

TRINITY COLLEGE, UNIVERSITY OF DUBLIN
School of Biochemistry & Immunology

This Thesis is Submitted to Trinity College, the University of Dublin
for the degree of Doctor of Philosophy in Biochemistry & Immunology



Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

**NMR SPECTROSCOPIC
CHARACTERIZATION of PROTEIN
LARGE and SMALL
Aggregates, Oligomers, and Peptides**

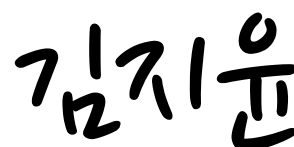
Ji Yoon Kim

2019

Republic of Ireland

Declaration

I hereby certify that this dissertation represents the sole work of the author except where stated, and has not been submitted in whole, or in part to any other university or institution, as part of a degree or other qualification. I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so my behalf, subject to the Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgment.

A handwritten signature in black ink, written in Korean characters, reading '김지윤' (Kim Ji-yoon).

Ji Yoon Kim

January 2019

to my family, especially those who are no longer with me

Acknowledgement

I would like to thank Prof Ken Mok for his support throughout this PhD. Your office door was always open to me for mentorship and guidance. You have encouraged and enhanced my ideas in science with in-depth discussion sessions. I have grown as a scientist under your watch and I am very grateful for that.

I would like to thank the wonderful Mok lab family, starting with Dr SoYoung Min, you have been a wonderful lab manager, supporting me at every juncture of my PhD. Anna, your sticky toffee pudding is the best in the world and having you in the same lab gave me the encouragement to carry on with my PhD. Pancham, thank you for all the NMR discussions we had, and for teaching me Hindi which I really enjoyed . . . Magda, thanks for all the chocolates you provided us during our lab meeting. BIG thanks go to Daniel, my science/food companion. I am deeply in debt for all your help and support. I truly enjoyed having you in our lab and for the endless salivating discussions on our favourite food items (ChiMac soon!). I would also like to thank former Mok lab member, Dr Youngjing Xia for her sound advice throughout my PhD.

Dr Manuel Ruether, my NMR God-Fairy. I wouldn't have been able to learn NMR without you. You have always been there no matter how my 'shim & lock' was up or down. Thank you so much. I also would like to thank UCD-Conway peeps, especially, Ciara and Hossam for helping me to obtain wonderful AFM images. Special thanks to Prof Andrew Bowie, Dr Andre Brodkorb, Aoife Buggy, and Prof Brian Rodriguez for the collaborations and the hard work we put together.

Dearest 5th floor TBSI peeps. If my PhD life was a book, without you guys it would have been plain black & white. You've all turned this into a wonderful colour book with all the lovely, fun, and joyful memories over the years. Nidhi & Niki, my wonderful science ladies, you girls have made my

PhD journey memorable, both inside and outside of TBSI. Katie, Emma, Aoife, Peter, and Aisling, all the lovely lunch we've had together, and the roof-top cocktail session, thank you for these lovely memories. Also, I would like to thank and welcome new members of 5th, especially James, Livia, and Dan, I feel more Italian than ever with all the Italian coffee we always had together. Roisin, I will not forget how to draw two-tone colour flowers and also all the sweat we poured out together in the gym. Natalie, Nadine, Andy, Ryan, and Paul, all the fun talk we had with a glass of wine (and more) is unforgotten. Darren, thank you for passing 'THE' chair to me. I will make sure it is passed on to 'THE' right person. Umar, one of my first TBSI friend outside of 5th floor, thank you for passing on your useful PhD life tips during all the coffee-breaks we've had.

I am in gratitude towards the Irish Research Council for providing me with a scholarship for the last 4 years. It has helped me in many ways to undertake and complete this PhD.

My Dublin Squad friends who became family to me, the precious joyful memories we made together throughout the years gave me the strength to push through with my PhD. Thank you for being there for me. Love you all.

Aunty Teresa and Uncle Tony, thank you for your support during the most intensive last month of my PhD. All the warm spicy food and countless cups of tea gave me the energy needed to function on all gears.

Ashton, who has been there for me regardless of me being happy or sad for the last 4 years and more. I still remember the time we jumped up-and-down after we found out that I received the IRC funding to pursue my PhD. You have been there for me when I have needed any kind of help. And you have always encouraged me and believed in me more than anyone. I would not have gotten here without you and your cakes (YES!). Words cannot describe how grateful I am (so, I will show you with my cooking skills). Thank you!

Huge thanks to my loving family for their unconditional support and love from Korea. My brother SungTae, your stupid jokes and funny selfies you kept sending me (even without me asking for it) improved my mood during this tumultuous time. You really are the best thing I ever asked for from our parents. Mummy2 and Daddy2, and my 2nd family for always being there for me since the day I was born. Also, mighty thanks to my four loving grandparents for their love and life wisdom.

Special thanks go out to my Éire-mum Siobhan who was there for all of my PhD years and cheered me with warm, fuzzy, real coal-lit fire and endless cups of coffee, chocolates and hugs (It always worked!).

My dearest mum and dad who have been my rock and provided me with endless support and unconditional love throughout my life. You also taught me that I have strength in me and with hard work I can overcome any life challenges and also, we can always always learn something, even from the mistakes. I could never have done this without you. Thank you for everything you've given me and done for me. This is your achievement as much as it is mine. Love you both very much. 엄마 아빠 사랑해요.

Abstract

Nuclear magnetic resonance (NMR) spectroscopy is one of the most powerful techniques for the investigation of structural, dynamic properties and interactions of biomolecules at the atomic level. In this thesis, by using biomolecular NMR spectroscopy combined with techniques such as atomic force microscopy (AFM), Thioflavin T (ThT) fluorescence, and circular dichroism (CD), we investigated biomolecules ranging from the biggest (various protein aggregates), to the smallest (an 11-amino acid residue peptide).

Chapter 3, using hen egg white lysozyme and β -lactoglobulin as model proteins, protein amyloid fibrils as well as early stage protein aggregation was characterised through further development and the application of diffusion NMR pulse sequences. ThT fluorescence was used to show the kinetics of fibrilization, CD was used to determine the changes in protein secondary structure during protein aggregation, and CR birefringence along with TEM was used for visual conformation of the amyloid fibril. With AFM, the morphology and physical characteristics of the early stage protein aggregates and of the amyloid fibril were investigated. Finally, an in-depth application of diffusion NMR was presented using protein molecules with non-uniform sizes occurring during the protein aggregation process and also with intentionally cross-linked protein with a cross-linker BS³.

Chapter 4, structural modifications of the β -Lg during thermal-denaturation at high concentration with neutral pH (pH 6.7) were investigated with ¹H-¹³C Heteronuclear Single Quantum Correction (HSQC) NMR spectroscopy combined with size exclusion HPLC (SE-HPLC). We employed current pasteurization conditions of 85 °C at high protein concentration (12 % *w/w*) at pH 6.7, and found that the β -Lg maintained its native-like structure with minor structural changes even after such heat treatment.

Lastly, in Chapter 5, the three-dimensional structure of the peptide with toll-like receptor 4 (TLR4) inhibitory property termed VIPER (Viral Inhibitory Peptide of TLR4) and its inactive mutant peptide VIPER L6AE10A were determined with various NMR spectroscopic techniques such as, ^1H - ^{15}N HSQC, ^1H - ^{13}C HSQC, Total correlated spectroscopy (TOCSY) and Nuclear Overhauser Effect Spectroscopy (NOESY). We were able to show the difference between VIPER and the mutant VIPER structure, although there were only minor side chain substitutions. Also, we compared the resulting structural ensembles of the two peptides with the 3D structure of the VIPER region in the vaccinia virus A46 protein and were able to show that as a 11 long *aa* peptide, VIPER still appears to retain helical-like structure, while the mutant peptide loses its helical-structure completely. This suggests that, the alpha-helical motif is important in binding with the TIR-domains.

These approaches provided relatively accurate estimate of the hydrodynamic volume of large biomolecules, the native-like structure of proteins even when they are assembled as aggregates, and also structural information on the important contributing residues in a peptide for TLR4 inhabitation.

List of Abbreviations

AFM	Atomic Force Microscopy
BAMLET	Bovine α -lactalbumin Made Lethal to Tumour cells
β -Lg	β -lactoglobulin
CD	Circular Dichroism
CR	Congo Red
DLS	Dynamic Light Scattering
DOSY	Diffusion Order Spectroscopy
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
ELOA	Equine Lysozyme with Oleic Acid
EM	Electron Microscopy
EMK	Ethyl Methyl Ketone
FTIR	Fourier Transform Infrared Spectroscopy
FT	Fourier Transformation
FID	Free Induction Decay
HAMLET	Human α -lactalbumin Made Lethal to Tumour cells
HEWL	Hen Egg White Lysozyme
HSQC	Heteronuclear Single Quantum Correlation
IMF	Infant Milk Formula
TLR	Toll-like Receptor
NMR	Nuclear Magnetic Resonance
nOe	Nuclear Over-Hauser Effect

NOESY	Nuclear Over-Hauser Effect Spectroscopy
PAMP	Pathogen Associated Molecular Patterns
PFG	Pulse Field Gradient
PGSE	Pulsed Gradient Spin Echo
ppm	parts per million
PRR	Pathogen Recognition Receptor
PTB	Phospho-Tyrosine-binding
RF	Radio Frequency
SSNMR	Solid-state NMR
TMS	Tetramethylsilane
ThT	Thioflavin T
TIRFM	Total Internal Reflection Fluorescence Microscopy
TOCSY	Total Correlation Spectroscopy
TSP	Trimethylsilylpropanoic acid
TEM	Transmission Electron Microscopy
VIPER	Viral Inhibitory Peptide of TRL4
SE-HPLC	Size Exclusion High-Performance Liquid Chromatography

Contents

Declaration	i
Acknowledgement	v
Abstract	ix
List of Abbreviations	xi
1 Introduction	23
1.1 Proteins	23
1.1.1 Protein folding	24
1.1.2 Protein misfolding	25
1.1.3 Alternatively folded proteins	26
1.2 Protein aggregation: amyloid fibrils & oligomers	26
1.2.1 Hen Egg White Lysozyme (HEWL)	30
1.2.2 Bovine β -lactoglobulin (β -lg)	31
1.2.3 Potential of amyloid fibrils as a nano-biomaterial	32
1.2.3.1 Branched amyloid fibril	35
1.3 Methodology for studying protein aggregates	36
1.3.1 The binding of ThT to amyloid fibrils	36
1.3.2 Protein secondary structure studies with far-UV CD	37
1.3.3 Dynamic light scattering (DLS)	38
1.4 Solution state nuclear magnetic resonance spectroscopy (NMR)	39
1.4.1 Spin properties of nuclei	39
1.4.2 Chemical shift scale (ppm) and NMR reference compounds	40
1.4.3 Biomolecular NMR spectroscopy	41
1.4.4 Defining the secondary structure with NMR	41
1.4.5 Diffusion NMR: Diffusion Ordered Spectroscopy (DOSY) NMR	42

1.4.5.1	Diffusion coefficients and molecule size	43
1.5	Objective and outcome of the research	45
2	Materials & Methods	47
2.1	Proteins	47
2.1.1	HEWL	47
2.1.2	Bovine β -lactoglobulin	47
2.1.3	Myoglobin	48
2.1.4	VIPER	48
2.1.5	Other proteins	48
2.2	Other chemicals	49
2.3	Lab materials	50
2.4	Protein aggregation: oligomerization & fibrillization	51
2.4.1	Preparation of HEWL oligomer & fibril	51
2.4.2	Preparation of cross-linked HEWL by BS ³	51
2.4.3	HEWL-BS ³ NMR sample preparation	52
2.4.4	Nano-structured films from amyloidogenic protein: HEWL	52
2.4.5	Preparation of bovine β -Lg fibrils	53
2.4.6	Preparation of apo–myoglobin (Apo–Mb) fibril	53
2.5	Amyloid fibril formation kinetics monitored by ThT assay	53
2.6	Congo Red (CR)	54
2.6.1	CR birefringence assay	54
2.6.2	CR spectroscopic assay	55
2.7	Dynamic Light Scattering (DLS)	55
2.8	Circular Dichroism (CD) spectroscopy	55
2.8.1	Far-UV CD: sample conditions & parameters	56
2.8.2	Near-UV CD: sample conditions & parameters	56
2.8.3	Estimation of secondary structure from CD data	57
2.9	SDS-PAGE denaturing gel electrophoresis	58
2.10	Transmission Electron Microscopy (TEM) imaging	59
2.11	Atomic Force Microscopy (AFM) imaging	59
2.11.1	Sample preparation on mica surface for AFM	59

2.11.2	AFM in-liquid measurement	60
2.12	Solution state NMR: diffusion NMR spectroscopy - protein aggregation	60
2.12.1	Pseudo real-time DOSY experiment to monitor the early stage of HEWL aggregation	61
2.12.1.1	Protein sample preparation	61
2.12.1.2	Pseudo real-time DOSY experiment	61
2.12.2	Hydrodynamic radii estimation from fitting diffusion NMR data	62
2.12.3	2-D solution NMR spectroscopy	64
2.12.3.1	NMR chemical shift assignment & 3-D model generation	65
2.12.3.2	Collaborative computational project for NMR (CCPN)	65
2.12.3.3	GeNMR (generate NMR structure)	66
2.12.4	NMR temperature calibration using 1,2-ethanediol	66
2.13	Bovine β -Lg heat stability experiment preparation	67
3	Characterising Protein Oligomerisation and Aggregation	69
3.1	Introduction	69
3.1.1	Investigation of protein aggregates with AFM	70
3.1.2	Biomolecular diffusion NMR	71
3.1.3	Aims	72
3.2	Result & Discussion	73
3.2.1	HEWL amyloid fibril formation	73
3.2.1.1	ThT fluorescence measurement	73
3.2.1.2	Congo red binds amyloid fibrils	74
3.2.1.3	Simultaneous measurement of aggregation with diffusion NMR (DOSY)	75
3.2.1.4	Tracing the early stage HEWL amyloid fibrilization with AFM	77
3.2.2	Nanostructured films from amyloidogenic proteins (with HEWL)	78
3.2.2.1	Kinetic measurement with ThT fluorescence assay	80
3.2.2.2	Particle size measurement using DLS	82
3.2.2.3	Casted hydrogel-film	85
3.2.3	Bovine β -Lg aggregation	86

3.2.3.1	ThT fluorescence measurement	86
3.2.3.2	Secondary structure study with CD: Far-UV	86
3.2.3.3	TEM imaging of native β -Lg protein and β -Lg fibril	89
3.2.3.4	Investigation of β -Lg fibril with AFM	91
3.2.4	Diffusion NMR of proteins with different sizes	91
3.2.5	Estimating the proportion of bovine β -Lg monomer-dimer equilibrium at low/high protein concentrations with diffusion NMR	93
3.2.6	Size estimation of cross-linked HEWL with BS ³ using diffusion NMR	96
3.2.6.1	SDS-PAGE electrophoresis	96
3.2.6.2	Diffusion NMR	97
3.2.6.3	Diffusion coefficient calculation from the decay profile with MATLAB	99
3.2.7	Pseudo real-time early stage amyloid oligomerization monitoring	100
3.2.7.1	Pseudo real-time monitoring of early stage amyloid oligomerization of HEWL with Diffusion NMR	100
3.2.7.2	Morphology of early stage amyloid oligomerization of HEWL with AFM	106
3.3	Conclusion	107
4	The Effect of β-Lg Concentration on its Properties of Heat Denaturation	111
4.1	Introduction	112
4.1.1	Protein HSQC NMR: characterisation of the bovine β -Lg stability	114
4.2	Results & Discussion	115
4.2.1	NMR temperature calibration using 1,2-ethanediol	115
4.2.2	Investigation of the effect of β -Lg concentration on denaturation and subsequent aggregation	117
4.2.2.1	Size exclusion HPLC (SE-HPLC) & reversed phase HPLC (RP-HPLC)	117
4.2.2.2	NMR: ¹ H- ¹³ C Heteronuclear Single Quantum Correlation (HSQC)	119
4.3	Conclusion	124
5	Solution Structure of 9R-VIPER and 9R-VIPER L6AE10A	127
5.1	Introduction	128
5.1.1	Toll-like Receptors (TLRs)	128

5.1.2	VACV protein A46	129
5.1.3	VIPER: inhibitory peptide derived from A46	129
5.1.4	Summary	131
5.2	Results & Discussions	132
5.2.1	Solution NMR Spectroscopy	132
5.2.2	Solution NMR based Peptide Structure Determination	139
5.3	Conclusion	146
6	Closing Summary	147
	Appendices	149
A.1	Hydrodynamic radius of biomolecule calculation by diffusion NMR	149
A.2	Estimates of hydrodynamic radius of HEWL- BS ³ oligomers obtained from SE-HPLC	152
A.3	3D structure and amino acid sequence of bovine β -Lg	153
A.4	Fiber of sickle cell hemoglobin	154
A.5	Full assigned chemical shift list of 9R-VIPER	155
A.6	Full assigned chemical shift list of 9R-VIPER L6AE10A	159
A.7	Publications	163
	Bibliography	194

List of Figures

1.1	Hierarchy of protein structure	24
1.2	Thermodynamics of protein folding process	25
1.3	Protein misfolding and subsequent conformations	27
1.4	Kinetics of nucleation-dependent amyloid fibril formation	28
1.5	Representation of 3-D structure of HEWL	31
1.6	Representation of NMR 3-D structure of Bovine β -lactoglobulin	33
1.7	Protein/peptide nanostructure-based applications	34
1.8	AFM image of branched amyloid fibril	36
1.9	The chemical structure of Thioflavin T	37
1.10	Characteristic of Far-UV CD curves of secondary structure elements	38
1.11	Typical example of Pseudo-2D-DOSY presentation	43
1.12	Diffusion decay profiles	44
3.1	Operating principle of AFM and AFM image of amyloid fibrils	71
3.2	ThT fluorescence spectra of HEWL protein	74
3.3	CR binds to HEWL at different states of fibrillisation	76
3.4	Different areas of the overlaid 1D NMR spectrum of the HEWL protein	77
3.5	DOSY spectra of different stages of HEWL aggregation during 21 days of incubation	78
3.6	AFM images obtained from 1 <i>mM</i> HEWL (pH 2) incubated at 62 °C for one, two, and three full days	79
3.7	AFM result of HEWL 1 <i>mM</i> (pH 2) incubated at 62 °C for six days	80
3.8	Structural analysis of HEWL after 6 days of incubation with AFM	81
3.9	ThT fluorescence of HEWL without and with seeding solution	82
3.10	Size distribution results of the pre-seeding stage from DLS measurement	83
3.11	The size distribution results of the post-seeding stage from DLS	84

3.12 HEWL hydrogel film with various concentrations of glycerol	86
3.13 Optical microscopy images of HEWL amyloid fibril in the hydrogel film	87
3.14 ThT fluorescence spectra of the β -Lg protein	88
3.15 Far-UV CD spectra of β -Lg protein	88
3.16 TEM images of Native β -Lg protein (A) and 6-hour incubated β -Lg (B,C)	90
3.17 Structural analysis of β -Lg fibrils with AFM	92
3.18 2D-DOSY spectra	94
3.19 Diffusion coefficient values of β -Lg and obtained hydrodynamic radius	96
3.21 Experimental 1D ^1H diffusion traces of HEWL-BS ³ 0.5 <i>mM</i> mixture sample	98
3.22 2D-DOSY spectrum of HEWL-BS ³ 0.5 <i>mM</i> cross-linked sample	99
3.23 Relative integrated spectral intensity as a function of gradient strength for HEWL-BS ³	101
3.24 Overlaid plot of normalised integrated intensity as a function of the gradient strength	102
3.25 Linear formatted diffusion NMR data of the native and 24 hours incubated HEWL	103
3.26 The normalised integrated intensity as a function of the gradient strength of HEWL (6,12,18,24hr) (Mono-Gaussian)	104
3.27 The normalised integrated intensity as a function of the gradient strength of HEWL (6,12,18,24hr) (Poly-Gaussian)	105
3.28 AFM result of HEWL 1 <i>mM</i> (pH 2) 6 hour and 24 hour incubated at 62 °C	107
4.1 ^1H - ^{15}N HSQC spectra of HEWL (I59T) at pH 1.2 recorded at different temperatures	115
4.2 NMR temperature calibration using 1,2-ethanediol	117
4.3 Heat treatment on 1 % β -Lg at different temperatures	118
4.4 SE:HPLC: size exclusion chromatography distribution of 1 and 12 % (<i>w/v</i>) β -Lg	119
4.5 ^1H - ^{13}C HSQC of native vs. heat treated β -Lg	120
4.6 Overlaid ^1H - ^{13}C HSQC spectrum of native 1 % β -Lg	121
4.7 Overlaid ^1H - ^{13}C HSQC spectrum of Native 12 % β -Lg	122
4.8 One-dimension traces of ^1H - ^{13}C HSQC peaks of Native 12 % β -Lg	123
5.1 Signalling adaptors involved in TLR4 and TLR2 activation of pro-inflammatory and interferon related genes	130
5.2 Superposition of ^1H - ^{15}N HSQC spectra of 9R-VIPER and 9R-VIPER L6AE10A	134
5.3 NOESY contour spectrum (NH-H α region) of the 9R-VIPER	135

5.4	NOESY contour spectrum (NH-H α region) of the 9R-VIPER L6AE10A	136
5.5	Sequential and NOE connectivities, and secondary chemical shifts of 9R-VIPER and 9R-VIPER L6AE10A	138
5.6	The 3-D structure of 9R-VIPER and 9R-VIPER L6AE10A	142
5.7	NMR-based structure ensemble of 9R-VIPER and 9R-VIPER L6AE10A	143
5.8	Superimposed models of the native VIPER in A46(75-227)	144
5.9	Composition of the 3D structure of VIPER region	145

List of Tables

2.1	Properties of proteins and tryptophan	50
2.2	HEWL-BS ³ mixture sample preparation	52
2.3	SDS-PAGE gel sample preparation ratio	58
2.4	Increase in gradient strength	61
3.1	Estimated hydrodynamic radius and the percentage proportion of HEWL-BS ³ oligomers via SE-HPLC and SDS-PAGE	100
3.2	Calculated hydrodynamic radius and percentage proportion of HEWL-BS ³ oligomers by poly-Gaussian decay function with three components	100
3.3	Hydrodynamic radius of the native and incubated (6,12,18, and 24 hours	104
3.4	Estimated hydrodynamic radius and ratio of the two components in 6,12,18, and 24 hours incubated HEWL	106
4.1	1,2-ethanediol calibrated temperature vs. NMR control panel temperature	116
5.1	¹ H- ¹³ C and ¹ H- ¹⁵ N chemical shift of 9R-VIPER and 9R-VIPER L6AE10A	133
5.2	NMR restraints table of 9RVIPER and 9R VIPER L6AE10A	137
5.3	Secondary structure prediction of 9R-VIPER and 9R-VIPER L6AE10A	140

Chapter 1

Introduction

1.1 Proteins

Proteins are one of the most fundamental biomolecules and are a crucial part of biological processes that include catalysis, signal transmission, and structural support. Proteins are linear, unbranched chains of amino acid monomers with a unique three-dimensional structure that determines its specific function. Amino acids are the building blocks of proteins and play a central role as intermediates in metabolism. Most natural polypeptide chains contain between 50–2000 amino acid residues and are commonly referred to as proteins, whilst polypeptide chains that contain less than 50 amino acids are referred to as peptides. Each protein has a unique amino sequence specified by a gene that determines its function and is directly dependent on its three-dimensional structure. Therefore, determining the structure is essential in understanding the functions and characteristics of proteins and peptides.^[1,2] Nuclear Magnetic Resonance (NMR) spectroscopy, X-ray crystallography, and electron microscopy are the principal techniques for determining three-dimensional structures of biomolecules.^[3–5]

¹ J. M. Berg, J. L. Tymoczko, and L. Stryer. *Biochemistry*. Basingstoke: W.H. Freeman, 2012.

² D. L. Nelson, M. M. Cox, and A. L. Lehninger. *Lehninger Principles of Biochemistry*. 7th ed. Houndmills, Basingstoke: Macmillan Higher Education, 2017.

³ M. P. Williamson. Peptide Structure Determination by NMR in: *Spectroscopic Methods and Analyses*. ed. by J. Jones, B. Mulloy, and A. Thomas. 17: 69–85, New Jersey: Humana Press, 1993. DOI: [10.1385/0-89603-215-9:69](https://doi.org/10.1385/0-89603-215-9:69)

⁴ K. Wüthrich. *Nat. Struct. Biol.*, 8: 923–925, 2001.

⁵ D. G. Reid. Protein NMR Techniques in: *Methods in Molecular Biology*. 60: New Jersey: Humana Press, 1997. DOI: [10.1385/0896033090](https://doi.org/10.1385/0896033090)

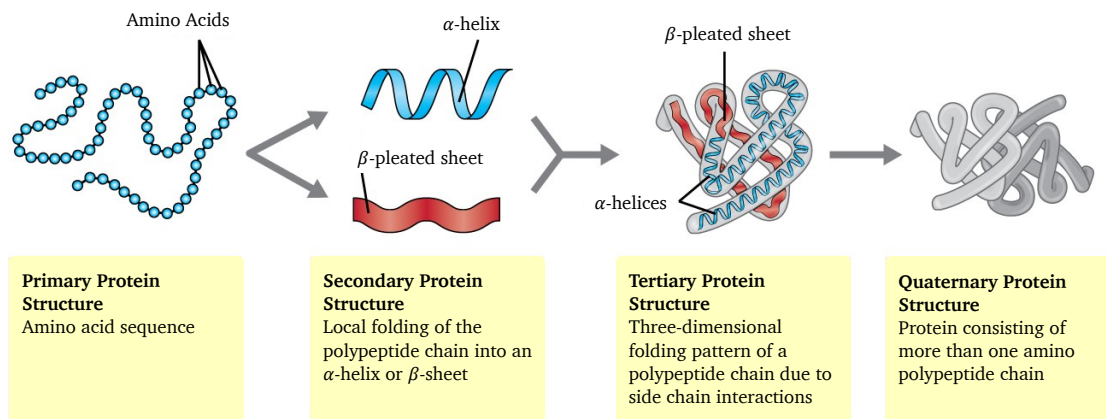


Figure 1.1: Hierarchy of the four levels of structure in a protein where the primary structure consists of a sequence of amino acids linked together by peptide bonds and includes any disulfide bonds. The resulting coiled polypeptides are units of secondary structure, such as the α -helix and the β -sheet. The formed secondary structure is a part of the tertiary structure of the folded polypeptide, which in itself can be one of the subunits of the quaternary structure of the multi-sub-unit protein, such as haemoglobin.^[2] / (Art: Modification of work by National Human Genome Research Institute)

1.1.1 Protein folding

Protein structure can be described in four hierarchical levels as shown in Figure 1.1. The primary structure refers to the amino acid residue sequences where amino acids are linked by peptide bonds to form polypeptide chains. A peptide bond is the link between the α -carboxyl group of one amino acid to the α -amino group of another amino acid. The secondary structure refers to the local conformation of the polypeptide chain. Certain amino acid sequences fold into regular structures such as the α helix, β sheet, turns, or loops. The α helix is a coiled structure stabilised by intra-chain hydrogen bonds whereas β sheets are stabilised by hydrogen bonding between polypeptide strands. Turns and loops are responsible for changing the direction of polypeptide chains. The tertiary structure represents the overall folding of the polypeptide chain and finally, the quaternary structure describes the specific association of multiple polypeptide chains to form a multi-subunit complex.^[1,2]

As mentioned earlier, protein structure correlates to a specific function, hence, loss of protein structure results in loss of function, also known as protein denaturation. This emphasises the importance of protein folding, however the protein folding mechanism is still not fully understood. Figure 1.2 shows the free energy funnel model of the protein folding which describes the protein

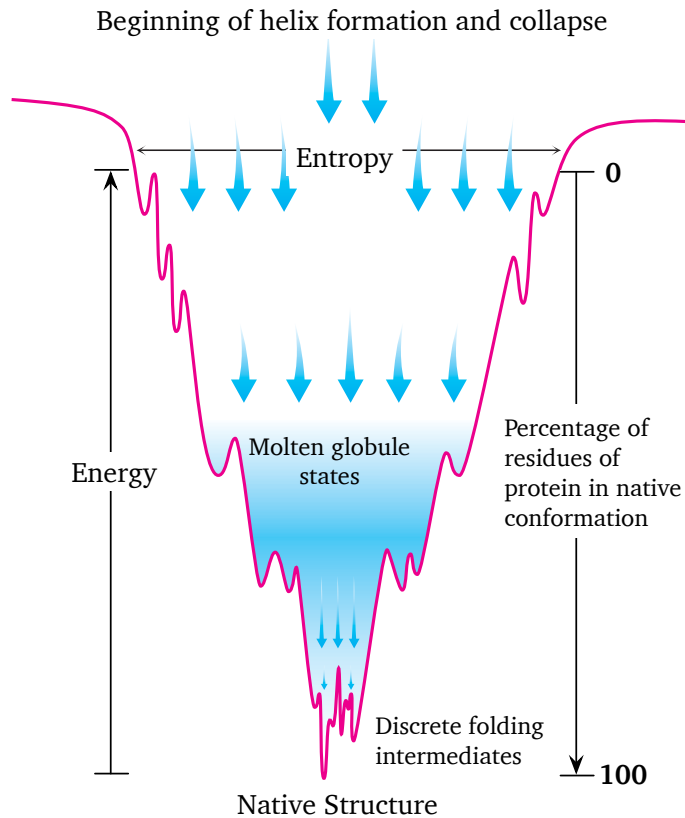


Figure 1.2: The thermodynamics of protein folding process presents as a free-energy funnel landscape. At the top, there are multiple accessible unfolded states and as folding progresses, the number of accessible conformations reduces while the free-energy also decreases which presented as funnel gets narrower. At the bottom of the funnel, native structure with well-folded conformation is achieved.^[1,2]

folding process in terms of a funnel shaped free-energy landscape. This funnel model suggests that instead of a fixed stepwise process from unfolded to well-folded native protein, there are multiple assessible conformations for unfolded protein and as the free energy of protein molecules decreases, the protein move down into narrower parts of the funnel with fewer accessible conformations. Finally, at the bottom of the funnel, accessible conformation reduces to a single native conformation, where energy is minimum.^[1,2]

1.1.2 Protein misfolding

As protein folding is an enormously complex and stochastic process, it would be remarkable if there were no errors occurring during the folding progression. Misfolded proteins are those that have adapted incomplete or incorrect overall protein structure due to various factors such as genetic

mutations, physical or chemical stress, temperature, pH, and exogenous compounds not usually found in the cytosol. As shown in Figure 1.3, depending on the circumstances, proteins go through different misfolding pathways and give rise to different conformations characterised by either a disordered aggregated form or an ordered aggregated form, commonly referred to as an amyloid fibril.^[6,7] Protein structure is closely related to the function of protein, hence, protein misfolding and protein aggregation are associated with some neurological diseases or so-called “conformational diseases” including Alzheimer’s disease, Parkinson’s disease, Huntington’s disease, and transmissible spongiform encephalopathies called prion disease *in vivo*.^[1,2,8]

1.1.3 Alternatively folded proteins

As previously mentioned, misfolded protein is recognised as a negative incident and is usually related with neurological diseases. However, Human α -lactalbumin Made Lethal to Tumour cells (HAMLET) and its related partially unfolded protein-fatty complexes have been shown to be effective against over 40 tumour cell lines, while leaving healthy, differentiated cells intact. HAMLET is a novel biomolecular nanoparticle that consists of partially folded α -lactalbumin bound to oleic acid. It is one of the few proteins that exhibits a well-defined function in its native state and is also capable of attaining a new and beneficial function after partial unfolding. There are other complexes similar to HAMLET, which consist of partially folded proteins and fatty acid that exhibit cytotoxic activities against cancer cell lines after binding. Examples of these complexes include Bovine α -lactalbumin Made Lethal to Tumour cells (BAMLET), Equine Lysozyme with Oleic Acid (ELOA) and other oleic acid complexes with β -lactoglobulin or α -lactalbumin.^[10]

1.2 Protein aggregation: amyloid fibrils & oligomers

Protein aggregation *in vivo* has been extensively associated with various human diseases. Amyloid fibrils are ordered aggregates of peptides or protein that are fibrillated into a β -sheet rich structure. The formation of amyloid fibrils is a multi-step process which begins with a native protein and ends

⁶ C. M. Dobson. *Nature*, **426**: 884–890, 2003.

⁷ C. M. Dobson. *Semin. Cell Dev. Biol.*, **15**: 3–16, 2004.

⁸ E. H. Koo, P. T. Lansbury, and J. W. Kelly. *Proc. Natl. Acad. Sci.*, **96**: 9989–9990, 1999.

¹⁰ S. Y. Min et al. *Biochem. Soc. Trans.*, **40**: 746–751, 2012.

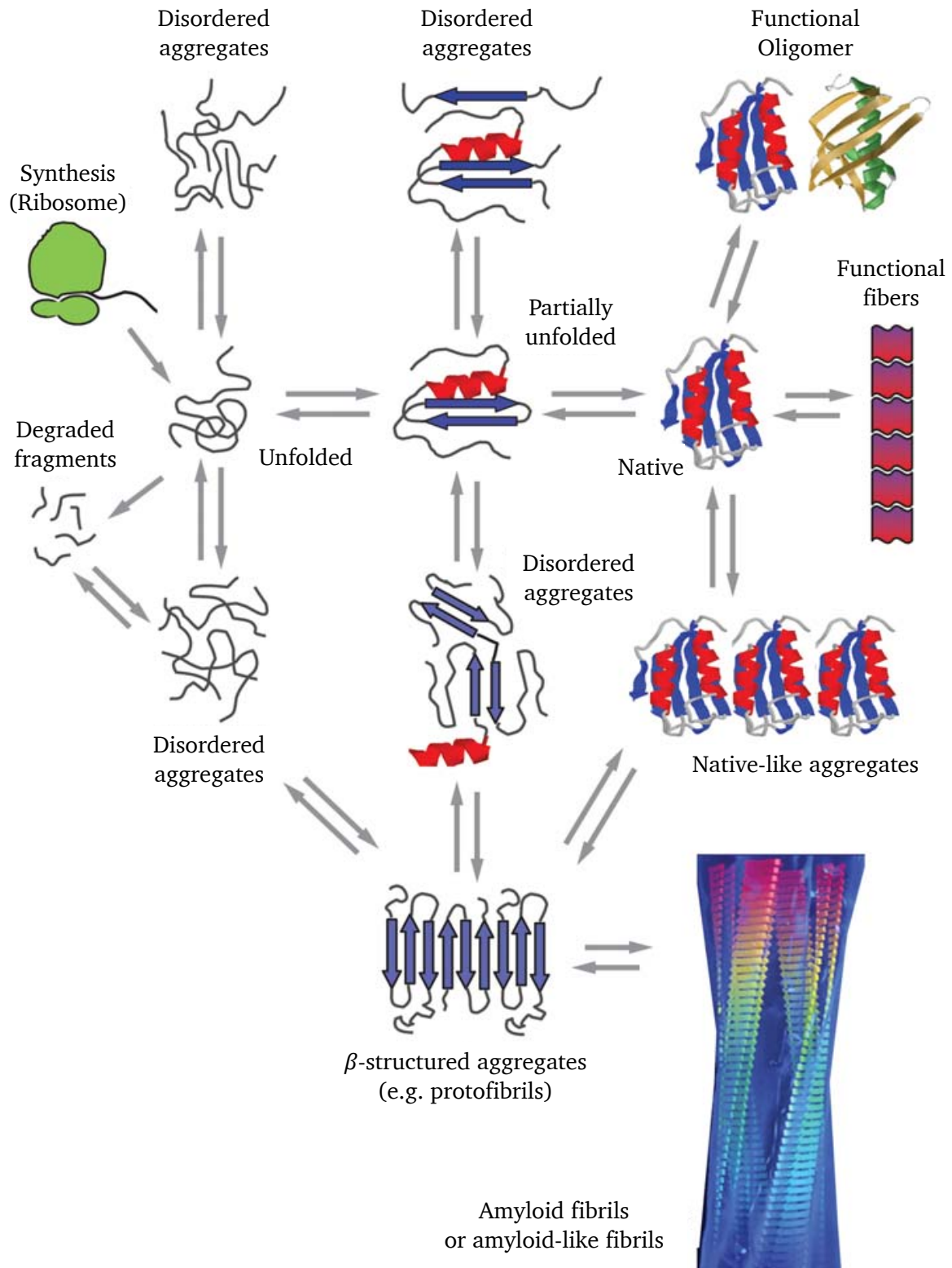


Figure 1.3: Different misfolding pathways give rise to different conformations. Aggregations exhibit different degrees of organisation, going from unstructured, disordered aggregated clumps to organised oligomers and fibrils. Amyloid fibrils are just one of the types of aggregated conformation that can occur due to misfolded proteins.^[6] Image adapted from .^[9]

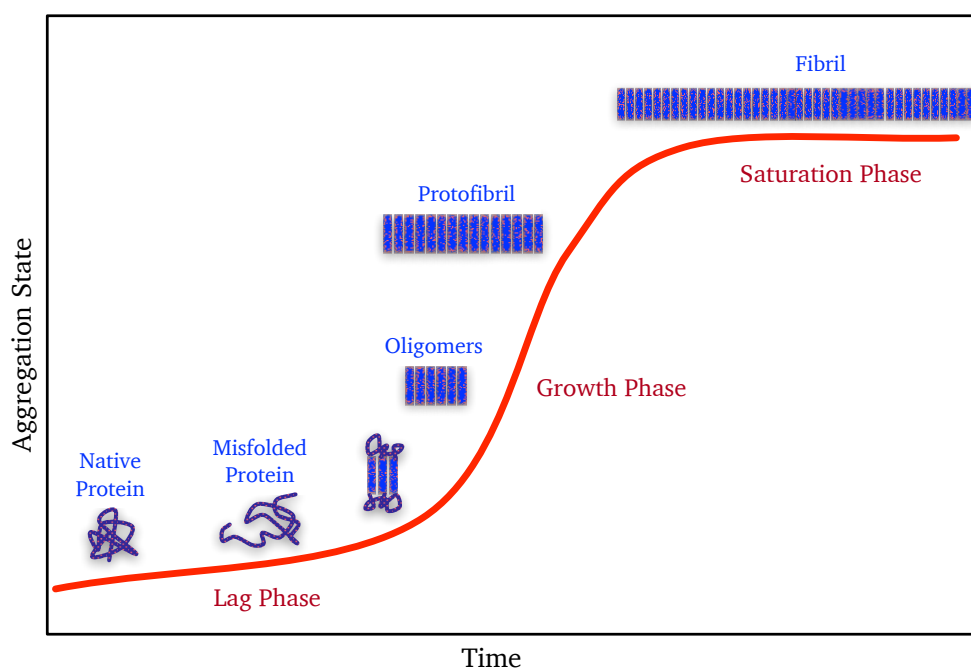


Figure 1.4: A variety of conditions such as pH, temperature, ionic strength and macromolecular crowding regulate the fibrillisation process.^[11]

with insoluble amyloid fibrils with many intermediate developmental stages such as the formation of monomers, oligomers, protofibrils, and fully mature fibrils (Figure 1.4). When hydrophobic residues (usually tightly packed within the core of the native protein) become exposed patches, protein fibrillisation can occur; the exposed patches will tend to aggregate in order to minimise interactions with the solvent. This aggregation leads to the formation of multimers (dimer, tetramer), oligomers and finally mature amyloid fibrils and usually follows nucleated polymerisation kinetics. A lag phase of the conversion of a peptide or protein into its fibrillar form is assumed to be the time required for nuclei to form and once a nucleus is formed, a rapid exponential growth phase follows by self-association or through the addition of either monomers or oligomers with the nucleus (Figure 1.4).^[9]

Amyloid fibrils have long been considered to be associated with neurodegenerative diseases ever since they were discovered. However, their true cytotoxic role is still in debate and each intermediate stage that occurs during the fibrillisation process has been investigated in order to identify and characterise the actual biological toxic species involved in those diseases.^[12] Furthermore, the

⁹ F. Chiti and C. M. Dobson. *Annu. Rev. Biochem.*, **75**: 333–366, 2006.

¹² R. M. Murphy. *Biochim. Biophys. Acta - Biomembr.*, **1768**: 1923–1934, 2007.

mature amyloid fibrils were originally thought to be the key cytotoxic component involved in the pathology of these degenerative disease. However, recent evidence suggests that the more soluble, intermediate oligomeric species may play more of a central cytotoxic role in the pathology of degenerative disease compared to the mature amyloid fibrils.^[9] A study by Kaye *et al.*^[13] showed that a specific antibody can bind to oligomer/protofibrillar species from different sources, but not to their corresponding monomeric or fibrillar form, which suggests that soluble amyloid oligomers do have some important role. Although soluble oligomers are considered an important cytotoxic component, mature fibrils also show evidence as cytotoxic agents and a recent study demonstrated that oligomers and fibrils display different cytotoxic capabilities with unlike toxic mechanisms.^[14]

Each species involved in the amyloid fibril formation have different characteristics: Monomers are defined as single protein or peptide chains and, are considered to be relatively non-toxic elements in the pathology of degenerative diseases. However, they are the first stage in the formation of mature amyloid fibrils by developing into oligomeric species. Oligomers are soluble aggregates which are formed during the amyloid growth process and appear as small spherical beads or a beaded chain. The sizes of different oligomeric species vary depending on the size of the original protein. Oligomers have a significantly lower rate of growth compared to amyloid fibrils, which suggest that they have different structure. It is a challenge to define the oligomers and smaller aggregates due to their extremely short-time of existence and difficulty to observe.^[15] Fibrils are insoluble, linear aggregates with a repetitive cross- β structure, and are generally unbranching structures around 10–15 nm in diameter.^[16] Although there are rare cases where branched amyloid fibril also can also occur.^[17,18]

To date, more than 20 proteins including A β peptides, α -synuclein, amyloid ANF, and PrP proteins have been identified to form fibrils and these amyloidogenic structures all appear to share a common β -sheet conformation of their polypeptide backbone.^[19,20] Recent evidence suggests the ability to aggregate and form fibrils is a generic property of possibly all polypeptide chains

¹³R. Kaye. *Science*, **300**: 486–489, 2003.

¹⁴N. P. Harte *et al.* *J. Biol. Chem.*, **290**: 28343–28352, 2015.

¹⁵P. Arosio, T. P. J. Knowles, and S. Linse. *Phys. Chem. Chem. Phys.*, **17**: 7606–7618, 2015.

¹⁶J. S. Pedersen, C. B. Andersen, and D. E. Otzen. *FEBS J.*, **277**: 4591–4601, 2010.

¹⁷I. V. Baskakov. *FEBS J.*, **274**: 3756–3765, 2007.

¹⁸C. B. Andersen *et al.* *Biophys. J.*, **96**: 1529–1536, 2009.

¹⁹M. Nilsson. *Methods*, **34**: 151–160, 2004.

²⁰R. N. Rambaran and L. C. Serpell. *Prion*, **2**: 112–117, 2008.

under specific denaturing conditions. Innocuous proteins that do not relate to neurological diseases have also been shown to assemble fibrils *in vitro* with the most notable one showing the fibril assembly with an all α -helical protein apo-myoglobin.^[21] HEWL has also been shown to form amyloid fibrils at acidic pH and elevated temperatures.^[22] These studies suggest that non-disease related proteins/peptides or α -helical protein/peptide (in native state) are capable of forming fibrils under specific denaturing conditions. Moreover, recent study carried out in our lab showed that both oligomer and fibrils of Hen Egg White Lysozyme (HEWL) which is not associated with degenerative diseases, to be toxic to cells.^[14] As a consequence, in-depth understanding of the biochemistry and biophysics properties of the protein aggregation processes will be important in understanding and unravelling fundamental mechanisms of degenerative disease as well as for developing therapeutic strategies.

1.2.1 Hen Egg White Lysozyme (HEWL)

HEWL is one of the most extensively studied proteins. Human lysozyme was discovered by Fleming in 1922 and egg white was used as a rich source of lysozyme outside of human tissues. HEWL is a single polypeptide chain containing 129 amino acids with 14.3 *kDa* of molecular weight and it shares 60 % sequence identity with human lysozyme (Figure 1.5). The primary structure of HEWL was elucidated in 1963 and the structure was determined by David Philips in 1965.^[23] The three-dimensional structure and folding of HEWL has been studied ever since^[24] and it is the model protein for protein structure and function studies.^[23,25,26] As mentioned in Section 1.2, recent studies have shown that possibly all proteins are able to form fibrils under specific denaturing conditions. HEWL forms amyloid fibrils under various conditions such as acidic pH with elevated temperature^[22] making it a popular protein model for investigations on the protein aggregations and amyloid research.

HEWL aggregation, unfolding, and fibrillisation has been intensively studied by various methods

²¹ M. Fändrich, M. A. Fletcher, and C. M. Dobson. *Nature*, **410**: 165–166, 2001.

²² M. R. H Krebs et al. *J. Mol. Biol.*, **300**: 541–549, 2000.

²³ R. Swaminathan et al. Lysozyme: A model protein for amyloid research in: *Advances in Protein Chemistry and Structural Biology*. ed. by R. Donev. **84**: 63–111, Elsevier, 2011. DOI: [10.1016/B978-0-12-386483-3.00003-3](https://doi.org/10.1016/B978-0-12-386483-3.00003-3)

²⁴ C. Redfield and C. M. Dobson. *Biochemistry*, **27**: 122–136, 1988.

²⁵ A. A. Meratan, A. Ghasemi, and M. Nemat-Gorgani. *J. Mol. Biol.*, **409**: 826–838, 2011.

²⁶ M. F. Rosario-Alomar et al. *J. Phys. Chem. B*, **119**: 1265–1274, 2015.

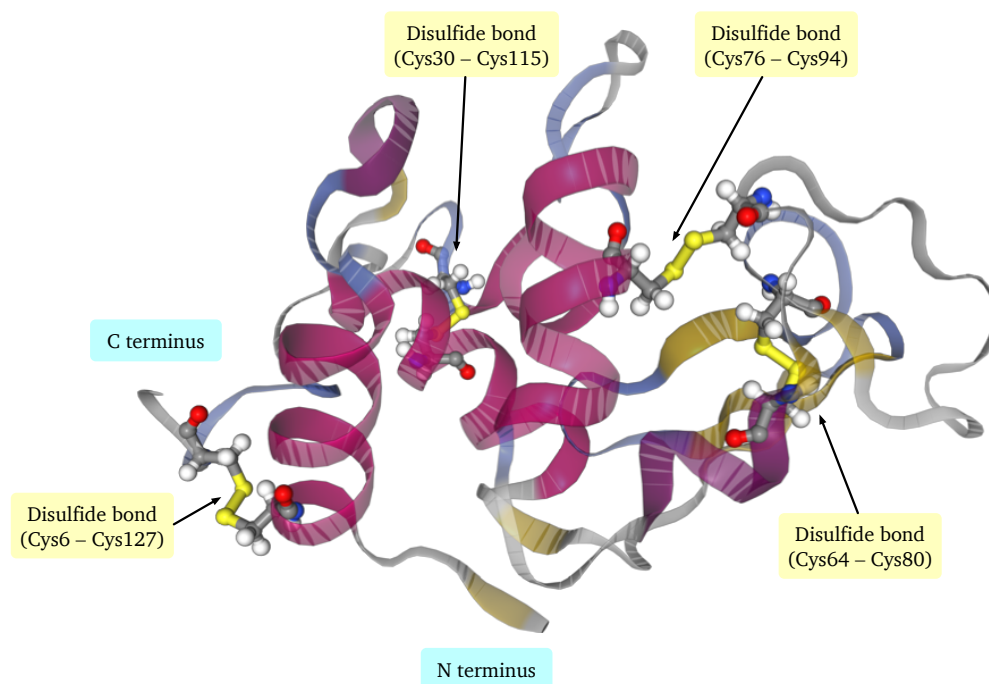


Figure 1.5: Representation of three-dimensional structure of Hen Egg White Lysozyme (HEWL), illustrated as a ribbon model. The loops are shown in white, the helices are shown in pink, the strands are shown in yellow, and the turns are shown in purple. Four disulfide bonds (Cys 6-Cys 127/Cys 30-Cys 115/Cys 64-Cys 80/Cys 76-Cys 94) are shown in ball-and-stick-model (PDB: 1GXX.^[27])

including X-ray diffraction,^[28] a two-step temperature-dependent Fourier Transform Infrared Spectroscopy (FTIR),^[29] and diffusion NMR.^[30] Morphological properties of amyloid fibrils has also been measured by using Atomic Force Microscopy (AFM) imaging.^[31] In addition to this, the effect of concentration, temperature and stirring,^[32] as well as agitation^[33] on HEWL amyloid fibril formation has been investigated.

1.2.2 Bovine β -lactoglobulin (β -lg)

Bovine β -lactoglobulin (β -Lg) is a major whey protein that offers various functional properties such as gel and foam formation when protein unfolding and denaturation occurs by heating to a

²⁸ N. Yagi, N. Ohta, and T. Matsuo. *Int. J. Biol. Macromol.*, **45**: 86–90, 2009.

²⁹ Ye Zou et al. *J. Phys. Chem. B*, **118**: 9834–9843, 2014.

³⁰ J. A. Jones et al. *J. Biomol. NMR*, **10**: 199–203, 1997.

³¹ J. F. Smith et al. *Proc. Natl. Acad. Sci.*, **103**: 15806–15811, 2006.

³² S. Y. Ow and D. E. Dunstan. *Soft Matter*, **9**: 9692, 2013.

³³ V. Sivalingam et al. *Biophys. Chem.*, **219**: 28–37, 2016.

sufficient temperature.^[34] β -Lg is a β -sheet predominant protein and consists of 162 amino acids and a molecular mass of $\sim 18,400$ Da. It has a β -barrel which is composed of eight antiparallel β -strands, one major α -helix at the C-terminal end of the molecule, and a final β -strand which forms part of the dimerization interface. At neutral pH, β -Lg resides in dimeric form, and each monomer has two disulphide bonds (Cys 66–Cys 160 and Cys 106–Cys 119) and one free cysteine (Cys 121).^[35,36] The free sulfhydryl group of Cys 121 is a reactive site, which can bind to other free reactive Cys residues, and lead to the formation of covalently bonded aggregates, hence, eliminating the possibility of reversing the denaturation process.^[37] At acidic pH, β -Lg exists in monomeric form, and the solution structure of monomeric β -Lg (Figure 1.6) was shown to be the same to that of the monomer of its dimeric form.^[38]

Due to the abundance and relative ease of purification as well as its ability to form amyloid fibrils under acidic pH with elevated temperature, Bovine β -Lg is a useful model protein for the investigation of protein aggregation and protein folding, stability and self-association studies. The aggregation of β -Lg has been studied intensively using various techniques such as, Transmission Electron Microscopy (TEM),^[39] small angle X-ray scattering, and AFM imaging.^[40] Also the effect of concentration, temperature, pH^[41,42] and urea denaturation^[43] on protein aggregation/stability has been investigated.

1.2.3 Potential of amyloid fibrils as a nano-biomaterial

Nanotechnology is an active area of research involving the study of the structure and property of substances on a nano-meter (*nm*) scale (usually between 1-100 *nm* in one to three dimensions). The aim of nanotechnology is to create unique and functional nanomaterials and nanodevices with

³⁴ N. Bansal and B. Bhandari. Functional Milk Proteins: Production and Utilization–Whey-based Ingredients in: *Advanced Dairy Chemistry*. ed. by P McSweeney and J O'Mahony. 67–98, New York, NY: Springer New York, 2016. DOI: [10.1007/978-1-4939-2800-2_3](https://doi.org/10.1007/978-1-4939-2800-2_3)

³⁵ S. Brownlow et al. *Structure*, **5**: 481–495, 1997.

³⁶ K. Sakurai and Y. Goto. *J. Mol. Biol.*, **356**: 483–496, 2006.

³⁷ T. V. Burova et al. *Protein Eng. Des. Sel.*, **11**: 1065–1073, 1998.

³⁸ K. Kuwata et al. *Protein Sci.*, **8**: 2541–2545, 2008.

³⁹ R.N. Zúñiga et al. *J. Food Sci.*, **75**: E261–E268, 2010.

⁴⁰ J. Adamcik et al. *Nat. Nanotechnol.*, **5**: 423–428, 2010.

⁴¹ A. K. Buggy et al. *Dairy Sci. Technol.*, **96**: 845–859, 2017.

⁴² A. K. Buggy et al. *J. Dairy Sci.*, **101**: 10819–10830, 2018.

⁴³ D. Hamada and C. M. Dobson. *Protein Sci.*, **11**: 2417–2426, 2009.

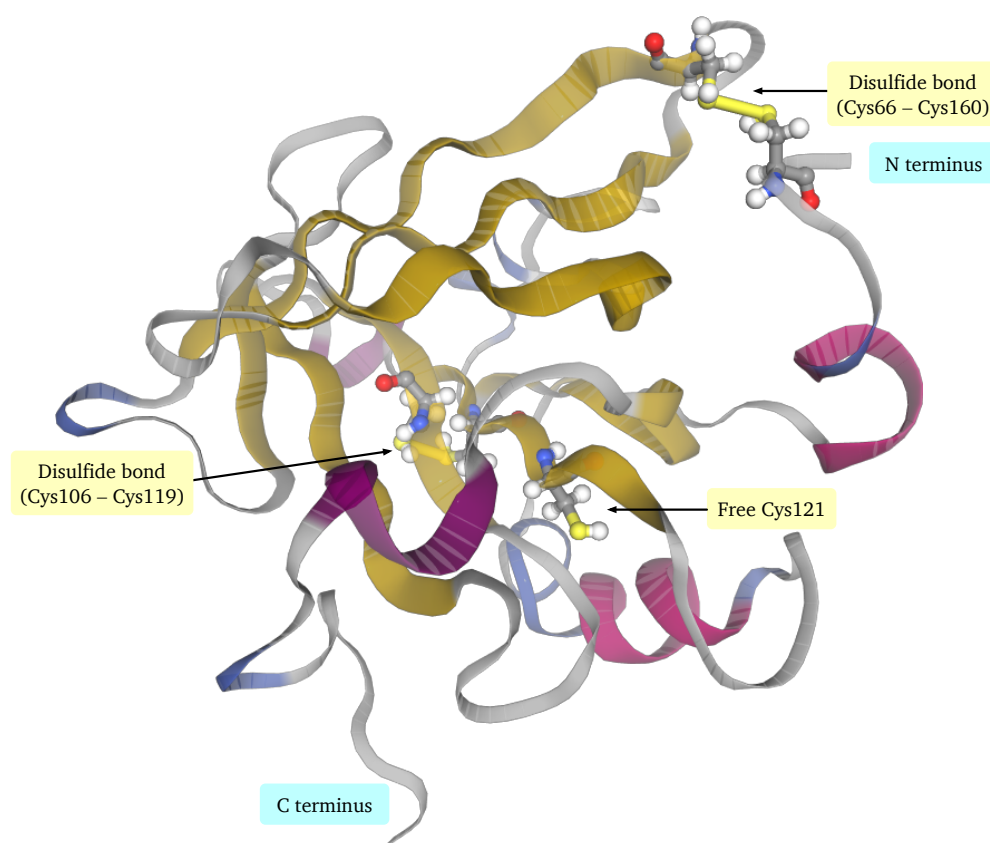


Figure 1.6: Representation of NMR three dimensional structure of Bovine β -lactoglobulin A at pH 2.0 and 45 °C. Here, the protein exists as native monomeric state, illustrated as a ribbon model. The loops are shown in white, the helices are shown in pink, the strands are shown in yellow, and the turns are shown in purple. Two disulfide bonds (Cys 106-Cys 119/Cys 66-Cys 160) and one free Cys 121 are shown in ball-and-stick-model (PDB: 1CJ5.^[38])

various nanoscale building blocks such as atoms and molecules. There are two main ways of creating the nanomaterials: Top-down and bottom-up techniques. Due to physical and technical restrictions of top-down techniques, bottom-up techniques are more and more spotlighted for nanostructure synthesis process. Biomolecules including protein, DNA, and peptides show potential for the preparation of functional nanomaterial and nanodevices.^[44] Similarly, amyloid fibrils have shown great potential as bio-nanomaterials based on their properties of extreme stability and resistance to degradation. They have been described as having a similar tensile strength to steel,^[31] as well as physical property that shares structural similarity with spider silk,^[45] which make amyloid fibrils highly attractive for future technological applications. Not only these physical properties, but also the self-assembling characteristic of amyloid fibrils have enhanced their status as nanomaterials. The

⁴⁴ L. Wang et al. *Materials*, **9**: 53, 2016.

⁴⁵ U. Slotta et al. *Macromol. Biosci.*, **7**: 183–188, 2007.

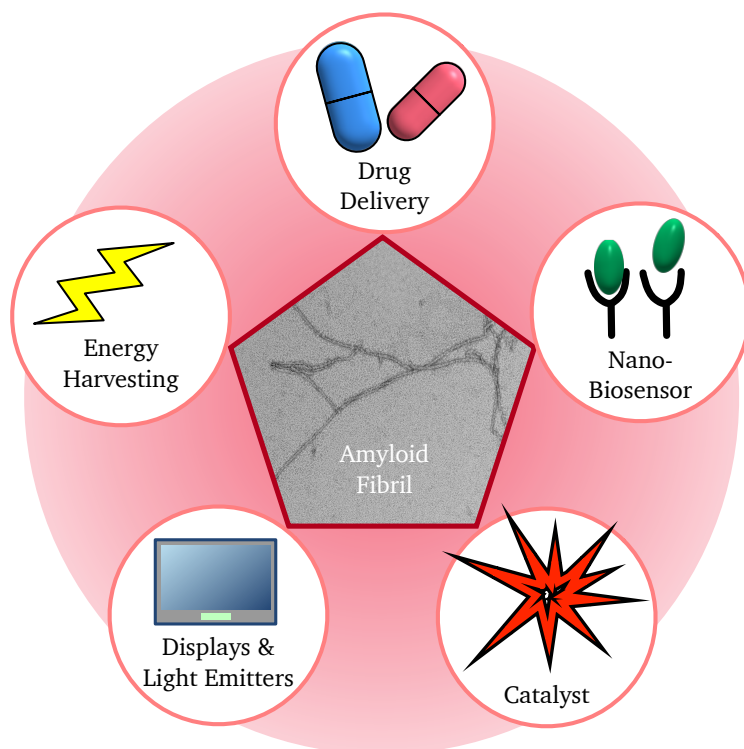


Figure 1.7: Amyloid fibrils have shown the potential to be an attractive bio-nanomaterial for use in various applications such as drug delivery, nano-biosensors, base material for energy harvesting, displays and light emitter devices, and as catalysts.

self-assembly process is a critical part for the “bottom-up” construction of nanostructures, and it is a common feature in many organisms allowing for the development of nano-scale three-dimensional structure.^[46] Physical characteristics such as length and morphology of amyloid fibrils can be varied depending on the choice of protein/peptide, assembly conditions such as concentration, seeding step and temperature control.^[20]

By controlling the synthesis process and biocompatibility, the amyloid fibril has shown potential to be a functional nano-structure in bio-nanodevices and bio-nano-sensors for use in various nano-therapeutics, nano-electronics and energy harvesting devices (Figure 1.7). Fibrils can be functionalised by conjugating proteins/peptides with nano particles (NPs) to form stable and functional nano-hybrid fibrils or they can be used as a platform for the immobilisation of functional groups, catalytic protein domains, antibodies, metals^[47] or non-biomaterials such as graphene.^[48]

⁴⁶ C. L. Chen and N. L. Rosi. *Angew. Chemie Int. Ed.*, **49**: 1924–1942, 2010.

⁴⁷ T. Aumüller and M. Fändrich. *Nat. Chem.*, **6**: 273–274, 2014.

⁴⁸ C. Li, J. Adamcik, and R. Mezzenga. *Nat. Nanotechnol.*, **7**: 421–427, 2012.

Functionalised catalytic amyloid fibrils have been described by various research groups. After generating a fusion protein between a globular protein and poly-alanine with cold shock protein B,^[49] then conjugating enzymatically active proteins, cross-linking was performed to the fully formed fibril.^[50] Furthermore, an enzyme-like amyloid from a short peptide that acts as a Zn^{2+} dependent esterase^[51] and recently, the characteristic zinc-binding structure of a catalytic amyloid by using Solid-state NMR (SSNMR) and structural bioinformatics have been determined.^[52]

1.2.3.1 Branched amyloid fibril

A nano-forest composed with vertically aligned arrayed fibrils are fascinating structures with a variety of applications due to its high-surface area. Similarly, branched amyloid fibrils can cater for high-surface area which can be referred to as a nano-tree structure. The amyloid fibril branching process requires the addition of a monomer to the lateral surface of an existing fibril, which was initially believed to be not formed during the fibril formation due to the lack of reported branched-structures presented by Electron Microscopy (EM) and AFM analysis. However, branched fibrils of sickle cell haemoglobin^[53] and insulin,^[54] have been observed. In addition, real-time branched amyloid fibrillisation of glucagon has been detected with Total Internal Reflection Fluorescence Microscopy (TIRFM).^[18] Despite these, the branching mechanism in fibrillisation is not fully understood and further investigation of branched amyloid fibrillisation process is still required to distinguish between secondary nucleation and true branching process, as well as to fully understand branching mechanism for further applications.^[55]

By combining the concepts of branching amyloid fibrils and nano-forests, along with the generation of the nano-forest composed of vertically aligned branched amyloid fibril, it is possible to achieve maximised surface area which can be used as a platform for the development of various applications such as energy storage, bio-sensors, and energy harvesting in the future.

