Investigating the Interactions of \(\alpha\)-synuclein with Fibrils and Lipid Bilayer and Structural Insights into the Process of Secondary Nucleation

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

Presented by

PRATIBHA KUMARI

Master of Science in Biology, NCBS, TIFR, Bangalore, India

Born on 29.01.1991

Citizen of India

Accepted on the recommendation of

Prof. Dr. Roland Riek, examiner

Prof. Dr. Paolo Arosio, co-examiner

Prof. Dr. Henning Stahlberg, co-examiner

2020
Acknowledgements

I would like to take this opportunity to thank all those people who have helped me during my Ph.D. It was a wonderful time here at ETH Zurich to discover my potential and capabilities. First and foremost, I would like to thank Prof. Roland Riek who was extremely helpful throughout my Ph.D. His friendly nature and accepting behavior gave me the liberty to try out many projects in the lab and learn a lot. I admire him for an outstanding scientist and more than that I admire him for a nice person. I do not think I can ever acknowledge him enough for the opportunities and the best advice he has given to me.

I would like to thank Kristina Comiotto for her support during the last four years. She is one of the most understanding persons I met, and she was always happy to help with everything from administrative work to find tasty cakes and chocolates in Switzerland. I did not even realize I came so far from my home country because of her attitude to make everything easy and well explained.

I thank Prof. Paolo Arosio and Prof. Henning Stahlberg for accepting to be in my thesis committee member. I would like to thank Prof. Henning Stahlberg and Prof. Paola Picotti for the opportunities to work with collaborative projects and discussions during the Synergia meeting.
I would like to thank Lukas Frey and Cedric Eichmann for being good friend and mentor. I spent a lot of time with them in the lab and it was full of happiness. I would like to thank Silvia Campioni for a short project together where I learn a few perfect ways to do experiments. I would like to thank Jason Greenwald, Witold Kwiatkowski, and Riccardo Cadalbert for being always ready to troubleshoot any kind of problem in the lab.

I would also like to thank Alexander Sobol, Julien Orts, Dean Strotz, Dhiman Ghosh, Juan Gerez, Natalia Prymaczok, Saroj Rout, Felix Torres, Harindranath Kadavath, Janina Kaderli, Radoslaw Bomba, Piotr Klukowski, Aditya, Peter Güntert, Dzmitry Ashkinadze and Yanick Fleischmann for being amazing colleagues and creating nice and efficient work environment.

I would like to thank Felix Torres, Dhiman Ghosh, Surabhi Choudhary, Nikhilesh Mishra, Saroj Rout, and Dzmitry Ashkinadze for the help in thesis correction.

I would like to thank some of my friends: Ankita Chavan, Shivani Gupta, Dhiman Ghosh, Mathias Steffen, Ursula, Saima Ahmed, Sonu Sahadevan, Jyoti Dhunna and Anu Nair with whom my weekends were relaxing and recharging.
I would like to thank my parents (Ramakar Choudhary, Kiran Choudhary) and God for their unconditional love. I would like to thank my sisters (Sangita, Kadambini, Priyanka, Surabhi) and brother (Ashutosh Bhaskar) for giving me support and courage whenever I needed.
Summary

Aggregation of α-synuclein (α-Syn) protein has been associated in several neurodegenerative diseases including Parkinson’s disease (PD), Alzheimer’s Disease (AD), Dementia, Lewy body, Multiple System Atrophy (MSA) and Amyotrophic lateral sclerosis (ALS). The loss of dopaminergic neurons in the substantia nigra (SN) region is a prominent hallmark of PD. Despite decades of studies on synuclein protein, its function and direct involvement in the disease pathogenesis is not clear. α-Syn is a 140 amino acid residues long unfolded protein. However, it adopts α-helical structure upon binding to lipids and converts into a well-ordered β-sheet rich structure upon aggregation/fibrillation. Amphipathic repeats (KTKEGV) in the N-terminal makes it compatible with lipid binding. α-Syn interaction with lipids has been suggested to be important for playing the essential native biological function (e.g. SNARE complex formation and synaptic plasticity) as well as in the modulation of aggregation and disease propagation.

In the first part of the thesis Chapter 2,3,4, and 5 α-Syn were studied to understand its function and disease. Starting from the modification of the NMR method provided the accurate timescale of the interaction of α-Syn with lipids and its aggregates at physiological condition (Chapter 2). Afterward, the lipoprotein particle was formed by variants of α-Syn to correlate with disease and function (Chapter 3). Further α-Syn monomer
interaction to preformed fibril was studied to understand the propagation and the secondary nucleation process (Chapter 4). The effect of physical and chemical properties of lipids was analyzed on α-Syn interaction. The goal was also to reveal the important residues of α-Syn involved in lipid interaction and also the effect of interaction on the modulation of aggregation. The way to disrupt the membrane and propagate PD pathology was the goal for deeper understanding in Chapter 5.

In the second part of the thesis (Chapter, 6 and 7) Apolipoprotein E (ApoE) and Serum amyloid A (SAA) were studied. Both proteins (SAA and ApoE) belong to the apolipoprotein family and found to be implicated in causing neurodegenerative diseases. More specifically the ApoE4 isoform has a higher risk for Alzheimer’s Disease. SAA is found to be involved in several chronic inflammatory diseases (amyloidosis, atherosclerosis and rheumatoid arthritis). In Chapter 6 ApoE4 homogeneous single population of nanodiscs were prepared for the N-terminal domain and full length, as well as further nanodiscs.

In Chapter 7, a stable functional high-density lipoprotein (HDL) bound SAA was prepared for the structural understanding. SAA was observed to be tightly bound to HDL. The transition from random coil to the secondary structure rich structure was observed for SAA upon HDL binding.
In conclusion, amino acid residues of α-Syn were revealed that is important for lipid interaction and fibrils interaction. The type of interactions was also possible to identify. In the future, these studies can be relevant in modulating α-Syn interaction with lipids and fibrils. Another apolipoprotein (SAA and ApoE) involved in aggregation and neurodegenerative diseases were also studied to reveal its functional structure.
Résumé

L'agrégation de la protéine α-synucléine a été associée à plusieurs maladies neurodégénératives dont la maladie de Parkinson, la maladie d'Alzheimer, la démence à corps de Lewy, l'atrophie à multisystématisée, la démence et la sclérose latérale amyotrophique. La perte des neurones dopaminergiques dans la région substantielle noire est une caractéristique importante de la maladie de Parkinson. Malgré des décennies d'études sur la protéine α-synucléine, sa fonction et son implication directe dans la pathogenèse de la maladie ne sont pas claires. La protéine α-synucléine comprend 140 résidus d'acides aminés et est intrinsèquement desordonnée. Cependant, elle adopte une structure α-helicoïdale au contact de membranes et se convertit en une structure secondaire de type feuillet β lors de l'agrégation/fibrillation. Les répétitions amphipathiques (KTKEGV) dans la région N-terminale la rendent compatible avec la liaison lipidique. L'interaction de la synucléine avec les lipides a été suggérée comme étant importante pour jouer un rôle physiologique (par exemple la formation du complexe SNARE ainsi que la plasticité synaptique) mais aussi pour la modulation de l'agrégation et la propagation des maladies.

Dans la première partie de cette thèse, les chapitres 2, 3, 4 et 5 rapportent l’étude de la fonction de la synucléine en conditions physiologiques et pathologiques. Une première partie relate le travail méthodologique qui a permis d’obtenir avec précision l’échelle de temps de l’interaction de
synucléine avec les lipides et ses agrégats dans des conditions physiologiques par RMN (Chapitre 2). Par la suite, les particules de lipoprotéines ont été formées par des variantes de synucléine pour établir une corrélation avec la maladie et la fonction (chapitre 3). L'interaction des monomères de synucléine avec les fibrilles préformées a été étudiée pour comprendre la propagation et le processus de nucléation secondaire. La nucleation secondaire joue un rôle clé dans la progression de la maladie (chapitre 4). L'effet des propriétés physiques et chimiques des lipides a été analysé lors de l'interaction avec la synucléine. Un objectif supplémentaire était également de révéler les résidus importants de α-Syn impliqués dans l'interaction lipidique, ainsi que l'effet de cette interaction sur la modulation de l'agrégation. Les perturbations de la membrane sont un facteur clé pour la compréhension de la propagation de la pathologie de la maladie de Parkison (chapitre 5).

Dans la deuxième partie de la thèse (chapitres 6 et 7), l'Apolipoprotéine E (ApoE) et l'amyloïde sérique A (SAA) ont été étudiées. Les deux protéines (SAA et ApoE) appartiennent à la famille des apolipoprotéines et sont impliquées dans les maladies neurodégénératives. Plus précisément, la présence de l'isoforme ApoE4 présente un risque plus élevé de développer la maladie d'Alzheimer. La SAA est impliquée dans plusieurs maladies inflammatoires chroniques (amyloïdose, athérosclérose et polyarthrite rhumatoïde). Au chapitre 6, ApoE4 a été préparé sous la forme d'une population homogène de nanodisques pour
le domaine N-terminal ainsi que la protéine entière. D’autres nanodisques ont également été caractérisés.

Au chapitre 7, une protéine SAA, fonctionnelle et stable, liée aux lipoprotéines de haute densité, a été préparé pour la compréhension structurelle. Nous avons observé que la SAA était étroitement liée aux lipoprotéines de haute densité. La transition de la structure aléatoire à la structure riche en structure secondaire a été observée pour la SAA lors de la liaison aux lipoprotéines de haute densité.

En conclusion, les acides aminés chargés et hydrophobes de la synucléine se sont révélés importants pour l’interaction lipidique et l’interaction fibrillaire. Les types d’interactions ont également pu être identifiés. A l’avenir, ces études pourront être pertinentes pour moduler l’interaction de la synucléine avec les lipides et les structures fibrillaires. D’autres apolipoprotéines (SAA et ApoE) impliquées dans l’agrégation et les maladies neurodégénératives ont également été étudiées pour révéler leur structure fonctionnelle.
Table of Contents

ACKNOWLEDGEMENTS ........................................................................................................ 1
SUMMARY .............................................................................................................................. 1
RÉSUMÉ ................................................................................................................................. 4
TABLE OF CONTENTS ........................................................................................................... 7
CHAPTER 1: INTRODUCTION .............................................................................................. 11
SYNUCLEIN (SYN) ............................................................................................................... 11
THE SYN GENES AND PROTEINS ..................................................................................... 11
SYNUCLEINOPATHIES ......................................................................................................... 12
α-SYN DOMAIN, STRUCTURE, AND FUNCTION ................................................................. 13
α-SYN–MEMBRANE INTERACTION ...................................................................................... 18
STRUCTURE OF α-SYN AGGREGATES ............................................................................. 18
MECHANISM OF α-SYN AGGREGATION AND TOXICITY ................................................ 21
MODULATION IN α-SYN AGGREGATION ......................................................................... 23
THERAPEUTIC PERSPECTIVE AND CHALLENGES ....................................................... 24
FAMILIES OF APOLIPOPROTEIN AND RESEMBLE PROTEINS ......................................... 27
APOLIPOPROTEIN E .......................................................................................................... 28
SERUM AMYLOID A (SAA). .................................................................................................. 30
SYN ....................................................................................................................................... 31
MOTIVATION OF THIS THESIS .......................................................................................... 32
CHAPTER 2: 15N TRANSVERSE RELAXATION MEASUREMENTS ARE DETERIORATED BY THE DEUTERIUM ISOTOPE EFFECT ON 15N RESULTING FROM SOLVENT EXCHANGE ............................................................................................... 34
ABSTRACT .......................................................................................................................... 34
INTRODUCTION .................................................................................................................... 35
RESULTS AND DISCUSSION ............................................................................................... 38
PULSE SEQUENCE FOR THE MEASUREMENT OF 15N R2 RELAXATION ............................. 38
15N R2 RELAXATION CONTRIBUTION BY THE DEUTERIUM ISOTOPE EFFECT ................ 41
THE USE OF AN EXTERNAL DEUTERIUM LOCK FOR 15N R2 RELAXATION MEASUREMENTS ......................................................................................................................... 46
# Table of Contents

**The effect of sample internal D$_2$O on the $^{15}$N R$_2$ relaxation measurements on the folded protein domain PDZ2** ................................................................. 48  
**The impact of internal D$_2$O on CPMG-based relaxation dispersion experiments** ....... 48  
**Discussion** ........................................................................................................... 53  
**Conclusion** ....................................................................................................... 56  
**Material and Methods** .................................................................................... 57  
**Supporting Information** .................................................................................. 62  
**CHAPTER 3: HIGH-DENSITY LIPOPROTEIN-LIKE PARTICLE FORMATION OF SYNUCLEIN VARIANTS** ................................................................. 67  
**Abstract** ........................................................................................................... 67  
**Introduction** ..................................................................................................... 68  
**Results and Discussion** ................................................................................... 70  
**Sequence similarity of the synuclein family with the apolipoprotein A-1** ............. 70  
**Lipoprotein particle formation of the synuclein family and the familial α-synuclein mutants** ............................................................................................. 70  
**Morphology of the synuclein lipoprotein particles** ............................................. 71  
**Discussion** ....................................................................................................... 76  
**Material and Methods** .................................................................................... 79  
**Supplementary Information** ............................................................................ 80  
**Synuclein nanodiscs formation by various lipids** .............................................. 80  
**CHAPTER 4: STRUCTURAL INSIGHTS INTO THE MONOMER-FIBRIL INTERACTION OF α-SYNUCLEIN** ................................................................. 81  
**Abstract** ........................................................................................................... 81  
**Introduction** ..................................................................................................... 82  
**Results and Discussion** ................................................................................... 84  
**Faster aggregation of α-Syn at pH 6.0 than 7.0 in the presence of pre-formed fibrils** 84  
**Transient interaction of α-Syn monomers with fibrils** ....................................... 85  
**Interaction of monomer with the fibrils observed by EM** ................................... 88  
**Mutagenesis-based structural investigations of the interaction of α-Syn monomers with the fibrils** ................................................................................. 88  
**Conclusion** ....................................................................................................... 91  
**Material and Methods** .................................................................................... 92  
**Supplementary Information** ............................................................................ 95
CHAPTER 5: α-SYNUCLEIN INTERACTION WITH VARIOUS LIPIDS CONTAINING NANODISCS AND ITS EFFECT ON AGGREGATION MODULATION ........................................................................................................ 97
ABSTRACT .................................................................................................................................. 97
INTRODUCTION ............................................................................................................................... 99
RESULTS AND DISCUSSION ......................................................................................................... 99
STRUCTURAL CHANGE OF α-SYN UPON LIPID INTERACTION ......................................................... 99
CHARGE OF LIPIDS IS ESSENTIAL FOR INTERACTION WITH α-SYN ............................................... 100
MODULATION OF α-SYN AGGREGATION IN THE PRESENCE OF NANODISCS ............................... 101
UPTAKE OF NANODISCS BY CELLS ............................................................................................... 101
THE PHASE CHANGE OF LIPIDS ALTERS THE NANODISC INTERACTION WITH α-SYN ................... 102
ATTEMPT TO OBSERVE THE CHANGE OF SECONDARY STRUCTURE .......................................... 107
α-SYN BOUND TO NANODISCS DIFFUSES SIMILARLY TO NANODISCS ALONE .............................. 109
INTERACTION OF α-SYN WITH 10% GM3 AND 90% DOPC LIPIDS CONTAINING NANODISCS ... 109
CONCLUSION .................................................................................................................................. 111
MATERIAL AND METHODS ........................................................................................................ 113
CHAPTER 6: STRUCTURAL AND MORPHOLOGICAL STUDIES OF APOE4 DIFFERENT CONSTRUCT BOTH IN FREE AND LIPID-BOUND FORM ........................................................................................................ 117
ABSTRACT .................................................................................................................................. 117
INTRODUCTION ............................................................................................................................... 118
RESULTS AND DISCUSSION ......................................................................................................... 119
SEQUENCE SIMILARITY WITHIN APOLIPOPROTEINS .................................................................... 119
FORMATION OF NANODISCS BY FULL-LENGTH APOE4 (1-299) AND N-TERMINAL DOMAIN (1-191) ........................................................................................................................................ 120
NANODISCS CONTAIN TWO SUBUNITS OF APOE PROTEIN ......................................................... 122
NANODISCS FORMATION BY NT DOMAIN WITHOUT MAJOR STRUCTURAL CHANGES ............... 123
NT DOMAIN AND CT DOMAIN ARE INVOLVED IN NANODISCS FORMATION ............................. 124
FORMATION OF NANODISCS BY WILDTYPE FULL LENGTHAPOE4 ............................................ 126
AN ATTEMPT TO DETERMINE HIGH-RESOLUTION STRUCTURAL DETAILS OF FULL-LENGTH APOE4 NANODISCS ........................................................................................................ 127
DIFFUSION COEFFICIENT REVEAL BIGGER SIZE WAS THE LIMITING FACTOR FOR STRUCTURAL DETERMINATION .................................................................................................................. 131
### Table of Contents

**ApoE4 Structure with Lipids (In Nanodiscs Structure) Versus ApoE4 in Detergent (Sodium Cholate)** ................................................................. 132
**Conclusion** ........................................................................................................ 133
**Material and Methods** ....................................................................................... 135

**CHAPTER 7: STRUCTURAL CHARACTERIZATION OF RECOMBINANT MURINE SERUM AMYLOID A (SAA) ISOFORM 1.1** ........................................... 140
**Abstract** ........................................................................................................... 140
**Introduction** ...................................................................................................... 141
**Results and Discussion** .................................................................................... 142

- Large-scale expression of SAA .......................................................................... 142
- NMR spectra of SAA monomer ......................................................................... 143
- Binding of monomer SAA to HDL ................................................................... 143
- Conditions tried for resolved NMR spectrum of HDL bound SAA ................ 145
- Spectral quality with temperature and time....................................................... 145
- SAA is bound to HDL .......................................................................................... 149
- SAA structural change from unfolded monomer to HDL bound and with detergent .............................................................................................................. 149
**Conclusion** ....................................................................................................... 151
**Material and Methods** ....................................................................................... 152

**CHAPTER 8: CONCLUSION AND OUTLOOK** .................................................. 153
**Investigation of Synuclein-lipid Interaction** ...................................................... 153
**Investigation of Synuclein-fibril Interaction** ...................................................... 155
**Investigation of the Proteins from the Apolipoprotein Family** ......................... 155
**Conclusion** ....................................................................................................... 157

**List of Abbreviations** ......................................................................................... 158
**Bibliography** ....................................................................................................... 160
Chapter 1

Introduction

Synuclein (Syn)

The Syn genes and proteins
The Syn genes SNCA, SNCB and SNCG map on human chromosome 4, 5 and 10 code for the respective three proteins - α-Syn, β-Syn and γ-Syn (1). The SNCA gene belongs to the family of PARK gene (Parkinson’s disease (PD) related genes) and is located on the long arm (q) of chromosome 4 at position 22, however the SNCB and SNCG genes which code for β-Syn and γ-Syn, respectively, do not belong to the PARK gene family. Duplication, triplication and seven distinct point mutations in the SNCA gene are known to cause autosomal dominant PD. These mutations include A53T (2, 3), A30P (4), E46K (5), G51D (6), H50Q (7), A53E (8) and A53V (9). SNCA is linked to both rare familial and sporadic PD (10) and diseases like multiple system atrophy (MSA) and Dementia (11, 12).

The synuclein protein was found to be mainly localized in presynaptic nerve terminals and in the nuclear envelope. Synucleins are exclusively found in vertebrates since no homologous are identified in *E. coli, S. cevevisiae, C. elegans, or D. melanogaster*. The three isoforms of Syn (α-
Syn, β-Syn and γ-Syn) differ in the amino acid sequences (1). The first report of Syn was in *Torpedo californica* (13) which was found to be similar to human γ-Syn (14) while in humans, the NAC region (hydrophobic stretch, 61-95) of Syn was first detected in the amyloid plaques of Alzheimer’s disease (AD) patients although it was not a part of Aβ and therefore was named as non Aβ component (NAC) (15)). α-Syn has been a protein of interest due to its role in PD and other synucleinopathies, including several neurodegenerative diseases AD, dementia Lewy body (DLB), multiple system atrophy (MSA), amyotrophic lateral sclerosis (ALS) (16) and many more. Physiological concentration of α-Syn in the brain is about 0.5-1.0% of the total protein (17) and is found to be localized in the pre-synaptic nerve terminals (13) in the proximity of synaptic vesicles (18).

**Synucleinopathies**

Aggregation of α-Syn is implicated in various neurodegenerative diseases which are termed as α-Synucleinopathies. α-Synucleinopathies include the following neurodegenerative diseases Parkinson’s disease (PD), Alzheimer’s Disease (AD), Dementia, Lewy body (DLB), Multiple System Atrophy (MSA), and Amyotrophic lateral sclerosis (ALS) (16). The most common disease associated with α-Synucleinopathies is PD followed by DLB and different symptoms are associated with these three different kinds of α- Synucleinopathies (Figure 1.1).

i) **PD**: is the second most common neurodegenerative disorder, after AD. Motor clinical symptoms includes bradykinesia, rigidity, resting
tremor and gait instability (19), non-motor symptoms include hallucinations, olfactory disturbances, sleep disturbances, depression and dementia (20). Loss of dopaminergic neurons in the substantia nigra (SN) is the main pathological cause, other than that the presence of α-Syn-positive inclusions in the cytoplasm and axons are also related α-Syn pathologies.

ii) DLB: Dementia is the predominant symptom in DLB (21). Pathological symptoms include Lewy body (LB) deposition in the SN and amygdale (22, 23).

iii) MSA: MSA has been classified either as, MSA-P or MSA-C based on the differences in the clinical symptoms. MSA-P symptoms are similar to PD whereas MSA-C symptoms are mainly related to a cerebellar syndrome (24).

**α-Syn domain, structure, and function**

α-Syn is a small (140 amino acid, ~14 kDa), soluble and intrinsically disordered protein (25). Its low intrinsic hydrophobicity and high net charge (-8.8) at neutral pH makes it intrinsically disordered (26), however, it adopts helical structure upon binding to membranes and micelles (27, 28). α-Syn consists of three domains - N-terminal domain, NAC domain and C-terminal domain (Figure 1.2).

**N-terminal domain**

It comprises a non-canonical 11 amino acid repeat with a fairly conserved consensus motif, KTKE(Q)GV (29).
Chapter 1

<table>
<thead>
<tr>
<th>Clinical</th>
<th>Parkinson’s disease (PD)</th>
<th>Dementia with Lewy bodies (DLB)</th>
<th>Multiple System Atrophy (MSA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of Onset (years)</td>
<td>~60</td>
<td>~70</td>
<td>~60</td>
</tr>
<tr>
<td>Duration (years)</td>
<td>~12</td>
<td>~5</td>
<td>6-9</td>
</tr>
<tr>
<td>Motor signs</td>
<td>All</td>
<td>Some</td>
<td>All</td>
</tr>
<tr>
<td>L-Dopa responsive</td>
<td>Yes</td>
<td>Variable</td>
<td>No</td>
</tr>
<tr>
<td>Non-motor signs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autonomic</td>
<td>Late</td>
<td>Late</td>
<td>Early</td>
</tr>
<tr>
<td>Hallucinations</td>
<td>Late</td>
<td>Early</td>
<td>No</td>
</tr>
<tr>
<td>Olfactory dysfunction</td>
<td>Early</td>
<td>Early</td>
<td>No</td>
</tr>
<tr>
<td>Dementia</td>
<td>Late</td>
<td>Early</td>
<td>No</td>
</tr>
<tr>
<td>Main synucleinopathies</td>
<td>NCI1s</td>
<td>NCI1s</td>
<td>GCI1s</td>
</tr>
</tbody>
</table>

**Figure 1.1: α-Synucleinopathies (PD, DLB, and MSA).** Figure adapted from McCann et al., (2014).

The N terminal domain of α-Syn is highly conserved among its variants as well as across different species. When bound to either negatively charged lipids containing small unilamellar vesicles (SUV), or micelles containing negatively charged head groups, the N-terminus region converts from random coil to alpha-helix rich structure (27, 30). α-Syn binding to lipids is facilitated through the conserved repeat motif. Interestingly, all seven
known diseases linked familial mutations in PD patients are all located in the N-terminus region (A30P, E46K, H50Q, G51D and A53T/V/E).

