Characterisation of Protein Structure and Interactions: Novel Applications to the Study of Bioactive Peptides

by

Antonio Nickolas Calabrese, B. Sc. (Hons.)

A thesis submitted for the Degree of
Doctor of Philosophy

in the
School of Chemistry and Physics
The University of Adelaide

THE UNIVERSITY of ADELAIDE

June 2013

Contents

Abstract

The studies of protein/peptide folding, misfolding, structure, and interactions are vital to understanding complex biological problems. The work presented in this thesis describes the development and application of a variety of biophysical techniques to investigate protein structure and interactions, with applications to the structure and function of several bioactive peptides.

Firstly, the development of a novel negative ion amenable chemical crosslinking-mass spectrometry (CX-MS) approach is described. CX-MS is a low-resolution technique to study protein structure and interactions. It involves covalent modification and tethering of a protein complex by a reactive reagent, followed by proteolytic digestion. The sites of the intra- and inter-molecular crosslinks provide distance restraints for modelling and enables conclusions to be drawn about the three-dimensional structure and binding interfaces within a protein complex. However, easy identification of crosslinks amongst the large quantity of proteolytic fragments remains challenging. In this study, the application of novel disulfide-based MS cleavable crosslinking reagents was investigated as a tool to easily identify crosslinked peptides by their highly reproducible and characteristic fragmentation patterns in the negative ion mode. MS3 analysis of the product anions allows easy sequencing and identification of crosslinking sites. Preliminary investigations validate these reagent as a tools to readily identify chemical crosslinks within proteins and their complexes, demonstrating that this approach is an effective and efficient means to determine aspects of the topologies of protein complexes of biological importance.

Secondly, the use of several biophysical methods is described to probe the structures of a variety of complexes involving the regulatory protein calmodulin (CaM) with bioactive amphibian peptides. CaM is ubiquitous in nature and plays a regulatory role in numerous biological processes, including some in amphibians and their predators; for example, it is involved in the upregulation of nitric oxide synthesis *in vivo*. Isothermal titration calorimetry was used to investigate the specific heats of the interactions, ion mobility-mass spectrometry was used to investigate the changes in collision cross section that occur as a result of complexation and nuclear magnetic resonance spectroscopy was used to track chemical shift changes upon binding. The results obtained confirm that these complexes adopt canonical collapsed structures and demonstrate the strength of the interaction between the peptides and CaM.

Next, work is presented which investigated the abilities of several bioactive amphibian peptides to inhibit fibril formation by disease related proteins. The peptide caerin 1.8 and several synthetic modifications were tested for their ability to inhibit fibril formation by the Alzheimer's related amyloid-β (1-42) peptide. The results obtained show that caerin 1.8 redirects the aggregation process of amyloid-β (1-42) toward the amorphous aggregation pathway. In addition, the self-assembly properties of the antimicrobial peptide uperin 3.5 were investigated using a variety of biophysical techniques, including transmission electron microscopy, ion mobility-mass spectrometry, circular dichroism, thioflavin T binding and cell viability assays. Similarities were observed between the fibrils formed by this peptide and those of disease related proteins, supporting the notion that information can be obtained about disease related amyloid fibril formation by studying amyloidogenic host-defence peptides.

Lastly, work detailing the effect of aspartic acid (Asp) isomerisation to isoAsp on the structure, activity and proteolytic cleavage susceptibility of three amphibian peptides, *Crinia* angiotensin II, uperin 1.1 and citropin 1.1 is presented. isoAsp formation has been shown to occur naturally as a result of age-related protein degradation, and is a consideration when preparing formulations of peptide therapeutics. isoAsp formation causes a 'kink' in the normally helical structure of citropin 1.1, as determined by nuclear magnetic resonance spectroscopy, which results in a reduction of its antimicrobial activity. The effect of this isomerisation process on the smooth muscle activities of *Crinia* angiotensin II and uperin 1.1 was different, with Asp isomerisation in *Crinia* angiotensin II causing a decrease in activity, and Asp isomerisation in uperin 1.1 causing greater contraction at lower concentrations. Proteolytic cleavage with trypsin was identical for each pair of Asp/isoAsp isomers, whilst cleavage with α-chymotrypsin was different for the two Asp/isoAsp citropin 1.1 isomers due to the presence of isoAsp adjacent to the cleavage site.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Antonio Nickolas Calabrese
18^{th} of June 2013

Acknowledgements

First and foremost I must thank my two supervisors Prof. John Bowie and Dr. Tara Pukala. Thank you for your hard work and dedication in supporting me throughout my studies. I am a much better scientist for having had the privilege of working with you.

In addition, the support of a number of collaborators must be recognised. Many thanks to Dr. Ian Musgrave from the School of Medical Sciences for providing the facilities to undertake the cell culture work. Thanks also to Assoc. Prof Grant Booker and his research group from the School of Molecular and Biomedical Science for allowing me to conduct protein expression in your laboratory, and for the advice when I was beginning my protein NMR project. Also, many thanks to Prof. John Carver for helpful discussions about protein aggregation. Much appreciation also goes to Prof. Lixin Zhang and his research group from the Institute of Microbiology, Chinese Academy of Sciences, for performing the antimicrobial testing. Thanks to Assoc. Prof. Lisa Martin and her research group at the School of Chemistry, Monash University, especially Stefania Piantavigna, for their helpful discussions, assistance, and for making me feel welcome during the time I spent in Melbourne. Finally, to Dr. Rico Tabor, also from the School of Chemistry, Monash University, for his help acquiring AFM images.

I also must thank all the technical staff in the Department of Chemistry at The University of Adelaide. In particular, Phil Clements for his help with NMR spectroscopy and mass spectrometry. Thanks also to Lyn Waterhouse at Adelaide Microscopy for her help with TEM imaging.

Past and present members of the Bowie and Pukala research groups must also be thanked for making Labs 4 and 6 great places to work. In particular, thanks must go to Dr Tianfang (Gavin) Wang, for his help with the computational aspects of the work presented in this thesis. Also, many thanks to Dr. Michael Maclean, Dr. Patrick Sherman, Dr. Rebecca Jackway, Hayley Andreazza, Dr. Yanqin Liu and Dr. Danielle Williams for their friendship, helpful discussions, and teaching me many laboratory techniques.

Thanks also to my friends both in and out of the Chemistry Department. Especially to Courtney, Jade, Claire, Jana, Joon-Yip, Vincent and Vinda. Your support and friendship during these past few, somewhat testing, years has been invaluable.

Finally, I must thank my family for their love and support throughout my life. To Mum, Dad and Evi, thank you for your encouragement and advice. I couldn't have done this without you.

Abbreviations

1D	one dimensional
2D	two dimensional
3D	three dimensional
$\Delta\delta$	secondary shift
ΔC_p	specific heat capacity
ΔH	enthalpy change
ΔS	entropy change
Å	Angstrom
Aβ	amyloid-β
AFM	atomic force microscopy
AMP	antimicrobial peptide
apoCaM	calcium-free calmodulin
ARIA	Ambiguous Restraints for Iterative Assignment
ATD	arrival time distribution
Ca^{2+}CaM	calcium-bound calmodulin
Ca_4^{2+}CaM	calcium-saturated calmodulin
CaM	calmodulin
CD	circular dichroism
CCS	collision cross-section
Chol	cholesterol
CID	collision induced dissociation
CMC	critical micelle concentration
CX	chemical cross-linking
CX-MS	chemical cross-linking mass spectrometry
COSY	correlation spectroscopy
CSI	chemical shift index
CSP	chemical shift perturbation
Da	Dalton
DC	direct current
DMPC	1,2-dimyristoyl-*sn*-glycero-3-phosphocholine
DMPG	1,2-dimyristoyl-*sn*-glycero-3-phospho-*rac*-(1-glycerol) (sodium salt)
DNA	deoxyribonucleic acid
DPC	dodecylphosphocholine
DQF	double-quantum filtered
DSA	dithiobis(succinimidyl acetate)

DSB	dithiobis(succinimidyl butanoate)
DSP	dithiobis(succinimidyl propionate)
E. coli	*Escherichia coli*
EDTA	ethylenediamine tetraacetic acid
EGTA	ethylene glycol tetraacetic acid
EGCG	(-)-epigallocatechin-3-gallate
EHSS	exact hard sphere scattering
eNOS	endothelial nitric oxide synthase
ESI	electrospray ionisation
ESI-MS	electrospray ionisation mass spectrometry
FPLC	fast protein liquid chromatography
FID	free induction decay
HIC	hydrophobic interaction chromatography
HPLC	high performance liquid chromatography
HSQC	heteronuclear single-quantum coherence
Hz	Hertz
I	nuclear spin quantum number
IM	ion mobility
IM-MS	ion mobility mass spectrometry
iNOS	inducible nitric oxide synthase
isoAsp	isoaspartic acid
LC-MS	liquid chromatography mass spectrometry
MALDI	matrix assisted laser desorption/ionisation
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MS^n	multi-stage mass spectrometry
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
m/z	mass to charge ratio
nanoESI	nanoelectrospray ionisation
NHS	N-hydroxysuccinimide
NMR	nuclear magnetic resonance
nNOS	neuronal nitric oxide synthase
NO	nitric oxide
NOS	nitric oxide synthase
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
PA	projection approximation
PAGE	polyacrylamide gel electrophoresis
PC12	pheochromocytoma-12
PDB	Protein Data Bank

PIR	protein interaction reporter
RF	radiofrequency
RMD	restrained molecular dynamics
RMSD	root-mean-square deviation
RNA	ribonucleic acid
SA	simulated annealing
SDS	sodium dodecylsulfate
SUV	small unilamellar vesicle
TEM	transmission electron microscopy
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
ThT	thioflavin T
TOCSY	total correlation spectroscopy
ToF	time of flight
TWIG	travelling wave ion guide
UV	ultraviolet

Chapter 1

Introduction

1.1 Overview

The central dogma of molecular biology was first proposed in 1958 [1, 2], and presents a framework to comprehend the transfer of information between the three biopolymers - DNA, RNA and protein. Proteins are the final product of the gene expression process and are essential for nearly all biochemical processes in the cell [3].

Proteins are synthesised *in vivo* by the ribosome, which translates the messenger RNA code into a protein sequence [4]. After translation, most proteins are required to adopt a well-defined, compact three-dimensional structure to perform their biological function [3]. This conformation is described by the four levels of protein structure, namely primary, secondary, tertiary and quaternary (Section 1.2) [4–7]. It is imperative that the correct conformation is adopted, as misfolding can lead to loss of function and numerous, potentially dire, pathologies [8–10].

Even though proteins must adopt a pre-determined conformation to perform their biological function, most of the tasks they perform also rely on molecular interactions [11]. This necessitates some inherent flexibility in the protein structure, as interactions often lead to a local or global reorganisation of conformation [12]. Consequently, understanding the structural basis of molecular interactions and their biological role necessitates a knowledge of the protein's conformation prior to and during complexation.

This thesis details the development and application of biophysical techniques to investigate the structure of native proteins, protein assemblies and also protein misfolding at the molecular level.

1.2 Protein and Peptide Structure

Proteins are all made from amino acids, with each comprising a specific polymeric sequence of amino acid residues [4–7]. The primary structure (Figure 1.1A) is this sequence of amino acids, and is governed by the DNA sequence of the encoding gene. The coded protein sequence contains all the information necessary for the chain to adopt subsequent levels of structure and fold into a fully functional conformation [13]. This is not always a spontaneous process, and in some instances molecular chaperones are required to facilitate the adoption of a native state (Section 1.3). Misfolding can occur as a result of mutations in the primary structure or pertubations in the cellular environment, which may lead to subsequent loss of function and/or a gain of toxic function (Section 1.3).

Secondary structure (Figure 1.1B) refers to the local arrangement of the protein to form regular structures, such as α-helices, β-strands, β-sheets and turns. This folding process is largely driven by the formation of extensive hydrogen bond networks, with different regions of a protein (termed protein domains) folding independently of each other. α-Helices (Figure 1.1B, left) result due to hydrogen bond formation between the backbone carbonyl oxygen of residue i and the amide hydrogen of residue $(i + 4)$, affording an α-helix with 3.6 amino acid residues per turn. Other helical structures may occur with different hydrogen bond interactions. β-Sheets form as a result of hydrogen bonding between parallel or antiparallel chains (Figure 1.1B, centre and right, respectively), with the resultant arrangement of α-carbon atoms forming a pleated pattern. Turn motifs are often used to make sharp turns in a polypeptide chain, such as the β-turn which forms as a result of hydrogen bond interactions between the backbone carbonyl oxygen of residue i and the amide hydrogen of residue $(i + 2)$. These turn structures require a *cis* arrangement about the peptide bond, instead of the more common *trans* configuration. Pro and Gly residues more easily assume *cis* arrangements about their peptide bonds so are often present in these regions of structure.

A

B

NOTE:
These figures/tables/images have been removed to comply with copyright regulations.
They are included in the print copy of the thesis held by the University of Adelaide Library.

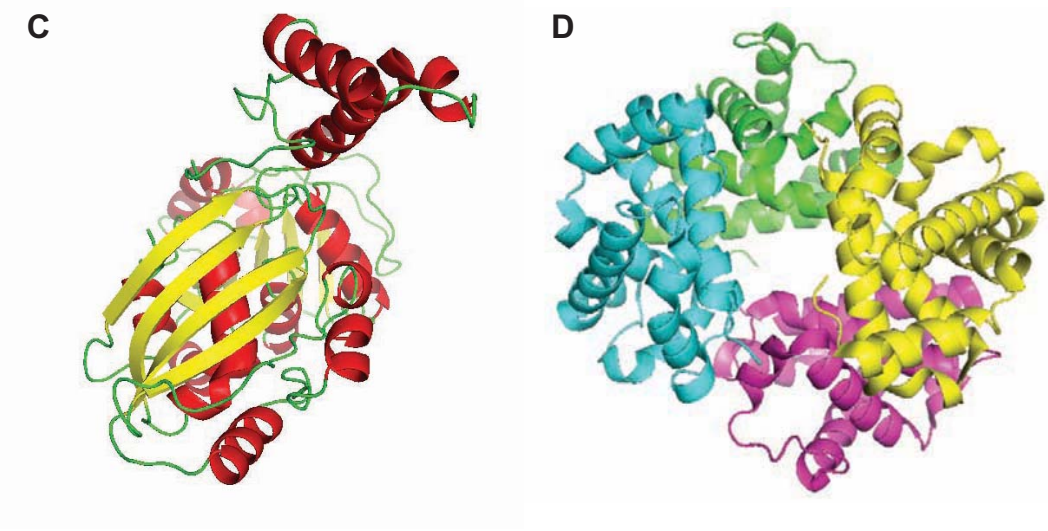


Figure 1.1: The levels of protein structure. (A) Primary structure, (B) common secondary structure elements, (C) tertiary structure (helices are shown in red, β-sheets are shown in yellow) (PDB: 1BRM), and (D) quaternary structure (in this case a four subunit complex) (PDB: 1GZX). Parts A and B are from reference [7].

Tertiary structure (Figure 1.1C) refers to the three-dimensional fold of a protein and is determined largely by non-covalent interactions between domains, but in some instances covalent disulfide-bonds are also present. Here, helices, sheets and turns are arranged in a specific fashion to form a globular structure. Often this arrangement results in the protein core containing hydrophobic amino acid residues. Some proteins may also require the association of non-protein ligands, such as metal ions (e.g. calmodulin requires Ca^{2+}) or prosthetic groups (e.g. myoglobin requires heme) to adopt a stable globular structure.

In many instances the tertiary structure is the biologically functional form of a protein, however, functionality may only be achieved in some instances by the association of multiple protein chains. This results in the formation of a protein complex, and in addition to the tertiary structure of each subunit, quaternary structure now exists (Figure 1.1D). This is the overall spatial arrangement that is adopted when multiple protein chains interact. Formation of these protein complexes is driven by non-covalent interactions between protein subunits.

It is vital that a protein folds in accordance with all these levels of protein structure, thereby adopting a biologically functional form. Understanding the basis of regulatory cellular processes, at a molecular level, necessitates a knowledge of the overall fold of a protein, and the mechanisms driving interactions with its binding partners.

1.3 Protein Folding

It has been demonstrated that the protein folding process is, for the most part, spontaneous and therefore determined by the protein's primary structure [13]. However, the mechanism by which the final folded conformation of a protein is encoded in the amino acid sequence, and how the folding process operates is still the subject of intense research. Typically, energy landscapes are used to describe the transition of a high energy unfolded structure down a funnel-like energy profile to yield a native, biologically functional state (Figure 1.2) [9, 10, 14, 15].

Figure 1.2: The energy landscape of both protein folding and aggregation. The blue surface shows that many intermediate conformations can lead to to the native state (which is in a local energy minimum) by the formation of intramolecular contacts. The purple surface shows the conformations that can lead to amorphous aggregation or amyloid fibril formation through intermolecular contacts. From reference [14].

It has been shown that most small proteins (<100 residues) assume their native structure in a one step process, where only the unfolded and native state are significantly populated [16]. Conversely, large multi-domain proteins (>100 residues) fold by a process which involves a variety of intermediate non-native states [14, 17]. The intermediate structures that form during the folding process may be on-pathway intermediates towards the native structure, or could be stable misfolded conformations that are trapped in local energy

minima (Figure 1.2). The native conformation of a protein is at a local minimum in the protein folding energy landscape [9, 10, 14, 15].

Aggregation (the association of two or more non-native protein molecules) can occur from partially folded or misfolded structures which expose hydrophobic amino acid residues that are ordinarily buried within the native structure [9, 10, 15, 18, 19]. These misfolded states can arise during folding, or as a result of cellular stress conditions. This can result in the protein leaving the folding pathway and entering the off-folding pathway, driven by hydrophobic interactions between folding intermediates (Figure 1.2 and 1.3). The off-folding pathway comprises two distinct routes and products. The amorphous aggregation pathway leads to the formation of disordered aggregates and protein precipitation. These amorphous aggregates are non-toxic and are able to be cleared by normal cellular degradation pathways i.e. the proteosome [18]. However, aggregation may also lead to the formation of thermodynamically stable amyloid [8–10, 15, 18–20]. These are highly ordered fibrillar structures which comprise a cross-β core (i.e. β-strands running perpendicular to the fibril axis) (discussed in Chapter 5). The oligomeric pre-fibrillar species formed on this off-folding pathway are considered to be highly toxic to eukaryotic cells, and are associated with the aetiology of several devastating pathologies including Alzheimer's and Parkinson's diseases.

Figure 1.3: Schematic of intracellular protein folding and aggregation.
Misfolding can occur as a result of cellular stresses, mutations or translational errors. Misfolded species can be refolded to a native conformation, be degraded, or enter the off-folding pathways to afford disordered or ordered aggregates. From reference [21].

The off-folding aggregation processes are enhanced in the crowded cellular environment. To overcome this, molecular chaperones [14, 18, 19, 22] are often present to sequester and assist in the rearrangement of a non-native structure, redirecting misfolded proteins towards their native conformations. Chaperones are involved at all stages of the folding process, and may also be required for *de novo* folding, the refolding of denatured proteins, disruption of oligomeric assemblies, protein transport and also degradation [14].

In many instances, the biological function of a protein necessitates interactions with other proteins (Section 1.4) or metabolites. Correct folding and the adoption of native structure is necessary for this to occur as it exposes specialised binding domains involved in molecular association processes.

1.4 Protein Interactions

Cellular functions are performed and regulated at the molecular level by the coordinated action of molecular complexes, which rely on the precise regulation of protein association in a spatial and temporal manner [23, 24]. In humans, disease can result when regulatory procedures break down and protein-protein and protein-metabolite interactions are perturbed [23, 24]. Additionally, cellular signalling pathways and the protein interactions they encompass, are targets for current and new therapeutics [25]. Consequently, analyses of macromolecular complexes contribute not only to the understanding of necessary cellular processes, but may lead to the development of new therapeutics and an understanding of their mode of action [23, 26].

Studies have identified a diverse array of specialised binding domains in proteins that mediate complexation [27]. However, in most cases the binding interfaces and structural motifs that confer recognition specificity are unknown [24]. Therefore, for novel protein interactions, it is vital to determine the specific structural determinants that lead to assembly.

The forces which maintain protein structures and complexes include electrostatic interactions, hydrogen bonding, hydrophobic interactions and other van der Waals forces [28]. The strength of these intermolecular forces determines the lifetime of the protein complex, with some assemblies being dynamic and transient whilst others are highly stable [11, 24, 29]. For example, signalling and regulatory networks comprise transient interactions, and complexes like the core of the ribosome are essentially permanent [11, 24, 29].

The fundamental importance of protein structure and protein assemblies in biochemistry has led to the development of numerous biophysical techniques to study protein structure and interactions (Chapter 2).

Chapter 2

Techniques for Studying Protein/Peptide Structure and Interactions

2.1 Introduction

As detailed in Sections 1.3 and 1.4, peptides and proteins are able to perform their specific biological function because of the precise three-dimensional structure they adopt *in vivo*, and the molecular interactions that they take part in. Therefore, studying the structures that these biomolecules adopt, both individually and as part of a complex can provide insight into their mechanism of action.

Numerous techniques have been developed to provide details of the structures of proteins and their complexes, including those which give atomic level resolution, such as X-ray crystallography [30] and nuclear magnetic resonance (NMR) spectroscopy [31]. However, unstable and/or very large assemblies are not amenable to study using such traditional techniques. Consequently, low resolution studies of these structures are predominantly used to provide significant insight into their structure and biological function. Techniques most useful for this purpose include circular dichroism (CD) spectroscopy [32], electron microscopy [33], small-angle X-ray scattering (SAXS) [34], molecular modelling [35] and isothermal titration calorimetry (ITC) [36]. More recently, mass spectrometry (MS) is emerging as a vital analytical tool, not only to study the primary structure of proteins and protein complexes by proteomics analyses, but also the native structures of dynamic

proteins and assemblies that are not amenable to study using conventional structural biology methods [28, 37].

This chapter details the major biophysical techniques used in this thesis to investigate protein structure and interactions, namely MS, NMR spectroscopy, ITC and CD.

2.2 Mass Spectrometry

MS relies on the formation of gas phase ions (positively and negatively charged) and their subsequent mass analysis and detection. The method has evolved rapidly in the past twenty years and today it is an indispensable analytical technique in many areas of research, especially the life and health sciences.

The first magnetic sector mass spectrometers were developed by Thomson and his student Aston in the early 1900s [38] and were utilised for the detection of isotopes. Since these early studies, MS instrumentation has dramatically evolved, leading to the development of new ionisation techniques and mass analysers to afford instrumentation with significantly enhanced resolution and capabilities.

Notably, the advent of proteomics triggered an explosion in the use of MS to study biological molecules [39]. As well as developments in MS instrumentation and methodologies, the interfacing of MS with other techniques, such as liquid chromatography-MS (LC-MS) [40, 41], ion mobility-MS (IM-MS) [42–44] as well as the arrangement of mass analysers in tandem array has allowed MS to find broad ranging applications.

Additionally, experimental evidence that protein conformation *in vacuo* is often not significantly affected by the loss of a solvation shell, has validated MS for the study of both native structure and molecular interactions. This has lead to the development of MS based structural biology methods [45, 46]. Consequently, integrative approaches, combining MS techniques with other low resolution data, are becoming commonplace in the development of structural models of proteins and assemblies with minimal predetermined, high resolution structural data. These studies are vital in the quest to characterise the many protein assemblies which regulate biological processes.

2.2.1 Mass Spectrometers

The research presented in this thesis was carried out using both the Q-ToF 2 and Synapt HDMS mass spectrometers.

The principle components of the Q-ToF 2 hybrid quadrupole time of flight mass spectrometer are shown in Figure 2.1. This instrument employs electrospray ionisation (ESI) (Section 2.2.2) to ionise sample molecules and contains both quadrupole and time of flight (ToF) mass analysers (Section 2.2.3) separated by a hexapole collision cell. This arrangement permits mass selection of ions in the quadrupole, followed by collisional induced dissociation (CID) in the hexapole collision cell (Section 2.2.6) and mass to charge ratio (m/z) determination in the ToF sector.

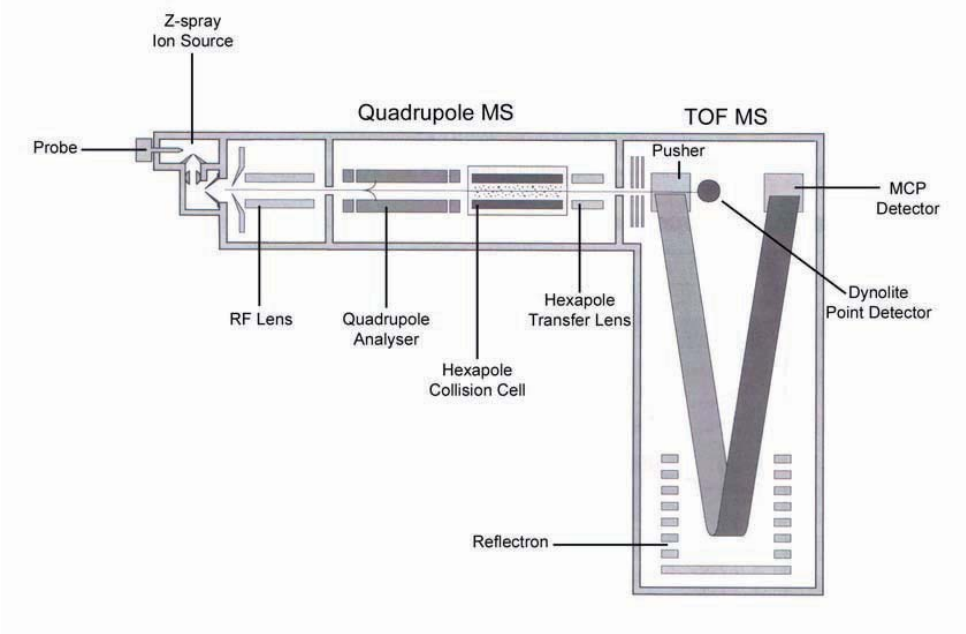


Figure 2.1: A schematic of the Q-ToF 2 mass spectrometer. The important components are labelled.

The Synapt HDMS quadrupole-ion mobility-time of flight mass spectrometer contains many similar features to the Q-ToF 2 (Figure 2.2). Again, ESI is employed to ionise sample molecules (Section 2.2.2), however, conventional Q-ToF technology (Section 2.2.3) is combined with travelling wave (T-Wave) IM separation (Section 2.2.4), allowing the analysis of gas phase ions both by their interaction with an inert buffer gas (their size/charge ratio), and their m/z [47]. CID is also able to be performed in the trap and transfer T-Wave sectors.

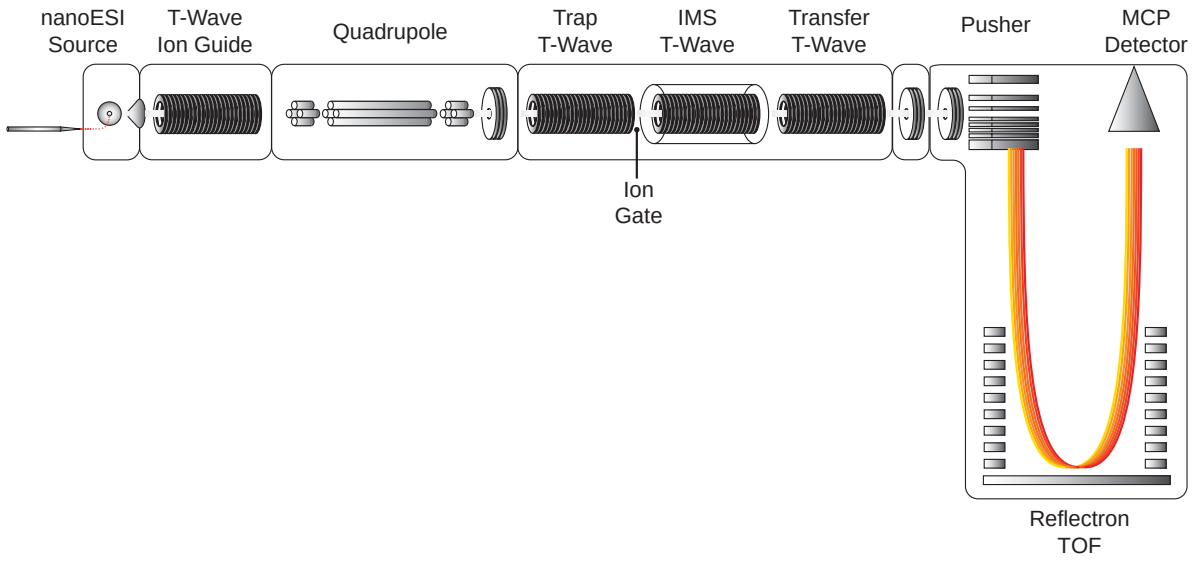


Figure 2.2: A schematic of the Waters Synapt HDMS Mass Spectrometer [48]. The important components are labelled.

2.2.2 Electrospray Ionisation

Early ionisation techniques, namely chemical and electron impact ionisation, limited the kinds of molecules that could be studied in the infancy of MS, as they are only suitable for the analysis of volatile samples. It was only in 1981 with the advent of fast-atom-bombardment that thermally labile molecules could be analysed by MS [49, 50]. Arguably the most significant methodological advancements have been the more recent introduction of matrix assisted laser desorption/ionisation (MALDI) [51] and ESI. These ionisation techniques have opened up the field of biological MS with characterisation of biological macromolecules in the megadalton range now relatively routine [52].

The development of ESI is attributed to the pioneering experiments of Dole *et al.* in 1968 [53]. However, it was not until 1984 that Fenn's group coupled ESI with MS for the analysis of biologically important compounds [54–58]. This ground-breaking work earnt Fenn a share of the 2002 Nobel prize for chemistry [59].

ESI is classified as a 'soft' ionisation technique as it imparts minimal internal energy on the system, permitting the retention of weak non-covalent interactions and minimising fragmentations in the ion source [60–63]. Additionally, multiply charged analyte ions can be produced, thus reducing the m/z of the formed species. This allows large macromolecules to be analysed within a m/z range accessible by commercial mass

spectrometers [40]. Additionally, for a protein or large peptide a series of multiply charged peaks are often observed and deconvolution of these data leads to determination of the molecular mass [64–66].

During the MS experiment, ionisation occurs at atmospheric pressure, and requires the analyte to be dissolved in a solvent where it becomes ionised to a certain extent, i.e. forming $(M+H)^+$ and $(M-H)^-$ ions [67]. This solution is then passed through a thin capillary to which a strong electric potential (1-4 kV) is applied, causing ions to gather at the tip of the needle forming an extended structure called a Taylor cone [68–70]. If the field applied is sufficiently high, the tip of the cone becomes unstable causing a fine mist of charged droplets to be sprayed [71, 72].

Heat and a nebulising gas (e.g. N_2) are applied to the spray, evaporating some of the solvent, reducing the size of the droplets. Once the droplets are sufficiently reduced in size, the Rayleigh stability limit is reached. This is the point where the repulsive Coulombic forces due to the charged nature of the droplet are equal to the attractive surface tension forces, given by Equation 2.1;

$$Q = 8\pi(\varepsilon_0 \gamma R^3)^{\frac{1}{2}} \tag{2.1}$$

where Q is the charge on the droplet, γ is the surface tension of the solvent, R is the radius of the droplet and ϵ_0 is the electrical permittivity [72–76]. Equation 2.1 provides an upper limit for the charge density necessary for droplet fission to occur and generally, Coulombic explosion occurs at approximately 65-110 % Q [77]. At this stage, the surface tension forces are overcome by the repulsive Coulombic forces and the droplets divide. Solvent evaporation and repeated droplet fission results in droplets which are small and contain few elementary charges (Figure 2.3) [72, 75].

Figure 2.3: The mechanism of electrospray ionisation [71].

From these small droplets, the ion is liberated into the gas phase in a method described by two models, namely the charge-residue [53] and ion-evaporation models [78]. Despite these two proposals, the precise mechanism of ion liberation is unknown and neither model is able to explain all observations [67, 72, 79, 80]. However, it is probable that the proposed mechanisms may operate individually or in tandem, depending on the analyte, to invoke ion formation [79].

Following ionisation, the generated ions are drawn through a sampling cone into the analyser portion of the mass spectrometer by application of both potential and pressure gradients [79]. The cone potential can be varied depending on the m/z of the ions to be analysed. Typically, ions with a larger m/z require increased cone voltages for optimal sensitivity [40]. Analyte ions are then guided into the analyser portion of the Synapt HDMS by a T-wave ion guide and by a radiofrequency (RF) lens in the Q-ToF 2.

ESI is extremely effective at producing gas phase ions from non-volatile analytes, unlike earlier ionisation techniques. Typically, aqueous organic solvents with a small quantity of acetic/formic acid (for positive ion mode) [81, 82] or triethylamine (for negative ion mode) are used as solvents in ESI-MS as they permit high ionisation efficiency. Additionally, the electrospray source may be directly coupled to a high performance liquid chromatography (HPLC) system and the analyte may be injected to the spectrometer using the solution conditions at elution. However, where it is necessary to retain native protein structure in the gas phase, these conditions are unsuitable as organic solvents can lead to denaturation [83–86]. Additionally, salts and buffer systems commonly used in molecular biology applications are incompatible with ESI-MS [87]. Consequently, ammonium acetate (NH_4OAc) buffered aqueous solutions are typically used as a solvent for the study of intact proteins and protein assemblies by ESI-MS [75, 88, 89].

Nanoelectrospray ionisation (nanoESI) was developed by Wilm and Mann [90, 91] in the mid 1990s and permits the use of significantly smaller sample volumes without compromising signal intensity. Conventional ESI sources require flow rates in the order of μL.min^{-1}, whilst nanoESI sources require flow rates in the order of nL.min^{-1}, and as such require sample volumes of only of 1-3 μL [79]. For nanoESI, the sample to be analysed is injected into a glass capillary that has been pulled to a fine tip and coated with a conductive metal such as platinum or gold. A potential is applied to the capillary (typically 1-2 kV)

via the metal coating, resulting in the formation of a Taylor cone, droplet fission and ion formation. The lower flow rate produces smaller droplets, greatly increasing desolvation efficiency [92], and the technique is more tolerant of salt contamination [91, 93] than conventional ESI.

In the study of native protein structure by MS, it is vital that the secondary forces which maintain native conformations, such as hydrogen bonding, hydrophobic interactions and electrostatic and van der Waals forces remain unaffected throughout ionisation and the pressurised regions of the spectrometer [60, 81]. For this reason, nanoESI is well suited to these studies as it is a gentler and more efficient ionisation technique, and conditions can be optimised so that native structure is retained upon vaporisation [60, 79, 81].

As a result, ESI-MS can be exploited to probe many aspects of protein structure and assembly, including the stoichiometry, stability and conformational changes involved in complexation. In the years since the advent of MS based structural biology, much research has been conducted using the technique to study biologically relevant protein assemblies and disease related protein aggregates [28, 94, 95].

2.2.3 Mass Analysis and Detection

Once ionisation has been achieved, it is necessary to separate ions based on their m/z for mass analysis and detection. A variety of instrumentation has been developed for this purpose including the magnetic sector, quadrupole [96], ion trap [97], Fourier transform ion cyclotron resonance [98], orbitrap [99] and time of flight (ToF) [100] mass analysers. The Q-ToF 2 and Synapt HDMS systems used in this work are fitted with quadrupole (Section 2.2.3.1) and ToF (Section 2.2.3.2) mass analysers and microchannel plate (MCP) detectors (Section 2.2.3.3). Increasingly, MS is being used to analyse intact native protein conformation, requiring optimisation of the ESI process and the control of several instrument voltages and pressures to ensure the transmission of relatively uncompromised structures [43, 60, 101, 102].

2.2.3.1 The Quadrupole Mass Analyser

Quadrupole mass analysers were developed by Paul in 1957 [103], who was awarded a share of the Nobel Prize in physics in 1989 for his work. They consist of four symmetrically arranged parallel rods, to which RF and direct current (DC) voltages are applied to generate an electric field [40, 71] (Figure 2.4). Rods opposite one another in the quadrupole have identical RF and DC fields applied, with the two pairs having RF fields 180° out of phase and opposite DC potentials [104]. This results in two pairs of rods, one pair is negatively charged and the other is positively charged. As ions enter the quadrupole they undergo an oscillating motion that is dependent on the electric field, and the applied quadrupole field forces the ions toward the centre of the device. In MS mode, the quadrupole simply operates as an ion guide by application of only an RF voltage, transferring all of the sample ions to the subsequent sectors of the spectrometer. When the instrument operates in MS/MS mode, an ion of interest can be selected by applying a constant electric field to the quadrupole [40, 71]. Variation of this applied electric field allows for ions to be selected based on their m/z for further analysis, whilst other ions are ejected from the quadrupole by striking the rods or the walls of the device [68, 96, 97, 105]. An ion that is selected can be passed into the hexapole collision cell (Q-ToF 2) or the trap/transfer T-Wave (Synapt HDMS) for CID (Section 2.2.6) [106–110].

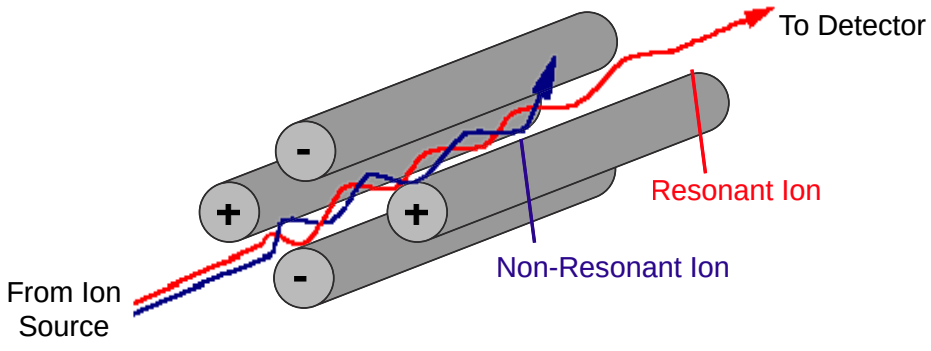


Figure 2.4: A quadrupole mass analyser. Red ions have a stable trajectory through the RF field, blue ions have an unstable trajectory and are ejected from the quadrupole.

2.2.3.2 The Time of Flight Mass Analyser

The ToF analyser was first described in 1946 by Stephens, who proposed that a mass spectrometer could be constructed without a magnetic field, by measuring the velocities of ions from a pulsed ion source [111]. Once the ions have entered the ToF analyser, an electric potential (V) from the pusher provides them all with an identical translational energy, given by Equation 2.2;

$$zV = \frac{mv^2}{2} \tag{2.2}$$

where m, z and v are the mass, charge and velocity of the ion, respectively. The ions then fly through a field-free region, with the time taken (t) to traverse this distance (d) given by Equation 2.3.

$$t = \left(\frac{m}{2Vz}\right)^{\frac{1}{2}} d \tag{2.3}$$

Ions with a smaller m/z will have a shorter flight time over the fixed distance between the pusher and the detector. As a result, the arrival time of an ion can be related to its m/z, with the instrument using this relationship to display a mass spectrum [68, 100, 111]. Once detection (Section 2.2.3.3) is complete, the next set of ions can be introduced into the ToF tube.

As a packet of ions enters the ToF tube, their location relative to the pusher determines the amount of kinetic energy they receive, with those closer to the pusher receiving more energy than those further away. Consequently, ions with the same m/z may obtain different energies and hence have different flight times, causing a reduction in resolution. This has been overcome by the development of the reflectron [112, 113], where a series of electrodes are used to create a retarding electric field at the end of the ToF tube which reverses the flight direction of the ions. Ions with higher kinetic energy will penetrate the reflectron more deeply than those with lower kinetic energy. This allows ions of the same m/z but different kinetic energies to catch up with one another and arrive at the detector simultaneously, thereby increasing resolution.

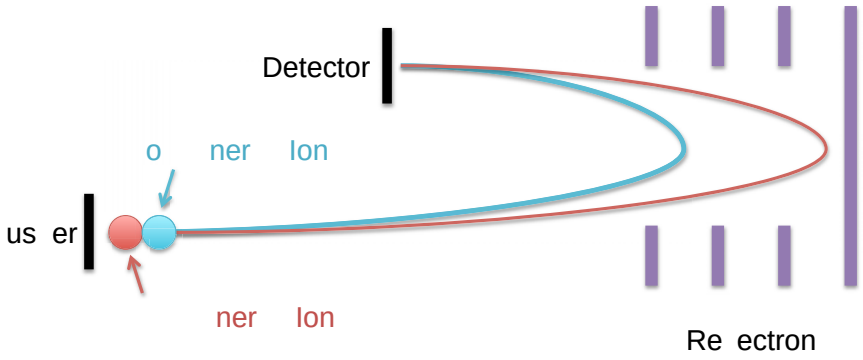


Figure 2.5: ToF mass analysis, showing two ions with the same m/z but different kinetic energies arriving together at the detector due to the presence of a reflectron ion mirror.

2.2.3.3 Ion Detection

Detection of ions is the last step in a MS experiment, and results in the production of a signal related to ion abundance. Both instruments used in this work utilise a microchannel plate (MCP) detector, which is a plate comprising an array of tiny tubes (microchannels) from one face to the other. Each microchannel functions as an electron multiplier, and an ion striking the MCP causes electrons to be emitted from the surface which are accelerated down the channel, striking the walls, and thereby releasing more electrons and amplifying the signal. From a single ion hitting the detector, up to 10^6 electrons can be produced by this process, giving a significant amplification of signal intensity [114, 115].

This amplified electron pulse is sent to the time to digital converter (TDC) and the arrival time relative to the ejection of ions from the pusher is recorded. Each packet of ions results in a spectrum of arrival times which are summed over the acquisition period to yield a mass spectrum [116].

2.2.4 Ion Mobility Spectrometry

IM has been utilised as an analytical tool for a number of decades, however, it was only in the 1990s that the coupling of IM with ESI and MALDI ion sources, as well as the development of theoretical methods enabled the technology to be utilised in the analysis of biomolecules [44, 117]. Traditional IM involves the separation of ions as they travel in a drift tube through a neutral buffer gas at different drift velocities (V_d) in the presence of an applied electric field (E), with V_d determined by the mobility of the ion (K), given by Equation 2.4.

$$V_d = KE \tag{2.4}$$

Smaller ions undergo fewer collisions with the neutral gas and thus pass through the drift tube faster than larger ions [43]. K, which defines the ion's drift velocity can therefore be related to the ion's collision cross-section (Ω or CCS). Typically, instead of directly determining K, the time required for the ions to drift though the distance of the drift tube (t_D) is measured. A direct relationship exists between t_D and Ω, given by Equation 2.5;

$$\Omega = \frac{(18\pi)^{\frac{1}{2}}}{16} \frac{q}{(k_b T)^{\frac{1}{2}}} \left[\frac{1}{m_l}\frac{1}{m_B}\right]^{\frac{1}{2}} \frac{t_D E}{L} \frac{760}{P} \frac{T}{273.2} \frac{1}{N} \tag{2.5}$$

where K_b is the Boltzmann constant, q is the charge on the ion, T is the absolute gas temperature, P is the buffer gas pressure, L is the length of the drift tube and m_l and m_B are the mass of the ion and buffer gas, respectively. This direct relationship allows experimentally determined drift time data to be converted to CCS, giving an indication about the size and topology of ions [117].

2.2.4.1 Travelling Wave Ion Mobility Mass Spectrometry

Traditional drift tube IM spectrometers have been coupled to mass spectrometers to record m/z and CCS simultaneously, but these instruments have proven difficult for the analysis of large protein complexes due to issues with sensitivity and maintaining non-covalent interactions. The recent development and commercial availability of T-wave IM has revolutionised the field and allowed the routine study of intact protein structures by IM-MS.

The IM sector of the Synapt HDMS consists of three travelling wave-enabled stacked ring ion guides (TWIGs). Each of the stacked ring ion guides consists of a series of ring shaped planar electrodes arranged in parallel (Figure 2.6). Opposite phases of an RF voltage are applied to adjacent electrodes producing a potential energy barrier that confines the ions to the centre of the guide, stabilising the ions' trajectory and preventing loss by radial diffusion. The ions are propelled through the ion guide in the presence of a background gas by applying a short DC voltage pulse to pairs of ring electrodes in sequence, resulting in a travelling wave voltage on which the ions can "surf". This ability of an ion to "surf" the wave through an inert gas is dependent on its mobility. Species with low mobility experience a greater drag and are unable to keep up with the travelling wave. This phenomenon causes ions of low mobility to slip behind the travelling wave (termed "roll over events") less often than those species of high mobility, thus allowing separation to occur on this basis, and giving arrival time distribution (ATD) information (Figure 2.7) [48, 118]. In the Synapt HDMS, three TWIGs form the T-Wave IM sector, the first is called the trap, the second is the mobility cell, where IM separation occurs, and the third is the transfer. IM separation is achieved in the mobility cell rather than the other two TWIGs due to an increased buffer gas pressure and the different electrical pulses that are applied to the ring electrodes.

