Mechanisms and prevention of protein aggregation

Michael Barber

Wadham College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2016

Nature is the proof of dialectics, and it must be said for modern science that it has furnished this proof with very rich materials increasing daily. Freidrich Engles, Anti-Duhring, 1877.

When I speak of reason or rationalism, all I mean is the conviction that we can learn through criticism of our mistakes and errors, especially through criticism by others, and eventually also through self-criticism. A rationalist is simply someone for whom it is more important to learn than to be proved right; someone who is willing to learn from others not by simply taking over another's opinions, but by gladly allowing others to criticize his ideas and by gladly criticizing the ideas of others. The emphasis here is on the idea of criticism or, to be more precise, critical discussion Karl Popper, All Life is Problem Solving, 1994

The challenge of modernity is to live without illusions and without becoming disillusioned. Antonio Gramsci, Analysis of social relations, 1937 [Antonio Gramsci, 1937].

Dedicated to my family, friends and teachers; curricular and extra-curricular.

Acknowledgements

It is my pleasure to thank several people and organizations, without which my D.Phil. would not have been possible.

I must start by thanking Gail Preston, James Wright and the Oxford Doctoral Training Partnership, alongside the Biotechnology and Biological Sciences Research Council (BBSRC), for giving me the opportunity to study in a world-renowned institution with high quality facilities and academics. Gail Preston especially has been supportive over the 4 years of my D.Phil and takes a great interest in her students' welfare.

I would also thank my supervisors, Robert Gilbert, Andrew Baldwin, and my rotation supervisor Justin Benesch. Andrew has provided good critical feedback throughout my DPhil, and taught me NMR and the uses of data. Robert has been very supportive throughout my D.Phil. and a great source of knowledge on all things crystallography related, as well as providing critical feedback. Justin introduced me to mass spectrometry and has very kindly provided feedback on some ideas/texts in this thesis. Most of all I am grateful for the intellectual freedom to explore hypotheses and ideas in my own way. Without Robert and Andrew I would

not be the independent scientist I am today.

No acknowledgment would be complete without thanking my fellow travelers in the Oxford academic system. Henrik Muller, Iva Pritisanac and Olga Tkachenko, my laboratory colleagues, who are responsible for a lot of great scientific debate and friendship. My fellow DTP/DTC students, Nathalie Willems, Samuel Demharter and Jonathon Brooks-Bartlett are responsible for many late night invigorating discussions about science and politics. These people have maintained my sanity throughout this odyssey.

My deepest acknowledgments are left for my family; a rabble of British Roma who remain the most introspective, thoughtful and intelligent people I've known. My father first taught me critical thinking and problem solving, my mother inspired ambition and adventure, and my sister, Teresa, patiently explained many complex ideas to a precocious young boy. She remains an inspiration, and I expect she will write one of these too some day.

Abstract

The deposition of amyloid in the central nervous system is associated with prevalent neurological disorders such as Alzheimer's and Parkinson's disease. This thesis studies the mechanisms and prevention of amyloid formation *in vitro*. We specifically focus on Parkinson's disease associated α-synuclein (α-syn). Using novel labeling methods we introduce NMR observable labels onto lysosomal protein glucocerebrosidase (GCase), a leading cause of Parkinson's disease. By introducing NMR active labels we are able to study GCase dynamics and screen potential drug therapeutics (chapter 3). Furthermore, we analyze the three way interaction between GCase, α-syn and lipids. We conclude that GCase is able to effectively chaperone α-syn under lysosomal conditions, both preventing amyloidogenesis and destabilizing mature amyloid fibrils (chapter 4). Additionally, a model chaperone-aggregate system is investigated to gain insight into the mechanisms of small heat shock protein chaperoning, and how such mechanisms prevent aggregation (chapter 5). Finally, a high resolution crystal structure of RNA editing enzyme Cid1 is presented, whilst not directly linked to aggregation, many of the techniques used in this thesis were first developed on Cid1 (chapter 7). Together, we utilize NMR, X-ray crystallography, electron microscopy and native mass spectrometry to elucidate aspects of protein aggregation mechanisms and prevention.

Contents

List of Figures

Key abbreviations

αlac - alpha lactalbumin

αSyn - alpha-synuclein

Aβ - Amyloid beta

AD - Alzheimer's disease

Cid - Caffeine induced death suppressor

CPMG - Carr-Purcell-Meiboom-Gill

EM - Electron microscopy

ER - Endoplasmic reticulum

ERAD - Endoplasmic reticulum associated degradation

GCase - Glucocerebrosidase

HMQC - Heteronuclear multiple quantum coherence

HSP - Heat shock protein

HSQC - Heteronuclear single quantum coherence

IMS - Ion mobility spectrometry

LCMSMS - liquid chromotography tandem mass spectrometry

MS - Mass spectrometry

nESI - Nano-electrospray ionization

NNDNJ - N-Nonyldeoxynojirimycin

NMR - Nuclear magnetic resonance

PC- Pharmacological chaperone

PD - Parkinson's disease

TROSY - Transverse relaxation optimized spectroscopy

TUT - Terminal uridyl transferases

XRC - X-ray crystallography

Chapter 1

Introduction: protein aggregation and misfolding

1.1 The amyloid conformation and toxicity

1.1.1 An overview

Most proteins fold into ordered three dimensional structures to carry out their biological functions. These folding pathways exist on an energy landscape consisting of multiple local minima (figure 1.1). One such minimum may represent the functional form of a protein. However, there exist other minima where proteins fold into alternative conformers. Most (if not all) proteins can form a conformer which, under certain conditions, represents the global energy minimum; amyloid. As a global minimum, the amyloid conformation represents the most thermodynamically stable conformer available to a dynamic protein, as such, the issue of amyloid formation is merely one of kinetics, and arguably, amyloid formation is inevitable for all proteins [Baldwin et al., 2011b].

Crucially, amyloid is recognized by both a linear, unbranched appearance

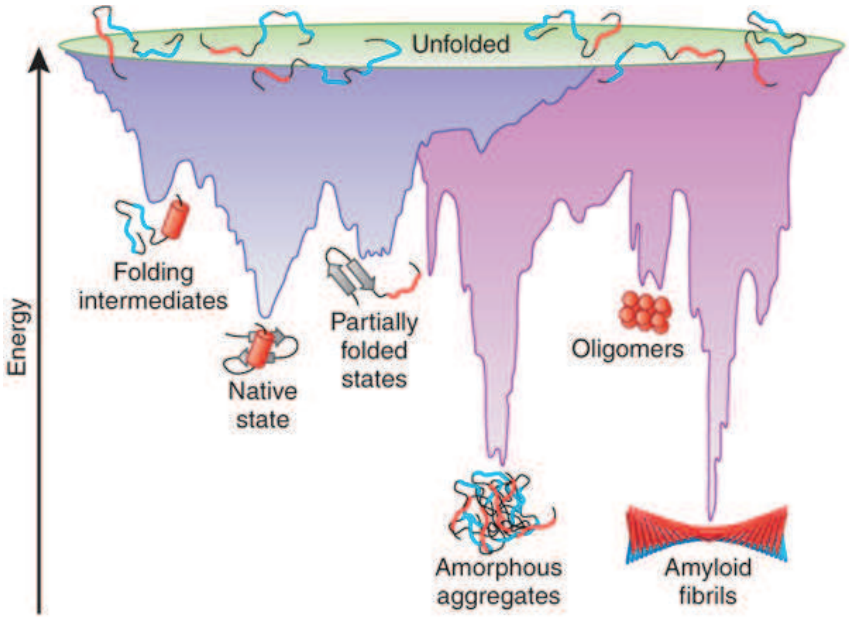


Figure 1.1: *A minimalist energy landscape schematic, adapted from [Hartl and Hayer-Hartl, 2009]. Blue shading represents on-pathway native intermediates and red areas represent on-pathway to aggregate folds.*

by electron microscopy, and the binding of histological dyes such as Thioflavin T and Congo Red. Detailed studies of amyloid fibrils by nuclear magnetic resonance (NMR) and X-ray diffraction have revealed a common compact cross β-sheet structure [Sawaya et al., 2007, Fitzpatrick et al., 2013, Wasmer et al., 2008, Fitzpatrick et al., 2013, Khurana et al., 2005, Trojanowski and Lee, 2002, Vilar et al., 2008, Cohen et al., 2015, Heise et al., 2005, Ecroyd et al., 2008, Heise et al., 2005, Heise et al., 2008, Lemkau et al., 2013, Zhou et al., 2012].

The deposition of amyloid in the central nervous system is associated with prevalent neurological disorders such as Alzheimer's (AD) and Parkinson's disease (PD) [Irvine et al.,]. Alzheimer's disease manifests as the cognitive decline of executive function and episodic memory [Tarawneh and Holtzman, 2012], and is projected to affect 80 million people globally by 2030 [Alzheimer

Disease International,]. Parkinson's disease is characterized by a different set of symptoms, most commonly; a resting tremor, slowed movement and cognitive decline. Estimates of Parkinson's disease are more complex due to mis-diagnosis, however, it is approximated that 5 million people globally are currently living with PD [World Health Organization, b, Buter et al., 2008]. Together, Alzheimer's and Parkinson's disease account for 1% of total global premature deaths [World Health Organization, a].

The association between insoluble amyloid deposits and disease is derived from *in vivo* and post mortem studies [*Knowleset al.*, 2014]. Causality between such diseases and amyloid was first suggested by the presence of familial mutations [Trojanowski and Lee, 2002, Levy et al., 1990]. PD is characterized by the aggregation of the short, intrinsically disordered protein α-synuclein (α-syn) into extracellular bodies known as Lewy Bodies [Trojanowski and Lee, 2002]. Naturally occurring mutations in α-syn significantly affect the age of onset for PD and may aggregate more rapidly *in vitro* [Recchia, 2004, Conway et al., 1998]; thus suggesting causality between amyloidogenesis and disease. Likewise, mutations in the enzymes responsible for processing amyloid-β (Aβ), the main constituent of Alzheimer's plaques, are associated with early onset AD [Levy et al., 1990].

Significant advances in understanding amyloid diseases have been made *in vitro*, however the molecular mechanisms associated with pathology *in vivo* are less well explored [Eisenberg and Jucker, 2012, Akter et al., 2016, Clark and Nilsson, 2004]. Recently, structural studies have attempted to identify conserved motifs in amyloid proteins to explain disease progression and uncover the relationship between structure and toxicity [Lue et al., 1999].

1.1.2 The structure and formation of amyloid

Amyloid displays a number of characteristic features elucidated over the years by a combination of NMR, X-ray diffraction, electron microscopy (EM) and other biophysical techniques. X-ray diffraction of *ex vivo* and synthetic (*in vitro*) amyloid fibrils revealed a general structure consisting of twisted cross β-sheets running perpendicular to the fibril axis [Sunde et al., 1997, Tycko, 2011]. The characteristic amyloid fibril diffraction pattern includes intense reflections at 4.7 and 10 Å. The former derive from generic (i.e. side chain independent) polypeptide backbone interactions perpendicular to the fibril axis and stabilized by hydrogen bonds, whilst the latter represent side chain dependent spacing of β-sheets perpendicular to the fibril axis. Such generic interactions along polypeptide backbones are responsible for the hypothesis that most, if not all, proteins are able to form amyloid [Baldwin et al., 2011b] (reviewed in [Eichner and Radford, 2011]).

Further study by Makin *et al.* utilizing electron and x-ray diffraction revealed the importance of aromatic-aromatic and charge-charge interactions in stabilizing the amyloid conformation [Makin et al., 2005]. Phenylalanine dependent π-π interactions were essential in the formation of sequence-designed (i.e. non-natural) amyloid fibrils and were found to be over-represented in amyloid disease associated proteins [Gazit, 2002, Marshall et al., 2011]. Taken together, this suggests that generic, backbone interactions are responsible for the universal amyloid structures observed, whilst charged or aromatic side chains effect aggregation kinetics [Marshall et al., 2011, Baldwin et al., 2011b].

The regular spacing achieved from backbone-backbone interactions may also explain the ability of amyloid fibrils to bind the fluorescent dye Thioflavin

T (ThT). The cross-β amyloid structure gives rise to repetitive side chain interactions on the exposed amyloid surfaces, these side chain interactions form 'cross-strand ladders' running parallel to the long fibril axis. When in solution, ThT aromatic groups are able to rotate freely and therefore quench any photon generation by excited states. However, when bound in the defined cross-strand ladder motif, it is likely this rotation is restricted, enhancing photon generation and generating fluorescence (reviewed [Biancalana and Koide, 2010]).

The most detailed structural studies of amyloid arise from synthetic, *in vitro*, amyloid preparations. Solid state magic angle spinning NMR is uniquely able to report on conformers of monomers within amyloid fibrils. The use of solid state NMR alongside complementary biophysical techniques such as electron microscopy and X-ray diffraction yields information on both local (tertiary) and global (quaternary) structure. Such approaches were successfully used to reveal that identical peptides could form alternative amyloid morphologies. Examination of Aβ(1-40) peptides, associated with Alzheimer's disease, revealed two distinct molecular morphologies; *striated ribbons* and *twisted pairs*. These conformers differed by 40% in density, and revealed two and three fold symmetry, for striated ribbon and twisted pair conformers, respectively [Petkova et al., 2006, Paravastu et al., 2008, Petkova, 2005]. Similar dual conformers have also been observed in insulin and amylin fibrils [Jiménez et al., 2002, Goldsbury et al., 2000]. The discovery of alternative amyloid conformers has been invoked to explain the poor correlations between amyloid load and disease progression for Alzheimer's, with some suggesting it is the *type* of fibril rather than the overall load that determines pathogenicity [Petkova, 2005].

Efforts to examine the heterogeneity of disease causing fibrils *in vivo* have

been most successful with the Aβ peptide. Notably, the Tycko group have employed solid state NMR and sparse ^{13}C labeling to derive structural restraints and long-distance information for Aβ fibrils seeded from *ex vivo* Alzheimer's disease samples. Such studies reveal spectra differing from both striated ribbon and twisted pair *in vitro* conformers [Paravastu et al., 2009]. Follow up studies comparing two patients found that amyloid structures were conserved across parietal, occipital and frontal lobe amyloid samples for each patient, but differed between patients. These structural models were also distinct to the synthetic fibrils that produced striated ribbon and twisted pair conformers. Whilst *ex vivo* seeded fibrils displayed a 3 fold symmetry and mass per length similar to that of synthetic twisted pair conformers, the structural models differed considerably [Lu et al., 2013, Paravastu et al., 2009].

A similar analysis for Parkinson's disease α-syn is yet to be carried out. However, Bousset *et al* generated two distinct strains of α-syn *in vitro* [Bousset et al., 2013]. Characterization of these morphologies through quantitative electron microscopy revealed either a straight, flat structure (termed 'fibrils') or a twisted structure with alternating wide and narrow sections (termed 'ribbons'). Most interestingly, the cytotoxicities, as measured by cell survival assays and apoptotic factors, differed between fibril conformers, with the straight, flat forms being more toxic than the twisted, ribbon-like forms [Bousset et al., 2013]. Likewise, a contemporary report from Guo *et al* revealed that α-syn strains could differentially promote the aggregation of Tau protein, a common aggregate observed in both AD and PD [Guo et al., 2013].

Together, these reports suggest that amyloid-disease associated sequence-identical proteins can form multiple conformers with differing toxicities. Unfortunately, structural information on these 'strains' is limited to low-resolution

information derived from electron microscopy [Bousset et al., 2013] or sparsely labeled solid state NMR [Paravastu et al., 2008] studies. The lack of detailed information on strain structure limits both an understanding of how strain structure effects toxicity and makes the accurate replication of strains by other scientific laboratories difficult. The recent report of a high resolution α-syn amyloid conformer (discussed in 1.2) may aid the structural characterization of α-syn strains [Tuttle et al., 2016].

Whilst the structure of amyloid fibrils themselves has received much attention, their association with disease has long been in question. This is mostly due to the poor correlation between disease progression and amyloid fibril load. More recently, it has been suggested that amyloid is a protective inclusion body, sequestering the real toxic species, small-scale oligomers [Lorenzen et al., 2014].

1.1.3 Targeting a toxic species

Aβ fibrils grown under various conditions reveal disparate morphologies and cytotoxicities [Petkova, 2005]. However, it is unclear if it is the fibrils themselves, or a dissociated oligomer, that is responsible for cytotoxicity. Evidence against the "toxic fibrils" has accumulated from animal and human studies. Most strikingly, a follow up of AD patients treated with antibodies specific for Alzheimer's associated amyloid-plaques revealed decreased plaque load, but continuing cognitive decline [Holmes et al., 2008]. Furthermore, total plaque load does not consistently correlate with neurone death or cognitive decline, and amyloid plaques are found in the brains of healthy subjects [Arriagada et al., 1992, Billings et al., 2005, Gómez-Isla et al., 1997, Terry, 1996, Delaère

et al., 1990, Westerman et al., 2002]. This has led many to suggest oligomeric intermediates as a toxic species in neurodegenerative diseases.

The effects exerted by toxic intermediates have thus been a major target of investigation recently in animal and cellular studies. Briefly, Aβ oligomers quantified by enzyme linked immunosorbent assays (ELISA) better predict AD pathology than amyloid deposits [Wang et al., 1999, McLean et al., 1999, Näslund et al., , Lue et al., 1999]. Injection of cell derived oligomers into rat brains or hippocampal slices has been shown to inhibit long term potentiation, an electrophysiological correlate of memory and, presumably, cognitive decline [Kamenetz et al., 2003, Townsend et al., 2006, Walsh et al., 2002]. Interestingly, the generation of soluble oligomers from proteins associated with other amyloid diseases, such as Parkinson's disease and diabetes, reveal shared structural features and pathological mechanisms [Kayed et al., 2003]. This shared structural motif may be responsible for the similar cytotoxicity observed from many disparate oligomeric proteins.

A number of studies have generated toxic amyloidogenic oligomers from proteins without disease association. E. *coli* protein HypF and human SH3 domain, neither associated with disease, revealed cytotoxicity when aggregated. These structures showed overlapping cross-reactivity to antibodies and a loosely packed hydrophobic core, again suggesting shared structural motifs [Campioni et al., 2010, Bucciantini et al., 2002, Kayed et al., 2003].

Whilst these studies lend significant support to a common structural motif and mechanism between amyloid intermediates and toxicity, there are a number of limitations. The broad range of species covered under the "oligomeric intermediates" umbrella, including, but not limited to, circular aggregates, SDS resistant aggregates, centrifugable aggregates, non-centrifugable aggre-

gates, protofibrils and paranuclei [Haass and Selkoe, 2007], and the lack of definition or specificity encouraged by such vagueness makes classification of oligomers into meaningful categories difficult. Furthermore, these oligomers will undergo further aggregation, generating higher molecular weight species with potentially different toxicities/properties; therefore it is necessary to use oligomers in a suitable time frame and with well-defined aggregation conditions. Finally, little distinction is made between "oligomeric intermediates" that are stable species occupying a local energy minima, and unstable transition states occupying maxima.

1.1.4 Kinetics of protein aggregation

To study the role of fibrils and oligomers *in vivo*, it is necessary to understand the kinetic processes which generate them. The ability of amyloid to bind ThT provides an easily accessible readout for growth of amyloid mass (see figure 1.2A). ThT binding curves can be divided into 3 main regions, a 'lag' phase, during which the total amyloid is relatively low, a growth phase, during which rapid amyloidogenesis is observed, and a stationary phase where equilibrium has been reached.

During each of the three phenomenologically observed phases several reaction processes are occurring simultaneously. Deconvolution of these processes can be estimated by combining robust experimental design (for example, dilution series) with careful, global data analysis aiming to reproducibly explain aggregation observables. Generally, simplified models of aggregation require a minimum of three species; free monomer, fibril and an intermediate nucleus capable of elongation. Nuclei are small aggregates with sufficient stability to

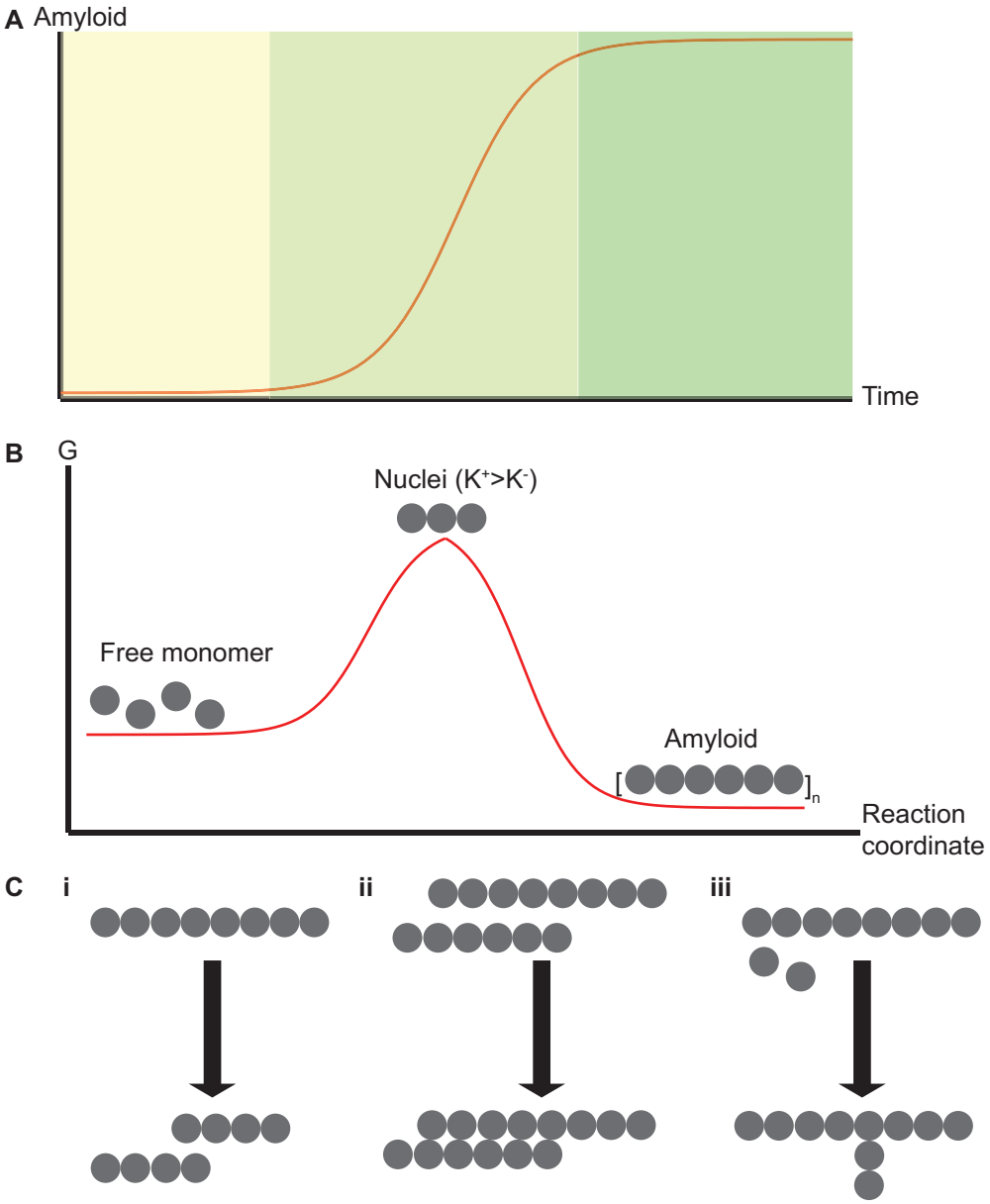


Figure 1.2: *Minimalist models of amyloid formation. A) a typical ThT curve observed for amyloidogenic proteins, lag phase (yellow) growth phase (light green) and stationary phase (dark green) are highlighted B) The simplest free energy diagram of species involved in amyloid formation, suggesting a low concentration of nuclei and thermodynamic stability of amyloid fibrils. In practice, the energy landscape is much more rugged than depicted here. C) Additional processes affecting amyloidogenesis, fragmentation (i), association (ii) and secondary nucleation (iii) are shown.*

allow monomer addition before monomer dissociation. Monomer and fibril are observed through biophysical techniques such as NMR, X-ray diffraction and ThT fluorescence (see section 1.1.2) whilst intermediate species are difficult

to observe directly due to their poor stability, low population and short lifetimes, instead intermediates are inferred from mass balancing or other analyses [Cohen et al., 2013].

As well as these three basic species, each defined by their own kinetics, other processes have been observed that contribute to amyloid formation. Given that elongation is end dependent, any process which alters the number of ends is capable of affecting amyloid formation rates. The most common processes are fibril fragmentation (figure 1.2Ci) [Xue et al., 2008], fibril association [Lee et al., 2012] (figure 1.2Cii) and secondary nucleation [Buell et al., 2014] (figure 1.2Ciii), where monomers attach to non-end surfaces of fibrils.

Each of these processes may be influenced by solution conditions such as pH, macromolecular crowding, agitation and hydrophobic interfaces (lipid membranes or air-water interfaces). For example, PD associated α-syn readily forms amyloid fibrils; this formation is dominated by monomer addition elongation reactions at neutral pH, and by both monomer addition elongation and secondary nucleation at acidic pH. This pH dependence may be due to the increasing hydrophobicity of α-syn C-terminus with acidity [Buell et al., 2014]. Similar enhancement of α-syn aggregation has been observed by modulating hydrophobic interfaces by introducing lipid membranes [Galvagnion et al., 2015, Pronchik et al., 2010]. Less well studied is the effect of macromolecular crowding on α-syn aggregation, however a correlation between crowding agent size and concentration with aggregation rates was observed for α-syn [Munishkina et al., 2004]. Notably, a 2 unit drop in pH or a 10 fold excess of amphipathic lipid increased α-syn amyloidogenesis by 10 and 1000 fold respectively [Galvagnion et al., 2015, Buell et al., 2014].

The difficulty of deconvoluting these amyloid curves into constituent parts

is represented in figure 1.3. Alteration of the rates for distinct microscopic processes, such as nucleation or elongation, can produce highly similar observations [Arosio et al., 2015].

In the case of Parkinson's disease, the role of kinetics in disease progression can be seen with the study of α-syn strains. Strains of α-syn generated either *de novo* or by repeated self-seeding were found to have distinct efficiencies in inducing hallmarks of Parkinson's disease and cell toxicity [Guo et al., 2013]. A second study produced distinct strains through alterations in the aggregation solution conditions. Thorough structural characterization through X-ray diffraction analysis, electron microscopy and proteinase K digestion confirmed the existence of distinct strains of α-syn. Once again a striking difference in cellular toxicity was observed [Bousset et al., 2013]. The relative balance of aggregation kinetics, which may favor the formation of one strain over another, is therefore pertinent to disease.

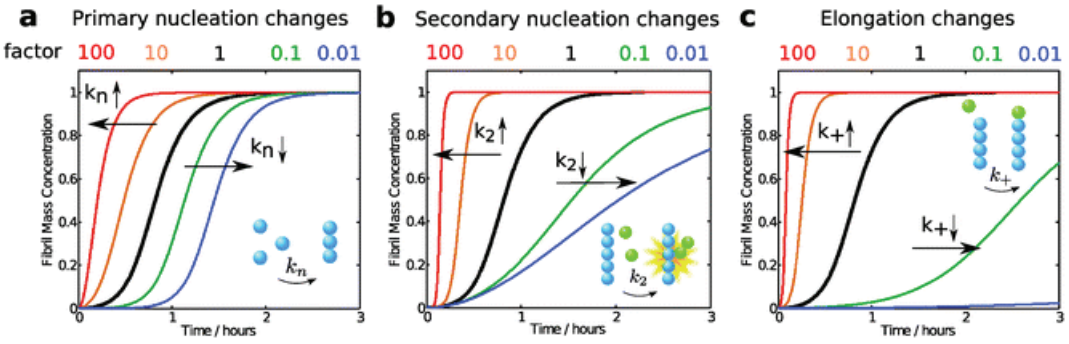


Figure 1.3: *Simulations of amyloid formation. Alterations in primary nucleation rates (a), secondary nucleation rates (b) and elongation rates (c) significantly affect the observed profile of amyloidogenesis. Image from [Arosio et al., 2015]*

1.1.5 Molecular chaperones in aggregation

In addition to the solution conditions outlined above, there are biological mechanisms to control and perturb protein aggregation. Specifically, molecular chaperones are responsible for delaying the inevitable aggregation of proteins into amyloid. These polypeptides are capable of intercepting and inhibiting substrate-protein misfolding and aggregation in adenosine triphosphate (ATP) independent or dependent manners (reviewed [Hilton et al., 2013] and [Saibil, 2013]). In practice, both ATP-dependent and independent chaperones work synergistically to maintain the integrity of the cell's protein network [Duennwald et al., 2012].

Chaperone function is therefore intimately linked to aggregation kinetics, the formation of amyloid is inevitable over long-periods, however, by binding aggregation nuclei, or sterically hindering fibril elongation, chaperones can reduce the growth of amyloid and thereby potentially the toxicity [Hilton et al., 2013].

Most notably, the small heat shock proteins (small HSPs) are a class of ATP-independent chaperones that form a first line of defense against protein misfolding and aggregation (reviewed [Cox et al., 2014]). Cell-based assays have suggested that human small HSPs are up-regulated in α-syn-oligomer induced stress scenarios, and are capable of reducing oligomer-associated toxicity [Outeiro et al., 2006]. Furthermore, small heat shock proteins have previously been observed bound to amyloid fibrils by electron microscopy, and their ability to delay aggregation is well-documented by ThT-based assays [Outeiro et al., 2006, Cox et al., 2014, Hochberg et al., 2014].

1.2 Aggregation in Parkinson's disease

The pathology of Parkinson's disease has been linked to the intrinsically disordered, 140 residue protein α-syn; the main focus of this thesis. Evidence for a casual link derives from PD associated inclusion bodies, known as Lewy Bodies (LB), which are mainly comprised of aggregated α-syn [Trojanowski and Lee, 2002]. Furthermore, early onset familial PD is linked to a number of mutations in α-syn (commonly, A30P, E46K and A53T) [Zarranz et al., 2004, Breydo et al., 2012]. In some mutations, α-syn aggregation kinetics are perturbed and result in more rapid aggregation at lower concentrations [Ahmad et al., 2012, Pandey et al., 2006, Li et al., 2001]. Furthermore the concentration of LBs in the *substantia nigra* correlates with cognitive decline in PD patients [Mattila et al., 2000, de Vos et al., 1995].

1.2.1 Structure, function and toxicity of alpha-synuclein

Structural studies of α-syn both *in vitro* and in-cell reveal an intrinsically disordered protein found in both intracellular and extracellular environments [Snead and Eliezer, 2014]. α-syn can be divided loosely into three regions; an N-terminal region associated with lipid interactions, a central region associated with amyloidogenesis and a disordered C-terminal region that remains free when lipid bound [Fusco et al., 2014]. In the presence of specific anionic lipids the N-terminus can fold into an α-helix and associate with lipids, a function associated with synaptic vesicle fusion [Snead and Eliezer, 2014]. Moreover, once associated, specific anionic lipids such as POPG (2-Oleoyl-1-palmitoyl-sn-glycero-3-phospho-rac–glycerol) or DMPS (1,2-Dimyristoyl-sn-glycero-3-phospho-L-serine) significantly modulate amyloidogensis of α-syn, whereby

high lipid:α-syn ratios preclude amyloidogenesis by sequestering monomeric α-syn, whilst low lipid:α-syn ratios enhance amyloidogenesis by providing a lipid bound population for nucleation [Wang et al., 2010, Lokappa and Ulmer, 2011, Drescher et al., 2008, Drescher et al., 2010, Braun et al., 2014, Galvagnion et al., 2015, Bodner et al., 2009].

The C-terminal region of α-syn reportedly remains free whilst lipid bound [Bodner et al., 2009, Eliezer et al., 2001, Fusco et al., 2014], displaying NMR relaxation properties consistent with a disordered polypeptide. Two enzymes are reported to interact with the C-terminus, Phospholipase D and Glucocerebrosidase. Both proteins process membrane lipids and are localized to the lysosome. Furthermore, interaction with α-syn inhibits the activity of both enzymes [Payton et al., 2004, Yap et al., 2013]. Finally, the central region regulates membrane affinity. Peptides of isolated α-syn central domain, lacking N and C termini, are capable of rapidly forming amyloid aggregates [Fusco et al., 2014]. The N and C termini form long range transient interactions shielding the hydrophobic core and possibly inhibiting amyloidogenesis [Dedmon et al., 2005].

Oligomers of α-syn have been reported to be the most toxic species, in line with other amyloid disease studies (see above). Mutations of α-syn that enhanced fibril formation or oligomer formation were found to be less and more toxic, respectively [Winner et al., 2011]. A major question underlying the toxicity of α-syn and other amyloid diseases is whether the toxic oligomeric species correspond to on-pathway amyloid products or by-products of this process [Lorenzen et al., 2014].

1.2.1.1 α-synuclein amyloid

Low resolution studies of α-syn amyloid using electron microscopy and antibody binding studies suggest that α-syn retains solvent accessibility in its C-terminus whilst sequestering its N-terminus and central domains into the amyloid core [Crowther et al., 2000]. Indeed, a short peptide derived from the central domain (residues 71-82) is capable of forming amyloid fibrils alone [Giasson et al., 2001].

Most recently, Tuttle *et al* reported a high resolution structure of α-syn amyloid derived from extensive solid state NMR study. Through careful preparation of α-syn fibrils, Tuttle *et al* were able to prepare multiple, differently labeled samples with a common structure. Combining data from each labeling scheme generated the first complete, high resolution structure of α-syn amyloid. Consistent with previous reports, the Tuttle *et al* structure reveals compact, folded N-terminal and central domains with a flexible C-terminus. Interestingly, as previously noted, common familial PD mutations (such as A30P, E46K and A53T) localize to the hydrophobic core of this folded domain [Tuttle et al., 2016, Comellas et al., 2011]. Whilst structures of A30P and A53T mutants are broadly similar to the wild-type structure presented by Tuttle *et al*, the E46K mutant reveals substantial chemical shift perturbations. These findings can be rationalized by the presence of a key wild-type E46-K80 salt bridge that is disrupted by the E46K mutation [Tuttle et al., 2016].

Structural characterization of monomeric, disordered α-syn may suggest pathological mechanisms for the other mutations. Notably, the central, pro-amyloidogenic, region of α-syn (including residues 71-82) was found to be occluded by monomer dynamics. Paramagnetic relaxation enhanced (PRE)

NMR studies of α-syn, with PRE tags on N-terminal, central and C-terminal residues, found that the central region of α-syn was in intramolecular transient contact with N and C termini [Bertoncini et al., 2005]. Subsequent FRET studies examined A30P and A53T familial PD mutations and found the average pairwise distances within the N-termini were longer for the A30P mutant compared to wild type [Lee et al., 2004b]. Taken together these results suggest that α-syn adopts a conformationally flexible ensemble of structures containing long range interactions, and that familial PD mutations may modulate these shielding interactions, promoting amyloidogenesis.

Some of these findings were recently confirmed by a landmark in-cell NMR study observing α-syn in mammalian cells. NMR studies revealed that α-syn is indeed disordered in the cell. However, macromolecular crowding also increases the shielding of the amyloidogenic central domain by partially collapsing the extended structure of α-syn [Theillet et al., 2016].

1.2.1.2 α-synuclein - Monomer or tetramer

A controversy over the native conformer of α-syn erupted recently with the suggestion that physiologically relevant α-syn exists as a soluble tetramer with N-terminal α-helical structure. Tetrameric α-syn was derived from human cells and studied by analytical ultracentrifugation and cross linking, with the authors reporting that human cell derived α-syn resisted aggregation more readily than recombinant bacterial derived α-syn and posited an unfolding step between tetrameric and disordered α-syn as necessary for aggregation (and therefore disease pathology) [Wang et al., 2011, Bartels et al., 2011].

Since these findings significant efforts have been made to identify native α-syn. In opposition to the tetrameric claim, an exhaustive multi-laboratory

study using a multitude of biophysical techniques on α-syn derived from multiple sources also concluded that α-syn was a disordered monomer natively [Fauvet et al., 2012]. Furthermore, human patient derived and bacterial recombinant α-syn revealed identical size exclusion profiles, eluting as a disordered monomer migrating at a rate equivalent to a folded 60 kDa protein. These findings are in line with previous observations by circular dichroism and small angle X-ray scattering suggesting α-syn to have a larger hydrodynamic radius (0.41nm) than a theoretical folded protein of equal mass (0.15nm), but slightly less that a completely unfolded amino acid chain (0.5nm)[Gast et al., 1995, Weinreb et al., 1996]. Finally, in human cell NMR of α-syn revealed the same disordered monomeric protein structure as predicted by previous in *in vitro* studies [Theillet et al., 2016]. Overall, this suggests that α-syn is natively monomeric and that chemical cross-linking studies must be interpreted with caution.

1.2.2 Parkinson's disease and Gaucher's disease: overlapping phenotypes, overlapping proteins

Recent clinical studies have linked PD to the lysosomal lipid lytic enzyme Glucocerebrosidase (GCase). GCase is a 60 kDa glycoprotein found in the acidic compartment of the lysosome, where it catalyzes the conversion of glucocerebroside to glucose and ceramide [Mehta, 2006]. Mutation of GCase causes an accumulation of glucocerebroside and Gaucher's disease - a lysosomal storage disorder with a spectrum of phenotypes ranging from organomegaly to neuropathy [Grabowski, 2012]. So far, over 300 mutations have been characterized for GCase: it is estimated that whilst homozygous disease mutations

are relatively rare (1/20,000), heterozygous mutation carriers are far more frequent[Hruska et al., 2008]. Of the 300 mutations studied to date, no clear correlation exists between structure and the severity of disease[Lieberman, 2011]. However, emerging evidence suggests that the fundamental disease phenotype is linked to the accelerated endoplasmic reticulum associated degradation (ERAD) of GCase [Ron and Horowitz, 2005, Bendikov-Bar et al., 2011]. This implies that mutants which destabilize GCase are cleared from the cell and degraded by the proteasome, reducing the overall cellular concentration of GCase and the total enzymatic activity.

The stability and activity of GCase is further affected by post translational modification. GCase has 4 N-linked glycosylation sites (N19, N59, N146, N270) and a putative 5th site at N462 [Brumshtein et al., 2006]. Ablation of the first site significantly perturbs enzymatic activity, whilst ablation of the other sites negatively effects enzyme stability [Berg-Fussman et al., 1993, Grace and Grabowski, 1990, Brumshtein et al., 2006]. The requirement for post translationally modified GCase has hindered the biophysical study of this disease-associated enzyme due to its incompatibility with E. *coli* expression systems. X-ray crystallography, however, has been able to derive GCase from eukaryotes, and there are a wealth of crystal structures of wild type GCase with various cofactors or amendments.

GCase is available from several pharmaceutical companies who produce it as part of their enzyme replacement therapeutic strategy. VPRIV (Shire PLC) is the only commercial GCase with fully native amino acid sequence. It is produced from human fibroblast cells treated with kifunensine, a mannosidase I inhibitor, producing GCase with carbohydrates favorable for internalization. VPRIV glycans include high mannose type glycans of six to nine

mannose units as the dominant glycan, and less dominant glycoforms (2%) include hybrid mannose-6-phosphate, mono-sialyated and fucosylated glycans [Brumshtein et al., 2010].

As well as a role in Gaucher's disease, mutation in GCase is the leading genetic risk factor for PD, with 5-10% of PD patients carrying one of hundreds of possible GCase mutations [Beavan and Schapira, 2013] (figure 1.4). Moreover, sporadic PD reveals wild type GCase colocalised with Lewy bodies and an overall depletion of GCase in the neuronal cells of the *substantia nigra*, a key PD associated brain structure. Overall, GCase mutation increases the risk of PD by up to 30 fold [Bultron et al., 2010, McNeill et al., 2012].

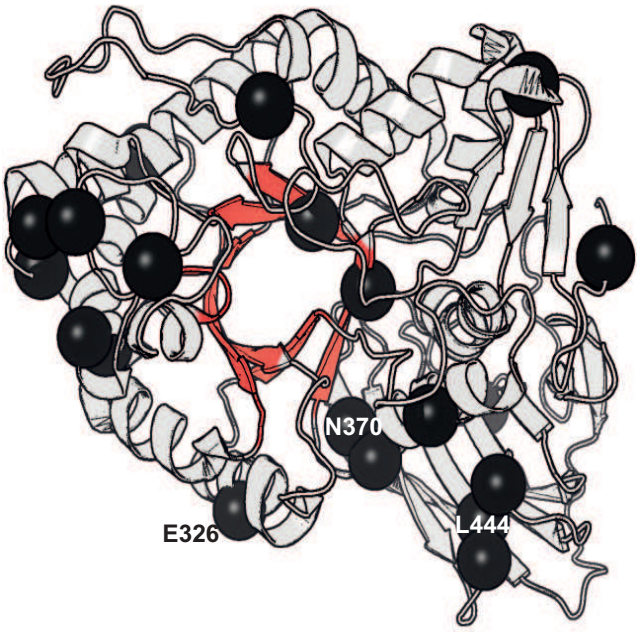


Figure 1.4: *A structure of GCase showing some mutation sites linked to PD (black), together with the active site (red). Key mutation sites (E326, N370, L444) for discussion are labeled (PDB code: 2V3E).*

Lysosomal GCase is depleted in both sporadic and non-sporadic PD [Chiasserini et al., 2015, Murphy et al., 2014, Bendikov-Bar et al., 2011, Ron and Horowitz, 2005, Mazzulli et al., 2011, Gegg et al., 2012, Winder-Rhodes et al., 2013, Bae et al., 2015, Nichols et al., 2009, Murphy and Halliday, 2014]. Healthy cells can clear α-syn and its oligomers through autophagic pathways which culminate in the lysosome [Wong et al., 2004, Lee et al., 2004a, Cuervo et al., 2004, Webb et al., 2003, Mak et al., 2010]. The intercellular transfer

of potentially toxic α-syn oligomers between lysosomes is also enhanced by GCase depletion [Mazzulli et al., 2011, Bae et al., 2014, Alvarez-Erviti et al., 2011]. Consistent with this, overexpression of GCase inhibits synucleinopathy in PD animal models [Rocha et al., 2015, Sardi et al., 2011, Sardi et al., 2013]. Finally, α-syn interacts with GCase directly at lysosomal pH to inhibit GCase activity [Yap et al., 2013]. Whilst these links stress the importance of GCase to PD, the protective ability of wild type GCase against PD has yet to be investigated.

The aggregation of α-syn into amyloid has been studied extensively *in vitro* [Giehm et al., 2010, Buell et al., 2014, Hoyer et al., 2002, Galvagnion et al., 2015]. Crucially, its aggregation into amyloid fibrils is accelerated 1000 fold in the presence of membrane mimicking anionic lipids such as DMPS and POPG (see section 1.1.4) [Galvagnion et al., 2015, Wang et al., 2010, Lokappa and Ulmer, 2011, Drescher et al., 2008, Drescher et al., 2010, Braun et al., 2014, Bodner et al., 2009]. Furthermore, acidic environments, like those found in the lysosome, further promote α-syn amyloidogenesis [Buell et al., 2010, Wu et al., 2009, Hoyer et al., 2002]. The link between GCase, acidity and lipids suggest the lysosome as a likely organelle for the onset of α-syn pathogenesis (figure 1.5) [Goker-Alpan et al., 2010, Mazzulli et al., 2011, Bae et al., 2015, Lee et al., 2004a, Mak et al., 2010, Alvarez-Erviti et al., 2011, Liu et al., 2015, Cuervo et al., 2004].

1.3 Thesis overview

This thesis studies the role of protein folding and aggregation in human health and disease. We begin with an introduction to NMR and other techniques used

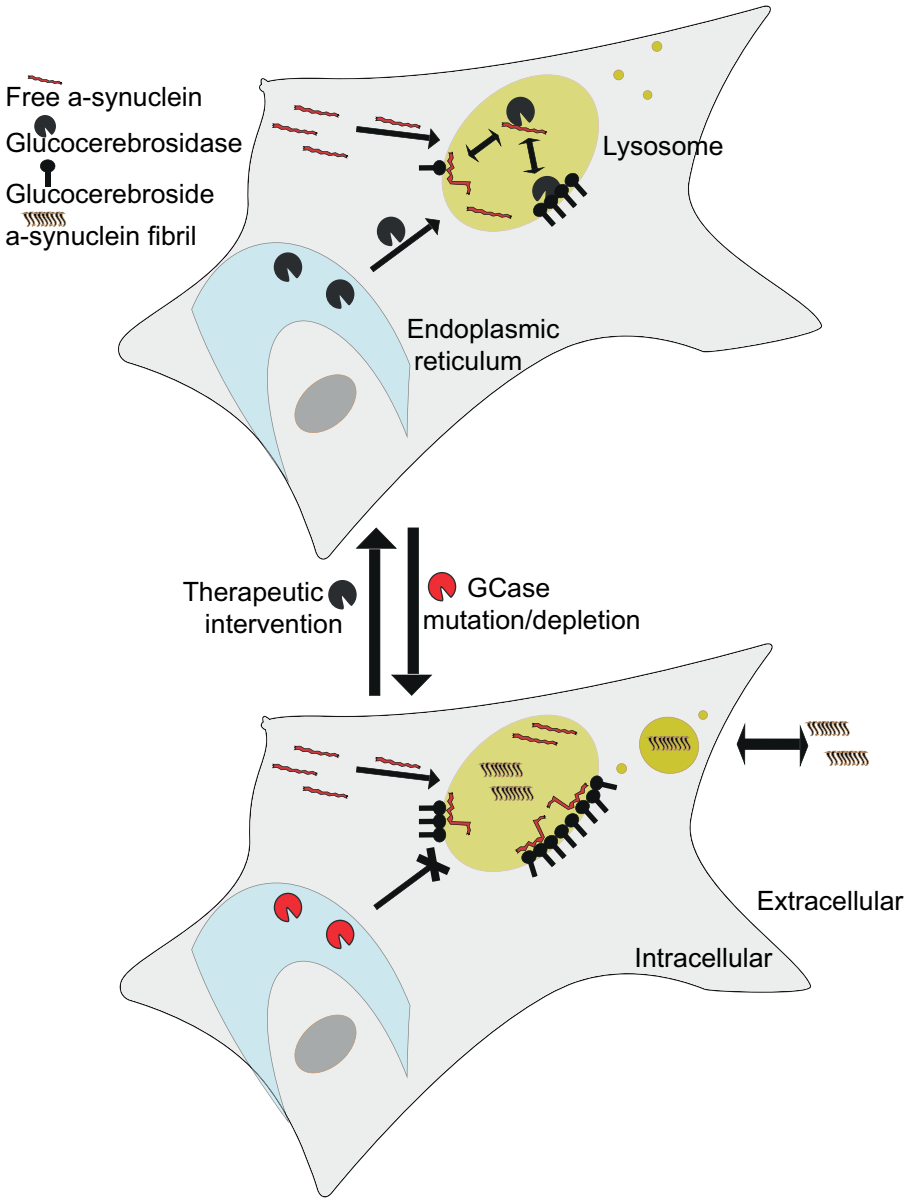


Figure 1.5: *GCase is produced in the endoplasmic reticulum where it is trafficked to the lysosome. Lysosomal GCase binds α-syn and also enzymatically degrades glucocerebroside. Mutation in GCase prevents it's trafficking to the lysosome, thereby allowing α-syn to aggregate and glucocerebroside to build up, these are the pathologies of PD and GD, respectively.*

in this thesis (chapter 2). In chapter 3, NMR labels are introduced onto GCase by reductive amine methylation with ^{13}C formaldehyde. These labels are partially assigned through altering reaction conditions and liquid chromatography tandem mass spectrometry. Labels are then exploited to observe human cell produced GCase, complete with post translational modifications, interacting

with potential GD therapeutics and in human cell lysate.

Chapter 4 builds on the work of chapter 3, utilizing labeled GCase to examine the three way interaction between GCase, α-syn and anionic lipids under lysosomal conditions. A combination of NMR spectroscopy, electron microscopy and ThT fluorescence assays are used alongside computational modeling to deduce a protective mechanism of action for GCase in rescuing α-syn from aggregation.

Chapter 5 examines the role of chaperones in protein aggregation. A model system, α-lactalbumin (α-lac) and small heat shock protein 16.5 (sHSP16.5) is introduced. Aggregation of αlac can be induced through reducing agents, providing a reliable and reproducible system to study. Likewise, sHSP16.5, derived from thermophillic archaea, is a model chaperone, producing high quality NMR spectra. This system is studied by mass spectrometry and NMR spectroscopy to provide insight into mechanisms of chaperoning and aggregation.

Finally, in chapter 6, the enhanced crystallization of Caffeine induced death suppressor 1 (Cid1) and the expression of homologous mammalian proteins are reported. Such proteins edit RNAs to achieve fine tuning of protein expression, and some of the techniques used in chapters 3-5 were initially developed here.

Chapter 2

Nuclear magnetic resonance and other techniques

2.1 Nuclear magnetic resonance: An introduction

Nuclear magnetic resonance (NMR) is a physical technique capable of providing information on the structure and dynamics of biomolecules at atomistic resolution. NMR is based on the intrinsic spin of sub-atomic particles; atomic nuclei with odd-numbered protons or neutrons exist in multiple orientations with degenerate energy levels in the absence of a magnetic field (orientations are termed spin-up or spin-down for spin $\frac{1}{2}$ nuclei). However, in the presence of a strong magnetic field such orientations are split into higher and lower energy levels, their populations given by the Boltzmann distribution. As nuclei transition between upper and lower energy levels following excitation they emit energy which is observable by NMR. Commonly used isotopes in biological

NMR include spin $\frac{1}{2}$ nuclei such as ^{1}H, ^{13}C and ^{15}N, alongside ^{31}P for nucleic acid NMR.

The difference between energy levels (ΔE) is given by:

$$\Delta E = -\frac{\gamma h}{2\pi} mB \tag{2.1}$$

Where γ is the gyro-magnetic ratio for a specific nuclei (e.g. $^{1}H = 2.675 \times 10^{8}$ *rad. s.* T^{-1}), h is Planck's constant, m is the magnetic quantum number and B is the magnetic field at the nucleus. A larger gyromagnetic ratio or magnetic field will yield a larger ΔE and therefore a larger difference in ground and excited populations; contributing to NMR signal intensity (the area under a peak).

Samples placed in a magnet will align with the static magnetic field (known as B_0) along the z-axis. Subsequent radio frequency pulses are then applied to rotate magnetization into x and y planes. The resulting magnetic moments precess about B_0 at a frequency that is dependent on local environment [Abragam and Goldman, 1982].

In NMR, the sum of this precession is observed as a free induction decay (FID) induced in the receiver coil of the NMR instrument. FIDs decay with time as they lose coherence; this decay rate is a key NMR observable known as relaxation (see below).

A Fourier transform of the FIDs yields an NMR spectrum where each nuclei

has four principal observables; chemical shift ($\delta\omega$), intensity (I), linewidth and J couplings. The chemical shift is exquisitely sensitive to local environment, yielding information on local structure (disorder, helicity etc). Linewidth contains measures of dynamics, intensity, coarsely, is a measure of the concentration of nuclei provided all other observables are equal (i.e. linewidth, chemical shift) and J-coupling can be exploited to yield information on chemical bonds. These observables are not simple to interpret as they convolve a number of phenomena, for example, chemical shift reports on disorder, torsion angles and the presence of charge simultaneously, and is further effected by chemical exchange.

2.2 Studying dynamics by NMR

The greatest strength of NMR lies in its ability to simultaneously study structure and dynamics at atomic resolution; many probes can be introduced non-perturbingly to a protein system, each probe can report on dynamics spanning timescales from seconds to picoseconds and additionally provides information on chemical environment. Finally, NMR is capable of studying transient and weak interactions between multiple partners simultaneously.

The most relevant dynamics observed in this thesis are those probed by Carr-Purcell Meiboom-Gill (CPMG) relaxation dispersion experiments and relaxation experiments. These experiments report on chemical exchange and local dynamics, respectively, yielding information covering microsecond-millisecond

(CPMG) and picosecond-nanosecond (nuclear spin relaxation experiments) timescales. Both experiments report on R^2 relaxation rates; the loss of transverse magnetization coherence

2.2.1 NMR and Chemical exchange

Chemical exchange is the process by which an NMR probe samples two or more distinct environments, such as a change in conformation. A simple exchanging system may look like equation 2.2, which depicts two site exchange. The apparent effect of chemical exchange is dependent on the size of the chemical shift ($\delta\omega$) and the k_{ex} ($k_{ex} = k_{on} + k_{off}$) (figure 2.1).

$$A \underset{k_off}{\overset{k_on}{\rightleftharpoons}} B \tag{2.2}$$

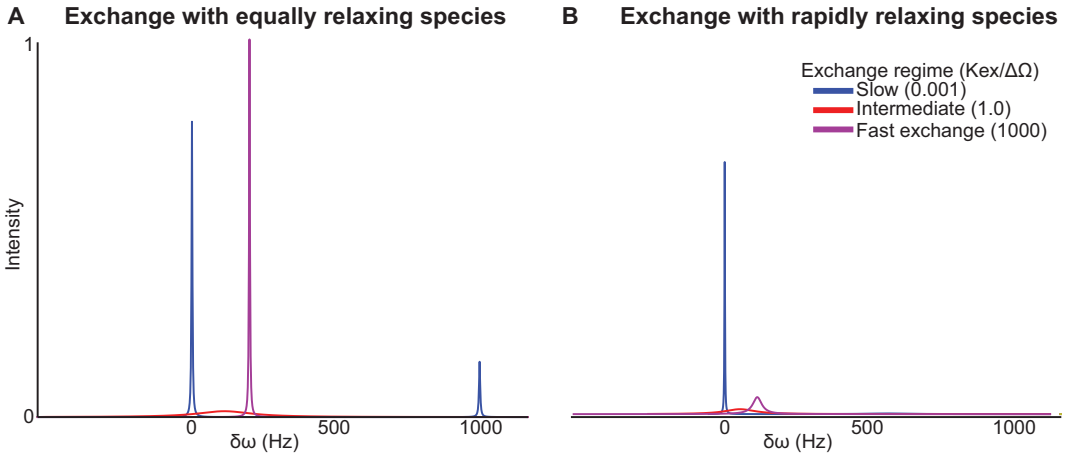


Figure 2.1: *A) Simulated chemical exchange between two species with equal relaxation properties. B) simulated exchange between two species with 100 fold difference in relaxation rates. In the case of intermediate or fast exchange, a clear peak broadening is expected. Slow exchange, however, displays a single peak attenuated in intensity. Simulations run with: minor population of 20%, and a $\delta\omega$ of $2\pi 1000$ rad s^{-1}.*

In a fast exchange regime (k_{ex} (s^{-1}) $>\delta\omega$ (rad/s)), a peak is observed that

has a weighted chemical shift and relaxation rate of the two environments (figure 2.1 Purple). In slow exchange (k_{ex} s^{-1} $<\delta\omega$ *(rad/s)*), two distinct peaks are observed at chemical shifts particular to their environment, the size and width of the peaks is determined by the relative relaxation rates of each species (figure 2.1 blue). Intermediate exchange (k_{ex} s^{-1} $=\delta\omega$ *(rad/s)*) leads to severe peak broadening (figure 2.1 red). Clearly, there also exist slow-intermediate and fast-intermediate regimes. Finally, spectra can be complicated by the second species being invisible to NMR, for example, when binding a large, slowly tumbling cofactor. In this case, a second peak is never seen, but broadening of the original peak allows differentiation between fast and slow exchange limits.

2.2.1.1 The Carr-Purcell-Meiboom-Gill (CPMG) NMR experiments

The CPMG experiment provides structural and kinetic information on protein dynamics occurring on the millisecond timescale, including events linked to domain reorganization, ligand binding and dimerization [Neudecker et al., 2009, Kempf and Loria, 2003].

Each CPMG experiment contains a variable number of spin-echo elements (also known as CPMG elements); which consists of a delay, 180° pulse, and another delay - over a fixed time period (figure 2.2). In the absence of exchange, magnetization is completely refocused and signal is independent of the frequency of spin-echos. However, in the presence of exchange, magnetization will not be refocused efficiently and signal will be lost. The so called 'CPMG effect' is the variation in apparent R^2 with the CPMG pulse frequency, and

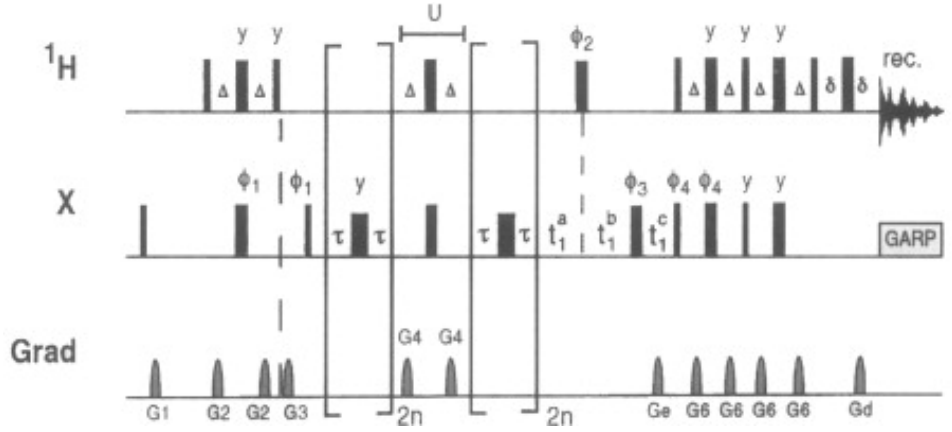


Figure 2.2: *The Relaxation-compensated CPMG pulse sequence reports on slow time scale dynamics through repeated 180° pulses (τ-180°-τ) and a relaxation period to mix in-phase and anti-phase coherences (U period) . Narrow and wide bars represent 90° and 180° pulses respectively. GARP pulse provides decoupling during acquisition, gradients are for coherence selection. Unlabeled pulses are x-axis. (From [Palmer et al., 2001])*

is dependent on the length of the fixed time period and the timescale of the exchange. This relationship between the frequency of spin-echoes and signal intensity is exploited to yield quantitative information on the rate of exchange (k_{ex}), the chemical shift difference ($\Delta\omega$), the relative population differences between major and minor conformers (named P_A and P_B, respectively) and the deconvoluted relaxation rates [Baldwin et al., 2009, Baldwin and Kay, 2012, Baldwin et al., 2010, Tollinger et al., 2001, McClendon et al., 2009].

In a system undergoing two-site exchange between a major (P_A) and a minor (P_B) conformer (as in equation 2.2) the relationship between the CPMG frequency and the signal intensity can be related via equation 2.3, where T_{rel} is the fixed time CPMG period and I and I^0 is the intensity in the presence and absence of a CPMG train, respectively.

$$R_2^{eff} = -\frac{1}{T_{rel}} \ln \frac{I(T_{rel})}{I^0} \tag{2.3}$$

Magnetization in the absence of pulses evolves according to equation 2.4, an adapted Bloch McConnel equation [Baldwin, 2014]:

$$\frac{d}{dt} \begin{bmatrix} I_G^+ \\ \\ I_E^+ \end{bmatrix} = R^+ \begin{bmatrix} I_G^+ \\ \\ I_E^+ \end{bmatrix} \tag{2.4}$$

Where G and E refer to ground and excited conformers, respectively, and the shift basis is used ($I_+ = I_x + iI_y$ and $I_- = I_x\ iI_y$). The matrix for spin evolution is given by equation 2.5:

$$R^+ = \begin{bmatrix} -k_{GE} - R_2^G & k_{EG} \\ \\ k_{GE} & -k_{EG} - R_2^E - i\Delta\omega \end{bmatrix} \tag{2.5}$$

Where k_{GE} and k_{EG} are the rates for ground-to-excited and excited-to-ground exchange, respectively (and k_{ex}=k_{EG}+k_{GE}). R^2G and R^2E represents the relaxation rates for ground and excited conformers, respectively, and $\Delta\omega$ represents the chemical shift difference between ground and excited conformers (in $rad\ s^{-1}$).