**Figure 1.2: Domain structure of α-Syn.** Positively charged amphipathic N-terminal domain followed by hydrophobic NAC (non-amyloid component) domain and then highly negatively charged C-terminal domain in blue, yellow and red colors respectively. Seven imperfect repeats of KTKE(Q)GV are shown in green. Five mutations sites (A30P, E46K, H50Q, G51D, and A53T/V/E) which were found in the patients of early-onset PD are shown in the N-terminal domain and highlighted by red arrows.

**NAC (Non-Aβ component) domain**

This domain was first identified in the amyloid fibrils of AD. It is the second major component of amyloid plaques in AD patients (15). This region is rich with hydrophobic residues and alone capable of forming β-sheet rich protein aggregates (31). The NAC domain is considered as aggregation causing region; deletion and mutations of amino acid residues from the NAC domain decrease its aggregation propensity both in vivo. However, N-terminal and C-terminal are considered as aggregation preventing domains (31,32).
**C-terminal domain**

This domain is polar and consists of many acidic amino acid residues. The C-terminal domain is not very conserved in length and sequence across different species, and different variants of Syn. The C-terminal domain plays a significant role in the regulation of nuclear localization, contains most of the post-translational modifications and is involved in interactions with small molecules, proteins, and ligands (25). The C-terminal domain contains seven phosphorylation sites (33). 90% of the α-Syn was found to be phosphorylated at S129 in LB, which suggests S129 phosphorylation as a hallmark of PD. α-Syn has a high propensity to get truncated at several C-terminal cleavage sites 101, 103, 115, 119, 122, 133, and 135 (34, 35). The removal of this region increases the aggregation propensity of α-Syn because this domain appears to serve as a chaperone-like entity (36). At neutral pH, the C-terminal domain forms long-range interactions (electrostatic interactions) with the N-terminal domain. Long-range interactions stabilize the monomer conformation and thereby disruption of these interactions by some external factors like polyamines induces misfolding and aggregation of α-Syn (37, 38).

α-Syn was shown to exist as the natively unfolded monomer. The unfolded structure was confirmed by Circular dichroism (CD) and Fourier transform infrared spectroscopy (FTIR) (26). However, at neutral pH, it was considered as more compact (radius of gyration, Rg 40 Å) than the expected value for completely unfolded protein (radius of gyration, Rg 52 Å) (26). On the other hand, at relatively low pH (2.5), it gets more compact
than at pH 7.0 (39). The compact structure of α-Syn has also been confirmed by NMR (40), EPR (41), and MD simulations (42). Long-range interactions between the C-terminal and N-terminal region have been identified by PRE (Paramagnetic relaxation enhancement) and analogous PRE (43, 44) NMR experiments. α-Syn’s native structure was considered and proven to belong to intrinsically disordered proteins (IDPs) since long but later it became in controversy after purification of α-Syn by using non-denaturing conditions from neuronal and non-neuronal cells. In these reports, α-Syn was found to be tetrameric in nature with majorly α-helical conformation as a ~58 kDa molecular weight. In vitro crosslinking, STEM and ultracentrifugation were used to confirm the tetrameric structure (45). By contrast, other studies although showed natively unfolded conformation in vivo (46, 47).

The essential physiological function of α-Syn is not clearly known yet. Several roles have been suggested on the basis of its localization and its interaction with some molecules. α-Syn is found to be expressed in both the central and peripheral nervous systems which suggest a role in neuronal function, synaptic plasticity, in learning and in the regulation of neurotransmitter release (17, 48, 49). α-Syn was found to be localized in the presynaptic terminals (50), describing its potential role in synaptic vesicle trafficking and its interaction with membrane suggested a role in the regulation of transmitter release (51, 52). Although normal survival of Syn (α-, β-, γ-) knockout mice suggests that Syn is not essential for neurotransmitter release (53). α-Syn also helps in the formation and
maintenance of the SNARE complex. Both the N and C-terminal regions of α-Syn act as chaperones for other proteins (54).

**α-Syn-membrane interaction**

α-Syn has been demonstrated to interact with lipids and membrane both for its native function, to induce aggregation and its propagation. α-Syn transforms its unfolded structure to α-helical secondary structure upon binding with negatively charged lipids containing lipid bilayer or SUV (27). Amphipathic repeats (KTKEGV) in the N-terminal makes it compatible with lipid binding. α-Syn also makes lipoprotein particle where 8-10 α-Syn monomer interacts with lipids, a mutation in single amino acid residues was observed to change the amount of lipoprotein particle formation (55, 56). When α-Syn binds to phospholipids it flattens the membrane curvature. Different familial mutants show the difference in the extent of membrane interaction. A30P was found to have less affinity for membrane however A53T and E46K mutation leads to increase the affinity of membrane binding (57). Lipid binding capacity and aggregation propensities of α-Syn depend on the physical and chemical properties (Phase, K_D, Tm, acyl chain length) of lipids (58).

**Structure of α-Syn aggregates**

α-Syn oligomers were found to be in many different secondary structures from disordered (59) to α-helical (60) or β-sheet (61), depending upon the aggregation/growth condition. Morphology, secondary structure, twist, and fold of the final fibrils were also found to be highly dependent on the
aggregation conditions, Syn variants and post-translational modifications. Aggregation of α-Syn was found to be salt-dependent in vitro and the presence of salt not only shortens the aggregation but also changes the morphology of aggregates, toxicities, seeding properties and interactions with the membrane (62, 63). Structural characterization of fibrils was revealed mostly using ssNMR (Solid-State NMR) and hydrogen/deuterium exchange NMR (HX-NMR). β-sheet rich core region (38-95), flexible C-terminal (96-140) and flexible the N-terminal region (1-22) was observed by ssNMR. A53T mutant was found to have more extended β-sheet core region (64)). Later, the core region of fibrils contains five β-strands (residues 30-110) were revealed by using HX-NMR, ssNMR and cryo-electron microscopy (65). The core region of oligomers was found to be similar to the core region fibrils (66). However, in some studies the core region of oligomer was found to be different and further suggested that structural difference can lead to the difference in toxicity (67). Later, β-sheet (parallel, in-register) structures were found to be similar in mouse and human α-Syn (68). Recently four different polymorphs of α-Syn were solved by CryoEM (68-70). The structures also reveal the possible reasons for early-onset diseases because of the presence of these mutants on the hydrophobic dimer interface (except A30P). Mutations H50Q and E46K compromise salt bridges and instead of creating charge repulsion which can lead to destabilization of the hydrophobic protofibrils dimer. Mutation G51D loss the flexibility, hydrophobicity, and insertion of a steric bulk at the edge of dimer interface which may lead to
destabilization of the Greek key structure. Mutations A53T/E lead to an increase hydrophilic residue at the interface which is proposed to weaken the hydrophobic packing at the interface. Mutation A53V leads to an increase in the hydrophobicity but insertion of bulky residues might cause steric hindrance at the interface. Overall all these mutations (E46K, H50Q, G51D, and A53T/E/V) leads likely to a destabilization on the hydrophobic zipper interface formed between two protofibrils which might be leading to a slower mature fibril formation and to long-lived toxic oligomers. The A30P mutation does not fall in the interface/core region as shown in polymorph1a (Figure 1.3).

However, in the new polymorph 2a and 2b, a steric zipper interface is between V16-E20 and I88-A91, which bring A30P in a more defined structure as compared to polymorph 1a and 1b. This might suggest the importance of A30P mutation in the early-onset familial mutations (Guerrero et al., 2019). Other possible explanation for toxicity of A30P can be a defect in the formation of the native fold or loss of functional interaction with membrane, other important proteins or small molecules. Polymorph 1a is also called a rod-like structure. It contains a Greek key topology and pre-NAC residues seem to be involved in the stabilization of the protofibril dimer interface. In contrast to polymorph1a, polymorph 1b called twister structure reveal a β-arch motif, here protofilaments are packed around the NAC residues, although both polymorph 1a and 1b contain two protofilaments with rotation symmetry about fibril axis. In
contrast, type 2 polymorphs (2a and 2b) contain a hydrophobic cleft, which is stabilized via intermolecular salt bridges (shown in Figure 1.4).

**Figure 1.3: 3D structure of a single layer within a mature α-Syn fibril.** Based on the CryoEM structure published by Li YW et al. 2018) displaying the formation of the ‘Greek Key’ topology with rotational symmetry about the axis of the fibril. The early-onset mutations (E46K, H50Q, G51D, A53T/E/V) are highlighted (pink) in addition to three key electrostatic interactions that are perturbed in early onset PD (K58-E61, E46-K80 and K45/H50-E57). Figure Adapted from (Meade RM et al 2019).

**Mechanism of α-Syn aggregation and toxicity**

Kinetic studies of α-Syn aggregation process reveal its polymerization to be nucleation dependent polymerization (NDP) as it has a lag phase which was observed to be dependent on the amount of seed concentration and it needs the critical initial concentration of monomer to aggregates (71). Acceleration in the formation of amyloid fibrils by the addition of
preformed aggregates (seeds) was explained by two mechanisms (i) elongation (addition of monomer at the terminal) and secondary

conditions lead to different rates for primary nucleation, elongation, secondary nucleation, and fragmentation. At relatively low pH (pH below 6) rate of secondary nucleation was found to be dominated as compared to a neutral pH and the rate of elongation was dominated over secondary nucleation.

α-Syn aggregates disrupt the membrane and cause toxicity (73). Extraction of lipids from the bilayer and accumulation of large clumps of lipids and fibrils might be the mechanism of toxicity (74). Oligomers were
Introduction

considered to be more toxic than fibrils both in vivo and in vitro (75-78). ssNMR has revealed oligomer contains distinct structures from both monomer and fibrils. SAXS has revealed wreath shapes of oligomer which makes it easier to make a channel in the membrane and create toxicity (79).

α-Syn was also found in biological fluids, blood plasma, and culture media which suggests it is also secreted extracellularly (80, 81). Monomer, as well as aggregates of α-Syn were found to be secreted by endocytic and non-classical exocytic pathways (Lee et al., 2005). α-Syn can also be transmitted across neuronal cells (82). However, there are several mechanisms (exocytosis, exosomal release, and direct penetration) which can transmit aggregates from the cell to cell and/or neurons (83-85), α-Syn was found to have a prion-like propagation when brain homogenates were injected (Intracerebral) from infected mice into a healthy mouse. It leads to inclusion formation as well as PD symptoms. It was also observed that the aggregates can cross the species barrier when the aggregate seeds of human were injected into a healthy mouse; mice develop PD symptoms (86, 87).

Modulation in α-Syn aggregation
Rate of aggregate formation depends on many factors such as post-translational modifications (PTM) (Table 1.1) different disease-linked mutants (88), change in pH and temperature (62, 89, 90), presence of anions (91), presence of small unilamellar vesicles (SUVs) (92), C-terminal truncation of α-Syn (38, 93, 94), presence of metals (95, 96), Dopamine
and Dequalinium (97), Chaperones (98, 99), Mutations in the NAC-region (100), fatty acids, anionic detergents (101), trifluoroethanol (101), SDS

Table 1.1: Effect of PTM on protein aggregation and degradation. (Table adapted from Zhang et al., 2019)

(102), arginine (103), presence and absence of air-water interface (104, 105) presence of secondary nucleation and fragmentation (72).

**Therapeutic perspective and challenges**

Aggregation of Syn is implicated in causing Parkinson’s Diseases (PD) and several other neurogenerative diseases. The ultimate goal is to understand the diseases and causes in order to discover the cure. Various methods have been described (106, 107); to reduce the toxicity of the α-Syn aggregates. Figure 1.5 shows a few methods to target PD. Therapeutic trials were tried at many different levels. First, the aggregation can be stopped by lowering the Syn monomer expression at the transcriptional level. Furthermore, before translation, it can be
modified (PTM) in order to reduce the aggregation and toxicity. Another way can be to remove the excess monomer, misfolded and aggregate Syn by degradation pathway more active. Immunotherapy was found to be a very promising therapeutic approach to target extracellular degradation of α-Syn. Further design of small molecules can be possible to prevent the monomer addition to aggregates and masking its toxicity once the structure can be understood for aggregates and monomer.

Reduction in α-Syn expression could be decreased by using small interfering RNA (siRNA), micro RNA (miRNA), antisense or other transcription inhibitors. Infusion of siRNA decreased the α-Syn around 40-50% compared to not treated siRNA. Infusion of short hairpin α-Syn inhibited the expression of endogenous α-Syn around 35% (108, 109).

Another possible target was observed to be β2-Adrenoreceptor (β2AR), which reduces the transcription of the SNCA gene (110).

Protein aggregation and toxicity can also be regulated by changing the PTM (e.g. phosphorylation, nitration, oxidation, and C-terminal cleavage). α-Syn uses autophagy-lysosomal pathways for its degradation, GBA enzyme is a lysosomal enzyme which is responsible for the metabolism of Syn. The mutation of GBA gene is considered as a risk factor of PD. Autophagy has been also induced by rapamycin.

Elevation in the dopamine levels was found to be a treatment strategy for PD (73). Some of the ligands were able to inhibit the aggregation of α-Syn by binding to the NAC-region of the α-Syn (111). Heat-shock proteins could also be used for PD treatment because when α-Syn was co-
expressed with the heat shock proteins, the toxicity was found to be lower in *D. melanogaster* (112).

Despite the many years of long research, there are still many challenges remaining with the α-Syn therapeutics. The most important reason might be the lack of knowledge about its exact essential function and heterogeneity among α-Syn structures. Variation in the symptoms, stage of diseases, propagation and aggregates structure can make it even more difficult to target. Another challenge is the lack of diagnosis at the early stage which is limited by the lack of potential biomarkers.

*Figure 1.5: Different steps to target α-Syn in the molecular pathogenesis.*

1. Reducing α-Syn synthesis, 2. preventing α-Syn aggregation, 3. promoting degradation of intracellular α-Syn aggregates, 4. increasing degradation of extracellular α-Syn. (Figure adapted from Fernández-Valle T et al., 2019)
Families of apolipoprotein and resemble proteins

Apolipoproteins

Lipids such as cholesterol and triglycerides are not soluble in water, so it needs protein for its circulation in body fluid. These proteins are called lipoprotein. Lipoprotein plays an important role in the absorption and transport of lipids (dietary) through body fluids to (and from) various organs. The structure of lipoprotein consist of a hydrophobic central core contains mostly non-polar lipids (such as cholesterol esters and triglycerides) further it is surrounded by hydrophilic lipids (such as phospholipids and non-esterified cholesterol) and then apolipoproteins. Plasma lipoproteins have been divided into five major classes on the basis of their densities. For these liposomes densities and size are inversely related.

Apolipoproteins are the protein part of the lipoprotein. It has four major functions including 1) structural role, 2) ligands receptors for lipoprotein, 3) guiding the formation of lipoproteins, and 4) modulator of enzymes involved in the metabolism of lipoproteins. Out of all Apolipoproteins, only ApoBs (including apoB-100, ApoB-48) were observed to be insoluble (113). ApoB-100 is also unique because of relatively a high β-sheet content. Rest all Apolipoproteins (ApoA, ApoC, and ApoE) are soluble and rich in helical conformation. Various Apolipoproteins are summarized in Table 1.2, with its primary source of secretion and function.
Apolipoprotein E

ApoE is a soluble apolipoprotein plays an important role in lipid metabolism. ApoE is a key regulator in maintaining the plasma and tissue lipid content in both plasma and central nervous system by binding to the cell surface lipoprotein receptor. Other than lipid metabolism it plays also a crucial role in neurobiology. ApoE has 3 major isoforms (ApoE2, ApoE3, ApoE4) which differ by only 2 residues in sequence, but it drastically changes the risk of developing atherosclerosis and neurodegenerative diseases (114). ApoE is also the most prominent genetic marker of Alzheimer’s disease (AD) and several other neurological conditions. Out of three different isoforms, a person containing ApoE4 homozygous has the highest risk of developing AD whereas ApoE2 isoforms are the safest and ApoE3 is with intermediate risk. ApoE is also involved in many other diseases such as Parkinson’s disease (PD), cardiovascular diseases (CVD), multiple sclerosis (MS), type 2 diabetes mellitus (T2DM), vascular dementia (VD), and ischemic (occlusive) stroke (IS) (115). Structure difference across these isoforms is expected to explain the drastic difference in the risk of diseases. Structure determination was not possible because of the formation of aggregates by full-length ApoE even at very low concentration (116). Since the C-terminal domain was found to be the reason for ApoE aggregation, the structure of the N-terminal domain alone of ApoE3 was solved and was found to contain a four-helical bundle (117, 118). Structural information about the C-terminal domain was obtained by prediction software, which reveals it to be highly helical which was also in agreement with Circular Dichroism (119). The full-
length ApoE3 protein was found to be stable when five bulky hydrophobic residues from the C-terminal domain were mutated to less bulky hydrophobic and hydrophilic residues. Furthermore, full length monomeric ApoE3 structure was determined by NMR (120), but it is still not clear how ApoE4, which is different from ApoE3 by one residue (C112R) is a significantly higher AD risk factor.

<table>
<thead>
<tr>
<th>Apolipoprotein</th>
<th>MW</th>
<th>Primary Source</th>
<th>Lipoprotein Association</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apo A-I</td>
<td>28,000</td>
<td>Liver, Intestine</td>
<td>HDL, chylomicrons</td>
<td>Structural protein for HDL, Activates LCAT</td>
</tr>
<tr>
<td>Apo A-II</td>
<td>17,000</td>
<td>Liver</td>
<td>HDL, chylomicrons</td>
<td>Structural protein for HDL, Activates hepatic lipase</td>
</tr>
<tr>
<td>Apo A-IV</td>
<td>45,000</td>
<td>Intestine</td>
<td>HDL, chylomicrons</td>
<td>Unknown</td>
</tr>
<tr>
<td>Apo A-V</td>
<td>39,000</td>
<td>Liver</td>
<td>VLDL, chylomicrons, HDL</td>
<td>Promotes LPL mediated TG lipolysis</td>
</tr>
<tr>
<td>Apo B-48</td>
<td>241,000</td>
<td>Intestine</td>
<td>Chylomicrons</td>
<td>Structural protein for chylomicrons</td>
</tr>
<tr>
<td>Apo B-100</td>
<td>512,000</td>
<td>Liver</td>
<td>VLDL, IDL, LDL, Lp (a)</td>
<td>Structural protein, Ligand for LDL receptor</td>
</tr>
<tr>
<td>Apo C-I</td>
<td>6,600</td>
<td>Liver</td>
<td>Chylomicrons, VLDL, HDL</td>
<td>Activates LCAT</td>
</tr>
<tr>
<td>Apo C-II</td>
<td>8,800</td>
<td>Liver</td>
<td>Chylomicrons, VLDL, HDL</td>
<td>Co-factor for LPL</td>
</tr>
<tr>
<td>Apo C-II</td>
<td>8,800</td>
<td>Liver</td>
<td>Chylomicrons, VLDL, HDL</td>
<td>Inhibits LPL and uptake of lipoproteins</td>
</tr>
<tr>
<td>Apo E</td>
<td>34,000</td>
<td>Liver</td>
<td>Chylomicron remnants, IDL, HDL</td>
<td>Ligand for LDL receptor</td>
</tr>
<tr>
<td>Apo (a)</td>
<td>250,000-800,000</td>
<td>Liver</td>
<td>Lp (a)</td>
<td>Inhibits plasminogen activation</td>
</tr>
</tbody>
</table>

*Table 1.2: Various class of apolipoproteins. (Table adapted from the book by Kenneth R Feingold 2018, Introduction to lipids and lipoprotein)*

To improve the understanding of the different behavior of ApoE isoforms a requirement is the 3D structures of wild-type ApoE4, ApoE3, and ApoE2 in their functional forms i.e. lipoprotein. Since the C-terminal is the hydrophobic region, which is considered to be important for lipid binding,
mutations in the C-terminal might deviate the structure from physiological structure.

**Serum amyloid A (SAA)**

SAA is a small (104 amino acids) proteins belongs to the family of apolipoproteins. The SAA name was chosen because of its presence in serum and because it was identified as a precursor of amyloid disease. The SAA sequence is highly conserved across mammals and birds \(^{(121, 122)}\). The normal concentration of SAA is 20-50 µg/ml in blood but it increases up to 1000-fold upon acute phase response (APR). Amino-terminal fragments of SAA are highly aggregation-prone and make insoluble fibrils \(^{(123)}\). Essential physiological role is not known although it has been implicated in lipid metabolism, cell-cell communication, and feedback in inflammation \(^{(124)}\). SAA protein lacks stability hence understanding the structure by NMR or crystallization becomes challenging. Based on the algorithm it was proposed to contain two α-helices and a β-sheet however when SAA monomer structure was solved by multi-wavelength anomalous dispersion it was found to contain four α-helices in a cone-shaped array and no β-sheet structure \(^{(125)}\). Later SAA was possible to purify without denaturation and structure was found to be hexamer formed by two identical trimers \(^{(126)}\). Based on its structure, HDL binding sites are predicted, but the actual HDL bound structure is not known. Amphipathic region of SAA and 95% of the circulating SAA is in the HDL fraction. This suggests its function is to be involved in lipid metabolism and cholesterol transport \(^{(123, 125)}\).
Association of SAA to HDL also leads to prevent the formation of aggregates and fragmentation.

To understand the physiological function of SAA and prevent the formation of toxic aggregates it is important to have the structural information of HDL bound SAA.

**Syn**

Recent studies of Syn started focusing on revealing the function of α-Syn which is predicted to be involved in synaptic functions and lipid metabolism, α, β, γ–Syn share structural resemblance to apolipoproteins (73, 127) other than amphipathic repeats it also contains a hydrophobic core and a cholesterol-binding domain at residues 67–78 (128).

Although α-Syn lacks a N-terminal signal peptide, still it can be secreted to extracellular space (body fluids, cerebrospinal fluid, and blood plasma) (129) where it can interact with the plasmatic apolipoproteins. A complex between α-Syn and ApoJ is also found in the intermediate-density lipoprotein (IDL) (130)). ApoE was also found to modulate α-Syn aggregation in a concentration-dependent manner (131).

α-Syn and it variant were able to form nanoparticles close to HDL with a diameter of approx. 25nm composed of eight α-Syn monomers and lipids. α-Syn nanoparticles were structurally characterized and found to be in a α-helical conformation (1-90) and a flexible C-terminal flexible domain (91-140). Cross-linking studies also showed a similar molecular weight species inside the cells that could lead to find the physiological function
of Syn as HDL (55, 56). α -Syn-related nanoparticles were suggested to play an important role in lipid transport in healthy conditions. α-Syn properties change upon binding to apolipoprotein metabolism (132), which can lead to create defects in lipid metabolism, aggregation formation, and propagation. α -Syn interaction to low-density lipoprotein might play a role in crossing the blood-brain barrier (133), which may reveal in further propagation of α -Syn and prion-like behavior. In a recent study, a crowded environment of membranes was found when the LB was characterized by light and electron microscopy from Parkinson’s disease patients (134), this again suggests the involvement of lipids in α -Syn related neurogenerative diseases.