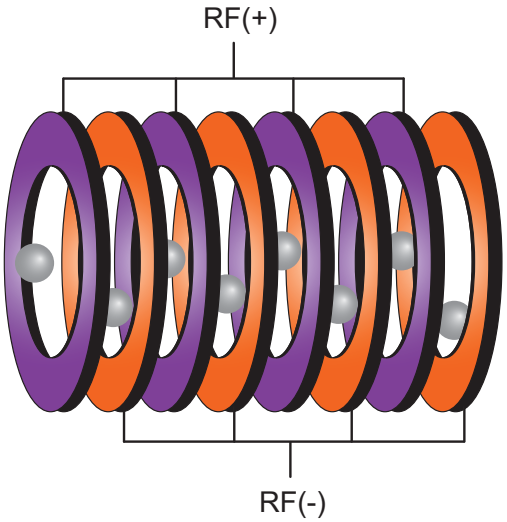


Figure 2.6: Schematic of the ring electrodes which comprise the TWIG. Adapted from [48].

Figure 2.7: Schematic of the operation of a TWIG in the propulsion of ions through the drift tube [48]. Ions of high mobility are shown in orange, ions of low mobility are shown in purple and fall behind the travelling wave.

The trap TWIG is used to accumulate ions during the previous mobility separation, and the final ring electrode is DC-only, allowing the gated release of packets of ions by modulating the applied voltage. Once released into the mobility cell, IM separation is performed as detailed above and shown in Figure 2.7. Pressures in this region can be optimised for the retention of native protein structure and interactions. Additionally, the travelling-wave height/voltage and travelling-wave velocity must be optimised in this region to achieve optimal IM separation. The travelling wave velocity is given in $m.s^{-1}$, and is determined by dividing the distance between electrode pairs by the time the pulse remains on each pair. The transfer TWIG has a constant travelling wave voltage to ensure that IM separation is maintained as the ions are transferred to the ToF for mass analysis. In order to obtain both IM data and m/z data simultaneously, the acquisition occurring in the ToF sector of the instrument is synchronised with the gated release of ions from the trap TWIG into the IM sector.

The instrument enables fragmentation of ions in either or both of the trap and transfer TWIGs as a result of CID (Section 2.2.6) [106–110]. These fragments are then analysed by the ToF sector of the spectrometer in the normal way, described above.

T-Wave IM spectrometry yields ATD information, but to obtain CCS values, calibration of the measured ATD data using standards of known CCSs is required. The IM-MS spectra of these calibrants must be recorded with the same IM parameters as the analyte (such as the height and velocity of the voltage waves), as these influence the obtained CCS [43]. In the analysis of proteins, typical calibrants include denatured cytochrome c, myoglobin and ubiquitin, which have published CCSs for a variety of charge states [43, 119].

Additionally, the data obtained must be corrected for mass-dependent flight time [43]. This is performed using Equation 2.6.

$$t_D^{'} = t_D \frac{c\sqrt{m/z}}{1000} \tag{2.6}$$

Where t_D and $t_D^{'}$ are the experimentally determined and corrected drift times (msec), respectively, m/z is the mass to charge ratio of the ion and C is a constant which varies from instrument to instrument.

The advent of IM-MS has significantly increased the amount of information that can be obtained from native MS data. Not only is it possible to characterise protein assemblies by their mass, but insight into the size and topology of a complex can be gained simultaneously. Low-resolution CCS information obtained from IM-MS data can be used in tandem with computational modelling and data from other techniques to gain an insight into the architecture of a protein assembly. Today, integrative approaches, combining aforementioned MS techniques with other low resolution data, are becoming commonplace in the development of structural models of proteins and assemblies with minimal predetermined, high resolution structural data.

2.2.5 Computational Approaches for Calculating CCS of Proteins

IM-MS studies of proteins and protein assemblies are often carried out concurrently with molecular modelling using both available structures (such as those deposited in the Protein Data Bank) or modelled structures from either course-grained projections or *ab initio* studies [35] to determine CCS. This can be achieved using readily available software, including the program MOBCAL [120, 121]. The program is capable of using three different theoretical methods to calculate cross-section, namely, the exact hard sphere scattering (EHSS) model, the hard sphere projection approximation (PA) model and the trajectory method. For large biomolecules, the EHSS and PA models are typically used due to the decreased computational time they require, and as such only these two models will be discussed here. CCS is influenced by the topology of an ion and therefore its interactions and collisions with a neutral buffer gas. These methodologies attempt to replicate the ion structure *in silico* and model its collisions with the buffer gas to calculate a theoretical CCS.

The PA methodology is the simpler of the two models used in CCS calculation. It involves replacing an ion's cross section with its projection, or shadow, and averaging the projection that is created by every possible orientation of the ion [122]. By this method, it assumes that ions are spherical with this average projection, and therefore multiple collisions with the buffer gas cannot occur. This is not always the case for large molecules with rough surfaces (like proteins) which may deflect gas molecules resulting in further collision(s) with the ion. Additionally, the method assumes a uniform charge distribution across the

surface and that there are no long-range interactions between the ions and the buffer gas [42]. This method generally provides a good approximation for small molecules, but often underestimates the cross-sections of proteins [122].

The EHHS model also takes into account the surface structure, and as a result is said to provide a better representation of larger ions with complex structures. Because of this, it correctly treats scattering between the ion and the buffer gas, but still disregards ion-buffer gas interactions, as does the PA method [42, 123]. This method is well known to overestimate CCS.

2.2.6 Collision Induced Dissociation of Peptides

Structural information about an ion can also be obtained using tandem mass spectrometry (MS/MS) experiments [106–110]. This process involves mass selection of a precursor ion in the quadrupole mass analyser after which it is passed into the hexapole collision cell (Q-ToF 2) or TWIG (Synapt). Here, CID is performed by energising the ion using collisions with a noble gas, such as He or Ar. This results in decomposition of the precursor ion to produce fragment ions which can provide structural information about the precursor.

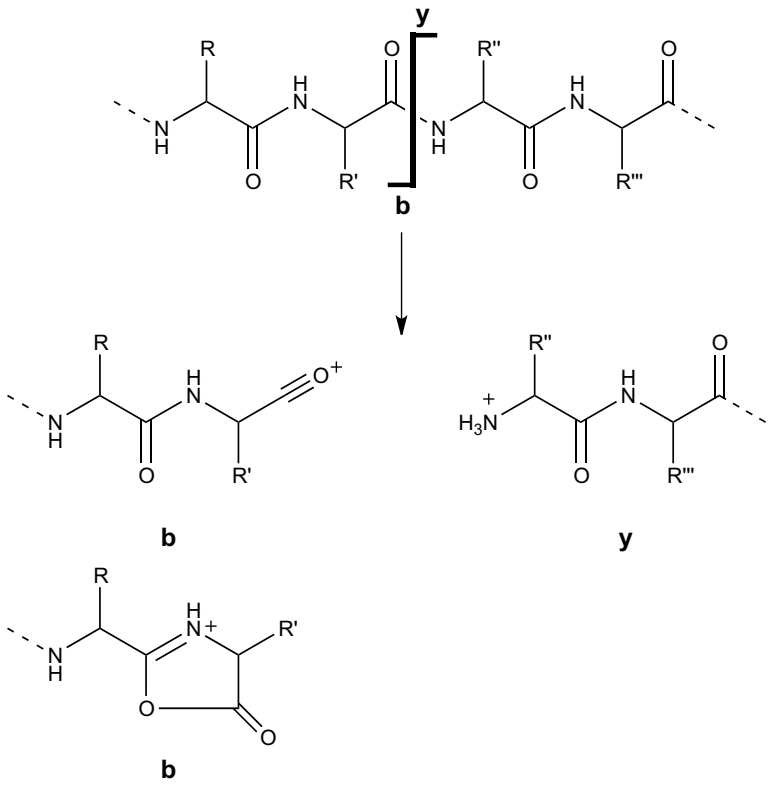
Analysis of the fragment ions produced upon CID during MS/MS experiments allows the amino acid sequence of peptides to be determined rapidly. Characteristic fragmentation processes occur in both positive and negative ion MS and are exploited to identify a peptide's primary structure.

2.2.6.1 Fragmentations in Positive Ion Mass Spectrometry

Collisional activation of linear peptides in the positive ion mode produces spectra that are dominated by the facile cleavage of peptide linkages between amino acid residues. The fragment cations produced allow easy sequencing of the peptide under investigation. Two fragmentations of the amide bond that can be used for sequencing are the **b** and **y** cleavages (Scheme 2.1) [40, 67, 124, 125]. The **b** fragmentations allow for sequencing of the peptide from the C-terminus, whilst the **y** cleavages allow sequencing from the N-terminus [126].

Mass differences between ions of the same fragmentation series are utilised to determine sequential residues in the peptide [124, 127].

Structures of the **b** and **y** fragment ions are illustrated in Scheme 2.1. In this scheme, the **b** ions are shown as both their originally proposed acylium cation structure and the more recently suggested cyclic oxazolones [128–132]. Other cyclic structures have also been proposed [133, 134].



Scheme 2.1: Positive ion cleavage of the peptide linkage to afford **b** and **y** fragment cations [124, 129–132]. Both the acylium and cyclic oxazolone cation structures of the **b** ion are shown.

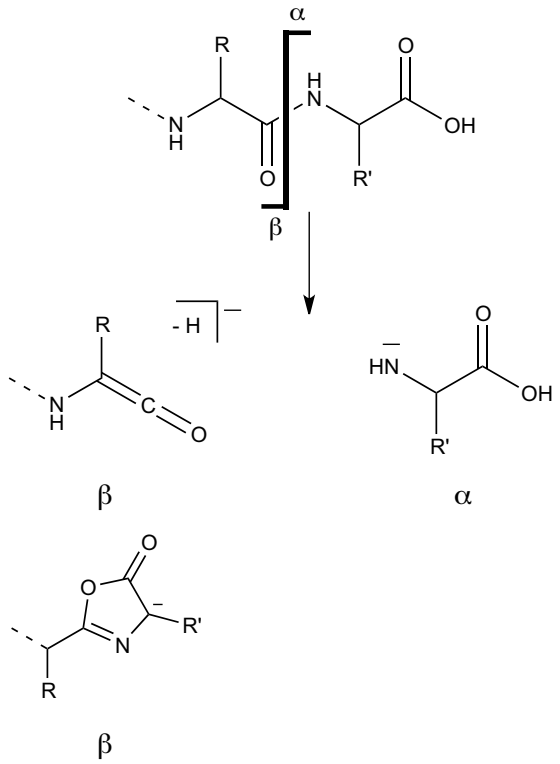
Utilising a combination of these fragmentations allows for the peptide identity to be determined in most cases. Isomeric Leu and Ile amino acids can be distinguished by application of techniques complementary to MS, such as Edman sequencing [135], while the isobaric amino acids Lys and Gln can be differentiated by instruments with high mass accuracy, by Edman sequencing, or by analysis in the negative ion mode (Section 2.2.6.2).

2.2.6.2 Fragmentations in Negative Ion Mass Spectrometry

Historically, negative ion MS sequencing of peptides has found limited application for the study of complex systems, as collisional activation of peptide anions affords mass spectra of increased complexity when compared with their positive ion counterparts. Many of the most facile fragmentation pathways in negative mode correspond to side chain induced cleavages, and not the amide backbone cleavages which are well suited to peptide sequencing. Consequently, these amide backbone fragmentation products are often of low relative abundance. Despite these limitations, negative ion MS can yield data that parallels those obtained by positive ion MS [136], and can even provide structural information that cannot be obtained in the positive mode.

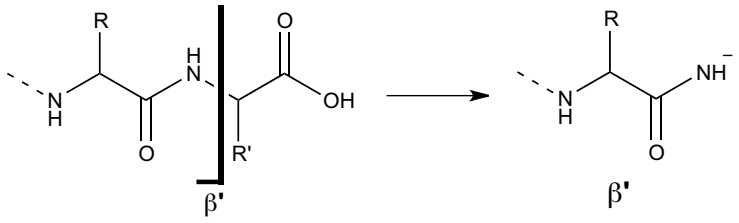
Identical information from the positive ion **b** and **y** cleavages are obtained from the negative ion β and α cleavage processes (Scheme 2.2). Thus, the β and α fragmentations allow for sequencing of the peptide from the N- and C-termini, respectively [136]. These processes are proposed to proceed from enolate anions adjacent to the amide carbonyl in the peptide backbone to yield β and α cleavage ions as shown in Scheme 2.2. Alternate cyclic structures, which form by different mechanisms, have also been proposed for the β fragment anions as shown in Scheme 2.2 [132, 133, 137].

The energetics of these processes have been investigated using *ab initio* calculations and reveal that the α and β cleavage processes are endothermic by 134 kJ.mol^{-1}and 285 kJ.mol^{-1} respectively, with a maximum transition state barrier of 397 kJ.mol^{-1} [138].



Scheme 2.2: Negative ion cleavage of the peptide linkage to afford α and β fragment anions [136, 137].

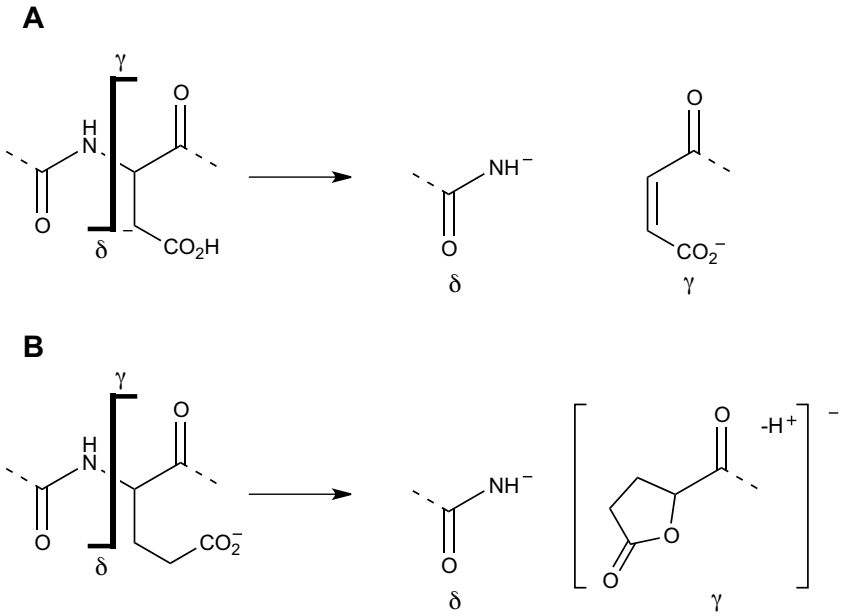
An additional negative ion backbone amide cleavage also occurs and has been called the β' fragmentation, which again provides sequencing information from the C-terminus [136, 138, 139]. The β' fragment anion has the structure shown in Scheme 2.3, and is formed by a mechanism involving S_Ni substitution/cleavage and cyclisation to liberate the fragment anion [139].



Scheme 2.3: Structure of the β' backbone fragment anion.

Backbone cleavage can also be effected by anions situated on the side chains of particular amino acids. These side chain induced backbone cleavages are caused by amino acids which can be deprotonated in a position adjacent to the backbone to form a resonance stabilised anion, such as in Asp, Asn, Glu, Gln, Phe and Tyr [136, 140]. Instead of cleavage of the amide bond, as occurs in the α, β and β' processes, side chain induced backbone

fragmentation occurs adjacent to the amide, with the NH-CH bond being cleaved [141]. CID of small peptides generally yields more side chain induced backbone fragmentation products (over amide bond fragmentations) from Phe, Tyr and His, when compared with larger peptides, whilst Asp and Asn fragmentations remain abundant for all peptides. The resultant products have been named the δ and γ ions (Scheme 2.4A) and are formed by mechanisms that have been reviewed elsewhere [136, 140]. Side chain induced backbone cleavage can also occur from Glu and Gln residues, even though the side chain enolate that forms is not directly adjacent to the backbone. Consequently, the mechanism of these backbone cleavages differs (Scheme 2.4B) [142, 143].



Scheme 2.4: Structures of the δ and γ ions formed by side chain induced backbone cleavage from (A) Asp and (B) Glu.

It is common to lose H_2O (18 Da) from the Asp and Gln γ ions and NH_3 (17 Da) from the Asn and Gln γ ions. These products often are the base peaks in these spectra and can be useful for distinguishing between the isobaric Lys and Gln amino acids on MS instruments without high resolution capability [144].

In the negative mode, characteristic side chain cleavages also occur (reviewed in [140]), a phenomenon not generally seen in the positive ion MS/MS spectra of peptides. Many of these processes are only observed in the spectra of short peptides (up to four amino acid residues in length) and only several are prevalent irrespective of the size of the peptide (indicated in Table 2.1). As the peptide decreases in size, these side chain cleavage ions

often dominate the negative ion spectrum, concealing the α, β and β' backbone cleavage ions. In some instances, α, β and β' cleavage ions are observed from side chain cleavage ions instead of the $(M\text{-}H)^-$ ion e.g. in Ser containing peptides they are seen from the $[(M\text{-}H)^-\text{-}CH_2O]^-$ fragment [136, 140].

Negative ion MS also shows many advantages over conventional positive ion MS in the study of post-translational modifications, such as disulfide bonds which undergo characteristic fragmentations (Chapter 3).

Table 2.1: Characteristic negative ion side chain fragmentations [140]. Fragmentations prevalent in peptides of all sizes are indicated by *.

Amino Acid	**Loss (or formation)**	**Mass (Da)**	
Ala	$Me^{\cdot}$	15	
Val	$Pr^{\cdot}$	42	
Leu (Ile)	$Bu^{\cdot}$	57	
Phe	$(PhCH_2^-)$	91	
Tyr	$(p\text{-}HOC_6H_4CH_2^-)$	107	
	$O{=}C_6H_4{=}CH_2$	106	
Trp	C_9H_7N	129	
Ser	CH_2O	30	*
Thr	MeCHO	44	*
Cys	H_2S	34	*
Met	MeSH	48	
	MeSMe	62	
	CH_2CH_2SMe	75	
Asp	H_2O	18	*
Asn	NH_3	17	*
Glu	H_2O	18	*
	$(C_5H_6O_3N^-)$	128	*
Gln	NH_3	17	*
Arg	$NH{=}C{=}NH$	42	

2.3 Nuclear Magnetic Resonance Spectroscopy

The first NMR spectrum of a protein was recorded in 1957 [145], some 11 years after the discovery of the NMR phenomenon [146, 147]. In spite of this achievement, it was another 28 years before instrumentation and methodological advances meant that that NMR could be used to solve the complete 3D solution structure of a protein [148]. The greatest advancements in the field include the introduction of multidimensional NMR [149], the advent of high field spectrometers (which provide increased resolution and sensitivity), as well as the development of Fourier transform spectroscopy. These along with the implementation of distance geometry calculations have all aided the development of NMR as a fundamental technique in structural biology, allowing the determination of both low and high resolution structural information about biological molecules and their assemblies [150–152].

2.3.1 The Fundamentals of NMR Spectroscopy

The NMR experiment is possible because nuclei possess an intrinsic angular momentum known as nuclear spin (given the quantum number I) [153]. Only nuclei with non-zero I values are active in NMR experiments. Spin active nuclei include ^{1}H, ^{13}C and ^{15}N, which all have $I = {}^{1}/_{2}$ [71, 154]. In the presence of an applied magnetic field $\mathbf{B}_0$, the orientations of the spins are influenced and the nuclei are able to adopt one of $(2\,I + 1)$ orientations within the field [71]. Thus, the spins of ^{1}H, ^{13}C and ^{15}N nuclei can adopt one of two orientations, aligning anti-parallel to $\mathbf{B}_0$ in a high energy state N_β, or in a low energy state N_α parallel to $\mathbf{B}_0$ [71]. The energy difference between N_α and N_β is given by;

$$\Delta E = \frac{h\gamma \boldsymbol{B_0}}{2\pi} \tag{2.7}$$

where γ is the gyromagnetic ratio, h is Planck's constant and $\mathbf{B}_0$ is the strength of the applied magnetic field. Nuclei can move between these two orientations by absorbing or emitting radiation corresponding to ΔE [153].

The two energy levels are unequally populated, with more spins in the N_α state, resulting in a net magnetic moment, **M** in the direction of $\mathbf{B}_0$ (Figure 2.8) [71, 153]. The number of nuclei initially populating each of the energy levels is described by the Boltzmann distribution [71], given by;

$$\frac{N_\beta}{N_\alpha} = e^{\frac{-\Delta E}{kT}} \tag{2.8}$$

where T is the temperature (in K), ΔE is the energy difference between N_α and N_β and k is the Boltzmann constant. Generally, ΔE is very small, but this increases when higher field strengths are used, and thus the population of nuclei in the high energy state decreases (Equations 2.7 and 2.8), increasing sensitivity [71].

When a radiofrequency (RF) pulse is applied to the sample, transitions are induced between the two energy levels, promoting those nuclei in the N_α state to N_β, altering **M** [71]. However, the energy of the RF pulse must correspond to ΔE in order for this to occur. This energy gap is related to the Larmor precession frequency (ν), the frequency at which the nuclei naturally precess in a magnetic field, which is given by Equation 2.9.

$$\nu = \frac{\gamma \boldsymbol{B_0}}{2\pi} \tag{2.9}$$

In modern NMR experiments, RF pulses of short duration are applied to the sample, with the magnetic component of the pulse, $\mathbf{B}_1$ (in the *xy* plane), projected 90° to the main magnetic field, $\mathbf{B}_0$ (along the *z* axis) [155] (Figure 2.8). Whilst the pulse is being applied, a force is exerted on **M**, tipping the magnetisation towards the *y*-axis, and as long as it continues the nuclei preferentially precess about its direction [156]. The angle by which the magnetisation is rotated increases with the length of application of $\mathbf{B}_1$. Normally, the irradiation with $\mathbf{B}_1$ is stopped once magnetisation reaches the y axis i.e. once a 90° pulse has been applied (Figure 2.8) [156]. After the pulse, the magnetisation in the *xy* plane is detected, during the detection phase (t_2) of the experiment. The magnetisation in this plane decreases with time as it relaxes back to the *z* axis by two relaxation processes, namely spin-spin relaxation (T_2) and spin-lattice relaxation (T_1) [156]. The reduction in magnetisation in the *y* plane is detected as a free induction decay (FID) (Figure 2.8) with a frequency corresponding to the Larmor procession frequency [153].

Each chemically distinct nucleus has a different resonant frequency. This is because the local magnetic environment contributes to the actual magnetic field experienced by the nucleus (Equations 2.7 and 2.9). In Fourier Transform NMR, the RF pulse contains radiation which spans a large frequency range in order to excite all the nuclei in the sample at once [155]. Fourier transformation of the obtained FID converts it from a function of time to a function of frequency, producing a 1D NMR spectrum (Figure 2.8) [154]. The FID obtained in t_2 contains multiple frequency components due to the different Larmor precession frequencies of chemically distinct nuclei. Therefore upon Fourier transformation, the NMR spectrum will show a unique signal for each of the chemically distinct nuclei of interest in the molecule.

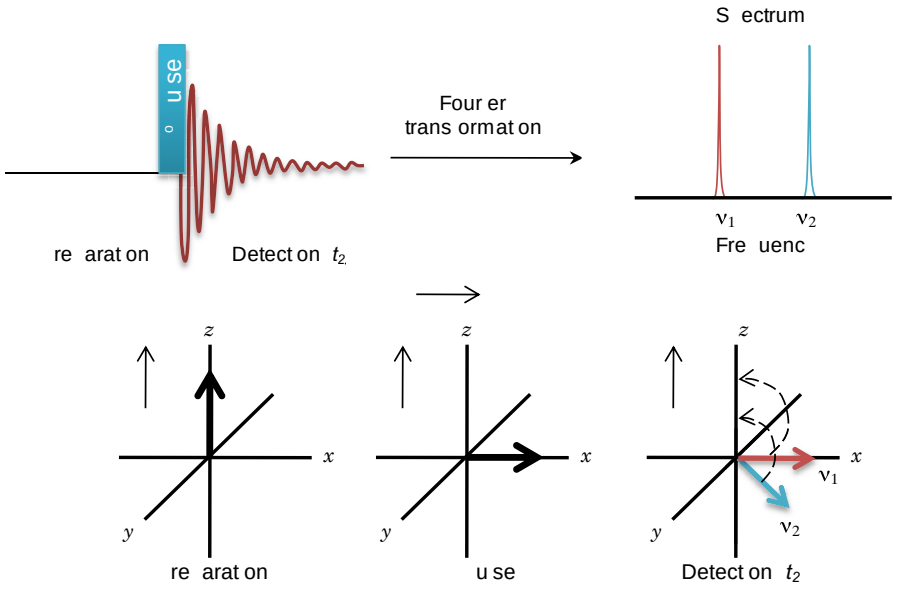


Figure 2.8: Schematic of the 1D NMR experiment and vector diagrams illustrating the change in orientation of the bulk magnetisation during an NMR experiment. [154]

2.3.2 Two and Three Dimensional NMR Spectroscopy

The standard 1D NMR experiment has only one frequency dimension. However, when analysis of large molecules is required, including for peptides and proteins, this method can result in very complicated spectra containing numerous overlapping signals [31]. 2D and 3D NMR experiments have a second and third frequency dimension, respectively, yielding spectra of increased resolution, so overlapping resonances can be separated and distinguished. The development of these multidimensional techniques along with

multinuclear experiments has allowed structural problems of significantly increased complexity to be addressed using NMR [157, 158].

In 1D NMR experiments, the preparation and t_2 phases occur with only the 90° pulse separating them (Figure 2.8). 2D NMR experiments contain two new elements, namely the evolution (t_1) and mixing (τ_m) periods which occur between the preparation and t_2 stages [71, 153, 154, 159, 160]. In order to obtain a 2D data set, a number of experiments are recorded with incremented t_1 periods. During τ_m, magnetisation transfer occurs (this process is different for each NMR experiment), determining which correlations are observed.

Each phase of the NMR experiment is made up of a series of delays and RF pulses, constituting the pulse sequence. Different pulse sequences allow different information to be obtained by permitting the transfer of magnetisation between dipolar coupled spins (nuclei that are close in space) or between nuclei that are spin-spin coupled (nuclei that are nearby in the molecule). These correlations can be used to deliver information about the primary structure of the molecule, its three-dimensional structure, or both. 2D Fourier transformation of the data gives a spectrum which is viewed as a contour plot to indicate the intensity of the peaks. A schematic of the general 2D NMR experiment is illustrated in Figure 2.9.

Whilst the increased resolution that 2D NMR provides is often sufficient for the study of peptides, 2D spectra of proteins with masses greater than approximately 10 kDa are often still of insufficient resolution [31]. This is due to, (i) the greater number of resonances present in spectra of these molecules, and, (ii) longer rotational correlation times which leads to faster transverse relaxation and broadened linewidths [31]. As a result, an additional frequency dimension is added, yielding a 3D NMR spectrum that provides increased spectral resolution [153, 159]. 3D NMR experiments are constructed by combining two 2D pulse sequences [31, 159]. As a result, the mixing period of one 2D experiment is combined with the preparation period of the second [31, 159]. The result is that the signal is recorded in t_3 as a function of not one, but two evolution times (t_1 and t_2), which like 2D spectra are successively incremented (independently) [31, 153, 159]. Following acquisition, a 3D Fourier transformation is applied to the data yielding a 3D frequency spectrum with dimensions F_1, F_2 and F_3 [31, 153, 159]. The data can be

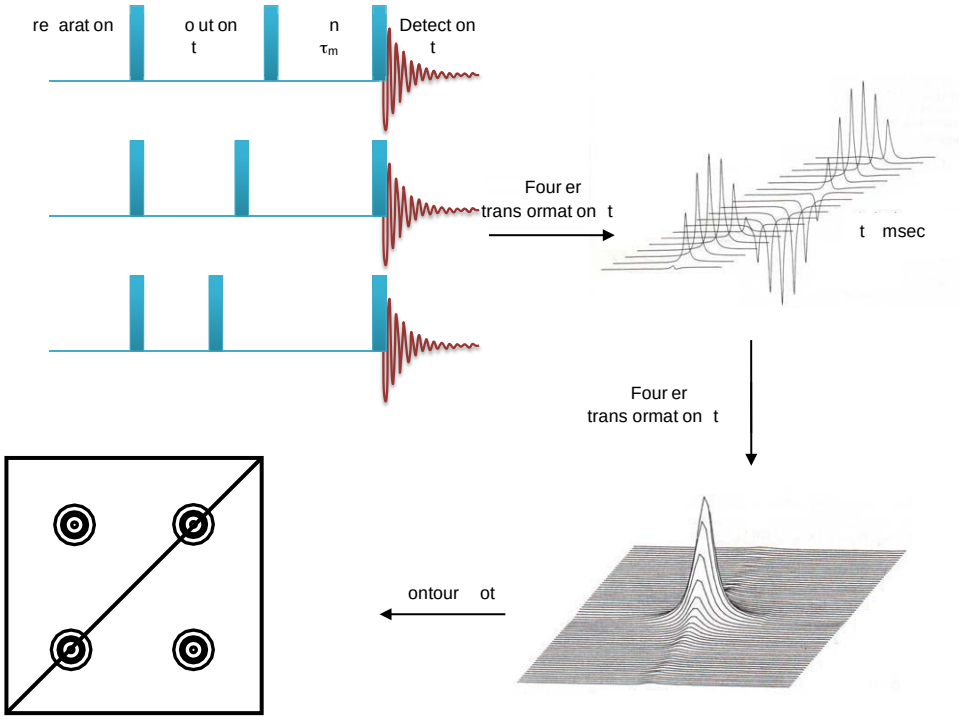


Figure 2.9: Schematic of the 2D NMR experiment [154, 159, 161].

represented as a cube, but practically, two dimensional slices of this cube are used in data analysis. The following sections describe both 2D and 3D NMR experiments necessary for the peptide and protein structure determination in this work.

2.3.3 Two Dimensional Spectroscopy of Peptides

In the assignment of resonances and subsequent 3D structure determination of a peptide, a number of 2D NMR experiments are used. These each differ in their pulse sequence and hence provide differing structural information. A detailed discussion of pulse sequences is not provided here, rather the focus is on the information each experiment can provide.

Total correlation spectroscopy (TOCSY) experiments show correlations between all proton resonances of a spin system. It exploits a phenomenon called cross-polarisation, where magnetisation is relayed between scalar coupled (through-bond) spins, to correlate all protons of a spin system [162–165]. Scalar coupling does not occur across the peptide bond, and as such, each amino acid residue constitutes a single spin system, although some contain two (e.g. Phe) and others three (e.g. Trp) due to the presence of aromatic

proton signals which exist as separate spin systems [71, 160]. Analysis of coupled spins in the TOCSY spectrum allows the amino acid identity to be determined, as each displays a characteristic pattern of connectivities [154]. By this mechanism, analysis of the TOCSY spectrum can lead to the the assignment of most proton resonances. If there are overlapping signals in the spectrum, correlated spectroscopy (COSY) experiments can be used to unambiguously assign resonances [166]. COSY spectra show interactions between protons that are two or three bonds apart, by a process known as spin-spin coupling [154, 167, 168].

Once the spin-systems have been identified, nuclear Overhauser effect spectroscopy (NOESY) experiments can be utilised to connect the spin systems together. The nuclear Overhauser effect (NOE) refers to the change in intensity of an NMR resonance, as a result of cross-relaxation (dipolar coupling), when a nearby resonance is irradiated [71, 154, 160, 169, 170]. NOESY experiments are utilised for the assignment of amino acid residues in a polypeptide and provide distance restraints for structure calculations (Sections 2.3.6 and 2.3.7). The relative intensity (I) of a NOESY cross-peak is related to the inter-atomic distance (r) by the relationship shown below.

$$I \propto r^{-6} \tag{2.10}$$

As a result, the NOE intensity dramatically decreases when the dipolar coupled spins are further apart. Consequently, the maximum observable spatial separation between spins is approximately 5 Å [71, 154].

2.3.4 Heteronuclear Correlation Spectroscopy

Heteronuclear correlation spectroscopy provides information about the coupling between a heteronucleus (such as ^{13}C or ^{15}N) and a proton. Coherence transfer is used in these experiments, like the COSY experiment, but the magnetisation is now transferred between a proton and a heteronucleus [154, 171, 172].

The most commonly encountered experiment of this type is the Heteronuclear Single-Quantum Coherence (HSQC) experiment which gives the identity of a heteronucleus

and its directly attached proton [154]. ^{13}C HSQC experiments can be used for assigning αC resonances in a peptide or protein whilst ^{15}N HSQC experiments are particularly useful in the analysis of ^{15}N enriched recombinant protein samples. In this spectrum, a single peak is observed for each amide nitrogen/hydrogen pair. Additionally, side chain NH groups are also detected, with the terminal amine group present in glutamine and asparagine residues displaying a further two peaks in the spectrum with the same nitrogen shift [31].

The ^{15}N HSQC is used extensively to track conformational changes in a protein upon complexation with a ligand of interest. Upon titration with a binding partner, the chemical shifts of the amide nitrogen and proton will be perturbed due to the change in conformation that binding induces [173]. These changes can be tracked during the titration series, with chemical shift changes greater than 0.5 ppm and 0.05 ppm in the ^{15}N and ^{1}H dimensions respectively often considered significant [174].

2.3.5 Three Dimensional Spectroscopy of Proteins

Triple resonance experiments are used to assign the resonances of large biological macromolecules. In 3D spectra, each peak is designated by three coordinates relating to the chemical shifts of three nuclei (for isotopically enriched proteins, a combination of ^{1}H, ^{15}N and ^{13}C). Each experiment exploits a different path of magnetisation transfer and thus provides different information. In this work, two experiments were utilised to assign protein backbone chemical shifts, namely the HNCACB and CBCA(CO)NH experiments.

The HNCACB experiment [175] correlates the $^{13}C\alpha$ and $^{13}C\beta$ resonances with the amide ^{1}H, ^{15}N resonances of both the same residue and the preceding residue (Figure 2.10) [31]. For many proteins, this experiment alone allows most, if not all, the ^{1}H, ^{15}N, $^{13}C\alpha$ and $^{13}C\beta$ resonances to be assigned. This is because the $^{13}C\alpha$ and $^{13}C\beta$ shifts provide information on the amino acid identity whilst sequential connectivities are also indicated in the spectrum. Interpretation of the spectrum is further aided by the fact that the $^{13}C\alpha$ resonances are 180° out of phase when compared with the $^{13}C\beta$ resonances, allowing these to be distinguished with ease and unambiguously assigned [175, 176].

HNCACB spectra, like all 3D NMR spectra, comprise a cube of data, but in practice, so-called ‘strips’ of the spectrum are taken and analysed, with each corresponding to a

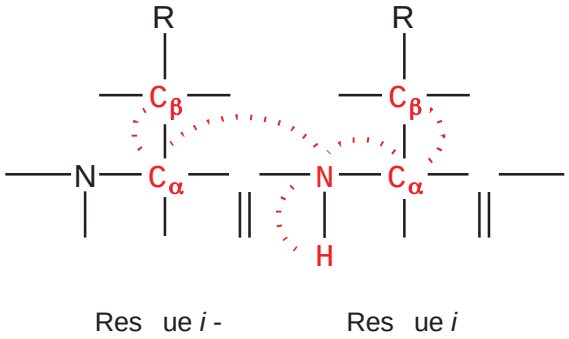


Figure 2.10: Illustration showing the correlations detected in the HNCACB experiment. The correlation network is illustrated in red.

cross-peak from the ^{15}N HSQC experiment (Figure 2.12A). Essentially, the experiment is ^{15}N HSQC edited, with the projection along the ^{13}C dimension being a ^{15}N HSQC. Each strip in the spectrum contains four peaks, one for each of the $^{13}C\alpha$ and $^{13}C\beta$ shifts of residue *i* and (*i -1*) (where applicable). Glycine residues are easily identified by this method as they lack a $^{13}C\beta$. Alanine, serine and threonine residues can also be easily identified due to their characteristic $^{13}C\alpha$ and $^{13}C\beta$ shifts [31].

The CBCA(CO)NH experiment [177] is similar to the HNCACB experiment, and together they are the two of the most vital experiments in backbone assignment. The data obtained from this experiment again comprise amide ^{1}H and ^{15}N correlations with $^{13}C\alpha$ and $^{13}C\beta$, however these are only between the carbons of the preceding residue (Figure 2.11). These inter-residue correlations are obtained by using the ^{13}C spin as a means of coherence transfer [31]. Therefore, each strip in the spectrum only contains two peaks, correlating the amide ^{1}H and ^{15}N with the $^{13}C\alpha$ and $^{13}C\beta$ of the preceding residue (Figure 2.12).

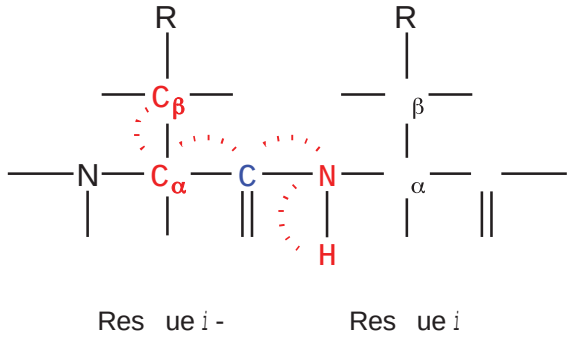


Figure 2.11: Illustration showing the correlations detected in the CBCA(CO)NH experiment. The correlation network is illustrated in red. The CO (shown in blue) is used to transfer the magnetisation only and is not observed in the spectrum.

Using both the HNCACB and CBCA(CO)NH spectra in concert can allow the assignment of most backbone amide ^{1}H, ^{15}N, ^{13}Cα and ^{13}Cβ resonances. Inter-residue connectivities in the HNCACB spectrum can be easily identified by overlaying the CBCA(CO)NH data as the same inter-residue peaks will appear in both spectra. Those peaks which do not appear in the CBCA(CO)NH must be due to intra-residue correlations (Figure 2.12C).

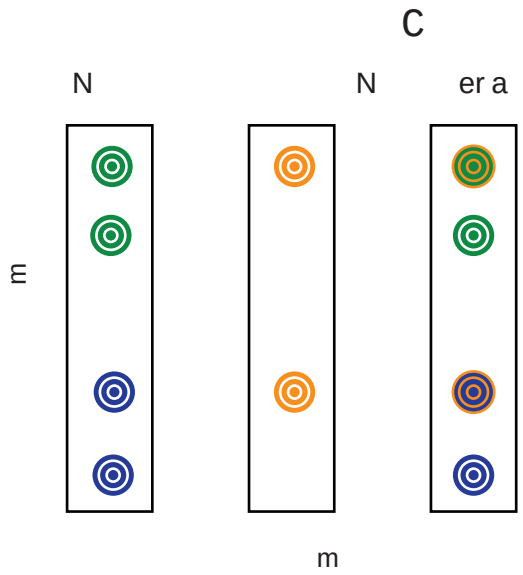


Figure 2.12: Typical strips from (A) HNCACB and (B) CBCA(CO)NH spectra. (C) Shows the overlaid spectra to demonstrate how they are typically used in sequential assignment. Peaks in both spectra can be identified as correlations with the preceding residue.

2.3.6 Resonance Assignment

In order to assign the NMR spectrum of a biomolecule, a combination of the 2D (for peptides) and 3D (for proteins) NMR experiments outlined in Sections 2.3.3 and 2.3.5 are used along with a combination of the sequential [160] and main chain directed assignment strategies [178].

Sequential assignment begins by determining the identity of each spin system by comparing their chemical shift signatures with published cross-peak patterns and random coil shifts [154, 179]. For small peptides, TOCSY experiments which can correlate all resonances within a spin system (Section 2.3.3) are particularly useful as they show correlations of all protons in the amino acid residue to the amide proton. COSY spectra are used in the event

of signal overlap as they show connectivities between nuclei up to three bonds apart and therefore permit ambiguity to be resolved.

After each spin system's identity has been confirmed, the position of each spin system in the sequence is determined. For smaller peptides, correlations in NOESY experiments, which denote spatially proximal protons, are used to determine sequential spin systems by the presence of cross-peaks corresponding to protons in successive residues. Particularly useful are correlations to the amide protons of residues, as shown in Figure 2.13 [161]. Alternatively, data from the HNCACB and CBCA(CO)NH spectra can be used to confirm sequential residues in the fashion outlined in Figure 2.12, even before the identity of each spin system has been determined.

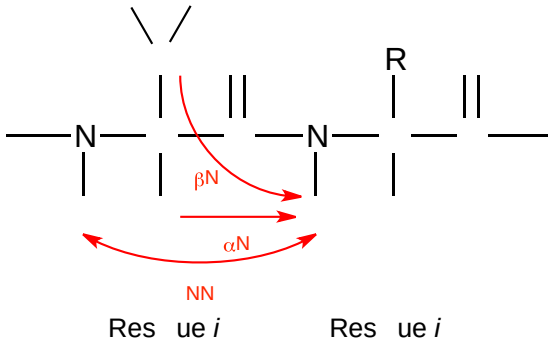


Figure 2.13: The NOE connectivities used to sequentially assign spin systems. Inter-residue NOE connectivities are denoted using the nomenclature $d_{AB(i,j)}$, which represents the distance between two protons A and B on residues i and j. Adapted from [160].

In the assignment of a larger protein using the 3D NMR spectra described (Section 2.3.5), this approach becomes more problematic as most residues have similar $^{13}C\alpha$ (typically 53-58 ppm) and $^{13}C\beta$ (typically 28-33 or 38-42 ppm) chemical shifts [180]. Consequently, only those residues with distinct and unique $^{13}C\alpha$ and $^{13}C\beta$ shifts can be tentatively assigned. These include, (i) Ser and Thr, which have $^{13}C\beta$ chemical shifts more downfield than those of their $^{13}C\alpha$, (ii) Val, Ile and Pro, which typically have $^{13}C\alpha$ chemical shifts greater than 60 ppm, (iii) Ala, with its characteristically upfield $^{13}C\beta$ chemical shift, and, (iv) Gly, which does not have a $^{13}C\beta$, and its $^{13}C\alpha$ chemical shift appears at about 45 ppm [31]. Consequently, the main chain directed assignment strategy is often used.

The main chain directed assignment strategy [178, 181] involves the assignment of sequential spin systems before determining their amino acid identity. After the spin

systems have been ordered, they can be aligned with and assigned to the known amino acid sequence. Where there are a large number of residues with similar chemical shift signatures, this approach is particularly useful.

2.3.7 Secondary Structure Analysis using NMR Spectroscopy

After assignment of the NMR spectra, the information-rich data which they contain allow structural conclusions to be drawn about the peptide or protein. These data can also be used in computational studies to determine the 3D structure. The data analysed in structure determination is discussed in the following sections.

2.3.7.1 Secondary Chemical Shifts

In NMR experiments, electrons surrounding a nucleus create a secondary magnetic field opposing $\mathbf{B}_0$. Consequently, in regions of high electron density, the applied field experienced by the nucleus is weaker and a higher field strength is required to bring the nucleus into resonance. This is termed shielding and causes resonances to occur at different frequencies (chemical shifts) in the NMR spectrum [71]. Shielding is influenced by primary structure and also molecular conformation. Thus, individual nuclei will have different chemical shifts due to their dissimilar chemical surrounds [71, 154, 182]. Consequently, chemical shift is highly sensitive to 3D conformation and can be used to predict regions of regular secondary structure.

^{1}H, ^{13}C and ^{15}N random coil chemical shift values have been determined for the twenty common amino acids in peptides with no defined secondary structure [179]. The difference between the observed chemical shift and the equivalent random coil resonance is termed the secondary shift ($\Delta\delta$). A positive $\Delta\delta$ value indicates a downfield shift from random coil values whilst a negative $\Delta\delta$ indicates the opposite.

When $\Delta\delta$ is plotted against the amino acid sequence, regions of regular secondary structure can be exposed by observing the pattern of $\Delta\delta$ [182, 183]. It has been shown that α-helices display α^{1}H and NH resonances which are upfield ($\Delta\delta < 0$) (α^{1}H shifts in the order of 0.39 ppm) and α^{13}C resonances downfield ($\Delta\delta > 0$) from random coil

values. β-sheet structures display the converse (α^1H shifts in the order of 0.37 ppm) [182, 184, 185]. Flexible or unstructured regions exhibit little deviation from random coil shifts ($\Delta\delta \approx 0$) [182, 183]. Smoothing of α^1H and α^{13}C $\Delta\delta$s over a window of $\pm$ 2 residues averages out the effects of any local electronic influences, allowing regions of secondary structure to be more easily identified [183]. Amphipathic α-helices show a different pattern of NH $\Delta\delta$s [184, 186]. Resonances from the hydrophobic face of the helix have $\Delta\delta > 0$, whereas those of the hydrophilic face demonstrate $\Delta\delta < 0.42$ [184]. Thus, a periodic variation is seen in these values over 3 to 4 residues.

2.3.7.2 NOE Connectivities

The NOE is the result of mutual cross-relaxation between dipolar coupled spins. Dipolar coupling is dependent on the distance between the two spins and will only be observed if the nuclei are close in space (typically less than 5 Å) [160]. Inter-residue NOEs are described using a standard notation $d_{AB(i,j)}$ which represents the distance between two protons A and B on residues i and j.

The most common NOEs are those observed between protons of the same amino acid residue and between protons of successive residues. Some medium range (2 - 4 residues apart) inter-proton distances are characteristically short in regions of regular secondary structure and can therefore be observed in the NOESY spectrum [160, 170, 187]. The pattern of NOEs in the spectrum can therefore be an indication of regular secondary structure, as summarised graphically in Figure 2.14 [149, 160, 161, 170].