The relationship in equations 2.4 and 2.5 can be used to derive approximate solutions to the analytical equations in figure 2.3. The most commonly

A

$$R_2\left(1/\tau_{cp}\right) = \frac{1}{2}\left(R_{2A}^0 + R_{2B}^0 + k_{ex} - \frac{1}{\tau_{cp}}\cosh^{-1}\left[D_+ \cosh\left(\eta_+\right) - D_- \cos\left(\eta_-\right)\right]\right)$$

$$D_\pm = \frac{1}{2}\left[\pm 1 + \frac{\Psi + 2\Delta\omega^2}{\left(\Psi^2 + \zeta^2\right)^{1/2}} \right]$$

$$\eta_\pm = \frac{\tau_{cp}}{\sqrt{2}}\left[\pm\Psi + (\Psi^2 + \zeta^2)^{1/2} \right]^{1/2}$$

$$\Psi = (R_{2A}^0 - R_{2B}^0 - p_A k_{ex} + p_B k_{ex})^2 - \Delta\omega^2 + 4 p_A p_B {k_{ex}}^2$$

$$\zeta = 2\Delta\omega(R_{2A}^0 - R_{2B}^0 - p_A k_{ex} + p_B k_{ex})$$

$$k_{ex} = k_{AB} + k_{BA}$$

$$p_B = \frac{k_{AB}}{k_{ex}}$$

$$p_A = \frac{k_{BA}}{k_{ex}}$$

B

$$R_2 = R_2^0 + \frac{p_A p_B \Delta\omega^2}{k_{ex}}\left(1 - \frac{4\nu_{CPMG}}{k_{ex}}\tanh\left(\frac{k_{ex}}{4\nu_{CPMG}}\right)\right)$$

C

$$R_2 = R_2^0 + k_{AB}\left(1 - \frac{\sin\left(\Delta\omega \cdot \tau_{CP}\right)}{\Delta\omega \cdot \tau_{CP}}\right)$$

Figure 2.3: *A) Carver Richards general equation for $R^2 eff$. τ_{cp} is the time between CPMG pulses, R^2 is the relaxation rate of each species (A or B), k_{ex} is the exchange rate, $\Delta\omega$ is the change in chemical shift and P_A is the population of conformer A. B) The Meiboom equation (notation as above). C) The Tollinger equation (notation as above) [Baldwin, 2014, Carver and Richards, 1972, Meiboom and Gill, 1958, Tollinger et al., 2001]*

used equations include the Meiboom equation - applicable to fast ($k_{ex} > \delta\omega$) exchange regimes, and the Tollinger equation - applicable to slow (k_{ex} *(rad/s)* $< \delta\omega$) exchange regimes. The Carver-Richards formula is robust for slow and fast exchange, and is able to extract all potential parameters from ideal CPMG data. However, the derived parameters can deviate substantially from the data when the population of the sparsely populated conformer (P_B) is <1%. Most recently, the Baldwin equation provides exact solutions to the analytical equation 2.4.

$$R_{2,eff} = R^{CR}_{2,eff} - \frac{1}{T_{rel}} \ln\left(\frac{1+y}{2} + \frac{(1-y)v_5}{2N\sqrt{v_{1c}^2-1}}\right)$$

$$R^{CR}_{2,eff} = \frac{R_2^G + R_2^E + k_{EX}}{2} - \frac{N_{cyc}}{T_{rel}} \cosh^{-1}(v_{1c})$$

$$h_3 = \frac{1}{\sqrt{2}}\sqrt{h_2 + \sqrt{h_1^2 + h_2^2}}$$

$$h_4 = \frac{1}{\sqrt{2}}\sqrt{-h_2 + \sqrt{h_1^2 + h_2^2}}$$

$$f_{00} = \frac{1}{2}(\Delta R_2 + k_{EX} - h_3) + \frac{i}{2}(\Delta\omega - h_4)$$

$$N = h_3 + ih_4$$

$$NN^* = h_3^2 + h_4^2$$

$$F_0 = (\Delta\omega^2 + h_3^2)/NN^*$$

$$F_2 = (\Delta\omega^2 - h_4^2)/NN^*$$

$$F_1^b = (h_4 + \Delta\omega)(\Delta\omega - ih_3)/NN^*$$

$$E_0 = 2h_3$$

$$|E_2| = 2h_4$$

$$E_1 = h_3 - ih_4$$

$$v_{1s} = F_0 \sinh(\tau_{cp}E_0) - iF_2 \sin(\tau_{cp}|E_2|)$$

$$v_4 = -2(k_{GE} - f_{00})F_1^b + k_{GE}(2\Delta\omega^2 - ih_1)/NN^*$$

$$v_5 = (\Delta R_2 + k_{EX} + I\Delta\omega)v_{1s} - 2v_4 \sinh(\tau_{cp}E_1)$$

$$y = \left(\frac{v_{1c} - (v_{1c}^2-1)^{1/2}}{v_{1c} + (v_{1c}^2-1)^{1/2}}\right)^{N_{cyc}}$$

Figure 2.4: *The 2014 Baldwin equation. T_{rel} is the fixed time CPMG period. τ_{CP} is the delay before and after the CPMG 180° element. R^2G and R^2E are the relaxation rates of the ground and excited conformers, respectively and ΔR^2 is the difference in relaxation rates. $\Delta\omega$ is the chemical shift difference in rad/s and P_B is the population of the minor conformer [Baldwin, 2014].*

The Baldwin equation is capable of providing exact solutions over a range of P_B values [Baldwin, 2014]. However, in practice, off-resonance and differential relaxation effects may lead to deviations from the exact solution [Baldwin, 2014]. For this reason numerical solutions (e.g. Catia [Hansen, 2008]) are preferable where there is no alternative, as well as when the problem to be solved is computationally intensive.

CPMG represents a robust method to observe millisecond exchange phenomena and has been previously used to characterize the relationship between structure and catalytic rates, for example, the prolyl isomerase CypA reveals dynamics in the free conformer that are associated with the substrate-bound conformer, suggesting the apo conformer samples the 'active' conformer [Eisenmesser et al., 2005]. Much of the evidence for such conformational selectivity derives from CPMG experiments [Baldwin and Kay, 2009].

In chapter 3 of this thesis, I use the analytical Baldwin solution to examine the CPMG observable dynamics of GCase. For GCase, the analytical solution is appropriate due to the extensive overlap in the NMR spectrum making numerical analysis unfeasible. Additionally, I apply numerical analysis (CATIA) to the well dispersed spectra for small heat shock protein 16.5 in chapter 5 [Hansen, 2008].

2.2.2 Relaxation

Unlike the CPMG experiment, which reports on multiple exchanging conformers, canonical relaxation experiments extract R^2 and R^1 parameters that report

protein dynamics *within* an energy well that is undergoing slow exchange with a secondary energy well. Such sampling of the energy well reflects picosecond-nanosecond timescale fluctuations.

Broadly, relaxation is the return to thermal equilibrium after the absorption of energy. This can be thought of as the recovery of magnetization aligned to the static magnetic field. Relaxation is stimulated by the interaction of adjacent magnetic fields; the magnetic moment associated with each spin oscillates as the environment changes, this oscillation in turn affects the magnetic moment of neighboring spins, leading to relaxation.

2.2.2.1 Nuclear spin relaxation NMR

In a nuclear spin relaxation experiment, we use these oscillating magnetic fields within the bulk solution to inform on local protein dynamics. Oscillation of local fields induced through chemical shift anisotropy (the dependence of bond orientation relative to B_0), dipolar coupling (through space interaction of nuclear spins) and paramagnetic events (unpaired electrons) stimulate relaxation in a manner dependent on local protein motion.

By relating the amplitude of these events to the total signal intensity we can derive relaxation rates that reflect local dynamics. In the simplest analysis, raw relaxation rates can be interpreted to provide insight into local motion, whilst more complex analysis (such as model free analysis methods) allow the derivation of order parameters which can be used to inform structural calculations.

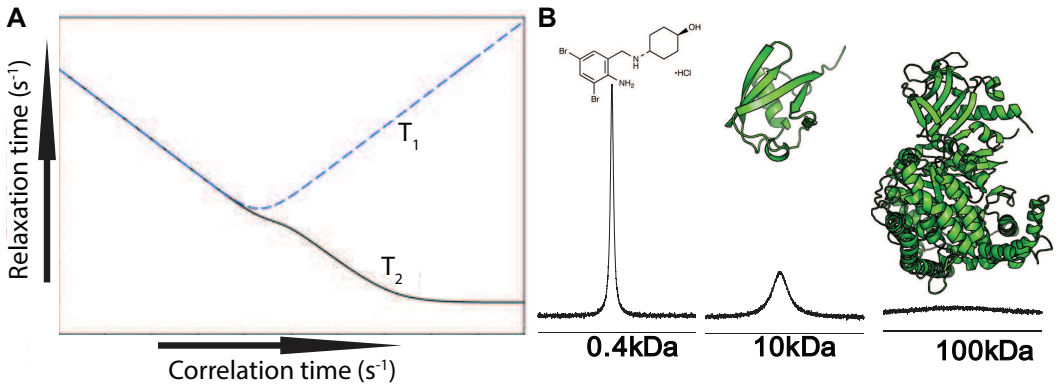


Figure 2.5: *Relaxation and molecular size A) Relaxation of nuclei as a function of correlation time, T_1=1/R_1 and T_2=1/R_2. Figure from [Cooper et al., 2013]. B) Simulated peak shapes for molecules of increasing size*

Relaxation rates are also influenced by the correlation time of a protein (figure 2.5); whereas R_2 increases with correlation time (t_c), R_1 has a maximum and subsequent decline (figure 2.5). The relation between correlation time and R^2 is one of the reasons NMR has historically been limited to smaller proteins, however novel advances have now pushed the limits of NMR to much larger protein systems.

2.2.3 NMR of large protein systems

Recent advances in NMR spectroscopy allow the examination of large (>30kDa) molecular weight species. The most revolutionary advance in technique came from the laboratory of Professor Lewis Kay (Toronto, Canada), who developed a robust method for introducing highly sensitive $^{13}CH_3$ groups into proteins. These methyl labels are highly sensitive due to favorable relaxation properties and the degeneracy of the 3 H groups, which all contribute to the same NMR resonance [Tugarinov et al., 2006, Goto et al., 1999]. Combination of selective labeling with methyl-TROSY pulse sequences allows the study of large com-

plex biological molecules, such as the proteasome and $\alpha\beta$-crystallin [Baldwin and Kay, 2012, Baldwin et al., 2010].

2.3 NMR and complementary techniques

In this thesis, NMR is used alongside a number of other techniques to elucidate structure, function and dynamics of biological systems. By combining the atomistic information on molecular dynamics from NMR with stoichometric measurements by mass spectrometry (MS), aggregate behavior by electron microscopy and high resolution data from X-ray crystallography, a more complete image of protein function can be ascertained. Specifically, the mechanisms and prevention of protein aggregation and Parkinson's disease amyloidogenesis.

Native protein mass spectrometry (MS) is a ground breaking tool allowing the quaternary structure and interactions of proteins to be studied. Whilst measurement occurs in the gas phase, structure and interactions are thought to be unperturbed due to the short timescale between ionization and detection. Native MS has been successfully used to characterize small heat shock proteins and a number of membrane proteins [Benesch and Ruotolo, 2011]. Whilst it is a technique with much potential, care must be taken in quantitative analysis. This is due to the intensity count being only relative and, more pervasively, the influence of uncontrolled factors (such as needle position) on observed species. Therefore, native MS data is presented here as a quantitative measure of mass, but a qualitative measure of populations. The combination of MS, NMR and

EM provides information on biological molecules from the atomistic to the quaternary structure. Together, these methods represent a complimentary, and highly effective, set of techniques to study biological systems.

Chapter 3

Glucocerebrosidase - nuclear magnetic resonance and native human proteins

Abstract

Glucocerebrosidase (GCase) is implicated in Gaucher's disease and Parkinson's disease. To accurately study GCase it is necessary to work with protein that has been glycosylated. Here, I describe a novel strategy to introduce NMR active nuclei onto large (>60kDa) human cell derived protein complete with relevant post translational modifications. This strategy is successfully employed to study millisecond GCase dynamics consistent with dimerization. This methodology is further utilized to assay GCase with potential therapeutics. I find that some potential therapeutics are capable of altering the dimerization rates of GCase, which may impact GCase stability.

3.1 GCase: a lysosomal protein implicated in human health and disease

Glucocerebrosidase (GCase) is a lysosomal protein responsible for the enzymatic processing of glucocerebroside lipids. Mutation in GCase is associated with Gaucher' disease (GD), a rare lysosomal storage disorder with a wide range of symptoms from organomegaly (over-sized organs) to neuropathy and thrombocytopenia [Hruska et al., 2008, Bendikov-Bar and Horowitz, 2012, Wong et al., 2004, Grabowski, 2012].

One of the current treatments for Gaucher's disease requires the intravenous infusion of recombinant GCase into patients [Barton et al., 1991]. There are multiple sources of recombinant GCase; plant cell, hamster cell and human cell derived GCase are all currently FDA approved. The main source of differentiation for these products is the presence of single point mutations and glycosylation patterns [Patten et al., 2007].

A pathological mutation in GCase is generally associated with its degradation via endoplasmic reticulum associated degradation (ERAD) and subsequent depletion in the lysosome [Bendikov-Bar et al., 2011, Ron and Horowitz, 2005]. Depletion or inactivation of GCase leads to a concomitant build up of GCase substrate glucocerebroside and the onset of GD symptoms [Hruska et al., 2008, Bendikov-Bar and Horowitz, 2012, Wong et al., 2004, Grabowski, 2012, Ron and Horowitz, 2005, Barton et al., 1991, Lieberman, 2011].

Many of the mutations in Gaucher's disease have subsequently been linked to Parkinson's disease (PD) (see section 1.2.2 and figure 3.1) [Siebert et al., 2014]. Mutation or 'sporadic' depletion of GCase is implicated causally in PD pathology [Mazzulli et al., 2011]. A build up of glucocerebroside may

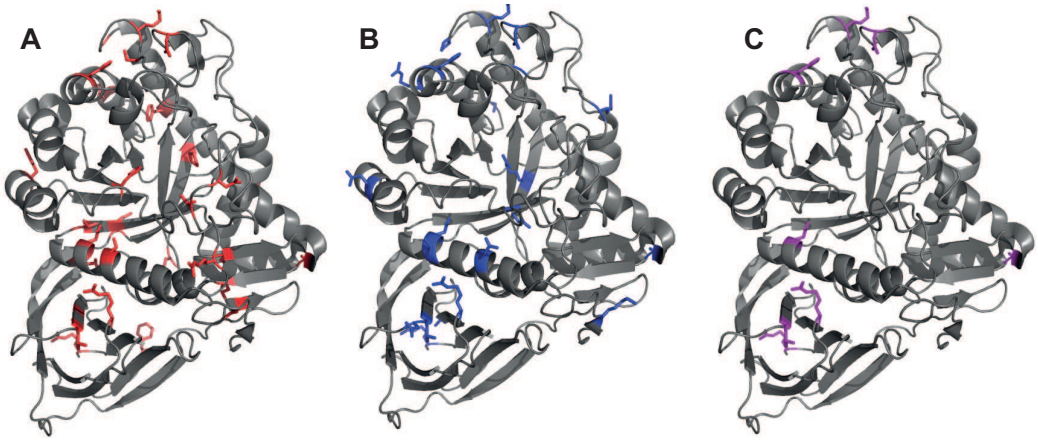


Figure 3.1: *Mutations associated with GCase. Mutations commonly associated with Gaucher's disease (A) and Parkinson's disease (B) often overlap (C) but show no structural localization. (PDB code 1OGS, mutation information from UNIPROT database.)*

itself contribute to PD pathology [Bae et al., 2015]. However, mutants such as E326K are linked to PD but not GD, suggesting that the GCase-PD links go beyond substrate accumulation [Park et al., 2002, Horowitz et al., , Siebert et al., 2014]. This hypothesis is further supported by the evidence of a direct protein-protein interaction between GCase and proteins associated with PD (see chapter 4) [Yap et al., 2011, Yap et al., 2013] and conflicting evidence on the role of GCase enzymatic activity on PD pathology [Liu et al., 2015, Dermentzaki et al., 2013].

The post translational modifications (PTMs) on GCase are known to be vital for GCase stability and activity. Specifically, 4 N-linked glycosylation sites at N19, N59, N149 and N270 are important for stability, trafficking and activity of GCase [Berg-Fussman et al., 1993, Grace and Grabowski, 1990, Brumshtein et al., 2006]. These PTMs are key to GCase studies and physiological relevance. To allow the study of GCase by NMR, it would theoretically be possible to label GCase in human cells [Banci et al., 2013], however, as well as being prohibitively expensive, the large molecular weight of GCase (60 KDa)

would result in broad peaks that are unsuitable for analysis. Therefore, approaches combining native mass spectrometry and novel NMR compatible labeling methods were employed to study human cell derived GCase and its interactions with binding partners.

Here, I have developed a novel labeling method capable of introducing NMR active labels onto human-cell produced GCase. I employ this method to observe GCase dynamics consistent with dimerization and small-molecule interactions with potential therapeutics. Together, this work represents a new strategy to study human produced proteins, as well as new insights into GCase.

3.2 Results: Studying native GCase by NMR

3.2.1 GCase reveals pH dependent dimer populations

Recombinant GCase (VPRIV) produced from human fibroblast cell lines was obtained from Shire PLC under material transfer agreement. VPRIV, unlike other recombinant GCase products, has a completely native sequence and is produced from human cell lines treated with kifunensine, yielding GCase glycoprotein with six to nine mannoses across its four N-linked glycosylation sites [Tekoah et al., 2013].

Initial observations of GCase by nano electrospray mass spectrometry (nESI MS) revealed a small pH dependent dimer population (figure 3.2) with a measured monomer mass of 63155±3 Da. nESI MS is a well established and versatile technique for studying protein interactions and conformations that are conserved from the solution to gas phases [Benesch and Ruotolo, 2011, Marcoux and Robinson, 2013, Ruotolo and Robinson, 2006].

The measured mass for the modal GCase species observed by nESI MS

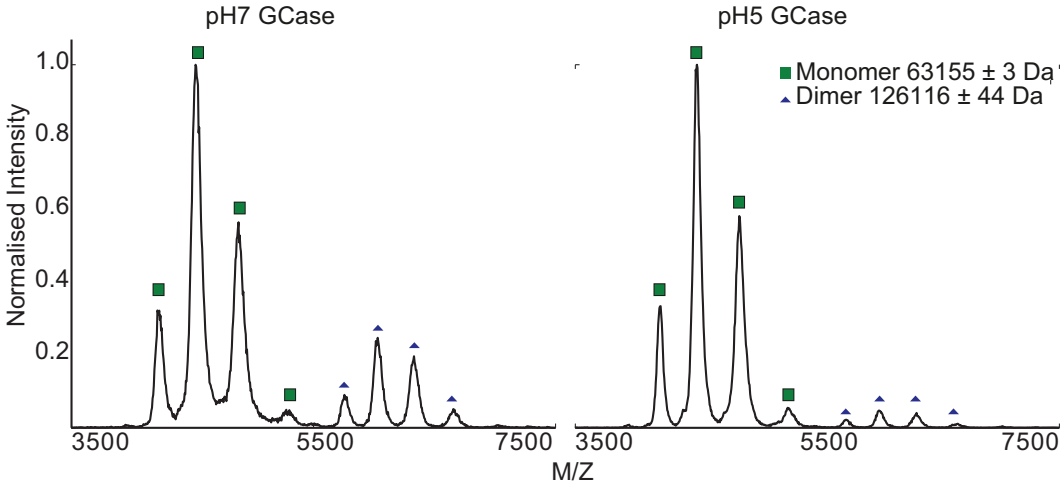


Figure 3.2: *GCase forms dimers at pH7 (left) but not significantly at pH5 (right) (10 μM GCase, pH7, 21°C)*

was 63155±3 Da by nESI MS (figure 3.2), this compares with a theoretical weight for the naked protein of 55597 Da. Thus, approximately 8KDa must be accounted for. This is most likely made up of PTMs such as the extended glycans groups. The predominant glycosyl groups incorporated consist of mannose (180 Da) and N-Acetylglucosamine (221 Da). A total of 36 mannose, 8 N-Acetylglucosamine per GCase are reported from liquid chromatography-ESI-MS studies, yielding a combined weight of 8 KDa, in good agreement with the molecular weight discrepancy between 'naked' and PTMed GCase [Tekoah et al., 2013].

3.2.2 NMR of post translationally modified GCase

NMR of large molecular weight complexes has recently become possible due to the pioneering work in the Professor Kay (Toronto, Canada) group [Tugarinov et al., 2006, Goto et al., 1999]. Their strategy is to introduce amino acid precursors containing $^{13}CH_3$ methyl groups on a select few amino acids (typically MILV). The degeneracy of methyl proton peaks provide a strong signal for

NMR, whilst the sparsity of these residues ensure the spectrum is not overly crowded.

A common technique in X-ray crystallography is to introduce CH_3 groups onto lysine and arginine side chains on the protein of interest, altering crystal contacts and potentially enabling crystallization of previously recalcitrant proteins [Walter et al., 2006, Rypniewski et al., 1993, Rayment, 1997] (see chapter 7). Combining these two separate ideas, it would be possible to introduce NMR active $^{13}CH_3$ labels *post hoc* to GCase. The favorable relaxation properties and sparse labeling of $^{13}CH_3$ lysines and $^{13}CH_3$ arginines should produce NMR observable peaks.

Methylation of GCase under standard conditions (see Methods and [Rayment, 1997]) revealed extensive modification as determined by native electrospray mass spectrometry and size exclusion chromotography. Both techniques confirmed an increase in molecular weight of GCase after methylation, as expected (figure 3.3). The increased breadth of the nESI MS peaks may reflect heterogeneous methylation or by-product adducts [Kurinov et al., 2000, Lund-Katz et al., 1988, Hattori et al., 2013]. The dimer population appears unperturbed by methylation (figure 3.3).

Analysis of methylated GCase by NMR (Varian 600MHz, 30 °C) reveals peaks centered at 46 and 3 ppm in ^{13}C and ^{1}H respectively (figure 3.4A). There are two sets of overlapping peaks visible, one at 47ppm and another at 45ppm in carbon. The identity of these peaks was deduced by ^{13}C edited ^{1}H pulsed field gradient (PFG) diffusion NMR and in-house analysis software coded in Python ([Drake and F.L., 2001]). The analysis software fits a linear regression to the exponential decay of signal to yield y-intercepts and decay rates. These are then used to model diffusion. Such diffusion rates are determined in cm^2

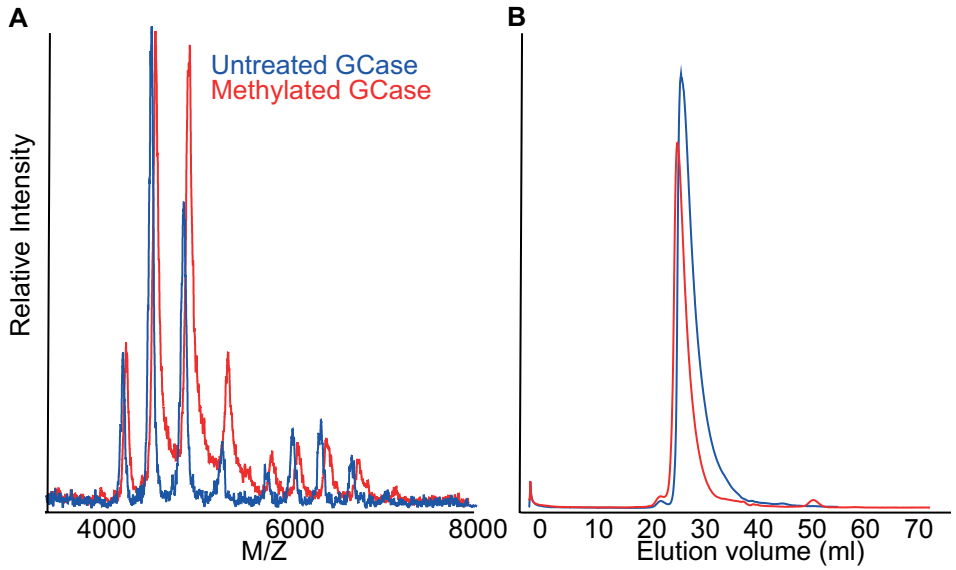


Figure 3.3: *Results of methylation reaction. A) Native electrospray MS of methylated and untreated GCase (10 μM GCase, pH7, 21°C), note the broadening and rightward shift of methylated GCase, consistent with an increase in molecular weight. B) Size exclusion chromatography profiles of treated and untreated GCase (100 μM GCase, pH7, 21°C), earlier elution volume of methylated GCase is consistent with a greater molecular size.*

s^{-1}; larger coefficients correspond to smaller species, and likewise, a more rapid signal decay corresponds to a smaller species.

The rapidly diffusing species at 47ppm carbon appears to be natural abundance ^{13}C citrate (a buffer component). Carbon-13 has a 1% natural abundance and the buffer contains 50mM citrate; it is therefore unsurprising that this species is observable by NMR. The peaks centered at 45ppm diffuse more slowly however, with a calculated diffusion coefficient of $7 \pm 1.11 \times 10^{-7}$ cm^2 s^{-1}, compared with in house measurement of $6 \pm 1 \times 10^{-7}$ cm^2 s^{-1} for albumin, a 66kDa protein. Therefore peaks at 45ppm have a diffusion coefficient consistent with GCase, a 63kDa protein (figure 3.4C).

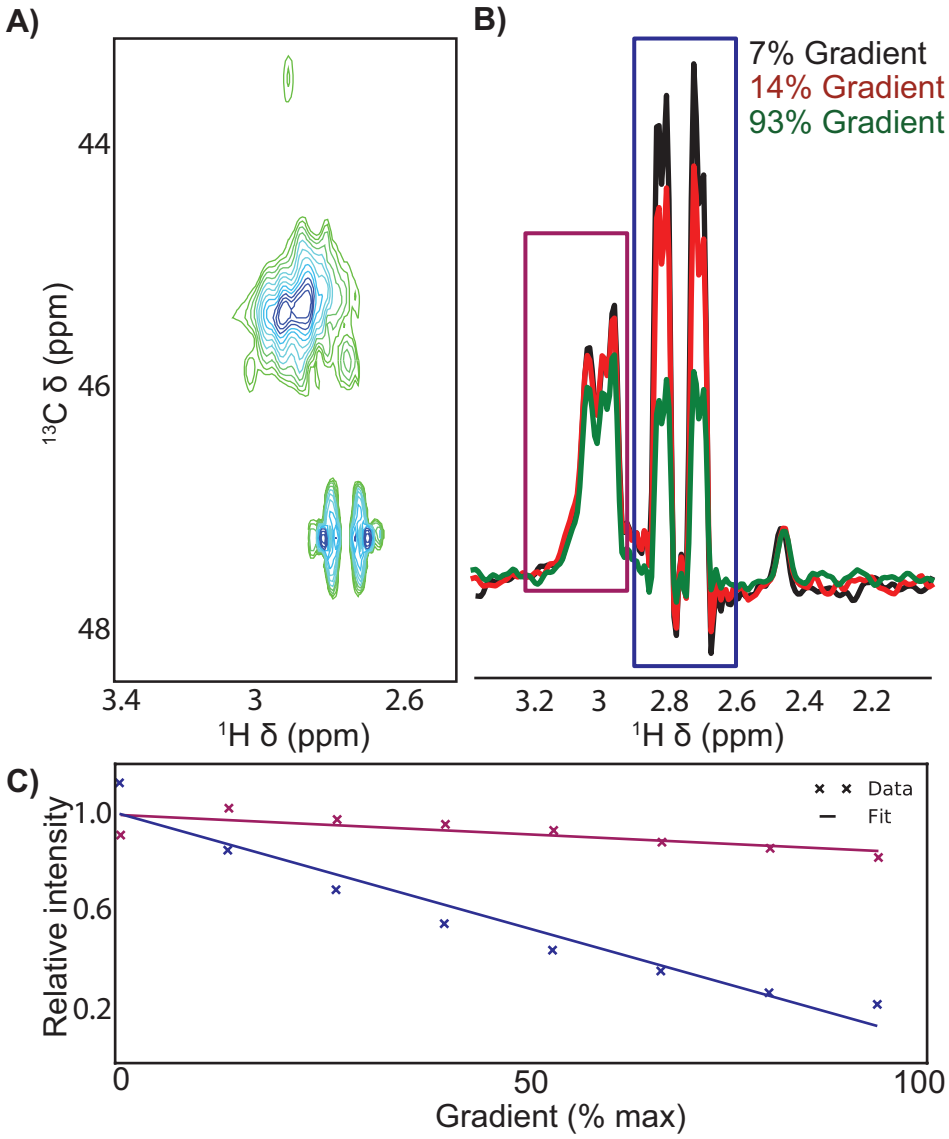


Figure 3.4: *Results of GCase methylation. A) 2D HMQC ^{13}C ^{1}H spectra (100 µM GCase, pH7, 30°C) of methylated GCase showing two clusters of peaks. B) 1D ^{1}H PFG diffusion spectra, varying the strength of gradient applied allows a diffusion coefficient to be determined C) I/I0 plots for peaks in (B), the two sets of peaks have drastically different diffusion behaviors*

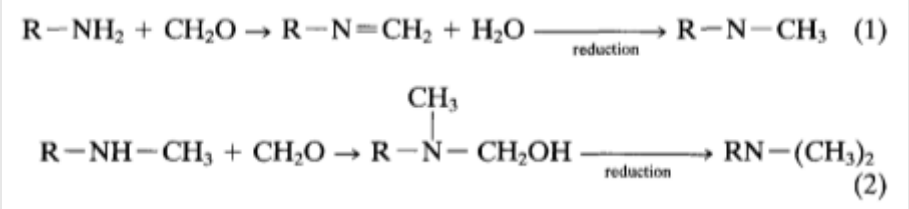


Figure 3.5: *Simplified schematic of reductive NH_2 methylation.*

3.2.2.1 Controlling methylation reaction conditions yields observable differences in NMR spectra

Maximal utility from NMR is dependent upon peak assignment. That is, knowing which resonances in a spectrum derive from which residues in a protein. Assignment of observable peaks was here achieved by probing the methylation reaction (figure 3.5). Alteration of pH and formaldehyde by 3 orders of magnitude effectively perturbed the methylation reaction, reducing efficiency at low substrate concentration and low pH (figure 3.6). These findings are consistent with collision theory and the role of pKa in methylation, respectively. It is interesting to note that even at a minimal concentration of 0.005M formaldehyde, reagent is still in excess over protein. (GCase has 44 lysines and arginines, each with 2 amenable NH groups, plus an NH N-terminus; therefore GCase has 0.0009M equivalents of amenable NH groups, compared to 0.005M formaldehyde.)

To investigate the effects of methylation conditions on NMR spectra large quantities of methylated GCase were prepared at pH5 and pH7. GCase was methylated for 20 hours with 1000 fold molar excess of formaldehyde. Structural changes in treated GCase were assayed by SDS-PAGE and native MS (nESI) (figure 3.7). There was no significant alteration in the SDS-PAGE presentation of GCase, indicating the absence of degradation and/or cross linking, likewise nESI did not reveal any degradation products or alterations

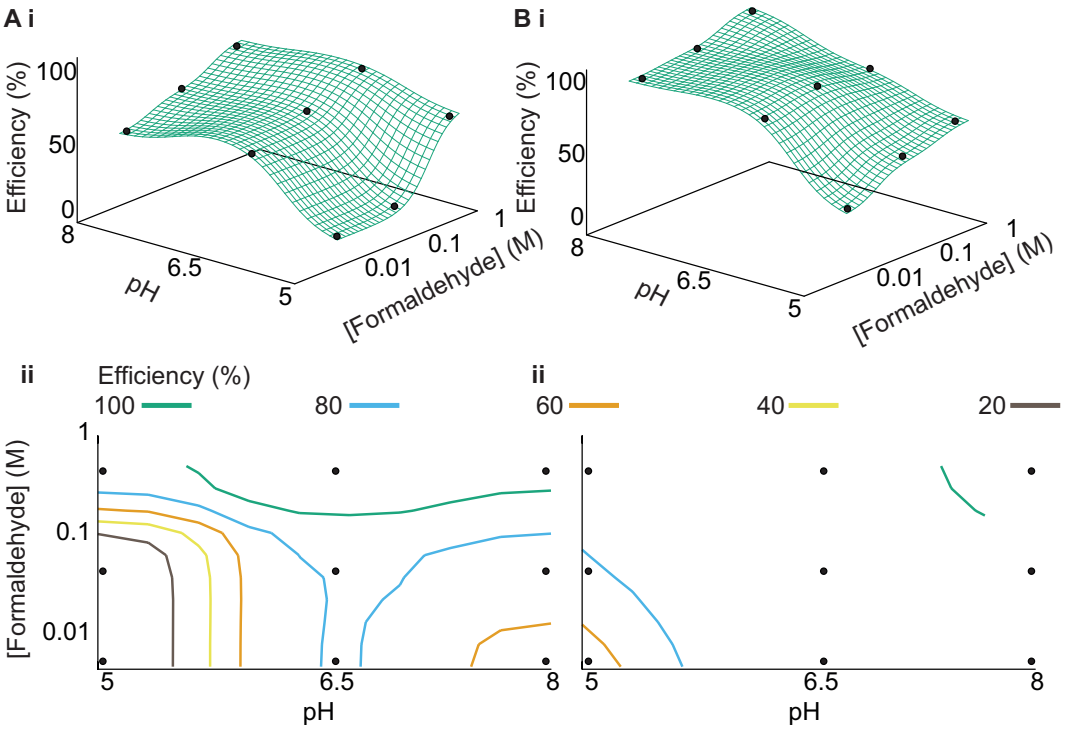


Figure 3.6: *GCase methylation efficiency is dependent on pH and formaldehyde concentration. A) GCase methylation at t=2 hours, presented in 3D (i) and 2D (ii) contours plots, as quantified by mass spectrometry (10 µM GCase, pH7, 21°C). B) GCase methylation at t=16 hours in 3D (i) and 2D (ii), data points (filled circles; 0.5, 0.05, 0.005 M formaldehyde, and pH5, pH6.5, pH8) are interpolated*

in dimersisation propensity (figure 3.7). Consistent with previous observations (figure 3.6) there was a lower shift in mass for pH5 treated GCase as compared to pH7 treated GCase (figure 3.7E), indicative of fewer modified side chains. Finally, the activity of treated GCase was assayed with synthetic substrate resorufin--D-glucopyranoside and fluorescence spectroscopy. Whilst no significant perturbation is observed after pH7 treatment, there is a slight (<5%) reduction for pH5 treated GCase (figure 3.9C).

The overall fold of GCase was also examined by ion mobility mass spectrometry (IM MS). Ion mobility mass spectrometry measures the time taken for an ionized gaseous protein to pass through a collision chamber. The time taken is proportional to the collisional cross section of the protein, thus providing information on the tertiary and quaternary structure of a protein. IM MS confirms that there is no significant difference in drift times after methylation (figure 3.8) - suggesting GCase retains it's native fold. The drift time of the second, higher m/z peak set is significantly longer than the first, confirming the presence of monomeric and dimeric GCase.

Inspection of ^{13}C ^{1}H 2D HMQC NMR spectra of GCase treated at pH5 and pH7 reveals distinct methylation patterns (figure 3.9 A, B). The most prominent difference in spectra is the presence of resonances at 35ppm at pH5. In line with the poor reaction efficiency, these extra resonances likely correspond to monomethylated lysine and arginine residues [Ashfield et al., 2000, Nott et al., 2015]. The central 45 ppm ^{13}C region shows an overall reduction in signal intensity (3.9 Aii), also consistent with poorer reaction efficiency. Most interestingly there is a novel, low intensity peak centered at ^{1}H ^{13}C 2.8/45.7 ppm (figure 3.9B).

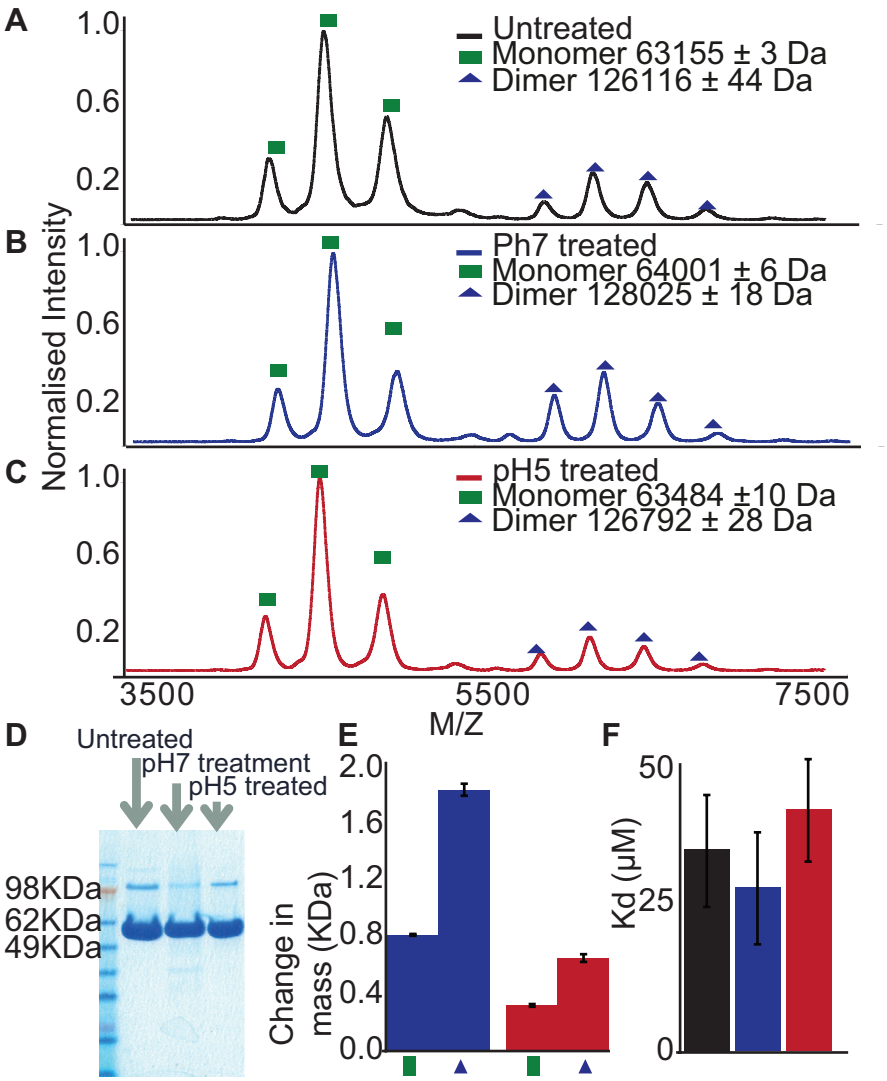


Figure 3.7: *GCase methylation at pH5 and pH7. A) nESI MS of untreated GCase reveals monomer and dimer populations (10 μM GCase, pH7, 21°C). B) pH7 treated GCase retains the ability to form dimers, C) likewise, pH5 treated GCase retains dimerization ability D) SDS PAGE analysis of treated GCase reveals no new bands, suggesting that no significant cross-linking or degradation has occurred. E) quantification of change in GCase mass under different treatment conditions, pH7 treatment results in more extensive methylation of GCase compared to pH5. Mass shifts correspond to 51 and 24 added methyls at pH 7 and 5, respectively F) Calculated Kd from MS (colors as A-C)*

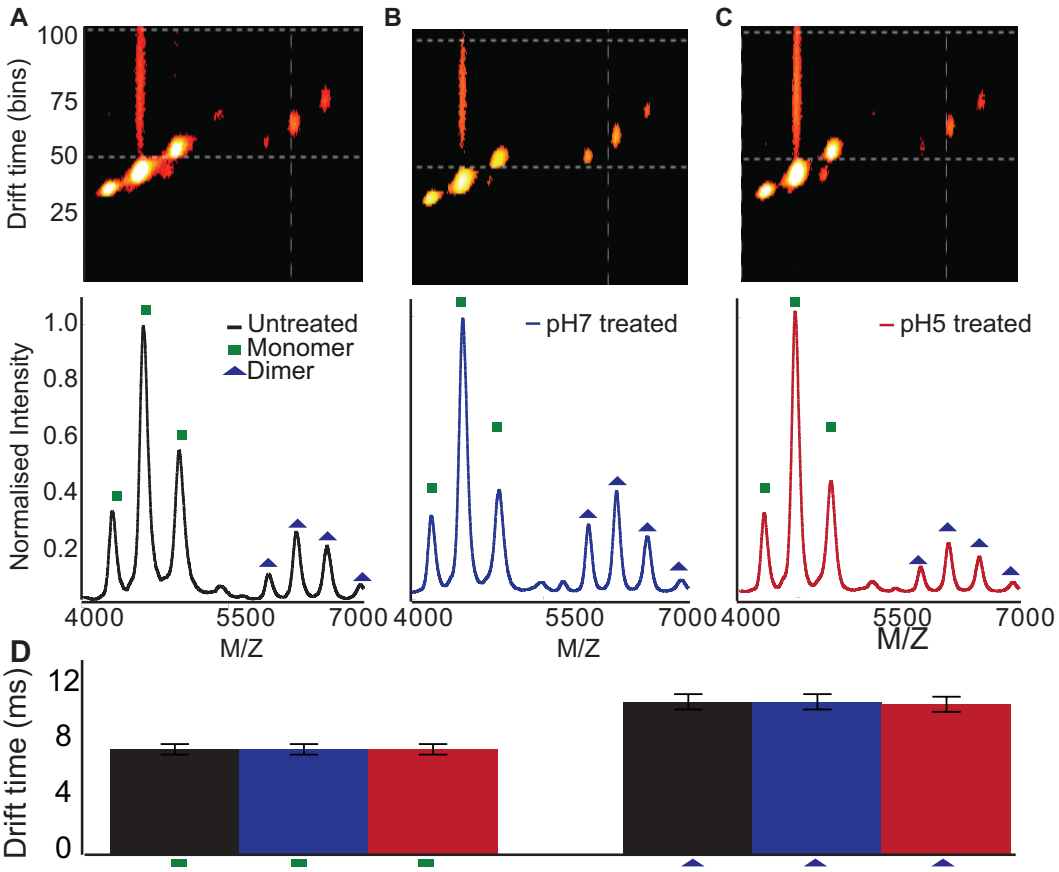


Figure 3.8: *Methylation of GCase does not effect GCase fold. Drift time measurements for GCase untreated (A), pH7 treated (B) and pH5 treated (C) (10 μM GCase, pH7, 21°C) reveal no significant difference for either monomer or dimer drift times (D).*

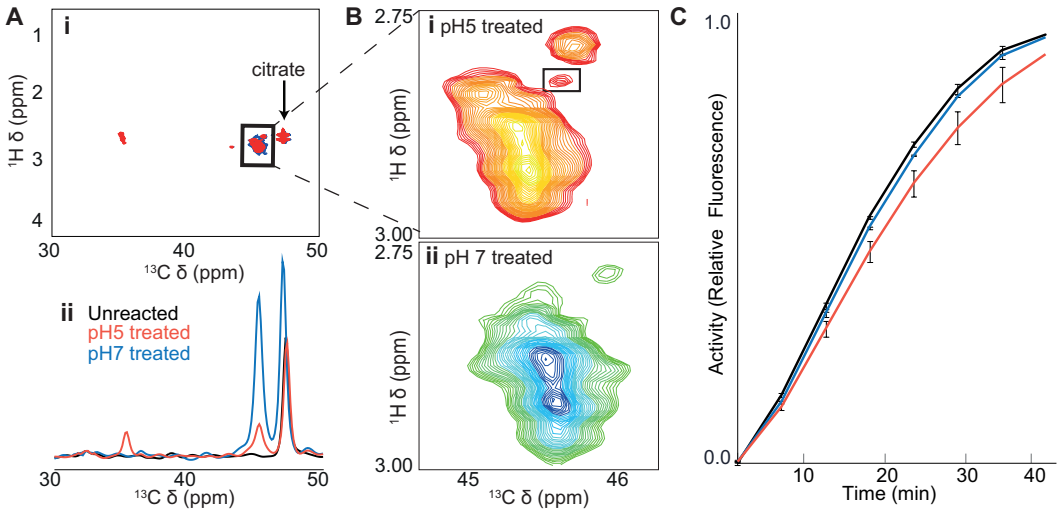


Figure 3.9: *GCase methylation with ^{13}C formaldehyde. All spectra shown were acquired at 100 μM GCase, pH5, 30°C, on a 600MHz spectrometer. A i) 2D ^{13}C ^{1}H HMQC of methylated GCase and ii) its projection on the ^{13}C axis. Incubating GCase with formaldehyde at pH 7 (blue) or pH 5 (red) results in distinguishable methyl spectra. B) The majority of labeled resonances largely overlap in the indicated region. Nevertheless, one specific resonance is observed in the pH 5 treated sample that is not observed after incubation at pH 7. C) GCase activity. The appearance of a fluorescent product was followed with time for untreated GCase and GCase treated at pH 5 (red) or pH 7 (blue).*

3.2.2.2 Assignment of methylated side chains

The gentleness of the methylation treatment outlined above suggested this reaction as a viable strategy to introduce NMR active isotopes. I rationalized that alterations in efficiency and NMR spectra may be exploited to assign peaks, if the identity of modified side chains could be confirmed. I therefore subjected methylated GCase to liquid chromotography tandem mass spectrometry (LCMSMS). Samples treated at either pH5 or pH7 were deglycosylated, digested and sprayed to identify lysines and arginines chemically modified with $^{13}CH_3$ methyl groups. LCMSMS analysis revealed more heterogeneity in methylation at pH5 than at pH7, with a mixture of mono and di methyl groups added onto GCase side chains. This is consistent with the NMR spectra which has a new cluster of peaks at 35ppm (expected for mono-methylated

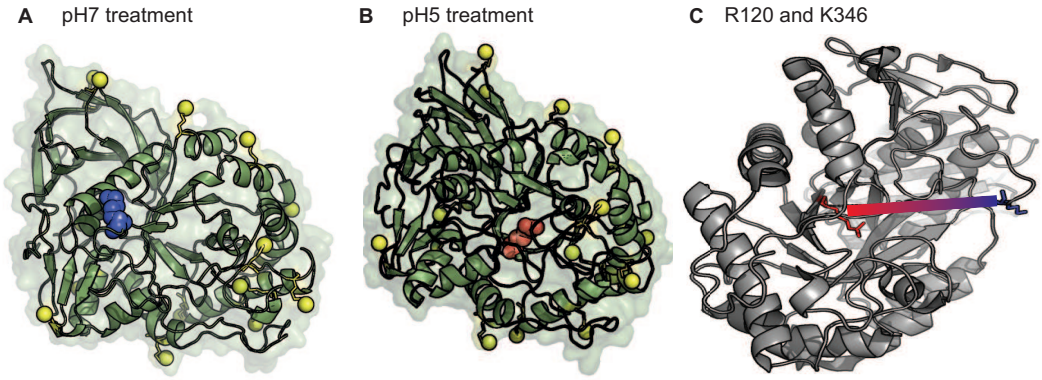


Figure 3.10: *GCase methylation at pH5 and pH7. LCMSMS analysis of treatments had a 87 and 93% coverage at pH7 and pH5 respectively. Sites methylated in both samples are indicated in yellow, sites methylated only at pH7 (A - K346) and pH5 (B - R120) are indicated in blue and red, respectively. C) Whilst both residues are located near to the active site β-barrel they are separated by 23 Å(PDB 1OGS).*

lysines). Additionally, two peaks were differentially methylated; pH5 treated GCase exclusively contained a label on R120, whilst pH7 treated GCase exclusively contained an additional methyl on K346.

The changes in NMR spectra may therefore be linked to differences in methylation sites. The changes between spectra must be consistent with the changes observed by LCMSMS. Alterations in methylation efficiency, leading to mono or di methylated side chains, are associated with a significant (10 ppm in carbon) chemical shift changes (figure 3.9Bi, inner box). I therefore posit that the additional peak visible in the 2D HMQC NMR experiment of pH 5 treated GCase is R120 (or, possibly an unidentified side chain outside of LCMSMS coverage). Similarly, I posit that the K346 methylation must lie in the central, heavily overlapped, region in the pH 7 NMR spectra. However, more experimental data would be required for certainty; clarity could be gained by methylating GCase with ^{12}C at pH7, and then ^{13}C at pH5, I would then expect only a single peak corresponding to the putative R120 peak. Both

R120 and K346 are proximal to the active site β barrel but separated by >23 Å(figure 3.10C).

3.2.3 Methylated GCase reveals slow dynamics consistent with dimerization

The introduction of $^{13}CH_3$ onto GCase provides a means to study GCase dynamics by NMR spectroscopy (figure 3.9). Of the many experiments available for studying dynamics, the CPMG is noted for being able to detect sparsely populated conformational states. On inspection, such sparsely populated conformers have been linked to enzyme catalysis, protein folding and other biological phenomena [Bouvignies et al., 2011, Vallurupalli et al., 2008, Neudecker et al., 2012, Korzhnev et al., 2010].

Briefly, the CPMG experiment consists of repeated 180° pulses (figure 2.2). Each pulse contains an echo period before and after the 180° pulse and the total number of pulses is varied over a fixed time period (the fixed time CPMG experiment). In the absence of exchange, magnetization remains in phase. However, in the presence of exchange magnetization is dephased, leading to peak broadening and a loss of signal intensity. This loss of intensity can be related to the frequency of CPMG pulses to derive a number of exchange parameters (specifically: $\delta\omega$, R_2 rates, k_{ex} and conformer population, see chapter 2 section 2.2.1.1).

The Baldwin formula is able to robustly deal with CPMG data provided that pulses are on resonance, magnetization is in phase and the system is in 2 site exchange (section 2.2.1.1). Figure 3.4 shows a typical GCase spectrum; consisting of peaks occupying a narrow ppm range in both carbon and proton,

it is therefore unlikely for pulses to be off-resonance. The need for in-phase magnetization is dealt with in the CPMG experiment itself, which mixes magnetization according to scalar coupling, and finally, inspection of the observed R_2 rates suggests that the system is indeed in exchange (figure 3.15B), thus satisfying the requirements for the Baldwin equation.

The poor peak dispersion for GCase means that fitting and numerically analyzing individual peaks (as is traditionally done) is not possible due to the high likelihood of peak fitting errors. To overcome this, a novel strategy was employed to fit 1D CPMG data (see methods). It was reasoned that partitioning the peaks into an arbitrary number of 'slices' (from 1, representing an unpartitioned peak, to 16, representing a fully partitioned peak), might represent a tractable alternative capable of extracting useful information, despite the overlapped spectrum.

In this approach, each partition would serve as a peak in traditional CPMG analysis. Each partition receives its own R_2 and $\delta\omega$, which can vary freely, as well as a P_B and k_{ex} enforced on all partitions. Data were fitted to the Baldwin equation using a Levenberg - Marquardt constrained least squares algorithm. A variety of physically reasonable initial values were used to ensure that global minimization was achieved, and errors were calculated using a bootstrapping method; sampling the data with replacement and recalculating fits.

The application of this partitioning strategy is illustrated in figure 3.11, for a peak treated as four or 16 partitions. There are substantial differences in error for each peak segment (figure 3.11Aii,Bii), suggesting partitioning as a sensible strategy. Whilst the fits are not noticeably improved by increased partitioning (figure 3.11Aiii, Biii), $\delta\omega$, R^2, P_B and k_{ex} reveal substantially narrower ranges, reflecting increased confidence in the fits and data (figure

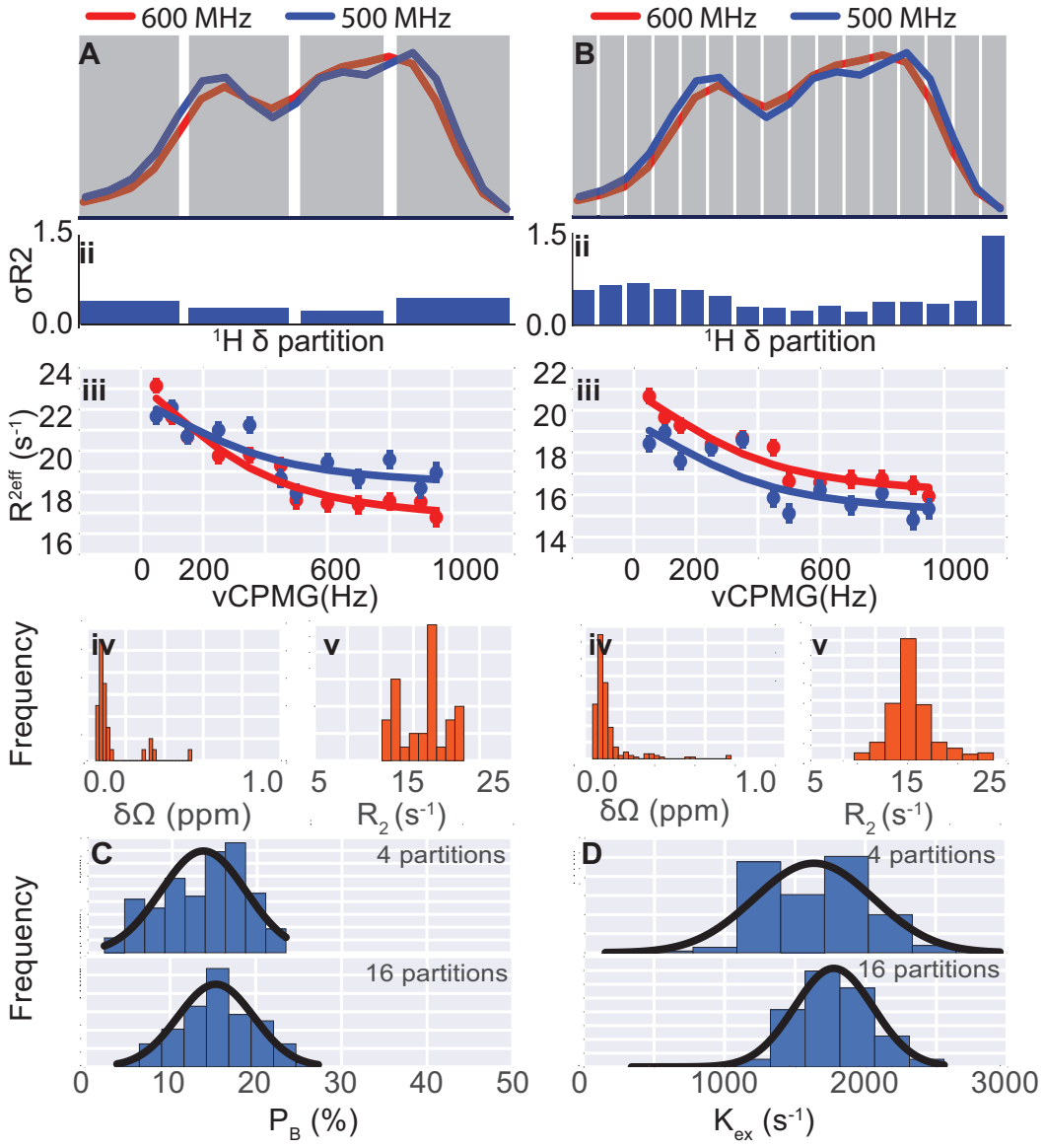


Figure 3.11: *CPMG analysis of pH 7 treated GCase (100 µM GCase, pH7, 30°C). To approximate peak analysis as in canonical CPMG, the spectrum was split into a number of equally sized partitions. Here, 4 partitions (A) and 16 partitions (B) are shown for illustration. Errors (ii) are calculated from 3 independent measurements of the same sample. Fits are shown in iii (lines represent fits and circles represent data points) for representative partitions. Similarly, iv and v display histograms for δω (ppm) and R^2 for representative partitions. C) histograms for global fits of P_B treated as 4 or 16 partitions, D) histograms for global fits of k_{ex}. Histograms were calculated by bootstrapping partitions and recalculating CPMG fits.*

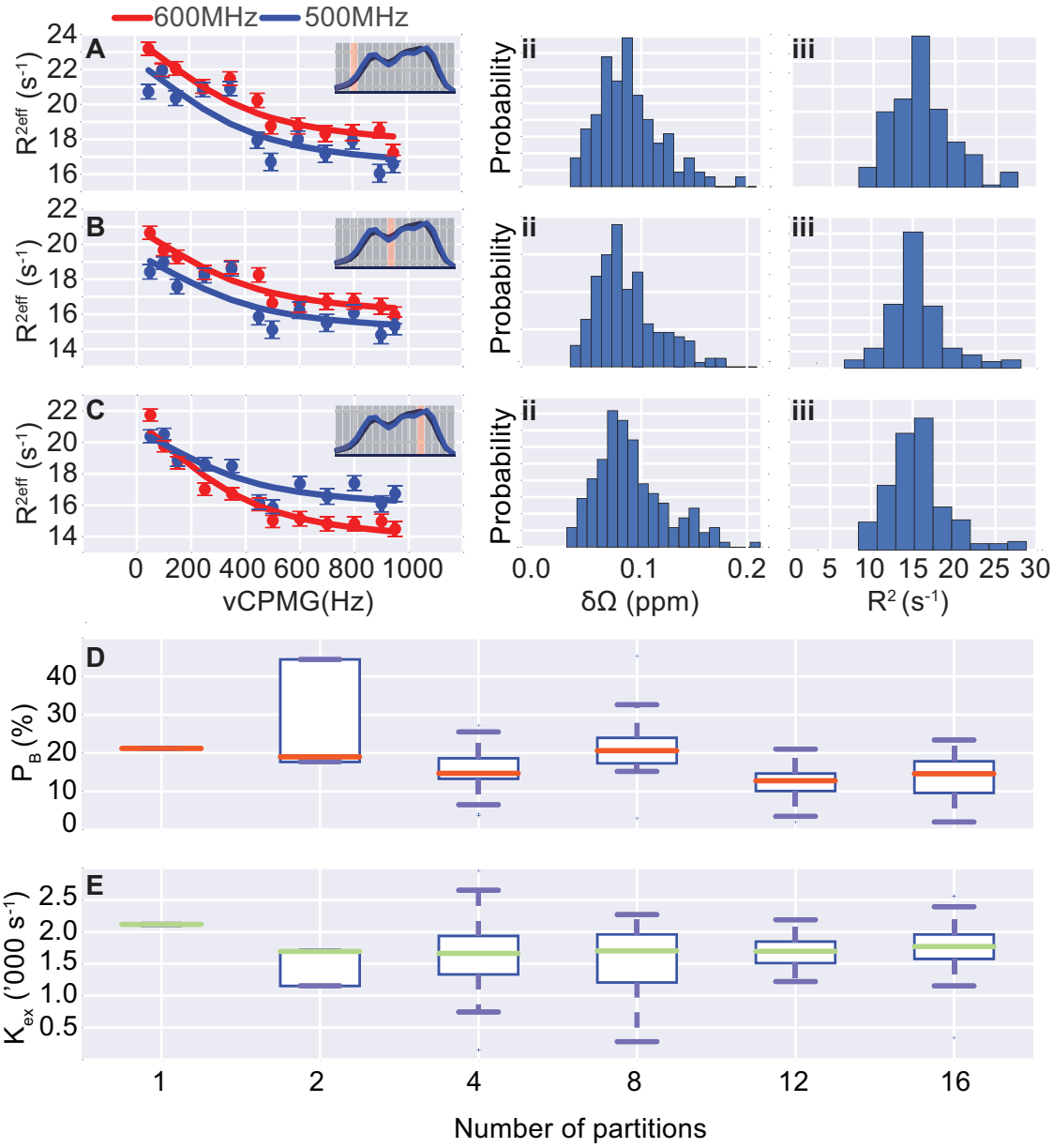


Figure 3.12: *CPMG derived distributions of P_B and k_{ex} when the peak is partitioned into 16 regions. The 3^{rd} (A) 8^{th} (B) and 12^{th} (C) partitions are shown for illustration, alongside R^{2eff} of each partition (i), delta omegas (ii) and R^2 rates (iii) for each partition. Finally P_B (D) and k_{ex} (E) are shown for when the peak is partitioned into a total of 1, 2, 4, 8, 12 or 16 peaks*

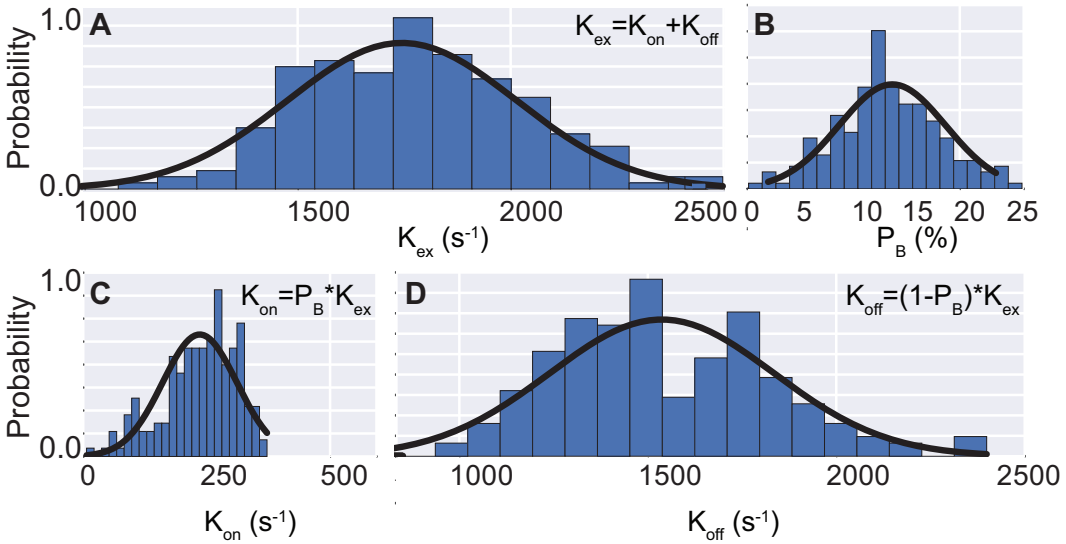


Figure 3.13: *Breakdown of k_{ex} rates for GCase dimerization (pH7), A) k_{ex} distributions for GCase CPMG effects, B) Population distributions for GCase, C) K_{on} rates distributions, D) K_{off} rate distributions. Distributions calculated from bootstrapped data*

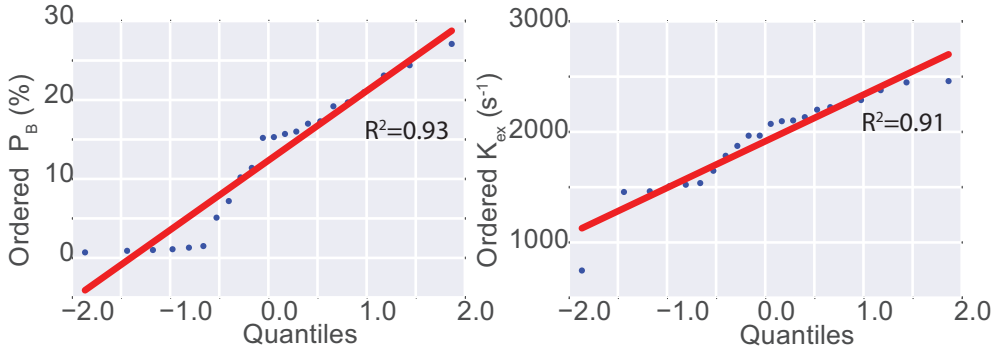


Figure 3.14: *QQ-plot comparing ideal normal distribution (red) and data (blue points). Both P_B (left) and k_{ex} (right) reveal normal distributions*

3.11Aiv,v, Biv,v, C, D).