Motivation of this thesis

Aggregation of the α-synuclein (α-Syn) protein has been implicated in several neurodegenerative diseases including Parkinson’s disease (PD), Alzheimer’s disease (AD), Dementia, Lewy body, Multiple System Atrophy (MSA) and Amyotrophic lateral sclerosis (ALS). The loss of dopaminergic neurons in the substantia nigra (SN) region is the most prominent hallmark of PD (135). Physiological and disease linked α-Syn conformation and function are still poorly understood after decades of studies. Although its essential physiological function is not known but some studies suggest α-Syn interact with lipids and membrane both for its native structure, function, to modulate aggregation and its propagation. In this thesis, the objectives were to find the basic important questions
about α-Syn native conformation, lipid interaction, aggregates formation and interconnection between them. Important goals were the following: Finding NMR methods that reveal the Syn interaction (residue level information) with lipids and its aggregates at physiological pH and temperature. Furthermore, How Syn forms a lipoprotein like particles and how the amount of formation of lipoprotein depends on the lipids and different variants of Syn?

How Syn interact with HDL? What properties of lipids are essential for binding and how they affect the formation of aggregates?

In order to understand the fast propagation of the disease: How Syn monomer interacts with preformed aggregates?

Another objective was to understand the structural detail of functional lipid-bound conformation of Apolipoproteins (ApoE) and Serum amyloid A (SAA) which are also found to be linked with the formation of aggregation and neurodegenerative diseases. ApoE and SAA are aggregation-prone less stable protein. The work started with finding the conditions where monomer can be stable, lipid ratio optimization for the formation of homogenous nanodiscs population and further structural characterization for monomer and lipid-bound nanodiscs.
Chapter 2

$^{15}$N transverse relaxation measurements are deteriorated by the deuterium isotope effect on $^{15}$N resulting from solvent exchange

Abstract

$^{15}$N $R_2$ relaxation measurements are key for the elucidation of the dynamics of both folded and intrinsically disordered proteins (IDPs). Here we show, on the example of the intrinsically disordered protein α-synuclein and the folded domain PDZ2, that at physiological pH and near
physiological temperatures amide – water exchange can severely skew Hahn-echo based $^{15}$N $R_2$ relaxation measurements as well as low frequency data points in CPMG relaxation dispersion experiments. The nature thereof is the solvent exchange with deuterium in the sample buffer, which modulates the $^{15}$N chemical shift tensor via the deuterium isotope effect, adding to the apparent relaxation decay which leads to systematic errors in the relaxation data. This results in an artificial increase of the measured apparent $^{15}$N $R_2$ rate constants – which should not be mistaken with protein inherent chemical exchange contributions, $R_{ex}$, to $^{15}$N $R_2$. For measurements of $^{15}$N $R_2$ rate constants of IDPs and folded proteins at physiological temperatures and pH, we recommend therefore the use of a very low D$_2$O concentration in the sample buffer, as low as 1%, or the use of an external D$_2$O reference along with a modified $^{15}$N $R_2$ Hahn-echo based experiment. This combination allows for the measurement of $R_{ex}$ contributions to $^{15}$N $R_2$ originating from conformational exchange in a time window from μs to ms.

**Keywords**
Intrinsically disordered proteins, NMR relaxation experiments, amide exchange, deuterium isotope effect, loop dynamics

**Introduction**
Proteins are inherently dynamic systems with motions that cover several orders of magnitude wide time scale from femtosecond to more than
seconds (136, 137). Such dynamics may be local, concerted, correlated or of anti-correlated nature (138-142). Nuclear magnetic resonance spectroscopy (NMR) is one of the major methods to study protein dynamics. A plethora of NMR experiments have been and are further being developed to elucidate protein motions (84, 137, 141, 143-148). One of the standard experiments are $^{15}$N $R_1$, $R_2$ relaxation measurements and the $^{15}$N NOE experiment for the detection of the rotational correlation time of the molecule under study as well as local fast dynamics at a residue-specific resolution (i.e. for each $^{15}$N-$^1$H moiety along the amino acid sequence) (149, 150). These measurements have been complemented with more sophisticated experiments and analyses to obtain also intermediate and slow time scale information from µs up to ms. This includes the $^{15}$N CPMG- or $^{15}$N $R_{1p}$ based relaxation dispersion experiments (151, 152), CEST or DEST measurements (144, 146) and alternatively $^{13}$C methyl relaxation measurements covering protein side-chain dynamics (153, 154). Towards a more comprehensive picture of dynamics, residual-dipolar couplings (52, 84, 155), cross-correlated relaxation (139, 142, 156), paramagnetic relaxation enhancement (PRE) (157-159) and eNOE-based (141, 160) data have been acquired and can be used in combination with molecular dynamics simulation (161, 162) or ensemble averaging (44, 84, 140, 160, 163) and chemical-shift based structural ensemble prediction (7, 164, 165).

For the investigation of µs-ms dynamics, $^{15}$N $R_2$ measurements are among the most frequently used experiments. The $^{15}$N $R_2$ rate constant, which
15N transverse relaxation measurements are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange.

describes the decay of 15N transverse magnetization as measured e.g. in a Hahn-echo experiment, has an exchange contribution, $R_{ex}$, due to conformational and chemical exchange that modulates the 15N chemical shift tensor (166, 167) that adds to the $R_{2,0}$ auto-relaxation rate constant: $R_2 = R_{2,0} + R_{ex}$.

It is probably surprising that the presented work identifies a systematic error in many 15N $R_2$ relaxation measurements that deteriorates the dynamics analysis of proteins and in particular intrinsically disordered proteins (IDP) and protein loops when measured under physiological conditions (i.e. pH ~7.4 and at a temperature of ~37 °C). The identified culprit is the fast exchange of the amide protons with water and simultaneously with the internal reference substance, D$_2$O, resulting in an exchange contribution induced by the deuterium-induced isotope shift of 15N, that becomes particularly acute at physiological pH and temperatures. We exemplify this effect using Hahn-echo based 15N $R_2$ measurements that do not refocus exchange contributions and CPMG relaxation dispersion measurements on α-synuclein, which is an IDP associated to Parkinson’s disease, as well as the PDZ2 domain of human phosphatase and provide a straightforward solution (i.e. the use of a very low D$_2$O concentration, as low as 1%, or, alternatively, the use of an external D$_2$O lock and the appropriate pulse sequence).
Results and discussion

**Pulse sequence for the measurement of $^{15}$N R$_2$ relaxation**

The average backbone amide exchange rate constant shows a strong pH dependence, with a minimum around pH 3 ($10^{-1}$ / min at 298 K) and a tenfold increase for each pH unit (168, 169), resulting in ca. $10^2$ / min at pH 6 (298 K) and roughly $10^3$ / min at pH 7.4 (298 K). Indeed, for α-synuclein at 288 K, measured amide solvent exchange rates varied between 2 and 20 s$^{-1}$ for different residues at low salt concentration (20 mM) and between 10 s$^{-1}$ and 80 s$^{-1}$ for high salt concentrations (300 mM) (170). With the emphasis to measure $^{15}$N relaxation of α-synuclein at physiological conditions including physiological temperature (i.e. 303 K) we have therefore selected an NMR pulse sequence that measures the transverse relaxation of $^{15}$N in-phase coherence, with proton decoupling applied during the relaxation period to alleviate the impact of exchange of the $^{15}$N-$^1$H moiety with water. By that, evolution into anti-phase $N_{x/y}H_z$ coherence is minimized (Figure 2.1). This is different from e.g. $^{15}$N TROSY-$R_2^\beta$ experiments using a Hahn-echo based pulse sequence element (Figure S2.1) (171, 172). Because if anti-phase $N_{x/y}H_z$ coherence is present or evolves during the Hahn-echo relaxation delay in presence of amide exchange, amide exchange will lead to decorrelation of two spin-order (173). This loss of the $N_{x/y}H_z$ coherence will lead to an artificial extra relaxation contribution to the measured $^{15}$N R$_2$, rate constant as illustrated in Fig. S2 (this artificial extra relaxation contribution is denoted $R_{ex,amide}$ in Figure S2.2). In Figure 2.1, a TROSY-based and Hahn-echo based
$^{15}$N $R_2$ experiment is shown that avoids this bias introduced by amide exchange. In details, anti-phase $^{15}$N magnetization generated after the first INEPT transfer is transferred further to in-phase $^{15}$N magnetization in the second step of the re-focused INEPT transfer (b). Therefore, at the beginning of the relaxation period, $N_x$ in-phase magnetization is present. After a z-filter (c), in-phase $N_x$ magnetization is subject to transverse $^{15}$N $R_2$ relaxation during the Hahn-echo element. Importantly, the generation of anti-phase magnetization is minimized by $^1$H decoupling (d). After a second z-filter (e), gradient as well as phase-cycling based echo / anti-echo encoding is achieved prior to $t_1$ evolution. After $t_1$ evolution (f), $N_{x/y} H^\beta$ coherence is transferred to $H_{x/y} N^\beta$ coherence during a TROSY-read out scheme, opening this pulse sequence also for large systems (174) (g), which then evolves during acquisition. Further, $^{15}$N magnetization, transferred from $^1$H during the TROSY read-out scheme is destroyed by a 90° pulse on $^{15}$N (h) (172, 175). Note that in this experiment, $^{15}$N $R_{2,0}$ auto-relaxation (plus $R_{ex}$ contribution) and therefore the average of fast and slowly relaxing NH doublet components is measured rather than the decay of the slowly relaxing $N_{x/y} H^\beta$ line, which is measured in the $^{15}$N TROSY- $R_2^\beta$ experiment (172). The relevance of selecting a $^{15}$N-inphase-based pulse sequence becomes apparent when comparing the $^{15}$N relaxation rate constants measured for the intrinsically disordered protein $\alpha$-synuclein using the pulse sequence of Figure 2.1 compared to those measured using the $^{15}$N TROSY- $R_2^\beta$ experiment (172) (see Figure S 2.3).
Figure 2.1. Pulse scheme for the $^{15}$N $R_2$ relaxation (Hahn-echo) experiment. $^1$H magnetization is transferred to $^{15}$N in-phase magnetization via a refocused INEPT transfer. After a z-filter, a Hahn echo ($\Delta - 180^\circ(N) - \Delta$) with the variable relaxation delay $\Delta$ is performed. $^1$H decoupling during the Hahn echo minimizes the evolution of the anti-phase term during the relaxation period. Echo/ anti-echo encoding for quadrature detection is performed prior to the $t_1$ evolution period. Narrow rectangles indicate hard 90° pulses and broader rectangles hard 180° pulses. The rectangular $^1$H pulses marked $-x$ are low power 90º pulses (1.2 ms at 600 MHz); shaped low power $^1$H pulses (1.9 ms) correspond to the center lobe of a $(\sin x)/x$ function, all serving to return the water magnetization to z prior to detection (176). For application to samples that also are enriched in $^{13}$C: durations of $^{13}$C pulses (all 180°) are equal to $\sqrt{\frac{\Omega}{2}}$ (47.4 μs at 600 MHz), where $\Omega$ is the frequency difference between $^{13}$C and $^{13}$C’. Delay durations are $\delta = 2.65$ ms and $\epsilon$ corresponds to the duration of the decoding gradient $G_4$ (60.8 μs; the slight offset ($\epsilon/2$) relative to the $^{15}$N 180° pulse enables insertion of the decoding gradient $G_4$ without introducing a linear phase error in the $^1$H dimension. Gradients: $G_0$ (1000 μs; 21 G/cm), $G_1$ (2650 μs; 0.7 G/cm), $G_2$ (2550 μs; 1.4 G/cm), $G_3$ (500 μs; 42 G/cm), $G_4$ (1000 μs; 7 G/cm), $G_5$ (300 μs; -23 G/cm), $G_6$ (300 μs; 7 G/cm), $G_7$ (1000 μs; 35 G/cm) and $G_8$ (60.8 μs; 23 G/cm) are sine-bell shaped. Phase cycling: $\varphi_1 = 8(y), 8(-y); \varphi_2 = y$; and $\varphi_3 = y, x, -y, -x, x, y; \varphi_4 = y; \varphi_5 = y$ and $\varphi_{rec} = y,-x,-y,x,-y,x,y,-x,-y,x,y,-x,-y,x.y, -x, y, -x, -y, x$. Quadrature detection is implemented using the Rance-Kay echo/anti-echo scheme (177), with the polarity of gradients $G_5$ and $-G_5$ inverted, and $\varphi_3 = y, -x, -y, x, -y, -x, y, x, \varphi_4 = -y$ and $\varphi_5 = -y$ for the second FID generated for each quadrature pair. The relaxation decay of $^{15}$N (in-phase) coherence is sampled at different delay durations $\Delta$ in an inter-leaved manner.
$^{15}$N $R_2$ relaxation contribution by the deuterium isotope effect

Measuring $^{15}$N relaxation of $\alpha$-synuclein at physiological conditions (i.e. pH 7.4 and 303 K) using the pulse sequence shown in Fig. 1, we noticed a variation of extracted rate constants, depending on the D$_2$O concentration in the sample buffer. This is demonstrated in Fig. 2, for which $^{15}$N $R_2$ rate constants of $\alpha$-synuclein were measured in 4 %, 10 %, and 50 % D$_2$O, respectively. The apparent rate constants measured are significantly elevated with increased D$_2$O, apart from C-terminal residues 110 to 140 that do not show any significant increase, due to exchange protection through hydrogen bond formation of acidic-side chains with amide groups (178).

![Figure 2.2: D$_2$O-dependent $^{15}$N relaxation rate constants.](image)

$^{15}$N $R_2$ rate constants of $^{15}$N-labeled $\alpha$-synuclein, measured using the pulse sequence shown in Figure 2.1 and in the presence of (A) 4% (black), 10% D$_2$O (red), or 50% D$_2$O (blue). The experimental temperature was 303 K and the pH was 7.4. The increase of the relaxation rates with increase of D$_2$O identifies D$_2$O as a culprit for $^{15}$N transverse relaxation measurements.
This finding points to a $R_{\text{ex}}$ contribution because of the deuterium isotope effect that modulates the $^{15}\text{N}$ chemical shift tensor due to exchange between protons and deuterons in the amide group at an exchange rate constant, $k_{\text{ex}}$, which is equal to the solvent exchange rate constant (179). This can be seen as follows: The residue-specific $k_{\text{ex}}$ is the sum of the forward and backward pseudo first order reaction rate constant, $k_{\text{HD}}$ and $k_{\text{DH}}$, respectively, where $k_{\text{HD}}$ is the product of the amide solvent exchange rate constant, $k_{\text{NH}}$, describing the exchange of amide protons with water, and the likelihood that an exchange to a deuteron takes place instead of a proton, which is equal to the population of D$_2$O in the sample buffer, $p_D$:

$$k_{\text{HD}} = k_{\text{NH}} \times p_D.$$  

Vice versa, the rate constant for the backward reaction is $k_{\text{DH}} = k_{\text{NH}} \times p_H$ where $p_H$ is the population of H$_2$O in the sample buffer and $k_{\text{ND}}$ the exchange of the amide deuterium with water. It is assumed that $k_{\text{ND}} = k_{\text{NH}}$ (180). This yield:

$$k_{\text{ex}} = k_{\text{HD}} + k_{\text{DH}} = k_{\text{NH}} (p_D + p_H) = k_{\text{NH}}$$  

Figure 2.3 illustrates the described process.

The exchange rate constant describing the modulation of the chemical shift tensor can also be obtained from a kinetic derivation as we shall see:

$$\text{NH} + D_2O \xrightarrow{k_{\text{HD}}^{(2)}} \text{ND} + HDO \quad (3a)$$

$$\text{ND} + H_2O \xrightarrow{k_{\text{DH}}^{(2)}} \text{NH} + HDO \quad (3b)$$

Equation (3a) ad (3b) are of pseudo first order, since both the H$_2$O and D$_2$O concentrations are much higher than the protein concentration.
15N transverse relaxation measurements are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange.

Figure 2.3: Chemical exchange between amide N-H and N-D moiety changes the resonance frequency of 15N nucleus by changing the chemical shift tensor via deuterium isotope effect. In equilibrium, the rate constant that describes conversion from N-H to N-D is described by the solvent amide exchange rate multiplied by population of D2O in sample buffer. The backward reaction from N-D to N-H is described by solvent amide exchange rate times population of H2O in the sample buffer. It is thereby assumed that the exchange of the amide deuterium with water is equal to the exchange of an amide proton with water (180).

Therefore, we obtain:

\[ \text{NH} \xrightarrow{k_{HD}^{(2)} [D_2O]} \text{ND} \]  \hspace{1cm} (4a)

\[ \text{ND} \xrightarrow{k_{DH}^{(2)} [H_2O]} \text{NH} \]  \hspace{1cm} (4b)

with the pseudo first order rate constants:

\[ k_{HD} = k_{HD}^{(2)} [D_2O] = k_{HD}^{(2)} [H_2O]_0 \times p_D = k_{NH} \times p_D \]  \hspace{1cm} (5a)

and

\[ k_{DH} = k_{DH}^{(2)} [H_2O] = k_{DH}^{(2)} [H_2O]_0 \times p_H = k_{NH} \times p_H \]  \hspace{1cm} (5b)

where \([H_2O]_0\) is the concentration of H2O in water in the absence of D2O; \(p_D\) is the population of D2O and \(p_H\) is the population of H2O in the sample.
buffer, resulting in an actual concentration $[H_2O] = [H_2O]_0 \times p_H$ of 
H$_2$O and $[D_2O] = [H_2O]_0 \times p_D$ of D$_2$O in the sample buffer.

For the rate equation describing the time evolution of the concentration 
of the amide group $[NH]$ we obtain:

$$\frac{d}{dt} [NH] = -k_{HD} [NH] + k_{DH} [ND]$$  \hspace{1cm} (6)

Solving the linear differential equation and using the initial condition $[NH]$
(t = 0) = $[NH]_0$ and $[ND]$ (t=0) = 0, as only $^{15}$N bound to protons will lead to
detectable signal after the refocused INEPT and start of the relaxation 
period, yields:

$$[NH](t) = \left(k_{HD} [NH]_0 e^{- (k_{HD} + k_{DH})t} + k_{DH} [NH]_0 \right) / (k_{HD} + k_{DH})$$  \hspace{1cm} (7a)

Using the relations described in equation (5) this yields:

$$[NH](t) = \left(k_{NH} p_D [NH]_0 e^{- k_{NH} (p_D + p_H)t} + k_{NH} p_H [NH]_0 \right) / \left(k_{NH} (p_D + p_H)\right)$$  \hspace{1cm} (7b)

which further simplifies with $p_D + p_H = 1$ to

$$[NH](t) = \left(p_D e^{- k_{NH} t} + p_H \right) [NH]_0$$  \hspace{1cm} (7c)

and thus $k_{ex} = k_{NH}$. That means that the chemical shift tensor gets
modulated at the amide solvent exchange rate.

The difference in the $^{15}$N chemical shift due to the deuterium isotope 
effect has been determined to be $\Delta \omega = N(NiD) = 687 \pm 35$ ppb (181) 
which amounts to ca. 40 Hz = 250 rad s$^{-1}$ at a magnetic field strength of 
14.1 T (corresponding to a proton Larmor frequency of 600 MHz). With 
an amide exchange rate constant, $k_{ex}$, in the range between 10 to 100 s$^{-1}$ 
at pH 7.4 and 303 K, the exchange process is neither in the fast exchange 
limit, $\Delta \omega \ll k$, nor in the slow exchange limit, $\Delta \omega \gg k$, but rather on an
15N transverse relaxation measurements are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange intermediate timescale. To estimate the exchange contribution on 15N $R_2$ as a result of solvent exchange in the sample buffer, the following formula was used:

$$R_{ex} \approx \frac{p_a p_b k_{ex}}{1 + \left(\frac{k_{ex}}{\Delta \omega}\right)^2} \quad (8).$$

as described in (182). For 4 % D$_2$O, described by $p_a = 0.96$ and $p_b = 0.04$, eq. (8) yields $R_{ex} = 3.31$ s$^{-1}$ for $k_{ex} = 100$ s$^{-1}$, $R_{ex} = 0.38$ s$^{-1}$ for $k_{ex} = 10$ s$^{-1}$ and $R_{ex} = 0.04$ s$^{-1}$ for $k_{ex} = 1$ s$^{-1}$. While for 50 % D$_2$O $R_{ex} = 21.55$ s$^{-1}$ for $k_{ex} = 100$ s$^{-1}$, $R_{ex} = 2.5$ s$^{-1}$ for $k_{ex} = 10$ s$^{-1}$ and $R_{ex} = 0.25$ s$^{-1}$ for $k_{ex} = 1$ s$^{-1}$ are estimated. Please note eq. (8) is strictly speaking no longer fulfilled in the latter case because $p_a = p_b = 0.5$ but can be used to get approximate values. Furthermore, effects like a different dipolar coupling interaction for 15N-D vs. 15N-H or the quadrupole moment of the deuteron have not been taken into consideration. Nonetheless, with this rough estimate, an idea on the order of magnitude of the exchange contribution caused by the deuterium isotope effect modulating the 15N chemical shift tensor as a result of chemical exchange between amide protons and deuterons is obtained. It is in good agreement with the experimentally observed D$_2$O dependency of the 15N $R_2$. There is an additional loss mechanism by solvent exchange from a 15N-H moiety to 15N-D moiety during the relaxation delay making the latter moiety impossible to detect by $^1$H acquisition (183). This effect scales linear with the D$_2$O concentration and can explain partly the observed increase in 15N $R_2$ rate constants for the sample containing 50% D$_2$O. As deuterium
is not decoupled during the relaxation period, also $^{15}$N-D anti-phase magnetization will evolve during the relaxation period and contribute by scalar relaxation of the second kind. This effect also scales linear with the D$_2$O concentration in the sample buffer; further an$^{15}$N-H spin pair will show a higher $^{15}$N R$_2$ rate constant than a $^{15}$N-D spin pair (184, 185). Scalar relaxation of the second kind induced by the exchange of amide protons can also be an additional loss mechanism in Hahn-echo based $^{15}$N R$_2$ measurements (186), however we did not observe any significant differences when changing the RF amplitude of the waltz64 $^1$H decoupling scheme from 2.5 kHz to 6 kHz (Figure S2.4).

The use of an external deuterium lock for $^{15}$N R$_2$ relaxation measurements
The findings discussed above request $^{15}$N relaxation R$_2$ measurements in absence of D$_2$O in the sample buffer. This is achieved by using a coaxial insert from Wilmad comprising D$_2$O inserted into a 5 mm thin-wall NMR tube containing the $^{15}$N-labeled α-synuclein in its D$_2$O-free buffer. The external D$_2$O reference is added by inserting a 2 mm capillary which leads to a loss of 16% effective sample volume for a 5 mm NMR tube. No line broadening as result of potential B$_0$ inhomogeneity was observed, however the quality of water suppression was slightly worse, and the spectral noise increased slightly.
This approach allows for $^{15}$N relaxation measurements using external D$_2$O as a lock substance. Figure 2.4 shows a comparison of the $^{15}$N R$_2$, relaxation rate constants of α-synuclein in presence of 4% D$_2$O and in the
15N transverse relaxation measurements are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange in the absence of any D$_2$O in the sample buffer, at two temperatures 283 K and 303 K. Interestingly, while rate constants at pH 7.4 and 283 K vary little (Figure 2.4a), at 303 K rate constants measured in the absence of D$_2$O are systematically lower than in the presence of only 4% of D$_2$O, with the exception of the last ~30 residues (Figure 2.4b).

