125
Figure 4.14: Alanine scan to identify residues in P1 region important for aggregation. ....	127
Figure 4.15: Aggregation kinetics of $\alpha$ Syn WT and alanine variants Y39A and S42A at acidic pH.....	128
Figure 4.16: AFM study of $\alpha$ Syn WT, Y39A, and S42A. ....	129

Figure 4.17: Exemplary AFM image of S42A to highlight fibril, oligomer and monomer species.....	130
Figure 4.18: Analysis of ThT end-point samples and efficiency of pelleting assay for $\alpha$ Syn WT, Y39A, and S42A. ....	131
Figure 4.19: Cell viability assays of Y39A and S42A oligomers.....	133
Figure 4.20: Cross-seeding $\alpha$ Syn variants with pre-formed fibrils of $\alpha$ Syn WT. ...	135
Figure 4.21: Aggregation assay of $\alpha$ Syn WT, $\beta$ Syn WT and $\gamma$ Syn WT demonstrating the low aggregation propensities of $\beta$ Syn and $\gamma$ Syn. ....	136
Figure 4.22: $\alpha$ Syn aggregation rate is dependent on the identity of residue 38. ..	138
Figure 4.23: Sequence alterations in P1 do not result in aggregation of $\gamma$ Syn. ....	139
Figure 4.24: Cross-seeding $\gamma$ Syn constructs with 10% (mol/mol) preformed $\alpha$ Syn WT seeds does not result in fibril formation. ....	140
Figure 4.25: Effect of single point mutations in $\alpha$ Syn and $\gamma$ Syn in the body wall muscle of <i>C. elegans</i> for puncta formation. ....	141
Figure 4.26: Effect of single point mutations in $\alpha$ Syn and $\gamma$ Syn in the body wall muscle of <i>C. elegans</i> for motility.....	142
Figure 4.27: Agarose gel of digested pPD30.38 vector and $\gamma$ Syn WT insert.....	143
Figure 4.28: Aggregation kinetics of chimeric proteins in which the N-terminal, NAC and C-terminal regions of $\alpha$ Syn and $\gamma$ Syn are swapped .....	144
Figure 4.29: Membrane binding of $\alpha$ Syn L38M and $\gamma$ Syn WT to DMPS liposomes. ....	146
Figure 4.30: Summary of all cryo-EM high resolution structures of $\alpha$ Syn solved to date. ....	148
Figure 5.1: Sequence alignment between $\alpha$ Syn WT and $\gamma$ Syn WT .....	152
Figure 5.2: <i>In silico</i> analysis of $\alpha$ Syn and $\gamma$ Syn aggregation propensity and solubility. ....	154
Figure 5.3: LC-ES-MS analysis of $\alpha$ NAC and $\gamma$ NAC synthesised peptides.....	155
Figure 5.4: Aggregation kinetics of NAC peptide (residues 61-95) for $\alpha$ Syn and $\gamma$ Syn at pH 4.5 and pH 7.5. ....	156
Figure 5.5: Net charge of full-length $\alpha$ Syn (WT) and $\gamma$ Syn (WT).....	156
Figure 5.6: Schematic of SNCG genome sequencing of the central European Russia population .....	157
Figure 5.7: Expression and purification of $\gamma$ Syn. ....	158
Figure 5.8: <i>In vitro</i> aggregation assay of genetic variants associated with ALS patients. ....	160
Figure 5.9: Inclusions formed by $\gamma$ Syn variants in human H4 cells .....	162
Figure 5.10: Expressing $\gamma$ I38 in worm muscle cells does not affect the toxicity ...	163
Figure 5.11: $\gamma$ I38 switches on aggregation with a strong buffer dependence.....	164
Figure 5.12: Aggregation kinetics for $\gamma$ WT and $\gamma$ I38 .....	166

Figure 5.13: The aggregation of $\gamma$ Syn is highly residue specific .....	167
Figure 5.14: Distinct seeding capacity for $\alpha$ Syn WT and $\gamma$ l38 pre-formed fibrils..	169
Figure 5.15: Protease K treatment results in distinct digestion patterns for $\alpha$ WT and $\gamma$ 38l fibrils.....	170
Figure 5.16: NMR PRE data of intramolecular interactions of $\gamma$ WT and $\gamma$ l38 with MTSL spin label at position A90C.....	171
Figure 6.1. Schematic of the findings in this thesis that N-terminal flanking regions are important for aggregation. ....	176
Figure 6.2: Schematic of predicting conformational ensembles and function/disease aspects using machine learning based on multidimensional data .....	180
Figure 6.3: Schematic of $\alpha$ Syn fibrillation on liposome surfaces.....	181
Figure 6.4: Conformational ensemble and small molecule binding to $\alpha$ Syn.....	182
Figure 6.5: Advantages and disadvantages of the most common model organisms used in biomedical research.....	183



# List of Tables

Table 1.1: Roles of the flanking regions in monomers of different amyloid protein sequences .....	17
Table 1.2: Summary of diseases associated with $\alpha$ Syn toxicity.....	27
Table 1.3: Summary of identified synucleopathy associated mutations in the SNCA gene causing missense mutations in $\alpha$ Syn .....	29
Table 1.4: Regions of intramolecular interactions within monomeric $\alpha$ Syn.....	33
Table 2.1: Sequence of peptides used in this study and their producers.....	46
Table 2.2: Oligonucleotides used in this study.....	47
Table 2.3: Temperature cycle for PCR for site directed mutagenesis.....	48
Table 2.4: SDS-PAGE gel compositions.....	52
Table 3.1: All purified $\alpha$ Syn variants, expected molecular masses and measured molecular masses of the proteins used in this chapter.....	75
Table 3.2: Aggregation kinetics of $\alpha$ Syn variants.....	80
Table 3.3: Aggregation kinetics for dimer locked proteins.....	81
Table 3.4: Aggregation kinetics of seeding reaction with 10 % $\alpha$ Syn WT seeds.....	83
Table 3.5: Aggregation kinetics of $\alpha$ Syn WT and $\Delta\Delta$ at pH 4.5 at low (20 mM) and high (200 mM) NaCl.....	97
Table 3.6: Expected and measured $\alpha$ -helical structure of $\alpha$ Syn WT, $\Delta\Delta$ and P1P2-GS in absence and presence of saturating conditions of DMPS liposomes (100:1 [DMPS]:[ $\alpha$ Syn]) .....	102
Table 3.7: Summary of processes involving the N-terminal region of $\alpha$ Syn highlighted in Figure 3.25.....	107
Table 4.1: Aggregation kinetics of $\Delta$ P1 .....	115
Table 4.2: Aggregation kinetics of $\alpha$ Syn WT, $\Delta$ P1, and $\Delta\Delta$ in the presence of P1- or P1-SG peptides.....	119
Table 4.3: Aggregation kinetics in the presence of N7 peptide.....	125
Table 4.4: All purified $\alpha$ Syn and $\gamma$ Syn variants, expected molecular masses and measured molecular masses of the proteins used in this chapter.....	126
Table 4.5: Aggregation kinetics of alanine scan of P1 region in $\alpha$ Syn.....	128
Table 4.6: Physical properties of aggregates imaged by AFM.....	130
Table 4.7: Aggregation kinetics of different synuclein family members ( $\alpha$ -, $\beta$ -, $\gamma$ Syn) and $\alpha$ Syn variants mutated at position 38 .....	138
Table 4.8: Aggregation kinetics for $\gamma$ Syn constructs.....	140
Table 4.9: Aggregation kinetics of chimeric synuclein variants.....	144
Table 4.10: Binding thermodynamics of synuclein variants to DMPS liposomes.....	145

Table 5.1: All purified $\alpha$ Syn and $\gamma$ Syn variants, expected molecular masses and measured molecular masses of the proteins used in this chapter.....	158
Table 5.2: Aggregation kinetics of ALS gene polymorphs M38/I38 and E110/V110. ....	161
Table 5.3: Aggregation kinetics of pH dependence of $\gamma$ Syn WT and $\gamma$ I38. ....	165
Table 5.4: Aggregation kinetics of substitutions at residue 38 in $\gamma$ Syn.....	167
Table 5.5: Aggregation kinetics of seeding reactions with $\alpha$ Syn WT and $\gamma$ I38 seeds. ....	169
Table 6.1: Overview of open questions with respect to this thesis and future experiments. ....	177

# List of abbreviations

<b>3D</b>	Three dimensional
<b>A<math>\beta</math></b>	Amyloid beta
<b>AD</b>	Alzheimer's Disease
<b>AEC</b>	Anion exchange chromatography
<b>AFM</b>	Atomic force microscopy
<b>ALS</b>	Amyotrophic Lateral Sclerosis
<b>APR</b>	Aggregation prone region
<b><math>\alpha</math>Syn</b>	Alpha-synuclein
<b>ATP</b>	Adenosine triphosphate
<b><math>\beta</math>2m</b>	Beta-2-microglobulin
<b><math>\beta</math>Syn</b>	Beta-synuclein
<b>BBPS</b>	Body bends per second
<b><i>C. elegans</i></b>	<i>Caenorhabditis elegans</i>
<b>CBD</b>	Corticobasal degeneration
<b>CD</b>	Circular dichroism
<b>CID</b>	Collision-induced dissociation
<b>CJD</b>	Creutzfeldt-Jakob disease
<b>CLEM</b>	Correlative light and electron microscopy
<b>CPEB</b>	Cytoplasmic polyadenylation element binding
<b>cryoEM</b>	Cryo-electron microscopy
<b>cryoET</b>	Cryo-electron tomography
<b>CSP</b>	Chemical shift perturbation
<b>CTE</b>	Chronic traumatic encephalopathy
<b>DNA</b>	Deoxyribonucleic acid
<b>DAPI</b>	4',6-Diamidino-2-phenylindole
<b>DLB</b>	Dementia with Lewy bodies
<b>DLS</b>	Dynamic light scattering
<b>DMPS</b>	1,2-Dimyristoyl-sn-glycero-3-phosphoserine
<b>DMS</b>	Deep mutational scanning
<b>DMSO</b>	Dimethyl sulfoxide
<b>DTT</b>	Dithiothreitol
<b><i>E. coli</i></b>	<i>Escherichia coli</i>
<b>EM</b>	Electron microscopy
<b>EPR</b>	Electron paramagnetic resonance

<b>ESI-MS</b>	Electrospray ionisation mass spectrometry
<b>FRET</b>	Fluorescence resonance energy transfer
<b>FTIR</b>	Fourier transform infrared spectroscopy
<b>HDX</b>	Hydrogen/deuterium exchange
<b>HP-SEC</b>	High-performance size exclusion chromatography
<b>HSQC</b>	$^1\text{H}$ - $^{15}\text{N}$ Heteronuclear single quantum coherence
<b>IAPP</b>	Islet amyloid polypeptide
<b>IDP</b>	Intrinsically disordered protein
<b>IDR</b>	Intrinsically disordered region
<b>IPTG</b>	Isopropyl $\beta$ - d-1-thiogalactopyranoside
<b>KLD</b>	Kinase, ligase, Dnpl
<b>LB</b>	Lysogeny broth
<b>LDH</b>	Lactate dehydrogenase
<b>MBD</b>	Microtubule binding domain
<b>MS</b>	Mass spectrometry
<b>MSA</b>	Multiple system atrophy
<b>MTSL</b>	S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl-oxyl) methyl methanesulfonothioate
<b>MTT</b>	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
<b>NAC</b>	Non-amyloid $\beta$ -component
<b>NGM</b>	Nematode growth medium
<b>NMR</b>	Nuclear magnetic resonance
<b>NTD</b>	N-terminal domain
<b>PCR</b>	Polymerase chain reaction
<b>PDB</b>	Protein data base
<b>PDD</b>	Parkinson's disease with dementia
<b>PFA</b>	Paraformaldehyde
<b>POPC</b>	1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
<b>POPS</b>	1-Palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine
<b>PRE</b>	Paramagnetic relaxation experiment
<b>PTM</b>	Post translational modification
<b>RDC</b>	Residual dipolar coupling
<b>Rg</b>	Radius of gyration
<b>Rh</b>	Hydrodynamic radius
<b>RNA</b>	Ribonucleic acid
<b>RRM</b>	RNA-recognition motif
<b>SAXS</b>	Small angle X-ray scattering



<b>SDS</b>	Sodium dodecyl sulfate
<b>SDS-PAGE</b>	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
<b>SEC</b>	Size exclusion chromatography
<b>SEC-MALS</b>	SEC multiple angle light scattering
<b>SILAC</b>	Stable isotope labelling by amino acids in cell culture
<b>SNARE</b>	Soluble N-ethylmaleimide-sensitive factor attachment protein receptor
<b>SN<sub>c</sub></b>	Substantia nigra pars compacta
<b>SNP</b>	Single nucleotide polymorphisms
<b>ssNMR</b>	Solid state NMR
<b>STED</b>	stimulated emission depletion
<b>SUV</b>	Small unilamellar vesicles
<b>T<sub>a</sub></b>	Annealing temperature
<b>TAE</b>	Tris-acetate-EDTA
<b>TEM</b>	Transmission electron microscopy
<b>ThT</b>	Thioflavin T
<b>Tris</b>	Tris(hydroxymethyl)aminomethane
<b>UPS</b>	Ubiquitin-proteasome system
<b>UV</b>	Ultraviolet
<b>VAMP</b>	Vesicle associated membrane protein
<b>WT</b>	Wildtype
<b>YFP</b>	Yellow fluorescent protein
<b>ySyn</b>	Gamma-synuclein



# List of amino acids

<b>A</b>	Ala	Alanine
<b>C</b>	Cys	Cysteine
<b>D</b>	Asp	Aspartate
<b>E</b>	Glu	Glutamate
<b>F</b>	Phe	Phenylalanine
<b>G</b>	Gly	Glycine
<b>H</b>	His	Histidine
<b>I</b>	Ile	Isoleucine
<b>K</b>	Lys	Lysine
<b>M</b>	Met	Methionine
<b>N</b>	Asn	Asparagine
<b>P</b>	Pro	Proline
<b>Q</b>	Gln	Glutamine
<b>R</b>	Arg	Arginine
<b>S</b>	Ser	Serine
<b>T</b>	Thr	Threonine
<b>V</b>	Val	Valine
<b>W</b>	Trp	Tryptophan
<b>Y</b>	Tyr	Tyrosine

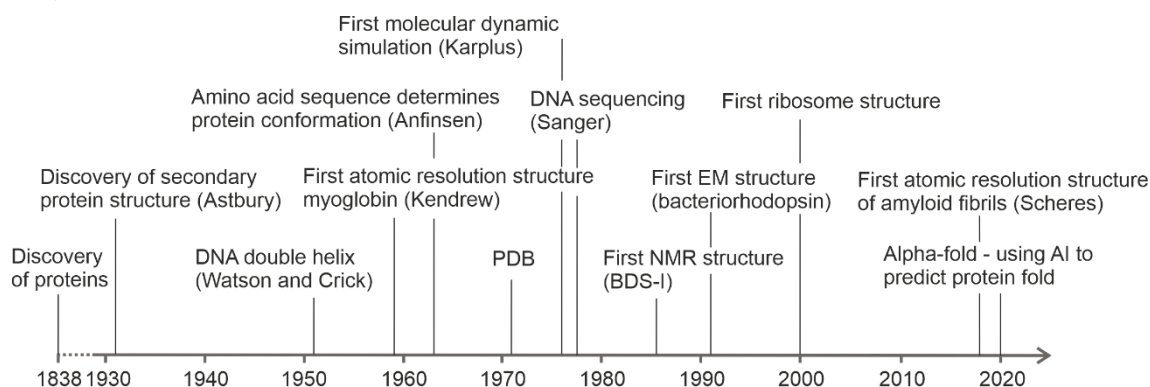


# Chapter 1

## Introduction

### 1.1 Protein Folding and Misfolding

William Astbury, who gave his name to the Astbury Centre, first pioneered structural studies of biological molecules nearly 100 years ago proposing the first models of protein structures in the three dimensional space (incorrect in details but accurate in essence). Based on his work on fibres, he suggested that proteins adopt two forms: the helical  $\alpha$ -form (today known as  $\alpha$ -helix) and the extended  $\beta$ -form (today named  $\beta$ -sheet)<sup>1-3</sup>. Later, Linus Pauling corrected his model and proposed  $\alpha$ -helix and  $\beta$ -sheet to be the secondary structure elements for all proteins<sup>4,5</sup>. Since then, huge milestones were reached in determining the 3D architectures of biomolecules. Examples being the structure of the DNA double helix by Watson and Crick in 1953<sup>6</sup>, the first atomic resolution structure of myoglobin by Kendrew *et al.* in 1960<sup>7</sup>, establishing the protein data bank (PDB) in 1971 by the Brookhaven National Laboratory<sup>8</sup>, using molecular dynamics<sup>9</sup>, and, in general, developing and improving methods for structural analysis including X-ray crystallography, NMR, electron microscopy, and structural prediction (reviewed in Ref<sup>10</sup>) allowing even *in situ* structural analysis via cryo-electron tomography or NMR<sup>11,12</sup> (**Figure 1.1**).

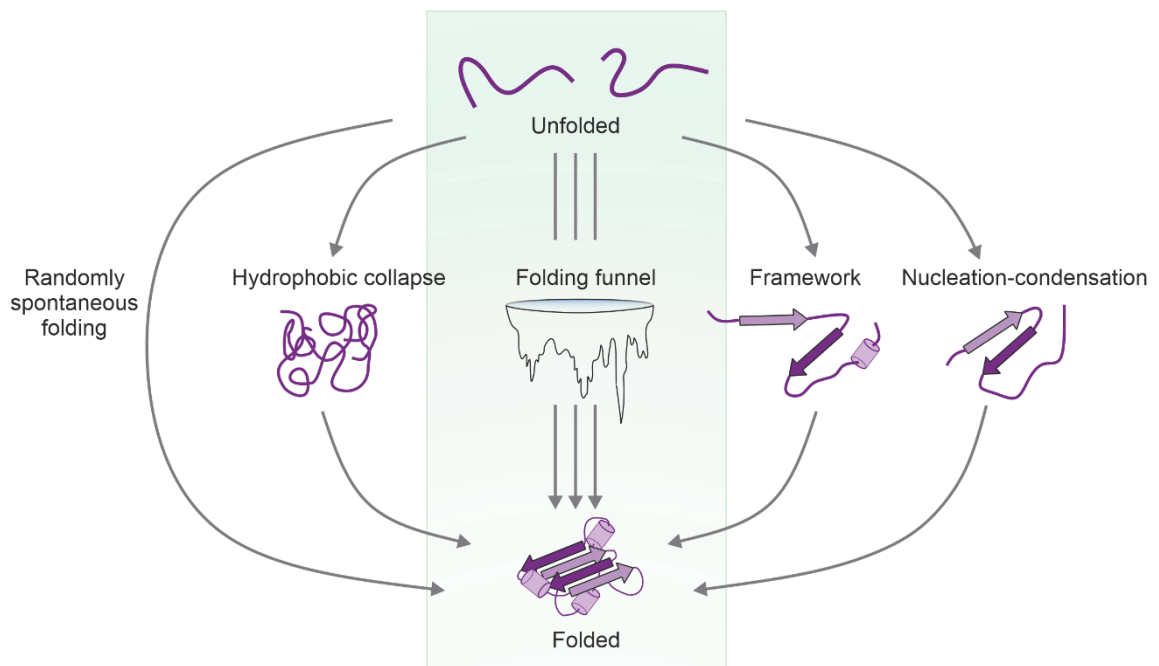


**Figure 1.1: Key events in the field of protein structures.** Redrawn from Ref<sup>10,13</sup>.

Also, understanding the molecular mechanisms underlying protein folding was always of high interest, leading to different proposed folding models (**Figure 1.2**). Anfinsen and co-workers suggested in 1961, based on experiments of the renaturation of Ribonuclease

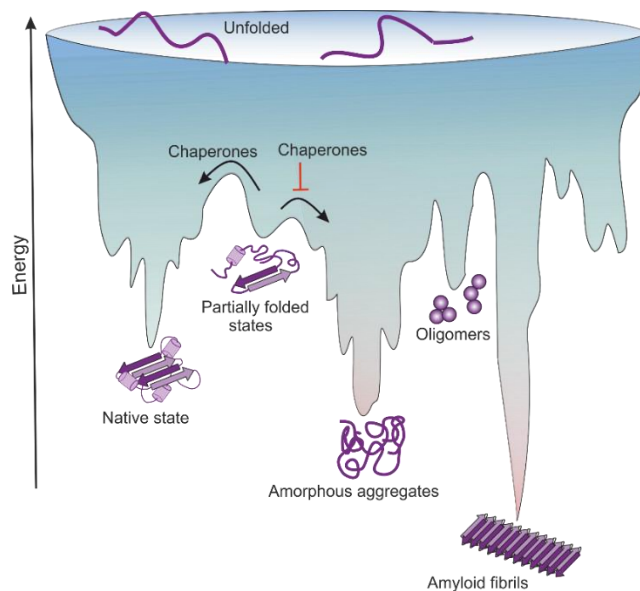
A, that the primary amino acid sequence encodes all the information required for a protein to fold into its native state by randomly searching for the lowest free energy and highest stability (spontaneous folding)<sup>14</sup>. In this model, however, even folding of a small 100 amino acid long protein would take over a billion years to test all possible conformations leading to the most stable native fold, described as the “Levinthal paradox”<sup>15</sup>. So, it was obvious, that there must be specific mechanisms that are kinetically controlled and lead the protein to follow the right folding pathway on a biologically relevant timescale<sup>15,16</sup>.

One of the first folding models, considering a specific folding pathway, was the nucleation-growth model in which the 3D native protein structure is built by the rapid formation of a nucleus of secondary structure elements followed by a hierarchical stepwise progression to the final structure<sup>17</sup>. This theory, however, does not take into account the presence of folding intermediates. Including these led to the framework model<sup>18,19</sup> and the equivalent diffusion-collision theory<sup>20,21</sup>, suggesting that secondary structure elements of the native structure can form independently from the tertiary structure and then diffuse until successful collision, adherent and coalescence to the native 3D architecture. Subsequently, the hydrophobic-collapse model was proposed, hypothesising that the protein collapses quickly around its hydrophobic regions and following rearranges to the final structure from the limited conformational space allowed by the molten globule intermediate<sup>22,23</sup>. Studies on the folding process of the chymotrypsin inhibitor-2 (a small protein with only 64 residues) indicated a two-state transition instead of a single intermediate<sup>24</sup>. Further analysis on this protein on the  $\Phi$ -value revealed that secondary and tertiary structures form in parallel<sup>25</sup>. This finding resulted in the nucleation-condensation model, proposing the formation of a small nucleus of little structure stabilised by long range interactions acting as transition state to form secondary and tertiary structures simultaneously<sup>26</sup>.



**Figure 1.2: Schematic of some proposed folding mechanisms discussed in this work.**  
Redrawn from Ref<sup>27</sup>.

Today, it is believed that there is not a single, specific folding pathway, as it was suggested in the early models described above<sup>28</sup>. In contrast, there is a multidimensional energy landscape, also described as a folding funnel, that describes the folding mechanism best (**Figure 1.3**). There are potentially many routes with different kinetic barriers from an unstructured high energy conformation to the final low energy state protein structure, determined by amino acid sequence, experimental conditions and topology<sup>27</sup>. It should be considered, that the native state of proteins is kinetically favoured but thermodynamically metastable<sup>29</sup>. Aggregates on the other hand, such as amorphous aggregates or amyloid fibrils, represent a thermodynamically and structurally highly stable conformation that often leads to loss of native protein function, as well as the onset of toxic gain-of-function that can lead to cell death and disease. For some proteins, however, the amyloid state can be the functional conformation of a protein<sup>30</sup>.



**Figure 1.3: Schematic energy landscape of folding and aggregation of proteins.** The surface shows the states unfolded proteins can adopt (folded and unfolded) throughout the folding funnel through intra- and inter-molecular contacts. The lowest energy can be found for amyloid fibrils. To overcome energy barriers to end in a folded state, chaperones can help as well as they block or disfavour aggregation. Figure redrawn from Ref<sup>31</sup>.

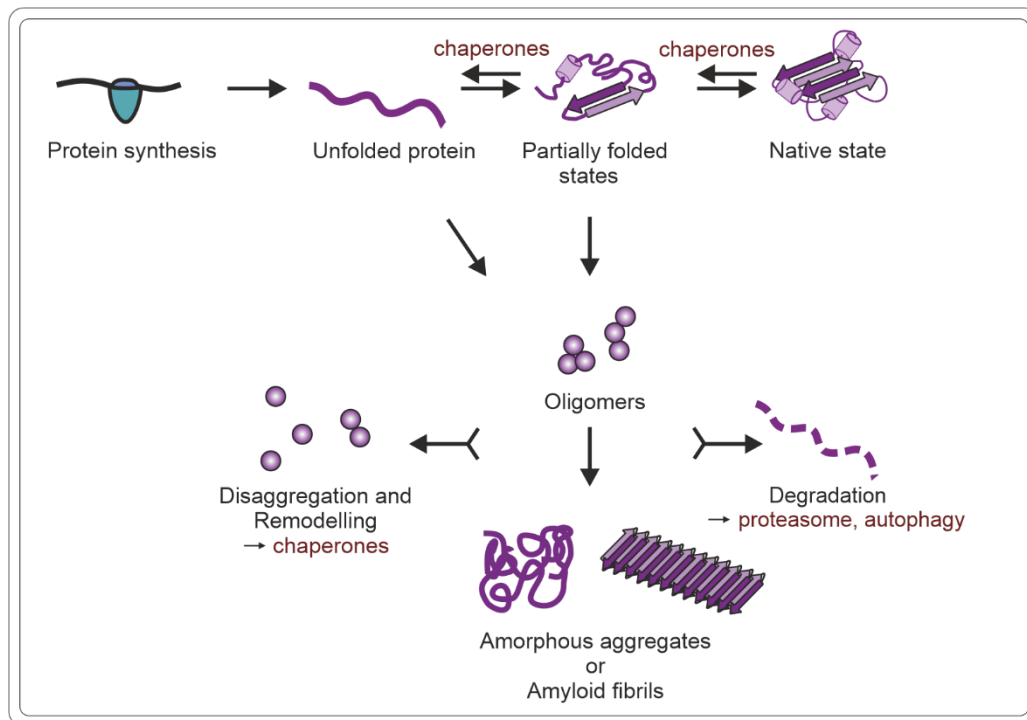
The folding process in a cellular environment is even more complex than *in vitro* due to various reasons. Firstly, protein synthesis on the ribosome and folding is closely coupled and folding can start even before the translation process is completed. This may result in incorrect folding of tertiary structures or exposure of hydrophobic surfaces (especially for bigger and multi-domain proteins)<sup>27,32</sup>. Further, proteins are envionred by a highly crowded cellular milieu with protein concentrations up to 300 g/L. The presence of many membranes and lipids further results in a higher risk of misfolding and aggregation<sup>33</sup>, and proteins are sometimes exposed to harsh conditions such as acidic pH or high temperatures<sup>34,35</sup>. Also, in some cases, genes are mutated resulting in the synthesis of a potentially more aggregation-prone protein variant<sup>36,37</sup>.

Fortunately, there is a complex proteostasis system in cells (**Figure 1.4**), regulating protein folding, stabilisation and degradation processes. The most important players of this system are molecular chaperones, first described to support protein folding about 45 years ago<sup>38,39</sup>. Today, there are more than 300 different chaperones identified in humans, orchestrating folding, refolding<sup>27</sup> and disaggregation<sup>40,41</sup>. There are also reports describing chaperone activity in the extracellular space<sup>42</sup>. The other key actors in the proteostasis network are the ubiquitin–proteasome system (UPS) and the autophagosomal–lysosomal machinery to degradate misfolded proteins and aggregates<sup>43</sup>.

However, sometimes the proteostasis network fails, with a higher risk to do so with increasing age, allowing the formation of toxic oligomers and aggregates. This leads to diseases such as cancer, metabolic syndromes, neurodegenerative diseases,



autoimmunity or inflammatory disorders<sup>44</sup>. Understanding the molecular mechanisms of folding, misfolding and aggregation is therefore crucial to develop new target strategies for drug developments.

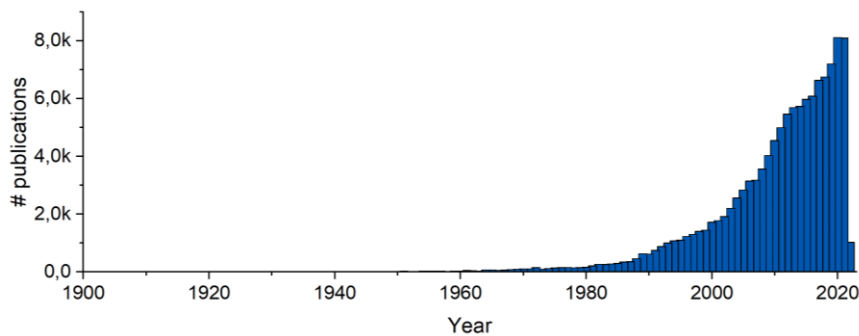


**Figure 1.4: Schematic presentation of the proteostasis network.** Proteins, synthesised at the ribosome fold to its native state supported by chaperones. Unfolded or partially folded proteins sometimes form oligomeric species leading to aggregation. The aggregation processes can be interrupted/reversed by protein degradation or disaggregation and remodelling.

## 1.2 Amyloid

Protein aggregates can be divided into distinct types including amorphous aggregates, in which the protein adopts no homogenous secondary structure and amyloids, which are aggregates that exhibit highly ordered, typically  $\beta$ -sheet rich, structures<sup>45,46</sup>. The interest in understanding the mechanism of conversion from soluble proteins and peptides into insoluble amyloid structures has increased over the last decades (**Figure 1.5**). This is driven by the association to amyloidogenic diseases which are one of the biggest health issues and death causes in upper income countries, together with heart diseases, cancer, and strokes. The most studied examples of amyloid diseases are neurodegenerative disorders such as Alzheimer's disease (caused by  $A\beta$ /Tau), Parkinson's disease (caused by  $\alpha$ Syn), and Huntington's disease (caused by huntingtin with polyQ expansion)<sup>47</sup>. However, some amyloid fibrils have biological functions and have been described in bacteria (e.g. curli in biofilm formation), fungi (e.g. HET-s, programmed cell death), yeast (e.g. Ure2p, nitrogen catabolism), plants (e.g., luminidependens, regulating flowering), and mammals (e.g. pmel17, melanin formation),

highlighting their wide range of functional roles<sup>30,48,49</sup>. Although being studied for decades now, the molecular mechanism of amyloid formation remains still unsolved.



**Figure 1.5: Number of publication in PubMed associated with 'amyloid' from 1900 to date.**

### 1.2.1 Historical perspective

The first time amyloid was described was in 1854 by the German physician Rudolf Virchow<sup>50</sup>. He identified that the application of iodine to abnormal macroscopic structures found in brains and spleen stained them blue. He named these structures “amylum” (Latin for starch) based on his misinterpretation that the recognised structures are built from carbohydrates (cellulose or starch)<sup>50,51</sup>. Only five years later, Friedrich and Kekule clarified that amyloid deposits were proteinaceous<sup>52</sup>.

Since the discovery of amyloid deposits and its correlation with disease in 1911<sup>53</sup>, the interest in these structures has expanded. Our understanding of amyloid has advanced with the growing numbers of available technologies such as light microscopy, the usage of histopathologic dyes including thioflavin and Congo red<sup>54</sup> and subsequent polarisation light microscopic studies<sup>55</sup>. These techniques revealed that amyloid consists of an ordered microscopic structure<sup>55</sup>. Later, electron microscopy and X-ray diffraction experiments showed a general common fibrillar ultrastructure<sup>56,57</sup> made up of a characteristic cross- $\beta$  architecture with  $\beta$ -strands being lined up in a 90° angle to the fibrillar axis<sup>58,59</sup>, properties that still define the characteristics of amyloids today<sup>60</sup>.

It was not until the 1970's that the heterogenic nature of amyloids started to be recognised<sup>61,62</sup>. Currently, there are more than 48 human soluble proteins that have been identified to form amyloid fibrils associated with disease (amyloidoses), presenting deposits distributed systemically or localised in specific organs<sup>63</sup>. In addition to disease related proteins there are more than 35 functional amyloid proteins<sup>64</sup>. The vast majority of amyloid associated with disease were linked with increasing age such as the neurodegenerative disorders Alzheimer's disease (AD) or Parkinson's disease (PD). In a population, in which the number of older people is growing, amyloidoses cause some of the most predominant diseases in the modern world and represent one of the greatest socio-economic burdens of our time<sup>65</sup>.

Therefore, the interest in understanding the nature of amyloid fibrils has never been greater. Today the amyloid field is immensely fast growing utilising a wide range of techniques. Structural methods used include X-ray fibre diffraction<sup>66</sup>, X-ray crystallography<sup>67</sup>, solid state nuclear magnetic resonance (ssNMR)<sup>68-70</sup>, and probably most importantly cryo-electron microscopy (EM)<sup>71-73</sup> (reviewed in Ref<sup>47,60</sup>). The devolvement of *in situ* analysis opens the door to see amyloid fibrils in their natural environment interacting with other molecules<sup>11,74-76</sup>.

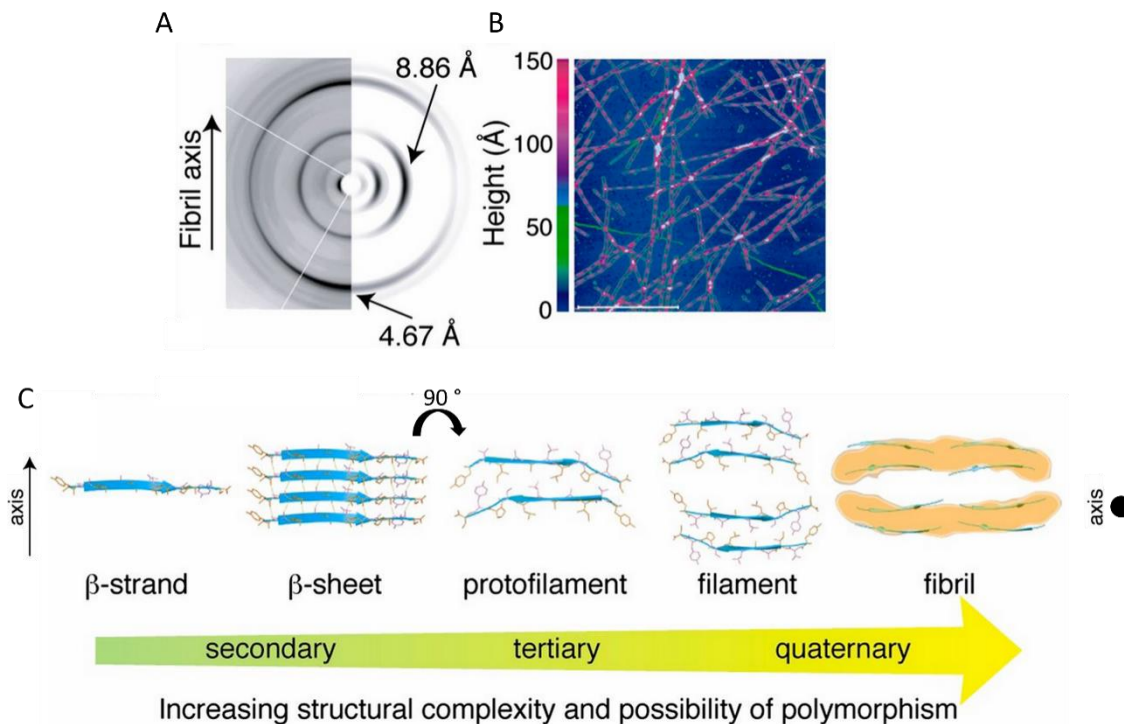
Understanding the mechanisms of amyloid formation and toxicity is driven by the improvement of computational tools allowing the extraction of complex fibril kinetics<sup>77</sup> and the prediction and simulation of aggregation characteristics and conformations of amyloid proteins<sup>78,79</sup>. Further, high resolution real-time observations of growing fibrils at a single-fibril level facilitated understanding the processes of fibril growth in more detail<sup>80,81</sup>. Also, the optimisation of disease models mimicking amyloid diseases in model organisms helps to study amyloid in an *in vivo* context<sup>82-85</sup>.

Despite all these successes in analysing amyloid proteins in more detail, the exact molecular mechanism of aggregation still remains unsolved and there is still no cure and only limited treatment (e.g. tafamidis<sup>86</sup>) for amyloidoses.

### 1.2.2 Structure of amyloids

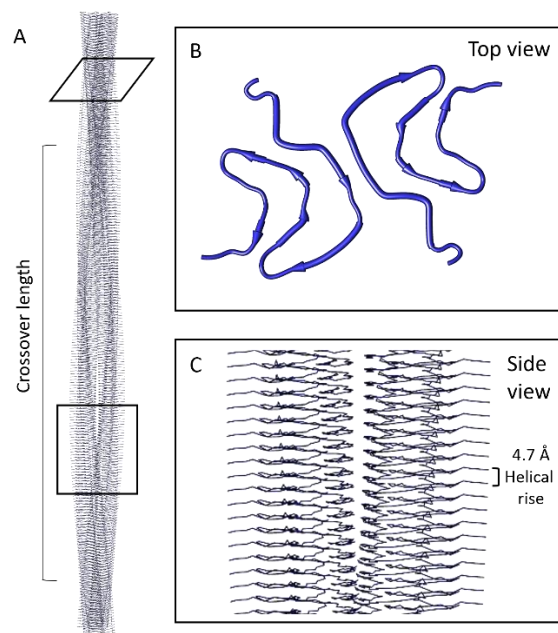
Amyloid is a conformational state that can be adopted by most, if not all proteins<sup>47,60</sup>. The primary sequence of a protein provides the information to adopt their native 3D fold<sup>87</sup> or, in the case of IDPs, remain dynamically unstructured<sup>88</sup>. But this sequence also contains the information for a kinetically and thermodynamically more stable (**Figure 1.3**) alternative structure(s) known as the 'amyloid fold'. Whereas the native fold of a protein presents various secondary structural elements ( $\alpha$ -helix,  $\beta$ -sheet, turn)<sup>89</sup>, amyloid fibrils are built nearly exclusively by  $\beta$ -strands (note that there are exceptions such as the bacterially secreted peptide PSM $\alpha$ 3 forming an unusual cross- $\alpha$  structure<sup>90</sup>). First atomic structural information of fibrils were collected by X-ray diffraction<sup>58,91</sup>. For amyloid, the diffraction patterns show two main reflections: one equatorial at  $\sim 10$  Å, thought to originate from the packing of  $\beta$ - sheets perpendicular to the fibril axis; and one meridional at 4.7 Å which arises from the packing of adjacent  $\beta$ -strands along the fibril axis (**Figure 1.6 A**)<sup>92</sup>. Assembly of  $\beta$ -sheets results in the formation of protofilaments which are usually twisted around each other to form the mature fibril presenting a highly hierarchical structure (**Figure 1.6 C**). Typically, amyloid fibrils are only a few nanometres in width but can extend up to micrometres in length (shown by atomic force microscopy (AFM) in **Figure 1.6 B**) and are formed by two protofilaments, though there are also structures with only one or up to four protofilaments<sup>93,94</sup>. Accumulation of these fibrillar structures,

together with oligomeric species and lipids<sup>76</sup>, are found in disease tissue and referred to as plaques<sup>95</sup>.



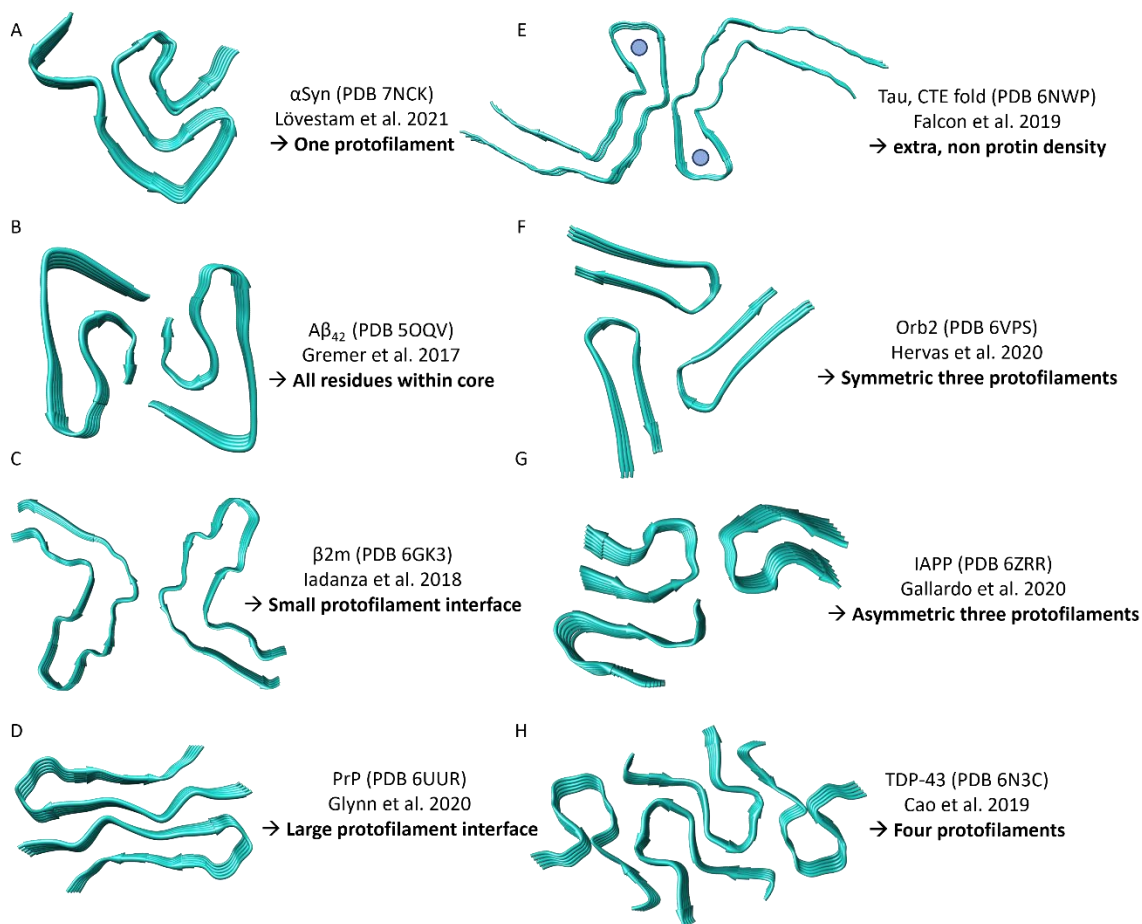
**Figure 1.6: Summary of fibril diffraction and morphology.** (A) Comparison of the high-resolution experimental (left) and simulated (right) X-ray diffraction pattern from TTR(105–115) fibrils. The fibril axis is vertical, with the incident beam directed orthogonally to the axis. The meridional reflection at 4.67 Å and the equatorial reflection at 8.86 Å are characteristic of cross- $\beta$  structure. (B) High-resolution AFM image of TTR fibrils. Scale bar 1  $\mu\text{m}$ . (C) Schematic of hierarchy of atomic-resolution motifs involved in the self-assembly of the amyloid fibrils and their polymorphism. Figure taken from Ref<sup>92</sup>.

Despite the fact that (nearly) all amyloid fibrils show the same main characteristics described above, high resolution structures of various fibrils suggest clear differences between the architectures of these fibres. For example, the conformation of the backbone (top view), as well as helical rise and crossover length can vary (**Figure 1.7**). Although many techniques provide information about the final fibril structure (e.g. X-ray, ssNMR) the field is now dominated by structures solved via cryoEM<sup>94,96,97</sup>.



**Figure 1.7: Schematic of fibril architecture on the example of  $\alpha$ Syn fibrils.** (A) Side view of the reconstructed 3D map highlighting the crossover length. (B) Cross-sectional view of the fibril and (C) side view of the 3D map illustrating the helical rise. PDB 6cu7<sup>72</sup>.

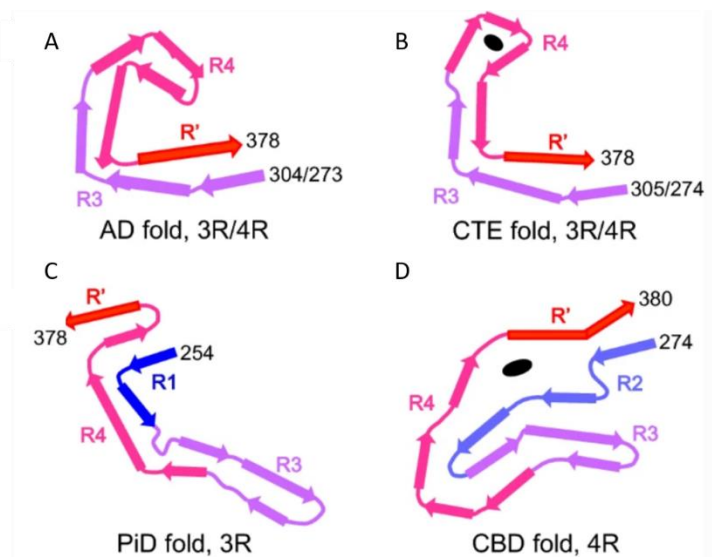
**Figure 1.8** highlights some of the solved high resolution cryoEM fibril structures of various amyloid proteins focussing on the large variability of their architectures. Fibrils can be formed by one, two, three or four protofilaments (**Figure 1.8 A,B,F,H**), all amino acids or only a small fraction of the sequence can be found in the rigid core (**Figure 1.8 B, Figure 1.10**), protofilament interfaces can include many or only a few residues (**Figure 1.8 C,D**), extra densities within the core from non-protein material can sometimes be observed (**Figure 1.8 E**), and structures can be symmetric or asymmetric (**Figure 1.8 F,G**).



**Figure 1.8: Example of wide variety of amyloid fibril structures from different proteins.**

Fibril architectures of  $\alpha$ Syn<sup>98</sup> (A),  $A\beta_{42}$ <sup>99</sup> (B),  $\beta 2m$ <sup>100</sup> (C), PrP<sup>101</sup> (D), Tau<sup>102</sup> (E), Orb2<sup>103</sup> (F), IAPP<sup>104</sup> (G), and TDP-43<sup>105</sup> (H). Structures are shown from a view along the fibril axis. Note that the structures shown here are not necessarily the predominant architecture found of these proteins. They were chosen to illustrate the wide range of identified amyloid fibril structures.

Interestingly, even the same protein sequence can result in distinct fibril morphologies depending on the growth conditions and can be an indicator for the disease type (i.e. structure of Alzheimer's Disease, Pick's disease, Chronic traumatic encephalopathy (CTE), and Corticobasal degeneration (CBD) fibrils are all different although all four are associated with Tau<sup>97</sup> (Figure 1.9)). Further, even small changes in the primary sequence by single point mutations or post-translational modification can result in completely changed fibril architectures compared to WT amyloid as seen for instance for  $\alpha$ Syn<sup>106</sup>, its familial associated PD mutations<sup>107,108</sup> and phosphorylated Y39<sup>94,109</sup>. Recently, Scheres and co-workers published a work reporting 76 cryoEM structures of *in vitro* grown Tau fibrils highlighting the immense polymorphism that can be observed in amyloid fibrils<sup>110</sup>.

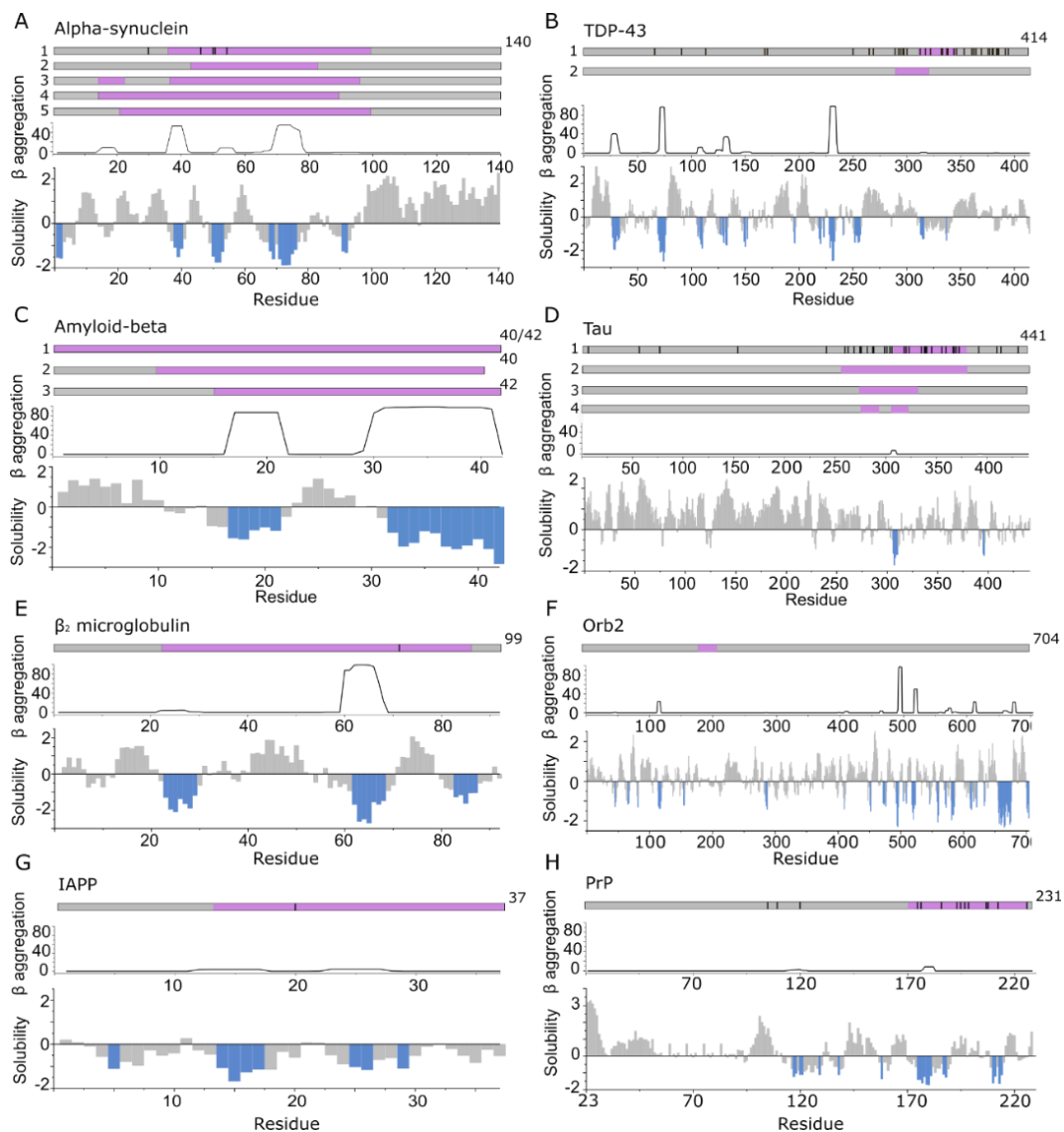


**Figure 1.9: Tau folds from fibrils extracted from patients with different diseases.**

Alzheimer's disease (A), Chronic traumatic encephalopathy (B), Pick's disease (C), and Corticobasal degeneration (D) associated fibril are shown. Different regions of the Tau protein are coloured in blue, purple, pink and red (from N- to C-term) and additional uncharacterized density found within the filament core of the CTE and CBD folds are illustrated by black dots.

Figure taken from Ref<sup>111</sup>.

All these structures determined by cryoEM (or ssNMR) have one big disadvantage: they only focus on the rigid, non-flexible fibril core and set aside the highly dynamic regions flanking the fibril core<sup>112</sup> (see Section 1.2.5). In fact, the amyloid core is often built by only a small fraction of the full-length protein. Online tools, predicting the aggregation-prone or insoluble sequences of a protein that most likely form the fibril core (e.g. TANGO<sup>113</sup> or CamSol<sup>114</sup>) do not always correspond to the fibril core determined experimentally (Figure 1.10). The predicted aggregation-prone regions (APRs) often only comprise a small fraction of the amyloid core, but on the other hand, there are also APRs that are not located within the solved fibril structures. This highlights how complex amyloid assembly and architectures can be and that complementary techniques are required to gain a full image of the fibril structure.

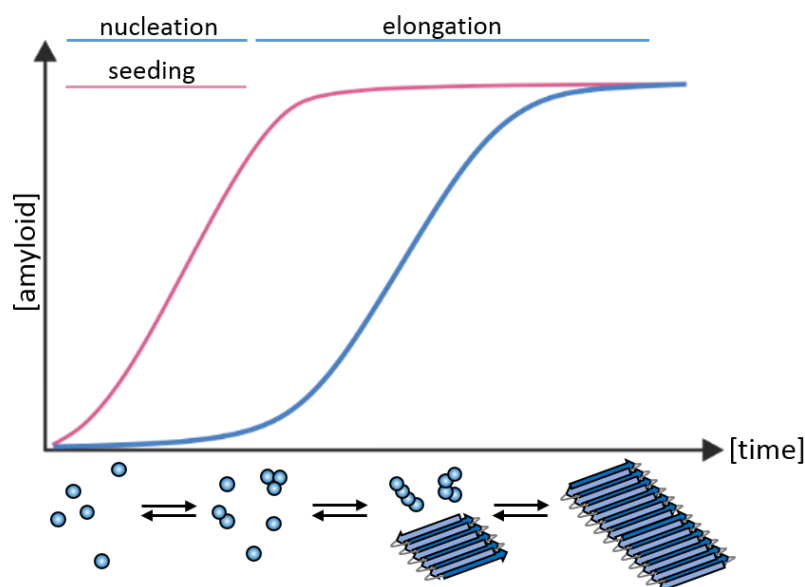


**Figure 1.10: APRs comprise only a small part of the amyloid core.** Top in **A–H**: location of fibril cores of  $\alpha$ Syn, TDP-43, A $\beta$ , Tau,  $\beta$ 2m, Orb2B, IAPP, and PrP defined by recent cryoEM or ssNMR fibril structures (purple). The positions of familial disease mutations are highlighted where appropriate as black lines. Bottom in **A–H**: regions with low solubility predicted by CamSol (below  $-1$  is aggregation promoting highlighted in blue)<sup>114</sup> and the  $\beta$ -aggregation potential of each sequence predicted using TANGO<sup>113</sup>. **(A)**  $\alpha$ Syn including polymorph 1a (1)<sup>72,106,115</sup> (core residues 37–99), 1b (2)<sup>72</sup>(core residues 43–83), 2a and b (3)<sup>116</sup> (core residues 14–24, 36–96) and the MSA ex vivo structures including residues 14–94 (for PF-IA and PF-IIA) (4) or residues 21–99 (for PF-IB and PF-IIB) (5) in the fibril core<sup>117</sup>. **(B)** TDP-43 cryoEM structure solved from C-terminal segments forming a dagger shaped core (1) (residues 312–346) or R-shaped core (2) (residues 288–319)<sup>105</sup>. **(C)** A $\beta$  structures solved (1) for A $\beta$ 42<sup>99</sup> and A $\beta$ 40 in which all residues comprise the core<sup>118,119</sup> (2) fibrils in which the core is formed by residues 10–40 for A $\beta$ 40 (including polymorphs 2A and 3Q)<sup>120,121</sup> and (3) for A $\beta$ 42 (core formed by residues 15–42)<sup>122,123</sup>. **(D)** Tau fibril structures PHF and SF from Alzheimer disease patients (1)<sup>73</sup> (core residues 306–378), NPF and WPF from Pick’s Disease (2)<sup>124</sup> (core residues 254–378) and heparin induced structures 4R-s and 3R formed *in vitro* (3) (core residues 272–330) and 4R-t and 4R-j (4) (core residues 274–292, 304–321, respectively)<sup>125</sup>. **(E)** The  $\beta$ 2m fibril core involves residues 22–85<sup>100</sup>. **(F)** The Orb2B fibril core consists of residues 176–206<sup>103</sup>. **(G)** Human IAPP forms fibrils with residues 13–37<sup>126</sup>, 14–37<sup>93</sup>, or 13–37<sup>104</sup>, with its early onset S20G variant adopting fibrils with two- and three filaments involving residues 15–37 in the core<sup>104</sup>. **(H)** PrP fibrils form fibril core with residues 170–229 revealed using cryoEM<sup>127</sup>. Figure taken from Ref<sup>128</sup>.



### 1.2.3 Mechanism of amyloid formation

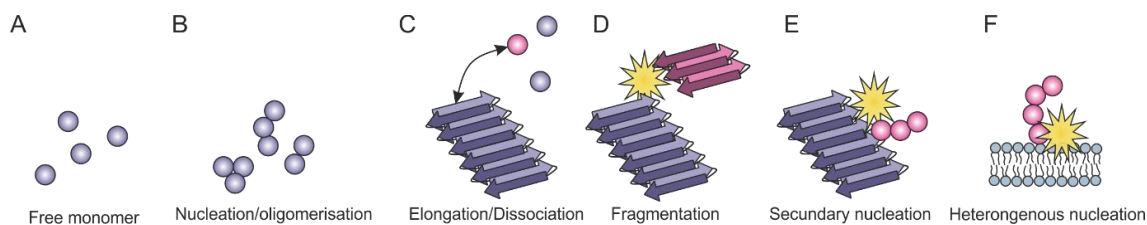
Amyloid formation is a complex process, involving the interplay between a variety of intermediates with different morphologies<sup>46,129</sup>. In general, fibril assembly can be described by a sigmoidal growth curve characterised by a nucleation phase (lag-phase) and an elongation phase (**Figure 1.11**). This nucleated growth reaction is rate limited by the formation of an aggregation-competent nucleus which is kinetically disfavoured. Further deposition of monomers occurring around this oligomeric structure leads to kinetically and thermodynamically highly stable amyloid fibrils<sup>130</sup> (**Figure 1.11** and **Figure 1.12 A-C**). The nucleation phase can be eliminated by the addition of pre-formed fibrils/seeds (**Figure 1.11**)<sup>131</sup> or increased/decreased by altering the growth conditions (e.g. adding additional surfaces or varying the pH)<sup>132,133</sup>. Protein-protein interactions play a major role in the lag phase of the aggregation mechanism. Transient interactions between molecules form species that have, or lack, the propensity to elongate, thus creating a heterogeneous pool of on- and off-pathway intermediates. During the amyloid assembly, a diverse mixture of aggregate types can be formed including oligomers, amorphous aggregates or other prefibrillar species<sup>131,134-136</sup>.



**Figure 1.11: Schematic of spontaneous/*de novo* (blue) and seeded (pink) amyloid formation.** Two distinguished phases can be observed: the lag phase, in which thermodynamically disfavoured nucleation events happen and a rapid, thermodynamically favoured elongation phase. The addition of seeds drastically shortens the lag phase of fibrils formation.

Fibril growth can experimentally be measured using the aromatic, amyloid specific fluorescent dye Thioflavin T (ThT). This small molecule binds to cross- $\beta$  structures changing its fluorescence characteristics and allowing the detection of the typical sigmoidal fibril growth curve (**Figure 1.11**)<sup>137</sup> (see Section 2.8.2). Although amyloid growth always follows similar overall processes (lag time, elongation phase, plateau

phase), at the molecular level they show significant differences due to various competing mechanisms (**Figure 1.12**). Self-assembly into amyloid fibrils are complex reactions with contributions of primary nucleation and elongation (**Figure 1.12 B,C**), but also secondary processes such as fragmentation, secondary nucleation and heterogenous nucleation (**Figure 1.12 D-F**)<sup>77,138</sup>. Secondary events can massively affect the aggregation kinetics by for example increasing the amount of fibril ends as a result of fragmentation<sup>139,140</sup> or changing aggregation rates by surface-catalysed processes that facilitate nucleation and aggregation which is especially important in the context of an *in vivo* setting of crowded cells<sup>141,142</sup>. This complex nature of fibril growth makes it challenging to understand the precise molecular mechanism of fibril assembly. *In vitro*, the development of mathematic fits of aggregation curves by Knowles and co-workers (named AmyloFit) has improved the interrogation of microscopic processes from macroscopic experimental measurements helping to understand which processes dominate the aggregation kinetics under specific conditions and determine individual aggregation rates<sup>77</sup>.



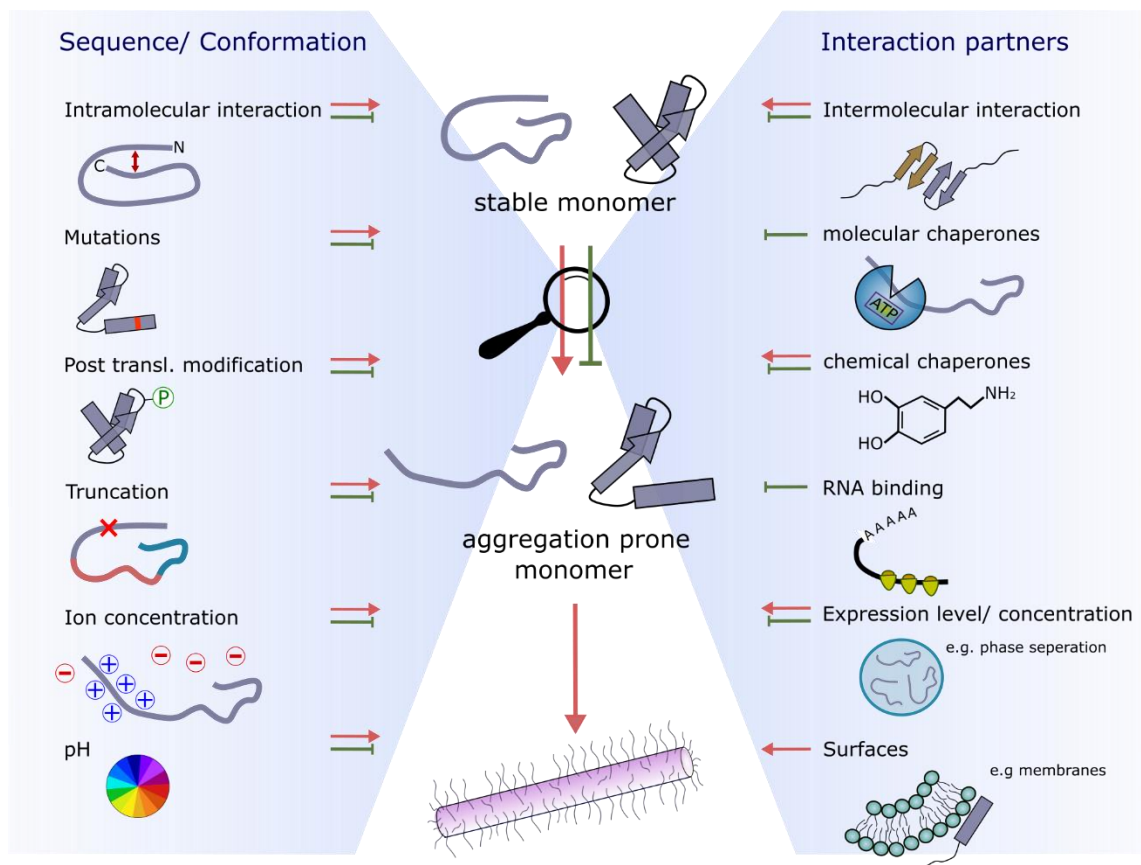
**Figure 1.12: Schematic illustrating the microscopic processes of fibril formation including secondary aggregation pathways.** (A) Monomeric proteins form nuclei and oligomers (B) through primary nucleation processes. (C) Amyloid fibrils grow linearly in a reversible manner. Secondary pathways include fragmentation (D) and monomer dependent secondary nucleation (E) and result in new fibril ends from pre-existing amyloid structures. (F) Heterogeneous nucleation on a surface (e.g. lipid bilayers) accelerates fibrils formation. Figure redrawn from Ref<sup>138</sup>.

Fibril formation of amyloidogenic proteins is also challenging to study as there are a great number of internal and external factors that can affect the aggregation kinetics of a protein. For example, the transition from stable monomers to monomeric intermediates, which is a vital early step for some amyloidogenic proteins<sup>143,144</sup>, is significantly affected by the sequence and conformation of the protein as well as their interaction partners (**Figure 1.13**). Consequently, side chain substitutions and post translational modifications (PTMs) can change the aggregation propensity of a protein, often observed for disease related polymorphism<sup>145,146</sup>. Further, varying buffer condition (or cellular environments) such as differences in ion concentration or pH, can drastically change the aggregation kinetics<sup>133,147,148</sup>. This effect is especially important for intrinsically disordered proteins as they have larger solvent exposed areas<sup>149,150</sup>.

Particularly in an *in vivo* context, the contribution of interaction partners is important. Binding to chaperones<sup>151-153</sup> or surfaces (e.g. membranes, other molecules)<sup>154</sup> can slow down or speed up the fibril formation process. Also, the local concentration is crucial,

determined by expression level<sup>155,156</sup> or processes such as phase separation<sup>157,158</sup> (**Figure 1.13**). All these factors can drastically change the kinetic landscape of individual proteins.

Taken together, although the methods of studying the molecular mechanism of aggregation have improved, the precise mechanism(s) are still unclear and a deeper, more systematic research is required.

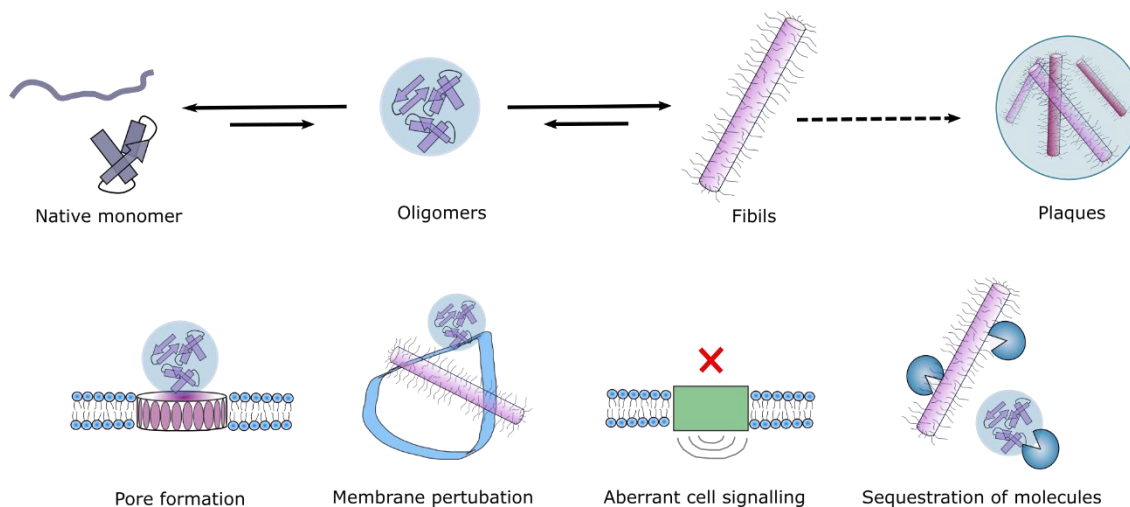


**Figure 1.13: Factors affecting the conformational properties and interactions with amyloid precursors that retard or accelerate fibril growth.** The two main influences are sequence/conformation-dependent (left-hand side) or occur as a result of binding to interaction partners (right-hand side). Figure taken from Ref<sup>128</sup>.

### 1.2.4 Toxicity in amyloid proteins

Both, amyloidogenic oligomers and fibrils have been demonstrated to contribute towards cytotoxicity through a variety of mechanisms (**Figure 1.14**)<sup>144,159-162</sup>. In general, the formation of amyloid assemblies can result in a toxic outcome driven by a loss of native function, gain of toxic function, or a combination of both. Especially for the toxic function, similar processes throughout most amyloid proteins can be observed. Oligomers and fibrils can interact with, and sequester a wide range of proteins<sup>163</sup>. These protein-protein interactions show large effects at the cellular level as seen for example for binding to chaperones leading to aggregation of other proteins by perturbing the cellular stress

response<sup>164</sup>. Amyloid assemblies have also been shown to interact and disturb membranes either by fibrils interacting with membrane surfaces<sup>11</sup> or pore formation caused by oligomers<sup>160-162</sup>. Further, binding to cell surface receptors has been linked with activation of signal transduction pathways leading to apoptosis<sup>165</sup>.



**Figure 1.14: Mechanisms of toxicity which have been observed for oligomers and fibrils of amyloid proteins.** Top: species formed from monomer to oligomers, fibrils and eventually plaques. Bottom: toxic processes such as pore formation in membranes, membrane disruption, aberrant of cell signalling, and sequestration of molecules such as chaperones caused by oligomers and/or fibrils.

An ongoing discussion in the field rises from which species, oligomers or fibrils, is the more toxic one. This question has not been finally answered and there is evidence for both species to be involved in cytotoxicity (**Figure 1.14**). An example of the high complexity of toxicity is demonstrated by the wide range of oligomer studies: Whilst pre-fibrillar oligomers formed during the lag phase are often described as the most toxic species<sup>161,166-168</sup>, other experiments show no impact on cells cultures<sup>169</sup> or a high dependence on conformational properties<sup>160,162</sup>.

Taken together, this highlights that inhibiting fibril formation does not automatically result in a reduction of toxicity<sup>168,170</sup>. In contrast, this could, in some cases, lead to an actual increase in cell death due to oligomeric species being accumulated. This is supported by disease linked mutations that show slower aggregation kinetics but involvement in disorders (e.g. A30P in  $\alpha$ Syn linked with PD)<sup>171,172</sup>. Therefore it is important to understand the origin of toxicity for individual amyloid proteins to guarantee effective treatment of amyloid associated diseases.

### 1.2.5 The role of flanking regions

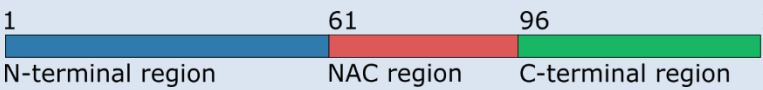
Driven by the early observation of the presence of ordered structure within amyloid fibrils<sup>50,55-57</sup> and the potential to develop inhibitors of their formation, a major goal of the amyloid field has been to elucidate the structure of the amyloid fold at atomic resolution.

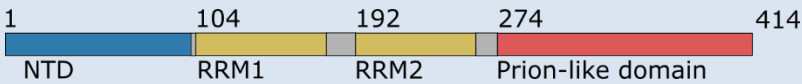
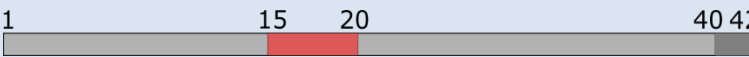
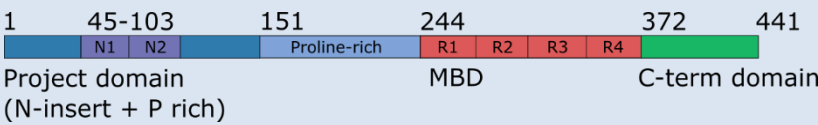
CryoEM studies, together with *in silico* methods, able to predict aggregation-prone regions (APRs) in protein sequences, have provided a wealth of information about the ordered fibril cores that comprise the amyloid fold (see Section 1.2.2).

Importantly, structural and kinetic analyses have shown that amyloidogenic proteins in their fibrillar form often contain less well-ordered sequences outside of the amyloid core, termed here as flanking regions. Modifying or changing (e.g. by deletion, mutation or PTM) these flanking regions can affect the fibril growth kinetics<sup>173,174</sup>, fibril morphology<sup>94,96,97</sup> or the formation of crucial contacts with interaction partners<sup>175,176</sup>. Flanking regions can play a key role at the monomeric or aggregated level.

The presence of so called ‘gatekeeper’ residues for example, which surround aggregation-prone regions, protect the monomeric protein from fibril formation<sup>177</sup>. Further, N- or C-terminal truncated variants often observed in disease (e.g.  $\Delta$ N6 in  $\beta$ -2 microglobulin ( $\beta$ 2m) associated with dialysis-related amyloidosis patients) significantly increase the aggregation propensity of the protein by destabilising its native state and increasing dynamics that drive amyloid formation<sup>178-181</sup>. Also, for a functional perspective, the regions outside the APR are crucial as seen for Orb2 (cytoplasmic polyadenylation element binding (CPEB) protein family) interacting with RNA to facilitate long term memory formation<sup>182</sup>. A detailed summary of contributions of flanking regions for aggregation and function of various amyloidogenic proteins is presented in **Table 1.1**.

**Table 1.1: Roles of the flanking regions in monomers of different amyloid protein sequences.** A schematic of each protein is shown to highlight important regions/domains, with sequences mainly involved in aggregation/core formation highlighted in red. Abbreviations: NAC (non-amyloid  $\beta$ -component), NTD (N-terminal domain), RRM (RNA-recognition motif), MBD (microtubule binding domain). Table taken from Ref<sup>128</sup>.

Protein	Residues/region	(dis-)function	Ref
$\alpha$ Syn	 <p>1 61 96 140 N-terminal region NAC region C-terminal region</p>		
	Residues 1-11/12	Monomer interacting with $\alpha$ Syn fibrils	183,184
	Residues 1-14	Membrane insertion	185
	Residues 1-25	Initial membrane binding	186
	Extreme N-terminus and region around Y39	Chaperone binding	175
	Residues 37-54	Forms $\beta$ -hairpin crucial for nucleation/oligomerisation processes	187
	C-terminal region (residues 91-140)	Protects protein from aggregation by shielding NAC region and/or $\beta$ -hairpin. C-terminal truncation (109–140) results in faster aggregation	188-191
	C-terminal region (residues 110-140)	Binding to chaperone-like protein SERF accelerates aggregation	192
	C-terminal region (residues 125-129)	Dopamine binding drives off-pathway oligomer formation.	193

TDP-43			
	Residues 3-183	Interactions initiate homo-dimerization important for polymerisation dependent splicing activity	194,195
	1-10 (especially Arg6, Val7, Thr8 and Glu9)	Mediates full-length TDP-43 oligomerisation important for splicing activity and key to initiate aggregate formation	196
	RRM1 (104-176), especially residues I107, D105, L111, W113, Q134, G146, F147, F149, R171, K176, N179 (RRM2 (residues 192-262))	Binds TG-rich DNA and UG-rich RNA for function (e.g. splicing, translation control, transport). RRM2 shows lower binding affinity.	197,198
	RRM1 (residues F147 and F149) and residues 208-441	Prevents aggregation by enhancing solubility when bound to single stranded RNA/DNA	199
	RRM1 (residues F147 and F149) and residues 321-366	Autoregulation of own protein expression by binding to its mRNA.	156
	Residues 320-340, especially W334, W385 and W412	Involved in liquid-liquid phase separation	200-202
A $\beta$			
	N-terminal domain (residues 1-17)	Binding to cystatin C (cysteine protease inhibitor)	203
	A $\beta$ <sub>40</sub> : central region (residues 25-29); part of the structured fibril core but solvent accessible	Disaggregase activity when binding Lipocalin-type Prostaglandin D synthase (L-PGDS)	204
Tau			
	Residues 1-202	Binding to plasma membrane	205
	N-terminal domain (1-150) interacts with proline rich domain (151-244)	Dimerization (head to tail), suggested to be the natural form for function and toxicity	206
	Residue 1-117 and 118-402	Electrostatic interactions between these regions drive phase separation	207
	Residues 114-193 (P-rich domain) and 198-278 (microtubule-binding domain)	Actin binding and promoting F-actin bundling and G-actin assembling	208
	N-terminal domain, proline-rich region and MBD	Chaperone binding	209
	Proline rich domain, MBD	Interaction and polymerization of tubulin	210,211
	Proline rich domain and C-terminal domain	Main locations of phosphorylation sites, but can be found throughout the whole sequence	212

	MBD (residues 295-305)	$\beta$ -hairpin formation that protects the aggregation-prone 306-311 region	211
	MBD (residues 275-280 and 306-311) and other regions	Heparin binding drives aggregation; MBD shows highest affinity to heparin	213
$\beta$ 2m			
	Residues 1-6	Stabilization of native structure; accelerates aggregation when deleted	214
	A and G strand (I7A, V9A and V93A)	Mutations drive fibril growth by destabilizing local tertiary structure and increasing dynamics	215
	A,B,E,F strand (6-11, 21-28, 64-70, 79-83)	Interaction with chaperone $\alpha$ B-crystalline preventing oligomerization and fibril formation	214
Orb2B			
	RNA binding domain	Interaction with RNA facilitates long term memory formation	182
IAPP			
	Residues 1-19	Membrane binding and disruption	216
	Residues 1-17 and/or 30-37	Liquid-liquid phase separation	217
PrP			
	N-terminal region (residues 23-90)	Interaction with Tau	218
	N-terminal region (residue 23-89)	Interaction with $\alpha$ Syn fibrils facilitating $\alpha$ Syn cell-to-cell spreading	219
	Residues 95-110	Receptor binding site for A $\beta$ 42-oligomers	167
	Hydrophobic region (residue 111-134)	Hydrophobically driven binding/insertion with anionic membranes, this interaction is important for (murine) PrP to gain C-terminal Proteinase K resistance and convert it to PrP <sup>Sc</sup>	220
	Octapeptide region in N-terminal domain	Increased numbers of octapeptides that bind Ca <sup>2+</sup> promotes fibril formation and disease development	221

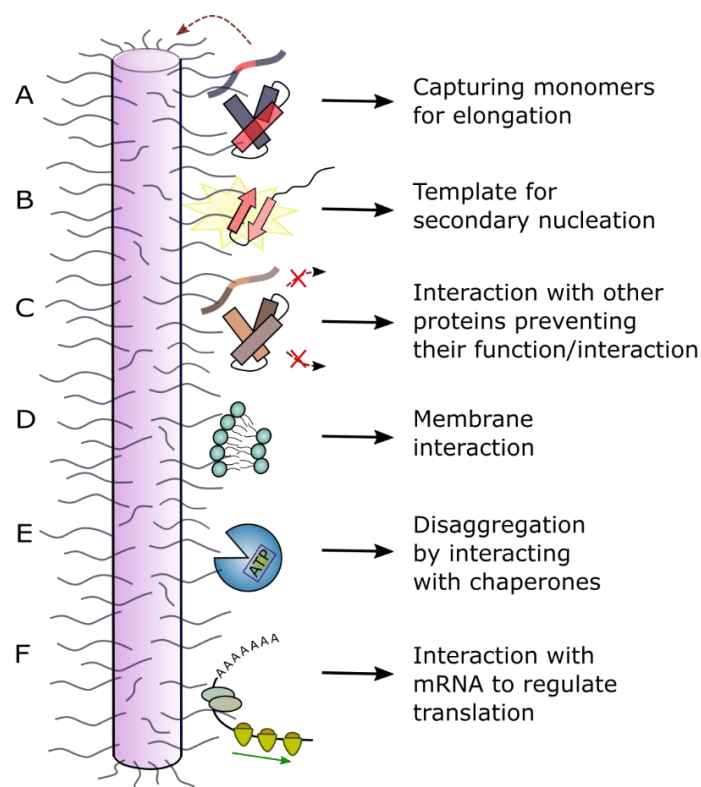
Flanking regions of fibrils, sometimes referred to as a ‘fuzzy coat’ are also crucial for diverse biological processes (**Figure 1.15**). It should be noted, that the ‘fuzzy coat’ can be structurally complex as described for example for Tau, which forms a ‘two layered polyelectrolyte brush’ surrounding the fibril core. This is formed by a dense and mechanically more rigid layer (residues ~173-243) and an N-terminal, less dense and more dynamic layer (residues ~1-172)<sup>222,223</sup>. The ‘fuzzy coat’ has been reported to be involved in elongation or secondary nucleation processes (e.g. first ~10 N-terminal residues of  $\alpha$ Syn are required for elongation and secondary nucleation)<sup>183,184</sup> (**Figure**

**1.15 A,B)** and being important for interactions with other proteins. As an example, A $\beta$  fibrils interact with at least 10 other disease-related amyloidogenic proteins (e.g. IAPP, Tau,  $\alpha$ Syn) supporting the “amyloid cascade hypothesis” as an underlying cause for Alzheimer’s disease<sup>224</sup> (**Figure 1.15 C**). Further, interactions with membranes<sup>11</sup>, chaperones<sup>225</sup> and mRNA<sup>103,226</sup> have been reported (**Figure 1.15 D-F**).

It should also be considered that flanking regions can have very different lengths. In the case of A $\beta_{40/42}$ , (nearly) the whole protein forms the fibril core (**Figure 1.10 C**)<sup>99,118</sup>, whilst in other proteins, e.g. Orb2, flanking regions >500 residues in length are observed (**Figure 1.10 F**)<sup>103</sup>. Tompa *et al.* hypothesised that the longer the flanking region, the more likely it can interact with other molecules by a ‘fly fishing mechanism’<sup>112</sup>. This might explain why functional amyloid fibrils such as Orb2 have long flanking regions able to interact with other proteins, RNA or surfaces, whilst pathological amyloid fibrils, with sequences of which have not evolved for functional reasons, may present shorter flanking regions.

Therefore, far from being passive bystanders, flanking regions may be as important in defining the physiological role and amyloid disease aetiology as the fibril cores themselves.



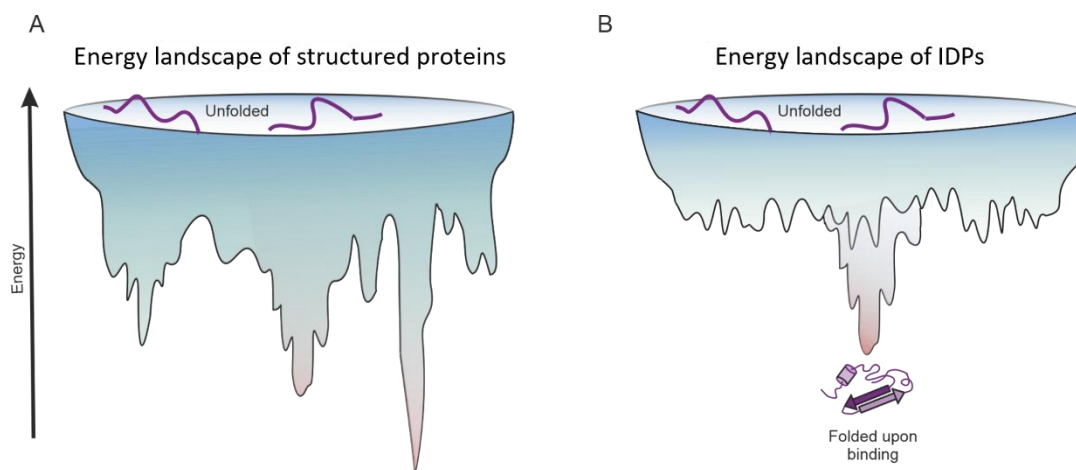


**Figure 1.15: Schematic of interactions between the dynamically disordered regions of fibrils and other molecules.** Dynamically disordered regions displayed on the surface of ordered amyloid fibrils can play roles in their function and cellular dysfunction. (A) The fibril “fuzzy coat” could capture amyloid precursors and facilitate their self-assembly into amyloid by (B) secondary nucleation or other molecular events. (C) Interaction with other (non-) amyloidogenic proteins could inhibit or alter their function. (D) Interaction with membranes can result in a toxic mechanism involving membrane disruption. (E) Interaction with chaperones can result in fibril depolymerisation. (F) Interaction with RNA has been observed in a functional context for the protein CPEB (cytoplasmic polyadenylation element binding) which regulates long-term memory<sup>103</sup>. In other cases, disruption of cellular RNA could enhance phase separation and lead to cellular toxicity and/or dysfunction. Figure taken from Ref<sup>128</sup>.

### 1.3 Intrinsically disordered proteins

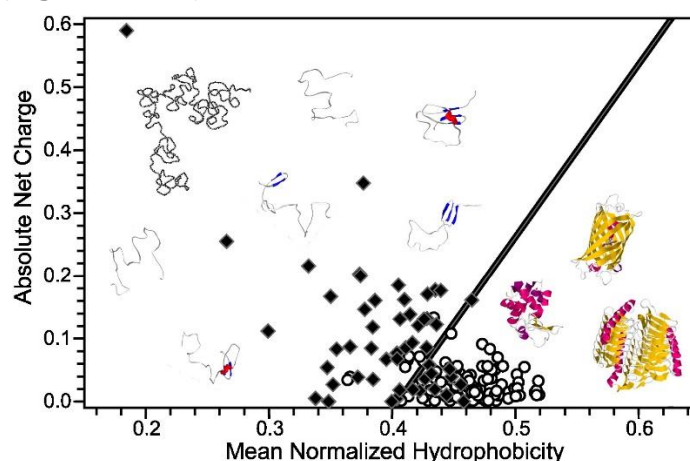
Decades of studies focused on understanding the mechanisms of protein folding of natively folded proteins<sup>19,23</sup>. However, today it is believed that over 50 % of the proteins in eukaryotic systems have long (> 30 residues) intrinsically disordered regions (IDR) or are completely intrinsically disordered proteins (IDP) (12% of eukaryotic proteins)<sup>112,227,228</sup>.

IDPs present a distinct energy landscape compared to natively folded proteins, being much flatter and lacking an obvious folded native state (**Figure 1.16**). These proteins are thermodynamically less stable compared with globular folded proteins, and have no single, well-defined 3D architecture. IDPs exist as heterogeneous ensembles of conformers in which no single set of coordinates or backbone dihedral angles is sufficient to describe their full conformational properties<sup>229</sup>.



**Figure 1.16: Energy landscape of a structured proteins (A) vs IDPs (B).** The folding funnel for IDPs lacks any obvious troughs populated by more structured polypeptides. A lower energy state exists that forms under IDPs binding to interaction partners.

The ability of proteins to fold or not to fold under physiological conditions is encoded in their amino acid sequence<sup>228</sup>. Sequences with a low hydrophobicity (low tendency for protein compaction) and high net charge (causing electrostatic repulsion) are usually indicators for disorder (**Figure 1.17**)<sup>230</sup>. Amino acids commonly known to cause order such as Ile, Leu, Val, Trp, Tyr, Phe, Cys, and Asn are usually reduced, whereas disorder-promoting residues including Ala, Arg, Gly, Gln, Ser, Glu, Lys, and Pro are enriched in IDPs and IDRs<sup>231</sup>. Computational analysis, single-molecule studies, and other structural biology techniques such as NMR and mass spectrometry (MS) have discovered that the amino acid sequence affects the IDP/IDR conformational ensemble. The side chain composition in addition with environmental conditions (e.g. salt, pH) determine which transient intra- and inter-molecular interactions are formed and if the protein adapts an extended or compact conformation resulting in a rough but shallow energy landscape<sup>232,233</sup> (**Figure 1.16 B**).

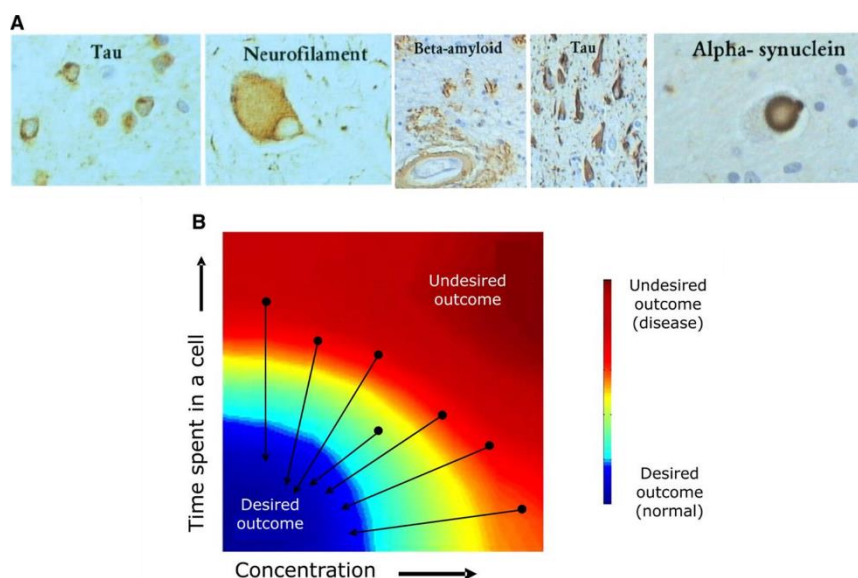


**Figure 1.17: Charge-hydropathy phase space.** Illustrated are the biophysical characteristics hydrophobicity and net charge of 275 folded (open circles) and 91 natively unfolded proteins (black diamonds) proteins. The solid line represents the border between extended IDPs and compact globular proteins. Examples of extended IDPs (native coils and native pre-molten globules) and some ordered globular proteins are shown. Figure taken from Ref<sup>230</sup>.

Interestingly, IDRs and IDPs are enriched in proteins with important regulatory or signalling functions<sup>112,234</sup> demonstrating the crucial role of flexible domains for biological processes. Disorder offers many biological advantages important for function such as connecting structural domains, mediating protein interactions via linear motifs or facilitating PTMs permitting encoding and decoding of information<sup>235,236</sup>. The lack of structure in IDPs and IDRs has been found to enable highly specific, but low affinity interactions with their targets that can easily be triggered by small perturbations (e.g. mutations or PTMs) of entropy or enthalpy of the binding process<sup>236</sup>. The large variety of binding partners and relatively easy and fine regulation of IDP interactions reasons why they are often classified as ‘hub-proteins’<sup>237</sup>.

Although defined by the lack of a unique 3D structure, IDPs and IDRs are not entirely and not under all circumstances flexible and disordered. For example, the adoption of distinct conformational states when interacting with different binding partners is often seen<sup>238-240</sup>. This folding upon binding of IDRs enables specific and weak interactions with their binding partners<sup>241</sup>. Also, structural transitions of IDRs (caused by for instance PTMs at specific side chains) have been shown to regulate protein functions<sup>232,240</sup>.

A high number of native IDRs and IDPs can be found in amyloid proteins, which are mainly known for their involvement in neurodegenerative disease such as Alzheimer’s Disease, Parkinson’s Disease, Amyotrophic Lateral Sclerosis or Huntington’s Disease<sup>47,51</sup> (**Figure 1.18 A**). Whether an IDP aggregates or not is highly regulated by its incubation time and concentration (**Figure 1.18 B**), therefore a functional proteostasis network in a cell is crucial (**Figure 1.4**). An example for a disease causing IDP when forming amyloid fibril structures is alpha-synuclein ( $\alpha$ Syn). This 140 amino acid protein is also the protein of interest in this thesis. It is mainly located in presynaptic termini of dopaminergic neurons and intrinsically disordered when in its free monomeric state in the cytoplasm<sup>242</sup>. Important for its function, it presents a repetitive motif (KTKEGV) involved in membrane binding, which induces the formation of an  $\alpha$ -helix<sup>243,244</sup>. Characteristically for amyloid proteins it forms fibrils with a high cross- $\beta$  sheet content under aggregation promoting conditions or disease<sup>47</sup>.  $\alpha$ Syn will be further discussed in Section 1.4.



**Figure 1.18: IDPs in disease. IDRs and disease.** (A) IDPs and IDRs are found in plaques and cellular deposits of patients with neurodegenerative disease such as FTDP (Tau, neurofilament), AD ( $A\beta$ , Tau) and PD ( $\alpha$ Syn)<sup>245</sup>. (B) Protein availability-outcome landscape. Tight regulation of proteins with IDRs (black arrow) ensures that they are present in the right amount and not longer than required (proteostasis network required). Figure taken from Ref<sup>236</sup>.

## 1.4 Alpha-synuclein

The synuclein family consists of three distinct genes (SNCA, SNCB, and SNCG) encoding the proteins alpha-synuclein ( $\alpha$ Syn), beta-synuclein ( $\beta$ Syn), and gamma-synuclein ( $\gamma$ Syn), respectively (see Section 1.5). The first synuclein to be discovered in 1988 was isolated from the Pacific electric ray *Torpedo californica* as a neuron-specific protein localised in presynaptic nerve termini and nuclei<sup>246</sup>. Although first described 35 years ago, it was only in 1997 that the association between  $\alpha$ Syn and neurodegenerative disorders was made: A mutation in the SNCA gene causing the missense mutation A53T in  $\alpha$ Syn was linked with early-onset familial PD<sup>247</sup>. To date, eleven more familial PD mutations have been identified (although only eight of them are commonly discussed in the literature) (see Section 1.4.2.1)<sup>248</sup>. Further,  $\alpha$ Syn aggregates and especially fibrils were identified to be the main components of Lewy bodies, intracellular inclusion bodies associated with PD<sup>249</sup>. The association of  $\alpha$ Syn with disease raised the scientific interest in this protein. But even extensive research over the past decades did not allow a full understanding of physiological function and toxic role<sup>250,251</sup>.

### 1.4.1 Physiological function

Alpha-synuclein is a 140 residue intrinsically disordered protein expressed exclusively in vertebrates<sup>252</sup>. It is predominantly located in the presynaptic termini of neurons (**Figure 1.19**) in the central nervous system<sup>253</sup> with concentrations up to  $20 \mu\text{M}$ <sup>254</sup>. Although extensively studied, the physiological function of  $\alpha$ Syn, is still unclear<sup>250,251</sup>. This is probably due to the small phenotypic effects in synuclein knockout experiments and

complex observations in overexpression tests.  $\alpha$ Syn knockout mice are not lethal, they show alterations in dopamine levels and impairments in synaptic responses due to changes in synaptic vesicle populations<sup>255,256</sup>. Triple knockout mice lacking all three paralogues ( $\alpha$ Syn,  $\beta$ Syn,  $\gamma$ Syn) show changes in synaptic structure and transmission as well as age-dependent neuronal dysfunction and slightly higher mortality rates (significant effect observed after 24 months)<sup>257</sup>.