⁴⁹ M. Sackewitz et al. *Protein Sci.*, **17**: 1044–1054, 2008.

⁵⁰ S. M. Pilkington et al. *Biotechnol. Prog.*, **26**: 93–100, 2009.

⁵¹ C. M. Rufo et al. *Nat. Chem.*, **6**: 303–309, 2014.

⁵² M. W. Lee et al. *Proc. Natl. Acad. Sci.*, **114**: 6191–6196, 2017.

⁵³ R. E. Samuel, E. D. Salmon, and Robin W. Briehl. *Nature*, **345**: 833–835, 1990.

⁵⁴ R. Jansen, W. Dzwolak, and R. Winter. *Biophys. J.*, **88**: 1344–1353, 2005.

⁵⁵ J. E. Gillam and C. E. MacPhee. *J. Phys. Condens. Matter*, **25**: 373101, 2013.

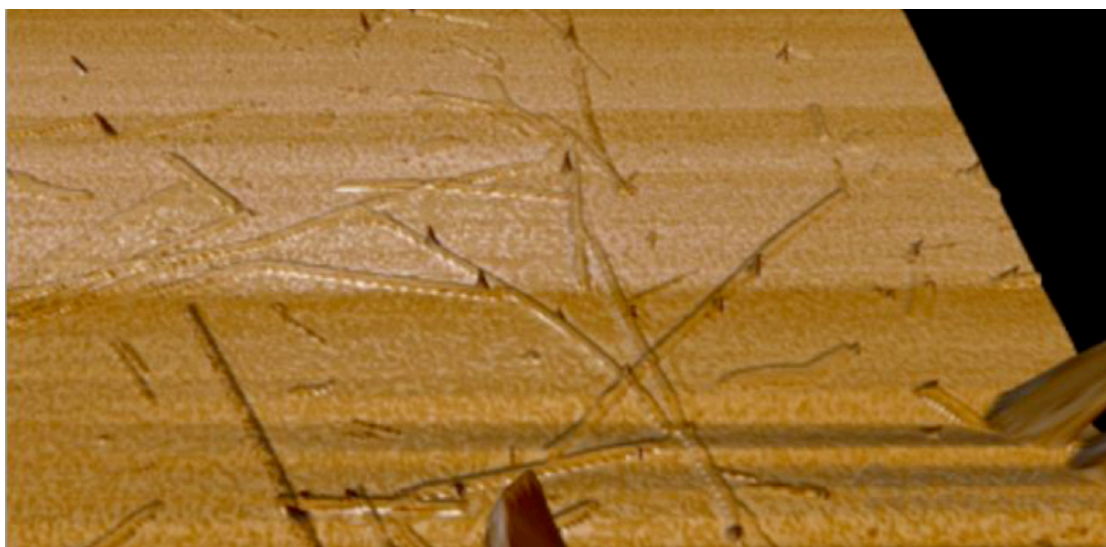


Figure 1.8: AFM image of branched amyloid fibril. (Image provided by Dr K. H. Mok, Trinity College Dublin)

1.3 Methodology for studying protein aggregates

A comprehensive range of biophysical and biochemical tools and techniques including FTIR, SSNMR, Circular Dichroism (CD) spectropolarimetry, Thioflavin T (ThT) assay, and Congo Red (CR) birefringence^[56] have all contributed to a better understanding of the structure of amyloid fibrils at a molecular level, and in undertaking kinetic studies of fibril formation *in vitro*.^[6,7] As amyloid fibrils are insoluble and heterogeneous, commonly used methods of structure determination are difficult. Hence, X-ray diffraction,^[28] TEM, and AFM are often used to study their morphological, structural and physical characteristics. The mature fibrils may possess twists in the structure and the number of twists may differ depending on assembly condition. In addition, the number of protofibrils can also vary.^[20,57–59]

1.3.1 The binding of ThT to amyloid fibrils

Amyloid fibril structure involves continuous β -sheets. ThT is a benzothiazole dye that exhibits enhanced fluorescence by binding rapidly with these β -sheet in the fibrils within seconds. Due to this

⁵⁶ W. E. Klunk, J. W. Pettegrew, and D. J. Abraham. *J. Histochem. Cytochem.*, **37**: 1273–1281, 1989.

⁵⁷ L. Hou et al. *J. Am. Chem. Soc.*, **126**: 1992–2005, 2004.

⁵⁸ H. Levine III. *Protein Sci.*, **2**: 404–410, 1993.

⁵⁹ J. Peng et al. *Sci. Rep.*, **5**: 10171, 2015.

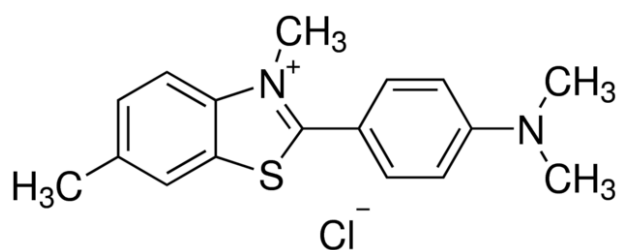


Figure 1.9: The chemical structure of Thioflavin T (Adapted from^[61])

property, the ThT fluorescence assay was developed for the detection of amyloid fibril formation.^[58] Investigation of the amyloid fibril formation process requires not only the ability to distinguish the characteristic amyloid β -sheet structure from amorphous aggregates or non-amyloid fibril forms, but also the quantification of the amyloid formed.^[60] The structure of ThT, shown in Figure 1.9, has a hydrophobic end and a polar benzothiazole group, and this combination creates the possibility for ThT molecules to form micelles in aqueous solution. There are suggestions that the charged nitrogen pointing toward the solvent, form hydrogen bonds with beta-sheets of fibril structures.^[61]

However, the ThT assay can be problematic when used to distinguish the presence of smaller oligomers or protofibrils in the sample due to ThT binding with channels in β -sheets without the protein being in a fibrillary state. Even with this limitation, ThT fluorescence assay is still mainly used as the key indicator of presence of β -sheet structure and for the kinetic study of fibrillisation.^[62]

1.3.2 Protein secondary structure studies with far-UV CD

Circular Dichroism is as an excellent tool to determine the secondary structure and folding properties of proteins. In the spectral region of 190-250 nm (far-UV region), when circularly polarised light impinges on a protein, the protein reflects two circularly polarised components of equal magnitude; one rotating anti-clockwise (left handed, L) and the other clockwise (right handed, R). The ratio between the left and right circularly polarised components is different between secondary structural elements, such as α -helix, β -sheet, and random coil. Hence, different structures give rise to the

⁶⁰H. Levine III. Quantification of β -sheet amyloid fibril structures with thioflavin T in: *Methods in Enzymology: Amyloid, Prions, and Other Protein Aggregates*. ed. by R. Wetzel. **309**: chap. 18, 274–284, Elsevier, 1999. DOI: [10.1016/S0076-6879\(99\)09020-5](https://doi.org/10.1016/S0076-6879(99)09020-5)

⁶¹R. Khurana et al. *J. Struc. Biol.*, **151**: 229–238, 2005.

⁶²M. R. H. Krebs, E. H. C. Bromley, and A. M. Donald. *J. Struc. Biol.*, **149**: 30–37, 2005.

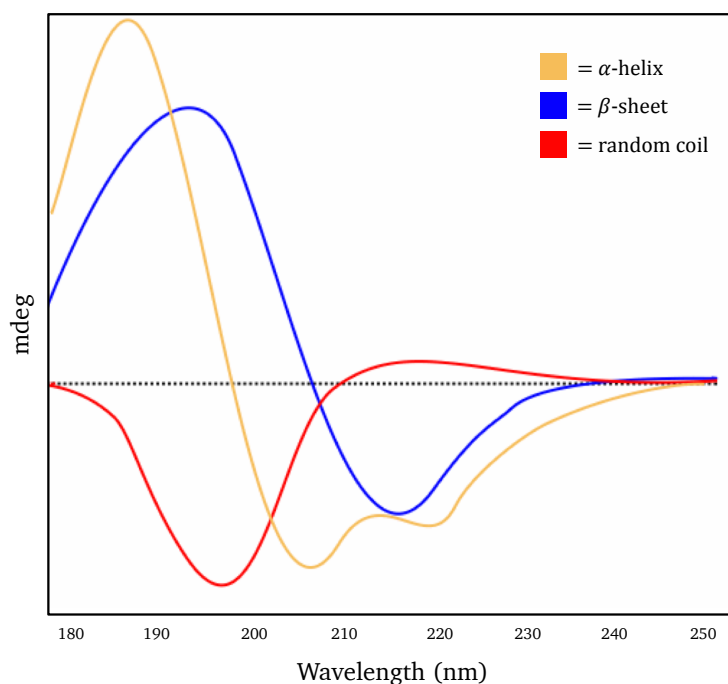


Figure 1.10: Characteristic of Far-UV CD curves of secondary structure elements. (Adapted from^[67])

characteristic bands in specific regions and thereby distinctive shapes and magnitudes in the CD spectrum are visible (Figure 1.10). The Alpha-helix far-UV CD spectra exhibits an intense positive band at 190 nm and negative bands at 208 nm and 222 nm. Beta-sheets are less well-defined in protein compared to the alpha helix, and they exhibit a broad negative band near 218 nm and a large positive band at 195 nm. Disordered random coil far-UV CD spectra shows a weak broad positive band near 217 nm and a large negative band near 200 nm. Because of the strong signals of the α -helix, the absence of α -helicity can also be monitored quite sensitively allowing any secondary structural transitions to β -sheet.^[63–66]

1.3.3 Dynamic light scattering (DLS)

Dynamic Light Scattering (DLS) is a technique that is used to determine particle size distributions in solution sample. When light from a laser source is scattered by particles in solution, light scattered

⁶³ S. M. Kelly, T. J. Jess, and N. C. Price. *Biochim. Biophys. Acta - Proteins Proteomics*, **1751**: 119–139, 2005.

⁶⁴ B. M. Bulheller, A. Rodger, and J. D. Hirst. *Phys. Chem. Chem. Phys.*, **9**: 2020, 2007.

⁶⁵ C. Louis-Jeune, M. A. Andrade-Navarro, and C. Perez-Iratxeta. *Proteins Struct. Funct. Bioinforma.*, **80**: 374–381, 2012.

⁶⁶ N. J. Greenfield. *Nat. Protoc.*, **1**: 2876–2890, 2006.

by each particle will interfere and enhance each other due to Brownian motion, leading to small fluctuations in scattered light intensity. Smaller particles cause the intensity to fluctuate more rapidly than larger particles and correlation takes a shorter time to decay than the correlation signal of larger particles. From the obtained diffusion coefficient information, particle size (hydrodynamic radius/diameter) can be calculated using the *Stokes-Einstein* equation (1.1).^[68–73]

1.4 Solution state nuclear magnetic resonance spectroscopy (NMR)

NMR is a powerful technique that exploits the magnetic properties of an atomic nucleus within a magnetic field and is used to determine the structure of peptide, proteins, and biomolecules,^[5,57,74] as well as protein-protein interactions (proteomics),^[75] and in metabolomics studies.^[76,77]

1.4.1 Spin properties of nuclei

Nuclei of an atom are composed of neutrons and protons. Nuclear spin (I) is the key element for the nuclei to be active in NMR. Nuclei with spin 0, have no NMR properties except for isotope shift on nearby nuclei (^{16}O , ^{12}C , and ^{32}S). NMR active nuclei are nuclei with spin $1/2$, ($^1\text{H}_1$, $^{13}\text{C}_6$, $^{15}\text{N}_7$, $^{31}\text{P}_{15}$, spin $3/2$, ($^7\text{Li}_3$, $^{11}\text{B}_5$) and spin $5/2$, ($^{27}\text{Al}_{13}$). For nuclei that have $1/2$ spin, when immersed in an external static magnetic field B_0 and exposed to a second oscillating magnetic field the spin $1/2$ property allows nuclei to have two different nuclear spin states; parallel or anti-parallel to the field. The population of nuclei in parallel and anti-parallel spin states is almost the same, however, there are small differences in the population that lead to a detectable NMR signal. The NMR signal consists of Radio Frequency (RF) waves with frequencies that match the energy difference between the two orientations of the individual nuclei. Over time the system will return to equilibrium by

⁶⁸ R. Pecora, ed. *Dynamic Light Scattering*. Boston, MA: Springer US, 1985. DOI: [10.1007/978-1-4613-2389-1](https://doi.org/10.1007/978-1-4613-2389-1)

⁶⁹ B. J. Berne and R. Pecora. *Dynamic light scattering : with applications to chemistry, biology and physics*. Mineola: Dover Publications, 2000.

⁷⁰ A. M. Streets et al. *PLOS ONE*, **8**: e54541, 2013.

⁷¹ E. Boisselier et al. *J. Am. Chem. Soc.*, **132**: 2729–2742, 2010.

⁷² S. E. Hill et al. *Biophys. J.*, **96**: 3781–3790, 2009.

⁷³ J. Camponovo et al. *J. Org. Chem.*, **74**: 5071–5074, 2009.

⁷⁴ M. R. Elkins et al. *J. Am. Chem. Soc.*, **138**: 9840–9852, 2016.

⁷⁵ M. R. O'Connell, R. Gamsjaeger, and J. P. Mackay. *Proteomics*, **9**: 5224–5232, 2009.

⁷⁶ J. L. Markley et al. *Curr. Opin. Biotechnol.*, **43**: 34–40, 2017.

⁷⁷ B. Al-Ani et al. *Sci. Rep.*, **6**: 38074, 2016.

allowing the lower energy state to transit to a higher energy state, hence, eventually the signal will decay and vanish. When nuclei are excited with a radio frequency pulse, the Free Induction Decay (FID) is recorded and through Fourier Transformation (FT) the NMR spectrum is obtained.^[5]

1.4.2 Chemical shift scale (ppm) and NMR reference compounds

The differences in resonance frequencies of the nuclei are referred to as chemical shifts and NMR gives the positions of the peak with a scale of parts per million (ppm). The reason for using such a shift scale is that the frequencies of NMR are directly proportional to the magnetic field strength and this causes difficulty in comparing absorption frequencies between spectrometers with different field strengths. The chemical shift scale, ppm is set up by agreeing to reference compounds and the same compounds are used world-wide with the same origin on their shift scales. The position of a peak in the spectrum is specified by measuring its frequency separation from the reference peak, and then dividing this difference by the frequency of the reference peak. In other words, the chemical shift scale peak is the ratio of frequency of sample and reference frequency. Therefore, the field dependence cancels out and this leads to a convenient way of comparing chemical shifts acquired under different field strengths. The chemical shift is very sensitive to environments of individual nuclei such as structural, electronic, magnetic and dynamic variables. This offers the possibility of distinguishing between otherwise identical nuclei in different environments. Hence, chemical shift is often referred to as the ‘fingerprint’ of elements.^[78,79]

As mentioned previously, there is a list of standard reference compounds which are commonly used worldwide, such as Tetramethylsilane (TMS), Trimethylsilylpropanoic acid (TSP), 1,4-dioxane, and 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS). Depending on the solvent conditioning, slight chemical shift dependences are observed for these compounds, particularly due to binding with the analytes (proteins and peptides).^[80,81] In this report, 1,4-dioxane is used as an internal chemical shift standard reference (chemical shift: 3.75 ppm) that is found not to interact with

⁷⁸ J. Keeler. *Understanding NMR Spectroscopy*. 2nd ed. Chichester: John Wiley & Sons, 2010.

⁷⁹ J. Cavanagh et al. *Protein NMR Spectroscopy: Principles and Practice*. 2nd Amsterdam: Elsevier, 2007. DOI: [10.1016/B978-0-12-164491-8.X5000-3](https://doi.org/10.1016/B978-0-12-164491-8.X5000-3)

⁸⁰ G. R. Fulmer et al. *Organometallics*, **29**: 2176–2179, 2010.

⁸¹ H. E. Gottlieb, V. Kotlyar, and A. Nudelman. *J. Org. Chem.*, **62**: 7512–7515, 1997.

proteins and partially-folded proteins.^[82–84]

1.4.3 Biomolecular NMR spectroscopy

Resonance assignment is the first step in the analysis of peptide/protein solution structural study. Total Correlation Spectroscopy (TOCSY) experiments were used to identify individual resonances associated with each spin system, and to identify the origin of amino acid residue. Nitrogen and carbon Heteronuclear Single Quantum Correlation (^1H - ^{15}N HSQC, ^1H - ^{13}C HSQC) were used to determine the nitrogen-proton and carbon-proton cross-peak signals in peptides. Nuclear Overhauser Effect Spectroscopy (NOESY) was used to yield correlations for pairs of protons that are close to each other.^[4,85] After the signal assignment, it is possible to calculate a detailed structure of the protein under investigation using sequence data, chemical shift assignment data, and NOE constraints data.^[4,5,79]

1.4.4 Defining the secondary structure with NMR

Studies clearly show that a strong relationship exists between α -CH ^1H NMR chemical shifts and protein secondary structure for 20 amino acids. As a result of these observations, methods for identifying secondary structure in proteins based on chemical shift information have been established by various researchers. Graphs with chemical shift differences (with respect to random coil average values) versus protein sequence are plotted, and with the obtained curve the identity and location of secondary structure is predicted. As a result, secondary structure can be assigned to α -helix, β -strand (β -sheet), or coil (loop, random coils, turns which are neither α -helix nor β -strand).^[86–90]

⁸² A. Shimizu, M. Ikeguchi, and S. Sugai. *J. Biomol. NMR*, **4**: 859–862, 1994.

⁸³ A. De Marco. *J. Magn. Reson.*, **26**: 527–528, 1977.

⁸⁴ J. Pettersson-Kastberg et al. *J. Mol. Biol.*, **394**: 994–1010, 2009.

⁸⁵ A Bax. *Annu. Rev. Biochem.*, **58**: 223–256, 1989.

⁸⁶ D. S. Weinstock et al. *Protein Sci.*, **17**: 950–954, 2008.

⁸⁷ Y. Wang. *Protein Sci.*, **11**: 852–861, 2002.

⁸⁸ G. Cornilescu, F. Delaglio, and A. Bax. *J. Biomol. NMR*, **13**: 289–302, 1999.

⁸⁹ D.S. Wishart, B.D. Sykes, and F.M. Richards. *J. Mol. Biol.*, **222**: 311–333, 1991.

⁹⁰ D. S. Wishart, B. D. Sykes, and F. M. Richards. *Biochemistry*, **31**: 1647–1651, 1992.

1.4.5 Diffusion NMR: Diffusion Ordered Spectroscopy (DOSY) NMR

The phenomenon of molecular diffusion in solution potentially offers a range of physical properties including molecular size, shape, aggregation, encapsulation, complexation, and hydrogen bonding, enabling the application of diffusion NMR in many areas of chemistry over the years. All modern diffusion NMR measurements are based on the application of field gradients to encode the physical location of a component of interest in solution and to characterise its diffusion rate along the direction of the applied field gradient.^[91]

Pulse Field Gradient (PFG) NMR is a well-established technique and the current method of choice for measuring diffusion in solutions of both chemical and biological systems. In principle, the diffusion coefficient of a specific molecular species depends on its molecular weight, size and shape, and under given conditions such as solvent and temperature. Therefore, diffusion can be used to determine interactions of intermolecular in-solution without disturbing the system, and it plays an important role in both chemistry and biology studies. Determining the intramolecular interaction is a key step for molecular recognition, supramolecular and combinatorial chemistry studies and more.^[92]

Morris KF and Johnson CS^[94,95] reported that diffusion spectra could be obtained by incrementing the areas of the gradient pulses (q) in PFG–NMR and transforming the NMR signal amplitudes with respect to q^2 , which results in Diffusion Order Spectroscopy (DOSY). DOSY is the pulsed field gradient spin echo-based experiment, which consist of a series of spin echo spectra with different pulsed field gradient strengths. The signal decays are analysed to extract a set of diffusion coefficients (diffusion constants) depending on the size of particles. A DOSY spectrum has conventional chemical shift spectra in one dimension and diffusion spectra in the other dimension. Due to the straightforward relationship between translational diffusion rates and particle sizes, DOSY is especially useful for complex mixtures with a large range of particle sizes, and it is possible to obtain the signals of individual compounds depending on the particle sizes by separating it

⁹¹ T. D.W. Claridge. Diffusion NMR Spectroscopy in: *High-Resolution NMR Techniques in Organic Chemistry*. 3rd ed. chap. 10, 381–419, Oxford: Elsevier, 2016. DOI: [10.1016/B978-0-08-099986-9.00010-5](https://doi.org/10.1016/B978-0-08-099986-9.00010-5)

⁹² Y. Cohen, L. Avram, and L. Frish. *Angew. Chemie Int. Ed.*, **44**: 520–554, 2005.

⁹⁴ K. F. Morris and C. S. Johnson. *J. Am. Chem. Soc.*, **114**: 3139–3141, 1992.

⁹⁵ K. F. Morris and C. S. Johnson. *J. Am. Chem. Soc.*, **115**: 4291–4299, 1993.

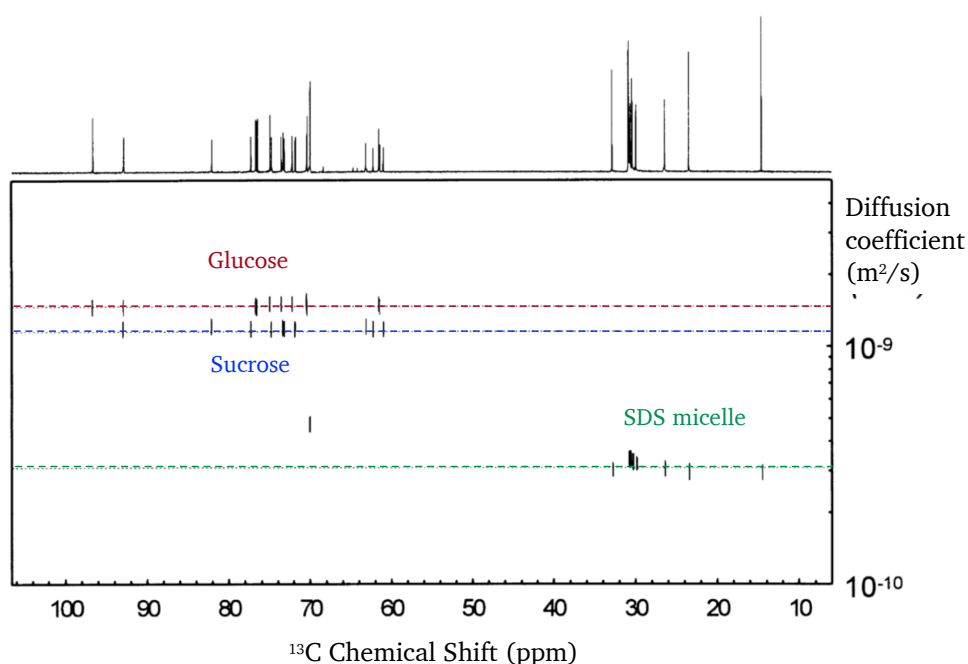


Figure 1.11: Pseudo-2D-DOSY plot obtained from a mixture containing glucose **Red**, sucrose **Blue**, and SDS micelle **Green** in D_2O . Dotted lines in the two-dimensional spectrum show average diffusion coefficients of each individual components.^[93]

into different rows as a pseudo-2D contour presentation (Figure 1.11). In this, one-dimension represents chemical shifts and the other, the diffusion coefficients of each components.^[94–96]

1.4.5.1 Diffusion coefficients and molecule size

Diffusion coefficients, D , have an inverse correlation with the molecular weight of the diffusion species. Diffusion coefficients can be calculated with the *Stokes-Einstein* equation (1.1). The larger the diffusing molecules or complexes, the smaller is the exhibited value of the diffusion coefficients (that is, larger molecules in solution diffuse slower than smaller molecules).

$$D = \frac{kT}{6\pi\eta r_s} \quad (1.1)$$

Where,

D = Diffusivity, η = Viscosity, T = Temperature, and r_s = Hydrodynamic radius

⁹⁶C. S. Johnson. *Prog. Nucl. Magn. Reson. Spectrosc.*, **34**: 203–256, 1999.

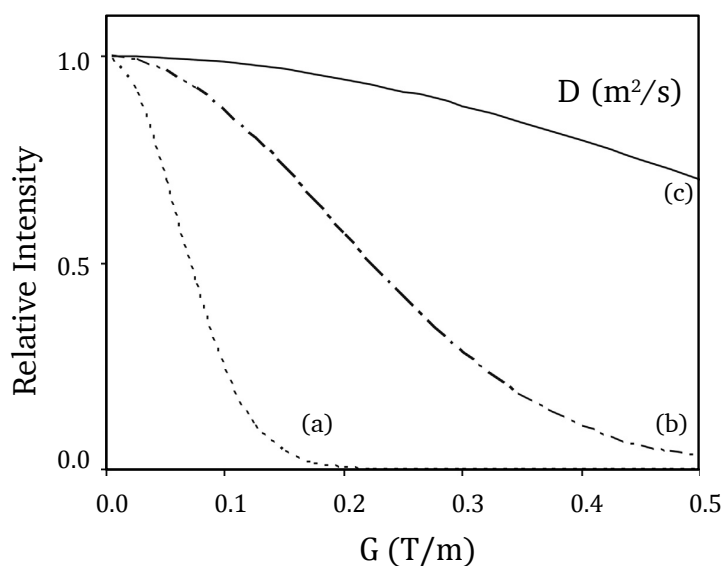


Figure 1.12: When signal attenuation is too fast (trace (a)) the latter data points do not contribute to the profile, when it is too slow (trace (c)) there is insufficient attenuation to provide an accurate determination of D . Trace (b) shows an optimum profile where all data points would contribute to characterisation of the decay. (Adapted from^[91])

Even though diffusion coefficient values are directly inversely related to the size of the diffusing species, it is affected by many other parameters, including the properties of the solvent such as, its viscosity, sample temperature and concentration; this means that the parameters for diffusion experiments have to be optimised on a sample-to-sample basis. The purpose of optimisation is to obtain a well-characterised decay profile as shown in Figure 1.12-(b) which allows for reliable data fitting of the curve.^[91]

From measurements of diffusion coefficients, calculation of hydrodynamic radius, molecular volumes, and even estimation of molecular weights is possible. Therefore, diffusion NMR has been applied in biochemistry research to address commonly asked questions pertaining to molecular size: Whether species exist as discrete monomers, dimers or higher oligomers, or whether biomolecules have a tendency to aggregate. A number of groups published in the relevant literature have applied diffusion NMR to biological studies. Price *et al.*^[97] have shown lysozyme aggregation and solution properties studies using Pulsed Gradient Spin Echo (PGSE) NMR diffusion measurements^[97] Also, Pulse field gradient NMR methods have been used to determine the effective hydrodynamic radii of a range of native and non-native protein conformations. The predicted values of the hydrodynamic

⁹⁷ W. S. Price, F. Tsuchiya, and Y. Arata. *J. Am. Chem. Soc.*, **121**: 11503–11512, 1999.

radii provide a framework that can be used to analyse the conformational properties of a range of non-native states of proteins.^[98]

As mentioned previously, the size of protein is often described by referring to its molecular weight (MW) or hydrodynamic radius (or Stokes radius) (R_h). Thus, in this thesis, hydrodynamic radius of proteins and aggregated proteins formed during the early stage aggregation is calculated with obtained diffusion coefficient values. Hydrodynamic radius is defined as the hypothetical hard sphere that diffuse in the same fashion at the identical experiment environment, reflects the apparent size adopted by the solvent (hydro-) and shape (-dynamic) of molecule.^[91] However, in reality proteins and aggregated proteins are not perfectly spheres. Also, protein aggregates with the same number of protein assembly can form into various shapes. Therefore, hydrodynamic radius calculated from obtained diffusion coefficient of the early stage aggregates species is an indication of the apparent size of the dynamic hydrated biomolecules.

1.5 Objective and outcome of the research

Using biomolecular NMR spectroscopy and other high-resolution biophysical techniques, we investigate various protein and peptides with the objective of

- (a) Characterising oligomerization and aggregation through the further development and the application of diffusion NMR pulse sequences,
- (b) Probing the structural details of β -Lg at neutral pH when it is heat-denatured at high concentration which is important for dairy industrial applications, and
- (c) Calculating the atomic-level three-dimensional structure of a TLR4 inhibiting peptide terms VIPER and its inactive mutant peptide.

⁹⁸D. K. Wilkins et al. *Biochemistry*, **38**: 16424–16431, 1999.

Chapter 2

Materials & Methods

2.1 Proteins

2.1.1 HEWL

Lysozyme from chicken egg white, in the form of lyophilised powder was purchased from Sigma-Aldrich (Sigma-Aldrich Co., St. Louis, MO, USA).

- Molecular weight (*kDa*) : 14.3
- No. of residues (*aa*) : 129
- Hydrodynamic radii. Rh (Å) : 20.5^[98]

2.1.2 Bovine β -lactoglobulin

Bovine β -lactoglobulin (β Lg) was used to synthesise fibrils in Chapter 3 and to study the heat stability with NMR in Chapter 5. Samples were provided by Aoife Buggy and Dr Andre Brodkorb (Teagasc Moorepark Research Centre). Whereas, β -Lg used for the diffusion NMR experiment in Chapter 3 was purchased from Sigma-Aldrich (Sigma-Aldrich Co., St. Louis, MO, USA).

- Molecular weight (*kDa*) : 18.4
- No. of residues (*aa*) : 162

⁹⁸D. K. Wilkins et al. *Biochemistry*, **38**: 16424–16431, 1999.

-
- Hydrodynamic radii. Rh (Å) : 22^[99]

2.1.3 Myoglobin

Myoglobin from equine skeletal muscle in the form of lyophilised powder was purchased from Sigma-Aldrich (Sigma-Aldrich Co., St. Louis, MO, USA).

- Molecular weight (*kDa*) : 16.9
 – No. of residues (*aa*) : 153

2.1.4 VIPER

9R-VIPER and 9R-VIPER L6AE10A peptides used in the structural study with solution NMR (Chapter 5) were manufactured from GenScript (Piscataway, USA) and provided by Prof Andrew Bowie's lab (School of Biochemistry and Immunology, Trinity College Dublin)

9R-VIPER	Sequence	:	RRRRRRRRRR-KYSFKLILAEY
	Lot no.	:	148039001101211wj
	Purity	:	99.0 %
	MW	:	2780.31
9R-VIPER L6AE10A	Sequence	:	RRRRRRRRRR-KYSFKAILAAY
	Lot no.	:	P17351405
	Purity	:	99.8 %
	MW	:	2680.19

2.1.5 Other proteins

Purchased from Sigma-Aldrich (Sigma-Aldrich Co., St. Louis, MO, USA).

- Thyroglobulin (bovine)
- γ -Globulin (bovine)

⁹⁹ R. Parker et al. *Biophys. J.*, **89**: 1227–1236, 2005.

- Human Serum Albumin
- Cytochrome C
- Tryptophan
- α -lactoglobulin
- Apo-Myoglobin (Apo-Mb) preparation from Myoglobin (Mb)

4~6 ml of Ethyl Methyl Ketone (EMK) was pipetted into two empty pre-chilled glass tubes, which were then placed in the ice bath. Lyophilised Myoglobin powder (7 mg) was transferred into a chilled glass tube and 6 ml H₂O was added to it and then pH was adjusted to pH 3 with 0.1 M HCl. Defined pH indicator paper was used to check pH. The pre-chilled EMK (6 ml) was added to Mb solution and vigorously shaken and then Mb solution was placed in the ice-bath for 1-2 minutes until Mb solution divided into two layers. Top layer is the EMK layer with Heme (red colour layer) and bottom layer is the H₂O layer with Apo-Mb (Pale yellow colour layer). The heme in EMK was carefully transferred into a glass tube and stored in fridge and Apo-Mb layer was transferred into a plastic tube and kept in the -20°C . Once the sample was frozen, the water was removed from the solution using a Free Zone 1 L Benchtop Freeze Dry System (Labconc, MO, USA). The solution was lyophilised after reaching a temperature of -50°C leaving a pure protein sample without water. Lyophilised Apo-Mb was stored in a -20°C freezer.^[100]

2.2 Other chemicals

- Thioflavin T and Congo Red were purchased from Sigma-Aldrich (Sigma-Aldrich Co, St. Louis, MO, USA).
- NuPAGE[®] Bis-Tris Mini Gels 12 %, NuPAGE[®] LDS sample buffer, NuPAGE[®] Antioxidant, MOPS SDS Running buffer and Spectra multicolour broad range protein ladder were purchased from Thermo Fisher Scientific (Waltham, MA USA).
- NativePAGE[™] Novex[®] gel (3-12 % Bis-Tris / 1.0 mm), NativePAGE[®] sample buffer, NativePAGE[®] running buffer, and NativePAGE[®] cathode buffer additive was purchased from Thermo Fisher Scientific (Waltham, MA, USA).
- Protein MW standards and NativePAGE[®] protein standard (Protein ladder) was purchased

¹⁰⁰F. W. J. Teale. *Biochim. Biophys. Acta*, **35**: 543, 1959.

Table 2.1: Molecular weight (*kDa*), ϵ_{280} ($M^{-1}cm^{-1}$), $\epsilon_{1\%}$ (10 mg/ml) $^{-1}cm^{-1}$ value of proteins and tryptophan

Protein	Molecular Weight (<i>kDa</i>)	ϵ_{280} ($M^{-1}cm^{-1}$)	$\epsilon_{1\%}$ (10 mg/ml) $^{-1}cm^{-1}$
Thyroglobulin (bovine)	675	675000	10
γ -Globulins (bovine)	150	343000	14
Human Serum Albumin	66.5	41435	6.2
β -Lactoglobulin (bovine)	18.4	17600	9.6
Myoglobin (Equine skeletal muscle)	16.9	13940	8.23
Hen Egg White Lysozyme (HEWL)	14.3	37970	26.5
Cytochrome C (Equine Heart)	12.3	13075	10.6
Tryptophan	0.2	5690	284.5

from Thermo Fisher Scientific (Waltham, MA USA).

- BS³ (bis(sulfosuccinimidyl)suberate): Cross-linker was purchased from Thermo Fisher Scientific (Waltham, MA USA).
- 1,4-Dioxane was purchased from Sigma-Aldrich (Sigma-Aldrich Co, St. Louis, MO, USA).
- D₂O was purchased from Sigma-Aldrich (Sigma-Aldrich Co, St. Louis, MO, USA).

2.3 Lab materials

- XCell SureLockTM Mini-Cell electrophoresis system was purchased from Thermo Fisher Scientific (Waltham, MA USA).
- Polystyrene 0.22 *mm* bottle top filter and the 0.20 *mm* sterile syringe filter were purchased from Corning (Corning Incorporated, NY, USA).
- Slide-A-Lyzer[®] Dialysis Cassettes were purchased from Thermo Fisher Scientific (Waltham, MA USA).
- AFM Cantilevers (DNP & SNL) were purchased from Bruker AFM Probes (Bruker, Camarillo,

CA, USA). Both DNP and SNL tips are soft Silicon Nitride tips with the tip radius (Nom) of 20 *nm* (DNP) and 2 *nm* (SNL).

2.4 Protein aggregation: oligomerization & fibrillization

2.4.1 Preparation of HEWL oligomer & fibril

HEWL was dissolved in a dH₂O/HCl solution at pH 2 (pH of HEWL in dH₂O altered with the addition of 4 *M* HCl). The final concentration of solution was amended to yield at 1 *mM* (MR = 14.3 *kDa*, extinction coefficient = 26.4 at 280 *nm* for a 1 % (10 *mg/ml*) HEWL solution) by using NanoDrop ND-1000 (NanoDrop, USA). Aliquots of this mixture were incubated at 65 °C for short period from 6 hours and longer time period to get dimers/oligomers and up to 21 days to allow sufficient fibrillisation to occur.

2.4.2 Preparation of cross-linked HEWL by BS³

HEWL 2 *mM* and BS³ 10 *mM* stock solution was prepared in 50 *mM* phosphate buffer at pH 7.4. Prepared stock solutions were mixed with various ratio as shown in Table 2.2 below. This was done to obtain a different size of multimers due to the ratio difference between cross-linker and protein. At this mixture preparation step, various volumes of dH₂O were first added to 1 *mL* of protein stock and then BS³ cross-linker stock solution was added. Once BS³ was added to protein, the HEWL-BS³ mixture solutions were incubated for 1 hour at room temperature and centrifuged for 20 mins at 4.4×1000 *rpm* to remove insoluble aggregates. Only supernatants were corrected and dialysis (3500 MWCO) in 20 *mM* phosphate buffer was performed overnight to remove free cross-linkers. The SDS-PAGE gel was then used with 10-fold diluted crosslinked-HEWL samples to determine the molecular weight of various crosslinked-HEWL present in the prepared mixture sample. Prepared samples contained various ratio of HEWL monomer, dimer, trimer, and bigger aggregates as a result of the concentration of added crosslinker.

Table 2.2: HEWL-BS³ mixture sample preparation

BS ³ final conc. [<i>mM</i>]	BS ³ 10 <i>mM</i> stock volume	HEWL 2 <i>mM</i> stock volume	dH ₂ O volume	Total volume
5	1.0 mL	1.0 mL	-	2.0 mL
2	400 μ L	1.0 mL	600 μ L	2.0 mL
1	200 μ L	1.0 mL	800 μ L	2.0 mL
0.5	100 μ L	1.0 mL	900 μ L	2.0 mL
0.1	20 μ L	1.0 mL	980 μ L	2.0 mL
0	-	1.0 mL	1.0 mL	2.0 mL

2.4.3 HEWL-BS³ NMR sample preparation

A total volume of 600 μ L cross-linked HEWL sample (H₂O / D₂O 10 % *v/v*, with 1 *mM* Dioxane) was transferred into the NMR tube. Dioxane was then added as internal chemical shift standard (3.75 *ppm*).

2.4.4 Nano-structured films from amyloidogenic protein: HEWL

In the pre-seeding stage, to prepare amyloid fibril seed solution, hen egg white lysozyme (Sigma Aldrich) 3 % *w/w* was incubated in aqueous 10 *mM* hydrochloric acid at 65 °C for 14 days to form the self-assembled amyloid fibrils. After 14 days, in the post-seeding stage, this pre-formed fibril suspension was added into the new HEWL 3 % *w/w* in aqueous 10 *mM* HCl, as seed material at the beginning of the incubation at 65 °C for 7 days. At the end of the fibril growth step, glycerol 0.8 % *v/v* (Sigma Aldrich) was added to the amyloid fibril hydro-gel as plasticizers. Films were then cast by transferring 0.5 *mL* of the amyloid fibril hydro-gel onto a flat Teflon non-stick liner (First4Spares). The solvent and the volatile acid were left to evaporate for +24 hours and the resulting amyloid fibril film (protein film) removed with tweezers.^[101]

¹⁰¹ T. P. J. Knowles et al. *Nat. Nanotechnol.*, 5: 204–207, 2010.

2.4.5 Preparation of bovine β -Lg fibrils

Lyophilised β -Lg powder was dissolved in water at pH 2 (+10 hours overnight) to allow complete hydration of the protein. The hydrated protein solution was centrifuged at 3900 *rpm* for 10 min at 4 °C using a Centrifuge 5810R (Eppendorf), and the supernatant was filtered with 0.22 μ m syringe filter to remove any insoluble particles. The pH 2 water adjusted with 4 M HCl was used to achieve a final concentration of 2 % *w/v*. The concentration of final β -Lg solution was checked with NanoDrop ND-1000 (NanoDrop, USA) by measuring the absorbance at 280 nm. β -Lg 2 % *w/v* sample was placed in the heat-block at 80 °C for 6 hours for fibrillisation. Once incubation was finished, fibrillisation in the sample was quenched by rapid cooling in an ice bath. The cooled samples were used freshly or kept at 5 °C to be used later.

2.4.6 Preparation of apo-myoglobin (Apo-Mb) fibril

Lyophilised Apo-Mb was used to prepare 1 *mg/mL* Apo-Mb solution with 50 *mM* sodium borate buffer (Borex), pH 9. Well dissolved Apo-Mb solution was filtered through a 0.22 μ m syringe filter to remove insoluble particles and then incubated in the heat-block at 65 °C for 5 days.^[21,102]

2.5 Amyloid fibril formation kinetics monitored by ThT assay

In order to detect the presence of amyloid fibril and to monitor the fibril formation kinetics, the ThT fluorescence assay was used. ThT assay is the most popular method to monitor amyloid fibril formation and growth.^[58,60,62]

- ThT buffer solution

Made in dH₂O with condition of 10 *mM* KH₂PO₄, 150 *mM* NaCl at pH 7.0. ThT buffer solution

²¹ M. Fändrich, M. A. Fletcher, and C. M. Dobson. *Nature*, **410**: 165–166, 2001.

¹⁰² M. Fändrich et al. *Proc. Natl. Acad. Sci.*, **100**: 15463–15468, 2003.

⁵⁸ H. Levine III. *Protein Sci.*, **2**: 404–410, 1993.

⁶⁰ H. Levine III. Quantification of β -sheet amyloid fibril structures with thioflavin T in: *Methods in Enzymology: Amyloid, Prions, and Other Protein Aggregates*. ed. by R. Wetzel. **309**: chap. 18, 274–284, Elsevier, 1999. DOI: [10.1016/S0076-6879\(99\)09020-5](https://doi.org/10.1016/S0076-6879(99)09020-5)

⁶² M. R. H. Krebs, E. H. C. Bromley, and A. M. Donald. *J. Struc. Biol.*, **149**: 30–37, 2005.

is then filtered through 0.22 μm bottle top filter and stored in the 250 mL glass bottle.

- ThT stock solution

Prepared by dissolving 8 mg ThT in 1 mL of ThT buffer solution and then filtering through a 0.22 μm syringe filter to remove insoluble particles. Prepared ThT stock solution was stored in 15 mL tube and covered with tin foil and stored in the refrigerator.

- ThT working solution

To generate a ThT working solution, ThT stock solution was diluted to 1 in 50 with ThT buffer solution. ThT working solution is freshly made before use.

- ThT fluorescence assay

Fluorescence emission spectra was acquired on a *JASCO FP-6200 spectrofluorometer* (Tokyo, Japan) by exciting at 440 nm and recording the fluorescence intensity at a wavelength range between 470 and 520 nm . Both the excitation and emission slit widths were set at 5 nm . In a typical assay, 1 mL ThT working solution was transferred into the cuvette and the blank curve was measured which is followed by the addition of 10 μL of protein to ThT working solution. The solution is mixed thoroughly, and measurement is acquired 3 times for each sample.

2.6 Congo Red (CR)

2.6.1 CR birefringence assay

Prepared ethanol 80 % / water 20 % with saturating amount of NaCl and then added saturating amount of CR. CR and NaCl were allowed to be dissolved in the solution and then filtered through a 0.22 μm syringe filter to remove insoluble particles. Once the CR stain solution was ready, 10 μL of protein/fibril sample was dropped on the glass and allowed to dry. Once the sample dried, 10 μL of CR stain solution was dropped on top of it and the excess solution was blotted away with a lint free kimwipe and allowed to dry. The solution was then examined under a polarised light

microscope at 90 °C!^[19,103,104]

2.6.2 CR spectroscopic assay

CR was dissolved (7.0 *mg/mL*) in a buffer consisting of 5 *mM* potassium phosphate, 150 *mM* NaCl with pH 7.4 (pH adjusted with NaOH) and filtered through a 0.22 μm syringe filter immediately before use. Prepared CR stock solution was then diluted 100-fold with the same buffer to make a working solution. A spectrum between 400 ~ 600 *nm* of working solution alone was acquired as blank data and fibril sample of 10 μL was added into the working solution and spectrum was acquired at the same range. A shift of the maximum CR absorbance from 480~490 *nm* to the region of 540 *nm* was taken as indicative of the presence of amyloid fibrils.^[56]

2.7 Dynamic Light Scattering (DLS)

Dynamic Light Scattering (DLS) is a technique which can be used to determine the size of particles such as proteins, promiscuous inhibitors, micelles, quantum dots, metallic nanoparticles and more in a solution^[69] In this study, DLS is used to detect the size distribution profile change occurring during the amyloid fibril formation. Prepared protein sample of 5 μL was transferred into a cuvette and DLS measurements were all performed on DynaPro-NanoStar (Wyatt Technology, CA, USA) at 25 °C. Cuvette with path length 2 *mm* was used (Eppendorf UVette).

2.8 Circular Dichroism (CD) spectroscopy

Circular dichroism (CD) is a well-established method of determining the secondary structure of proteins and folding properties by defining the unequal absorption of left-handed and right-handed circularly polarized light. In the spectral region 190-250 which is referred to as Far-UV

¹⁹ M. Nilsson. *Methods*, **34**: 151–160, 2004.

¹⁰³ A. J. Howie et al. *Lab. Investig.*, **88**: 232–242, 2008.

¹⁰⁴ R. Khurana et al. *J. Biol. Chem.*, **276**: 22715–22721, 2001.

⁵⁶ W. E. Klunk, J. W. Pettegrew, and D. J. Abraham. *J. Histochem. Cytochem.*, **37**: 1273–1281, 1989.

⁶⁹ B. J. Berne and R. Pecora. *Dynamic light scattering : with applications to chemistry, biology and physics*. Mineola: Dover Publications, 2000.

region, CD detects unique characteristic shape and magnitude of α -helix, β -sheet, and random coil conformation. Hence, CD can examine the presence and formation of β -sheet rich structure amyloid fibrils in the protein solution.^[63–66]

All CD measurements were performed on *JASCO J-815 spectroscopy* (Tokyo, Japan) at room temperature (21 °C). A quartz glass cuvette with path length of 1 *mm* was used for far-UV measurements (Hellma Analytics, USA), whereas a path length of 5 *mm* was used for near-UV measurements (CHANDOS Intercontinental, USA).

2.8.1 Far-UV CD: sample conditions & parameters

For far-UV CD measurement, protein concentration between 50 $\mu\text{g/mL}$ $\sim$ 1 mg/mL was used and 400 μL of sample was used.

- **Setting parameters**

Sensitivity	:	Standard	(100 <i>mdeg</i>)
Region	:	185-260	(<i>nm</i>)
Data pitch	:	1	(<i>nm</i>)
Scanning mode	:	Continuous	
Scanning speed	:	50	(<i>nm/min</i>)
Response	:	0.5	(<i>sec</i>)
Bandwidth	:	1	(<i>nm</i>)
Accumulation	:	8	(scans)

2.8.2 Near-UV CD: sample conditions & parameters

For near-UV CD measurement, protein concentration between 0.25 $\sim$ 2 mg/mL was used and a volume between 800 $\sim$ 1000 μL was used.

⁶³ S. M. Kelly, T. J. Jess, and N. C. Price. *Biochim. Biophys. Acta - Proteins Proteomics*, **1751**: 119–139, 2005.

⁶⁴ B. M. Bulheller, A. Rodger, and J. D. Hirst. *Phys. Chem. Chem. Phys.*, **9**: 2020, 2007.

⁶⁵ C. Louis-Jeune, M. A. Andrade-Navarro, and C. Perez-Iratxeta. *Proteins Struct. Funct. Bioinforma.*, **80**: 374–381, 2012.

⁶⁶ N. J. Greenfield. *Nat. Protoc.*, **1**: 2876–2890, 2006.

- **Setting parameters**

Sensitivity	:	Standard	(100 <i>mdeg</i>)
Region	:	240-360	(<i>nm</i>)
Data pitch	:	1	(<i>nm</i>)
Scanning mode	:	Continuous	
Scanning speed	:	50	(<i>nm/min</i>)
Response	:	0.5	(<i>sec</i>)
Bandwidth	:	1	(<i>nm</i>)
Accumulation	:	8	(scans)

2.8.3 Estimation of secondary structure from CD data

The obtained CD spectra is the averaged data of all scans performed. CD data are presented in ellipticity $[\theta]$. The mean residue ellipticity $[\theta]_{MRW}$ ($deg.cm^2.dmol^{-1}$) was calculated using Equation (2.1) and Equation (2.2).^[63–66]

In this thesis, online analysis software K2D3 DichroWeb (London, UK) was also used to process CD data and estimate protein secondary structure.^[105,106]

$$\theta_{MRW} = \left(\frac{\theta_{obs}}{10} \right) \left(\frac{MRW}{l * c} \right) \quad (2.1)$$

Where,

θ_{obs} = the observed ellipticity in degree,

MRW = mean residue weight (shown in Equation (2.1),

l = the light path length in *cm*,

c = the protein concentration in *g/ml*

$$MRW = \frac{M}{N - 1} \quad (2.2)$$

Where,

¹⁰⁵ L. Whitmore and B. A. Wallace. *Nucleic Acids Res.*, **32**: W668–W673, 2004.

¹⁰⁶ L. Whitmore and B. A. Wallace. *Biopolymers*, **89**: 392–400, 2008.

Table 2.3: SDS-PAGE gel sample preparation ratio

Reagent	Reduced Sample	Non-reduced Sample
Sample	10 μL	10 μL
NuPAGE® LDS sample buffer	5 μL	5 μL
DTT (1 M): Reducing Agent	2 μL	0 μL
Deionized water	3 μL	5 μL
Total sample volume	20 μL	20 μL

M = the molecular mass of the polypeptide chain in Da ,

N = the number of amino acids in the chain,

$N-1$ = the number of peptide bonds

2.9 SDS-PAGE denaturing gel electrophoresis

To prepare SDS-PAGE Gels, sample NuPAGE® LDS sample buffer and DTT (1M, reducing agent) added to protein sample. Sample preparation ratio followed is described in Table 2.3. Vortexed samples were placed in heat-block at 80 °C for 10 minutes. In the meantime, 1X SDS running buffer was prepared by diluting 50 mL 20X NuPAGE® LDS sample buffer with 950 mL of deionized water. Once the sample heating process was finished, the gel (NuPAGE® Bis-Tris Mini Gel) was placed in the electrophoresis system and then 10 μL of protein sample and 12 μL of the protein ladder was loaded on the gel. Both upper and lower buffer chamber of the electrophoresis system were filled with prepared 1X SDS running buffer, and when the reduced sample is loaded in the gel, 500 μL of antioxidant (NuPAGE® Antioxidant) was added in the upper buffer chamber. Finally, the electrophoresis system was connected to the power supply with voltage setting of 120 V and run for 120 minutes.

The resolved gel was then washed with deionized water and stained for minimum 4 hours and placed in the de-stain overnight. A broad range protein ladder used in this thesis shows 10 different protein bands at 10, 17, 26, 34, 43, 55, 72, 95, 130, and 180 kDa , hence, referring to these ten protein ladder bands, it was possible to determine the molecular weights of protein sample.

2.10 Transmission Electron Microscopy (TEM) imaging

Transmission electron microscopy (TEM) was used to confirm the presence of and determine the morphology of amyloid fibril. Amyloid fibril samples were visualized by using a *JEOL JEN-1400 TEM* (Tokyo, Japan) operating at 120 *kV* LaB6 with microscope camera XR80.

The β -Lg protein sample which prepared using the method described in Section 2.4.5 was used for the TEM imaging. β -Lg protein sample without applying the heating process was used to achieve the native β -Lg protein image, and the β -Lg protein incubated for 6 hours was used to obtain the evidence of fibril formation.

A 20 μ L droplet of protein sample was applied to fresh clean para-film and 300 mesh Carbon formvar coated Copper grid was inverted onto the droplet and allowed to incubate for 15 mins. The grid was then washed and transferred (inverted) to fresh droplet of filtered dH₂O for 15 mins. Excess water was carefully wicked away with a filter paper (Whatman 50) wedge and the grid was transferred to a fresh filtered 2 % uranyl acetate droplet for 60 seconds. Excess solution was wicked away with a filter paper wedge and the grid was allowed to dry before examination in TEM.

2.11 Atomic Force Microscopy (AFM) imaging

All AFM imaging was obtained in-liquid with *Asylum Research Olympus MFP 30 AFM* (Oxford Instruments company, Tubney Woods, Abingdon, UK) at the Conway Institute (University College Dublin).

2.11.1 Sample preparation on mica surface for AFM

Unmodified AFM-grade mica (diameter = 1 *cm*) was glued on the glass slide. Before use, the mica surface was exfoliated by peeling off the layers with sticky tape. It was then immediately placed in the UV Ozone cleaner for 10 mins for removal of any organic contaminants and to enhance the hydrophilicity of mica surface. After UV Ozone treatment, 10 μ L of diluted protein sample (concentration of sample varies for each sample) was deposited on the mica and allowed to incubate

for 10 mins in a dust-free atmosphere by covering the whole glass slide. Lastly, the glass-slide with the sample was tilted to 45 degree and then gently washed (twice) with 200 μL Milli-Q water to remove any excess sample. For the in-air measurement, excess water was dried with air/ N_2 , however, because the in-liquid measurement was used in this study, extra 200 μL Milli-Q water was directly applied on the mica before the measurement.

2.11.2 AFM in-liquid measurement

Each cantilever has a different spring constant value which is required for AFM measurement. For the calibration, a cantilever (DNP or SNL) was fitted in the cantilever holder and a clean glass slide was positioned on the AFM sample stage. The cantilever holder was then fitted in the AFM head and the head was carefully placed over the sample stage. The spring constant of the cantilever was measured at the contact mode and once the spring constant value was gained after the calibration, the glass slide was removed. Prepared sample on the mica was placed on the sample stage and 200 μL Milli-Q water was directly applied on the sample for the in-liquid measurement. Obtained AFM images were then analysed with *Gwyddion* (Czech Metrology Institute, Czech)

2.12 Solution state NMR: diffusion NMR spectroscopy - protein aggregation

All diffusion NMR spectroscopy was recorded on an *Agilent Technologies 18.8 T (800 MHz) DD2 Premium Compact* spectrometer with a triple-resonance, 5 mm enhanced cold probe. 1,4-Dioxane was added to all NMR sample, which is often used as a chemical shift standard in protein NMR because it is known to behave independently of protein concentration and the folded state of the protein.^[82–84] All diffusion NMR spectrum was analysed with the *VnmrJ* (Agilent Technologies, Santa Clara, CA, USA) and MATLAB.

⁸² A. Shimizu, M. Ikeguchi, and S. Sugai. *J. Biomol. NMR*, **4**: 859–862, 1994.

⁸³ A. De Marco. *J. Magn. Reson.*, **26**: 527–528, 1977.

⁸⁴ J. Pettersson-Kastberg et al. *J. Mol. Biol.*, **394**: 994–1010, 2009.

Table 2.4: Gradient strength (G/cm) increasing from 1200 to 30000 in 60 increment steps

Array no.	G/cm	Array no.	G/cm	Array no.	G/cm
1	1200	21	7200	41	15600
2	1500	22	7500	42	16100
3	1800	23	7800	43	16600
4	2100	24	8100	44	17100
5	2400	25	8400	45	17600
6	2700	26	8700	46	18100
7	3000	27	9000	47	18600
8	3300	28	9300	48	19100
9	3600	29	9600	49	19600
10	3900	30	10100	50	20600
11	4200	31	10600	51	21600
12	4500	32	11100	52	22600
13	4800	33	11600	53	23600
14	5100	34	12100	54	24600
15	5400	35	12600	55	25600
16	5700	36	13100	56	26600
17	6000	37	13600	57	27600
18	6300	38	14100	58	28600
19	6600	39	14600	59	29600
20	6900	40	15100	60	30000

2.12.1 Pseudo real-time DOSY experiment to monitor the early stage of HEWL aggregation

2.12.1.1 Protein sample preparation

To monitor early stage aggregation of HEWL with pseudo real-time DOSY NMR experiment, HEWL 1 *mM* (pH 2) solution was prepared as described in Section 2.4.1 and divided into five Eppendorf and kept in -80°C and defrosted before the DOSY experiment. A total volume of 600 μL HEWL sample ($\text{H}_2\text{O}/\text{D}_2\text{O}$ 10 % v/v , with 1 *mM* Dioxane) was transferred into the NMR tube.

2.12.1.2 Pseudo real-time DOSY experiment

The Dbppste DOSY pulse sequence was used, which consists of ‘Biopolar Pulse Pair Stimulated Echo’ method. A spectral window of 20 ppm was used, with a relaxation delay of 3 *s*. The FIDs were collected with 32000 complex data points with 16 scans. The gradient pulse strength was increased

Table 2.5: Parameter of pseudo real-time DOSY NMR experiment

Parameter	values
Total number of scans	16
Number of increment	60
Lowest gradient value (G/cm)	1200
Highest gradient value (G/cm)	30000
Diffusion gradient length (δ)	5
Diffusion delay (Δ)	100

from 1200 G/cm to 30000 G/cm in 60 increment steps as shown in Table 2.4. A diffusion delay (Δ) of 100 ms and diffusion gradient length (δ , bipolar half-sine-shaped gradient pulses) of 5 ms was applied (Table 2.5).

With the parameter described, DOSY experiment was performed at temperature 25 °C. After DOSY measurement at 25 °C finished, the temperature of the NMR probe was increased from 25 °C to 65 °C (calibration of NMR probe temperature: Section 2.12.4) and the protein sample was left to be incubated inside of the NMR probe for 6 hours. After 6 hours, the temperature was dropped back to 25 °C and the sample was cooled down and equilibrated for 30 mins, and the DOSY experiments was run three times with the same parameter setting. Following this, the DOSY experiment was run with different incubation time of 12 hours, 18 hours, and 24 hours.

2.12.2 Hydrodynamic radii estimation from fitting diffusion NMR data

Diffusion coefficients, D , of a diffusion species can be calculated with the *Stokes-Einstein equation* (1.1). However, as mentioned in Section 1.4.5.1 diffusion coefficient values are affected by temperatures and solvent conditions, such as concentration and viscosity, which makes it difficult to obtain the absolute diffusion coefficient value. To overcome this problem, the diffusion coefficient of a protein ($D_{protein}$) is measured by use of another molecule as the reference molecule with known hydrodynamic radius present in the same sample.^[30,107] As the protein and the reference compound is present in the same solution. The hydrodynamic radius of the protein is directly inversely

³⁰ J. A. Jones et al. *J. Biomol. NMR*, **10**: 199–203, 1997.

¹⁰⁷ A. Chen, D. Wu, and C. S. Jr. Johnson. *J. Phys. Chem.*, **99**: 828–834, 1995.

proportional to the hydrodynamic radius of the reference compound, because the environmental factors are supposed to be the same for both molecules. The hydrodynamic radius of the protein ($R_h^{protein}$) can be calculated using Equation (2.3).

$$R_h^{protein} = \frac{D_{ref}}{D_{protein}} \times R_h^{ref} \quad (2.3)$$

Where,

$R_h^{protein}$ = Hydrodynamic radius of protein,

R_h^{ref} = Hydrodynamic radius of the reference compound,

$D_{protein}$ = Diffusion coefficient value of protein,

D_{ref} = Diffusion coefficient value of the reference compound

As internal hydrodynamic radius reference compound, 1,2-dioxane, which is known to behave independently of protein concentration and the folded state of the protein was used ($R_h = 2.12 \text{ \AA}^{[98]}$ / ^1H : $3.75 \text{ ppm}^{[80]}$).

Peaks were chosen from the aliphatic region of protein NMR spectra, which is far from the water peaks and the amide region where proton exchange may occur and cause difficulty in D calculation. The peak of the reference compound 1,4-dioxane was also chosen for the following fitting procedure. Chosen peaks were then integrated and the series of integrated intensity value from each series of 1D NMR traces recorded with increasing gradient amplitudes were extracted for the further fitting analysis by using *MATLAB: Curve Fitting Tool*.

The integrated spectral intensity of diffusing species is decaying gradually with the increase of gradient strength, which followed a Gaussian decay function (Equation (2.4)). Fitting process was started from single components and, more components were added to obtain the optimum fitting and examples of mono/poly-Gaussian decay function are shown below.

$$I(g) = I_0 \exp(-Dg^2) \quad (2.4)$$

⁸⁰ G. R. Fulmer et al. *Organometallics*, **29**: 2176–2179, 2010.