**Figure 2.4:** The importance of using a D$_2$O concentration as low as 1% or an external deuterium lock for the measurement of $^{15}$N relaxation rates of $^{15}$N-labeled α-synuclein: $^{15}$N relaxation rates in the absence of D$_2$O versus a D$_2$O content of 4% and 1% in the sample buffer. Hahn-echo based $^{15}$N $R_2$ rate constants of α-synuclein measured with the pulse sequence shown in Figure 2.1. Rate constants measured in the presence of 4% D$_2$O (black) are compared to those measured without D$_2$O (light blue) in the sample buffer at pH 7.4 and temperatures of (A) 283 K and (B) 303 K (using an external deuterium lock). When using a D$_2$O concentration of 1% (red), even at 303 K the effect is small.
Similar observations have been made with a D2O-free sample that lacked an external locking substance and was thus measured without locking the magnetic field (data not shown).

The effect of sample internal D$_2$O on the $^{15}$N R$_2$ relaxation measurements on the folded protein domain PDZ2
To illustrate that the documented deuterium exchange effects are visible not only for IDPs as illustrated above for α-synuclein, relaxation measurements on the $^{15}$N-labeled PDZ2 domain of human phosphatase (187) were performed at pH 8.0 and a temperature of 303 K. The impact of the presence of D$_2$O in the sample buffer on the measured $^{15}$N R$_2$ rate constants for the PDZ2 domain is illustrated in Figure 2.5. Some residues in loop regions (i.e. Asn16, Gly19, Gly24, Gly25, Gly34, Gly50, and Gly63) show a systematic increase in the $^{15}$N R$_2$ rate constants when measured in the presence of only 4% D$_2$O in the sample buffer compared to the sample without any D$_2$O in the sample buffer, using an external D$_2$O reference. Glycine residues appear thereby to be overrepresented which is attributed to their overall fast intrinsic amide-water exchange (188).

The impact of internal D$_2$O on CPMG-based relaxation dispersion experiments
Because of the significant $R_{ex}$ contribution on the measured $^{15}$N R$_2$ rate constant caused by D$_2$O in the sample buffer, we simulated the anticipated $R_{ex}$ contributions in a CPMG relaxation dispersion experiment.
15N transverse relaxation measurements are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange

$s^{-1}, (170)$ depending on the extent of solvent exposure of the respective residue. At pH 7.4 and 25°C (298 K), for solvent-exposed residues the amide exchange rate will assume values in the order of $k_{NH} = 10 \text{ s}^{-1}$ to $k_{NH} = 100 \text{ s}^{-1}$ per residue. Calculations for different amounts of D$_2$O in the sample buffer (1%, 4% and 10%) are shown in Figure 2.6.

As illustrated in Figure 2.6A, in the presence of 10% D$_2$O for a residue showing fast amide exchange with $\gamma = 100 \text{ s}^{-1}$ the $R_{ex}$ contribution due to D$_2$O in the sample buffer is present at CPMG frequencies less than $100 \text{ s}^{-1}$, is significantly reduced for CPMG frequencies $\nu_{cpmg} > 100 \text{ s}^{-1}$, and fully averaged out for a CPMG frequency $\nu_{cpmg} = 500 \text{ s}^{-1}$. Even for a D$_2$O concentration as low as 1%, the maximum $R_{ex}$ contribution goes up to $1 \text{ s}^{-1}$.
\textsuperscript{1} (Figure 2.6E). While this may be negligible for the structured part of a large globular protein with an $R_2$ rate constants of e.g. 50 s\textsuperscript{-1}, it amounts to an error of 50\% for an IDP with a rate constant of e.g. 2 s\textsuperscript{-1}. Overall, the $R_{ex}$ contribution roughly scales linear with the percentage of D\textsubscript{2}O in the sample buffer and the given amide exchange rate constant $k_{NH}$. Furthermore, at lower pH < 6.5 and temperatures around or below room temperature, where the amide exchange rate will usually be less than $k_{NH}<$ 10 s\textsuperscript{-1}, at 1\% of D\textsubscript{2}O in the sample buffer the $R_{ex}$ contribution by D\textsubscript{2}O can be safely ignored (Figure 2.6F). When approaching physiological pH and temperature however, the amide exchange rate constants for many residues can approach values of 100 s\textsuperscript{-1}(170). Then, for $\nu_{cpme}<$ 100 s\textsuperscript{-1} the $R_{ex}$ contribution by D\textsubscript{2}O in the sample buffer can add a significant systematic error on measured $R_2$ rate constants of IDPs that have low $R_{2,0}$ auto-relaxation constants.
15N transverse relaxation measurements are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange.

**Figure 2.6:** Simulated CPMG relaxation dispersion curves in presence of different amounts of D$_2$O in the sample buffer. The $R_{ex}$ contribution to $^{15}$N $R_2$ is shown as a function of the applied CPMG frequency, with $\nu_{\text{cpmg}} = 1 / (2 \tau_{\text{cp}})$ and $\tau_{\text{cpmg}}/2 - 180^\circ$ (N) $- \tau_{\text{cpmg}}/2$ constituting the basic CPMG block. Data are shown for 10% D$_2$O in the sample buffer and an amide exchange rate constant of (A) $k_{NH} = 100$ s$^{-1}$ and (B) = 10 s$^{-1}$, for 4% D$_2$O and (C) $k_{NH} = 100$ s$^{-1}$ and (D) = 10 s$^{-1}$ as well as for 1% and (E) $k_{NH} = 100$ s$^{-1}$ and (F) = 10 s$^{-1}$. See Materials and Methods for further details.
We have tested the impact of D$_2$O in sample buffer on the extracted CPMG-based $^{15}$N $R_2$ rate constants experimentally with $\alpha$-synuclein. Indeed, at a CPMG frequency of 100 Hz, the exchange contribution induced by D$_2$O appears to be refocused. However, at a low CPMG frequency of 20 Hz we observe substantial $R_{ex}$ contributions, leading to increased $^{15}$N $R_2$ rate constants in the presence of 10% D$_2$O (Figure 2.7). Finally, we measured also standard $^{15}$N $R_{1\rho}$ relaxation measurements (172) with a spin-lock RF amplitude of 2 kHz on $\alpha$-synuclein (pH 7.4, 303 K). As expected, when comparing $^{15}$N $R_{1\rho}$ rate constants in the presence of 10% D$_2$O and absence of D$_2$O in the sample buffer, we observe only little differences (Figure 2.8). Therefore, the $R_{ex}$ contribution induced by D$_2$O in the sample buffer appears to be refocused for a standard $^{15}$N $R_{1\rho}$ experiment employing a spin-lock RF amplitude of 2 kHz.
15N transverse relaxation measurements are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange.

**Figure 2.7:** Impact of D$_2$O on CPMG-based $^{15}$N R$_2$ rate constants of α-synuclein with weak CPMG frequencies (i.e. 20 Hz and 100 Hz). CPMG-based $^{15}$N R$_2$ rate constants of α-synuclein were measured at pH 7.4 and at 303 K in the presence of 10% D$_2$O (black), 1% D$_2$O (red) and in the absence of D$_2$O in the sample buffer (light blue). The CPMG frequency was (A) 20 Hz and (B) 100 Hz.

**Discussion**

The presented data shows that at near physiological pH (i.e. pH 7.4) and physiological temperatures of 30-37 °C, solvent exchange of the amide protons with deuterium in the sample buffer can impact Hahn-echo based $^{15}$N R$_2$ measurements significantly due to the deuterium isotope effect even at low concentration of D$_2$O in the sample buffer (as low as 4%). This effect is pronounced for several loop residues in the folded protein.
domain PDZ2 but is most prominent in the intrinsically disordered protein α-synuclein. As many IDPs show very low $^{15}$N $R_2$ rate constants (< 5 s$^{-1}$) due to their high intrinsic flexibility, even a small systematic artifactual $R_{ex}$ contribution of e.g. 1 s$^{-1}$ can lead to a large error in the data. Therefore, for Hahn-echo based $^{15}$N $R_2$ measurements the use of only a very low D$_2$O concentration in the sample buffer, as low as 1%, is necessary or, alternatively, the use of an external D$_2$O lock using a coaxial capillary insert.

Since IDPs form a large part of the human proteome (30-40%) and play an essential role in cellular signaling and regulation of many biomolecular interactions (189-191), over the last two decades solution-state NMR provided important insights to characterize secondary structure propensity, conformational space (192, 193) and non-local and local dynamics of IDPs using mainly $^{15}$N CPMG based relaxation dispersion experiments (149-152, 191, 193-203). Several experimental strategies

Figure 2.8: Lack of impact of D$_2$O on $^{15}$N $R_{1p}$ rate constants of α-synuclein under a spin-lock frequency (i.e. 2 kHz). $^{15}$N $R_{1p}$ rate constants for α-synuclein (pH 7.4, 303 K) were measured using a spin-lock RF amplitude of 2 kHz in absence (red) and presence (black) of 10% D$_2$O.

189-191
192, 193
149-152, 191, 193-203
15N transverse relaxation measurements are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange have been designed to allow the recording of \(^1\text{H}-^{15}\text{N}\) correlation spectra (204) and CPMG relaxation experiments of IDPs under physiological conditions and obviate the influence of amide exchange (183, 204), but the adverse impact of D\(_2\)O through the isotope effect has to our knowledge escaped attention. At physiological pH and near physiological temperatures, we observed a substantial \(R_{\text{ex}}\) contribution induced by D\(_2\)O in the sample buffer that is not refocused for a low CPMG frequency of 20 kHz. That finding agrees with simulated data that predict a substantial contribution for CPMG frequencies < 100 Hz. At a CPMG frequency of 100 Hz the \(R_{\text{ex}}\) contribution induced by D\(_2\)O in the sample buffer appears however refocused, in agreement between experimental and simulated data. Therefore, low frequency data points in CPMG relaxation experiments < 100 Hz will be affected by \(R_{\text{ex}}\) contribution induced by D\(_2\)O in the sample buffer, unless a very low D\(_2\)O concentration in the sample buffer, as low as 1%, or an external D\(_2\)O reference is used. For standard \(^{15}\text{N} R_2\) measurements, employing a CPMG frequency of at least 100 Hz as well as proton decoupling (183, 204) (to counteract the adverse effect of chemical exchange), the effect is however refocused and will not lead to an artificial increase of the \(^{15}\text{N} R_2\) rate constants. Also, in standard \(^{15}\text{N} R_{1p}\) experiments that spin-lock \(^{15}\text{N}\) transverse magnetization, deuterium isotope effects will be refocused, as long as the spin-lock RF amplitude, given in frequency units, is significantly faster than the amide exchange rate constants – which is usually the case, e.g. for an RF amplitude of 2 kHz and an amide exchange rate in the order of 100 s\(^{-1}\).
For relaxation dispersion measurements on IDPs using CPMG-based NMR pulse sequences (such as suggested by (183) and (204)) the CPMG frequency for the first, low frequency data points, can however be lower than the solvent-exchange rate, depending on the settings for the minimal frequency of the CPMG block in relaxation dispersion experiments, and thus exchange with deuterons in the sample buffer may adversely affect the accuracy of the extracted results and may lead to artificial $R_{ex}$ effects, which originate from deuterium exchange and a modulation of the $^{15}$N chemical shift tensor through the deuterium isotope effect rather than conformational dynamics. For high precision CPMG-based relaxation measurements, we therefore recommend also the use of a very low D$_2$O concentration, as low as 1%, or, alternatively, the use of an external deuterium reference, which is easily possible using commercially available NMR tube inserts.

**Conclusion**

The determination of the $^{15}$N $R_2$ relaxation rate constants is a standard NMR experiment in the evaluation of the dynamics of proteins, including both folded and intrinsically disordered protein entities. While measurements at low pH (< 6.5) or low temperatures (< 10 °C) are usually uncritical because of low solvent amide exchange rates, at physiological pH and temperatures, effects related to solvent amide exchange can lead to artifactual $R_{ex}$ contributions.
The presented results show that the presence of D$_2$O > 1% in the sample buffer can deteriorate the accuracy of the rates constants measured using a Hahn-echo based $^{15}$N $R_2$ experiments and also for low CPMG-frequency data points (< 100 Hz) in CPMG relaxation dispersion experiments. For CPMG frequencies > 100 Hz as well as for $^{15}$N $R_{1p}$ experiments that apply a high-power spin-lock RF amplitude, of e.g. 2 kHz, the effect is refocused, even in the presence of larger amounts of D$_2$O in the sample buffer. For Hahn-echo based $^{15}$N $R_2$ measurement or CPMG-based $^{15}$N $R_2$ relaxation dispersion experiments at near physiological conditions, we however recommend the use of a very low D$_2$O concentration in the sample buffer, as low as 1% or, alternatively, the use of an external deuterium reference. This applies both to in vitro or in-cell NMR experiments (47, 205-209).

**Material and Methods**

*Protein expression and purification*

Acetylated α-synuclein was expressed using co-expression of the N-terminal acetyltransferase B (NatB) complex and the α-synuclein plasmid (pRK172), as described earlier (210). Expression and purification were performed as described earlier (211), with some modification. Briefly, after transformation, colonies containing both plasmids (NatB and pRK172) were grown at 37 °C in 10 ml Lysogeny Broth (LB) medium overnight and were then transferred into 1 L of LB media. After reaching an OD$_{600}$ of around 0.5, cells were harvested by centrifugation and resuspended into 1 L M9 minimal media containing $^{15}$NH$_4$Cl and grown till
an OD$_{600}$ of 1.0 was reached. Protein expression was carried out overnight at 37°C, after induction with 1 mM IPTG. Cells were harvested by centrifugation and α-synuclein, present in the periplasm, was purified using ion exchange chromatography and hydrophobic interaction chromatography as described earlier (105).

The PDZ2 domain from human phosphatase (hPTP1E) was encoded into a pET21 expression system with a T7 promoter and Histidine tag. Expression and purification were performed as described earlier (187), with some modifications. After transformation, a single colony was inoculated overnight in 10 ml LB medium at 37 °C and then transferred into 1 L M9 minimal media containing $^{15}$NH$_4$Cl and grown till an OD$_{600}$ of 0.5 was reached. Protein expression was induced by adding 1 mM IPTG and cells were harvested by centrifugation after 5 h. A Ni-affinity column (HisTrap FF) was used for purification of protein and the histidine-tag was cleaved with Human Rhinovirus 3C (HRV 3C) followed by another Ni-affinity column purification step.

**NMR Measurements**

NMR spectra were recorded with 500 µM of $^{15}$N-labeled acetylated α-synuclein dissolved in 20 mM Tris (pH 7.4) and 100 mM NaCl, unless indicated differently. Spectra of $^{15}$N-labeled PDZ domain, dissolved in 50 mM sodium phosphate buffer (pH 8.0) and 150 mM NaCl, were recorded at an experimental temperature of 303 K. $^{15}$N TROSY- $R_2^*$ rate constants
15N transverse relaxation measurements are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange.

were measured by applying the NMR experiment described earlier (172). 15N R2 experiments, applying proton decoupling during the relaxation delay period (167), were recorded using the pulse sequence described in Figure 2.1. For proton decoupling, waltz64 with an RF amplitude of 2.5 kHz was applied. 15N R1ρ rate constants were recorded using the NMR experiment described in (172). CPMG-based 15N R2 rate constants were determined using a proton-decoupled CPMG experiment, similar to the one described by Yuwen and Skrynnikov (204), however using waltz64 with an RF amplitude of 2.5 kHz for proton decoupling rather than DIPSI2. 15N R2 rate constants were measured for the two CPMG frequencies, 20 Hz and 100 Hz. All NMR experiments were performed on a Bruker 600 MHz Avance III HD spectrometer equipped with cryogenic probe. Spectral dimensions were $\Omega(^{1}H) \times \Omega(^{15}N) = 14.014$ ppm x 35 ppm. 512 complex points were recorded in the direct dimension ($^{1}H$) and 80 complex points in the indirect dimension ($^{15}N$), resulting in an acquisition time of 60.08 ms in the direct and 37.5 ms in the indirect dimension, respectively. The $^{1}H$ carrier was set to 4.7 ppm and the $^{15}N$ carrier to 118 ppm, respectively. The magnetization decay was recorded using four different relaxation decay periods, in an inter-leaved manner (172). For $\alpha$-synuclein, R2 relaxation delays were set to 0 ms, 200 ms, 100 ms, and 50 ms and for the 15N R1ρ experiments, delays were set to 1 ms, 120 ms, 60 ms, and 20 ms, respectively. The spin-lock RF field strength in the $^{5}$N R1ρ measurement was set to 2 kHz. For experiments using the 15N TROSY- R2 sequence decay periods differed slightly; at pH 7.4 and at a temperature of 303 K or
283 K delays were 0, 100, 50, and 20 ms. Spectral intensities for the different decay periods were recorded in an inter-leaved manner, 16 scans were recorded for each decay period. The total experimental time was 4.75 h. For the PDZ domain, relaxation delays were set to 0 ms, 50 ms, 30 ms, and 10 ms. For the CPMG experiments, a fixed relaxation delay of 200 ms was used. Relaxation data were recorded for two different CPMG frequencies, 100 Hz and 20 Hz by adjusting the number of 180°(N) pulses and the inter-pulse delay accordingly.

The software Sparky 3.115 (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco, USA) and Bruker Topspin 3.5pl7 (Bruker, Inc., Billerica, MA, USA) were used for analyzing the spectra and extracting the rate constants.

NMR spectra recorded on samples containing D₂O in the sample buffer were measured using a regular Shigemi tube (Sigma-Aldrich, Merck KGaA, Darmstadt, Germany). For samples without D₂O in the sample buffer, D₂O was added as an external reference using a Wilmad coaxial insert (stem length 50 mm, 2 mm diameter) and the sample was kept in a thin wall 5 mm NMR tube (Wilmad NMR tubes, 5 mm diam., precision, Sigma-Aldrich, Merck KGaA, Darmstadt, Germany). The coaxial insert containing D₂O was inserted into the 5 mm tube thin-wall NMR tube containing the sample.

**Simulation of CPMG relaxation dispersion curves**

CPMG relaxation dispersion curves were calculated using the formula:
15N transverse relaxation measurements are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange

\[ R_{ex} = p_{H}p_{D} \Delta \omega^2 k_{ex} / \left[ k_{ex}^2 + \left( p_{H}^2 \Delta \omega^4 + 144/\tau_{CP} \right)^{1/2} \right] \]  

as described in (212, 213), with the basic CPMG element \( \tau_{CP}/2 - 180^\circ \) \(-\tau_{CP}/2 \). The inter-pulse delay \( \tau_{CP} \) relates to the CPMG frequency \( \nu_{CPMG} \) via

\[ \nu_{CPMG} = 1/(2\tau_{CP}). \Delta \omega = N(N_{1}D) = 687 \pm 35 \text{ ppb (181)} \] which amounts to ca. 40 Hz = 250 rad s\(^{-1}\) at a magnetic field strength of 14.1 T (corresponding to a proton Larmor frequency of 600 MHz) and \( k_{ex} \) is equal to the assumed amide solvent exchange rate constant (see Results section for derivation); \( p_{D} \) is the population of deuterium in the sample buffer (e.g. 0.1 for 10% D\(_{2}\)O) and \( p_{H} \) the population of H\(_{2}\)O in the sample buffer.
Supporting Information

Figure S2.1. Pulse scheme used for measurement of the relaxation rates of the slowly relaxing $^{15}$N TROSY component (i.e. $R_2^β$) using a Hahn echo ($Δ - 180°(N) - Δ$) element during the echo/anti-echo gradient encoding step (172). The rectangular $^1$H pulses marked $-x$ are low power 90° pulses (1.2 ms at 600 MHz); shaped low power $^1$H pulses (1.9 ms) correspond to the center lobe of a (sinx)/x function, all serving to return the water magnetization to z prior to detection (176). For application to samples that also are enriched in $^{13}$C: durations of $^{13}$C pulses (all 180°) are equal to $\frac{\sqrt{Ω}}{2π}$ (47.4 μs at 600 MHz), where Ω is the frequency difference between $^{13}$C$^α$ and $^{13}$C$'$. Delay durations are δ = 2.65 ms and ε corresponds to the duration of the decoding gradient $G_4$ (60.8 μs; the slight offset (ε/2) relative to the $^{15}$N 180° pulse enables insertion of the decoding gradient $G_4$, without introducing a linear phase error in the $^1$H dimension. Gradients: $G_1$ (200 μs; 1.4 G/cm), $G_2$ (300 μs; 7 G/cm), $G_3$ (1 ms; 35 G/cm), $G_4$ (60.8 μs; 23 G/cm), $G_5$ (300 μs; -23 G/cm) and $G_6$ (1 ms; 21 G/cm) are sine-bell shaped. Phase cycling: $φ_7$ = y,x,-y,-x; $φ_1$ = y; $φ_{rec}$ = y,-x,-y,x. Quadrature detection is implemented using the Rance-Kay echo/anti-echo scheme (177) with the polarity of gradients $G_5$ and $-G_5$ inverted, and phase $φ_7$ = y,-x,-y,x and $φ_1$ = -y for the second FID generated for each quadrature pair. The relaxation decay of the slowly relaxing TROSY component $R_2^β$ is sampled at different delay durations $Δ$ in an interleaved manner.

\[ R_2^β \]
15N transverse relaxation measurements are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange.

\[ R_2 = R_{2,0} + R_{2,ex} + R_{ex, amide} \]

\[ R_{ex, amide} : \]

**Figure S2.2:** Mechanism how amide water exchange contributes to increased relaxation rates. If anti-phase \( Nx/yHz \) coherence is present during the Hahn-echo relaxation delay and subject to amide exchange, loss of coherence will lead to an artificial extra contribution, \( R_{ex, amide} \), to the measured 15N R2 rate constant. The effect of amide exchange is illustrated for the \( NxHβ \) operator, selected during the TROSY-readout scheme: (a) in the absence of amide exchange. b) The \( NxHβ \) coherence is destroyed by amide exchange and cannot be transferred to detectable magnetization, therefore leading to an additional decay of 15N R\( _2^\beta \) magnetization, reflected by a skewed (too high) measured 15N R2 rate constant. This effect applies not only to TROSY sequences but to all 15N R2 measurements where anti-phase \( Nx/yHz \) coherence is generated.
Figure S2.3: Pulse sequence-dependent effects on the transverse $^{15}$N $R_2$ relaxation rate constants measured on α-synuclein attributed to amide hydrogen – water exchange. Comparison between $^{15}$N TROSY-$R_2^0$ rate constants measured using the Hahn-echo based $^{15}$N TROSY-$R_2^0$ experiment (black) subject to amide exchange contributions on the apparent $^{15}$N TROSY-$R_2^0$ rate constant using the pulse sequence described earlier (172) and $^{15}$N $R_2$ rate constants measured using the pulse sequence shown in Figure 2.1 (blue). The experiments were measured on α-synuclein at a pH of 7.4 and at a temperature (A) 283 K and (B) 303 K, respectively. Due to amide – water exchange the $R_2^0$ rate constants are significantly higher than the $R_2$ rates. The effect is even more pronounced at 303 K when compared to 283 K as the amide – water exchange is faster at higher temperature.
15N transverse relaxation measurements are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange.