Overexpression of  $\alpha$ Syn in mice suggested the decrease of neurotransmitter release at the synapse causing synaptic dysfunction<sup>258-260</sup>. Further, there is evidence, that an increased concentration of  $\alpha$ Syn reduces the size of the recycling pool of vesicles<sup>259</sup>, therefore a function in maintaining the synaptic vesicle recycling pool homeostasis is expected<sup>260</sup> (**Figure 1.19**). It should be noted that overexpression of  $\alpha$ Syn also results in the formation of toxic aggregates complicating the distinction between loss-of-function or gain-of-toxic-function.

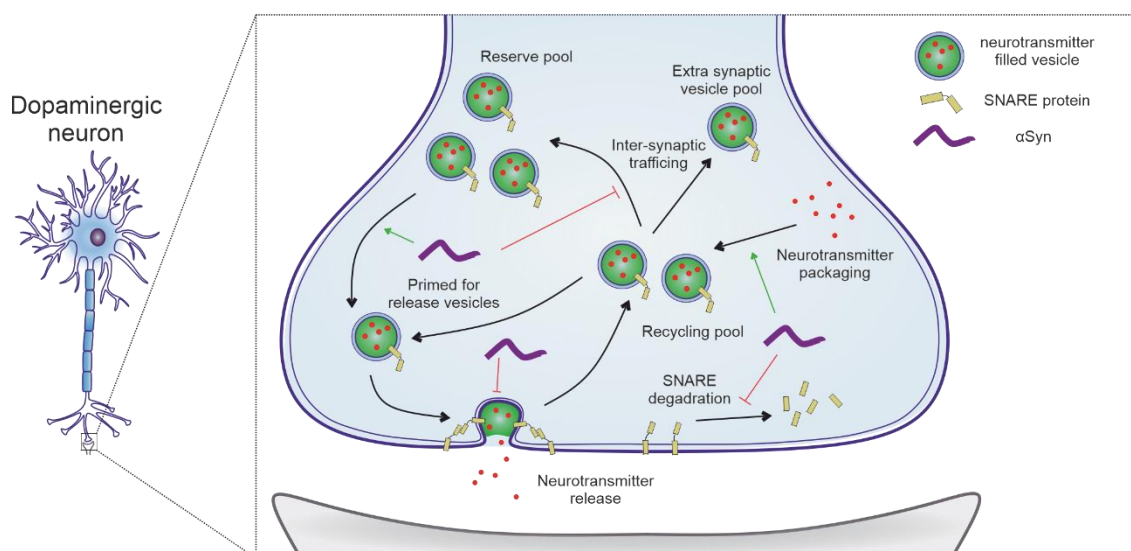
As discussed in Section 1.3, IDRs and IDPs are often enriched in proteins with important signalling or regulatory functions indicating a crucial role for flexible, non-globular domains in biological processes<sup>227,234,261</sup>. Intrinsic disorder is especially beneficial where a fine 'volume' control is needed rather than a binary on/off switch.  $\alpha$ Syn being an IDP therefore accords with the hypothesis that its function might be involved in signalling and regulation. A proteomics study with dopaminergic neurons using SILAC (stable isotope labelling by amino acids in cell culture) identified 324 proteins that form complexes with  $\alpha$ Syn. The interaction partners included mostly signalling proteins, chaperones and metabolism associated proteins<sup>262</sup>. 141 of the identified proteins displayed changes in the relative abundance after treatment with rotenone, a chemical that induces PD like toxicity<sup>262</sup>. Further, *in silico* approaches were used to identify interaction partners of  $\alpha$ Syn. A recent study used STRING<sup>263</sup> and BioGRID3.5<sup>264</sup> showing the main interaction partners to be kinases and ubiquitin proteins<sup>265</sup>. A combination of peptide pulldown assays and mass spectrometry demonstrated that interaction partners can vary after post-translational modifications. Whilst phosphorylation at residue S129 of  $\alpha$ Syn (often found in aggregated species and disease<sup>266</sup>) drives interactions with serine protein kinases, cytoskeletal and vesicular traffic proteins, non-phosphorylated  $\alpha$ Syn showed contacts with oxidative phosphorylation proteins<sup>267</sup>. All these assays demonstrate the wide range of possible functional interactions of  $\alpha$ Syn.

Another important aspect to help define the function of  $\alpha$ Syn is its location. Located in the presynaptic termini of dopaminergic neurons,  $\alpha$ Syn can be found soluble in the cytoplasm but also bound to membranes<sup>268</sup>. The binding equilibrium is tightly regulated with an estimated 15% of  $\alpha$ Syn bound to membranes<sup>269</sup>. This supports the hypothesis of  $\alpha$ Syn being responsible for the packing of neurotransmitter vesicles in the distal reserve pool of synaptic vesicles<sup>270,271</sup> and subsequent trafficking to the site of synaptic vesicle

release<sup>255,256</sup> (**Figure 1.19**). *In vitro* studies using liposomes or rat vesicles showed that  $\alpha$ Syn can bind membranes with two parts of its sequences<sup>272,273</sup> (see Section 1.4.4). This allows binding of two distinct liposomes/vesicles resulting in clustering and assembling of the liposomes<sup>272</sup>. It could further be shown that vesicle clustering by  $\alpha$ Syn is dependent on the presence of negatively charged liposomes and the vesicle-associated membrane protein VAMP2 (vesicle associated membrane protein)<sup>274</sup>.

Binding to membranes further allows  $\alpha$ Syn to act as a chaperone, promoting the SNARE (soluble N-ethylmaleimide-sensitive factor attachment protein receptor) complex to assemble and supporting vesicle fusion<sup>275</sup> (**Figure 1.19**).

Overall, the physiological function of  $\alpha$ Syn seems non-essential. However, it does interact with a wide range of proteins and is most likely involved in regulating synaptic vesicle homeostasis.



**Figure 1.19: Physiological role of  $\alpha$ Syn at pre-synaptic termini illustrating the various regulatory functions  $\alpha$ Syn exerts over synaptic vesicles.**  $\alpha$ Syn is shown in purple,  $\alpha$ Syn mediated inhibitory regulatory activity is shown by a red flat arrow head and  $\alpha$ Syn mediated regulation is shown by a green arrow.  $\alpha$ Syn promotes neurotransmitter packaging, maintains the reserve pool of synaptic vesicles and promotes subsequent trafficking to the site of vesicle release. Chaperone activity has been observed with members of the SNARE complex where it increases the stability of target SNARE proteins. Adapted and redrawn from Ref<sup>251</sup>.

## 1.4.2 Synucleopathies

Neurodegenerative disorders associated with the accumulation of aggregated  $\alpha$ Syn are summarised under the umbrella term synucleopathies<sup>276</sup>. The most occurring synucleopathy is PD, which is also the second most common neurodegenerative disorder worldwide after Alzheimer's disease (mainly caused by aggregated  $A\beta$ <sup>277</sup>). The connection between  $\alpha$ Syn and PD was made when identifying  $\alpha$ Syn deposits in PD patients, termed Lewy bodies<sup>249</sup> and the association with the familial PD mutation A53T<sup>247</sup>.

Nowadays, it is known that  $\alpha$ Syn is involved in many neurodegenerative diseases, such as Multiple system atrophy (MSA) or Dementia with Lewy bodies (DLB) (summarised in **Table 1.2**), hence adding to the complexity in determining the mechanism of toxicity during the misfolding and aggregation of this protein. CryoEM as well as using fluorescent probes, NMR spectroscopy and electron paramagnetic resonance (EPR) analysis of fibril structures from MSA and PD patients suggest that different fibril morphologies correlate with distinct disease developments<sup>117,278</sup>.

The following sections will give more insights into the genetics and pathology of PD.

**Table 1.2: Summary of diseases associated with  $\alpha$ Syn toxicity.** Table taken from Ref<sup>279</sup>.

Disease	Symptoms	Pathology
Parkinson's disease (PD)	<ul style="list-style-type: none"> <li>○ Parkinsonism (bradykinesia, muscular rigidity, resting tremors and postural instability)</li> <li>○ Nonmotor symptoms (constipation, impaired olfaction and rapid-eye movement sleep behaviour disorder)</li> <li>○ Cognitive impairment</li> </ul>	<ul style="list-style-type: none"> <li>○ Substantia nigra compacta (SNc) dopaminergic degeneration</li> <li>○ Variable neuron loss in areas including locus coeruleus, dorsal motor nucleus of the vagus and the olfactory bulb</li> <li>○ <math>\alpha</math>Syn Lewy body and neurite pathology in neurons</li> </ul>
Parkinson's disease with dementia (PDD) or dementia with Lewy bodies (DLB)	<ul style="list-style-type: none"> <li>○ Parkinsonism and dementia</li> <li>○ In DLB versus PDD: fewer resting tremors, bilateral parkinsonism</li> </ul>	<ul style="list-style-type: none"> <li>○ Cholinergic/SNc (substantia nigra pars compacta) dopaminergic degeneration</li> <li>○ <math>\alpha</math>Syn Lewy body and neurite pathology in neurons</li> <li>○ A<math>\beta</math> amyloid plaques and tau tangles</li> </ul>
Multiple system atrophy (MSA)	<ul style="list-style-type: none"> <li>○ Parkinsonism, cerebellar ataxia, autonomic failure</li> <li>○ Nonmotor symptoms (sexual dysfunction, urinary incontinence and rapid-eye movement sleep)</li> </ul>	<ul style="list-style-type: none"> <li>○ SNc/olivopontocerebellar degeneration</li> <li>○ <math>\alpha</math>Syn pathology in oligodendrocytes</li> </ul>
Gaucher's disease	<ul style="list-style-type: none"> <li>○ Type I adult-onset (thrombocytopenia, anaemia, hepatosplenomegaly and bone pain)</li> <li>○ Type II (infant)/III (juvenile) neuropathic form (seizures, cognitive impairment and oculomotor problems)</li> </ul>	<ul style="list-style-type: none"> <li>○ <math>\alpha</math>Syn Lewy body and neurite pathology in some patients</li> </ul>
Additional lysosomal storage disorders	<ul style="list-style-type: none"> <li>○ Multisystem disorder</li> </ul>	<ul style="list-style-type: none"> <li>○ <math>\alpha</math>Syn Lewy body and neurite pathology in some patients</li> </ul>
Neurodegeneration with brain iron accumulation	<ul style="list-style-type: none"> <li>○ Variable symptoms that may include dystonia, muscle rigidity, spasticity and ataxia</li> </ul>	<ul style="list-style-type: none"> <li>○ Iron accumulation in the globus pallidus and SNc</li> <li>○ <math>\alpha</math>Syn Lewy body and neurite pathology in some patients</li> </ul>
Alzheimer's disease (AD)	<ul style="list-style-type: none"> <li>○ Progressive memory loss</li> <li>○ Cognitive impairment</li> </ul>	<ul style="list-style-type: none"> <li>○ Cortical and CA1 hippocampal degeneration</li> <li>○ A<math>\beta</math> amyloid plaques and tau tangles</li> <li>○ <math>\alpha</math>Syn Lewy body and neurite pathology in some patients</li> </ul>

### 1.4.2.1 Genetics of PD

Although PD was long considered a non-genetic disorder of ‘sporadic’ origin, 5–10% of patients are now known to have monogenic forms of the disease<sup>280</sup>. Today, there are 12 known familial PD mutations in the SNCA gene encoding  $\alpha$ Syn (the most common and best analysed A30P<sup>171</sup>, A30G<sup>281</sup>, E46K<sup>282</sup>, H50Q<sup>283</sup>, G51D<sup>284</sup>, A53T<sup>247</sup>, A53E<sup>285</sup>, A53V<sup>286</sup> as well as M5T<sup>287</sup>, T72M<sup>248</sup>, G93A<sup>287</sup>, and P117S<sup>287</sup>) expanded by few missense mutation with no characterised familial history but association with PD (L8I<sup>288</sup>, V15A<sup>289</sup>, V15D<sup>290</sup>, A18T, A29S<sup>291,292</sup>, E57D<sup>293</sup>, and M127I<sup>290</sup>). Recently, also a DLB linked mutation was found (E83Q<sup>294</sup>) (**Table 1.3**). Further causes of autosomal dominant forms of familial PD are genomic duplications or triplications that contain the  $\alpha$ Syn locus<sup>248,295,296</sup>. Genome wide association studies have also identified single nucleotide polymorphisms (SNPs) in  $\alpha$ Syn as risk factors that increase the susceptibility to sporadic PD<sup>297</sup>.

Interestingly, not all PD associated protein variants increase the fibrillation capacity of the protein<sup>145,298</sup>. This hints to the idea that not only fibrils, but also oligomeric species are neurotoxic and involved in disease (Section 1.2.4) and further highlights the complexity of amyloid linked disease<sup>159</sup>. In addition to this, also the distinct membrane binding (note that all familial PD mutations besides P117S are within the membrane binding site) and surface induced aggregation characteristics change for the different variants<sup>172</sup> as well as final fibril morphologies<sup>94,96</sup>. Despite intense study, the pathological role(s) of all these mutations remains unknown.

It should be noted that polymorphism not only in the SNCA gene, but also in genes such as LRRK2, Parkin or PINK1 is involved in PD<sup>280</sup>.

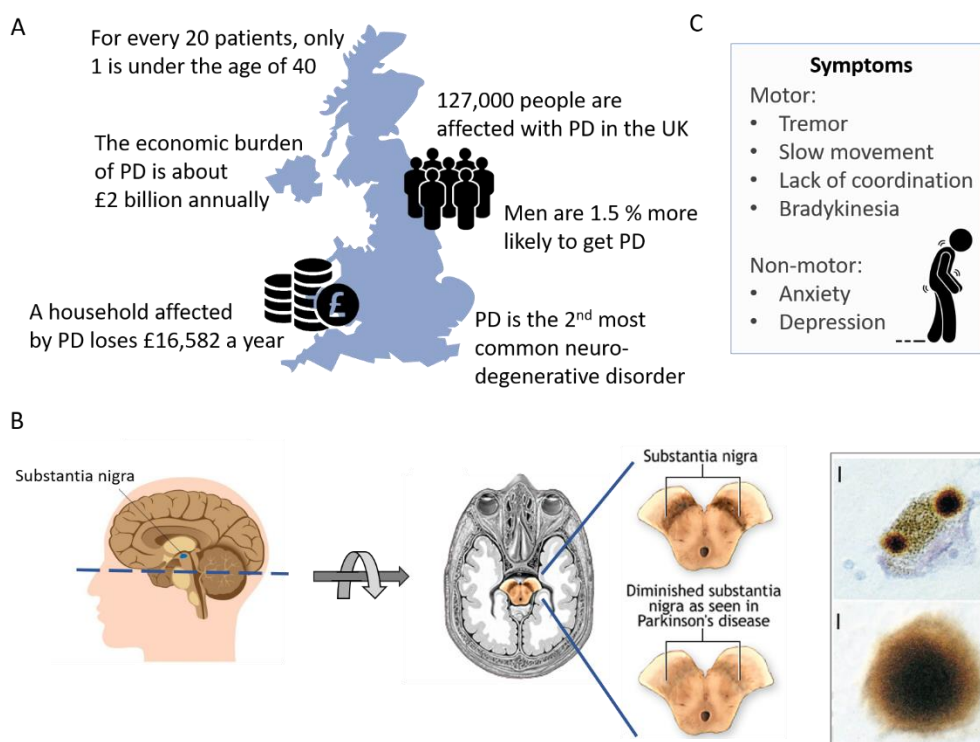


**Table 1.3: Summary of identified synucleopathy associated mutations in the SNCA gene causing missense mutations in  $\alpha$ Syn.** Most analysed variants are highlighted in light blue. Origin, disease association and influence on fibril formation is indicated.

Mutation	Origin	Disease	Influence on fibril formation compared to WT	Literature
M5T	Chinese	Familial PD	No data	287
A30P	German	Familial PD	retard	171
A30G	Greek	Familial PD	no significant change	281
E46K	Spanish	Familial PD	accelerate	282
H50Q	British	Familial PD	accelerate	283
G51D	British	Familial PD	retard	284
A53E	Finnish	Familial PD	retard	285
A53T	Italian	Familial PD	accelerate	247
A53V	Japanese	Familial PD	retard	286
T72M	Turkish	Familial PD	No data	248
G93A	Chinese	Familial PD	No data	287
P117S	Chinese	Familial PD	No data	287
L8I	Chinese	Sporadic PD	No data	288
V15A	Chinese	Sporadic PD	No data	289
V15D	Chinese	Sporadic PD	No data	290
A18T	Polish	Sporadic PD	accelerate	291
A29S	Polish	Sporadic PD	accelerate	291
E57D	Korean	Sporadic PD	No data	293
M127I	Chinese	Sporadic PD	No data	290
E83Q	-	Familial DLB	accelerate	294,299

#### 1.4.2.2 Pathology in PD

PD affects 10 million people worldwide with men having a 1.5-fold higher risk than women<sup>300,301</sup>. 1 % of individuals over the age of 60 and 5 % of the population over the age of 85 suffer from PD illustrating the impact that advancing age has on the risk of developing PD<sup>302</sup> (**Figure 1.20 A**). In PD, dopaminergic neurons of the substantia nigra pars compacta (SNc) are lost in the basal ganglia (**Figure 1.20 B**), an area of the brain responsible for co-ordinating fine motor control, which ultimately leads to the onset of Parkinsonism symptoms such as bradykinesia, muscle rigidity, resting tremors and postural instability<sup>303</sup> (**Figure 1.20 C**). Within the few surviving dopaminergic neurons, Lewy bodies and Lewy neurites can be found which show a high concentration of aggregated  $\alpha$ Syn<sup>304</sup>. Using CLEM (correlative light and electron microscopy) and STED (stimulated emission depletion) microscopy, Lauer and co-workers recently showed a crowded environment within Lewy bodies built by lipid membrane fragments and distorted organelles together with a non-fibrillar form of  $\alpha$ Syn<sup>76</sup>. However, if  $\alpha$ Syn is aggregated in Lewy bodies or not remains inconclusive and there is still an ongoing discussion in the field whether Lewy bodies are the cause or a consequence of PD and whether they include  $\alpha$ Syn fibrils or not<sup>305</sup>.



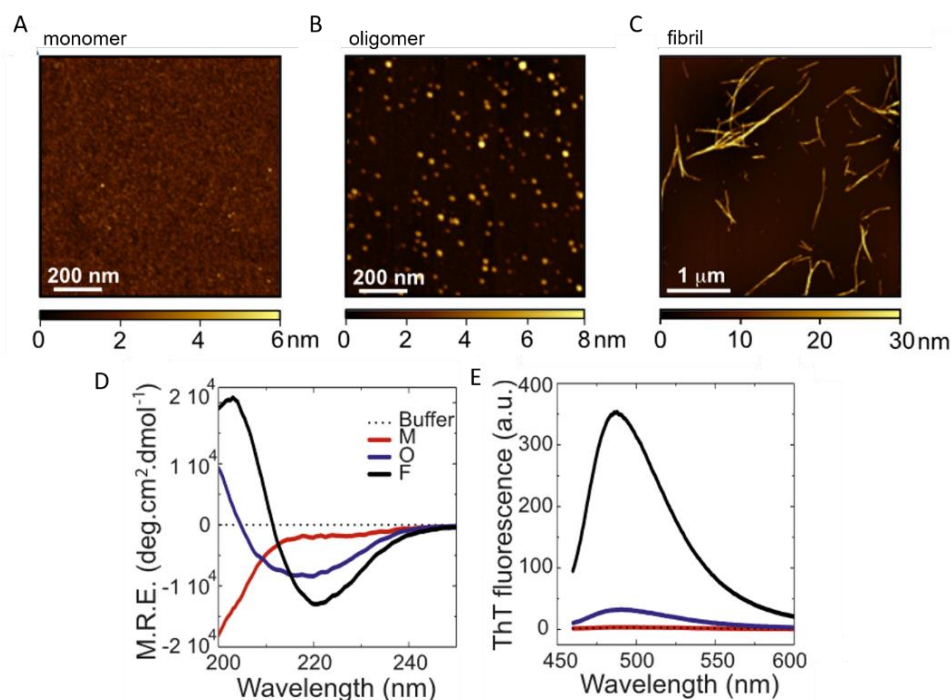
**Figure 1.20: Significance and symptoms of Parkinson's disease.** (A) Statistical overview of the number of patients and financial aspects of PD in the UK highlighting the importance of research in this field. (B) Left: Pathological symptoms of PD which includes in the loss of neurons in the substantia nigra, the formations of Lewy bodies and an accumulation of aggregated  $\alpha$ Syn in the few surviving cells. Right: Pigmented nerve cells with  $\alpha$ Syn-positive Lewy bodies. Upper scale bar 8  $\mu$ m; lower scale bar 4  $\mu$ m. Figure taken from<sup>249</sup>. (C) List of the motor and non-motor symptoms of PD taken from<sup>306</sup>.

To study the pathology of synucleopathies (including PD) and associated cellular processes, it is crucial to have good model organisms mimicking the diseases. Although *in vitro* assays using heterogeneously expressed  $\alpha$ Syn protein facilitate the well-controlled analysis of aggregation kinetics, it is missing the crucial factor of the cellular environment including buffer changes, interactions with other proteins and the complexity of a cell metabolism. Therefore various models have been used such as cell cultures (ranging from yeast over primary neurons to patient derived cell lines)<sup>307</sup>, worms<sup>83</sup>, zebrafish<sup>308</sup>, flies<sup>309</sup> or mice<sup>270</sup>. Importantly, vertebrate systems are beneficial as they naturally express  $\alpha$ Syn, however studies using these organisms might be more expensive, complex and time intensive. It is therefore vital to find the right model for the biological questions to answer.

### 1.4.3 Structural properties

$\alpha$ Syn can adopt a large conformational ensemble ranging from different monomeric structures (e.g. intrinsically disordered or helical when bound to membranes), to various oligomeric species and fibrils. The transition from monomers to fibrils is highly complex and can be followed by using a variety of techniques. As the characteristics of

monomers, oligomers, and fibrils drastically change (e.g. size, morphology, secondary structure elements, ThT binding (**Figure 1.21**)), different experimental techniques are required to characterise their structures in more detail.

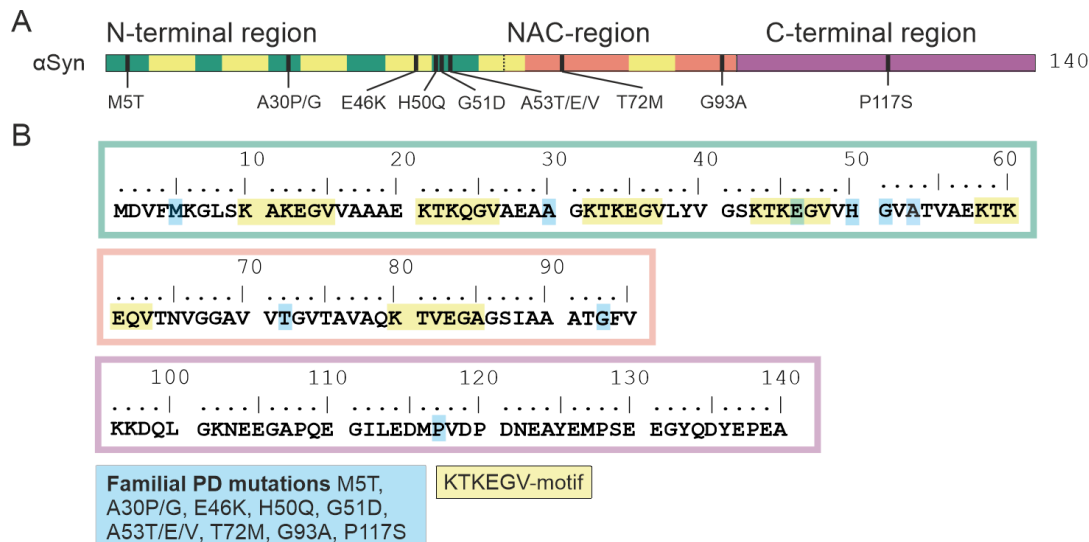


**Figure 1.21: Morphological and structural characterisation of  $\alpha$ Syn monomers, oligomers and fibrils.** Examples of AFM images of monomeric (A), oligomeric (B), and fibrillar (C)  $\alpha$ Syn species are shown. Far-UV CD (D) and ThT fluorescence (E) for all three  $\alpha$ Syn species, monomer (red), oligomer (blue) and fibril (black). Figure adapted from Ref<sup>166</sup>.

#### 1.4.3.1 Monomer

Alpha-synuclein is a 140 amino acid long, intrinsically disordered protein (**Figure 1.22 A**). Its primary sequence can be divided into three regions: the N-terminal region (residues 1-60), the non-amyloid  $\beta$  component (NAC) region (residues 61-95) and the C-terminal region (residues 96-140). These three parts have defined biochemical characteristics supporting the physiological function and aggregation propensity of this protein. The amphipathic N-terminal region consists of five imperfect repeats with the consensus sequence KTKEGV (there is a sixth one located in NAC) separated by inter-repeat regions (**Figure 1.22 B**). This region, is especially involved in membrane binding<sup>310</sup>, important for the physiological function of  $\alpha$ Syn (see Sections 1.4.1 and 1.4.4). As a result of Eisenberg and co-worker's work on amyloid like crystals formed by peptides spanning residues 47-56, this region is sometimes individually identified as preNAC<sup>311</sup>. The central NAC region is highly hydrophobic and was originally discovered through its co-purification with A $\beta$  from amyloid plaques in AD patients, hence its appellation<sup>312</sup>. There is a large interest in NAC as it has been shown to be crucial in the misfolding and aggregation of  $\alpha$ Syn<sup>312,313</sup> and forms the core of all amyloid structures to

date<sup>96,117</sup>. Further, NAC was also demonstrated to be involved in lipid binding in synergy with the N-terminal region<sup>310,314</sup>. Last but not least, the C-terminal domain is enriched with acidic and proline residues (**Figure 1.22 B**), providing high flexibility and protection against aggregation<sup>190</sup>. This region is also mainly involved in ion binding, which strongly affects the aggregation propensity of full-length  $\alpha$ Syn<sup>315</sup>.



**Figure 1.22: Sequence identity of  $\alpha$ Syn.** The protein comprises three regions: the amphipathic N-region (green), the amyloidogenic NAC region (red) and the acidic C-terminal region (purple). The KTKEGV repeat motif is coloured in yellow, positions of familial PD mutations is indicated below the schematic. **(B)** Amino acid sequence of  $\alpha$ Syn (regions coloured according to **(A)**). Positions of familial PD mutations of  $\alpha$ Syn are highlighted in blue and each of the KTKEGV motifs involved in membrane binding are highlighted in yellow.

In general,  $\alpha$ Syn exists as an IDP lacking extensive secondary or tertiary structure<sup>316</sup> (**Figure 1.21 D**). Studies on  $\alpha$ Syn have shown that it had a larger hydrodynamic radius ( $R_h$ ) (measured by size exclusion chromatography (SEC)) and sediments more slowly (measured by a sucrose gradient ultracentrifugation) compared to globular proteins of similar masses (14 kDa)<sup>252</sup>. CD and Fourier transform infrared (FTIR) spectroscopy confirmed the absence of a defined 3D structure<sup>317</sup>.

Interestingly, small angle X-ray scattering (SAXS) experiments indicated that  $\alpha$ Syn has a smaller radius of gyration ( $R_g$ ) than expected of a random coil polypeptide of the same length (40 Å compared to a theoretical  $R_g$  of ~52 Å)<sup>317</sup>. This suggests that  $\alpha$ Syn forms some transient intramolecular structure as it is more compact than a ‘perfect’ random coil protein<sup>317</sup>. This observation of a relative compactness of  $\alpha$ Syn has now been confirmed via multiple approaches and is believed to be driven by clustering of hydrophobic residues and transient long range interactions<sup>318-323</sup> (**Table 1.4**).

**Table 1.4: Regions of intramolecular interactions within monomeric  $\alpha$ Syn.** Summarised are the characterised interactions, the positions of the label (if required), the buffer conditions and protein concentration of the experiment, and the used techniques. Adapted from Ref<sup>189</sup>.

Interaction region 1	Interaction region 2	Label position	Conditions	technique	Notes	Lit.
5-26 and 40-55	93-140	-	20 mM NaP, pH 4.0	HDX-MS	Using degree of protection as indicator	324
30-100	C-term., 120-140	Q24, S42, Q62, S87, N103	10 mM NaP, pH 7.4, 100 mM NaCl 100 $\mu$ M protein	Spin label NMR	Binding MTSL to induced Cys variants (see labelling position)	325
22-93	C-term.		25 mM Tris-HCl, pH 7.4, 0.1 M NaCl 100 $\mu$ M protein	HSQC NMR	Polyamine binding enhances aggregation and leads to structural changes	326
12-26	C-term.	A18, A90, A140	25 mM Tris-HCl, pH 7.4, 1 M NaCl	RDC (Residual dipolar coupling) and PRE NMR	Polyamine binding and/or temperature increase interrupts interactions	321
105-115	120-130	A140				321
NAC	125, 133, 136	-	20 mM Tris, pH 7.2	Met oxid., MS	oxidation of methionines by H <sub>2</sub> O <sub>2</sub> inhibited fibrillation	327
Met1	Met116, 127	-				327
1-10	124-138	-	pH 7.4	Comp. analysis	analysed NMR PRE previously published	328
38-56, 74-80, 90-100	104-130	-				328
6-10	136-140	-	-	Simulation, RDC NMR	Measure short- and long-range interactions	329
3	124, 140	V3, V71, A124, A140	20 mM Gly-NaOH, pH 7.4 100 $\mu$ M protein	Tyr fluorescence	Also tested the effect of PD mutations which are more compact	330
19	60-140	A19, A90, G132	PBS, pH 7.4 250 $\mu$ M protein	PRE NMR	low pH shows a more homogeneous conformational ensemble	331
132	70-140	A19, A90, G132	10 mM NaP, 140 mM NaCl, pH 2.5 250 $\mu$ M protein			
N-term.	C-term.	S9, E20, E61, A85, E110, E130	pH 7.4 70 $\mu$ M protein	RDC and PRE NMR	N and C contacts remain, C term. collapses. Hydroph. interactions drive agg. rather than charge alteration.	332
N-term	C-term.		100 mM NaCl, 10 mM NaP, pH 3.0, 70 $\mu$ M protein			
20 ~10 and 90	Around 128 120	E20, A85, P120	10 mM NaP, pH 7.4, 100 mM NaCl 70 $\mu$ M protein	PRE NMR	Also analysis of $\beta$ - and $\gamma$ Syn	333
N-term.	109-137		pH 6.0 70 $\mu$ M protein			
N-term	C-term.	-	20 mM NaP, pH 7.5	HSQC NMR	Comparison in CSPs of WT with familial PD mutations	334
3-6	124-127, 133-135	-	20 mM NaP, pH 6.0			
N-term.	C-term.	A18, A76, A90, A140	pH 7.4 100 $\mu$ M protein	PRE NMR	Rigidification and compaction of C-term at acidic pH	335
NAC	120-140		20 mM Na acetate, 100 mM NaCl, pH 3.0 100 $\mu$ M protein			

The disordered nature of  $\alpha$ Syn is a consequence of its relatively low hydrophobicity and high net charge (see Section 1.3). Therefore, the conformational ensemble of  $\alpha$ Syn is highly sensitive to the surrounding environmental conditions. Salt concentrations, ion binding, pH and the protein's charge have large effects on the transient intra- and intermolecular interactions of  $\alpha$ Syn as it has been shown mainly by NMR and computational approaches (**Table 1.4**), reviewed in Ref<sup>189</sup>. Overall, at neutral pH the N-terminal and C-terminal regions form long-range interactions protecting the aggregation-prone NAC region from solvent exposure and fibril formation<sup>321,335</sup>. Under acidic conditions, the Asp-/Glu-rich C-terminal region collapses, intramolecular interactions are perturbed, and the NAC region is less protected resulting in accelerated fibril assembly<sup>189,332,335</sup>. This is in consensus with *in vitro* aggregation studies using Thioflavin T (ThT) fluorescence assays to detect fibril growth at different pH values<sup>133</sup> or research with C-terminal truncated  $\alpha$ Syn variants that speed up aggregation due to an exposed NAC region<sup>190</sup>. Some studies also describe the increased hydrophobicity of the protein (and collapsed C-terminus) as the main contribution for the altered aggregation kinetics when decreasing the pH, rather than the release of long-range contacts<sup>332</sup>.

Determination of structural ensembles of IDPs is challenging but became more feasible during the past years due to improvements in computational modelling and simulations by, for example, optimising force fields for dynamic, flexible proteins<sup>336</sup>. Importantly, in most cases the computationally calculated conformations overlap with independent experimental observations but describe an ensemble rather than a static conformation<sup>336,337</sup>. For  $\alpha$ Syn, the large effect of pH on the monomer was evaluated using computational approaches confirming the collapse of the C-terminal region at acidic pH but also suggesting that the overall end-to-end distance within the molecule are maintained due to internal reorganisation<sup>338</sup>. Further, the conformational state of different familial PD mutations was analysed without the need of expression, purification and complex laboratory experiments, highlighting that the dynamic equilibrium between different  $\alpha$ Syn monomer conformations is modified by the missense mutations in a subtle way<sup>339</sup>. MD simulations were also shown to evaluate small-molecule binding to  $\alpha$ Syn monomer, providing a large benefit for drug development in synucleopathies treatment<sup>340</sup>.

Monomeric  $\alpha$ Syn can adopt clearly defined secondary structural elements under specific conditions. When binding to micelles or membranes such as vesicles, the N-terminal region and NAC form an  $\alpha$ -helix<sup>186,314</sup> (Section 1.4.4).  $\beta$ -hairpin formation in the N-terminal region of the protein is believed to be important for fibril assembly<sup>188,341</sup> and can be observed when binding certain molecules such as  $\beta$ -wrapin, a small protein inhibiting  $\alpha$ Syn aggregation<sup>187,342</sup>.

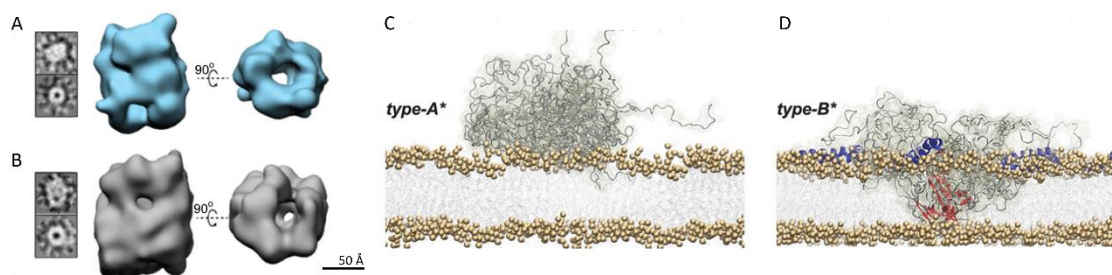
Taken together, the structure of monomeric  $\alpha$ Syn is complex and strongly depends on its environment. A large conformational ensemble can be observed with transient interactions favouring or disfavouring the formation of higher ordered structures leading to fibril formation. The knowledge about exact conformations and interactions involved in amyloid assembly remains incomplete.

### 1.4.3.2 Oligomers

Oligomeric species are often considered as being the cytotoxic state of  $\alpha$ Syn<sup>159</sup> and therefore their composition and architecture is of highest interest. The structural characterisation is challenging, as oligomeric samples can be heterogeneous, highly dynamic, and stable for only a short period of time at low concentrations. Classical structural techniques such as cryo-EM, NMR or X-ray scattering struggle to describe these small aggregates. Yet, experiments using transmission electron microscopy (TEM), AFM, SAXS or HDX-MS helped to analyse the structure of oligomers in more detail.

Due to differences in the way oligomers are generated and purified, they are reported to form a large range of structures<sup>343</sup>. However, it is suggested, based on HDX-MS experiments, that oligomers are relatively ordered assemblies with elongated  $\beta$ -strands formed by the N-terminal region and NAC and an exposed C-terminus<sup>344</sup>. The probably best characterised oligomers are Type-A(\*) and Type-B(\*), their structures were determined via cryoEM with resolutions of 18 Å and 19 Å, respectively (**Figure 1.23 A,B**). These two oligomeric species present distinct conformations and cytotoxicities (e.g. Type-B\* has more  $\beta$ -sheet content shown by ssNMR studies and disrupts SH-SY5Y cells and rat primary cortical neuron membranes more efficiently)<sup>160,166</sup>. Whereas Type-A\* oligomers bind on the surface of membranes without much of an effect, Type-B\* assemblies were shown to bind and disrupt membranes (**Figure 1.23 C,D**), being more toxic for cells by increasing intracellular reactive oxygen species and reducing mitochondrial activity in neurons<sup>160</sup>.

It should be noted that although oligomers are the precursor species of fibrils, not all oligomeric structures can be elongated to mature fibrils. It is distinguished between on- and off-pathway oligomers<sup>159</sup>.



**Figure 1.23: Oligomer structures and effect on membrane binding.** Typical side view and end-on view of Type-A (A) and type-B (B) oligomers determined by cryoEM. Schematic of interactions of  $\alpha$ S oligomers (Type-A\* (C) and Type-B\* (D)) with lipid bilayers. Figures taken from Ref<sup>160,166</sup>.

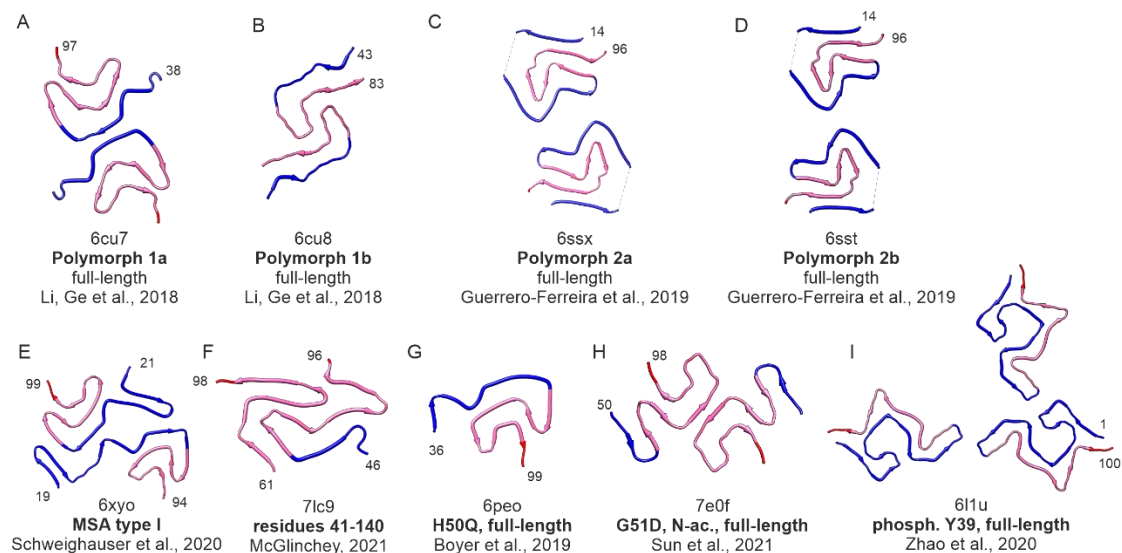
### 1.4.3.3 Fibrils

Amyloid fibrils present the thermodynamically and kinetically most stable conformation. The fast improvement of cryoEM techniques over the past few years has allowed the determination of high resolution structures of many  $\alpha$ Syn architectures (Figure 1.24). Fibrils are polymorphic and present distinct structures for *in vitro* or *in vivo* grown assemblies. Full-length  $\alpha$ Syn fibrils grown *in vitro* show four main structures named 1a, 1b, 2a and 2b<sup>72,116</sup>. They all form the rigid amyloid core with a part of the N-terminal region and NAC, whilst the extreme N-terminus (~residues 1-35) and the C-terminal region (~residues 96-140) are flexible and dynamic. A mature fibril usually consists of two protofilament twisted around each other. The residues involved in the interface between protofilaments and the backbone structure vary, though (Figure 1.24 A-D). Interestingly, fibril architectures from samples extracted from MSA patients were found to have a different structure compared to *in vitro* grown  $\alpha$ Syn WT fibrils<sup>117</sup> (Figure 1.24 E) and different diseases show distinct fibril morphologies<sup>117,278</sup>.

Further, posttranslational modifications, often observed in patients, or single residue substitutions associated with PD result in altered fibril structures (Figure 1.24 F-I). Fibril morphologies of N- and C-terminal truncations result in the 1a polymorph (deleting the C-terminus, not shown here) or structures unrelated to WT architectures solved so far<sup>106,345,346</sup>. Amyloid from PD mutations or phosphorylation at residue Y39 present structures with only one, two or even three protofilaments and also often represent distinct backbone conformations compared to the WT structures (Figure 1.24 G,I)<sup>107-109,347-349</sup>.

Altogether,  $\alpha$ Syn fibrils can be highly polymorphic and their structure(s) seems to be an indicator for disease. Even small sequence changes can completely alter the morphology, demonstrating an important role of single residues for aggregation.





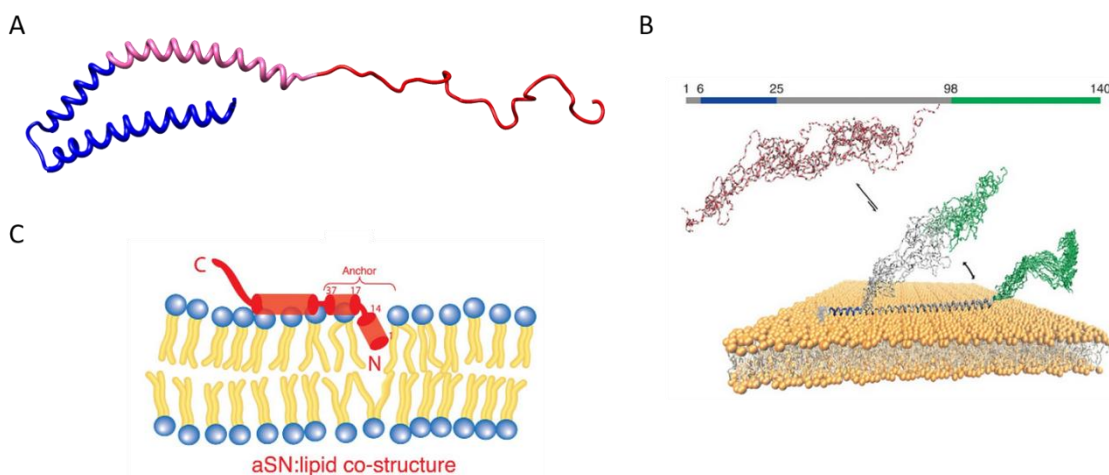
**Figure 1.24: Examples of  $\alpha$ Syn fibril structures determined by cryoEM.** Full-length structures form polymorph 1a (A), 1b (B), 2a (C), 2b (D). Fibrils extracted from MSA patients (E), or those formed from proteins with truncations (F), disease associated point mutations (G,H) or phosphorylation (I) can result in distinct architectures compared to the WT protein. Note that this figure does not illustrate all solved fibril structures and focusses on highlighting the high polymorphism of  $\alpha$ Syn fibril structures. The N-terminal region is coloured in blue, NAC in pink and the C-terminal region in red. If regions are not shown, they were not resolved and not part of the rigid fibril core, residues within the core are indicated. PDB file, special characteristics and references are indicated below each fibril structure.

#### 1.4.4 Lipid interactions

About 15% of  $\alpha$ Syn is not free in the cytosol, but bound to lipid membranes within the cell<sup>269</sup>. The interactions between  $\alpha$ Syn and membranous structures is crucial as being involved in physiological function<sup>255,259,350</sup> as well as fibril formation<sup>269,351</sup>. As mentioned above (Section 1.4.3.1),  $\alpha$ Syn forms an  $\alpha$ -helical structure when binding to micelles or lipid bilayers<sup>185,314,352</sup>. A ssNMR study on  $\alpha$ Syn interacting with SDS (Sodium-dodecyl-sulfate) micelles identified that residues 3-37 and 45-92 form curved  $\alpha$ -helices, connected by a well ordered, extended linker, while the C-terminal region remained disordered<sup>314</sup> (**Figure 1.25 A**). By contrast, structural characterisations using EPR and vesicles formed by POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) and POPS (1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine) lipids defined a membrane-bound region featuring a single extended helix spanning the first 97 N-terminal residues<sup>353</sup>. There is also evidence for interchanging structures between these two models instead of exclusive conformations<sup>354,355</sup>.

Despite these differences in helix morphology, the key feature for membrane binding in a biological context seems to be a highly dynamical behaviour. Residues 1-25 of the  $\alpha$ Syn sequence have a fundamental relevance for the overall equilibrium between membrane bound and unbound states, as they act as an anchor to initiate the adhesion to lipid vesicles<sup>186,352,356</sup>. The rest of the N-terminal region and NAC (residues 26-97) exist

in an equilibrium between membrane tethered conformations whereas the C-terminal region was not shown to bind to the membrane surface (**Figure 1.25 B**)<sup>186,352,356</sup>. A different study from Cholak *et al.* suggested that membrane binding is defined by two processes: 1. The N-terminal tail insertion (first 14 residues) to the lipid head group layer and folding to a helix when binding at the membrane surface for N-terminal region and NAC (**Figure 1.25 C**)<sup>185</sup>.

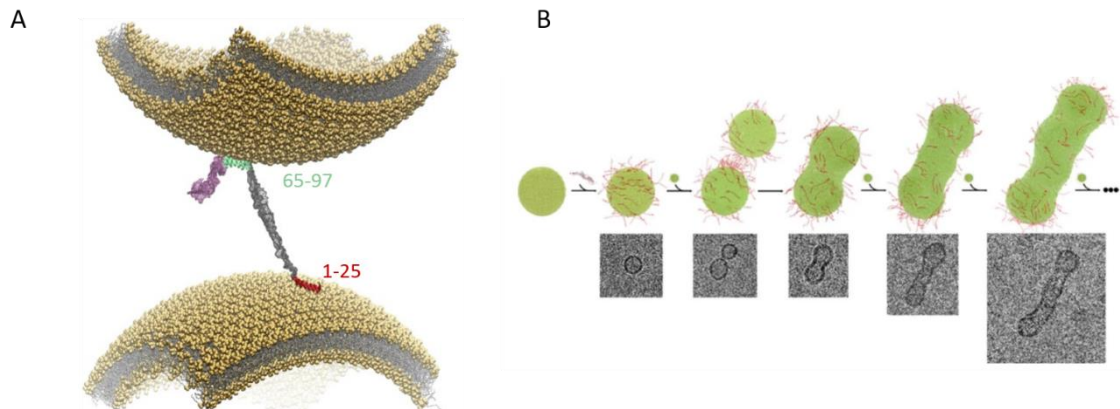


**Figure 1.25: Summary of  $\alpha$ Syn binding to membranes.** (A) NMR structure of  $\alpha$ Syn bound to SDS micelles (PDB 1xq8)<sup>314</sup>. N-terminal region is coloured in blue, NAC in pink and C-terminal region in red. (B) Illustration of the highly dynamic nature of  $\alpha$ Syn membrane binding. The first 25 N-terminal residues (blue) initiate membrane binding whereas residues 26-97 (grey) are in equilibrium between bound and unbound and the C-terminal region (green) is not bound. Figure taken from Ref<sup>186</sup>. (C) Schematic of N-terminal insertion to membranes with residues 1-14, figure taken from Ref<sup>185</sup>.

It should be further noticed that the membrane binding characteristics (e.g. binding affinity, free energy) strongly depend on the membrane composition and curvature<sup>357,358</sup>. It was demonstrated that  $\alpha$ Syn preferably binds to small unilamellar vesicles over large ones (>15-fold increase in affinity for 50 nm diameter vesicles over 180 nm diameter). The protein also predominantly interacts with anionic, highly charged lipids (e.g. phosphatidic acid head groups)<sup>358</sup>.

For the physiological function of  $\alpha$ Syn the ability to cluster synaptic vesicles is important. An *in vitro* study revealed the 'double-anchor mechanism' where  $\alpha$ Syn binds to different liposomes with its N-terminus (residues 1-25) and region 65-97. This enables a single  $\alpha$ Syn molecule to simultaneously bind two vesicles that are up to 150 Å apart (**Figure 1.26 A**)<sup>272</sup>. SUVs (small unilamellar vesicles) initially assemble together in dimeric, trimeric, tetrameric, and higher order states, before fusing to form larger vesicles (**Figure 1.26 B**). Fused liposomes were observed to mainly add further liposomes at the termini probably due to the higher affinity of  $\alpha$ Syn for significantly curved membrane surfaces<sup>242</sup>. The interaction between  $\alpha$ Syn and membranes as well as vesicle clustering can be affected by the presence of ions. The ability of  $\alpha$ Syn to remodel membranes rises as the

concentration of divalent metal ions increases<sup>359</sup>. For example, it was shown that calcium mediates the localisation of  $\alpha$ Syn at the pre-synaptic terminal, accelerating aggregation and vesicle tethering *in vitro* and *in vivo*<sup>273</sup>.



**Figure 1.26: Schematic of liposome clustering by  $\alpha$ Syn.** (A) Molecular details of the double-anchor mechanism. The N-terminal anchor region is coloured in red, region 65–97 is highlighted in green. C-terminal region and linker region are coloured in pink and grey, respectively. (B) Stepwise mechanism of vesicles assembly as probed from images obtained *in vitro* by cryo-EM (bottom). Vesicles are illustrated in green,  $\alpha$ Syn in red. Figure adapted from Ref<sup>272</sup>.

Membrane binding is not only associated with physiological function, but also with aggregation and toxicity. Lipid bilayers offer a surface, promoting  $\alpha$ Syn fibril formation via primary nucleation<sup>351</sup>. The molecular processes behind this might be driven by accumulating  $\alpha$ Syn on the surface of the liposomes and/or via conformational changes towards more aggregation-prone states. Additionally, deep mutational scanning studies indicated that the helical, membrane bound state of  $\alpha$ Syn is the most toxic conformation in yeast cells as mutations interrupting membrane binding protect from cell death<sup>360,361</sup>.

In addition to the interaction between lipids and monomeric  $\alpha$ Syn, the correlation with oligomeric species and fibrils has been analysed. As mentioned above (Section 1.4.3.2), some oligomers were shown to bind and disrupt membranes<sup>160</sup>. In contrast, *in situ* analysis using cryo-electron tomography (cryoET) focussing on  $\alpha$ Syn fibrils indicated that  $\alpha$ Syn fibrils do not contact membranes directly and are not involved in vesicle clustering<sup>75</sup>.

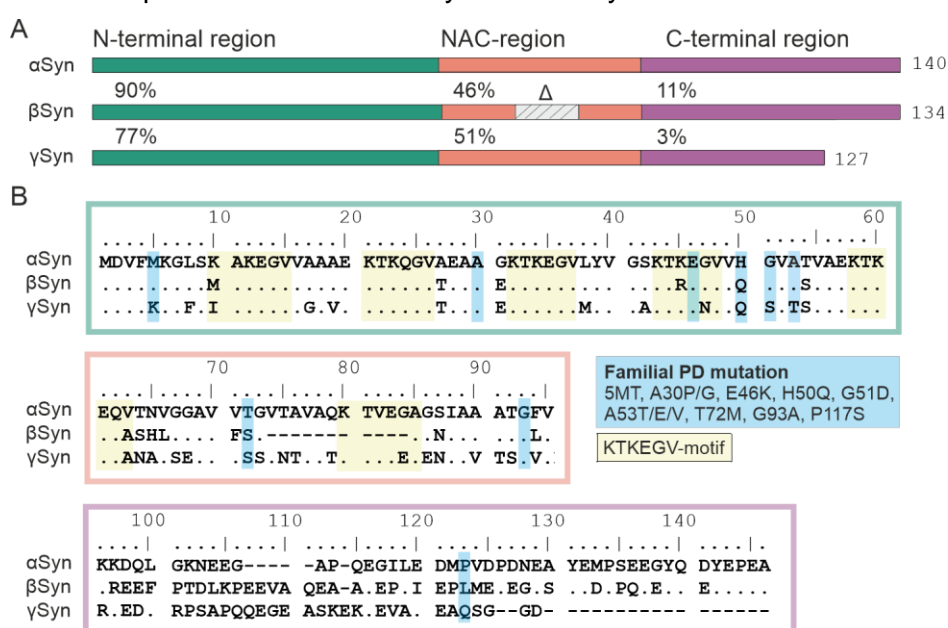
## 1.5 Synuclein paralogues

$\alpha$ Syn is part of the synuclein family, consisting of two other IDP paralogues, beta-synuclein ( $\beta$ Syn) and gamma-synuclein ( $\gamma$ Syn). Whilst  $\alpha$ - and  $\beta$ Syn are mainly expressed in the central nervous system in the pre-synaptic termini of neurons throughout the brain,  $\gamma$ Syn is primarily localised in more differentiated nerve cells in the peripheral nervous system (e.g. motor neurons)<sup>362</sup>.  $\alpha$ Syn is known to be involved in neurodegenerative disorders (**Table 1.2**), but neither  $\beta$ Syn nor  $\gamma$ Syn have been convincingly linked to diseases of the nervous system with only a few studies showing the presence of these

proteins in PD or DLB deposits, hitherto<sup>363,364</sup>. *In vitro* studies supported these findings as  $\beta$ - and  $\gamma$ Syn are both significantly less aggregation-prone compared to  $\alpha$ Syn<sup>365,366</sup>. In fact, the paralogues even inhibit assembly of the aggregation-prone  $\alpha$ Syn variant when co-incubated<sup>365,367</sup>.

Interestingly, the sequence similarity between the full-length paralogues and  $\alpha$ Syn is high with 78% and 60% for  $\beta$ Syn and  $\gamma$ Syn, respectively (**Figure 1.27**). Especially the N-terminal region is highly conserved with a decreasing sequence identity for NAC and the C-terminal region.

The reduced aggregation propensity for  $\beta$ Syn may be explained by the lack of eleven residues in the central NAC region<sup>333</sup>. For  $\gamma$ Syn, however, there are no obvious sequence alterations that explain the lower tendency to form amyloid fibrils.



**Figure 1.27: Sequence alignment between the synuclein family members  $\alpha$ Syn,  $\beta$ Syn and  $\gamma$ Syn.** (A) Each protein comprises three regions: the amphipathic N-region (green), the amyloidogenic NAC region (red) and the acidic C-terminal region (purple). The sequence identity of  $\beta$ Syn and  $\gamma$ Syn to  $\alpha$ Syn for each region is shown. The light grey region in the centre of NAC for  $\beta$ Syn depicts residues in the highly aggregation NAC core that is deleted in this sequence. The number of residues in each protein are shown (right). (B) Sequence alignment for each of the paralogues (regions coloured according to (A)). Positions of familial PD mutations of  $\alpha$ Syn are highlighted in blue and each of the KTKEGV motifs are highlighted in yellow.

## 1.5.1 Gamma-synuclein

### 1.5.1.1 Function and disease

$\gamma$ Syn, which will also be investigated in this thesis, is significantly less studied than  $\alpha$ Syn in the literature, therefore there are many deficiencies in understanding the biology and biochemistry of this protein.  $\gamma$ Syn was discovered in 1997/1998 by two individual

groups<sup>368,369</sup>. Different to  $\alpha$ Syn,  $\gamma$ Syn is only 127 amino acid long (lacking 13 residues in the C-terminal region) and encoded by the *SNCG* gene localized on chromosome 10q23<sup>370</sup>.

The function of  $\gamma$ Syn is not yet clarified<sup>363</sup>, but it is believed to interact with heat-shock proteins where it acts as a co-chaperone to stimulate estrogen receptor signalling<sup>371</sup> and it was shown to interact with BubR1 (a gene encoding a serine/threonine-protein kinase) to regulate cell cycle checkpoints<sup>372</sup>. Related to these functions,  $\gamma$ Syn is involved in disease and is overexpressed in breast and ovarian cancer cells<sup>368,373,374</sup>. Recently, a correlation with the development of motor neuron pathology in amyotrophic lateral sclerosis (ALS) was found by Buchman and co-workers. For one third of 31 ALS linked human tissue samples analysed, aggregated  $\gamma$ Syn was found within the descending axons of the corticospinal tract in the dorsolateral column of the spinal cord<sup>375</sup>. Further, overexpression of mouse  $\gamma$ Syn in neurons of transgenic mice supported the hypothesis of this protein being involved in neuron pathology as the animals developed middle-age onset progressive motor neuron pathology that recapitulated many key characteristics of ALS, and died prematurely<sup>376,377</sup>. The exact role of  $\gamma$ Syn for neuronal diseases however remains unclear and usually the aggregation of the amyloid protein TDP-43 is predominantly associated with ALS<sup>174</sup>.

### 1.5.1.2 Structural properties

As mentioned above,  $\gamma$ Syn has been investigated in much less detail compared to  $\alpha$ Syn. However, there is some structural analysis, mainly focussing on the monomeric protein. In consensus with  $\alpha$ Syn,  $\gamma$ Syn can be separated into three regions (N-terminal, NAC, and C-terminal region) and shows a repetitive motif in the N-terminal region of the protein (**Figure 1.27**). Due to  $\gamma$ Syn lacking 13 residues in the C-terminal region it presents a significantly less charged C-terminus with a net charge of -5 compared to -12 for  $\alpha$ Syn at neutral pH<sup>378</sup>. This results in disrupted intramolecular interactions between N- and C-terminal region<sup>333</sup>. Also,  $\gamma$ Syn has been found to have an increased  $\alpha$ -helical propensity in the NAC region possibly stabilising the protein and interrupting fibril assembly processes<sup>366</sup>.

$\gamma$ Syn has further been described to bind membranes and adopting an  $\alpha$ -helical structure in its N-terminal region and NAC during this process<sup>379</sup>. Similar to  $\alpha$ Syn, it shows preferred binding to anionic lipids and curved membranes<sup>380</sup>. There is no evidence for membrane clustering though, which might be attributed to the lack of such studies or sequence changes in the NAC region (residues 65-97) required for the 'double-anchor-mechanism'<sup>272</sup>.

$\gamma$ Syn has been demonstrated to have a higher propensity to form oligomeric species compared to  $\alpha$ Syn *in vitro*<sup>365</sup>. Nevertheless, fibril formation is significantly slower than

$\alpha$ Syn<sup>365,381</sup>. This hints that oligomers formed by  $\gamma$ Syn might be off-pathway and are not elongated to mature fibril structures efficiently. In addition to this incubation of  $\gamma$ Syn oligomers with binding partners such as membranes has been shown to lead to disruption of the oligomers into monomeric protein<sup>382</sup>.

To date, there is no high resolution structural information for  $\gamma$ Syn fibrils.

## 1.6 Thesis aim

The introduction of this thesis clearly demonstrates that, although there has been massive progress over the past decade(s) in understanding the role of amyloids and synucleins specifically for function and disease, the molecular mechanisms are often not clarified in detail. The work in this thesis aims to understand the early processes of synuclein aggregation on a molecular level. Focussing on the amino acid composition and their role for function and in particular amyloid assembly could help to shine light on the mechanisms underlying disease development. To gain insight into the aggregation processes, *in vitro* and *in vivo*, a wide variety of techniques was used including ThT fluorescence aggregation assay (*de novo* growth, seeded reactions, and surface catalysed aggregation), electron microscopy, NMR PRE experiments and toxicity assays in the model organism *C. elegans*.

The objectives of each individual chapter will be briefly summarised below:

### **Chapter 3: A motif in the N-terminal region crucial for aggregation and function of $\alpha$ -synuclein**

This chapter uses the primary amino acid sequence of  $\alpha$ Syn to find aggregation hotspots especially prone to aggregation via *in silico* techniques. In addition to NAC, also three regions in the N-terminal region were identified that are analysed in more detail in this work by using deletion variants. Aggregation assays confirm that the identified regions indeed are crucial for fibril formation *in vitro* and *in vivo* and functional assay with liposomes further highlight their role in membrane remodelling. NMR PRE experiment allow the interpretation of transient long-range intra- and inter-molecular interactions involving these regions to be important for driving aggregation of  $\alpha$ Syn.

### **Chapter 4: Single residue switches in the N-terminal P1-region change aggregation of $\alpha$ -synuclein**

Following up on the work in Chapter 3, the role of each individual residue of one of the identified regions which showed the most drastic effect (P1: residues 36-42) was analysed in large detail. An alanine scan and sequence alignments with  $\gamma$ Syn were used to find the residues that contribute most to the aggregation process within this region pointing to residue L38, Y39 and S42 that are key factors for fibril formation. Substituting

these residues with other aliphatic side chains demonstrated the power of individual residues to control aggregation with a high sequence specificity.

### **Chapter 5: An ALS-linked sequence variant of $\gamma$ Syn in P1 induces aggregation**

Chapter 5 focusses on the role of  $\gamma$ Syn. In collaboration with Vladimir Buchman, University of Cardiff, an ALS linked variant in  $\gamma$ Syn was identified (M38I). In this work the biophysical characteristics of this substitution were evaluated *in vitro* and its aggregation potential analysed using cell cultures and *C. elegans*. The identity of this disease associated variant further supports the crucial role of P1 in synucleins and the contribution of individual residues for fibril formation.

Taken together, this study focusses on the role of specific N-terminal regions of  $\alpha$ Syn for aggregation and function. It highlights the important involvement of regions flanking the central aggregation hotspots (e.g. NAC for  $\alpha$ Syn) for function and toxicity in amyloid proteins. This will be further discussed in Chapter 6 – Concluding Remarks and Future Perspective.





## Chapter 2

# Materials and Methods

## 2.1 Chemicals

All chemicals were purchased from Sigma-Aldrich (Poole, Dorset, UK) and were of analytical grade unless specified otherwise.  $^{15}\text{N}$  labelled ammonium chloride and  $^{13}\text{C}$  labelled glucose for isotopic labelling of  $\alpha\text{Syn}$  for NMR experiments were purchased from Cambridge Isotope Laboratories (Massachusetts, USA). S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl-oxyl)methyl methanesulfonothioate (MTSL) was purchased from Cambridge Isotope Laboratories (Massachusetts, USA). 1,2-Dimyristoyl-sn-glycero-3-phosphoserine (DMPS) lipids were purchased from Avanti® Polar Lipids and peptides were ordered from Severn Biotech Ltd., Peptide Synthesis or GeneScript.

## 2.2 Bacterial strains

In this study, the following *Escherichia coli* (*E. coli*) strains were used:

**DH5 $\alpha$**  (Invitrogen), Genotype: *E. coli* F-  $\phi$ 80lacZ $\Delta$ M15  $\Delta$ (lacZYA-argF)U169 recA1 endA1 hsdR17(rk-, mk+) phoA supE44 thi-1 gyrA96 relA1  $\lambda$ -

**BL21 (DE3)**, Genotype: *E. coli* B F<sup>-</sup> dcm ompT hsdS(rB<sup>-</sup> mB<sup>-</sup>) gal  $\lambda$ (DE3)

**HT115 (DE3)**, Genotype: *E. coli* F-, mcrA, mcrB, IN(rrnD-rrnE)1, rnc14::Tn10(DE3 lysogen: lacUV5 promoter -T7 polymerase) (IPTG-inducible T7 polymerase) (RNase III minus).

DH5 $\alpha$  cells were used for molecular biology experiments and BL21 (DE3) cells were used for protein expression. HT115 (DE3) were used for *C. elegans* maintenance.

## 2.3 Vectors

For the recombinant expression of  $\alpha$ - and  $\gamma\text{Syn}$  in *E. coli*, the pET23a plasmid (expression under control of T7 promoter, ampicillin resistance) was used. The  $\alpha\text{Syn}$  encoding plasmid was kindly provided by Prof Jean Baum (Department of Chemistry and

Chemical Biology, Rutgers University, NJ, USA).  $\gamma$ Syn was purchased from Eurofins and cloned into the pET23a vector by Matthew Jackson. Genes for chimeric proteins ( $\alpha$ - and  $\gamma$ Syn mixed constructs) were purchased from Twist Bioscience in a pET29b plasmid (expression under control of T7 promoter, kanamycin resistance).

For synuclein expression in *C. elegans*, the  $\alpha$ Syn wildtype gene in vector pPD30.38 was used and provided as a generous gift from Prof Ellen Nollen, UMCG, Groningen, The Netherlands. The codon optimised (for *C. elegans*)  $\gamma$ Syn construct was ordered from Eurofins and cloned into the pPD30.38 vector. Variants of interest in this study were generated by Q5 site-directed mutagenesis (Section 2.5.1).

## 2.4 Peptides

Peptides were used to study the aggregation kinetics when peptides are added *in trans* and self-assembly of the NAC peptide in isolation. All peptides used in this thesis are listed in **Table 2.1**.

**Table 2.1: Sequence of peptides used in this study and their producers.** Peptides were generated at >98 % purity. NAC peptides were synthesised by Martin Walko, University of Leeds.

<b>P1-peptide</b>	Acetyl-KTKEGVLYVGSKTKE-amide	Severn Biotech
<b>P1-Cys-peptide</b>	Acetyl-CKTKEGVLYVGSKTKE-amide	Peptide Synthesis
<b>P1-SG-peptide</b>	Acetyl-KTKESGSGSGSKTKE-amide	Severn Biotech
<b>N7-peptide</b>	NH <sub>2</sub> -MDVFMKGKTKE-amide	GeneScript
<b><math>\alpha</math>NAC-peptide</b>	NH <sub>2</sub> -EQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFV-amide	Synthesised by M.W.
<b><math>\gamma</math>NAC-peptide</b>	NH <sub>2</sub> -EQANAVSEAVVSSVNTVATKTVEEAENIAVTSGVV-amide	Synthesised by M.W.

## 2.5 Molecular Biology

### 2.5.1 Site directed mutagenesis

Various substitution and deletion variants of  $\alpha$ - and  $\gamma$ Syn were created in this project using the NEB Q5 site-directed mutagenesis kit. Primers were designed using the NEB online tool (<http://nebasechanger.neb.com/>) and further reviewed for potential dimer formation with the Multiple Primer Analyzer (<https://www.thermofisher.com/>). Primers were supplied by Eurofins MWG Operon and are listed in **Table 2.2**. The Q5 mutagenesis was performed by a polymerase chain reaction (PCR) according to the manufacturer's instructions (see Section 2.5.1.1) followed by a kinase, ligase, Dnpl (KLD) treatment Section 2.5.1.2.

Table 2.2: Oligonucleotides used in this study.

Construct	Primer sequence
<b>αSyn A18C</b>	Forward: 5' TGTCGTAGCA <sub>tgt</sub> GCGGAAAAGAC 3' Reverse: 5' CCTTCTTTCGCTTTCGAC 3'
<b>αSyn A90C</b>	Forward: 5' CTCCATTGCA <sub>tc</sub> GCGACTGGCTTTGTGAAG 3' Reverse: 5' CCAGCACCTCCACCGTT 3'
<b>αSyn A140C</b>	Forward: 5' CGAACCCAGAA <sub>tgt</sub> TAAAAGCTTGCGGCCG 3' Reverse: 5' TAGTCCTGGTAGCCTTCT 3'
<b>αSyn S129C</b>	Forward: 5' TGAGATGCCG <sub>tgt</sub> GAAGAAGGCT 3' Reverse: 5' TAGGCCTCATTATCCGGATC 3'
<b>αSyn V40C</b>	Forward: 5' TGTGCTGTAT <sub>tg</sub> CGGCAGCAAG 3' Reverse: 5' CCTTCCTTCGTTTTACCG 3'
<b>αSyn ΔN7</b>	Forward: 5' CTGTCGAAAAGCGAAAGAAG 3' Reverse: 5' CATATGATATCTCCTTCTTAAAG 3'
<b>αSyn ΔC1</b>	Forward: 5' AAGACCAAACAAGGCGTAG 3' Reverse: 5' TTCTTTCGCTTTCGACAG 3'
<b>αSyn ΔP1</b>	Forward: 5' AAGACCAAAGAAGGCGTTG 3' Reverse: 5' TTCCTTCGTTTTACCGGC 3'
<b>αSyn ΔP2</b>	Forward: 5' AAAACGAAAGAGCAGGTGAC 3' Reverse: 5' GGTCTTGCTGCCGACATA 3'
<b>αSyn ΔP1ΔP2</b>	Forward: 5' AAAACGAAAGAGCAGGTG 3' Reverse: 5' GGTCTTTCCTTCGTTTTAC 3'
<b>αSyn P1-GS</b>	Forward: 5' cggtagcGGCAGCAAGACCAAAGAAG 3' Reverse: 5' ctaccgctTTCCTTCGTTTTACCGGC 3'
<b>αSyn P1P2-GS</b>	Forward: 5' atcaggatcaggatcaggaAAAACGAAAGAGCAGGTGACAAAC 3' Reverse: 5' ccactcccactcccactcccGGTCTTGCTGCCGCTACC 3'
<b>αSyn G36A</b>	Forward: 5' AACGAAGGAA <sub>gcg</sub> GTGCTGTATGTCGGC 3' Reverse: 5' TTACCGGCTGCTTCGGCT 3'
<b>αSyn V37A</b>	Forward: 5' GAAGGAAGGT <sub>gcg</sub> CTGTATGTCG 3' Reverse: 5' GTTTTACCGGCTGCTTCG 3'
<b>αSyn L38A</b>	Forward: 5' GGAAGGTGT <sub>gcg</sub> TTATGTCGGAG 3' Reverse: 5' TTCGTTTTACCGGCTGCT 3'
<b>αSyn L38M</b>	Forward: 5' GGAAGGTGT <sub>gatg</sub> TATGTCGGCA 3' Reverse: 5' TTCGTTTTACCGGCTGCTC 3'
<b>αSyn L38I</b>	Forward: 5' GGAAGGTGT <sub>gatt</sub> TATGTCGGCAG 3' Reverse: 5' TTCGTTTTACCGGCTGCT 3'
<b>αSyn Y39A</b>	Forward: 5' AGGTGTGCT <sub>gcg</sub> GTCGGCAGCA 3' Reverse: 5' TCTTCGTTTTACCGGCT 3'
<b>αSyn V40A</b>	Forward: 5' TGTGCTGTAT <sub>gcg</sub> GGCAGCAAGAC 3' Reverse: 5' CCTTCCTTCGTTTTACCG 3'
<b>αSyn G41A</b>	Forward: 5' GCTGTATGTC <sub>gcg</sub> AGCAAGACCAAAG 3' Reverse: 5' ACACTTCCTTCGTTTTAC 3'

<b>αSyn S42A</b>	Forward: 5' GTATGTCGGCgcgAAGACCAAAGAAGGC 3' Reverse: 5' AGCACACCTTCCTTCGTT 3'
<b>γSyn M38I</b>	Forward: 5' AAGGCGTCATtTATGTGGGTG 3' Reverse: 5' CTTTGGTCTTTTCGGCTG 3'
<b>γSyn M38L</b>	Forward: 5' AGAAGGCGTCctgTATGTGGGTG 3' Reverse: 5' TTGGTCTTTTCGGCTGCC 5'
<b>γSyn M38A</b>	Forward: 5' AGAAGGCGTCgcgTATGTGGGTGC 3' Reverse: 5' TTGGTCTTTTCGGCTGCC 3'
<b>γSyn M38V</b>	Forward: 5' AGAAGGCGTCctgTATGTGGGTG 3' Reverse: 5' TTGGTCTTTTCGGCTGCC 3'
<b>γSyn A42S</b>	Forward: 5' GTATGTGGGTagcAAAACGAAAGAGAACGTTG 3' Reverse: 5' ATGACGCCTTCTTTGGTC 3'
<b>γSyn E110V</b>	Forward: 5' CAGGAAGGCGtgGCCTCCAAGGAG 3' Reverse: 5' TTGCGGTGCTGATGGACG 3'

### 2.5.1.1 Polymerase chain reaction

For the PCR, the Q5 hot start high fidelity DNA polymerase was used with the template DNA and the mutagenic primers in the following reaction:

Q5 enzyme mix	1x
Forward primer	0.5 μM
Reverse primer	0.5 μM
Template DNA	25 ng
Water	To 25 μL

The reaction mixture was transferred to a thermocycler and the PCR cycling conditions outlined in **Table 2.3** were performed

**Table 2.3: Temperature cycle for PCR for site directed mutagenesis.**

Step	Temperature (°C)	Time (s)
Initial denaturation	98	30
Denaturation	98	10
Annealing	$T_a$	30
Elongation	72	30 per kb
Repeat denaturation, annealing, and elongation for 25 cycles		
Final extension	72	120

The annealing temperature ( $T_a$ ) was calculated using **Equation 2.1**:

#### Equation 2.1

$$T_a = (nAT \times 2) + (nGC \times 4)$$

### 2.5.1.2 Kinase, ligase and DpnI (KLD) treatment

Following the PCR, the amplified product is subject to treatment with kinase, ligase and DpnI (KLD) enzymes to allow efficient phosphorylation, intramolecular ligation of the plasmid DNA and to remove template DNA respectively. The KLD reaction outlined below was assembled and incubated at 25 °C for 5 min. Following incubation, 5 µL of the reaction was transformed into DH5α competent cells (Section 2.5.3).

Q5 PCR product	1 µL
KLD reaction buffer	5 µL
KLD enzyme mix	1 µL
Water	3 µL

### 2.5.2 Subcloning

For the *C. elegans* codon optimised ySyn gene, subcloning was performed to insert the gene in the pPD30.38 vector. The gene was PCR amplified with additional flanking sites including the NheI and AgeI restriction site. Vector and insert were cut by the enzymes, gel purified and ligated in a mixture 1:3 (M:M) using the T4 DNA ligase.

### 2.5.3 Transformation

The generated plasmids were transformed in *E. coli* DH5α cells for sequencing experiments or in BL21 (DE3) cells for protein expression. Competent *E. coli* cells were incubated with 1 µg plasmid for 30 min on ice before heat shocking them at 42 °C for 45 s. After resting on ice for another 5 min, cells were plated on a 100 µg/mL carbenicillin containing LB agar plate. For DH5α cells, additionally 950 µL SOC media was added after the heat shock and the mixture incubated 1 h at 37 °C, 200 rpm before plating. Colonies grown over night at 37°C were further used.

### 2.5.4 Sequencing

5 mL 100 µg/mL containing LB overnight cultures were inoculated from a single colony obtained from the transformation in DH5α cells and incubated at 37 °C and 200 rpm. The next day, the plasmid DNA was isolated using the Wizard® Plus SV Minipreps DNA Purification System kit. DNA concentrations were determined using the nanodrop 1000 spectrophotometer and 15 µL sample (~100 ng/µL) was sent to Genewiz or Eurofins for sequencing to verify successful mutagenesis.

## 2.6 Protein expression and purification

### 2.6.1 $\alpha$ Syn

#### 2.6.1.1 Expression

Pre-cultures were grown in 5 mL LB medium with 100  $\mu$ g/mL carbenicillin inoculated with a single colony of the BL21 (DE3) cells containing the plasmid of interest. The next day, 500  $\mu$ L culture was mixed with the same volume of sterile 30 % (v/v) glycerol and stored at -80 °C until usage.

For the overnight cultures 100 mL LB medium with 100  $\mu$ g/mL carbenicillin was inoculated with 100  $\mu$ L of the glycerol stock. Cultures were incubated at 37 °C and 200 rpm shaking overnight. The next day, 10 mL of the overnight cultures were used to inoculate 5 to 10 flasks with 1 L  $^{14}$ N LB medium. The cultures were grown at 37 °C and shaking with 200 rpm until an OD<sub>600</sub> of 0.6 was reached. Protein expression was then induced by adding 1 mM isopropyl- $\beta$ -D-thio-galactopyranoside (IPTG). The cultures were grown for an additional 4 to 5 hours before the cells were harvested.

For expression of  $^{15}$ N (and  $^{13}$ C) labelled protein, the protocol above was followed but using only 0.5 L medium per flask and expression was performed in  $^{15}$ N/ $^{13}$ C HCDMI medium (7.5 g Na<sub>2</sub>HPO<sub>4</sub>, 10 g K<sub>2</sub>PO<sub>4</sub>, 9 g K<sub>2</sub>SO<sub>4</sub>, 10 g KH<sub>2</sub>PO<sub>4</sub>, 1 g NH<sub>4</sub>Cl ( $^{15}$ N labelled in 1L)) supplemented with 1 mM MgCl<sub>2</sub>, 100  $\mu$ M CaCl<sub>2</sub>, 0.8 % (w/v) glucose ( $^{13}$ C labelled if needed), and 100  $\mu$ g/mL carbenicillin. Expression was induced by adding IPTG when an OD<sub>600</sub> of 0.6 was reached and cells were harvested after 5 hours of protein production.

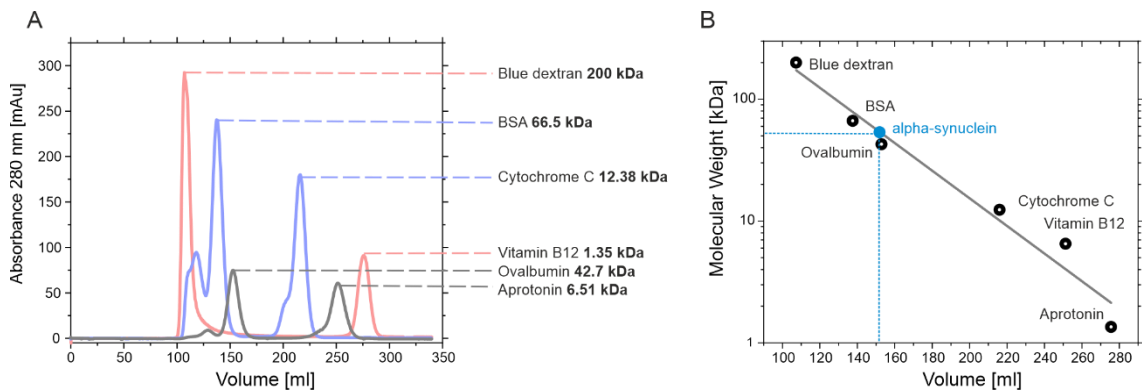
#### 2.6.1.2 Purification

Cell harvesting was performed by centrifugation at 6000 rpm for 20 min (rotor JA 8.1) at 4 °C. The cell pellet was resuspended and homogenized in 15 mL/litre culture lysis buffer and incubated for 30 min on a roller to disrupt the cells. Samples were then heated to 80 °C for 10 min and centrifuged for 30 min at 35000 xg. The supernatant was further used for an ammonium sulfate (AS) precipitation. 29.1 g AS per 100 mL protein solution was added and incubated 30 min at 4 °C. The precipitated protein was pelleted by centrifuging for 30 min at 35000 x g. Then, the pellet was washed with 50 % (w/v) AS in 50 mL water and centrifuged another time (35,000 x g, 4 °C, 30 min).

Finally, the protein pellet was resuspended in 300-500 mL wash buffer (20 mM Tris-HCl, pH 8.0) and loaded on a self-made 300 mL anion exchange column (Q-sepharose resin). The protein was eluted over a 0 – 500 mM NaCl gradient over a volume of 500-1000 mL.  $\alpha$ Syn containing fractions were combined and dialysed against 5 L of 50 mM ammonium bicarbonate pH 8, the buffer was exchanged at least three times and incubated at 4 °C.

The partially purified  $\alpha$ Syn was lyophilised and stored at  $-20\text{ }^{\circ}\text{C}$  until further purification by size exclusion chromatography (SEC). For the size exclusion an HiLoad<sup>TM</sup> 26/60 Superdex 75 prep grade gel filtration column was used with a flow rate of 2.6 mL/min. Filtered ( $0.22\text{ }\mu\text{m}$ )  $\alpha$ Syn was loaded and eluted in 50 mM ammonium bicarbonate pH 8, lyophilized and stored at  $-20\text{ }^{\circ}\text{C}$  until usage.

Molecular weight calibrations for the used SEC column are shown below in **Figure 2.1**.



**Figure 2.1: Molecular weight calibrants for size exclusion chromatography. (A)** SEC profiles for molecular weights calibrants blue dextran (200 kDa), BSA (66.5 kDa), ovalbumin (42.7 kDa), cytochrome C (12.38 kDa), Aprotonin (6.51 kDa), vitamin B12 (1.35 kDa). **(B)** Molecular weight (log-scale) of calibrants plotted against their elution volume. The elution volume of  $\alpha$ Syn is highlighted in blue, the mass seems bigger than it is due to its conformational state. The data were kindly provided by Nicolas Guthertz, University of Leeds.

The presence of the protein of interest was followed by running a sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) through all purification steps (Section 2.7.1). The identity of the final protein solution was confirmed by electrospray ionisation mass spectrometry (ESI-MS) performed by the Mass Spec facility.

### 2.6.2 $\gamma$ Syn

$\gamma$ Syn protein expression and purification was performed similarly to the methods described for the  $\alpha$ Syn protocol (Section 2.6.1). To lyse the cells, the cell suspension was additionally cell disrupted at 40 KPSI.

### 2.6.3 Chimeric proteins

The chimeric proteins were expressed and purified as described above (Section 2.6.1). Due to the use of a different plasmid (Section 2.3), 50  $\mu\text{g}/\text{mL}$  kanamycin instead of carbenicillin was added to LB plates and media.

## 2.7 General biochemical techniques

### 2.7.1 SDS-PAGE

To determine the size and purity of a protein, a sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed. It separates proteins based on their molecular weight. Stacking and resolving gels (**Table 2.4**) were freshly prepared and poured in a sealed clean casting chamber. Before running the gel, rubber seals were removed and the gel was placed in an electrophoresis cell. The cathode chamber was filled with cathode buffer (100 mM Tris, 100 mM Tricine, 0.1 % (w/v) SDS) and the anion chamber with anion buffer (200 mM Tris-HCl pH 8.9). For Western blots, pre-poured 4-20 % (v/v) gradient gels from BioRad were used.

**Table 2.4: SDS-PAGE gel compositions.**

Solution component	Resolving gel [mL]	Stacking gel [mL]
30 % (w/v) Acrylamide	7.5	0.83
3M Tris HCl, 0.3 % SDS (w/v), pH 8	5.0	1.55
H <sub>2</sub> O	0.48	3.72
Glycerol	2.0	-
10 % (w/v) ammonium persulfate	0.2	0.2
Tetramethylethylenediamine	0.02	0.02

Protein samples were diluted 1:1 with SDS-PAGE loading buffer (2x stock : 2 % (w/v) SDS, 10 % (v/v) glycerol, 0.1 % (w/v) bromophenol blue, 100 mM DTT) and boiled for 10 min before centrifuging for 1 min at 13000 x g. 15 µL of each sample was loaded on the gel, 5 µL Precision plus protein dual colour standards protein ladder (BioRad) was loaded as a reference. The gel was run constantly at 30 mA until the resolving gel was reached, then the electric current was increased to 60 mA. The electric potential was limited to 200 V. Finally, the gel was visualised by incubating it with Quickblue Coomassie stain and imaged at the imaging machine UVITEC, Cambridge.

### 2.7.2 Agarose gel electrophoresis

Agarose gels were used to separate DNA fragments. For this, the gel was prepared by dissolving 1 % (w/v) agarose in 100 mL Tris-acetate-EDTA (TAE) buffer (40 mM Tris-HCl, pH 8, 20 mM acetic acid (glacial), 1 mM EDTA, pH 7.5). A final concentration of 0.5 µg/mL of ethidium bromide was added prior to pouring the gel into a 15 cm gel tray. Once the gel had set, it was transferred to the TAE buffer filled electrophoresis unit. DNA samples were diluted in 6x Purple gel loading dye before loading to the gel along with 5 µL of 1 kb and 100 bp DNA ladders (New England Biolabs) to allow size determination. The electrophoresis was performed at 100 V until the DNA fragments were suitably



resolved. Gels were visualised using ultraviolet (UV) transillumination and imaged using Alliance Q9 Advanced gel doc system.

### 2.7.3 Immunoblotting analysis

Immunoblotting was used to visualise the expression level of the synuclein variants in *C. elegans*. Nematodes were collected from agar plates, washed in M9 buffer (5.8g Na<sub>2</sub>HPO<sub>4</sub>-7H<sub>2</sub>O, 3g KH<sub>2</sub>PO<sub>4</sub>, 5g NaCl, 1mM MgSO<sub>4</sub> in 1 L H<sub>2</sub>O), and resuspended in worm lysis buffer (20 mM Tris, pH 7.5; 10 mM β-mercaptoethanol; 0.5% (v/v) Triton X-100; supplemented with complete protease inhibitor (Roche)). After shock freezing in liquid nitrogen, three freeze-thaw cycles were performed before the worm pellet was grinded with a motorized pestle. After lysis on ice, the sample was centrifuged for 1 min at 1000 rpm to pellet the carcasses. The protein concentration was determined using the Bradford assay (Bio-Rad).

Worm samples were mixed in a 1:1 ratio with SDS loading buffer and boiled for 10 min. A final protein amount of 25 µg was loaded onto a 4-20% Tris-HCl gel (Bio-Rad). After separating the proteins by SDS-PAGE they were blotted onto a PVDF membrane (BioRad), αSyn, γSyn, and tubulin were visualised using syn211 (1:5000) (Sigma), anti-GFP (1:1000) (ThermoFisher) or a monoclonal mouse anti-tubulin antibody (1:5000) (Sigma), respectively, followed by an anti-mouse horse-radish peroxidase-coupled secondary antibody (1:5000). The bands were visualised using the SuperSignal West Pico Plus Chemiluminescence Substrate (Thermo).

### 2.7.4 Preparation of disulfide locked dimeric α-synuclein species

Disulfide locked αSyn species were generated to analyse aggregation kinetics of αSyn dimers. To guarantee the formation of disulfate bonds, monomeric cysteine substitution variants were generated, expressed and purified and then incubated at a concentration of 400 µM in 100 mM Tris HCl, pH 8.4, for 2 h at room temperature. The protein samples were run on a HiLoad™ 26/60 Superdex 75 preparative grade gel filtration column (GE Healthcare) in 50 mM ammonium bicarbonate, pH 8.0 to separate monomers and dimers. Fractions containing the dimerised αSyn were collected, lyophilised and stored at -20 °C until usage. This experiment was performed by Ciaran Doherty.

### 2.7.5 SEC MALS

SEC-MALS (multiple angle light scattering) analysis was used to estimate the mass of the dimeric αSyn constructs. A TOSOH G200SWXL column was equilibrated with 20 mM Tris-HCl, 200 mM NaCl, pH 7.5 and 50 µL of 30 µM αSyn sample was loaded onto the column. The eluted fractions were run through a Wyatt miniDawnTreos system with three-angle detection, and the data analysed using the Astra 6.0.3 software that was

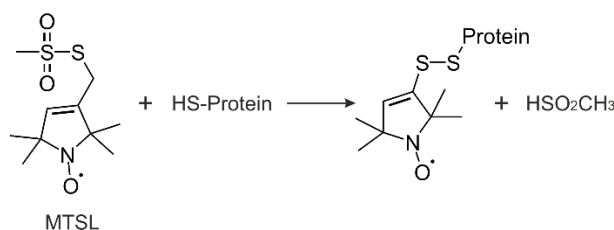
supplied with the instrument. This experiment was performed by Ciaran Doherty and Leon Willis, University of Leeds.

### 2.7.6 Dynamic Light Scattering

Dynamic light scattering (DLS) was used to assess the size of the prepared liposomes (Section 2.7.8). For this, 250  $\mu\text{L}$  of 100  $\mu\text{M}$  samples (lipid concentrations) were injected into a Wyatt miniDawnTreos system (equipped with an additional DLS detector). The data were analysed using the Astra 6.0.3 software supplied with the instrument. Filtered (0.22  $\mu\text{m}$ ) and de-gassed buffer, kept cool on ice to minimise bubble formation inside the instrument, was used to obtain 5-min baselines before and after sample injection. A 3 min sample window was used for the analysis by the software. Using this analysis, the liposomes were found to have a diameter of  $\sim 160$  nm.

### 2.7.7 MTSL-labelling

MTSL (S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl-oxy)methyl) labelling at cysteine residues of synucleins allows the detection of transient interactions in proteins by paramagnetic relaxation enhancement (PRE) experiments. Due to the absence of naturally occurring cysteines, substitution variants of  $\alpha\text{Syn}$  allowed a selective labelling at positions of interest with the introduced cysteines<sup>383</sup>. The labelling reaction is illustrated in **Figure 2.2**.



**Figure 2.2: MTSL labelling reaction.**

For the labelling reaction, 20 mg lyophilised  $^{15}\text{N}$  labelled or  $^{14}\text{N}$   $\alpha\text{Syn}$  with an integrated cysteine was dissolved in 8 mL 20 mM Tris-HCl buffer, 200 mM NaCl, pH 7.5. A final concentration of 5 mM DTT was added to the protein solution and incubated for  $\geq 30$  min at room temperature to reduce all cysteine residues. DTT was then removed by using a Zeba Spin Desalting Column (Thermo scientific) where only the protein was eluted into an empty falcon tube already containing 40 molar excess MTSL. The protein sample with MTSL was incubated at 4  $^{\circ}\text{C}$  overnight. To remove the excess MTSL the protein was again run through a Zeba Spin Desalting Column (Thermo scientific).  $\alpha\text{Syn}$  was eluted in 20 mM Tris-HCl buffer pH 4.5 with low (20mM) or high (200 mM) NaCl concentration. Successful labelling was confirmed by ESI-MS. If required, proteins were concentrated using a Vivaspin 20 (Sartorius) prior to the PRE NMR experiment.

For MTSL labelling of the P1-Cys peptide (N-terminal cysteine for labelling reaction), 3 mg peptide was incubated for 30 min in the presence of 5 mM DTT in 20 mM Tris HCl, 200 mM NaCl, pH 7.5. After removing access DTT using a Zeba spin column (PD10 column, GE Healthcare), the peptide sample was immediately labelled with a 40-fold molar excess of MTSL for 8 h at 25 °C in 20 mM Tris-HCl, 200 mM NaCl, pH 7.5. Excess MTSL and labelled peptide were separated by HPLC and complete modification of the peptide with MTSL confirmed by ESI-MS (performed by Yong Xu, University of Leeds).

### 2.7.8 Liposome Preparation

Liposomes were used to analyse the physiological function of synuclein of membrane binding and remodelling vesicles in more detail. 1,2-dimyristoyl-sn-glycero-3-phospho-L-serine (sodium salt DMPS, Avanti Polar Lipids) lipids were dissolved in 20 mM sodium phosphate buffer, pH 6.5 and stirred for 2h at 45 °C. The lipid solution was then frozen in dry ice and thawed in a 45 °C warm water bath 5-times. The preparation of liposomes was then finalised by sonication in a bath sonicator (U50 ultrasonic bath, Ultrawave) for 1 h. The size distribution of liposomes was determined using DLS (Section 2.7.6). Liposomes were used within two days storing at 4 °C if necessary.

### 2.7.9 Peptide synthesis

NAC peptides (residues 61-95 of  $\alpha$ Syn and  $\gamma$ Syn) were synthesized on CEM Liberty Blue peptide synthesizer with microwave assistance using default coupling cycles for the first 15 residues and double couplings for the rest of the peptide. The synthesis was performed on 0.1 mmol scale using Valine preloaded HMPB-NovaPeg resin (0.56 mmol/g), DMF as a solvent, 20% piperidine in DMF for the deprotection and DIC and OXYMA pure for couplings. Cleavage from resin was accomplished using TFA:H<sub>2</sub>O:TIS:EDT, 92.5:2.5:2.5:2.5 (5mL x 3 h) and peptides were precipitated using cold ether. Pure peptides were obtained after preparative HPLC purification on Kinetex EVO 5  $\mu$ m C18 100Å 21.2 x 250 mm reverse phase column using a 10 - 30% acetonitrile gradient with 0.1% ammonia followed by lyophilisation. The synthesis was kindly performed by Martin Walko, University of Leeds.

### 2.7.10 Limited proteolysis

Limited proteolysis with proteinase K was used to gain information on the core-stability of fibrils formed from different variants. The predominant cleavage site of Proteinase K is the peptide bond adjacent to the carboxyl group of aliphatic and aromatic amino acids. Pre-formed fibrils were spun down in a table centrifuge for 5 min at maximum speed. The pellet was dissolved in 20 mM Tris-HCl, 200 mM NaCl, pH 7.5 to a final concentration of 50  $\mu$ M monomer equivalent. Proteinase K was added in a 1:100 ratio (enzyme:protein) and aliquots at various time points were taken. The digestion reaction was stopped by

boiling the samples for 15 min in the presence of SDS-loading buffer before running on a SDS-PAGE gel. Different digestion patterns between different fibrils are an indicator of distinct fibril architectures.

## 2.8 Aggregation assays

### 2.8.1 Seed preparation

Fibril growth *de novo* or in the presence of liposomes, peptides or pre-formed fibril seeds was characterised. To generate synuclein seeds, 500  $\mu\text{L}$  of a 600  $\mu\text{M}$   $\alpha\text{Syn}$  solution in Tris-HCl, pH 7.5, 20 mM or 200 mM NaCl was stirred with a magnetic stirrer at 1200 rpm for 48 h at 45 °C. Alternatively, 500  $\mu\text{L}$  of a 600  $\mu\text{M}$   $\gamma\text{Syn}$  M38I solution in sodium acetate, pH 4.5, 200 mM NaCl was used. The solutions were aliquoted and stored at -20 °C. Before usage, the pre-formed fibrils were sonicated twice for 30 s with a break of 30 s at 40% maximum power using a Cole-Parmer-Ultraprocessor-Sonicator and then added to the monomeric protein (10 % seeds (mol:mol)) to measure the synuclein aggregation.