Where,

$I(g)$ = the observed intensity at each gradient strength ' g '

D = the relevant diffusion coefficient of the diffusion species

I_0 = the initial intensity

Example:

One component : $I(g) = I_1 \exp(-D_1 g^2)$

Two components : $I(g) = I_1 \exp(-D_1 g^2) + I_2 \exp(-D_2 g^2)$

Three components : $I(g) = I_1 \exp(-D_1 g^2) + I_2 \exp(-D_2 g^2) + I_3 \exp(-D_3 g^2)$

2.12.3 2-D solution NMR spectroscopy

Two-dimensional solution NMR experiments were carried out to determine the 3D structure of 9R-VIPER and 9R-VIPER L6AE10A. A total of 600 μL each of 2 $m\text{M}$ 9R-VIPER and 5 $m\text{M}$ 9R-VIPER L6AE10A were prepared with $\text{H}_2\text{O} / \text{D}_2\text{O}$ 10 % v/v (pH 4.5), 20 $m\text{M}$ deuterated sodium acetate using 0.5 $m\text{M}$ DSS as internal standard. ^1H - ^{15}N HSQC, ^1H - ^{13}C HSQC, TOCSY, NOESY solution NMR experiments were performed at 25 °C on an Agilent Technologies 18.8 T (800 MHz) DD2 Premium Compact spectrometer with a triple-resonance, 5 mm enhanced cold probe and the spectra collected from NMR experiments were used for sequence assignment and NOE-distance restraints calculation.

^1H - ^{15}N HSQC of both 9R-VIPER and 9R-VIPER L6AE10A were acquired with a 1 s relaxation delay (D1) and the acquired data sizes were 512×4096 points in the F_1 and F_2 dimensions while the $J(^{15}\text{N} - ^1\text{H})$ coupling was set at 93. ^1H - ^{13}C HSQC of both 9R-VIPER and 9R-VIPER L6AE10A were acquired with 2 s relaxation delay (D1) and acquired data sizes were 2048×4096 points in the F_1 and F_2 dimensions. TOCSY spectra of 9R-VIPER was acquired with 2 s relaxation delay (D1) and 80 ms spin lock time and dipsi2 spin lock pattern was used. The acquired data sizes were 2048×4096 points in the F_1 and F_2 dimensions. TOCSY spectra of 9R-VIPER L6AE10A was acquired with 1 s relaxation delay (D1) and 0.12 ms spin lock time and dipsi2 spin lock pattern was used. The acquired data sizes were 1024×4096 points in the F_1 and F_2 dimensions. NOESY spectra of 9R-VIPER was acquired with 2 s relaxation delay (D1) and mixing time of 300 ms, and the acquired

data size was 2048×4096 points in the F_1 and F_2 dimensions. NOESY spectra of 9R-VIPER L6AE10A was acquired with 1 s relaxation delay (D1) and 0.12 ms spin lock time and the acquired data size was 1024×2048 points in the F_1 and F_2 dimensions.

Other data were collected with different mixing time setting but not used for chemical shift assignment or NOE data. NMR data were processed and analysed with the *VnmrJ* version 4.2 (Agilent Technologies, Santa Clara, CA, USA) and further assignments were processed with *Collaborative Computing Project for the NMR community (CCPN)* software (Section 2.12.3.2).^[108]

2.12.3.1 NMR chemical shift assignment & 3-D model generation

Chemical shift assignments, NOE distance restraints, and dihedrals angle prediction were carried out using CCPN as described in previous section.^[108] Assignment panel reference table was used to compare identified chemical shifts at pH 4.5, and temperature of 25 °C.^[109] Atom types included in ^1H - ^{15}N HSQC, ^1H - ^{13}C HSQC, TOCSY, and NOESY were HN, HA, HB1, HB2, CA, CB, C, and N. Sequence data, NMR chemical assignments and calculated NOE-based distance restraints of 9R-VIPER and 9R-VIPER L6AE10A (VIPER mutant) were analysed from CCPN, and used as input data to generate 3D structure of protein through GeNMR (Generate NMR structure) web server (Section 2.12.3.3). The web server produces an ensemble of PDB coordinates for the protein, depending on model complexity and completeness of experimental restraints. CCPN provides three methods of volume calculation and ‘truncated box sum’ has been used in this study.^[110]

2.12.3.2 Collaborative computational project for NMR (CCPN)

The collaborative computational project for NMR (CCPN) is a graphics-based interactive NMR spectrum visualisation, resonance assignment and data analysis program. CCPN has a user-friendly graphic interface and enables reading of spectra from multiple data formats including Agilent/Varian, Bruker, NMRPipe, and NMRView, without the need for a file format conversion.^[108] In this report, CCPN was used for the assignment of NMR spectrum including ^1H - ^{15}N HSQC, ^1H - ^{13}C HSQC, TOCSY, and NOESY of 9R-VIPER and 9R-VIPER L6AE10A (VIPER mutant). All NMR

¹⁰⁸ W. F. Vranken et al. *Proteins Struct. Funct. Bioinforma.*, **59**: 687–696, 2005.

¹⁰⁹ D. S. Wishart et al. *J. Biomol. NMR*, **5**: 67–81, 1995.

¹¹⁰ M. Berjanskii et al. *Nucleic Acids Res.*, **37**: W670–W677, 2009.

spectra obtained from Agilent NMR 800 *MHz* (Agilent Technologies, Santa Clara, CA, USA) in Varian format were uploaded, and sequence of peptides was added. Based on the input sequence, assignment panel was created for each peptide and, chemical shifts were assigned based on ^1H - ^{15}N HSQC, ^1H - ^{13}C HSQC, and TOCSY spectrum, and NOE distance restraints were assigned with NOESY spectrum. The NOE signal intensity obtained from NOESY is directly related to the inter-distance between protons. In CCPN, NOE intensities are classified into strong (1.72–2.56 Å), medium (1.72–3.84 Å), and weak (1.72–5.5 Å).

2.12.3.3 GeNMR (generate NMR structure)

GeNMR (Generate NMR structure) is a web server for generating 3D protein structure using sequence data, NMR chemical shifts and NOE-derived distance restraints as input. The web server then produces an ensemble of PDB coordinates for the protein.^[110] In this report, the 3D structure of 9R-VIPER and 9R-VIPER L6AE10A (VIPER mutant) were generated through GeNMR. Once assignments were completed in CCPN, the obtained data was then exported to Microsoft excel or Microsoft notepad for reformatting into GeNMR uploadable format. The sequence of the peptide in FASTA format, NMR chemical shifts in BMRB NMR-STAR 2.1 format, and NOE&H-bond distance restraints in CYANA format were uploaded to the server. Generating the ensemble of PDB coordinates were completed within 15-30 mins depending on the size of input data, and the obtained 3D models were then visualised on PyMOL (Schrödinger, NY, USA).

2.12.4 NMR temperature calibration using 1,2-ethanediol

This calibration is required at the high temperature, especially when temperature value showing at the control panel and actual temperature of the probe have differences. Calibrating the temperature of NMR was carried out with 1,2-ethanediol sample provided in NMR tube from Agilent, NMR manufacturer.

Once 1,2-ethanediol was placed in the NMR probe set to 25 °C, the NMR is tuned and matched and 1D Proton spectra acquired. In the acquired spectra, 1,2-ethanediol showed two peaks and delta ppm between two peaks was measured which will be used to calculate the real temperature of the NMR probe with Equation (2.5). Increased 5 degrees from 25 °C to 60 °C and increased

1 degree from 60 °C to 70 °C and once the temperature increased, a minimum wait time of 10 mins is required for the sample to equilibrate. Re-shimmed each time after equilibration to new temperature was completed and repeated the same steps acquired at 25 °C.

$$T[K] = -108\Delta\delta + 460.41 \quad (2.5)$$

Where,

$T[K]$ = Real temperature of NMR probe in the Kelvin scale,

$\Delta\delta$ = ppm difference between two peaks obtained from 1,2-ethanediol

2.13 Bovine β -Lg heat stability experiment preparation

Lyophilised β -Lg powder was dissolved in dH_2O / D_2O (90 % / 10 %) overnight (+10 hours) to allow complete hydration of the protein. The hydrated β -Lg solution was centrifuged at 3900 *rpm* for 10 min at 4 °C using a Centrifuge 5810R (Eppendorf), and the supernatant was filtered with 0.22 μm syringe filter to remove any insoluble particles. Filtered β -Lg solution was amended to yield at 1 % w/v (10 mg/mL) or 12 % w/v (extinction coefficient of 1 % β -Lg = 9.63) and final concentration was checked with NanoDrop. Once concentration of protein solution is finalised, the prepared samples were then incubated depending on the experiment requirement at 62 °C (20 hours) or 85 °C (5 mins). Dioxane is then added as internal chemical shift standard (3.75 ppm).

^1H - ^{13}C HSQC solution NMR experiments were performed at 20 °C on Agilent NMR 800 *MHz* with triple resonance cold probe and spectra collected from NMR experiments were then analysed with VnmrJ (Agilent Technologies, Santa Clara, CA, USA). ^1H - ^{13}C HSQC (gHSQCAD pulse sequence) experiments with 512 increment were acquired with 2 *s* relaxation delay (D1) and acquired data sizes were 2048 $\times$ 4096 points in the F_1 and F_2 dimensions. All spectra were calibrated relative to the dioxane peak in both dimensions followed by baseline and phase correction.

Chapter 3

Characterising Protein Oligomerisation and Aggregation

3.1 Introduction

As mentioned previously in Section 1.2, protein aggregation has been associated with various human diseases. Aggregation of proteins can occur via different pathways that lead to different conformational outcomes, such as amyloid fibrils and its intermediate form oligomers. Both amyloid fibrils and oligomeric species are considered to be species that are cytotoxic in neurodegenerative diseases^[9,12] showcasing different cytotoxic mechanisms.^[14] Also, amyloid fibrils have shown their potential to be nano-materials for various applications such as bio-nanodevices, bio-nano-sensors, nano-electronics, and energy harvesting devices. Functionalisation of fibrils can be achieved via different approaches such as, use of fibrils as a platform for other functional groups, or by producing a hybrid nanostructure of protein and functional groups.^[47,48] Therefore, understanding the morphology, physical and chemical characteristics of both amyloid fibrils and oligomeric species is important not only for disease related research, but also for future applications as a nano-bio material.

⁹ F. Chiti and C. M. Dobson. *Annu. Rev. Biochem.*, **75**: 333–366, 2006.

¹² R. M. Murphy. *Biochim. Biophys. Acta - Biomembr.*, **1768**: 1923–1934, 2007.

¹⁴ N. P. Harte et al. *J. Biol. Chem.*, **290**: 28343–28352, 2015.

⁴⁷ T. Aumüller and M. Fändrich. *Nat. Chem.*, **6**: 273–274, 2014.

⁴⁸ C. Li, J. Adamcik, and R. Mezzenga. *Nat. Nanotechnol.*, **7**: 421–427, 2012.

In this chapter, HEWL and bovine β -Lg were used as model proteins to study the aggregate properties of protein. Characterisation of amyloid aggregates as well as early stage protein aggregation were investigated by diffusion NMR combines using different techniques.

A variety of methods are used to determine the presence and kinetics of amyloid fibril formation. Of these, the ThT fluorescence assay and CR assay are the most well-known methods. However, these have limitations in distinguishing the presence of smaller oligomers or protofibrils from amyloid fibril/mature amyloid fibril. Therefore, further classification is required using other techniques to determine whether one is dealing with mature amyloid fibrils or amyloid oligomers or protofibrils.^[58,60,62]

Amyloid fibril is rich in β -sheet, hence any secondary structural transition that occurs during fibril formation can be observed by far-UV CD, FTIR spectroscopy, and NMR imaging.^[57] Since amyloid fibrils are insoluble and heterogeneous structure determination is difficult, TEM and AFM are often used for the conformation of fibril formation, and to study the morphology and physical characteristics of formed fibrils.^[59,111]

3.1.1 Investigation of protein aggregates with AFM

Atomic force microscopy (AFM) was invented in 1986^[112] and ever since it has been applied in many biological studies to characterise amyloid fibrils and investigate its morphology and physical properties.^[33,40,54,113–115] The benefit of AFM is that it can be used in a liquid environment which is essential in studying proteins, protein aggregates and amyloid fibrils. The operating principle of

⁵⁸ H. Levine III. *Protein Sci.*, **2**: 404–410, 1993.

⁶⁰ H. Levine III. Quantification of β -sheet amyloid fibril structures with thioflavin T in: *Methods in Enzymology: Amyloid, Prions, and Other Protein Aggregates*. ed. by R. Wetzel. **309**: chap. 18, 274–284, Elsevier, 1999. DOI: [10.1016/S0076-6879\(99\)09020-5](https://doi.org/10.1016/S0076-6879(99)09020-5)

⁶² M. R. H. Krebs, E. H. C. Bromley, and A. M. Donald. *J. Struc. Biol.*, **149**: 30–37, 2005.

⁵⁷ L. Hou et al. *J. Am. Chem. Soc.*, **126**: 1992–2005, 2004.

⁵⁹ J. Peng et al. *Sci. Rep.*, **5**: 10171, 2015.

¹¹¹ M. Reches and E. Gazit. *Nat. Nanotechnol.*, **1**: 195–200, 2006.

¹¹² G. Binnig, C. F. Quate, and Ch. Gerber. *Phys. Rev. Lett.*, **56**: 930–933, 1986.

³³ V. Sivalingam et al. *Biophys. Chem.*, **219**: 28–37, 2016.

⁴⁰ J. Adamcik et al. *Nat. Nanotechnol.*, **5**: 423–428, 2010.

⁵⁴ R. Jansen, W. Dzwolak, and R. Winter. *Biophys. J.*, **88**: 1344–1353, 2005.

¹¹³ F. S. Ruggeri et al. *Nat. Commun.*, **9**: 3890, 2018.

¹¹⁴ J. Adamcik and R. Mezzenga. *Curr. Opin. Colloid Interface Sci.*, **17**: 369–376, 2012.

¹¹⁵ B. Moores et al. *PLOS ONE*, **6**: e25954, 2011.

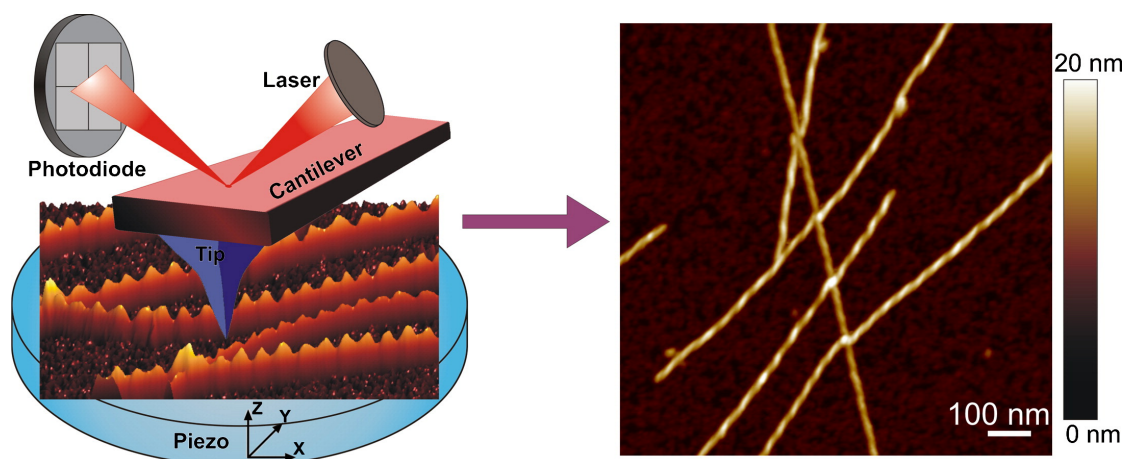


Figure 3.1: Operating principle of AFM and AFM image of amyloid fibrils. AFM works by sensing the forces between the sample surface and a cantilever tip. The sample is mounted on a piezo-tube scanner that moves in the X-Y-Z direction. The laser beam is reflected at the top of the cantilever and a position-sensitive photodiode detects the changes in reflection due to the sample surface. (Adapted from Adamcik J. *et al.*^[114])

AFM is shown in Figure 3.1.

3.1.2 Biomolecular diffusion NMR

Diffusion NMR gives information of physical molecular properties including molecular size, shape, hydrogen bonding, and aggregation. It especially shows the correlation between diffusion coefficient values and the size of the diffusing species.

In this chapter, diffusion NMR has been extensively applied in the investigation of the early stage protein aggregation, and for the determination of protein hydrodynamic radii. As mentioned in Section 1.4.5.1, the diffusion coefficient is affected by various parameters, including solvent viscosity, temperatures and the concentration of the sample. Due to these factors, measurements of absolute values of the diffusion coefficient is difficult. This issue can be overcome if the diffusion coefficient of a diffusion species is measured relative to the diffusion coefficient of the reference molecule with known hydrodynamic radius present in the same sample. With this method, the hydrodynamic radius of the protein ($R_h^{protein}$) is directly inversely proportional to the hydrodynamic radius of the reference molecule (R_h^{ref}), as other factors affecting sample environments are considering to be the same for all diffusing compounds in sample. In this thesis, 1,4-dioxane with known hydrodynamic radius of 2.12 Å is used as the reference compound, as prior investigations show that

it does not interact with the protein. The diffusion coefficient value of the molecule in question, and 1,4-dioxane are obtained via Gaussian decay function (Equation (2.4)). Once the diffusion coefficient values are found, the $R_h^{protein}$ can be calculated with the Equation (2.3) as described in Section 2.12.2.

3.1.3 Aims

Amyloid fibril growth is preceded by the formation of a wide range of aggregates such as unstructured oligomer and structured protofibril. Amyloid fibril has been extensively studied due to its association with neurodegenerative disease as well as due to its potential as a future nano-biomaterial. However, amyloid fibril properties at the molecular level, especially the structure of the oligomers and amyloid protofibrils still requires detailed studies even though they have been identified as the key player in the pathogenesis of protein deposition diseases. Biophysical techniques are only able to provide average information about the population ensemble without offering any observations of the individual population presence in a heterogenous sample.

Here, we characterised amyloid fibril formation process with the model protein HEWL and β -Lg. We also present diffusion NMR along with biophysical techniques such as AFM and SDS-PAGE as a supporting methodology for determining the hydrodynamic radii of the larger biomolecules such as proteins and protein oligomers/aggregates, which may applicable for unveiling amyloid fibril transient species.

This chapter aims to:

- Investigate protein amyloid formation with ThT, Congo-Red, CD, TEM, AFM, and diffusion NMR techniques.
- Synthesise the nanostructured film with HEWL and visually confirm the presence of the fibril in the film. Also, evaluate the kinetics of protein aggregation/fibrillisation during the synthesis.
- Diffusion NMR was applied to HEWL-Tyr-Dioxane mixture to show its capability of separating them base on their hydrodynamic radius.
- Synthesise cross-linked HEWL oligomer mixture with various hydrodynamic radius (R_h) and

estimate and verify the R_h of each component of the cross-linked oligomer mixture with SDS-PAGE and SE-HPLC (data not shown in this chapter, but available in the publication section (A.7)). Following this, we will calculate the R_h of the components based on the diffusion NMR data and then compare it with the estimated R_h value to establish diffusion NMR as a proof-of-principle methodology for the protein size determination.

- Apply diffusion NMR to determine the ratio between the monomer and dimeric β -Lg with acidic & neutral pH at low and high protein concentration.
- Apply diffusion NMR to monitor early stage aggregates of the HEWL and estimate the R_h of oligomers/aggregates formed during the aggregation process.

3.2 Result & Discussion

3.2.1 HEWL amyloid fibril formation

3.2.1.1 ThT fluorescence measurement

The HEWL protein in pH 2 incubated at 65 °C has been shown to convert the native protein into a fibrillary form.^[14,22] The thiazine dye, Thioflavin T (ThT) associates rapidly with β -sheet in the aggregated proteins. Due to this property, the ThT fluorescence assay has been used for the detection of the formation of amyloid fibrils.^[60] HEWL 1 *mM* was incubated in pH 2 for 21 days at 65 °C and ThT fluorescence was measured as described in the Section 2.5. Incubated HEWL showed a gradual increase in ThT emission over 21 days as ThT indicating more β -sheet formed along formation of fibrils. In a typical growth profile of a wide range of protein fibrillation phenomena in vitro, an initial lag phase is usually observed, during which no measurable aggregation occurs and the effective growth rate increases significantly until it reaches the inflection point of the saturation phase.^[11,55] HEWL fibril formation showed that there is no fibrillisation that occurred between day 0 and day 4, which can be referred to being the lag phase. After the lag phase, the growth phase followed between day 4 and day 10, and between this period, exponential fibril growth was observed. This significant increase finally reached the inflection point at day 10 and entered into

²² M. R. H Krebs et al. *J. Mol. Biol.*, **300**: 541–549, 2000.

¹¹ C. Iannuzzi et al. *Int. J. Mol. Sci.*, **14**: 14287–14300, 2013.

⁵⁵ J. E. Gillam and C. E. MacPhee. *J. Phys. Condens. Matter*, **25**: 373101, 2013.

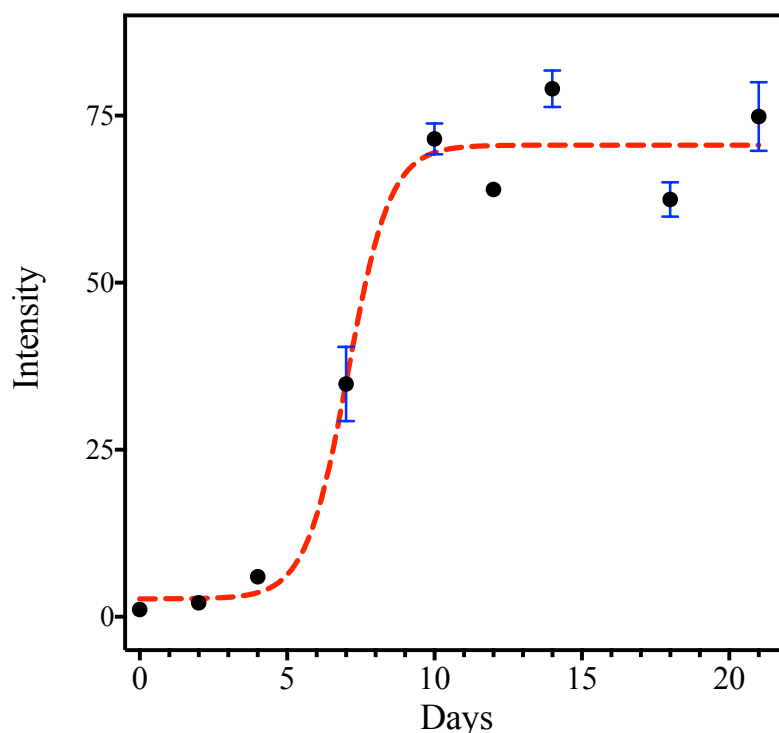


Figure 3.2: ThT fluorescence spectra of HEWL protein. Fluorescence intensity of 1 *mM* HEWL (pH 2) incubated at 65 °C was measured by excitation at 440 *nm* and emission at 482 *nm* for 21 days. At each time points, three ThT measurements were performed.

the saturation phase (Figure 3.2).

3.2.1.2 Congo red binds amyloid fibrils

Congo Red (CR) binds to fibrillary β -sheets that result in birefringence – a yellow/apple-green colour – when visualised using 90 ° cross polarisers.^[19,103,104,116] The presence of HEWL fibril was also detected using a CR birefringence assay. HEWL samples were incubated for fibrillisation (Section 2.4.1) and CR birefringence is performed for visual conformation of the HEWL fibrils by staining the sample with CR stain solution under a polarised light microscope using the procedure described in Section 2.6.1.

HEWL incubated for 7 days showed short fibril-like structure (Figure 3.3:A), with prefibrillar

¹⁹ M. Nilsson. *Methods*, **34**: 151–160, 2004.

¹⁰³ A. J. Howie et al. *Lab. Investig.*, **88**: 232–242, 2008.

¹⁰⁴ R. Khurana et al. *J. Biol. Chem.*, **276**: 22715–22721, 2001.

¹¹⁶ H. Puchtler, F. Sweat, and M. Levine. *J. Histochem. Cytochem.*, **10**: 355–364, 1962.

formation detected on day 11 (Figure 3.3:B). Mature fibrils are then seen on day 15 (Figure 3.3:C) with the presence of twists among distinct fibrils. This result demonstrates that fibrillisation occurs when HEWL is incubated at 65 °C and mature fibril is formed in 15 days. This result also supports the ThT assay of HEWL, whereby HEWL fibrillisation enters into the saturation phase after day 11 (Figure 3.2). In Figure 3.3: A2, B2, and C2 fibrils detected under 90 ° cross polarisers showed yellow-orange colours as opposed to a bright apple-green colour described in the literature.^[19,103,104,116]

3.2.1.3 Simultaneous measurement of aggregation with diffusion NMR (DOSY)

HEWL was used as a model system to investigate fibril formation with diffusion NMR, which allows for the monitoring of the early stage fibril formation process. Diffusion NMR measurements were carried out with native HEWL protein sample as control at day 0, and once the incubation of HEWL protein sample in the heat block (at 65 °C) started, a total of four diffusion NMR experiments were performed with HEWL incubated for 3, 8, 11, and 21 days with the parameter setting described in Section 2.12.

Diffusion NMR provides an assignment of the diffusion correlation peaks associated with the molecular size present in the mixture sample. The diffusion coefficient of a molecule depends on its molecular weight, size, and shape under specific environment such as solvents and temperature, and hence it is used to study the intermolecular interaction in both chemical and biological research.^[92] Based on the relation between translational diffusion rates and the size of particles, we attempt to utilise diffusion NMR to obtain simultaneous measurements of HEWL protein fibrillisation, as well as to achieve a size exclusion chromatography-like effect with different size of proteins.^[94,95]

Figure 3.4 is the overlaid spectrum of five NMR spectra obtained at day 0, 3, 8, 11 and 21 showing different areas of the spectra. The intensity, shape and chemical shift of peaks offers a multitude of information about the protein such as structural conformation, identification, and concentration. The intensity decrease of the peaks mean that the part of the protein which gives specific chemical shift (ppm) has had its environment changed and has lost its structure, hence the areas where purple arrows are pointing indicate that HEWL gradually aggregates during incubation

⁹²Y. Cohen, L. Avram, and L. Frish. *Angew. Chemie Int. Ed.*, **44**: 520–554, 2005.

⁹⁴K. F. Morris and C. S. Johnson. *J. Am. Chem. Soc.*, **114**: 3139–3141, 1992.

⁹⁵K. F. Morris and C. S. Johnson. *J. Am. Chem. Soc.*, **115**: 4291–4299, 1993.

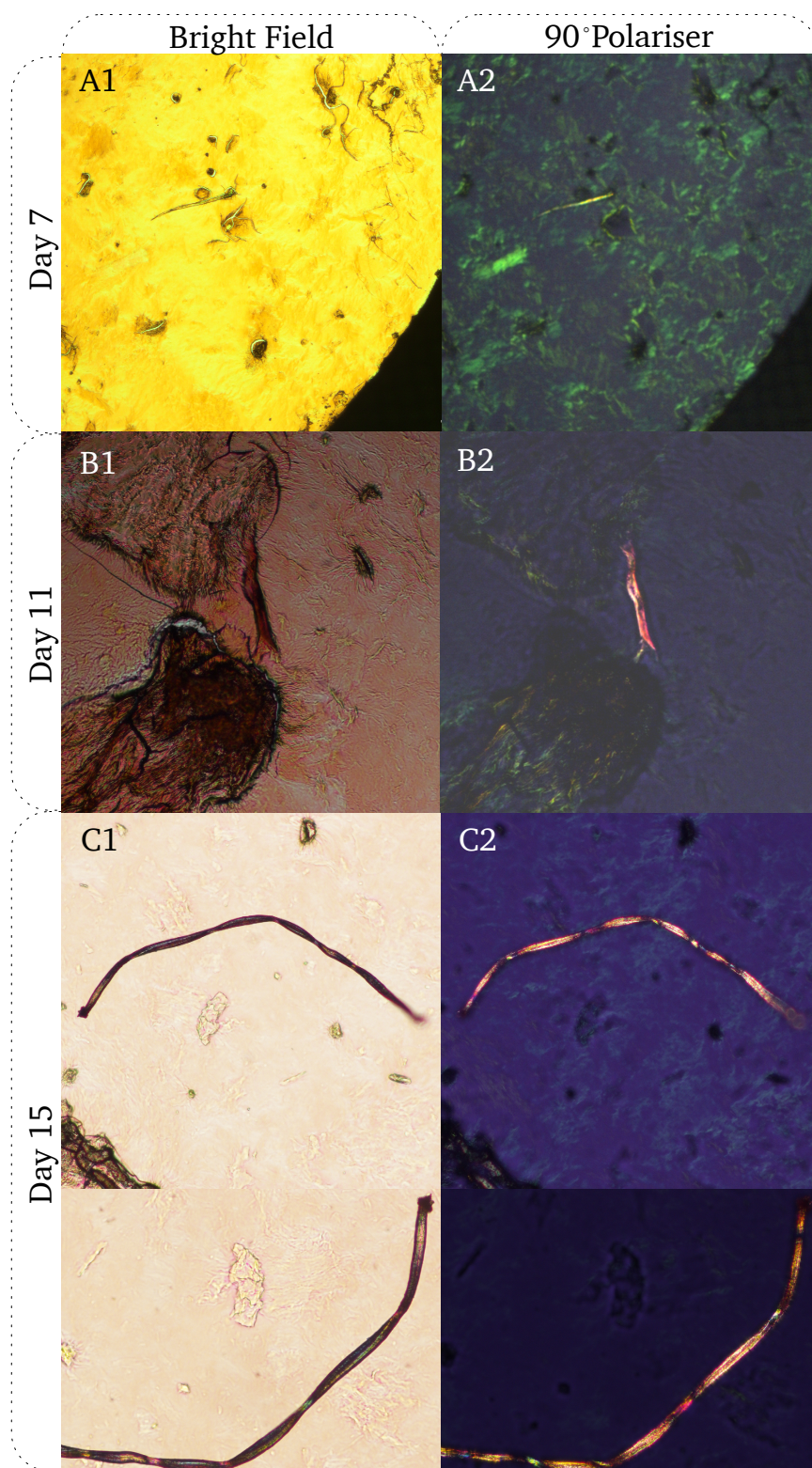


Figure 3.3: CR binds to HEWL at different states of fibrillisation. CR birefringence assay of HEWL. (A) HEWL incubated for 7 days, (B) Formed prefibrillar detected on day 11, (C) Mature fibrils shown on day 15. Images A1~C1 were analysed without cross polarisers and images A2~C2 were with cross polariser at 90°.

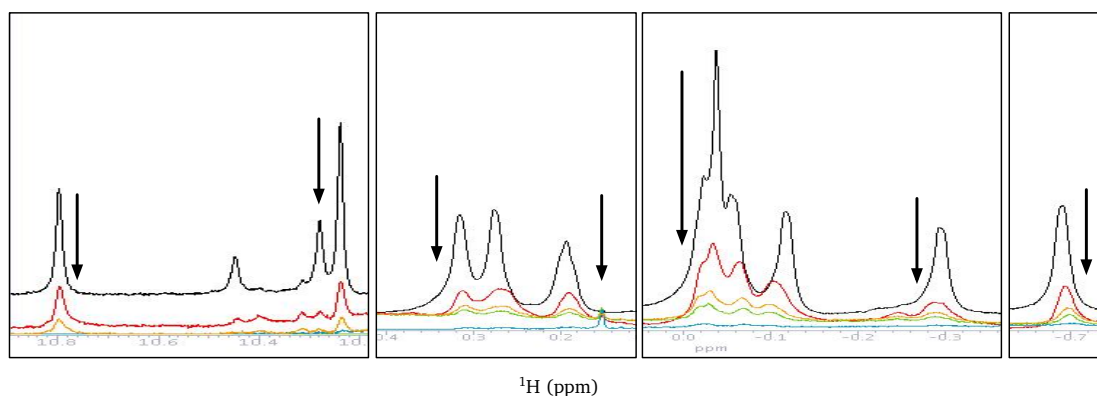


Figure 3.4: Different areas of the overlaid 1D NMR spectrum of the HEWL protein. Different areas of the overlaid 1D NMR spectra before the DOSY processing shows the intensity decrease. This indicates that HEWL protein is losing the structure due to aggregation occurring during the incubation period at 65 °C. (A: Day 0, B: Day 3, C: Day 8, D: Day 11, E: Day 21)

at 65 °C for 21 days.

Obtained DOSY data in 1D NMR form were then further processed to DOSY spectrum with the diffusion coefficient as the y-axis and the chemical shift value (ppm) as the x-axis (Figure 3.5). When the particle size is larger, the diffusion coefficient value is lower, and based on the relationship between the size of the particle and the diffusion coefficient, the signals obtained from each particle are separated into different rows of 2D matrix DOSY.^[94,95] The DOSY spectra shows a gradual decrease in diffusion coefficients during the course of the fibril's formation. There is no significant decrease in diffusion coefficient from Day 0 (Figure 3.5:A) and Day 3 (Figure 3.5:B). Furthermore, significant decrease of the diffusion coefficient was observed from Day 3 (Figure 3.5:B) to Day 8 (Figure 3.5:C), which is expected as Day 8 falls in the middle of the growth phase of the HEWL fibrillation. Once the fibril formation is completed on Day 11 (Figure 3.5:D), the diffusion coefficient value did not decrease on Day 21 (Figure 3.5:E), but more noise on the spectra was noticed, which is an expected result caused by the aggregation of HEWL. This result is coordinating with the trend shown in the HEWL ThT fluorescence assay data (Figure 3.2).

3.2.1.4 Tracing the early stage HEWL amyloid fibrilization with AFM

AFM was used to follow the early stages of HEWL amyloid fibrilization at different time points during 6 days of incubation. Size distribution of protein aggregates after one, two, and three days of incubation, as well as AFM images of surface-absorbed HEWL were acquired in liquid (Figure

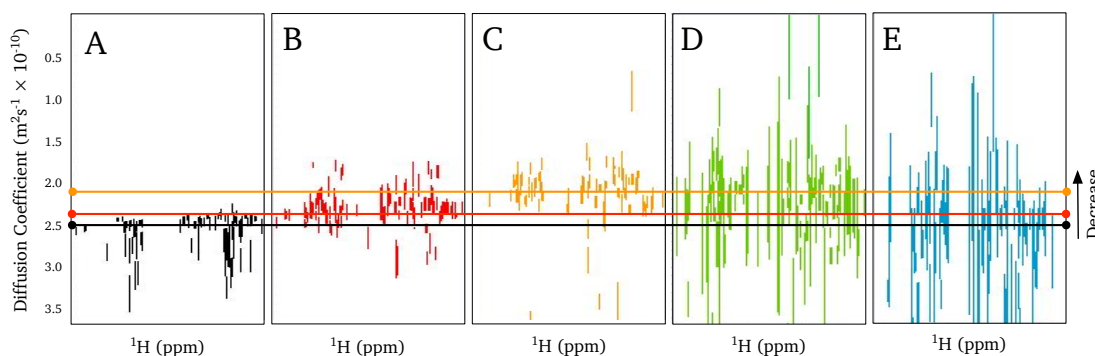


Figure 3.5: DOSY spectra of different stages of HEWL aggregation during 21 days of incubation. (A: Day 0, B: Day 3, C: Day 8, D: Day 11, E: Day 21). Diffusion coefficients decreasing trend is shown from Day 0 (A), Day 3 (B) to Day 8 (C), but once the fibril formation is completed on Day 11 (D), diffusion coefficient value did not change on Day 21 (E). However, more noise on the spectra was noticed, which is an expected result due to the aggregation of HEWL.

3.6). During three days of incubation oligomerization but not fibrillation occurred, as shown by the particle size distribution (Figure 3.6:A), which shows an increase in the hydrodynamic radii over time. The AFM image of HEWL incubated for six days acquired in liquid as well as the size distribution of protein aggregates after one, two, three & six days of incubation are presented in Figure 3.7. The presence of amyloid protofibrils and much larger population of aggregates arising during the additional of 3 days of incubation time was observed.

Also we observed that the overall size increase during the second three days of incubation was more significant compared to the size increase that occurred during the first three days as we expected from the ThT (Figure 3.2). As observed from the ThT, the first three days of incubation fall within the lag phase (during which aggregation occurs slowly), however, the additional three days of incubation fall between the end of lag phase and the start of the growth phase, during which larger oligomers and protofibrils are formed. The location of various cross-sections of protofibrils formed after 6 days of incubation is shown in Figure 3.8 with their height profile. The height of protofibrils were between 3.2~6.5 nm.

3.2.2 Nanostructured films from amyloidogenic proteins (with HEWL)

As mentioned in Section 1.2.3, amyloid fibril has shown its potential to be a nano-biomaterial. In this part of the thesis, nanostructured film were synthesised from HEWL protein by incubating at 65 °C

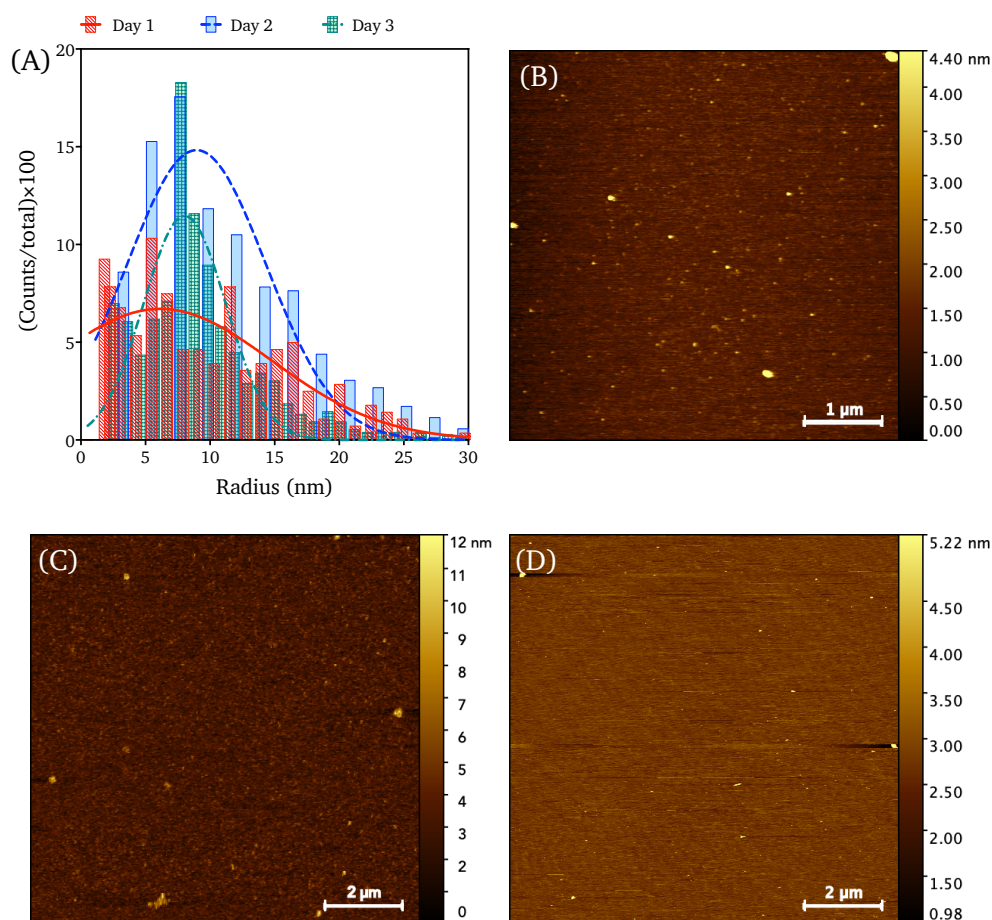


Figure 3.6: AFM images obtained from 1 *mM* HEWL (pH 2) incubated at 62 °C for one, two, and three full days. (A) Particle radius (*nm*) distribution of HEWL after one, two, three days of incubation. AFM image of HEWL incubated for one (B), two (C), and three days (D).

(Section 2.4.1).^[101] Incubation took a total of 21 days including, preparation of the HEWL fibril seeding solution (14 days), and incubating the HEWL protein with prepared seeding solution (7 days). ThT fluorescence assay and dynamic light scattering detection are simultaneously performed to probe fibril formation and aggregation during the incubation period.^[70] After the incubation step is completed, glycerol was added to the hydrogel-film solution and transferred onto a flat surface for casting. Casted films are then stained with Congo-Red and visualised under optical microscopy with polarised light microscopy.^[103,116]

¹⁰¹ T. P. J. Knowles et al. *Nat. Nanotechnol.*, **5**: 204–207, 2010.

⁷⁰ A. M. Streets et al. *PLOS ONE*, **8**: e54541, 2013.

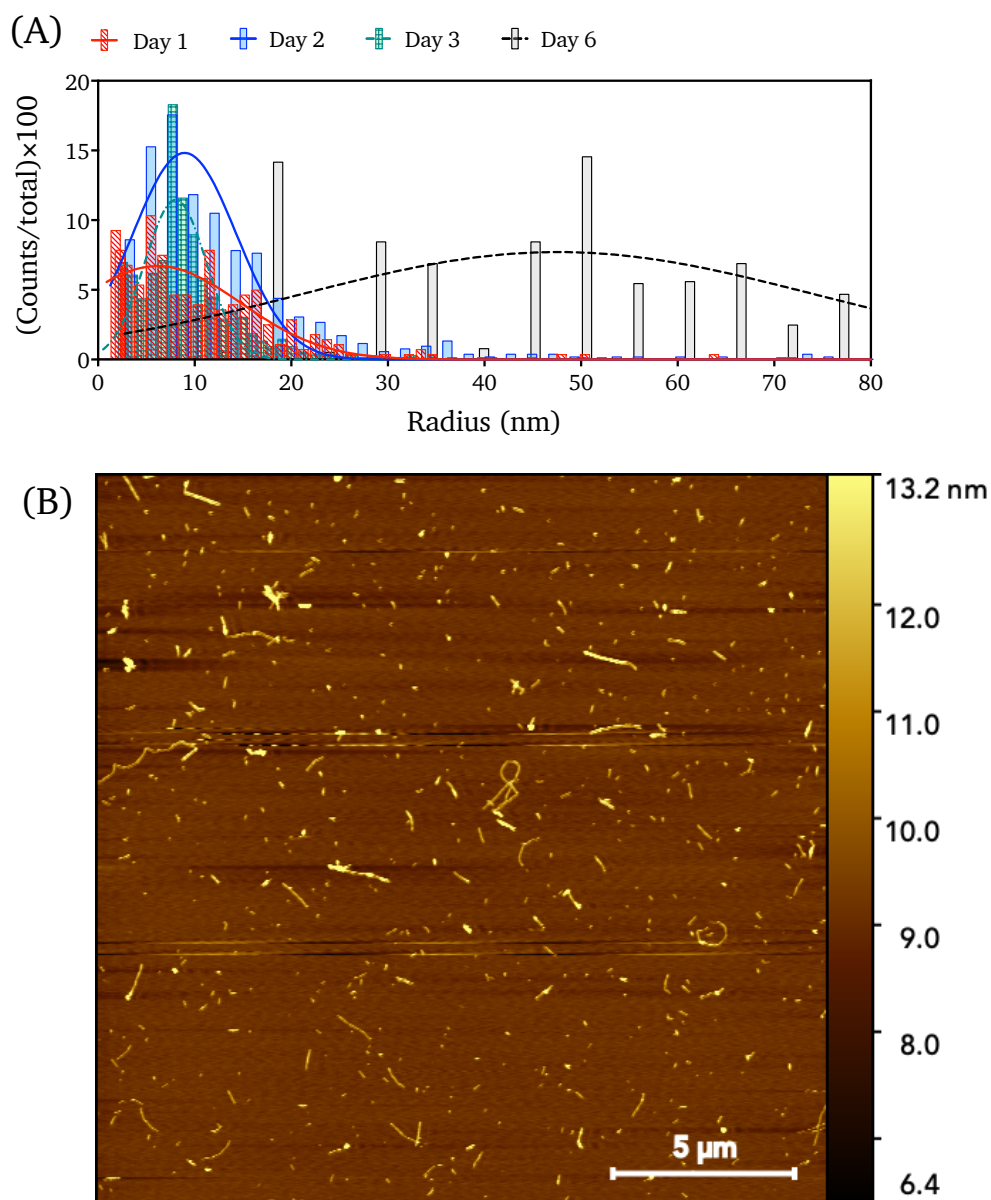


Figure 3.7: AFM result of HEWL 1 *mM* (pH 2) incubated at 62 °C for six days. (A) Particle radius (*nm*) distribution of HEWL (B) AFM image of HEWL incubated for six days.

3.2.2.1 Kinetic measurement with ThT fluorescence assay

The HEWL 3 % *w/w* protein in 10 *mM* HCl is incubated at 65 °C for 14 days to prepare the seed solution, and this prepared seed solution is then added to the freshly prepared HEWL 3 % *w/w* solution with a concentration of 5 % *v/v*. ThT fluorescence was measured at 482 *nm* with excitation at 450 *nm* during 21 days of incubation (Section 2.5). Figure 3.9 shows the HEWL fibril formation

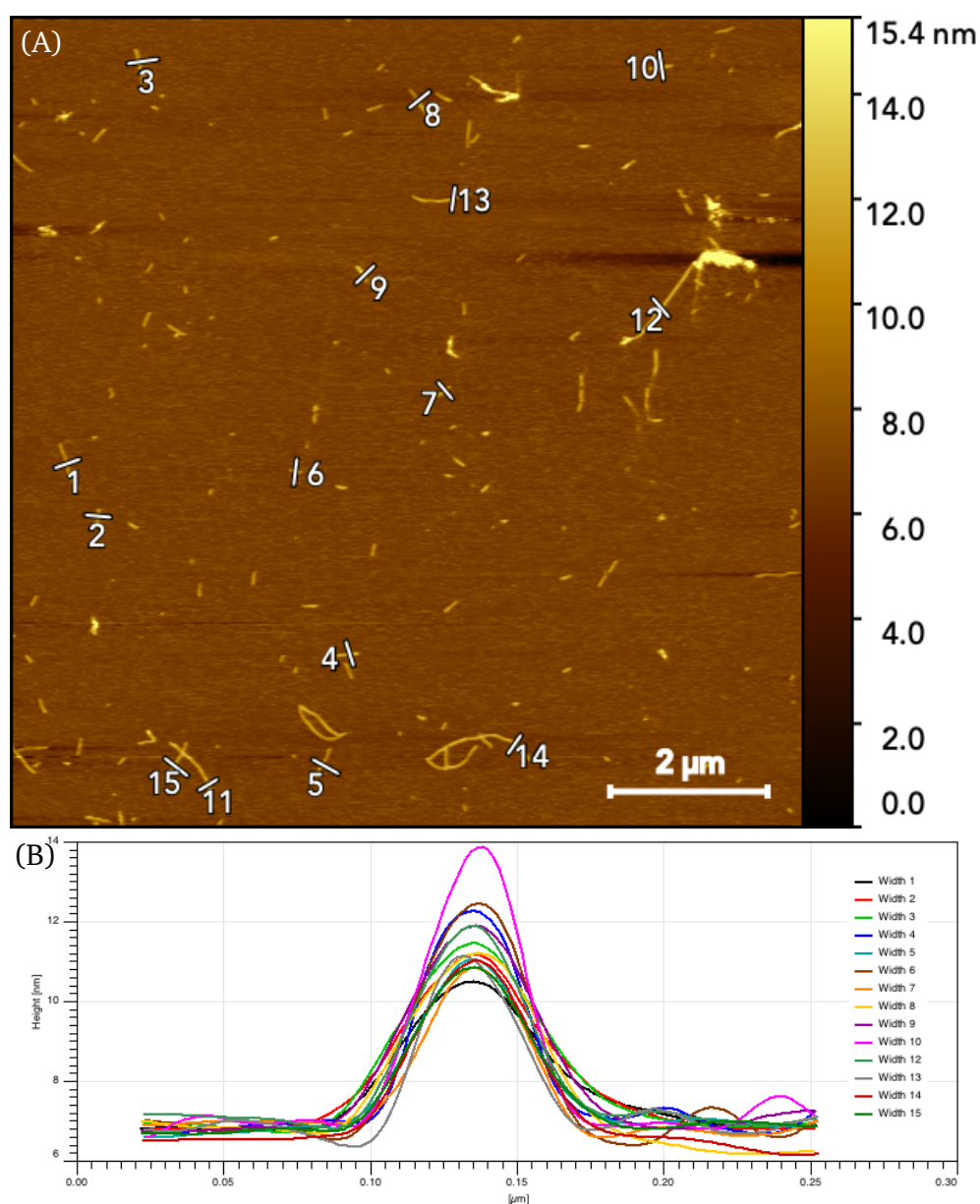


Figure 3.8: Structural analysis of HEWL after 6 days of incubation with AFM. (A) AFM image of HEWL aggregate with location of cross-sections. (B) Height profile of various sections of different protofibrils.

during 14 days and is referred to as the pre-seeding stage. The fibril formation of HEWL with the addition of the seed solution prepared at pre-seeding stage referred to as post-seeding stage (HEWL 3 % w/w with 5 % v/v seed) also shown in Figure 3.9. Pre-seeding stage showed no significant increase in ThT emission during day 0 and day 4 (lag phase), and at day 14 ThT emission has been increased which is an expected figure due to the ThT binding with β -sheets in generated fibrils. Post-seeding stage on the other hand showed a gradual increase in ThT over 7 days of incubation

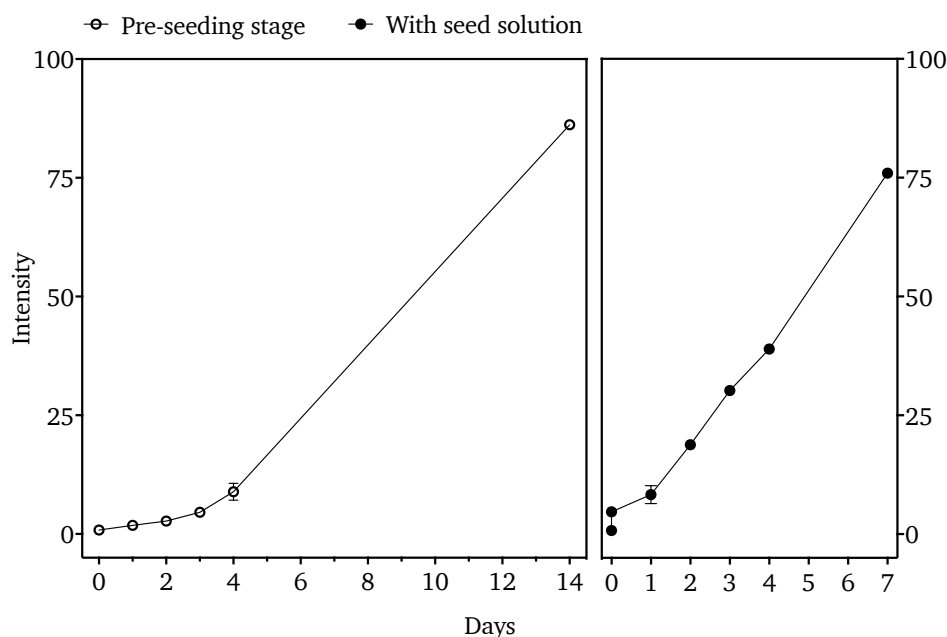


Figure 3.9: ThT fluorescence of HEWL without and with seeding solution. Pre-seeding stage (left): Fluorescence intensity of HEWL 3 % (w/w), HCl 10 mM incubated at 65 °C was measured at 482 nm with excitation at 450 nm , for 14 days. Green arrow: Prepared HEWL fibril solution at Day 14 is added to freshly made HEWL 3 % (w/w), HCl 10 mM protein sample. Post-seeding stage (right): Fluorescence intensity of HEWL 3 % (w/w) with 5 % (v/v) seed was measured at 482 nm with excitation at 450 nm , during 7 days of incubation at 65 °C.

without the lag phase, even though initial HEWL solution used for both steps were prepared under the sample condition and incubated at the same temperature. This suggest that the addition of fibril seed encourages the fibrillisation of protein.

3.2.2.2 Particle size measurement using DLS

DLS is an optical analytical method to determine the size distribution of particles in the nano-meter (nm) to micro-meter (μM) size range and possesses the potential to provide growth kinetics. Based on this competence, DLS has been used in numerous investigations in the past!^[70–73]

In this chapter, DLS is used to detect the particle size increase which occurs due to the fibrillisation of HEWL as a complimentary technique to ThT florescence assay method. This method is utilised to detect the particle size increase that the ThT florescence assay cannot detect!^[70] The

⁷¹ E. Boisselier et al. *J. Am. Chem. Soc.*, **132**: 2729–2742, 2010.

⁷² S. E. Hill et al. *Biophys. J.*, **96**: 3781–3790, 2009.

⁷³ J. Camponovo et al. *J. Org. Chem.*, **74**: 5071–5074, 2009.

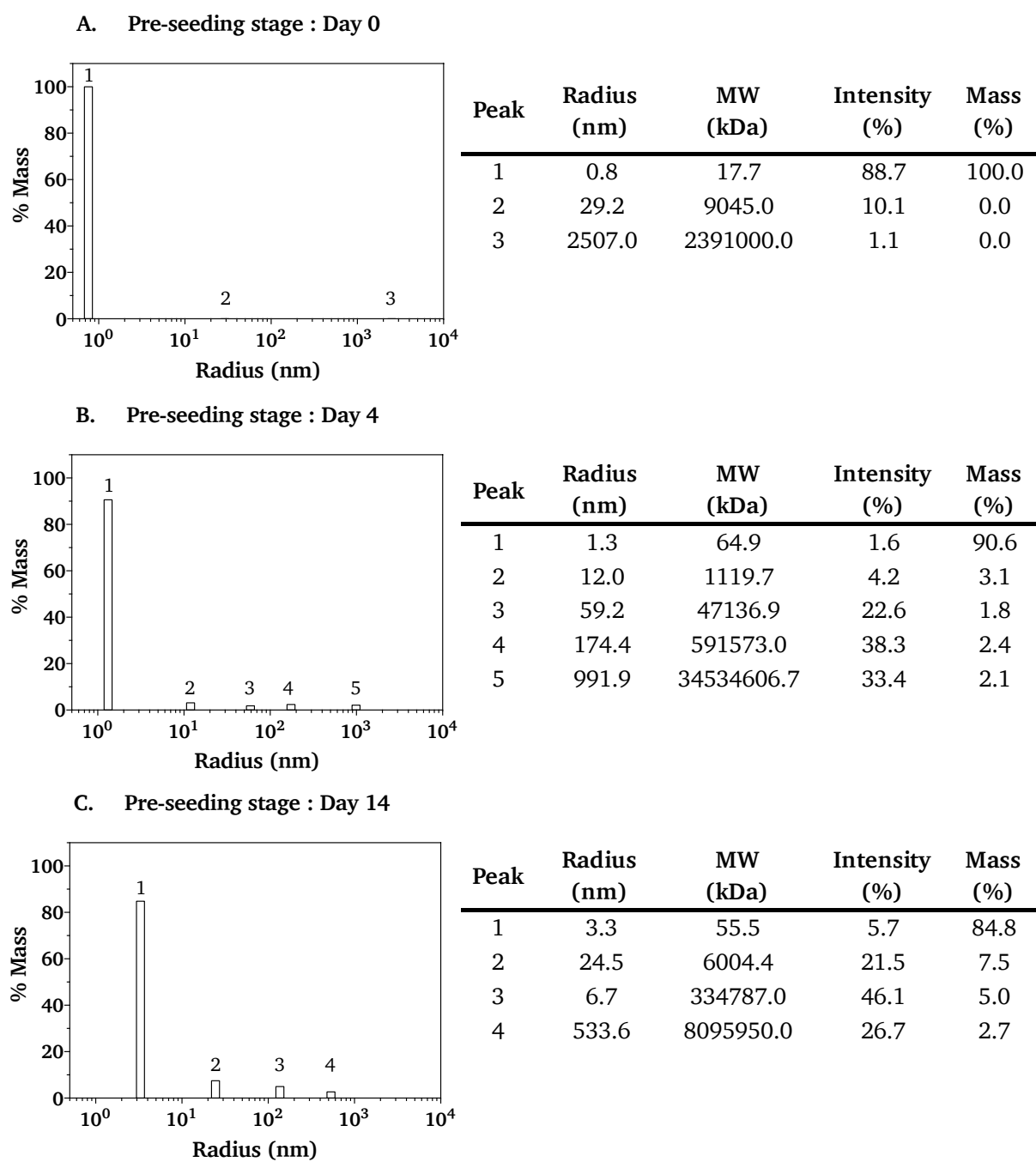


Figure 3.10: TSize distribution results of the pre-seeding stage from DLS measurement. HEWL 3 % (w_v) protein sample incubated in 65 °C for preparation of the HEWL seed solution for 14 days. The particle size distribution of the native HEWL (A), HEWL sample incubated for 4 days (B), and HEWL sample incubated for 14 days (C) were derived by DLS. Histogram on the left-hand side shows % mass of particles with different sizes, and tables on the right-hand side contains the radius (nm), molecular weight (kDa), % intensity, and % mass calculated per each peak.

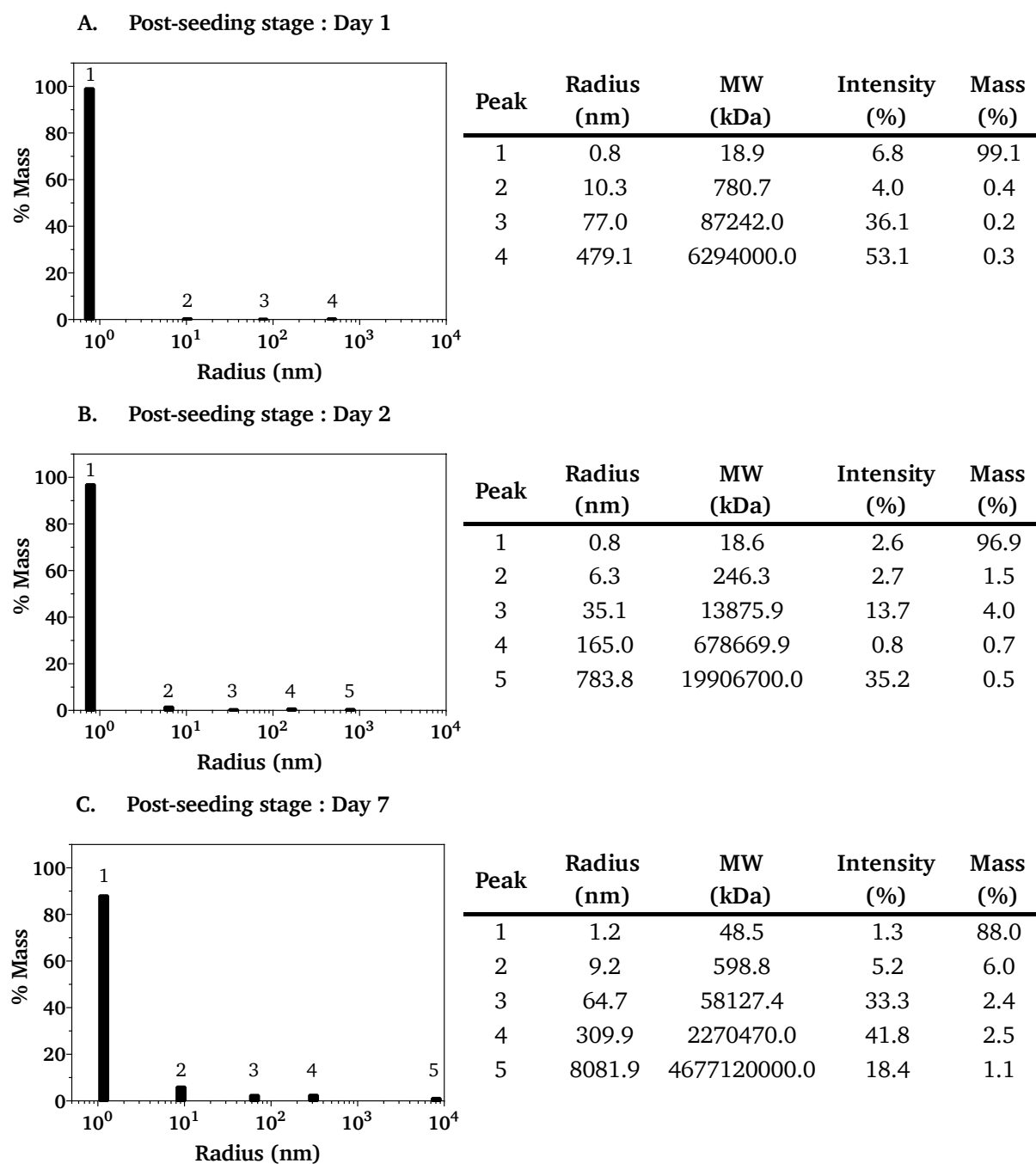


Figure 3.11: The size distribution results of the post-seeding stage from DLS. HEWL fibril seed prepared from the pre-seeding stage was added to HEWL 3 % (w/w) and incubated at 65 °C for 7 days. The particle size distribution of the HEWL+seed incubated for 1 day (A), 2 days (B) and 7 days (C) were derived by DLS. Histogram on the left-hand side shows % mass of particles with different sizes, and tables on the right-hand side contains the radius (nm), molecular weight (kDa), % intensity, and % mass calculated per each peak obtained.

