Conformational transitions governing function and aggregation of enzymes

Jayneil Patel

Imperial College London
Department of Life Sciences

Thesis for the degree of Doctor of Philosphy

February 2017

ABSTRACT

In certain cases, the dynamical and conformational properties of enzymes, fundamental to their function, may lead to adverse effects such as loss of activity and aggregation. Studying these changes will help to further our understanding of how these structural transitions affect enzyme activity and solubility. Two exemplar enzymes are studied; Lipex is an industrially used lipase present in laundry detergent products, and HypF-N is used as a model system for protein aggregation. HypF-N can aggregate to form non-pathogenic amyloid fibrils, and under certain solution conditions can form toxic or non-toxic oligomers. NMR methods, suited to studying dynamical and conformational changes, were used. The aims of this thesis set to elucidate how solution conditions affect enzyme conformations. Using Lipex to understand how structural transitions affect functionality in gain or loss of activity. And through HypF-N understanding the process of aggregation and elucidating the origins of toxic gain of function.

Lipex was expressed and purified in an *Escherichia coli* host. Activity assays demonstrated presence of surfactants and calcium can have positive and negative effects on catalytic activity. 2D HSQC NMR experiments showed differences in the conformation of Lipex between high-active and low-active conditions. NMR relaxation measurements indicated that a slower tumbling protein is linked to a more catalytically active enzyme. It was also determined that anionic surfactants cause destabilising effects to the tertiary structure of Lipex and should be reduced in laundry products to increase protein stability.

NMR techniques demonstrated that aggregation occurs for HypF-N in folded and unfolded states. Despite these different initial conformations, a similar aggregation process exists across both toxic and non-toxic conditions. Solid-state NMR revealed a similar core structure in both toxic and non-toxic HypF-N oligomers. Differences in the mobility of the oligomer were detected using solid-state NMR and we hypothesise that certain flexible regions, and the interactions with the cell membrane, define toxicity.

ACKNOWLEDGEMENTS

I am incredibly indebted and grateful to my supervisor Dr. Alfonso De Simone for providing the support, knowledge and patience to help guide and encourage me through this PhD.

I am thankful to a number of people for their help and support over the course of the PhD:

From Unilever, Dr. Dietmar Lang, Dr. Massimo Noro, Ms. Ilaria Sambi and Dr. Mark Thompson for their patience, guidance and expertise in lipases and biotechnological studies of enzymes.

Prof. Fabrizio Chiti for the opportunity to collaborate on such an interesting and exciting project with HypF-N. I am grateful for the, discussions and insights.

Dr. Yingqi Xu for his invaluable help and advice in NMR data collection, processing and analysis. Likewise, to Dr. Geoff Kelly at NIMR for setting up NMR experiments.

To Dr. Justin Yeoman, Dr. Rhian Jones and Dr. James Jarvis for proofreading.

I feel very lucky to have shared my time in the lab with many dear friends: Mostafa, Sarah, Rhian, Seb, William, Maximo, Sabine and Xeni amongst many others.

Most significantly to BBSRC and Unilever for funding opportunities.

To Nesha, for her love and warm support. She reminds me that there is a wonderful life outside of the lab!

Finally to my family, to my brother who inspires me to always be better and to my parents for their love, confidence and encouragement, without which, I would not have made it to the end.

For my parents

Declaration of Originality

I declare that the authorship of this thesis and the research presented is my own work, unless otherwise stated, and appropriately referenced. This work has been carried out in the Division of Molecular Biosciences, Department of Life Scienes, Imperial College London between September 2012 and September 2016, and has not been submitted for any other degree.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-commerical No Derivatives Licence. Researchers are free to copy, distributre or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do no alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licenece terms of this work.

ABBREVIATIONS

°C: degrees celsius
°fH: French water hardness
6HIS: 6-hisitine (poly-histidine) tag
AFM: Atomic force microscopy
BMGY: Buffered minimal glycerol media
BMMY: Buffered minimal methanol media
Ca^{2+}: Calcium
Chol: Cholesterol
CMC: Critical micelle concentration
DARR: Dipolar assisted rotational resonance
DEST: Dark-state exchange saturation
DOPC: 18:1 (Δ9-Cis) 1,2-dioleoyl-*sn*-glycero-3-phosphocholine
DTT: Dithiothreitol
E. coli: Escherichia coli
FRET: Förster resonance energy transfer
g: Grams
g: Relative centrifugal force
GM1: Ganglioside 1
h: Hours
HSQC: Heteronuclear Single Quantum Coherence
HypF-N: N terminal domain of HypF
INEPT: Insensitive nuclei enhanced by polarization transfer
IPTG: Isopropyl β-D-1-thiogalactopyranoside
ITC: Isothermal titration calorimetry
K: Kelvin
kDa: Kila daltons
L: Litre
LAS: Linear alkyl benzene sulfonate
LB: Lysogeny (luria) broth
M: Molar
MAS: Magic angle spinning
Mg^{2+}: Magnesium
min: Minutes
ml: Millilitre
ms: Milliseconds
N': N terminal domain
NaAc: Sodium acetate
NaCl: Sodium chloride
NI: Neodol 25-7
NMR: Nuclear magnetic resonance
P. pastoris: Pichia pastoris
PCMB: p-Chloromercuribenzoic acid
PRE: Paramagnetic Relaxation Experiment
R: Rhamnolipid
SDS-PAGE: Sodium dodecyl sulphate polyacrylamide gel electrophoresis

sec: Size exclusion chromatography
SLES: Sodium laureth sulfate
SM: Sphingomyelin
ssNMR: Solid-state nuclear magnetic resonance
SUV: Small unilammar vesicle
T. lanuginosus: *Thermomyces lanuginosus*
T_2: Transverse relaxation
TFA: Trifluoroacetic acid
TFE: Tetrafluoroethylene
TlL: *Thermomyces lanuginosus* lipase
TROSY: Transverse relaxation-optimized spectroscopy
RF: Radio frequency

CONTENTS

Abstract **2**
Acknowledgments **3**
Abbreviations **6**
Contents **8**
List of figures **12**
List of tables **17**
Chapter 1: NMR **18**
1.1 Principles of Nuclear Magnetic Resonance 19
1.1.1 Nuclear spin 19
1.1.2 Longitudinal and transverse magnetisation 20
1.1.3 Energy levels 21
1.1.4 Chemical environments 23
1.1.5 INEPT 24
1.2 TROSY 26
1.3 T_2 relaxation measurements 28
1.4 Dark-state Exchange Saturation Transfer (DEST) 29
1.5 Solid state NMR (ssNMR) 31
1.5.1 Heteronuclear decoupling 34
1.5.2 Cross polarisation 33
1.5.3 DARR 33
1.5.4 INEPT in solids 34

Chapter 2: Introduction to Lipex **36**
Chapter summary 37
2 Protein structure and dynamics 37
2.1 Solubility of enzymes 38
2.2 Aggregation 39
2.3 Lipases 39
2.4 Industrial lipases 43
2.5 Detergent industry 44
2.6 Thermomyces lanuginose lipase (TlL) 44
2.7 Lipex 48
2.8 Interfacial activation 49
2.9 Lipase dimers; self aggregation at interfaces 52
2.10 Structural conformations and re-organizations of TlL 52
2.11 Significant residues of TlL 55
2.12 Surfactants 57
2.13 pH 63
2.14 Calcium 64
2.15 Aims 66

Chapter 3: Lipex Methods and Materials 69
3.1 Yeast 70
3.1.1 Constructs 70
3.1.2 Strains 71
3.1.3 Transformation of E. coli cells with Lipex_pPICZα vector 72
3.1.4 Lipex_pPICZα_A vector linearization 72
3.1.5 Ethanol precipitation of plasmid DNA 72
3.1.6 Electroporation of P. pastoris cells 73
3.2 Expression trials 74
3.2.1 Expression trials in Rich Buffered Glycerol Complex Media (BMGY), Buffered Methanol Complex Media (BMMY) and Basal Salt Medium 74
3.2.2 Expression trials in Minimal Buffered Methanol Complex Media BMMY (mBMMY) 74
3.2.3 Expression trials in FM-22 75
3.3 Western blot 75
3.4 E. coli 77
3.4.1 Construct design 77
3.4.2 Strain 78
3.4.3 Lipex_pET21a vector 78
3.5 Expression of Lipex in E. coli 79
3.5.1 Expression of unlabeled Lipex in E. coli 79
3.5.2 Expression of isotopically labelled ^{15}N and or ^{13}C Lipex in E. coli 80
3.6 Purification of Lipex expressed in E. coli 80
3.7 Natural Source Lipex 81
3.8 Activity Assay 81
3.9 Protein Concentration Estimation 82
3.10 NMR 82
3.10.1 Solution NMR 82
3.10.2 1D ^{1}H NMR 83
3.10.3 2D ^{15}N-^{1}H TROSY-HSQC 83
3.10.4 Relaxation experiments 84

Chapter 4: Lipex Results 85
4.1 Chapter overview 86
4.2 Expression trials 86
4.2.1 Initial expression in *P. pastoris* 86
4.2.2 Proteolytic degradation 90
4.2.3 Expression in *E. coli* 94
4.2.3.1 Strain screening and construct optimisation 95
4.2.3.2 Expression optimisation 100
4.2.4 Purification 102
4.2.5 Concluding remarks 106
4.3 Preliminary NMR characterisation of Lipex and catalytic activity under industrially relevant conditions 107
4.3.1 NMR 107
4.3.2 Comparisons to Unilever glycosylated Lipex 109
4.3.3 NMR buffer optimisation 111
4.3.4 Enzyme stability 113
4.3.5 Lipex activity under industrially relevant conditions 114
4.3.6 Secondary structure of Lipex under industrially relevant conditions as measured by Circular Dichroism (CD) 122

4.3.7 Effect of temperature on Lipex conformation measured by NMR 125
4.3.8 Effect of calcium on Lipex activity and conformation measured by NMR 130
4.4 Conformation and dynamics of Lipex in formulations 134
4.4.1 Conformation and dynamics of Lipex in high active formulations 134
4.4.2 Effect of calcium on Lipex conformation and dynamics in high and low active formulations 138
4.4.3 Studying Lipex activity with surfactants and increasing calcium levels 143
4.5 Surfactant induced unfolding of Lipex 148
4.5.1 Lipex unfolded under storage conditions 149
4.5.2 Effect of individual surfactants on Lipex structure 154
4.5.2.1 LAS 155
4.5.2.2 SLES and NI 157
4.5.3 Destabilising effects of surfactants on Lipex structure 160
4.6 Chapter summary 166
4.6.1 Concluding remarks 166
4.6.2 Future work 167

Chapter 5: Introduction to HypF-N and context to research 168
Chapter summary 169
5.1 Protein aggregation in human health and disease 169
5.2 Architecture of amyloid fibrils 171
5.3 Toxicity arises from prefibrillar aggregates 173
5.4 Architecture of oligomers 174
5.5 Aggregation pathway 175
5.6 Key features of oligomer cytotoxicity 177
5.7 HypF-N 178
5.8 HypF-N aggregation 180
5.9 Differences in initial structure of HypF-N 182
5.10 Toxicity arises from exposure of hydrophobic regions 183
5.11 HypF-N and acylphosphatase aggregation 185
5.12 Mechanisms of oligomer and membrane interaction 187
5.13 Biophysical techniques for studying membrane interactions 188
5.14 Membrane physiology can modulate toxicity 189
5.15 Hypothesis and Aims 191

Chapter 6: HypF-N Materials and Methods 193
6.1 Constructs and strains 194
6.1.1 Construct design 194
6.1.2 Strain 195
6.2 Expression and purification 195
6.2.1 Expression of unlabelled HypF-N 195
6.2.2 Purification 196
6.2.3 Expression of ^{15}N and $^{15}N^{13}C$ labelled HypF-N 196
6.4 Oligomer formation 196
6.5 Tapping Mode – Atomic Force Microscopy (TM-AFM) 197
6.6 NMR 197
6.6.1 Solution NMR 197
6.6.2 Backbone Assignment 198
6.6.2.1 Native 198

6.6.2.2 Toxic condition 198
6.6.2.3 Non-toxic condition 199
6.6.3 DEST 200
6.6.3.1 Toxic condition 200
6.6.3.2 Non-toxic condition 200
6.6.4 Solid state NMR 201
6.6.4.1 Oligomer only samples for ssNMR 201
6.6.4.2 Lipid SUV formation 201
6.6.4.3 Oligomer + Lipid samples 202
6.6.4.4 PRE lipid samples 202
6.6.4.5 DARR 202
6.6.4.6 INEPT 203

Chapter 7: HypF-N Results 204
7.1 Expression and Purification 205
7.2 Solution-state NMR 208
7.2.1 NMR characterisation 208
7.2.1 Backbone assignment 213
7.2.3 Native 213
7.2.4 Toxic 218
7.2.5 Non-toxic 219
7.3 Structural predictions of HypF-N under oligomerizing conditions 224
7.4 Oligomerization 225
7.5 Dark state exchange saturation transfer (DEST) 226
7.5.1 Sample optimisation 227
7.5.1.1 Toxic 227
7.5.1.2 Non-toxic 229
7.5.2 Nucleation 233
7.5.3 Aggregation occurs via the β2 strand 235
7.5.4 β2 strand and the β-loop are present in the core of both oligomers 237
7.5.5 Comparison to previous data 238
7.5.6 Concluding remarks 239
7.6 Solid-state NMR 239
7.6.1 Toxic and non-toxic oligomers have similar core structure 240
7.6.2 Differences in oligomers lie within the flexible regions 243
7.6.3 Vesicle interactions 244
7.6.4 GM1 content 247
7.6.5 Position of oligomers in the membrane 250
7.6.6 Concluding remarks 258
7.7 Future work 258

References 260

Supplementary Material 281

LIST OF FIGURES

Figure 1.1 Nuclear spin precession 20
Figure 1.2 Orientation of states against B_0 23
Figure 1.3 Pulse sequence of INEPT 25
Figure 1.4 Simple line-selective 180° pulse sequence 28
Figure 1.5 Transverse relaxation over time as the coherence of spins is lost in the xy plane 29
Figure 1.6 Simple spin echo 29
Figure 1.7 Representation of magic-angle spinning of a sample 32
Figure 2.1 Lipase hydrolysis of triglyceride molecule 40
Figure 2.2 Different reactions catalysed by lipases 41
Figure 2.3 Detailed mechanism of lipase catalytic activity. 42
Figure 2.4 TlL secondary structure topology 46
Figure 2.5 TlL structure, catalytic residues highlighted 46
Figure 2.6 Helical lid domain of TlL in open and closed conformation 47
Figure 2.7 Enzyme kinetic profile during interfacial activation 50
Figure 2.8 Lipid interface allow for interfacial activation 51
Figure 2.9 Structural change of TlL from open and closed conformation 53
Figure 2.10 Structural rearrangement of key residues of TlL from open to closed conformation 54
Figure 2.11 Tryptophan 89 is a critical amino acid in the lid domain 56
Figure 2.12 Different surfactant configurations 58
Figure 2.13. Linear Alkyl Benzene Sufonate (LAS) (C_{12} alkyl chain) 60
Figure 2.14 Sodium Lauryl Ether Sulphate (SLES) 60
Figure 2.15 Neodol 25-9 (NI) (C_{13} alkyl chain) 60
Figure 2.16 Model mechanism of anionic and non-ionic surfactant interactions 61
Figure 2.17 Rhamnolipid JBR425 structure of mono and di-rhamnolipid 62
Figure 3.1 pPICZα vector map 71
Figure 3.2 pET21a vector map 77
Figure 4.1 Diagram representation of Lipex construct for Lipex_pPICZa 87
Figure 4.2 Expression and secretion of Lipex in X-33 cells containing Lipex_pPICZα in BMMY culture 87
Figure 4.3 Media trial for expression of Lipex in X-33 cells 89

Figure 4.4 Proteolytic degradation of Lipex expressed in X-33 cells in minimal BMGY/BMMY media 90
Figure 4.5 Expression trial of Lipex using combination of protease inhibition 93
Figure 4.6 Lemo21 expresses Lipex in insoluble form..... 96
Figure 4.7 Soluble expression of Lipex in SoluBL21 cells 97
Figure 4.8 Diagram representation of Lipex construct for Lipex_pET21a 98
Figure 4.9 Periplasmic expression and purification with negative controls of Lipex expression analysed by 12% SDS-PAGE 99
Figure 4.10 Activity assay of periplasmic purified soluble protein samples on pNp valerate substrate 100
Figure 4.11 Electrospray-ionization Time of Flight (ES-ToF) mass spectrometry profile of periplasm purified fraction 101
Figure 4.12 Ion exchange chromatogram of Lipex 102
Figure 4.13 12% SDS-PAGE gel of the sucrose gradient periplasmic purification and ion exchange of Lipex 103
Figure 4.14 Precipitate of Lipex with contaminating protein during heat incubation 104
Figure 4.15 SEC Gel filtration purification of Lipex after ion exchange 105
Figure 4.16 SDS-PAGE gel of fractions taken from gel filtration of Lipex 105
Figure 4.17 1D ^{1}H spectrum of Lipex 107
Figure 4.18 2D ^{15}N TROSY-HSQC spectrum of Lipex 108
Figure 4.19 1D ^{1}H NMR spectrum of Lipex in glycosylated and non-glycosylated form..... 110
Figure 4.20 Glycosylation does not affect activity of Lipex 110
Figure 4.21 Lipex stability measured by 2D NMR..... 113
Figure 4.22 Surfactant formulation screening on Lipex activity at increasing water hardness levels 116
Figure 4.23 Effects of weakly anionic surfactant mixture on Lipex activity..... 118
Figure 4.24 Effects of non-ionic and weakly anionic surfactant mixture on Lipex activity 119
Figure 4.25 Effects of anionic surfactants on Lipex activity in presence of non-ionic surfactant..... 120
Figure 4.26 Far-UV CD spectra of Lipex at different surfactant conditions at increasing water hardness levels 123
Figure 4.27 2D ^{15}N TROSY-HSQC overlay of Lipex at a range of temperatures 125
Figure 4.28 T_2 relaxation measurements for Lipex as function of temperature 128
Figure 4.29 Effect on calcium on Lipex activity 131
Figure 4.30 2D ^{15}N TROSY-HSQC overlay of Lipex in absence and presence of calcium at 50 °fH 132

Figure 4.31 Effect of calcium on the T_2 relaxation of Lipex 133
Figure 4.32 2D ^{15}N TROSY-HSQC of Lipex in high active formulations 136
Figure 4.33 2D ^{15}N TROSY-HSQC of 50 degrees French hardness with Lipex and with Lipex + surfactant mixture NI:R 138
Figure 4.34 Changes in Lipex dynamics in low active and high active formulations 142
Figure 4.35 2D ^{15}N TROSY-HSQC overlay of Lipex with surfactants and increasing calcium 144
Figure 4.36 Residues involved in binding to surfactants LAS 145
Figure 4.37 Effects of LAS:R 1:1 at 50 °fH on Lipex dynamics..... 146
Figure 4.38 NMR observed unfolding of Lipex in Forumulation X..... 149
Figure 4.39 2D ^{15}N TROSY-HSQC overlay of Lipex in absence and presence of 5% (v/v) formulation Y 151
Figure 4.40 Activity assay of Lipex with increasing concentrations of Formulation X (FX)..... 153
Figure 4.41 Lipex with LAS at 0.0066%..... 155
Figure 4.42 Activity assay of Lipex with increasing concentrations of LAS 156
Figure 4.43 Lipex with NI and SLES at 0.0066% exhibit same conformation 157
Figure 4.44 Activity assay of Lipex with anionic and non ionic mixtures..... 159
Figure 4.45 NI inhibits Lipex activity where it is retained for SLES and LAS 159
Figure 4.46 Effects of SLES on Lipex structure 161
Figure 4.47 Partially unfolded conformation of Lipex under SLES and LAS 162
Figure 4.48 2D ^{15}N TROSY-HSQC overlay of surfactants at 0.1%..... 164
Figure 5.1 Molecular model of 4-protofilament insulin fibril 172
Figure 5.2 NMR derived β- stranded structures and packing of protofilaments 172
Figure 5.3. Fragment of human prion protein in an oligomeric structure determined by X-ray crystallography 174
Figure 5.4 Aggregation pathway of amyloidogenic proteins 176
Figure 5.5 HypF-N structure and sequence 179
Figure 5.6 Regions predicted to form core of HypF-N oligomers 183
Figure 5.7 Hydrophobicity and aggregation propensity of HypF-N sequence..... 185
Figure 5.8 Structural similarities between HypF-N and AcpDro2 186
Figure 6.1. pQE30 vector map..... 194
Figure 7.1 Diagram representation of HypF-N construct 205
Figure 7.2 Expression of HypF-N..... 206
Figure 7.3 SEC gel filtration profile of HypF-N 207
Figure 7.4 12% SDS-PAGE gel of SEC-GF purification of HypF-N 207

Figure 7.5 2D ^{15}N-HSQC of HypF-N 208
Figure 7.6 2D ^{15}N-HSQC of HypF-N under toxic buffer condition 209
Figure 7.7 2D ^{15}N-HSQC of HypF-N under non-toxic buffer condition 210
Figure 7.8 Monomeric HypF-N exhibits different level of structure and conformation in toxic and non-toxic conditions 212
Figure 7.9 1H strips for CACB and COCACB spectra collected for assignment of HypF-N 214
Figure 7.10 2D ^{15}N-HSQC HypF-N peak shifts with addition of phosphate 216
Figure 7.11 Sodium phosphate has an effect at the anion pocket and loop regions of HypF-N 216
Figure 7.12 ^{15}N-HSQC of assigned backbone of HypF-N under native condition 217
Figure 7.13 ^{15}N-HSQC of assigned backbone of HypF-N under toxic condition 219
Figure 7.14 ^{15}N-HSQC spectra of the unfolded conformation of HypF-N is maintained with addition of 8M urea at pH 1.7 220
Figure 7.15 HncocaHN backbone amide 'walk' connecting *i* and -*i* backbone amide resonances 222
Figure 7.16 ^{15}N-HSQC of assigned backbone of HypF-N unfolded by 8 M urea 223
Figure 7.17 Secondary structure prediction for HypF-N under non-toxic (red) and toxic (green) conditions calculated by δ2D 224
Figure 7.18 TM-AFM images collected on HypF-N toxic (A) and non-toxic (B) oligomers 225
Figure 7.19 Assigned spectrum of toxic DEST 228
Figure 7.20 Factors affecting aggregation propensity of non-toxic HypF-N 230
Figure 7.21 Assigned spectrum of non-toxic DEST 232
Figure 7.22 DEST profile of HypF-N in toxic and non-toxic conditions 234
Figure 7.23 β2 strand is in close proximity to the α1 helix 235
Figure 7.24 Residues that form the core of toxic (A) and non-toxic (B) HypF-N oligomers 237
Figure 7.25 HypF-N toxic and non-toxic oligomers have a similar core structure 242
Figure 7.26 HypF-N toxic oligomers are less flexible that non-toxic 243
Figure 7.27 Structure of DOPC and SM 245
Figure 7.28 HypF-N toxic oligomers are more mobile in a membrane environment 245
Figure 7.29 GM1 reduces the mobility of HypF-N toxic oligomers 248
Figure 7.30 Mobility of non-toxic oligomers with GM1 is similar to that of toxic oligomers in absence of GM1 249
Figure 7.31 Structures of PRE-spin labeled lipids 250

Figure 7.32 PRE effects from ^{13}C-^{13}C DARR spectra of toxic oligomers in presence and absence of GM1 252

Figure 7.33 PRE effects from ^{1}H-^{13}C INEPT spectra of toxic oligomers in presence and absence of GM1 253

Figure 7.34 PRE effects from ^{13}C-^{13}C DARR spectra of non-toxic oligomers in presence and absence of GM1 255

Figure 7.35 PRE effects from ^{1}H-^{13}C INEPT spectra of non-toxic oligomers in presence and absence of GM1 256

LIST OF TABLES

Table 1. Water hardness levels assessed by Unilever 65

Table 2. Time constant values (ms) for peaks which experience increases in signal intensity in presence of calcium 134

CHAPTER ONE

NMR:

PRINCIPLES AND TECHNIQUES

1.1 Principles of Nuclear Magnetic Resonance

A general report of the basic principles of Nuclear Magnetic Resonance (NMR), and a more detailed description of the specific NMR experiments performed in the thesis are described. The application of NMR in this thesis is used to study the conformation and dynamics of proteins based on solution conditions.

1.1.1 Nuclear spin

All spectroscopy is based on the transition between different energy states; NMR measures the difference in energy between states of nuclei orientated within an externally applied magnetic field (B_0). Nuclei possess an intrinsic property of spin, alongside mass, electrical charge and magnetism. The nuclear spin is dependent on the protons and neutrons that compose the nucleus, and has a quantum mechanical property: nuclear spin angular quantum number, or spin quantum number (I), which is a natural property of nuclei. I can have an integer or half integer values. Nuclei containing unpaired electrons (^{1}H, ^{15}N, ^{13}C) have non-zero ½ I values, along with nuclei with spin quantum values greater than ½ values such as ^{2}H, exhibit nuclear magnetic moments. The nuclear magnetic moment is described as a vector and defines its interactions with magnetic fields; this forms the basis of NMR spectroscopy. The magnetic moment associated with a particular nuclei (i) is proportional to its spin angular momentum:

$$\vec{\mu} = \gamma \vec{S}$$

where γ is the gyromagnetic ratio of the nuclei, which is a proportionality constant and an individual property of nuclear isotopes, and $\vec{S}$ is the spin angular momentum of the nucleus, another feature of nuclei.

Both magnetic moments and the spin angular momentum cause the nuclear spins to have a dynamical precession, the direction of the precession depends on direction of the spin.

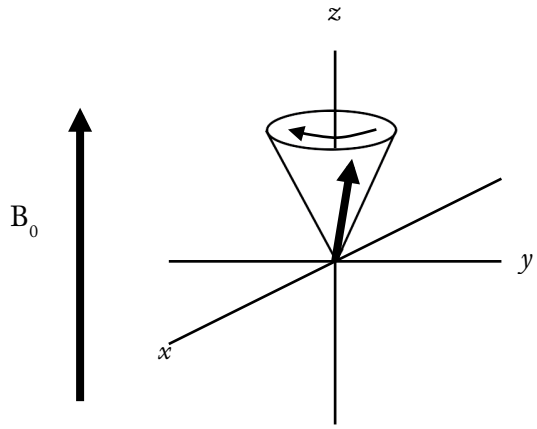


Figure 1.1 Nuclear spin precession

Nuclear spin precession about the direction of external magnetic field (B_0) following a cone at a constant angle to the z-axis, in this example the spin is aligned with the field.

NMR utilises the magnetic moments and nuclear spin, in an applied external magnetic field the precession of spin can be manipulated to study many characteristics of the sample. The atomistic detail and wealth of information obtainable makes it a highly popular biophysical technique to study biomolecular samples, as discussed further. This frequency of precession is termed the Larmor Frequency and can be defined as:

$$\omega_0 = -\gamma B_0$$

The Larmor Frequency describes the frequency of precession of nuclear spins and is dependent on both the gyromagnetic ratio (γ) specific to the nuclei and the magnetic field strength (B_0).

1.1.2 Longitudinal and transverse magnetisation

On a microscopic level, each individual nuclear spin is slightly fluctuating at a different magnitude and direction to its neighbour. Small fluctuations in the field due to the thermal environment cause a gradual breakdown of the spin precession, and over time, the angle between the spin angular momentum and the external field varies. At this state of thermal equilibrium there is a slight anisotropic bias for states to be

aligned parallel to the external field than parallel. This is due to the Boltzmann distribution, and leads to a polarisation of spins and a net macroscopic magnetisation moment (M) aligned parallel to the external field, termed longitudinal magnetisation. If this thermal equilibrium of spins is perturbed then the time taken for the spins to relax back to their populated distributions at thermal equilibrium is termed the longitudinal or spin-lattice relaxation (T_1).

Longitudinal magnetisation is almost undetectable, to circumvent this problem, NMR measures the nuclear spin magnetisation perpendicular to the field, rather than along it. To achieve this an external radio-frequency (RF) pulse is applied with a determinant length of time (90° pulse) to rotate M to be in the *xy* plane, perpendicular to B_0, the magnetic moments now precess, with coherence, around the *xy* plane where it is termed transverse magnetisation. Transverse magnetisation precesses under the influence of B_0 at the Lamor Frequency. Similar to T_1, transverse magnetisation relaxes; after the RF pulse is turned off the nuclear magnetic moments begin to lose their coherence and synchrony. Over time, the spin precession in the *xy* plane is lost as they return to equilibrium. This phenomenon is termed transverse, or spin-spin relaxation (T_2). It is the oscillating electrical current of the transverse magnetisation that is detected in the receiver detection coil as a signal of free induction decay (FID) as the NMR signal. A Fourier transform function is applied which converts the FID from units of time to frequency, giving the observable NMR spectrum.

1.1.3 Energy levels

The work discussed in the thesis does not require a quantum mechanical understanding of NMR principles, however certain principles are fundamental to NMR and must be addressed, they are presented below in a simplistic form.

Since NMR is based on the changes in wavefunctions from the transverse magnetisation oscillation, the properties of wave functions must be expressed as observable functions. Operators are mathematical descriptions of observables for any system. The operator for energy in a nuclear spin system is $\hat{H}$, called the Hamiltonian. The Hamiltonian function allows for the quantum mechanical expressions of kinetic

energy of motion. In an applied magnetic field (B_0), the Hamiltonian for a nuclear spin is termed the Zeeman interaction:

$$\widehat{H_z} = -\gamma\hbar\hat{I}_z B_0$$

where the energy of the system, and the wavefunction for all possible state of the spin system is dependent on the gyromagnetic ratio of the nuclei and planks constant ($\hbar$) and the applied field strength. $\hat{I}_z$ is the spin operator that represents the interaction between the spin angular momentum term (I) related to the z-component with B_0, which encompass the chemical shift and chemical shift anisotropic values. The total Hamiltonian ($\widehat{H_T}$) consists of contributions from chemical shift, chemical shift anisotropies, anisotropic dipolar couplings, anisotropic quadrupolar couplings and isotropic and anisotropic J-couplings (indirect dipole-dipole interactions mediated through bonding electrons between indirectly connected individual nuclear spins).

The eigenvalues of $\widehat{H}$ are the energy values associated with the possible states of the nuclei. Spin ½ nuclei can exist in two possible eigenvalues with two different energy levels at +½ (Nα) and -½ (Nβ), these states are referred to as Zeeman eigenstates and are aligned with or against the applied field. The energy level of each is dictated by:

$$E_{\frac{1}{2}\pm\frac{1}{2}} = \mp\frac{1}{2} - \gamma\hbar B_0$$

The excitation energy required to move one state (α) to the other (β), or the difference between the two energy levels is defined as:

$$\Delta E = \gamma\hbar B_0$$

As previously mentioned at equilibrium these states are distributed with a slightly greater population in the ½ (Nα) state in an applied field according to the Boltzmann's distribution. The populations of each eigenstate ($p\beta$ or $p\alpha$) can be written as an equation:

$$\frac{p\beta}{p\alpha} = exp\left(\frac{-\hbar\gamma B_0}{K_B T}\right)$$

where K_B is the Boltzmann's constant and T is the temperature.

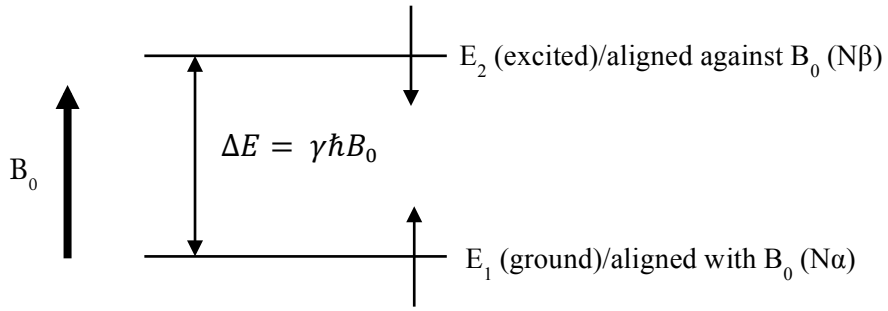


Figure 1.2 Orientation of states against B_0

The Lamor Frequency is proportional to the local magnetic field strength experienced by the nucleus (B_{local}), the strength of B_{local} can be affected by electron clouds surrounding the nucleus that chemically shield the nucleus thereby reducing the strength of B_{local} from B_0. Factors such as: chemical shielding, orientation of molecules to the applied field (known as chemical shift anisotropy), and chemical interactions such as through bond scalar connectivities (*J*-couplings) or through space dipolar connectivities (dipolar couplings) between proximal spins all influence the B_{local} and the eigenvalues for Nα and Nβ states. Different levels of chemical shielding from neighbouring groups, varying electron densities, hydrogen bonding and polar group interactions with polar solvents will cause an extrinsic shift in the signals detected and leads to the determination of structural information of samples (Keeler, 2014).

1.1.4 Chemical environments

The local chemical environment is unique for each nucleus, and thus its nuclear moment, and can be affected from intermolecular and intramolecular interactions. The NMR signal is therefore an excellent probe for detecting changes in nuclei based on the surrounding interactions and is a powerful tool for studying biomolecular systems at an atomistic level.

The factors discussed are utilised in the basis of all NMR experiments. NMR experiments build upon these basic principles by applying different selective or non-selective RF pulses; such as 90° pulses that creates coherence of the ±½ state functions (coherence of spins), and 180° pulses that invert the populations. Applying varied and multiple RF pulses can lead to complex manipulation of the nuclear magnetic properties and detection of specific parameters.

1.1.5 INEPT

A normal one-dimensional (1D) experiment provides information such as the chemical shift and spin-spin couplings. Attaining additional information from dipolar couplings, or *J*-couplings requires multiple RF pulses to be applied in multidimensional NMR experiments that use selective irradiation of particular resonances.

2D experiments are set up with a general pulse sequence involving periods of:

1) Preparation – delay allowing for thermal equilibration of sample
2) Evolution – coherence of nuclear spins in *xy* plane is achieved after one or several RF pulses.
3) Mixing – once the system has established correct state coherence is transferred between spins; determining the frequency pairs ω_1 and ω_2 (which define the 2D axes, e.g. $\omega_1 = {}^{1}H$ and $\omega_2 = {}^{15}N$ for a standard ${}^{1}H$-${}^{15}N$ HSQC)
4) Detection – the final step is detection of the FID during t_2, to build up a 2D spectrum t_1 is incremented between measurements to obtain a 2D data set

An example of a commonly used pulse sequence component and its application in a widely used 2D experiment is demonstrated below with INEPT and HSQC NMR.

The INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) pulse sequence was designed for solution NMR to give higher sensitivity, in particular for nuclei with low γ ratios; by transferring the coherence to a larger equilibrium magnetisation of a higher γ ratio nucleus (Morris & Freeman, 1979). Generally protons are used as high γ ratio nuclei, and the their magnetisation is transferred to ${}^{15}N$ or ${}^{13}C$ nuclei. The

transfer of polarisation is based on the scalar through bond *J*-couplings between the nuclear spins.

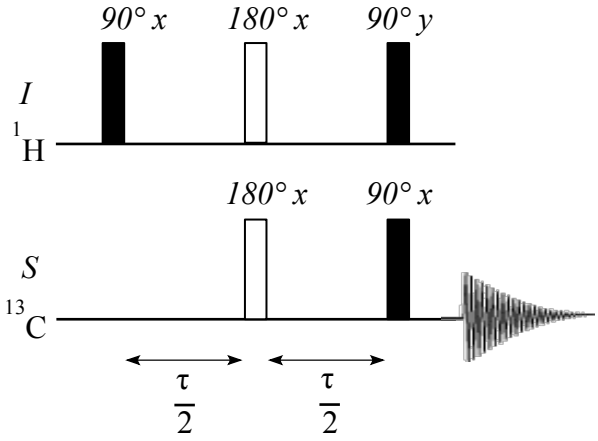


Figure 1.3 Pulse sequence of INEPT

During the first step the magentisation is evolved on the *I* spin by an initial 90° pulse along the *x*-axis as a spin echo pulse, this is transferred to the *S* state by two 90° pulses about the *y*-axis to the *I* spin. The final phase is another spin echo which gives an observable anti-phase signal on the *S* spin in the transverse plane, which is detected. The transfer occurs unambiguously via orbital overlaps of the electrons involved in covalent through bond connectivities, through the *J*-couplings. It is a fundamental pulse sequence for multidimensional heteronuclear correlations and connectivities. A universal experiment, the 2D Heteronuclear Single Quantum Correlation (HSQC) is based on the INEPT pulse sequence, and is often used to provide signal from ^{1}H and $^{15}N/^{13}C$ correlations. HSQC uses INEPT to transfer magnetisation from abundant *I* spins to the high-γ *S* spins, where it is evolved over a time period, and transferred back to the *I* spin for detection. The cross peak detected is from the heteronuclear coupling and correlation between ^{1}H and $^{15}N/^{13}C$, used to determine the chemical shifts of such nuclei. The high signal-to-noise and information accessible from HSQC experiments make it a widely used experiment across NMR groups.

The following sections discuss the specific NMR experiments used in the thesis; these experiments are not routinely used for NMR investigations. A brief introduction into

some fundamental aspects of solid state NMR (ssNMR) is described, along with an account of the two types of ssNMR experiments used within the thesis.

1.2 TROSY

Solution NMR spectroscopy of large proteins, generally above 25 kDa, is difficult even with the magnetic field strengths of 800 MHz. Large proteins are slower tumbling and have faster T_2, this leads to line broadening and poor spectral sensitivity. Larger proteins have shorter T_2 values; which is inversely proportional to the resonance line widths, exhibiting broader resonances and line widths (Fernández & Wider, 2003). T_2 relaxation is due to the loss of coherence of spins in the transverse (*xy*) plane over time, this is linked to the rotational diffusion of the molecule, characterised by the correlation time (τ_c) where:

$$\tau_c \propto 1/T_2$$

T_2 relaxation is a natural process of magnetisation returning to the equilibrium. The T_2 rates (R_2) are largely dependent on the chemical shift anisotropy (CSA) interactions between ^{1}H, ^{15}N and ^{13}C nuclei and the dipole-dipole interactions between ^{15}N and ^{1}HN of the amide or remote protons.

Dipolar relaxation occurs due to magnetic dipolar field fluctuations causing relaxation of nearby spins in a tumbling molecule. The strength of relaxation depends on the magnetic moment of the nuclei and the distance between nuclei (r^{-6}). Protons contribute the greatest source of dipolar relaxation, which is reduced in deuterated molecules. CSA relaxation is due to the fluctuations in the local magnetic field of nuclei. CSA involves only one spin so it cannot cause cross relaxation by transfer of the magnetisation to neighbouring spins. The rate of relaxation is linked to the constant $\gamma(B_0)^2$ so higher strength fields and gyromagnetic ratios nuclei are affected to a greater extent; along with R_2 increases with molecular weight.

For the case of a spin pair (αβ), and the energy transitions of a spin pair (αα, βα, αβ and ββ), cross correlation results in an asymmetric linewidth of the two peaks in both transitions, one becoming broader than the other. Cross correlations are interference

effects in both the longitudinal and transverse relaxation of coupled spins. The degree of asymmetry increases as the relaxation rates becomes greater. These effects combine and molecules with greater R_2 values lead to broader peaks.

The cross correlation acts to increase the relaxation rate of one spin pair state and decrease the other. The dipolar coupling and CSA interferes with the transverse relaxation, and resultantly the relaxation rates of individual components of α and β spins may be different. TROSY selects slower relaxing spin states over faster relaxing ones from these scalar coupled nuclei. Selecting the transitions with the sharpest linewidths can be achieved using two line-selective 180° pulses to manipulate the effect of cross correlation *(*Pervushin *et al*., 1997).

TROSY selects the decreased rate, and therefore has reduced levels of asymmetry and reduced line broadening. The correlation between the dipolar couplings and CSA act to cancel themselves out which is an effect of the cross correlation leading to only one spin state being selected giving a sharper line width and reduced resonance broadening.

Selective labeling can decrease the level of overlapping resonances and deuteration can reduce transverse relaxation rates by reducing the dipole-dipole relaxation between amide ^{1}HN and remote protons and minimising magnetisation transfer losses. These strategies are often costly and require extensive sample preparation. Transverse Relaxation-Optimized SpectroscopY (TROSY) is a very powerful and useful technique to reduce transverse relaxation rates of very large proteins without the requirement of deuteration or selective labeling. Its introduction has increased the number of NMR studies of proteins < 25 kDa (Pervushin K., 2000, Fernández & Wider, 2003).

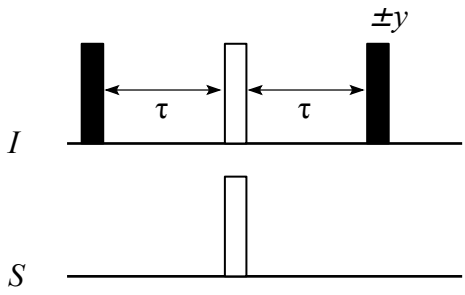


Figure 1.4 Simple line-selective 180° pulse sequence

Filled bars represent 90 ° pulses and unfilled bars represent 180° pulses.
Under this pulse sequence, one line of the I-spin doublet is selected. In a HSQC type experiment the 180° selective pulse is applied to both I- and S- spin doublets.

In a standard HSQC pulse sequence this is not applied, instead cross correlation is collapsed by applying an 180° pulse to the *I* spin for ω_1 in the τ_1 phase, and decoupling steps for *S* spins during signal detection for the ω_2 dimension.

1.3 T_2 relaxation measurements

NMR is a powerful technique to study the dynamical properties of biomolecular samples, ranging from timescales of picoseconds to seconds. It is the nuclear relaxation of spins in a magnetic field that is detected, commonly T_1 and T_2, providing relaxation data, chemical exchange and dynamical properties. T_1 or longitudinal relaxation is the spin-lattice relaxation, and corresponds to the reestablishment of the normal Boltzmann distribution of the α and β spin states to B_0. T_2 or transverse relaxation is the spin-spin relaxation and corresponds to the loss of coherence of spins in the transverse (*xy*) plane over time as magnetisation returns to the *z*-plane. Both forms of relaxation are linked since an increase in *z*-magnetisation cannot occur without a decrease in T_2, therefore $T_2 \leq T_1$.

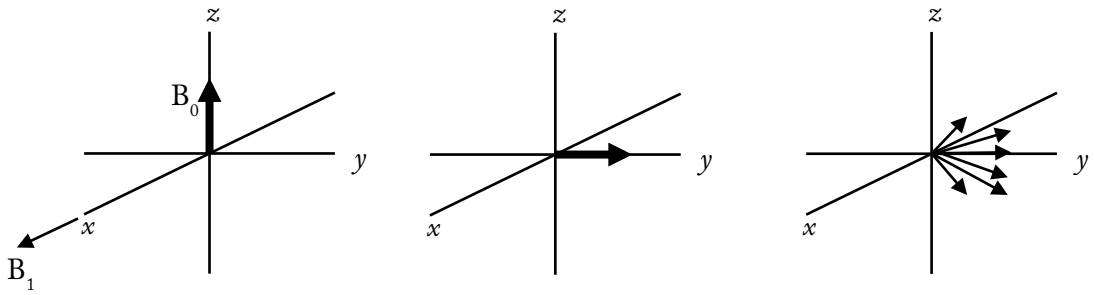


Figure 1.5 Transverse relaxation over time as the coherence of spins is lost in the xy plane

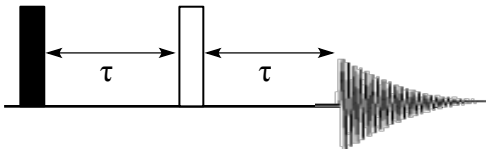


Figure 1.6 Simple spin echo

The spin echo refocuses the magnetisation so that when acquisition starts the size of the signal is dependent on only the transverse relaxation, which has occurred during 2τ. The peak height and intensity is therefore reflective of the transverse relaxation-induced decay. The decay of transverse relaxation can be determined from the peak height in the recorded FID of relaxation experiments. Analysis is based on the reduction of signal as a function of increasing delay times (~ 0 – 10 ns for T_1 and ~ 0 – 200 ms for T_2) to give time constant values, which represent the relaxation properties. Assigned spectra can indicate residues with relaxation properties, and can be mapped to tertiary structure. The linewidths and the time constants for T_2 are a very sensitive probe for nuclear motions on the microsecond timescale.

1.4 Dark-state Exchange Saturation Transfer (DEST)

Dark-state Exchange Saturation Transfer (DEST) is a 2D NMR experiment involving the transfer of saturation between two states, exploiting the exchange process and reduction in reorientational motions occurring between the binding of an easily observed monomeric state and an NMR invisible state (in this thesis the NMR invisible species is a protein protofibril) (Fawzi *et al*., 2010, Fawzi *et al*., 2011). NMR

invisible states are entities such as large molecular complexes, which have slow tumbling rates and rapid transverse relaxation properties that are orders of magnitude greater than monomeric proteins giving broad resonances.

^{15}N longitudinal magnetisation is prepared for the free observable monomeric protein species by applying a refocused INEPT pulse; this is transferred by chemical exchange as the monomer binds to the protofibril. The signal is usually unobserved in standard solution NMR methods due to the fast T_2 rates. To access this information a weak radio frequency (RF) offset from the ^{15}N carrier frequency is applied to allow for partial saturation of the ^{15}N longitudinal magnetization, which does not affect the monomer. During this period the monomeric species experiences a protofibril bound state. The partial saturation and magnetisation of the protofibril is transferred back to the monomer by chemical exchange as it dissociates and can be detected as an NMR observable signal from the free monomer species.

The magnetisation is initiated in the *z*-plane on the amide proton (^{1}HN) of the monomer. The first INEPT transfer creates a longitudinal two spin order, a second INEPT transfer gives an in-phase N magnetisation in the *z*-plane (N_2). Next a 90° proton pulse purges any residual antiphase second order states before a constant wave (CW) saturation pulse is applied at specified RF offsets from the ^{15}N frequency. During the CW phase, when the monomer samples the protofibril the Nz magnetisation is partially saturated, the rapid transverse relaxation of this state precludes it from direct observation. The saturation reduces the N_2 magnetisation to a level and rate, which depends on the strength and frequency of the ^{15}N saturation pulse along with the kinetic behavior between the oligomer-monomer exchange process.

The signal detected is from monomeric species only, since INEPT magnetisation transfer in a protofibril state from ^{1}HN to ^{15}N nuclei is very inefficient due to the rapid transverse relaxation rates of the protofibril. Detected signal is only from monomers that bind to the protofibril where transfer of saturation occurs, and dissociate within the saturation time. DEST allows for the detection of exchange process between monomeric and protofibril species on an atomistic level despite having a very large difference in the transverse relaxation rates between the two

species. The signal detected from the monomer is dependent on the transverse relaxation rates of the monomer bound to the protofibril: the stronger the decrease in linewidth intensity, the stronger the interaction to protofibril. The analysis is based on the attenuation of the cross peak signal from NMR observable monomer relative to a reference spectrum obtained without any saturation offset applied. Attenuation of signal indicates residues involved in the interaction between monomer and protofibril, this can be mapped to the sequence and structure to determine the interaction and potential site of interaction and nucleation from the monomeric state. These observations are residue specific, and not a global change consistent throughout the protein. Since there is sequence variation at each residue point we can be confident that changes are due to sequence dependent interactions with the invisible state entity. The presence of low molecular weight protofibrils would cause effects to DEST profile, which would be uniform throughout the sequence, these are not observed.

1.5 Solid state NMR (ssNMR)

In solution state samples, the isotropic re-orientations occur on a faster timescale than the relaxation processes measurable. This allows for an average of anisotropic contributions to be measured, which gives single peaks, dependent on chemical shifts and indirect spin-spin *J*-couplings. The anisotropic contributions are averaged out over the molecular orientations, and are therefore rarely observed. In contrast, solid samples are orientated with respect to B_0, and have an orientation dependent NMR signal. Powdered and solid samples used in ssNMR contain anisotropic interactions that possess orientational order due to the restricted motion in solids compared to solution NMR samples, and result in broad, convoluted peaks with poor resolution called powdered patterns. The anisotropic interactions arise from numerous contributions, which include chemical shielding, *J*-couplings, spin-rotation couplings, chemical shift anisotropies (CSA) and dipolar and quadrupolar couplings. It is the anisotropic contributions from molecular motions, which defines the broad distortion of the spectrum in solids, where broad pake powder patterns are observed from static powdered/solid samples (Duer, 2004). Complex methods are required to analyse and access the wealth of NMR interactions and structural information within solids. Magic angle spinning is a popular method to average anisotropic interactions by

rotating the sample at rapid speeds around an axis, rotating at an angle of 54.74° in respect to B_0 (Andrew *et al*., 1958, Andrew *et al*., 1959, Lowe, 1959).

The molecular orientation dependence of a sample can be defined as $(3\cos^2\theta - 1)$ where θ is the angle of the orientation of the spin tensor with respect to B_0. In solutions the molecular orientation is averaged out to zero due to the rapid isotropic tumbling. In a solid or powdered sample the orientation of the chemical shielding tensor can be any possible angle, since in powdered samples all molecular orientations are represented. Spinning a sample about an axis inclined at an angle θ_R to B_0 then θ varies with time as the sample rotates around θ_R. Spinning the sample at θ_R 54.74º, the 'magic angle', averages the anisotropic contributions and homogenous interactions such as homonuclear dipolar couplings to zero. Spinning at this angle, the angular dependence $(3\cos^2\theta_R - 1)$ is equal to 0 and averages the orientation dependent anisotropic interactions (Andrew *et al*., 1958, Andrew *et al*., 1959, Lowe, 1959).

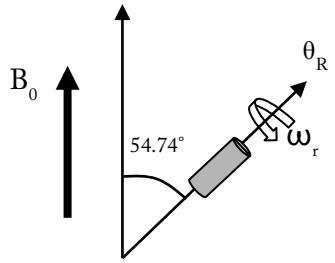


Figure 1.7 Representation of magic-angle spinning of a sample.
B_0 is the external magnetic field, and the sample spins at the angle 54.74° between B_0 and the axis of rotation (θ_R) at a spinning frequency of ω_r.

The anisotropy is averaged out, and peak splitting remain dependent on the isotropic interactions and chemical shift, resulting in single peaks comparable to that of solution NMR. The extent to which anisotropic interactions are averaged and the degree of line narrowing depends on the frequency of rotation about the magic angle. At rapid speeds of spinning (where ω_r is in the order of kHz) spinning sidebands and anisotropic contributions can be averaged, providing sharper linewidths.

1.5.1 Heteronuclear decoupling

A common problem for NMR is studying low abundant dilute spins such as ^{13}C or ^{15}N nuclei, strong heteronuclear dipolar couplings between these nuclei and abundant nuclei such as ^{1}H broadens the signal and leads to poor resolution. Even at high MAS frequencies these contributions cannot be fully averaged out to zero. A common method to circumvent this problem is to apply decoupling, by applying an RF irradiation to remove the couplings. This involves a high power continuous wave irradiation set at the Lamor Frequency of the ^{1}H spins. Examples include schemes such as TPPM (Bennett *et al*., 1995), SPINAL64 (Fung *et al*., 2000) or XiX (Detken *et al*., 2002) amongst others. The irradiation of the ^{1}H is transmitted to all ^{1}H spins across the network of dipolar couplings.

1.5.2 Cross polarisation

Cross polarisation (CP) is a key component of any solid state experiments and acts to utilise the strong dipolar couplings between two spins (*I* and *S*) such as abundant ^{1}H and dilute spins such as ^{13}C or ^{15}N with low γ nuclei. The problem observing low abundance spins such as ^{13}C is that they give low signal-to-noise ratios and fast relaxation times, these attributes give poor NMR sensitivity and long delays between scans. The MAS sensitivity of studying lowly populated nuclei and low-γ nuclei such as ^{13}C can be improved from exploiting the strong dipolar coupling between ^{1}H and low-γ nuclei, through cross polarisation (Pines *et al*., 1972, 1973).

CP is achieved by first applying a 90°_{x} pulse on the high abundant *I* spin ^{1}H, so that the ^{1}H magnetisation is alone the *y*-axis. A spin-lock field on-resonance contact pulse (B_1) is applied to the ^{1}H magnetisation acting to maintain it along the *y*-axis (called $B_1(I)$). This is similar to B_0, and maintains the magnetisation of the *I* ^{1}H spins in the absence of any RF pulses. Analogous to the ^{1}H spins, the same contact pulse is applied to the low abundant spin *S* with spinlock field B_1S present. The dipolar coupling network between the *I* and *S* spins maintains the energy and the angular momentum of the *I*–*S* spin system. The polarisation is transferred from I to S spins by matching the Lamor Frequencies of both *I* and *S* spins to match the Hartmann-Hahn condition listed below (Hartmann & Hahn, 1962):

$$\gamma_I B_1(I) = \gamma_S(B_1(S)$$

The signal enhancement from CP is proportional to $\gamma_I/\ \gamma_S$, for ^{1}H to ^{13}C the maximum theoretical enhancement is by a factor of 4.

1.5.3 DARR

The overlapping ^{13}C 1D spectra of solid samples can be resolved further in homonuclear dipolar recoupling experiments for ^{13}C nuclei. Experiments such as proton driven spin diffusion (PDSD) (Szeverenyi *et al*., 1982) and dipolar assisted rotational resonance (DARR) (Takegoshi *et al*., 2001) can be performed on solid samples using MAS to determine homonuclear ^{13}C-^{13}C correlated cross peaks. The polarisation utilises ^{13}C-^{1}H dipolar interactions and is a combination of rotor driven and ^{1}H driven mechanisms. Recoupling methods are used to reintroduce the weak homonuclear dipolar couplings, which are lost or attenuated under MAS conditions. In DARR ^{13}C-^{1}H dipolar interactions are recoupled by applying a continuous RF irradiation pulse on ^{1}H.

The pulse sequence for DARR consists of a CP step to transfer polarisation to the ^{13}C from ^{1}H spins through dipolar couplings before an evolution period, t_1, is set. A 90° pulse is applied to the ^{13}C spins creating ^{13}C magnetisation along the *z*-axis. This is followed by a delay time (mixing time) during which a DARR field is applied to ^{1}H spins at the MAS frequency. Afterwards a final 90° pulse results in the ^{13}C magnetisation being in the transverse *xy*-plane where the signal is detected. Increasing the mixing time increases the range to which correlations are observed. At short mixing times the transfer occurs mainly between adjacent nuclei, providing short-range intra-residual connectivites (<~3Å). Increasing the mixing time increases the range to inter-residual connectivities (>~3-4Å). When implemented into 2D experiments, a 2D ^{13}C-^{13}C DARR spectrum can yield information pertaining to carbon (^{13}CO, $^{13}C\alpha$ and $^{13}C\beta$) resonances and can indicate inter- and intra- residue correlations.

1.5.4 INEPT in solids

Heteronuclear ^{1}H-^{13}C/X INEPT experiments can be performed in solid-state samples under MAS conditions. Analogous to solution experiments it involves the basic coherence transfer of nuclear spin from the *I* to *S* state via through bond connectivities and *J*-couplings. MAS does not eliminate *J*-couplings therefore they do not rely on MAS averaging out these contributions as is the case for CSA. Transference of solution INEPT experiments to solid sample applications was initially difficult to achieve due to the rapid decay of spin coherences in solids. This was reduced by applying strong homonuclear ^{1}H decoupling and refocused INEPT pulse sequences, to allow for well resolved multidimensional experiments such as INEPT ^{1}H-^{13}C HSQC (Bax & Morris, 1981).

CHAPTER TWO

LIPEX:

INTRODUCTION

Chapter summary

The content of this thesis is focused on studying the conformational states of proteins under different solution conditions. The effects of protein structure, dynamics and conformational flexibility are examined using two different enzyme systems, selected to study different aspects and consequences in enzyme functionality. The first project utilises a bioindustrial lipase enzyme called Lipex, in collaboration with an industrial partner, Unilever. The effects of Lipex conformation, structure, stability, dynamics and activity will be studied under industrially relevant solution conditions. The second project studies the effects of protein solubility on an enzyme HypF-N under specific solution conditions that leads to non-pathogenic cytotoxic aggregation.

This chapter details the contextual research and background knowledge of Lipex. Lipase enzymes are introduced, as well as the structural and conformational changes that occur within this class of enzymes, which relate to catalytic activity, and their application to industrially used detergent lipases. Solution conditions and factors influencing lipase activity are discussed, and the project aims are listed at the end.

2. Protein structure and dynamics

Structural dynamics play a significant role in many fundamental aspects of protein science; backbone and side chains fluctuations are crucial for folding, function and thermodynamic stability. Over the last decade substantial evidence has emerged supporting the critical role conformational flexibility, spatial and temporal changes play in a number of protein molecular functions; from ligand or substrate binding (Beach *et al*., 2005) to molecular recognition and allostery (Weber, 1992) (Creighton, 1993). Enzyme catalysis, a highly dynamical process, is functionally dependent on conformational changes for ligand binding and release, substrate affinity, signal transduction and atomic fluctuations for new bond formations within the catalytic regions and rate of product release (Boeh *et al*., 2006). The determination of the structural dynamics of proteins at atomic-resolution is key for a comprehensive understanding of the molecular mechanisms governing biological process.

The depths of our knowledge and understanding of protein dynamics are limited to

the few experimental techniques available. Whilst large-scale conformational changes, such as haemoglobin binding oxygen in the blood (Fersht, 1999), have been determined from static X-ray structures, understanding the pathways and transition states is highly difficult as these species are short-lived and lowly populated. Moreover, it is clear that highly dynamical systems require an ensemble description rather than static structural determinations. There is a wealth of information within these conformational ensembles (Alexandrescu *et al*., 1993); however, accessing it poses a difficult question. Ensemble-averaged experiments, common to structural biology, have been shown to elude significant conformations (Rajagopalan, 2002), potentially losing valuable structural information. A challenge is therefore defined to extract information on individual conformational species and populations from ensemble averaged measurements. An alternative approach includes single molecule measurements such as fluorescence measurements (Wang & Geva, 2005, Chang & Rosenthal, 2013), however these techniques have limited spatial resolution. Among the techniques available to ascertain information on protein dynamics, NMR offers a powerful approach, as it provides information at an atomic level, on different timescales (from 10^{-12} to 10^{5} seconds) and without the need for bulky and potentially interfering probes, e.g. as those associated with Förster Radius Energy Transfer (FRET) fluorescence measurements or electroparamagnetic resonance. NMR structural studies of proteins that undergo conformation changes are of known difficulty due to poor ability to measure nuclear overhauser effect (NOEs) and the lack of efficacy of dihedral restraints from chemical shifts (Camilloni *et al*., 2012). Studying these proteins therefore requires novel approaches to infer structural dynamics.

2.1 Solubility of enzymes

Proteins with increased dynamical and conformational flexibility properties are more susceptible to sampling conformations that may promote aggregation or reduce solubility. This is deleterious for proteins, and enzyme functionality and catalytic activity can be compromised. A more significant and harmful gain of function can be the formation of toxic pathological aggregates from the misfolding and aggregation of proteins such as Alzheimer's β-peptide or α-synculein responsible for the pathological diseases Alzheimer's and Parkinsons. Studying the effects of solubility requires

biophysical techniques that can measure protein dynamics, to track the changes in conformation during aggregation.

2.2 Aggregation

Enzymes exhibit very high activity under varying experimental conditions such as temperature, pressure and aqueous medium; conditions which are exploited within industrial applications to ensure optimal performance of both enzyme and product. These conditions, however all can have an effect on protein solubility and rates of aggregation.

The link between protein solubility, aggregation and function is explored within this thesis. Two separate protein examples are used to study the effects of solution conditions of protein conformation, leading to a) modulation of enzyme activity using a bioindustrial lipase enzyme (Lipex) and b) non-pathogenic aggregation of a well-studied protein (HypF-N). Both projects explore the crucial dependency of protein functionality on solubility and how aqueous conditions can affect this fine balance. Studies on Lipex will investigate how protein functionality and enzyme catalytic activity is dependent on solubility within the context of an industrial biocatalyst. Studies on HypF-N will focus on the role of protein aggregation within amyloidogenic proteins relating to pathogenic amyloid diseases, providing insights on a significant field of human health and disease. From both projects we can observe the effects of solubility directly on the structure, conformational states, dynamics, function, activity and aggregation.

2.3 Lipases

Lipases are ubiquitous soluble triacylglycerolacyl hydrolases (E.C. 3.1.1.3) that catalyse the hydrolysis carboxyl ester bonds of long-chain insoluble triglycerides to form di- and mono- glycerides, glycerol and free fatty acids (Brockman, 1984) (Figure 2.1).

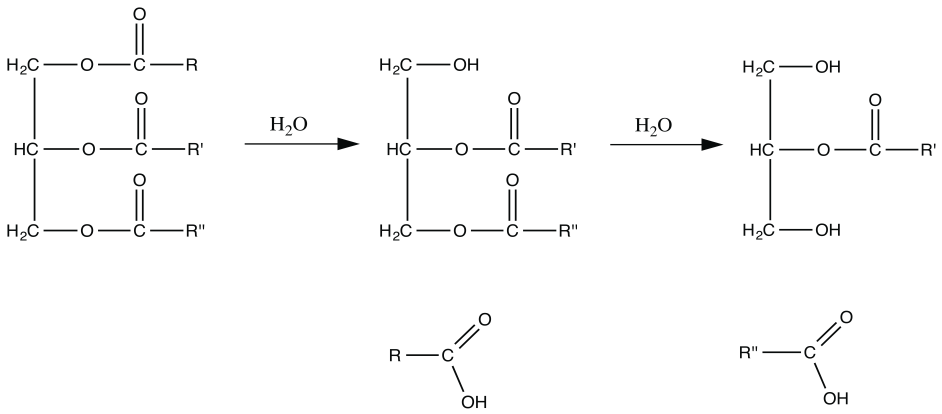


Figure 2.1 Lipase hydrolysis of triglyceride molecule

Simplified two step schematic of lipase with sn 1,3 specificity catalysed hydrolysis of a triglyceride molecule to a digylceride and free fatty acid group with R group, and subsequent hydrolysis to monoglyceride and free fatty acid with R'' group. Where R is the (CH_2)n chain

Lipases can also catalyse the transesterification of other esters leading to the synthesis of esters, and the acyl- or deacylation of a range of unnatural substrates. Their range of biocatalytic abilities (Figure 2.2) with high specificity, stability, selectivity and efficiency, has made them an extremely useful and popular multifaceted enzyme for industrial usage. These applications range across biodiesel, agrochemical, food, cosmetics, detergents, chemical and pharmaceutical reactions (Haki & Rakshit, 2003).

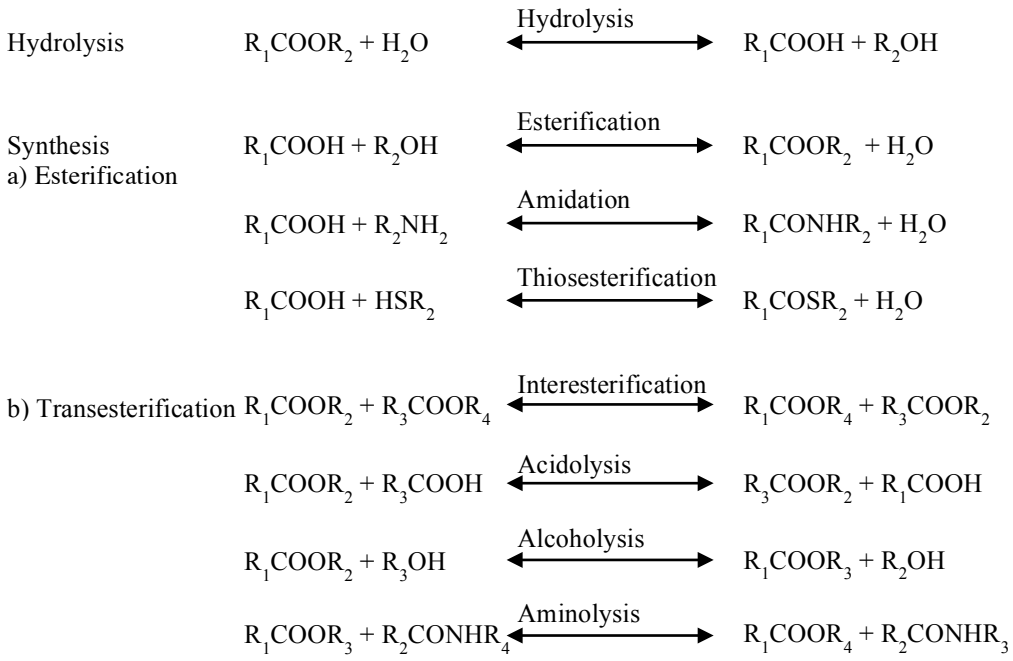


Figure 2.2 Different reactions catalysed by lipases

Lipases share a similar α/β hydrolase tertiary scaffold fold, a common fold across numerous enzymes, which would suggest common evolutionary relationships, despite showing little amino acid sequence similarity (Nardini & Dijkstra, 1999). The fold provides a stable but flexible conformation for a range of enzymes to perform their catalytic activities. This is significant for lipases, which require large conformational changes and wide substrate selectivity. A key exception to variability and something conserved across all α/β fold hydrolase enzymes is the positioning of the third α-helix in the sequence, which is centrally located within the structure and plays a role in arranging the catalytic nucleophile (Schrag & Cygler, 1997).

The active site of lipases is formed of a catalytic triad and contains the same handedness in all lipases; it is located within a nucleophilic motif present on the 3rd helix. The triad consists of Ser-Acid-His where the serine acts as nucleophile, the acidic residue is usually aspartate and the histidine is a present as a base. The triad is part of a conserved pentapeptide sequence of Sm-X-Nu-X-Sm (where Sm denotes a small amino acid usually glycine but can be alanine or serine, X any amino acid, and Nu is the nucleophile: serine or sometimes aspartate or cysteine) (Lawson *et al.*, 1994). The pentapeptide forms a strand-turn-helix motif and residues +/- 2 and 6

positions from the nucleophile are structurally constrained to be small hydrophobic residues as to allow for hydrophobic interactions packing of the β-strand against the α-helix. The hydrolytic activity of lipases on triglyceride substrates is shown in detail in Figure 2.3.

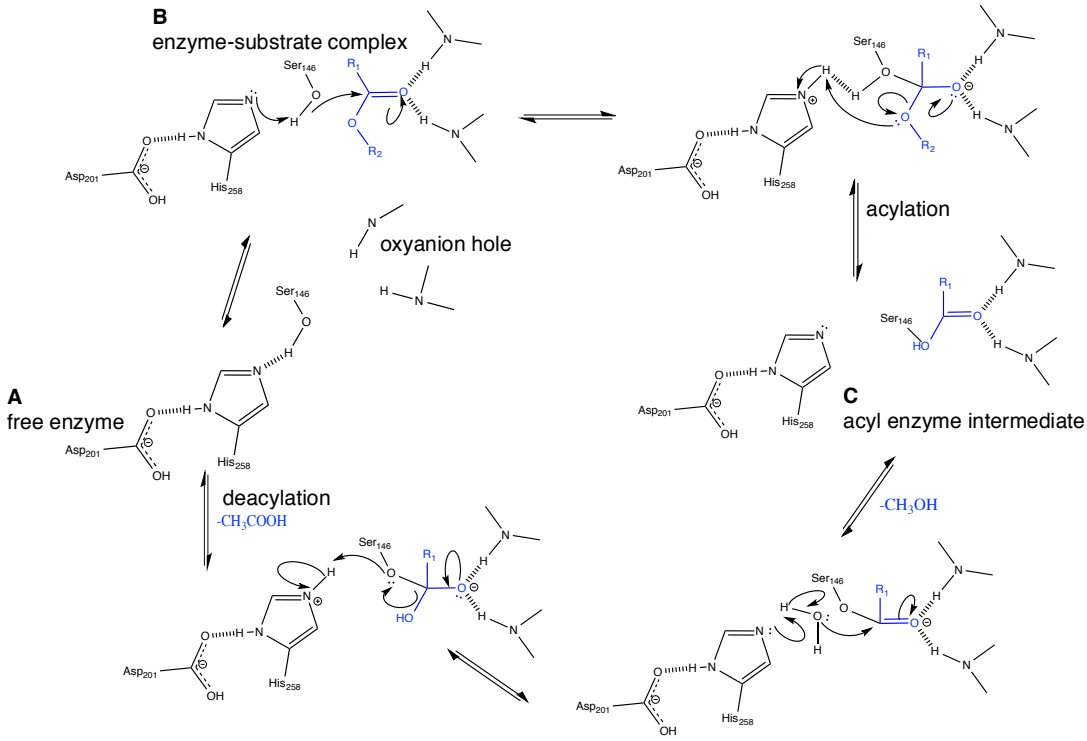


Figure 2.3 Detailed mechanism of lipase catalytic activity

Residue numbers shown correspond to the catalytic triad of TlL. Figure adapted from Rauwerdink & Kazlauskas, 2015. The starting point (A) depicts the catalytic triad of the enzyme (Asp201-His258-Ser146) in its free form and a vacant oxyanionic hole from two main chain amide groups. The substrate binds to the active site to form an enzyme-substrate complex (B). The ester first interacts with Ser146 (the most exposed catalytic residue, Brzozowski et al 2000) and the His258 acts as base to deprotonate the serine. The reformation of the ester carbon-oxygen bond double bond releases methanol and forms the acyl enzyme intermediate (C). Water binds to the acyl enzyme intermediate and the His258 again as a base deprotonates the water that can then attack the carbonyl carbon of the acyl enzyme intermediate. Deacylation occurs as the last step with the release of acetic acid as the carob-oxygen double bond is reformed and the free enzyme state is established.

2.4 Industrial lipases

Lipases hold numerous specific attributes against other enzymes that make them highly suited to industrial usage such as high catalytic activity and chemical, entantio- and regioselectivity, mild reaction conditions, low environmental impact, and unlike most enzymes, are catalytically active against wide range of substrates (Chen & Sih, 1989).

A range of lipases are used as biocatalysts within industry, lipases from the fungal genus *Aspergillus* such as *A. oryzae* and *A. niger* are used to produce esters in the production of foodstuffs (Contesini *et al*., 2010). Lipases are particularly useful in pharmaceutical industry and within the manufacturing process of synthetic chemical compounds since they have a high enantioselectivity for racemic and prochiral compounds (Gupta *et al*., 2007). Within the cosmetic industry the esters of carboxylic acid and alcohol moieties are used as additives and lubricants (Haki & Rakshit 2003), and the lipase Novosym 435 from *A.* Niger is used in the biodiesel industry to produce fatty acid methyl esters from rapeseed oils (Linko *et al*., 1994). The lipase of this study originating from *Thermomyces lanuginose (T. lanuginose)* is used in detergent and laundry.

Whilst lipases can be located within the world's flora and fauna, industrially used lipases are produced by different sources including plants, mammalian and microbial such as bacterial, fungal and yeast origins. The microbial forms, especially extracellular glycoprotein bacterial lipases, are most commonly used in biotechnological applications due to easier bulk production through fermentation methods (Singh & Mukhopadhyay, 2011). The bioindustrial applications of enzymes, such as lipases, are reported to be multibillion-dollar industry, making them a heavily researched commercial product within industry (Joseph *et al*., 2008).

Generally most industrial processes are performed at temperatures above 45 °C, requiring a heat stable lipase with optimum activity at around 50 °C (Sharma, 2002). Whilst thermostability is an important attribute, the search for highly active lipases that exhibit stability across pH, ionic strength and against organic solvents is ever present through either new sources or bioengineered variants (Lee & Rhee, 1993).

2.5 Detergent industry

Hydrolytic lipases are an essential component of detergents for laundry washing given their ability to hydrolyse fats; those used in commercial industry have low substrate specificity, high proteolytic resistance, and are stable across a range of pH, temperature and damaging surfactants known to inhibit enzyme activity (Sharma *et al*., 2001, Macrae & Hammond, 1985). Lipases used in detergents account for almost 32% of all total lipase production and sales (Jaeger & Reetz, 1998). For viable usage within an industrial application, biocatalysts must offer high process efficiency along with the ability to be produced in mass and in a cost effective process. It is estimated that almost a hundred thousand tons of lipase is produced each year to be used in detergents, such a vast amount is reflective upon both the high global usage of detergents and the short half-lives of lipases in harsh solution and conditions encountered within detergent usage. Most biocatalyst lipases are of microbial and fungal origins and expressed through fermentation using fungal hosts to accommodate for the high amounts required.

In addition to lipases, detergents contain a number of enzymes such as proteases, amylases and cellulases that are present to cleave protein, starch and cotton respectively. Lipases are unable to digest most fatty deposits on their own, the presence of surfactants in the formulations enhance the capabilities of lipases to achieve fully water soluble products (Macrae & Hammond, 1985). Consequently lipases used in detergents need to be stable and active in the presence of other enzymes and surfactants, along with solution condition factors associated with laundry such as temperature and pH previously alluded to. Studying the structural changes of lipases under these industrial conditions will provide better understanding of the stability, activity and functionality of these enzymes.

2.6 Thermomyces lanuginose lipase (TlL)

The industrial lipase studied within this thesis named Lipex, is a modified version of a lipase originating from the fungal organism *Thermomyces* (formerly *Humicola*) *lanuginose (T. Lanuginose*) (TlL). TlL is a mono-gylcosylated 269 amino acid extracellular lipase (Boel & Huge-Jensen, 1988) with a pI of 4.4 (Omar *et al*., 1987a,

Omar *et al*., 1987b, Jha *et al*., 1999). TlL is capable of degrading a range of substrates, via interfacial activation (Martinelle *et al*., 1995), and preferentially hydrolyses fatty acid chains from the *sn1* and *sn3* positions of the triglyceride molecule (Lawson *et al*., 1994b). TlL has greater activity on increasing carbon chain lengths substrates from 2-8 (Martinelle *et al*., 1994) but the preferred substrate for TlL is trilaurate, a glyceride composed of three 12-carbon fatty acids (Novo-Nordisk, unpublished data). The structure of TlL was determined in 1994 by X-ray crystallography at 2.7 Å (Lawson *et al*., 1994b), and has since been refined to 1.84 Å (Derewendra *et al*., 1994), in a number of conformations varying in lid open (Figure 2.6), lid closed, at high and at low ionic strength buffer (Brzozowski *et al*., 2000). The structure has three disulphide bonds (Lawson *et al*., 1994b), with one being structurally significant (Cys22 - Cys268) (Figure 2.10) (Brzozowski *et al*., 2000).

TlL is roughly spherical ellipsoid with a size of 35 Å x 45 Å x 50 Å and an average radius of 23 Å, and is structurally composed of 10 β-sheet strands with 6 interconnecting α-helices. The helical lid occluding the active site is located from amino acids 86 – 92 (Glu-Asn-Trp-Ile-Gly-Asn). The lipase catalytic triad Ser – His – Asp (Glu) is highly conserved and is usually shielded by the helical lid from the external environment (Lawson *et al*., 1994, Brzozowski *et al*., 2000). The lid region undergoes structural changes during interfacial activation, with lid closed or open conformations both present in dynamic equilibrium (Lawson *et al*., 1994a). The tryptophan at position 89 is linked to catalytic activity, where cationic residues are responsible for controlling the interfacial activation (Skjold-Jørgensen *et al*., 2015).

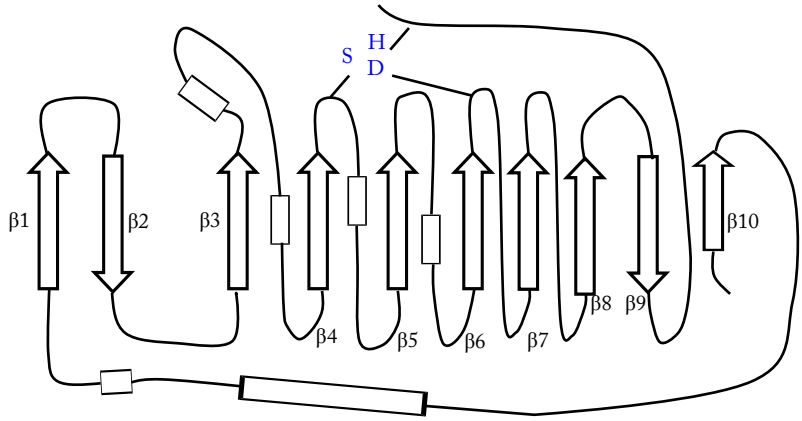


Figure 2.4 TlL secondary structure topology

Catalytic residues shown in blue and β-strands numbered

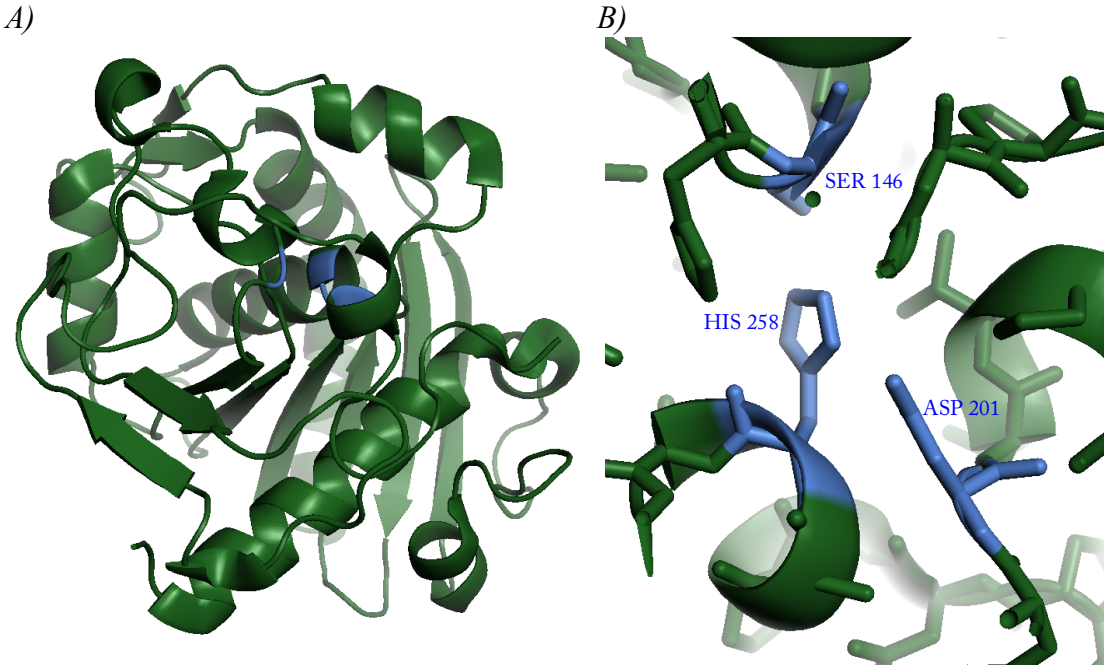


Figure 2.5 TlL structure, catalytic residues highlighted

TlL in an open state with catalytic domain highlighted in blue, shown in A) as full enzyme and in B) the zoomed catalytic region and residues shown in blue. PDB code: 1GT6 (Brzozowski et al., 2000).

The active centre and catalytic triad of the lipase is shielded from the external environment by a helical lid (Bourne *et al*., 1994, Noble *et al*., 1993). The lid varies in size between lipases from different origins. The lid is mobile and in the presence of a

hydrophobic substrate, or even any hydrophobic interface, the lid moves to allow a hydrophobic interaction to occur between the hydrophobic substrate with hydrophobic residues of the catalytic core. This breaks the hydrophobic interaction that is formed between the lid and lipase active centre resulting in a thermodynamic penalty. Lid movements and structural rearrangements are discussed further in section 2.10.

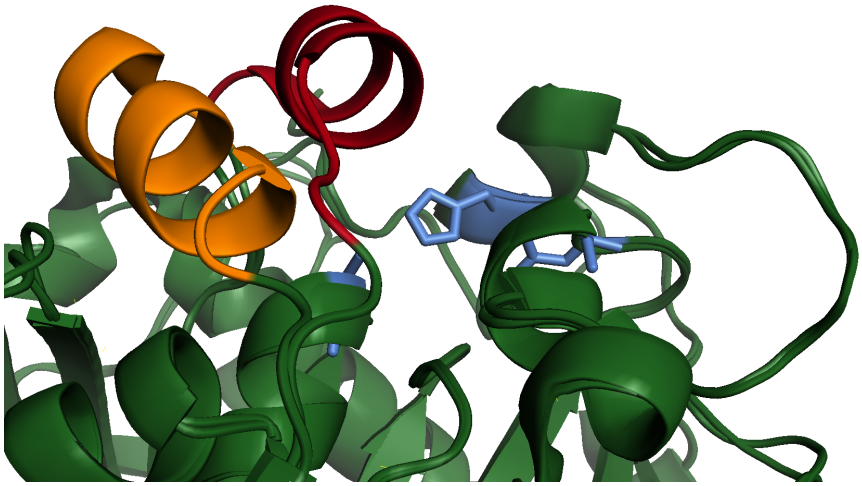


Figure 2.6 Helical lid domain of TlL in open and closed conformation

Helical lid domain of TlL in open (PDB code 1GT6) (orange) and closed (PDB code 1DT3) (red) conformation, the latter blocks access to the catalytic residues shown in blue. Structures determined by Brzozowski et al., 2000

TlL was the first lipase to be used in laundry tablets in 1988 by Unilever, selected over alternative lipases for its collection of unique attributes. Originating from a mildly thermostable organism, the mesophilic protein has a high melting temperature of 74 °C at pH 8 (Zhu *et al*., 2001), and retains 100% activity up to 65 °C (Omar *et al*., 1987a). The thermostable property of the protein allows for the study of changes in structure, function and catalysis over a much greater range of temperature. In addition to temperature stability, TlL exhibits high kinetic stability against anionic surfactants (surface active agent) allowing it to function in conditions that would denature alternative lipases, such as *Yarrowia liplytica*, which undergoes interfacial denaturation (Aloulou *et al*., 2007). The kinetic stability of TlL is attributed to a high activation barrier between native and denatured states, despite a change in overall Gibbs free energy favouring the unfolded state (Wang *et al*., 2013). These properties make the enzyme an ideal candidate for usage in industrial detergents.

TlL is industrially produced in enormous scales of approximately several hundred tonnes per year. There has been much research into the immobilization of the enzyme, often preferred for industrial biocatalysts for usage in chemical reactions since it can be reused over several reactions, however in the detergent usage only the enzyme is not immobilized and not reusable. TlL has been strongly researched, given its thermal and kinetic stability it makes it a very good candidate for use in the detergent industry used in washing powder for stain removals, TlL has a commercial name of Lipolase ® within the detergent industry. It has undergone directed evolution and protein engineering to improve its catalysis and stability (from development of Lipolase ®, Lipolase Ultra ® and LipoPrime ® to Lipex ®) (Svendsen *et al*., 1997, Houde *et al*., 2004) and to increase its acid resistance for *in vivo* gastro-intestinal usage (Wang *et al*., 2013). These changes include removal of negative surface charges, improving lid opening and activation, stability and activity, which resulted in comparable wash results at reduced temperatures (Svendsen *et al*., 1997, Houde *et al*., 2004). Engineering industrial enzymes for higher activity at lower temperatures is of critical importance to detergent industry given the environmental and global impact.

2.7 Lipex

Lipex, manufactured by Novozyme and registered under Unilever copyright patents, is the latest generation of modified TlL, the third variant from directed evolution of LipoLase®. Lipex has cleaning capabilities at wash temperature as low as 20 °C and can trap and break down fat stains entrapped in cloth fibres (Houde *et al*., 2004). Lipex is currently not commercially used, but current research performed by Unilever is performed on the Lipex variant, to better understand the enzyme system and potential advancements in product design before its commercial introduction. Unilever have not set a release date for commercial introduction of Lipex, but our studies will focus on the improved Lipex variant of TlL. Lipex is an industrially sensitive enzyme, and specific details cannot be disclosed. To this end, much of the available information presented is taken from published material on the TlL enzyme, which retains significant similarities to Lipex. Since the sequence is industrially sensitive material it is not disclosed. Lipex is a 29.4 kDa (~32 kDa when glycosylated) protein, with all conserved functionally and structurally significant

residues of TlL, directed evolution has led to several key amino acid mutations not disclosed.

2.8 Interfacial activation

Lipases are classed as a specific class of esterases with a unique interfacial activation (Sardra & Desnuelle, 1958) consisting of slight structural changes during activation in presence of lipids, at a water-lipid interface, which increases their activity (Schrag & Cygler, 1997). Interfacial activation is defined as the enhanced activity of lipases on insoluble substrate compared with similar substrate in its monomeric form. Colleagues Sarda and Desnuelle defined the term interfacial activation, and much of the pioneering aspects, in the 1960s, however it is a complicated mechanism still not fully understood, The role of interfacial activation is, in sense, an evolutionary advantage as lipase substrates are biophysically prone to molecular aggregation in aqueous dispersions.

Interfacial activity is still not fully understood, as it does not conform entirely to the simplicity of standard enzyme theories. This may possibly be due to the dynamic nature of the process and the majority of structural data of lipases being generated through static X-ray crystallographic studies. The dynamic structural fluctuations are difficult to distinguish in X-ray studies, NMR, however, is able to detect the conformational and lowly populated transient species as an average within an ensemble. The lid domain for TlL has poor electron density from residues 82 to 99, and considered to be significantly disordered within electron mapping (Derewenda *et al*., 1994).

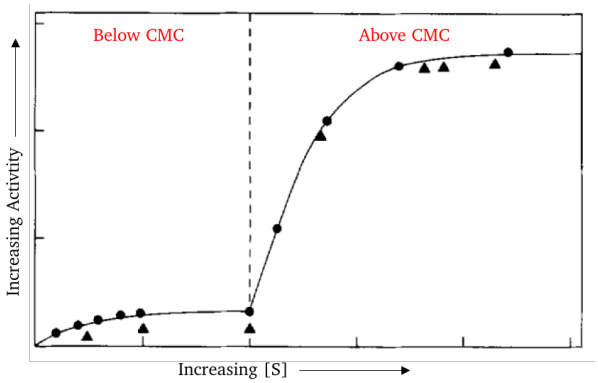


Figure 2.7 Enzyme kinetic profile during interfacial activation

An increase in substrate concentration in monomeric form leads to low increase in activity when below its CMC value. When the substrate concentration exceeds the CMC, lipases can bind to the lipid-water interface and reach the fully activated state. Figure formally adapted from Verger R. (1997). Interfacial activation of lipases: Facts and artifacts. TIBS TECH ***15****, 32–38.*

Whilst most enzymes are active at interfaces, i.e. at the surface of membranes, interfacial activation is a phenomenon and property of most lipases, where the activity of lipases is enhanced upon insoluble substrates. This enzymatic mechanism is not to be regarded as an increase in catalytic activity on the enzyme; it deviates from standard Michaelis-Menten kinetics where the Michaelis constant K_m, which has a volume concentration, cannot be expressed with an insoluble substrate, and any noted increase in reaction rate is due to the increase in interfacial activity. The rich heterogeneity, dispersion and complexity of the chemical and physical properties of lipid interfaces makes them unsuitable for Michaelis-Menten kinetics which apply to enzymology in more homogenous solutions. There is still debate as to whether the kinetic properties displayed for lipases undergoing interfacial activation is specifically due to the conformational changes of the lid region of lipases (Aloulou *et al*., 2006).