### 2.8.2 Thioflavin T assay

Thioflavin T (ThT) assays were performed to determine the aggregation kinetics of synuclein variants *in vitro*. When cross  $\beta$ -sheet rich fibrils form, ThT binds to these highly ordered structures, which changes its fluorescence characteristics, which can be detected at an emission wavelength of 482 nm<sup>137</sup>. *De novo* fibril growth, elongation (seeding) and surface induced aggregation were analysed using this technique. All graphs in this figure show data from one biological repeat performed in at least duplicates.

#### 2.8.2.1 *De novo* aggregation

To measure *de novo* fibril growth, a final concentration of 100  $\mu\text{M}$  filtered synuclein was mixed in a sealed 96-well non-binding surface flat bottom assay plate (Corning) with 20  $\mu\text{M}$  thioflavin T (ThT) in a final volume of 100  $\mu\text{L}$ . Different buffer conditions were tested including 20 mM Tris pH 7.5 at high (200mM) and low (20mM) salt concentration and 20 mM acetate buffer pH 4.5 at high (200mM) and low (20mM) salt concentration. The samples were incubated shaking at 600 rpm at 37 °C for 100 h to 250 h in a BMG Labtech FLUOstar optima plate reader. During the time of the experiment, the fluorescence was detected by exciting ThT at 440 nm and monitoring the emission at 480 nm with a gain set to 600. At the end of a ThT assay the presence of fibrils was confirmed by transmission electron microscopy (TEM) (Section 2.12) and fibril yields were determined.

### 2.8.2.2 Seeded aggregation

Seeding experiments were performed to analyse the ability of synuclein variants to elongate  $\alpha$ Syn WT or  $\gamma$ Syn M38I fibrils. By adding seeds, the nucleation step is eliminated and aggregation is driven by elongation processes. The experiment was performed as described above (Section 2.8.2.1). In addition to the 100  $\mu$ M monomeric protein, 10 % (mol/mol) seeds were supplemented. The aggregation was followed at 37 °C under quiescent conditions at neutral or acidic pH for 42 h.

### 2.8.2.3 Surface induced aggregation

Aggregation of synuclein on the surface of liposomes was measured. For this, the protocol described by Galvagnion *et al.* was followed<sup>351</sup>. In brief, a protein solution with a concentration of 50  $\mu$ M in 20 mM sodium phosphate buffer, pH 6.5 was mixed with varying amounts of liposomes (DMPS, Avanti) (lipid concentrations: 0; 0.4 and 3 mM). Samples were incubated at 30 °C in a 96 well plate under quiescent conditions. After finishing the aggregation assay, samples were imaged by TEM to observe whether synuclein induced liposome merging had occurred.

### 2.8.2.4 Aggregation with peptides

Aggregation assays in the presence of peptides (**Table 2.1**) were measured similar as described above. 100  $\mu$ M (1:1) or 1 mM (1:10) peptide (dissolved in 20 mM Tris, pH 7.5, 200 mM NaCl) were added to 100  $\mu$ M monomeric protein. As a control, the aggregation of peptide alone was followed and validated by CD and TEM.

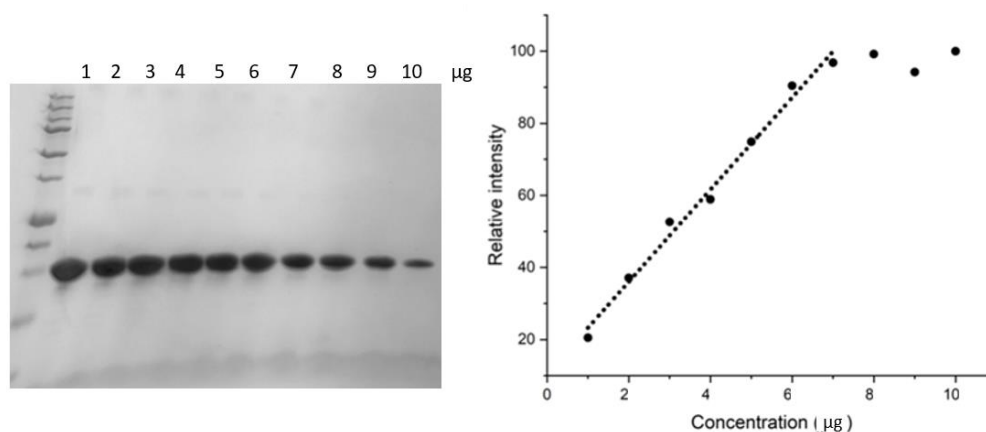
## 2.8.3 ThT assay data analysis

The ThT fluorescence curves were normalised so that the highest value, when reaching a plateau, was equal to 100 %. In the case of incomplete aggregation, data were normalised to a control (usually  $\alpha$ Syn WT). The OriginPro software (OriginPro 2018b 64Bit) was used to determine lag time and elongation rate by fitting a linear gradient to the elongation phase (normalised fluorescence between 40 % and 60 % of the ThT-aggregation curve). The elongation rate is the slope and the lag time is the intersection with the x-axis of the fitted linear curve.  $t_{50}$  times were calculated by GraphPad Prism 9 using a sigmoidal fit with  $\frac{1}{2}$  max. absorbance representing the  $t_{50}$  timepoint. Average values and standard deviation were calculated for at least three repeat measurements.

## 2.8.4 Pelleting assay

The amount of pelletable material was analysed to verify the results of ThT aggregation assays. Endpoint material from ThT aggregation assays were centrifuged for 30 min at 13,000 rpm (Microfuge SN 100/90). The remaining soluble material and the starting

material were run on a SDS-PAGE gel (sample dilution 1:1), imaged on the Alliance Q9 Imager (Uvitec), and the band intensities were determined using ImageJ software. The ratio between band intensities of supernatant and starting material give information about the remaining free monomer (and possible other species not pelleted) in solution. Repeat experiments and loading controls gave an error of ~5-10 % and a saturation in intensities > 7  $\mu\text{g}$  (equivalent to ~70  $\mu\text{M}$   $\alpha\text{Syn}$ ) (**Figure 2.3**) when quantifying band intensities.



**Figure 2.3: Loading control for fibril yield analysis.** 1  $\mu\text{g}$  to 10  $\mu\text{g}$   $\alpha\text{Syn}$  were loaded onto an SDS-PAGE gel to evaluate the concentration range for a linear dependence between band intensity and loaded protein concentration. Based on this experiment concentrations between 1  $\mu\text{g}$  and 7  $\mu\text{g}$  are within the linear range. For all fibril yield analysis experiments a final amount of 5  $\mu\text{g}$  synuclein was loaded on the gel per well. This experiment was performed by Emily Prescott, University of Leeds.

### 2.8.5 Characterisation of supernatant and pellet from pelleting assay

The percentage of amount of pelletable material (% pellet) was determined by SDS-PAGE (see Section 2.7.1). For further investigation of the ThT endpoint samples and the efficiency of the pelleting assay, the whole sample, SN and resuspended pellet (in 90  $\mu\text{L}$   $\text{H}_2\text{O}$ ) was imaged via negative stain TEM (samples prepared as described in Section 2.12). To test if all of the aggregated material is in the pellet after centrifugation, the whole sample and supernatant were analysed by high-performance SEC (HP-SEC) with assistance from Leon Willis, University of Leeds. For this, samples were sealed into 300  $\mu\text{L}$  polypropylene conical insert HPLC vials (VWR) and incubated at 5° C in the SIL-40C autosampler (Shimadzu). 25  $\mu\text{L}$  of sample was loaded onto a TOSOH SWXL guard column (4.0 x 60 mm) connected directly to a TOSOH G3000SWXL HPLC column. Columns were equilibrated in 20 mM Tris-HCl, pH 7.5, 200 mM NaCl. Samples were eluted isocratically with a Shimadzu Nexera LC-40 HPLC system at 0.5 ml/min, with the column kept at 25 °C. The column was calibrated with Blue Dextran (>2000 kDa), ADH (147 kDa), BSA (66 kDa), Cytochrome C (12 kDa), and Vitamin B12 (1.3 Da), with additional SEC-MALS measurements obtained using Wyatt miniDAWN Treos and

Optilab T-rex. SEC data were analysed using LabSolutions software (Shimadzu), with SEC-MALLS data analysed using Astra 6.1 (Wyatt).

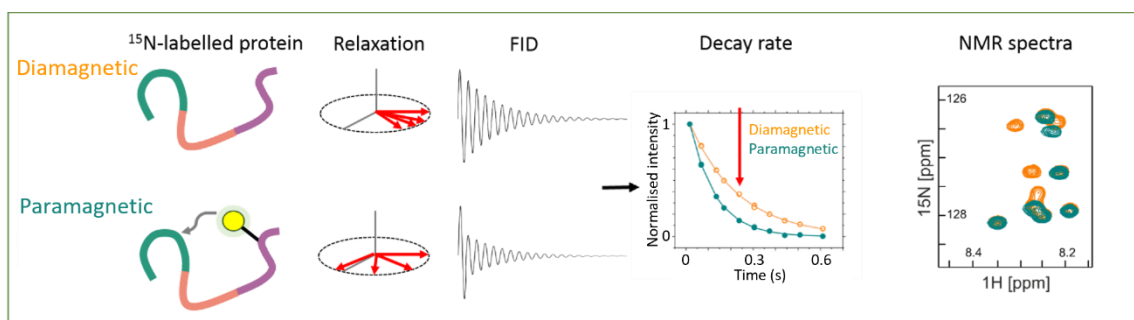
## 2.9 NMR experiments

### 2.9.1 NMR backbone assignment

Backbone assignments of  $\alpha$ Syn WT (20 mM sodium acetate, pH 4.5, 20 mM NaCl (BMRB:27900) and 20 mM Tris-HCl, pH 7.5, 20 mM NaCl (BMRB:51120)),  $\alpha$ Syn  $\Delta\Delta$  (20 mM sodium acetate, pH 4.5, 20 mM NaCl (BMRB:27901)),  $\alpha$ Syn P1P2-GS (20 mM sodium acetate, pH 4.5, 20 mM NaCl (BMRB:27902)) and  $\Delta$ P1 (20 mM Tris-HCl, pH 7.5, 20 mM NaCl (BMRB: 51121)) were kindly performed by Roberto Maya-Martinez, University of Leeds. Data were obtained on a Varian-Inova 750 MHz spectrometer equipped with a cryogenic probe at 15 °C. A triple resonance technique (HNCO, HNC $\alpha$ CO, HNC $\alpha$ C $\beta$ , HNC $\alpha$ NNH, HN(CO)C $\alpha$ NNH) was used with  $^{13}\text{C}/^{15}\text{N}$  uniformly labelled samples containing 200  $\mu\text{M}$  synuclein in the buffer of interest with 10% (v/v) D $_2$ O and 0.02% (w/v) sodium azide.

### 2.9.2 Paramagnetic relaxation enhancement experiments

Paramagnetic Relaxation Enhancement (PRE) experiments were performed to gain insight into the intra- and inter-molecular interactions of  $\alpha$ Syn variants under different buffer conditions, similarly to described previously in the literature<sup>331,335</sup>. PRE experiments require a paramagnetic centre which is often introduced by spin-labelling the analysed protein with MTSL<sup>335,384</sup> (Section 2.7.7). MTSL has an unpaired electron, which increases transverse relaxation rates ( $R_2$  relaxation) and results in a decreased NMR signal (**Figure 2.4**). The effect on the relaxation rate is distance dependent ( $r^{-6}$ ) and allows the determination of distances between label and measured nuclei<sup>385</sup>. Distances up to 25 Å can be detected<sup>386</sup>.



**Figure 2.4: Schematic of PRE experiment.**  $\alpha$ Syn is drawn in green (N-terminal region), red (NAC region) and purple (C-terminal region). MTSL spin label is illustrated in yellow. For the paramagnetic experiment (bottom) the  $R_2$  relaxation of spins is increased (red arrows), resulting in a changed FID compared to the diamagnetic experiment. As a result, the decay rate for the paramagnetic experiment (green) is increased and peak broadening is observed compared to the diamagnetic setup (orange).

### 2.9.2.1 Intramolecular interactions

To enable calibration of the NMR, a final amount of 5 % (v/w) D<sub>2</sub>O and 0.02 % (v/w) sodium azide as biocide were added to a final volume of 500-550  $\mu$ L of 100  $\mu$ M protein solution (20 mM sodium acetate buffer, low (20mM) and high (200 mM) salt, pH 4.5). To detect intramolecular interactions all of the proteins were MTSL and <sup>15</sup>N labelled and <sup>1</sup>H-<sup>15</sup>N BTROSY spectra were obtained on an AVANCE III Bruker spectrometer (600 MHz) with a triple-channel QCI-P cryogenic probe at 15 °C. After measuring the MTSL labelled  $\alpha$ Syn paramagnetic spectrum, a final concentration of 2 mM ascorbic acid was added to the sample in the NMR tube and incubated for 30 min to reduce the spin label before measuring the diamagnetic spectrum (note that small changes in chemical shift occur upon adding the acid to the protein spectra). Data were analysed in Topspin (Bruker) using CCPN<sup>387</sup>, the relative intensity ratios (paramagnetic/diamagnetic) were calculated, and presented as high peaks for each amino acid. For a better overview, data were smoothed in Origin using Percentile Filtering (50 %) over five data points. This approach replaces signal values at each point by the median of a group of neighboured points to eliminate shot noise. Intermolecular interactions at the tested conditions were ruled out by performing control experiments in which 50  $\mu$ M <sup>15</sup>N  $\alpha$ Syn and <sup>14</sup>N  $\alpha$ Syn-MTSL were mixed and showed no PRE. Further, PRE effects arising from nonspecific binding of the hydrophobic MTSL probe to the protein was excluded by performing experiments in which 100  $\mu$ M free MTSL was added to 100  $\mu$ M <sup>15</sup>N  $\alpha$ Syn WT (lacking cysteine) in which no PRE effects were observed. The minimal signal to noise value was calculated as described in **Equation 2.2**.

#### Equation 2.2

$$\text{signal to noise} = \frac{\text{intensity}}{\text{noise level}}$$

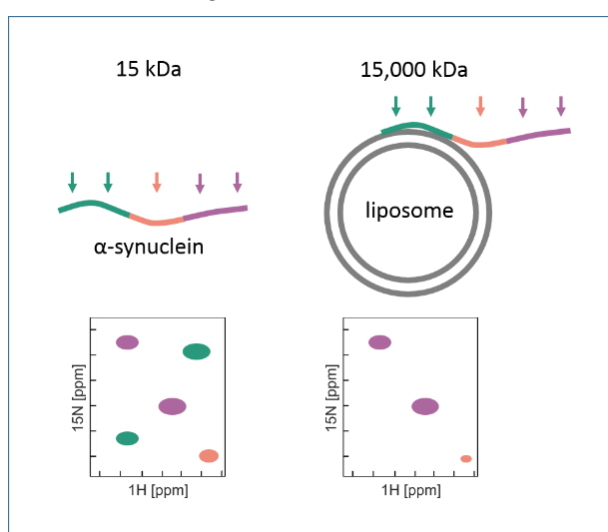
### 2.9.2.2 Intermolecular interactions

Sample preparation and experiments were performed similarly to the procedure described to measure intramolecular interactions (Section 2.9.2.1) besides the following differences: To detect intermolecular interactions, 250  $\mu$ M protein was <sup>14</sup>N-MTSL labelled and mixed with 250  $\mu$ M <sup>15</sup>N labelled (but not MTSL-labelled) protein with a total protein concentration of 500  $\mu$ M. Intermolecular data were recorded on an AVANCE III Bruker spectrometer (600 MHz) using <sup>1</sup>H-<sup>15</sup>N correlation based pulse sequences with 10 time points (16.96 – 610.56 ms). Data for both oxidised and reduced (addition of 2 mM ascorbic acid) samples were collected and R<sub>2</sub> rates were determined by fitting the relaxation data to single exponentials using NMRPINT<sup>388</sup>. The relative relaxation rate changes for each amino acid were determined and illustrated in a histogram.



### 2.9.3 NMR experiments on liposomes

NMR experiments in the presence of liposomes were performed to identify residues of  $\alpha$ Syn involved in the binding process (**Figure 2.5**). For this,  $^1\text{H}$ - $^{15}\text{N}$  BTROSY spectra were obtained using  $^{15}\text{N}$ -labelled  $\alpha$ Syn in the absence and presence of liposomes in a 60:1 [DMPS]:[ $\alpha$ Syn] (mol/mol) ratio. 25  $\mu\text{M}$  synuclein in 20 mM sodium phosphate, pH 6.5 containing 10% (v/v)  $\text{D}_2\text{O}$ , 0.02% (w/v) sodium azide were used and analysed on an AVANCE III Bruker spectrometer (600 MHz) equipped with a cryogenic probe at 20  $^\circ\text{C}$ . Assignments from the literature were used to analyse the data (BMRB 16543)<sup>356</sup> and the spectra were processed in Topspin (Bruker) and examined in CCPN analysis. The peak heights of synuclein in the presence of liposomes relative to synuclein in isolation was calculated and illustrated as a histogram.



**Figure 2.5: Schematic of liposome binding experiment.**  $\alpha$ Syn is drawn in green (N-terminal region), red (NAC region) and purple (C-terminal region). Due to its small size ( $\sim 15$  kDa) and hydrodynamic radius, each residue correlates with a peak. In the presence of liposomes (right), the mass is increased by  $\sim 1000$  x resulting in an increased hydrodynamic radius. This leads to peak broadening effects for peaks correlating with residues involved in the binding process.

### 2.9.4 NMR experiments in the presence of peptides

For all NMR experiments in the presence of P1 or P1-SG peptide,  $^1\text{H}$ - $^{15}\text{N}$  BTROSY spectra were obtained using 100  $\mu\text{M}$   $^{15}\text{N}$  spin-labelled  $\alpha$ Syn in 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 15  $^\circ\text{C}$ . For chemical shift perturbation analysis 0, 500  $\mu\text{M}$  or 1mM peptide (P1-peptide or P1-SG peptide) was added and data acquired using a Bruker AVANCE III 750 MHz spectrometer. Spectra were processed in Topspin (Bruker). Peak positions and intensities were extracted using ccpNMR-analysis, and HN-CSP were calculated using **Equation 2.3**:

**Equation 2.3**

$$\Delta\delta = \sqrt{(5 * \delta^1H)^2 + (\delta^{15N})^2}$$

For the comparison of chemical shifts at different pH values, 100  $\mu\text{M}$   $\alpha\text{Syn}$  in 20 mM Tris-HCl, pH 7.5, 20 mM NaCl or in 20 mM sodium acetate, pH 4.5, 20 mM NaCl were measured, and peak positions analysed as described above (**Equation 2.3**).

For PRE NMR experiments, 100  $\mu\text{M}$   $^{15}\text{N}$  spin-labelled protein with 100  $\mu\text{M}$   $^{14}\text{N}$  MTSL labelled peptide was used. The diamagnetic spectra were acquired 30 min after adding 1 mM ascorbic acid. Data were collected using a Bruker AVANCE III 950 MHz spectrometer and data were processed as described above, the peak heights being used to calculate intensity ratios (paramagnetic/ diamagnetic).

#### 2.9.4.1 Data acquisition, processing and analysis

HN-BTROSSY experiments were acquired on an AVANCE III Bruker spectrometer (600 MHz or 750 MHz) under the following conditions: 1274 and 256 points for the direct and indirect dimensions were measured, respectively. Four scans per increment were obtained. The spectral width was set to 27ppm, the transmitted offset frequency was set to 4.7 and 120 ppm, for direct and indirect dimensions, respectively. With the used parameter set up a fid resolution of 15.24 and 15.98 Hz, for direct and indirect dimension, respectively was measured.

Data were analysed in Topspin (Bruker) using CCPN. Each cross-peak was identified as a specific residue based on the previous assignment (see Section 2.9.1). In case of overlapping peaks, these data were not included in the analysis. Following, peak height (for PRE) and position (for CSPs) were extracted and evaluated.

## 2.10 Circular dichroism

Circular Dichroism (CD) is a powerful technique to determine the secondary structure content of proteins<sup>389</sup>. It uses the fact that different structural elements have characteristic CD spectra caused by an unequal absorption of left-handed and right-handed circularly polarized light<sup>390</sup>. A relevant example for this work is that  $\alpha$ -helical proteins have negative bands at 222 nm and 208 nm and a positive band at 193 nm<sup>391</sup>.

### 2.10.1 Binding to liposomes

Synuclein forms an  $\alpha$ -helix structure in the N-terminal and NAC region when binding lipid membranes or micelles<sup>314</sup>, CD can be used to monitor this binding.

To determine the binding affinity between synucleins and DMPS liposomes (see Section 2.7.8 for liposome preparation) based on the gain of secondary structure, CD spectroscopy was performed by incubating 25  $\mu\text{M}$   $\alpha\text{Syn}$ ,  $\gamma\text{Syn}$  or variants with different concentrations of liposomes (0x to 150x molar excess lipids) in 20 mM sodium phosphate buffer, pH 6.5. Far-UV (190-260 nm) CD spectra were acquired in quartz cuvettes

(Hellma) with 1 mm path length using a Chirascan<sup>TM</sup> plus CD Spectrometer (Applied Photophysics). The experiments were performed at 30 °C using a 2 nm bandwidth and 1 s time steps. Data were collected at 1 nm increments and an average of 3 scans (190-260 nm) was taken per sample. The data were fitted to secondary structure content using Dichroweb<sup>392</sup>.

Binding affinity ( $K_d$ ) and stoichiometry ( $L$ ) were calculated as described before by Galvagnion *et al.*<sup>351</sup>. The measured CD spectra can be described as shown in **Equation 2.4** in which  $x_B$  and  $x_F$  are the fractions of bound  $\alpha$ Syn to the membrane and free synuclein, respectively and  $CD_B$  and  $CD_F$  are the detected CD signals of the bound and free protein. It is assumed that the sum of  $x_B$  and  $x_F$  equals 1 and that the CD signals are given by the measurement in buffer (for  $CD_F$ ) or in the presence of saturating liposomes (for  $CD_X$ ).

**Equation 2.4**

$$CD_{obs} = x_B CD_B + x_F CD_F$$

The fraction of synuclein bound to the liposomes was determined for each sample using **Equation 2.5**.

**Equation 2.5**

$$x_B = \frac{CD_{obs} - CD_F}{CD_B - CD_F}$$

The binding characteristics were analysed by fitting the CD data to the fitting function shown in **Equation 2.6**.

**Equation 2.6**

$$x_B = \frac{\left( \left( [\alpha\text{Syn}] + \frac{[\text{DMPS}]}{L} + K_D \right) - \sqrt{\left( \left( [\alpha\text{Syn}] + \frac{[\text{DMPS}]}{L} + K_D \right)^2 - \frac{4[\text{DMPS}][\alpha\text{Syn}]}{L} \right)} \right)}{2[\alpha\text{Syn}]}$$

Here,  $x_B$  describes the fraction of bound synuclein to the liposomes,  $L$  is the number of DMPS molecules involved in binding one synuclein molecule and can be described as shown in **Equation 2.7**.

**Equation 2.7**

$$[\text{DMPS}] = L([\text{DMPS}_L] + [x_B(\text{DMPS}_L)])$$

The python script to perform the fitting was kindly provided by Bob Schiffrin, University of Leeds.

### 2.10.2 Binding to peptides

To detect if peptides formed amyloid structures when incubated over time, far UV CD spectra of peptides P1 and P1-SG (20  $\mu$ M, 20 mM Tris-HCl, pH 7.5, 200 mM NaCl) were acquired in quartz cuvettes (Hellma) with 1 mm path length, using a 2 nm bandwidth, 1 s time steps and 1 nm increments at 25 °C using a Chirascan™ plus CD Spectrometer (Applied Photophysics). Three scans ranging from 190 to 260 nm were measured for each sample and averaged.

## 2.11 Mass Spectrometry analysis

To investigate the binding mode between synuclein and the P1-peptide,  $\alpha$ Syn WT and  $\Delta$ P1 samples with a final concentration of 20  $\mu$ M were prepared in 20 mM aqueous ammonium acetate buffer (pH 7.5). The P1 peptide was diluted to a final molar ratio of  $\alpha$ Syn and P1-peptide of 1:10. Native ESI-MS analysis was performed on a Synapt G1 HDMS instrument (Waters Corp., Wilmslow, UK). All samples were analysed using positive ionisation ESI with a spray capillary voltage of 1.2 kV. The following instrumental parameters were used: source temperature 30 °C; sampling cone 30 V; backing pressure 2.25 mbar; extraction cone 1 V; trap collision energy 5 V; trap DC bias 30 V; transfer collision energy 2 V. The system was calibrated with NaI cluster ions from a 2  $\mu$ g/ $\mu$ L 50:50 2-propanol:water solution. Data were acquired over the  $m/z$  range of 100-4000 and processed by using MassLynx V4.1 supplied with the mass spectrometer. CID MS/MS experiments were conducted in the trap cell of the Synapt G1 mass spectrometer with argon gas, collision energy was applied to the trap cell in steps from 5 V to 60 V. Experiments were performed by Emily Byrd, University of Leeds.

## 2.12 Transmission Electron Microscopy

For negative stain TEM images, end-point samples from ThT assays (usually after 100 h for *de novo* growth and 40 h for seeding experiments) (Section 2.8.2) were diluted 5x in 18 M $\Omega$  H<sub>2</sub>O. 5  $\mu$ L samples were put on a carbon coated copper grid and incubated for 20-30 s, dried with filter paper and then the grid was washed three times with water in a drop wise fashion with drying steps in-between each wash. To natively stain the fibrils, 5  $\mu$ L 1 % (w/v) uranyl-acetate was then applied and blotted off twice before imaging the samples on the FEI Tecnai T12 or Joel JEM-1400 electron microscope.

## 2.13 Atomic force microscopy (AFM)

Mica was freshly cleaved before being treated to create a positive surface charge by adding poly-L-lysine (70-150 kDa) at 15  $\mu$ g/mL for 10 sec followed by drying with nitrogen.

A sample volume of 90  $\mu\text{L}$  of protein (WT  $\alpha\text{Syn}$ , L38M, Y39A or S42A) was taken at the end point of a fibril growth assay (Section 2.8.2) before being deposited at a concentration of 30  $\mu\text{M}$  onto poly-L-lysine treated mica and allowed to incubate for 4 min. The mica surface was then rinsed with buffer (50 mM sodium phosphate buffer, 300 mM KCl, pH 7.5) via fluid exchange, maintaining the samples in a liquid environment. AFM observations were performed in liquid in tapping mode using a Dimension FastScan Bio with FastScan-D-SS probes (Bruker) in the same buffer. The force applied by the tip on the sample was minimised by maximising the set point whilst maintaining tracking of the surface. Heights of single particles were measured automatically using routines written in MATLAB (<https://github.com/George-R-Heath/Particle-Detect>). Heights and lengths of fibrils were measured either automatically using MATLAB (<https://github.com/George-R-Heath/Correlate-Filaments>) or manually in ImageJ for densely packed overlapping fibrils. AFM imaging and analysis were performed by George Heath, University of Leeds.

## 2.14 Computational techniques

### 2.14.1 *In silico* identification of aggregation-prone regions

The aggregation propensity, solubility and ability to form a steric zipper were analysed by using the online tools Zyggregator<sup>393</sup>, Camsol<sup>114</sup> and ZipperDB<sup>394</sup> at pH 4.0 or 7.0. Note, that tools can only be used at integral pH values (e.g. not pH 4.5 and 7.5).

### 2.14.2 ConSurf analysis

ConSurf is a bioinformatic tool to estimate the evolutionary conservation of primary protein sequence based on the phylogenetic relations between homologous sequences (<https://consurf.tau.ac.il/overview.php>)<sup>395</sup>. For this study, the  $\alpha\text{Syn}$  WT sequence was used and compared to the 150 closest homologues (automatically picked) in the UniProt database.

### 2.14.3 Genomic analysis of $\gamma\text{Syn}$

Genomic DNA was extracted from blood leukocytes using Wizard Genomic DNA Purification Kit (Promega). The whole coding region of the SNCG gene was analysed by direct sequencing. Nucleotide sequence analysis was performed on a capillary genetic analyser ABI Prism 3130 (Applied Biosystems). Data Collection Software, (version v3.0), Sequencing Analysis Software (version v5.2) and SeqScape Software (version v2.5) were used for data analysis. The following primers flanking the SNCG exons with adjacent intronic regions (at least 50 nucleotide pairs from each end) were used: 5'-TGGAGGAAGGTGAGGCTGA-3' and 5'-ACATAGGTGTGCACAGGGC-3' for exon I, 5'-CCCCACATTCTGTCCTGTC-3' and 5'-GCGCTCAGTGGGTACTGAAA-3' for exons

II-III, 5'-TCATCAGAGCCCCGGGTATT-3' and 5'-CCAGGTCCTCCCAGGAACA-3' for exons IV-V. The genomics analysis was kindly performed by the Illarioshkin lab, Research Center of Neurology, Russia.

## 2.15 Biological techniques

### 2.15.1 Cellular viability assays

SH-SY5Y cells were used for the cell toxicity assays and experiments were performed by Madeline Brown, University of Leeds. 100 000 cells per well were seeded in a 96-well plate and grown for 24 hours before incubating them with a final concentration of 1 or 10  $\mu\text{M}$  monomer or ThT endpoint sample (Section 2.8.2). Controls included untreated cells and 2 % (w/v)  $\text{NaN}_3$  treated cells. The cells were further incubated for 48 h. Then, various cell viability assays were performed to qualify the toxic effect of  $\alpha\text{Syn}$  samples on SH-SY5Y cells.

#### 2.15.1.1 MTT assay

MTT-assay (identifies mitochondrial activity): After exposing the cells to  $\alpha\text{Syn}$  samples, 5 mg/mL 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) in PBS was added to each well. The plates were incubated for 2 h hours at 37 °C. Then, the medium was removed from each well and 100  $\mu\text{L}$  DMSO was added. The solution was incubated for 10 min before measuring the absorbance at 540 nm and 650 nm on a FluoStar OPTIMA or Fluostar OMEGA fluorescence plate reader.

#### 2.15.1.2 LDH assay

LDH assay (identifies cell disruption): The LDH assay was performed using the Pierce LDH Cytotoxicity Assay Kit. 50  $\mu\text{L}$  cell medium was transferred to a new plate and mixed with the same volume of Reaction Mixture. After incubating for 30 min at room temperature, the reaction was stopped by adding the Stop Solution. Finally the absorbance at 490 and 680 nm was measured on a FluoStar OPTIMA or Fluostar OMEGA fluorescence plate reader.

#### 2.15.1.3 ATP assay

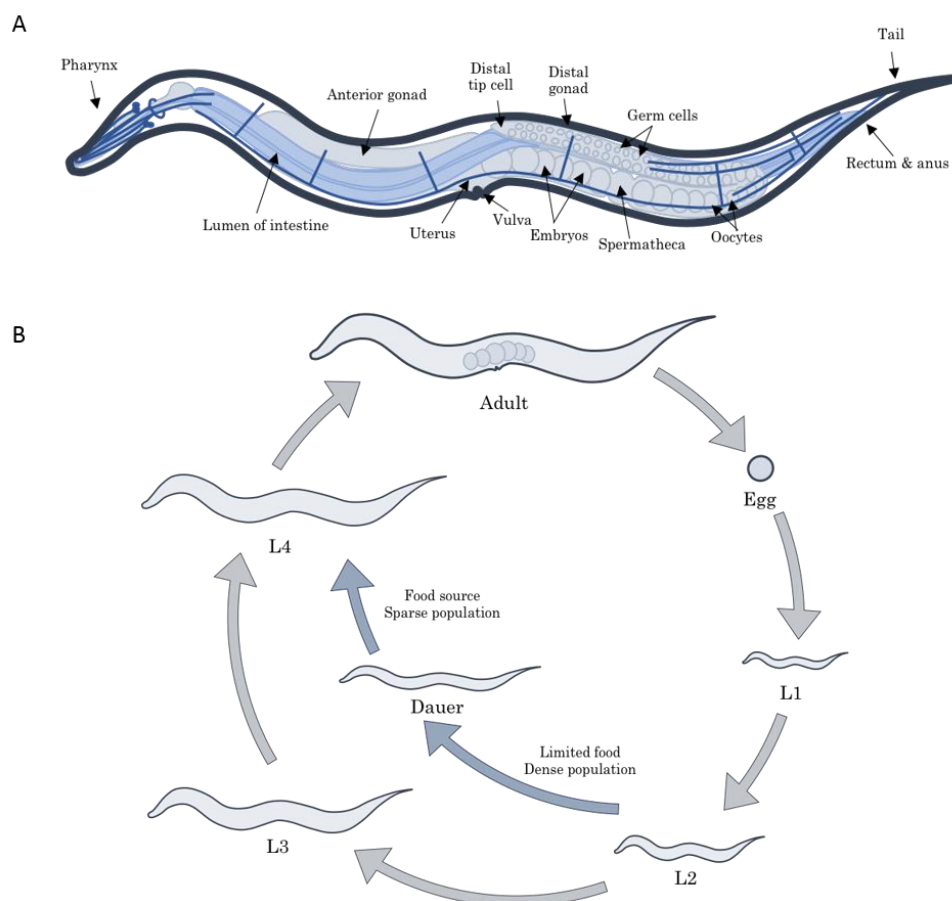
ATP assay (identifies inhibition of cell growth): For the ATP assay the ATPlite Luminescence Assay System Kit (PerkinElmer) was used. Cells were mixed with 50  $\mu\text{M}$  lysis buffer before incubating for 5 min, 600 rpm. Then, 50  $\mu\text{L}$  of a luciferase and D-luciferin mix was added and the luminescence was measured on a FluoStar OPTIMA or Fluostar OMEGA fluorescence plate reader.

### 2.15.2 Inclusion formation in cells

Human neuroglioma cells (H4) were maintained in Opti-MEM I Reduced Serum Medium (Life Technologies- Gibco, Carlsbad, CA, USA) supplemented with 10% Fetal Bovine Serum Gold (FBS) (PAA, Cölbe, Germany) and 1% Penicillin-Streptomycin (PAN, Aidenbach, Germany). The cells, expressing  $\gamma$ Syn, were grown at 37°C in an atmosphere of 5% CO<sub>2</sub>. For imaging inclusion formation, cells were plated on 13 mm glass coverslips in 24-well plates. Forty-eight hours after transfection with the  $\gamma$ Syn constructs, H4 cells were washed with PBS and fixed with 4% paraformaldehyde (PFA) for 20 minutes at 25 °C, followed by a permeabilisation step with 0.5% Triton X-100 (Sigma-Aldrich, St. Louis, MO, USA) for 20 minutes. To visualise  $\gamma$ Syn, cells were incubated with a primary antibody (mouse anti- $\gamma$ -synuclein (1:500, Santa-Cruz Biotechnology)), over-night at 4 °C and secondary antibody (Alexa Fluor 488 donkey anti-mouse IgG (Life Technologies- Invitrogen, Carlsbad, CA, USA)). Nuclei were stained with 4'6'-diamidino-2-phenylindol (DAPI, Sigma-Aldrich, D8417) (1:5000 in DPBS) for 10 minutes. After a final wash, coverslips were mounted by using Mowiol (Sigma Aldrich, St. Louis, MO, USA) and subjected to fluorescence microscopy. The proportion of cells with  $\gamma$ Syn inclusions was determined by analysing at least 150 cells per experiment. Images were acquired using a 63x objective, and analysed using LAS AF v.2.2.1 (Leica Microsystems) software. Statistical analyses of the data were performed using ANOVA. Experiments and analysis of the data were performed by the Outeiro lab, University of Göttingen, Germany.

### 2.15.3 Experiments with *C. elegans*

The ~1 mm long *C. elegans* worm is a good model organism to study neurodegenerative disease associated with amyloid formation<sup>83,396-398</sup>. A rapid reproductive cycle and a short lifecycle (~2 to 3 weeks)<sup>399</sup>, a complete mapping of the entire cell lineage<sup>400</sup>, and a fully sequenced genome offer many advantages (**Figure 2.6**). Also, the nematodes are transparent, allowing the use of fluorescent proteins to label proteins of interest to investigate biological processes *in vivo*. Around 60-80% of all human genes and 42% of human disease genes possess an ortholog in *C. elegans*<sup>401</sup>.



**Figure 2.6: Schematic of tissues and lifecycle of *C. elegans*.** (A) Adult hermaphrodite *C. elegans* worm with selected anatomical features highlighted. (B) Life cycle of the self-fertilising *C. elegans* hermaphrodite. Adult worms lay eggs, which hatch into the L1 stage. If conditions are favourable, the larvae continue through developmental stages into reproductive adults. Under unfavourable conditions they enter the dauer stage, an alternative L3 stage. Figure taken from Sarah Good, University of Leeds.

### 2.15.3.1 Strain generation

The synuclein genes were fused at its C-terminus to yellow fluorescent protein (YFP) in vector pPD30.38, expressing the protein in the body wall muscle cells under the *unc-54* promoter. The following strains were used: N2 (control strain, no expression of synuclein nor YFP),  $\alpha$ WT,  $\Delta$ P1,  $\Delta\Delta$ ,  $\alpha$ L38M,  $\alpha$ Y39A,  $\alpha$ S42A,  $\gamma$ WT, and  $\gamma$ M38I. The  $\alpha$ WT::YFP expressing worms were created using gene bombardment and were kindly provided by E. Nollen (University of Groningen),  $\alpha$ Syn variants were generated by Q5 site-directed mutagenesis (Section 2.5.1). The  $\gamma$ Syn gene was purchased from Eurofins and cloned into the pPD30.38 vector (Section 2.5.2). Transgenic *C. elegans* were then generated by microinjection of the construct into the germline of N2 nematodes by Dovile Milonaityte, University of Leeds. The expression level was evaluated by a Western blot analysis (Section 2.7.3).



### 2.15.3.2 *C. elegans* maintenance

*C. elegans* strains were maintained at 20°C on 60 mm Nematode growth medium (NGM) agar plates (1.7% (w/v) agar, 50 mM NaCl and 0.25% (w/v) bactopectone, autoclaving and then adding 25 mM KPO<sub>4</sub> pH 6.0, 1 mM CaCl<sub>2</sub>, 1 mM MgSO<sub>4</sub>, 5 µg/mL cholesterol, and 200 µg/mL streptomycin) seeded with HT115 (DE3) *Escherichia coli* as a food source. All worm strains were maintained on NGM plates with plenty of food source and transferred to a new plate every few days using a platinum wire pick.

Bleaching was performed to age synchronise worms. For this, a worm population was grown on NGM plates until a large population of eggs and non-starved adults were present. The animals were washed from the plate using M9 buffer (5.8g Na<sub>2</sub>HPO<sub>4</sub>·7H<sub>2</sub>O, 3g KH<sub>2</sub>PO<sub>4</sub>, 5g NaCl, 1mM MgSO<sub>4</sub> in 1 L H<sub>2</sub>O) into a 15 mL microcentrifuge tube and pelleted by centrifugation at 1000 rpm for 1 minute. 10 mL of bleach solution (1% sodium hypochlorite (v/v), 0.25M NaOH) was added to the nematodes-containing pellet and the tubes were inverted for 5 minutes (worm bodies dissolve, eggs remain) before centrifugation for 1 min at 1000 rpm. The pellet was washed three times with 15 mL M9 buffer to remove any trace of bleach solution. Finally, the worm pellet was resuspended in 1 mL of M9 buffer and placed on a rotator overnight for the eggs to hatch. The next day the hatched larvae were added to NGM plates to further develop. Worms were transferred every second day to a new NGM plate until the desired stage was reached.

### 2.15.3.3 *C. elegans* phenotypic analysis

To measure the amount of inclusions formed over time, age synchronised *C. elegans* were imaged using a Zeiss LSM880 confocal fluorescent microscope through a 20 × 1.0 or a 40 × 1.0 numerical aperture objective with a 514-nm line for excitation of YFP. Before imaging, worms were anesthetized using 5 mM sodium azide solution and mounted on 2% (w/v) agar pads. The number of αSyn::YFP foci were then counted, and the mobility of all foci in at least ten animals per time point and in three independent cultures of *C. elegans* (biological replicates) was determined using fluorescence recovery after photobleaching (FRAP), as described previously<sup>83</sup>.

As the synuclein constructs were expressed in the body wall muscle cells, the motility of the worms was analysed (in body bends per second (BBPS)) to investigate the phenotypic effect. For this, a total of 10 age-synchronised animals were used for each assay, and each experiment was repeated at least three times. Animals were moved into M9 buffer at indicated time points (day 0 through to day 13 of adulthood), and thrashing rates were measured by counting body bends for 15 s using the wrMTrck plugin for ImageJ (available at <http://www.phage.dk/plugins/wrmtck.html>)<sup>402</sup>. Error bars represent standard error of the mean of three biological replicates. Worm experiments in Chapter 3 were performed by Sarah Good and Jemma Makepeace, University of Leeds.



## Chapter 3

# A motif in the N-terminal region crucial for aggregation and function of $\alpha$ -synuclein

### 3.1 Introduction

$\alpha$ -Synuclein ( $\alpha$ Syn) is a 140 amino acid, intrinsically disordered protein involved in synucleopathies including PD, DLB and MSA<sup>403</sup>. Although being discovered more than 30 years ago<sup>246</sup> and being aware of the correlation with neurodegenerative disease since 1997<sup>247</sup>, the function and pathogenicity of  $\alpha$ Syn are still not understood in detail<sup>250,251,404</sup>. On a primary structural level,  $\alpha$ Syn can be divided into three regions: the N-terminal region (residues 1-60), the NAC region (residues 61-95) and the C-terminal region (residues 96-140) (Figure 1.22). So far, research was predominantly focused on the NAC region which is hydrophobic, highly aggregation-prone and forms the core of all solved fibril structures to date<sup>96,117,378</sup>. Deletion experiments with a removed NAC region resulted in inhibition of amyloid growth<sup>313</sup>, whilst aggregation assays with only the NAC peptide presented rapid fibril formation<sup>405</sup>, highlighting the crucial role of this domain for aggregation.

More recently, the importance of regions flanking these highly aggregation-prone and core-forming regions in amyloid have become more prevalent<sup>112,128</sup>. In the case of  $\alpha$ Syn, it could be demonstrated that this IDP adopts an ensemble of conformations, strongly dependent on the status of N- and C-terminal regions (e.g. charge, post-translational modifications or interaction partners)<sup>189,378</sup>. Exemplary, the charge pattern of  $\alpha$ Syn with 12 basic (including 1 His) and 15 acidic residues in the N- and C-terminal regions, respectively, imply a high pH and ionic strength dependence for its conformational properties<sup>319,323</sup>. Indeed, NMR experiments at neutral pH have shown that N- and C-terminal regions interact and consequently protect the NAC region from aggregation, whilst at acidic pH, the C-terminal region collapses and exposes the NAC region,

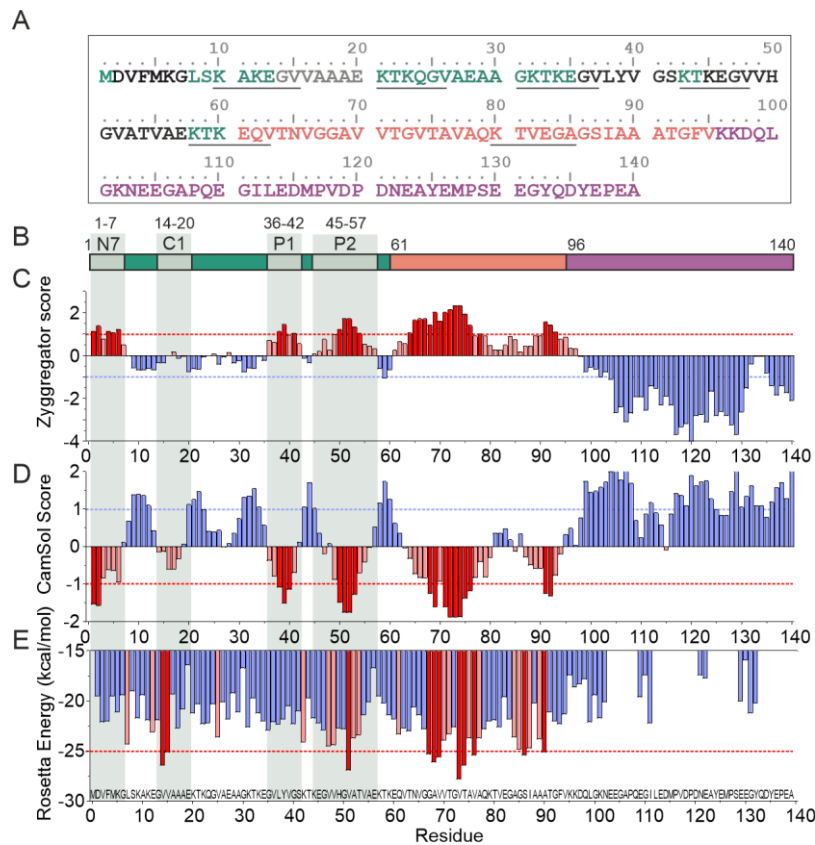
accelerating fibril formation<sup>335</sup>. Both these conditions are biologically relevant as  $\alpha$ Syn can be found in various environments, including the cytoplasm and the endo-lysosomal pathway of a cell during its lifetime<sup>34,189</sup>. The effects of sequence changes on the conformational properties and aggregation rates of IDPs have been widely studied<sup>406-409</sup>. Interestingly for  $\alpha$ Syn, out of the twelve known familial point mutations that lead to PD, nine are located in the N-terminal region between residues 5 to 53 that flank the NAC region (**Table 1.3**). This leads to an increased interest in the N-terminal region of  $\alpha$ Syn. Whilst the C-terminus is considered to be disordered and highly dynamic, the N-terminal region can adopt multiple secondary structural elements. When binding lipids (in the form of membranes, liposomes or micelles), important for the physiological function of  $\alpha$ Syn<sup>272,274</sup>,  $\alpha$ -helix formation is observed throughout the first 97 N-terminal amino acids<sup>185,314</sup>. Residues 36-55 were observed to form a  $\beta$ -hairpin *in vitro*<sup>341</sup> and *in silico*<sup>188</sup>, and finally Eisenberg and co-workers showed that the isolated preNAC peptide (residues 47–56) forms amyloid-like structures<sup>311</sup>. These observations allow the hypothesis that the N-terminal region of  $\alpha$ Syn might play a critical role in both, fibril self-assembly and physiological function.

In this chapter, *in silico* analysis of the primary sequence of  $\alpha$ Syn was used to identify aggregation-prone regions throughout the full-length protein. The identified regions were examined via numerous techniques to understand their role and involvement in aggregation and function in an *in vitro* and *in vivo* context. Aggregation assays and NMR experiments gave insights into key interactions required to initiate fibril formation, and experiments with liposomes shine light on the functional contribution of these N-terminal motifs.

### 3.2 Identification of a N-terminal aggregation-prone motif

Computational approaches were used to predict the presence of insoluble/aggregation-prone regions in full length human  $\alpha$ Syn based on its amino acid sequence. The Zyggregator score, providing information about the tendency to form amyloid structures based on the propensity of polypeptide chains to form proto-fibrillar assemblies<sup>393</sup>, CamSol which predicts the local solubility of a protein<sup>114</sup> and the Rosetta energy, developed by Eisenberg and co-workers calculating the propensity to form steric  $\beta$ -zipper, a common structural feature considered in amyloid<sup>394,410</sup> were used. As expected, the NAC region (residues 61-95) is predicted to be aggregation-prone, hydrophobic and able to form a steric zipper (**Figure 3.1**). Interestingly, three additional sequences in the N-terminal region were identified to fulfil these criteria (not all of them showing steric zipper formation): residues <sup>1</sup>MDVFMKGL<sup>7</sup> (named N7), <sup>36</sup>GVLYVGS<sup>42</sup>

(named P1) and  $^{45}$ KEGVVHGVATVAE $^{57}$  (named P2) (**Figure 3.1**). The C-terminal region is predicted to be highly soluble and non-aggregation-prone.



**Figure 3.1: Aggregation and solubility profiles of  $\alpha$ Syn.** (A) Human  $\alpha$ Syn sequence. The N-terminal region (1–60), NAC region (61–95) and C-terminal region (96–140) are coloured in green, red, and purple, respectively. N7-region, control region 1 (C1), P1 and P2 regions are coloured pale gray and black. The imperfect KTKEGV repeats are underlined in black. (B) Schematic of regions of  $\alpha$ Syn highlighting the N7, P1, P2 and C1 sequences. (C-E) Zyggregator, CamSol and ZipperDB profiles for the  $\alpha$ Syn sequence. Red bars indicate aggregation-prone/low-solubility regions. Light red bars indicate residues with a higher than average aggregation propensity/low solubility that do not meet the threshold and blue bars show non aggregation-prone/low-solubility regions. Red dashed lines indicate the low-solubility/high-aggregation propensity threshold, and blue dashed lines show threshold values for high-solubility/low-aggregation propensity.

The NAC region was focus of many studies in the past<sup>311,313,365,378</sup> and the role of the N-terminal residues 1 to 30 in fibril formation was also analysed in detail previously, using deletion and insertion protein variants<sup>411,412</sup>. Deleting the N-terminal 10 to 30 residues resulted in an increased lag time measured by ThT fluorescence assays in some studies<sup>411</sup>, but others observed a faster aggregation compared with  $\alpha$ Syn WT when deleting residues 1 to 36 of the N-terminus<sup>345,413</sup>. This might be explained by differences in experimental conditions (all mentioned assays were performed at pH 7.4, but using different buffers (e.g. Tris vs NaPi) and using diverse salt concentrations and additives such as SDS) and/or final fibril architectures formed<sup>345</sup>. Experiments focussing on the N-terminal imperfect repeats (**Figure 3.1 A**) found faster fibril growth when deleting two repeat motifs (residues 9-30) whilst when two additional motifs were added (by

duplicating residues 9-30) aggregation was observed to be slowed down<sup>412</sup>. Eisenberg and co-workers identified the pre-NAC region (residues 47-56), which is similar to the P2 region (residues 45-57), as able to form amyloid-like fibrils in isolation and therefore predicted it to be important for aggregation of the full length protein<sup>311</sup>. The P2 region also includes six of the twelve known familial PD mutations and forms the protofilament interface between some (but not all)  $\alpha$ Syn fibril structures determined to date (Section 1.4.3.3, **Figure 1.24**)<sup>96,117</sup> further highlighting its important role for amyloid growth. In addition, it could be shown experimentally (*in vitro* and *in vivo*) and computationally that the combined P1 and P2 region (spanning residues 36-55) form a  $\beta$ -hairpin structure that drives aggregation<sup>187,188,414,415</sup>. The hairpin structure forms intermolecular interactions important for nucleation<sup>188</sup> that can be inhibited *in vitro*, in *Drosophila* and in primary neurones by binding to a  $\beta$ -wrapin (engineered binding protein) to this region<sup>187,342,414,415</sup>. Still, the exact role of the N-terminal sequence, especially P1, for fibril formation of the full-length protein is not clarified and will be further analysed in this chapter using deletion and substitution variants and various biophysical and biochemical techniques. The majority of the results have been published in Ref<sup>416</sup>.

### 3.3 Engineering recombinant $\alpha$ Syn

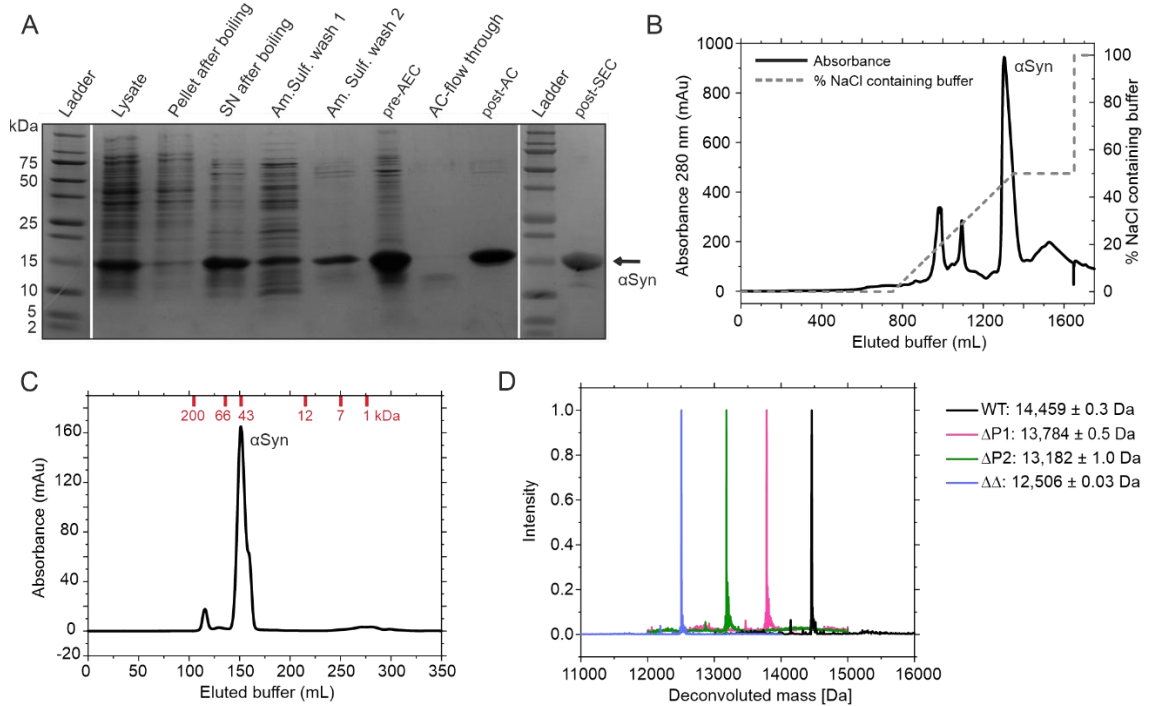
#### 3.3.1 Molecular biology

The  $\alpha$ Syn-pET23a plasmid was kindly provided by Prof. Jean Baum (Department of Chemistry and Chemical Biology, Rutgers University, NJ, USA). To investigate the role of P1 and P2 in the full-length  $\alpha$ Syn protein, deletion and substitution variants were generated using the England Biolabs Q5<sup>®</sup> mutagenesis protocol. For PRE NMR experiments, cysteine residues were inserted at various positions throughout the protein. The mutated genes were sent for sequencing (Eurofins) to confirm a successful mutagenesis. Detailed methods and used primers are provided in Sections 2.5.1, 2.5.3, 2.5.4, and **Table 2.2**.

#### 3.3.2 Expression and purification

All  $\alpha$ Syn variants were expressed in the same way unless mentioned otherwise.  $\alpha$ Syn was expressed in *E. coli* BL21 (DE3) cells and purified by cell lysis, boiling, ammonium sulphate precipitation, anion exchange chromatography (AEC) and size exclusion chromatography (SEC) (see Section 2.6 for detailed methods). An exemplary SDS-PAGE gel for  $\alpha$ Syn WT, to demonstrate a typical purification, is shown in **Figure 3.2 A** with an  $\alpha$ Syn band at  $\sim$ 15 kDa. AEC and SEC chromatograms are illustrated in **Figure 3.2 B and C**, respectively. Interestingly,  $\alpha$ Syn WT eluted at a higher molecular mass than expected for a globular monomeric protein. This is most likely due to  $\alpha$ Syn being an intrinsically disordered protein with an extended structure resulting in a larger

hydrodynamic radius. Finally, all proteins were analysed by ESI-MS (main proteins analysed in this study shown as an example in **Figure 3.2 D**. For all used protein variants and molecular masses see **Table 3.1**) to confirm the correct mass and purity of the sample. Protein production yielded in typically 10-15 mg/L growth medium.

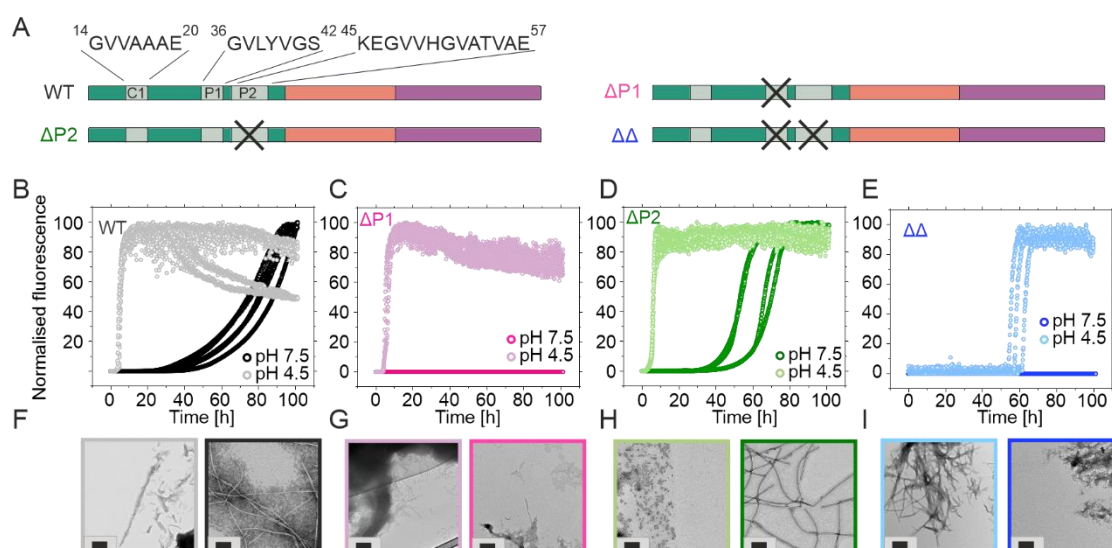


**Table 3.1: All purified  $\alpha$ Syn variants, expected molecular masses (top) and measured molecular masses (via MS, bottom) of the proteins used in this chapter.  $^{15}$ N labelled proteins are indicated with a superscript 15.**

Variant	Mass	Variant	Mass	Variant	Mass
WT	14,460 Da	P1-GS	14,304 Da	WT	14,476 Da
	14459 $\pm$ 0.3 Da		14,304 $\pm$ 0.9 Da	S129C	14,476 $\pm$ 0.2 Da
$^{15}$ WT	14,626 Da	P1P2-GS	13,948 Da	WT	14,492 Da
	14,621 $\pm$ 1.2 Da		13,947.4 $\pm$ 0.06 Da	A140C	14,491 $\pm$ 0.4 Da
$\Delta$ P1	13,784 Da	WT A18C	14,492 Da	$^{15}$ WT	14,658 Da
	13,784 $\pm$ 0.5 Da		14,490 $\pm$ 0.05 Da	A140C	14,653 $\pm$ 1.1 Da
$\Delta$ P2	13,183 Da	$^{15}$ WT	14,658 Da	$^{15}$ $\Delta\Delta$	12,682 Da
	13,182 $\pm$ 1.0 Da	A18C	14,651 $\pm$ 0.8 Da	A18C	12,677 $\pm$ 0.5 Da
$\Delta\Delta$	12,506 Da	WT V40C	14,464 Da	$^{15}$ $\Delta\Delta$	12,682 Da
	12,506 $\pm$ 0.03 Da		14,463 $\pm$ 0.3 Da	A90C	12,679 $\pm$ 1.5 Da
$^{15}$ $\Delta\Delta$	12,649 Da	WT V52C	14,464 Da	$\Delta\Delta$ S129C	12,523 Da
	12,645 $\pm$ 1.9 Da		14,464 $\pm$ 0.1 Da		12,523 $\pm$ 0.3 Da
$\Delta$ N7	13,782 Da	WT A90C	14,492 Da	$^{15}$ $\Delta\Delta$	12,682 Da
	13,781 $\pm$ 0.2 Da		14,493 $\pm$ 0.09 Da	A140C	12,680 $\pm$ 0.9 Da
$\Delta$ C1	13,862 Da	$^{15}$ WT	14,658 Da	$^{15}$ P1P2-	14,143 Da
	13,861 $\pm$ 0.6 Da	A90C	14,656 $\pm$ 0.9 Da	SG A90C	14,137 $\pm$ 0.8 Da

### 3.4 N-terminal motif controls $\alpha$ Syn aggregation *in vitro*

To determine the role of the identified N-terminal-aggregation-prone regions N7, P1 and P2 in aggregation, ThT fluorescence assays were performed to measure the aggregation kinetics for  $\alpha$ Syn WT,  $\Delta$ N7 (residues 2-7 deleted),  $\Delta$ P1 (residues 36-42 deleted),  $\Delta$ P2 (residues 45-57 deleted), and  $\Delta\Delta$  (residues 36-42 and 45-57 deleted) (**Figure 3.3 A-E**, **Figure 3.4 A**, **Table 3.2**). Fibril growth was analysed in 20 mM Tris HCl, pH 7.5, mimicking a cytosolic environment or 20 mM sodium acetate, pH 4.5, similar to conditions found in lysosomes, each with 200 mM NaCl. Aggregation was followed for 100 h at 37 °C under shaking conditions (600 rpm) in a plate reader (FLUOstar Omega plate reader (BMG Labtech)), before determining the amount of pelletable material by centrifugation and SDS-PAGE gel analysis (**Table 3.2**), and end products were imaged by TEM (**Figure 3.3 F-I**, **Figure 3.4 B**). Detailed experimental protocols are given in Sections 2.8.2.1, 2.8.4, and 2.12, respectively.

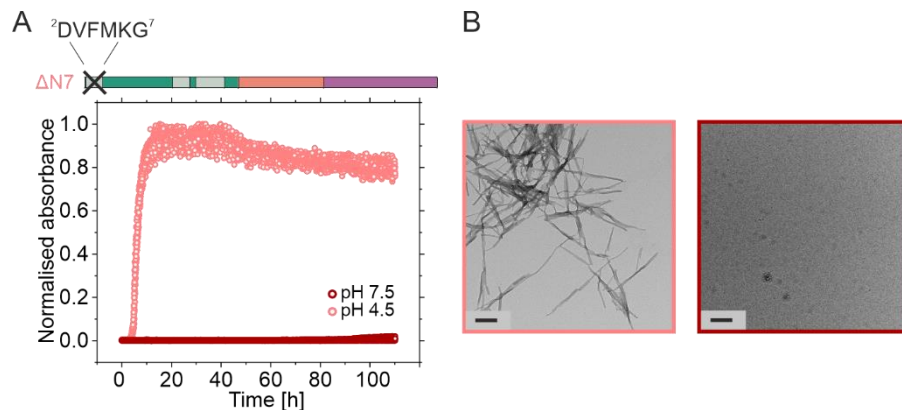


**Figure 3.3: The aggregation kinetics of 100  $\mu$ M WT  $\alpha$ Syn (B);  $\Delta$ P1 (C);  $\Delta$ P2 (D) or  $\Delta\Delta$  variants (E).**  $\alpha$ Syn schematics are shown in (A) with the N-terminal region coloured in green, NAC coloured in red, C-terminal region coloured in purple and C1, P1 and P2 coloured in grey. Sequences of the motifs of interest are highlighted and deleted motifs are indicated with black crosses. Dark and light colours denote incubations carried out at pH 7.5 (20 mM Tris HCl, 200 mM NaCl) or pH 4.5 (20 mM sodium acetate, 200 mM NaCl), respectively. All experiments were carried out at 37 °C with agitation at 600 rpm and measured at least in triplicates. **F-I** show negative stain TEM images of ThT end-point samples after 100 h incubation, scale bar 200 nm. Lag times, elongation rates and % pellet were determined and are listed in **Table 3.2**.

In general, at acidic pH (pH 4.5), fibril growth was observed to be faster for all  $\alpha$ Syn variants compared with neutral pH, as has been described for  $\alpha$ Syn WT before in the literature<sup>133</sup>. The lag time for  $\alpha$ Syn WT increases from 11.4 h  $\pm$  1.9 h to 45.3  $\pm$  4.2 h when changing the pH conditions from 4.5 to 7.5. Similar kinetics were observed for  $\Delta$ P2 (lag time 14.2  $\pm$  0.2 h and 45.4  $\pm$  16.3 h for pH 4.5 and 7.5, respectively) (**Table 3.2**). Amyloid formation for  $\Delta$ N7 and  $\Delta$ P1 is highly pH dependent, with slightly faster aggregation



kinetics compared with  $\alpha$ Syn WT at acidic pH (lag time  $3.9 \pm 0.2$  h and  $19.5 \pm 3.0$  h, respectively), but no fibril formation at pH 7.5 for at least 100 h (**Figure 3.3 C and Figure 3.4 A**). Interestingly,  $\Delta\Delta$ , which lacks 20 residues in the N-terminal region (P1 + P2) did not form fibrils for 100 h at both acidic and neutral pH (**Figure 3.3 E**). The presence or absence of fibrils was confirmed by TEM (**Figure 3.3 F-I, Figure 3.4 B**) and pelleting experiments (**Table 3.2**). These experiments suggest a dominant role of N7 and P1 at neutral pH in controlling fibril formation of the intrinsically disordered protein  $\alpha$ Syn with a synergistic effect at pH 4.5 for P1 and P2.



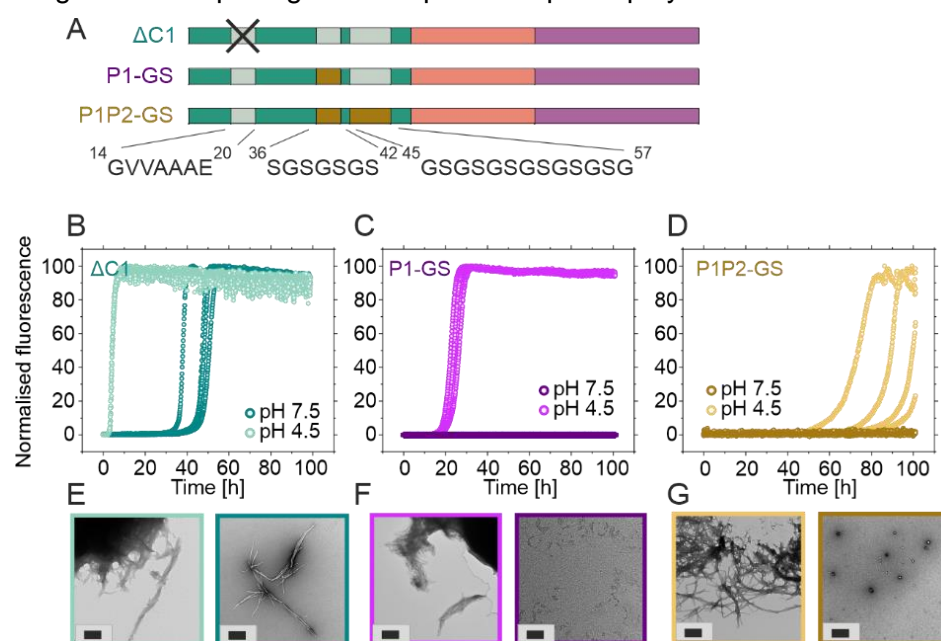
**Figure 3.4: The aggregation kinetics of 100  $\mu$ M  $\Delta$ N7 (A).** Schematic of  $\Delta$ N7 construct with the N-terminal region coloured in green, NAC coloured in red, C-terminal region coloured in purple and N7, P1 and P2 coloured in grey. Dark and light colours denote incubations carried out at pH 7.5 (20 mM Tris HCl, 200 mM NaCl) or pH 4.5 (20 mM sodium acetate, 200 mM NaCl), respectively. All experiments were carried out at 37 °C with agitation at 600 rpm and measured at least in triplicate. (B) show TEM images of ThT end-point samples after 100 h incubation, scale bar 200 nm. Lag times, elongation rates and % pellet were determined and are listed in **Table 3.2**.

As the extreme N-terminal region of  $\alpha$ Syn has been studied and found to be important before in various contexts including *de novo* growth<sup>345,411,413</sup>, seeding reactions<sup>107,183</sup>, membrane binding<sup>185,272</sup>, and chaperone interaction<sup>175</sup>, the work described here was mainly focused on the P1 and P2 regions that received no or limited attention in the literature before and present a novel motif(s) controlling protein aggregation.

The N-terminal region of  $\alpha$ Syn includes five imperfect repetitive KTKEGV sequences (plus one in NAC) that have been shown before to be important for aggregation<sup>244,412</sup>. To evaluate the possibility that the effects of deleting P1 and/or P2 may result from changes in the spatial organisation of the repeats, a control variant  $\Delta$ C1 (deletion of residues 14-20) was designed and aggregation kinetics determined (**Figure 3.5 A,B**).  $\Delta$ C1 was picked to mimic the general characteristics of P1 as closely as possible: the deleted sequence was of the same length, seven residues, and similar positioning between imperfect repeats as P1. In contrast to  $\Delta$ P1,  $\Delta$ C1 showed no significant inhibitory effect on fibril growth at all tested conditions (pH 4.5 and 7.5) and rather accelerates aggregation compared with  $\alpha$ Syn WT. Fibrils were observed after 100 h via TEM, and

pelleting assays demonstrate most of the sample being insoluble at the end of the experiment (**Figure 3.5 B,E and Table 3.2**).

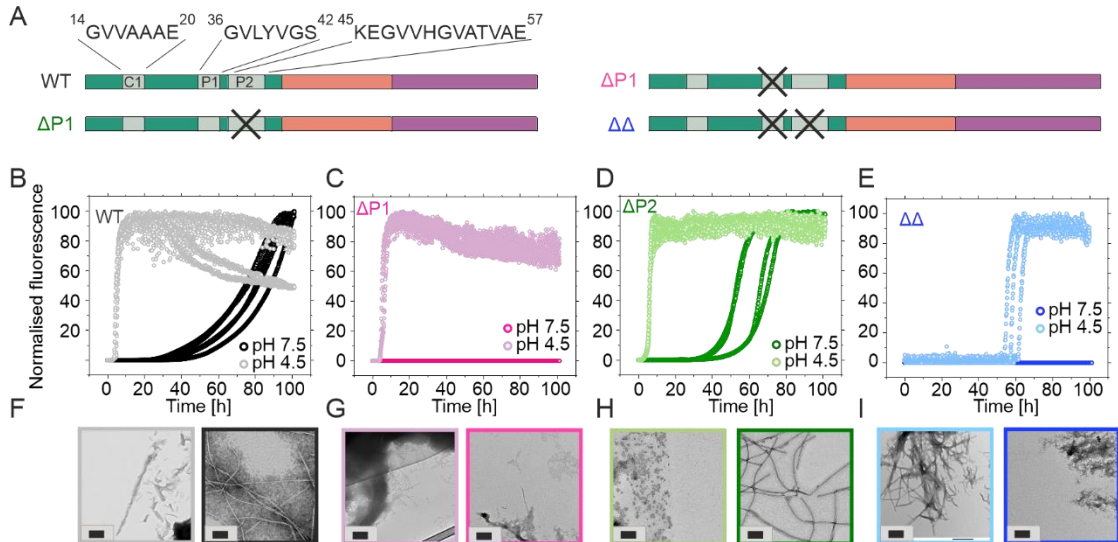
A second type of control variants was analysed in which P1 or P1/P2 were replaced with a glycine/serine linker (named P1-GS or P1P2-GS) preserving the spacing of the imperfect repeats (**Figure 3.5 A**). This experiment was designed to show whether altered aggregation in  $\Delta P1$  and  $\Delta\Delta$  is caused by the change in spacing, or whether it depends on precise protein sequences. For P1-GS similar aggregation kinetics were observed compared with  $\Delta P1$ : at pH 4.5 a lag time of  $21.2 \pm 1.4$  h was measured, no aggregation was observed at neutral pH for 100 h (**Figure 3.5 C,F, Table 3.2**). P1P2-GS did not/very slowly aggregate, comparable to the  $\Delta\Delta$  variant (**Figure 3.5 D,G, Table 3.2**). Together, these data show that P1 and P2 play a critical role in regulating the aggregation of  $\alpha$ Syn primarily controlled by the sequence of these motifs. Alterations in the length of the N-terminal region or the spacing of the imperfect repeats plays a minor role.



**Figure 3.5: Aggregation kinetics of 100  $\mu$ M  $\Delta C1$  (B), P1-GS (C), and P1P2-GS (D).** (A) schematic of  $\Delta C1$ ,  $\Delta\Delta$  and P1P2-GS  $\alpha$ Syn variants with the N-terminal region coloured in green, the NAC region coloured in red, the C-terminal region coloured in purple and the C1, P1 and P2 region coloured in grey. Dark and light colours denote incubations carried out at pH 7.5 (20 mM Tris HCl, 200 mM NaCl) or pH 4.5 (20 mM sodium acetate, 200 mM NaCl), respectively. All experiments were carried out at 37 °C with agitation at 600 rpm and measured at least in triplicates. (E-G) show TEM images of ThT end-point samples after 100 h incubation, scale bar 200 nm. Lag times, elongation rates and % pellet were determined and are listed in **Table 3.2**.

In addition to the aggregation kinetics measured with 200 mM NaCl, also the fibril growth kinetics in the presence of a lower salt concentration (20 mM) was evaluated (**Figure 3.6**). Salt is known to have a large effect on IDPs such as  $\alpha$ Syn, which present a high amount of charged residues<sup>189,230,417</sup>. Aggregation of  $\alpha$ Syn was affected by ionic strength at pH 4.5, with assembly into amyloid occurring more rapidly at low (20 mM added NaCl) compared with high (200 mM added NaCl) salt for all tested protein variants (WT,  $\Delta P1$ ,

$\Delta$ P2,  $\Delta\Delta$ ). Only marginal changes were observed at neutral pH between different ionic strengths (Table 3.2). As an increased salt concentration (slowing down aggregation at pH 4.5) would be expected to disrupt electrostatic intra- and inter-molecular interactions<sup>417</sup>, specific electrostatic interactions, rather than hydrophobic contacts, might accelerate aggregation and/or self-assembly into amyloid structures.



**Figure 3.6: The aggregation kinetics of 100  $\mu$ M WT  $\alpha$ Syn (B);  $\Delta$ P1 (C);  $\Delta$ P2 (D) or  $\Delta\Delta$  variants (E) at low salt concentration (20 mM NaCl).  $\alpha$ Syn schematics are shown in (A) with the N-terminal region coloured in green, NAC coloured in red, C-terminal region coloured in purple and C1, P1 and P2 coloured in grey. Sequences of the motifs of interested are highlighted and deleted motifs are indicated with black crosses. Dark and light colours denote incubations carried out at pH 7.5 (20 mM Tris HCl, 20 mM NaCl) or pH 4.5 (20 mM sodium acetate, 200 mM NaCl), respectively. All experiments were carried out at 37 °C with agitation at 600 rpm and measured at least in triplicates. F-I show TEM images of ThT end-point samples after 100 h incubation, scale bar 200 nm. Lag times, elongation rates and % pellet were determined and are listed in Table 3.2.**

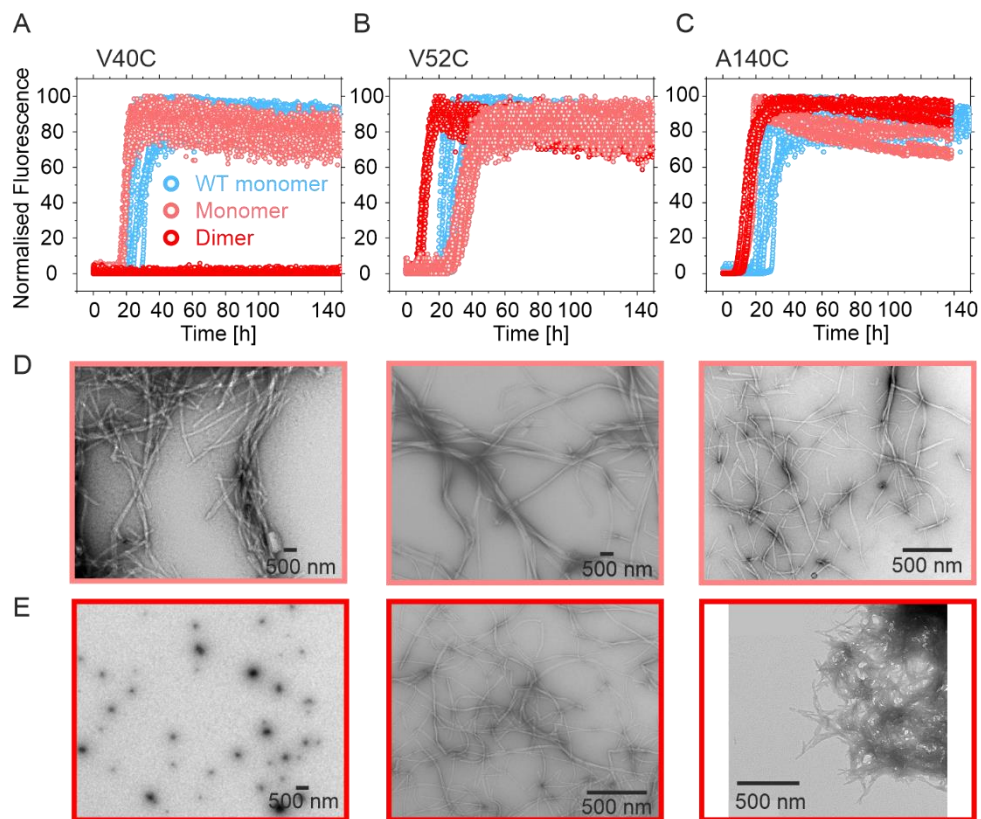
**Table 3.2: Aggregation kinetics of  $\alpha$ Syn variants. Lag times, elongation rates,  $t_{50}$  and % pellet for  $\alpha$ Syn variants.** The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 100 h incubation is indicated by “-“. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in  $10^4$ [RFU/h]. Assays were performed in 20 mM sodium acetate, pH 4.5 or 20 mM Tris-HCl, pH 7.5 with 200 mM NaCl unless stated otherwise (low salt corresponds with 20 mM NaCl).

	$\alpha$ WT, pH 4.5	$\alpha$ WT, pH 7.5	$\Delta$ P1, pH 4.5	$\Delta$ P1, pH 7.5
Lag time [h]	11.4 $\pm$ 1.9	45.3 $\pm$ 4.2	19.5 $\pm$ 3.0	-
Elongation rate*	3.3 $\pm$ 0.2	0.4 $\pm$ 0.02	1.7 $\pm$ 0.3	-
$t_{50}$ [h]	13.7 $\pm$ 1.5	55.3 $\pm$ 9.8	21.3 $\pm$ 1.4	-
% pellet	90	100	100	0
	$\Delta$ P2, pH 4.5	$\Delta$ P2, pH 7.5	$\Delta\Delta$ , pH 4.5	$\Delta\Delta$ , pH 7.5
Lag time [h]	14.2 $\pm$ 0.5	45.4 $\pm$ 16.3	-	-
Elongation rate*	0.7 $\pm$ 0.2	0.6 $\pm$ 0.07	-	-
$t_{50}$ [h]	28.7 $\pm$ 1.8	57.3 $\pm$ 9.4	-	-
% pellet	100	100	20	15
	$\Delta$ N7, pH 4.5	$\Delta$ N7, pH 7.5	$\Delta$ C1, pH 4.5	$\Delta$ C1, pH 7.5
Lag time [h]	3.9 $\pm$ 0.2	-	3.1 $\pm$ 0.1	43.2 $\pm$ 4.3
Elongation rate*	4.2 $\pm$ 0.3	-	3.8 $\pm$ 0.1	2.0 $\pm$ 0.5
$t_{50}$ [h]	6.5 $\pm$ 0.9	-	3.8 $\pm$ 0.2	49.1 $\pm$ 3.1
% pellet	85	10	90	70
	P1-GS, pH 4.5	P1-GS, pH 7.5	P1P2-GS, pH 4.5	P1P2-GS, pH 7.5
Lag time [h]	21.2 $\pm$ 1.4	-	82.8 $\pm$ 12.7	-
Elongation rate*	1.3 $\pm$ 0.2	-	0.01 $\pm$ 0.001	-
$t_{50}$ [h]	29.1 $\pm$ 10.1	-	88.1 $\pm$ 9.9	-
% pellet	60	5	20	0
	$\alpha$ WT, pH 4.5, low NaCl	$\alpha$ WT, pH 7.5, low NaCl	$\Delta$ P1, pH 4.5, low NaCl	$\Delta$ P1, pH 7.5, low NaCl
Lag time [h]	4.0 $\pm$ 0.5	49.7 $\pm$ 7.2	4.1 $\pm$ 0.5	-
Elongation rate*	3.5 $\pm$ 0.5	0.2 $\pm$ 0.02	3.0 $\pm$ 0.9	-
$t_{50}$ [h]	6.5 $\pm$ 0.8	71.9 $\pm$ 6.3	5.9 $\pm$ 0.5	-
% pellet	90	100	100	0
	$\Delta$ P2, pH 4.5, low NaCl	$\Delta$ P2, pH 7.5, low NaCl	$\Delta\Delta$ , pH 4.5, low NaCl	$\Delta\Delta$ , pH 7.5, low NaCl
Lag time [h]	4.8 $\pm$ 0.4	36.2 $\pm$ 1.7	57.3 $\pm$ 3.6	-
Elongation rate*	4.4 $\pm$ 0.8	0.7 $\pm$ 0.2	2.3 $\pm$ 0.6	-
$t_{50}$ [h]	7.1 $\pm$ 1.0	60.4 $\pm$ 8.2	62.2 $\pm$ 6.1	-
% pellet	100	80	90	20

In a next step, the importance of P1 and P2 was examined by determining the aggregation kinetics of disulfide-crosslinked dimers of  $\alpha$ Syn (**Figure 3.7**). Dimers were created using cysteine substitutions at positions V40C (in P1), V52C (in P2), and V140C (at the C-terminus) and samples were incubated for 2 h without reducing agent. Separation of dimers from monomer left in the reaction was performed by analytical and preparative SEC and a successful dimerization demonstrated using SEC-MALS experiments (**Figure 3.8 A-D**) (see Section 2.7.5 for detailed methods).

Interestingly, whilst all  $\alpha$ Syn variants (V40C, V52C, and A140C) aggregate similarly to  $\alpha$ Syn WT in their monomeric states (when adding 2 mM dithiothreitol (DTT) to reduce the disulphide bond), dimerization of V40C inhibited fibril formation for over 140 h (**Figure 3.7 A**). Dimers of V52C and A140C did not significantly affect the aggregation kinetics compared with their respective monomer (**Figure 3.7 B,C**). These findings were validated by TEM and pelleting analysis (**Figure 3.7 D,E and Table 3.3**). This

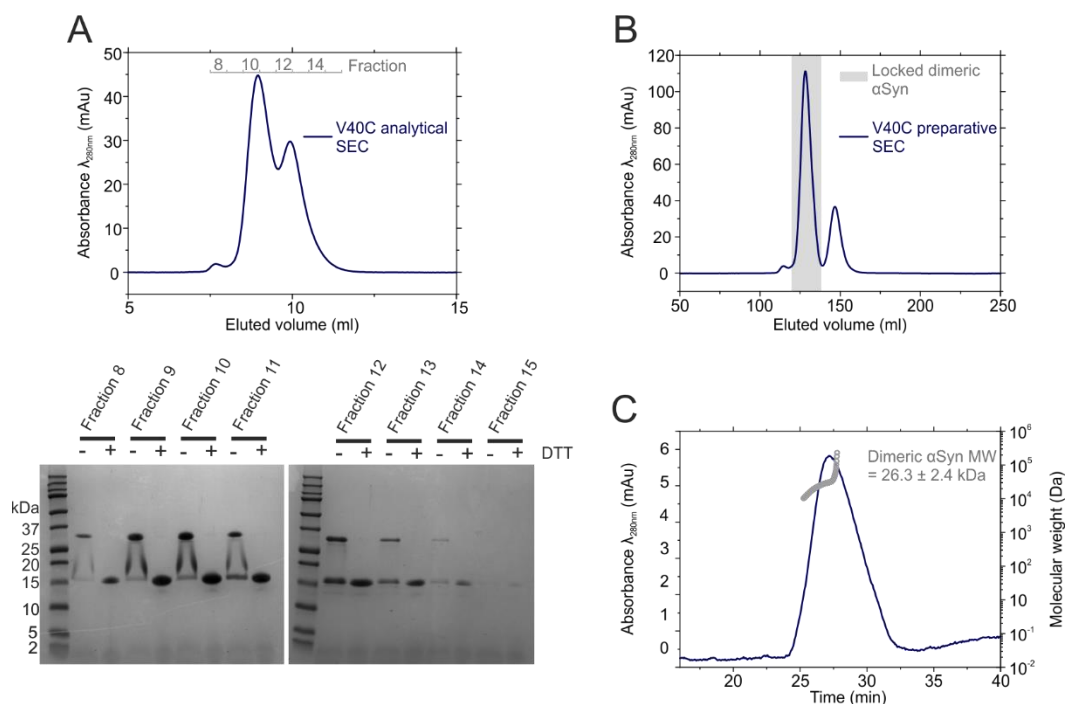
observation further supports the hypothesis that P1 is important for fibril formation and might contribute to key interactions driving aggregation.



**Figure 3.7: Aggregation kinetics and negative-stain TEM images of endpoint (140 h) aggregates of 100  $\mu$ M V40C (A), V52C (B) or A140C (C) monomers or dimers.** Incubations of monomeric or disulfide-locked dimers of  $\alpha$ Syn are shown in light and dark red, respectively. Aggregation kinetics of  $\alpha$ Syn WT are shown in blue. The same data for the WT  $\alpha$ Syn are shown overlaid for all three variants. All experiments were measured at least in triplicate. TEMs in (D) show endpoint images of reduced samples, and (E) show endpoint images of disulfide-bonded dimers. Each image was collected from a representative sample for each condition. All assays were carried out in 20 mM sodium acetate buffer containing 200 mM NaCl, pH 4.5 (including 2 mM DTT for reduced samples) at 37  $^{\circ}$ C with agitation at 600 r.p.m. This experiment was performed by Ciaran Doherty, University of Leeds.

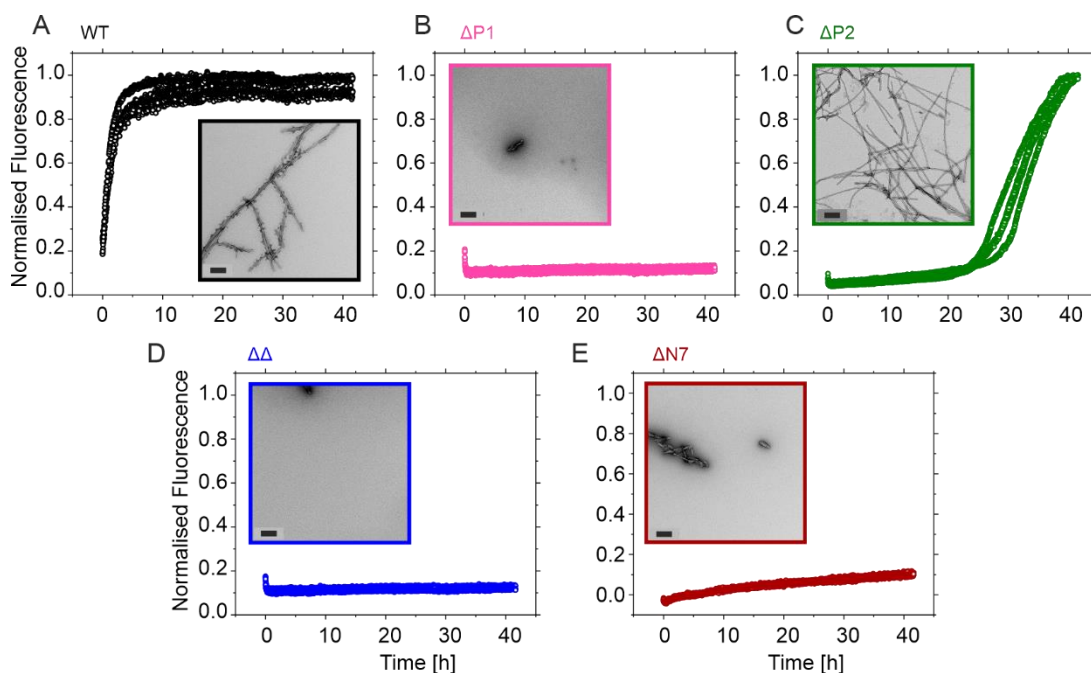
**Table 3.3: Aggregation kinetics for dimer locked proteins.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\alpha$ Syn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 140 h incubation is indicated by “-”. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in  $10^*$ [RFU/h].

	$\alpha$ V40C, monomer	$\alpha$ V40C, dimer	$\alpha$ V52C, monomer	$\alpha$ V52C, dimer
Lag time [h]	17.1 $\pm$ 0.9	-	6.9 $\pm$ 1.3	27.9 $\pm$ 2.7
Elongation rate*	1.8 $\pm$ 0.3	-	1.0 $\pm$ 0.04	0.6 $\pm$ 0.2
$t_{50}$ [h]	21.2 $\pm$ 0.6	-	11.8 $\pm$ 0.9	36.7 $\pm$ 3.1
% pellet	90	5	90	80
	$\alpha$ A140C, monomer	$\alpha$ A140C, dimer	$\alpha$ WT, monomer	
Lag time [h]	12.7 $\pm$ 1.1	11.2 $\pm$ 2.0	21.1 $\pm$ 1.2	
Elongation rate*	1.0 $\pm$ 0.04	0.9 $\pm$ 0.02	1.5 $\pm$ 0.2	
$t_{50}$ [h]	15.3 $\pm$ 1.0	14.0 $\pm$ 1.2	26.4 $\pm$ 1.0	
% pellet	95	85	85	



**Figure 3.8: Preparation of disulfide locked  $\alpha$ Syn variants.** (A) Analytical SEC of  $\alpha$ Syn V40C incubated at room temperature for 2 h in 100 mM Tris HCl, pH 8.4 in the absence of reducing agent. The fraction numbers are indicated in grey. SDS-PAGE of the corresponding fractions in the absence or presence of 2 mM DTT are shown below. (B) Preparative SEC of  $\alpha$ Syn after incubation for 2 h in 100 mM Tris HCl, pH 8.4. The dimeric species taken forward for experiments is highlighted in grey. (C) SEC-MALS of dimeric  $\alpha$ Syn V40C, a mass of  $26.3 \pm 2.4$  kDa was recorded corresponding clearly to a theoretical dimeric mass of 28.9 kDa. This experiment was performed by Ciaran Doherty, University of Leeds. Detailed description of experimental setups can be found in Sections 2.7.4 and 2.7.5.

To test the capability of the newly designed deletion variants  $\Delta$ P1,  $\Delta$ P2,  $\Delta$  $\Delta$  and  $\Delta$ N7 to elongate  $\alpha$ Syn WT fibrils, a seeding assay with 10 % (mol/mol) pre-formed  $\alpha$ Syn WT seeds was performed (see Section 2.8.2.2). Whilst  $\alpha$ Syn WT could be self-seeded (**Figure 3.9 A**), cross-seeding  $\Delta$ N7,  $\Delta$ P1 and  $\Delta$  $\Delta$  monomer with  $\alpha$ Syn WT seeds was inhibited at pH 7.5 (**Figure 3.9 B,D,E**).  $\Delta$ P2 formed fibrils slowly (**Figure 3.9 C**, **Table 3.4**). The presence of fibrils for WT and  $\Delta$ P2 was confirmed by TEM (**Figure 3.9, inserts**). The data indicate a structural incompatibility of N7, P1 and P2 deletion variants with fibril seeds formed from the WT protein demonstrating elongation to be the inhibited mechanistic step. In the case of  $\Delta$ N7, the lack of aggregation could also be explained by crucial interactions between monomer and fibril being eliminated when deleting the first seven residues as residues 1-10 are the main interaction sites between these two species as demonstrated recently using NMR and ThT aggregation assays<sup>107,183</sup>.



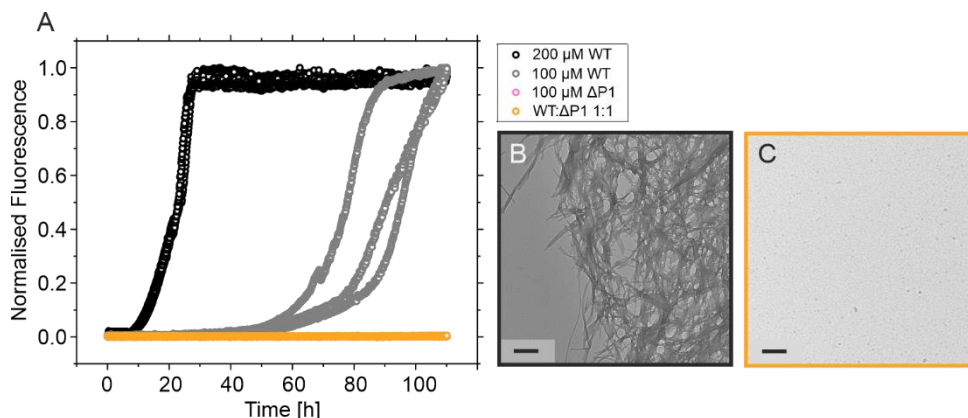
**Figure 3.9: Cross-seeding  $\alpha$ Syn variants using seeds created from  $\alpha$ Syn WT.** ThT fluorescence assays of  $\alpha$ Syn variants (100  $\mu$ M) WT (A),  $\Delta$ P1 (B),  $\Delta$ P2 (C),  $\Delta\Delta$  (D), and  $\Delta$ N7 (E) seeded with 10 % (mol/mol) WT  $\alpha$ Syn fibril seeds formed at pH 7.5. Seeding assays were performed at pH 7.5, 20 mM NaCl, 37  $^{\circ}$ C, quiescent. End point (42 h) TEM images of representative samples of fibrils from the seeding experiments are shown in inserts. Scale bars, 200 nm.