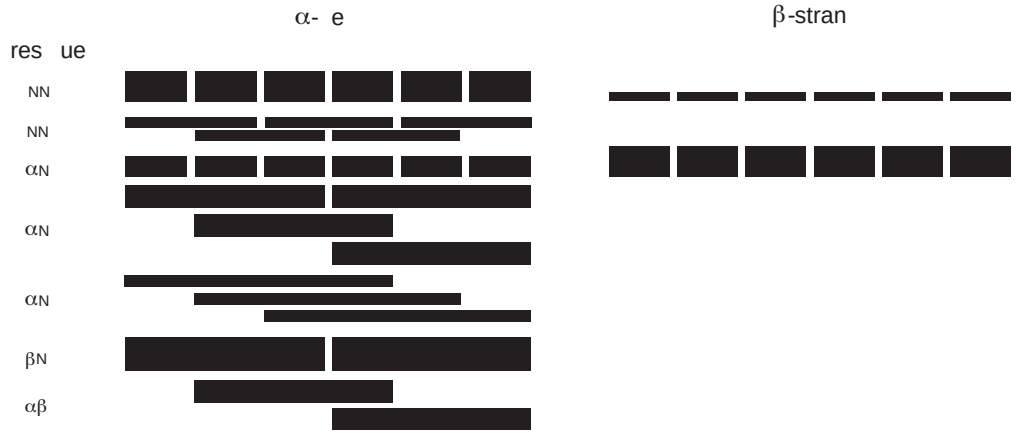


Figure 2.14: Standard pattern of NOEs for ideal α-helices and β-strands. The thickness of the bands indicates the relative intensity of the NOE signal.

2.3.7.3 Coupling Constants

Scalar coupling is observed between nuclei that are separated by a small number of covalent bonds. In the 1D NMR spectrum, the coupled signal will appear as a multiplet with the number of lines equal to (2I + 1) (where I is the spin quantum number of the neighbouring nucleus), and the spacing between the lines will equal the coupling constant, J [71, 160, 188].

The peptide backbone dihedral angles ϕ and ψ are defined by the atoms shown in Figure 2.15 [154]. The size of the vicinal coupling constant between NH and α^1H ($^3J_{NH\alpha H}$) is related to the ϕ angle by the Karplus relationship below (Equation 2.11).

$$^3J_{NH\alpha H} = 6.4cos^2(\phi - 60°) - 1.4cos(\phi - 60°) + 1.9 \quad (2.11)$$

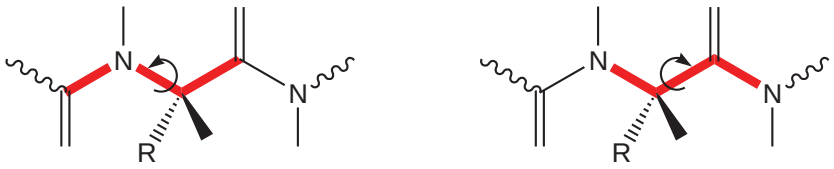


Figure 2.15: The (A) phi (ϕ) and (B) psi (ψ) backbone dihedral angles (shown in red).

The well-defined backbone dihedral angles present in elements of regular secondary structure mean that the magnitude of $^3J_{NH\alpha H}$ can be correlated with the presence of any local structural elements [149, 154]. Generally, $^3J_{NH\alpha H}$ values of less than 6 Hz are associated with α-helices and values greater than 8 Hz are associated with β-sheet structures [160]. $^3J_{NH\alpha H}$ values between 6 and 8 Hz are seen in random structure. These values are perturbed by the presence of Pro residues [154].

Coupling constants can be measured from high-resolution 1D ^{1}H-NMR spectra [154]. In larger peptides or proteins where signal overlap occurs, they can also be measured from phase sensitive COSY spectra by measuring the distance between the peak and trough of the anti-phase signal of the NH-α^1H cross-peak [189]. However, coupling constants cannot be reliably determined by this method if the $^3J_{NH\alpha H}$ value is less than the signal's line width [154, 160].

2.3.8 Structure Calculations

In the calculation of 3D structure, an initial ensemble of structures, typically with random conformation is generated. Distance and dihedral angle restraints (Sections 2.3.8.1 and 2.3.8.3) are applied from the experimental data, and calculations are performed yielding a series of structures which are consistent with the NMR data. The quality of the structure ensemble obtained is then analysed to ensure it is a reasonable representation of the solution structure. The structure calculations in this body of work were performed using the program ARIA (Ambiguous Restraints for Iterative Assignment, version 1.2) [190, 191] implemented with CNS (Crystallography and NMR System, version 1.1) [192]. The following sections describe the structure calculation process in further detail.

2.3.8.1 Distance Restraints

It is possible to estimate interproton distance (r) between two protons (i and j) from the NOE signal intensity (I), according to the following equation;

$$I_{ij} = \alpha r_{ij}^{-6} \tag{2.12}$$

where α is a calibration factor dependent on the system being investigated. α is calculated by ARIA in an iterative calculation process [193]. However, a number of other chemical factors also influence NOE signal intensity, and therefore affect the measured distance restraints [154]. Because of this, ARIA uses an approach whereby a distance error (δ) is calculated using the simple equation shown below [191, 193].

$$\delta = 0.12 \times r_{ij}^{2} \tag{2.13}$$

This yields upper (U) and lower (L) bounds for the measured restraint, as shown below [191, 193]:

$$\begin{aligned} U &= r_{ij} + \delta \\ L &= r_{ij} - \delta \end{aligned} \tag{2.14}$$

These distance restraints along with the lower and upper bounds are incorporated into structure calculations ensuring that minimal violations are observed in the final structures.

2.3.8.2 Ambiguous NOEs

The probability of coincident resonances in the NMR spectrum greatly increases with the size of the molecule. This prevents the unambiguous assignment of all cross-peaks in the NOESY spectrum [194]. Ambiguity in resonance assignment greatly reduces structure quality and may be resolved by applying an iterative approach to structure calculation. Preliminary calculations are performed excluding these ambiguous NOEs, with the resultant structures analysed to determine which possible NOEs are not feasible considering other unambiguous restraints. In this way, ambiguities in which the inter-proton distances appear unrealistic are excluded from further calculations. This process of refinement continues so that the ambiguous distances are assigned to specific proton pairs, where possible [194, 195]. The successive assignment of ambiguous resonances continues until no further ambiguity can be resolved. These remaining ambiguous assignments are still utilised in structure calculation by application of sum-averaging [194, 196]. This involves the ambiguous peak being treated as a combination of each possible contributing NOE signal, thereby defining multiple restraints. Although less accurate, overall structure quality is improved as it allows information in ambiguous restraints to be utilised and not simply discarded [194].

2.3.8.3 Dihedral Angle Restraints

Equation 2.11 in Section 2.3.7.3 details the relationship between the ϕ dihedral angles and the $^{3}J_{NH\alpha H}$ coupling constant [197]. Consequently, coupling constant data can be used in addition to NOE data to provide additional restraints for structure calculations. Dihedral angles are typically restrained over ranges according to: 5 Hz $< {}^{3}J_{NH\alpha H} <$ 6 Hz, ϕ -60° ± 40°; $^{3}J_{NH\alpha H} <$ 5 Hz, ϕ -60° ± 30°; $^{3}J_{NH\alpha H} >$ 8 Hz, ϕ -120° ± 40°. ϕ angles are unrestrained when 6 Hz $< {}^{3}J_{NH\alpha H} <$ 8 Hz [198, 199]. Ranges are used for these restraints as coupling constants can be affected by local rotation about N-Cα bonds [154].

2.3.8.4 Restrained Molecular Dynamics and Simulated Annealing

ARIA employs restrained molecular dynamics (RMD) and simulated annealing (SA) calculations to find the lowest energy structures of peptides and proteins from NMR data [194, 200–204]. RMD aims to locate the minimum of a potential energy function which combines experimentally derived conformational restraints from NMR analysis with terms describing the forces acting on the atoms of the system [159, 205–207]. The RMD protocol applies the experimentally derived restraints and simulates the motion of a system of atoms according to classical laws [194, 200]. Throughout this process, the force on every atom is calculated using Newton's equation of motion as they evolve over a series of positions and velocities;

$$F_i = m_i a_i \tag{2.15}$$

where F_i is the force on atom i, m_i is the mass and a_i is the acceleration [153, 205, 207]. An alternate method can also be employed to calculate F_i [154];

$$\frac{dV}{dr_i} = m_i \frac{d^2 r_i}{dt_i^2} \tag{2.16}$$

where F_i is expressed as the derivative of the potential energy (V) with respect to the position of the atom, r_i and a_i is expressed as the second derivative of position with respect to time, t_i. Using an appropriate potential energy and the known masses m_i, Equation 2.16 is solved numerically for future positions in time, finding the coordinates and velocities of each atom in the Cartesian plane at successive time periods, t_i [154, 207]. Excess potential energy within the system is converted to kinetic energy throughout the calculations and removed from the system by coupling to a theoretical heat bath [206]. This process allows the energy minima of the system (favourable conformers) to be located.

The total potential energy function utilised in the RMD protocol is given by Equation 2.17 [149, 161, 206].

$$V_{total} = V_{covalent} + V_{repel} + V_{NOE} \tag{2.17}$$

The $V_{covalent}$ term ensures bond lengths, angles, chirality and aromatic planarity are maintained within the molecule. When deviations occur from ideal values, higher energy structures result. The potential energies of the van der Waals forces and Coulombic interactions are also incorporated into the function, given by V_{repel}. NOE and dihedral angle restraints (Sections 2.3.8.1 and 2.3.8.3), derived from NMR data, are also integrated, defined by V_{NOE}. This ensures that protons with measured distance restraints are separated by distances close to the restraint value, and that dihedral angles are close to the restraint values [149, 154, 161, 206].

The RMD procedure is combined with a SA protocol to overcome any local energy minima. SA involves the system being theoretically heated to a high temperature after which slow cooling occurs [208]. This process allows local energy minima to be overcome (due to the additional kinetic energy), allowing the global minimum of the potential energy function to be reached [161, 206]. Consequently, a more accurate representation of the structure is obtained.

2.3.8.5 Structure Quality

Once a data set has been utilised to gain an experimentally derived structure, the quality of the structure ensemble obtained must be examined to ensure it is truly representative of the solution conformation.

Firstly, restraint violations are examined to indicate if the final structures and the NMR data concur. The calculated inter-atomic distances are analysed, to ensure calculated distances agree with restraints from the NOESY experiment, along with the calculated dihedral angles, to ensure they agree with the experimentally determined dihedral restraints. Distance restraint violations should be minimal and structures with violations of less than 0.3-0.5 Å are generally considered acceptable [209, 210]. Significant restraint violations are often observed when there are inconsistencies in the input data [209, 210]. Additionally, if there are few restraints incorporated into structure calculation, there may be insufficient data to characterise the true structure. Furthermore, when there is a lack of defined structure, only a few NOEs may be observed in the spectrum. Nonetheless, random structure cannot be assumed [161, 211].

The root-mean-squared deviation (RMSD) describes the precision of the overall structure, by providing a measure of structure convergence. The RMSD measures the closeness of the individual conformers of the ensemble to the average structure. A lower RMSD indicates that the structure ensemble likely resembles the actual structure (assuming the structure is well-defined). Generally, the upper limit for the RMSD is 2 Å [211].

Angular order parameters (S) can also be calculated to describe how well-defined the dihedral angles are over the structure ensemble. The backbone dihedrals ϕ and ψ (Figure 2.15) are commonly analysed in this fashion. S values are calculated for all dihedral angles of the structure ensemble. For an angle which does not change over the ensemble, S=1, while S=0 if the dihedral angles across the structure ensemble are random, indicating a disordered conformation [212]. Dihedral angles are considered to be well-defined when $S > 0.9$ [212]. In addition to dihedral angles, the calculated structures should also be analysed to ensure that other angles, bond lengths and impropers are consistent with ideal covalent geometries.

Finally, the dihedral angles ϕ and ψ for each residue are analysed to ensure that the combination calculated is sterically allowed. This is done using a Ramachandran plot [213]. Only certain combinations of ϕ and ψ are permitted, corresponding to regions of regular secondary structure (Pro and Gly residues exhibit atypical dihedral angles so are generally excluded from this analysis) [211]. The Ramachandran plot is divided into favourable, allowed and generous regions corresponding to the ϕ/ψ combinations present in regular secondary structure. High quality structures contain many residues in the favourable and allowed regions (typically >95 % for well-defined structures) [211].

2.3.9 Solvent Systems for the Structure Determination of Peptides

The biologically important conformation of a peptide which binds membrane-embedded receptors, or is a membrane active antimicrobial, is the structure which it adopts in a lipid environment [214–216]. Additionally, many peptides are unstructured in solution, but upon binding a protein target adopt secondary structure, such as α-helices, due to the change in chemical environment. As such, solution structures of these peptides are often obtained in media that induce secondary structure if it has a propensity to

form. The organic solvent 2,2,2-trifluoroethanol (TFE) [217, 218] and micelles of dodecylphosphocholine (DPC) or sodium dodecyl sulfate (SDS) [219, 220], are examples of these types of media. Pure water cannot be used as it destabilises the intramolecular hydrogen bonding responsible for maintaining secondary structure by forming its own hydrogen bonds with the peptide [221].

This work uses the zwitterionic, amphipathic lipid DPC, which is capable of forming micelles once the critical micelle concentration (CMC) of 1 mM is exceeded [222–224]. Approximately forty lipid molecules associate to form a micelle and a number of these are displaced from the complex upon binding by the peptide [222, 224]. In micelles, the hydrophilic head groups aggregate forming the surface of the spherical structure, whilst the hydrophobic tails extend towards the centre of the complex. Micelle solutions may mimic a membrane environment better than TFE, due to their lipid nature which allows for mimicry of the lipid-water interface [225], but the large degree of membrane curvature may result in the disruption of any tertiary structure that is present [226].

2.4 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) is a well established technique used to characterise the kinetics and thermodynamics of biomolecular interactions [36, 227–230], by exploiting the fact that most chemical reactions or physical changes (such as binding events) involve a change in heat or enthalpy. ITC directly monitors heat changes, which are then used to determine binding affinity (K_a), enthalpy changes (ΔH) and stoichiometry (n). Other thermodynamic parameters (Gibbs free energy changes, ΔG, and entropy changes, ΔS) are then determined mathematically using the relationship [36];

$$\Delta G = -RTlnK_a = \Delta H - T\Delta S \tag{2.18}$$

where R is the gas constant and T is the temperature (Kelvin).

A typical isothermal titration calorimeter contains two coin shaped cells, a sample cell and a reference cell, enclosed within an adiabatic chamber (Figure 2.16). The reference cell is filled with deionised water and the sample cell contains a buffered solution of one

binding partner. The second binding partner is loaded in a syringe and is titrated into the cell. The syringe is fitted with a stirrer to ensure efficient mixing in the sample cell. The entire system, i.e. both the cells and adiabatic chamber, are maintained at an identical temperature throughout the experiment by using thermcouple circuits [229].

The thermocouple circuits detect temperature differences between the two cells and also between the cells and the adiabatic chamber. Thus, the heat change involved in a binding event can be is directly determined by measuring the amount of power required by heaters to maintain isothermal conditions (Figure 2.16) [36, 229, 231]. In the event of an endothermic binding event occurring in the sample cell, the temperature decrease is detected triggering an increase in the power supplied to the sample cell heater to maintain isothermal conditions. The opposite is true for an exothermic binding event, where the temperature increase will deactivate the power supplied to the sample cell and its heating (Figure 2.16). The initial additions of titrant will cause large endothermic or exothermic signals as all the ligand binds to the binding partner in the cell. Hence, subsequent injections result in decreased signal intensity due to saturation of the binding partner [232] (Figure 2.16).

The heat (q) evolved or absorbed as a result of each injection of ligand into the sample cell is measured using the heat compensation principle [232], which measures the power (units of μcal.sec^{-1}) necessary to maintain isothermal conditions. Thus, the integral of the area under each injection's corresponding peak is used to calculate the heat change. For a single-site binding interaction (i.e. n=1), the heat change can be correlated with a variety of thermodynamic parameters using the equation [227];

$$q = \frac{[M_T]V_0\Delta H_B K_a[L_F]}{1 + K_a[L_F]} \tag{2.19}$$

where $[M_T]$ is the concentration of the macromolecule in the sample cell, V_0 is the volume of the cell, ΔH_B is the enthalpy change of binding per mole of ligand, K_a is the binding constant, and $[L_F]$ is the concentration of free ligand. However, experimentally it is the total concentration of ligand that is known ($[L_T]$), which is the sum of $[L_F]$ and the amount of bound ligand ($[L_B]$). Consequently, more complicated relationships are used to correlate q with thermodynamic and kinetic parameters, by calculating $[L_F]$ from known, measured quantities, as detailed elsewhere [227, 229, 231, 233]. Algorithms in the supplied software

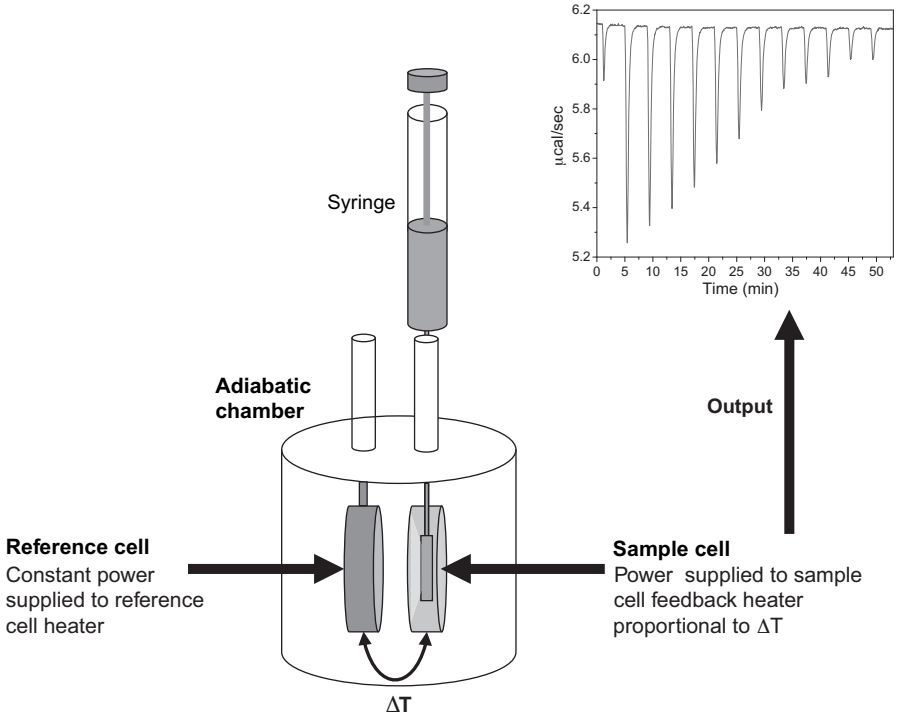


Figure 2.16: Schematic diagram of a typical ITC apparatus. The syringe plunger injects precise volumes of titrant, and acts as a stirrer during the experiment. Power is applied to the sample cell to ensure it maintains the same temperature as the reference cell, which gives the instrument signal. A typical output from an exothermic interaction is shown. Adapted from [230].

use these equations, and others, to find values of K_a and ΔH_B that are in agreement with the experimentally determined q values for each injection [233].

Repetition of the titration at different temperatures allows the change in specific heat capacity (ΔC_p) of the binding reaction to be determined, by exploiting the relationship [232, 234];

$$\Delta C_p = \frac{\delta \Delta H_B}{\delta T} \tag{2.20}$$

where $\delta\ \Delta\ H_B$ is the rate of change in ΔH_B and δ T is the rate of change of temperature. ΔC_p is a vital thermodynamic parameter as it determines the magnitude of ΔH_B and ΔS. It can also be used to give an indication of the type of interaction occurring. For example, large and negative ΔC_p values which are coupled with favourable entropy changes are often an indicator of hydrophobic binding interactions [235, 236].

2.5 Circular Dichroism Spectroscopy

Circular dichroism (CD) spectroscopy gives structural information about peptides, proteins and other molecules by exploiting the fact that they are chiral, and that all chiral molecules absorb left and right circularly polarised light to different extents [237, 238]. The technique can be utilised to investigate protein structure by measuring this difference in absorption of left and right circularly polarised light as a function of wavelength. Typically, the ranges 320-260 nm (near-UV, not discussed further) and 240-180 nm (far-UV) are used to give information about tertiary and secondary structure, respectively [239].

To measure a CD spectrum, plane polarised light is passed through an optically active substance [32, 238–240]. This light is considered to be comprised of equal parts left and right circularly polarised light (Figure 2.17A). If one of these components is absorbed more than the other due to the chiral nature of the sample, the light emerging will be elliptically polarised (Figure 2.17B). Additionally, due to differences in their refractive indices, the two components of the light travel through the sample at different velocities. Consequently, the plane polarised light is rotated by an angle α, an effect termed optical rotation (Figure 2.17B).

The ellipticity (θ, measured in degrees) can be measured by a CD spectropolarimeter by determining the difference (ΔA) between the absorbance of the left (A_L) and right (A_R) circularly polarised light [32]. This information is converted to θ, using the relationship below [32].

$$\theta = 32.98 \Delta A \tag{2.21}$$

Alternatively, θ can be measured directly as the arctangent of the ratio of minor to major axes of the ellipse (Figure 2.17C) [32].

Generally, CD data are reported as mean residue ellipticity (θ_{MRW}, units deg.cm^2.dmol^{-1}), to correct for any differences in concentration, path length, and the number of residues in the peptide/protein being analysed, using the equation [32];

$$\theta_{MRW} = \frac{\theta \times M \times 10}{l \times c \times N_r - 1} \tag{2.22}$$

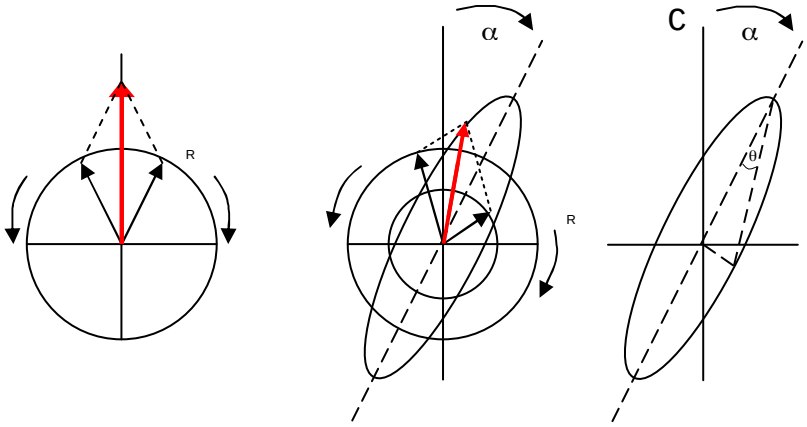


Figure 2.17: (A) The electric field vector (**E**) of linearly polarised light can be represented by left ($\mathbf{E}_L$) and right ($\mathbf{E}_R$) circularly polarised components. (B) Passing through an optically active sample results in $\mathbf{E}_L$ and $\mathbf{E}_R$ being absorbed differently. Adding the two remaining components gives the resultant electric field vector **E**, which traces out an ellipse that has been rotated through an angle α. (C) The ellipticity θ is the arctangent of the ratio of the major and minor axes of the ellipse. Adapted from [32, 239].

where M is the molecular weight (g.mol^{-1}), l is the path length (cm), c is the concentration (g.mL^{-1}), N_R is the number of residues and θ is the measured ellipticity (degrees). In this way, the normalised CD spectra for a given sample can be quantitatively and qualitatively compared.

In the far-UV region (180-240 nm), the spectra of proteins/peptides are dominated by transitions of carbonyl group in the peptide bond. There is a weak but broad n-π* transition at about 210 nm, and an intense π-π* transition at approximately 190 nm [239]. Other chromophores, including aromatic side chains, disulfide bonds and the terminal carboxylic acid group, do not significantly affect the far-UV spectrum [32].

Regular secondary structural elements give characteristic far-UV CD spectra. Representative spectra for these secondary structural elements are displayed in Figure 2.18. α-Helices give rise to a curve with a maximum at approximately 192 nm (π-π* transitions) and two minima at 208 and 222 nm (n-π* transitions) [238]. β-Sheet structures can be recognised by their characteristic spectra with a maximum at 198 nm and a minimum at 215 nm (n-π* and π-π* transitions, respectively), but these can vary [238]. Finally, random coil structures give rise to a minimum at 195 nm (n-π* transitions) [238]. Proteins and peptides that contain multiple secondary structural elements give spectra that are a combination of these representative spectra, allowing computational

deconvolution to be performed to estimate the amount of each structural component in the protein, if required [241].

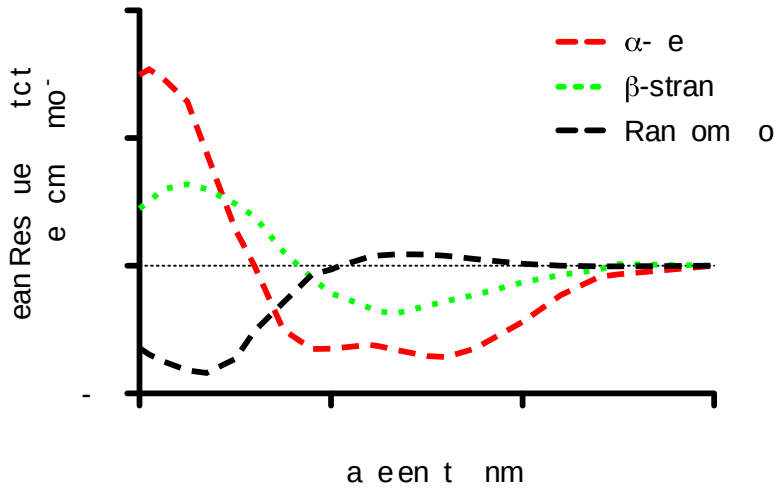


Figure 2.18: Representative far-UV CD spectra associated with various types of secondary structure. Data are from [242].

Chapter 3

Negative Ion Mass Spectrometry Amenable Chemical Cross-Linking Reagents

3.1 Introduction

Chemical cross-linking (CX) has been used for decades in the study of protein structure and interactions. It involves the addition of a reactive CX reagent to a native protein or protein assembly that covalently traps its conformational state. In the early years of CX, initial studies were conducted in tandem with gel electrophoresis to identify protein binding partners by their mass [243, 244]. However, it was the development of peptide MS techniques [39, 68] that provided the catalyst to enhance and refine CX methodologies. This led to the birth of CX-MS, allowing the cross-linked proteins to be digested with proteases, efficiently identified and the chemical modifications involved tracked with residue-level specificity (Section 3.1.1, reviewed in [245–252]). By this technique, the sites of the cross-links are used to determine maximum distances between amino acid side chains and develop proximity maps [253]. From these data it is possible to draw conclusions about a structure of a protein and protein assembly with otherwise limited structural information [254]. The topologies of numerous protein complexes, both small and large, have been probed using CX-MS [37].

CX-MS exploits several advantages associated with proteomics-type MS analysis [245, 248, 255, 256]. Firstly, in a CX-MS analysis, peptides produced by proteolytic digestion are analysed. Consequently, there is theoretically no limit to the size of the proteins or

protein assemblies that can be investigated. Additionally, only small amounts of protein are required, and analysis is relatively fast. Moreover, the studies are conducted in solution, unlike X-ray crystallography, and protein assemblies, such as membrane proteins, can be investigated which were often considered otherwise intractable.

Broad application of the CX-MS approach has been hampered primarily by analytical challenges [245, 247–249, 255, 256]. Notably, it is difficult to assign the identity of cross-linked peptides by mass alone, and detection and identification of the intermolecular cross-linked peptides, which provide the most structural information, remains difficult due to their relatively low abundance.

Furthermore, peptide identification and location of the specific amino acids within the protein that contain the cross-linker modification is hindered by often complex MS fragmentation behaviour. In addition, there is a need for the development of automated analysis systems to examine the large quantity of data produced in CX-MS experiments [247].

Consequently, the development of new methodologies for the identification of cross-linked peptides is of fundamental importance. Some of these issues can be in part addressed by CX-MS reagents specifically designed for the desired downstream detection and analysis application (Section 3.1.2).

3.1.1 Cross-Linking Strategy

There are many CX reagents available commercially, and many others have been reported in the literature (Section 3.1.2). Despite their large number, the reagents mostly comprise the same fundamental building blocks of a spacer arm (the bridge formed between the two cross-linked sites after the reaction) separating two reactive groups [245, 249, 256–258]. This type of reagent is termed bifunctional (Figure 3.1A). When a spacer arm is not present, the reagent is termed 'zero-length,' as it covalently joins proximal residues (<3 Å) [245, 256, 258].

All bottom-up CX-MS studies share a common workflow, with minor deviations in the MS analysis protocol, depending on the CX-MS reagent used (Section 3.1.2) [245–252]. Briefly, the protein or protein complex is treated in its native form with the CX reagent, covalently trapping its conformational state. Optimisation of reaction conditions, including concentration, reaction times, and buffers must be performed to ensure a high-yield of cross-linked material without the disruption of native structure. The cross-linked protein is digested with a protease, usually trypsin, resulting in a complex mixture of both modified (by the CX-MS reagent) and unmodified proteolytic peptides. These peptides are analysed by MS, often LC-MS, to identify the peptide products and determine the identity of any cross-linked adducts (Figure 3.1A).

The aim of any cross-linking reaction is to form covalent bonds between proximal residues, however, other reaction products are possible either because both reactive groups in the linker may not come into contact with an appropriate residue or the reactive groups may by deactivated prior to reacting (e.g. by hydrolysis). Consequently, the cross-linking reaction may afford several products, namely dead-end (type 0), intramolecular (type 1) and intermolecular (type 2) cross-linked peptides, or higher order adducts (Figure 3.1B) [248, 253, 259].

Each adduct type can provide some structural information. For example, dead-end (type 0) adducts yield information about solvent accessibility, intramolecular (type 1) cross-links may give information about local structural elements i.e. secondary and tertiary structure, whilst intermolecular (type 2) cross-links can give distance information that may be used to define binding interfaces [37, 247, 248, 260]. Together, this information can be used to

generate distance constraints, allowing the structure of a protein or protein complex to be approximated, as a pair of cross-linked residues must be within a certain distance of each other, as determined by the spacer arm length.

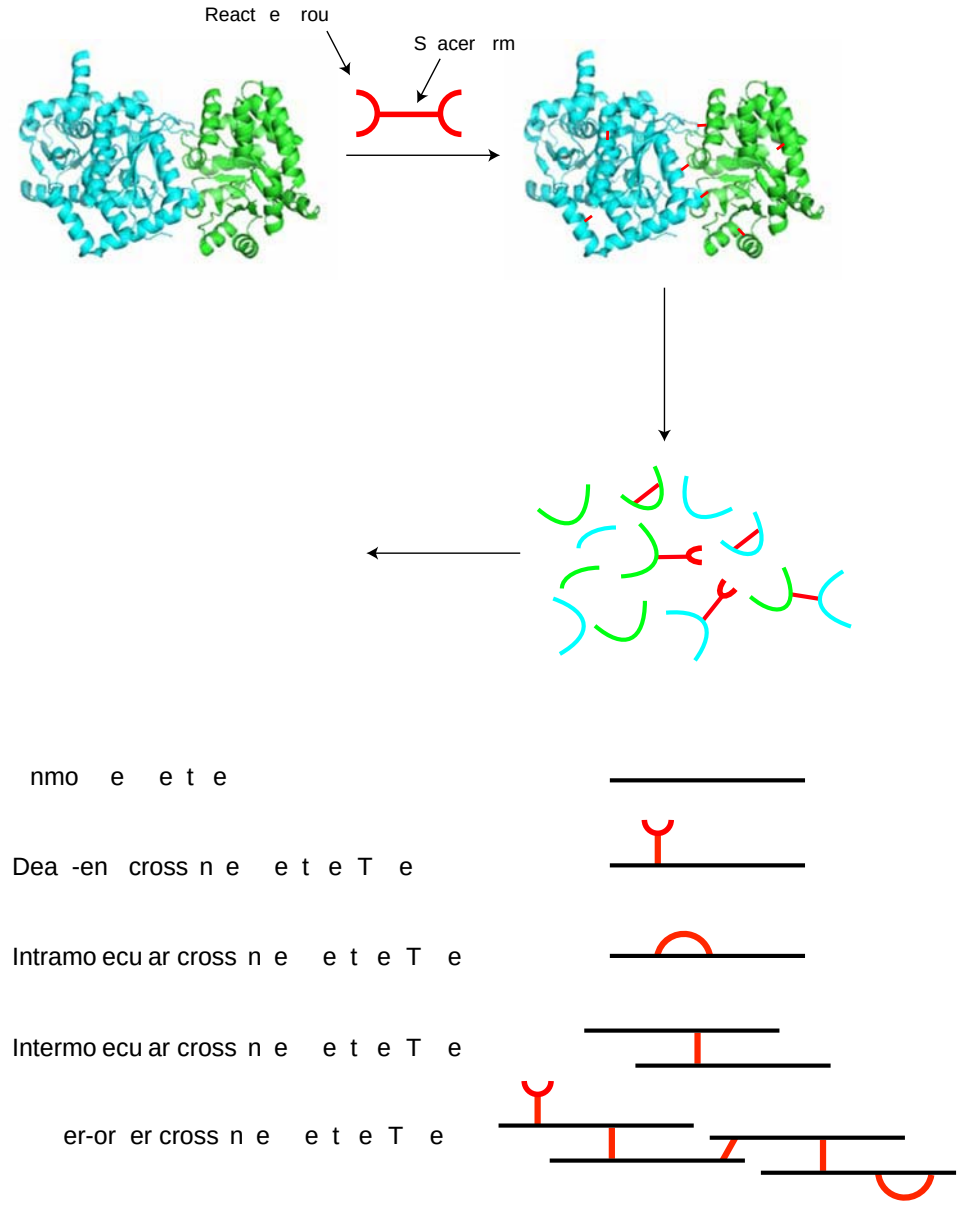


Figure 3.1: (A) The CX-MS workflow showing the general structure of a CX-MS reagent (PDB:1KFK [261]). (B) Peptides that can be observed after cross-linking and proteolytic digestion [248].

3.1.2 Cross-Linking Reagents

As previously described (Section 3.1.1), the most commonly used CX-MS reagents are bifunctional and contain two reactive groups (Section 3.1.2.1) separated by a spacer arm of defined length. Even though CX-MS reagents often contain the same basic structural elements, additional functionality is often included in their design to enable successful downstream MS analysis (Section 3.1.2.1).

3.1.2.1 Cross-Linker Reactivity and Design

There are numerous CX reagents available commercially and others have been described in the literature. However, the reactions of all these CX reagents with proteins are based on a small number of organic chemical reactions [245, 249, 255–257].

The most commonly encountered reactive group amongst the many studied CX reagents is the *N*-hydroxysuccinimide (NHS) ester (Figure 3.2). These activated carboxylic acid moieties are acylating agents [245, 249, 253, 255–258], and have been utilised as CX reagents for over 35 years [262, 263]. Most NHS ester CX reagents are insoluble in buffered aqueous solutions, so are dissolved in an organic solvent before being diluted into a solution containing the protein(s) of interest [245]. Alternatively, water-soluble sulfonated derivatives have been developed to avoid the use of organic co-solvents [264]. The NHS moiety is highly labile, with a half life of several hours at physiological conditions, and reacts with nucleophilic sites on proteins resulting in the release of NHS (Figure 3.2) [245, 256–258]. These reactions afford stable amide and imide bonds when reacting with primary or secondary amines respectively, including those at the protein N-terminus, and on Lys side chains. It has also been shown that NHS esters are susceptible to nucleophilic attack by the hydroxyl groups of Ser, Thr and Tyr residues to form stable ester linkages [265, 266], although these processes are less common and often dependent on reaction conditions.

There are many alternate amine reactive groups that are incorporated into CX reagents, including aldehydes, imidoesters, carbodiimides, and other activated esters, however, these

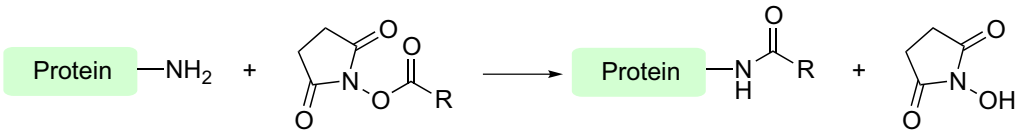


Figure 3.2: The general reaction of an NHS-ester with a free amine on a protein.

are less commonly used due to their decreased cross-linking efficiency [245, 249, 253, 255–258].

Other residue specific reactive groups may also be incorporated into CX reagents, such as the maleimide functionality which selectively reacts with the sulfhydryl moiety of Cys residues [245, 249, 253, 255–258]. Alternatively, aryl azides, diazirines or benzophenones may be incorporated into cross-linker designs, which form reactive species upon photoactivation that non-specifically insert into chemical bonds [245, 249, 253, 255–258].

Finally, zero-length oxidative cross-linking techniques have been developed that are mediated by high-valent Ru metal complexes. This technique is called photo-induced cross-linking of unmodified proteins (PICUP) [267] and has been particularly useful in the study of oligomerisation processes that precede disease related fibril formation [268, 269].

Homobifunctional CX reagents contain two of the same reactive groups, and allow cross-linking between identical functional groups in a protein. However, they can produce a large quantity of unwanted side products and can potentially form high molecular weight aggregates in solution [245, 256, 258]. Consequently, reaction conditions must be optimised to ensure the 3D structure is not perturbed whilst obtaining a high yield of cross-linked peptides.

Heterobifunctional CX reagents contain two different reactive groups and therefore target two different functional groups on the protein. These reagents are often used in a two step cross-linking protocol, where one protein is activated with the cross-linker, excess reagent is removed, and the binding partner added. This reduces the number of unwanted cross-linked peptides that are produced, however it is not a suitable approach for CX reagents containing extremely labile reactive groups [245, 256, 258]. Often, heterobifunctional reagents contain both NHS ester and photoreactive groups, thus the

NHS ester can be allowed to react with the protein and excess cross-linker removed before photoactivation [245, 256, 258].

When three functionalities are present in a cross-linker it is termed trifunctional. Often these comprise two different reactive groups in addition to a third substituent that is used for affinity purification, allowing enrichment of the cross-linked peptides and consequently easier detection [270–277].

3.1.2.2 Identification of Cross-Linked Peptides

Detection and identification of cross-linked peptides by MS remains the primary analytical challenge associated with the technique. These issues arise primarily due to the complexity of the peptide mixture, the relatively low abundance of cross-linked peptides and their often complex fragmentation behaviour. Consequently, various approaches have been developed to aid this identification process. These include the development of trifunctional CX-MS reagents (Section 3.1.2.1) and the use of various isotope labelling strategies [278–282]. These methods, however, often yield complex, difficult to analyse MS/MS data, as the intermolecular cross-links consist of two peptides covalently joined and product ions from both peptides overlap in the spectrum [283].

To address issues with sequencing by MS/MS, CX-MS reagents have been developed which can be cleaved either by CID [275, 283–288], or chemically prior to MS analysis [274, 279]. CID cleavable cross-linkers may afford products corresponding to a reporter ion of known m/z [275, 285], or cleavage of the cross-linker at a defined position may give product ions corresponding to each of the two linked peptides which can each be sequenced independently by MS^n [284, 286, 287, 289, 290]. In certain designs, both cleavage strategies are used [275, 291]. An example of each type of CX-MS reagent is discussed below.

Dithiobis[succinimidyl propionate] (DSP, also known as Lomant's reagent) (Figure 3.3) has been shown to fragment about the disulfide when subjected to CID [288]. As a result, disulfide cleavages are observed in addition to ordinary backbone fragmentations upon collisional activation. The fragmentation produces two ions with a characteristic mass

spacing (66 Da) that can be used for cross-link identification. This is not an ideal approach as high collision energies are required to effect this cleavage and consequently sequencing is hampered by the complex fragmentation behaviour of the cross-linked peptides. In fact, in this previously published study, sequencing was achieved by first chemically reducing the disulfide prior to a second MS/MS analysis [288].

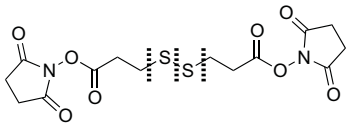


Figure 3.3: The cleavable CX-MS reagent DSP.

It is anticipated that a negative ion approach can be applied for DSP as fragmentation of a disulfide differs in the negative ion mode, which may prove more useful to identify intermolecular cross-links by DSP. This is the subject of investigation in this chapter (Section 3.1.3).

The CX-MS reagent disuccinimidyl sulfoxide (DSSO) (Figure 3.4) contains two labile bonds that cleave prior to the backbone, on either side of the sulfoxide group [284]. These cleavages are favourable and occur at low energies, thus releasing the two peptide halves that are involved in the cross-link. This allows each peptide to be sequenced individually by MS^n, as cleavage of the C-S bond occurs at collision energies lower than those required to fragment the peptide backbone. The ions produced are abundant and easily identifiable.

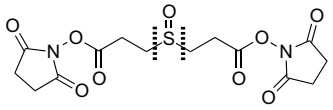


Figure 3.4: The cleavable CX-MS reagent DSSO.

Finally, an example of an MS cleavable CX-MS reagent which ejects a reporter ion of known m/z, is shown in Figure 3.5 [275]. This particular compound contains two NHS-ester reactive groups separated by a specially designed spacer arm based around two RINK moieties, and is referred to as the protein interaction reporter (PIR) [292]. After conjugation and proteolysis of the cross-linked proteins, mass spectrometric analysis of the cross-linked peptides is conducted, and as the reagent contains two labile bonds which can be cleaved by CID (Figure 3.5). Intermolecular cross-linked peptides are easily recognised

as they eject a central reporter group (m/z 711, m/z 828 for a dead-end peptide). This process also separates the two peptide halves so they can be individually sequenced.

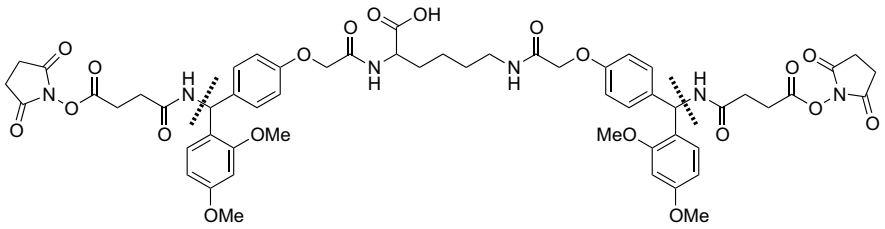


Figure 3.5: The cleavable PIR CX-MS reagent. The MS cleavable bonds are indicated [275].

3.1.3 Negative Ion Fragmentations of the Natural Cystine Disulfide

The negative ion MS cleavages of underivatized peptides have been well studied and are outlined in Section 2.2.6.2. It has been shown that CID of symmetrical intermolecular disulfide bonded peptides in the negative ion mode results in facile fragmentation, affording a maximum of four product ions (denoted A to D), corresponding to cleavages at or adjacent to the disulfide bond (Figure 3.6, Scheme 3.1) [140, 293, 294]. These cleavages are effected by either an enolate anion or by an anion situated directly adjacent to the disulfide, and are amongst the most energetically favourable of all the negative ion peptide cleavages studied [294]. Consequently, these fragmentations occur in preference to backbone fragmentation at low collision energies.

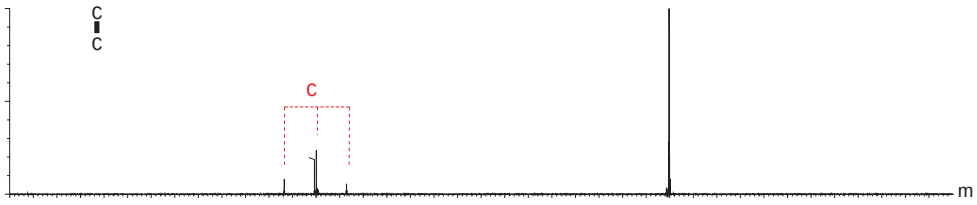
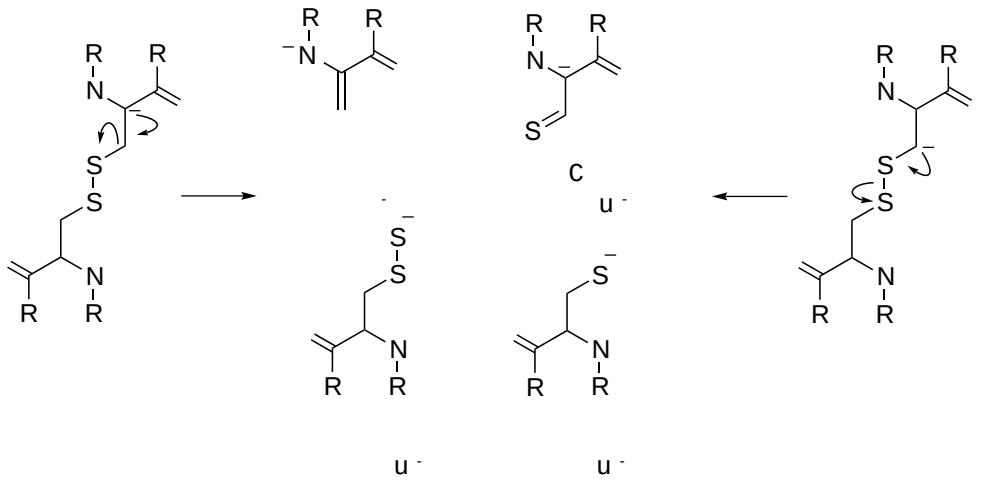


Figure 3.6: Low energy MS/MS spectrum of intermolecular cystine disulfide linked Ac-AACA ($[M\text{-}H]^-$ m/z 749). Product ions A to D are observed in the spectrum (Scheme 3.1). Q-Tof 2 mass spectrometer.