HEWL 3 % (w/w) protein in 10 mM HCl is incubated for 14 days to prepare the seed solution and the prepared seed solution is then added to HEWL 3 % (w/w) solution. DLS was then measured along with the ThT fluorescence assay for monitoring the amyloid fibrillisation which occurs during incubation.

The particle size distribution histogram of HEWL 3 % (w/w) before the incubation and after 14 days of incubation (Figure 3.10) as well as HEWL 3 % (w/w) with 5 % (w/v) seed solution incubated for 7 days (Figure 3.11) are presented with tables describing the radius (nm), molecular weight (kDa), % intensity, and % mass calculated per each peak obtained. HEWL protein without the application of heat remained as a monomer with a hydrodynamic radii of 0.8 nm (Figure 3.10:A). This obtained hydrodynamic radii of monomer is much smaller than the actual HEWL monomer which is 2.05 nm .^[98] Also, it should be noted that the radii of the minor components (peak 2-5 of Figure 3.10 and Figure 3.11) are not very convincing. These may appeared due to the multiple scattering occurs by the high concentration of sample where the light scattered from a single particle is re-scattered by others before reaching the detector.^[117]

However, we were able to detect the increase of hydrodynamic radius detected over the incubation time for both HEWL 3 % (w/w) with and without seed solution.

3.2.2.3 Casted hydrogel-film

At the end of the fibril formation, 0.8 %, 1.0 %, 1.2 %, and 1.5 % v/v glycerol were respectively added to 300 μL each of the hydrogel-film solution. The hydrogel-film solution with four different concentrations of glycerol were then cast by transferring 500 μL onto a flat PTFE (polytetrafluoroethylene) coated film. The solvent was left to evaporate (± 72 hours) and the resulting amyloid fibril/protein films were removed with tweezers. The fibril films were very fragile and easily broke into small pieces. It was noticed that with the addition of 1.2 % v/v (Figure 3.12:C) and 1.5 % v/v (Figure 3.12:D) glycerol, the fibril/protein films stayed intact without breaking into smaller pieces and were easier to remove using tweezers. The dimension of the biggest film obtained with 1.5 % v/v (Figure 3.12:D) glycerol was approximately 1.5×1.5 cm .

⁹⁸ D. K. Wilkins et al. *Biochemistry*, **38**: 16424–16431, 1999.

¹¹⁷ E. Farrel and J. L. Brousseau. *Guide for DLS sample preparation*. Brookhaven Instruments, a Nova Instruments company, 2014.

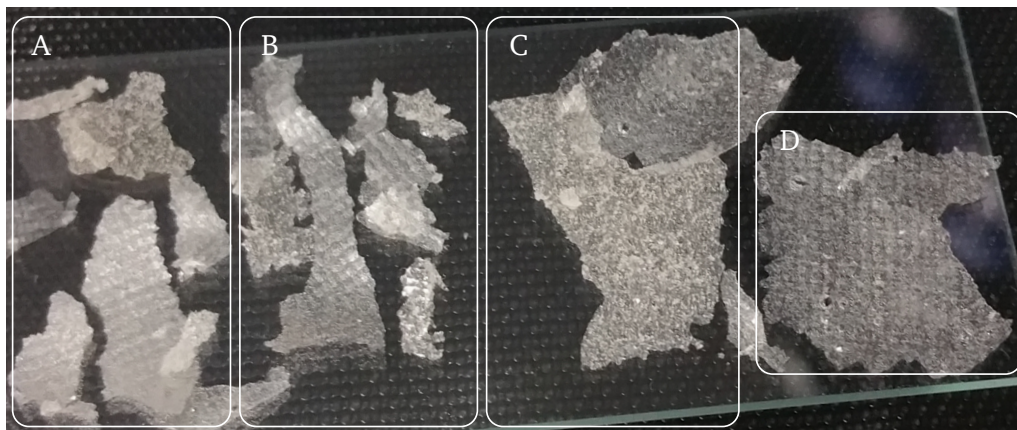


Figure 3.12: HEWL hydrogel film with various concentrations of glycerol. Concentration of (A) 0.8 % v/v , (B) 1.0 % v/v , (C) 1.2 % v/v , and (D) 1.5 % v/v . Addition of more glycerol (1.2 % v/v and 1.5 % v/v) clearly shows that the nature of the film stays intact without disintegrating into smaller pieces as seen in 0.8 % v/v and 1.0 % v/v films.

3.2.3 Bovine β -Lg aggregation

3.2.3.1 ThT fluorescence measurement

β -Lg 2 % (w/v) with pH 2 was incubated at 80 °C for 7 hours as described in Section 2.4.5. ThT fluorescence was measured by excitation at 440 nm at an interval of 1 hour for 7 hours as described in Section 2.5. ThT emission from β -Lg was increasing gradually over 7 hours as ThT binds to β -sheets formed during amyloid formation (Figure 3.14). β -Lg fibril formation showed that there is no fibrillisation that occurred in the first three hours (0-3 hours), which can be referred to as the lag phase. After the lag phase, the growth phase follows in the next 3 hours (3-6 hours) of incubation and between this period exponential fibril growth is observed. This significant increase reaches the inflection point after 6 hours of incubation and enters into the saturation phase (6-7 hours).

3.2.3.2 Secondary structure study with CD: Far-UV

Far-UV CD spectral measurement was performed to detect the conformational changes of the β -Lg proteins. The β -Lg 2 % (w/v) sample with pH 2 was prepared and incubated for 7 hours at 80 °C as described in Section 2.4.5. Far-UV CD spectra measurements were taken every one hour during 7 hours of incubation under the parameters described in Section 2.8.1. Figure 3.15 shows the far-UV

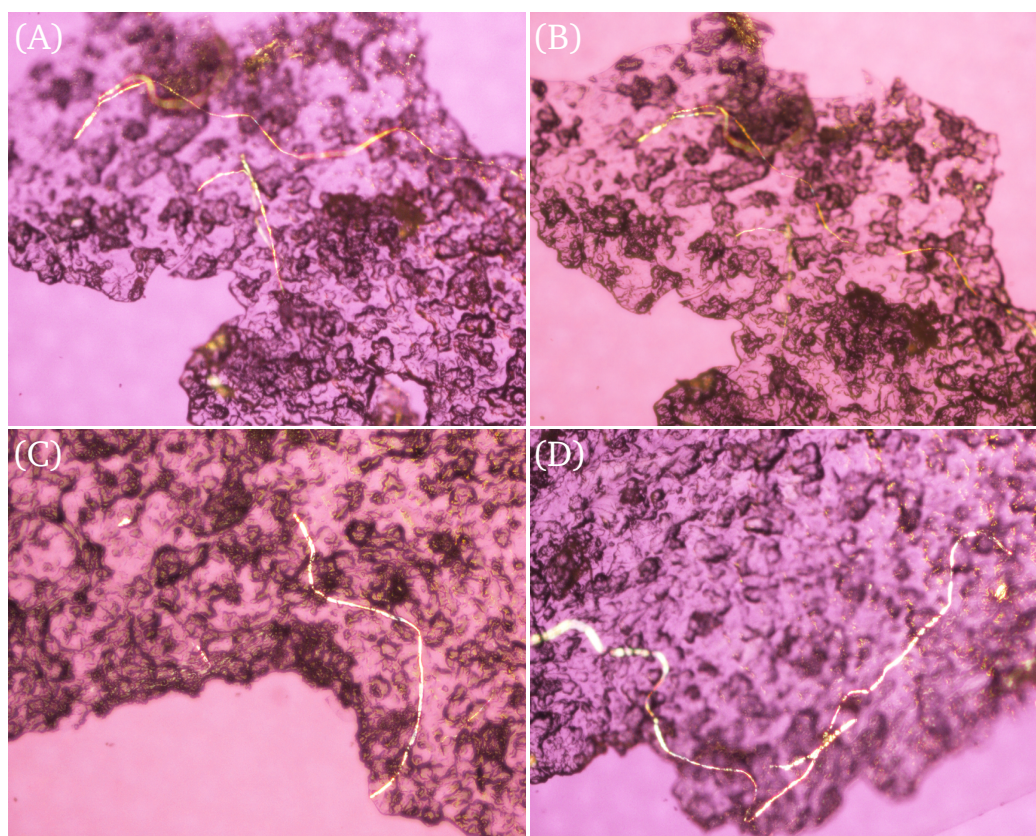


Figure 3.13: Optical microscopy images of HEWL amyloid fibril in the hydrogel film. Hydrogel films contained fibrils were stained with Congo-Red under polarised light microscopy at 90° . The samples are indicated as found identically in Figure 3.12. (Image taken with Nikon ECLIPSE E400 POL with Nikon DS-Fi2 camera).

CD spectra of β -Lg protein at different stages of the timeline. Native β -Lg protein without heating procedure is shown as a red straight line and the shape of Native β -Lg CD spectra does not have the typical features of α -helix with negative bands at 208 nm and 222 nm , nor the typical features of β -sheet with a broad negative band near 218 nm (Figure 1.10), but a broad single band around $208\text{-}218\text{ nm}$. This suggests that secondary structures from the native β -Lg are not just single type of secondary structure element but a mix of multiple secondary elements, and the shape of native β -Lg is more similar to that of the characteristic of β -sheet CD spectra as it has a single native band even if it is very broad. β -Lg protein after 7 hours of incubation at 80°C shown as blue straight line has a native broad band at 218 nm which is the typical feature of β -sheet CD spectra, however it still shows signs of negative band around 208 nm , which is one of negative bands of α -helix (Figure 3.15). This suggest that after 7 hours of incubation, the ratio between α -helical and β -sheet conformation changed towards β -sheet richer structure.

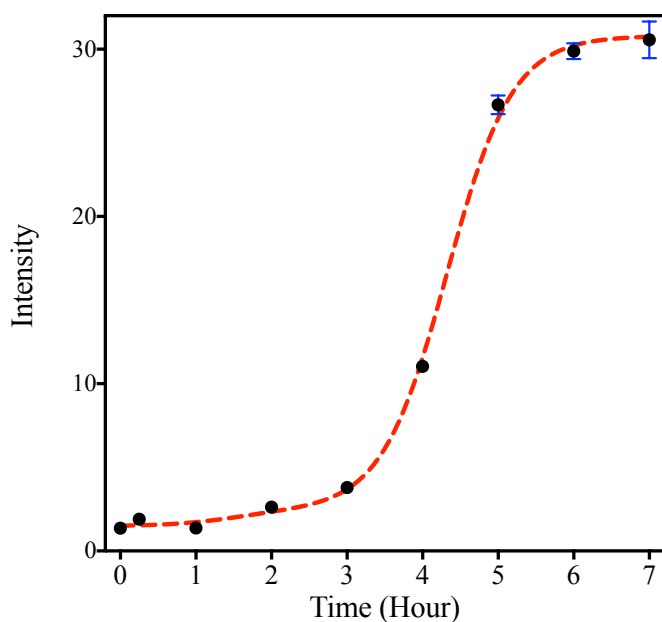


Figure 3.14: ThT fluorescence spectra of the β -Lg protein. Fluorescence intensity of β -Lg 2 % (w/v) incubated at 80 °C was measured by excitation at 440 nm and emission at 482 nm at an interval of 1 hour for a total of 6 hours. Each time points, three ThT measurement were performed.

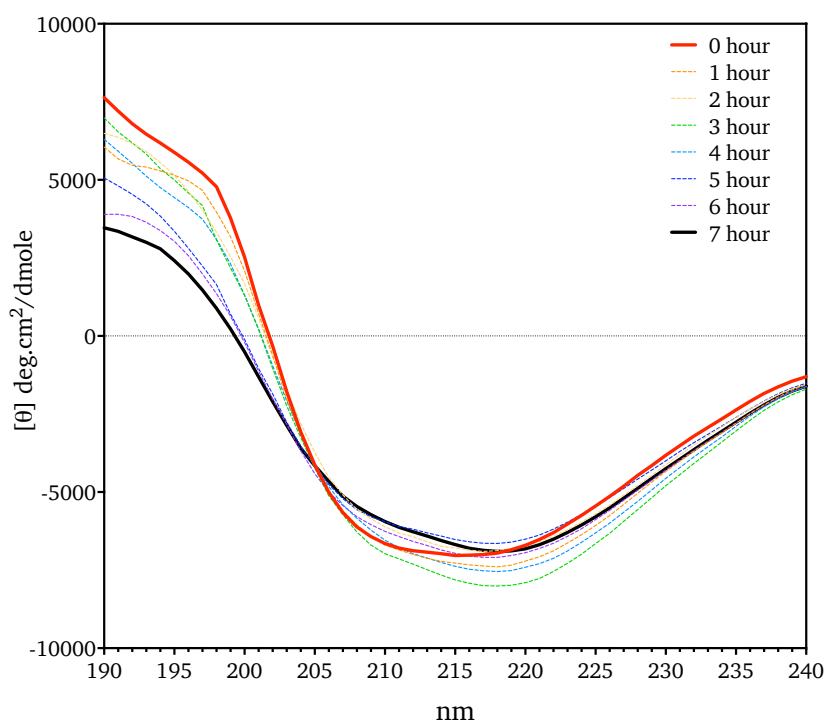


Figure 3.15: Far-UV CD spectra of β -Lg protein. Far-UV spectra of β -Lg 2 % (w/v) incubated at 80 °C measured at wavelength 190 nm to 240 nm at an interval of 1 hour for a total of 7 hours. Red line (0 hour) represents native β -Lg without heat treatment and black line (7 hour) shows the heat-treated β -Lg.

Interestingly, the secondary structure calculations performed by K2D3^[65]ⁱ do not show a dramatic increase in β -sheet structure as the incubation time is increased. Even though mature fibrils are clearly observed with AFM and TEM. In this case, we think that the β -barrel structures of the native molecules is not as perturbed as those found in proteins with a large proportion of α -helical conformation.

Other dot lines between two straight lines are representing the incubated β -Lg protein between 1 hour and 6 hours.^[63–66]

3.2.3.3 TEM imaging of native β -Lg protein and β -Lg fibril

Amyloid fibril or amyloid-like fibrils are distinguished from amorphous aggregates by their ordered β -sheet rich structure. Fibrillization of protein is generally observed with the ThT assay. A limitation of the ThT assay is in its inherent difficulty to distinguish between oligomers, fibril-like structure, and mature fibrils. Therefore, TEM imaging is often necessary and is used as a final conformational test to determine the presence of mature fibrils. TEM is a useful and powerful tool for characterising the morphology of fibrils and fibrillary structures such as early aggregates, oligomers, protofibrils, mature fibrils, and amorphous aggregates.^[19]

As described in Section 2.10, both native β -Lg protein and 6-hour incubated β -Lg protein sample were deposited onto carbon coated grids and negatively stained. Amyloid fibrils and proteins that are negatively stained can be visualised at high resolution TEM. Native β -Lg protein (control) before applying the heating process show no signs of aggregates when observed with TEM at different scales (Figure 3.16:A). However, as expected from ThT fluorescence and CD measurements, presence of β -Lg amyloid fibrils in the 6-hour incubated β -Lg sample was detected under TEM (Figure 3.16:B&C). Mature β -Lg amyloid fibrils show extensive fibrillary growth in a relatively short time (6 hours) with many long unbranched fibrillary structures present in the sample.

⁶⁵ C. Louis-Jeune, M. A. Andrade-Navarro, and C. Perez-Iratxeta. *Proteins Struct. Funct. Bioinforma.*, **80**: 374–381, 2012.

⁶³ S. M. Kelly, T. J. Jess, and N. C. Price. *Biochim. Biophys. Acta - Proteins Proteomics*, **1751**: 119–139, 2005.

⁶⁴ B. M. Bulheller, A. Rodger, and J. D. Hirst. *Phys. Chem. Chem. Phys.*, **9**: 2020, 2007.

⁶⁶ N. J. Greenfield. *Nat. Protoc.*, **1**: 2876–2890, 2006.

ⁱ Webinterface for estimating protein secondary structure from CD spectra

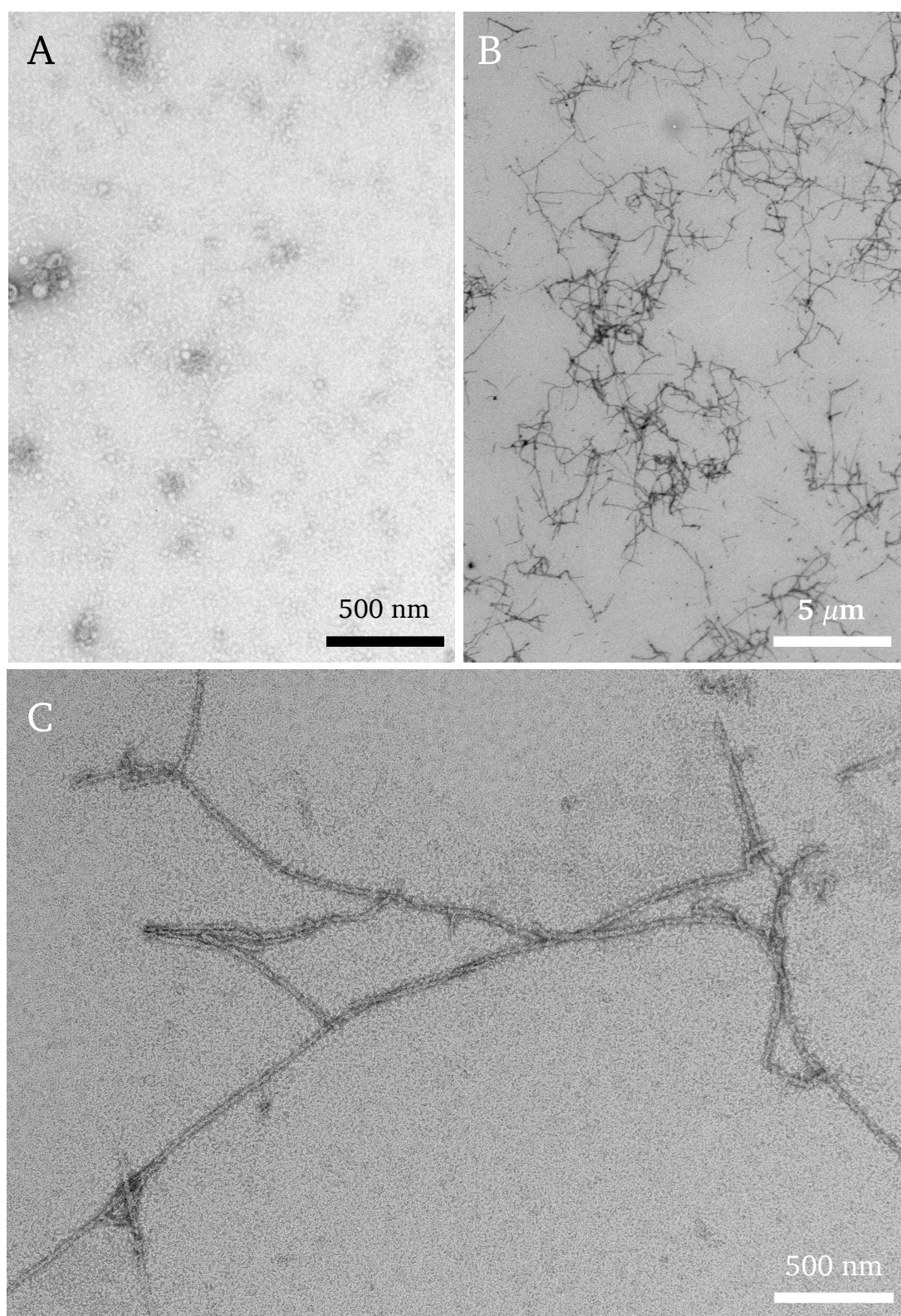


Figure 3.16: TEM images of Native β -Lg protein (A) and 6-hour incubated β -Lg (B,C). Native β -Lg protein and 6-hour incubated β -Lg sample were examined using TEM. Native β -Lg (A) images do not have any distinguishable aggregations. Images of β -Lg protein incubated for 6 hours (B,C) show the presence of fine fibrillary structure.

3.2.3.4 Investigation of β -Lg fibril with AFM

Atomic force microscopy (AFM) is a powerful technique to investigate the morphology and physical characteristics of amyloid fibrils and their intermediate species. TEM imaging is useful to confirm the presence of mature fibrils, however, it is not able to give further physical properties such as height and width of fibrils, which AFM can provide. Here, AFM is used to investigate the morphology of amyloid fibrils formed from β -Lg 2 % (w/v) at pH 2 after 6 hours of incubation at 80 °C.

By performing AFM, we were able to obtain 2D/3D images of surface-adsorbed β -Lg fibrils, acquired in liquid, as shown in Figure 3.17:A&B, as well as cross-section of fibrils at various location Figure 3.17:C&C1-12. Helical twists of amyloid fibrils are indicated with arrowheads (Figure 3.17:A&B). Height profile of two fibrils were compared at the location shown in Figure 3.17:C. Even under the same condition, formation of fibrils with different physical properties were noticed. Fibril I (C1-6) has a maximum height of 8 nm (C3/5) and minimum height of 6 nm (C2/4/6) with a period (distance between two helical twists) of 150 nm. Fibril II (C7-12) has a maximum height of 3.5 nm (C9/11) and minimum height of 2 nm (C8/10/12) with a period of 75 nm. Compared to previous studies on β -Lg fibrils,^[40] our results observed and resolved the direct correlation between period and height of the fibrils. When the height of the fibril increases, the distance between neighbouring helical twists also increases.

3.2.4 Diffusion NMR of proteins with different sizes

Diffusion NMR enables separation of signals in a 2D matrix based on particle size.^[94,95] Lokeren L. Y. *et al.*^[118] used DOSY NMR for characterisation of dispersed titanium dioxide nanoparticles and Boisselier E. *et al.*^[71] used DOSY NMR for size diversity of different sizes of gold NPs-triethyleneglycol (TEG) composite dendrimers.

In this section, two diffusing biomolecules HEWL (14.3 kDa) and L-Tryptophan (L-Trp) (0.2 kDa) were selected to demonstrate the separation of signals. The molecule 1,4-dioxane was also added to each NMR sample as an internal chemical shift standard (3.75 ppm).

¹¹⁸L. Van Lokeren *et al.* *Chem. - A Eur. J.*, **13**: 6957–6966, 2007.

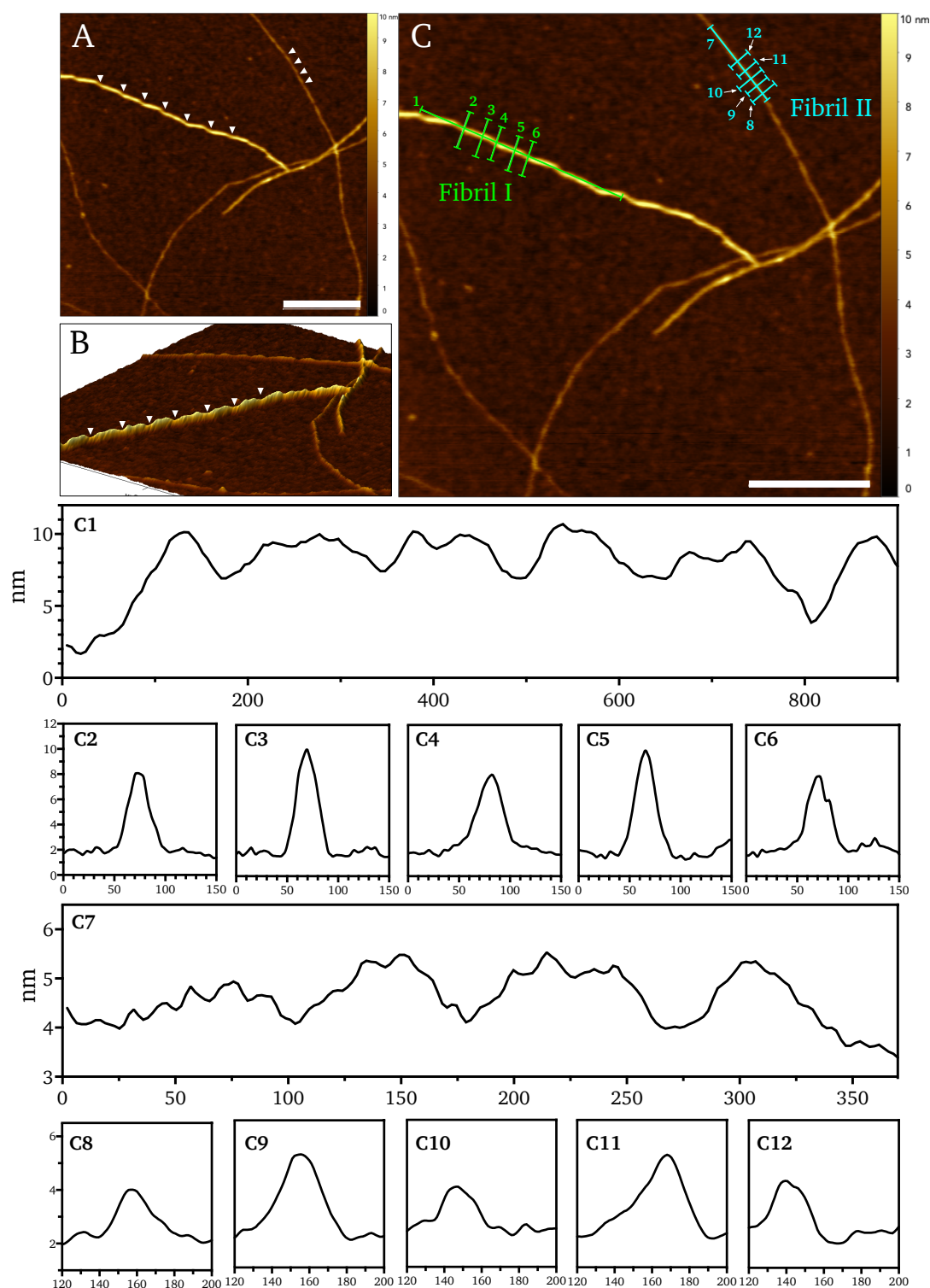


Figure 3.17: Structural analysis of β -Lg fibrils with AFM. (A) AFM image of β -Lg amyloid fibril absorbed on mica. Arrowheads indicate the location and orientation of the helical twist. (B) Three-dimensional rendering of the surface topography of β -Lg fibrils. Arrowheads indicate helical twist. (C) & (C1-C12) Height profile of various area of fibrils and their locations. Scale bars: 500 nm.

Four different combinations of protein sample were prepared and DOSY NMR was acquired under the same diffusion NMR settings (Section 2.12). The diffusion coefficient analysis resulting from the 2D DOSY spectrum obtained from the combined sample of HEWL, Trp, and 1,4-dioxane (Figure 3.18:A1-Red), Tryptophan with 1,4-dioxane (Figure 3.18:A2-Green), HEWL with 1,4-dioxane (Figure 3.18:A3-Blue), and 1,4-dioxane on its own (Figure 3.18:A4-Purple) are shown along with the overlay of all four of the spectra obtained (Figure 3.18:B).

Figure 3.18 clearly reveals the presence of three components with different diffusion coefficient values. Comparing the four spectra, HEWL has the lowest diffusion coefficient whilst Trp has a slightly smaller diffusion coefficient than 1,4-dioxane. Furthermore, dioxane has the largest diffusion coefficient of the three molecules. Based on their size, signals from three components are separated clearly on the diffusion NMR spectra.

3.2.5 Estimating the proportion of bovine β -Lg monomer-dimer equilibrium at low/high protein concentrations with diffusion NMR

The conformation and association behaviour of β -Lg are pH dependent as previously mentioned in Section 1.2.2. At acidic condition, β -Lg exists in monomeric form while at neutral pH, β -Lg is associated with both monomeric and dimeric forms.

Here, the hydrodynamic radii of 1 % and 12 % (w/w) β -Lg at pH 2 and pH 7 were calculated. The diffusion coefficient value of four samples; 1 % β -Lg (pH 2), 1 % β -Lg (pH 7), 12 % β -Lg (pH 2), and 12 % β -Lg (pH 7) as well as the reference molecule 1,4-dioxane of each sample was obtained from DOSY NMR as shown in Figure 3.19. With these obtained diffusion coefficients, the average hydrodynamic radius of β -Lg with four different conditions were calculated using the Equation (2.3) (an example calculation is provided in Appendix A.1).

The calculated hydrodynamic radius of 1 % β -Lg at pH 2 (monomer) is 22.90 Å and the average hydrodynamic radius of 1 % β -Lg at pH 7 (monomer-dimer) is 28.89 Å. By approximating the dimer as the interaction of two touching hard-sphere molecules, the hydrodynamic radius ratio between the dimer and the monomer is estimated to be 1.33!^[119] The estimated β -Lg dimer radii from the

¹¹⁹P. Aymard, D. Durand, and T. Nicolai. *Int. J. Biol. Macromol.*, **19**: 213–221, 1996.

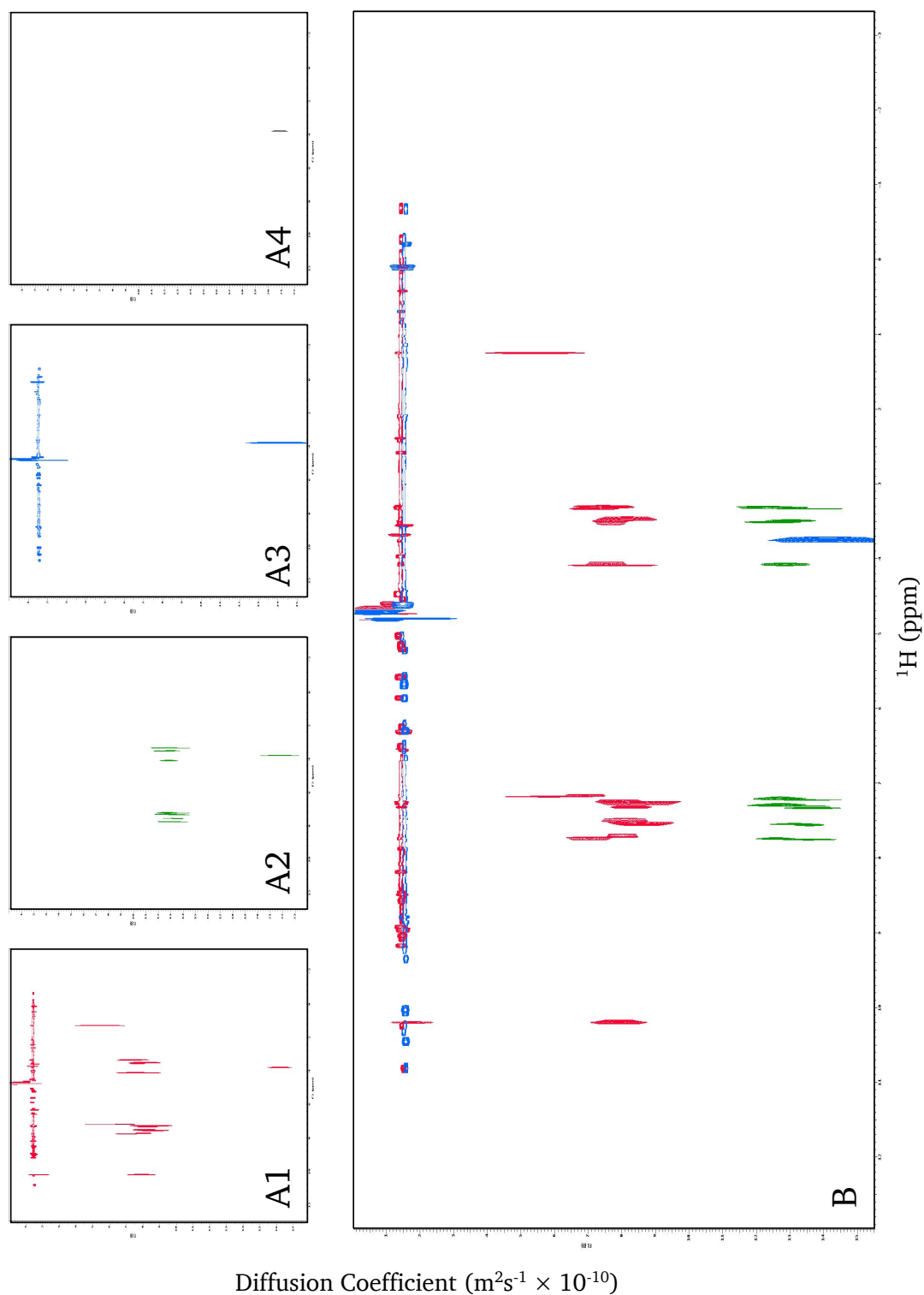


Figure 3.18: 2D-DOSY spectra; (A) HEWL+Trp+Dioxane (A1:Red), Trp+Dioxane (A2:Green), HEWL+Dioxane (A3:Blue), and Dioxane (A4:Purple). (B) Overlapped spectrum of HEWL+Trp+Dioxane (A1), Trp+Dioxane (A2), and HEWL+Dioxane (A3) (Direct imported image for VnmrJ software).

obtained monomer radii is 30.46 Å ($1.33 \times 22.90 \text{ Å} = 30.46 \text{ Å}$). With the obtained radii/radius values, the percentage proportion of monomer and dimer at pH 7 was calculated as;

$$R_h^{pH7} = (R_h^{monomer}) \times (R_h^{dimer})(1-x)$$

$$28.89 = (22.90) \times (30.46)(1-x)$$

$$x = 0.21$$

Therefore, in β -Lg pH 7, the calculated percentage proportion of the monomer is 21 % and the dimer is 79 %. This estimated percentage proportion corresponds well with the recent monomer-dimer equilibrium study presented by Khan *et al.*^[120] (monomer/dimer: 25/75 % at pH 7.0) Further to this, the calculated average hydrodynamic radius of 1 % β -Lg at pH 2 and pH 7 was 22 Å (2.2 nm) and 29.3 Å (2.93 nm)^[99] respectively. These results agreed with the β -Lg monomer & dimer hydrodynamic radius presented by Parker *et al.*,^[99] where the obtained β -Lg hydrodynamic radii was 2.2 nm at low pH (2.0-3.0) and the range was 2.6 to 4.9 nm at neutral pH (pH 6.0-8.0).

As shown, at low concentration of 1 % w/w , β -Lg exists as a monomer at pH 2. However, when the β -Lg concentration was increased to 12 % w/w (pH 2), the hydrodynamic radius increased to 32.45 Å. Similarly, the hydrodynamic radius of 12 % β -Lg at pH 7 increased to 36.01 Å from that of 1 % β -Lg at pH 7. The hydrodynamic radii of 12 % β -Lg at both pHs (2 and 7) are larger than the estimated hydrodynamic radius for a dimer, suggesting the presence of higher order oligomers such as, tetramers and octamers, even though β -Lg is known to exist as a monomer at acidic conditions. An increase in the hydrodynamic radius of β -Lg is observed when the pH is increased from 2 to 7, indicating that the proportion of the higher order oligomer present at pH 7 is higher. The oligomerization of β -Lg at higher concentration (lower than 12 % (w/w)) has been shown in previous studies.^[120,121] Verheul *et al.*^[121] obtained structures larger than dimers at pH 6.9 with high concentration β -Lg ($> 30 \text{ mg/mL}$), and Khan *et al.*^[120] presented an increase in concentration from 1 to 2 mg/mL , where the weight of monomer-dimer percentage proportion shifted toward the dimer at both pH 2 and 7.

¹²⁰ S. Khan *et al.* *Biomacromolecules*, **19**: 2905–2912, 2018.

⁹⁹ R. Parker *et al.* *Biophys. J.*, **89**: 1227–1236, 2005.

¹²¹ M. Verheul *et al.* *Biopolymers*, **49**: 11–20, 1999.

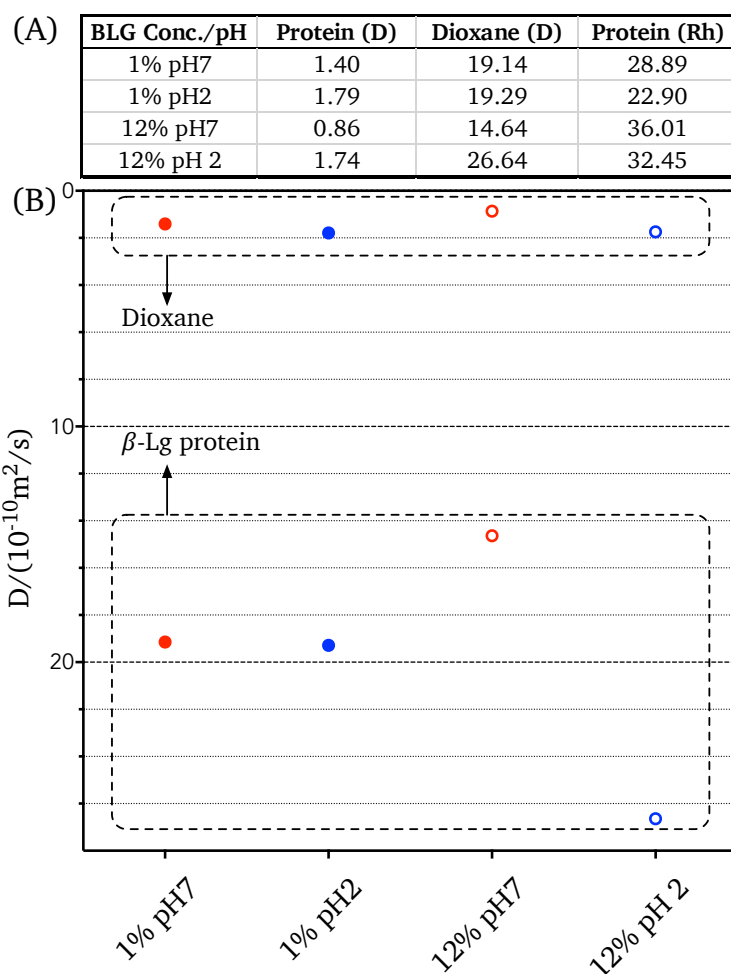


Figure 3.19: Diffusion coefficient values of β -Lg 1 %/pH 7, β -Lg 1 %/pH 2, β -Lg 12 %/pH 7, and β -Lg 1 %/pH 2 and obtained hydrodynamic radius. (A) Average relevant diffusion coefficient values obtained from 2D-DOSY plot of β -Lg 1 %/pH 7, β -Lg 1 %/pH 2, β -Lg 12 %/pH 7, and β -Lg 1 %/pH 2 as well as dioxane was used to calculate the hydrodynamic radius ($\text{\AA}$) of protein in table. (B) Diffusion coefficient value plot. The Doneshot DOSY pulse sequence used (Diffusion gradient length = 5 ms/Diffusion delay = 100 ms)

3.2.6 Size estimation of cross-linked HEWL with BS³ using diffusion NMR

3.2.6.1 SDS-PAGE electrophoresis

Molecular weight of the prepared cross-linked HEWL using various BS³ concentrations was determined with SDS-PAGE as described in Section 2.9. The HEWL 1 mM – BS³ 0.5 mM sample contained monomer, dimer, and trimer HEWL populations as shown in Figure 3.20, and this sample was used in the following diffusion NMR experiments.

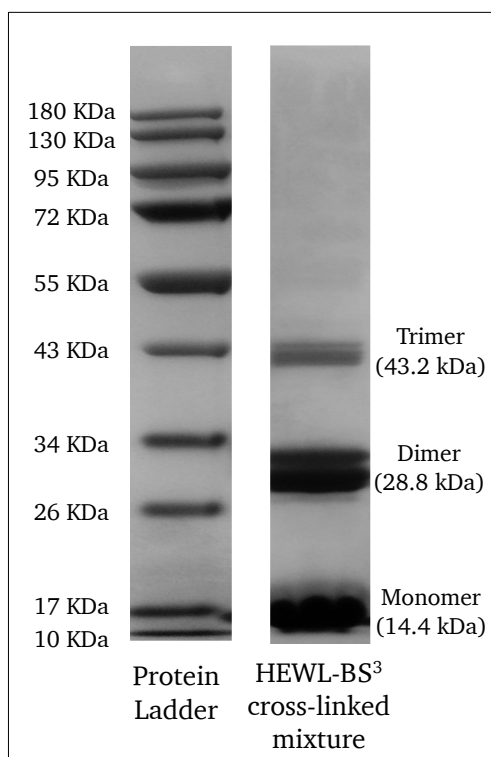


Figure 3.20: SDS-PAGE gel of the HEWL (1 *mM*)–BS³ (0.5 *mM*) cross-linked mixture sample

3.2.6.2 Diffusion NMR

The 1-dimension ¹H diffusion trace over the increasing gradient strength of the HEWL-BS³ 0.5 *mM* sample are shown in Figure 3.21 with their 1D-¹H NMR spectra. Gradual decaying of dioxane (3.75 ppm) & cross-linker BS³ (3.80-3.78 ppm) signal (Figure 3.21:A) and the aliphatic region signal (3.0~0.0 ppm) of HEWL-BS³ sample (Figure 3.21:B & C) is observed with the increase of gradient strength. The rate of decay of the cross-linked sample signal was much slower compare to the rate of decay of dioxane and BS³, signals which signifies that the cross-linked HEWL sample has, on average, a much larger hydrodynamic radius than dioxane and BS³.

The obtained diffusion NMR data can be represented as a pseudo-2D contour (Figure 3.22), termed ‘diffusion-ordered spectroscopy (DOSY)’ in which one dimension represents the chemical shifts and the other the diffusion coefficients of each component population. However, due to small size differences between the cross-linked components; monomer, dimer and trimer HEWL-BS³ as SDS-PAGE showed (Figure 3.20), as well as the fact that all components would maintain the same

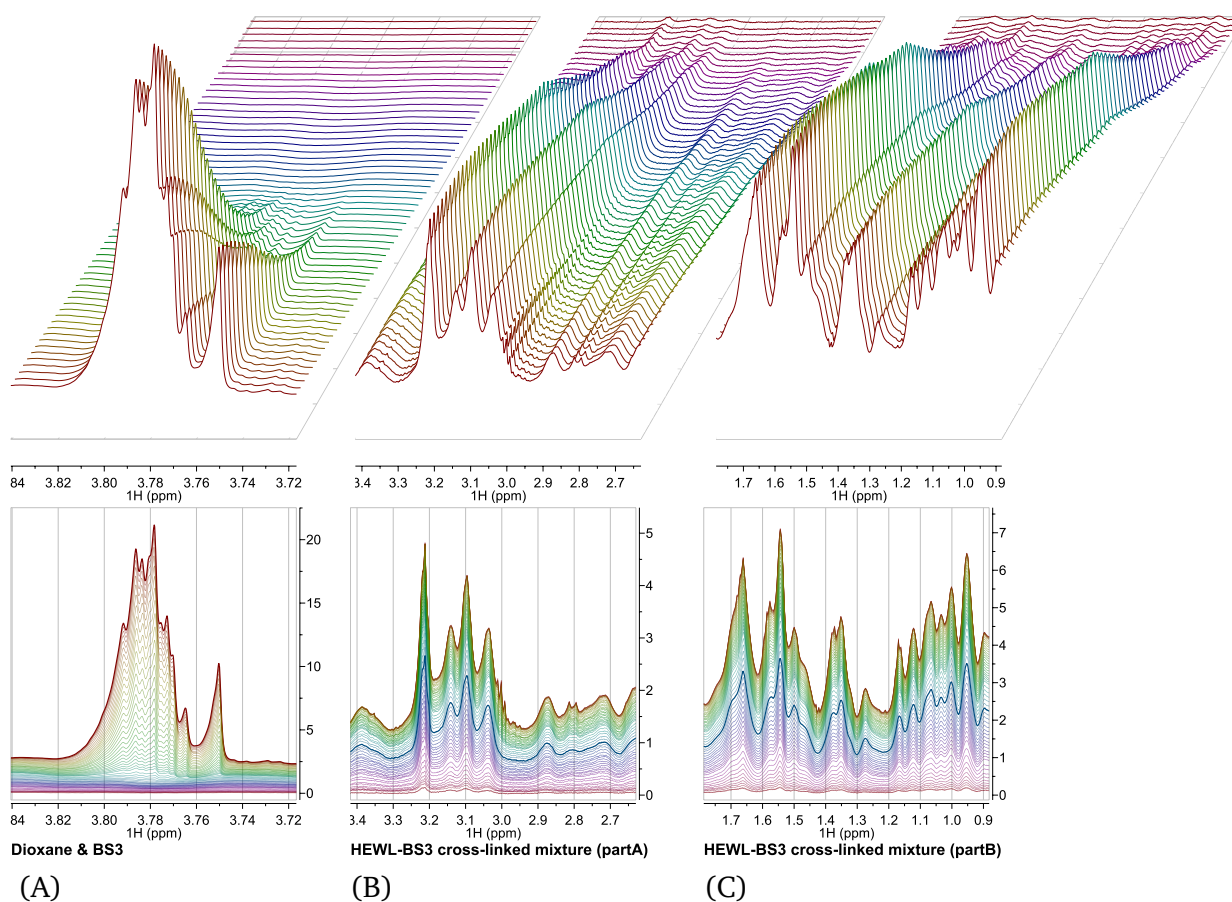


Figure 3.21: Experimental 1D ^1H diffusion traces of HEWL-BS³ 0.5 mM mixture sample showing progressive intensity decay as a function of the increasing gradient strength with 1D ^1H spectrum. (A): Dioxane (3.75 ppm) & Cross-linker BS³ (3.80~3.78 ppm)/(B)-(C): HEWL-BS³ mixture sample peaks shown in two different ranges

chemical shift. Therefore DOSY cannot distinguish each of the crosslinked HEWL-BS³ components and only an average diffusion coefficient of the HEWL-BS³ mixture can be obtained from DOSY. Having obtained the diffusion coefficient of the HEWL-BS³ mixture ($2.116 \pm 0.141 \times 10^{-10} \text{ m s}^{-1}$) and of 1,4-dioxane ($18.75 \times 10^{-10} \text{ m s}^{-1}$), the average hydrodynamic radius of the HEWL-BS³ mixture was calculated through the Equation (2.3). An example of this calculation is displayed in Appendix A.1. Calculated population average hydrodynamic radius of HEWL-BS³ mixture was $18.79 \pm 1.34 \text{ \AA}$.

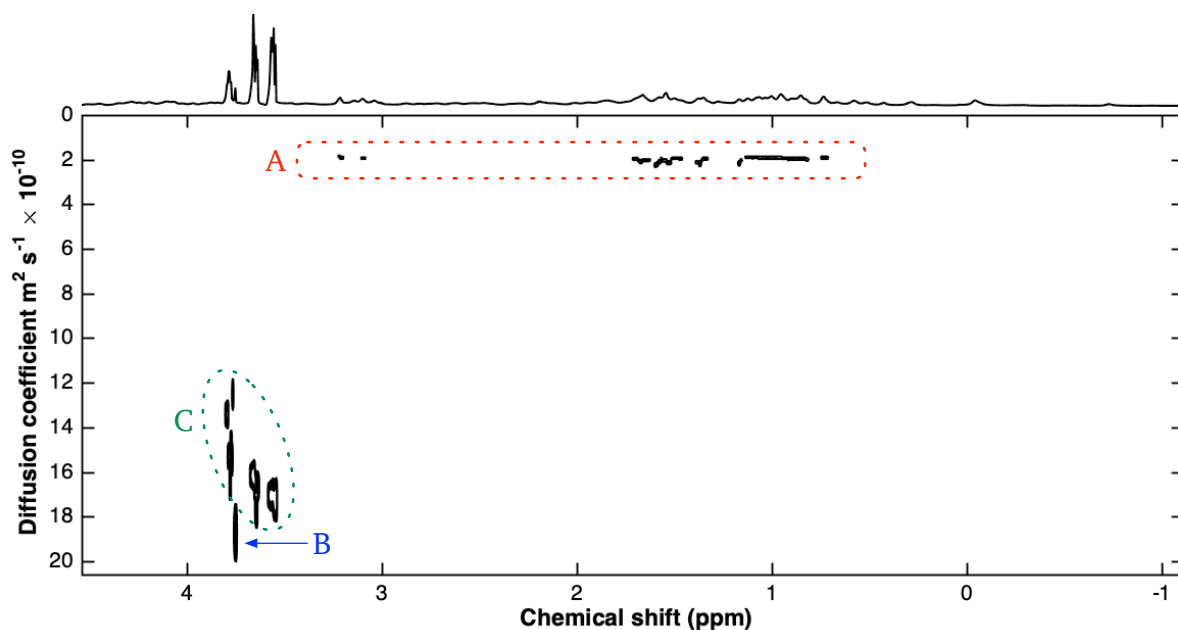


Figure 3.22: 2D-DOSY spectrum of HEWL-BS³ 0.5 mM cross-linked sample. (A): Cross-linked HEWL-BS³ ($D_{HEWL-BS^3} = 1.975 \sim 2.257 \text{ m}^2 \text{ s}^{-1} \times 10^{-10}$), (B): Dioxane (3.75 ppm / $D_{Dioxane} = 18.75 \text{ m}^2 \text{ s}^{-1} \times 10^{-10}$), and (C): Cross-linker BS³ (Spacer Arm: 11.4 Å)

3.2.6.3 Diffusion coefficient calculation from the decay profile with MATLAB

The normalized integrated spectral intensity as a function of gradient strength of the aliphatic region of the HEWL-BS³ mixture and the dioxane signal (3.75 ppm) was plotted as shown in Figure 3.23.

Firstly, combining the estimated hydrodynamic radius of HEWL-BS³ populations from SE-HPLC (Appendix A.2) and the estimated ratio between each components from SDS-PAGE (Table 3.1), the poly-Gaussian decay function with three components was set as below.

$$I(g) = I_1 e^{-D_1 g^2} + I_2 e^{-D_2 g^2} + I_3 e^{-D_3 g^2}$$

$$I(g) = 0.5 e^{-4.58 \times 10^{-9} g^2} + 0.45 e^{-3.98 \times 10^{-9} g^2} + 0.047 e^{-3.27 \times 10^{-9} g^2}$$

Fitting obtained by using the above poly-Gaussian decay function is displayed using a green solid line in Figure 3.23:A.

Secondly, the poly-Gaussian decay function (Equation (2.4)) with the three diffusing components was employed to measure the diffusion NMR experimental hydrodynamic radius of the HEWL-BS³

Table 3.1: Estimated hydrodynamic radius and the percentage proportion of HEWL-BS³ oligomers via SE-HPLC and SDS-PAGE

HEWL peak	R_h (Å)	Proportion percentage (%)
Monomer	20	50.0
Dimer	23	45.0
Trimer	28	4.7

Table 3.2: Calculated hydrodynamic radius and percentage proportion of HEWL-BS³ oligomers by poly-Gaussian decay function with three components

Diffusing component	R_h (Å)	Proportion percentage (%)
Component 1	18.68	57.5
Component 2	20.88	35.3
Component 3	34.54	8.5

mixture present in the sample, as well as to estimate the ratio between the monomer/oligomer population by using the MATLAB curve fitting tool. The fitting of HEWL-BS³ is shown as a black solid line while that of 1,4-dioxane is shown as a red solid line in Figure 3.23:B. The diffusion coefficient value of three HEWL-BS³ components and 1,4-dioxane collected from the fitting was then used for the hydrodynamic radius calculation through Equation (2.3) (calculation example is shown in Appendix A.1). Also, from the intensity of each component, the proportion of each population can be estimated. The calculated hydrodynamic and population ratio is displayed in Table 3.2.

3.2.7 Pseudo real-time early stage amyloid oligomerization monitoring

3.2.7.1 Pseudo real-time monitoring of early stage amyloid oligomerization of HEWL with Diffusion NMR

As shown in Figure 3.24, decay rates as a function of gradient strength of the native HEWL and HEWL oligomers incubated for 6, 12, 18, and 24 hours decreased over the incubation time. From this overlaid figure we can instantly identify the average hydrodynamic radius/volume of the diffusion species increased during the incubation period from 0 to 24 hours.

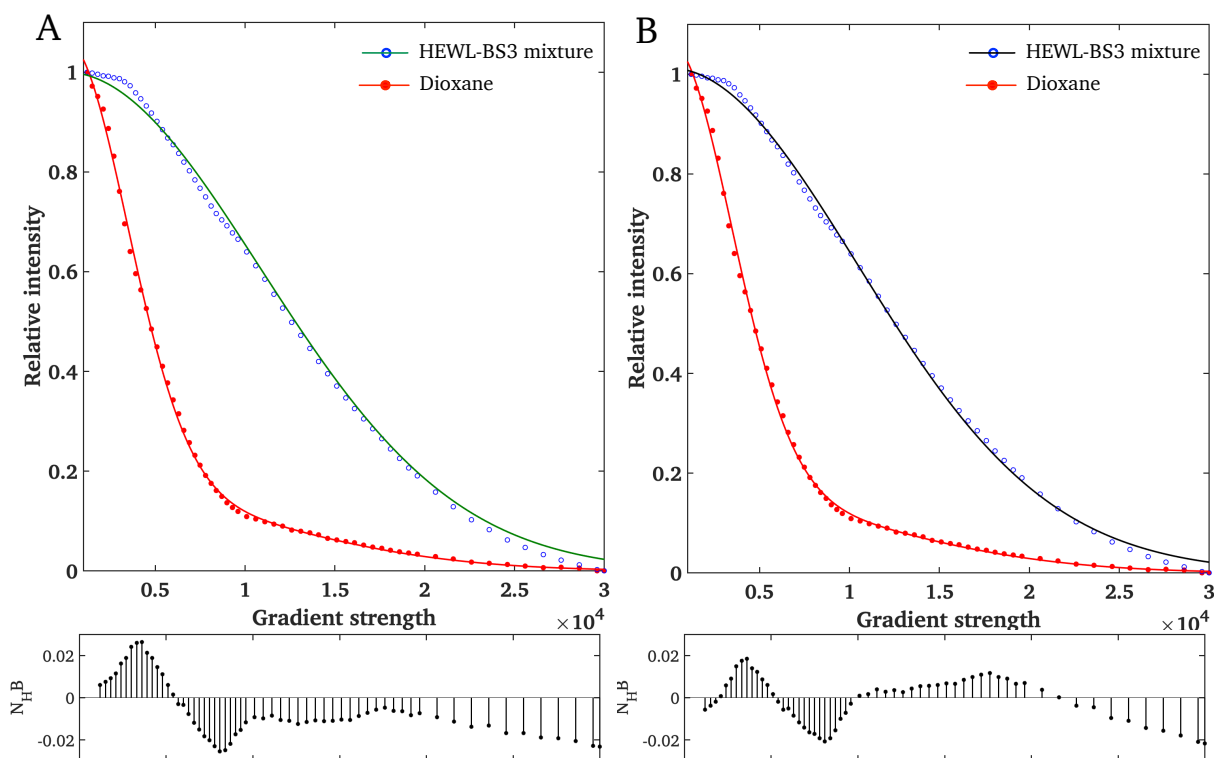


Figure 3.23: Plot of the relative integrated spectral intensity as a function of gradient strength for the aliphatic region of the HEWL-BS³ mixture and the dioxane signal (3.75 ppm). Solid lines represent the fitting curve obtained via Gaussian curve function. (A) Fitting with an estimated R_h obtained from SE-HPLC with residual plot at the bottom (RMSE=0.0149) (B) Fitting with poly-Gaussian curve function with three components with residual plot at the bottom (RMSE=0.0110) and RMSE between plot (A) and (B) is 0.00005717

Also, the significance of the difference between diffusion coefficients of the native HEWL and the 24 hours incubated HEWL was tested. Firstly, a linear transformation was performed on the Gaussian decay function (Equation (2.4)).

$$\ln(I) = \ln(I_0) - Dg^2$$

The values obtained were plotted as shown in Figure 3.25. The slopes of the linear fits of native HEWL and the 24 hours incubated HEWL were respectively $-5.602 \times e^{-9}$ (95%CI = $-5.5 \times e^{-9}$, $-5.362 \times e^{-9}$) and $-4.223 \times e^{-9}$ (95%CI = $5.339 \times e^{-9}$, $-4.107 \times e^{-9}$). Equality of the slopes was then tested and found to be different to a high level of significance ($p < 0.0001$).

Mono-Gaussian decay function with a single component

The integrated spectral intensity of the aliphatic region signals decayed gradually with the

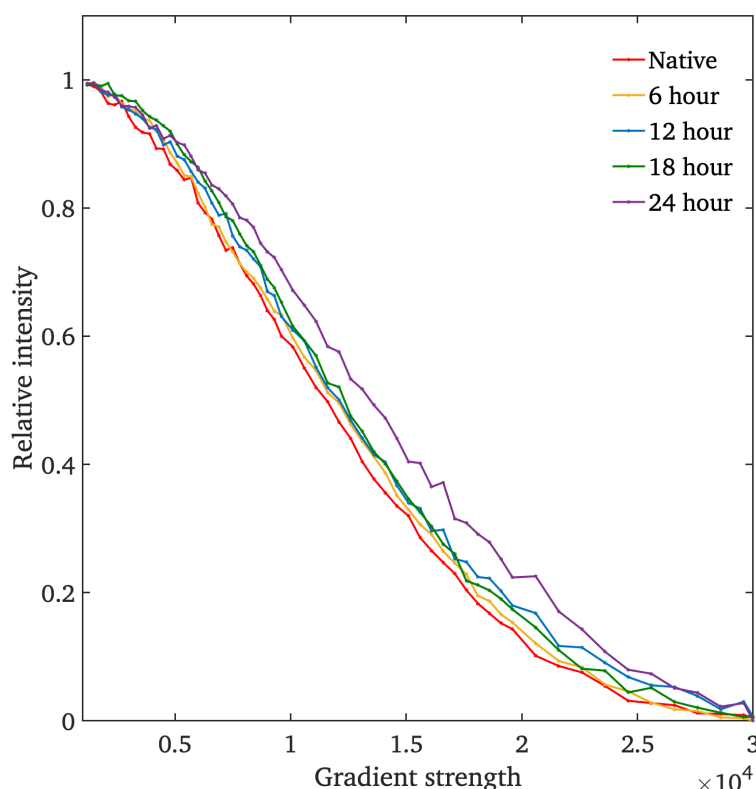


Figure 3.24: Overlaid plot of the normalised integrated intensity as a function of the gradient strength at the aliphatic region signal (3.0-0.0 ppm) of the native HEWL and incubated HEWL for 6, 18, 21, 24 hours at 65 °C.

increase of gradient strength, which followed a Gaussian decay function (Equation (2.4)). Firstly, the fitting procedure was attempted using a mono-Gaussian decay function (single component) as shown in Figure 3.26. Then, the hydrodynamic radius of native HEWLⁱⁱ and the average hydrodynamic radius of the HEWL oligomers were calculated through Equation (2.3) using the diffusion coefficient value of protein oligomer ($D_{protein}$) and dioxane ($D_{dioxane}$) obtained from the fitting process. The calculated hydrodynamic radius of the native HEWL and HEWL oligomer samples (6, 12, 18, and 24 hours) are shown in Table 3.3, and a calculation example is available in Appendix A.1.

The calculated hydrodynamic radii of monomer HEWL (Native HEWL) 19.20 Å aggregates well with the hydrodynamic radii value presented by Hill *et al.*^[72] (DLS = 19 Å / AFM = 18±3 Å) and slightly smaller than the hydrodynamic radii presented by Wilkins *et al.*^[98] (PFG NMR = 20.5

ⁱⁱ Fitting of native HEWL not shown in the thesis but hydrodynamic radius was calculated

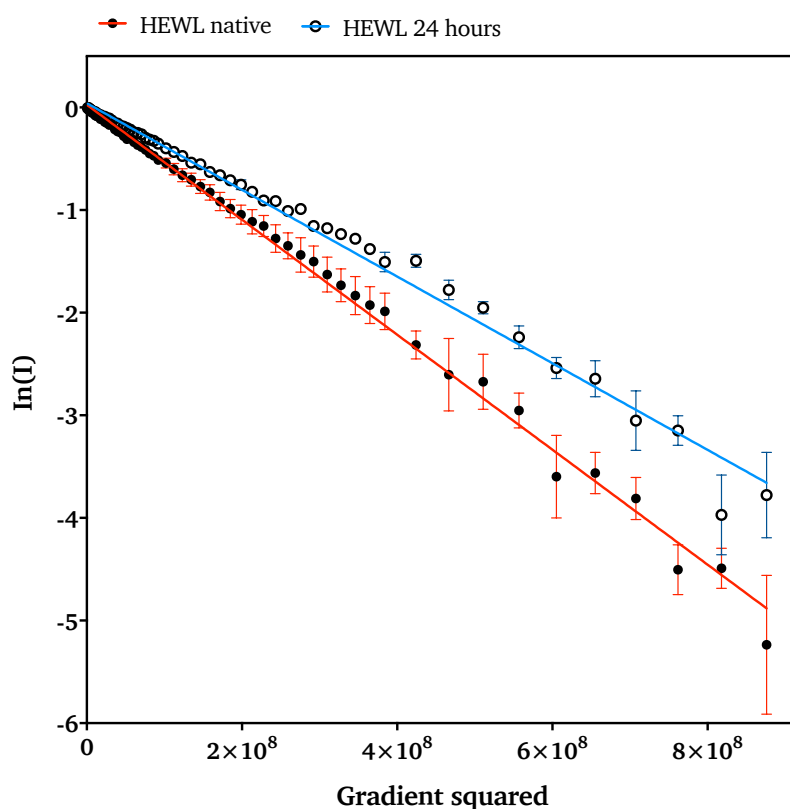


Figure 3.25: Linear formatted diffusion NMR data of the native and 24 hours incubated HEWL. Slope of the native HEWL is -5.602×10^{-9} and of the HEWL incubated for 24 hours is -4.223×10^{-9} . (ANCOVA outcomes: $F = 177.6$, $P < 0.0001$)

Å) As shown in Table 3.3, the average hydrodynamic radius of HEWL oligomer increased over 24 hours of incubation period which we expected from the overlaid decay profile (Figure 3.24). Only a small increase in the hydrodynamic radius of HEWL during the first 18 hours was observed as it increased from 19.20 Å to 21.76 Å which is a 13 % increase from the monomer. However after an additional 6 hours of incubation at 65 °C, the average hydrodynamic radius saw an increase of 36 % from the monomeric form of HEWL 19.20 Å to 26.11 Å. Once the HEWL sample is incubated, the sample loses its homogeneity in size as the oligomers form. Therefore, the average increase in the hydrodynamic radius is evidence of the presence of a population of soluble oligomers arising with the incubation time. This information obtained indicates that diffusion NMR can be use as a detection tool for early stage protein aggregation/amyloid oligomerization.