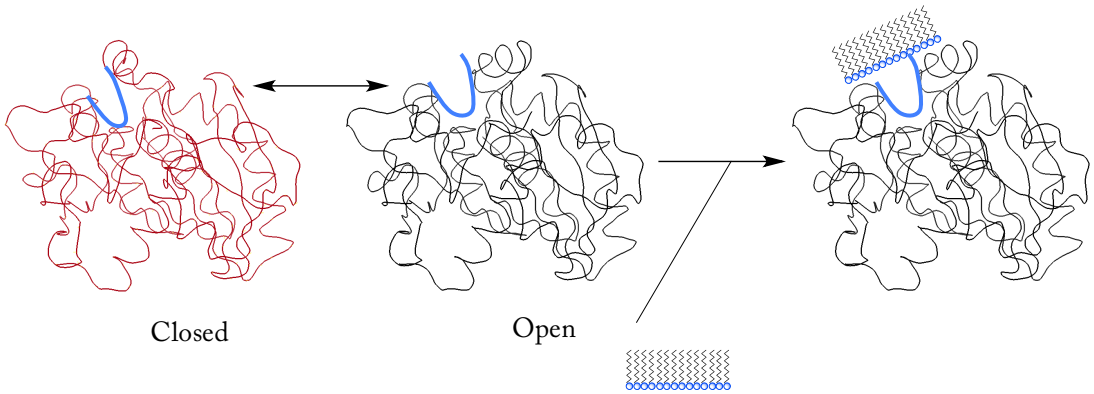


Figure 2.8 Lipid interface allow for interfacial activation

TlL structure is in equilibrium between the open (black) (PDB code: 1GT6) and closed (red) (PDB code: 1DT3) state, presence of substrates or surfactants in a micelle state stabilize the open conformation allowing for interfacial activation.

The activation of lipases increases the accessible hydrophobic surface and active site, during lid displacement, water molecules bound to the surface of the enzyme allows binding to a lipid interface. This interaction is greatly complex and involves free energy changes including hydrogen bonding, van der Walls, electrostatic interactions, solvation and hydrophobic interactions. The extent of the free energy changes from removing polar residues and increasing exposure of apolar residues is dependent on the lipids involved and the interface. Expectantly, conformational changes, disordering water molecules, rotations and vibrations provide an entropic cost (Peters *et al*., 1998).

The mechanism of interfacial activation, and lid domain movement has been studied previously for TIL, generally through co-crystallization with detergents or substrate analogues (Brzozowski *et al*., 1991, Brzozowski *et al*., 2000, Hedin *et al*., 2002, Derewenda *et al*., 1994 Lawson *et al*., 1994, Aloulou *et al*., 20007, Martinelle *et al*., 1995) and electron spin resonance (Hedin *et al*., 2005). The conformational change in TlL, specifically the lid domain movement, has also been studied using fluorescence studies (Yapoudjian *et al*., 2002, Jutila *et al*., 2004, Skjold-Jørgensen *et al*., 2015, Skjold-Jørgensen *et al*., 2016). Specific structural details are discussed further in section 2.11.

2.9 Lipase dimers; self aggregation at interfaces

TlL, along with lipases from *Candida rugosa* and *Mucor miehei* can form soluble aggregates with itself, with the large hydrophobic surface at the active core at the centre of the interface, to form biomolecular aggregates, interfacially activating each other (Palomo *et al*., 2003). These dimer aggregates exhibit altered functional properties since their active sites are partially occluded.

It would appear that there is a functional relevance for this phenomenon, since the activity of the enzyme is critically linked to the open conformation it seems evolutionarily irrational to obscure the active centre. The dimer formation has a higher stability; at lower concentrations (0.09 mg/ml) where the enzyme is in a predominantly monomeric and closed form, the residual activity reduces from 100% to 40% over a time of 30 hours (h) incubated at 50 °C; whereas at higher concentrations in a dimeric state maintained 80% activity over 30 h at 50 °C (Palomo *et al*., 2003). Whilst the dimer formation does show increased stability TlL is expressed in its monomeric form and dimer formation is not noted in literature in the conditions generally used for studying TlL. High ionic strength conditions can increase the propensity for aggregated lipases (Palomo *et al*., 2003) Dimer formation has been noted in other fungal lipases such as *Rhizomucor miehei* where only purification at pH 4.5 yielded the monomeric form (Boel *et al*., 1988). It may be that for TlL dimer formation is not induced under physiological or experimentally constructed conditions.

2.10 Structural conformations and re-organisations of TlL

It is likely that lipases exist in aqueous solution in a dynamic equilibrium between open and closed states (Brzozowski *et al*., 2000), given the flexible and dynamic nature of the lid domain. During transition from open to the closed conformations the protein undergoes structural changes. The extent of the lid movement is highlighted in Figure 2.10, where TlL is shown in open (orange) and closed (red) structures, the two structures are overlaid illustrating there are not any other major changes to the structure aside from the lid movement.

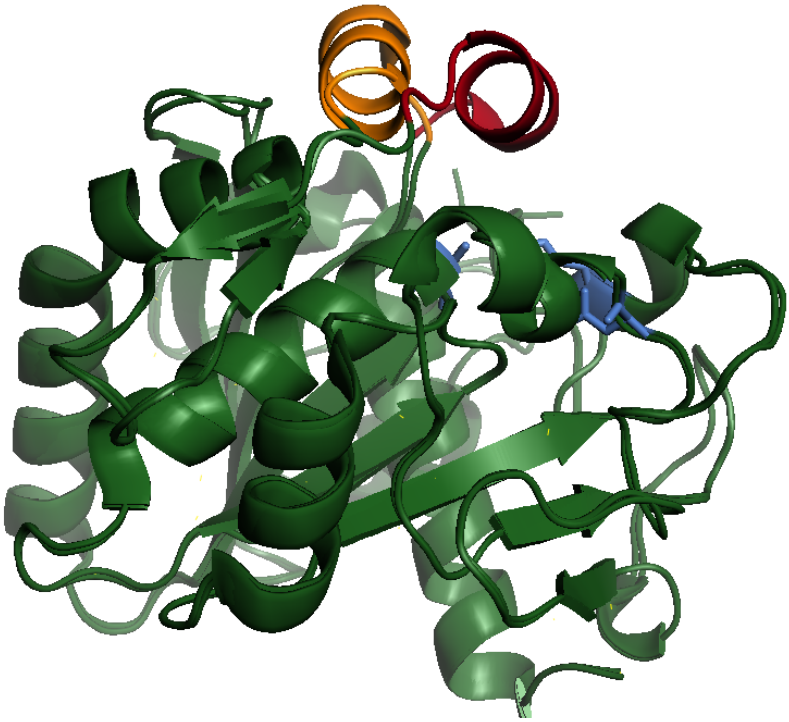
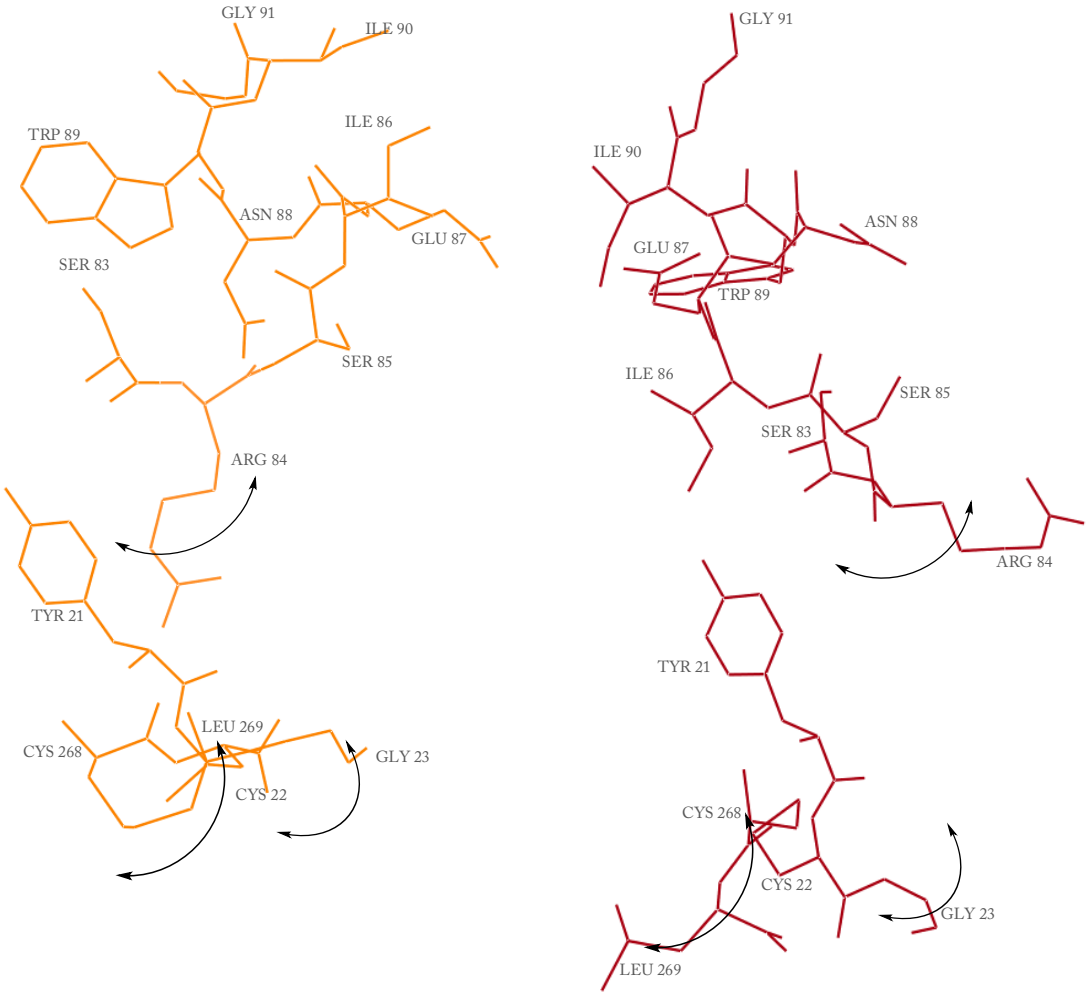


Figure 2.9 Structural change of TlL from open and closed conformation

Overlay of TlL in open (PDB code: 1GT6) and closed (PDB code: 1DT3) conformations To highlight the extent of the structural change between open and closed conformations both structures are overlaid in green with the lid domain residues 82 – 99 coloured for open (orange) and closed (red). The catalytic triad are coloured blue for geographical identification.

During the structural change from open to closed there is disulphide cis-trans isomerization between the cysteine residues Cys22 and Cys268 from a standard left handed *ng⁻ g⁻* disulphide to *pg⁻t* right handed form which mediates much of the changes in TlL. This is accompanied with a new hydrogen bond formation that provides structural stability in this region of the enzyme (Brzozowski *et al*., 2000). This disulphide bond between Cys22 and Cys268 is intimately linked to Arg84, which undergoes the greatest movement during interfacial activation (Figure 2.10), flipping about its Cα atom and changing orientation by ~ 120° (Brzozowski *et al*., 2000). In the closed state the NH_2 group of Arg84 can form hydrogen bonds with Cys268, and disrupt the neighbouring Ser85 residue interaction with Asn88 which is considered the most important interaction in the open state (Brzozowski *et al*., 2000). Arg84, along with other cationic residues Arg81 and Lys98 are responsible for forming

electrostatic interactions during lid opening with the negatively charged substrate or lipid (Berg *et al*., 1998, Cajal *et al*., 2000, Hedin *et al*., 2002). These changes, whilst being < 5 Å distances in structural and conformational rearrangements may be subtle but are functionally significant to TlL. The pH can affect the charged state of Arg84 and may disrupt the hydrogen bond between Arg84 and Asp57, inducing the Arg84 switch and movement from pH rather than interfacial activation (Derewenda *et al*., 1994).



OPEN CLOSED

Figure 2.10 Structural rearrangement of key residues of TlL from open to closed conformation

The structural changes occurring through transition between open and closed states, specifically residues Arg84, Cys268 and Gly23 (the latter two as part of disulphide bond isomerization). Arrows indicate the most significant structural movements.

The site that predominantly blocks access to the active site is part of the hydrophobic region composed of residues His258 (from the catalytic triad); Trp89 and Leu93 of the lid domain and Val203, Leu206 and Ile255 from the core (Lawson *et al*., 1994). The region surrounding the active site in the open conformation is mostly non-polar, but there are distinct hydrophobic areas, which are well defined in crystal structures, more so than the non-polar regions (Lawson *et al*., 1994, Derewenda *et al*., 1994). It is hypothesised that the substrate interfaces interact with the non-polar regions, and the acyl chains bind to the hydrophobic patches; the polar charged and main chain carbonyl groups prevent tight binding of acyl chains to these sites. The exposure of these hydrophobic patches is aided by the orientation of the catalytic Ser146 being in close proximity (Hedin *et al*., 2002). Whilst there is inherent flexibility of the substrate upon binding hydrolysis at *sn1* and *sn3* positions of the triglyceride may be preferential to TlL since there is steric clashing at the *sn2* position from the catalytic Ser146 being in close proximity (Lawson *et al*., 1994).

2.11 Significant residues of TlL

There are several key residues within the TlL sequence, which provide functional significance for catalytic activity. The surface tryptophan at position 89 (Trp89) is involved at the interface between lipid and enzyme; the hydrophobic part of the amino acid can interact with the lipid whilst the ring nitrogen can interact with the carbonyl regions of either the substrate, in particular carbonyl oxygen atoms at sn2 or sn3 chains in the active site or the main-chain of TlL (Lawson *et al*., 1994). It has been shown to influence substrate binding at the active site in a qualitative rather than quantitative manner (Holmquist *et al*., 1994). The lid tryptophan is generally well conserved across fungal lipases such as *R. Miehei* and *Penicillium camembertii* (Derewende *et al*., 1994, Thirsrup *et al*., 1994). It is not however essential to lipolytic function, the lipase from *R. delemar* has evolved without a tryptophan residue in its lid domain (Martinelle *et al*., 1996).

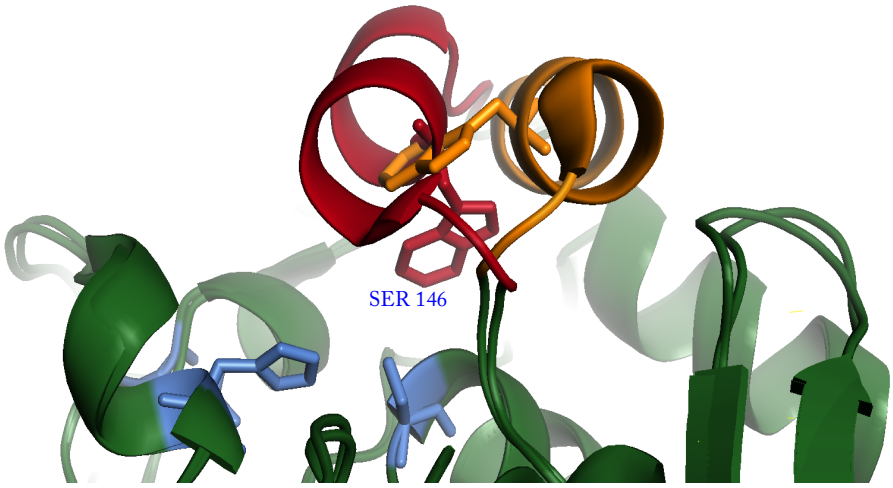


Figure 2.11 Tryptophan 89 is a critical amino acid in the lid domain

Tryptophan residue 89 on helical lid domain indicated for open (orange) (PDB code: 1GT6) and closed (red) (PDB code: 1DT3) conformations of TlL, with catalytic residues in blue and Ser146 labelled

Substitution of Trp89 to the more hydrophobic, but less reactive, phenylalanine showed that activity to different length trigylcerides were affected; with shorter chains that may not be able to form the close interactions with the phenylalanine being more affected (Martinelle *et al*., 1995). Mutating Trp89 to leucine affects the structural stability of the enzyme, presumably from changing the hydrophobicity and the π-cation interactions that may be formed between tryptophan and proximal basic amino acids. Trp89 also behaves differently under thermal unfolding and chemical unfolding with guandadine hydrochloride (Zhu *et al*., 2001).

Since the lid can move freely when TlL is in complex with chlorophosphonate ethyl ester (Lawson *et al*., 1994) it is believed that there are no intermolecular contacts formed between Trp89 and the substrate, but it is considered that Trp89 plays a role in guiding the acyl chain into the correct position in the active site for catalysis (Martinelle *et al*., 1996). Mutating Trp89 to leucine reduces the thermal stability of TlL from 74.4 °C to 68.5 °C, increasing the number of conformational fluctuations with temperature and decreasing the number of conformations with a functional catalytic triad (Zhu *et al*., 2001).

Another significant residue within the lid domain is the glutamate at position Glu87, which stabilizes the lid in the open conformation forming electrostatic interactions in the hydrophobic environment between its carboxyl group and the surface of the protein at the exposed active site; it also stabilises the transition state during reaction (Holmquist *et al*., 1995). Mutations of Glu87 (to alanine) cause a significant loss of activity in all substrates regardless to acyl chain length, affecting K_M and the release of the product (Martinelle *et al*., 1995).

The Ser83 residue of the catalytic triad acts to form a hydrogen bond with its hydroxyl group and amide nitrogen and the tetrahedral oxyanion transition state when TlL is in its open conformation (Lawson *et al*., 1994). It is a crucial residue with all fungal lipases having a hydroxyl group at the position of Ser83 for TlL (Derewendra *et al*., 1994); mutation to alanine forms an inactive enzyme (Mogensen *et al*., 2005). Fluorescence spectroscopy and MD simulations on the Ser146Ala mutant show that the loop region is more flexible, but the lid hinges are more flexible in the wild type enzyme, leading to significantly different conformations, and resultantly different binding affinities. (Peters *et al*., 1998).

2.12 Surfactants

Surfactants (surface active agents) are tensioactive chemicals, whereby they form a mono-molecular layer at the surface interface and cause changes in the surface tension. They are water-soluble at concentrations of micro- to millimolar ranges above their critical micelle concentration (CMC), forming micelles around 20 nm in length and soluble crystals (Sammalkorpi *et al*., 2009). Detergents play a crucial role in lipase activity in industrial applications, acting to stabilize (Jutilia *et al*., 2000), increase lipase activity (Fernandez-Lorente *et al*., 2007) and even act as inhibitors for TlL (Mogensen *et al*., 2005). The work presented herein looks to further understand the structural basis and mechanism of these interactions and the role of interfacial activation on a dynamical and functional level, to improve the usage of Lipex under conditions relevant within the detergent industry. Another popular method for increasing catalytic efficiency and thermal stability is through protein engineering strategies and direct evolution and informed site-directed mutagenesis (Shinkai *et al*., 1996, Li *et al*., 2014, Bornscheuer *et al*., 2012). The work will focus on the non-

covalent interactions of Lipex with surfactants, which can potentially aid in the next phase of site-directed mutagenesis to further optimise these interactions and enzymatic activity.

Surfactants can vary in a number of terms including hydrophobicity, charge, CMC and ionic charge; all of which can have effects on their interaction with lipase adsorption rates and substrate hydrolysis activity. They consist of a polar head group that can vary in charge, and a hydrophobic chain, and are generally classed on charge: anionic, cationic, zwitterionic and non-ionic. Due to the inherent conformational flexibility within the active centre of lipases, environmental conditions have a noticeable effect on lipase, structurally and functionally. Surfactant binding to lipases and proteins and complex formation is generally driven by hydrophobic interactions, and electrostatic and hydrogen bonding (Antipova *et al*., 2001, Aloulou *et al*., 2007).

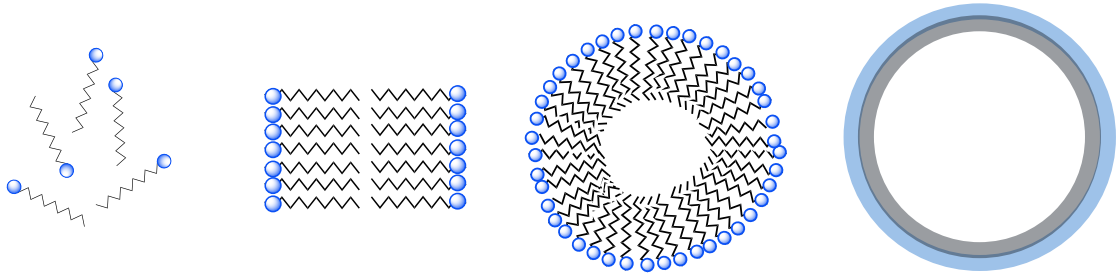


Figure 2.12 Different surfactant configurations

Surfactants can form a number of configurations from monomeric state (A), to bilayer (B), micelle (C) and microemulsions (D). These can interact with lipases and proteins in different modes of mechanism and cause different responses in activity and stability.

Interactions between proteins and surfactants are multifaceted, dependent on the protein surface and potential, surfactant structure, charge, monomeric or micellular state. Additionally solution conditions such as pH, temperature and ionic strength can affect the interactions between surfactant and protein. The level of ions can control the association and disassociation of molecules in micelles by affecting the electrostatic interactions (Miller *et al*., 2000). Low ionic strength solutions decrease the propensity for ionic micelles to aggregate (Sammalkorpi *et al*., 2009). Surfactants can affect the solubility and co-solubility of proteins and their self-association

properties in aqueous media reducing the propensity for aggregation (Antipova *et al*., 2001). Furthermore, surfactants can alter the interfacial tension thus affecting lipase activity (Sonesson *et al*., 2006). This effect arises from replacing the protein at the surface of the interface and, in turn solubilising the protein (Wahlgren & Arnebrant, 2003). Interactions with surfactants can affect conformational stability and tertiary structure, interfacial activation, hydrophobicity of the protein surface, and adsorption rates. Surfactants have a remarked effect on the surface activity of proteins (Aloulou *et al*., 2007, Miller *et al*., 2000, Antipova *et al*., 2001) and since the catalytic activity of lipases is intimately linked to the surface and lid domains, it would be predicted surfactants to have an influence, negative or positive, on lipase activity. TlL has a very low surface pressure threshold, above which the enzyme activity is detectable, which is equivalent to high surface free energy where proteins may be unfolding; indicating that TlL maybe undergoing conformation changes and to a degree of unfolding at interfaces (Rogalska *et al*., 1995). This is a factor, which may be reduced or inhibited by the presence of detergents (Aloulou *et al*., 2007). Unlike other lipases such as from *Yarrowia lipolytica*, TlL exhibits a decrease in lipolytic activity as detergent concentrations increase, even at levels below their CMC (Aloulou *et al*., 2007). Since detergents can have an effect below their CMC, the monomeric form may form a complex with the lipase via its hydrophobic parts, solubilising it (Aloulou *et al*., 2007, Antipova *et al*., 2001). The inhibition occurs from the detergent blocking the binding of the enzyme at the lipid substrate interface; it is unknown whether the enzymatic inhibition is due to the interaction of enzyme to detergent, or the reduced rates of enzyme adsoption, or a combination of the two (Aloulou *et al*., 2007).

There are a number of advanced techniques used by numerous groups to study interactions and adsorption rates for surfactant and protein complexes, for both industrial, medicinal and academic knowledge (i.e. mediation of proteins through bilayers) (Otzen, 2011). Techniques such as harmonic perturbations, isothermal titration calorimetry (Otzen *et al*., 2009), fluorescence measurements such as FRET (Sonesson *et al*., 2006), small angle X-ray scattering (Giehm *et al*., 2010) and molecular modelling can be used to study the dynamical, adsorption, relaxation and thermodynamic properties of such interactions (Miller *et al*., 2000). There have been NMR based studies performed to study the time-resolved unfolding mechanism of α-

lactalbumin in presence of SDS (Malmendal *et al*., 2010), however this was limited to 1D ^{1}H NMR analysis.

Increasing concentrations of ionic and non-ionic surfactants can have a negative effect on adsorped proteins, and proteins may become less surface active due to competitive adsorption from the surfactant displacing the complex from the interface (Miller *et al*., 2000). Enzymes, even highly thermostable versions of TIL, such as Lecutase Ultra (LU) which is a modified phospholipase of TlL with the starting 284 amino acids, can be unfolded by forming clusters with detergents such as the anionic sodium dodecyl sulphate (SDS), even at concentrations well below its CMC (Madsen *et al*., 2015), where SDS has a significant effect on secondary and tertiary structure. For the case of TIL, increasing surfactant concentration to values at which micelles would form, regardless of ionic or non-ionic form, caused increases in inhibiting activity compared to monomeric detergent values (Aloulou *et al*., 2007).

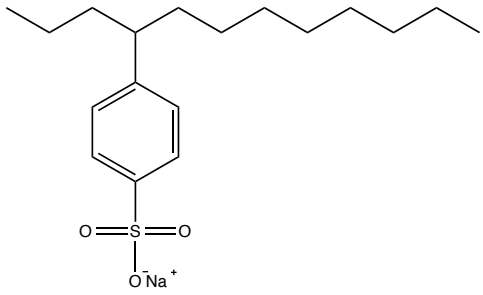


Figure 2.13 Linear Alkyl Benzene Sufonate (LAS) (C_{12} alkyl chain)

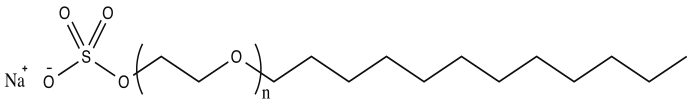


Figure 2.14 Sodium Lauryl Ether Sulphate (SLES)

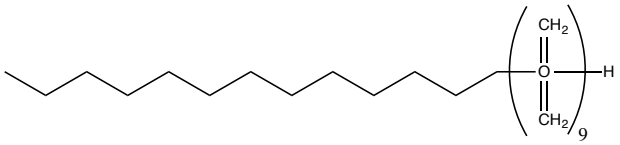


Figure 2.15 Neodol 25-9 (NI) (C_{13} alkyl chain)

Unilever use a combination of anionic (LAS, Figure 2.13 and SLES, Figure 2.14) and non-ionic (Neodol (NI), Figure 2.15) surfactants within their products to achieve the

optimal catalytic activities of TlL variants. Within the liquid form of washing detergents, surfactants are already able to interact with detergent enzymes, potentially either inactivating them via denaturation or unfolding, or increasing solubility and stability (Madsen *et al*., 2015). This is reduced in powder form as enzymes are prepared as dust free particles, which are protected by sugars and a hydrophilic environment to prevent interactions and damage (Otzen, 2011). Understanding the interactions of surfactants and lipase enzymes in the storage condition and in during the wash are two different fields of research, both require detecting changes in the structure, stability and activity of the enzyme.

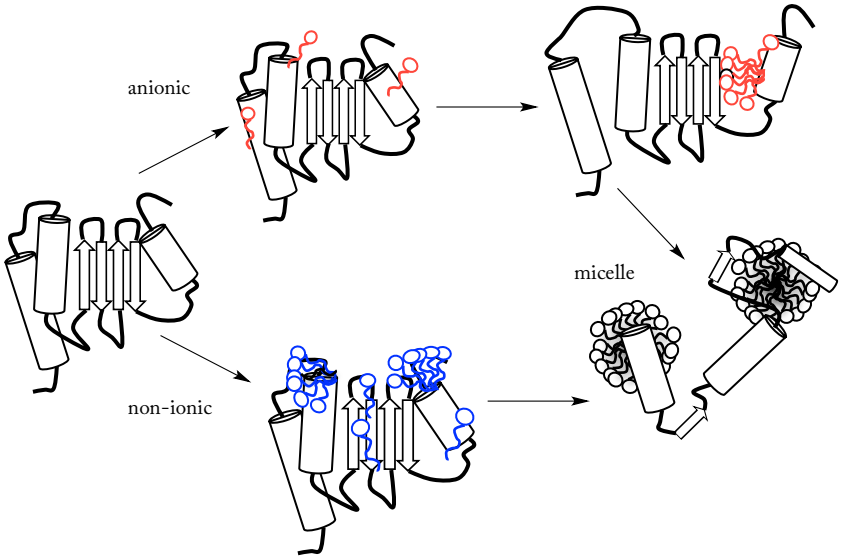


Figure 2.16 Model mechanism of anionic and non-ionic surfactant interactions

Model mechanism of anionic (red) and non-ionic (blue) surfactant binding to protein in monomeric states and inducing unfolding in micelle stae (black). Anionic surfactants have a marked effect on structure, compares to uncharged surfactatns and monomeric states, however micelles are combined much faster for non-ionic where unfolding can then be instigated. Figure adapted from Otzen, 2011

Non-ionic surfactants such as TWEEN 20, triton X100 or Neodol form complexes with proteins through hydrophobic interactions only by competitive adsorption between the molecules, and at certain concentrations will exclusively cover the interface (Miller *et al*., 2000). Non-ionic surfactants can bind in monomeric form and can promote and cooperate micelles to bind and increase denaturation rates (Otzen, 2011). Anionic surfactants interact with proteins via their negatively charged sulphate

headgroup, and through hydrophobic and electrostatic interactions (at cationic hot-spots) across its alkyl chain allowing binding to occur across numerous sites of the protein (Otzen, 2011). The binding of anionic surfactants to proteins proceeds via a number of steps, at monomeric states they can bind and illicit conformational changes via electrostatic interactions with cationic side chains such as lysine, arginine and histidine. All of which are present in the proximity of the lid domain of TlL (Lys98, Arg84 and His258). The alkyl chain will bind to hydrophobic patches, in the case for TlL the newly exposed hydrophobic region upon lid opening. Increasing the concentration of surfactant can cause clusters to form and induces unfolding (Figure 2.16) (Otzen, 2011). The concentrations vary on the type of surfactant and the stability of the protein. The production and synthesis of anionic surfactants such as LAS is demanding and often inefficient.

One facet of research by Unilever is the implementation and introduction of biosurfactants into their products for greater sustainability and environmental goals. One class of biosurfactants of potential interest and consideration is the anionic rhamnolipid (R).

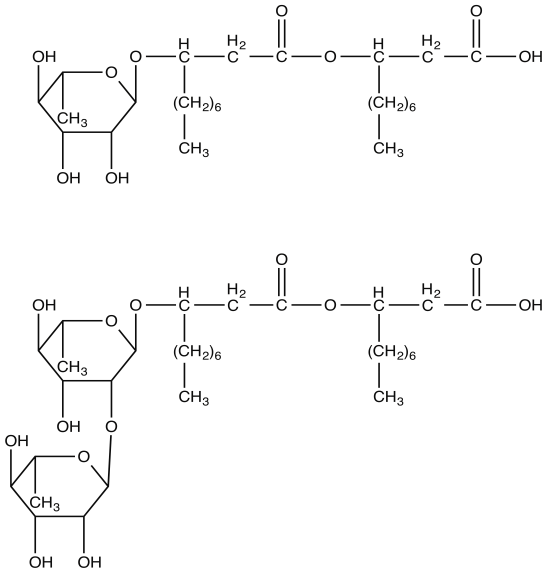


Figure 2.17 Rhamnolipid JBR425 structure of mono and di-rhamnolipid

Rhamnolipids are a glycosylated biosurfactant with a carboxylated anionic group from *Pseudomononas aeruginosa*, and can be present in mono- and di- forms with MWs of 504 and 605 Da respectively. They contain a L-(X)-rhamnose and β-

hydroxyalkanoic acid unit (Müller *et al*., 2012). Generally, most surfactants are derived from complex chemical reactions using petrochemical recourses, which is both non-renewable and expensive, given the increasing rise in oil prices. Biosurfactants are a promising new class and next generation of surfactants, sourced mainly from microorganisms offering a cheaper, globally accessible and renewable synthetic production, reducing costs and environmental impacts. They are non-toxic, have low biodegradability and allergenicity, and are able to be customized by chemo-enzymatic reactions to form modified derivatives which is advantageous within a cosmetic, food and laundry industrial application (Patel *et al*., 2003, Müller *et al*., 2012). They also have higher tolerances to pH, temperature and salinity; which is a significant attribute for use in detergents in hard water conditions (Chen *et al*., 2013). Rhamnolipids are a very promising class of biosurfactants compared to others such as mannosylerythritol lipids, sophorolipids and surfactin lipopeptides; the rising number of applications and cost effective production has been reflected with rhamnolipids having the highest reported number of publications and patent applications of all glycosylated biosurfactants (Müller *et al*., 2012, Leitermann *et al*., 2010). Generally anionic surfactants, such as sodium dodecyl sulphate (SDS), which whilst they are highly effective surfactant with great surface activity, are also considered more aggressive towards proteins and enzymes, inducing unfolding or denaturation of tertiary and secondary structure due to the disruption of electrostatic interactions (Madsen *et al*., 2015). Rhamonolipids appear to be a class of anionic surfactants which do not destabilise protein structures, and can even increase thermal stability in both monomeric and micelle states (Madsen *et al*., 2015). This is not concurrent with enzyme activity, at monomeric rhamnolipid concentrations activity of Lecitasae Ultra was reduced to 40%, where as at higher values exceeding its CMC, activity was increased up to 200%. These changes have been predicted to arise from the competition for binding at the hydrophobic exposed regions at the active site between substrate and rhamnolipid (Madsen *et al*., 2015).

2.13 pH

Lipases are affected greatly by physico-chemical factors such as temperature, pH and the presence of lipids and inorganic salts (Brune & Gotz, 1992, Aires- Barros *et al*., 1994). TlL has an optimal pH range between 7 – 8.5, but is active at higher pH values,

this is particularly useful for detergent industry applications since many other fungal lipases from *C. rugosa* and *R. oryzae* hosts exhibit a rapid decrease in activity above pHs of 8.1 and 8.5 respectively (Zheng *et al*., 2011, Minning *et al*., 1998). Most lipases exhibit optimum activity at pH values of 7, which corresponds to a pK_a values of 6.5 for the ionized catalytic histidine residue. At higher pH values the ionization state of fatty acids are increased leading to an enhancement in lipase activity, which is advantageous to TIL. Changes in pH effects both the enzyme turnover rates, and the adsorption step during interfacial activation (Aloulou *et al*., 2007); affecting conformational changes in the lid domain and the ionization state of free fatty acids, which to a lesser extent may trigger lipase adsorption.

Whilst pH is a macroscopic property affecting the entire protein molecule and not just the active site; the pH of the solution will affect the protonation state of the catalytic His258, it is a complex process since the closure of the lid creates a lack of chemical communication between the active site and the bulk pH. In the closed state, by either lid closure or aggregation, it has been postulated that the cavity is only large enough to accommodate only one positively charged entity; making analysis of pH changes difficult to accurately report (Poulsen *et al*., 2005).

2.14 Calcium

The hardness of water can play a crucial role to the enzymatic activity of lipases; the level of multivalent cations present defines water hardness, with calcium and magnesium divalent cations the most predominant ions originating from calcium and magnesium bicarbonate sources. There are numerous ways of quantifying water hardness, all based on the total concentrations of calcium and magnesium ions expressed in either parts per million, mg/L, or English (°e) or French degrees (°fH). Unilever prefer to report measurements in French hardness so for consistency hardness of water will be reported as both concentration (mg/L) and degrees of French hardness (°fH).

The addition of calcium demonstrated an increased catalytic activity up to 141% for a similar lipase (LN) from *T. lanuginosus*, whereas monovalent and divalent ions such as magnesium, lithium, sodium and barium ions had no effect (Zheng *et al*., 2011).

Calcium ions affect the ionization states and surface charge of ionic and zwitterionic surfactant micelles; for anionic surfactants, e.g. LAS, the stronger binding of Ca^{2+} leads to a greater impact on the self-assembly of mixed micelles with LAS and rhamnolipids. The strength of interaction is surfactant dependent, as for the less strongly anionic rhamnolipid, Ca^{2+} binds much weakly (Chen *et al*., 2013). Multivalent ions such as calcium are known to bind strongly to anionic surfactants at low ionic strength, inducing micellar growth of anionic surfactants such as SLES and multilayer formation at air-water interfaces; the latter may play a significant role in lipase activity and interfacial activation (Petkov *et al*., 2010). For non-ionic surfactants, Ca^{2+}, as would be predicted, has little effect on their adsorption and self-assembly even at high pH (Chen *et al*., 2013), and can supress micelle formation of non-ionic surfactants. Ca^{2+} ions mediate stable salt bridges between neighbouring headgroups of ionic and anionic surfactants, these micelle aggregates are more stable, compact and structurally rigid than in Na^{+} environments (Sammalkorpi *et al*., 2009). It may be that the more ordered and rigid micelles formed in the presence of calcium has an effect on the interaction with lipase; and more compact micelles may have a positive or negative effect on interfacial activation and catalytic activity.

Table 1. Water hardness levels assessed by Unilever

	Very Soft	Soft	Moderately Hard	Hard	Very Hard
Concentration range	< 30 mg/L < 15 mg/L	30 – 50 mg/L 15 – 25 mg/L	50 – 100 mg/L 25 – 50 mg/L	100 – 150 mg/L 50 – 75 mg/L	> 30 mg/L > 15 mg/L
French Hardness (°fH) range	< 7.50	7.50 – 12.50	12.50 – 25.00	25.00 – 37.50	> 37.50
Actual concentration used	20 mg/L 10 mg/L		75 mg/L 37.5 mg/L		200 mg/L 100 mg/L
Actual French Hardness (°fH) used	5.00		18.75		50.00

Table reports the class of 5 groups of increasing water hardness used in this thesis ranging from very soft to very hard with increasing level of calcium and magnesium cations. These

values have been taken from our collaborators Drs D. Lang, M. Noro and Ms. I. Sambi from Unilever. Values for magnesium concentrations have been calculated from calcium levels since the ratio between Ca^{2+}:Mg^{2+} is generally regarded to be equal to 2:1. French hardness was calculated as x0.1 of the concentration of $CaCO_3$ (mg/L) present.

Calcium chloride is known to activate hydrolysis of all triglyceride substrates of pancreatic lipase (Schønheyder & Volqvartz, 1945) and other mesophilic lipases (Benzonana, 1968) by forming calcium soaps which are formed from calcium binding to free fatty acid groups liberated after hydrolysis. The mechanism behind this is still not fully understood. Mutants of TlL that were lid variants showed lower activity in calcium rich buffers on pNp decanoate and pNp butyrate (Skjold-Jørgenesen *et al*., 2014), demonstrating the pivotal role of the lid domain in governing calcium dependency. It is also known that cutinase-like enzymes that share common structural elements and catalytic domains to lipases undergo calcium induced activation and conformational changes (Miyakawa *et al*., 2015). Calcium and divalent cations also play a role in the activity of phospholipases (Berg *et al*., 2001), however their role is critical to catalytic activity and reaction progress. It is unknown whether calcium binding elicits a conformational change or affects the allosteric properties of the enzyme.

Stability of the thermostabile lipase from *Bacillus stearothermophilus* is increased by up to 10 °C upon calcium concentrations of 10 and 100 mM, postulated to be due to the binding of calcium and stabilizing effect on the tertiary structure as measured by tryptophan fluorescence emission (Kim *et al*., 2000).

2.15 Aims

Recent studies have shown surfactants can affect both positively and adversely the catalysis of lipases through a number of mechanisms: binding to the lipase, affecting lipase surface area or affecting the interfacial activation (Delorme *et al*., 2011). Under certain surfactant conditions the activity of Lipex is considerably increased (Unilever; personal communication); the effects of the surfactant and solution conditions on Lipex structure, conformation, dynamics and enzyme functionality and catalysis at a biomolecular level are however, unclear. Whilst the work previously performed and

presented has established an insight into the mechanisms of TlL, there is still much to be determined. These conditions are considered classical parameters in interfacial catalysis, however, the interfacial activation, conformational changes, interactions and characteristics of Lipex at a dynamical level are still not fully understood (Chahinian *et al*., 2006).

A range of biophysical techniques such as crystallography (Brzozowski *et al*., 2000, Lawson *et al*., 1994), fluorescence spectroscopy (Mogensen *et al*., 2005, Skjold-Jørgensen *et al*., 2016), electron paramagnetic resonance (Sonesson *et al*., 2006, Hedin *et al*., 2005), and molecular dynamic simulations (Lawson *et al*., 1994, Santini *et al*., 2009) have been used to explore many facets of the behaviour and functionality of TIL. Crystallographic data provides only static images and MD simulations are only guidelines. Whilst the field of intrinsic and extrinsic fluorescent probe labelling has provided much information on TlL and its interactions with lipid micelles (Hedin *et al*., 2002, Skjold Jørgensen *et al*., 2015, 2016) the incorporation of fluorescent probes, within the protein sequence, may interfere with the structure, activity and conformational behavioural properties of the protein. NMR offers a non-invasive technique to study these changes, at picosecond timescales and at 1 - 10 Å movements, and new FRET tryptophan quenching measurements can only analyse changes 7 – 15 Å (Mansoor *et al*., 2005).

NMR has the power to provide a novel investigation into the function of Lipex at a residual and dynamical level, in industrially relevant conditions. NMR based techniques will be employed to study these effects on enzyme catalysis, structure and dynamics. I will attempt to determine what dynamical and conformational changes occur in the properties of Lipex, which are caused by surfactants to infer on the functional role of lipase in the catalytic cycles and during interfacial activation. To my knowledge, there are no reported studies of TlL using multidimensional NMR, so these insights will provide a novel exploration of its activity and function. The use of NMR allows assessment of individual amino acids during the dynamical process of catalytic mechanism and interfacial activation; a factor which is very difficult to achieve alluded to in relation to site directed mutagenesis of Glu87Ala and Trp89Phe mutations (Holmquist *et al*., 1995). Simulations run with TlL have shown that the transition state conformation experienced during interfacial activation was highly

unstable and sampled a number of configurations; illustrating the difficulty in studying transient dynamic conformations, a factor that may be addressed using NMR based methods.

NMR methods involving 2D HSQC and dynamic measurements will allow insights into the conformational changes of Lipex at the residual level upon surfactant and calcium addition. Lipolytic activity assays can be performed in parallel, to ascertain if specific changes in the structure is related to activity.

The observations from these studies could potentially lead to improved design of bio-industrial lipase enzymes. The affect of factors such as surfactants, solubility, water hardness, pH and temperature on Lipex structure, stability catalysis, dynamics and function will be studied and could also help for design or improvement of *de-novo* enzymes. Specifically, the aims of this project are to determine and understand the interactions between industrially selected surfactants and Lipex, under different solution conditions (of industrial relevance) and the impact of activity, at a residual level using NMR techniques. The preliminary aims of the project will be to accomplish:

- Expression and purification of isotopically labelled Lipex for NMR studies
- NMR characterisation and optimisation of samples
- Backbone ^{1}H-^{15}N HSQC assignment
- Identification of solution conditions optimal to NMR studies that remain industrially relevant and maintain lipolytic activity.
- Analysis of NMR 2D spectral changes mapped to protein structure under addition of surfactants and different solution conditions (i.e. water hardness)
- Dynamical NMR studies to elucidate regions of dynamic sensitivity and significance
- Determine effects of surfactants on enzyme structure and stability using NMR methods

CHAPTER THREE

LIPEX

METHODS AND MATERIALS

Chemical suppliers

All chemicals and substances were purchased from Sigma Aldrich unless otherwise stated.

3.1 Yeast

3.1.1 Constructs

The gene for Lipex was codon optimised for expression in yeast and inserted into the pPICZα_A vector using EcoRI and NotI restriction enzyme sites in frame with the C-terminal peptide containing the *c-myc* epitope and the N′ terminal polyhistidine tag and synthesised by Invitrogen, Gene Art. The pPICZα vector contains the *Sh ble* gene from *Streptoalloteichus hindustanus* conferring antibiotic resistance to Zeocin^{TM}. The pPICZα vector contains the AOX1 (alcohol oxidase enzyme) gene, which is tightly regulated and induced under addition of methanol. The vector pPICZα also contains a α-factor which is a prepro secretion signal sequence peptide taken from *Saccharomyces Cerevisiae* which is exported out of the cell into the culture media. It is cleaved from the gene of interest by the KEX2 enzyme in the yeast membrane.

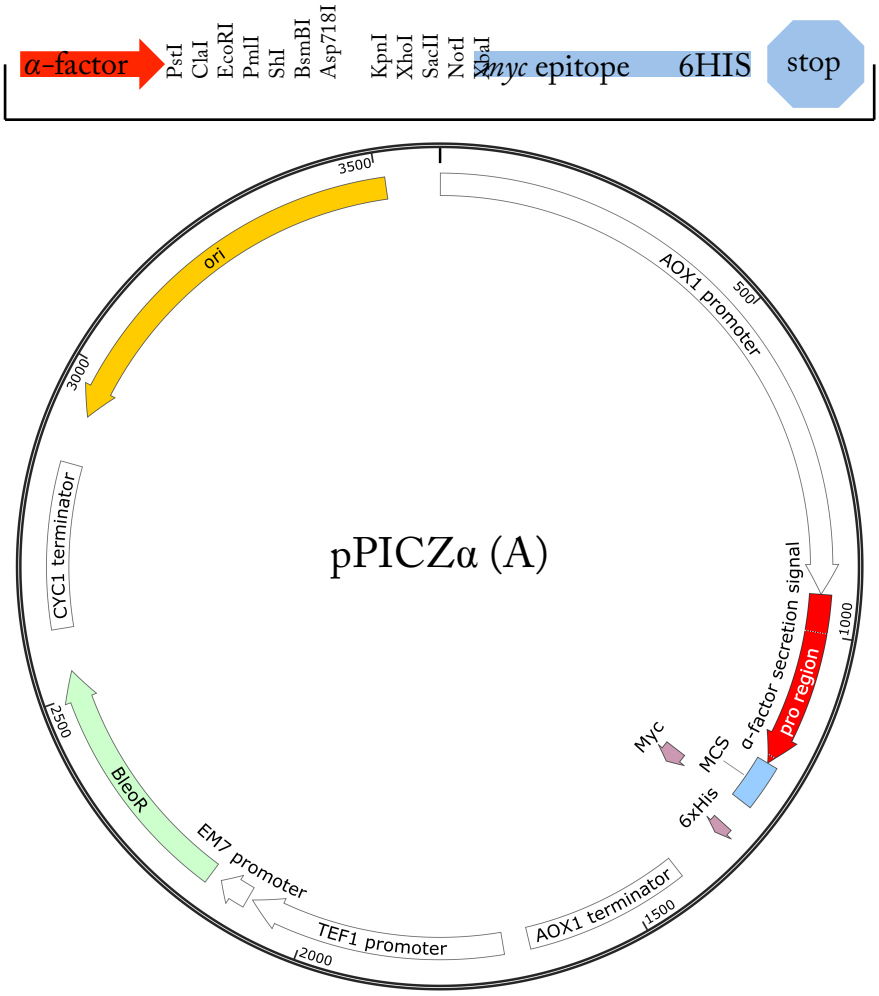


Figure 3.1 pPICZα vector map

pPICZα vector map illustrating restriction enzyme sites and major features of the vector. Multiple cloning site is shown in detail. Figure made by Snapgene

3.1.2 Strains

Cloning and DNA amplification was performed in Top10 *E. coli* cells (Agilent), which is endonuclease (*endA*) and recombination (*recA*) deficient increasing quality of mini-prep DNA and reducing plasmid multimer formation; furthermore Top10 confers resistance to Zeocin™.

Expression was first trialled in X-33 *P. pastoris* cells (*Mut*$^+$ *His*$^+$) due to its common usage and wild-type genotype. All colonies produced are expected to *Mut*$^+$ genotype and does not need supplementation with histidine.

Following issues with protein degradation SMD1168 was selected, it is a protease deficient strain that has *pep4* genotype, and is deficient of the proteinase A, a vacuolar aspartyl protease which activates other vacuolar proteases such as carboxypeptidase Y and proteinase B. SMD1168 is also Mut^+ His^+.

3.1.3 Transformation of E. coli cells with Lipex_pPICZα vector

The plasmid containing the pPICZα vector and Lipex gene (Lipex_pPICZα) was synthesized by Invitrogen. 5 µl of plasmid DNA *Lipex_pPICZα* at 100 ng/µl concentration was combined with 20 µl of Top10 cells and left on ice for 30 minutes (min) with periodic gentle agitation, after which it was subjected to 42 °C for 45 s and returned on ice for a further 10 min. Next, 500 µl of pre-warmed LB (10 g/L tryptone, 5 g/L yeast extract, 10 g/L sodium chloride, pH 7.5) was added and the cells were left shaking at 200 rpm at 37 °C for 1 h; 150 µl of which was plated under sterile conditions on LB agar plates (10 g/L tryptone, 5 g/L yeast extract, 10 g/L sodium chloride, 15 g/L agar, pH 7.5) against 25 mg/ml Zeocin™ antibiotic resistance. Plates were incubated at 37 °C overnight.

3.1.4 Lipex_pPICZα_A vector linearization

Transformants of TOP10 containing Lipex_pPICZα vector were selected and grown in 10 mls LB containing 25 µg/ml Zeocin™ for 8 h at 37 °C, 1 ml of this was used to inoculate 100 mls LB containing 25 µg/ml Zeocin™ and grown overnight at 37 °C. Midi-prep (Qiagin) kits were used to extract the DNA following manufacturer instructions.

DNA containing Lipex_pPICZα_A vector was linearized using Sac1 restriction enzyme and NEBuffer 1 (Invitrogen). Reaction was carried out overnight at 37 °C and stopped by heat inactivation at 65 °C for 30 min.

3.1.5 Ethanol precipitation of plasmid DNA

3 M sodium acetate at pH 5.2 was added to the DNA at a volume of 10% of sample volume, along with 2.5 volumes of ice-cold ethanol and the sample was left at -80 °C

for 30 min. The sample was centrifuged to pellet the DNA at 17,000 *g* for 25 min at 4 °C, the supernatant was decanted and pellet was air dried under N_2 stream. The pellet was washed in 1 ml of ice-cold 70% ethanol, centrifuged again under same conditions for 5 min, the supernatant was removed and the pellet dried again. Finally the DNA was re-suspended in a low volume of (~30 µl) H_2O, a low volume does not dilute cells for electroporation increasing transformation efficiency.

3.1.6 Electroporation of P. pastoris cells

P. pastoris cells (X-33, SMD1168) (Invitrogen) were streaked from a glycerol stab onto YPD agar (20 g/L tryptone, 10 g/L yeast extract, 20 g/L agar and 20 g/L glucose (added separately after autoclaving media as a solution of 50% (w/v))) and grown for 24 h at 28 °C. A single colony of X-33 cells was added to 5 ml YPD media (20 g/L tryptone, 10 g/L yeast extract, 20 g/L glucose) (Sherman, 2002) and grown for 16 h at 28 °C. Next 90 µl of the pre-culture was used to inoculate 500 ml YPD and grown 16 h at 28 °C to an OD_{600} of 1.0. The cells were then gently centrifuged at 1,000 *g* for 5 min at 4 °C and re-suspended in 500 ml sterile ice-cold water. This wash process was repeated three times and the pelleted cells were re-suspended in 250 ml sterile ice-cold water, 20 ml 1 M sorbitol and finally 1 ml 1 M sorbitol to total volume of approximately 1.5 ml. Fresh competent cells were made for each electroporation.

Lipex_pPICZα vector was prepared from sections 3.1.3 – 3.1.5, 10 µl of 1 mg/µl Lipex_pPICZα vector was added to 80 µl of *P. pastoris* cells in a 0.2 cm gap electroporation cuvette (Biorad) and left on ice for 5 min. Electroporation was performed as manufacturer instructions described for *P. pastoris* on a Bio-rad Gene Pulser II machine using 1.5 kV voltage, 400 ohms resistance, 25 mF capacitance and a time constant of 8.0 ms, 1 ml of 1 M sorbitol was added immediately after electroporation. The cells were left at 28 °C in a non-shaking incubator for 1 h, before addition of 1 ml YPD and another hour incubation non-shaking at 28 °C. The cells were then plated on YPDS plates (20 g/L tryptone, 10 g/L yeast extract, 20 g/L glucose, 1 M sorbitol) containing 500 µg/ml Zeocin™ antibiotic. The plates were left at 28 °C for 2 - 5 days, until yeast colonies had formed.

3.2 Expression trials

3.2.1 Expression trials in Rich Buffered Glycerol Complex Media (BMGY), Buffered Methanol Complex Media (BMMY) and Basal Salt Medium

Multiple colonies of X-33 cells containing Lipex_pPICZα plasmid were inoculated in a 25 ml pre-culture of BMGY (10 g/L yeast extract, 20 g/L tryptone, 34 g/L yeast nitrogen base (YNB) without amino acids and ammonium sulphate, 10 g/L ammonium sulphate, 10 g/L glycerol, 21.2 g/L potassium phosphate, pH 6.0) at 28 °C for 16 h to an OD_{600} of ~6. No form of antibiotic was used. The culture was centrifuged at 2,000 *g*, 15 min, 4 °C and the pellet re-suspended in 100 ml BMMY (10 g/L yeast extract, 20 g/L tryptone, 34 g/L YNB without amino acids and ammonium sulphate, 10 g/L ammonium sulphate, 10 ml/L methanol, 21.2 g/L potassium phosphate, pH 6.0) to an OD_{600} of 1.0. The switch between glycerol in BMGY and methanol in BMMY culture media is what induces expression of the *aox1* promoter in the pPICZα vector, inducing expression of Lipex. Cultures were grown at 28 °C and culture aliquots of expression were taken at specified time-points, flash frozen immediately in liquid nitrogen and stored at –80 °C until analysed further. Lipex is expected to be secreted to the media therefore samples of both cells and culture were taken for analysis.

Expression trials in basal salt were performed in parallel. X-33 colonies containing the Lipex_pPICZα plasmid were grown as for BMGY cultures, after reaching an OD_{600} of ~6 cells were centrifuged at 2,000 *g*, 15 min, 4 °C and the pellet re-suspended in 100 ml basal salt medium medium (26.7 ml 85% H_3PO_4, 0.93 g $CaSO_4.2H_2O$, 18.2 g K_2SO_4, 14.9 g $MgSO_4.7H_2O$, 40.0 g Glycerol, pH 5.0 with 30% NH_4OH). Expression was induced by adding 20 ml/24 hrs 100% methanol.

3.2.2 Expression trials in Minimal Buffered Methanol Complex Media BMMY (mBMMY)

Expression of isotopically labelled proteins needs to be performed in media where the carbon and nitrogen sources can be selectively chosen. Minimal BMMY allows this for culturing yeast strains such as X-33 and SMD1168H.

Cultures were set up as described for rich BMMY expression in section 3.2.1 with minimal BMGY (34 g/L YNB without ammonium sulphate, 10 g/L ammonium sulphate, 10 g/L glycerol, 100 mM potassium phosphate, pH 6.0) and minimal BMMY (34 g/L YNB without ammonium sulphate, 10 g/L ammonium sulphate, 10 ml/L methanol, 21.2 g/L potassium phosphate, pH 6.0). ^{15}N labelled attempts used ^{15}N labelled ammonium sulphate in minimal BMMY media.

3.2.3 Expression trials in FM-22

An alternative culture media FM-22 was also trialed for optimising expression and reproducibility. Multiple colonies of X-33 cells containing Lipex_pPICZα vector were inoculated in a 25 ml pre-culture of YPD containing 25 μg/ml Zeocin at 28 °C for 16 h. The culture was centrifuged at 2,000 *g*, 5 min, 4 °C and the pellet re-suspended in 1 L of FM-22 growth media (42.9 g/L potassium phosphate monobasic, 11.7 g/L magnesium sulphate heptahyrdate (Duchefa Biochemie), 40.0 g/L glycerol, 14.9 g/L potassium sulphate, 1.25 g/L calcium sulphate dihydrate, 5.0 g/L ammonium chloride (VWR), 400 μg/L biotin). This culture was grown at 28 °C during the growth phase till an OD_{600} of 2 - 6 was reached where it was pelleted at 2,000 *g*, 15 min, 4 °C and re-suspended in 200 ml fresh FM-22 media. The culture was left at 28 °C for 1 h to allow cells to acclimatize to new environment before expression was induced with methanol. Methanol was introduced slowly at 5 ml/h for 4 h to allow cells to adjust to expression environment. Induction was continued using 20 mls methanol was added every 24 h afterwards. Culture aliquots of were taken at specified time-points, flash frozen immediately in liquid nitrogen and stored at – 80 °C until analysed further.

Variations of FM-22 in increased 1.4 × original (60.0 g/L potassium phosphate monobasic) or reduced 0.6 × original (25.74 g/L potassium phosphate monobasic) potassium phosphate amounts were used in attempt to reduce proteolysis. No other changes to the expression protocol were made.

3.3 Western blot

To check for expression of proteins, 4 ml aliquots of the expression trial cultures were collected at specified time points; the sample was centrifuged at 3750 *g* for 5 min at

room temperature to pellet the cells. The supernatant was removed and concentrated to 0.5 ml on a 10 kDa MW cut-off concentrator, 20 µl of protein sample was added to 5 µl of 5× SDS loading buffer and analysed for extracellular secreted protein expression.

The pellet was saved and intracellular expression was analysed. The pellet was re-suspended in 50 mM sodium phosphate, 25 mM NaCl, pH 8.0 and sonicated on a Sonoplus HD 2070 sonicator using a KE76 probe (Sonoplus, Bandelin) for 1 min at 50% amplitude. The sample was then centrifuged at 17,000 *g* for 5 min at room temperature to pellet the cell debris. The supernatant was checked for soluble cytoplasmic protein expression. Gels were run using 20 µl of protein sample was added to 5 µl of 5× SDS loading buffer.

Samples were run on a 12% SDS-PAGE gel using a 2–212 kDa broad range pre-stained marker (New England Biolabs) at 180 V for 45 min. The bands were transferred to a nitrocellulose 0.2 µm pore size membrane (Invitrogen) using a Xcell II Blot Module (Invitrogen) in a buffer of 12 mM Tris-HCl, 96 mM Glycine, 20% methanol, pH 7.4 at 35 V for 180 min. The membrane was blocked in 50 ml phosphate buffered saline (PBS) (137 mM sodium chloride, 2.7 mM potassium chloride, 10 mM sodium phosphate monobasic, 1.8 mM potassium phosphate monobasic, pH 7.4) with 5 g skim milk powder for 60 min at room temperature, then incubated with 1:4000 ratio anti-polyhistidine mAb antibody (Sigma) in PBS with 0.5% (v/v) Tween-20 (PBST) for 1 h. The membrane was washed 3 times for 5 min in PBST, and SuperSignal West Pico Chemiluminescent substrate (ThermoScientific) was added following manufacturer instructions and incubated for 5 min before being viewed on a Fujifilm LAS 3000 using a 30 s exposure time.

3.4 E. coli

3.4.1 Construct design

The Lipex gene containing an N′ terminus pel B periplasmic secretion signal peptide was codon optimised for expression in *E. coli* and cloned into pET21a vector using Xho1 and Xba1 restriction enzymes.

Pel B signal peptide and construct was selected after discussion with Steve Wilson (Unilever, Colsworth). The ligated vector was transformed into SoluBL21 (DE3) *E. coli* cells and grown overnight on LB agar plates containing 100 ug/ml carbenicillin for antibiotic resistance selection.

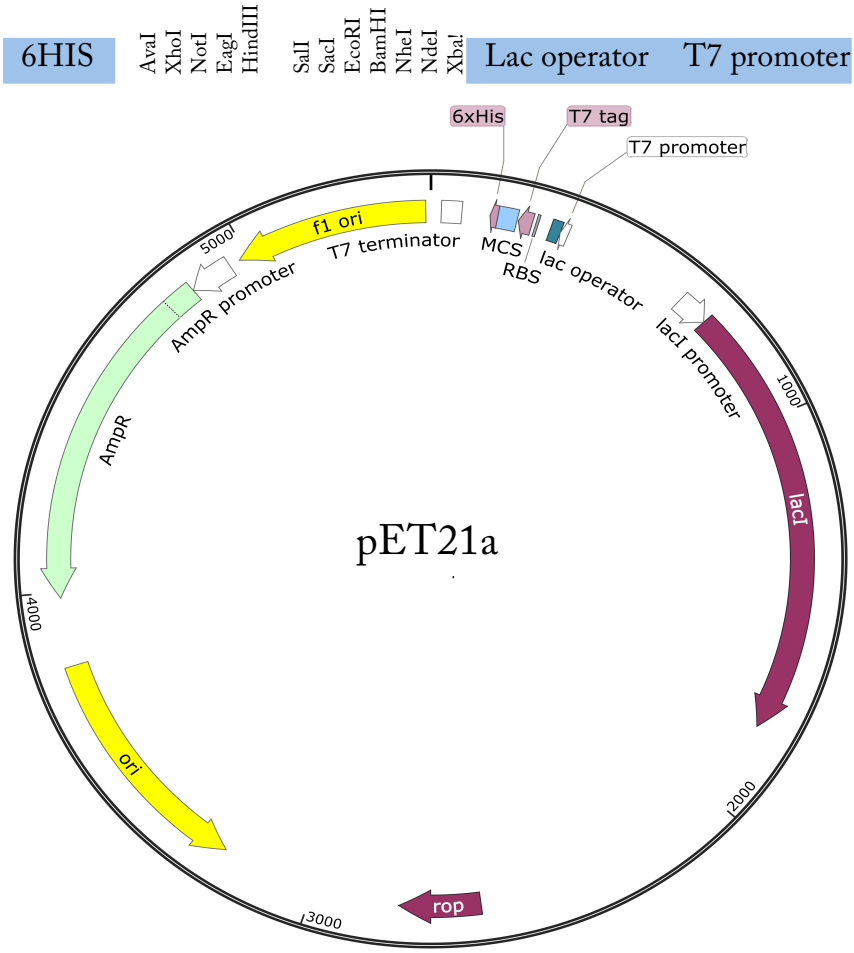


Figure 3.2. pET21a vector map

pET21a vector map illustrating restriction enzyme sites and major features of the vector. Figure made by Snapgene

[F]**ATG**AAATACCTGCTGCCGACCGCTGCTGCTGGCCTGCTGCTGCTGGCTGC
CCAACCGGCGATGGCG

N terminus pelB periplasmic signal peptide sequence, F refers to forward frame of DNA

3.4.2 Strain

SoluBL21(DE3) (Genlantis) encodes T7 RNA polymerase for high level transcription, and is deficient in *ompT* protease resulting in higher expression levels of intact recombinant proteins. It also has tighter control of methylated DNA with *hsd SB* (r_B- m_B-) and *dcm* genotypes. There are more mutations unspecified by the manufacturer present within the strain that help it to express insoluble proteins in a soluble form.

3.4.3 Lipex_pET21a vector

The gene for Lipex was codon optimised for expression in *E. coli* and synthesized by Invitrogen Gene Art with Xho1 and Xba1 restriction enzyme sites at the N′ and C′ termini sites of the gene respectively. A pMAT vector containing the Lipex gene was supplied by Invitrogen (pMAT_Lipex) and transformed into Top10 cells as described in section 3.1.3 for DNA amplification by miniprep.

5 µg of pMAT_Lipex DNA was cut with Xho1 and Xba1 restriction enzymes and NEBuffer 4 under a double restriction enzyme digestion reaction at 37 °C for 48 h, the reaction was stopped by heat inactivation at 50 °C for 20 min.

The pET21a vector was purchased from Invitrogen and transformed into Top10 cells as described in section 3.1.3 and 3.1.4 for DNA amplification by miniprep. The vector underwent the same double digestion reaction as for the insert with Xho1 and Xba1 restriction enzymes.

The DNA was separated on a 1% agarose gel run at 100 V till all bands were separated. For a more efficient ligation the gel was run until the difference between double digested pET21a vector and uncut pET21a could be distinguished. The Lipex insert and pET21a double digested vector was excised from the DNA gel by a DNA extraction kit (Qiagen) following manufacturer instructions.

Xho1 and Xba1 digested pET21a and Lipex insert were mixed in a concentration ratio of 1:1 and 2 µl of ligation Master mix (Invitrogen) was added, the ligation was incubated for 30 min at 25 °C. The reaction was then transformed into DH5α (Invitrogen) cells using the protocol described in section 3.1.3 and plated on LB plates against antibiotic carbenicillin used at 100 µg/ml. A single colony was selected and grown overnight in LB with carbenicillin at 100 µg/ml at 37 °C for miniprep. Successful cloning of the Lipex_pET21a vector was confirmed by DNA sequencing (Source Bioscience) using T7 forward and reverse promoters supplied by Source Bioscience.

SoluBL21(DE3) cells were transformed with the Lipex_pET21a vector as described in section 3.1.3. Transformed colonies of SoluBL21(DE3) cells containing the Lipex_pET21a vector were selected on LB plates against carbenicillin antibiotic selection at 100 µg/ml.

3.5 Expression of Lipex in E. coli

3.5.1 Expression of unlabeled Lipex in E. coli

A single colony was selected and cultured in 100 ml LB media at 37 °C overnight with 100 µg/ml carbenicillin. The following day 10 mls of this pre-culture was used to inoculate 1 litre (L) of LB media, the culture was grown at 25 °C and induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.5 mM when the OD_{600} reached 0.6. Protein expression was carried out at 25 °C overnight. The cells were harvested by centrifugation at 3750 *g* for 15 min at 4 °C. The supernatant was removed and the cells pellet were stored at -20 °C.

3.5.2 Expression of isotopically labelled ^{15}N and or ^{13}C Lipex in E. coli

A single colony was selected and cultured in 100 ml M9 minimal media (6 g/L sodium phosphate dibasic, 3 g/L potassium phosphate monobasic and 0.5 g/L sodium chloride, pH 7.4, the following compounds were added after being filtered using 2 μm filter (Sartorious): 1 ml 1M magnesium sulphate, 1 ml 40 mg/ml thiamine, 1 ml 50 mg/ml carbenicillin, 200 μl 1 M $CaCl_2$ and 10 ml trace elements containing 3 $x10^{-6}$ M ammonium molybdate, 4 x 10^{-4} M boric acid, 3 x 10^{-5} M cobalt chloride, 1 x 10^{-5} M copper sulphate, 8 $x10^{-5}$ M manganese chloride, 1 x 10^{-5} M zinc sulphate, 3 g/L glucose and 0.7 g/L ammonium chloride) at 37 °C overnight. For isotopically labelled cultures 3 g/L ^{13}C labelled glucose (Cambridge Isotope Laboratories) and 0.7 g/L ^{15}N labelled ammonium chloride were used instead of ^{12}C glucose and ^{14}N ammonium chloride. The following day 30 mls of this pre-culture was used to inoculate 1 L of M9 minimal media, the culture was grown at 25 °C and induced with IPTG at a final concentration of 0.5 mM when the OD_{600} reached 0.6. Protein expression was carried out at 25 °C overnight. The cells were harvested as for unlabeled preparation.

3.6 Purification of Lipex expressed in E. coli

With the pelB signal peptide, Lipex is secreted and to the periplasm. A protocol for periplasmic purification was followed by lysing the periplasm by using a sucrose gradient inducing osmotic stress in the cell membrane.

Cell pellets containing Lipex were thawed at 37 °C for 20 min, and re-suspended in 25% sucrose, 30 mM Tris-HCl, 1 mM EDTA, pH 8.0 and left on ice for 20 min to increase osmotic gradient across the cell membrane. The re-suspended media was centrifuged at 8,000 *g* for 20 min at 10 °C. The supernatant was decanted and the cell pellet was re-suspended in ice-cold H_2O containing 5 mM magnesium chloride, 5 mM sodium chloride and left on ice for 20 min, disrupting the membrane and releasing proteins within the periplasm. A final centrifugation step at 10,000 *g* for 25 min at 4 °C was applied to pellet the cells, the supernatant was decanted, filtered on 0.2 μm membrane (Sartorious), brought to pH 9.0 and kept on ice.

A high-trap Q FF ion exchange column (GE healthcare) was equilibrated in 5 mM sodium chloride, 25 mM Tris-HCl, pH 9.0 (equilibration buffer). Periplasmic supernatant was loaded on the column at a flow rate of 2 ml/min. The column was washed in equilibration buffer. Lipex was eluted in a gradient of 100 steps to 100% 1 M sodium chloride, 25 mM Tris-HCl, pH 9.0 at a flow rate of 4 ml/min. Fractions containing Lipex were pooled and concentrated down to 4 mls and incubated at 55 °C for 1 h. The heat step precipitated any contaminating protein in the sample, which was removed by centrifugation at 17, 000 *g* for 10 min at room temperature, the supernatant was removed and loaded onto a S75 size exclusion chromatography gel filtration column (GE healthcare) equilibrated in 25 mM sodium chloride, 25 mM Tris-HCl, pH 8.0. Lipex was eluted at a flow rate of 1 ml/min at a t_R of ~90 min. A ÄKTA FPLC machine (GE healthcare) was used for both ion exchange and SEC-GF chromatographic steps.

3.7 Natural Source Lipex

Unlabelled Lipex samples expressed by Novozyme® (donated by Dr. Ilaria Sambi, Unilever; Port Sunlight) was purified by SEC gel filtration on a Superdex S75 column (GE healthcare) equilibrated with 25 mM Tris-HCl, 50 mM NaCl, pH 8.0. Gel filtration was performed on a ÄKTA FPLC machine (GE healthcare). A pH of 8.0 was used to avoid the pI of 5.4; at this pH the protein is negatively charged and less likely to aggregate. Lipex was eluted at a flow rate of 1.0 ml min/ml. To further improve purification, specifically removal of lipid and components from the storage buffer used by Novozyme® eluted fractions were colluded and heated at 55 °C for 1 h before gel filtration was repeated and the eluted sample was heated at 55 °C for 1 h and filtered on a 0.2 μm membrane (Sartorious) in attempt to remove residual lipid and proteins from the Novozyme® enzyme storage buffer formulation.

3.8 Activity Assay

The protocol for measuring lipase activity assay was donated by Ms. Ilaria Sambi (Unilever; Port Sunlight). The substrate was made with at a 8 mM stock solutions of pNp 4-nitrophenyl (valerate or dodecanoate) in 10 ml methanol and diluted to 1 mM in water at pH 4.5. Due to the instability of the substrate these solutions were prepared

fresh daily. Lipase was diluted to a range of 1 – 20 ng/ml in 50 mM Tris-HCl pH 8.5. Measurements were recorded in 96 well plate black with clear bottom plates (Grenier). In each well 20 µl of enzyme was added to 100 µl of Tris-HCl, 60 µl H_2O, and 20 µl of substrate at 1 mM. For samples containing surfactant, the surfactant mixes were diluted 10 fold in 25 mM Tris-HCl pH 8.5 and used in place of the 60 µl of H_2O. The measurements were recorded and absorbance values were measured at wavelength 405 nm on a Spectramax M2e fluorescence plate reader. End point values were taken after 15 min. For kinetic measurements, data points were recorded at 15 s intervals over 1 h. All samples were run in triplicate and for each sample, controls and blanks where 25 mM Tris-HCl pH 8.5 was used to replace enzyme, substrate and surfactant. Different water hardness conditions were established by using magnesium chloride and calcium chloride at different concentrations: completely soft (Tris-HCl), very soft (20 mg/L $CaCl_2$, 10 mg/L $MgCl_2$), moderately hard (75 mg/L $CaCl_2$, 37.5 mg/L $MgCl_2$) and very hard (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$).

3.9 Protein Concentration Estimation

The concentration of purified protein was estimated using a NanoDrop 1000 device (NanoDrop). The molecular weights and extinction coefficients were calculated by the online prediction tool Protparam (http://web.expasy.org/protparam/) according to the method of Gill and von Hippel (Gill & Vonhippel, 1989). This protocol was used for all protein samples worked on within this thesis.

3.10 NMR

3.10.1 Solution NMR

All NMR was performed with 10% (v/v) ^2H_2O for signal lock to deuterium. 5 mm tubes (Norrell) were used as standard with a sample volume of 500 µl. 5 mm microscale tubes (Shigemi) have a minimum sample volume of 300 µl were used where sample volume was limited.

Solution NMR for Lipex was performed on a triple resonance (^1H, ^{15}N, ^{13}C) 5.0 mm probe on a 800 MHz Bruker Avance spectrometer.

All multidimensional spectra were processed using NMRpipe (Delagio *et al*., 1995).

3.10.2 1D ^{1}H NMR

1D ^{1}H NMR was performed with a Hahn-Echo pulse sequence (Hahn, 1950), with Shaka excitation sculpting (Shaka & Hwang, 1995) for water suppression. Spectra were acquired with 256 scans.

	^{1}H (F1)
SW	20.024
Time Domain	8192
Acquisition time (sec)	0.25
Pulse length (μs)	8.10
Acquisition mode	DQD

3.10.3 2D ^{15}N-^{1}H TROSY-HSQC

2D ^{15}N-^{1}H TROSY-HSQC experiments were performed with a 2D [^{15}N-^{1}H] TROSY pulse sequence (Nietlispach, 2005). Spectra were centred at 4.74 and 118.60 ppm in ^{1}H and ^{15}N dimensions and acquired with 48 scans.

	^{1}H (F2)	^{15}N (F1)
SW	13.94	34.00
Time Domain	2048	100
Acquisition time (sec)	0.09	0.03
Pulse length (μs)	8.10	
Acquisition mode	DQD	Echo-Antiecho

Surfactants were added to NMR samples as a percentage of the sample volume. To avoid enzyme dilution 10% surfactant stock solutions were made using 25 mM sodium chloride, 25 mM Tris-HCl, pH 8.0.

3.10.4 Relaxation experiments

^{15}N TROSY-HSQC T_2 relaxation experiments were performed on 800 MHz (Imperial College London) and 950 MHz (National Institute of Medical Research, Mill Hill) spectrometers. 2D TROSY-HSQC versions were performed with 48 scans. (Zhu *et al*., 2000).

Delay times of 0, 20, 40, 60, 80, 100 and 120 ms were used for T_2 measurements. The delay times were non-sequential to prevent any erroneous data collection. T_2 relaxation measurements performed on the 800 MHz spectrometer were performed as individual experiments with varying T_2 delay times. For the 950 MHz spectrometer measurements were run as a single interleaved 3D experiment, where the first increment was collected at all time delays before collecting the second increment at all time delays thus averaging any effects observed from sample aggregation or precipitation.

	^{1}H (F2)	^{15}N (F1)
SW	13.94	34.00
Time Domain	2048	100
Acquisition time (sec)	0.09	0.03
Pulse length (μs)	8.10	
Acquisition mode	DQD	Echo-Antiecho

Analysis was performed using either Sparky (Goddard & Kneller) or CCPN Analysis (Vranken *et al*., 2005).

CHAPTER FOUR

LIPEX:

RESULTS AND DISCUSSION

4.1 Chapter overview

This chapter entails the, ultimately unsuccessful, results from expression and purification of Lipex in *P. pastoris*. Proteolytic degradation and instability of the recombinant protein in its glycosylated form are presented, and efforts to overcome these difficulties and establish reproducibility in a yeast expression system are outlined. A number of expression strategies that potentially may alleviate the complications with degradation and instability are offered, however due to time constraints an alternative host expression system was sought. The success of Lipex expression in *E. coli* was optimised along with a purification protocol, and isotopically labelled material was produced enabling NMR studies.

NMR experiments were performed to study and characterise the interaction between Lipex and surfactants. The NMR based results are predominantly assessed by comparing changes to the ^{15}N-HSQC spectrum of the protein, noting changes in peak position and intensity, which can be insights into interactions; additionally T_2 CPMG experiments shed light on the dynamical properties of the enzyme. The NMR results are split into two sections, first detailing the NMR characterisation of Lipex, as well as preliminary studies into industrially relevant parameters such as temperature and calcium ion concentration, coupled with the activity of the enzyme in the presence of surfactants and changes in water hardness. The second section studies the structural aspects of the protein under selected conditions relevant to Unilever where the tertiary structure is destabilised and begins to unfold and collapse. Initial investigations to determine the contributing factors are shown.

4.2 Expression trials

4.2.1 Initial expression of Lipex in *Pichia pastoris*

A major portion of this work was directed towards expression and purification of Lipex. For NMR studies, an expression system needed to be designed to be reproducible, with large yields, and with the ability to be grown in ^{13}C and ^{15}N isotopically labelled culture media at a relatively low cost. Lipex used by Unilever is produced by a company Novozyme using fermentation in a yeast host and is therefore

glycosylated. The glycosylation occurs via a single site at residue Asn33 by *N*-acetylglucosamine using mannose sugars to give a monoglycosylated enzyme (Boel *et al*., 2002). Glycosylation does not appear to affect catalytic activity, but has a slight effect on adsorption rates (Peters *et al*., 2002). Glycosylation does not significantly affect functionality, but provides stability to the enzyme (Unilever; personal communication). In an effort to correlate our findings with Unilever's own research it was decided that expressing the enzyme in yeast would be beneficial.

Recombinant protein expression occurs in *P. pastoris* cells such as X-33 in the cytoplasm, and can be purified by lysing cells by standard cell lysis techniques such as french press or sonication. An alternative method for purifying expressed proteins is by transporting the protein to the supernatant. This is achieved by using a pPICZα vector, where an N terminus α-factor acts as a signal peptide to export the recombinant protein out of the cell into the supernatant. The α-factor is cleaved by a KEX-2 enzyme located in the outer membrane of the cell, resulting in only the full-length protein in the supernatant media. Secretion to the media is advantageous since there are fewer proteases in the supernatant than in the cytoplasm, reducing the chance of proteolytic degradation. Additionally, since fewer proteins are natively secreted, purification from the media rather than from the cytoplasm requires less techniques that reduces the overall yield loss during purification.

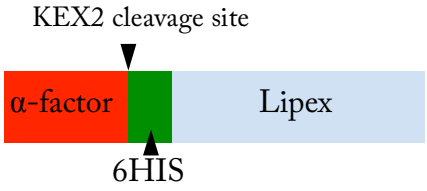


Figure 4.1 Diagram representation of Lipex construct for Lipex_pPICZa

Figure drawn starting from the N terminus

BMGY and BMMY were selected as they have have been a standard media for unlabelled proteins in yeast for decades. BMGY is used to grow the cells to a certain OD_{600} and the media is switch to BMMY where the AOX1 promoter and expression is induced by presence of methanol. SDS-PAGE gels and Western blots using

antibody to 6HIS tags were performed on culture media to confirm expression and secretion of a 6HIS-tagged protein after each expression trial.

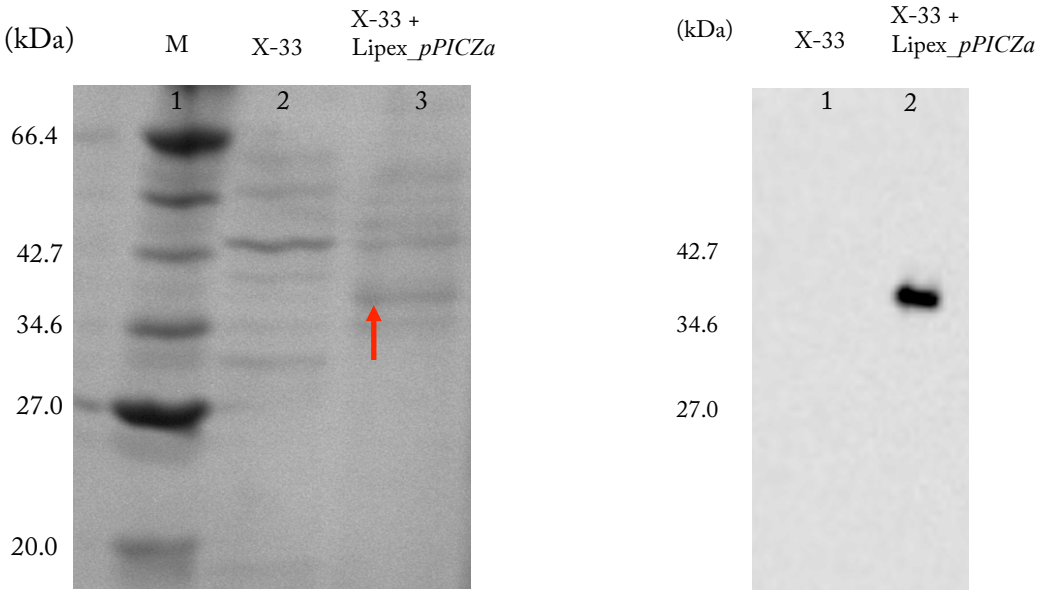


Figure 4.2 Expression and secretion of Lipex in X-33 cells containing Lipex_pPICZα in BMMY culture

(A) 12% SDS-PAGE and (B) western blot of culture media from negative control X-33 cells only and X-33 cells containing the Lipex_pPICZα vector.

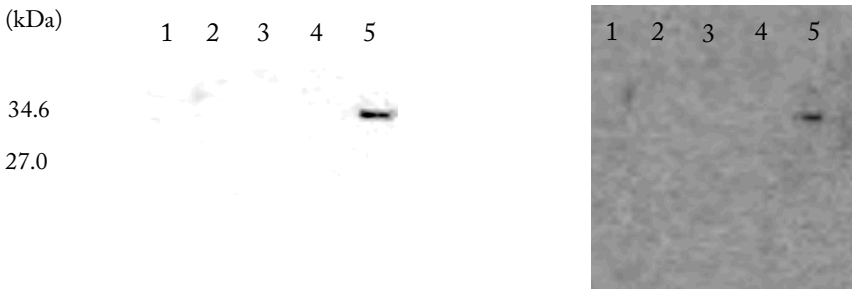
Cells were centrifuged at 4,500 rpm for 5 min and fractions from the supernatant were taken from a culture of BMMY cells at 24 h post induction with 1% methanol.

M denotes Marker in lane 1 (A). The red arrow in lane 3 (A) highlights a band at the predicted MW of mono-glycosylated Lipex at 37.2 kDa. This band is absent in the negative control confirming protein is result of Lipex_pPICZα vector.

Lipex was successfully expressed from *P. Pastoris* X-33 cells grown in rich BMGY/BMMY culture as confirmed by SDS-PAGE and western blot analysis (Figures 4.2 (A) and (B)). Lipex was expressed at 24 h post induction. *P. pastoris* X-33 cells natively secrete a protein with a MW of ~40 kDa, which is close to the 37 kDa MW size of 6HIS-tagged Lipex. Figure 4.2 (A) shows the secreted proteins present in *P. pastoris* X-33 cells with and without the Lipex_pPICZα plasmid. It is difficult to determine if Lipex is expressed on SDS-PAGE analysis alone, western blots were performed after each expression trial to validate findings.

The band present in Figure 4.2 (B) provides confidence that 6HIS-Lipex is expressed given the MW size, presence of a single band and detection using 6HIS antibody. This single band is present in samples that only contain the Lipex_pPICZa plasmid providing support that it is recombinantly expressed Lipex.

Successful expression of Lipex was observed, however it could not be reproduced. Numerous attempts of electroporation and colony formation failed to yield any colonies with correct morphology or any that expressed Lipex in rich BMGY/BMMY cultures as previous. Culture media was changed to FM-22 cultures, a medium popular for isotopic labelling for NMR studies, did not resolve the issue and no expression was noted. A number of conditions were trialled with FM-22 cultures: varying temperature, pH, growth phases and dextrose addition, however none showed success. Based on observations by Sinha *et al*., (2005) a range of culture media compositions that exhibited reduced protease activity was trialled. Two alternative media compositions basal salt medium and minimal BMGY/BMMY for isotopic labelled was trialled. Cells are grown as for rich cultures but the composition is altered to allow for isotopic labelling for ^{15}N and ^{13}C nuclei using ammonium sulphate and glucose sources as the sole presence of ^{15}N and ^{13}C nuclei.



(A) *(B)*

Figure 4.3 Media trial for expression of Lipex in X-33 cells

Western blot using 6HIS-antibody of secreted proteins of X-33 cells with Lipex_pPICZa grown in varying culture media. Samples taken at (A) 24 h and (B) 72 h after induction

1 = FM-22 high phosphate (1.4 × original) *4 = Basal Salt Medium*

2 = FM-22 low phosphate (0.6 × original) *5 = minimal BMGY/BMMY*

3 = FM-22 original

An expression trial was initiated varying several culture compositions, with only minimal BMGY/BMMY successfully expressing Lipex, confirmed by western blot (Figure 4.3). For minimal BMGY/BMMY expressed is noted at 72 h showing promising results for Lipex expression and stability. Based on these results minimal BMGY/BMMY was selected as culture media of choice. Unlabelled and ^{15}N labelled minimal BMGY/BMMY cultures were grown for expression of unlabelled and ^{15}N labelled Lipex to begin NMR studies.

4.2.2 Proteolytic degradation

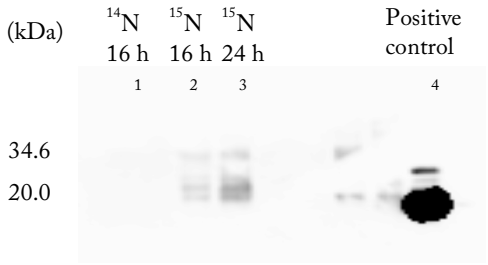


Figure 4.4 Proteolytic degradation of Lipex expressed in X-33 cells in minimal BMGY/BMMY media

Western blot using 6HIS-antibody of secreted proteins of X-33 cells with Lipex_pPICZa grown in ^{14}N unlabelled and ^{15}N labelled minimal BMGY/BMMY media. Samples taken at 16 h and 24 h after induction

Positive control (0.06 mg) is shown to confirm detection of 6HIS tagged proteins using 6HIS antibody.

Protease activity was detected at a rate previously not encountered. Two minimal BMGY/BMMY cultures were performed in parallel, one using unlabelled ^{14}N and the other labelled ^{15}N ammonium sulphate. The supernatant was collected at 16 and 24 h and a western blot was performed to check for the presence of 6HIS-Lipex. Lipex was expressed at ~10 fold less than the control,

The detection of proteolytic degradation was noted as a smeared band in the western blot (Figure 4.4). The smears detected after 24 h post induction in ^{15}N labelled BMMY culture in Figure 4.4 are indicative of proteolytic degradation. It is unknown where the cleavage occurs. The second observation is that Lipex is not expressed in

the unlabelled culture in lane 1 exemplifying poor reproducibility in protein expression. This identifies a problem in the media and expression of the protein.

The enzyme is reportedly very stable and can withstand up to 30 μg/ml trypsin digest at 37 °C (Wang *et al*., 2013, Unilever, personal communication). High kinetic stability is a characteristic required of industrial enzymes; it is surprising that the recombinant glycosylated Lipex is prone to such a degree of proteolysis in a yeast expression system. It is unreported to what degree Lipex undergoes proteolytic degredation when expressed by Novozyme in yeast fermenter production, it has been noted that TlL is highly prone to proteolytic activity, almost 20 fold higher compared to fungal lipases from *Rhizomucor miehei* and *Candida rugosa I* and *II* (Schmind & Verger, 1989).