Figure S2.4: $^{15}$N $R_2$ rate constants of α-synuclein measured using the proton-decoupled Hahn-echo sequence described in Figure S2.1, using an RF amplitude of 6 kHz (black) compared to 2.5 kHz for the 1H waltz 64 decoupling scheme used in this study.
Chapter 3

High-density lipoprotein-like particle formation of synuclein variants

Abstract

α-Synuclein (α-Syn) is an intrinsically disordered protein in solution whose fibrillar aggregates are the hallmark of Parkinson’s disease (PD). Although the specific function of α-Syn is still unclear, its high structural plasticity is key in understanding basic interactions of α-Syn with biological membranes. Recently, it has been observed that α-Syn is able
to form high-density lipoprotein-like (HDL-like) particles reminiscent of self-assembling phospholipid bilayer nanodiscs. Here we extended our preparation method for the production of α-Syn lipoprotein particles to the β- and γ-Syn variants, and the PD-related familial α-Syn mutants. We show that the entire family of human Syns is able to define stable and homogeneous populations of HDL-like particles with distinct morphologies. Our results characterize the impact of the individual Syns on the formation capacity of these particles and indicate that Syn HDL-like particles are neither causing toxicity nor a toxicity-related loss of α-Syn in PD.

**Keywords**
alpha Synuclein (α-Synuclein), α-Synuclein lipoprotein particles, neurodegenerative disease, apolipoprotein, lipid transport, membrane.

**Introduction**

α-, β-, and γ-Synuclein (Syn) form a family of charged proteins that are exclusively found in neurons. α-Syn plays a key role in Parkinson’s disease (PD) (214, 215) by initiating amyloid fibril formation of β-sheet structure (64, 65, 216). The natural state of α-Syn inside the cell (47) and in an aqueous solution is monomeric and largely unstructured (25, 217, 218). However, membrane mimicking environments containing anionic phospholipids (25, 27, 92) or anionic detergents (28) trigger a
conformational transition within the seven imperfect N-terminal repeat region (Figure 3.1) towards a helical state reminiscent of the amphipathic helices found in apolipoproteins (27, 217, 219).

Multiple in vivo functions ranging from vesicle pool maintenance (51, 52), regulation of dopamine neurotransmission (48, 220), transport of lipids and fatty acids (132, 221-224), membrane trafficking (60, 140, 225), synaptic plasticity (27, 226, 227) to the assistance in SNARE complex formation (228-231) are based on observed α-Syn-membrane interactions. However, there is evidence that membranes can also play a role in the aggregation of α-Syn (74, 101, 232-234).

Recent studies detail the observation of high-density lipoprotein-like (HDL-like) particles consisting of multiple α-Syn molecules that encapsulate anionic lipid bilayers (231, 235). In these protein-lipid entities, residues 1-100 of α-Syn adopt a broken helical conformation to form the core structure, while the 40 C-terminal residues remain in a random-coil state (38). The close resemblance in size between cross-linked in vitro-derived α-Syn lipoprotein nanoparticle sand a cross-linked species of endogenous α-Syn from SH-SY5Y cells (235) may point to a potential functional relevance of these protein complexes in vivo during lipid transport events or membrane remodeling.

In the present study we show that all three variants of Syn (i.e. α-Syn, β-Syn, and γ-Syn) and the PD-related familial α-Syn mutants are able to form HDL-like particles and describe their impact on the lipoprotein particle formation.
Results and Discussion

Sequence Similarity of the Synuclein Family with the Apolipoprotein A-1
As previously observed, there is a significant sequence similarity between α-Syn and the Apolipoprotein A-1 (231). Here, we expanded this sequence alignment to include β- and γ-Syn. Applying a multiple protein sequence alignment of the Syn family members α-, β-, and γ-Syn with the Apolipoprotein A-1 using Clustal Omega (236) shows that the highest similarity is found for β-Syn (21.5%) followed by α- (19.1%), and γ-Syn (18.7%) (Figure 3.1).

Lipoprotein Particle Formation of the Synuclein Family and the Familial α-Synuclein Mutants
Stable nanometer-sized lipoprotein particles composed of the cell membrane lipid sphingomyelin and members of the Syn family were produced following the protocol described by Eichmann et al. (235). Specifically, α-Syn and its N-terminally acetylated form, the non-acetylated α-Syn mutants, A30P, E46K, H50Q, G51D and A53T, as well as non-acetylated β- and γ-Syn were mixed with sphingomyelin dissolved in sodium cholate at a molar protein-to-lipid ratio of 1 to 40. Subsequently, lipid-bound Syn species were detected by SEC at a retention volume of ~8 ml (α-Syn(H50Q) and γ-Syn) and ~10 ml (α- and β-Syn) separated from monomeric Syn (retention volume~15 ml) (Figures 3.2 and 3.3). The extent of lipoprotein particle formation was quantified from the
integrated areas of the two SEC elution profile peaks (Table 1). In the case of \( \alpha \)-Syn, N-terminally acetylated \( \alpha \)-Syn, \( \beta \)-, and \( \gamma \)-Syn the SEC elution profiles indicate that the propensity of lipoprotein particle formation increases in the above-specified order with values of the populated lipoprotein particle species of 68, 85, 87, and 96\%, respectively. Comparing the populations of the lipid-bound \( \alpha \)-Syn mutants shows that A30P favors lipoprotein particle formation only to 52\% while higher populations are observed for E46K (62\%), H50Q (68\%), and A53T (78\%). Full conversion from a monomeric state into a lipid-protein entity is exclusively detected for the \( \alpha \)-Syn familial variant G51D (Figure 3.3, Table 3.1).

**Morphology of the Synuclein Lipoprotein Particles**
The morphology of the Syn sphingomyelin lipoprotein particles eluting at \(~8\) and \(10\ \text{ml}\) in the SEC were assessed by negative-stain EM (Figures 3.2 and 3.3). In line with our previous findings (38), population of circular particles were observed for \( \alpha \)- and \( \beta \)-Syn as well as the \( \alpha \)-Syn mutants A30P, E46K, G51D, H50Q, and A53T with diameters of \(~25\ \text{nm}\) (Figures 3.2 and 3.3). In contrast, sphingomyelin lipoprotein particles formed by \( \gamma \)-Syn are substantially larger (\(~60\ \text{nm}\)) and are characterized by an elongated elliptical shape (Figures 3.2).
Figure 3.1: Sequence alignment between various Syn variants and Apolipoprotein A-1 as indicated. The seven predicted membrane-binding helices of interest are highlighted and labeled I-VII. Amino acid residue identity is indicated by ‘*’, ‘:’, and ‘.’ for fully, strongly, and weakly conserved residues. The sequence alignment has been done using the software Clustal Omega (41).
High-density lipoprotein-like particle formation of synuclein variants

Figure 3.2: The α-, β-, and γ-Syn variants form HDL-like particles in presence of sphingomyelin. Size-exclusion gel chromatography (Superdex 200 10/300GL) of α-, β-Syn, and γ-Syn sphingomyelin lipoprotein particles. The size-exclusion gel chromatogram of N-terminally acetylated α-Syn sphingomyelin lipoprotein particles is shown in red. HDL-like particles elute at ~8-10 ml, while the monomeric Syns elute at ~15 ml. Negatively stained electron micrographs of the fraction taken at ~8-10 ml containing the maxima at the absorbance of 280 nm are shown as insets (scale bar 50 nm). Both top and side views of the nanoparticles are visible. The nanoparticles tend to stack.
Figure 3.3: The familial mutants of α-Syn form sphingomyelin lipoprotein particles. Size-exclusion gel chromatography (Superdex 200 10/300GL) of all known familial α-Syn variants in presence of sphingomyelin as indicated (i.e. A30P, E46K, H50Q, G51D, and A53T). The HDL-like particles elute at ~8-10 ml, while monomeric Syns elute at ~15 ml. Negatively stained electron micrographs of the fraction taken at ~8-10 ml containing the maxima at the absorbance of 280 nm are shown as insets (scale bar 50 nm). Both top and side views of the nanoparticles are visible. The nanoparticles tend to stack.
Figure 3.4: Relative amount of sphingomyelin lipoprotein particle formation within the familial mutants of α-Syn as a function of the relative predicted α-helix propensity.

The relative amount of sphingomyelin particle formation compared to free monomeric protein within the familial mutants of α-Syn obtained from the corresponding integrated size-exclusion gel chromatogram peak areas is shown as a function of the predicted relative α-helical propensity using established (237) with 0 for WT α-Syn.
Discussion

The seven imperfect repeats in the N-terminal region of the Syn primary sequences, which are reminiscent of the amphipathic helices found in apolipoproteins, mediate binding to membranes \textit{in vitro} while the C-terminal residues remain substantially unstructured (27, 217, 219, 235). Comparing the degree of protein sequence similarity

<table>
<thead>
<tr>
<th>Protein</th>
<th>Retention volume [ml]</th>
<th>Relative amount of Syn sphingomyelin particle formation compared to free Syn [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetylated (\alpha)-Syn</td>
<td>10.1</td>
<td>85</td>
</tr>
<tr>
<td>(\alpha)-Syn</td>
<td>9.6</td>
<td>68</td>
</tr>
<tr>
<td>(\beta)-Syn</td>
<td>10.1</td>
<td>87</td>
</tr>
<tr>
<td>(\gamma)-Syn</td>
<td>8.1</td>
<td>96</td>
</tr>
<tr>
<td>A30P</td>
<td>9.6</td>
<td>52</td>
</tr>
<tr>
<td>E46K</td>
<td>9.8</td>
<td>62</td>
</tr>
<tr>
<td>H50Q</td>
<td>8.3</td>
<td>68</td>
</tr>
<tr>
<td>G51D</td>
<td>9.5</td>
<td>100</td>
</tr>
<tr>
<td>A53T</td>
<td>9.3</td>
<td>78</td>
</tr>
</tbody>
</table>

\textbf{Table 3.1.} Retention volume and relative amount of Syn sphingomyelin lipoprotein particle formation compared to free monomeric protein obtained from the corresponding integrated size-exclusion gel chromatogram peak areas in Figures 3.2 and 3.3.
between the different members of the Syn family with the Apolipoprotein A-1 reveals that the propensity of Syn sphingomyelin lipoprotein particle formation is independent from the sequence similarity but is dictated by length of the acidic C-terminus, i.e. γ-Syn 96% lipoprotein particle formation (18.7% identity, total number of amino acids 127), β-Syn 87% (21.5% identity, total number of amino acids 134), and α-Syn 68% (19.1% identity, total number of amino acids 140). In addition, the SEC data show an increase of α-Syn sphingomyelin lipoprotein particle formation by 17% upon N-terminal acetylation of α-Syn, a post-translational modification present in all mammalian cell systems (34, 45, 47), which in solution increases the population of a transient helix within the N-terminus and enhances thereby lipid-binding (238). These findings indicate that the extent of the amphipathic helix formation is key in the formation of Syn HDL-like particles. This notion finds support by plotting the degree of HDL-like particle formation of the familial α-Syn variants versus their relative helix propensities (Figure 3.4). For example, the amino acid replacement Gly to Asp in the G51D variant increases the helical propensity of α-Syn by +0.45, which is reflected by an almost complete incorporation of α-Syn (G51D) into sphingomyelin lipoprotein particles, while for the variant A30P having a decreased predicted helix propensity of -0.86 only about half of the protein is forming HDL-like particles. The only exception within this correlation is the data of α-Syn (A53T), which has a mutation outside the helical region (between predicted helix IV and V in Figure 3.1). Thus, it is expected that this mutation does not have an influence on the
formation of HDL-like particles. Indeed, α-Syn (A53T) shows the same degree of sphingomyelin lipoprotein particle formation as WT α-Syn (Figures 3.1, 3.2, 3.3, and 3.4). Considering that some of the familial variants display a decreased (i.e. A30P, E46K) while others an increased level (i.e. G51D) in HDL-like particle formation compared to WT α-Syn, it can be rationalized that in PD the HDL-like particles are neither causing toxicity nor a toxicity-related loss of α-Syn function. Finally, it is worth mentioning that in addition to the propensity variation of lipoprotein particle formation among the Syn family and its variants, distinct morphologies differences are observed as revealed by negative stain EM ranging from roundish (α-Syn) to more elongated elliptical (γ-Syn) shapes.

In conclusion, our presented data show that the entire family of human Syns including the familial mutants of α-Syn is able to form HDL-like particles with different morphologies. However, single point mutations in the amino acid sequence of α-Syn can have a significant impact on the capacity to form HDL-like particles. It remains to be demonstrated that these Syn lipoprotein particles are present in vivo with a specific biological function.
Material and Methods

Expression of the Synuclein Family and the Familial α-Synuclein Mutants

Wild-type (WT) α-, β-, and γ-Syn and the familial α-Syn mutants A30P, E46K, H50Q, G51D and A53T were overexpressed in the Escherichia coli strain BL21 Star™ (DE3) pLysS (Invitrogen) and purified as described previously (105).

Preparation of Synuclein Lipoprotein Particles

Syn lipoprotein particles consisting of sphingomyelin (brain, porcine, Avanti Polar Lipids) were produced following the protocol by Eichmann et al. (235) at a protein-to-lipid ratio of 1:40. Subsequently, the reactions were subjected to size-exclusion chromatography (SEC) using a Superdex 200 10/300GL gel filtration column (GE Healthcare) and eluted at a flow rate of 0.5 ml/min in 20 mM Bis-Tris-HCl pH 7.0 containing 150 mM NaCl.

Negative-Stain Electron Microscopy (EM)

5 µl of Syn lipoprotein particles at a concentration of 10 µM in 20 mM Bis-Tris-HCl pH 7.0, 150 mM NaCl were deposited for 1 min on a glow-discharged carbon-coated copper grid, which was blotted and washed twice with drops of H2O before staining with 2% w/v uranyl acetate for 15 s. EM micrographs were then collected using a FEI Morgagni 268 transmission electron microscope operated at 100 kV.
Supplementary Information

Synuclein nanodiscs formation by various lipids

<table>
<thead>
<tr>
<th>Lipids (Tm, chain length, double bond, net charge)</th>
<th>Lipoprotein particle formation as compared to free monomer</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOPS (Tm -11 °C ) 18:1, -1</td>
<td>93%</td>
</tr>
<tr>
<td>DOPG (Tm -18 °C ) 18:1, -1</td>
<td>92%</td>
</tr>
<tr>
<td>DOPC (Tm -17 °C ) 18:1, 0</td>
<td>52%</td>
</tr>
<tr>
<td>(Δ9-Cis) PC (Tm -4 °C) 14:1, 0</td>
<td>49%</td>
</tr>
<tr>
<td>DMPG (Tm 23 °C ) 14:0, -1</td>
<td>28%</td>
</tr>
<tr>
<td>DMPC (Tm 24 °C ) 14:0, 0</td>
<td>23%</td>
</tr>
</tbody>
</table>

Table S3.1: Effect of lipids on the amount of nanodiscs formation.

Both presences of negative charge and unsaturated hydrophobic lipids were found to be important in nanodiscs formation. Loosing either of this (charge or degree of unsaturation) was losing the formation of nanodiscs by 50%. PC head group was either making big nanodiscs or there was some aggregates formation because it was eluting earlier than others. Hydrophobic tail length does do not has a significant effect on the formation of amount of nanodiscs.
Chapter 4

Structural insights into the monomer-fibril interaction of α-synuclein

Abstract

Protein aggregation into amyloid fibrils is associated with several dozens of diseases including Parkinson’s disease. In the aggregation kinetics, the free monomer gets recruited onto the fibrils via secondary nucleation and elongation, which has been recognized as an important aggregation accelerator. Here, we investigated the secondary nucleation mechanism of α-synuclein by studying the interaction of monomeric α-synuclein with its fibrillar counterpart at two different pH (6 and 7). It was demonstrated that positively charged N-terminal segment of monomeric α-synuclein interacts transiently in micro to millisecond timescale with the negatively charged flexible C-terminal end of the fibril at pH 7.0. The transient interaction between fibrils and monomer is electrostatic which decreases upon addition of salt (NaCl). However, at pH 6.0, the interaction between monomer and fibrils is much stronger and was not altered upon the addition of salts. Moreover, interaction at pH 6.0 also leads to faster
aggregation even at $10^\circ$C without air-water interface. Therefore, it was concluded that transient interaction enhanced the local concentration of $\alpha$-synuclein monomers, which further helped to align several monomeric $\alpha$-synuclein molecules close in space in a parallel fashion yielding oligomer that are well prepared for aggregating into amyloids.

**Keywords**

$\alpha$-synuclein, protein aggregation, amyloids, Parkinson’s disease, NMR, secondary nucleation, aggregation kinetics.

**Introduction**

Synucleinopathies including Parkinson's disease (PD) and several other neurodegenerative diseases are associated with the accumulation of intracellular neuronal aggregates termed as Lewy bodies and Lewy neurites that contain a high concentration of the protein $\alpha$-synuclein ($\alpha$-Syn) in an aggregated state (239). The disease-relevant role of $\alpha$-Syn is further highlighted by mutations in the $\alpha$-Syn gene (SNCA) that cause familial PD (i.e. A30P (4), E46K (5), H50Q (7), G51D (6), A53E (8), and A53T (2), and the duplication or triplication of the gene encoding $\alpha$-Syn that lead to early-onset PD in affected families. $\alpha$-Syn comprises 140 amino acid residues. The N-terminus of $\alpha$-Syn (residues 1-60) is rich in lysine residues and contains KTK lipid-binding motif repeats associated with vesicle binding (50, 226). It is also the region that contains all known SNCA familial PD mutations. The central region (residues 61-95) is the non-
amyloid-ß component (NAC region) (15), which is essential for α-Syn aggregation while the C-terminus (residues 96-140) is rich with highly negatively charged residues (65).

α-Syn is an intrinsically disordered protein in solution (25, 47) forms a helical structure upon membrane interactions and several polymorphs of cross-β-sheet motifs when form aggregates (69, 65, 63). The in vitro-prepared fibrils induce PD-like pathology when injected in the animal model. In concert with the PD pathology described above, Synucleinopathies belong to the class of amyloid diseases for which it is stated that the protein aggregation path is the key for the disease development. A detailed analysis of the aggregation kinetic pathway of α-Syn into amyloids is therefore important towards understanding the toxic mechanisms relevant for Synucleinopathies.

α-Syn amyloid formation is sensitive to solution conditions including pH (90), temperature (96), and salt concentration (91) and requires the presence of an interface such as the air-liquid interface (105) or membrane interaction (58). Amyloid fibril formation by α-Syn occurs through a nucleation-dependent polymerization reaction (71, 99, 75), where the formation of the primary nucleus takes longer but further secondary nucleation via α-Syn monomer addition to the pre-formed aggregates (seeds) does not take long time. Furthermore, there are two secondary processes: (i) amyloid fibrils multiply by fragmentation and (ii) amyloid fibrils catalyze the formation of new amyloids from monomer onto their surface. The latter process was described by Ferrone and
Cohen in the case of sickle cell anemia already 40 years ago. Both processes are fibril mass-dependent and therefore may give rise to highly accelerated aggregation kinetics. During α-Syn aggregation under the quiescent condition, there is no fragmentation and thus the secondary nucleation process and elongation may occur. Indeed, detailed kinetics reveals that secondary nucleation is more prominent below pH 6 (90, 99). To elaborate on the structural properties of this secondary nucleation process, α-Syn monomer - fibril interactions were studied by solution-state nuclear magnetic resonance spectroscopy (NMR). The studies reveal that α-Syn monomers align with each other by transient binding via their positive charged N-terminus to the negatively charged C-terminus of the fibril being well prepared for aggregation.

Results and Discussion

Faster aggregation of α-Syn at pH 6.0 than 7.0 in the presence of pre-formed fibrils
Around half of the monomer, protein was found to be aggregated at pH 6.0 however there was no aggregation at pH 7.0 (Figure S4.1). For aggregation, 250 µM monomer α-Syn was incubated with 1.35 mM preformed fibril (Fibrils were not sonicated, to prevent the elongation process and observe secondary nucleation as the majority) in 5 h at 10 °C in 25 mM Sodium phosphate without salt and without air-water interface (AWI).
Transient interaction of α-Syn monomers with fibrils

To study the interaction between α-Syn monomer and its fibrils, the experimental conditions were kept constant as discussed above. The temperature of 10°C was selected to slow down the aggregation process and to enhance the signal intensity of the NMR experiment (by limiting the HN exchange). Freshly prepared 15N labeled α-Syn monomer solution was added to approximate 3 weeks old mature unlabeled fibrils of α-Syn (grown under 25 mM Sodium Phosphate, pH 7.0, with no salts) in various ratios where the concentration of monomer was kept at 250 µM. Immediately after preparation a [15N,1H]-HSQC spectrum was measured (with various fibrils concentration; Figure S4.2). The intensities of the cross-peaks were compared between a sample containing 1350 µM fibrils and a sample reference lacking any fibrils (Figure 4.1). The results illustrate sequence-specific signal loss at pH 6 and pH 7 in the presence of fibrils. Mostly, the N-terminal residues (1-60) are influenced by the presence of fibrils at acidic pH 6, while the entire sequence is influenced at pH 7.0 These findings indicate transient interaction between the N-terminus of α-Syn monomer with the fibrils. The 15N R2 relaxation data shown in Figure 4.2 and Figure S4.3 further strengthen these findings and indicate that the transient nature of the interaction is in the micro to the millisecond time range. The R2 relaxation measurements were further done with different fibril and monomer concentration with a maximal fibril concentration of 1800 µM showing a saturation-like effect (Figure 4.2). This titration study allowed to determine the Michaelis Menten constant of binding to be ~ 1 mM (Figure 4.2C).
Figure 4.1: $\left[ ^{15}N, ^1H \right]$-HSQC Intensity difference in the absence and presence of fibrils at pH 7.0 and pH 6.0. Full-length (all domain) of monomer protein was found to be interacting with fibrils at pH 7.0 however only the N-terminal domain was interacting with fibrils in case of pH 6.0.
Figure 4.2: Transient interaction of 250 µM α-synuclein monomer with various concentrations of α-Syn fibrils as evidenced by R2 relaxation rates. The Value of R2 increased mostly at the N-terminal with a high concentration of fibrils indicate transient interaction of the N-terminal region to fibrils. (A) pH 7.0 (B) pH 6.0 (C) Binding kinetics of α-Syn monomer to its fibril at pH 7. The R2 values of Leu8 were fitted by a hyperbolic binding function $r = a + b*n[\alpha\text{Syn fibril}]^n/(K_d + [\alpha\text{Syn fibril}]^n)$ yielded the Kd value of 0.78 mM.
Interaction of monomer with the fibrils observed by EM
Further, we recorded negative-stained electron microscopy images (Figure 4.3) of the fibril interaction. While at neutral pH the fibrils are clean and well defined, but when incubated with monomer at pH 7.0 and 6.0, the fibrils are decorated with newly formed aggregates in line with the secondary nucleation concept where α-Syn monomers bind transiently to the fibril.

![Figure 4.3: Negative-stained EM images of α-Syn fibrils showing surface material](image)

(A) α-Syn fibrils in the presence of α-Syn monomer pH 7.0, (B) pH 6.0 (C). The presence of premature small aggregates was found on the surface of aggregates as expected for secondary nucleation.