Poly-Gaussian decay function with multiple components

As mentioned earlier, the mechanisms of early stage protein aggregation is still not fully

Table 3.3: Hydrodynamic radius of the native and incubated (6, 12, 18, and 24 hours) HEWL

Sample name	Rh (Å)
Native HEWL (monomer)	19.20 ± 0.44
HEWL 6 hr heated	19.96 ± 0.46
HEWL 12 hr heated	21.61 ± 0.50
HEWL 18 hr heated	21.76 ± 0.50
HEWL 24 hr heated	26.11 ± 0.60

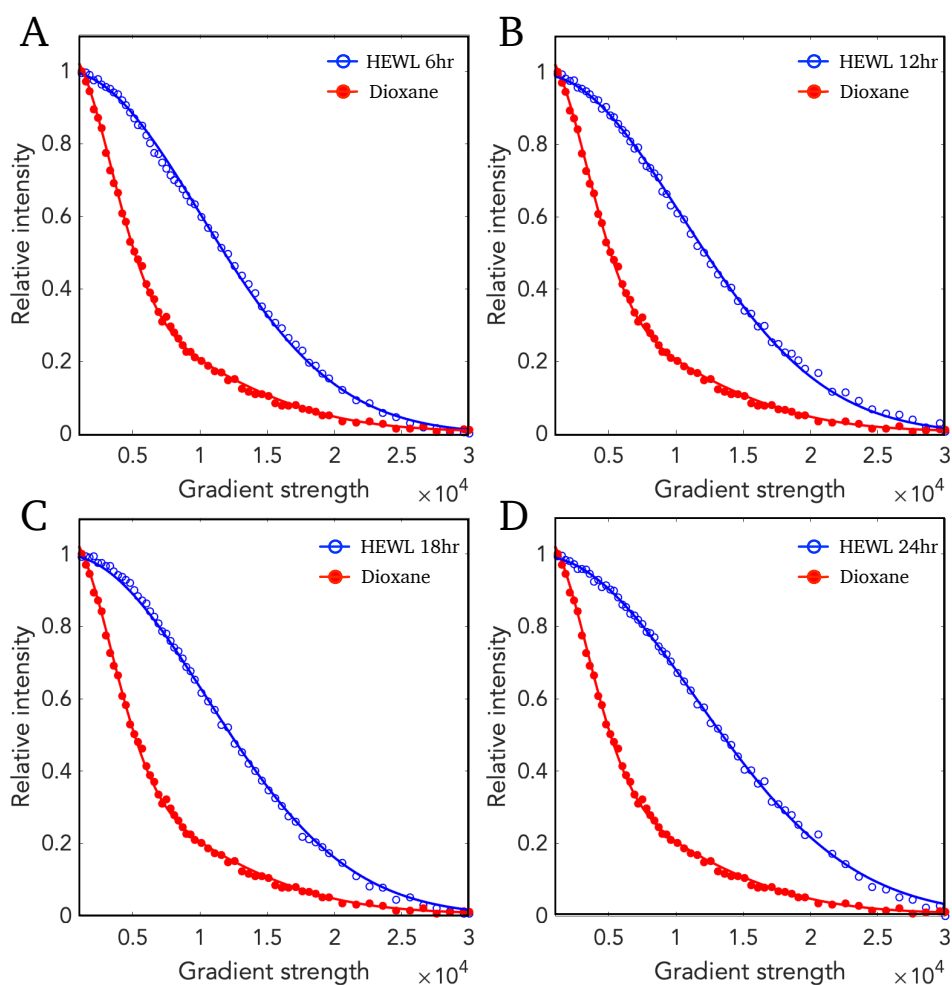


Figure 3.26: Plot of the normalised integrated intensity as a function of the gradient strength at the aliphatic region signal (3.0-0.0 ppm) of the proteins and the dioxane signal (3.75 ppm). The solid lines represent the best fit of a mono-Gaussian decay function with single-component to the experiment data. (Single diffusion component) HEWL incubated after (A) 6 hours (RMSE:0.0113), (B) 12 hours. (RMSE:0.0100), (C) 18 hours. (RMSE:0.0104), and (D) 24 hours. (RMSE:0.0100).

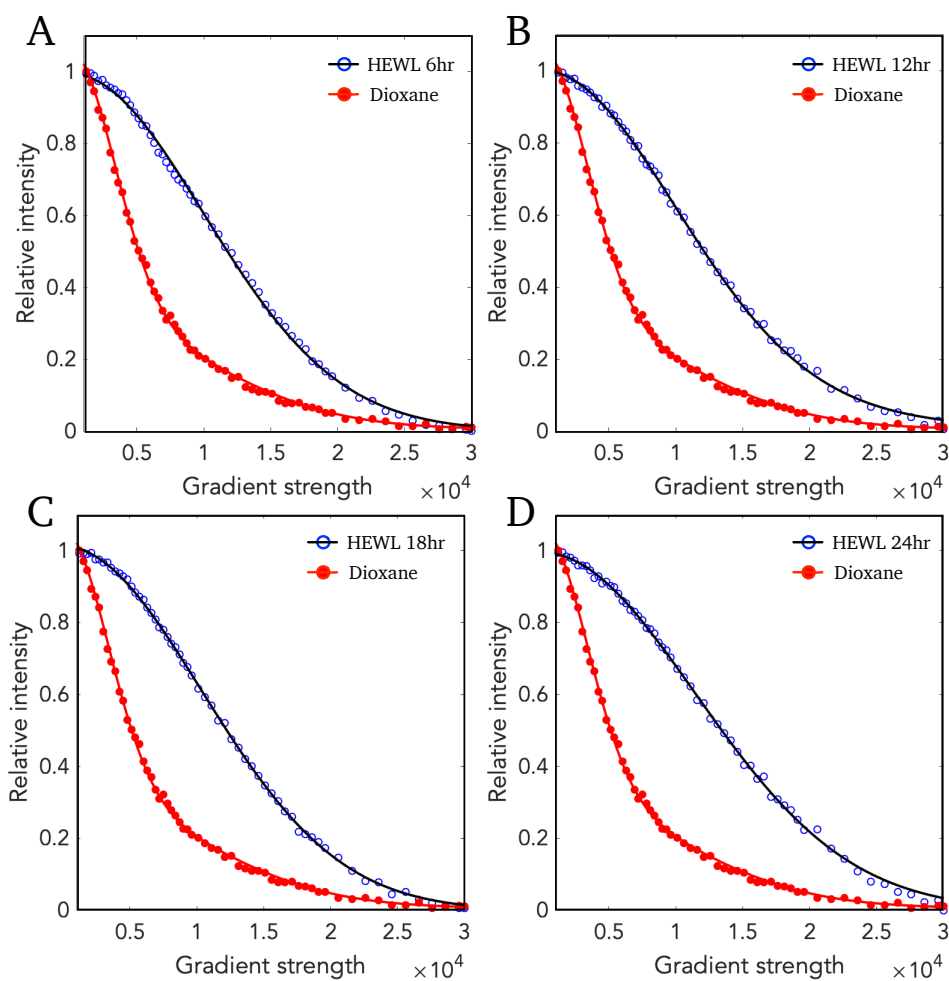


Figure 3.27: Plot of the normalised integrated intensity as a function of the gradient strength at the aliphatic region signal (3.0-0.0 ppm) of the proteins and the dioxane signal (3.75 ppm). The solid lines represent the best fit of a poly-Gaussian decay function with two-components to the experiment data. (Double diffusion components) HEWL incubated after (A) 6 hours (RMSE:0.0110), (B) 12 hours. (RMSE:0.0094), (C) 18 hours. (RMSE:0.0078), and (D) 24 hours. (RMSE:0.0118).

understood and is a challenge for biophysical techniques as they only probe average information about the population ensemble and cannot provide further information due to sample heterogeneity.

In the present section, the integrated spectral intensity of the aliphatic region signals over the increase of gradient strength was fitted by applying poly-Gaussian decay functions (Equation (2.4)) with multiple diffusing components as described in Section 2.12.2.

We estimated the hydrodynamic radii of the oligomer population presence in the sample through the poly-Gaussian decay function with two components are presented in Figure 3.27. In

Table 3.4: Estimated hydrodynamic radius and ratio of the two components in 6, 12, 18, and 24 hours incubated HEWL

HEWL Samples	1 st component (monomer)		2 nd component	
	Rh (Å)	Proportional percentage (%)	Rh (Å)	Proportion percentage (%)
HEWL 6 hr heated	19.2	65	21.85 ± 0.51	34.5
HEWL 12 hr heated	19.2	52.3	24.6 ± 0.57	47.1
HEWL 18 hr heated	19.2	59.8	24.99 ± 0.58	41
HEWL 24 hr heated	19.2	36.2	30.10 ± 0.69	64.2

the fitting process, the first component is assumed as monomer, therefore, the diffusion coefficient obtained from HEWL native sample fitting was set as the constant value. With obtained relevant diffusion coefficients of second component, hydrodynamic radius of second component was calculated with the method described in Appendix A.1. Also from the intensity of each component, the proportion percentage of monomer/oligomer population was estimated. Obtained value of hydrodynamic radius and proportion of components are displayed in the Table 3.4.

3.2.7.2 Morphology of early stage amyloid oligomerization of HEWL with AFM

Atomic force microscopy (AFM) was used to investigate the early stages of HEWL amyloid oligomer formation at pH 2. Samples were imaged over 24 hours at different time points (6 & 24 hours). From AFM, we were able to obtain 2D images of surface-absorbed HEWL oligomers acquired in liquid, as shown in Figure 3.28:B&C as well as the size distribution of oligomers after 6 and 24 hours of incubation. The solid line is the fitted distribution.

As the size distribution shows, even after 6 hours of incubation, oligomers with hydrodynamic radii larger than 4 nm have been formed, however, most of the oligomer population has hydrodynamic radii between 2-4 nm. After 24 hours of incubation, the proportion of oligomers with a hydrodynamic radius between 2-4 nm decreased while the population of oligomers with hydrodynamic radii larger than 4 nm increased.

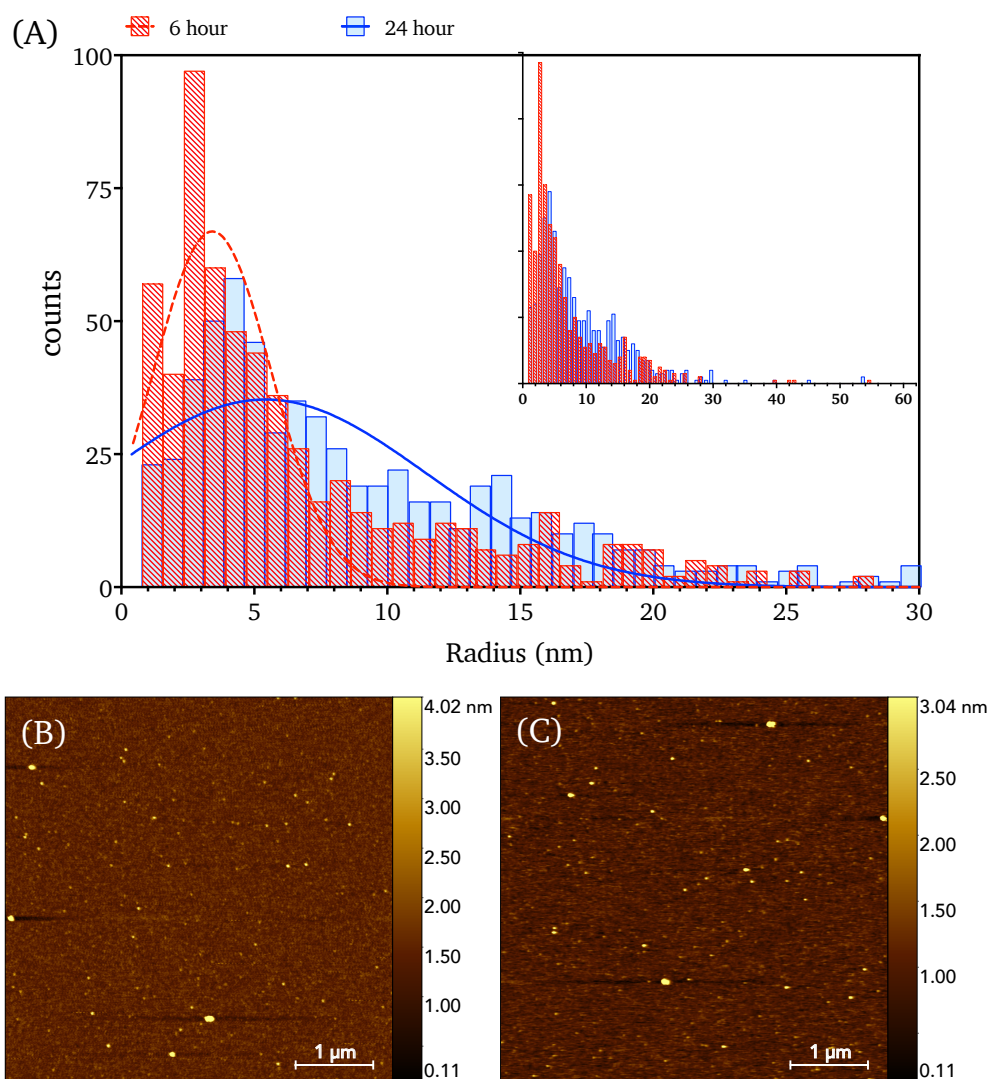


Figure 3.28: AFM result of HEWL 1 *mM* (pH 2) 6 hour and 24 hour incubated at 62 °C sample. (A) Particle radius (*nm*) distribution of 6 hours & 24 hours incubated HEWL. (B) AFM image of 6-hour incubated HEWL. (C) AFM image of 24-hour incubated HEWL.

3.3 Conclusion

Protein oligomers and fibrils have been intensively studied due to its relation to neurodegenerative disease as well as their potential nano-biomaterial applications. However, due to its insolubility and heterogeneous nature, the study of protein aggregates is difficult, and the mechanism of oligomerization/fibrilization is yet to be fully understood.

In this chapter, protein fibrillization and oligomerization were investigated with various

techniques such as ThT fluorescence, Congo-Red, CD, AFM, TEM, and diffusion NMR. Further to this, diffusion NMR is proposed as a potential supporting tool for estimating the hydrodynamic radius of oligomers and their proportional presence in the sample.

Firstly, the mixture of HEWL, Tyr, and Dioxane was used to show the capability of diffusion NMR in separating diffusing components (in this case, three components with different hydrodynamic radii) based on their hydrodynamic radius (Section 3.2.4). Diffusion NMR was then applied to estimate the proportion of bovine β -Lg monomer-dimer equilibrium at 1 and 12 % (w/w) under acidic and neutral pH conditions (Section 3.2.5). The hydrodynamic radii of monomer and the percentage proportion of monomer-dimer were calculated, and the obtained values corresponded well with the results presented in a recent study.^[120] Also, we were able to observe the presence of higher order β -Lg oligomer at pH 2 (where β -Lg is known to exist as a monomer) and at pH 7 when the concentration of protein was 12 % (w/w). The oligomerization of β -Lg at high concentration has been shown in previous studies.^[120–122]

Diffusion NMR was then applied to estimate the hydrodynamic radius of cross-linked HEWL-BS³ with different sizes (Section 3.2.6). The hydrodynamic radius of cross-linked species and the proportional percentage of each species was calculated through Gaussian decay function fitting method. The calculated hydrodynamic radius, proportion and fitting were similar to the hydrodynamic radius, proportion and fitting curve of the cross-linked species (monomer-dimer-trimer) estimated based on the SDS-PAGE and SE-HPLC results. Finally, the hydrodynamic radius of HEWL oligomers –formed after being incubated for different length of time (6, 12, 18, and 24 hours)– was estimated from diffusion NMR data through the Gaussian decay function method (Section 3.2.7). The decrease of diffusion decay rate over incubation time was detected, and the significant difference between native HEWL and 24 hour incubated HEWL was realised. The increase in overall hydrodynamic radius over incubation time was also recognised and diffusion components of each samples were estimated.

Overall, the obtained information indicates that diffusion NMR can be used as a useful tool for investigating the early stage protein aggregation. With additional development, we hope that diffusion NMR can contribute to the understanding of in-depth biophysical characteristics of protein

¹²² A. L. Jansma et al. *J. Biol. Chem.*, **285**: 14424–14437, 2010.

aggregation, which is considered to be essential in understanding and unravelling the mechanism of neurodegenerative disease, and for the development of therapeutic approaches in the future.

Chapter 4

The Effect of β -Lg Concentration on its Properties of Heat Denaturation

The thermal stability of milk/whey proteins is an important aspect to consider due to the inevitable but essential process of pasteurisation (application of heat in order to reduce bacterial load) during the production of milk/whey products. The globular structure of milk/whey proteins are known to begin to unfold and denature at temperatures greater than 64 °C. Such unfolding of the protein can expose reactive sites which are otherwise inaccessible within the native globular structure. The binding of two or more of these exposed reactive sites can lead to further aggregation of protein or create precipitation. This can be problematic during manufacture of milk/whey products such as infant formulations as these may lead to changes in the protein functionality (as defined by the dairy industry). Hence, controlling and understanding the thermal denaturation and aggregation kinetics of milk/whey proteins during the thermal process is important to maintain the functionality of the protein.

In this collaborative study, the effect of bovine β -lactoglobulin concentration on heat-induced molecular rearrangements were investigated using ^1H - ^{13}C HSQC NMR, while research at Teagasc Moorepark Food Research Institute was carried out using size exclusion high performance liquid chromatography (SE-HPLC) and FTIR. β -lactoglobulin concentrations of 1 % (w/v) and 12 % (w/v) were used, and protein samples were heated to 62 °C and 85 °C to clarify the correlation between protein denaturation and protein concentration.

From the ^1H - ^{13}C HSQC spectrums of native 1 % β -Lg, heat treatment at 62 °C for 20 hours, and β -Lg cooled down to room temperature from 62 °C, we were able to show that β -Lg did not denature upon heat treatment at 62 °C even after 20 hours of treatment time, which was also observed through SE-HPLC result.

The findings of pre-and- post 85 °C heat treatment suggests that while increasing the temperature resulted in increased protein aggregation, structural changes determined by NMR showed only minor changes in chemical shifts. Increasing the protein concentration from 1 % (w/v) to 12 % (w/v) resulted in more of these minor changes in the structure of protein as there were more chemical shift changes shown in the ^1H - ^{13}C HSQC spectra. This result indicates that although the β -lactoglobulin molecules were heated at 85 °C for 10 mins and formed multimers/aggregates, protein still behaved native-like.

From the results presented in this chapter, we were able to conclude that the current thermal-heating process method in the dairy industry is acceptable, as the proteins maintain their native-like structure.

4.1 Introduction

Proteins perform specific biological functions that strongly depend on their three-dimensional structure. Protein denaturation is the net effect of alterations in the biological, chemical, and physical properties of the protein when its structure is disrupted by various stresses such as temperature, pH, and pressure. The denaturation process is prominent when proteins are exposed to high temperatures in a variety of real-life situations predominantly in the food & dairy industry. Therefore, understanding the effect of thermal process on protein structure and biological activity is crucial. Multiple studies have been carried out to show the effect of high temperature on proteins such as, HEWL,^[123] β -lactoglobulin,^[39,124] and enzymes.^[125]

Based on the development of techniques in dairy science, bovine milk has become the main

¹²³ A. Dhulesia et al. *J. Am. Chem. Soc.*, **132**: 15580–15588, 2010.

³⁹ R.N. Zúñiga et al. *J. Food Sci.*, **75**: E261–E268, 2010.

¹²⁴ C. Bhattacharjee et al. *Protein J.*, **24**: 27–35, 2005.

¹²⁵ D. M. Roy, D. Mark, and H. H. Petach. *Biochem. J.*, **317**: 1–11, 1996.

source of protein for various milk products, especially Infant Milk Formula (IMF). Heating is an important process in the manufacture of many dairy products, and it is used to modify functional properties as well as to ensure food safety. Hence the physical stability of milk/whey protein during the thermal treatment is an important factor.^[126] Bovine milk contains whey proteins (β -lactoglobulin & α -lactalbumin), specific nutrients, minerals, and an amino acid profile that is required for growth and development of an infant. Whey proteins play an important role in determining the stability of IMF, thus, substantial research has been carried out investigating their properties.^[41] Bovine β -lactoglobulin (β -Lg) is the major whey protein that offers various functional properties such as gel and foam formation when protein unfolding and denaturation occurs by heating to a sufficient temperature.^[34] Therefore, understanding the thermal effect on β -Lg is important in the dairy industry, and the denaturation of β -Lg has been extensively studied under various conditions.^[127]

β -Lg is a β -sheet predominant protein and consists of 162 amino acids with a molecular mass of $\sim 18,400$ Da, with the β -barrel composed of eight antiparallel β -strands, one major α -helix at the C-terminal end of the molecule, and a final β -strand which forms part of the dimerization interface. At neutral pH, β -Lg resides in dimeric form, and each monomer has two intramolecular disulphide bonds (Cys 66 – Cys 160 and Cys 106 – Cys 119) and one free cysteine (Cys 121) which is buried within the protein structure on the β -strand H.^[35,36,128] The free sulfhydryl group of Cys 121 is a reactive site, which can bind to other free reactive Cys residues, and lead to the formation of covalently bonded aggregates, hence, eliminating the possibility of reversing the denaturation process.^[37]

Heat treatment of β -Lg leads to structural changes to the unfolded form from its native globular

¹²⁶ O. Hernell. Human Milk vs. Cow's Milk and the Evolution of Infant Formulas in: *Milk and Milk Products in Human Nutrition*. ed. by R. Clemens, O. Hernell, and K. Michaelsen. 17–28, Basel: Nestec Ltd., Vevey/S. Karger AG, 2011. DOI: [10.1159/000325572](https://doi.org/10.1159/000325572)

⁴¹ A. K. Buggy et al. *Dairy Sci. Technol.*, **96**: 845–859, 2017.

³⁴ N. Bansal and B. Bhandari. Functional Milk Proteins: Production and Utilization–Whey-based Ingredients in: *Advanced Dairy Chemistry*. ed. by P McSweeney and J O'Mahony. 67–98, New York, NY: Springer New York, 2016. DOI: [10.1007/978-1-4939-2800-2_3](https://doi.org/10.1007/978-1-4939-2800-2_3)

¹²⁷ S. M. Loveday. *Int. Dairy J.*, **52**: 92–100, 2016.

³⁵ S. Brownlow et al. *Structure*, **5**: 481–495, 1997.

³⁶ K. Sakurai and Y. Goto. *J. Mol. Biol.*, **356**: 483–496, 2006.

¹²⁸ T. Croguennec et al. *Biochem. Biophys. Res. Commun.*, **301**: 465–471, 2003.

³⁷ T. V. Burova et al. *Protein Eng. Des. Sel.*, **11**: 1065–1073, 1998.

form, further allowing for interactions between the unfolded state and other β -Lg monomers and/or dimers.^[39] Conformational changes in β -Lg monomers and dimers result in the exposure of one free sulfhydryl group and two intramolecular disulphide bonds, which can lead to further aggregation through sulfhydryl/disulphide exchange. During heating process at neutral pH, conformation changes occurs when the dimers dissociate into monomers, which form a reactive monomer with a partially disrupted helix which result in the exposure of the one free sulfhydryl group, two intramolecular disulphide bonds, as well as hydrophobic amino acids which can lead to further formation of larger species through sulfhydryl/disulphide exchange.^[41] In particular, the effect of free sulfhydryl group exposure during the thermal process was shown in the absence or presence of thiol-blocking agent NEM, and no protein aggregation was observed in the presence of NEM.^[128] Also preservation of the dimer association of β -Lg limits the exposition of buried residues including free cysteine,^[129] which again suggests that conservation of free reactive Cys is presumably reduces the possibility of protein aggregate formation. Most research has been carried out in relation to the structure of β -Lg at low concentration, whereby very few experiments carried out have shown the effect of protein concentration on the structure of the protein, particularly in post-thermal treatment.

The objective of this collaborative study is to determine if protein concentration has an effect on the size of whey protein aggregates formed after thermal treatment, and whether these differences in aggregated size, aggregate formation etc., has an effect on the thermal stability of a protein solution upon further heating with various techniques. The findings of this study should give further insight into the relationship between protein concentration, heat treatments, formation of aggregates and thermal stability during processing of whey products, including infant milk formula.

4.1.1 Protein HSQC NMR: characterisation of the bovine β -Lg stability

Heteronuclear Single Quantum Correlation (HSQC) NMR spectroscopy gives signals of correlation between two different types of nuclei usually proton with carbon (^1H - ^{13}C HSQC) or proton with nitrogen (^1H - ^{15}N HSQC), separated by one bond. The HSQC experiment is routinely used to show changes of the local structure in protein, because ^1H - ^{13}C HSQC and ^1H - ^{15}N HSQC chemical shifts

¹²⁹J. A. Seo et al. *J. Phys. Chem. B*, **114**: 6675–6684, 2010.

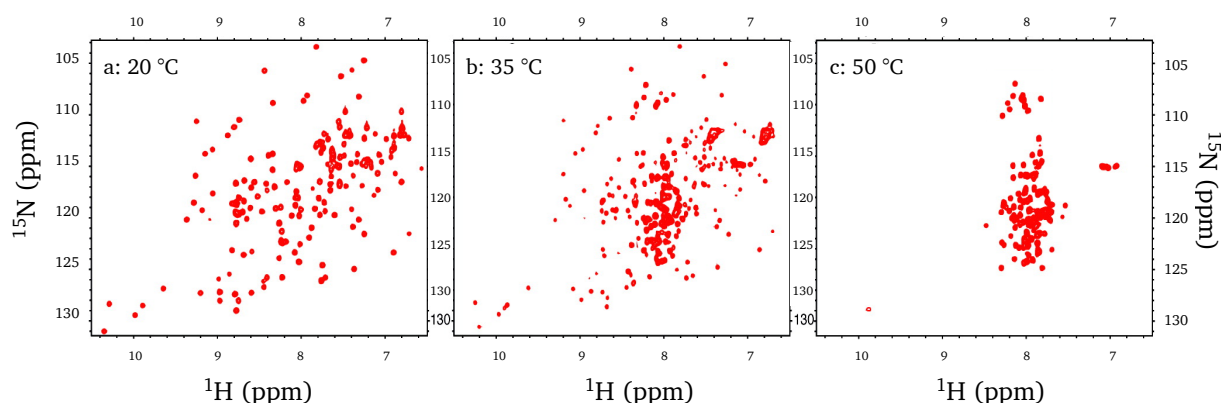


Figure 4.1: ^1H - ^{15}N HSQC spectra of HEWL (I59T) at pH 1.2 recorded at different temperatures. At 20 °C, the protein is fully folded (a), whereas at 50 °C, the protein is fully unfolded (c). Near the midpoint of the unfolding transition, at 35 °C, the protein shows resonances typical of both folded and unfolded states (b). (Adopted from Dhulesia *et al.*^[123])

are extremely sensitive to the neighbouring structure.^[79]

Due to its sensitivity, HSQC spectroscopy has been used in a number of studies to showcase protein structural conformation changes, including the monitoring of native & unfolding states of Phospho-Tyrosine-binding (PTB) domain. These include, Dok1 protein under various urea concentration,^[130] the pH-dependent conformational transitions of bovine β -Lg,^[36] as well as thermal denaturation of human lysozyme.^[123] Figure 4.1 is the typical example which shows the difference in chemical shifts between the folded protein and unfolded protein.^[123]

In this chapter, ^1H - ^{13}C HSQC was used to detect the denaturation of bovine β -Lg during the thermal treatment.

4.2 Results & Discussion

4.2.1 NMR temperature calibration using 1,2-ethanediol

Because heat treatment on proteins at specific temperatures is essential in this experiment, the NMR probe was calibrated before performing the actual experiments with the β -Lg sample. The real temperature within the NMR probe was measured using 1,2-ethanediol (see Figure 4.2 and Table

⁷⁹ J. Cavanagh *et al.* *Protein NMR Spectroscopy: Principles and Practice*. 2nd Amsterdam: Elsevier, 2007. DOI: [10.1016/B978-0-12-164491-8.X5000-3](https://doi.org/10.1016/B978-0-12-164491-8.X5000-3)

¹³⁰ S. Gupta and S. Bhattacharjya. *PLOS ONE*, **9**: e90557, 2014.

Table 4.1: Calibrated temperature using 1,2-Ethenediol compared to temperature displayed on the NMR temperature control panel.

Temp. on NMR control panel	delta ppm	Real temp.	Real Temp.	Temp. Difference	Temp. Difference
[°C]	[ppm]	[K]	[°C]	[°C]	[%]
25	1.513	296.51	23.35	-1.65	7
30	1.47	301.16	28	-2	7
35	1.428	305.71	32.55	-2.45	7
40	1.385	310.37	37.21	-2.79	7
45	1.338	315.46	42.3	-2.7	6
50	1.294	320.23	47.07	-2.93	6
55	1.248	325.21	52.05	-2.95	5
60	1.204	329.98	56.82	-3.18	5
61	1.194	331.06	57.9	-3.1	5
62	1.187	331.82	58.66	-3.34	5
63	1.178	332.8	59.64	-3.36	5
64	1.169	333.77	60.61	-3.39	5
65	1.16	334.75	61.59	-3.41	5
66	1.151	335.72	62.56	-3.44	5
67	1.141	336.81	63.65	-3.35	5
68	1.132	337.78	64.62	-3.38	5
69	1.122	338.86	65.7	-3.3	5
70	1.114	339.73	66.57	-3.43	5

4.1) and based on these, the variable temperature (VT) unit was set accordingly. 1,2-ethenediol is a molecule that exhibits an easily identifiable relationship of the chemical shifts of its protons to the temperature of the environment. The peak shift differences of OH–CH₂ decreases as the sample temperature increases due to the reduced intermolecular hydrogen-bonding.^[91] Our results show that real temperature in the cold probe is 5~7 % lower than the temperature value readings on the control panel. In this study, the VT unit was set as 65.5 °C to achieve the required temperature, 62 °C.

⁹¹ T. D.W. Claridge. Diffusion NMR Spectroscopy in: *High-Resolution NMR Techniques in Organic Chemistry*. 3rd ed. chap. 10, 381–419, Oxford: Elsevier, 2016. DOI: [10.1016/B978-0-08-099986-9.00010-5](https://doi.org/10.1016/B978-0-08-099986-9.00010-5)

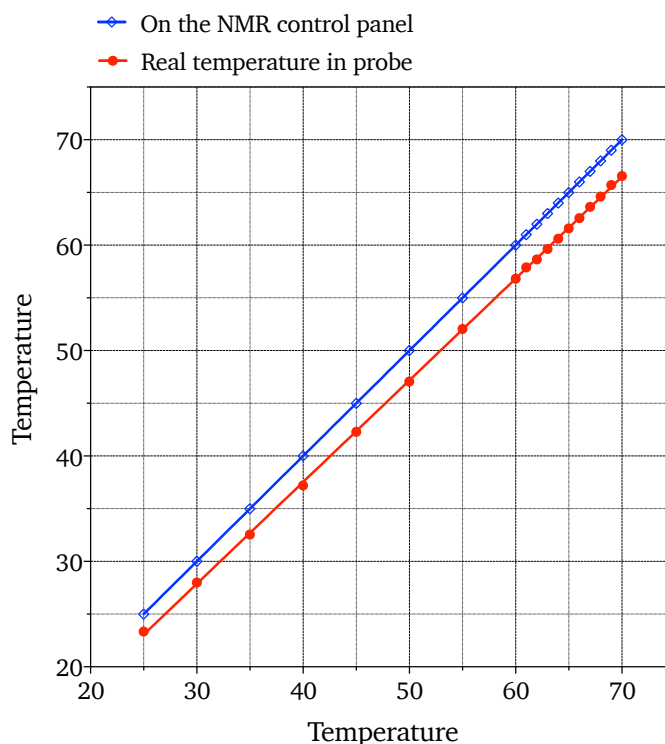


Figure 4.2: NMR Temperature calibration using 1,2-Ethanediol. The blue line represents the temperature displayed on the NMR temperature control panel, and the red line represents the measured real temperature of the NMR probe.

4.2.2 Investigation of the effect of β -Lg concentration on denaturation and subsequent aggregation

4.2.2.1 Size exclusion HPLC (SE-HPLC) & reversed phase HPLC (RP-HPLC)

Size exclusion HPLC (SE-HPLC) was used to observe any correlation between protein aggregate formation and temperature. β -Lg 1 % samples were heated for a period of 20 hours between 62 °C and 70 °C (pH 6.7) (Figure 4.3). Results show that when the temperature of the heat treatment is increased above 62 °C, more dimeric, oligomeric and aggregated proteins are formed. However, at 62 °C, no oligomeric / aggregated proteins were observed.

Size exclusion HPLC (SE-HPLC) was then used to monitor the relationship between protein aggregate formation and concentration of protein (Figure 4.4). No significant changes were noticed in the amount of monomeric and dimeric protein in both 1 % β -Lg heat-treated at 62 °C for 1200 minutes (= 20 hours) and 12 % β -Lg heat-treated at 62 °C for 180 minutes. This result correlates

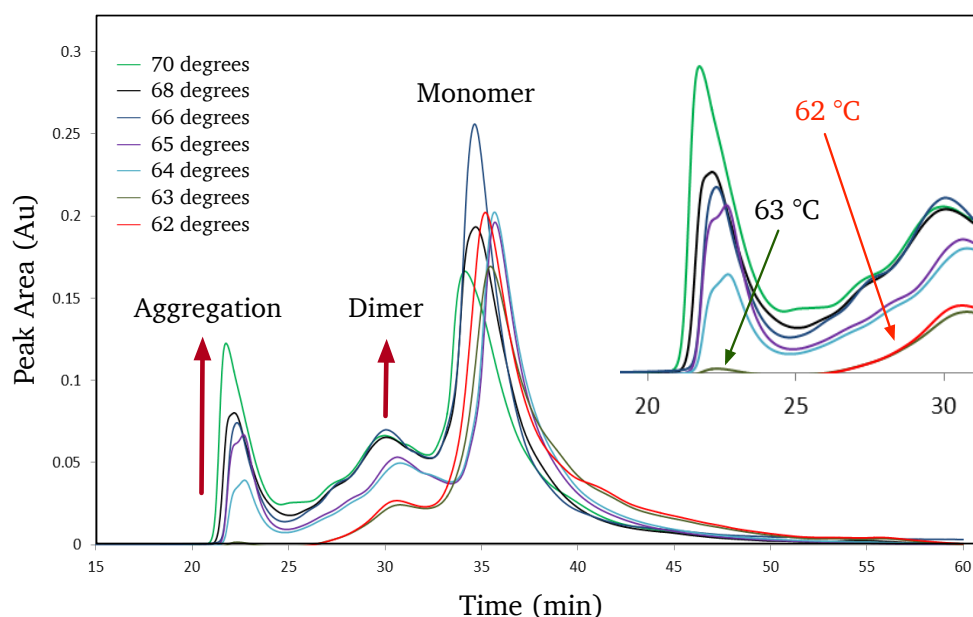


Figure 4.3: Heat treatment at different temperatures between 62 °C and 70 °C on 1 % β -Lg (pH 6.7) for 20 hours. (Data provided by Aoife K. Buggy, Teagasc Food Research Centre, Moorepark, Co. Cork, Ireland.)

with the results shown in Figure 4.2, indicating the heat treatment at 62 °C was deemed to be too low to induce protein denaturation, and therefore also too low to induce subsequent protein aggregation.

In contrast, both 1 % and 12 % β -Lg samples heat-treated at 85 °C showed significant protein aggregation within 5 minutes of the heating process. After the heat treatment at 85 °C for 5 minutes, reduction in the quantity of monomeric and an increase of dimeric forms of the 1 % β -Lg sample was noted (approximately 11 % of aggregated protein). The percentage of oligomeric/aggregated protein had increased significantly to 82 % after 10 minutes of the heating process. It is clearly shown that the rate of formation of aggregated protein for the equivalent 12 % sample, heated at 85 °C, were extensively aggregated at 97 % and 98 % respectively, and this indicates that there was no significant difference in protein aggregation shown after the additional 5 minutes of heating. Previous research has demonstrated that the β -Lg concentration of a sample has an effect on the rate of denaturation, where an increase in denaturation is noticed with a higher protein concentrations.^[131]

¹³¹ D. J. Oldfield, H. Singh, and M. W. Taylor. *J. Dairy Res.*, **72**: 369–378, 2005.

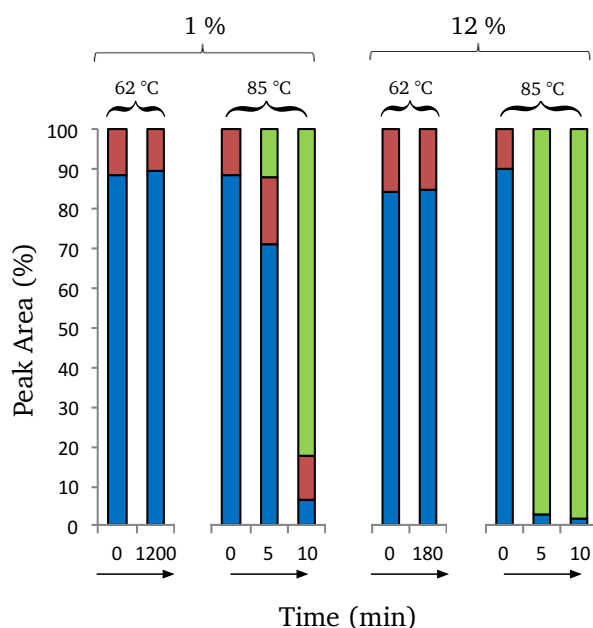


Figure 4.4: SE-HPLC: Size exclusion chromatography distribution of 1 and 12 % (w/v)- β -lactoglobulin; monomeric (Blue), dimeric (Red) and oligomeric/aggregated (Green) protein peak area (%) at 62 °C (0~1200 mins and 0~180 mins) and 85 °C (0, 5, 10 mins). (Data provided by Aoife K. Buggy, Teagasc Food Research Centre, Moorepark, Co. Cork, Ireland)

4.2.2.2 NMR: ^1H - ^{13}C Heteronuclear Single Quantum Correlation (HSQC)

Heat treatment at 62 °C

Each peak in the ^1H - ^{13}C HSQC NMR spectrum represents the specific link between carbon and hydrogen, and the chemical shifts in both carbon and proton dimensions are extremely sensitive to the surrounding chemical environment of two nuclei, as well as the physical conditions such as temperature and pH value.

Figure 4.5 shows the overlaid ^1H - ^{13}C HSQC NMR spectrum of 1 % native β -Lg (Figure 4.5: A), the 1 % β -Lg heat treated at 62 °C for 20 hours (Figure 4.5: B, The temperature was set at 62 °C for the duration of the ^1H - ^{13}C HSQC measurement), and the cool-downed 1 % β -Lg heat treated at 62 °C for 20 hours then further cooled down to 20 °C (Figure 4.5: C). The peaks (i – 4.85 ppm, 20 °C: red), (ii – 4.45 ppm, 62 °C: orange), and (iii – 4.85 ppm, 20 °C: cyan) are the artefact from the water signal. The chemical shift value of water is directly reflected to the temperature of the sample and it is good indicator to check if temperature of sample is decreased to 20 °C from 62 °C.

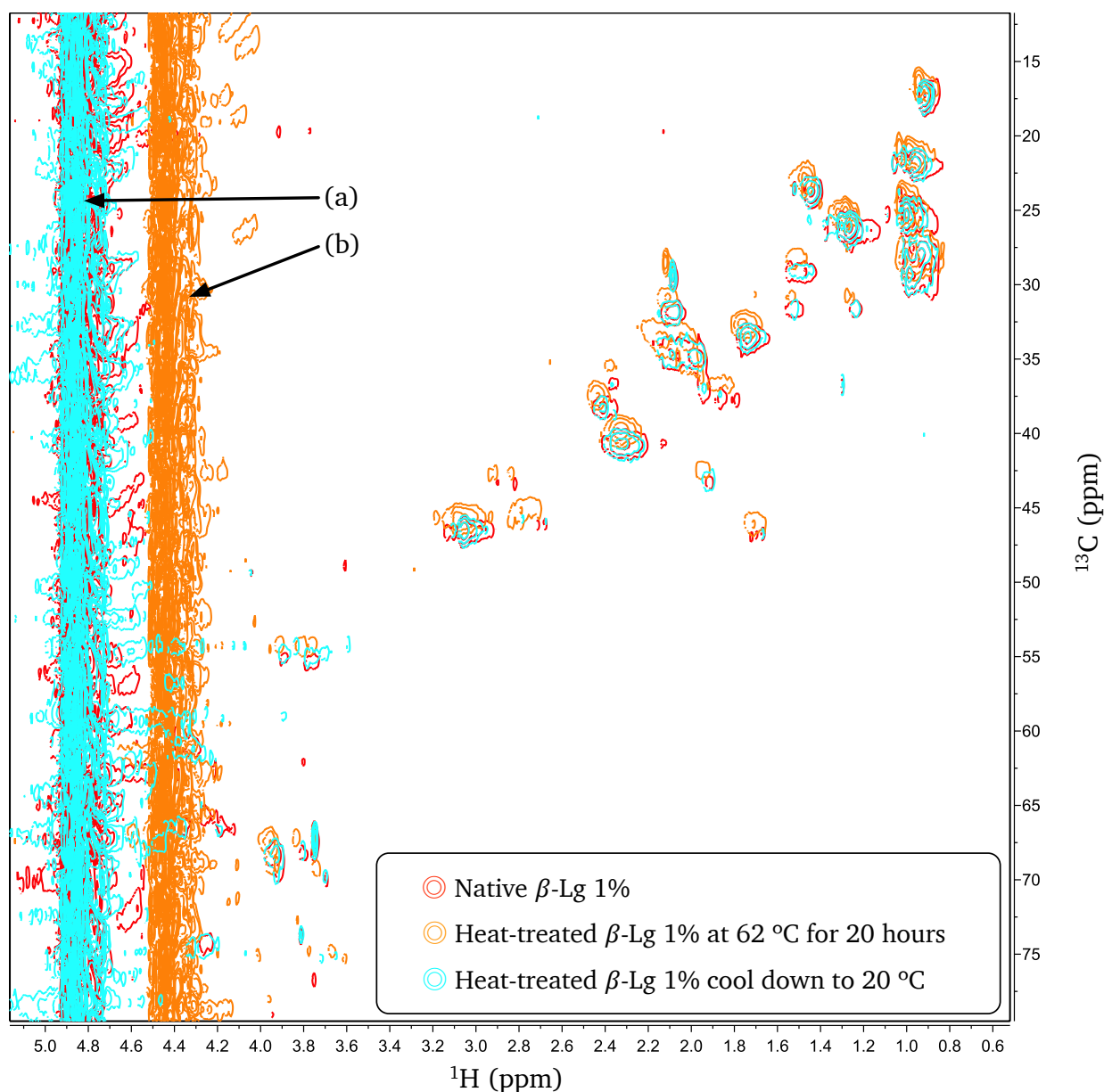


Figure 4.5: ^1H - ^{13}C HSQC of native β -Lg 1 % (A, Red), the β -Lg 1 % heat treated at 62 °C for 20 hours (NMR spectrum acquired at 62 °C) (B, Orange), and the β -Lg heat treated at 62 °C for 20 hours then cooled to 20 °C (C, Cyan). (i), (ii), and (iii): water peaks)

There are only minor changes after a 20-hour heat treatment at 62 °C (B), however, even after the heating process, there is no significant chemical shift changes monitored between the control sample (A) and the cooled sample (C). This suggests that heat treatment at 62 °C did not induce protein denaturation and only minor chemical shift changes can be observed in sample (B), as the unfolded proteins appeared to reversibly renature to the native form (A & C).

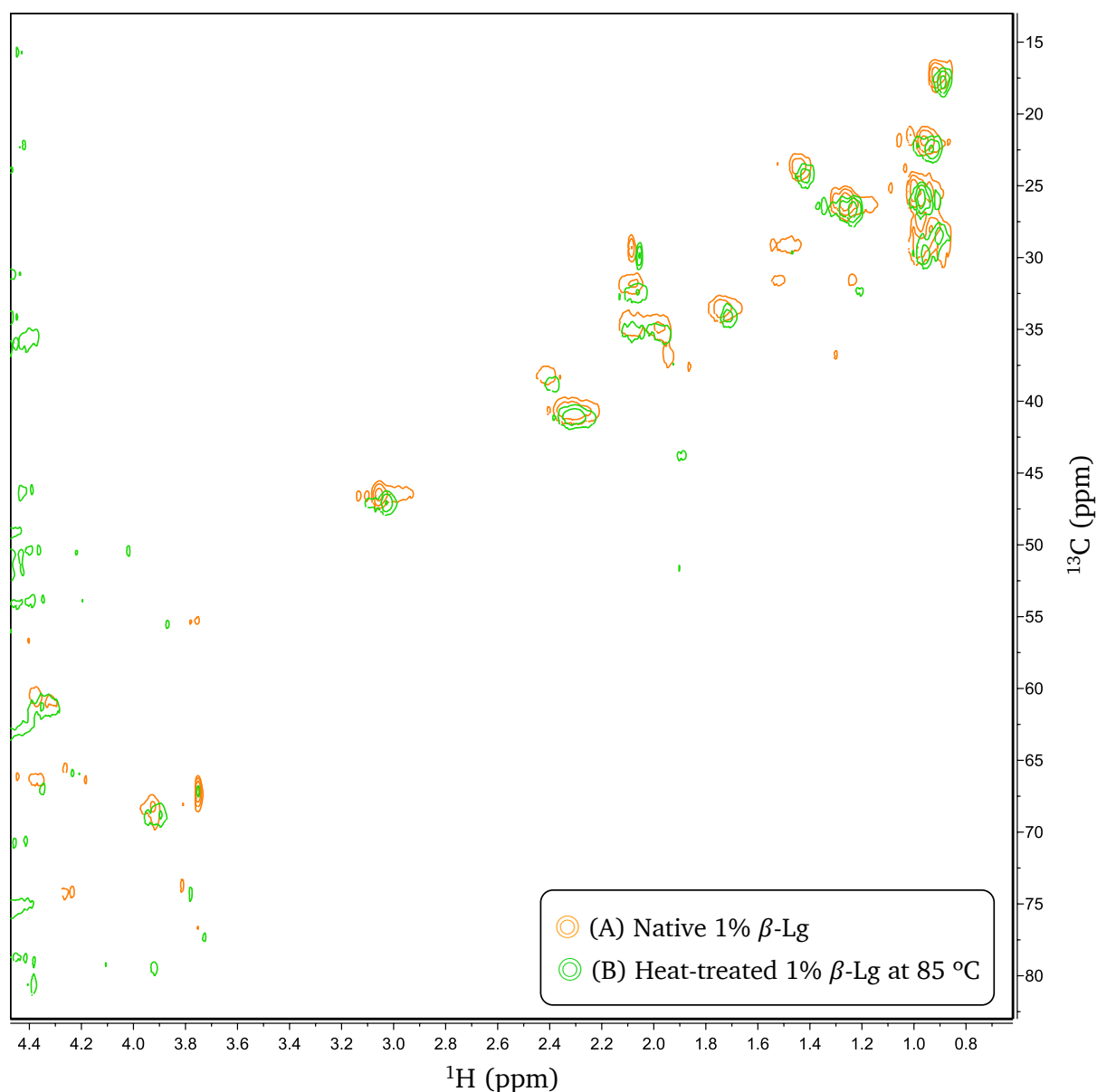


Figure 4.6: Overlaid ^1H - ^{13}C HSQC spectrum of Native 1 % β -Lg (Orange), Heat-treated 1 % β -Lg at 85 °C (Green)

Heat treatment at 85 °C

The ^1H - ^{13}C HSQC NMR was then used to determine whether the heat treatment at 85 °C has an effect on the structure of β -Lg protein, and to determine the correlation between the concentration of protein and protein structure in association with SE-HPLC spectroscopy results (Figure 4.4). The ^1H - ^{13}C HSQC NMR experiment was carried out on native 1 % and 12 % β -Lg as well as on the heated 1 % and 12 % β -Lg at 85 °C for 10 mins.

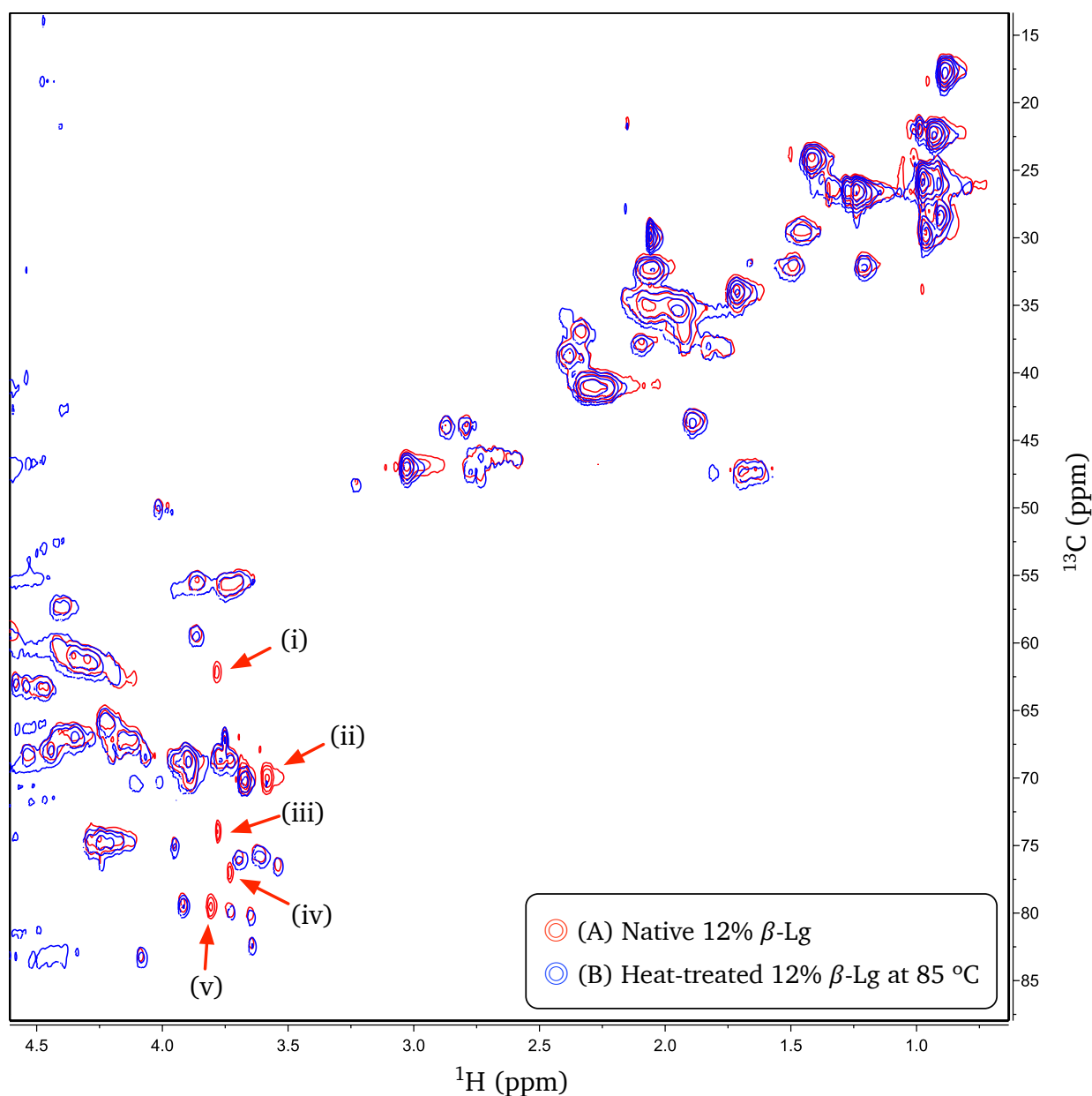


Figure 4.7: Overlaid ^1H - ^{13}C HSQC spectrum of Native 12 % β -Lg (Red), Heat-treated 12 % β -Lg at 85 °C (Blue).

Overlaid ^1H - ^{13}C HSQC NMR spectrum of native 1 % β -Lg (Figure 4.6:A) with 1 % β -Lg heat-treated at 85 °C (Figure 4.6: B) is shown in Figure 4.6. There were only minor changes in the spectrum between the control sample and 85 °C heat-treated 1 % β -Lg sample. As mentioned earlier in Section 4.1.1, the chemical shifts of HSQC reflect sensitively on protein structural changes occurring during the heat treatment process (Figure 4.1). This suggests that the structure of protein has not denatured even after the heat treatment at 85 °C at 1 % concentration.

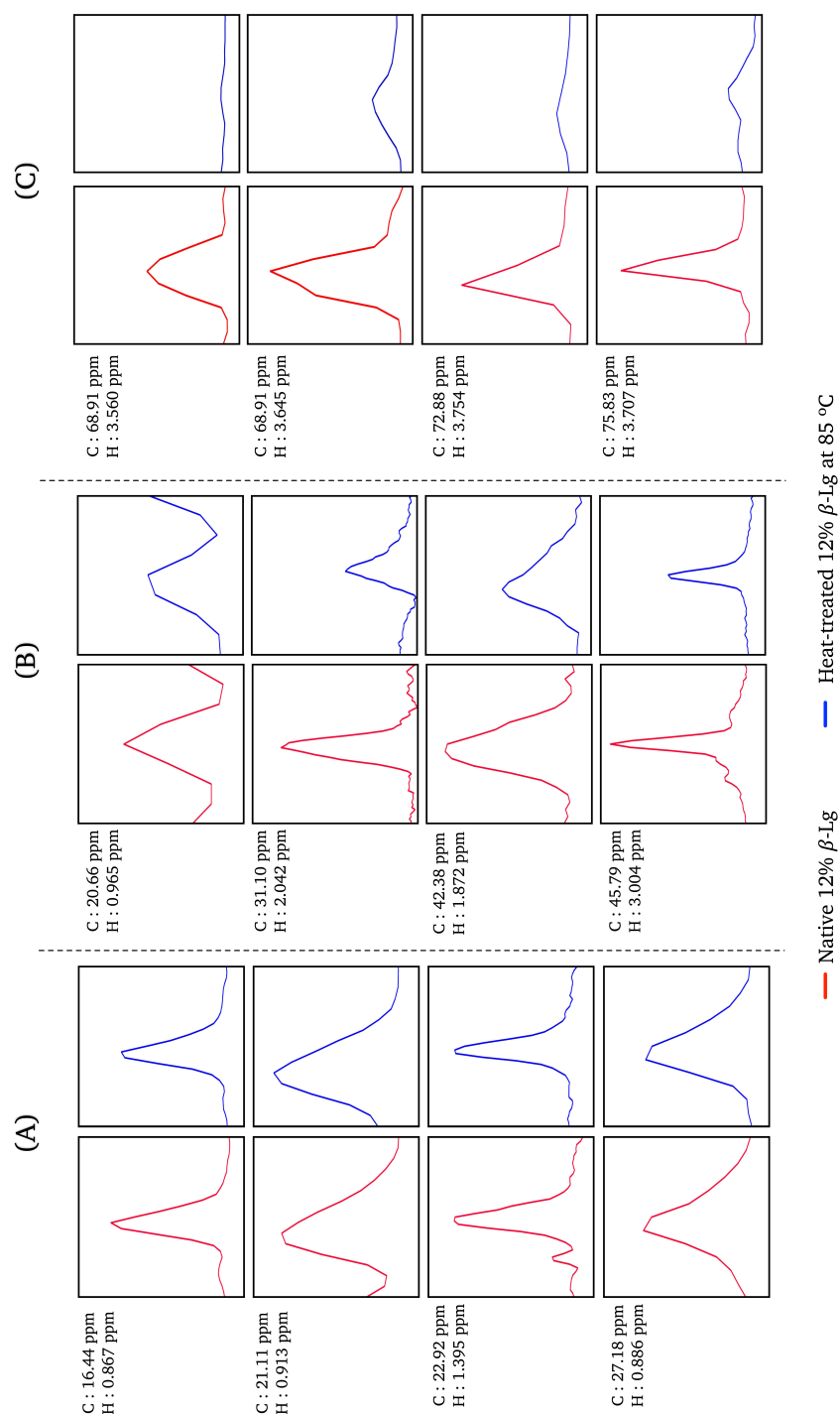


Figure 4.8: (A): One-dimension traces of ^1H - ^{13}C HSQC peaks of Native 12 % β -Lg (Red), Heat-treated 12 % β -Lg at 85 °C (Blue). The intensity of peaks remains fully equivalent even after the heat treatment. (B): One-dimension traces of ^1H - ^{13}C HSQC peaks of Native 12 % β -Lg (Red), Heat-treated 12 % β -Lg at 85 °C (Blue). The intensity of peaks remains even after the heat treatment. (C): One-dimension traces of ^1H - ^{13}C HSQC peaks of Native 12 % β -Lg (Red) and peaks of 12 % β -Lg which disappeared after heat treatment at 85 °C (Blue).

The overlaid ^1H - ^{13}C HSQC NMR spectrum of native 12 % β -Lg (Figure 4.7:A) and 12 % β -Lg heat-treated at 85 °C (Figure 4.7: B) is shown in Figure 4.7. We are able to see that most of the peaks remain at the same chemical shift even after 85 °C heat treatment. The one-dimensional traces of these remaining peaks are shown in Figure 4.8. However, there were noticeable absences of five peaks (Figure 4.7: (i), (ii), (iii), (iv), and (v)) which only exists in native 12 % β -Lg HSQC (Figure 4.7: A) and which disappeared after heat treatment at 85 °C. In the case where peaks are diminished, the one-dimension traces of these peaks are shown as in Figure 4.7.

4.3 Conclusion

At neutral pH, NMR peak broadening occurs due to monomer-dimer equilibrium, limiting NMR measurements of wild-type β -Lg at neutral pH value. There are many NMR studies carried out at acidic pH condition where monomeric formation is preferred relative to dimeric formation,^[38,132] along with mutant β -Lg which does not give an HSQC signal broadening effect even when it forms a dimer.^[36] Although these conditions have their own merits, they are quite irrelevant to the physical condition actually used in the dairy industry. Similar to the kinetics study of β -Lg heat denaturation described by Loveday S. M.,^[127] there are multiple studies carried out under various conditions with variants such as temperature, pH, as well as buffers, which are then investigated by a number of physical/chemical methods including NMR, FT-IR, CD, and SE-HPLC. However, so far, the NMR study of wild-type β -Lg at neutral pH value with high concentration (12 %) upon thermal-treatment at 85 °C has not been done yet.

In this thesis chapter, we carried out ^1H - ^{13}C HSQC NMR combined with chromatography techniques to determine the effect of heat treatment at 85 °C on 1 % and 12 % β -Lg. Most of the chemical shifts remained even after the heat treatment. However, there were five noticeable absence of peaks after the heat treatment at 85 °C for 10 mins. Except for those five peaks determined, the observed chemical shift of ^1H - ^{13}C HSQC spectrum for both 1 % and 12 % β -Lg at pH 6.7 remain more-or-less intact. This result indicates that β -Lg did not denature or unfold to random coil during the thermal treatment conditions of the experiment. Hence β -Lg remained in native-like

³⁸ K. Kuwata et al. *Protein Sci.*, **8**: 2541–2545, 2008.