A more extensive presentation of the 16 partition method is presented in figure 3.12.As the number of partitions is increased the global fits for P_B and k_{ex} become more confident (figure 3.12D,E). The increasing stability of mean values with more partitions, stabilizing at 12 and 16 partitions, suggests this analysis to be robust for CPMG analysis.

Data were found to be normally distributed (figure 3.14). Consistently, P_B

and k_{ex} were determined with a P_B of 14.1±6.3 % and a k_{ex} of 1640±269 s^{-1} (figure 3.12).

The implementation of a novel 1D CPMG analysis suggests the presence of a minor populated species observable by CPMG. I sought to determine the source of these dynamics. Given that it was previously observed that GCase forms dimers at pH7 (figure 3.2 and 3.8), I investigated the possibility that the CPMG event is a dimerization event. To test this hypothesis, GCase was similarly assayed by CPMG at pH5 where dimer populations are minimal (figure 3.2). A lack of CPMG dispersions at pH5 supports the hypothesis that the CPMG event is dimerization (figure 3.15).

As a bimolecular process, dimerization should be sensitive to protein concentration, whereas unimolecular dynamics should not. To this end, a dilution series was carried out, diluting GCase at pH7 and recording CPMG data. The results from this concentration series are presented in figure 3.16. There is a clear concentration dependence, with the rapidly relaxing species observed at low CPMG frequency reducing with concentration. Notably the later species, where I observe exchange free relaxation rates, reveal similar relaxation rates across dilutions. Together, these data supports the hypothesis that the CPMG event is the dimerization event.

Examination of GCase crystal structures obtained at pH7 reveal the crystallographic repeating unit comprises of either tetrameric or dimeric GCase. These crystal units may suggest possible dimerization interfaces for GCase; dividing the tetramer yields two possible dimers (figure 3.17). Of the two possible dimers identified, dimer A (figure 3.17) has more extensive dimer contacts and a larger dimer interface. The theoretical cross sectional areas for dimers A and B are highly similar, at 5175 Å and 4970 Å respectively. Such a differ-

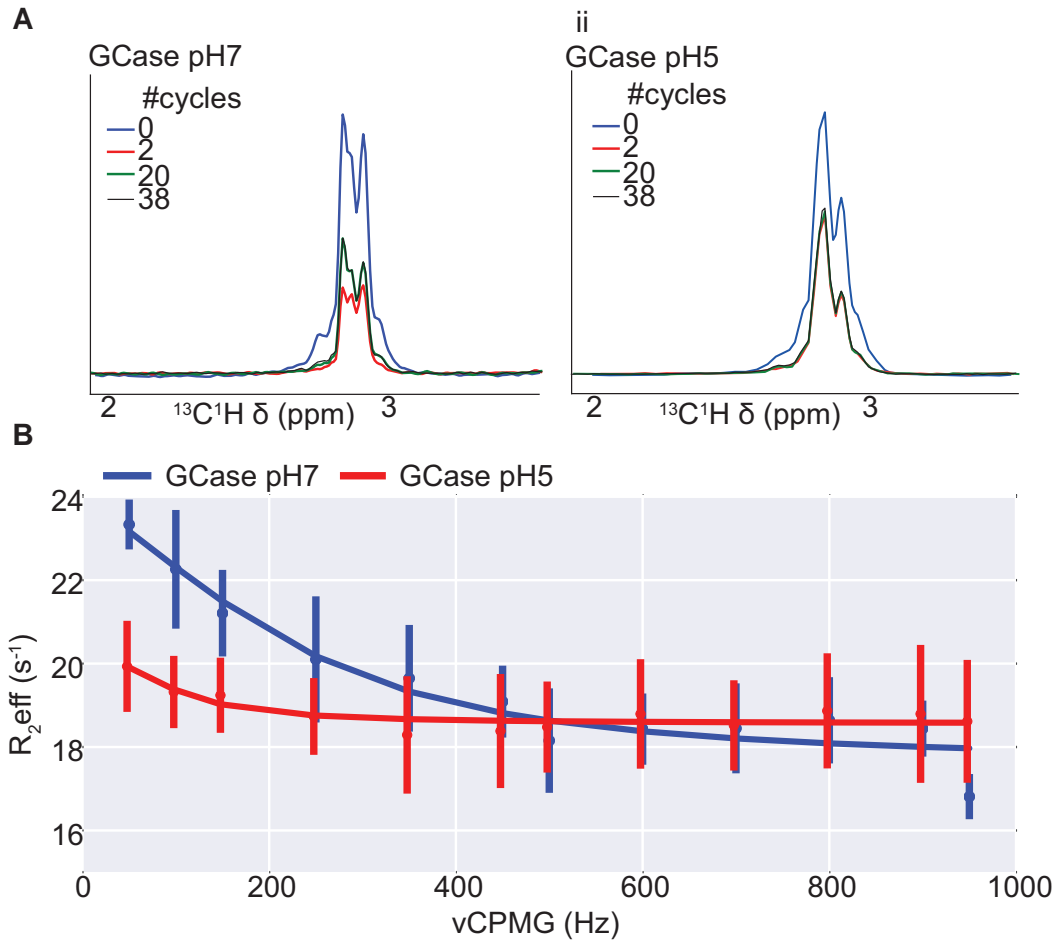


Figure 3.15: *CPMG study of GCase, A) 1D CPMG ^{13}C edited ^{1}H spectrum of GCase at selected CPMG frequency (expressed as number of cycles per fixed time) at pH7 (i) and pH5 (ii). B) CPMG analysis and fits using the Baldwin equation. Lines represent fits, circles and error bars represent experimental error (standard deviations of 3 measurements)*

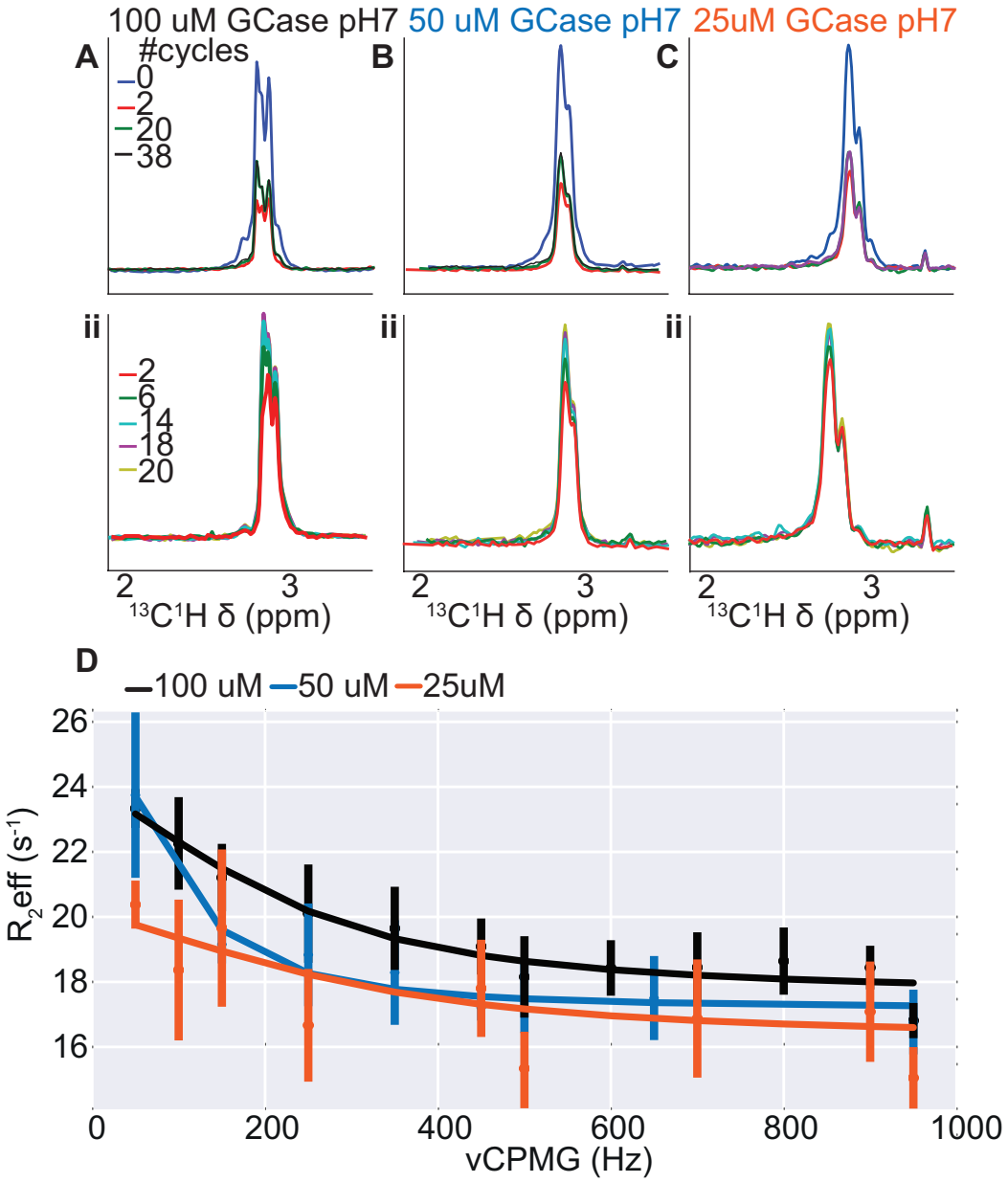


Figure 3.16: *Dilution series of GCase (pH7, 30°C). GCase was assayed by 1D ^{13}C edited ^{1}H CPMG. Sample CPMG frequencies at 100(A) 50 (B) and 25(C) μM are presented. D) CPMG analysis and fits using the Baldwin equation. Lines represent fits, circles and error bars represent experimental error (standard deviations of 3 measurements)*

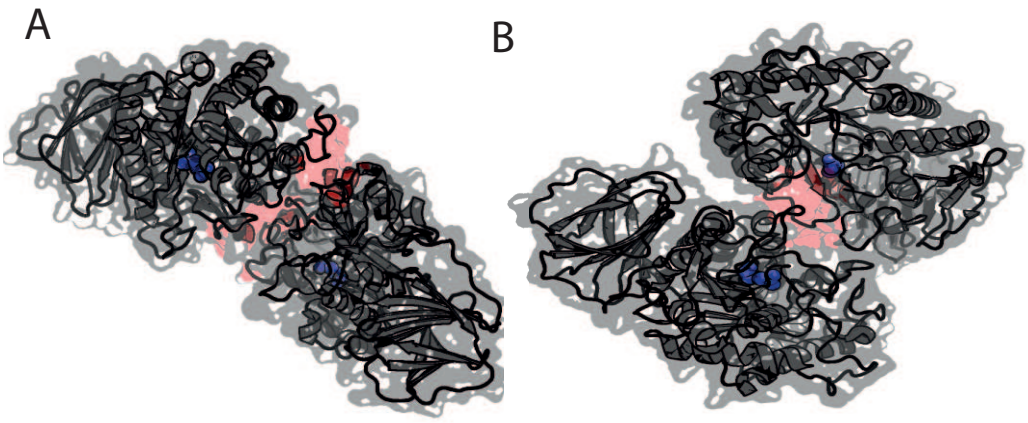


Figure 3.17: *Possible dimers observable in crystal structures. Blue spheres represent active site residue E340 and red areas represent intermolecular distances compatible with hydrogen bonding. Theoretical cross section areas of 5175Å(A) and 4970Å(B) were calculated. Structural data on carbohydrate moieties is lacking from crystallographic data, however, it is possible that glycans may sterically favor certain dimer conformers*

ence would not be detectable with IMS MS [Marklund et al., 2015]. However, hydrogen-deuterium exchange MS may be able to probe this interface further.

Taken together, the data provide significant insight into the dimerization of GCase; Baldwin equation fits yielded a P_B of 14.1±6.3 % and a k_{ex} of 1640±269 s^{-1}. Suggesting that the dimer represents a significant population at neutral (endoplasmic reticulum) pH, but not at lysosomal (acidic) pH.

3.2.4 Labeled GCase is a tool for drug discovery

The ability to introduce highly intense $^{13}CH_3$ groups onto human cell derived protein allows the characterization of PTM containing high molecular weight GCase. NMR can then be employed to study transient protein-protein (chapter 4) and protein drug interactions.

Industrial and biomedical interest in GCase derives from its role in Gaucher's disease (section 1.2.2). Briefly, Gaucher's disease is thought to derive from the degradation of GCase in the endoplasmic reticulum (ER) in a process termed

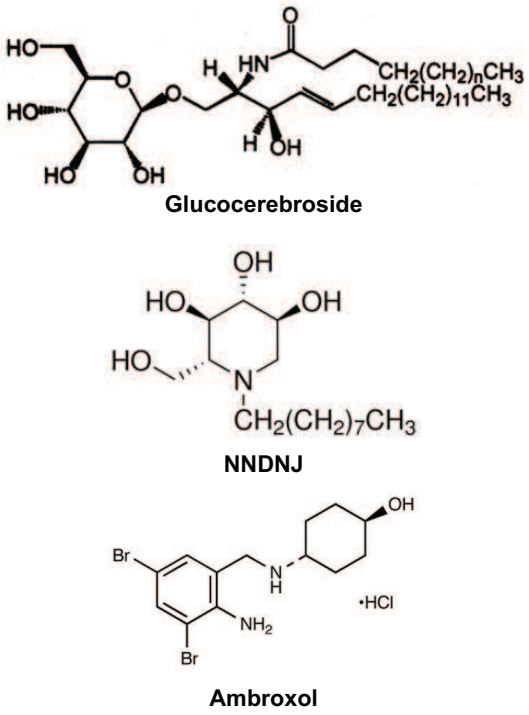


Figure 3.18: *GCase natural substrate (glucosylceramide) and potential pharmacological chaperones NNDNJ and ambroxol*

endoplasmic reticulum associated degradation (ERAD). Mutant GCase often displays enzymatic activity (albeit at a reduced level) when stabilized by acidic environments. However, glycosylation of GCase in the neutral pH ER acts as a 'molecular timer', if GCase does not reach its correct fold in sufficient time it is targeted for degradation [Bendikov-Bar and Horowitz, 2012, Ron and Horowitz, 2005, Zimmer et al., 1999, Bendikov-Bar et al., 2011]. The kinetics of folding are more troublesome for mutant GCase than wild type, leading to GCase degradation and a loss of activity in the lysosome. A possible therapeutic intervention, therefore, is to stabilize GCase sufficiently in the ER to allow it's trafficking to the lysosome, where residual activity prevents the build up of substrate and associated pathology [Parenti, 2009, Kolter and Wendeler, 2003, Yu et al., 2007].

Small molecules capable of stabilizing GCase are termed pharmacological

chaperones (PCs). Of the many possible PCs available, two were chosen for study by solution state NMR, namely, N-Nonyldeoxynojirimycin (NNDNJ) and ambroxol. On first inspection these PCs appear paradoxical, they are competitive GCase inhibitors that act to rescue GCase deficiency. However, the main role of PCs therapeutically is to stabilize GCase in the endoplasmic reticulum, mitigating ERAD processes and facilitating trafficking to the acidic lysosome where it may carry out its enzymatic function in an energetically favorable environment. Thus, PCs increase the amount of GCase delivered to the lysosome, limiting substrate accumulation [Yu et al., 2007, Parenti, 2009, Powers et al., 2009].

NNDNJ and its derivatives represent one of the better studied PCs, and a significant effort is underway to synthesize and assay derivatives for wide spectrum activity against the maximum number of GCase mutations [Yu et al., 2007]. Ambroxol, by contrast, is relatively under studied but posses the distinct advantage of being approved for humans as a pre-existing flu medication [Luan et al., 2013]. Furthermore, ambroxol reportedly binds GCase preferably at pH7 (ER pH) and not at pH5 (lysosomal pH) and therefore has the potential to selectively chaperone GCase only in the ER stages where it is vulnerable to degradation [Bendikov-Bar et al., 2011, Bendikov-Bar et al., 2013, Zimran et al., 2013, McNeill et al., 2014].

3.2.4.1 NNDNJ as a potential therapeutic

The effects of PCs on $^{13}CH_3$ containing GCase were investigated by NMR. Incubation of GCase treated at pH7, but not GCase treated at pH5, revealed significant peak intensity changes for NNDNJ GCase mixtures (figure 3.19A B). This may be due to the production of different methylation patterns at pH7

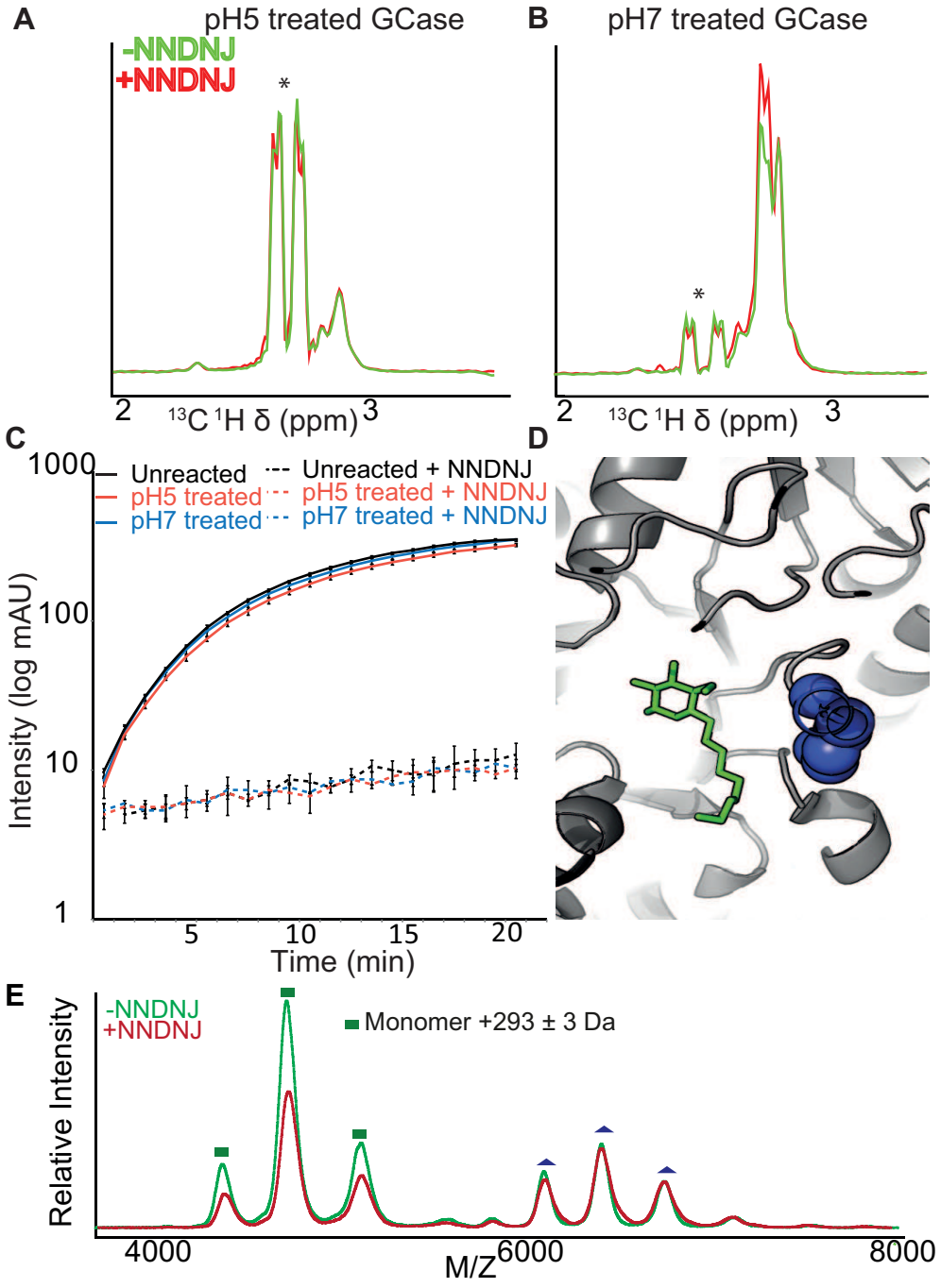


Figure 3.19: *GCase and NNDNJ A) pH5 treated GCase with and without 100 μM NNDNJ (100 μM GCase, pH7, 30°C), B) pH7 treated GCase with 100 μM NNDNJ reveals changes in peak intensity (100 μM GCase, pH7, 30°C), * denotes natural abundance citrate buffer. C) Activity assay of treated GCase plus NNDNJ revealed no significant differences in NNDNJ inhibition. D) GCase with NNDNJ (green) and K346 (blue) (PDB 2V3E) E) nESI MS of GCase with and without NNDNJ (10 μM GCase, pH7, 21°C).*

and pH5, one possibility being that K346 is specifically reporting on NNDNJ binding. To ensure intensity differences were not due to methyl groups perturbing activity/inhibition, an assay was carried out with synthetic substrate β-resofluroin and NNDNJ. This assay revealed GCase to be equally inhibited under both treatment regimes by NNDNJ (figure 3.19C). NNDNJ could also be seen bound to GCase by nESI MS (figure 3.19E).

Interestingly, examination of GCase intensity with and without NNDNJ (figure 3.19) reveals a significant *increase* in signal intensity with NNDNJ. An increase in signal intensity, whilst unusual, could derive from a number of sources, namely; a faster correlation time, altered chemical exchange or slower relaxation.

GCase displays pH dependent dimerization (figure 3.2 and 3.2.3) which could affect both correlation time and chemical exchange (figure 2.5 and section 2.2.1). A shift in the monomer-dimer population equilibrium could alter the average correlation time, whilst a change in k_{ex} could also affect signal if the exchange regime were shifted from intermediate into fast or slow exchange.

To probe these possibilities a CPMG experiment was carried out as above (section 3.2.3). The constant time CPMG experiment can be used to examine overlapped peaks in 1D ^{13}C edited ^{1}H spectra. As well as probing for slow (millisecond) exchange, CPMG pulse sequences can examine exchange free relaxation rates. If NNDNJ were stabilizing certain GCase conformers, or altering dimerization patterns, then one would expect to see differences in chemical exchange.

Data were fitted using the same methodology employed previously (section 3.2.3). Fitting data to the Baldwin equation reveals that the sparsely populated dimer conformer populates 21.5±5.4 % with a k_{ex} of 675±85 s^{-1} with

NNDNJ (figure 3.20), compared to values of 14.1±6.3 % (P_B) and 1640±269 s^{-1} (k_{ex}) for GCase alone. k_{ex} and P_B are significantly effected by NNDNJ addition. It is therefore possible that the increased signal intensity derives from a NNDNJ induced reduction in chemical exchange (section 2.2.1). Indeed, simulating a 1D spectrum undergoing chemical exchange, using the $\delta\omega$, R^2, P_B and k_{ex} derived from the fitted data, recapitulates the increased signal intensity observed upon NNDNJ addition (figure 3.21).

The increased signal intensity with pH7 treated GCase (but not with pH5 treated GCase) must suggest that there are extensive differences between methylation patterns on GCase treated at different pHs and that the alterations in signal intensity and CPMG profile are due to multiple labels on the GCase molecule (figure 3.22).

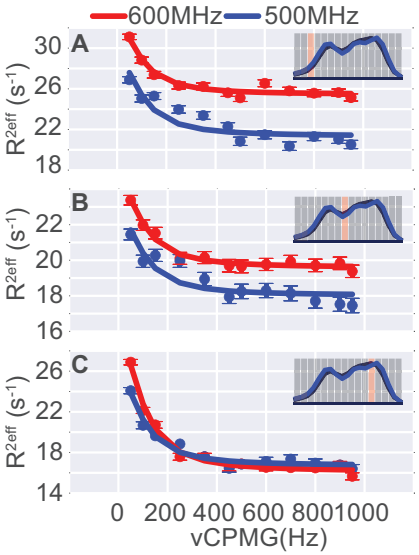


Figure 3.20: *Relaxation dispersion fits of GCase plus NNDNJ, the 3rd (A), 8th (B) and 12th (C) partitions are shown for illustration*

These insights suggest that the signal increase observed at pH7 with NNDNJ does derive from an alteration in dimerization rates. However, the presence of NNDNJ induced intensity increases at pH 5, may suggest that additional mechanisms are at play 3.22F), or that NNDNJ is capable of inducing dimerization even at pH 5.

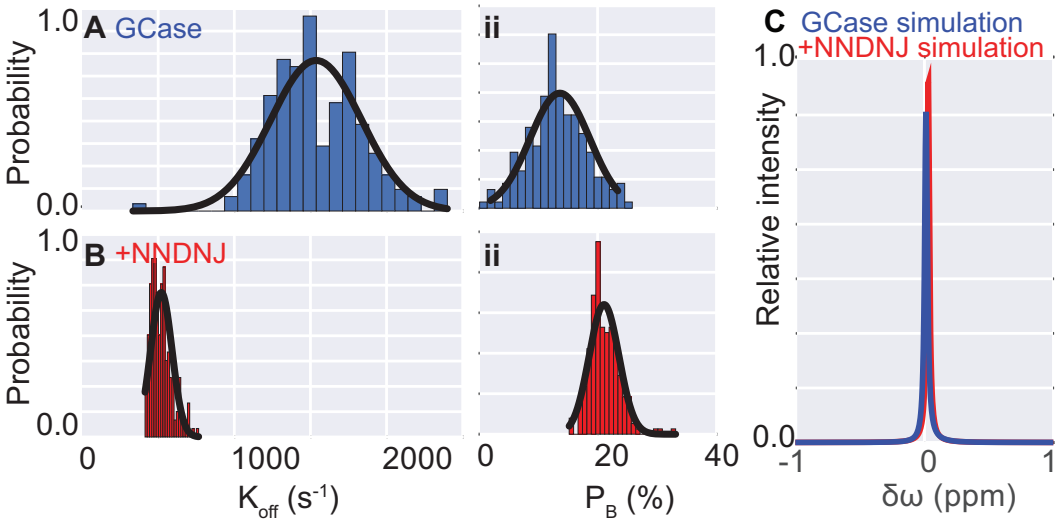


Figure 3.21: *Histograms of pH 7 treated GCase and NNDNJ CPMG analysis at pH 7 (100 μM GCase, 30°C). A) K_{off} rates and P_B (ii) of GCase, B) K_{off} rates and P_B (ii) for GCase+ 100 μM NNDNJ, K_{off} rates show significant reduction, suggesting the dimer is stabilized upon NNDNJ addition. C) Simulated spectra, using parameters derived from CPMG analysis, of GCase with and without NNDNJ*

3.2.4.2 Ambroxol binds GCase under endoplasmic reticulum conditions

The PC ambroxol was also examined by NMR. 1D HMQC spectra of ambroxol reveal an increase in signal intensity upon ambroxol addition similar to that observed with NNDNJ (figure 3.23). Interestingly this increase is attenuated at pH5, confirming previous observations that ambroxol interacts with GCase in a pH dependent manner [Siebert et al., 2014, Maegawa et al., 2009]. Furthermore, CPMG analysis suggests that whilst P_B is unaffected (16.8±4.8 %), k_{ex} is once again reduced from control values of 1640±269 s^{-1} to ambroxol induced values of 1160±370 s^{-1} (figure 3.24). This may suggest a common mechanism of action for both ambroxol and NNDNJ.

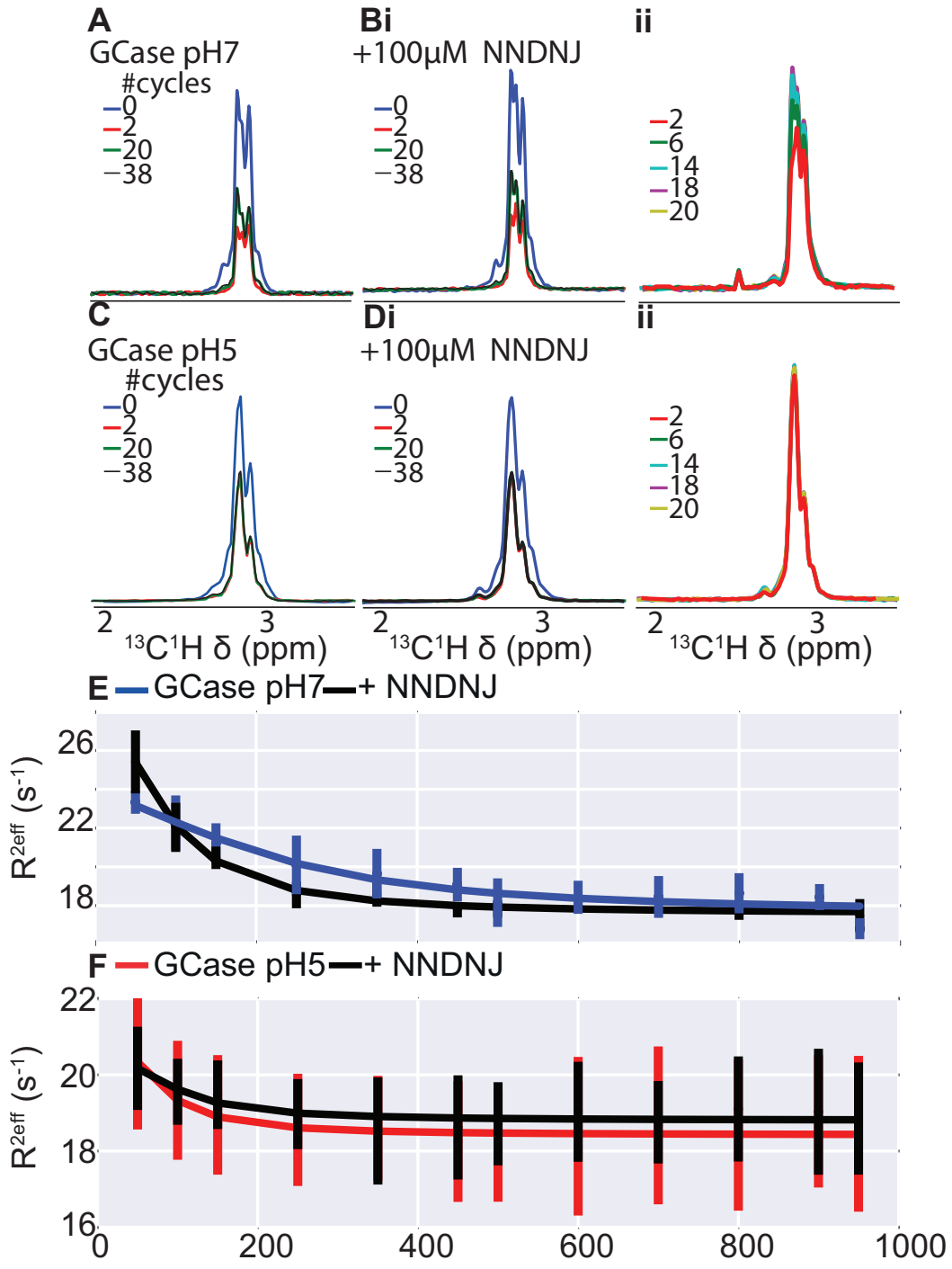


Figure 3.22: *^{13}C edited ^{1}H CPMG of GCase and NNDNJ at pH 5 and 7 (100 µM GCase, 30°C). A) CPMG of methyl labeled GCase alone, B) GCase + 100 µM NNDNJ with 0 to 38 (i) and 2 to 20 (ii) cycles shows substantial CPMG effects which reach a maximum at 20 cycles. C) GCase alone at pH5, D) GCase + 100 µM NNDNJ with 0 to 38 (i) and 2 to 20 (ii) cycles show no substantial CPMG effects.E) Quantification of CPMG effects at pH 7 and (F) pH 5. Error bars represent the standard deviation of three measurements.*

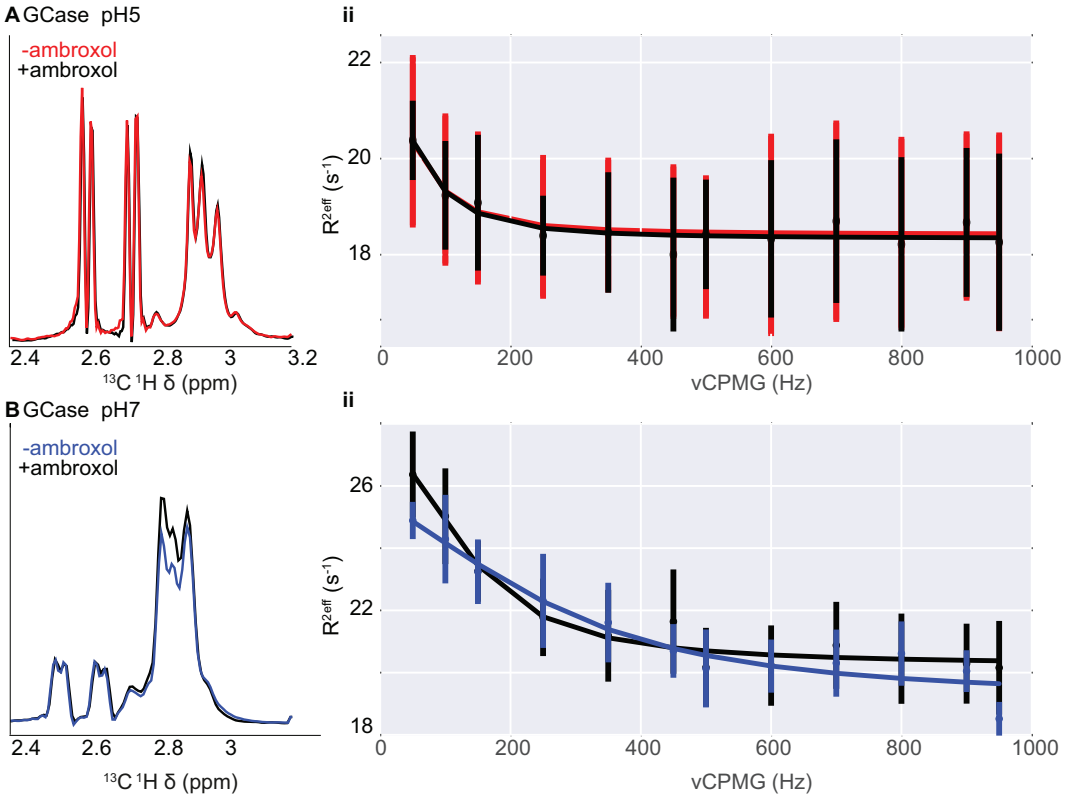


Figure 3.23: *^{13}C edited ^{1}H HMQC of pH 7 treated GCase with ambroxol at pH5 (A) and pH7 (B) and corresponding CPMG dispersion profiles (Aii, Bii) (100 μM GCase, 30°C)*

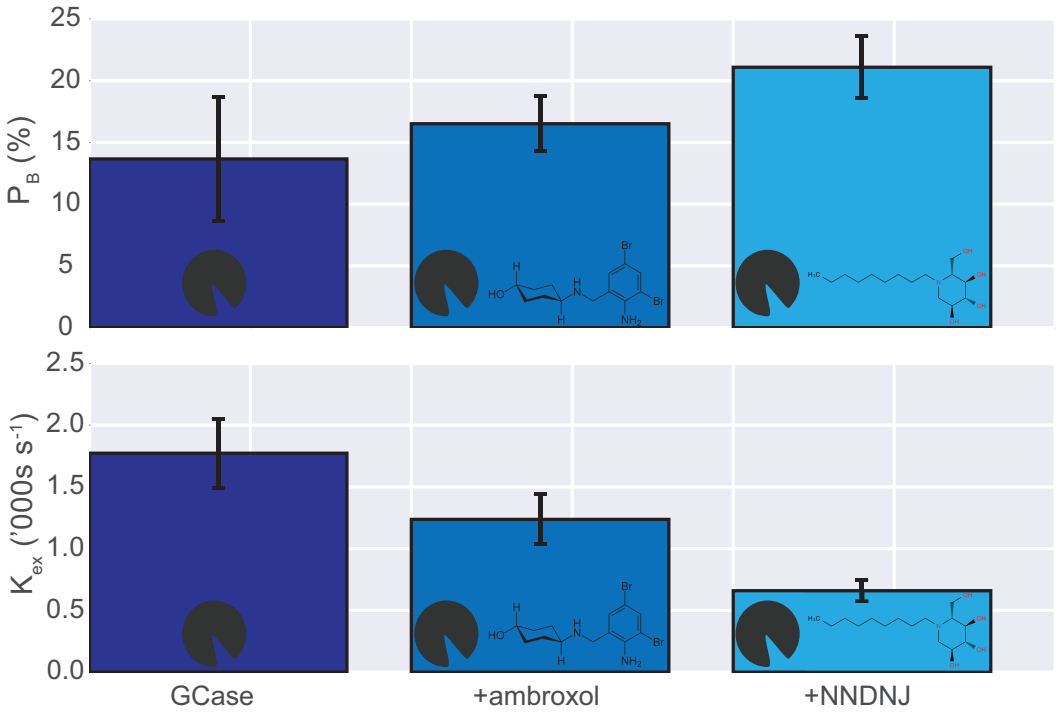


Figure 3.24: *Summary of pharmaceutical chaperone induced effects on GCase dynamics, the addition of ambroxol or NNDNJ at a 1:1 molar ratio reduces k_{ex} and stabilizes dimer*

3.2.4.3 GCase dynamics persist in cell-like environments

To study the GCase CPMG event in a more physiologically relevant environment, GCase was diluted to 50μM, reintroduced into 90 % (w/v) human cell lysate derived from human epidermoid carcinoma cells (A431 cells) and studied by NMR. Whilst this environment lacks the acidic pH of the lysosome, it recapitulates the macromolecular crowding and serves as a first step in achieving in-organelle NMR.

Lysate buffered GCase is presented in figure 3.25. Measurements over 22 hours resulted in some minor degradation of GCase (figure 3.25A), however lysate studies are feasible for shorter timescale experiments. Most interestingly, relaxation effects are conserved in lysate and phosphate buffered GCase. However, time dependent degradation effects clearly increase the noise and

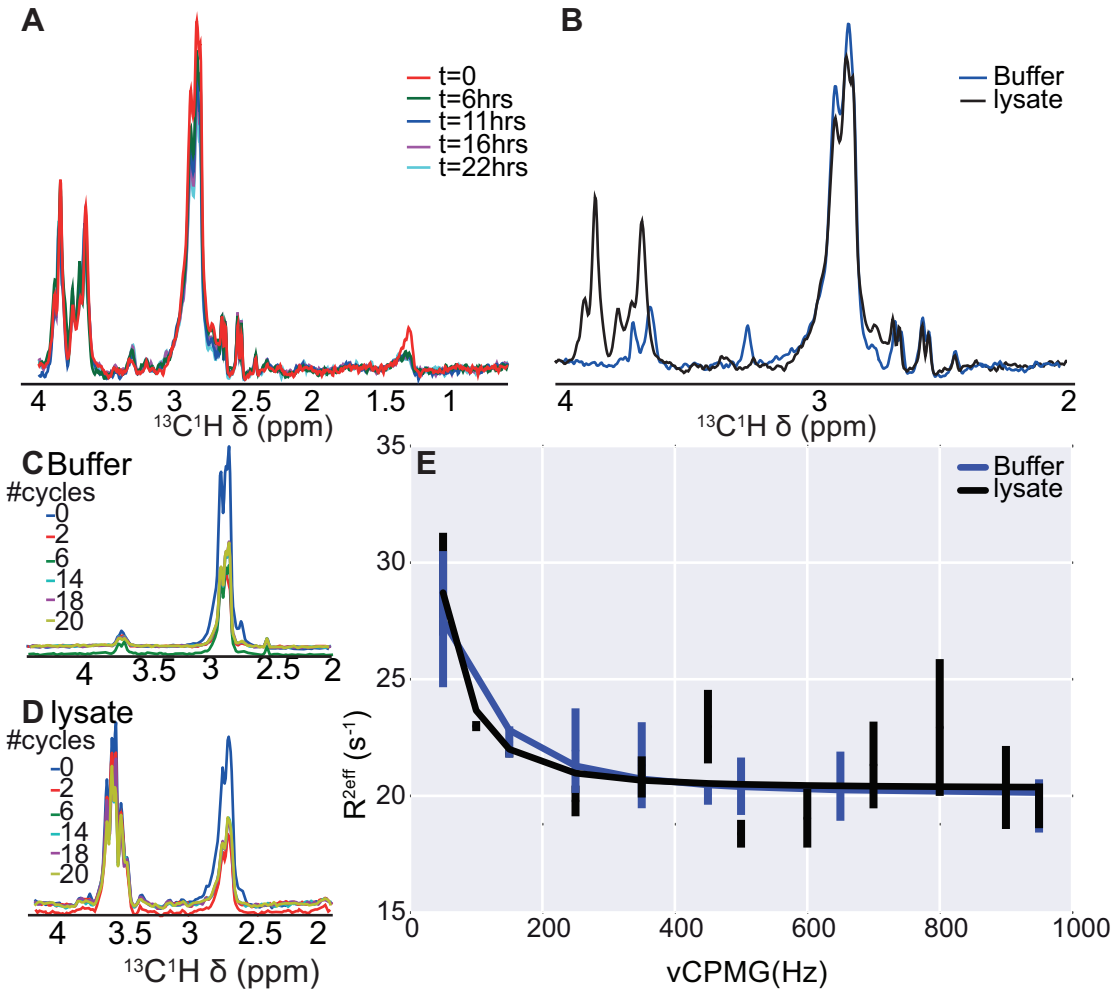


Figure 3.25: *Comparison of lysate (pH 7) and buffered (pH 7) GCase (50 μM GCase, 30°C). A) Stability profile of GCase assayed by serial ^{13}C edited ^{1}H 1D HMQC. B) Comparison of ^{13}C edited ^{1}H 1D HMQC from lysate and buffer GCase. Comparison of CPMG effects between lysate (C) and buffer (D) GCase reveal similar profiles. E) GCase in buffer and lysate reveal similar relaxation dispersion measurements.*

error of longer experiments, this is seen in the CPMG experiment (>4 hours) (figure 3.25E). Randomizing CPMG frequencies may prevent time dependent systematic error (as done here), but cannot reduce overall error. However, relaxation dispersion effects consistent with dimerization are clearly observed in human cell lysate (section 3.2.3).

3.3 Discussion

GCase is a lysosomal enzyme key to both Gaucher's and Parkinson's disease. A number of mutations destabilize GCase and predispose to one, or both of these diseases (figure 1.4), often through mechanisms involving endoplasmic reticulum clearance. Whilst the current therapy for Gaucher's disease is enzyme replacement therapy, the injection of GCase to restore activity and clear accumulated substrate, novel approaches are being trialled with so called pharmacological chaperones. Pharmacological chaperones have the potential to stabilize faulty GCase and rescue it from degradation. This may impact not only Gaucher's disease but also PD (see chapter 4), and simplify treatment regimes by providing a single oral capsule for treatment rather than a course of injections [Parenti, 2009, Kolter and Wendeler, 2003, Wright et al., 2015].

Here, I have presented a novel method for labeling and interrogating human cell derived GCase, complete with relevant PTMs, by solution state NMR spectroscopy. The addition of isotopically enriched methyl labels through reductive amine methylation (similar to the approach employed in chapter 7) permitted the study of structure, dynamics and molecular interactions of GCase at high resolution. Speculative peak assignment was achieved by complimenting NMR with LCMSMS (figure 3.6) to successfully assign two peaks, K346 and R120 (figure 3.10).

Methylation of GCase did not greatly perturb it's structure or activity (figures 3.9 and 3.7), consistent with other studies examining the effect of methylation on protein structure [Kurinov et al., 2000, Lund-Katz et al., 1988, Hattori et al., 2013, Walter et al., 2006, Larda et al., 2012, Rayment, 1997,

Gerken et al., 1982].

Investigation of methylated GCase revealed a novel pH dependent dimerization, not previously reported (figure 3.2). Mass spectrometry and CPMG NMR analysis was employed to characterize this process. Dimerization occurs on the millisecond timescale and the dimer population represents 14.1±6.3 % of total GCase. The pH sensitivity of dimer formation may be significant as GCase is trafficked from the neutral pH endoplasmic reticulum to the acidic lysosome where it carries out its function (figure 1.5). Interestingly the trafficking of GCase is itself pH dependent, due to the interaction with trafficking protein LIMP2 [Zhao et al., 2014]. Speculatively, it is possible that there is overlap in these two phenomena.

Further investigation revealed the effects of pharmacological chaperone binding using two model PCs, NNDNJ and ambroxol. Binding of either PC results in an increased signal intensity, an unexpected observation. To understand the source of this intensity change CPMG data were acquired over a range of conditions. Together, the data suggest that NNDNJ and ambroxol reduce chemical exchange rates of GCase, boosting signal intensity through altering chemical exchange regimes and stabilizing GCase dimers (figure 3.22). Whilst I can explain the increased intensity at pH7 through chemical exchange, the presence of increased intensity at pH5, where there is little dimer population and no observable CPMG event, suggests other mechanisms may be at play (figure 3.22).

The ability of pH7 treated, but not pH5 treated GCase to report on NNDNJ binding suggests that changes in methylation patterns may contribute to increased signal intensity at pH 5. Inspection of previously published GCase crystal structures with and without NNDNJ (PDB 1OGS and 2V3E respec-

tively) reveals a significant shift in β factors for residues neighboring K346 (figure 3.26). Flexibility around K346 is lower in the NNDNJ bound structure than the apo structure. The likely scenario, therefore, seems that at pH5 NNDNJ induces local flexibility proximal to K346, contributing to an increased NMR signal.

Like NNDNJ, ambroxol induced an increase in signal intensity. However this increase occurred only at pH7, confirming previous reports. Furthermore ambroxol addition revealed similar effects upon k_{ex} values. Finally, a cell-like environment was reconstituted by combining human cell derived GCase with human cell lysate. GCase is clearly visible in cell lysate at low concentration (50 μM) and CPMG relaxation dispersion experiments suggest that millisecond dynamics observed in buffer are also present in cell-like environments. This suggests that these measurements are transferable to cellular environments.

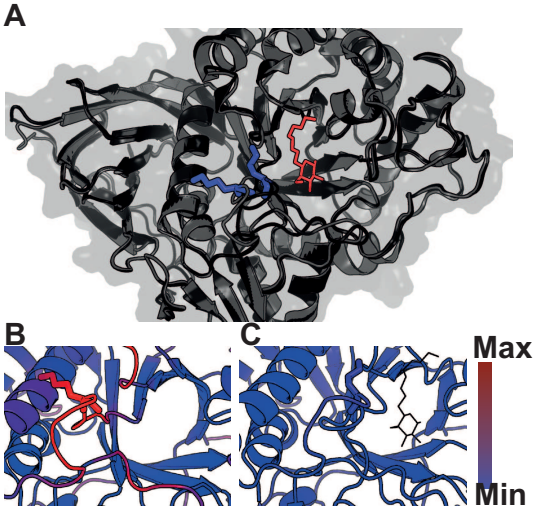


Figure 3.26: *GCase and NNDNJ crystal structures. A) GCase with and without (black and grey) NNDNJ with K346 (blue) and NNDNJ (red) highlighted. B) β factors for apo GCase, C) β factors for NNDNJ bound GCase using blue-red (low-high) spectrum*

I have here developed a non-intrusive labeling method, derived from crystallographic methodologies (first investigated in chapter 7), capable of observing large, PTMed proteins in a native-like environment. I have used this to examine properties of GCase and potential drug-protein interactions. From these studies I observe slow exchange dynamics of GCase at ER pH7 but not

at lysosomal pH5. I further observe the binding of NNDNJ and a concomitant increase in signal intensity associated with a loss of exchange broadening.

3.4 Methods

3.4.1 GCase sourcing

GCase (VPRIV) was sourced from Shire Pharmaceuticals PLC under material transfer agreement. VPRIV is an enzyme replacement therapeutic for Gauchers disease. VPRIV is produced from human cell lines supplemented with kifunensine resulting in high mannose glycans favourable for enzyme replacement therapy [Brumshtein et al., 2010].

3.4.2 Methylation of GCase

It was necessary for this work to devise a labeling protocol for human cell line derived GCase. Labeling of arginine and lysine residues was accomplished by incubating GCase with ^{13}C formaldehyde, using methods previously described [Rayment, 1997]. GCase was buffered in 50mM citric acid, 100mM sodium phosphate, 2mM sodium azide (pH5) or 50mM citric acid, 100mM sodium phosphate, 2mM sodium azide (pH7). Borane dimethylamine complex (Sigma Aldrich) was added to dilute (10μM) GCase at 4°C to a final concentration of 20μM. Immediately, ^{13}C formaldehyde (Sigma-Aldrich) was added to a final concentration of 40μM and the solution incubated on ice at 4°C for 2 hours. This was repeated once and incubated for a further two hours. Finally, Borane dimethylamine is added to a final concentration of 10μM and the reaction incubated overnight at 4°C. The reaction was quenched by the addition of Trizma (pH7) (Trizma, sigma Aldrich) to a final concentration of 100mM. Reaction

byproducts are then removed using Amicon concentrators (Millipore), washing with a buffer containing fresh 50mM citric acid, 100mM sodium phosphate, 2mM sodium azide (pH5) buffer.

3.4.3 LCMSMS and methyl assignment

In order to assign the observed residues, LCMSMS experiments were performed. Peptides were re-suspended in 5% formic acid and 5% DMSO. They were separated on an Ultimate 3000 UHPLC system (Thermo Fischer Scientific) and electrosprayed directly into an QExactive mass spectrometer (Thermo Fischer Scientific) through an EASY-Spray nano-electrospray ion source (Thermo Fischer Scientific). The peptides were trapped on a C18 PepMap100 pre-column (300μm i.d. x 5 mm, 100 Å, Thermo Fisher Scientific) using solvent A (0.1% Formic Acid in water) at a pressure of 500 bar. The peptides were separated on an in- house packed analytical column (75μm i.d. packed with ReproSil-Pur 120 C18-AQ, 1.9μm, 120 Å, Dr. Maisch GmbH) using a linear gradient (length: 60 minutes, 7% to 28% solvent B (0.1% formic acid in acetonitrile), flow rate: 250 nL/min). The raw data was acquired on the mass spectrometer in a data-dependent mode (DDA). Full scan MS spectra were acquired in the Orbitrap (scan range 350-2000 m/z, resolution 70000, AGC target 3e6, maximum injection time 100 ms). After the MS scans, the 20 most intense peaks were selected for HCD fragmentation at 30% of normalised collision energy. HCD spectra were also acquired in the Orbitrap (resolution 17500, AGC target 5e4, maximum injection time 120 ms) with first fixed mass at 180 m/z.

NMR spectra acquired at pH 5 of GCase (VPRIV) pre-treated with formalde-

hyde at either pH 5 or pH7 under otherwise identical conditions yielded different spectra. In both cases, only one methyl addition was identified by LCMSMS to be present under each condition as described in the text. This led to unambiguous assignment of R120, the only isolated, well dispersed peak structurally proximal to the active site E340 and E235 residues. K346, the other methyl identified by LCMSMS led to no obvious change in the NMR spectrum, suggesting its resonance overlays significantly in the center of the spectrum.

3.4.4 GCase activity

GCase activity was performed in triplicate in 96-well clear flat bottom black polystyrene microplates (Product 3904, Corning) with 120 μM substrate (resorufin--D-glucopyranoside (Sigma)) and 50nM GCase buffered in 176mM K2HPO4, 10mM sodium taurocholate, 50mM citrate, 0.01% tween 20, pH 5.9. The final well volume was 0.1ml. Fluorescence of hydrolysed substrate was measured on a Flurostar omega (584/630nm excitation and emission respectively) at 37°C [Motabar et al., 2012]. NNDJ and ambroxol were purchased from Sigma-Aldrich, USA.

3.4.5 NMR

NMR experiments were performed on a Varian 600 MHz spectrometer with a 5-mm z-axis gradient triple resonance room temperature probe at 30 °C. Isotopically labeled ^{13}C GCase was buffered in 50mM citric acid, 100mM sodium phosphate, 2mM sodium azide (pH 5 or pH7) and concentrated to 100 μM using Amicon concentrators (Millipore). 1D HMQCs spectra were recorded

with 128 scans per transient with acquisition times of 67.5 ms over and a 1.5 s relaxation delay. 2D 13C-1H HMQC were recorded with 64 scans per transient and an acquisition time of 0.064 s, over 100 complex points and a 1.5 s relaxation delay.

Diffusion pulsed field gradient experiments were carried out over 11 gradient strengths with 256 scans per transient and an acquisition time of 0.5 s. The maximum gradient strength was 60 G cm^{-1}. The gradient duration was 0.002 s with a 0.15 s delay between gradients.

CPMG experiments were carried out with 0-38 cycles and a relaxation time of 40 ms. NMRPipe was used to analyze spectra and Sparky was used to pick peaks and prepare figures.

3.4.6 Mass spectrometry

Ion mobility (drift time) measurements were carried out on an Ion Mobility QToF (Synapt G1; Waters) modified to incorporate a linear drift tube. Nanoelectrospray mass spectra were acquired under conditions optimized for the preservation of noncovalent interactions and with the following instrument parameters: sample cone, 15 V; extraction cone, 5 V; trap, 20 V; trap gas (argon), 2.6 mL/min; drift-cell pressure (helium), 2.14 Torr. Unless stated otherwise, samples were analyzed at 10μM concentration in 200mM ammonium acetate (pH 7 or pH 5). Drift time measurement followed a single charge species for each protein to allow comparison across conditions.

3.4.7 CPMG analysis

One-dimensional CPMG data was analyzed using Python with SciPy and NMRglue modules [Drake and F.L., 2001, et al. Eric Jones, Travis Oliphant, 2001, Helmus and Jaroniec, 2013]. Overlapped peaks in the NMR spectra were extracted and partitioned into the maximum number of partitions (a total of 16). Each partition had an R^{2eff} calculated and was fitted using a Levenberg Marquardt constrained least squares machine learning algorithm with physically reasonable initial values (R^2=20 s^{-1}, $\delta\omega$=1 ppm, P_B=0.01, k_{ex}=1000 s^{-1}). Fits for representative partitions are presented in figures 3.12 and 3.20. Summary plots (figures 3.23, 3.25, 3.16, 3.15) were made by calculating R^{2eff} for each partition, averaging $R2^{eff}$s and plotting with partition $R2^{eff}$ standard deviation (error bars).

Fitting errors were calculated by bootstrapping the data, that is, resampling the original data with replacement. A normal distribution was subsequently fitted to this bootstrap distributions and mean and standard deviation values derived from this fitted Gaussian.

Chapter 4

Parkinson's disease - α-synuclein and glucocerebrosidase: a tug of war for your mental health

Abstract

The aggregation of the protein α-synuclein (αSyn) into amyloid fibrils, or amyloidogenesis, is associated with Parkinsons disease (PD) and greatly enhanced in the presence of anionic lipids. Recently, the lysosomal protein glucocerebrosidase (GCase) has been identified as a major risk factor in both genetic and sporadic PD. Here, I use solution state NMR to elucidate the protective role of human cell derived GCase in αSyn amyloidogenesis. I demonstrate that GCase can directly inhibit lipid induced αSyn amyloidogenesis. Structurally, I show that the mechanism for this requires competition between lipids and GCase for αSyn, binding the N and C termini respectively. The affinity of GCase for the C-terminus of αSyn is such that not only does it inhibit lipid induced amyloid formation, but also it destabilizes mature αSyn amyloid fibrils. These results reveal a molecular tug of war for αSyn by GCase and lipid and provide a mechanistic link between the clinically observed changes in GCase abundance and Parkinsons disease.

4.1 α-synuclein and GCase in Parkinson's disease

Parkinsons disease (PD) and other synucleinopathies are a group of neurodegenerative disorders characterized by the aggregation of α-synuclein (αSyn) into amyloid rich pathological hallmarks known as Lewy Bodies [Trojanowski and Lee, 2002] (chapter 1). Recent clinical studies have linked PD to the lysosomal lipid lytic enzyme Glucocerebrosidase (GCase). Mutation in GCase is the leading genetic risk factor for PD, with 5-10 % of PD patients carrying one of hundreds of possible GCase mutations [Beavan and Schapira, 2013] (Figure 1.4). Moreover, sporadic PD reveals wild type GCase co-localized with Lewy bodies and an overall depletion of GCase in the neuronal cells of the *substantia nigra*, a key PD associated brain structure [Alcalay et al., 2015, Chiasserini et al., 2015, Murphy et al., 2014, Goker-Alpan et al., 2010, Wong et al., 2004]. Overall, GCase mutation increases the risk of PD by up to 30 fold [Bultron et al., 2010, McNeill et al., 2012].

Lysosomal GCase is depleted in both sporadic and non-sporadic PD [Bendikov-Bar et al., 2011, Ron and Horowitz, 2005, Mazzulli et al., 2011, Gegg et al., 2012, Winder-Rhodes et al., 2013, Chiasserini et al., 2015, Bae et al., 2015, Nichols et al., 2009, Murphy et al., 2014, Murphy et al., 2014]. Healthy cells can degrade both αSyn and αSyn oligomers through a number of pathways including chaperone mediated autophagy, microautophagy and macroautophagy, these pathways culminate in the lysosome [Goker-Alpan et al., 2010, Lee et al., 2004a, Cuervo et al., 2004, Webb et al., 2003, Mak et al., 2010] (figure 1.5). The intercellular transfer of potentially toxic αSyn oligomers between lysosomes is also enhanced by GCase depletion [Mazzulli et al., 2011, Bae et al.,

2014, Alvarez-Erviti et al., 2011]. Finally, αSyn interacts with GCase directly at lysosomal pH to inhibit GCase activity [Yap et al., 2013]. Whilst these links stress the importance of GCase to PD, the protective ability of wild type GCase against PD has yet to be investigated.

Cellular and animal studies also further reveal the importance of GCase to PD. Animal studies of PD demonstrate that knock down of wild type GCase or over expression of mutant GCase causes accumulation of αSyn. Likewise, over expression of wild type GCase can reverse αSyn related pathological and behavioral features in mouse models [Rocha et al., 2015, Liu et al., 2015, Sardi et al., 2011, Sardi et al., 2013]. Taken together, these results suggest a molecular relationship between GCase and αSyn that is pertinent to PD.