Proteomics studies generally shy away from using negative ion MS for peptide sequencing, as collisional activation of peptide anions generally produces dominant side chain induced fragmentations (Section 2.2.6.2) [136, 140]. Conversely, in the positive ion mode, collisional activation yields predominantly characteristic backbone cleavage cations, which



Scheme 3.1: Characteristic fragmentations of the natural cystine disulfide [294]. Masses of the ions are indicated relative to fragment anion A (denoted by the arbitrary mass, M). Structures of product anions B and D can be rationalized directly from the proposed reaction mechanisms, whilst anions A and C are formed by charge transfer via an ion-neutral complex.

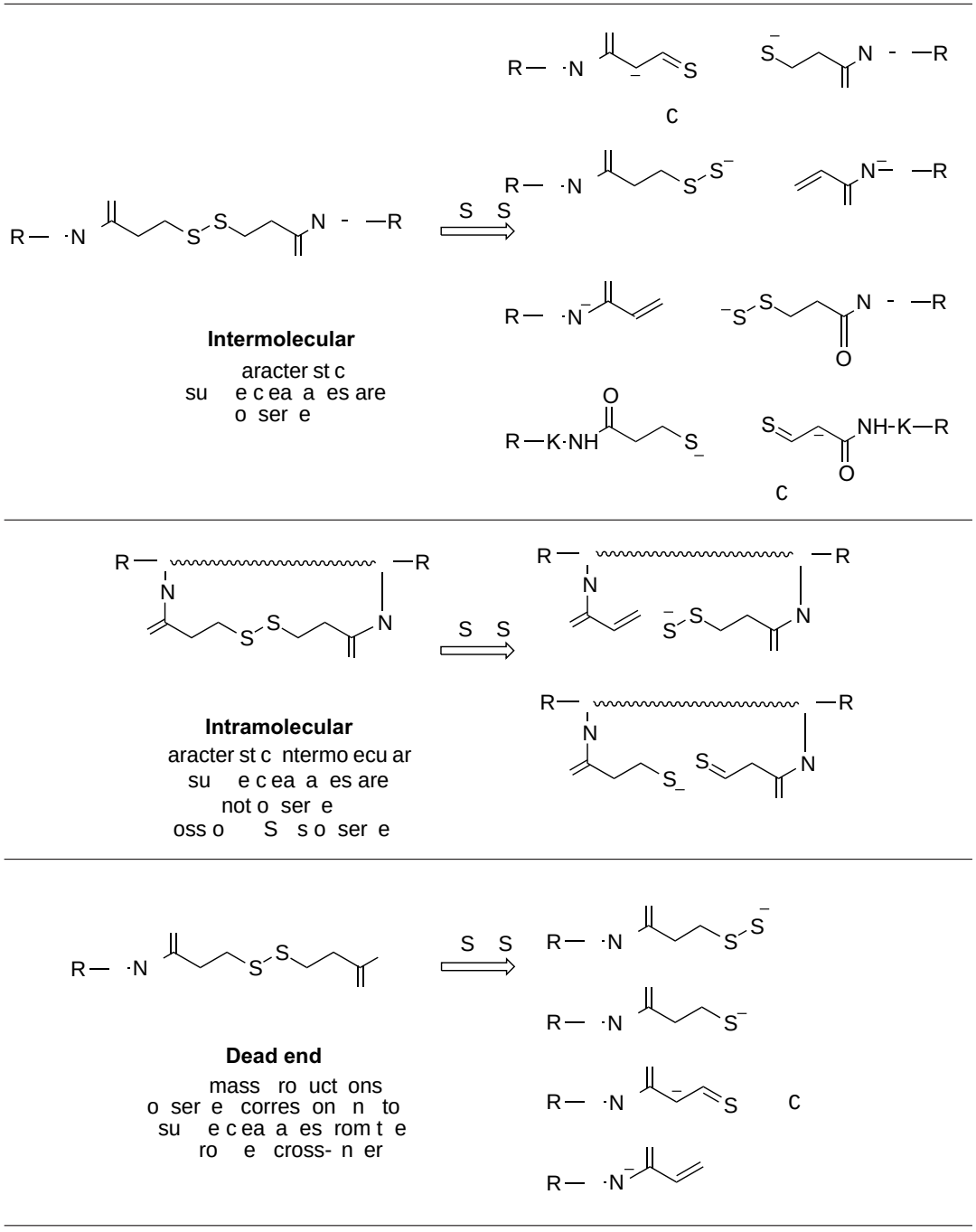
are well suited to peptide sequencing. Despite this, similar information can be obtained from both ionisation modes, and notably, it has been shown that negative ion MS is a superior to positive ion MS in the mapping of disulfide bonds and other post-translational modification in peptides and proteins [293–296].

It is proposed that the facile cleavages of the cystine disulfide (Scheme 3.1) could form the basis of a negative ion cleavable cross-linking approach. Consequently a CX-MS reagent with similar structural features to the natural cystine is required, such as the previously discussed DSP (Figure 3.3), which contains a core structure analogous to the natural cystine. Consequently, it should fragment in a similar fashion (Scheme 3.1), with both the initiating enolate and -$^{-}$CH-S-S- anion sites situated on the linker. Exploiting the facile cleavages in this approach is ideal as breakage of the cross-link can be initiated by up to four anions (two on each side of the disulfide), yielding four ions for each peptide involved in the cross-link. These processes should occur at collision energies lower than those required to fragment the peptide backbone, be easily recognizable by the characteristic product ions produced (even when all four possible products are not observed), and allow for the exact location of the cross-linking site to be identified by standard negative ion sequencing of the resultant anions (by MS^3), making data analysis reliable and fast.

This reagent has been chosen for this study as it has well-established cross-linking

characteristics. Additionally, MS/MS analyses of this reagent in the positive mode have been utilized to identify peptide cross-links [288] (Section 3.1.2.2) and it is intended to demonstrate that negative ion MS is superior to this positive ion approach.

A summary of the fragmentations expected at low collision energies for the cross-linked adducts of DSP shown in Scheme 3.2. In brief, intermolecular cross-links are expected to fragment like the natural disulfide, with each peptide of the dimer yielding up to four product ions upon CID. Intramolecular cross-links cannot fragment in this manner, as disulfide cleavage will not yield a product ion of changed mass. Instead, characteristic loss of H_2S_2 is expected [293], in addition to side chain induced backbone cleavages. Dead-end adducts, where one end of the cross-linker reacts with the protein and the other end is hydrolysed, will give low mass losses upon collisional activation, corresponding to the facile cleavages of the hydrolysed disulfide cross-linker. To differentiate the cleavages of the cross-linker from those of the natural cystine disulfide (Scheme 3.1), ions from the cross-linked adducts analogous to those of the natural cystine are labelled according to the nomenclature in Scheme 3.2.



Scheme 3.2: Summary of the expected fragmentations for DSP cross-linked peptides in the negative ion mode.

3.2 Aims

In order to utilise cross-linking reagents that contain CID labile groups to aid in cross-link identification, it is imperative that the diagnostic fragmentation pathways of the cross-linked products are understood. Consequently, one aim of this chapter was to investigate the negative ion MS fragmentations of several disulfide-containing CX-MS reagents, to ensure that their characteristic fragmentations enable the reliable identification of cross-linked adducts.

Additionally, there are many other side chain specific low energy fragmentations unique to the negative ion mode, that may compete with CX-MS fragmentation. Consequently, an additional aim of this work was to perform a systematic study to identify any of these processes that may interfere with cross-link identification in a peptide adduct of unknown sequence.

The combination of these two studies is necessary to validate the use of negative ion MS as a suitable analytical technique to identify cross-linked adducts and map protein interactions with residue level specificity.

3.3 Results

3.3.1 Cross-Linking of Ac-IR7 with DSP

To experimentally ascertain the low-energy CID fragmentation of intermolecular cross-links introduced by DSP, cross-linking reactions were performed on a model peptide system named Ac-IR7 (Ac-IEAEKGR, 843.94 g.mol^{-1}), under solution conditions known to generate high yields of the cross-linked adducts. This peptide contains only one free amine group that is available for cross-linking, meaning only one intermolecular cross-linked product is possible. Whilst this is not a biologically relevant system, it is an established model to validate NHS-ester containing cross-linking strategies and to define the characteristic fragmentations of the resultant adducts [284, 297]. As expected, following the cross-linking reaction, intermolecular cross-linked peptides, dead-end products and unreacted peptide were present. Intermolecular cross-linked peptides were identified by the expected mass (m/z 1860), and MS/MS data were recorded at low collision energy (Figure 3.7).

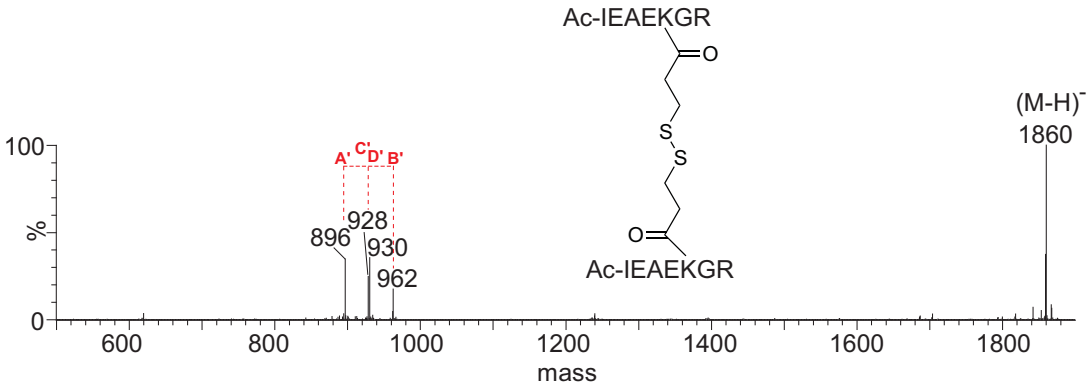
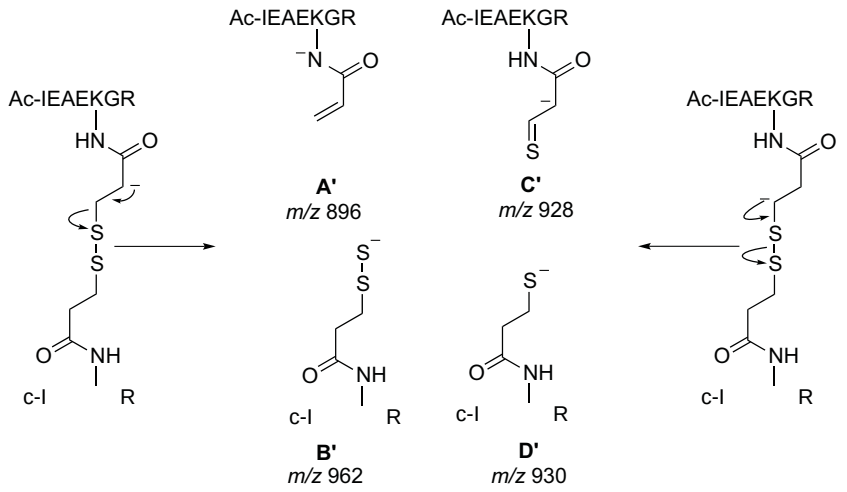


Figure 3.7: Low energy MS/MS of DSP cross-linked Ac-IR7 ([M-H]$^-$ m/z 1860) results in the production of four major product ions corresponding to cleavage about the disulfide bond (Structures of product ions A′ to D′, as indicated, are described in Scheme 3.3).

It can be seen that intermolecular cross-links introduced by DSP fragment in an analogous fashion to the natural cystine disulfide (Scheme 3.1), with all four possible product ions observed in this case (Figure 3.7), as described in Schemes 3.2 and 3.3. The four characteristic product ions, A′ to D′, are easily recognizable in this instance, allowing rapid identification of the intermolecular cross-linked peptides.



Scheme 3.3: Structures of Ac-IR7-DSP adducts. Fragment anions A′ to D′ correspond to anions A-D in Scheme 3.1.

For comparison, the positive ion electrospray MS/MS data are shown in Figure 3.8. Two peaks with a 66 Da mass difference are observed in this spectrum as a result of fragmentation about the disulfide bond. This fragmentation has been previously been used to identify DSP cross-links using positive ion MALDI MS [288]. However, these signals are difficult to observe since they occur in addition to the abundant backbone fragmentations which dominate the spectrum.

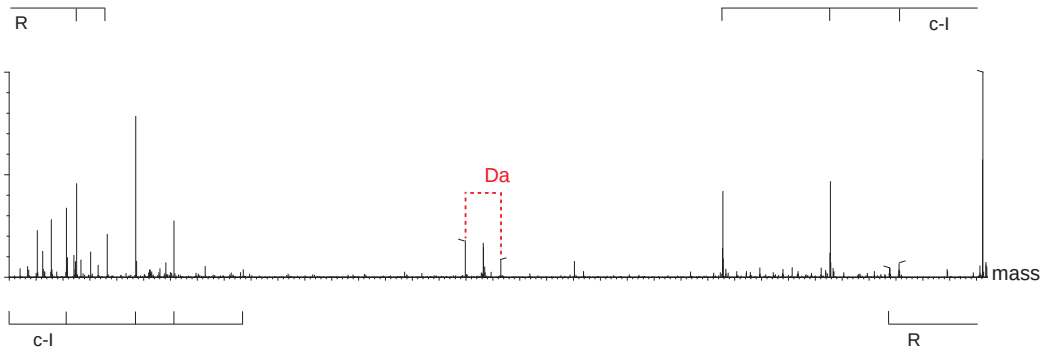


Figure 3.8: Positive ion MS/MS spectrum of DSP cross-linked Ac-IR7 (*m/z* 1862) is more complex than the corresponding negative ion mode spectrum. Two peaks with a 66 Da mass difference are observed as a result of fragmentation about the disulfide bond. However, these signals are easily overlooked due to the abundant backbone fragmentations which dominate the spectrum.

It is possible to sequence the intermolecular cross-linked peptides on the Q-ToF 2 by performing 'pseudo MS^3 experiments' i.e. MS/MS experiments on product ions created by capillary-skimmer dissociation in the ion source. Characteristic modifications of the lysine side chains (where the linker had bound) were observed in the spectrum, which would allow for the exact location of cross-linking sites to be determined in protein systems of unknown structure. Analysis of the MS^3 data from the lowest mass DSP adduct (*m/z* 896) (A′ in Figure 3.7 and Scheme 3.3) gave data that allowed easy interpretation and sequencing by the normal backbone cleavages in the negative ion mode (Figure 3.9). It is also possible to sequence the peptide from the higher mass adducts B′, C′ and D′ (Figure 3.7 and Scheme 3.3). As an example, the 'pseudo MS^3' spectrum for the B′ adduct is shown in Figure 3.10. In all cases, the spectra are dominated by side chain induced backbone cleavages from Glu (the γ and δ cleavages), in addition to the normal backbone fragmentations (the α and β cleavages) [136]. Prominent fragmentations to yield lower mass cross-linker adducts, or complete removal of the cross-linker moiety is also observed (Figure 3.10).

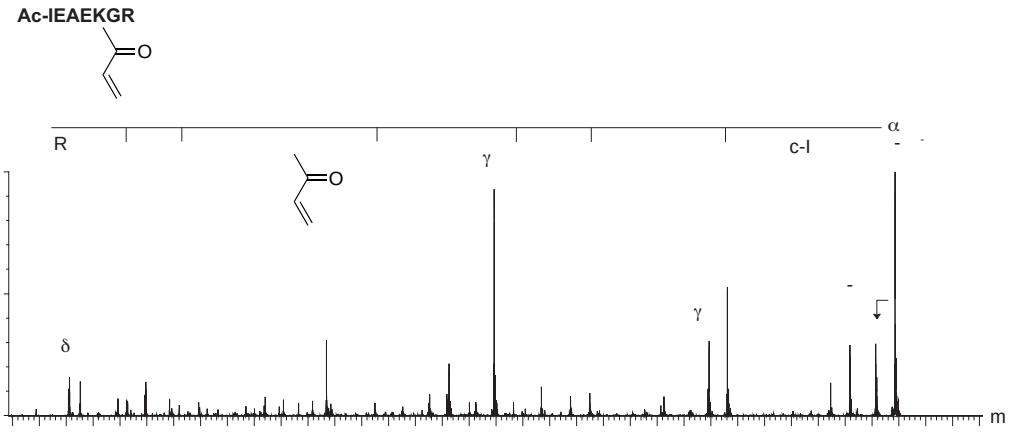


Figure 3.9: MS/MS spectrum of the A′ product ion formed by capillary-skimmer dissociation (*m/z* 896) from DSP cross-linked AcIR7 gives peptide sequence information and cross-link location. Spectrum is magnified as follows; 50 - 880 (x3). Standard backbone cleavages as well as side chain induced backbone cleavages from Glu (γ and δ cleavages) are observed as indicated.

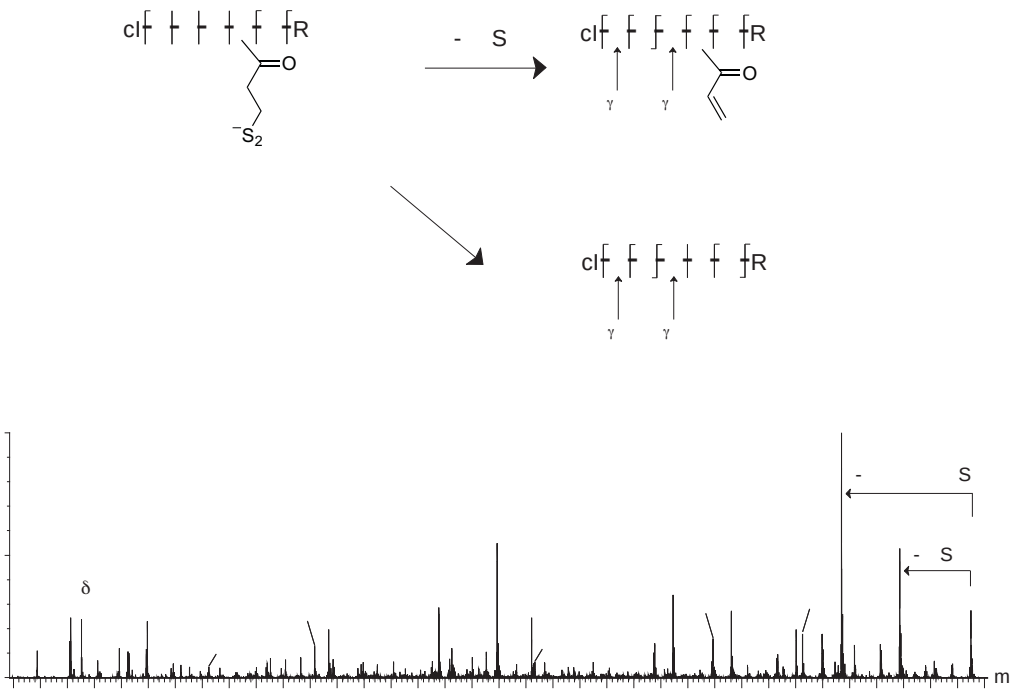


Figure 3.10: MS/MS spectrum of the B′ product ion formed by capillary-skimmer dissociation (*m/z* 962) from DSP cross-linked AcIR7 gives peptide sequence and cross-link location. The scheme above the spectrum shows the prominent fragmentations to yield lower mass cross-linker adducts, or complete removal of the cross-linker moiety. Standard backbone cleavages as well as side chain induced backbone cleavages from Glu (γ and δ cleavages) are observed as indicated.

3.3.2 Cross-Linking of Ubiquitin

To further validate DSP as a negative ion MS cleavable CX reagent, the model protein ubiquitin was cross-linked. Following tryptic digestion, intermolecular, intramolecular and dead-end cross-links were identified by their *m/z* (consistent with previous reports [284]) and the characteristic fragmentation behaviour of the DSP cross-linked peptides. Examples of these characteristic fragmentations are discussed further in subsequent sections based on cross-link type.

3.3.2.1 Intermolecular Cross-Linked Peptides

MS/MS and 'pseudo MS^3' data of a symmetrical intermolecular cross-linked species from ubiquitin are shown in Figure 3.11. Again, these cross-linked tryptic peptides fragment in

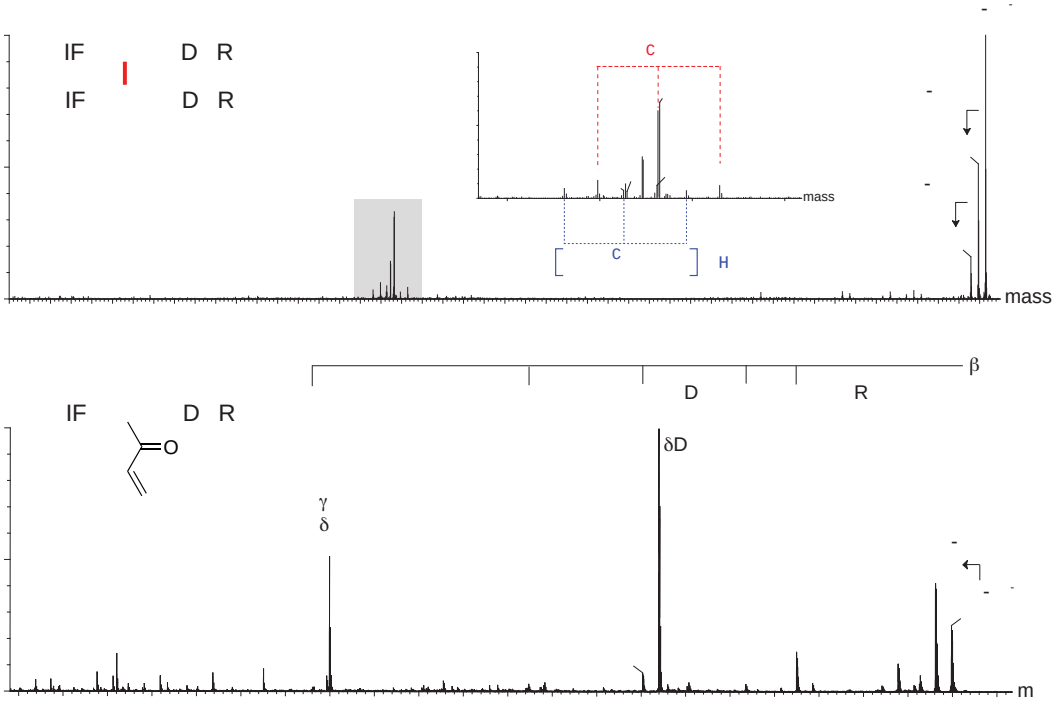


Figure 3.11: MS/MS spectrum of a symmetrical intermolecular DSP cross-linked peptide from ubiquitin ($[M\text{-}H]^-$ *m/z* 2864). (A) Low energy MS/MS results in product ions corresponding to cleavage about the disulfide bond (inset shows an enlargement of the relevant region of the spectrum, with the disulfide cleavages indicated by A′ to D′); and, (B) MS/MS of the A′ product ion formed by capillary-skimmer dissociation ($[M\text{-}H]^-$ *m/z* 1399) gives partial peptide sequence information and cross-link location.

an identical fashion to the natural cystine disulfide, showing four product ions indicative of a symmetrical intermolecular cross-link (c.f. Scheme 3.2). In this case, there is an additional, overlapping set of four diagnostic peaks observed in the MS/MS spectrum (denoted A′-H_2O to D′-H_2O). This is likely due to the prominent side chain cleavage of Asp [136], which results in the loss of H_2O (18 Da) from the product ions which contain one Asp residue. This phenomenon is investigated further in Section 3.3.4.

Partial sequence information could be obtained for the peptide fragment corresponding to *m/z* 1399 (Figure 3.11B), conclusively assigning this intermolecular cross-link. Again, side chain induced backbone cleavages from Asp and Glu residues dominate the spectrum.

An example of a MS/MS spectrum of a non-symmetrical intermolecular cross-link (*m/z* 3646) is given in Figure 3.12. This again fragmented analogously to a natural cystine disulfide, and produced A′ to D′ product ions for both peptide halves. Again, side chain induced loss of water was seen from each of these product ions as a result of the single Asp

residue present in each of the peptide halves [136]. The characteristic fragmentations and the mass of the precursor ion enabled the identification of this intermolecular cross-linked product.

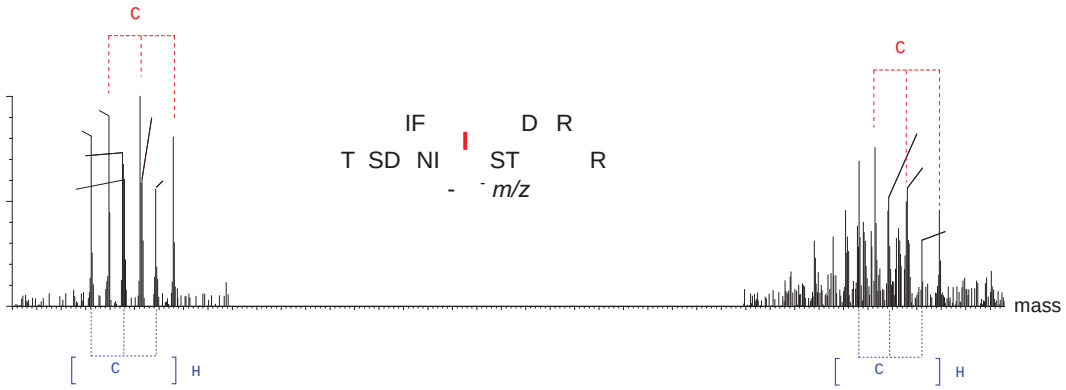


Figure 3.12: Low energy MS/MS spectrum of a non-symmetrical intermolecular DSP cross-linked peptide from ubiquitin ($[M\text{-}H]^-$ m/z 3646) shows major product ions corresponding to cleavage about the disulfide bond (A′ to D′ and A′-H_2O to D′-H_2O as indicated), and allows for identification of the cross-linked peptides.

3.3.2.2 Dead-End Cross-Linked Peptides

A significant advantage of this negative ion approach is that dead-end cross-linking products should be readily identifiable as, under low energy CID conditions, they should produce only small mass losses corresponding to fragmentations of the labile, hydrolysed disulfide cross-linker. An example of this is given in Figure 3.13, for a dead-end cross-linked peptide from ubiquitin. Products resulting from cleavages of the hydrolysed cross-linker's disulfide bond are observed, as detailed in Scheme 3.2, with a complete description of these ions given in Scheme 3.4. Again, side chain induced loss of water was seen as a result of the single Asp residue present in each of the peptide chains [136]. Additionally, in this case a fragment corresponding to the side chain induced δ backbone cleavage of Asp10 is observed (m/z 1208). This process appears to occur competitively with disulfide cleavage, and consequently cross-linker fragmentations analogous to those from the molecular ion are also noted from this δD^{10} product ion.

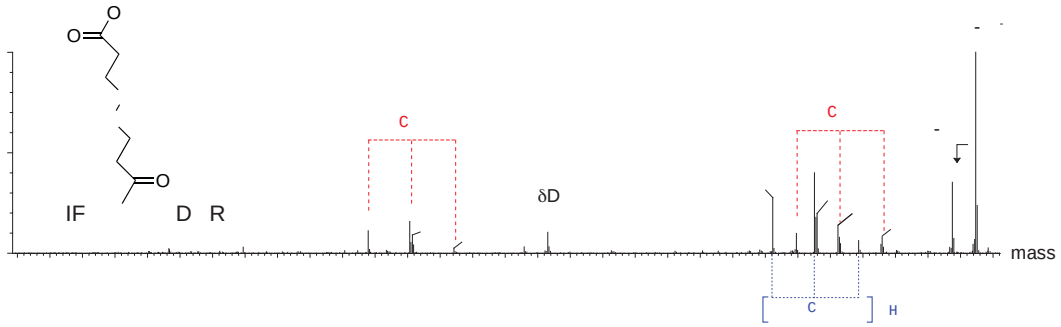
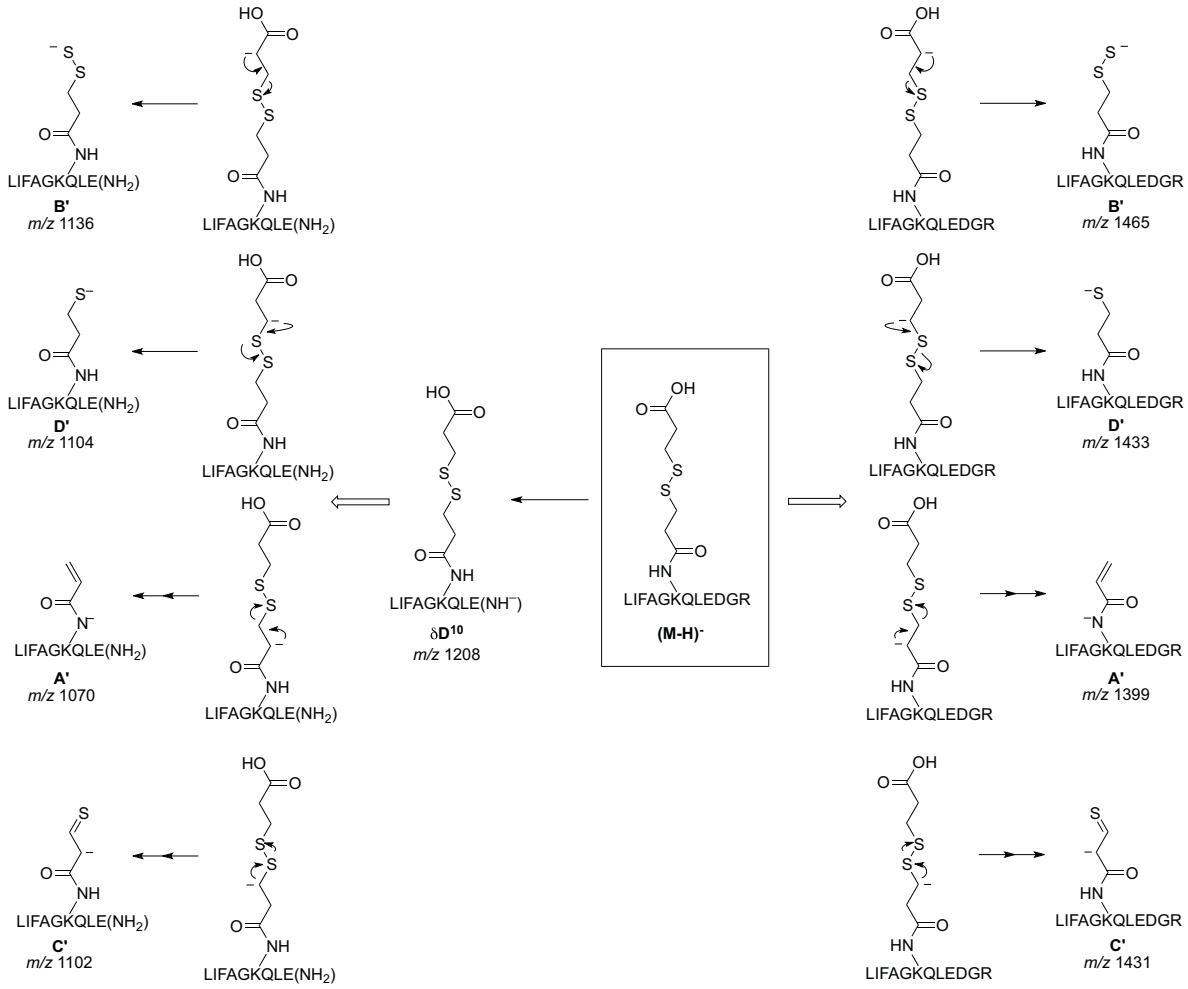


Figure 3.13: Low energy MS/MS spectrum of a dead-end DSP cross-linked peptide product from ubiquitin shows characteristic fragmentation of the hydrolysed linker ($[M\text{-}H]^-$ *m/z* 1537). Product ions are labelled A′ to D′ corresponding to Scheme 3.2. An additional set of peaks corresponding to A′-H_2O to D′-H_2O are also present. Analogous processes occur from the ion δD^{10} ion *m/z* 1208 and are labelled accordingly. For a detailed schematic of these cleavage processes refer to Scheme 3.4.



Scheme 3.4: Structures of product ions produced by disulfide cleavage of a dead-end DSP adduct from ubiquitin, as seen in the low energy MS/MS spectrum shown in Figure 3.13. Product ions are labelled A′ to D′ corresponding to the structures in Scheme 3.2. Product ions containing Asp (those on the right hand side) also show a side chain induced loss of water, so an additional, overlapping set of 4 peaks 18 Da lower in mass are observed. Anions A′ and C′ are formed by charge transfer via an ion-neutral complex.

3.3.2.3 Intramolecular Cross-Linked Peptides

Intramolecular DSP cross-links, such as the example given from ubiquitin in Figure 3.14, yield much more complex MS/MS data (summarised in Scheme 3.5). In this example, the side chain induced backbone cleavage at Asp12 occurs to give *m/z* 1367. Loss of water from this species yields *m/z* 1349. Side chain induced backbone cleavage at Asp5 results in what is essentially an unsymmetrical intermolecular cross-link, and the disulfide behaves as such, resulting in fragment anions being produced as outlined in Scheme 3.5. Not all the possible product ions are observed in the spectrum with the smaller peptide producing only the C′ and D′ type fragmentations in the spectrum (*m/z* 542 and 544, respectively). In all of the examples investigated, this is the only case where fewer than the four possible cleavage products are observed.

This spectrum is complex, yet clearly diagnostic of an intramolecular cross-link, providing significant information about the identity of the peptide involved.

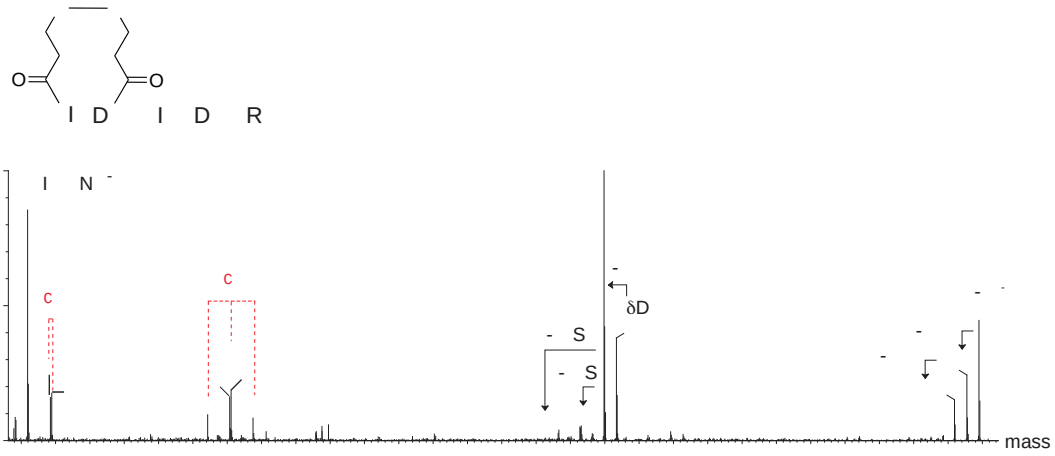
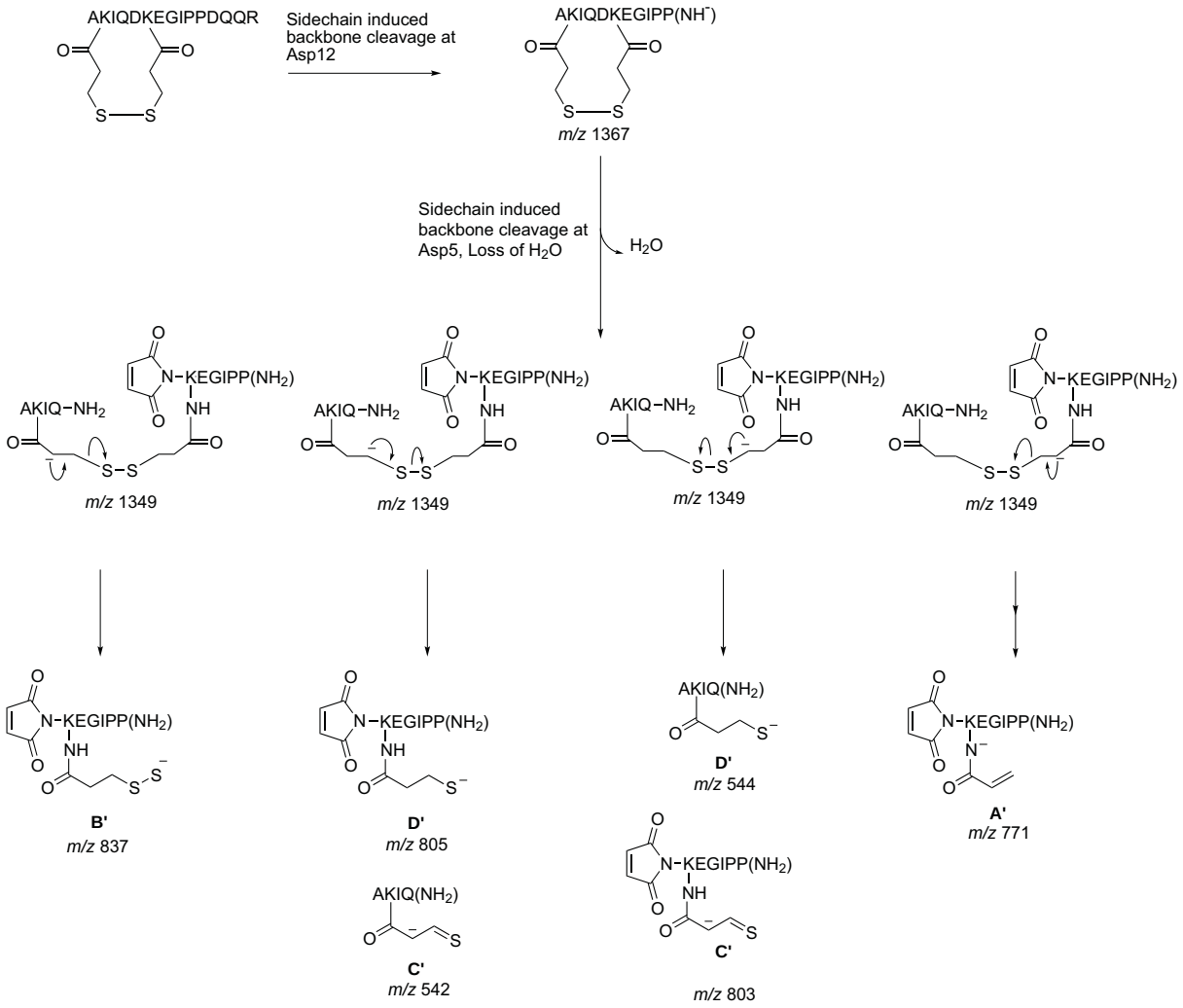


Figure 3.14: Low MS/MS CID spectrum of an intramolecular DSP cross-linked peptide product from ubiquitin ($[M\text{-}H]^-$ *m/z* 1895) shows complex fragmentation behaviour.



Scheme 3.5: Structures of product ions produced by MS/MS of an intramolecular DSP cross-linked adduct from ubiquitin. Product ions are labelled A′ to D′ corresponding to the structures in Scheme 3.2. Anions A′ and C′ are formed by charge transfer via an ion-neutral complex.

3.3.3 The Effect of Altering Spacer Arm Length

Reagents of different spacer-arm length are often utilised in CX-MS studies so that variable distance restraints can be used in molecular modelling. To address this, the disulfide containing reagents dithiobis(succinimidyl)acetate (DSA) and dithiobis(succinimidyl)butanoate (DSB) (Figure 3.15) were synthesised by esterification reactions starting from the corresponding diacid. These reagents differ from DSP by the number of methylene groups in the spacer arm. It was anticipated that the similar structural features of the three cross-linking reagents, DSP, DSA and DSB, mean they will have similar reactivity and fragmentation behaviour. The fragmentations of these novel CX reagents are discussed in this section, to determine if they are suitable for this negative ion CX-MS approach.

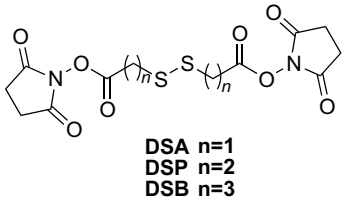


Figure 3.15: Structures of the disulfide containing CX-MS reagents DSA, DSP and DSB.

To determine the fragmentation behaviour of the intermolecular cross-links of these reagents, CX-MS studies were performed on a simple model peptide (Ac-AAKA, 401.5 $g.mol^{-1}$). Figure 3.16 displays the low energy CID MS/MS spectra of Ac-AAKA cross-linked with the three reagents. DSP cross-linked peptides fragment as previously discussed (Section 3.3.1), whilst DSA cross-linked peptides display much simpler fragmentation behaviour. Low energy CID of these species produces only two product ions 2 Da apart, corresponding to the cross-linker breaking symmetrically at the S-S bond to give C′ and D′ type ions (Scheme 3.6). Low energy activation of the DSB cross-linked peptides did not lead to fragmentation and application of higher collision energies was required. However, backbone cleavages were also seen, obscuring the disulfide cleavages (Scheme 3.7). This is a similar situation to that observed for DSP in the positive ion mode [288].

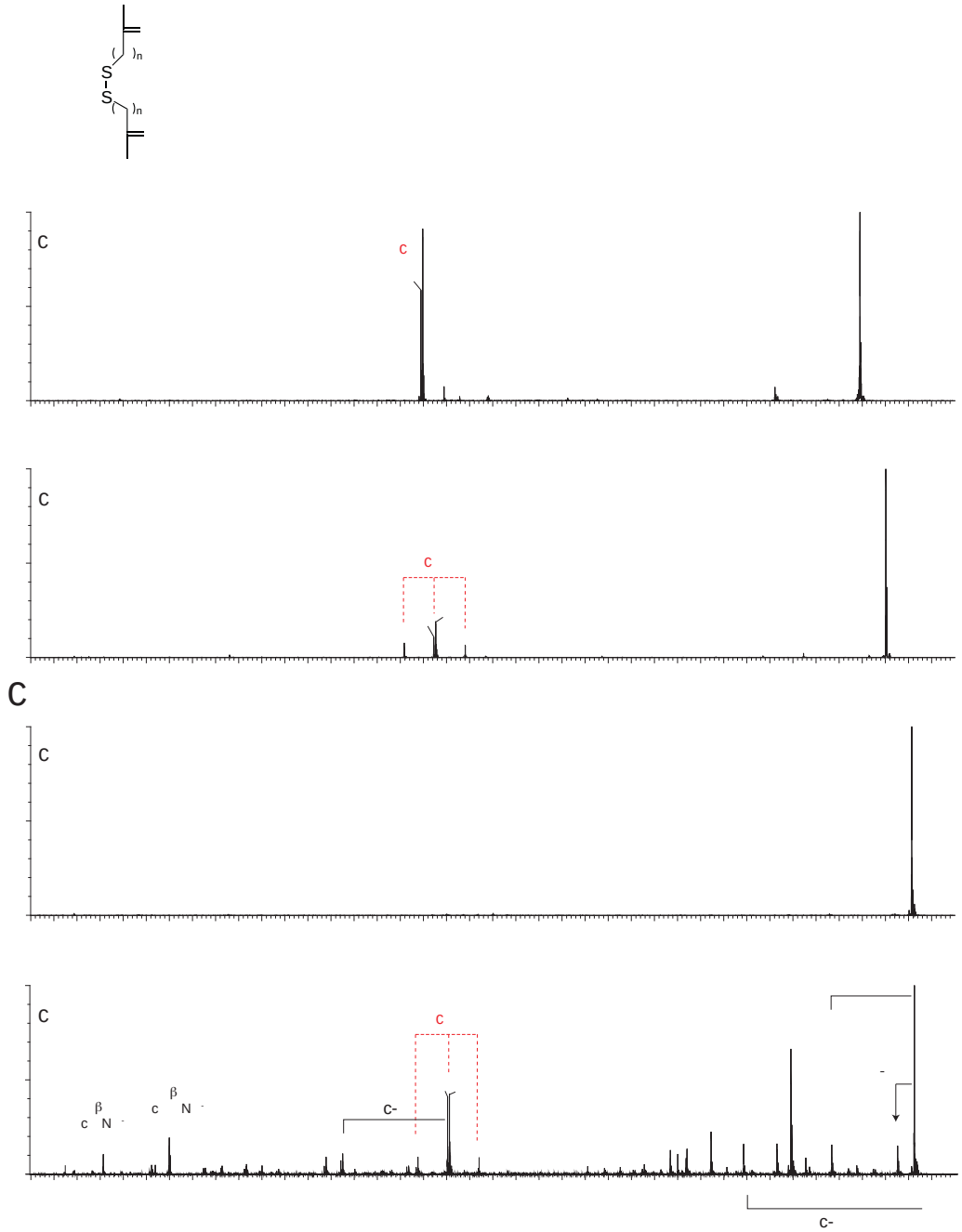
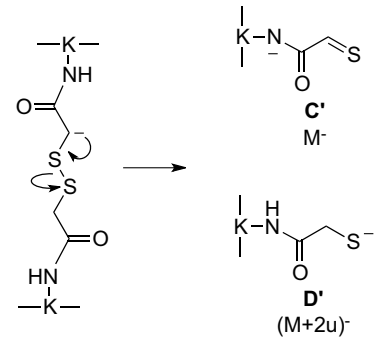
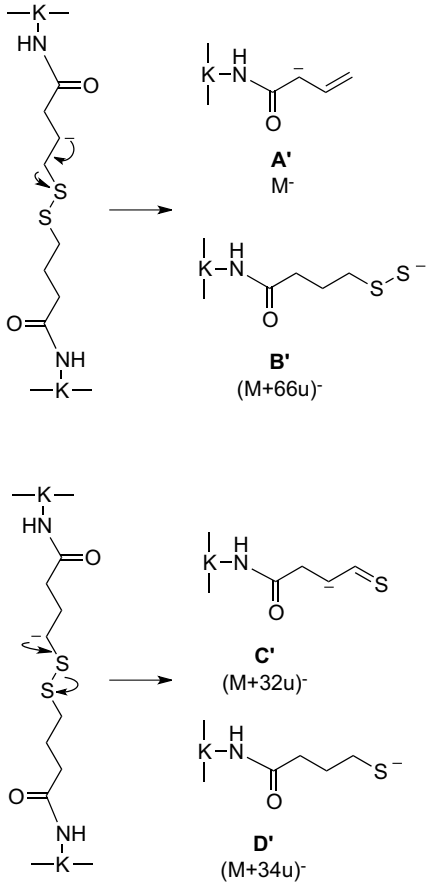


Figure 3.16: Low energy MS/MS spectra of (A) DSA, (B) DSP and (C) DSB intermolecular cross-linked linked Ac-AAKA (401.5 Da) and, (D) high collision energy MS/MS spectrum of DSB intermolecular cross-linked Ac-AAKA. Product ions A′ to D′ are observed in the spectra as indicated, corresponding to the structures in Schemes 3.2, 3.6 and 3.7. Q-Tof 2 mass spectrometer.