¹³² F. Fogolari et al. *FEBS Lett.*, **436**: 149–154, 1998.

structure upon cooling to room temperature after being heated at 85 °C. The result is similar to those observed by Bhattacharjee *et al.*,^[124] Qi *et al.*,^[133] and Casal *et al.*,^[134] indicating that β -Lg preserves much of its structure even when it is heated to 90 °C.

As mentioned in Section 1.1.2, protein denaturation and protein unfolding occur under various conditions includes generic mutations, physical or chemical stress, temperature, and pH values. When a protein goes through these conditions, different forms of protein aggregates are observed. This is usually a disordered aggregate or an ordered aggregate form, also known as an amyloid fibril.^[6,7] In the previous chapter, amyloid fibril formed by β -Lg protein was presented, where β -Lg unfolded and lost its secondary structures and refolded to β -sheet rich structure and became an amyloid fibril. In contrast, in this chapter, the same protein β -Lg aggregated while remaining in its native-like structure. This demonstrates that regardless of the same protein, depending on the condition, the outcome of the protein aggregate can be different. Some examples of this include, wild-type β -Lg heated at 85 °C , and sickle haemoglobin (HbS), both of which maintain its native-like structure even when they form into a fibre under deoxygenated conditions.^[135] (Appendix A.4)

It follows that if the desired outcome is to keep the integrity of the overall protein structure while producing high-concentrated milk protein solution in the dairy industry, then the current method of heating is acceptable, as the resulting proteins are not denatured to a large extent and remains as native-like structure.

¹³³ X. L. Qi *et al.* *Biochem. J.*, **324**: 341–346, 1997.

¹³⁴ H. L. Casal, U. Köhler, and H. H. Mantsch. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.*, **957**: 11–20, 1988.

⁶ C. M. Dobson. *Nature*, **426**: 884–890, 2003.

⁷ C. M. Dobson. *Semin. Cell Dev. Biol.*, **15**: 3–16, 2004.

¹³⁵ F. A. Ferrone, J. Hofrichter, and W. A. Eaton. *J. Mol. Biol.*, **183**: 611–631, 1985.

Chapter 5

Determination of the solution structure of 9R-VIPER and its inactive mutant (9R-VIPER L6AE10A) with NMR Spectroscopy

Toll-like receptors (TLRs) play a role in viral detection leading to cytokine and IFN induction, and as such they are targeted by viruses for immune evasion. The poxviral protein A46 has been identified to inhibit TLR signalling by interacting with TIR domain-containing proteins of the receptor complex to collectively inhibit all TLR adaptor proteins that positively regulate transcription-factor activation. One 11 *aa* peptide (KYSFKLILAEY) termed Viral Inhibitory Peptide of TRL4 (VIPER) was reported to retain the inhibitory properties of full length A46 against TLR4 signalling. A 9R homo-polymer delivery sequence at the C-terminus provided delivery of the peptide into cells. 9R-VIPER showed its efficacy in a TLR4-dependent flu induced lung damage model. Structural comparisons are presented between 9R-VIPER, which is active in preventing TLR4-dependent cytokine induction in cell culture, and a mutant that exhibited loss of function (9R-VIPER L6AE10A) through solution NMR spectroscopy. Assignments of the resonances were accomplished in ^1H - ^{15}N HSQC, ^1H - ^{13}C HSQC, TOCSY and NOESY with an 800 *MHz* NMR. Chemical shift assignment and Nuclear Over-Hauser Effect (nOe) results were used to probe peptide conformation and generate 3D structures. We find that both peptides are not fully folded but have intermittent helical-like turns, and despite a relatively minor sequence difference, the loss of hydrophilicity and negative electrostatic interactions, as well as structure changes, result in subtle but potentially significant

differences in the region of the peptide proposed to interface with TLR4.

5.1 Introduction

5.1.1 Toll-like Receptors (TLRs)

Viruses are obligate parasites that rely on host proteins for their sustaining own life cycle and require strategies to evade and modulate the immune response. Toll-like Receptor (TLR) have an anti-viral role in detecting viruses that lead to cytokine and IFN induction in the innate immune system. The detection happens through Pathogen Recognition Receptor (PRR)s by responding to Pathogen Associated Molecular Patterns (PAMP)s.^[136] Among the PRRs, the TLRs are well-studied, and each set of TLRs respond to particular pathogen components and initiates the production of pro-inflammatory cytokines and chemokines.^[137] Signal transduction by TLRs requires the interaction of particular signalling adaptors such as myeloid differentiation primary-response gene 88 (MyD88)^[138] and TIR domain-containing adapter inducing IFN- β (TRIF).^[139]

TLR5, TLR7, TLR8 and TLR9 use the signal adaptor protein MyD88. TLR3 depends on TRIF for signal transduction. The heterodimers of TLR2/1 and TLR2/6 utilise both MyD88 and Mal adaptors for functional signalling. TLR4 is unique, whereby it is engaged by both MyD88-dependent and TRIF-dependent signalling pathways making TLR4 the most important member of TLR protein family.^[140] Kagan *et al.*^[141] proposed that TLR4 engages these two signal transducers at distinct cellular locations, sequentially, such that signalling of MyD88 is at the plasma membrane and TRIF signalling is in the endosomes. A schematic model of TLR4 and TLR2 signalling is shown in Figure 5.1.

Many viruses are capable of interfering with innate signalling pathways through intracellular events.^[142] These mechanisms can function to inhibit host signalling to prevent viral detection and,

¹³⁶ S. Akira, S. Uematsu, and O. Takeuchi. *Cell*, **124**: 783–801, 2006.

¹³⁷ T. Kawai and S. Akira. *Immunity*, **34**: 637–650, 2011.

¹³⁸ R. Medzhitov *et al.* *Mol. Cell*, **2**: 253–258, 1998.

¹³⁹ M. Yamamoto *et al.* *J. Immunol.*, **169**: 6668–6672, 2002.

¹⁴⁰ J. A. Owen *et al.* *Kuby Immunology*. 7th ed. New York, NY: W.H. Freeman, 2013.

¹⁴¹ J. C. Kagan *et al.* *Nat. Immunol.*, **9**: 361–368, 2008.

¹⁴² A. G. Bowie and L. Unterholzner. *Nat. Rev. Immunol.*, **8**: 911–922, 2008.

in some cases, promote inflammatory responses to facilitate viral dissemination.

5.1.2 VACV protein A46

VACV protein A46 was one of the earliest characterised viral inhibitors of TLR signalling.^[143–146] Hutchens M. A.^[147] found that inhibition of TLR4 signalling increased viral replication, hypothermia, and mortality. A46 is capable of inhibiting the TLR4 signalling of both MyD88-dependent and TRIF-dependent pathways. Stack J. and Bowie A. G.^[148] showed that A46 antagonises MyD88-dependent signalling by inhibiting TLR4-Mal interaction, as well as TRIF-dependent signalling by inhibiting TLR4-TRAM interactions.

5.1.3 VIPER: inhibitory peptide derived from A46

The inhibitory peptide derived from VACV A46 has been determined to retain the inhibitory properties of TLR4 pro-inflammatory signalling (Stack et al. 2005) and hence prevent receptor-adaptor interactions, whilst having a TRAM-specific interaction motif.^[148] Due to the ability of A46 to inhibit all TLRs,^[148] different variations of peptide fragments derived from this protein sequence was tested for its ability to inhibit TLR signalling. One inhibitory 11 *aa* peptide (KYSFKLILAEY), termed VIPER (Viral Inhibitory Peptide of TLR4), was reported to retain the properties of full length A46 against TLR4 signalling.^[149] Although A46 inhibits multiple TLRs by interacting with MyD88, TRIF, and MAL, VIPER was shown to specifically inhibit only TLR4 and not IL-1 nor other TLRs.^[148] VIPER represents the first example of a short virally derived peptide inhibitor of a specific TLR signalling pathway. Not only does this offer insight into mechanisms of TLR signalling, but may provide therapeutic benefit since VIPER inhibits TLR4 signalling at much lower concentrations and with a lower molecular weight than comparable BB loop peptides.^[150–152] TLR4 has shown its

¹⁴³ A. Bowie et al. *Proc. Natl. Acad. Sci.*, **97**: 10162–10167, 2000.

¹⁴⁴ I. R. Haga and A. G. Bowie. *Parasitology*, **130**: S11, 2005.

¹⁴⁵ G. L. Smith et al. *J. Gen. Virol.*, **94**: 2367–2392, 2013.

¹⁴⁶ J. Stack et al. *J. Exp. Med.*, **201**: 1007–1018, 2005.

¹⁴⁷ M. A. Hutchens et al. *PLOS Pathog.*, **4**: e1000153, 2008.

¹⁴⁸ J. Stack and A. G. Bowie. *J. Biol. Chem.*, **287**: 22672–22682, 2012.

¹⁴⁹ T. Lysakova-Devine et al. *J. Immunol.*, **185**: 4261–4271, 2010.

¹⁵⁰ W. Piao et al. *Proc. Natl. Acad. Sci.*, **110**: 19036–19041, 2013.

¹⁵¹ W. Piao, S. N. Vogel, and V. Y. Toshchakov. *J. Immunol.*, **190**: 2263–2272, 2013.

¹⁵² V. Y. Toshchakov, M. J. Fenton, and S. N. Vogel. *J. Immunol.*, **178**: 2655–2660, 2007.

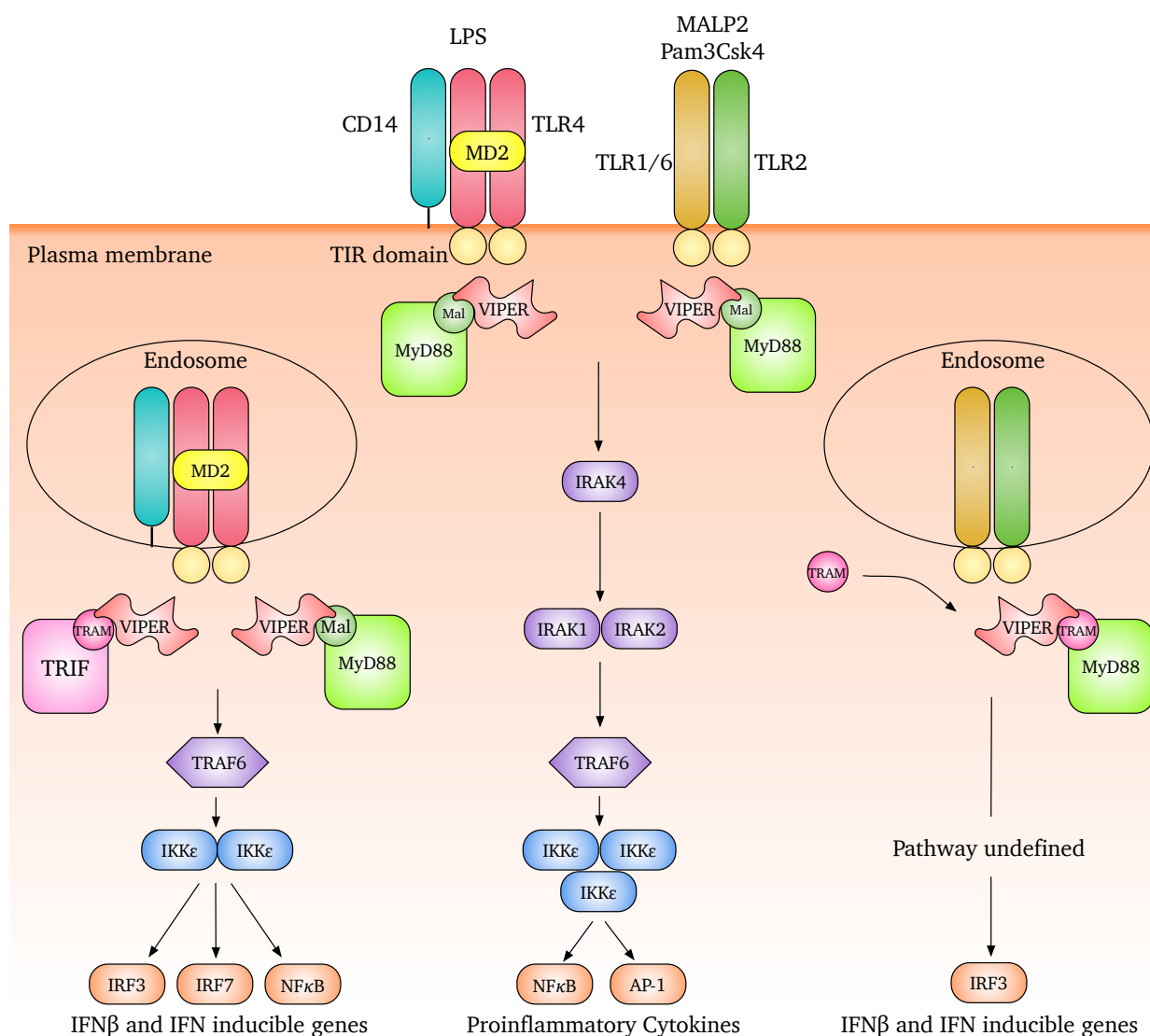


Figure 5.1: Signalling adaptors involved in TLR4 and TLR2 activation of pro-inflammatory and interferon related genes. Mal recruits MyD88 to the plasma membrane, promoting interaction with the cytosolic region of TLR4. Three discrete sites of the MyD88 TIR domain are important for TLR4 signalling, two of which mediate direct binding to the Mal TIR domain. TRAM is thought to be endocytosed along with the TLR4 complex after Mal-dependent signalling. Both MyD88-dependent and TRIF-dependent pathways use TRAF6 to activate the TAB/TAK complex for production of proinflammatory cytokines or the TBK1/IKK complex for production of IFN β and IFN inducible genes. TLR2 TIR domain is bound by Mal at the plasma membrane and thereby recruits MyD88 for induction of pro-inflammatory genes.

role in a number of conditions such as sepsis,^[153] acute pancreatitis,^[154,155] acute liver failure,^[156]

¹⁵³ A. Savva and T. Roger. *Front. Immunol.*, **4**: 387, 2013.

¹⁵⁴ J. Xue and A. Habtezion. *J. Clin. Invest.*, **124**: 437–447, 2014.

¹⁵⁵ X. Zhang et al. *Pancreas*, **39**: 819–824, 2010.

¹⁵⁶ N. Shah et al. *Liver Transplant.*, **19**: 751–761, 2013.

ischemia/reperfusion injury,^[157–159] rheumatoid arthritis,^[160] endoplasmic reticulum (ER) and stress.^[161] Therefore, targeting TLR4 directly would be more efficient and hence reduce the offshoot effects of pervasive immune suppression. This has been successfully shown in number of research studies such as, the 9R-VIPER efficacy in a TLR4-dependent flu induced lung damage model,^[162] the neuronal TLR4 signalling induces brain endothelial activation and neutrophil transmigration study in vitro,^[163] and the paraventricular nucleus attenuates blood pressure and inflammatory response in a generic model of hypertension.^[164] A schematic model of signalling adaptors involved in TLR4 and TLR2 activation and TLR inhibition of VIPER is shown in Figure 5.1.

A 9R homo-polymer delivery sequence at the N-terminus provided delivery of the peptide into cells, and in this thesis, structure of 9R-VIPER and a mutant peptide 9R-VIPER L6AE10A that exhibited loss of function was compared through Nuclear Magnetic Resonance (NMR) spectroscopy.

The amino acid sequence of 9R-VIPER and 9R-VIPER L6AE10A is shown:

9R-VIPER	9R-	K	Y	S	F	K	L	I	L	A	E	Y
9R-VIPER L6EE10A	9R-	K	Y	S	F	K	A	I	L	A	A	Y
without 9R-		1	2	3	4	5	6	7	8	9	10	11
with 9R-	1~9	10	11	12	13	14	15	16	17	18	19	20

5.1.4 Summary

In this study, the wild type and mutated residue peptides were compared to predict the key amino acid for protein binding, and hence determine the 3D structure of 9R-VIPER and a mutant form that exhibited loss of function termed 9R-VIPER L6AE10A with solution state NMR spectroscopy. The NMR structure of these peptides will be applied in moving to design peptidomimetics based on VIPER, and this project will contribute to the development of a novel virally-derived TLR4 inhibitor

¹⁵⁷ T. Dziudzio, M. Biebl, and J. Pratschke. *Curr. Opin. Organ Transplant.*, **19**: 108–114, 2014.

¹⁵⁸ J. W. Lee et al. *Biochem. Biophys. Res. Commun.*, **444**: 121–127, 2014.

¹⁵⁹ H. Meimei et al. *Crit. Care Med.*, **42**: e123–e131, 2014.

¹⁶⁰ S. Savic et al. *J. Autoimmun.*, **50**: 59–66, 2014.

¹⁶¹ G. S. Masson et al. *PLOS ONE*, **10**: e0122850, 2015.

¹⁶² K. A. Shirey et al. *Mucosal Immunol.*, **9**: 1173–1182, 2016.

¹⁶³ S. Leow-Dyke et al. *J. Neuroinflammation*, **9**: 698, 2012.

¹⁶⁴ R. B. Dange et al. *J. Neuroinflammation*, **12**: 31, 2015.

which may ultimately have applications in TLR4-dependent human disease.

5.2 Results & Discussions

5.2.1 Solution NMR Spectroscopy

Figure 5.2 is the overlapped ^1H - ^{15}N HSQC of 9R-VIPER (Blue) and 9R-VIPER L6AE10A (Red) peptides and the same residues of two peptides are connected with line to show difference in peak location. The substituted two residues (L15AE19A) are highlighted with yellow in the spectra which indicate that HN-peaks of 9R-VIPER L6AE10A were significantly shifted from those of 9R-VIPER. The carbon-proton interaction assignments were obtained with ^1H - ^{13}C HSQC, and the assigned nitrogen and carbon atom chemical shifts are summarised in Table 5.1.

By using ^1H - ^{15}N correlation and ^1H - ^{13}C 2D correlation as foundation, we were able to observe through bond correlations via spin-spin coupling along the peptides. As expected, the residue substitutions of L15 and E19 to alanine in the mutant peptide 9R-VIPER L6AE10A made assignment of side chains easier due to better chemical shift dispersion in this region. The full chemical shift assignment of 9R-VIPER and 9R-VIPER L6AE10A was achieved and the complete chemical shift values are given in Appendix A.5 and Appendix A.6.

Individual protons in the peptide act as dipoles, and two dipoles that are close in space have an effect on the neighbouring proton. The closer these protons are, the stronger the dipolar coupling nuclear over-Hauser effect (nOe's) and this appears as a stronger peak on NOESY spectra. Figure 5.3 shows nOe cross-peaks observed in the NOESY experiment of 9R-VIPER and Figure 5.4 shows 9R-VIPER L6AE10A peptides. Figure 5.3 and Figure 5.4 are nOe cross peak between the NH-H α region of the backbone area, and lines crossing nOe-peaks showcase the backbone walking along intra-residues and inter-residues correlation. A total of 276 nOe cross-peaks were assigned in 9R-VIPER and 158 nOe cross-peaks were assigned in 9R-VIPER L6AE20A.

As shown in Table 5.2, most of the nOes were intra-residue (mostly cross-peaks and sequential) for both cases. Residues of 9R-VIPER nOe connections were all weak connections and did not show any strong long-range nOe cross-peaks. On the other hand, 9R-VIPER L6AE10A has less nOe

Table 5.1: ^1H - ^{13}C and ^1H - ^{15}N chemical shift of 9R-VIPER and 9R-VIPER L6AE10A

9R-VIPER					9R-VIPER L6AE10A				
Residue	N	C α	C β	other C	Residue	N	C α	C β	other C
1ARG	-	-	-	-	1ARG	-	-	-	-
2ARG	124	-	30.98	γ 26.21 δ 43.23	2ARG	124.1	-	30.95	γ 26.21 δ 43.23
3-7ARG	-	-	-	-	3-7ARG	-	-	-	-
8ARG	123.31	-	31.16	γ 27.15 δ 43.42	8ARG	123.57	56	31.07	γ 27.15 δ 43.42
9ARG	122.78	55.98	31.38	γ 26.99 δ 43.31	9ARG	123.11	56	31.16	γ 27.06 δ 43.39
10LYS	122.57	56.54	33.32	γ 24.77 δ 29.23 ϵ 42.04	10LYS	122.91	56.48	33.23	γ 24.76 δ 29.10 ϵ 42.05
11TYR	121.11	-	39.17	δ 133.29 ϵ 118.16	11TYR	121.56	-	39.17	δ 133.33 ϵ 118.01
12SER	116.98	57.91	63.88	-	12SER	117.42	57.8	63.91	-
13PHE	121.46	-	39.37	δ 131.88 ϵ 131.50 ζ 129.98	13PHE	112.34	-	39.57	δ 131.89 ϵ 131.46 ζ 129.95
14LYS	121	56.38	33.25	γ 24.81 δ 29.12 ϵ 42.11	14LYS	122.51	56.18	33.1	γ 24.69 δ 29.19 ϵ 42.10
15LEU	122.8	55.18	42.36	δ_1 24.71 δ_2 24.80	15ALA	125.27	52.45	19.19	-
16ILE	122.55	60.86	38.77	γ_1 17.52 γ_2 27.1 δ 12.76	16ILE	120.22	61.16	38.63	γ_1 17.48 γ_2 27.26 δ 12.82
17LEU	126.13	54.89	42.51	δ_1 23.57 δ_2 23.5	17LEU	126.04	54.9	42.34	δ_1 24.86 δ_2 23.48
18ALA	124.61	52.18	19.8	-	18ALA	124.86	52.16	-	-
19GLU	119.89	-	30.68	γ 35.88	19ALA	123.44	52.31	19.45	-
20TYR	125.1	59.13	39.56	δ 133.22 ϵ 117.95	20TYR	123.24	58.82	39.34	δ 133.23 ϵ 118.14

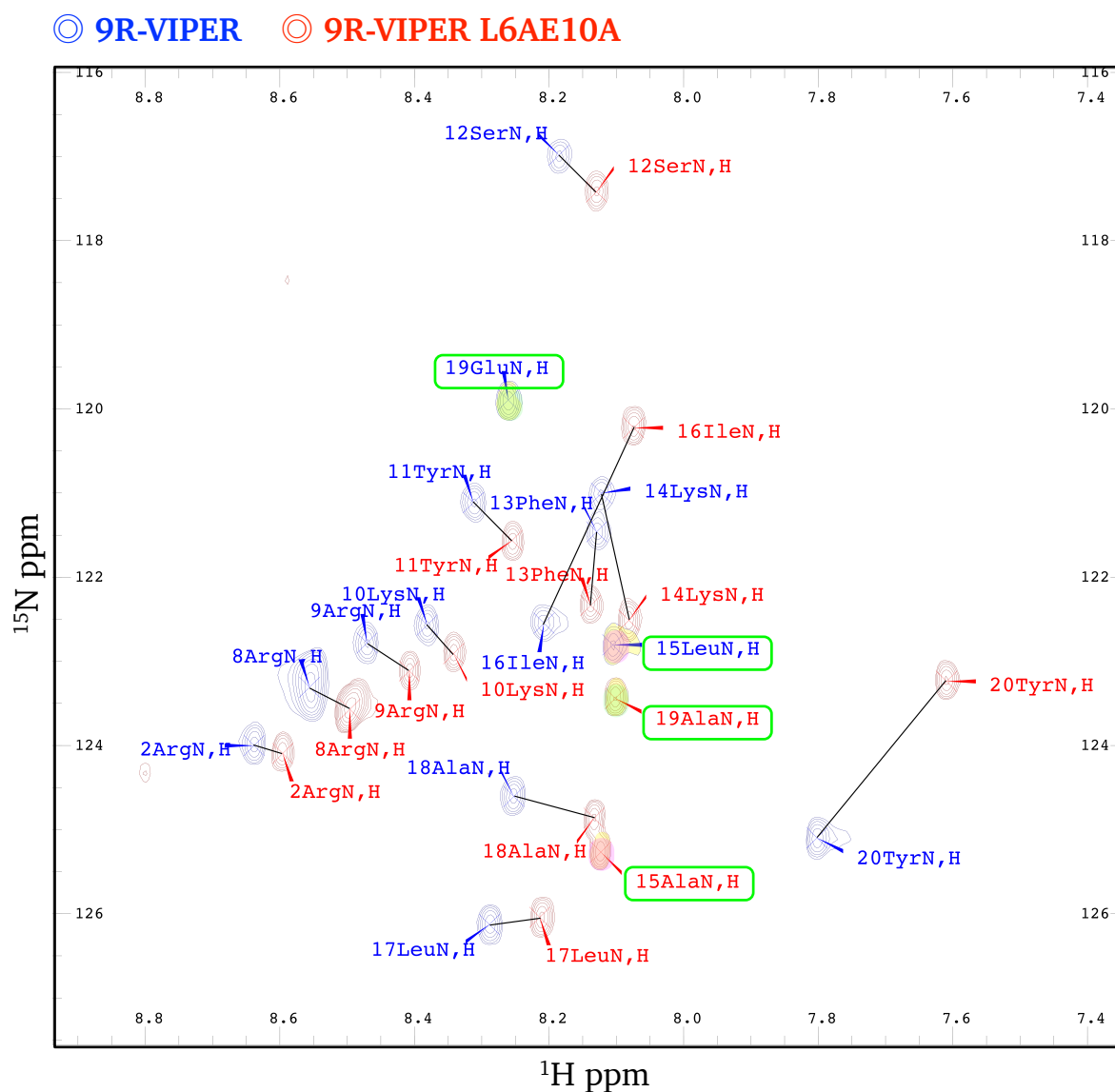


Figure 5.2: Superposition of ^1H - ^{15}N HSQC spectra of 9R-VIPER (Blue) and 9R-VIPER L6AE10A (Red) peptides. Green boxed residues are the amino acid - substituted residues.

connections when compared to 9R-VIPER. Also, few long range nOes were noticed in 9R-VIPER, however, 9R-VIPER L6AE10A did not have any strong long-range nOe cross-peaks, but had more medium strength connections (Table 5.2).^[165]

Interestingly, despite the E19A substitution, the nOe distribution near residue number 19 was comparable. In terms of the NH — H β connectivities, the grouping of the interactions (all weak)

¹⁶⁵ S. Berger and B. Siegmars. *200 and More NMR Experiments: A Practical Course*. Weinheim: Wiley-VCH, 2004.

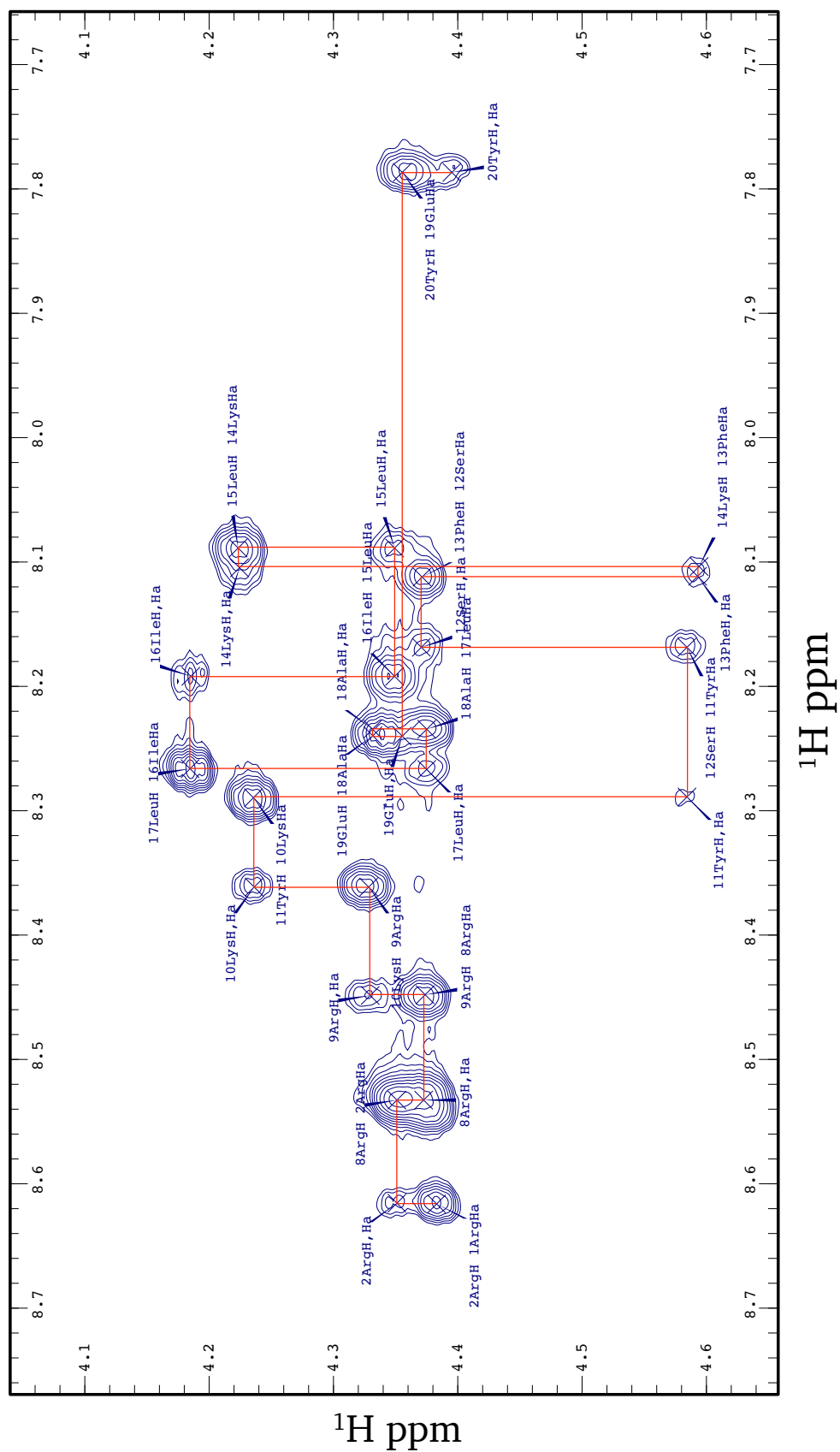


Figure 5.3: NOESY contour spectrum (NH-Hα region) of the 9R-VIPER peptide

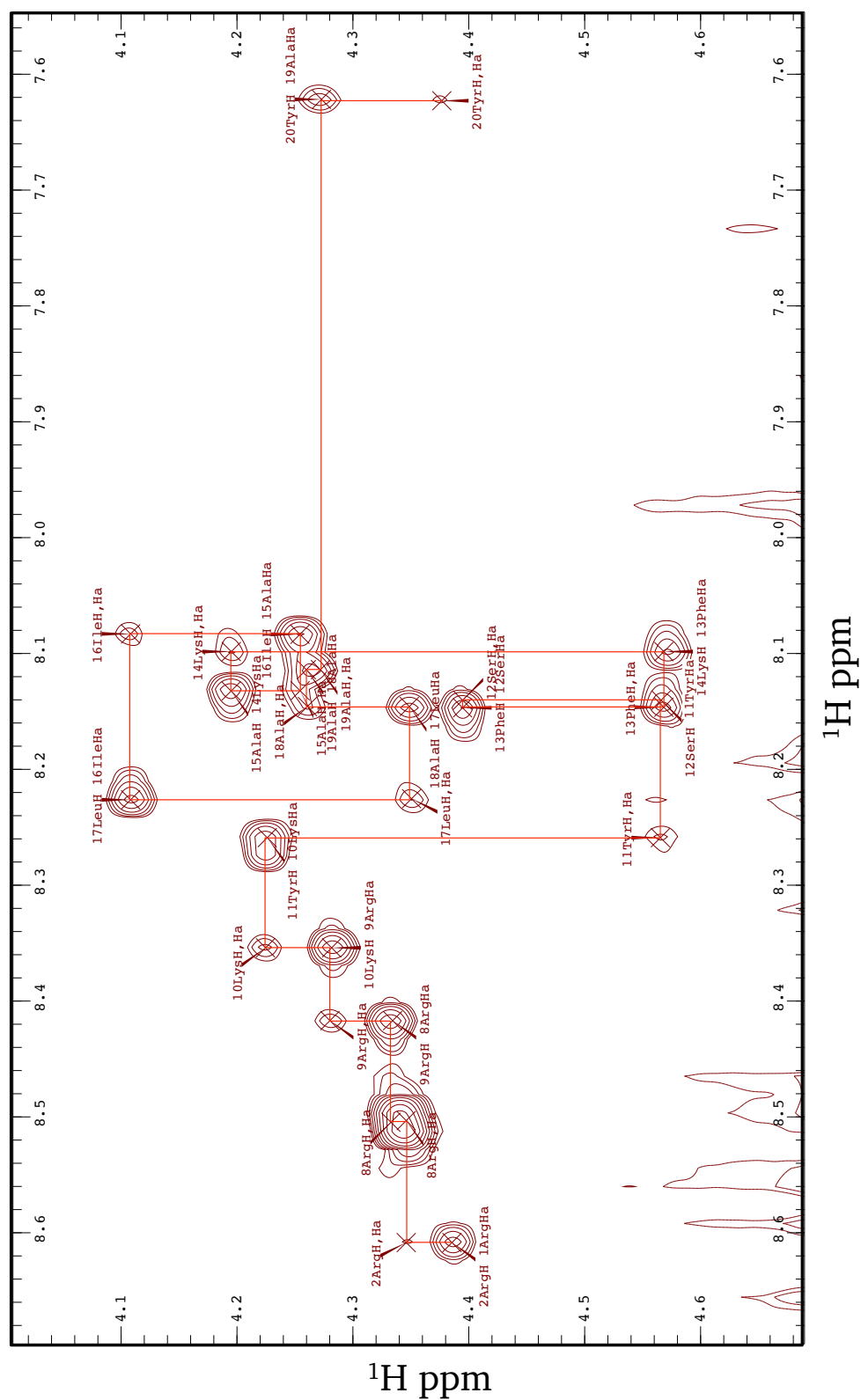


Figure 5.4: NOESY contour spectrum (NH-H α region) of the 9R-VIPER L6AE10A

Table 5.2: NMR restraints table of 9RVIPER and 9R VIPER L6AE10A

		Through-Space Connectivity					
9RVIPER	9RVIER L6AE10A	alpha(i)-amide(i+1)		amide(i)-amide(i+1)		beta(i)-amide(i+1)	
		9RVIPER	9RVIPER L6AE10A	9RVIPER	9RVIPER L6AE10A	9RVIPER	9RVIPER L6AE10A
1 Arg	1 Arg	Weak	Weak				
2 Arg	2 Arg						
3 Arg	3 Arg						
4 Arg	4 Arg						
5 Arg	5 Arg						
6 Arg	6 Arg						
7 Arg	7 Arg						
8 Arg	8 Arg	Weak	Medium	Weak	Weak	Weak	Weak
9 Arg	9 Arg	Weak	Medium	Weak	Weak	Weak	Weak
10 Lys	10 Lys	Weak	Weak	Weak	Weak	Weak	Weak
11 Tyr	11 Tyr	Weak		Weak		Weak	
12 Ser	12 Ser	Weak	Medium			Weak	Weak
13 Phe	13 Phe	Weak	Weak				Weak
14 Lys	14 Lys	Weak	Weak				Weak
15 Leu	15 Ala	Weak	Medium	Weak		Weak	Weak
16 Ile	16 Ile	Medium	Medium	Weak	Weak	Weak	Weak
17 Leu	17 Leu	Weak	Weak			Weak	
18 Ala	18 Ala	Weak	Weak				
19 Glu	19 Ala	Weak	Weak	Weak	Weak	Weak	Weak
20 Tyr	20 Tyr						

would be seen to differ slightly. Whereas there were weak interactions in the 8R–12S and 15L–17L stretches in the wild type, connectivities for the L6AE10A mutant were 8R–10K and 12S–16I stretches.

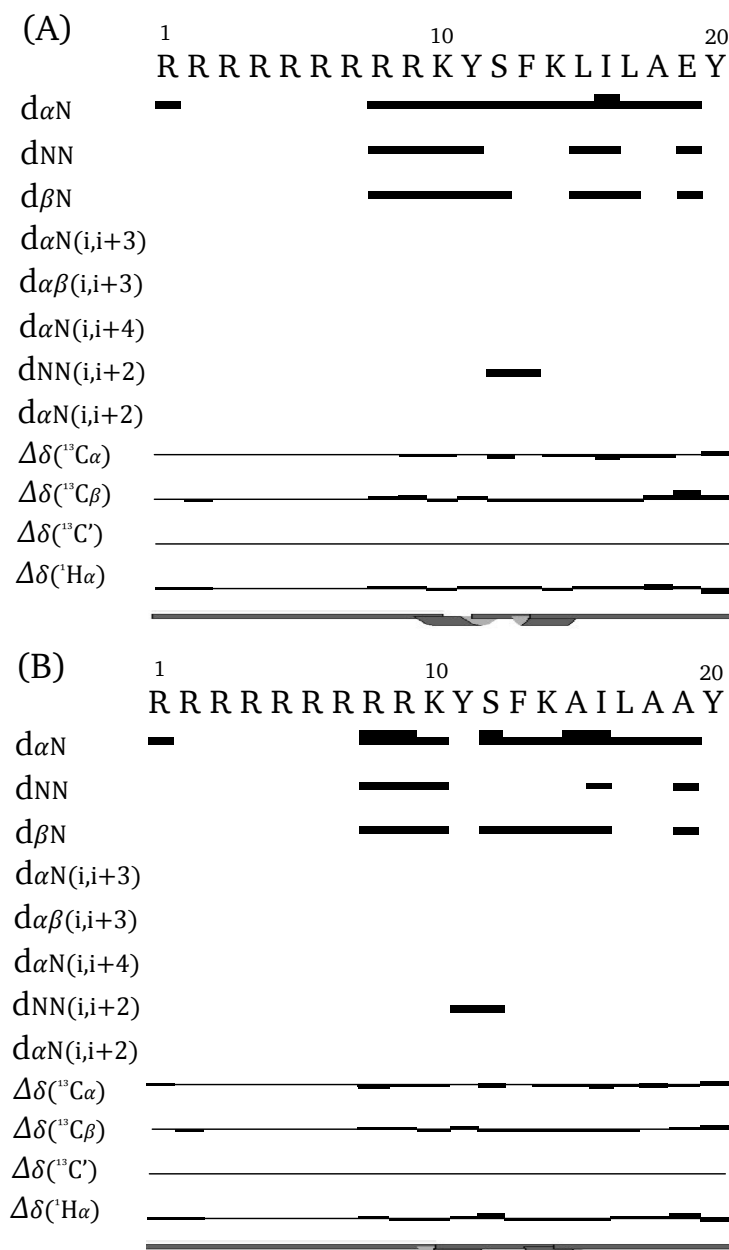


Figure 5.5: Summary of the sequential and NOE connectivities, and secondary chemical shifts of 9R-VIPER (A) and 9R-VIPER L6AE10A (B). $d\alpha N$ is intensity classification for the observed $H\alpha(i)$ to $HN(i+1)$ NOE, and respectively dNN is $HN(i)$ to $HN(i+1)$ NOE, $d\beta N$ is $H\beta(i)$ to $HN(i+1)$ NOE, $d\alpha N(i,i+3)$ is $H\alpha(i)$ to $HN(i+3)$ NOE, $d\alpha\beta(i,i+3)$ is $H\alpha(i)$ to $H\beta(i+3)$ NOE, $d\alpha N(i,i+4)$ is $H\alpha(i)$ to $HN(i+4)$ NOE, $dNN(i,i+2)$ is $HN(i)$ to $HN(i+2)$ NOE, and $d\alpha N(i,i+2)$ is $H\alpha(i)$ to $HN(i+2)$ NOE. The thickness of the bars is proportional to the NOE intensity classified as either strong, medium, or weak. $\Delta\delta(^{13}C\alpha)$ is secondary chemical shifts difference to random coil of $C\alpha$, and $\Delta\delta(^{13}C\beta)$ is of $C\beta$, $\Delta\delta(^{13}C')$ is of CO , and $\Delta\delta(^1H\alpha)$ is of $H\alpha$. The thickness of the bars indicates differences in chemical shift to random coil value and the thicker the bar the larger difference in chemical shift. Δ 9R-VIPER shows a partial α -helical conformation at Lys¹⁰, Tyr¹¹, Ser¹², and Lys¹⁴. 9R-VIPER L6AE10A shows a partial α -helical conformation at Tyr¹¹, Ser¹², Lys¹⁴, and Ala¹⁵.

5.2.2 Solution NMR based Peptide Structure Determination

Because of the success in showing loss of TLR4 signal inhibition when peptide residues were mutated, the mutant peptide 9R-VIPER L6AE10A structure was also analysed with NMR to compare the structure of 9R-VIPER. Surprisingly, Figure 5.5 is based on chemical shift assignment and nOe distance restraints, showing little difference in order to secondary structures. Secondary structure was predicted via CCPN/DANGLE.^[166]ⁱⁱⁱ 9R-VIPER shows a partial α -helical conformation at Lys¹⁰, Tyr¹¹, Ser¹², and Lys¹⁴, whereas 9R-VIPER L6AE10A shows a partial α -helical conformation at Tyr¹¹, Ser¹², Lys¹⁴, and Ala¹⁵ (Table 5.3). However, the level of certainty of this prediction is low between helical conformation and random coil for those residues predicted to be helical conformation. This is to be expected for short peptides (<30 aa), which is not likely observed with the secondary structure. Although, this minor difference can possibly affect the overall structure of 9R-VIPER and 9R-VIPER L6AE10A and was found to be due to the lack of Glu¹⁹ and Leu¹⁵. Glu¹⁹ and Leu¹⁵ of 9R-VIPER had more nOe connections with other residues than the replaced Ala¹⁵ and Ala¹⁹ of mutant 9R-VIPER. This has an effect on minor changes in secondary structure, specifically the location of α -helixes, as well as the side chains of Lys and Tyr made both peptides overall in loop (Ω) like conformation.

Figure 5.6 shows the best scored 3D structure of 9R-VIPER and 9R-VIPER L6AE10A with solvent accessible surface based on NMR data obtained as shown in the assignment table with reference against expected chemical under the same circumstances as the experimentation.^[3] Models displayed in Figure 5.6 contain only 8-20th residues without first seven arginine (7R) which we were unable to assign due to chemical shift overlaps. In these models, detailed insight into the backbone and side-chain configurations were found to explain the difference of partial α -helix

¹⁶⁶ M. S. Cheung et al. *J. Magn. Reson.*, **202**: 223–233, 2009.

³ M. P. Williamson. Peptide Structure Determination by NMR in: *Spectroscopic Methods and Analyses*. ed. by J. Jones, B. Mulloy, and A. Thomas. **17**: 69–85, New Jersey: Humana Press, 1993. DOI: [10.1385/0-89603-215-9:69](https://doi.org/10.1385/0-89603-215-9:69)

ⁱⁱⁱ Dihedral ANgles from Global Likelihood Estimates (DANGLE) estimates the backbone dihedral angles ϕ and ψ of each residue in peptide/protein by using the chemical shift data. Dihedral angle ϕ (phi) is the angle of rotation between N and C $_{\alpha}$ atom, and ψ (psi) is the angle of rotation between the C $_{\alpha}$ and C $_N$ atom. Two-dimensional plot of ϕ (x-axis) and ψ (y-axis) is known as a Ramachandran diagram and each of secondary structural conformation has unique regions that appear in the diagram.^[11] Based on the estimated dihedral angle values (ϕ , ψ), combined with the nOe values, the conformation preferences of each amino acid type can be predicted.^[166]

Table 5.3: Secondary structure prediction of 9R-VIPER and 9R-VIPER L6AE10A generated via CCPN/DANGLE. (Alpha: α -helical conformation, Coil: Secondary structure which is not both α -helix nor β -strand, Undefined: Unable to define the secondary structure).

9R-VIPER		9R-VIPER L6AE10A	
Residue	Secondary Structure	Residue	Secondary Structure
1 ARG	Undefined	1 ARG	Undefined
2 ARG	Undefined	2 ARG	Undefined
3 ARG	Coil	3 ARG	Undefined
4 ARG	Undefined	4 ARG	Undefined
5 ARG	Undefined	5 ARG	Undefined
6 ARG	Undefined	6 ARG	Undefined
7 ARG	Coil	7 ARG	Coil
8 ARG	Coil	8 ARG	Coil
9 ARG	Coil	9 ARG	Coil
10 LYS	Alpha	10 LYS	Coil
11 TYR	Alpha	11 TYR	Alpha
12 SER	Alpha	12 SER	Alpha
13 PHE	Coil	13 PHE	Coil
14 LYS	Alpha	14 LYS	Alpha
15 LEU	Coil	15 ALA	Alpha
16 ILE	Coil	16 ILE	Coil
17 LEU	Coil	17 LEU	Coil
18 ALA	Coil	18 ALA	Coil
19 GLU	Undefined	19 ALA	Undefined
20 TYR	Undefined	20 TYR	Undefined

conformation location in 9R-VIPER and 9R-VIPER L6AE10A. This is due to the loss of the long side chains of Glu¹⁹ and Leu¹⁵ whose α , β , γ , and δ protons in particular were used stabilise the more complex conformation of 9R-VIPER. Chemical-shift based secondary structure prediction (Alpha in red in Table 5.3) show that the position of the α -helical conformation is different between the two peptides (Alpha in red in Table 5.3). These differences after residue substitution may explain the

loss of inhibitory activities in mutant peptide 9R-VIPER L6AE10A *in vitro*.^[148,167]

As a consequence, Figure 5.7 shows the ensemble of backbone chains of the NMR-based structures of 9R-VIPER (Figure 5.7:A) and 9R-VIPER L6AE10A (Figure 5.7:B) with solvent assessable surface, generated via CCPNGrid/ARIA.^[108,169–171] Figure 5.7 is the superposition of the two respective peptides after disregarding the loose, random-coil-like 9R N-terminal sequences (residues 10th to 20th) without nine arginine (9R) delivery sequence. The overall value of the RMSD of 9R-VIPER and 9R-VIPER L6AE10A were similar.

Finally, the 3D structure of the VIPER region in the vaccinia virus A46 protein (residues no. 88th–98th) is shown (Figure 5.8) (PDB ID code: 4M0S)^[168] and subsequently superimposed with the VIPER and VIPER mutant peptide structural ensembles (Figure 5.9). VIPER region in A46 has a defined α -helical structure, while both VIPER and VIPER mutant peptides are not fully folded or possess any defined secondary structure but contain some fluctuating helical-like turns which is expected as secondary structure is not typically observed for short peptides. Although the VIPER peptide appears to be in general agreement in the structural properties with the counterpart region in the A46 protein, the helical-like register is still conserved and hence, placement of the side chains can occur native-like (Figure 5.9:A). In contrast, a larger structural change is observed when compared with the VIPER L6AE10A peptide, particularly in the central part where a loop is found for the mutant peptide (Figure 5.9:B). This supports the suggestion that the 11-residue VIPER peptide exhibits A46-like activity only for the inhibition of MAL and TRAM, whereas the native protein is capable of inhibiting other TIR-domains like MyD88, TRIF, MAL and TRAM.^[172] This binding activity is completely diminished for the mutant VIPER peptide suggesting that two alanine replacements affected the location of helical-like turns. This influenced the overall conformation of the peptide which is considered important in the binding with the TIR-domains.^[148,167] Hence, we are able to show strong corroboration of the cell biology experiments with the present structural studies.^[148,149,172]

¹⁶⁷ J. Y. Kim et al. *Manuscript in Preparation*, , 2019.

¹⁰⁸ W. F. Vranken et al. *Proteins Struct. Funct. Bioinforma.*, **59**: 687–696, 2005.

¹⁶⁹ R. H. Fogh et al. *Bioinformatics*, **21**: 1678–1684, 2005.

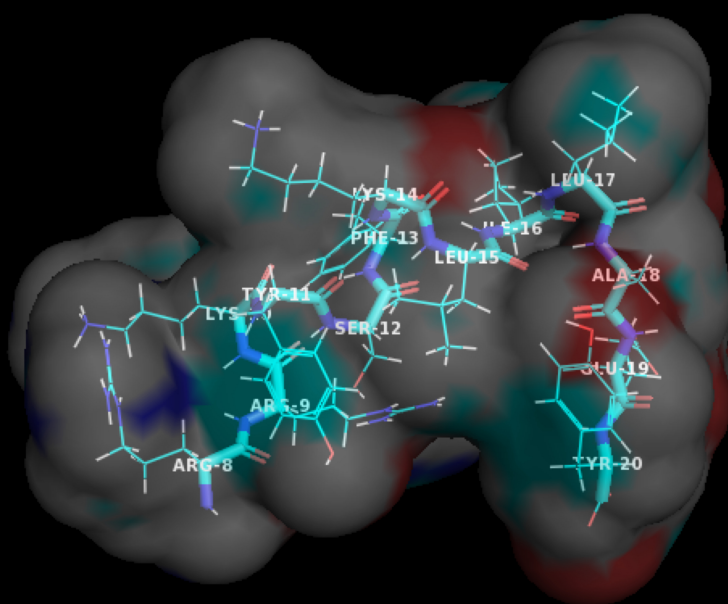
¹⁷⁰ M. Nilges et al. *Structure*, **16**: 1305–1312, 2008.

¹⁷¹ W. Rieping et al. *Bioinformatics*, **23**: 381–382, 2007.

¹⁶⁸ Y. W. Kim et al. *Protein Sci.*, **23**: 906–914, 2014.

¹⁷² S. I. Oda, E. Franklin, and A. R. Khan. *Mol. Immunol.*, **48**: 2144–2150, 2011.

(A)



(B)

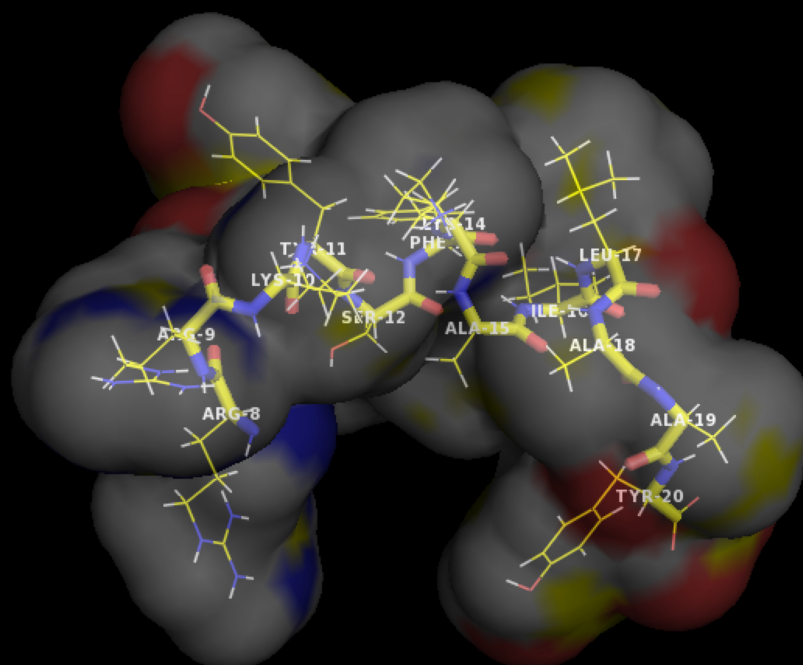


Figure 5.6: The 3-D structure of 9R-VIPER (A), 9R-VIPER L6AE10A (B). In this model, residues 8 to 20 are shown without the seven arginine (7R) at N-terminus and models show solvent accessible surface. (Models were generated from GeNMR^[110])

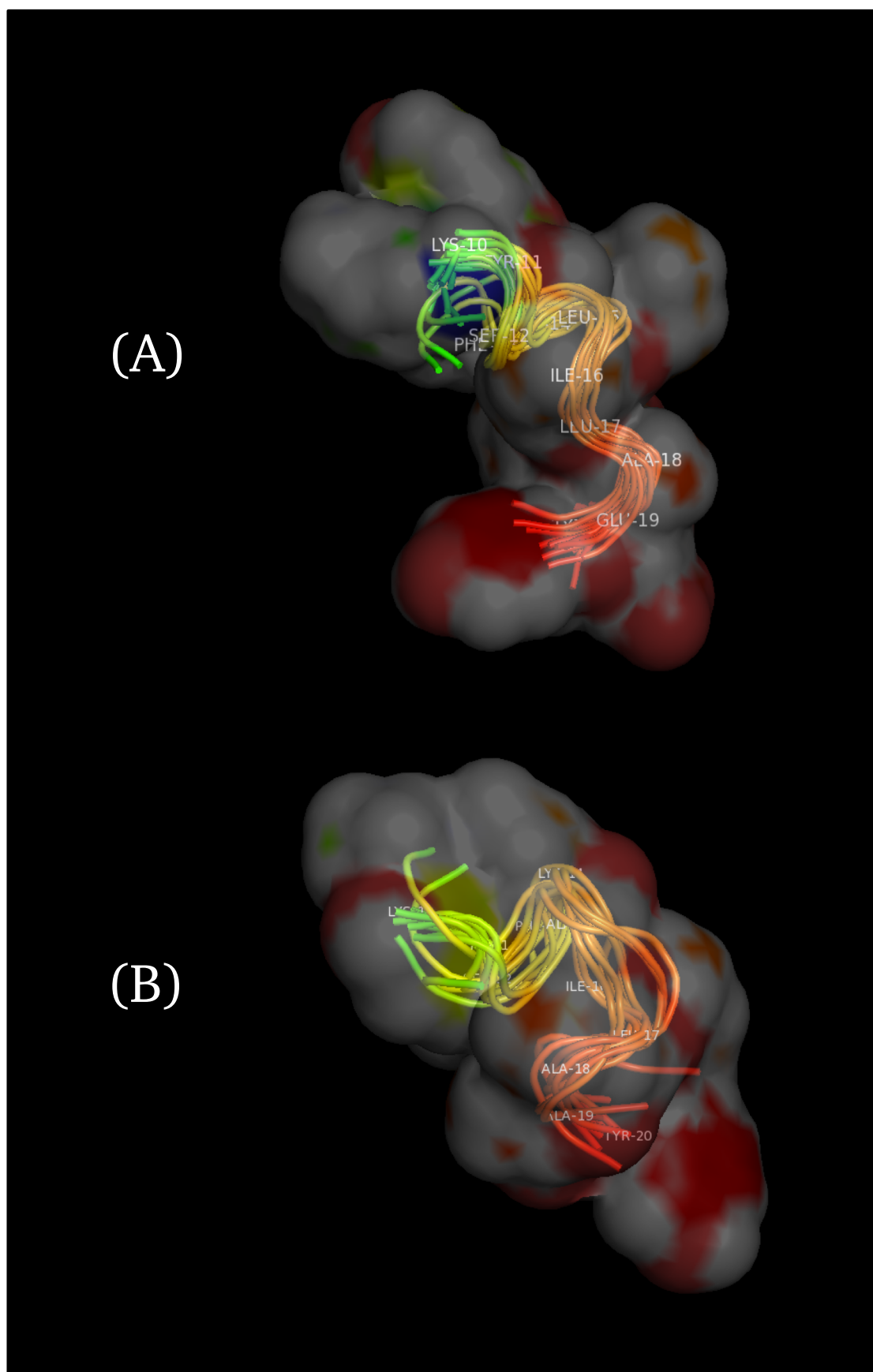


Figure 5.7: NMR-based structure ensemble of 9R-VIPER (A) and 9R-VIPER L6AE10A (B) with solvent assessable surface. Only the backbone traces are shown and the 9R segment at the N terminus is not shown. The Structures are coloured from green at the N terminus to red at the C terminus.

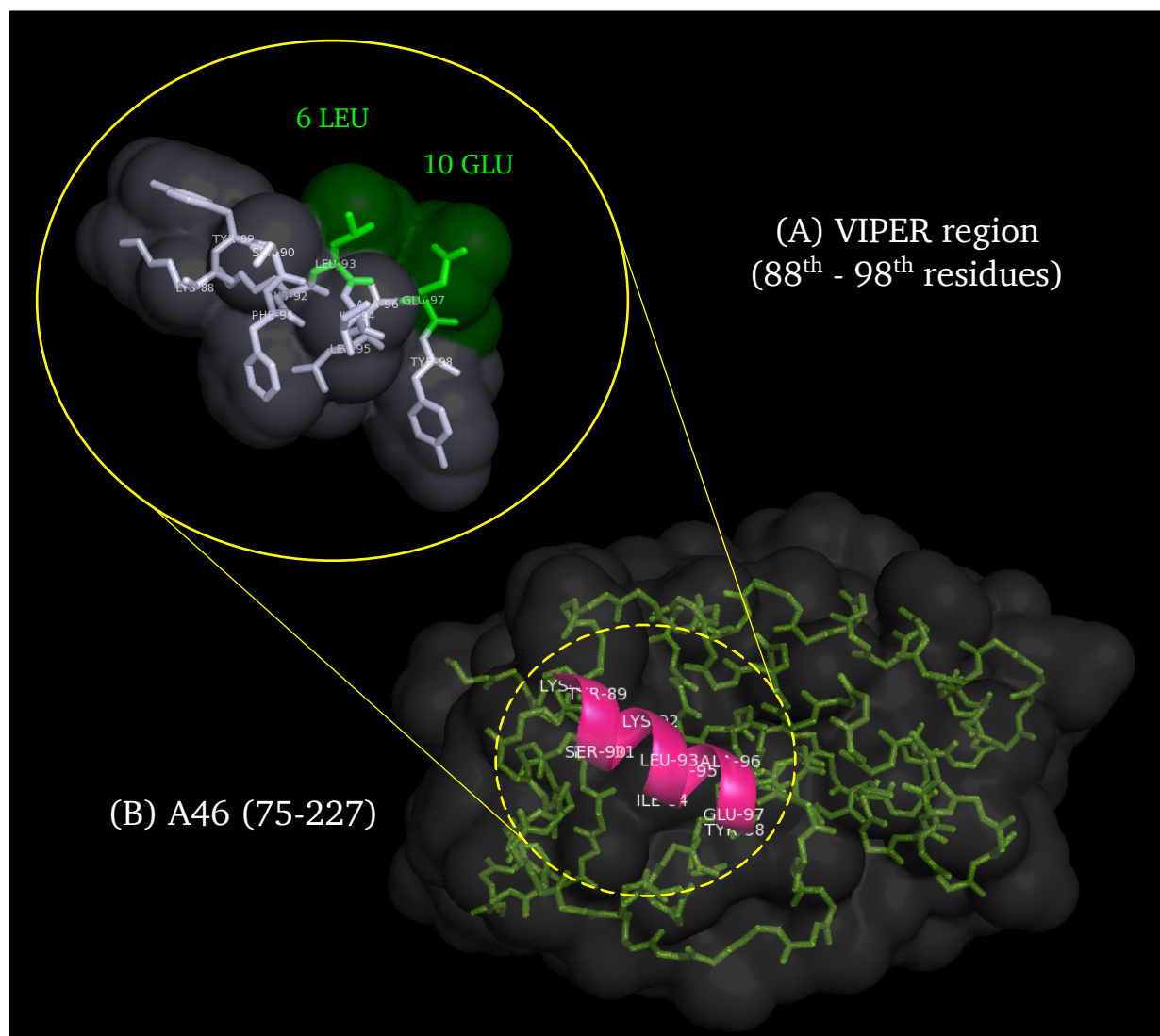


Figure 5.8: Superimposed models of the native VIPER in A46(75-227) (PDB ID code: 4M0S) (Pink) with VIPER (Green) (A) and mutant peptide VIPER L6AE10A (Yellow) (B).