On a molecular level, intrinsically disordered αSyn plays a role in synaptic vesicle fusion [Snead and Eliezer, 2014]. The structure of αSyn can be divided into three distinct regions, an N-terminal lipid binding domain, a central pro-amyloidogenic domain and a C-terminal domain that remains free when lipid bound [Stefanis, 2012, Fusco et al., 2014] (see 1.2.1 for more detail). The aggregation of αSyn into amyloid has been studied extensively in vitro [Giehm et al., 2010, Buell et al., 2014, Hoyer et al., 2002, Galvagnion et al., 2015] (see 1.2.1.1 and 1.2). Crucially, its aggregation into amyloid fibrils, characterized by their ThT binding and cross β sheet structure [Vilar et al., 2008, Heise et al., 2008, Zhou et al., 2012, Ecroyd et al., 2008], is accelerated 1000 fold in the presence of membrane mimicking anionic lipids such as DMPS and POPG [Wang et al., 2010, Lokappa and Ulmer, 2011, Drescher et al., 2008, Drescher et al., 2010, Braun et al., 2014, Galvagnion et al., 2015, Bodner et al., 2009]. Furthermore, acidic environments, like those found in the lysosome, further promote αSyn amyloidogenesis [Wu et al., 2009, Hoyer et al., 2002, Buell et al.,

2014]. The link between GCase, acidity and lipids suggest the lysosome as a likely organelle for the onset of αSyn pathogenesis [Lee et al., 2004a, Alvarez-Erviti et al., 2011, Bae et al., 2015, Mak et al., 2010, Mazzulli et al., 2011, Cuervo et al., 2004, Liu et al., 2015, Goker-Alpan et al., 2010].

GCase is a 60kDa glycoprotein found in the acidic compartment of the lysosome, where it catalyses the conversion of glucocerebroside into glucose and ceramide [Mehta, 2006]. Post translational modifications are pivotal for GCase activity and stability, ablation of its four N-linked glycosylation sites *in vitro* leads to an inactive and unstable enzyme [Berg-Fussman et al., 1993]. In order to understand the interplay between GCase and αSyn it is therefore necessary to work with human derived GCase (see 3.1 and 1.2.2).

To understand how GCase can influence amyloidosis, I investigated the relationship between lipids, human cell derived GCase and full length αSyn under acidic conditions. Solution state nuclear magnetic resonance (NMR) allows the examination of multiple binding partners with transient interactions at high resolution. I have therefore developed a novel labelling method for isotope incorporation and assignment of large (60 KDa) human cell derived proteins for NMR (see chapter 3 and section 3.2.2). I employ a range of lipids and detergents from micelle forming SDS, to physiologically relevant DMPS (summarized in figure 4.1). Such phosphatidylserines (PS) form micelles and are found in lysosomal inner leaflets and synaptic vesicles [Takamori et al., 2006, Mirnikjoo et al., 2009]. These lipids and detergents cover a range of charges and chain lengths previously used to study αSyn and allow the generalization of observations [Drescher et al., 2010, Braun et al., 2014, Drescher et al., 2008, Wang et al., 2010, Lokappa and Ulmer, 2011, Takamori et al., 2006, Mirnikjoo et al., 2009].

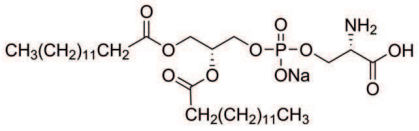
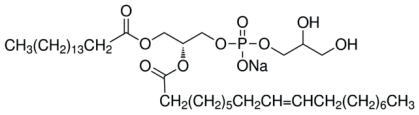
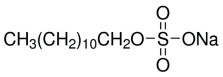
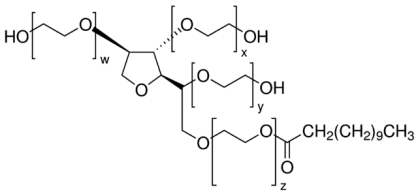
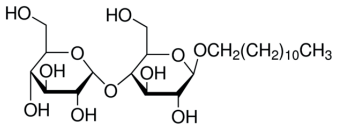
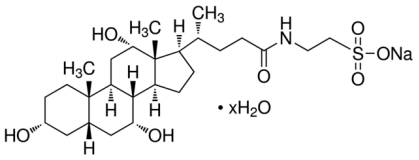
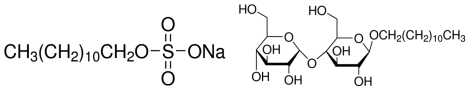
Lipid	Charge (pH5)	Chemical Structure	CMC (mM)	Macro structure	Size
DMPS	-1		0.01	SUV	<100nm
POPG	-1		0.3	SUV	<100nm
SDS	-1		8	Micelle	3.5nm
Tween	0		0.4	Micelle	5.4nm
DDM	0		0.17	Micelle	6.4nm
Taurocholate	-1		10	Micelle	-
1:1 SDS/DDM	-1		0.2	Micelle	-

Figure 4.1: *Summary of lipids under consideration in this study together with their expected charge, the type of assembly, the critical concentration for assembly and where known, the average size of the aggregate. The abbreviations are n-dodecyl--D-maltoside (DDM), 1-hexadecanoyl-2-(9Zoctadecenoyl)-sn-glycero-3-phospho-(1'-rac-glycerol) (POPG), 1,2-ditetradecanoyl-sn-glycero-3-phospho-L-serine (DMPS) and sodium dodecyl sulphate (SDS), 1,2-Dimyristoyl-sn-glycero-3-phospho-L-serine (DMPS). DDM was mixed with POPG and SDS in this study to provide kinetically stable micelles amenable for NMR studies.*

By means of NMR, I elucidate a competition for αSyn between GCase and anionic lipids that rescues αSyn from pro-amyloidogenic pathways under healthy lysosomal conditions, but which promotes amyloidogenesis upon GCase depletion. Furthermore, under these conditions GCase can destabilize pre-formed αSyn fibrils. I reveal that αSyn N-terminus binds anionic lipids whilst the C-terminus does not. By contrast the C-terminus of αSyn binds GCase whereas the N-terminus does not. The differential affinity for these binding partners allows GCase to rescue αSyn from lipid binding and subsequent amyloid formation. Furthermore, αSyn contacts GCase near its active site, revealing a structural mechanism for GCase inhibition. Taken together, these results reveal that αSyn amyloidogenesis is accelerated under acidic conditions, and that GCase is a potent molecular chaperone inhibiting aggregation of αSyn and protecting against PD.

4.2 Results: GCase inhibits αSyn amyloidogenesis under acidic conditions

To mimic the amyloidogenesis of αSyn in the lysosome, αSyn aggregation was monitored using the fibril specific dye thioflavin-T (ThT) at pH 5 in the presence and absence of both GCase and assemblies of anionic lipids (Figure 4.2) [Takamori et al., 2006, Galvagnion et al., 2015]. The aggregation of αSyn at pH 5 was markedly faster in the presence of DMPS, reducing the lag time from 3 days (red), to 1 (blue) (Figure 4.2). Although the aggregates produced bound ThT and possessed fibrillar morphological aspects, discrete fibrils were not observed in electron micrographs (Figure 4.2A). Remarkably, adding GCase effectively prevented aggregation both in the presence (cyan)

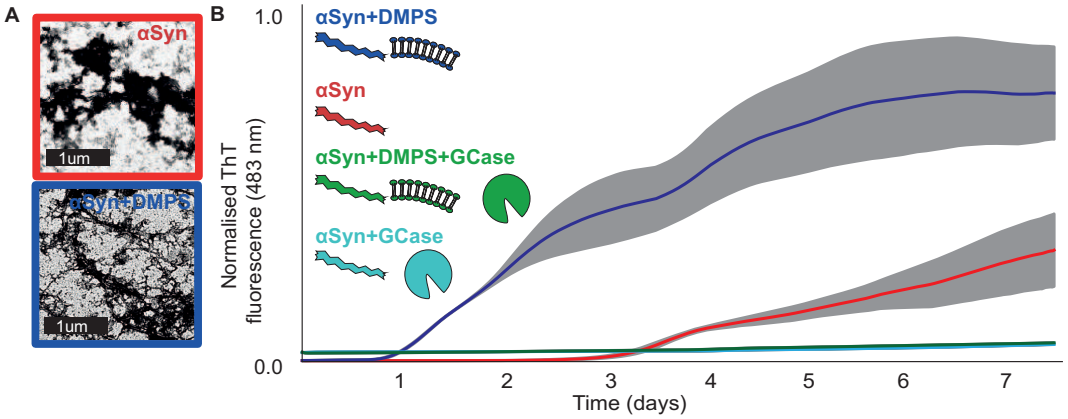


Figure 4.2: *Aggregation rates of αSyn are effected by GCase and lipid, (100 μM GCase, 100 μM αSyn, 1mM DMPS, pH 5, 30°C) A) Negatively stained electron micrographs reveal aggregation of S and S+DMPS samples at pH 5 after 7 days. Though they stain with ThT and the aggregates have a fibrillar aspect, discrete fibrils were not observed. B) αSyn aggregation kinetics were monitored by a thioflavin-T (ThT) assay at pH 5 with DMPS (blue), GCase (cyan) and both (green). Aggregation of αSyn (red) is promoted by DMPS (blue) and inhibited by GCase both with and without DMPS (cyan/green). The grey band shows the standard deviation of 3 independent measurements.*

and absence (green) of DMPS, demonstrating that it is a potent inhibitor of αSyn amyloidogenesis under acidic conditions.

4.2.1 GCase and anionic lipids compete for opposite termini of monomeric αSyn

In order to understand the molecular basis of this interaction, I recorded fingerprint ^{15}N-^{1}H HSQC NMR spectra of αSyn in absence and presence of DMPS, POPG, SDS, DDM, and tween under otherwise identical conditions. Assignment of αSyn was achieved by comparison to previously published assignments (4.4.6 [Bodner et al., 2009]). In a ^{1}H^{15}N HSQC spectrum, one resonance is observed for each distinct HN environment in the molecule. ^{1}H NMR resonances fell predominantly in the region 8-8.5 ppm, confirming the predominantly disordered character of αSyn (Figure 4.3).

In the presence of the anionic taurochlorate, DMPS and 1:1 mixtures of DDM with SDS and POPG (figure 4.1) a residue specific reduction in signal intensity was observed, with no significant change in either resonance frequencies or observed linewidths (figures 4.3 and 4.4). In line with previous observations, N-terminal αSyn residues showed a greater decrease in signal intensity than those in the C-terminus (figure 4.4). Moreover, a higher lipid ratio decreased the N-terminal NMR signal intensities further. By contrast, micelles formed from the non-ionic detergents DDM and tween resulted in no change in signal intensity (figure 4.5). In the case of the anionic lipids, the observed diffusion coefficients of αSyn were significantly reduced (figure 4.6). Together, the loss in signal intensity can be explained by the formation of αSyn lipid complexes, where any structured portion of the molecule tumbles too slowly to be observed directly by solution state NMR.

Next I determined the effect of GCase on αSyn in a 1:1 mixture of both components. Contrasting with lipid binding, a signal intensity reduction was observed localized to the C-terminus of αSyn (Figure 4.3). Incubation of αSyn in the presence of lipids and GCase (figure 4.3) resulted in a recovery of signal intensity in both the N- and C-termini. Furthermore, NMR diffusion coefficients of αSyn returned to values similar to those measured for the isolated form (Figure 4.6). These results indicate that GCase substantially rescues αSyn from its lipid-bound amyloidogenic state.

To allow for a quantitative comparison of the effects at the N- and C-termini of αSyn, the signal intensities of the first and last 20 residues were integrated for each of the mixtures of αSyn with GCase and lipid micelles. Although the magnitudes of the interaction of αSyn with different lipids vary, GCase was always found to rescue the signal intensity of αSyn (figure 4.7).

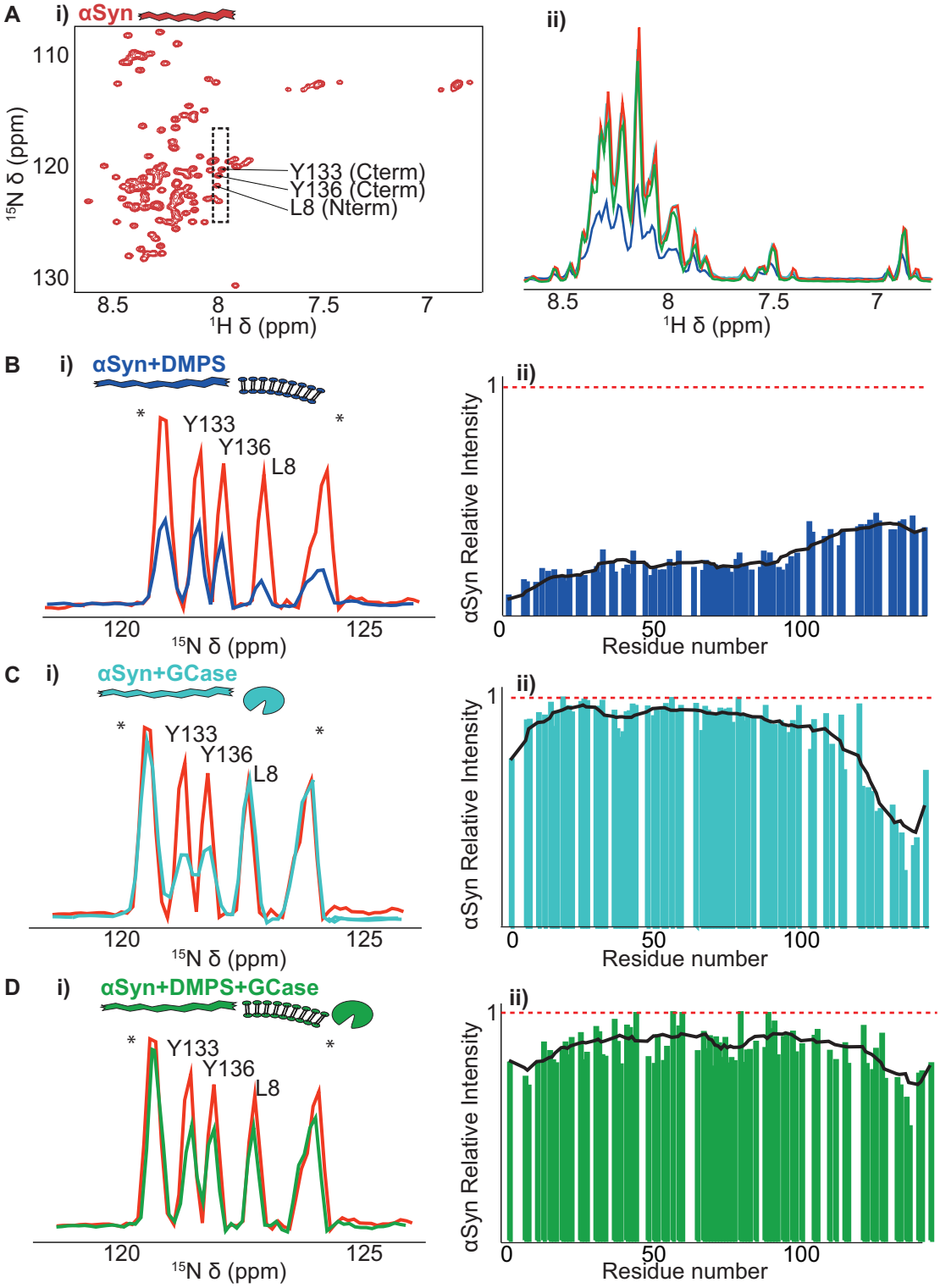


Figure 4.3: *GCase rescues αSyn from lipid binding under acidic conditions at 1:1 molar ratios (pH7, 30°C). A) i) A lack of αSyn signal dispersion in a 2D 1 H-^{15}N HSQC spectrum confirms its intrinsically disordered character. ii) Incubating 100 μM αSyn (red) with 1mM of the anionic lipid DMPS (blue) attenuates the intensity of all signals, whereas the addition of 100 μM GCase (green) recovers it. In B, C, and D, the (i) same 2D 1 H-^{15}N HSQC region (see box in Ai) and (ii) corresponding αSyn signal intensities are shown in absence or presence of DMPS or GCase, respectively, relative to the signal intensity of free αSyn.*

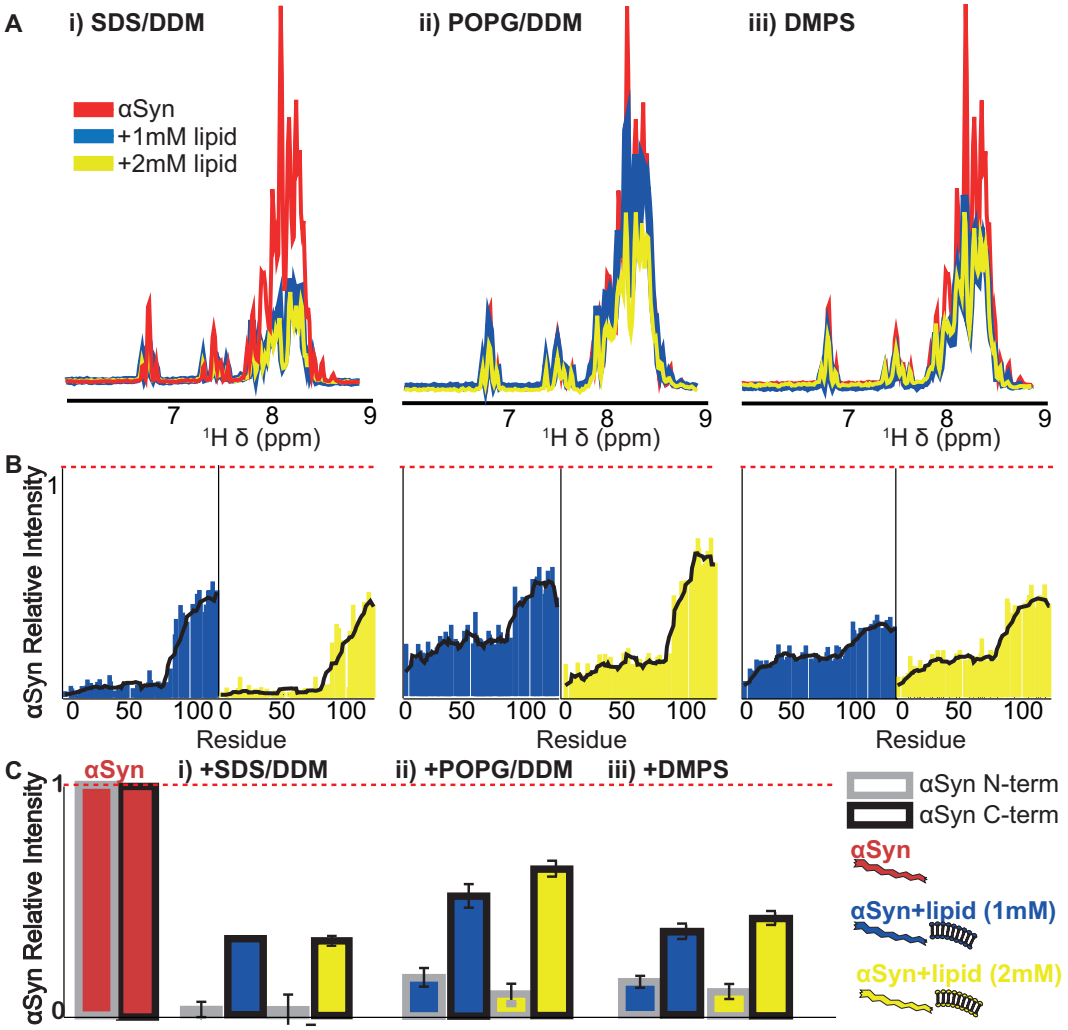


Figure 4.4: *The effect of increasing lipid concentration on αSyn interaction with anionic lipids and detergents. A) 1D $^{15}N^{1}H$ HSQCs of αSyn in the absence (red) and presence (blue 1mM, 1:10 αSyn/lipid and yellow 2mM 1:20 αSyn/lipid) of SDS+DDM (i) POPG+DDM (ii) and DMPS (iii). B) The relative intensity of resonances coming from individual αSyn residues upon addition of the different lipids. C) The integrated signal intensity of the first and last 20 residues. The data from 1 mM DMPS is shown in figures 2B and 3B.*

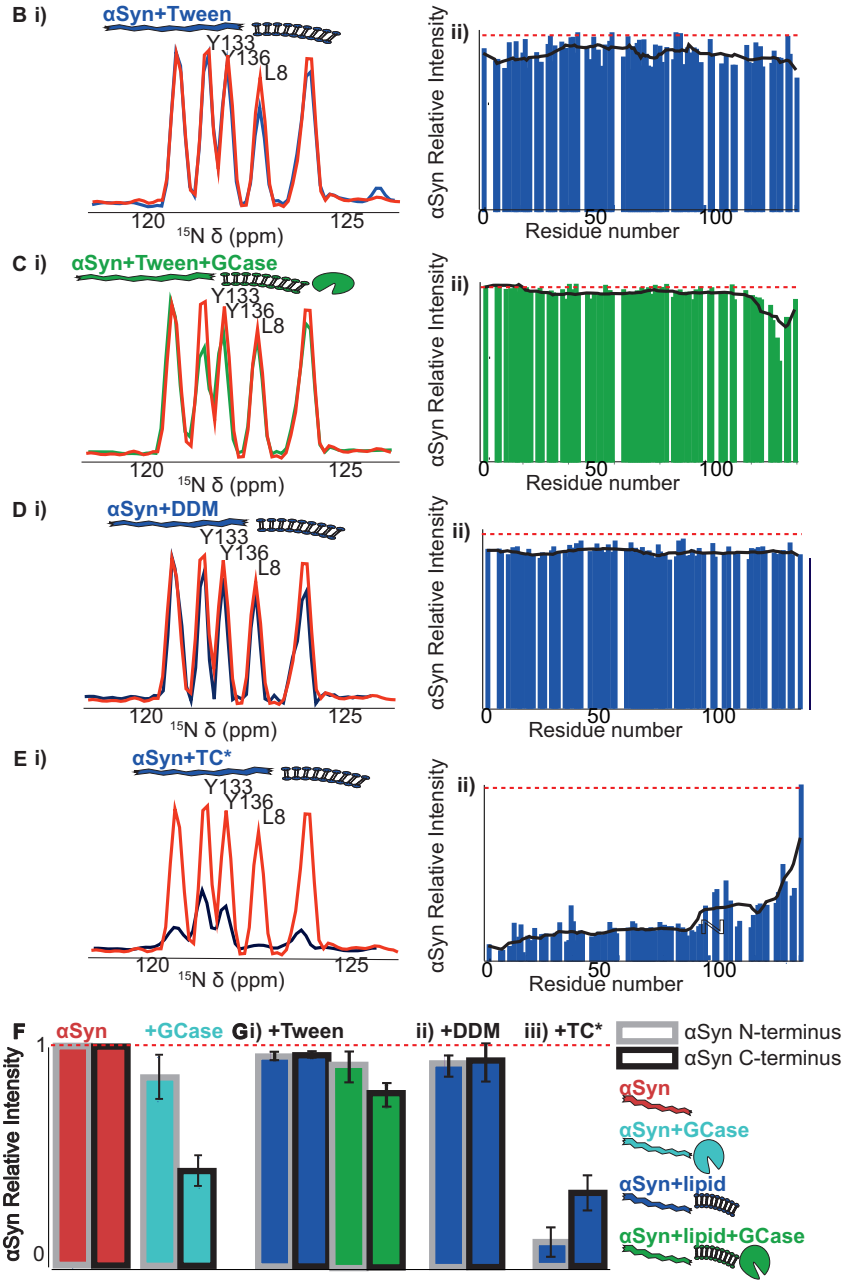


Figure 4.5: *αSyn does not bind assemblies formed from the non-ionic detergents tween (10mM) and DDM (1mM), but does bind assemblies of the anionic detergent taurochlorate (TC, 1mM) (100 μM αSyn, pH5, 30°C). A i) 2D $^{15}N^{1}H$ HSQC of free αSyn. ii) 1D $^{15}N^{1}H$ HSQC of αSyn with both lipids, and GCase. B-E i) extracted 1D region from 2D ^{15}N ^{1}H HSQC spectra showing effects of lipid addition on resonances coming from specific N and C terminal residues. ii) Relative intensity of individual resonances from αSyn upon lipid addition. F Integrated signal intensity from the first and last 20 residues of αSyn (red) upon addition of GCase (cyan), and the non-ionic detergents tween and DDM, and the anionic detergent taurochlorate. Tween and DDM have no appreciable effect on αSyn. In the presence of tween, the interaction of GCase with αSyn is attenuated.*

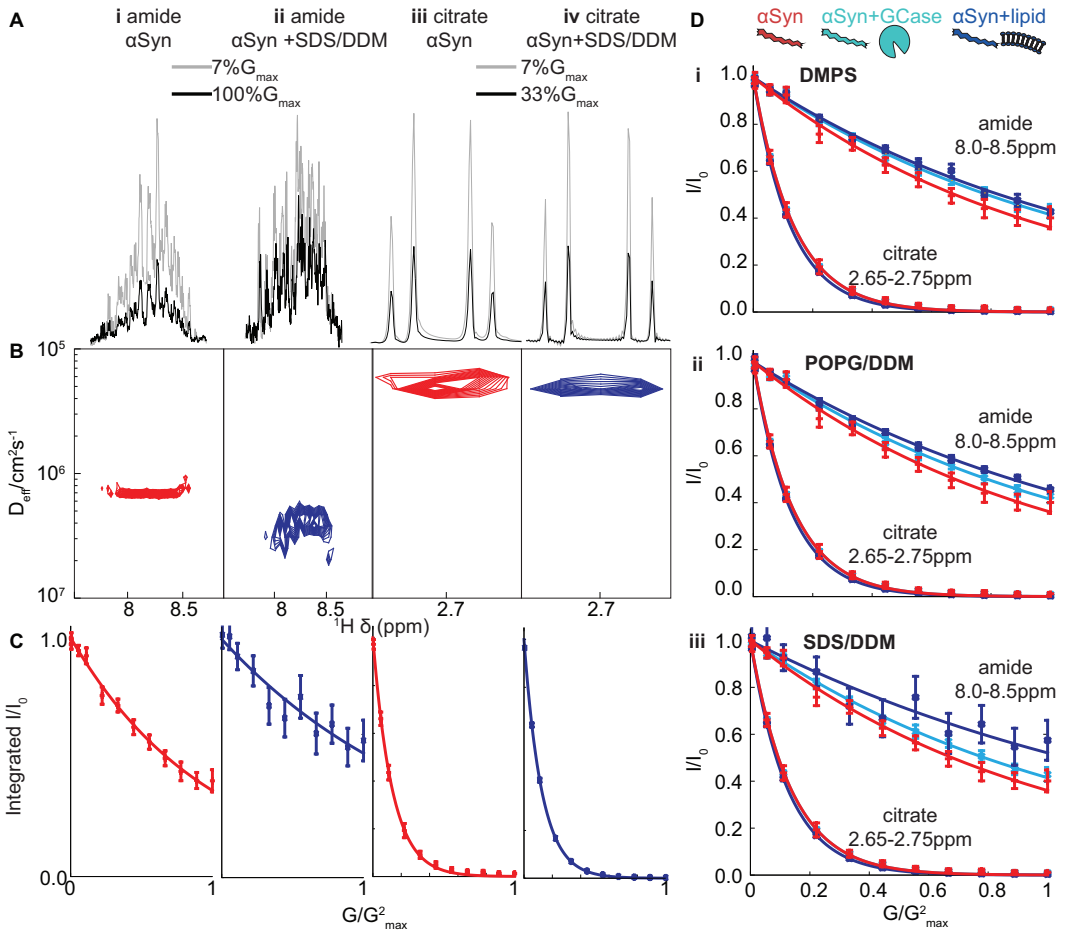


Figure 4.6: *Translational diffusion measurements (100 μM αSyn, pH5, 30°C).A) PGFSE NMR spectra of αSyn with and without 1mM SDS/DDM are shown at high (black) and low (gray) gradient strengths. i) Signals in the amide region that originate only from αSyn. ii) After addition of 1mM (total) SDS/DDM, the attenuation at high gradient strength is reduced, indicating a reduction in diffusion coefficient. iii) In the same spectrum, resonances from citrate can be observed in the region of 2.7 ppm providing an internal control for changes in viscosity. iv) The decay of the citrate resonances is unaffected by the addition of SDS/DDM, revealing that the viscosity of the sample is largely unchanged by the addition. B) The diffusion coefficient of each identified chemical in the spectrum can be obtained. The collection of resulting values is presented here as a 2D histogram quantifying the change in observed diffusion coefficients on changing solution conditions. C) The regions indicated are integrated and shown as a function of the individual gradient strengths applied. The error bars come from taking the standard deviation of the normalized signal intensity I/I0 over the integrated region, where I0 is the fitted intensity at zero gradient strength for each individual chemical shift. D) the relative decay in signal intensity is shown for the amide region of αSyn, and the citrate resonances in the presence and absence of GCase and lipids (all 1mM)*

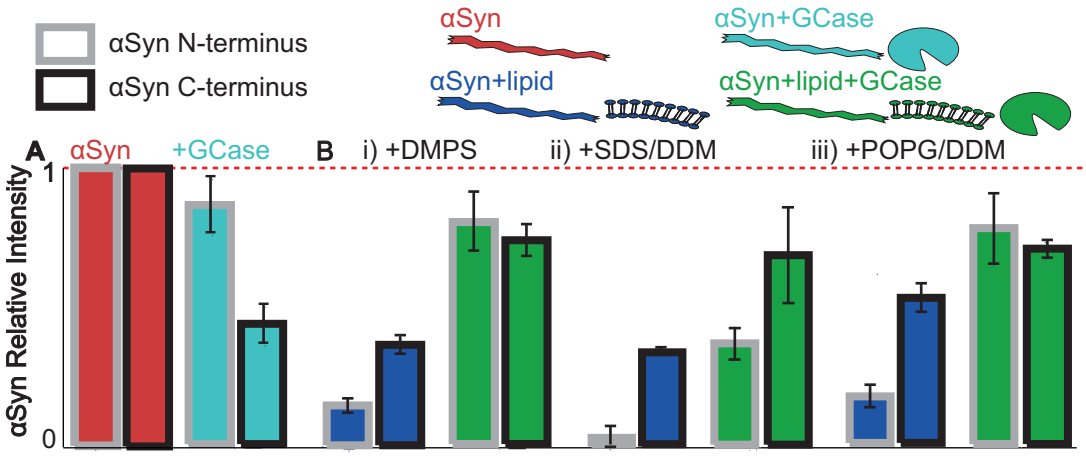


Figure 4.7: *GCase rescues αSyn from a variety of anionic lipids (1:1:10 molar ratios of GCase, αSyn and (monomeric) lipid). Integrated signal intensities of the 20 most N- (grey) and C- (black) terminal αSyn residues in the presence of GCase (cyan), anionic lipid micelles formed from DMPS, SDS/DDM or POPG/DDM (blue) and their combination (green). In each case, there is substantial recovery in the observed signal intensity by adding GCase (green).*

4.2.2 GCase destabilizes mature αSyn amyloid fibrils

I sought to examine the effect of GCase on pre-formed amyloid fibrils. At elevated temperatures and concentrations (200 μM, 45°C) ^{15}N labeled αSyn showed increased ThT fluorescence (figure 4.8), and discrete amyloid fibrils were observed by EM (figure 4.8Aii) [Galvagnion et al., 2015]. A ^{1}H-^{15}N HSQC NMR spectrum of these samples at pH 5 revealed a small population of flexible C-terminal residues observable by NMR, a finding in agreement with proteolysis and solid state NMR studies [Heise et al., 2005, Vilar et al., 2008, Comellas et al., 2012]. Most remarkably, addition of GCase to these samples recovered signal from N-terminal residues (figure 4.8C), indicating a destabilization of the amyloid form and partial rescue of free αSyn by GCase.

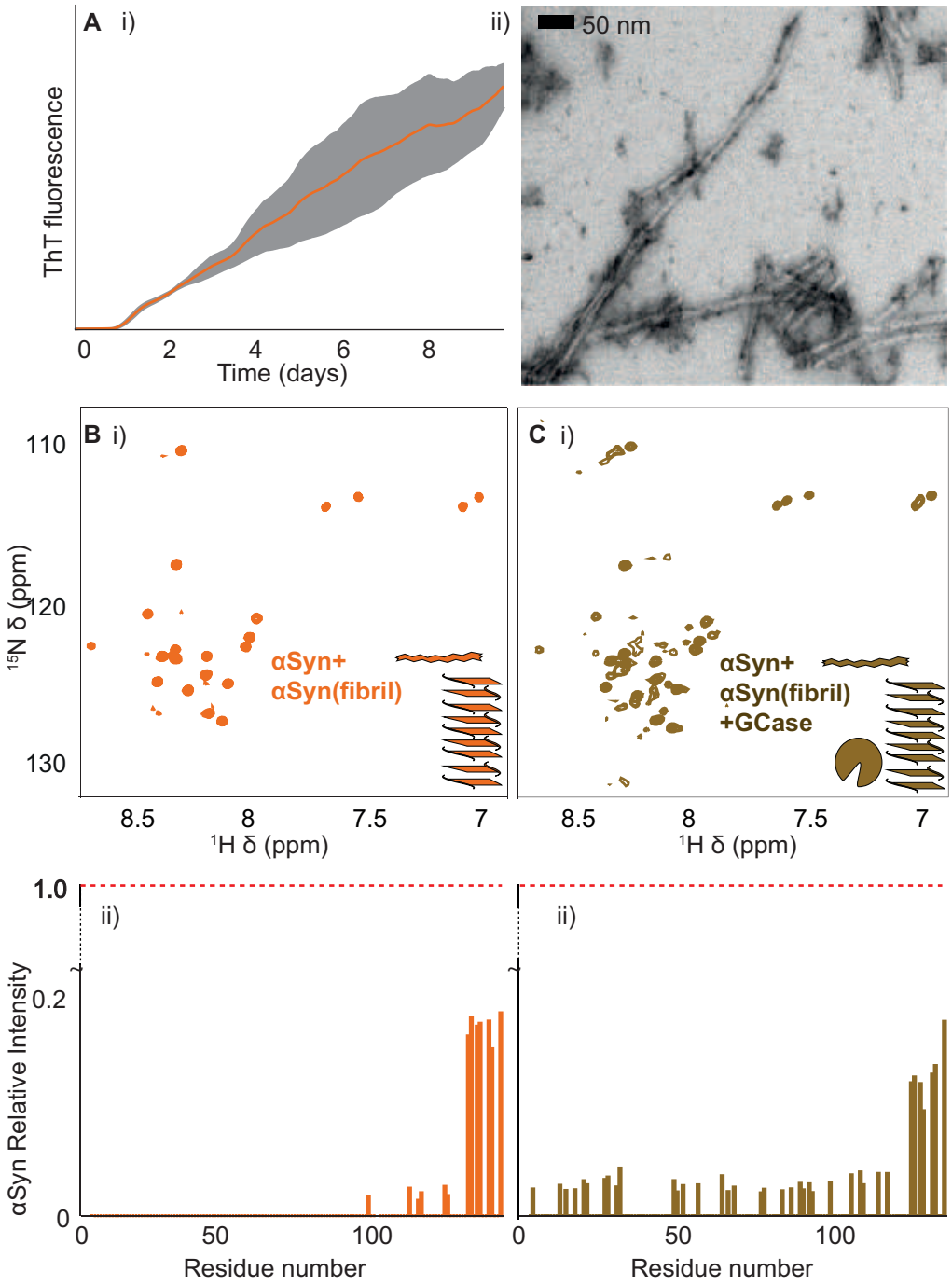


Figure 4.8: *A) i) The αSyn aggregation kinetics were monitored by a thioflavin-T fluorescence assay at an elevated temperature and concentration (45°C, 200 μM) at pH5. ii) These conditions result in discrete amyloid fibrils, as observed by EM (recorded after 9 days (Ai)). B/C) i) 1 H-^{15}N HSQC spectra of purified amyloids (100 μ monomeric concentration) before (B,orange) and after incubation with 100 μM GCase (C,brown) under acidic conditions (pH 5, also after 9 days). ii) The relative intensity (vs. apo αSyn) of observed resonances as a function of residue position. Addition of GCase rescues αSyn from its fibrillar form, increasing N-terminal intensity, whilst suppressing C-terminal intensity due to GCase-αSyn complex formation.*

4.2.3 GCase interacts with αSyn near its active site

GCase (VPRIV, Shire PLC) was produced from human cell lines and incubated with ^{13}C formaldehye to produce lys^{13}C $^{1}H_3$ and arg^{13}C $^{1}H_3$ enriched GCase (chapter 3).

The ^{13}C $^{1}H_3$ methyl groups were used to monitor the effects of both lipids and αSyn on GCase (figure 4.10 4.9). The addition of either αSyn or DMPS resulted in decreased GCase signal intensity. Most notably, a modest yet significant change in chemical shift of the peak arising from R120 was observed on mixing with αSyn, whereas the resonance frequencies of the remaining residues were largely unchanged. This suggests that the interaction site is in the vicinity of this location. As R120 is close to the active site, this observation is consistent with αSyn reducing the activity of GCase [Yap et al., 2013, Mazzulli et al., 2011] and with FRET data [Yap et al., 2013].

4.2.4 GCase prevents αSyn aggregation through a competitive equilibrium

The data presented here facilitates the study of interactions between αSyn, lipids and GCase from the perspective of αSyn and GCase. The loss of signal intensities in both αSyn and GCase were both reversible, and dependent on the concentrations of the three species, suggesting that these two species are in competition with lipids. I sought to determine whether a simple equilibrium model could explain the data. To do this I needed to analyze the possible origin of intensity losses in the NMR spectra (see 2.2.1 and 2.2.2).

As I do not directly observe either micelles or GCase in ^{1}H NMR experiments due to their slow tumbling, I would not expect to observe structured

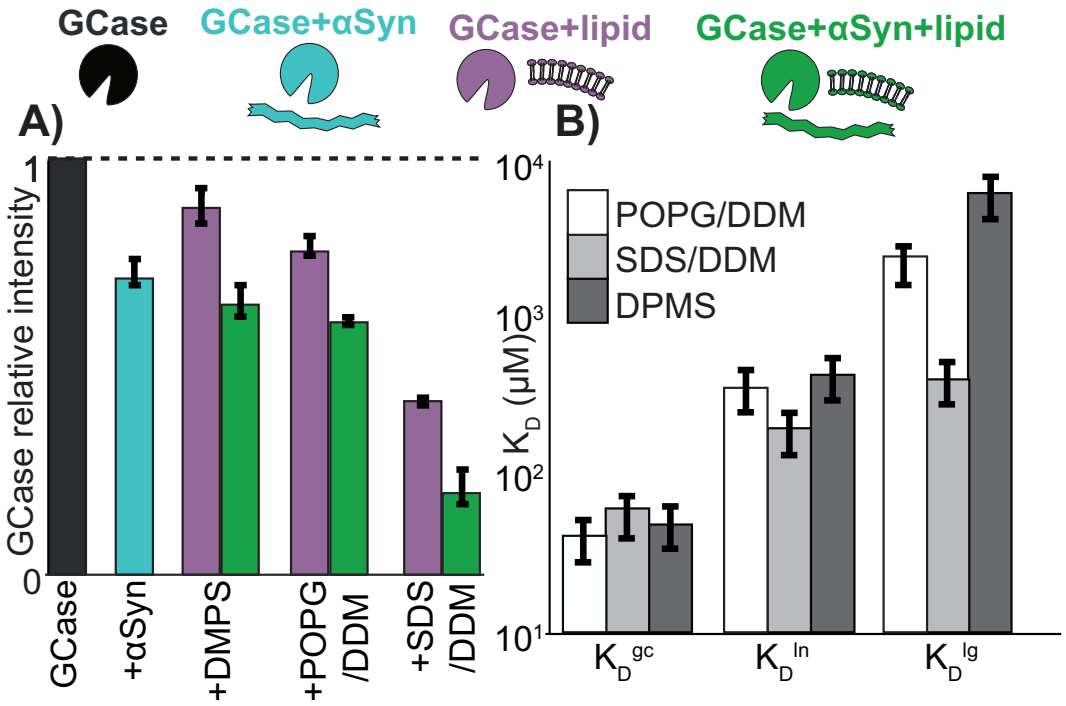


Figure 4.9: *Summary of the changes in resonance intensity of $^{13}C\,^{1}H_3$ labeled GCase on incubation with 1:1 (100 μM) αSyn only (cyan), the lipids DMPS, POPG/DDM, SDS/DDM (purple - 1mM), and a mixture of both αSyn and lipid (green). B) Quantitative Kd values for the three processes from fitting the data. The dissociation constant for GCase/αSyn complex formation was independent of lipid choice. The specific Kds for both αSyn and GCase lipid binding are highly dependent on the choice of lipid, and are highly correlated. Lipid Kds are given for monomeric lipid. The strongest binding was observed with SDS/DDM, and the weakest with DMPS. Kd^{gc} describes the interaction between the C-terminus of αSyn and GCase, Kd^{ln} describes the interaction between the N-terminus of αSyn and lipids, and Kd^{lg} describes the interaction between GCase and lipids.*

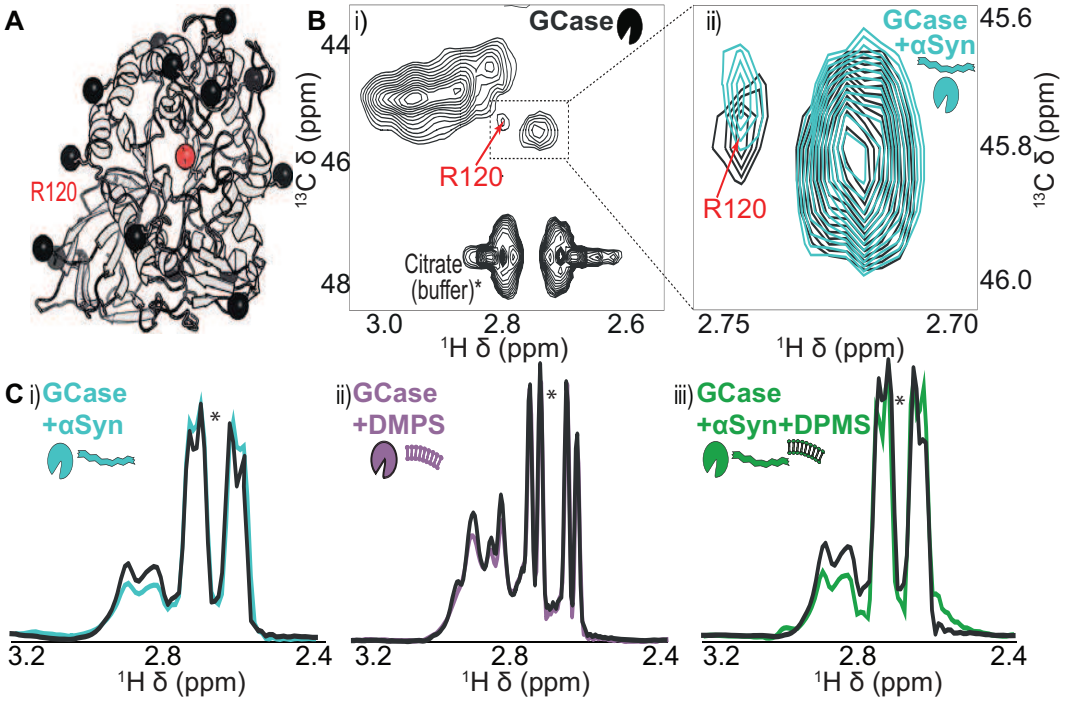


Figure 4.10: *A) Lysines and arginines on the surface of GCase (PDB 1OGS) were modified by incubation with ^{13}C labeled formaldehyde (black spheres). B)i) 2D $^{1}H^{13}C$ HMQC NMR spectrum of the labeled residues. Natural abundance of ^{13}C in citrate results in sharp resonances in the indicated positions (*). While the majority of the resonances overlap, one distinct peak from R120 could be distinguished and unambiguously assigned using LCMSMS (Fig S5). ii) After incubation with 100 μM αSyn, the locations of all the resonances were largely unchanged. Only the resonance corresponding to R120 was found to move suggesting a direct interaction between this site and αSyn. C)^{1}H projections of 2D spectra after arginine methylation (black) and mixed (cyan/purple/green). As GCase samples were prepared independently, small fluctuations in spectra between samples were observed (black i,ii,iii). The small differences in signal intensity from the citrate in projection are due to truncation artefacts in the spectra, and do not reflect changes in citrate concentration. The uniform loss of signal intensity from resonances originating from GCase after incubation with αSyn, DMPS or a mixture of both indicates a formation of high weight species that are not directly observed in a solution-state NMR spectrum.*

αSyn bound to either lipids or GCase, and so complex formation will result in the loss of signal intensity. By contrast, if for example the N-terminus of αSyn bound either a lipid aggregate or GCase, and the C-terminus remained flexible, I would still expect it to contribute to the spectrum albeit with an attenuated signal intensity. Moreover, I do not directly observe broadening or chemical shift movement in spectra of αSyn [Fusco et al., 2014, Bodner et al., 2009, Eliezer et al., 2001] as well as no variation in transverse relaxation rates in CPMG experiments. Taken together, these results suggest that chemical exchange within slow-exchange limits [Fusco et al., 2014, Bodner et al., 2009, Eliezer et al., 2001]. I account for these effects in the NMR spectra by allowing regions of αSyn that are expected to retain significant flexibility to contribute to the NMR spectrum in a manner that is proportional to the concentration of complex formed, thought reduce from that of the free state (section 4.4.8). This exchange regime is also consistent with the reduced diffusion coefficients upon binding (figure 4.6). If both lipid-bound and free αSyn were observable by NMR then I would expect a bi-exponential diffusion relation, however, as the lipid-bound form is invisible to NMR (likely due to slow tumbling), I do not observe bi-exponential behavior.

Subsequently, I tested a wide range of possible binding models (figure 4.11) and used statistical tests to identify the simplest model explaining all of the data, taking care to avoid over-fitting (section 4.4.8). The simplest model tested considers only binary complexes and does not distinguish between binding to the N- and the C-termini of αSyn, whereas the most complex model included all possible ternary complexes (figure 4.11). The optimal model for binding of DMPS, POPG/DDM and SDS/DDM allows for both GCase and the N-terminus of αSyn to bind lipids, and the C-terminus of αSyn to bind

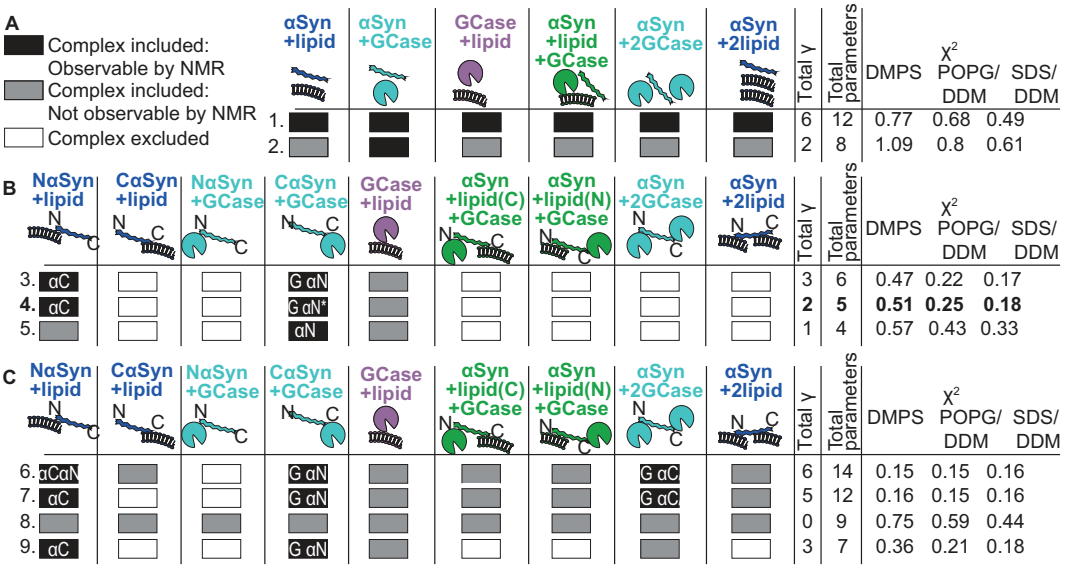


Figure 4.11: *9 of the models tested, enabling identification of the most statistically significant model. 21 measurements of signal intensity were recorded for each lipid, comprising measurements of αSyn/N, αSyn/C or GCase signal intensity with varying concentrations of the three components. A number of binding models were constructed, varying in both the number of species considered (affecting the number of Kds) and the number of species that contribute to the different signal intensities, albeit reduced from the free form by a factor γ, approximating the effects of chemical exchange. Formation of a complex can contribute directly to αSyn/N, αSyn/C or GCase signal intensity, denoted by N, C or G respectively. N/C indicates the species contributes to the signal intensity of both the N- and C- termini of αSyn, with the same γ factor. An asterisk indicates the γ factor is 1, such that the contribution to the NMR spectrum is identical to that of the free state. A) The simplest models do not account for a difference in the N- and C-termini binding affinities and uniformly produce large χ2 values, indicating a poor fit. B) Models were tested including only binary complexes, and a limited number of species were assumed to contribute to the change in signal intensity. C) Models with both a large number of NMR observable species and complexes gave generally lower χ2 values. No appreciable decrease in χ2 was observed with more complex models. one of these values were sufficiently low to pass an F-test (probability of <5%) over the simplified models identified in B. Overall, the most statistically significant model was found to be 4. In this model, there are Kds for the 3 complexes; the C-terminus contributes to the spectrum when the N-terminus is bound to lipids with reduced signal intensity, as does GCase when bound to the C-terminus of αSyn. Finally, the N-terminus contributes to the spectrum when the C-terminus is bound to GCase with no reduction in intensity versus the free state.*

A	Av. error	GCase-αSyn(C) γ_{GCase}	GCase-αSyn(C) K_d^{gc}	GCase-Lipid K_d^{lg}	αSyn(N)-Lipid $\gamma_{\alpha Syn(C)}$	αSyn(N)-Lipid K_d^{ln}
DMPS	18%	0.3	49 μM	5800 μM	0.4	453 μM
POPG/DDM	17%	0.6	41 μM	2300 μM	0.5	362 μM
SDS/DDM	14%	0.4	57 μM	400 μM	0.4	194 μM

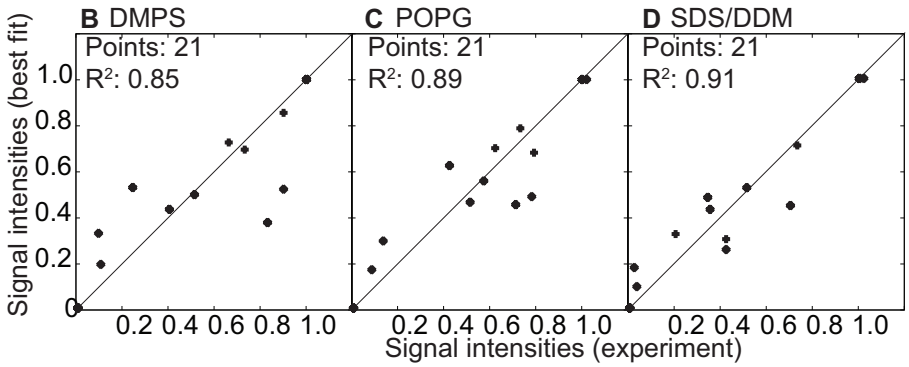


Figure 4.12: *Summary of Kd values and attenuation constants (γ) for the optimal model (figure 4.11). γ GCase is the attenuation of signal for GCase when bound to the C-terminus of αSyn, γ αSynC is the attenuation factor for signal contributing the C-terminus of αSyn when the N-terminus is bound to lipids. The N-terminus of αSyn when the C-terminus is bound to GCase contributes to the signal to the same degree as free αSyn ($\gamma\alpha Syn$ = 1). The average error on each data point was determined by assuming the reduced χ^2 is equal to 1, then Av.Err= (χ2/(-p))1/2 where N is the number of data points and P the number of parameters in the fit. B-D) The Pearsons correlation coefficient between the experimental signal intensities and those predicted by the model is shown.*

GCase. This model includes three dissociation constants for the formation of complexes between GCase and lipids (Kd^{lg}), the C-terminus of αSyn and GCase (Kd^{gc}) and the N-terminus of αSyn and lipids (Kd^{ln}) (figure 4.11 4.12). It also includes an attenuated contribution to the NMR signal from an αSyn termini where the opposite end is involved in complex formation. This model quantitatively explains all of the data (figure 4.12) and is consistent with the decreased diffusion coefficients (figure 4.6).

Overall, this simple equilibrium binding confirms that the N- and C-termini of αSyn are in competition with lipids and GCase, respectively (figures 4.12

4.11 4.3). The Kdgc for the GCase/αSyn(C) interaction was ca. 60 μM for all lipids, whereas Kd^{ln} for the αSyn(N)/lipid followed the degree of charge on the lipid head group in the sequence DMPS, POPG/DDM, SDS/DDM. It is interesting to note that the lipids that bound GCase more tightly also had a higher affinity for αSyn suggesting the two compete for binding sites on the surface of lipids.

All of the data can be quantitatively summarized on a free energy surface (figure 4.13). Free αSyn (red) will form amyloids (orange) if left for sufficient time. This process is greatly accelerated by anionic lipids (blue). GCase effectively competes with lipids for binding of αSyn (cyan). By reducing the population of lipid bound complexes GCase prevents the build-up of amyloidogenic precursors to inhibit amyloid formation. If left for long enough, stable amyloid formation is expected to become inevitable [Baldwin et al., 2011a]. Yet the formation of rapid and relatively weak contacts between GCase and αSyn greatly curtails amyloidogenesis.

4.3 Discussion

Here, I demonstrate that GCase directly inhibits amyloidogenesis of αSyn both in the presence and absence of anionic lipids under acidic solution conditions (figure 4.2). Quantitative analysis of the changes in both GCase and αSyn reveals a molecular tug of war mechanism in which GCase and anionic membrane mimics are in competition for the C- and N-termini of αSyn respectively (figure 4.7). The strongest interaction identified here between GCase and the C-terminus of αSyn inhibits amyloidogenesis, revealing a protective role for GCase in PD pathology. Complexes between αSyn and GCase inhibit the

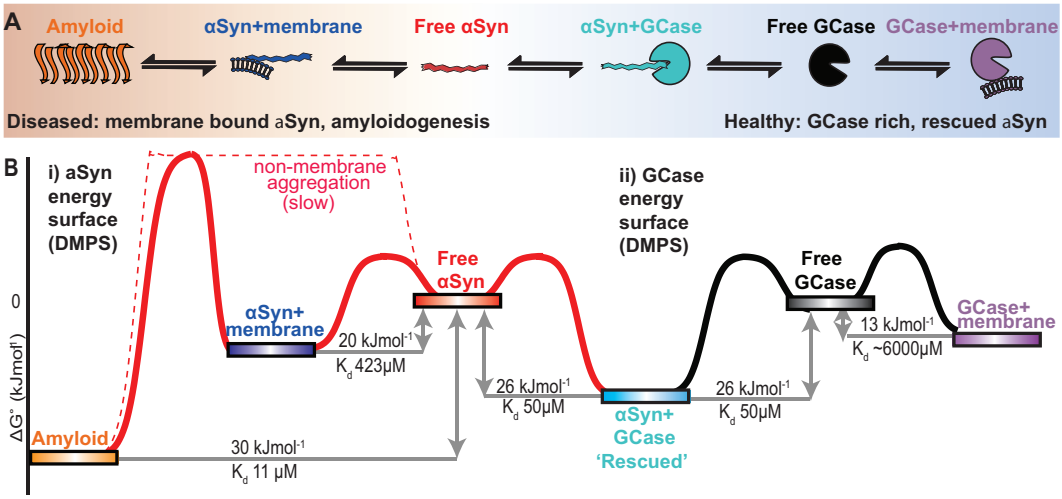


Figure 4.13: *A quantitative equilibrium competition model between αSyn, GCase and lipid well explains all of the data. αSyn aggregation is accelerated by the presence of anionic lipids. GCase binding greatly inhibits this process, inhibiting αSyn amyloidogenesis. B) A schematic free energy surface for DMPS displays how all complexes are related. The presence of αSyn- GCase complexes effectively rescues αSyn by preventing their build up on the surface of membranes. The energy landscape is shown from the perspective of αSyn (i) and from GCase (ii). All binding reactions came to equilibrium at maximum on the timescale of a few minutes, whereas αSyn aggregation occurs on a timescale of days under these conditions.*

formation of αSyn aggregation nuclei both on and off lipids. These results demonstrate that GCase can function as a molecular chaperone, reversibly binding αSyn and both inhibiting the formation and reducing the stability of amyloid formed from αSyn.

Central to the determination of this mechanism was the need to study human cell derived GCase complete with relevant post-translational modifications. The methylation methodology I employ is generally applicable for studies of biologically relevant proteins without the need to be prepared in lower organisms.

A healthy lysosome, rich in GCase, is able to degrade both monomeric αSyn and αSyn oligomers [Goker-Alpan et al., 2010, Cuervo et al., 2004, Mak et al., 2010]. In PD, depletion of GCase in the lysosome is associated with the build-up of amyloid and enhanced susceptibility to Parkinsons disease [Lee et al., 2004a, Mak et al., 2010, Cuervo et al., 2004, Murphy et al., 2014, Bultron et al., 2010, Liu et al., 2015, Gegg et al., 2012, Sidransky, 2005, Neumann et al., 2009, Mazzulli et al., 2011]. αSyn amyloidogenesis in the lysosome is enhanced by both GCase depletion, and the presence of anionic lipids [Mazzulli et al., 2011, Lee et al., 2004a, Mak et al., 2010, Cuervo et al., 2004, Murphy et al., 2014, Chiasserini et al., 2015, Yap et al., 2011]. GCase depletion also enhances the transfer of αSyn oligomers between cells by the lysosomal system [Mazzulli et al., 2011, Bae et al., 2014, Alvarez-Erviti et al., 2011]. The presence of anionic lipids and mutations that deplete the lysosomal GCase concentration are both significant risk factors in PD. The model explains these observations, predicting a shift in equilibrium favouring increased lipid-induced amyloid formation in the case of GCase depletion under lysosomal conditions. Furthermore, GCases destabilising role on amyloid fibrils explains why de-

pleted GCase leads to enhanced transfer of αSyn oligomers between lysosomes of adjacent cells [Mazzulli et al., 2011, Bae et al., 2014, Alvarez-Erviti et al., 2011].

Taken together, this quantitative analysis explains the protective effect of GCase on PD models and why its genetic and sporadic depletion results in an increased risk of PD [Schapira and Gegg, 2013, Sardi et al., 2013]. I demonstrate that GCase can function as a molecular chaperone, inhibiting αSyn amyloid formation under acidic conditions. These results provide a molecular framework to interpret the potential therapeutic benefits of GCase in Parkinsons disease.

4.4 Methods

4.4.1 Expression and purification of α-synuclein

Human sequence αSyn plasmid was kindly provided by Ad Bax (NIH). Recombinant ^{15}N enriched αSyn was expressed from BL21(DE3) star cells (Thermofisher) containing kanamycin resistant, T7, IPTG inducible plasmids (pET 41, Novagen Inc). Cells were grown in M9 minimal media; consisting of 6 g/L Na2HPO4.7H2O, 3 g/L KH2PO4, 0.5 g/L NaCl, 1 mM MgSO4, 300 μM CaCl2 supplemented with 0.5g/L $^{15}NH_4Cl$. Cells were induced at OD_{600} 0.7 with 1mM IPTG (Sigma) for 3 hours at 37°C. After induction cells were pelleted by centrifugation (Avanti J26XP, Beckman Coulter) and stored at -80°C. Cell pellets were suspended in 50 mM Tris-HCl, pH7.4, 0.5 M NaCl containing protease inhibitors (EDTA free protease inhibitor cocktail, Roche) and heated to 90°C for 10 minutes in a water bath (VWB2, VWR). Resulting lysate was clarified by centrifugation at 20,000g for 30 minutes at 4°C and diluted into

buffer A (50mM Tris-HCl, pH 7.4). Diluted lysate was loaded onto a 5 mL Hitrap Q column pre-equilibrated with buffer A. Protein was eluted with a gradient of buffer B (50mM Tris-HCl, pH 7.4., 1 M NaCl). Fractions with a high absorbance at 280 nm were assayed by SDS-gel electrophoresis, pooled, concentrated and run on a Superdex S75 size exclusion column (GE Healthcare Life Sciences). Elution of αSyn occurred at a position equivalent to a 60 kDa folded protein, which is consistent with a disordered protein. Supporting this observation is the diffusion coefficient and chemical shifts observed by NMR, both of which are consistent with monomeric disordered protein. Subsequent purity was assayed by SDS-gel electrophoresis. Pure fractions were pooled, buffer exchanged and concentrated into 50 mM citric acid, 100 mM sodium phosphate, 2 mM sodium azide, pH 5 with Amicon concentrators (Millipore), and stored at -80°C for future use.