Mutagenesis-based structural investigations of the interaction of α-Syn monomers with the fibrils
In order to elaborate on the structural nature of the interaction between α-Syn monomers and fibrils a low-resolution investigation using mutants with an anticipated interference effect on the monomers-fibril interaction have been performed based on the hypothesis that the intermolecular contact is based on charged-charged interactions between the positively charged residues on the monomer with the
negatively charged C-terminus on the fibril side. The C-terminal domain is highly flexible in the known 3D structures of all the polymorphs determined by cryo-EM (69). Hence, a truncation variant of \( \alpha \)-Syn (1-121) lacking the last 19 residues was produced and fibrilized at pH 7.0. Following the approach presented above, \(^{15}\text{N,}^{1}\text{H}\)-HSQC and \(^{15}\text{N}\) R2 relaxation experiments were performed on \(^{15}\text{N}\)-labeled wild type \( \alpha \)-Syn at pH 7.0 in the presence and absence of the \( \alpha \)-Syn (1-121) fibrils (Figure 4A). The lack of distinct differences as observed in the case of wild-type \( \alpha \)-Syn fibrils indicates the absence of significant transient interactions of wild-type \( \alpha \)-Syn monomer with the fibrils comprising the variant \( \alpha \)-Syn (1-121). Altogether these findings support the hypothesis that the N-terminal positively charged segment of monomeric \( \alpha \)-Syn interacts transiently with the flexible negatively charged C-terminal segment of \( \alpha \)-Syn fibrils.

**Effect of addition of 100 mM NaCl on monomer-fibril interaction**

To understand the type of interaction between monomer and fibrils, 100 mM NaCl was added to the monomer-fibril sample at both pH values (6 and 7). The loss of the N-terminal enhanced relaxation in the presence of NaCl shows that at pH 7.0 the transient interaction is mostly based on electrostatic nature. However, at pH 6 the transient interaction was not broken by the addition of NaCl. Monomer and fibrils are interacting more tightly at pH 6.0 than at pH 7 in the presence of salt (Figure 4A, 4B).
Figure 4.4: α-Syn monomers interact with its positively charged N-terminal segment to the flexible C-terminal part of the negatively charged fibrils. (A) $^{15}$N $R_2$ relaxation measurement of $^{15}$N-labeled α-Syn monomers (250 µM) in presence of the C-terminal truncated construct α-Syn (1-121) fibrils (1350 µM) relative to wildtype α-Syn fibrils at pH 7.0 (B) and (C) Effect of the addition of 100 mM NaCl on the interaction of fibrils with monomer at (B) pH 7.0 and (C) pH 6.0.
Conclusion

Model 4.1: Electrostatic interaction between monomer and fibrils at pH 7.0. Positively charge N-terminal of monomer interacts with the negatively charged C-terminal domain of fibrils.

Taken together, our NMR data indicate that in vitro α-Syn monomer interacts with fibrils at both pH 6 and 7 transiently through electrostatic forces between the positively charged N-terminal segment of the monomer with the negatively charged flexible C-terminal part of the fibrils. While the addition of 100 mM NaCl interferes with these transient interactions at pH 7, at pH 6 they are still preserved.
Material and Methods

**Preparation of samples**

**Protein expression and purification:** Recombinant acetylated full length and C-terminal deleted (1-121) acetylated α-Synuclein were expressed using the periplasmic purification protocol as described earlier (211). Briefly, α-synuclein (pRK172) plasmids were co-expressed with N-terminal acetyltransferase B (NatB) complex in *E. coli* BL21 (DE3 STAR), further colonies containing both plasmids (NatB and pRK172) were used for expression of acetylated α--Synuclein in Lysogeny Broth (LB) medium (in case of unlabeled protein) and M9 medium containing N\textsubscript{15} ammonium chloride and C\textsubscript{13} glucose (with 80% D\textsubscript{2}O in case of isotopically labeled) at 37 °C. After growing till OD\textsubscript{600} of 1.0 it was induced with 1mM IPTG for 5-6 h. Bacterial cells were harvested and proteins were extracted by osmotic shock methods. Further protein was purified using ion-exchange chromatography (Hi Trap Q FF column, GE Healthcare) and hydrophobic interaction chromatography (Hi Prep Phenyl FF as described earlier (105). Protein was dialyzed in milli-Q water and stored at -80°C after lyophilized for storage.

**Amyloid fibril preparation both at pH 6.0 and 7.0:** 500 μM (1 ml in 1.5 ml Eppendorf tubes) of unlabeled α-Synuclein monomer was incubated at 37°C in 25 mM sodium phosphate buffer at pH 7.0 with 0.05% azide and
kept on orbital rotation for 2-3 weeks. Aggregation condition was the same for C-terminal truncated (1-121) α-Synuclein.

**Negative-Stain Electron Microscope (EM):** To observe the monomer interaction with fibrils, 30 µM fibrils were incubated with 60 µM monomers in 300 ul of sodium phosphate buffer of pH 6.0 and pH 7.0 at 4°C for 18h. Further 10 ul of the sample was deposited on carbon film 150 square mesh copper grids CF150-Cu) for 1 min and then washed twice with Milli-Q water further stained with 2% Uranyl acetate. EM micrographs were then collected with a FEI Morgagni 268 transmission electron microscope operated at 100 kV.

**NMR experiments:** All NMR experiments were performed on a Bruker 600 MHz Avance III HD spectrometer equipped with a cryogenic probe at 10 °C. The software Sparky 3.115 (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco, USA) and Bruker Topspin 3.5pl7 (Bruker, Inc., Billerica, MA, USA) were used for analyzing the spectra and extracting the rate constants.

Fibrils were centrifuged in 5417R bench top Eppendorf centrifuge with 20,000 rpm for 5 minutes, the pellet was resuspended and centrifuged in 25 mM sodium phosphate buffer, pH 6.0 and pH 7.0 as required. Further fibrils (unlabeled) was mixed with the monomer (labeled) in the different ratio on ice to prevent fast aggregation.250 µl of total volume (Fibrils and monomer) was filled in Shigemi tube and R₂ was measured with a
modified pulse sequence as described earlier (Chapter 2) with 4% D$_2$O. HSQC measurement was performed before and after $R_2$ measurements, to monitor the aggregation process during the relaxation experiment. $^{15}$N $R_{1\rho}$ rate constants were recorded using the NMR experiment described in (201). The magnetization decay was recorded using four different relaxation decay periods, in an interleaved manner (172). $R_2$ relaxation delays were set to 0 ms, 200 ms, 100 ms, and 50 ms, for the $^{15}$N $R_{1\rho}$ experiments, delays were set to 1 ms, 120 ms, 60 ms, and 20 ms, respectively and for $^{15}$N $R_1$ experiments delays were 0 ms, 600 ms, 400 ms, 200 ms. The spin-lock RF field strength in the $^5$N $R_{1\rho}$ measurement was set to 2 kHz.
Supplementary Information

**Figure S4.1:** [\(^{15}\)N,\(^1\)H]-HSQC Intensity decrease with time in the presence of fibrils. Intensity decrease indicates the formation of aggregates in the presence of 1350 \( \mu \)M fibrils at 10 °C in the absence of AWI (A) pH 6.0 (B) pH 7.0 (C) loss of Intensities after 5h at both pH values. At pH 7.0, protein does not aggregate in the presence of fibrils however at pH 6.0 half of the protein form aggregates in 5h.
Figure S4.2: $[^{15}\text{N},^{1}\text{H}]$-HSQC spectra of $^{15}\text{N}$-labeled $\alpha$-synuclein in the presence of fibrils. At both pH 6.0 and 7.0, the addition of more fibrils leads to the disappearance of those peaks which were interacting with fibrils. A few peaks which have major reduction has been shown by the blue arrow.

Figure S4.3: Relaxation measurement of 250 $\mu$M synuclein monomer with 1200 $\mu$M Fibrils at pH7.0 $R_{1p}$ with spin lock 2 kHz, refocusing Rex dynamics slower than 80 $\mu$s contributions. Elevated $R_2$ exchange results from conformational exchange on micro- to milliseconds timescale.
Chapter 5

α-Synuclein interaction with various lipids containing nanodiscs and its effect on aggregation modulation

Abstract

α-Synuclein (α-Syn) is involved in Parkinson’s Disease (PD) and many other neurodegenerative diseases. The interaction of α-Syn with the membrane is suggested to play an important role in both (α-Syn) function and disease pathology. The α-Syn interaction to the membrane has been suggested to disrupt the membrane and propagate the toxic aggregates from one cell to another. Exosomes containing ganglioside lipids (GM1 or GM3) were found to be involved in aggregation acceleration and transmission of toxicity. In this study, the properties of the lipids which were important for α-Syn-lipid interaction and in the modulation of aggregation were investigated. Various membrane scaffold proteins (MSP) of different sizes were used to mimic the high-density lipoproteins (HDL) of diameter 6-25 nm. We found that negative charge, high degree of unsaturation and membrane fluidity were the important properties for lipids to interact with α-Syn. The degree of lipid saturation was found to
modulate α-Syn aggregation. Saturated lipids were inhibiting aggregation while unsaturated lipids were inducing. HDL like particles were prepared with mixed lipids (10% GM3, 90% DOPC) similar to exosomes composition. N-terminal residues (1-60) of α-Syn were found to be involved in lipid interaction. Hydrophobic residues followed by charged residues from the N-terminal were found to have the maximum interaction with the membrane. This finding supports the importance of exosomal lipids for understanding α-Syn pathology. This reveals the reason for aggregation induction and membrane disruption by exosomes (GM3) lipids. We also showed that nanodiscs can be successfully uptaken by cells. Hence, this provides the potential for further in-cell studies of the α-Syn lipid interactions.

**Keywords**

Lipids; 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC); 1,2-dioleoyl-sn-glycero-3-phospho-(1′-rac-glycerol) (sodium salt) (DOPG); 1,2-dimyristoyl-sn-glycero-3-phospho-(1′-rac-glycerol) (sodium salt) (DMPG); 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC); 1,2-dioleoyl-sn-glycero-3-phospho-L-serine (sodium salt) (DOPS); Ganglioside lipids (GM3). Proteins MSP: membrane scaffold protein; MSPD1 (full length with half of the first helix); MSPDH5 (deleted helix 5 in MSPD1), MSPDH4-H6 (deleted helix 4,5 and 6 in MSPD1), MSPDH5 R2 (MSPDH5 and additionally 2 times H4 and H6); α-Synuclein (α-Syn). High-density lipoprotein (HDL)


Introduction

The interaction of α-Syn with the membrane has a physiological function; it also modulates aggregation and propagation of disease by membrane disruption. α-Syn transforms from an unfolded native monomer structure to an α-helical secondary structure upon binding with negatively charged lipids (27). The presence of amphipathic repeats (KTKEGV) in the N-terminal domain is mainly responsible for lipid binding. PD related familial α-Syn mutations affect binding affinity to the membrane. A30P α-Syn mutant was found to have less affinity for the membrane whereas A53T and E46K mutations led to increased affinity to the membrane (57). Furthermore, lipid binding and aggregation modulation of α-Syn depend on the physical and chemical properties (Physical state, $K_D$, $T_m$, acyl chain length) of lipids (58). Moreover, exosomes have been involved in cell to cell transmission of α-Syn and acceleration of aggregation by nucleation formation. In this study, we show how α-Syn interaction with the membrane depends on lipid properties and the effect of the interaction on the modulation of aggregation.

Results and discussion

Structural change of α-Syn upon lipid interaction

Many lipids and MSP (of different length) were used to make HDL particles of various sizes (Table 5.1). Conformation of α-Syn modulated by interaction with lipid containing nanodiscs was sampled in listed
conditions with the help of TROSY NMR. There was no condition exhibiting any chemical shift perturbation, but some properties of lipids were found to be regulating the interaction with α-Syn.

<table>
<thead>
<tr>
<th>Conditions for synuclein-lipid interaction</th>
<th>Syn ND (25 nm), MSPDH5 R2 ND (12 nm), MSPD1 ND (10 nm), MSPDH5 ND (8 nm), MSPDH(H4-H6) (6 nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Different size of Nanodiscs (ND)</td>
<td></td>
</tr>
<tr>
<td>Lipids</td>
<td>Phospholipids Charge Degree of saturation Tm (°C)</td>
</tr>
<tr>
<td></td>
<td>DMPG (14:0)  -1 Saturated 23</td>
</tr>
<tr>
<td></td>
<td>DMPC (14:0)  0 Saturated 24</td>
</tr>
<tr>
<td></td>
<td>DOPS (18:1)  -1 Unsaturated -11</td>
</tr>
<tr>
<td></td>
<td>DOPC (18:1e9)  0 Unsaturated -17</td>
</tr>
<tr>
<td></td>
<td>DOPG (18:1)  -1 Unsaturated -18</td>
</tr>
<tr>
<td>Sphingolipids</td>
<td>Ganglioside GM3  -1 Unsaturated not calculated</td>
</tr>
<tr>
<td>Temperature</td>
<td>37 °C, 30 °C, 25 °C, 10 °C</td>
</tr>
</tbody>
</table>

**Table 5.1: Experimental conditions.** The diameter of nanodiscs is shown in the table. Several lipids were used with different physical and chemical properties. Different temperatures were tried to observed the effect of phase change of lipids on Syn interaction.

**Charge of lipids is essential for interaction with α-Syn**

The positively charged N-terminal residues of α-Syn were observed to interact with nanodiscs containing negatively charged lipids (DOPS). This interaction was observed based on the peak intensity reduction in HSQC spectra with and without nanodiscs. However, no interaction could be observed for zwitterionic lipids (DOPC and DMPC). This shows that the presence of charge is critical for α-Syn affinity to the membrane. Measurement was performed with non-acetylated α-Syn, at 30°C with 100 mM NaCl (Figure 5.1). NMR relaxation measurements were performed to know the dynamics of α-Syn-lipid interaction.¹⁵ N R1p
relaxation measurement revealed α-Syn interacts with lipid in μs-ms timescale. In the case of DOPS-lipids, all residues of the N-terminal domain (1-60) interact whereas in the case of DOPC-lipids only half of the N-terminal domain (1-30) interacts (Figure 5.2).

**Modulation of α-Syn aggregation in the presence of nanodiscs**

The aggregation kinetics was monitored using Thioflavin T (ThT) dye in the absence and presence of various lipids containing nanodiscs. ThT fluorescence increases upon binding to cross β structure. Its fluorescence is often used for the observation of aggregation kinetics. Unsaturated lipids, both without and with charge (DOPS and DOPC), were able to induce aggregation whereas saturated lipids (DMPC and DMPG) have shown the inhibitory effect on aggregation (Figure 5.3A). Total amount of fibrils varied significantly among samples. Saturated lipids were able to inhibit the aggregation efficiently even when aggregation was accelerated with agitation at 500 rpm (Figure 5.3B). These results were verified by Electron microscopy (EM). Fibrils were detected in the presence of unsaturated lipids containing sample whereas no fibrils were observed with saturated lipids (Figure 5.3C).

**Uptake of nanodiscs by cells**

To observe the modulation of aggregation inside the cells by various lipids containing nanodiscs, it was important to observe if nanodiscs can be taken up by cells. To that extent, MSPDH5 protein was labeled with an iridium complex ([Ir(ppy)2(H2O)2] OTf, excitation wavelength 365 nm).
The presence of green color by confocal microscopy is confirming that nanodiscs were inside the COS-7 cell lines (Figure 5.3D).

The phase change of lipids alters the nanodisc interaction with α-Syn
Each lipid bilayer has its own melting temperature (Tm) (Table 5.1). For the temperatures below the Tm, membranes will be mainly in solid-phase. For the temperatures above the Tm, it will be mostly in fluid form. To observe the effect of the phase of the membranes on the interaction of α-Syn different lipids were used at two different temperatures. DOPG has Tm of -18°C, and was then in fluid form at both 10°C and 30°C. DMPG has a Tm of 23°C, hence it was in solid-phase at 10°C and in fluid-phase at 30°C. The interaction was observed on HSQC (\(^{15}\)N, \(^1\)H) spectra by the disappearance of the peaks. Only the lipids in the fluid form were able to interact with α-Syn (Figure 5.4). Presumably the hydrophobic area is more exposed in the fluid form as compared to the solid form which was also discussed before (58). Charge of lipid is an essential property for α-Syn interaction.

In this study, as compared to previous (Figure 5.1), α-Syn was N-terminal acetylated, NaCl concentration was set to 25 mM and zwitterionic lipids were used.\(^{15}\)N-HSQC spectra were measured to observe intensity decrease (or disappearance) of the α-Syn N-H signals. DMPG and DMPC lipids were compared for the interaction at 30°C (above Tm) and it was observed that charge is essential (Figure 5.5).
α-Synuclein interaction with various lipids containing nanodiscs and its effect on aggregation modulation

Figure 5.1. HSQC (15 N, 1H) Spectra of α-Syn in the presence of various lipids-containing MSPDH5 nanodiscs. Mostly N-terminal residues of α-Syn were found to be interacting with negatively charged (DOPS) lipids. Residues are highlighted in the α-Syn spectrum which decreased the intensity upon interaction with DOPS. There was no interaction when lipids were zwitterionic.
Figure 5.2. 15N transverse (R1ρ) relaxation rates of α-Syn-lipid interaction. α-Syn alone (white) with DMPC (green), with DOPC (blue), with DOPS (red). Spin lock frequency was 2 kHz, refocuses Rex dynamics slower than 80μs contributions so interaction observed here are in μs-ms timescale.
α-Synuclein interaction with various lipids containing nanodiscs and its effect on aggregation modulation

**Figure 5.3. Aggregation of Syn in the presence of various lipid-containing nanodiscs and uptake of nanodiscs (ND) inside the cells.** Thioflavin T fluorescence measured in a plate reader with the static condition (A) in Eppendorf with manual measuring at Fluorometer with mixing at 500 rpm to facilitate the aggregation (B). EM images were taken at the end of the aggregation (C). Confocal image of cells after uptake of protein and nanodiscs. Green color shows the nanodiscs inside the cells (D).
Figure 5.4. Comparison of HMQC ($^{15}$N, $^1$H) peaks of acetylated $\alpha$-Syn with various lipids containing MSPDH5 R2 nanodiscs at different temperatures. The loss of the NMR peaks for residues suggests its interaction with lipids. $\alpha$-Syn interacts with DMPG at 30$^\circ$C which is above the Tm (23$^\circ$C) however it does not interact below Tm (10$^\circ$C). $\alpha$-Syn interacts with DOPG at both temperatures (10$^\circ$C, 30$^\circ$C) since the Tm of DOPG is low -18$^\circ$C and it will in fluid form at both temperatures.
α-Synuclein interaction with various lipids containing nanodiscs and its effect on aggregation modulation

**Figure 5.5.** Comparison of HMQC (15 N, 1H) peaks of acetylated α-Syn with various lipids containing MSPDH5 R2 nanodiscs at 30°C. The loss of the NMR peaks for residues suggests its interaction with lipids. α-Syn is interacting with DMPG which contains the negatively charged head group, where the interaction is present whereas interaction is not present with DMPC which contains zwitterion.

**Attempt to observe the change of secondary structure**
Many different conditions were tried to observe the structural change in the α-Syn residues upon lipid interaction (summarized in Table 5.1). However, there was no success in the observation of any structural change within the NMR spectra. Suspected reasons were the big size of nanodiscs (around 100 kDa of MSPDH5) make tumbling rate slower and shorter T2 relaxation time leading to vanishing NMR signal.
Spectra were acquired at high temperatures (37°C) in order to make tumbling faster but still, no new peak was observed, which would have witness of slow conformational exchange (Figure 5.6). Furthermore, 8.5% Trifluoroethanol (TFE) was used to stabilize the secondary structure on the membrane. TFE was chosen as a stabilizer because it has been found to stabilize the partially helical monomer α-Syn conformation.

Figure 5.6. TROSY spectra of α-Syn interacting with DMPG lipid-containing nanodiscs at 37°C. Different size of nanodiscs were used in order to observe the structural change after binding to nanodiscs.
Stabilization by 8.5% TFE led to disappearing of the residues on the HSQC spectra. When more than 8.5% TFE was used, a visible precipitation appeared, and spectrum could not be measured (data not shown).

**α-Syn bound to nanodiscs diffuses similarly to nanodiscs alone**

2D Diffusion ordered spectroscopy (DOSY) measurement reveals that the diffusion coefficient of α-Syn bound to MSPDH5 nanodiscs (blue) is similar to the diffusion coefficient of MSPDH5 nanodiscs (green), thus suggesting that α-Syn binds to nanodiscs. The diffusion coefficient of α-Syn monomer without binding is significantly lower. (Figure 5.7).

![Figure 5.7. 2D Diffusion experiments at 303 K with MSPDH5 (DMPG) nanodiscs with 50 mM NaCl, pH 7.0. Blue line: $^1$H coupled to $^{15}$N (bound α-Syn) has diffusion coefficient of $5.7 \times 10^{-5}$ cm$^2$s$^{-1}$, green line: $^1$H not coupled to $^{15}$N (only MSPDH5) has diffusion coefficient of $5.4 \times 10^{-5}$ cm$^2$s$^{-1}$ and black line: α-Syn monomer without nanodiscs has diffusion coefficient of $8.1 \times 10^{-5}$ cm$^2$s$^{-1}$.](image)

**Interaction of α-Syn with 10% GM3 and 90% DOPC lipids containing nanodiscs**

Hydrophobic residues followed by positively charged residues of the N-terminal domain interact strongly with lipids as it is shown by intensity changes (Figure 5.8). This suggests that it is important to have the weak
interaction between the positively charged N-terminal residues and the lipids (anchoring) which can bring the hydrophobic residues closer to lipids for strong interaction or insertion inside the lipid bilayer. It is not yet clear which type of structure they adopt since no additional/shifted peak was observed in $^{15}$N-HSQC spectra.

**Figure 5.8.** α-Syn interaction with MSPD1 nanodiscs containing 90% DOPC and 10% GM3 lipids

A) Lipids structure  
B) HSQC Intensity change ratio ($I/I_0$) in the presence and absence of nanodiscs (full length)  
C) Residues level analysis for the major change in the initial 50 residues of N-terminal domain.
α-Synuclein interaction with various lipids containing nanodiscs and its effect on aggregation modulation

Lipid composition
90% DOPC + 10% GM3

Model 5.1: Showing the binding of α-Syn on the nanodiscs.

In order to observe the structural change for the residues that interact with lipids, TROSY-HSQC spectrum can be measured at relatively high temperature with a smaller size of nanodiscs. Different ratios of charged and zwitterionic lipids can be tried to find the optimum interaction to find the structural change.

Conclusion

α-Syn -lipid binding and its effect on aggregates formation were studied using several lipids (different physical and chemical properties). Many conditions (presence of different concentration of salts, temperature and
different sizes of nanodiscs) that can modulate the α-Syn-lipid interaction were screened. α-Syn was observed to have the diffusion value similar to nanodiscs which confirms its complexation with nanodiscs.

α-Syn-lipid interaction was more pronounced with the negatively charged lipids, a higher degree of unsaturation.

Different lipids containing nanodiscs were found to modulate aggregation. Saturated lipids (DMPG and DMPC) were found to be a strong inhibitor however unsaturated lipids (DOPS and DOPC) were found to be an inducer of α-Syn aggregation. Less packed arrangement of unsaturated lipids can make it easier for the insertion and accumulation of proteins which can lead to faster aggregation.