**Table 3.4: Aggregation kinetics of seeding reaction with 10 %  $\alpha$ Syn WT seeds.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\alpha$ Syn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 42 h incubation is indicated by “-“. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in  $10^*$ [RFU/h].

	$\alpha$ WT	$\Delta$ P1	$\Delta$ P2
Lag time [h]	0	-	$24.5 \pm 1.8$
Elongation rate*	$2.4 \pm 0.4$	-	$7.2 \pm 0.7$
$t_{50}$ [h]	$1.6 \pm 0.1$	-	$38.3 \pm 6.5$
% pellet	85	0	60
	$\Delta\Delta$	$\Delta$ N7	
Lag time [h]	-	-	
Elongation rate*	-	-	
$t_{50}$ [h]	-	-	
% pellet	0	5	

We next examined the effect of monomeric  $\alpha$ Syn WT and  $\Delta$ P1 being co-incubated in the same reaction to investigate whether  $\Delta$ P1 can affect the aggregation of  $\alpha$ Syn WT *in trans*.  $\alpha$ Syn WT and  $\Delta$ P1 were mixed in a 1:1 ratio (mol/mol) and fibril formation was followed by a ThT fluorescence assay (Figure 3.10 A). Whilst  $\alpha$ Syn WT aggregates with a lag time of  $62.9 \pm 5.8$  h at 100  $\mu$ M and  $9.8 \pm 0.3$  h at 200  $\mu$ M, the addition of 100  $\mu$ M  $\Delta$ P1 to 100  $\mu$ M  $\alpha$ Syn WT inhibits aggregation for at least 110 h (Figure 3.10 A). The presence of fibrils for  $\alpha$ Syn WT alone and the absence of amyloid structures when incubated with  $\Delta$ P1 was confirmed by TEM (Figure 3.3 F,G and Figure 3.10 B,C). This

experiment indicates that the WT and  $\Delta P1$  proteins are interacting and that this prevents fibril formation (at least within the experimental time scale). Whether this happens by interactions of monomers or oligomeric states cannot be answered and further experiments are required to gain better insight into the molecular mechanism by which  $\Delta P1$  inhibits aggregation of the WT protein.



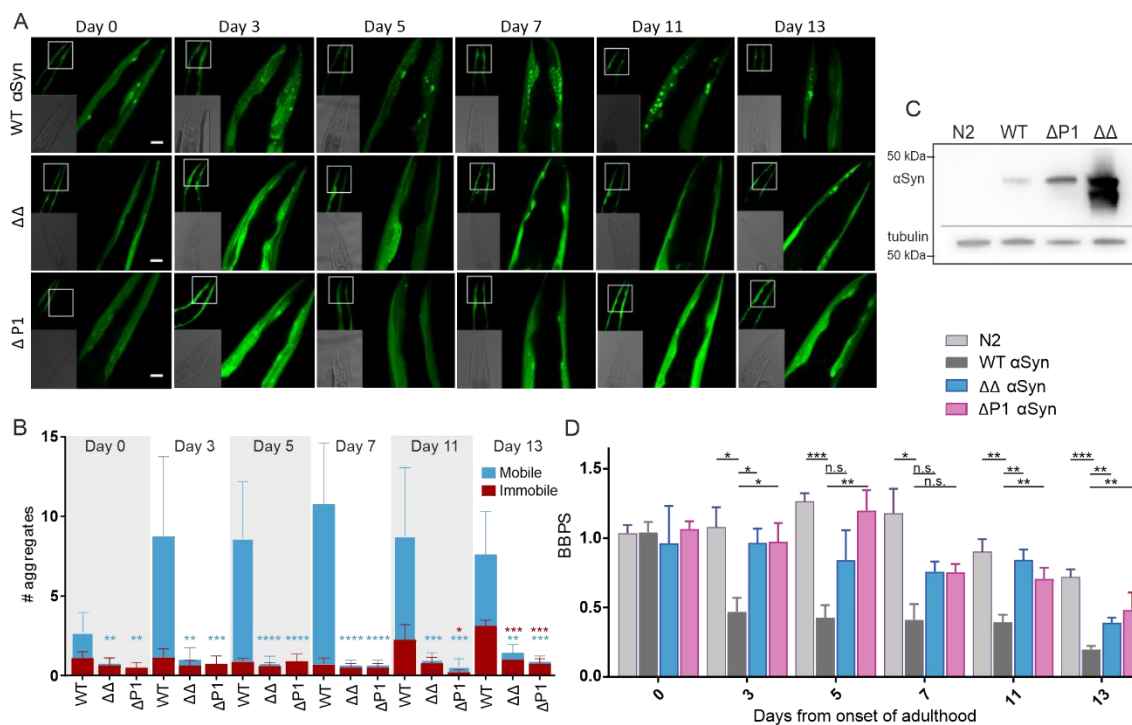
**Figure 3.10: Aggregation kinetics of co-incubated  $\alpha\text{Syn}$  WT and  $\Delta P1$  monomer.** (A) ThT assay of  $\alpha\text{Syn}$  WT alone (black: 200  $\mu\text{M}$ , grey 100  $\mu\text{M}$ ),  $\Delta P1$  alone (pink, behind orange line) and  $\alpha\text{Syn}$  WT and  $\Delta P1$  mixed in a 1:1 ratio (100 $\mu\text{M}$ :100  $\mu\text{M}$ ) (orange). The experiment was performed in 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 37  $^{\circ}\text{C}$ , 600 rpm. Note that the pink data are below the orange data points. End-points were imaged by negative stain TEM for  $\alpha\text{Syn}$  WT 200  $\mu\text{M}$  (B) and  $\alpha\text{Syn}$  WT: $\Delta P1$  (C). TEM images for  $\alpha\text{Syn}$  WT 100  $\mu\text{M}$  and  $\Delta P1$  100  $\mu\text{M}$  are shown in Figure 3.3 F,G. Scale bar 200 nm.

### 3.5 The N-terminal motifs control aggregation *in vivo*

After demonstrating a dramatic effect on aggregation by removing P1 and P2 *in vitro*, the effect of the deletion variants was evaluated *in vivo* using the model organism *C. elegans*.  $\alpha\text{Syn}$  WT,  $\Delta P1$  and  $\Delta\Delta$  were fused C-terminally to YFP and expressed in the body wall muscle cells of the animals as described before for the WT protein<sup>83</sup>. As expected, the WT  $\alpha\text{Syn}$ -expressing animals showed puncta formation/ inclusions in L4 larvae (day 0), increasing in number when aging probably due to a declined protease network<sup>418,419</sup>, and reaching a plateau from day 3 to day 13 of adulthood (Figure 3.11 A,B). Remarkably, in contrast with this, animals expressing  $\Delta P1::\text{YFP}$  or  $\Delta\Delta::\text{YFP}$  formed only a few, if any, visible aggregates throughout aging (Figure 3.11 A,B), even though the expression levels of these proteins were higher than in WT  $\alpha\text{Syn}$  nematodes as detected by a Western Blot analysis (Figure 3.11 C) (Section 2.7.3 for detailed methods). The reason for the differences in expression levels between WT  $\alpha\text{Syn}$  (obtained from the Nollen lab, University of Groningen) and the  $\Delta P1$  and  $\Delta\Delta$  constructs (commercially obtained) could (in part) arise from the method of gene transformation (gene bombardment and microinjection, respectively (Section 2.15.3.1) which results in different copy numbers of each gene in the different organisms. Notably, in contrast to WT  $\alpha\text{Syn}::\text{YFP}$ , the total number of  $\Delta P1::\text{YFP}$  or  $\Delta\Delta::\text{YFP}$  foci did not increase during



aging, with only few aggregates observed even at day 13 (**Figure 3.11 B**). The percentage of immobile WT  $\alpha$ Syn::YFP aggregates, measured by fluorescence recovery after photobleaching (FRAP) experiments (see Section 2.15.3.3 for detailed methods), increased approximately four-fold from day 7 to day 13 of adult-hood (**Figure 3.11 B**). In comparison,  $\Delta$ P1 and  $\Delta\Delta$   $\alpha$ Syn::YFP formed few aggregates (one or two foci per animal) that were immobile (**Figure 3.11 B**). Since  $\alpha$ Syn was expressed in the muscle cells of the nematodes, the phenotypic effect on the worms could be analysed by testing the motility of the animals by measuring body bends per second (BBPS) (see Section 2.15.3.3). *C. elegans* expressing WT  $\alpha$ Syn::YFP showed an age-dependent decline of motility between days 3 and 13 compared with the control strain N2 (**Figure 3.11 D**). In contrast, motility remained similar to that of healthy N2 animals between days 0 and 7 (**Figure 3.11 D**), with slightly reduced (two-fold) thrashing rates observed in  $\Delta\Delta$ ::YFP and  $\Delta$ P1::YFP at day 13 of adulthood (**Figure 3.11 D**). Thus, deletion of P1 or both P1 and P2 prevents age-dependent aggregation of  $\alpha$ Syn *in vivo* and suppresses aggregation-induced proteotoxicity. Note that these experiments were performed and data analysed by Sarah Good and Jemma Makepeace, University of Leeds.



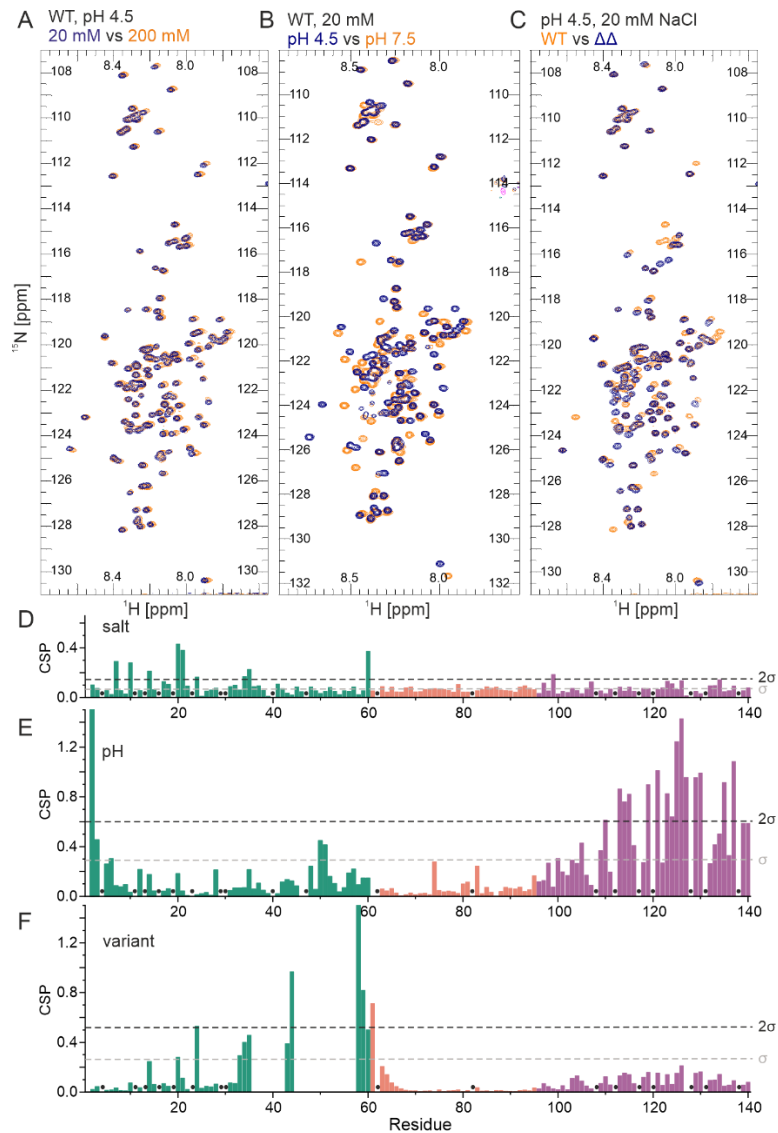
**Figure 3.11: Deletion of P1 or both P1 and P2 in *C. elegans* expressing  $\alpha$ Syn::YFP suppresses aggregation and proteotoxicity.** (A) Confocal microscopy images showing the head region of transgenic *C. elegans* expressing WT  $\alpha$ Syn,  $\Delta\Delta$  or  $\Delta$ P1 tagged C-terminally to YFP in the body wall muscle during aging (day 0 to day 13 of adulthood). Scale bars, 10  $\mu$ m. (B) Numbers of mobile and immobile inclusions larger than  $\sim 2 \mu\text{m}^2$  per animal between the tip of the head and pharyngeal bulb during aging, determined by FRAP. Data shown are the mean and s.e.m. for three independent experiments (biological replicates); in each experiment, ten worms ( $n = 10$ ) were assessed for each time point. Blue asterisks indicate significance between the number of mobile aggregates of animals expressing WT  $\alpha$ Syn or the  $\Delta$ P1 or  $\Delta\Delta$  variants. Red asterisks indicate significance between the number of immobile aggregates exhibited in animals expressing WT  $\alpha$ Syn compared with mutant animals. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ . ANOVA was used in all cases. (C) Western blot analysis of protein extracts isolated from N2, WT  $\alpha$ Syn::YFP,  $\Delta$ P1::YFP and  $\Delta\Delta$ ::YFP animals using an anti- $\alpha$ Syn antibody. Tubulin was used as a loading control. The loading control (anti-tubulin) was run on a different gel or membrane loaded with the same protein sample and treated and analysed in the same manner. The images are cropped, showing all relevant bands. (D) Number of body bends per second (BBPS) of N2, WT  $\alpha$ Syn::YFP,  $\Delta$ P1::YFP and  $\Delta\Delta$ ::YFP animals from day 0 (L4 stage) through day 13 of adulthood. Data shown are mean and s.e.m. for three independent experiments; in each experiment, ten worms were assessed for each time point.  $n = 10$  for each experiment, and error bars represent s.e.m. of three biological replicates. n.s. = not significant; \*\*\* $P < 0.001$ ; \*\* $P < 0.01$ ; \* $P < 0.05$  ANOVA was used. Worm images, FRAP experiments and motility assays were performed by Sarah Good and Jemma Makepeace supervised by Patricija van Oosten-Hawle, University of Leeds.

### 3.6 Assignment and Evaluation of NMR experiment conditions for analysis of $\alpha$ Syn variants

It is known, that IDPs contain some sequence-specific conformational preferences and their structure is not completely random<sup>88,337,420,421</sup>. Indeed, that is also the case for  $\alpha$ Syn, according to small angle X-ray scattering (SAXS) data, the actual radius of gyration is smaller ( $\sim 40$  Å) than the calculated radius for a 140-residue random coil ( $52$  Å)<sup>317</sup>. Beside the commonly known secondary structure elements when interacting with membranes ( $\alpha$ -helix)<sup>186,314</sup> or forming amyloid fibrils ( $\beta$ -sheets)<sup>96</sup>, recent work in the literature identified preferred conformations in monomeric  $\alpha$ Syn<sup>189,335</sup>. It was found, that the conformational ensemble of  $\alpha$ Syn is pH dependent<sup>335</sup>, which is in consensus with the different aggregation behaviour at acidic and neutral pH (**Figure 3.3 B** and ref<sup>422</sup>). To determine whether P1 and P2 affect the conformational properties of  $\alpha$ Syn monomers that alter their ability to assemble into amyloid, WT  $\alpha$ Syn and  $\Delta\Delta$  were examined using NMR paramagnetic relaxation enhancement (PRE) experiments (see Section 2.9.2). First, both proteins were expressed  $^{15}\text{N}/^{13}\text{C}$  labelled and purified (see Section 2.6.1) before resonances were assigned under conditions used in this study (20 mM sodium acetate, pH 4.5, 20 mM NaCl) (**Figure 3.12 A, B** for  $\alpha$ Syn WT and  $\Delta\Delta$ , respectively) by Roberto Maya-Martinez, University of Leeds (see Section 2.9.1). Importantly, residues M1, P108, P117, P120, P128 and P138 could not be assigned. Note that an acidic pH and a low salt concentration were chosen for these experiments as this is the condition with the most dramatic effects on aggregation ( $\alpha$ Syn WT aggregates within a few hours whilst  $\Delta\Delta$  does not form fibrils for over 100 h (**Figure 3.6 B,E**)). Further pH 4.5 can be found in lysosomes, and lysosomal entry of  $\alpha$ Syn is well known in the literature to be important in disease initiation spreading<sup>34,423</sup>.

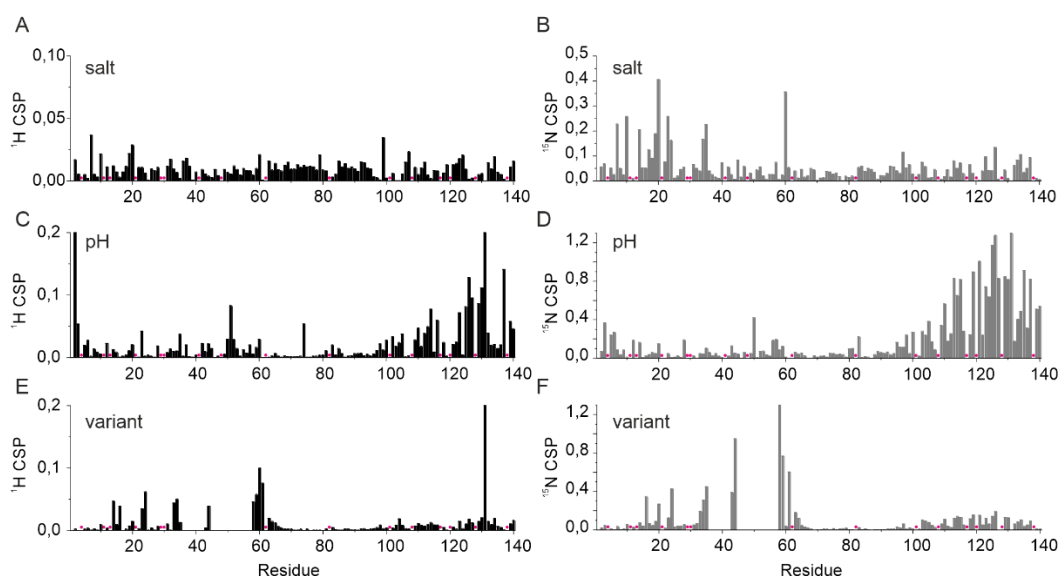


changing the ionic strength conditions shows only small HN-CSPs throughout the whole protein sequence with the largest changes in the N-terminal 25 residues (**Figure 3.13 A,D**), changing the pH has a significant effect on the N-terminus (residues 2-6) and the C-terminal region (residues 109-140) (**Figure 3.13 B,E**). Also HN-CSPs between  $\alpha$ Syn WT and  $\Delta\Delta$  were analysed indicating that, although having contrasting aggregation propensities, most changes are located close to the deleted P1- and P2-regions without significantly affecting NAC or C-terminal region (**Figure 3.13 C,F**).



**Figure 3.13: Comparison of chemical shifts for different buffer conditions and synuclein variants.** (A)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra for  $\alpha$ Syn WT A18C at high (200 mM) and low salt (20 mM), pH 4.5, 15 °C. (B)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra for  $\alpha$ Syn WT pH 4.5 and 7.5, 20 mM NaCl. (C)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra for  $\alpha$ Syn WT A18C vs  $\Delta\Delta$  A18C. Data were collected at pH 4.5, 200 mM NaCl, 15 °C. **D,E,F** show the HN-CSPs for the conditions described in **A,B,C** calculated using **Equation 2.2** (Section 2.9.4). The N-terminal region is coloured in green, NAC in red and the C-terminal region in purple. Dashed lines indicate the standard deviation  $\sigma$  (grey) and  $2\sigma$  (black).

Residues 4, 11, 13, 16, 19, 23, 29, 30, 41, 47, 62, 82, and 112 were not included because of overlapping peaks; residues 1, 108, 117, 120, 128, and P138 were not assigned, these residues are indicated with black dots.



**Figure 3.14: Comparison of separate  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts for different buffer conditions and synuclein variants shown in Figure 3.13. (A, B)  $^1\text{H}$  CSP and  $^{15}\text{N}$  CSP for  $\alpha\text{Syn}$  WT A18C at high (200 mM) and low salt (20 mM), pH 4.5, 15 °C. (C, D)  $^1\text{H}$  CSP and  $^{15}\text{N}$  CSP for  $\alpha\text{Syn}$  WT pH 4.5 and 7.5, 20 mM NaCl. (E, F)  $^1\text{H}$  CSP and  $^{15}\text{N}$  CSP for  $\alpha\text{Syn}$  WT A18C vs  $\Delta\Delta$  A18C. Residues 4, 11, 13, 16, 19, 23, 29, 30, 41, 47, 62, 82, and 112 were not included because of overlapping peaks; residues 1, 108, 117, 120, 128, and P138 were not assigned.**

## 3.7 Intra- and inter-molecular interactions involving N-terminal motifs promote aggregation

### 3.7.1 Intramolecular interactions with P1 and P2 determine aggregation

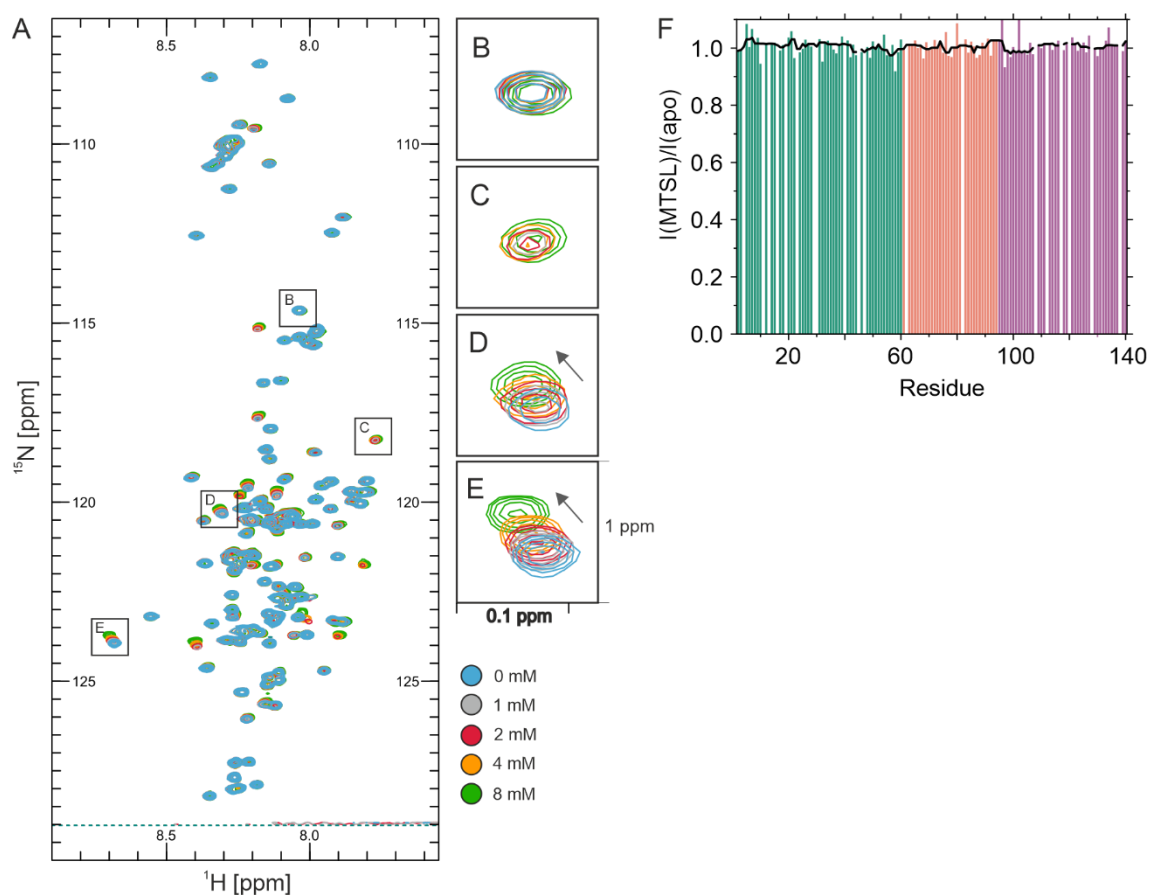
As mentioned above, NMR PRE experiments were next used to discover the transient conformation(s) of  $\alpha\text{Syn}$  that could be important in controlling its aggregation. This approach allows protein dynamics to be investigated in the  $\mu\text{s}$  timescale for rare (0.5-5 % population) and transient interactions with distances of  $\sim 20\text{-}40$  Å between a nucleus and spin label with an unpaired electron<sup>424,425</sup>. Previous studies used PRE NMR to investigate the intramolecular interactions of  $\alpha\text{Syn}$  WT at various pH values including pH 2.5, 3.0, 6.0, 7.4 and 7.5 and using different protein concentrations ranging from 100  $\mu\text{M}$  to 650  $\mu\text{M}$  at 15 °C<sup>321,325,331,335,426,427</sup> (see **Table 1.4**). Further, some of the familial PD mutations (A30P, A53T)<sup>428</sup> as well as  $\beta\text{Syn}$  and  $\gamma\text{Syn}$ <sup>333</sup> have been investigated using this experimental setup.

To analyse how the deletion of the P1 and P2 motifs affects the conformational properties of  $\alpha\text{Syn}$  monomers,  $^{15}\text{N}$ -labeled WT  $\alpha\text{Syn}$  or  $\Delta\Delta$  containing a single cysteine introduced at position 18 ( $\alpha\text{Syn}$  A18C), 90 ( $\alpha\text{Syn}$  A90C), or 140 ( $\alpha\text{Syn}$  A140C) were expressed and purified (**Table 3.1**). As  $\alpha\text{Syn}$  does not have a natural cysteine in its sequence, specific single residue labelling is possible at the desired positions. Each protein was covalently

labelled with the paramagnetic spin label S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl-methanesulfonylthioate (MTSL). For the experimental setup, the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of the  $\alpha\text{Syn}$  sample with covalently bound MTSL was collected (paramagnetic) before reducing the MTSL label by adding ascorbic acid, which eliminates the effect of the spin label (diamagnetic) (see Section 2.9.2.1 for experimental details). To identify the ideal ascorbic acid concentration to fully reduce MTSL and recover peak intensities, but to minimise chemical shifts caused by pH changes when adding the acid, a titration was performed observing the effect of adding ascorbic acid to  $\alpha\text{Syn}$  WT (**Figure 3.15 A-E**). Different cross-peaks are affected to different extents by the addition of ascorbic acid, however, for future experiments 2 mM ascorbic acid were used to measure the diamagnetic spectra (**Figure 3.16**).

Further, whether free MTSL in solution has an effect on the PRE NMR data was checked by mixing  $^{15}\text{N}$   $\alpha\text{Syn}$  WT (no spin label) with free MTSL and analysing the peak intensities of the resulting  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra (**Figure 3.15 F**). For this, no PRE effects could be seen indicating that potential free MTSL after the labelling process would not affect the measured intramolecular interactions between  $\alpha\text{Syn}$  monomers.

Note that these control experiments have been performed at pH 3, as they were done early on in this project, in which NMR PRE data were compared to literature results showing that the experimental setup is successful in principal before moving on to conditions relevant in this thesis. The pH swap should not have an effect on the outcome of the control experiments.



**Figure 3.15: Ascorbic acid titration to identify ideal conditions for reducing MTSL in PRE**

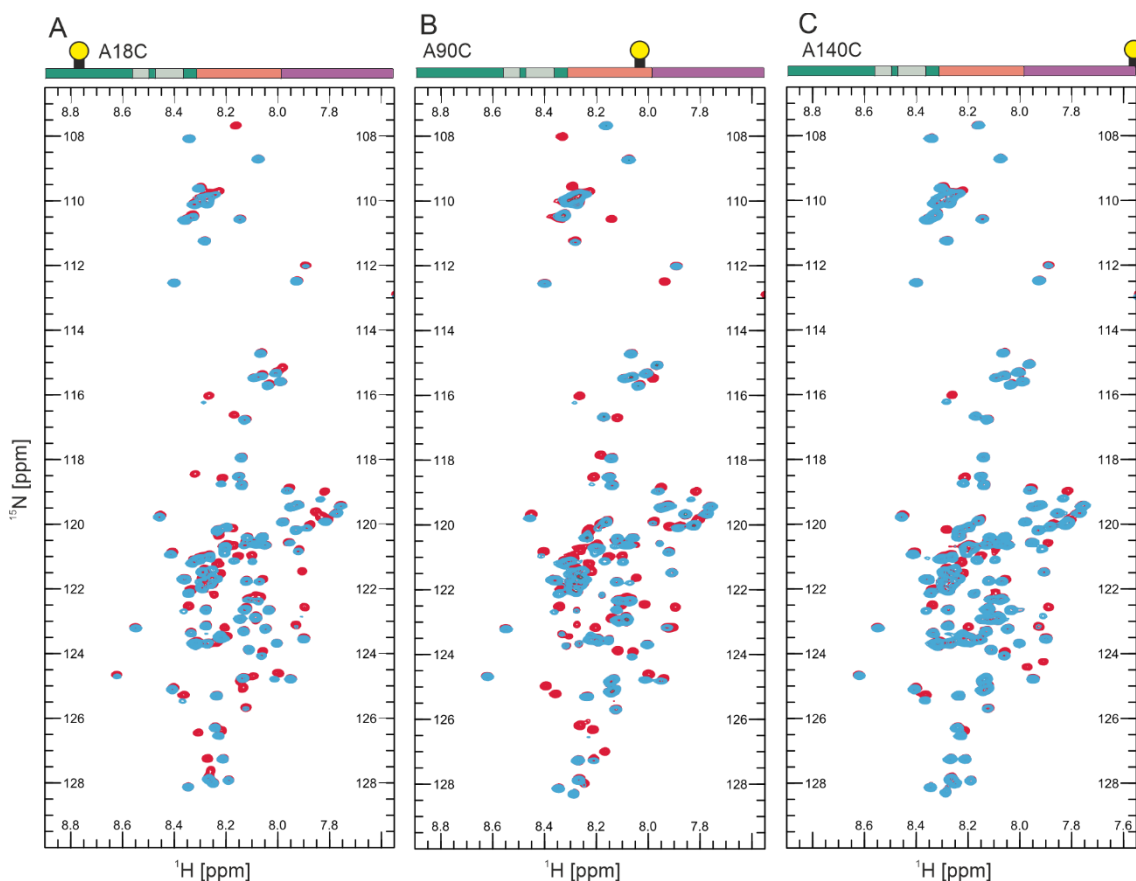
**experiments and effect of free MTSL. (A)**  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of MTSL-labelled  $\alpha\text{Syn}$  A18C

(20 mM sodium acetate buffer, pH 3, 100 mM NaCl, 15°C) titrated with ascorbic acid at concentrations ranging from 0 mM (paramagnetic, blue) to 8 mM (diamagnetic, green). **B-E:**

highlight different effects on the HSQC spectra. **(B)** no changes observable between different ascorbic acid concentrations. **(C)** Signal completely disappeared in paramagnetic spectra, increasing amounts of ascorbic acid increase diamagnetic signal. **(D, E)** Signal intensity stays constant with an increasing ascorbic acid concentration but the signal shows clear shifts (arrow). A 2 mM ascorbic acid concentration was used for all further experiments. **(F)** Control

experiment of free MTSL (100  $\mu\text{M}$ ) with 100  $\mu\text{M}$   $^{15}\text{N}$   $\alpha\text{Syn}$  WT (not spin labelled) in 20 mM sodium acetate buffer, pH 4.5, 20 mM NaCl to show that the free spin label does not have an effect on PRE data. Residues 4, 11, 13, 16, 19, 23, 29, 30, 41, 45, 47, 62, 82, and 112 were not included because of overlapping peaks; residues 1, 108, 117, 120, 128, and P138 were not assigned. The signal to noise level is  $>3.52$  (dia) and  $>2.97$  (para).

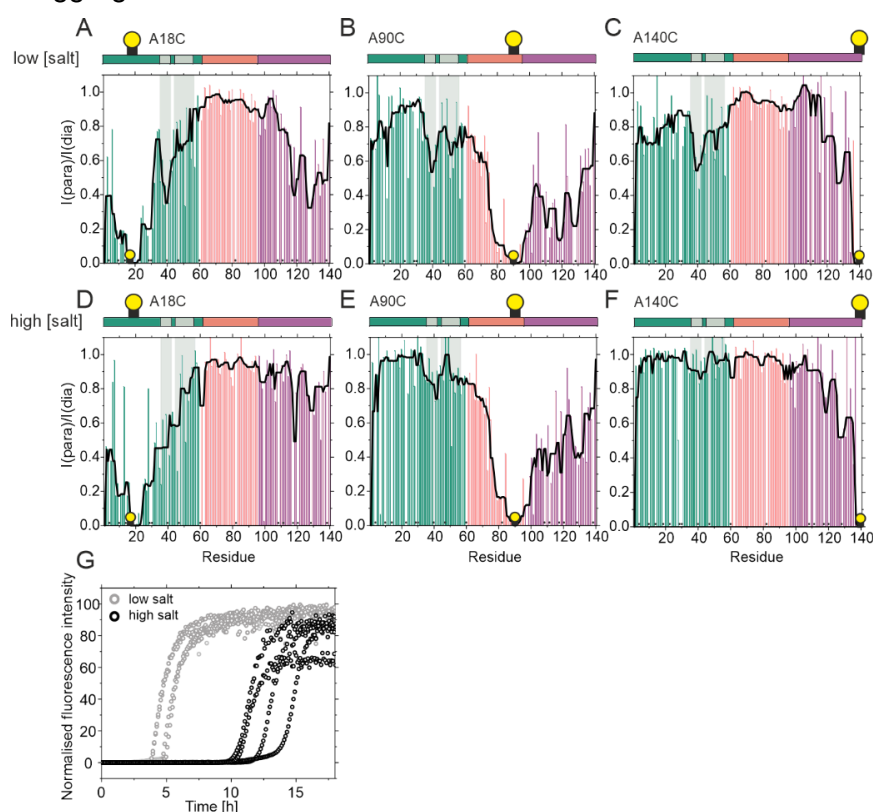




**Figure 3.16: Overlaid paramagnetic (blue) and diamagnetic (red) spectra for WT  $\alpha$ Syn labelled at positions A18C (A), A90C (B) or A140C (C).** Schematics are shown above each spectrum with the N-terminal (green), NAC (red) and C-terminal (purple) regions highlighted. The location of the spin label is indicated by a yellow circle. Note that small chemical shift changes are observed upon reduction with ascorbic acid, which can be attributed to small changes in pH (see **Figure 3.15 A-E**).

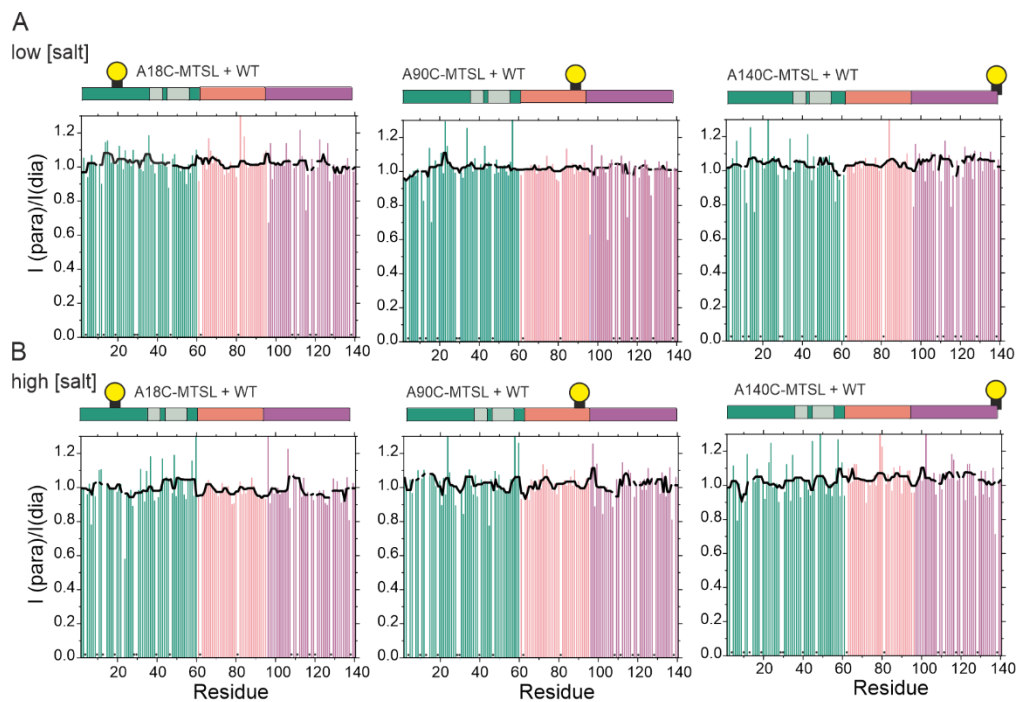
The intramolecular interactions of  $\alpha$ Syn WT with the MTSL-label at position A18C, A90C and A140C were measured using 100  $\mu$ M monomeric  $^{15}$ N labelled  $\alpha$ Syn in 20 mM sodium acetate, pH 4.5 with 20 mM (low salt) or 200 mM (high salt) NaCl (**Figure 3.17 A-F**). Conditions that promote aggregation, pH 4.5, 20 mM NaCl, compared to higher salt concentrations (200 mM) which slows aggregation (**Figure 3.17 G**) show specific long-range intramolecular interactions between defined regions for all tested labelling positions (**Figure 3.17 A-C**). In particular, interactions between the N-terminal and C-terminal regions could be observed. This can be detected when putting the MTSL label at residue A18C and PRE effects are seen at the C-terminus (residues 110 – 140) (**Figure 3.17 A**) or with a slightly smaller reciprocal N-terminal PRE effect when MTSL labelling at A140C (**Figure 3.17 C**). These interactions are consistent with previous results at other pH values<sup>189,321,335</sup>. Also consistent with previous work in the literature at acidic pH<sup>331,335</sup> are the measured significant interactions between the NAC- and C-terminal regions (**Figure 3.17 B**). Excitingly, transient intramolecular interactions between the P1-region (and some residues in the P2 region) and residues A18C, A90C and A140C were observed in these experiments which have not been detected in other

studies<sup>189</sup>. Interestingly, meta-analysis of 11 PRE NMR studies has provided further evidence for the P1, P2 regions to be an interaction hub<sup>328</sup>. Remarkably, performing the same experiments under conditions that slow down the aggregation kinetics for  $\alpha$ Syn WT (pH 4.5, 200 mM NaCl (high salt), **Figure 3.17 G**) showed a significant decrease in magnitude for these detected interactions with the P1 (and P2) region, especially when putting the spin label at residue 90 or 140 (**Figure 3.17 D-F**). This observation is consistent with effects on electrostatic interactions when adding salt, possible involving residues K45, E46, E57 or H50 in P2, or residues that juxtapose P1 (K32, K34, E35, K43) and/or P2 (E61). The finding that stronger transient intramolecular interactions with P1 and P2 result in faster fibril formation suggests that these interactions are critical in defining the aggregation rate.



**Figure 3.17: NMR PRE data of  $\alpha$ Syn WT with MTSL labels at position A18C, A90C and A140C to observe intramolecular interactions through the protein. A-F, Intramolecular PRE intensity ratios of amide protons (paramagnetic/diamagnetic) for WT  $\alpha$ Syn variants with MTSL spin labels at A18C (A,D), A90C (B,E) or A140C (C,F) at low (A-C) or high (D-F) ionic strengths, at 100  $\mu$ M protein concentration, 15  $^{\circ}$ C. Green, red and purple bars show intensity ratios for the N-terminal, NAC and C-terminal regions, respectively. Data highlighted with green panels indicate the position of the P1–P2 region. Black line shows a rolling window (over five residues for easier comparison) of the PRE effects. Schematics are shown above each plot with a corresponding colour scheme. The locations of the spin label is indicated by yellow circles. (G) Aggregation kinetics (note the short time scale depicted) of WT  $\alpha$ Syn (100  $\mu$ M in 20 mM sodium acetate, pH 4.5, at low (20 mM added NaCl) or high (200 mM added NaCl) ionic strength at 37  $^{\circ}$ C). Residues 4, 11, 13, 16, 19, 23, 29, 30, 41, 45, 47, 62, 82, and 112 were not included because of overlapping peaks; residues 1, 108, 117, 120, 128, and P138 were not assigned. The signal to noise level for the diamagnetic spectra is >3.37, >2.11, >3.98, >4.77, >2.83, and >3.15 and for the paramagnetic spectra >3.15, >3.64, >4.91, >4.56, >2.01, and >3.25 for A-F, respectively.**

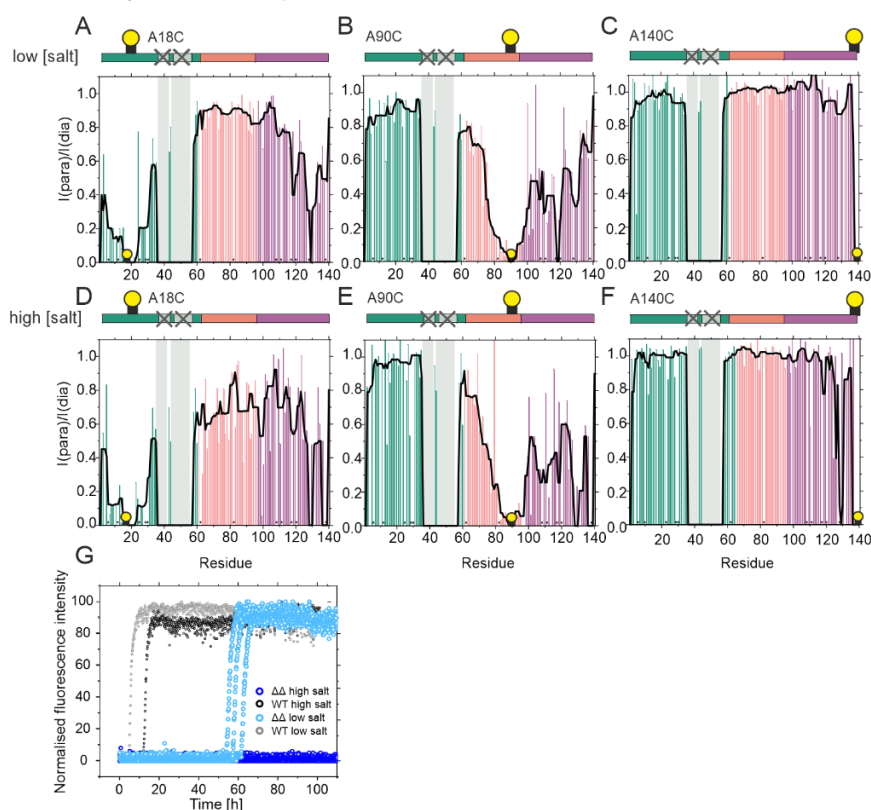
To ensure that the PRE effects observed in **Figure 3.17** are solely driven by intramolecular interactions and no intermolecular interactions are contributing, control experiments with 50  $\mu\text{M}$   $^{15}\text{N}$  labelled  $\alpha\text{Syn}$  WT (no MTSL) and 50  $\mu\text{M}$   $^{14}\text{N}$  MTSL labelled protein (at positions A18C, A90C and A140C) were performed under the same experimental conditions as described above (**Figure 3.18 A,B**). Both tested conditions (low and high NaCl) show no PRE effects, indicating no contribution of intermolecular interactions at the protein concentrations used.



**Figure 3.18: Control intermolecular PRE experiments of WT  $\alpha\text{Syn}$  in low (20 mM) and high (200 mM) salt (NaCl) at pH 4.5.** Intermolecular PRE intensity ratios of amide protons (paramagnetic/diamagnetic) for  $^{15}\text{N}$  WT  $\alpha\text{Syn}$  (50  $\mu\text{M}$ ) incubated with  $^{14}\text{N}$   $\alpha\text{Syn}$  variants (50  $\mu\text{M}$ ) with MTSL spin labels at A18C, A90C and A140C in low (A) and high (B) ionic strength (20 mM sodium acetate, pH 4.5, 15  $^{\circ}\text{C}$ ). Black lines show the median value over a rolling window of 5 residues. Schematics of the sequence of  $\alpha\text{Syn}$  are shown above each plot (green = N-terminal region, red = NAC region, purple = C-terminal region). The location of spin labels are denoted by a yellow circle. Residues 4, 11, 13, 16, 19, 23, 29, 30, 41, 45, 47, 62, 82, and 112 were not included because of overlapping peaks; residues 1, 108, 117, 120, 128, and P138 were not assigned. The signal to noise level for the diamagnetic spectra is >2.32, >3.41, >2.98, >3.56, >2.58, 3.77 and for the paramagnetic spectra is >2.22, >4.17, >3.62, >2.09, >2.99, >4.35 for A and B, respectively.

Next, the effect of removing the P1 and P2 region on the conformational properties of  $\alpha\text{Syn}$  was analysed by performing NMR PRE experiments for the  $\Delta\Delta$  variant. The NMR experiments were repeated under identical conditions compared to  $\alpha\text{Syn}$  WT (pH 4.5 at low and high ionic strength) (**Figure 3.19 A-F**). Under these conditions,  $\Delta\Delta$  aggregates significantly slower (for low ionic strength, lag time of  $57.3 \pm 3.6$  h) compared with  $\alpha\text{Syn}$  WT or does not show any fibril formation for over 110 h (high ionic strength) (**Figure 3.19 G**). The NMR PRE data show that contacts observed for  $\alpha\text{Syn}$  WT between N- and C-

terminal regions, as well as between NAC and C-terminal regions were mostly maintained for the deletion variant  $\Delta\Delta$ . Interactions with the P1 and P2 region on the other hand were removed for obvious reasons as the P1 and P2 motif is deleted in  $\Delta\Delta$ . This might further suggest an important role of interactions with the P1 (and P2) region for fibril formation as  $\Delta\Delta$  shows strongly decreased aggregation kinetics. It should be noticed that contacts between the N-terminal region (MTSL label at residue 18) and the NAC and C-terminal regions and between the NAC region (MTSL label at residue 90) and the N- and C-terminal regions are similar in  $\alpha$ Syn WT and  $\Delta\Delta$ , whilst transient interactions between the C-terminal region (MTSL label at residue 140) and the N-terminal region are smaller in  $\Delta\Delta$ , which is indicative of a complex interplay of interactions that depends closely on the sequence and solution conditions.

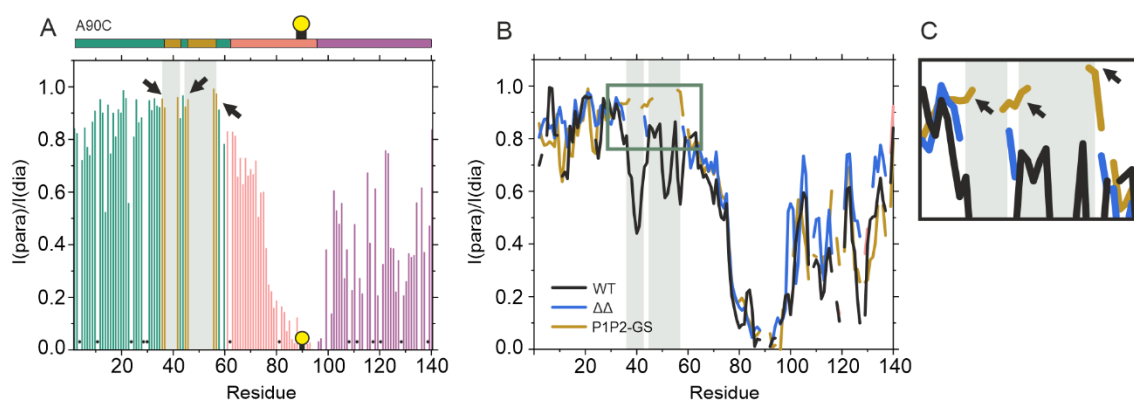


**Figure 3.19: NMR PRE data of  $\alpha$ Syn  $\Delta\Delta$  with MTSL labels at position A18C, A90C and A140C to observe intramolecular interactions through the protein. (A-F)** Intramolecular PRE intensity ratios of amide protons (paramagnetic/diamagnetic) for  $\Delta\Delta$   $\alpha$ Syn variants with MTSL spin labels at A18C (A,D), A90C (B,E) or A140C (C,F) at low (A-C) or high (D-F) ionic strengths, at 100  $\mu$ M protein concentration, 15  $^{\circ}$ C. Green, red and purple bars show intensity ratios for the N-terminal, NAC and C-terminal regions, respectively. Data highlighted with green panels indicate the position of the P1–P2 region which are deleted in this variant. Black line shows a rolling window (over five residues for easier comparison) of the PRE effects. Schematics are shown above each plot with a corresponding colour scheme. The locations of spin labels are indicated by yellow circles, deleted regions are shown by black crosses. (G) Aggregation kinetics of WT  $\alpha$ Syn and  $\Delta\Delta$  in comparison (100  $\mu$ M in 20 mM sodium acetate, pH 4.5, at low (20 mM added NaCl) or high (200 mM added NaCl) ionic strength at 37  $^{\circ}$ C). Residues 4, 11, 13, 16, 19, 23, 29, 30, 41, 45, 47, 62, 82, and 112 were not included because of overlapping peaks; residues 1, 108, 117, 120, 128, and P138 were not assigned. Residues 4, 11, 13, 16, 19, 23, 29, 30, 62, 82, and 112 were not included because of overlapping peaks; residues 1, 108, 117, 120, 128, and P138 were not assigned.

**Table 3.5: Aggregation kinetics of  $\alpha$ Syn WT and  $\Delta\Delta$  at pH 4.5 at low (20 mM) and high (200 mM) NaCl.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\alpha$ Syn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 110 h incubation is indicated by “-”. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in  $10^*$ [RFU/h].

	$\alpha$ WT, pH 4.5, low	$\alpha$ WT, pH 4.5, high	$\Delta\Delta$ , pH 4.5, low	$\Delta\Delta$ , pH 4.5, high
Lag time [h]	4.0 $\pm$ 0.5	11.4 $\pm$ 1.9	57.3 $\pm$ 3.6	-
Elongation rate*	3.6 $\pm$ 0.5	3.3 $\pm$ 0.3	2.3 $\pm$ 0.6	-
$t_{50}$ [h]	3.3 $\pm$ 0.4	13.7 $\pm$ 1.5	60.2 $\pm$ 5.8	-
% pellet	90	90	90	15

To further investigate the importance of the monomeric conformational states that present transient long-range interactions with the P1 and P2 regions for aggregation, intramolecular PREs were measured for P1P2-GS. This low-aggregation-prone  $\alpha$ Syn variant (**Figure 3.5 D**) was MTSL labelled at position A90C (**Figure 3.20 A**). Backbone assignment (performed by Roberto Maya-Martinez, University of Leeds) and NMR analysis for this variant were more challenging especially in the N-terminal region as 20 additional serines and glycines were introduced with very similar HN-chemical shifts due to their nearly identical chemical properties and environment resulting in strongly overlapping peaks. Therefore, in addition to the proline residues (108, 117, 120, 128, and 138), many peaks within the P1 and P2 region ( $^{36}$ SGSGSGS $^{42}$  and  $^{42}$ GSGSGSGSGSGSG $^{57}$ ) could not be assigned including resonances from residues 38-41 and 47-55 (see BMRB file 28045). However, PRE data showed a similar interaction trend to  $\Delta\Delta$  with long-range interactions remaining between N-and C-terminal region as well as NAC and C-terminal region and hint that the PRE effect on the P1 and P2 regions observed for WT  $\alpha$ Syn, is significantly reduced in P1P2-GS (**Figure 3.20 B,C**). Accordingly, deleting or replacing P1 and P2 does not change the compaction of the  $\alpha$ Syn protein, yet a significant reduction in aggregation is observed, demonstrating the crucial role of the P1 and P2 sequences in controlling the aggregation of  $\alpha$ Syn by interacting with other regions of the protein.

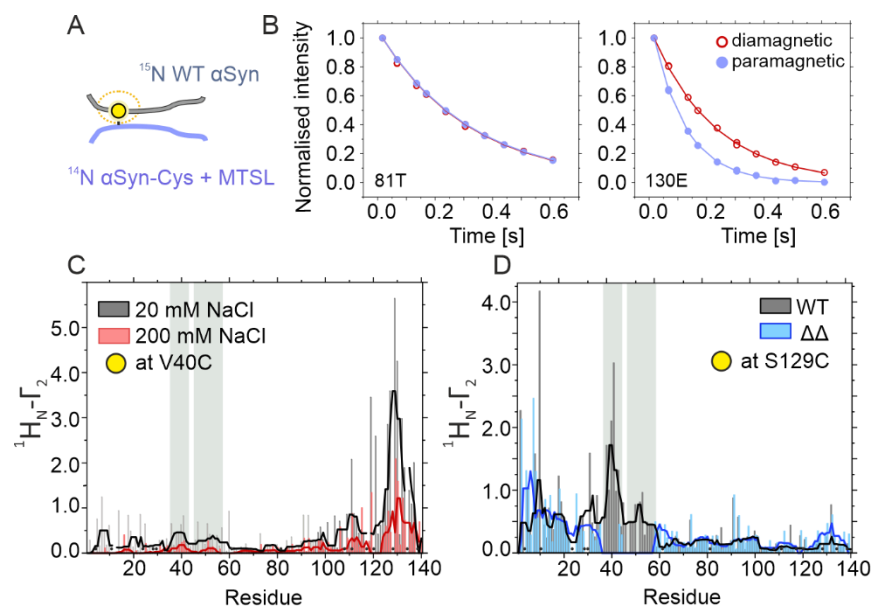


**Figure 3.20: Intramolecular interactions for  $\alpha$ Syn P1P2-GS.** (A) Intramolecular PRE intensity ratios of amide protons (paramagnetic/diamagnetic) for P1P2-GS  $\alpha$ Syn with the MTSL spin label at A90C at low ionic strength (20 mM NaCl), 15 °C, pH 4.5. Green, red and purple bars show intensity ratios for residues in the N-terminal, NAC and C-terminal regions, respectively. Orange bars highlight residues in the P1 and P2 regions that could be assigned and measured. The grey boxes mark the P1 and P2 regions. Black arrows show only a small PRE effect is observed in the P1-P2 region for P1P2-GS. Due to the repeating glycine and serine residues in the P1-P2 sequence, not all residues could be assigned. (B) Comparison of a rolling window (over five residues for easier comparison) of the PRE effects for WT (black),  $\Delta\Delta$  (blue) and P1P2-GS (orange)  $\alpha$ Syn. The black box is zoomed out in (C) to highlight residues in the P1-P2 region. Residues 4, 11, 13, 16, 19, 23, 29, 30, 62, 82, and 112 were not included because of overlapping peaks; residues 1, 108, 117, 120, 128, and P138 were not assigned. The signal to noise level is  $>3.92$  (dia) and  $>2.87$  (para).

### 3.7.2 Intermolecular interactions with P1 determine aggregation

Next, the role of the P1 and P2 motifs on intermolecular interactions was assessed. Literature data have shown that at high protein concentrations (500  $\mu$ M, normal brain concentration of  $\alpha$ Syn  $\sim 20$   $\mu$ M<sup>254</sup>) weak transient intermolecular interactions are formed between residues 38–45 (similar to the P1 region (residues 36–42)) and the C-terminal region, residues 124–140 (head to tail) and also homo-typic contacts (head to head) ( $K_D \approx 500$   $\mu$ M) at pH 6.0 and low ionic strength (10 mM MES, no additional NaCl)<sup>429</sup>. Intermolecular PREs were performed with  $\alpha$ Syn WT and  $\Delta\Delta$  to investigate if removal of the P1 and P2 regions disrupts these interchain interactions. For this experimental setup R2 relaxations were measured as these provide a more accurate detection of distances to the spin label considering only R2 relaxation processes (instead of both, R1 and R2) which contribute when measuring intensity ratios of para- and diamagnetic samples<sup>168</sup> (Figure 3.21 A,B) (see Section 2.9.2.2 the NMR setup for R2 measurements). 250  $\mu$ M <sup>14</sup>N- $\alpha$ Syn was spin labelled at residue 40 (V40C, within the P1 region) and incubated with the same amount of <sup>15</sup>N WT  $\alpha$ Syn at pH 4.5 at low and high ionic strengths (Figure 3.21 C). The data showed that indeed, residue 40 forms intermolecular interactions primarily with residues in the negatively charged C-terminal region of  $\alpha$ Syn WT, but also weakly with the N-terminal region, consistent with the literature<sup>429</sup>. These interactions were shown to be electrostatically mediated, as under high salt conditions (retarded aggregation), the <sup>1</sup>HN- $\Gamma_2$  rates decrease, further highlighting that these contacts are

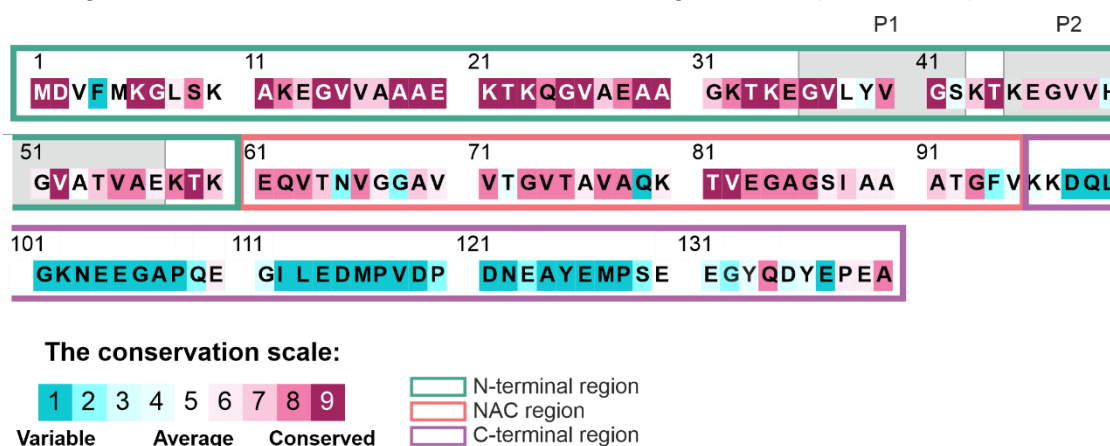
important in early stages of aggregation. As the  $\Delta\Delta$  variant cannot be MTSL-labelled at position 40, the experiment was repeated with a spin label at residue S129C for  $\alpha$ Syn WT and  $\Delta\Delta$  to see if similar interactions between P1 and the C-terminal region can be detected (**Figure 3.21 D**). These experiments showed a significant PRE effect for the P1 and P2 region as well as the first  $\sim 20$  N-terminal residues in  $\alpha$ Syn WT. Importantly, contacts with the extreme N-terminus (residues 1-15) were maintained in the deletion variant  $\Delta\Delta$ , whilst interactions with the P1 and P2 motifs were no longer possible. This indicates, that intermolecular contacts between the N- and C-terminal regions are independent of the presence of P1 and P2 regions. In summary, intra- and intermolecular interaction analysis at acidic pH underlines the importance of the P1 and P2 motifs in controlling  $\alpha$ Syn aggregation, not only by their local insolubility and high aggregation propensity, but also because they regulate the conformational ensemble of the monomeric IDP that defines its ability to aggregate into amyloid fibrils.



**Figure 3.21: Intramolecular interactions between P1, P2 region and C-terminal region control aggregation.** (A) Schematic of intermolecular PRE experiments in which  $^{14}\text{N}$  and  $^{15}\text{N}$   $\alpha$ Syn are illustrated as blue and grey chains, respectively. MTSL is shown as a yellow circle. (B) exemplary decay rates for residue 81T (left hand site), no PRE effect or 130E (right hand site), strong PRE effect. Blue colours indicate decay of paramagnetic, red curve of diamagnetic spectra. (C)  $\text{HN-}\Gamma_2$  rates for  $\alpha$ Syn WT labelled with MTSL at position 40 (A40C) at pH 4.5 in low-salt (20 mM added NaCl) (black) or high-salt (200 mM added NaCl) (red) conditions, 15  $^\circ\text{C}$ . Bars depict residue-specific  $\text{HN-}\Gamma_2$  rates. (D)  $\text{HN-}\Gamma_2$  rates at pH 4.5 under low-salt conditions (20 mM added NaCl) for WT (black) or  $\Delta\Delta$  (blue)  $\alpha$ Syn, labelled at position 129 (S129C). Bars depict residue-specific  $\text{HN-}\Gamma_2$  rates. Residues 4, 11, 13, 16, 19, 23, 29, 30, 62, 82, and 112 were not included because of overlapping peaks; residues 1, 108, 117, 120, 128, and P138 were not assigned.

### 3.8 P1 and P2 are conserved in synucleins

An analysis for conserved regions in synucleins using ConSurf<sup>395</sup> with the  $\alpha$ Syn WT amino acid sequence revealed that the N-terminal region is highly conserved with a ConSurf grade of 7.5 (1=variable; 9=conserved, see **Figure 3.22**), the NAC presents a ConSurf value of 6.7 and the C-terminal region differs most between synuclein variants with a ConSurf grade of 2.6. This is in agreement with the literature<sup>378</sup> as also  $\beta$ Syn and  $\gamma$ Syn were considered in this search. The P1-P2 regions show a higher than average (5.7 for the full-length protein) overall conservation with values of 6.9 (P1) and 6.5 (P2) (**Figure 3.22**). A possible explanation for keeping P1 and P2 in the sequence, although highly aggregation-prone, could be that amyloid formation might have a protective effect against neurotoxic stress<sup>430</sup>. Alternatively, the aggregation-prone motifs might be required to fulfil the protein's physiological function centred on the stabilisation, sequestration and fusion of pre-synaptic vesicles<sup>255,274</sup>. The latter was further investigated in this work as described in the following sections (Section 3.9).



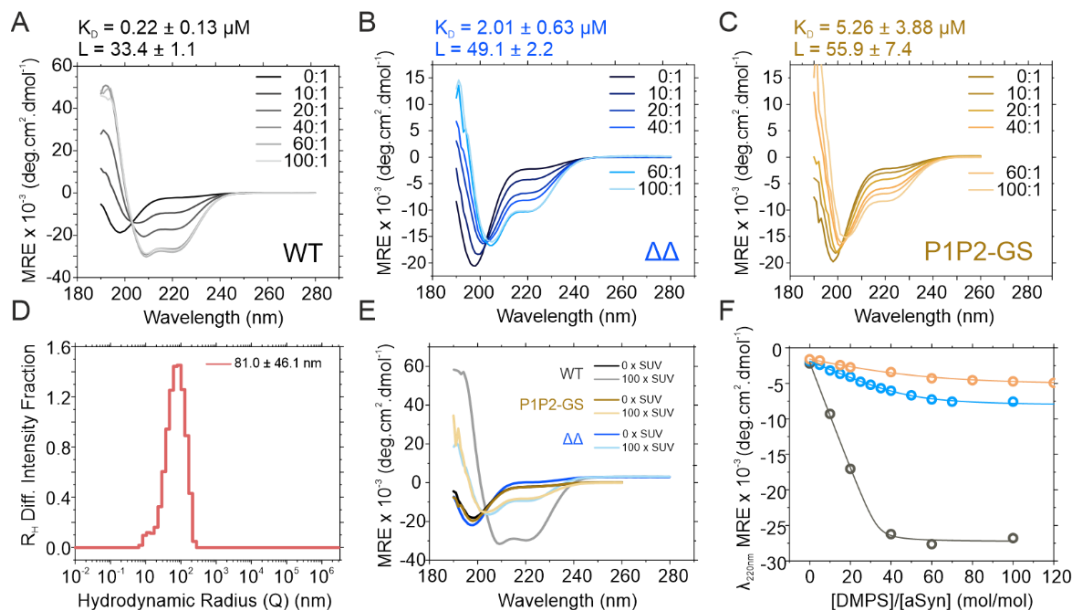
**Figure 3.22: Sequence conservation within the synuclein family calculated using ConSurf<sup>395</sup>.** The grade of conservation is colour coded from blue (variable) to red (very conserved). N-terminal region, NAC region and C-terminal region are highlighted in green, red and purple, respectively. The P1 and P2 region are highlighted in grey.

### 3.9 Importance of P1 and P2 for membrane remodelling

It is known that the N-terminal region (~100 residues), including the P1 and P2 region, is involved in micelle, liposome and membrane binding<sup>314,431</sup>, important for the physiological function of  $\alpha$ Syn of remodelling membrane vesicles<sup>272,274,356</sup>. Therefore, the role of P1 and P2 in  $\alpha$ Syn function was evaluated by performing liposome binding experiments with the variants  $\Delta\Delta$  and P1P2-GS.  $\alpha$ Syn WT has been shown to form an  $\alpha$ -helical secondary structure in its N-terminal region (residues 1–97) upon membrane binding<sup>186,432</sup>. Especially the extreme N-terminus (residues 1-25) was previously shown to play a key role in initiating membrane binding<sup>185</sup>. Membrane binding enhances aggregation into



amyloid fibrils<sup>351</sup> and the resulting helical conformation has been described as being the toxic state in yeast cells<sup>360,361</sup>. First, it was assessed whether  $\Delta\Delta$  and P1P2-GS can still adopt a helical conformation when binding liposomes (**Figure 3.23 A-C**). Liposomes were prepared from 1,2-dimyristoyl-sn-glycero-3-phospho-L-serine (DMPS), one of the major lipids in synaptic vesicles, with a diameter of  $\sim 160$  nm as determined via DLS (**Figure 3.23 D**).  $\alpha$ Syn monomeric protein and DMPS liposomes were mixed and the secondary structure of  $\alpha$ Syn was monitored using far-UV circular dichroism (CD) (**Figure 3.23 A-C**) (see Section 2.10.1 for detailed description of methods).



**Figure 3.23: CD binding assays of  $\alpha$ Syn WT,  $\Delta\Delta$  and P1P2-GS to DMPS liposomes. (A-C)** Far-UV CD spectra of 25  $\mu\text{M}$  WT  $\alpha$ Syn (**A**),  $\Delta\Delta$  (**B**) or P1P2-GS (**C**) incubated with increasing ratios of [DMPS]:[protein].  $K_D$  and  $L$  values were calculated from the change in MRE at  $\lambda_{222}$  nm fitted to a single-step binding model<sup>351</sup> (see Section 2.10.1. and **Equation 2.3** for detailed methods). (**D**) Dynamic light scattering of DMPS liposomes (see Section 2.7.6) showing they have a hydrodynamic radius ( $R_h$ ) of  $81.0 \pm 46.1$  nm. (**E**) Far-UV CD spectra of 25  $\mu\text{M}$  WT  $\alpha$ Syn (black),  $\Delta\Delta$  (blue) or P1P2-GS (orange) incubated in the absence or presence of liposomes (100:1 (mol/mol) DMPS: $\alpha$ Syn). (**F**) Change of CD signal of WT  $\alpha$ Syn (black),  $\Delta\Delta$  (blue) or P1P2-GS (orange) at 220 nm as a function of [DMPS]/[ $\alpha$ Syn] ratio. Data were fitted (solid lines) to a single-step binding model, yielding the affinity ( $K_D$ ) and stoichiometry value ( $L$ , the number of DMPS molecules in the bilayer that are involved in binding to one molecule of  $\alpha$ Syn) (see Section 2.10.1. and **Equation 2.3**).

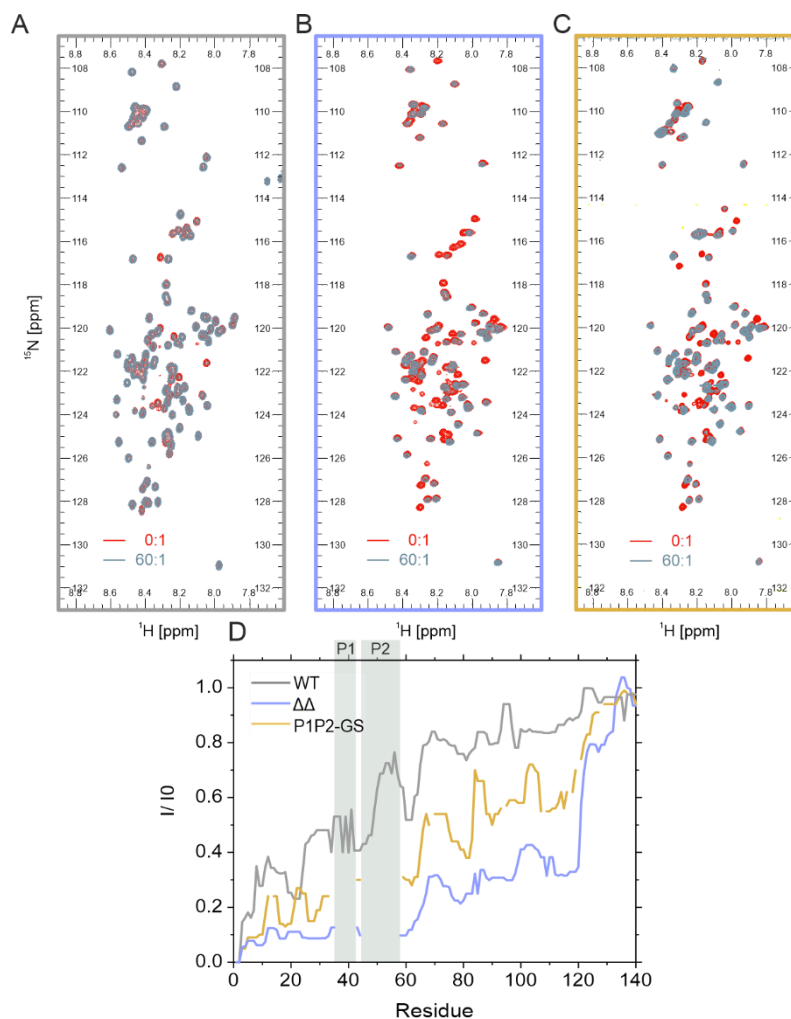
Both,  $\Delta\Delta$  and P1P2-GS are able to bind liposomes but adopt only 30 % and 13 % helical structure, respectively. This is significantly less helical content than the expected 64 % for these variants assuming similar helix formation to that of  $\alpha$ Syn WT (**Figure 3.23 E and Table 3.6**). The binding affinity between DMPS liposomes and  $\alpha$ Syn variants ( $\alpha$ Syn WT,  $\Delta\Delta$  and P1P2-GS) was measured by a lipid titration experiment via CD (see Section 2.10.1. for detailed methods) and showed a 10-fold or even 20-fold weaker affinity for  $\Delta\Delta$  and P1P2-GS compared to  $\alpha$ Syn WT ( $K_D = 0.22 \pm 0.13 \mu\text{M}$ ;  $2.01 \pm 0.63 \mu\text{M}$  and  $5.26 \pm 3.88$  for WT,  $\Delta\Delta$  and P1P2, respectively) (**Figure 3.23 A-C, F**). However, the stoichiometry ( $L$ ), providing information about the total number of DMPS molecules in

the bilayer involved in binding one  $\alpha$ Syn molecule, is similar for all three protein variants. The fact that  $\Delta\Delta$  and P1P2-GS show similar binding characteristics indicates that reduced helical formation is sequence specific and does not result from changing the spacing of the imperfect repeats.

**Table 3.6: Expected and measured  $\alpha$ -helical structure of  $\alpha$ Syn WT,  $\Delta\Delta$  and P1P2-GS in absence and presence of saturating conditions of DMPS liposomes (100:1 [DMPS]:[ $\alpha$ Syn]).** Expected values denote the amount of  $\alpha$ -helical structure observed in previous studies for WT<sup>351</sup> (69%, 97/140 residues),  $\Delta\Delta$  was presumed to have a lower percentage of  $\alpha$ -helical structure based on the deletion region being part of the structured stretch of  $\alpha$ Syn (64%, 77/120 residues). Measured data denote the average population of structure calculated from Dichroweb<sup>392</sup> using experimental CD data. Note that for P1P2-GS the expected  $\alpha$ -helical structure ranges from 64-69 % as P1P2 are replaced with a Ser-Gly linker less likely to form helical structures compared to the WT sequence.

Protein variant	Presence of liposomes	Expected $\alpha$ -helical structure (%)	Measured $\alpha$ -helical structure (%)
WT	+	69	74
	-	0	10
$\Delta\Delta$	+	64	30
	-	0	7
P1P2-GS	+	64-69	13
	-	0	8

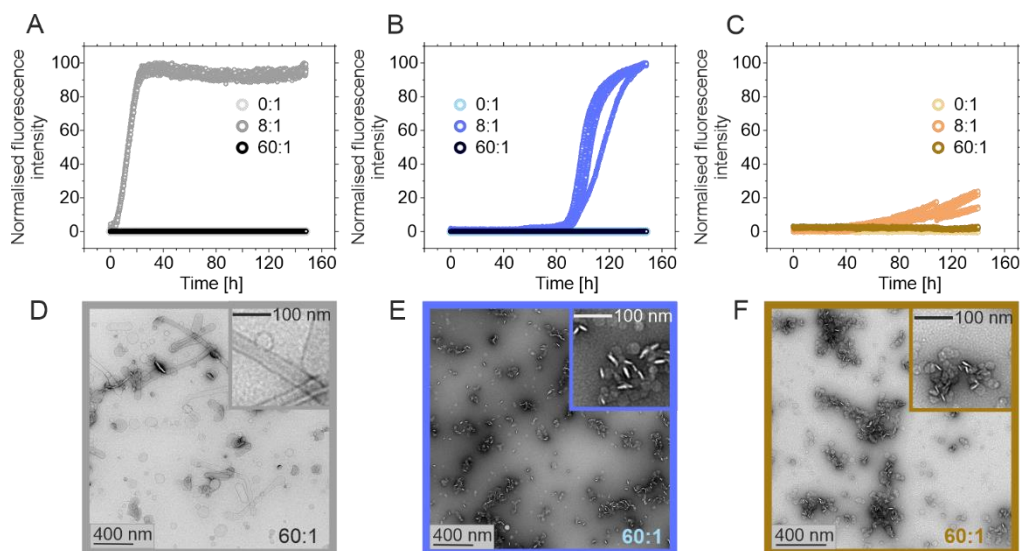
Next, it was examined which residues of  $\alpha$ Syn WT,  $\Delta\Delta$  and P1P2-GS are involved in the binding process to DMPS liposomes using NMR (see Section 2.9.3). For this, <sup>15</sup>N-<sup>1</sup>H HSQC NMR spectra of  $\alpha$ Syn variants in the presence or absence of liposomes was measured obtaining a residue specific binding information (**Figure 3.24 A-C**). As the used liposomes have a molecular mass ~1000 x bigger than  $\alpha$ Syn resulting in an increased hydrodynamic radius, they present slower tumbling rates compared with the ~15 kDa  $\alpha$ Syn. Due to this, in a binding event, resonances of residues that strongly bind to DMPS liposomes are reduced in intensity compared to non-binding residues<sup>433</sup>. Peak height intensity ratios (excess liposomes present (60x excess of lipid molecules compared with  $\alpha$ Syn) / no liposomes present) therefore give an information of which amino acids are involved in the binding. It should be considered that differences in peak intensities might also be a result of the varying helical propensities when the synuclein variants bind liposomes. Formation of secondary structural elements could cause peak broadening. Interestingly, whilst for  $\alpha$ Syn WT only the extreme N-terminal region builds the binding interface (residues 1-25)<sup>186</sup>,  $\Delta\Delta$  shows binding with all but the last ~20 C-terminal residues. The experiment with P1P2-GS indicated intermediate binding events suggesting that both the sequence as well as the relative positions of the P1P2 motif play an important role in lipid binding (**Figure 3.24 D**).



**Figure 3.24: Residues involved in binding to DMPS liposomes investigated by NMR.**  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of WT  $\alpha$ Syn (A),  $\Delta\Delta$  (B) and P1P2-GS (C) in the presence (grey) or absence (red) of a 60:1 ratio of [DMPS]:[protein]. (D) Intensity ratios (presence/absence of liposomes) of cross-peaks for WT  $\alpha$ Syn (grey),  $\Delta\Delta$  (blue) and P1P2-GS (orange) are shown by illustrating the median value over a rolling window of five residues determined using OriginPro. The positions of P1 and P2 are highlighted with gray bars. Note that residues 36–42 and 45–57 are deleted in  $\Delta\Delta$ , and these residues (replaced with (SG)<sub>3</sub>S (P1) and (GS)<sub>6</sub>G (P2)) could not be assigned for P1P2-GS.

DMPS liposomes are known to drive  $\alpha$ Syn aggregation by promoting heterogeneous primary nucleation<sup>351</sup>. Therefore the effect on amyloid formation of liposomes for  $\Delta\Delta$  and P1P2-GS was tested using a ThT fluorescence assay (Figure 3.25 A-C). Different [DMPS]:[ $\alpha$ Syn] ratios were employed, with 0x, 8x or 60x (mol/mol) excess of lipid molecules over  $\alpha$ Syn molecules under quiescent conditions (20 mM sodium phosphate, pH 6.5) following literature procedures<sup>351</sup>. As expected,  $\alpha$ Syn WT did not form fibrils in the absence of liposomes (due to quiescent conditions). However, adding 8x molar excess of DMPS drastically accelerated fibril formation. When adding an excess of lipids (60x (mol/mol)), aggregation is prevented due to the depletion of free monomeric  $\alpha$ Syn that is not bound to liposomes and required for elongation of nuclei/oligomers (Figure

**3.25 A)** and ref<sup>351</sup>. For  $\Delta\Delta$ , a similar effect could be observed with no aggregation at 0x and 60 x molar excess of lipids, but heterogeneous nucleation (in the presence of 8x lipid) was slowed down by a factor of 20 compared with  $\alpha$ Syn WT (lag times =  $4.9 \pm 0.3$  h and  $93.0 \pm 2.6$  h for WT and  $\Delta\Delta$ , respectively) (**Figure 3.25 B**). The increased lag-time could be due to reduced helical structure in  $\Delta\Delta$  when bound to lipids. Also, P1P2-GS shows a significantly reduced aggregation in the presence of liposomes (**Figure 3.25 C**). TEM images of end-points of these samples incubated with 60x molar excess of lipid were taken (**Figure 3.25 D-F**). Strikingly, whilst  $\alpha$ Syn WT remodels liposomes into long tubes as reported previously<sup>272</sup>, images of  $\Delta\Delta$  and P1P2-GS incubated with liposomes did not show this effect. Instead, these proteins resulted in formation of small, prefibrillar-like aggregates which associated with the liposome surfaces and seemed to cause liposome fission, releasing smaller spherical liposomes. In summary, the P1 and P2 regions not only regulate aggregation, but also control the lipid-binding properties of distal regions of the  $\alpha$ Syn sequence and structure in the lipid-bound state and perturb membrane remodelling, without preventing binding to DMPS liposomes.



**Figure 3.25: Liposome induced aggregation of  $\alpha$ Syn and membrane remodelling.** (A-C) Aggregation kinetics of 50  $\mu$ M  $\alpha$ Syn WT (A),  $\Delta\Delta$  (B) or P1P2-GS (C) incubated with 0:1, 8:1 or 60:1 [DMPS]:[protein] (20 mM sodium phosphate, pH 6.5; 30  $^{\circ}$ C, no shaking). (D-F) TEM images of representative samples of WT  $\alpha$ Syn (D),  $\Delta\Delta$  (E) or P1P2-GS (F) at the endpoint of the incubations (150 h) in the presence of 60:1 [DMPS]:[protein].

### 3.10 Discussion

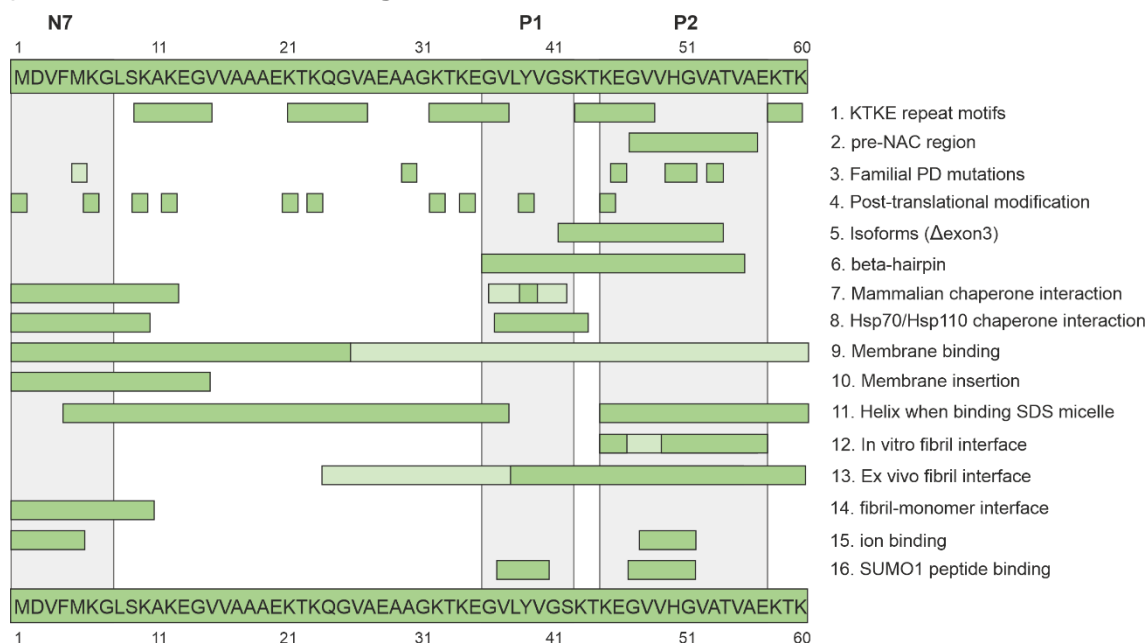
In this chapter the important role of short motifs in the N-terminal region of  $\alpha$ Syn in controlling aggregation and function was identified. Specifically, N7 (residues 1-7) and P1 (residues 36-42) were shown to be key regions in determining the ability of  $\alpha$ Syn to self-assemble into amyloid fibrils. The extreme N-terminal region, including the first seven residues, has been described previously to be crucial for aggregation, especially for the processes of elongation and secondary nucleation<sup>107,183</sup>, and also its contribution to membrane binding has been studied intensively, highlighting its indispensable role in the physiological function of membrane binding and remodelling<sup>185,186,272</sup>.

Therefore, the work in this chapter focussed mainly on the ability of P1 (and P2) to regulate aggregation of  $\alpha$ Syn *in vitro* and *in vivo* and their effect on function of  $\alpha$ Syn. Remarkably, when deleting or replacing (with a Gly-Ser-linker) the P1 motif, fibril formation at neutral pH is abolished (over 100 h) and *C. elegans* expressing  $\Delta$ P1 (or  $\Delta\Delta$ ) in the muscle cells is protected against a phenotypic effect, although the highly aggregation-prone NAC region<sup>313</sup> is still present and unaltered. Aggregation assays and NMR PRE experiments enabled the correlation of alterations in monomer conformation caused by sequence changes in the P1 and P2 regions with aggregation propensity. Based on these observations, the P1 region was defined as a 'master controller' of  $\alpha$ Syn aggregation, in that this region controls  $\alpha$ Syn fibril formation at neutral pH, synergistically with the P2 (preNAC) region<sup>311</sup> under acidic conditions. NMR PRE studies revealed that P1 and P2 exert their control by fine-tuning intra- and inter-molecular transient long-range interactions. This results in an aggregation-prone conformational ensemble for  $\alpha$ Syn monomers when P1 and P2 are present, but shifts to self-assembly protecting conformations when P1 and P2 are deleted or substituted with a Gly-Ser-linker. Presumably, the aggregation propensity of conformers is determined by the exposure or sequestration of the important NAC region<sup>189,190,313</sup> defined by the described long-range intra- and inter-molecular interactions involving the P1 and P2 regions. The fact that  $\Delta$ P1 and  $\Delta\Delta$  are able to form amyloid structures under specific conditions (e.g. low pH or in the presence of liposomes) indicates that the P1 and P2 regions enforce a kinetic control over the thermodynamically favourable process of amyloid formation, although the precise molecular mechanism of this process, including the contribution of P1 and P2, remains to be elucidated. Co-incubation and cross-seeding experiments with  $\alpha$ Syn WT and  $\Delta$ P1 suggest inhibition of elongation, but deletion of P1 and P2 might also affect the structure and/or aggregation competence on the nuclei or oligomeric level. Whatever the precise mechanism of fibril formation is, the pH and ionic strength dependent aggregation indicates that P1 and P2 control  $\alpha$ Syn aggregation by a fine balance of hydrophobicity and charge, such that amyloid formation becomes highly sensitive to the solution conditions. This might explain why other studies on  $\alpha$ Syn below<sup>258,331,335</sup> or

above<sup>321,325,335,426,427</sup> pH 4.5 did not detect interactions involving P1 and P2. Taken together, the data presented in this chapter demonstrate that despite NAC being necessary and sufficient for aggregation<sup>313</sup>, flanking regions also play a critical role in modulating the aggregation propensity of  $\alpha$ Syn shown by an only seven residue long motif in the N-terminal region able to define its aggregation kinetics, even ablating aggregation (at least for long time periods) at neutral pH.

The crucial role of flanking regions of aggregation hotspots in controlling aggregation has been seen for other amyloid proteins before. For example, the polyproline region (P17) in exon 1 of Huntington has broadly suppressive effects on fibril formation<sup>407,408,434</sup>, the N-terminal region (residues 11-16) of amyloid  $\beta$  ( $A\beta_{40}$ ) enhanced fibril stability at acidic pH<sup>435</sup> and residues 306-311 destabilise the local structure of tau and trigger spontaneous aggregation<sup>211</sup>. Further, the N-terminal six residues of  $\beta$ 2-microglobulin are known to speed up aggregation of the protein when deleted<sup>214</sup>.

The discovery that the P1P2 motif is essential for the function of  $\alpha$ Syn in vesicle remodelling further indicates that the N-terminal region of  $\alpha$ Syn is crucial for both its physiological function and its disease etiology<sup>274,436</sup>. There is additional evidence in the literature supporting that the motifs N7, P1, and P2 are important in various biological processes summarised in **Figure 3.26** and **Table 3.7** and further discussed below.



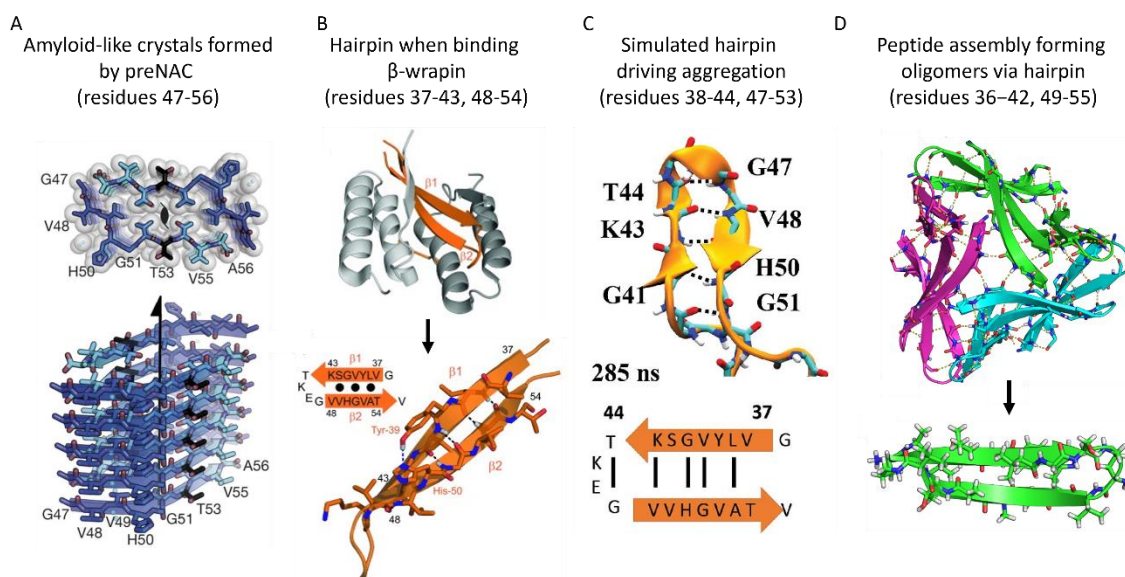
**Figure 3.26: Summary of processes involving the N-terminal region of  $\alpha$ Syn.** Shown is only the sequence of the N-terminal region. N7, P1 and P2 regions are highlighted with grey bars. Green bars indicate residues involved in the described process (listed on the left-hand side), light green boxes show residues less involved/more dynamic in the described process. More detail is provided in **Table 3.7** numbered accordingly to this figure.

**Table 3.7: Summary of processes involving the N-terminal region of  $\alpha$ Syn highlighted in Figure 3.26.** Numbers are accordingly to Figure 3.26. Interesting feature(s), involved residues, notes and references are listed.

#	feature	residues	notes	ref
	N7, P1, P2	1-7, 36-42, 45-57	These regions were identified to be crucial for aggregation and function as part of this thesis	
1	Repeat motif	10-15, 21-26, 32-37, 43-48, 58-63, 80-85	KTKEGV repeat motif, involved in membrane binding	310
2	preNAC	47-56	preNAC forms amyloid like crystals in isolation and was identified by Eisenberg and co-workers (crystallisation)	311
3	Familial PD mutations	M5T, A30P/G, E46K, H50Q, G51D, A53T/E/V	PD mutations show increased or decreased aggregation kinetic but are all involved in disease	145, 298
4	PTMs	M1; K6,10,12,21,23,32,34, 43; Y39	Acetylation (M1,K6,K10), ubiquitination (all lysines) and phosphorylation (Y39)	437
5	Isoforms ( $\Delta$ exon3)	41-54	Isoforms are observed with deleted region 41-54 and/or 103-130 in disease patients	438
6	$\beta$ -hairpin (see Figure 3.26)	36-42, 49-55	Peptide self-assembles to form oligomers via $\beta$ -hairpin formation	341
		37-43, 48-54	Hairpin formation as a result of $\beta$ -wrapin binding, inhibiting aggregation	187
		38-44, 47-53	Simulations find hairpin formation that accelerates aggregation	188
7	chaperone interaction	1-12, Y39 and six residues around	Chaperone binding prevents aggregation, phosphorylation at Y39 decreases binding	175
8	Hsp70, Hsp110	1-10, 37-43	Interaction for fibril dis-assembly process	225
9	Membrane binding	1-25, 26-97	1-25 responsible for initial binding, rest dynamic interaction	186
10	Membrane insertion	1-14	1-14 insert into membrane, rest of N-term. and NAC binds to surface	185
11	Helix formation	3-37, 45-92	NMR structure when binding SDS-micelles	314
12	<i>In vitro</i> interface	50-57 (1a), 66-78 (1b), 45-57 (2a), 45-46 (2b)	See Figure 1.24	96
13	<i>Ex vivo</i> interface	~24-55 and ~38-64	Non-identical protofilaments (see Figure 1.24), K43, K45 and H50 form cavity	117
14	Fibril-monomer interface	1-11/12	N-terminus of monomer interacts with C-terminus in fibrils, important for seeding and elongation processes	183, 184
15	Ion binding	1-5, 48-52	Most ion binding is seen to the C-terminal region, $\text{Cu}^{2+}$ , $\text{Zn}^{2+}$ and $\text{Fe}^{2+}$ are also observed to bind N-terminus	437
16	SUMO1 peptide binding	37-40, 48-52	SUMO1(15-55) inhibits aggregation <i>in vitro</i> and <i>in vivo</i> of $\alpha$ Syn, binds either P1 or P2	439

The P2 region (45-57) includes six of the twelve known familial PD mutations<sup>145,248</sup>. Further, this region was also shown to form the protofilament interface of some *in vitro* and *ex vivo* fibril architectures<sup>72,98,106,115,117</sup> highlighting its involvement in fibril formation. Moreover, the pre-NAC region (similar to P2) was demonstrated to form amyloid-like structures in isolation<sup>311</sup> (Figure 3.27 A). Work with a  $\beta$ -wrapin selected to bind  $\alpha$ Syn<sup>187</sup>

showed that  $\alpha$ Syn aggregation can be inhibited *in vitro*, in cell culture and *in vivo* by binding of this small protein to residues 37-54 of  $\alpha$ Syn<sup>187,342,415</sup>. Binding of this  $\beta$ -wrapin to  $\alpha$ Syn causes  $\beta$ -hairpin formation of residues <sup>37</sup>VLYVGSK<sup>43</sup> and <sup>48</sup>VVHGVAT<sup>54</sup> of  $\alpha$ Syn<sup>187</sup> (**Figure 3.27 B**) which is also observed in MD simulations (in the absence of the  $\beta$ -wrapin)<sup>188</sup> (**Figure 3.26** and **Figure 3.27 C**). Further, engineering an intra-molecular disulfide bond between residues 41 and 48 inhibited  $\alpha$ Syn self-assembly without the formation of a  $\beta$ -hairpin structure<sup>414</sup> possibly by perturbing required interactions with the P1 and P2 region which are demonstrated here to be important to initiate aggregation. Parts of the P1 and P2 region were also shown to be the interaction hubs of SUMO1 (15-55) peptide binding (residues 37-40 and 48-52). Interestingly, binding, which inhibits  $\alpha$ Syn aggregation *in vitro* and in *Drosophila*, occurs to either P1 or P2 and no hairpin formation is observed. In fact, the hairpin was shown to reduce the binding affinity to SUMO1 and  $\beta$ -hairpin formation was competing against SUMO1 binding<sup>439</sup>. In addition to this, a peptide similar to P1P2 (residues 36-55) was revealed to form a  $\beta$ -hairpin structure that self-assembles into cytotoxic oligomers (**Figure 3.27 D**), allowing the authors to make the assumption that this region, rather than NAC, drives nucleation into oligomers<sup>341</sup>.