The proteases responsible for degradation were not identified. Proteolytic degradation may have occurred from a number of different sources. Proteases in the media would originate from either being secreted by *P. pastoris* X-33 cells, or are of cytoplasmic origin and are present in the supernatant being released from lysed cells. *P. pastoris* natively secretes two forms of proteases, active at low (3) and neutral (5) pH levels (Brierley *et al*., 1994). This exemplifies the need for a strong buffering agent, usually potassium phosphate, within the culture media to maintain pH culture levels above pH 5. BMGY/BMMY cultures were grown at pH 6.0 and avoid both types of proteases. During expression the pH levels of BMGY/BMMY cultures did not drift beyond pH 5.5. Additionally, it has been shown that methanol as a carbon source leads to release and misdirection of vacuolar proteases into the culture medium, via a mechanism that is not yet fully understood (Sinha *et al*., 2005). Using methanol for induction of protein expression might have triggered the proteolysis. Findings by Sinha *et al*., (2003) showed that proteases were only detected after 48 h of methanol induction of *P. pastoris* X-33 cells, and not during glycerol growth phases where glycerol acts as the sole carbon source. This does not correlate to findings reported here where proteolytic activity was noted only 16 h after induction. Proteinase A, B, carboxylpeptidase Y and aminopeptidase are the four most prevalent vacuolar proteases. The first three have higher levels after methanol induction (Sinha *et al*., 2005). It is possible to have release of vacuolar proteases from the cytosol and eventually to the extracellular culture media (Stevens *et al*., 1986).

If not from natively secreted extracellular proteases, proteolytic degradation could indicate that the cells have lysed, possibly from shearing from incubator shaking which has been noted even at low speeds of 90 rpm (Taylor, 1989). Shearing of cells can lead to the release of intracellular proteases into the extracellular medium not normally found in the media. This could explain the cause of protein loss after 12 h. Clare *et al*., (1999) demonstrated that the addition of 5 g/L casamino acids inhibited extracellular proteases; these observations were also noted by Inan *et al*., (2005) where protease activity was reduced by 66% and Sinha *et al*., (2004) with 72% reduced protease activity. Unfortunately the introduction of casamino acids would affect the success of ^{15}N and ^{13}C isotopic labelling required for NMR measurements and labelled casamino acids are not commercially available, therefore this is not a viable option for NMR studies. It was reported that yeast cultures of TlL in *Aspergillus niger* cells expressed sulphydryl proteases, a likely candidate responsible for degradation, which was inhibited upon addition of p-chloromercuribenzoic acid (PCMB) (Shenolikar & Stevenson 1982, Ong & Gaucher, 1973). Taylor (1989) noted an increase in 50% yield, stabilization and activity of TlL when PCMB was added during purification. The addition of PCMB in combination with a cocktail of yeast proteases (containing AEBSF, E-64, pepstatin and 1,10-phenanthroline, for inhibition of serine, cysteine, aspartic and metalloproteases SIGMA code p8215) was not investigated since SMD1168H cultures gave best expression of Lipex without any additional protease inhibitors. The study performed by Sinha *et al*., (2004) was performed using a New Brunswick Scientific Edison GA 5L Bioflo 111/3000 fermenter tank where pH, dissolved oxygen and carbon levels can be accurately monitored and regulated as compared to performing cultures in 2.5 L baffled flasks which was the method of choice used within this thesis for cost and accessibility (Section 3.2.1). The excessive proteolysis observed could be attributed to a lack of precise culture media regulation. Since there were a number of different proteolytic sources the first step was to attempt expression in a protease deficient strain SMD1168H. Fermenter cultures would require a lot of time and high cost to establish optimal conditions. Eliminating the number of proteases present offered a more efficient and viable option for attaining successful and stable expression.

Cultures of minimal BMGY/BMMY were attempted with a new protease deficient strain SMD1168H. Three cultures were set up in parallel using X-33 cells with the addition of a yeast protease inhibitor cocktail (Sigma code p8215), SMD1168H cells only and SMD1168H cells with the yeast protease inhibitor cocktail in attempt to establish an optimal culture. Western blots require a lot of time and the transfer step to the nitrocellulose membrane can be inefficient and result in false negative results, an alternative method for detecting expression of implemented by performing a nickel purification on Ni^{2+}-NTA resin of the 6HIS-tagged Lipex from the media of cell cultures. A rapid purification protocol reduces the length of time proteolytic activity can occur.

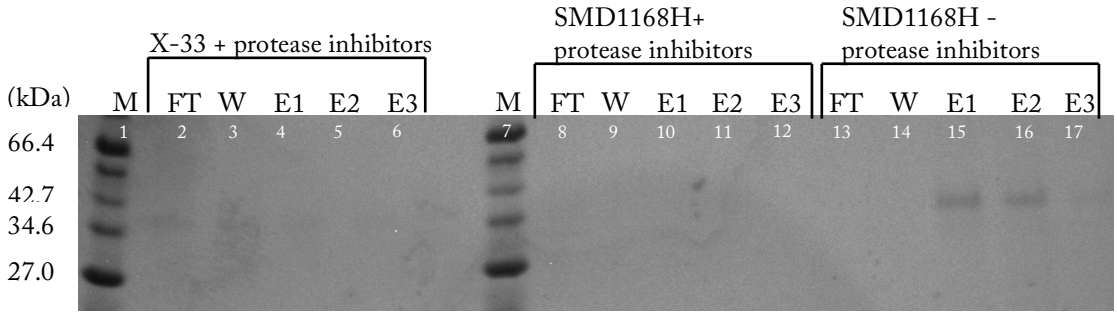


Figure 4.5 Expression trial of Lipex using combination of protease inhibition

Nickel purification of 6HIS-tagged Lipex expressed and secreted from in X-33 (lanes 2-6) or SMD1168H strains with (lanes 8-12) and without protease inhibitor cocktail (lanes 13-17) run on 12% SDS PAGE. All cell strains grown in minimal BMGY/BMMY cultures, samples taken at 48 h.

The flow through (FT), wash (W) and eluted fractions (E1, E2, E3) for each nickel purification are shown for each culture condition.

It is clear that the protease inhibitor cocktail is ineffective for minimal BMGY/BMMY cultures and, quite surprisingly, in combination with SMD1168H. Protein bands for correct Lipex MW are only noted in Figure 4.5 lanes 15 – 17, depicting that the optimal culture condition is SMD1168H without additional protease inhibitors.

Cultures were reproduced in SMD1168H cells without additional protease inhibitors to confirm reproducibility and establish a purification protocol. Cells were grown for

48 h to increase yields of protein expression. Figure 4.5 shows Lipex can be expressed in SMD1168H, and was stable up to 48 h.

Figure 4.5 shows that Lipex was successfully expressed and purified but the yield was too low to collect a 1D ^{1}H NMR spectrum. Despite this initial success and optimisation, numerous attempts failed to reproduce results in SMD1168H cells grown in either unlabelled or ^{15}N labelled minimal BMGY/BMMY. Given the extensive time spent working with *P. pastoris* it was felt that an alternative expression host should be sought for successful expression.

There are a number of factors which may explain why expression in yeast was ultimately unsuccessful; whilst preventative measures were taken at all times as best as possible provided by the working laboratory, the risk of contamination to yeast by *E. coli* sources could have been a major contributor to failed expression tests. Published data from a number of different sources expressing TlL in yeast use differing vectors such as pBoel960 and cell strains *A. niger* NW297 (Boel & Huge-Jensen, 1988, Prathumpai *et al*., 2004, Aloulou *et al*., 2007), a difference which may be crucial in expression. These vectors and cell strains are under industrial patents and are not accessible to academia. Modern expression strains and vectors were selected for assumed improved beneficiary aspects; secretion to the media by pPICZα was assumed to aid purification, since only a fraction of natively expressed proteins including proteases are also secreted aiding stability of recombinant expression; however the opposite effect of proteolysis was noted. The expression attempts and findings for Lipex in *P. pastoris* cells were ultimately halted, since reproducible expression was quickly established in *E. coli*.

4.2.3 Expression in *E. coli*

After *P. pastoris* expression encountered numerous problems the expression host *E. coli* was investigated. Bacterial hosts such as *E. coli* are often the preferred system for protein expression. They offer many advantages against yeast or mammalian expression systems such as cheaper reagents for cell culturing and chemicals for isotopic labelling, faster rates of colony growth and culture times and reduced risk of contamination. An industrial company, Novozyme, produces the Lipex used by

Unilever. Novozyme expresses Lipex using a yeast expression system. Whilst the polypeptide sequences are identical, Novozyme produces a glycosylated version. It is crucial to our investigations that the two protein variants are catalytically and structurally equivalent. Switching to a bacterial expression system will lose the glycosylation effects noted from yeast system. This was the only concern when deciding to trial *E. coli* expression. It has, however, been shown that glycosylation has no effect on the adsorption of TlL, or its orientation and secondary structure upon adsorption. The Gibbs free energy of unfolding is similar to non-glycosylated variants leading to an indication that the glycosylation plays no major role in activity, or interfacial activation of TlL (Pinholt *et al*., 2010). From these results we were confident the loss of glycosylation would not have a significant effect on enzyme functionality, activity or structure and proceeded with expression trials in *E. coli*. The activity of Lipex on 4C substrate valerate was shown to be consistent between non-glycosylated and glycosylated Lipex confirming our predictions (Figure 4.20(A)).

4.2.3.1 Strain screening and construct optimisation

Initially the Lipex gene was codon optimised for expression in *E. coli* and inserted into a pMAT vector as a standard vector for protein expression with an N terminal 6HIS tag for purification (GeneArt, Invitrogen). Expression was tested in a number of cell strains including BL21(DE3), SHuffle, M15[prep4], SHuffle T7 (DE3), C41 and Lemo21. Additional to cell strain, the concentration of IPTG used for induction (1.0, 0.5 and 0.1 mM), cell growth temperature (37, 30 and 25 °C) and expression temperature (30, 25 and 18 °C respectively) were varied without any successful soluble over-expression across any strain or culture condition.

The cell strain BL21(DE3) is commonly used as a standard across many labs for protein expression in *E. coli*. Since there are no published reports of expression of TlL in the literature initial expression of Lipex were performed using BL21(DE3) cells, however this cell strain failed to yield any Lipex expression. It is often the case that screening of a number of different strains and culture conditions leads to identifying successful candidates for recombinant protein expression. A number of alternative cell strains were screened, such as Lemo21 and C41, which are designed and optimised for expression of toxic proteins. One significant factor may be the presence

of 8 disulphide bonds within Lipex, BL21(DE3) cells are not adapted for expression of cytoplasmic proteins containing disulphide bonds so the BL21(DE3) derivative strain SHuffle T7 (DE3) was screened. SHuffle T7 (DE3) cells have mutations in the glutathione reductase and cytoplasmic thioredoxin genes that make the cytoplasm a non-reducing environment for better expression of disulphide bond containing proteins.

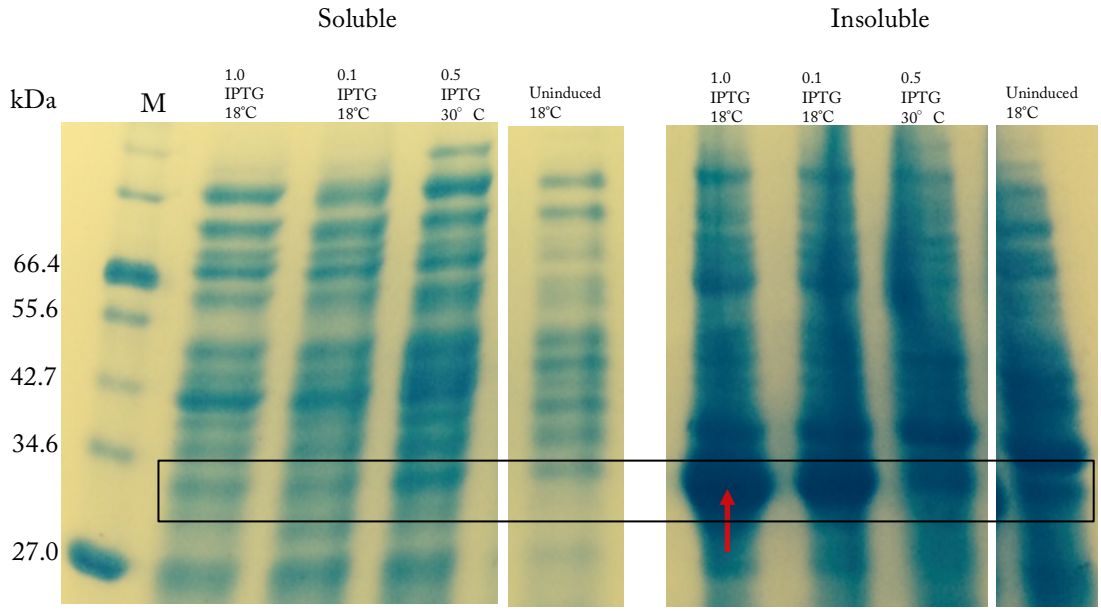


Figure 4.6 Lemo21 expresses Lipex in insoluble form

Lipex_pMAT vector grown in Lemo21 cells and induced with IPTG. An uninduced culture was grown alongside as a control. Cells were lysed by sonication and the soluble and insoluble fractions were run on a 12% SDS-PAGE gel. M denotes the Marker for protein band MW size.

It appears that Lipex might be expressed in a soluble form using Lemo21 cells, these levels are however very low and not substantial for the large quantities required for NMR studies. Lipex was however, over-expressed in vast amounts but in an insoluble form (Figure 4.6). The large band at ~30 kDa indicated in Figure 4.6 observed in the induced fractions is not present in the uninduced fractions and is therefore most likely to be Lipex which is non glycosylated and has a MW of 29.4 kDa.

Lipex expressed in an insoluble form cannot be used for any NMR or other biophysical studies we would like to perform. Insoluble proteins can potentially be

refolded under a number of different refolding techniques. Gogoy *et al*. successfully refolded a cysteine containing thermostable lipase from *Geobacillus thermocatenulatus.* After incubation in 8 M guanidine, 25 mM DTT the enzyme was immobilised onto a glyoxyl-agarose resin and refolded by removing guanidine by washing in aqueous buffer, retaining 95% enzyme activity (Gogoy *et al*., 2011). Despite over-expression of the protein, this method was not sought for Lipex. Any slight changes in activity, stability or tertiary conformation would compromise results.

Soluble expression in high yields of the protein is a fundamental requirement for solution NMR studies, so further strain screening was performed. Greater yields of soluble expression were achieved for Lipex in pMAT vector in the cell strain SoluBL21.

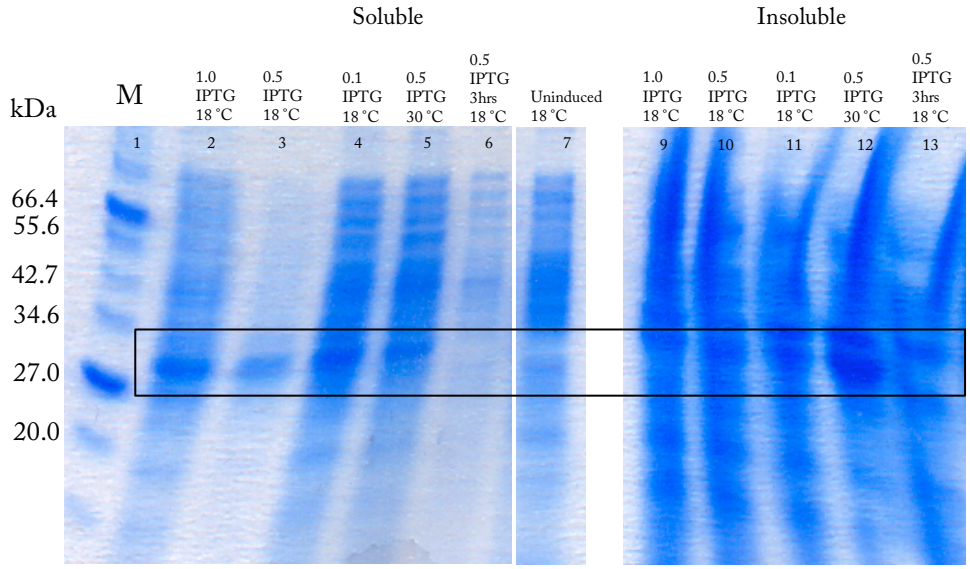


Figure 4.7 Soluble expression of Lipex in SoluBL21 cells

Lipex_pMAT vector grown in SoluBL21 cells and induced with different concentrations of IPTG and expressed at different temperatures. An uninduced culture was grown along side as a control.

Successful expression of Lipex is shown by a strong overexpressed band at ~30 kDa in Figure 4.7 corresponding to the 29.4 kDa MW of non-glycosylated Lipex. This band is observed in all IPTG induced samples and unobserved in the negative control in lane 7, indicating it is a recombinant expressed protein. It is also noted from

comparing 3 h induction (lane 6) and overnight induction (lanes 2–5) that expression yields are increased over 3 h following IPTG induction, leading to the assumption that the protein is not toxic or undergoes proteolysis during expression overnight. The band for Lipex is also observed in all insoluble fractions (Figure 4.7). Insoluble expression was reduced by firstly changing the vector to pET21a which has had shown successful expression of lipases similar to TlL (Steve Wilson, Unilever, Colworth; personal communication), and secondly by optimising the construct with the addition of an N terminal pelB signal sequence to direct the protein for periplasmic expression. The oxidising milieu of the periplasm in *E. coli* bacterial cells provides an environment suitable for the formation of disulphide bonds. This allows for the protein to be expressed with its correct tertiary fold in a soluble and stable form. This 6HIS tag was also removed since it may interfere with enzyme structure and functionality and Lipex can be purified without the presence of any purification tags. A more detailed expression trial with more controls was performed to confirm the expression of Lipex.

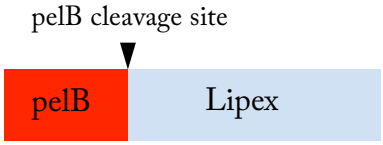


Figure 4.8 Diagram representation of Lipex construct for Lipex_pET21a

Figure drawn starting the N terminus

To determine with high confidence that the protein expressed is Lipex, a negative control screen was performed using a different vector containing a different gene of interest (HypF-N) and an empty pUC19 vector, and multiple single colonies to ensure reproducibility.

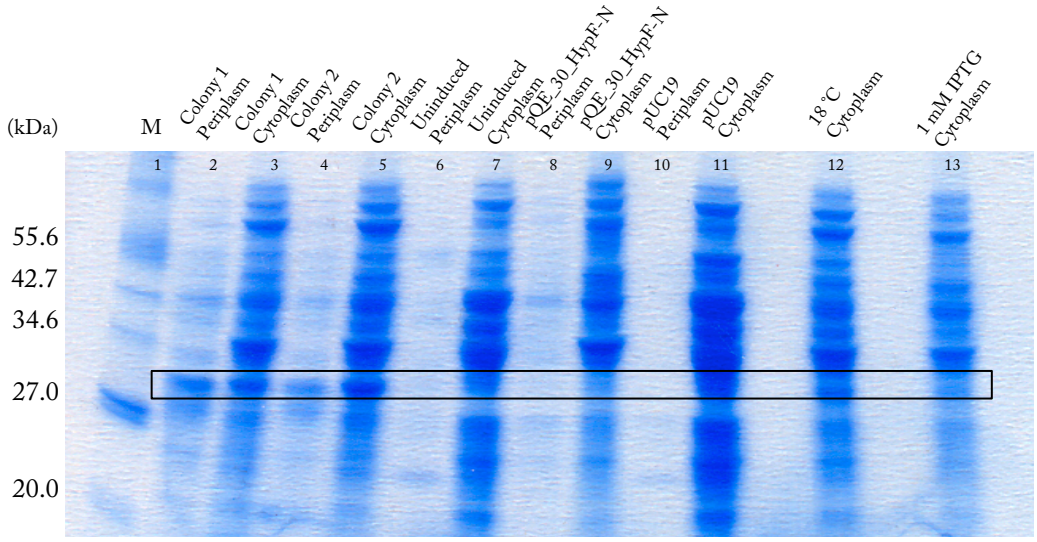


Figure 4.9 Periplasmic expression and purification with negative controls of Lipex expression analysed by SDS-PAGE

Single colonies of SoluBL21 cells containing the Lipex_pET21a plasmid were cultured and induced. Cells were lysed to check for periplasmic and cytoplasmic soluble expression. Numerous negative controls were performed containing different vectors (pQE_30_Hyp-FN and pUC19) and uninduced cultures to confirm recominbant expression of Lipex. Lipex has a molecular weight of 29.4 kDa, approximated by the black box.

Soluble expression of Lipex is noted in the periplasm (Figure 4.9, lanes 2 and 4) as expected, but also in cytoplasm fractions (Figure 4.9, lanes 3 and 5). The enzyme purified from the cytoplasm will still have the pelB signal peptide sequence present on its N terminus, having not been cleaved by the processing unit responsible within the periplasm.

The presence of a band at ~30 kDa in only the lanes containing the Lipex_pET21a vector strongly supports that the expressed protein is only from the Lipex_pET21a vector and therefore must be a recombinant protein. This is further supported by the absence of this band in Figure 4.9 lanes 6 and 7 where no IPTG was added. Since all vectors were transformed into SoluBL21 cells, the band cannot be from a native protein expressed from this cell strain. The periplasmic fractions of the controls were tested in a lipase activity assay. The lipolytic assay uses a using pNp valerate as a substrate carbon chain substrate with a pNp nitrophenol group attached, lipase

enzymes cleave the nitrophenol group from the substrate, and the free nitrophenol group leads to an change in colour in the solution, the change in colour can be quantified and measured at λ = 405. The increase in absorbance at λ = 405 is attributed to lipase activity, increase in absorbance and faster rates indicate a more catalytically active lipase enzyme. Figure 4.10 provides confidence that the expressed protein assumed to be Lipex, is indeed a lipase.

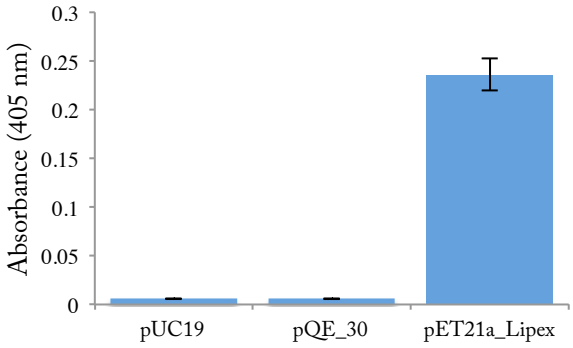


Figure 4.10 Activity assay of periplasmic fraction soluble protein samples on pNp valerate substrate

SoluBL21 cells were grown and induced containing negative control vectors pUC19 and pQE_30, periplasm fraction was isolated as described previously and 20 µl was used in an activity assay. Absorbance at 405 nm was read after 15 min at 37 °C *as an end point value, n = 3.*

4.2.3.2 Expression optimisation

Screening for optimal expression conditions was performed in M9 minimal media (results not shown) where growth and expression temperatures were varied (37, 30, 25 and 37, 30, 25 and 18 °C respectively) along with IPTG concentrations for induction (0.1, 0.5 and 1.0 mM). Across the screen it was found that growth and expression at 25 °C with 0.5 mM IPTG concentration was optimal, consistent with manufacturer instructions of the SoluBL21 strain (Genlantis). Most optimisation of M9 isotopically labelled cultures are limited to temperature, induction concentrations and expression temperatures, but there are other factors that can still be tested such as media composition (inclusion of glycerol) and types of flask where aeration can affect growth rates (Steve Wilson, Unilever; Colworth, personal communication). These types of expression optimisation have not yet been performed since the yield of

protein is acceptable from single labelled cultures (~7 mg/L) is both reproducible and economically viable.

Mass spectrometry was performed on a supernatant sample to confirm the exact MW of the protein assumed to be Lipex.

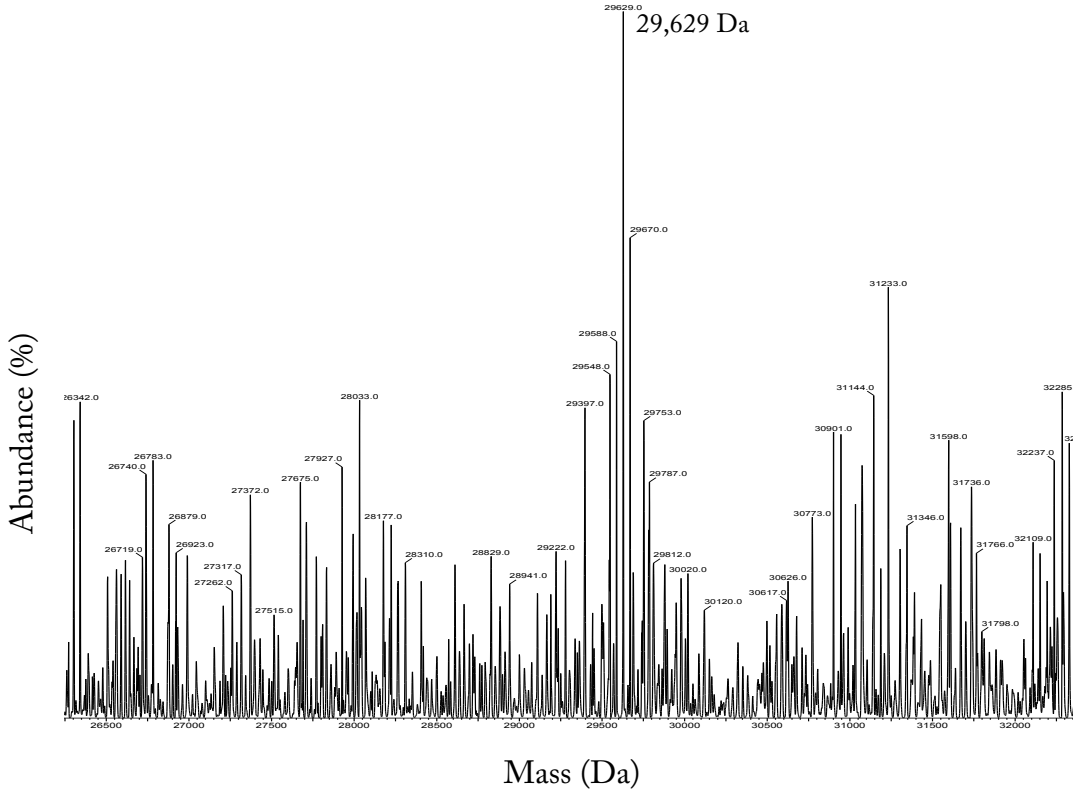


Figure 4.11 Electrospray-ionisation Time of Flight (ES-ToF) mass spectrometry profile of periplasm purified fraction.

Supernatant lysed by periplasmic purification was passed over a desalting PD-10 column equilibrated in H_2O. Elution was performed with 1 ml H_2O and sample was run for ES-ToF.

The mass spectrometry results are quite noisy due to the sample preparation. The main protein sample detected is at 29,629 Da, which corresponds to the molecular weight of Lipex, calculated at 29,615 Da. Differences in 14 Da could be due to the presence of salt of Lipex, but 14 Da is within the accepted range of standard deviation for the sample purity and experimental set up. Combining all results gives strong confidence that the protein band highlighted in Figure 4.9 is Lipex.

4.2.4 Purification

The purification protocol consists of lysing only the periplasm of cells using a sucrose gradient. Briefly, an osmotic gradient was created by re-suspending induced overnight cells in high sucrose media, the cells are pelleted and then re-suspended in water. This ruptures and lyses the periplasm only. The supernatant was brought to a pH of 9.0 that gives Lipex an overall negative charge. It was loaded onto an anionic ion exchange column (GE healthcare) at low salt concentrations and eluted with an increasing concentration gradient of sodium chloride (Figure 4.12). The fractions containing Lipex were heated at 55 °C for one hour to precipitate a bound contaminating protein and the soluble sample was finally gel filtered to give Lipex at purities suitable for NMR studies. Since, to the best of our knowledge, this is the first time Lipex has been expressed in an *E. coli* system all steps of purification have been presented.

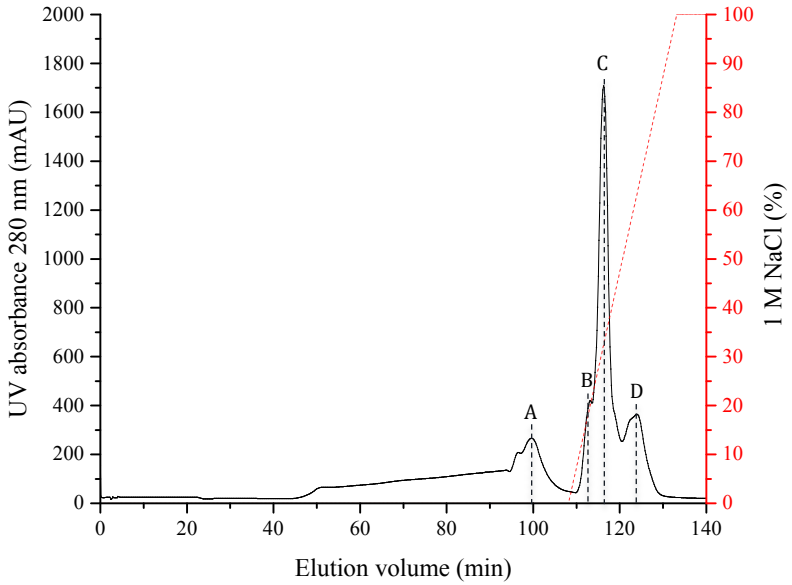


Figure 4.12 Ion exchange chromatogram of Lipex.

Following periplasmic cell lysis the supernatant was taken and loaded onto a high-trap Q FF ion exchange column (GE). Lipex was eluted in a gradient of 100 steps to 100% 1 M sodium chloride, 25 mM sodium phosphate, pH 9.0 and fractions taken are shown here. Protein is measured by following the absorbance at λ 280nm.

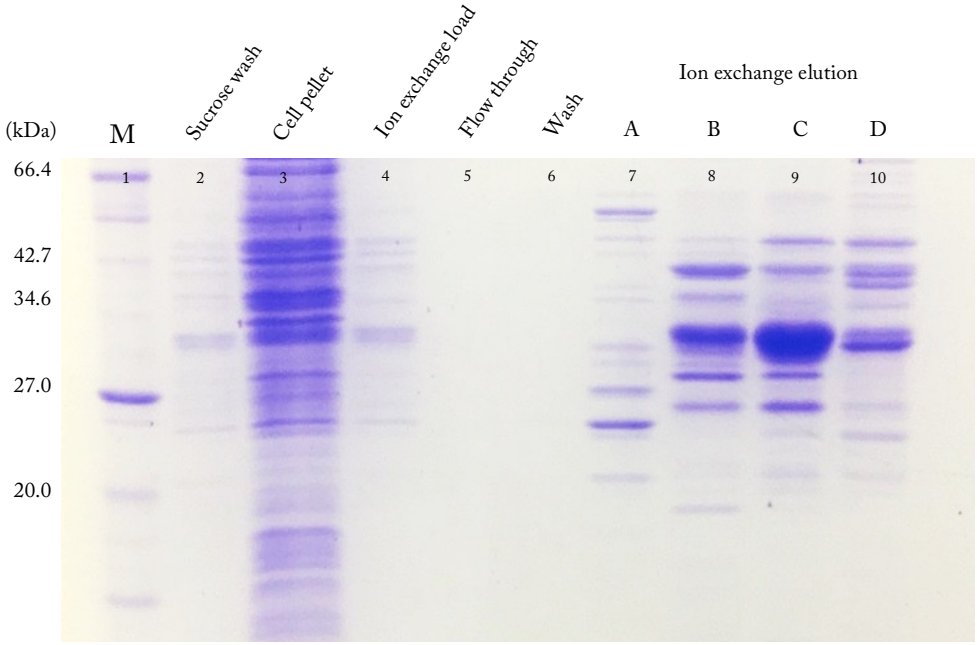


Figure 4.13 SDS-PAGE gel of the sucrose gradient periplasmic purification and ion exchange of Lipex

Fractions taken from specified points of the periplasmic purification. Fractions A-D from the ion exchange correspond to the full chromatogram shown in Figure 4.12. A band corresponding to protein of correct MW to Lipex and is highly expressed is noted as a strong band in fractions B and C from ion exchange.

It is noted that some Lipex is found in the supernatant after cells have been resuspended in 25% sucrose (Figure 4.13, lane 2). There should not be any bands present in the sample since no cell lysis steps have been performed. The faint band appears to be the same MW of Lipex. The source of this could be from cells that have lysed from shearing during re-suspension or centrifugation. Furthermore we note that there is Lipex present in the cell pellet after the periplasm is lysed, however this is most likely Lipex expressed and present in the cytoplasm. The soluble fraction loaded on the ion exchange shows a faint band for Lipex since it is diluted in ~200 ml. Lipex is not lost from the flow through or the wash indicating a strong interaction with the ion exchange column and optimal conditions for ion exchange. The ion exchange elution (Figure 4.13 lanes 8–10) shows that further purification is required. The contaminants were removed by heat denaturation and size exclusion (SEC) gel filtration.

There is natural variability between each culture and purification, whilst these slight changes do not affect expression yields in any significant manner, one issue observed with purifying Lipex after ion exchange was the presence of a contaminating protein.

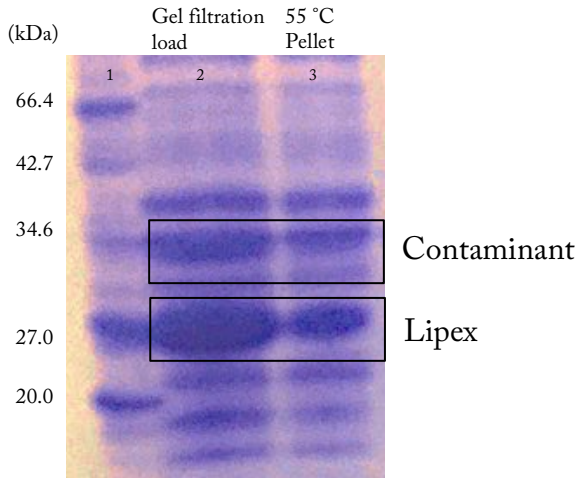


Figure 4.14 Precipitate of Lipex with contaminating protein during heat incubation

Lipex purified by ion exchange is concentrated to 5 ml volume and incubated at 55 °C for one hour to precipitate a contaminating protein. Although this is removed some Lipex is lost in an insoluble pellet indicated in lane 3.

1 = Marker

2 = Soluble fraction after heat incubation

3 = Insoluble fraction after heat incubation

A harsh thermal denaturation step of incubating the sample at 55 °C was applied to intentionally precipitate the contaminating protein, leaving the mildly thermostable Lipex unaffected. Whilst the contaminant could be removed after precipitation and centrifugation, it is noted that some of the Lipex remains bound to the unfolded contaminant and is lost from the soluble fraction when pelleted, reducing the final yield of purified protein. The amount of bound contaminant varied from each culture but acts as a major step in loss of yield, Figure 4.14 illustrates the extent of Lipex lost from precipitating out the complex. SEC gel filtration was performed as the final step to give pure Lipex.

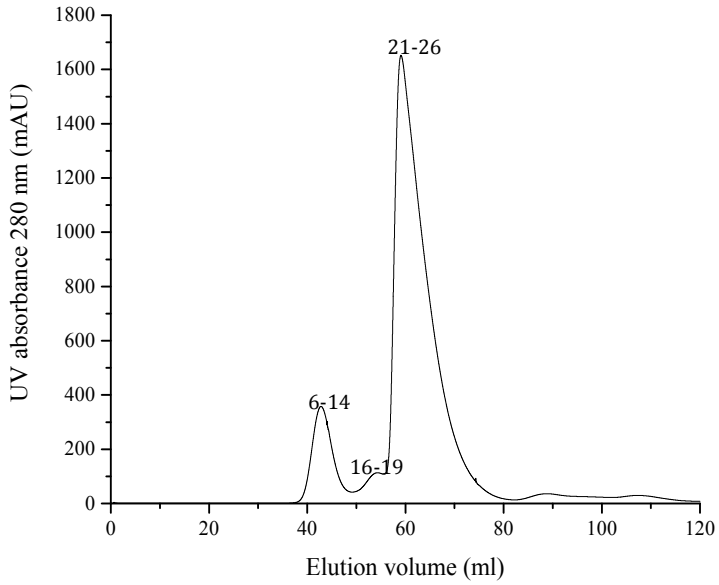


Figure 4.15 SEC gel filtration purification of Lipex after ion exchange.

Collected fractions from ion exchange containing Lipex were pooled and run on a S75 SEC gel filtration column. The absorbance at 280nm was measured.

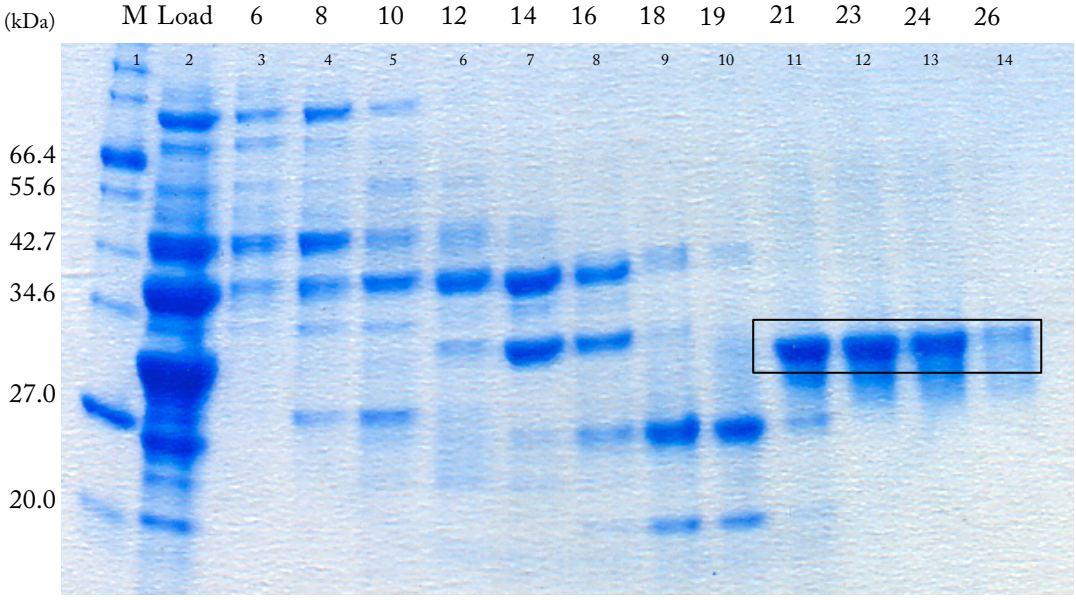


Figure 4.16 SDS-PAGE gel of fractions taken from gel filtration of Lipex

Lanes 3 to 14 contain fractions taken from the gel filtration of which the full chromatogram and corresponding fractions shown in Figure 4.15. The box highlights pure Lipex at ~30 kDa.

It is noted in Figure 4.15 that the peaks preceding Lipex during SEC gel filtration are complex formations of aggregated proteins; the first peak (Figure 4.15 fractions 6 – 14) elutes very close to the column void volume of 40 ml, and consists of a number of

proteins. Since they are not separated by MW size by SEC gel filtration they must have formed a complex (Figure 4.16, lanes 3–7). Bands noted in fractions 18 and 19 (Figure 4.16, lanes 9 & 10) have lower MW sizes than Lipex but elute from the gel filtration column before Lipex indicating that they form a complex species. Autolytic degradation is not observed with purified Lipex over any course of time, so it is unlikely that these bands are from degraded Lipex. Ion exchange, heat denaturation and gel filtration was sufficient for purifying Lipex from the periplasm of SoluBL21 cells.

NMR experiments require high concentrations of protein (> 200 μM for > 25 kDa proteins). Lipex expression yields are 12, 9 and 4.5 mg/L for ^{14}N, ^{15}N and $^{13}C^{15}N$ labelled cultures respectively, which are suitable and economically reasonable for NMR studies. The purification can be optimised further, particularly at the heat incubation step. An alternative and popular method for purifying lipases is through hydrophobic interaction chromatography (HIC). HIC is similar to ion exchange and involves binding of lipases to an immobilised Sepharose column; elution is also dependent on salt used, ionic strength and pH (Jürgens & Huser, 1981). Precipitation, commonly by ammonium sulphate or ethanol, is another alternative step, often preferred for industrial usage as a first stage of purification given the large quantities of material to purify and its rapid inexpensive method (Saxena *et al*., 2003). TlL has previously been purified by acetone precipitation and successive chromatographic steps on SEC gel filtration, DEAE-Sepharose ion exchange and hydroxyapatite columns (Omar *et al*., 1987b) and by adsorption to acrylic microporous membranes (Taylor, 1989). It may be that precipitation methods remove the interaction between Lipex and the binding contaminant protein; provided that activity and structure is unaffected these provide a likely alternative step that may be trialled in the future.

4.2.5 Concluding remarks

The difficulties and problems encountered from expression of Lipex in yeast and *P. pastoris* were not overcome despite our best efforts. We describe an optimised expression and purification protocol for Lipex in *E. coli*. Yields are sufficient for NMR studies and the protein can be easily isotopically labelled with ^{2}H, ^{13}C and ^{15}N nuclei with no change to the purification protocol.

4.3 Preliminary NMR characterisation of Lipex and catalytic activity under industrially relevant conditions

4.3.1 NMR

Since this is the first reported NMR study on Lipex, initial characterisation was performed to assess the viability of NMR directed studies on a 30 kDa protein. The 1D ^{1}H and 2D ^{15}N-^{1}H HSQC–TROSY NMR spectra were collected and the sample buffer was optimised.

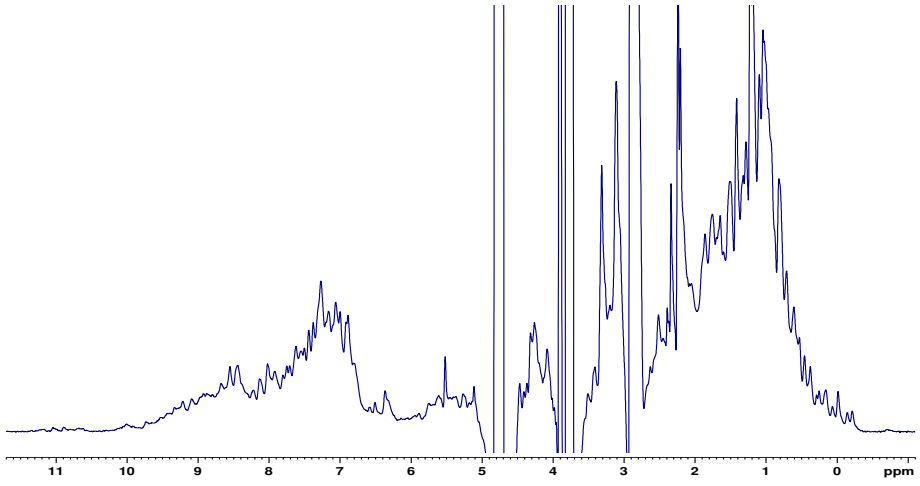


Figure 4.17 1D ^{1}H spectrum of Lipex

Spectrum recorded with 250 µM Lipex in 25 mM sodium chloride, 20 mM Tris-HCl, pH 8.0, recorded at 303 K on an 800 MHz spectrometer

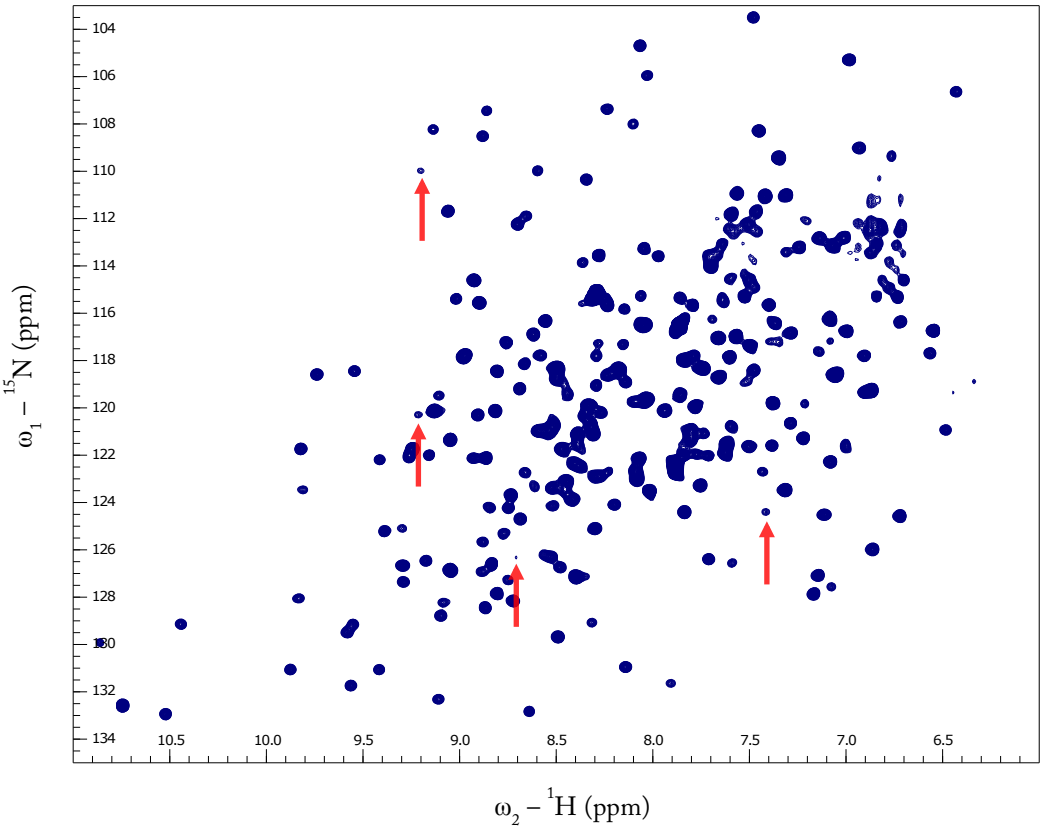


Figure 4.18 2D ^{15}N TROSY-HSQC spectrum of Lipex

Spectrum recorded with 250 µM Lipex in 25 mM sodium chloride, 20 mM Tris-HCl, pH 8.0, recorded at 303 K on an 800 MHz spectrometer

Arrows show examples of peaks with lower signal intensity that may originate from a less populated conformation.

The 1D ^{1}H spectrum depicts a folded protein (indicated by methyl peaks at < 0 ppm) (figure 4.17). For a protein with a MW of 29.4 kDa consisting of 269 amino acids, the 2D ^{15}N TROSY–HSQC spectrum shows well resolved and dispersed peaks, indicating a well folded protein (Figure 4.18). We observe ~210 backbone amide peaks, peaks missing may be in conformational exchange. There is little line broadening and peak overlapping observed. The ^{15}N TROSY-HSQC spectra offered much greater signal intensity and resolution than non-TROSY versions (detailed further in Section 1.2); therefore all HSQC spectra were collected as TROSY versions. The SEC gel filtration and NMR data suggest that the protein is monomeric in solution. The SEC gel filtration and spectrum are concordant with a monomeric protein with a 30 kDa MW. Purifications were repeated using 2 mM DTT concentrations and no changes were recorded in the gel filtration chromatograms or HSQC spectrum (data not shown).

There are a number of peaks across the spectrum highlighted in Figure 4.18 that are at weaker signal intensity than the average. This may indicate a protein undergoing conformational exchange between states with the weaker peaks originating from a less populated transient species. For Lipex only samples, approximately 200 peaks can be observed, with ~60 peaks assumed missing (excluding proline residues); some may be within the overcrowding regions. More peaks are observed, however, with the addition of surfactants such as LAS or the addition of surfactant and calcium mixtures, and to a lesser extent by addition of calcium and changing temperature, as described later. These unobserved peaks could be from the regions undergoing conformational changes between lid open and closed forms of the lipase. The intermediate exchange and dynamics of this structural inhomogeneity could lead to these peaks being unobserved through resonance broadening. These factors are significant to the functional activity of Lipex, within an industrial usage for detergent applications; it is these results that are the primary focus of this thesis. The additives bound to the lipase could stabilise a particular conformation, giving rise to the observation of new peaks. It may also be that mobility is crucial to enzymatic activity, and the locked conformation may be less active due to this constraint. These peaks missing in the Lipex only spectrum may therefore be of functional significance.

4.3.2 Comparisons to Unilever glycosylated Lipex

The Lipex used by Unilever is produced industrially as a glycosylated enzyme. Since Lipex used in this study is from an *E. coli* host, our NMR studies are performed entirely on a non-glycosylated version. We performed experiments to determine the effect of glycosylation on the 1D ^{1}H NMR spectrum and lipolytic activity. The glycosylated Lipex supplied by Unilever is unlabelled and natural abundance of labelled ^{15}N and ^{13}C are too low, preventing any heteronuclear 2D NMR experiments.

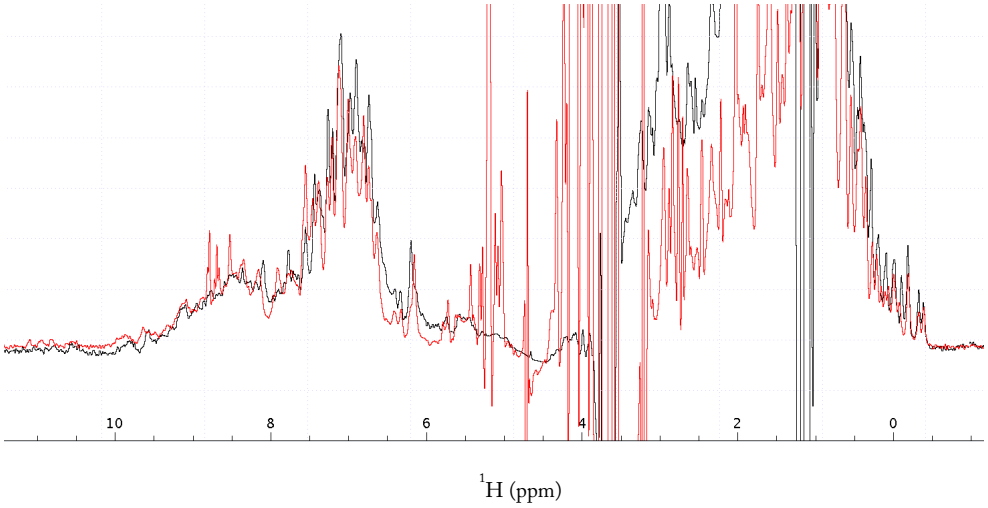


Figure 4.19 1D 1H NMR spectrum of Lipex in glycosylated and non-glycosylated form

^{1}H 1D spectrum of glycosylated Lipex from yeast expression donated by Unilever (red) and non-glycosylated Lipex from E. coli expression (black).

Both spectra were collected with 250 µM Lipex in 25 mM sodium chloride, 20 mM sodium phosphate, pH 8.0, recorded at 313.15 K on a 600 MHz spectrometer.

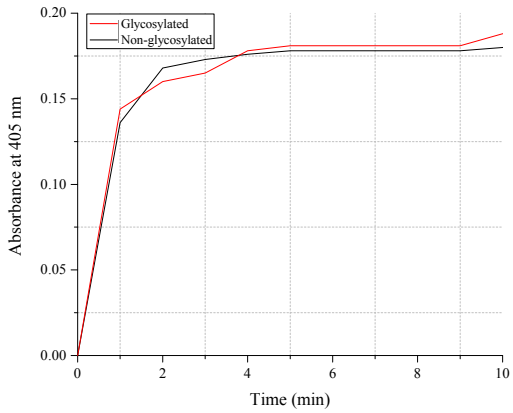


Figure 4.20 Glycosylation does not affect activity of Lipex

Activity of glycosylated (red) and non-gylcosylated Lipex (black) on pNp valerate at 303.15 K.

The spectra in Figure 4.19 overlay very well, particularly in the amide (5-10 ppm) and methyl region from CH_3 groups at (0-2 ppm). The aliphatic region (2-5 ppm) is slightly different, the large peaks found between ~ 3.5 - 5.5 ppm in the glycosylated red spectrum are predominantly from the glycosylation along with buffering agents

and residual chemicals that remain after the purification protocol described for Unilever supplied Lipex (as determined by CD and DLS). This suggests that the glycosylation does not alter the enzyme conformation or structural properties.

Protein glycosylation is a common post-translational modification of native proteins used in pharmaceutical industries, where it can increase solubility, stability and reduce aggregation (Elliot *et al*., 2003). Any glycosylation from expression in a yeast host is lost when switching to an *E. coli* system. Glycosylation does have an effect on increasing Lipex stability (Unilever, personal communication), the NMR experiments performed do not require the additional stability provided by glycosylation, unless the activity and interactions are affected our results should not be affected by the loss of glycosylation. The lipolytic activity on pNp valerate and butyrate substrates are consistent and comparable with non-glycosylated Lipex expressed in *E. coli* and glycosylated Lipex supplied by Unilever (Figure 4.20). Additionally it is more advantageous to NMR studies for the protein to be non-glycosylated since it reduces the MW by approximately 5 kDa, thereby reducing rotational correlation times and increasing T_2 rates. Whilst there are protocols for removing glycosylation, heterogeneous samples with a mixture of non- and glycosylated Lipex will mar any NMR derived functional results obtained.

4.3.3 NMR buffer optimisation

The choice of buffer can affect the sensitivity of NMR signal detection by up to 50% (Kelly *et al*., 2002); factors such as pH, conductivity and ion concentration and mobility are required to provide protein stability and prevent aggregation but also need to be optimised for enhanced sensitivity.

Lipex exhibits similar stability levels across all buffering agents; whilst sodium phosphate is a preferred buffering agent for biomolecular NMR since it does not give signal in the proton dimension, the phosphate group can react with calcium chloride to form calcium phosphate salt $Ca_3(PO_4)OH$. This has detrimental effects for studying the lipase at different water hardness levels. After screening protein stability, with calcium and surfactant additives in a range of buffers (sodium acetate, sodium

phosphate, HEPES and TRIS) and pH values, the optimal buffer selected for NMR experiments was 25 mM Tris-HCl, 20 mM sodium chloride at pH 8.0.

Performing NMR at slightly acidic pH values, around 5.5, reduces the chemical exchange rate between amide protons and water thus increasing sensitivity. Aggregation has previously been observed for TlL at pH 5 in 50 mM sodium acetate at concentrations exceeding 1 mg/ml (30 μM) (Rodriguez-Larrea *et al*., 2006). Aggregation also occurred for Lipex at concentrations exceeding 50 μM in 50 mM sodium acetate in pH 5. The unfolding and precipitation may likely be due to histidine protonation and resultant loss of rigidity and increased flexibility, along with the electrostatic interactions of Lipex and formation of ion pairs (Jaenicke & Bohm, 1998). Small changes between pH 8 and 6 where the net charge decreases from -12 to -5 are enough to reduce the kinetic stability of Lipex (Wang *et al*., 2013). Since Lipex has an isoelectric point of 5.0 (Mogensen *et al*., 2005), it is expected that at pH 5 and lower the surface potential is drastically altered and highly positive which could promote electrostatic interactions and mediate unfolding, aggregation and precipitation. Alternative buffers such as BIS-Tris or MES that are NMR preferable (Kelly *et al*., 2002) and suitable for buffering in the pH range of 5.5 – 6.5 were not tested. Whilst performing NMR at acidic pH would not benefit studies with surfactants, it could provide increased signal intensity in 3D spectrum recorded for backbone assignment. No change in signal intensity or peak detection was observed changing pH from 7 – 9, since the enzyme is most active at pH 8.0 (Unilever, personal communication) this value was used as standard for NMR experiments. Interactions of TlL with surfactants such as SDS can change with pH (Wang et al 2013), where acidic pHs below 6 led to increasing instability and an increased propensity for unfolding.

It is preferential for NMR that the ionic strength of the sample is kept low (<100 mM), to reduce the conductivity of the sample particularly when measuring signal in multidimensional experiments using a cryogenic probe (Flynn *et al*., 2000). This may not always be possible since proteins are more prone to aggregate at low salt concentrations, however for Lipex the protein is stable at 20 mM sodium chloride. The factors more significant to the bioindustrial usage of Lipex (surfactant and calcium ion concentration) were assessed first. Whilst the ionic strength of the buffer can have an effect on the interactions between surfactants and protein (Helenius *et al*.,

1979), along with temperature, these two factors are not used as variables since the complexity of the system would expand beyond what time frame was available in the PhD.

4.3.4 Enzyme stability

The stability of the non-glycosylated variant was assessed by recording a 2D ^{15}N-HSQC spectrum and comparing after a 90 day period. It is important to understand the effects of losing glycosylation on the stability of the enzyme from an NMR perspective. Whilst activity should not be affected, adverse effects to the enzyme stability will reduce the shelf life of the protein.

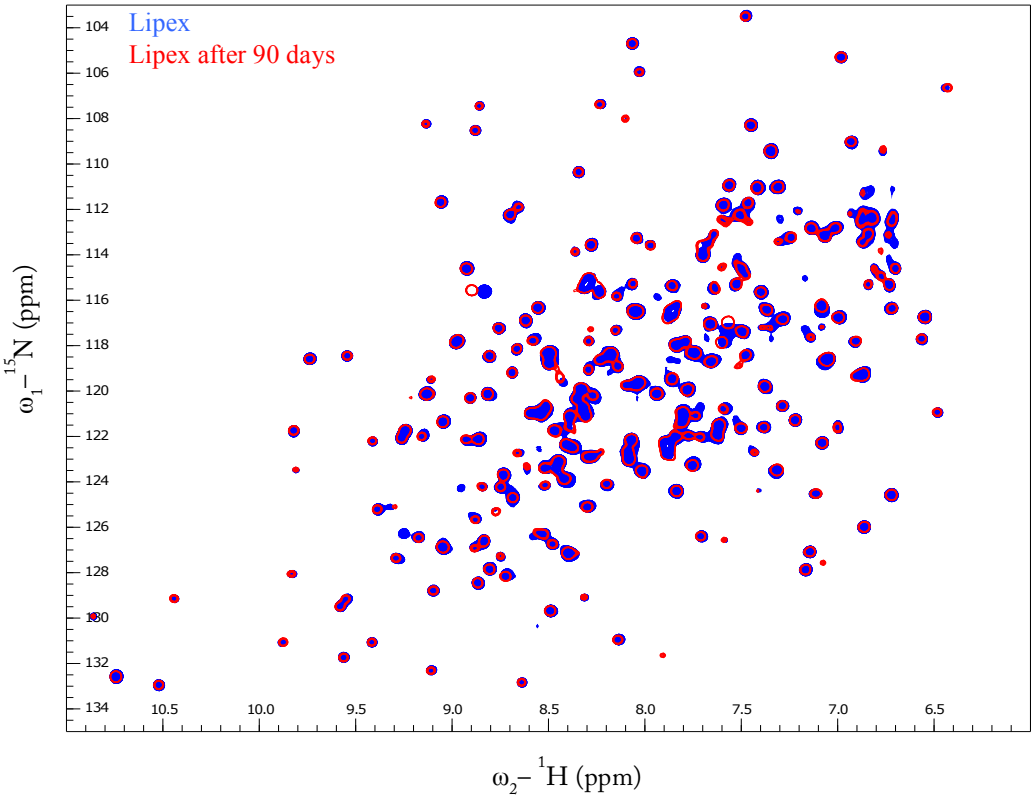


Figure 4.21 Lipex stability measured by 2D NMR

2D ^{15}N TROSY HSQC spectrum of Lipex at 0 hrs (blue) and 90 days (red) stored at 4 °C. Spectra recorded on the same sample of Lipex at 250 µM in 25 mM sodium chloride, 20 mM Tris-HCl, pH 8.0, recorded at 303 K on an 800 MHz spectrometer

The protein was stored at 4 °C in 25 mM Tris-HCl, 20 mM sodium chloride at pH 8.0 over 3 months. There was no detection in visible precipitation or change of protein concentration as measured by absorbance at 280 nm, or from NMR signal intensity.

No changes in the spectra indicative of significant aggregation or degradation are detected. These data demonstrated that the non-glycosylated enzyme is highly stable. Additionally the enzyme can be concentrated to values > 1.5 mM avoiding precipitation, and avoids degradation after extensive time in a low ionic strength buffer. The protein is also stable after 3D NMR experiments were performed at 323 K over a 7 day period; as measured by comparing ^{15}N-HSQC spectra collected before and after the 3D experiments, showing a high degree of thermal stability without glycosylation.

4.3.5 Lipex activity under industrially relevant conditions

As detailed in section 1.12, Unilever use a range of surfactants in their detergent to modulate lipase activity. There is little published and accessible information on TlL activity in the presence of a combination of surfactants. A condition screen was performed varying ratios of each surfactant to identify catalytically interesting combinations for NMR studies. Conditions could be quickly and cheaply identified where Lipex is highly active and where activity is abolished; these functionally interesting conditions could then be studied by NMR methods. We also used the activity assays as a platform for understanding the effects of individual surfactants on Lipex activity. One particular interest was the introduction of the biosurfactant rhamnolipid (R), and its affect on surfactant formulations and Lipex activity. The percentages of surfactant levels in activity assays ranges slightly but are ~0.001% which is indicative of conditions in the washing machine during cleaning. The activity assay screen was performed using glycosylated Lipex donate from Unilever since expression of unlabelled material had not yet been achieved. The assay developed by Unilever uses a *p*-nitrophenol substrate with a carbon chain to emulate a lipid structure. Unilever commonly use pNp valerate, a 5 carbon chain substrate for measuring the catalytic activity of Lipex. The cleavage of the *p*-nitrophenol group from the carbon chain leads to a colour formation in the sample which is detected by measuring absorbance at $\lambda = 405$ nm.

Autodegradation is a common problem for *p*-nitrophenol substrates especially in the presence of strong alkalis such as sodium hydroxide (Gupta *et al*., 2003). The usage of *p*-nitrophenol substrates has a number of inherent problems such as poor

reproducibility, but is still a common and standard method for measuring lipase activity since it is a very efficient and cheap assay. A number of conditions were sampled varying the ratios of surfactants, and assessed at four water hardness levels. The four water hardness levels chosen are at 0, 5, 18.75 and 50 degrees French water hardness and cover a range from very soft to very hard water levels; further details are shown in Table 1 (2.14). Studying different water hardness levels will illustrate the effects from using certain surfactant formulations in different regions within the U.K. and across different countries. Numerous conditions experienced loss of activity, however some interesting mixtures exhibited increase in lipolytic activity, which increased further as a function of water hardness.

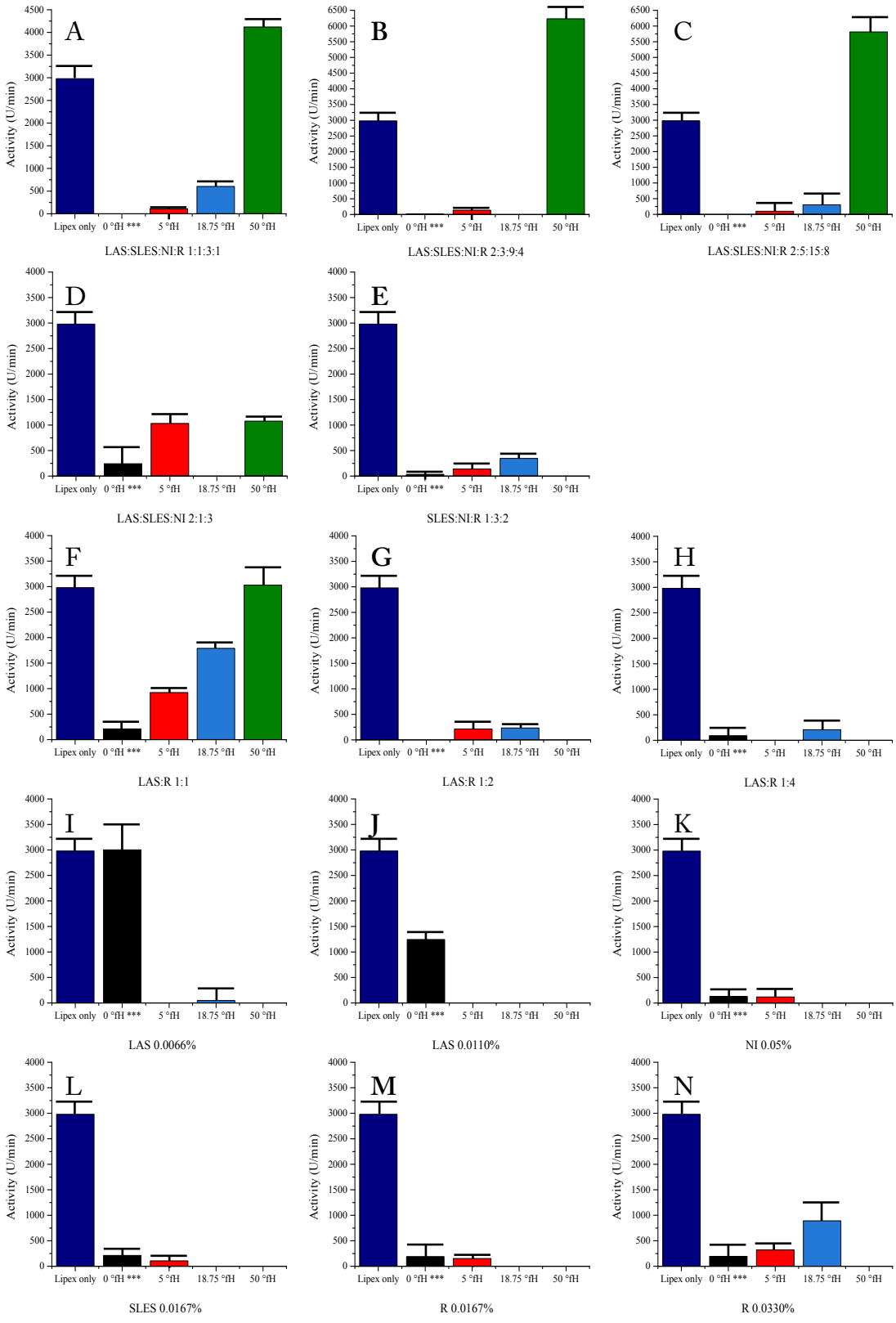


Figure 4.22 Surfactant formulation screening on Lipex activity at increasing water hardness level.

Activity assays were performed on glycosylated version of Lipex donated by Unilever on pNp valerate substrate at 303 K. Four water hardness levels were assessed: 0 degrees French

hardness (Tris-HCl) (green), 5 degrees French hardness (20 mg/L $CaCl_2$, 10 mg/L $MgCl_2$) (yellow), 18.75 degrees French hardness (75 mg/L $CaCl_2$, 50 mg/L $MgCl_2$) (orange) and 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$) (red). Absorbance at 405 nm was read as an end point value after 15 min. The control without Lipex was subtracted from each sample reading, and a mean was taken (n=3). The first 3 charts have an extended Y-axis than the rest. Excluded datasets are designated due to negative slope (*), insignificant slope (**) and negative activity after baseline subtraction (***).

The activity screen in Figure 4.22 surveyed a number of conditions and identified some interesting results despite only the data being reported as an average activity in U/min, which gives limitations in comparing data (an advanced fluoresce plate reader capable of kinetic studies was not available till the later stages of the PhD). Magnesium alone

Activity is significantly reduced or abolished in a large number of conditions tested, indicating a much stronger negative effect of surfactant on lipases. There are a few conditions were the activity of Lipex is enhanced in the presence of surfactants; particularly when calcium is present at 50 °fH, in the case for LAS:SLES:NI:R in 2:3:9:4 (Figure 4.22 B) and increasing the surfactants expect for LAS by 2 to give a ratio of 2:5:15:8 (Figure 4.22 C). These two results, along with LAS:SLES:NI:R 1:1:3:1 (Figure 4.22 A) indicate how it is within mixtures of all surfactants, with high concentrations of calcium that the activity is enhanced by up to a factor of 2. This exemplifies a major significance of calcium in Lipex activity when surfactants are present. Anionic surfactants can precipitate in hard water conditions (Sachdev & Krishnan, 1997), which may affect readings, over 50% of the activity assays performed at 50 °fH were discarded due to erroneous slopes, indicating a major problem, however results reported are baseline corrected with control samples in absence of Lipex. Calcium additionally plays a positive defining role in LAS:R at 1:1 (Figure 4.22 F), but only the absence of calcium with LAS at 0.0066% enhances the activity, suggesting a role of calcium on rhamnolipid (R). Interestingly LAS:R at 1:2 (Figure 4.22 G) and 1:4 (Figure 4.22 H) do not elicit the same activity trends; implying a dependence on the ratios of surfactants, anionic strength and calcium. Calcium binding at anionic surfactant micelles such as LAS and R can displace the sodium ions of LAS, changing the electrostriction and configurational freedom

(Stoner *et al*., 2005). This could have effects on the electrostatic interactions with Lipex, and may affect Lipex adsorption to surfactants and therefore activity. The following activity assays were performed using a plate reader measuring kinetic data that gave more reproducible readings.

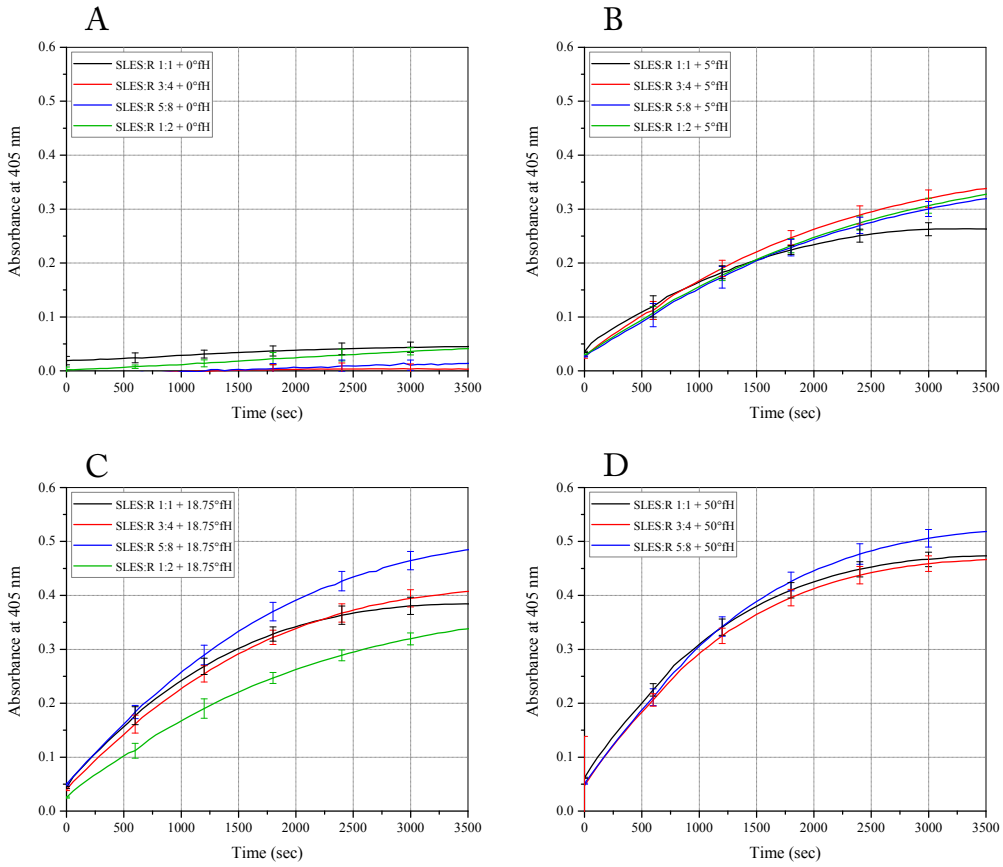


Figure 4.23 Effects of weakly anionic surfactant mixture on Lipex activity

Activity assay of Lipex with SLES:R in ratio of 1:1, 3:4, 5:8 and 1:2 on pNp butyrate at different water hardness at 30 °C. Values plotted are a mean of 3 experiments. SLES:R at 1:2 at 50 °fH was omitted due to negative baseline.

Four water hardness levels were assessed: 0 degrees French hardness (Tris-HCl), 5 degrees French hardness (20 mg/L $CaCl_2$, 10 mg/L $MgCl_2$), 18.75 degrees French hardness (75 mg/L $CaCl_2$, 50 mg/L $MgCl_2$) and 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$). Absorbance at 405 nm was recorded at 60 s intervals. The control without Lipex was subtracted from each sample reading, and a mean was taken (n=3).

Introducing R into the mixture containing LAS:SLES:NI leads to much higher activity rates at 50 °fH (Figure 4.22 A, B, C). To better understand the role of R in the surfactant mixture, kinetic activity measurements with NI or SLES individually with R were tested. We aimed to determine which surfactants and their interactions lead to

the activity increase at 50 °fH. All ratios of SLES:R tested saw an increase in activity across increasing water hardness levels (Figure 4.23 A-D). Unlike LAS:R 1:1 where only equal amounts of surfactants gives activity increases as a positive effect of calcium (Figure 4.22 F), when combined with SLES the concentration of R can be increased almost double to a ratio of 5:8 which demonstrates the highest activity at both 18.75 and 50 °fH (data omitted for SLES:R 1:2 at 50 °fH due to erroneous curve). This is not observed for LAS:R at 1:2 (Figure 4.22 G), indicating a more positive effect between SLES and R than for LAS and R.

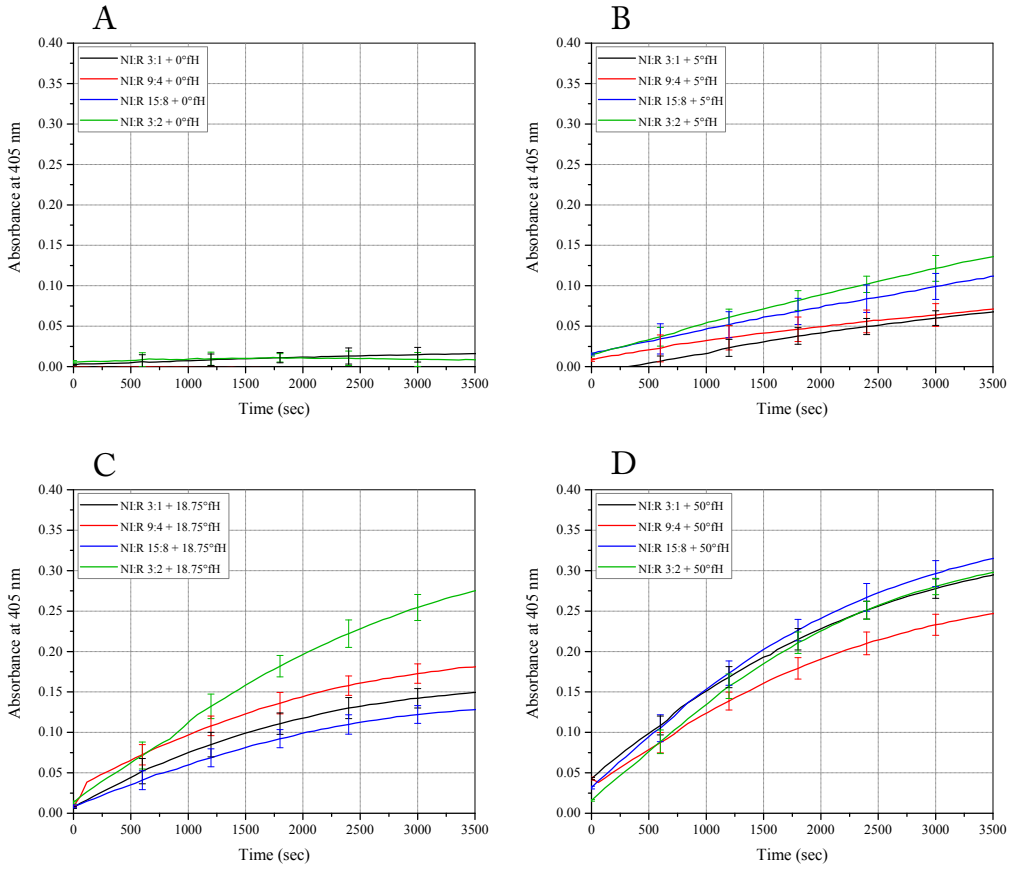


Figure 4.24 Effects of non-ionic and weakly anionic surfactant mixture on Lipex activity

Activity assay of Lipex with NI:R in ratio of 3:1, 9:4, 15:8 and 3:2 on pNp butyrate at different water hardness at 30 °C.

Four water hardness levels were assessed: 0 degrees French hardness (Tris-HCl), 5 degrees French hardness (20 mg/L $CaCl_2$, 10 mg/L $MgCl_2$), 18.75 degrees French hardness (75 mg/L $CaCl_2$, 50 mg/L $MgCl_2$) and 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$). Absorbance at 405 nm was recorded at 60 s intervals. The control without Lipex was subtracted from each sample reading, and a mean was taken (n=3).

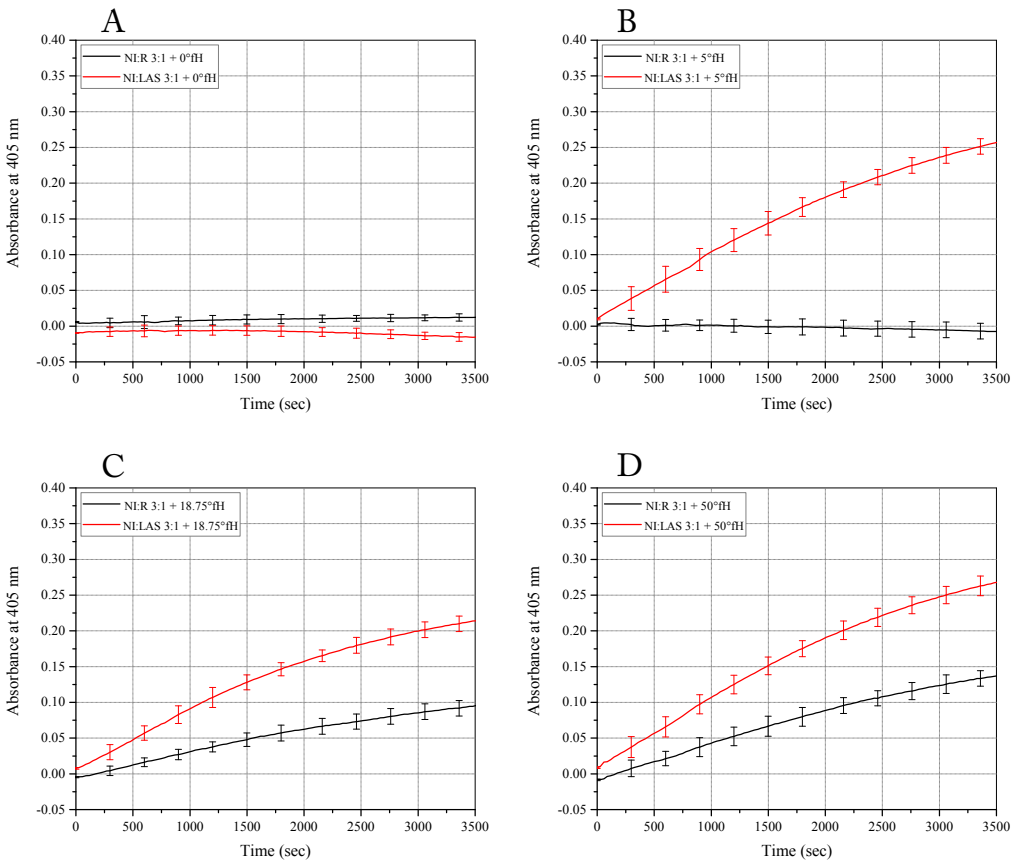


Figure 4.25 Effects of anionic surfactants on Lipex activity in presence of non-ionic surfactant

Activity assay of Lipex with NI:R in ratio of 3:1, and NI:LAS in ratio of 3:1 on pNp butyrate at different water hardness at 30 °C.

Four water hardness levels were assessed: 0 degrees French hardness (Tris-HCl), 5 degrees French hardness (20 mg/L $CaCl_2$, 10 mg/L $MgCl_2$), 18.75 degrees French hardness (75 mg/L $CaCl_2$, 50 mg/L $MgCl_2$) and 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$). Absorbance at 405 nm was recorded at 60 s intervals. The control without Lipex was subtracted from each sample reading, and a mean was taken (n=3).

The activity of Lipex increases in a surfactant mixture of NI:R in all ratios as a function of increasing water hardness (Figure 4.24 A-D and Figure 4.25 A-D). This is a similar trend to SLES:R but the maximum activity reached for NI:R is less. It may be that the anionic nature of SLES gives a greater increase of activity relative to NI when R is present. Both SLES:R and NI:R increase in activity in all ratios as calcium levels increase, implying a beneficial role of calcium with these surfactants. Anionic and non-ionic surfactant mixtures generally provide greater stability than solely anionic surfactants. This is due to two separate reasons, directly there is a reduced

concentration of surfactant present to bind and cause surfactant induced unfolding. Secondly the presence of non-ionic head-groups increases the spacing between anionic groups, thereby indirectly reducing the strength that Ca^{2+} ions bind to the surfactants (Stoner *et al*., 2005). Whilst the stability might be improved, the activity with anionic/non-ionic mixtures is reduced compared to anionic only (Figure 4.25).

The system demonstrated is highly complex with factors such as charge, size and surfactant ratios playing major contributing factors. Since calcium ions bind to anionic surfactants (Petkov *et al*., 2010, Chen *et al*., 2013) changes in Lipex activity with LAS, SLES and R present are likely due to this interaction. The differences noted in Figure 4.25 may be attributed to the effect of calcium on LAS and R. At NI:LAS 5 °fH activity is present, however for NI:R the activity is diminished (Figure 4.25 (B)). It is known that R, along with SLES as weakly anionic surfactants, chelate calcium ions; it could be that the free calcium ions below 18.75 °fH are sequestered by SLES and R when these surfactants are present, and the positive influence of calcium is not seen (Figure 4.25 (C,D)) until the concentration of calcium surpasses what can be chelated by SLES and R. Since LAS is strongly anionic, it is likely to be stronger at chelating calcium and may explain why activity is noted for NI:LAS 3:1 but devoid in NI:R 3:1 at 5 °fH (Figure 4.25 (B)).

In addition, the lipids present form a micelle at these concentrations above their CMC values, the size of the micelle and its interactions with Lipex may also affect activity, perhaps the SLES:LAS micelle is too compact and interacts with Lipex in a manner not to promote, but reduce interfacial activation. Experiments such as diffusion-ordered NMR spectroscopy, isothermal titration calorimetry (ITC) and interfacial activation assays could provide further information on how calcium affects the micelle properties and its interactions with Lipex.

4.3.6 Secondary structure of Lipex under industrially relevant conditions as measured by Circular Dichroism (CD)

CD was performed on a number of conditions used for screening activity (Figure 4.26). Interfacial activation is a phenomenon made possible through structural rearrangements of lipases. It was hypothesised that conditions increasing lipolytic activity would be concurrent with Lipex structural changes as interfacial activation is increased. CD was used as a technique for rapidly measuring secondary structural changes of unlabelled Lipex with addition of surfactants across different water hardness levels.

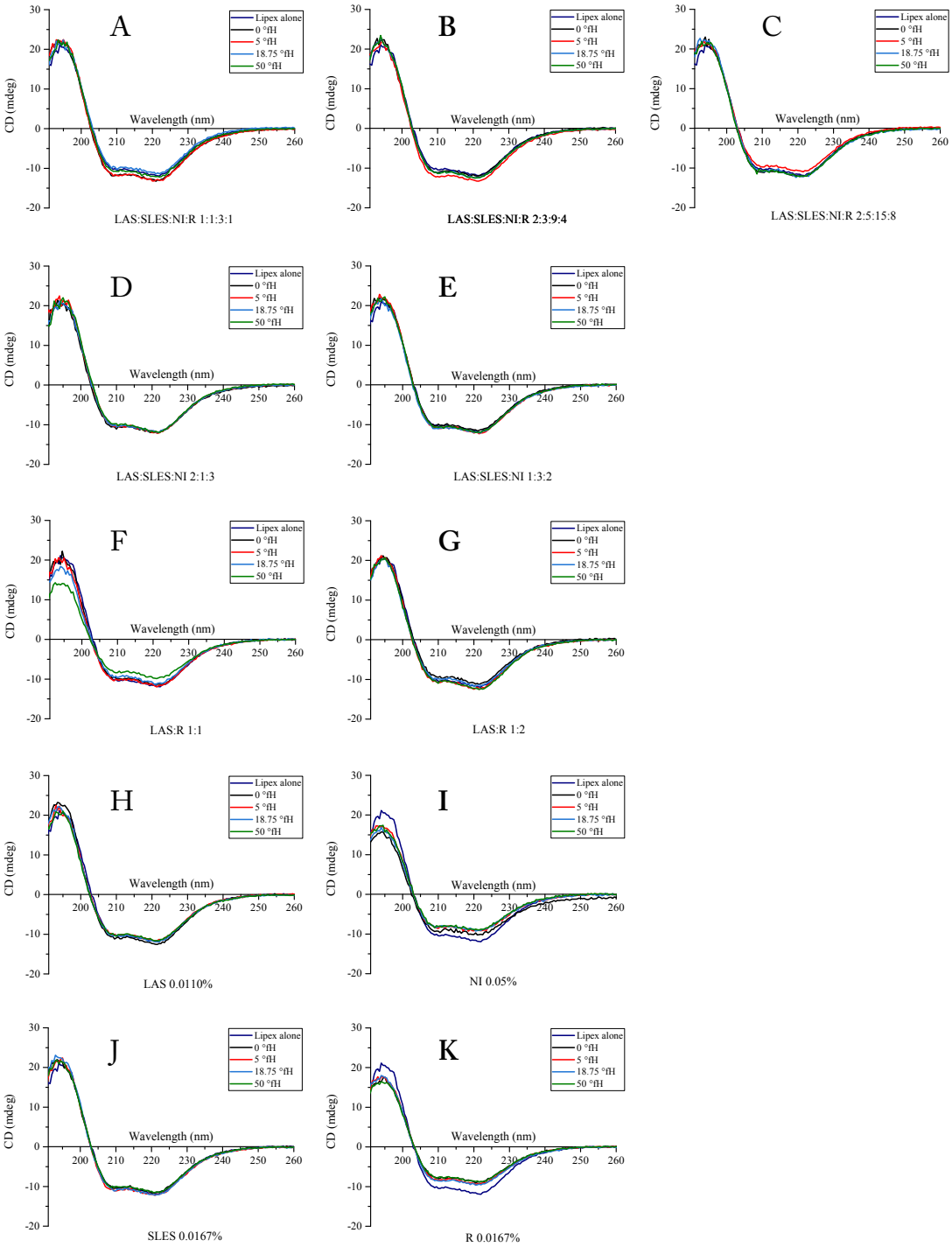


Figure 4.26 Far-UV CD spectra of Lipex at different surfactant conditions at increasing water hardness levels.