Scheme 3.6: Fragmentation of DSA cross-linked peptides. Masses of product ions are indicated relative to anion C′ (denoted by the arbitrary mass, M). Structures of product anion D′ can be rationalised directly from the proposed reaction mechanisms, whilst anion C′ is formed by charge transfer via an ion-neutral complex.



Scheme 3.7: Fragmentation of DSB cross-linked peptides. Masses of DSB product ions are indicated relative to anion A′ (denoted by the arbitrary mass, M). Structures of product anions B′ and D′ can be rationalised directly from the proposed reaction mechanisms, whilst anions A′ and C′ are formed by charge transfer via an ion-neutral complex.

3.3.4 Competitive Fragmentation Processes

The negative ion MS/MS spectra of peptides are often dominated by several facile side chain specific fragmentation pathways. This is a unique feature of negative ion MS. Consequently, it is vital to study these low energy fragmentations which may compete with cross-linker fragmentation thereby potentially hampering cross-link identification.

Loss of water from the Asp side chain is an extremely facile process in the negative ion mode [136, 298]. In fact, this process often produces some of the most abundant peaks in the negative ion spectra of peptides. In addition, Asp often produces side chain induced backbone fragmentation products, termed δ and γ ions [136]. These facile Asp-induced fragmentations were observed in the cross-linked peptides of ubiquitin (Section 3.3.2). Analogous processes also occur from the Glu, Gln and Asn side chains (Asn and Gln lose NH_3 from the side chain) [136]. In this section, a systematic study is performed using model peptides to identify if other unique negative ion side chain fragmentations are competitive with those of the cross-linker disulfide. To begin with, Asp containing model peptides with both a natural intermolecular disulfide-bond (Ac-ACADA, Figure 3.17) and intermolecular cross-links with DSA and DSP (Ac-AKADA, Figure 3.18) were synthesised.

Competitive loss of water is observed in the MS/MS spectrum of the intermolecular cystine disulfide containing peptide (Figure 3.17), with 8 prominent product ions observed in the spectrum, i.e. product ions A to D and A-H_2O to D-H_2O. The same phenomenon is also observed in the MS/MS spectra of Asp containing peptides cross-linked with DSP (Figure 3.18B), with an additional, overlapping set of four fragment peaks (A′-H_2O to D′-H_2O) observed in the spectrum (analogous to the natural disulfide case). This has also been previously observed in the cross-linked tryptic peptides of ubiquitin (Section 3.3.2). An additional set of peaks are also observed in this spectrum (m/z 383, 415, 417 and 449) corresponding to disulfide cleavages after side chain induced backbone cleavage at Asp i.e. disulfide cleavage from the Asp δ ion.

Shortening the spacer-arm length does not eliminate the competitive loss of H_2O from the Asp side chain completely, and again, cleavage of DSA (Figure 3.18A) occurs competitively with the loss of water from the Asp side chain. However, the relative abundance of the two sets of peaks indicates that cleavage of the DSA disulfide is more favourable than that of

DSP. Additionally, product ions from the Asp δ ion are not observed in the spectrum, again indicating that the cleavages of DSA must be more facile than those of DSP.

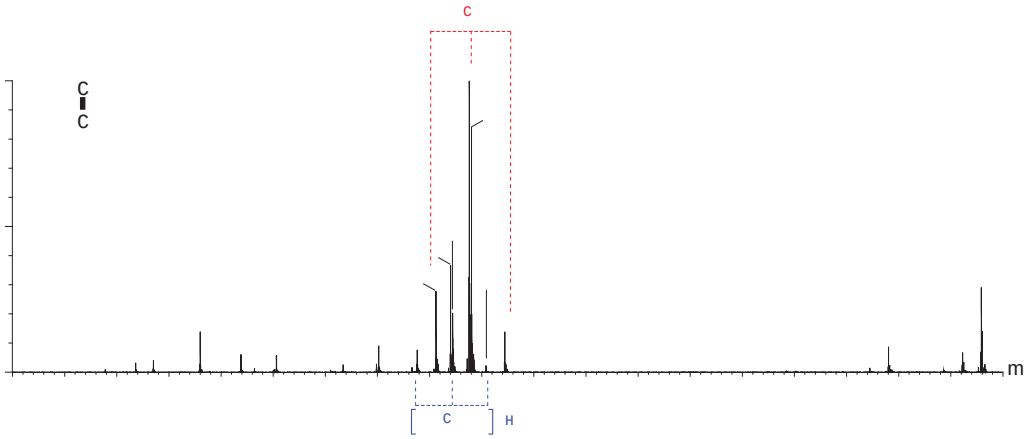


Figure 3.17: Low energy MS/MS spectrum of intermolecular cystine disulfide linked Ac-ACADA ($[M\text{-}H]^-$ m/z 979). Product ions A to D are observed in the spectrum (c.f. Figure 3.6 and Scheme 3.1) in addition to a second set of peaks, A-H_2O to D-H_2O.

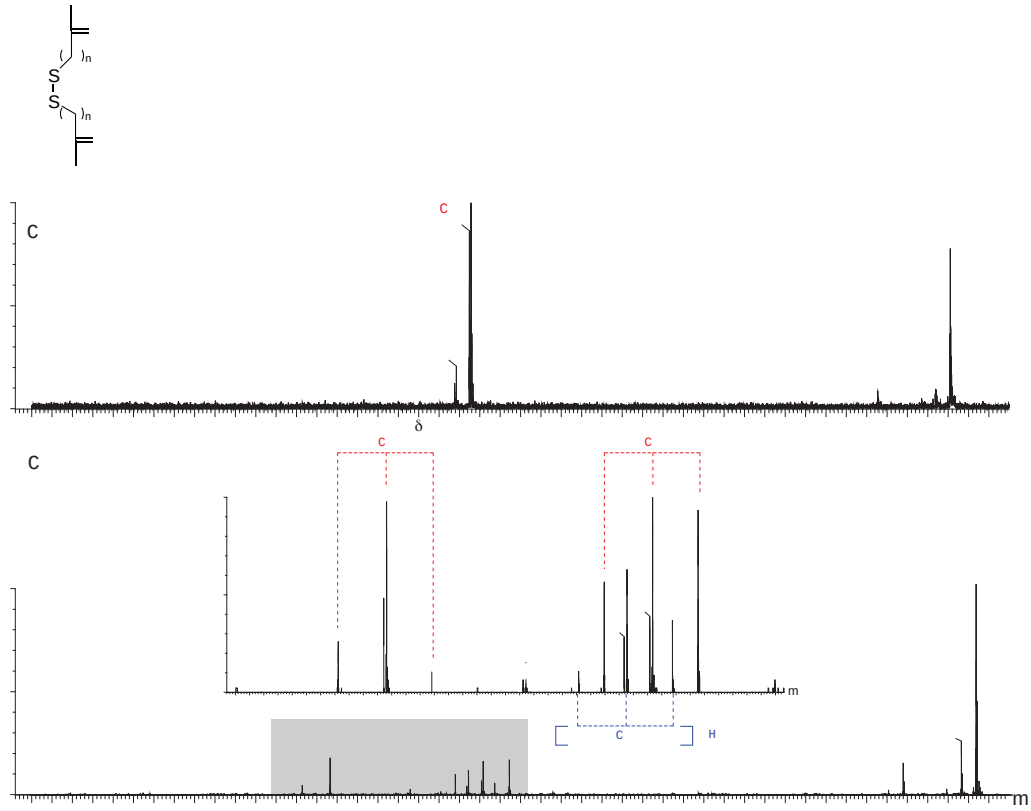


Figure 3.18: Low energy MS/MS spectra of (A) DSA ($[M\text{-}H]^-$ m/z 1178) and (B) DSP ($[M\text{-}H]^-$ m/z 1207) intermolecular cross-linked linked Ac-AKADA. Product ions A′ to D′ are indicated in the spectra where observed (Schemes 3.2 and 3.6) as well as an additional set of peaks corresponding to A′-H_2O to D′-H_2O and those following Asp side chain induced backbone cleavage.

The side chain loss of water from Glu, and loss of NH_3 from Asn do not appear to compete with any of the DSP or DSA cross-linker fragmentations (Figures 3.19 and 3.20). In these instances, the spectra only show A′ to D′ type product ions at low collision energies, and no side chain losses nor side chain induced backbone cleavages are observed.

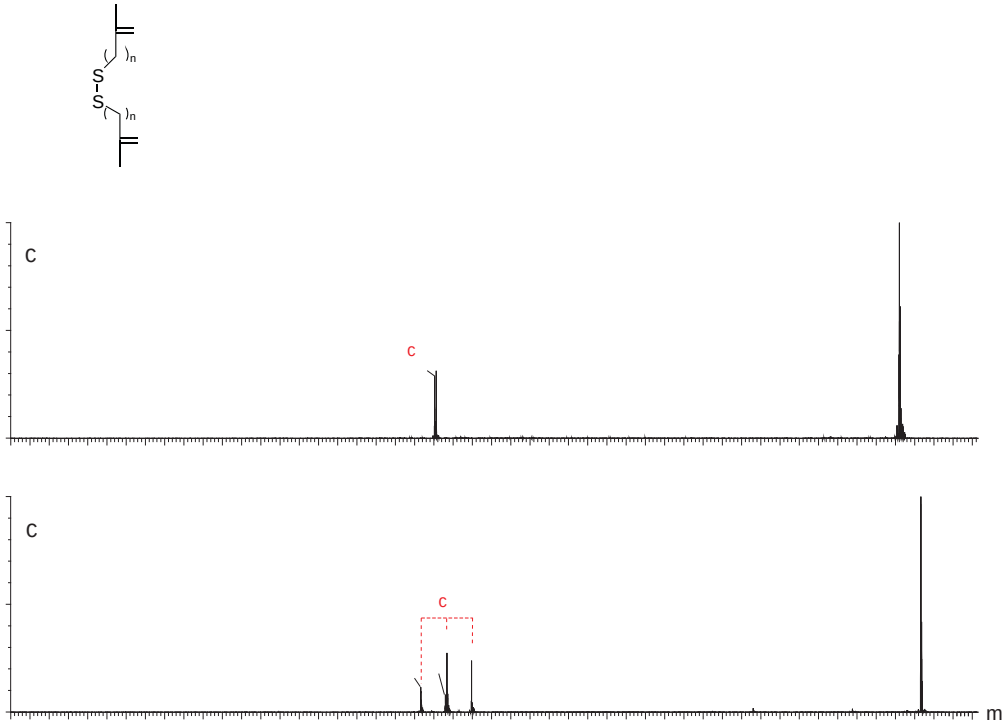


Figure 3.19: Low energy MS/MS spectra of (A) DSA ($[M\text{-}H]^-$ *m/z* 1205) and (B) DSP ($[M\text{-}H]^-$ *m/z* 1233) intermolecular cross-linked linked Ac-AKAEA. Product ions A′ to D′ are indicated in the spectra where observed (Schemes 3.2 and 3.6).

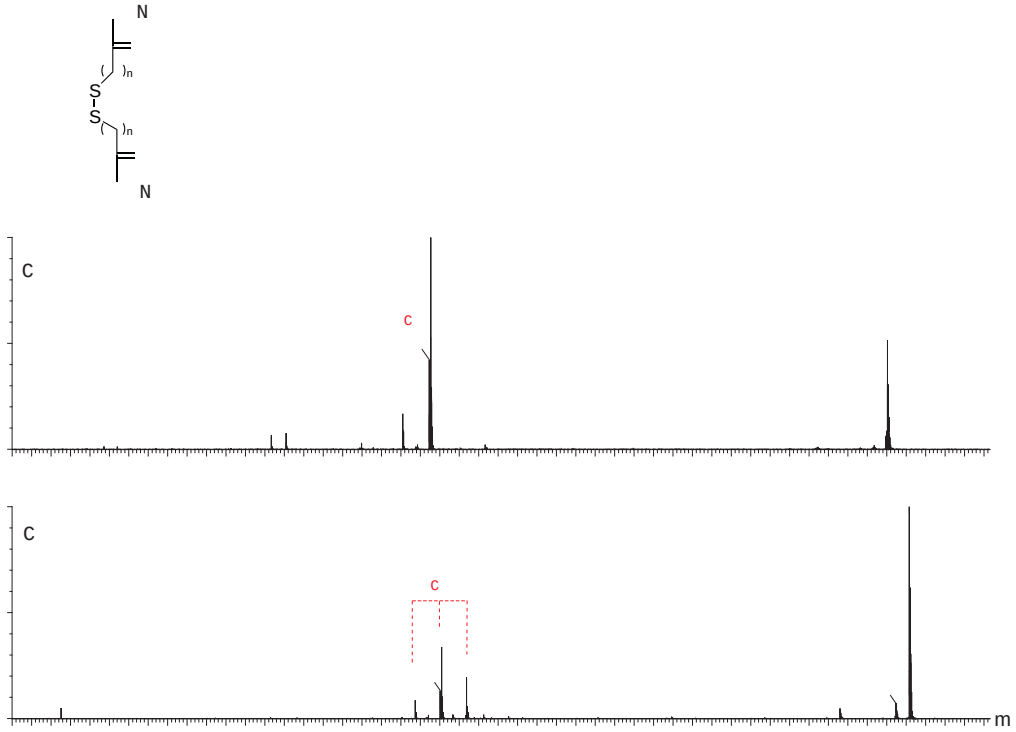


Figure 3.20: Low energy MS/MS spectra of (A) DSA ($[M\text{-}H]^-$ *m/z* 1176) and (B) DSP ($[M\text{-}H]^-$ *m/z* 1204) intermolecular cross-linked linked Ac-AKANA. Product ions A′ to D′ are indicated in the spectra where observed (Schemes 3.2 and 3.6).

Another negative ion fragmentation that could potentially compete with the low energy disulfide cleavages of DSA and DSP is the loss of CH_2O from the Ser side chain (converting Ser to Gly). In fact, the MS/MS spectra of Ser containing peptides often show backbone fragmentations occurring from the $[(M\text{-}H)^-\text{-}CH_2O]$ fragment anion rather than the parent ion $(M\text{-}H)^-$, demonstrating that this loss is more facile than backbone fragmentations [140, 144]. To determine if Ser side chain fragmentations are competitive with those of the CX-MS reagents, the model peptide Ac-AKASA was cross-linked with DSA and DSP. Competitive loss of CH_2O is not observed from the DSA cross-linked model peptide (Figure 3.21A). However, low energy activation of the DSP cross-linked peptide afforded the four product ions as expected, and relatively low abundance peaks corresponding to the loss of CH_2O (Figure 3.21B).

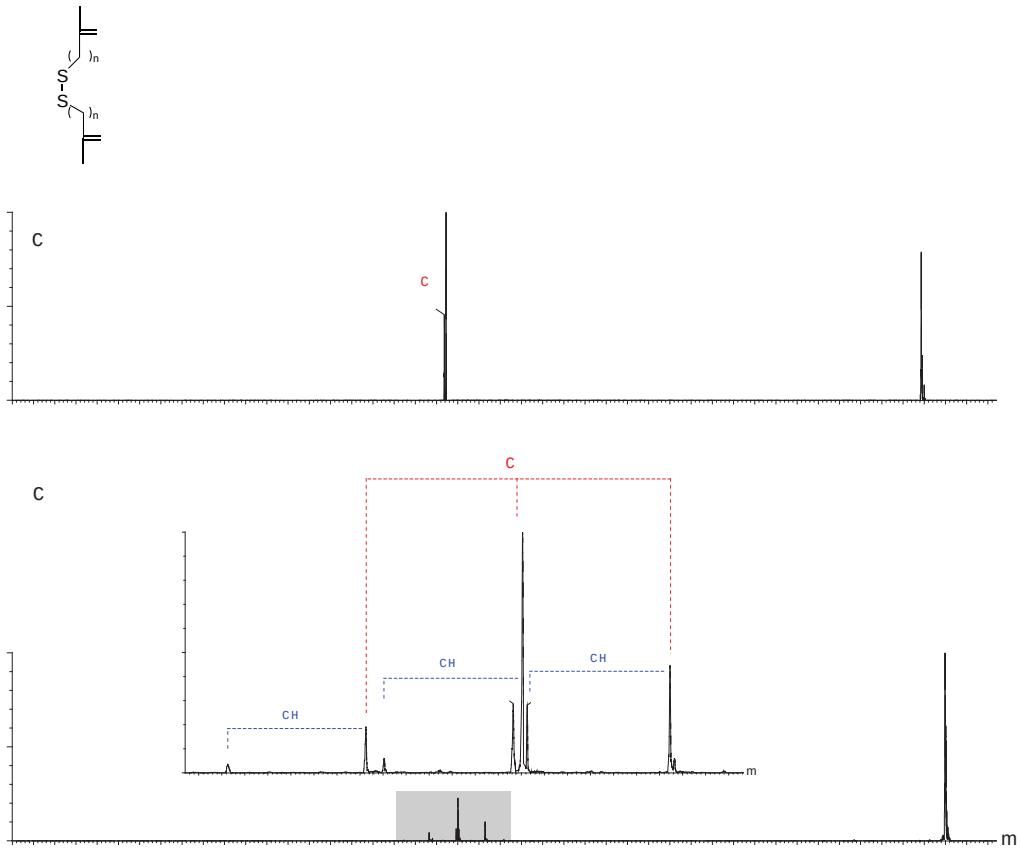


Figure 3.21: Low energy MS/MS spectra of (A) DSA ($[M\text{-}H]^-$ *m/z* 1122) and (B) DSP ($[M\text{-}H]^-$ *m/z* 1150) intermolecular cross-linked linked Ac-AKASA. Product ions A′ to D′ are indicated in the spectra where observed (Schemes 3.2 and 3.6) as well as an additional set of peaks corresponding to A′-CH_2O to D′-CH_2O, as indicated.

3.3.5 Calculations

The calculations presented in this section were performed by Dr. Tianfang Wang (School of Chemistry and Physics, The University of Adelaide).

It should be noted that the reaction coordinate profiles in this section all commence at a nominal energy of 0 kJ.mol^{-1}. However, calculations have shown that a carboxylate anion is more stable than the enolate, and approximately 105-115 kJ.mol^{-1} of energy is required for this interconversion to occur by proton transfer [139]. Additionally, conversion of the carboxylate to an amide anion requires approximately 40 kJ.mol^{-1} of energy [299]. Consequently, the data presented here assume that there is sufficient energy present during the MS experiment for hydrogen transfer to occur, forming the anions which initiate these processes. In principle, negative ion cleavages may be charge-remote processes [300], i.e. not influenced by a charged site. However, no major negative ion cleavages have been confirmed to occur in this fashion [136, 140], suggesting that these processes proceed by charge-directed reaction mechanisms as described.

3.3.5.1 Energetics of DSP cleavages

The reaction coordinate profile of the cleavage reaction adjacent to the disulfide in a model DSP system is shown in Figure 3.22. These calculations were performed to confirm the relative energetics of the disulfide cleavage processes in DSP compared with the normal backbone fragmentations, and to allow structures for the resultant anions to be proposed [301]. The process to yield a B′ anion is exothermic with only a slight transition state barrier. The data suggest that following bond cleavage, the anion R-S_2^- approaches the amide hydrogen of the Lys side chain to form a stable ion-neutral complex, resulting in deprotonation at this site to afford an A′ type anion (in a slightly unfavourable process). The reaction coordinate profile of the cleavage reaction between the sulfur atoms of the disulfide is shown in Figure 3.23. The formation of D′ is slightly unfavourable, and again, ion-neutral complex formation drives charge transfer by deprotonation at the enolate position to yield C′ in an overall exothermic process.

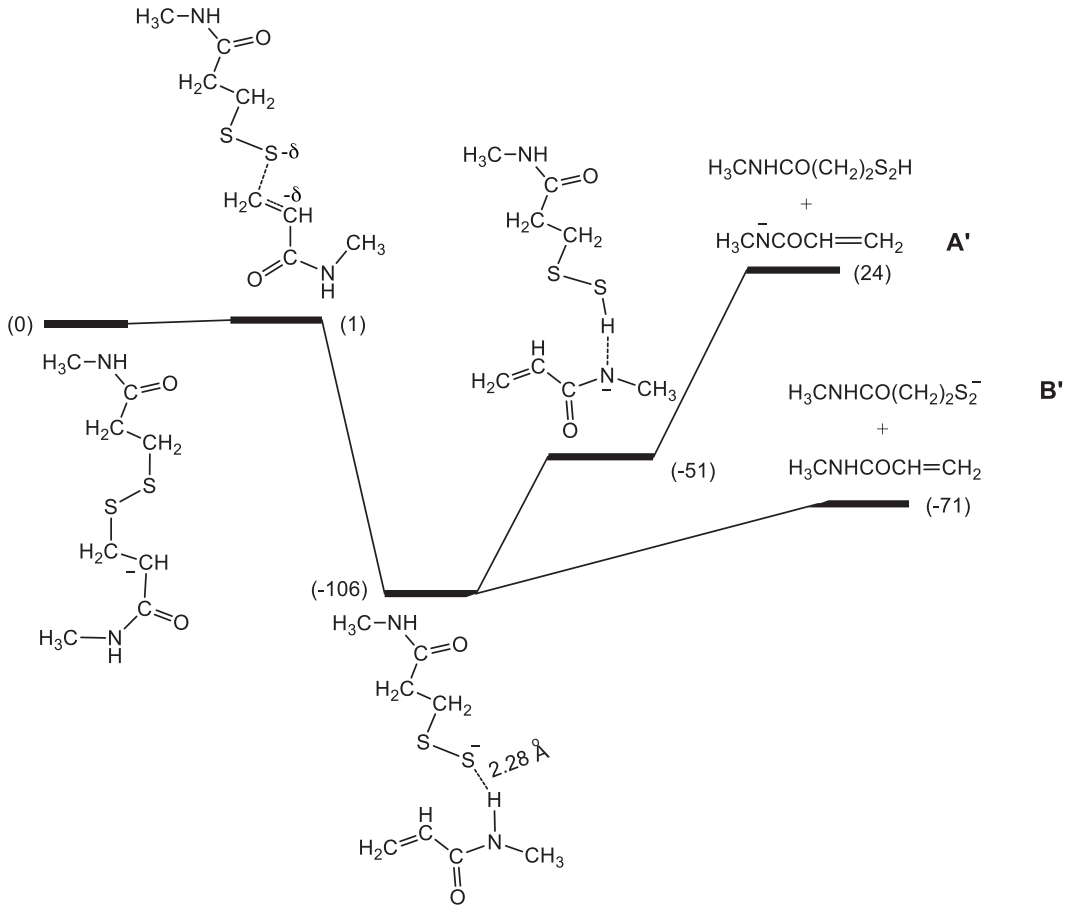


Figure 3.22: Reaction coordinate profiles of processes that give A′ and B′ anions (Scheme 3.2), which proceed via the enolate anion. CAM-B3LYP/6-311++g(d,p) level of theory. Relative ΔG values are in kJ.mol^{-1} [301].

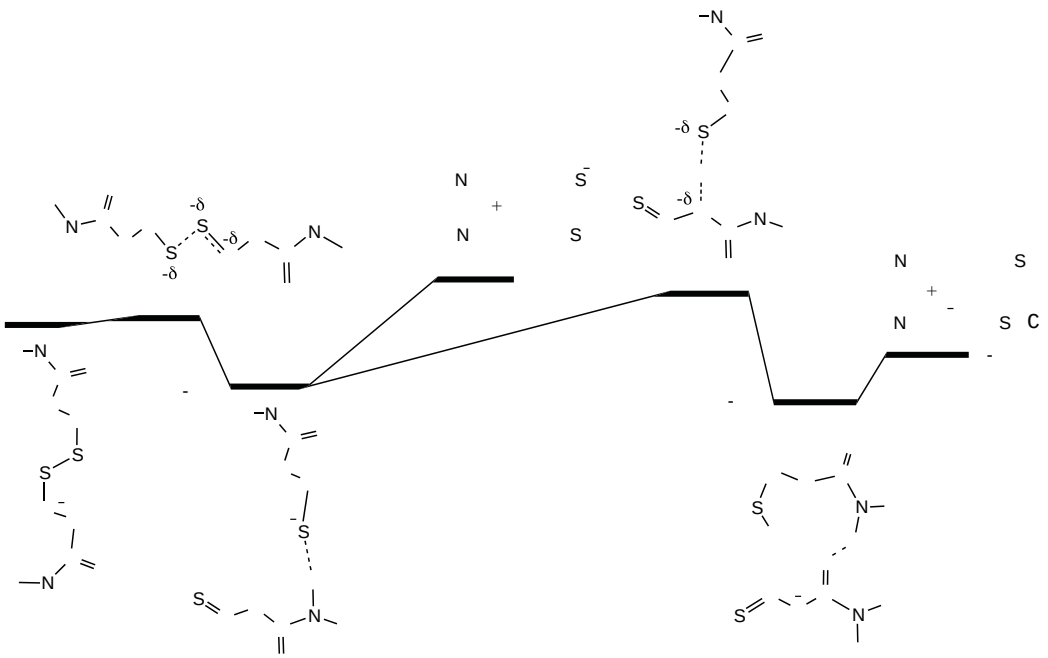


Figure 3.23: Reaction coordinate profiles of processes that give C′ and D′ (Scheme 3.2), which proceed via the anion located adjacent to the disulfide on the cross-linker. CAM-B3LYP/6-311++g(d,p) level of theory. Relative ΔG values are in kJ.mol^{-1} [301].

3.3.5.2 Energetics of DSA cleavage

Symmetrical cleavage of the DSA linker between the sulfur atoms yields two characteristic product ions that can be used to easily identify cross-linked peptides by low energy collisional activation of the precursor ion (Figure 3.16A, Scheme 3.6). Again, the low energy nature of this cleavage process has been probed using calculations beginning with a model system deprotonated at the enolate position, to give the C′ and D′ ions seen experimentally (Figure 3.24) [302]. The process is only slightly exothermic and a transition state could not be found.

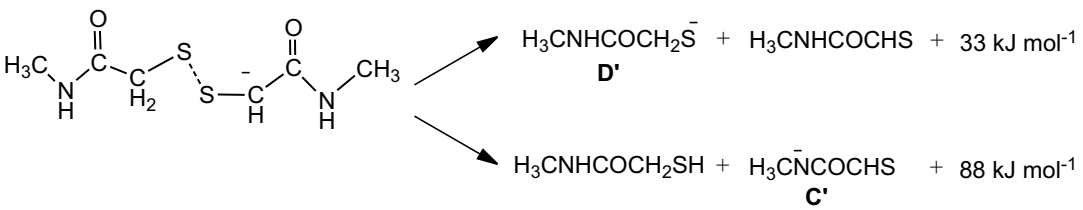


Figure 3.24: Reaction energetics for cleavage of the disulfide bond in a DSA model system. CAM-B3LYP/6-311++g(d,p) level of theory. ΔG relative values are in kJ.mol^{-1} [302].

3.3.5.3 Energetics of DSB cleavage

The low energy CID MS/MS data show that DSB is not a suitable candidate cross-linker for this negative ion CX-MS approach, as disulfide cleavage does not occur before fragmentation of the backbone (Figure 3.16C/D, Scheme 3.7). Consequently, detection of intermolecular cross-links from these data is cumbersome, and sequencing would be more difficult by on-line MS^n methods.

Again, calculations have been performed to examine the energetics of these cleavage processes [302]. Fragmentation between the sulfur atoms of the disulfide bond proceeds from an enolate anion, via a seven-centre hydrogen transfer transition state (with a modest energy barrier) to form C′ and D′ type product ions in an overall exothermic processes (Figure 3.25). For the cleavage process adjacent to the disulfide bond, hydrogen transfer is again required to induce fragmentation (with a modest energy barrier), to give fragment anions A′ and B′ in overall exothermic processes (Figure 3.26).

Figure 3.25: Reaction coordinate profiles for the cleavage between the sulfur atoms of the disulfide bond in a DSB model system. CAM-B3LYP/6-311++g(d,p) level of theory. ΔG relative values are in $kJ.mol^{-1}$ [302].

Figure 3.26: Reaction coordinate profiles for the cleavage adjacent the disulfide bond in a DSB model system. CAM-B3LYP/6-311++g(d,p) level of theory. ΔG relative values are in kJ.mol^{-1} [302].

The energetics of the DSB cleavage pathways shown in Figures 3.25 and 3.26 suggest that DSB fragmentation should be more energetically favourable than that of the backbone, however, this was not observed in the MS/MS data. Consequently, this difference must be due to alternate factors. It has been shown that this cleavage does not proceed due to the energy required for deprotonation to occur at the enolate position in DSB [302]. It can be seen that the Gibbs Free Energy of deprotonation (ΔG_{acid}) at the enolate position of DSB is significantly higher than that for DSA and DSP (Scheme 3.8). In addition, DSA and DSP enolate formation requires similar energy to the deprotonation process at the backbone enolate position in a model Gly tripeptide (Scheme 3.8). This suggests that the population of this DSB enolate anion would be significantly lower than that of the backbone enolate anion, and the conversion (proton transfer) between these anions would be very unfavourable. This explains why cleavages of the disulfide bond in DSB do not occur at the same low collision energies.

Scheme 3.8: Gibbs Free Energy of deprotonation (ΔG_{acid}) at the enolate position of DSA, DSP, DSB and a peptide (GlyGlyGly) model. ΔG_{acid} values are in kJ.mol^{-1}. CAM-B3LYP/6-311++g(d,p) level of theory [302].

3.3.5.4 Energetics of H_2O loss from the Asp side chain

The most prominent competitive side chain fragmentations observed in this work are from Asp residues. Asp-containing peptides have been shown to eliminate H_2O from the Asp side chain and Asp has also been shown to trigger side chain induced backbone cleavage [136, 298]. In this work, a loss of water that is competitive with the cross-linker disulfide fragmentation in DSP is particularly prominent. Deuterium labelling studies have indicated that the loss of water is most likely a consequence of a cyclisation process initiated by an amide anion of the subsequent residue [136, 298]. This cleavage process occurs from an amide anion (Figure 3.27) and is slightly exothermic with a modest transition state barrier [302]. Thus, this cleavage process is similar in energy to that of the disulfide fragmentations of all three cross-linker systems studied (Figures 3.22 to 3.26), and explains the competitive nature of these processes.

Figure 3.27: Reaction coordinate profile for water elimination from the Asp side chain. CAM-B3LYP/6-311++g(d,p) level of theory. ΔG relative values are in kJ.mol^{-1} [302].

3.4 Discussion

This work presents the the first examples of the structural analysis of cross-linked peptides by negative ion MS. Together, these results indicate that the disulfide-containing reagents DSA and DSP show promise as CID cleavable cross-linking reagents with analysis in the negative ion mode, and that this approach can effectively distinguish dead-end, intra- and inter-molecular cross-linked peptides.

3.4.1 MS Analysis of Cross-Linked Peptides

The data presented herein indicate that this negative ion approach offers a clear advantage over traditional positive ion methodologies for the identification of disulfide containing cross-linked peptides from complex mixtures. The cross-linker disulfide in DSA and DSP cleaves prior to the backbone resulting in the production of product ions that are characteristic and easy to recognise (c.f. Figures 3.7 and 3.16A/B). This is a result of the low energy barriers associated with disulfide cleavage (Section 3.3.5) when compared with those of the backbone α and β fragmentations (Section 2.2.6.2), which are endothermic and must overcome a maximum transition state barrier in the order of 400 kJ.mol^{-1} [136].

Despite the somewhat reduced sensitivity that results from using negative ion MS for the analysis of peptides, the detectability of disulfides increases substantially due to these facile, characteristic fragmentations. As a consequence, cross-link identification can occur with relative ease and the identity of the peptides involved confirmed by sequencing them individually in the same experiment (by MS^3).

All examples of symmetrical intermolecular cross-links by DSP that have been investigated produce four product ions upon collisional activation (8 ions if unsymmetrical), as well as those following Asp induced H_2O losses (Figures 3.7, 3.11 and 3.12). These four product ions A$'$ to D$'$ are diagnostic, easily recognisable and are amenable to sequencing by MS^3 approaches. Thus the presence of some or all of these four ions may be used as a marker for the identification of intermolecular cross-links, and can provide the exact site of cross-linking.

In comparison, dead-end cross-linked peptides show low mass losses corresponding to cleavage of the hydrolysed disulfide (Figure 3.13). These losses again correspond to fragmentation processes producing anions of type A′ to D′, which can be utilised to easily recognise a dead-end modified peptide (Schemes 3.2 and 3.4).

Intramolecular cross-linked peptides show much more complex spectra, with side chain induced backbone cleavages dominating the MS/MS spectrum (Figure 3.14 and Scheme 3.5). However, these spectra show neither low mass losses nor the characteristic disulfide pattern of product ions upon CID. This complication is not necessarily disadvantageous as intermolecular cross-links provide little information about the quaternary structures of the complex and are often not investigated. Additionally, the example given here is unusual as it contains an Asp residue between the two cross-linked Lys residues, causing intermolecular disulfide cleavage to be observable following breakage of the peptide backbone at Asp. Furthermore, in this instance a loss of H_2S_2 is observed in the spectrum (albeit not from the molecular ion), which is characteristic of an intramolecular disulfide (Scheme 3.2) [293].

It should be noted that in a previous study where the negative ion cleavages of disulfide containing proteins were investigated [295], it was shown that all four possible product ions were not evident in all cases. This indicates that some care may need to be taken in analysing the low energy MS/MS spectra, as all four possible cleavage products may not be observed. This may have an effect on the development of high throughput data analysis methods using this approach, however, the detection of at least two of the possible fragment anions should provide sufficient diagnostic data to indicate the presence of a cross-linked peptide.

It should also be noted that positive ion MS has been utilized to identify both naturally occurring cystine disulfides and cross-linked products of cysteine-based regents [288, 303–305]. In these instances, disulfide fragmentations occur simultaneously to the backbone fragmentations used for sequencing [288, 303–305]. Data of this kind make identification of disulfide-containing peptides time-consuming and difficult. Additionally, it has been reported that disulfide cleavage may be favoured from the $[MH]^+$ ion but not from those that are multiply charged [306] and therefore this positive ion approach has limited applicability. In contrast, this negative ion approach to identify intermolecular or dead-end

cross-linked peptides yields MS/MS data which are simpler, containing primarily fragment anions corresponding to the facile cleavages of the disulfide, without the complications of backbone fragmentation. This makes identification and sequencing (by MS^3) of cross-linked peptides easier than these other positive ion methodologies. Furthermore, negative ion MS has been shown to have other advantages in peptide sequencing, such as the detection of other post-translational modifications and the further development of negative ion methodologies is vital in this endeavour [307].

3.4.2 Competitive Fragmentation Processes

CID labile CX-MS reagents of the type discussed in this chapter rely on the cross-linker fragmentation to occur at lower energies than any of the backbone cleavages. This allows the peptides involved in any intermolecular cross-link to be separated and sequenced independently. However, in the negative ion mode there are a variety of residue-specific fragmentations, many of which are more favourable than backbone fragmentations (as detailed in Section 2.2.6.2). Therefore, there is potential for these other interfering fragmentations to complicate any MS/MS analysis of cross-linked peptides in the negative ion mode.

The comparable energies of the cross-linker fragmentation and loss of H_2O from Asp (Section 3.3.5) suggest that identification of intermolecular disulfide-linked peptides in the negative ion mode that contain Asp is likely to be somewhat complicated by the concomitant water loss and disulfide cleavages. However, it may also prove diagnostic for the presence of Asp within a disulfide containing peptide and in many cases this, along with the measured m/z, may be enough to tentatively confirm the identity of the peptide.

Additionally, the data presented herein indicate that loss of CH_2O from the Ser side chain (converting Ser to Gly) is competitive with the low energy disulfide cleavages in DSP intermolecular cross-linked peptides. The energetics of this side chain loss show (at a modest level of theory) that it is more energetically favourable than backbone cleavages [140, 144], and must surpass a maximum barrier of 149 $kJ.mol^{-1}$. The losses of CH_2O observed in the spectrum here are not as abundant as the side chain losses present in the MS/MS spectra of Asp containing peptides, and consequently should not interfere with any

analysis in this negative ion CX-MS approach. Again, the presence of these peaks may prove diagnostic for the presence of Ser within a disulfide containing peptide and may assist with assigning the identity of the peptide.

3.5 Experimental Procedures

3.5.1 Materials

Chemicals and proteins were purchased from Sigma Aldrich (St. Louis, MO, USA) unless otherwise specified, and used as received. *N*-9-Fluorenylmethoxycarbonyl (*N*-Fmoc) protected amino acids and 2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) were purchased from GL Biochem (Shanghai, China). *N,N′*-Dimethylformamide (DMF) (AR grade) and piperidine (reagent grade) were purchased from Merck (Darmstadt, Germany). *N,N′*-Diisopropylethylamine (DIPEA) and trifluoroacetic acid (TFA) were purchased from Alfa Aesar (Heysham Lancashire, UK). Acetonitrile (HPLC grade) was purchased from Scharlau (Sentmenat, Barcelona, Spain). Diethyl ether and dimethylsulfoxide (DMSO) were purchased from ChemSupply (Gillman, SA, Australia). Dithiobis(succinimidyl propionate) (DSP) was purchased from Thermo Scientific (Rockford, IL, USA). Sequencing grade modified trypsin was purchased from Promega (Madison, WI, USA).

3.5.2 General Procedure: Peptide Synthesis

Peptides were synthesised by standard Fmoc solid phase methods [308] using 2-chlorotrityl chloride resin (GL Biochem, Shanghai, China) on a 0.05 mmol scale. Sequential amino acid couplings were performed by incubating the resin in a solution containing 10 equivalents of the appropriate *N*-Fmoc protected amino acid (0.5 mmol), 10 equivalents of HATU [309] (from a 0.5 M solution in DMF, 1 mL, 0.5 mmol) and 20 equivalents of DIPEA (1 mmol) in DMF (2 mL) for 20 min at 25°C. The Fmoc group was removed by treating the resin with 20% v/v piperidine in DMF for 2 × 10 min.

Amino acid attachment and Fmoc deprotection were confirmed using the 2,4,6-trinitrobenenesulfonic acid (TNBS) test [310]. This was achieved by removing a sample of the resin, adding equal quantities of 5 % DIPEA in DMF and 1 % TNBS in DMF (25 μL) and mixing the sample for 2 min. If after coupling the test was positive, the *N*-Fmoc protected amino acid was coupled again under the same conditions.

The final peptide was released from the resin, with simultaneous removal of protecting groups, by treatment with TFA/triisopropylsilane (TIPS)/H_2O/3,6-dioxa-1,8-octanedithiol (DODT) (90:3.3:3.3:3.3) at 25 °C for 3 hr. The crude peptide was precipitated by addition of cold diethyl ether, and isolated by centrifugation. The solid material was washed with cold diethyl ether, dissolved in 30 % acetonitrile/water and then lyophilised.

3.5.3 General Procedure: Synthesis of Cross-linkers

The appropriate diacid (1.4 mmol), N-hydroxysuccinimide (0.41 g, 3.6 mmol) and N,N′-dicyclohexylcarbodiimide (0.83 g, 4.0 mmol) were dissolved in 1:1 anhydrous acetone:dichloromethane (15 mL) and the reaction mixture was stirred overnight at 25 °C. The resultant precipitate was removed by gravity filtration and the filtrate concentrated *in vacuo* to afford a powder, which was resuspended in acetone (approx. 2 mL). The insoluble material was collected and washed with acetone to afford the cross-linker as a white powder.

Dithiobis(succinimidyl acetate) (DSA). Dithiodiglycolic acid (0.26 g, 1.4 mmol) was used and DSA was produced as a white powder (0.39 g, 74 %); HRMS *m/z* calc'd for $[C_{12}H_{12}N_2O_8S_2+H]^+$: 377.0108, found 377.0110. ^{1}H NMR (300 MHz, DMSO-d_6) δ 2.83 (s, 8H), 4.24 (s, 4H).

Dithiobis(succinimidyl butanoate) (DSB). Dithiodibutyric acid (0.34 g, 1.4 mmol) was used and DSB was produced as a white powder (0.51 g, 84 %); HRMS *m/z* calc'd for $[C_{16}H_{20}N_2O_8S_2+H]^+$: 433.0734, found 433.0745. ^{1}H (300 MHz, DMSO-d_6) δ 2.00 (q, J=7.0 Hz, 4H), 2.80 (t, J=7.0 Hz, 8H), 2.81 (s, 8H).

3.5.4 General Procedure: Cross-Linking of Model Peptides

The appropriate peptide (200 nmol) was dissolved in DMSO (200 μL), to give a concentration of 1 mM. 1 Equivalent of both the appropriate cross-linker (20 μL, from a 10 mM solution in DMSO) and DIPEA (20 μL, from a 10 mM solution in DMSO) were added. The mixture was incubated at 25 °C for 5 hr with periodic mixing, before being concentrated to dryness using a SpeedVac (Savant, Farmingdale, NY). These conditions

produce high yields of cross-linked products for analysis [284, 297]. The residue was dissolved in water and desalted by ZipTip (Millipore, Bedford, MA, USA) before MS analysis.

3.5.5 Cross-Linking of Ubiquitin

A 200 μM stock solution of bovine ubiquitin was prepared by dissolving the protein (1.7 mg) in phosphate buffered saline (PBS) (1 mL). An aliquot (40 μL) was removed and a ten-fold molar excess of DSP was added from a 10 mM solution in DMSO (4 μL). The reaction mixture was incubated at 25 °C for 1 hr and digested overnight with trypsin, which was added in a 1:50 ratio (ubiquitin:trypsin, w/w). The digest was then lyophilised and desalted by ZipTip (Millipore, Bedford, MA, USA) before MS analysis.

3.5.6 Mass Spectrometry

Nano-ESI mass spectra were acquired using a Micromass Q-ToF 2 (Waters/Micromass, Manchester, UK) orthogonal acceleration time-of-flight mass spectrometer. Samples were introduced into the spectrometer using platinum-coated borosilicate capillaries prepared in-house. Conditions were typically as follows: capillary voltage 1.4 kV, source temperature 80 °C, and cone voltage 30-50 V. Ar collision gas energies of approximately 15-30 eV were used for 'low energy' CID, while 'high energy' CID utilized collision gas energies of 30-50 eV. Product ions created by capillary-skimmer dissociation were generated by applying cone voltages of typically 70 V and were then mass selected for further MS/MS analysis (described in the chapter as 'pseudo MS^3 experiments'). Spectra were deconvoluted using the maximum entropy algorithm in MassLynx software (Waters, Manchester, UK).

High-resolution mass spectra were acquired using an LTQ Orbitrap XL ETD hybrid mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) equipped with an nano-ESI source. Samples were introduced into the spectrometer using platinum-coated borosilicate capillaries prepared in-house, using a spraying voltage of typically 1.5 kV. A mass resolution of 30 000 (at m/z 400) was used.

Chapter 4

Amphibian Peptides That Inhibit Neuronal Nitric Oxide Synthase

4.1 Introduction

NO is a labile gaseous free-radical that is produced *in vivo* by NO synthase (NOS) enzymes [311]. In low concentrations, NO is important as a biological second messenger, however, at higher concentrations, its major role as a second messenger is overcome by its cytotoxic properties [312]. NO regulates many important biological processes of the mammalian immune, circulatory and central nervous systems [313]. These include controlling vasomotor tone, neurotransmission, cell adhesion, platelet aggregation, and vascular smooth muscle proliferation [314–317]. Its hydrophobic nature ensures it can easily permeate membranes and trigger many biological outcomes. This is unlike other neurotransmitter molecules, which function by exocytosis rather than simple diffusion.