4.4.2 Source and assignment of GCase

GCase (VPRIV) was a gift from Shire Pharmaceuticals PLC. VPRIV is an enzyme replacement therapeutic for Gauchers disease. VPRIV is produced from human cell lines supplemented with kifunensine resulting in high mannose glycans favourable for enzyme replacement therapy. See chapter 3 for details.

4.4.3 Lipid and detergent preparation

Micelles were produced as previously outlined [Galvagnion et al., 2015]. It is important to have a preparation that does not precipitate and is kinetically stable under the solution conditions investigated. Stock solutions of DMPS, POPG, SDS and taurocholate were prepared in 50 mM citric acid, 100 mM

sodium phosphate, 2 mM sodium azide (pH 5). DDM was mixed in with POPG and SDS to reduce the affinity of αSyn providing more signal intensity in the spectrum. Stock solutions were mixed at 50°C in a Thermomixer (Fisher scientific) for 4 hours, freeze thawed with a combination of a 50°C water bath (VWB2, VWR) and dry ice 5 times, before sonication for 5 minutes. Stock solutions were then immediately diluted to their working concentration. All lipid concentrations are stated where they were used in the manuscript, and all are above the critical concentration of the specified lipid/detergent.

4.4.4 GCase activity assays

GCase activity was performed in triplicate in 96-well clear flat bottom black polystyrene microplates (Product #3904, Corning) with 120 μM substrate (resorufin--D-glucopyranoside (Sigma)) and 50 nM GCase buffered in 176 mM K_2HPO_4, 10 mM sodium taurocholate, 50 mM citrate, 0.01% tween 20, pH 5.9. The final well volume was 0.1ml. Fluorescence of hydrolysed substrate was measured on a Flurostar Omega (584/630nm excitation and emission respectively) at 37°C [Motabar et al., 2012].

4.4.5 Aggregation assays

Protein aggregation was observed via thioflavin-T (ThT) fluorescence in 96 microwell plates and a Fluostar Optima plate reader (BMG Lab Technologies). Microwell plates containing αSyn were aggregated either at 200 μM, 45°C or 100 μM, 30°C pH 5 in 50 mM citric acid, 100 mM sodium phosphate, 2 mM sodium azide pH 5. An in-situ ThT binding assay with 10 μM ThT and 440/490 nm excitation/emission filters monitored aggregation. ThT

assays were repeated in triplicate yielding standard deviations for error measurements.

4.4.6 NMR experiments

NMR experiments were performed on a Varian 600 MHz spectrometer with a 5-mm z-axis gradient triple resonance room temperature probe at 30°C. Isotopically labeled ^{15}N αSyn and ^{13}C GCase were buffered in 50 mM citric acid, 100 mM sodium phosphate, 2 mM sodium azide (pH 5) and concentrated to 100 μM using Amicon concentrators (Millipore). 2D ^{15}N-^{1}H HSQC spectra were recorded with 128 scans per transient with acquisition times of 67.5 ms, over 100 complex points and a 1 s relaxation delay. 2D ^{13}C-^{1}H HMQC were recorded with 64 scans per transient and an acquisition time of 0.064 s, over 100 complex points and a 1.5 s relaxation delay.

Diffusion pulsed field gradient experiments were used to characterize diffusion behavior. The translational diffusion of specific resonances was determined using ^{1}H pulsed field gradient NMR experiments for all mixtures of αSyn, GCase and lipids/detergents described in the paper. In such experiments, signal intensity is attenuated according to the strength of an applied gradient (expressed here with respect to the maximum applied gradient strength) in a manner defined by the translational diffusion coefficient. Increased signal attenuation reveals a larger diffusion coefficient. PFG experiments were carried out over 11 gradient strengths with 256 scans per transient and an acquisition time of 0.5 s. The maximum gradient strength was 60 G cm^{-1}. The gradient duration was 0.002 s with a 0.15 s delay between gradients. NMRPipe was used to analyse spectra [T. D. Goddard and D. G. Kneller,]

and Sparky was used to pick peaks and prepare figures. Peaks were assigned by comparison to previously published assignments of αSyn [Bodner et al., 2010]. As not all peaks were assigned due to overlap, I have in total a 60% assignment of resonances that are clearly isolated, whose coverage that uniformly spans the entire sequence. Peak intensity was calculated from peak heights obtained using Sparky. C-terminal and N-terminal intensities represent the mean and standard deviation of the 20 most C or N terminal residues.

Isotope (^{15}N) enriched αSyn fibrils were pooled and diluted in 50 mM citric acid, 100 mM sodium phosphate, 2 mM sodium azide (pH 5) to a final 100 μM concentration. Stock samples were then split into two equal working samples. Labeled (^{13}CH3) GCase was added to a single sample to a final concentration of 100 μM. 2D ^{1}H^{15}N HSQC and ^{1}H^{13}C HMQC data were collected as described above.

4.4.7 Electron microscopy

Carbon coated grids (Electron Microscopy Sciences) were incubated with diluted microwell sample (10 μM) for 2 minutes and post stained with 2% aqueous uranyl acetate (10 s) and washed with water droplets. Grids were then imaged on a transmission electron microscope (FEI Tecnai 12) operated at 120kV using a digital CCD camera (Ultrascan US1000, Gatan).

4.4.8 Thermodynamic modelling

The reduction in signal intensity was recorded for the first and last 20 residues of αSyn, together with GCase as a function of changes in concentration of the three components in the system, αSyn, GCase and lipids. The changes in signal

intensity were concentration dependent and so I sought an equilibrium-binding model that accounted for all data. The total concentrations were varied of each such that I had 7 measurements of both GCase, N and C-terminal αSyn signal, providing 21 intensity datapoints for each of the anionic lipids and detergents DMPS, SDS/DDM and POPG/DDM. There are two challenges to account for these observations. The first is to identify which species are forming, and the second is to account for how the formation of these species affects the NMR signal intensity. I sought to determine the model that gives the best fit to the data, subject to the restraint that improved fitting from more complex models is tested for statistical significance by applying an F-test and determining the probability that this improvement came by chance. A <5% probability level was taken as the condition for statistical significance.

The problem involves three species, lipid micelles, αSyn and GCase. From the three components, there are 5 binary combinations that can be formed. Should I choose to distinguish independent sites for binding on the N and C termini, the formation of each of these can be characterized by the following association equilibrium constants and definitions:

$$
\begin{aligned}
[Ls-] &= K_{ln}[-s-][L] \\
[-sL] &= K_{lc}[-s-][L] \\
[GL] &= K_{\mathrm{lg}}[L][G] \\
[-sG] &= K_{gc}[-s-][G] \\
[Gs-] &= K_{gn}[-s-][G]
\end{aligned}
$$

Where the representation -S- for αSyn denotes either an empty N/C termini. Replacing one of the symbols with either an L (lipid) or G (GCase) implies occupancy. There are 4 corresponding ternary complexes, which can

be similarly defined in terms relevant equilibrium constants:

$$[LsG] = K_{n\mathrm{lg}}[Ls-][G]$$
$$[GsL] = K_{c\mathrm{lg}}[-sL][G]$$
$$[GsG] = K_{sgg}([-sG]+[Gs-])[G]$$
$$[LsL] = K_{sll}([-sL]+[Ls-])[L]$$

The equilibria are restrained by total numbers of molecules through the following three mass balancing equations:

$$S_{tot} = [-s-]+[Ls-]+[-sL]+[Gs-]+[-sG]+[LsG]+[GsL]$$
$$G_{tot} = [G]+[Gs-]+[-sG]+[GL]+[GsL]+[LsG]+2[GsG]$$
$$L_{tot} = [L]+[Ls-]+[-sL]+[GL]+[GsL]+[LsG]+2[LsL]$$

Where S_{tot}, G_{tot} and L_{tot} are the total concentrations of αSyn, GCase and lipids respectively. For known values of the total concentrations, and trial values of the relevant equilibrium constants, I can self-consistently solve these equations to obtain the corresponding mole fractions of the individual species. By optimising the equilibrium constants against experimental measures of mole fractions, I can obtain their most probable values.

To do this, it is necessary to obtain expressions for the observed signal intensity. There are a number of ways in which this can be formulated depending on the precise details of timescales of the interactions. The timescale together with the specific structural changes that accompany the association will affect the signal intensities through chemical exchange. As described in the text, as I observe neither appreciable line broadening nor change in the chemical shift of resonances, I expect slow or slow-intermediate exchange. Moreover, no variation in the R^2 of the residues was observed with 180° pulse frequency in CPMG experiments, suggesting no species interconversion on the milli-second exchange. If a species rigidly tumbles like either GCase or micelle, I would not

expect it to contribute directly to the solution NMR spectra. Nevertheless, if part of αSyn retains some of its conformational flexibility when bound, I would expect it to contribute to the spectrum, although with a reduced signal intensity. To account for this type of interaction, and to first order to account for the effects of chemical exchange, the relaxation correction factors γ are introduced such that I can express the observed signal intensities as:

$$\left(\frac{I}{I_0}\right)_{\alpha Syn,N} = \frac{[-s-]+[-sL]\gamma_1+[-sG]\gamma_2}{S_{tot}}$$

$$\left(\frac{I}{I_0}\right)_{\alpha Syn,C} = \frac{[-s-]+[Ls-]\gamma_3+[Gs-]\gamma_4}{S_{tot}}$$

$$\left(\frac{I}{I_0}\right)_{GCase} = \frac{[G]+[-sG]\gamma_5+[Gs-]\gamma_6}{G_{tot}}$$

The most complex model described by this picture is parameterized by 19 values (9 equilibrium constants including the ternary species and 10 attenuation constants). This model was found to provide the best fit to the data. But as I have 21 individual signal intensity measurements, to apply this model carries a strong risk of over-fitting. I tested a number of models both including and excluding ternary complexes, and with varying numbers of species included that are proposed to be NMR observable (non zero γ factors). Following the principles of sensible fitting, I sought to determine the minimally complex model where the reduction in $\chi 2$ is to more complex models is determine to be statistically insignificant by application of an F-test. The $\chi 2$ was defined as:

$$\chi^2 = \sum_{concentrations} \sum_{\substack{\alpha Syn,N \\ \alpha Syn,C \\ GCase}} \left(\left(\frac{I}{I_0}\right)_{calculated} - \left(\frac{I}{I_0}\right)_{measured}\right)^2$$

Such testing revealed that no significant improvement was obtained if I incorporate 1) lipids binding the C-terminus of αSyn, 2) GCase binding the N-terminus of GCase, nor 3) any of the ternary complexes. Moreover, there was no statistical benefit to including observation of GCase signal when it is bound to lipids. This results in the following simplified model that I conclude gives the best statistical fit to the data, with three equilibrium constants:

$$[Ls-] = K_{ln}[-s-][L]$$
$$[GL] = K_{lg}[L][G]$$
$$[-sG] = K_{gc}[-s-][G]$$

And three attenuation exchange factors whose origin comes from chemical exchange (γ):

$$\left(\frac{I}{I_0}\right)_{\alpha Syn,N} = \frac{[-s-]+[-sG]\gamma_{\alpha SynN}}{S_{tot}}$$
$$\left(\frac{I}{I_0}\right)_{\alpha Syn,C} = \frac{[-s-]+[Ls-]\gamma_3}{S_{tot}}$$
$$\left(\frac{I}{I_0}\right)_{GCase} = \frac{[G]+[-sG]\gamma_{GCase}}{G_{tot}}$$

This model (figure 4.11 model 3) is not significantly improved over a model where I set $\gamma 3$ to 1, which amounts to the N-terminus of αSyn contributing to the observed signal intensity equally to free αSyn. Overall this model (figure 4.11 model 4) is parametrized by 5 constants and predicts the existence of 3 binary complexes in addition to the 3 free species. This model gives an excellent description of all signal intensities recorded over the wide range of combinations of concentrations tested. Incorporating additional parameters did not result in a statistically significant improvement in the fitting. The formulation of the signal intensities successfully account for the observed increase in diffusion

coefficients (figure 4.6) as it predicts bound species contribute the spectrum albeit with reduced signal intensity when compared to the free species. The code to perform this analysis was written in python, where it was determined that a simulated annealing approach was necessary to find the global optimum. Uncertainties in fitting parameters were derived from a bootstrapping procedure, with on average a 30% error in each parameter. The uncertainty is relatively high. The trends in these parameters however span orders of magnitude (figure 4.12) and so I obtain meaningful mechanistic insight with this approach. The high uncertainties also imply that incorporation of additional information into the model, for example a more detailed description of chemical exchange parameterized with additional data could further refine the model. The optimal model does not include ternary complexes. I note that this does not mean ternary complexes cannot form, just that I see no statistically significant reason in the data to justify their inclusion to quantitatively explain the data, suggesting if they are formed they are highly unstable.

Chapter 5

Small heat shock protein 16.5 - chaperones in aggregation

Abstract

Small heat shock protein 16.5 is an archaeum derived cellular chaperone involved in protein homeostasis. Under physiological conditions, 16.5 assembles into large 24-mer oligomers. Here, I report the use of a truncated 16.5 construct, capable only of forming dimers, for NMR assignment and chaperone function studies. Using NMR on both the 24mer and the dimer I reveal an extensive structural conservation. Furthermore it is revealed that dimeric 16.5 is capable of effective chaperoning, suggesting that HSP16.5 dimers are sufficient to encode both the structure and function of the oligomer. By following aggregation by NMR I suggest truncated 16.5 to interact preferentially with misfolded protein oligomers that appear at early times. Finally, the extensive overlap between dimeric and oligomeric HSP16.5 was exploited to assign the oligomer and investigate slow and fast scale dynamics of a 400 kDa protein consistent with subunit exchange.

5.1 Molecular chaperones: an introduction

Long lived and unstable proteins will naturally sample conformers that are distinct to their native state (figure 1.1). Whilst many of these conformers are important for protein function (see section 2.2), some may be on-pathway to aggregation (chapter 4) or non-functional conformers ('misfolded', chapter 3) in need of rescue.

The responsibility for rescuing proteins from such pathological misfolding events falls to a network of cellular systems including the proteasome and molecular chaperones, which maintain the integrity of the cell's proteome. The importance of chaperones is observed in the cases where familial mutation predisposes patients to a wide range of diseases, from cataracts to myocardial infarction, depending on which chaperone is mutated [Bakthisaran et al., 2015].

Interestingly, whilst the human genome consists of approximately 25,000 genes, only a few hundred of these encode chaperones. Chaperones must therefore be promiscuous proteins capable of sensing when their substrate adopts non-native folds. The basis for this sensing is thought to be exposed hydrophobic regions or newly disordered regions, however little is known about the precise mechanisms of chaperoning[Powers et al., 2009].

A particular class of molecular chaperones, known misleadingly as the small heat shock proteins(small HSPs), display a remarkable ability to oligomerize into large quaternary structures. These ATP-independent chaperones have subunit masses of 10-40 kDa, but, when assembled, can form complexes exceeding 1 MDa (figure 5.1) [Hilton et al., 2013]. Moreover they are found in all domains of life, from archaea to metazoa [Franck et al., 2004].

Structurally, the small HSPs consist of a highly conserved central domain,

measuring 100 residues in length and termed the α-crystallin domain, flanked by more variable N and C termini thought to play key roles in the stabilization of larger oligomers [Franck et al., 2004, Kappé et al., 2003, Baldwin et al., 2011a].

The small HSPs show a remarkable ability to inhibit protein aggregation, including amyloidogenesis (chapter 1). Specifically, human αB-cystallin and HSP20 inhibit the amyloidogenesis of Alzheimer's disease Aβ peptides and their associated cytotoxicity [Wilhelmus et al., 2006, Hochberg et al., 2014]. Likewise, αB-cystallin is reported to inhibit α-synuclein aggregation and to bind to and destabilize mature α-synuclein fibrils [Waudby et al., 2010, Rekas et al., 2004]. Most interestingly, a synergistic relationship between the ATP-independent small HSPs and the ATP-dependent HSPs has been observed, where small HSPs potentiate the disaggregation of α-synuclein aggregates [Duennwald et al., 2012].

The conservation of structure amongst small HSPs suggests that studies on homologous proteins from other domains of life may elucidate human HSP function. HSP16.5, derived from *Methanococcus jannaschii* - a thermophillic archaeum, is able to effectively inhibit amyloidogenesis [Xi et al., 2011] and forms large (24mer) homooligomers from 16 kDa subunits.

Here, I report high quality NMR spectra of both oligmeric (24mer) HSP16.5 and a truncated construct capable of only forming dimers. Using the smaller construct I am able to assign several resonances on the oligomer and examine their dynamics. Furthermore, the structural similarity of dimeric HSP16.5 to oligomeric 16.5 suggests a common building block, whilst the ability of the truncated core domain to effectively chaperone model substrate α-lactalbumin (α-lac) suggests that dimer also encodes the function of the oligmer. Finally,

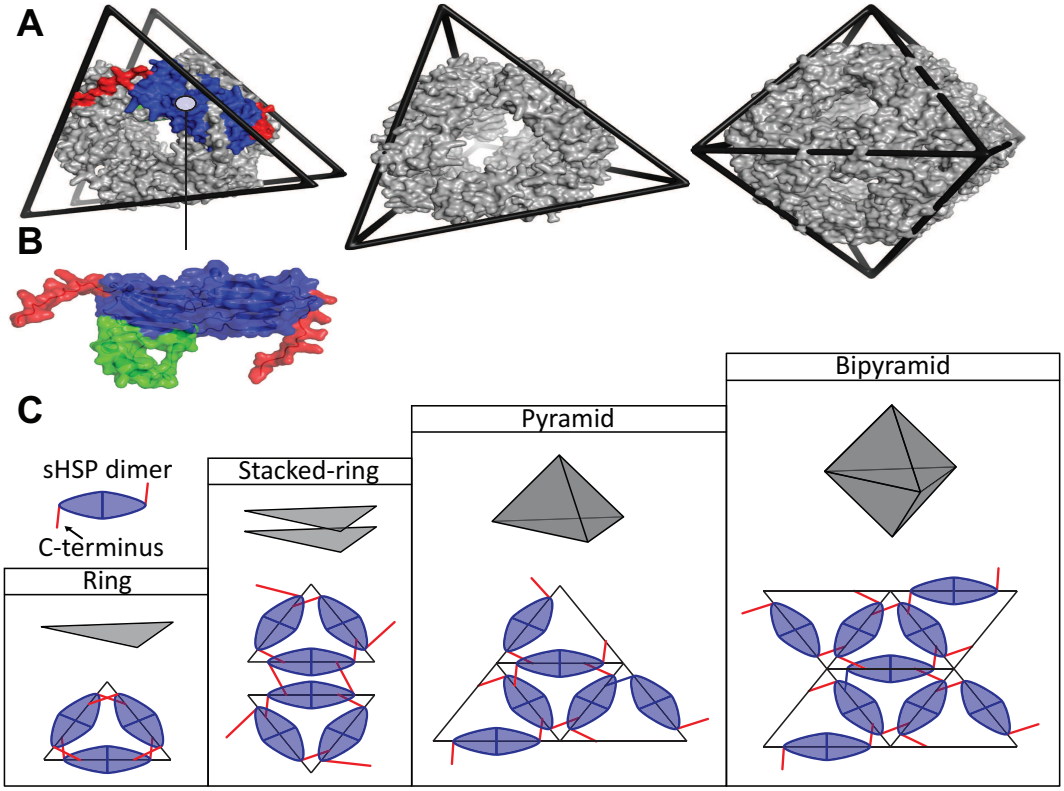


Figure 5.1: *HSPs form large oligomers. From left to right, HSP16.9 (12mer - stacked ring), Acr1 (12mer - pyramid) and HSP16.5 (24mer - bipyramid). Schematics (A) dimers (B) and organisation (C) are shown for each example. Image adapted from [Hilton et al., 2013]*

NMR timecourse studies of aggregation suggest HSP16.5 to preferentially interact with substrate oligomers formed at early time scales on pathway to aggregation.

5.2 Results: HSP16.5 - a model chaperone

Human small HSPs are implicated in a multitude of pathologies [Hilton et al., 2013]. However, the most notable member of this group, αB-crystallin is difficult to observe by NMR [Baldwin et al., 2011a] due to the sparsity of observable peaks from the oligomer. A model chaperone should ideally yield high quality NMR data for peak assignment, dynamics and functional studies.

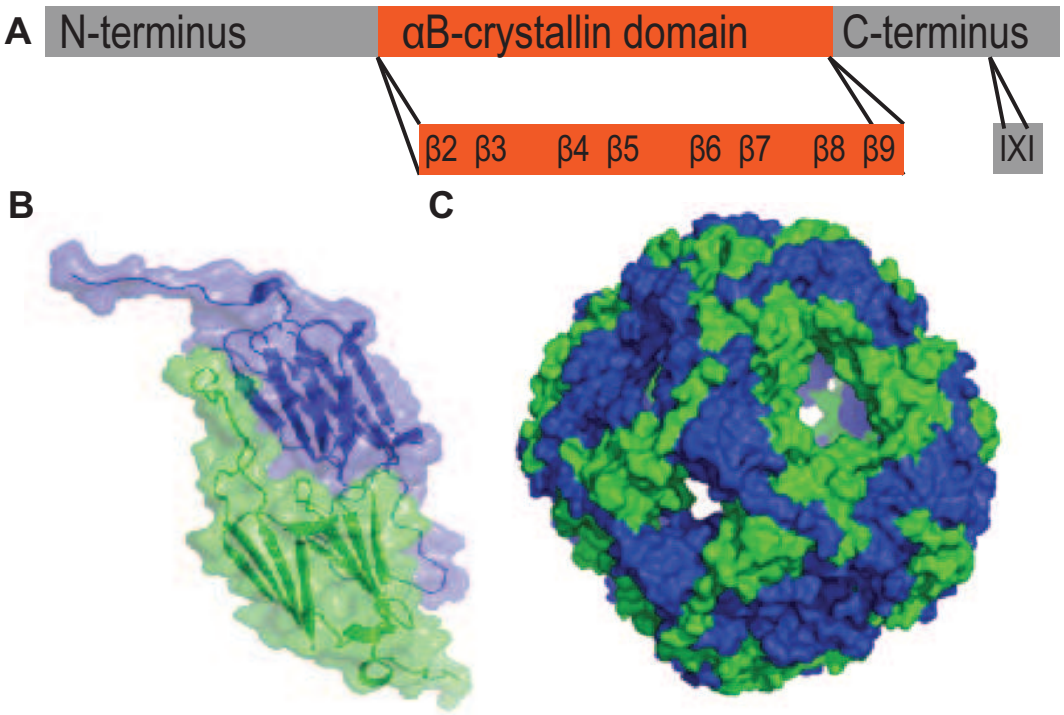


Figure 5.2: *HSP16.5 sequence and structure. A) The small HSP family proteins are organized into three domains, a central α-crystallin domain (red) flanked by N and C terminal domains (gray). The central domain contains a number of βsheets (β2-9 shown), whilst the C-termini contains a palindromic IXI sequence. B) Structure of HSP16.5 dimer with monomers colored blue and green, the 'core domain'(PDB 1SHS) C) structure of oligomeric HSP16.5 (PDB 1SHS). D) Sequence of HSP16.5 with the α-crystallin domain in red, and the core domain underlined. Note that the core domain contains an additional N-terminal glycine and serine from the TEV tag*

To examine the suitability of HSP16.5 for NMR studies, isotopically enriched protein was produced from *E. coli* and minimal media enriched in ^{13}C precursors of methionine, isoleucine, leucine and valine [Sprangers and Kay, 2007, Ruschak and Kay, 2010, Baldwin et al., 2009, Goto et al., 1999, Tugarinov et al., 2006]. Additionally, a uniformly labelled ^{13}C ^{15}N ^{1}H HSP16.5 'core' construct with truncated N and C termini was produced (sequence SGMFGRDPFDSL FERMFKEFFA TPMTGTTMIQ SSTGIQISGK GFMPISIIEG DQHIKVIAWL PGVNKEDIIL NAVGDTLEIR AKRSPLMITE SERIIYSEIP EEEEIYRTIK LPATVKEENA SAKFENGVLS VILPKAESSI KKGINIE). Like other truncated HSPs, this core domain is only able to form dimers (figure 5.2) [Hilton et al., 2013].

Examination of the oligomer NMR spectra reveal that 55 $^{13}C^{1}H_3$ resonances can be observed for the oligomer spectra (figure 5.3). Given the large size of this oligomer (>400 kDa), it is unexpected to observe so many well dispersed peaks.

However, by employing methyl-TROSY NMR, alongside selective amino acid labeling [Sprangers and Kay, 2007], we are able to exploit countervailing relaxation effects to reduce spectral crowding and peak breadth. Furthermore, the thermophillic nature of HSP16.5 allows spectra to be recorded at temperatures exceeding 80°C, where molecular motion is rapid and relaxation minimized. Together, these effects yield high quality NMR spectra.

The substantial overlap in the spectra of the core and oligomer suggest a building block common to both constructs. Chemical shift is an exquisite reporter of local environment and so a similarity of spectra suggests a similarity of structure (figure 5.3). I therefore conclude that core and oligomer HSP16.5 share a common structural unit, namely, the dimer.

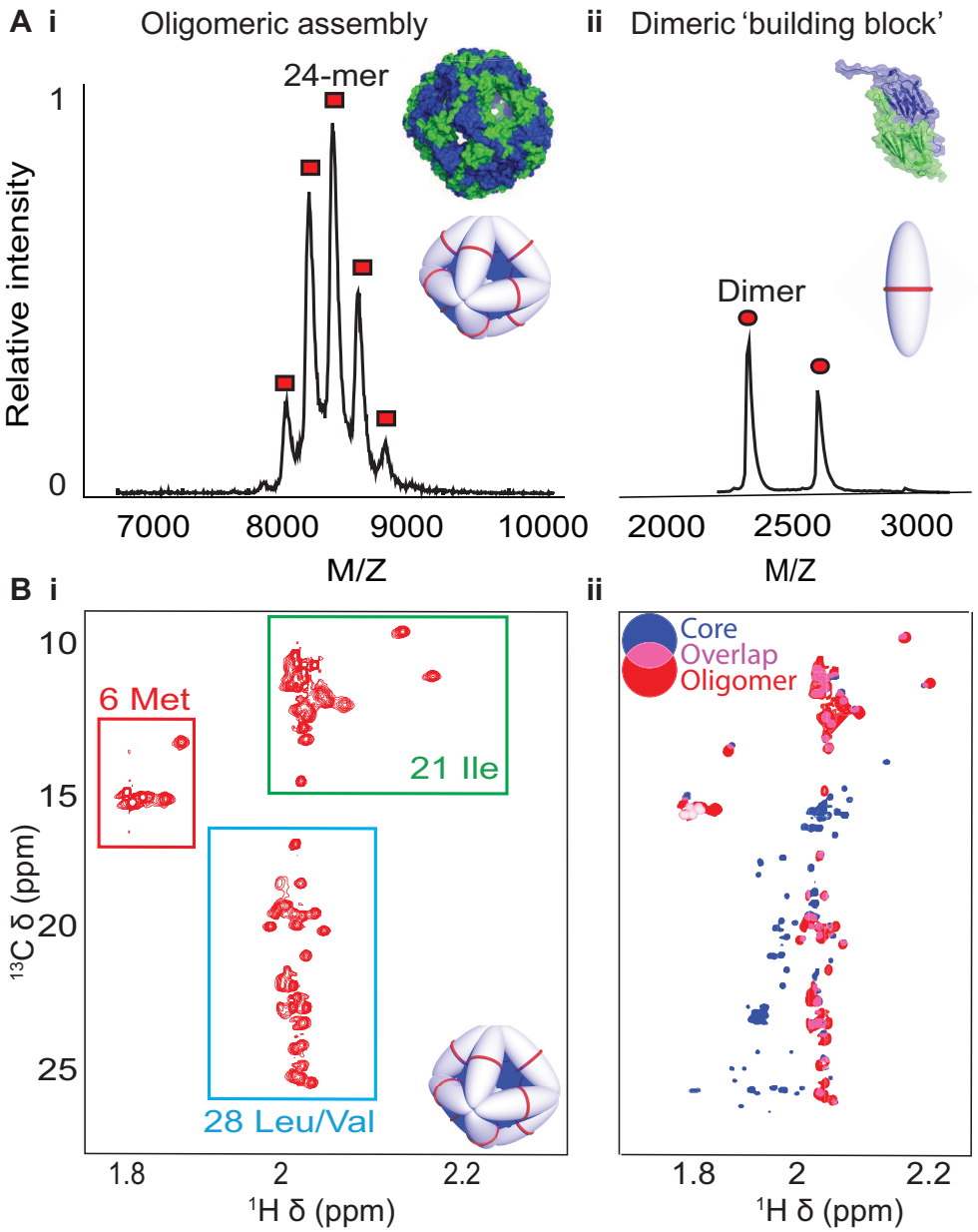


Figure 5.3: *HSP16.5. A) nESI MS (10 μM, 21°C, pH 7.4) spectra of oligomer (i) and core (ii). Oligomeric mass corresponds to a 24mer. B i) ^{13}C ^{1}H spectra (800 μM (monomeric equivalent), 60°C, pH 7.4) of oligomer, using $^{13}CH_3$ TROSY NMR. ii) Overlay of oligomer methyl TROSY spectra(red) and core ^{13}C ^{1}H HMQC (blue 800 μM, 60°C, pH 7.4). Note the overlap where MILV side chains appear in both spectra (pink - numbering approximately 44 resonances)*

The high quality spectra produced provide two possibilities; first, using the extensive overlap between core and oligomeric ^{13}C spectra to assign oligomeric 16.5 spectra, thus assigning a 400 kDa protein for NMR studies. Secondly, the extensive overlap in structure suggests that core HSP16.5 may also be able to effectively chaperone, thus encoding structure and function of the oligomer.

5.2.1 Assigning HSP16.5

Canonical NMR methods for assignment require the transfer of magnetization over several chemical bonds and are generally only applicable to smaller (>30 kDa) proteins. This is due to the enhanced relaxation associated with the longer correlation times of larger proteins (see section 2.2.2 and figure 2.5). The oligomeric 16.5, with a molecular weight of approximately 400 kDa, is much too large for these methods. However, the core construct is well within this size limit; the monomer core is only 12.4 kDa (truncated from 16.5 kDa) and dimer only 24.8 kDa.

Canonical NMR assignment experiments involve transferring magnetization along the polypeptide backbone and side chains, recording H,C and N data (figure 5.4). The thermophillic nature of HSP16.5 facilitated assignment by enabling high temperature (60°C) data acquisition; increasing molecular tumbling to yield intense resonances (figure 5.5, section 2.2.2 for further information).

A total of 93 well dispersed backbone $^{15}N^{1}H$ resonances can be observed in the HSQC, this compares with the 107 theoretically available (i.e. excluding prolines) NHs. Of the 107 residues 75 have been assigned (representing 70%) (figure 5.6). As expected for an NMR experiment, the signal to noise is in-

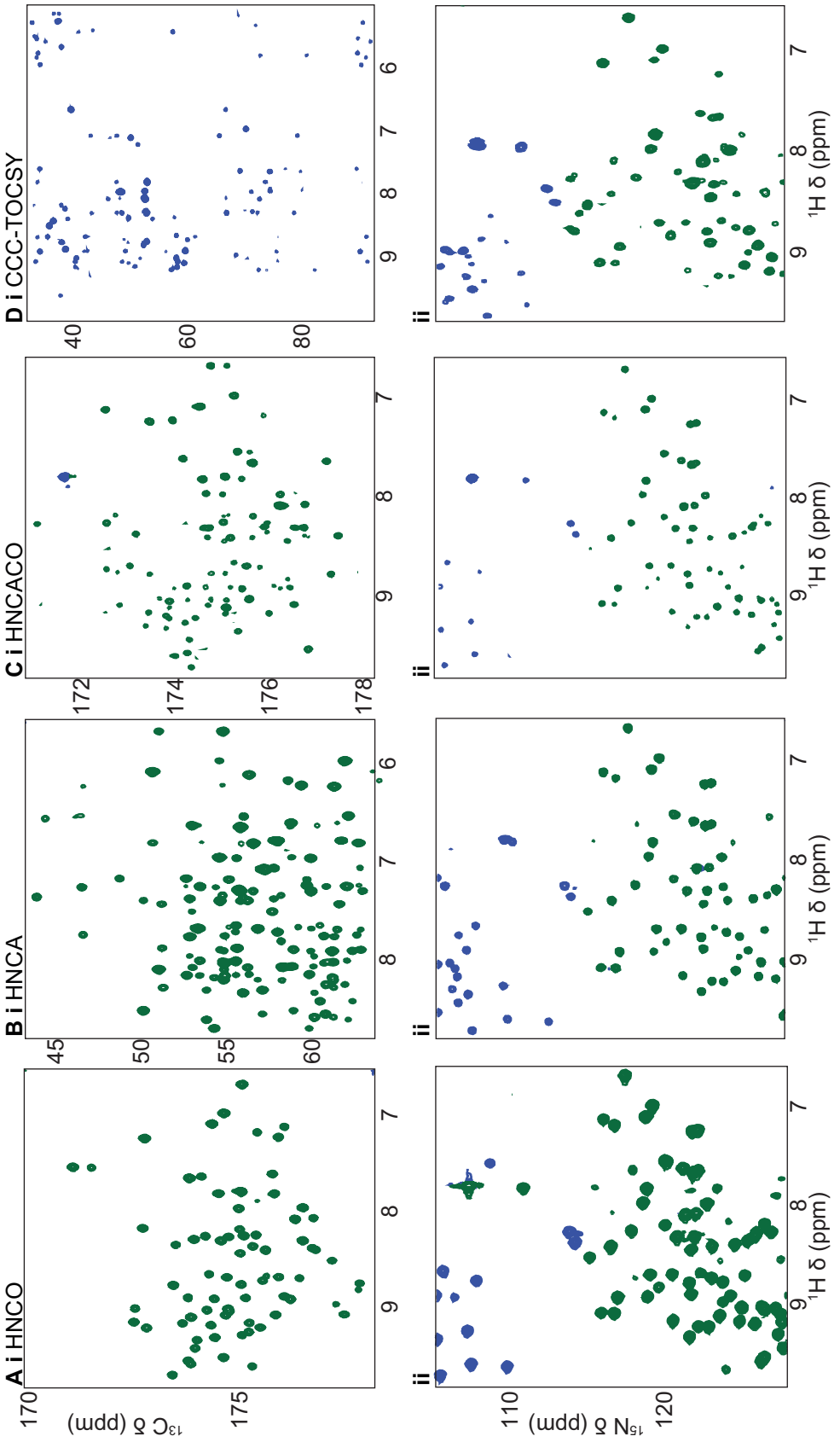


Figure 5.5: *Assignment of HSP16.5 core construct, 800 µM, 60°C, pH 7.4. 2D projections, either $^{1}H^{13}C$ (i) or $^{1}H^{15}N$ (ii), for each 3D experiment are shown. Blue resonances are folded in the spectra*

versely proportional to the number of nuclei recorded and the magnetization transfer steps (figure 5.6B), likely due to increased relaxation effects (section 2.2.2).

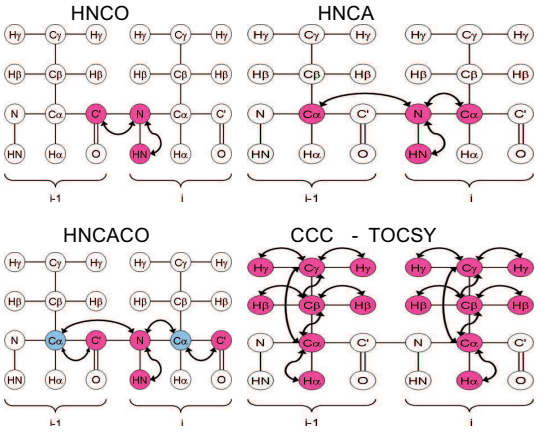


Figure 5.4: *Assignment of HSP16.5 core construct. Assignment experiments employed here; arrows represent transfer of magnetization, pink circles represent the acquisition of chemical shift data.*

It was reasoned, given the substantial overlap between oligomer and dimer HSP16.5 NMR resonances, that assigned peaks could be transferred from the dimer to the oligomer, thus assigning a 400 kDa complex for NMR studies. To this end, CCC-TOCSY (figure 5.4) data were acquired on the dimeric construct to assign core domain side chains. I successfully assigned 23 MILV side chains and obtained 9 resonances that correspond to Cγ on methionine, isoleucine, leucine and valine residues. These ^{13}Cγ resonances were then used to assign 6 of the HSP16.5 oligomer ^{1}H^{13}C resonances (figure 5.7).

5.2.2 HSP16.5 reveals slow (ms) and fast(ns) time scale dynamics

Dynamics studies of the HSP16.5 oligomer revealed substantial CPMG dispersions for several residues as well as some differences in R^2 rates (figure 5.8). Such dynamics studies report on millisecond (CPMG) and nanosecond dynamics events (section 2.2.2). Fitting of CPMG data using the Catia [Hansen

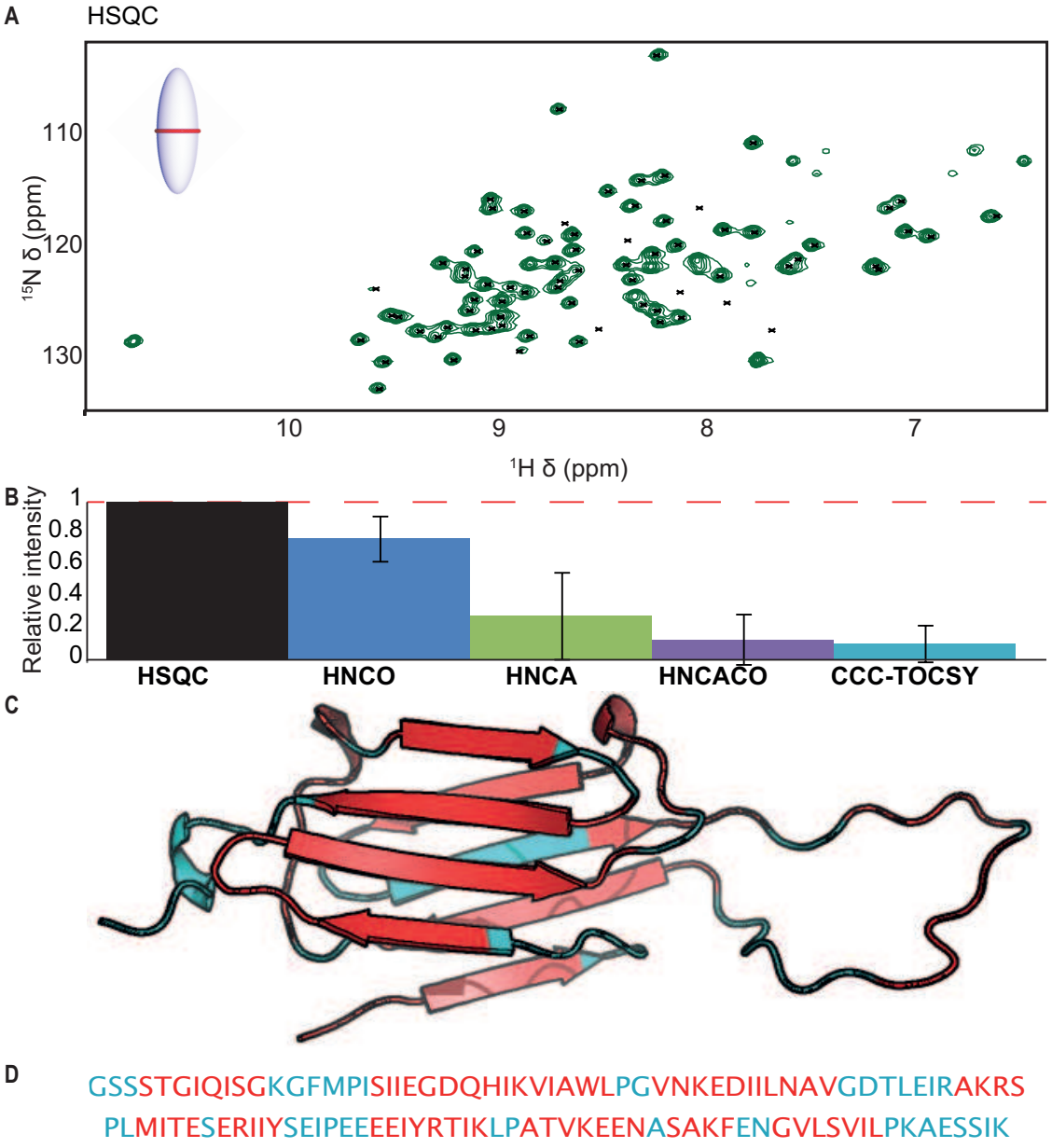


Figure 5.6: *Assignment of HSP16.5 core construct. A) $^{1}H^{15}N$ HSQC (800 μM, 60°C, pH 7.4) of 16.5 core domain, assigned resonances are indicated with a black cross (**x**). B) Bar chart displaying the average intensity of each peak when normalised by transients, baseline noise and compared to HSQC (=1), error bars represent standard deviation across all resonances C) HSP16.5 colored by assigned (red) and unassigned (blue) residues (PDB 1SHS) D) Sequence of HSP16.5 core construct colored as C. There are a total of 113 residues, of which 6 are prolines. Excluding prolines, HSP16.5 was 70% assigned*

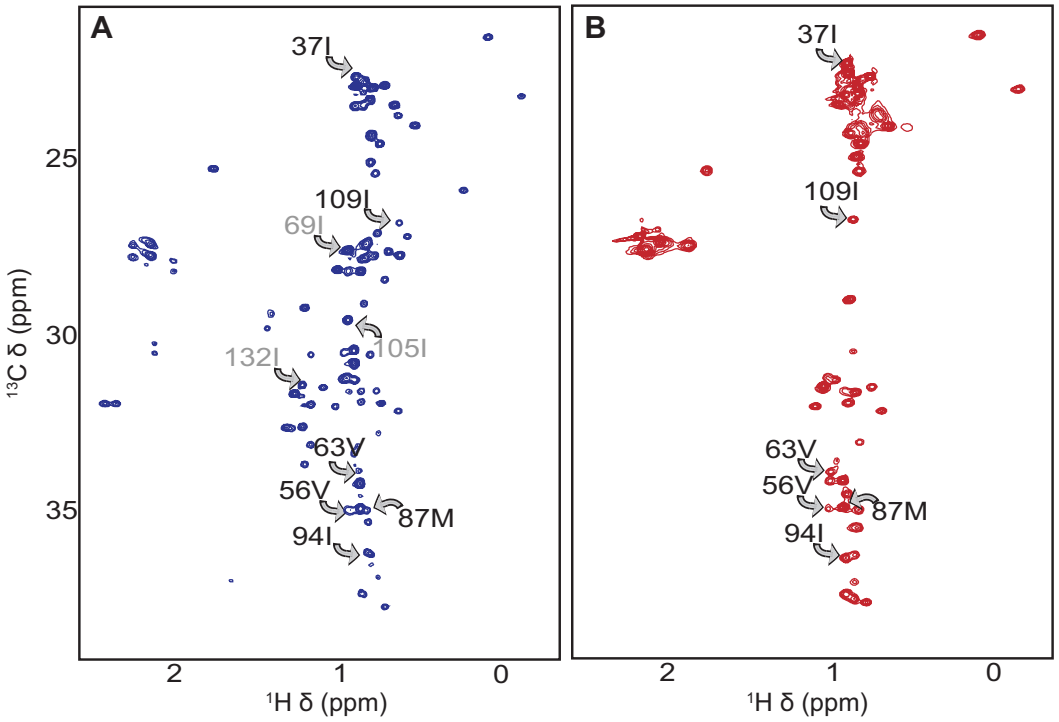


Figure 5.7: *$^{1}H^{13}C$ spectra of core (A) and oligomer (B) (800 μM, 21°C, pH 7.4). A number of assignments from the core could be transferred to the oligomer (black) whilst some remained ambiguous (light grey).*

et al., 2008] software at 4 fields yielded an exchange rate of 900±135 s^{-1} and a minor populated state of 1.6±0.69%.

By transferring the assignments from the core to the oligomer it may be possible to elucidate the structural role of these dynamics. Examination of CPMG dispersion curves for 109I reveal that dispersions occur along the dimer interface (figure 5.9).

The intersubunit contact observed in the crystal structure (figure 5.9B) is likely the CPMG observable event, as previously noted, CPMG experiments probe timescales consistent with large scale molecular motion such as fold rearrangement or oligomerization events (section 2.2.1.1 and 3.2.3). Speculatively; the formation and breaking of intersubunit contacts is being observed along the dimer interface of a 400 kDa assembly by CPMG.

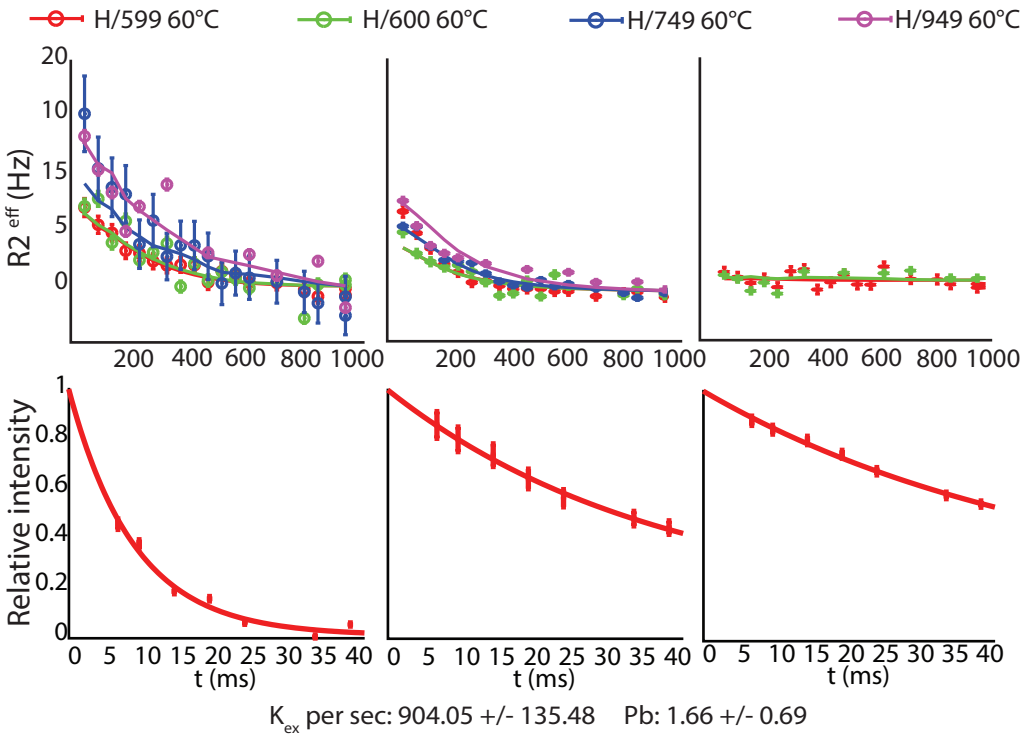


Figure 5.8: *HSP16.5 oligomer displays a range of dynamics. R^2 experiments (bottom panel) report on ns-µs dynamics whilst CPMG (top panel) report on ms dyanmics. Selected unassigned residues are shown to illustrate the absence (left) and presence (right of dynamics). CPMG fits yielded a K_{ex} of 904±135 and a Pb of 1.66±0.69% (With Andrew Baldwin)*

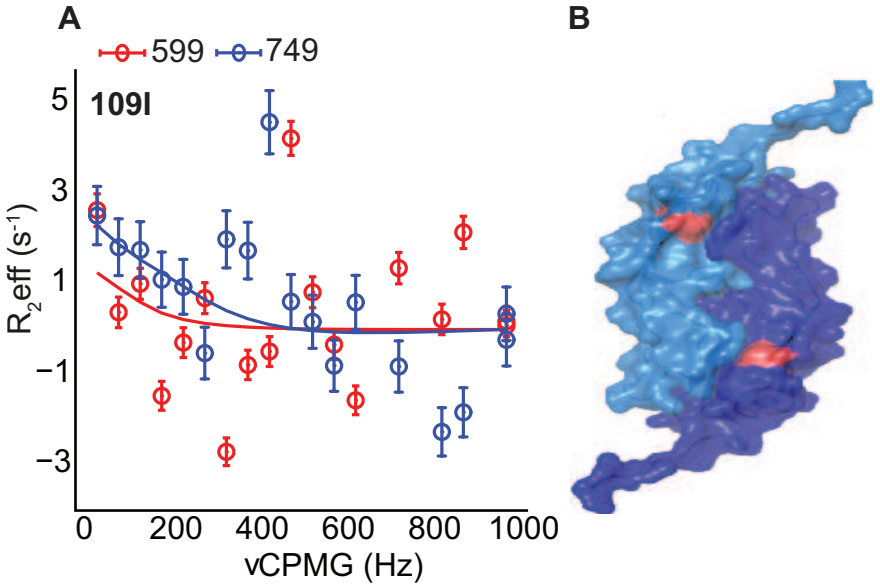


Figure 5.9: *A) CPMG data of the assigned residue 109I at 599 MHz and 749 MHz field strengths. B) 109I (red) lies along the dimer interface, speculatively, the CPMG observable dynamics may be related to monomer-dimer dynamics (two-shade blue) (PDB 1SHS) (With Andrew Baldwin).*

5.2.3 HSP16.5 and α-lactalbumin aggregation

The ability to obtain high quality NMR data of a small heat shock protein suggested HSP16.5 as a model chaperone for aggregation assays. Probing the chaperoning ability of 16.5 core domain may provide insight into small HSPs in human health and disease and elucidate whether the dimer alone encodes some chaperoning ability.

To complement the idealized aggregation/chaperoning system α-lactalbumin (α-lac) was employed as a model substrate. Aggregation of milk α-lac can be induced through the addition of DTT, furthermore the rate of aggregation can be fine tuned with DTT concentration [Bumagina et al., 2010, Bettelheim et al., 1999], providing a reliable, easily accessible model substrate. I sought to explore the interactions between substrate and chaperone through light scattering and NMR experiments.

Light scattering provides information on the total number of large aggregates. Investigation of 16.5 core at a 1:1 monomeric molar ratio with 200μM α-lac by light scattering revealed HSP16.5 core to be a potent molecular chaperone at equimolar concentrations. Interestingly, this suggests that the minimal construct, limited to dimers, encodes both structural (figure 5.3) and functional properties of the oligomer.

Examination of α-lac aggregates by electron microscopy revealed short, tubular aggregates (figure 5.10B), as opposed to the long, straight and unbranched amyloid aggregates observed in previous chapters (chapter 4).

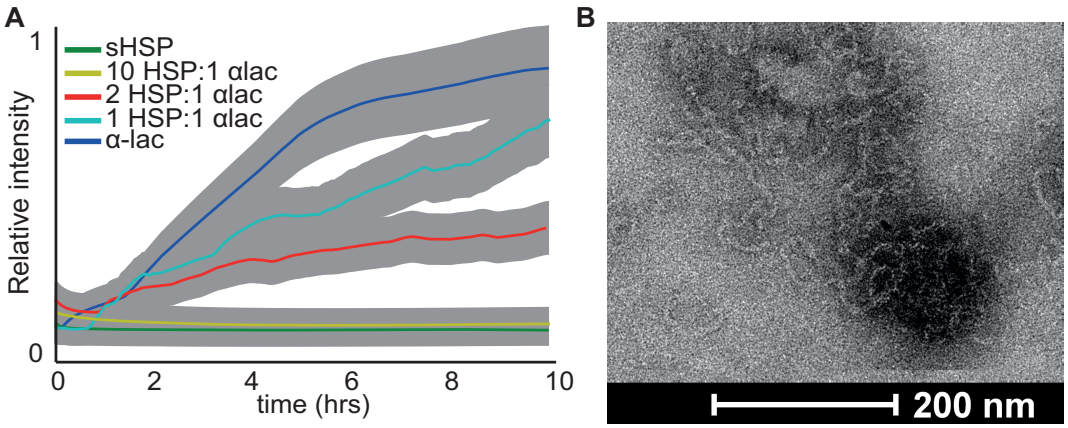


Figure 5.10: *The aggregation of α-lac. A) Light scattering timecourse of 200μM α-lac with variable ratio of truncated HSP16.5 (monomeric concentration). B) EM micrograph at end timepoint for α-lac (dark blue), there were no clear differences in EM micrographs between samples.*

5.2.4 Following aggregation by NMR

The aggregation of α-lac is clearly perturbed by HSP16.5 core (figure 5.10). To gain insight into specific mechanisms and interactions linking these proteins I employed solution NMR. Whilst light scattering experiments probe the average size of aggregates within the reaction, NMR has the potential to observe the specific contributions to signals from each component. By deconvoluting the signal from chaperone and substrate it is possible to observe aggregation in detail.

NMR signal is dependent on molecular size (section 2.2.2), therefore, as molecules bind or aggregate their overall contribution to signal should decrease. With this in mind, NMR reports on the loss of free monomer from solution.

In the spectra presented here, ^{1}H signals report on chaperone and substrate behavior, whilst ^{15}N spectra report only on isotopically enriched ^{15}N HSP16.5 core construct. Sections of 1D spectra were integrated from a time series, normalized to the first time point and plotted to report on time dependent intensity changes (figures 5.11, 5.12, 5.13, 5.14). Inspection of α-lac aggrega-

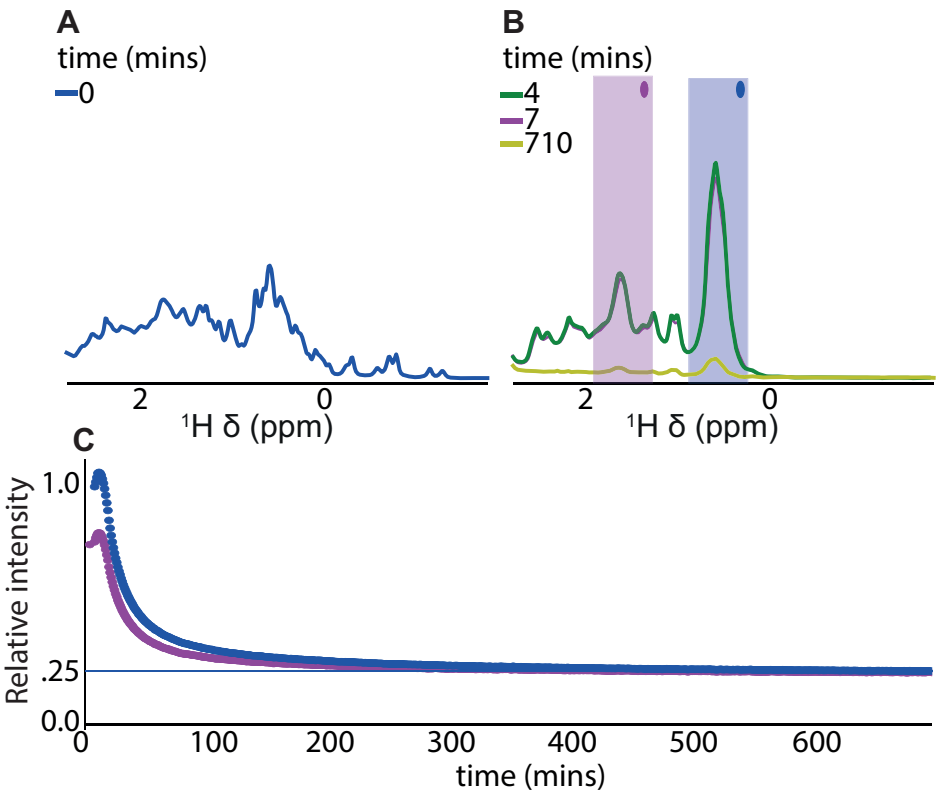


Figure 5.11: *α-lac aggregation. A) ^1H watergate spectrum of folded (no DTT -100 μM, 30°C, pH 7.4) α-lac B) selected ^1H watergate spectra at selected time points (with DTT) (100 μM, 30°C, pH 7.4). C) Change in intensity of α-lac after DTT addition, points colored as B. Note that the spectra in B (after DTT) are more intense than those in A (folded protein), and that spectra in B (and C) are more intense at early times. Data from O. Tkachenko*

tion alone reveals a decrease of signal intensity over time, consistent with the loss of monomer as α-lac aggregates into larger species, however, the signal increase observed at early time scales (figures 5.11C;0-15 minutes) may suggest an unfolding step or domain rearrangement associated with α-lac aggregation. Few chemical shift dispersions are observed after the initial unfolding step.

The addition of chaperone immediately after the induction of α-lac aggregation (t=0) revealed attenuated signal loss over time when compared to α-lac alone. Chaperone action can be deduced from the relative intensities in figures 5.11 and 5.12; in the latter figure (panel C) ^{1}H intensity plateaus at 0.4, whilst the former plateaus at 0.15, less than 5% of this change can be attributed to convolved chaperone signal (5.12Cii). Further studies using dynamic light scattering may be able to ascertain aggregate size distributions.

This attenuation confirmed that HSP16.5 acts as a molecular chaperone targeting α-lac, and that this chaperoning is observable by NMR. Examination of the ^{15}N edited ^{1}H spectra, which reports only on chaperone, revealed a significant drop in intensity in the first 50 minutes, followed by a slight recovery. This initial signal loss followed by recovery suggested that 16.5 may interact preferentially with small oligomers, I sought to test this hypothesis.

The aggregation of α-lac is induced by the addition of DTT (section 5.2.3). To examine the effect of oligomer size on chaperone signal intensity (and therefore interaction) the addition of chaperone was delayed to either 90 or 960 minutes post DDT addition. At 90 minutes post induction, α-lac aggregation is towards the bottom of the exponential decline, whilst at 960 minutes, α-lac signal is stable and the reaction has reached equilibrium (figure 5.11).