Each graph contains baseline corrected CD spectrum of Lipex as a control (n=3) and Lipex with surfactant at 4 different water hardness levels (n=1). Four water hardness levels were assessed: 0 degrees French hardness (Tris-HCl), 5 degrees French hardness (20 mg/L $CaCl_2$, 10 mg/L $MgCl_2$), 18.75 degrees French hardness (75 mg/L $CaCl_2$, 50 mg/L $MgCl_2$) and 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$).

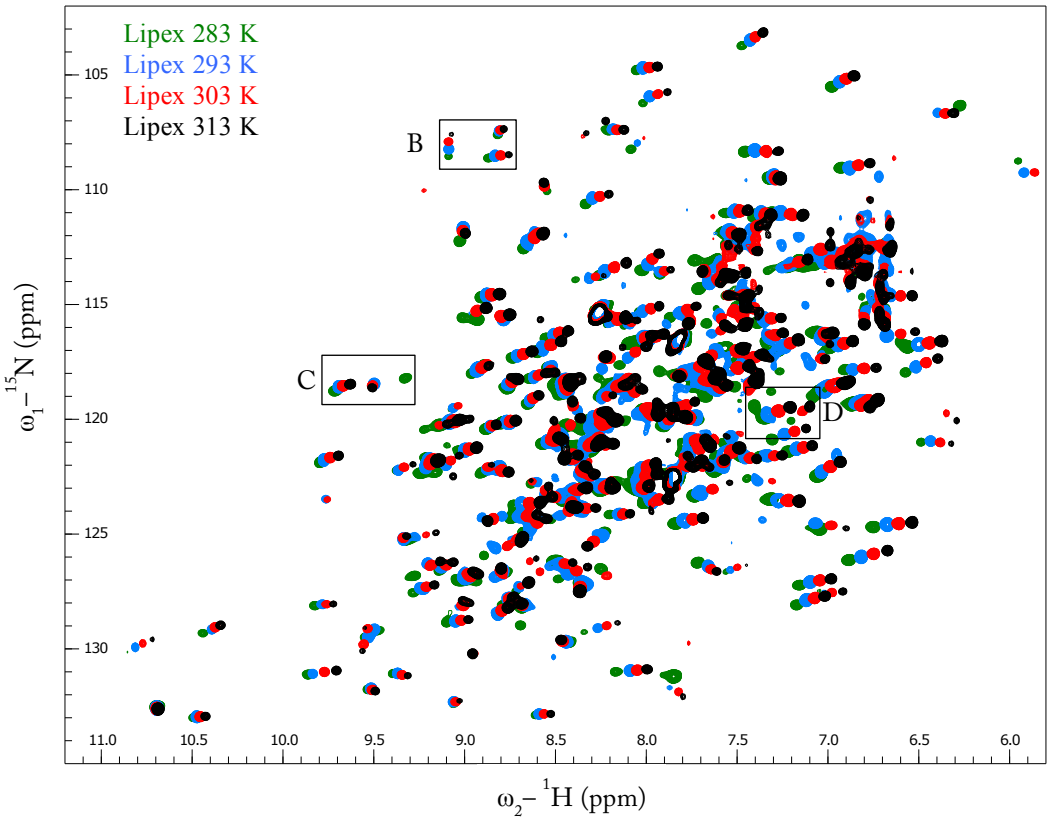
There are slight changes noted in the CD spectrum of Lipex when added to certain surfactants (Figure 4.26 F, I and K). The majority of conditions do not illict any change in Lipex CD spectrum (Figure 4.26 A, B, C, D, E, G, H, J). Changes in α-helicity are identified by measuring the minima detected at 208 and 222 nm λ, for condition LAS:R 1:1 (Figure 4.26 F) a decrease in α-helicity is detected, as a function of increasing calcium. There is a shift noted for NI and R individually at 0.05% (Figure 4.26 I) and 0.0167% (Figure 4.26 K) respectively, the reduction in signal may be due to loss of α-helicity or secondary structure elements. The change in α-helicity may be linked to the lid domain of Lipex, where the lid domain is fixed and reduces solvent access to the catalytic domain; but changes in CD spectra cannot be assigned to specific regions of the tertiary structure. For LAS:R 1:1, the increases in activity may be due structural changes in the lid domain, increasing the accessibility of the active site by displacing the helical lid, thereby increasing lipolytic activity. It is interesting that for the high active conditions of LAS:SLES:NI:R in 2:3:9:4 at 50 °fH there is no change in CD spectra or α-helicity content. There may be a different mechanism between how LAS:R at 1:1 and LAS:SLES:NI:R in 2:3:9:4, both increase catalytic activity at 50 °fH.

At concentrations up to 0.01% it is apparent that the surfactants have no significant individual effect on the secondary structure of Lipex. Through CD analysis, however, it is not clear if any changes in the helical lid domain are experienced through addition of surfactants. CD spectra were not collected for samples containing individual surfactants at greater than 0.01%, these conditions analysed by 2D ^{15}N-HSQC NMR indicate secondary and tertiary structural unfolding of the enzyme.

CD analysis does not provide residue level information on the structure of proteins. For more detailed results on enzyme structure and conformation NMR is required. The following sub-sections present NMR derived results to determine structural and conformational changes in Lipex against industrially relevant variables such as temperature, calcium and surfactants. Understanding the role of surfactants and calcium and defining a conformational state of a highly active formulation is attempted. Since we cannot screen all conditions at all variables, we selected formulations that exhibited low, high or increasing activity as a function of water hardness.

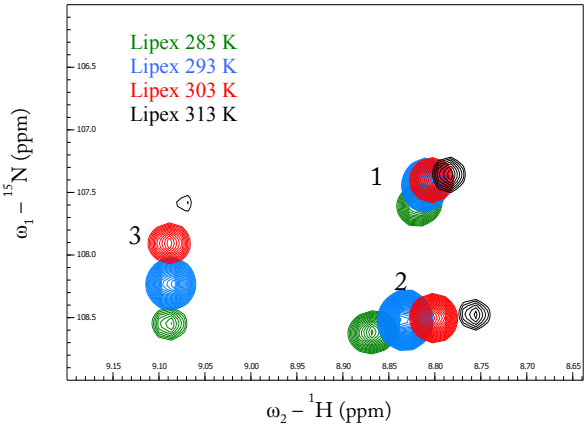
4.3.7 Effect of temperature on Lipex conformation measured by NMR

Lipex is a mildly thermophilic enzyme; temperature has a significant effect on the activity of the lipase, increasing activity as a function of temperature. Dynamical and transient conformational states of the enzyme, which can be studied by NMR, may elucidate how its thermophilic properties are exhibited. It is therefore of interest to understand these mechanisms, potentially providing crucial information to aid the directed evolution of the next generation of enzymes. There is an industrial relevance to studying the effects of temperature on Lipex since laundry can be performed at a range of different temperatures. One long-term goal of Unilever is to generate enzymes and formulations capable for maximum performance at lower temperatures. This offers improvements to the environmental and economical sustainability of laundry tablet production and usage.

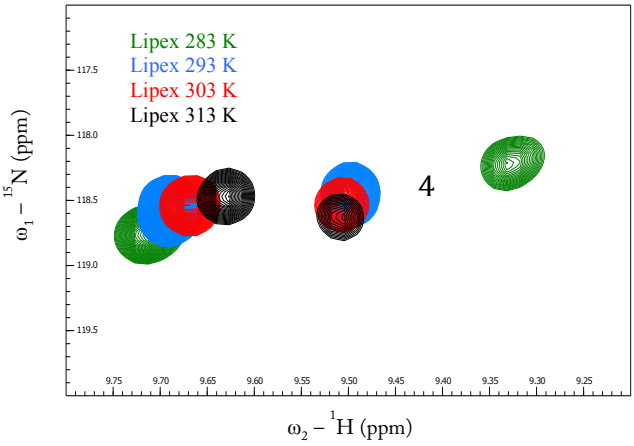


(A)

(B)



(C)



(D)

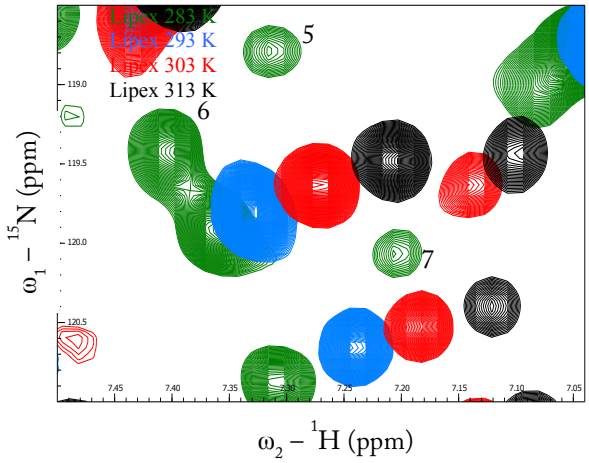


Figure 4.27 2D ^{15}N TROSY-HSQC overlay of Lipex at a range of temperatures

Lipex at 283 K (blue), 293 K (red), 303 K (green) and 313 K (black). All spectra collected in 25 mM Tris-HCl, 20 mM NaCl, pH 8.0 on 800 MHz spectrometer

The local chemical environment of each residue can affect its chemical shift position detected in the 2D HSQC spectrum. Slight changes in the conformational state of the residue, influenced by factors such as the position and orientation of the residue, intermolecular bonding and changes in electronegativity and polarity in the local environment can lead to chemical shift perturbations in the NMR spectrum. The intensity of each peak may also provide information on the dynmical state of the residue, or provide indications if is involved in bonding. These changes in the HSQC spectrum can be used to infer conformational changes in the protein influenced by the solution conditions.

Figure 4.27 (B, C and D) show changes in peak intensities and chemical shifts across temperature. Peak 1 exhibits little change in intensity or chemical shift, where as peak 2 exhibits chemical shift changes usually observed in increasing temperature, and peak 3 the signal intensity is the same from 283 to 303 K, but is significantly reduced at 313 K. In Figure 4.27 (C) peak 4 experiences dramatic chemical shift perturbation, this significant change could be indicative of a change in conformation for the unassigned residue occurring from 293 K to 283 K. Peaks 5, 6 and 7 in Figure 4.27 (D) are observed only at 283 K. These results indicate a slight conformational change, at the residue level between 283 and 293 K, and between 303 and 313 K. These are not major structural changes but the conformational exchange and dynamics of certain residues are affected.

^{15}N T_2 relaxation data was collected for Lipex by performing ^{15}N-^{1}H TROSY-HSQC experiments with varying delay times (0, 20, 40, 60, 80, 100 and 120 ms). During the delay time transverse relaxation is detected and residues with short or long relaxation rates can be determined. The time constant values for the protein are indicative of the tumbling of the protein molecule in the chemical environment. As mobility increases the tumbling becomes slower, and the time constant values become shorter. Without an assignment the analysis is performed on the protein as a globular entity, noting changes in the overall trend, based on industrially relevant solution conditions (temperature, calcium and presence of surfactants) are discussed.

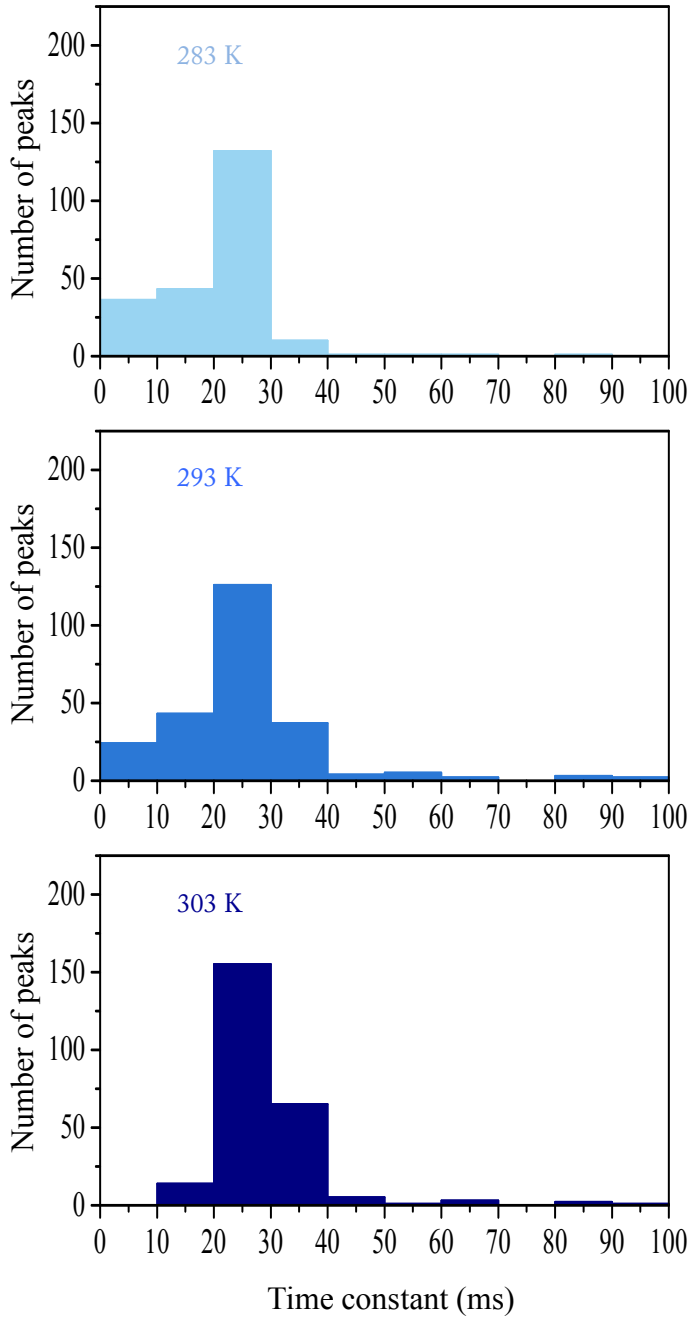


Figure 4.28 T_2 Relaxation measurements for Lipex as function of temperature

Histogram plots of the time constants of Lipex at 10 °C (light blue), 20 °C (blue) and 30 °C (navy)

2D ^{15}N-HSQC spectra and T_2 relaxation data for Lipex alone at different temperatures shows that new peaks are observed at reduced temperatures, and that time constant values increase as temperature increases (Figure 4.28). Whilst it would be expected that the general distribution of time constant values would increase as temperature increases, observing new peaks in the ^{15}N-HSQC spectra indicates that these peaks

may play a role in the thermal stability and activity of the protein. Reducing the temperature can slow the rotational correlation times and also reduce the conformational and chemical exchange of certain residues to a level where the exchange between dynamical states is reduced and can be detected. It may be that at lower temperatures these residues are in a more rigid, fixed position rather than freely mobile, enabling them to be detected by NMR. It is well known that conformational flexibility is a significant attribute of thermophilic proteins (Závodszky *et al*., 1998). The loss of activity of thermophilic enzymes at lower temperatures may be due to the reduced chemical exchange.

The peak shifts and new peaks observed in ^{15}N-HSQC spectra from reduced temperatures (Figure 4.27 A, B) are likely to be due a reduced level chemical exchange rather than reduced time constant values. Reduced conformational exchange of certain peaks, brought about by decreasing the thermal energy of the system, is most likely what has led to the detection of new peaks at 283 K. The new peaks identified at 283 K (Figure 4.27 (D)) would need to be assigned in order to identify the structural elements of the protein most susceptible to changes in temperature. The newly observed peaks might be from amino acids of the protein that undergo dynamical conformational changes at higher temperatures; only single or few amino acid changes are often required for thermal stability (Zhu *et al*., 2001). These peaks could be tryptophan residues, which are required for thermal stability (Zhu *et al*., 2001), possibly Trp89, 221 or 260 or a neighbouring basic residue with which it interacts. The tryptophan residues may be significant due to their hydrophobicity and ability to form π–cation interactions with the π-electrons in the side chain indol group and nearby basic amino acids. 3D data unfortunately cannot assign the amino acid spin system, since the Cα and Cβ peaks do not show an unambiguous amino acid spin system.

We decided to keep temperature a constant in all further experiments, since there is an extensive range of variables already present. All NMR experiments were performed at 303 K where Lipex gives good NMR signal and is an average temperature for laundry.

4.3.7 Effect of calcium on Lipex activity and conformation measured by NMR

Aside from the independent variable of temperature during laundry, another variable is the water hardness. Water hardness levels vary across the country and globe; the defining characteristic of water hardness is the level of calcium. The interplay between laundry enzymes such as Lipex, surfactants and the hardness of water during washing cycles is complex and multifactorial. Understanding the role of calcium in this system requires firstly observing the effects of calcium on Lipex alone, and then in addition with surfactants.

Calcium has a positive effect on Lipex activity, without any addition of surfactant. The activity of Lipex on pNp butyrate increases as a function of calcium (Figure 4.29). It is known that a number of mesophilic lipases exhibit increases in lipolytic activity in presence of calcium (Benzonana, 1968). The mechanism behind this is still not fully understood, but it is believed that calcium plays a role in aiding the removal of fatty acid products and increases activity in a pseudo-dependent manner. The positive divalent charge of the ion may be a driving role in the interplay between calcium and fatty acids, since monovalent ions such as sodium and potassium show no effect. Whilst the structures for TlL show no calcium-binding site (Skjold-Jørgensen *et al*., 2014), it is evident that calcium plays a role in the activity of Lipex, most likely in a pseudo-dependent manner.

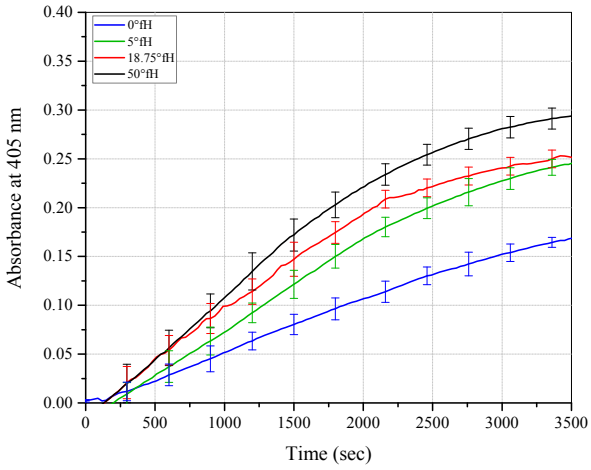
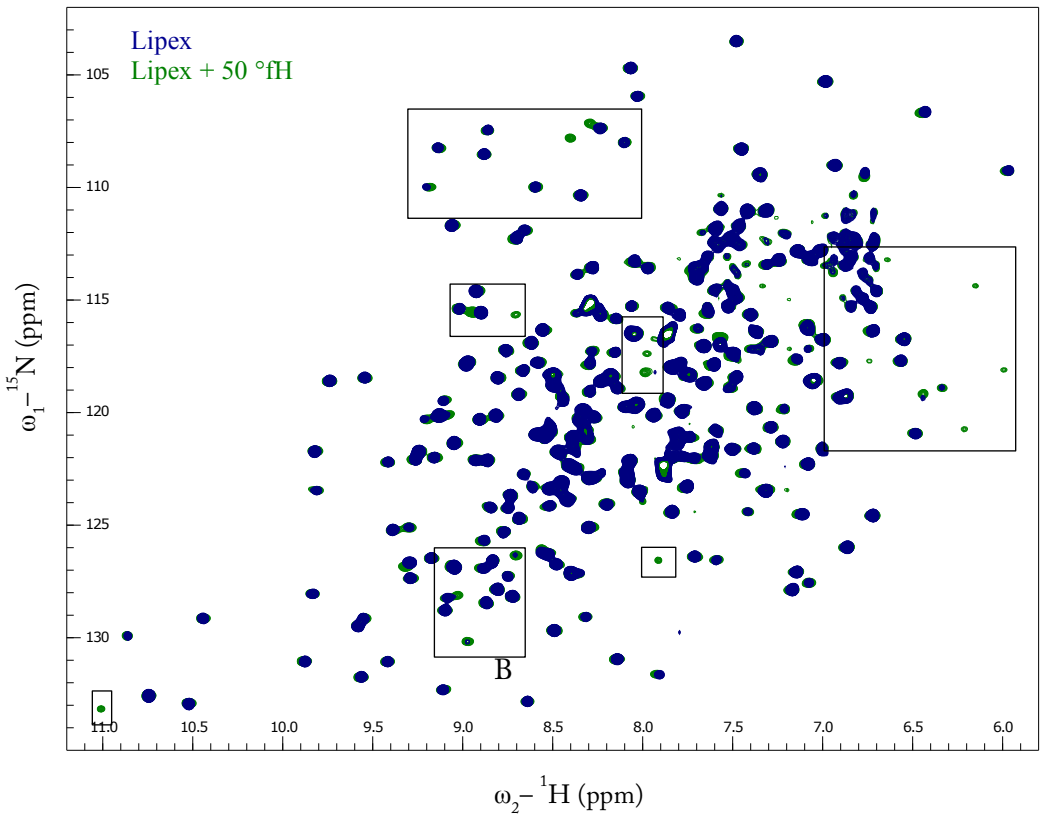


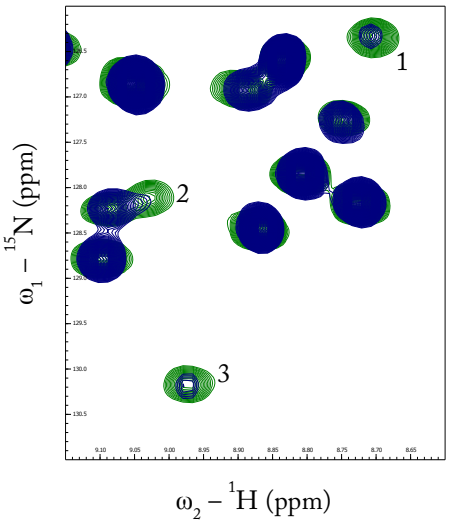
Figure 4.29 Effect on calcium on Lipex activity

Activity assay of Lipex on pNp butyrate at different water hardness levels at 30 °C. Four water hardness levels were assessed: 0 degrees French hardness (Tris-HCl), 5 degrees French hardness (20 mg/L $CaCl_2$, 10 mg/L $MgCl_2$), 18.75 degrees French hardness (75 mg/L $CaCl_2$, 50 mg/L $MgCl_2$) and 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$). Absorbance at 405 nm was recorded at 60 s intervals. The control without Lipex was subtracted from each sample reading, and a mean was taken (n=3).

^{15}N TROSY-HSQC NMR were performed on Lipex with 50 °fH calcium levels to identify any conformational changes in Lipex in the presence of calcium.



(A)



(B)

Figure 4.30 2D ^{15}N TROSY-HSQC overlay of Lipex in absence and presence of calcium at 50 °fH

(A) Lipex alone (dark blue) and Lipex at 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$) (green). (B) Selected region of (A)

The boxed regions show examples of new peaks observed at 50 degrees French hardness. Both spectra collected in 25 mM Tris-HCl, 20 mM NaCl, pH 8.0 at 303 K on 800 MHz spectrometer.

Changes are observed in the ^{15}N TROSY-HSQC spectra of Lipex in presence of calcium chloride (Figure 4.30). Peaks labelled 1, 2 and 3 in Figure 4.30 (B) indicate examples of peaks where the signal intensity increases upon addition of calcium. There is no calcium-binding domain within the Lipex structure, so these changes are not from calcium binding to the protein structure. It is possible the newly observed peaks (boxed in Figure 4.30 (A)) experience reduced chemical exchange in presence of calcium and therefore gives increase signal, possibly via an electrostatic interaction from the positively charged calcium ions. The changes noted show that calcium might affect the conformation of Lipex through either electrostatics or non-specific interactions, eliciting small, but functionally significant changes. These residues may therefore be of functional importance.

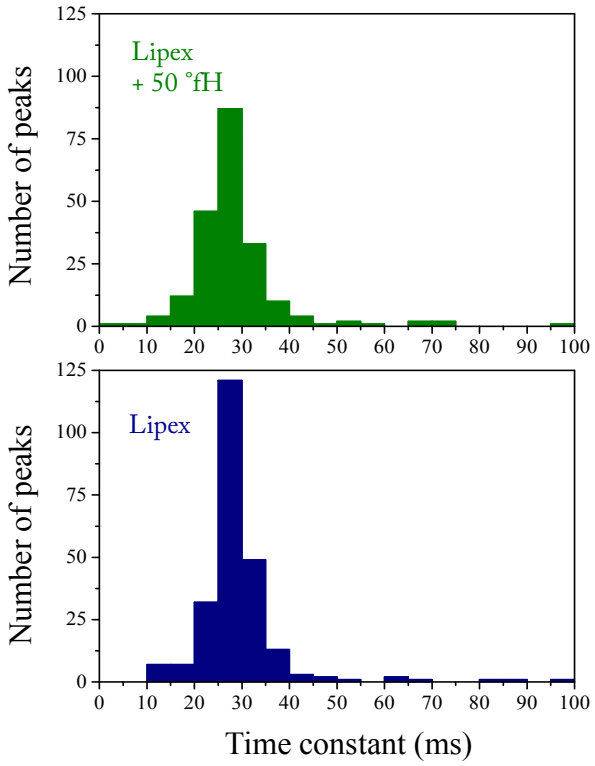


Figure 4.31 Effect of calcium on the T_2 relaxation of Lipex

Time constant histogram of Lipex in presence and absence of calcium at 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$) at 303 K.

Table 2. Time constant values (ms) for peaks which experience increases in signal intensity in presence of calcium

Peak number	Without calcium	With calcium
1	27.27	24.50
2	28.69	22.15
3	33.63	32.98

The presence of calcium at 50 °fH does not signifcantly alter the dynamics of Lipex. The increase in signal does not appear to be correlated to the dynamics. Some of the new peaks observed (Figure 4.30 (B)) have lower time constant values upon addition of calcium (Table 2, peak 1 & 2), however this is not consistent with all new peaks (Table 2, peak 3). This indicates that the presence of calcium on these residues (Table 2, peak 1 & 2) must be in their local dynamics rather than a global effect on protein tumbling (Figure 4.31) Without an assignment it is difficult to ascertain if the peaks that experience changes in dynamical properties are isolated to a specfic region in the structure, such as the surface, or if the phenomenon observed is effecting the global dynamics of the protein.

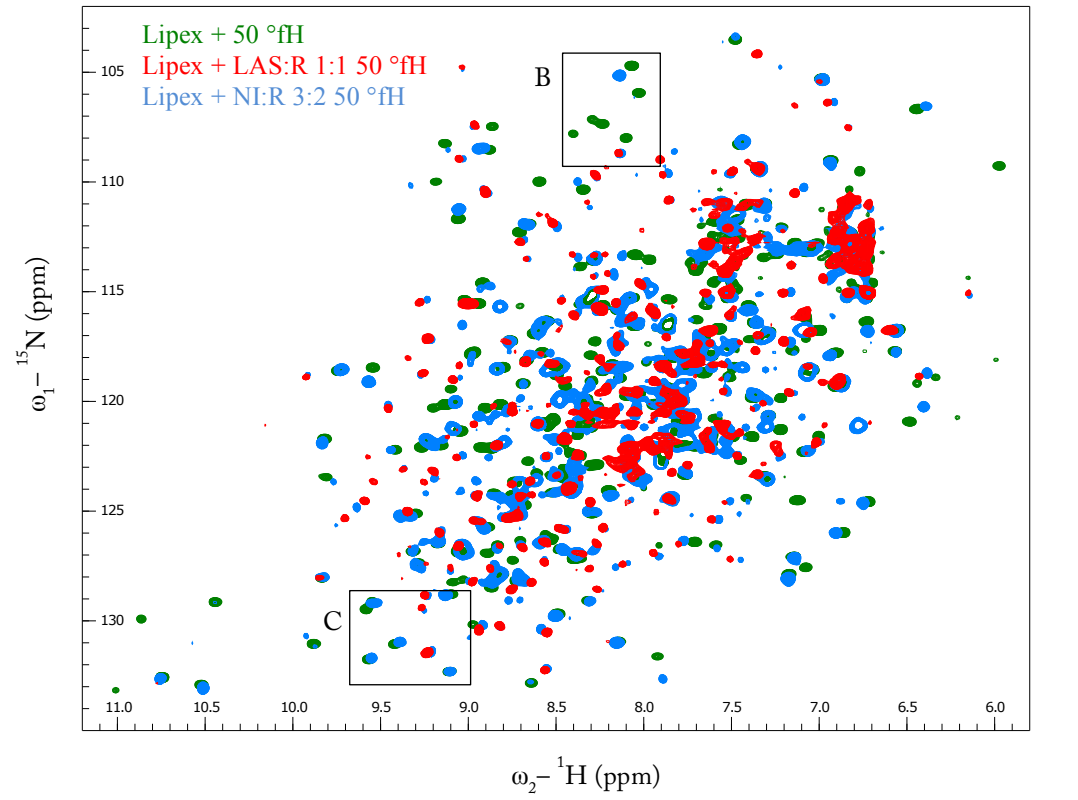
4.4 Conformation and dynamics of Lipex in formulations

4.4.1 Conformation and dynamics of Lipex in high active formulations

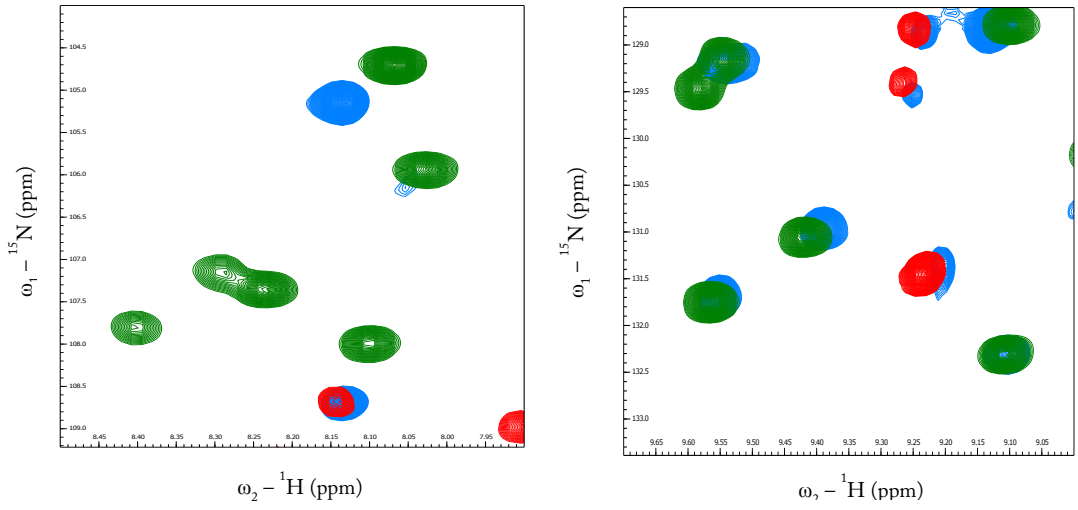
More complex systems were studied by introducing surfactant mixtures as an additional variable to the NMR sample. Initial investigations of surfactant mixtures were selected based on changes in Lipex enzyme activity. Formulations that resulted in increasing lipolytic activity as a function of water hardness from activity assays previously performed (LAS:R 1:1 Figure 4.22 F and NI:R 3:2 Figure 4.23 A-D) were selected as interesting systems to study the effects of both surfactant properties and calcium using NMR.

Studying surfactant interactions with Lipex with an unassigned backbone gives limited understanding and analysis. Analysis is restricted to observing changes in chemical shift perturbations and intensities, which can be observed in the ^{15}N

TROSY-HSQC spectra. These changes are indirect effect of surfactants and calcium and give inference to the conformation of Lipex; however the precise location of these changes on the Lipex structure remains elusive.



(A)



(B) *(C)*

Figure 4.32 2D ^{15}N TROSY-HSQC of Lipex in high active formulations

(A)

Lipex at 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$) (green)

Lipex + LAS:R 1:1 at 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$) (red)

Lipex + NI:R 3:2 at 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$) (light blue).

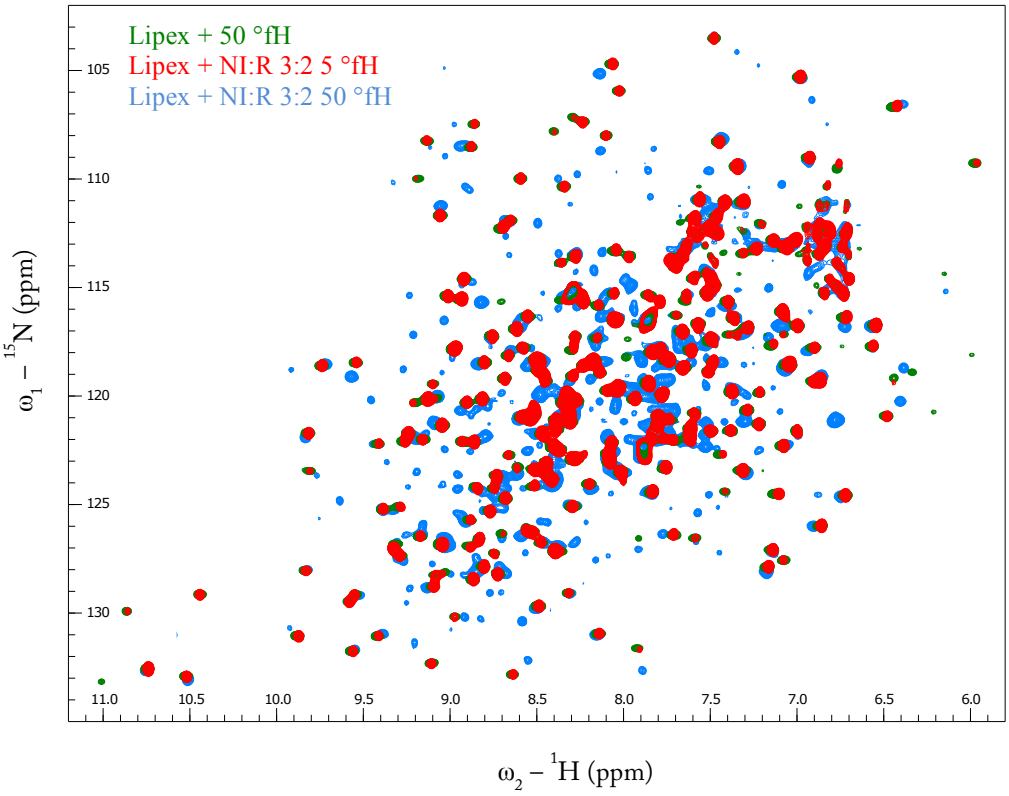
(B) and (C) Selected regions of (A)

Different chemical shift pertubations can be noted in both conditions. Indicating that mixtures of LAS:R and NI:R interact differently with Lipex, and the enzyme exhibits a different conformation in both samples. Lipex with 50 °fH gives 220 peaks, which increases to 237 upon addition of LAS:R 1:1 and increases further to 250 peaks with NI:R 3:2. The increase of peaks potentially exemplifies the different conformational stability of Lipex in presence of surfactants and calcium. A number of these newly observed peaks in presence of surfactants are peak pairs, which are likely to be the same residue but in a conformational exchange between two states; examples of these peaks are depicted in Figure 4.33 (B). Between the three spectra, 77 peaks overlay between Lipex with 50 °fH and Lipex with LAS:R 1:1 with 50 °fH with the majority of 160 peaks experiencing a shift. 50 peaks overlay between Lipex with 50 °fH and Lipex with NI:R 3:2 with 50 °fH with 200 peaks experiencing a shift. Comparing the two surfactant conditions LAS:R 1:1 and NI:R 3:2 both at 50 °fH there are 45 peaks that overlay, and 205 peaks which experience peak shifts. These results, whilst limited in their analysis, illustrates the significant effect surfactants and calcium can have on the conformational state of Lipex; furthermore different surfactant mixtures can have significantly different effects on Lipex.

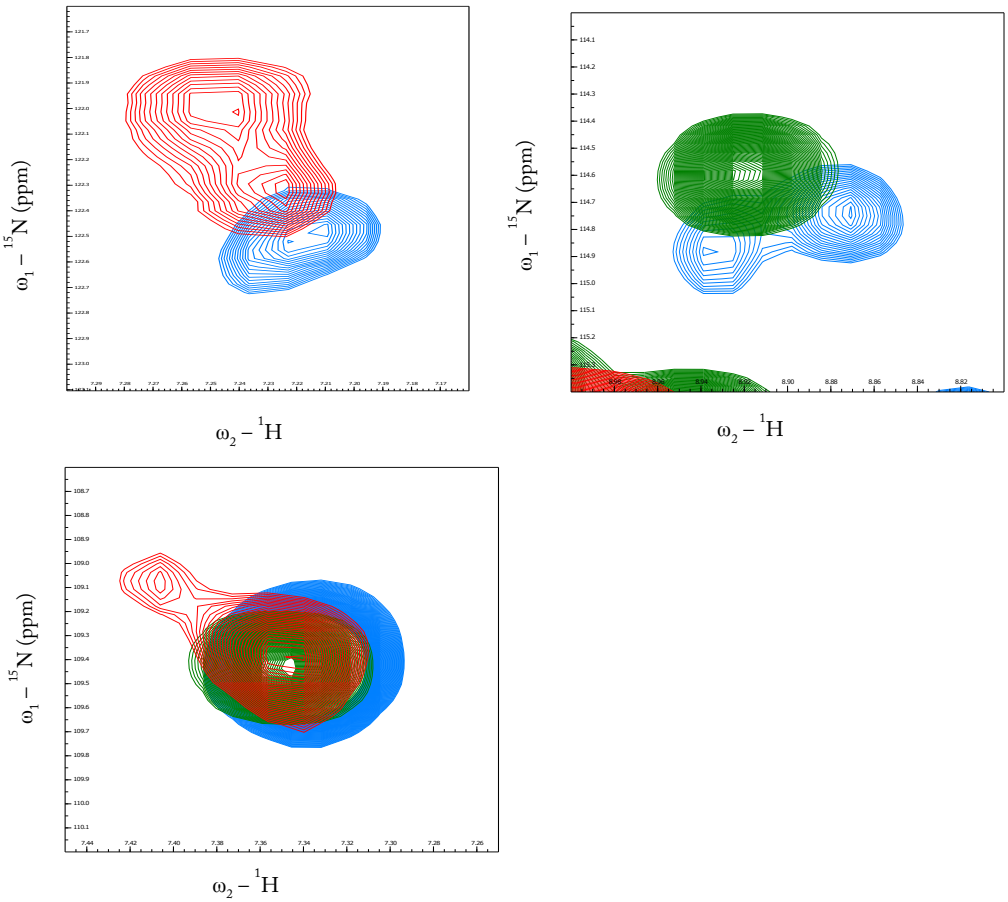
Some of newly observed peaks for NI:R 3:2 overlap with peaks observed for Lipex with calcium at 50 °fH as shown in Figure 4.32 (C) (light blue vs green peaks). These peaks are seen only for NI:R 3:2, not in the sample containing LAS:R 1:1. The effects of calcium on Lipex conformation are still observed with NI but not with LAS, which supports a role of LAS in chelating calcium ions, more so than NI. These peaks are likely to represent residues where calcium in the presence of surfactants either dependently or pseudo-dependently interacts with Lipex. The peaks lost after the introduction of surfactants in Figure 4.32 (B), may indicate regions of the enzyme where surfactants are interacting with Lipex, or affecting a change. It is interesting to note that spectral changes upon surfactant addition highlighted in Figure 4.32 (B) are located within a similar region of the spectrum upon addition of calcium (Figure 4.30 (A)). The two spectra exhibit differences and appear not to correlate to Lipex exhibiting a single high active conformation. This demonstrates there are further factors which define the enzyme activty.

4.4.2 Effect of calcium on Lipex conformation and dynamics in high and low active formulations

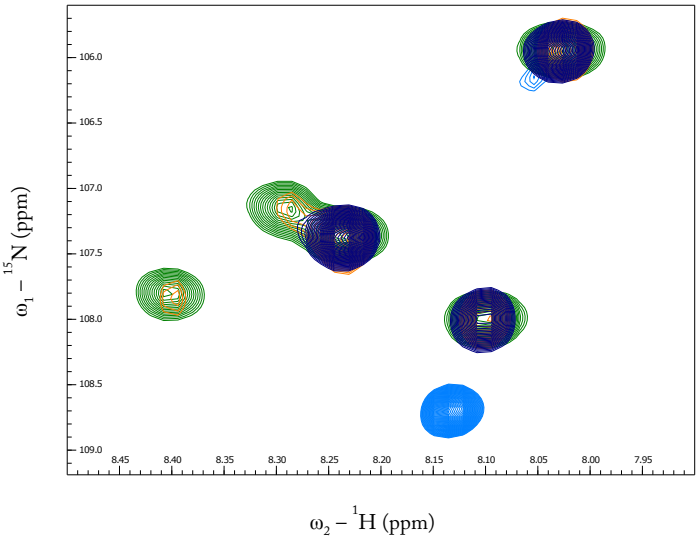
Lipex was studied with NI:R 3:2 at both 5 and 50 °fH to observe the effects of calcium on a sample containing Lipex and surfactants. Increasing calcium from 5 to 50 °fH increases the catalytic activity, therefore changes in NMR spectra may indicate how calcium affects the conformation of Lipex in presence of surfactants to give an enzymatically more active state.



(A)



(B)



(C)

Figure 4.33 2D ^{15}N TROSY-HSQC of 50 degrees French hardness with Lipex and with Lipex + surfactant mixture NI:R

(A) 2D ^{15}N- TROSY-HSQC overlay of

Lipex at 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$) (green)

Lipex + NI:R 3:2 at 5 degrees French hardness (20 mg/L $CaCl_2$, 10 mg/L $MgCl_2$) (red)
(B) Selected peaks of (A) showing peak pairs at both 5 and 50 degrees French hardness
(C) Selected region of 2D ^{15}N- TROSY-HSQC overlay of
Lipex alone (dark blue)
Lipex at 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$) (green)
Lipex + LAS:R 1:1 at 50 degrees French hardness (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$) (red)
Lipex + NI:R 3:2 at 50 degrees French hardness (20 mg/L $CaCl_2$, 10 mg/L $MgCl_2$) (light blue)
Lipex + NI:R 3:2 at 5 degrees French hardness (20 mg/L $CaCl_2$, 10 mg/L $MgCl_2$) (orange)
Spectra collected in 25 mM Tris-HCl, 20 mM NaCl, pH 8.0 at 303 K on 800 MHz spectrometer.

We observe 65 new peaks present from increasing the water hardness from 5 to 50 °fH for Lipex with NI:R 3:2 (Figure 4.33 (A)). We also observe ~15 peak shifts upon increasing water hardness from 5 to 50 °fH. These new peaks are not observed for Lipex alone, supporting a hypothesis that a particular conformation is stabilised. Many new peaks observed appear to be a pair, examples of which are illustrated in Figure 4.33 (B), and are likely to be the same residue but in two different slowly exchanging conformations. This would indicate that whilst the protein is stabilised, these residues are still dynamic and in chemical exchange and equilibrium between two states.

It is interesting that it is not the surfactant, but the level of calcium that leads to observing new peaks. At 5 °fH the spectrum is very different to Lipex alone, which is the effect of the surfactants. However, increasing calcium to 50 °fH dramatically changes the spectrum, indicating a role of calcium on surfactant + enzyme conformation at this water hardness. This indicates a significant role of calcium. As discussed further the dynamics of Lipex with the same surfactants but at different water hardness levels are altered. Increases in calcium can affect the states of surfactant micelles, the changes observed between 5 and 50 °fH may be due to Lipex interacting with a different form of micelle in terms of size, aggregation or dynamical characteristic (Sammalkorpi *et al.*, 2009). Further biophysical studies using techniques such as measuring surface tension would elucidate changes in the surfactant properties.

New peaks that are noted for Lipex in 50 °fH, but not present in Lipex alone, are weakly observable in Lipex with NI:R at 3:2 at 5 °fH, but are not detected when increasing the water hardness to 50 °fH (or Lipex + LAS:R 1:1 at 50 °fH) (Figure 4.33 (C)). Without surfactants these two peaks are visible at 50 °fH, when surfactants are added they are only visible at 5 °fH. Whilst a titration of calcium would provide more information, from the data at present it appears that for these peaks, the effect of calcium is removed when surfactants are added. This illustrates that for the conformational state of these residues, surfactants have a more dominant effect than calcium. These differences portray a complex interplay between calcium and surfactants. The changes in the conformational state of Lipex may have an effect on the relaxation properties and dynamics of the protein. To study this we performed T_2 relaxation measurements on both conditions.

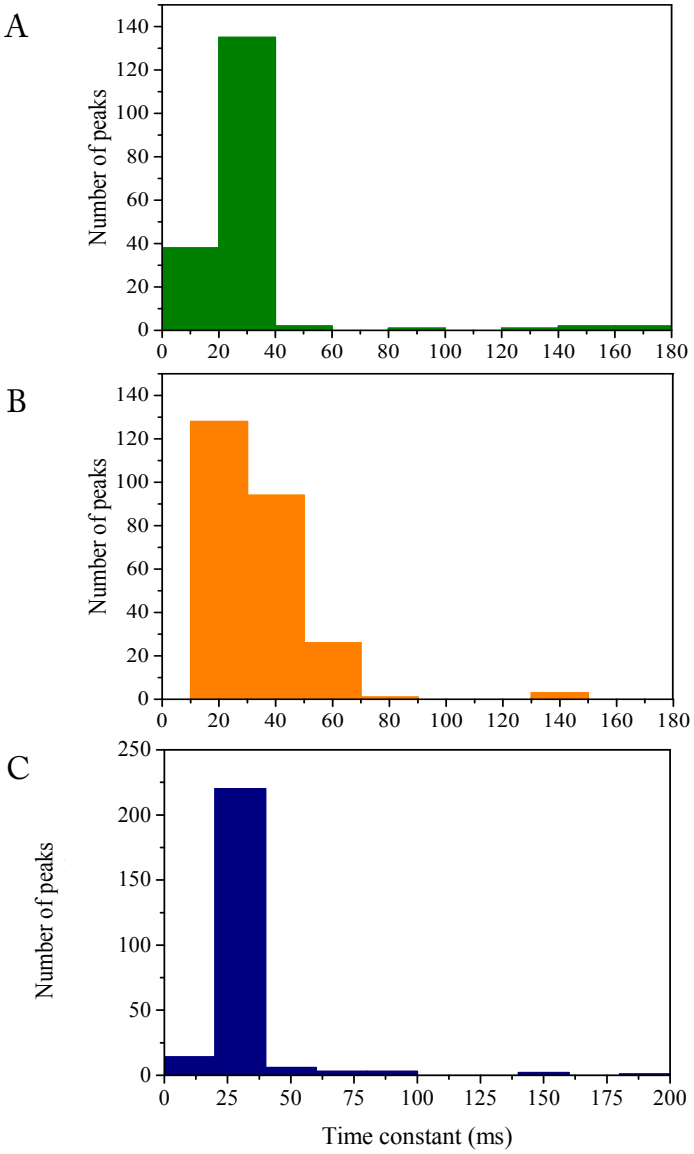


Figure 4.34 Changes in Lipex dynamics in low active and high active formulations

Histogram of T_2 time constant values for:

A) Lipex + NI:R 3:2 at 5 degrees French hardness (20 mg/L $CaCl_2$, 10 mg/L MgCl2) (green)

B) Lipex + NI:R 3:2 at 50 degrees French hardness (100 mg/L $CaCl_2$, 200 mg/L $MgCl_2$) (orange)

C) Lipex alone

The addition of calcium at 5 and 50 °fH has an interesting effect on the protein dynamics with surfactants NI:R 3:2. At 5 °fH (Figure 4.34 (A)) the distribution of T_2 time constants are similar to Lipex alone (Figure 4.34 (C)), but at 50 °fH, the distribution is shifted to higher time constant values (Figure 4.34 (B)). There is an

increased shift in the global trend and distribution of the time constant values from 5 to 50 °fH, this change in the global dynamical property may explain why we observe new peaks for Lipex with NI:R 3:2 at 50 °fH where the relaxation of the protein is much slower. When considering changes in tumbling rates of a protein in the presence of surfactants is important to be aware that the presence of surfactants in the sample may change the viscosity and surface tension of the sample, creating a change in environment. Changes observed in Figure 4.34 may due to the presence of calcium in the sample. Increases in calcium can affect the dynamics and structure of ionic surfactant micelles (Sammalkorpi *et al*., 2009), and may result in a change of dynamic behaviour when bound to Lipex. Multiple protein conformations have been sampled for TlL in presence of SDS by stopped-flow fluorescence experiments (Morgensen *et al*., 2005), confirming conformational ensembles are experienced between TlL and surfactants and provide support for the dynamical changes obtained through NMR.

The spectral changes in presence of surfactant, and more significantly the dynamical properties, between 5 and 50 °fH relate to an enzyme in a low and high active state respectively. The distribution of time constant values are similar between Lipex only and Lipex with NI:R 3:2 at 5 °fH. This distribution is shifted for Lipex with NI:R 3:2 at 50 °fH. Based on the distribution of T_2 time constants, it appears that highly active states are where the enzyme is less dynamic. Further investigations determining the T_1 and the rotational correlation times (τ_c) would provide clearer insights into the dynamical properties and the effects of surfactants and calcium on the molecular tumbling rates of Lipex. These experiments were not performed in the scope of the PhD due to time constraints. T_2 measurements were taken as an overview to provide an initial insight into which conditions would be interesting for further in-depth studies.

4.4.3 Studying Lipex activity with surfactants and increasing calcium levels

The composition of Lipex with LAS:R in a 1:1 ratio exhibits increasing activity at increasing water hardness levels (Figure 4.22 F). This provided a formulation containing surfactants with similar, or greater activity levels compared to Lipex only.

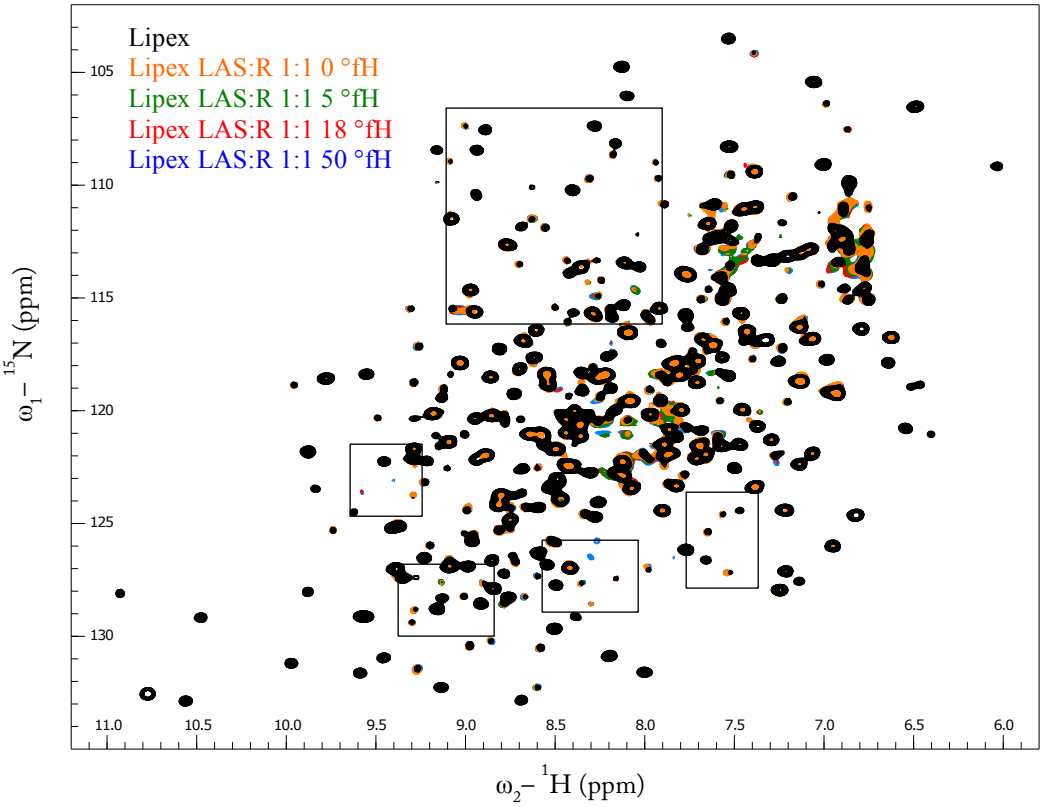


Figure 4.35 2D ^{15}N TROSY-HSQC overlay of Lipex with surfactants and increasing calcium

Lipex alone (black),

Lipex + LAS:R 1:1 at 0 French degrees (orange)

Lipex + LAS:R 1:1 at 5 French degrees (20 mg/L $CaCl_2$, 10 mg/L $MgCl_2$) (dark green)

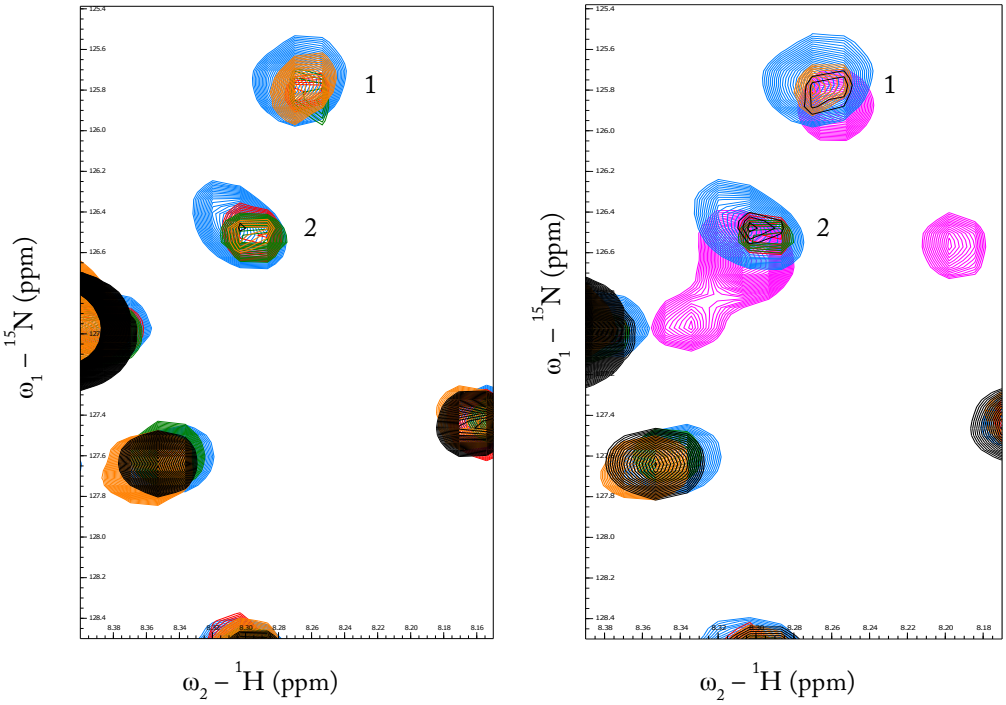
Lipex + LAS:R 1:1 at 18 French degrees (100 mg/L $CaCl_2$, 50 mg/L $MgCl_2$) (red)

Lipex + LAS:R 1:1 at 50 French degrees (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$) (light blue)

All spectra collected in 25 mM Tris-HCl, 20 mM NaCl, pH 8.0 at 303 K on 800 MHz spectrometer

It is interesting that for this formulation the spectrum does not significantly change between addition of surfactants, and again through addition of calcium as seen for NI:R 3:2. There are new peaks observed, highlighted in Figure 4.35. Additionally there are a few peaks, which exhibit increases in signal intensity, relating to either a less dynamic state of the residue or a reduction in the rotational correlation time. This signifies a crucial role dynamical property plays in the activity and function of Lipex, with surfactants and calcium. We do not observe the same increase in new peaks as for NI:R 3:2, and we observe more peaks at 0 and 5 °fH than at higher values. This implies a different mechanism for LAS:R than for NI:R, potentially through the

differences in how anionic and non-ionic surfactants interact with lipases, and a sequestering of calcium.



(A) *(B)*

Figure 4.36 Residues involved in binding to surfactants LAS

(A) Zoomed region of Figure 4.35 (B) Zoomed region of Figure 4.35 with overlay of Lipex + LAS at 0.0066%

Lipex alone (black)

Lipex + LAS:R 1:1 at 0 French degrees (orange)

Lipex + LAS:R 1:1 at 5 French degrees (20 mg/L $CaCl_2$, 10 mg/L $MgCl_2$) (dark green)

Lipex + LAS:R 1:1 at 18.75 French degrees (100 mg/L $CaCl_2$, 50 mg/L $MgCl_2$) (red)

Lipex + LAS:R 1:1 at 50 French degrees (200 mg/L $CaCl_2$, 100 mg/L $MgCl_2$) (light blue)

Lipex + LAS at 0.0066% (pink)

It is apparent that the peaks labelled 1 and 2 in Figure 4.36 (A) have greater signal intensity than for Lipex only samples, increasing the concentrations of calcium leads to greater signal intensity for these peaks. The addition of calcium increases the signal of peak 1 and 2, the stronger the signal the more active the lipase, thereby inferring a possible causality between these residues and Lipex catalytic activity. These peaks are also observed upon addition of LAS at 0.0066% (Figure 4.36 (B)), indicating that they are involved in binding to LAS. The fact that not all peaks noted in presence of

0.0066% LAS is reproduced with addition of LAS:R 1:1 implies that the peaks not observed may be involved in the activity of Lipex (

4.36 (B)). These results illustrate how it is difficult to determine a high active conformation based on ^{15}N TROSY-HSQC spectra alone. Peaks are observed, and lost between high active samples indicating a complex system in terms of enzymatic conformational ensembles.

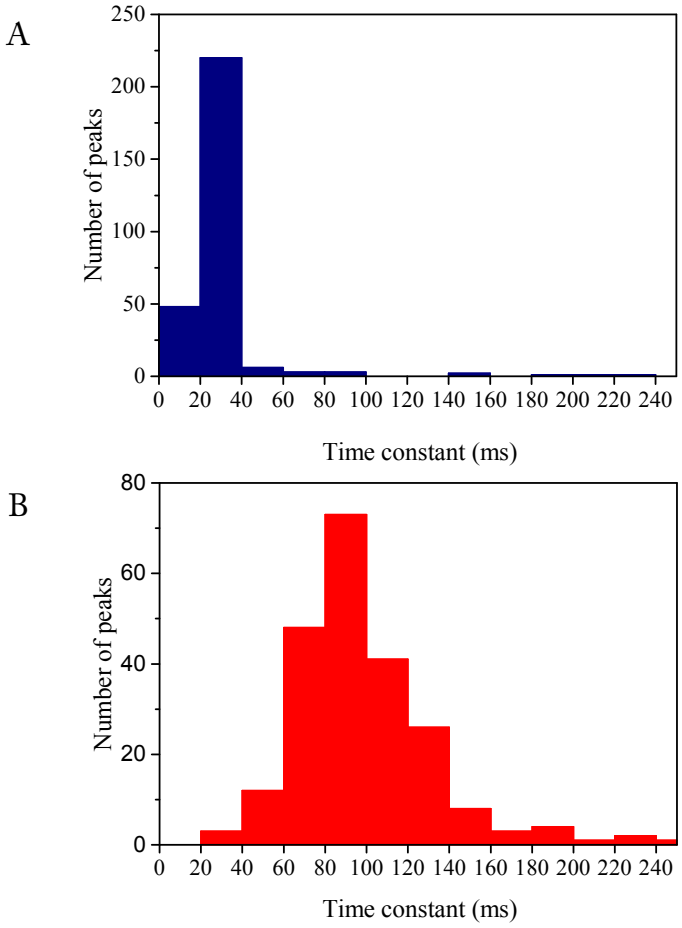


Figure 4.37 Effects of LAS:R 1:1 at 50 °fH on Lipex dynamics

Histogram plots of the time constants of Lipex at 30 °C (A) (navy) and with LAS:R 1:1 at 50 degrees French hardness (B) (red)

T_2 relaxation rates of Lipex LAS:R at 1:1 ratio at 50 °fH illustrate a longer T_2 time constant values, and therefore a faster tumbling protein than without surfactants, and slower than that for Lipex + NI:R 3:2 at 50 °fH, exemplifying a different state of the protein under different surfactant conditions. This may be due to a larger complex size based on different surfactants. DLS experiments were attempted to ascertain information pertaining to the complex size of Lipex with a range of surfactants

however the presence of lipids obscured measurement readings, Diffusion Orderded Spectroscopy (DOSY)-NMR could be an alternative and viable alternative technique, but due to time constraints was not investigated. Calcium ions are known to make anionic SDS micelles more compact and ordered, affecting the dynamical properties (Sammalkorpi *et al*., 2009). This may explain the shifts in dynamics between NI:R 3:2 and LAS:R 1:1 both at 50 °fH where LAS and R, both anionic, are more compact and ordered than NI and R.

The dynamics of Lipex in both NI:R 3:2 and LAS:R 1:1 at 50 °fH are shifted to longer time constant values, the increases in activity for these conditions is likely be within the dynamical changes of the protein. These sets of data imply that the dynamics of the system plays a significant role in the activity and functionality of the lipase. These results demonstrate how ^{15}N TROSY-HSQC spectra cannot infer activity, the spectra between LAS:R 1:1 and NI:R 3:2 both at 50 °fH are not similar yet both are highly active. It has previously been suggested that the dynamics of the protein measured by fluorescence spectroscopy, specifically the active site lid may influence binding affinities and be significant in activity (Peters *et al*., 2002). The activity of the protein is intimately linked to its dynamical properties; the lid domain displacement and the orientation of the lipase in aqueous media during interfacial activation are both dynamical processes critical in lipase activity (Peters *et al*., 2002). This supports the theory presented here, that the overall relaxation properties, linked to shift in distribution of T_2 time constant values of the protein may have a role in the activity of the enzyme.

These results are based on unassinged residues, but NMR can detect changes on a residual scale, something which is not detectable using techniques such as ITC or CD, nor would the dynamcial data be accessible from crystallography, it within the scope of these results that exemplify and validate the choice of using NMR to study this system. Further complimentary experiments such as ITC would provide data for strength of interactions between surfactant mixtures and Lipex. This would provide supporting information relating to which surfactant is causing changes detected in the ^{15}N TROSY-HSQC NMR when surfactant mixtures are studied. Performing an interfacial activation assay where adsorption of the lipase is measured on a microtiter plate, a method initially proposed by Martinelle *et al*., 1995 may provide more insight

into whether interfacial activation and adsorption rates are, presumably, being affected by the increased addition of calcium. These experiments were considered, however the scope of our work alongside Unilever is primarily focused on NMR.

Without an assigned ^{15}N HSQC spectra it is difficult to ascertain confidently if new peaks and peaks experiencing chemical shift perturbation are from residues involved in catalysis. The results presented provide a preliminary view to how surfactants and calcium interact with Lipex, causing changes to the conformation of the enzyme detected by 2D ^{15}N TROSY-HSQC NMR. Whilst these changes cannot infer a high active state they can highlight certain residues involved in surfactant binding and potentially enzyme functionality. Additionally the relaxation properties of the sample indicate that slower relaxation rates indicate a more active enzyme.

4.5 Surfactant induced unfolding of Lipex

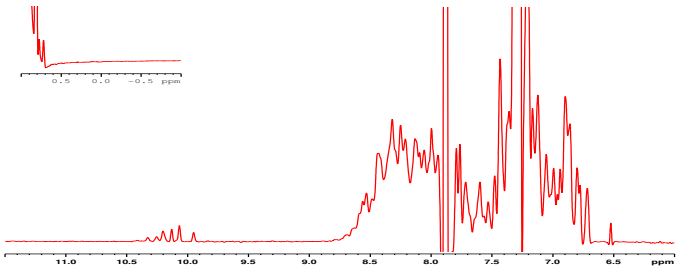
The following section covers the unfolding of Lipex under industrially and commercially used conditions. Since surfactants and detergent enzymes can interact in aqueous solutions during storage in liquid detergents it is crucial that surfactants within the formulations are selected carefully. Deactivating surfactants such as SDS (Mogensen *et al*., 2005) can have a deleterious effect, increasing the rate of enzyme deactivation especially in the presence of water. This has a devastating effect on reducing the shelf life of the product. Preliminary investigations into determining the cause of unfolding and the most deleterious components are presented.

There are two states of the enzyme, which are industrially relevant for studying: under storage and under wash conditions. At storage the enzyme is contained in a concentrated and viscous media with high (~20% v/v) levels of surfactants in formulations designed by Unilever, along with numerous stabilising chemical additives such as glycerol and sorbitol. This condition is difficult to study through NMR since the high concentration of surfactants leads to unfolding, line broadening and potentially aggregation of the protein, as discussed further. During laundry the sample is diluted by a factor of ~570 from the addition of ~20 litres of water in a standard washing cycle (Unilever; personal communication). The dilution brings the surfactants to a working concentration of ~0.001% (v/v) which is suitable for NMR

studies, and is therefore where most of the work presented previously has been focused.

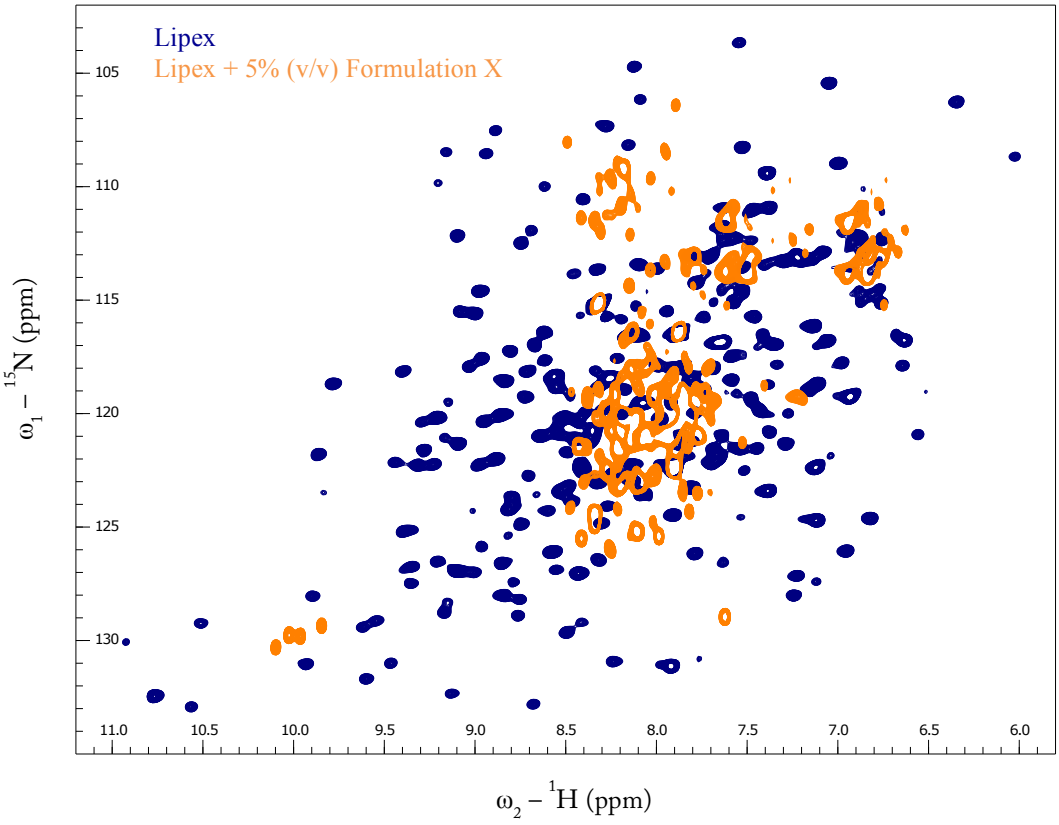
4.5.1 Lipex unfolded under storage conditions

Formulations with varying percentages of surfactants were specifically designed by Unilever to be tested by NMR to understand the changes caused to the Lipex structure by the presence of surfactants. Aside from stating that LAS, SLES and NI surfactants are used in these formulations, to avoid disclosing industrially sensitive material the specific ratios and additional components of the formulations have not been disclosed. Within the time frame two formulations (Formulation X and Y) were studied to determine the effects of high surfactant percentage during storage on Lipex.



^{1}H (ppm)

(A)



(B)

Figure 4.38 NMR observed unfolding of Lipex in Formulation X

(A) 1D ^{1}H NMR spectrum of Lipex with 5% (v/v) Formulation X. Selected amide region from 6.5 to 11.0 ppm (and 0 ppm inlay)

(B) 2D ^{15}N TROSY-HSQC overlay of Lipex alone (dark blue) and Lipex with 5% (v/v) Formulation X.

Spectra collected in 25 mM Tris-HCl, 20 mM NaCl, pH 8.0 at 323 K on 800 MHz spectrometer

The 1D ^{1}H NMR spectrum indicates that the protein is unfolded in 5% (v/v) Formulation X. The loss of folded methyl peaks at positions less than 0 ppm (Inlay Figure 4.38 (A)), and loss of dispersion with the collapse of peaks within the amide region to a narrow band between 7.5 and 8.5 ppm is indicative of an unfolded state.

The sample was prepared diluted at 95% than what is currently used in industry; the results presented are a fraction of the eventual effect. It is not known if the tertiary structure can be refolded and activity is restored after dialysis and dilution of the surfactants from 20% to 0.001%.

The ^{15}N TROSY-HSQC spectra of Lipex in 5% Formulation X Figure 4.38 (B) depicts an unfolded conformation. When the tertiary structure collapses the protein is unfolded, at this state the protein is not in a well defined structure and the residues are not in their own local chemical environment which gives rise the good dispersion of peaks noted for Lipex in its natively folded state. Residues experiencing the same chemical environment from the lack of defined secondary and tertiary structure causes the collapse of peaks to a narrow band. There may be residual secondary structure remaining that CD could be more acute to detecting. These results were unobserved in any combination of surfactants previously studied, but perhaps the surfactant concentrations tested previously were not great enough to illicit this effect. Formulation Y comprises the same surfactants but in a reduced total surfactant ratio compared to Formulation X.

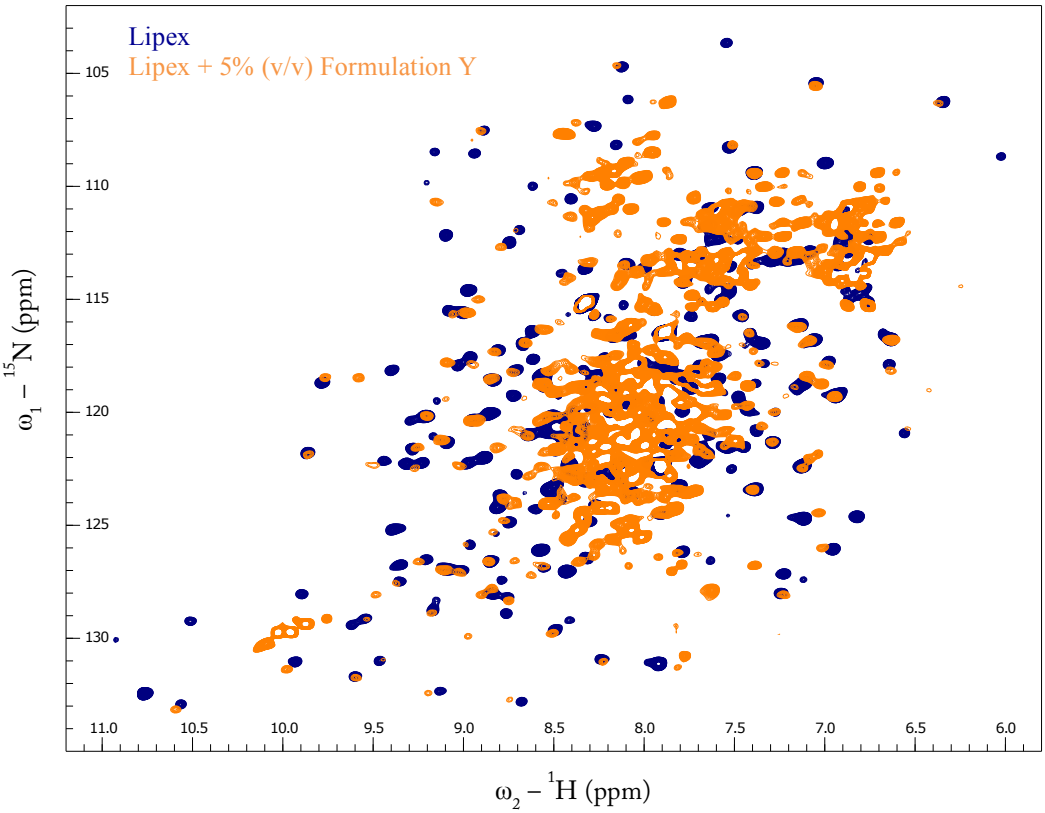


Figure 4.39 2D ^{15}N TROSY-HSQC overlay of Lipex in absence and presence of 5% (v/v) formulation Y

Lipex alone (dark blue) and Lipex with 5% (v/v) formulation Y (details undisclosed).

Spectra collected in 25 mM Tris-HCl, 20 mM NaCl, pH 8.0 at 323 K on 800 MHz spectrometer

Reducing the levels of surfactants reduces the degree of unfolding, and peaks corresponding to both natively folded and unfolded states are observed for Lipex + 5% (v/v) Formulation Y. Faint peaks which overlay to the native structure are observed. This depicts an enzyme in both unfolded and native conformation. The natively folded peaks are less intense indicating that the native conformation is less populated than unfolded.

The activity of the enzyme is compromised in an unfolded state, and offers no usage to an industrial application. Whilst most proteins are able to unfold and refold reversibly, the time frame for this to be achievable, and retaining activity upon refolding depends on the protein and factors such as its kinetic stability. The longer the protein is left in an unfolded state, the higher the propensity for aggregation to occur. The effect of irreversible aggregation would be deleterious to detergent usage, with no gain of alternative function as an aggregate. At the washing condition Formulation X does not inhibit Lipex activity. Formulation X at the percentage that would be present in laundry conditions (0.0001%), gives activity comparable to Lipex alone Figure 4.40). Since the activity is present under washing conditions, the defining factor is how much enzyme is refolded after being stored in 95% Formulation X/Y.

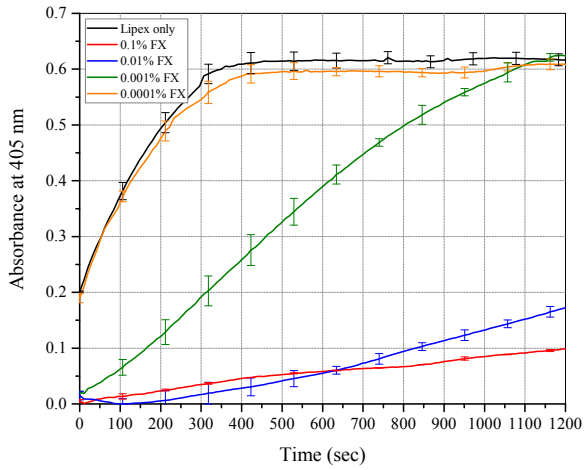


Figure 4.40 Activity assay of Lipex with increasing concentrations of Formulation X (FX)

Activity assay of Lipex with increasing percentages of Formulation X (FX) on pNp butyrate at 30 °C. Five increasing percentages of FX were assessed: 0.1, 0.01, 0.001, 0.0001%. Absorbance at 405 nm was recorded at 60 s intervals. The control without Lipex was subtracted from each sample reading, and a mean was taken (n=3). Lipex only sample is shown as a control.

It is the high kinetic stability of Lipex that leads to it remaining active in concentrations of surfactants that would entirely denature less stable proteins. This depends quite significantly on the thermodynamic and kinetic barriers between an unfolded denatured state and the native structure, across any transitional states. Small amounts of surfactants such as SDS are known to increase the thermodynamic barrier between denatured and transition states, stabilizing the denatured state (Wang *et al.*, 2013). Certain proteases and amylases used as detergent enzymes exhibit stability at 37 °C over 14 day incubation periods in surfactant blends of 30% non-ionic alcohol ethoxysulfonate (AE) 23-7. These percentages are similar to what is present in Formulations X and Y. However Formulation X and Y contain other surfactants than non-ionic ones such as anionic LAS and SLES, which have deleterious effects to enzyme structure. Increasing the levels of LAS with AE 23-7 reduces the stability and deactivates the protease and amylase enzymes. These effects were reduced upon addition of stabilizing compounds such as propylene glycol and sodium formate, with more success in formulations containing higher LAS ratios (Kravetz & Guin, 1985).

4.5.2 Effect of individual surfactants on Lipex structure

Understanding the mechanism, factors responsible for unfolding and its reversibility and effect on enzyme stability and functionality, is of significance to Unilever. Additionally the interactions of Lipex with surfactants that induce this structural collapse will enhance future formulation designs. Additives and stabilizers such as polyols and formate acid salts present in detergent formulations aid refolding of detergent enzymes (Kravetz & Guin, 1985), however the effects on Lipex are studied without additional variables for simplicity. To fully understand and determine the origin of unfolding, each surfactant should be studied individually to determine their interaction with Lipex. Each surfactant is studied at a monomeric state at 0.0066%, the percentage close to laundry wash conditions, increasing percentages of surfactant is added changes to Lipex conformation are measured as determined by ^{15}N TROSY-HSQC.

4.5.2.1 LAS

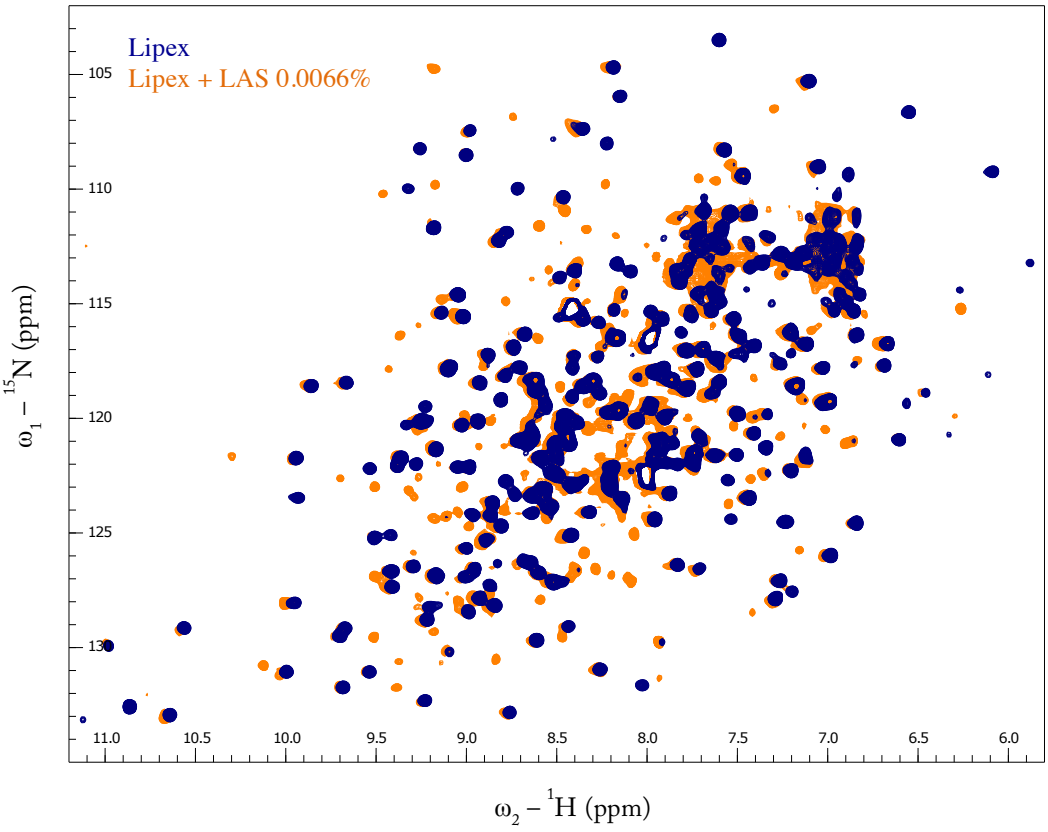


Figure 4.41 Lipex with LAS at 0.0066%

2D ^{15}N TROSY-HSQC overlay of Lipex alone (dark blue) and Lipex + 0.0066% LAS (v/v) (orange).

Spectra collected in 25 mM Tris-HCl, 20 mM NaCl, pH 8.0 at 323 K on 800 MHz spectrometer

The addition of 0.0066% LAS, is at a concentration significantly less than what is tested by Unilever (personal communication, Unilever), it is below its CMC value of 0.65g/L reported by Smulders 2002 and therefore in a monomeric state. Whilst these values are a factor of 10 less than what is present in the washing condition the results show differences between how surfactants in their monomeric interact with Lipex. The study of Lipex with monomeric surfactants is of importance; monomeric binding can lead to activation of lipases through non-interfacial mechanisms by inducing catalytically open conformations in the absence of an interface provided by miceller states (Bañó *et al*., 2003). A number of new peaks are observed that were previously

unobserved for Lipex alone; these new peaks are located across the both spectral dimensions. This indicates that these peaks are not localised to a particular section of sequence. Perhaps LAS in a monomeric form binds and stabilises the entire Lipex structure, or many LAS monomers could bind in multiple regions of Lipex. Slight conformational changes to a cutinase enzyme from *Humicula insolens* (HiC) are observed in addition of small amounts of anionic SDS surfactant, caused by two SDS monomers binding to the active site increasing the hydrophobic surface area (Nielsen *et al*., 2005). Given the similarities between lipases and cutinases it is likely a similar mechanism occurs with Lipex and LAS. There is limited published literature on the interactions of TlL with LAS, SLES or NI, most presumably since the information is industrially sensitive.

It was thought that the new peaks observed were arising from regions of the lipase involved in interfacial activation, or within the lid domain. The alkyl chains of LAS are likely to bind to the active site of Lipex because of their hydrophobic moieties; inducing a conformational change within lipase to facilitate substrate binding.

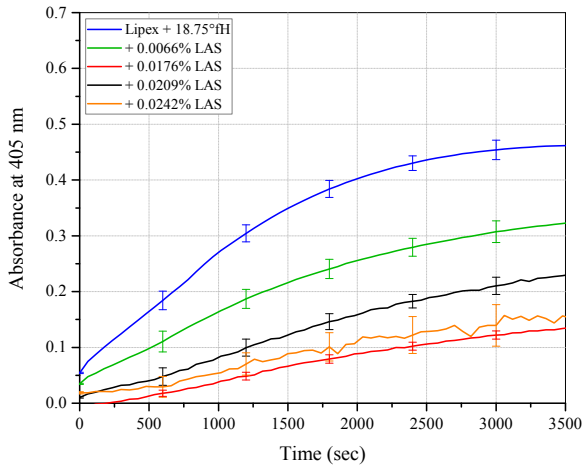
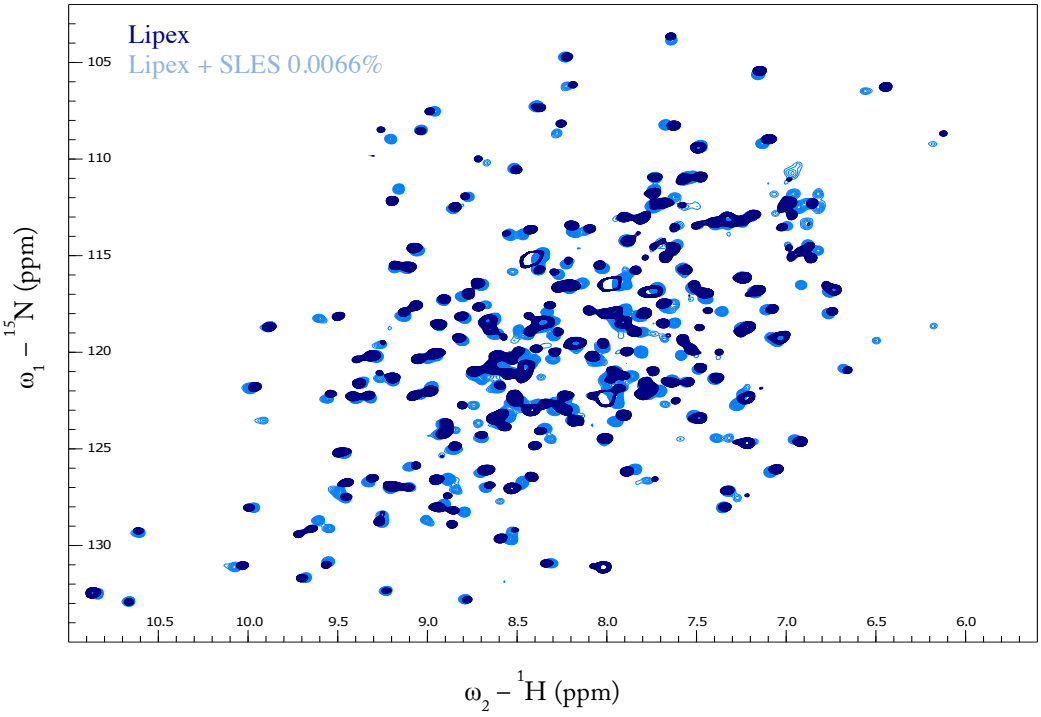


Figure 4.42 Activity assay of Lipex with increasing concentrations of LAS

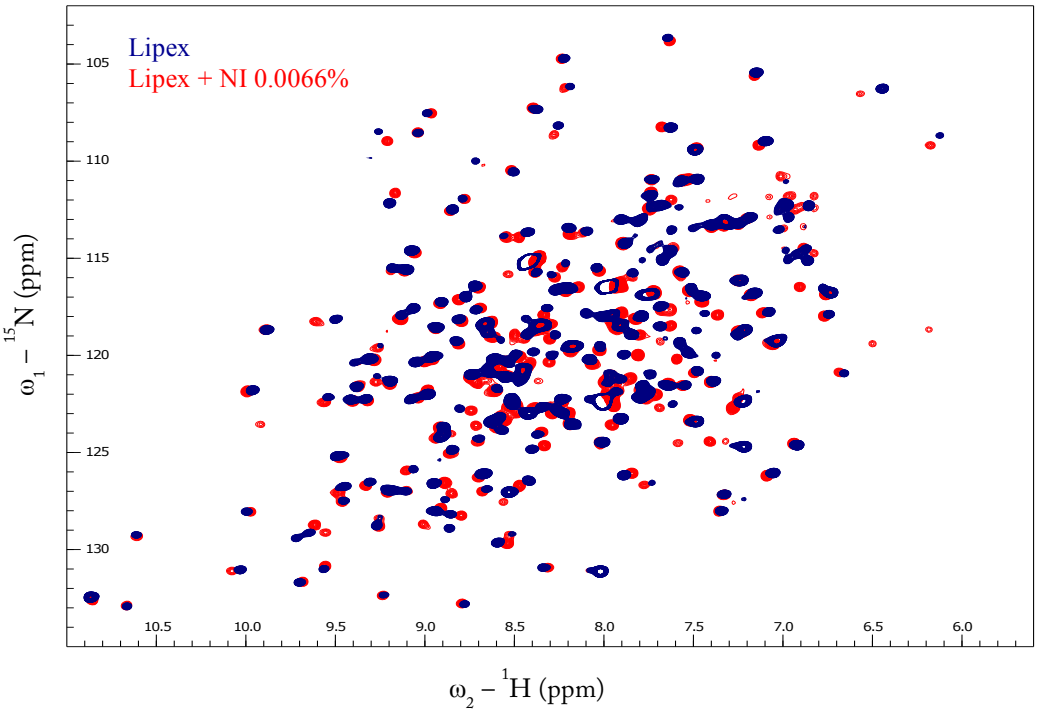
Activity assay of Lipex with increasing percentages of LAS on pNp butyrate at 18.75 degrees French hardness (75 mg/L $CaCl_2$, 50 mg/L $MgCl_2$) at 30 °C. Four increasing percentages of LAS were assessed: 0.0066, 0.0176, 0.0209 and 0.0242%. Absorbance at 405 nm was recorded at 60 s intervals. The control without Lipex was subtracted from each sample reading, and a mean was taken (n=3).