NO has an extremely short half-life ($\sim$ 5 sec), making it only available to nearby or adjacent cells [318, 319]. It is rapidly degraded by oxidation to NO_2/NO_3, scavenged by superoxide-generating compounds and can react with superoxide radicals to form the bioactive peroxynitrite anion ($ONOO^-$) [320, 321]. This all indicates that NO signalling is not regulated by storage of the signalling molecule, but rather at the synthesis stage itself [314, 319].

4.1.1 Nitric Oxide Synthesis

The labile, highly reactive nature of NO means that its synthesis by NOS must be carefully regulated. Three distinct NOS isoforms exist, namely neuronal NOS (nNOS), endothelial NOS (eNOS) and inducible NOS (iNOS) [314, 322]. These differ from each other by their expression mode, primary sequence and calcium dependence [323]. eNOS and nNOS are generally constitutively expressed and both operate in a Ca^{2+}-dependent manner, but iNOS irreversibly binds Ca^{2+}, and therefore its activity is mostly independent of intracellular Ca^{2+} levels [315]. Isoform expression differs between tissue types and interestingly, the eNOS isoform contains two post-translational modifications allowing it to localise at the membrane [312].

Each isoform exists as a homodimer [324], with each subunit 130-160 kDa in mass [311, 314, 325]. Each isoform is encoded on a different chromosome [326], but despite this, there is a high degree of sequence similarity both between the different isoforms and also between different species [327–329]. Each isoform contains the same basic domain structure, with an N-terminal catalytic oxygenase domain and a C-terminal reductase domain, separated by a calmodulin (CaM) binding region [330]. The C-terminal domain contains binding sites for nicotinamide adenine dinucleotide phosphate (NADPH), flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN), whilst the N-terminal region contains binding sites for heme (iron protoporphyrin IX), tetrahydrobiopterin (BH_4) and L-arginine (L-Arg) (Figure 4.1) [311, 314, 315, 328, 330].

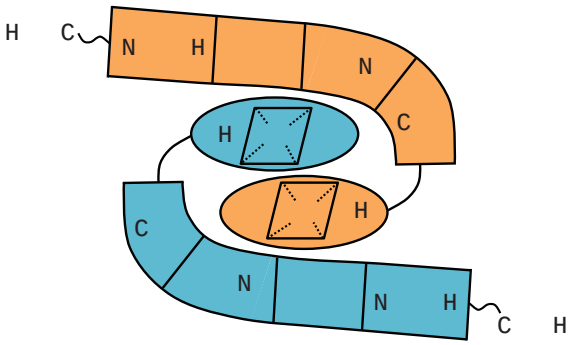


Figure 4.1: Schematic of the NOS dimer. The oxygenase domains are illustrated as ellipses. Binding sites are indicated. Adapted from [315, 330, 331].

NOS catalyses a two-step, five electron oxidation of L-Arg, producing L-citrulline and a stoichiometric quantity of NO. This process occurs though the stable intermediate

N^{ω}-hydoxyl-L-arginine (Figure 4.2) and utilises electrons from NADPH [312, 325, 329]. It has been shown that the nitrogen atom of NO comes from a guanidino nitrogen of L-Arg and the oxygen comes from O_2 [314]. Each NOS domain performs a different catalytic function. The reductase domain transfers electrons from NADPH to the oxygenase domain of the partner subunit, via the flavins FAD and FMN, where the heme domain is located [314]. The arrival of two electrons at the heme domain triggers the hydroxylation of L-Arg to the intermediate N^{ω}-hydoxyl-L-arginine. The remainder of the reaction mechanism is less well understood, but it is known that a further three electrons are required along with molecular oxygen [314]. Dimerisation of NOS enzymes is necessary for their activity, with the dimerisation interfaces located on the oxygenase domain of each subunit [330]. It is proposed that heme incorporation promotes dimerisation and that BH_4 is necessary for iNOS dimerisation, but it may not be essential for nNOS or eNOS dimerisation where it may act as a stabilising molecule [330]. CaM binding to NOS is essential for activation of the enzyme as it triggers electron transfer to the heme unit by altering the conformation of the reductase domain, thereby increasing the rate of electron transfer [313, 332–334].

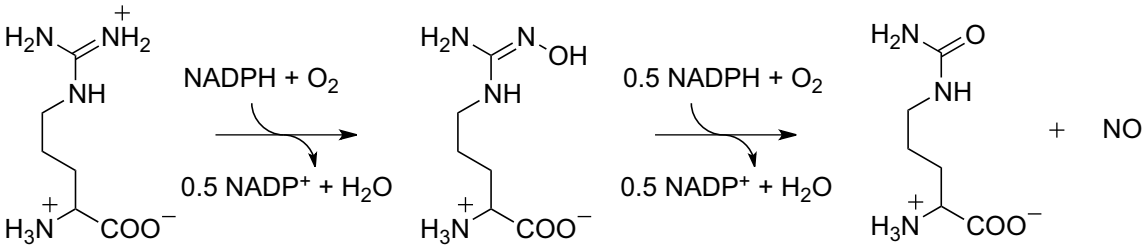


Figure 4.2: Mechanism of NO production by NOS enzymes. Adapted from [311, 312].

NO production is stringently regulated, albeit by different mechanisms for each NOS isoform. Both eNOS and nNOS are constitutively expressed and mostly regulated by intracellular Ca^{2+} levels [317, 320, 324]. Basal Ca^{2+} levels are insufficient for NOS activity, and influx of Ca^{2+} from the extracellular environment or from intracellular stores is quickly bound by CaM, triggering a conformational change in the protein which enables it to bind to NOS and activate the enzyme [315, 329]. This makes NO production largely the result of signalling cascades that result in Ca^{2+} influx. Conversely, CaM binds irreversibly to iNOS, making its activity largely independent of intracellular Ca^{2+} levels, however, increased intracellular Ca^{2+} levels do promote a slightly increased activity [315]. As a result, the

activity of this enzyme is generally regulated at the stage of transcription, which is induced by the receipt of a suitable stimulus [328, 335]. It is also thought that phosphorylation of nNOS may also play a role in regulating its activity [336], and it is proposed that when intracellular Ca^{2+} levels remain high for extended periods NO production is downregulated by phosphorylation of nNOS [336].

4.1.2 Calmodulin

CaM is a small, ubiquitous, highly conserved regulatory protein that is involved in many Ca^{2+}-dependent processes in most eukaryotic cells [337–340]. It has the ability to bind and activate numerous target proteins such as myosin light chain kinases, CaM-dependent protein kinases, calcineurin, NOS, the plasma membrane Ca^{2+} pump and a variety of ion channels [337, 341–343]. It is currently estimated that over 300 proteins contain CaM-binding motifs [344, 345]. Activation of these targets enables CaM to regulate physiological processes such as cell cycle control, muscle contraction, transcription, cyclic nucleotide metabolism, nitric oxide production and ion movement [346, 347].

CaM is able to perform its regulatory function due to its ability to act as a cytosolic Ca^{2+} sensor and bind four Ca^{2+} ions in a cooperative fashion with high affinity (K_d from 10^{-7} to 10^{-11} M) [337, 348]. Upon Ca^{2+} influx from the extracellular space or intracellular stores (such as the endoplasmic reticulum) the intracellular Ca^{2+} concentration increases from its resting level of 10^{-7} M to approximately 10^{-6} - 10^{-5} M resulting in Ca^{2+} binding at the four EF hand (Ca^{2+} binding domain) regions located at the lobes of CaM (to afford holo-CaM) [342]. This triggers a conformational change to the active configuration (Figure 4.3).

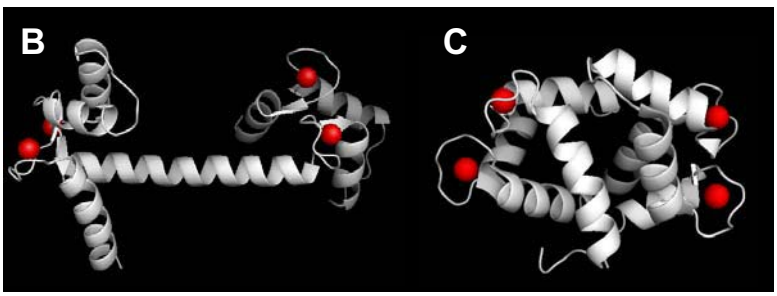


Figure 4.3: Schematic representation of the 3D structure of (A) apo-CaM (PDB: 1CFD) [349], (B) the dumbbell structure of holo-Ca_4^{2+}CaM (PDB: 1CLL) [350] and (C) a globular structure of holo-Ca_4^{2+}CaM (PDB: 1PRW) [351].

The first crystal structure of Ca_4^{2+}CaM was published in 1985 [352] and since this time numerous X-ray and NMR structures of the protein from a variety of species have been reported [350, 353–355]. A high degree of structural similarity is observed in all the reported structures, with many describing the characteristic dumbbell structure with N- and C- terminal globular domains separated by a central helical region (Figure 4.3B) [337, 356]. X-ray studies had initially shown that this central region adopted a complete 8 turn α-helix, however, there is significant deviation from normal α-helical dihedral angles in the centre of this helix and kinks have also been observed in some of the reported structures, suggesting a degree of flexibility. The presence of a central uncoiled linker region was confirmed by a number of techniques, including limited trypsin digestion, small angle X-ray scattering, hydrogen-deuterium exchange, nuclear Overhauser effect analysis, chemical shift analysis and NMR relaxation studies [357–362]. Additionally, structures of Ca_4^{2+}CaM have also been reported with more compact, globular conformations, demonstrating the wide array of conformational flexibility inherent in the protein (Figure 4.3C) [351].

CaM plays an important regulatory role in NO production by nNOS. Ca_4^{2+}CaM binds in the central amphipathic α-helical region connecting the reductase and oxygenase domains of the nNOS monomer (Figure 4.1) [325, 363, 364]. This is vital for activity, as binding increases both the rate of electron transfer from NADPH to the flavins in the reductase domain and the rate of electron receipt by heme in the oxygenase domain [313, 332, 333, 365]. The result is increased NO production.

Ca_4^{2+}CaM typically recognises a binding domain of 20-25 amino acid residues in its target proteins [342]. In addition to these proteins, CaM also binds numerous peptides and toxins with high affinity [342, 343]. These domains display no sequence similarity but are typically positively charged, contain hydrophobic residues and exhibit a propensity to form amphipathic α-helices upon binding to CaM [337, 338, 366, 367]. Hydrophobic interactions are the major associative force between CaM and its target molecule, with large hydrophobic residues present in conserved positions in the binding motif, typically at positions 1-5-10 or 1-8-14 [366]. These residues all lie on the same face of the helix which forms upon CaM association. Electrostatic interactions between the basic, positively charged peptide and the negatively charged protein assist in stabilising the interaction

[337, 366]. In addition, interactions between abundant Met residues on CaM and the peptide play a stabilising role [366]. These residues are particularly important because of the flexible nature of the side chain and the increased van der Waals interactions that result from the sulfur atom having a larger polarisability than carbon [366].

Significant conformational changes are observed in both Ca_4^{2+}CaM and its target upon complexation [368]. The complex which forms typically adopts either a so-called collapsed or extended structure. The canonical collapsed conformer is more compact than Ca_4^{2+}CaM [366] and its formation is largely driven by hydrophobic interactions between the anchor residues of the binding partner and the hydrophobic binding cavity formed when the N- and C- terminal regions of CaM come together in space [366, 369–371]. The binding channel formed contains charged outlets with the C-terminal end having a negative charge and the N-terminal end having both positively and negatively charged residues. These contribute to the binding by providing electrostatic interactions and also determine the orientation of the bound target [366]. A number of complexes of this type have been reported in the literature [157, 372–374], one example is illustrated in Figure 4.4A.

Complexes of Ca_4^{2+}CaM do not always collapse into an ellipsoidal conformation, but some remain in an extended structure [366]. For example, in the case of the CaM binding peptide from the plasma membrane Ca^{2+}-pump (C20W), the peptide is bound by the C-terminal lobe in an α-helical conformation by hydrophobic interactions (Figure 4.4B) [368]. In this instance, the two lobes of Ca_4^{2+}CaM are not in contact with each other and the binding partner is only bound at one of the lobes. A number of different CaM:peptide complexes have been shown to adopt this conformation, or some variant of it [366].

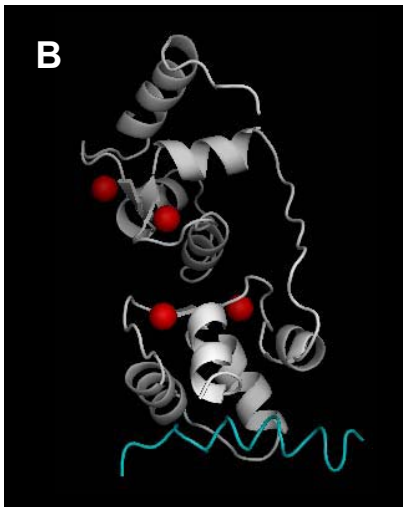


Figure 4.4: Schematic representation of Ca_4^{2+}CaM (white, Ca^{2+} are shown in red) bound to small peptides (blue) in (A) a collapsed conformation (PDB: 2BBM) [157] and (B) an extended conformation (PDB: 1CFF) [368].

4.1.3 Anuran Skin Secretions

The anuran (frog and toad) skin is a complex organ that performs numerous functions necessary for the animal's survival, including respiration, thermoregulation and self-defence [375]. Important for these functions is the presence of two distinct classes of cutaneous glands, namely mucous and granular, which are primarily distributed on the dorsal surface of the dermal layer, and release their contents through secretory ducts [376]. The animal controls skin moisture, lubrication and temperature through secretions of the mucous glands, whilst the contents of the granular glands is generally discharged in response to infection or predation [375–378]. Both types of glands synthesise and store the compounds of the secretion, the release of which is under the control of the sympathetic nervous system [379]. Secretions of the granular glands are an important component of the animal's defence arsenal, and are a vital to ensure the animal can survive in the hostile environment in which they live.

Anuran skin secretions contain a variety of diverse chemical components, including biogenic amines [380], bufadienolides [381], alkaloids [382] and peptides [383, 384], with broad ranging activities. These include antimicrobial, hormone, analgesic and neurotransmitter functions [384, 385]. The peptide components of the secretions are particularly known for their antimicrobial action but have also been shown to have

anticancer, antifungal and antiviral activities, demonstrating the potential broad-ranging pharmaceutical applications of self-defence compounds. Several treatments for a variety of medical conditions have been developed, or are in development, from both peptide and alkaloid components of anuran skin secretions [386–389]. Additionally, the development of peptide based antimicrobials is emerging as a vital area of research to combat resistance to current therapies and the amphibian skin represents a promising resource for lead compounds [390].

4.1.4 Peptides from Australian Frogs

The first extensive study pertaining to the skin secretions of Australian amphibians was reported in 1984 by Erspamer and co-workers [391]. In this study, the peptides in the skin extracts from 100 species of amphibians from Australia and Papua New Guinea were tested for their smooth muscle contraction activity and systemic blood pressure modulation. Of the species studied, the most abundant peptides were found to be those of the caerulein, bombesin and tachykinin families [391].

Since this seminal work, the skin secretions of many other species of Australian amphibians have been investigated, and their peptide components identified and tested [384, 392–394]. In particular, species from the genus *Litoria* have been studied in addition to those from the genera *Limnodynastes, Uperoleia* and *Crinia*. Most of the secretions that have been investigated contain a broad range of peptide host-defence compounds, many with multifaceted activities, including neuropeptides, narrow and broad spectrum antimicrobials and nNOS inhibitory peptides. Peptides with antiviral, anticancer, antifungal and pheromone activities have also been identified, whilst the activity of numerous peptides remain unknown [384, 392–394]. The sequences and activities of peptides isolated from Australian frogs are multifaceted and highly variable, as shown in Table 4.1.

Table 4.1: Selected peptides isolated from the skin secretions of Australian amphibians. Adapted from [384, 392].

Name	**Sequence**	**Species**[1]	**Activity**[2]
Aurein 1.1	GLFDIIKKIAESI-NH_2	a	1,2
Aurein 1.2	GLFDIIKKIAESF-NH_2	a	1,2
Aurein 2.1	GLLDIVKKVVGAFGSL-NH_2	a,b	1,2
Caeridin 1.1	GLLαDGLLGTGL-NH_2	c,d,e,f,g	
Caeridin 1.2	GLLβDGLLGTGL-NH_2	e	
Caeridin 1.4	GLLαDGLLGGLGL-NH_2	f,g	
Caeridin 1.5	GLLβDGLLGGLGL-NH_2	f,g	
Caerin 1.1	GLLSVLGSVAKHVLPHVVPVIAEHL-NH_2	c,d,e	1,2,3,4,5
Caerin 1.4	GLLSSLGSVAKHVLPHVVPVIAEHL-NH_2	d,e	1
Caerin 1.8	GLFKVLGSVAKHLLPHVVPVIAEKL-NH_2	g	1,2,3,5
Caerin 1.9	GLFGVLGSIAKHVLPHVVPVIAEKL-NH_2	g	1,2,3,4,5
Caerulein 1.1	pEQDY(SO_3)TGWMDF-NH_2	h	6
Caerulein 1.2	pEQDY(SO_3)TGWFDF-NH_2	c,i	6
Citropin 1.1	GLFDVIKKVASVIGGL-NH_2	i	1,2,3,5
Citropin 1.2	GLFDIIKKVASVVGGL-NH_2	i	1,2,3,5
Citropin 1.3	GLFDIIKKVASVIGGL-NH_2	i	1,2,3,5
Dahlein 1.1	GLFDIIKNIVSTL-NH_2	j	1
Dahlein 5.6	GLLASLGKVFGGYLAEKLKPK-OH	j	5
Frenatin 3	GLMSVLGHAVGNVLGGLFKPKS-OH	k	5
Lesueurin	GLLDILKKVGKVA-NH_2	l	5
Maculatin 1.1	GLFGVLAKVAAHVVPAIAEHF-NH_2	m	1,2,3,5
Maculatin 1.3	GLLGLLGSVVSHVVPAIVGHF-NH_2	n	1,2
Maculatin 2.1	GFVDFLKKVAGTIANVVT-NH_2	n	1,2
Riparin 1.1	RLCIPVIFPC-OH	o	6
Riparin 2.1	IIEKLVNTALGLLSGL-NH_2	o	1
Signiferin 1	RLCIPYIIPC-OH	p	6
Signiferin 2.1	IIGHLIKTALGMLGL-NH_2	p	1
Signiferin 3.1	GIAEFLNYIKSKA-NH_2	p,o	5
Splendipherin	GLVSSIGKALGGLLADVVKSKGQPA-OH	c,d	7
Tryptophyllin L 1.3	pEFPWL-NH_2	q	6
Uperin 1.1	pEADPNAFYGLM-NH_2	r	6
Uperin 2.1	GIVDFAKKVVGGIRNALGI-NH_2	r	1
Uperin 3.5	GVGDLIRKAVSVIKNIV -NH_2	s	1
Uperolein	pEPDPNAFYGLM-NH_2	t	6

[1]**Species: (a)** *Litoria raniformis* [395]; **(b)** *Litoria aurea* [395]; **(c)** *Litoria splendida* [396]; **(d)** *Litoria caerulea* [397]; **(e)** *Litoria gilleni* [398]; **(f)** *Litoria xanthomera* [399]; **(g)** *Litoria chloris* [400]; **(h)** many species from the genus *Litoria* [378, 401]; **(i)** *Litoria citropa* [402]; **(j)** *Litoria dahlii* [403]; **(k)** *Litoria infrafrenata* [404]; **(l)** *Litoria lesueri* [405]; **(m)** *Litoria genimaculata* [406]; **(n)** *Litoria eucnemis* [407]; **(o)** *Crinia riparia* [408]; **(p)** *Crinia signifera* [409]; **(q)** *Litoria rubella* [410]; **(r)** *Uperoleia inundata* [411]; **(s)** *Uperoleia mjobergii* [412]; **(t)** many species of the genus *Uperoleia* [378].

[2]**Activity: (1)** antimicrobial; **(2)** anticancer; **(3)** antifungal; **(4)** antiviral; **(5)** nNOS inhibitor; **(6)** neuropeptide; **(7)** pheromone.

It has been shown that glandular secretions of many Australian frogs contain at least one wide-spectrum antibiotic peptide along with numerous other peptides with narrow-spectrum activities [384, 392]. These ensure effective protection against the plethora of bacteria that are rife in the hostile environment in which they live. Some of the most potent antimicrobial peptides isolated are caerin 1.1, maculatin 1.1 and aurein 1.2 [384, 392]. Many of these antimicrobials also show anticancer and antiviral activities. This is discussed in further detail in Chapter 5.

Neuropeptides also form an integral part of the host-defence arsenal in frogs of the genus *Litoria* and toadlets of the genus *Uperoleia*, in addition to playing a regulatory role in the animal [384, 392, 394]. Amongst the most potent of this class are the caeruleins and uperins 1. Caerulein, which contains a sulfated Tyr residue, pyroglutamate and amidated C-terminus, has many activities, including smooth muscle contraction, thermoregulation and analgesic properties [384, 392]. Peptides which inhibit NO production by nNOS (discussed in Section 4.1.5) are also found in the secretions [413], along with numerous peptides with as yet unknown activities, such as the caeridins.

4.1.5 Amphibian Peptides that Inhibit nNOS

Most amphibians of the genera *Litoria* and *Crinia* that have been studied thus far produce one or more peptide(s) that inhibit nNOS as part of their glandular secretion, with current studies identifying more than 50 such peptides [384, 392]. It is thought that peptides of this type may play a regulatory role in the animals, as it is known that NO is involved in anuran sight, gastric modulation and reproduction [414–416]. Additionally, they may also form an integral part of the animal's host-defence mechanisms by interfering with NO signalling in an attacking predator or pathogen [413].

The nNOS inhibitory peptides isolated thus far can be divided into three inhibitory groups based on their common structural motifs. The first of these, the caerin 1 peptides, adopt a helix-hinge-helix structure and often contain a Phe residue at position 3 [385, 417]. Second are the aurein/citropin peptides, which are short linear amphipathic α-helices, and third are the frenatin/dahlein type peptides which have a characteristic C-terminal free

acid and a Lys-X-Lys motif near their C-terminus [384]. A number of other peptides also exist which cannot be categorised by this classification system (Table 4.2).

Evidence suggests that the amphibian peptides inhibit NO production by interacting with Ca^{2+} CaM [413]. This includes *in vitro* studies where dose-response curves showed a Hill slope greater than one, indicating a non-competitive interaction is causing nNOS inhibition [418]. Also, it has been shown that the addition of excess CaM is able to partially rescue nNOS activity in an amphibian peptide treated sample, and that these peptides also inhibit the function of calcineurin, another CaM regulated enzyme [413]. Finally, NMR and MS data have shown that the peptides complex with CaM in a high-affinity fashion [413, 419]. Together, these data suggest that attenuation of NO production is achieved by the peptide binding to CaM triggering a conformational change, and thereby rendering the protein unable to associate with nNOS.

Table 4.2: nNOS inhibition activities of selected amphibian peptides and synthetic derivatives [384, 392, 413].

Name	Species	Sequence	IC_{50} (μM)	Charge
Caerin 1 Peptides				
Caerin 1.1	*L. caerulea* *L. splendida* *L. gilleni*	GLLSVLGSVAKHVLPHVVPVIAEHL-NH_2	36.6	+1
Caerin 1.6	*L. chloris*	GLFSVLGAVAKHVLPHVVPVIAEKL-NH_2	8.5	+2
Caerin 1.8	*L. chloris*	GLFKVLGSVAKHLLPHVVPVIAEKL-NH_2	1.7	+3
Caerin 1.8.1	*L. chloris*	GSVAKHLLPHVVPVIAEKL-NH_2	Inactive	+2
Caerin 1.8 (mod. 2)		GLFKVLGSVAKHLLPHVVP-NH_2	1.5	+3
Caerin 1.8 (mod. 3)		GLFKVLGSVAKHLLPHV-NH_2	1.7	+3
Caerin 1.8 (mod. 4)		GLFKVLGSVAKHLLP-NH_2	3.1	+3
Caerin 1.8 (mod. 5)		GLFKVLGSVAKHL-NH_2	3.7	+3
Caerin 1.8.11		GLFKVLGSVAK-NH_2	3.3	+3
Caerin 1.8 (mod. 8)		GLFKVLGS-NH_2	> 850	+2
Caerin 1.19	*L. gracilenta*	GLFKVLGSVAKHLLPHVAPIIAEKL-NH_2	4.1	+3
Caerin 1.19.3	*L. gracilenta*	GSVAKHLLPHVAPIIAEKL-NH_2	Inactive	+2
Citropin/Aurein Peptides				
Citropin 1.1	*L. citropa*	GLFDVIKKVASVIGGL-NH_2	8.2	+2
Citropin 1.1 d		GlfdvikkvasviGGL-NH_2	30.7	+2
Citropin 1.1 (mod. 13)		GLFDVIKKVASVIKKL-NH_2	2.0	+4
Citropin 1.1 (mod. 18)		GLFAVIKKVASVIKKL-NH_2	1.2	+5
Aurein 1.1	*L. aurea*	GLFDIIKKIAESI-NH_2	33.9	+1
Aurein 2.2	*L. aurea*	GLFDIVKKVVGALGSL-NH_2	4.3	+2
Aurein 2.4	*L. aurea*	GLFDIVKKVVGTLAGL-NH_2	2.1	+2
Aurein 2.3	*L. aurea*	GLFDIVKKVVGIAGSL-NH_2	1.8	+2
Lesueurin	*L. lesueuri*	GLLDILKKVGKVA-NH_2	16.2	+3
Frenatin/Dahlein Peptides				
Frenatin 3	*L. infrafrenata*	GLMSVLGHAVGNVLGGLFKPKS-OH	6.8	+3
Frenatin 3 (mod.3)		GLMRVLGHAVGNVLGGLFKPKS-OH	1.4	+3
Dahlein 5.1	*L. dahlii*	GLLGSIGNAIGAFIANKLKP-OH	3.2	+2
Dahlein 5.2	*L. dahlii*	GLLASIGKVLGGYLAEKLKP-OH	1.2	+2
Dahlein 5.3	*L. dahlii*	GLLASLGKVFGGYLAEKLKP-OH	1.4	+2
Dahlein 5.6	*L. dahlii*	GLLASLGKVFGGYLAEKLKPK-OH	1.6	+3
Splendipherin	*L. splendida*	GLVSSIGKALGGLADVVKSKGQPA-OH	9.0	+2
Miscellaneous Peptides				
Fallaxidin 3.1	*L. fallax*	GLLASLGKVFGGYLAEKLKPK-OH	15.4	+1
Dahlein 4.2	*L. dahlii*	GLWQFIKDKIKDAARTGLVTGIQS-NH_2	11.1	+2
Deserticolin 1	*C. deserticola*	GLADFLNKAVGGKVVDFVKS-NH_2	2.4	+1

4.2 Aims

In order to design improved nNOS inhibitors based on the amphibian peptides, it is vital to gain further insight into their binding properties and the mechanism by which they inhibit nNOS. The aims of this chapter were to use a variety of biophysical techniques (Chapter 2), including ITC, CD, IM-MS and NMR spectroscopy to afford low resolution structural information about the complexes of Ca_4^{2+}CaM with the amphibian peptides caerin 1.8, caerin 1.8.11, citropin 1.1 (mod. 13) and dahlein 5.6 (Table 4.2 and 4.3). These peptides represent each of the three classes of nNOS active amphibian peptides. Caerin 1.8.11 was investigated as it is the smallest fragment of caerin 1.8 which retains significant nNOS inhibition activity.

These data will provide significant insight into the mechanism by which these peptides interact with Ca_4^{2+}CaM, in terms of the strength of the interactions and the overall conformation of the complexes.

Table 4.3: nNOS inhibition activities of the amphibian peptides studied [384, 392, 413].

Name	**Sequence**	**IC_{50} (μM)**	**Charge**
Caerin 1.8	GLFKVLGSVAKHLLPHVVPVIAEKL-NH_2	1.7	+3
Caerin 1.8.11	GLFKVLGSVAK-NH_2	3.3	+3
Citropin 1.1 (mod. 13)	GLFDVIKKVASVIKKL-NH_2	2.0	+4
Dahlein 5.6	GLLASLGKVFGGYLAEKLKPK-OH	1.6	+3

4.3 Results

4.3.1 Expression and Purification of Calmodulin

CaM was either obtained commercially (for MS studies) or expressed recombinantly (for ITC and NMR studies). CaM was expressed in *E. coli* BL21 DE3 cells with no fusion partner or affinity tags and purified by hydrophobic interaction chromatography (HIC), according to literature procedures [420, 421]. HIC is an established technique to purify CaM and other Ca^{2+} binding proteins, as apoCaM has a relatively hydrophilic surface but upon binding Ca^{2+} (to form holoCaM) a hydrophobic pocket is exposed [422].

In order to purify the recombinantly expressed CaM, holoCaM was introduced to the HIC column, with the hydrophobic surface allowing it to bind strongly. Other native hydrophilic proteins, which do not bind to the hydrophobic media were washed from the column. Elution of CaM was achieved by chelating Ca^{2+} with ethylene glycol-bis(β-aminoethylether)-N,N'-tetraacetic acid (EGTA), triggering a conformational change in CaM, hiding the hydrophobic pocket. This procedure yielded CaM of > 95 % purity (as determined by SDS-PAGE and MS).

4.3.2 Circular Dichroism Spectroscopy

Ca_4^{2+}CaM is a highly helical protein, comprising eight helical domains. It is well recognised that when a Ca_4^{2+}CaM binds its peptide targets, it induces amphipathic α-helical structure in the bound peptide [366, 423–425]. Conversely, the secondary structure of Ca_4^{2+}CaM remains relatively unaffected by binding to its peptide targets.

The conformational changes induced in the amphibian peptides caerin 1.8.11, dahlein 5.6 and citropin 1.1 (mod. 13) were studied by far-UV CD measurements. In solution the peptides themselves have CD spectra which are characteristic of a random coil conformation with a minimum at approximately 195 nm (data not shown).

The spectrum of Ca_4^{2+}CaM (Figure 4.5) in solution shows two characteristic minima at approximately 208 and 222 nm. In the presence of one equivalent of peptide, the minima

at these wavelengths increase in intensity. These data indicate there is an increase in helical content upon binding, which can be attributed to the induced secondary structure of the peptide targets.

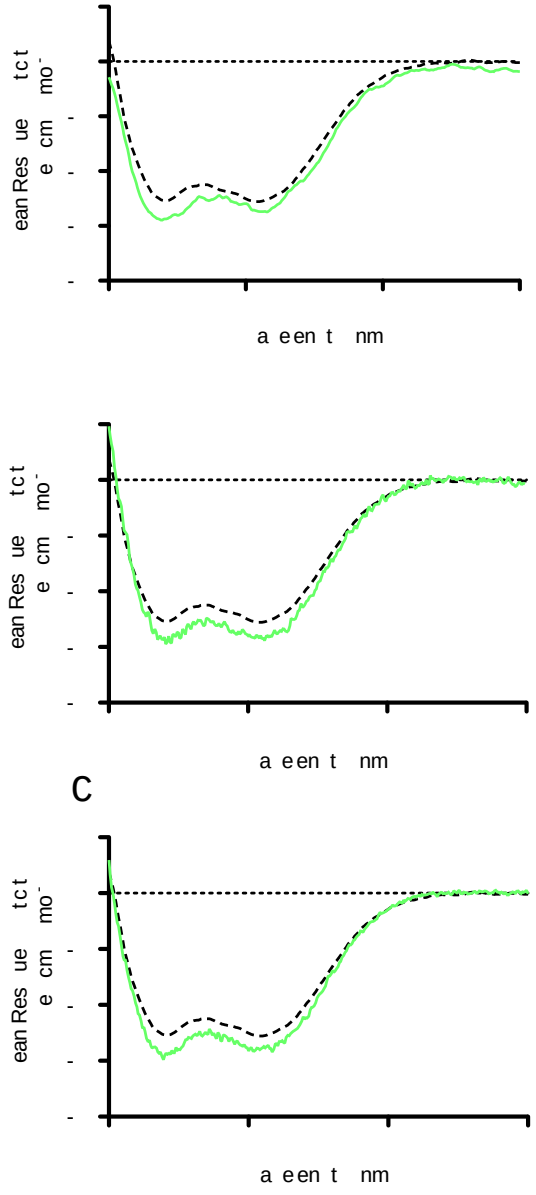


Figure 4.5: Far-UV CD spectra of Ca_4^{2+}CaM (black traces, 2 μM, in 10 mM Tris-HCl, 10 mM KCl, 5 mM $CaCl_2$, pH 6.3) and in the presence of (A) caerin 1.8.11, (B) dahlein 5.6 and (C) citropin 1.1 (mod. 13) (green traces, 1 molar equivalent).

4.3.3 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) was employed to determine thermodynamic and kinetic parameters for the binding events between CaM and three amphibian peptides, namely caerin 1.8.11, dahlein 5.6 and citropin 1.1 (mod. 13). Each titration was repeated at four different temperatures and the enthalpy change of the interaction (ΔH) was determined at each temperature.

Figures 4.6 to 4.8 display the binding data for the interactions of caerin 1.8.11, dahlein 5.6 and citropin 1.1 (mod. 13) with Ca_4^{2+}CaM at 40 ^{o}C. The enthalpy of dilution was subtracted from the data by performing a buffer into protein control experiment (data not shown). Tables 4.4 to 4.6 display the thermodynamic and kinetic parameters obtained by curve fitting to the measured data.

The data were fit using a single site binding model and in all cases indicate a 1:1 stoichiometry of binding, with fitted stoichiometries between 0.97 and 1.23 for all titrations. In the case of citropin 1.1 (mod. 13) (Figure 4.8, Table 4.6), at low temperatures the binding reaction is endothermic, becoming exothermic with increasing temperatures. A large favourable entropy change drives the binding interaction at lower temperatures. The interactions of caerin 1.8.11 (Figure 4.6, Table 4.4) and dahlein 5.6 (Figure 4.7, Table 4.5) with Ca^{2+} CaM are driven by large, favourable enthalpy changes at all the tested temperatures. This is consistent with other calorimetric studies of Ca_4^{2+}CaM binding in the literature [426–428].

Fitted dissociation constants (K_ds) for the three peptides at all temperatures were within the range of 20-70 nM, demonstrating the strength of the binding interaction. These data show that the interactions of the amphibian peptides with CaM are very strong, with low nanomolar K_d values.

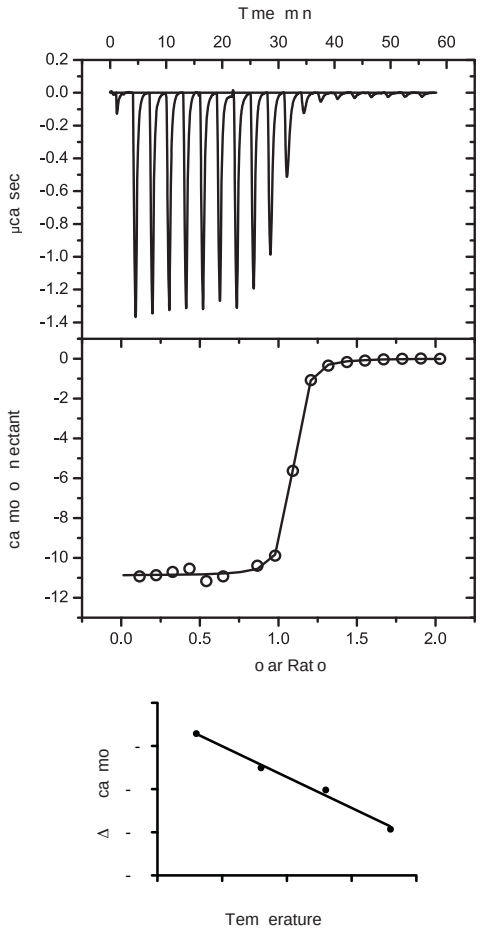


Figure 4.6: Calorimetric analysis of the binding of caerin 1.8.11 to Ca_4^{2+}CaM. (A) Top panel: The calorimetric titration at 40 ^{o}C. Lower panel: Binding isotherm obtained by integrating the above data. (B) Plot of ΔH as a function of temperature.

Table 4.4: Thermodynamic and kinetic parameters for the titration of Ca_4^{2+}CaM with caerin 1.8.11.

Temp (K)	K_d (μM)	ΔH (kcal/mol)	-TΔS (kcal/mol)	n
303.15	0.03 $\pm$ 0.008	-3.56 $\pm$ 0.03	-6.9	1.03 $\pm$ 0.004
308.15	0.05 $\pm$ 0.010	-7.52 $\pm$ 0.05	-2.8	1.05 $\pm$ 0.004
313.15	0.04 $\pm$ 0.006	-10.88 $\pm$ 0.05	0.23	1.04 $\pm$ 0.003
318.15	0.03 $\pm$ 0.002	-14.64 $\pm$ 0.03	3.8	1.04 $\pm$ 0.001

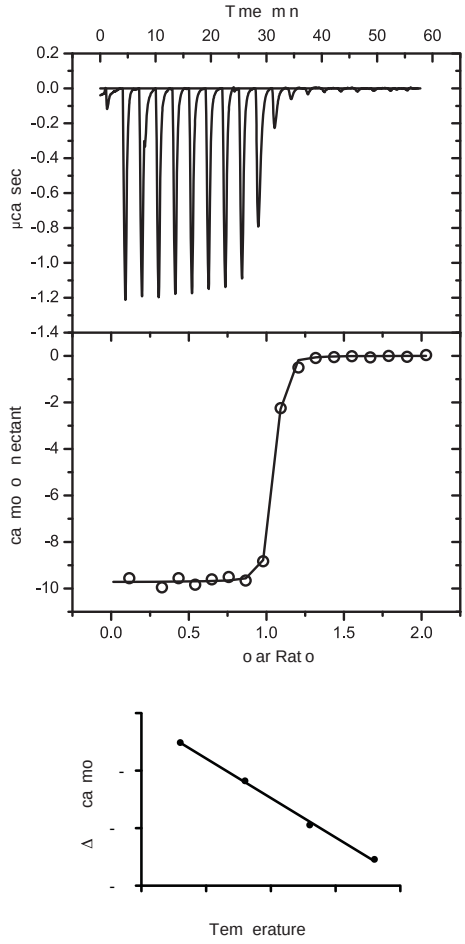


Figure 4.7: Calorimetric analysis of the binding of dahlein 5.6 to Ca_4^{2+}CaM. (A) Top panel: The calorimetric titration at 40 ^{o}C. Lower panel: Binding isotherm obtained by integrating the above data. (B) Plot of ΔH as a function of temperature.

Table 4.5: Thermodynamic and kinetic parameters for the titration of Ca_4^{2+}CaM with dahlein 5.6.

Temp (K)	K_d (μM)	ΔH (kcal/mol)	-TΔS (kcal/mol)	n
303.15	0.03 ± 0.004	-2.55 ± 0.04	-9.2	1.11 ± 0.07
308.15	0.02 ± 0.006	-5.88 ± 0.05	-5.1	1.07 ± 0.004
313.15	0.01 ± 0.003	-9.72 ± 0.05	-1.6	0.99 ± 0.002
318.15	0.03 ± 0.009	-12.71 ± 0.15	1.7	1.10 ± 0.06

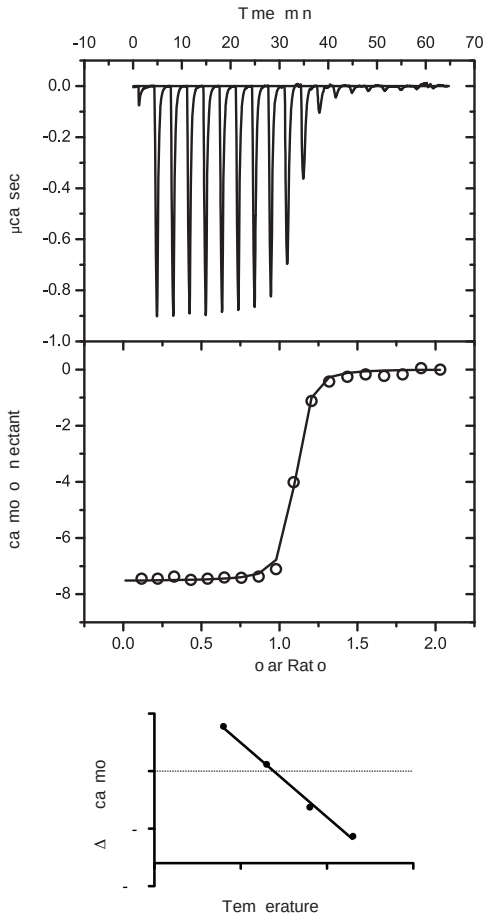


Figure 4.8: Calorimetric analysis of the binding of citropin 1.1 (mod. 13) to Ca_4^{2+}CaM. (A) Top panel: The calorimetric titration at 40 ^{o}C. Lower panel: Binding isotherm obtained by integrating the above data. (B) Plot of ΔH as a function of temperature.

Table 4.6: Thermodynamic and kinetic parameters for the titration of Ca_4^{2+}CaM with citropin 1.1 (mod. 13).

Temp (K)	K_d (μM)	ΔH (kcal/mol)	-TΔS (kcal/mol)	n
298.15	0.07 $\pm$ 0.02	3.89 $\pm$ 0.08	-12.2	1.23 $\pm$ 0.02
303.15	0.06 $\pm$ 0.02	0.59 $\pm$ 0.01	-9.2	0.91 $\pm$ 0.02
308.15	0.02 $\pm$ 0.004	-3.32 $\pm$ 0.02	-7.6	0.97 $\pm$ 0.003
313.15	0.05 $\pm$ 0.02	-5.65 $\pm$ 0.05	-4.8	1.09 $\pm$ 0.008

After plotting ΔH as a function of temperature (Figures 4.6B, 4.7B and 4.8B), the heat capacity change upon binding (ΔC_p) was estimated by measuring the slope of the resultant graph. The ΔC_p of an interaction can provide information about the binding mechanism [427]. Table 4.7 displays the estimated ΔC_ps for the binding event between Ca_4^{2+}CaM and the studied amphibian peptides. These ΔC_p values for the three peptide systems studied are similar, indicating a common binding mechanism.

Table 4.7: Approximated ΔC_p values calculated from ITC experiments.

Peptide	**ΔC_p (kJ/mol)**
Caerin 1.8.11	-3.00 ± 0.21
Dahlein 5.6	-2.87 ± 0.09
Citropin 1.1 (mod. 13)	-2.71 ± 0.15

4.3.4 Ion Mobility-Mass Spectrometry

To gain further insight into changes in the conformations of CaM, its Ca^{2+} adducts and the amphibian peptide:CaM complexes, IM-MS studies were performed to determine the changes in CCS that occur upon complexation. As a reference, Table 4.8 summarises the *m/z* of ions expected in the ESI-MS spectra of CaM and the peptide:CaM complexes which are studied in subsequent sections.

Table 4.8: Ions expected for the mass spectrometric investigations of CaM and peptide:CaM complexes.

Charge	***m/z***							
	CaM	**Ca_2^{2+}CaM**	**Ca_4^{2+}CaM**	**Ca_4^{2+}CaM + caerin 1.8**	**Ca_4^{2+}CaM + caerin 1.8.11**	**Ca_4^{2+}CaM + citropin 1.1 (mod. 13)**	**Ca_4^{2+}CaM + dahlein 5.6**	**Ca_4^{2+}CaM + Plasma Pump C20W**
0	16790	16866	16942	19604	18059	18697	19129	19362
1+	16791	16867	16943	19605	18060	18698	19130	19363
2+	8396	8434	8472	9803	9031	9350	9566	9682
3+	5598	5623	5648	6536	6021	6233	6377	6455
4+	4199	4218	4237	4902	4516	4675	4783	4842
5+	3359	3374	3389	3922	3613	3740	3827	3873
6+	2799	2812	2825	3268	3011	3117	3189	3228
7+	2400	2410	2421	2802	2581	2672	2734	2767
8+	2100	2109	2119	2452	2258	2338	2392	2421
9+	1867	1875	1883	2179	2008	2078	2126	2152
10+	1680	1688	1695	1961	1807	1871	1914	1937
11+	1527	1534	1541	1783	1643	1701	1740	1761
12+	1400	1407	1413	1635	1506	1559	1595	1615
13+	1293	1298	1304	1509	1390	1439	1472	1490
14+	1200	1206	1211	1401	1291	1337	1367	1384
15+	1120	1125	1130	1308	1205	1247	1276	1292
16+	1050	1055	1060	1226	1130	1170	1197	1211
17+	989	993	998	1154	1063	1101	1126	1140
18+	934	938	942	1090	1004	1040	1064	1077
19+	885	889	893	1033	951	985	1008	1020
20+	841	844	848	981	904	936	957	969

4.3.4.1 Calmodulin structure by IM-MS

The positive ion nanoESI IM-MS spectrum of CaM (10 μM, in 20 mM ammonium acetate, 40 μM calcium acetate) is shown in Figure 4.9. The experimentally determined mass of CaM was 16790 Da, which is in agreement with the expected mass based on the primary sequence with an acetylated N-terminus and trimethyl-Lys residue at Lys115. Two gaussian charge-state distributions can be clearly seen within the spectrum centred around charge states $[M+13H]^{13+}$ and $[M+7H]^{7+}$, indicating two different conformational states of Ca_n^{2+}CaM are present. As mild electrospray conditions were used to maintain native protein structures (low voltages and source temperatures, see Experimental Procedures), it can be assumed that the multiple conformational states seen in the gas phase reflect the solution-phase flexibility of CaM, which allows it to adopt a variety of conformations. This assumption is consistent with other studies [419, 429–431]. Above the conventional mass spectrum is a drift plot, showing the drift time of the ions present in the sample (Figure 4.9). Again, two different conformational states can be observed, based on the two distributions in the drift plot.