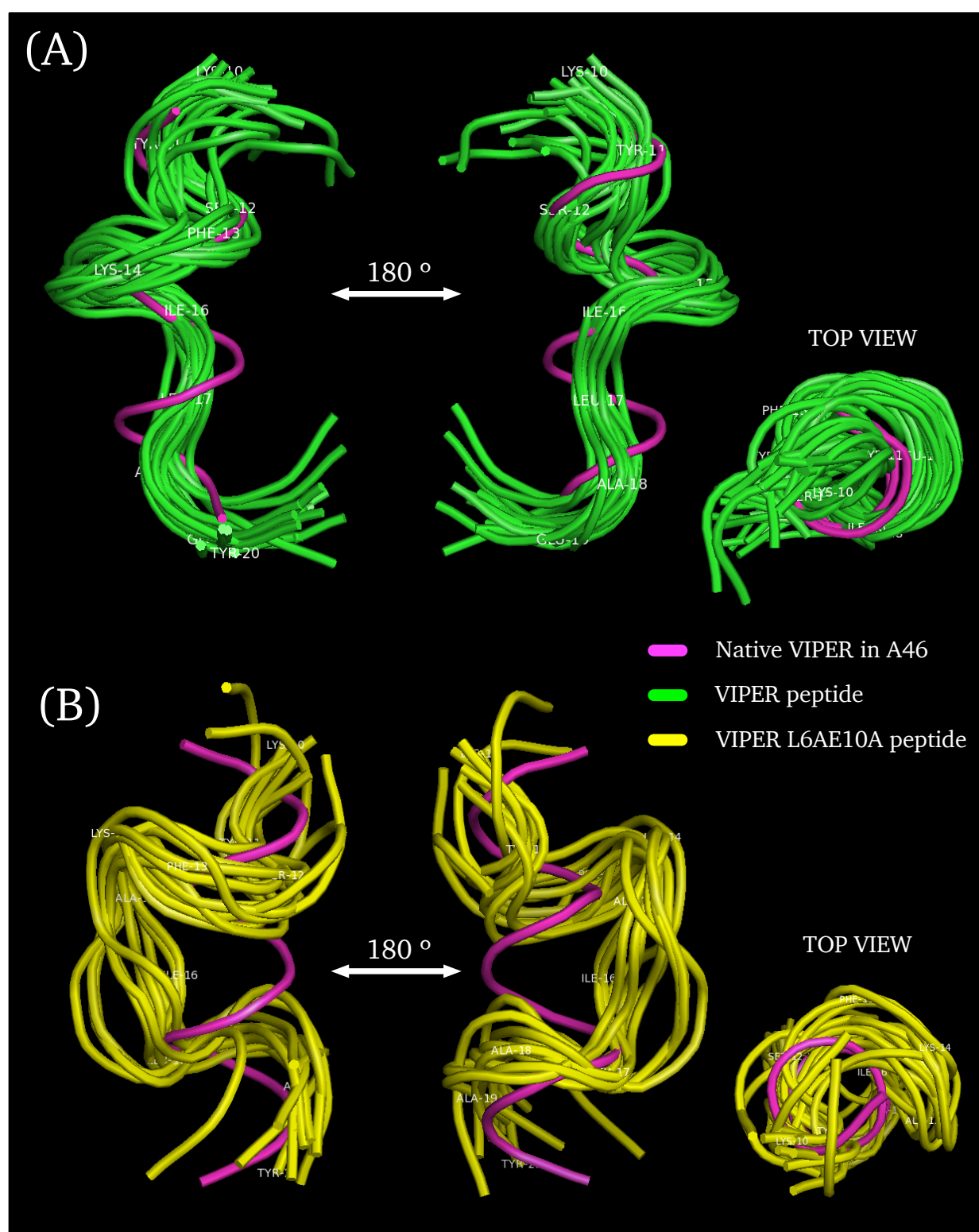


Figure 5.9: Composition of the 3D structure of VIPER region the 11aa-residue fragment in the A46 protein (A) and the A46(75-227) protein (PDB ID code: 4M0S)^[168] (B)

5.3 Conclusion

Previous studies confirm that 9R-VIPER as more effective at TLR4 and TLR2 signal inhibition than VIPER. Novel mechanisms of TLR2-driven interferon induction have been shown to be affected by 9R-VIPER and VIPER. The longevity of TLR4 inhibition by VIPER and 9R-VIPER were determined *in vitro* with 9R-VIPER showing greater overall activity with room for modification to assist biological availability time. Wild type and peptides with mutated residues predicted to be required for protein binding were compared, and residues L15 and E19 were confirmed to be of significant importance to inhibitory activity of 9R-VIPER. This finding drove us to investigate the structure of 9R-VIPER and the mutant peptide L6AE10A by NMR.^[167]

In this study, we were able to generate the 3D structure of 9R-VIPER and 9R-VIPER L6AE10A. Both peptides are not fully folded but consist of helical-like turns at different location which lead to the difference between 9R-VIPER and 9R-VIPER L6AE10A structure. Despite minor amino acid chain substitutions, which is directly reflected in the bioactivity and the loss of activity when key residues Leu¹⁵ and Glu¹⁹ are changed to Ala. This highlights that the molecular activity of the peptide requires a specific shape and a characteristic distribution of functionalities at the binding surface.

This information can be used to design peptidomimetics based on 9R-VIPER and can contribute to the development of a novel virally-derived TLR4 inhibitors which may ultimately assist in treating TLR4-dependent human diseases in the future.

Chapter 6

Closing Summary

Biomolecular NMR is a powerful technique which is used to extract various information of biomolecules such as the 3D-structure, dynamic properties, and interactions between biomolecules.

In this thesis, NMR spectroscopy is combined with other biophysical techniques and then applied to estimate the hydrodynamic radii of large biomolecules present in the same sample. The technique is also used to investigate the native-like structure of a milk protein after thermal-treatment, and to determine structural information of TLR4 inhibiting peptide terms VIPER.

In Chapter 3, protein aggregation and oligomerisation were investigated with diffusion NMR and other techniques. We were able to further develop diffusion NMR as a tool to detect protein aggregation, as well as to estimate sizes of the oligomer species. Through the Gaussian decay function method, the proportional presence of the obtained oligomer species and its hydrodynamic radius was also estimated.

In Chapter 4, the structural modification of bovine β -Lg was investigated after applying thermal treatment at 85 °C. Heat treatment was applied to a high concentration of β -Lg (12 % w/w) at neutral pH (pH 6.7), identical to the current pasteurization condition used in the dairy industry. We observed β -Lg form into aggregates after the thermal treatment, which however, still remains in its native-like structure.

In Chapter 5, with various NMR spectroscopy techniques, the 3D-structure of VIPER peptide

with TLR4 inhibiting behaviour (which is the fraction of the vaccinia virus A46) and its inactive mutant peptide VIPER L6AE10A were determined. The calculated ensemble structures of peptides were then superimposed with the VIPER region in A46 with alpha-helical structure. Both VIPER and mutant VIPER are not fully folded but have helical-like turns at different location which resulted in a difference between the overall structure of the two peptides. This suggests that two residues (Leu⁶ and Glu¹⁰) replaced in mutant peptides are important not only for the overall structure but also for the TLR4 inhibitory property.

We have applied various NMR spectroscopy methods to interrogate the biophysical property of interest. NMR is a well-established method in determining the structure of protein/peptide as applied in Chapter 3 and Chapter 4. We are beginning to see that NMR is becoming a useful supporting tool for investigating protein aggregation.

Appendices

A.1 Hydrodynamic radius of biomolecule calculation by diffusion NMR

The integrated intensity of protein and reference compound 1,4-dioxane were recorded while increasing gradient strength and were subsequently extracted. The integrated intensity of the native HEWL aliphatic region (3.0~0.0 ppm) and 1,4-dioxane (chemical shift: 3.75 ppm) as well as their normalized values are shown in the Table 1 as an example.

Table 1: Integrated spectral intensity of the aliphatic region signals of the native HEWL (3.0~0.0 ppm) and 1,4-dioxane (3.75 ppm)

Gradient strength	Native HEWL		1,4 dioxane	
	Integrated intensity	Normalised intensity	Integrated intensity	Normalised intensity
1200	60.6195	0.9925094	1.43225	1
1500	61.802	0.9897728	1.40875	0.97056
1800	60.05225	0.9813069	1.35375	0.944653
2100	59.17475	0.9629715	1.2905	0.893906
2400	59.52625	0.9607118	1.284	0.871096
2700	59.00025	0.966699	1.20525	0.842338
3000	57.61225	0.9425104	1.1275	0.774168
3300	55.58575	0.9256095	1.024	0.726316
3600	55.35225	0.9176287	0.9885	0.690579
3900	54.3305	0.9157397	0.93225	0.664232
4200	53.309	0.8928161	0.84725	0.60696
4500	54.7185	0.8921927	0.85025	0.583357
4800	51.6985	0.8681074	0.73325	0.529424
5100	51.27175	0.8582058	0.71175	0.502289
5400	50.32675	0.8438968	0.669	0.480725
5700	51.57925	0.8473215	0.6755	0.462818
6000	48.333	0.8075589	0.56875	0.412927

Table 1: Integrated spectral intensity of the aliphatic region signals of the native HEWL (3.0~0.0 ppm) and 1,4-dioxane (3.75 ppm)

Gradient strength	Native HEWL		1,4 dioxane	
	Integrated intensity	Normalised intensity	Integrated intensity	Normalised intensity
6300	48.2315	0.7930522	0.54525	0.388843
6600	47.90575	0.7826401	0.53825	0.370195
6900	45.90175	0.7569435	0.467	0.33573
7200	44.8635	0.7342043	0.436	0.30931
7500	45.5955	0.7382519	0.46	0.322355
7800	44.473	0.7132091	0.4265	0.296138
8100	43.21375	0.6940691	0.38975	0.278628
8400	42.73175	0.6813317	0.38525	0.262797
8700	41.615	0.6636626	0.35775	0.244755
9000	40.3045	0.6398914	0.3305	0.225729
9300	38.633	0.6257852	0.30425	0.22513
9600	36.95275	0.6003478	0.27275	0.210391
10100	36.523	0.5834746	0.2925	0.201524
10600	34.90925	0.5503736	0.2895	0.187058
11100	32.51075	0.520099	0.24125	0.172936
11600	31.178	0.4979624	0.246	0.168766
12100	29.29725	0.4659306	0.22225	0.147422
12600	27.695	0.4404137	0.2105	0.150934
13100	25.63575	0.4043753	0.17825	0.123435
13600	24.462	0.3773847	0.17475	0.115379
14100	23.50625	0.3558891	0.17125	0.109964
14600	22.8445	0.3349948	0.18	0.109789
15100	21.6545	0.3196781	0.17625	0.103891
15600	19.557	0.2858224	0.14675	0.084931
16100	18.01475	0.2654043	0.12475	0.078318
16600	17.4305	0.2470213	0.13675	0.078021
17100	15.8185	0.2297451	0.12475	0.078843
17600	14.09825	0.2039948	0.10475	0.068489
18100	12.554	0.1831315	0.09625	0.065219
18600	11.4925	0.1675997	0.083	0.060593
19100	11.0055	0.1522269	0.0975	0.050452
19600	10.3095	0.1430703	0.09025	0.051134
20600	7.4545	0.1014504	0.05175	0.034259
21600	7.52875	0.0855453	0.07	0.03083
22600	6.18275	0.0755189	0.06	0.034058
23600	4.205	0.0543881	0.029	0.027508
24600	2.81125	0.031348	0.00425	0.014793

Table 1: Integrated spectral intensity of the aliphatic region signals of the native HEWL (3.0~0.0 ppm) and 1,4-dioxane (3.75 ppm)

Gradient strength	Native HEWL		1,4 dioxane	
	Integrated intensity	Normalised intensity	Integrated intensity	Normalised intensity
25600	3.69025	0.0276347	0.04125	0.01408
26600	2.47625	0.0244307	0.01875	0.020063
27600	1.4285	0.0118447	0.0045	0.006607
28600	1.85975	0.0105766	0.026	0.009785
29600	1.87625	0.0089396	0.034	0.011736
30000	1.9315	0.0053076	0.0295	0.010447

The normalized values of integrated intensity of the native HEWL and 1,4-dioxane were fitted to a Gaussian decay function derived from the *Stejskal-Tanner Equation* (2.3), and Equation (2.4)) using MATLAB. From the fitting process, the relative diffusion values of diffusion species were obtained as described in Section 2.12.2.

The relative diffusion coefficient of the native HEWL (monomer) is:

$$D_{nativeHEWL} = 5.17 \times 10^{-9} (\pm 0.584 \times 10^{-9})$$

And the relative diffusion coefficient for the reference compound, 1,4-dioxane is:

$$D_{dioxane} = 4.68 \times 10^{-8} (\pm 0.108 \times 10^{-8})$$

The hydrodynamic radius of the protein ($R_h^{protein}$) is directly proportional to the hydrodynamic radius of the reference compound 1,4-dioxane as the other effecting factors on sample solvent are the same for all diffusing species. The hydrodynamic radius of 1,4-dioxane is known as 2.12 Å ($R_h^{dioxane}$) and $R_h^{protein}$ can be calculated as shown below.

$$R_h^{protein} = (D_{dioxane} / D_{protein}) \times (R_h^{dioxane})$$

$$= (4.68 \times 10^{-8} / 5.17 \times 10^{-9}) \times (2.12)$$

$$= 19.20 \text{ Å}$$

Therefore, the calculated hydrodynamic radius of native HEWL is 19.20 Å.

A.2 Estimates of hydrodynamic radius of HEWL- BS³ oligomers obtained from SE-HPLC

The size of the synthesised oligomers of HEWL-BS³ cross linked mixture was estimated through SE-HPLC and the estimated monomer, dimer, and trimer HEWL hydrodynamic radii were respectively 20, 23, and 28 Å. (Data adopted from A.7). From the estimated hydrodynamic radii, the diffusion coefficient of each HEWL were calculated with the obtained 1,4-dioxane diffusion coefficient from diffusion NMR data fitting process by using Equation (2.3). The obtained relative diffusion coefficient of 1,4-dioxane ($R_h^{dioxane} = 2.12$ Å) in HEWL-BS³ mixture sample is given below:

$$D_{dioxane} = 4.32 \times 10^{-8} (\pm 0.158 \times 10^{-8})$$

$$D_{protein} = (D_{dioxane} \times R_h^{dioxane}) / (R_h^{protein})$$

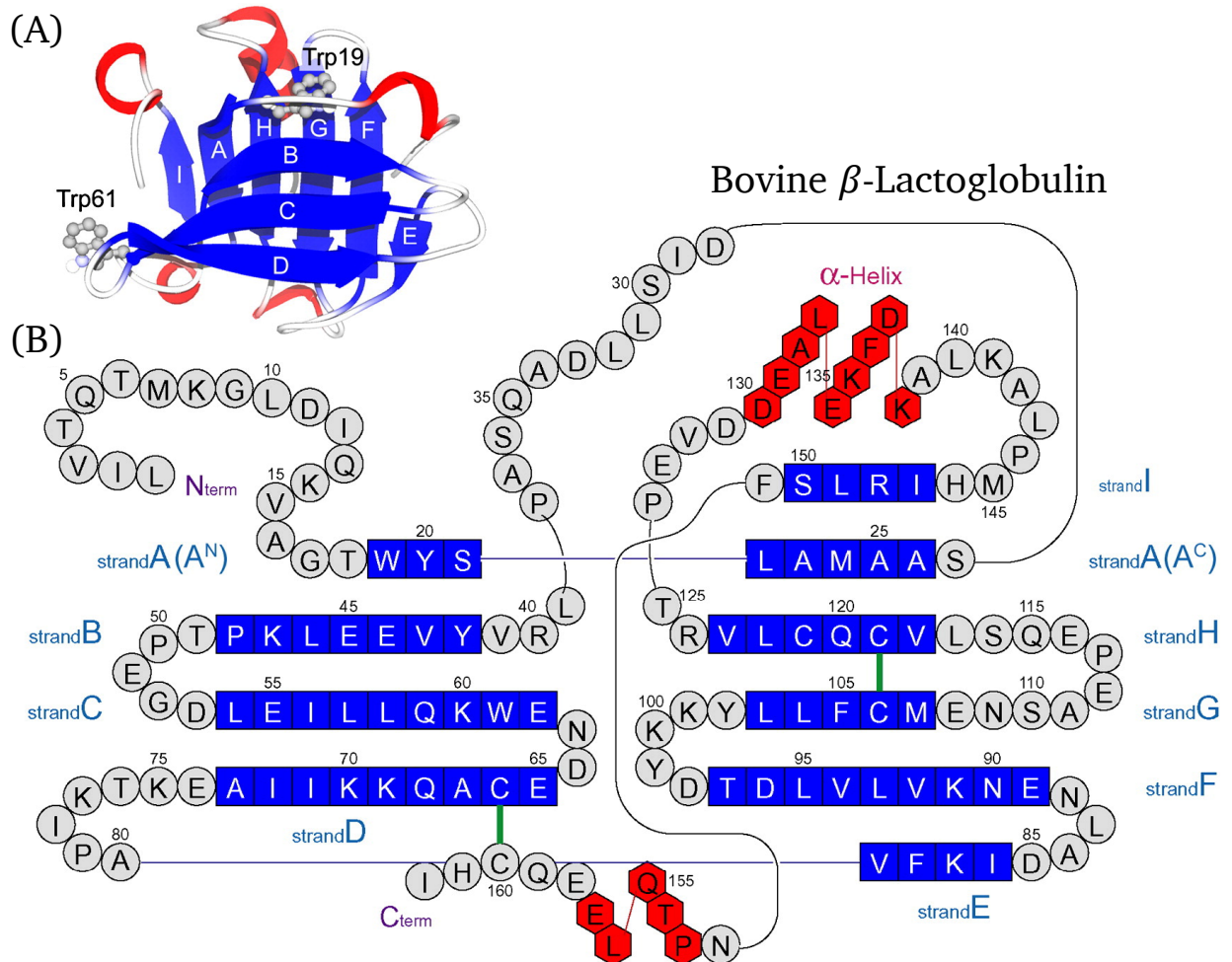
$$D_{monomer} = (4.32 \times 10^{-8} / 20) \times (2.12) = 4.58 \times 10^{-9}$$

Therefore, the relative diffusion coefficient of the HEWL monomer is 4.58×10^{-9} . Diffusion coefficient values of the dimer and trimer were also calculated through the same method and are shown in the following table.

Table 2: Estimated hydrodynamic radius and diffusion coefficients of HEWL-BS³ oligomers

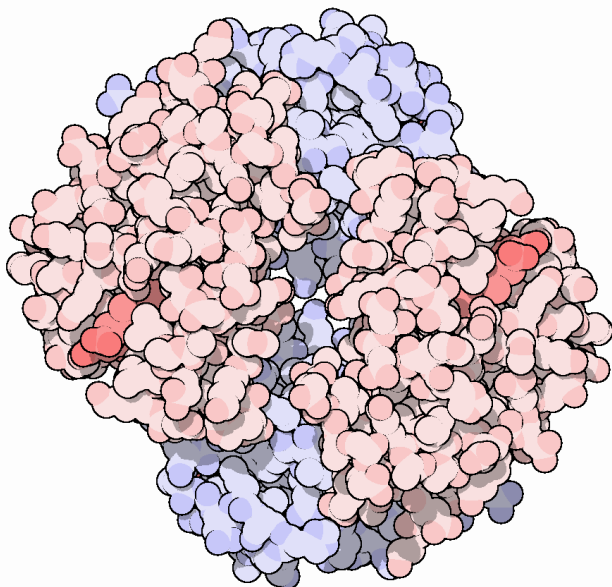
	R_t (min)	R_h (Å)	Calculated $D_{protein}$
Trimer	36.5	28	3.27×10^{-9}
Dimer	39	23	3.98×10^{-9}
Monomer	41	20	4.58×10^{-9}

A.3 3D structure and amino acid sequence of bovine β -Lg

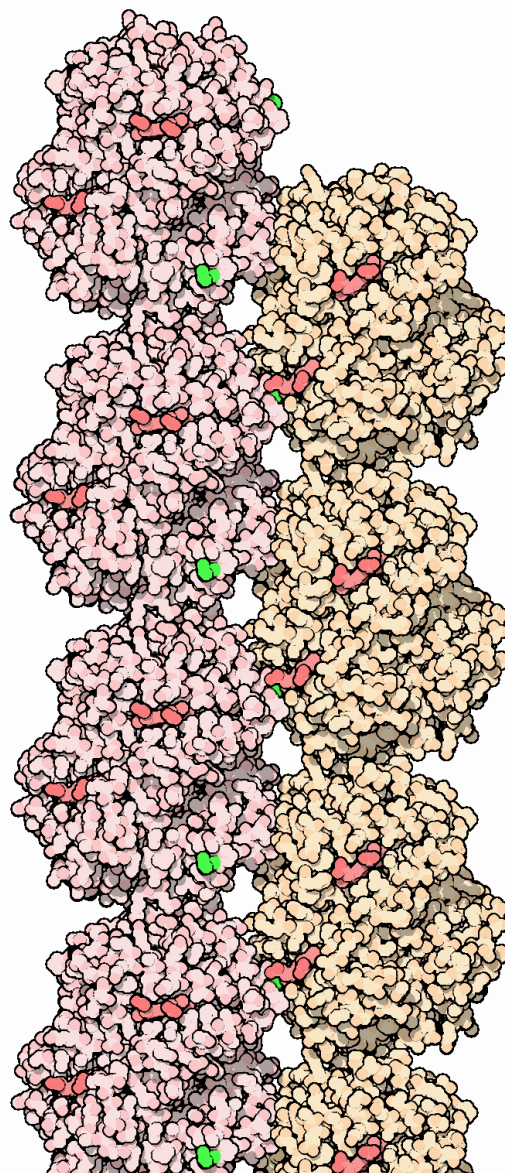


The 3D structure and amino acid sequence of bovine β -Lactoglobulin (A) Ribbon diagram of a single subunit of bovine β -Lg protein (PDB code: 1BEB). The β -strands are labelled (A–I). (B) A schematic representation of the amino acid residues of the β -Lg sequence. Residues making up α -helix, β -sheet, and loop are represented by hexagons in red, squares in blue, and circles in grey, respectively. Green lines indicate the position of disulfide bonds. β -Lg has two β -sheets; The B–D strands and N-terminal half of the A strand consist of one and the E–H strands and C-terminal half of the A strand consist of the other. (Adopted from K. Sakurai et. al (Sakurai et al. 2009))

A.4 Fiber of sickle cell hemoglobin



Hemoglobin (Above)



Fibre of sickle hemoglobin (Right)

Adapted from - WEBPAGE (<https://pdb101.rcsb.org/motm/41>)

A.5 Full assigned chemical shift list of 9R-VIPER

Residue sequence code	Residue label	Atom name	Atom type	Chemical shift value (ppm)
1	ARG	H	Ha	4.383
2	ARG	H	H	8.619
2	ARG	H	Ha	4.349
2	ARG	H	Hda	3.184
2	ARG	H	Hdb	3.189
2	ARG	H	Hga	1.625
2	ARG	H	Hgb	1.65
2	ARG	H	Hba	1.751
2	ARG	H	Hbb	1.817
2	ARG	C	Cb	30.983
2	ARG	C	Cg	26.207
2	ARG	C	Cd	43.231
2	ARG	H	He	7.215
2	ARG	N	Ne	134.624
2	ARG	N	N	124.002
8	ARG	H	Ha	4.366
8	ARG	H	H	8.535
8	ARG	H	Hda	3.154
8	ARG	H	Hdb	3.163
8	ARG	H	Hgb	1.628
8	ARG	H	Hga	1.596
8	ARG	H	Hba	1.732
8	ARG	H	Hbb	1.802
8	ARG	C	Cb	31.163
8	ARG	C	Cg	27.149
8	ARG	C	Cd	43.423
8	ARG	H	He	7.211
8	ARG	N	Ne	134.647
8	ARG	N	N	123.308
9	ARG	C	Ca	55.977
9	ARG	H	H	8.452
9	ARG	H	Ha	4.328
9	ARG	H	Hdb	3.132
9	ARG	H	Hda	3.124
9	ARG	H	Hbb	1.725
9	ARG	H	Hba	1.693
9	ARG	H	Hgb	1.56
9	ARG	H	Hga	1.535
9	ARG	C	Cb	31.376
9	ARG	C	Cg	26.985
9	ARG	C	Cd	43.312
9	ARG	H	He	7.163
9	ARG	N	Ne	134.603
9	ARG	N	N	122.782
10	LYS	H	Ha	4.236

Residue sequence code	Residue label	Atom name	Atom type	Chemical shift value (ppm)
10	LYS	C	Ca	56.544
10	LYS	H	H	8.364
10	LYS	H	Hea	2.91
10	LYS	H	Heb	2.924
10	LYS	C	Ce	42.044
10	LYS	H	Hbb	1.645
10	LYS	H	Hgb	1.296
10	LYS	H	Hga	1.182
10	LYS	C	Cg	24.772
10	LYS	H	Hba	1.626
10	LYS	C	Cb	33.324
10	LYS	C	Cd	29.227
10	LYS	H	Hdb	1.592
10	LYS	H	Hda	1.575
10	LYS	N	N	122.569
11	TYR	H	H	8.293
11	TYR	H	Ha	4.587
11	TYR	H	Hbb	2.94
11	TYR	H	Hba	2.862
11	TYR	C	Cb	39.172
11	TYR	H	Hd*	7.062
11	TYR	C	Cd*	133.289
11	TYR	H	He*	6.787
11	TYR	C	Ce*	118.162
11	TYR	N	N	121.114
12	SER	H	Hbb	3.787
12	SER	C	Cb	63.882
12	SER	H	Ha	4.372
12	SER	C	Ca	57.915
12	SER	H	H	8.171
12	SER	H	Hba	3.734
12	SER	N	N	116.984
13	PHE	H	H	8.115
13	PHE	H	Ha	4.596
13	PHE	H	Hbb	3.148
13	PHE	H	Hba	3.091
13	PHE	C	Cb	39.366
13	PHE	H	H _z	7.304
13	PHE	C	C _z	129.977
13	PHE	H	He*	7.363
13	PHE	C	Ce*	131.5
13	PHE	H	Hd*	7.251
13	PHE	C	Cd*	131.876
13	PHE	N	N	121.461
14	LYS	H	Ha	4.225
14	LYS	C	Ca	56.375
14	LYS	H	H	8.109
14	LYS	H	Hea	2.946

Residue sequence code	Residue label	Atom name	Atom type	Chemical shift value (ppm)
14	LYS	H	Heb	2.961
14	LYS	C	Ce	42.109
14	LYS	H	Hbb	1.767
14	LYS	H	Hgb	1.305
14	LYS	H	Hga	1.271
14	LYS	C	Cg	24.808
14	LYS	H	Hba	1.679
14	LYS	C	Cb	33.249
14	LYS	C	Cd	29.122
14	LYS	H	Hda	1.619
14	LYS	H	Hdb	1.635
14	LYS	N	N	121.002
15	LEU	H	Ha	4.349
15	LEU	C	Ca	55.175
15	LEU	H	H	8.092
15	LEU	H	Hbb	1.585
15	LEU	H	Hba	1.566
15	LEU	H	Hg	1.542
15	LEU	H	Hdb*	0.875
15	LEU	H	Hda*	0.846
15	LEU	C	Cb	42.36
15	LEU	C	Cda	24.875
15	LEU	C	Cdb	24.706
15	LEU	N	N	122.804
16	ILE	H	Ha	4.189
16	ILE	C	Ca	60.863
16	ILE	H	H	8.194
16	ILE	H	Hb	1.854
16	ILE	C	Cg1	27.17
16	ILE	H	Hg1b	1.456
16	ILE	H	Hg1a	1.149
16	ILE	H	Hg2*	0.878
16	ILE	H	Hd1*	0.835
16	ILE	C	Cg2	17.519
16	ILE	C	Cd1	12.764
16	ILE	C	Cb	38.773
16	ILE	N	N	122.547
17	LEU	H	Ha	4.377
17	LEU	C	Ca	54.895
17	LEU	H	H	8.269
17	LEU	H	Hbb	1.585
17	LEU	H	Hba	1.537
17	LEU	C	Cb	42.509
17	LEU	H	Hg	1.509
17	LEU	H	Hdb*	0.864
17	LEU	H	Hda*	0.828
17	LEU	C	Cdb	23.875
17	LEU	C	Cda	23.567

Residue sequence code	Residue label	Atom name	Atom type	Chemical shift value (ppm)
17	LEU	C	Cdb	23.567
17	LEU	N	N	126.128
18	ALA	H	Ha	4.336
18	ALA	C	Ca	52.179
18	ALA	H	H	8.236
18	ALA	H	Hb*	1.33
18	ALA	C	Cb	19.801
18	ALA	N	N	124.606
19	GLU	H	Ha	4.361
19	GLU	H	H	8.243
19	GLU	H	Hga	2.172
19	GLU	H	Hgb	2.253
19	GLU	C	Cg	35.881
19	GLU	H	Hbb	1.99
19	GLU	H	Hba	1.861
19	GLU	C	Cb	30.678
19	GLU	N	N	119.894
20	TYR	H	Ha	4.395
20	TYR	C	Ca	59.134
20	TYR	H	H	7.788
20	TYR	H	Hbb	3.046
20	TYR	H	Hba	2.892
20	TYR	C	Cb	39.564
20	TYR	H	Hd*	7.063
20	TYR	C	Cd*	133.224
20	TYR	H	He*	6.754
20	TYR	C	Ce*	117.948
20	TYR	N	N	125.098

A.6 Full assigned chemical shift list of 9R-VIPER L6AE10A

Residue sequence code	Residue label	Atom name	Atom type	Chemical shift value (ppm)
1	ARG	H	Ha	4.389
1	ARG	C	Ca	56.236
1	ARG	H	H*	8.811
1	ARG	H	Hda	3.208
1	ARG	H	Hdb	3.208
1	ARG	H	Hba	1.801
1	ARG	H	Hbb	1.801
1	ARG	H	Hga	1.647
1	ARG	H	Hgb	1.647
2	ARG	H	H	8.606
2	ARG	H	Ha	4.352
2	ARG	N	N	124.095
2	ARG	H	He	7.242
2	ARG	N	Ne	134.692
2	ARG	H	Hda	3.209
2	ARG	H	Hdb	3.209
2	ARG	C	Cd	43.233
2	ARG	C	Cb	30.945
2	ARG	H	Hbb	1.824
2	ARG	H	Hba	1.766
2	ARG	C	Cg	26.211
2	ARG	H	Hga	1.662
2	ARG	H	Hgb	1.662
8	ARG	H	H	8.502
8	ARG	H	Ha	4.344
8	ARG	C	Ca	56.005
8	ARG	N	N	123.57
8	ARG	H	He	7.197
8	ARG	N	Ne	134.692
8	ARG	H	Hdb	3.189
8	ARG	H	Hda	3.188
8	ARG	C	Cd	43.419
8	ARG	H	Hbb	1.821
8	ARG	H	Hba	1.756
8	ARG	C	Cb	31.071
8	ARG	C	Cg	27.152
8	ARG	H	Hga	1.613
8	ARG	H	Hgb	1.657
9	ARG	H	H	8.416
9	ARG	H	Ha	4.284
9	ARG	C	Ca	56.005
9	ARG	N	N	123.114
9	ARG	N	Ne	134.673
9	ARG	H	He	7.118
9	ARG	H	Hda	3.141

Residue sequence code	Residue label	Atom name	Atom type	Chemical shift value (ppm)
9	ARG	H	Hdb	3.141
9	ARG	C	Cd	43.387
9	ARG	C	Cb	31.156
9	ARG	H	Hba	1.715
9	ARG	H	Hbb	1.734
9	ARG	H	Hga	1.544
9	ARG	H	Hgb	1.565
9	ARG	C	Cg	27.065
10	LYS	C	Ca	56.48
10	LYS	H	Ha	4.227
10	LYS	H	H	8.351
10	LYS	N	N	122.908
10	LYS	H	Hea	2.947
10	LYS	H	Heb	2.947
10	LYS	C	Ce	42.054
10	LYS	H	Hda	1.638
10	LYS	H	Hdb	1.638
10	LYS	C	Cd	29.099
10	LYS	H	Hba	1.668
10	LYS	H	Hbb	1.668
10	LYS	C	Cb	33.228
10	LYS	H	Hga	1.229
10	LYS	H	Hgb	1.229
10	LYS	C	Cg	24.756
11	TYR	H	H	8.26
11	TYR	H	Ha	4.567
11	TYR	N	N	121.562
11	TYR	H	Hbb	2.943
11	TYR	H	Hba	2.87
11	TYR	C	Cb	39.171
11	TYR	H	Hd*	7.075
11	TYR	C	Cd*	133.333
11	TYR	H	He*	6.791
11	TYR	C	Ce*	118.006
12	SER	H	H	8.139
12	SER	H	Ha	4.397
12	SER	C	Ca	57.8
12	SER	N	N	122.336
12	SER	N	N	117.421
12	SER	N	N	117.421
12	SER	H	Hbb	3.774
12	SER	H	Hba	3.741
12	SER	C	Cb	63.911
13	PHE	H	H	8.147
13	PHE	H	Ha	4.573
13	PHE	N	N	122.336
13	PHE	C	Cb	39.568
13	PHE	H	Hbb	3.114

Residue sequence code	Residue label	Atom name	Atom type	Chemical shift value (ppm)
13	PHE	H	Hba	3.074
13	PHE	H	He*	7.358
13	PHE	C	Ce*	131.456
13	PHE	H	Hd*	7.256
13	PHE	C	Cd*	131.889
13	PHE	H	H _z	7.3
13	PHE	C	C _z	129.95
14	LYS	H	Ha	4.198
14	LYS	C	Ca	56.181
14	LYS	H	H	8.094
14	LYS	N	N	122.512
14	LYS	H	Hea	2.963
14	LYS	H	Heb	2.963
14	LYS	C	Ce	42.105
14	LYS	H	Hdb	1.636
14	LYS	H	Hda	1.634
14	LYS	C	Cd	29.195
14	LYS	H	Hbb	1.752
14	LYS	H	Hba	1.656
14	LYS	C	Cb	33.103
14	LYS	C	Cg	24.686
14	LYS	H	Hga	1.315
14	LYS	H	Hgb	1.315
15	ALA	H	H	8.132
15	ALA	H	Ha	4.255
15	ALA	N	N	125.268
15	ALA	C	Ca	52.45
15	ALA	H	Hb*	1.352
15	ALA	C	Cb	19.189
16	ILE	H	H	8.082
16	ILE	H	Ha	4.11
16	ILE	C	Ca	61.157
16	ILE	N	N	120.225
16	ILE	H	Hb	1.847
16	ILE	C	Cb	38.633
16	ILE	H	Hg1b	1.48
16	ILE	H	Hg1a	1.178
16	ILE	C	Cg1	27.26
16	ILE	H	Hg2*	0.882
16	ILE	C	Cg2	17.476
16	ILE	C	Cd1	12.818
16	ILE	H	Hd1*	0.854
17	LEU	H	Ha	4.35
17	LEU	C	Ca	54.897
17	LEU	H	H	8.222
17	LEU	N	N	126.042
17	LEU	H	Hba	1.578
17	LEU	H	Hbb	1.578

Residue sequence code	Residue label	Atom name	Atom type	Chemical shift value (ppm)
17	LEU	C	Cb	42.339
17	LEU	C	Cdb	24.856
17	LEU	C	Cda	23.48
17	LEU	H	Hda*	0.843
17	LEU	H	Hdb*	0.9
18	ALA	H	H	8.143
18	ALA	H	Ha	4.263
18	ALA	N	N	124.864
18	ALA	C	Ca	52.163
18	ALA	H	Hb*	1.322
19	ALA	H	H	8.112
19	ALA	H	Ha	4.272
19	ALA	C	Ca	52.31
19	ALA	N	N	123.441
19	ALA	H	Hb*	1.309
19	ALA	C	Cb	19.454
20	TYR	H	Ha	4.378
20	TYR	C	Ca	58.817
20	TYR	H	H	7.62
20	TYR	N	N	123.238
20	TYR	H	Hbb	3.065
20	TYR	H	Hba	2.91
20	TYR	C	Cb	39.345
20	TYR	H	Hd*	7.093
20	TYR	C	Cd*	133.229
20	TYR	H	He*	6.803
20	TYR	C	Ce*	118.139

A.7 Publications

Size Determination of Protein Oligomers/Aggregates Using Diffusion NMR Spectroscopy

Pancham S. Kandiyal, Ji Yoon Kim, Daniel L. Fortunati and K.H. Mok*

Running Title: DOSY NMR of Protein Oligomers/Aggregates

Affiliation (all authors):

Trinity Biomedical Sciences Institute (TBSI), School of Biochemistry & Immunology,
Trinity College Dublin,
Dublin 2, Ireland

* Corresponding Author: mok1@tcd.ie

Abstract

Diffusion-ordered spectroscopy (DOSY) is a widely used NMR technique for the identification of different chemical moieties/compounds contained in mixtures and has been successfully employed for the separation of small molecules based on hydrodynamic radii. Herein we show that DOSY can also be applied for the size determination of larger biomolecules such as proteins and protein oligomers/aggregates. Proof-of-principle is first shown with a cross-linked oligomeric protein mixture where the hydrodynamic volumes of each component are estimated and subsequently verified with size-exclusion HPLC and SDS polyacrylamide gel electrophoresis. We then determine the sizes of protein oligomers contained in a protein solution subjected under amyloid fibrillogenesis conditions. These studies aim to provide insight into the kinetics behind protein aggregation involved in amyloidoses as well as to determine the hydrodynamic radii of proteins within the mixture.

Key words

DOSY, Hydrodynamic radii, Protein, Oligomers, Aggregation, Diffusion coefficient, Pulsed field gradient (PFG), Stokes-Einstein equation.

1. Introduction

Nuclear magnetic resonance (NMR) spectroscopy is used for characterization of conformational and atomic level details of small molecules, peptides and proteins in solution. Availability of various multi-dimensional, multi-nuclear experiments and isotopically labelled techniques have made life easier to study the 3D structure of macromolecules using NMR spectroscopy (1). Diffusion is an intrinsic property of all molecules and varies on the basis of a particle's size, shape, solubility, charge distribution, etc (2). The theoretical and experimental demonstration for self-diffusion property of molecules was first given by Stejskal and Tanner in 1965 (3). Advancement in hardware and software technology in NMR have made the performing of diffusion experiments straightforward for small molecules, leading to 1) the determination of the mobility of a compound, 2) binding of small molecules to protein, 3) aggregation of molecules and 4) identification of individual compounds in a mixture (Fig.1) (4).

One of the key technical features in a diffusion NMR experiment pulse sequence is the utilization of pulsed-field gradients (PFGs) that spatially-encode the solutes, the magnetizations of which are then re-focused through either a spin echo or stimulated-echo of the spins (PFGSE or PFGSTE) (5). For the past two decades, DOSY has been implemented for the characterization of mixtures containing small molecules (6,7). In contrast, its application for macromolecules has been less due to the shorter relaxation times and the overlapping of chemical shifts, despite the fact that in principle, DOSY NMR can be very helpful to understand the rate of protein aggregation and oligomer formation. While previous work with DOSY NMR has been able to show the separation of proteins and lipids, samples containing a mixture of different size proteins have not been shown with any frequency (8). Here we have implemented 2D DOSY NMR to determine the size of an intentionally-prepared, chemically-cross-linked oligomeric mixture of a protein by addition of a cross linker (9). For our experiment, we used DOSY Bipolar Pulse Pair STimulated Echo (Dbppste) pulse sequence (10) and optimized the experiment by varying different parameters such as the diffusion gradient length, gradient diffusion delay time, relaxation delay and gradient recovery time.

When samples contain a mixture of different proteins, it is hard to readily distinguish all of the different sized-proteins according as a function of their diffusion coefficient value (D value) due to the large overlapping of their chemical shifts. For our cross-linked sample contained a mixture of overlapping signals originating from monomers, dimers and trimers of a protein, extracting the diffusion coefficients from requires deconvolution of the data. By plotting the intensities/peak areas obtained against the square of the gradient strength (G^2) according to the Stejskal-Tanner equation and fitting the data points with multiple gaussian curves, it was possible to capture more subtle differences in diffusion coefficients (D) which would not be accessible using a less involved analysis. The D value used further for calculating the hydrodynamic radii using the Stokes-Einstein equation (11).

From these data and the exponential/ gaussian fitting of them, we were able to extract three different sizes of component protein *i.e.* monomer, dimer and trimer for lysozyme, in spite that these proteins did not separate in 2D DOSY experiment on the basis of their diffusion coefficient value.

2. Materials

All reagents should be prepared using ultra-pure water.

1. Proteins; Hen egg-white lysozyme (HEWL; lyophilized powder, protein $\geq 90\%$, $\geq 40,000$ units/mg protein), bovine thyroglobulin, bovine γ -globulins, human serum albumin, β -lactoglobulin, horse cytochrome c, L-tryptophan ,

2. Deuterium oxide (D_2O)

3. 1,4-dioxane.

4. The water-soluble, homobifunctional crosslinker bis(sulfosuccinimidyl) substrate (BS^3). BS^3 contains sulfo-NHS (N-hydroxysuccinimide) esters on both carboxylate ends of the molecule, providing greater solubility due to the negatively-charged sulfonate groups, making the molecule very hydrophilic.

5. Chromatography was performed on a Dionex Ultimate HPLC 3000 Standard System running Chromeleon 6 software (Dionex, Thermo Scientific), and the size exclusion-HPLC (SE-HPLC) was a Superdex 200 10/300 GL column (10 mm x 30 cm, Particle size 13 μm , GE Healthcare).

6. SDS Gel Electrophoresis gels, tank and power supply.

7. NMR experiments were carried out on an Agilent Technologies 18.8 T (800 MHz) DD2 Premium Compact spectrometer with a triple-resonance, 5 mm enhanced cold probe, however any high-field magnet with pulsed-field gradient capabilities will be sufficient. (The interest is more in monitoring the decay of signal as a function of applied gradient than the resolution of the spectra.).

8. HPLC elution buffer; 0.05 M sodium phosphate at pH 7.0 containing 0.1 M Na_2SO_4 , filtered through a 0.45 μm (or 0.2 μm) pore-size membrane and degassed.

9. Conjugation buffer and; 0.1 M sodium phosphate buffer at pH 7.0 containing 0.15 M NaCl, filtered through a 0.45 μm filter.

10. Quenching buffer for crosslinking with BS^3 ; 1M Tris•HCl, pH 7.5 filtered through a 0.45 μm (or 0.2 μm) pore-size membrane.

11. Solutions for NMR; 0.05 M sodium phosphate buffer at pH 7.0 filtered through a 0.45 μm (or 0.2 μm) pore-size membrane. Other buffers are also acceptable provided that the buffer does not contain components (usually of aliphatic nature) that have strong, unexchangeable 1H spectral signals. All solutions for NMR should contain at least 10% of D_2O that will be used as a deuterium lock signal.

3. Methods

3.1 Initial diffusion coefficient value estimation from SE-HPLC data

An initial, rough estimate of the expected diffusion coefficients of the cross-linked HEWL species can be obtained from analysis of size exclusion-HPLC (SE-HPLC) data.

1. Dissolve the calibration standards at a concentration of 0.1 mM with the elution buffer and inject onto the SE-HPLC column eluting with 0.05 M sodium phosphate buffer pH 7.0 containing 0.1 M Na₂SO₄ at a flow rate of 0.50 mL/min with detection at 280 nm.
2. Generate a standard calibration curve using a suitable range of protein molecular weight standards (for example, bovine thyroglobulin, bovine γ-globulins, human serum albumin, β-lactoglobulin, horse cytochrome c, L-tryptophan) including human serum albumin (HSA), which has a hydrodynamic radius (R_H) of 40 Å (Figs. 2 and 3). The calibration curve is a plot of R_H values of these proteins vs the elution volume obtained from the SE-HPLC.
3. Prepare cross-linked forms of HEWL using BS³. For a HEWL concentration of 5 mg/mL, use a 10-fold molar excess of BS³ crosslinker and react at room temperature for 3 hours. Quench the reaction using the quenching buffer to a final concentration of 20-50 mM Tris. Dialyse the protein solution with 0.05 M sodium phosphate buffer pH 7.0 (or any buffer that will keep the protein well-solubilized).
4. Run a SDS-PAGE to observe the extent of cross-linking.
5. Inject an approximate 0.1 mM concentration sample of cross-linked HEWL onto the SE-HPLC and measure the retention times (see **Note 1**).
6. Obtain the R_H values by using the linear relationship found on the standard calibration curve (Fig 3). From the Stokes-Einstein equation, a rough estimate of the diffusion coefficients (D) for each species is as follows:

$$D = \frac{K_b T}{6\pi\eta R_H}$$

where K_b is Boltzmann's constant; T is the absolute temperature; η is the dynamic viscosity; and R_H is the radius of the spherical particle.

Typical retention times and corresponding hydrodynamic radii/diffusion coefficients obtained shown in Table 1.

3.2. DOSY-NMR and the identification of protein oligomer/aggregate components in a mixture

Background

The diffusion coefficient (D) of a single species of molecule can be extracted from DOSY-NMR data by plotting the Intensity of the peaks (I) obtained against the square of the gradient strength (g^2), according to the Stejskal-Tanner equation (3):

$$I = I_0 e^{-D\gamma^2 g^2 \delta^2 (\Delta - \delta/3)} \quad (1)$$

which takes the general form:

$$y = A e^{-Qx^2} \quad (2)$$

where Q corresponds to the diffusion coefficient multiplied by $(\gamma^2 \delta^2 (\Delta - \delta/3))$. (Here, γ is the gyromagnetic ratio, Δ is diffusion time given, and δ is the time for the gradient pulse.)

For a mixture of different components, the graph will represent a summation of different Gaussian curves (polygaussian), so the equation will become:

$$\sum i = A_1 e^{-Q_1 x^2} + A_2 e^{-Q_2 x^2} + \dots + A_i e^{-Q_i x^2} \quad (3)$$

To fit the DOSY-NMR data of a mixture of components with an equation of this type, a sequential approach is taken.

1. Dissolve the HEWL samples in 50 mM sodium phosphate buffer (pH 7.4, 90 % H_2O , 10 % D_2O), but the pH and buffer of the sample should be optimized to the conditions necessary for the oligomers/aggregates (see **Note 2**).

2. Perform diffusion-ordered spectroscopy (DOSY) measurements at 293 K. The DgscteSL_dpfgsc DOSY pulse program (Agilent VNMR) is used, which consists of gradient compensated stimulated echo with spin lock using the excitation sculpting solvent suppression method (12). A spectral window of 13020 Hz was used, with an acquisition time of 2.46 s with a relaxation delay of 3 s. The FIDs were collected with 32000 complex data

points with 64 scans. Logarithmically the gradient pulse strength was increased from 3 % to 86 % of the maximum strength of 32767 G/cm in 60 steps. A diffusion time (Δ) of 100 ms and bipolar half-sine-shaped gradient pulses (δ) of 5 ms were applied. 1,4-Dioxane, which is known to behave independently of protein concentration and the folded state of the protein, was used as an internal chemical shift reference and hydrodynamic radius calibration reference (3.75 ppm; $R_H = 2.12 \text{ \AA}$) (11, 13). Three replicate acquisitions were given for each sample, and the resulting diffusion coefficient (D) values calculated.

3. Upon acquisition of the NMR spectra, fit first a sample of neat dioxane with a single Gaussian to obtain the diffusion value for dioxane (Fig. 4). By referencing to dioxane it is then possible to obtain relative diffusion values/hydrodynamic radii for other species without having to correct for all the other factors contained in Q (which is instead necessary to obtain absolute measurements) (see **Note 3**).

4. Running the identical experiment and observing dioxane in a protein mixture sample clearly results in a sum of different poly-gaussian curves (Fig. 5). By fixing the diffusion coefficient of one of the values, it is possible to extract the diffusion coefficient for the second component, which (proportionally to the pure dioxane value) is comparable to that of lysozyme. The mixture of cross-linked HEWL was fit by sequentially adding curves to an initial (imperfect) fit, fixing values as fits got better and narrowing the range towards the values obtained through SE-HPLC (Fig. 6).

4. Notes

1. Compared to the SDS PAGE results, which reflect disulfide bridge-reduced and fully-denatured proteins (Fig. 7), the separation of monomer, dimer and trimer of HEWL showed that the hydrodynamic radii of these species (in their non-denatured/reduced state) were not as widely distributed (Figs. 8 and 9). Nevertheless, not only was it possible to distinguish different species in the shoulders of the peak obtained (Figs. 9 and 10), one is able to identify a further-compacted monomer which appeared to have resulted from intramolecular covalent cross-linkages.

2. When performing NMR, it would be advantageous to use a buffer which does not have interfering signals with the sample of interest.

3. The D (diffusion coefficient) values for 1,4-dioxane are slightly variable dependent upon the co-solute, which rightly reflects the different solution micro-environment conditions that both solutes are mutually experiencing for each sample.

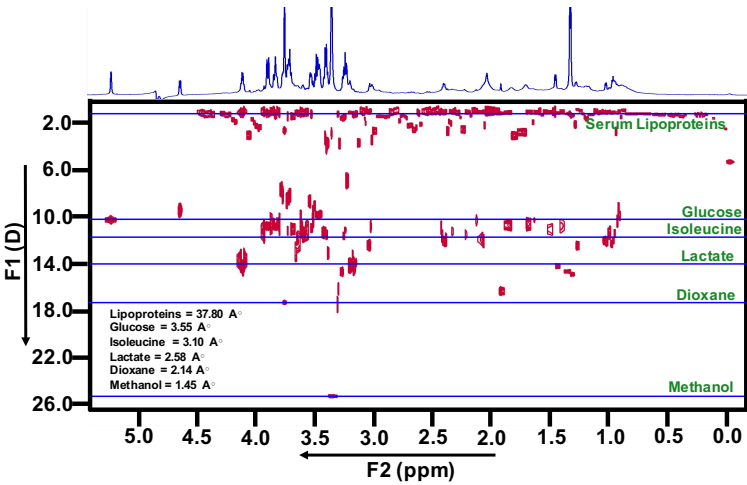


Figure 1. Separation of metabolites and lipoproteins in a serum sample using DOSY NMR

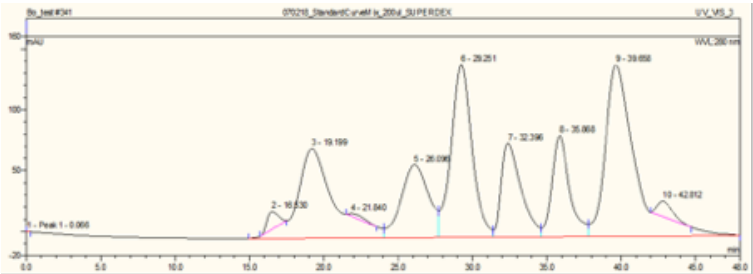


Figure 2. Protein calibration standard chromatogram using size-exclusion HPLC

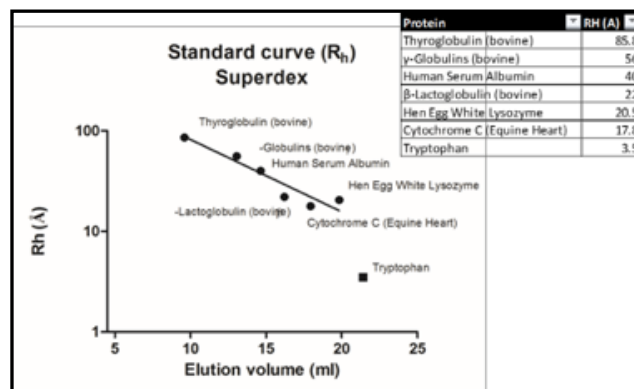
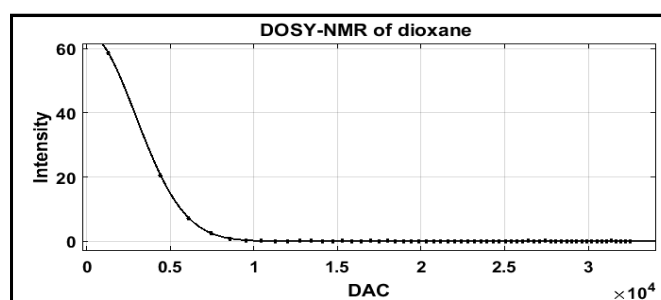


Figure 3. SE-HPLC standard curve

Figure 4. DOSY-NMR fitting of pure 1,4-dioxane. Pure dioxane fit with a single curve (equation 2). The diffusion coefficient (un-adjusted) obtained from this fit was 5.868×10^{-8} .

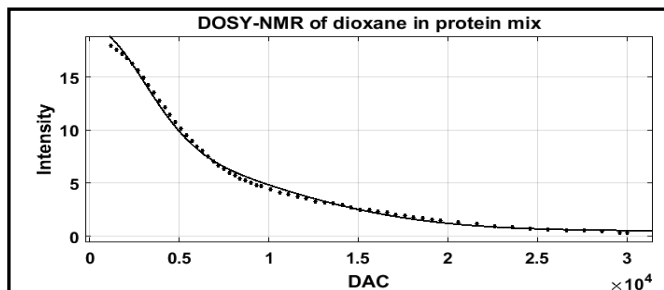


Figure 5. DOSY-NMR fitting of 1,4-dioxane in a protein solution. Dioxane from the protein mix, fit with a poly-gaussian (equation 3). In this case, it is clear that the fit could not be a single Gaussian like in the case of the pure dioxane sample, probably due to protein peaks overlapping in the dioxane region. By fixing the diffusion value of one of the components of the poly-gaussian, it was possible to obtain a diffusion values for a second component, equal to 6.137×10^{-9}

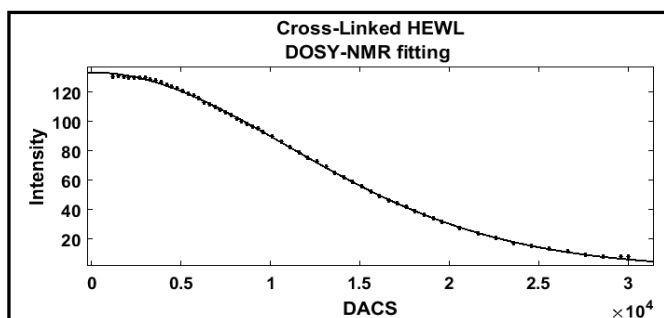


Figure 6. DOSY-NMR fitting of cross-linked hen egg white lysozyme. Poly-gaussian fitting of cross-linked HEWL DOSY-NMR data. A four component polygaussian was chosen to approximate the population of crosslinked species present in the sample (although probably more different species than this were present, the SE-HPLC and SDS PAGE results suggested 4 major species to be present in the sample). The initial Gaussian was fixed at the value obtained for the second component of the dioxane with HEWL sample, and other components were added sequentially, increasing the quality of the fit while narrowing the range of values of the possible diffusion coefficients used. Diffusion coefficients obtained are presented in Table 2.

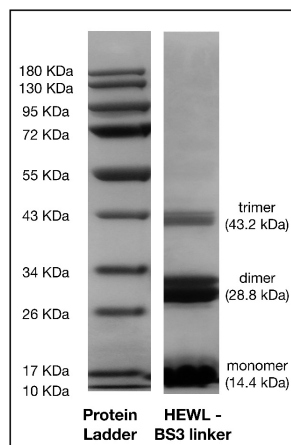


Figure 7. SDS-PAGE Gel of the selected HEWL-BS3 cross-linked mixture sample

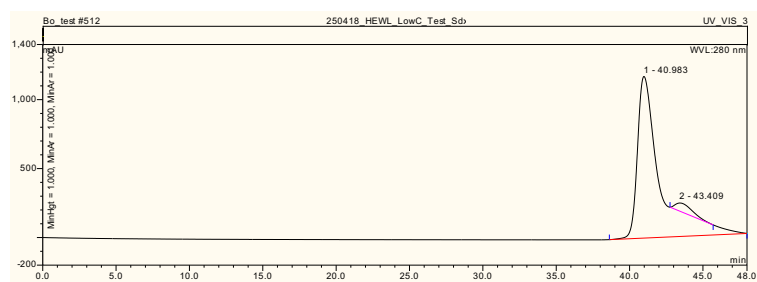


Figure 8. Size-exclusion chromatogram of monomeric HEWL



Figure 9. Size-exclusion chromatogram of a mixture of cross-linked HEWL (trimer, dimer, monomer)



Figure 10. Size-exclusion chromatogram of mixture of cross-linked HEWL (trimer & dimer enriched via centrifugal concentration)

Table 1. Rough estimates of hydrodynamic radii/diffusion coefficients obtained from SE-HPLC results

HEWL Peak	R_t (min)	R_h (Å)	D ($10^{10} \text{ m}^2 \text{ s}^{-1}$)
Trimer	36.5	28	0.87
Dimer	39.0	23	1.10
Monomer	41.0	20	1.21
Compacted Monomer	42.9	17	1.42

Table 2. Diffusion coefficients/hydrodynamic radii obtained from DOSY-NMR

Species	$D \times 10^9$ (a.u.)	R_h (Å)	I_0 (a.u.)
Dioxane	58.7	2.12	64.8
HEWL1 (Trimer)	3.0	41.4	82.5
HEWL2 (Dimer)	4.8	25.9	16.19
HEWL3 (Mon)	6.1	20.3	34.97
HEWL4 (Compact Monomer)	6.5	19.1	0.4

5. References

1. Cavanagh J, Fairbrother WJ, Palmer III AG, Rance M, Skelton NJ (2006) Protein NMR Spectroscopy (2nd Ed): Principles and Practice. Academic Press, New York.
2. Li D, Keresztes I, Hopson R, Williard PG (2009) Characterization of reactive intermediates by multinuclear diffusion-ordered NMR spectroscopy (DOSY). *Acc Chem Res* 42(2): 270-280.
3. Stejskal EO, Tanner JE (1965) Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *J Chem Phys* 42: 288-292.
4. Dehner A, Kessler H (2005) Diffusion NMR Spectroscopy: Folding and Aggregation of Domains in p53. *ChemBioChem*. 6, 1550-1565.
5. Pagès G, Gilard V, Martinob R, Malet-Martino M (2017) Pulsed-field gradient nuclear magnetic resonance measurements (PFG NMR) for diffusion ordered spectroscopy (DOSY) mapping. *Analyst*. 142, 3771.
6. Morris KF, Johnson Jr CS (1992) Diffusion-Ordered Two-Dimensional Nuclear Magnetic Resonance Spectroscopy. *JACS* 114: 3139-3141.
7. Barjat H, Morris GA, Smart SC, Swanson AG, Williams SCR (1995) High Resolution Diffusion Ordered 2D Spectroscopy (HR-DOSY)-A New Tool for the Analysis of Complex Mixtures. *J Magn Reson B* 108: 170-1.
8. Balayssac S, Delsuc M-A, Gilard V, Prigent Y, Malet-Martino M (2009) Two-dimensional DOSY experiment with Excitation Sculpting water suppression for the analysis of natural and biological media. *J Magn Reson* 196: 78-83.
9. Arora B, Tandon R, Attri P and Bhatia R (2017) Chemical Crosslinking: Role in Protein and Peptide Science. *Curr Protein Pept Sci* 18(9): 946-955.
10. D. Wu, A. Chen and C. S. Johnson, Jr (1995) An improved diffusion-ordered spectroscopy experiment incorporating bipolar-gradient pulses. *J Magn Reson A* 115(2): 260-226.
11. Jones JA, Wilkins DK, Smith LJ, CM (1997) Characterization of protein unfolding by NMR diffusion measurement. *J Biomol NMR* 10: 199-203.