Monomer SDS binding to TlL causes a conformational change in the lid and shifts the lipase population to a completely open conformation (Mogensen *et al*., 2005). Upon LAS binding perhaps these residues have reduced mobility and dynamical properties and the enzyme is locked in a particular, presumably a non-active conformation. T_2 relaxation experiments might provide insight into this. The interaction of LAS with Lipex limits enzyme activity (Figure 4.42) displaying an inhibitory and destabilising effect. The decreasing activity with increasing LAS percentages may be due to monomer LAS binding to the Lipex at hot spot regions and forming premicellar aggregates which can occur below the CMC of surfactants, affecting the activity (Ibel *et al*., 1990).

4.5.2.2 SLES and NI



(A)



(B)

Figure 4.43 Lipex with NI and SLES at 0.0066% exhibit same conformation

2D ^{15}N TROSY-HSQC overlay of (A) Lipex alone (dark blue) and Lipex + 0.0066% SLES (light blue) and (B) Lipex alone (dark blue) and Lipex + 0.0066% NI (red).

Spectra collected in 25 mM Tris-HCl, 20 mM NaCl, pH 8.0 at 303 K on 800 MHz spectrometer

At 0.0066% the effects of SLES (Figure 4.43 (A)) and NI (Figure 4.43 (B)) are not the same as LAS (Figure 4.41). Interestingly the same changes in ^{15}N TROSY-HSQC spectra are observed from additions of equal percentages of NI and SLES to Lipex. We do not observe the same number of new peaks as for LAS. There are some newly observed peaks with SLES and NI, which are similarly observed with LAS. Many peaks change in chemical shift position and intensity when compared to Lipex alone, these chemical shift perturbations are consistent in both SLES and NI. These results indicate a different conformation of Lipex observed with LAS. Perhaps this is a result between the differences in anionic strength between LAS and SLES, which dictates the mode of binding and interaction to Lipex. Since the same spectra for SLES as for NI at 0.0066% is observed it is very likely they interact in the same manner to Lipex. At these concentrations it is surprising that a weakly anionic and a non-ionic surfactant bind via the same interaction. For NI we would expect not to observe a

drastic change in the conformation given that non-ionic surfactants do not extensively interact with lipase surfaces due to their lack of charge.

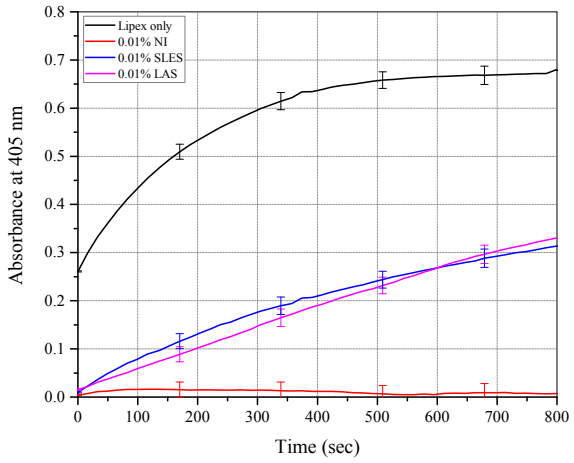


Figure 4.44 Activity assay of Lipex with anionic and non ionic mixtures

Activity assay of Lipex with surfactants at 0.01% in mixtures of SLES:LAS (red), SLES:NI (blue) and LAS:NI (pink) on pNp butyrate at 30 °C. Absorbance at 405 nm was recorded at 60 s intervals. The control without Lipex was subtracted from each sample reading, and a mean was taken (n=3). Lipex only (black) sample is shown as a control.

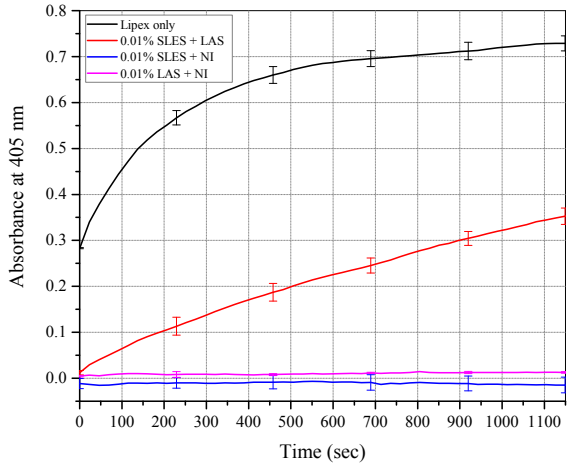


Figure 4.45 NI inhibits Lipex activity where it is retained for SLES and LAS

Activity assay of Lipex with surfactants at 0.01% NI (red), SLES (blue) and LAS (pink) on pNp butyrate at 30 °C. Absorbance at 405 nm was recorded at 60 s intervals. The control without Lipex was subtracted from each sample reading, and a mean was taken (n=3). Lipex only (black) sample is shown as a control.

Both LAS and SLES in the presence of NI cause a loss of lipolytic activity (Figure 4.44). Higher levels of NI surfactant can bind to the enzyme due to a lack of repulsive charges present for anionic surfactants, thereby inhibiting Lipex with a stronger efficiency by blocking the active site. These effects have been observed for TlL with non-ionic surfactants octyl β-glucoside and dodecyl β-maltoside (Mogensen *et al*., 2005). Additionally it is known that non-ionic surfactants bind via hydrophobic interactions and competitive adsorption. It is therefore likely that it will bind to the interface, or active site of Lipex, thereby reducing its capabilities for interfacial activation. Lipase activity with NI alone is significantly reduced and may provide supporting evidence for this (Figure 4.45).

4.5.3 Destabilising effects of surfactants on Lipex structure

A titration of individual surfactants was performed to determine the effects of increasing the surfactant concentration on Lipex and the NMR signal. Identifying the key surfactant interactions will aid designing formulations that reduce destabilisation of enzymes, not just specifically lipases but proteases and cellulases present in the laundry detergent. This will lead to maintaining the tertiary structure in both the storage and wash conditions, leading to higher levels of catalytic activity and product performance.

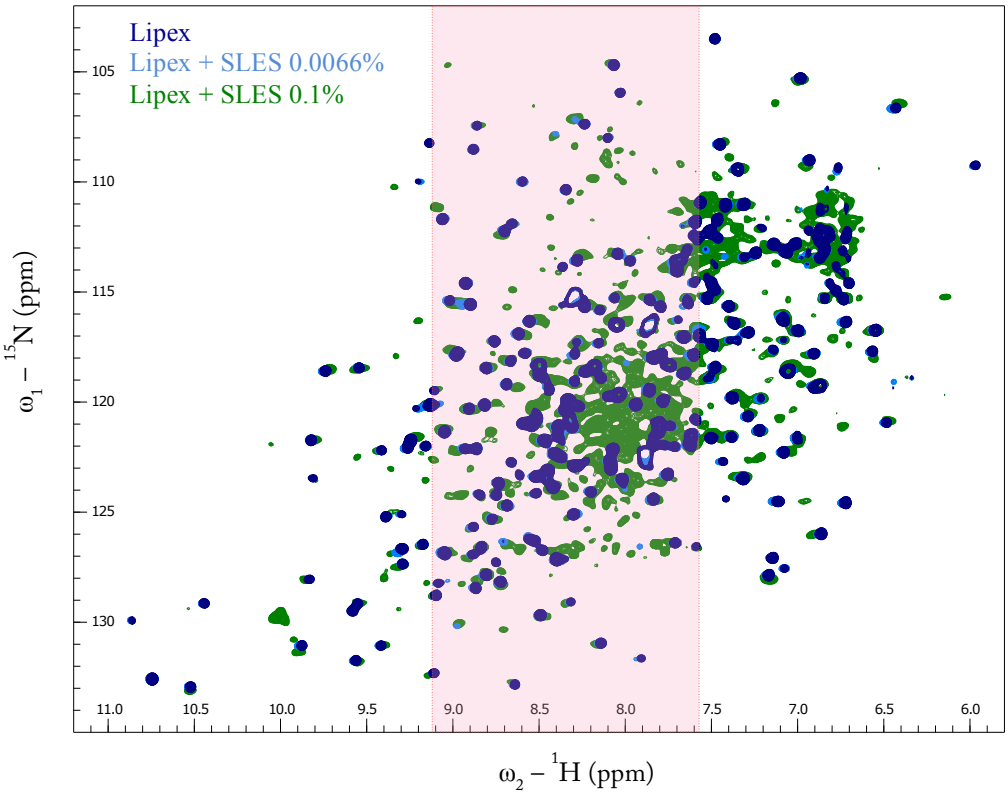


Figure 4.46 Effects of SLES on Lipex structure

2D ^{15}N TROSY-HSQC overlay of

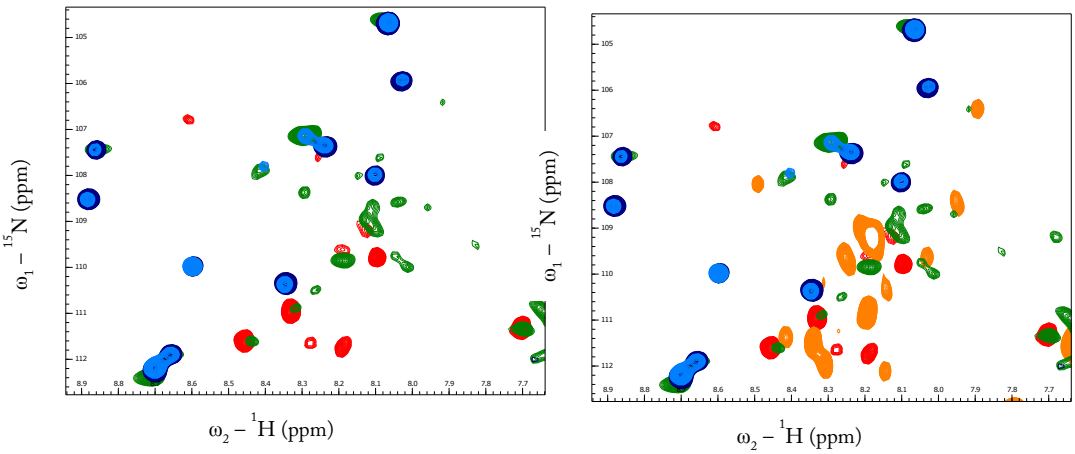
Lipex alone (dark blue)

Lipex + 0.0066% SLES (light blue)

Lipex + 0.1% SLES (green).

The band highlights region of spectra where peak collapse from well dispersed to a narrow band in unfolded proteins. All spectra collected in 25 mM Tris-HCl, 20 mM NaCl, pH 8.0 at 303 K on 800 MHz spectrometer

It was determined that at increasing percentages SLES and LAS had a major effect on the Lipex conformation, inducing partial unfolding and increasing line broadening or chemical exchange which may indicate aggregation. One aspect not studied due to time constraints was the effects of R on Lipex, which could replace SLES or LAS and maintain activity and increase Lipex stability in a native folded conformation.



(A) (B)

Figure 4.47 Partially unfolded conformation of Lipex under SLES and LAS

Selected region of 2D ^{15}N TROSY-HSQC overlay of

(A) *Lipex alone (dark blue)* *Lipex + 0.1% LAS (red)*
Lipex + 0.1% SLES (green) *Lipex + 0.1% NI (light blue)*

(B) *Lipex alone (dark blue)* *Lipex + 0.1% LAS (red)*
Lipex + 0.1% SLES (green) *Lipex + 0.1% NI (light blue)*
Lipex + 5% formulation 2 (orange).

Spectra collected in 25 mM Tris-HCl, 20 mM NaCl, pH 8.0 at 303 K on 800 MHz spectrometer

At 0.1% SLES and LAS begin to exhibit detrimental effects to the structure of Lipex, where partial unfolding is observed. At 0.1% SLES elements of protein unfolding are detected, not found in Lipex alone or with SLES at less than 0.1%. These new peaks are distinguished at regions at ~8 ppm in ^{1}H dimension, in particular those at ~109 ppm in ^{15}N dimension, are most likely glycine residues which characteristically have chemical shifts within this region of the ^{15}N-HSQC spectrum in unfolded proteins (Figure 4.47). The detection of peaks within this band is common of unfolded protein where the dispersion of peaks due to a natively folded structure is lost, with peaks collapsing to a narrow band around 8 ppm in ^{1}H (highlighted in Figure 4.46). The glycine peaks are noted with 5% Formulation X where the protein is entirely unfolded do not overlay with those observed for LAS and SLES, implying a different conformation, most likely complete and partially unfolded states (Figure 4.47 (B)). More peaks are noted for SLES; therefore it is believed SLES exhibits a stronger

denaturing effect on Lipex than LAS. These results indicate that both SLES and LAS play major roles in unfolding the protein. This is further supported since both Formulation X and Y contain only SLES and LAS surfactants (Figure 4.38 and 4.39). Interactions between Lipex and anionic surfactants are most likely mediated through charged side chains of Lipex, and this is likely to explain the differences in spectra observed in LAS and SLES with NI. Charged surfactants can disrupt crucial electrostatic interactions in enzyme structures leading to destabilization of tertiary structure (Nielsen *et al*., 2005). Anionic sulphated surfactants including LAS are known to strongly destabilize detergent enzymes (Kravetz & Guin, 1985). LAS and SLES have similar structural properties with a long hydrophilic carbon chain, and a hydrophobic charged sulphated head group. The large benzene ring of LAS most likely affects how it binds to Lipex, and possibly by steric hindrance has a different effect to SLES.

Both LAS and SLES interact with Lipex in a manner to aid collapse of tertiary structure through intermediate conformations visible in NMR. Destabilization is thought to be global across the structure, and confirmed by these results, since we do not observe specific regions of the HSQC collapsing, but a global change. The binding mechanisms and behaviour of these surfactants are still not entirely understood, it is known that surfactants such as LAS compete for hydrophobic binding of TlL (Sonesson *et al*., 2006) affecting the adsorption rates of the lipase, but the manner of binding and specific sites is still not known; a question that an assigned Lipex ^{15}N-HSQC would begin to answer. Since at 0.1% we begin to observe unfolding in SLES we studied each surfactant at this percentage.

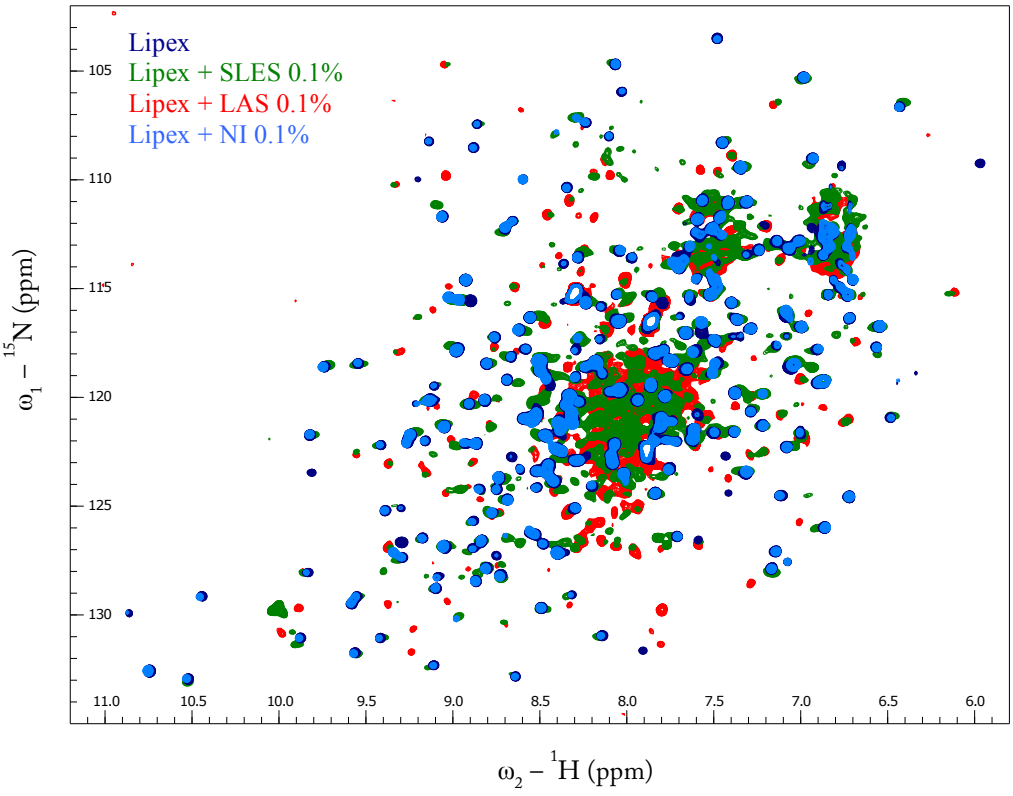


Figure 4.48 2D ^{15}N TROSY-HSQC overlay of surfactants at 0.1%

Lipex alone (dark blue) *Lipex + 0.1% LAS (red)*

Lipex + 0.1% SLES (green) *Lipex + 0.1% NI (light blue)*

Spectra collected in 25 mM Tris-HCl, 20 mM NaCl, pH 8.0 at 303 K on 800 MHz spectrometer

NI does not elicit the same response as for LAS and SLES. At 0.1% NI the spectra remains well folded and devoid of any line broadening (Figure 4.48). The lack of negative charge on NI may affect how Lipex binds at a monomeric or micelle state and may indicate why the same response is not exhibited. The line broadening noticed in LAS and SLES may be due to interactions between surfactant and Lipex charged side chains, promoting either reduced solubilisation, leading to destabilisation or aggregation of the enzyme. The charge between Lipex-LAS complexes may lead to dimer formation, similarly noted between cutinase-SDS complexes (Nielsen *et al*., 2005). It is known that non-ionic surfactants do not alter the protein structure as much as ionic surfactants such as SDS (Helenius & Simons, 1975), and do not affect the stability of TlL, (Mogensen *et al*., 2005); as evidenced here with a different mode of interaction. Whilst the activity is altered, the presence of NI alone does not elicit the same denaturation effect of anionic LAS and SLES. This change in activity noted for

NI only formulations (Figure 4.22 K NI at 0.05%) may not be a response from unfolding, but perhaps the occlusion of the active site from the presence of NI surfactant.

Another factor is LAS and SLES both exhibit large levels of line broadening at 0.1% (Figure 4.48), this may be due to the partial unfolding, or at these percentages the surfactants exceed their CMC values (0.65g/L and 0.02g/L respectively) and are present as a micelle. The CMC for NI is not available since it is kept under industrial patents. Determining the micelle state for these surfactants is difficult since they are given as industrial grade chemicals, the MW and purity can vary creating ambiguity in calculations. Techniques such as conductometry, fluorescence spectroscopy and tensiometry could be performed on each surfactant to accurately the micelle state at given percentages and water hardness (Chakraborty *et al*., 2011). Binding to the micelles may be different to that of surfactant monomers and may be the defining factor for initiating protein unfolding. The size and conformation of micelles can bind to enzymes in either a protective or non-protective manner. The micelles may be a microemulsion that form instantaneously and are often used to extract and isolate proteins (Vasudevan *et al*., 1994), thereby affecting the protein structure. This supports the results seen for SLES, which has a lower CMC value for LAS, therefore the unfolding at 0.1% may be due to a micelle interaction.

The effects of unfolding were not investigated as a factor of water hardness, something that would be of industrial significance. Calcium can add beneficial effects on detergent enzymes such as amylases reducing their susceptibility of surfactant induced unfolding (Stoner *et al*., 2005). The effects of increasing activity at increasing water hardness concentrations may be due to a stabilising effect of calcium on Lipex.

Formulations currently used and under consideration for commercial usage store the enzyme in an unfolded state. The degree of refolding upon dilution during laundry wash cycle is still to be determined. The interactions of each surfactant vary at different percentages; at a monomeric state LAS interacts differently to SLES and NI, which appear to be the same. At 0.1% SLES has shown to be have destabilising effects on Lipex structure, more so than LAS. At 0.1% NI shows to have no negative

effect on the enzyme. These investigations provide consideration into how future formulations are designed, taking Lipex stability and structure into account.

4.6 Chapter summary

4.6.1 Concluding remarks

The concluding remarks from this project are limited to an unassigned ^{15}N-HSQC spectra. The success of the project was the optimised expression and purification of Lipex in *E. coli* for NMR studies. The successful expression and purification of the Lipex enzyme took a considerable amount of time, leading to insufficient time left within the PhD timeframe to study aspects of the enzyme using NMR, to a completed level. As a result, much of this work has been a preliminary insight and investigation to the NMR characterisation of Lipex across a range of conditions. These conditions were industrially relevant to enzyme activity under different washing conditions and formulation mixtures; and academically relevant to protein thermal stability, and solubility in respect to maintaining function and structure.

New surfactant compositions were found with increased lipolytic activity in low water hardness and high water hardness conditions respectively. These formulations can be further studied to provide specifically designed tablets for given water hardness regions or countries. Initial insights were gained from NMR investigations using surfactants individually, and as a mixture, with the additional variable of water hardness. We have preliminary understanding of the link between the conformational states, dynamics and activity of Lipex in the presence of surfactants and calcium as determined by ^{15}N-HSQC spectra and the T_2 measurements. Whilst the interfacial activation and mechanism is still to be fully realised, these insights provide new information on the system. Further to this we have began to investigate and determine the causes of protein unfolding under individual surfactants identifying the presence of anionic, in particular SLES as a significant factor. New formulations can be designed with this information to reduce protein unfolding and increase activity and shelf-life of products. The enzyme system is highly complex and multifactorial; coupling both academically interesting and industrially relevant fields of research the scope for the project can take many intriguing paths.

4.6.2 Future work

It still remains to fully understand the roles of surfactants on interfacial activation. The analysis of NMR results has been limited by the partial backbone assignment, work that is still in progress. Complete assignment will provide a much more in-depth and informative analysis of results obtained, where chemical shift perturbations and changes in intensity described here can be mapped structurally. Additional to this, Unilever are interested in using NMR techniques in a number of different investigations on Lipex. One field of research is determining the interactions between Lipex and other enzymes such as proteases present within the detergent mix. Determining the site of proteolytic degradation of Lipex and knowing the level of proteolysis would help engineer variants to combat this effect allowing for more effective product design.

Beyond the scope of this work solid-state investigations to how lipase absorbs to lipid micelles, in presence of surfactants could elucidate structural changes in the lipase during interfacial activation. A study already published by Ullrich *et* al., in 2011 combines solid state NMR with enzyme activity to understand both the structural and kinetic aspects of enzymatic reactions. Solid state provides a technique to study the effects of lipase in a powdered environment, allowing characterisation of the enzyme within its storage as a tablet form, and on solid material, better replicating washing environments. Solid-state NMR has been investigated for Lipex, and preliminary spectrum show promising scope to study this enzyme under this powerful technique.

Another field of research is the comparison between thermal adapted and mesophilic lipases, in effort to form next generation industrial enzymes that are catalytically active at lower temperatures but possess the stability of thermal proteins. This is of environmental and financial interest to the detergent industry. Understanding the structural and kinetic differences between these lipases and implementing changes to the Lipex amino acid sequence can lead to the next generation of industrial enzymes.

CHAPTER FIVE

HYPFN:

INTRODUCTION

Chapter summary

The previous project studied the changes to enzyme conformation and structure in relation to catalytic activity on an industrial application. The following project is focused on similar themes, but studying the effects of protein conformations in the view of aggregation and amyloidogenesis occurring in human health and disease. Using a model protein, the N terminal domain of HypF (HypF-N), we investigate how the changes in structural conformations lead to different levels of cytotoxic aggregates. HypF is an *E. coli* protein involved in the maturation and regulation of nickel containing hydrogenases, the biological role of HypF-N is unrelated to its ability to form amyloid fibrils under specific solution conditions. Using HypF-N we can study how under separate solution conditions the protein can form two types of oligomers, which can be toxic and non-toxic to cells. This chapter introduces the background of pathogenic amyloids and contextual research. Current views of the aggregation pathway are described, linking structural features to cytotoxicity. HypF-N is introduced along with the previous research, which has elucidated its aggregative properties. Finally the hypothesis and aims are discussed in the final section.

5.1 Protein aggregation in human health and disease

An important functional requisite of all proteins, especially enzymes, is their solubility. To this end, nature has, through evolution, selected protein sequences with specific requirements for maintaining solubility (Otzen *et al*., 2000, Richardson & Richardson, 2002). In rare cases, conformational flexibility can be detrimental to proteins *in vivo*, playing an elemental role in the conversion of a few usually soluble, functional proteins into insoluble aggregates. Understanding the origin of this pathway can be significant to control the solubility of proteins in industrial processes, as well as to boost the research in neurodegenerative diseases. Indeed, some fibrillar protein aggregates, namely amyloids, are linked to over 20 pathological diseases (Chiti & Dobson, 2006), that can vary from neurodegenerative diseases such as Alzheimer's and Parkinson's, localised non-neuropathic such as Diabetes Mellitus type II or systemic like fibrinogen amyloidosis. Whilst the polypeptide sequences and secondary structure between the causative proteins do not share any identity or

homology, the amyloid fibrils formed all exhibit a common well-defined structure (Sunde *et al*., 1997), composed of a protofilament sub-structure (Serpell *et al*., 2000). Polypeptide sequences involved in pathogenic amyloids include a range of structurally dissimilar full-length proteins (lysozyme), biological peptides (amylin), protein fragments (Alzheimer β-peptide (Aβ)) and degradative protein fragments (polyQ extensions) (Stefani & Dobson 2003).

Neurodegenerative amyloid diseases alone affect an ever increasing number of patients, whilst mostly connected to the elderly, these diseases are global and indiscriminate against most factors except for age. Research undertaken in 2015 showed that the number of patients suffering from Alzheimer's disease alone is estimated to be on a global scale of ~50 million, and predictions show this number will rise to 75 million by 2030 (Prince *et al*., 2015). The estimated economical cost of dementia, in treatments and patient care currently is 620 million pounds, rising to 1.5 trillion pounds by 2030 (Prince *et al*., 2015). As advancements and progression is made into human health and disease, along with all other factors of life, the average life span will continue to increase; leading to greater prevalence of these neurodegenerative diseases. It is therefore of critical significance to understand the biochemical and biomolecular processes and mechanisms occurring during these diseases to identify better drug and treatment therapies.

Aggregation of polypeptide sequences into amyloids in nature is not always associated with disease, hydrophobins (Sunde *et al*., 2008), spider silk (Gosline *et al*., 1986), biofilm formation (Sauer *et al*., 2007) and peptide storage in mammals (Maji *et al*., 2009, Narayanaswamy *et al*., 2009) all utilise the process of protein aggregation for an advantageous functional role. Furthering our knowledge on aggregation will help understand pathological and natural processes in nature, and could potentially lead to applications and advancements in biotechnology (Lewis, 2006).

The ability to misfold and aggregate is believed to be a property intrinsically inherent to all polypeptide chains (Dobson, 2001). Natural selection and evolution has led to the formation and preservation of structures and sequences with reduced propensity for aggregation. Despite evolutionary steps taken to reduce aggregation (Richardson & Richardson, 2002) and the presence of molecular chaperones, protein misfolding

can still occur, and under certain solution and environmental conditions this can lead to pathological results. Not all protein aggregation is disease related; many proteins have been discovered to form amyloid fibrils *in vitro* that are structurally similar to disease-associated polypeptides, naturally (Guijarro *et al*., 2005) or by manipulation of the solution conditions (Chiti *et al*., 1999, Chiti *et al*., 2001). These nonpathogenic examples can, *in vitro*, form fibrils and it has been shown that their pre-fibrillar state can induce cytotoxic effects to cultured cells when added to the culture medium (Guijarro *et al*., 2005, Chiti *et al*., 2001). These cytotoxic responses are comparable to pathogenic proteins such as Aβ (Bucciantini *et al*., 2002). Whilst not related to disease, these proteins demonstrate a common feature during amyloid formation that results in cytotoxic effects. These amyloidogenic proteins, both pathogenic and non-pathogenic share little sequence homology but result in structurally similar fibrils, supporting a hypothesis that formation of amyloid fibrils can occur for any given polypeptide chain. The peptide backbone structure appears to be the singular commonality between all proteins studied (Chiti *et al*., 1999, Dobson 2001, 2002). There are however, certain attributes that appear to be dominant characteristics in the polypeptide backbone of proteins susceptible to aggregation such as increased hydrophobicity, high propensity for β-sheet secondary structure and low overall charge (Chiti *et al*., 2002).

5.2 Architecture of amyloid fibrils

Amyloid fibrils are composed of a cross β-structure with the protein backbone strands running perpendicular to the long axis of the fibril in an extended conformation. One or more protofibril units are present, usually wrapping around each other to form the ordered fibril. Examples of fibrillar structures are shown in a model system and from NMR derived methods in Figures 5.1 and 5.2. The protofibril strands are interconnected and stabilised by hydrogen bonding between the β-strands to form a rigid stable structure (Eanes & Glenner, 1968, Serpell *et al*., 1999). The cross β-fold has shown to be significant for functional amyloids such as hormone peptide storage (Maji *et al*., 2009). Amyloid fibrils share a common structure despite different initial protein structures and conformations. The well-documented protein lysozyme forms fibrils that are structurally similar despite different conformational non-native states under low pH and temperature (Krebs *et al*., 2000) and denaturant guanidine

hydrochloride (Vernaglia *et al*., 2004). These results highlight a common pathway for amyloid formation. The structural features of fibrils are visible through methods such as electron microscopy and atomic force microscopy (AFM) depicting a straight unbranched morphology up to several μm in length and 60 – 120 Å in diameter. High-resolution atomic models of Aβ has been elucidated by solid state NMR (ssNMR) and X-ray crystallography methods which depicts a triple parallel β-sheet motif, providing insights into the propagation of the fibril and the formation and intermolecular contacts that stabilise the oligomer species of the $A\beta_{1-42}$ peptide (Xiao *et al*., 2015).

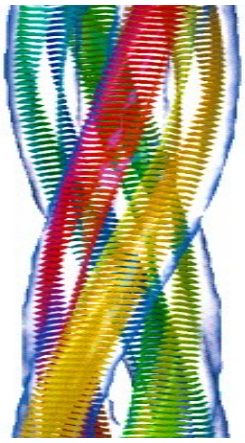
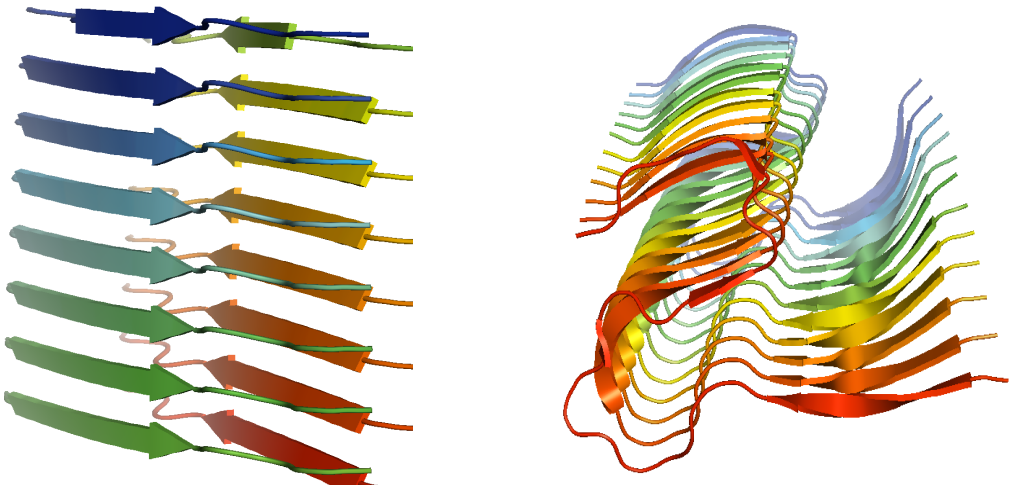


Figure 5.1 Molecular model of 4-protofilament insulin fibril

Image taken from (http://people.cryst.bbk.ac.uk/~ubcg16z/amyloid/insmod.jpg)



(A) *(B)*

Figure 5.2 NMR derived β- stranded structures and packing of protofilaments

(A) Cross β-amyloid fibril derived from combinational methods using solid state NMR (Fitzpatrick et al., 2013) (PDB code: 2M5N)

(B) Aβ fibrillar structure determined by solid state NMR Xiao et al., 2015 (PDB code: 2MXU)

5.3 Toxicity arises from prefibrillar aggregates

Amyloid deposits can be further assembled into larger aggregative plaques and proteinaceous deposits. These deposits are visible in the tissues of patients suffering from pathological amyloid diseases, suggesting a causative link between aggregative formations and pathological symptoms, termed the amyloid hypothesis (Kelly, 1998). These deposits are commonly found extracellularly, but can be intracellular, localised in the cytoplasm or nucleus. This hypothesis is of intense debate, the specific cause of cellular dysfunction and pathology is argued to lie elsewhere to amyloid deposits (Lambert *et al*., 1998, Conway *et al*., 2000). One major doubt over the amyloid hypothesis was the emergence that the levels of Aβ amyloid deposition and the fibril densities did not correlate to the severities of dementia related symptoms in Alzheimer's patients (Dickson, 1995). Recently it has been shown that the mature ordered amyloid fibrils do not elicit the same levels of cell impairment and reduction of cell viability as compared to the early stage aggregates; casting further doubt over the amyloid hypothesis (Bucciantini *et al.,* 2002). An alternative theory is that the mature fibrils, which show reduced toxicity compared to the early stage aggregates, are in fact formed as a protective mechanism; therefore the formation of amyloid fibrils can be viewed as a cellular response to removing and depleting the cytotoxic prefibrillar aggregate species (Caughey *et al*., 2003, Volles & Lansbury, 2003).

Increasing evidence suggests that the harmful species in amyloid diseases is the pre-fibrillar and diffusible oligomer aggregates, also referred to as protofibrils, which have been shown to interact with phospholipid bilayers (Merlini & Bellotti, 2003, Kayed *et al.,* 2003). This is found to be consistent in both pathogenic proteins such as Aβ (Bhatia *et al*., 2000) and α-synculein (Conway *et al*., 2000) and in non-pathogenic proteins such as HypF-N (Bucciantini *et al*., 2002). These results indicate that common structural aspects of these pre-fibrillar aggregates are the cause of cytotoxicity, rather than the specific amino acid sequence. Structural determination of these intermediates is currently a top challenge in biochemistry. This challenge is complicated by the inherent structural dynamics, transient life span and structural heterogeneity of oligomers.

5.4 Architecture of oligomers

Structural studies of oligomers are a challenging field given their size, homogeneity and transient life span. A number of techniques have been employed to study the structure of oligomers and fibrils such as ssNMR, X-ray crystallography, transmission electron microscopy (TEM), small angle X-ray scattering and analytical ultracentrifugation, allowing for experimentally derived observations compared to previously drawn data from molecular simulations. Inherent properties of fibrils and oligomers make them difficult subjects for crystallisation and artefacts may occur during sample preparation for TEM; ssNMR is an increasingly popular method for high-resolution structural studies of homogenous amyloid fibrils and precursors.

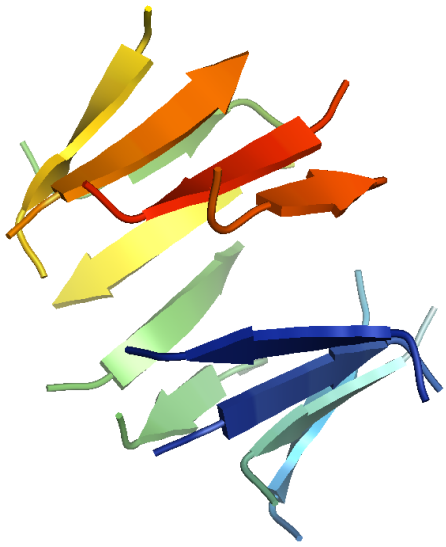


Figure 5.3. Fragment of human prion protein in an oligomeric structure determined by X-ray crystallography

Apostol et al., 2013 (PDB code: 4E1H)

The pre-fibrillar oligomeric species is composed generally of a stable core with an increased β-sheet structure (as compared to the monomeric structure of the protein), but reduced β-sheet content than the fibril (Breydo *et al.*, 2016). Antibodies synthesised for binding to Aβ protofibrils also bind to protofibrils of different proteins, indicating similar structural features (Kayed *et al.*, 2003). There is a degree of heterogeneity between oligomers in terms of structure, stability and size; different forms of oligomers exhibiting antiparallel or parallel β-strands have been observed for Aβ (Huange *et al.*, 2015, Chimon & Ishii, 2005). An example of a β-stranded

oligomer formed from a fragment of the toxic human prion protein is shown in Figure 5.3, depicting an ordered β-sheet rich oligomer for a non-amyloidogenic protein. Oligomers containing α-helical structure have been observed for α–synuclein after the oligomers bind to a lipid membrane demonstrating structural differences (Varkey *et al*., 2013). However, most commonly, all oligomers comprise of some level of β-strand structure, generally anti-parallel at the secondary structure level (Cerf *et al*., 2009).

5.5 Aggregation pathway

The pathway of protein aggregation follows certain patterns for all amyloidogenic proteins. The initial native conformations of proteins are in an equilibrium state between a range of conformational ensembles, under certain conditions aggregation-prone populations are stabilised and this can initiate the aggregation pathway. Alternatively, misfolding can lead to off-pathway aggregation into disordered aggregates or degraded species. Aggregation can, in certain cases, lead to the formation of prefibrillar aggregates that are predominantly composed of a β-stranded protofilament structure. These early stage aggregates self-assemble to from larger amyloid fibrils, which have the same overall structural composition, discussed in Section 5.2.

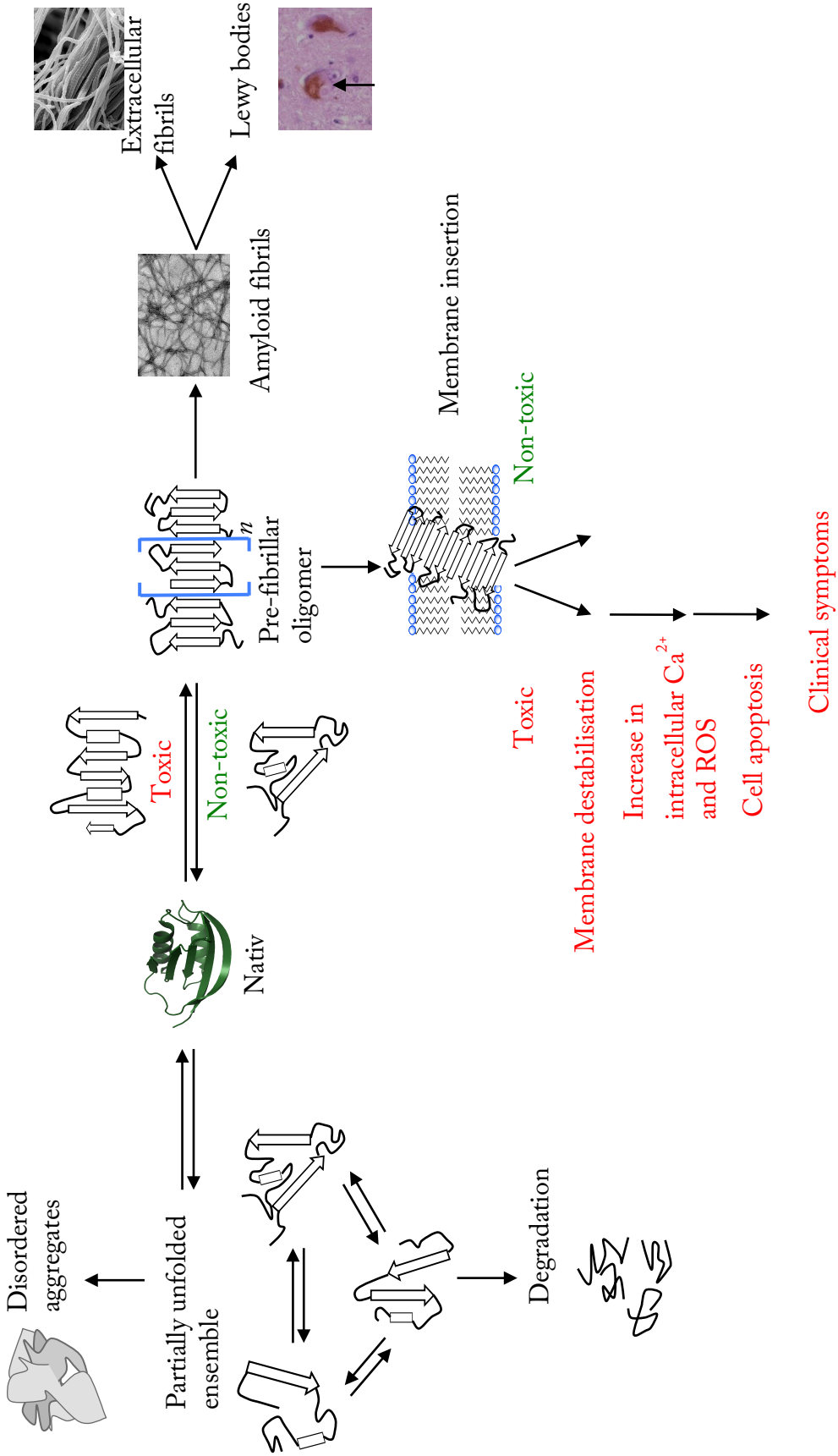


Figure 5.4 Aggregation pathway of amyloidogenic proteins

HypF-N used as a example, with additional notes

Figure 5.4 presents a general overview of pathogenic protein aggregation, the initial conformations of proteins can vary; α-synuclein exhibits an unfolded and disordered monomeric state whilst the toxic HypF-N is more ordered, in a native conformation. Whilst fibrils of amyloidogenic proteins are composed of a β-stranded core (Serpell, 2000), β-stranded structure, and secondary structure, is not always a pre-requisite for aggregation; regions of Aβ peptide and α-synuclein protein that lack a structural presence have been shown to promote aggregation (Pawar *et al*., 2005, Bemporad *et al*., 2006). Both these pathogenic proteins are largely intrinsically unfolded in their monomeric state. Additionally the formation of amyloid fibrils from apomyoglobin has been shown to be dependent on a disordered polypeptide sequence despite its α-helical structure (Fändrich *et al*., 2003).

The primary steps in aggregation for all proteins is likely to be governed by the physiochemical factors such as hydrophobicity, net charge and aggregation propensity (Chiti *et al*., 2003); however, the precise nucleation process during prefibrillar and protofilament formation is likely to vary, to a degree, between proteins based on these differences in the initial conformations. Recent advancements in biophysical techniques have allowed ultrastructure investigations revealing polymorphisms between fibril structures (Bubeck, 2015). However, the general principles and route of aggregation is similar and is shared across all amyloidogenic proteins.

5.6 Key features of oligomer cytotoxicity

The molecular basis of toxic oligomers appears to have commonalities between different proteins and polypeptide chains. Exposure of the hydrophobic regions in proteins that are usually maintained within the core of the protein, has been shown to be a significant factor in cytotoxic oligomers; this alteration promotes destabilising interactions with cell membranes (Bychkova *et al*., 1988, Volles & Lansbury, 2001, Kourie & Shorthouse 2000, Kourie & Henry, 2002). The toxic gain of function from monomer to protofibril appears to be through this feature and reveals the significance of the initial state, and the nucleation process on forming toxic aggregates. Different forms of α-synuclein have shown varied levels of cytotoxicity, the differences in oligomer morphology appear to lie in a higher degree of flexibility and a less dense core, implying greater exposure of hydrophobic regions (Cremades *et al*., 2012,

Paslawski *et al*., 2014). The structural integrity and ordered stability of mature fibrils is remarkably different to the more flexible oligomer species, furthermore, the higher level of structural organisation in the fibril state reduces the exposure of hydrophobic regions (Rahimi *et al*., 2008). This is evidenced structurally between Figures 5.1 and 5.2, which depict a loss of defined order between fibril and oligomer respectively. It is also considered that the high mobility of oligomers is another factor linked to their cytotoxicity.

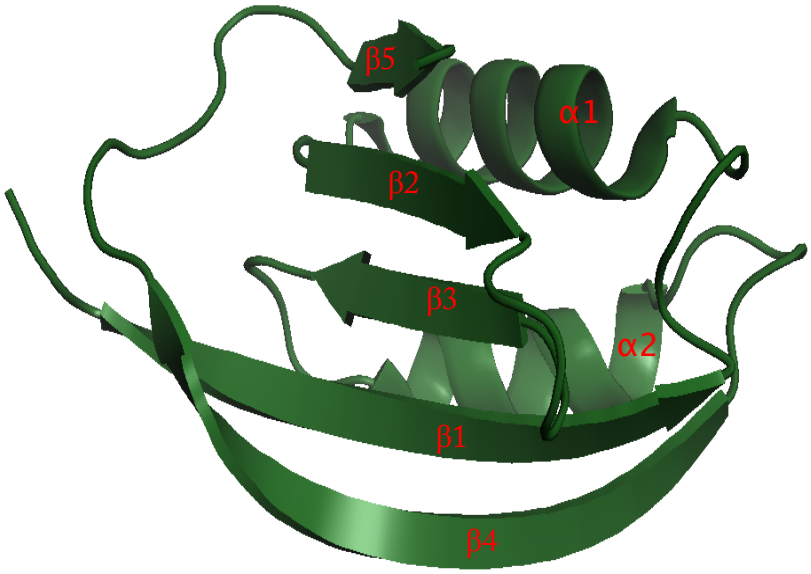
5.7 HypF-N

HypF-N is a non-pathogenic protein with the ability to form amyloid fibrils (Chiti *et al*., 2001). Based on the solution conditions HypF-N can uniquely form two different types of oligomers that are toxic and non-toxic to cell cultures (Campioni *et al*., 2010). The background and previous research of this protein is presented below.

HypF is part of a family of *E. coli* enzymes (HypA-F) that have a functional role in the maturation and regulation of nickel containing hydrogenases in *E. coli* bacteria. Hydrogenases are key enzymes involved in the oxidation and metabolism of hydrogen in prokaryotes (Colbeau *et al*., 1998). HypF is an 82 kDa auxiliary protein involved in the nickel insertion of active [NiFe] hydrogenase enzymes (Maier & Böck, 1996, Binder *et al*., 1996, Colbeau *et al*., 1998), and also plays a role in post-translational activation of the regulatory hydrogenase HupUV proteins (Yamamoto *et al*., 1990). The native structure of HypF has been determined by X-ray crystallography at 1.27Å resolution (Rosano *et al*., 2002) and has been shown to contain two identical zinc finger motifs (Yamamoto *et al*., 1990) and an acyl phosphatase motif (Wolf *et al*., 1998). HypF consists of three domains: an N terminal domain and the two zinc finger motifs (Wolf *et al*., 1998, Casalot & Rousset, 2001).

The N-terminal domain (HypF-N) is a 91 amino acid sequence with a MW of 10.46 kDa, and has been shown to contain an acylphosphatase motif (Rosano *et al*., 2002). It shares 22% and 50% sequence identity to human and *E. coli* acylphosphatase enzymes and has structural similarities to acylphospatase domains and the ferredoxin superfamily (Rosano *et al*., 2002). HypF-N is a small α/β protein with a diameter of

~3.5 nm (Campioni *et al*., 2008) and has an overall topology occupying a volume of 43 Å × 28 Å × 27 Å, consisting of 5 β-sheet strands and 2 α-helical regions in a secondary structure composition of βαββαβ (Figure 5.5). The β-sheets are arranged as five slightly twisted antiparallel sheets, with one side facing the antiparallel α-helices and the other side being completely solvent exposed. The hydrophobic residues of the α-helices are packed against the inner face of the β-sheet. Each secondary structure element is connected via a loop region giving six loops (Rosano *et al*., 2002). HypF-N contains a surface exposed cysteine residue (Cys8), which can form disulphide bonds with Cys8′ of neighbouring HypF-N molecules, dimerization can be prevented by addition of reducing agents such as DTT to keep the protein in its functional monomeric native state, shown in Figure 5.5.



MAKNTSCGVQLRIRGKVQGVGFRPFVWQLAQQLNLHGDVCNDGDGVE
VRLREDPETFLVQLYQHCPPLARIDSVEREPFIWSQLPTEFTIR

Figure 5.5 HypF-N structure and sequence

(Rosano et al., 2002) (PDB code: 1GXT)

HypF-N can aggregate *in vitro* into amyloid fibrils under specific solution conditions (Chiti *et al*., 2001). HypF-N is one of many non-pathogenic proteins that form amyloid fibrils, as discussed further many acylphosphatase enzymes share this feature, along with proteins such as lysosomal enzymes, muscle myoglobin and

kinase domains (Glenner *et al.,* 1974, Chiti *et al.,* 1999, Fandrich *et al*., 2001, Guijarro *et al.,* 2005).

HypF-N (as for the other non-pathogenic amyloidogenic proteins), forms amyloid fibrils under *in vitro* conditions, which destabilise the native globular fold, but do not affect non-covalent bonds. Protein aggregation occurs via the existence of misfolded conformations (Stefani & Dobson 2003), for the case of HypF-N this is initiated by the solution conditions (Chiti *et al*., 2001). HypF-N amyloid formation was first discovered using 10-40% (v/v) trifluoroethanol (TFE) at neutral pH (Chiti *et al.,* 2001). TFE is used as an organic solvent, which denatures stable proteins (Luo & Baldwin, 1998) but also stabilizes secondary structure and promotes main chain hydrogen bonds by providing a low di-electric constant environment (Roccatano *et al.,* 2002). The solution conditions play an important role in increasing the propensity of aggregation, promoting a conformation that has a greater aggregation propensity (Chiti *et al.,* 2001), as does the polypeptide sequence (López de la Paz *et al.,* 2002). Although HypF-N is not associated to any pathogenic amyloid disease, its mature fibrils formed *in vitro* share morphological, tinctorial, physiochemical (Marcon *et al.,* 2005, Campioni *et al.,* 2008, Relini *et al.,* 2004) and behavioural similarities to pathological amyloid fibrils (Bucciantini *et al.,* 2004, Tatini *et al.,* 2013); the two being considered undistinguishable (Relini *et al.,* 2004).

5.8 HypF-N aggregation

The unique characteristic of HypF-N is its ability to form oligomers both toxic and non-toxic to cell cultures using differing environmental conditions: toxic conditions (50 mM sodium acetate, 12% (v/v) TFE, 2 mM dithiothreitol (DTT), pH 5.5) and non-toxic (20 mM trifluoroacetic acid (TFA), 330 mM sodium chloride pH 1.7) (Campioni *et al.,* 2010). Both sets of oligomers are morphologically similar as determined by AFM, with an overall spherical conformation, and slight size diameter difference (toxic are ~7-9 nm, non-toxic are ~9-11nm). Both oligomers bind to thioflavin T (ThT) and give increased fluorescence indicative of the formation of rich intermolecular β-sheet structures and amyloid fibrils (LeVine, 1999, Campioni *et al*., 2010). It has been shown, that toxicity originates from the oligomer species, not the mature fibril (Cecchi *et al*., 2005). Both forms of oligomers are stable up to 24 h in

solution and do not undergo any structural reorganisations, they also can be trapped in their oligomeric state allowing experiments to be performed selectively, on the protofibril structures (Campioni *et al*., 2010).

The toxicity of each oligomer was assessed by 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) reduction inhibition assay, often used to measure cellular viability (Mosmann, 1983). Toxic oligomers reported substantial decreases in MTT reduction after being added to cell cultures of human neuroblastoma and mouse endothelial cells. These reductions were comparable to Aβ and IAPP responsible for cell death in Alzheimer's disease and type II diabetes respectively. The non-toxic oligomers showed no effect on either cell line, and results were comparable to untreated cells, and cells treated with native HypF-N (Campioni *et al*., 2010). These results were supported by Hoechst 33342 staining and release of intracellular cytosolic calcium level which are both indicative markers of apoptotic cell death caused by toxic oligomers (Demuro *et al*., 2005). Both types of oligomers bind to the cell membrane, but only the toxic oligomer crosses the bilayer and causes influx of extracellular calcium into the cytosol (Campioni *et al*., 2010).

Cell apoptosis in the presence of toxic HypF-N oligomers has been attributed to increased reactive oxygen species and calcium ion dysregulation and dyshomeostasis (Cecchi *et al*., 2005, Canale *et al*., 2006, Campioni *et al*., 2010, Zampagni *et al*., 2011). These results have been determined across a range of cell lines and are factors known to trigger apoptotic cell death via the caspase-3 pathway in cells exposed to deleterious protein aggregates (Demuro *et al*., 2005). These effects are unobserved for the non-toxic HypF-N oligomers (Zampagni *et al*., 2011), but are shared by pathogenic amyloid protein aggregates, which supports a common mode of action for all cytotoxic oligomers (Bucciantini *et al.,* 2004, Kayed *et al*., 2003). There is strong evidence to indicate that HypF-N aggregates cause cell death through membrane pore-formation. Pre-fibrillar aggregates have been noted to disrupt lipid vesicles leading to membrane permeabilization, leading to an influx of calcium ions and finally cell toxicity and death (Relini *et al*., 2004). Similar modes of action have been observed for Aβ peptides (Kwahara & Kuroda, 2000) and α-synuclein (Volles & Lansbury, 2002) causing neurodegenerative diseases. Cell lines such as IMR90

fibroblasts are highly resistant to cytotoxic effects of protein aggregates; the higher resistance has been identified in the cellular biochemistry and machinery that reduce the causes responsible for cell apoptosis and reduce the activation of apoptotic pathways (Cecchi *et al*., 2006).

Different experimental conditions can give rise to oligomers and fibrils that vary in toxicity for other proteins such as α-synuclein, however these oligomers are highly unstable; with HypF-N the difference in toxicity is definitive between non-toxic and toxic, rather than a degree of toxicity (Lee *et al*., 2007, Kayed *et al*., 2009). HypF-N therefore provides a unique and exploitative opportunity to study the structural and functional mechanisms behind pre-fibrillar oligomer toxicity within a single system.

5.9 Differences in initial structure of HypF-N

For the native globular conformation of HypF-N in the toxic condition, it appears that HypF-N aggregates through inherent fluctuations of a conformational ensemble that is transitional from the native state (Marcon *et al*., 2005). The structure of HypF-N under 12% TFE as measured by far UV CD is indistinguishable to its native state, but HypF-N aggregates most likely through a transient partially unfolded conformation that is in dynamic equilibrium with the native state, the population of this conformation can be as low as 1%, and still initiate aggregation (Marcon *et al*., 2005).

Under the non-toxic condition the protein maintains a degree of secondary structure and hydrophobic packing as shown by far UV CD but adopts a premolten globular state (Calloni *et al.,* 2008, Campioni *et al.,* 2008). The premolten globule conformation is defined as a species with residual structure between molten globule and fully disordered.

In many cases of protein aggregation *in vitro* the solution conditions can be the initiating cause of aggregation based on conformational changes to the protein structure (Chiti *et al*., 1999, Chiti *et al*., 2001, Zurdo *et al*., 2001, Fezoui & Teplow 2002). There are an increasing number of amyloidogenic proteins being determined *in vitro* than naturally occurring *in vivo*; this is most likely due to the fact that more environmental and solution conditions can be controlled *in vitro* (Stefani & Dobson, 2003). Additionally structural elements of protein, such as protecting β-edge strands,

have evolved to reduce the propensity of aggregation *in vivo*; it is often these features, which are destabilized within *in vitro* conditions (Richardson & Richardson, 2002). The common pathway that seems consistent through all oligomers studies (Figure 5.4), allows findings from HypF-N and other such proteins to be applicable to the basis of pathogenic proteins such as Aβ, which are linked to human health and disease.

5.10 Toxicity arises from exposure of hydrophobic regions

Fluorescent measurements have shown that both the toxic and non-toxic oligomer forms via the regions of the sequence 22 – 34, 55 – 59 and 75 – 87 which are at the surface of the protein and more likely to be contact with the oligomer, with a predicted higher tendency to be involved in nucleation (Figure 5.7) (Campioni *et al.*, 2010). These regions contain a number of hydrophobic residues, which is believed to drive the nucleation. These regions differ from predictions based on the high aggregation propensity regions (9 – 15, 27 – 35, 46 – 48 and 58 – 60) (Calloni *et al.*, 2008). For the unstructured conformation in the non-toxic condition, it was initially proposed that regions of high aggregation propensity drove the nucleation process (Calloni *et al.*, 2008).

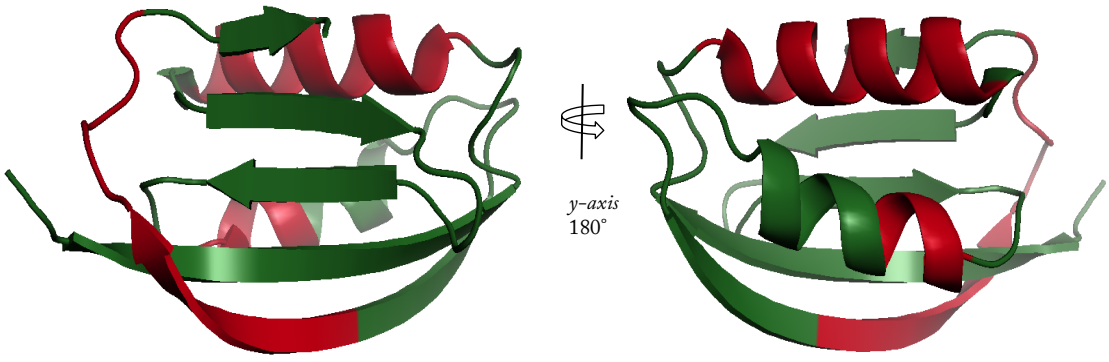


Figure 5.6 Regions predicted to form core of HypF-N oligomers

Regions coloured red are predicted to form the core of HypF-N oligomers as determined by fluorescent measurements by Campioni et al., 2010.

The same regions determined for the non-toxic condition are involved for the toxic condition; however, the reduced intermolecular packing these regions leads to a less

dense core and greater surface exposure of the hydrophobic regions in the toxic condition (Campioni *et al*., 2010). This exposure and the structural plasticity for toxic oligomers, are unobserved in the non-toxic oligomer species benign to cultured cells, and is the attributing cause for HypF-N toxicity (Campioni *et al*., 2010, Zampagni *et al.,* 2011).

It is not uncommon for a protein to aggregate in the absence of an unfolded conformation, even in highly populated structures with α-helical content amyloid formation can occur, with conversion to the β-stranded topology occurring late in the pathway (Fezoui & Teplow 2002, Krkitadze *et al*., 2001).

In the disordered premolten globule state there is a high exposure of hydrophobic residues to the solvent (Calloni *et al*., 2008), this supports a mechanism for the non-toxic nucleation via these regions, incorporating them into the core of the oligomer to a greater and stronger degree than the toxic condition (Campioni *et al*., 2010). The reduced packing, and consequential solvent exposure of hydrophobic regions in the toxic oligomer is hypothesised to be the crucial attribution in HypF-N oligomer toxicity (Campioni *et al*., 2010).

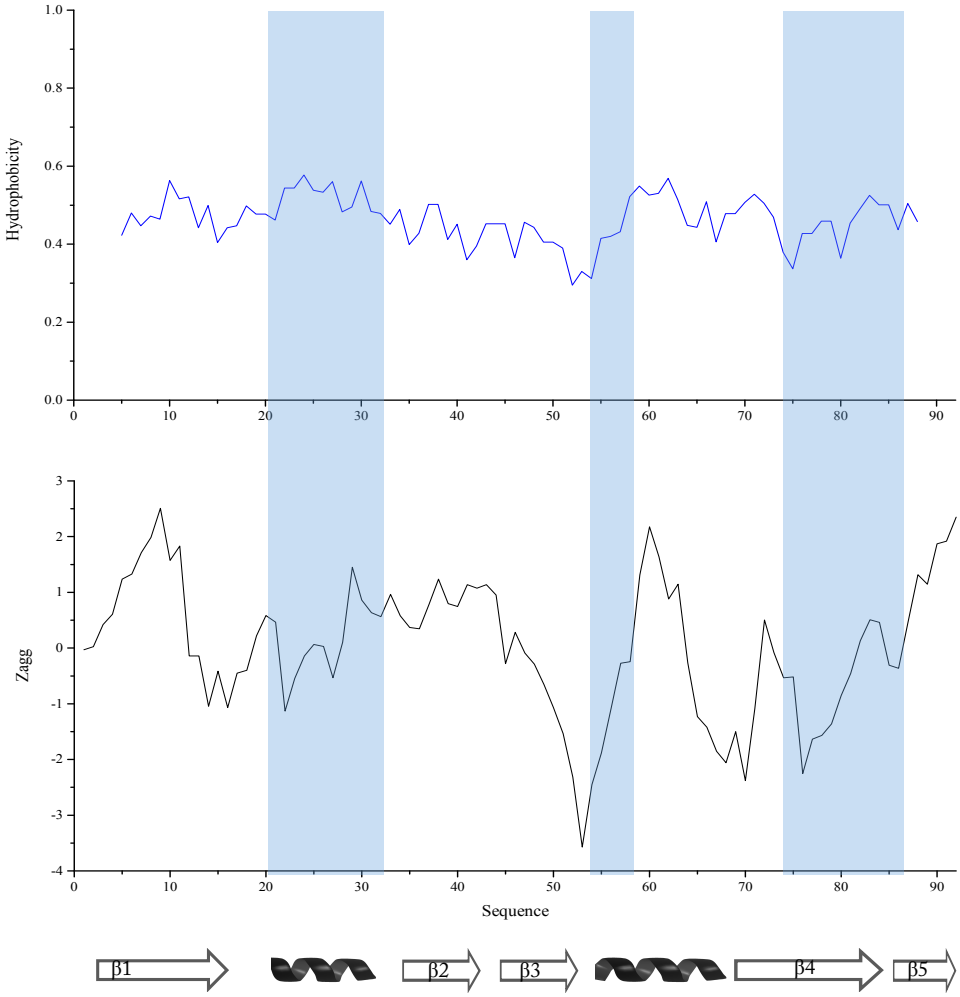


Figure 5.7 Hydrophobicity and aggregation propensity of HypF-N sequence

Hydrophobicity plot using Kyte-Doolittle prediction (Kyte & Doolittle, 1982) based on HypF-N sequence. Zagg scores predicted by Zaggregtor prediction tool (DuBay et al., 2004) based on HypF-N in the native structure.

Highlighted regions correspond to those predicted to form the core of both toxic and non-toxic oligomer (Campioni et al., 2010).

5.11 HypF-N and acylphosphatase aggregation

Acylphosphatases from *Drosophilia melanogaster* (AcPDro2), human muscle acylphosphatase (mAcp), and *Sulfolobus solfataricus* (ssoAcp) all share a similar characteristic with HypF-N, forming amyloid fibrils in varying concentrations of TFE (Plakoutsi *et al*., 2004, Chiti *et al*., 1999, Soldi *et al*., 2005). In all cases the protein maintains a monomeric native-like conformation, which is a βαββαβ topology

identical to HypF-N (De Simone *et al*., 2011, Calamai *et al*., 2005, Plakoutsi *et al*., 2005). Similarities between the proteins originate within their similar primary polypeptide backbone and structure (Figure 5.8), which acts as a major determinant in amyloid structures through the formation of hydrogen bonds (Chamberlain *et al*., 2000). There are observed differences between the proteins, where the rate and mechanism of aggregation between ssoAcp and HypF-N differ (Bemporad & Chiti, 2009). ssoAcp aggregates significantly faster than it unfolds, via fluctuations in an ensemble of the native state (Plakoutsi *et al*., 2004); however HypF-N appears to aggregate through fluctuations of an ensemble that is transitional to the native state (Marcon *et al*., 2005). This marks a difference between the two proteins which otherwise would share similar aggregation characteristics. HypF-N appears to aggregate closer to mAcp, both via a partially unfolded conformation induced by the presence of TFE (Calamai *et al*., 2005).

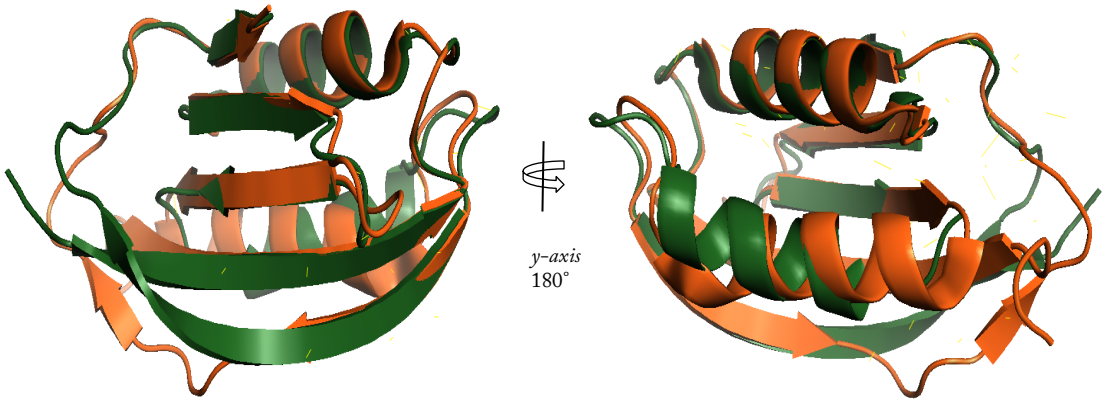


Figure 5.8 Structural similarities between HypF-N and AcpDro2

HypF-N shown in green (PDB code: 1GXT) (Rosano et al., 2002)

AcpDro2 shown in orange (PDB code: 1URR) (Zuccotti et al., 2003)

For both HypF-N and mAcp stabilisation of the α-helical content did not increase the intrinsic aggregation propensities (Taddei *et al*., 2001, Ahmed *et al*., 2011). This implies that aggregation is likely to occur across the β-sheet secondary structure, or via hydrophobic interactions as evidenced by Campioni *et al*., 2010. It is believed that generally, aggregation occurs via the amyloidogenic regions of the sequence, which exhibits high aggregation propensity, based on their charge, β-sheet propensity and hydrophobicity, particularly the β-sheet secondary structure (Ahmed *et al*., 2011).

SsoAcp and AcpDro2 both aggregate via β-strands, through interactions with an unstructured N terminal domain and the β4 strand in ssoAcp and solvent exposed β2 and β1 strands for AcpDro2 (Plakoutsi *et al*., 2006, De Simone *et al*., 2011). The structure of AcPDro2 in 5% TFE (v/v) was calculated by a range of NMR methods. It was determined that in 5% TFE (v/v) the β5 strand was displaced, offering solvent access between the β2 and β1 strands. This was modelled and supported by simulation studies that fibrils formed between the β2 and β1 strand interaction, made possible by the displacement of β5 (De Simone *et al*., 2011). The β5 strand was defined as an element of negative design (Richardson & Richardson, 2002), present in the structure not for any functional advantage but to prevent the aggregation of the protein and maintain solubility. It is known that TFE can affect the tertiary state of the protein, however, it is unclear as to the direct effect of TFE on the β5 strand, and how the displacement is caused.

The differences between HypF-N and acylphosphatases can be attributed to the slight structural differences, amino acid composition and resulting side chain interactions and the unfolding processes, which may influence aggregation.

5.12 Mechanisms of oligomer and membrane interaction

The formation pathway of the prefibrillar species has been discussed, the following section details the interactions between the cytotoxic aggregate species and cell membranes. This interaction results in membrane destabilisation and dyshomeostasis of intracellular ions, significantly calcium, leading to an increase in reactive oxygen species, which triggers cell apoptosis.

It is believed that both pathogenic and non-pathogenic aggregates cause cytotoxic effects through impairing the same biochemical processes; implying a similar mode of action for toxic prefibrillar aggregates (Bucciantini *et al*., 2002, Kayed *et al*., 2003, Evangelisti *et al*., 2012, Tatini *et al*., 2013). Cell impairment and loss of viability from the presence of toxic oligomers is attributed to the dysregulation of fundamental cellular processes. Oligomers may activate or promote signal transduction pathways for apoptosis through cellular receptor interactions (Soto, 2003), or bind to cellular components such as transcription factors and cytoskeletal entities that are crucial for

cell viability, disrupting their physiological purpose (Ii *et al*., 1997, Cummings *et al*., 1998,). The same hallmarks of cell apoptosis are noted across cellular studies with toxic oligomers. A likely hypothesis and mechanism, which is gaining increasing supporting evidence, is that the interaction with oligomer and membrane leads to membrane disruption and potentially pore formation. This disruption promotes changes and increases of free intracellular calcium ion levels (Butterfield *et al*., 2001, Milhavet & Lehmann, 2002, Campioni *et al*., 2010). Changes to the intracellular redux status affects calcium and ion homeostasis and gives rise to an increased population of reactive oxygen and nitrogen species leading to oxidative stress, and dysregulation of signal transductions and ion homeostasis (Soto, 2003). The changes are most likely due to the interactions between oligomers and the lipid membrane. Current advancements in microscopy and other biophysical techniques have provided informative results on the interaction between oligomer and membrane, and how cytotoxicity is initiated; however, the mechanisms behind how the oligomers elicit this effect is still not fully understood.

5.13 Biophysical techniques for studying membrane interactions

It is widely considered that amyloid diseases cause cell degeneration via interactions between the prefibrillar aggregate species and the plasma membrane of cells (Kagan *et al*., 2004). The cellular pathways subsequently induced have been elucidated (Kayed *et al*., 2003, Bucciantini *et al*., 2004, Kagan *et al*., 2004, Cecchi *et al*., 2005); however the mechanisms behind initiating these pathways are still not fully understood (Andreasen *et al*., 2015). Oligomer interaction within the membrane of cells is complex and multifactorial, making it a highly challenging system to study. Understanding the interaction will provide more effective potential therapeutic strategies.

A number of different biophysical techniques can be used to study the interactions between protein aggregates and cell membranes, either in model vesicles or cells. Techniques such as macroscopic conductance measurements of lipid bilayers (Lin *et al*., 1996), confocal microscopy (Lorenzen *et al*., 2014), vesicle leakage and influx of ions (Stöckl *et al*., 2012) can measure the changes to the membrane, and channel

formation. These outcomes infer changes to the membrane physiology due to the structural aspects of the oligomers. AFM and transmission electron microscopy (TEM) are powerful methods for detecting and describing pore formation (Lorenzen *et al*., 2014), however artefacts may occur during sample preparation as measurements are recorded on dried samples. Complimentary to these methods, alternative techniques can study directly the prefibrillar species in the membrane environment. Small angle X-ray scattering has been used to study oligomer structures of α-synuclein in solution to provide *in vitro* model systems (Giehm *et al*., 2010). Additionally, tryptophan fluorescence has been used to elucidate site-specific regions of oligomer binding (van Rooijen *et al*., 2009), however fluorescence measurement methods involve changes to the protein sequence and chemical modifications to the sample preparation, which may adversely affect results. Methionine oxidation and tyrosine nitration of α-synuclein form two different forms of oligomers, disordered and a β-sheet stabilised oligomer respectively, that differ in cytotoxicity (Zhou *et al*., 2010, Uversky *et al*., 2005). Adverse changes in oligomer preparation present a level of unreliability to the results; it is therefore crucial to avoid any adverse changes in oligomer preparation. Advancements in ssNMR provide a platform to study the oligomeric entity within a model vesicle at an atomistic level without any potentially adverse changes required to the protein. ssNMR is preferred for phospholipid membrane studies over techniques such as electron paramagnetic resonance (EPR) as large nitroxide spin labels for EPR are thought to perturb the lipid structure (Lee, 1975), unlike isotopic labelling (Seeling and Seeling, 1974). Conformational variants and structural aspects can be determined, along with dynamical information and revealing interactions at a residual level (Bousset *et al*., 2012, Fusco *et al*., 2014). These techniques are complimentary and collectively provide greater understanding of how cell cytotoxicity is attributed for prefibrilliar aggregates.

5.14 Membrane physiology can modulate toxicity

Different cell lines can exhibit different affects from being incubated with the same cytotoxic oligomer. This is likely due to the individual cell biochemical and physiochemical composition (Cecchi *et al*., 2005, 2006, Evangelisti *et al*., 2012). Understanding the susceptibilities and the significant interactions that cause membrane derangement, which initiates cell apoptosis, is of crucial importance, as

developing our understanding of such pathogenic diseases at these crucial steps will allow for the design of more advanced and efficient potential therapeutic strategies.

Two key components within cell membranes: ganglioside GM1 and cholesterol have been shown to modulate toxicity of HypF-N oligomers (Evangelisti *et al*., 2012). These two components are important in the lipid raft composition (Lingwood & Simons, 2010) and are implicated in Aβ binding to cell membranes (Ehehalt *et al*., 2003, Molander-Melin *et al*., 2005). GM1 is a major ganglioside found within brain lipid rafts, known to bind to Aβ (Iwamoto *et al*., 1990) and even promote β-sheet formation (Choo-Smith *et al*., 1997) and induce polymorphisms into Aβ oligomers forming a more cytotoxic species (Yamamoto *et al*., 2008). Cholesterol alters the fluidity of membrane bilayer and may destabilise aggregates by reducing the number of interactions they may form (Yip *et al*., 2001). Patients suffering from Alzheimer's disease show higher levels of GM1 and reduced cholesterol (Fujita *et al*., 2007, Molander-Melin *et al*., 2005), with high-density clusters of GM1 being identified as a site of Aβ deposition (Yamamoto *et al*., 2008).

Increasing the content of GM1 leads to increasing the cytotoxic effects of non-toxic HypF-N oligomers, to levels comparable to the toxic condition. Decreasing the amount of cholesterol has a similar effect, increasing the cytotoxicity of the non-toxic oligomers. Cholesterol alone has been shown to increase the resistance of cell membranes to toxic HypF-N oligomers; suspected to be via impairing adsorption of oligomers to the plasma membrane (Cecchi *et al*., 2005). The lipid composition and presence of cholesterol is thought to reduce the destabilisation of the membrane upon aggregate interaction. Conversely, for the toxic condition, decreasing GM1 and increasing cholesterol affects the cytotoxicity, with oligomers eliciting no significant cytotoxic effects. When both cholesterol and GM1 are present together, it is the presence of GM1 that has the dominant contribution to cytotoxicity (Evangelisti *et al*., 2012). The results indicate a difference between the oligomers and their interactions with key components in the cell membrane. ssNMR methods would allow exploratory insights into how the structure and dynamics of the oligomer if affected in the membrane, with GM1 and cholesterol. Elucidating the molecular mechanisms governing GM1 and oligomer interaction may help define oligomer cytotoxicity.

5.15 Hypothesis and Aims

It is the solution conditions that cause changes to the native structure of the protein, forming a conformational state that is prone to aggregation (Chiti *et al*., 1999). Determining the factors that are responsible during this structural change are key to understanding how proteins maintain their soluble and functional states, for enzymes this is of critical importance since activity, function and structure are intimately linked. HypF-N is used as a model system, unique in its ability to form stable, toxic and non-toxic oligomers. Whilst HypF-N is non-pathogenic, results are applicable to pathogenic examples such as Aβ and α-synculein, where parallels in data and mechanisms have already been observed (Evangelisti *et al*., 2012 Tatini *et al*., 2013). Developing our understanding of oligomer formation, and the key factors that define toxicity has major beneficial applications for pathogenic amyloidogenic proteins. It is through improving our knowledge on the molecular mechanisms behind protein aggregation that allows for better therapeutic advancements (Hammarström *et al*., 2003).

Factors affecting the aggregation of HypF-N such as the kinetics and contributing factors such as ionic strength and pH are well studied (Bucciantini *et al*., 2002, Relini *et al*., 2004, Marcon *et al*., 2005, Canale *et al*., 2006, Campioni *et al*., 2008, Ahmed *et al*., 2010). Additionally research has elucidated the cellular responses and effects from oligomer toxicity, drawing similarities to pathogenic oligomers Aβ and α-synculein (Cecchi *et al*., 2005, 2006, Zampagni *et al.,* 2011, Evangelisti *et al*., 2012). However, there are still unknown factors relating to HypF-N aggregation. The two oligomers of HypF-N dictate different toxicities, but we do not know the detailed structural differences between them. Understanding the structural elements of the protein under each condition could help understand the process of aggregation and nucleation; most structural studies on HypF-N aggregation have been performed using far UV CD, this technique does not provide the atomistic and conformational detail achievably through NMR. NMR is a powerful technique with the potential to study the transient lowly populated conformations of proteins, particularly unfolded, or partially unfolded ensembles, making this an ideal method to examine HypF-N under the non-toxic condition.

The nucleation pathway has not been studied on an unmodified protein structure and sequence, furthermore the interactions within the membrane are not fully understood. This project will use advanced NMR based methods to develop our understanding on the protein system, elucidating the origins of toxicity, supplementing what is already known. Specifically, the aggregation and nucleation process will studied using an innovative NMR technique Dark State Exchange Saturation (DEST) (Fawzi *et al*., 2011), as well as the interaction between toxic and non-toxic oligomers with membrane vesicles using non-purturbing ssNMR methods previously shown to successful for α-synculein (Fusco *et al*., 2014, 2016). Elucidating differences in the origin of toxicity, and oligomer interactions with GM1 on a molecular level will help develop our understanding of cytotoxic amyloid formation. A more detailed understanding of the mechanisms behind oligomer interactions at the cell membrane will help identify potential targets for drug design and therapeutic techniques in pathogenic proteins.

Our work looks to use firstly solution NMR methods as a non-invasive tool for studying the nucleation of monomeric HypF-N under each condition, and ssNMR to study the differences in topology of the oligomers. These studies will not require specific solvent conditions or chemical modifications of the sequence that may affect the natural progression and formation of HypF-N oligomers. The project specific aims are listed below:

- Expression and purification of isotopically labelled HypF-N for NMR studies
- NMR characterisation of HypF-N under oligomerizing conditions
    - ^{1}H-^{15}N HSQC backbone assignment under toxic and non-toxic condition
- Determine regions and residues of the sequence involved in the nucleation step of oligomer formation in both toxic and non-toxic condition using DEST
- Preliminary characterisation of the oligomer using ssNMR methods
- Studying the interaction between oligomer and lipid vesicles as model cell membranes using ssNMR methods
- Observing effects of GM1 and cholesterol on oligomer interactions in lipid vesicles

CHAPTER SIX

HYPF-N

METHODS AND MATERIALS

Chemical suppliers

All chemicals and substances were purchased from Sigma Aldrich unless otherwise stated.

6.1 Constructs and strains

6.1.1 Construct design

The HypF-N gene was cloned into a pQE30 vector modified to contain an N terminal thrombin cleavage site preceding the 6-HIS purification tag (pQE30-Th). The construct was made and generously provided by Prof. F. Chiti, Florence. The vector carries a T5 promoter and is inducible under IPTG, a 6HIS tag for purification across nickel resin and a thrombin cleavage site to cleave the 6HIS tag from the protein.

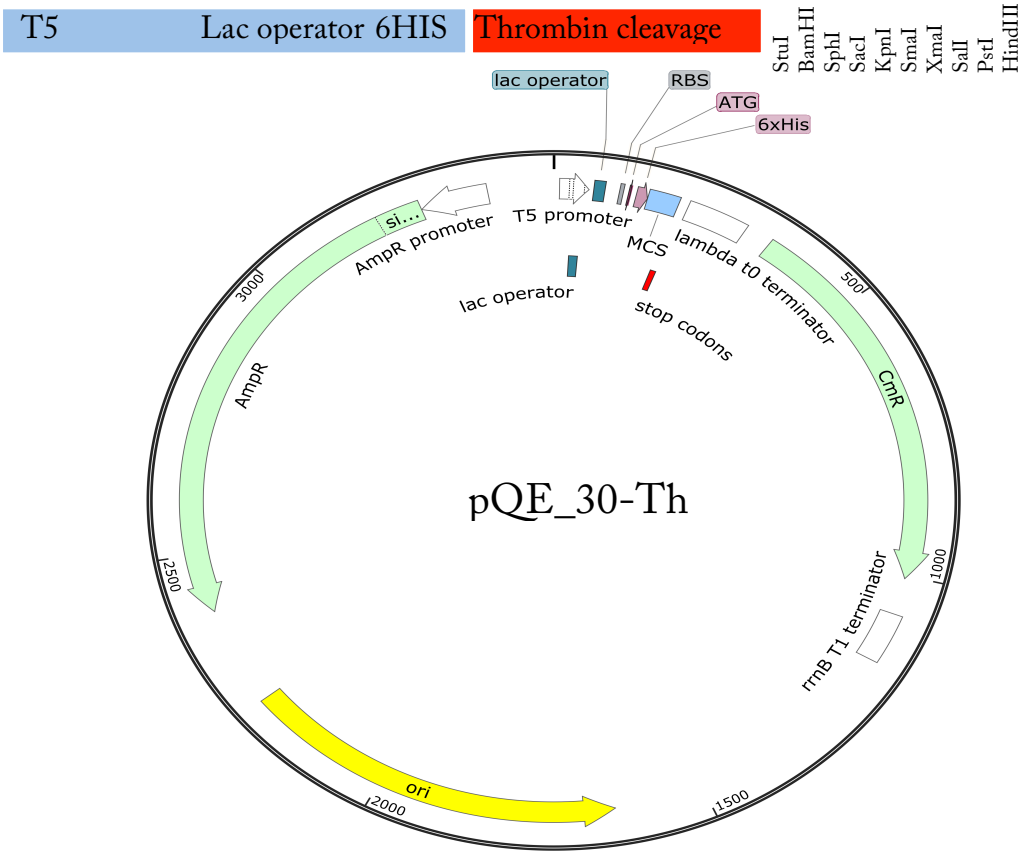


Figure 6.1 pQE30 vector map

pQE30 vector map illustrating restriction enzyme sites and major features of the vector. Figure made by Snapgene

6.1.2 Strain

The cell strain XL1-Blue (New England Biolabs) was used for DNA amplification and storage, XL1-Blue has a genotype mutated for DNA work rather than protein expression with a genotype containing *recA1* for reducing recombination of DNA and *hsdR17* which prevents cleavage of cloned DNA insert leading to more stable insert and *endA1* for endonuclease deficiency improving quality of miniprep DNA.

The *E. coli* cell strain M15 (Qiagen) was used to express HypF-N, M15 cells lack the M15 in the *lacZ* gene and complemented by the α–fragment of the β galactosidase, and is under T5 promoter induction from pQE vectors. It is $recA^+$ for reduced DNA recombination. It is resistant to infection of lambda phages.

6.1 Expression and purification

6.2.1 Expression of unlabelled HypF-N

Expression and purification of HypF-N was performed by methods previously published (Chiti *et al*., 2001).

5 µl of HypF-N plasmid DNA (donated by Prof. F. Chiti, Florence) was combined with 20 µl of M15 [PREP4] cells (Qiagen) and left on ice for 30 min with periodic gentle agitation, after which it was subjected to 42 °C for 45 seconds and returned on ice for a further 10 min. 500 µl of pre-warmed LB was added and the cells were left shaking at 200 rpm at 37 °C for one hour; 150µl of which was plated under sterile conditions on LB agar plates with 50 mg/ml carbenicillin antibiotic resistance. Plates were incubated at 37 °C overnight.

Several colonies of transformed M15 [PREP4] cells were picked and grown overnight at 37 °C in a pre-culture of 200 ml LB with 50 mg/ml carbenicillin. The pre-culture was diluted 1:10 into LB and grown at 37 °C until its OD_{600} reached ~ 0.6, where it was induced with 1 mM IPTG and left overnight at 25 °C.

6.2.2 Purification

Overnight induced cells were pelleted by centrifugation at 8,000 *g* for 15 min at 4 °C, the resulting cell pellet was lysed in 50 mM sodium phosphate, 300 mM sodium chloride, 10mM imidazole, pH 8.0. The cell lysis was sonicated using a Sonoplus HD 2070 sonicator and KE76 probe (Sonoplus, Bandelin) for 5 min at a cycle of 30 seconds on and 30 seconds off at 50% amplitude. The sample was then centrifuged at 17,000 *g* for 45 min at 4 °C. At 4 °C the clear lysate was filtered using a 0.45 μm filter and passed over a $Nickel^{2+}$-NTA resin (Qiaging) column equilibrated with 50 mM phosphate and 50 mM sodium chloride at pH 8.0. The clear lysate was passed over the column four times to ensure saturation of the nickel binding sites. The resin was washed with 50 mM sodium phosphate, 300 mM sodium chloride, 20 mM imidazole at pH 8.0 before 50 units of thrombin in 50 mM phosphate and 50 mM sodium chloride at pH 8.0 was added in a 500 μl volume. The resin was left to incubate at 37 °C for one hour, and moved to 4 °C overnight. 1 ml fractions were eluted using 50 mM sodium phosphate, 50 mM sodium chloride, 10 mM imidazole at pH 8.0 and run on 15% SDS PAGE gel for analysis. Eluted fractions containing cleaved HypF-N were combined and concentrated to a 10 ml sample.

This was dialysed with a 3 kDa MW cut off dialysis membrane against 5 mM sodium acetate, 2 mM dithiothreitol (DTT), pH 5.5 using acetic acid overnight at 4 °C.

6.2.3 Expression of ^{15}N and $^{15}N^{13}C$ labelled HypF-N

Multiple colonies of transformed cells were grown overnight at 37 °C in a pre-culture of 100 ml M9 media with 50 μg/ml carbenicillin, with labelled compounds as described in section 3.5.2. The pre-culture was diluted 1:50 into M9 minimal media and grown at 25 °C until its OD_{600} reached ~ 0.8, where it was induced with 1 mM IPTG and left overnight at 25 °C. Labelled HypF-N purification was carried out as described for unlabelled LB media.

6.4 Oligomer formation

Toxic pre-fibrillar oligomers were formed by incubating HypF-N at a concentration of

48 μM in 12% 2,2,2- trifluoroethanol (v/v) (TFE) (VWR), 50 mM sodium acetate, 2 mM DTT, pH 5.5 for 4 hours at 25 °C.

Non-toxic pre-fibrillar oligomers were formed by incubating HypF-N at a concentration of 48 μM in 20 mM trifluoroacetic acid (TFA), 330 mM sodium chloride, pH 1.7 for 4 hours at 25 °C.