**Figure 3.27: Schematic of structural features of  $\alpha$ Syn including the P1 and P2 regions.** (A) Amyloid-like crystal structure formed by the preNAC peptide (residues 47-56). Figure taken from Ref<sup>311</sup>. (B)  $\beta$ -hairpin structure of residues 37-54 when binding to a  $\beta$ -wrapin determined by NMR. Top: structure of hairpin (orange) and  $\beta$ -wrapin (white), bottom: detailed structure of  $\beta$ -hairpin highlighting side chains and H-bonds. Figure taken from Ref<sup>187</sup>. (C)  $\beta$ -hairpin structure found in simulations to accelerate aggregation including residues 38-53. Figure taken from Ref<sup>188</sup>. (D) Crystal structure of oligomeric species formed by  $\alpha$ Syn peptides including residues 36-55. Three copies of a triangular tetramer built by  $\beta$ -hairpin subunits (bottom) form an oligomeric structure. Figure taken from Ref<sup>341</sup>.

Interestingly, residue Tyr39 has been reported previously to play a key role in  $\alpha$ Syn assembly, as the substitution Y39A prevents aggregation<sup>440</sup> and Y39 (in combination



with F94) was shown to be responsible for interacting with small molecules able to decelerate aggregation<sup>433</sup>. Further, the P1 region of monomeric  $\alpha$ Syn is involved in binding several chaperones (centred on residue Y39) that prevent fibril formation<sup>175</sup>, and binding of chaperones to P1 can result in fibril disaggregation<sup>225</sup>.

In summary, the data presented here in combination with work from the literature highlight the vital role of the P1 and P2 motif in controlling  $\alpha$ Syn aggregation. This draws the attention to the importance of regions that flank NAC in aggregation, and demonstrates that aggregation in  $\alpha$ Syn is not solely initiated by NAC. Disrupting the crucial interactions made by P1 and/or P2 using small molecules or other reagents that target these sites, may therefore help to controlling  $\alpha$ Syn self-assembly.

Although P1 and P2 are highly aggregation-prone, evolution retained these sequences. A possible explanation for this could be their involvement in membrane binding and the linked physiological function of this protein. Despite the fact that the exact physiological role of  $\alpha$ Syn is not clarified<sup>250</sup>, stabilisation, sequestration and fusion of presynaptic vesicles are thought to be involved in its repertoire of functions<sup>256,273,274</sup>. In this Chapter, it could be shown that indeed, deleting the P1 and P2 regions or replacing them with a Gly-Ser linker eliminates the functional activity of  $\alpha$ Syn in membrane remodelling. Instead, liposome morphologies that are different from the large fused tubular structures formed by the  $\alpha$ Syn WT protein are formed. This result points out the frustration between aggregation and function involving the P1-P2 regions. This delicate balance between function and simultaneously enhancing fibril formation rationalises why PD-associated single point mutations such as A53T, E46K and others<sup>145</sup> enhance disease onset by simultaneously causing a gain-of-toxic function activities and loss-of-physiological function. For IDPs in general, the aggregation propensity of such aggregation-prone, yet functionally important, motifs cannot be protected by the framework of a folded tertiary native structure and are therefore at especially high risk to cause disease. So, it is not surprising that over a third of the 48 currently known human amyloidogenic proteins are at least in parts intrinsically disordered<sup>47</sup>. Despite such 'master controller' regions enabling dangerous liaisons, they also offer promising potential to regulate self-assembly by binding small molecules, chaperones, or other agents to these regions. Due to the fine balance of weak intra- and inter-molecular interactions that define the early stages of amyloid formation in  $\alpha$ Syn, minor changes in the state of the energetic landscape could interrupt aggregation without significantly perturbing the function of the protein.

Taken together, in this Chapter a 'master controller' region for aggregation and function was identified providing evidence for the important role of these flanking regions in amyloid disease. The role of individual residues within this newly discovered seven residue motif (P1) in defining  $\alpha$ Syn aggregation into amyloid will be investigated in Chapter 4.



## Chapter 4

# Single residue switches in the N-terminal P1-region change aggregation of $\alpha$ -synuclein

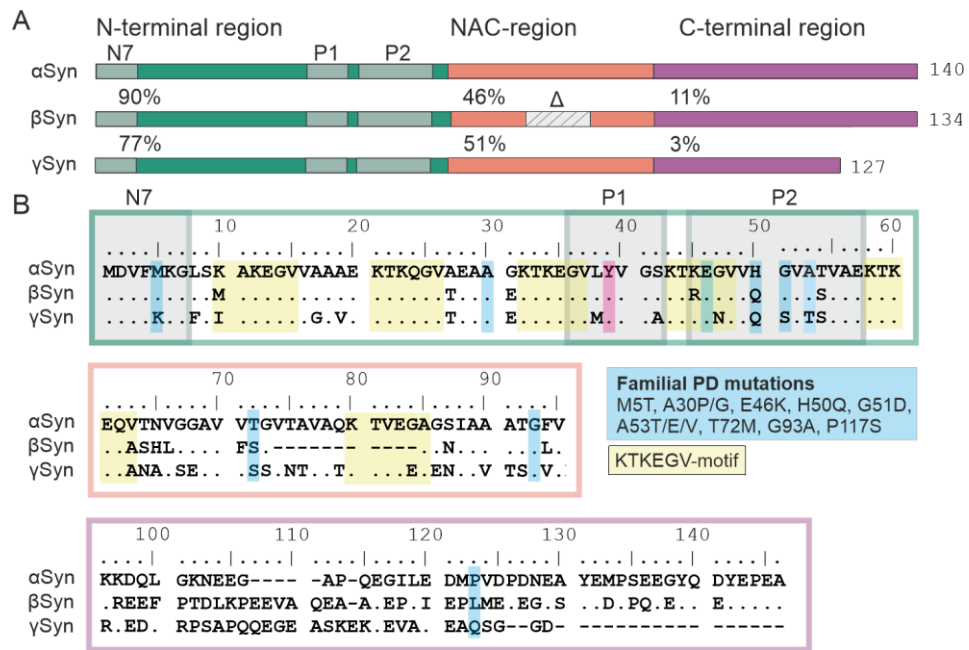
### 4.1 Introduction

Despite the long-known importance of the central NAC region in  $\alpha$ Syn aggregation<sup>313,405</sup>, it is becoming increasingly apparent that motifs that flank the NAC region, can influence the aggregation kinetics of the full length protein<sup>345,413,441,442</sup>. In Chapter 3 it was shown that the N-terminal region of  $\alpha$ Syn, especially the P1 motif (<sup>36</sup>GVLYVGS<sup>42</sup>), plays a critical role in controlling aggregation and function *in vitro* and *in vivo*. However, the exact role of each individual amino acid in the P1 sequence remained unclear and are further investigated in this Chapter.

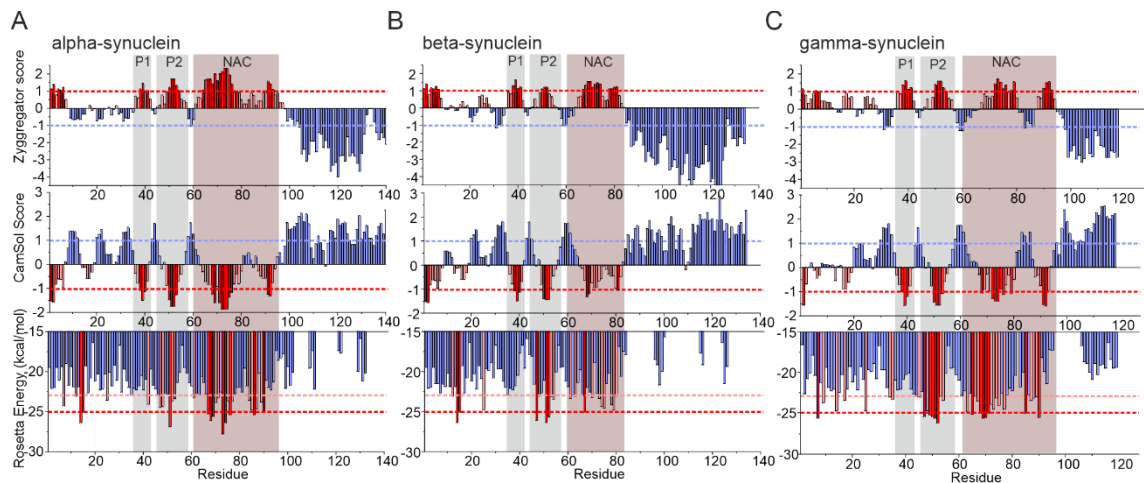
Single residue substitutions in  $\alpha$ Syn can drastically affect the aggregation propensity of the protein *in vitro*, in cells or *in vivo* and alter its toxicity and role in disease, the most common example probably being the PD-associated mutations A30P/G, E46K, H50Q, G51D, A53A/T/V (familial)<sup>145,172,268</sup> as well as A18T, and A29S<sup>291,292</sup> (no family history), and the DLB-associated variant E83Q<sup>294,299</sup>. There are also many other  $\alpha$ Syn point mutations (not associated with disease) that have been studied, including substitutions in the N-terminal- NAC- and C-terminal regions of the protein accelerating, retarding or not altering oligomer and fibril formation and having different effects on toxicity measured in cells and *in vivo*<sup>298,361,409,440,443-445</sup>. A more systematic study of amino acid changes was performed by Newberry *et al.* using a deep mutational scanning (DMS) approach in yeast. This work suggested the  $\alpha$ -helical form of  $\alpha$ Syn, formed when bound to membranes, to be the toxic species, and mutations perturbing helix formation were shown to be protective against cytotoxicity<sup>360,361</sup>. Also, post-translational modifications of individual residues have been shown to change the aggregation rate and pathological

processes of  $\alpha$ Syn<sup>446,447</sup>. Within the P1 region, especially one amino acid, Tyr39, stands out. Phosphorylated  $\alpha$ Syn at position Y39 is enriched in brain tissues and Lewy bodies in the substantia nigra and striatum of PD patients<sup>266</sup>. This post-translational modification results in distinct fibril architectures compared to  $\alpha$ Syn WT structures determined by cryoEM<sup>109</sup>. Also, Y39 is involved in chaperone binding (e. g. SecB, Skp) that protects  $\alpha$ Syn from self-assembly into amyloid structures, a process perturbed when Y39 is phosphorylated<sup>175</sup>. The same region is required for interactions with Hsp70 causing fibril disassembly<sup>225</sup>. Further, mutating Y39 to an alanine has been shown to inhibit fibril formation at neutral pH<sup>440</sup>.

Residue Y39 is conserved within the synuclein family, a family consisting of three proteins:  $\alpha$ -,  $\beta$ -, and  $\gamma$ Syn. In general, these proteins show high sequence similarities in the N-terminal region, which is of particular interest in this thesis (90% and 77% sequence identity to  $\alpha$ Syn for  $\beta$ - and  $\gamma$ Syn, respectively)<sup>369</sup> (**Figure 4.1**). Whilst  $\alpha$ Syn is aggregation-prone,  $\beta$ - and  $\gamma$ Syn show no self-assembly into amyloid at neutral pH<sup>365,366</sup> and even interrupt  $\alpha$ Syn aggregation when the proteins are co-incubated<sup>365,367</sup>. For  $\beta$ Syn, this might be explained by the lack of 11 residues within the NAC region<sup>313</sup>, however, there is no obvious sequence difference in  $\gamma$ Syn that explains this low ability to form fibrils. Although 17 of 35 residues within NAC are different between  $\alpha$ - and  $\gamma$ Syn (**Figure 4.1**), the overall aggregation propensity is predicted to be similar to the one observed in  $\alpha$ Syn (**Figure 4.2**). Therefore the altered aggregation kinetics of these proteins may be explained by changes in the N- or C-terminal flanking regions.



**Figure 4.1: Sequence alignment between the synuclein family members  $\alpha$ Syn,  $\beta$ Syn and  $\gamma$ Syn.** (A) Each protein comprises three regions: the amphipathic N-region (green), the amyloidogenic NAC region (red) and the acidic C-terminal region (purple). The previously identified regions N7 (residues 1-7), P1 (residues 36-42) and P2 (residues 45-57) are indicated in grey. The sequence identity of  $\beta$ Syn and  $\gamma$ Syn to  $\alpha$ Syn for each region is shown. The light grey region in the centre of NAC for  $\beta$ Syn depicts residues in the highly aggregation-prone NAC core that is deleted in this sequence. The number of residues in each protein are shown (right). (B) Sequence alignment for each of the paralogues (regions coloured according to (A)). Positions of familial PD mutations of  $\alpha$ Syn are highlighted in blue, each of the KTKEGV motifs involved in membrane binding are highlighted in yellow, and the conserved Y39 is highlighted in pink.

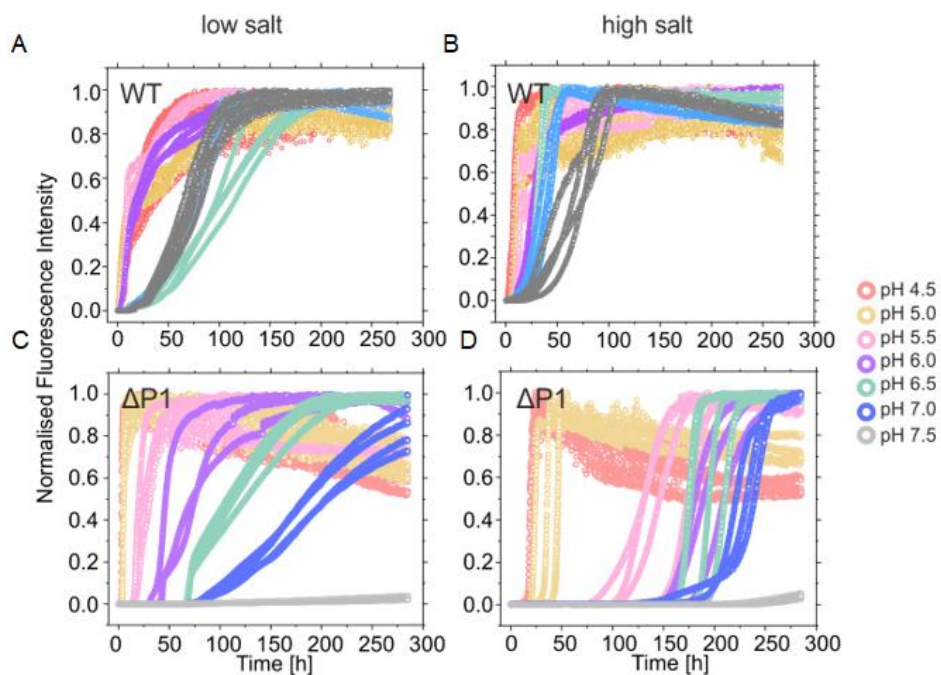


**Figure 4.2: *In silico* analysis of the  $\alpha$ -,  $\beta$ - and  $\gamma$ Syn sequences.** Zyggregator, CamSol and ZipperDB profiles for (A)  $\alpha$ Syn, (B)  $\beta$ Syn and (C)  $\gamma$ Syn. Bright red bars indicate aggregation-prone/low solubility regions. Light red bars indicate residues with a higher than average aggregation propensity/low solubility, but which do not meet the threshold and blue bars represent residues with low aggregation propensity. Red dashed lines indicate the low solubility/high aggregation propensity threshold, while blue dashed lines show threshold values for high solubility/low aggregation propensity. For Zipper DB, the red dashed lines show the threshold value of residues with a high probability of  $\beta$ -zipper formation. P1 and P2 regions are highlighted in grey and NAC region in pale red.

In this Chapter, the role of individual residues within the P1 region in the aggregation of  $\alpha$ Syn was evaluated *in vitro* and *in vivo*. Adding a synthetic peptide with the sequence of P1 *in trans* and performing an alanine scan shine light on the sequence-specificity of this motif. A complex interplay of interactions in  $\alpha$ Syn and  $\gamma$ Syn was observed, indicating that P1 synergises with residues in the NAC region and C-terminus to create conformers capable of initiating aggregation. The results presented here highlight that early interactions control the fibrillation process of synucleins and are remarkably specific, despite their weak and transient nature. For P1, aggregation vitally depends on both the location of the residues within the sequence and the identity of the sidechain at specific sites.

## 4.2 The P1 sequence drives aggregation

In the previous chapter (Chapter 3), the P1 region (residues <sup>36</sup>GVLVYVGS<sup>42</sup>) was identified as a 'master controller' of aggregation, having a significant inhibitory effect on fibril formation *in vitro* and *in vivo* (in *C. elegans*) when deleted ( $\Delta$ P1). A pH dependent aggregation assay of  $\Delta$ P1 in comparison to  $\alpha$ Syn WT at different ionic strengths (20 mM and 200 mM NaCl) was first performed over 300 h to determine how general the effect of P1 on aggregation is under different conditions (**Figure 4.3**). Aggregation has been shown to increase with reducing the pH for both tested variants ( $\alpha$ Syn WT and  $\Delta$ P1) and ionic strengths, and fibril formation is overall faster at lower salt concentrations (especially for  $\Delta$ P1) (**Table 4.1** and **Figure 4.4**). Interestingly,  $\Delta$ P1 fibrillates significantly slower than  $\alpha$ Syn WT at all tested buffer environments with a more drastic effect at increasing pH (no amyloid formation detected at pH 7.5; high salt within 300 h). This result further highlights the impact of the P1 region for aggregation especially at neutral pH and suggests a kinetic rather than a thermodynamic interruption of fibril formation as amyloids are being formed eventually but with longer lag-times than measured for  $\alpha$ Syn WT (**Figure 4.3**, **Table 4.1**). The presence of fibrils for  $\Delta$ P1 was confirmed by TEM imaging and fibril yield analysis (**Figure 4.5**, **Table 4.1**). Replacing the P1 region with a Ser-Gly-linker (P1-SG) also abolishes aggregation, providing additional evidence for the hypothesis that the P1 sequence rather than the altered spacing in  $\Delta$ P1 plays a key role in regulating the aggregation of this 140 amino acid long IDP (see **Figure 3.5 C**). From these experiments it can be concluded that the amino acid sequence that comprises the P1 region must play a vital role in controlling the aggregation of  $\alpha$ Syn.

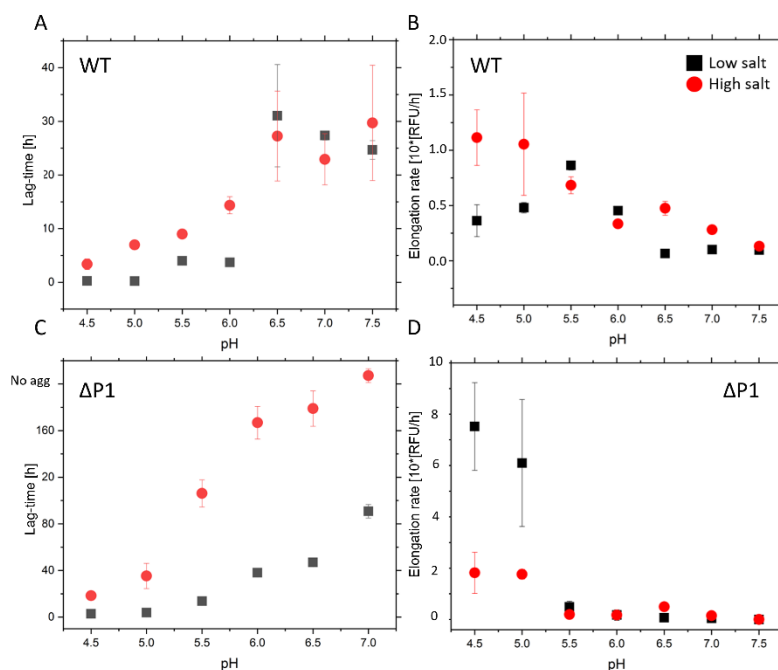


**Figure 4.3: pH dependent aggregation of  $\alpha$ Syn WT and  $\Delta$ P1.** Aggregation kinetics of  $\alpha$ Syn WT at low (20 mM) (A) and high (200 mM) (B) ionic strength at pH 4.5 to 7.5 in 0.5 steps. (C) and (D) show aggregation kinetics for  $\Delta$ P1 at low and high salt concentrations, respectively. ThT fluorescence assay was performed at 20 mM sodium acetate (pH 4.5; 5.0; 5.5), MES (pH 6.0; 6.5; 7.0) or Tris-HCl (pH 7.5) accordingly with 100  $\mu$ M protein at 37  $^{\circ}$ C, 600 rpm. Note that a constant ionic strength was considered for all buffers. Note timescale for ThT assays is 300 h. Aggregation kinetics are illustrated in Figure 4.4 and TEM images at experiment endpoints are shown in Figure 4.5.

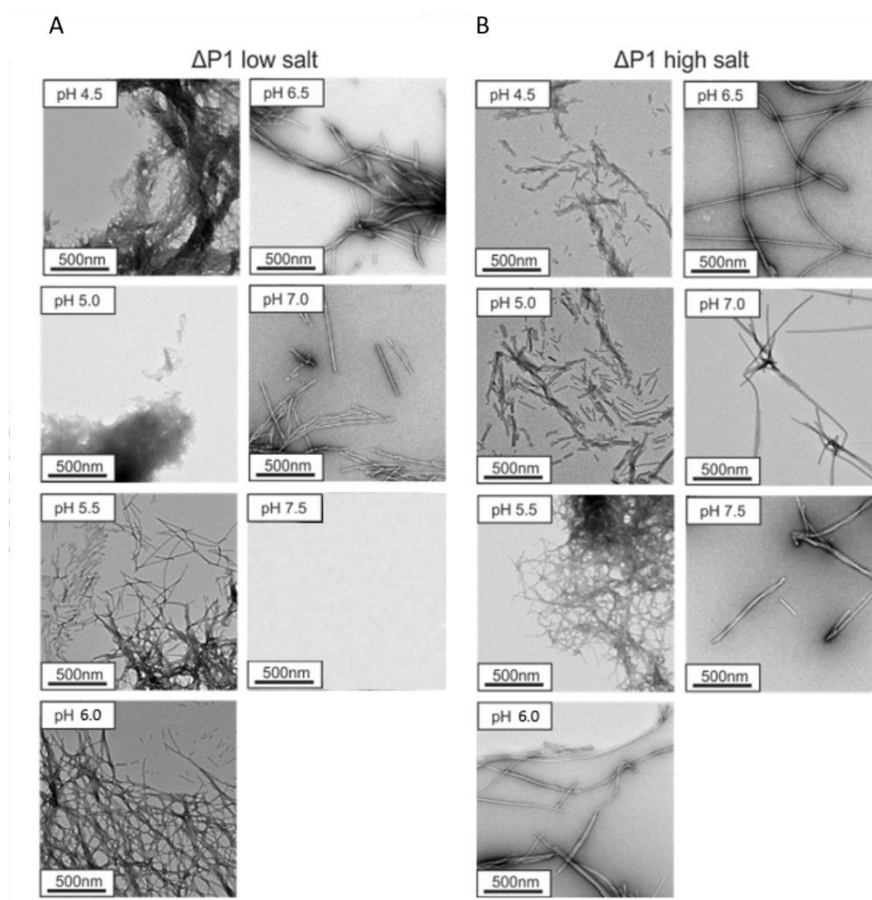
**Table 4.1: Aggregation kinetics of  $\Delta$ P1.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\Delta$ P1. The rates of aggregation in each condition were measured in at least triplicate measurements.

The errors show the standard deviation of the mean of the replicates. No aggregation after 280 h incubation is indicated by “-“. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in  $10^*$ [RFU/h].

$\alpha$ Syn variant	$\Delta$ P1				
pH; salt	4.5 20 mM	4.5 200 mM	5.0 20 mM	5.0 200mM	5.5 20 mM
Lag time [h]	$2.8 \pm 0.4$	$18.4 \pm 0.8$	$3.9 \pm 0.1$	$35.3 \pm 10.8$	$13.8 \pm 2.7$
Elongation rate*	$7.5 \pm 2.0$	$2.0 \pm 0.2$	$6.0 \pm 2.5$	$2.0 \pm 0.2$	$0.5 \pm 0.2$
$t_{50}$ [h]	$3.3 \pm 0.4$	$20.8 \pm 1.4$	$4.8 \pm 0.4$	$38.0 \pm 10.7$	$24.0 \pm 3.4$
% pellet	95	95	80	95	90
$\alpha$ Syn variant	$\Delta$ P1				
pH; salt	5.5 200 mM	6.0 20 mM	6.0 200mM	6.5 20 mM	6.5 200 mM
Lag time [h]	$106.1 \pm 11.7$	$38.1 \pm 1.3$	$166.8 \pm 14.0$	$47.0 \pm 2.0$	$178.9 \pm 15.2$
Elongation rate*	$0.2 \pm 0.02$	$0.2 \pm 0.2$	$0.2 \pm 0.02$	$0.1 \pm 0.02$	$0.5 \pm 0.08$
$t_{50}$ [h]	$144.2 \pm 21.2$	$80.6 \pm 11.0$	$193.9 \pm 14.9$	$115.3 \pm 9.0$	$189.2 \pm 16.6$
% pellet	65	40	50	40	20
$\alpha$ Syn variant	$\Delta$ P1				
pH; salt	7.0 20 mM	7.0 200mM	7.5 20 mM	7.5 200 mM	
Lag time [h]	$90.9 \pm 5.8$	$207.2 \pm 5.9$	-	-	
Elongation rate*	$0.05 \pm 0.005$	$0.1 \pm 0.02$	-	-	
$t_{50}$ [h]	$195.2 \pm 5.9$	$230.4 \pm 9.4$	-	-	
% pellet	25	5	0	0	



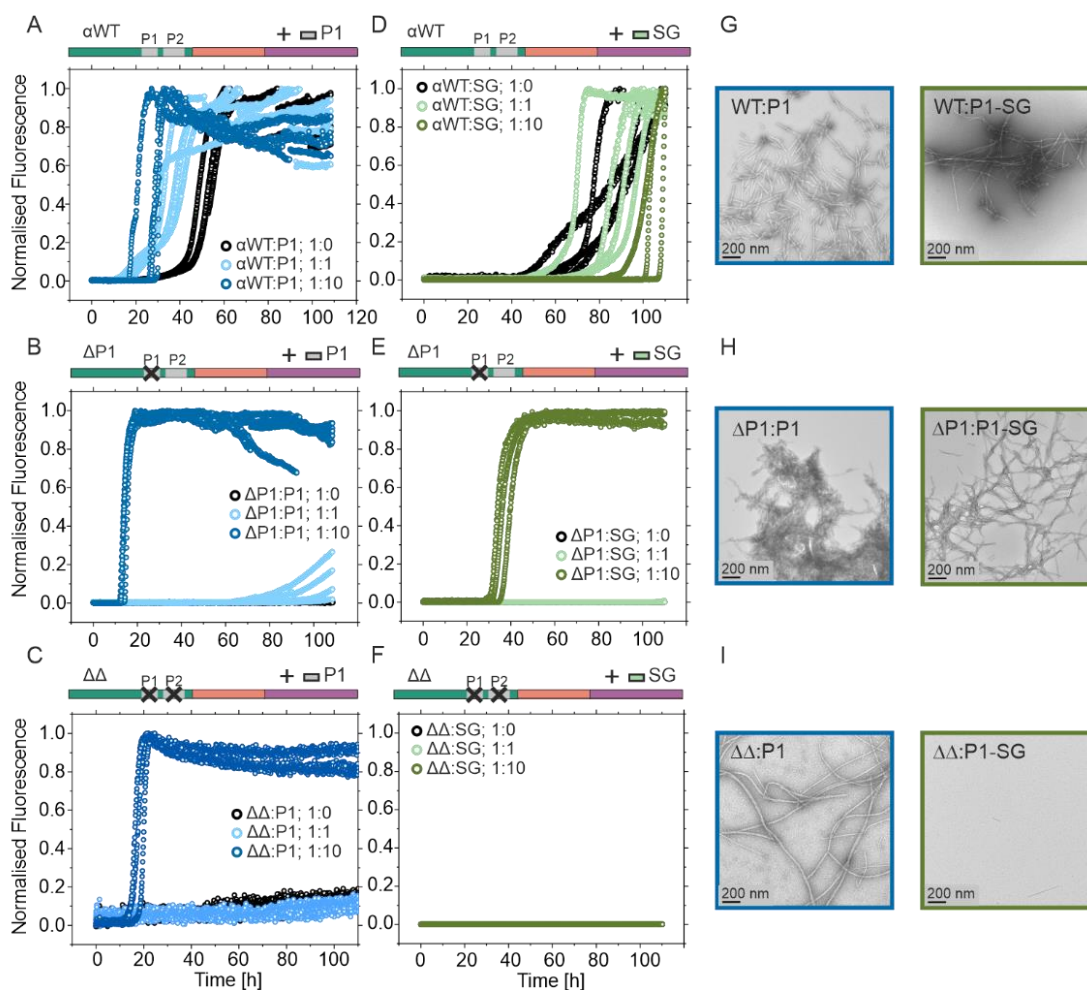
**Figure 4.4: Aggregation kinetics of pH titration for  $\alpha$ Syn WT and  $\Delta P1$ .** Lag time (A,C) and elongation rate (B,D) of WT (A,B) and  $\Delta P1$  (C,D). Data extracted from Figure 4.3 using the OriginPro software. Black datapoint show kinetics at low ionic strength (20 mM) and red points at high ionic strength (200 mM). Error bars for at least triplicates are shown.



**Figure 4.5: TEM images at the endpoint of ThT aggregation assay with  $\Delta P1$ .** TEM images for  $\Delta P1$  aggregation at low (A) and high (B) salt shown in Figure 4.3 were taken at the end of the ThT experiment, scale bar 500 nm.



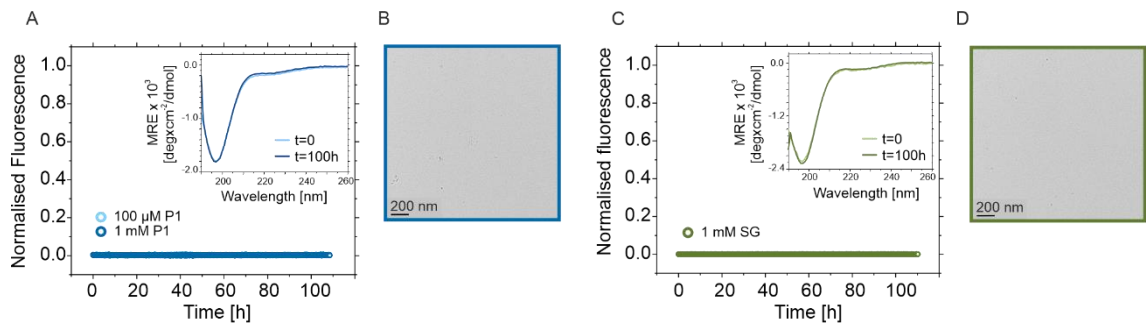
To gain a better insight into the role of the P1-sequence and how P1 exerts its effects on fibril formation, a synthetic P1-peptide was added to  $\alpha$ Syn WT,  $\Delta$ P1 and  $\Delta\Delta$  *in trans* and the aggregation kinetics were measured using a ThT fluorescence assay. As the P1 peptide alone (residues 36-42) is highly hydrophobic and was found to be insoluble in 20 mM Tris-HCl, 200 mM NaCl buffer, a N-and C-terminal extended peptide with the sequence taken from the natural  $\alpha$ Syn sequence and with an acetylated N-terminus and amidated C-terminus was used in this study (Ac-KTKE-GVLYVGS-KTKE-NH<sub>2</sub>). The ThT fluorescence assay in which the P1 peptide is mixed (1:1 or 1:10 (protein:peptide)) with  $\alpha$ Syn WT shows only a marginal effect (slight (up to ~2-fold) acceleration) on the aggregation kinetics (lag times of  $48.6 \pm 3.6$ ,  $30.6 \pm 2.2$ , and  $27.4 \pm 1.8$  h with 0, 1:1 or 1:10  $\alpha$ Syn:P1-peptide) (**Figure 4.6 A, Table 4.2**). Adding 10x excess of P1 peptide to  $\Delta$ P1 by contrast, has a drastic effect on fibril formation, changing the lag time from no aggregation in the absence of P1 peptide (20 mM Tris-HCl, pH 7.5, 200 mM NaCl) to  $11.9 \pm 0.1$  h (**Figure 4.6 B, Table 4.2**). For  $\Delta\Delta$ , the addition of 10-fold molar excess P1 peptide *in trans* also resulted in rapid aggregation (lag time =  $16.3 \pm 1.8$  h) compared to no aggregation in the absence of the peptide for the experimental time of 110 h (**Figure 4.6 C, Table 4.2**).



**Figure 4.6: Aggregation kinetics of WT  $\alpha$ Syn,  $\Delta$ P1 and  $\Delta\Delta$  in the presence of the peptides P1 or P1-SG.** Aggregation kinetics of (A,D) WT  $\alpha$ Syn, (B,E)  $\Delta$ P1 and (C,F)  $\Delta\Delta$  in the presence of different concentrations of P1-peptide (A-C) or peptide P1-GS (D-F). All experiments were carried out using 100  $\mu$ M  $\alpha$ Syn and peptide concentrations of 0  $\mu$ M, 100  $\mu$ M or 1 mM, at pH 7.5, 200 mM NaCl, 37  $^{\circ}$ C, 600 rpm. Note that under conditions of no or low aggregation, data points for different conditions overlay. Negative stain TEM images of samples with 10-fold molar excess of peptide taken at the end point of the experiment (110 h) are shown (G-I).

The end point samples of the ThT assays were used for a pelleting assay and TEM imaging and confirmed the presence of fibrils for all three experiments ( $\alpha$ Syn WT,  $\Delta$ P1 and  $\Delta\Delta$ ) when 10x P1-peptide is added (Figure 4.6 G-I, left side, Table 4.2). The morphology, however, is distinct from  $\alpha$ Syn aggregating by itself with short and clumpy fibrils, indicating a potentially different mechanism of aggregation. Control experiments with the P1 peptide in isolation showed no amyloid formation measured via ThT fluorescence, far UV CD and negative stain TEM (Figure 4.7 A,B). These data demonstrate that the P1-peptide is able to accelerate  $\alpha$ Syn aggregation possibly by interacting with the protein, replacing the effect of the P1 sequence on intra-/inter-molecular interactions described in Chapter 3 *in trans*. Further control experiments were performed with a peptide in which the P1 sequence is replaced with a Ser-Gly-linker, P1-SG (Ac-KTKE-SGSGSGS-KTKE-NH<sub>2</sub>). Likewise, this peptide does not aggregate in isolation (Figure 4.7 C,D) but, in contrast to the P1-peptide, it shows less of an effect

when added to  $\alpha$ Syn WT,  $\Delta$ P1 or  $\Delta\Delta$  *in trans* (Figure 4.6 D-F, Table 4.2). For  $\alpha$ Syn WT no change (possibly a slight retardment) in aggregation is detected, in the case of  $\Delta$ P1 less efficient induction of aggregation was observed (lag time of  $30.6 \pm 1.1$  and  $11.9 \pm 0.1$  h, for P1-SG and P1-peptide, respectively) (Table 4.2) and adding P1-SG to  $\Delta\Delta$  did not induce fibrils to form within the experimental time of 110 h. These observations add further weight to the role of the specific sequence of P1 in driving aggregation, with the rate of aggregation depending both on the sequence of the peptide and the protein to which it was added.



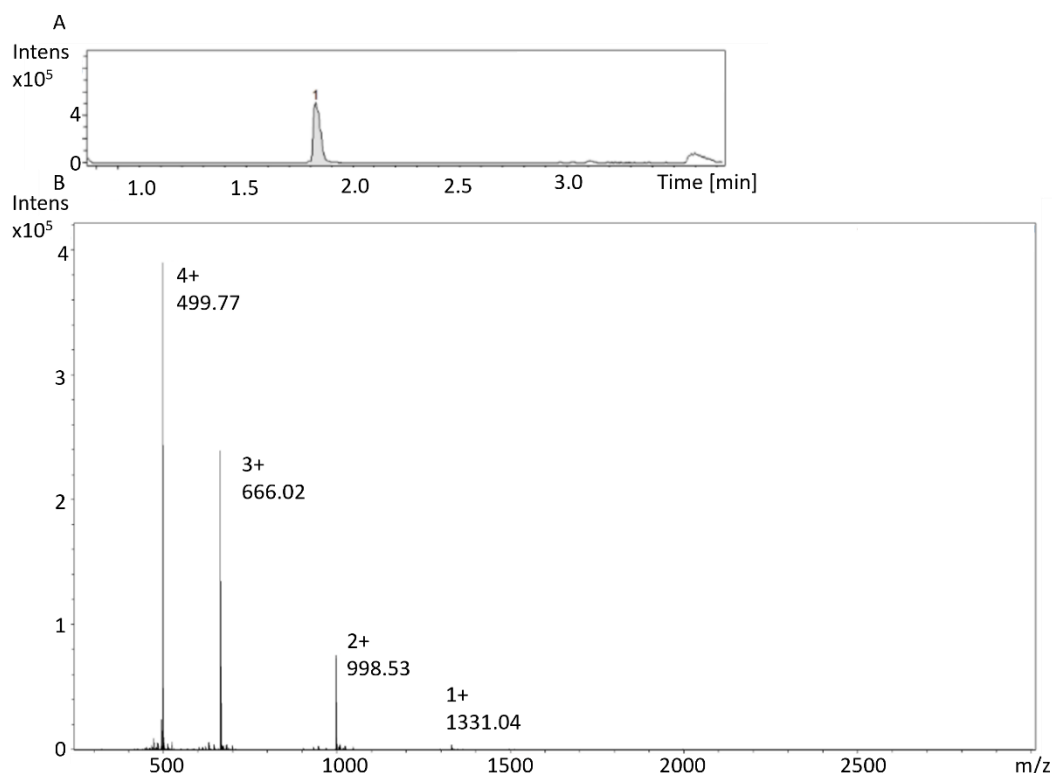
**Figure 4.7: The synthetic peptides P1 and P1-SG do not aggregate in isolation under the conditions employed.** Analysis of (A,B) the P1 peptide or (C,D) the peptide P1-SG in isolation (100  $\mu$ M or 1 mM (pH 7.5, 200 mM NaCl, 37°C, 600 rpm)) by ThT fluorescence (A,C), far UV CD (inset) and negative stain TEM (B,D) shows that the peptides do not aggregate in isolation under the conditions employed. Note: data points for 100  $\mu$ M and 1mM P1 peptide overlay in (A).

**Table 4.2: Aggregation kinetics of  $\alpha$ Syn WT,  $\Delta$ P1, and  $\Delta\Delta$  in the presence of P1- or P1-SG peptides.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\alpha$ Syn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 110 h incubation is indicated by “-”. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in  $10^*$ [RFU/h].

	WT: peptide P1 1:0	WT: peptide P1 1:1	WT: peptide P1 1:10	WT: peptide P1-SG 1:1	WT: peptide P1-SG 1:10
Lag time [h]	$48.6 \pm 3.6$	$30.6 \pm 2.2$	$27.4 \pm 1.8$	$74.8 \pm 7.7$	$91.7 \pm 8.1$
Elongation rate*	$0.38 \pm 0.07$	$0.7 \pm 0.1$	$3.1 \pm 2.1$	$0.6 \pm 0.1$	$1.0 \pm 0.3$
$t_{50}$ [h]	$51.3 \pm 2.2$	$40.1 \pm 7.2$	$26.6 \pm 3.9$	$81.6 \pm 9.2$	$80.0 \pm 10.9$
% pellet	65	40	65	65	15
	$\Delta$ P1: peptide P1 1:0	$\Delta$ P1: peptide P1 1:1	$\Delta$ P1: peptide P1 1:10	$\Delta$ P1: peptide P1-SG 1:1	$\Delta$ P1: peptide P1-SG 1:10
Lag time [h]	-	-	$11.9 \pm 0.1$	-	$30.6 \pm 1.1$
Elongation rate*	-	-	$2.2 \pm 0.4$	-	$0.9 \pm 0.04$
$t_{50}$ [h]	-	-	$14.4 \pm 0.7$	-	$37.0 \pm 2.5$
% pellet	0	0	70	0	50
	$\Delta\Delta$ : peptide P1 1:0	$\Delta\Delta$ : peptide P1 1:1	$\Delta\Delta$ : peptide P1 1:10	$\Delta\Delta$ : peptide P1-SG 1:1	$\Delta\Delta$ : peptide P1-SG 1:10
Lag time [h]	-	-	$16.3 \pm 1.8$	-	-
Elongation rate*	-	-	$2.9 \pm 0.5$	-	-
$t_{50}$ [h]	-	-	$20.1 \pm 1.5$	-	-
% pellet	5	10	90	5	5

### 4.3 Binding mode between $\alpha$ Syn and P1-peptide

In a next step, the binding mode of  $\alpha$ Syn and P1-peptide, as well as the mechanism by which the peptide increases the rate of fibril formation, was investigated in more detail. To define the interaction hubs between  $\alpha$ Syn WT and P1-peptide  $^1\text{H}$ - $^{15}\text{N}$ -HSQC NMR experiments were performed with  $^{15}\text{N}$  labelled  $\alpha$ Syn and (non-labelled)  $^{14}\text{N}$  P1-peptide. The interaction interface(s) were determined using HN-chemical shift perturbation (HN-CSP) of protein alone versus in the presence of peptide (**Figure 4.9 A, E**). HN-CSP analysis is a common technique to determine binding sites with ligands<sup>448</sup> (see Section 2.9.4). When adding a 10-fold molar excess of P1-peptide to  $\alpha$ Syn WT significant HN-CSPs were measured, corresponding to residues in the extreme N-terminus (~15 residues), the P2 region (residues 45-57) and the C-terminal 40 residues. Performing the experiment with P1-SG peptide instead, showed a similar HN-CSP pattern but with significantly reduced intensities especially for the C-terminal region (**Figure 4.9 B**). The observed changes, which are also dependent on peptide concentration (**Figure 4.9 A**), could arise from direct binding of the peptides to  $\alpha$ Syn at these identified sites, or from indirect effects such as conformational changes in regions distant from the binding site. For a better interpretation of the interaction site, NMR PRE experiments were performed with  $^{15}\text{N}$  labelled  $\alpha$ Syn WT and MTSL-labelled ( $^{14}\text{N}$ ) P1-peptide. For this, the P1 peptide with an additional cysteine at the N-terminus (Acetyl-**C**-KTKE-GVLYVGS-KTKE-amide) was labelled with MTSL before separating the labelled peptide from excess MTSL by HPLC and confirming complete modification of the peptide with MTSL by ESI-MS (**Figure 4.8**) (Section 2.7.7). The purification and characterisation of the labelled peptide was performed by Yong Xu, University of Leeds.

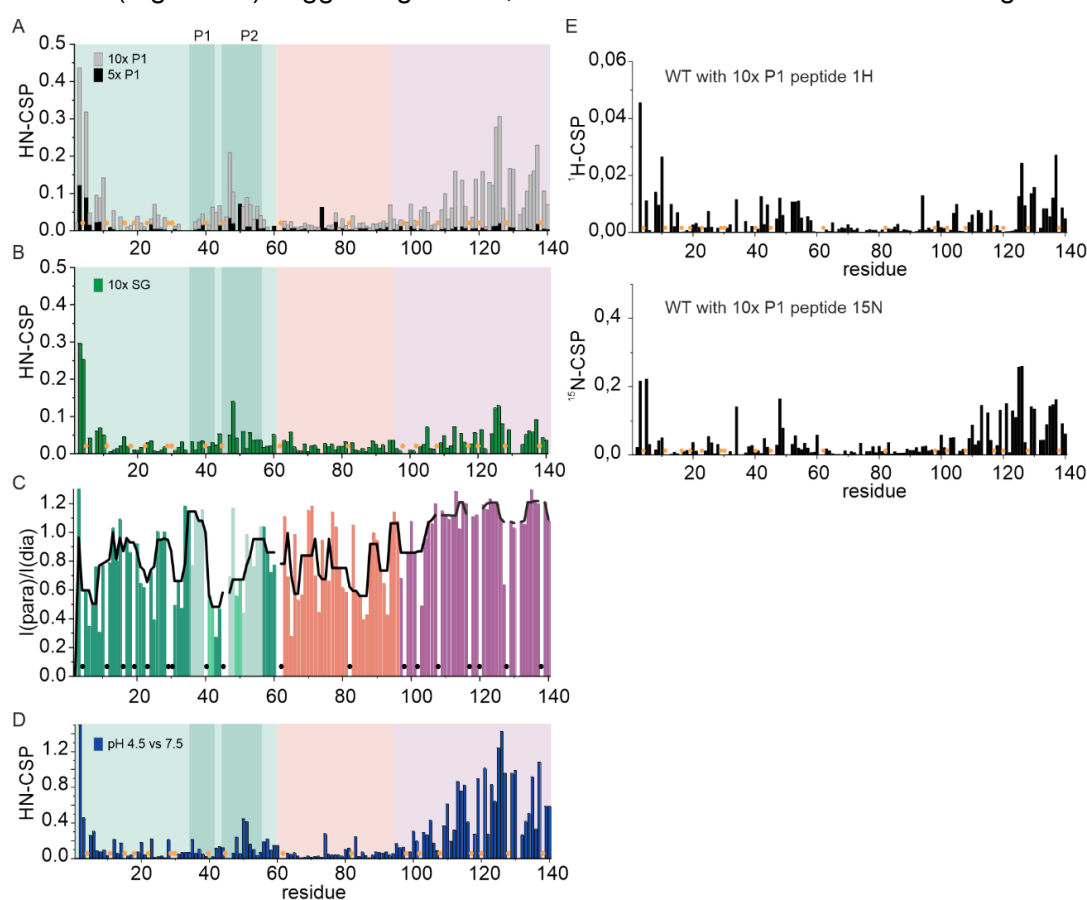


**Figure 4.8: Purification and characterisation of MTSL-labelled Cys-P1 peptide by LC-MS.**

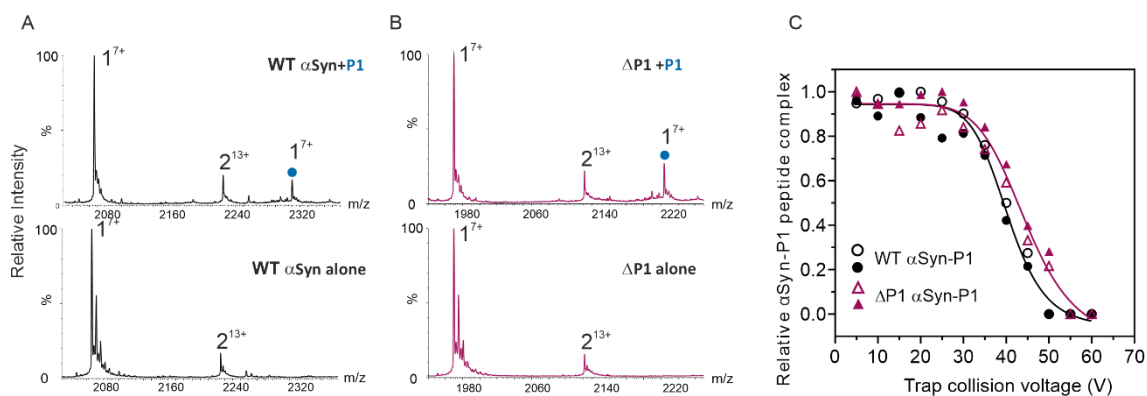
(A) The Cys-P1 peptide (Acetyl-C-KTKE-GVLYVGS-KTKE-amide) was purified from excess MTSL by separating the labelled peptide via HPLC (see 2.7.7) and (B) complete modification of the peptide with MTSL was confirmed by ESI-MS. The mass of the purified peptides by ESI-MS was 1995.07 Da (expected = 1995.03 Da). The purification end characterisation was performed by Yong Xu, University of Leeds.

As discussed in Chapter 2 and 3, NMR PRE experiments are ideal for detecting transient interactions between a nucleus and a spin label with an unpaired electron with distance sensitivity of  $\sim 20\text{-}40 \text{ \AA}$ <sup>424,425</sup>. NMR PRE experiments revealed that increased relaxation was found in the first  $\sim 100$  N-terminal residues when MTSL labelled peptide was mixed with equimolar  $\alpha$ Syn WT, suggesting multiple possible binding sites in N-terminal and NAC region of  $\alpha$ Syn (**Figure 4.9 C**). Especially in NAC a periodicity can be observed in the PRE data that might hint to a helix formation when binding the peptide. For a final conclusion for that, further experiments would be required such as CD. To evaluate the binding stoichiometry, native nESI-MS experiments were performed (**Figure 4.10 A**) (experiment performed by Emily Byrd, University of Leeds). A 1:1 binding mode was observed between  $\alpha$ Syn WT and P1 peptide (mass of the protein-peptide complex  $16,167.54 \pm 0.17$  Da; theoretical mass 16,168 Da), but no higher order binding processes were detected. This indicates that only one peptide per protein is binding at different locations within the first  $\sim 100$  residues. Interestingly, the PRE experiments did not show any binding in the C-terminal region, although one might expect binding between the mainly positively charged KTKE motifs (added to enhance solubility of the peptides) and the acidic C-terminus. Therefore, changes in HN-CSPs must be a consequence of alterations in long-range intramolecular interactions between the N-terminal and/or NAC

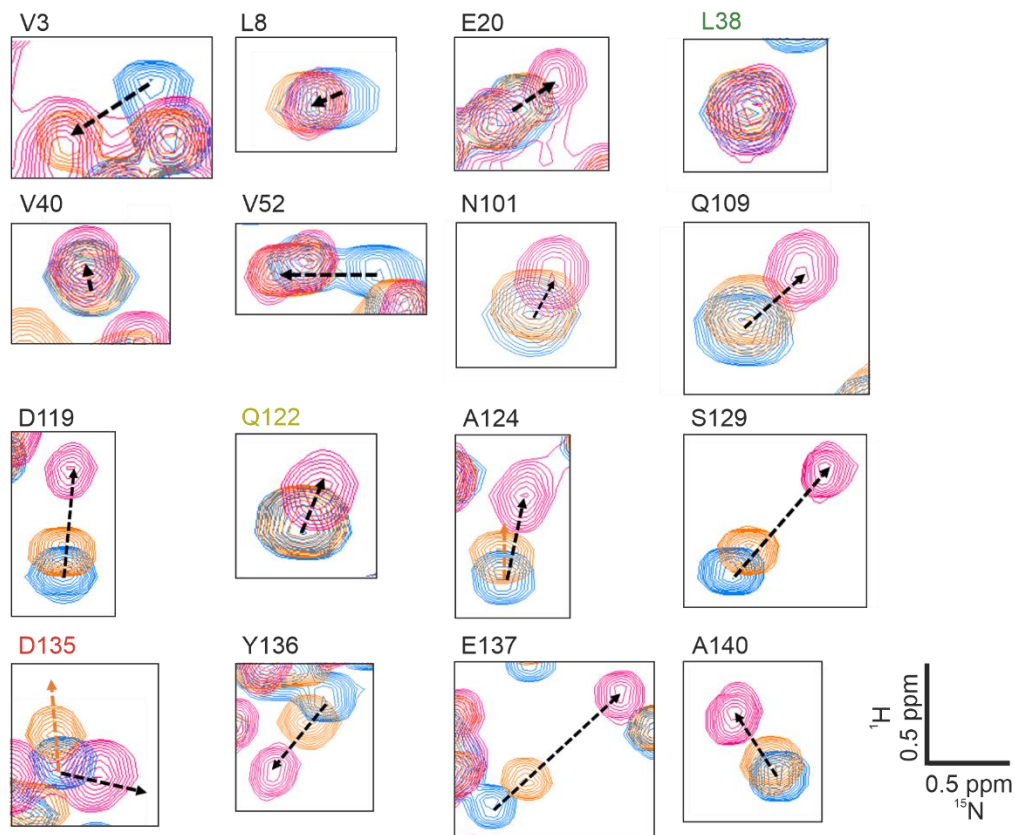
region with the C-terminus upon peptide binding. Note that at the concentrations used here (100  $\mu\text{M}$   $\alpha\text{Syn}$ ) no intermolecular interactions are detected (see **Figure 3.17**). The important role of interactions between the N-terminal, NAC and C-terminal regions for regulating aggregation has been described before<sup>189,321,335</sup> and the observations here are hence consistent with the literature. Strikingly, HN-CSPs when  $\alpha\text{Syn}$  WT binds the P1-peptide are very similar to changes observed when decreasing the pH from 7.5 to 4.5 (**Figure 4.9 D**) which also accelerates fibril formation (**Figure 4.3**). Further analysis of the HN-CSPs of individual residues (**Figure 4.11**) show that most cross-peaks move on the same vector when comparing the CSPs with pH or upon P1 addition. However, there are exceptions in which no HN-CSPs are observed for the addition of the peptide, but are observed upon changes in pH (e.g. Q122), or where chemical shifts move in opposite directions (e.g. D135) suggesting similar, but not identical conformational changes.



**Figure 4.9: The binding of P1-peptide alters the conformational ensemble of WT  $\alpha\text{Syn}$ .** HN-CSPs of WT  $\alpha\text{Syn}$  upon the addition of (A) a 5- or 10-fold molar excess of peptide-P1, (B) a 10-fold molar excess of peptide P1-SG and (C) NMR PREs of  $^{15}\text{N}$ - $\alpha\text{Syn}$  upon the addition of equimolar MTSL-labelled peptide-P1. The black line represents the median value over a rolling window of 5 residues. The signal to noise level is  $>2.11$  (dia) and  $>2.28$  (para). (D) Difference in chemical shifts of WT  $\alpha\text{Syn}$  at pH 7.5 and 4.5. (E) Separate  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts for WT with 10x peptide shown in combination in (A). All experiments were performed in 20 mM Tris-HCl buffer, pH 7.5, 200 mM,  $15^\circ\text{C}$ . The N-terminal region is shaded green, NAC is in red and the C-terminal region is in purple. The P1 and P2 regions are highlighted in darker green. Residues 4, 11, 13, 16, 19, 23, 29, 30, 41, 47, 62, 82, and 112 were not included because of overlapping peaks; residues 1, 108, 117, 120, 128, and P138 were not assigned.

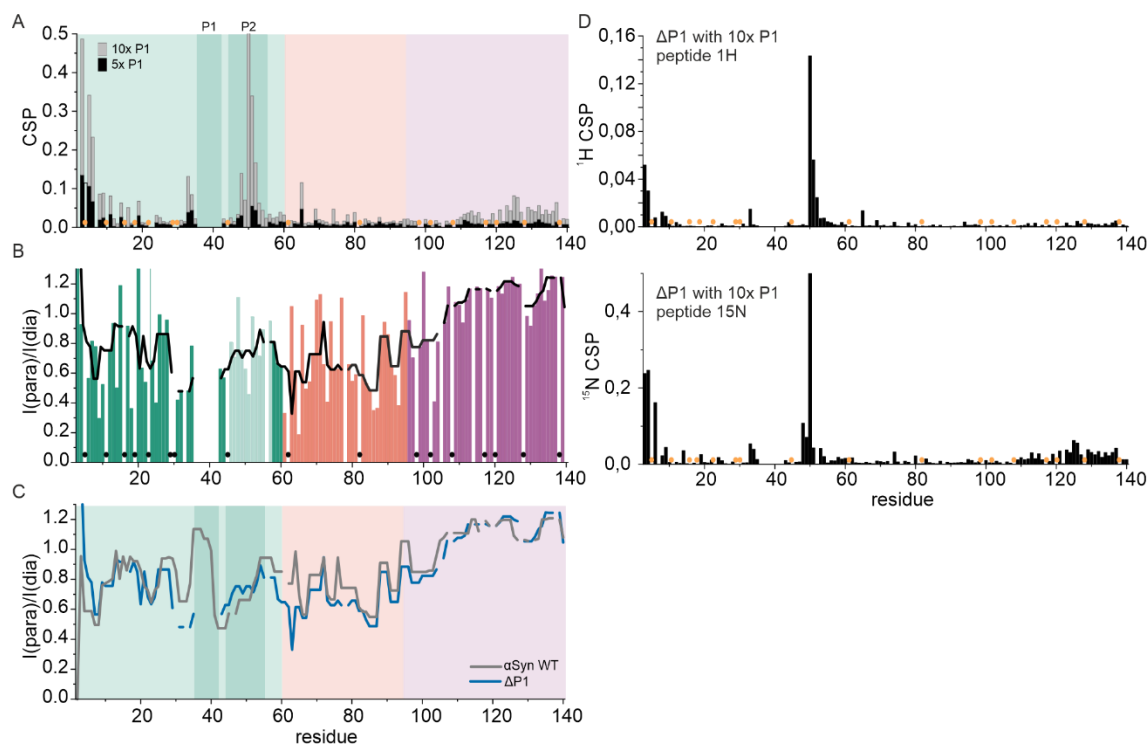


**Figure 4.10: Interaction between  $\alpha$ Syn WT and  $\Delta$ P1 with the P1 peptide characterised by native nESI-MS. (A,B)** Positive-ion nESI mass spectra of  $\alpha$ Syn WT (A) and  $\Delta$ P1 (B) in the presence of P1-peptide (top) and without (bottom). (C) Collision induced dissociation of WT  $\alpha$ Syn-P1 complex (black) and  $\Delta$ P1-P1 complex (pink) with the intensity of the complex relative to the apo-protein plotted versus the trap cell voltage. The experiment was performed in duplicates (as indicated in the inset legend) and the data fitted to a sigmoid function (solid line). This experiment was performed and analysed by Emily Byrd, University of Leeds.



**Figure 4.11: Cross-peaks of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra to identify differences between pH titration and P1-peptide addition for  $\alpha$ Syn WT.** Peaks for  $\alpha$ Syn WT in 20 mM Tris HCl pH 7.5 (blue),  $\alpha$ Syn WT in 20 mM sodium acetate pH 4.5 (pink) and  $\alpha$ Syn WT with 10x P1 peptide (orange). Black arrows indicate direction of HN-CSPs. Each box is labelled with the corresponding residue, black residue label: changes occur along a similar vector, green residue label: no change for any of the conditions, yellow: no change for one of the conditions but for the other, red: distinct HN-CSPs for the pH titration and peptide addition. These data indicate that similar, but not identical, conformational changes occur under the two conditions.

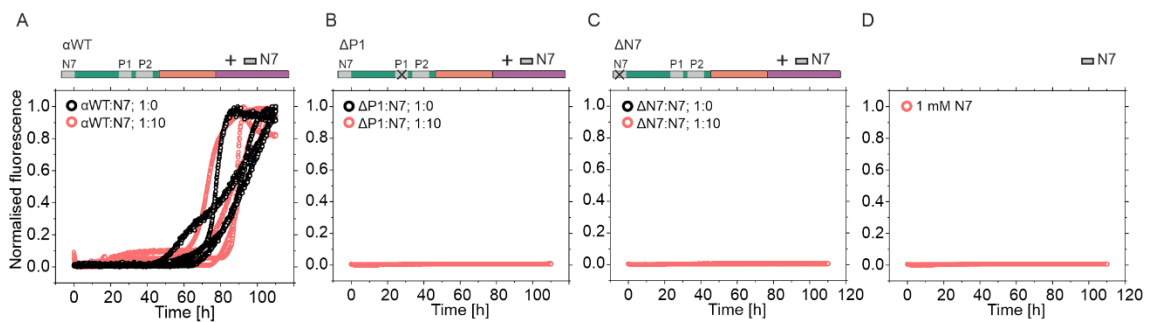
The P1 peptide has a more drastic effect on  $\Delta$ P1 aggregation compared with  $\alpha$ Syn WT (**Figure 4.6**). Hence the binding mode of  $\Delta$ P1 and P1-peptide was investigated in more detail by NMR and nESI-MS. Detecting the HN-CSPs indicated a similar pattern of chemical shifts upon binding peptide P1 compared with  $\alpha$ Syn WT with changes in the N-terminal 15 amino acids, P2 region and C-terminal region, but with distinct relative amplitudes. Whilst HN-CSPs in the P2 region are increased, effects observed in the C-terminal region are decreased compared with binding with  $\alpha$ Syn WT (compare **Figure 4.9 A** and **Figure 4.12 A, D**). PRE NMR and nESI-MS analysis further revealed binding of the P1 peptide throughout the N-terminal and NAC region (~100 residues) of  $\Delta$ P1 with a 1:1 binding mode as seen for  $\alpha$ Syn WT (observed mass for the complex of  $15,492.23 \pm 0.25$  Da; theoretical mass 15,492 Da) (**Figure 4.12 B,C** and **Figure 4.10 B**). Further, collision-induced dissociation (CID) experiments were used to determine the interaction strength of the protein-peptide complex and showed a slightly tighter binding of P1-peptide to  $\Delta$ P1 compared to  $\alpha$ Syn WT with  $IC_{50}$  values of 44  $\mu$ M and 40  $\mu$ M, respectively (**Figure 4.10 C**).



**Figure 4.12: Binding of peptide P1 alters the conformational ensemble of  $\Delta$ P1.** (A) HN-CSPs and (B)  $^1$ H- $^{15}$ N PRE effects upon addition of peptide-P1 to 100  $\mu$ M  $^{15}$ N-labelled  $\Delta$ P1 (pH 7.5, 200 mM, 15°C). The black line represents the median value over a rolling window of 5 residues (here using equimolar peptide P1: $\Delta$ P1). The signal to noise level is  $>3.37$  (dia) and  $>1.86$  (para). (C) Comparison of the median value PREs of WT  $\alpha$ Syn and  $\Delta$ P1 upon the addition of equimolar P1-peptide are compared to highlight their similar profiles. (D) Separate  $^1$ H and  $^{15}$ N chemical shifts for  $\Delta$ P1 with 10x peptide shown in combination in (A). Note that residues 36-42 are deleted in  $\Delta$ P1. The N-terminal region is shaded light green, NAC is in red and the C-terminal region in purple. The P1 and P2 regions are highlighted in darker green. Residues 4, 11, 13, 16, 19, 23, 29, 30, 47, 62, 82, and 112 were not included because of overlapping peaks; residues 1, 108, 117, 120, 128, and P138 were not assigned.



In Chapter 3 it was also shown that the deletion of the N-terminal seven residues of  $\alpha$ Syn (residues 2-7) results in inhibition of aggregation at neutral pH, similar to the effect observed for  $\Delta$ P1. Therefore, it was tested whether the addition of a peptide designed accordingly to the N7 sequence (N7-peptide: NH<sub>2</sub>-MDVFMKG-KTKE-amide) *in trans* is also able to enhance aggregation of  $\alpha$ Syn WT,  $\Delta$ P1 or  $\Delta$ N7 (**Figure 4.13 A-C, Table 4.3**). Interestingly, none of the protein variants were affected by the presence of 10-fold molar excess N7-peptide and no aggregation was observed for the peptide in isolation (**Figure 4.13 D**) further highlighting the important, exceptional role of the P1-sequence in controlling protein assembly. This result also hints to different mechanisms of inhibition between  $\Delta$ N7 and  $\Delta$ P1. A more detailed analysis of the N7-peptide would be required to gain further information of its binding and differences compared to the P1-peptide.



**Figure 4.13: Aggregation kinetics of WT  $\alpha$ Syn,  $\Delta$ P1 and  $\Delta$ N7 in the presence of the N7-peptide.** Aggregation kinetics of (A) WT  $\alpha$ Syn, (B)  $\Delta$ P1 and (C)  $\Delta$ N7 in the presence of 10-fold molar excess of N7-peptide. (D) Aggregation kinetics of N7-peptide in isolation. Experiments were carried out using 100  $\mu$ M  $\alpha$ Syn and peptide concentrations of 0  $\mu$ M or 1 mM, at pH 7.5, 200 mM NaCl, 37  $^{\circ}$ C, 600 rpm. Note that under conditions of no or low aggregation, data points for different conditions overlay.

**Table 4.3: Aggregation kinetics in the presence of N7 peptide.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\alpha$ Syn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 110 h incubation is indicated by “-”. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in 10\*[RFU/h].

	WT: peptide N7 1:0	WT: peptide N7 1:10	$\Delta\Delta$ : peptide N7 1:0	$\Delta\Delta$ : peptide N7 1:10	$\Delta$ N7: peptide N7 1:0	$\Delta$ N7: peptide N7 1:10
Lag time [h]	67.3 $\pm$ 5.2	71.1 $\pm$ 9.3	-	-	-	-
Elongation rate*	0.5 $\pm$ 0.2	0.7 $\pm$ 0.1	-	-	-	-
$t_{50}$ [h]	81.7 $\pm$ 10.8	79.0 $\pm$ 3.4	-	-	-	-
% pellet	70	65	0	0	5	0

In summary, these findings show that the P1-peptide switches on amyloid formation of  $\Delta$ P1 and  $\Delta\Delta$  when adding in 10-fold excess and accelerates aggregation for  $\alpha$ Syn WT, presumably by competing with long-range intra-molecular interactions between the N- and C-terminal regions of  $\alpha$ Syn that have been shown to protect the protein from aggregation in the literature<sup>189,335</sup>. Variations in the chemical shifts of  $\alpha$ Syn WT versus  $\Delta$ P1 with P1-peptide might suggest distinct molecular mechanisms, but nevertheless the

P1 sequence could be shown to drastically increase fibril formation for  $\alpha$ Syn when added *in trans* in contrast to P1-SG (mild effect) and especially the N7-peptide (no effect).

## 4.4 Specific residues within P1 are responsible for controlling aggregation in $\alpha$ Syn

After identifying and studying the role of the P1 sequence in  $\alpha$ Syn aggregation, the next focus was put on the contribution of individual residues in this seven residue motif. Literature data have shown before that single residues throughout the sequence of  $\alpha$ Syn can play important roles in modulating fibril formation *in vitro*, in cells and *in vivo*<sup>145,298,409</sup> as well as alter their membrane binding properties<sup>268,361</sup>.

### 4.4.1 Alanine scan to identify key residues in P1

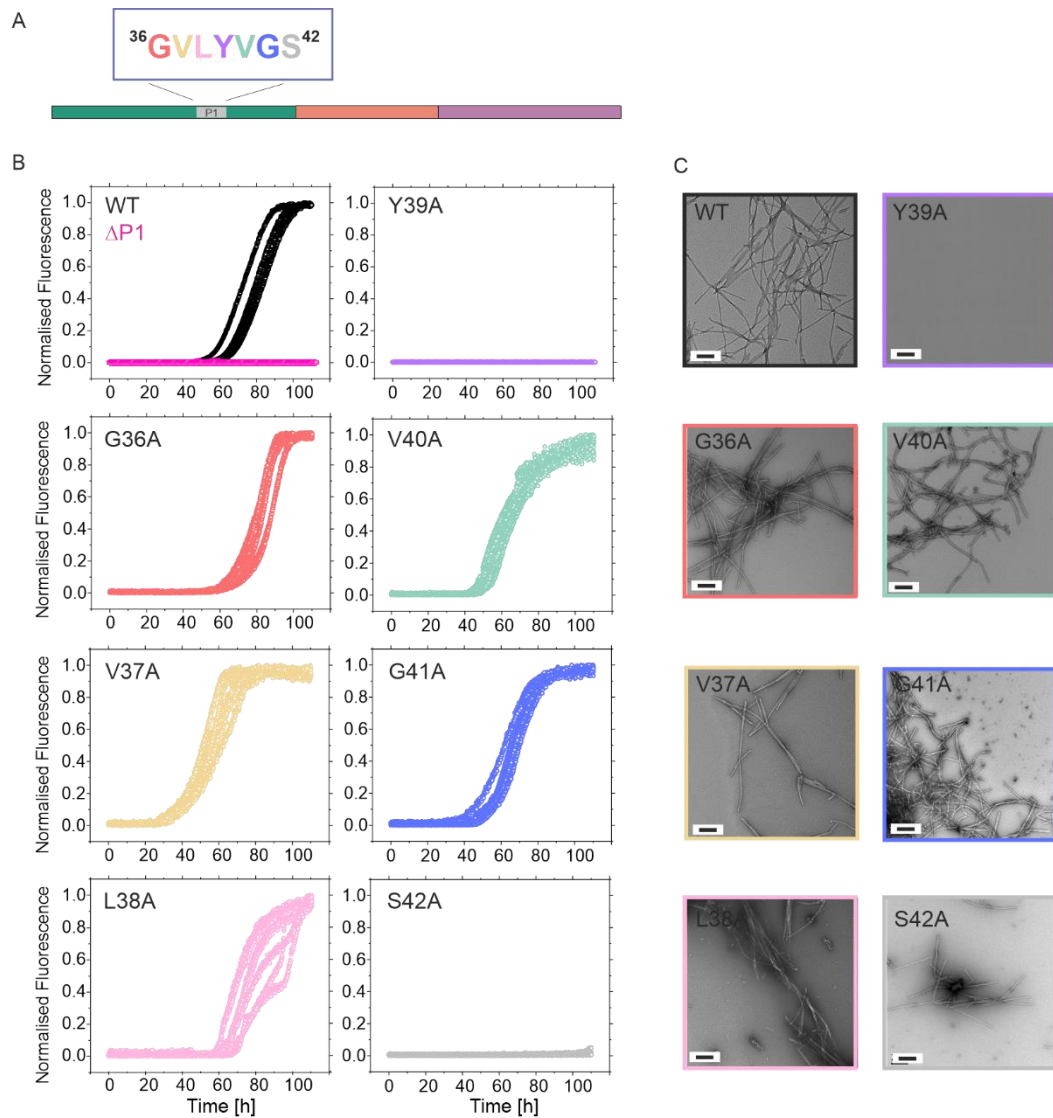
For a closer investigation of the role of the P1 region in  $\alpha$ Syn aggregation, an alanine scan was performed, mutating each residue (<sup>36</sup>GVLVYVGS<sup>42</sup>) individually to an alanine. All alanine scan variants were generated as described in Section 2.5.1 and expressed and purified as described for the  $\alpha$ Syn WT protein (see Section 2.6.1) confirming successful protein production by MS (Rachel George, MS facility, University of Leeds) (Table 4.4).

**Table 4.4: All purified  $\alpha$ Syn and  $\gamma$ Syn variants, expected molecular masses (top) and measured molecular masses (via MS, bottom) of the proteins used in this chapter.**

Variant	Mass	Variant	Mass	Variant	Mass
$\alpha$ WT	14,460 Da	$\alpha$ L38A	14,418 Da	$\gamma$ WT	13,331 Da
	14459 $\pm$ 0.3 Da		14,417 $\pm$ 0.05 Da		13,330 $\pm$ 0.2 Da
$\Delta$ P1	13,784 Da	$\alpha$ L38M	14,478 Da	$\gamma$ M38L	13,313 Da
	13,784 $\pm$ 0.5 Da		14,478 $\pm$ 0.7 Da		13,313 $\pm$ 0.09 Da
$\Delta\Delta$	12,506 Da	$\alpha$ L38I	14,460 Da	$\gamma$ A42S	13,347 Da
	12,507 $\pm$ 0.03 Da		14,461 $\pm$ 0.1 Da		13,346 $\pm$ 0.3 Da
$\Delta$ N7	13,782 Da	$\alpha$ Y39A	14,368 Da	$\gamma$ M38L/A42S	13,329 Da
	13,781 $\pm$ 0.2 Da		14,366 $\pm$ 0.03 Da		13,329 $\pm$ 0.03 Da
P1-GS	14,304 Da	$\alpha$ V40A	14,432 Da	YAA	14,688 Da
	14,303 $\pm$ 0.3 Da		14,430 $\pm$ 0.1 Da		14,686 $\pm$ 0.7 Da
$\alpha$ G36A	14,474 Da	$\alpha$ G41A	14,474 Da	YAY	13,088 Da
	14,473 $\pm$ 0.1		14,473 $\pm$ 0.2 Da		13,089 $\pm$ 0.1 Da
$\alpha$ V37A	14,432 Da	$\alpha$ S42A	14,444 Da	YYA	14,932 Da
	14,431 $\pm$ 0.2 Da		14,444 $\pm$ 0.5 Da		14,932 $\pm$ 0.4 Da

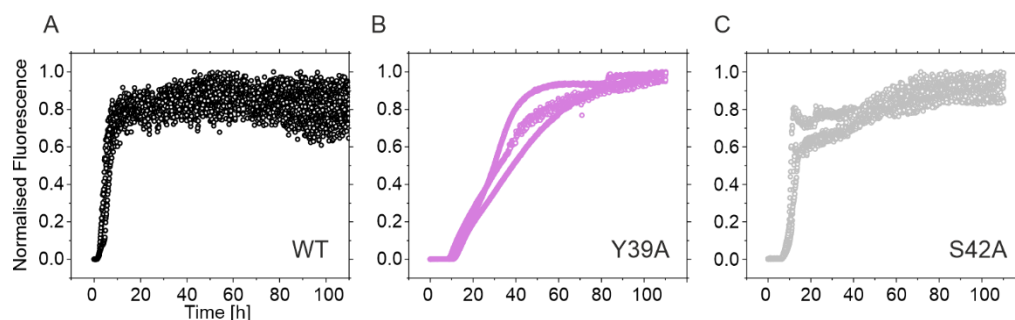
ThT fluorescence aggregation assay was performed to measure the aggregation kinetics of the seven created alanine scan variants in comparison to  $\alpha$ Syn WT and  $\Delta$ P1 in 20 mM Tris-HCl, pH 7.5, 200 mM NaCl (Figure 4.14). Five of the tested variants had no/little effect on the aggregation rates resulting in fibril growth with similar kinetics (G36A, L38A) or slightly faster aggregation (V37A, V40A, and G41A) than  $\alpha$ Syn WT. Y39A on the other

hand did not aggregate within 110 h, consistent with previous literature reports<sup>440</sup> (**Figure 4.14**, purple). Interestingly, there was a second mutant inhibiting aggregation, S42A, which had not been described hitherto (**Figure 4.14**, grey). The presence or absence of fibrils was confirmed by analysis of the amount of pelletable material and negative stain TEM with the end-point samples of the ThT assay (**Figure 4.14** and **Table 4.5**).



**Figure 4.14: Alanine scan to identify residues in P1 region important for aggregation. (A)** Schematic showing the sequence of the P1 region of  $\alpha$ Syn. **(B)** Aggregation kinetics of the seven Ala variants in the P1 region of  $\alpha$ Syn determined using 100  $\mu$ M  $\alpha$ Syn at 37°C, 20 mM Tris-HCl, 200 mM NaCl, pH 7.5, 600 rpm, coloured as in **(A)**. Data for  $\alpha$ Syn WT and  $\Delta$ P1 are shown for comparison. **(C)** TEM images taken at the end point of each experiment. Scale bar = 200 nm.

The previous experiments clearly highlight that single residue substitution to an alanine at two positions (Y39A and S42A) within the P1 region abolish aggregation at neutral pH (at least for the experimental time of 110 h). A ThT fluorescence assay at acidic pH demonstrates that these variants do form fibrils at pH 4.5 (**Figure 4.15**), showing a similar pH dependence observed for  $\Delta$ P1 (**Figure 4.3**).



**Figure 4.15: Aggregation kinetics of  $\alpha$ Syn WT and alanine variants Y39A and S42A at acidic pH.** ThT fluorescence assay of  $\alpha$ Syn WT (A), Y39A (B) and S42A (C) was performed at 20 mM sodium acetate, pH 4.5, 200 mM NaCl, 37 °C, 600 rpm.

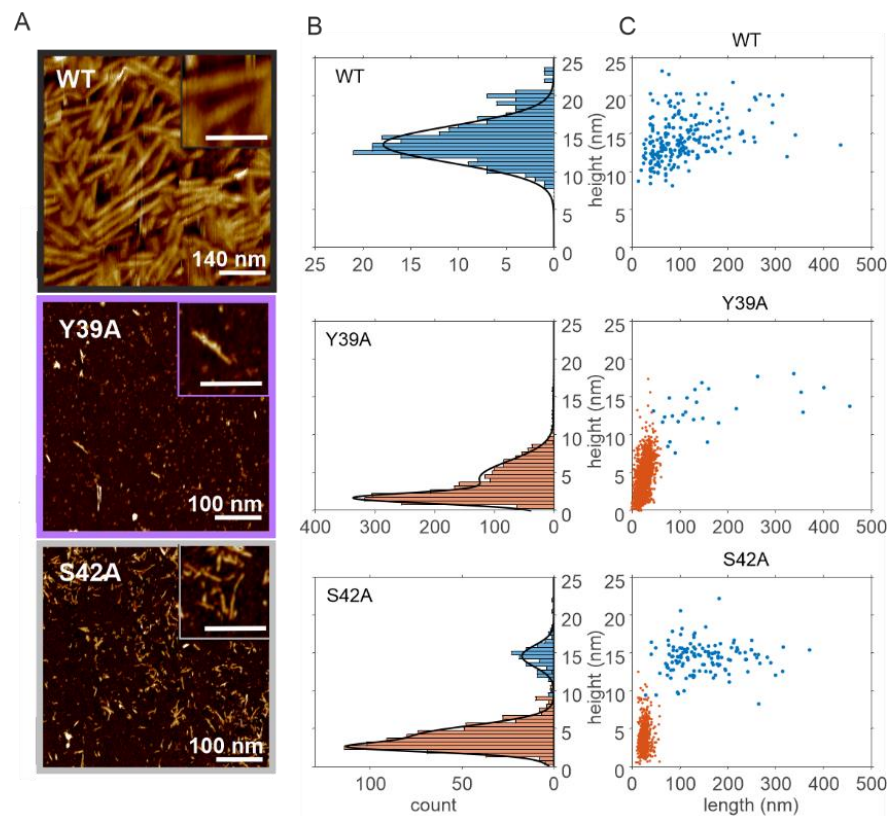
**Table 4.5: Aggregation kinetics of alanine scan of P1 region in  $\alpha$ Syn.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\alpha$ Syn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 110 h incubation for *de novo* growth or 42 h for seeded reactions is indicated by “-“. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in  $10^*$ [RFU/h].

De novo growth, pH 7.5								
	WT	G36A	V37A	L38A	Y39A	V40A	G41A	S42A
Lag time [h]	65.0 ± 4.0	72.6 ± 4.8	40.2 ± 3.4	56.7 ± 6.7	-	42.5 ± 4.2	53.6 ± 3.5	-
Elongation rate*	0.38 ± 0.07	0.51 ± 0.1	0.33 ± 0.09	0.30 ± 0.07	-	0.24 ± 0.03	0.4 ± 0.03	-
$t_{50}$ [h]	79.8 ± 6.0	83.3 ± 4.0	57.5 ± 7.2	74.6 ± 10.0	-	61.8 ± 5.5	66.1 ± 3.0	-
% pellet	65	65	75	75	30	40	65	15
De novo growth, pH 4.5								
	WT	Y39A	S42A					
Lag time [h]	2.3 ± 0.1	11.2 ± 0.9	8.3 ± 1.4					
Elongation rate*	3.1 ± 0.2	0.1 ± 0.05	0.5 ± 0.1					
$t_{50}$ [h]	6.8 ± 0.2	39.0 ± 5.1	17.5 ± 2.0					
Seeded growth ( $\alpha$ Syn WT seeds)								
	WT	G36A	V37A	L38A	Y39A	V40A	G41A	S42A
Lag time [h]	0	0	0	0	0	0	0	0
Elongation rate*	2.1 ± 0.2	2.9 ± 1.0	0.1 ± 0.07	0.87 ± 0.08	2.2 ± 0.2	0.52 ± 0.04	1.7 ± 0.1	0.6 ± 0.01
$t_{50}$ [h]	1.5 ± 0.1	1.5 ± 0.5	4.8 ± 0.3	3.6 ± 0.3	1.5 ± 0.2	9.2 ± 2.3	2.4 ± 0.1	4.4 ± 0.6
% pellet	85	75	80	90	85	50	85	85
	$\Delta$ P1, seed	P1-SG- $\alpha$ Syn, seed						
Lag time [h]	-	-						
Elongation rate*	-	-						
$t_{50}$ [h]	-	-						
% pellet	20	20						

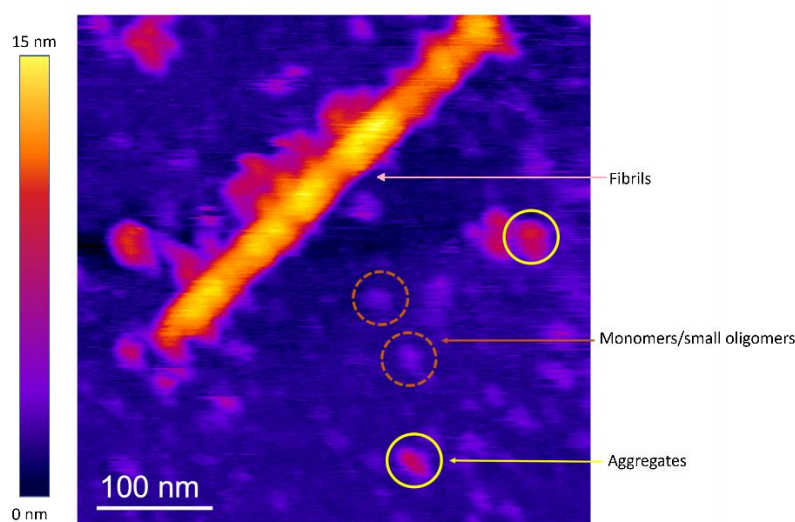
#### 4.4.2 The variants Y39A and S42A form toxic oligomers

Interestingly, although Y39A and S42A do not show a positive ThT signal when incubated at neutral pH for 110 h, there are some fibrils identified by negative stain TEM for S42A and 30% and 15 % of the material is found in the pellet for Y39A and S42A, respectively (Figure 4.14 and Table 4.5). The ThT assay end-products of  $\alpha$ Syn WT and the variants Y39A and S42A (formed at neutral pH) were therefore closer investigated by atomic force microscopy (AFM) after 110 h of incubation (Figure 4.16) (see Section

2.13 for experimental details, experiments performed by George Heath, University of Leeds). Whilst for  $\alpha$ Syn WT only fibrils with a height of  $13.6 \pm 3.4$  nm could be detected, Y39A and S42A formed more heterogeneous samples with fibrils, oligomers, and monomers being present (**Figure 4.16** and **Table 4.6**). An example of typical morphologies for monomers, oligomers, and fibrils imaged by AFM is illustrated in **Figure 4.17**. More precisely, Y39A showed mainly spherical particles with broad height distributions centred around  $1.6 \pm 1.0$  nm (monomer) and  $4.3 \pm 3.3$  nm (oligomers) and a small fraction (1.1%) of short fibrils with a height of  $13.5 \pm 3.8$  nm. The S42A variant also yielded spherical particles with heights centred around  $2.5 \pm 0.8$  nm and  $3.9 \pm 2.1$  nm, and a minor population (~15 %) of fibrils comparable in height to those formed by  $\alpha$ Syn WT (heights of  $14.6 \pm 2.0$  nm and lengths  $153 \pm 74$  nm) (**Figure 4.16** and **Table 4.6**). These findings explain the presence of some fibrillar structures in TEM and pellet material for Y39A and S42A.



**Figure 4.16: AFM study of  $\alpha$ Syn WT, Y39A, and S42A.** (A) AFM images of  $\alpha$ Syn WT (top), Y39A (middle) and S42A (bottom) at the endpoint (110 h) of the fibril growth experiment. The inset shows an expanded scale (scale bar = 50 nm). (B) Height and (C) length/height distributions of the AFM samples (WT= 232; Y39A= 2355; S42A= 898 counts). Orange datapoints represent monomers/oligomers, blue datapoints represent fibrillar species. AFM imaging and analysis was performed by George Heath, University of Leeds.

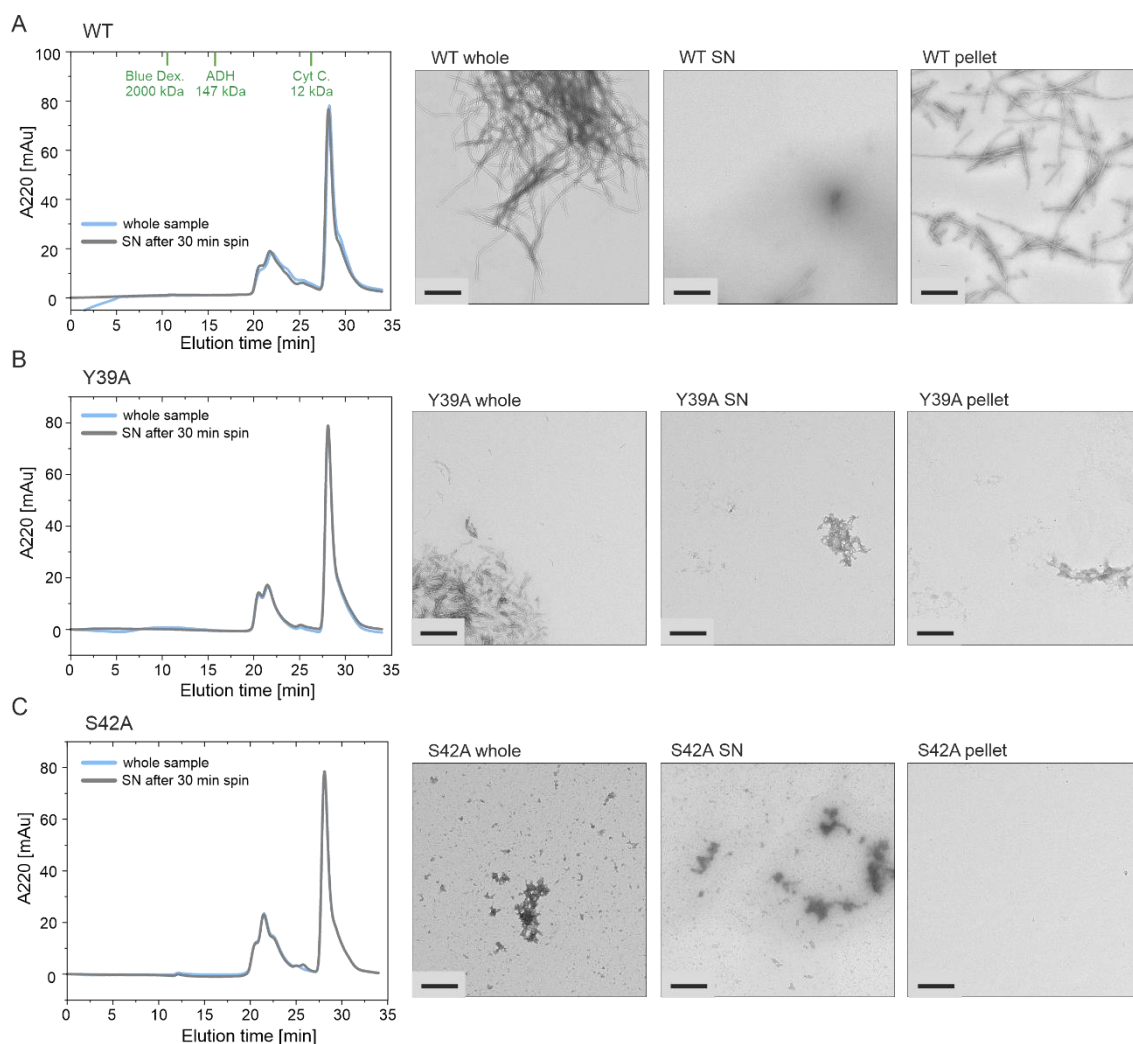


**Figure 4.17: Exemplary AFM image of S42A to highlight fibril, oligomer and monomer/small oligomer species.** Colour code indicates measured height from background state (dark blue) to 15 nm (yellow). Monomers/small oligomers are indicated by dotted orange circles, oligomers by yellow circles. This image was taken by George Heath, University of Leeds.

**Table 4.6: Physical properties of aggregates imaged by AFM.** Fibril lengths are averages with errors showing the standard deviation of the mean ( $n$  = number of fibrils measured). Height and height errors show the peak position and standard deviation of Gaussian fitting to height histograms. % values give number proportion of each feature. – indicates no particles of that type were observed. AFM imaging and analysis was performed by George Heath, University of Leeds.

Variant	Fibril Length [nm]	Fibril Height [nm]	Oligomer Height [nm]	Monomer height [nm]	Number measured (n)
WT	$106 \pm 73$ ( $n = 232$ )	$13.6 \pm 3.4$ (100%)	- (0%)	- (0%)	232
Y39A	$170 \pm 115$ ( $n = 27$ )	$13.5 \pm 3.8$ (1.1%)	$4.3 \pm 3.3$ (61%)	$1.6 \pm 1.0$ (38%)	2,355
S42A	$153 \pm 74$ ( $n = 135$ )	$14.6 \pm 2.0$ (15%)	$3.9 \pm 2.1$ (65%)	$2.5 \pm 0.8$ (20%)	898
L38M	- (0%)	- (0%)	$2.6 \pm 1.7$ (45%)	$1.4 \pm 0.8$ (55%)	2,502

To gain better insight into what species are in the supernatant or pellet of the pelleting assay (see Section 2.8.4), the whole ThT end-point sample, supernatant, and pellet (resuspended in H<sub>2</sub>O) were imaged by negative stain TEM and whole sample versus supernatant were characterised by analytical SEC for  $\alpha$ Syn WT, Y39A, and S42A (**Figure 4.18**) (see Section 2.8.5 for experimental details). Both techniques show the presence of higher order species in the supernatant in addition to monomers (which are too small to detect by TEM) which are apparently not spun down under the used conditions (30 min, 13,000 spin).

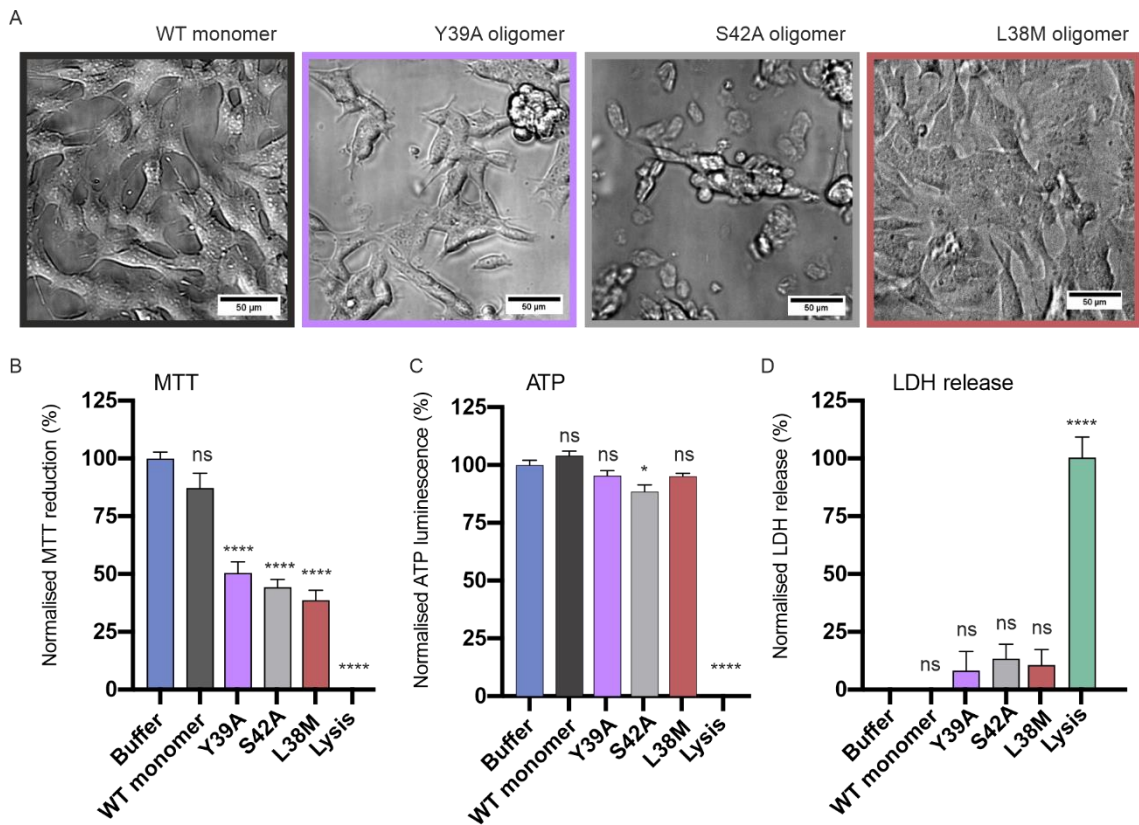


**Figure 4.18: Analysis of ThT end-point samples and efficiency of pelleting assay for  $\alpha$ Syn WT, Y39A, and S42A.** SEC trace and TEM imaging of  $\alpha$ Syn WT (A), Y39A (B), and S42A (C) for the whole sample (including fibrils, oligomers and monomers) versus the supernatant (SN) of the spun sample (30 min at 13,000 rpm). For the SEC analysis fibrils got stuck on the gate column (see Section 2.8.5). However, oligomeric species are present after spinning indicating that there is higher order material present that does not end up in the pellet under the used conditions. Column was equilibrated as indicated in (A). Experiment performed with help from Leon Willis, University of Leeds. Blue trace shows the whole sample, black trace shows the SN after spinning. For the TEM imaging, the pellet was resuspended in 90  $\mu$ L H<sub>2</sub>O before imaging. Scale bar 500 nm.