N-terminal of α-Syn domain were found to be involved in interacting with nanodiscs with mixed lipid composition (10% GM3 and 90% DOPC). This interaction might lead to an increase of the local concentration of monomer in/on the membrane and in the acceleration of aggregation. Furthermore, studies can be performed inside the cells, since nanodiscs were taken inside by COS-7 cells.

**Future direction**

Perdeuterated lipids and deuterated apolipoprotein can be used to make nanodiscs for reduction of the relaxation rates. Furthermore, smaller construct of the nanodiscs (MSPDH4H5H6/MSPDH5) can be produced in
order to slow down the relaxation of the magnetization. Mixed lipids (with charged lipids) can be used for the intermediate binding. Charged lipids are required for the essential anchoring of α-Syn. Charged membrane surface can inhibit the insertion of hydrophobic residues. High temperatures up to 40°C can increase the tumbling. These steps can increase the possibility to observe the secondary structure change of α-Syn upon binding to lipids.

Material and Methods

α-Syn Purification

Recombinant acetylated and not acetylated full length (1-140) α-Syn were expressed using the periplasmic purification protocol as described earlier (211). Briefly for expression of acetylated α-Syn, α-Syn and N-terminal acetyltransferase B (NatB) plasmids were co-transformed in *E. coli* BL21 (DE3 STAR) and for nonacetylated protein only α-Syn plasmids were transformed. Further colonies containing both plasmids (NatB and Syn) were selected for the expression of acetylated α-Syn in the Lysogeny Broth (LB) medium in case of unlabeled and N<sup>15</sup> ammonium chloride containing M9 medium in case of isotopically labeled. When OD<sub>600</sub> reached 1.0 expression was induced with 1 mM IPTG for 5-6h. Bacterial cells were harvested, and proteins were released by osmotic shock method. Further protein was purified using ion-exchange chromatography and hydrophobic interaction chromatography as described earlier (105).
Expression of the various construct of ApolipoproteinA1 (ApoA1)

ApoA1 (MSPD1, MSPDH5, MSPDH4H5H6, MSPDH5 R2) genes of various length was inserted in a pET28a vector containing N-terminal His6-tag and a TEV (Tobacco Etch Virus) protease recognition site (ENLYFQ), which was received from addgene (Cambridge, US). Briefly, plasmids were transformed into E. coli strain then plated on kanamycin antibiotic containing LB-Agar plate. Few colonies were added in 100 mL of LB (preculture) followed by 1-liter TB-Media and grown at 37°C until the (OD600) reached 0.8. Afterward protein expression was induced using 0.5 mM IPTG. After 1h, the temperature was lowered to 28°C for another 4-5h and then the cells were harvested by centrifugation for 30 minutes at 4000 g. Pellet was resuspended in 20 mM Tris pH 8, 500 mM NaCl and to prevent protein degradation one protease inhibitor tablet/100mL buffer was added to the solution and mixed for 1h at 4°C. After mixing cell were broken by passing three times through microfluidizer. Then cellular parts were separated from soluble protein by high-speed centrifugation 36900 g (20000 rpm) for 20 minutes. Prepacked Histrap TM excel (Ni-column) was used to isolate ApoE protein (contain His-tag). Elution was done using 300 mM imidazole. Imidazole was removed by using a PD-10 desalting column. Further, if it was required to concentrate, Amicon Ultra Centrifugal Filter (10 kDa 15 ml) was used. Protein was stored at -20°C. For isotope-labeled sample M9 media was used (containing $^{15}$N ammonium chloride and $^{13}$Cglucose) (in case of isotopically labeled) in place of LB media rest purification protocol remain the same.
Formation of HDL (nanodiscs) particle

Lyophilized lipids (powder) were bought from Avanti Polar Lipids Inc. and were kept in -20°C. 100mM lipids stock was made in 200 mM Sodium Cholate, the ratio of lipid to cholate was always kept 1:2, these stocks were made in a glass vial. Further solubilization of lipids was done by freezing (with liquid Nitrogen) and thawing (water bath of 60 °C). For the formation of nanodiscs, the various ratio of lipid to protein was tried for the single uniform peak of nanodiscs in size exclusion chromatography. Briefly different amount of protein and lipids were mixed (cholate was always 2X of lipids) and incubated at room temperature for overnight with gentle mixing, after overnight mixing 100% (w/v) Bio Beads SM Resin was added for 2h and then it was centrifuged to separate Bio Beads and aggregates if formed. The soluble fraction was injected to Superdex 200 size exclusion chromatography (SEC) to isolated pure fraction of nanodiscs. Further nanodiscs were concentrated by Eppendorf Amicon of 50kDa.

Thioflavin T (ThT) measurement of the fibril formation

1 ml of 300μM α-Syn was incubated with various concentration of nanodiscs at 37°C with 100 mM NaCl, pH 7.4 and 0.05% Sodium azide (NaN₃). ThT measurement was done by both manual (Fluoromax-4 Spectrofluorometer, Horiba Jobin) and automated (Plate reader, PHERAstar plus BMG LABTECH) fluorescence measurement. In plate reader the aggregation process was slower. However, when eppendorf tube was
agitated with 500 rpm the aggregation was faster. For the manual measurement 10 μl of sample was taken out from the Eppendorf at each time point and mixed with 490 μl of 10 μM of ThT. For the automated plate reader, ThT was prior to the experiment in ThT: protein 1:20 ratio.

**Negative-Stain Electron Microscope (EM)**

10 μl of the aggregated α-Syn sample was deposited at on carbon film 150 square mesh copper grids CF150-Cu) for 1 min and then washed twice with Milli-Q water further grid was stained with 2% Uranyl acetate. EM micrographs were then collected with a FEI Morgagni 268 transmission electron microscope operated at 100 kV.

**NMR experiments**

All NMR experiments were performed on a Bruker 600 MHz Advance III HD spectrometer equipped with a cryogenic probe at 10 °C. The software Sparky 3.115 (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco, USA) and Bruker Topspin 3.5pl7 (Bruker, Inc., Billerica, MA, USA) were used for analyzing the spectra and extracting the rate constants.

HSQC for monomers and TROSY-HSQC for nanodiscs were acquired at different temperatures with various lipids containing nanodiscs.¹⁵N R₁ experiments delays were set to 1ms, 120ms, 60ms, and 20ms, respectively with the spin-lock RF field strength in the ¹⁵N R₁ measurement was set to 2 kHz.
Chapter 6

Structural and morphological studies of ApoE4 different construct both in free and lipid-bound form

Abstract

The ApoE4 isoform is the most prominent genetic biomarker for Alzheimer's diseases compared to other ApoE isoforms (ApoE3, ApoE2). Therefore, elucidating the structure of ApoE4 in lipoproteins would be beneficial to understand its role in disease pathogenesis. In the current study, different constructs of ApoE4 were purified and a homogenous single population of lipoprotein (nanodiscs) was prepared for the morphological and structural characterization. Similar to well-studied ApoA1 (MSPDH5), two subunits of ApoE4 proteins are involved in one nanodiscs formation, revealed by a cross-linking experiment. Interestingly, 2D NMR spectra revealed there is no significant structural change in the N-terminal domain after nanodiscs formation. After the characterization of N-terminal nanodiscs, full-length ApoE4 nanodiscs were prepared but because of large size, only flexible residues were observed on the spectra. ApoE full length with 5 mutations (M5),
nanodiscs were stable at high temperature and high concentration whereas wildtype ApoE full-length nanodiscs were not stable at high concentrations with various lipids. The structural change with detergent was found to be significantly different than the lipoproteins. This suggests using detergent as lipid mimicking system can mislead the real structural change.

**Keywords**

Apolipoprotein E (ApoE), Lipoproteins (nanodiscs) Alzheimer’s diseases (AD), Mass-spectrometry (MS), Size Exclusion Chromatography (SEC), NMR Experiments, DMPC (14:0, 1,2-dimyristoyl-sn-glycero-3-phosphocholine), DMPG (14:0, 1,2-dimyristoyl-sn-glycero-3-phospho-(1′-rac-glycerol).

**Introduction**

The apolipoprotein E (APOE) gene is located on chromosome 19 and encodes a glycoprotein that consists of 299 amino acids. (~34 kDa protein) (241) Three different major isoforms are present in human ApoE, which are ApoE2, ApoE3, apoE4. These isoforms differ only by one/two amino acid residues at positions 112 and 158, respectively. However, alteration at these two places was suggested to change protein stability, charge and functions significantly. ApoE is the major apolipoprotein of very-low-density lipoprotein (VLDL) (241) ApoE is also involved in many diseases such as Alzheimer’s disease (AD), Parkinson’s disease (PD), cardiovascular
diseases (CVD), multiple sclerosis (MS), type 2 diabetes mellitus (T2DM), vascular dementia (VD), and ischemic (occlusive) stroke (IS) (115). ApoE contains three different domains, N-terminal domain which is connected to the C-terminal domain via the hinge region. Wild-type full-length ApoE protein was found to be tetramer below 10 µM and formed big aggregates at high protein concentration in the absence of lipids (114). Because of the stability issue, the full-length wild-type protein structure was not studied, however, the structure of the N-terminal domain was known for a long time (117). Later on, full-length ApoE stable monomer was prepared after the introduction of five mutations (F257A, W264R, V269A, L279Q, V287E) at the hydrophobic C-terminal (120). While the structural elucidation of the soluble lipid-free ApoE is an important step towards a structure-function relationship of ApoE, we concentrated here on lipided ApoE, since it is the functionally relevant conformational species.

Results and discussion

Sequence similarity within Apolipoproteins
Both ApoE and ApoA belong to the apolipoprotein family. The major difference is the ApoE is the major apolipoprotein of very-low-density lipoprotein (VLDL) and ApoA is the major apolipoprotein of high-density lipoprotein (HDL) of blood plasma. The sequence similarity is shown in Figure 6.1.
Formation of nanodiscs by full-length ApoE4 (1-299) and N-terminal domain (1-191)

Because of the large sequence similarity between ApoE and ApoA (Figure 6.1) the formation of ApoE nanodiscs was followed closely the established procedure for ApoA nanodiscs as described (235). Soluble recombinant, ApoE was mixed to different lipid ratios to obtain a single homogeneous peak.

**Figure 6.1: Sequence similarity of ApoE4 with ApoA1, shows around 40%, sequence similarities. Orange and blue lines (alternate helix) shows the helix (1 to 10) in ApoA1. ApoE contains additional amino acid residues around helix4 which might lead an additional helix in ApoE nanodiscs as compared to ApoA.**

In detail, full-length ApoE monomer with five mutations (M5) to lipid at a ratio of 1:80 resulted in a homogenous single nanodiscs population in the size exclusion column (SEC) whereas 1:40 protein to lipid ratio was required for the N-terminal (NT) domain comprising residues 1-191 (Figure 6.2). In contrast, the C-terminal (CT) domain comprising residues
Structural and morphological studies of ApoE4 different construct both in free and lipidd-bound form

(216-299) was aggregated both in the presence and absence of lipids and was thus excluded for further studies (Figure 6.2).

Figure 6.2: ApoE4 nanodiscs preparation using DMPC lipids by full length (M5), NT domain, CT domain and size estimation on the basis of size exclusion chromatography (SEC) profile. (A) SEC profile (Superdex 200, 10/300 GL) after formation of nanodiscs with various protein to lipid ratio for full length (M5), N-terminal and C-terminal domain (B, C) Expected size (diameter) on the basis of elution volume of SEC.
Nanodiscs contain two subunits of ApoE protein
Similar to the MSPDH5 nanodiscs, both full-length (M5) and NT domain of ApoE4 showed a prominent dimer band in a SDS gel after crosslinking, which suggests two proteins are surrounded to make one nanodisc (Figure 6.3). The CT domain (216-299, M5) could neither form nanodiscs nor it was stable as a monomer.

*Figure 6.3: Cross-linking of ApoE4 and ApoA1 nanodiscs with Disuccinimidyl glutarate (DSG) Lane 1, 2, 3 is Full-length (M5), N-terminal and C-terminal domain of ApoE4 nanodiscs without cross-linking. 4,5,6, is after cross-linked. Result suggests two rings are involved in the formation of one nanodisc of full length and NT domain however CT domain is aggregated. A similar result was seen in MSPDH5 nanodiscs Lane: A without cross-linker Lane B. with cross-linker.*
Nanodiscs formation by NT domain without major structural changes
NT ApoE4 nanodiscs are significantly bigger than NT monomer (Figure 6.4) The spectral quality of 2D $[^{15}\text{N},^{1}\text{H}]$-TROSY change for NT nanodiscs at different temperature (Figure 6.5) also indicates about the size difference. However, no significant structural changes were observed in the $[^{15}\text{N},^{1}\text{H}]$-TROSY spectra of NT monomer versus NT nanodiscs at 316 K (Figure 6.6).

**Figure 6.4:** Characterization of NT (1-191) monomer and nanodiscs of ApoE4. (A) SEC of N-terminal nanodiscs (elutes at 12.65 ml) and N-terminal monomer (elutes at 15.5 ml). (B) 1D $^{1}\text{H}$ NMR Spectra shows the presence of DMPC lipids in the Nanodiscs. (C) Diffusion experiment for the size difference between monomer versus nanodiscs (D) EM images of NT nanodiscs.
Figure 6.5: \([^{15}N,^1H]\)-TROSY spectra of NT monomer and nanodiscs at different temperatures. The quality of spectra decreases for the nanodiscs sample stronger at low temperature (290 K) when compared with the soluble NT monomer because of larger size.

NT domain and CT domain are involved in nanodiscs formation

ApoE contains three domains (NT domain 1-191, hinge region 192-215 and CT domain 216-299). Digestion followed by bottom-up mass-spectrometry revealed amino acid residues 16-206 (by Proteinase K) and 2-251 (by Trypsin) are involved in nanodiscs formation (Figure 6.7) PK is a non-specific enzyme having slightly more preference for hydrophobic
Structural and morphological studies of ApoE4 different construct both in free and lipid-bound form

Figure 6.6: Overlay $[^{15}N, ^{1}H]$-TROSY spectra of NT monomer (red) and NT nanodiscs (blue) at 316K. There is no significant spectral change after the formation of nanodiscs except at the tryptophan (9.9-10.5) region.

residues and trypsin cleave carboxyl end of lysine and arginine, except when either of this followed by proline. ApoE contains 36 Arginine and 12 Lysine residues uniformly throughout the protein sequence. PK digestion suggested mostly N-terminal ApoE is protected (16-206) whereas, trypsin digestion result suggested N-terminal, as well as half of the C-terminal (2-251), are involved in nanodiscs formation.
Figure 6.7: Bottom-up mass-spectrometry analysis of ApoE4 full length (M5) nanodiscs.

(A) Nanodiscs were made with DMPC lipids further was digestion with PK and trypsin, 20 kDa band was found to be stable and protected for both digestive enzymes which was cut and send for the bottom-up (MS/MS) analysis.

Formation of nanodiscs by wildtype full length ApoE4

Wild-type full-length ApoE4 protein is known to form aggregates (tetramer and big aggregates) so it was not possible to isolate as a monomer in SEC (Figure 6.8 a, black). However, the mutation in 5 residues of full-length ApoE (M5) resulted in stable monomer as shown above (Figure 6.8 a, blue). Interestingly, with wild-type full-length ApoE4 protein after nanodiscs preparation, aggregates dissociated and only the nanodiscs peak appeared which was in a similar position to the nanodiscs peak with ApoE full length (with five mutations). The formation of aggregates by ApoE4 wild type did not perturb the formation of nanodiscs. Full length with five mutations, ApoE4 forms a stable
monomer and shows well-resolved peaks in the 2D NMR spectrum (Figure 6.9). However, it was not possible to take a good spectrum of full-length WT (data not shown), as big aggregates were formed immediately and did not appear in the NMR signal.

**Figure 6.8: Peaks of ApoE4 WT versus with ApoE4 M5.** a) Before nanodiscs preparation ApoE4 (WT) shows aggregates b) After nanodiscs formation peak elutes at the same volume 11.6 for both WT and M5.

An attempt to determine high-resolution structural details of full-length ApoE4 nanodiscs
Because of the higher stability of the nanodiscs of ApoE4 (M5) over ApoE4 WT at high concentration, the ApoE4 (M5) nanodiscs were prepared and tried to get the structural insights with different lipids (Figure 6.10). However, only flexible segments were observed on the NMR spectrum. EM images reveal the elliptical morphology of nanodiscs with 10-20 nm of diameter. Next, ApoE4 WT nanodiscs were made with either DMPG or
PC14:1 lipid and 2-dimensional (2D) spectrum was acquired at high concentration and temperature (Figure 6.11, Figure 6.12). For both WT and M5 variants, almost similar spectra were found containing only a few cross-peaks of residues, which might be present in flexible regions (not directly involved in nanodiscs formation).

Figure 6.9: 2D [$^{15}$N, $^1$H]-TROSY Spectrum of full-length ApoE4 monomer (M5). Shows the spectrum as expected for a non-aggregated monomer sample.
Figure 6.10: $[^{15}\text{N},^{1}\text{H}]$-TROSY spectra of full-length ApoE4 (M5) nanodiscs and its morphology. (A) Only flexible residues (around 30-50) are visible in the case of both lipids containing nanodiscs. (B) The morphological characterization of nanodiscs was observed under the Electron microscope (EM).
Figure 6.11: [$^{15}$N,$^1$H]-TROSY of full-length ApoE4 WT nanodiscs versus ApoE4 M5 nanodiscs with DMPG (charged) lipids. More residues are in the flexible region (visible peaks) when it is mutated protein (M5), they might be those five hydrophobic residues which is no longer able to bind with lipid tightly because of losing/reducing the hydrophobicity. Significantly more tryptophan peaks are visible in WT nanodiscs spectra. Residues involved in nanodiscs formation might be not visible in the spectra possibly because of the large size of the nanodiscs or conformational dynamics.
Diffusion coefficient reveal bigger size was the limiting factor for structural determination

To understand the disappearance of the residues (ApoE4 nanodiscs) in 2D $^{15}$N,$^1$H-TROSY spectra of nanodiscs, Diffusion Ordered Spectroscopy (DOSY) experiments were performed for several samples including the ApoA1 (MSPDH5) nanodiscs which the structure was previously determined (235). The diffusion coefficient of ApoE4 nanodiscs (made of either DMPC or DMPG lipids) was found to be approximately two times less than MSPDH5 nanodiscs (Figure 6.13), which signifies it to be rather large for solution-state NMR.

![Figure 6.12: $^{15}$N,$^1$H-TROSY spectrum of full-length WT ApoE4 nanodiscs with different lipids (A) with DMPG lipids (B) PC14:1. It is mentioned that significant visible aggregates were formed upon NMR measurement at concentration of 500 M and temperature of 310 K.](image)
ApoE4 structure with lipids (in nanodiscs structure) versus ApoE4 in detergent (sodium cholate)

In a final experiment on the topic of ApoE4 nanodisc preparation for solution-state NMR, apoE4 WT were prepared in detergents and the 2D fingerprint \(^{15}\text{N},^{1}\text{H}\)-TROSY spectrum was measured (Figure 6.14). The comparison with the corresponding spectrum of ApoE4 WT in nanodiscs (Figure 6.11) indicates that detergent cannot mimic the membrane or lipids as expected.

*Figure 6.13: Diffusion measurement by 1D NMR for many samples. ApoE4 WT and M5 have significantly low value of diffusion coefficient than apoE4 free form and ApoA1(MSPDH5) nanodiscs.*
Figure 6.14: $^{15}$N,$^1$H]-TROSY spectrum of full-length ApoE4 was measured with sodium cholate. Detergent leads a completely different structure as compared to in nanodiscs. Although the sample was stable for a long time in high concentration at 310K.

Conclusion

In the present work, various proteins to lipids ratios were used to get a single homogenous population of lipoprotein particles by the various constructs of ApoE4. N-terminal (NT) domain (composed of residues 1-191) was able to form stable nanodiscs. The formation of NT nanodiscs was characterized by several techniques (SEC, 1D NMR Spectra, diffusion NMR measurements, 2D fingerprint NMR spectra at different
temperatures, and EM) and it was concluded that the size of NT nanodiscs is bigger than NT monomer. Crosslinking studies also suggested that two NT proteins are involved in the formation of one nanodisc. Interestingly, NMR spectral comparison indicated that there was no significant structural change in NT nanodiscs versus NT monomer.

In an attempt to get further structural insights into the ApoE nanodiscs limited proteolysis was done indicating that residues 16-206 are rather protected in the ApoE4(M5) construct indicating the important role of the NT in the formation of lipoprotein particles.

In both mutant M5 and wild-type full-length ApoE4 samples, only a few flexible (i.e. ~47) residues were visible on the 2D [$^{15}$N, $^1$H]-TROSY spectra. Diffusion experiments indicated the rather large size of the nanodiscs indicating that the low signal in the NMR spectra could be attributed at least in part to the size of the system. If the signals are attributed to a single segment with high flexibility in combination with the limited proteolysis data, one would assign them to the C-terminal residues 252-299.

Further many different constructs that contain the C-terminal domain but lack the N-terminal segments (i.e. constructs 62-299, 73-299, 165-299, etc) were also expressed. However, most of them did not express well, which may be another indicator of the important role of the NT.
In the future, the presented data are the first step towards a structure determination of ApoE lipoprotein particles. One aim might be to determine the lipoprotein structure of ApoE4 NT by solution-state NMR. Another approach is the use of solid-state NMR or cryo-EM for the structure determination of full-length ApoE4 apolipoproteins.

Material and Methods

Expression of the different construct of ApolipoproteinE4 (ApoE4) and ApolipoproteinA1 (ApoA1)

ApoE genes with various length and mutants were inserted in a pET28a vector containing N-terminal His6-tag and a TEV (Tobacco Etch Virus) protease recognition site (ENLYFQ), which was received from GenScript Biotech (Netherlands) B.V. ApoA1 plasmids were ordered from addgene (ID: 71714; 2006, Cambridge, US). Plasmids were transformed into E. coli (BL21DE3STAR) strain (Thermo Fisher Scientific, Waltham, US) then plated on kanamycin antibiotic containing LB-Agar plate. Few colonies were added to 100 ml of LB Broth media and grown overnight. Afterward 10 ml of preculture were added to 1 lit of TB-Media and grown at 37°C till the (OD600) reached 0.8, afterward protein expression was induced using 0.5mM IPTG. After 1h, the temperature was lowered to 28°C for another 4-5h and then the cells were harvested by centrifugation for 30 minutes at 4000g. Pellet was resuspended in 20mM Tris pH 8, 500mM NaCl and to prevent protein digestion one complete Protease Inhibitor tablet/ 100mL buffer was added to the solution and was mixed for 1h at 4°C, after mixing
cell were broken by passing three times through microfluidizer after that cellular parts were separated from soluble protein by high-speed centrifugation 36900 g (20000 rpm) for 20 minutes. Prepacked Histag TM excel (Ni-column) was used to isolate ApoE protein (contain His-tag). Elution was done using 300 mM imidazole. The high concentration of imidazole was removed by using a PD-10 desalting column. Further, if it was required to concentrate, Amicon Ultra Centrifugal Filter (10kDa, 15 ml) was used. Protein was stored at -20°C. For isotope-labeled sample M9 media was used (containing N\textsuperscript{15} ammonium chloride and C\textsuperscript{13} glucose) (in case of isotopically labeled) in place of LB media rest purification protocol remain the same.