Addition of chaperone at t=90 minutes revealed a short decline in ^{1}H signal, followed by stabilization. Examination of ^{15}N signal, revealed a smaller drop in

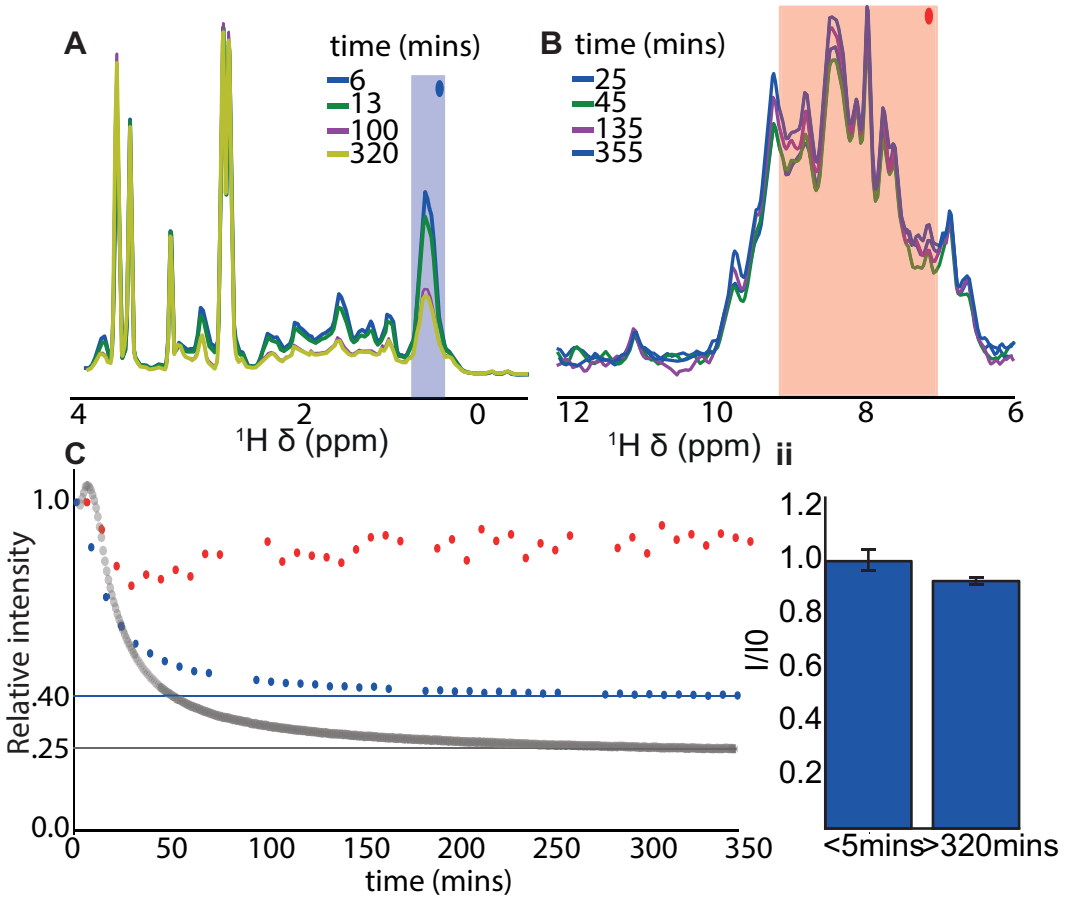


Figure 5.12: *Truncated HSP16.5 and α-lac aggregation assay. HSP16.5 was added at t=0 minutes and the reaction followed by NMR. A) Selected* ^{1}H *HSQC spectra, B) Selected 1D* $^{1}H^{15}N$ *HSQC spectra (reporting on chaperone only), C) intensity plots for chaperone and* ^{1}H *signal over time red points correspond to chaperone and blue points to mixed chaperone/substrate signal, grey points represent data from figure 5.11. Horizontal lines at 0.4 and 0.25 represent approximate equilibrium with and without chaperone, respectively. Points represent a single experiment and therefore error calculations are needed. Cii) quantification of chaperone intensity at early and late times. The insignificant drop in chaperone intensity over long times (Cii) suggests that changes in* ^{1}H *signal intensity are predominantly due to αlac*

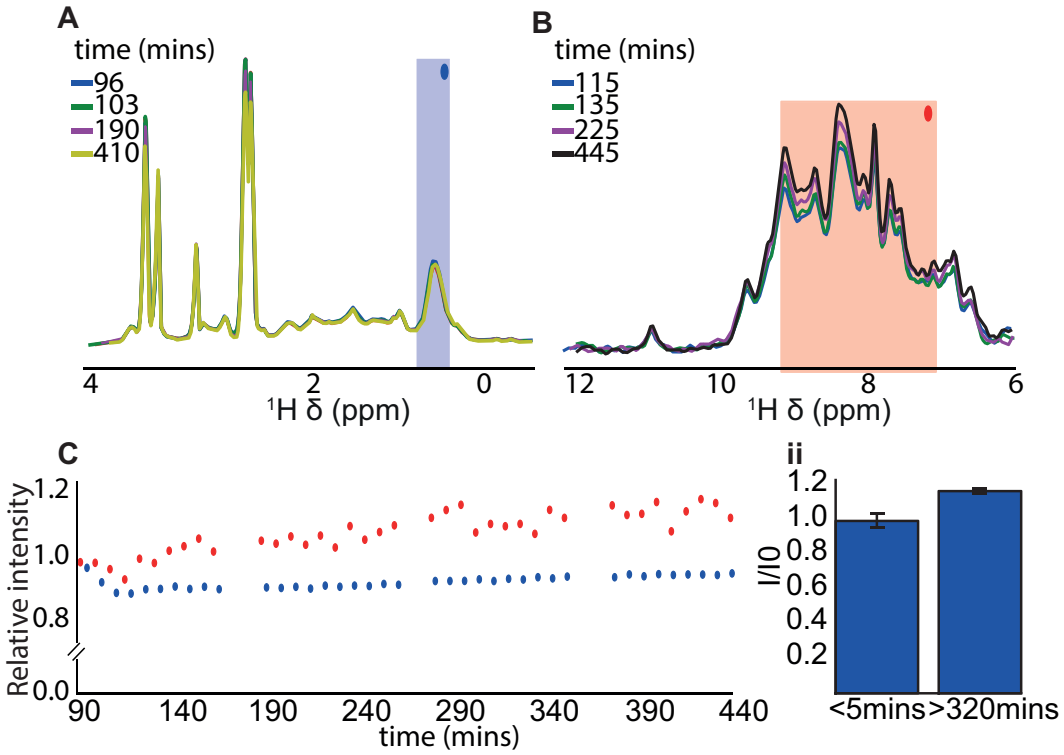


Figure 5.13: *Truncated HSP16.5 and α-lac aggregation assay. HSP16.5 was added at t=90 minutes (i.e. 90 minutes after DTT addition) and the subsequent reaction followed by NMR. A) Selected ^{1}H HSQC spectra, B) Selected 1D $^{1}H^{15}N$ HSQC spectra (reporting on chaperone only), C) intensity plots for chaperone and ^{1}H signal over time, Cii) quantification of chaperone intensity at early and late times*

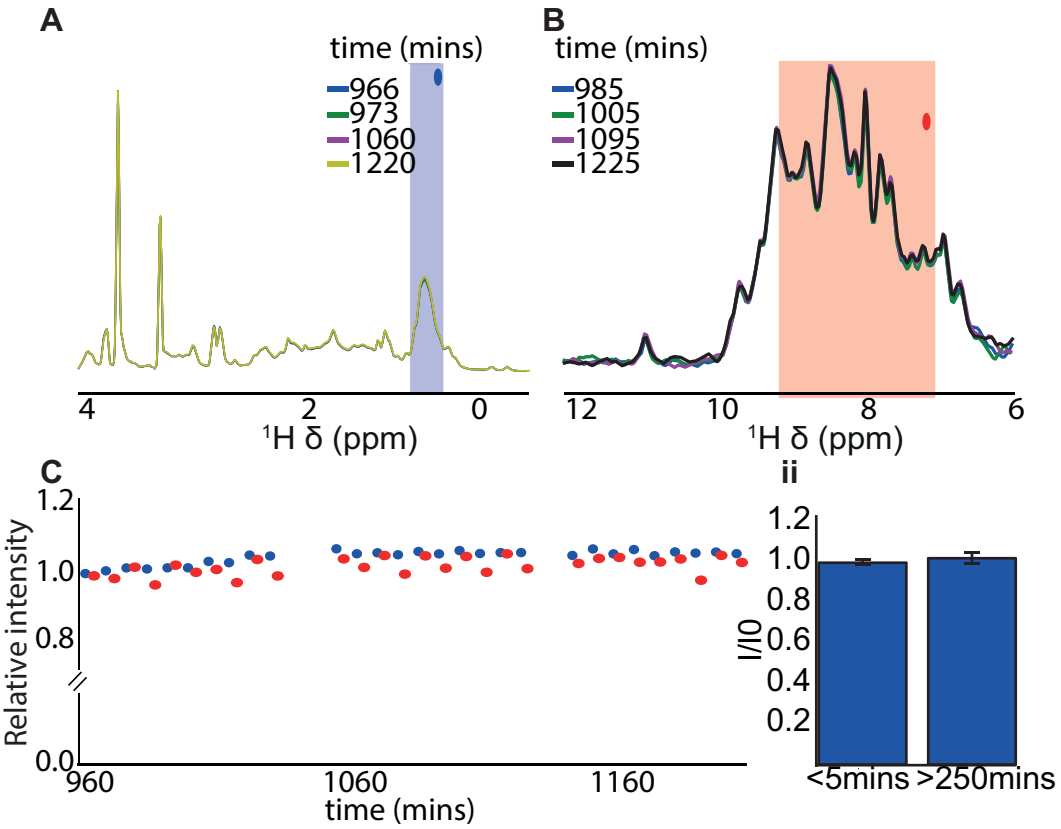


Figure 5.14: *Truncated HSP16.5 and α-lac aggregation assay. HSP16.5 was added at t=960 (i.e. 960 minutes after DTT addition) minutes and the subsequent reaction followed by NMR. A) Selected ^{1}H HSQC spectra, B) Selected 1D $^{1}H^{15}N$ HSQC spectra (reporting on chaperone only), C) intensity plots for chaperone and ^{1}H signal over time, Cii) quantification of chaperone intensity at early and late times*

chaperone signal followed by a rapid recovery (5.12B and C). Interestingly, the start of the chaperone recovery phase coincides perfectly with the end of the ^{1}H signal decay. This further supports the hypothesis that 16.5 core interacts preferentially with small scale aggregates.

Finally, addition of HSP16.5 960 minutes after DDT addition, during the flat, equilibrium phase of α-lac aggregation resulted in no significant change in chaperone signal intensity. Taken together, our data suggest that HSP16.5 interacts preferentially with smaller species on pathway to aggregation, this can be seen clearly by the effect on chaperone intensity of adding chaperone at different time points (figure 5.17). Further studies to obtain oligomer sizes by dynamic light scattering would clarify the role of HSP16.5 in chaperoning αlac.

An examination of the chemical shift perturbations (CSPs) also yields insight into HSP16.5 function. Chemical shift perturbations reflect an alteration in an environment sampled by NMR active nuclei, for example substrate binding, they are therefore often used in ligand binding studies (reviewed in [Williamson, 2013]). CSP measurements for HSP16.5 aggregation experiments (figures 5.12 to 5.14) are presented in figure 5.15. Under conditions where HSP16.5 is added shortly (0 or 90 minutes) after DDT addition, the same three residues are show significant (>3 standard deviations) CSPs, however these CSPs are not notable when HSP is added after aggregation has completed, consistent with intensity timecourse data which suggests minimal interaction between HSP16.5 and α-lac once HSP has completed aggregation (figure 5.14). Interestingly, these residues lie along the outermost β-strand of the α-crystallin domain. These CSPs derived from the core domain construct correspond to the outer face of the oligomeric HSP bipyramid structure (figure

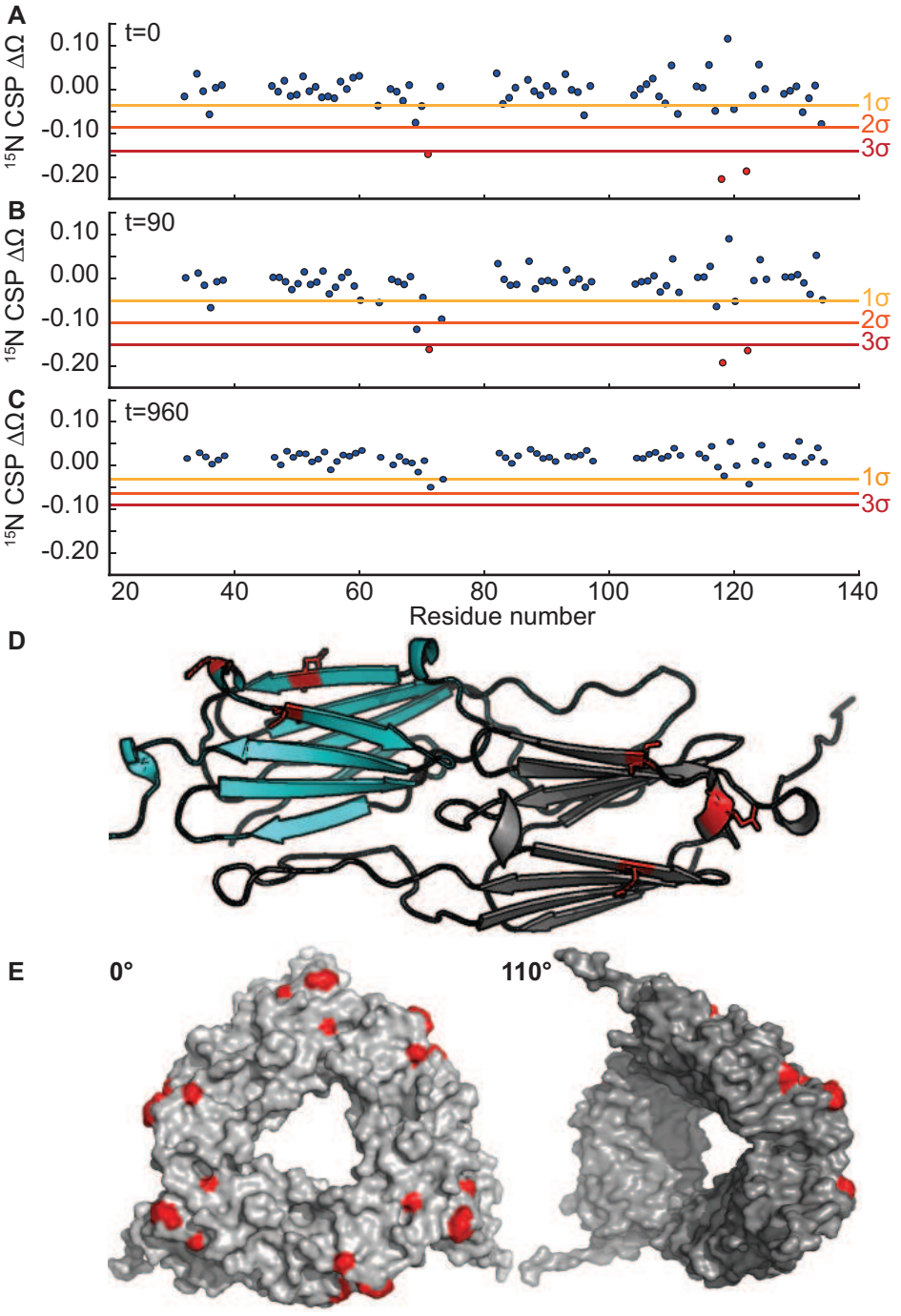


Figure 5.15: *Truncated HSP16.5 chemical shift perturbations (CSPs) at endpoints of respective experiments. A) CSPs for HSP16.5 added to α-lac concurrently as DDT, B) CSPS for HSP16.5 added to α-lac 90 minutes after DDT addition. C) CSPs for HSP16.5 added 960 minutes after DDT addition. Sigmas represent standard deviation multiples across all peaks.D) Residues which show the greatest CSPs (red) lie along β-strands at the edges of αcrystallin domains, E) likewise, CSP residues (red) identified on the core domain correspond to the outer face of oligomeric bipyramids (PDB 1SHS)*

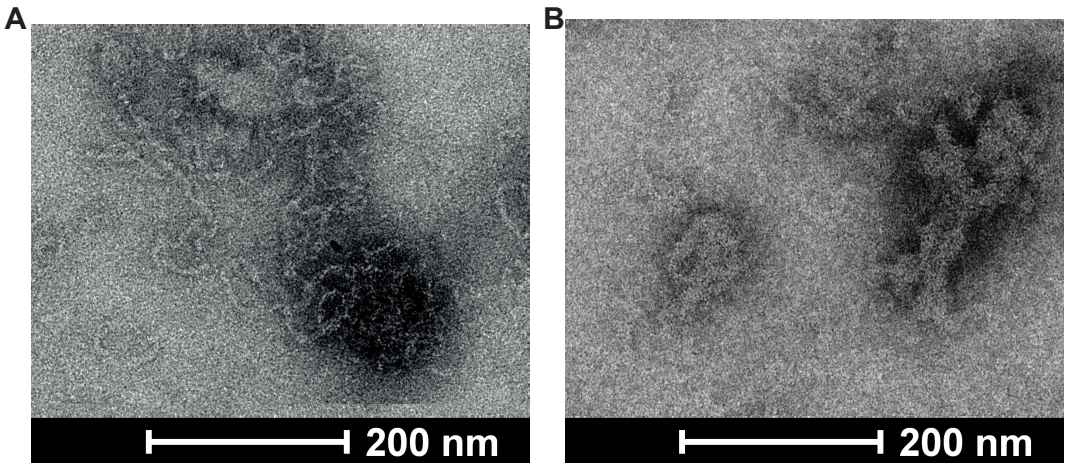


Figure 5.16: *α-lac aggregate morphology was unperturbed by sub-equimolar concentrations of truncated HSP16.5, A) α-lac aggregates (200μM), B) α-lac aggregated in the presence of HSP16.5 (200μM α-lac, 100μM HSP16.5 (monomeric concentration))*

5.15E).

Interestingly, HSP16.5's interaction with these species effectively slows aggregation, even when the substrate (α-lac) is derived from another domain (mammalian milk protein). Whilst a clear effect on aggregation can be observed through light scattering and NMR studies, EM comparisons of α-lac aggregated with and without HSP16.5 revealed no clear morphological differences (figure 5.16).

The ability of HSP16.5 to chaperone substrates across species may suggest a conservation of mechanisms across lifeforms, and thus provide insight into mammalian chaperones involved in health and disease [Hilton et al., 2013].

5.3 Discussion

Protein chaperones are key to resisting protein misfolding and protein aggregation [Lindberg et al., 2015]. ATP dependent chaperones, such as HSP104, have been reported to reverse amyloid formation and eliminate toxic oligomers [De-

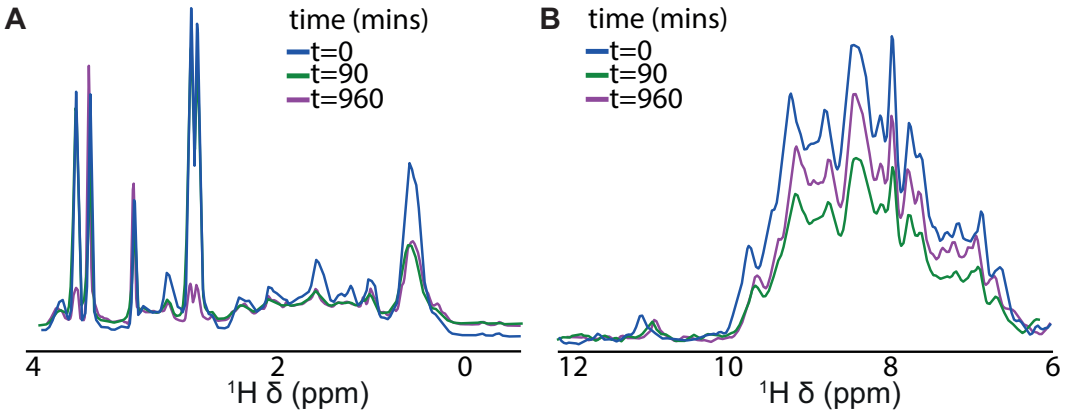


Figure 5.17: *Summary of aggregation experiments.* ^{1}H *(A) and* ^{15}N *edited* ^{1}H *(B) spectra are shown, corresponding to chaperone added immediately (t=0), 90 minutes (t=90) or 960 minutes (t=960) after DDT induced* α*-lac aggregation. The chaperone signal is attenuated less by addition at t=960 than t=90.*

santis et al., 2012, Lo Bianco et al., 2008]. Moreover, upregulation of HSP70 in animal models of Parkinson's disease resulted in reduced amyloidogenesis and neurotoxicity [Putcha et al., 2010, McLean et al., 2004].

The ATP-independent small HSPs have likewise been found to have amyloid suppressing capabilities [Hochberg et al., 2014]. A wealth of information on HSPs has elucidated their biological roles and subunit exchange mechanisms, however, the mechanism of chaperoning is still to be uncovered [Hilton et al., 2013].

Here, I have reported the study of both oligomeric small HSP16.5 and a truncated construct capable of only forming dimers, the so-called 'core domain'. Investigation of the core domain revealed spectra highly similar to the oligomeric variant, suggesting that there is substantial structural overlap between dimeric and oligomeric HSP16.5, I interpret this result as suggesting that the HSP16.5 dimer is the main NMR observable building block of the oligomer (figure 5.3).

Extensive analysis of multi-dimensional NMR data allowed for the assign-

ment of 70% of the core construct (figure 5.6). Due to the similarity in spectra between the core and oligomeric protein, I was able to transfer a number of core assignments to the oligomer, thus assigning 6/55 of the oligomeric peaks (figure 5.7). These assignments were then used to observe fast and slow scale dynamics on the 400 kDa 24mer HSP16.5 (figure 5.8). Slow scale CPMG dynamics of the oligomer localize along the dimer interface and suggest that I can observe dimerization by CPMG NMR with an exchange rate of 900±130 s^{-1}.

The establishment of the HSP16.5 system allowed for the study of chaperone-substrate interactions. To this end, a model system was established employing α-lactalbumin and HSP16.5 core. Oligomeric HSP16.5 is a notoriously poor chaperone, generally requiring very high temperatures and/or pressures to achieve significant chaperoning activity [Xi et al., 2014, Tölgyesi et al., 2004]. It is therefore remarkable that light scattering reveled the core domain to be able to effectively chaperone α-lac at equimolar ratios and mesophillic (37°C) temperatures (figure 5.10). Furthermore several authors have suggested that HSP16.5 oligomers chaperone by forming cage-like structures around substrates [Shi et al., 2013], whilst I cannot rule out multiple core domains assembling on a single, central substrate, I can posit that HSP16.5 24-mers are not necessary for chaperoning.

The high quality and assignment of the HSP16.5 spectra, coupled with the reproducible and malleable α-lac kinetics, provided a model system for small HSP chaperone study. Analysis of aggregation curves obtained by NMR (figure 5.11) allowed the deconvolution of signal arising from chaperone and substrate (figure 5.17), timecourse studies examining these signals revealed the strongest HSP16.5 - α-lac interactions to occur at early aggregation times, suggesting

that HSP16.5 interacts preferentially with smaller aggregating species.

I also observed significant CSPs on the edge of the immunoglobulin sandwich, which would correspond to the outer face of an oligomeric HSP16.5 (figure 5.15). It is difficult to reconcile this finding with the cage-like function proposed for HSP16.5 oligomer; if HSP16.5 forms cages around aggregating substrates then I would expect substrate interacting regions on the inner pyramid surface. Speculatively, it is possible these CSP rich regions are sensitive to substrate, and may have additional structural roles in the assembled HSP16.5

Taken together, I present the establishment of a model chaperone-substrate system for investigating small heat shock protein mechanisms and dynamics. Further investigation with a greater range of chaperone:substrate ratios and mechanistic modeling (as in chapter 4) may yield more specific information on which early time scale species are targeted by HSP16.5. Importantly, the system and approaches I have developed will enable us to probe the relationship between dynamics and chaperoning. It is possible that understanding the ability of small proteins to inhibit aggregation may be of use biomedically for amyloid and misfolding related diseases.

5.4 Methods

5.4.1 Protein purification

Isotopically enriched proteins were produced with E. *coli* BL21 strain cells in minimal media (m9) supplemented with $^{15}NH_4Cl$ and $^{13}C_6H_{12}O_6$. Core construct HSP16.5 protein was expressed with N-terminal tobacco etch virus (TEV) protease-cleavable His-tags and purified by nickel-affinity chromatography. The N-terminal His-tag was removed with TEV protease, and the

protein was further purified using a Superdex 75 gel filtration column (GE Healthcare). Protein purity was assayed by SDS-PAGE and found to be 99% pure after purification. α-lac was purchased from Sigma.

5.4.2 NMR

NMR experiments were performed on a Varian 600 MHz spectrometer with a 5-mm z-axis gradient triple resonance room temperature probe at 37°C (aggregation) or 60°C (assignment and dynamics). Isotopically labeled ^{15}N and ^{13}C HSP16.5 were buffered in 50 mM NaPi, 100mM NaCl, 2 mM sodium azide (pH 7.0) and concentrated to 200 μM (aggregation assays) or 900μM (assignment sample) (monomeric concentration) using Amicon concentrators (Millipore). 2D ^{15}N-^{1}H HSQC and HMQC spectra were recorded with 128 scans per transient with acquisition times of 67.5 ms, over 100 complex points and a 1 s relaxation delay. 2D ^{13}C-^{1}H HMQC were recorded with 64 scans per transient and an acquisition time of 0.064 s, over 100 complex points and a 1.5 s relaxation delay.

5.4.3 Aggregation assays

Light scattering assays were carried out in sealed 96-microwell plates and a Fluostar Optima plate reader (BMG Lab Technologies). The amorphous aggregation of α-lac (Sigma-Aldrich, 100μM), incubated at 37°C in 50 mM NaPi, 100mM NaCl, 2 mM sodium azide (pH 7.0), was initiated by the addition of DTT (20 mM). Chaperones were added at the molar ratios stated. Aggregation was monitored by measuring the change in apparent absorbance caused by light scattering at 340 nm, which was negligible in the absence of α-lac

NMR aggregation assays were carried out at 37°C in 50 mM NaPi, 100mM NaCl, 2 mM sodium azide (pH 7.0) in a Varian 600MHz spectrometer. Spectra were recorded at time points shown. Integrals of the regions were plotted (also shown, figure 5.12). ^{1}H water gate spectra were acquired with 32 scans per transient with acquisition times of 67.5 ms and a 1 s relaxation delay. Likewise, ^{1}H ^{15}N spectra were acquired with 32 scans per transient, acquisition times of 67.5 ms, over 1 complex point and a 1 s relaxation delay.

5.4.4 Electron microscopy

Carbon coated grids (Electron Microscopy Sciences) were incubated with diluted microwell sample (10 μM) for 2 minutes and post stained with 2% aqueous ammonium molybdate (10 s) and washed with water droplets. Grids were then imaged on a transmission electron microscope (FEI Tecnai 12) operated at 120kV using a digital CCD camera (Ultrascan US1000, Gatan).

Chapter 6

Conclusions

6.1 Thesis conclusions

Together, protein folding and aggregation represent one of the greatest challenges to human health in the 21^{st} century. In this thesis, we have employed a combination of solution state nuclear magnetic resonance, mass spectrometry, X-ray crystallography and associated biophysical techniques to gain insight into the mechanisms and causes of amyloid disease.

I have here introduced a novel labeling method to study post-translationally modified GCase by NMR spectroscopy (chapter 3). By introducing ^{13}C enriched formaldehyde onto GCase post-production I am able to examine a large human cell-derived protein with human glycosylation patterns. The favorable relaxation properties of methyl groups allow the study of large proteins at low concentrations (here, a 60 kDa protein at 25 μM). However, the NMR spectra produced from such methods contain heavily overlapped peaks, due to highly similar chemical environments (solve exposed amine groups). Whilst this is not ideal for NMR spectra analysis, there are currently few alternatives available for introducing such labels onto human proteins. Furthermore, unlike

selective MILV labeling using amino acid precursors [Tugarinov et al., 2006], the formaldehyde method is incredibly cheap. Future work may look at possible improvements to the methylation strategy, such as part-^{132}C, part ^{13}C labelling, both to assign NMR spectra and reduce crowding of peaks.

The work presented in chapter 3, detailing pharmacological chaperone interactions with GCase suggests that combining ^{13}C with interesting biological questions can potentially yield insight. The pharmacological chaperone work needs additional techniques (for example analytical ultra-centrifugation) to corroborate the effects on dimerization seen by NMR and mass spectrometry. Additional pharmacological chaperones at a greater range of concentrations could also be employed to elucidate whether the stabilization of GCase dimers is a shared property. It is of interest that the trafficking of GCase from the ER to the lysosome is pH dependent [Zhao et al., 2014], dimer population is pH dependent, and that pharmacological chaperones may alter dimer populations. Speculatively, the ability of PCs to stabilize GCase dimers may underpin their therapeutic mechanism of action *in vivo*.

Subsequent chapters expand upon the study of protein misfolding and aggregation. Chapter 4 utilizes labeling methods developed in chapter 3 to posit a molecular 'tug-of-war' between GCase and lipids for αsyn. I believe this to be a major discovery and the most important impact from my D.Phil. The lysosome is emerging as a key organelle in Parkinson's disease pathology, responsible for degrading endogenous αsyn [Cuervo et al., 2004], trafficking exogenous αsyn between cells [Bae et al., 2014] (and thus, presumably, accounting for the ability of αsyn to spread across the brain [Li et al., 2008]). Furthermore, lysosomal conditions, namely, low pH and anionic lipid rich membranes have been shown to increase αsyn aggregation by several orders of magnitude [Galvagnion

et al., 2015, Buell et al., 2014]. GCase mutation (or depletion in sporadic PD [Gegg et al., 2012]) would thus diminish the protective role of GCase in the lysosome and facilitate αsyn aggregation. Furthermore, recent experiments on Parkinson's disease animal models by Sardi. *et al.* reveal that wild type GCase infusion is capable of reversing PD associated pathology [Sardi et al., 2013, Sardi et al., 2011]. I believe my work is capable of explaining these multiple phenomena, and, speculatively, I would expect infusion of wild-type GCase into human Parkinson's disease patients to relieve their underlying pathology. I would further expect pharmacological chaperone treatments of GCase mutation associated PD patients to also alleviate PD pathologies.

This work could be expanded in a number of directions. First, the effect of mutant GCase on the interaction with αsyn could be examined. There is some data suggesting that the relatively common N370S mutant has a lower affinity for αsyn than wild-type (consistent with my theory of protection outlined in chapter 4). I believe the most insight would be from the E326K mutant, which is associated with PD but not Gaucher's disease. I would hypothesize that this mutant would also have a reduced ability to rescue αsyn from lipid membranes. Secondly, a *in vivo* study could be employed, treating GCase mutant mice with pharmacological chaperone and observing the prevalence of PD pathology and memory-impairment. Finally, phospholipase D is also reported to interact with αsyn C-terminus [Payton et al., 2004], further probing of this interaction may yield insights into a conserved protective mechanism.

In chapter 5 I examined the protective role of canonical chaperones, such as small heat shock protein (HSP) 16.5, in protein aggregation. Whilst the chaperone of choice (HSP16.5) is not of biological interest, it was thought that mechanistic insight may be generalized to other small heat shock proteins

involved in the chaperoning of neurodegenerative-disease associated proteins [Hochberg et al., 2014]. Further examination is required to yield substantive results, as a priority, oligomer distributions of aggregating protein α-lac could be examined by dynamic light scattering to corroborate the speculative results suggesting that HSP16.5 has a preference for smaller-scale oligomers. A fuller comparison of the chaperoning abilities of the shortened 'core' construct and the wild type HSP16.5 could also be employed and may yield insight into the relative role of HSP16.5 oligomers and dimers in chaperoning substrate.

Finally, the study of RNA editing enzymes, critical to post transcriptional regulation of protein levels, is reported. A combination of crystallography and mass spectrometry revealed Cid1 to be capable of binding RNA alone, and expression trials of mammalian homologues of Cid1 have yielded promising results for future studies (chapter 7). Methods first developed here became integral to efforts aiming to study GCase by NMR spectroscopy.

This thesis has gained most of its biophysical insights through the application of nuclear magnetic resonance and careful data analysis. The main advantage of NMR lies in its ability to simultaneously study multiple transiently interacting partners and their dynamics. Together, this represents a powerful system for deconvoluting complex biophysical phenomena. After deconvolution, data can be analyzed using simple algorithms to extract as much information as possible and make falsifiable predictions for further model validation.

Together, our data here suggest multiple mechanisms for improving the cell's ability to maintain proteome integrity. Small molecule therapy with pharmacological chaperones (chapter 3), the balancing of protein levels within aggregating systems through protein replacement therapy (chapter 4), and the

up-regulation of innate molecular chaperones (chapter 5) may all play a role in combating protein folding diseases.

6.1.1 Publications

Much of the work in this thesis is in preparation for publication. Chapter 4 is currently in review with PNAS. Chapter 7 contributed to a publication in Acta F. crystallography [Yates et al., 2015a], whilst chapter 3 is currently in preparation for submission. Finally, chapter 5 will be prepared for a biochemical journal in the near future.

Chapter 7

Appendix: Crystallographic studies of RNA editing enzymes

Abstract

Aberrant RNA metabolism can lead to a wide range of diseases such as cancer, motor neurone disease, and diabetes. There are many overlapping regulatory mechanisms for RNA metabolism; however an emergent and pervasive mechanism regulating both mRNA and miRNA maturation is 3' terminal uridylylation; which is catalyzed by terminal uridyl transferases (TUTases) such as *Schizosaccharomyces pombe* Cid1 or mammalian proteins TUT4 and TUT7. This chapter reports the successful purification, crystallization and diffraction of Cid1 to at least 2.1 $\mathring{A}$, an improvement on the 3 Åstructure previously solved in-group. I also report the first use of surface entropy reduction strategies which later gained traction in chapters 3 and 4, and the expression of mammalian TUT constructs for future NMR study. Together, I have established a solid foundation for future work combining X-ray crystallography and NMR.

7.1 RNA editing enzymes and crystallography: An introduction

Francis Crick's central dogma has underpinned our understanding of biology for decades [Crick, 1970], and includes RNA as the messenger between gene (information) and protein (effector). Recent evidence, however, has challenged this simplistic view by uncovering diverse RNA molecules with many roles. Three important discoveries have changed the way we view the genome and Crick's central dogma. The first was the elucidation of a tRNA structure in 1973 [Kim et al., 1973, Holley et al., 1965] (giving the characteristic cloverleaf structure known to all students of biochemistry today) which suggested that RNAs could have roles other than as simple messengers. The second great discovery, in 1982, was of the importance of RNA enzymes (ribozymes). The most famous RNA enzyme being the ribosome [Kruger et al., 1982, Ban et al., 2000]. But perhaps the most profound series of discoveries began in the 1990s, with microRNA (miRNA), first discovered in *Caenorhabditis elegans* - miRNA is capable of exquisite regulation of protein levels post-transcriptionally [Lee et al., 1993]. Together, these discoveries have opened immense new fields for research beyond the scope of this chapter. Here, I focus on post-transcriptional regulation of mRNA and the elegant role played by miRNA in this process.

7.1.1 The many roles of short RNA

Short RNAs add a layer of complexity to the central dogma by exerting influence over diverse processes; for example host defence, where piwiRNA (piRNA) in germline cells protect against parasitic transposons and are implicated in sterility when disrupted [Aravin et al., 2006, Tam et al., 2008, Malone et al.,

2009, Malone and Hannon, 2009]. Furthermore, prokaryotes have evolved their own RNA based defense against bacteriophage known as CRISPR. The CRISPR system incorporates small segments of foreign DNA into the host genome to then produce anti sense RNA. In this way, the invading pathogen's gene products are silenced by base pairing to CRISPR RNA [Sabin et al., 2013]. Indeed, RNA has become so central to our understanding of biology that it has even been posited as the progenitor of life on earth (the so called 'RNA world' hypothesis) [Talini et al., 2009, Neveu et al., 2013, Müller, 2006].

More recently, short RNA molecules have become known for their role in post-transcriptional regulation [Fabian et al., 2010]. Briefly, short (22nt) miRNA sequences can base pair with complementary mRNA sequences to inhibit translation and thus silence gene expression - miRNA-RNA complexes are degraded by the 'RNA Induced Silencing Complex' (RISC, reviewed in [Paroo et al., 2007]). miRNA mechanisms complement established systems of RNA regulation, such as 5' capping and 3' polyadenylation (figure 7.3)[Kapp and Lorsch, 2004], to control RNA turnover and protein expression (figure 7.2). Such miRNA systems exist in all higher organisms (and there are homologous systems in simpler life), however the first miRNAs, lin4 and let7, were discovered in model organism *C. elegans* [Boehm and Slack, 2005, Sokol, 2012, Roush and Slack, 2008], and subsequently, their role in tumor suppression was elucidated (reviewed in [Hayes et al., 2014]).

The biogenesis of miRNAs is outlined in figure 7.1. Briefly, long pri-miRNAs are generated from intronic or intergenic regions of the genome which undergo numerous processing steps which trim the RNA to a 22 nucleotide, single stranded, core sequence capable of binding 3' mRNA regions [Lee et al., 2003]. Such trimming steps can be dependent or independent on cleavage

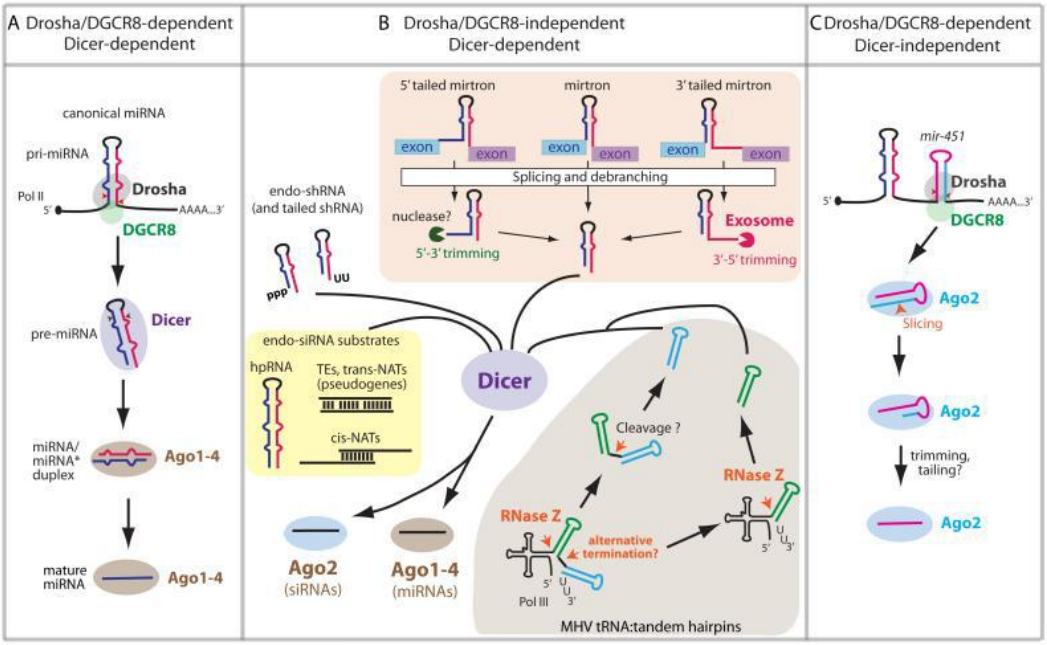


Figure 7.1: *miRNA processing is a complex system of pathways, however there are two key nodes, Drosha, a nuclear localised RNase III, and Dicer, a cytoplasmic RNase III. miRNA processing is defined as either Drosha or Dicer independent/dependent. A) Drosha and DGCR8 (Pasha) dependent processing (canonical). B) Drosha independent but Dicer dependent (the so called mirtrons, miRNA generated from introns). C) Drosha dependent, Dicer indeprendent processing, this is most clearly seen with miRNA451. Where the Ago2 protein ('Slicer') cleaves the pre-miRNA to the correct size. Image from [Yates et al., 2013]*

protein Dicer[Weinberg et al., 2011, MacRae and Doudna, 2007].

7.1.1.1 Regulating the regulators

The pivotal role of RNA in the central dogma provides an obvious impetus for research into RNA metabolism, especially as aberrant RNA metabolism can lead to a wide range of diseases [Lu et al., 2008, Amiel et al., 2012], including cancer [Farazi et al., 2013], motor neurone disease [Lemmens et al., 2010] and diabetes [Kornfeld et al., 2013]. In this chapter I am concerned with an emergent, pervasive, mechanism of RNA regulaltion, namely, 3' terminal uridylylation catalyzed by terminal uridyl transferases (TUTases) [Choi et al., 2012, Norbury, 2010, Rissland and Norbury, 2008].

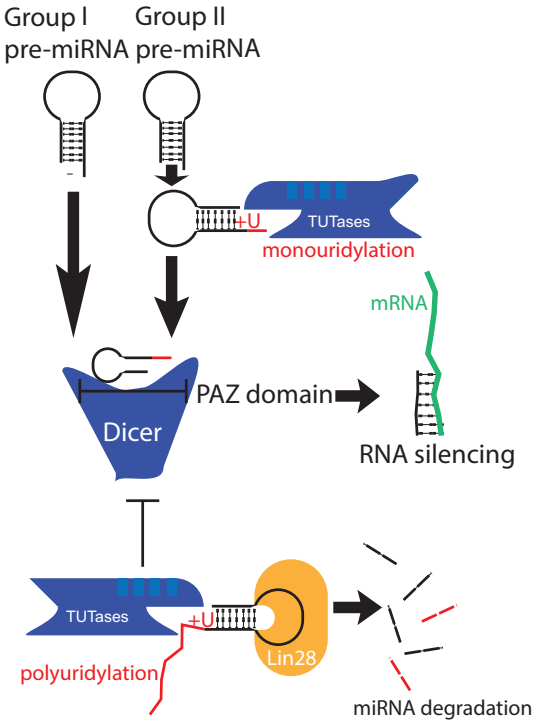


Figure 7.2: *TUTases in miRNA. Group II miRNAs require TUTase dependent monouridylation to enable binding to Dicer's PAZ domain (a sort of molecular ruler). TUTases, however, can also polyuridylate miRNA to signal their degradation*

Uridylylation of miRNA (figure 7.2) demonstrates the complexity of oligo(U) regulatory mechanisms as it can both enhance and inhibit miRNA maturation. Initial processing of pri-miRNAs generates a pre-miRNA with a two nucleotide 3'overhang that can be preferentially processed by Dicer [Park et al., 2011]. However, a subset of pre-miRNAs (the so called group II miRNAs) are instead left with a single nucleotide overhang, making them sub optimal for Dicer.

Recently, it was shown that the group II single nucleotide overhang can be mono-uridylated by TUTases to provide the necessary 2 nucleotide overhang [Heo et al., 2012]. siRNA knockdown of several TUTases specifically inhibited the maturation of several group II miRNAs (including the tumour suppressing let-742). Most interestingly, these same components can be used to inhibit miRNA maturation. The addition of Lin28 to an in vitro mixture of pre-let-7 miRNA and TUT4 leads to the 3' oligouridylylation of the miRNA and the inhibition of maturation [Heo et al., 2009]. Finally, miRNA directed silencing of mRNAs itself involves uridylylation. Cloning of the decaying mRNA intermediates in Arabidopsis thaliana found a non-templated stretch of 1-9 uridines downstream of the cleavage site [Shen and Goodman, 2004]. This oligouridine was purported to recruit degradation machinery to the mRNA. This has recently been validated by two independent reports [Rissland and Norbury, 2009, Mullen and Marzluff, 2008].

The first of these reports concerned histone mRNAs. Tight regulation of histone mRNA is essential for cell cycle control. Mullen *et al.* (2008) reported that the histone mRNA of cells arrested in S-phase contained a 3' oligouridyl tail and proposed that this tail recruited decapping and degredation machinery [Mullen and Marzluff, 2008]. Further support was provided by Rissland and Norbury who, in 2009, examined multiple mRNAs from S. pombe and found mono/di uridylylation of these transcripts [Rissland and Norbury, 2009]. Uridylylation was Cid1 dependent and preceded decapping [Rissland and Norbury, 2009]. A final added complexity to mRNA degradation was contributed by Sement *et al.* who have reported (in an S. pombe model) that 3' uridylylation protects RNA from deadenylation and 3'-5' degradation, whilst encouraging 5'-3' degradation. Therefore uridylylation is thought to give RNA

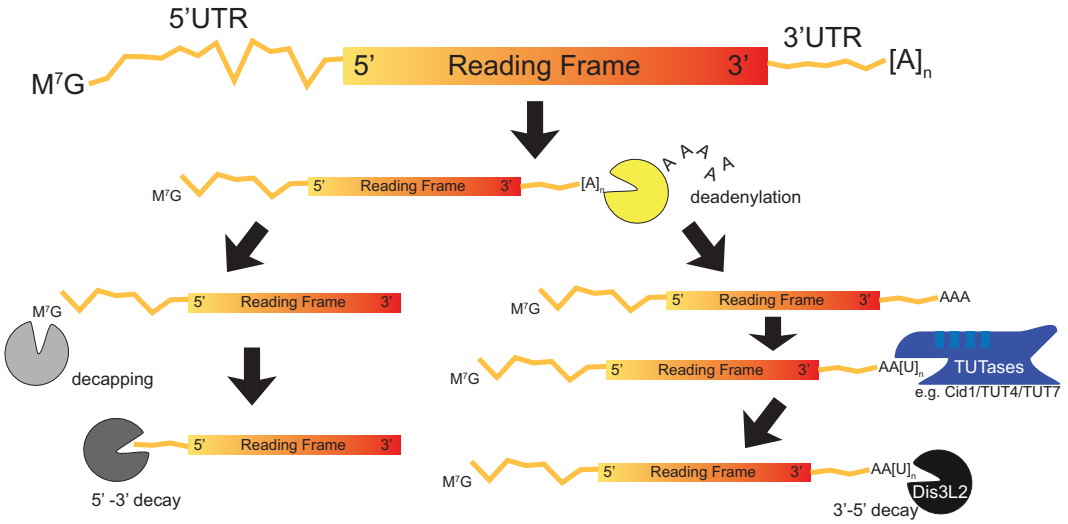


Figure 7.3: *RNA degradation pathways. 3' uridylylation has been found to play a novel role in RNA regulation and the 'uridylylation pathway' is beginning to be elucidated. Thus far, it seems that uridylylation can co-opt existing mechanisms as well as use those that are newly discovered. The diagram above shows the multiple pathways to RNA degradation and the role of TUTases in those pathways.*

degradation its polarity [Sement et al., 2013], which may have importance for co-translational degradation.

7.1.1.2 Terminal uridyl transferases

The elucidation of 3' uridylylation as a key regulator of RNA metabolism has generated considerable interest in the proteins that catalyze uridylylation, which are collectively known as TUTases, ZCCHCs, or Cid1-like proteins.

The best characterized of these proteins is the founding member Cid1 (caffeine induced death suppressor 1); a yeast protein that shares significant homology with the catalytic domain of mammalian orthologues (figure 7.5). Cid1 was first discovered in a screen for factors involved in the cell cycle [Wang et al., 1999] and is part of the polymerase B superfamily. Its uridylyl nucleotidyltransferase activity was initially reported as an adenylyl nucleotidyltransferase activity that was involved in polyA tail maintenance [Stevenson and Norbury,

2006]. However, further investigation revealed that Cid1 had a much higher affinity for UTP than ATP (around a 15:1 incorporation ratio at physiologically relevant levels [Rissland and Norbury, 2008]) and in-group crystallographic structures have now elucidated the mechanism of selectivity (figure 7.4). A single residue, His336, differs from the trypanosomal TUTases but is conserved in mammalian orthologues. Strikingly, mutation of this residue to an alanine confers polyadenine polymerase (PAP) activity to Cid1. It was proposed that a single water molecule, coordinated by the His336 and uracil (of UTP) also aids selectivity of UTP over ATP. Additionally, the employment of electrophoretic mobility shift assays (EMSA) and crystallography have identified an RNA binding groove on the surface of Cid1 [Yates et al., 2012].

A partial elucidation of the Cid1 catalytic cycle was achieved through the observation of both UTP bound and APO Cid1 structures in the crystal asymmetric unit. Structures differ by the presence of two ordered anti-parallel β strands, termed the β trap door, present in the UTP bound state and absent in the APO state. It is speculated that this structural motif may be partially responsible for UTP selectivity and correlated to alternate His336 conformations [Yates et al., 2012]. Interestingly, whilst structure-based phylogenetic analysis suggest Cid1 to be similar to the last common TUTase ancestor, this β trap door is lacking in Cid1 orthologues from other species [Rissland and Norbury, 2008, Yates et al., 2012].

Most cell biological work on terminal uridylylation has been carried out in yeast models, however there is evidence of extensive conservation in mammals (figure 7.5) [Hagan et al., 2009, Jones et al., 2009]. Most metazoan research has focused on typanosomal TUTases [Aphasizhev and Aphasizheva, 2011, Aphasizhev and Aphasizheva, 2008] and mammalian TUTases involved in let-

7 maturation (specifically TUT4 and TUT7). Such research efforts may open up new therapies for trypanosomiasis (African sleeping sickness) and cancer, respectively.

Mammalian TUTases possess a catalytic domain with a high degree of homology to *S. pombe's* Cid1 (figure 7.5). However, TUT4 and TUT7 also contain several zinc knuckle and zinc finger domains as well as an N-terminal catalytically inactive Cid1-like nucleotidyltransferase domain [Blahna et al., 2011]. TUT4 and TUT7 serve to redundantly block pre-let-7 maturation via terminal uridylylation in mammalian cell lines and in vitro[Hagan et al., 2009, Thornton et al., 2012]. The exact mechanism for this TUTase-Lin28-RNA complex is unknown, however reconstitution experiments have identified an N-terminal C2H2 zinc finger domain as responsible for interaction in vitro, suggesting a nucleic acid mediated interaction between these proteins [Thornton et al., 2012]. Real time single molecule approaches have suggested Lin28 as a processivity factor for TUTases, prolonging the time TUTases spend bound to the RNA to allow oligouridylylation [Yeom et al., 2011]. This is consistent with previous reports which differentiated the mono and oligouridylating abilities of TUTases [Heo et al., 2009, Heo et al., 2012]. Recently TUT4 has been implicated in immune modulation and TUT4 knockout mice show diminished neonatal survival via compromised IGF-1 expression [Jones et al., 2009, Jones et al., 2012].

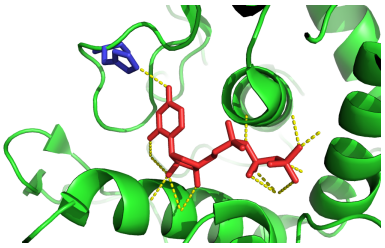


Figure 7.4: *Cid1 UTP binding site. Cid1 has a much higher affinity for UTP than ATP, crystallographic studies revealed the mechanism for this; His336 (blue) forms specific contacts with UTP (red). Further local contacts (yellow dashed lines) are established between UTP and Cid1. (PDB code: 4E8F.)*

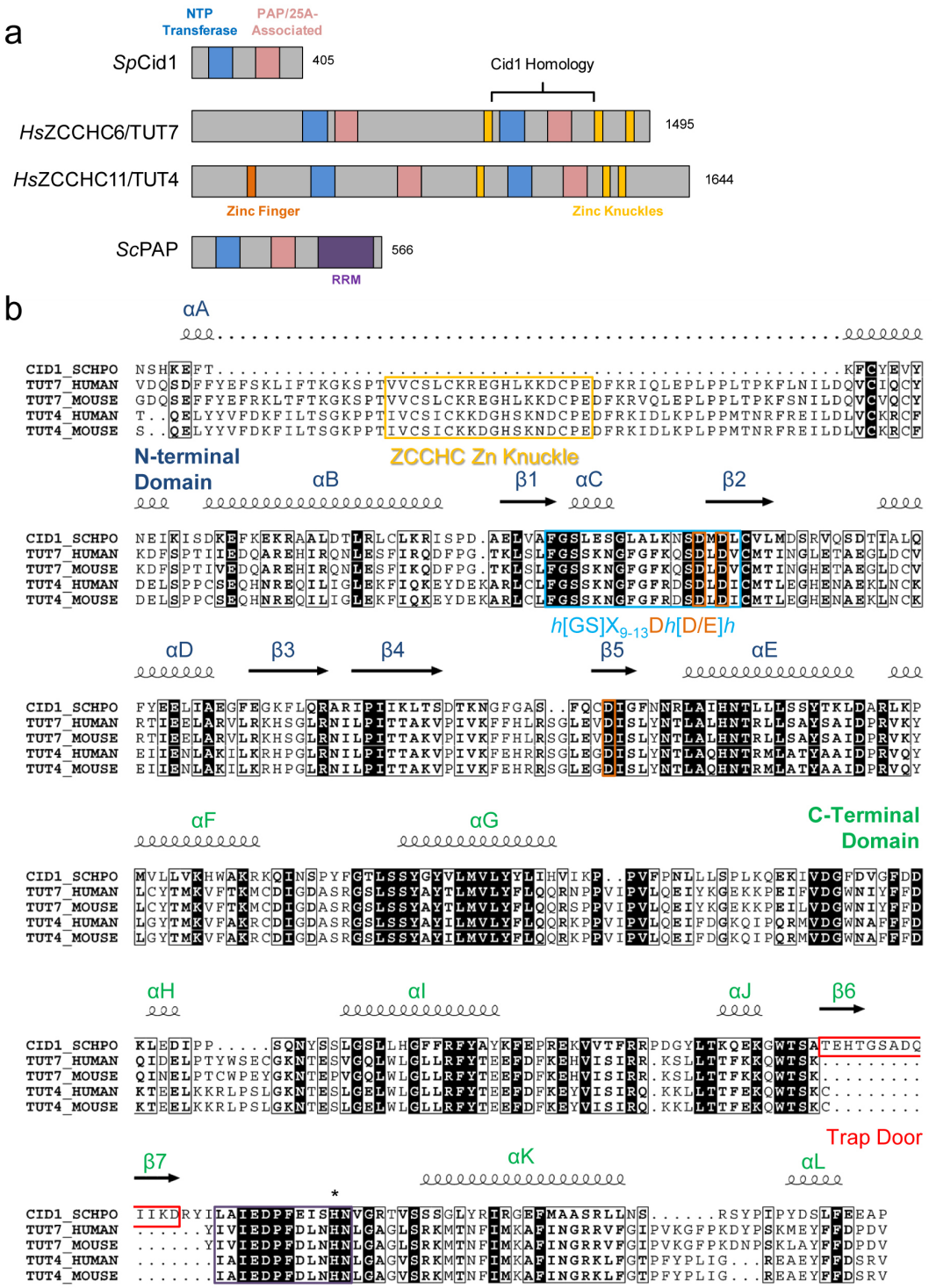


Figure 7.5: *Conservation of TUTases. a) Domain organisation of Cid1 compared to human TUT4 and 7. Human TUTases are much more complex, with an enzymatically inactive Cid1 like domain and several zinc fingers/knuckles. b) Sequence analysis of Cid1 and mouse/human TUT 4 and 7 reveals a relatively high degree of similarity, h represent hydrophobic residues, His336 (responsible for UTP selectivity) is indicated by an asterisk (*). From [Yates et al., 2012]*

The importance of TUTases to mammalian systems, and the difficulty of its expression and crystallization, has made Cid1 the best model for the structure and function of TUTases. The Cid1 structure was recently solved by several groups [Yates et al., 2012, Munoz-Tello et al., 2012, Lunde et al., 2012] at a resolution of up to 2 Å. However, the exact mechanism of catalysis, and the full catalytic cycle is still unclear, high resolution crystallographic analysis would provide a detailed understanding of this process. There is also much to be uncovered about the mode of RNA binding, and possible cooperativity effects between RNA binding and catalysis [Stagno et al., 2007] however it has been demonstrated that RNA-binding is essential for enzymatic activity [Yates et al., 2012]. Therefore it is necessary to obtain high resolution x-ray diffraction data of Cid1 with its substrates, including RNA, and cofactors.

Here, I report the improved crystallization of Cid1. Examination of crystal structures and comparison to previous structures reveals some differences, notably in the β-trapdoor motif which is absent here but present in previous structures from within the group. Cid1 was also studied by NMR and MS and it's ability to bind RNA confirmed. Finally, the expression of soluble domains of mammalian TUTases is reported. Together this work led to the methods utilized in chapters 3 and 4, and represents a foundation for future study of TUTases combining NMR and X-ray crystallography.

7.2 Results: The enhanced crystallization of Cid1

Previous in-group crystallization trials had yielded crystallographic data of 3-3.2 Å[Yates et al., 2012], whilst other groups have reported data with res-

olutions up to 2 Å[Lunde et al., 2012]. The original in-group crystals were obtained through analysis of proteolytic digestion fragments of Cid1 and the construct named truncated Cid, or tCid. However, higher resolution structures were desired for insights into small molecule and RNA binding mechanisms. To this end, a combination of two strategies were employed to yield resolution enhancements. First, residues associated with areas of low electron density in the tCid data set were removed. These corresponded to flexible N and C termini and yielded a new minimal Cid construct termed minimal Cid (mCid). Secondly, surface entropy reduction methods were employed and anticipated to yield high quality crystals through reducing the flexibility of surface exposed amino acid side chains [Walter et al., 2006]. I carried out over 3000 crystallization trials (including optimizations) and report the crystallization and diffraction of mCid to 2 Å(see methods for details).

7.2.1 Production and purification of mCid

N-terminal hexahistidine tagged mCid was produced from autoinducing TB media. Expression trials across a range of IPTG concentrations, temperatures, media and induction cell densities (in multi factorial experiments) produced negligible protein yield. However, a TB-autoinducing strategy yielded significant (10 mg/L) mCid. Initial autoinduction trials were carried out at 37°C (to replicate standard autoinducing media conditions), whilst this yielded significant amounts of protein, the size exclusion trace (120ml Superdex s75 -figure 7.6) was asymmetric, indicative of different geometries. I hypothesized that the high temperature produced mCid was partially unfolded or misfolded. To remedy this, a more optimal temperature was found (18°C) which yielded a

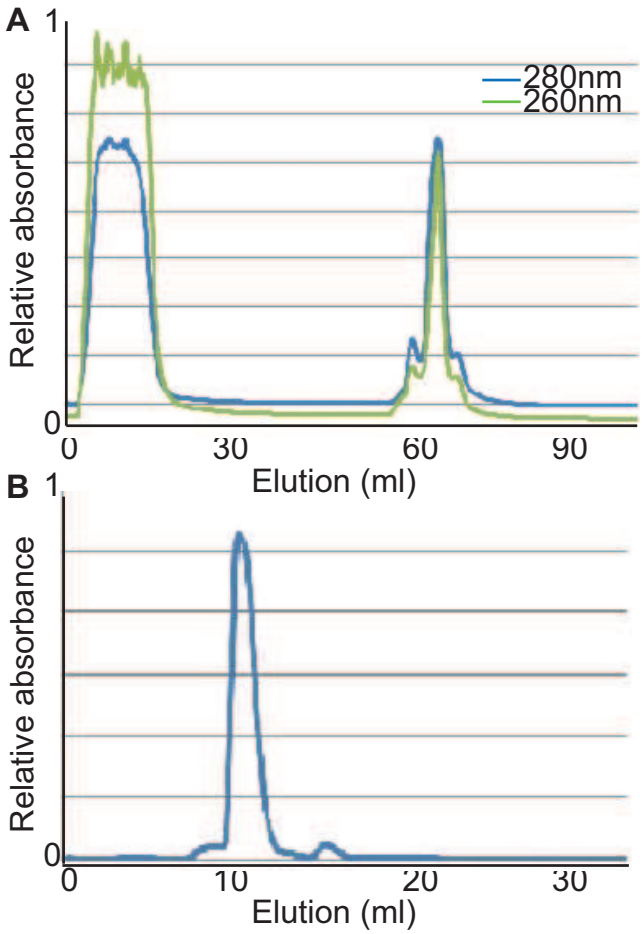


Figure 7.7: *mCid purification. A) Heparin affinity purification was carried out on an AKTA FPLC with a 260 and 280 nm wavelength UV source, to reveal presence of nucleic acid and protein respectively. It can be observed that most nucleic acid contamination is lost in the flow through. B) size exclusion chromotography was employed to assay protein quality and as a final purification step, a symmetrical central peak is observed*

smoother and more symmetric size exclusion trace (figure 7.7). The full purification procedure consists of three steps (methods); histidine tag purification, heparin affinity purification and size exclusion chromatography (figure 7.7). These steps resulted in pure, homogenous mCid. Interestingly, mCid elutes from heparin columns at a higher salt concentration than tCid, I reasoned that this may suggest it binds RNA with a higher affinity, and so complexed pure mCid with an 11mer of RNA for additional crystallization trials (see methods).

7.2.2 mCid surface entropy reduction strategies

It is not possible to rationally predict the ideal conditions for protein crystallization. Contemporary approaches to rational screening generally involve exploring an individual protein's phase space through high throughput microfluidics [Anderson et al., 2006], or large scale mutagenesis efforts [Derewenda, 2004]. However, these are of limited success and have not been widely introduced in crystallography thus far.

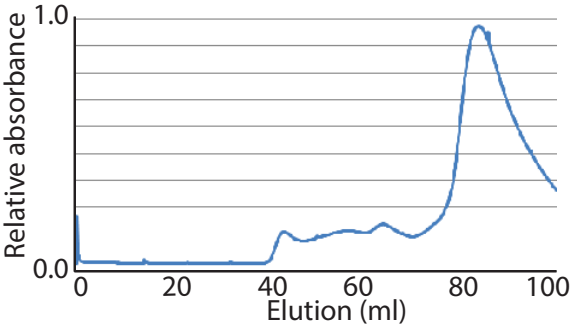


Figure 7.6: *mCid produced at a higher temperature shows an asymmetric size exclusion chromotography profile, indicating variable geometry. This may be due to partially misfolded protein.*

Instead, the most common strategy for crystallization trials is to employ stochastic searches covering a wide range of crystallization space. Any alteration to protein or buffer potentially explores more of this crystallization space and therefore may yield crystals of appreciable resolution. One such alteration is the introduction of methyl groups onto the side chains of proteins. This method has become popular due to the extensive evidence suggesting that the native fold of the protein of interest is unperturbed by methylation [Kurinov et al., 2000, Lund-Katz et al., 1988, Hattori et al., 2013, Walter et al., 2006, Larda et al., 2012, Rayment, 1997, Gerken et al., 1982].

The methylation of proteins utilizes formaldehyde and borane dimethylamine complex to introduce CH_3 groups onto lysine and arginine side chains through reductive amine methylation [Walter et al., 2006]. Whilst these labels are thought not to perturb biological function, they do reduce the surface en-

tropy of a protein and thereby may enhance the crystallization of recalcitrant proteins.