In a next step, the cytotoxicity of the oligomers formed by Y39A and S42A was studied by incubating SH-SY5Y neuroblastoma cells with the mentioned  $\alpha$ Syn samples (Figure 4.19) (see Section 2.15.1). These experiments were performed by Chalmers Chau and Madeline Brown, University of Leeds. Phase contrast microscopy imaging showed that incubating the cells for only 6 h with 10  $\mu$ M  $\alpha$ Syn sample (final concentration), taken from the ThT assay end-point, resulted in cell morphology changes, clumping and cell death for Y39A and S42A, whilst incubation with  $\alpha$ Syn WT monomer did not have any effect on the cell morphology (Figure 4.19 A). Cell viability was analysed using a MTT<sup>449</sup>, ATP<sup>450</sup> and LDH<sup>451,452</sup> assay (Figure 4.19 B-D). For this, SH-SY5Y cells were incubated for 48 h with 10  $\mu$ M  $\alpha$ Syn ( $\alpha$ Syn WT monomer and ThT end point sample for Y39A and

S42A) before performing the assays. Especially the MTT assay, detecting metabolic activity, showed a significant decrease in MTT reduction for the Y39A and S42A variants compared with incubation with monomer (**Figure 4.19 B**). ATP and LDH assays (providing information about metabolic activity and cell disruption, respectively) were not affected by the addition of  $\alpha$ Syn samples with the exception of S42A in the ATP assay that shows a significantly lower metabolic activity (**Figure 4.19 C,D**). At this point it is not understood why endo-lysosomal cycling (measured by MTT) is reduced when incubated with the variants, but ATP and LDH assays are unaffected in most cases despite obvious morphology changes. The lack of correlation between MTT, LDH, ATP and cell morphology is an atypical observation compared to literature data analysing the effect of  $\alpha$ Syn on SH-SY5Y cells<sup>453</sup>. More work would be required to gain better insight about what is happening with the cells in this study, such as incubation with  $\alpha$ Syn WT oligomers or fibrils, a time- and concentration dependent analysis and investigation of cellular uptake of the samples as control experiments. Taken together, the work presented here highlights that although Y39A and S42A are significantly less aggregation-prone compared to  $\alpha$ Syn WT *in vitro*, oligomers are generated under the used conditions that have some effect on SH-SY5Y cells which needs to be further evaluated.



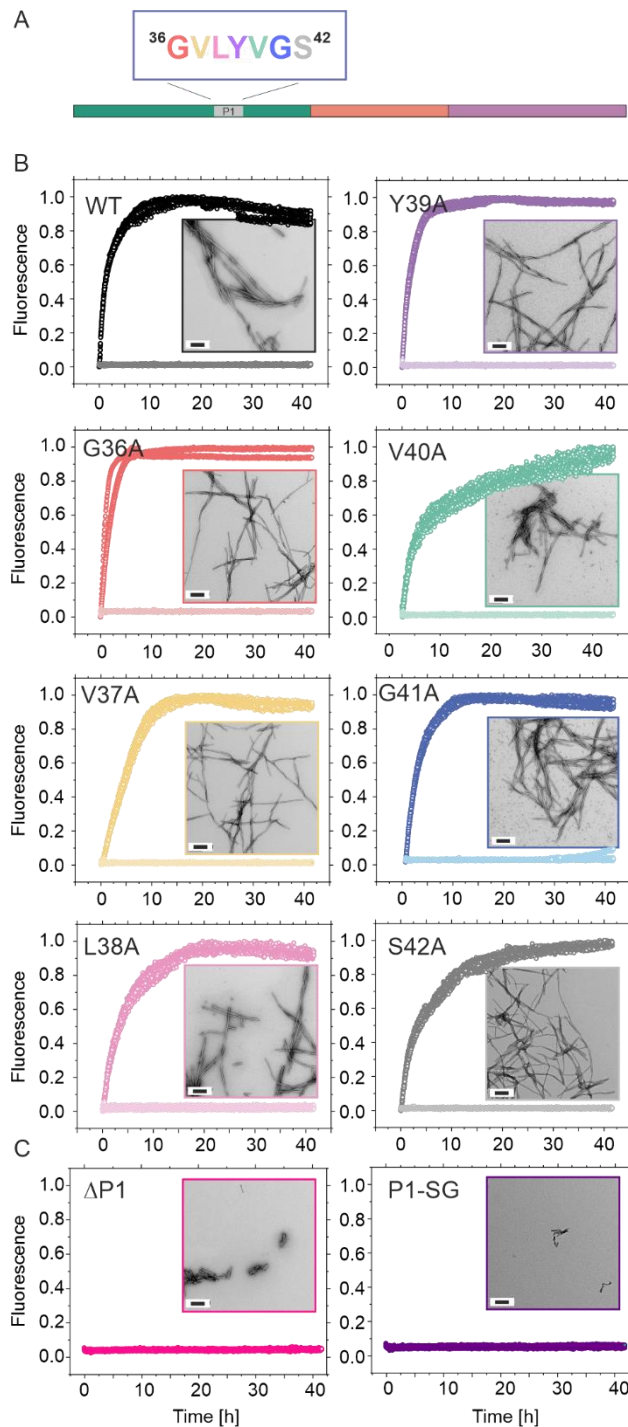


**Figure 4.19: Cell viability assays of Y39A and S42A oligomers.** (A) Phase contrast microscopy of adding 10  $\mu$ M ThT endpoint samples Y39A (purple), S42A (grey) and L38M (red) (Figure 4.14) and  $\alpha$ Syn WT (black) to SH-SY5Y cells after 6 h incubation. The scale bar is 50 nm. Imaging was done by Chalmers Chau, University of Leeds. MTT (B), ATP (C) and LDH release (D) assay after adding 10  $\mu$ M final concentration of  $\alpha$ Syn samples to SH-SY5Y cells and incubating for 48 h. Buffer (blue) and a lysis control ( $\text{NaN}_3$ ) (green) were used as controls. Stars indicate significance between the buffer sample and all other  $\alpha$ Syn samples. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ . ANOVA was used in all cases. Colour code for synuclein variants as described in (A). Assays were performed by Madeline Brown, University of Leeds. Note that variant L38M will be discussed in Section 4.5.

#### 4.4.3 Capacity to seed $\alpha$ Syn WT fibrils for Ala-scan variants

To gain more insight into the mechanism of inhibition of fibrillation of Y39A and S42A, the seeding capacity of the alanine scan variants (monomeric G36A, V37A, L38A, Y39A, V40A, G41A or S42A; **Figure 4.20 A**) with pre-formed  $\alpha$ Syn WT seeds were tested. Adding 10 % (mol/mol) seeds (see Section 2.8.1 for seed preparation protocol) overcomes the nucleation phase and only detects elongation processes if the monomer is seed-elongation-competent under quiescent conditions. Aggregation kinetics were measured by a ThT fluorescence assay and end-point samples were imaged by TEM (**Figure 4.20 B**). The data demonstrate that all tested alanine variants were able to be cross-seeded by  $\alpha$ Syn WT pre-formed fibrils with slight variations in kinetics (**Table 4.5**). This is especially interesting for the Y39A and S42A variants which shown significantly decreased aggregation propensities *de novo* (without seeds) but were able to cross-seed

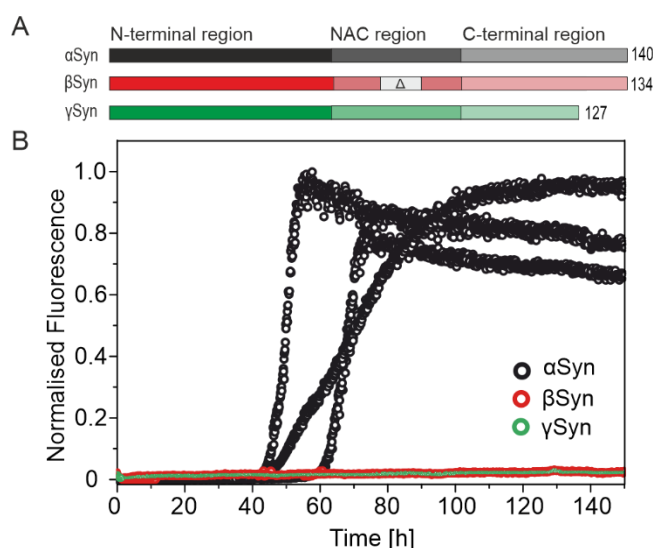
$\alpha$ Syn WT seeds. Therefore amyloid formation of these two variants is most likely blocked/slowed down at an early aggregation state (e.g. nucleation/oligomerisation). So, within the experimental timescale used here (110 h), oligomeric species are accumulated and less fibrils formed as observed by AFM (**Figure 4.14 C-E**). Experiments with no seeds present under quiescent conditions did not show any fibril formation, as expected (**Figure 4.20 B, light colours**). Strikingly, cross-seeding  $\Delta$ P1 or P1-SG with  $\alpha$ Syn WT seeds did not result in successful fibril growth (**Figure 4.20 C, Table 4.5**). This indicates that these variants, in contrast to the single substitution ones, are not able to nucleate fibril growth but also cannot elongate  $\alpha$ Syn WT seeds probably because their sequence is incompatible with the  $\alpha$ Syn WT fibril morphology. The P1 region is intimately or peripherally involved in nearly all cores of determined  $\alpha$ Syn WT fibril architectures to date including *in vitro* grown<sup>96</sup> and patient (MSA) extracted amyloid<sup>117</sup> (see Section 4.10 for more detail). This further points out the P1 sequence to be a key region for nucleation of fibril growth and for stabilising the cross- $\beta$  structure of  $\alpha$ Syn amyloids fibrils.



**Figure 4.20: Cross-seeding  $\alpha$ Syn variants with pre-formed fibrils of  $\alpha$ Syn WT.** (A) Schematic of  $\alpha$ Syn showing the sequence of P1. (B,C) The ability of the different  $\alpha$ Syn variants to elongate seeds formed from  $\alpha$ Syn WT at pH 7.5 was assessed using ThT fluorescence. For each sample 10% (mol/mol) preformed fibril seeds of WT  $\alpha$ Syn formed at pH 7.5 were added to monomers (100  $\mu$ M) of each variant and fibril growth was monitored at pH 7.5; 200 mM NaCl, 37°C, quiescent (no shaking). Fibril growth was monitored in the presence (dark colours) or absence (light colours) of fibril seeds. TEM images after each seeding experiment (42 h) are shown in inset (scale bar 200 nm). Note that the data points with/without seeding in (C) overlay.

## 4.5 Residue 38 in P1 is a key regulator of $\alpha$ Syn aggregation

The above data (Section 4.4) demonstrate that individual residues in P1 can regulate the aggregation propensity of  $\alpha$ Syn, turning off fibrillation on the timescale analysed here (110 h) dependent on the identity of the side chains at position 39 and 42. It is commonly known, that  $\alpha$ Syn is the most aggregation-prone paralogue of the synuclein family.  $\beta$ Syn and  $\gamma$ Syn are significantly less aggregation-prone<sup>365</sup>, including under the buffer conditions used in this study (20 mM Tris-HCl, pH 7.5, 200 mM NaCl) (**Figure 4.21**, **Table 4.7**). The precise reason for the limited aggregation propensity of the paralogues is not fully understood, however, as  $\beta$ - and  $\gamma$ Syn present a 64 % and 54 % sequence similarity, respectively to  $\alpha$ Syn, differences in their aggregation might be explained by residue changes in the primary sequence (**Figure 4.1**). For  $\beta$ Syn, 11 residues in the central NAC region are missing, which is known to be crucial for fibril formation<sup>313</sup>, possibly explaining the lower aggregation propensity.  $\gamma$ Syn on the other hand has a complete NAC region (17 out of 35 residues substituted compared to  $\alpha$ Syn) but lacks 13 residues in the C-terminal region which is considered to protect the protein from forming fibrils in  $\alpha$ Syn<sup>190</sup>, leaving open an explanation for the reduced aggregation propensity in  $\gamma$ Syn.

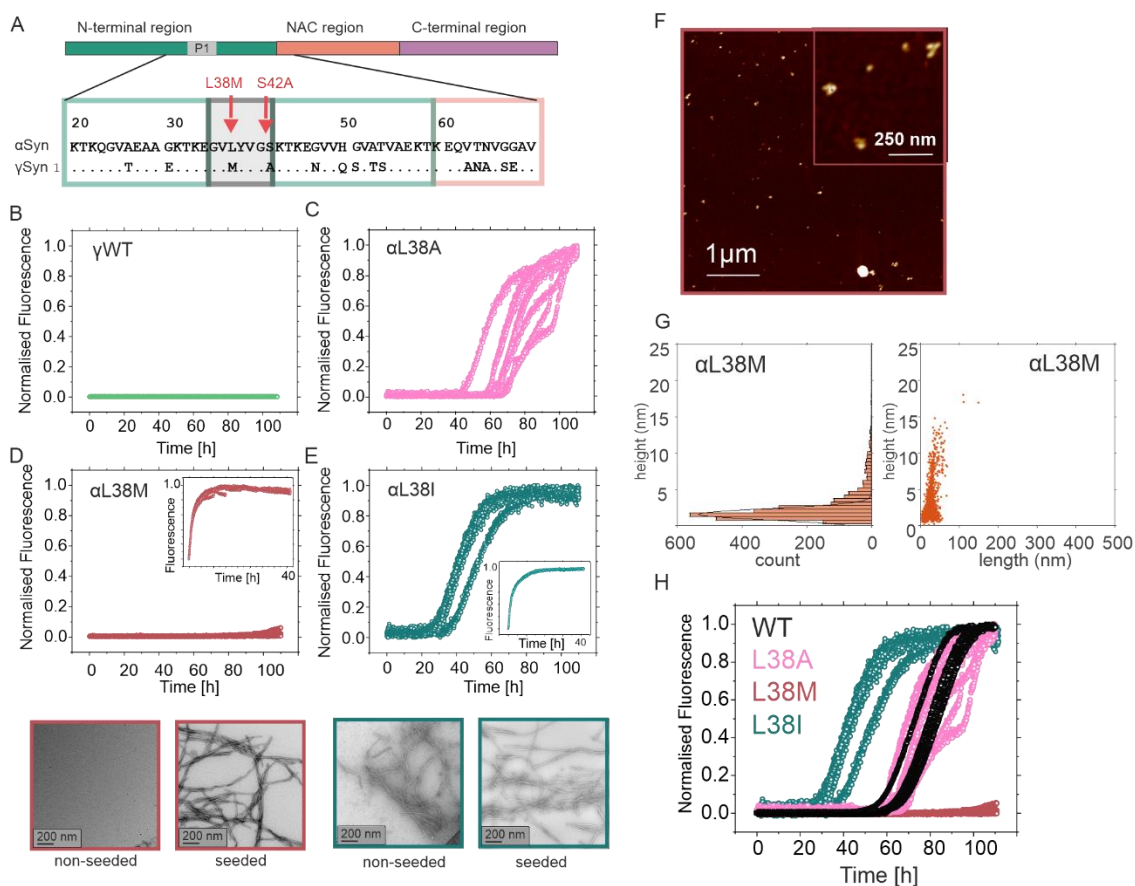


**Figure 4.21: Aggregation assay of  $\alpha$ Syn WT,  $\beta$ Syn WT and  $\gamma$ Syn WT demonstrating the low aggregation propensities of  $\beta$ Syn and  $\gamma$ Syn.** (A) Schematic of the three synuclein family members.  $\alpha$ Syn,  $\beta$ Syn and  $\gamma$ Syn are shown in black, red, green, respectively and N-terminal region, NAC and C-terminal region are coloured with increasing transparency towards the C-terminus. Grey box in  $\beta$ Syn labelled  $\Delta$  represents the 11 missing residues in the  $\beta$ Syn sequence in NAC. (B) ThT fluorescence aggregation assay performed at 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 37 °C, 600 rpm using 100  $\mu$ M monomer. This experiment was performed by Ciaran Doherty, University of Leeds.

The aggregation propensities of  $\alpha$ -,  $\beta$ -, and  $\gamma$ Syn were also analysed using *in silico* techniques (Zygggregator (amyloid propensity<sup>393</sup>), CamSol (local solubility<sup>114</sup>) and

ZipperDB ( $\beta$ -zipper propensity<sup>394</sup>) (**Figure 4.2**). All three proteins present similar overall patterns across their sequence with an aggregation-prone/insoluble N7 motif, as well as P1 and P2 region, a high aggregation propensity for the NAC region (although significantly shorter for  $\beta$ Syn) and a non-aggregation-prone/soluble C-terminal region (shortened for  $\gamma$ Syn). Therefore, there is no obvious sequence-based explanation (**Figure 4.1**) for the altered aggregation kinetics in  $\gamma$ Syn.

A closer look to the sequence of the P1 region for all three variants (**Figure 4.1**), shows that two residues in this seven residue motif are mutated between  $\alpha$ Syn and  $\gamma$ Syn (while  $\beta$ Syn is 100 % identical to  $\alpha$ Syn). Therefore, the focus was further put on understanding the self-assembly of  $\gamma$ Syn and what role the substitutions within P1 play. Leu38 is switched to a Met (L38M) and Ser42 is mutated to an Ala (S42A) in  $\gamma$ Syn (**Figure 4.22 A**). Interestingly, the S42A mutation was shown to significantly reduce aggregation of  $\alpha$ Syn in the alanine scan presented above (**Figure 4.14 B**) and could therefore also contribute to the inability of  $\gamma$ Syn to form fibrils (**Figure 4.22 B**). Based on this finding, a L38M variant in  $\alpha$ Syn was created and its aggregation also investigated. Strikingly, whilst the  $\alpha$ Syn L38A substitution does not change the aggregation kinetics compared with  $\alpha$ Syn WT (lag times of  $56.7 \pm 6.7$  h (L38A) compared to  $\alpha$ Syn WT with  $57.1 \pm 4.7$  h) (**Figure 4.14 B, Figure 4.22 C, Table 4.7**),  $\alpha$ Syn L38M inhibited amyloid formation for at least 110 h (**Figure 4.22 D**). This could be confirmed by TEM (**Figure 4.22 D, bottom**), and AFM analysis showed the formation of oligomeric species with an average height of  $3.1 \pm 2.5$  nm at the endpoint of aggregation (**Figure 4.22 F**). Based on the striking finding that the identity of residues at position 38 makes such a dramatic difference for aggregation, the effect of replacing Leu38 with another aliphatic residue Ile (L38I) was investigated (**Figure 4.22 E**). Strikingly, this substitution results in even faster aggregation compared with  $\alpha$ Syn WT (lag time of  $57.1 \pm 4.7$  h and  $28.1 \pm 3.6$  h for  $\alpha$ Syn WT and L38I, respectively) (**Table 4.7**). Swapping residues at position 38 in  $\alpha$ Syn to different hydrophobic residues can therefore drastically effect the aggregation kinetics speeding it up or slowing it down (**Figure 4.22 H**). Cross-seeding experiments with L38M or L38I monomer and  $\alpha$ Syn WT pre-formed fibrils (see Section 2.8.2.2) resulted in successful elongation, indicating that whilst L38M disrupts amyloid formation, L38I promotes this process and has no/little effect on elongation (**Figure 4.22 D,E insets**). Taken together, these results highlight the specificity of the interactions of residues within the P1 region at the early stages of *de novo* aggregation.



**Figure 4.22: αSyn aggregation rate is dependent on the identity of residue 38. (A)**

Sequence alignment showing that the P1 regions of αSyn and γSyn differ at two positions, residues 38 and 42. Aggregation kinetics of **(B)** αSyn, **(C)** αSyn L38A, **(D)** αSyn L38M and **(E)** αSyn L38I. The inset shows seeding of L38M and αL38I with 10 % (mol/mol) preformed WT αSyn seeds. TEM images of αSyn L38M and αL38I incubated in the absence (non-seeded) or presence (seeded) of 10% (mol/mol) preformed seeds of WT αSyn are shown in D,E, bottom.

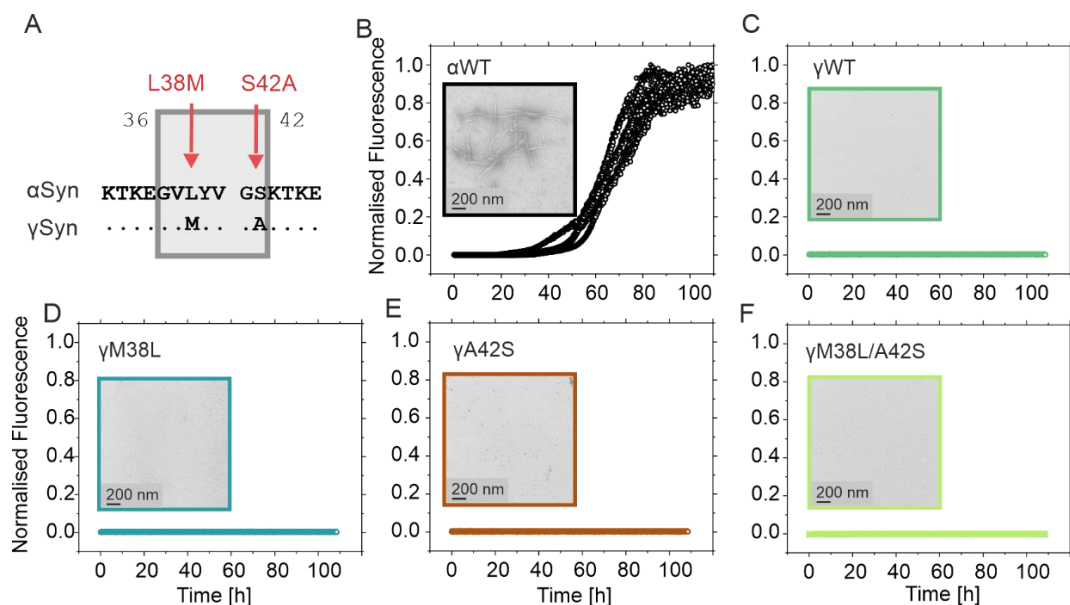
**(F,G)** AFM image and height/length distribution analysis ( $n=2,167$ ) of the products of aggregation of L38M (without seeds). **(H)** Aggregation kinetics of αSyn WT, αL38A, αL38M and αL38I overlaid for direct comparison. All reactions were carried out at pH 7.5, 200 mM NaCl, 37°C, shaking (600 rpm) (*de novo* growth) or quiescent (seeded growth) each using 100 μM synuclein.

**Table 4.7: Aggregation kinetics of different synuclein family members (α-, β-, γSyn) and αSyn variants mutated at position 38.** Lag times, elongation rates,  $t_{50}$  and % pellet for αSyn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 110 h incubation is indicated by “-”. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in 10\*[RFU/h].

	αWT	βWT	γWT	L38A	L38M	L38M, seed	L38I	L38I, seed
Lag time [h]	57.1 ± 4.7	-	-	56.7 ± 6.7	-	0	28.1 ± 3.6	0
Elongation rate*	0.35 ± 0.1	-	-	0.30 ± 0.07	-	2.3 ± 0.4	0.2 ± 0.04	1.5 ± 0.1
$t_{50}$ [h]	63.2 ± 4.9	-	-	74.6 ± 10.0	-	1.4 ± 0.1	44.3 ± 5.8	1.9 ± 0.08
% pellet	65	0	0	75	50	90	65	75

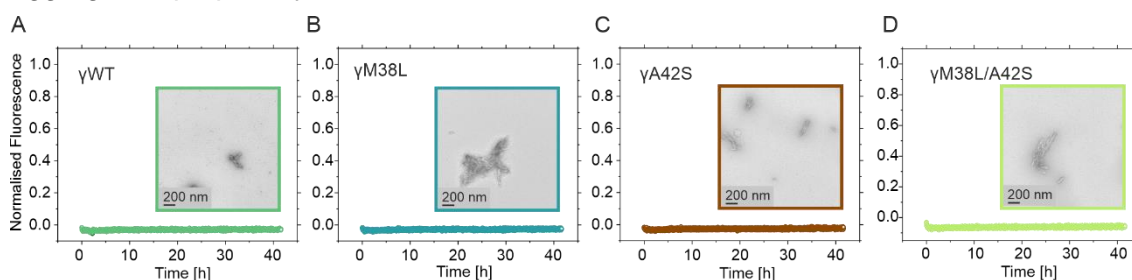
## 4.6 The role of the P1 region in $\gamma$ Syn aggregation

The study of  $\alpha$ Syn aggregation exposed that fibril formation can be switched on or off by changing single residues in the P1 region, despite the aggregation-prone and core-forming NAC region being unchanged. Both,  $\alpha$ Syn and  $\gamma$ Syn NAC region are predicted to be aggregation-prone with an only slightly more soluble region in  $\gamma$ Syn for residues 80-86 (**Figure 4.2 A,C**, more detail in Chapter 5). Particularly the central NAC core (spanning residues 65 to 79<sup>311,441,454</sup>) is calculated to be highly aggregation-prone for both paralogues (**Figure 4.2 A,C**). Thus, the distinct aggregation kinetics of  $\alpha$ - and  $\gamma$ Syn might be driven by sequence differences in the P1 region that was shown to be a key controller in  $\alpha$ Syn aggregation. To test whether substitutions in P1 alter the aggregation kinetics of  $\gamma$ Syn,  $\gamma$ Syn variants M38L, A42S and the double mutant M38L/A42S were created in each case swapping the residue in  $\gamma$ Syn to that in  $\alpha$ Syn (**Figure 4.23 A**). As shown above (Section 4.5, **Figure 4.21**),  $\alpha$ Syn WT aggregates into fibrils under the tested conditions (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 37 °C, 600 rpm), whereas  $\gamma$ Syn WT does not (**Figure 4.23 B,C**)<sup>455</sup>.  $\gamma$ M38L,  $\gamma$ A42S and  $\gamma$ M38L/A42S did not form fibrils *de novo* within the experimental time (110 h) and conditions (**Figure 4.23 D-F**, **Table 4.8**). Also, in the presence of 10 % (mol/mol) pre-formed  $\alpha$ Syn WT seeds, no aggregation could be observed for any of the tested  $\gamma$ Syn variants, demonstrating that the  $\gamma$ Syn constructs are not seed-elongation-competent with  $\alpha$ Syn WT fibrils (**Figure 4.24**, **Table 4.8**).



**Figure 4.23: Sequence alterations in P1 do not result in aggregation of  $\gamma$ Syn.** (A) Sequence alignment of  $\alpha$ Syn and  $\gamma$ Syn focussing on the P1 region (boxed). Residues that differ are highlighted with red arrows and labels. Aggregation kinetics for (B)  $\alpha$ Syn WT, (C)  $\gamma$ Syn WT, (D)  $\gamma$ Syn M38L, (E)  $\gamma$ Syn A42S and (F)  $\gamma$ Syn M38L/A42S. In each case the relevant residue in  $\gamma$ Syn is replaced with the equivalent residue in  $\alpha$ Syn, demonstrating that these amino acid substitutions do not switch on aggregation under the tested conditions (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 37°C, 600 rpm shaking). The insets show TEM images of each sample taken at the end of the reaction (110 h).

The results indicate that fibril formation must be regulated by a complex interplay of interactions between residues in the P1 region and the NAC and/or C-terminal region which share only 51 % and 3 % sequence identity, respectively between  $\alpha$ - and  $\gamma$ Syn (**Figure 4.1**). This is consistent with previous data shown in Section 3.7 (**Figure 3.16**) where intra- and inter-molecular long-range interactions were measured between P1 and other parts of the  $\alpha$ Syn sequence driving aggregation. Also, previous NMR PRE studies in the literature demonstrated that there are fewer interactions between the N- and C-termini of  $\gamma$ Syn compared with  $\alpha$ Syn at neutral pH possibly explaining the lower aggregation propensity<sup>333</sup>.



**Figure 4.24: Cross-seeding  $\gamma$ Syn constructs with 10% (mol/mol) preformed  $\alpha$ Syn WT seeds does not result in fibril formation.** Aggregation kinetics of (A)  $\gamma$ Syn WT, (B)  $\gamma$ Syn M38L, (C)  $\gamma$ Syn A42S and (D)  $\gamma$ Syn M38L/A42S seeded with 10% (mol/mol) seeds of preformed  $\alpha$ Syn WT fibrils formed at pH 7.5, 200 mM NaCl. Each experiment was performed using 100  $\mu$ M monomeric  $\gamma$ Syn at pH 7.5 (200 mM NaCl, 37°C) quiescent. The inserts show negative stain TEM images taken at the endpoint of the experiment (42 h).

**Table 4.8: Aggregation kinetics for  $\gamma$ Syn constructs.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\alpha$ Syn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 110 h incubation for *de novo* growth or 42 h for seeded reactions is indicated by “-“. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in 10\*[RFU/h].

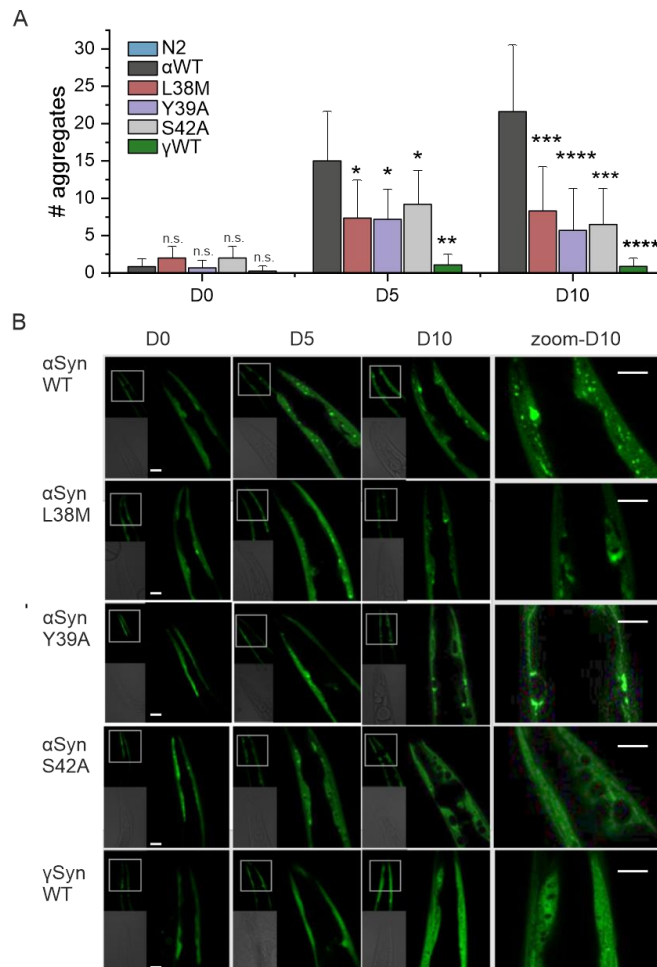
	$\gamma$ M38L	$\gamma$ A42S	$\gamma$ M38L/A42S	$\gamma$ WT, seed	$\gamma$ M38L, seed	$\gamma$ A42S, seed	$\gamma$ M38L/A42S, seed
Lag time [h]	-	-	-	-	-	-	-
Elongation rate*	-	-	-	-	-	-	-
$t_{50}$	-	-	-	-	-	-	-
% pellet	0	0	0	10	5	5	15

## 4.7 Single residues in $\alpha$ Syn fine-tune aggregation *in vivo*

Expressing  $\alpha$ Syn WT in the body wall muscle cells of *C. elegans*, is commonly used as an *in vivo* model for studying pathogenesis and disease, allowing to investigate the effect of amyloid formation over lifespan in a living organism<sup>83</sup>. Previous experiments in Section 3.5 have shown that the expression of  $\Delta$ P1::YFP protects the worms from forming aggregates and developing a phenotypic effect (reduction in body bends) compared with  $\alpha$ Syn WT::YFP. Therefore, the impact of the single substitutions in  $\alpha$ Syn that inhibit aggregation *in vitro* (L38M, Y39A and S42A), were tested in this worm model (**Figure 4.25**). Puncta formation and changes in motility of the nematodes were measured over



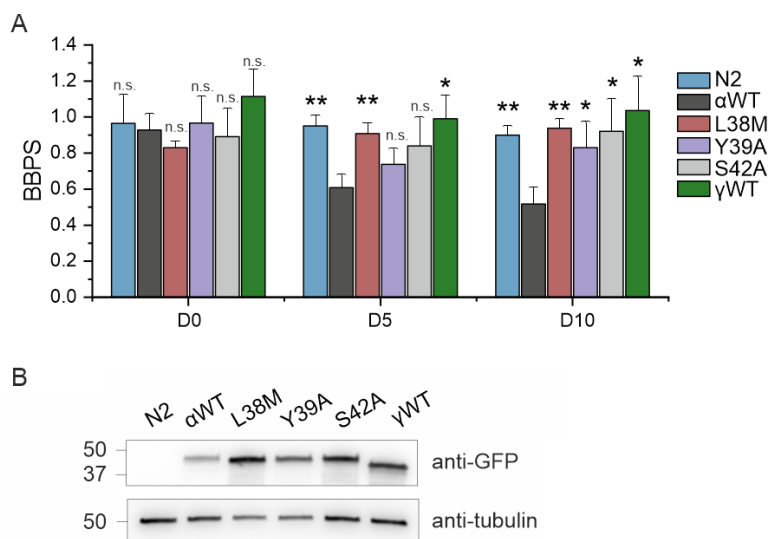
their adult lifespans when expressing these protein variants with a C-terminally fused YFP construct in the body wall muscle cells of *C. elegans*. For all worms, very low numbers of inclusions were observed at the larvae stage (day zero (D0)). When aging, the number of inclusions significantly increases for  $\alpha$ Syn WT::YFP expressing animals measured at D5 and D10 as seen before (**Figure 4.25 A,B, Figure 3.11**)<sup>83</sup>. Remarkably, the single substitution variants showed a two-fold reduction of inclusion formation, indicating lower aggregate formation.



**Figure 4.25: Effect of single point mutations in  $\alpha$ Syn and  $\gamma$ Syn in the body wall muscle of *C. elegans* for puncta formation. (A)** Number of inclusions (larger than  $\sim 2 \mu\text{m}^2$  per animal). Data shown are the mean and s.e.m. for 10 worms ( $n = 10$ ) that were assessed for each time point. Stars indicate significance between the number of aggregates of  $\alpha$ Syn WT expressing worms with all other constructs. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ . ANOVA was used in all cases. Note that for N2 worms no data were collected as they do not express YFP. **(B)** Confocal microscopy images showing the head region of transgenic *C. elegans* expressing WT  $\alpha$ Syn,  $\alpha$ Syn L38A,  $\alpha$ Syn Y39A,  $\alpha$ Syn S42A or WT  $\gamma$ Syn (each fused to YFP at their C-terminus) in the body wall muscle during ageing (Day 0, Day 5 and Day 10 of adulthood). Small images on left hand site show light microscopy and fluorescence image highlighting the zoomed in area shown on the right hand site. D10 is further zoomed in for easier visualisation of formed puncta. Scale bar, 10  $\mu\text{m}$ .

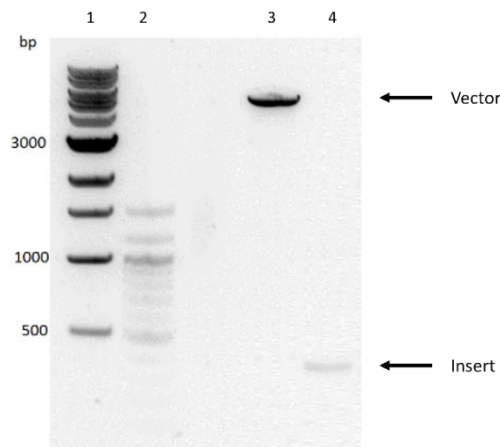
As the proteins are expressed in the body wall muscle cells, their phenotypic effect can be evaluated by a trashing assay in which the body bends per second (BBPS) are counted (Section 2.15.3.3). In agreement with the timing and extent of inclusion

formation, all *C. elegans* strains presented similar motilities at D0 (**Figure 4.26 A**). As expected, worms expressing  $\alpha$ Syn WT showed a significant reduction in BBPS. By contrast, the motility of the mutant-expressing nematodes was unaffected up to day 10 of adulthood, further highlighting the protective effect of the M38L, Y39A and S42A substitutions *in vivo*. It should be noted, that expression levels for all protein variants were similar as determined via western blot analysis by Kathrine Dewison, University of Leeds (**Figure 4.26 B**).



**Figure 4.26: Effect of single point mutations in  $\alpha$ Syn and  $\gamma$ Syn in the body wall muscle of *C. elegans* for motility.** (A) Number of body bends per second (BBPS) of N2, WT  $\alpha$ Syn::YFP,  $\alpha$ Syn L38M::YFP,  $\alpha$ Syn Y39A::YFP,  $\alpha$ Syn YS42A::YFP and WT  $\gamma$ Syn::YFP animals from Day 0 (L4 stage), Day 5 and Day 10 of adulthood. Data shown are mean and s.e.m. for three independent experiments; in each experiment, >10 worms were assessed for each time point. \*\*P<0.01; \*P<0.05, ANOVA was used comparing WT with synuclein variants or N2 worms. (B) Western Blot analysis of synuclein worm extracts with anti-GFP and anti-tubulin antibodies performed by Kathrine Dewison, University of Leeds.

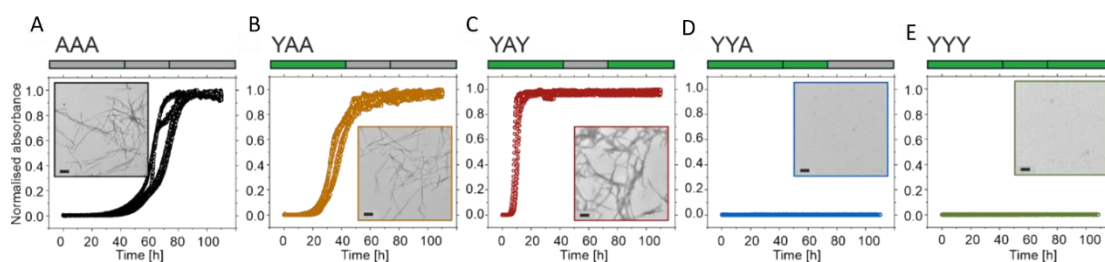
In addition, the consequence of expressing  $\gamma$ Syn WT::YFP in *C. elegans* was studied. For this, the  $\gamma$ Syn WT gene (gene sequence optimised for *C. elegans*, Eurofins) was cloned into the pPD30.38 vector for expression in the worm muscle cells. PCR amplification of the insert and enzyme digestion of the plasmid (Section 2.5.2) resulted in DNA bands accordingly to the sizes of vector and insert (**Figure 4.27**). After ligation of insert and vector (resulting in  $\gamma$ Syn::YFP in pPD30.38 vector) and generation of worm strains with the new  $\gamma$ Syn gene by microinjection (kindly performed by Dovile Milonaityte, University of Leeds) (Section 2.15.3.1) the phenotypic effect on *C. elegans* was analysed. The results showed only few inclusions even at day 10 (**Figure 4.25**) and body bends similar to the control worms (N2) not expressing synuclein (**Figure 4.26 A**). This suggests that  $\gamma$ Syn is not proteotoxic, even in nematodes at advanced age (D10), supporting the *in vitro* findings that the protein is less aggregation-prone compared with  $\alpha$ Syn WT.



**Figure 4.27: Agarose gel of digested pPD30.38 vector and  $\gamma$ Syn WT insert.** Line 1 and 2: 1 kb and 100 bp DNA ladder, respectively, line 3: vector (pPD30.38), line 4:  $\gamma$ Syn WT construct.

## 4.8 Interactions between N- and C-terminal regions control synuclein aggregation

The failure of changing residues 38 and 42 in the P1 region of  $\gamma$ Syn to those in  $\alpha$ Syn to switch on its aggregation could arise from sequence differences in the NAC region, the absence of the complementary interaction sites for residues 38 and 42 in other regions of the  $\gamma$ Syn sequence<sup>333</sup>, or both. To examine these possibilities chimeric proteins were created in which the N-terminal (residues 1-60), NAC (residues 61-95) and C-terminal regions of  $\alpha$ Syn and  $\gamma$ Syn (residues 96 to 140/127) were interchanged, creating a series of chimeric proteins including  $\alpha$ Syn WT (named AAA) (N-terminal-NAC-C-terminal each containing the sequence of  $\alpha$ Syn), WT  $\gamma$ Syn (named YYY) (N-terminal-NAC-C-terminal regions each containing the sequence of  $\gamma$ Syn), and the variants YAA, YAY and YYA (**Figure 4.28**) and the rate of aggregation of each protein was measured at pH 7.5. The results show a clear correspondence of aggregation with the identity of NAC, with all constructs containing  $\alpha$ Syn NAC aggregating rapidly (with aggregation reaching completion within < 20 h) (**Figure 4.28 A-C**), while those containing  $\gamma$ Syn NAC (YYA and YYY) failed to aggregate over the time course of these experiments (**Figure 4.28 D,E**). Interesting, the rate of aggregation of the constructs that contain an  $\alpha$ Syn NAC region is dependent on the nature of the flanking regions, with YAY aggregating more rapidly than YAA, and AAA (WT  $\alpha$ Syn) aggregating the most slowly of these three proteins (lag times of  $48.9 \pm 0.9$ ;  $22.3 \pm 1.9$  and  $8.6 \pm 0.4$  h for AAA, YAA and YAY, respectively) (**Table 4.9**). Previous reports have shown that replacing the N- and C-terminal regions of  $\alpha$ Syn with those from  $\beta$ Syn did not show acceleration of aggregation<sup>456</sup>, presumably because of the high sequence similarity of  $\beta$ Syn to  $\alpha$ Syn, especially in the N-terminal region (90% identity).



**Figure 4.28: Aggregation kinetics of chimeric proteins in which the N-terminal, NAC and C-terminal regions of  $\alpha$ Syn and  $\gamma$ Syn are swapped.** Fibril growth kinetics of (A) AAA ( $\alpha$ Syn WT), (B) YAA, (C) YAY, (D) YYA and (E) YYY ( $\gamma$ Syn WT) (100  $\mu$ M protein, 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 37°C, 600 rpm shaking). Bars above indicate regions' identity being alpha-synuclein (grey) or gamma-synuclein (green). The results demonstrate the importance of the identity of the NAC region and the flanking regions in determining the rate of aggregation into amyloid. The insets show TEM images of samples taken at the end of the reaction for each sample (110 h).

**Table 4.9: Aggregation kinetics of chimeric synuclein variants.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\alpha$ Syn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 110 h incubation is indicated by "-". For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in  $10^4$ [RFU/h].

	AAA	YAA	YAY	YYA	YYY
Lag time [h]	48.9 $\pm$ 0.9	22.3 $\pm$ 1.9	8.6 $\pm$ 0.4	-	-
Elongation rate*	0.2 $\pm$ 0.05	0.3 $\pm$ 0.02	2.0 $\pm$ 0.07	-	-
$t_{1/2}$	67.6 $\pm$ 3.3	36.0 $\pm$ 2.7	9.7 $\pm$ 1.2	-	-
% pellet	65	70	60	0	0

These results add further weight to the conclusion that the interactions that drive  $\alpha$ Syn aggregation are highly specific, with different residues in the P1 region (L38, Y39 and S42) synergising with residues in the NAC and C-terminal regions to fine-tune the rate of aggregation, presumably by favouring or disfavouring conformer(s) that define its ability to aggregate into amyloid. The precise molecular mechanism by which this is accomplished remains to be elucidated. Thus, while NAC is necessary and sufficient for  $\alpha$ Syn aggregation, the ability to prevent aggregation at pH 7.5 by removal or substitution of a single, specific, residue in the seven-residue P1 sequence provides a striking demonstration of the crucial effect of this region in amyloid formation.

## 4.9 Membrane binding of L38M and $\gamma$ Syn WT

As described in Section 3.9 the P1 and P2 regions (residues 36-42 and 45-57) in  $\alpha$ Syn were found to be key motifs to allow membrane remodelling, and changes number and conformation (helix proportion) of residues involved in DMPS liposome binding (shown by deleting or replacing the P1P2 region). Previous studies demonstrated that the binding and aggregation properties of familial PD mutations with DMPS liposomes are affected, showing a drastic inhibition of surface induced aggregation for E46K, H50Q or G51D (all located within P2)<sup>172</sup>. As aggregation can be controlled by substituting only a

single residue in P1 (shown for L38M, Y39A or S42A), the role of the point mutations for DMPS liposome interaction was also investigated. All three  $\alpha$ Syn variants are rather located at the membrane exterior (**Figure 4.29 A**)<sup>361</sup>. As L38 is closest to the lipid bilayer head groups and therefore believed to have the largest effect on membrane binding, the L38M variant was chosen for further investigation (**Figure 4.29 B-E**). First, the interaction with DMPS liposomes was monitored using the formation of the  $\alpha$ -helical structure in the N-terminal part of the protein as an indicator for liposome binding measured by far-UV-CD (**Figure 4.29 B,C**). This experiment indicated binding between  $\alpha$ L38M and DMPS liposomes with a  $K_D = 0.7 \pm 0.4 \mu\text{M}$  and a stoichiometry of  $L = 38.2 \pm 2.2$  (**Table 4.10**). These values are similar to  $\alpha$ Syn WT binding ( $K_D = 0.22 \pm 0.13 \mu\text{M}$  and  $L = 33.4 \pm 1.1$ ), and also the resulting helical conformation is comparable between  $\alpha$ Syn WT and L38M (74 % and 77 %, respectively) (**Table 4.10**) suggesting an analogous binding mode. Monitoring the heterogenous primary nucleation of  $\alpha$ L38M by incubation with liposomes under quiescent conditions showed successful fibril formation in the presence of 8x lipids (mol/mol) as seen before for  $\alpha$ Syn WT<sup>351</sup>. However, in contrast to  $\alpha$ Syn WT, a two-phase growth curve is observed for L38M (**Figure 4.29 D, Figure 3.25**). Strikingly, the end-point sample of  $\alpha$ L38M with 60x lipids (mol/mol) did not show the formation of liposome tubes (**Figure 4.29 E**), therefore the Leu38Met substitution most likely interrupts the physiological function of  $\alpha$ Syn of remodelling vesicles, although still enabling binding to the liposomes, similar to the observations for  $\alpha$ Syn  $\Delta\Delta$  (at least considering liposome remodelling).

**Table 4.10: Binding thermodynamics of synuclein variants to DMPS liposomes.**

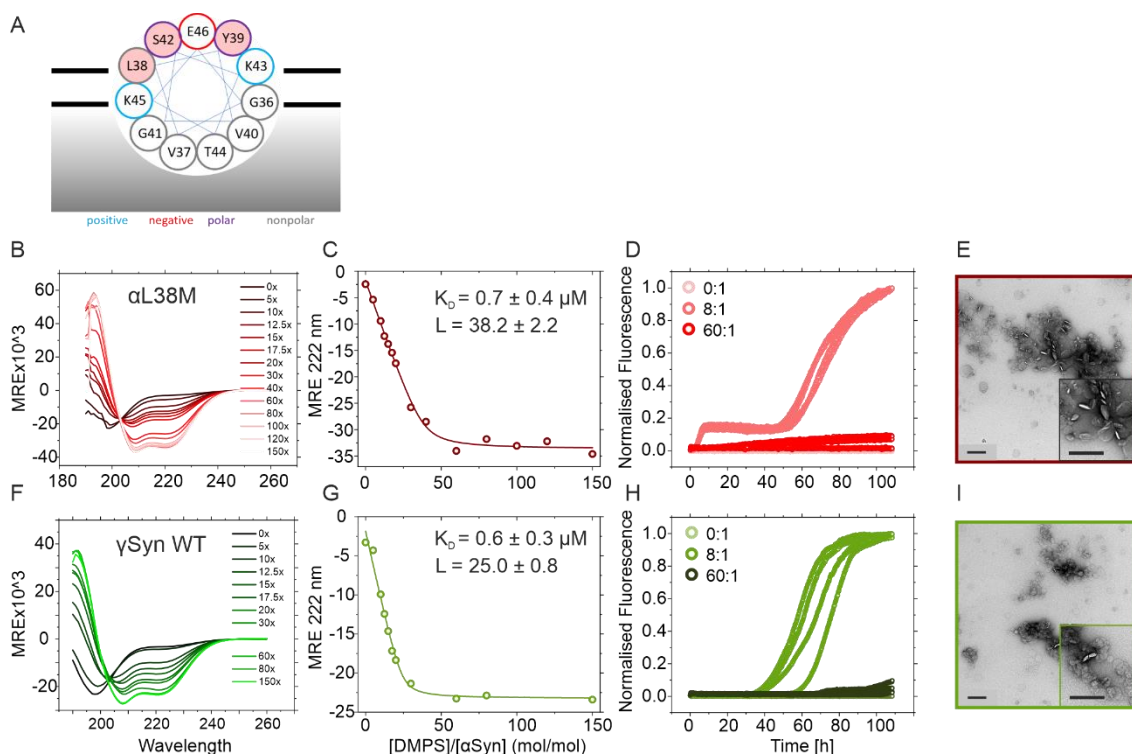
Parameters were calculated from the change in MRE at  $\lambda_{222 \text{ nm}}$  fitted to a single-step binding model<sup>351</sup> (Section 2.10.1). % helix denote the population of helical structure calculated from Dichroweb<sup>392</sup> using experimental CD data. CD data for  $\alpha$ Syn WT,  $\Delta\Delta$ , and P1P2-GS are shown in **Figure 3.23**.

Variant	$K_D$ [ $\mu\text{M}$ ]	L (stoichiometry)	Helix [%]
$\alpha$ Syn WT	$0.22 \pm 0.13$	$33.4 \pm 1.1$	74
$\alpha$ Syn $\Delta\Delta$	$2.01 \pm 0.63$	$49.1 \pm 2.2$	30
$\alpha$ Syn P1P2-GS	$5.26 \pm 3.88$	$55.9 \pm 7.4$	13
$\alpha$ Syn M38L	$0.7 \pm 0.4$	$38.2 \pm 2.2$	77
$\gamma$ Syn WT	$0.6 \pm 0.3$	$25.0 \pm 0.8$	55

It is known that also  $\gamma$ Syn WT interacts with liposomes and forms an  $\alpha$ -helical structure in the N-terminal and NAC region upon binding<sup>379,380</sup>. Thus, the binding characteristics for  $\gamma$ Syn were also analysed (**Figure 4.29 F-I**). The binding affinity measured by far-UV-CD is comparable to  $\alpha$ Syn WT and  $\alpha$ L38M with a  $K_D = 0.6 \pm 0.3$  and stoichiometry of  $L = 25.0 \pm 0.8$ , though the latter shows slightly less lipids being involved in the binding process (**Figure 4.29 F,G and Table 4.10**). A reduced helix formation was detected with only 55 % of the protein forming this secondary structure at saturation (60x excess lipid). Surface induced nucleation was observed in the presence of DMPS liposomes with significantly increased aggregation kinetics compared with *de novo* fibril growth (lag time

of  $43.1 \pm 2.9$  h in the presence of 8x excess lipid), with inhibition of fibril growth at 60x excess lipids (**Figure 4.29 H**), and no remodelling of liposomes could be shown (**Figure 4.29 I**).

Taken together these results indicates that  $\gamma$ Syn binds DMPS liposomes similar to  $\alpha$ Syn WT but adopts less helical structure (similar to the effect observed for  $\Delta\Delta$  and P1P2-GS) and cannot merge liposomes into tube-like structures. The lack of this function of  $\gamma$ Syn might be explained by its localisation away from the synaptic termini to the centrosome<sup>457</sup> where interactions with vesicles might not be required for its physiological function (see Section 1.5.1).



**Figure 4.29: Membrane binding of  $\alpha$ Syn L38M and  $\gamma$ Syn WT to DMPS liposomes. (A)**

Schematic of  $\alpha$ -helix bound to a membrane showing which residues face towards or away from the membrane surface. Residues shown to slow down aggregation are highlighted in red.

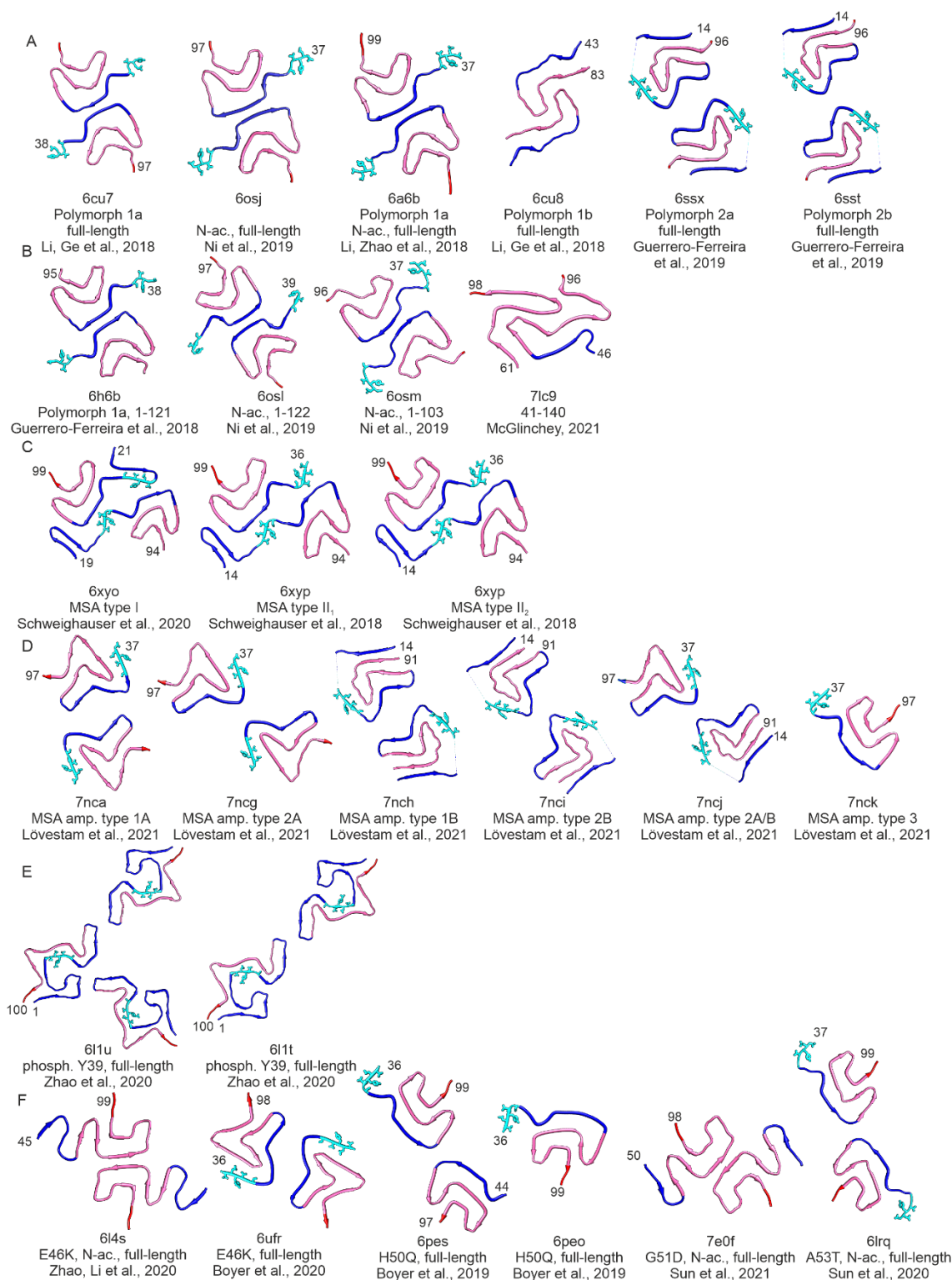
Figure adopted from Ref<sup>361</sup>. L38M (**B**) or  $\gamma$ Syn WT (**F**) incubated with increasing ratios of [DMPS]:[protein] (mol/mol). (**C,G**)  $K_D$  and  $L$  values were calculated from the change in MRE at  $\lambda_{222\text{nm}}$  fitted to a single-step binding model<sup>351</sup>. (**D,H**) Aggregation kinetics of 50  $\mu$ M L38M (**D**) or  $\gamma$ Syn WT (**H**) incubated with 0:1, 8:1 or 60:1 [DMPS]:[protein] (20 mM sodium phosphate, pH 6.5; 30  $^\circ$ C, no shaking). (**E,I**) TEM images at the endpoint of the incubations (110 h) in the presence of 60:1 [DMPS]:[protein] (mol/mol). Scale bar = 200 nm.

## 4.10 Discussion

Having identified the P1 (residues 36-42) and P2 (residues 45-57) regions in  $\alpha$ Syn to be ‘master controllers’ of aggregation and functional activity in Chapter 3, here, the role of individual residues in the P1 region of  $\alpha$ Syn and  $\gamma$ Syn was analysed in detail.

The data presented in this chapter suggest a high specificity of the sequence of P1 in regulating self-assembly into amyloid fibrils of  $\alpha$ Syn. Single point amino acid substitutions identified by an alanine scan, Y39A and S42A, were shown to significantly reduce the aggregation kinetics at neutral pH and also retard puncta formation and toxicity in *C. elegans*. Similar alanine substitutions within the P1 region (G36A, V37A, L38A, V40A, and G41A) did not show an effect on amyloid formation compared with  $\alpha$ Syn WT in *in vitro* studies. Remarkably, whilst L38A was shown to have no effect, L38M inhibits fibril formation *in vitro* and *in vivo* and the mutation L38I speeds up aggregation (although this substitution was not studied in *C. elegans*). This suggests that specific sidechain-sidechain interactions in the early stages of self-assembly depend on the identity of the amino acid at this site. How and why these specific residue changes affect fibril formation remains unclear, requiring a more in-depth analysis of the conformational ensemble of the monomers and more information about the oligomeric species formed. Factors such as changed intramolecular interactions by altering hydrogen bonding with the sidechain hydroxyl of Y39 and/or S42, as well as steric effects by, for example swapping L38 to a longer methionine, or L38 to the  $\beta$ -branched Ile, could affect aggregation by changing transient secondary structure or long range/local contacts within the dynamic IDP. Further, the contribution of toxicity from oligomers versus fibrils should be investigated in more detail, as some familial PD mutations (e.g. A30P) have been shown to slow down fibril formation, but to increase oligomer concentrations, and the data presented here on cellular toxicity in SH-SY5Y cells is inconclusive.

All variants with substitutions in P1 analysed in this study were able to elongate pre-formed fibril seeds formed by  $\alpha$ Syn WT, indicating that these variants are compatible with the  $\alpha$ Syn WT fibril architecture. This would be expected based on the high stability of the amyloid fold<sup>96</sup>, and same turn sequence (residues 44-47) which can form a  $\beta$ -hairpin structure thought to be involved in the initial processes of aggregation within the P1P2 region<sup>458</sup>. Interestingly, whereas *in vitro*  $\alpha$ Syn fibril structures show no protofilament interfaces involving the P1-region (although P1 is part of the rigid core in most structures) (**Figure 4.30 A,B**)<sup>96</sup>, *ex vivo* structures extracted from MSA patients present an interface involving residue V40 and Y39 (**Figure 4.30 C**)<sup>117</sup>. It is known that  $\alpha$ Syn point mutations or post-translational modifications can result in altered fibril architectures (**Figure 4.30 E,F**)<sup>96</sup>, therefore the  $\alpha$ Syn variants that form fibrils either unseeded, or seeded with  $\alpha$ Syn WT pre-formed fibrils, could have different fibril stabilities and preferred structures, changing thermodynamic and kinetic processes towards the amyloid state.



**Figure 4.30: Summary of all cryo-EM high resolution structures of  $\alpha$ Syn solved to date.** Fibril structures are illustrated from the top view. Structures of *in vitro* grown full length  $\alpha$ Syn WT (**A**), *in vitro*  $\alpha$ Syn truncated architectures (**B**), MSA patient extracted structures (**C**), MSA fibril amplified structures (**D**), *in vitro* structure with post-translational modification (phosphorylated Y39) (**E**) and familial PD associated variants grown *in vitro* (**F1**). For all structures, the N-terminal region is coloured in blue, the P1 region is highlighted in cyan (sidechains are shown), the NAC region is pink, and the C-terminal region is red. Note that for some structures the P1 region (as well as other residues in N- and/or C-terminus) do not form the stable core and are therefore not illustrated here. First and last residue solved are indicated. Below each fibril structure, the PDB file, important characteristics and the associated publication are provided.



Single residue substitutions in  $\alpha$ Syn are associated with the development of familial PD<sup>172</sup> (Section 1.4.2.1). Three of the twelve known familial mutations are located N-terminal to P1 (MT5, A30P/G), two in NAC (T72M, G93A), one in the C-terminal region (P117S), and the remaining six substitution sites occur in the P2 region (E46K, H50Q, G51D, A53E/T/V). To date, there are no known natural mutations in the P1 region, although the work presented here has now clearly demonstrated its crucial contribution to fibril formation. This might be explained by rare findings of new familial mutations or by the fact that the P1 region might be required for physiological function. Consistent with the observations reported here, the effect of the familial PD mutations is also highly residue-specific. For instance, while the familial PD mutation A30P decreases the rate of amyloid formation, the aggregation kinetics of A30G are unchanged compared with  $\alpha$ Syn WT<sup>281</sup>. Similar, A53T and A53V have been shown to fibrillate faster than A53E<sup>459</sup>.

It is known that a complex interplay of intra- and inter-molecular interactions between the N-terminal region, NAC and C-terminal region determine the conformational ensemble of  $\alpha$ Syn and affecting the aggregation rates of the protein (see Sections 1.4.3.1 and 3.7)<sup>189,335</sup>. Therefore it is not surprising that mutations in other parts of  $\alpha$ Syn can also affect the assembly rate, such as V70G/E, V74G/E, V76E/N or S87N in NAC that reduce the lag-time of fibril formation<sup>313,365,409</sup> or Y133A in the C-terminal region slowing down the aggregation kinetics<sup>440</sup> as well as D121A showing a reduced fibrillation in the presence of 2.5 mM CaCl<sub>2</sub> compared with  $\alpha$ Syn WT (probably due to charge differences resulting in distinct Ca<sup>2+</sup> binding and monomer conformations)<sup>460</sup>. These substitutions likely modulate the population of monomers with an exposed NAC region towards less aggregation-prone conformations via subtle changes in this dynamically disordered IDP<sup>325,461-463</sup>. This hypothesis is consistent with the results presented in this chapter that the addition of the P1 sequence *in trans* to  $\alpha$ Syn WT and  $\Delta$ P1 speeds up fibril formation by interrupting interactions between the N-terminal and C-terminal region upon peptide binding.

The results presented in this chapter indicate that the early stages of self-assembly involve interactions that crucially depend on the location and identity of individual sidechains at defined locations in the P1 region of this 140-residue IDP. Further investigations will be needed to clarify the origin of this specificity in more detail using techniques such as cross-linking, single molecule FRET, NMR and other biophysical methods combined with MD simulations. This might help to gain an understanding of the atomistic models of the fluctuating ensembles of monomers and early aggregates<sup>325,461</sup>. Further, other approaches including determining the energy landscape with small-molecules added non-covalently or via tethering<sup>464,465</sup>, deep mutational scanning of the full length protein or focussed on P1 with suitable selection screens in different organisms<sup>79,360,361,466</sup>, or a detailed comparison of synuclein variants such as those

generated here with natural paralogues with different aggregation characteristics (e.g.  $\alpha$ Syn,  $\gamma$ Syn, and chimeras)<sup>333,456</sup>. A combination of these experiments could help to tease apart the intra- and inter-molecular interactions required for fibril formation.

Interestingly, residue Y39 was already described in the literature to be vital for  $\alpha$ Syn aggregation as well as other biological progresses in contexts such as chaperone binding<sup>175</sup> or disease (phosphorylation in disease brains)<sup>266</sup>. In contrast, residues L38 and S42 have not been identified to be important hitherto, and the effect of the substitutions to methionine and alanine, respectively on  $\alpha$ Syn function at the synapse, in membrane binding and in chaperone function remain to be explored. The fact that all three identified mutation sites and all but one familial PD mutations are outside the aggregation-prone NAC region highlights the requirement of analysing these flanking regions in more detail to gain a better understanding of the molecular mechanisms of aggregation of the full-length protein.

The importance of residues outside the aggregation hotspot is not only crucial for self-association of  $\alpha$ Syn but also for other amyloid proteins, that often present disease-causing mutations distal to the most aggregation-prone regions (see Section 1.2.5, **Figure 1.10**). These 'master controller' regions of aggregation (and function) could be good targets for the development of reagents to combat amyloid formation by binding small molecules, chaperones, biologics, or other agents. In the case of  $\alpha$ Syn,  $\beta$ -wrapin has been shown to bind to the P1-P2 region of  $\alpha$ Syn, resulting in prevention of aggregation *in vitro*, in *Drosophila* and in neurons<sup>187,342,415</sup>. Further, a SUMO-1 derived peptide binds regions 37-42 (in P1) or 48-52 (in P2) and thereby inhibits  $\alpha$ Syn aggregation *in vitro* and *in vivo*<sup>439</sup>. The formation of the  $\beta$ -hairpin structure was shown to compete against SUMO1 binding which does not induce a hairpin formation and therefore functions through a distinct mechanism compared with  $\beta$ -wrapin (see Section 3.10 for  $\beta$ -hairpin discussion). These two examples offer a proof-of-principle of the potentials of an approach targeting the P1 and/or P2 region(s).

## Chapter 5

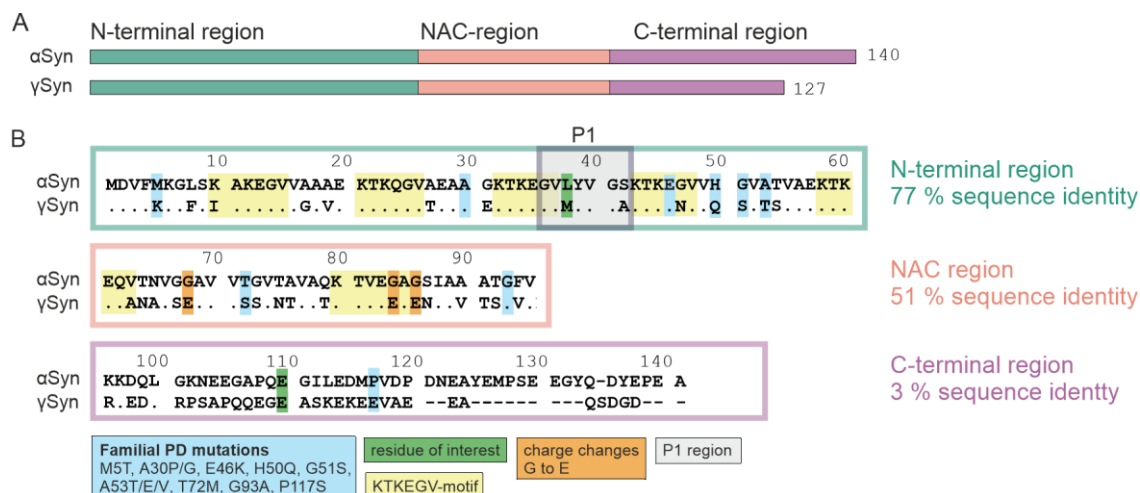
# An ALS-linked sequence variant of $\gamma$ Syn in P1 induces aggregation

### 5.1 Introduction

The results described in Chapter 3 and Chapter 4 analysed the role of the N-terminal region, and especially the P1 region (residues 36-42) in controlling the rate of aggregation of  $\alpha$ Syn in great detail. Whether the P1 region plays a critical role in its paralogue  $\gamma$ Syn however, remained unclear and will be further investigated in this chapter.

$\gamma$ Syn, exclusively expressed in vertebrates<sup>257</sup>, shows a significantly lower aggregation propensity compared with  $\alpha$ Syn<sup>378</sup>. Similar to  $\alpha$ Syn,  $\gamma$ Syn has a N-terminal region, a NAC region and a truncated C-terminal region (13 residues shorter than  $\alpha$ Syn) (**Figure 5.1 A**). The exact details for its resilience towards fibril formation are not known in detail, but it is believed that factors such as a higher tendency to form a helix structure in NAC of  $\gamma$ Syn might stabilise the protein on a monomeric level<sup>366</sup>, and important transient intramolecular interactions driving self-assembly are altered in this protein<sup>333</sup>. Changes in fibrillation might therefore be caused by sequence changes between the paralogues and the resulting altered conformational ensembles of the monomer.

The data presented in Chapter 4 together with different literature reports (for instance the presence of different single point mutations causing disease<sup>248</sup> (Section 1.4.2.1), highlight the fact that even a single residue swap can have a significant effect on aggregation kinetics and/or toxicity of the protein. In the case of  $\alpha$ Syn, 20 neurodegenerative disease associated missense mutations (familial and sporadic) in the SNCA gene are recognised (see Table 1.3). In contrast, for  $\gamma$ Syn no such links were made to date.



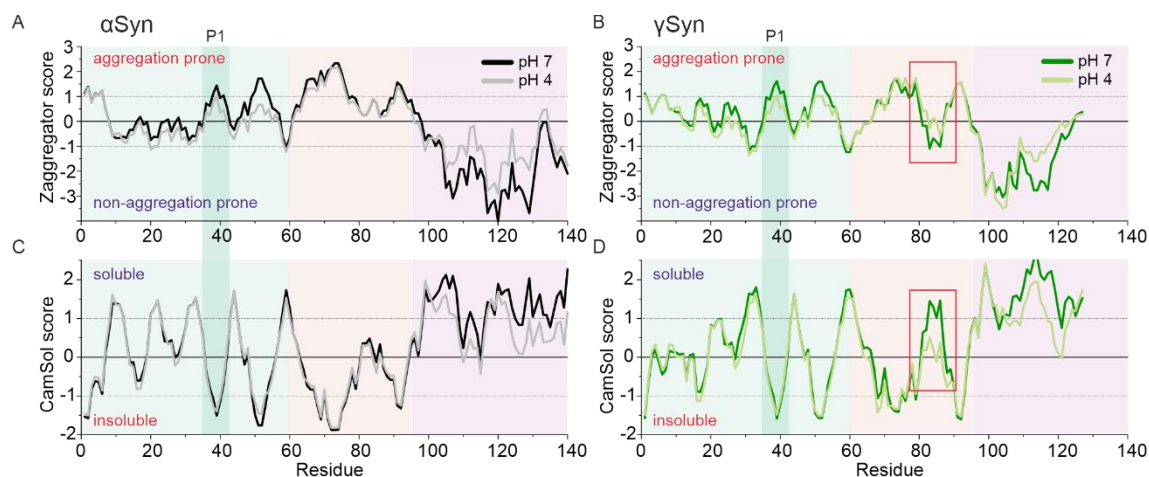
**Figure 5.1: Sequence alignment between αSyn WT and γSyn WT.** (A) Schematic of α- and γSyn protein. Boxes show N-terminal region (green), NAC region (red) and C-terminal region (purple). (B) Illustrates the sequence alignment between αSyn and γSyn. The colour code is accordingly to (A). Familial PD mutations are highlighted in blue, residues of interested in this study are coloured in green, the repetitive KTKEGV motif is highlighted in yellow, charge changes in the NAC region between αSyn and γSyn are coloured in orange (G → E) and the P1 region is highlighted in grey. Sequence identities of each region are indicated next to the sequence, ‘.’ = identity; ‘-’ = deletion of an amino acid.

However, accumulation of γSyn aggregates has been shown to contribute to the development of motor neuron pathology in amyotrophic lateral sclerosis (ALS). Distinct profiles containing fibrillated γSyn were found in ALS patients<sup>375</sup> and overexpression of γSyn in mice models resulted in the progression of middle-age onset motor neuron pathology that recapitulates many key characteristics of ALS, and premature death<sup>376,377</sup>. It should be considered that ALS is usually linked with amyloid formation of the protein TDP-43<sup>174</sup>. As it is the case for most neurodegenerative diseases, though, the disease is not only linked with one protein but also the accumulation of other molecules such as FUS and SOD1<sup>467</sup>. Therefore a correlation of γSyn aggregation and ALS, although only little research data have been obtained to date, is conceivable.

This chapter will focus on understanding the aggregation propensity of γSyn WT and its correlation to neurodegenerative diseases in more detail. Sequencing the SNCG genome of a small cohort of ALS patients identified a single residue missense mutation leading to γSyn M38I located in P1 and E110V in the C-terminal region (data from the Illarioshkin group, Research Centre of Neurology, Russia). By using *in vitro* aggregation assays, cell culture experiments (performed by the Outeiro group, Max Planck Institute for Experimental Medicine, Germany) and the model organism *C. elegans*, the role of P1 and especially residue 38 will be further investigated and a high sequence specificity of this motif for controlling aggregation of γSyn is identified.

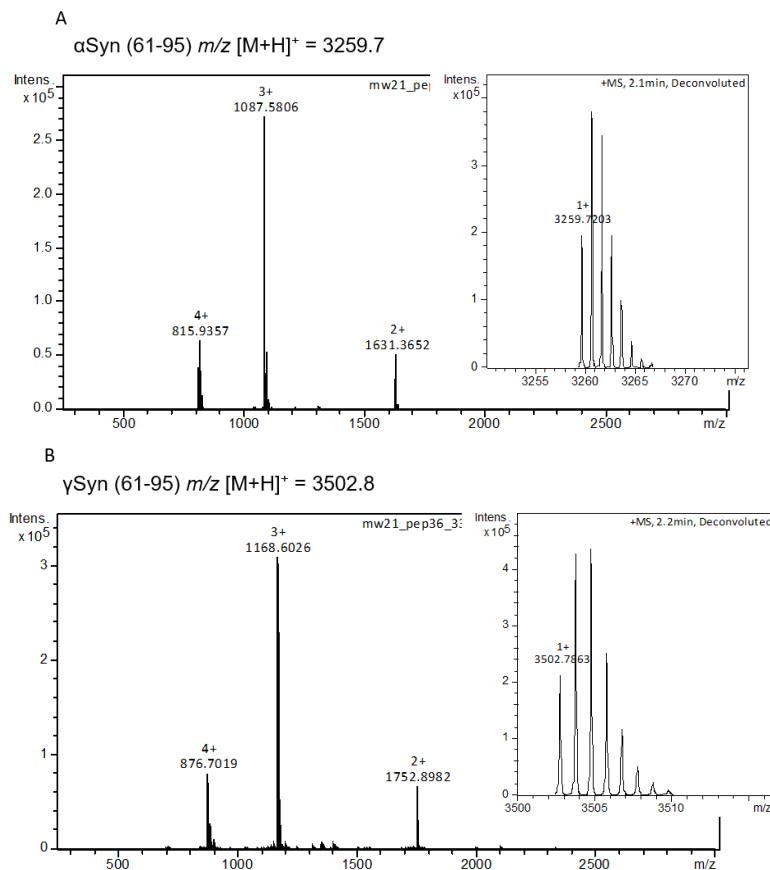
## 5.2 Reduced aggregation propensity for $\gamma$ Syn might be driven by sequence changes in NAC

Full length  $\alpha$ Syn is significantly more aggregation-prone compared with  $\gamma$ Syn under most buffer conditions (see Section 1.5 and **Figure 4.21**)<sup>365,378</sup>. A sequence alignment between these two paralogues reveal a 77 %, 51 % and 3 % sequence identity between N-terminal, NAC, and C-terminal regions, respectively (**Figure 5.1 B**). Despite differences in the sequence, the overall predicted aggregation propensity (Zygggregator score<sup>393</sup>) and solubility (CamSol<sup>114</sup>) is similar for  $\alpha$ - and  $\gamma$ Syn, with an aggregation-prone/insoluble N7 (1-7), P1 (36-42), P2 (45-57), and NAC (61-95) regions and a non-aggregation-prone/ soluble C-terminal region (96-140 or 96-127 for  $\alpha$ - and  $\gamma$ Syn, respectively) (**Figure 5.2**). However, there are some subtle differences as for example a higher pH dependence for  $\gamma$ Syn NAC in the predictions, indicating a lower aggregation propensity at neutral pH versus acidic conditions for residues 80-90. The central NAC region, is known to be sufficient and necessary for aggregation of  $\alpha$ Syn (residues 71-82)<sup>313</sup>, forms the core of all  $\alpha$ Syn fibril structures determined to date<sup>96,117</sup> (Section 1.4.3.3) and its identity was shown to play a crucial role in the chimeric experiments mixing  $\alpha$ - and  $\gamma$ Syn regions (Section 4.8, **Figure 4.28**). Substitutions within NAC might therefore be responsible for changes in fibril formation propensity of  $\alpha$ - and  $\gamma$ Syn. This could be observed for the low aggregation-prone  $\beta$ Syn, where changing residues 63–66 and 71–72 to the  $\alpha$ Syn sequence switched on fibrillation of the full-length  $\beta$ Syn protein<sup>468</sup>. A deeper look into the sequence of  $\gamma$ Syn NAC ( $\gamma$ NAC) reveals three Gly to Glu substitutions compared with  $\alpha$ Syn (**Figure 5.1 B**, highlighted in orange), which might be responsible for the reduced aggregation propensity of  $\gamma$ Syn at neutral pH and  $\gamma$ Syn being more sensitive to pH changes in the *in silico* analysis (**Figure 5.2 B**, red boxes). Note that both, pH 4.5 and 7.5 are important in a biological context as acidic pH mimics a lysosomal environment and neutral pH imitates a cytosolic milieu.



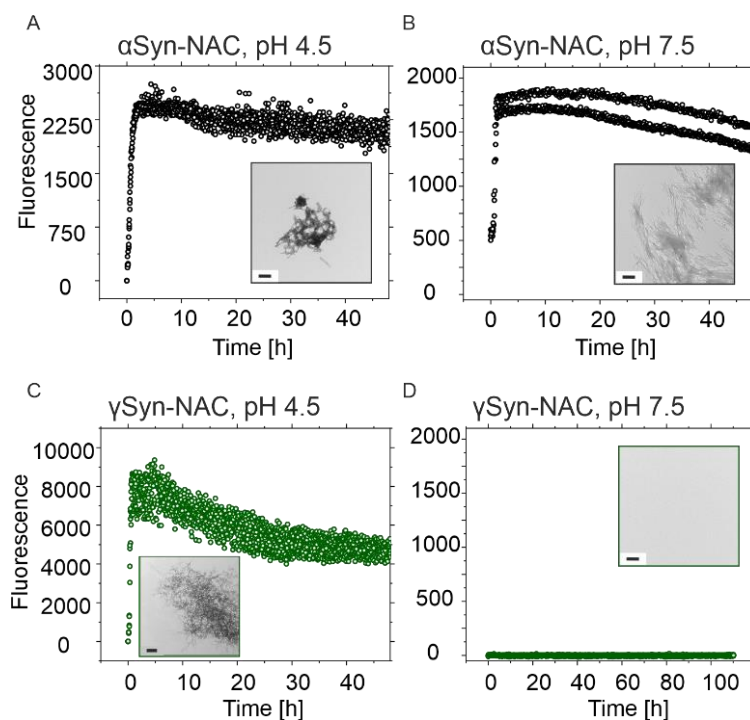
**Figure 5.2: *In silico* analysis of  $\alpha$ Syn and  $\gamma$ Syn aggregation propensity and solubility.** Aggregation propensity was analysed using the Zyggregator score for  $\alpha$ Syn (A) and  $\gamma$ Syn (B) at pH 4.0 (grey, light green for  $\alpha$ Syn and  $\gamma$ Syn, respectively) and pH 7.0 (black, dark green for  $\alpha$ Syn and  $\gamma$ Syn, respectively). Aggregation-prone regions have a positive Zyggregator score whereas non-aggregation-prone regions show negative values. N-terminal, NAC, C-terminal, and P1 regions are coloured accordingly in light green, red, purple, and dark green respectively. Analysis of solubility using CamSol of  $\alpha$ Syn (C) and  $\gamma$ Syn (D) at pH 4.0 (grey, light green for  $\alpha$ Syn and  $\gamma$ Syn, respectively) and pH 7.0 (black, dark green for  $\alpha$ Syn and  $\gamma$ Syn, respectively). Soluble regions have a positive CamSol score whereas insoluble regions show negative values. The red box in B,D (around residues 80-90) highlights the stronger pH dependence of  $\gamma$ Syn within the NAC region.

To validate how important the NAC region is for  $\alpha$ Syn and  $\gamma$ Syn aggregation (independently from the N- and C-terminal regions), ThT fluorescence aggregation assays with the NAC peptide (residues 61-95) were performed.  $\alpha$ NAC and  $\gamma$ NAC peptides were synthesised (with N- and C-terminal amide, see **Table 2.1**) and characterised via LC-MS by Martin Walko, University of Leeds (**Figure 5.3**).

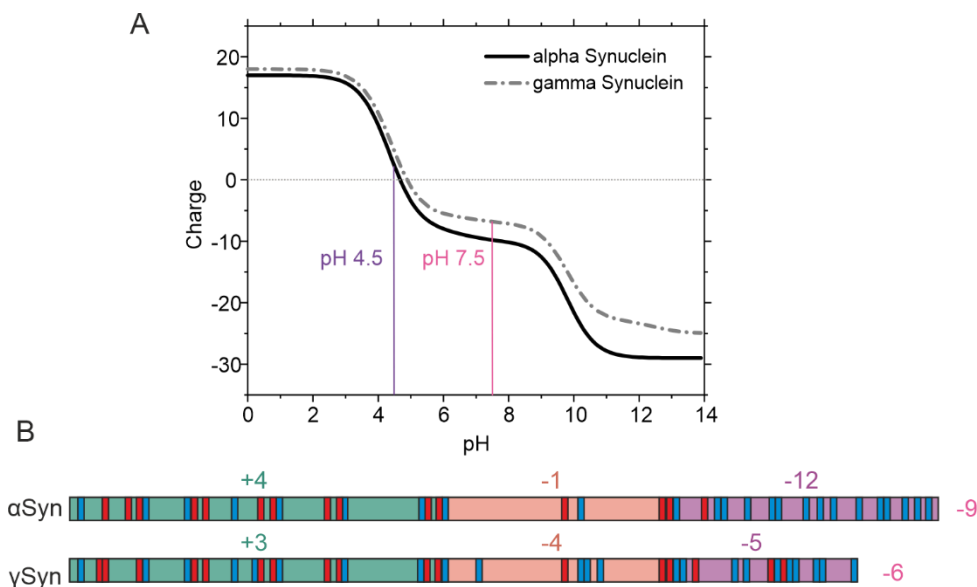


**Figure 5.3: LC-ES-MS analysis of  $\alpha$ NAC and  $\gamma$ NAC synthesised peptides.**  $\alpha$ NAC: expected mass = 3260.6 Da; measured mass = 3259.7 Da.  $\gamma$ NAC: expected mass = 3503.8 Da; measured mass = 3502.8 Da. Synthesis and MS analysis were performed by Martin Walko, University of Leeds.

The *in vitro* aggregation assays revealed that, similar to the full-length protein (**Figure 4.21**),  $\gamma$ NAC is fibrillating significantly slower (no amyloid material after 110 h incubation) compared with  $\alpha$ NAC at pH 7.5 and further shows a stronger pH dependence compared with  $\alpha$ NAC. While  $\alpha$ NAC forms fibrils at pH 4.5 and 7.5 immediately after initiating the experiment, the  $\gamma$ Syn peptide only self-assembles into amyloid fibrils at acidic pH (amyloid driving conditions) but does not form fibrils for at least 110 h at pH 7.5 (**Figure 5.4**). These data indicate that it could be the NAC region controlling the aggregation propensity of  $\gamma$ Syn, especially at neutral pH. Interestingly, the net charge of the NAC region (as well as C-terminal region) changes significantly, resulting in a less negative overall net-charge for the full-length protein compared with  $\alpha$ Syn (-9 and -6 for  $\alpha$ Syn and  $\gamma$ Syn, respectively) (**Figure 5.5**), which might affect the aggregation propensity, since the charge of the protein is known to have a large effect on its aggregation propensity<sup>189,233</sup>.



**Figure 5.4: Aggregation kinetics of NAC peptide (residues 61-95) for  $\alpha$ Syn (black) (A,B) and  $\gamma$ Syn (green) (C,D) at pH 4.5 (A,C) and pH 7.5 (B,D).** Aggregation assays were performed in 20 mM sodium acetate (pH 4.5) or 20 mM Tris-HCl (pH 7.5) in the presence of 200 mM NaCl, at 37°C, and 600 rpm shaking. Scale bar for TEM is 200 nm.

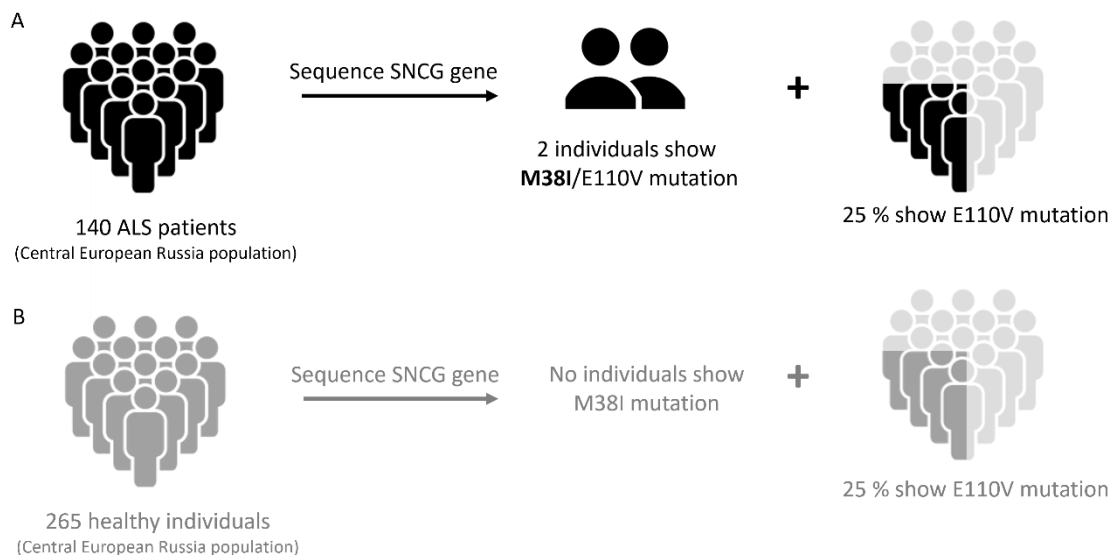


**Figure 5.5: Net charge of full-length  $\alpha$ Syn (WT) and  $\gamma$ Syn (WT).** (A) Charge calculated using Prot pi (<https://www.protpi.ch/Calculator/ProteinTool>). pH 4.5 and 7.5 are highlighted for easier orientation in purple and pink, respectively. (B) Schematic of  $\alpha$ Syn (top) and  $\gamma$ Syn (bottom) with their charged residues indicated (at neutral pH). The N-terminal region, NAC and C-terminal region are coloured in green, light red and purple, respectively. Positively charged residues (lysine) are highlighted in red, negatively charged residues (aspartate and glutamate) are highlighted in blue. The net charge at neutral pH of the three regions is indicated above each region, the net charge of the full-length protein is written in pink next to the schematic.



### 5.3 Identification of an ALS-associated substitution outside NAC

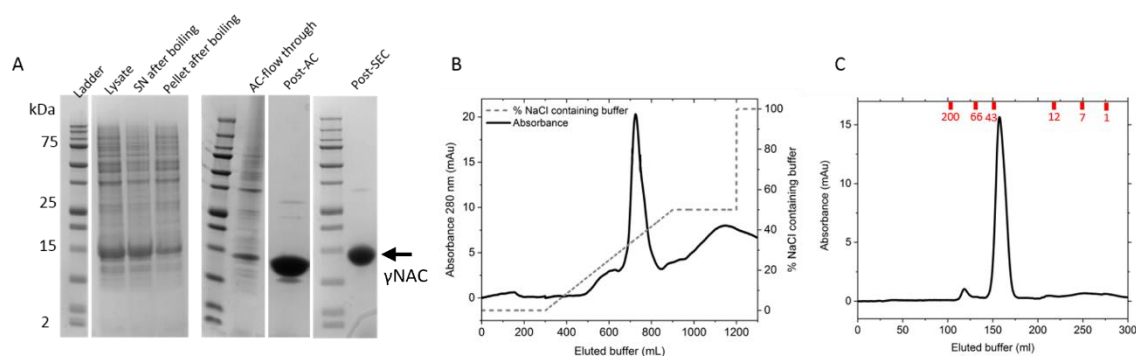
Exon sequencing of the SNCG gene was performed by the Illarioshkin group, Research Centre of Neurology, Russia and carried out in cohorts of ALS patients and healthy control individuals from the Central European Russia population. A single nucleotide G→A substitution in the first exon of the gene in two patients with young spinal onset (35-year-old male and 36-year-old female) of total 140 assessed ALS patients (minor allele frequency (MAF) 0.00714) was identified (**Figure 5.6 A**). This genetic variant corresponding to SNP rs148591902, has not been found in 265 healthy individuals from the same population (**Figure 5.6 B**) and it is rare in the publicly available databases (e.g. MAF is 0.00024 for the European population and 0.00012 overall in gnomAD (international Genome Aggregation Database)). This nucleotide substitution leads to a M38I amino acid substitution (located within the P1 region) in the encoded  $\gamma$ Syn protein. Both patients carrying this substitution were also heterozygous for the previously described SNP rs9864 (MAF is 0.23-0.27 in all available databases for ALS patients and general population) in fourth exon, which minor allele encodes valine instead of a more common glutamic acid at position 110 (located in C-terminal region) of  $\gamma$ Syn sequence<sup>469</sup>. This E110V substitution can be found in ~25% of SNCG alleles in all studied populations. Interestingly, both these substitutions are located outside the important NAC region.



**Figure 5.6: Schematic of SNCG genome sequencing of the central European Russia population. (A)** The genome of 140 ALS patients was sequenced identifying that 2 individuals carry a M38I/E110V mutation. 25 % of the population show the E110V variation. **(B)** the healthy control group (265 individuals) do not have the identified M38I mutation. Also 25 % present the E110V variation. Data were provided by the Illarioshkin group, Research Centre of Neurology, Russia.

## 5.4 Protein expression and purification

The finding of these disease-associated genetic variants of  $\gamma$ Syn raises the question of whether the resulting mutations (M38I and/or E110V) change the aggregation kinetics of the full length protein. To evaluate the aggregation propensity *in vitro*,  $\gamma$ Syn protein variants were expressed and purified as described before (see Section 2.6). An example of the resulting purification gel, anion exchange chromatogram, and size exclusion chromatogram is shown in **Figure 5.7** for  $\gamma$ Syn WT. As seen for  $\alpha$ Syn (**Figure 3.2**),  $\gamma$ Syn also elutes at a higher apparent molecular mass than predicted for a 13 kDa globular protein due to its intrinsically disordered conformation. It should be noted that the extinction coefficient of  $\gamma$ Syn WT is significantly lower than for  $\alpha$ Syn ( $\epsilon=1490 \text{ M}^{-1}\text{cm}^{-1}$  and  $\epsilon=5960 \text{ M}^{-1}\text{cm}^{-1}$ , respectively) resulting in lower absorbances (at 280 nm) in the chromatograms.  $\gamma$ Syn yields were in general lower than for  $\alpha$ Syn with approximately 5 mg pure protein per litre growth. **Table 5.1** summarises all generated proteins in this chapter with their expected and measured molecular mass (performed by the MS facility).



**Figure 5.7: Expression and purification of  $\gamma$ Syn.** (A) Example SDS-PAGE gel of purification steps of  $\gamma$ Syn WT.  $\gamma$ Syn band is at  $\sim 13$  kDa. (B) AEX chromatogram of  $\gamma$ Syn WT. Absorbance at 280 nm and concentration of salt-containing buffer (20 mM Tris-HCl, pH 8.0, 1M NaCl) are plotted against the elution volume.  $\gamma$ Syn eluted after  $\sim 700$  mL. (C) SEC chromatogram of  $\gamma$ Syn WT. Absorbance at 280 nm is plotted against elution volume.  $\alpha$ Syn eluted after  $\sim 150$  mL. Calibrants (Figure 2.1) are shown in red. Detailed expression and purification methods are provided in Section 2.6.