ATDs were extracted from the spectra for CaM species with differing Ca^{2+} occupation. The drift times were converted to CCS using a standard calibration protocol [43] and are displayed in Table 4.9. The ATDs observed in these experiments were often broad, likely due to the inherent structural flexibility of CaM.

To minimise errors in the conversion process, CCSs were averaged over all the recorded IM-MS spectra, which were each performed using different IM parameters (wave heights were varied between 6 - 9 V whilst maintaining the wave velocity at 400 ms^{-1}). The error associated with the measured CCSs is estimated to be approximately 8-10 %, taking into consideration errors associated with calibrant cross-sections, the calibration curve and the relative precision of replicate measurements [43].

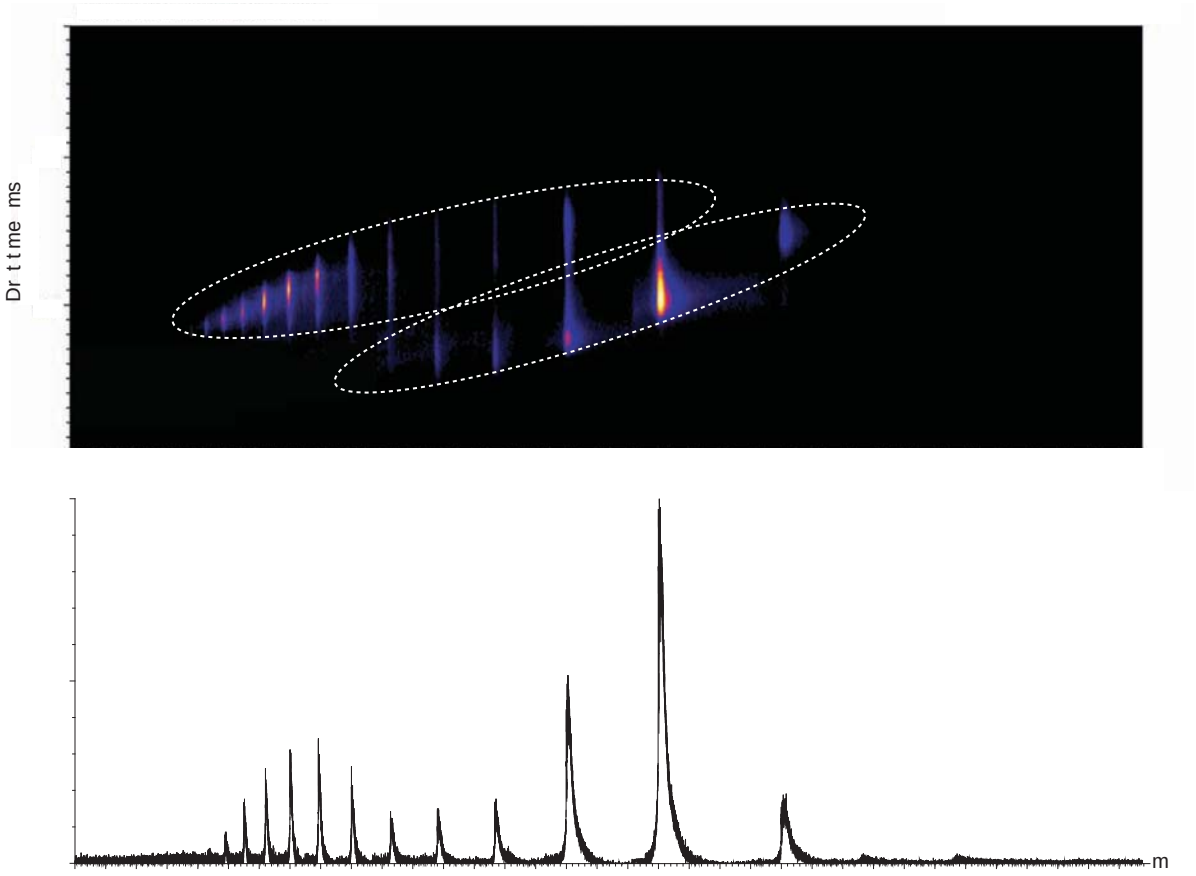


Figure 4.9: IM-MS spectrum of Ca^{2+}CaM (10 μM CaM, in 20 mM ammonium acetate, 40 μM calcium acetate). The lower panel shows the conventional mass spectrum, whilst the top panel shows the IM drift plot indicating the drift times of individual ions. The two distribution are indicated in the drift plot.

Table 4.9: Experimental collision cross-sections for Ca_n^{2+}CaM complexes as a function of charge state.Two CCSs are reported for ATDs with two major features. An estimate of the error in CCS measurement is 8-10 % [43]. Where no CCS is given, ATDs were not detected or of poor signal-to-noise.

Charge	**CaM CCS (Å^2)**				
	Apo	**1 Ca^{2+}**	**2 Ca^{2+}**	**3 Ca^{2+}**	**4 Ca^{2+}**
7+	1954	1954	1938	1938	1856
8+	1990	1974	1961	1962	1951
9+	2178	2085	2055	2034	2131
9+	2391	2380	2341	2358	2221
10+	2605	2610	2616		2639
11+	2776	2772	2781	2754	2754
12+	3017	2983	2972	2891	2914
13+	3106	3082	3079	3049	3059
14+	3199	3183	3148	3155	3148
15+	3316	3274	3263	3267	3250
16+	3384	3369	3376	3373	

4.3.4.2 Structural Analysis of CaM Complexes by IM-MS

In a similar fashion, complexes of CaM with caerin 1.8, caerin 1.8.11, dahlein 5.6, citropin 1.1 (mod. 13) and the 20 residue binding domain (C20W) of the plasma membrane Ca^{2+} pump (sequence shown below) were investigated by IM-MS, as shown in Figure 4.10. In all cases, the complexes exhibited a 1:1 stoichiometric ratio with at least four Ca^{2+} ions required for binding, except for C20W in which only two Ca^{2+} ion were required for binding. The experimentally determined masses of all the complexes compare favourably with the theoretically determined values (as outlined in Table 4.8). It is evident from the narrow charge state distributions of the protein complexes that they are much more well-defined in structure than free Ca_4^{2+}CaM, which adopts a variety of charge states.

C20W LRRGQILWFRGLNRIQTQIK-OH

Again, ATDs were extracted from the spectra for CaM:peptide complexes and the drift times were converted to CCS [43]. The ATDs observed in these experiments were much better defined and narrower than those of CaM and its Ca^{2+} adducts. This indicates the structures of the complexes are less flexible than those of CaM. The calibrated CCSs are shown in Table 4.10.

Table 4.10: Experimental collision cross-sections for Ca^{2+} CaM: peptide complexes as a function of charge state. An estimate of the error in CCS measurement is 8-10 % [43].

Peptide	**Charge**	**CCS (Å^2)**
Caerin 1.8	7+	2211
	8+	2191
Caerin 1.8.11	7+	2022
	8+	2036
Citropin 1.1 (mod. 13)	7+	2085
	8+	2096
Dahlein 5.6	7+	2136
	8+	2134
Plasma Pump C20W	8+	2186
	9+	2212

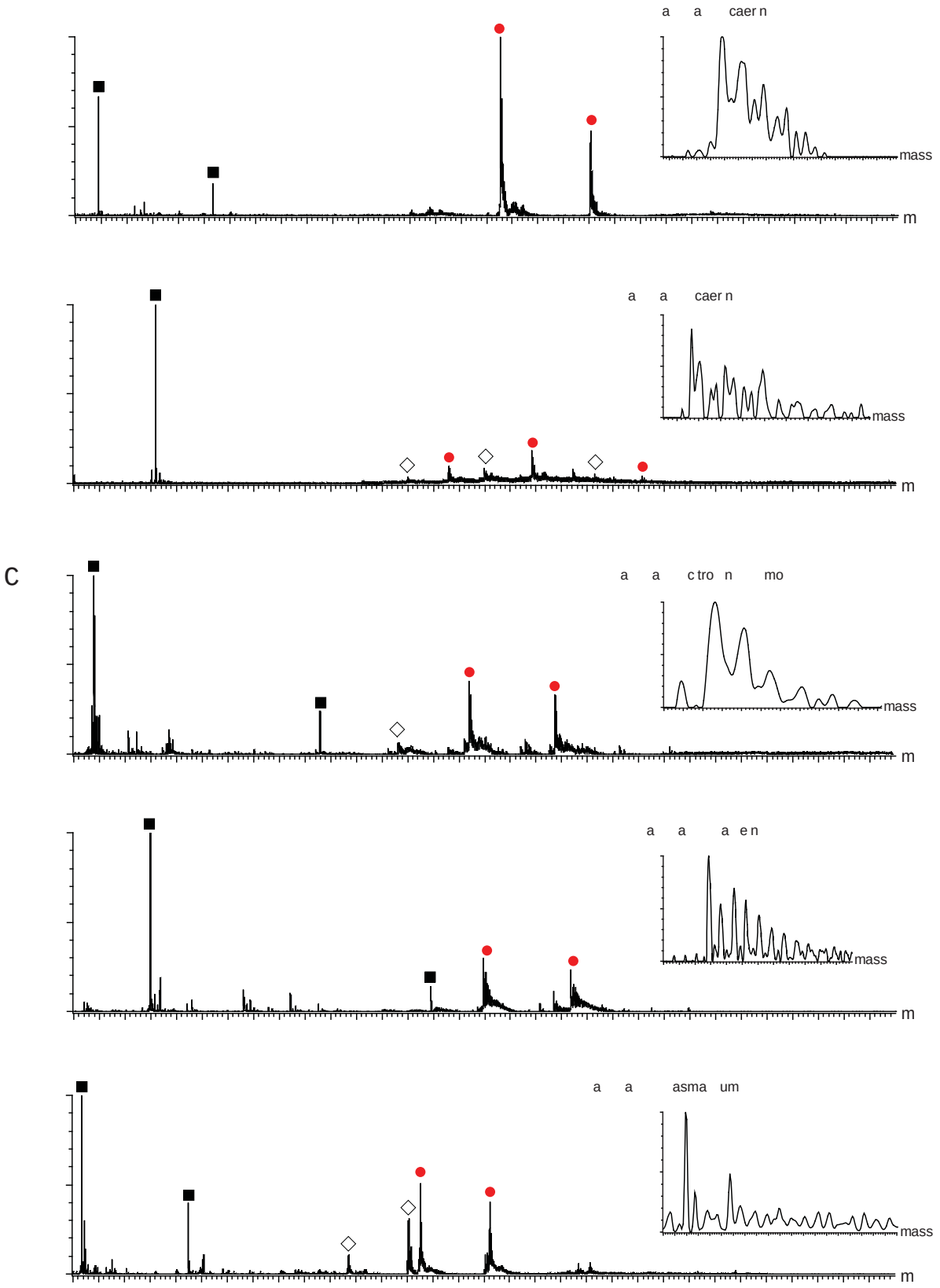


Figure 4.10: MS spectra of Ca^{2+} CaM with (A) caerin 1.8, (B) caerin 1.8.11, (C) citropin 1.1 (mod. 13), (D) dahlein 5.6, and (E) plasma pump C20W. Free peptide (■); free CaM (◇); peptide-CaM complexes (●). The insets show the transformed spectra. Synapt HDMS Mass Spectrometer.

4.3.4.3 Theoretical Calculation of CaM Cross-Sections

The program MOBCAL was used to calculate theoretical CCSs of high resolution protein structures from coordinates in the PDB for comparison with experimental values (Table 4.11) [120, 121]. The PA and EHSS methodologies were utilised for this purpose. Both values are reported here as it is known that the PA method generally underestimates CCS whilst the EHSS often overestimates the actual CCS [42, 122]. Consequently, these values provide lower and upper bounds for the experimentally determined CCS.

Table 4.11: Computationally determined CCS of CaM and selected complexes.

Complex	**PDB**	**Calculated CCS (Å^2)**		**Binding**	**Ref.**
		PA	**EHSS**	**Mode**	
apoCaM	1DMO	1627	2048	-	[432]
holoCaM (dumbbell)	1CLL	1622	2020	-	[350]
holoCaM (globular)	1PRW	1311	1650	-	[351]
CaM:NR1C1	2HQW	1421	1791	Collapsed	[433]
CaM:C28	2KNE	1820	2313	Collapsed	[434]
CaM:MLCK Peptide	2BBM	1770	2257	Collapsed	[157]
CaM:CaMKK Peptide	1CKK	1564	1985	Collapsed	[374]
CaM:C20W Peptide	1CFF	1862	2361	Extended	[368]

4.3.5 NMR Spectroscopy

To further investigate the complexation of Ca_4^{2+}CaM with caerin 1.8.11, NMR studies on both the free peptide and non-covalent complex were performed.

4.3.5.1 NMR Spectroscopy of Unbound Caerin 1.8.11

Upon binding CaM, the 11 residue caerin 1.8.11 is expected to adopt an amphipathic α-helical structure. In aqueous solution, peptides of this type adopt random coil structures. Consequently, the high resolution structure of this α-helical form of the peptide was determined in DPC micelles, a solvent system known to induce secondary structure, where it has the propensity to form.

4.3.5.2 NMR Assignment

NMR spectra of caerin 1.8.11 were recorded in DPC micelles at pH 6.0. The proton resonances were assigned using the sequential assignment strategy outlined by Wüthrich [160] and a combination of TOCSY, DQF-COSY and NOESY spectra. These assignments are indicated in the partial TOCSY and NOESY spectra in Figure 4.11.

There were several coincident amide resonances, creating some ambiguity in the assignment process. DQF-COSY data and NOE connectivities were used to unambiguously assign the chemical shifts. The NH region of the NOESY spectrum was utilised to sequentially assign the amide proton chemical shifts. Where sequential d_{NN} NOE cross-peaks were not present, $d_{\alpha N}$ correlations were utilised to obtain reliable resonance assignments.

A ^{13}C HSQC spectrum was used to assign the α^{13}C resonances that were located in the α^1H/α^{13}C region (spectrum not shown). The α^{13}C chemical shifts for the two Gly residues were not observed in the spectrum and could not be assigned.

Table 4.12 summarises the assigned ^{1}H and ^{13}C resonances.

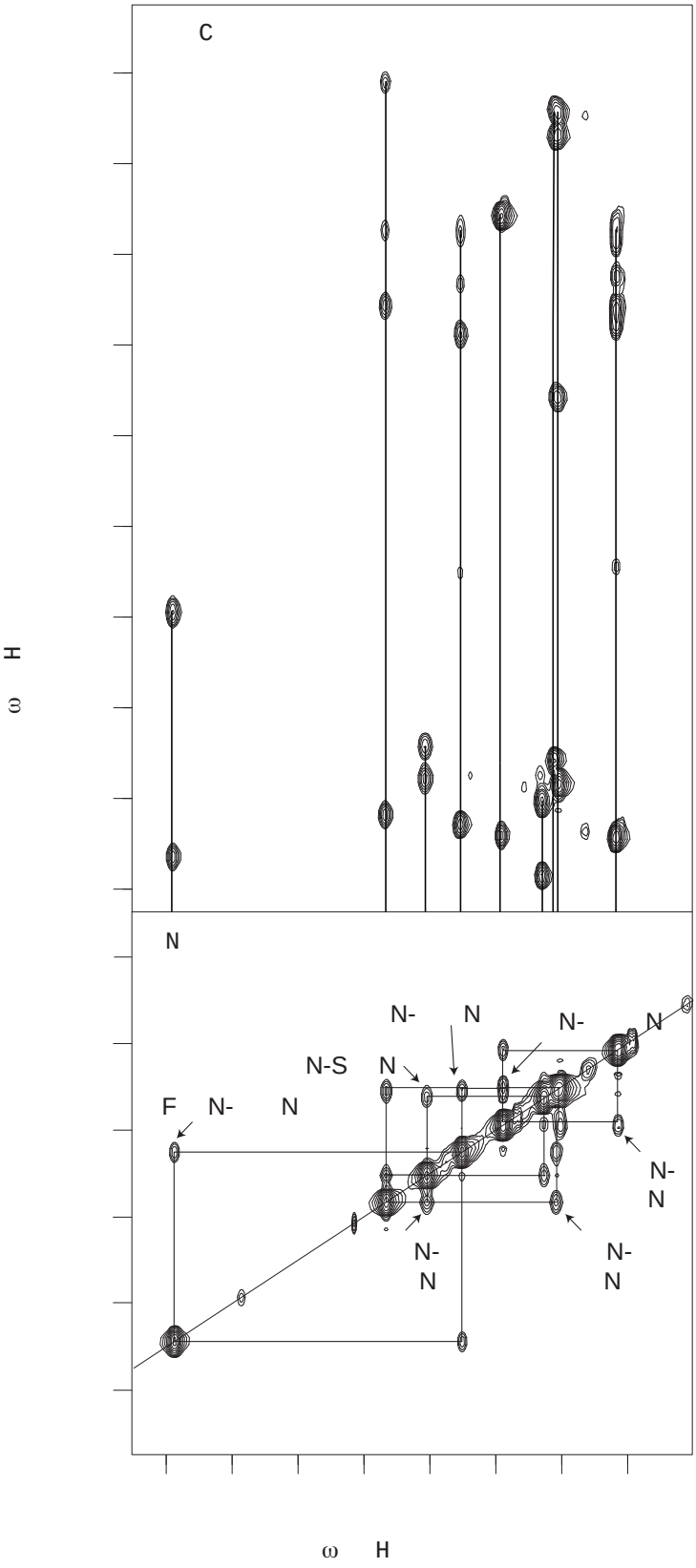


Figure 4.11: Partial NOESY and TOCSY spectra of caerin 1.8.11 in DPC micelles. In the TOCSY spectrum, lines connect resonances of the same amino acid residue. NOEs between sequential amide protons are shown in the NOESY spectrum.

Table 4.12: Assigned ^{1}H and ^{13}C chemical shifts for caerin 1.8.11 in DPC micelles. n.o. indicates resonance was not observed.

	Chemical Shift (ppm)				
Residue	**NH**	**αH**	**βH**	**Other H**	**αC**
Gly1	n.o.	4.07, 3.89			n.o.
Leu2	7.81	4.00	1.94	γH - 1.55	57.8
				δH - 1.08, 0.90	
Phe3	8.98	4.25	3.18	n.o	60.8
Lys4	8.10	4.11	1.95	γH - 1.50	58.6
				δH - 1.74	
				ϵH - 3.01	
Val5	7.82	3.83	2.23	γH - 1.07, 0.96	65.3
Leu6	8.33	4.07	1.83	γH - 1.49	57.1
				δH - 0.95,0.84	
Gly7	8.21	3.91,3.77			n.o.
Ser8	7.86	4.34	4.01		60.5
Val9	7.81	3.94	2.23	γH - 1.08, 0.97	63.9
Ala10	7.98	4.16	1.34	53.1	
Lys11	7.63	4.17	1.91, 1.83	γH - 1.48,1.55	56.2
				δH - 1.70	
				ϵH - 2.98	
				$CONH_2$ - 7.30,7.20	

4.3.5.3 Secondary Shifts

$\Delta\delta$s describe the chemical shift variations for the amino acids in the peptide from their corresponding random coil values (as determined in water and outlined in [435]). The $\Delta\delta$'s of the α^{1}H and α^{13}C were smoothed over $\pm$ 2 residues, while the N^{1}H $\Delta\delta$s are plotted unsmoothed. All $\Delta\delta$s were plotted against the amino acid sequence to allow extraction of structural information (Figure 4.12). As discussed in Section 2.3.7.1, negative $\Delta\delta$ values indicate an upfield shift from random coil values, whilst positive values indicate a downfield shift.

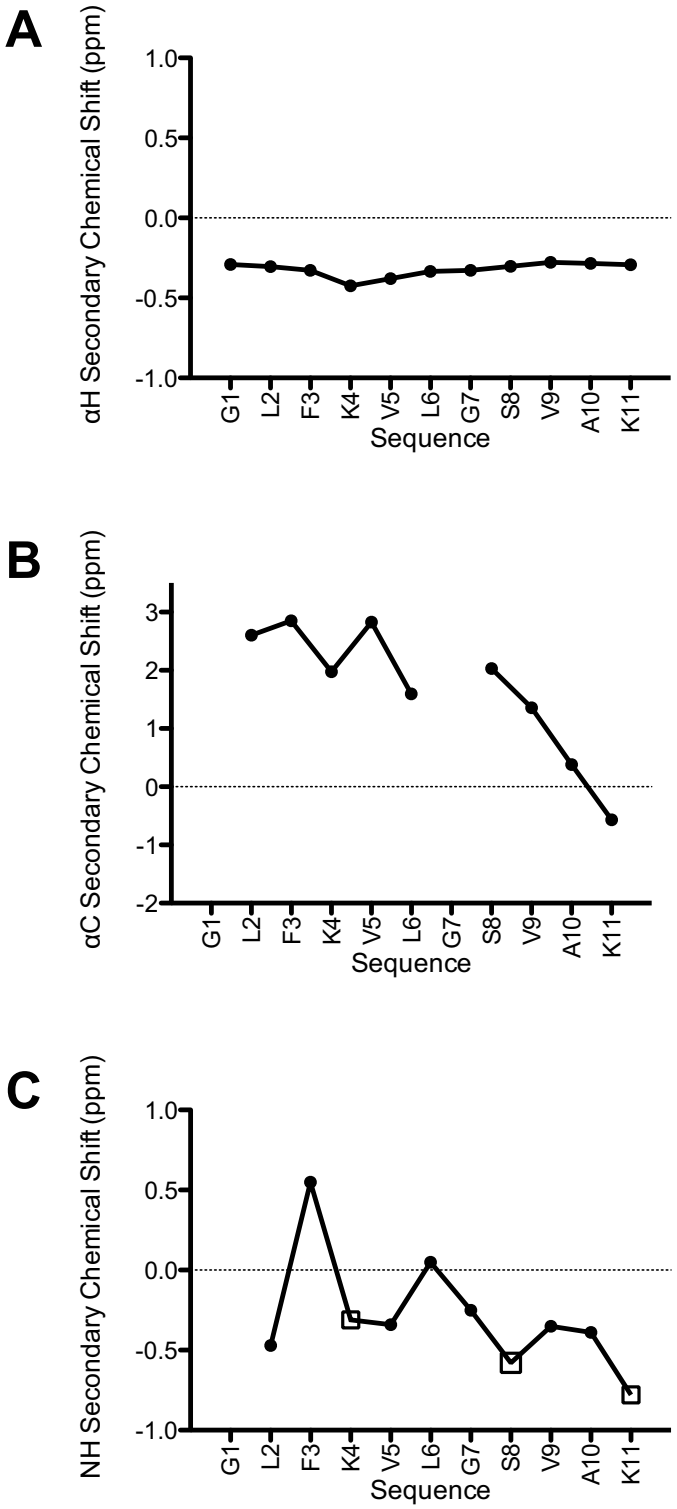


Figure 4.12: (A) α^{1}H and (B) α^{13}C secondary shifts of caerin 1.8.11 in DPC micelles, smoothed over n ± 2 residues. (C) N^{1}H secondary shifts of caerin 1.8.11. Negative values indicate an upfield shift from random coil values, whilst positive values indicate a downfield shift. Hydrophilic residues are indicated as open symbols.

The α^1H resonances (Figure 4.12A) display an upfield shift from random coil values as indicated by the negative $\Delta\delta$s. This indicates that the peptide is adopting a helical structure encompassing the length of the peptide [182, 185]. This is supported by the α^{13}C resonances (Figure 4.12B) which exhibit a downfield trend across all of the peptide, apart from the final residue.

The N^{1}H $\Delta\delta$s (Figure 4.12C) display a degree of periodicity across the length of the peptide, with the three hydrophilic residues having a more upfield chemical shift from random coil values when compared with hydrophobic residues. This suggests a degree of amphipathicity [184, 186] is present in the helix. Additionally, at the N-terminus the N^{1}H $\Delta\delta$ values show an upfield shift, characteristic of an α helical secondary structure. Together, these $\Delta\delta$ values lend support to the notion that there is a degree of helical structure present in the peptide.

4.3.5.4 NOE Connectivities

Diagnostic NOE connectivities used in structure calculations for caerin 1.8.11 in DPC micelles are displayed in Figure 4.13. Sequential d_{NN}, $d_{\alpha N}$ and $d_{\beta N}$ NOEs are observed along most of the sequence. Additionally, medium range NOEs between residues that are three or four apart are observed. Notably, $d_{\alpha N(i,i+3)}$ and $d_{\alpha N(i,i+4)}$ connectivities are observed throughout the sequence, whilst $d_{\alpha\beta(i,i+3)}$ connectives are observed at the C-terminus.

Taken together these data suggest that caerin 1.8.11 adopts an α helical structure in DPC micelles. The presence of additional medium-range connectivites at the C-terminus suggests that the helical structure is more well defined in this region and that at the N-terminus a more flexible structure may be present.

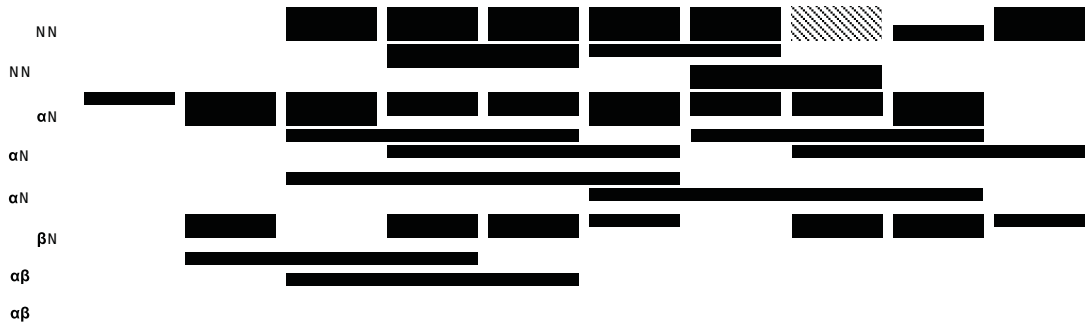


Figure 4.13: Diagnostic NOEs used in structure calculations for caerin 1.8.11. The thickness of the bar indicates the relative strength of the NOE.

4.3.5.5 Structure Calculations

The NOESY spectrum for caerin 1.8.11 was assigned and the volume of each peak converted to distance restraints using the procedures outlined by Xu *et al.* [436]. A total of 214 distance restraints were generated, and are summarised in Table 4.13. Only a small number of these distance restraints were ambiguous.

Table 4.13: Distance restraints obtained from the NOESY spectrum of caerin 1.8.11 in DPC micelles utilised in structural calculations.

	Number of Restraints
Sequential NOEs	76
Medium-range NOEs	41
Long-range NOEs	-
Intra-residue NOEs	92
Ambiguous NOEs	5
Total	**214**

The RMD and SA protocol of ARIA generated sixty final structures, of which the twenty with the lowest potential energy were selected for analysis. Figure 4.14 shows these twenty lowest energy structures of caerin 1.8.11 superimposed over the well defined residues. This structure calculation result confirms the conclusions drawn from analysis of the NOE and secondary shift data from the NMR. It can be seen that caerin 1.8.11 adopts a helical structure along its length, however, it appears to be more flexible at the termini.

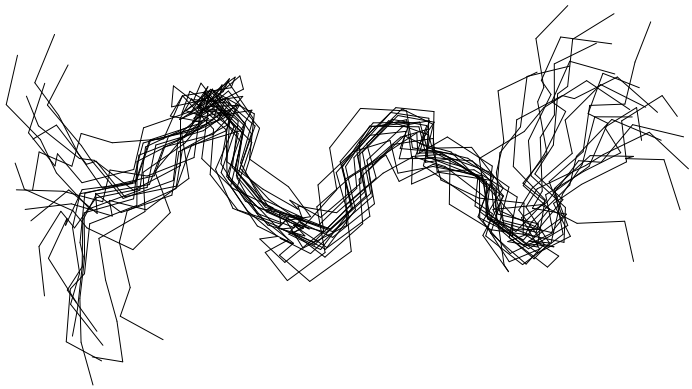


Figure 4.14: 20 lowest energy structures of caerin 1.8.11 in DPC micelles superimposed over the backbone atoms of well-defined residues.

The 20 lowest energy structures show considerable convergence over the entire sequence, but the RMSD values (Table 4.14) are slightly higher than longer, previously reported peptide structures (c.f. in [437]). This may indicate that the helical structure exhibits a degree of conformational flexibility, consistent with its short sequence containing the helix-breaking Gly residue. Additionally, the final structures had no violations from idealised covalent geometries, supporting the quality of the obtained structures.

Table 4.14: Structural statistics for caerin 1.8.11 following RMD/SA calculations.

Energy (kcal.mol^{-1})	
E_{total}	5.60 ± 0.39
E_{bond}	0.03 ± 0.01
E_{angle}	3.90 ± 0.05
$E_{improper}$	0.10 ± 0.01
E_{VDW}	1.54 ± 0.33
E_{NOE}	0.01 ± 0.03
E_{cdih}	0.00
Well-defined residues	3-9
RMSD from mean geometry (Å)	
Backbone atoms of well-defined residues	0.55 ± 0.22
Heavy atoms of well-defined residues	1.35 ± 0.30
All backbone atoms	1.31 ± 0.24
All heavy atoms	1.81 ± 0.23

Analysis of the angular order parameters (AOPs) of the twenty lowest energy structures demonstrated that seven of the residues were well-defined (Table 4.14). A Ramachandran plot showing the average ϕ and ψ angles for caerin 1.8.11 shows that all the residues

fall within the favoured region for α-helical structure (Figure 4.15), further supporting the quality of the calculated structures in representing the solution structure of caerin 1.8.11.

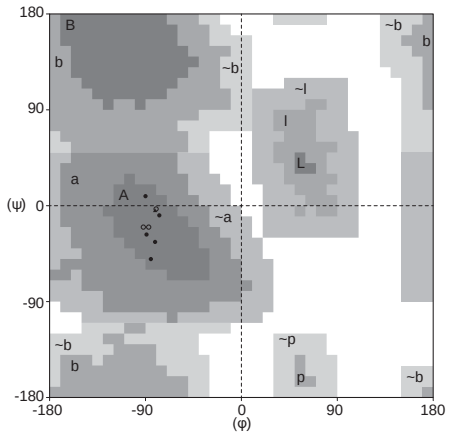


Figure 4.15: Ramachandran plot for caerin 1.8.11 in DPC micelles. Well-defined residues are indicated by filled symbols. Favourable regions are labelled A and B for α helical and β strand structures, respectively.

The most stable calculated structure of caerin 1.8.11 in DPC micelles is shown in Figure 4.16. The amphipathic helical structure along the length of the peptide is clearly evident.

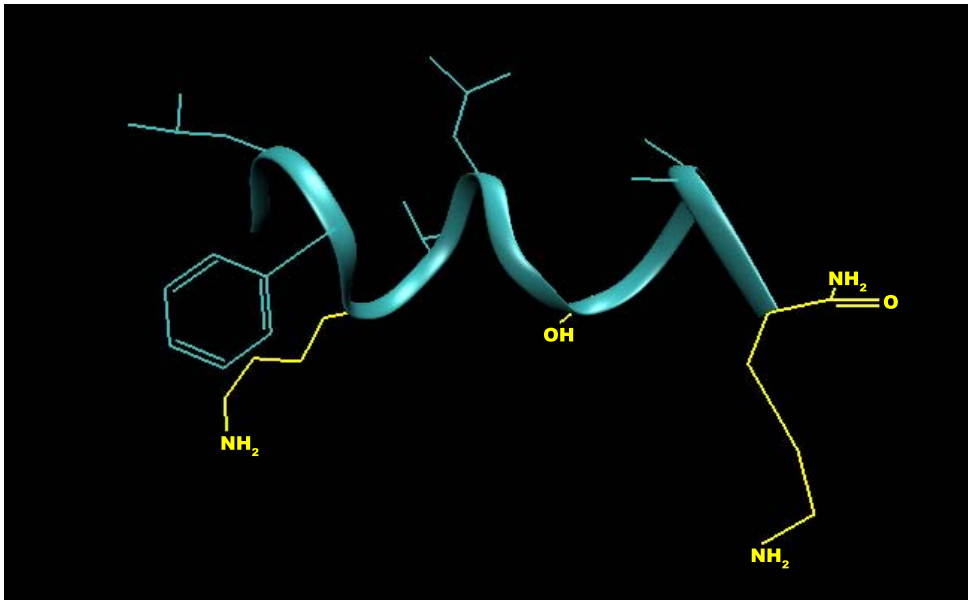


Figure 4.16: Lowest calculated potential energy structure of caerin 1.8.11 in DPC micelles. Hydrophobic groups are shown in green, hydrophilic groups are shown in yellow.

4.3.5.6 ^{15}N HSQC Titration

In order to investigate the nature of the non-covalent complex that forms between Ca_4^{2+}CaM and caerin 1.8.11, titration experiments were performed by adding successive quantities of unlabelled caerin 1.8.11 to ^{15}N-labelled Ca_4^{2+}CaM. After each addition, a ^{15}N-^{1}H HSQC spectrum was recorded and chemical shift changes were monitored by overlaying the processed spectra (Figure 4.17). Chemical shift changes were considered significant if greater than 0.05 ppm in the ^{1}H dimension and 0.5 ppm in the ^{15}N dimension [174].

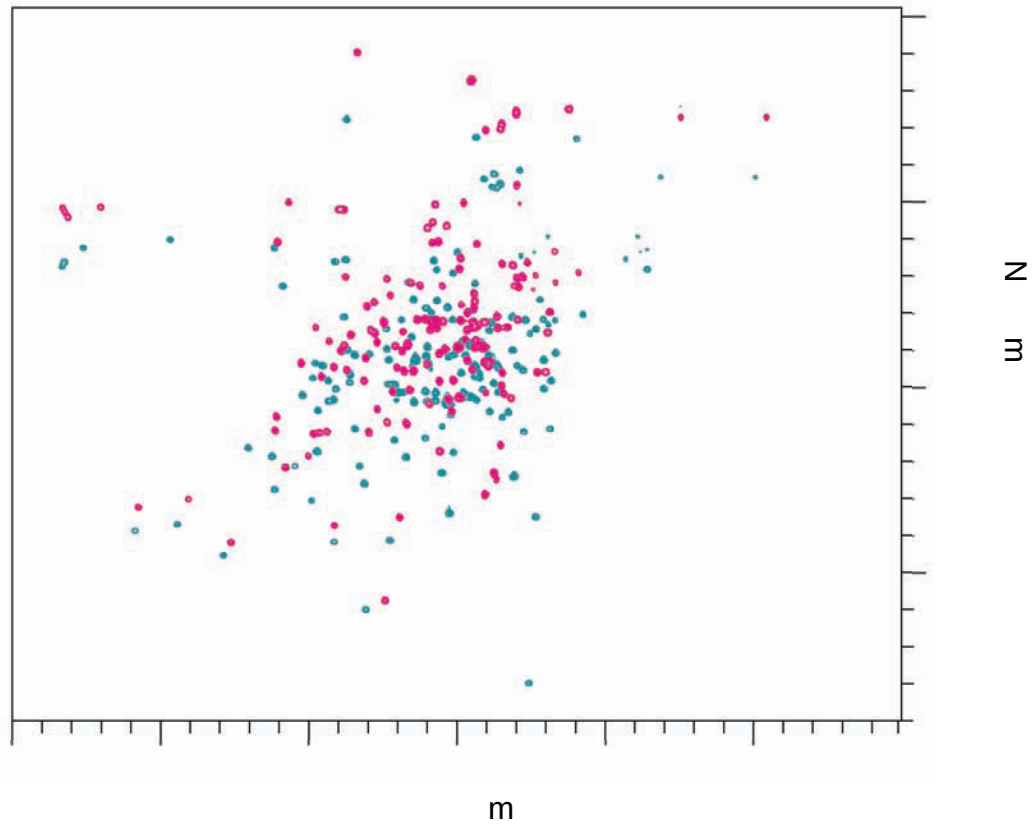


Figure 4.17: Overlaid ^{15} HSQC spectra of ^{15}N-labelled Ca_4^{2+}CaM (red) and after the addition of 1 molar equivalent of caerin 1.8.11 (green).

It can also be seen in Figure 4.17 that complexation does not cause any significant line broadening (as a result of the modest increase in molecular weight), and that the resonances are mostly well-dispersed. However, the central region of the spectrum between ~ 119 - 123 ppm in the ^{15}N dimension and ~ 8.5 - 7.5 ppm in the ^{1}H does display some signal overlap in both the bound and unbound species. Additionally, Figure 4.17 shows that significant chemical shift changes were observed for most residues in the titration, suggesting a global reorganisation of protein structure. This is consistent with the presence

of the canonical collapsed complex structure. Chemical shift changes were not observed as a function of concentration, instead a second set of peaks with a different chemical shift signature corresponding to the protein-peptide complex, were observed after 0.4 equivalents of peptide were added. This is consistent with the complex being in the slow exchange regime. Addition of an equimolar quantity of peptide resulted in only one set of peaks being observed in the spectrum, corresponding to the complex, and addition of further peptide did not have any additional effect on the observed chemical shifts.

4.3.5.7 Backbone Chemical Shift Assignment

It is apparent from the ^{15}N-HSQC spectra in Figure 4.17 that there are significant chemical shift changes across the entirety of Ca_4^{2+}CaM upon binding caerin 1.8.11. Consequently, a combination of the main chain directed and sequential assignment strategies was used to assign the backbone chemical shifts of Ca_4^{2+}CaM when bound to caerin 1.8.11. This process was mainly achieved using HNCACB and CBCA(CO)NH spectra and was assisted by a variety of NMR investigations that have been previously performed to study calmodulin and its complexes [368, 438, 439]. As an example, Figure 4.18 shows overlaid regions of the CBCA(CO)NH and HNCACB spectra, illustrating sequential connectivities from residues 58 to 61.

A complete list of the assigned chemical shifts is given in Appendix A.

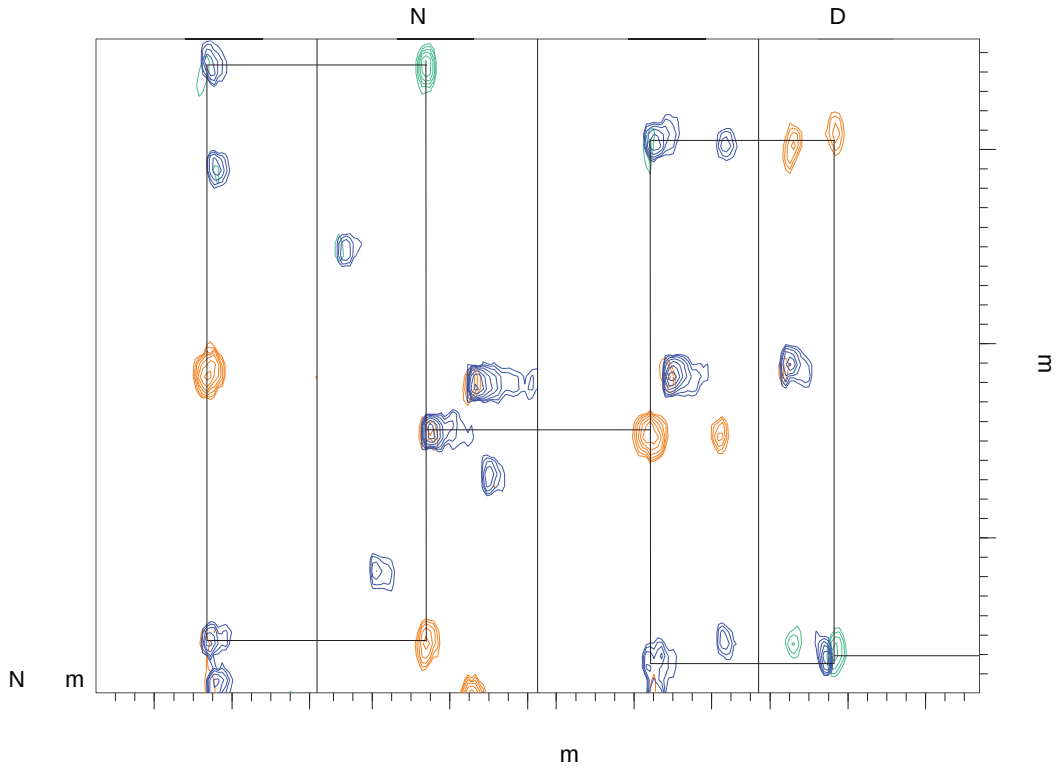


Figure 4.18: Overlayed HNCACB (orange, green) and CBCA(CO)NH (blue) spectra of Ca_4^{2+}CaM bound to caerin 1.8.11. Sequential connectivities for four amino acid residues are shown.

4.3.5.8 Secondary Chemical Shifts

The structural changes in Ca_4^{2+}CaM associated with binding a peptide result in chemical shift perturbations. Consequently, these data can be used to indicate which regions of Ca_4^{2+}CaM experience significant changes in chemical environment upon binding. A full structural analysis using NOE information would be required to determine the overall structure of Ca_4^{2+}CaM in complex with a peptide. Nevertheless, significant structural information can be obtained from the assigned backbone chemical shifts.

The $^{13}C\alpha$ $\Delta\delta$s (Figure 4.19A) and chemical shift index (CSI) values [440, 441], indicate downfield ($\Delta\delta$ >0) or upfield ($\Delta\delta$ <0) shifts from random coil values (with the CSI only indicating the direction, not size of the shift). These values can be used to infer regions of regular secondary structure, however this structural analysis approach requires a grouping of four or more consecutive residues with the same CSI value to define a region as having a particular secondary structure [441]. As outlined in Section 2.3.7.1, α-helical

and β-sheet regions of secondary structure are indicated by downfield and upfield $^{13}C\alpha$ shifts, respectively.

Ca_4^{2+}CaM is a well-studied protein that is known to have a high helical content, comprising eight α-helices separated by several loop motifs (Table 4.15). In the X-ray structure, helices IV and V appear to combine into one extended helix possibly due to crystal packing effects, however, NMR and other data indicate that these are discrete structural motifs and their separation results in a centrally located flexible hinge region [355, 359]. These known helical regions are indicated by the shaded regions in Figure 4.19.

Table 4.15: Location of α-helices in Ca_4^{2+}CaM as determined by NMR spectroscopy [359] and X-ray crystallography [355].

α-helix	**NMR**	**X-ray**
I	E6-F19	T5-F19
II	T29-S38	T29-S38
III	E45-E54	E45-V55
IV	F65-K77	F65-F92
hinge	K77-E82	
V	E82-D93	
VI	A102-N111	A102-N111
VII	D118-E127	D118-A128
VIII	Y138-T146	Y138-S147

The $\Delta\delta$ information in Figure 4.19 demonstrates the loss of helicity about the central flexible hinge region (residues 77-82), as indicated by the negative $\Delta\delta$s in this region. However, the eight helical regions of Ca_4^{2+}CaM are largely retained upon binding to caerin 1.8.11, indicated by the eight sustained regions of positive $\Delta\delta$s and positive CSIs. This demonstrates that the large chemical shift changes observed as a result of binding are likely due to rearrangement of the helices, rather than a dissolution of any secondary structure elements.

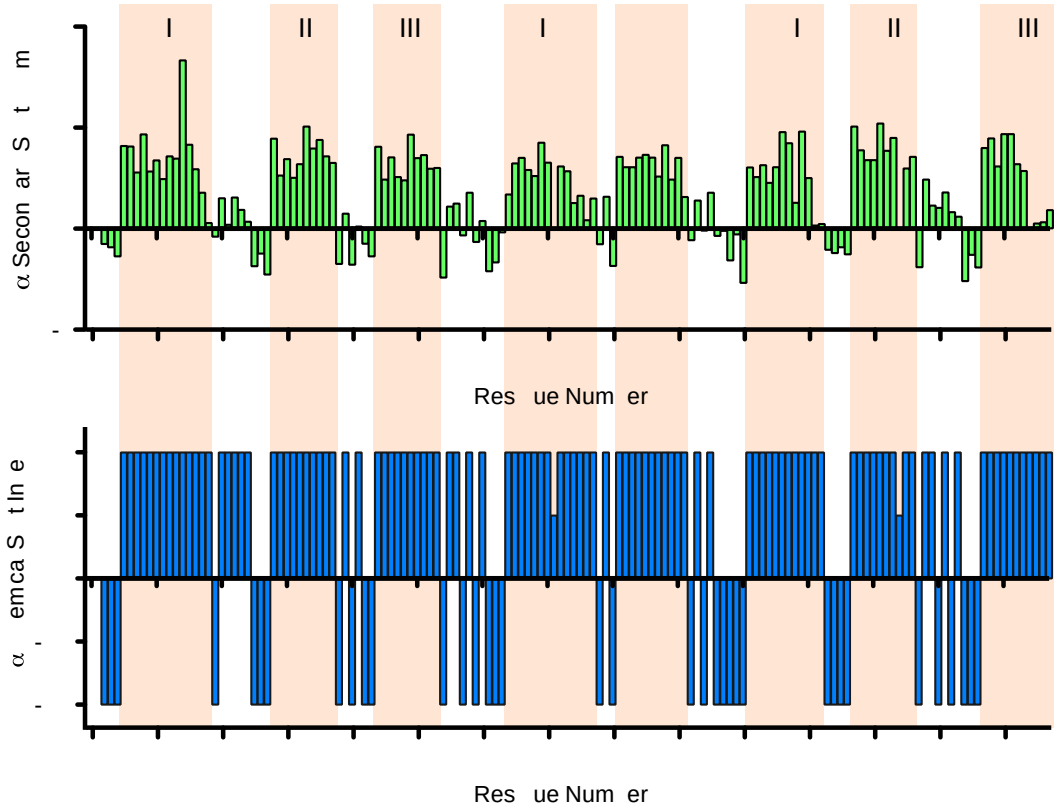


Figure 4.19: (A) αC ^{13}C secondary shifts and, (B) the corresponding CSI plot of Ca_4^{2+}CaM bound to caerin 1.8.11. Where the resonance was not assigned, the CSI is arbitrarily plotted as 0.5 to differentiate from zero and negative CSIs.