6.5 Tapping Mode – Atomic Force Microscopy (TM-AFM)

Oligomers were prepared as detailed above. For optimised imaging toxic oligomers were diluted 5 times into distilled H_2O. Non-toxic oligomers were diluted 500 times into distilled H_2O. 10 μl drops were placed on freshly cleaved mica substrate (Agar Scientific) and air-dried. TM-AFM measurements were collected in intermittent contact (air) with a drive frequency of 112 kHz and line rate of 1.0 Hz. Measurements were collected on a JPK NanoWizard II instrument (Germany). An 8 nm tip radius silicon cantilever with 40° cone angle and tip height of 15 μm was used (μmasch). Data was analysed using Gwyddion software. Oligomer sizes were calculated from the height in cross section of the topographic AFM images. Since the samples were dried under air and not vacuum no shrinking factors were applied.

6.6 NMR

6.6.1 Solution NMR

Solution NMR for HypF-N was performed on a triple resonance (^{1}H, ^{15}N, ^{13}C) 5.0 mm probe on a 600 MHz Bruker Avance spectrometer.

1D ^{1}H experiments were formed using the same pulse sequence and water suppression described in section 3.10.2.

	^{1}H (F1)
SW	16.019
Time Domain	8192

Acquisition time (sec)	0.106
Acquisition mode	DQD

2D ^{1}H-^{15}N HSQC experiments were recorded at 298 K unless otherwise stated.

	^{1}H (F2)	^{15}N (F1)
SW	14.03	30.00
Time Domain	2048	256
Acquisition time (sec)	0.121	0.035
Acquisition mode	DQD	Echo-Antiecho

6.6.2 Backbone Assignment

6.6.2.1 Native

Standard NMR protocols for assigning protein backbone ^{13}Cα, ^{13}Cβ, ^{13}CO, ^{15}NH, ^{1}Hα nuclei were used. Complementary 3D spectra HNCACB, HNCOCACB, HNCO and HNCACO, HNHA were collected on HypF-N under its native condition at a concentration of 250 μM in 2 mM DTT, 5 mM NaAc, pH 5.5 at 298 K.

The acquisition times are listed below.

Experiment	^{1}H (F3)	^{15}N (F2)	^{13}C (F1)
HNCACB	0.121	0.024	0.006
HNCOCACB	0.121	0.024	0.006
HNCO	0.106	0.024	0.108
HNCACO	0.106	0.010	0.037
HNHA	0.106	0.008	0.027

6.6.2.2 Toxic condition

The majority of the protein backbone ^{15}NH nuclei in the toxic condition were assigned by transferring the native assignment to the toxic condition, where peak changes could be easily tracked in the ^{15}N HSQC. To confirm assignment and measure ^{13}Cα,

$^{13}C\beta$, ^{13}CO, $^{1}H\alpha$ nuclei resonances complementary 3D spectra HNCACB, HNCOCACB, HNCO and HNCACO, HNHA were collected on HypF-N a concentration of 50 μM under the toxic condition in 10% TFE (v/v), 2 mM DTT, 5 mM NaAc, pH 5.5 at 283 K.

Acquisition times were same as listed for the native condition (section 6.6.2.1)

6.6.2.3 Non-toxic condition

NMR backbone assignment of the non-toxic condition was performed by denaturing the protein by adding 8M Urea to native buffer conditions at pH 1.7 (5 mM NaAc, 2 mM DTT pH 5.5). The acquisition times are listed below for each 3D experiment.

Experiment	^{1}H (F3)	^{15}N (F2)	^{13}C (F1)
HNCACB	0.106	0.024	0.005
HNCOCACB	0.106	0.024	0.005
HNCO	0.106	0.024	0.108
HNCACO	0.106	0.024	0.108
HNHA	0.106	0.010	0.037

The non-toxic condition required further data collection using HNCOCANH (Sun *et al*., 2005). HSQC spectra were collected before and after to check for any changes in protein state. Details of the HNCOCANH experiment are listed below:

	^{1}H (F3)	^{15}N (F2)	^{15}N (F1)
SW	13.94	20.00	20.00
Time Domain	2048	1441	146
Acquisition time (sec)	0.09	0.04	0.04
Acquisition mode	DQD	Echo-Antiecho	States-TPPI

6.6.3 DEST

A more in-depth explanation into DEST is provided in section X.X. The experimental parameters used are described below.

6.6.3.1 Toxic condition

DEST experiments were performed on samples with 10% TFE, 50 mM NaAc, 2 mM DTT, 100 μM HypF-N, pH 5.5 at 283.15 K on a 600 MHz Bruker Avance spectrometer. 2D ^{15}N-DEST spectra were collected in sets of 5 interleaved experiments with the weak ^{15}N weak saturation pulse applied at varying RF offset and a standard reference set at 10,000 kHz for 0.9 s. RF offsets were collected at +/- 0.5, 1, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 7.5 and 10.0 kHz, the offset is from the ^{15}N carrier frequency set at 118.0 ppm at the centre of the monomer spectrum. RF field strengths were set to 300 Hz in all experiments. Experiments were performed with 24 scans.

	^{1}H (F2)	^{15}N (F1)
SW	16.02	20.00
Time Domain	2048	220
Acquisition time (sec)	0.106	0.090
Pulse length (μs)	7.35	
Acquisition mode	DQD	Echo-Antiecho

6.6.3.2 Non-toxic condition

DEST experiments were performed on a sample with 2 mM TFA, 50 mM NaCl, 120 μM HypF-N, pH 1.7 at 283.15 K on a 600 MHz Bruker Avance spectrometer. Due to rapid aggregation and loss of signal ^{15}N-DEST spectra were collected in sets of 2 interleaved experiments with the weak ^{15}N weak saturation pulse applied at varying RF offset and a standard reference set at 10,000 kHz for 0.9 s. RF offsets were collected at +/- 0.5, 1, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 7.5 and 10.0 kHz, the offset is from the ^{15}N carrier frequency set at 118.0 ppm at the centre of the monomer spectrum. RF field strengths were set to 300 Hz in all experiments. Experiments were performed with 64 scans.

	^{1}H (F2)	^{15}N (F1)
SW	16.02	20.00
Time Domain	2048	220
Acquisition time (sec)	0.106	0.090
Pulse length (μs)	11.75	
Acquisition mode	DQD	Echo-Antiecho

6.6.4 Solid state NMR

Calibration of the 54.7° magic angle was achieved by overlaying the spectrum obtained from a potassium bromide standard. The probe was shimmed before each spectrum to ensure optimum measurements, 90° pulses on ^{1}H, ^{15}N, ^{13}C nuclei and cross-polarization steps between ^{1}H - ^{13}C and ^{1}H - ^{15}N pulses were calibrated and verified for each sample. All experiments were performed at a MAS spinning frequency of 10 kHz at 277.15 K. A triple resonance (^{1}H, ^{15}N, ^{13}C) 3.2mm CP/MAS probe was used on a 600 MHz Bruker Avance spectrometer.

6.6.4.1 Oligomer only samples for ssNMR

Oligomers were prepared as previously reported in section 6.4, and pelleted at 16,100 *g* for 10 min. The pellet was dried under a N_2 stream and packed into a thin-walled 3.2mm MAS rotor (Bruker) by centrifugation at 3750 *g*.

6.6.4.2 Lipid SUV formation

Lipids (Avanti Polar Lipids Inc.) dissolved in chloroform were measured by volume into a glass vial and gently dried under a stream of N_2 to evaporate off all chloroform. DOPC:SM lipids were used in a molar ratio of 2:1. The lipids were then lypholized for a minimum of one hour to ensure complete chloroform removal. The lipids were then re-hydrated in an aqueous Dulbecco's Modified Eagle's Medium (DMEM) buffer, along with any additional substrate such as cholesterol or ganglioside GM1 lipid added as a powder. The sample was placed in a water bath sonicator for 2 min to homogenize the sample. The solution was transferred to an eppendorf and sonicated at

10% power for 3 min cycles until the solution was transparent, and then centrifuged at 17,000 *g* for 30 min at 4°C to pellet residual metal from the sonicator tip. The lipids then under went 5 rounds of freeze-thaw cycles in liquid nitrogen and under warm water to form small unilammelar vesicles (SUV).

6.6.4.3 Oligomer + Lipid samples

SUVs formed under section 6.6.4.2 were added to dried oligomer pellets formed in section 6.4 and mixed by pipetting until the oligomers were completely re-suspended in solution. The mixture was pelleted by ultracentrifugation at 70,000 *g* for 90 min at 4°C using a Beckman Coulter centrifuge and a TLA 100.2 fixed angle rotor. The supernatant was quickly removed and the pellets dried again under N_2. The pellets were then transferred to a thin-walled 3.2mm MAS rotor (Bruker) by centrifugation at 3750 *g*.

6.6.4.4 PRE lipid samples

2% molar ratio of the spin-labelled lipids in chloroform or powder were added to the mix of chloroform mix of lipids (DOPC:SM in a 2:1 ratio). Additional components such as GM1 or cholesterol were added whilst the sample was in solution in chloroform. SUVs were prepared in the same manner described in section 6.6.4.2.

6.6.4.5 DARR

Samples were packed into a thin-walled 3.2mm MAS rotor (Bruker). For ^{13}C-^{13}C correlation experiments using DARR mixing MAS frequencies of 10 kHz were used. Experiments were performed with 20 and 100ms DARR mixing times to observe short-range and long-range contacts, respectively. Experiments were performed with 64 scans per t_1 increment and 440 complex increments. 100 kHz ^{1}H decoupling was achieved using swftppm decoupling sequence with a recycle time between experiments set at 3s.

	^{13}C (F2)	^{13}C (F1)
SW	301.23	256.08
Time domain	2048	440
Acquisition time (sec)	0.011	0.005
Pulse width (μs)	4.67	
Acquisition mode	DQD	States TPPI

6.6.4.6 INEPT

Samples were packed into a thin-walled 3.2mm MAS rotor (Bruker). MAS frequencies of 10 kHz were used. Experiments were performed with 128 scans per t_1 increment and 180 complex increments. A proton SPINAL64 decoupling was applied at $\omega_{RF}/(2\pi)$ = 104 kHz.

	^{1}H (F2)	^{13}C(F1)
SW	301.25	8.50
Time domain	1024	180
Acquisition time (sec)	0.011	0.015
Pulse width (μs)	2.4	3.3
Acquisition mode	DQD	States TPPI

CHAPTER SEVEN

HYPF-N:

RESULTS AND DISCUSSION

7.1 Expression and Purification

The protocol for expression and purification of unlabelled and labelled HypF-N was kindly given by collaborator Prof. F. Chiti and are published in Chiti *et al*., 2001 and Calloni *et al*., 2008 respectively. Briefly, the pQE_30 vector containing the gene for HypF-N with an N′ 6HIS tag is transformed into M15[PREP4] cells. Cells are grown in either rich LB media or M9 media for unlabelled or isotopically labelled of ^{15}N or ^{13}C nuclei respectively. HypF-N is expressed in the cytoplasm of M15[PRE4] cells after the pQE_30 T5 promoter is induced with addition of IPTG. Cells are lysed by sonication and the insoluble fractions are removed by centrifugation. The soluble fraction is loaded onto a Ni^{2+}-NTA resin to which the 6HIS tag of HypF-N binds. The 6HIS tag is cleaved by incubating with thrombin to allow elution of pure untagged HypF-N. Thrombin is removed and buffer is exchanged into optimal storage conditions using gel filtration. The conditions for expression and purification have been optimised and the protein expresses in good yield for NMR requirements; ~12 and 9 mg/L for unlabelled and double labelled cultures (^{13}C and ^{15}N) respectively.

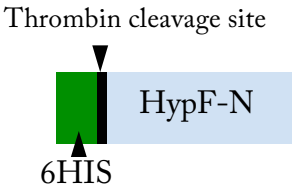


Figure 7.1 Diagram representation of HypF-N construct

Figure drawn starting the N terminus

The expression of unlabelled HypF-N in rich LB media is briefly presented below to demonstrate the expression yields and purification success using the methods outlined. No changes in the expression or purification were noticed by changing to M9 cultures for expression of isotopically labeled ^{15}N or $^{13}C^{15}N$ labelled HypF-N.

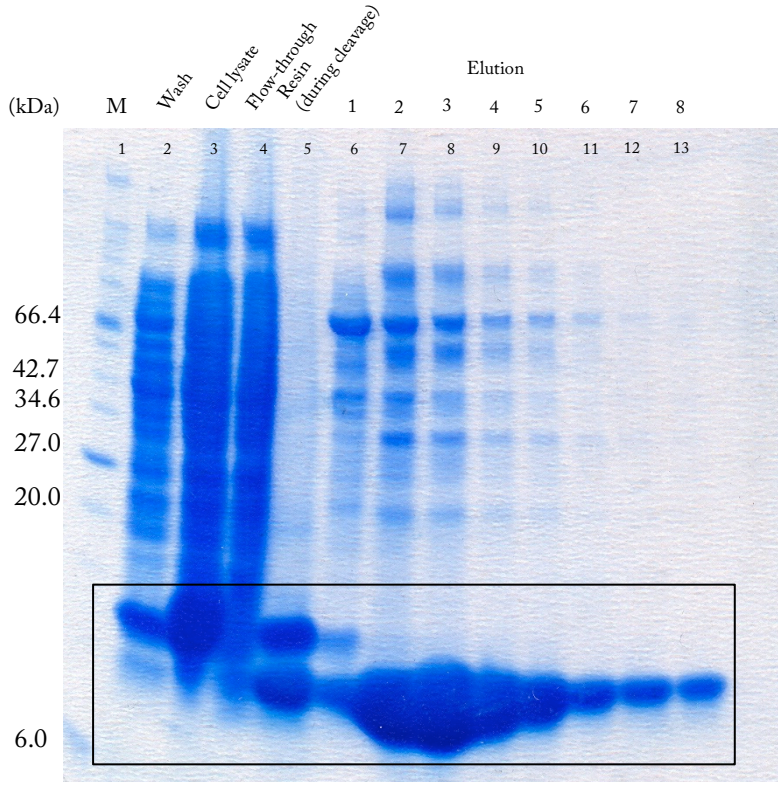


Figure 7.2 Expression of HypF-N

12% SDS-PAGE gel of Nickel purification of HypF-N protein. M15[prep4] cells grown with the pQE_30_HypF-N plasmid were induced with 1 mM IPTG and expressed at 25 °C overnight. Cells were pelleted and lysed by sonication. The clear lysate after centrifugation was loaded onto a Ni^{2+}-NTA resin. The column was washed and incubated with thrombin for 1 hour at 37 °C then overnight at 4 °C. Cleaved HypF-N was eluted.

It is evident that some HypF-N is eluted during the wash; using a lower concentration of imidazole in the wash buffer can reduce this loss of protein. The elution shows strong band for HypF-N indicating good over expression of the recombinant protein.

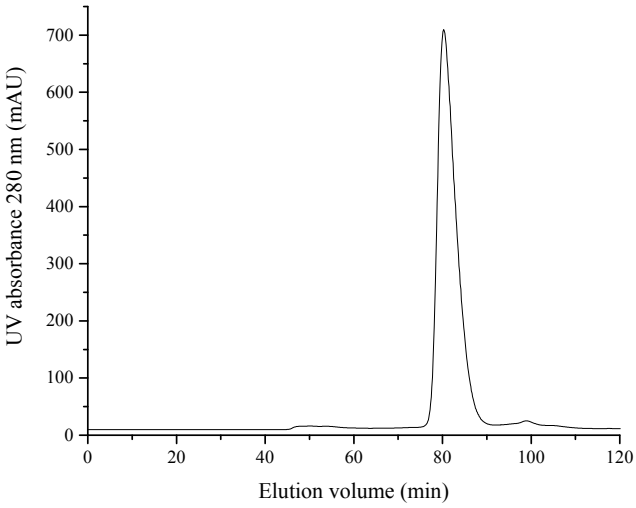


Figure 7.3 SEC gel filtration profile of HypF-N

Fractions collected from the nickel elution were run on a SEC-gel filtration column to remove thrombin and contaminating proteins following nickel purification and allow buffer exchange into suitable storage buffer of HypF-N.

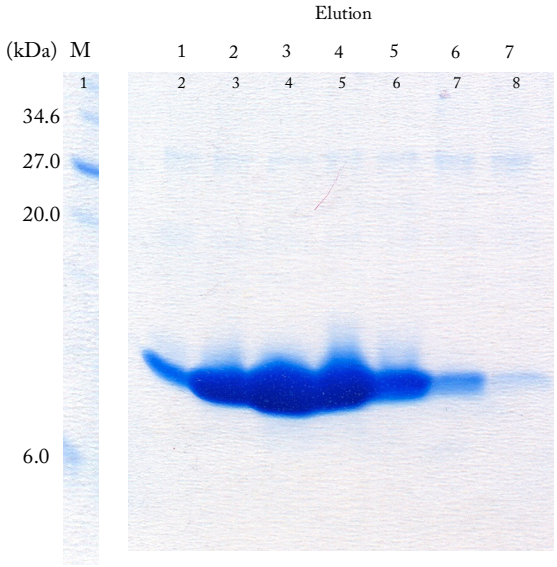


Figure 7.4 12% SDS-PAGE gel of SEC-GF purification of HypF-N

Fractions taken from the single main peak of the gel filtration shown in Figure 7.3 were run on a 12% SDS-PAGE to assess purity.

The SEC gel filtration chromatogram shows a single symmetric peak, containing only HypF-N as confirmed by SDS-PAGE analysis (Figure 7.4). The sample is 95% pure based on the SDS-PAGE and suitable for NMR studies.

7.2 Solution-state NMR

7.2.1 NMR characterisation

The first step in the amyloidogenesis pathway outlined in Figure 5.4 is the monomeric state of the protein. To study this ^{15}N-HSQC spectra were collected of HypF-N under the native, toxic and non-toxic solution conditions. The ^{15}N-HSQC provides an informative insight into the conformational state of proteins.

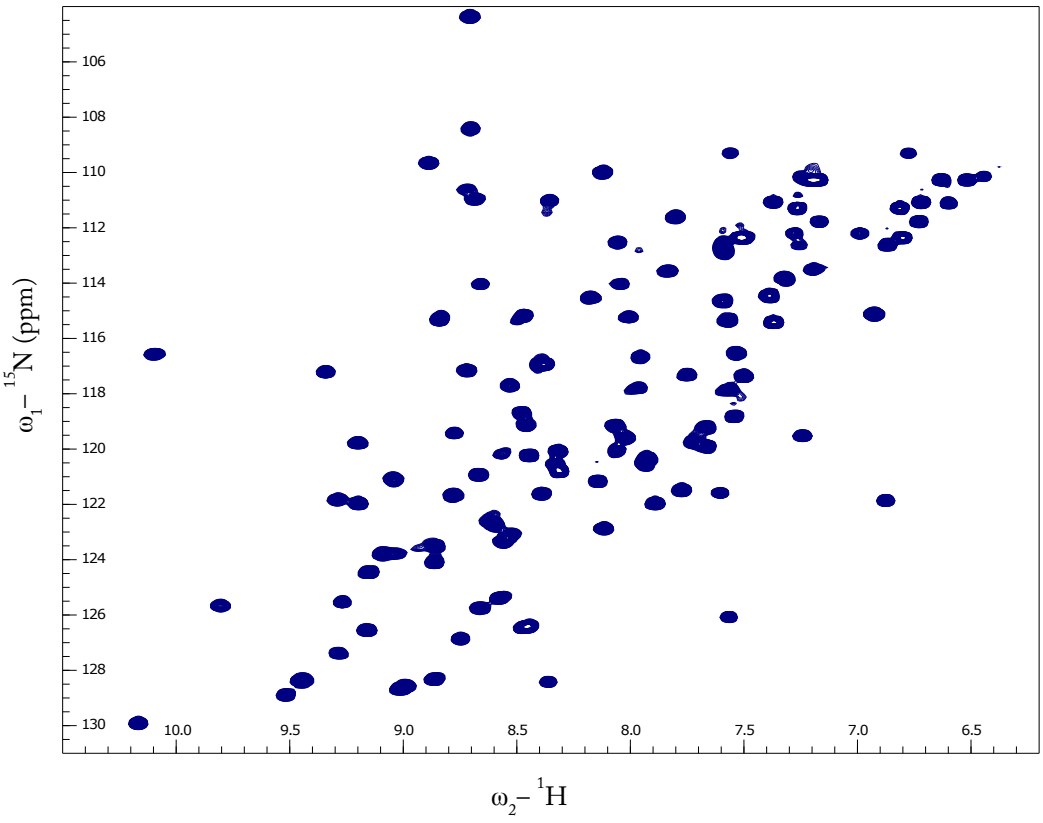


Figure 7.5 2D ^{15}N-HSQC of HypF-N

Spectrum was collected at 200 µM HypF-N in 5 mM NaAc, 2 mM DTT pH 5.5 at 298 K on a 600 MHz spectrometer.

HypF-N in the native condition gives dispersed peaks, indicative of a folded protein. The native buffer for HypF-N (5 mM NaAc, 2 mM DTT pH 5.5) is optimal for NMR given the low ionic strength and slightly acidic pH values (Kelly *et al*., 2002). 81 backbone peaks are observed within the ^{15}N-HSQC accounting for most of non-

proline amide cross peaks, of which there are 87. Some residues were not detected until after addition of 50 mM sodium phosphate to stabilise the flexible regions of the protein, as discussed further in section 7.2.3.

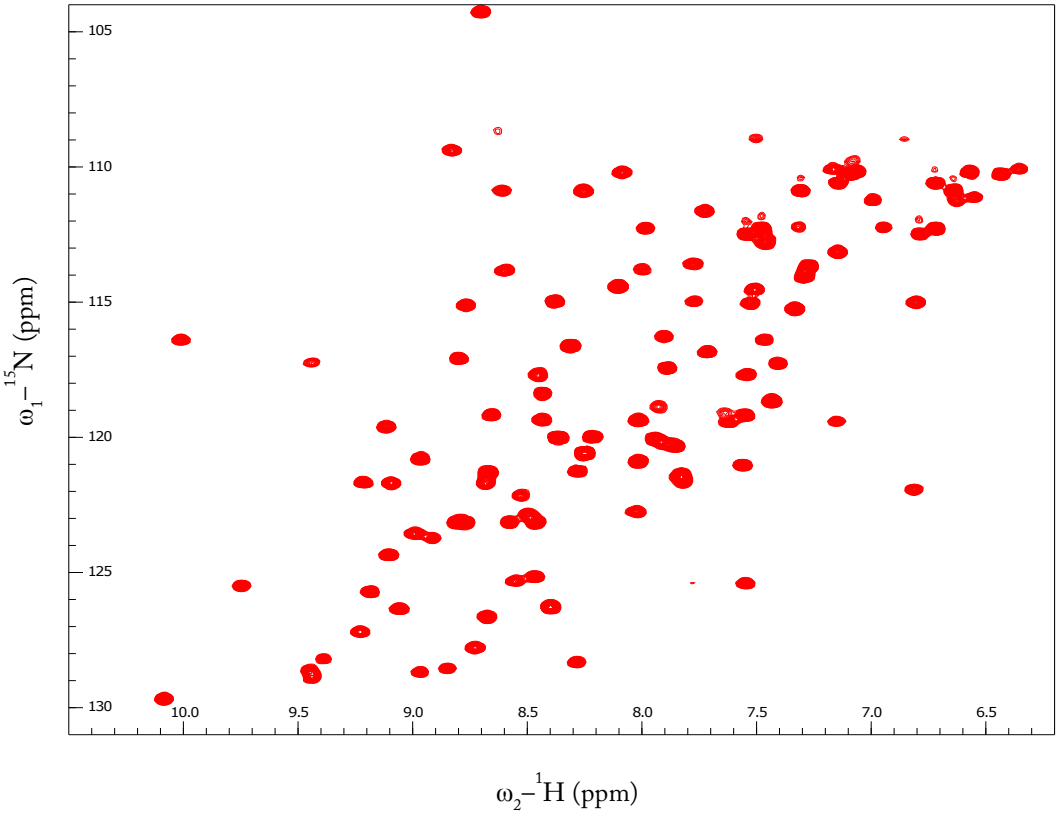


Figure 7.6 2D ^{15}N-HSQC of HypF-N under toxic buffer condition

Spectrum was collected by diluting HypF-N in the native condition to 48 μM in 50 mM NaAc, 2 mM DTT, 12% TFE (v/v) pH 5.5 at 298 K on a 600 MHz spectrometer.

HypF-N under the toxic condition (12% TFE (v/v), 50 mM NaAc, 2 mM DTT pH 5.5) remains folded and exhibits a conformation similar to that of its native state. Despite the presence of 12% TFE and increase in ionic strength the ^{15}N HSQC spectrum for HypF-N under the toxic condition is very similar to the native state. Slight global peak shifts are noted in the ^{1}H dimension due to the presence of TFE within the sample (Figure 7.6). TFE is known to strengthen hydrogen bonding of secondary structures, particularly α–helices but also β-sheets (Goodman *et al*., 1971), increasing the overall stability of proteins. TFE has often been used as a cosolvent in NMR samples to record measurements for marginally stable proteins, given the increased resolution of resonances in presence of TFE (Merutka *et al*., 1995). Whilst

the secondary structure is stabilized, the tertiary and quaternary structure may be negatively affected by TFE disrupting hydrophobic interactions (Slupsky *et al*., 1995). From 2D ^{15}N-HSQC, HypF-N under the toxic condition appears to be very similar and largely unchanged from its native conformation. The spectrum for both native and toxic conditions depicts well dispersed peaks indicative of a folded protein, where the residues are distinctly folded and are each within their own chemical environment. Given that many chemical shifts overlap between the native and toxic, we can infer that HypF-N exhibits a similar folded structure in both conditions.

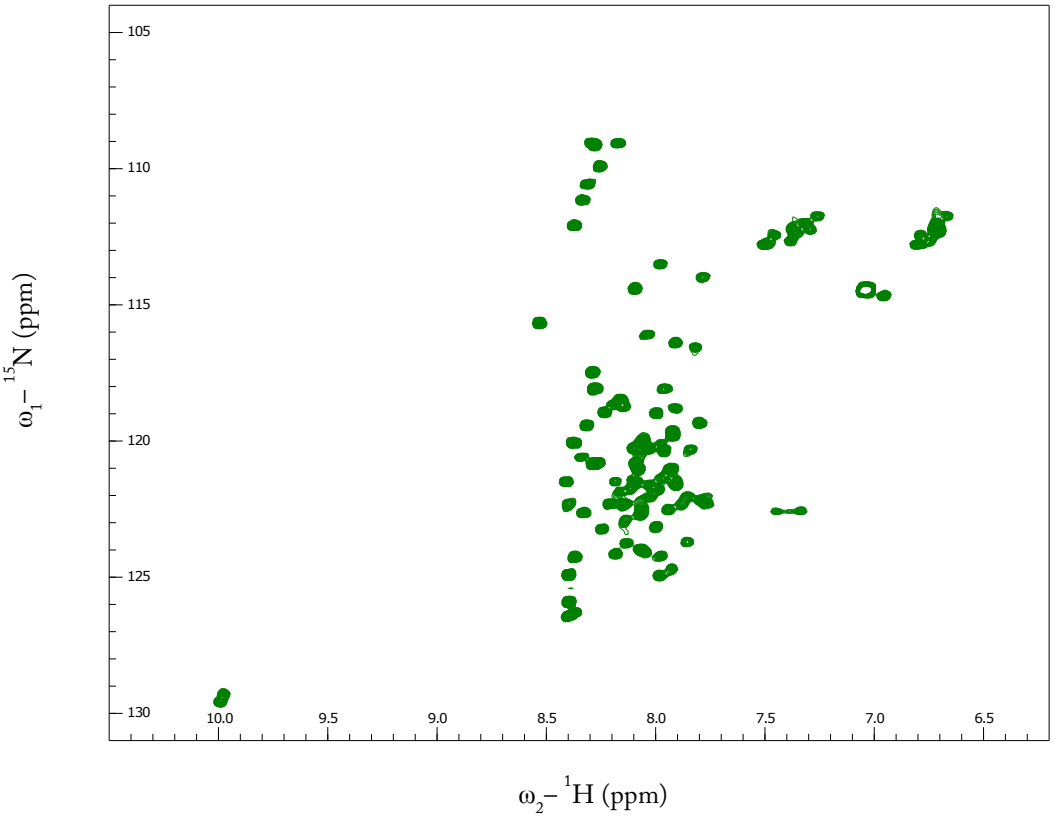


Figure 7.7 2D ^{15}N-HSQC of HypF-N under non-toxic buffer condition

Spectrum was collected by diluting HypF-N in the native condition to 48 µM in 330 mM NaCl, 20 mM TFA pH 1.7 at 298 K on a 600 MHz spectrometer.

The spectrum for HypF-N under the non-toxic buffer condition (20 mM TFA, 330 mM NaCl, pH 1.7) is indicative of an unfolded, unstructured protein. This represents a conformation of the protein where all the residues experience a similar chemical environment, and therefore exhibit similar chemical shifts. This is identified by the collapse of the dispersed peaks observed in the native folded conformation to a

narrow band centered ~ 8 ppm in the proton dimension. The well dispersed peaks noted in the native and toxic condition is not the case for unfolded proteins.

The conformation of HypF-N under the non-toxic condition has been characterised as a partially unfolded, pre-molten globule state using far UV CD and NMR. At low pH HypF-N is in an ensemble state of fluctuating conformations composed with elements of nonrandom structure, some residual structure is maintained (Campioni *et al*., 2008, Calloni *et al*., 2008).

Whilst the high ionic strength leads to increases in structure in the unfolded state as measured by far UV CD (Campioni *et al*., 2008); the TFA present acts as very strong chemical acidic denaturant creating a very acidic pH environment, which induces the unfolding. The spectrum shown in Figure 7.7 surprisingly depicts a high number of well-resolved peaks. Previous studies for HypF-N under the same condition did not yield the same number of observed peaks (Calloni *et al*., 2008). These results were achieved initially in the PhD, but could not be replicated; despite much of the project being focused in replicating similar NMR ^{15}N-HSQC spectra.

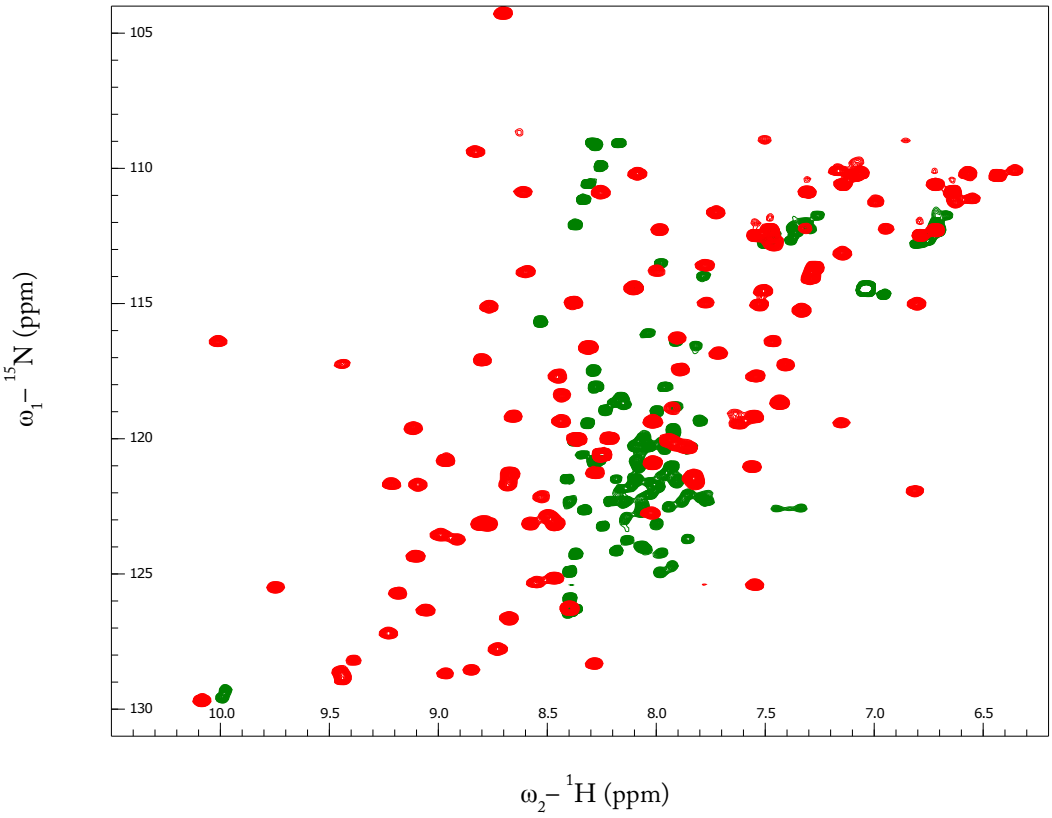


Figure 7.8 Monomeric HypF-N exhibits different level of structure and conformation in toxic and non-toxic conditions

Spectrum was collected by diluting HypF-N in the native condition to 48 μM in 50 mM NaAc, 2 mM DTT, 12% TFE (v/v) pH 5.5 for the toxic (red) and to 48 μM in 330 mM NaCl, 20 mM TFA pH 1.7 for the non-toxic condition. Spectrum recorded at 298 K on a 600 MHz spectrometer.

At the monomeric state, HypF-N exhibits major differences in its structural state under the two different conditions. Figure 7.8 illustrates the differences between folded (toxic; red) and unfolded (non-toxic; green) ^{15}N HSQC, and the dispersion of peaks, or lack of, in each condition. This acts as the first difference between the two conditions under the amyloidogenic pathway presented in Figure 5.4.

7.2.2 Backbone assignment

To probe the conformation of the protein by NMR on a residual level each backbone peak within ^{15}N-HSQC spectra must be assigned. This is achieved by collecting ^{13}C nuclei resonances using 3D NMR experiments which allows for determination for individual amino acid based on specific ^{13}Cα, ^{13}Cβ and ^{13}CO resonances. The position of the residue in the sequence is determined by analysis of the previous residues connected in the polypeptide chain, to allow for sequential assignment. The ^{15}N-HSQC peaks for HypF-N under the three solution conditions (native, toxic and non-toxic) were assigned.

HypF-N was assigned under all conditions using CcpNmr Analysis (Vranken *et al*., 2005), given the size of the protein manual peak picking and assignment was performed with unambiguous assignment confirmed using MARS analysis (Jung & Zweckstetter, 2004) with a high confidence across the entire sequence. Figure 7.9 illustrates the backbone sequential assignment by connecting ^{13}Cα, ^{13}Cβ and ^{13}CO resonances between *i* and *–i* states.

7.2.3 Native

Backbone assignment of HypF-N under its native conformation was achieved using standard 3D pulses HNCACB, HN(CO)CACB, HNCO and HN(CA)CO. ^{15}N-HSQC peaks were assigned in a sequential manner based on ^{13}Cα, ^{13}Cβ and ^{13}CO resonances. ^{15}N-HSQC experiments were performed before and after each 3D spectrum pair and showed no change in protein conformation.

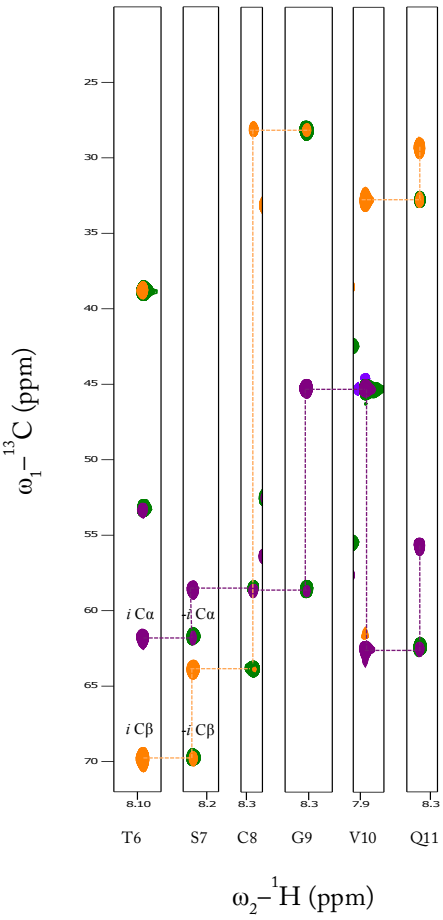


Figure 7.9 1H strips for CACB and COCACB spectra collected for assignment of HypF-N

Assignment for residues amino acid stretch Thr6-Ser7-Cys8-Gly9-Val10-Gln11. CACB is shown with Cα peaks connected in purple and Cβ connected in yellow. HN(CO)CACB peaks are shown in green. Spectra taken from HypF-N under the urea induced unfolded state since signal was greatest for this condition for ease of visualization. The stripes shows are representative of the quality of data observed throughout the 3D spectra.

In HNCACB experiments the transfer of magnetization starts from the ^{1}Hα to ^{13}Cα and ^{1}Hβ to ^{13}Cβ then ^{13}Cα. From here it is transferred, bypassing the ^{13}CO, to the ^{15}N and detected at the ^{1}HN. Given the bypass of ^{13}CO, each ^{1}HN receives ^{13}Cα and ^{13}Cβ of the intraresidue (*i*) and the preceding residue (i_{-1}) (Grzesiek & Bax 1992b). In HN(CO)CACB a signal is detected from the chemical shift evolved from the transfer of magnetization from the ^{1}Hα and ^{13}Cα and ^{1}Hβ to ^{13}Cβ then ^{13}Cα. From here it is transferred via the ^{13}CO to the ^{15}N and detected at the ^{1}HN. The resulting signal provides connectivity information between the amide ^{1}HN and ^{15}N of the intraresidue

i and, through *J* coupling between the backbone and side chain resonances, both the ^{13}Cα and ^{13}Cβ of the preceding residue *i-1* (Grzesiek & Bax 1992a). Complementary to these experiments HNCO and HN(CO)CA experiments were performed for ^{13}CO chemical shifts, these additional spectra help resolve ambiguities caused by ^{1}Hα-^{13}Cα pair chemical shift degeneracies. HNCO is an extremely sensitive triple resonance experiment that transfers magnetization from the ^{1}HN to ^{15}N, and selectively to the ^{13}CO via ^{15}NH- ^{13}CO *J* coupling. The ^{13}CO chemical shift of the preceding residue *i-1* is detected as the magnetization is passed back to the ^{1}HN via the ^{15}N (Kay et al., 1990). Analogous to the HN(CO)CACB, the HN(CO)CA provides ^{13}CO chemical shift values for the intraresidue *i* and the preceding residue *i-1*. Magnetization originates from the ^{1}HN to ^{15}N and then via the ^{15}NH- ^{13}Cα *J* coupling to the ^{13}Cα, from here it is transferred to the ^{13}CO via the ^{13}Cα-^{13}CO *J* coupling. It follows the same route back to the ^{1}HN for detection. Since the amide ^{15}N is coupled to *i* and *i-1* residues we detect ^{13}CO chemical shifts from both *i* and *i-1* residues. Intraresidue transfer is often stronger, resulting in a more intense ^{13}CO *i* peak than for ^{13}CO *i-1* (Kay, Ikura, Tschudin 1990). The combination of all spectra reliably allows for sequential assignment of ^{15}N and ^{13}C isotopically enriched proteins using ^{13}Cα, ^{13}Cβ and ^{13}CO resonances.

The backbone assignment for HypF-N native condition is referenced in the Supplementary Material in Table S1.

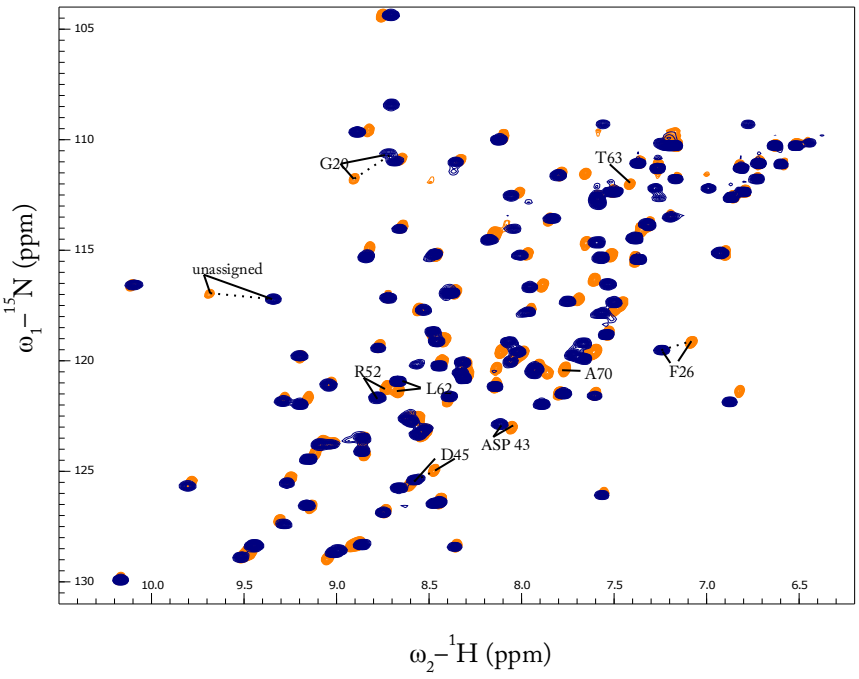


Figure 7.10 2D ^{15}N-HSQC HypF-N peak shifts with addition of phosphate

Spectrum was collected with 200 µM HypF-N at 298 K on a 600 MHz spectrometer. Significant peak changes noted from the addition of 50 mM sodium phosphate are indicated

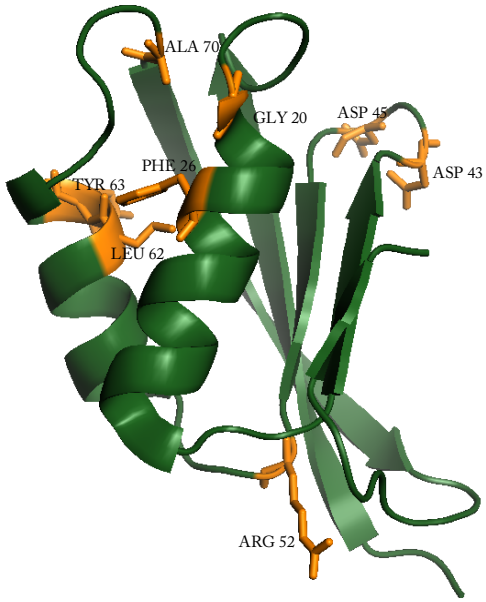


Figure 7.11 Sodium phosphate has an effect at the anion pocket and loop regions of HypF-N

It is clear that most loop regions are affected by the addition of sodium phosphate with residues Gly20, Asp43, Asp45, Arg52 and Ala70 all found within loop regions connecting the α-helices and β-sheets. Figure created using PyMol. PDB code: 1GXT (Rosano et al., 2002)

Initial attempts under the native condition yielded assignment of ~70% of peaks, this was increased to 94% after the addition of phosphate to the sample. Similarly to human muscle acylphosphatase (Fusco *et al*., 2012), HypF-N can bind to free phosphate ions at its catalytic domain that reduces the flexibility and mobility of the protein, particularly the loop regions. HypF is involved in the dephosphorylation of carbamoyl phosphates (Paschos *et al*., 2002), phosphate and sulphate ion bonding interactions are observed in the N terminal domain at residues Gly19 – Arg23 (Rosano *et al*., 2002). We do observe chemical shift perturbations for residues Gly20 and Phe26 in proximity to this binding site. Addition of 50 mM sodium phosphate led to new peaks Ala70 and Tyr63 being assigned which are in proximal to the region Gly19 – Arg23. It is likely the stabilisation at region Gly19 – Arg23 upon addition of phosphate is propagated to residues Ala70 and Tyr63 through molecular interactions and bonding.

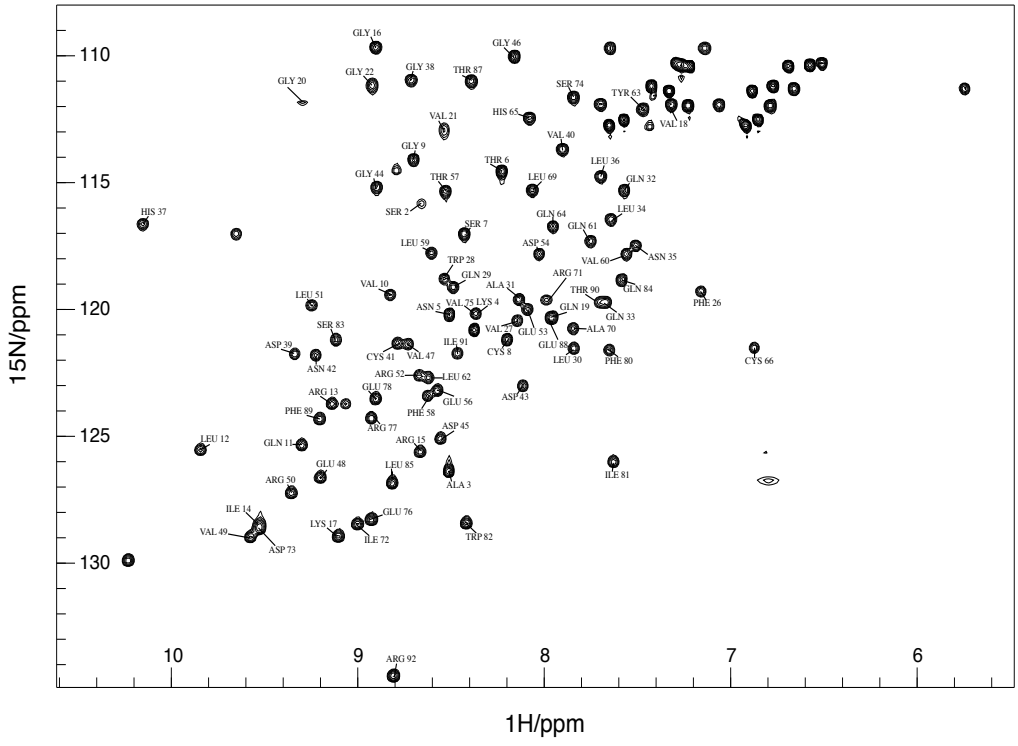


Figure 7.12 ^{15}N-HSQC of assigned backbone of HypF-N under native condition

HypFN assigned at 250 µM, under 5 mM NaAc, 2 mM DTT, pH 5.5. Peaks are labeled with amino acid residue and number in the sequence.

The peak at ^{1}H chemical shift of 10.23 ppm is the indol group of a tryptophan residue within the HypF-N sequence. 95% of all non-proline backbone peaks are present in

the ^{15}N-HSQC (Figure 7.12); but some unassigned peaks did not contain any ^{13}C nuclei resonances in the 3D data sets. The dynamic nature or conformational exchange of these residues may have had an impact at the 3D level of NMR, where given the multiple magnetization transfer steps signal in fast exchange between two states is lost.

7.2.4 Toxic

The ^{15}N-HSQC of HypF-N in the native and in the toxic condition overlays well. Most peaks were easily transferable given the slight changes in chemical shifts, and confirmed with data collected from HNCACB and HN(CO)CACB spectra collected with HypF-N under the toxic condition. Due to absence of the 50 mM in the sample residues close to the catalytic domain and neighbouring residues (19-25), and Tyr63 in the α2 helix were unassigned. Overall 90% of non-proline peaks were assigned. Given the rapid oligomerization rates of HypF-N under the toxic condition 3D experiments were collected in < 18 hours, this was not problematic since the data was used as validation and cross referenced against the native conformation. The assignment for HypF-N under the toxic condition is shown in Figure 7.13, and the full chemical shift assignment is shown in Supplementary Material in Table S2.

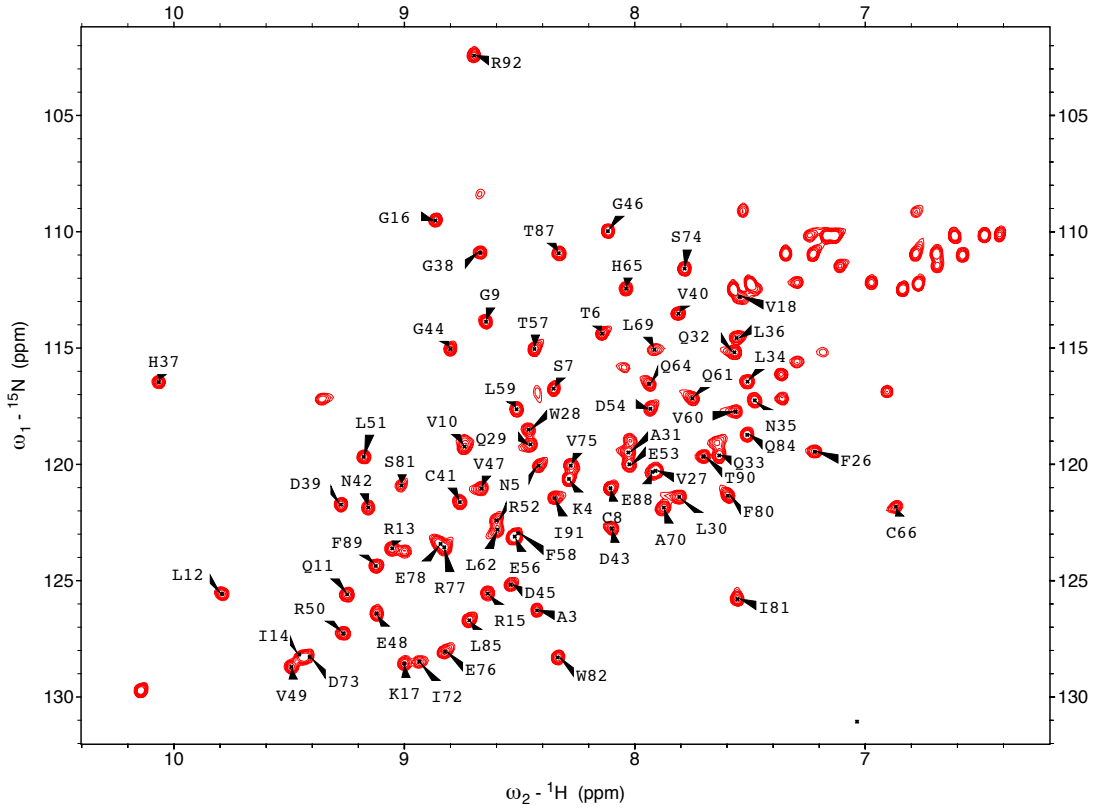


Figure 7.13 ^{15}N-HSQC of assigned backbone of HypF-N under toxic condition

HypFN assigned at 50 µM, under 12% TFE (v/v), 50 mM NaAc, 2 mM DTT, pH 5.5. Peaks are labeled with amino acid residue and number in the sequence. Figure created using Sparky

7.2.5 Non-toxic

The non-toxic condition posed a greater problem for obtaining a complete backbone assignment. The harsh acidic chemical environment leads to the collapse of dispersed peaks, resulting in a spectrum with high resonance overlap and poor resolution, which is propagated to the 3D experiments.

An alternative method to assign the protein under the non-toxic oligomerizing condition was to unfold the protein by adjusting the pH to 1.7, but with addition of 8 M urea (Figure 7.14).

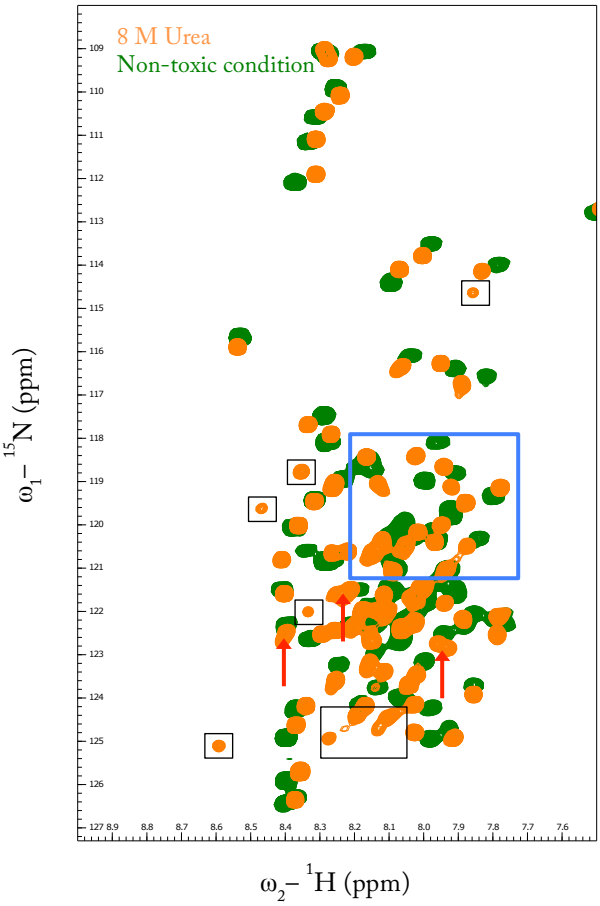


Figure 7.14 ^{15}N-HSQC spectra of the unfolded conformation of HypF-N is maintained with addition of 8M urea at pH 1.7

HypF-N at 200 µM HypF-N in 5 mM NaAc 2 mM DTT, 8 M urea, pH 1.7 (orange). HypF-N in non-toxic condition at 48 µM, 330 mM NaCl, 20 mM TFA, pH 1.7 (green). Both spectra were collected at 298 K on a 600 MHz spectrometer

Peaks boxed in black are newly observed in 8 M urea, and not under the non-toxic condition.

Peaks boxed in blue required ^{13}Cα and ^{13}Cβ resonances to confirm residue assignment.

Red arrows indicate peak doublets illustrating residues with a high level of conformational exchange between two states.

At pH 1.7 HypF-N in unfolded to give a ^{15}N-HSQC that overlays with that under the non-toxic condition (330 mM NaCl, 20 mM TFA, pH 1.7), the presence of urea does not alter the conformational state (Calloni *et al.*, 2008). Urea acts as a strong denaturant, when bound to the unfolded HypFN urea reduces protein-protein interaction by their strong repulsive charge preventing HypFN from oligomerizing. This decreases line broadening and signal loss as observed from the non-toxic condition. Unfolded HypF-N stabilised in the monomeric state leads to stronger signal

detection and better peak resolution within the 3D datasets, highly valuable given the overlapping resonances within the ^{15}N-HSQC (Figure 7.15). A titration of urea at 2, 4, 6 and 8 M was performed, however all peaks and sharp resonances were only observable at 8M.

The urea unfolded spectra, whilst missing 5 backbone amide peaks shows more peaks than for TFA. This condition, whilst suitable for backbone assignment cannot be used under conditions that require oligomerization, such as DEST measurements. 9 new peaks are observed in the urea-unfolded spectrum, boxed in black in Figure 7.14; these peaks are likely to be involved in conformational exchange and therefore not observable in the non-toxic condition, however many do not have 3D data sets and were unassigned. There are also a number of peak doublets (indicated by red arrows in Figure 7.14), these peaks are likely to be in exchange between two conformational states, exemplifying the high level of conformational exchange of HypF-N under these conditions. HypF-N under the non-toxic condition was assigned using a solution condition of 5 mM NaAc, 2 mM DTT, 8 M urea, pH 1.7,l where the sample is stable and does not aggregate over 48 h. The majority of the peaks between the two unfolded conditions can be tracked and assignment transferred from the urea sample to the TFA non-toxic sample, however the trajectory of some peaks shifts are difficult to follow (boxed in blue in Figure 7.14), these peaks were confirmed by running short 3D HNCACB and HN(CO)CACB experiments to confirm ^{13}Cα and ^{13}Cβ resonances on the non-toxic condition. The full chemical shift assignment is shown in Supplementary Material in Table S3.

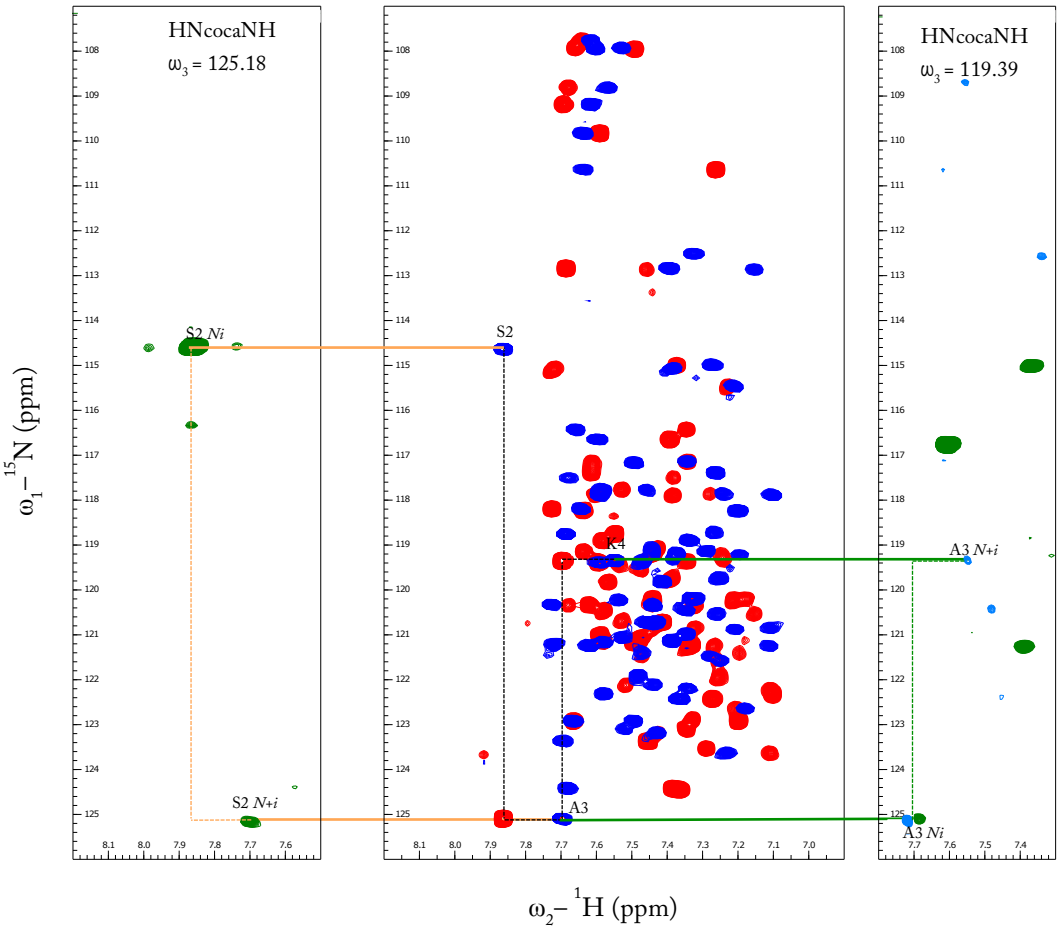


Figure 7.15 HncocaHN backbone amide 'walk' connecting i and -i backbone amide resonances

HncocaHN spectra used for assigning HypF-N in the unfolded state denatured by urea. The central plane are the HncocaNH (red) and hNcocaNH (blue) spectra. The black dashed lines correspond to the NH 'walk' connecting the residues Ser2-Ala3-Lys4 in a sequential manner. The left panel (orange lines) corresponds to the strip of the HncocaNH and resonances for Ser2, connecting the NH of Ser with residue Ser + i NH (Ala3). The right panel (green lines) corresponds to the same for Ala3 NH peak and Ala3 + i NH (Lys4).

Starting at Ser2, we find its NH amide peak in the 3D HncocaNH spectrum (left panel). The stronger intensity peak corresponds to its amide resonance, this resonance is also present in the blue hNcocaNH spectra. Within the HncocaNH spectra (red) the +i residue for this peak has the same proton chemical shift. Upon determining this peak within the HncocaNH spectra, we can then find the resonance with the same ^{15}N chemical shift to determine the +i residue. In this example it is Ala3. The process is repeated again for Ala3, finding its +i residue in the same manner with Lys4, and the sequential backbone assignment is achieved using the amide resonances without the need for ^{13}C resonance data sets.

Approximately 70% of backbone amide peaks were assigned under the urea induced unfolded state using the standard 3D HNCACB and HN(CO)CACB, HNCO and HN(CA)CO spectra. Unambiguous assignment was improved to 92% by recording an HNcocaHN experiment (Figure 7.16) that allows for a 'backbone amide HN walk', preferable for proteins with high degrees of resonance overcrowding (Sun *et al.*, 2005).

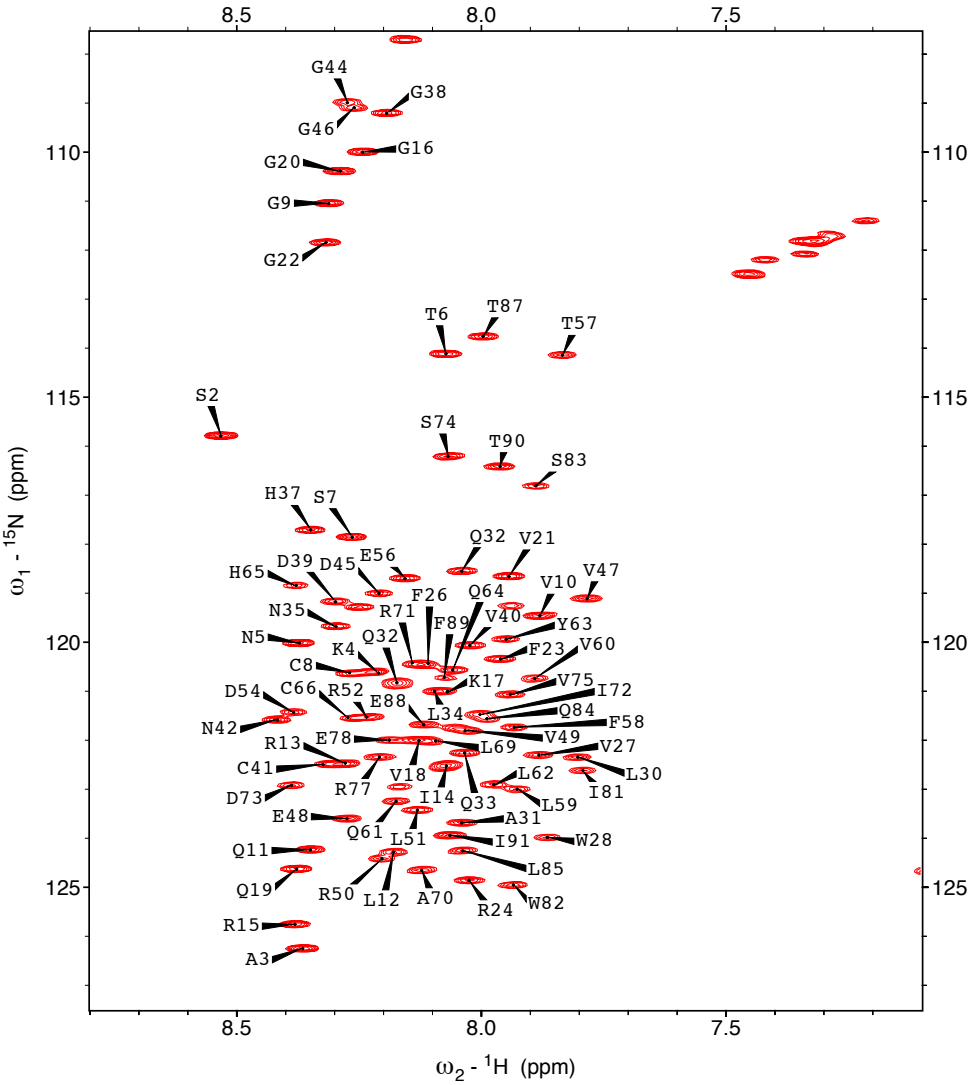


Figure 7.16 ^{15}N-HSQC of assigned backbone of HypF-N unfolded by 8 M urea

HypFN assigned at 250 µM, under 8M urea, 5 mM NaAc, 2 mM DTT, pH 1.7. Peaks are labeled with amino acid residue and number in the sequence. Backbone assignment was made with HNCACB, HN(CO)CACB, HNCO and HN(CA)CO spectra. Figure created using Sparky

7.3 Structural predictions of HypF-N under oligomerizing conditions

Structural determination of HypF-N under each condition was not performed. The rapid oligomerization rates of HypF-N under each condition makes it unfeasible to collect NMR NOESY and side chain experiments given their requirements for long experimental times. To circumvent this problem, secondary structure predictions were made based on experimentally derived measurements obtained from the chemical shifts of HypF-N under the toxic and non-toxic conditions.

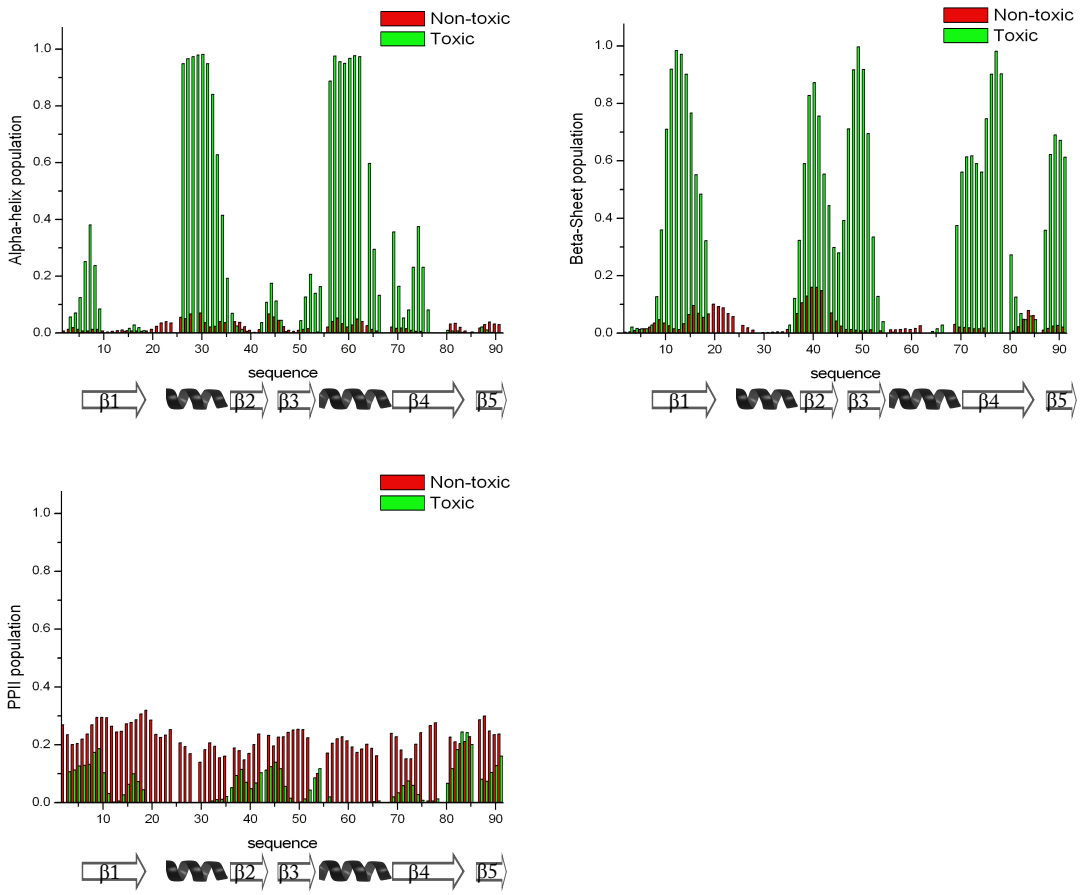


Figure 7.17 Secondary structure prediction for HypF-N under toxic and non-toxic conditions calculated by δ2D.

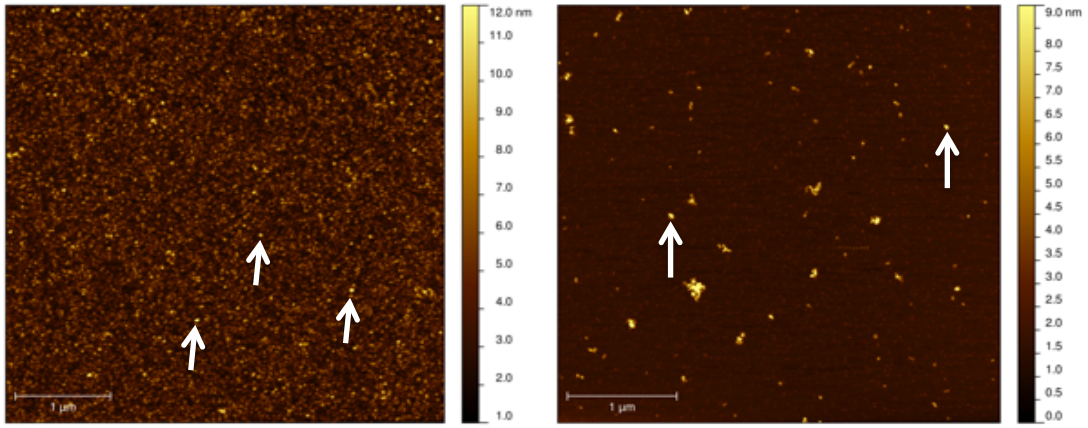
Predictions are based on chemical shift values ($^{13}C\alpha$, $^{13}C\beta$, ^{13}CO, ^{15}NH and $^{1}H\alpha$) experimentally measured from HypF-N under toxic (green) and non-toxic (red) conditions using NMR methods. The polyproline population (PPII) is shown as a negative control. δ2D prediction software developed by Camilloni et al., 2012.

Chemical shift values (Cα, Cβ, carbonyl, amide and Hα) were taken from the 3D experiments collected on HypF-N under the toxic and non-toxic conditions. The

results show that for the toxic condition, all structural motifs are maintained in a βαββαβ topology. For HypF-N under the non-toxic condition elements of secondary structure are maintained. The α-helices are lost, but some portions of the more stable β-strand structures remain, presumably due to the strong hydrogen bonding within these structural elements. The β2, and to a lesser extent the β1 strand are maintained as residual structure. HypF-N in a pre-molten globule state under the non-toxic condition was predicted to contain residual structure (Calloni *et al*., 2008), which can be concluded to consist of the β2 and β1 strands.

7.4 Oligomerization

The presence and formation of oligomers were confirmed using Atomic Force Microscopy (AFM). HypF-N was incubated in the respective oligomerizing buffer conditions and oligomer formation was examined after 4 hours.



(A) *(B)*

Figure 7.18 TM-AFM images collected on HypF-N toxic (A) and non-toxic (B) oligomers

Oligomers incubated for 4 hours at 25 °C and were diluted × 500 in H_2O for toxic and × 100 for non-toxic samples. The level of noise in the toxic condition is an artefact of the dilution factor.

The toxic oligomers are ~7 – 9 nm and the non-toxic are larger at ~9 – 11 nm. These results correspond to reported data (Chiti *et al*., 2001, Campioni *et al*., 2010), based on the spherical morphology and size determined by AFM we can be confident HypF-N is forming the correct oligomers under each respective condition.

At the monomeric level, HypF-N exhibits significantly different structural conformations as determined by NMR under the different solution conditions for toxic and non-toxic oligomer formation. Within the amyloidogensis pathway outlined in section 5.5, the nucleation of monomeric subunits (in their aggregative conformation) to form the oligomer species is the first step of the pathway. Nucleation is the addition of monomeric subunits to the oligomer composed of *n* subunits. It is the process by which the oligomer is formed, and defines its core structure. To study the nucleation process we employ an advanced solution NMR technique called DEST to directly determine the residues of the monomeric protein responsible for nucleation. (Fawzi *et al*., 2011). The assignment of the ^{15}N-HSQC spectra and conformational determination of HypF-N under each condition is a prerequisite for analysis of DEST experiments. The next section details the nucleation process of HypF-N under the toxic and non-toxic conditions respectively as measured by NMR.

7.5 Dark state exchange saturation transfer (DEST)

DEST has previously identified the residues responsible for nucleation in $A\beta_{1\text{-}40}$ and $A\beta_{1\text{-}42}$ mutant, (Fawzi *et al*., 2011), as well defining regions of α-synculein which bind to membrane surfaces (Fusco *et al*., 2014). DEST provides a non-invasive method for studying oligomerization at a residual level, on a protein that does not require potentially obstructive chemical modifications or changes to the amino acid sequence. Performing DEST does have optimal sample requirements, the transfer of magnetisation from monomer to protofibril and back during chemical exchange is very inefficient given the size of the protofibril; therefore spectra collected often require a longer amount of time to achieve the intensities and resolution of standard ^{15}N-HSQC experiments. For a protein undergoing oligomerization a condition is required to reduce oligomerization and the lag phase to a rate where DEST spectra can be collected within a reasonable time frame. Secondly, there needs to be both association and disassociation constants of the monomer to the protofibril for the signal to be detected, if the nucleation is not transient the signal cannot be detected. Much time was spent on optimising the conditions for HypF-N oligomerization,

resulting in a sample that formed oligomers within a reasonable time frame, gave good attenuation of DEST signal, and gave a good NMR signal.

7.5.1 Sample optimisation

7.5.1.1 Toxic

HypF-N oligomerizes into toxic oligomers much more rapidly than non-toxic. At 12% TFE (v/v), 50 mM NaAc, pH 5.5 the rate of oligomerization occurs instantaneously and the sample shows signs of visible precipitation. Within 8 h of this condition the NMR ^{1}H 1D spectrum is entirely diminished due to the monomeric protein being depleted to form oligomers. Obtaining data from 3D or DEST spectra, both requiring < 8 h acquisition times under this condition is very challenging.

Since the detected FID signal of DEST is reliant on the presence of monomeric HypF-N and the detection of NMR signal from this species different solution conditions varying the amount of TFE and temperature were set up in parallel and the 1D ^{1}H NMR spectrum was collected. These two variables could be manipulated without adversely affecting the aggregation process. The percentage of TFE present has an effect on the aggregation kinetics of AcP, increasing the percentage from 5 to 25% reduces the rate of aggregation (Calamai *et al*., 2005). Increasing TFE in the solution condition increases the α-helical content of AcP and provides an environment that favours strong intermolecular interactions, possibly promoting the rate of aggregation. The rate of oligomerization was measured by recording the loss of signal from 1D ^{1}H NMR, providing insights into maximum experimental acquisition times for DEST could potentially be set. Reducing the TFE to 10% and the temperature to 283 K reduced the kinetics of aggregation. The effects of changing the ionic strength for the toxic condition have not been investigated, since we do not want to alter any factors, which could potentially affect the mechanism of aggregation the percentage of TFE and temperature were assessed only. Similarly, pH could not be altered and the presence of DTT could not be removed since these components prevent structural collapse and dimerization respectively. The protein concentration was increased from 48 μM to 100 μM increasing the signal intensity without altering the aggregation.

This reduces the experiment time needed for each offset, allowing multiple offsets to be recorded off a single sample; therefore reducing errors and inconsistencies.

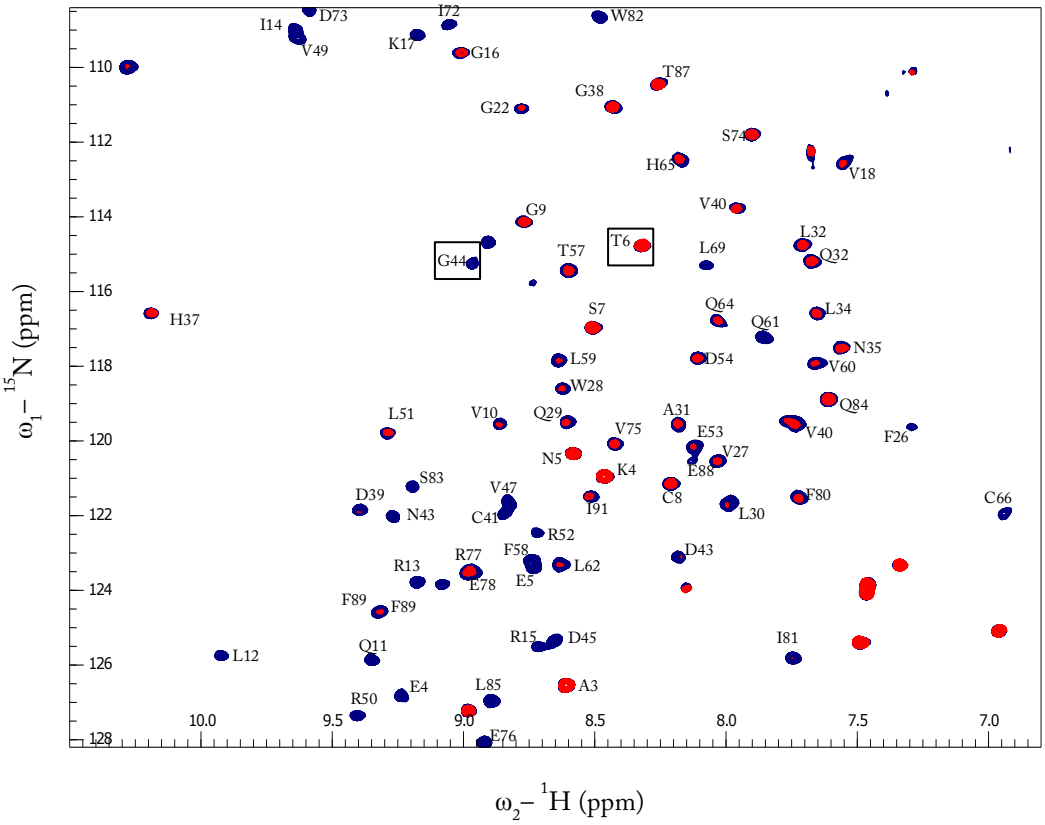


Figure 7.19 Assigned spectrum of toxic DEST

The reference at +100 kHz is shown in blue, and offset +0.5 kHz is shown in red. Residues experiencing signal attenuation (such as G44) and no signal attenuation (such as T6) are boxed. The spectral widths (^{1}H 16.02 ppm and ^{15}N 20.00 ppm) have been altered from previous ^{15}N-HSQC spectra collected for HypF-N (^{1}H 14.03 ppm and ^{15}N 30.00 ppm) to reduce the experimental times.

Analysis of DEST is based on the attenuation of signal observed between the partially saturated spectra at a specified RF offset from the ^{15}N carrier frequency set at 118 ppm (0.5 kHz) and the reference spectrum (10.0 kHz). Figure 7.19 illustrates differences in signal strength and residue specific attenuation. The reduction of signal is due to the faster T_2 relaxation rates of residues interacting at the surface of the protofibril, the greater transfer of magnetization through chemical exchange between the monomer and protofibril leads to greater attenuation of signal. The signal is detected after the monomer has bound and dissociated from the protofibril, residues

associated with the interaction will have a greater loss of attenuation relative to the reference spectrum.

The ^{15}N-HSQC of HypF-N under the toxic DEST condition (10% TFE (v/v), 50 mM NaAc, 2 mM DTT, pH 5.5, 100 μM HypF-N) loses the signal detection of 12 backbone residues compared to the native conformation. These residues include Val21 and Gly22 near to the phosphate binding domain and Tyr63 and Ala70, due to the absence of phosphate in the sample. Under the toxic condition 87% of the sequence is covered and can be analysed from the ^{15}N-HSQC spectra collected for DEST measurements.

7.5.1.2 Non-toxic

Unfolded states are often more challenging to study by NMR given the heterogeneity and ensemble of interconverting states and overlapping resonances (Baum *et al*., 1989, Alexandrescu *et al*., 1994, Dyson & Wright, 2004). The non-toxic condition posed a more difficult challenge to optimise for DEST. It is a complex, multifaceted system with pH, ionic strength, temperature and protein concentration all playing factors (Calloni *et al*., 2008, Campioni *et al*., 2010, Campioni *et al*., 2012). Figure 7.20 describes the interplay between these factors.

A major problem is the protein concentration used as standard for oligomerization is 48 μM, relatively low for a technique such as NMR, where protein concentrations above 100 μM are often much preferable. For the non-toxic condition it was determined that increasing the protein concentration leads to increases in the rate of oligomerization. Increasing the concentration of HypF-N increases the NMR signal, allowing for detection of resolved peaks at a faster rate, however the rate of aggregation is also increased, cancelling any positive effects to the NMR signal from the monomeric state. At 100 μM a balance is reached between aggregation rates and NMR signal.

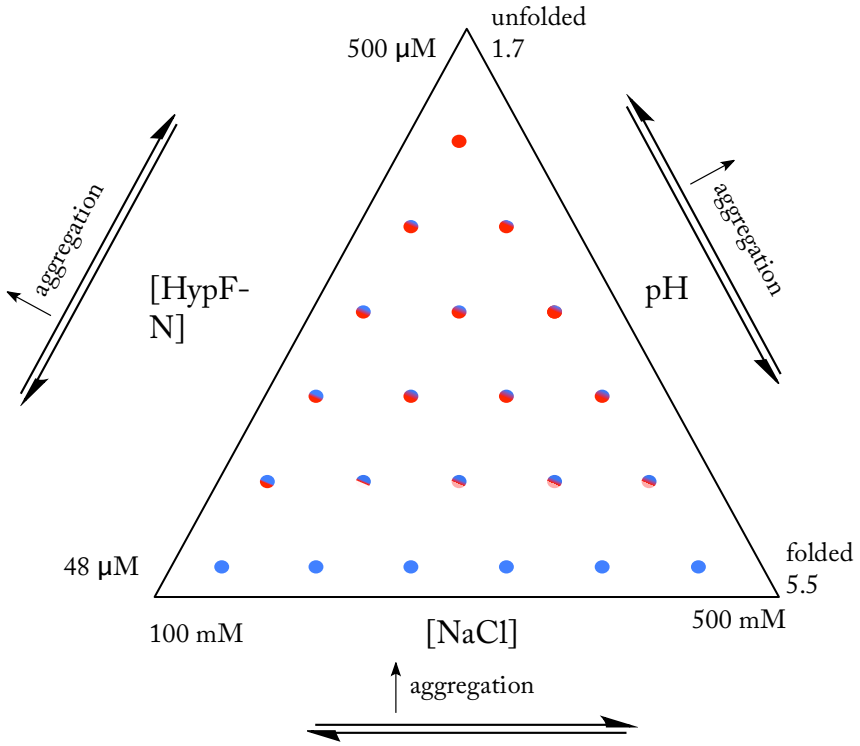


Figure 7.20 Factors affecting aggregation propensity of non-toxic HypF-N

The degree of red colour of the circles indicates a higher propensity for aggregation, as concentration of HypF-N and NaCl increases, and pH decreases the rate of aggregation increases.

The concentration of salt and the salt ion can influence the non-toxic aggregation kinetics of HypF-N. The rate of HypF-N aggregation in NaCl is lower than for alternative ions such as sodium iodide and sodium sulphate, additionally reducing the ionic strength from 330 mM NaCl to 100 mM NaCl reduces the rate of aggregation to almost × 22 slower, without changing the pre-molten globule state of HypF-N in the solution condition (Campioni *et al*., 2012). Decreasing the ionic strength from 330 mM NaCl to 100 mM reduces aggregation propensity and improves NMR signal. Increasing salt concentrations, to at least 100 mM leads to increases in the residual structure of the unfolded conformation of HypF-N (Campioni *et al*., 2008). This gain of structure may be accompanied by enhancing the hydrophobic packing. To maintain the initial residual structure, which is significant to oligomer nucleation, DEST was performed at 100 mM NaCl.

Acidic pH is not ideal for NMR samples, however this factor cannot be altered since the unfolded pre-molten globule state is dependent on the acidic denaturation of HypF-N. 20 mM TFA is used for forming the non-toxic oligomer (Campioni *et al*., 2010), this reduces the pH of the solution to 1.7. Experiments determined that reducing the concentration of TFA to 2 mM significantly reduces the rate of oligomerization, but maintaining the pH at 1.7 using HCl maintains the conformation of the unfolded protein as determined by ^{15}N-HSQC.

Whilst there are many factors contributing to defining the solution condition of non-toxic oligomerization, many of the factors ([HypF-N] and pH) are dictated and cannot be altered; reducing temperature, [NaCl] and [TFA] reduces the kinetics of aggregation without altering the conformational state of the protein as researched previously (Calloni *et al*., 2008, Campioni *et al*., 2012).

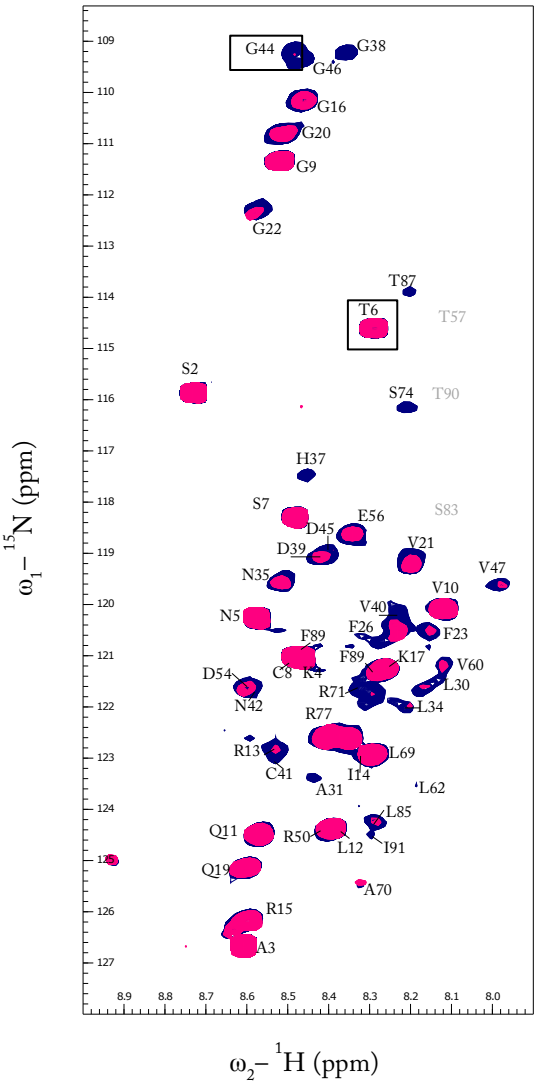


Figure 7.21 Assigned spectrum of non-toxic DEST

The reference at +100 kHz is shown in blue, and the offset +0.5 kHz is shown in pink. Examples of residues experiencing signal attenuation (G44) and no signal attenuation (T6) are boxed. Residues T57, T90 and S83 are not present at the depicted contour levels which shows good signal to noise ratio

The optimised non-toxic sample for DEST does not exhibit all backbone peaks detected under the urea induced unfolded conformation studied by ^{15}N-HSQC. The structural fluctuations and conformational exchange between the unfolded states on a microsecond to millisecond timescale is likely to lead to the absence of certain peaks. Additionally, the spectrum still retains a strong degree of resonance overlap dictated by the unfolded conformation HypF-N.

In the acid induced unfolded state we observe ~30% peak loss compared to HypF-N unfolded in urea, potentially from chemical exchange occurring between the monomer and protofibrils. It is hypothesised that the cross peaks absent in the non-toxic condition ^{15}N-HSQC are from regions of HypF-N that retain some degree of structure, leading to the line broadening effects (Calloni *et al*., 2008). Secondary structure predictions based on chemical shifts support this, since many residues that form the β4 strand (residues 72 – 81) are not observable in the ^{15}N-HSQC and shown to still maintain a residual β-strand structure. This is detrimental to analysing DEST results, given that more confident conclusions can be made with complete sequence coverage. Under the non-toxic condition for collecting DEST 72% of the sequence is observable, the loss of peaks and sequence coverage does have a slight impact on analysing data. The residues with observable DEST data are shown in Figure 7.22. Strong peak intensities are noted for the first 7 residues in the N terminal domain; based on δ2d structural predictions they are predicted to lack a defined structure and are therefore flexible.