The application of surface entropy reduction to mCid *via* amine methylation, however, led to extreme precipitation (>75%). The remaining soluble protein was analyzed by size exclusion chromatography (figure 7.8) and showed a significantly asymmetric peak (which may be comprised of two separate peaks, representing alternate morphologies). Based on this analysis I reasoned that the remaining soluble methyl-mCid may be resistant to aggregation simply because it had been inefficiently methylated. Liquid chromatography tandem mass spectrometry analysis confirmed this; only 2 lysines were found to be methylated with an LC-MSMS sequence coverage of 53%. These results led to the discontinuation of this strategy, however, it is possible that surface entropy reduction utilizing other protocols [Shaw et al., 2007], reagents or buffers [Lasker et al., 2006] may prove better suited to mCid.

Whilst crystallization efforts to reduce Cid surface entropy by amine methylation patently failed, the approach first trialled here was later employed alongside isotopically enriched ^{13}C formaldehyde to introduce spin $\frac{1}{2}$ labels for the NMR study of Glucocerebrosidase (chapters 3 and 4).

7.2.3 mCid crystallization trials

After the failure of the methylation trials, I turned to high throughput screening and optimization of crystallization conditions (see methods). Initial screening of mCid yielded thin, needle-like crystals (550μm long- figure 7.9) amongst large amounts of precipitate. These crystals formed after 7 days in 10% PEG3350, 0.2M HEPES pH7.5, 0.2Mproline (21°C). PEG optimization in-

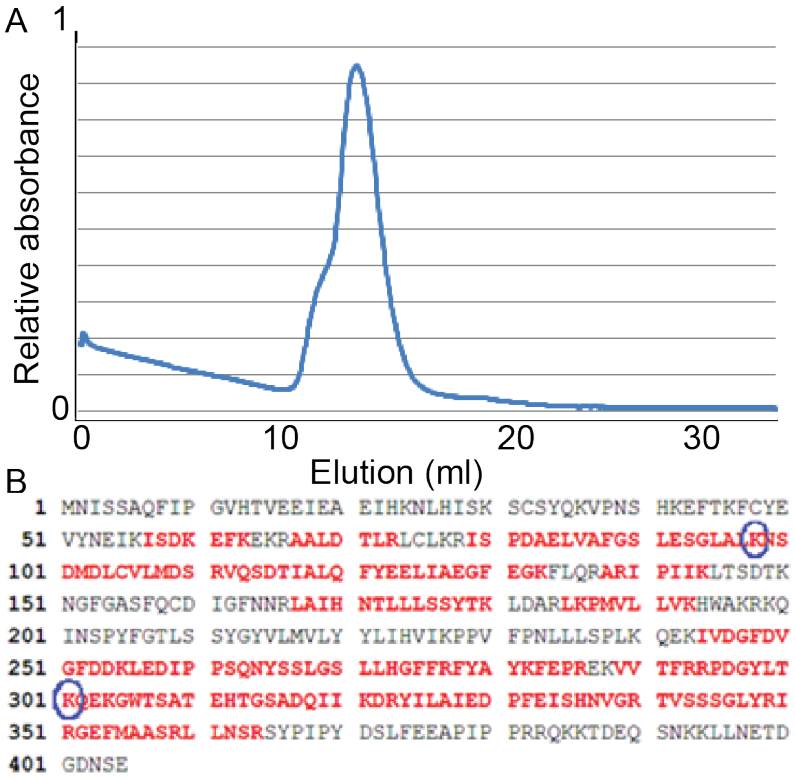


Figure 7.8: *Methylated mCid production. A) size exclusion chromotography analysis of methylate mCid revealed an asymmetric trace with 2 species. b) LC MS-MS results from methylation trials revealed a 53% coverage (red) and two lysines that had been successfully methylated (blue circles).*

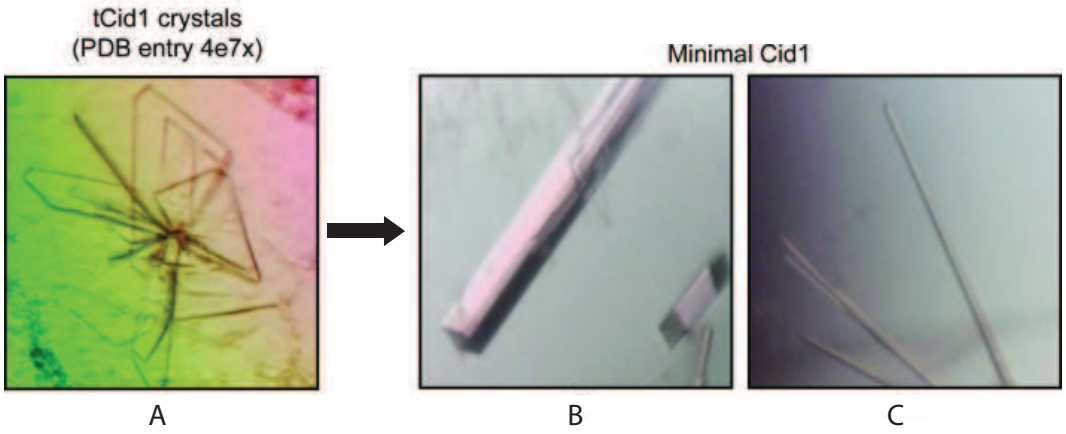


Figure 7.9: *A) Original tCid crystals. B) step by step optimization of mCid yielded better quality crystals. C) The best crystals were obtained through the dilution screen*

creased the clarity of the drop and speeded the formation of crystals from 7 days to 2-4 days. These crystals were somewhat shorter (250 μm long) and were obtained in a 1:2 ratio of protein: reservoir, with 10.5% PEG3350, 0.2M HEPES pH7.5, 0.2M proline. Observation under a light microscope revealed a few short, rectangular crystals. Concurrent with this optimization, a three way dilution screen was carried out (methods). This screen yielded larger, more regular crystals (figure 7.9) grown in 10% PEG3350, 0.2M HEPES pH7.5, 0.2M proline (21°C) diluted 21% with distilled water at a 1:1 ratio of protein to reservoir. Crystals were grown in 3-5 days and diffracted at Diamond synchrotron beamline I03.

Interestingly, the initial crystallization 'hit' contains proline, which was used as an additive in the crystallization of tCid also, to yield a 2.6 Å structure [Yates et al., 2012]. The mechanism of action for this is not clear; a recent additive screening exercise which combined 200 proteins and several hundred additives failed to find proline addition to be significant [McPherson and Cudney, 2006]. However, an informative study by Samuel et al. [Samuel et al., 2000] found proline to inhibit protein aggregation and to trap their protein

of interest (lysozyme) in one conformation, the group also proposed that at higher concentrations proline may stabilize amphipathic protein regions. This was also true for tCid1, where proline stabilized an unusual conformation of the enzyme, which may represent a product ejection state [Yates et al., 2012]. It is possible that the propensity of Cid1 to crystallize in the presence of proline is due to this role as a molecular chaperone. I therefore set up a further 700 trials with 0.2M proline in the crystallization buffer.

Finally, screening of the mCid RNA complex also yielded crystals at 4°C in two conditions; 1) 0.2 M Sodium Chloride, 0.1 M Citrate pH 5.5, 1.0 M di-Ammonium Hydrogen Phosphate, pH 7.6 and 2) 50% 2-Methyl-2,4-Pentanediol, 0.2 M Ammonium di-Hydrogen Phosphate, 0.1 M Tris-HCl pH 8.5. However, upon diffraction these crystals were revealed to be composed of salt.

7.2.4 mCid diffraction

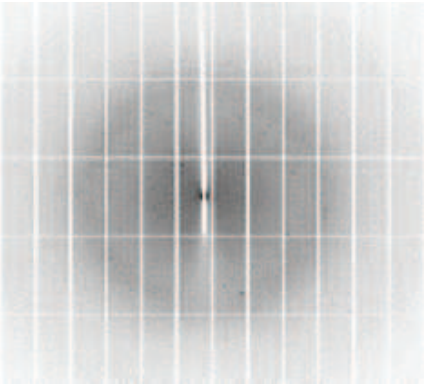


Figure 7.10: *Diffraction pattern of mCid crystals from Diamond Light Source beamline I03, UK*

Crystallographic data analysis of mCid revealed an improved resolution of up to 2.1 Å resolution data and that crystals are of space group C2 and unit cell parameters a=157.11, b=54.84, c=119.56 , α=γ=90° and β=125°. From these data I observe that the β-trapdoor present in tCid crystals is lacking in mCid structures (figure 7.11). Interestingly this β-trapdoor was also absent in the previously published Cid structures [Lunde

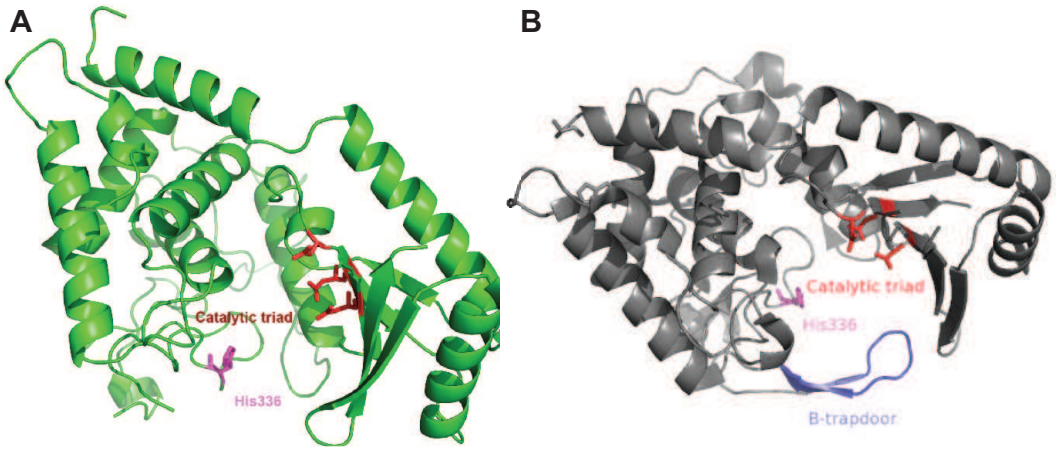


Figure 7.11: *mCid (A) and tCid (B)) structures. The mCid (green) structure does not contain a β-trapdoor, unlike the tCid.*

et al., 2012], and may suggest that it is inherently flexible/unstable. The conformation observed in mCid crystals appears to be the catalytically closed form, which may not be suitable for crystal soaking, however, I have described the improved diffraction of Cid, and present a foundation for future studies of Cid1 and RNA binding [Yates et al., 2015a].

The approach I have employed here; consisting of proteolytic fragmentation Cid1 (tCid) interest, followed by its crystallization and finally re-engineering the construct (to mCid, based on observable electron density), may be more widely applicable for rational approaches to protein crystallization. Indeed, this approach was later employed successfully in-group to determine crystals of Cid in an open conformation. Surface entropy reduction was successfully employed through mutagenesis of Cid RNA binding domains to yield high resolution crystals of 1.7 Å [Yates et al., 2015a, Yates et al., 2015b].

7.3 NMR of mCid

The difficulty in obtaining crystal structures of Cid with highly mobile RNA moved the investigation to alternative techniques complimentary to X-ray crys-

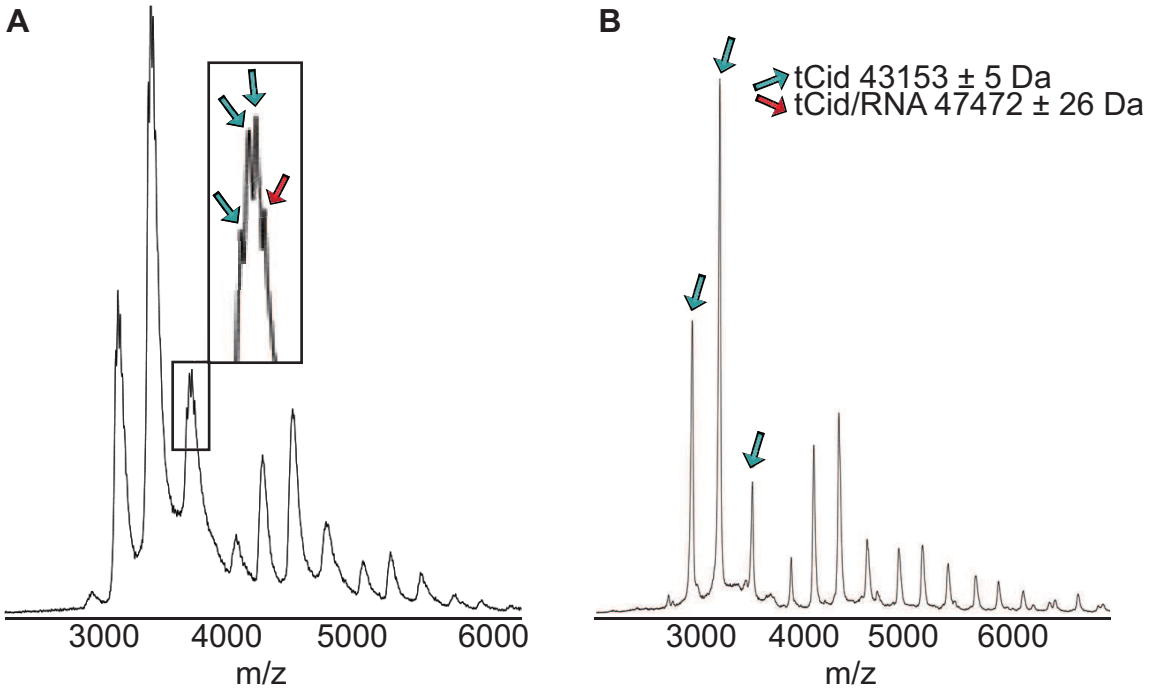


Figure 7.12: *MS studies of Cid1. A Cid/RNA complex is observed by MS, increased collision energy (B) causes this complex to dissociate.*

tallography. To begin studies of Cid it was first necessary to find conditions where Cid would be stable by NMR. To this end, crystallization conditions were re-examined for environments where Cid *did not* crystallize or precipitate. A number of factors were repeatably seen to preclude the precipitation of Cid, most notably, succinate and magnesium. After identifying conditions under which Cid was stable, the search was then fine tuned by employing a multi-factorial, high throughput stability screen. Such a screen employs hydrophobic region binding dye SYPROorange (Bio-rad, California, USA) to report on exposed hydrophobic residues (indicative of unfolding), temperature is increased over time and the level of fluorescence produced coarsely corresponds to the degree of unfolding in a technique known as differential scanning fluorimetry. By varying salt, magnesium, succinate and pH together it is possible to ex-

plore the energy landscape of Cid and find conditions where it is most stable. Inspection of figure 7.13A suggests that Cid is most stable in 100mM NaCl, 200mM succinate, 200mM Magnesium formate, pH7.6 at temperatures below 60°C. However, mixing of Cid1 and RNA under these conditions led to the rapid precipitation of Cid1 and spectra that are uninterpretable.

Native MS study of Cid1 and complexed RNA (see methods) revealed co-complexes observable by MS. Increasing collision energy destabilized these complexes yielding M/Z peaks for apo Cid1 and RNA, consistent with their theoretical mass (figure 7.12) and confirming the ability of Cid1 to bind RNA.

7.4 Expression and purification of mammalian TUTase domains

Cid1 displays significant homology to the more complex mammalian TUTases, however, the basic RNA binding patch on Cid1 appears to be RNA promiscuous when compared to the highly specific zinc fingers and zinc knuckles found in the mammalian TUTases. Therefore, to study mammalian TUTase function and structure it is necessary to produce soluble protein. Given the small and dynamic nature of zinc fingers and their neighboring regions (see figure 7.5), the structural study of TUTase sub-domains may be better suited to NMR.

A total of 8 constructs were made, 4 from TUT4 and 4 from TUT7 (figure 7.14). Each TUT had a construct corresponding to a isolated zinc knuckle, a double zinc knuckle linked by a disordered region, and finally the double zinc knuckle, disordered region and Cid1-like region together with the N-terminal zinc knuckle (see methods). A range of conditions were then assayed for the production of soluble protein (table 7.1). The only condition which yielded sig-

Media: Terrific Broth, Lysogeny Broth, Minimal media (M9)
Additives: Trace metals, Zinc chloride, Magnesium chloride
Inductions temperatures: 37°C, 25°C, 18°C
Induction systems:: IPTG, auto-induction
Cell lines: BL21-Gold, BL21-Origami, BL21-Star, BL21-SHuffle
Purification tags : Hexahis, Maltose binding protein, Glutathione, Periplasmic expression tags

Table 7.1: *Factors varied in expression screening for soluble TUTase domains*

nificant amounts of soluble protein was auto-inducing LB media supplemented with 5mM ZnCl. This yielded soluble, purifiable, hexa-his tagged protein (figure 7.14).

Whilst the LB expression of TUT domains yields soluble, purifiable protein, the minimal media expression fails to produce soluble protein. Further optimization of conditions in minimal media may result in appropriate conditions to produce sample for NMR studies.

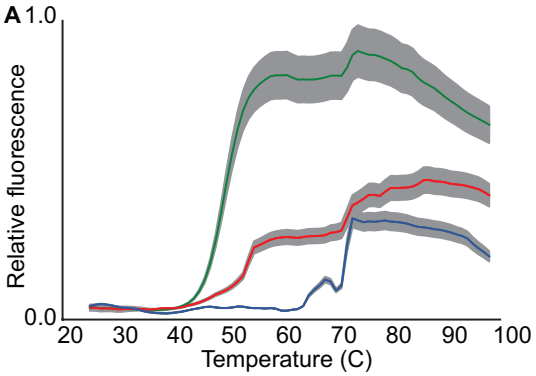
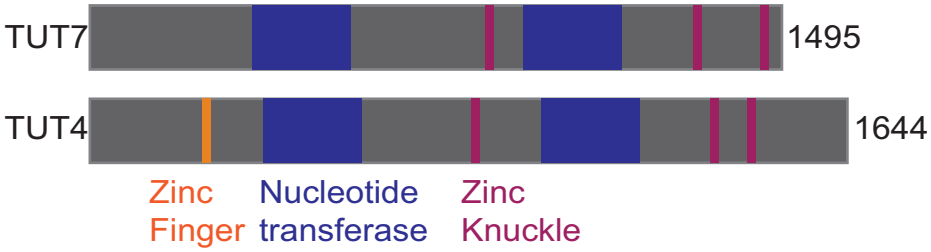


Figure 7.13: *Stability studies of Cid A) Thermofluor analysis was employed to find buffers suitable for NMR studies. Over 500 conditions were trialled, 3 of which are shown for illustration: 250mM NaCl (green) pH6, 250mM NaCl 50mM succinate pH8 (red) and 100mM NaCl, 200mM succinate, 200mM Magnesium formate, pH7.6 (purple). Grey areas represent error.*

7.5 Discussion

Over 3000 crystallization trials of mCid with various cofactors and conditions have yielded crystals for diffraction. It was thought that obtaining high quality x-ray diffraction data would allow the assignment of key water molecules involved in UTP binding and catalysis. However, whilst improved crystallographic data was

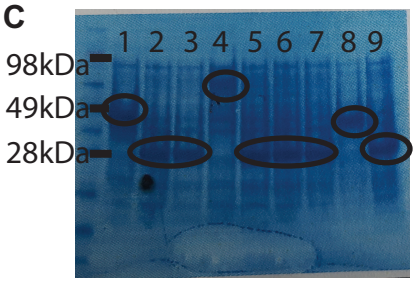
A



B

Construct	Zn knuckles #	Cid-like domains	Size (kDa)	Expresses?	Purifiable?
TUT7 #1	1 (1st)	0	4	Y	N
TUT7 #2	1 (2nd)	0	4	Y	N
TUT7 #3	2 (1st,2nd)	0	24	Y	Y
TUT7 #4	3 (1st,2nd,3rd)	1	60	Y	N
TUT4 #1	1 (1st)	0	4	Y	N
TUT4 #2	1 (2nd)	0	4	Y	N
TUT4 #3	2 (1st,2nd)	0	26	Y	Y
TUT4 #4	3 (1st,2nd,3rd)	1	62	Y	Y

C



D

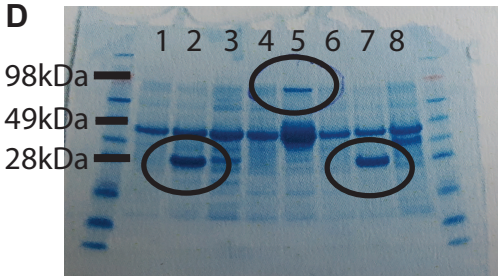


Figure 7.14: *TUT constructs and purification A) Overall structure of TUT4 and TUT7, constructs are outlined in black. B) Table of constructs made, components and size (knuckles labeled C to N term.). C) SDS PAGE of successful expression trials of GST-fused constructs (black ellipses) D) SDS-PAGE of successful purification (black ellipses) from LB media of both tandem zinc knuckles and the larger catalytic fragment (corresponding to TUT7#4), tandem constructs were hexahistidine fused whilst the larger construct was GST and hexahistidine fused. Selected molecular weight markers are shown as black bars*

obtained, it did not elucidate the catalytic role of water in Cid1.

The rational strategies employed here to improve diffraction data consisted of truncating the enzyme in a manner consistent with the observed tCid electron density. This strategy be more widely applicable [Yates et al., 2015a]. Furthermore, the presence and absence of the β-trapdoor in tCid and mCid (and other published structures) may represent different functional states of Cid. If the β-trapdoor represents a mechanism to hold substrates, as has been previously suggested [Yates et al., 2012], then it is possible that Cid1 states associated with pre and post reaction are being observed.

The methodology here, specifically the attempts at Cid1 methylation, were adapted and applied successfully to Glucocerebrosidase (see chapter 3). Rather than methylating amine side chains to reduce surface entropy and improve crystallization conditions, as common in crystallography [Rayment, 1997], I used ^{13}C reagents to introduce NMR observable nuclei onto a large (60 kDa) human produced protein complete with post translational modifications. Adaptation of this technique was the basis for chapters 3 and 4, as well as a research article currently in review (Barber et al. PNAS. 2016).

Future work on the TUTase project will entail the determination high resolution Cid structures with various ligands. A potential use of Cid1 is in screening drug candidates for the homologous human proteins TUT4 and TUT7, implicated in several cancers [Hayes et al., 2014, Molenaar et al., 2012, Heo et al., 2009, Hagan et al., 2009, Heo et al., 2008, Wang et al., 2000]. This may require further optimization of crystallization conditions for soaking experiments. Similarly, further work improving the minimal media expression of TUT domains should yield protein for NMR studies of TUTase associated zinc knuckles and their interaction with RNA .

7.6 Methods

7.6.1 Cloning and expression of recombinant Cid1

Cid1 is an S. pombe nucleotidytransferase of 451 residues. A previously used 'truncated Cid' (tCid) construct was crystallized and solved, using these data a 'minimal Cid' (mCid) construct was built containing only the residues seen in the crystallographic structure (residues 41-377). mCid was cloned using the InFusion (Clontech) ligation independent system into the Oxford protein Production Facility (OPPF) designed pOPINF vector [Berrow et al., 2007] that encodes an N-terminal hexahistidine tag with a 3C cleavage site. The pOPINF-mCid1 construct was subsequently transformed into BL21 (DE3) Star E. coli (Novagen). Transformants were selected for by overnight incubation at 37°C on Luria broth (LB) 66 agar containing 50 g ml Ampicillin. Single colonies were picked and used to inoculate 10ml liquid (with 50 g ml Ampicillin) shaking at 37 °C, 200rpm for 8 hours. This LB broth was then used to inoculate a starter-culture of 300ml Terrific Broth (TB) (with 50mg/ml Ampicillin) which was grown overnight at 37°C, 200 rpm. Finally, 50ml aliquots of starter-culture were used to inoculate 1L TB (1:20 dilution) supplemented with 50mg/ml Ampicillin and autoinduction solutions (5g/L glycerol, 0.5g/L glucose, 2g/L lactose). Cultures were grown at 37°C for 5 hours at 200 rpm, where upon temperature and speed were reduced to 18°C and 180rpm respectively for 12-16 hours.

TB grown E. coli cultures were pelleted with a Beckman (USA) centrifuge at 7,000 x g, at 4°C. Cell pellets were resuspended in 50mM HEPES, 150mM NaCl, 20mM imidazole, pH7 with one EDTA free protease inhibitor tablet, 2000U DNase I and 1mg of hen egg lysozyme per 8g cells (equivalent to the

cell mass from 1L TB). Cell disruption was achieved by sonication (10 cycles of 10s pulses, 20s cooling, repeated three times) whilst cells were cooled in an ice bath. Resulting lysate was clarified by high speed centrifugation (60,000 x g, 30 minutes, at 4°C) and syringe filtering. Clarified supernatant was incubated with pre-equilibrated nickel chelated sepharose beads overnight at 4°C (with magnetic stirrer). Beads were washed with 15× column volumes of HEPES buffer (50mM HEPES, 150mM NaCl, 20mM imidazole, pH7) on a gravity flow column (BioRad) and then eluted with a stepwise gradient of imidazole (steps of 50mM ,100 mM, 300 mM, 500 mM and 800 mM), Elutions were fractionated and these fractions were analysed by SDS PAGE for protein content.

Subsequently, elution fractions containing mCid1 were pooled and concentrated to <3ml for heparin affinity chromatography using an AKTA FPLC and a 5ml HP Heparin column (GE Healthcare, UK). mCid was adsorbed to the heparin matrix and eluted with a linear NaCl gradient from 150mM NaCl to 2M NaCl (at 38mM/ml/min). mCid1 elution occurred at 800mM NaCl. To obtain homogenous mCid and assess hydrodynamic behavior of the protein, mCid1 was applied to an S75 (10/30) size exclusion column in 20mM HEPES, pH 7.0, 150mM NaCl using an AKTA FPLC at a flow rate of 0.4ml/min. Eluant corresponding to mCid was analysed by SDS PAGE, pooled, buffer-exchanged (into crystallization buffer;12mM HEPES, 150mM NaCl, 1mM TCEP, pH 7.0) and concentrated to 9mg/ml (225mM). Concentrated protein was then divided into aliquots and flash frozen by plunge freezing in liquid nitrogen, and stored at - 80°C until further use.

The truncated Cid variant was also produced and purified as previously [Yates et al., 2012].

Hampton:	Crystal screens I and II, PEG Ion, Grid Screen 6000, Grid Screen Ammonium Sulphate, Natrix, Crystal Screen Cryo, Grid screen PEG LiCl, Grid Screen Sodium Chloride, Grid Screen MPD, Quick Screen, Index Screen, SaltRX screen
Emerald:	Wizard Screen 1 and 2
Jena Bioscience:	NucPro Screen
Molecular dimensions:	PACTpremier

Table 7.2: *Commercially available screens used in the enhanced crystallization of Cid1*

7.6.2 Crystallography rationale for Cid1

Many matrices from many companies exist to explore crystallization space in an attempt to find conditions in which homogenous, singular crystals may form. Common parameters include pH, salt (concentration and compound), temperature and precipitant. As well as exploring crystallization space through crystallization trials, a protein directed strategy can be employed to maximize probability of crystallization. Both strategies were employed in this study and are described below.

7.6.2.1 Crystallography matrices

Initial crystallization screening (sitting drop vapour diffusion) utilized commercially available screens (table 7.2) at both 4 and 21°C (totaling 1500 conditions) in 96-well plates (Greiner Bio-one's (Austria) CrystalQuick plates). Thus, 100nl mother liquor was added to 100nl sitting drops of 9mg/ml mCid and equilibrated against a 100μl reservoir. This was achieved using a high throughput methodology developed at the OPPF [Mayo et al., 2005, Walter et al., 2005]. Briefly, reservoirs were dispensed with a Hydra 96 microdispenser (Matrix Technologies, USA) robot. A Cartesian Technologies Mi-

crosys MIC400 nanodrop robot (Genomic Solutions, UK) was then employed to dispense protein and reservoir thus forming the sitting drop. Crystallization plates were subsequently sealed with VIEWSEAL (Greiner io-one, Austria) adhesive covers. Crystallization trails were then held and imaged using an automated storage vault (The Automation Partnership) or a Rock Imager 1000 (Formulatrix, SA) at 21°C and 4°C, respectively. The automated storage vaults have the advantage of automated drop imaging; these images are then deposited into XtalPIMs (OPPF).

A further 500 conditions were trialled at 4°C only for a mCid/RNA complex. To form a complex between mCid and RNA, 200mM mCid (in crystallization buffer), 1mM UMPcPP (a non hydrolysable TP analogue) and 5mM magnesium was incubated overnight at 4 C, followed by the addition of 240mM of 13-mer RNA (5-GAUUUGACGUUGGUUUUGACG-3) at a 1:1.2 molar ratio together with RNase inhibitor (RNasin, Promega, USA) and left to equilibrate at room temperature for 2-3 hours prior to crystallization screening.

7.6.2.2 Protein directed crystallization strategies

Additionally, I applied surface entropy reduction (SER) strategies to mCid. SER has been previously employed to yield high quality crystallographic data or to crystallize proteins recalcitrant to crystallization [Rayment, 1997, Walter et al., 2006, Cooper et al., 2007]. SER can alter space groups and other properties of crystals to yield better data, it also allows the exploration of crystallization space from a new starting location and can improve resolution by reducing the thermal motion of exposed side chains. To these ends I employed two distinct strategies. First, mCid methylation, using an established protocol from the OPPF; 5mg of mCid (at 0.5mg/ml) was incubated

with step wise addition of 20μM dimethylamine borane complex (ABC) and 40μM formaldehyde for two sets of two hours and then overnight. The reaction mixture was then centrifuged to remove precipitant and the reaction quenched with 20mM Tris, 200mM NaCl, pH7 on an AKTA FPLC size exclusion column (see results). Methylation of surface lysines increases their hydrophobicity and therefore reduce thermal motion, promoting crystallization.

A second strategy utilized was site directed mutagenesis. The previously determined crystal structure revealed several basic patches for RNA binding and these were mutated for biochemical studies during this study [Yates et al., 2012]. However, these mutants may also provide SER mutants. Therefore constructs containing mutations of the RNA binding sites were used as templates to make two mutant mCid variants, mCid1/3 (corresponding to the first and third putative RNA binding domains - K133A, R137A, R277A, K282A) and mCid 2/3 (corresponding to the second and third putative RNA binding domains -R277A, K282A, K321A, R323A).

The cloning of these mutants was achieved with mCid forward primer (CACAAGGAATTTACGAAGTTTTGC) and reverse primer (GGCCTCCT-CAAATAATGAATCATAA) on a mutated DNA. Mutant DNA (100ng) was incubated with 100ng of pure, cut, linear POPINF (cut with HindIII and KpnI at 37°C for 1 hour) and one aliquot of lyophilized InFusion enzyme (Clontech, USA) in a 100μl reaction volume for 30 minutes at 42°C. The reaction mixture is then diluted with Tris-EDTA (TE) buffer (10mM Tris, 1mM EDTA) and transformed in D5α cells which are incubated at 4°C overnight on LB agar plates with 50μg/ml Ampicillin, 0.5mM X-GAL and 1mM IPTG. White colonies were then picked and grown overnight in liquid media at 37°C, 200rpm. Plasmid was then extracted using QIAgen's (Netherlands) qiaprep miniprep

kit according to manufacturer instructions. This plasmid could then be transformed into BL21 (DE3) cells and purified with the same method as mCid. The rationale behind mutation of surface exposed charged groups is similar to that of methylation; it reduces surface entropy allowing for different crystallization conditions.

7.6.2.3 Optimization and seeding of crystals

High throughput screening of crystallization conditions is a coarse grained method of exploring crystallization space, therefore optimization methods are often applied to find finer conditions. Several optimization strategies exist; some of the more popular include additive screening (the addition of small molecules to crystallization conditions to alter protein-protein contacts, viscosity, charges etc. [Jancarik et al., 2004]) dilution screening (altering the ratio of protein to reservoir, or the concentration of the reservoir to move around the phase diagram [Chayen and Saridakis, 2008]. A final method is factorial optimization, where conditions that yielded crystals have their component parts (pH, salt, precipitant) altered to produce a very fine grain factorial trial. Of these popular methods the first two have been implemented [Chayen and Saridakis, 2008].

The additive screen used is available commercially (Hampton Research) and was used for two conditions that yielded crystals, Hampton's Index 6 (10% PEG3350, 0.1M HEPES, 0.2M proline) and Hampton's PEG Ion 10 (20% PEG3350 0.2M Sodium iodide). The second optimization strategy undertaken consisted of a dilution screen. The ratio of protein:reservoir was varied (1:1, 1:2, 2:1), as was the concentration of the reservoir (100-64% in steps of 3%) in a 96 well sitting drop plate, utilizing 36 wells. To maximize exploration

of the phase diagram mCid, dilution screens were implemented with mCid at both 9mg/ml and 6mg/ml. A final optimization strategy consisted of a partial factorial experiment where PEG3350 concentration was varied from 10.5% to 0.5% in steps of five, this also used variable protein: reservoir ratios (1:1, 1:2, 2:1). I identified proline as a possible aid to crystallization from initial screening (see crystallization matrix strategy) and previous observation of tCid, therefore a further 700 conditions were trialled with 0.2M proline added to the crystallization buffer, with mCid at 6mg/ml.

Seeding was achieved with the cat's whisker approach (an established technique [Luang et al., 2010]). First, the screening condition was altered (0.1M HEPES pH 7.5, 0.2M proline, 5% PEG3350) to reduce precipitant and therefore discourage spontaneous crystallization. The ratio of protein: reservoir was also explored (1:1, 1:2, 1:3, 1:4) and this was left to equilibrate overnight. A cat's whisker was then used to transfer needle-like mCid crystals from a stock to the sitting drops by gently streaking the whisker across the stock solution and then the experimental conditions. This also served to explore dilution space as the whisker was refreshed with crystals only after every 8 wells.

7.6.3 Mass spectrometry

LCMSMS and MS collection and analysis was as previous (see chapter 3 methods).

7.6.4 TUTase expression

TUTase constructs were cloned using the InFusion (Clontech) ligation independent system into the Oxford protein Production Facility (OPPF) designed

pOPINF vector [Berrow et al., 2007] that encodes either an N-terminal hexahistidine tag, a maltose binding protein tag, or a glutathione tag, each with a 3C cleavage site and optional periplasmic targeting sequence. Cloned constructs were introduced into a variety of cell lines, sourced from New England Biolabs, Mass, USA, through heat shock transformation methods.

Expression trials were carried out in TB, LB or minimal media (M9) following the EMBL procedure [EMBL,]. Components listed in 7.1 were assayed over a coarse grid search and evaluated by SDS-PAGE on whole cell fractions, soluble and insolble fractions. Promising conditions were then purified. Cells were lysed with BPER reagent [Information, 2007] and applied to nickel chelated sepharose beads in a batch method, washed, eluted with high concentration imidazole (1M) and run on SDS-PAGE for analysis.

Bibliography

[Abragam and Goldman, 1982] Abragam, a. and Goldman, M. (1982). Nuclear magnetism: Order and disorder. *Nuclear magnetism: Order and disorder*, pages +626.

[Ahmad et al., 2012] Ahmad, B., Chen, Y., and Lapidus, L. J. (2012). Aggregation of α-synuclein is kinetically controlled by intramolecular diffusion. *Proceedings of the National Academy of Sciences of the United States of America*, 109(7):2336–41.

[Akter et al., 2016] Akter, R., Cao, P., Noor, H., Ridgway, Z., Tu, L.-H., Wang, H., Wong, A. G., Zhang, X., Abedini, A., Schmidt, A. M., and Raleigh, D. P. (2016). Islet Amyloid Polypeptide: Structure, Function, and Pathophysiology. *Journal of diabetes research*, 2016:2798269.

[Alcalay et al., 2015] Alcalay, R. N., Levy, O. A., Waters, C. C., Fahn, S., Ford, B., Kuo, S.-H., Mazzoni, P., Pauciulo, M. W., Nichols, W. C., Gan-Or, Z., Rouleau, G. A., Chung, W. K., Wolf, P., Oliva, P., Keutzer, J., Marder, K., and Zhang, X. (2015). Glucocerebrosidase activity in Parkinson's disease with and without GBA mutations. *Brain : a journal of neurology*, 138(Pt 9):2648–58.

[Alvarez-Erviti et al., 2011] Alvarez-Erviti, L., Seow, Y., Schapira, A. H., Gardiner, C., Sargent, I. L., Wood, M. J. A., and Cooper, J. M. (2011). Lysosomal dysfunction increases exosome-mediated alpha-synuclein release and transmission. *Neurobiology of disease*, 42(3):360–7.

[Alzheimer Disease International,] Alzheimer Disease International. World Alzheimer Report 2015: The Global Impact of Dementia — Alzheimer's Disease International.

[Amiel et al., 2012] Amiel, J., de Pontual, L., and Henrion-Caude, A. (2012). Chapter One - miRNA, Development and Disease. In Theodore Friedmann, J. D. and Goodwin, S., editors, *Advances in Genetics*, volume Volume 80, pages 1–36. Academic Press.

[Anderson et al., 2006] Anderson, M. J., Hansen, C. L., and Quake, S. R. (2006). Phase knowledge enables rational screens for protein crystallization. *Proceedings of the National Academy of Sciences of the United States of America*, 103(45):16746–51.

[Antonio Gramsci, 1937] Antonio Gramsci (1937). *An analysis of the relations of force.*

[Aphasizhev and Aphasizheva, 2008] Aphasizhev, R. and Aphasizheva, I. (2008). Terminal RNA uridylyltransferases of trypanosomes. *Biochimica et biophysica acta*, 1779(4):270–280.

[Aphasizhev and Aphasizheva, 2011] Aphasizhev, R. and Aphasizheva, I. (2011). Uridine Insertion/Deletion Editing In Trypanosomes: A Playground for RNA-guided Information Transfer. *Wiley interdisciplinary reviews. RNA*, 2(5):669–685.

[Aravin et al., 2006] Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M. J., Kuramochi-Miyagawa, S., Nakano, T., Chien, M., Russo, J. J., Ju, J., Sheridan, R., Sander, C., Zavolan, M., and Tuschl, T. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*, 442(7099):203–207.

[Arosio et al., 2015] Arosio, P., Knowles, T. P. J., and Linse, S. (2015). On the lag phase in amyloid fibril formation. *Physical chemistry chemical physics : PCCP*, 17(12):7606–18.

[Arriagada et al., 1992] Arriagada, P. V., Marzloff, K., and Hyman, B. T. (1992). Distribution of Alzheimer-type pathologic changes in nondemented elderly individuals matches the pattern in Alzheimer's disease. *Neurology*, 42(9):1681–8.

[Ashfield et al., 2000] Ashfield, J. T., Meyers, T., Lowne, D., Varley, P. G., Arnold, J. R., Tan, P., Yang, J. C., Czaplewski, L. G., Dudgeon, T., and Fisher, J. (2000). Chemical modification of a variant of human MIP-1alpha; implications for dimer structure. *Protein science : a publication of the Protein Society*, 9(10):2047–53.

[Bae et al., 2015] Bae, E.-J., Yang, N. Y., Lee, C., Lee, H.-J., Kim, S., Sardi, S. P., and Lee, S.-J. (2015). Loss of glucocerebrosidase 1 activity causes lysosomal dysfunction and α-synuclein aggregation. *Experimental & molecular medicine*, 47(3):e153.

[Bae et al., 2014] Bae, E.-J., Yang, N.-Y., Song, M., Lee, C. S., Lee, J. S., Jung, B. C., Lee, H.-J., Kim, S., Masliah, E., Sardi, S. P., and Lee, S.-J.

(2014). Glucocerebrosidase depletion enhances cell-to-cell transmission of α-synuclein. *Nature communications*, 5:4755.

[Bakthisaran et al., 2015] Bakthisaran, R., Tangirala, R., and Rao, C. M. (2015). Small heat shock proteins: Role in cellular functions and pathology. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1854(4):291–319.

[Baldwin, 2014] Baldwin, A. J. (2014). An exact solution for R2,eff in CPMG experiments in the case of two site chemical exchange. *Journal of Magnetic Resonance*, 244:114–124.

[Baldwin et al., 2009] Baldwin, A. J., Hansen, D. F., Vallurupalli, P., and Kay, L. E. (2009). Measurement of methyl axis orientations in invisible, excited states of proteins by relaxation dispersion NMR spectroscopy. *Journal of the American Chemical Society*, 131(33):11939–11948.

[Baldwin et al., 2011a] Baldwin, A. J., Hilton, G. R., Lioe, H., Bagnéris, C., Benesch, J. L. P., and Kay, L. E. (2011a). Quaternary dynamics of αb-crystallin as a direct consequence of localised tertiary fluctuations in the C-terminus. *Journal of Molecular Biology*, 413(2):310–320.

[Baldwin and Kay, 2009] Baldwin, A. J. and Kay, L. E. (2009). NMR spectroscopy brings invisible protein states into focus. *Nature chemical biology*, 5(11):808–814.

[Baldwin and Kay, 2012] Baldwin, A. J. and Kay, L. E. (2012). Measurement of the signs of methyl 13C chemical shift differences between interconverting ground and excited protein states by R 1rho: An application to aB-crystallin. *Journal of Biomolecular NMR*, 53(1):1–12.

[Baldwin et al., 2011b] Baldwin, A. J., Knowles, T. P. J., Tartaglia, G. G., Fitzpatrick, A. W., Devlin, G. L., Shammas, S. L., Waudby, C. A., Mossuto, M. F., Meehan, S., Gras, S. L., Christodoulou, J., Anthony-Cahill, S. J., Barker, P. D., Vendruscolo, M., and Dobson, C. M. (2011b). Metastability of native proteins and the phenomenon of amyloid formation. *Journal of the American Chemical Society*, 133(36):14160–14163.

[Baldwin et al., 2010] Baldwin, A. J., Religa, T. L., Hansen, D. F., Bouvignies, G., and Kay, L. E. (2010). 13CHD2 Methyl group probes of millisecond time scale exchange in proteins by 1H relaxation dispersion: An application to proteasome gating residue dynamics. *Journal of the American Chemical Society*, 132(32):10992–10995.

[Ban et al., 2000] Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. a. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science*, 289(August):905–920.

[Banci et al., 2013] Banci, L., Barbieri, L., Bertini, I., Luchinat, E., Secci, E., Zhao, Y., and Aricescu, a. R. (2013). Atomic-resolution monitoring of protein maturation in live human cells by NMR. *Nature Chemical Biology*, 9(5):297–299.

[Bartels et al., 2011] Bartels, T., Choi, J. G., and Selkoe, D. J. (2011). α-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature*, 477(7362):107–10.

[Barton et al., 1991] Barton, N. W., Brady, R. O., Dambrosia, J. M., Di Bisceglie, A. M., Doppelt, S. H., Hill, S. C., Mankin, H. J., Murray, G. J., Parker, R. I., and Argoff, C. E. (1991). Replacement therapy for inherited en-

zyme deficiency–macrophage-targeted glucocerebrosidase for Gaucher's disease. *The New England journal of medicine*, 324(21):1464–70.

[Beavan and Schapira, 2013] Beavan, M. S. and Schapira, A. H. V. (2013). Glucocerebrosidase mutations and the pathogenesis of Parkinson disease. *Annals of medicine*, 45(8):511–21.

[Bendikov-Bar and Horowitz, 2012] Bendikov-Bar, I. and Horowitz, M. (2012). Gaucher disease paradigm: From ERAD to comorbidity. *Human Mutation*, 33(10):1398–1407.

[Bendikov-Bar et al., 2013] Bendikov-Bar, I., Maor, G., Filocamo, M., and Horowitz, M. (2013). Ambroxol as a pharmacological chaperone for mutant glucocerebrosidase. *Blood Cells, Molecules, and Diseases*, 50(2):141–145.

[Bendikov-Bar et al., 2011] Bendikov-Bar, I., Ron, I., Filocamo, M., and Horowitz, M. (2011). Characterization of the ERAD process of the L444P mutant glucocerebrosidase variant. *Blood cells, molecules & diseases*, 46(1):4–10.

[Benesch and Ruotolo, 2011] Benesch, J. L. P. and Ruotolo, B. T. (2011). Mass spectrometry: Come of age for structural and dynamical biology.

[Berg-Fussman et al., 1993] Berg-Fussman, A., Grace, M. E., Ioannou, Y., and Grabowski, G. A. (1993). Human acid beta-glucosidase. N-glycosylation site occupancy and the effect of glycosylation on enzymatic activity. *The Journal of biological chemistry*, 268(20):14861–6.

[Berrow et al., 2007] Berrow, N. S., Alderton, D., Sainsbury, S., Nettleship, J., Assenberg, R., Rahman, N., Stuart, D. I., and Owens, R. J. (2007). A

versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Research*, 35(6).

[Bertoncini et al., 2005] Bertoncini, C. W., Jung, Y.-S., Fernandez, C. O., Hoyer, W., Griesinger, C., Jovin, T. M., and Zweckstetter, M. (2005). Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. *Proceedings of the National Academy of Sciences of the United States of America*, 102(5):1430–5.

[Bettelheim et al., 1999] Bettelheim, F. A., Ansari, R., Cheng, Q. F., and Zigler, J. S. (1999). The mode of chaperoning of dithiothreitol-denatured alpha-lactalbumin by alpha-crystallin. *Biochemical and biophysical research communications*, 261(2):292–7.

[Biancalana and Koide, 2010] Biancalana, M. and Koide, S. (2010). Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochimica et biophysica acta*, 1804(7):1405–12.

[Billings et al., 2005] Billings, L. M., Oddo, S., Green, K. N., McGaugh, J. L., and LaFerla, F. M. (2005). Intraneuronal Abeta causes the onset of early Alzheimer's disease-related cognitive deficits in transgenic mice. *Neuron*, 45(5):675–88.

[Blahna et al., 2011] Blahna, M. T., Jones, M. R., Quinton, L. J., Matsuura, K. Y., and Mizgerd, J. P. (2011). Terminal Uridyltransferase Enzyme Zcchc11 Promotes Cell Proliferation Independent of Its Uridyltransferase Activity. *Journal of Biological Chemistry*, 286(49):42381–42389.

[Bodner et al., 2009] Bodner, C. R., Dobson, C. M., and Bax, A. (2009). Multiple tight phospholipid-binding modes of alpha-synuclein revealed by solution NMR spectroscopy. *Journal of molecular biology*, 390(4):775–90.

[Bodner et al., 2010] Bodner, C. R., Maltsev, A. S., Dobson, C. M., and Bax, A. (2010). Differential phospholipid binding of alpha-synuclein variants implicated in Parkinson's disease revealed by solution NMR spectroscopy. *Biochemistry*, 49(5):862–871.

[Boehm and Slack, 2005] Boehm, M. and Slack, F. (2005). A developmental timing microRNA and its target regulate life span in C. elegans. *Science (New York, N.Y.)*, 310(5756):1954–1957.

[Bousset et al., 2013] Bousset, L., Pieri, L., Ruiz-Arlandis, G., Gath, J., Jensen, P. H., Habenstein, B., Madiona, K., Olieric, V., Böckmann, A., Meier, B. H., and Melki, R. (2013). Structural and functional characterization of two alpha-synuclein strains. *Nature communications*, 4:2575.

[Bouvignies et al., 2011] Bouvignies, G., Vallurupalli, P., Hansen, D. F., Correia, B. E., Lange, O., Bah, A., Vernon, R. M., Dahlquist, F. W., Baker, D., and Kay, L. E. (2011). Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature*, 477(7362):111–114.

[Braun et al., 2014] Braun, A. R., Lacy, M. M., Ducas, V. C., Rhoades, E., and Sachs, J. N. (2014). α-Synuclein-induced membrane remodeling is driven by binding affinity, partition depth, and interleaflet order asymmetry. *Journal of the American Chemical Society*, 136(28):9962–72.

[Breydo et al., 2012] Breydo, L., Wu, J. W., and Uversky, V. N. (2012). Alpha-synuclein misfolding and Parkinson's disease. *Biochim Biophys Acta*, 1822(2):261–285.

[Brumshtein et al., 2010] Brumshtein, B., Salinas, P., Peterson, B., Chan, V., Silman, I., Sussman, J. L., Savickas, P. J., Robinson, G. S., and Futerman, A. H. (2010). Characterization of gene-activated human acid-beta-glucosidase: crystal structure, glycan composition, and internalization into macrophages. *Glycobiology*, 20(1):24–32.

[Brumshtein et al., 2006] Brumshtein, B., Wormald, M. R., Silman, I., Futerman, A. H., and Sussman, J. L. (2006). Structural comparison of differently glycosylated forms of acid-beta-glucosidase, the defective enzyme in Gaucher disease. *Acta crystallographica. Section D, Biological crystallography*, 62(Pt 12):1458–1465.

[Bucciantini et al., 2002] Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M., and Stefani, M. (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*, 416(6880):507–11.

[Buell et al., 2010] Buell, A. K., Dobson, C. M., Knowles, T. P. J., and Welland, M. E. (2010). Interactions between Amyloidophilic dyes and their relevance to studies of amyloid inhibitors. *Biophysical Journal*, 99(10):3492–3497.

[Buell et al., 2014] Buell, A. K., Galvagnion, C., Gaspar, R., Sparr, E., Vendruscolo, M., Knowles, T. P. J., Linse, S., and Dobson, C. M. (2014). Solution conditions determine the relative importance of nucleation and growth pro-

cesses in α-synuclein aggregation. *Proceedings of the National Academy of Sciences of the United States of America*, 111(21):7671–7676.

[Bultron et al., 2010] Bultron, G., Kacena, K., Pearson, D., Boxer, M., Yang, R., Sathe, S., Pastores, G., and Mistry, P. K. (2010). The risk of Parkinson's disease in type 1 Gaucher disease. *Journal of inherited metabolic disease*, 33(2):167–73.

[Bumagina et al., 2010] Bumagina, Z. M., Gurvits, B. Y., Artemova, N. V., Muranov, K. O., Yudin, I. K., and Kurganov, B. I. (2010). Mechanism of suppression of dithiothreitol-induced aggregation of bovine a-lactalbumin by a-crystallin. *Biophysical Chemistry*, 146(2-3):108–117.

[Buter et al., 2008] Buter, T. C., van den Hout, A., Matthews, F. E., Larsen, J. P., Brayne, C., and Aarsland, D. (2008). Dementia and survival in Parkinson disease: a 12-year population study. *Neurology*, 70(13):1017–22.

[Campioni et al., 2010] Campioni, S., Mannini, B., Zampagni, M., Pensalfini, A., Parrini, C., Evangelisti, E., Relini, A., Stefani, M., Dobson, C. M., Cecchi, C., and Chiti, F. (2010). A causative link between the structure of aberrant protein oligomers and their toxicity. *Nature chemical biology*, 6(2):140–7.

[Carver and Richards, 1972] Carver, J. P. and Richards, R. E. (1972). A general two-site solution for the chemical exchange produced dependence of T2 upon the carr-Purcell pulse separation. *Journal of Magnetic Resonance (1969)*, 6(1):89–105.

[Chayen and Saridakis, 2008] Chayen, N. E. and Saridakis, E. (2008). Protein crystallization: from purified protein to diffraction-quality crystal. *Nature Methods*, 5(2):147–153.

[Chiasserini et al., 2015] Chiasserini, D., Paciotti, S., Eusebi, P., Persichetti, E., Tasegian, A., Kurzawa-Akanbi, M., Chinnery, P. F., Morris, C. M., Calabresi, P., Parnetti, L., and Beccari, T. (2015). Selective loss of glucocerebrosidase activity in sporadic Parkinson's disease and dementia with Lewy bodies. *Molecular neurodegeneration*, 10(1):15.

[Choi et al., 2012] Choi, Y. S., Patena, W., Leavitt, A. D., and McManus, M. T. (2012). Widespread RNA 3'-end oligouridylation in mammals. *RNA (New York, N.Y.)*, 18(3):394–401.

[Clark and Nilsson, 2004] Clark, A. and Nilsson, M. R. (2004). Islet amyloid: a complication of islet dysfunction or an aetiological factor in Type 2 diabetes? *Diabetologia*, 47(2):157–69.

[Cohen et al., 2015] Cohen, S. I. A., Arosio, P., Presto, J., Kurudenkandy, F. R., Biverstål, H., Dolfe, L., Dunning, C., Yang, X., Frohm, B., Vendruscolo, M., Johansson, J., Dobson, C. M., Fisahn, A., Knowles, T. P. J., and Linse, S. (2015). A molecular chaperone breaks the catalytic cycle that generates toxic Aβ oligomers. *Nature structural & molecular biology*, 22(3):207–13.

[Cohen et al., 2013] Cohen, S. I. A., Linse, S., Luheshi, L. M., Hellstrand, E., White, D. A., Rajah, L., Otzen, D. E., Vendruscolo, M., Dobson, C. M., and Knowles, T. P. J. (2013). Proliferation of amyloid-β42 aggregates occurs

through a secondary nucleation mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, 110(24):9758–63.

[Comellas et al., 2011] Comellas, G., Lemkau, L. R., Nieuwkoop, A. J., Kloepper, K. D., Ladror, D. T., Ebisu, R., Woods, W. S., Lipton, A. S., George, J. M., and Rienstra, C. M. (2011). Structured Regions of α-Synuclein Fibrils Include the Early-Onset Parkinson's Disease Mutation Sites. *Journal of Molecular Biology*, 411(4):881–895.

[Comellas et al., 2012] Comellas, G., Lemkau, L. R., Zhou, D. H., George, J. M., and Rienstra, C. M. (2012). Structural intermediates during α-synuclein fibrillogenesis on phospholipid vesicles. *Journal of the American Chemical Society*, 134(11):5090–5099.

[Conway et al., 1998] Conway, K. A., Harper, J. D., and Lansbury, P. T. (1998). Accelerated in vitro fibril formation by a mutant alpha-synuclein linked to early-onset Parkinson disease. *Nature medicine*, 4(11):1318–20.

[Cooper et al., 2013] Cooper, C. L., Cosgrove, T., van Duijneveldt, J. S., Murray, M., and Prescott, S. W. (2013). The use of solvent relaxation NMR to study colloidal suspensions. *Soft Matter*, 9(30):7211–7228.

[Cooper et al., 2007] Cooper, D. R., Boczek, T., Grelewska, K., Pinkowska, M., Sikorska, M., Zawadzki, M., and Derewenda, Z. (2007). Protein crystallization by surface entropy reduction: optimization of the SER strategy. *Acta crystallographica. Section D, Biological crystallography*, 63(Pt 5):636–645.

[Cox et al., 2014] Cox, D., Carver, J. A., and Ecroyd, H. (2014). Preventing α-synuclein aggregation: The role of the small heat-shock molecular chaperone

proteins. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1842(9):1830–1843.

[Crick, 1970] Crick, F. (1970). Central dogma of molecular biology. *Nature*, 227(5258):561–563.

[Crowther et al., 2000] Crowther, R., Daniel, S. E., and Goedert, M. (2000). Characterisation of isolated α-synuclein filaments from substantia nigra of Parkinson's disease brain.

[Cuervo et al., 2004] Cuervo, A. M., Stefanis, L., Fredenburg, R., Lansbury, P. T., and Sulzer, D. (2004). Impaired degradation of mutant alpha-synuclein by chaperone-mediated autophagy. *Science (New York, N.Y.)*, 305(5688):1292–1295.

[de Vos et al., 1995] de Vos, R. A., Jansen, E. N., Stam, F. C., Ravid, R., and Swaab, D. F. (1995). 'Lewy body disease': clinico-pathological correlations in 18 consecutive cases of Parkinson's disease with and without dementia. *Clinical neurology and neurosurgery*, 97(1):13–22.

[Dedmon et al., 2005] Dedmon, M. M., Lindorff-Larsen, K., Christodoulou, J., Vendruscolo, M., and Dobson, C. M. (2005). Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *Journal of the American Chemical Society*, 127(2):476–7.

[Delaère et al., 1990] Delaère, P., Duyckaerts, C., Masters, C., Beyreuther, K., Piette, F., and Hauw, J. J. (1990). Large amounts of neocortical beta A4 deposits without neuritic plaques nor tangles in a psychometrically assessed, non-demented person. *Neuroscience letters*, 116(1-2):87–93.

[Derewenda, 2004] Derewenda, Z. S. (2004). Rational Protein Crystallization by Mutational Surface Engineering.

[Dermentzaki et al., 2013] Dermentzaki, G., Dimitriou, E., Xilouri, M., Michelakakis, H., and Stefanis, L. (2013). Loss of β-Glucocerebrosidase Activity Does Not Affect Alpha-Synuclein Levels or Lysosomal Function in Neuronal Cells. *PLoS ONE*, 8(4).

[Desantis et al., 2012] Desantis, M. E., Leung, E. H., Sweeny, E. A., Jackrel, M. E., Cushman-Nick, M., Neuhaus-Follini, A., Vashist, S., Sochor, M. A., Knight, M. N., and Shorter, J. (2012). Operational plasticity enables Hsp104 to disaggregate diverse amyloid and nonamyloid clients. *Cell*, 151(4):778–793.

[Drake and F.L., 2001] Drake, G. v. R. and F.L. (2001). Python language. *Python Reference Manual,*, PythonLabs.

[Drescher et al., 2008] Drescher, M., Godschalk, F., Veldhuis, G., van Rooijen, B. D., Subramaniam, V., and Huber, M. (2008). Spin-label EPR on alpha-synuclein reveals differences in the membrane binding affinity of the two antiparallel helices. *Chembiochem : a European journal of chemical biology*, 9(15):2411–6.

[Drescher et al., 2010] Drescher, M., van Rooijen, B. D., Veldhuis, G., Subramaniam, V., and Huber, M. (2010). A stable lipid-induced aggregate of alpha-synuclein. *Journal of the American Chemical Society*, 132(12):4080–2.

[Duennwald et al., 2012] Duennwald, M. L., Echeverria, A., and Shorter, J. (2012). Small heat shock proteins potentiate amyloid dissolution by protein disaggregases from yeast and humans. *PLoS Biology*, 10(6):e1001346.

[Ecroyd et al., 2008] Ecroyd, H., Koudelka, T., Thorn, D. C., Williams, D. M., Devlin, G., Hoffmann, P., and Carver, J. a. (2008). Dissociation from the oligomeric state is the rate-limiting step in fibril formation by κ-casein. *Journal of Biological Chemistry*, 283(14):9012–9022.

[Eichner and Radford, 2011] Eichner, T. and Radford, S. (2011). A Diversity of Assembly Mechanisms of a Generic Amyloid Fold. *Molecular Cell*, 43(1):8–18.

[Eisenberg and Jucker, 2012] Eisenberg, D. and Jucker, M. (2012). The amyloid state of proteins in human diseases.

[Eisenmesser et al., 2005] Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. a., Skalicky, J. J., Kay, L. E., and Kern, D. (2005). Intrinsic dynamics of an enzyme underlies catalysis. *Nature*, 438(7064):117–121.

[Eliezer et al., 2001] Eliezer, D., Kutluay, E., Bussell, R., and Browne, G. (2001). Conformational properties of alpha-synuclein in its free and lipid-associated states. *Journal of molecular biology*, 307(4):1061–1073.

[EMBL,] EMBL. EMBL Minimal media recipe.

[et al. Eric Jones, Travis Oliphant, 2001] et al. Eric Jones, Travis Oliphant (2001). SciPy: Open Source Scientific Tools for Python.

[Fabian et al., 2010] Fabian, M. R., Sonenberg, N., and Filipowicz, W. (2010). Regulation of mRNA Translation and Stability by microRNAs. *Annual Review of Biochemistry*, 79(1):351–379.

[Farazi et al., 2013] Farazi, T. A., Hoell, J. I., Morozov, P., and Tuschl, T. (2013). MicroRNAs in human cancer. *Advances in experimental medicine and biology*, 774:1–20.

[Fauvet et al., 2012] Fauvet, B., Mbefo, M. K., Fares, M.-B., Desobry, C., Michael, S., Ardah, M. T., Tsika, E., Coune, P., Prudent, M., Lion, N., Eliezer, D., Moore, D. J., Schneider, B., Aebischer, P., El-Agnaf, O. M., Masliah, E., and Lashuel, H. A. (2012). α-Synuclein in central nervous system and from erythrocytes, mammalian cells, and Escherichia coli exists predominantly as disordered monomer. *The Journal of biological chemistry*, 287(19):15345–64.

[Fitzpatrick et al., 2013] Fitzpatrick, A. W. P., Debelouchina, G. T., Bayro, M. J., Clare, D. K., Caporini, M. A., Bajaj, V. S., Jaroniec, C. P., Wang, L., Ladizhansky, V., Muller, S. A., MacPhee, C. E., Waudby, C. A., Mott, H. R., De Simone, A., Knowles, T. P. J., Saibil, H. R., Vendruscolo, M., Orlova, E. V., Griffin, R. G., and Dobson, C. M. (2013). Atomic structure and hierarchical assembly of a cross-Beta amyloid fibril. *Proceedings of the National Academy of Sciences of the United States of America*, 110(14):5468–5473.