**Lipoparticles (also called Nanodiscs) formation**

Lyophilized lipids (powder) were bought from Avanti Polar Lipids Inc. and were kept in -20°C. 100 mM lipids stock was made in 200 mM Sodium Cholate, the ratio of lipid to cholate was always kept 1:2, these stocks were made in a glass vial, Further solubilization of lipids was done by freezing (with liquid Nitrogen) and thawing (water bath of 60 °C). For the formation of nanodiscs, the various ratio of lipid to protein was tried for the single uniform peak of nanodiscs in size exclusion chromatography, it was well explained by group of Sligar and Wagner. Briefly different amount of protein and lipids were mixed (cholate was always 2X of lipids) and incubated at room temperature for overnight with gentle mixing, after overnight mixing 100% (w/v) Bio Beads SM Resin was added for 2h
and then it was centrifuged to separate Bio Beads and aggregates if formed. The soluble fraction was injected into Superdex 200 size exclusion chromatography (SEC) to the isolated pure fraction of nanodiscs. Further nanodiscs were concentrated by Eppendorf Amicon of 50kDa.

**SDS gel after cross-linking (CL)**

Disuccinimidylgluturate (DSG) contains two NHS ester (7.7 Å spacer length) forms irreversible amide bonds with lysine residues and primary amines. DSG was bought from ThermoFisher Scientific. DSG stock (25 mM) was made after dissolving in DMSO. Further 100 uM Apolipoprotein were incubated with 2 mM cross-linker in 20 mM sod. Phosphate buffer, pH 7.0 and kept at room temperature for 30 minutes further it was quenched with 1M Tris and kept for 15 minutes, later it was run on SDS page NuPageTM 4-12% Bis-Tris Gel (Invitrogen Thermo Fisher Scientific, Waltham, US) and stained with InstantBlueTM (Expedeon Ltd., Cambridgeshire, UK)

**Negative-Stain Electron Microscope (EM)**

10 µl of nanodiscs (10uM) were deposited on carbon film 150 square mesh copper grids CF150-Cu) for 1 min and then washed with Milli-Q water further grid was stained with 2% Uranyl acetate. EM micrographs were then collected with a FEI Morgagni 268 transmission electron microscope operated at 100 kV.
**NMR experiments**

All NMR (TROSY, HSQC, 1D and DOSY) experiments were performed on a Bruker 600 MHz Avance III HD or 700 MHz spectrometers equipped with a cryogenic probe. Sparky 3.115 (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco, USA) and Bruker Topspin 3.5pl7 (Bruker, Inc., Billerica, MA, USA) software were used for analyzing the spectra and extracting the rate constants.

Mass-spectrometry analysis of nanodiscs

**Mass-spectrometry**

ApoE4 Full length (M5) Nanodiscs were made with DMPC lipids further it was digested with two enzymes with various concentrations and duration and it was run on SDS-gel after running it was stained in a new container to prevent the presence of any other protein. Further the consistent band (protected from digestion) was given to mass-spectrometry bottom-up analysis at Functional Genomics center Zurich (FGCZ).
Chapter 7

Structural characterization of recombinant murine serum amyloid A (SAA) isoform 1.1

Abstract

Serum Amyloid A (SAA) belongs to the Apolipoprotein family. 95% of the SAA protein was found to be bound to high-density lipoprotein (HDL). Hence suggesting its role in lipid metabolism and clearance. Binding to HDL also prevents SAA to get fragmented and aggregated. SAA has been involved in several chronic inflammatory diseases (amyloidosis, atherosclerosis and rheumatoid arthritis). It is known that it forms α-helical structure upon HDL binding, but residues level structural information is still not known for the HDL-bound SAA. In this study, a stable functional HDL bound SAA was prepared for the structural understanding. SAA was observed to be tightly bound to HDL. The transition from random coil to the folded structure was observed for SAA upon HDL binding. Spectra was found to more resolved at high temperatures (315K), pH 6.8 and with charged lipids (DMPG). Rotational correlation time ($\tau_c$) observed to be decreasing above melting temperature (Tm) of the lipid bilayer which suggests SAA is bound to the...
lipid bilayer. Decreasing temperature below the melting temperature (Tm) increases the $\tau_c$ which suggests it may not be going inside the lipid bilayer like $\beta$ barrel proteins but most likely binds on the surface.

Introduction

SAA is a small (104 amino acids) protein belongs to the apolipoprotein’s family. SAA name was given because of its presence in the serum and being precursor of amyloid disease. SAA sequence is highly conserved in mammals and birds. (240). The normal concentration of SAA is 20-50 $\mu$g/ml in blood but upon acute phase response (APR) concentration increases by 1000 times. N-terminal fragments of SAA have been observed to form aggregates. The essential physiological role is not known clearly although it has been involved in lipid metabolism, cell-cell communication, role in inflammation (124). Understanding the structure by NMR or Crystallization was challenging because of the instability of SAA. Based on the sequence-based algorithm SAA was proposed to contain two $\alpha$-helices and a $\beta$-sheet however when its monomer structure was solved by multiwavelength anomalous dispersion four $\alpha$-helices in a cone-shaped array and no $\beta$-sheet structure was observed (125). Hexamer formed by two identical trimers were found when SAA was purified without denaturation (211). The amphipathic region of SAA and 95% of the circulating SAA is in the HDL fraction. Therefore, this suggests its function to be in lipid metabolism and cholesterol transport.
In this study we found the good condition to have resolved the structural change of SAA upon HDL binding.

**Results and discussion**

**Large-scale expression of SAA**

Fermenter (LABFORS 3) was used for 1.8-liter media. A high OD$_{600}$ of approx. 25 was possible to reach approx. 80g of pellet weight at the end (Figure 7.1). A similar way of expression was performed for unlabeled, deuterated, specifically unlabeled and specifically labeled NMR samples.

*Figure 7.1: Expression of SAA (N15, C13, 90% D2O) with specifically unlabeled amino acids (Alanine, isoleucine, leucine, and valine) (A) Measured OD$_{600}$ with time (1.8 L M9 media in the fermenter) (B) SDS-PAGE: before and after induction of protein expression with various dilution*
NMR spectra of SAA monomer
Measuring the spectrum of the monomer revealed itself to be challenging. This is due to its limited solubility. To dissolve the lyophilized protein, it was important to increase the pH to 11-12. Once the protein was dissolved; the pH was adjusted to 7.6. It was not possible to measure the spectra from pH 3.5 to 7.5 range because of the formation of visible precipitates. NMR signal did not get better by decreasing the temperature as expected for unfolded protein, reason might be the formation of partially helical structure (125). The maximum solubility for the monomer was found to be 40 µM. Protein also tends to degrade and aggregate with time, therefore spectra were measured just after sample preparation.

Binding of monomer SAA to HDL
SAA-HDL bound complex was formed by two different methods. Method 1: Nanodiscs were prepared with MSPDH5 (membrane scaffold protein with deleted helix 5) and lipids, it was purified by size exclusion chromatography (SEC). Furthermore, diluted SAA was added to already formed nanodiscs. Further, it was concentrated via using 10 kDa filter, injected again to confirm binding (Figure 7.3). Method 2: SAA, MSPDH5, and lipids were incubated together and then peak was isolated by SEC containing SAA bound nanodiscs. SAA-HDL complexes prepared in different manners looked similar on NMR spectra and SDS page. Hence, the 1st method was used to prepare nanodiscs bound SAA.
Figure 7.2: 2D HSQC Spectra of monomer SAA at different pH. Protein concentration of 40 µM at 298K in 50 mM sodium phosphate buffer. Spectra suggest SAA is unfolded at this condition.
Structural characterization of recombinant murine serum amyloid A (SAA) isoform 1.1

Figure 7.3: MSPDH5 (DMPG) nanodiscs bound SAA (prepared by the 1st method). (A) SEC elution profile (Superdex 200, 10/300 GL) of a bound sample (B) SDS-PAGE shows both MSPDH5 and SAA are prominent in the same fraction leads to indicate nanodiscs bound SAA.

Conditions tried for resolved NMR spectrum of HDL bound SAA
Resolved spectra with structural changes and the stable sample were found with both MSPDH5 and MSPDH4-H6 at high temperature (315K) and at relatively low pH with negatively charged lipids. Adding Protease inhibitor and EDTA was also found to be essential to control protein fragmentation (Table 7.1).

Spectral quality with temperature and time
$^{15}$N -HSQC spectra were acquired at several temperatures and processed in the same way, which reflects the significant effect of temperature on spectra (Figure 7.4). DMPG lipids have a melting temperature ($T_m$) of 23°C (296K) and the reason for the poor spectral resolution might be the lack of interaction of SAA with the solid-like nanodiscs.
Table 7.1: $^{15}$N-HSQC Spectra were acquired for several conditions for the well-resolved spectra. Two different lengths of ApoA1 (MSPDH5 and MSPDH4-H6) proteins were used to make nanodiscs with various lipids containing different physical and chemical properties.

Another reason might be that SAA was trapped inside the lipid and tumbling was affected. $\tau_c$ decreased at high temperature (315K) which indicates faster tumbling of nanodiscs bound SAA (Figure 7.5). Spectra quality decreased significantly over time because of the aggregation of SAA (Figure 7.6).
Figure 7.4. 2D $^{15}$N-HSQC spectra, intensity, and linewidth at different temperature A). HSQC spectra at various temperature B). Line-width at half height (LWHH) decreases with temperature increase.
Figure 7.5. Rotational correlational time ($\tau_c$) of SAA when bound to MSPDH5 (DMPG) nanodiscs. $\tau_c$ decreases by increasing temperature from 310 to 293 which indicates bound SAA tumble faster at high temperature. However, below Tm, $\tau_c$ decreases which indicates SAA might get unbound. The integration of peak was done from 6.5-11 ppm.

Figure 7.6. Sample stability over time. $^{15}$N-HSQC spectra after 10 days show SAA aggregates with time.
SAA is bound to HDL
SAA bound to HDL has a similar diffusion coefficient value $5 \times 10^{-5}$ cm$^2$s$^{-1}$ as HDL alone $4.7 \times 10^{-5}$ cm$^2$s$^{-1}$, which confirms SAA is bound to nanodiscs.

![Figure 7.7. 2D Diffusion experiments at 310K with MSPDH5 (DMPG) nanodiscs with 100mM NaCl, pH 7.4. The blue line indicates $^1$H coupled to $^{15}$N (Bound SAA), the green line shows $^1$H not coupled to $^{15}$N (only HDL).]

SAA structural change from unfolded monomer to HDL bound and with detergent
Structural change was observed for the SAA when it was bound to Fos-Choline (FC-12) detergent. SAA peaks were well resolved when it was bound to detergent. However, the structural change of SAA was completely different in detergent from the HDL bound structure. This suggests detergent cannot be used to study the structural change upon HDL binding.
Figure 7.8: $^{15}\text{N}$-HSQC spectra of SAA with HDL versus detergent. A) SAA monomer spectrum versus SAA in detergent B) SAA with HDL particle.
Conclusion

SAA monomer was observed to be unstable, therefore it was not possible to reach a high concentration (aggregation-prone and fragmentation). However, when SAA was bound to HDL it was possible to attain a higher SAA concentration (approx. 2 mM). Furthermore, the interaction of SAA with HDL was confirmed by the bigger retention time in SEC and by the decrease in the diffusion coefficient. In order to have resolved 15N-HSQC spectra of SAA bound to nanodiscs, several lipids, nanodiscs size, and solvent conditions have been tried. A well-resolved spectrum was observed at relatively high temperature (315K), pH 6.8 with negatively charged lipids (Table 7.1). The structure of SAA changed significantly upon HDL binding. However, the structural change was not similar in the presence of detergent. Further detail about the structure could not be achieved due to the bigger size of the HDL bound form; only 20% of the residues were possible to assign. In future, the formation of HDL with deuterated lipids and protein and choosing the smaller nanodisc construct (MSPDH4H5H6) can help in solving the 3D structure. Since several SAA units can hypothetically bind to one nanodisc, a condition where a 1:1 ratio is achieved could reduce the system size. Moreover, the SAA is thought to adopt different conformations (data not shown) in contact with the lipid membrane. Hence, a 1:1 ratio might favor the most thermodynamically stable conformation and thus simplify the resulting spectra.
Material and Methods

Nanodiscs formation, size exclusion chromatography, NMR measurements were performed as explained in Chapter 5 and 6.
Chapter 8

Conclusion and Outlook

Investigation of Synuclein-lipid interaction
Interaction of α-Syn with lipid bilayer is essential for the understanding of both physiological and pathological functions of α-Syn. To study the interaction of α-Syn with lipid bilayer at physiological condition a relaxation pulse sequence was required which exclude the contribution from the solvent exchange and give the correct timescale for protein interaction. In Chapter 2 15N the R2 relaxation pulse sequence was modified as well as the amount of D2O was reduced to exclude the misleading increase in relaxation value because of amide-water exchange. Further studies were performed using this modified pulse sequence of NMR relaxation measurement. In the future, the modified pulse sequence, as well as reduction of D2O amount, can lead to an accurate measurement of protein interaction in-cells at the physiological conditions.

α-Syn has been shown to make lipoprotein (nano) particles where 8-10 synuclein monomer interacts with lipids. In Chapter 3 the goal was to study the correlation of disease and function to the amount and morphology of nanoparticle formation. The nanoparticle formation
triggered by different synuclein variants (α-Syn, β Syn, and γ Syn) as well as disease-associated mutants (A30P, E46K, A53T, H50Q, G51D) of α-Syn were characterized. Distinct differences in the morphologies were observed (by EM) ranging from roundish (α-Syn) to more elongated elliptical (β-Syn) shapes of nanoparticles. There was a drastic effect on the amount of nanoparticle formation upon single point mutations. The lowest amount of nanoparticle formation was by A30P (~50%) to the highest amount of nanoparticles formation by G51D (~100%). In the future, finding Syn nanoparticles in vivo will be very interesting to understand the function in lipid storage and transport.

Additionally, in Chapter 5 the interaction of α-Syn with the different lipids (have different Phase, Tm, charge) was studied using ApoA1 nanodiscs as a membrane model. α-Syn-lipid interaction (especially exosomes) involved in the propagation of PD pathology and in inducing the α-Syn aggregation was understood. α-Syn-lipid interaction was more pronounced with the negatively charged lipids with a high degree of unsaturation in the liquid phase (above Tm). Hydrophobic residues of the N-terminal were found to be strongly involved in membrane interaction in the case of exosomes lipid composition. This reveals the possible mechanism of membrane disruption and the increase in the local monomer concentration within the membrane. In the future, membrane-bound α-Syn structure can be characterized. Finding the ways against the
formation of membrane-bound α-Syn conformation can be protective for membrane disruption and aggregation acceleration.

Investigation of Synuclein-fibril interaction
Most of the neurodegenerative diseases are late-onset, appears in the late 5th decade or 6th decade of life. However once detected around 70% of the neurons are damaged and thereafter, the progression is relatively fast. Studies in Chapter 4 reveals the understanding of the interaction between α-Syn monomer and fibrils (secondary nucleation). The important results conclude the positively charge N-terminal domain of the α-Syn monomer interacts with negatively charged C-terminal domain of the fibrils in the millisecond timescale. At pH 7.0 the transient interaction between fibrils and monomer is mostly electrostatic which decreases upon the addition of salt however at pH 6.0 interaction between monomer and fibrils is much tighter and not changed upon salt addition. In the future, drugs could be designed to inhibit the interaction between the monomer-fibril and therefore lead to inhibit the fast formation and accumulation of fibrils in PD patients.

Investigation of the proteins from the apolipoprotein family
Apolipoprotein E (ApoE) belongs to the apolipoprotein family. It has been reported that people containing ApoE4 isoform (rather than ApoE3, ApoE2) have a higher risk for Alzheimer’s Disease (AD). In Chapter 6, structural insights about the functional form of ApoE4 were characterized. ApoE contains two folded domains (N-terminal and C-
terminal) connected to the hinge region. Both N-terminal and C-terminal domains are rich in helical structure. Different domains and mutated constructs of ApoE4 were purified and homogenous lipoprotein-like particle was formed. N-terminal domain was found to form lipoprotein like structure without any structural change. For full-length ApoE lipoprotein, only 30-40 residues were found on the NMR spectra. However, all the other residues involved in the lipoprotein formation were too big to be observed by solution-state NMR. In the future, the combination of structural biology technics such as Cryo-EM, EPR, and solid-state NMR studies could help to get the structural detail.

Another important protein Serum Amyloid A was studied, which also belongs to the apolipoprotein family. Serum Amyloid A (SAA) binds to high-density lipoprotein (HDL) to reroutes HDL and maintain lipid metabolism and clearance. SAA is found to be involved in several chronic inflammatory diseases (amyloidosis, atherosclerosis and rheumatoid arthritis). It is known that it forms -helical structure after binding to HDL, but residues level structural information is still not known for the HDL-bound SAA. In Chapter 7 the Various size of HDL, different lipids and solvent conditions were screened to observe a well-resolved spectrum. Relatively resolved spectrum for HDL-bound SAA was found at 315K, pH 6.8 with negatively charged lipids. The structure of HDL bound SAA was found to be drastically different from the unbound form and with detergent. Although the assignment of the HDL bound SAA was not
Conclusion and Outlook

possible because of the size limitation. In the future, using deuterated and smaller HDL might help in achieving the 3D structural detail of SAA upon HDL binding.

Conclusion
Interestingly, both positively charged and hydrophobic residues from the N-terminal of α-Syn were found to be important in the interaction with lipids and pre-formed aggregates. It was also interesting to find the formation of very stable lipoprotein like particle which can be the potential functional form of α-Syn in vivo. There were also progress in the purification and preparation of a homogeneous samples of ApoE HDL like particle and HDL bound SAA. Solvent conditions and lipids were screened for well-resolved spectra. However, 3D structural detail is still sought.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-Syn</td>
<td>Alpha-synuclein/α-Synuclein</td>
</tr>
<tr>
<td>AD</td>
<td>Alzheimer’s Disease</td>
</tr>
<tr>
<td>ALS</td>
<td>Amyotrophic Lateral Sclerosis</td>
</tr>
<tr>
<td>ApoA</td>
<td>Apolipoprotein A</td>
</tr>
<tr>
<td>ApoE</td>
<td>Apolipoprotein E</td>
</tr>
<tr>
<td>AWI</td>
<td>Air Water Interface</td>
</tr>
<tr>
<td>CL</td>
<td>Crosslinker</td>
</tr>
<tr>
<td>CT</td>
<td>C-terminal</td>
</tr>
<tr>
<td>DLB</td>
<td>Dementia Lewy Body</td>
</tr>
<tr>
<td>DMPC</td>
<td>1,2-dimyristoyl-sn-glycero-3-phosphocholine</td>
</tr>
<tr>
<td>DMFG</td>
<td>1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (sodium salt)</td>
</tr>
<tr>
<td>DOPC</td>
<td>1,2-dioleoyl-sn-glycero-3-phosphocholine</td>
</tr>
<tr>
<td>DOPG</td>
<td>1,2-dioleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (sodium salt)</td>
</tr>
<tr>
<td>DOPS</td>
<td>1,2-dioleoyl-sn-glycero-3-phospho-L-serine (sodium salt)</td>
</tr>
<tr>
<td>DOSY</td>
<td>Diffusion Ordered Spectroscopy</td>
</tr>
<tr>
<td>DSG</td>
<td>Disuccinimidygl glutarate</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EM</td>
<td>Electron Microscopy</td>
</tr>
<tr>
<td>GM3</td>
<td>Ganglioside</td>
</tr>
<tr>
<td>HDL</td>
<td>High-Density Lipoprotein</td>
</tr>
<tr>
<td>HMQC</td>
<td>Heteronuclear Multiple-Quantum Correlation</td>
</tr>
<tr>
<td>HSQC</td>
<td>Heteronuclear Single Quantum Coherence</td>
</tr>
<tr>
<td>IDL</td>
<td>Intermediate Density Lipoprotein</td>
</tr>
<tr>
<td>IDP</td>
<td>Intrinsically Disorder Protein</td>
</tr>
<tr>
<td>LB</td>
<td>Lewy Body</td>
</tr>
<tr>
<td>LDL</td>
<td>Low Density Lipoprotein</td>
</tr>
<tr>
<td>MS</td>
<td>Mass Spectrometry</td>
</tr>
<tr>
<td>MSA</td>
<td>Multiple System Atrophy</td>
</tr>
<tr>
<td>MSP</td>
<td>Membrane Scaffold Protein</td>
</tr>
<tr>
<td>MSPD1</td>
<td>MSP with half of the first helix</td>
</tr>
<tr>
<td>MSPDH4-H6</td>
<td>MSPD1 with deleted helix 4, 5, 6</td>
</tr>
<tr>
<td>MSPDH5</td>
<td>MSPD1 with deleted helix 5</td>
</tr>
<tr>
<td>MSPDH5 R2</td>
<td>MSPDH3 with 2 X (H4 and H6)</td>
</tr>
<tr>
<td>NAC</td>
<td>Non-Alpha component</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium Chloride</td>
</tr>
<tr>
<td>ND</td>
<td>Nanodiscs</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>NT</td>
<td>N-terminal</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PD</td>
<td>Parkinson’s Disease</td>
</tr>
<tr>
<td>PTM</td>
<td>Post translational modification</td>
</tr>
<tr>
<td>SAA</td>
<td>Serum amyloid A</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SEC</td>
<td>Size exclusion chromatography</td>
</tr>
<tr>
<td>SN</td>
<td>Substantia Nigra</td>
</tr>
<tr>
<td>SNCA</td>
<td>Alpha-synuclein gene</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>SNCB</td>
<td>Beta-synuclein gene</td>
</tr>
<tr>
<td>SNCG</td>
<td>Gamma-synuclein gene</td>
</tr>
<tr>
<td>ssNMR</td>
<td>Solid State NMR</td>
</tr>
<tr>
<td>SUV</td>
<td>Small Unilamellar vesicles</td>
</tr>
<tr>
<td>TFE</td>
<td>Trifluoroethanol</td>
</tr>
<tr>
<td>ThT</td>
<td>Thioflavin T</td>
</tr>
<tr>
<td>Tm</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>TROSY</td>
<td>Transverse relaxation optimized spectroscopy</td>
</tr>
<tr>
<td>VLDL</td>
<td>Very Low-Density Lipoprotein</td>
</tr>
<tr>
<td>WT</td>
<td>Wild type</td>
</tr>
</tbody>
</table>
Bibliography


synuclein mutation causes a novel parkinsonian-pyramidal syndrome, Annals of neurology 73, 459-471.


Alzheimer's disease amyloid is a presynaptic protein of the central nervous system, Neuron 14, 467-475.


interactions potentiates aggregation of natively unstructured alpha-synuclein, Proceedings of the National Academy of Sciences of the United States of America 102, 1430-1435.


52. Cabin, D. E., Shimazu, K., Murphy, D., Cole, N. B., Gottschalk, W., McIlwain, K. L., Orrison, B., Chen, A., Ellis, C. E., Paylor, R., Lu, B., and Nussbaum,


hydrogen/deuterium exchange monitored by mass spectrometry, Biochemistry 52, 9097-9103.


synuclein variants, Biochemical and biophysical research communications 386, 165-169.


Time Scale Motions in Proteins Revealed by High-Resolution NMR Relaxometry, Journal of the American Chemical Society 135, 18665-18672.


intrinsically disordered protein by phosphorylation as a regulatory switch, Nature 519, 106-U240.