7.5.2 Nucleation

2D ^{15}N-^{1}H DEST HSQC spectra are collected at increasing RF offsets; magnetisation of signal is initiated on monomeric protein molecules which then bind to the protofibril, chemical exchange of the magnetisation occurs and it is detected from monomer species that have dissociated from the protofibril. The difference in signal intensities observed is indicative of greater chemical exchange and identify residues of the monomer species interacting with the protofibril. These residues and regions are the sites of monomer binding and nucleation of the protofibril. Attenuation of the signal identifies residues responsible for nucleation.

DEST experiments were performed at selected offsets and processed using nmrPIPE (Delaglio *et al*., 1995). Intensity values of each cross peak were obtained using CCPNAnalysis (Vranken *et al*., 2005). Each assigned peak in each offset was compared to its reference spectrum at +100 kHz to give I/I_0 values and plotted across the sequence. Residues exhibiting attenuation of signal resultantly have a low I/I_0 value, and are sites of the protein that in its monomeric form bind to the oligomer.

These residues are therefore involved in the nucleation process of the oligomeric species, and compose the core of the oligomer structure.

DEST results are analysed as a trend across the HypF-N sequence. The average signal for both conditions at ±1.5 kHz is shown. This offset shows good DEST signal for HypF-N in both conditions, above this offset the attenuation of signal begins to be lost across the sequence. Conclusions between the toxic and non-toxic conditions are drawn and comparisons to previous research on HypF-N and homologous proteins using alternative techniques are made (Campioni *et al.*, 2010, De Simone *et al.*, 2011).

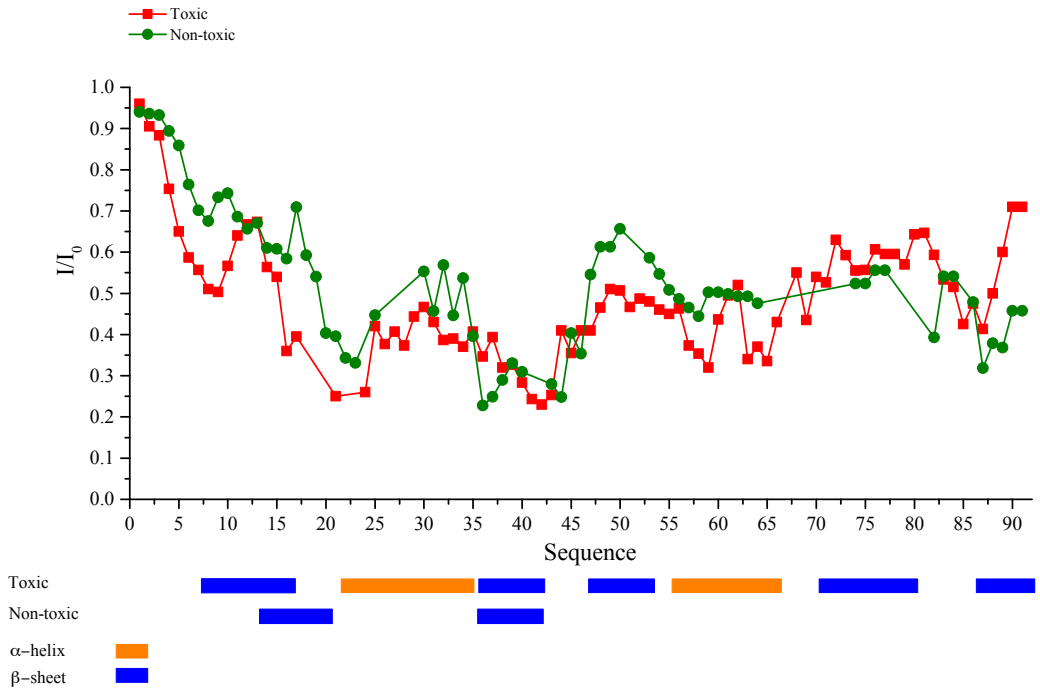


Figure 7.22 DEST profile of HypF-N in toxic and non-toxic conditions

Overlay of DEST of toxic (red) and non-toxic (green) oligomers at 1.5 kHz offset

Data derived from DEST analysis are plotted as dots (non-toxic) and squares (toxic); backbone residues missing from the assignment and spectra including proline residues are not marked.

Secondary structure predictions based on NMR derived chemical shifts are indicated for HypF-N under each condition.

7.5.3 Aggregation occurs via the β2 strand

The profiles for HypF-N under both conditions overlay very well. It appears that residues 15 – 24, 36 – 44 and 85 – 87 are consistent between both conditions, and that both forms of oligomers form via a similar mechanism, despite different monomeric conformation and level of folded structure.

In both cases we observe attenuation of signal for regions 15 – 24, which form the preceding loop region and the start of the α1 helix. It is unlikely that the α1 helix plays are role since it has been shown that residual α-helical content does not influence the aggregation of the non-toxic oligomer (Ahmed *et al*., 2011). Both the toxic and non-toxic conditions do not involve the β1 strand; whilst there is a shift in the β1 strand position in the non-toxic condition the prediction score is weak, and attenuation of signal falls at the end of the secondary structure element. It is most likely that the effects of the β2 strand are propagated across the intermolecular bonds to the α1 helix in the toxic condition. For the non-toxic condition a similar mechanism is likely to occur, the signal is still propagated across proximal residues despite the loss of the α-helical structure.

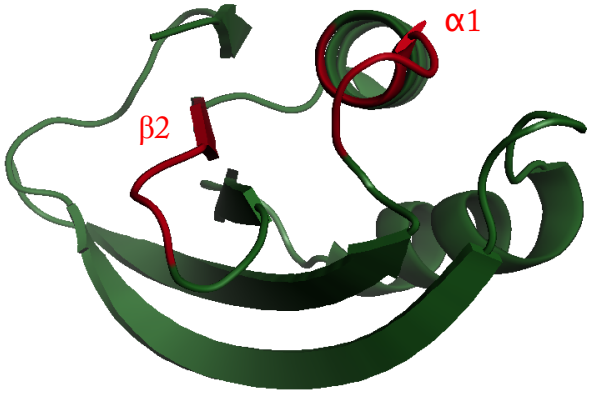


Figure 7.23 β2 strand is in close proximity to the α1 helix

Regions coloured red represent regions involved in the nucleation of toxic and non-toxic oligomers as derived from DEST measurements.

Hydrogen bonded salt bridges are formed between residues of the β1 and β2 strands (Arg12-Asp42, Arg14-Asp44) with hydrogen bonding occurring between Gln18-

Asn41/Gly43 (Rosano *et al*., 2003). This provides evidence for this in the toxic case where the intermolecular contacts are maintained. The attenuation of signal for the residues 15 – 23 in both conditions is likely to be indirectly caused from the β2 strand on the α1 helix and the loop region following the helix. The proximity of both structural elements are shown in Figure 7.23.

Residues 36 – 44 form the β2 strand and part of the loop region that connects the β2 and β3 strand. This structure is maintained in both conditions, with residual β2 strand structure present in the premolten globule conformation under the non-toxic condition. DEST results depict that aggregation is propagated across this structure in both cases. The attenuation of signal extends to the loop region connecting the β2 and β3 strands; this is likely to be due to the effect of the β2, the loop is mobile, consisting of small flexible residues such as Gly44 and Gly46. These residues have not been previously identified as regions of unfolded high aggregative propensity (Calloni *et al*., 2008), or involved in the core as determined by fluorescence measurement (Campioni *et al*., 2010).

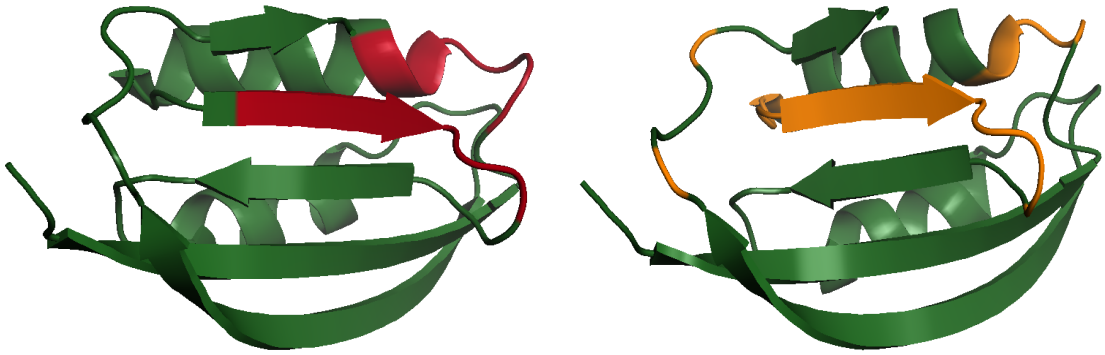
In both conditions we observe another reduction in signal at residue 85 – 87, this is likely to be due to the changes to the β5 strand in during aggregation. The displacement of this strand, similar to the mechanism of AcPDro2 (De Simone *et al*., 2011), may cause attenuation of signal in the loop region and β5 strand. The increased flexibility or mobility due to the displacement of the strand in the toxic condition, and loss of structure in the non-toxic condition could implicate its involvement in nucleation. This effect is not likely to drive or initiate the aggregation, rather an effect from the displacement. Additionally this region is in proximity to the β2 strand, and it could be that the effects observed are a consequence of the effects from HypF-N aggregation across the β2 strand.

A difference between the toxic and non-toxic results is that residues 63 - 70 appear to be involved in the toxic nucleation. This does not correspond to a hydrophobic region, or a high aggregation propensity of the residue. These residues are located in the loop region at the end of the α2 helix, connecting the α2 and β4 strands. This region does exhibit strong hydrophobicity; these observations may be from an interaction occurring across the hydrophobic residues during aggregation. It is not observed in

the non-toxic condition, potentially because the hydrophobic regions are collapsed in the unfolded state, it has been shown previously that hydrophobic regions are more accessible in the toxic case (Campioni *et al*., 2010).

7.5.4 β2 strand and the β-loop are present in the core of both oligomers

Despite very different initial monomeric states, there is a consistent trend in the DEST profile for both toxic and non-toxic condition. Results indicate that the β2 strand is involved in the nucleation of both toxic and non-toxic oligomers. This suggests that the β2 strand is a key structural motif in nucleation of HypF-N under both toxic and non-toxic oligomers.



(A) *(B)*

Figure 7.24 Residues that form the core of toxic (A) and non-toxic (B) HypF-N oligomers

Residues identified from I/I_0 values taken from 1 kHz offset of both toxic and non-toxic condition

From secondary structure predictions using experimentally obtained NMR chemical shift values, the β2 and, to lesser degree the β1 strands are maintained in the non-toxic condition. The residual structure of the β2 strand is functionally relevant in the oligomerization process of the non-toxic oligomer. Although the residual structure is likely to present in a low population, aggregation can occur across conformations that represent 1% of an ensemble of the protein structure (Marcon *et al*., 2005). This highlights the significance of these structural elements and the steps evolution has taken to reduce aggregation prone conformations in polypeptide structures.

The β2 strand is composed of an amino acid stretch Gly38-Asp39-Val40-Cys41-Asn42; the region does not show high levels of hydrophobicity, charge or aggregation propensity. However an increase in solvent accessibility could increase the propensity for aggregation across hydrophobic residues such as Gly38 and Val40 and potentially stabilised by Cys41 side chain. The changes in initial conformations induced by the presence of TFE or by reduction in pH, whilst different, both conformational states of the protein could have increased solvent accessibility of the β2 strand, which appears to be responsible for aggregation.

7.5.5 Comparison to previous data

Fluorescence measurements have shown that the hydrophobicity of the sequence may play a role in the aggregation pathway with sequences 22 – 34, 55 – 59 and 75 – 87 being shown to form the core of the oligomer of the non-toxic condition, and to a lesser extent under the toxic condition (Campioni *et al*., 2010). These regions of the sequence correspond to the most hydrophobic regions of the sequence. The results derived from DEST experiments indicate that the hydrophobicity does not drive the aggregation of HypF-N. DEST is performed on an unaltered protein structure, without any invasive probes such as *N*-(1-pyrene)maleimide (PM) used in fluorescence measurements. These probes require changes to the amino acid sequence, and the large pyrene chromophore may perturb the enzyme structure and cause slight conformational changes that may have adverse effects on the oligomerization rates or aggregation mechanism of HypFN. This may explain the differences between results observed here with those of published work (Campioni *et al*., 2010).

Results attained using NMR methods correspond more the homologous AcPDro2 acylphosphotase enzyme. It was determined that in for AcPDro2 in 5% TFE (v/v) the β5 strand was displaced, offering solvent access between the β2 and β1 strands. This was modelled and supported by simulation studies (De Simone *et al*., 2011). It is known that TFE can affect the tertiary state of the protein, and this may be the effect observed for both AcPDro2 and HypF-N in its toxic condition; suggesting a similar mechanism for the two structurally homologous proteins. For the non-toxic case, the residual presence of the β2 strand is required for aggregation to occur, despite the loss of structure in other regions of the protein.

7.5.6 Concluding remarks

These results provide new insights into HypF-N aggregation, and offers an alternative view contrary to previous data which supported a theory that the residual structure did not play a significant determinant in HypF-N aggregation, particularly for the non-toxic condition (Ahmed *et al*., 2011). Our results differ to findings based on fluorescence data measurements (Campioni *et al*., 2010), and are based on aggregation occurring across the β2 strand, similar to AcPDro2 (De Simone *et al*., 2011).

Factors such as hydrophobicity, net charge and β-sheet propensity are considered to play a greater role in aggregation than the nature of the initial conformation of proteins (Calamai *et al*., 2005, Chiti *et al*., 2003). Data collected from DEST supports this hypothesis. HypF-N under both toxic and non-toxic conditions exhibits different initial conformations, however the β-sheet structures that are maintained, and solvent accessible initiate the aggregation in both conditions.

It appears that toxicity cannot be defined by the nucleation of each oligomer. Both oligomers appear to nucleate via the same mechanism. It does not appear that the increased surface exposure of hydrophobic regions of the toxic condition relates to its toxicity, as previously hypothesised (Campioni *et al*., 2010). Results presented suggest that an alternative difference between the oligomers defines the differences in their respective cytotoxicity. These findings are supported in the next section, which details solid-state NMR studies on the oligomeric state of the protein.

7.6 Solid-state

HypF-N oligomers cannot be studied by solution state NMR, the slow molecular tumbling of such a large species broadens the resonances to a level effectively invisible to solution NMR. An alternative technique to study such a sample is ssNMR. Advantageous to crystallography, ssNMR can provide information on structure, dynamics, conformational states and interactions of biomolecular macromolecular complexes. Additionally ssNMR can be used to study oligomers in a lipid membrane

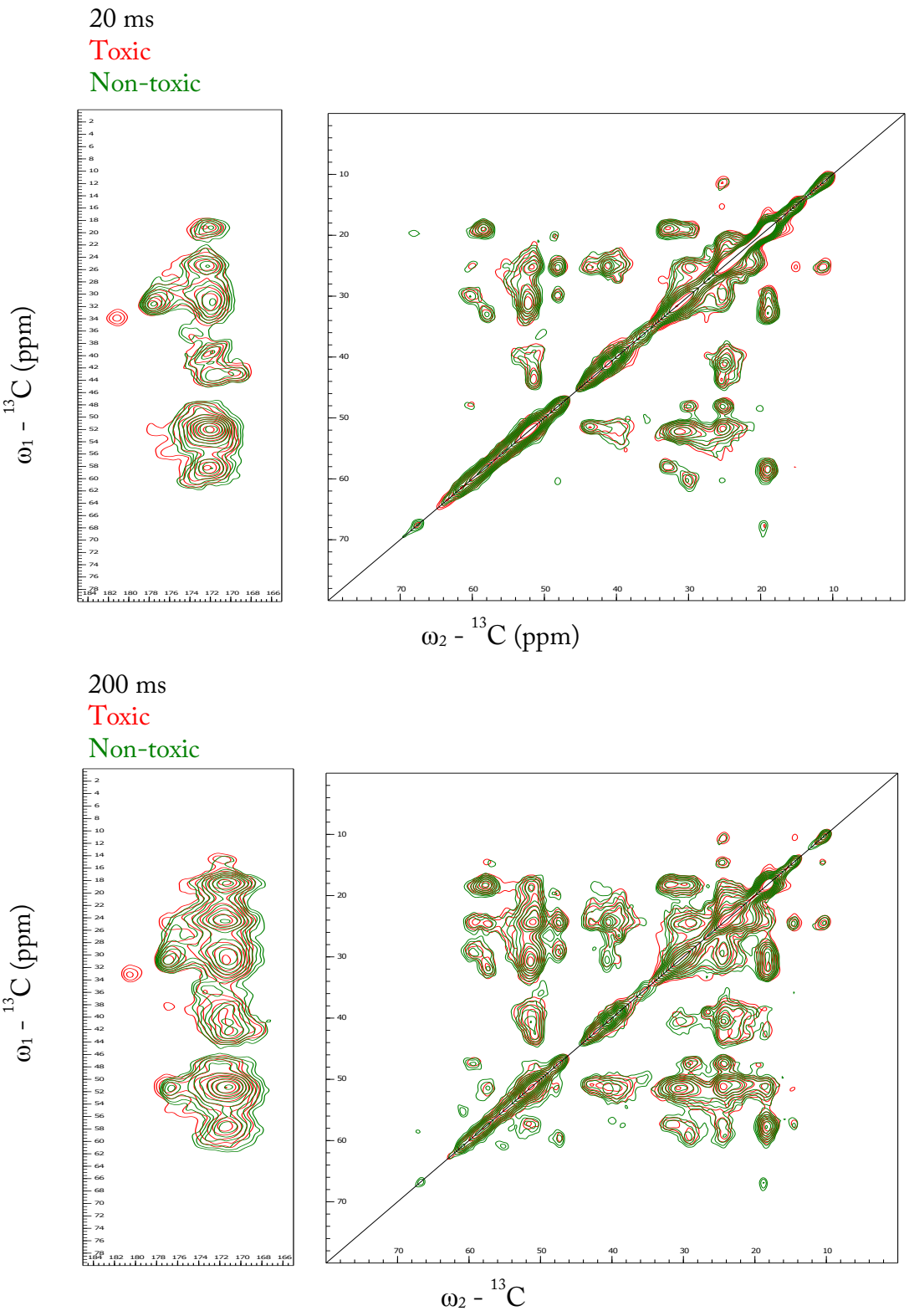
environment, where the physiological effects can be characterised in detail. It is known that the cell membrane and its physiochemical properties play a major role in the toxicity of oligomers (Cecchi *et al*., 2005, Cecchi *et al*., 2006, Evangelisti *et al*., 2012). ssNMR was performed to study the oligomer conformation and its interactions with model cell membranes, varying crucial components of the cell membrane.

HypF-N fibrils are morphologically and structurally similar with only slight variations in size (Chiti *et al*., 2001). However, the oligomeric pre-fibrillar species structure can vary, HypF-N protofibrils can differ in population and mechanical properties despite similar incubation conditions (Relini *et al*., 2010). These differences in protofibril population and polymorphisms are noted elsewhere in literature, the structural elements of Aβ oligomers that vary in β-sheet content exhibit different cytotoxic levels (Ahmed *et al*., 2010). The structural determination and characterisation of oligomers identify key regions that potentially elucidate cytotoxic origins and key interactions to cell membranes (Ahmed *et al*., 2010, Fusco *et al*., 2014, Huang *et al*., 2015). A number of these observations have been made possible through ssNMR and the rapidly occuring advancements of ssNMR techniques, magnet field strengths, probes, higher MAS frequencies, pulse sequences, smaller rotor sizes and better sample preparation. AcP has previously shown to exhibit similar aggregative states detected by transmission EM despite different ordered and disordered initial states which varied in the content of secondary structure (Calamai *et al*., 2005). To fully understand the toxicity of HypF-N, it is pivotal to understand and determine the structural and characteristics of the oligomeric state of HpyF-N in the presence of cell membrane environment.

7.6.1 Toxic and non-toxic oligomers have similar core structure

Two different experiments were performed, ^{13}C-^{13}C DARR and ^{1}H-^{13}C INEPT to ascertain information pertaining to the rigid and flexible regions of the oligomer respectively. ^{13}C-^{13}C DARR correlate both main chain and side chain ^{13}C resonances in the protein through homonuclear dipolar recoupling. The signal detected originates from the rigid moieties of the sample. DARR is often used to detect interactions, where protein regions become more rigid upon binding to surfaces (Traaseth *et al*., 2009). In combination to DARR, ^{1}H-^{13}C INEPT spectra were collected where signal

is detected from the most mobile and dynamic regions of the sample. In INEPT experiments the transfer of magnetisation is via *J*-couplings observable where the dipolar couplings are averaged by the mobility of the protein, signal detected is from regions and residues of the sample, which exhibit a high level of mobility and flexibility, not from solid core regions. This complementary system allows quick and detailed information on two key structural aspects of the oligomers. All spectra were collected at 4 °C where signal-to-noise ratios were highest, no changes in the spectra were observed at -20 °C or 25 °C.



(B)

Figure 7.25 HypF-N toxic and non-toxic oligomers have a similar core structure

^{13}C-^{13}C DARR collected for HypF-N toxic (red) and non-toxic (green) oligomers with 20 (A) and 200 ms (B) mixing times.

^{13}C-^{13}C DARR spectra of toxic and non-toxic oligomers at different contact times of 20 and 200 ms overlay very well. The short (20 ms) and long-range (200 ms) connectivities and resonances are largely consistent between both toxic and non-toxic oligomers. This indicates that the core regions of both oligomers are structurally similar. We do observe some differences; a CO peak with ~180 ppm in the ω_2 ^{13}C dimension is detected in toxic oligomer unobserved for non-toxic. These results add support to the DEST data, where both oligomers appear to nucleate via similar residues and would resultantly have similar inner core structure. There are a few pekas in the toxic spectrum unobserved for the non-toxic, at both mixing times however these differences do not suggest major, global changes in the fibril structure.

7.6.2 Differences in oligomers lie within the flexible regions

To study the flexible regions, which are expected to be present on the surface of the oligomer, and the dynamic properties of the oligomers, ^{1}H-^{13}C INEPT measurements were performed. Signal detected from the INEPT spectra is from the flexible, mobile regions of the solid sample, and may elucidate the differences in the exposed surface of the oligomer, which are able to form interactions with cellular components.

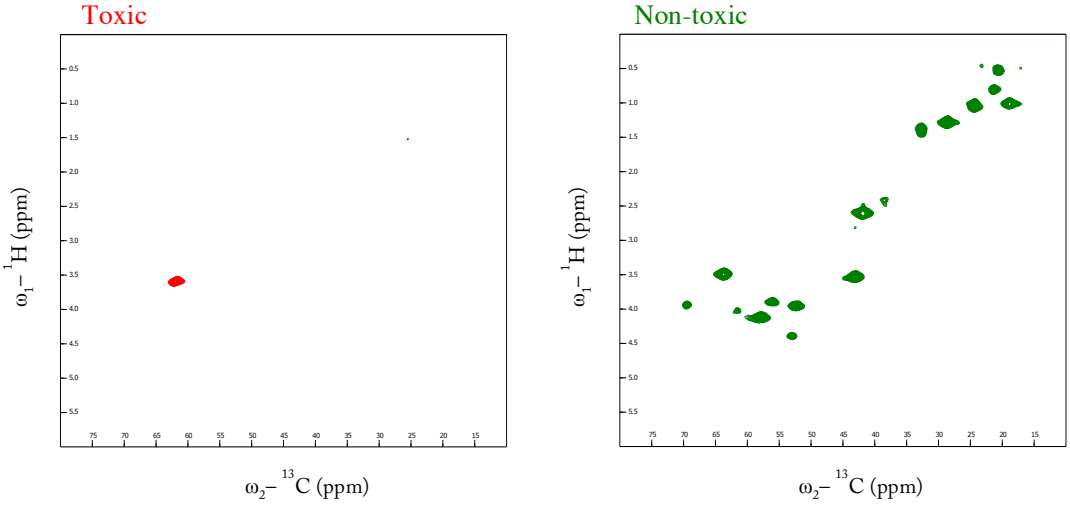


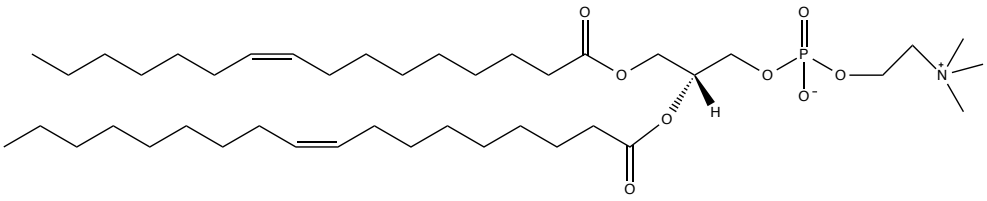
Figure 7.26 HypF-N toxic oligomers are less flexible that non-toxic
^{1}H-^{13}C INEPT collected for HypF-N toxic (red) and non-toxic (green) oligomer only samples.

Only a single peak is noted for the toxic oligomer from the ^{1}H-^{13}C INEPT spectra, which may possibly be from residual buffer (20 mM TFA, 50 mM NaAc, 2 mM DTT), however a significantly higher number of peaks (18) are observed for the non-

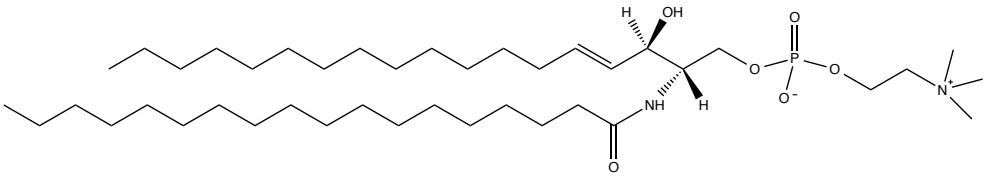
toxic (Figure 7.26). This indicates that certain residues of the non-toxic oligomer exhibit greater mobility and flexibility, compared to the toxic oligomer. These residues may be within a certain sequence or region of the polypeptide chain. Aside from the monomeric state of the protein, this is the first major difference detected between the oligomers and may be the pivotal difference in toxicity. Comparisons to solution-state ^{1}H-^{13}C HSQC NMR may provide similarities, these experiments have to be performed.

7.6.3 Vesicle interactions

The lipid bilayer composition plays a crucial role in protein interactions, factors such as phospholipid charge, acyl chain ordering, Tm, acidity and sterol inclusion can alter and modulate effects such as binding orientation, specificity and permeability of proteins and peptides (Feng-Liu *et al*., 2002). Accurate phospholipid composition for model bilayers is essential to elucidate results and assumptions that can be replicated within *in-vitro* studies. To study oligomer behavior in a suitable membrane environment we incubated the oligomers with small unilamellar vesicles (SUV) composed of the zwitterionic dioleoylphosphatidylcholine (DOPC) and sphingomyelin (SM) lipids at a ratio 2:1 as a model for cell membranes. DOPC and SM are common components of the cell membrane and are often used to mimic the general phospholipid bilayer environment. The zwitterionic DOPC is the most abundant lipid present in neural membranes (Norton *et al*., 1975), studying DOPC only reduces the complexities observed when studying mixed lipid vesicles; anionic charged headgroups such as DOPS have been shown to influence oligomer interactions (Elliott *et al*., 2005). SM has crucial roles in its interactions with sterols such as cholesterol and forms ordered domains in the lipid bilayer (Patton, 1970, Slotte & Ramstedt, 2007). Whilst the entire plethora of cellular environment cannot be entirely replicated, the ability to select each component within the SUV sample allows for focused and specifically designed experiments.



18:1 (Δ9-Cis) 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC)



Sphingomyelin (SM)

Figure 7.27 Structure of DOPC and SM lipids

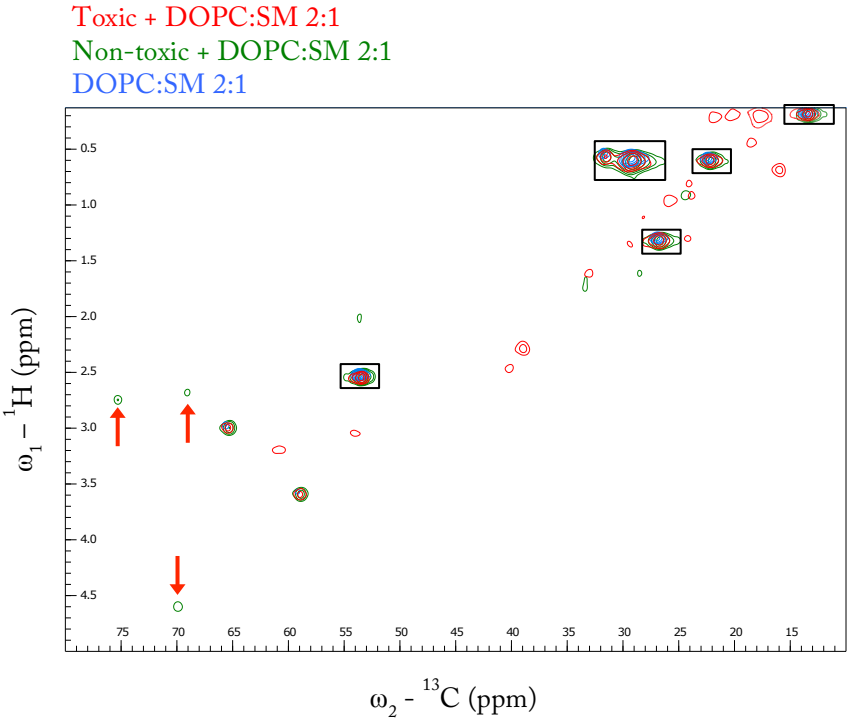


Figure 7.28 HypF-N toxic oligomers are more mobile in a membrane environment

^{1}H-^{13}C INEPT collected for HypF-N toxic (red) and non-toxic (green) oligomers in DOPC:SM (2:1) SUVs. Control peaks from DOPC:SM (light blue). Arrowed peaks are likely to be serine or threonine residues based on their ^{13}C resonances. Boxed peaks are peaks arising from the lipid component of the sample, not from protein.

The toxic oligomer gives 18 peaks when incorporated into an SUV of DOPC:SM 2:1, this is compared to 9 from the non-toxic oligomer (Figure 7.28). Since ^{1}H-^{13}C INEPT spectra detect signal from the mobile, flexible residues this may indicate that residues

within the toxic oligomers exhibit an increase in dynamical properties and flexibility when in a lipid membrane environment, or that the level of conformational exchange in reduced in the lipid membrane environment. Whilst these peaks are unassigned, some tentative assignment can be made, we observe a number of serine or threonine peaks for the non-toxic case, (indicated in Figure 7.28). These could be from the N terminus at positions Ser6 and Thr7, which is believed to be unstructured (Figure 7.17) with a high level of flexibility, given that these residues give very sharp and strong cross peaks in solution-state ^{15}N HSQC spectra of the partially unfolded monomeric HypF-N under the non-toxic condition (Figure 7.16). We detect an increase in the number of peaks in the INEPT spectrum for toxic oligomers in SUV sample of DOPC:SM 2:1. This indicates that the behavior and dynamical properties of the toxic oligomers are affected when in a lipid membrane. Previous research has indicated that a potential structural flexibility and disruptive interaction with cell membrane is noted for the toxic oligomer over the non-toxic (Campioni *et al*., 2010), this could be due to the increased mobility observed. Conversely we observe fewer peaks for the non-toxic; those detected are from the lipids in the control sample. This indicates that unlike the toxic oligomer, the residues in the non-toxic oligomer becomes less mobile, and exhibits less flexibility. It may be that the peaks lost in the non-toxic sample are due to interactions formed between the surface of the oligomer and the lipid structures. The lipid vesicle appears to affect and inverse the flexibility that each oligomer exhibits; toxic oligomers give more peaks, which can be linked to an increased dynamic or flexibility of certain regions of the oligomer. The non-toxic oligomer depicts, the opposite, reduced dynamics. The changes in the dynamical aspects of each oligomer may be a defining factor in their interactions with cell membranes, thus eliciting different cytotoxic effects. It has previously been noted that different forms of α-synculein oligomers with increased flexibility are more cytotoxic than the normal form (Cremades *et al*., 2012, Paslawski *et al*., 2014). These results imply a role of flexibility and possibly a link between mobility of the exposed hydrophobic region in forming interactions and destabilising the membrane. These results imply that a common mechanism, mediated through flexibility, may be applicable to all cytotoxic oligomers.

No significant differences are detected for the DARR spectra for oligomers in presence of lipid SUVs (Supplementary Material Figure S8, S9), indicating that core

of the oligomer in each condition are largely similar. This adds further support to the hypothesis that the surface of the oligomer is most likely the site of cellular interactions.

7.6.4 GM1 content

The SUV samples are composed of the phospholipid DOPC and SM lipid, as a model membrane. Whilst they do provide a reliable simplistic model system, they do not fully replicate the plethora and complexities of cellular membranes. Ganglioside GM1 and cholesterol have been shown to modulate toxicity of HypF-N oligomers (Evangelisti *et al*., 2012). Increasing GM1 leads to increasing the cytotoxic effects of non-toxic HypF-N oligomers, to levels comparable to the toxic condition. Decreasing the amount of cholesterol has a similar effect, increasing the cytotoxicity of the non-toxic oligomers. Conversely, for the toxic condition, decreasing GM1 and increasing cholesterol affects the cytotoxicity, with oligomers eliciting no significant cytotoxic effects. When both cholesterol and GM1 are present together, it is the presence of GM1 that has the dominant contribution to cytotoxicity (Evangelisti *et al*., 2012). Our experiments attempt to identify what effect GM1 has on the oligomer, and its interaction and behavior in the membrane. The amounts of GM1 (10%) and cholesterol (5%) were based on previous findings (Evangelisti *et al*., 2012) and through personal communications with collaborator Prof. F. Chiti.

The addition of 5% cholesterol to DOPC:SM (2:1) SUVs exhibited no changes in the DARR or INEPT spectra for either the toxic or non-toxic case (Supplementary Material Figure S10). This validates previous findings that GM1 is the dominant component to cytotoxic effects (Evangelisti *et al*., 2012). The following sections details the changes in ssNMR spectra, in the presence of 10% GM1 for both types of oligomer.

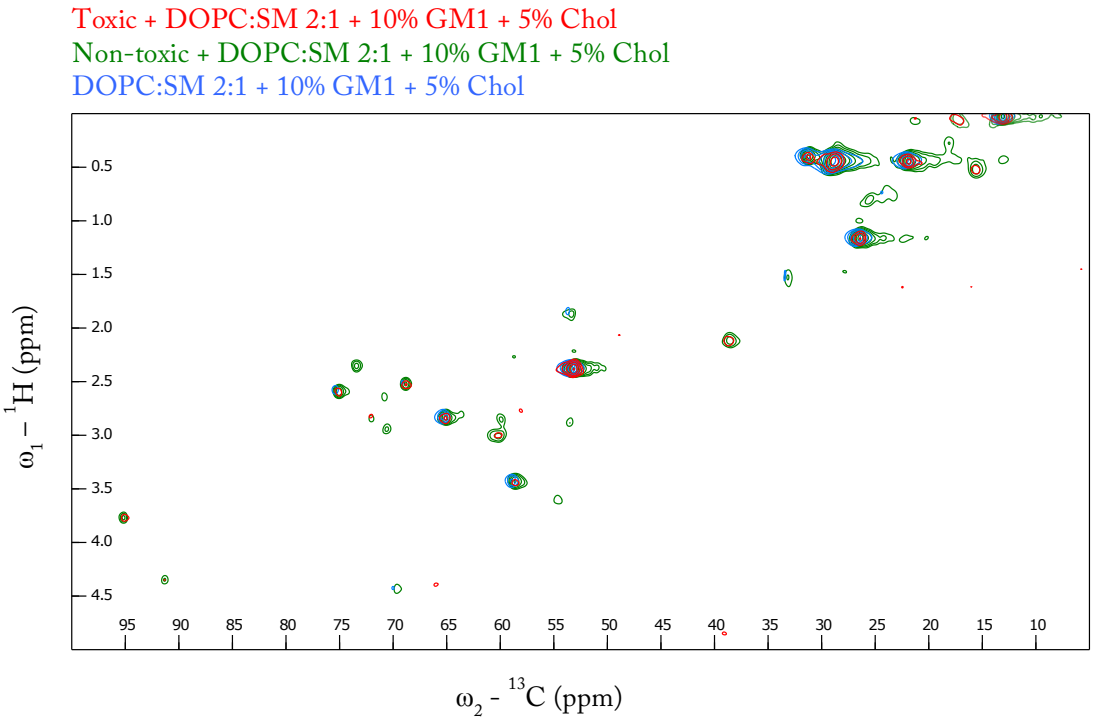


Figure 7.29 GM1 reduces the mobility of HypF-N toxic oligomers

^{1}H-^{13}C INEPT collected for HypF-N toxic (red) and non-toxic (green) oligomers in DOPC:SM (2:1) + 10% GM1 + 5% cholesterol SUVs. Control peaks from DOPC:SM + 10% GM1 + 5% cholesterol (light blue).

The addition of GM1 reduces the number of peaks for the toxic oligomer in the INEPT spectrum. This could indicate either an interaction between the toxic oligomer and GM1, or that the presence of GM1 reduces the mobility of the oligomer by altering the physiochemical environment of the lipid vesicle. The presence of GM1 has a clear effect on the toxic oligomer, reducing the number of flexible peaks detected in the INEPT spectrum from 18 to 6. For the non-toxic oligomer, new peaks are observed, which overlay for the toxic oligomer without GM1, and do not appear in the control spectra.

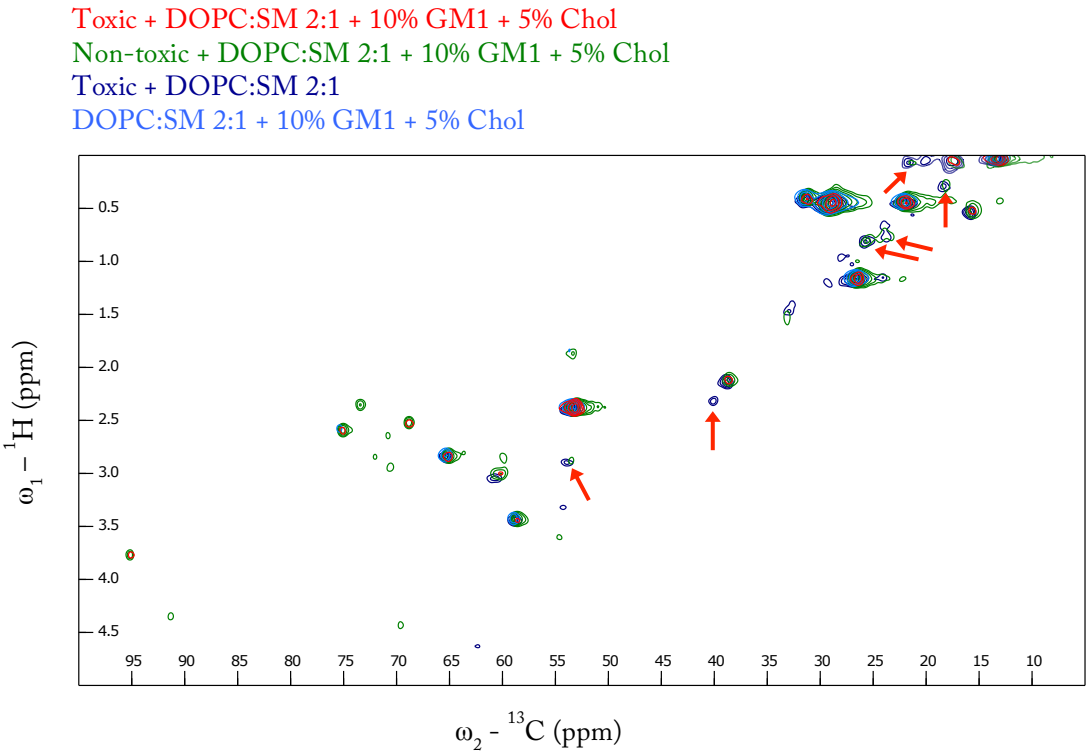


Figure 7.30 Mobility of non-toxic oligomers with GM1 is similar to that of toxic oligomers in absence of GM1

^{1}H-^{13}C INEPT collected for HypF-N toxic (red) and non-toxic (green) oligomers in DOPC:SM (2:1) + 10% GM1 + 5% cholesterol SUVs. Toxic oligomer with DOPC:SM (2:1) only (navy). Control peaks from DOPC:SM + 10% GM1 + 5% cholesterol (light blue). Arrows indicate peaks that are observed in both non-toxic with GM1 (green) and toxic without GM1 (dark blue).

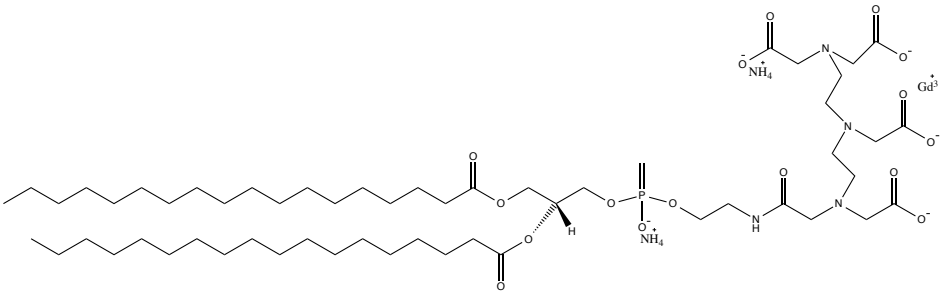
We observe new peaks for the non-toxic oligomer in presence of GM1. Some new peaks detected for non-toxic with GM1 overlay with those observed for the toxic oligomer without GM1. With GM1 the non-toxic exhibits flexible regions noted for the toxic oligomer without GM1. Since these two conditions both correspond to cytotoxic effects this could indicate a relationship between the mobility of oligomers and their toxicity. However, this mechanism is not observed or maintained for the toxic oligomer in presence of GM1.

No changes were observed in the DARR spectrum upon addition of cholesterol and GM1 at any delay time of 3, 20, 50 and 200 ms. This provides support that the flexible regions on the surface of the oligomer are significant in the cellular interactions and cytotoxic responses, rather than the core of the oligomer.

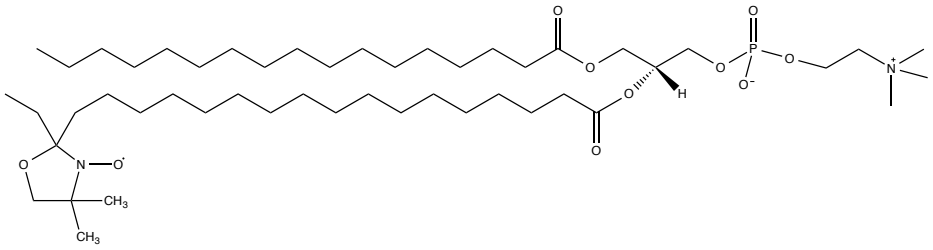
7.6.5 Position of oligomers in the membrane

Aside from determining the internalization of oligomers within the cell it is difficult to ascertain in detail the molecular mechanisms occurring between the oligomer and the cellular membrane from techniques such as cellular toxicity assays, and fluorescence confocal microscopy. ssNMR can provide atomistic information at the residue level of the oligomer, providing a much more detailed understanding of the mechanism of oligomer membrane interaction, in relation to the position and topology of the oligomer.

To gain insights into the position of oligomers within the membrane of the SUV sample, paramagnetic relaxation enhancement (PRE) experiments were performed. Small percentages (2%) of spin-labelled phosphoethanolamine lipids with paramagnetic groups present in the head and tail of the aliphatic chain were incorporated into the SUV composition used. The unpaired electrons within the spin groups increase the transverse relaxation rates of nuclei in proximity, this leads to line broadening and the signal originating from regions of the protein near the spin-label are affected.



DTPA (PE-DTPA (Gd)18:0)



Doxyl (16:0)

Figure 7.31 Structures of PRE-spin labeled lipids

1-palmitoyl-2-stearoyl-(16-doxyl)-*sn*-glycero-3-phosphocholine (Doxyl) has a spin label located on the tail of hydrophobic hydrocarbon chain. The presence of this spin-label will reduce signal from the sample that is located in proximity to the bottom of the SUV. Conversely incorporation of the phospholipid 1,2-distearoyl-*sn*-glycero-3-phosphoethanolamine-N-diethylenetriaminepentaacetic acid (DTPA) that has a spin-label at the hydrophilic head of the lipid will indicate peaks that are in proximity to the top, the surface of the SUV. The use of PRE samples provides general and approximate results; differences in the behavior and location of the oligomer may elucidate how oligomers behave within the cellular membrane, and how these observations respond to cytotoxic effects.

The effect of the PRE spin label arises from an unpaired electron with an isotropic *g*-tensor such as the *S*-state gadolinium ion (Gd^{3+}). The *g*-tensor is defined as the anisotropy and magnitude of the magnetic moment. For the case of Gd^{3+} the electron and nuclear spin interactions are very small. The magnetic dipoles between the unpaired electrons of Gd^{3+} and a neighbouring nucleus result in an increase in nuclear longitudinal and transverse relaxation rates, resulting in the PRE effect and line broadening. The distance dependence of the PRE between the spin label centre and the nucleus of interest is set at r^{-3}, and is therefore often manipulated in studying dynamic and structural aspects of biological macromolecules. The effects of the PRE can extend up to 35 Å, greater than that of NOEs (~6 Å), making it a powerful tool in the study of protein interactions, dynamics and structural studies (Clore & Iwahara 2009).

Analysis is based on signal reduction between reference spectra and PRE spin-labelled DOPC:SM (2:1) SUV samples to ascertain information on the positioning of the oligomer within the bilayer. The addition of cholesterol and GM1 were assessed to determine any changes in oligomer position that related to cytotoxicity. Since the spectra are not assigned we view the oligomer as a single entity and are limited to stating general positions of the oligomer, either at the top or bottom of the vesicle.

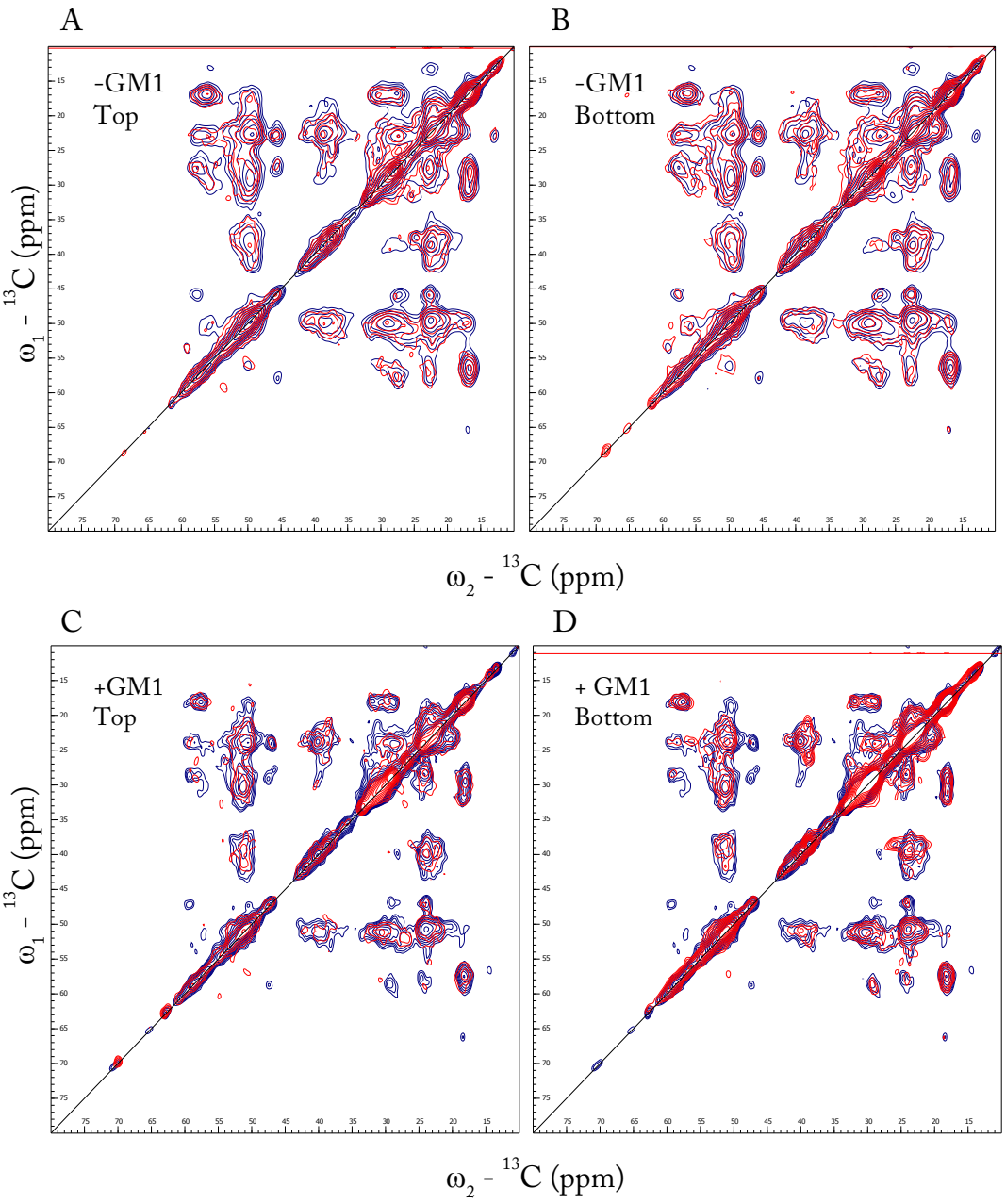


Figure 7.32 PRE effects from ^{13}C-^{13}C DARR spectra of toxic oligomers in presence and absence of GM1

^{13}C-^{13}C DARR with 100 ms mixing time collected for HypF-N toxic oligomers in a DOPC:SM (2:1) SUV with PRE spin labeled lipids in the absence of GM1 (A & B) and presence of GM1 (C & D). Spin labeled lipid samples are shown in red over a control sample prepared without any spin labeled lipids in navy.

(A) 3% DTPA spin lipid top of SUV -GM1 *(B) 2% Doxyl spin lipid bottom of SUV -GM1*

(C) 3% DTPA spin lipid top of SUV +GM1 *(D) 2% Doxyl spin lipid bottom of SUV +GM1*

Attenuation of signal of a number of peaks is noted between Doxyl containing

samples in Figure 7.32 panels B and D (top of SUV), compared to DTPA containing samples in Figure 7.32 panels A and C (bottom of SUV). This indicates that the toxic oligomer appears to be located near to the top of the bilayer, rather than the bottom. We do observe greater attenuation of signal between DTPA and doxyl samples in presence of GM1 (Figure 7.32 (C and D)), indicating that the toxic oligomer is orientated to a greater extent at the top of the bilayer when GM1 is present. These results are consistent in the ^{1}H-^{13}C INEPT spectra (Figure 7.33).

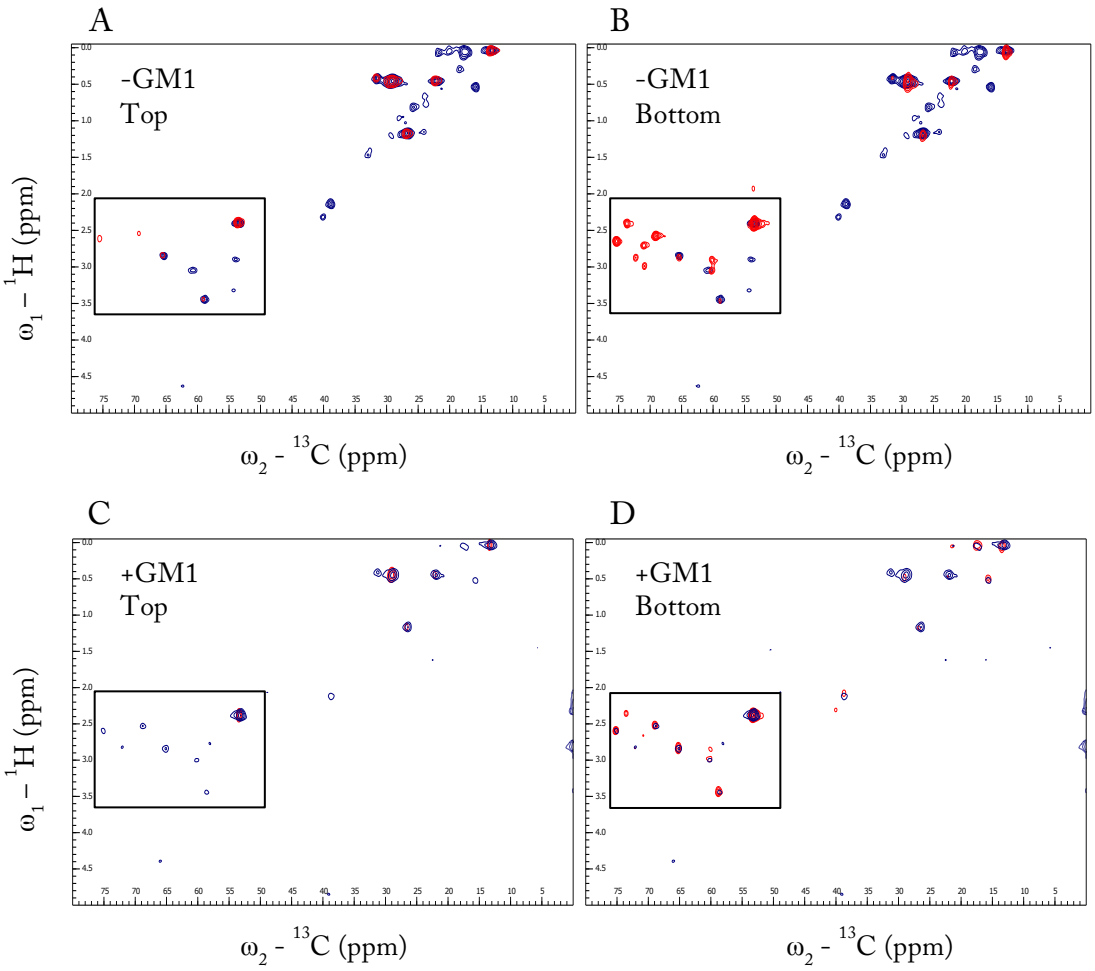


Figure 7.33 PRE effects from ^{1}H-^{13}C INEPT spectra of toxic oligomers in presence and absence of GM1

^{1}H-^{13}C INEPT collected for HypF-N toxic oligomers in a DOPC:SM (2:1) SUV with PRE spin labeled lipids in the absence of GM1 (A & B) and presence of GM1 (C & D). Spin labeled lipid samples are shown in red over a control sample prepared without any spin labeled lipids in navy. Boxed peaks are from protein oligomer.

(A) 3% DTPA spin lipid top of SUV -GM1 *(B) 2% Doxyl spin lipid bottom of SUV -GM1*

(C) 3% DTPA spin lipid top of SUV +GM1 *(D) 2% Doxyl spin lipid bottom of SUV +GM1*

The group of peaks in the ^{1}H-^{13}C INEPT spectra at ~ 0–20 ppm in the ^{13}C dimension are mainly from the lipid components. The majority of the peaks boxed in Figure 7.34 are protein cross-peaks (Figure 7.28); from this region the attenuation of signal between doxyl (bottom) (Figure 7.33 B & D) and DTPA (top) (Figure 7.33 A & C) is easily observed. Without GM1, the PRE effect from DTPA causes greater signal loss compared to doxyl, indicating the oligomers are present at the top of the bilayer. When GM1 is added the attenuation of signal is much greater, corresponding to the results obtained from the DARR spectra (Figure 7.33).

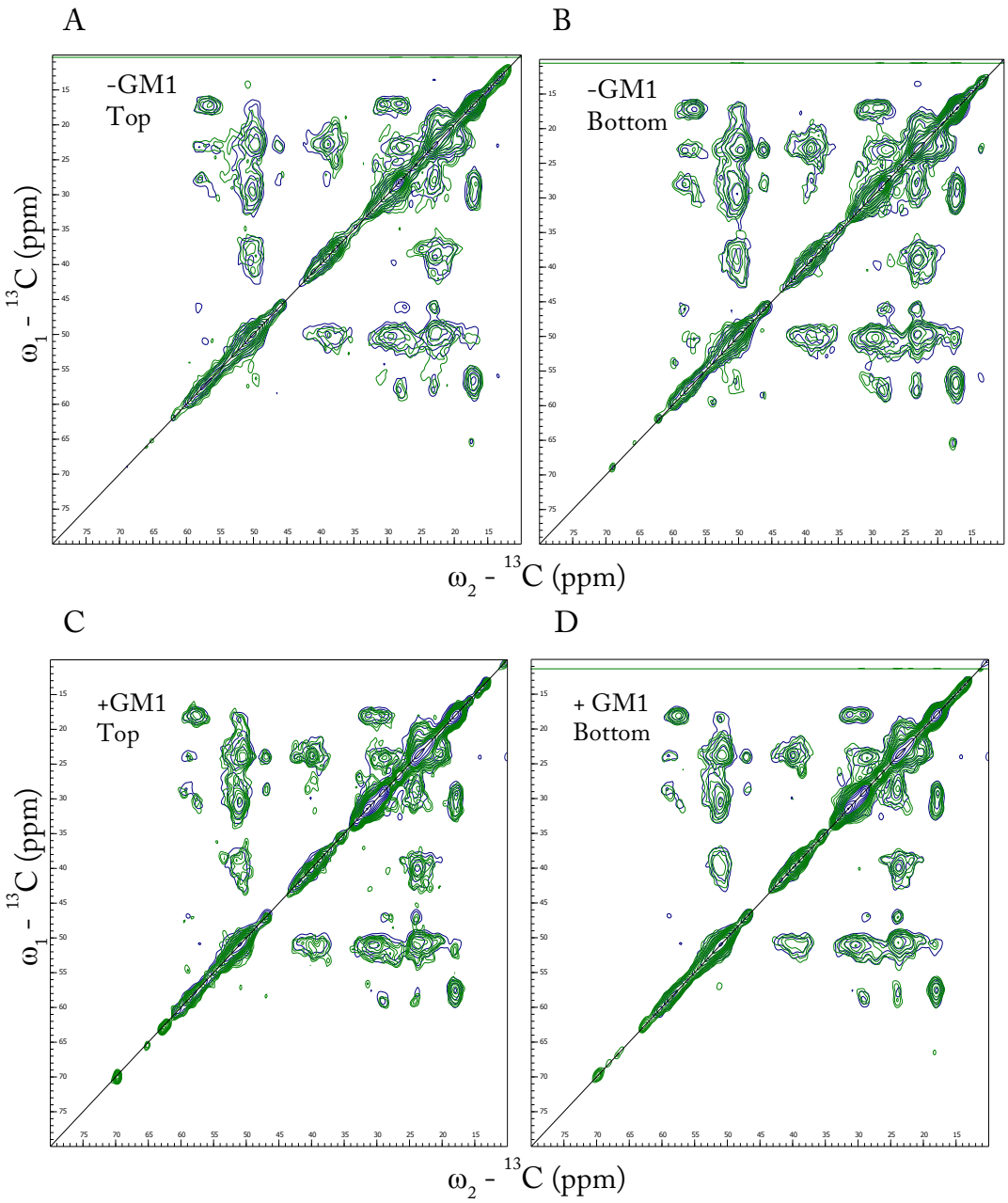


Figure 7.34 PRE effects from ^{13}C-^{13}C DARR spectra of non-toxic oligomers in presence and absence of GM1

^{13}C-^{13}C DARR with 100 ms mixing time collected for HypF-N non-toxic oligomers in a DOPC:SM (2:1) SUV with PRE spin labeled lipids in the absence of GM1 (A & B) and presence of GM1 (C & D). Spin labeled lipid samples are shown in green over a control sample prepared without any spin labeled lipids in navy.

(A) 3% DTPA spin lipid top of SUV -GM1 *(B) 2% Doxyl spin lipid bottom of SUV -GM1*

(C) 3% DTPA spin lipid top of SUV +GM1 *(D) 2% Doxyl spin lipid bottom of SUV +GM1*

For the non-toxic oligomer, there is slight attenuation between DTPA and doxyl

samples (comparing Figure 7.34 (panels A & B and C & D)), indicating that in both absence and presence of GM1 the oligomer is located more at the top of the SUV. The presence of GM1 does not appear to influence the position of the oligomer. We do not observe any significant attenuation of signal upon addition of GM1 (comparing Figure 7.34 (panels A & B to C & D)).

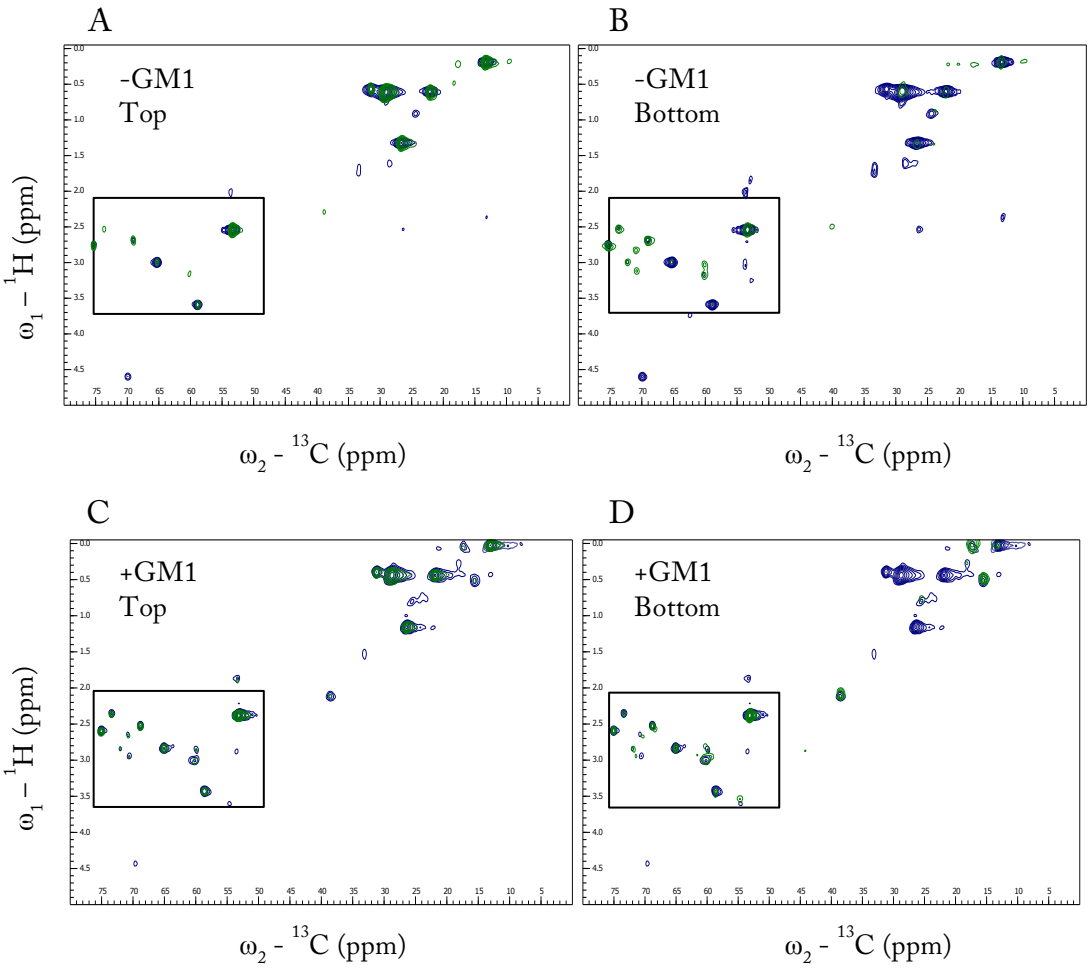


Figure 7.35 PRE effects from ^{1}H-^{13}C INEPT spectra of non-toxic oligomers in presence and absence of GM1

^{1}H-^{13}C INEPT collected for HypF-N non-toxic oligomers in a DOPC:SM (2:1) SUV with PRE spin labeled lipids in the absence of GM1 (A & B) and presence of GM1 (C & D). Spin labeled lipid samples are shown in green over a control sample prepared without any spin labeled lipids in navy. Boxed peaks are from protein oligomer.

(A) 3% DTPA spin lipid top of SUV -GM1 *(B) 2% Doxyl spin lipid bottom of SUV -GM1*

(C) 3% DTPA spin lipid top of SUV +GM1 *(D) 2% Doxyl spin lipid bottom of SUV +GM1*

The ^{1}H-^{13}C INEPT results for the non-toxic oligomer are consistent with the ^{13}C-^{13}C spectra. In the absence of GM1 we observe slight attenuation of signal between doxyl (bottom) (Figure 7.35 B) and DTPA (top) (Figure 7.35 A). This indicates that oligomers are situated more at the top of the bilayer, than at the bottom. The addition of GM1 does not appear to have a significant effect on oligomer positioning as shown by ^{1}H-^{13}C INEPT.

The negatively charged sialic acid head group of GM1 has been shown to be the crucial chemical factor of GM1 in its interactions with oligomers (Evangelisti *et al*., 2012). For the toxic oligomer, the shift in position of the oligomer, to the top of the membrane upon addition of GM1 can possibly be due to the oligomer binding to the sialic acid head group, located the surface of the vesicle. This however is not observed in the non-toxic condition.

It is difficult to infer physiological behavior of the oligomers from the ssNMR data collected. Without GM1 the toxic oligomers interact more at the top of the bilayer compared to non-toxic oligomers; and this may be an indication of how cell apoptosis occurs for the toxic HypF-N oligomer. However, in the presence of GM1 the positioning of the oligomer cannot be correlated to the cytotoxic response; non-toxic oligomers show no significant difference in position, despite eliciting cytotoxicity. The addition of GM1 may alter the physiochemical nature of the phospholipid vesicle, and it may be that the effects of this environmental change on the oligomer are observed from the ssNMR spectra.

These results indicate that with the inclusion of GM1 both toxic and non-toxic oligomers interact at the top of the membrane, and a greater shift in the toxic position is noted; however the changes in position do not to correlate to cytotoxic response. These results may imply a more complex mechanism for which oligomers initiate apoptosis from interactions with the cell membrane. Further studies are required to determine the detailed interactions between oligomer and GM1; potentially phosphorous (^{31}P) NMR measuring changes to the lipid bilayer may provide such insights.

7.6.6 Concluding remarks

Using NMR techniques we have determined that HypF-N oligomerizes to form toxic and non-toxic oligomers via the same structural elements. Despite different chemical environments and monomeric folded states of the protein the β2 strand is shown to be significantly important for oligomerization. Data presented indicates that it is not the formation and nucleation process of the oligomer that contributes to its toxicity. Using ssNMR we have also shown that the core of both oligomers is similar, and that the differences in toxicity are most likely defined at the oligomer surface, and its interactions to components in the cell membrane. The dynamic nature of the oligomer appears to be a defining factor. In a lipid membrane environment toxic oligomers exhibit a more dynamic and flexible state than non-toxic, and with GM1 non-toxic oligomers exhibit an increase in dynamical nature, which is concurrent with increasing cytotoxicity. The dynamical regions have not been identified at a residual level. The interplay between GM1 and toxic oligomers is however, not consistent with increased dynamics, implying an alternative mechanism defining the toxicity between toxic oligomers and GM1. A single unifying concept that describes the cytotoxic origin of HypF-N oligomers remains to be determined, however our research has provided insights into the factors that may be responsible.

7.7 Future work

Much of the aggregation pathway of HypF-N has been heavily researched, these results alongside data from Aβ and α-synuclein indicate that the oligomeric species are the crucial step in the cytotoxic aggregation pathway; and focus should be directed to the structural determination and biomolecular understanding of cellular membrane interactions of these aggregates.

ssNMR is a rapidly progressing and advancing powerful technique to study amyloid and oligomer samples. It is within this field of NMR that our future work is directed to gain further insights into the defining characteristics of HypF-N oligomer toxicity.

Assignment of the ssNMR spectra of the oligomers will provide much more detailed

information on the surface exposed residues that we believe to be crucial to toxicity. Work is currently under progress within the lab to assign the oligomer ssNMR DARR and INEPT spectrum, and determine the dynamical region and residues identified, which is a challenging task.

We have only presented data pertaining to the oligomer, recording ^{31}P ssNMR on lipid vesicles with unlabeled HypF-N oligomers could provide further information relating to the cell membrane, and how the oligomers affect the membrane. Advanced PRE experiments and ^{31}P ssNMR could give insights into the topology and orientation of the oligomer within the membrane which will add to our understanding and may provide detailed differences in how toxic and non-toxic insert in membranes. In addition to this ssNMR experiments have been recorded for CHO cells incubated with oligomers. Performing NMR of oligomers within a cellular membrane provides more realistic results to physiological conditions, which gives more in-depth understanding and knowledge as to how the oligomers are behaving in cell toxicity assays performed (Evangelisti *et al*., 2012). Preliminary spectra correlate to the observations noted in lipid vesicle membrane samples, validating our findings.

REFERENCES

Ahmad B., Vigliotta I., Tatini F., Campioni S., Mannini B., Winkelmann J., Tiribilli B. and Chiti F. (2011). The induction of α-helical structure in partially unfolded HypF-N does not affect its aggregation propensity. *Protein Engineering Design and Selection* **24**(7), 553–563.

Ahmed B., Winkelmann J., Tiribilli B. and Chiti F. (2010). Searching for conditions to form stable protein oligomers with amyloid-like characteristics: The unexplored basic pH. *Biochemica et Biophysica Acta* **1804**, 223–234.

Aires-Barros M.R., Taipa M.A. and Cabral J.M.S. (1994). Isolation and purification of lipases. In: Wooley P., and Petersen S.B. (Eds) Lipases - their structure, biochemistry and application. *Chemistry University Press.*

Akke M., Liu J., Cavanagh J., Erickson H. and Palmer III A.G. (1998). Pervasive conformational fluctuations on microsecond time scales in a fibronectin type III domain. *Nature Structural Biology.* **5**(1), 55–59.

Alexandrescu A.T., Abeygunawardana C. and Shortle D. (1994). Structure and dynamic of a denatured 131 residue fragment of staphylococcal nuclease: a heteronuclear study. *Biochemistry* **33**, 1063–1072.

Alexandrescu, A.T. Evans P.A., Pitkeathly M., Baum J. and Dobson C.M. (1993). Structure and dynamics of the acid denatured molten globule state of α-lactalbumin: a two-dimensional NMR study. *Biochemistry* **32**, 1707–1718.

Aloulou A., Puccinelli D., De Caro A., Leblond Y. and Carrière F. (2007). A comparative study on two fungal lipases from *Thermomyces lanuginosus* and *Yarrowia lipolytica* shows the combined effects of detergents and pH on lipase adsorption and activity. *Biochemica et Biophysica Acta* **1771**, 1446–1456.

Andreason M., Lorenzen N. and Otzen D. (2015). Interactions between misfolded protein oligomers and membranes: A central topic in neurodegenerative diseases? *Biochemica et Biophysica Acta* **1848**(9), 1897–1907.

Andrew E. R., Bradbury A. and Eades R.G. (1958). Nuclear Magnetic Resonance spectra from a crystal rotated at high speed. *Nature* **182**, 1659–1659.

Andrew E. R., Bradbury A. and Eades R.G. (1959). Removal of dipolar broadening of Nuclear Magnetic Resonance spectra of solids by specimen rotation. *Nature* **183**, 1802–1803.

Apostol M.I., Perry K. and Surewicz W.K. (2013). Crystal structure of a human prion protein fragment reveals a motif for oligomer formation. *J. Am. Chem. Soc.* **135**, 10202–10205.

Baum J., Dobson C. M., Evans P. A. and Hanley C. (1989). Characterization of a partly folded protein by NMR methods: studies on the molten globule state of guinea pig α-lactalbumin. *Biochemistry* **28**, 7–13.

Bax A., and Morris G.A. (1969). An improved method for heteronuclear chemical shift correlation by two-dimensional NMR. *Journal of Magnetic Resonance* **42**(3), 501–505.

Beach H., Cole R., Gill M.L. and Loria J.P. (2005). Conversion of μs – ms enzyme motions in the apo- and substrate-mimicked state. *J. Am. Chem. Soc.* **127**, 9167–9176.

Bemporad F. and Chiti F. (2009). "Native-like aggregation" of the acylphosphatase from *Sulfolobus solfataricus* and its biological implications. *FEBS Letters* **583**, 2630–2638.

Bemporad F., Calloni G., Campioni S., Plakoutsi G., Taddei N. and Chiti F. *(*2006). Sequence and structural determinants of amyloid fibril formation. *Acc. Chem. Res.* ***39***, 620–627.

Bennett A.E., Rienstra C.M., Auger M., Lakshmi K.V., and Griffin R.G. (1995). Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* **103**, 6951–6958.

Binder U., Maier T., and Böck A. (1996). Nickel incorporation into hydrogenase 3 from *Escherichia coli* requires the precursor form of the large subunit. *Arch. Microbiol.* **165**, 69–72.

Boehr D., McElheny D., Dyson H.J. and Wright P.E. (2006). The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **313**, 1638–1642.

Boel E. and Huge-Jensen I.B., (1988). European Patent Application EP 3052-16B1.

Boel E., Christensen T., and Woeldike H.F. (2009). Process for the production of protein products in *Aspergillus*. Novo Nordisk A/S. Patent 08435557[7517668].

Boel E., Christensen T., Woldike H.F. and Huge-Jensen I.B. (1998). *Humicola* lipase produced in aspergillus. US Patent 5,766,912.

Bornscheuer U.T., Huisman G.W., Kazlauskas R.J., Lutz S., Moore J.C. and Robins K. (2012). Engineering the third wave of biocatalysis. *Nature* **485**, 185–194.

Bousset L., Pieri L., Ruiz-Arlandis G., Gath J., Jensen P.H., Habenstein B., Madiona K., Olieric V., Böckmann A., Mir B.H., and Melki R. (2013). Structural and functional characterization of two α-synuclein strains. *Nature Communications* **4**(2575).

Breydo L., Kurouski D., Rasool S., Milton S., Wu J.W., Uversky V.N., Lednev I.K. and Glabe C.G. (2016). Structural differences between amyloid β oligomers. *Biochem. Biophys. Res. Communs.* **477**(4), 700–705.

Brierley R.A., Davis G.R. and Holtz G.C. (1994). Production of insulin-like growth factor-1 in methylotrophic yeast cells. *United States Patent 5,324,639.*

Brocca S., Schmidt-Dannert C., Lotti M., Alberghina L. and Schmid R.D. (1998). Design, total synthesis, and functional overexpression of the *Candida rugosa lip1* gene coding for a major industrial lipase. *Protein Science* **7**, 1415–1422.

Brockman H.L. (1984). General features of lipolysis: reaction scheme, interfacial structure and experimental approaches. In: Borgström B., and Brockman H.L. (Eds) Lipases. *El Sevier Amsterdam,* 3–46.

Brozozowski A.M., Savage H., Verma C.S., Turkenburg J.P., Lawson D.M., Svendsen A. and Patkar S. (2000). Structural Origins of the Interfacial Activation in *Thermomyces (Humicola) lanuginosa* Lipase. *Biochemistry* **39**(49), 15071–15082.

Brune A.K. and Gotz F. (1992). Degradation of lipids by bacterial lipases. In: Winkelman G (Ed) Microbial degradation of natural products. VCH: *Weinhein* 243–266.

Bubeck D. (2015). Unraveling structural polymorphism of amyloid fibers. *Structure* **23**(1), 10–11.

Bucciantini M., Calloni G., Chiti F., Formigli L., Nosi D., Dobson C.M. and Stefani M. (2004). Prefibrillar amyloid protein aggregates share common features of cytotoxicity. *J. Biol. Chem.* **279**(30), 31374–31382.

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**, 507–511.

Butterfield D.A., Drake J., Pocernich C. and Castegna A. (2001). Evidence of oxidative damage in Alzheimer's disease brain: central role for amyloid β-peptide. *Trends Mol. Med.* **7**, 548–554.

Bychkova V.E., Pain R.H. and Ptitsyn O.B. (1988). The molten globule state is involved in the translocation of protein across membranes. *FEBS Letters* **238**, 231–234.

Calamai M., Chiti F. & Dobson C.M. (2005) Amyloid fibril formation can proceed from different conformations of a partially unfolded protein. *Biophys J.* **89**(6) 4201-4210.

Calloni G., Lendel C., Campioni S., Giannini S., Gliozzi A., Relini A., Vendruscolo M., Dobson C.M., Salvatella X. and Chiti F. (2008). Structure and dynamics of a partially folded protein are decoupled from its mechanism of aggregation. *J. Am. Chem. Soc.* **130**(39), 13040–13050.

Camilloni C. Schaal D., Schweimer K., Schwarzinger S. and De Simone A. (2012). Energy landscape of the prion protein helix 1 probed by metadynamics and NMR. *Biophysical Journal* **102**, 158-167.

Camilloni C., De Simone A., Vranken W.F. and Vendruscolo M. (2012). Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry* **51,** 2224–2231.

Campioni S., Mannini B., Zampagni M., Penselfini 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 Chem. Biol.* **6**, 140–147.

Campioni S., Mossuto M.F., Torrassa S., Calloni G., de Laureto P.P., Relini A., Fontana A. and Chiti F. (2008). Conformational properties of the aggregation precursor state of HypF- N. *J. Mol. Biol.* **379**, 554–567.

Canale C., Torrassa S., Rispoli P., Relini A., Rolandi R., Bucciantini M., Stefani M. and Gliozzi A. (2006). Natively folded HypF-N and its early amyloid aggregates interact with phospholipid monolayers and destabilize supported phospholipid bilayers. *Biophys. J.* **91**(12), 4575–4588.

Carr, H.Y., and Purcell, E.M. (1954). Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.* **94**, 630–638.

Casalot L. and Rousset M. (2001). Maturation of the [NiFe] hydrogenases. T*rends Microbiol.* **9**, 228–237.

Caughey B., and Lansbury P.T. (2003). Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annu. Rev. Neurosci.* **26**, 267–298.

Cecchi C., Baglioni S., Fiorillo C., Pensalfini A., Liguri G., Nosi D., Rigacci S., Bucciantini M. & Stefani M. (2005). Insights into the molecular basis of the differing susceptibility of varying cell types to the toxicity of amyloid aggregates. *J. Cell Sci.* **118**, 3459-3470.

Cecchi C., Pensalfini A., Baglioni S., Fiorillo C., Caporale R., Formigli L., Liguri G. & Stefani, M. (2006). Differing molecular mechanisms appear to underlie early toxicity of prefibrillar HypF-N aggregates to different cell types. *FEBS Journal*. **273**, 2206–2222.

Cerf, E., Sarroukh R., Tamamizu-Kato S., Breydo L., Derclaye S., Dufrêne Y.F., Narayanaswami V., Goormaghtigh E., Ruysschaert J.M and Raussens V. (2009). Antiparallel β-sheet: a signature structure of the oligomeric amyloid β-peptide. *Biochem. J.* **421**, 415–423.

Chahinian H., Snabe T., Attias C., Fojan P., Petersen S.B and Carrière F. (2006). How gastric lipase, an interfacial enzyme with a Ser-His-Asp catalytic triad, acts optimally at acidic pH. *Biochemistry* **45**, 993-1001.

Chamberlain A.K., MacPhee C.E., Zurdo J., Morozova-Roche L.A., Hill H.A., Dobson C.M. and Davis J.J. (2000). Ultrastructural organization of amyloid fibrils by atomic force microscopy. *Biophys. J.* **79**, 3282–3293.

Chang J.C. and Rosenthal S.J. (2013). Quantum dot-based single-molecule microscopy for the study of protein dynamics. *NanoBiotechnology Protocols* **1026**, 71–84.

Chen C.S. and Sih C. J. (1989). General aspects & optimization of enantioselective biocatalysis in organic solvents: The use of lipases. *Angew. Chem., Int. Ed. Engl.* **28**, 695–708.

Chenal A., Vernier G., Savarin P., Bushmarina N.A., Gèze A., Guillain F., Gillet D. and Forge V. (2005). Conformational states and thermodynamics of α-lactalbumin bound to membranes: a case study of the effects of pH, calcium, lipid membrane curvature and charge. *J. Mol. Biol.* **349**(4), 890–905.

Chimon S. and Ishii Y. (2005). Capturing intermediate structures of Alzheimer's β-amyloid, Aβ (1-40), by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **127**, 13472–13473.

Chiti F. and Dobson C.M. (2006). Protein misfolding, functional amyloid and human disease. *Annu Rev. Biochem.* **75**, 333-366.

Chiti F., Bucciantini M., Capanni C., Taddei N., Dobson C.M. and Stefani M. (2001). Solution conditions can promote formations of either amyloid protofilaments or mature fibrils from the HypF-N terminal domain. *Protein Science* **10**, 2541–2547.

Chiti F., Calamai M., Taddei N., Stefani M., Ramponi G. and Dobson C.M. (2002). Studies on the aggregation of mutant proteins *in vitro* provide insights into the genetics of amyloid diseases. *Proc. Natl. Acad. Sci. USA.* **99**(4), 16419–16426.

Chiti F., Stefani M., Taddei N., Ramponi G. and Dobson C.M. (2003). Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* **424**, 805–808.

Chiti F., Webster P., Taddei N., Clark A., Stefani M., Ramponi G. and Dobson C.M. (1999). Designing conditions for *in vitro* formation of amyloid protofilaments and fibrils. *Proc. Natl. Acad. Sci. USA.* **96**, 3590–3594.

Choo-Smith L.-P., Garzon-Rodriguez W., Glabe C. and Surewicz W. (1997). Acceleration of amyloid fibril formation by specific binding of Aβ(1–40) peptide to ganglioside-containing membrane vesicles. *J. Biol. Chem.* **272**, 22987–22990.

Christensen T., Woeldike H., Boel E., Mortensen S.B., Hjortshoej K., Thim L. and Hansen M.T. (1988). High level expression of recombinant genes in *Aspergillus oryzae*. *Biotechnology* **6**, 1419–1422.

Clare J.J., Romanos M.A., Rayment F.B., Rowedder J.E., Smith M.A., Payne M.M., Sreekrishna K. and Henwood C.A. (1991). Production of epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strains containing multiple gene copies. *Gene* **105**, 205–212.

Clore G.M. and Iwahara J. (2009). Theory, practice and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem. Rev.* **109**, 4108–4139.

Clubb R.T., Thanabal V. and Wagner G. (1992). A constant-time three dimensional triple-resonance pulse scheme to correlate intraresidue $^{1}H^{N}$, ^{15}N and $^{13}C'$ chemical shifts in ^{15}N-^{13}C labeled proteins. *Journal of Magnetic Resonance* **97**, 213–217.

Colbeau A., Toussaint B., Caballero F.J., Elster C., Delphin C., Smith R.L., Chabert J. and Vignais P.M. (1993). Organization of the genes necessary for hydrogenase expression in *Rhodobacter capsulatus*. Sequence analysis and identification of two *hyp* regulatory mutants. *Mol. Microbiol.* **8**(1), 15–29.

Conway K.A., Lee S.J., Rochet J.C., Ding T.T., Williamson R.E. and Lansbury P.T. (2000). Acceleration of oligomerization not fibrillization is a shared property of both α-synuclein mutations linked to early-onset Parkinson's disease. Implication for pathogenesis and therapy. *Proc. Natl. Acad. Sci. USA*. **97**, 571–576.

Cregg J.M., Cereghino J.L., Shi J. and Higgins D.R., (2000). Recombinant Protein Expression in *Pichia pastoris*. *Molecular Biotechnology* 16, 23–52.

Cregg, J. M., Tolstorukov I., Kusari A., Sunga J., Madden K. and Chappell T. (2009). Expression in the yeast *Pichia pastoris*. *Methods Enzymol.* **463**, 169–189.

Creighton T.E. (1993 Proteins: Structures and Molecular Properties. (2nd Ed). W.H. Freeman and Co (Eds) *New York* 1–507.

Cremades N., Cohen S.I.A., Deas E., Abramov A.Y., Chen A.Y., Orte A., Sandal M., Clarke R.W., Dunne P., Aprile F.A., Bertoncini C.W., Wood N.W., Knowles T.P.J., Dobson C.M. and Klenerman D. (2012). Direct observation of the interconversion of normal and toxic forms of α-synuclein. *Cell* **149**(5), 1048–1059.

Cummings C.J., Mancini M.A., Antalffy B., DeFranco D.B., Orr H.T. and Zoghbi H.Y. (1998). Chaperone suppression of aggregation and altered subcellular proteasome localization imply protein misfolding in SCA1. *Nat. Genet.* **19**(2), 148–154.

Cygler M. and Schrag J.D. (1997). Structure as basis for understanding interfacial properties of lipases. *Lipases, Part A* **284**, 3–27.

Czapinska H., Otlewski J., Krzywda S., Sheldrick G.M. and Jaskolski M. (2000). High-resolution structure of bovine pancreatic trypsin inhibitor with altered binding loop sequence. *J. Mol. Biol. 295*, 1237–1249.

Daly R. and Hearn M.T. (2006). Expression of the human activin type I and II receptor extracellular domains in *Pichia pastoris*. *Protein Expr. Purif.* **46**(2), 456–467.

Danielsen S., Eklund M., Deussen H.J., Gräslund T., Nygren P.A. and Borchert T.V. (2001). *In vitro* selection of enzymatically active lipase variants from phage libraries using a mechanism-based inhibitor. *Gene* **272**, 267–274.

De Simone A., Dhulesia A., Soldi G., Vendruscolo M., Hsu S.T., Chiti F. and Dobson C.M. (2011). Experimental free energy surfaces reveal the mechanisms of maintenance of protein solubility. *Proc. Natl. Acad. Sci. USA.* **108**(52), 21057–2106.

De Simone A., Kitchen C., Kwan A.H,. Sunde M., Dobson C.M. and Frenkel D. (2012). Intrinsic disorder modulates protein self-assembly and aggregation *Proc. Natl. Acad. Sci. USA.* **109**(18), 6951–6956.

De Simone A., Richter B., Salvatella X. and Vendruscolo M. (2009). Toward an accurate determination of free energy landscapes in solution states of proteins. *J. Am. Chem. Soc.* **131**(11), 3810–3811.

Dedmon M.M., Lindorff-Larsen K., Christodoulou J., Vendruscolo M. and Dobson C.M. (2005). Mapping long-range interactions in α-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *J. Am. Chem. Soc.* **127**(2), 476-477.

Delaglio F., Grzesiek S., Vuister G.W., Zhu G., Pfeifer J. and Bax A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol. NMR* **6**(3), 277–293.