**Table 5.1: All purified  $\alpha$ Syn and  $\gamma$ Syn variants, expected molecular masses (top) and measured molecular masses (via ESI MS, bottom) of the proteins used in this chapter.**

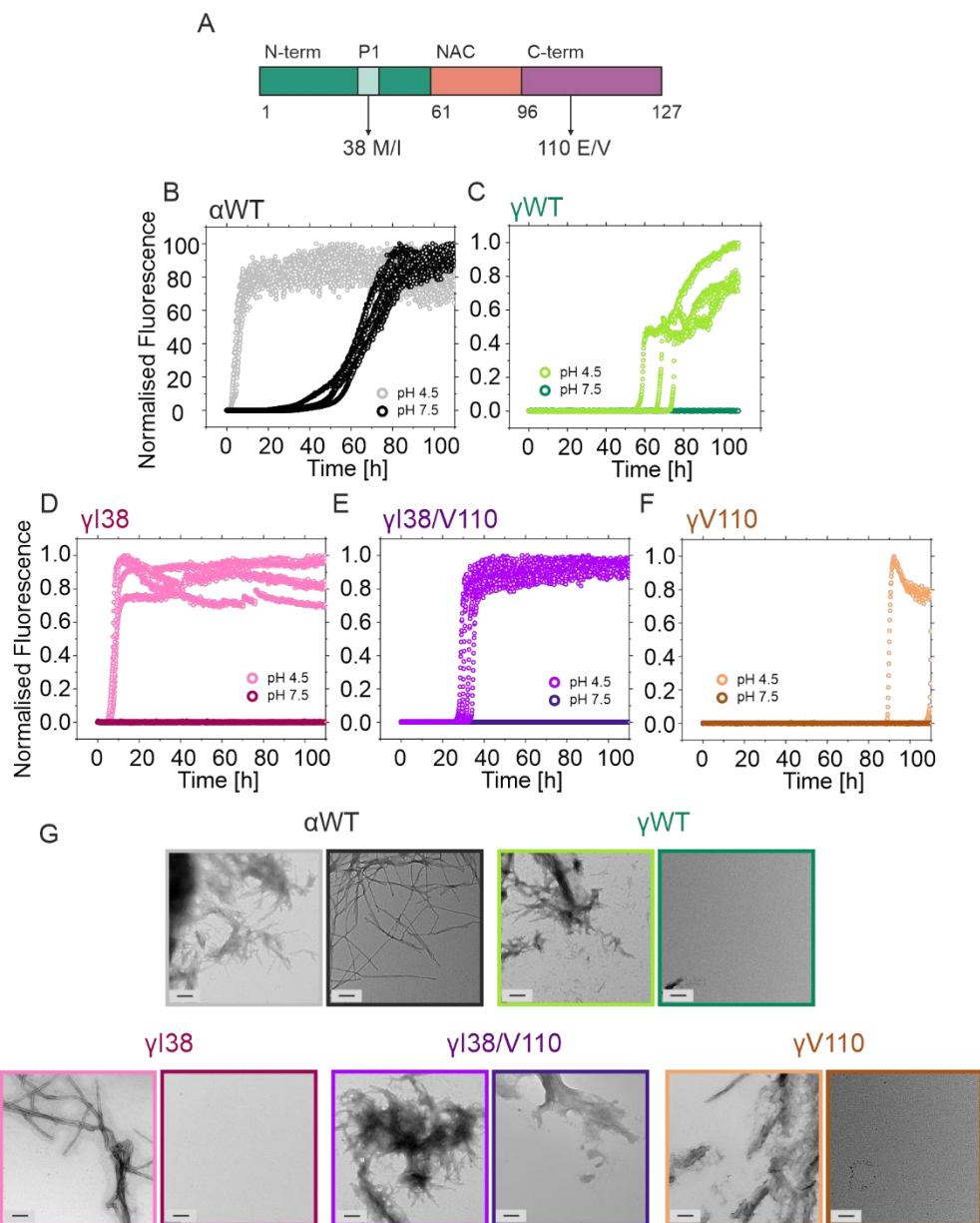
<sup>15</sup>N labelled proteins are indicated with a superscript 15.

Variant	Mass	Variant	Mass
$\alpha$ WT	14,460 Da	$\gamma$ L38	13,313 Da
	$14459 \pm 0.3$ Da		$13,313 \pm 0.09$ Da
$\gamma$ WT	13,331 Da	$\gamma$ V38	13,299 Da
	$13,330 \pm 0.2$ Da		$13,300 \pm 0.1$ Da
<sup>15</sup> $\gamma$ WT A90C	13,521 Da	$\gamma$ A38	13,271 Da
	$13,517 \pm 0.8$ Da		$13,270 \pm 0.06$ Da
$\gamma$ I38	13,313 Da	$\gamma$ I38/V110	13,283 Da
	$13,313 \pm 0.5$ Da		$13,282 \pm 0.4$ Da
<sup>15</sup> $\gamma$ I38 A90C	13,503 Da	$\gamma$ V110	13,301 Da
	$13,495 \pm 0.7$ Da		$13,301 \pm 0.1$ Da

## 5.5 Presence of isoleucine at position 38 dramatically increase $\gamma$ Syn propensity to form amyloid structures *in vitro* and in cell culture

The *in vitro* aggregation kinetics of  $\alpha$ Syn WT,  $\gamma$ Syn WT and the ALS linked  $\gamma$ Syn variants ( $\gamma$ I38,  $\gamma$ I38/V110, and  $\gamma$ V110) were measured by a ThT fluorescence assay at acidic (pH 4.5) and neutral pH (pH 7.5) followed by imaging via negative stain TEM and determining the amount of pelletable material (**Figure 5.8 and Table 5.2**) (see Section 2.8.2, 2.8.4 and 2.12 for detailed methods).

In accordance with data observed before (**Figure 3.3 and Figure 4.3**),  $\alpha$ Syn aggregates with lag times of  $0.9 \pm 0.7$  h and  $50.2 \pm 1.9$  h for pH 4.5 and 7.5, respectively (at high salt; 200 mM NaCl), while  $\gamma$ Syn WT shows a significantly reduced aggregation propensity with fibrils being formed after  $66.0 \pm 8.0$  h at acidic pH and no detectable fibrillation at neutral pH within the experimental time of 100 h (**Figure 5.8 B,C and Table 5.2**). The M38I substitution shows a drastic increase in aggregation kinetics at acidic pH (lag time =  $6.7 \pm 0.6$  h) and TEM imaging and pelleting assay analysis confirmed the presence of amyloid fibrils at the end of the ThT fluorescence assay (**Figure 5.8 D,G and Table 5.2**). The double amino acid swap found in 1.5 % of the analysed ALS patients (M38I + E110V) (see **Figure 5.6**) reduces the aggregation propensity compared with the single substitution  $\gamma$ I38, but is still significantly more aggregation-prone than  $\gamma$ WT with a lag time of  $30.7 \pm 2.6$  h at acidic pH (**Figure 5.8 E and Table 5.2**).  $\gamma$ V110 (seen in 25 % of the population<sup>469</sup>) shows no fibril formation under the tested conditions (**Figure 5.8 F,G**). These data point to an important role of regions flanking NAC for the aggregation of  $\gamma$ Syn. A single residue substitution within the P1 region can completely change the aggregation kinetics of this IDP although the NAC region remains unchanged. This provides first evidence that the P1 region plays a key role in  $\gamma$ Syn self-assembly into amyloid structures and regulating fibril formation of the full-length protein similar to the role described for P1 in  $\alpha$ Syn in Chapter 3 and 4.



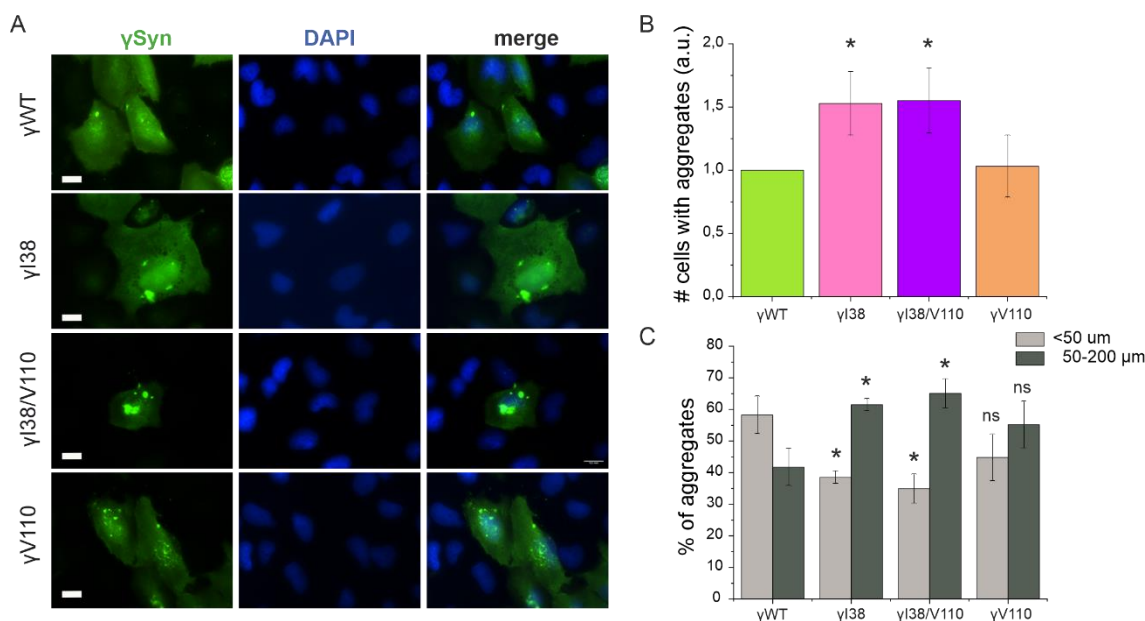
**Figure 5.8: *In vitro* aggregation assay of genetic variants associated with ALS patients.** (A) Schematic of  $\gamma$ Syn construct used for the aggregation assay. N-terminal region (green), NAC region (red), C-terminal region (purple) and P1 region (light green) are highlighted. The residue variants 38M/I and 110E/V are indicated. (B-F) Normalised ThT fluorescence for aggregation assays with  $\alpha$ L38/E110 (B),  $\gamma$ M38/E110 (C),  $\gamma$ I38/E110 (D),  $\gamma$ I38/V110 (E) and  $\gamma$ M38/V110 (F). Aggregation assays were performed in 20 mM Tris-HCl, pH 7.5 or 20 mM sodium acetate, pH 4.5 with 200 mM NaCl. (G) TEM images for endpoint samples from ThT-assay at pH 4.5 and at pH 7.5, scale bar 200 nm.

**Table 5.2: Aggregation kinetics of ALS gene polymorphs M38/I38 and E110/V110.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\gamma$ Syn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 110 h incubation for the *de novo* growth is indicated by “-”. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in  $10^4$ [RFU/h]. Assays were performed in 20 mM sodium acetate, pH 4.5 or 20 mM Tris-HCl, pH 7.5 with 200 mM NaCl.

	$\alpha$ WT, pH 4.5	$\alpha$ WT, pH 7.5	$\gamma$ WT, pH 4,5	$\gamma$ WT, pH 7.5	$\gamma$ I38, pH 4.5	$\gamma$ I38, pH 7.5
Lag time [h]	0.9 $\pm$ 0.7	50.2 $\pm$ 1.9	66.0 $\pm$ 8.0	-	6.7 $\pm$ 0.6	-
Elongation rate*	1.3 $\pm$ 0.5	0.3 $\pm$ 0.06	0.02 $\pm$ 0.01	-	2.3 $\pm$ 0.5	-
$t_{50}$	5.0 $\pm$ 1.0	67.6 $\pm$ 3.3	74.7 $\pm$ 5.2	-	8.3 $\pm$ 0.9	-
% pellet	100	65	95	0	100	5
	$\gamma$ I38/V110, pH 4.5	$\gamma$ I38/V110, pH 7.5	$\gamma$ V110, pH 4.5	$\gamma$ V110, pH 7.5		
Lag time [h]	30.7 $\pm$ 2.6	-	-	-		
Elongation rate*	4.2 $\pm$ 0.6	-	-	-		
$t_{50}$	34.3 $\pm$ 3.0	-	-	-		
% pellet	80	5	5	0		

To confirm the *in vitro* observations in a cellular context, the  $\gamma$ Syn variants were expressed in human H4 neuroglioma cells, a model that is commonly used to assess the aggregation of  $\alpha$ Syn<sup>298,470</sup>. The cell culture experiments were performed and analysed by the Outeiro group, Max Planck Institute for Experimental Medicine, Germany. Consistent with the *in vitro* data, cells expressing  $\gamma$ I38 or  $\gamma$ I38/V110 formed approximately 1.5 times more microscopically-visible  $\gamma$ Syn inclusions than cells expressing  $\gamma$ WT or  $\gamma$ V110 (**Figure 5.9 A,B**). Moreover, in cells expressing  $\gamma$ WT only about 35% of inclusions were large (>50  $\mu$ m in diameter), whereas there were nearly twice as many (~ 60 %) of these large aggregates in cells expressing variants with isoleucine ( $\gamma$ I38 and  $\gamma$ I38/V110) (**Figure 5.9 C**).

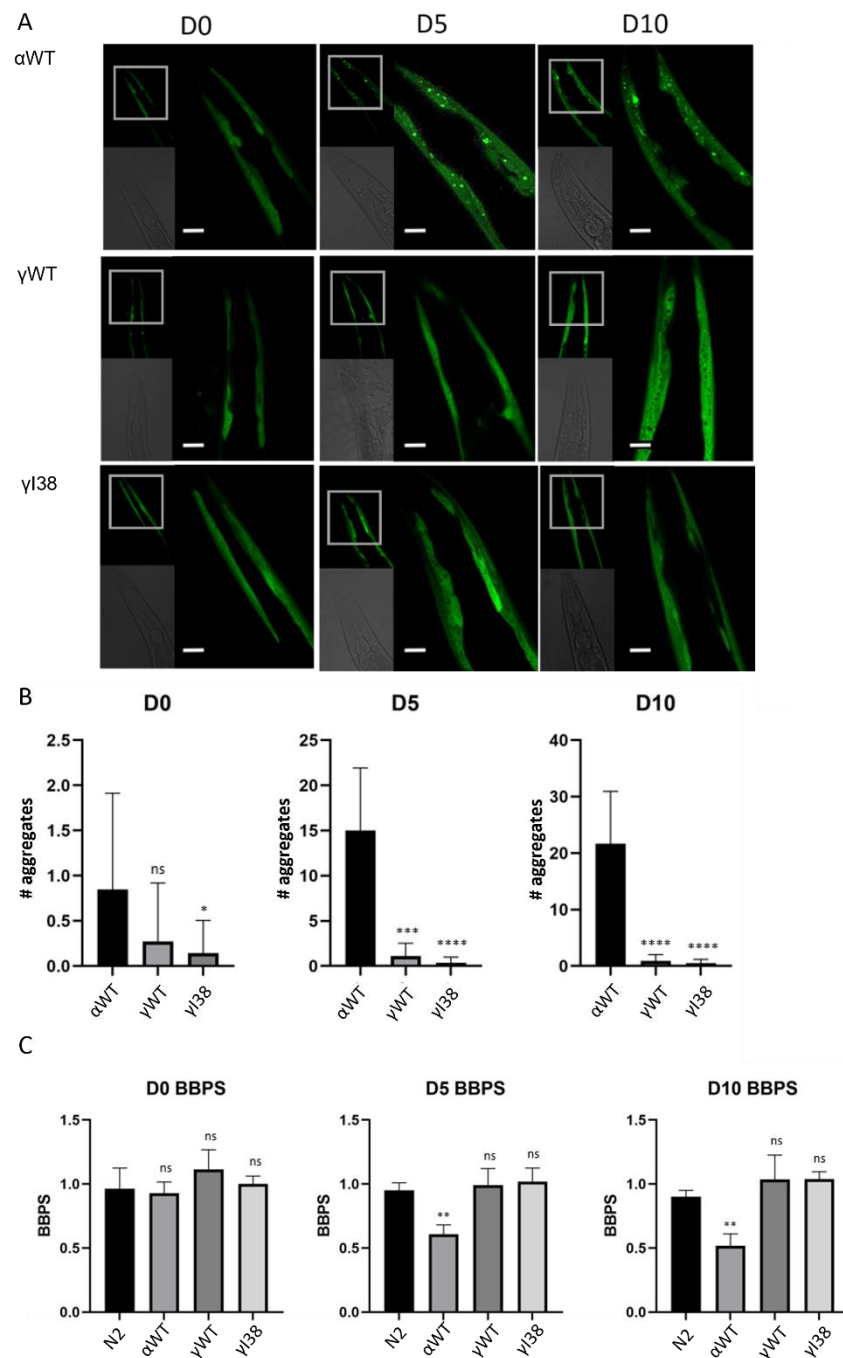
Taken together, the *in vitro* and cell culture experiments highlight the increased aggregation propensity for  $\gamma$ Syn variants with a Met38Ile substitution. As this substitution has been found in ALS patients but not healthy individuals from the same population it could be speculated that the increased propensity to form aggregates relates to the development or progression of the disease as it has been observed for some familial PD point mutations in  $\alpha$ Syn<sup>145,471,472</sup>.



**Figure 5.9: Inclusions formed by  $\gamma$ Syn variants in human H4 cells.** (A) Following transfection with the respective expression plasmids, cells were immunostained with anti- $\gamma$ Syn antibody, nuclei were counterstained with DAPI. Scale bar = 0.5  $\mu$ m (B) Quantification revealed significant increase in percent of transfected cells that formed inclusions after expression of  $\gamma$ I38 or  $\gamma$ I38/V110 variants compared with expression of the  $\gamma$ WT variant (\* $p < 0.05$ , Student's t test,  $n=4$  independent experiments). (C) Analysis of inclusion sizes reveal that in cells expressing  $\gamma$ I38 or  $\gamma$ I38/V110 larger inclusions are more common than in cells expressing  $\gamma$ WT. The bar chart shows the mean  $\pm$  SD of percent of small (diameter  $<50$  nm) and large (diameter 50-200 nm) aggregates in transfected cells (\* $p < 0.05$ ,  $n=4$  independent experiments). The data presented here were provided by the Outeiro group, Max Planck Institute for Experimental Medicine, Germany.

## 5.6 Met38Ile substitution does not change toxicity in *C. elegans*

In a next step the effect of the identified M38I substitution was evaluated in the model organism *C. elegans*. As described in Section 2.15.3 the protein was expressed in the body wall muscle cells of the worms with a C-terminally attached YFP for visualisation.  $\alpha$ Syn WT,  $\gamma$ Syn WT and  $\gamma$ I38 (generated by Q5 mutagenesis, see Section 2.5.1) expressing animals were investigated by determining the number of aggregates in the head region and measuring the phenotypic effect via a motility assay (BBPS) (Figure 5.10). As shown in Figure 3.11 and Figure 4.25, the  $\alpha$ WT::YFP expressing strains present an increase in inclusion formation over time (0.8, 15, and 20 inclusions per worm at day 0, 5, and 10, respectively) (Figure 5.10 A,B) resulting in a loss of motility as the body bends per second decrease from 1.0 to 0.6 and 0.5 for days D0, D5, and D10 (Figure 5.10 C). The  $\gamma$ WT::YFP and  $\gamma$ I38::YFP expressing animals, in contrast show no increase in puncta formation even at D10 and no significant decrease in the motility was observed compared to the control worms N2 (no expression of synuclein) (Figure 5.10).



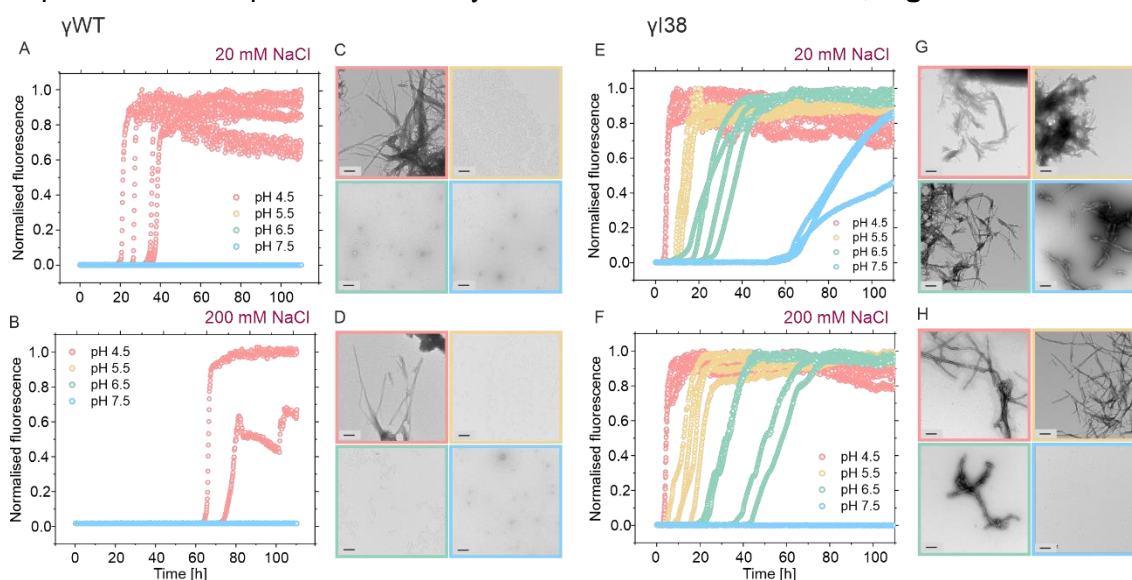
**Figure 5.10: Expressing  $\gamma$ I38 in worm muscle cells does not affect the toxicity. (A)**

Confocal microscopy images showing the head region of transgenic *C. elegans* expressing  $\alpha$ Syn WT,  $\gamma$ Syn WT or  $\gamma$ I38 tagged C-terminally to YFP in the body wall muscle during aging (day 0, day 5 and day 10 of adulthood). Scale bars, 10  $\mu$ m. (B) Numbers of inclusions larger than  $\sim 2 \mu\text{m}^2$  per animal between the tip of the head region and pharyngeal bulb during aging. Data shown are the mean and s.e.m. for three independent experiments (biological replicates); in each experiment, ten worms ( $n = 10$ ) were assessed for each time point. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ . ANOVA was used in all cases. (D) Number of body bends per second (BBPS) of N2,  $\alpha$ Syn WT::YFP,  $\gamma$ Syn WT::YFP and  $\gamma$ I38::YFP animals from day 0 (L4 stage), day 5 and day 10 of adulthood. Data shown are mean and s.e.m. for three independent experiments; in each experiment, ten worms were assessed for each time point.  $n = 10$  for each experiment, and error bars represent s.e.m. of three biological replicates. n.s. = not significant; \*\*\* $P < 0.001$ ; \*\* $P < 0.01$ ; \* $P < 0.05$ , ANOVA was used.

The lack of an effect in an *in vivo* system between  $\gamma$ WT and  $\gamma$ I38 could have various reasons including a chaperone network in worms affecting amyloid formation<sup>153</sup>, or fibril formation being a protective mechanism of toxic oligomeric species<sup>159,160</sup>. Indeed, the familial disease-associated variant A30P in  $\alpha$ Syn also shows no increased toxicity in *C. elegans* compared with the WT protein using the same assays used in this study<sup>473</sup>. Further, it should be considered that *C. elegans* does not naturally express synucleins and therefore a loss-of-function aspect cannot be assessed using this model organism. Other organisms, such as zebrafish or mice might offer better biological systems to study the complexity of amyloid formation of  $\gamma$ Syn (Section 1.4.2.2). Also, expressing the  $\alpha$ - or  $\gamma$ Syn variants in the neurons (e.g. pan-neuronal cells) instead of the muscle cells potentially improves the outcome of this assay.

## 5.7 Ile38 aggregates under all biologically relevant conditions

Since  $\alpha$ Syn and  $\gamma$ Syn are intrinsically disordered proteins. Their aggregation is highly dependent on their buffer environment as shown in many studies for  $\alpha$ Syn with e.g. changing pH<sup>133,422</sup>, added ions<sup>189,315</sup> or different protein concentrations<sup>474</sup>. To further examine the aggregation kinetics of the newly identified  $\gamma$ Syn variant with the M38I substitution, fibril formation of  $\gamma$ WT and  $\gamma$ I38 was compared at different pH values (pH 4.5 to 7.5) and different ionic strengths (20 mM and 200 mM NaCl) (**Figure 5.11**). Similar experiments were performed for  $\alpha$ Syn WT and  $\Delta$ P1 in Section 4.2, **Figure 4.3**.



**Figure 5.11:  $\gamma$ I38 switches on aggregation with a strong buffer dependence.** pH dependent *de novo* aggregation of  $\gamma$ WT (**A,B**) and  $\gamma$ I38 (**E,F**). *In vitro* aggregation measured by ThT fluorescence assay at 20 mM (**A,D**) or 200 mM (**B,E**) NaCl. Note that for  $\gamma$ WT data for pH 5.5; 6.5 and 7.5 overlap. ThT fluorescence assay was performed at 20 mM sodium acetate (pH 4.5; 5.5), MES (pH 6.5) or Tris-HCl (pH 7.5) accordingly with 20 mM or 200 mM NaCl and 100  $\mu$ M protein at 37  $^{\circ}$ C, 600 rpm. Note that a constant ionic strength was considered for all buffers. TEM images show ThT assay end-point samples for  $\gamma$ WT (**C,D**) or  $\gamma$ I38 (**G,H**); scale bar 200 nm.



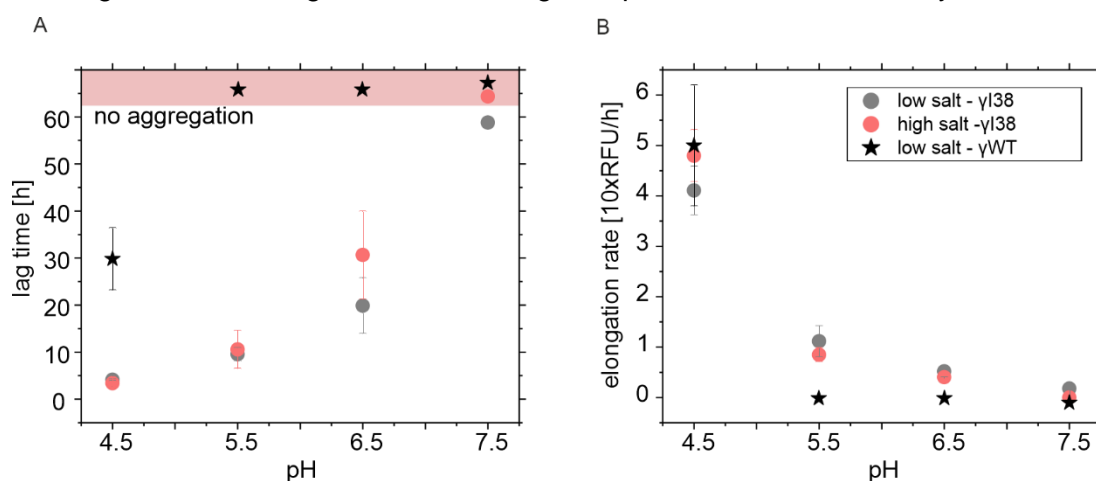
**Table 5.3: Aggregation kinetics of pH dependence of  $\gamma$ Syn WT and  $\gamma$ I38.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\gamma$ Syn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 110 h incubation for the *de novo* growth is indicated by “-“. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in  $10^*$ [RFU/h].

	$\gamma$ WT, 4.5, low	$\gamma$ WT, 5.5, low	$\gamma$ WT, 6.5, low	$\gamma$ WT, 7.5, low	$\gamma$ WT, 4.5, high	$\gamma$ WT, 5.5, high
Lag time [h]	29.8 $\pm$ 6.6	-	-	-	-	-
Elongation rate*	5.0 $\pm$ 1.2	-	-	-	-	-
$t_{50}$	30.7 $\pm$ 6.7	-	-	-	-	-
% pellet	95	10	0	10	75	0
	$\gamma$ WT, 6.5, high	$\gamma$ WT, 7.5, high	$\gamma$ I38, 4.5, low	$\gamma$ I38, 5.5, low	$\gamma$ I38, 6.5, low	$\gamma$ I38, 7.5, low
Lag time [h]	-	-	4.1 $\pm$ 0.1	9.6 $\pm$ 1.4	20.0 $\pm$ 5.9	58.8 $\pm$ 1.2
Elongation rate*	-	-	4.1 $\pm$ 0.5	1.1 $\pm$ 0.3	0.5 $\pm$ 0.1	0.2 $\pm$ 0.03
$t_{50}$	-	-	5.2 $\pm$ 0.2	13.7 $\pm$ 1.4	29.7 $\pm$ 4.9	80.0 $\pm$ 3.2
% pellet	5	0	100	100	70	15
	$\gamma$ I38, 4.5, high	$\gamma$ I38, 5.5, high	$\gamma$ I38, 6.5, high	$\gamma$ I38, 7.5, high		
Lag time [h]	3.4 $\pm$ 0.1	10.6 $\pm$ 4.0	30.6 $\pm$ 9.3	-		
Elongation rate*	4.8 $\pm$ 0.5	0.9 $\pm$ 0.1	0.4 $\pm$ 0.07	-		
$t_{50}$	4.4 $\pm$ 0.1	16.3 $\pm$ 3.3	42.7 $\pm$ 11.3	-		
% pellet	100	85	60	5		

As expected,  $\gamma$ WT displayed a low aggregation propensity under all tested conditions. Fibril formation could only be observed under acidic conditions (pH 4.5, 20 mM or 200 mM NaCl) and no amyloid structures were detected at pH 5.5; 6.5 or 7.5 measured via ThT fluorescence assay, TEM imaging and pelleting analysis (**Figure 5.11 A-D and Table 5.3**). This result fits with observation from the  $\gamma$ Syn-NAC aggregation measurement (**Figure 5.4**) in which fibril assembly was only observed at acidic pH but not at neutral pH. Interestingly, at low ionic strength (20 mM), the ALS linked variant  $\gamma$ I38 forms fibrils under all tested pH values within the experimental timescale of 110 h (**Figure 5.11 E**). Also, at 200 mM NaCl aggregation was measured at pH 4.5; 5.5, and 6.5 (**Figure 5.11 F**) indicating that this  $\gamma$ Syn variant is on the verge of aggregation under all tested biologically relevant buffer conditions. In all experiments resulting in a positive ThT signal, fibrils were observed by TEM (**Figure 5.11 G,H**). This is an exciting observation considering that  $pK_a$ , net-charge and charge pattern of the protein remain unchanged between  $\gamma$ WT and  $\gamma$ I38. Differences to  $\alpha$ Syn aggregation on the other hand might, at least partially, be explained by changes in net charge at neutral pH (**Figure 5.5**).

A more detailed analysis of the aggregation kinetics (lag time and elongation rate) determined by fitting a linear fit through the data in the elongation phase (from 40 % to 60 % maximum intensity) using Origin pro (Section 2.8.3) reveal a correlation between fibril assembly kinetics and pH value (**Figure 5.11, Figure 5.12 and Table 5.3**). Interestingly, at pH 4.5, lag times between  $\gamma$ WT and  $\gamma$ I38 are changed by a factor of 10, whereas the elongation rates are not massively affected ( $5.0 \pm 1.2$   $10^*$ RFU/h and  $4.1 \pm 0.5$   $10^*$ RFU/h for  $\gamma$ WT and  $\gamma$ I38, respectively at pH 4.5, low NaCl). This suggests that

$\gamma$ I38 probably acts predominantly at the early stages of aggregation in which nuclei and small oligomers are being formed, the elongation phase is less disturbed by the mutation.



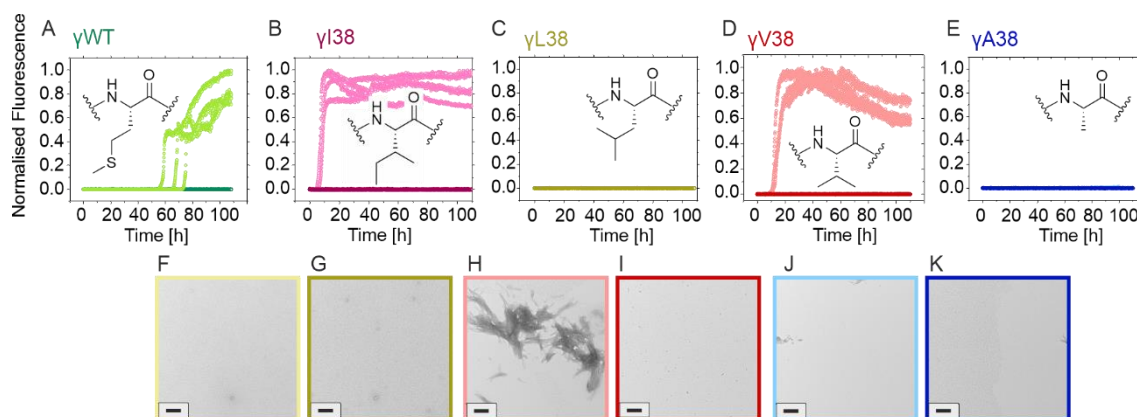
**Figure 5.12: Aggregation kinetics for  $\gamma$ WT and  $\gamma$ I38.** Lag Time (A) and elongation rate (B) at different pH values and ionic strengths (20 mM vs 200 mM NaCl) measured from data shown in **Figure 5.11** highlighting the strong buffer condition dependence. Both parameters were measured using the OriginPro software. Note that for  $\gamma$ WT the kinetic values are only shown at low salt as aggregation does not occur consistently at high ionic strength.

## 5.8 High sequence specificity is required to accelerate $\gamma$ Syn fibril formation

Sequence alignments between  $\alpha$ Syn and  $\gamma$ Syn show that the key residue identified in this study M/I38, switching off/on aggregation in  $\gamma$ Syn (**Figure 5.13 A,B**) is different to the more aggregation-prone  $\alpha$ Syn which has a leucine at residue 38 (**Figure 5.1 B**). It was shown in Chapter 4 that single residues in the P1 region of  $\alpha$ Syn are crucial for fibril formation. Surprisingly, the hypothesis that  $\gamma$ L38 might switch on its aggregation, as the sequence is more similar to the aggregation-prone  $\alpha$ Syn, was not confirmed, as  $\gamma$ L38 did not form fibrils at acidic or neutral pH, confirmed by TEM and pelleting assay (**Figure 5.13 C,F,G, Table 5.4** and **Figure 4.23**). Since it could be shown in Chapter 4 that the identity of the residue 38 in P1 of  $\alpha$ Syn plays a crucial role for the aggregation propensity of the full-length protein (see Section 4.5, **Figure 4.22**) with  $\alpha$ L38,  $\alpha$ A38,  $\alpha$ M38,  $\alpha$ I38 showing distinct aggregation kinetics), the effect of introducing other aliphatic residues in  $\gamma$ Syn at residue 38 was tested. Strikingly, replacing methionine 38 with a valine ( $\gamma$ V38) also induced aggregation, with similar kinetics compared with  $\gamma$ I38 at acidic pH (lag time of  $6.7 \pm 0.6$  h and  $11.1 \pm 0.8$  h for I38 and V38, respectively) (**Figure 5.13 D** and **Table 5.4**). This hints that the structural feature of a beta-branched amino acid sidechain, that preferably adopts  $\beta$ -sheet structures, might be important for fibril formation (see inserts in **Figure 5.13**). Finally, another hydrophobic residue ( $\gamma$ A38) was introduced at residue 38. Remarkably,  $\gamma$ A38 did not form fibrils under any conditions studied here (**Figure 5.13 E**), further highlighting the important role of a steric effect and/or tendency of the amino

acid to form a  $\beta$ -strand to accelerate aggregation. All ThT fluorescence assay results were confirmed by TEM imaging and pelleting analysis of the end-point samples (**Figure 5.8 C,D,G, Figure 5.13F-K and Table 5.4**).

Taken together, these data demonstrate a high sequence specificity for residue 38 in  $\gamma$ Syn to switch on aggregation at acidic pH possibly by promoting conformations driving nucleation/oligomerisation processes.



**Figure 5.13: The aggregation of  $\gamma$ Syn is highly residue specific.** Aggregation kinetics for  $\gamma$ WT (A),  $\gamma$ I38 (B),  $\gamma$ L38 (C),  $\gamma$ V38 (D), and  $\gamma$ A38 (E) at pH 4.5 (light colours) and 7.5 (dark colours), in the presence of 200 mM NaCl, at 37°C, and 600 rpm shaking. Note that if only a flat line is presented both conditions were measured, and no fibril formation was observed under both conditions. **F-K:** TEM images at ThT endpoints at acidic (F, H, J) and neutral (G, I, K) pH. Scale bar: 200 nm. TEMs for data shown in (A) and (B) are shown in **Figure 5.8 G**.

**Table 5.4: Aggregation kinetics of substitutions at residue 38 in  $\gamma$ Syn.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\gamma$ Syn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 110 h incubation for the *de novo* growth is indicated by “-”. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in  $10^*$ [RFU/h].

	$\gamma$ A38, pH 4.5	$\gamma$ A38, pH 7.5	$\gamma$ L38, pH 4.5	$\gamma$ L38, pH 7.5	$\gamma$ V38, pH 4.5	$\gamma$ V38, pH 7.5
Lag time [h]	-	-	-	-	$11.1 \pm 0.8$	-
Elongation rate*	-	-	-	-	$1.5 \pm 0.7$	-
$t_{50}$	-	-	-	-	$14.3 \pm 1.2$	-
% pellet	15	5	0	0	100	0

## 5.9 $\gamma$ I38 shows a distinct fibril architecture compared with $\alpha$ Syn WT

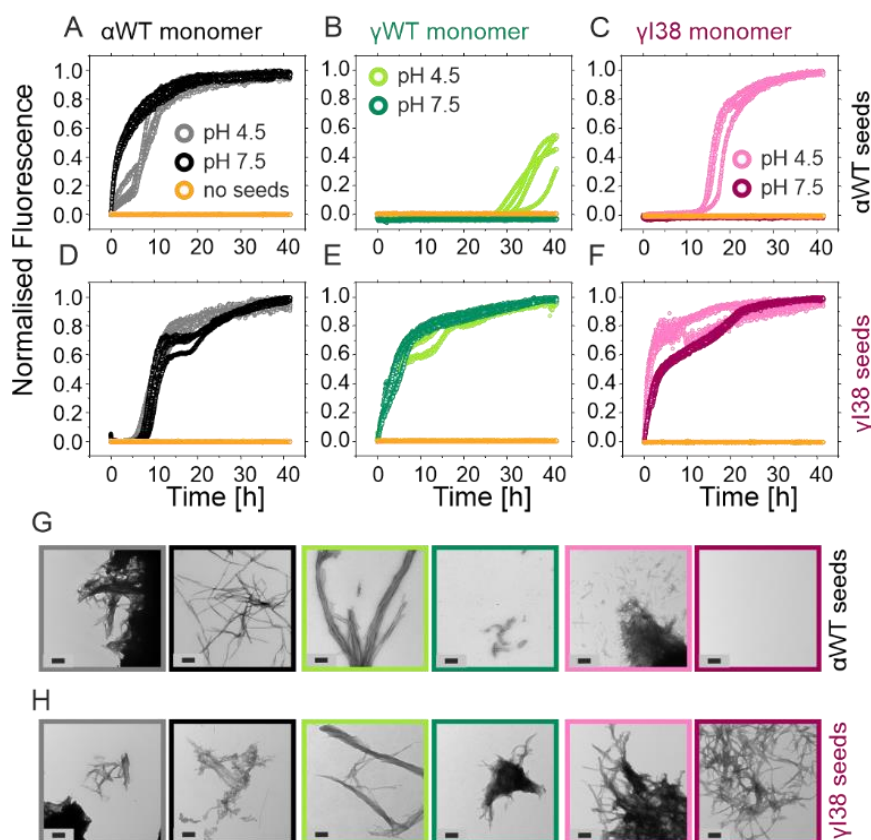
After demonstrating the crucial role of residue 38 to form fibril assemblies in both  $\alpha$ Syn and  $\gamma$ Syn *de novo*, the morphology and biochemical characteristics of the ALS-associated  $\gamma$ I38 fibrils formed *in vitro* were investigated in more detail. As stated in Section 1.2.3, fibril formation of amyloidogenic proteins is typically defined by a sigmoidal growth curve described by a nuclei- and oligomer-forming lag-phase, an elongation phase, which is dominated by fibril growth and secondary nucleation events; and a

stationary/plateau phase where mature fibrils are found in equilibrium with soluble monomer/oligomers<sup>138</sup>. The addition of pre-formed seeds obviates the need for nucleation and removes the rate-limiting nucleation phase if the monomer is seed-elongation-competent (**Figure 1.11**).

Whether the monomeric  $\gamma$ Syn variants ( $\gamma$ WT and  $\gamma$ I38) can elongate  $\alpha$ Syn WT seeds (10 % (mol/mol) monomer equivalent) (**Figure 5.14 A-C**) and if  $\alpha$ WT and  $\gamma$ WT monomers are seed-compatible with  $\gamma$ I38 fibrils (**Figure 5.14 D-F**) was next examined. As shown before in Chapter 3 and 4 (**Figure 3.9** and **Figure 4.20**) and the literature<sup>133</sup>,  $\alpha$ Syn WT fibrils efficiently self-seeded  $\alpha$ Syn monomers at both tested pH values (lag times of  $1.4 \pm 0.8$  h and 0 h for pH 4.5 and 7.5, respectively) (**Figure 5.14 A**). Interestingly, cross-seeding the  $\gamma$ Syn variants with  $\alpha$ Syn seeds was only successful at acidic pH and with significantly delayed lag times compared with self-seeding kinetics (lag times of  $31.7 \pm 4.1$  h and  $14.3 \pm 1.0$  h for  $\gamma$ WT and  $\gamma$ I38 monomers, respectively), and no elongation was observed at pH 7.5 (**Figure 5.14 B,C** and **Table 5.5**). It should be considered, that aggregation under acidic conditions might be affected by secondary nucleation, in which the seeds do not elongate at the fibril end but the fibril sides which provide a catalysing surface for aggregation. Buell *et al.* has shown that secondary nucleation is the dominant process for  $\alpha$ Syn aggregation below pH 5.8<sup>141</sup>.

The results for seeding with 10 % (mol/mol)  $\gamma$ I38 pre-formed fibrils indicated that all of the synuclein variants are able to elongate  $\gamma$ I38 seeds at pH 4.5 and 7.5, resulting in similar rates of elongation (**Figure 5.14 D-F** and **Table 5.5**). Note that all these experiments are performed under quiescent conditions (no shaking) where  $\alpha$ - and  $\gamma$ Syn do not fibrillate without the addition of aggregation accelerating material (e.g. seeds, membranes or glass beads)<sup>133,148</sup>. Control experiments with no seeds present did not result in fibril formation (**Figure 5.14 A-F**, orange datapoints). ThT endpoint samples were imaged by TEM and confirmed the presence of amyloid structures in ThT positive samples (**Figure 5.14 G,H**).

Taken together, the cross-seeding experiments provide the first evidence for distinct fibril architectures between  $\alpha$ Syn WT and  $\gamma$ I38 fibrils, as different seeding capacities for the fibril variants were observed.

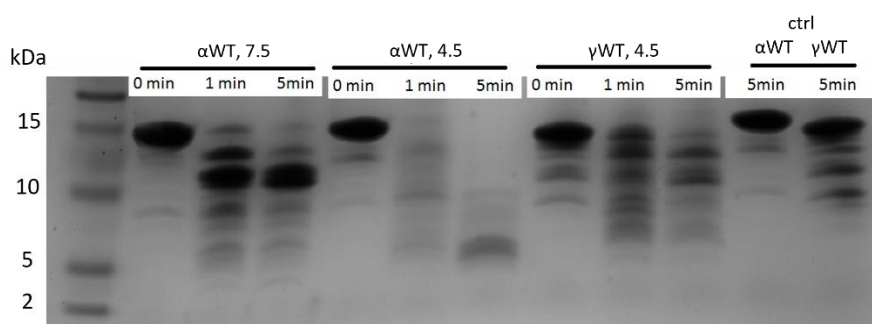


**Figure 5.14: Distinct seeding capacity for  $\alpha$ Syn WT and  $\gamma$ I38 pre-formed fibrils.** Seeded aggregation assays with pre-formed  $\alpha$ WT (grown at neutral pH) (A-C) and  $\gamma$ I38 (grown at acidic pH) (D-F) seeds.  $\alpha$ WT (A,D),  $\gamma$ WT (B,E), and  $\gamma$ I38 (C,F) monomer were cross-seeded with 10 % seeds ((mol/mol) monomer equivalent) in 20 mM Tris-HCl or ammonium acetate, 200 mM NaCl, at 37°C, quiescent conditions (no shaking compared to *de novo* fibril growth in **Figure 5.8**). Orange curves show aggregation kinetics without the addition of seeds. TEM of endpoints from  $\alpha$ WT seeding (G) or  $\gamma$ I38 seeding (H), scale bar 200 nm.

**Table 5.5: Aggregation kinetics of seeding reactions with  $\alpha$ Syn WT and  $\gamma$ I38 seeds.** Lag times, elongation rates,  $t_{50}$  and % pellet for  $\gamma$ Syn variants. The rates of aggregation in each condition were measured in at least triplicate measurements. The errors show the standard deviation of the mean of the replicates. No aggregation after 42 h incubation for the seeded reactions is indicated by “-”. For errors in estimating % pellet via SDS PAGE subsequent to centrifugation (see Section 2.8.4) the values were rounded to the nearest 5%. \*Elongation rate in  $10^*$ [RFU/h].

	$\alpha$ WT, $\alpha$ -seeds, pH 4.5	$\alpha$ WT, $\alpha$ -seeds, pH 7.5	$\gamma$ WT, $\alpha$ -seeds, pH 4.5	$\gamma$ WT, $\alpha$ -seeds, pH 7.5	$\gamma$ I38, $\alpha$ -seeds, pH 4.5	$\gamma$ I38, $\alpha$ -seeds, pH 7.5
Lag time [h]	1.4 $\pm$ 0.9	0	31.7 $\pm$ 4.1	-	14.3 $\pm$ 1.0	-
Elongation rate*	0.1 $\pm$ 0.06	0.13 $\pm$ 0.02	0.06 $\pm$ 0.01	-	0.15 $\pm$ 0.02	-
$t_{50}$	8.6 $\pm$ 0.5	3.8 $\pm$ 0.8	34.5 $\pm$ 1.4	-	17.8 $\pm$ 1.3	-
% pellet	75	100	50	0	60	0
	$\alpha$ WT, $\gamma$ I38-seeds, pH 4.5	$\alpha$ WT, $\gamma$ I38-seeds, pH 7.5	$\gamma$ WT, $\gamma$ I38-seeds, pH 4.5	$\gamma$ WT, $\gamma$ I38-seeds, pH 7.5	$\gamma$ I38, $\gamma$ I38-seeds, pH 4.5	$\gamma$ I38, $\gamma$ I38-seeds, pH 7.5
Lag time [h]	5.6 $\pm$ 0.3	6.9 $\pm$ 0.4	0	0	0	0
Elongation rate*	0.3 $\pm$ 0.5	0.1 $\pm$ 0.04	0.1 $\pm$ 0.01	0.1 $\pm$ 0.00	0.2 $\pm$ 0.05	0.1 $\pm$ 0.03
$t_{50}$	10.3 $\pm$ 0.7	11.7 $\pm$ 1.5	3.1 $\pm$ 0.1	4.5 $\pm$ 0.9	1.3 $\pm$ 0.5	2.8 $\pm$ 0.6
% pellet	100	100	100	85	100	100

The hypothesis of distinct fibril architectures for  $\alpha$ Syn WT and  $\gamma$ Syn I38 fibrils was further validated by performing a limited proteolysis experiment with proteinase K (**Figure 5.15**). This enzyme cleaves non-specifically with a preference for peptide bonds next to the carboxyl group of aliphatic and aromatic residues. In this assay different digestion patterns correlate with distinct fibril morphologies due to variabilities in accessible residues and stable cores as described in the literature<sup>157,475</sup>. Incubation of  $\alpha$ Syn WT or  $\gamma$ I38 fibrils with proteinase K reveals substantial differences in digestion resistance.  $\alpha$ Syn WT fibrils grown at pH 7.5 present a stable core with a clear band at ~10 kDa, whereas fibrils of the same protein assembled at acidic pH show less proteinase K resistant morphologies with a diffuse pattern and a core at ~5 kDa after 5 min incubation. It should be noted that the enzyme reaction was performed at the same pH (pH 7.5), as fibrils were spun down and resuspended in the same buffer prior to the experiment as described in Section 2.7.10 (fibrils are stable when resuspended in the pH neutral buffer as shown in the control experiments with no enzyme present). Therefore, distinct band patterns are not a consequence of different proteinase K reactivities based on buffer changes. In contrast to  $\alpha$ Syn WT,  $\gamma$ I38 fibrils (grown at acidic pH) are less degraded after 5 minutes and a broad distribution of polypeptides with apparent molecular weight ranging from 13 kDa to 5 kDa are generated. Similar differences in limited proteolysis were observed for fibril versus ribbon  $\alpha$ Syn structures<sup>475</sup>. Although the interpretation of the data at this point is difficult (e.g. controls testing the equilibrium between fibril and monomer should be tested and ESI-MS should be used to identify the digestion products), the data provide further evidence that different fibril morphologies might be formed. It also highlights that the seeding assay would have better been done with fibrils grown at the same pH as for  $\alpha$ Syn WT as the pH difference seems to have a huge effect on fibril morphology. For a conclusive and more detailed analysis of the architectures of the fibrils formed in this work, higher resolution techniques such as cryoEM fibril structures would be required.

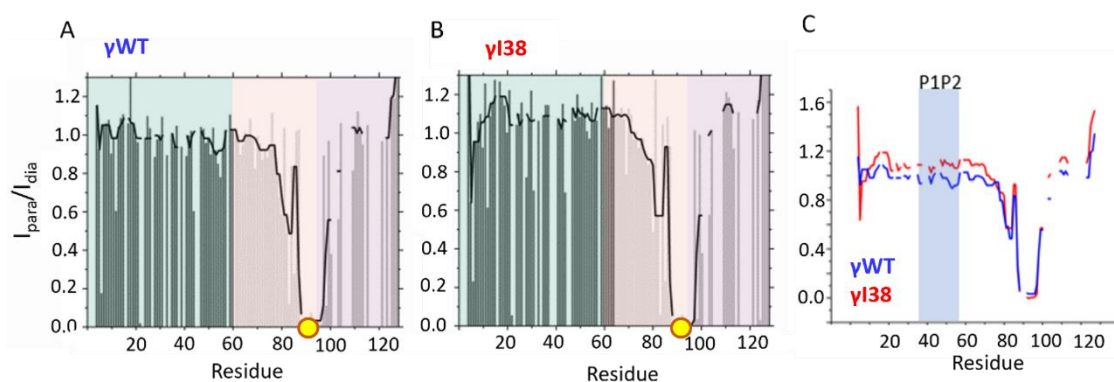


**Figure 5.15: Protease K treatment results in distinct digestion patterns for  $\alpha$ WT and  $\gamma$ 38I fibrils.**  $\alpha$ Syn WT seeds grown at pH 4.5 or 7.5 and  $\gamma$ 38I, pH 4.5 were incubated with proteinase K for 0, 1, and 5 min before boiling for 15 min to stop the reaction and loading on an SDS-PAGE gel. As a control (ctrl), variants were treated the same way as the other samples but with no enzyme addition.

## 5.10 PRE NMR to investigate conformational changes between $\gamma$ WT and $\gamma$ I38

In Chapter 3, PRE NMR was used as a technique to gain insight into the conformational space of  $\alpha$ Syn and to analyse how P1 affects self-assembly into higher order amyloid structures. It was shown that transient intra- and inter-molecular interactions with the P1 region of  $\alpha$ Syn drive fibril formation (see Section 3.7). As the identified ALS linked substitution  $\gamma$ M38I is located within P1, the hypothesis that nucleation of  $\gamma$ Syn might be affected by conformational changes of the monomeric protein accelerating aggregation arose.

To test if there are indeed intramolecular contacts that determine  $\gamma$ I38 aggregation which are not present in  $\gamma$ WT, a NMR PRE experiment was performed with  $^{15}\text{N}$   $\gamma$ WT and  $^{15}\text{N}$   $\gamma$ I38, both spin labelled with MTSL at position A90C (**Figure 5.16 A,B**). The experiment was performed at pH 4.5, where  $\gamma$ WT does not form amyloid whilst  $\gamma$ I38 rapidly forms fibrils (**Figure 5.12**). Interestingly, by contrast with the experiments performed on  $\alpha$ Syn,  $\gamma$ WT and  $\gamma$ I38 showed no significant changes of transient interactions despite their distinct aggregation propensity (**Figure 5.16 C**). Therefore, differences in aggregation cannot be explained on a monomeric level using this experimental setup. A more detailed analysis with different MTSL labelling positions and additional techniques such as cross linking MS and single molecule FRET assays will be needed to understand the self-assembly processes in more detail.



**Figure 5.16: NMR PRE data of intramolecular interactions of  $\gamma$ WT and  $\gamma$ I38 with MTSL spin label at position A90C.** Intramolecular PRE intensity ratios of amide protons (paramagnetic/diamagnetic) for  $\gamma$ WT (**A**) and  $\gamma$ I38 (**B**) are shown. Experiments were performed at 20 mM sodium acetate, pH 4.5, 200 mM NaCl, 15°C. Green, red and purple bars show intensity ratios for the N-terminal, NAC and C-terminal regions, respectively. Black line shows the rolling window (over five residues for easier comparison) of the PRE effects. MTSL label position is indicated with a yellow circle at residue A90C. (**C**) Overlay of rolling window from (**A**) and (**B**) ( $\gamma$ WT, blue and  $\gamma$ I38, red) for easier comparison of the data. The signal to noise level is >3.75 (dia), >2.70 (para), >1.99 (dia), and >2.34 (para) for  $\gamma$ WT and  $\gamma$ I38, respectively.

## 5.11 Discussion

In this Chapter a new, rare SNCG gene variant was revealed associated with ALS and encoding the aggregation-prone Met38Ile  $\gamma$ Syn variant. This is interesting in the context of Chapter 3 and 4 as this substitution is located in P1 and highlights that this NAC flanking region is not only involved in modulating the aggregation of  $\alpha$ Syn but also plays a critical role for  $\gamma$ Syn.

Although  $\gamma$ I38 cannot be put into a familial context, its significantly increased occurrence in ALS patients suggests its involvement in this disease. The potential association with neurodegeneration is further supported by the finding that this single residue substitution causes increased fibril growth *in vitro* and enhanced inclusion formation in cell culture. Gene variations causing single residue substitutions involved in disease have been seen before for other amyloidogenic proteins such as  $\alpha$ Syn in PD (e.g. A30P, A53T)<sup>145</sup> (see Section 1.4.2.1 for more detail), IAPP in type II diabetes (S20G)<sup>476</sup>, or  $\beta$ 2m in hereditary amyloidosis (D76N)<sup>477</sup>. Strikingly though, this is the first substitutional variant observed for  $\gamma$ Syn that significantly affects the protein aggregation properties with a potential pathogenic effect.

Usually, the NAC region (residues 61-95) has been considered as the highest aggregation-prone, fibril-core forming region in synucleins<sup>313</sup> and the work presented here shows that, indeed, this region plays a critical role in self-assembly of  $\alpha$ - and  $\gamma$ Syn. Interestingly, the Met38Ile substitution described in this Chapter is located outside the NAC region in the N-terminal flanking P1 region which has shown to be crucial for  $\alpha$ Syn function and toxicity. Remarkably, similar to  $\alpha$ Syn, substitution of only a single residue (in this case Met38 to Ile38) in P1, whilst not changing NAC at all, significantly affects the aggregation kinetics of the full-length protein. Therefore P1, and especially residue 38 are key regulators of fibril formation for both  $\alpha$ Syn and  $\gamma$ Syn. This highlights the important role of flanking regions of aggregation-prone regions to affect aggregation in amyloidogenic proteins (see Section 1.2.5 for more details on flanking regions in other proteins).

Despite the molecular details of why and how isoleucine contributes to a more aggregation-prone protein variant remaining unclear and further investigations are needed to fully understand the aggregation processes, experiments substituting other aliphatic residues (L,V,A) suggest beta-branched amino acid to play a crucial role in inducing aggregation. The formation of a  $\beta$ -hairpin in the P1 and P2 region of  $\alpha$ Syn has been shown to be crucial for fibrillation<sup>187,188,341</sup>. The increased aggregation kinetics for  $\gamma$ I38 and  $\gamma$ V38 might therefore be driven by introducing substitutions favouring  $\beta$ -sheet formation, and as a result the built-up of the  $\beta$ -hairpin.



The data demonstrate a striking sequence specificity to modulate aggregation and similar effects could be observed in other biologic systems such as the Hsp90 and Hsp70 proteins where a C-terminal Met to Ile substitution (**MEEVD to IEEVD**) drastically affects binding affinities to CHIP (Ubiquitin-protein ligase)<sup>478</sup>. Further, a sequence dependent aggregation could be observed for PrP<sup>Sc</sup> where in the pathogenic D178N substitution in combination with M129 is known to be involved in Fatal familial insomnia (FFI), whilst the V129 mutation causes familial Creutzfeldt-Jakob disease (CJD)<sup>479</sup>. So, although structurally very similar, different hydrophobic residues could result in distinct stable oligomeric or fibril states favouring or inhibiting aggregation.

The cross-seeding and limited proteolysis experiments with  $\alpha$ Syn and  $\gamma$ l38 fibrils suggest that distinct fibril morphologies result from aggregation of these protein variants. In general, aggregation is usually only marginally affected when cross-seeded with variants with only a single substitution<sup>458</sup>. This has been seen for the PD-related  $\alpha$ Syn variants H50Q, G51D, and A53T when seeded with fibrils formed from  $\alpha$ Syn WT<sup>172</sup> or non-disease related single point mutations located in the pre-NAC region<sup>458</sup>. This fits the observations proposed here, in which seeding  $\gamma$ l38 fibrils with  $\gamma$ WT monomer was efficient. However, literature data suggest that larger sequence variations often result in low cross-seeding efficiency as shown for human and hen lysozyme (60% sequence identity) or human  $\alpha$ -lactalbumin and hen lysozyme (36% sequence identity)<sup>480</sup>.  $\alpha$ Syn and  $\gamma$ l38 show a sequence identity of 54 %, so low cross-seeding would be expected. Surprisingly,  $\gamma$ l38 fibrils are able to cross-seed  $\alpha$ Syn WT monomer at acidic and neutral pH, whilst  $\alpha$ Syn fibrils can only cross-seed  $\gamma$ l38 monomer at acidic pH which might be driven by secondary nucleation processes. A possible explanation for the successful cross-seeding reaction in only one direction ( $\gamma$ l38 fibrils seed  $\alpha$ WT monomer, but  $\alpha$ WT fibrils do not seed  $\gamma$ l38 monomer at neutral pH) might be the fact that monomeric  $\gamma$ Syn is more stable than  $\alpha$ Syn (caused by a higher propensity to form  $\alpha$ -helix structures in NAC<sup>366</sup>) hindering the conformational transition possibly required to elongate  $\alpha$ Syn fibril ends. Similar observations were made for the sequence related amyloid- $\beta$  (A $\beta$ ) and islet amyloid polypeptide (IAPP). Whilst IAPP seeds can efficiently cross-seed A $\beta$  monomers, the reverse experiment resulted in no fibril growth<sup>126,481</sup>. A final conclusion on fibril morphology for  $\alpha$ Syn and  $\gamma$ l38 fibrils will only be measured, however, by high resolution approaches such as cryoEM.

In conclusion, it was demonstrated here that the Met38Ile amino acid substitution in human  $\gamma$ Syn increases its aggregation propensity *in vitro* and in cultured human cells and is linked with the development of ALS. These consequences of the sequence swap in  $\gamma$ Syn further demonstrate the key functional role of the N-terminal P1 region in controlling aggregation of synucleins with a high sequence specificity, and opens the

door to a new series of experiments using structural methods to elucidate in more detail how P1 exerts its remarkable effects.

## Chapter 6

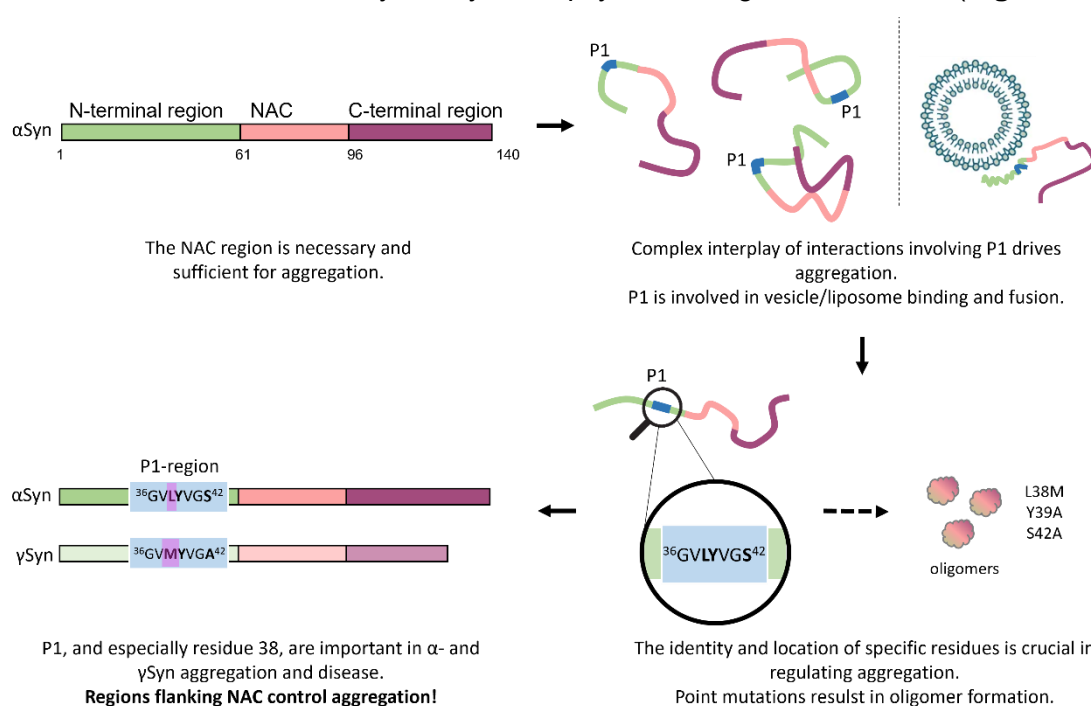
# Concluding Remarks and Future Perspective

Amyloid diseases such as Alzheimer's or Parkinson's disease present one of the greatest health issues in our modern society resulting in enormous social and economic burdens<sup>482,483</sup>. Today, there are more than 10 million people diagnosed with PD and over 55 million individuals have AD (or other forms of dementia) worldwide<sup>300,482</sup>. With an aging population, this problem is expected to increase over the next decades. The need for drug therapies is therefore continuously growing and understanding the molecular mechanisms that lead to self-assembly into toxic amyloidogenic oligomers and fibrils is essential.

More than 48 proteins are currently known to form amyloid fibrils associated with human disease<sup>63</sup>. Of these, 17 proteins are IDPs, or comprise intrinsically disordered regions including  $\alpha$ Syn<sup>47</sup>. Such sequences are of high risk to aggregate due to their intrinsic amyloid potential being exposed, unabridged by the protection of a native structure. This raises fundamental questions about how these initially dynamically disordered proteins self-assemble and gain order, and how this self-assembly process yields different cross- $\beta$  fibril structures from similar, or identical, precursor sequences<sup>441,484</sup>. While many protein-protein interactions are highly specific, as exemplified by antigen-antibody recognition, virus capsid assembly and many of the protein complexes that control essential cellular reactions (such as the ribosome and nucleosomes)<sup>485-487</sup>, others are more promiscuous, as shown by molecular chaperones that bind a diverse array of non-native protein molecules<sup>175,488</sup>. Whether the initiating stages of amyloid formation depend on specific interactions, or whether these interactions are more promiscuous for IDPs, however, remained unclear and was investigated in this thesis on the example of  $\alpha$ - and  $\gamma$ -synucleins.

A wide array of techniques, including *in silico* aggregation predictions, *in vitro* fibril formation assays, TEM and AFM imaging, experiments in the presence of liposomes, NMR (PRE and CSP), cell studies, and *in vivo* analysis using the model organism

*C. elegans* allowed the identification of an N-terminal motif (P1, residues 36-42) to be crucial to control fibril assembly in  $\alpha$ Syn and  $\gamma$ Syn on a single residue level (**Figure 6.1**).



**Figure 6.1. Schematic of the findings in this thesis that N-terminal flanking regions are important for aggregation.** From the initial idea that NAC controls amyloid formation (top, left), Chapter 3 identified P1 to play an important role in driving aggregation and function (top, right).

Chapter 4 further highlighted the role of individual residues in the P1 sequence that form oligomers (bottom, right) and Chapter 5 also considered the role of  $\gamma$ Syn for fibril formation (bottom, left), key residues (L38, Y39, S42) shown in bold font. The N-terminal region is coloured in green, NAC in red and the C-terminal region in purple. The P1 region is highlighted in blue.

In more detail, in Chapter 3, the P1 motif (together with N7 (residues 1-7) and P2 (residues 45-57)) was identified to play a vital role in aggregation and function of  $\alpha$ Syn. The results also highlighted that the NAC region alone does not determine the aggregation propensity of  $\alpha$ Syn, in fact, P1 forms crucial intra- and inter-molecular interactions allowing self-assembly into amyloid fibrils. However, deleting this key motif results in loss of function in *in vitro* assays, demonstrating a functional role as well as a key role in driving aggregation.

Chapter 4 further characterised the role of individual residues within P1 in aggregation. An alanine scan and sequence alignment with  $\gamma$ Syn identified three residues (L38, Y39, and S42A) to be the main contributors for the aggregation propensity of  $\alpha$ Syn. It was highlighted that the amino acid identity and position are important to regulate amyloid formation and early stages of fibrillation are highly specific.

Last but not least, Chapter 5 focusses on the role of  $\gamma$ Syn. It was found that P1 plays a critical role also in this paralogue since the M38I substitution (linked with ALS patients) can switch on aggregation of  $\gamma$ Syn *in vitro* and in cells. This points out that regions

flanking the highly aggregation-prone NAC region, are crucial probably throughout the whole synuclein family.

Although this work significantly contributed to understanding the molecular processes of amyloid formation in more detail and putting emphasis on the important role of the N-terminal region directly flanking NAC for synuclein aggregation, there still remain open questions. **Table 6.1** summarises the key questions with respect to this thesis, overall approaches to target them and detailed experimental setups that might help bringing forward the understanding of amyloid formation and toxicity that will be further discussed below. Importantly, now that the role of P1 is understood in more detail, it is possible to interpret the literature on a new level, allowing the design of target orientated experiments that highlight the potentials of P1 in developing anti-aggregation strategies.

**Table 6.1: Overview of open questions with respect to this thesis and future experiments.**

What is the molecular mechanism of aggregation and how do single residues contribute to it?	
Approaches	Experiments
Aggregation/ fibril formation	Deep mutational scanning
	Further characterisation of P2 by mutations (e.g. ThT assay, TEM imaging)
	Explore natural deletion variants in $\alpha$ Syn (e.g. exon3) <sup>438</sup>
	Analyse co-incubation of variants with $\alpha$ Syn WT ( <i>in vivo</i> and <i>in vitro</i> )
Structural analysis	Evaluate single residues in NAC or C-terminus that change the aggregation propensity
	Measure intra- and inter-molecular interactions by NMR PRE, MS, and cross-linking especially of variants (e.g. $\Delta$ P1, L38M) and at various pH values (e.g. pH 7.5)
	Solve high resolution fibril structures of $\alpha$ Syn and $\gamma$ Syn variants (using cryo-EM)
Can we regulate function independently from aggregation?	
Approaches	Experiments
Liposome interactions	Use extracted vesicles instead of DMPS liposomes
	Test physiological function of membrane remodelling for synuclein variants (incubation and TEM imaging, CD, NMR)
	Investigate conformational ensemble and intermediates in the presence of liposomes/vesicles
Aggregation vs function	Identifying molecules that retard fibril formation but allow physiological function (e.g. via small molecules)
How can we improve direct disease related research?	
Approaches	Experiments
Better <i>in vivo</i> systems	Express synuclein in neurons instead of muscle cells of <i>C. elegans</i>
	Experiments with vertebrates (e.g. zebrafish, mice) that naturally express synucleins
Define toxic state	Investigate if the non-aggregation variants are also less toxic (compare to $\alpha$ Syn A30P) (e.g. oligomer characterisation via cell assays in Chapter 4)
	Analyse chaperone response to synuclein variants (e.g. <i>in vivo</i> assays in worms)
Disease associated genetic polymorphism	Gain insight into familial history of $\gamma$ M38I variant
	Genome sequencing of larger/different population groups to identify further variants

## What is the molecular mechanism of $\alpha$ Syn aggregation and how do single residues contribute to it?

In this thesis it was found that not only the residue position, but also the side chain identity, play a crucial role in controlling  $\alpha$ Syn aggregation. To evaluate single residues that are key regulators, an alanine scan was performed. Interestingly, on the example of residue 38 in  $\alpha$ Syn and  $\gamma$ Syn, it was shown that different amino acids (although chemically very similar) can either speed up, slow down or not change the self-assembly of the full-length protein (Section 4.5 and Section 5.8). To better assess the molecular mechanism of aggregation, a more systematic approach would be required for instance by using deep mutational scanning (DMS) approaches. Newberry *et al.* have performed a DMS study in yeast screening for toxicity induced by membrane binding<sup>360,361</sup>. In these experiments, the readout is crucial, since selecting for (cyto-)toxicity, aggregation, or specific structural features may result in different outcomes. The Radford/Brockwell lab, for example, developed and optimised the tripartite  $\beta$ -lactamase assay, an assay screening for aggregation resistance in *E. coli*<sup>79,466,489</sup> which would offer an alternative approach to present studies.

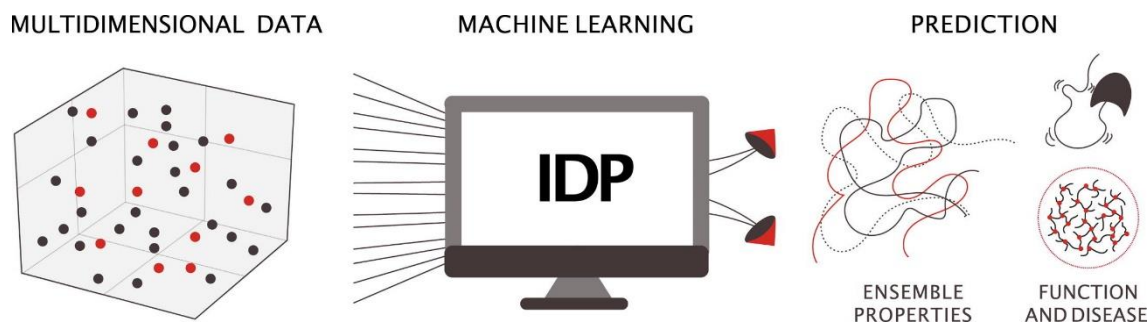
In the work presented in this thesis, P2 (residues 45-57) was identified in addition to P1 as synergising to control fibril assembly. This sequence acts in synergy with P1 (especially at acidic pH) but did not change the aggregation kinetics when deleted in isolation (**Figure 3.3**). Interestingly, this P2 region comprises the location of six of the twelve known familial PD mutations (E46K, H50Q, G51D, A53E/V/T). Although the focus was predominantly put on P1 here, we have learnt that P1 and P2 closely interact with each other. This was observed when adding the P1 peptide to  $\alpha$ Syn *in trans* (Section 4.3) and described in the literature in which P1 and P2 were shown to form a  $\beta$ -hairpin structure vital for self-assembly into amyloids<sup>187,188,415</sup> or SUMO1<sub>15-55</sub> binding to either P1 or P2, competing against the  $\beta$ -hairpin formation and inhibiting fibril formation<sup>439</sup>. Therefore, P2 most likely is important on a single residue level and familial PD mutations possibly interrupt crucial interactions with P1 (and the rest of the protein) resulting in conformational changes exposing the aggregation-prone NAC region. Further characterisation of P2 might therefore be beneficial to especially understand the role of disease linked substitutions.

Deletion variants of  $\alpha$ Syn, similar to the ones described in Chapter 3 ( $\Delta$ P1,  $\Delta$ P2), are also naturally observed in  $\alpha$ Syn. There is evidence that the regulation of the SNCA gene isoform expression is altered in the brains of patients with PD or other Lewy body disorders resulting in  $\alpha$ Syn variants lacking residues 41-54 (exon 3) and/or 103-130 (exon 5)<sup>438,490</sup>. The aggregation kinetics and membrane binding characteristics of these isoforms are only little analysed, but changes in fibril formation and membrane localisation have been reported<sup>438,491,492</sup>. Interestingly, as a consequence of different

expression levels, co-localisation of full-length  $\alpha$ Syn with the deletion variants is observed. It would be exciting to assess if adding these isoforms to  $\alpha$ Syn WT *in trans* acts as a protective mechanism or accelerates the formation of toxic amyloid species. Co-aggregation, in general, is an interesting factor to consider. Further investigation on incubating  $\alpha$ Syn WT together with the identified single point mutation or deletion variants aborting fibril formation might offer new insights into the inhibitory mechanism of the variants. These assays could be performed *in vitro* or *in vivo* (in neurons, *C. elegans*, *Drosophila* or more).

It should be considered that, although finding the N-terminal region to be vital for fibril formation in this thesis, this does not exclude the importance of other individual residues in NAC or C-terminal region. Many other studies have shown a role of residues in NAC or the C-terminus of  $\alpha$ Syn in aggregation. This includes reduced fibrillation for the substitution variants Y125A and Y133A<sup>440</sup>, as well as D121A and phosphorylated S129 especially in the context of calcium binding due to changes in charge<sup>460</sup>. Also, substitutions in NAC to make the sequence more similar to  $\beta$ Syn (residues 63-66 and 71-72 from VTNV, VT ( $\alpha$ Syn) to ASHI, FS ( $\beta$ Syn))<sup>468</sup> or single amino acid substitutions such as V70G/E, V74G/E, V76E/N or S87N<sup>313,365,409</sup> slow down amyloid formation. The changes in the aggregation kinetics are probably due to an altered conformational ensemble of the monomeric protein interrupting the complex interplay of contacts between the N-terminal region with NAC and C-terminal region<sup>189,460</sup>.

As mentioned before, the key to understand the molecular mechanism of self-assembly into amyloid structures is to recognise the ensemble of different conformations that the intrinsically disordered protein  $\alpha$ Syn can adopt under various buffer conditions and evaluate the free energy landscape<sup>493</sup>. For this challenging task, it is vital to combine various techniques including NMR (PRE) to determine transient interactions<sup>189,319,335,463</sup>, SAXS or DLS to investigate radius of gyration (Rg) or hydrodynamic radius (Rh)<sup>494</sup>, cross-linking and HDX to detect exposed surfaces and interactions<sup>460-462</sup>, single molecule FRET do evaluate distances between labels<sup>495</sup>, using small molecules to identify individual conformational states<sup>340,496</sup>, and computational modelling as well as machine learning to bring together the different observations and find the lowest energy states<sup>320,328,421</sup> (**Figure 6.2**). Further investigation and differentiation between  $\alpha$ Syn WT and the variants discussed in this thesis will allow a deeper understanding of the initial processes driving amyloid formation in particular in IDPs.



**Figure 6.2: Schematic of predicting conformational ensembles and function/disease aspects using machine learning based on multidimensional data.** Figure taken from Ref<sup>421</sup>.

Comprehension of the full mechanism of aggregation from monomer to mature fibrils also requires the analysis of oligomers and fibrils. This is especially important as recent studies highlighted the correlation between  $\alpha$ Syn monomer conformational state and fibril morphology<sup>462,463</sup>. Due to oligomeric species being less stable and diverse, little structural information of synuclein oligomers is available<sup>160,166,497</sup>. However, the accumulation of oligomeric species in the less aggregation-prone variants found in this work (**Figure 4.16** and **Figure 4.22**, L38M, Y39A, S42A) hints to differences in oligomer structures compared to  $\alpha$ Syn WT. This could result in higher or lower affinities to elongate into mature fibrils. There is more information on fibrillar structures (see **Figure 4.30** for a summary) highlighting how single residue substitutions or modifications can completely alter the fibril architecture (possibly by changing the monomeric conformational energy landscape<sup>462,463</sup>). Literature studies have shown that the fibril morphology might have an effect on toxicity and disease development<sup>117,278</sup>. The consequence of changes in the primary sequence on fibril stability, secondary nucleation, and morphology might therefore be crucial and should be further investigated using ideally cryoEM.

### Can we regulate function independently from aggregation?

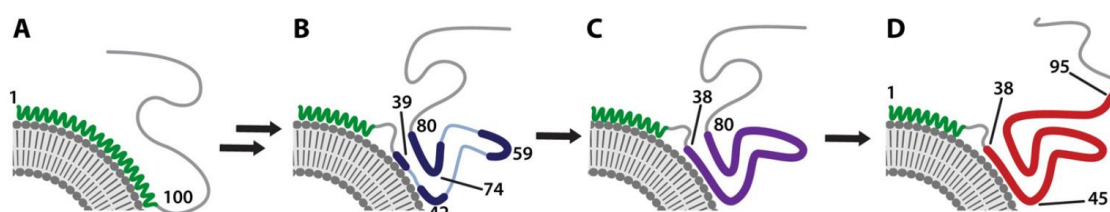
This study emphasised the opposed roles of the P1P2 region for function and toxicity (Section 3.4, Section 3.5 and Section 3.9). Targeting protein aggregation without negatively affecting the function of remodelling membranes of  $\alpha$ Syn<sup>272</sup> might therefore be a challenging task.

In a first step, the role for membrane binding of individual residues in the P1 and P2 region should be evaluated in more detail. Preliminary data on  $\alpha$ Syn L38M in Section 4.9 (**Figure 4.29**) and studies from Flagmeier *et al.*<sup>172</sup> on the lipid induced aggregation of the familial PD mutations demonstrate a large effect of single residue swaps for surface-induced fibril formation and remodelling of membranes. These experiments, however, are lacking some information and a systematic analysis of membrane binding and fusion via ThT fluorescence assays, CD, NMR, and TEM would be required. It should be further considered, that DMPS liposomes were used in these studies which might not be the best choice to mimic *in vivo* conditions. It is known that membranes in neurons are built



by a mixture of different lipids (33 % phosphatidylcholine, 30 % cholesterol, 8 % phosphatidylethanolamine and 29 % others (including for example sphingomyelin, phosphatidylserine, and phosphatidylinositol) in mice CNS brain regions)<sup>498</sup> and that presynaptic vesicles, that  $\alpha$ Syn binds to, have numerous proteins bound to their surface<sup>499</sup> possibly changing the interaction of  $\alpha$ Syn with these organelles. Other studies have consequently used lipid mixtures (DOPE, DOPS and DOPC in 5:3:2 molar ratios)<sup>186,272</sup> or vesicles extracted from rat brains<sup>273</sup>.

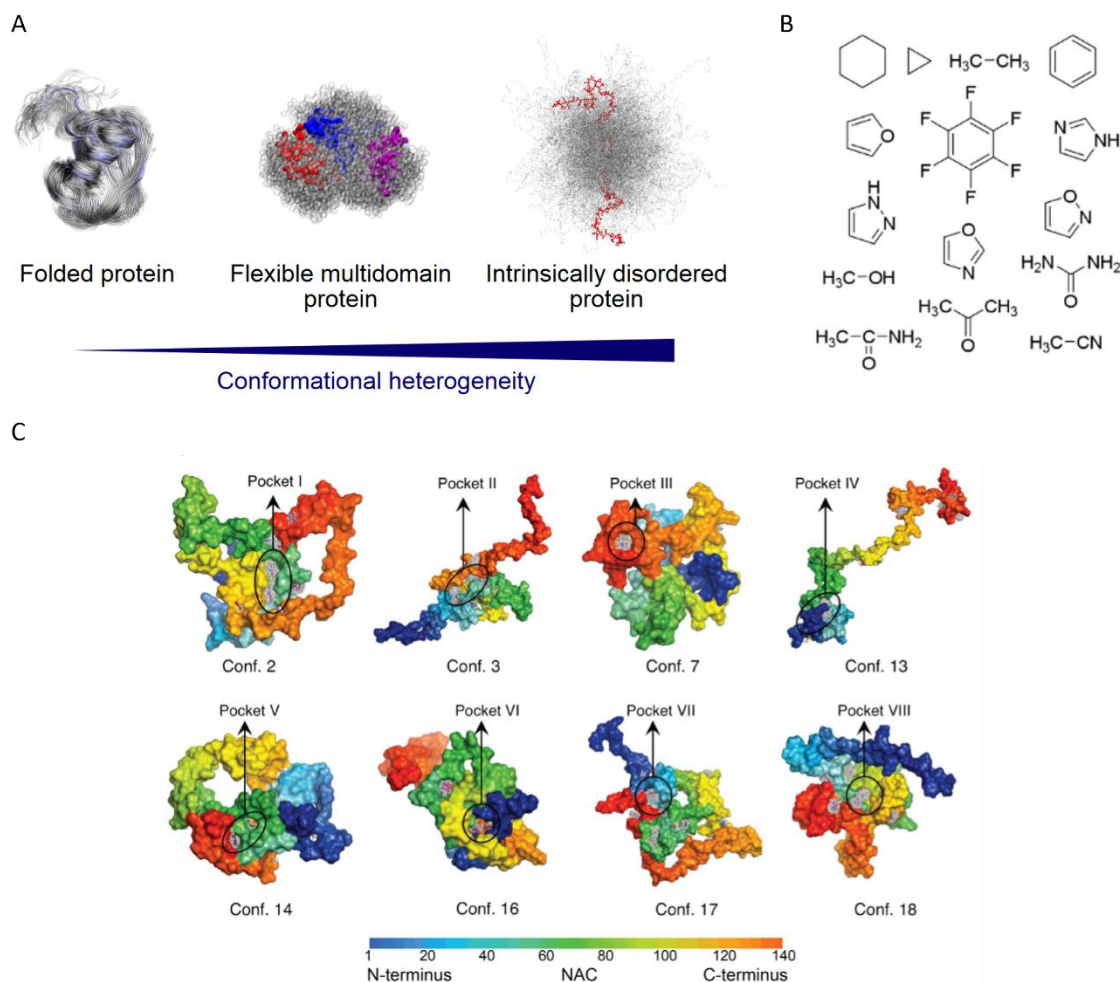
Interestingly, in some studies the membrane bound state is described as the toxic one<sup>361</sup> possibly by inducing surface catalysed aggregation or increased monomer concentrations at the membranes<sup>500</sup>. There is evidence for helical intermediates in amyloid formation of IDPs<sup>143</sup> and more precisely for  $\alpha$ Syn<sup>501</sup> in which residues 1-100 form helical structures when binding membranes, transitioning to 'β-sheet loops' for residues 57 to 61 and 73 to 80 whereas the N-terminal domain remains bound to the membrane and assists folding before eventually forming mature fibrils (**Figure 6.3**). Investigating these dynamic intermediates might help understanding the role of lipid binding for amyloid formation vs function.



**Figure 6.3: Schematic of  $\alpha$ Syn fibrillation on liposome surfaces.** Green:  $\alpha$ -helix, blue:  $\beta$ -sheet loops, purple:  $\beta$ -sheet of fibril-like conformation, red:  $\beta$ -sheet of mature fibril structure, grey: unstructured. Figure taken from Ref<sup>501</sup>.

Small molecules have been found to be promising molecules to interrupt oligomer and fibril formation of disease-associated amyloid proteins<sup>464,502</sup>. In contrast to natively folded proteins, the intrinsically disordered protein  $\alpha$ Syn can adopt a wide range of conformational states in its monomeric form (**Figure 6.4 A**). Therefore it can be challenging to recognise binding molecules due to the inherent structural heterogeneity and the absence of persistent structural elements<sup>503</sup>. Also, targeting fibrillation without affecting physiological function might be difficult. There have been promising approaches to overcome these challenges including an *in silico* screen based on structural information (NMR PRE) and computational modelling<sup>340,496</sup> (**Figure 6.4 B,C**) as well as biophysical approaches using surface plasmon resonance high-throughput screens<sup>504</sup>. Possible binding pockets to  $\alpha$ Syn considering the wide distribution of conformational states are shown in **Figure 6.4 C**. Targeting specific conformations allows the shift to explicit conformations enabling modulation of cellular malfunctions associated with  $\alpha$ Syn overexpression (e.g. restore vesicular dynamics, block cell-to-cell transmission)

independently of inhibitory effects on aggregation in solution as shown in the described studies<sup>496,504</sup>. So, there are promising experimental setups already developed for discovering small molecules binding to IDPs. Other approaches used for globular proteins might be possible to optimise for disordered proteins such as covalently tethered small molecule fragments to stabilise specific oligomers<sup>464</sup>.



**Figure 6.4: Conformational ensemble and small molecule binding to  $\alpha$ Syn.**

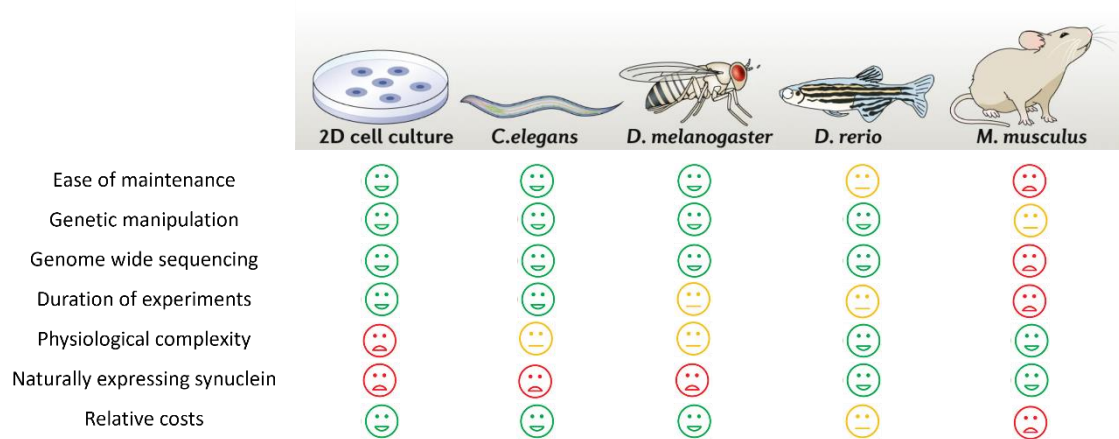
(A) Conformational heterogeneity in folded, flexible domain, and intrinsically disordered proteins on the example of ubiquitin, TIA-1 and  $\alpha$ Syn, respectively. Figure taken from Ref<sup>505</sup>. (B) Chemical structures of small molecules used as fragment probes in the *in silico* screen<sup>496</sup>. (C) 3D representation of the binding pockets identified within  $\alpha$ Syn conformations filled with clusters of small molecules resulting from the fragment probe mapping calculations. Residues are coloured according to the amino acid sequence shown below. Figure B and C taken from Ref<sup>496</sup>.

### How can we improve direct disease related research?

The main model organism used in this thesis was *C. elegans*. Van Ham *et al.* have shown that the expression of  $\alpha$ Syn::YFP in the body wall muscle cells of *C. elegans* provides a good model for PD in a living organism<sup>83</sup>. Although producing synuclein in the muscle cells has many advantages such as easily showing visible protein inclusions that increase with age, functional assays such as body bends, or facilitating biochemical analysis and fluorescent protein interaction studies, it should be considered that

expression in the muscle cells might not mimic the natural environment of a neuronal cell. Therefore, expression in dopaminergic neurons, motor neurons or pan-neuronal cells might be beneficial<sup>506,507</sup>. This can be challenging though, as  $\alpha$ Syn WT (which is used as a positive control in the work of this thesis) sometimes fails to induce a phenotypic effect in these cells and alternatives such as the more aggregation-prone A53T variant have to be used as comparison<sup>508</sup>.

Further, model organisms naturally expressing  $\alpha$ Syn might be beneficial. Whilst *C. elegans* facilitates the analysis of the effects of mutations in amyloid proteins by easily creating new strains and being able to assess effects in the absence of endogenous  $\alpha$ Syn, it also means that any findings may be an artefact of the expression of a non-native protein and loss-of-function aspects are difficult to evaluate. The usage of transgenic zebrafish or mice might therefore be needed but demands more time for experimental outputs at higher costs (Section 1.4.2.2) (**Figure 6.5**).



**Figure 6.5: Advantages and disadvantages of the most common model organisms used in biomedical research** (cell culture, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Danio rerio* and *Mus musculus*). Green smiley indicates suitable conditions, yellow smiley represents average conditions and red smiley stands for less/not suitable conditions. Figure adapted from Ref<sup>509</sup>.

Further, working in an *in vivo* system would allow to analyse the effect(s) of the chaperone response network in correlation with synuclein expression. This has not been done in this study but should be investigated, eventually, as it is known that  $\alpha$ Syn interacts with chaperones and can have a large effect on fibril assembly and disassembly<sup>175,225</sup>. The fact that  $\gamma$ Syn-I38::YFP expressing animals did not show a phenotypic effect (although aggregating *in vitro*) might be a result of a highly active proteostasis network, (over-) expressing molecules rescuing  $\gamma$ I38 from aggregation into toxic species<sup>510</sup>. Also, the disaggregation via Hsp70 and Hsp110 has been shown to result in harmless oligomers *in vitro*<sup>176</sup> but were demonstrated to result in toxic oligomeric species *in vivo*, highlighting the complexity of an *in vivo* system with various proteins/molecules and organelles being present<sup>511</sup>.

Another interesting aspect to take into account is the identity of the toxic species especially in an *in vivo* context. Whether oligomers or fibrils are the disease causing assembly is still unclear<sup>159</sup>. In this study, it was observed that the single point mutations L38M, Y39A, and S42A form oligomers within the experimental time of 110 h *in vitro* (**Figure 4.16** and **Figure 4.22**). Adding these oligomers to SH-SY5Y cells had a (partial) toxic effect (i.e. on cell morphology but not in metabolic assays), but expression in worms of the same variants protected from a phenotypic outcome. Similar seemingly controversial observations have been made for  $\alpha$ Syn A30P, a familial PD mutation that is less aggregation-prone than WT but forms more oligomers *in vitro*<sup>512</sup>, shows less toxicity in *C. elegans*<sup>473</sup>, but strikingly is nevertheless causing human disease<sup>171</sup>. Defining which factor(s) cause disease and how to identify the 'dangerous' species is therefore crucial.

Last, but not least, the genetic analysis in this study, but also in general, could be optimised. Identifying the ALS linked M38I substitution in Chapter 5 was important, but having familial background information would be beneficial. Further, it would be interesting to evaluate, if disease causing mutations (including  $\alpha$ Syn PD mutations) occur in people without developing the associated neurodegenerative disease. Interestingly, most PD linked substitutions (nine out of twelve) have been found in the N-terminal region, whereas the DLB associated variant E83Q is located in NAC<sup>294</sup>. Identifying more (familial) mutations, or protective mutations if they exist, might help to gain better insight in the role individual amino acid swaps play in disease progression.

Taken together, although not every question can be answered at this stage, this thesis has helped recognising that the N-terminal flanking region of NAC is an important regulator of  $\alpha$ Syn aggregation and function. Interestingly, over the time of this four year PhD thesis, there has been further growing evidence for the N-terminal region (in particular N7, P1, and P2) to be involved in toxicity and biological processes (see **Figure 3.25** and **Table 3.7**). This allows the assumption that also in other amyloidogenic or intrinsically disordered proteins, hotspots for aggregation might not solely be limited to the most aggregation-prone, core forming sequence. In contrast, other parts of the proteins might have large effects on the overall conformational ensemble and interactions with important binding partners that define its aggregation propensity. Targeting these flanking regions for drug discovery and development consequently offers a new approach and might provide promising results on the years ahead.

# References

1. Astbury, W.T. X-ray studies of protein structure. *Nature* **137**, 803–805 (1936).
2. Astbury, W. & Woods, H. The molecular weights of proteins. *Nature* **127**, 663–665 (1931).
3. Astbury, W.T. & Street, A. X-ray studies of the structure of hair, wool, and related fibres. *Philos. Trans. Royal Soc.* **230**, 75–101 (1931).
4. Pauling, L. & Corey, R.B. Atomic coordinates and structure factors for two helical configurations of polypeptide chains. *Proc. Natl. Acad. Sci. U. S. A.* **37**, 235 (1951).
5. Pauling, L. & Corey, R.B. Configurations of polypeptide chains with favored orientations around single bonds: two new pleated sheets. *Proc. Natl. Acad. Sci. U. S. A.* **37**, 729 (1951).
6. Watson, J.D. & Crick, F.H. Genetical implications of the structure of deoxyribonucleic acid. *Nature* **171**, 964–967 (1953).
7. Kendrew, J.C. *et al.* Structure of myoglobin: a three-dimensional Fourier synthesis at 2 Å resolution. *Nature* **185**, 422–427 (1960).
8. Bank, P.D. Protein data bank. *Nature New Biol* **233**, 223 (1971).
9. McCammon, J.A., Gelin, B.R. & Karplus, M. Dynamics of folded proteins. *Nature* **267**, 585–590 (1977).
10. Berman, H.M., Kleywegt, G.J., Nakamura, H. & Markley, J.L. The Protein Data Bank at 40: reflecting on the past to prepare for the future. *Structure* **20**, 391–396 (2012).
11. Bäuerlein, F.J. *et al.* *In situ* architecture and cellular interactions of PolyQ inclusions. *Cell* **171**, 179–187. e110 (2017).
12. Nishida, N., Ito, Y. & Shimada, I. In situ structural biology using in-cell NMR. *Biochim Biophys Acta Gen Subj* **1864**, 129364 (2020).
13. Fersht, A.R. From the first protein structures to our current knowledge of protein folding: delights and scepticisms. *Nat Rev Mol Cell Biol* **9**, 650–654 (2008).
14. Anfinsen, C.B., Haber, E., Sela, M. & White Jr, F. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.* **47**, 1309 (1961).
15. Levinthal, C. Are there pathways for protein folding? *Journal de chimie physique* **65**, 44–45 (1968).
16. Englander, S.W. & Mayne, L. The case for defined protein folding pathways. *Proc. Natl. Acad. Sci.* **114**, 8253–8258 (2017).