[Franck et al., 2004] Franck, E., Madsen, O., Van Rheede, T., Ricard, G., Huynen, M. A., and De Jong, W. W. (2004). Evolutionary diversity of vertebrate small heat shock proteins. *Journal of Molecular Evolution*, 59(6):792–805.

[Fusco et al., 2014] Fusco, G., De Simone, A., Gopinath, T., Vostrikov, V., Vendruscolo, M., Dobson, C. M., and Veglia, G. (2014). Direct observa-

tion of the three regions in α-synuclein that determine its membrane-bound behaviour. *Nature communications*, 5:3827.

[Galvagnion et al., 2015] Galvagnion, C., Buell, A. K., Meisl, G., Michaels, T. C. T., Vendruscolo, M., Knowles, T. P. J., and Dobson, C. M. (2015). Lipid vesicles trigger α-synuclein aggregation by stimulating primary nucleation. *Nature chemical biology*, 11(3):229–234.

[Gast et al., 1995] Gast, K., Damaschun, H., Eckert, K., Schulze-Forster, K., Maurer, H. R., Müller-Frohne, M., Zirwer, D., Czarnecki, J., and Damaschun, G. (1995). Prothymosin alpha: a biologically active protein with random coil conformation. *Biochemistry*, 34(40):13211–8.

[Gazit, 2002] Gazit, E. (2002). A possible role for pi-stacking in the self-assembly of amyloid fibrils. *The FASEB Journal*, 16(1):77–83.

[Gegg et al., 2012] Gegg, M. E., Burke, D., Heales, S. J. R., Cooper, J. M., Hardy, J., Wood, N. W., and Schapira, A. H. V. (2012). Glucocerebrosidase deficiency in substantia nigra of parkinson disease brains. *Annals of neurology*, 72(3):455–463.

[Gerken et al., 1982] Gerken, T. A., Jentoft, J. E., Jentoft, N., and Dearborn, D. G. (1982). Intramolecular interactions of amino groups in 13C reductively methylated hen egg-white lysozyme. *The Journal of biological chemistry*, 257(6):2894–900.

[Giasson et al., 2001] Giasson, B. I., Murray, I. V. J., Trojanowski, J. Q., and Lee, V. M.-Y. (2001). A Hydrophobic Stretch of 12 Amino Acid Residues in the Middle of -Synuclein Is Essential for Filament Assembly. *Journal of Biological Chemistry*, 276(4):2380–2386.

[Giehm et al., 2010] Giehm, L., Oliveira, C. L. P., Christiansen, G., Pedersen, J. S., and Otzen, D. E. (2010). SDS-Induced Fibrillation of α-Synuclein: An Alternative Fibrillation Pathway. *Journal of Molecular Biology*, 401(1):115–133.

[Goker-Alpan et al., 2010] Goker-Alpan, O., Stubblefield, B. K., Giasson, B. I., and Sidransky, E. (2010). Glucocerebrosidase is present in α-synuclein inclusions in Lewy body disorders. *Acta neuropathologica*, 120(5):641–9.

[Goldsbury et al., 2000] Goldsbury, C., Goldie, K., Pellaud, J., Seelig, J., Frey, P., Müller, S. A., Kistler, J., Cooper, G. J., and Aebi, U. (2000). Amyloid fibril formation from full-length and fragments of amylin. *Journal of structural biology*, 130(2-3):352–62.

[Gómez-Isla et al., 1997] Gómez-Isla, T., Hollister, R., West, H., Mui, S., Growdon, J. H., Petersen, R. C., Parisi, J. E., and Hyman, B. T. (1997). Neuronal loss correlates with but exceeds neurofibrillary tangles in Alzheimer's disease. *Annals of neurology*, 41(1):17–24.

[Goto et al., 1999] Goto, N. K., Gardner, K. H., Mueller, G. A., Willis, R. C., and Kay, L. E. (1999). A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated 15N-, 13C-, 2H-labeled proteins. *Journal of Biomolecular NMR*, 13(4):369–374.

[Grabowski, 2012] Grabowski, G. A. (2012). Gaucher disease and other storage disorders. *ASH Education Program Book*, 2012(1):13–18.

[Grace and Grabowski, 1990] Grace, M. E. and Grabowski, G. A. (1990). Human acid beta-glucosidase: glycosylation is required for catalytic activity. *Biochemical and biophysical research communications*, 168(2):771–777.

[Guo et al., 2013] Guo, J. L., Covell, D. J., Daniels, J. P., Iba, M., Stieber, A., Zhang, B., Riddle, D. M., Kwong, L. K., Xu, Y., Trojanowski, J. Q., and Lee, V. M. Y. (2013). Distinct α-synuclein strains differentially promote tau inclusions in neurons. *Cell*, 154(1):103–17.

[Haass and Selkoe, 2007] Haass, C. and Selkoe, D. J. (2007). Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. *Nature reviews. Molecular cell biology*, 8(2):101–12.

[Hagan et al., 2009] Hagan, J. P., Piskounova, E., and Gregory, R. I. (2009). Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nature structural & molecular biology*, 16(10):1021–5.

[Hansen et al., 2008] Hansen, D. F., Vallurupalli, P., Lundström, P., Neudecker, P., and Kay, L. E. (2008). Probing chemical shifts of invisible states of proteins with relaxation dispersion NMR spectroscopy: how well can we do? *Journal of the American Chemical Society*, 130(8):2667–75.

[Hansen, 2008] Hansen, F. (2008). CATIA - CPMG, Anti-TROSY, and TROSY Intelligent Analysis. *Online - only*.

[Hartl and Hayer-Hartl, 2009] Hartl, F. U. and Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and in vivo. *Nature structural & molecular biology*, 16(6):574–581.

[Hattori et al., 2013] Hattori, Y., Furuita, K., Ohki, I., Ikegami, T., Fukada, H., Shirakawa, M., Fujiwara, T., and Kojima, C. (2013). Utilization of lysine C-methylation NMR for protein-protein interaction studies. *Journal of biomolecular NMR*, 55(1):19–31.

[Hayes et al., 2014] Hayes, J., Peruzzi, P. P., and Lawler, S. (2014). MicroRNAs in cancer: biomarkers, functions and therapy. *Trends in Molecular Medicine*, 20:460–469.

[Heise et al., 2008] Heise, H., Celej, M. S., Becker, S., Riedel, D., Pelah, A., Kumar, A., Jovin, T. M., and Baldus, M. (2008). Solid-state NMR reveals structural differences between fibrils of wild-type and disease-related A53T mutant alpha-synuclein. *Journal of molecular biology*, 380(3):444–50.

[Heise et al., 2005] Heise, H., Hoyer, W., Becker, S., Andronesi, O. C., Riedel, D., and Baldus, M. (2005). Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. *Proceedings of the National Academy of Sciences of the United States of America*, 102(44):15871–15876.

[Helmus and Jaroniec, 2013] Helmus, J. J. and Jaroniec, C. P. (2013). Nmrglue: an open source Python package for the analysis of multidimensional NMR data. *Journal of Biomolecular NMR*, 55(4):355–367.

[Heo et al., 2012] Heo, I., Ha, M., Lim, J., Yoon, M.-J., Park, J.-E., Kwon, S. C., Chang, H., and Kim, V. N. (2012). Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs. *Cell*, 151(3):521–532.

[Heo et al., 2008] Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V. N. (2008). Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. *Molecular Cell*, 32(2):276–284.

[Heo et al., 2009] Heo, I., Joo, C., Kim, Y.-K. K., Ha, M., Yoon, M.-J. J., Cho, J., Yeom, K.-H. H., Han, J., and Kim, V. N. (2009). TUT4 in Con-

cert with Lin28 Suppresses MicroRNA Biogenesis through Pre-MicroRNA Uridylation. *Cell*, 138(4):696–708.

[Hilton et al., 2013] Hilton, G. R., Lioe, H., Stengel, F., Baldwin, A. J., and Benesch, J. P. (2013). Small heat-shock proteins: Paramedics of the cell. *Topics in Current Chemistry*, 328:69–98.

[Hochberg et al., 2014] Hochberg, G. K. a., Ecroyd, H., Liu, C., Cox, D., Cascio, D., Sawaya, M. R., Collier, M. P., Stroud, J., Carver, J. a., Baldwin, A. J., Robinson, C. V., Eisenberg, D. S., Benesch, J. L. P., and Laganowsky, A. (2014). The structured core domain of αB-crystallin can prevent amyloid fibrillation and associated toxicity. *Proceedings of the National Academy of Sciences of the United States of America*, 111(16):E1562–70.

[Holley et al., 1965] Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., Penswick, J. R., and Zamir, A. (1965). Structure of a Ribonucleic Acid. *Science*, 147(3664):1462–1465.

[Holmes et al., 2008] Holmes, C., Boche, D., Wilkinson, D., Yadegarfar, G., Hopkins, V., Bayer, A., Jones, R. W., Bullock, R., Love, S., Neal, J. W., Zotova, E., and Nicoll, J. A. R. (2008). Long-term effects of Abeta42 immunisation in Alzheimer's disease: follow-up of a randomised, placebo-controlled phase I trial. *Lancet (London, England)*, 372(9634):216–23.

[Horowitz et al.,] Horowitz, M., Pasmanik-Chor, M., Ron, I., and Kolodny, E. H. The enigma of the E326K mutation in acid β-glucocerebrosidase. *Molecular genetics and metabolism*, 104(1-2):35–8.

[Hoyer et al., 2002] Hoyer, W., Antony, T., Cherny, D., Heim, G., Jovin, T. M., and Subramaniam, V. (2002). Dependence of α-synuclein aggre-

gate morphology on solution conditions. *Journal of Molecular Biology*, 322(02):383–393.

[Hruska et al., 2008] Hruska, K. S., LaMarca, M. E., Scott, C. R., and Sidransky, E. (2008). Gaucher disease: mutation and polymorphism spectrum in the glucocerebrosidase gene (GBA). *Human Mutation*, 29(5):567–583.

[Information, 2007] Information, I. P. (2007). B-PER bacterial protein extraction reagent. *Tanpakushitsu kakusan koso. Protein, nucleic acid, enzyme*, 52(1):88–89.

[Irvine et al.,] Irvine, G. B., El-Agnaf, O. M., Shankar, G. M., and Walsh, D. M. Protein aggregation in the brain: the molecular basis for Alzheimer's and Parkinson's diseases. *Molecular medicine (Cambridge, Mass.)*, 14(7-8):451–64.

[Jancarik et al., 2004] Jancarik, J., Pufan, R., Hong, C., Kim, S. H., and Kim, R. (2004). Optimum solubility (OS) screening: An efficient method to optimize buffer conditions for homogeneity and crystallization of proteins. *Acta Crystallographica Section D: Biological Crystallography*, 60(9):1670–1673.

[Jiménez et al., 2002] Jiménez, J. L., Nettleton, E. J., Bouchard, M., Robinson, C. V., Dobson, C. M., and Saibil, H. R. (2002). The protofilament structure of insulin amyloid fibrils. *Proceedings of the National Academy of Sciences of the United States of America*, 99(14):9196–201.

[Jones et al., 2012] Jones, M. R., Blahna, M. T., Kozlowski, E., Matsuura, K. Y., Ferrari, J. D., Morris, S. A., Powers, J. T., Daley, G. Q., Quinton, L. J., and Mizgerd, J. P. (2012). Zcchc11 Uridylates Mature miRNAs to

Enhance Neonatal IGF-1 Expression, Growth, and Survival. *PLoS Genetics*, 8(11).

[Jones et al., 2009] Jones, M. R., Quinton, L. J., Blahna, M. T., Neilson, J. R., Fu, S., Ivanov, A. R., Wolf, D. A., and Mizgerd, J. P. (2009). Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nature Cell Biology*, 11(9):1157–1163.

[Kamenetz et al., 2003] Kamenetz, F., Tomita, T., Hsieh, H., Seabrook, G., Borchelt, D., Iwatsubo, T., Sisodia, S., and Malinow, R. (2003). APP processing and synaptic function. *Neuron*, 37(6):925–37.

[Kapp and Lorsch, 2004] Kapp, L. D. and Lorsch, J. R. (2004). The Molecular Mechanics of Eukaryotic Translation. *Annual Review of Biochemistry*, 73(1):657–704.

[Kappé et al., 2003] Kappé, G., Franck, E., Verschuure, P., Boelens, W. C., Leunissen, J. A. M., and de Jong, W. W. (2003). The human genome encodes 10 alpha-crystallin-related small heat shock proteins: HspB1-10. *Cell stress & chaperones*, 8(1):53–61.

[Kayed et al., 2003] Kayed, R., Head, E., Thompson, J. L., McIntire, T. M., Milton, S. C., Cotman, C. W., and Glabe, C. G. (2003). Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science (New York, N.Y.)*, 300(5618):486–9.

[Kempf and Loria, 2003] Kempf, J. G. and Loria, J. P. (2003). Protein dynamics from solution NMR: theory and applications. *Cell biochemistry and biophysics*, 37(3):187–211.

[Khurana et al., 2005] Khurana, R., Coleman, C., Ionescu-Zanetti, C., Carter, S. A., Krishna, V., Grover, R. K., Roy, R., and Singh, S. (2005). Mechanism of thioflavin T binding to amyloid fibrils. *Journal of Structural Biology*, 151(3):229–238.

[Kim et al., 1973] Kim, S. H., Quigley, G. J., Suddath, F. L., McPherson, a., Sneden, D., Kim, J. J., Weinzierl, J., and Rich, a. (1973). Three-dimensional structure of yeast phenylalanine transfer RNA: folding of the polynucleotide chain. *Science (New York, N.Y.)*, 179(70):285–288.

[Knowles et al., 2014] Knowles, T. P. J., Vendruscolo, M., and Dobson, C. M. (2014). The amyloid state and its association with protein misfolding diseases. *Nature reviews. Molecular cell biology*, 15(6):384–96.

[Kolter and Wendeler, 2003] Kolter, T. and Wendeler, M. (2003). Chemical chaperones–a new concept in drug research. *Chembiochem : a European journal of chemical biology*, 4(4):260–4.

[Kornfeld et al., 2013] Kornfeld, J.-W., Baitzel, C., Könner, A. C., Nicholls, H. T., Vogt, M. C., Herrmanns, K., Scheja, L., Haumaitre, C., Wolf, A. M., Knippschild, U., Seibler, J., Cereghini, S., Heeren, J., Stoffel, M., and Brüning, J. C. (2013). Obesity-induced overexpression of miR-802 impairs glucose metabolism through silencing of Hnf1b. *Nature*, 494(7435):111–115.

[Korzhnev et al., 2010] Korzhnev, D. M., Religa, T. L., Banachewicz, W., Fersht, A. R., and Kay, L. E. (2010). A transient and low-populated protein-folding intermediate at atomic resolution. *Science*, 329(5997):1312–1316.

[Kruger et al., 1982] Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982). Self-splicing RNA: Autoex-

cision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell*, 31(1):147–157.

[Kurinov et al., 2000] Kurinov, I., Mao, C., Irvin, J., and Uckun, F. (2000). X-Ray Crystallographic Analysis of Pokeweed Antiviral Protein-II after Reductive Methylation of Lysine Residues. *Biochemical and Biophysical Research Communications*, 275(2):549–552.

[Larda et al., 2012] Larda, S. T., Bokoch, M. P., Evanics, F., and Prosser, R. S. (2012). Lysine methylation strategies for characterizing protein conformations by NMR. *Journal of biomolecular NMR*, 54(2):199–209.

[Lasker et al., 2006] Lasker, M. V., Kuruvilla, S. M., Gajjar, M. M., Kapoor, A., and Nair, S. K. (2006). Metal Ion-Mediated Reduction in Surface Entropy Improves Diffraction Quality of Crystals of the IRAK-4 Death Domain. *Journal of Biomolecular Techniques : JBT*, 17(2):114–121.

[Lee et al., 2012] Lee, C. F., Bird, S., Shaw, M., Jean, L., and Vaux, D. J. (2012). Combined effects of agitation, macromolecular crowding, and interfaces on amyloidogenesis. *The Journal of biological chemistry*, 287(45):38006–19.

[Lee et al., 2004a] Lee, H.-J., Khoshaghideh, F., Patel, S., and Lee, S.-J. (2004a). Clearance of alpha-synuclein oligomeric intermediates via the lysosomal degradation pathway. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 24(8):1888–1896.

[Lee et al., 2004b] Lee, J. C., Langen, R., Hummel, P. A., Gray, H. B., and Winkler, J. R. (2004b). Alpha-synuclein structures from fluorescence energy-transfer kinetics: implications for the role of the protein in Parkinson's

disease. *Proceedings of the National Academy of Sciences of the United States of America*, 101(47):16466–71.

[Lee et al., 1993] Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, 75(5):843–854.

[Lee et al., 2003] Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., and Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature*, 425(6956):415–419.

[Lemkau et al., 2013] Lemkau, L. R., Comellas, G., Lee, S. W., Rikardsen, L. K., Woods, W. S., George, J. M., and Rienstra, C. M. (2013). Site-specific perturbations of alpha-synuclein fibril structure by the Parkinson's disease associated mutations A53T and E46K. *PloS one*, 8(3):e49750.

[Lemmens et al., 2010] Lemmens, R., Moore, M. J., Al-Chalabi, A., Brown Jr., R. H., and Robberecht, W. (2010). RNA metabolism and the pathogenesis of motor neuron diseases. *Trends in Neurosciences*, 33(5):249–258.

[Levy et al., 1990] Levy, E., Carman, M. D., Fernandez-Madrid, I. J., Power, M. D., Lieberburg, I., van Duinen, S. G., Bots, G. T., Luyendijk, W., and Frangione, B. (1990). Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type. *Science (New York, N.Y.)*, 248(4959):1124–1126.

[Li et al., 2001] Li, J., Uversky, V. N., and Fink, A. L. (2001). Effect of Familial Parkinson's Disease Point Mutations A30P and A53T on the Structural

Properties, Aggregation, and Fibrillation of Human α-Synuclein . *Biochemistry*, 40(38):11604–11613.

[Li et al., 2008] Li, J.-Y., Englund, E., Holton, J. L., Soulet, D., Hagell, P., Lees, A. J., Lashley, T., Quinn, N. P., Rehncrona, S., Björklund, A., Widner, H., Revesz, T., Lindvall, O., and Brundin, P. (2008). Lewy bodies in grafted neurons in subjects with Parkinson's disease suggest host-to-graft disease propagation. *Nature medicine*, 14(5):501–3.

[Lieberman, 2011] Lieberman, R. L. (2011). A Guided Tour of the Structural Biology of Gaucher Disease: Acid-β-Glucosidase and Saposin C. *Enzyme Research*, 2011.

[Lindberg et al., 2015] Lindberg, I., Shorter, J., Wiseman, R. L., Chiti, F., Dickey, C. a., and McLean, P. J. (2015). Chaperones in Neurodegeneration. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 35(41):13853–9.

[Liu et al., 2015] Liu, G., Chen, M., Mi, N., Yang, W., Li, X., Wang, P., Yin, N., Li, Y., Yue, F., Chan, P., and Yu, S. (2015). Increased oligomerization and phosphorylation of α-synuclein are associated with decreased activity of glucocerebrosidase and protein phosphatase 2A in aging monkey brains. *Neurobiology of aging*, 36(9):2649–59.

[Lo Bianco et al., 2008] Lo Bianco, C., Shorter, J., Regulier, E., Lashuel, H., Iwatsubo, T., Lindquist, S., and Aebischer, P. (2008). Hsp104 antagonizes alpha-synuclein aggregation and reduces dopaminergic degeneration in a rat model of Parkinson disease. *Journal of Clinical Investigation*, 118(9):3087–3097.

[Lokappa and Ulmer, 2011] Lokappa, S. B. and Ulmer, T. S. (2011). Alpha-synuclein populates both elongated and broken helix states on small unilamellar vesicles. *The Journal of biological chemistry*, 286(24):21450–7.

[Lorenzen et al., 2014] Lorenzen, N., Nielsen, S. B., Buell, A. K., Kaspersen, J. D., Arosio, P., Vad, B. S., Paslawski, W., Christiansen, G., Valnickova-Hansen, Z., Andreasen, M., Enghild, J. J., Pedersen, J. S., Dobson, C. M., Knowles, T. P. J., and Otzen, D. E. (2014). The role of stable α-synuclein oligomers in the molecular events underlying amyloid formation. *Journal of the American Chemical Society*, 136(10):3859–68.

[Lu et al., 2013] Lu, J.-X., Qiang, W., Yau, W.-M., Schwieters, C. D., Meredith, S. C., and Tycko, R. (2013). Molecular structure of β-amyloid fibrils in Alzheimer's disease brain tissue. *Cell*, 154(6):1257–68.

[Lu et al., 2008] Lu, M., Zhang, Q., Deng, M., Miao, J., Guo, Y., Gao, W., and Cui, Q. (2008). An Analysis of Human MicroRNA and Disease Associations. *PLoS ONE*, 3(10).

[Luan et al., 2013] Luan, Z., Li, L., Higaki, K., Nanba, E., Suzuki, Y., and Ohno, K. (2013). The chaperone activity and toxicity of ambroxol on Gaucher cells and normal mice. *Brain and Development*, 35(4):317–322.

[Luang et al., 2010] Luang, S., Cairns, J. R. K., Streltsov, V. A., and Hrmova, M. (2010). Crystallisation of Wild-Type and Variant Forms of a Recombinant Plant Enzyme β-D-Glucan Glucohydrolase from Barley (Hordeum vulgare L.) and Preliminary X-ray Analysis. *International Journal of Molecular Sciences*, 11(7):2759–2769.

[Lue et al., 1999] Lue, L. F., Kuo, Y. M., Roher, A. E., Brachova, L., Shen, Y., Sue, L., Beach, T., Kurth, J. H., Rydel, R. E., and Rogers, J. (1999). Soluble amyloid beta peptide concentration as a predictor of synaptic change in Alzheimer's disease. *The American journal of pathology*, 155(3):853–62.

[Lund-Katz et al., 1988] Lund-Katz, S., Ibdah, J., Letizia, J., Thomas, M., and Phillips, M. (1988). A 13C NMR characterization of lysine residues in apolipoprotein B and their role in binding to the low density lipoprotein receptor. *J. Biol. Chem.*, 263(27):13831–13838.

[Lunde et al., 2012] Lunde, B. M., Magler, I., and Meinhart, A. (2012). Crystal structures of the Cid1 poly (U) polymerase reveal the mechanism for UTP selectivity. *Nucleic acids research*, 40(19):9815–9824.

[MacRae and Doudna, 2007] MacRae, I. J. and Doudna, J. A. (2007). Ribonuclease revisited: structural insights into ribonuclease III family enzymes. *Current Opinion in Structural Biology*, 17(1):138–145.

[Maegawa et al., 2009] Maegawa, G. H., Tropak, M. B., Buttner, J. D., Rigat, B. A., Fuller, M., Pandit, D., Tang, L., Kornhaber, G. J., Hamuro, Y., Clarke, J. T., and Mahuran, D. J. (2009). Identification and characterization of ambroxol as an enzyme-enhancement agent for Gaucher disease. *J Biol Chem*, 284(35):23502–23516.

[Mak et al., 2010] Mak, S. K., McCormack, A. L., Manning-Bog, A. B., Cuervo, A. M., and Di Monte, D. a. (2010). Lysosomal degradation of α-synuclein in vivo. *Journal of Biological Chemistry*, 285(18):13621–13629.

[Makin et al., 2005] Makin, O. S., Atkins, E., Sikorski, P., Johansson, J., and Serpell, L. C. (2005). Molecular basis for amyloid fibril formation and sta-

bility. *Proceedings of the National Academy of Sciences of the United States of America*, 102(2):315–20.

[Malone et al., 2009] Malone, C. D., Brennecke, J., Dus, M., Stark, A., McCombie, W. R., Sachidanandam, R., and Hannon, G. J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. *Cell*, 137(3):522–535.

[Malone and Hannon, 2009] Malone, C. D. and Hannon, G. J. (2009). Small RNAs as Guardians of the Genome. *Cell*, 136(4):656–668.

[Marcoux and Robinson, 2013] Marcoux, J. and Robinson, C. V. (2013). Twenty years of gas phase structural biology.

[Marklund et al., 2015] Marklund, E. G., Degiacomi, M. T., Robinson, C. V., Baldwin, A. J., and Benesch, J. L. P. (2015). Collision cross sections for structural proteomics. *Structure*, 23(4):791–799.

[Marshall et al., 2011] Marshall, K. E., Morris, K. L., Charlton, D., O'Reilly, N., Lewis, L., Walden, H., Serpell, L. C., and O'Reilly, N. (2011). Hydrophobic, Aromatic, and Electrostatic Interactions Play a Central Role in Amyloid Fibril Formation and Stability - Biochemistry (ACS Publications). *Biochemistry*, 50(12):2061–2071.

[Mattila et al., 2000] Mattila, P. M., Rinne, J. O., Helenius, H., Dickson, D. W., and Röyttä, M. (2000). Alpha-synuclein-immunoreactive cortical Lewy bodies are associated with cognitive impairment in Parkinson's disease. *Acta Neuropathologica*, 100(3):285–290.

[Mayo et al., 2005] Mayo, C. J., Diprose, J. M., Walter, T. S., Berry, I. M., Wilson, J., Owens, R. J., Jones, E. Y., Harlos, K., Stuart, D. I., and Esnouf, R. M. (2005). Benefits of automated crystallization plate tracking, imaging, and analysis. *Structure (London, England: 1993)*, 13(2):175–182.

[Mazzulli et al., 2011] Mazzulli, J. R., Xu, Y.-H., Sun, Y., Knight, A. L., McLean, P. J., Caldwell, G. A., Sidransky, E., Grabowski, G. A., and Krainc, D. (2011). Gaucher Disease Glucocerebrosidase and α-Synuclein Form a Bidirectional Pathogenic Loop in Synucleinopathies. *Cell*, 146(1):37–52.

[McClendon et al., 2009] McClendon, S., Rospigliosi, C. C., and Eliezer, D. (2009). Charge neutralization and collapse of the C-terminal tail of alpha-synuclein at low pH. *Protein science : a publication of the Protein Society*, 18(7):1531–40.

[McLean et al., 1999] McLean, C. A., Cherny, R. A., Fraser, F. W., Fuller, S. J., Smith, M. J., Bush, A. I., and Masters, C. L. (1999). Soluble pool of beta amyloid as a determinant of severity of neurodegeneration in Alzheimer's disease. *Annals of Neurology*, 46(6):860–866.

[McLean et al., 2004] McLean, P. J., Klucken, J., Shin, Y., and Hyman, B. T. (2004). Geldanamycin induces Hsp70 and prevents alpha-synuclein aggregation and toxicity in vitro. *Biochemical and Biophysical Research Communications*, 321(3):665–669.

[McNeill et al., 2012] McNeill, A., Duran, R., Hughes, D. A., Mehta, A., and Schapira, A. H. V. (2012). A clinical and family history study of Parkinson's disease in heterozygous glucocerebrosidase mutation carriers. *Journal of neurology, neurosurgery, and psychiatry*, 83(8):853–4.

[McNeill et al., 2014] McNeill, A., Magalhaes, J., Shen, C., Chau, K. Y., Hughes, D., Mehta, A., Foltynie, T., Cooper, J. M., Abramov, A. Y., Gegg, M., and Schapira, A. H. V. (2014). Ambroxol improves lysosomal biochemistry in glucocerebrosidase mutation-linked Parkinson disease cells. *Brain*, 137(5):1481–1495.

[McPherson and Cudney, 2006] McPherson, A. and Cudney, B. (2006). Searching for silver bullets: an alternative strategy for crystallizing macromolecules. *Journal of structural biology*, 156(3):387–406.

[Mehta, 2006] Mehta, A. (2006). Epidemiology and natural history of Gaucher's disease. *European Journal of Internal Medicine*, 17, Supple:S2–S5.

[Meiboom and Gill, 1958] Meiboom, S. and Gill, D. (1958). Modified spin-echo method for measuring nuclear relaxation times. *Review of Scientific Instruments*, 29(8):688–691.

[Mirnikjoo et al., 2009] Mirnikjoo, B., Balasubramanian, K., and Schroit, A. J. (2009). Suicidal membrane repair regulates phosphatidylserine externalization during apoptosis. *The Journal of biological chemistry*, 284(34):22512–6.

[Molenaar et al., 2012] Molenaar, J. J., Domingo-Fernández, R., Ebus, M. E., Lindner, S., Koster, J., Drabek, K., Mestdagh, P., van Sluis, P., Valentijn, L. J., van Nes, J., Broekmans, M., Haneveld, F., Volckmann, R., Bray, I., Heukamp, L., Sprüssel, A., Thor, T., Kieckbusch, K., Klein-Hitpass, L., Fischer, M., Vandesompele, J., Schramm, A., van Noesel, M. M., Varesio, L., Speleman, F., Eggert, A., Stallings, R. L., Caron, H. N., Versteeg, R.,

and Schulte, J. H. (2012). LIN28B induces neuroblastoma and enhances MYCN levels via let-7 suppression. *Nature Genetics*, 44(11):1199–1206.

[Motabar et al., 2012] Motabar, O., Goldin, E., Leister, W., Liu, K., Southall, N., Huang, W., Marugan, J. J., Sidransky, E., and Zheng, W. (2012). A high throughput glucocerebrosidase assay using the natural substrate glucosylceramide. *Analytical and Bioanalytical Chemistry*, 402(2):731–739.

[Mullen and Marzluff, 2008] Mullen, T. E. and Marzluff, W. F. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes & development*, 22(1):50–65.

[Müller, 2006] Müller, U. F. (2006). Re-creating an RNA world.

[Munishkina et al., 2004] Munishkina, L. A., Cooper, E. M., Uversky, V. N., and Fink, A. L. (2004). The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *Journal of molecular recognition : JMR*, 17(5):456–64.

[Munoz-Tello et al., 2012] Munoz-Tello, P., Gabus, C., and Thore, S. (2012). Functional implications from the Cid1 poly(U) polymerase crystal structure. *Structure (London, England: 1993)*, 20(6):977–986.

[Murphy et al., 2014] Murphy, K. E., Gysbers, A. M., Abbott, S. K., Tayebi, N., Kim, W. S., Sidransky, E., Cooper, A., Garner, B., and Halliday, G. M. (2014). Reduced glucocerebrosidase is associated with increased α-synuclein in sporadic Parkinson's disease. *Brain : a journal of neurology*, 137(Pt 3):834–48.

[Murphy and Halliday, 2014] Murphy, K. E. and Halliday, G. M. (2014). Glucocerebrosidase deficits in sporadic Parkinson disease. *Autophagy*, 10(7):1350–1.

[Näslund et al.,] Näslund, J., Haroutunian, V., Mohs, R., Davis, K. L., Davies, P., Greengard, P., and Buxbaum, J. D. Correlation between elevated levels of amyloid beta-peptide in the brain and cognitive decline. *JAMA*, 283(12):1571–7.

[Neudecker et al., 2009] Neudecker, P., Lundstr??m, P., Kay, L. E., Lundstrom, P., and Kay, L. E. (2009). Relaxation dispersion NMR spectroscopy as a tool for detailed studies of protein folding. *Biophysical Journal*, 96(6):2045–2054.

[Neudecker et al., 2012] Neudecker, P., Robustelli, P., Cavalli, a., Walsh, P., Lundstrom, P., Zarrine-Afsar, a., Sharpe, S., Vendruscolo, M., and Kay, L. E. (2012). Structure of an Intermediate State in Protein Folding and Aggregation. *Science*, 336(6079):362–366.

[Neumann et al., 2009] Neumann, J., Bras, J., Deas, E., O'Sullivan, S. S., Parkkinen, L., Lachmann, R. H., Li, A., Holton, J., Guerreiro, R., Paudel, R., Segarane, B., Singleton, A., Lees, A., Hardy, J., Houlden, H., Revesz, T., and Wood, N. W. (2009). Glucocerebrosidase mutations in clinical and pathologically proven Parkinson's disease. *Brain*, 132(7):1783–1794.

[Neveu et al., 2013] Neveu, M., Kim, H.-J., and Benner, S. a. (2013). The "strong" RNA world hypothesis: fifty years old. *Astrobiology*, 13(4):391–403.

[Nichols et al., 2009] Nichols, W. C., Pankratz, N., Marek, D. K., Pauciulo, M. W., Elsaesser, V. E., Halter, C. A., Rudolph, A., Wojcieszek, J., Pfeiffer, R. F., and Foroud, T. (2009). Mutations in GBA are associated with familial Parkinson disease susceptibility and age at onset. *Neurology*, 72(4):310–316.

[Norbury, 2010] Norbury, C. J. (2010). 3' Uridylation and the regulation of RNA function in the cytoplasm. *Biochemical Society transactions*, 38(4):1150–1153.

[Nott et al., 2015] Nott, T. J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T. D., Bazett-Jones, D. P., Pawson, T., Forman-Kay, J. D., and Baldwin, A. J. (2015). Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Molecular Cell*, 57(5):936–947.

[Outeiro et al., 2006] Outeiro, T. F., Klucken, J., Strathearn, K. E., Liu, F., Nguyen, P., Rochet, J.-C., Hyman, B. T., and McLean, P. J. (2006). Small heat shock proteins protect against alpha-synuclein-induced toxicity and aggregation. *Biochemical and biophysical research communications*, 351(3):631–8.

[Palmer et al., 2001] Palmer, A. G., Kroenke, C. D., and Patrick Loria, J. (2001). [10] Nuclear Magnetic Resonance Methods for Quantifying Microsecond-to-Millisecond Motions in Biological Macromolecules. In *Methods in Enzymology*, volume 339, pages 204–238.

[Pandey et al., 2006] Pandey, N., Schmidt, R. E., and Galvin, J. E. (2006). The alpha-synuclein mutation E46K promotes aggregation in cultured cells. *Experimental neurology*, 197(2):515–20.

[Paravastu et al., 2008] Paravastu, A. K., Leapman, R. D., Yau, W.-M., and Tycko, R. (2008). Molecular structural basis for polymorphism in Alzheimer's beta-amyloid fibrils. *Proceedings of the National Academy of Sciences of the United States of America*, 105(47):18349–54.

[Paravastu et al., 2009] Paravastu, A. K., Qahwash, I., Leapman, R. D., Meredith, S. C., and Tycko, R. (2009). Seeded growth of beta-amyloid fibrils from Alzheimer's brain-derived fibrils produces a distinct fibril structure. *Proceedings of the National Academy of Sciences of the United States of America*, 106(18):7443–8.

[Parenti, 2009] Parenti, G. (2009). Treating lysosomal storage diseases with pharmacological chaperones: From concept to clinics.

[Park et al., 2011] Park, J.-E., Heo, I., Tian, Y., Simanshu, D. K., Chang, H., Jee, D., Patel, D. J., and Kim, V. N. (2011). Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature*, 475(7355):201–205.

[Park et al., 2002] Park, J. K., Tayebi, N., Stubblefield, B. K., LaMarca, M. E., MacKenzie, J. J., Stone, D. L., and Sidransky, E. (2002). The E326K mutation and Gaucher disease: mutation or polymorphism? *Clinical genetics*, 61(1):32–4.

[Paroo et al., 2007] Paroo, Z., Liu, Q., and Wang, X. (2007). Biochemical mechanisms of the RNA-induced silencing complex. *Cell Research*, 17(3):187–194.

[Patten et al., 2007] Patten, S. M. V., Hughes, H., Huff, M. R., Piepenhagen, P. A., Waire, J., Qiu, H., Ganesa, C., Reczek, D., Ward, P. V., Kutzko, J. P., and Edmunds, T. (2007). Effect of mannose chain length on targeting

of glucocerebrosidase for enzyme replacement therapy of Gaucher disease. *Glycobiology*, 17(5):467–478.

[Payton et al., 2004] Payton, J. E., Perrin, R. J., Woods, W. S., and George, J. M. (2004). Structural determinants of PLD2 inhibition by alpha-synuclein. *Journal of molecular biology*, 337(4):1001–9.

[Petkova, 2005] Petkova, A. T. (2005). Self-Propagating, Molecular-Level Polymorphism in Alzheimer's -Amyloid Fibrils. *Science*, 307(5707):262–265.

[Petkova et al., 2006] Petkova, A. T., Yau, W.-M., and Tycko, R. (2006). Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils. *Biochemistry*, 45(2):498–512.

[Powers et al., 2009] Powers, E. T., Morimoto, R. I., Dillin, A., Kelly, J. W., and Balch, W. E. (2009). Biological and chemical approaches to diseases of proteostasis deficiency. *Annual review of biochemistry*, 78:959–91.

[Pronchik et al., 2010] Pronchik, J., He, X., Giurleo, J. T., and Talaga, D. S. (2010). In vitro formation of amyloid from alpha-synuclein is dominated by reactions at hydrophobic interfaces. *Journal of the American Chemical Society*, 132(28):9797–803.

[Putcha et al., 2010] Putcha, P., Danzer, K. M., Kranich, L. R., Scott, A., Silinski, M., Mabbett, S., Hicks, C. D., Veal, J. M., Steed, P. M., Hyman, B. T., and McLean, P. J. (2010). Brain-permeable small-molecule inhibitors of Hsp90 prevent alpha-synuclein oligomer formation and rescue alpha-synuclein-induced toxicity. *The Journal of pharmacology and experimental therapeutics*, 332(3):849–857.

[Rayment, 1997] Rayment, I. (1997). Reductive alkylation of lysine residues to alter crystallization properties of proteins. *Methods in Enzymology*, 276:171–179.

[Recchia, 2004] Recchia, A. (2004). alpha-Synuclein and Parkinson's disease. *The FASEB Journal*, 18(6):617–626.

[Rekas et al., 2004] Rekas, A., Adda, C. G., Andrew Aquilina, J., Barnham, K. J., Sunde, M., Galatis, D., Williamson, N. A., Masters, C. L., Anders, R. F., Robinson, C. V., Cappai, R., and Carver, J. A. (2004). Interaction of the molecular chaperone alphaB-crystallin with alpha-synuclein: effects on amyloid fibril formation and chaperone activity. *Journal of molecular biology*, 340(5):1167–83.

[Rissland and Norbury, 2008] Rissland, O. S. and Norbury, C. J. (2008). The Cid1 poly(U) polymerase. *Biochimica et biophysica acta*, 1779(4):286–294.

[Rissland and Norbury, 2009] Rissland, O. S. and Norbury, C. J. (2009). Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. *Nature structural & molecular biology*, 16(6):616–623.

[Rocha et al., 2015] Rocha, E. M., Smith, G. A., Park, E., Cao, H., Brown, E., Hayes, M. A., Beagan, J., McLean, J. R., Izen, S. C., Perez-Torres, E., Hallett, P. J., and Isacson, O. (2015). Glucocerebrosidase gene therapy prevents α-synucleinopathy of midbrain dopamine neurons. *Neurobiology of disease*, 82:495–503.

[Ron and Horowitz, 2005] Ron, I. and Horowitz, M. (2005). ER retention and degradation as the molecular basis underlying Gaucher disease heterogeneity. *Human Molecular Genetics*, 14(16):2387–2398.

[Roush and Slack, 2008] Roush, S. and Slack, F. J. (2008). The let-7 family of microRNAs. *Trends in Cell Biology*, 18(10):505–516.

[Ruotolo and Robinson, 2006] Ruotolo, B. T. and Robinson, C. V. (2006). Aspects of native proteins are retained in vacuum.

[Ruschak and Kay, 2010] Ruschak, A. M. and Kay, L. E. (2010). Methyl groups as probes of supra-molecular structure, dynamics and function. *Journal of biomolecular NMR*, 46(1):75–87.

[Rypniewski et al., 1993] Rypniewski, W. R., Holden, H. M., and Rayment, I. (1993). Structural consequences of reductive methylation of lysine residues in hen egg white lysozyme: an X-ray analysis at 1.8-A resolution. *Biochemistry*, 32(37):9851–8.

[Sabin et al., 2013] Sabin, L., Delás, M., and Hannon, G. (2013). Dogma Derailed: The Many Influences of RNA on the Genome. *Molecular Cell*, 49(5):783–794.

[Saibil, 2013] Saibil, H. (2013). Chaperone machines for protein folding, unfolding and disaggregation. *Nature Reviews Molecular Cell Biology*, 14(10):630–642.

[Samuel et al., 2000] Samuel, D., Kumar, T. K., Ganesh, G., Jayaraman, G., Yang, P. W., Chang, M. M., Trivedi, V. D., Wang, S. L., Hwang, K. C., Chang, D. K., and Yu, C. (2000). Proline inhibits aggregation during protein refolding. *Protein science: a publication of the Protein Society*, 9(2):344–352.

[Sardi et al., 2011] Sardi, S. P., Clarke, J., Kinnecom, C., Tamsett, T. J., Li, L., Stanek, L. M., Passini, M. a., Grabowski, G. a., Schlossmacher, M. G., Sidman, R. L., Cheng, S. H., and Shihabuddin, L. S. (2011). CNS expression of glucocerebrosidase corrects alpha-synuclein pathology and memory in a mouse model of Gaucher-related synucleinopathy. *Proceedings of the National Academy of Sciences of the United States of America*, 108(29):12101–12106.

[Sardi et al., 2013] Sardi, S. P., Clarke, J., Viel, C., Chan, M., Tamsett, T. J., Treleaven, C. M., Bu, J., Sweet, L., Passini, M. A., Dodge, J. C., Yu, W. H., Sidman, R. L., Cheng, S. H., and Shihabuddin, L. S. (2013). Augmenting CNS glucocerebrosidase activity as a therapeutic strategy for Parkinsonism and other Gaucher-related synucleinopathies. *Proceedings of the National Academy of Sciences of the United States of America*, 110(9):3537–3542.

[Sawaya et al., 2007] Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J. W., McFarlane, H. T., Madsen, A. Ø., Riekel, C., and Eisenberg, D. (2007). Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature*, 447(7143):453–7.

[Schapira and Gegg, 2013] Schapira, A. H. V. and Gegg, M. E. (2013). Glucocerebrosidase in the pathogenesis and treatment of Parkinson disease. *Proceedings of the National Academy of Sciences*, 110(9):3214–3215.

[Sement et al., 2013] Sement, F. M., Ferrier, E., Zuber, H., Merret, R., Alioua, M., Deragon, J.-M., Bousquet-Antonelli, C., Lange, H., and Gagliardi, D.

(2013). Uridylation prevents 3 trimming of oligoadenylated mRNAs. *Nucleic Acids Research*.

[Shaw et al., 2007] Shaw, N., Cheng, C., and Liu, Z.-J. (2007). Procedure for reductive methylation of protein to improve crystallizability. *Protocol Exchange*.

[Shen and Goodman, 2004] Shen, B. and Goodman, H. M. (2004). Uridine Addition After MicroRNA-Directed Cleavage. *Science*, 306(5698):997.

[Shi et al., 2013] Shi, J., Koteiche, H. A., McDonald, E. T., Fox, T. L., Stewart, P. L., and McHaourab, H. S. (2013). Cryoelectron microscopy analysis of small heat shock protein 16.5 (Hsp16.5) complexes with T4 lysozyme reveals the structural basis of multimode binding. *The Journal of biological chemistry*, 288(7):4819–30.

[Sidransky, 2005] Sidransky, E. (2005). Gaucher disease and parkinsonism. *Molecular genetics and metabolism*, 84(4):302–4.

[Siebert et al., 2014] Siebert, M., Sidransky, E., and Westbroek, W. (2014). Glucocerebrosidase is shaking up the synucleinopathies. *Brain : a journal of neurology*, 137(Pt 5):1304–22.

[Snead and Eliezer, 2014] Snead, D. and Eliezer, D. (2014). Alpha-synuclein function and dysfunction on cellular membranes. *Experimental neurobiology*, 23(4):292–313.

[Sokol, 2012] Sokol, N. S. (2012). Small temporal RNAs in animal development. *Current opinion in genetics & development*, 22(4):368–373.

[Sprangers and Kay, 2007] Sprangers, R. and Kay, L. E. (2007). Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature*, 445(7128):618–22.

[Stagno et al., 2007] Stagno, J., Aphasizheva, I., Aphasizhev, R., and Luecke, H. (2007). Dual role of the RNA substrate in selectivity and catalysis by terminal uridylyl transferases. *Proceedings of the National Academy of Sciences of the United States of America*, 104(37):14634–14639.

[Stefanis, 2012] Stefanis, L. (2012). α-Synuclein in Parkinson's disease. *Cold Spring Harbor perspectives in medicine*, 2(2):a009399.

[Stevenson and Norbury, 2006] Stevenson, A. L. and Norbury, C. J. (2006). The Cid1 family of non-canonical poly(A) polymerases. *Yeast*, 23(13):991–1000.

[Sunde et al., 1997] Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. (1997). Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology*, 273(3):729–739.

[T. D. Goddard and D. G. Kneller,] T. D. Goddard and D. G. Kneller. Sparky. *University of California, San Francisco*.

[Takamori et al., 2006] Takamori, S., Holt, M., Stenius, K., Lemke, E., Grønborg, M., Riedel, D., Urlaub, H., Schenck, S., Brügger, B., Ringler, P., Müller, S., Rammner, B., Gräter, F., Hub, J., De Groot, B., Mieskes, G., Moriyama, Y., Klingauf, J., Grubmüller, H., Heuser, J., Wieland, F., and Jahn, R. (2006). Molecular anatomy of a trafficking organelle. *Cell*, 127(4):831–846.

[Talini et al., 2009] Talini, G., Gallori, E., and Maurel, M. C. (2009). Natural and unnatural ribozymes: Back to the primordial RNA world. *Research in Microbiology*, 160(7):457–465.

[Tam et al., 2008] Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R. M., and Hannon, G. J. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature*, 453(7194):534–538.

[Tarawneh and Holtzman, 2012] Tarawneh, R. and Holtzman, D. M. (2012). The clinical problem of symptomatic Alzheimer disease and mild cognitive impairment. *Cold Spring Harbor perspectives in medicine*, 2(5):a006148.

[Tekoah et al., 2013] Tekoah, Y., Tzaban, S., Kizhner, T., Hainrichson, M., Gantman, A., Golembo, M., Aviezer, D., and Shaaltiel, Y. (2013). Glycosylation and functionality of recombinant β-glucocerebrosidase from various production systems. *Bioscience reports*, 33(5):771–781.

[Terry, 1996] Terry, R. D. (1996). The pathogenesis of Alzheimer disease: an alternative to the amyloid hypothesis. *Journal of neuropathology and experimental neurology*, 55(10):1023–5.

[Theillet et al., 2016] Theillet, F.-X., Binolfi, A., Bekei, B., Martorana, A., Rose, H. M., Stuiver, M., Verzini, S., Lorenz, D., van Rossum, M., Goldfarb, D., and Selenko, P. (2016). Structural disorder of monomeric α-synuclein persists in mammalian cells. *Nature*, advance on.

[Thornton et al., 2012] Thornton, J. E., Chang, H.-M., Piskounova, E., and Gregory, R. I. (2012). Lin28-mediated control of let-7 microRNA expres-

sion by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *RNA*, 18(10):1875–1885.

[Tölgyesi et al., 2004] Tölgyesi, E., Böde, C. S., Smelleri, L., Kim, D. R., Kim, K. K., Heremans, K., and Fidy, J. (2004). Pressure activation of the chaperone function of small heat shock proteins. *Cellular and molecular biology (Noisy-le-Grand, France)*, 50(4):361–9.

[Tollinger et al., 2001] Tollinger, M., Skrynnikov, N. R., Mulder, F. A. A., Forman-Kay, J. D., and Kay, L. E. (2001). Slow dynamics in folded and unfolded states of an SH3 domain. *Journal of the American Chemical Society*, 123(46):11341–11352.

[Townsend et al., 2006] Townsend, M., Shankar, G. M., Mehta, T., Walsh, D. M., and Selkoe, D. J. (2006). Effects of secreted oligomers of amyloid beta-protein on hippocampal synaptic plasticity: a potent role for trimers. *The Journal of physiology*, 572(Pt 2):477–92.

[Trojanowski and Lee, 2002] Trojanowski, J. Q. and Lee, V. M. Y. (2002). Parkinson's disease and related synucleinopathies are a new class of nervous system amyloidoses. *Neurotoxicology*, 23(4-5):457–60.

[Tugarinov et al., 2006] Tugarinov, V., Kanelis, V., and Kay, L. E. (2006). Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nature Protocols*, 1(2):749–754.

[Tuttle et al., 2016] Tuttle, M. D., Comellas, G., Nieuwkoop, A. J., Covell, D. J., Berthold, D. A., Kloepper, K. D., Courtney, J. M., Kim, J. K., Barclay, A. M., Kendall, A., Wan, W., Stubbs, G., Schwieters, C. D., Lee, V. M. Y., George, J. M., and Rienstra, C. M. (2016). Solid-state NMR

structure of a pathogenic fibril of full-length human α-synuclein. *Nature Structural & Molecular Biology*, 23(February):1–9.

[Tycko, 2011] Tycko, R. (2011). Solid-state NMR studies of amyloid fibril structure. *Annual review of physical chemistry*, 62:279–99.

[Vallurupalli et al., 2008] Vallurupalli, P., Hansen, D. F., and Kay, L. E. (2008). Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 105(33):11766–11771.

[Vilar et al., 2008] Vilar, M., Chou, H.-T., Lührs, T., Maji, S. K., Riek-Loher, D., Verel, R., Manning, G., Stahlberg, H., and Riek, R. (2008). The fold of alpha-synuclein fibrils. *Proceedings of the National Academy of Sciences of the United States of America*, 105(25):8637–8642.

[Walsh et al., 2002] Walsh, D. M., Klyubin, I., Fadeeva, J. V., Cullen, W. K., Anwyl, R., Wolfe, M. S., Rowan, M. J., and Selkoe, D. J. (2002). Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo. *Nature*, 416(6880):535–9.

[Walter et al., 2005] Walter, T. S., Diprose, J. M., Mayo, C. J., Siebold, C., Pickford, M. G., Carter, L., Sutton, G. C., Berrow, N. S., Brown, J., Berry, I. M., Stewart-Jones, G. B. E., Grimes, J. M., Stammers, D. K., Esnouf, R. M., Jones, E. Y., Owens, R. J., Stuart, D. I., and Harlos, K. (2005). A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization. *Acta crystallographica. Section D, Biological crystallography*, 61(Pt 6):651–657.

[Walter et al., 2006] Walter, T. S., Meier, C., Assenberg, R., Au, K.-F., Ren, J., Verma, A., Nettleship, J. E., Owens, R. J., Stuart, D. I., and Grimes, J. M. (2006). Lysine methylation as a routine rescue strategy for protein crystallization. *Structure (London, England : 1993)*, 14(11):1617–22.

[Wang et al., 2010] Wang, G.-F., Li, C., and Pielak, G. J. (2010). 19F NMR studies of α-synuclein-membrane interactions. *Protein science : a publication of the Protein Society*, 19(9):1686–91.

[Wang et al., 1999] Wang, J., Dickson, D. W., Trojanowski, J. Q., and Lee, V. M. (1999). The levels of soluble versus insoluble brain Abeta distinguish Alzheimer's disease from normal and pathologic aging. *Experimental neurology*, 158(2):328–37.

[Wang et al., 2000] Wang, S. W., Toda, T., MacCallum, R., Harris, A. L., and Norbury, C. (2000). Cid1, a fission yeast protein required for S-M checkpoint control when DNA polymerase delta or epsilon is inactivated. *Mol Cell Biol*, 20(9):3234–3244.

[Wang et al., 2011] Wang, W., Perovic, I., Chittuluru, J., Kaganovich, A., Nguyen, L. T. T., Liao, J., Auclair, J. R., Johnson, D., Landeru, A., Simorellis, A. K., Ju, S., Cookson, M. R., Asturias, F. J., Agar, J. N., Webb, B. N., Kang, C., Ringe, D., Petsko, G. A., Pochapsky, T. C., and Hoang, Q. Q. (2011). A soluble α-synuclein construct forms a dynamic tetramer. *Proceedings of the National Academy of Sciences of the United States of America*, 108(43):17797–802.

[Wasmer et al., 2008] Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A., Riek, R., and Meier, B. H. (2008). Amyloid Fibrils of the HET-s(218-289)

Prion Form a {beta} Solenoid with a Triangular Hydrophobic Core. *Science*, 319(5869):1523–1526.

[Waudby et al., 2010] Waudby, C. A., Knowles, T. P. J., Devlin, G. L., Skepper, J. N., Ecroyd, H., Carver, J. A., Welland, M. E., Christodoulou, J., Dobson, C. M., and Meehan, S. (2010). The interaction of alphaB-crystallin with mature alpha-synuclein amyloid fibrils inhibits their elongation. *Biophysical journal*, 98(5):843–51.

[Webb et al., 2003] Webb, J. L., Ravikumar, B., Atkins, J., Skepper, J. N., and Rubinsztein, D. C. (2003). α-synuclein Is Degraded by Both Autophagy and the Proteasome. *Journal of Biological Chemistry*, 278(27):25009–25013.

[Weinberg et al., 2011] Weinberg, D. E., Nakanishi, K., Patel, D. J., and Bartel, D. P. (2011). The Inside-Out Mechanism of Dicers from Budding Yeasts. *Cell*, 146(2):262–276.

[Weinreb et al., 1996] Weinreb, P. H., Zhen, W., Poon, A. W., Conway, K. A., and Lansbury, P. T. (1996). NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry*, 35(43):13709–15.

[Westerman et al., 2002] Westerman, M. A., Cooper-Blacketer, D., Mariash, A., Kotilinek, L., Kawarabayashi, T., Younkin, L. H., Carlson, G. A., Younkin, S. G., and Ashe, K. H. (2002). The relationship between Abeta and memory in the Tg2576 mouse model of Alzheimer's disease. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 22(5):1858–67.

[Wilhelmus et al., 2006] Wilhelmus, M. M. M., Boelens, W. C., Otte-Höller, I., Kamps, B., de Waal, R. M. W., and Verbeek, M. M. (2006). Small heat

shock proteins inhibit amyloid-beta protein aggregation and cerebrovascular amyloid-beta protein toxicity. *Brain research*, 1089(1):67–78.

[Williamson, 2013] Williamson, M. P. (2013). Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 73:1–16.

[Winder-Rhodes et al., 2013] Winder-Rhodes, S. E., Evans, J. R., Ban, M., Mason, S. L., Williams-Gray, C. H., Foltynie, T., Duran, R., Mencacci, N. E., Sawcer, S. J., and Barker, R. A. (2013). Glucocerebrosidase mutations influence the natural history of Parkinson's disease in a community-based incident cohort. *Brain*, 136(2):392–399.

[Winner et al., 2011] Winner, B., Jappelli, R., Maji, S. K., Desplats, P. A., Boyer, L., Aigner, S., Hetzer, C., Loher, T., Vilar, M., Campioni, S., Tzitzilonis, C., Soragni, A., Jessberger, S., Mira, H., Consiglio, A., Pham, E., Masliah, E., Gage, F. H., and Riek, R. (2011). In vivo demonstration that alpha-synuclein oligomers are toxic. *Proceedings of the National Academy of Sciences of the United States of America*, 108(10):4194–9.

[Wong et al., 2004] Wong, K., Sidransky, E., Verma, A., Mixon, T., Sandberg, G. D., Wakefield, L. K., Morrison, A., Lwin, A., Colegial, C., Allman, J. M., and Schiffmann, R. (2004). Neuropathology provides clues to the pathophysiology of Gaucher disease. *Molecular genetics and metabolism*, 82(3):192–207.

[World Health Organization, a] World Health Organization. WHO — Neurological Disorders: Public Health Challenges.

[World Health Organization, b] World Health Organization. WHO — The global burden of disease: 2004 update.

[Wright et al., 2015] Wright, M. A., Aprile, F. A., Arosio, P., Vendruscolo, M., Dobson, C. M., and Knowles, T. P. J. (2015). Biophysical approaches for the study of interactions between molecular chaperones and protein aggregates. *Chemical communications (Cambridge, England)*, 51(77):14425–34.

[Wu et al., 2009] Wu, K. P., Weinstock, D. S., Narayanan, C., Levy, R. M., and Baum, J. (2009). Structural Reorganization of α-Synuclein at Low pH Observed by NMR and REMD Simulations. *Journal of Molecular Biology*, 391(4):784–796.

[Xi et al., 2011] Xi, D., Dong, X., Deng, W., and Lai, L. (2011). Dynamic behavior of small heat shock protein inhibition on amyloid fibrillization of a small peptide (SSTSAA) from RNase A. *Biochemical and biophysical research communications*, 416(1-2):130–4.

[Xi et al., 2014] Xi, D., Wei, P., Zhang, C., and Lai, L. (2014). The minimal α-crystallin domain of Mj Hsp16.5 is functional at non-heat-shock conditions. *Proteins: Structure, Function, and Bioinformatics*, 82(7):1156–1167.

[Xue et al., 2008] Xue, W.-F., Homans, S. W., and Radford, S. E. (2008). Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. *Proceedings of the National Academy of Sciences of the United States of America*, 105(26):8926–31.

[Yap et al., 2011] Yap, T. L., Gruschus, J. M., Velayati, A., Westbroek, W., Goldin, E., Moaven, N., Sidransky, E., and Lee, J. C. (2011). α-Synuclein

Interacts with Glucocerebrosidase Providing a Molecular Link between Parkinson and Gaucher Diseases. *The Journal of Biological Chemistry*, 286(32):28080–28088.

[Yap et al., 2013] Yap, T. L., Velayati, A., Sidransky, E., and Lee, J. C. (2013). Membrane-bound α-synuclein interacts with glucocerebrosidase and inhibits enzyme activity. *Molecular genetics and metabolism*, 108(1):56–64.

[Yates et al., 2013] Yates, L., Norbury, C., and Gilbert, R. (2013). The Long and Short of MicroRNA. *Cell*, 153(3):516–519.

[Yates et al., 2015a] Yates, L. A., Durrant, B. P., Barber, M., Harlos, K., Fleurdépine, S., Norbury, C. J., and Gilbert, R. J. C. (2015a). Improved crystallization and diffraction of caffeine-induced death suppressor protein 1 (Cid1). *Acta crystallographica. Section F, Structural biology communications*, 71(Pt 3):346–53.

[Yates et al., 2015b] Yates, L. A., Durrant, B. P., Fleurdépine, S., Harlos, K., Norbury, C. J., and Gilbert, R. J. C. (2015b). Structural plasticity of Cid1 provides a basis for its distributive RNA terminal uridylyl transferase activity. *Nucleic acids research*, 43(5):2968–79.

[Yates et al., 2012] Yates, L. A., Fleurdépine, S., Rissland, O. S., De Colibus, L., Harlos, K., Norbury, C. J., and Gilbert, R. J. C. (2012). Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase. *Nature structural & molecular biology*, 19(8):782–787.

[Yeom et al., 2011] Yeom, K.-H., Heo, I., Lee, J., Hohng, S., Kim, V. N., and Joo, C. (2011). Single-molecule approach to immunoprecipitated protein

complexes: insights into miRNA uridylation. *EMBO reports*, 12(7):690–696.

[Yu et al., 2007] Yu, Z., Sawkar, A. R., and Kelly, J. W. (2007). Pharmacologic chaperoning as a strategy to treat Gaucher disease.

[Zarranz et al., 2004] Zarranz, J. J., Alegre, J., Gómez-Esteban, J. C., Lezcano, E., Ros, R., Ampuero, I., Vidal, L., Hoenicka, J., Rodriguez, O., Atarés, B., Llorens, V., Gomez Tortosa, E., del Ser, T., Muñoz, D. G., and de Yebenes, J. G. (2004). The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia. *Annals of neurology*, 55(2):164–73.

[Zhao et al., 2014] Zhao, Y., Ren, J., Padilla-Parra, S., Fry, E. E., and Stuart, D. I. (2014). Lysosome sorting of β-glucocerebrosidase by LIMP-2 is targeted by the mannose 6-phosphate receptor. *Nature communications*, 5:4321.

[Zhou et al., 2012] Zhou, D. H., Nieuwkoop, A. J., Berthold, D. A., Comellas, G., Sperling, L. J., Tang, M., Shah, G. J., Brea, E. J., Lemkau, L. R., and Rienstra, C. M. (2012). Solid-state NMR analysis of membrane proteins and protein aggregates by proton detected spectroscopy. *Journal of biomolecular NMR*, 54(3):291–305.

[Zimmer et al., 1999] Zimmer, K. P., le Coutre, P., Aerts, H. M., Harzer, K., Fukuda, M., O'Brien, J. S., and Naim, H. Y. (1999). Intracellular transport of acid beta-glucosidase and lysosome-associated membrane proteins is affected in Gaucher's disease (G202R mutation). *J Pathol*, 188(4):407–414.

[Zimran et al., 2013] Zimran, A., Altarescu, G., and Elstein, D. (2013). Pilot study using ambroxol as a pharmacological chaperone in type 1 Gaucher disease. *Blood Cells, Molecules, and Diseases*, 50(2):134–137.