Delorme V. Dhouib R., Canaan S., Fotiadu F., Carrière F. and Cavalier J.F. (2011). Effects of surfactants on lipase structure, activity, and inhibition. *Pharm. Res.* **28**(8), 1831–1842.

Demuro A., Mina E., Kayed R., Milton S.C., Parker I. and Glabe C.G. (2005). Calcium dysregulation and membrane disruption as a ubiquitous neurotoxic mechanism of soluble amyloid oligomers. *J. Biol. Chem.* **280**(17), 17294–17300.

Denton H., Smith M., Husi H., Uhrin D., Barlow P.N., Batt C.A. and Sawyer L. (1998). Isotopically labeled bovine β-lactoglobulin for NMR studies expressed in *Pichia pastoris*. *Protein Expression and Purification.* **14**(1), 97–103.

Derewendra U., Swenson L., Green R., Wei Y., Dodson G.G., Yamaguchi S., Haas M.J. and Derewenda Z.S. (1994). An unusual buried polar cluster in a family of fungal lipases. *Nat. Struct. Biol.* **1**(1), 36–47.

Detken A., Hardy E.H., Ernst M., and Meier B.H. (2002) Simple and efficient decoupling in magic-angle spinning solid-state NMR: the XiX scheme. *Chemical Physics Letters* **356**, 298-304

Dickinson E. and Rodriguez Patino J.M. (1999). Food emulsions and foams: interfaces, interactions and stability. Cambridge: *Royal Society of Chemistry* **227**.

Dobson C.M. (2001). The structural basis of protein folding and its links with human disease. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **356**(1406), 133–145.

Dobson C.M. (2002). Getting out of shape. *Nature* **418,** 729–730.

DuBay K.F., Pawar A.P., Chiti F., Zurdo J., Dobson C.M. and Vendruscolo M. (2004). Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. *J. Mol. Biol.* **341**, 1317–1326.

Duer M.J. (2004). Solid-state NMR spectroscopy. *Blackwell Publishing Ltd.*

Dyson H.J. and Wright P.E. (2004). Unfolded proteins and protein folding studied by NMR. *Chem. Rev.* **104**(8), 3607–3622.

Eanes E.D. and Glenner G.G. (1968). X-ray diffraction studies on amyloid filaments. *J. Histochem. Cytochem.* **16**(11), 673–677.

Ehehalt R., Keller P., Haass C., Thiele C. and Simons K. (2003). Amyloidogenic processing of the Alzheimer β-amyloid precursor protein depends on lipid rafts. *J. Cell Biol.* **160**, 113–123.

Elliott J.I., Surprenant A., Marelli-Berg F.M., Cooper J.C., Cassady-Cain R. L., Wooding C., Linton K., Alexander D.R. and Higgins C.F. (2005). Membrane phosphatidylserine distribution as a non-apoptotic signaling mechanism in lymphocytes. *Nature Cell Biology* **7**, 808–816.

Elliott S., Lorenzini T., Asher S., Aoki K., Brankow D., Buck L., Busse L., Chang D., Fuller J., Grant J., Hernday N., Hokum M., Hu S., Knudten A., Levin N., Komorowski R., Martin F., Navarro R., Osslund T., Rogers G., Rogers N., Trail G., Egrie J. (2003). Enhancement of therapeutic protein *in vivo* activities through glycoengineering. *Nature Biotechnology* **21**, 414–421.

Evangelisti E., Cecchi C., Cascella R., Sgromo C., Becatti M., Dobson C.M., Chiti F. & Stefani M. (2012). Membrane lipid composition and its physicochemical properties define cell vulnerability to aberrant protein oligomers. *J. Cell Sci.* **125,** 2416-2427.

Fandrich M., Fletcher M.A. and Dobson C.M. (2001). Amyloid fibrils from muscle myoglobin. *Nature* **410**, 165–166.

Fändrich M., Forge V., Buder K., Kittler M., Dobson C.M. and Diekmann S. (2003). Myoglobin forms amyloid fibrils by association of unfolded polypeptide segments. *Proc. Natl. Acad. Sci. USA.* **100**, 15463–15468.

Fawzi N.L., Ying J., Ghirlando R., Torchia D.A. and Clore G.M. (2011). Atomic resolution dynamics on the surface of amyloid β protofibrils probed by solution NMR. *Nature*. **480**(7376), 268–272.

Feng L.R., Lewis N.A.H., Hodges R.S. and McElhaney R.N. (2002). Effect of variations in the structure of a polyleucine-based α-helical transmembrane peptide on its interaction with phosphatidylcholine Bilayers. *Biochemistry,* **41***,* 9197–9207.

Fernández C. and Wider G. (2003). TROSY in NMR studies of the structure and function of large biological macromolecules. *Current Opinion in Structural Biology* **13**, 570–580.

Fersht A.R. (1999). Structure and Mechanism in Protein Science: A guide to Enzyme Catalysis and Protein Folding. (3rd Ed) Freeman: *New York.*

Fezoui Y. and Teplow D.B. (2002). Kinetic studies of amyloid β-protein fibril assembly. Differential effects of α-helix stabilization. *J. Biol. Chem.* **277**, 36948–36954.

Friedrich B. and Schwartz E.B. (1993). Molecular biology of hydrogen utilisation in aerobic chemolithotrophs. *Annu. Rev. Microbiol.* **47**, 351–383

Fujita, A., Cheng, J., Hirakawa, M., Furukawa, K., Kusunoki, S., and Fujimoto, T. (2007). Gangliosides GM1 and GM3 in the living cell membrane form clusters

susceptible to cholesterol depletion and chilling. *Mol. Biol. Cell* **18**, 2112–2122.

Fung B,M., Khitrin A.K., and Ermolaev K. (2000). An improved broadband decoupling sequence for liquid crystals and solids. *J. Magn. Reson.* **142**, 97–101.

Fusco G., De Simone A., Arosio P., Venduscolo M., Veglia G., and Dobson C.M. (2016). Structural ensembles of membrane-bound α-synuclein reveal the molecular determinants of synaptic vesicle affinity. *Nature Scientific Reports* **6**(27125).

Fusco G., De Simone A., Gopinath T., Vostikov V., Venduscolo M., Dobson C.M. and Veglia G. (2014). Direct observation of the three regions in α-synuclein that determine its membrane-bound behavior. *Nature Comms*. **5**(3827).

Fusco G., De Simone A., Hsu S.T., Bemporad F. Vendruscolo M., Chiti F. and Dobson C.M. (2012). ^{1}H, ^{13}C and ^{15}N resonance assignments of human muscle acylphosphatase. *Biomol. NMR Assign.* **6**(1), 27–29.

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. *J. Mol. Biol.* **401**, 115–133.

Giehm L., Svergun D.I., Otzen D.E. and Vestergaard B. (2011). Low-resolution structure of a vesicle disrupting α-synuclein oligomer that accumulates during fibrillation. *Proc. Natl. Acad. Sci. USA*. **108**, 3246–3251.

Gill S.C. and Vonhippel P.H. (1989). Calculation of protein extinction coefficients from amino-acid sequence data. *Anal. Biochem.* **182**, 319–326.

Glenner G., Eanes E.D., Bladen H.A., Linke R.P. and Termine J.D. (1974). β-pleated sheet fibrils. A comparison of native amyloid with synthetic protein fibrils. *J. Histochem. Cytochem.* **22**(12), 1141–1158.

Goddard T. D., and Kneller D. G., SPARKY 3, University of California, San Francisco.

Goodman M., Naider F. & Toniolo C. (1971). Circular dichroism studies of isoleucine oligopeptides in solution. *Biopolymers* **10**, 1719-1730.

Gosline J.M., DeMont M.E. and Denny M.W. (1986). The structure and properties of spider silk. *Endeavour* **10**, 37–43.

Grzesiek S. and Bax A. (1992a). Correlation of backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114**, 6291–6293.

Grzesiek S. and Bax A. (1992b). An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *Journal of Magnetic Resonance.* **99**, 201–207.

Guijarro J.I., Sunde M., Jones J.A., Campbell I.D. and Dobson C.M. (2005). Amyloid fibril formation by an SH3 domain. *Proc. Natl Acad. Sci. USA*. **95**, 4224–4228.

Gupta N., Shai V. and Gupta R. (2007). Alkaline lipase from a novel strain *Burkholderia multivorans*: Statistical medium optimization and production in a bioreactor. *Process Biochem.* **42**, 518–526.

Gupta R., Rathi P., Gupta N. and Bradoo S. (2003). Lipase assays for conventional and molecular screening: An overview. *Biotechnol. Appl. Biochem*. **37**, 63–71.

Haki G.D. and Rakshit S.K. (2003). Developments in industrially important thermostable enzymes: a review. *Bioresource Technology* **89**, 17–34.

Hammarström P., Wiseman R.L., Powers E.T. and Kelly J.W. (2003). Prevention of transthyretin amyloid disease by changing protein misfolding energetics. *Science* **299**, 713–716.

Hartman S. and Hahn E. (1962). Nuclear double resonance in the rotating frame. *Physical Review* **128**, 2042–2053.

Hasnain S., Adeli K. and Storer A.C. (1992). Purification and characterization of an extracellular thiol-containing serine proteinase from *Thermomyces lanuginosus*. *Biochem. Cell Biol.* 70, 117–122.

Hedin E.M.K., Høyrup P., Patkar S.A., Vind J., Svendsen A., and Hult K. (2005). Implications of surface charge and curvature for the binding orientation of *Thermomyces lanuginosus* lipase on negatively charged or zwitterionic phospholipid vesicles as studied by ESR spectroscopy. *Biochemistry* **44**, 16658−16671.

Helenius A. and Simons K. (1975). Solubilization of membranes by detergents. *Biochimica et Biophysica Acta.* **415**, 29–79.

Helenius A., McCaslin D.R., Fries E. and Tanford C. (1979). Properties of Detergents. *Methods in Enzymology.* **11**, 734−749.

Helms L.R. & Wetzel R. (1996). Specificity of abnormal assembly in immunoglobulin light chain deposition disease and amyloidosis. *J. Mol. Biol.* **257**, 77–66.

Huang D., Zimmerman M.I., Martin P.K., Nix A.J., Rosenberry T.L. and Paravastu A. K. (2015). Antiparallel β-sheet structure within the C-terminal region of 42-Residue Alzheimer's amyloid-β peptides when they form 150-kDa oligomers. *J. Mol. Biol.* **427**, 2319–2328.

Ibel K., May R. P., Kirschner K., Szadkowski H., Mascher E. and Lundahl P. (1990). Protein-decorated micelle structure of sodium-dodecyl-sulfate-protein complexes as determined by neutron scattering. *Eur. J. Biochem.* **190**, 311–318.

Ii K., Ito H., Tanaka K. and Hirano A. (1997). Immunocytochemical co-localization of the proteasome in ubiquitinated structures in neurodegenerative diseases and the elderly. *J. Neuropathol. Exp. Neurol.* **56**(2), 125–131.

Inan M., Meagher M.M., Plantz B.A. and Sinha J. (2005). Causes of proteolytic degradation of secreted recombinant proteins produced in methylotrophic yeast *Pichia*

pastoris: Case study with recombinant ovine Interferon-T. *Biotecnol. Bioeng.* **89**(1), 102-112.

Iwamoto N., Suzuki Y., Makino Y., Haga C., Kosaka K. and Iizuka R. (1990). Cell membrane changes in brains manifesting senile plaques: an immunohistochemical study of GM1 membranous ganglioside. *Brain Res.* **522**, 152–156.

Jaeger K.E. and Reetz T.M. (1998). Microbial lipases from versatile tools for biotechnology. *Trends Biotechnol.* **16**, 396 – 403.

Jaenicke, R. and Bohm, G. (1998). The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* **8**, 738−748.

Joseph B., Ramteke P.W. and Thomas G. (2008). Cold active microbial lipases: some hot issues and recent developments. *Biotechnology Advances*, **26**, 457–470.

Jung Y.S. and Zweckstetter M. (2004). Mars - robust automatic backbone assignment of proteins. *J. Biomol. NMR* **30**, 11–23.

Jürgens D. and Huser H. (1981). Large Scale Purification of Staphylococcal lipase by Hydrophobic Interaction Chromatography. *J. Chromatography* **216**, 216–301.

Jutila A., Zhu K., Tuominen E.K.J. and Kinnunen P.K.J. (2004). Fluorescence spectroscopic characterization of *Humicola lanuginosa* lipase dissolved in its substrate. *Biochimica et Biophysica Acta.* **1702**, 181−189.

Kagan, B. L., Azimov, R. and Azimova, R. (2004). Amyloid peptide channels. *J. Membr. Biol.* **202**, 1–10.

Karplus M. and Kuriyan J. (2005). Molecular dynamics and protein function. *Proc. Natl Acad. Sci. USA*. **102**(19), 6679–6685.

Kawahara M. and Kuroda Y. (2000). Molecular mechanism of neurodegeneration induced by Alzheimer's β-amyloid protein: channel formation and disruption of calcium homeostasis. *Brain Res. Bull.* **53**(4), 389–397.

Kay L., Ikura M., Tschudin R. and Bax A. (1990). Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. **89**(3), 496–514.

Kayed R., Head E., Thompson J.L., McIntire T.M., Milton S.C., Cotman C.W. and Glabe C. (2003). Common structures of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* **300**(5618), 486–489.

Kayed R., Pensalfini A., Margol L., Sokolov Y., Sarsoza F., Head E., Hall J. and Glabe C. (2009). Annular protofibrils are a structurally and functionally distinct type of amyloid oligomer. *J. Biol. Chem.* **284**, 4230–4237.

Keeler J. (2014). Understanding NMR spectroscopy. *John Wiley & Sons, Ltd*.

Kelly J. (1998). Alternative conformation of amyloidogenic proteins and their multi-step assembly pathways. *Curr. Opin. Struct. Biol.* **8**, 101–106.

Kim T.R., Goto Y., Hirota N., Kuwata K., Denton H., Wu S.Y., Sawyer L. and Batt C.A. (1997). High-level expression of bovine β-lactoglobulin in *Pichia pastoris* and characterization of its physical properties. *Protein Engineering* **10**(11), 1339–1345.

Kirkitadze M.D., Condron M.M. and Teplow D.B. (2001). Identification and characterization of key kinetic intermediates in amyloid β-protein fibrillogenesis. *J. Mol. Biol.* **312**, 1103–1119.

Kourie J.I. and Henry C.L. (2002). Ion channel formation and membrane-linked pathologies of misfolded hydrophobic proteins: the role of dangerous unchaperoned molecules. *Clin. Exp. Pharmacol. Physiol.* **29**, 741–753.

Kourie J.I. and Shorthouse A.A. (2000). Properties of cytotoxic peptide-induced ion channels. *Am. J. Physiol. Cell Physiol.* **278**(6), 1063–1087.

Krebs, M.R., Wilkins D.K., Chung E.W., Pitkeathly M.C., Chamberlain A.K., Zurdo J., Robinson C.V. and Dobson C.M. (2000). Formation and seeding of amyloid fibrils from wild-type hen lysozyme and a peptide fragment from the β-domain. *J. Mol. Biol.* **300**, 541–549.

Kyte J. and Doolittle R. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132.

Lambert M.P, Barlow A.K., Chromy B.A., Edwards C., Freed R., Liosatos M., Morgan T.E., Rozovsky I., Trommer B., Viola K.L., Wals P., Zhang C., Finch C.E., Krafft G.A. and Klein W.L. (1998). Diffusible nonfibrillar ligands derived from Aβ-42 are potent central nervous system neurotoxins. *Proc. Natl. Acad. Sci. USA*. **95**, 6448–6453.

Lawson D.M., Brzozowski A.M., Dodson G.G., Hubbard R.E., Huge-Jensen B., Boel E. and Derewenda Z.S. (1994). The three-dimensional structure of two fungal lipases. In: Woolley P. and Petersen S.B. (Eds) Lipases—Their Structure, Biochemistry and Applications. *Cambridge University Press*, Cambridge 77–94.

Lawson D.M. Brzozowski A.M., Rety S., Verma C. and Dodson G.G. (1994b). Probing the nature of substrate binding in *Humicola lanuginosa* lipase through X-ray crystallography and intuitive modeling. *Protein Engineering* **7**(4), 543–550.

Lee S. and Rhee J. (1993). Production and partial purification of a lipase from *Pseudomonas putida 3 SK*. *Enzyme Microbiol. Technol.* **15**, 617–623.

Lee S., Fernandez E.J. and Good T.A. Role of aggregation conditions in structure, stability, and toxicity of intermediates in the Aβ fibril formation pathway. *Protein Sci.* **16**, 723–732 (2007).

LeMaster D.M. (1994). Isotope labelling in solution protein assignment and structural analysis *Prog. Nucl. Magn. Reson. Spectrosc.* **26,** 371–419.

LeVine H. (1999). Quantification of β-Sheet Amyloid Fibril Structures with Thioflavin T. *Methods in Enzymology* **309**, 274–284.

Levitt M.H. (2008). Spin dynamics: Basics of nuclear magnetic resonance. *John Wiley & Sons, Inc.*

Lewis R.V. (2006). Spider silk: ancient ideas for new biomaterials. *Chem. Rev.* **106**(9), 3762–3774.

Lin M.C., Mirzabekov T. and Kagan B.L. (1997). Channel formation by a neurotoxic prion protein fragment. *J. Biol. Chem.* **272**, 44–47.

Linko Y., Lamsa M., Huhatala A. and Linko, P. (1994). Lipase catalyzed transesterification of rapeseed oil snf 2-ethyl-1-hexanol. *JAOCS* **71**, 1411–1414.

López de la Paz M., Goldie K., Zurdo J., Lacroix E., Dobson C.M., Hoenger A. and Serrano L. (2002). *De novo* designed peptide-based amyloid fibrils. *Proc. Natl Acad. Sci.* **99**(25), 16052–16057.

Lorenzen N., Nielsen SB., Yoshimura Y., Vad B.S., Andersen C.B., Betzer C., Kaspersen J.D., Christiansen G., Pedersen J.S., Jensen P.H., Mulder F.A.A. and Otzen D.E. (2014). How epigallocatechin gallate can inhibit α-synuclein oligomer toxicity *in vitro*. *J. Biol. Chem.* **289**(31), 21299–21310.

Lowe I. (1959). Free induction decays of rotating solids. *Physical Review Letters* **2**, 285–287.

Luo Y. and Baldwin R.L. (1998). Trifluoroethanol stabilizes the pH 4 folding intermediate of sperm whale apomyoglobin. *J. Mol. Biol.* 279: 49–57.

MacRae A.R. and Hammond R.C. (1995). Present and Future Applications of Lipases *Biotechnol. Genet. Eng. Rev.* **3**, 193–217.

Maheshwari R., Bharadwaj G. and Bhat M.K. (2000). Thermophilic fungi: their physiology and enzymes. *Mol. Biol. Rev. Microbiol.* **64**, 461–488.

Maier T. and Böck A. (1996). Nickel incorporation into hydrogenases. In: Hausinger R.P., Eichhorn G.L., Marzilli L.G. (Eds) Advances in inorganic biochemistry: mechanisms of metallocenter assembly. VCH Publishers Inc., New York, pages 173–192.

Maji S.K., Perrin M.H., Sawaya M.R., Jessberger S., Vadodaria K., Rissman R.A., Singru P.S., Nilsson K.P., Simon R., Schubert D., Eisenberg D., Rivier J., Sawchenko P., Vale W. and Riek R. (2009). Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* **325**(5938), 328–332.

Malmendal A., Underhaug J., Otzen D.E. and Nielsen N.C. (2010). Fast mapping of global protein folding states by multivariate NMR: a GPS for proteins. *PLoS ONE* **5** e10262.

Mansoor S.E., DeWitt M.A. and Farrens D.L. (2010). Distance mapping in proteins using fluorescence spectroscopy: The Tryptophan-Induced Quenching (TrIQ) method. *Biochemistry* **49**, 9722−9731.

Marcon G., Plakoutsi G., Canale C., Relini A., Taddei N., Dobson C.M., Ramponi G.

and Chiti F. (2005). Amyloid Formation from HypF-N under conditions in which the protein is initially in its native State. *J. Mol. Biol.* **347**, 323–335.

Martinelle M., Holmquist M. and Hult K. (1995). On the interfacial activation of *Candida antarctica* lipase A and B as compared with *Humicola lanuginosa* lipase. *Biochimica et Biophysica Acta.* **1258** 272–276.

Martinelle M., Holmquist M. and Hult K. (2005). Activation, inhibition, and destabilization of *Thermomyces lanuginosus* lipase by detergents. *Biochemistry,* **44**, 1719–1730.

Merlini G. and Bellotti V. (2003). Molecular mechanisms of amyloidosis. *N. Engl. J Med.* **349**, 583–596.

Merutka G., Dyson, H.J. & Wright P.E. (1995). 'Random coil' 1H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. *J. Biomol. NMR* **5,** 14-24.

Milhavet O. and Lehmann S. (2002). Oxidative stress and the prion protein in transmissible spongiform encephalopathies. *Brain Res. Brain Res. Rev.* **38**, 328–339.

Minning S., Schmidt-Dannert C. and Schmid R.D. (1998). Functional expression of *Rhizopus oryzae* lipase in *Pichia pastoris*: high-level production and some properties *J. Biotechnol.* **66**, 147–156.

Mittermaier A.K. and Lewis E.K. (2009). Observing biological dynamics at atomic resolution using NMR. *Cell Press, Trends in Biochemical Sciences.* **34**(12), 601–611.

Molander-Melin M., Blennow K., Bogdanovic, N., Dellheden B., Månsson J. E. and Fredman P. (2005). Structural membrane alterations in Alzheimer brains found to be associated with regional disease development; increased density of gangliosides GM1 and GM2 and loss of cholesterol in detergent-resistant membrane domains. *J. Neurochem.* **92**, 171–182.

Montalvao R.W., De Simone A. and Vendruscolo M. (2012). Determination of structural fluctuations of proteins from structure-based calculations of residual dipolar couplings. *J. Biomol. NMR*. **53**, 281-292.

Morris G.A., and Freeman R. (1979). Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc.* **101**(3), 760–762.

Mosmann T. (1983). Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J. Immunol. Methods* **65**, 55–63.

Mossuto M.F., Dhulesia A., Devlin G., Frare E., Kumita J.R., de Laureto P.P., Dumoulin M., Fontana A., Dobson C.M. and Salvatella X. (2010). The non-core regions of human lysozyme amyloid fibrils influence cytotoxicity. *J. Mol. Biol.* **402**, 783–796.

Narayanaswamy R., Levy M., Tsechansky M., Stovall G.M., O'Connell J.D., Mirrielees J., Ellington A.D and Marcotte E.M. (2009). Widespread reorganization of

metabolic enzymes into reversible assemblies upon nutrient starvation. *Proc. Natl. Acad. Sci. USA*. **106**(25), 10147–10152.

Nardini M. and Dijkstra B.W. (1999). α/β Hydrolase fold enzymes: the family keeps growing. *Current Opinion in Structural Biology* **9**, 732–737.

Neudecker P., Robustelli P., Cavalli A., Walsh P., Lundström 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.

Norton W.T., Abe T., Poduslo S.E., and DeVries G.H. (1975). The lipid composition of isolated brain cells and axons. *J. Neurosci. Res.* **1**, 57–75.

Omar I.C., Hayashi M. and Nagai S. (1987a). Purification and Some Properties of a Thermostable Lipase from *Humicola lanuginosa* No. 3. *Agric. Biol. Chem.* **51**, 37–45.

Omar I.C., Nishio N. and Nagai S. (1987b). Productions of a thermostable lipase by *Humicola lanuginosa* grown on sorbitol-corn steep liquor medium. *Agric. Biol. Chem.* **51**, 2145–2151.

Ong P.S. and Gaucher G. M. (1973). Protease production by thermophilic fungi. *Can. J. Microbial.* **19**, 129–133.

Otzen D. (2011). Protein-surfactant interactions: A tale of many states. *Biochemica et Biophysica Acta.* **1814**, 562–591.

Otzen D.E., Kristensen O. and Oliverberg M. (2000). Designed protein tetramer zipped together with a hydrophobic Alzheimer homology: a structural clue to amyloid assembly. *Proc. Natl. Acad. Sci. USA*. **97**, 9907–9912.

Palmer III A.G., Williams J. and McDermott A. (1996). Nuclear magnetic resonance studies of biopolymer dynamics. *J. Phys. Chem.* **100**, 13293–13310

Paschos A., Bauer A., Zimmermann A., Zehelein E. and Böck A. (2002). HypF, a carbamoyl phosphate-converting enzyme involved in [NiFe] hydrogenase maturation. *J. Biol. Chem.* **277**(51), 49945–49951.

Paslawski W., Mysling S., Thomsen K., Jorgensen T.J. and Otzen D.E. (2014). Co-existence of two different α-synuclein oligomers with different core structures determined by hydrogen/deuterium exchange mass spectrometry *Angew. Chem. Int. Ed. Engl.* **53**(29), 7560–7563.

Patel M. (2003). Surfactants based on renewable raw materials. *J. Ind. Ecol.* **7**, 47–62.

Patton S. (1970). Correlative relationship of cholesterol and sphingomyelin in cell membranes. *J. Theor. Biol.* **29**, 489–491.

Pawar A.P., Dubay, K.F., Zurdo J., Chiti F., Vendruscolo M. and Dobson C.M. (2005). Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol.* **350**, 379–392.

Pervushin K., Riek R., Wider G. and Wüthrich K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy

indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. USA*. **94**, 12366–12371.

Pervushin K. (2000). Impact of Transverse Relaxation Optimized Spectroscopy (TROSY) on NMR as a technique in structural biology. *Q. Rev. Biophys.* **33**(2), 161–197.

Pines A., Gibby M.G. and Waugh J.S. (1972). Proton-enhanced nuclear induction spectroscopy, A method for high resolution NMR of dilute spins in solids. *J. Chem. Phys.* **56**, 1176–1777.

Pines A., Gibby M.G. and Waugh J.S. (1973). Proton-enhanced NMR of dilute spins in solids. *J. Chem. Phys.* **59**, 569–590.

Plakoutsi G., Bemporad F., Calamai M., Taddei N., Dobson C.M. and Chiti F. (2005). Evidence for a mechanism of amyloid formation involving molecular reorganisation within native-like precursor aggregates *J. Mol. Biol.* **351**, 910–922.

Plakoutsi G., Taddei N., Stefani M. and Chiti F. (2004). Aggregation of the acylphosphatase from *Sulfolobus solfataricus*: the folded and partially unfolded states can both be precursors for amyloid formation. *J. Biol. Chem.* **279**, 14111–14119.

Prince M., Wimo A., Guerchet M., Ali G.M., Wu Y.T. and Prina M. (2015). World Alzheimer Report 2015: The global impact of dementia, an analysis of prevalence, incidence, cost and trends. *Alzheimer's Disease International.*

Rahimi F., Shanmugam A. and Bitan G. (2008). Structure-function relationships of pre-fibrillar protein assemblies in Alzheimer's disease and related disorders. *Curr. Alzheimer Res.* **5**, 319–341.

Rajagopalan P.T., Zhang Z., McCourt L., Dwyer M., Benkovic S.J. and Hammes G.G. (2002). Interaction of dihydrofolate reductase with methotrexate: Ensemble and single-molecule kinetics. *Proc. Natl. Acad. Sci. USA.* **99**(21), 13481–13486.

Relini A., Torrassa S., Rolandi R., Gliozzi A., Rosano C., Canale C., Bolognesi N., Plakoutsi G., Bucciantini N., Chiti F. and Stefani M. (2004). Monitoring the process of HypF fibrillization and liposome permeabilization by protofibrils. *J. Mol. Biol.* **338**(5), 943–957.

Richardson J.S. and Richardson D.C. (2002). Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation. *Proc. Natl. Acad. Sci. USA.* **99**(5), 2754–2759.

Roccatano D., Colombo G., Fioroni M. and Mark A.E. (2002). Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: a molecular dynamics study. *Proc. Natl. Acad. Sci. USA*. **99**, 12179–12184.

Rogalska E., Nury S., Douchet I. and Verger R. (1995). Lipase stereoselectivity and regioselectivity toward three isomers of dicaprin: a kinetic study by the monomolecular film technique, *Chirality* **7**, 505–515.

Romdhane I.B., Frikha F., Maalej-Achouri I., Gargouri A. and Belghith H. (2012).

Gene cloning and molecular characterizations of the *Talaromyces thermophilus* lipase catalyzed efficient hydrolysis and synthesis of esters. *Gene* **494**, 112–118.

Rosano C., Zuccotti S., Bucciantini M., Stefani M., Rampioni G. and Bolognesi M. (2002). Crystal structure and anion binding in the prokaryotic hydrogenase maturation factor HypF acylphosphatase-like domain. *J. Mol. Biol.* **321**(5), 785–796.

Sachdev A. and Krishnan S. (1997). Heavy-duty liquid detergents. In: Liquid Detergents. Lai K.Y. (Eds) *Marcel Dekker*: New York 261–324.

Salzmann M., Pervushin K., Wider G., Senn H. and Wüthrich K. (1998). TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins. *Proc. Natl. Acad. Sci. USA* **95**, 13585–13590.

Sardra L. and Desnuelle P. (1958). Actions of pancreatic lipase on esters in emulsions. *Biochimica et Biophysica Acta.* **30**, 513–521.

Sauer K., Rickard A.H., and Davies D.G. (2007). Biofilms and biocomplexity. *Microbe* **2**, 347–353.

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-β spines reveal varied steric zippers. *Nature*, **447**(7143), 453–457.

Saxena R.K., Sheoran A., Giri B. and Davidson W.S. (2003). Purification strategies for microbial lipases. *J. Microbiological Methods* **52**, 1–18.

Schrag J.D. and Cygler M. (1997). Lipases and α/β hydrolase fold. *Methods Enzymol.* **284**(4), 85–107.

Seeling J. and Seeling A., (1974). Deuterium magnetic resonance studies of phospholipid bilayers. *Biochem. Biophys. Res. Commun.,* **57***,* 406–411.

Serpell L., Sunde M., Benson M.D., Tennent G.A., Pepys M.B. and Fraser P.E. (2000). The protofilament substructure of amyloid fibrils. *J. Mol. Biol*. **300**(5), 1033–1039.

Serpell L.C., Fraser P.E. and Sunde M. (1999). X-ray fiber diffraction of amyloid fibrils. *Methods Enzymol.* **309**, 526–536.

Sharma R., Chisti Y. and Banerjee U.C. (2001). Production, purification, characterization, and applications of lipases. *Biotechnology Advances* **19**, 627–662.

Sharma R., Soni S., Vohra R., Gupta L., Gupta J., (2002). Purification and characterization of a thermostable alkaline lipase from a new thermophilic *Bacillus sp.* RSJ-1. *Proc. Biochem.* **37**, 1075–1084.

Shenolikar S. and Stevenson K.J. (1982). Purification and partial characterization of a thiol proteinase from the thermophilic fungus *Humicola lanuginosus. Biochem. J.* **205**, 147–152.

Sherman F. (2002). Getting started with yeast. *Meth Enzymol.* **350**, 3–41.

Simons J.W., van Kampen M.D., Riel S., Götz F., Egmond M.R. and Verheij H.M. (1998). Cloning, purification and characterisation of the lipase from *Staphylococcus epidermidis*: Comparison of the substrate selectivity with those of other microbial lipases. *Eur. J. Biochem.* **253**(3), 675–683.

Slotte J.P., and Ramstedt B. (2007). The functional role of sphingomyelin in cell membranes. *Eur. J. Lipid Sci. Technol.* **109**, 977–981.

Slupsky C.M., Kay C.M., Reinach F.C., Smillie L.B. & Sykes B.D. (1995). Calcium-Induced Dimerization of Troponin C: Mode of Interaction and Use of Trifluoroethanol as a Denaturant of Quaternary Structure. *Biocehmistry.* **34**, 7365-7375.

Smulders E. (2002). Laundry detergents. *Wiley* VCH Verlag, Weinheim.

Soldi G., Bemporad F., Torrassa S., Relini A., Ramazzotti M., Taddei N. and Chiti F. (2005). Amyloid formation of a protein in the absence of initial unfolding and destabilization of the native state. *Biophys. J.* **8**, 4234–4244.

Sonesson A.W., Elofsson U.M., Brismar H., Callisen T.H. (2006). Adsorption and mobility of a lipase at a hydrophobic surface in the presence of surfactants. *Langmuir* **22**, 5810–5817.

Soto C. (2003). Unfolding the role of protein misfolding in neurodegenerative diseases. *Nat. Rev. Neurosci.* **4**(1), 49–60.

Stefani M. (2010). Biochemical and biophysical features of both oligomer/fibril and cell membrane in amyloid cytotoxicity. *FEBS J.* **277**(22), 4602−4613.

Stöckl M., Claessens M.M. and Subramaniam V. (2012). Kinetic measurements give new insights into lipid membrane permeabilization by α-synuclein oligomers. *Mol. BioSyst.* **8**(1), 338–345.

Stratton J., Chiruvolu V. and Meagher M. (1998). High cell-density fermentation. *Methods in Molecular Biology* **103**, 107-120.

Stryer L., and Haugland R.P. (1967). Energy transfer: A spectroscopic ruler. *Proc. Natl. Acad. Sci. USA.* **58**, 719−726.

Sun Z.Y., Frueh D.P., Selenko P., Hoch J.C. and Wagner G. (2005). Fast assignment of ^{15}N-HSQC peaks using high-resolution 3D HNcocaNH experiments with non-uniform sampling. *Journal of Biomolecular NMR*, **33**(1), 43–50.

Sunde M., Kwan A.H., Templeton M.D., Beever R.E. and Mackay J.P. (2008). Structural analysis of hydrophobins. *Micron.* **37**(7), 773–784.

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. *J. Mol. Biol.* **273**(3), 729–739.

Sunna A. and Antranikian G. (1997). Xylanolytic enzymes from fungi and bacteria. *Crit. Rev. Biotechnol.* **17**, 39–67.

Svendsen A., Clausen I. G., Patkar S. A., Borch K., and Thellersen M. (1997). Protein Engineering of Microbial Lipases of Industrial Interest. *Methods Enzymol.* **284**(19), 317–340.

Szeverenyi N.M., Sullivan M.J., and Maciel G.E. (1982). Observation of spin exchange by two-dimensional fourier-transform C-13 cross polarization-magic-angle spinning. *Journal of Magnetic Resonance* **47**(3), 462–475.

Tan S.Y. and Pepys M.B. (1994). Amyloidosis. *Histopathology* **25**(5), 403–414.

Tatini A., Pugliese A.M., Traini C., Niccoli S., Maraula G., Ed Dami T., Mannini B., Scartabelli T., Pedata F., Casamenti F. and Chiti F. (2013). Amyloid-β oligomer synaptotoxicity is mimicked by oligomers of the model Protein HypF-N. *Neurobiology of Aging* **34**(9), 2100–2109.

Takegoshi, K., S. Nakamura, and T. Terao, (2001) C-13-H-1 dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chemical Physics Letters*. **344**(5-6), 631–637.

Taylor F. (1989 Stabilization of lipase from *Thermomyces lanuginosus* with *p*-chloromercuribenzoic acid. *Journal of Fermentation and Bioengineering* **68**(2), 141–143.

Traaseth N. J., Shi L., Verardi R., Mullen D.G., Barany G., and Veglia G. (2009). Structure and topology of monomeric phospholamban in lipid membranes determined by a hybrid solution and solid-state NMR approach. *Proc. Natl. Acad. Sci. USA.* **106**, 10165–10170.

Uversky V.N., Yamin G., Munishkina L.A., Karymov M.A., Millett I.S., Doniach S., Lyubchenko Y.L. and Fink, A.L. (2005). Effects of nitration on the structure and aggregation of α-synuclein. *Brain Res. Mol. Brain Res.* **134**(1), 84–102.

Vallurupalli P., Bouvignies G., and Kay L.E. (2012). Studying "invisible" excited protein states in slow exchange with a major state conformation *J. Am. Chem. Soc.* **134**(19), 8148–816.

Varkey J., Mizuno N., Hegde B.G., Cheng N., Steven A.C. and Langen R. (2013). α-synuclein oligomers with broken helical conformation form lipoprotein nanoparticles. *J. Biol. Chem.* **288**, 17620–17630.

Vendruscolo M. and Dobson C.M. (2006). Dynamic visions of enzymatic reactions. Science **313**, 1586–1587.

Verger R. (1997). Interfacial activation of lipases: Facts and artifacts. *TIBS TECH* **15**, 32–38.

Vernaglia B.A., Huang J. and Clark E.D. (2004). Guanidine hydrochloride can induce

amyloid fibril formation from hen egg-white lysozyme. *Biomacromolecules* **5**, 1362–1370.

Volles M.J. and Lansbury P.T. (2002). Vesicle permeabilization by protofibrillar α-synuclein is sensitive to Parkinson's disease-linked mutations and occurs by a pore-like mechanism. *Biochemistry* **41**, 4595–4602.

Volles M.J. and Lansbury P.T. (2003). Zeroing in on the pathogenic form of α-synuclein and its mechanism of neurotoxicity in Parkinson's disease *Biochemistry* **42** (26), 7871–7878.

Vranken W.F., Boucher W., Stevens T.J., Fogh R.H., Pajon A., Llinas M., Ulrich E.L., Markley J.L., Ionides J. and Laue E.D. (2005). The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins* **59**, 687–696.

Wahlgren M. and Arnebrant T. (1991). Protein adsorption to solid surfaces, *Trends Biotechnol.* **9**, 201–208.

Wales T.E. and Engen J.R. (2006). Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrometry Reviews.* **25**, 158–170.

Wang D., and Geva E. (2005). Protein structure and dynamics from single-molecule Fluorescence Resonance Energy Transfer. *J. Phys. Chem. B* **109**(4), 1626–1634.

Wang H., Hagedorn J., Svendsen A., Borch K. and Otzen D.E. (2013). Variant of the *Thermomyces lanuginosus* lipase with improved kinetic stability: A candidate for enzyme replacement therapy. *Biophysical Chemistry* **172**, 43–52.

Weber G. (1992). Protein Interactions. Chapman and Hall (Ed.) *New York.*

Wolf I., Buhrke T., Dernedde J., Pohlmann A. and Friedrich B. (1998). Duplication of *hyp* genes involved in maturation of [NiFe] hydrogenases in *Alcaligenes eutrophus* H16. *Arch. Microbiol.* **170**(6), 451–459.

Wüthrich K. (1986). NMR of proteins and nucleic acids. *John Wiley & Sons, Inc.* Canada.

Yamamoto N., Matsubara T., Sato T. and Yanagisawa K. (2008). Age-dependent high-density clustering of GM1 ganglioside at presynaptic neuritic terminals promotes amyloid b-protein fibrillogenesis. *Biochemica et Biophysica Acta* **1778**, 2717–2726.

Yamamoto T., Tomiyama M., Mita H., Sode K. and Karube I. (1990). Identification of proteins encoded in *Escherichia coli hydA*, *hydB* and analysis of the *hydA* locus. *FEMS Microbiol. Lett.* **66**, 187–192.

Yang L. and Bahar I. (2005). Coupling between catalytic site and collective dynamics: A requirement for mechanochemical activity of enzymes. *Structure* **13**, 893–904.

Yapoudjian S., Ivanova M. G., Brzozowski A. M., Patkar S. A., Vind J., Svendsen A. and Verger R. (2002). Binding of *Thermomyces* (*Humicola*) *lanuginosa* lipase to the mixed micelles of cis-parinaric acid/NaTDC: Fluorescence resonance energy transfer and crystallographic study. *Eur. J. Biochem.* **269**, 1613−1621.

Yip C.M., Elyton E.A., Darabie A.A., Morrison M.R. and McLaurin J. (2001). Cholesterol, a modulator of membrane-associated Aβ-fibrillogenesis and neurotoxicity. *J. Mol. Biol.* **311**, 723–734.

Zampagni M., Cascella R., Casamenti F., Grossi C., Evangelisti E., Wright D., Becatti M., Liguri G., Mannini B., Campioni S., Chiti F. and Cecchi C. (2011). A comparison of the biochemical modifications caused by toxic and non-toxic protein oligomers in Cells. *J. Cell Mol. Med.* **15**(10), 2106–2116.

Zheng Y.Y., Guo X.H., Song N.N. and Li D.C. (2011). Thermophilic lipase from *Thermomyces lanuginosus*: Gene cloning, expression and characterization. *Journal of Molecular Catalysis B: Enzymatic.* **69**, 127–132.

Zhou W., Long C., Reaney S.H., Di Monte D.A., Fink A.L. and Uversky V.N. (2010). Methionine oxidation stabilizes non-toxic oligomers of α-synuclein through strengthening the auto-inhibitory intra-molecular long- range interactions. *Biochemica et Biophysica Acta* **1802**, 322–330.

Zhu K., Jutila A., Tuominen E.K., Patkar S.A., Svendsen A. and Kinnunen P.K. (2001). Impact of the tryptophan residues of *Humicola lanuginosa* lipase on its thermal stability. *Biochemica et Biophysica Acta* **1547**, 329–338.

Zuccotti S., Rosano C., Ramazzotti M., Degl'Innocenti D., Stefani M., Manao G. and Bolognesi M. (2004). Three-dimensional structural characterization of a novel *Drosophila Melanogaster* acylphosphatase. *Acta Crystallogr. Sect. D* **60**, 1177–1179.

SUPPLEMENTARY MATERIAL

Table S1. Chemical shift values for HypF-N under native condition as measured by NMR.

Grey boxes without data are unassigned or proline residues

Residue	H	NH	Cα	Cβ	CO	Cα-i	Cβ-i	CO-i	Hα	Hα (2)
GLY 1										
SER 2	8.658	115.825	58.230	64.010		43.370		170.231		
ALA 3	8.512	126.387	52.780	19.140	177.695	58.070	64.018	174.230	4.392	
LYS 4	8.377	120.159	56.480	32.994	176.320	52.720	19.114	177.703	4.344	
ASN 5	8.509	120.198	53.457	38.807	175.723	56.372	32.991	176.324	4.830	
THR 6	8.228	114.555	62.393	69.829	174.758	53.491	38.827	175.728	4.447	
SER 7	8.428	117.020	58.912	64.021	173.522	62.275	69.838	174.766	4.672	
CYS 8	8.199	121.189	59.356	28.044	174.251	58.982	63.946	173.534	4.713	
GLY 9	8.701	114.080	45.311		173.868	59.377	28.062	174.246	4.877	4.193
VAL 10	8.826	119.419	60.678	35.832	171.754	45.312		173.856	5.070	
GLN 11	9.300	125.352	54.111	31.991	173.907	60.683	35.824	171.743	5.155	
LEU 12	9.843	125.545	52.466	44.157	175.696	54.151	31.920	173.902	5.571	
ARG 13	9.138	123.722	55.840	31.066	175.228	52.443	44.155	175.700	5.052	
ILE 14	9.529	128.612	61.160	39.366	176.546	55.819	31.061	175.177	4.784	
ARG 15	8.665	125.624	54.431	33.429	175.475	61.041	39.380	174.535	4.994	
GLY 16	8.904	109.659	45.346		172.716	54.346	33.420	175.468	4.615	3.880
LYS 17	9.104	128.959	55.146	31.864	175.560	45.526		172.705	4.787	
VAL 18									2.903	
GLN 19	7.960	120.313	55.180	29.519	176.223	59.972	32.358		4.933	
GLY 20										
VAL 21	8.537	112.931				46.159				
GLY 22	8.922	111.162	46.356		176.526			175.712		
PHE 23										
ARG 24										
PRO 25										
PHE 26	7.159	119.288	60.452	38.263	176.170	66.046	30.705	178.485	4.542	
VAL 27	8.145	120.445	67.185	31.255	176.622	60.349	38.274	176.170	4.303	
TRP 28	8.535	118.787	62.284	29.300	177.222	67.121	31.276	176.617	3.969	
GLN 29	8.487	119.114	59.337	28.032	178.815	62.286	29.279	177.246	3.581	
LEU 30	7.841	121.540	57.638	41.926	178.395	59.227	27.952	178.811	3.788	
ALA 31	8.134	119.607	55.025	16.559	179.845	57.589	42.025	178.385	4.736	
GLN 32	7.571	115.297	57.748	27.543	179.699	55.024	16.568	179.855	3.789	
GLN 33	7.674	119.719	58.716	28.238	177.825	57.710	27.493	179.661	3.891	
LEU 34	7.642	116.438	54.448	42.356	175.188	58.760	28.300	177.837	4.217	
ASN 35	7.509	117.478	54.170	36.866	173.099	54.277	42.191	175.193	4.130	
LEU 36	7.696	114.749	52.886	42.745	176.235	54.147	36.857	173.088	4.355	
HIS 37	10.152	116.628	55.928	30.943	172.704	52.778	42.570	176.256	4.558	
GLY 38	8.714	110.956	45.121		171.449	55.941	30.860	172.697	5.754	
ASP 39	9.338	121.763	53.182	44.200	173.570	45.440	44.133	171.412	4.496	

VAL 40	7.902	113.681	58.556	34.870	171.294	53.020	44.253	173.572	5.178	
CYS 41	8.787	121.339	56.258	32.459	173.205	58.445	34.770	171.294	5.182	
ASN 42	9.226	121.802	53.538	38.486	175.116	56.242	32.423	173.196	5.244	
ASP 43	8.115	123.021	53.901	43.446	175.623	53.232	38.299	175.113	4.778	
GLY 44	8.899	115.176	47.104		174.980	53.264	43.333	175.608	3.704	4.017
ASP 45	8.557	125.101	54.153	40.732	176.221	47.094		174.973	4.757	
GLY 46	8.160	110.026	44.823		173.182	53.967	40.670	176.230	4.535	4.017
VAL 47	8.730	121.367	61.089	36.507	173.723	44.723		173.181	5.006	
GLU 48	9.200	126.631	54.326	31.900	175.541	60.921	36.469	173.716	5.642	
VAL 49	9.574	128.977	60.725	35.017	174.233	54.284	31.908	175.537	5.179	
ARG 50	9.357	127.246	55.398	32.553	175.684	60.749	34.764	174.233	5.481	
LEU 51	9.247	119.835	53.331	47.586	178.954	55.274	32.481	175.682	5.900	
ARG 52	8.671	122.611	58.248	31.282	175.226	53.346	47.673	178.958	4.599	
GLU 53	8.089	119.988			174.640			175.233	3.303	
ASP 54	8.027	117.808	52.520	42.400	177.275	58.330	30.064	174.653	4.724	
PRO 55										
GLU 56	8.573	123.196	62.037	28.566	178.844	63.619	33.045	175.944	4.036	
THR 57	8.529	115.339	66.736	68.203	176.150	61.789	28.609	178.850	4.048	
PHE 58	8.624	123.413	61.719	38.337	175.645	66.721	68.104	176.132	3.735	
LEU 59	8.604	117.771	58.158	41.888	178.398	61.730	38.198	175.647	3.725	
VAL 60	7.559	117.810	67.083	31.831	179.637	58.192	41.865	178.420	3.760	
GLN 61	7.750	117.297	58.496	28.817	177.419	67.076	31.939	179.623	3.954	
LEU 62	8.622	122.689	58.190	42.303	177.426	58.613	28.646	177.426	3.680	
TYR 63	7.471	112.102	61.682	38.461	179.729	58.164	42.326	177.425	3.089	
GLN 64	7.953	116.725	58.814	29.360	177.554	61.597	38.370	179.710	4.019	
HIS 65	8.080	112.443	54.519	28.729	171.454	58.774	29.293	177.565	4.859	
CYS 66	6.872	121.513	58.040	27.275	172.547	54.348	28.563	171.440	3.940	
PRO 67										
PRO 68										
LEU 69	8.063	115.284	55.811	41.610	176.697	64.779	31.983	177.712	4.254	
ALA 70	7.845	120.746	50.298	22.314	175.701	55.690	41.398	176.673	4.880	
ARG 71	7.987	119.622	56.019	33.490	174.208	50.272	22.314	175.689	4.427	
ILE 72	9.001	128.468	61.755	38.632	175.094	56.108	33.643	174.193	3.886	
ASP 73	9.531	128.356	55.718	42.120	176.546	61.455	38.533	175.169	4.900	
SER 74	7.843	111.632	58.266	64.845	171.998	55.682	41.854	176.540	4.642	
VAL 75	8.367	120.162	62.090	34.781	175.273	58.010	64.859	171.997	4.997	
GLU 76	8.927	128.298	54.805	32.172	175.033	61.709	34.832	175.273	4.785	
ARG 77	8.926	124.287	54.555	33.758	175.010	54.734	32.081	174.887	5.443	
GLU 78	8.904	123.519	54.188	32.476	172.653	54.540	33.581	174.848	5.064	
PRO 79										
PHE 80	7.651	121.598	57.608	42.766	172.034	63.307	32.516	175.592	4.510	
ILE 81	7.629	126.012	59.543	37.779	175.157	57.573	42.640	172.036	4.126	
TRP 82	8.418	128.437	56.966	29.384	176.959	59.495	37.782	175.159	4.263	
SER 83	9.117	121.194	60.631	62.773		56.530	29.409	176.941	4.336	
GLN 84	7.585	118.837	54.732	31.037	175.204	60.690	62.637	174.596	4.511	
LEU 85	8.815	126.849	53.404	40.910	174.758	54.446	30.946	175.204	3.936	

PRO 86										
THR 87	8.391	110.996		69.097	173.900	62.124	31.283	176.790	4.113	
GLU 88	7.977	120.332	54.753	33.117	174.471	61.608	69.081	173.915	2.831	
PHE 89	9.204	124.328	57.645	40.508	175.351	54.673	33.105	174.449	5.480	
THR 90	7.711	119.719	60.778	72.375	172.665	57.543	40.361	175.406	4.903	
ILE 91	8.465	121.736	61.548	38.809	175.227	61.548	72.361	172.646	4.446	
ARG 92	8.808	134.446	56.794	32.418	179.993	61.314	38.592	175.227	4.443	

Table S2. Chemical shift values for HypF-N under toxic condition as measured by NMR.

Grey boxes without data are unassigned or proline residues.

Arg92 is has a shifted ^{15}NH position since the spectral widths were altered

Residue	H	NH	Cα	Cβ	CO	Cα-i	Cβ-i	Hα (1)	Hα (2)
GLY 1									
SER 2									
ALA 3	8.425	126.277	52.511	19.115	174.084	58.112	63.980	4.366	
LYS 4	8.286	120.622	56.344	32.996	177.550	52.510	19.116	4.287	
ASN 5	8.416	120.061	53.180	38.661	176.168	56.368	32.971	4.809	
THR 6	8.142	114.378	62.166	69.596	175.544	53.167	38.674	4.436	
SER 7	8.351	116.747	63.795	58.469	174.602	62.187	69.606	4.655	
CYS 8	8.105	121.016	59.302	27.906	173.397	63.765	58.616	4.709	
GLY 9	8.646	113.870	45.042		174.136	59.304	27.892	5.068	
VAL 10	8.739	119.241	60.451	35.650	173.460	45.019		5.064	
GLN 11	9.248	125.587	53.943	31.930	171.752	60.409	35.681	5.165	
LEU 12	9.791	125.565	52.242	44.022	173.770	53.951	31.896	5.589	
ARG 13	9.053	123.613	55.588	30.999	175.689	52.227	44.028	5.062	
ILE 14	9.454	128.263	61.117	39.456	175.556	55.568	31.045	4.991	
ARG 15	8.637	125.554	54.092	33.232	174.406	61.164	39.214	4.988	
GLY 16	8.864	109.507	45.296		175.512	54.131	33.183	4.617	
LYS 17	9.000	128.558	54.678	31.771	172.695	45.315		4.782	
VAL 18	7.544	112.788	59.815	32.242	176.453	54.861	31.952	4.026	
GLN 19									
GLY 20									
VAL 21									
GLY 22									
PHE 23									
ARG 24									
PRO 25									
PHE 26	7.218	119.455	60.112	37.863	178.070	66.369	30.684		
VAL 27	7.909	120.260	66.957	31.095	175.777	60.197	37.896		
TRP 28	8.461	118.514	61.930	29.173	176.580	66.993	31.086	3.878	
GLN 29	8.453	119.147	59.200	27.991	176.977	61.880	29.218	3.570	
LEU 30	7.807	121.397	57.560	41.598	178.664	59.109		3.736	

ALA 31	8.028	119.481	54.916	16.394	178.136	57.568	41.693	3.295	
GLN 32	7.567	115.174	57.697	27.506	179.740	16.447	54.857	3.783	
GLN 33	7.634	119.616	58.283	28.108	175.290	57.790	27.482	3.881	
LEU 34	7.511	116.447	54.100	42.173	177.658	58.287	28.170	3.319	
ASN 35	7.479	117.249	53.972	36.775	175.073	54.336	42.122	4.138	
LEU 36	7.557	114.556	52.682	42.498	173.003	36.867	53.138	4.369	
HIS 37	10.065	116.462	55.581	30.695	176.146	52.642	42.433	4.561	
GLY 38	8.671	110.900	44.965		172.670	55.598	30.760	5.774	
ASP 39	9.275	121.729	52.938	44.061	171.363	45.181		4.496	
VAL 40	7.811	113.518	58.133	34.857	173.455	52.980	44.100	5.718	
CYS 41	8.759	121.613	56.016	32.323	171.206	58.127	34.976	5.023	
ASN 42	9.156	121.858	53.411	38.525	173.157	56.017	32.381	5.294	
ASP 43	8.099	122.751	53.425	43.221	175.073	53.604	38.528	4.764	
GLY 44	8.800	115.033	46.992		175.348	53.444	43.192	3.708	3.993
ASP 45	8.538	125.176	53.960	40.606	174.852	46.988		4.762	
GLY 46	8.116	109.963	44.630		176.166	40.608		4.109	
VAL 47	8.666	121.030	60.745	36.474	173.031	45.613		5.178	
GLU 48	9.121	126.422	54.034	31.974	173.629	60.813	36.500	5.652	
VAL 49	9.489	128.698	60.491	34.951	175.533	54.033	31.994	5.190	
ARG 50	9.265	127.267	55.123	32.501	174.123	60.444	34.950	5.485	
LEU 51	9.175	119.680	53.141	47.511	175.617	55.053	32.460	5.911	
ARG 52	8.600	122.404	57.933	31.333	178.896	53.198	47.498	4.620	
GLU 53	8.024	119.998	58.087	30.049	175.071	57.810	30.816	4.310	
ASP 54	7.932	117.608	52.257	42.040	174.420	58.077	30.069	4.719	
PRO 55									
GLU 56	8.520	123.370	61.752	28.568	179.100	63.503	33.102	4.002	
THR 57	8.435	115.034	66.544	67.945	178.459	62.027	28.421	4.025	
PHE 58	8.526	122.780	61.632	38.184	175.891	66.486	67.926	3.733	
LEU 59	8.513	117.631	58.022	41.621	175.474	61.136	38.260	3.707	
VAL 60	7.563	117.734	66.694	31.657	178.278	58.005	42.019	3.715	
GLN 61	7.750	117.156	58.428	28.413	179.406	66.918	31.718	4.069	
LEU 62	8.596	122.810	58.158	41.797	177.573	58.238	28.405		
TYR 63									
GLN 64	7.938	116.529	58.396	29.127	179.608			3.979	
HIS 65	8.038	112.448	54.527	28.612	177.496	58.461		4.789	
CYS 66	6.867	121.827	58.022	27.174	171.654				
PRO 67									
PRO 68									
LEU 69	7.914	115.061	55.499	40.834	177.387			4.279	
ALA 70	7.874	121.857	50.111	22.194	176.219			4.881	
ARG 71	7.625	118.850	56.089	33.560	175.214	50.023	22.235	4.780	
ILE 72	8.934	128.485	61.096	38.208	173.983			3.862	
ASP 73	9.415	128.115	55.441	41.972	175.457	61.058	38.281	4.900	
SER 74	7.783	111.584	64.643	58.068	176.463	42.075		4.653	
VAL 75	8.277	120.052	61.699	34.927	171.915	64.657	57.993	4.959	
GLU 76	8.842	128.065	54.657	32.205	175.081	61.621	34.700	4.772	

ARG 77	8.827	123.567	54.493	33.594	174.955	54.528	32.305	5.495	
GLU 78	8.843	123.416	53.680	32.524	174.766	54.288	33.452	5.040	
PRO 79									
PHE 80	7.595	121.345	57.116	42.520	175.411	62.985	32.338	4.520	
ILE 81	7.554	125.788	59.482	37.683	171.919	57.117	42.526	4.103	
TRP 82	8.334	128.299	56.664	29.382	175.065	59.454	37.733	4.312	
SER 83	9.013	120.905	60.507	62.638	176.796	56.572	29.505	4.335	
GLN 84	7.512	118.727	54.567	30.956	174.431	56.219	62.605	4.489	
LEU 85	8.718	126.696	53.050	40.851	175.018	54.569	30.971	3.941	
PRO 86									
THR 87	8.324	110.919	61.781	68.892	176.738	62.024	31.251	4.115	
GLU 88	7.909	120.281	54.499	33.014	173.747	61.586	68.933		
PHE 89	9.123	124.360	57.550	40.205	174.355	54.485	33.038	5.484	
THR 90	7.701	119.660	60.607	72.123	175.290	57.447	40.183	4.919	
ILE 91	8.347	121.446	61.123	38.690	172.538	61.322	72.151	4.501	
ARG 92	8.697	102.427	56.603	32.372	175.002	61.178	28.645	4.467	

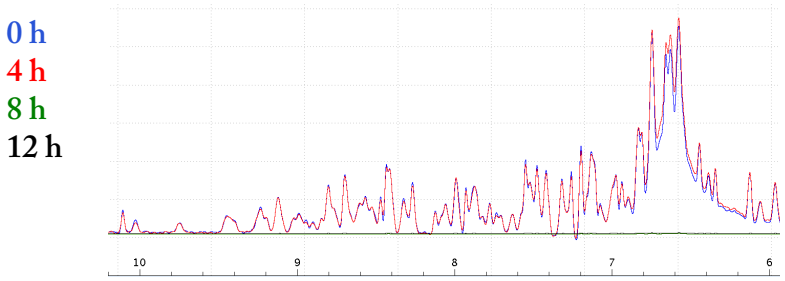
Table S3. Chemical shift values for HypF-N under urea induced unfolded condition as measured by NMR.

Grey boxes without data are unassigned or proline residues

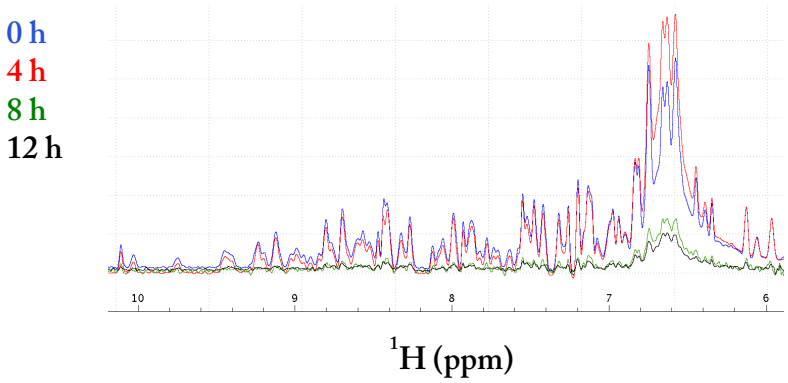
Residue	H	NH	Cα	Cβ	CO	Cα-i	Cβ-i	CO-i	Hα
GLY 1									
SER 2	8.52272	115.7155	58.278	64.155	174.138	43.460		170.147	4.473
ALA 3	8.34387	126.17346	52.566	19.193	177.591	58.115	64.185	174.163	4.318
LYS 4	8.21624	120.56804	56.417	33.161	176.251	52.576	19.265	177.614	4.249
ASN 5	8.36317	120.00057	53.308	38.785	175.496	56.317	33.163	176.263	4.745
THR 6	8.07468	114.10499	61.815	69.752	174.665	53.298	38.765	175.474	4.365
SER 7	8.25343	117.79486	58.545	63.877	174.456	61.845	69.662	174.726	4.485
CYS 8	8.26094	120.61397	58.900	28.191	174.817	58.509	63.934	174.449	4.512
GLY 9	8.30286	111.03511	45.310		173.845	58.515	28.205	174.874	3.946
VAL 10	7.87677	119.432	62.556	32.769	176.061	45.386		173.879	4.051
GLN 11	8.34912	124.37613	55.743	29.438	175.665	62.505	32.768	176.088	4.298
LEU 12	8.18794	124.36729	55.339	42.505	177.029	55.680	29.386	175.654	4.287
ARG 13	8.27948	122.50499	55.967	30.698	175.963	55.183	42.518	177.037	4.334
ILE 14	8.07619	122.60917	61.061	38.889	175.919	55.990	30.749	175.961	
ARG 15	8.3766	125.83684	56.273	30.843	176.449	61.018	38.806	176.064	4.283
GLY 16	8.23773	110.03287	44.980		173.500	56.235	30.860	176.438	
LYS 17	8.08255	120.89305	56.207	33.452	176.525	44.975		173.492	4.305
VAL 18	8.12034	121.99694	62.426	32.599	175.913	52.232	38.400	176.520	
GLN 19	8.3673	124.5839	55.883	29.759	176.026	62.340	32.712	175.924	4.305
GLY 20	8.28283	110.40481	45.151		173.889	55.742	29.761	176.101	3.930

VAL 21	7.94192	118.60351	62.484	32.755	176.598	45.170		173.886	4.091
GLY 22	8.30922	111.85024	45.036		173.482	62.478	32.738	176.607	3.858
PHE 23	7.95268	120.17232	57.736	39.606	175.145	45.055		173.450	
ARG 24	8.02069	124.9441	53.590	30.563	173.326	57.768	39.752	175.164	4.463
PRO 25									
PHE 26	8.12147	120.51328	57.842	39.449	175.753	62.967	32.017	176.264	
VAL 27	7.44387	122.57583	62.569	32.741	175.556	57.804	39.487	175.764	4.180
TRP 28	7.87115	124.17943	57.550	29.361	176.329	62.595	32.701	175.522	4.492
GLN 29									
LEU 30	7.80211	122.56558	55.433	42.308	177.234	56.212	29.503	175.768	4.152
ALA 31	8.03732	123.82722	52.872	18.966	177.846	55.413	42.315	177.227	4.170
GLN 32	8.05602	118.678	56.217	29.442	176.132	52.840	19.022	177.883	4.240
GLN 33	8.18176	120.90413	55.440	28.875	175.950	56.210	29.478	176.095	4.230
LEU 34	8.1018	121.0034	55.605	38.557		55.859	28.991	175.884	4.651
ASN 35	8.3139	119.6211	53.470	38.206	175.520	55.243	38.854	174.883	4.690
LEU 36									
HIS 37	8.34085	117.69194	55.428	28.763	174.726	55.373	42.153	177.293	4.656
GLY 38	8.18711	109.12918	45.279		173.647	55.413	28.750	174.736	3.915
ASP 39	8.24848	118.97956	52.957	38.370	175.046	45.282		173.642	4.627
VAL 40	8.02415	120.19913	62.579	32.768	175.905	52.938	38.336	175.025	
CYS 41	8.30965	122.62524	58.592	28.012	174.162	62.547	32.711	175.899	4.452
ASN 42	8.4077	121.54243	53.421	38.800	174.879	58.579	28.097	174.169	4.710
ASP 43									
GLY 44	8.2660	108.7770	45.674		174.082	53.050	38.156	175.543	3.916
ASP 45	8.15533	118.34671	53.086	38.298	175.626	45.648		174.103	4.254
GLY 46	8.2636	109.02268	45.581		174.018	53.098	38.168	175.632	3.916
VAL 47	7.76384	119.01239	62.659	32.631	176.097	45.548		174.016	4.041
GLU 48	8.26248	123.72208	55.972	28.766	176.180	62.670	32.602	175.939	4.361
VAL 49	8.03406	121.95571	62.697	32.656	175.912	55.892	28.753	176.179	
ARG 50	8.21483	124.67924	56.084	30.742	176.081	62.671	32.682	175.908	4.281
LEU 51	8.12414	123.51179	55.310	42.461	177.225	56.018	30.759	176.097	4.287
ARG 52	8.23566	121.62945	56.241	30.656	176.052	55.311	42.452	177.248	
GLU 53									
ASP 54	8.42658	121.05943	51.238	38.176	173.095	55.519	29.040	175.246	4.913
PRO 55									
GLU 56	8.15536	119.179	56.324	28.507	176.237	63.560	32.045	176.893	4.705
THR 57	7.84541	114.35637	62.442	69.734	174.172	56.294	28.565	176.203	4.179
PHE 58	7.95286	121.9571	57.851	39.533	175.375	62.443	69.745	174.173	4.572
LEU 59	7.94663	123.29917	55.342	42.395	175.383	57.832	39.574	176.966	4.264
VAL 60	7.90907	120.99257	62.958	32.642	176.121	55.323	42.341	176.968	4.089
GLN 61	8.17973	123.3647	55.979	29.292	175.729	62.941	32.607	176.114	4.315
LEU 62	7.99212	123.06202	55.652	42.495	177.066	56.052	29.290	175.714	4.174
TYR 63	7.94886	119.48413	57.670	38.539	175.740	55.537	42.486	177.061	
GLN 64	8.07405	120.77419	55.986	29.550	175.540	57.591	38.535	175.743	
HIS 65	8.39593	119.02534	55.420	28.882	173.921	55.592	29.537	175.570	4.641
CYS 66	8.27393	121.64497	56.660	27.455	171.894	55.382	28.872	173.927	

PRO 67									
PRO 68									
LEU 69	8.1024	121.99694	55.371	42.404	177.146	62.997	31.934	176.801	
ALA 70	8.12754	124.76765	52.405	19.296	177.503	55.296	42.468	177.117	4.241
ARG 71	8.13055	120.52236	56.153	30.710	176.264	52.349	19.300	177.498	
ILE 72	8.0264	121.433	61.239	38.779	175.863	56.097	30.690	176.265	
ASP 73	8.41444	122.7901	53.153	38.450	174.929	61.203	38.778	175.872	4.711
SER 74	8.07672	116.37447	58.517	63.875	174.349	53.040	38.367	174.977	4.417
VAL 75	7.91945	121.0325	62.525	32.585	175.987	58.483	63.869	174.378	4.085
GLU 76									
ARG 77	8.22814	122.46848	55.940	30.790	175.812	55.534	28.838	175.637	4.587
GLU 78	8.18876	122.04137	53.661	28.312	173.994	55.973	30.832	175.853	
PRO 79									
PHE 80									
ILE 81	7.782	122.7597	60.965	38.621	175.491	57.678	39.572	175.388	4.048
TRP 82	7.93842	125.05073	57.419	29.601	176.134	60.902	38.722	175.488	4.559
SER 83	7.87053	116.88342	58.365	63.957	173.859	57.417	29.605	176.126	4.282
GLN 84	7.99478	121.58251	55.732	29.508	175.505	58.375	63.891	173.858	
LEU 85	8.04704	124.29249	53.318	41.711	175.295	55.688	29.578	175.477	4.496
PRO 86									
THR 87	7.99216	113.75463	62.140	69.803	174.622	63.221	32.042	176.983	4.191
GLU 88	8.09821	121.60967	55.925	28.766	175.343	62.118	69.719	174.603	4.276
PHE 89	8.05822	120.76452	57.571	39.617	175.540	55.876	28.720	175.349	
THR 90	7.95193	116.41076	61.862	70.071	173.828	57.636	39.641	175.536	4.281
ILE 91	8.05142	123.82722	61.600	38.997	175.853	61.911	70.006	173.836	4.121
ARG 92									



(A)



(B)

Figure S4 1D ^{1}H NMR of HypF-N in toxic oligomer buffer

Optimisation for DEST sample, HypF-N was incubated in the toxic condition (A) and 1D ^{1}H NMR collected at specified time points. A modified condition (B) maintained presence of some monomeric protein after 12 hours.

(A) 12% TFE, 5 mM NaAc, 2 mM DTT, pH 5.5 at 298 K
(B) 10% TFE, 5 mM NaAc, 2 mM DTT, pH 5.5 at 283 K

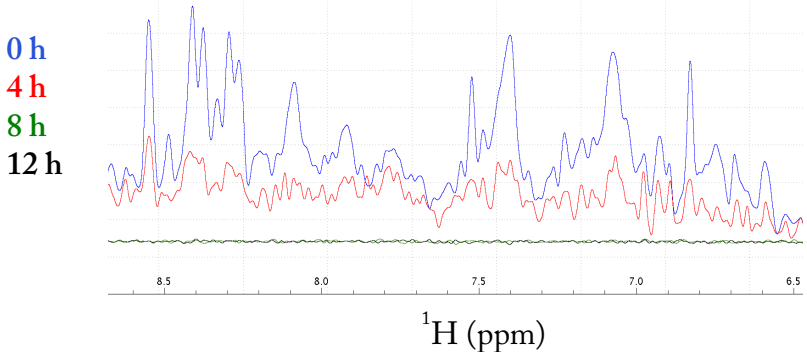


Figure S5 1D ^{1}H NMR of HypF-N in optimised non-toxic oligomer buffer

HypF-N was incubated in the optimised toxic condition 1D ^{1}H NMR collected at specified time points. We still note loss of monomeric signal after between 4 and 8 h.

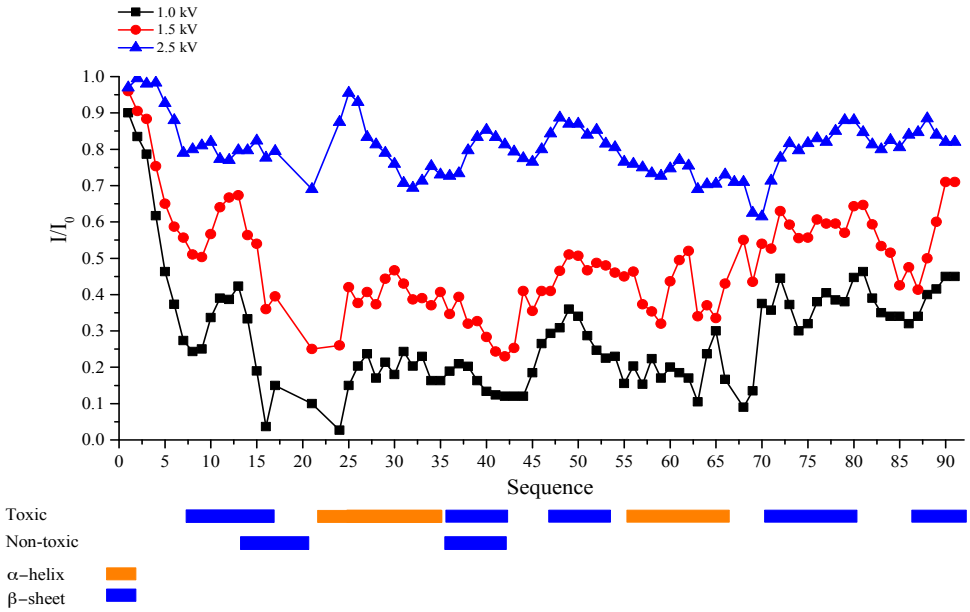


Figure S6 DEST profiles for HypF-N toxic oligomerization

DEST profile across HypF-N sequence under the optimized toxic condition. Average score for offsets 1.0, 1.5, 2.5 are shown in black, red and blue respectively. Secondary structure predictions are shown below.

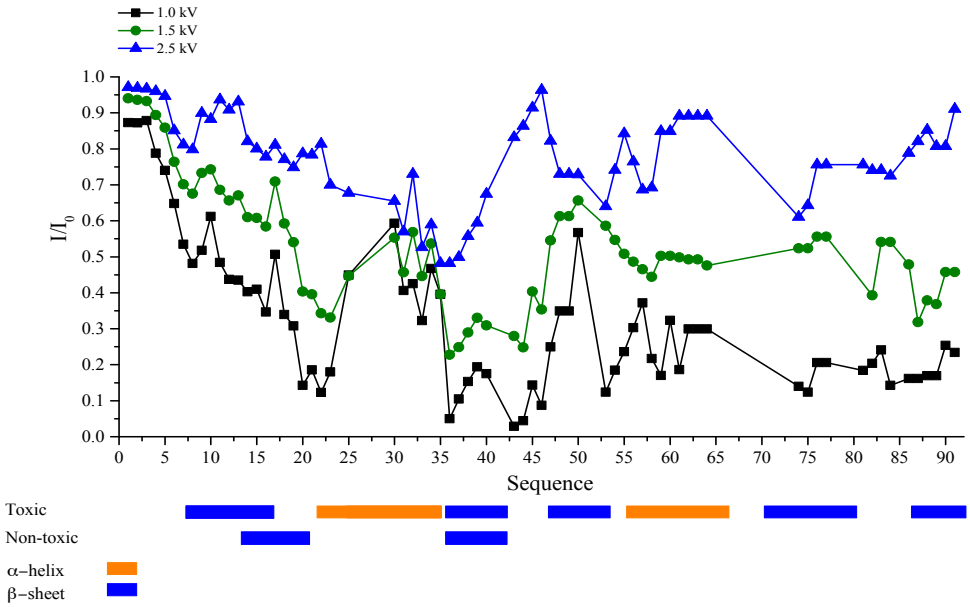
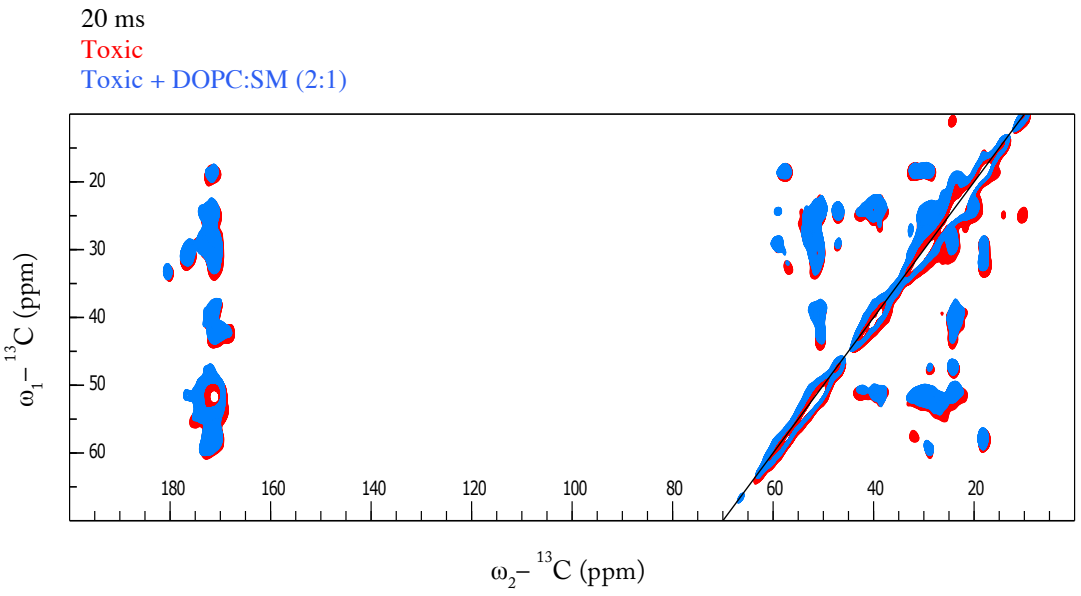
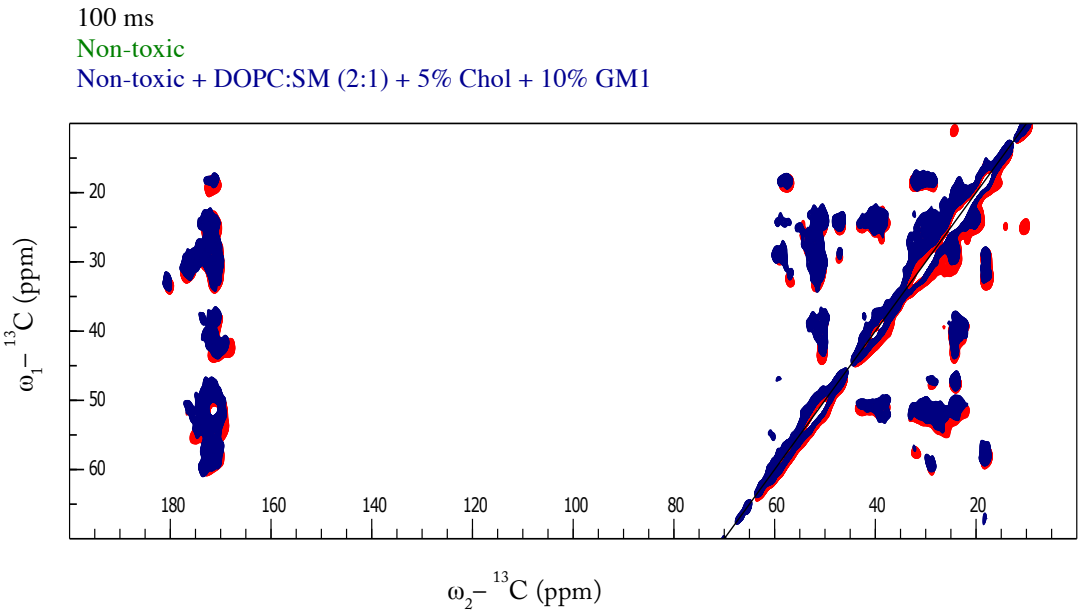


Figure S7 DEST profiles for HypF-N non-toxic oligomerization

DEST profile across HypF-N sequence under the optimized toxic condition. Average score for offsets 1.0, 1.5, 2.5 are shown in black, green and blue respectively. Secondary structure predictions are shown below.



(A)

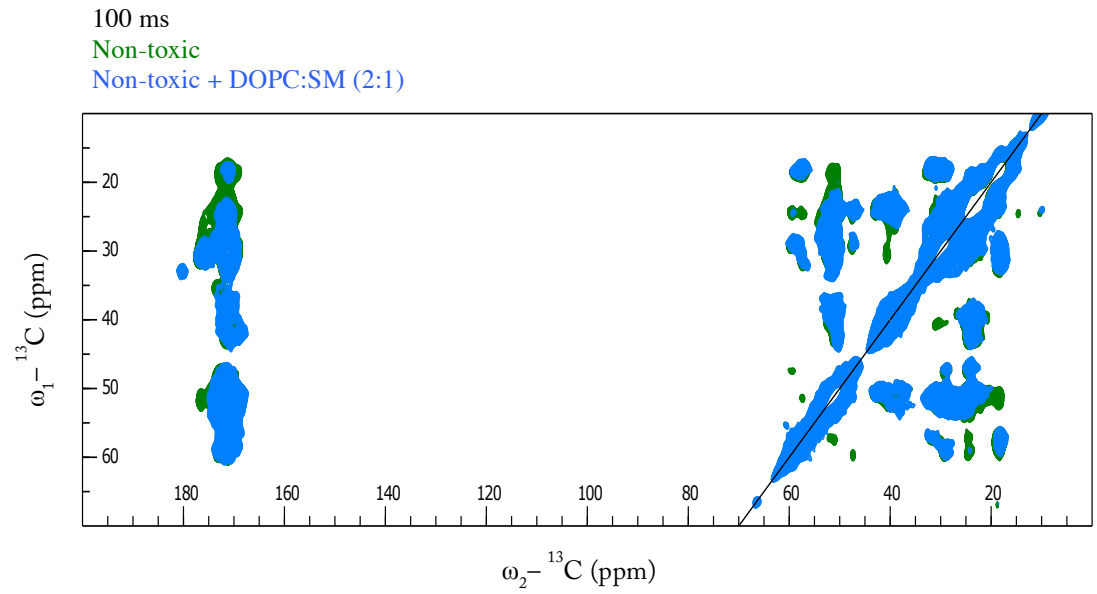


(B)

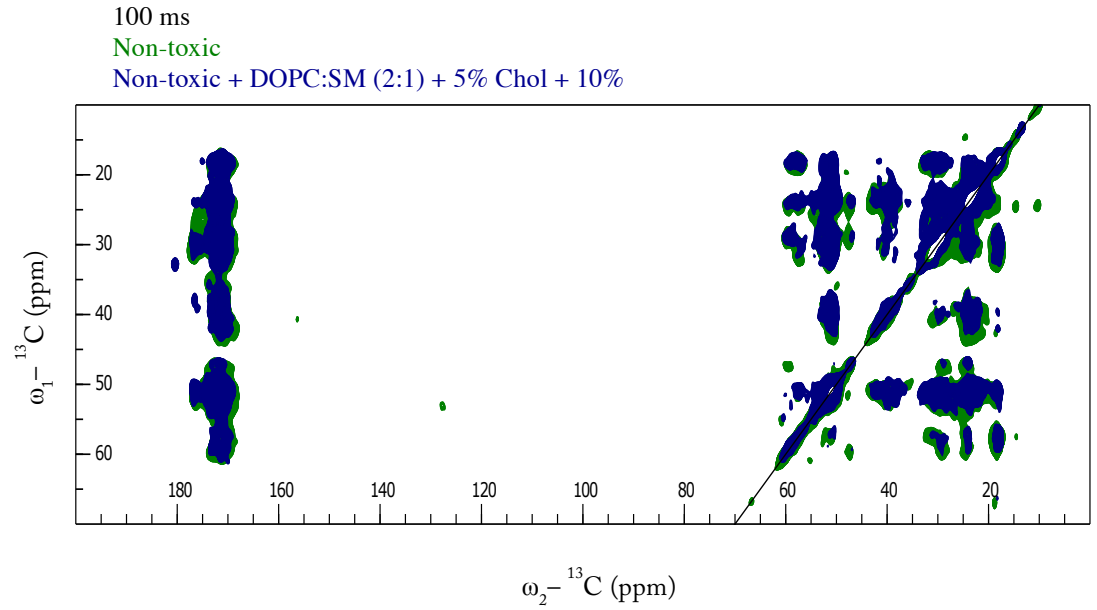
Figure S8 Core structure of toxic oligomers remain largely unchanged in presence of SUV and GM1

(A) ^{13}C-^{13}C DARR with 100 ms mixing time of toxic HypF-N oligomer in presence and absence of DOPC:SM (2:1) SUV

(B) ^{13}C-^{13}C DARR with 100 ms mixing time of toxic HypF-N oligomer in presence and absence of DOPC:SM (2:1) + 5% Chol + 10% GM1 SUV



(A)

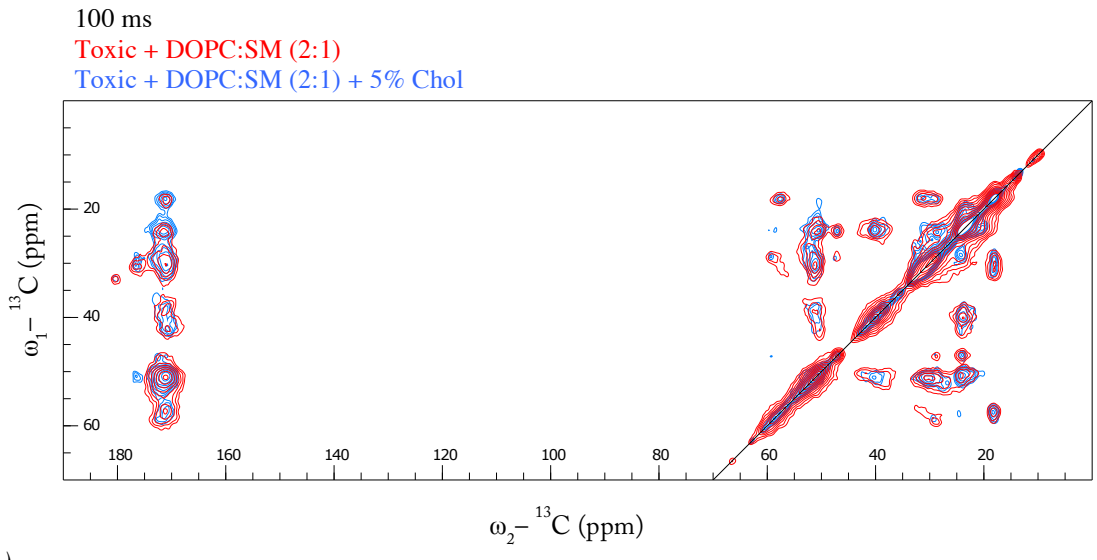


(B)

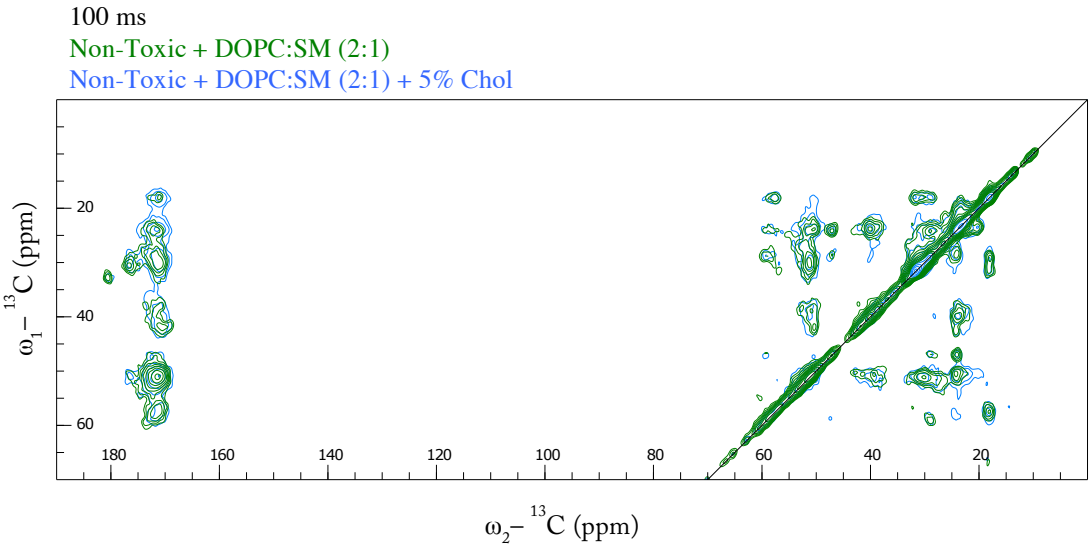
Figure S9 Core structure of non-toxic oligomers remain largely unchanged in presence of SUV and GM1

(A) ^{13}C-^{13}C DARR with 100 ms mixing time of non-toxic HypF-N oligomer in presence and absence of DOPC:SM (2:1) SUV

(B) ^{13}C-^{13}C DARR with 100 ms mixing time of non-toxic HypF-N oligomer in presence and absence of DOPC:SM (2:1) + 5% Chol + 10% GM1 SUV



(A)



(B)

Figure S10 Core structure of toxic and non-toxic oligomers remain largely unchanged upon addition of 5% cholesterol

(A) ^{13}C-^{13}C DARR with 100 ms mixing time of toxic HypF-N oligomer in DOPC:SM (2:1) SUV in absence (red) and presence (blue) of 5% cholesterol

(B) ^{13}C-^{13}C DARR with 100 ms mixing time of non-toxic HypF-N oligomer in DOPC:SM (2:1) SUV in absence (red) and presence (blue) of 5% cholesterol