17. Wetlaufer, D.B. Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Natl. Acad. Sci.* **70**, 697-701 (1973).
18. Kim, P.S. & Baldwin, R.L. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* **51**, 459-489 (1982).
19. Baldwin, R.L. The nature of protein folding pathways: the classical versus the new view. *J. Biomol. NMR* **5**, 103-109 (1995).
20. Karplus, M. & Weaver, D.L. Diffusion-collision model for protein folding. *Biopolymers: Original Research on Biomolecules* **18**, 1421-1437 (1979).
21. Karplus, M. & Weaver, D.L. Protein folding dynamics: The diffusion-collision model and experimental data. *Protein Sci.* **3**, 650-668 (1994).
22. Tanford, C. Contribution of hydrophobic interactions to the stability of the globular conformation of proteins. *J. Am. Chem. Soc.* **84**, 4240-4247 (1962).
23. Baldwin, R.L. How does protein folding get started? *Trends Biochem. Sci.* **14**, 291-294 (1989).
24. Jackson, S.E. & Fersht, A.R. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428-10435 (1991).
25. Itzhaki, L.S., Otzen, D.E. & Fersht, A.R. The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J. Mol. Biol.* **254**, 260-288 (1995).
26. Fersht, A.R. Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. *Proc. Natl. Acad. Sci.* **92**, 10869-10873 (1995).
27. Radford, S.E. Protein folding: progress made and promises ahead. *Trends Biochem. Sci.* **25**, 611-618 (2000).
28. Fersht, A.R. Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.* **7**, 3-9 (1997).
29. Baldwin, A.J. *et al.* Metastability of native proteins and the phenomenon of amyloid formation. *J. Am. Chem. Soc.* **133**, 14160-14163 (2011).
30. Jackson, M.P. & Hewitt, E.W. Why are functional amyloids non-toxic in humans? *Biomolecules* **7**, 71 (2017).
31. Jahn, T.R. & Radford, S.E. The Yin and Yang of protein folding. *FEBS J* **272**, 5962-5970 (2005).
32. Agashe, V.R. & Hartl, F.-U. Roles of molecular chaperones in cytoplasmic protein folding. *Semin. Cell Dev. Biol.* **11**, 15-25 (2000).
33. van den Berg, B., Ellis, R.J. & Dobson, C.M. Effects of macromolecular crowding on protein folding and aggregation. *EMBO J* **18**, 6927-6933 (1999).

34. Jackson, M.P. & Hewitt, E.W. Cellular proteostasis: degradation of misfolded proteins by lysosomes. *Essays Biochem.* **60**, 173-180 (2016).
35. Gershenson, A. & Gierasch, L.M. Protein folding in the cell: challenges and progress. *Curr. Opin. Struct. Biol.* **21**, 32-41 (2011).
36. Olzmann, J.A. *et al.* Familial Parkinson's disease-associated L166P mutation disrupts DJ-1 protein folding and function. *J. Biol. Chem.* **279**, 8506-8515 (2004).
37. Kanapathipillai, M. Treating p53 mutant aggregation-associated cancer. *Cancers (Basel)* **10**, 154 (2018).
38. Laskey, R., Honda, B., Mills, A. & Finch, J. Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. *Nature* **275**, 416-420 (1978).
39. Ellis, R.J. Discovery of molecular chaperones. *Cell Stress Chaperones* **1**, 155 (1996).
40. Brehme, M. *et al.* A chaperome subnetwork safeguards proteostasis in aging and neurodegenerative disease. *Cell Rep.* **9**, 1135-1150 (2014).
41. 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).
42. Gallotta, I. *et al.* Extracellular proteostasis prevents aggregation during pathogenic attack. *Nature* **584**, 410-414 (2020).
43. Pohl, C. & Dikic, I. Cellular quality control by the ubiquitin-proteasome system and autophagy. *Science* **366**, 818-822 (2019).
44. Popovic, D., Vucic, D. & Dikic, I. Ubiquitination in disease pathogenesis and treatment. *Nat. Med.* **20**, 1242-1253 (2014).
45. Chiti, F. & Dobson, C.M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333-366 (2006).
46. Knowles, T.P., Vendruscolo, M. & Dobson, C.M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell. Biol.* **15**, 384-396 (2014).
47. Iadanza, M.G. *et al.* A new era for understanding amyloid structures and disease. *Nat. Rev. Mol. Cell Biol.* **12**, 755-773 (2018).
48. Perrett, S., Pham, C.L., Kwan, A.H. & Sunde, M. Functional amyloid: widespread in Nature, diverse in purpose. *Essays Biochem.* **56**, 207-219 (2014).
49. Shanmugam, N. *et al.* Microbial functional amyloids serve diverse purposes for structure, adhesion and defence. *Biophys. Rev.* **11**, 287-302 (2019).
50. Virchow, R. Weitere Mittheilungen über das Vorkommen der pflanzlichen Cellulose beim Menschen. *Archiv für pathologische Anatomie und Physiologie und für klinische Medicin* **6**, 268-271 (1854).

51. Sipe, J.D. & Cohen, A.S. History of the amyloid fibril. *J. Struct. Biol.* **130**, 88-98 (2000).
52. Friedreich, N. & Kekulé, A. Zur Amyloidfrage. *Archiv für pathologische Anatomie und Physiologie und für klinische Medizin* **16**, 50-65 (1859).
53. Alzheimer, A. Über eigenartige Krankheitsfälle des späteren Alters. *Zeitschrift für die gesamte Neurologie und Psychiatrie* **4**, 356 (1911).
54. Bennhold, H. Eine spezifische Amyloidfärbung mit Kongorot. *Meunch. Med. Wocheschr.* **69**, 1537-1538 (1922).
55. Divry, P. & Florkin, M. The optic properties of amyloid. *CR Seances Soc Biol Fil* **97**, 1808-1810 (1927).
56. Cohen, A.S. & Calkins, E. Electron microscopic observations on a fibrous component in amyloid of diverse origins. *Nature* **183**, 1202-1203 (1959).
57. Cohen, A.S., Frensdorff, A., Lamprecht, S. & Calkins, E. A study of the fine structure of the amyloid associated with familial Mediterranean fever. *Am. J. Clin. Pathol.* **41**, 567 (1962).
58. Geddes, A., Parker, K., Atkins, E. & Beighton, E. "Cross- $\beta$ " conformation in proteins. *J. Mol. Biol.* **32**, 343-358 (1968).
59. Eanes, E. & Glenner, G. X-ray diffraction studies on amyloid filaments. *J. Histochem. Cytochem.* **16**, 673-677 (1968).
60. Gallardo, R., Ranson, N.A. & Radford, S.E. Amyloid structures: much more than just a cross- $\beta$  fold. *Curr. Opin. Struct. Biol.* **60**, 7-16 (2020).
61. Glenner, G. *et al.* Amyloid fibril proteins: proof of homology with immunoglobulin light chains by sequence analyses. *Science* **172**, 1150-1151 (1971).
62. Benditt, E.P. & Eriksen, N. Chemical classes of amyloid substance. *Am. J. Clin. Pathol.* **65**, 231 (1971).
63. Benson, M.D. *et al.* Amyloid nomenclature 2020: update and recommendations by the International Society of Amyloidosis (ISA) nomenclature committee. *Amyloid* **27**, 217-222 (2020).
64. Balistreri, A., Goetzler, E. & Chapman, M. Functional amyloids are the rule rather than the exception in cellular biology. *Microorganisms* **8**, 1951 (2020).
65. Gauthier, S., Rosa-Neto, P., Morais, J. & Webster, C. World Alzheimer Report 2021: Journey through the diagnosis of dementia. *Alzheimer's Disease International* (2021).
66. Morris, K.L. & Serpell, L.C. X-ray fibre diffraction studies of amyloid fibrils. *Amyloid Proteins* 121-135 (Springer, 2012).
67. Sawaya, M.R. *et al.* Atomic structures of amyloid cross- $\beta$  spines reveal varied steric zippers. *Nature* **447**, 453-457 (2007).



68. Nelson, R. *et al.* Structure of the cross- $\beta$  spine of amyloid-like fibrils. *Nature* **435**, 773-778 (2005).
69. Qiang, W. *et al.* Structural variation in amyloid- $\beta$  fibrils from Alzheimer's disease clinical subtypes. *Nature* **541**, 217-221 (2017).
70. Tuttle, M.D. *et al.* Solid-state NMR structure of a pathogenic fibril of full-length human alpha-synuclein. *Nat. Struct. Mol. Biol.* **23**, 409-415 (2016).
71. Radamaker, L. *et al.* Cryo-EM structure of a light chain-derived amyloid fibril from a patient with systemic AL amyloidosis. *Nat. Commun.* **10**, 1-8 (2019).
72. Li, B. *et al.* Cryo-EM of full-length  $\alpha$ -synuclein reveals fibril polymorphs with a common structural kernel. *Nat. Commun.* **9**, 3609 (2018).
73. Fitzpatrick, A.W. *et al.* Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **547**, 185-190 (2017).
74. Binolfi, A., Theillet, F.-X. & Selenko, P. Bacterial in-cell NMR of human  $\alpha$ -synuclein: a disordered monomer by nature? (Portland Press Limited, 2012).
75. Trinkaus, V.A. *et al.* *In situ* architecture of neuronal  $\alpha$ -Synuclein inclusions. *Nat. Commun.* **12**, 1-10 (2021).
76. Shahmoradian, S.H. *et al.* Lewy pathology in Parkinson's disease consists of crowded organelles and lipid membranes. *Nat. Neurosci.* **22**, 1099-1109 (2019).
77. Meisl, G. *et al.* Molecular mechanisms of protein aggregation from global fitting of kinetic models. *Nat. Protoc.* **11**, 252-272 (2016).
78. Ma, B. & Nussinov, R. Simulations as analytical tools to understand protein aggregation and predict amyloid conformation. *Curr. Opin. Chem. Biol.* **10**, 445-452 (2006).
79. Ebo, J.S., Guthertz, N., Radford, S.E. & Brockwell, D.J. Using protein engineering to understand and modulate aggregation. *Curr. Opin. Struct. Biol.* **60**, 157-166 (2020).
80. Andersen, C.B. *et al.* Branching in amyloid fibril growth. *Biophys. J.* **96**, 1529-1536 (2009).
81. Pinotsi, D. *et al.* Direct observation of heterogeneous amyloid fibril growth kinetics via two-color super-resolution microscopy. *Nano letters* **14**, 339-345 (2014).
82. Jeibmann, A. & Paulus, W. *Drosophila melanogaster* as a model organism of brain diseases. *Int. J. Mol. Sci.* **10**, 407-440 (2009).
83. Van Ham, T.J. *et al.* *C. elegans* model identifies genetic modifiers of  $\alpha$ -synuclein inclusion formation during aging. *PLoS Genet.* **4**, e1000027 (2008).
84. Rencus-Lazar, S. *et al.* Yeast models for the study of amyloid-associated disorders and development of future therapy. *Front. Mol. Biosci.* **6**, 15 (2019).

85. Myers, A. & McGonigle, P. Overview of transgenic mouse models for Alzheimer's disease. *Curr. Protoc. Neurosci.* **89**, e81 (2019).
86. Rapezzi, C. *et al.* Efficacy of tafamidis in patients with hereditary and wild-type transthyretin amyloid cardiomyopathy: further analyses from ATTR-ACT. *Heart Fail.* **9**, 115-123 (2021).
87. Bowman, G.R., Voelz, V.A. & Pande, V.S. Taming the complexity of protein folding. *Curr. Opin. Struct. Biol.* **21**, 4-11 (2011).
88. Babu, M.M., van der Lee, R., de Groot, N.S. & Gsponer, J. Intrinsically disordered proteins: regulation and disease. *Curr. Opin. Struct. Biol.* **21**, 432-440 (2011).
89. Smolarczyk, T., Roterman-Konieczna, I. & Stapor, K. Protein secondary structure prediction: a review of progress and directions. *Curr. Bioinform.* **15**, 90-107 (2020).
90. Tayeb-Fligelman, E. *et al.* The cytotoxic *Staphylococcus aureus* PSM $\alpha$ 3 reveals a cross- $\alpha$  amyloid-like fibril. *Science* **355**, 831-833 (2017).
91. Blake, C. & Serpell, L. Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is a continuous  $\beta$ -sheet helix. *Structure* **4**, 989-998 (1996).
92. Fitzpatrick, A.W. *et al.* Atomic structure and hierarchical assembly of a cross- $\beta$  amyloid fibril. *Proc. Natl. Acad. Sci.* **110**, 5468-5473 (2013).
93. Cao, Q. *et al.* Cryo-EM structure and inhibitor design of human IAPP (amylin) fibrils. *Nat. Struct. Mol. Biol.* **27**, 653-659 (2020).
94. Creekmore, B.C., Chang, Y.-W. & Lee, E.B. The cryo-EM effect: structural biology of neurodegenerative disease aggregates. *J. Neuropathol. Exp. Neurol.* **80**, 514-529 (2021).
95. Taylor, J.P., Hardy, J. & Fischbeck, K.H. Toxic proteins in neurodegenerative disease. *Science* **296**, 1991-1995 (2002).
96. Guerrero-Ferreira, R., Kovacic, 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).
97. Scheres, S.H., Zhang, W., Falcon, B. & Goedert, M. Cryo-EM structures of tau filaments. *Curr. Opin. Struct. Biol.* **64**, 17-25 (2020).
98. Lövestam, S. *et al.* Seeded assembly in vitro does not replicate the structures of  $\alpha$ -synuclein filaments from multiple system atrophy. *FEBS Open bio* **11**, 999-1013 (2021).
99. Gremer, L. *et al.* Fibril structure of amyloid- $\beta$  (1-42) by cryo-electron microscopy. *Science* **358**, 116-119 (2017).

100. Iadanza, M.G. *et al.* The structure of a  $\beta$  2-microglobulin fibril suggests a molecular basis for its amyloid polymorphism. *Nat. Commun.* **9**, 1-10 (2018).
101. Glynn, C. *et al.* Cryo-EM structure of a human prion fibril with a hydrophobic, protease-resistant core. *Nat. Struct. Mol. Biol.* **27**, 417-423 (2020).
102. Falcon, B. *et al.* Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **568**, 420-423 (2019).
103. Hervas, R. *et al.* Cryo-EM structure of a neuronal functional amyloid implicated in memory persistence in *Drosophila*. *Science* **367**, 1230-1234 (2020).
104. Gallardo, R. *et al.* Fibril structures of diabetes-related amylin variants reveal a basis for surface-templated assembly. *Nat. Struct. Mol. Biol.* **27**, 1048-1056 (2020).
105. Cao, Q. *et al.* Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat. Struct. Mol. Biol.* **26**, 619-627 (2019).
106. Guerrero-Ferreira, R. *et al.* Cryo-EM structure of alpha-synuclein fibrils. *Elife* **7**, e36402 (2018).
107. Sun, Y. *et al.* The hereditary mutation G51D unlocks a distinct fibril strain transmissible to wild-type  $\alpha$ -synuclein. *Nat. Commun.* **12**, 1-10 (2021).
108. Sun, Y. *et al.* Cryo-EM structure of full-length  $\alpha$ -synuclein amyloid fibril with Parkinson's disease familial A53T mutation. *Cell Res.* **30**, 360-362 (2020).
109. Zhao, K. *et al.* Parkinson's disease-related phosphorylation at Tyr39 rearranges  $\alpha$ -synuclein amyloid fibril structure revealed by cryo-EM. *Proc. Natl. Acad. Sci.* **117**, 20305-20315 (2020).
110. Lövestam, S. *et al.* Assembly of recombinant tau into filaments identical to those of Alzheimer's disease and chronic traumatic encephalopathy. *bioRxiv* (2021).
111. Zeng, Y. *et al.* The structure and phase of tau: From monomer to amyloid filament. *Cell. Mol. Life Sci.* **78**, 1-14 (2020).
112. Tompa, P. Structural disorder in amyloid fibrils: its implication in dynamic interactions of proteins. *FEBS J.* **276**, 5406-5415 (2009).
113. Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J. & Serrano, L. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* **22**, 1302-1306 (2004).
114. Sormanni, P., Aprile, F.A. & Vendruscolo, M. The CamSol method of rational design of protein mutants with enhanced solubility. *J. Mol. Biol.* **427**, 478-490 (2015).
115. Li, Y. *et al.* Amyloid fibril structure of  $\alpha$ -synuclein determined by cryo-electron microscopy. *Cell Res.* **28**, 897 (2018).

116. Guerrero-Ferreira, R. *et al.* Two new polymorphic structures of alpha-synuclein solved by cryo-electron microscopy. *bioRxiv*, 654582 (2019).
117. Schweighauser, M. *et al.* Structures of  $\alpha$ -synuclein filaments from multiple system atrophy. *Nature* **585**, 464-469 (2020).
118. Lu, J.-X. *et al.* Molecular structure of  $\beta$ -amyloid fibrils in Alzheimer's disease brain tissue. *Cell* **154**, 1257-1268 (2013).
119. Kollmer, M. *et al.* Cryo-EM structure and polymorphism of A $\beta$  amyloid fibrils purified from Alzheimer's brain tissue. *Nat. Commun.* **10**, 1-8 (2019).
120. Petkova, A.T. *et al.* A structural model for Alzheimer's  $\beta$ -amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Acad. Sci.* **99**, 16742-16747 (2002).
121. Paravastu, A.K., Leapman, R.D., Yau, W.-M. & Tycko, R. Molecular structural basis for polymorphism in Alzheimer's  $\beta$ -amyloid fibrils. *Proc. Natl. Acad. Sci.* **105**, 18349-18354 (2008).
122. Colvin, M.T. *et al.* Atomic resolution structure of monomorphic A $\beta$ 42 amyloid fibrils. *J. Am. Chem. Soc.* **138**, 9663-9674 (2016).
123. Wälti, M.A. *et al.* Atomic-resolution structure of a disease-relevant A $\beta$  (1-42) amyloid fibril. *Proc. Natl. Acad. Sci.* **113**, e4976-e4984 (2016).
124. Falcon, B. *et al.* Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature* **561**, 137 (2018).
125. 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).
126. Röder, C. *et al.* Cryo-EM structure of islet amyloid polypeptide fibrils reveals similarities with amyloid- $\beta$  fibrils. *Nat. Struct. Mol. Biol.* **27**, 660-667 (2020).
127. Wang, L.-Q. *et al.* Cryo-EM structure of an amyloid fibril formed by full-length human prion protein. *Nat. Struct. Mol. Biol.* **27**, 598-602 (2020).
128. 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**, 1216 (2020).
129. 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).
130. Jahn, T.R. & Radford, S.E. Folding versus aggregation: polypeptide conformations on competing pathways. *Arch. Biochem. Biophys.* **469**, 100-117 (2008).
131. Narayan, P. *et al.* The extracellular chaperone clusterin sequesters oligomeric forms of the amyloid- $\beta$  1-40 peptide. *Nat. Struct. Mol. Biol.* **19**, 79-83 (2012).

132. Stefani, M. Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1739**, 5-25 (2004).
133. Buell, A.K. *et al.* Solution conditions determine the relative importance of nucleation and growth processes in  $\alpha$ -synuclein aggregation. *Proc. Natl. Acad. Sci.* **111**, 7671-7676 (2014).
134. Cremades, N. *et al.* Direct observation of the interconversion of normal and toxic forms of  $\alpha$ -synuclein. *Cell* **149**, 1048-1059 (2012).
135. Nettleton, E.J. *et al.* Characterization of the oligomeric states of insulin in self-assembly and amyloid fibril formation by mass spectrometry. *Biophys. J.* **79**, 1053-1065 (2000).
136. Karamanos, T.K. *et al.* A Population shift between sparsely populated folding intermediates determines amyloidogenicity. *J. Am. Chem. Soc.* **138**, 6271-6280 (2016).
137. Biancalana, M. & Koide, S. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochim. Biophys. Acta - Proteins Proteom.* **1804**, 1405-1412 (2010).
138. Cohen, S.I. *et al.* Nucleated polymerization with secondary pathways. I. Time evolution of the principal moments. *J. Chem. Phys.* **135**, 08B615 (2011).
139. Xue, W.-F. & Radford, S.E. An imaging and systems modeling approach to fibril breakage enables prediction of amyloid behavior. *Biophys. J.* **105**, 2811-2819 (2013).
140. Knowles, T.P. *et al.* An analytical solution to the kinetics of breakable filament assembly. *Science* **326**, 1533-1537 (2009).
141. Buell, A.K., Dobson, C.M. & Knowles, T.P. The physical chemistry of the amyloid phenomenon: thermodynamics and kinetics of filamentous protein aggregation. *Essays Biochem.* **56**, 11-39 (2014).
142. Zhu, M. *et al.* Surface-catalyzed amyloid fibril formation. *J. Biol. Chem.* **277**, 50914-50922 (2002).
143. Abedini, A. & Raleigh, D.P. A role for helical intermediates in amyloid formation by natively unfolded polypeptides? *Phys. Biol.* **6**, 015005 (2009).
144. Fändrich, M. Oligomeric intermediates in amyloid formation: structure determination and mechanisms of toxicity. *J. Mol. Biol.* **421**, 427-440 (2012).
145. Mehra, S., Sahay, S. & Maji, S.K.  $\alpha$ -Synuclein misfolding and aggregation: Implications in Parkinson's disease pathogenesis. *Biochim. Biophys. Acta - Proteins Proteom.* **1867**, 890-908 (2019).

146. Schaffert, L.-N. & Carter, W.G. Do post-translational modifications influence protein aggregation in neurodegenerative diseases: A systematic review. *Brain Sci.* **10**, 232 (2020).
147. Peng, C. *et al.* Cellular milieu imparts distinct pathological  $\alpha$ -synuclein strains in  $\alpha$ -synucleinopathies. *Nature* **557**, 558-563 (2018).
148. Horvath, I., Kumar, R. & Wittung-Stafshede, P. Macromolecular crowding modulates  $\alpha$ -synuclein amyloid fiber growth. *Biophys. J.* **120**, 3374-3381 (2021).
149. Stephens, A.D. & Schierle, G.S.K. The role of water in amyloid aggregation kinetics. *Curr. Opin. Struct. Biol.* **58**, 115-123 (2019).
150. Pramanik, S. & Ahmad, B. Exposure of aggregation-prone segments is the requirement for amyloid fibril formation. *Curr. Protein Pept. Sci.* **19**, 1024-1035 (2018).
151. Muchowski, P.J. Protein misfolding, amyloid formation, and neurodegeneration: a critical role for molecular chaperones? *Neuron* **35**, 9-12 (2002).
152. Hervás, R. & Oroz, J. Mechanistic insights into the role of molecular chaperones in protein misfolding diseases: from molecular recognition to amyloid disassembly. *Int. J. Mol. Sci.* **21**, 9186 (2020).
153. Wentink, A., Nussbaum-Krammer, C. & Bukau, B. Modulation of amyloid states by molecular chaperones. *Cold Spring Harb. Perspect. Biol.* **11**, a033969 (2019).
154. Grigolato, F. & Arosio, P. The role of surfaces on amyloid formation. *Biophys. Chem.* **270**, 106533 (2021).
155. Buratti, E. & Baralle, F.E. The multiple roles of TDP-43 in pre-mRNA processing and gene expression regulation. *RNA Biol.* **7**, 420-429 (2010).
156. Ayala, Y.M. *et al.* TDP-43 regulates its mRNA levels through a negative feedback loop. *The EMBO journal* **30**, 277-288 (2011).
157. de Oliveira, G.A., Cordeiro, Y., Silva, J.L. & Vieira, T.C. Liquid-liquid phase transitions and amyloid aggregation in proteins related to cancer and neurodegenerative diseases. *Adv. Protein Chem. Struct. Biol.* **118**, 289-331 (2019).
158. Bolognesi, B. *et al.* A concentration-dependent liquid phase separation can cause toxicity upon increased protein expression. *Cell Rep.* **16**, 222-231 (2016).
159. Alam, P., Bousset, L., Melki, R. & Otzen, D.E.  $\alpha$ -synuclein oligomers and fibrils: a spectrum of species, a spectrum of toxicities. *J. Neurochem.* **150**, 522-534 (2019).
160. Fusco, G. *et al.* Structural basis of membrane disruption and cellular toxicity by  $\alpha$ -synuclein oligomers. *Science* **358**, 1440-1443 (2017).

161. Serra-Batiste, M. *et al.* A $\beta$ 42 assembles into specific  $\beta$ -barrel pore-forming oligomers in membrane-mimicking environments. *Proc. Natl. Acad. Sci.* **113**, 10866-10871 (2016).
162. Birol, M., Kumar, S., Rhoades, E. & Miranker, A.D. Conformational switching within dynamic oligomers underpins toxic gain-of-function by diabetes-associated amyloid. *Nat. Commun.* **9**, 1-12 (2018).
163. Olzscha, H. *et al.* Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell* **144**, 67-78 (2011).
164. Park, S.-H. *et al.* PolyQ proteins interfere with nuclear degradation of cytosolic proteins by sequestering the Sis1p chaperone. *Cell* **154**, 134-145 (2013).
165. Du Yan, S. *et al.* Receptor-dependent cell stress and amyloid accumulation in systemic amyloidosis. *Nat. Med.* **6**, 643-651 (2000).
166. Chen, S.W. *et al.* Structural characterization of toxic oligomers that are kinetically trapped during alpha-synuclein fibril formation. *Proc. Natl. Acad. Sci. U. S. A.* **112**, e1994-2003 (2015).
167. Laurén, J. *et al.* Cellular prion protein mediates impairment of synaptic plasticity by amyloid- $\beta$  oligomers. *Nature* **457**, 1128-1132 (2009).
168. Abedini, A. *et al.* Time-resolved studies define the nature of toxic IAPP intermediates, providing insight for anti-amyloidosis therapeutics. *Elife* **5**, e12977 (2016).
169. Campioni, S. *et al.* A causative link between the structure of aberrant protein oligomers and their toxicity. *Nat. Chem. Biol.* **6**, 140-147 (2010).
170. Bieschke, J. *et al.* Small-molecule conversion of toxic oligomers to nontoxic  $\beta$ -sheet-rich amyloid fibrils. *Nat. Chem. Biol.* **8**, 93-101 (2012).
171. Krüger, R. *et al.* Ala30Pro mutation in the gene encoding  $\alpha$ -synuclein in Parkinson's disease. *Nat. Genet.* **18**, 106-108 (1998).
172. Flagmeier, P. *et al.* Mutations associated with familial Parkinson's disease alter the initiation and amplification steps of  $\alpha$ -synuclein aggregation. *Proc. Natl. Acad. Sci.* **113**, 10328-10333 (2016).
173. Benilova, I. *et al.* The Alzheimer disease protective mutation A2T modulates kinetic and thermodynamic properties of amyloid- $\beta$  (A $\beta$ ) aggregation. *J. Biol. Chem.* **289**, 30977-30989 (2014).
174. Johnson, B.S. *et al.* TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *J. Biol. Chem.* **284**, 20329-20339 (2009).
175. Burmann, B.M. *et al.* Regulation of  $\alpha$ -synuclein by chaperones in mammalian cells. *Nature* **577**, 127-132 (2020).

176. Gao, X. *et al.* Human Hsp70 disaggregase reverses Parkinson's-linked  $\alpha$ -synuclein amyloid fibrils. *Mol. Cell* **59**, 781-793 (2015).
177. Reumers, J., Maurer-Stroh, S., Schymkowitz, J. & Rousseau, F. Protein sequences encode safeguards against aggregation. *Hum. Mutat.* **30**, 431-437 (2009).
178. Chiti, F. *et al.* A partially structured species of  $\beta$ 2-microglobulin is significantly populated under physiological conditions and involved in fibrillogenesis. *J. Biol. Chem.* **276**, 46714-46721 (2001).
179. Eichner, T. *et al.* Conformational conversion during amyloid formation at atomic resolution. *Mol. Cell* **41**, 161-172 (2011).
180. Jahn, T.R., Parker, M.J., Homans, S.W. & Radford, S.E. Amyloid formation under physiological conditions proceeds via a native-like folding intermediate. *Nat. Struct. Mol. Biol.* **13**, 195-201 (2006).
181. Karamanos, T.K. *et al.* Structural mapping of oligomeric intermediates in an amyloid assembly pathway. *Elife* **8**, e46574 (2019).
182. Krüttner, S. *et al.* Drosophila CPEB Orb2A mediates memory independent of its RNA-binding domain. *Neuron* **76**, 383-395 (2012).
183. Kumari, P. *et al.* Structural insights into  $\alpha$ -synuclein monomer–fibril interactions. *Proc. Natl. Acad. Sci.* **118**, e2012171118 (2021).
184. Yang, X. *et al.* NMR unveils an N-terminal interaction interface on acetylated- $\alpha$ -synuclein monomers for recruitment to fibrils. *Proc. Natl. Acad. Sci.* **118**, e2017452118 (2021).
185. Cholak, E. *et al.* Avidity within the N-terminal anchor drives  $\alpha$ -synuclein membrane interaction and insertion. *FASEB J.* **34**, 7462-7482 (2020).
186. Fusco, G. *et al.* Direct observation of the three regions in  $\alpha$ -synuclein that determine its membrane-bound behaviour. *Nat. Commun.* **5**, 3827 (2014).
187. Mirecka, E.A. *et al.* Sequestration of a beta-hairpin for control of alpha-synuclein aggregation. *Angew. Chem. Int. Ed. Engl.* **53**, 4227-4230 (2014).
188. Yu, H., Han, W., Ma, W. & Schulten, K. Transient  $\beta$ -hairpin formation in  $\alpha$ -synuclein monomer revealed by coarse-grained molecular dynamics simulation. *J. Chem. Phys.* **143**, 12B623\_621 (2015).
189. Stephens, A.D., Zacharopoulou, M. & Kaminski Schierle, G.S. The cellular environment affects monomeric  $\alpha$ -synuclein structure. *Trends Biochem. Sci.* **5**, 453-466 (2018).
190. Hoyer, W., Cherny, D., Subramaniam, V. & Jovin, T.M. Impact of the acidic C-terminal region comprising amino acids 109-140 on  $\alpha$ -synuclein aggregation in vitro. *Biochemistry* **43**, 16233-16242 (2004).



191. Hong, D.-P., Xiong, W., Chang, J.-Y. & Jiang, C. The role of the C-terminus of human  $\alpha$ -synuclein: intra-disulfide bonds between the C-terminus and other regions stabilize non-fibrillar monomeric isomers. *FEBS Lett.* **585**, 561-566 (2011).
192. Falsone, S.F. *et al.* SERF protein is a direct modifier of amyloid fiber assembly. *Cell Rep.* **2**, 358-371 (2012).
193. Herrera, F.E. *et al.* Inhibition of  $\alpha$ -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).
194. Afroz, T. *et al.* Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. *Nat. Commun.* **8**, 1-15 (2017).
195. Shiina, Y., Arima, K., Tabunoki, H. & Satoh, J.-i. TDP-43 dimerizes in human cells in culture. *Cell. Mol. Neurobiol.* **30**, 641-652 (2010).
196. Zhang, Y.-J. *et al.* The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Hum. Mol. Genet.* **22**, 3112-3122 (2013).
197. Lukavsky, P.J. *et al.* Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat. Struct. Mol. Biol.* **20**, 1443-1449 (2013).
198. Kuo, P.-H. *et al.* The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG-and TG-rich nucleic acids. *Nucleic Acids Res.* **42**, 4712-4722 (2014).
199. Huang, Y.-C. *et al.* Inhibition of TDP-43 aggregation by nucleic acid binding. *PLoS One* **8**, e64002 (2013).
200. Conicella, A.E., Zerze, G.H., Mittal, J. & Fawzi, N.L. ALS mutations disrupt phase separation mediated by  $\alpha$ -helical structure in the TDP-43 low-complexity C-terminal domain. *Structure* **24**, 1537-1549 (2016).
201. Sun, Y. & Chakrabarty, A. Phase to phase with TDP-43. *Biochemistry* **56**, 809-823 (2017).
202. Li, H.-R. *et al.* TAR DNA-binding protein 43 (TDP-43) liquid-liquid phase separation is mediated by just a few aromatic residues. *J. Biol. Chem.* **293**, 6090-6098 (2018).
203. Sastre, M. *et al.* Binding of cystatin C to Alzheimer's amyloid  $\beta$  inhibits in vitro amyloid fibril formation. *Neurobiol. Aging* **25**, 1033-1043 (2004).
204. Kannaian, B. *et al.* Abundant neuroprotective chaperone Lipocalin-type prostaglandin D synthase (L-PGDS) disassembles the Amyloid- $\beta$  fibrils. *Sci. Rep.* **9**, 1-17 (2019).

205. Brandt, R., Léger, J. & Lee, G. Interaction of tau with the neural plasma membrane mediated by tau's amino-terminal projection domain. *J. Cell. Biol.* **131**, 1327-1340 (1995).
206. Rosenberg, K.J. *et al.* Complementary dimerization of microtubule-associated tau protein: Implications for microtubule bundling and tau-mediated pathogenesis. *Proc. Natl. Acad. Sci.* **105**, 7445-7450 (2008).
207. Boyko, S. *et al.* Liquid–liquid phase separation of tau protein: the crucial role of electrostatic interactions. *J. Biol. Chem.* **294**, 11054-11059 (2019).
208. He, H.J. *et al.* The proline-rich domain of tau plays a role in interactions with actin. *BMC Cell Biol.* **10**, 1-12 (2009).
209. Mok, S.-A. *et al.* Mapping interactions with the chaperone network reveals factors that protect against tau aggregation. *Nat. Struct. Mol. Biol.* **25**, 384-393 (2018).
210. McKibben, K.M. & Rhoades, E. Independent tubulin binding and polymerization by the proline-rich region of Tau is regulated by Tau's N-terminal domain. *J. Biol. Chem.* **294**, 19381-19394 (2019).
211. Chen, D. *et al.* Tau local structure shields an amyloid-forming motif and controls aggregation propensity. *Nat. Commun.* **10**, 1-14 (2019).
212. Liu, F. *et al.* Site-specific effects of tau phosphorylation on its microtubule assembly activity and self-aggregation. *Eur. J. Neurosci.* **26**, 3429-3436 (2007).
213. Sibille, N. *et al.* Structural impact of heparin binding to full-length Tau as studied by NMR spectroscopy. *Biochemistry* **45**, 12560-12572 (2006).
214. Esposito, G. *et al.* Removal of the N-terminal hexapeptide from human  $\beta$ 2-microglobulin facilitates protein aggregation and fibril formation. *Protein Sci.* **9**, 831-845 (2000).
215. Jones, S., Smith, D.P. & Radford, S.E. Role of the N and C-terminal strands of beta 2-microglobulin in amyloid formation at neutral pH. *J. Mol. Biol.* **330**, 935-941 (2003).
216. Brender, J.R. *et al.* Amyloid fiber formation and membrane disruption are separate processes localized in two distinct regions of IAPP, the type-2-diabetes-related peptide. *J. Am. Chem. Soc.* **130**, 6424-6429 (2008).
217. Pytowski, L. *et al.* Liquid–liquid phase separation of type II diabetes-associated IAPP initiates hydrogelation and aggregation. *Proc. Natl. Acad. Sci.* **117**, 12050-12061 (2020).
218. Han, J. *et al.* Study on interaction between microtubule associated protein tau and prion protein. *Science in China Series C: Life Sciences* **49**, 473-479 (2006).
219. Aulić, S. *et al.*  $\alpha$ -synuclein amyloids hijack prion protein to gain cell entry, facilitate cell-to-cell spreading and block prion replication. *Sci. Rep.* **7**, 1-12 (2017).

220. Wang, F. *et al.* Role of the highly conserved middle region of prion protein (PrP) in PrP– lipid interaction. *Biochemistry* **49**, 8169-8176 (2010).
221. Goldfarb, L.G. *et al.* Transmissible familial Creutzfeldt-Jakob disease associated with five, seven, and eight extra octapeptide coding repeats in the PRNP gene. *Proc. Natl. Acad. Sci.* **88**, 10926-10930 (1991).
222. Wegmann, S., Medalsy, I.D., Mandelkow, E. & Müller, D.J. The fuzzy coat of pathological human Tau fibrils is a two-layered polyelectrolyte brush. *Proc. Natl. Acad. Sci.* **110**, e313-321 (2013).
223. Sillen, A. *et al.* Regions of tau implicated in the paired helical fragment core as defined by NMR. *Chembiochem* **6**, 1849-1856 (2005).
224. Luo, J., Wärmländer, S.K., Gräslund, A. & Abrahams, J.P. Cross-interactions between the Alzheimer disease amyloid- $\beta$  peptide and other amyloid proteins: a further aspect of the amyloid cascade hypothesis. *J. Biol. Chem.* **291**, 16485-16493 (2016).
225. Wentink, A.S. *et al.* Molecular dissection of amyloid disaggregation by human HSP70. *Nature* **587**, 483-488 (2020).
226. Khan, M.R. *et al.* Amyloidogenic oligomerization transforms Drosophila Orb2 from a translation repressor to an activator. *Cell* **163**, 1468-1483 (2015).
227. Ward, J.J. *et al.* Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **337**, 635-645 (2004).
228. Dunker, A.K. *et al.* Intrinsic protein disorder in complete genomes. *Genome informatics* **11**, 161-171 (2000).
229. Dunker, A.K. *et al.* What's in a name? Why these proteins are intrinsically disordered: Why these proteins are intrinsically disordered. *Intrinsically Disord. Proteins* **1**, e24157 (2013).
230. Uversky, V. Intrinsically disordered proteins from A to Z. . *Int J Biochem Cell Biol.* **43**, 1090-1103 (2011).
231. Williams, R. *et al.* The protein non-folding problem: amino acid determinants of intrinsic order and disorder. *Biocomputing* **2001**, 89-100 (2000).
232. Bah, A. *et al.* Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519**, 106-109 (2015).
233. Mao, A.H. *et al.* Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc. Natl. Acad. Sci.* **107**, 8183-8188 (2010).
234. Coletta, A. *et al.* Low-complexity regions within protein sequences have position-dependent roles. *BMC Syst. Biol.* **4**, 43 (2010).
235. Buljan, M. *et al.* Alternative splicing of intrinsically disordered regions and rewiring of protein interactions. *Curr. Opin. Struct. Biol.* **23**, 443-450 (2013).

236. Babu, M.M. The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochem. Soc. Trans.* **44**, 1185-1200 (2016).
237. Wang, J., Cao, Z., Zhao, L. & Li, S. Novel strategies for drug discovery based on intrinsically disordered proteins (IDPs). *Int. J. Mol. Sci.* **12**, 3205-3219 (2011).
238. Rogers, J.M. *et al.* Interplay between partner and ligand facilitates the folding and binding of an intrinsically disordered protein. *Proc. Natl. Acad. Sci.* **111**, 15420-15425 (2014).
239. Kragelund, B.B. *et al.* DSS1/Sem1, a multifunctional and intrinsically disordered protein. *Trends Biochem. Sci.* **41**, 446-459 (2016).
240. Davey, N.E. The functional importance of structure in unstructured protein regions. *Curr. Opin. Struct. Biol.* **56**, 155-163 (2019).
241. Zhou, H.-X. Intrinsic disorder: signaling via highly specific but short-lived association. *Trends Biochem. Sci.* **37**, 43-48 (2012).
242. Theillet, F.-X. *et al.* Structural disorder of monomeric  $\alpha$ -synuclein persists in mammalian cells. *Nature* **530**, 45 (2016).
243. Davidson, W.S., Jonas, A., Clayton, D.F. & George, J.M. Stabilization of  $\alpha$ -synuclein secondary structure upon binding to synthetic membranes. *J. Biol. Chem.* **273**, 9443-9449 (1998).
244. Bartels, T. *et al.* The N-terminus of the intrinsically disordered protein  $\alpha$ -synuclein triggers membrane binding and helix folding. *Biophys. J.* **99**, 2116-2124 (2010).
245. Hardy, J. & Gwinn-Hardy, K. Genetic classification of primary neurodegenerative disease. *Science* **282**, 1075-1079 (1998).
246. Maroteaux, L., Campanelli, J.T. & Scheller, R.H. Synuclein: a neuron-specific protein localized to the nucleus and presynaptic nerve terminal. *J. Neurosci.* **8**, 2804-2815 (1988).
247. Polymeropoulos, M.H. *et al.* Mutation in the  $\alpha$ -synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045-2047 (1997).
248. Fevga, C. *et al.* A new alpha-synuclein missense variant (Thr72Met) in two Turkish families with Parkinson's disease. *Parkinsonism Relat. Disord.* **89**, 63-72 (2021).
249. Spillantini, M.G. *et al.*  $\alpha$ -Synuclein in Lewy bodies. *Nature* **388**, 839 (1997).
250. Villar-Piqué, A., Lopes da Fonseca, T. & Outeiro, T.F. Structure, function and toxicity of alpha-synuclein: the Bermuda triangle in synucleinopathies. *J. Neurochem.* **139**, 240-255 (2016).

251. Lashuel, H.A., Overk, C.R., Oueslati, A. & Masliah, E. The many faces of  $\alpha$ -synuclein: from structure and toxicity to therapeutic target. *Nat. Rev. Neurosci.* **14**, 38 (2013).
252. Weinreb, P.H. *et al.* NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry* **35**, 13709-13715 (1996).
253. Iwai, A. *et al.* The precursor protein of non A $\beta$  component of Alzheimer's Disease amyloid is a presynaptic protein of the central-nervous-system. *Neuron* **14**, 467-475 (1995).
254. Wilhelm, B.G. *et al.* Composition of isolated synaptic boutons reveals the amounts of vesicle trafficking proteins. *Science* **344**, 1023-1028 (2014).
255. Abeliovich, A. *et al.* Mice lacking  $\alpha$ -synuclein display functional deficits in the nigrostriatal dopamine system. *Neuron* **25**, 239-252 (2000).
256. Cabin, D.E. *et al.* Synaptic vesicle depletion correlates with attenuated synaptic responses to prolonged repetitive stimulation in mice lacking  $\alpha$ -synuclein. *J. Neurosci.* **22**, 8797-8807 (2002).
257. Greten-Harrison, B. *et al.*  $\alpha\beta\gamma$ -Synuclein triple knockout mice reveal age-dependent neuronal dysfunction. *Proc. Natl. Acad. Sci.* **107**, 19573-19578 (2010).
258. Wu, N. *et al.* Alpha-synuclein overexpression in mice alters synaptic communication in the corticostriatal pathway. *J. Neurosci. Res.* **88**, 1764-1776 (2010).
259. Nemani, V.M. *et al.* Increased expression of  $\alpha$ -synuclein reduces neurotransmitter release by inhibiting synaptic vesicle reclustering after endocytosis. *Neuron* **65**, 66-79 (2010).
260. Scott, D.A. *et al.* A pathologic cascade leading to synaptic dysfunction in  $\alpha$ -synuclein-induced neurodegeneration. *J. Neurosci.* **30**, 8083-8095 (2010).
261. Tompa, P., Dosztányi, Z. & Simon, I. Prevalent Structural Disorder in *E. coli* and *S. cerevisiae* proteomes. *J. Proteome Res.* **5**, 1996-2000 (2006).
262. Jin, J. *et al.* Identification of novel proteins associated with both  $\alpha$ -synuclein and DJ-1. *Mol. Cell. Proteomics* **6**, 845-859 (2007).
263. Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, 362-368 (2016).
264. Oughtred, R. *et al.* The BioGRID interaction database: 2019 update. *Nucleic Acids Res.* **47**, 529-541 (2019).

265. Bernal-Conde, L.D. *et al.* Alpha-synuclein physiology and pathology: a perspective on cellular structures and organelles. *Front. Neurosci.* **13**, 1399 (2020).
266. Brahmachari, S. *et al.* Activation of tyrosine kinase c-Abl contributes to  $\alpha$ -synuclein-induced neurodegeneration. *J. Clin. Investig.* **126**, 2970-2988 (2016).
267. McFarland, M.A., Ellis, C.E., Markey, S.P. & Nussbaum, R.L. Proteomics analysis identifies phosphorylation-dependent  $\alpha$ -synuclein protein interactions. *Mol. Cell. Proteomics* **7**, 2123-2137 (2008).
268. McLean, P.J., Kawamata, H., Ribich, S. & Hyman, B.T. Membrane association and protein conformation of  $\alpha$ -synuclein in intact neurons effect of Parkinson's Disease-linked mutations. *J. Biol. Chem.* **275**, 8812-8816 (2000).
269. Lee, H.-J., Choi, C. & Lee, S.-J. Membrane-bound  $\alpha$ -synuclein has a high aggregation propensity and the ability to seed the aggregation of the cytosolic form. *J. Biol. Chem.* **277**, 671-678 (2002).
270. Kahle, P.J. *et al.* Subcellular localization of wild-type and Parkinson's disease-associated mutant  $\alpha$ -synuclein in human and transgenic mouse brain. *J. Neurosci.* **20**, 6365-6373 (2000).
271. Zhang, L. *et al.* Semi-quantitative analysis of  $\alpha$ -synuclein in subcellular pools of rat brain neurons: an immunogold electron microscopic study using a C-terminal specific monoclonal antibody. *Brain Res.* **1244**, 40-52 (2008).
272. Fusco, G. *et al.* Structural basis of synaptic vesicle assembly promoted by alpha-synuclein. *Nat. Commun.* **7**, 1-12 (2016).
273. Lautenschläger, J. *et al.* C-terminal calcium binding of  $\alpha$ -synuclein modulates synaptic vesicle interaction. *Nat. Commun.* **9**, 712 (2018).
274. Diao, J. *et al.* Native  $\alpha$ -synuclein induces clustering of synaptic-vesicle mimics via binding to phospholipids and synaptobrevin-2/VAMP2. *Elife* **2**, e00592 (2013).
275. Burré, J. *et al.*  $\alpha$ -Synuclein promotes SNARE-complex assembly in vivo and in vitro. *Science* **329**, 1663-1667 (2010).
276. Goedert, M., Jakes, R. & Spillantini, M.G. The synucleinopathies: twenty years on. *J. Parkinsons Dis.* **7**, 51-69 (2017).
277. Hardy, J. Amyloid, the presenilins and Alzheimer's disease. *Trends Neurosci.* **20**, 154-159 (1997).
278. Strohäker, T. *et al.* Structural heterogeneity of  $\alpha$ -synuclein fibrils amplified from patient brain extracts. *Nat. Commun.* **10**, 1-12 (2019).
279. Wong, Y.C. & Krainc, D.  $\alpha$ -synuclein toxicity in neurodegeneration: mechanism and therapeutic strategies. *Nat. Med.* **23**, 1-13 (2017).

280. Lesage, S. & Brice, A. Parkinson's disease: from monogenic forms to genetic susceptibility factors. *Hum. Mol. Genet.* **18**, 48-59 (2009).
281. Liu, H. *et al.* A Novel SNCA A30G mutation causes familial Parkinson's Disease. *Mov. Disord.* **36**, 1624-1633 (2021).
282. Zarranz, J.J. *et al.* The new mutation, E46K, of  $\alpha$ -synuclein causes parkinson and Lewy body dementia. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* **55**, 164-173 (2004).
283. Proukakis, C. *et al.* A novel  $\alpha$ -synuclein missense mutation in Parkinson disease. *Neurology* **80**, 1062-1064 (2013).
284. Kiely, A.P. *et al.*  $\alpha$ -Synucleinopathy associated with G51D SNCA mutation: a link between Parkinson's disease and multiple system atrophy? *Acta Neuropathol.* **125**, 753-769 (2013).
285. Pasanen, P. *et al.* A novel  $\alpha$ -synuclein mutation A53E associated with atypical multiple system atrophy and Parkinson's disease-type pathology. *Neurobiol. Aging* **35**, e2181-2180 (2014).
286. Yoshino, H. *et al.* Homozygous alpha-synuclein A53V in familial Parkinson's disease. *Neurobiol. Aging* **57**, e247-248 (2017).
287. Zhao, Y. *et al.* The role of genetics in Parkinson's disease: a large cohort study in Chinese mainland population. *Brain* **143**, 2220-2234 (2020).
288. Chen, Y. *et al.* Evaluating the role of SNCA, LRRK2, and GBA in Chinese patients with early-onset Parkinson's Disease. *Mov. Disord.* **35**, 2046-2055 (2020).
289. Cali, F. *et al.* Interpreting genetic variants: hints from a family cluster of Parkinson's disease. *J. Parkinsons Dis.* **9**, 203-206 (2019).
290. Zheng, R. *et al.* Analysis of rare variants of autosomal-dominant genes in a Chinese population with sporadic Parkinson's disease. *Molecular genetics & genomic medicine* **8**, e1449 (2020).
291. Hoffman-Zacharska, D. *et al.* Novel A18T and pA29S substitutions in  $\alpha$ -synuclein may be associated with sporadic Parkinson's disease. *Parkinsonism Relat. Disord.* **19**, 1057-1060 (2013).
292. Kumar, S. *et al.* Role of sporadic Parkinson disease associated mutations A18T and A29S in enhanced  $\alpha$ -synuclein fibrillation and cytotoxicity. *ACS Chem. Neurosci.* **9**, 230-240 (2018).
293. Youn, J. *et al.* Genetic variants of PARK genes in Korean patients with early-onset Parkinson's disease. *Neurobiol. Aging* **75**, e229-224 (2019).
294. Kapasi, A. *et al.* A novel SNCA E83Q mutation in a case of dementia with Lewy bodies and atypical frontotemporal lobar degeneration. *Neuropathology* **40**, 620-626 (2020).

295. Singleton, A. *et al.* alpha-synuclein locus triplication causes Parkinson's disease. *Science* **302**, 841-842 (2003).
296. Chartier-Harlin, M.-C. *et al.*  $\alpha$ -synuclein locus duplication as a cause of familial Parkinson's disease. *The Lancet* **364**, 1167-1169 (2004).
297. Simon-Sanchez, J. *et al.* Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat. Genet.* **41**, 1308-1312 (2009).
298. Lázaro, D.F. *et al.* Systematic comparison of the effects of alpha-synuclein mutations on its oligomerization and aggregation. *PLoS Genet.* **10**, e1004741 (2014).
299. Kumar, S.T. *et al.* A novel mutation (E83Q) unlocks the pathogenicity of human alpha-synuclein fibrils and recapitulates its pathological diversity. *bioRxiv* (2021).
300. Ping, L. *et al.* Global quantitative analysis of the human brain proteome in Alzheimer's and Parkinson's Disease. *Scientific data* **5**, 180036 (2018).
301. Haaxma, C.A. *et al.* Gender differences in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **78**, 819-824 (2007).
302. Reeve, A., Simcox, E. & Turnbull, D. Ageing and Parkinson's disease: why is advancing age the biggest risk factor? *Ageing Res. Revs.* **14**, 19-30 (2014).
303. Lang, A.E. & Lozano, A.M. Parkinson's disease. *N. Engl. J. Med.* **339**, 1130-1143 (1998).
304. Spillantini, M.G. *et al.* alpha-synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 6469-6473 (1998).
305. Lashuel, H.A. Do Lewy bodies contain alpha-synuclein fibrils? and Does it matter? A brief history and critical analysis of recent reports. *Neurobiol. Dis.* **141**, 104876 (2020).
306. Martinez-Martin, P. *et al.* Parkinson symptoms and health related quality of life as predictors of costs: a longitudinal observational study with linear mixed model analysis. *PLoS One* **10**, e0145310 (2015).
307. Delenclos, M. *et al.* Cellular models of alpha-synuclein toxicity and aggregation. *J. Neurochem.* **150**, 566-576 (2019).
308. Ünal, İ. & Emekli-Alturfan, E. Fishing for Parkinson's Disease: a review of the literature. *J. Clin. Neurosci.* **62**, 1-6 (2019).
309. Aryal, B. & Lee, Y. Disease model organism for Parkinson disease: *Drosophila melanogaster*. *BMB reports* **52**, 250 (2019).
310. Bussell Jr, R. & Eliezer, D. A structural and functional role for 11-mer repeats in  $\alpha$ -synuclein and other exchangeable lipid binding proteins. *J. Mol. Biol.* **329**, 763-778 (2003).



311. Rodriguez, J.A. *et al.* Structure of the toxic core of alpha-synuclein from invisible crystals. *Nature* **525**, 486-490 (2015).
312. Ueda, K. *et al.* Molecular-cloning of cDNA-encoding an unrecognized component of amyloid in Alzheimer-Disease. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 11282-11286 (1993).
313. Giasson, B.I., Murray, I.V., Trojanowski, J.Q. & Lee, V.M. A hydrophobic stretch of 12 amino acid residues in the middle of alpha-synuclein is essential for filament assembly. *J. Biol. Chem.* **276**, 2380-2386 (2001).
314. Ulmer, T.S., Bax, A., Cole, N.B. & Nussbaum, R.L. Structure and dynamics of micelle-bound human  $\alpha$ -synuclein. *J. Biol. Chem.* **280**, 9595-9603 (2005).
315. Moons, R. *et al.* Metal ions shape  $\alpha$ -synuclein. *Sci. Rep.* **10**, 1-13 (2020).
316. Eliezer, D., Kutluay, E., Bussell, R. & Browne, G. Conformational properties of alpha-synuclein in its free and lipid-associated states. *J. Mol. Biol.* **307**, 1061-1073 (2001).
317. Uversky, V.N., Li, J. & Fink, A.L. Evidence for a partially folded intermediate in alpha-synuclein fibril formation. *J. Biol. Chem.* **276**, 10737-10744 (2001).
318. Cho, M.K. *et al.* Amino acid bulkiness defines the local conformations and dynamics of natively unfolded alpha-synuclein and tau. *J. Am. Chem. Soc.* **129**, 3032 (2007).
319. Allison, J.R., Varnai, P., Dobson, C.M. & Vendruscolo, M. Determination of the free energy landscape of alpha-synuclein using spin label nuclear magnetic resonance measurements. *J. Am. Chem. Soc.* **131**, 18314-18326 (2009).
320. Salmon, L. *et al.* NMR characterization of long-range order in intrinsically disordered proteins. *J. Am. Chem. Soc.* **132**, 8407-8418 (2010).
321. Bertocini, C.W. *et al.* Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 1430-1435 (2005).
322. Rao, J.N. *et al.* A Combinatorial NMR and EPR approach for evaluating the structural ensemble of partially folded proteins. *J. Am. Chem. Soc.* **132**, 8657-8668 (2010).
323. Lee, J.C. *et al.* alpha-Synuclein tertiary contact dynamics. *J. Phys. Chem.* **111**, 2107-2112 (2007).
324. Stephens, A.D. *et al.* Different structural conformers of monomeric  $\alpha$ -synuclein identified after lyophilizing and freezing. *Anal. Chem.* **90**, 6975-6983 (2018).
325. Dedmon, M.M. *et al.* Mapping long-range interactions in  $\alpha$ -synuclein using spin-label NMR and ensemble molecular dynamics simulations. *J. Am. Chem. Soc.* **127**, 476-477 (2005).

326. Fernandez, C.O. *et al.* NMR of  $\alpha$ -synuclein–polyamine complexes elucidates the mechanism and kinetics of induced aggregation. *EMBO J.* **23**, 2039-2046 (2004).
327. Zhou, W. *et al.* Methionine oxidation stabilizes non-toxic oligomers of  $\alpha$ -synuclein through strengthening the auto-inhibitory intra-molecular long-range interactions. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1802**, 322-330 (2010).
328. Esteban-Martín, S., Silvestre-Ryan, J., Bertoncini, C.W. & Salvatella, X. Identification of fibril-like tertiary contacts in soluble monomeric  $\alpha$ -synuclein. *Biophys. J.* **105**, 1192-1198 (2013).
329. Bernadó, P. *et al.* Defining long-range order and local disorder in native  $\alpha$ -synuclein using residual dipolar couplings. *J. Am. Chem. Soc.* **127**, 17968-17969 (2005).
330. Sahay, S. *et al.* Familial Parkinson disease-associated mutations alter the site-specific microenvironment and dynamics of  $\alpha$ -synuclein. *J. Biol. Chem.* **290**, 7804-7822 (2015).
331. Wu, K.P. *et al.* Structural reorganization of alpha-synuclein at low pH observed by NMR and REMD simulations. *J. Mol. Biol.* **391**, 784-796 (2009).
332. McClendon, S., Rospigliosi, C.C. & Eliezer, D. Charge neutralization and collapse of the C-terminal tail of alpha-synuclein at low pH. *Protein Sci.* **18**, 1531-1540 (2009).
333. Sung, Y.-h. & Eliezer, D. Residual structure, backbone dynamics, and interactions within the synuclein family. *J. Mol. Biol.* **372**, 689 (2007).
334. Ranjan, P. & Kumar, A. Perturbation in long-range contacts modulates the kinetics of amyloid formation in  $\alpha$ -synuclein familial mutants. *ACS Chem. Neurosci.* **8**, 2235-2246 (2017).
335. Cho, M.K. *et al.* Structural characterization of  $\alpha$ -synuclein in an aggregation prone state. *Protein Sci.* **18**, 1840-1846 (2009).
336. Ahmed, M.C. *et al.* Refinement of  $\alpha$ -synuclein ensembles against SAXS data: Comparison of force fields and methods. *Front. Mol. Biosci.* **8**, 216 (2021).
337. Nguyen, P.H. & Derreumaux, P. Structures of the intrinsically disordered A $\beta$ , tau and  $\alpha$ -synuclein proteins in aqueous solution from computer simulations. *Biophys. Chem.* **264**, 106421 (2020).
338. Lindsay, R.J., Mansbach, R.A., Gnanakaran, S. & Shen, T. Effects of pH on an IDP conformational ensemble explored by molecular dynamics simulation. *Biophys. Chem.* **271**, 106552 (2021).
339. Guzzo, A. *et al.* Missense mutations modify the conformational ensemble of the  $\alpha$ -synuclein monomer which exhibits a two-phase characteristic. *Front. Mol. Biosci.* **8**(2021).

340. Robustelli, P. *et al.* Molecular basis of small-molecule binding to  $\alpha$ -synuclein. *bioRxiv* (2021).
341. Salveson, P.J., Spencer, R.K. & Nowick, J.S. X-ray crystallographic structure of oligomers formed by a toxic  $\beta$ -hairpin derived from  $\alpha$ -synuclein: trimers and higher-order oligomers. *J. Am. Chem. Soc.* **138**, 4458-4467 (2016).
342. Falkenburger, B.H. *et al.* A  $\beta$ -wrapin targeting the N-terminus of  $\alpha$ -synuclein monomers reduces fibril-induced aggregation in neurons. *Front. Neurosci.* **15**, 751 (2021).
343. Stöckl, M.T., Zijlstra, N. & Subramaniam, V.  $\alpha$ -Synuclein oligomers: an amyloid pore? *Mol. Neurobiol.* **47**, 613-621 (2013).
344. Mysling, S., Betzer, C., Jensen, P.H. & Jorgensen, T.J. Characterizing the dynamics of  $\alpha$ -synuclein oligomers using hydrogen/deuterium exchange monitored by mass spectrometry. *Biochemistry* **52**, 9097-9103 (2013).
345. McGlinchey, R.P. *et al.* The N terminus of  $\alpha$ -synuclein dictates fibril formation. *Proc. Natl. Acad. Sci.* **118**, e2023487118 (2021).
346. Ni, X., McGlinchey, R.P., Jiang, J. & Lee, J.C. Structural insights into  $\alpha$ -synuclein fibril polymorphism: Effects of Parkinson's disease-related C-terminal truncations. *J. Mol. Biol.* **431**, 3913-3919 (2019).
347. Zhao, K. *et al.* Parkinson's disease associated mutation E46K of  $\alpha$ -synuclein triggers the formation of a distinct fibril structure. *Nat. Commun.* **11**, 1-9 (2020).
348. Boyer, D.R. *et al.* The  $\alpha$ -synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure. *Proc. Natl. Acad. Sci.* **117**, 3592-3602 (2020).
349. Boyer, D.R. *et al.* Cryo-EM structures of  $\alpha$ -synuclein fibrils with the H50Q hereditary mutation reveal new polymorphs. *bioRxiv*, 738450 (2019).
350. Vargas, K.J. *et al.* Synucleins regulate the kinetics of synaptic vesicle endocytosis. *J. Neurosci.* **34**, 9364-9376 (2014).
351. Galvagnion, C. *et al.* Lipid vesicles trigger alpha-synuclein aggregation by stimulating primary nucleation. *Nat. Chem. Biol.* **11**, 229-234 (2015).
352. Fusco, G., Sanz-Hernandez, M. & De Simone, A. Order and disorder in the physiological membrane binding of  $\alpha$ -synuclein. *Curr. Opin. Struct. Biol.* **48**, 49-57 (2018).
353. Cheng, C.-Y. *et al.* Hydration dynamics as an intrinsic ruler for refining protein structure at lipid membrane interfaces. *Proc. Natl. Acad. Sci.* **110**, 16838-16843 (2013).
354. Georgieva, E.R. *et al.* The lipid-binding domain of wild type and mutant  $\alpha$ -synuclein: compactness and interconversion between the broken and extended helix forms. *J. Biol. Chem.* **285**, 28261-28274 (2010).

355. Pfefferkorn, C.M., Jiang, Z. & Lee, J.C. Biophysics of  $\alpha$ -synuclein membrane interactions. *Biochim. Biophys. Acta. Biomembr.* **1818**, 162-171 (2012).
356. Bodner, C.R., Dobson, C.M. & Bax, A. Multiple tight phospholipid-binding modes of  $\alpha$ -synuclein revealed by solution NMR spectroscopy. *J. Mol. Biol.* **390**, 775-790 (2009).
357. Kjaer, L., Giehm, L., Heimburg, T. & Otzen, D. The influence of vesicle size and composition on  $\alpha$ -synuclein structure and stability. *Biophys. J.* **96**, 2857-2870 (2009).
358. Middleton, E.R. & Rhoades, E. Effects of curvature and composition on  $\alpha$ -synuclein binding to lipid vesicles. *Biophys. J.* **99**, 2279-2288 (2010).
359. Pandey, A.P., Haque, F., Rochet, J.-C. & Hovis, J.S. Clustering of  $\alpha$ -synuclein on supported lipid bilayers: role of anionic lipid, protein, and divalent ion concentration. *Biophys. J.* **96**, 540-551 (2009).
360. Newberry, R.W. *et al.* Robust sequence determinants of  $\alpha$ -synuclein toxicity in yeast implicate membrane binding. *ACS Chem. Biol.* **15**, 2137-2153 (2020).
361. Newberry, R.W. *et al.* Deep mutational scanning reveals the structural basis for  $\alpha$ -synuclein activity. *Nat. Chem. Biol.* **16**, 653-659 (2020).
362. Akopian, A.N. & Wood, J.N. Peripheral nervous system-specific genes identified by subtractive cDNA cloning. *J. Biol. Chem.* **270**, 21264-21270 (1995).
363. George, J.M. The synucleins. *Genome Biol.* **3**, 3002 (2001).
364. Galvin, J.E., Uryu, K., Lee, V.M.-Y. & Trojanowski, J.Q. Axon pathology in Parkinson's disease and Lewy body dementia hippocampus contains  $\alpha$ -,  $\beta$ -, and  $\gamma$ -synuclein. *Proc. Natl. Acad. Sci.* **96**, 13450-13455 (1999).
365. Uversky, V.N. *et al.* Biophysical properties of the synucleins and their propensities to fibrillate inhibition of  $\alpha$ -synuclein assembly by  $\beta$ - and  $\gamma$ -synucleins. *J. Biol. Chem.* **277**, 11970-11978 (2002).
366. Marsh, J.A., Singh, V.K., Jia, Z. & Forman-Kay, J.D. Sensitivity of secondary structure propensities to sequence differences between  $\alpha$ - and  $\gamma$ -synuclein: Implications for fibrillation. *Protein Sci.* **15**, 2795-2804 (2006).
367. Hashimoto, M. *et al.*  $\beta$ -Synuclein inhibits  $\alpha$ -synuclein aggregation: a possible role as an anti-parkinsonian factor. *Neuron* **32**, 213-223 (2001).
368. Ji, H. *et al.* Identification of a breast cancer-specific gene, BCSG1, by direct differential cDNA sequencing. *Cancer Res.* **57**, 759-764 (1997).
369. Buchman, V.L. *et al.* Persyn, a member of the synuclein family, has a distinct pattern of expression in the developing nervous system. *J. Neurosci.* **18**, 9335-9341 (1998).

370. Lavedan, C. *et al.* Identification, localization and characterization of the human  $\gamma$ -synuclein gene. *Hum. Genet.* **103**, 106-112 (1998).
371. Jiang, Y., Liu, Y.E., Goldberg, I.D. & Shi, Y.E.  $\gamma$  synuclein, a novel heat-shock protein-associated chaperone, stimulates ligand-dependent estrogen receptor  $\alpha$  signaling and mammary tumorigenesis. *Cancer Res.* **64**, 4539-4546 (2004).
372. Gupta, A. *et al.* Breast cancer-specific gene 1 interacts with the mitotic checkpoint kinase BubR1. *Oncogene* **22**, 7593 (2003).
373. Wu, K. *et al.* Expression of neuronal protein synuclein gamma gene as a novel marker for breast cancer prognosis. *Breast Cancer Res. Treat.* **101**, 259-267 (2007).
374. Bruening, W. *et al.* Synucleins are expressed in the majority of breast and ovarian carcinomas and in preneoplastic lesions of the ovary. *Cancer: Interdisciplinary International Journal of the American Cancer Society* **88**, 2154-2163 (2000).
375. Peters, O.M. *et al.* Gamma-synuclein pathology in amyotrophic lateral sclerosis. *Ann. Clin. Transl. Neurol.* **2**, 29-37 (2015).
376. Ninkina, N. *et al.*  $\gamma$ -Synucleinopathy: neurodegeneration associated with overexpression of the mouse protein. *Hum. Mol. Genet.* **18**, 1779-1794 (2009).
377. Peters, O.M. *et al.* Selective pattern of motor system damage in gamma-synuclein transgenic mice mirrors the respective pathology in amyotrophic lateral sclerosis. *Neurobiol. Dis.* **48**, 124-131 (2012).
378. Uversky, V.N. & Fink, A.L. Amino acid determinants of  $\alpha$ -synuclein aggregation: putting together pieces of the puzzle. *FEBS Lett.* **522**, 9-13 (2002).
379. Sung, Y.h. & Eliezer, D. Secondary structure and dynamics of micelle bound  $\beta$ - and  $\gamma$ -synuclein. *Protein Sci.* **15**, 1162-1174 (2006).
380. Ducas, V.C. & Rhoades, E. Quantifying interactions of  $\beta$ -synuclein and  $\gamma$ -synuclein with model membranes. *J. Mol. Biol.* **423**, 528-539 (2012).
381. Jain, M.K., Singh, P., Roy, S. & Bhat, R. Comparative analysis of the conformation, aggregation, interaction, and fibril morphologies of human  $\alpha$ -,  $\beta$ -, and  $\gamma$ -synuclein proteins. *Biochemistry* **57**, 3830-3848 (2018).
382. Golebiewska, U., Zurawsky, C. & Scarlata, S. Defining the oligomerization state of  $\gamma$ -synuclein in solution and in cells. *Biochemistry* **53**, 293-299 (2014).
383. Melanson, M., Sood, A., Török, F. & Török, M. Introduction to spin label electron paramagnetic resonance spectroscopy of proteins. *Biochem. Mol. Biol. Educ.* **41**, 156-162 (2013).
384. Volkov, A.N., Worrall, J.A., Holtzmann, E. & Ubbink, M. Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase

- determined by paramagnetic NMR. *Proc. Natl. Acad. Sci.* **103**, 18945-18950 (2006).
385. Gottstein, D., Reckel, S., Dötsch, V. & Güntert, P. Requirements on paramagnetic relaxation enhancement data for membrane protein structure determination by NMR. *Structure* **20**, 1019-1027 (2012).
386. Matei, E. & Gronenborn, A.M. <sup>19</sup>F paramagnetic relaxation enhancement: a valuable tool for distance measurements in proteins. *Angew. Chem. Int. Ed.* **55**, 150-154 (2016).
387. Fogh, R. *et al.* The CCPN project: an interim report on a data model for the NMR community. *Nat. Struct. Biol.* **9**, 416-418 (2002).
388. Ahlner, A., Carlsson, M., Jonsson, B.-H. & Lundström, P. PINT: a software for integration of peak volumes and extraction of relaxation rates. *J. Biomol. NMR* **56**, 191-202 (2013).
389. Johnson Jr, W.C. Protein secondary structure and circular dichroism: a practical guide. *Proteins: Struct. Funct. Genet.* **7**, 205-214 (1990).
390. Greenfield, N.J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876 (2006).
391. Holzwarth, G. & Doty, P. The ultraviolet circular dichroism of polypeptides<sup>1</sup>. *J. Am. Chem. Soc.* **87**, 218-228 (1965).
392. Whitmore, L. & Wallace, B. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32**, 668-673 (2004).
393. Tartaglia, G.G. & Vendruscolo, M. The Zyggregator method for predicting protein aggregation propensities. *Chem. Soc. Rev.* **37**, 1395-1401 (2008).
394. Thompson, M.J. *et al.* The 3D profile method for identifying fibril-forming segments of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4074-4078 (2006).
395. Ashkenazy, H. *et al.* ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, 529-533 (2010).
396. Link, C.D. Expression of human beta-amyloid peptide in transgenic *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* **92**, 9368-9372 (1995).
397. Diomedea, L. *et al.* *C. elegans* expressing human  $\beta$ 2-microglobulin: a novel model for studying the relationship between the molecular assembly and the toxic phenotype. *PLoS One* **7**, e52314 (2012).
398. Morley, J.F., Brignull, H.R., Weyers, J.J. & Morimoto, R.I. The threshold for polyglutamine-expansion protein aggregation and cellular toxicity is dynamic and

- influenced by aging in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* **99**, 10417-10422 (2002).
399. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71-94 (1974).
400. Sulston, J. Neuronal cell lineages in the nematode *Caenorhabditis elegans*. *Cold Spring Harbor symposia on quantitative biology* **48**, 443-452 (Cold Spring Harbor Laboratory Press, 1983).
401. Markaki, M. & Tavernarakis, N. Modeling human diseases in *Caenorhabditis elegans*. *Biotechnol. J.* **5**, 1261-1276 (2010).
402. Nussbaum-Krammer, C.I. *et al.* Investigating the spreading and toxicity of prion-like proteins using the metazoan model organism *C. elegans*. *J. Vis. Exp.* **95**, e52321 (2015).
403. Brás, J., Gibbons, E. & Guerreiro, R. Genetics of synucleins in neurodegenerative diseases. *Acta Neuropathol.* **141**, 471-490 (2021).
404. Perez, R.G. The protein alpha-synuclein: its normal role (in neurons) and its role in disease. *Front. Neurosci.* **14**, 116 (2020).
405. El-Agnaf, O.M. *et al.* Aggregates from mutant and wild-type  $\alpha$ -synuclein proteins and NAC peptide induce apoptotic cell death in human neuroblastoma cells by formation of  $\beta$ -sheet and amyloid-like filaments. *FEBS Lett.* **440**, 71-75 (1998).
406. Das, M. *et al.* Amyloidogenic mutations in human apolipoprotein A-I are not necessarily destabilizing—a common mechanism of apolipoprotein A-I misfolding in familial amyloidosis and atherosclerosis. *FEBS J.* **281**, 2525-2542 (2014).
407. Hoop, C.L. *et al.* Polyglutamine amyloid core boundaries and flanking domain dynamics in huntingtin fragment fibrils determined by solid-state nuclear magnetic resonance. *Biochemistry* **53**, 6653-6666 (2014).
408. Lucato, C.M., Lupton, C.J., Halls, M.L. & Ellisdon, A.M. Amyloidogenicity at a distance: how distal protein regions modulate aggregation in disease. *J. Mol. Biol.* **429**, 1289-1304 (2017).
409. Koo, H.-J., Choi, M.Y. & Im, H. Aggregation-defective  $\alpha$ -synuclein mutants inhibit the fibrillation of Parkinson's disease-linked  $\alpha$ -synuclein variants. *Biochem. Biophys. Res. Commun.* **386**, 165-169 (2009).
410. Goldschmidt, L., Teng, P.K., Riek, R. & Eisenberg, D. Identifying the amyloids, proteins capable of forming amyloid-like fibrils. *Proc. Natl. Acad. Sci.* **107**, 3487-3492 (2010).
411. Terada, M. *et al.* The effect of truncation on prion-like properties of  $\alpha$ -synuclein. *J. Biol. Chem.* **293**, 13910-13920 (2018).

412. Kessler, J.C., Rochet, J.-C. & Lansbury, P.T. The N-terminal repeat domain of  $\alpha$ -synuclein inhibits  $\beta$ -sheet and amyloid fibril formation. *Biochemistry* **42**, 672-678 (2003).
413. Gallardo, J., Escalona-Noguero, C. & Sot, B. Role of  $\alpha$ -synuclein regions in nucleation and elongation of amyloid fiber assembly. *ACS Chem. Neurosci.* **11**, 872-879 (2020).
414. Shaykhalishahi, H. *et al.* Contact between the beta1 and beta2 segments of alpha-synuclein that inhibits amyloid formation. *Angew. Chem. Int. Ed. Engl.* **54**, 8837-8840 (2015).
415. Agerschou, E.D. *et al.* An engineered monomer binding-protein for  $\alpha$ -synuclein efficiently inhibits the proliferation of amyloid fibrils. *Elife* **8**, e46112 (2019).
416. Doherty, C. *et al.* A short motif in the N-terminal region of  $\alpha$ -synuclein is critical for both aggregation and function. *Nat. Struct. Mol. Biol.* **27**, 249-259 (2020).
417. Ramis, R. *et al.* Unraveling the NaCl concentration effect on the first stages of  $\alpha$ -synuclein aggregation. *Biomacromolecules* **21**, 5200-5212 (2020).
418. Ben-Zvi, A., Miller, E.A. & Morimoto, R.I. Collapse of proteostasis represents an early molecular event in *Caenorhabditis elegans* aging. *Proc. Natl. Acad. Sci.* **106**, 14914-14919 (2009).
419. Labbadia, J. & Morimoto, R.I. Repression of the heat shock response is a programmed event at the onset of reproduction. *Mol. Cell* **59**, 639-650 (2015).
420. Dyson, H.J. & Wright, P.E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell. Biol.* **6**, 197-208 (2005).
421. Lindorff-Larsen, K. & Kragelund, B.B. On the potential of machine learning to examine the relationship between sequence, structure, dynamics and function of intrinsically disordered proteins. *J. Mol. Biol.* **433**, 167196 (2021).
422. Hoyer, W. *et al.* Dependence of  $\alpha$ -synuclein aggregate morphology on solution conditions. *J. Mol. Biol.* **322**, 383-393 (2002).
423. Schapira, A.H.V. Glucocerebrosidase and Parkinson disease: recent advances. *Mol. Cell. Neurosci.* **66**, 37-42 (2015).
424. Kleckner, I.R. & Foster, M.P. An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta - Proteins Proteom.* **1814**, 942-968 (2011).
425. Silvestre-Ryan, J. *et al.* Average conformations determined from PRE data provide high-resolution maps of transient tertiary interactions in disordered proteins. *Biophys. J.* **104**, 1740-1751 (2013).



426. Wu, K.-P. & Baum, J. Detection of transient interchain interactions in the intrinsically disordered protein  $\alpha$ -synuclein by NMR paramagnetic relaxation enhancement. *J. Am. Chem. Soc.* **132**, 5546-5547 (2010).
427. Wu, K.-P., Kim, S., Fela, D.A. & Baum, J. Characterization of conformational and dynamic properties of natively unfolded human and mouse  $\alpha$ -synuclein ensembles by NMR: implication for aggregation. *J. Mol. Biol.* **378**, 1104-1115 (2008).
428. Bertocini, C.W. *et al.* Familial mutants of  $\alpha$ -synuclein with increased neurotoxicity have a destabilized conformation. *J. Biol. Chem.* **280**, 30649-30652 (2005).
429. Janowska, M.K., Wu, K.-P. & Baum, J. Unveiling transient protein-protein interactions that modulate inhibition of alpha-synuclein aggregation by beta-synuclein, a pre-synaptic protein that co-localizes with alpha-synuclein. *Sci. Rep.* **5**, 15164-15164 (2015).
430. Lee, H.-g. *et al.* Emerging evidence for the neuroprotective role of alpha-synuclein. *Exp. Neurol.* **200**, 1 (2006).
431. Emamzadeh, F.N. Alpha-synuclein structure, functions, and interactions. *J. Res. Med. Sci.* **21**, 1-9 (2016).
432. Jao, C.C. *et al.* Structure of membrane-bound  $\alpha$ -synuclein from site-directed spin labeling and computational refinement. *Proc. Natl. Acad. Sci.* **105**, 19666-19671 (2008).
433. Fonseca-Ornelas, L. *et al.* Small molecule-mediated stabilization of vesicle-associated helical  $\alpha$ -synuclein inhibits pathogenic misfolding and aggregation. *Nat. Commun.* **5**, 1-11 (2014).
434. Bugg, C.W. *et al.* Structural features and domain organization of huntingtin fibrils. *J. Biol. Chem.* **287**, 31739-31746 (2012).
435. Brännström, K. *et al.* The N-terminal region of amyloid  $\beta$  controls the aggregation rate and fibril stability at low pH through a gain of function mechanism. *J. Am. Chem. Soc.* **136**, 10956-10964 (2014).
436. Goedert, M. Alpha-synuclein and neurodegenerative diseases. *Nat. Rev. Neurosci.* **2**, 492 (2001).
437. González, N. *et al.* Effects of alpha-synuclein post-translational modifications on metal binding. *J. Neurochem.* **150**, 507-521 (2019).
438. Bungeoth, M. *et al.* Differential aggregation properties of alpha-synuclein isoforms. *Neurobiol. Aging* **35**, 1913-1919 (2014).

439. Liang, Z., Chan, H.Y.E., Lee, M.M. & Chan, M.K. A SUMO1-derived peptide targeting SUMO-interacting motif inhibits  $\alpha$ -synuclein aggregation. *Cell. Chem. Biol.* **28**, 180-190 (2021).
440. Ulrih, N.P., Barry, C.H. & Fink, A.L. Impact of Tyr to Ala mutations on  $\alpha$ -synuclein fibrillation and structural properties. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1782**, 581-585 (2008).
441. Khammari, A., Arab, S.S. & Ejtehad, M.R. The hot sites of  $\alpha$ -synuclein in amyloid fibril formation. *Sci. Rep.* **10**, 1-14 (2020).
442. Houben, B. *et al.* Autonomous aggregation suppression by acidic residues explains why chaperones favour basic residues. *EMBO J.* **39**, e102864 (2020).
443. Zhou, W. & Freed, C.R. Tyrosine-to-cysteine modification of human  $\alpha$ -synuclein enhances protein aggregation and cellular toxicity. *J. Biol. Chem.* **279**, 10128-10135 (2004).
444. Sode, K., Usuzaka, E., Kobayashi, N. & Ochiai, S. Engineered  $\alpha$ -synuclein prevents wild type and familial Parkin variant fibril formation. *Biochem. Biophys. Res. Commun.* **335**, 432-436 (2005).
445. Waxman, E.A., Mazzulli, J.R. & Giasson, B.I. Characterization of hydrophobic residue requirements for  $\alpha$ -synuclein fibrillization. *Biochemistry* **48**, 9427-9436 (2009).
446. Zhang, J., Li, X. & Li, J.-D. The roles of post-translational modifications on  $\alpha$ -synuclein in the pathogenesis of Parkinson's diseases. *Front. Neurosci.* **13**, 381 (2019).
447. He, S. *et al.* Effects of  $\alpha$ -synuclein-associated post-translational modifications in Parkinson's Disease. *ACS Chem. Neurosci.* **12**, 1061-1071 (2021).
448. Williamson, M.P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1-16 (2013).
449. Van Meerloo, J., Kaspers, G.J. & Cloos, J. Cell sensitivity assays: the MTT assay. *Cancer cell culture* 237-245 (Springer, 2011).
450. Kurbacher, C.M. *et al.* Use of an *ex vivo* ATP luminescence assay to direct chemotherapy for recurrent ovarian cancer. *Anticancer Drugs* **9**, 51-57 (1998).
451. Smith, S.M., Wunder, M.B., Norris, D.A. & Shellman, Y.G. A simple protocol for using a LDH-based cytotoxicity assay to assess the effects of death and growth inhibition at the same time. *PLoS One* **6**, e26908 (2011).
452. Korzeniewski, C. & Callewaert, D.M. An enzyme-release assay for natural cytotoxicity. *J. Immunol. Methods* **64**, 313-320 (1983).

453. Martínez, M.-A. *et al.* Use of human neuroblastoma SH-SY5Y cells to evaluate glyphosate-induced effects on oxidative stress, neuronal development and cell death signaling pathways. *Environ. Int.* **135**, 105414 (2020).
454. 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).
455. Biere, A.L. *et al.* Parkinson's disease-associated  $\alpha$ -synuclein is more fibrillogenic than  $\beta$ - and  $\gamma$ -synuclein and cannot cross-seed its homologs. *J. Biol. Chem.* **275**, 34574-34579 (2000).
456. Williams, J.K. *et al.* Multi-pronged interactions underlie inhibition of  $\alpha$ -synuclein aggregation by  $\beta$ -synuclein. *J. Mol. Biol.* **430**, 2360-2371 (2018).
457. Surguchov, A., Palazzo, R.E. & Surgucheva, I. Gamma synuclein: subcellular localization in neuronal and non-neuronal cells and effect on signal transduction. *Cell Motil. Cytoskeleton* **49**, 218-228 (2001).
458. Agerschou, E.D. *et al.*  $\beta$ -Turn exchanges in the  $\alpha$ -synuclein segment 44-TKEG-47 reveal high sequence fidelity requirements of amyloid fibril elongation. *Biophys. Chem.* **269**, 106519 (2021).
459. Mohite, G.M. *et al.* Comparison of kinetics, toxicity, oligomer formation, and membrane binding capacity of  $\alpha$ -synuclein familial mutations at the A53 site, including the newly discovered A53V mutation. *Biochemistry* **57**, 5183-5187 (2018).
460. Stephens, A.D. *et al.* Extent of N-terminus exposure of monomeric alpha-synuclein determines its aggregation propensity. *Nat. Commun.* **11**, 1-15 (2020).
461. Brodie, N.I. *et al.* Conformational ensemble of native  $\alpha$ -synuclein in solution as determined by short-distance crosslinking constraint-guided discrete molecular dynamics simulations. *PLoS Comput. Biol.* **15**, e1006859 (2019).
462. Seetaloo, N. *et al.* Local structural dynamics of alpha-synuclein correlate with aggregation in different physiological conditions. *bioRxiv* (2022).
463. Aguirre, C. *et al.* Conformational change in the monomeric alpha-synuclein imparts fibril polymorphs. *bioRxiv* (2022).
464. Cawood, E.E. *et al.* Modulation of amyloidogenic protein self-assembly using tethered small molecules. *J. Am. Chem. Soc.* **142**, 20845-20854 (2020).
465. Cawood, E.E., Karamanos, T.K., Wilson, A.J. & Radford, S.E. Visualizing and trapping transient oligomers in amyloid assembly pathways. *Biophys. Chem.* **268**, 106505 (2021).
466. Ebo, J.S. *et al.* An in vivo platform to select and evolve aggregation-resistant proteins. *Nat. Commun.* **11**, 1-12 (2020).

467. Ling, S.-C., Polymenidou, M. & Cleveland, D.W. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79**, 416-438 (2013).
468. Roodveldt, C. *et al.* A rationally designed six-residue swap generates comparability in the aggregation behavior of  $\alpha$ -synuclein and  $\beta$ -synuclein. *Biochemistry* **51**, 8771-8778 (2012).
469. Nishioka, K. *et al.* Association of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -synuclein with diffuse Lewy body disease. *Arch. Neurol.* **67**, 970-975 (2010).
470. Masaracchia, C. *et al.* Molecular characterization of an aggregation-prone variant of alpha-synuclein used to model synucleinopathies. *Biochim. Biophys. Acta - Proteins Proteom.* **1868**, 140298 (2020).
471. Ghosh, D. *et al.* The Parkinson's disease-associated H50Q mutation accelerates  $\alpha$ -synuclein aggregation in vitro. *Biochemistry* **52**, 6925-6927 (2013).
472. Fredenburg, R.A. *et al.* The impact of the E46K mutation on the properties of  $\alpha$ -synuclein in its monomeric and oligomeric states. *Biochemistry* **46**, 7107-7118 (2007).
473. Perni, M. *et al.* Comparative studies in the A30P and A53T  $\alpha$ -synuclein *C. elegans* strains to investigate the molecular origins of Parkinson's Disease. *Front. Cell Dev. Biol.* **9**, 492 (2021).
474. Van Raaij, M.E. *et al.* Concentration dependence of  $\alpha$ -synuclein fibril length assessed by quantitative atomic force microscopy and statistical-mechanical theory. *Biophys. J.* **95**, 4871-4878 (2008).
475. Bousset, L. *et al.* Structural and functional characterization of two alpha-synuclein strains. *Nat. Commun.* **4**, 2575 (2013).
476. Sakagashira, S. *et al.* Missense mutation of amylin gene (S20G) in Japanese NIDDM patients. *Diabetes* **45**, 1279-1281 (1996).
477. Valleix, S. *et al.* Hereditary systemic amyloidosis due to Asp76Asn variant  $\beta$ 2-microglobulin. *N. Engl. J. Med.* **366**, 2276-2283 (2012).
478. Assimon, V.A., Southworth, D.R. & Gestwicki, J.E. Specific binding of tetratricopeptide repeat proteins to heat shock protein 70 (Hsp70) and heat shock protein 90 (Hsp90) is regulated by affinity and phosphorylation. *Biochemistry* **54**, 7120-7131 (2015).
479. Goldfarb, L.G. *et al.* Fatal familial insomnia and familial Creutzfeldt-Jakob disease: disease phenotype determined by a DNA polymorphism. *Science* **258**, 806-808 (1992).
480. Krebs, M.R. *et al.* Observation of sequence specificity in the seeding of protein amyloid fibrils. *Protein Sci.* **13**, 1933-1938 (2004).

481. O'Nuallain, B., Williams, A.D., Westermark, P. & Wetzel, R. Seeding specificity in amyloid growth induced by heterologous fibrils. *J. Biol. Chem.* **279**, 17490-17499 (2004).
482. Wiley, J. Alzheimer's disease facts and figures. *Alzheimers Dement* **17**, 327-406 (2021).
483. Ou, Z. *et al.* Global trends in the incidence, prevalence, and years lived with disability of Parkinson's disease in 204 countries/territories from 1990 to 2019. *Frontiers in public health* **9**, 1994 (2021).
484. Eisenberg, D.S. & Sawaya, M.R. Structural studies of amyloid proteins at the molecular level. *Annu. Rev. Biochem.* **86**, 69-95 (2017).
485. Ivarsson, Y. & Jemth, P. Affinity and specificity of motif-based protein-protein interactions. *Curr. Opin. Struct. Biol.* **54**, 26-33 (2019).
486. Schreiber, G. Protein-protein interaction interfaces and their functional implications. *Protein-Protein Interaction Regulators*, 1-24 (2020).
487. Holland, D.O., Shapiro, B.H., Xue, P. & Johnson, M.E. Protein-protein binding selectivity and network topology constrain global and local properties of interface binding networks. *Sci. Rep.* **7**, 1-14 (2017).
488. Bose, D. & Chakrabarti, A. Substrate specificity in the context of molecular chaperones. *IUBMB life* **69**, 647-659 (2017).
489. Saunders, J.C. *et al.* An in vivo platform for identifying inhibitors of protein aggregation. *Nat. Chem. Biol.* **12**, 94-101 (2016).
490. Brudek, T. *et al.* Altered  $\alpha$ -synuclein, parkin, and synphilin isoform levels in multiple system atrophy brains. *J. Neurochem.* **136**, 172-185 (2016).
491. Valastyan, J.S., Termine, D.J. & Lindquist, S. Splice isoform and pharmacological studies reveal that sterol depletion relocalizes  $\alpha$ -synuclein and enhances its toxicity. *Proc. Natl. Acad. Sci.* **111**, 3014-3019 (2014).
492. Manda, K.M. *et al.* The chaperone-like activity of  $\alpha$ -synuclein attenuates aggregation of its alternatively spliced isoform, 112-synuclein in vitro: plausible cross-talk between isoforms in protein aggregation. *PLoS One* **9**, e98657 (2014).
493. Strodel, B. Energy landscapes of protein aggregation and conformation switching in intrinsically disordered proteins. *J. Mol. Biol.* **433**, 167182 (2021).
494. Ullman, O., Fisher, C.K. & Stultz, C.M. Explaining the structural plasticity of  $\alpha$ -synuclein. *J. Am. Chem. Soc.* **133**, 19536-19546 (2011).
495. Nath, A. *et al.* The conformational ensembles of  $\alpha$ -synuclein and tau: combining single-molecule FRET and simulations. *Biophys. J.* **103**, 1940-1949 (2012).

496. Toth, G. *et al.* Targeting the intrinsically disordered structural ensemble of  $\alpha$ -synuclein by small molecules as a potential therapeutic strategy for Parkinson's disease. *PLoS One* **9**, e87133 (2014).
497. Giehm, L., Svergun, D.I., Otzen, D.E. & Vestergaard, B. Low-resolution structure of a vesicle disrupting  $\alpha$ -synuclein oligomer that accumulates during fibrillation. *Proc. Natl. Acad. Sci.* **108**, 3246-3251 (2011).
498. Fitzner, D. *et al.* Cell-type-and brain-region-resolved mouse brain lipidome. *Cell Rep.* **32**, 108132 (2020).
499. Takamori, S. *et al.* Molecular anatomy of a trafficking organelle. *Cell* **127**, 831-846 (2006).
500. Auluck, P.K., Caraveo, G. & Lindquist, S.  $\alpha$ -Synuclein: membrane interactions and toxicity in Parkinson's disease. *Annu. Rev. Cell Dev. Biol.* **26**, 211-233 (2010).
501. Antonschmidt, L. *et al.* Insights into the molecular mechanism of amyloid filament formation: Segmental folding of  $\alpha$ -synuclein on lipid membranes. *Sci. Adv.* **7**, eabg2174 (2021).
502. Cohen, F.E. & Kelly, J.W. Therapeutic approaches to protein-misfolding diseases. *Nature* **426**, 905-909 (2003).
503. Metallo, S.J. Intrinsically disordered proteins are potential drug targets. *Curr. Opin. Chem. Biol.* **14**, 481-488 (2010).
504. Tóth, G. *et al.* Novel small molecules targeting the intrinsically disordered structural ensemble of  $\alpha$ -synuclein protect against diverse  $\alpha$ -synuclein mediated dysfunctions. *Sci. Rep.* **9**, 1-14 (2019).
505. Thomasen, F.E. & Lindorff-Larsen, K. Conformational ensembles of intrinsically disordered proteins and flexible multidomain proteins. *Biochemical Society Transactions* **50**, 541-554 (2022).
506. Lakso, M. *et al.* Dopaminergic neuronal loss and motor deficits in *Caenorhabditis elegans* overexpressing human  $\alpha$ -synuclein. *J. Neurochem.* **86**, 165-172 (2003).
507. Gaeta, A.L., Caldwell, K.A. & Caldwell, G.A. Found in translation: The utility of *C. elegans* alpha-synuclein models of parkinson's disease. *Brain sciences* **9**, 73 (2019).
508. Kuwahara, T. *et al.* Familial Parkinson mutant  $\alpha$ -synuclein causes dopamine neuron dysfunction in transgenic *Caenorhabditis elegans*. *J. Biol. Chem.* **281**, 334-340 (2006).
509. Kim, J., Koo, B.-K. & Knoblich, J.A. Human organoids: model systems for human biology and medicine. *Nat. Rev. Mol. Cell. Biol.* **21**, 571-584 (2020).

510. Voisine, C., Pedersen, J.S. & Morimoto, R.I. Chaperone networks: tipping the balance in protein folding diseases. *Neurobiol. Dis.* **40**, 12-20 (2010).
511. Tittelmeier, J. *et al.* The Hsp110/Hsp70 disaggregation system generates spreading-competent toxic  $\alpha$ -synuclein species. *EMBO J.* **39**, e103954 (2020).
512. Li, J., Uversky, V.N. & Fink, A.L. Effect of familial Parkinson's disease point mutations A30P and A53T on the structural properties, aggregation, and fibrillation of human  $\alpha$ -synuclein. *Biochemistry* **40**, 11604-11613 (2001).