12. Pelta MD, Barjat H, Morris GA, Davis AL, Hammond SJ (1998) Pulse sequences for high-resolution diffusion-ordered spectroscopy (HR-DOSY). *Magn Reson Chem* 36: 706-714.
13. Shimizu A, Ikeguchi M, Sugai S (1994) Appropriateness of DSS and TSP as internal references for (^1H) NMR studies of molten globule proteins in aqueous media. *J Biomol NMR* 4: 859-862

Bibliography

- [1] J. M. BERG, J. L. TYMOCZKO, and L. STRYER. *Biochemistry*. Basingstoke: W.H. Freeman, 2012. (see pp. [23–26](#), [139](#))
- [2] D. L. NELSON, M. M. COX, and A. L. LEHNINGER. *Lehninger Principles of Biochemistry*. 7th ed. Houndmills, Basingstoke: Macmillan Higher Education, 2017. (see pp. [23–26](#))
- [3] M. P. WILLIAMSON. Peptide Structure Determination by NMR in: *Spectroscopic Methods and Analyses*. ed. by J. JONES, B. MULLOY, and A. THOMAS. **17**: 69–85, New Jersey: Humana Press, 1993. DOI: [10.1385/0-89603-215-9:69](#) (see pp. [23](#), [139](#))
- [4] K. WÜTHRICH. The way to NMR structures of proteins. *Nature Structural Biology*, **8**: 923–925, 2001. DOI: [10.1038/nsb1101-923](#) (see pp. [23](#), [41](#))
- [5] D. G. REID. Protein NMR Techniques in: *Methods in Molecular Biology*. **60**: New Jersey: Humana Press, 1997. DOI: [10.1385/0896033090](#) (see pp. [23](#), [39–41](#))
- [6] C. M. DOBSON. Protein folding and misfolding. *Nature*, **426**: 884–890, 2003. DOI: [10.1038/nature02261](#) (see pp. [26](#), [27](#), [36](#), [125](#))
- [7] C. M. DOBSON. Principles of protein folding, misfolding and aggregation. *Seminars in Cell & Developmental Biology*, **15**: 3–16, 2004. DOI: [10.1016/j.semcd.2003.12.008](#) (see pp. [26](#), [36](#), [125](#))
- [8] E. H. KOO, P. T. LANSBURY, and J. W. KELLY. Amyloid diseases: Abnormal protein aggregation in neurodegeneration. *Proceedings of the National Academy of Sciences*, **96**: 9989–9990, 1999. DOI: [10.1073/pnas.96.18.9989](#) (see p. [26](#))
- [9] F. CHITI and C. M. DOBSON. Protein Misfolding, Functional Amyloid, and Human Disease. *Annual Review of Biochemistry*, **75**: 333–366, 2006. DOI: [10.1146/annurev.biochem.75.101304.123901](#) (see pp. [27–29](#), [69](#))
- [10] S. Y. MIN, J. MEEHAN, L. M. SULLIVAN, N. P. HARTE, Y. J. XIE, G. P. DAVEY, C. SVANBORG, A. BRODKORB, and K. H. MOK. Alternatively folded proteins with unexpected beneficial functions. *Biochemical Society Transactions*, **40**: 746–751, 2012. DOI: [10.1042/BST20120029](#) (see p. [26](#))

-
- [11] C. IANNUZZI, R. MARITATO, G. IRACE, and I. SIRANGELO. Misfolding and Amyloid Aggregation of Apomyoglobin. *International Journal of Molecular Sciences*, **14**: 14287–14300, 2013. DOI: [10.3390/ijms140714287](https://doi.org/10.3390/ijms140714287) (see pp. 28, 73)
- [12] R. M. MURPHY. Kinetics of amyloid formation and membrane interaction with amyloidogenic proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, **1768**: 1923–1934, 2007. DOI: [10.1016/j.bbamem.2006.12.014](https://doi.org/10.1016/j.bbamem.2006.12.014) (see pp. 28, 69)
- [13] R. KAYED. Common Structure of Soluble Amyloid Oligomers Implies Common Mechanism of Pathogenesis. *Science*, **300**: 486–489, 2003. DOI: [10.1126/science.1079469](https://doi.org/10.1126/science.1079469) (see p. 29)
- [14] N. P. HARTE, I. KLYUBIN, E. K. MCCARTHY, S. Y. MIN, S. A. GARRAHY, Y. J. XIE, G. P. DAVEY, J. J. BOLAND, M. J. ROWAN, and K. H. MOK. Amyloid Oligomers and Mature Fibrils Prepared from an Innocuous Protein Cause Diverging Cellular Death Mechanisms. *Journal of Biological Chemistry*, **290**: 28343–28352, 2015. DOI: [10.1074/jbc.M115.676072](https://doi.org/10.1074/jbc.M115.676072) (see pp. 29, 30, 69, 73)
- [15] P. AROSIO, T. P. J. KNOWLES, and S. LINSE. On the lag phase in amyloid fibril formation. *Physical Chemistry Chemical Physics*, **17**: 7606–7618, 2015. DOI: [10.1039/C4CP05563B](https://doi.org/10.1039/C4CP05563B) (see p. 29)
- [16] J. S. PEDERSEN, C. B. ANDERSEN, and D. E. OTZEN. Amyloid structure - one but not the same: the many levels of fibrillar polymorphism. *FEBS Journal*, **277**: 4591–4601, 2010. DOI: [10.1111/j.1742-4658.2010.07888.x](https://doi.org/10.1111/j.1742-4658.2010.07888.x) (see p. 29)
- [17] I. V. BASKAKOV. Branched chain mechanism of polymerization and ultrastructure of prion protein amyloid fibrils. *FEBS Journal*, **274**: 3756–3765, 2007. DOI: [10.1111/j.1742-4658.2007.05916.x](https://doi.org/10.1111/j.1742-4658.2007.05916.x) (see p. 29)
- [18] C. B. ANDERSEN, H. YAGI, M. MANNO, V. MARTORANA, T. BAN, G. CHRISTIANSEN, D. E. OTZEN, Y. GOTO, and C. RISCHER. Branching in Amyloid Fibril Growth. *Biophysical Journal*, **96**: 1529–1536, 2009. DOI: [10.1016/j.bpj.2008.11.024](https://doi.org/10.1016/j.bpj.2008.11.024) (see pp. 29, 35)
- [19] M. NILSSON. Techniques to study amyloid fibril formation in vitro. *Methods*, **34**: 151–160, 2004. DOI: [10.1016/j.ymeth.2004.03.012](https://doi.org/10.1016/j.ymeth.2004.03.012) (see pp. 29, 55, 74, 75, 89)
- [20] R. N. RAMBARAN and L. C. SERPELL. Amyloid fibrils: Abnormal protein assembly. *Prion*, **2**: 112–117, 2008. DOI: [10.4161/pri.2.3.7488](https://doi.org/10.4161/pri.2.3.7488) (see pp. 29, 34, 36)
- [21] M. FÄNDRICH, M. A. FLETCHER, and C. M. DOBSON. Amyloid fibrils from muscle myoglobin. *Nature*, **410**: 165–166, 2001. DOI: [10.1038/35065514](https://doi.org/10.1038/35065514) (see pp. 30, 53)
- [22] M. R. H. KREBS, D. K. WILKINS, E. W. CHUNG, M. C. PITKEATHLY, A. K. CHAMBERLAIN, J. ZURDO, C. V. ROBINSON, and C. M. DOBSON. Formation and seeding of amyloid fibrils from wild-type hen lysozyme and a peptide fragment from the β -domain. *Journal of Molecular Biology*, **300**: 541–549, 2000. DOI: [10.1006/jmbi.2000.3862](https://doi.org/10.1006/jmbi.2000.3862) (see pp. 30, 73)

-
- [23] R. SWAMINATHAN, V. K. RAVI, S. KUMAR, M. V. S. KUMAR, and N. CHANDRA. Lysozyme: A model protein for amyloid research in: *Advances in Protein Chemistry and Structural Biology*. ed. by R. DONEV. **84**: 63–111, Elsevier, 2011. DOI: [10.1016/B978-0-12-386483-3.00003-3](https://doi.org/10.1016/B978-0-12-386483-3.00003-3) (see p. 30)
- [24] C. REDFIELD and C. M. DOBSON. Sequential proton NMR assignments and secondary structure of hen egg white lysozyme in solution. *Biochemistry*, **27**: 122–136, 1988. DOI: [10.1021/bi00401a020](https://doi.org/10.1021/bi00401a020) (see p. 30)
- [25] A. A. MERATAN, A. GHASEMI, and M. NEMAT-GORGANI. Membrane Integrity and Amyloid Cytotoxicity: A Model Study Involving Mitochondria and Lysozyme Fibrillation Products. *Journal of Molecular Biology*, **409**: 826–838, 2011. DOI: [10.1016/j.jmb.2011.04.045](https://doi.org/10.1016/j.jmb.2011.04.045) (see p. 30)
- [26] M. F. ROSARIO-ALOMAR, T. QUIÑONES-RUIZ, D. KUROWSKI, V. SEREDA, E. B. FERREIRA, L. D. JESÚS-KIM, S. HERNÁNDEZ-RIVERA, D. V. ZAGOREVSKI, J. LÓPEZ-GARRIGA, and I. K. LEDNEV. Hydrogen Sulfide Inhibits Amyloid Formation. *The Journal of Physical Chemistry B*, **119**: 1265–1274, 2015. DOI: [10.1021/jp508471v](https://doi.org/10.1021/jp508471v) (see p. 30)
- [27] M. REFAEE, T. TEZUKA, K. AKASAKA, and M. P. WILLIAMSON. Pressure-dependent Changes in the Solution Structure of Hen Egg-white Lysozyme. *Journal of Molecular Biology*, **327**: 857–865, 2003. DOI: [10.1016/S0022-2836\(03\)00209-2](https://doi.org/10.1016/S0022-2836(03)00209-2) (see p. 31)
- [28] N. YAGI, N. OHTA, and T. MATSUO. Structure of amyloid fibrils of hen egg white lysozyme studied by microbeam X-ray diffraction. *International Journal of Biological Macromolecules*, **45**: 86–90, 2009. DOI: [10.1016/j.ijbiomac.2009.04.007](https://doi.org/10.1016/j.ijbiomac.2009.04.007) (see pp. 31, 36)
- [29] Ye ZOU, Wenying HAO, Haoyi LI, Yachao GAO, Ying SUN, and Gang MA. New Insight into Amyloid Fibril Formation of Hen Egg White Lysozyme Using a Two-Step Temperature-Dependent FTIR Approach. *The Journal of Physical Chemistry B*, **118**: 9834–9843, 2014. DOI: [10.1021/jp504201k](https://doi.org/10.1021/jp504201k) (see p. 31)
- [30] J. A. JONES, D. K. WILKINS, L. J. SMITH, and C. M. DOBSON. Characterisation of protein unfolding by NMR diffusion measurements. *Journal of Biomolecular NMR*, **10**: 199–203, 1997. DOI: [10.1023/A:1018304117895](https://doi.org/10.1023/A:1018304117895) (see pp. 31, 62)
- [31] J. F. SMITH, T. P. J. KNOWLES, C. M. DOBSON, C. E. MACPHEE, and M. E. WELLAND. Characterization of the nanoscale properties of individual amyloid fibrils. *Proceedings of the National Academy of Sciences*, **103**: 15806–15811, 2006. DOI: [10.1073/pnas.0604035103](https://doi.org/10.1073/pnas.0604035103) (see pp. 31, 33)
- [32] S. Y. OW and D. E. DUNSTAN. The effect of concentration, temperature and stirring on hen egg white lysozyme amyloid formation. *Soft Matter*, **9**: 9692, 2013. DOI: [10.1039/c3sm51671g](https://doi.org/10.1039/c3sm51671g) (see p. 31)

- [33] V. SIVALINGAM, N. L. PRASANNA, N. SHARMA, A. PRASAD, and B. K. PATEL. Wild-type hen egg white lysozyme aggregation in vitro can form self-seeding amyloid conformational variants. *Biophysical Chemistry*, **219**: 28–37, 2016. DOI: [10.1016/j.bpc.2016.09.009](https://doi.org/10.1016/j.bpc.2016.09.009) (see pp. [31](#), [70](#))
- [34] N. BANSAL and B. BHANDARI. Functional Milk Proteins: Production and Utilization–Whey-based Ingredients in: *Advanced Dairy Chemistry*. ed. by P. MCSWEENEY and J. O’MAHONY. 67–98, New York, NY: Springer New York, 2016. DOI: [10.1007/978-1-4939-2800-2_3](https://doi.org/10.1007/978-1-4939-2800-2_3) (see pp. [32](#), [113](#))
- [35] S. BROWNLOW, J. H. M. CABRAL, R. COOPER, D. R. FLOWER, S. J. YEW DALL, I. POLIKARPOV, A. C. T. NORTH, and L. SAWYER. Bovine β -lactoglobulin at 1.8 Å resolution—still an enigmatic lipocalin. *Structure*, **5**: 481–495, 1997. DOI: [10.1016/S0969-2126\(97\)00205-0](https://doi.org/10.1016/S0969-2126(97)00205-0) (see pp. [32](#), [113](#))
- [36] K. SAKURAI and Y. GOTO. Dynamics and Mechanism of the Tanford Transition of Bovine β -Lactoglobulin Studied using Heteronuclear NMR Spectroscopy. *Journal of Molecular Biology*, **356**: 483–496, 2006. DOI: [10.1016/j.jmb.2005.11.038](https://doi.org/10.1016/j.jmb.2005.11.038) (see pp. [32](#), [113](#), [115](#), [124](#))
- [37] T. V. BUROVA, Y. CHOISET, V. TRAN, and T. HAERTLE. Role of free Cys121 in stabilization of bovine beta-lactoglobulin B. *Protein Engineering Design and Selection*, **11**: 1065–1073, 1998. DOI: [10.1093/protein/11.11.1065](https://doi.org/10.1093/protein/11.11.1065) (see pp. [32](#), [113](#))
- [38] K. KUWATA, S. ERA, M. HOSHINO, V. FORGE, Y. GOTO, and C. A. BATT. Solution structure and dynamics of bovine β -lactoglobulin A. *Protein Science*, **8**: 2541–2545, 2008. DOI: [10.1110/ps.8.11.2541](https://doi.org/10.1110/ps.8.11.2541) (see pp. [32](#), [33](#), [124](#))
- [39] R.N. ZÚÑIGA, A. TOLKACH, U. KULOZIK, and J.M. AGUILERA. Kinetics of Formation and Physicochemical Characterization of Thermally-Induced β -Lactoglobulin Aggregates. *Journal of Food Science*, **75**: E261–E268, 2010. DOI: [10.1111/j.1750-3841.2010.01617.x](https://doi.org/10.1111/j.1750-3841.2010.01617.x) (see pp. [32](#), [112](#), [114](#))
- [40] J. ADAMCIK, J. M. JUNG, J. FLAKOWSKI, P. DE LOS RIOS, G. DIETLER, and R. MEZZENGA. Understanding amyloid aggregation by statistical analysis of atomic force microscopy images. *Nature Nanotechnology*, **5**: 423–428, 2010. DOI: [10.1038/nnano.2010.59](https://doi.org/10.1038/nnano.2010.59) (see pp. [32](#), [70](#), [91](#))
- [41] A. K. BUGGY, J. J. MCMANUS, A. BRODKORB, N. M. CARTHY, and M. A. FENELON. Stabilising effect of α -lactalbumin on concentrated infant milk formula emulsions heat treated pre- or post-homogenisation. *Dairy Science & Technology*, **96**: 845–859, 2017. DOI: [10.1007/s13594-016-0306-1](https://doi.org/10.1007/s13594-016-0306-1) (see pp. [32](#), [113](#), [114](#))

-
- [42] A. K. BUGGY, J. J. MCMANUS, A. BRODKORB, S. A. HOGAN, and M. A. FENELON. Pilot-scale formation of whey protein aggregates determine the stability of heat-treated whey protein solutions—Effect of pH and protein concentration. *Journal of Dairy Science*, **101**: 10819–10830, 2018. DOI: [10.3168/jds.2017-14177](https://doi.org/10.3168/jds.2017-14177) (see p. 32)
- [43] D. HAMADA and C. M. DOBSON. A kinetic study of β -lactoglobulin amyloid fibril formation promoted by urea. *Protein Science*, **11**: 2417–2426, 2009. DOI: [10.1110/ps.0217702](https://doi.org/10.1110/ps.0217702) (see p. 32)
- [44] L. WANG, Y. SUN, Z. LI, A. WU, and G. WEI. Bottom-Up Synthesis and Sensor Applications of Biomimetic Nanostructures. *Materials*, **9**: 53, 2016. DOI: [10.3390/ma9010053](https://doi.org/10.3390/ma9010053) (see p. 33)
- [45] U. SLOTTA, S. HESS, K. SPIESS, T. STROMER, L. SERPELL, and T. SCHEIBEL. Spider Silk and Amyloid Fibrils: A Structural Comparison. *Macromolecular Bioscience*, **7**: 183–188, 2007. DOI: [10.1002/mabi.200600201](https://doi.org/10.1002/mabi.200600201) (see p. 33)
- [46] C. L. CHEN and N. L. ROSI. Peptide-Based Methods for the Preparation of Nanostructured Inorganic Materials. *Angewandte Chemie International Edition*, **49**: 1924–1942, 2010. DOI: [10.1002/anie.200903572](https://doi.org/10.1002/anie.200903572) (see p. 34)
- [47] T. AUMÜLLER and M. FÄNDRICH. Catalytic amyloid fibrils. *Nature Chemistry*, **6**: 273–274, 2014. DOI: [10.1038/nchem.1904](https://doi.org/10.1038/nchem.1904) (see pp. 34, 69)
- [48] C. LI, J. ADAMCIK, and R. MEZZENGA. Biodegradable nanocomposites of amyloid fibrils and graphene with shape-memory and enzyme-sensing properties. *Nature Nanotechnology*, **7**: 421–427, 2012. DOI: [10.1038/nnano.2012.62](https://doi.org/10.1038/nnano.2012.62) (see pp. 34, 69)
- [49] M. SACKEWITZ, S. von EINEM, G. HAUSE, M. WUNDERLICH, F. X. SCHMID, and E. SCHWARZ. A folded and functional protein domain in an amyloid-like fibril. *Protein Science*, **17**: 1044–1054, 2008. DOI: [10.1110/ps.073276308](https://doi.org/10.1110/ps.073276308) (see p. 35)
- [50] S. M. PILKINGTON, S. J. ROBERTS, S. J. MEADE, and J. A. GERRARD. Amyloid fibrils as a nanoscaffold for enzyme immobilization. *Biotechnology Progress*, **26**: 93–100, 2009. DOI: [10.1002/btpr.309](https://doi.org/10.1002/btpr.309) (see p. 35)
- [51] C. M. RUFO, Y. S. MOROZ, O. V. MOROZ, J. STÖHR, T. A. SMITH, X. HU, W. F. DEGRADO, and I. V. KOREDOVYCH. Short peptides self-assemble to produce catalytic amyloids. *Nature Chemistry*, **6**: 303–309, 2014. DOI: [10.1038/nchem.1894](https://doi.org/10.1038/nchem.1894) (see p. 35)
- [52] M. W. LEE et al. Zinc-binding structure of a catalytic amyloid from solid-state NMR. *Proceedings of the National Academy of Sciences*, **114**: 6191–6196, 2017. DOI: [10.1073/pnas.1706179114](https://doi.org/10.1073/pnas.1706179114) (see p. 35)
- [53] R. E. SAMUEL, E. D. SALMON, and Robin W. BRIEHL. Nucleation and growth of fibres and gel formation in sickle cell haemoglobin. *Nature*, **345**: 833–835, 1990. DOI: [10.1038/345833a0](https://doi.org/10.1038/345833a0) (see p. 35)

-
- [54] R. JANSEN, W. DZWOLAK, and R. WINTER. Amyloidogenic Self-Assembly of Insulin Aggregates Probed by High Resolution Atomic Force Microscopy. *Biophysical Journal*, **88**: 1344–1353, 2005. DOI: [10.1529/biophysj.104.048843](https://doi.org/10.1529/biophysj.104.048843) (see pp. 35, 70)
- [55] J. E. GILLAM and C. E. MACPHEE. Modelling amyloid fibril formation kinetics: mechanisms of nucleation and growth. *Journal of Physics: Condensed Matter*, **25**: 373101, 2013. DOI: [10.1088/0953-8984/25/37/373101](https://doi.org/10.1088/0953-8984/25/37/373101) (see pp. 35, 73)
- [56] W. E. KLUNK, J. W. PETTEGREW, and D. J. ABRAHAM. Quantitative evaluation of congo red binding to amyloid-like proteins with a beta-pleated sheet conformation. *Journal of Histochemistry & Cytochemistry*, **37**: 1273–1281, 1989. DOI: [10.1177/37.8.2666510](https://doi.org/10.1177/37.8.2666510) (see pp. 36, 55)
- [57] L. HOU et al. Solution NMR Studies of the A β (140) and A β (142) Peptides Establish that the Met35 Oxidation State Affects the Mechanism of Amyloid Formation. *Journal of the American Chemical Society*, **126**: 1992–2005, 2004. DOI: [10.1021/ja036813f](https://doi.org/10.1021/ja036813f) (see pp. 36, 39, 70)
- [58] H. LEVINE III. Thioflavine T interaction with synthetic Alzheimer's disease β -amyloid peptides: Detection of amyloid aggregation in solution. *Protein Science*, **2**: 404–410, 1993. DOI: [10.1002/pro.5560020312](https://doi.org/10.1002/pro.5560020312) (see pp. 36, 37, 53, 70)
- [59] J. PENG, Y. XIONG, Z. LIN, L. SUN, and J. WENG. Few-layer bismuth selenides exfoliated by hemin inhibit amyloid- β 1–42 fibril formation. *Scientific Reports*, **5**: 10171, 2015. DOI: [10.1038/srep10171](https://doi.org/10.1038/srep10171) (see pp. 36, 70)
- [60] H. LEVINE III. Quantification of β -sheet amyloid fibril structures with thioflavin T in: *Methods in Enzymology: Amyloid, Prions, and Other Protein Aggregates*. ed. by R. WETZEL. **309**: chap. 18, 274–284, Elsevier, 1999. DOI: [10.1016/S0076-6879\(99\)09020-5](https://doi.org/10.1016/S0076-6879(99)09020-5) (see pp. 37, 53, 70, 73)
- [61] R. KHURANA, C. COLEMAN, C. IONESCU-ZANETTI, S. A. CARTER, V. KRISHNA, R. K. GROVER, R. ROY, and S. SINGH. Mechanism of thioflavin T binding to amyloid fibrils. *Journal of Structural Biology*, **151**: 229–238, 2005. DOI: [10.1016/j.jsb.2005.06.006](https://doi.org/10.1016/j.jsb.2005.06.006) (see p. 37)
- [62] M. R. H. KREBS, E. H. C. BROMLEY, and A. M. DONALD. The binding of thioflavin-T to amyloid fibrils: localisation and implications. *Journal of Structural Biology*, **149**: 30–37, 2005. DOI: [10.1016/j.jsb.2004.08.002](https://doi.org/10.1016/j.jsb.2004.08.002) (see pp. 37, 53, 70)
- [63] S. M. KELLY, T. J. JESS, and N. C. PRICE. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, **1751**: 119–139, 2005. DOI: [10.1016/j.bbapap.2005.06.005](https://doi.org/10.1016/j.bbapap.2005.06.005) (see pp. 38, 56, 57, 89)
- [64] B. M. BULHELLER, A. RODGER, and J. D. HIRST. Circular and linear dichroism of proteins. *Physical Chemistry Chemical Physics*, **9**: 2020, 2007. DOI: [10.1039/b615870f](https://doi.org/10.1039/b615870f) (see pp. 38, 56, 57, 89)

- [65] C. LOUIS-JEUNE, M. A. ANDRADE-NAVARRO, and C. PEREZ-IRATXETA. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins: Structure, Function, and Bioinformatics*, **80**: 374–381, 2012. DOI: [10.1002/prot.23188](https://doi.org/10.1002/prot.23188) (see pp. 38, 56, 57, 89)
- [66] N. J. GREENFIELD. Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols*, **1**: 2876–2890, 2006. DOI: [10.1038/nprot.2006.202](https://doi.org/10.1038/nprot.2006.202) (see pp. 38, 56, 57, 89)
- [67] V. I. DODERO, Z. B. QUIROLO, and M. A. SEQUEIRA. Biomolecular studies by circular dichroism. *Frontiers in Bioscience*, **16**: 61–73, 2011. DOI: [10.2741/3676](https://doi.org/10.2741/3676) (see p. 38)
- [68] R. PECORA, ed. *Dynamic Light Scattering*. Boston, MA: Springer US, 1985. DOI: [10.1007/978-1-4613-2389-1](https://doi.org/10.1007/978-1-4613-2389-1) (see p. 39)
- [69] B. J. BERNE and R. PECORA. *Dynamic light scattering : with applications to chemistry, biology and physics*. Mineola: Dover Publications, 2000. (see pp. 39, 55)
- [70] A. M. STREETS, Y. SOURIGUES, R. R. KOPITO, R. MELKI, and S. R. QUAKE. Simultaneous Measurement of Amyloid Fibril Formation by Dynamic Light Scattering and Fluorescence Reveals Complex Aggregation Kinetics. *PLOS ONE*, **8**: e54541, 2013. DOI: [10.1371/journal.pone.0054541](https://doi.org/10.1371/journal.pone.0054541) (see pp. 39, 79, 82)
- [71] E. BOISSELIER, A. K. DIALLO, L. SALMON, C. ORNELAS, J. RUIZ, and D. ASTRUC. Encapsulation and Stabilization of Gold Nanoparticles with “Click” Polyethyleneglycol Dendrimers. *Journal of the American Chemical Society*, **132**: 2729–2742, 2010. DOI: [10.1021/ja909133f](https://doi.org/10.1021/ja909133f) (see pp. 39, 82, 91)
- [72] S. E. HILL, J. ROBINSON, G. MATTHEWS, and M. MUSCHOL. Amyloid Protofibrils of Lysozyme Nucleate and Grow Via Oligomer Fusion. *Biophysical Journal*, **96**: 3781–3790, 2009. DOI: [10.1016/j.bpj.2009.01.044](https://doi.org/10.1016/j.bpj.2009.01.044) (see pp. 39, 82, 102)
- [73] J. CAMPONOV, C. HADAD, J. RUIZ, E. CLOUTET, S. GATARD, J. MUZART, S. BOUQUILLON, and D. ASTRUC. “Click” Glycodendrimers Containing 27, 81, and 243 Modified Xylopyranoside Termini. *The Journal of Organic Chemistry*, **74**: 5071–5074, 2009. DOI: [10.1021/jo900554b](https://doi.org/10.1021/jo900554b) (see pp. 39, 82)
- [74] M. R. ELKINS, T. WANG, M. NICK, H. JO, T. LEMMIN, S. B. PRUSINER, W. F. DEGRADO, J. STÖHR, and M. HONG. Structural Polymorphism of Alzheimer’s β -Amyloid Fibrils as Controlled by an E22 Switch: A Solid-State NMR Study. *Journal of the American Chemical Society*, **138**: 9840–9852, 2016. DOI: [10.1021/jacs.6b03715](https://doi.org/10.1021/jacs.6b03715) (see p. 39)
- [75] M. R. O’CONNELL, R. GAMSJAEGER, and J. P. MACKAY. The structural analysis of protein–protein interactions by NMR spectroscopy. *Proteomics*, **9**: 5224–5232, 2009. DOI: [10.1002/pmic.200900303](https://doi.org/10.1002/pmic.200900303) (see p. 39)

- [76] J. L. MARKLEY, R. BRÜSCHWEILER, A. S. EDISON, H. R. EGHBALNIA, R. POWERS, D. RAFTERY, and D. S. WISHART. The future of NMR-based metabolomics. *Current Opinion in Biotechnology*, **43**: 34–40, 2017. DOI: [10.1016/j.copbio.2016.08.001](https://doi.org/10.1016/j.copbio.2016.08.001) (see p. 39)
- [77] B. AL-ANI et al. Changes in urinary metabolomic profile during relapsing renal vasculitis. *Scientific Reports*, **6**: 38074, 2016. DOI: [10.1038/srep38074](https://doi.org/10.1038/srep38074) (see p. 39)
- [78] J. KEELER. *Understanding NMR Spectroscopy*. 2nd ed. Chichester: John Wiley & Sons, 2010. (see p. 40)
- [79] J. CAVANAGH, W. J. FAIRBROTHER, A. G. III PALMER, M. RANCE, and N. J. SKELTON. *Protein NMR Spectroscopy: Principles and Practice*. 2nd Amsterdam: Elsevier, 2007. DOI: [10.1016/B978-0-12-164491-8.X5000-3](https://doi.org/10.1016/B978-0-12-164491-8.X5000-3) (see pp. 40, 41, 115)
- [80] G. R. FULMER, A. J. M. MILLER, N. H. SHERDEN, H. E. GOTTLIEB, A. NUDELMAN, B. M. STOLTZ, J. E. BERCAW, and K. I. GOLDBERG. NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist. *Organometallics*, **29**: 2176–2179, 2010. DOI: [10.1021/om100106e](https://doi.org/10.1021/om100106e) (see pp. 40, 63)
- [81] H. E. GOTTLIEB, V. KOTLYAR, and A. NUDELMAN. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *The Journal of Organic Chemistry*, **62**: 7512–7515, 1997. DOI: [10.1021/jo971176v](https://doi.org/10.1021/jo971176v) (see p. 40)
- [82] A. SHIMIZU, M. IKEGUCHI, and S. SUGAI. Appropriateness of DSS and TSP as internal references for ¹H NMR studies of molten globule proteins in aqueous media. *Journal of Biomolecular NMR*, **4**: 859–862, 1994. DOI: [10.1007/BF00398414](https://doi.org/10.1007/BF00398414) (see pp. 41, 60)
- [83] A. DE MARCO. pH dependence of internal references. *Journal of Magnetic Resonance (1969)*, **26**: 527–528, 1977. DOI: [10.1016/0022-2364\(77\)90104-4](https://doi.org/10.1016/0022-2364(77)90104-4) (see pp. 41, 60)
- [84] J. PETTERSSON-KASTBERG, A. K. MOSSBERG, M. TRULSSON, Y. J. YONG, S. Y. MIN, Y. H. LIM, J. E. O'BRIEN, C. SVANBORG, and K. H. MOK. α -Lactalbumin, Engineered to be Nonnative and Inactive, Kills Tumor Cells when in Complex with Oleic Acid: A New Biological Function Resulting from Partial Unfolding. *Journal of Molecular Biology*, **394**: 994–1010, 2009. DOI: [10.1016/j.jmb.2009.09.026](https://doi.org/10.1016/j.jmb.2009.09.026) (see pp. 41, 60)
- [85] A. BAX. Two-Dimensional NMR and Protein Structure. *Annual Review of Biochemistry*, **58**: 223–256, 1989. DOI: [10.1146/annurev.bi.58.070189.001255](https://doi.org/10.1146/annurev.bi.58.070189.001255) (see p. 41)
- [86] D. S. WEINSTOCK, C. NARAYANAN, J. BAUM, and R. M. LEVY. Correlation between ¹³C α chemical shifts and helix content of peptide ensembles. *Protein Science*, **17**: 950–954, 2008. DOI: [10.1110/ps.073365408](https://doi.org/10.1110/ps.073365408) (see p. 41)
- [87] Y. WANG. Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Science*, **11**: 852–861, 2002. DOI: [10.1110/ps.3180102](https://doi.org/10.1110/ps.3180102) (see p. 41)

- [88] G. CORNILESCU, F. DELAGLIO, and A. BAX. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *Journal of Biomolecular NMR*, **13**: 289–302, 1999. DOI: [10.1023/a:1008392405740](https://doi.org/10.1023/a:1008392405740) (see p. 41)
- [89] D.S. WISHART, B.D. SYKES, and F.M. RICHARDS. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *Journal of Molecular Biology*, **222**: 311–333, 1991. DOI: [10.1016/0022-2836\(91\)90214-Q](https://doi.org/10.1016/0022-2836(91)90214-Q) (see p. 41)
- [90] D. S. WISHART, B. D. SYKES, and F. M. RICHARDS. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry*, **31**: 1647–1651, 1992. DOI: [10.1021/bi00121a010](https://doi.org/10.1021/bi00121a010) (see p. 41)
- [91] T. D.W. CLARIDGE. Diffusion NMR Spectroscopy in: *High-Resolution NMR Techniques in Organic Chemistry*. 3rd ed. chap. 10, 381–419, Oxford: Elsevier, 2016. DOI: [10.1016/B978-0-08-099986-9.00010-5](https://doi.org/10.1016/B978-0-08-099986-9.00010-5) (see pp. 42, 44, 45, 116)
- [92] Y. COHEN, L. AVRAM, and L. FRISH. Diffusion NMR Spectroscopy in Supramolecular and Combinatorial Chemistry: An Old Parameter–New Insights. *Angewandte Chemie International Edition*, **44**: 520–554, 2005. DOI: [10.1002/anie.200300637](https://doi.org/10.1002/anie.200300637) (see pp. 42, 75)
- [93] D. WU, A. CHEN, and C. S. JOHNSON, JR. Heteronuclear-Detected Diffusion-Ordered NMR Spectroscopy through Coherence Transfer. *Journal of Magnetic Resonance, Series A*, **123**: 215–218, 1996. DOI: [10.1006/jmra.1996.0239](https://doi.org/10.1006/jmra.1996.0239) (see p. 43)
- [94] K. F. MORRIS and C. S. JOHNSON. Diffusion-ordered two-dimensional nuclear magnetic resonance spectroscopy. *Journal of the American Chemical Society*, **114**: 3139–3141, 1992. DOI: [10.1021/ja00034a071](https://doi.org/10.1021/ja00034a071) (see pp. 42, 43, 75, 77, 91)
- [95] K. F. MORRIS and C. S. JOHNSON. Resolution of discrete and continuous molecular size distributions by means of diffusion-ordered 2D NMR spectroscopy. *Journal of the American Chemical Society*, **115**: 4291–4299, 1993. DOI: [10.1021/ja00063a053](https://doi.org/10.1021/ja00063a053) (see pp. 42, 43, 75, 77, 91)
- [96] C. S. JOHNSON. Diffusion ordered nuclear magnetic resonance spectroscopy: principles and applications. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **34**: 203–256, 1999. DOI: [10.1016/S0079-6565\(99\)00003-5](https://doi.org/10.1016/S0079-6565(99)00003-5) (see p. 43)
- [97] W. S. PRICE, F. TSUCHIYA, and Y. ARATA. Lysozyme Aggregation and Solution Properties Studied Using PGSE NMR Diffusion Measurements. *Journal of the American Chemical Society*, **121**: 11503–11512, 1999. DOI: [10.1021/ja992265n](https://doi.org/10.1021/ja992265n) (see p. 44)
- [98] D. K. WILKINS, S. B. GRIMSHAW, V. RECEVEUR, C. M. DOBSON, J. A. JONES, and L. J. SMITH. Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques. *Biochemistry*, **38**: 16424–16431, 1999. DOI: [10.1021/bi991765q](https://doi.org/10.1021/bi991765q) (see pp. 45, 47, 63, 85, 102)

-
- [99] R. PARKER, T. R. NOEL, G. J. BROWNSEY, K. LAOS, and S. G. RING. The Nonequilibrium Phase and Glass Transition Behavior of β -Lactoglobulin. *Biophysical Journal*, **89**: 1227–1236, 2005. DOI: [10.1529/biophysj.105.064246](https://doi.org/10.1529/biophysj.105.064246) (see pp. 48, 95)
- [100] F. W. J. TEALE. Cleavage of the haem-protein link by acid methylethylketone. *Biochimica et Biophysica Acta*, **35**: 543, 1959. DOI: [10.1016/0006-3002\(59\)90407-X](https://doi.org/10.1016/0006-3002(59)90407-X) (see p. 49)
- [101] T. P. J. KNOWLES, T. W. OPPENHEIM, A. K. BUELL, D. Y. CHIRGADZE, and M. E. WELLAND. Nanostructured films from hierarchical self-assembly of amyloidogenic proteins. *Nature Nanotechnology*, **5**: 204–207, 2010. DOI: [10.1038/nnano.2010.26](https://doi.org/10.1038/nnano.2010.26) (see pp. 52, 79)
- [102] M. FÄNDRICH, V. FORGE, K. BUDER, M. KITTLER, C. M. DOBSON, and S. DIEKMANN. Myoglobin forms amyloid fibrils by association of unfolded polypeptide segments. *Proceedings of the National Academy of Sciences*, **100**: 15463–15468, 2003. DOI: [10.1073/pnas.0303758100](https://doi.org/10.1073/pnas.0303758100) (see p. 53)
- [103] A. J. HOWIE, D. B. BREWER, D. HOWELL, and A. P. JONES. Physical basis of colors seen in Congo red-stained amyloid in polarized light. *Laboratory Investigation*, **88**: 232–242, 2008. DOI: [10.1038/labinvest.3700714](https://doi.org/10.1038/labinvest.3700714) (see pp. 55, 74, 75, 79)
- [104] R. KHURANA, V. N. UVERSKY, L. NIELSEN, and A. L. FINK. Is Congo Red an Amyloid-specific Dye? *Journal of Biological Chemistry*, **276**: 22715–22721, 2001. DOI: [10.1074/jbc.M011499200](https://doi.org/10.1074/jbc.M011499200) (see pp. 55, 74, 75)
- [105] L. WHITMORE and B. A. WALLACE. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research*, **32**: W668–W673, 2004. DOI: [10.1093/nar/gkh371](https://doi.org/10.1093/nar/gkh371) (see p. 57)
- [106] L. WHITMORE and B. A. WALLACE. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers*, **89**: 392–400, 2008. DOI: [10.1002/bip.20853](https://doi.org/10.1002/bip.20853) (see p. 57)
- [107] A. CHEN, D. WU, and C. S. JR. JOHNSON. Determination of the Binding Isotherm and Size of the Bovine Serum Albumin-Sodium Dodecyl Sulfate Complex by Diffusion-Ordered 2D NMR. *The Journal of Physical Chemistry*, **99**: 828–834, 1995. DOI: [10.1021/j100002a054](https://doi.org/10.1021/j100002a054) (see p. 62)
- [108] W. F. VRANKEN, W. BOUCHER, T. J. STEVENS, R. H. FOGH, A. PAJON, M. LLINAS, E. L. ULRICH, J. L. MARKLEY, J. IONIDES, and E. D. LAUE. The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins: Structure, Function, and Bioinformatics*, **59**: 687–696, 2005. DOI: [10.1002/prot.20449](https://doi.org/10.1002/prot.20449) (see pp. 65, 141)
- [109] D. S. WISHART, C. G. BIGAM, A. HOLM, R. S. HODGES, and B. D. SYKES. ^1H , ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *Journal of Biomolecular NMR*, **5**: 67–81, 1995. DOI: [10.1007/BF00227471](https://doi.org/10.1007/BF00227471) (see p. 65)

-
- [110] M. BERJANSKII et al. GeNMR: a web server for rapid NMR-based protein structure determination. *Nucleic Acids Research*, **37**: W670–W677, 2009. DOI: [10.1093/nar/gkp280](https://doi.org/10.1093/nar/gkp280) (see pp. 65, 66, 142)
- [111] M. RECHES and E. GAZIT. Controlled patterning of aligned self-assembled peptide nanotubes. *Nature Nanotechnology*, **1**: 195–200, 2006. DOI: [10.1038/nnano.2006.139](https://doi.org/10.1038/nnano.2006.139) (see p. 70)
- [112] G. BINNIG, C. F. QUATE, and Ch. GERBER. Atomic Force Microscope. *Physical Review Letters*, **56**: 930–933, 1986. DOI: [10.1103/PhysRevLett.56.930](https://doi.org/10.1103/PhysRevLett.56.930) (see p. 70)
- [113] F. S. RUGGERI, J. CHARMET, T. KARTANAS, Q. PETER, S. CHIA, J. HABCHI, C. M. DOBSON, M. VENDRUSCOLO, and T. P. J. KNOWLES. Microfluidic deposition for resolving single-molecule protein architecture and heterogeneity. *Nature Communications*, **9**: 3890, 2018. DOI: [10.1038/s41467-018-06345-4](https://doi.org/10.1038/s41467-018-06345-4) (see p. 70)
- [114] J. ADAMCIK and R. MEZZENGA. Study of amyloid fibrils via atomic force microscopy. *Current Opinion in Colloid & Interface Science*, **17**: 369–376, 2012. DOI: [10.1016/j.cocis.2012.08.001](https://doi.org/10.1016/j.cocis.2012.08.001) (see pp. 70, 71)
- [115] B. MOORES, E. DROLLE, S. J. ATTWOOD, J. SIMONS, and Z. LEONENKO. Effect of Surfaces on Amyloid Fibril Formation. *PLOS ONE*, **6**: e25954, 2011. DOI: [10.1371/journal.pone.0025954](https://doi.org/10.1371/journal.pone.0025954) (see p. 70)
- [116] H. PUCHTLER, F. SWEAT, and M. LEVINE. On the Binding of Congo Red by Amyloid. *Journal of Histochemistry & Cytochemistry*, **10**: 355–364, 1962. DOI: [10.1177/10.3.355](https://doi.org/10.1177/10.3.355) (see pp. 74, 75, 79)
- [117] E. FARREL and J. L. BROUSSEAU. *Guide for DLS sample preparation*. Brookhaven Instruments, a Nova Instruments company, 2014. (see p. 85)
- [118] L. VAN LOKEREN, G. MAHEUT, F. RIBOT, V. ESCAX, I. VERBRUGGEN, C. SANCHEZ, J. C. MARTINS, M. BIESEMANS, and R. WILLEM. Characterization of Titanium Dioxide Nanoparticles Dispersed in Organic Ligand Solutions by Using a Diffusion-Ordered Spectroscopy-Based Strategy. *Chemistry - A European Journal*, **13**: 6957–6966, 2007. DOI: [10.1002/chem.200601722](https://doi.org/10.1002/chem.200601722) (see p. 91)
- [119] P. AYMARD, D. DURAND, and T. NICOLAI. The effect of temperature and ionic strength on the dimerisation of β -lactoglobulin. *International Journal of Biological Macromolecules*, **19**: 213–221, 1996. DOI: [10.1016/0141-8130\(96\)01130-0](https://doi.org/10.1016/0141-8130(96)01130-0) (see p. 93)
- [120] S. KHAN, R. IPSEN, K. ALMDAL, B. SVENSSON, and P. HARRIS. Revealing the Dimeric Crystal and Solution Structure of β -Lactoglobulin at pH 4 and Its pH and Salt Dependent Monomer–Dimer Equilibrium. *Biomacromolecules*, **19**: 2905–2912, 2018. DOI: [10.1021/acs.biomac.8b00471](https://doi.org/10.1021/acs.biomac.8b00471) (see pp. 95, 108)

- [121] M. VERHEUL, J. S. PEDERSEN, S. P. F. M. ROEFS, and K. G. de KRUIF. Association behavior of native β -lactoglobulin. *Biopolymers*, **49**: 11–20, 1999. DOI: [10.1002/\(SICI\)1097-0282\(199901\)49:1<11::AID-BIP2>3.0.CO;2-1](https://doi.org/10.1002/(SICI)1097-0282(199901)49:1<11::AID-BIP2>3.0.CO;2-1) (see pp. 95, 108)
- [122] A. L. JANSMA, J. P. KIRKPATRICK, A. R. HSU, T. M. HANDEL, and D. NIETLISPACH. NMR Analysis of the Structure, Dynamics, and Unique Oligomerization Properties of the Chemokine CCL27. *Journal of Biological Chemistry*, **285**: 14424–14437, 2010. DOI: [10.1074/jbc.M109.091108](https://doi.org/10.1074/jbc.M109.091108) (see p. 108)
- [123] A. DHULESIA, N. CREMADES, J. R. KUMITA, S. T. D. HSU, M. F. MOSSUTO, M. DUMOULIN, D. NIETLISPACH, M. AKKE, X. SALVATELLA, and C. M. DOBSON. Local Cooperativity in an Amyloidogenic State of Human Lysozyme Observed at Atomic Resolution. *Journal of the American Chemical Society*, **132**: 15580–15588, 2010. DOI: [10.1021/ja103524m](https://doi.org/10.1021/ja103524m) (see pp. 112, 115)
- [124] C. BHATTACHARJEE, S. SAHA, A. BISWAS, M. KUNDU, L. GHOSH, and K. P. DAS. Structural Changes of β -Lactoglobulin during Thermal Unfolding and Refolding—An FT-IR and Circular Dichroism Study. *The Protein Journal*, **24**: 27–35, 2005. DOI: [10.1007/s10930-004-0603-z](https://doi.org/10.1007/s10930-004-0603-z) (see pp. 112, 125)
- [125] D. M. ROY, D. MARK, and H. H. PETACH. The denaturation and degradation of stable enzymes at high temperatures. *Biochemical Journal*, **317**: 1–11, 1996. DOI: [10.1042/bj3170001](https://doi.org/10.1042/bj3170001) (see p. 112)
- [126] O. HERNELL. Human Milk vs. Cow's Milk and the Evolution of Infant Formulas in: *Milk and Milk Products in Human Nutrition*. ed. by R. CLEMENS, O. HERNELL, and K. MICHAELSEN. 17–28, Basel: Nestec Ltd., Vevey/S. Karger AG, 2011. DOI: [10.1159/000325572](https://doi.org/10.1159/000325572) (see p. 113)
- [127] S. M. LOVEDAY. β -Lactoglobulin heat denaturation: A critical assessment of kinetic modelling. *International Dairy Journal*, **52**: 92–100, 2016. DOI: [10.1016/j.idairyj.2015.08.001](https://doi.org/10.1016/j.idairyj.2015.08.001) (see pp. 113, 124)
- [128] T. CROGUENNEC, S. BOUHALLAB, D. MOLLÉ, B. T. O'KENNEDY, and R. MEHRA. Stable monomeric intermediate with exposed Cys-119 is formed during heat denaturation of β -lactoglobulin. *Biochemical and Biophysical Research Communications*, **301**: 465–471, 2003. DOI: [10.1016/S0006-291X\(02\)02997-2](https://doi.org/10.1016/S0006-291X(02)02997-2) (see pp. 113, 114)
- [129] J. A. SEO, A. HÉDOUX, Y. GUINET, L. PACCOU, F. AFFOUARD, A. LERBRET, and M. DESCAMPS. Thermal Denaturation of Beta-Lactoglobulin and Stabilization Mechanism by Trehalose Analyzed from Raman Spectroscopy Investigations. *The Journal of Physical Chemistry B*, **114**: 6675–6684, 2010. DOI: [10.1021/jp1006022](https://doi.org/10.1021/jp1006022) (see p. 114)
- [130] S. GUPTA and S. BHATTACHARJYA. NMR Characterization of the Near Native and Unfolded States of the PTB Domain of Dok1: Alternate Conformations and Residual Clusters. *PLOS ONE*, **9**: e90557, 2014. DOI: [10.1371/journal.pone.0090557](https://doi.org/10.1371/journal.pone.0090557) (see p. 115)

-
- [131] D. J. OLDFIELD, H. SINGH, and M. W. TAYLOR. Kinetics of heat-induced whey protein denaturation and aggregation in skim milks with adjusted whey protein concentration. *Journal of Dairy Research*, **72**: 369–378, 2005. DOI: [10.1017/S002202990500107X](https://doi.org/10.1017/S002202990500107X) (see p. 118)
- [132] F. FOGOLARI, L. RAGONA, L. ZETTA, S. ROMAGNOLI, K. G. DE KRUIF, and H. MOLINARI. Monomeric bovine β -lactoglobulin adopts a β -barrel fold at pH 2. *FEBS Letters*, **436**: 149–154, 1998. DOI: [10.1016/S0014-5793\(98\)00936-3](https://doi.org/10.1016/S0014-5793(98)00936-3) (see p. 124)
- [133] X. L. QI, C. HOLT, D. McNULTY, D. T. CLARKE, S. BROWNLOW, and G. R. JONES. Effect of temperature on the secondary structure of β -lactoglobulin at pH 6.7, as determined by CD and IR spectroscopy: a test of the molten globule hypothesis. *Biochemical Journal*, **324**: 341–346, 1997. DOI: [10.1042/bj3240341](https://doi.org/10.1042/bj3240341) (see p. 125)
- [134] H. L. CASAL, U. KÖHLER, and H. H. MANTSCH. Structural and conformational changes of β -lactoglobulin B: an infrared spectroscopic study of the effect of pH and temperature. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, **957**: 11–20, 1988. DOI: [10.1016/0167-4838\(88\)90152-5](https://doi.org/10.1016/0167-4838(88)90152-5) (see p. 125)
- [135] F. A. FERRONE, J. HOFRICHTER, and W. A. EATON. Kinetics of sickle hemoglobin polymerization: II. A double nucleation mechanism. *Journal of Molecular Biology*, **183**: 611–631, 1985. DOI: [10.1016/0022-2836\(85\)90175-5](https://doi.org/10.1016/0022-2836(85)90175-5) (see p. 125)
- [136] S. AKIRA, S. UEMATSU, and O. TAKEUCHI. Pathogen Recognition and Innate Immunity. *Cell*, **124**: 783–801, 2006. DOI: [10.1016/j.cell.2006.02.015](https://doi.org/10.1016/j.cell.2006.02.015) (see p. 128)
- [137] T. KAWAI and S. AKIRA. Toll-like Receptors and Their Crosstalk with Other Innate Receptors in Infection and Immunity. *Immunity*, **34**: 637–650, 2011. DOI: [10.1016/j.immuni.2011.05.006](https://doi.org/10.1016/j.immuni.2011.05.006) (see p. 128)
- [138] R. MEDZHITOV, P. PRESTON-HURLBURT, E. KOPP, A. STADLEN, C. CHEN, S. GHOSH, and C. A. JANEWAY. MyD88 Is an Adaptor Protein in the hToll/IL-1 Receptor Family Signaling Pathways. *Molecular Cell*, **2**: 253–258, 1998. DOI: [10.1016/S1097-2765\(00\)80136-7](https://doi.org/10.1016/S1097-2765(00)80136-7) (see p. 128)
- [139] M. YAMAMOTO, S. SATO, K. MORI, K. HOSHINO, O. TAKEUCHI, K. TAKEDA, and S. AKIRA. Cutting Edge: A Novel Toll/IL-1 Receptor Domain-Containing Adapter That Preferentially Activates the IFN- Promoter in the Toll-Like Receptor Signaling. *The Journal of Immunology*, **169**: 6668–6672, 2002. DOI: [10.4049/jimmunol.169.12.6668](https://doi.org/10.4049/jimmunol.169.12.6668) (see p. 128)
- [140] J. A. OWEN, J. PUNT, S. A. STRANFORD, P. P. JONES, and J. KUBY. *Kuby Immunology*. 7th ed. New York, NY: W.H. Freeman, 2013. (see p. 128)
- [141] J. C. KAGAN, T. SU, T. HORNG, A. CHOW, S. AKIRA, and R. MEDZHITOV. TRAM couples endocytosis of Toll-like receptor 4 to the induction of interferon- β . *Nature Immunology*, **9**: 361–368, 2008. DOI: [10.1038/ni1569](https://doi.org/10.1038/ni1569) (see p. 128)

- [142] A. G. BOWIE and L. UNTERHOLZNER. Viral evasion and subversion of pattern-recognition receptor signalling. *Nature Reviews Immunology*, **8**: 911–922, 2008. DOI: [10.1038/nri2436](https://doi.org/10.1038/nri2436) (see p. 128)
- [143] A. BOWIE, E. KISS-TOTH, J. A. SYMONS, G. L. SMITH, S. K. DOWER, and L. A. J. O’NEILL. A46R and A52R from vaccinia virus are antagonists of host IL-1 and toll-like receptor signaling. *Proceedings of the National Academy of Sciences*, **97**: 10162–10167, 2000. DOI: [10.1073/pnas.160027697](https://doi.org/10.1073/pnas.160027697) (see p. 129)
- [144] I. R. HAGA and A. G. BOWIE. Evasion of innate immunity by vaccinia virus. *Parasitology*, **130**: S11, 2005. DOI: [10.1017/S0031182005008127](https://doi.org/10.1017/S0031182005008127) (see p. 129)
- [145] G. L. SMITH, C. T. O. BENFIELD, C. MALUQUER DE MOTES, M. MAZZON, S. W. J. EMBER, B. J. FERGUSON, and R. P. SUMNER. Vaccinia virus immune evasion: mechanisms, virulence and immunogenicity. *Journal of General Virology*, **94**: 2367–2392, 2013. DOI: [10.1099/vir.0.055921-0](https://doi.org/10.1099/vir.0.055921-0) (see p. 129)
- [146] J. STACK, I. R. HAGA, M. SCHRÖDER, N. W. BARTLETT, G. MALONEY, P. C. READING, K. A. FITZGERALD, G. L. SMITH, and A. G. BOWIE. Vaccinia virus protein A46R targets multiple Toll-like–interleukin-1 receptor adaptors and contributes to virulence. *The Journal of Experimental Medicine*, **201**: 1007–1018, 2005. DOI: [10.1084/jem.20041442](https://doi.org/10.1084/jem.20041442) (see p. 129)
- [147] M. A. HUTCHENS, K. E. LUKER, J. SONSTEIN, G. NÚÑEZ, J. L. CURTIS, and G. D. LUKER. Protective Effect of Toll-like Receptor 4 in Pulmonary Vaccinia Infection. *PLOS Pathogens*, **4**: e1000153, 2008. DOI: [10.1371/journal.ppat.1000153](https://doi.org/10.1371/journal.ppat.1000153) (see p. 129)
- [148] J. STACK and A. G. BOWIE. Poxviral Protein A46 Antagonizes Toll-like Receptor 4 Signaling by Targeting BB Loop Motifs in Toll-IL-1 Receptor Adaptor Proteins to Disrupt Receptor:Adaptor Interactions. *Journal of Biological Chemistry*, **287**: 22672–22682, 2012. DOI: [10.1074/jbc.M112.349225](https://doi.org/10.1074/jbc.M112.349225) (see pp. 129, 141)
- [149] T. LYSAKOVA-DEVINE, B. KEOGH, B. HARRINGTON, K. NAGPAL, A. HALLE, D. T. GOLENBOCK, T. MONIE, and A. G. BOWIE. Viral Inhibitory Peptide of TLR4, a Peptide Derived from Vaccinia Protein A46, Specifically Inhibits TLR4 by Directly Targeting MyD88 Adaptor-Like and TRIF-Related Adaptor Molecule. *The Journal of Immunology*, **185**: 4261–4271, 2010. DOI: [10.4049/jimmunol.1002013](https://doi.org/10.4049/jimmunol.1002013) (see pp. 129, 141)
- [150] W. PIAO, L. W. RU, K. H. PIEPENBRINK, E. J. SUNDBERG, S. N. VOGEL, and V. Y. TOSHCHAKOV. Recruitment of TLR adapter TRIF to TLR4 signaling complex is mediated by the second helical region of TRIF TIR domain. *Proceedings of the National Academy of Sciences*, **110**: 19036–19041, 2013. DOI: [10.1073/pnas.1313575110](https://doi.org/10.1073/pnas.1313575110) (see p. 129)
- [151] W. PIAO, S. N. VOGEL, and V. Y. TOSHCHAKOV. Inhibition of TLR4 Signaling by TRAM-Derived Decoy Peptides In Vitro and In Vivo. *The Journal of Immunology*, **190**: 2263–2272, 2013. DOI: [10.4049/jimmunol.1202703](https://doi.org/10.4049/jimmunol.1202703) (see p. 129)

-
- [152] V. Y. TOSHCHAKOV, M. J. FENTON, and S. N. VOGEL. Cutting Edge: Differential Inhibition of TLR Signaling Pathways by Cell-Permeable Peptides Representing BB Loops of TLRs. *The Journal of Immunology*, **178**: 2655–2660, 2007. DOI: [10.4049/jimmunol.178.5.2655](https://doi.org/10.4049/jimmunol.178.5.2655) (see p. 129)
- [153] A. SAVVA and T. ROGER. Targeting Toll-Like Receptors: Promising Therapeutic Strategies for the Management of Sepsis-Associated Pathology and Infectious Diseases. *Frontiers in Immunology*, **4**: 387, 2013. DOI: [10.3389/fimmu.2013.00387](https://doi.org/10.3389/fimmu.2013.00387) (see p. 130)
- [154] J. XUE and A. HABTEZION. Carbon monoxide-based therapy ameliorates acute pancreatitis via TLR4 inhibition. *Journal of Clinical Investigation*, **124**: 437–447, 2014. DOI: [10.1172/JCI71362](https://doi.org/10.1172/JCI71362) (see p. 130)
- [155] X. ZHANG, C. ZHU, D. WU, and X. JIANG. Possible Role of Toll-Like Receptor 4 in Acute Pancreatitis. *Pancreas*, **39**: 819–824, 2010. DOI: [10.1097/MPA.0b013e3181ca065c](https://doi.org/10.1097/MPA.0b013e3181ca065c) (see p. 130)
- [156] N. SHAH, M. MONTES DE OCA, M. JOVER-COBOS, K. TANAMOTO, M. MUROI, K. SUGIYAMA, N. A. DAVIES, R. P. MOOKERJEE, D. K. DHAR, and R. JALAN. Role of Toll-Like Receptor 4 in Mediating Multiorgan Dysfunction in Mice With Acetaminophen Induced Acute Liver Failure. *Liver Transplantation*, **19**: 751–761, 2013. DOI: [10.1002/lt.23655](https://doi.org/10.1002/lt.23655) (see p. 130)
- [157] T. DZIODZIO, M. BIEBL, and J. PRATSCHKE. Impact of brain death on ischemia/reperfusion injury in liver transplantation. *Current Opinion in Organ Transplantation*, **19**: 108–114, 2014. DOI: [10.1097/MOT.000000000000061](https://doi.org/10.1097/MOT.000000000000061) (see p. 131)
- [158] J. W. LEE, S. C. KIM, Y. S. KO, H. Y. LEE, E. Y. CHO, M. Y. KIM, S. K. JO, W. Y. CHO, and H. K. KIM. Renoprotective effect of paricalcitol via a modulation of the TLR4-NF- κ B pathway in ischemia/reperfusion-induced acute kidney injury. *Biochemical and Biophysical Research Communications*, **444**: 121–127, 2014. DOI: [10.1016/j.bbrc.2014.01.005](https://doi.org/10.1016/j.bbrc.2014.01.005) (see p. 131)
- [159] H. MEIMEI, M. DEJIN, C. ERZHEN, S. MINMIN, J. SONGYAO, L. JIANFANG, and C. HAO. Inhibiting the Toll-Like Receptor 4 Toll/Interleukin-1 Receptor Domain Protects Against Hepatic Warm Ischemia and Reperfusion Injury in Mice*. *Critical Care Medicine*, **42**: e123–e131, 2014. DOI: [10.1097/CCM.0b013e3182a647c1](https://doi.org/10.1097/CCM.0b013e3182a647c1) (see p. 131)
- [160] S. SAVIC et al. TLR dependent XBP-1 activation induces an autocrine loop in rheumatoid arthritis synoviocytes. *Journal of Autoimmunity*, **50**: 59–66, 2014. DOI: [10.1016/j.jaut.2013.11.002](https://doi.org/10.1016/j.jaut.2013.11.002) (see p. 131)
- [161] G. S. MASSON, A. R. NAIR, R. B. DANGE, P. P. SILVA-SOARES, L. C. MICHELINI, and J. FRANCIS. Toll-Like Receptor 4 Promotes Autonomic Dysfunction, Inflammation and Microglia Activation in the Hypothalamic Paraventricular Nucleus: Role of Endoplasmic Reticulum Stress. *PLOS ONE*, **10**: e0122850, 2015. DOI: [10.1371/journal.pone.0122850](https://doi.org/10.1371/journal.pone.0122850) (see p. 131)

- [162] K. A. SHIREY et al. Novel strategies for targeting innate immune responses to influenza. *Mucosal Immunology*, **9**: 1173–1182, 2016. DOI: [10.1038/mi.2015.141](https://doi.org/10.1038/mi.2015.141) (see p. [131](#))
- [163] S. LEOW-DYKE, C. ALLEN, A. DENES, O. NILSSON, S. MAYSAMI, A. G. BOWIE, N. J. ROTHWELL, and E. PINTEAUX. Neuronal toll-like receptor 4 signaling induces brain endothelial activation and neutrophil transmigration in vitro. *Journal of Neuroinflammation*, **9**: 698, 2012. DOI: [10.1186/1742-2094-9-230](https://doi.org/10.1186/1742-2094-9-230) (see p. [131](#))
- [164] R. B. DANGE, D. AGARWAL, R. TERUYAMA, and J. FRANCIS. Toll-like receptor 4 inhibition within the paraventricular nucleus attenuates blood pressure and inflammatory response in a genetic model of hypertension. *Journal of Neuroinflammation*, **12**: 31, 2015. DOI: [10.1186/s12974-015-0242-7](https://doi.org/10.1186/s12974-015-0242-7) (see p. [131](#))
- [165] S. BERGER and B. SIEGMAR. *200 and More NMR Experiments: A Practical Course*. Weinheim: Wiley-VCH, 2004. (see p. [134](#))
- [166] M. S. CHEUNG, M. L. MAGUIRE, T. J. STEVENS, and R. WILLIAM BROADHURST. DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *Journal of Magnetic Resonance*, **202**: 223–233, 2009. DOI: [10.1016/j.jmr.2009.11.008](https://doi.org/10.1016/j.jmr.2009.11.008) (see p. [139](#))
- [167] J. Y. KIM, D. LAWLESS, M. BARAN, M. RUETHER, A. BOWIE, and K. H. MOK. Exploring the Therapeutic Potential of a Peptide Derived from a Poxviral Immune Evasion Protein A46: With NMR Determination of the Solution Structure of VIPER and Its Inactive Mutant. *Manuscript in Preparation*, 2019. (see pp. [141](#), [146](#))
- [168] Y. W. KIM, H. LEE, L. HEO, C. SEOK, and J. W. CHOE. Structure of vaccinia virus A46, an inhibitor of TLR4 signaling pathway, shows the conformation of VIPER motif. *Protein Science*, **23**: 906–914, 2014. DOI: [10.1002/pro.2472](https://doi.org/10.1002/pro.2472) (see pp. [141](#), [145](#))
- [169] R. H. FOGH, W. BOUCHER, W. F. VRANKEN, A. PAJON, T. J. STEVENS, T. N. BHAT, J. WESTBROOK, J. M. C. IONIDES, and E. D. LAUE. A framework for scientific data modeling and automated software development. *Bioinformatics*, **21**: 1678–1684, 2005. DOI: [10.1093/bioinformatics/bti234](https://doi.org/10.1093/bioinformatics/bti234) (see p. [141](#))
- [170] M. NILGES, A. BERNARD, B. BARDIAUX, T. MALLIAVIN, M. HABECK, and W. RIEPING. Accurate NMR Structures Through Minimization of an Extended Hybrid Energy. *Structure*, **16**: 1305–1312, 2008. DOI: [10.1016/j.str.2008.07.008](https://doi.org/10.1016/j.str.2008.07.008) (see p. [141](#))
- [171] W. RIEPING, M. HABECK, B. BARDIAUX, A. BERNARD, T. E. MALLIAVIN, and M. NILGES. ARIA2: Automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics*, **23**: 381–382, 2007. DOI: [10.1093/bioinformatics/btl589](https://doi.org/10.1093/bioinformatics/btl589) (see p. [141](#))
- [172] S. I. ODA, E. FRANKLIN, and A. R. KHAN. Poxvirus A46 protein binds to TIR domain-containing Mal/TIRAP via an α -helical sub-domain. *Molecular Immunology*, **48**: 2144–2150, 2011. DOI: [10.1016/j.molimm.2011.07.014](https://doi.org/10.1016/j.molimm.2011.07.014) (see p. [141](#))