4.3.5.9 Chemical Shift Perturbations

It is possible to map chemical shift perturbations (CSPs) that occur in Ca_4^{2+}CaM upon complexation with caerin 1.8.11, to determine regions where significant CSPs occur as a result of binding. These data essentially depict the difference in chemical shift between the unbound and bound forms of Ca_4^{2+}CaM, and are summarised for the α^{13}C, and backbone amide ^{15}NH and N^{1}H in Figure 4.20. Chemical shifts of unbound Ca_4^{2+}CaM were obtained from a previous study [442]. There are significant CSPs across the protein backbone, but generally speaking, the most significant perturbations are observed in the central region of the protein (around residue 81) and at the C-terminus.

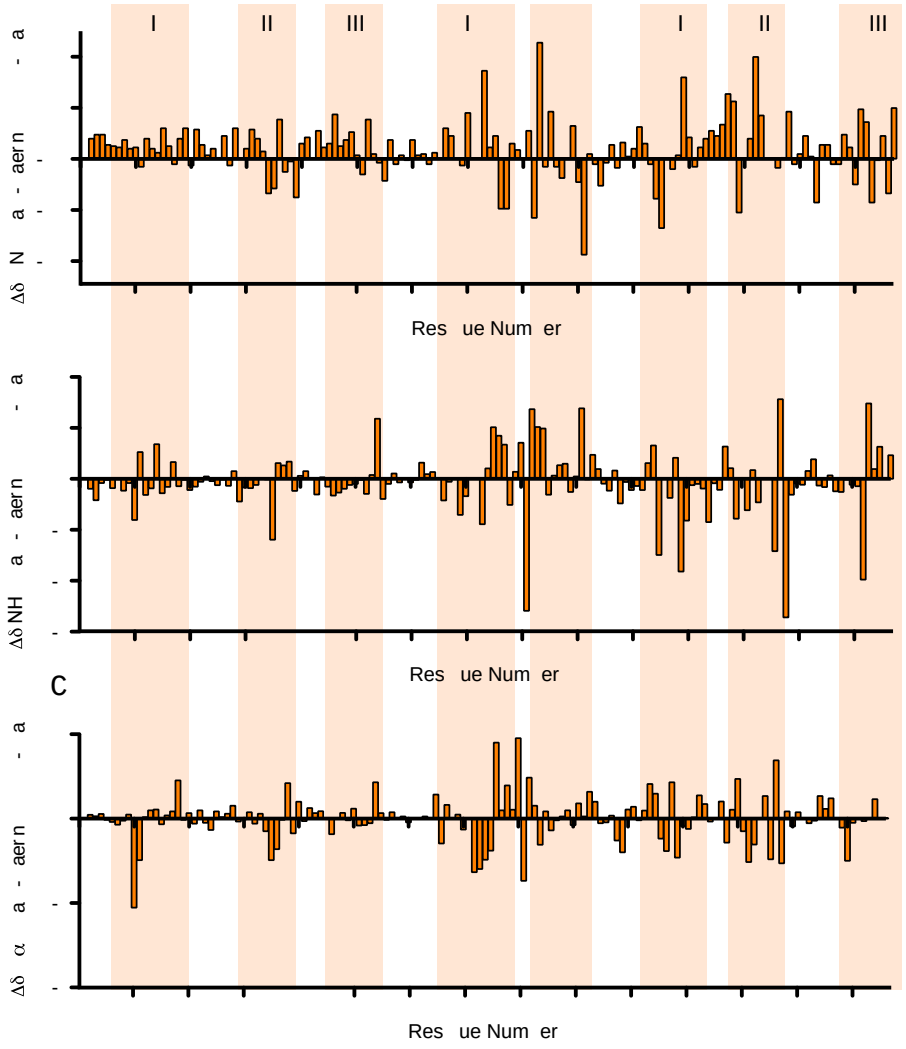


Figure 4.20: Chemical shift perturbations in the (A) amide N^{1}H, (B) amide ^{15}NH, and (C) α^{13}C resonances of Ca_4^{2+}CaM as a result of binding to caerin 1.8.11.

4.4 Discussion

4.4.1 Insights from Circular Dichroism

It is well established that the binding of Ca_4^{2+}CaM to its peptide targets induces helical structure in the peptide, without significantly affecting the secondary structural elements of the protein [366, 423–425]. The amphibian peptides studied herein represent each of the three classes of amphibian peptide inhibitors. They also have other multifaceted activities, and in addition to their Ca_4^{2+}CaM binding properties they are potent antimicrobial peptides (AMPs, Section 5.1.3) [384, 392]. This AMP activity requires them to adopt an α-helical conformation to trigger membrane disruption.

It can be seen from the data presented in Figure 4.5 that addition of the peptide substrate to a solution of Ca_4^{2+}CaM results in the induction of further α-helical structural elements in the sample. Whilst the increase in helical signal may be due to conformation change and stabilisation of the CaM structure, the likely reason is due to helix formation by the peptide. As the peptide itself in solution adopts a random coil structure (data not shown), with a different characteristic CD spectrum (a minimum at 195 nm), it can be inferred that addition of peptide results in binding, and a structural conversion of the peptide to the favoured α-helical conformation. This is evidenced by the enhanced CD signals at 208 and 222 nm.

4.4.2 Insights from Isothermal Titration Calorimetry

The binding isotherms obtained from the ITC experiments were all indicative of very strong interactions, as evidenced by their steepness (Figures 4.6, 4.7 and 4.8). Optimisation of the experimental conditions was attempted in order to try and obtain isotherms better suited to more reliably determine kinetic data. However, this was unsuccessful, presumably due to the strength of the binding interactions. Nevertheless, the data were fit to a single site binding model, and all fitted K_d values were in the low nM range, demonstrating the strength of the interaction. The binding of the peptides was also shown to be stoichiometric, with fitted n-values in the range 0.91-1.23.

The ΔC_ps of binding have also been determined (Table 4.7), with these values lying in the range of -2.7 to -3.0 kJ.mol^{-1}.K^{-1}. A number of Ca_4^{2+}CaM:peptide complexes have been analysed by ITC, and their associated ΔC_p values have been previously reported [427]. The previously studied peptides are divided into two groups, depending on the ΔC_p of their complexation with Ca_4^{2+}CaM. It has been determined that a complex which adopts the canonical collapsed conformation exhibits a ΔC_p in the order of - 3.2 kJ.mol^{-1}.K^{-1}, whilst extended conformations, where only the C-terminal lobe of Ca_4^{2+}CaM is involved in binding, exhibit a ΔC_p in the order of -1.6 kJ.mol^{-1}.K^{-1} [427]. This significant difference in magnitude can be attributed to the decreased solvent accessible surface area when the complex adopts the canonical collapsed conformation [427]. Comparison of the results obtained here with previously published ΔC_p values lends support to the notion that the studied amphibian peptides force Ca_4^{2+}CaM to adopt the canonical collapsed conformation upon binding, with the peptide encapsulated between the N- and C- terminal lobes of the protein.

4.4.3 Insights from Ion Mobility-Mass Spectrometry

4.4.3.1 Conformational Analysis of Calmodulin and Ca^{2+} Binding

Often negative ion mode mass spectrometry is used to study CaM and its complexes [419, 429, 431]. This is primarily to minimise non-specific Ca^{2+} adduction, which can occur more readily in the positive ion mode [443, 444]. In this work, positive ion mode travelling wave IM-MS was used, as it is a well-characterised technique to probe the CCSs of proteins and protein complexes [43]. Here, conditions were carefully controlled to minimise non-specific Ca^{2+} adduction. Negative ion mode travelling wave IM-MS is rarely utilised to calibrate for drift time, due to the more limited number of calibrants available. However, preliminary work in this area suggests calibrated CCSs of CaM in the positive and negative ion mode are comparable (data not shown) [445]. It should also be noted that CaM has been shown to behave differently in the positive and negative ion modes when studied by ESI-MS [446].

Recent reports have challenged traditional thinking about the dumbbell structure of Ca_4^{2+}CaM. In particular, NMR evidence suggests that Ca^{2+} has a destabilising effect on the dumbbell structure, and a recent X-ray crystal structure has shown that holo-CaM can exist in a more globular form [349, 351, 357]. It is apparent from the IM-MS spectrum of CaM (Figure 4.9) that there are two gaussian charge state distributions and two populations in the IM-MS drift plot, consistent with the presence of two gas phase CaM conformers. The distribution centred around $[M+13H]^{13+}$ is likely to correspond to more extended, unfolded structures [447–449], whilst the other around $[M+7H]^{7+}$ is likely to correspond to the more globular form of CaM. The relative abundance of the two charge state distributions indicate that the globular form predominates. As the mass spectrometric conditions used were controlled to retain native protein structures, it is likely that these two populations reflect the solution phase conformational flexibility of CaM, consistent with previous MS [430, 431] and NMR [349, 357] data. The IM data does not show a significant population of ions with structures intermediate between the two conformations, indicating that these may be unstable in the gas phase on the time-scale of the experiment.

The calibrated CCSs of CaM at a variety of calcium occupation states are summarised in Table 4.9, and are consistent with reports studying CaM complexes by both positive and negative ion mode IM-MS [429, 450]. It is evident from these data that significant increases in CCS occur at and above the $[M+9H]^{9+}$ charge state, indicating these higher charge state ions are of a different conformational state. This is the case for apoCaM and all of the Ca^{2+} adducts. The difference between the CCSs of the two conformational states observed in the IM-MS spectra (Table 4.9) and those calculated for the dumbbell and globular Ca_4^{2+}CaM (Table 4.11) compare favourably, suggesting the maintenance of these two solution phase conformations.

The differences in CCS between apoCaM and Ca_4^{2+}CaM indicate that rearrangement to a more compact structure takes place to incorporate all four Ca^{2+} ions. The data also indicate that the first and third Ca^{2+} binding events do not result in significant CCS changes. Conversely, the CCS trends indicate that binding of Ca^{2+} two and four lead to a more pronounced decrease in CCS. This is especially evident at lower charge states (below $[M+9H]^{9+}$), that is, for the more compact CaM structures. This suggests that binding of the first Ca^{2+} to each lobe results in a local structural change rather than a global reorganisation

of protein structure. Only once both EF hand domains in a lobe have bound Ca^{2+} does a significant structural change occur.

4.4.3.2 Calmodulin-Peptide Binding

This work presents the first example of IM-MS analysis of Ca_4^{2+}CaM:amphibian peptide complexes, and in addition to this, the first IM-MS analysis of a CaM:peptide complex of extended structure. The CaM:C20W complex studied has been shown by NMR to adopt this less common extended conformation, in which the first 12 residues of the peptide interact with only the C-terminal lobe of CaM [368]. This binding event has been shown to only require 2 equivalents of Ca^{2+}, unlike those which require both lobes of Ca_4^{2+}CaM [419, 445].

It can be seen in Figure 4.10 that the charge state distributions of the CaM:peptide complexes are relatively narrow and are at low charge states (high m/z), suggesting the complexes are relatively compact and of well-defined structure. This is distinct from the conformational flexibility in apo- and holo-CaM, and the consequent large number of charge states observed (Figure 4.9). Additionally, it can be seen that the Ca_4^{2+}CaM:amphibian peptide complexes adopt charge states of $[M+8H]^{8+}$ and $[M+7H]^{7+}$, whilst the Ca_2^{2+}CaM:C20W complex adopts charge states $[M+8H]^{8+}$ and $[M+9H]^{9+}$ (Figure 4.10). The higher charge states adopted by the CaM:C20W complex is consistent with it having a more extended structure than the other complexes studied.

The calibrated CCSs for the complexes studied at the charge states observed are summarised in Table 4.10. Comparison of the measured CCSs at the $[M+8H]^{8+}$ charge state demonstrate that, generally speaking, the Ca_4^{2+}CaM:amphibian peptide complexes exhibit smaller CCSs than that of the Ca_2^{2+}CaM:C20W complex. This is consistent with these complexes adopting the more compact, canonical structure upon binding. Notably, the Ca_4^{2+}CaM:caerin 1.8 and Ca_4^{2+}CaM:C20W complexes have comparable CCS's (within the estimated error of 8 - 10 %), which are larger than the CaM complex with the truncated variant caerin 1.8.11. NMR experiments have previously shown that only the N-terminal portion of caerin 1.8 is encapsulated by CaM, with the remaining C-terminal region excluded from the hydrophobic binding pocket [451]. The slightly larger CCS of the

Ca_4^{2+}CaM:caerin 1.8 could be explained by these data, with the remaining peptide excluded from the complex, so it is unstructured and flexible in solution resulting in a larger CCS. Conversely, data has indicated that the entirety of the dahlein 5.6 peptide interacts with Ca_4^{2+}CaM [452], and is encapsulated by the protein, providing an explanation as to why the complex has a smaller CCS than that involving caerin 1.8. Similar experiments have not been performed for citropin 1.1 (mod. 13), although the measured CCS, which is similar to that of the caerin 1.8.11 complex, suggests the peptide is completely engulfed within Ca_4^{2+}CaM upon complexation. This observation should be further probed using complementary techniques, including NMR spectroscopy.

Theoretical CCSs were calculated for model canonical compact and extended complex structures (Table 4.11). Both PA and EHSS values were calculated as a guide to upper and lower bounds for CCS. These theoretical CCSs demonstrate that the canonical collapsed complex structures exhibit a reasonable variation in CCS. However, it is clearly evident that the Ca_2^{2+}CaM:C20W complex exhibits a larger CCS, consistent with its extended conformation, where only one lobe of CaM is involved in binding the peptide. Comparison of the theoretical CCS of the Ca_2^{2+}CaM:C20W complex with those measured from IM-MS show a good agreement between the two. In addition, CCSs of the amphibian peptide complexes all show good agreement with those calculated from structures in the PDB which adopt the compact conformation (except for the caerin 1.8 complex, due to the reasons outlined above). This provides further support for the notion that these Ca_4^{2+}CaM:amphibian peptide complexes are of this canonical structure.

4.4.4 Insights from NMR Spectroscopy

4.4.4.1 NMR Spectroscopy of Unbound Caerin 1.8.11

It is well-known that in order for an interaction with Ca_4^{2+}CaM to occur, the binding partner must comprise an overall positively charged sequence which adopts an amphipathic α-helical structure when bound [337, 338, 366, 367]. It has been shown by CD that caerin 1.8.11 also adopts this structure when bound to Ca_4^{2+}CaM. To determine the helical structure of caerin 1.8.11 when bound, a solvent system was used to induce

this secondary structure in bulk solution, namely the membrane mimicking DPC micelles (Section 4.3.5.1). The peptide is highly likely to adopt this structure when bound to CaM. This amphipathic character is further illustrated in the Schiffer-Edmundson helical wheel projection (Figure 4.21B), and it is evident that this amphipathicity is more pronounced in the truncated peptide, when compared with full length caerin 1.8 (Figure 4.21A).

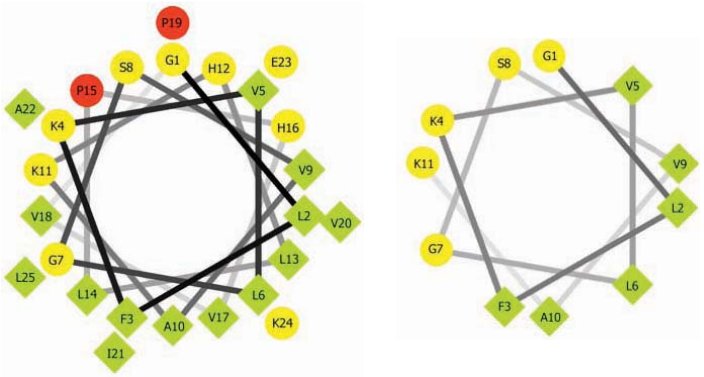


Figure 4.21: Schiffer-Edmundson helical wheel representation of (A) caerin 1.8 and (B) caerin 1.8.11. Hydrophobic residues are coloured green. Hydrophilic residues are coloured yellow. Pro residues are illustrated in red.

In addition to forming this positively charged amphipathic α-helical structure upon binding, it is generally accepted that large hydophobic residues must be in conserved positions, on the same face of the α-helix. The presence of these conserved residues at positions 1-5-10 or 1-8-14 determines the interactions with Ca_4^{2+}CaM [366]. The 11 residue caerin 1.8.11 has large hydrophobic groups Leu2, Leu6, Val9, corresponding to a 1-5-8 arrangement of hydrophobic anchor residues in the peptide. This differs from most of the recognition motifs in peptides which form the canonical compact type structure upon binding. Interestingly, the C20W peptide, which forms an extended complex with Ca_4^{2+}CaM exhibits this motif of anchor residues [368]. However, in this instance, a significant amount of structural data supports the notion that Ca_4^{2+}CaM adopts the canonical collapsed conformation when bound to caerin 1.8.11. This demonstrates that the binding motif alone cannot be used to determine the conformation of the resultant peptide:Ca_4^{2+}CaM complex.

4.4.4.2 NMR Spectroscopy of the CaM:Caerin 1.8.11 Complex

It is well documented that peptide binding to Ca_4^{2+}CaM results in minimal disruption to the helical structure of the protein, but induces secondary structure in the binding target

[366, 423–425]. It is apparent from the NMR data that a similar situation is observed here for the complex of Ca_4^{2+}CaM with caerin 1.8.11. The eight sustained regions of downfield αC ^{13}C $\Delta\delta$s (positive CSIs), as indicated in Figure 4.19, conform with the known eight helical domains as determined by NMR studies of the unbound Ca_4^{2+}CaM [359].

These data are also consistent with previous observations that the central helix of Ca_4^{2+}CaM, as observed in the X-ray crystal structure, is disrupted at the known hinge region in solution (residues K77-E82) [355, 359]. This flexible hinge is necessary for the conformational rearrangement that occurs when binding peptide targets in the canonical compact conformation.

In addition, the data obtained from NMR again supports the notion that the whole protein is involved in binding, but the enhanced CSPs at the C-terminus suggest the peptide more strongly associates with this region (Figure 4.22), as outlined for several other CaM binding peptides which adopt the canonical complex structure [424, 438].

To obtain a clearer, overall picture of the CSPs that are occurring upon binding to caerin 1.8.11, Figure 4.22 displays the overall CSPs of the amide group of Ca_4^{2+}CaM, calculated using the relationship below [453];

$$CSP(ppm) = \sqrt{0.102 \times \Delta\delta(^{15}N)^2 + \Delta\delta(^{1}H)^2} \tag{4.1}$$

where $\Delta\delta(^{15}$N) and $\Delta\delta(^{1}$H) are the ^{15}N and ^{1}H CSPs, respectively. CSPs greater than 0.25 ppm are considered significant, whilst those greater than 0.5 ppm are considered dramatic [424]. These data show the effects of binding more clearly, and confirm that the largest CSPs are occurring at the C-terminus of Ca_4^{2+}CaM, with more dramatic CSPs in this region, but there are significant CSPs across the whole sequence [424].

These CSPs can be mapped to residues of the hydrophobic pockets of CaM, or those residues nearby (Figure 4.23). This hydrophobic core region is considered to be formed by the first and fourth helices of each domain (i.e helices I and IV form the N-terminal hydrophobic domain and helices V and VIII form the C-terminal hydrophobic domain [351]). Notably, significant or dramatic CSPs are observed for numerous hydrophobic residues in and around the hydrophobic core region. This indicates that hydrophobic interactions are vital in the binding of Ca_4^{2+}CaM to caerin 1.8.11. Additionally, many of the Met residues located

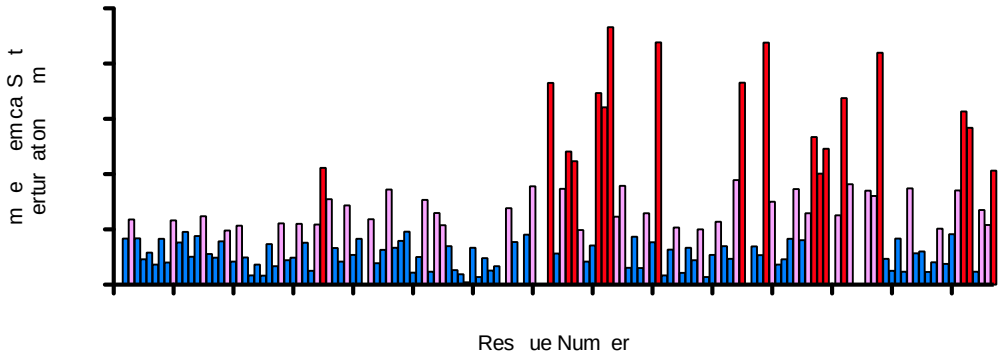


Figure 4.22: Average amide CSP in Ca_4^{2+}CaM as a result of binding to caerin 1.8.11. CSPs >0.25 are considered significant (pink) and CSPs >0.5 are considered dramatic (red) [424].

in both of the N- and C- terminal hydrophobic pockets of CaM (such as Met 36, 71, 109, 124, 144 and 155) show dramatic CSPs upon binding (Figures 4.22 and 4.23). These Met residues are well known to play an important role in binding [366]. Together, these data indicate that both hydrophobic regions of Ca_4^{2+}CaM are involved in binding to caerin 1.8.11, confirming that the complex is adopting the canonical compact arrangement.

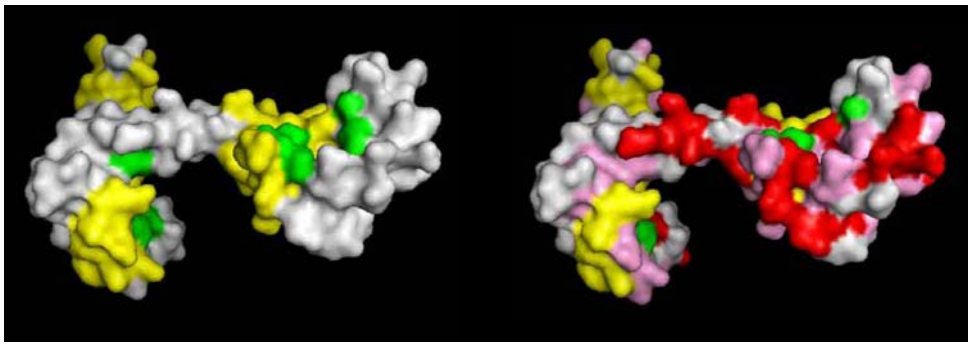


Figure 4.23: (A) Surface representation of the dumbbell structure of Ca_4^{2+}CaM (PDB: 1CLL). The hydrophobic domains are shown in yellow, and Met residues are shown in green. (B) Residues that exhibited significant (CSP >0.25 ppm) and dramatic (CSP >0.5 ppm) chemical shift changes are shown in pink and red, respectively.

The experiments presented in this chapter do not allow for a complete high-resolution structural characterisation of the Ca_4^{2+}CaM complexes. For this to occur, further multidimensional NMR experiments would need to be performed to determine intermolecular NOEs between CaM and the peptides. Even so, the wealth of low resolution structural data presented provides a substantial basis for the understanding and further characterisation of the interactions between Ca_4^{2+}CaM and these amphibian

peptides which inhibit nNOS. Overall, the results overwhelmingly demonstrate that all the Ca_4^{2+}CaM:amphibian peptide complexes studied adopt the canonical collapsed structure.

4.5 Experimental Procedures

4.5.1 Materials

Unless specified, reagents were purchased from Sigma Aldrich (St. Louis, MO, USA). ^{15}N-ammonium chloride and ^{13}C-glucose were purchased from Cambridge Isotope Laboratories (Andover, MA, USA). Peptides were synthesised using L-amino acids and the standard N-α-Fmoc methodology by GenScript Corp. (Piscataway, NJ, USA). Samples for NMR spectroscopy were shown to be greater than 90 % pure by HPLC and ESI-MS.

4.5.2 Protein Gels

The Mini-Protean Tetra Cell system was used along with precast TGX Gels (Bio-Rad, Gladesville, NSW) and TGS running buffer (25 mM tris, 192 mM glycine, 0.1 % (w/v) SDS, pH 8.3). Protein standards were purchased from Bio-Rad and all gels were stained with Coomassie blue stain solution (0.1 % Coomassie blue, 30 % methanol, 10 % glacial acetic acid) and destained with a solution of 50 % methanol, 5 % glacial acetic acid.

4.5.3 Expression of Calmodulin

CaM was expressed by inoculating Lurea broth (LB, 1 % bacto-tryptone, 0.5 % yeast extract, 1 % sodium chloride, adjusted to pH 7) (10 mL) with a single colony of the BL21 (DE3) strain of *E. coli* containing the pET28 vector with the CaM gene inserted (a kind gift from Joachim Krebs, Max Planck Institute for Biophysical Chemistry, Germany). This was incubated overnight at 37 °C, with shaking. A 3 % subculture was made the following morning into LB and the cells were grown until an OD_{600} of 0.6-0.8 was reached. Protein expression was induced by adding IPTG to a concentration of 50 mM. The cells were left to express for 2 hr before being harvested by centrifugation at 3000 $\times g$ for 15 min at 4 °C. Cell pellets were frozen at -20 °C until required.

^{15}N/^{13}C-labelled CaM was prepared by inoculating Min A medium (60 mM dipotassium phosphate, 33 mM potassium dihydrogen phosphate, 1.7 mM sodium citrate, autoclaved,

then added 15 mM ^{15}N-ammonium chloride, 0.005 % thiamine, 0.2 % ^{13}C-glucose, 0.8 mM magnesium sulfate, 10 mL) with a single colony of the BL21 (DE3) strain of *E. coli* containing the pET28 vector with the CaM gene inserted. This was incubated overnight at 37 °C, with shaking. A 3 % subculture was made the following morning into Min A medium containing ^{15}N-ammonium chloride and ^{13}C-glucose as the sole nitrogen and carbon sources. Protein expression was completed as described above.

4.5.4 Purification of Calmodulin by Hydrophobic Interaction Chromatography

CaM was purified using a procedure previously outlined [420, 421, 454]. Briefly, the cells were suspended in 20 mL of lysis buffer (10 mM HEPES pH 7.5, 1 mM EDTA, 1 mM NaN_3, 10 mM DTT, 1 mM PMSF) and lysed by repeated freeze-thawing using liquid N_2. The lysate was centrifuged (13, 000 $\times g$, 20 min) and the supernatant heated at 80 °C in a water bath for 10 min and centrifuged again. 5 mM $CaCl_2$ was added to the sample and it was further centrifuged.

The supernatant was filtered through a 0.45 μm syringe filter and the soluble proteins were loaded onto a 50 mL Phenyl Sepharose column (GE Life Sciences, Rydalmere, NSW) equilibrated with 10 mM HEPES pH 7.4, 4 mM $CaCl_2$. Unbound material was removed from the column by washing with equilibration buffer. The column was then washed with 3 column volumes of equilibration buffer supplemented with 500 mM NaCl and then reequilibrated in the initial buffer. CaM was eluted with 10 mM HEPES pH 7.5, 5 mM EGTA. Fractions were analysed by SDS-PAGE and those which contained CaM were pooled and concentrated by centrifugal ultrafiltration using Amicon centrifugal filtration units with a 10 kDa molecular weight cut-off (Millipore, Bedford, MA, USA).

4.5.5 Circular Dichroism Spectroscopy

CD spectra were acquired in the far-UV range using a Jasco-815 spectropolarimeter (Jasco, MD, USA) at room temperature using a 0.1 cm path length quartz cuvette. Each spectrum was obtained by averaging five scans and subtracting the buffer contribution. A 2 μM solution of CaM in 10 mM Tris-HCl, 10 mM KCl, 5 mM $CaCl_2$, pH 6.3 was used. The concentration of CaM was determined spectrophotometrically by measuring the absorbance at 277 nm and using an extinction coefficient of 3300 $M^{-1}{\cdot}cm^{-1}$ [455]. Peptides (1 molar equivalent) were added from a stock solution (400 μM) in the same buffer.

4.5.6 Isothermal Titration Calorimetry

CaM was expressed and purified as previously described. The protein was extensively dialysed against the ITC buffer (10 mM HEPES, pH 7.5, 100 mM NaCl, 10 mM $CaCl_2$) to ensure complete buffer exchange. The concentration of CaM was determined spectrophotometrically by measuring the absorbance at 277 nm and using an extinction coefficient of 3300 $M^{-1}{\cdot}cm^{-1}$ [455].

Peptide stock solutions were prepared in mQ water and concentrations determined by amino acid analysis. Aliquots were taken and lyophilised before being resuspended in the ITC buffer, used to dialyse the protein, prior to analysis.

Binding of the peptides to Ca_4^{2+}CaM was studied using a VP-ITC microcalorimeter (MicroCal, Norhampton, MA). The peptide solutions (300 μM) were injected into the sample cell containing CaM (30 μM). The first injection (2 μL) was followed by 19 injections of the peptide solution (15 μL). Experiments were conducted at four different temperatures. The heat of dilution was measured by titrating buffer into the protein solution and was found to be minimal, nonetheless this was subtracted from the experimental data (not shown). The raw data were fitted to a single site binding model using Origin software (Version 5.0, MicroCal).

4.5.7 Ion Mobility-Mass Spectrometry

4.5.7.1 Sample Preparation

CaM (Sigma Aldrich, MO, USA) was dissolved in water at a concentration of 40 μM and dialysed against 4× 2L 10 mM ammonium acetate, 2 mM EDTA; 4× 2L 10 mM ammonium acetate, pH 6.8. The concentration of CaM was determined spectrophotometrically by measuring the absorbance at 277 nm and using an extinction coefficient of 3300 $M^{-1}{\cdot}cm^{-1}$ [455], and the solution was diluted to a final concentration of 10 μM. Ca_4^{2+}CaM was prepared by adding 2.5 molar equivalents of Ca^{2+} ions in the form of calcium acetate (from a stock solution of 5 mM calcium acetate in 10 mM ammonium acetate, pH 6.8) to observe CaM with its full complement of four Ca^{2+} ions. Peptide solutions were prepared by dissolving the solid in buffer (10 mM ammonium acetate, pH 6.8) at a concentration of 1-3 mM. Small aliquots were added to the CaM solutions to achieve the desired peptide:CaM molar ratios (typically 1:1).

4.5.7.2 Mass Spectrometry

IM-MS spectra were acquired on a Synapt HDMS system (Waters, UK) [48], using nanoESI in the positive ion mode. The sample was introduced using platinum-coated borosilicate capillary needles that were prepared in-house. Instrument parameters were optimised to remove adducts whilst preserving non-covalent interactions, and were typically as follows; capillary voltage, 1.8 kV; cone voltage, 40-80 eV; trap collision energy, 10 V; source temperature, 50 °C; backing pressure, 5 mBar; IMS cell pressure (N_2), 0.5 mBar; travelling wave velocity, 400 ms^{-1}; travelling wave height, 6-9 V. The MS data were processed using the program MassLynx (Waters, UK, version 4.1) and deconvoluted using the maximum entropy algorithm incorporated in the software.

4.5.7.3 Data Analysis

Drift-time measurements obtained from the Synapt HDMS were normalised for charge state and a nonlinear correction function was applied for calibrant ions such that their relative differences mirror those previously observed for the same ions [43, 122, 456, 457]. CCSs of the reference samples were taken from the literature using values for ubiquitin, myoglobin and cytochrome c [119]. The CCS calibration procedure is described in detail elsewhere [43].

CCSs of model protein structures were calculated using the program MOBCAL (developed by Jarrold and co-workers) [120, 121]. Structural coordinates from previous studies using NMR and X-ray crystallography were obtained from the Protein Data Bank (PDB) with accession numbers indicated in the text.

4.5.8 NMR Spectroscopy

4.5.8.1 NMR Spectroscopy of Unbound Caerin 1.8.11

Caerin 1.8.11 (2.8 mg, 2.5 μmol) and DPC (39 mg, 0.1 mmol) were dissolved in an aqueous solution of 10 % D_2O and NaH_2PO_4 buffer (50 mM) giving a final peptide concentration of 5 mM in a volume of 0.5 mL. This concentration of DPC was selected as it greatly exceeds the CMC [223] and it produces a solution of approximately one peptide molecule per micelle [222, 224, 458]. The pH was adjusted to pH 5.98 (using NaOH).

NMR spectra were recorded using a Varian Inova-600 NMR spectrometer, with a ^{1}H frequency of 600 MHz and a ^{13}C frequency of 150 MHz. Experiments were carried out at 25 °C and referenced to water (4.82 ppm). The water signal in both the TOCSY and NOESY experiments was suppressed by presaturation. This was achieved by application of a low power RF pulse at the frequency of the water resonance during the 1.1 s relaxation delay between scans. The water signal in the DQF-COSY spectra was suppressed using gradient methods [459].

TOCSY, DQF-COSY and NOESY experiments were acquired in the phase-sensitive mode, with time proportional phase incrementations in t_1 [460]. For each experiment, 8 to 16

time-averaged scans were acquired per increment, with a total of 200 t_1 increments for TOCSY and NOESY spectra and 512 t_1 increments for DQF-COSY spectra. The FID in t_2 consisted of 2878 data points over a spectral width of 9592.3 Hz. NOESY spectra were acquired with a mixing time of 150 ms and the TOCSY pulse sequence included a spin-lock of 80 ms. HSQC experiments were recorded for the assignment of α^{13}C resonances using correlations with the attached α^{1}H. These experiments consisted of 8 time-averaged scans over a total of 512 t_1 increments. 2878 data points were collected over a spectral width of 9592.3 Hz in the ^{1}H dimension (F_2).

2D spectra were processed using VNMR software (VNMRJ, Version 2.1, Revision B). The data matrices were multiplied by a Gaussian function in both dimensions then zero-filled before Fourier transformation. The final processed 2D NMR matrices consisted of 4096 × 4096 real data points.

4.5.8.2 Structure Calculations

^{1}H resonances in the NOESY, DQF-COSY and TOCSY spectra of caerin 1.8.11 were assigned using Sparky software (version 3.111) by means of the sequential assignment procedure [160]. The volumes of the larger of the two symmetrical cross-peaks were converted to distance restraints by the methods of Nilges *et al.* [190]. $^{3}J_{NH\alpha H}$ coupling constants were measured from the high-resolution 1D spectra obtained. From this analysis, dihedral angles were restrained in accordance with the guidelines outlined in Section 2.3.8.3.

Structures were generated from random starting conformations, using the standard RMD and SA protocol of ARIA (Version 1.2) [461], implemented with CNS (Version 1.1) [462]. Each ARIA run consisted of 8 iterations of structure calculations. The standard ARIA parameters were utilised, however, better convergence was achieved using parameters based on those used by Pari *et al.* [463] and Kang *et al.* [464]. The final iteration of calculations generated 60 structures from which the 20 lowest calculated potential energy structures were selected for analysis. The programs VMD (Version 1.8.2) [465] and MOLMOL (Version 2k.2) [466] were used to visualise the 3D structures.

4.5.8.3 Sample Preparation for NMR Titration

$^{15}N/^{13}C$-labelled CaM (3.16 mg, 1.89 x 10^{-7} mol) was dissolved in a 500 μL aqueous solution of 10% D_2O, potassium chloride (100 mM) and calcium chloride (40 mM). The solution was adjusted to pH 6.3 by the addition of small quantities of dilute hydrochloric acid or sodium hydroxide, as required. Sodium azide (0.02 %) was added as a preservative. Caerin 1.8.11 (1.69 mg, 1.51 x 10^{-6} mol) was dissolved in water (500 μL) and divided into aliquots such that successive additions would give the desired mole ratio of peptide:CaM (Table 4.16). The aliquots were lyophilised and the dried peptide portions were added to the solution in sequence prior to recording each spectrum. The pH was readjusted to 6.3 as required.

Table 4.16: Caerin 1.8.11 quantities used in the titration with $^{15}N/^{13}C$-labelled CaM.

Step	Caerin 1.8.11 (mg)	Caerin 1.8.11 (x 10^{-7} mol)	Peptide: CaM ratio
0	0	0	0:1
1	0.042	0.38	0.2:1
2	0.084	0.76	0.4:1
3	0.127	1.14	0.6:1
4	0.168	1.52	0.8:1
5	0.211	1.89	1:1
6	0.844	3.78	2:1
7	1.689	7.56	4:1

4.5.8.4 ^{15}N HSQC NMR Titration

All NMR spectra were recorded on a Varian Inova-600 NMR Spectrometer with a ^{1}H frequency of 600 MHz and a ^{13}C frequency of 150 MHz. Experiments were conducted at 25 °C. The standard gNhsqc [467] pulse sequence from the VnmrJ library was used. 128 increments, each consisting of 8 transients were acquired over 2048 data points. In the ^{1}H dimension, a spectral width of 12019.2 Hz was used, whilst in the ^{15}N dimension, a spectral width of 1944.3 Hz was used. Spectra were processed using NMRPipe [468] and viewed with CCPNMR Analysis (Version 2.1) [469].

4.5.8.5 Three-Dimensional NMR Spectroscopy

All NMR spectra were recorded on a Varian Inova-600 NMR Spectrometer with a ^{1}H frequency of 600 MHz and a ^{13}C frequency of 150 MHz. Experiments were conducted at 25 °C. The ^{1}H frequency domain was referenced to DSS at 0.0 ppm, whilst the heteronuclear dimensions were referenced indirectly [470]. The standard gHN_CACB [176] and gCBCA_CO_NH [177] pulse sequences from the VnmrJ library were used.

Both spectra were acquired with 64 (t_1) and 32 (t_2) increments, each consisting of 8 transients over 2048 data points. The carrier frequencies for ^{1}H, ^{13}C and ^{15}N were set to 4.773, 47.362 and 118.861 ppm, respectively and spectral widths of 12019.2 Hz (^{1}H), 12062.0 Hz (^{13}C) and 1944.3 Hz (^{15}N) were used.

The spectra were processed using NMRPipe [468] and viewed with CCPNMR Analysis (Version 2.1) [469]. Assignment of the NMR spectra was performed using a combination of the main-chain directed [178] and sequential [160] assignment procedures.

Chapter 5

Amphibian Peptides That Inhibit Fibril Formation and Self-Assemble

5.1 Introduction

A number of human disease states have the hallmark histological feature of extracellular or intracellular amyloid deposits [8, 471]. These include neurodegenerative diseases, where protein aggregation occurs in the brain, and nonneuropathic localised or systemic amyloidoses, where aggregation occurs in a single tissue (not the brain), or in multiple tissues, respectively (Table 5.1). Some of these conditions arise in individuals with a genetic predisposition, alternatively they can occur sporadically. The proteins found in the aggregates vary between disease states and share no obvious functional similarity, sequence identity or structural homology [8].

The high economic and social impacts [472] of amyloid diseases has led to much research into fibril formation by disease related proteins, in an attempt to identify new targets for therapeutic intervention [471, 473, 474]. This is especially true for proteins associated with age related neurodegenerative diseases, including Amyloid-β (Aβ, Alzheimer's disease) [475, 476] and α-synuclein (Parkinson's disease) [477, 478], and nonneuropathic amyloidoses, such as IAPP (Type II diabetes) [479–481] and β2-microglobulin (hemodialysis-related amyloidosis) [482–484].

Table 5.1: Selected diseases associated with amyloid formation. The associated aggregated proteins/peptides are indicated. Adapted from [8].

Many non-disease related proteins also form these amyloid structures as a requirement for their normal biological activities [8, 485, 486]. Several examples of proteins which form amyloid fibrils with non-pathological functions are detailed in Table 5.2. A number of host-defence peptides have also been shown to form fibrils, again a functional link is proposed (Section 5.1.3). Finally, as amyloid structures can also be formed *in vitro* from proteins not known to cause disease, it has also been proposed that fibrillar species may potentially be useful as a new class of nanomaterials [487–491].

Table 5.2: Proteins that form functional, non-pathological amyloid fibrils. Adapted from [8].

5.1.1 Amyloid Fibril Formation, Structure and Toxicity

The two process of protein folding and aggregation are competitive (Section 1.3), with environmental and physicochemical conditions dictating which pathway is favoured. For ordered (amyloidogenic) aggregation to occur from a folded protein, a partial unfolding of the native structure must occur to afford an intermediate on the amyloid pathway [8]. Conversely, intrinsically disorded proteins must partially fold to form an intermediate structure [492]. These intermediates then interact in a specific manner to afford the β-sheet containing nuclei upon which the fibrillar structure is assembled [8, 493, 494]. The formation of these nuclei (or 'seeds') is often the rate determining step in fibril formation.

All amyloid fibrils share similar structural characteristics, irrespective of the protein from which they are derived [495]. Typically, they are long filamentous structures (often 6-12 nm in diameter and up to microns in length) which contain a 'cross-β' fibrillar core that is resistant to proteolysis [496–499] (Figure 5.1). In many instances, only certain regions of the protein are incorporated into the core, whilst the remainder of the chain remains flexible, and susceptible to proteolytic degradation [500]. This core is characterised by the presence of β-strands oriented perpendicular to the fibril axis which are separated by

a distance of approximately 4.7 Å. Stacking of these β-strands results in β-sheets which extend laterally along the length of the fibril, with an inter-sheet spacing of approximately 9-11 Å [8, 496, 501]. These stacked β-sheet structures are termed protofilaments and the mature fibril comprises several of these structures which are interwoven [496, 502, 503].

Figure 5.1: Schematic diagram illustrating the substructure of an amyloid fibril (based on data from an SH3 domain amyloid fibril). (A) The fibril comprising intertwined protofibrils. (B) Zoomed view of one protofibril, showing the stacked β-strands. (C) Top and (D) side view of the mature fibril illustrating the packing of the β-sheets about the central core. From [504].

Amyloid fibrils are highly ordered, thermodynamically stable structures that are resistant to extreme conditions such as dehydration, pH, temperature and pressure [487, 505, 506]. This stability often leads to many of the issues associated with amyloid diseases, as the fibrillar species are unable to be cleared and accumulate in the affected tissue(s) [506, 507]. These deposits are termed amyloid plaques, and their presence and composition in a relevant tissue can be used to diagnose amyloid diseases, often post-mortem [508].

It is proposed that the most toxic species on the amyloid folding pathway are the oligomeric and pre-fibrillar aggregates [8, 509–515]. The fibrils themselves are considered less cytotoxic or may in some cases be an inert, protective accumulation of the toxic species. However, in some cases fibril cytotoxicity has been observed [516–518], and fibril assembly on membranes has been shown to lead to cell death [519]. Several mechanisms to

rationalise the cytotoxicity of the amyloidogenic species have been proposed. These include the oligomeric species forming pores in cellular membranes (in a similar way to AMPs, Section 5.1.3) leading to an influx of Ca^{2+} from the extracellular space [520]. Alternatively, the oligomeric species may bind to and activate Ca^{2+} ATPases either on the cell surface or the mitochondrial membrane (after oligomer endocytosis) [510]. The overall increase in intracellular Ca^{2+} results in oxidative stress, which leads to the enhancement of protein misfolding and impairment of ATPase activity [521]. Alternatively, the oligomers may interact with the endoplasmic reticulum and mitochondrial membranes triggering opening of the mitochondrial permeability transition pore, cytochrome c release and caspase 12 activation, leading to apoptosis [509]. The engulfed oligomers may also overwhelm the proteasome leading to a build up of intracellular misfolded proteins [509]. Data also indicate that the mature fibrils may interact with the membrane leading to extrinsic apoptosis by a variety of mechanisms [516, 518].

5.1.2 Inhibitors of Amyloid Formation

Inhibition of misfolding and fibril formation presents an attractive target to treat a variety of disease states which have the hallmark aetiological feature of amyloid deposits [473, 474, 476, 522]. It is proposed that the neurodegenerative effects of amyloids are related to the conformation that they adopt [509–511, 514, 515], and a variety of strategies have been investigated to prevent the formation of this toxic fold. These include stabilisation of the native conformation in a structured protein, thus preventing the structural changes required for amyloid formation. Alternatively, natively unstructured proteins may be sequestered by the addition of a binding partner to prevent aggregation. Strategies have also been developed to target aggregates, redirecting the aggregation pathway (to afford non-toxic amorphous aggregates), preventing fibril elongation, or evoking fibril disassembly [473, 474, 522]. Numerous small molecules, peptides and proteins have been identified and developed which display inhibition characteristics, and many show potential as treatments for amyloid disorders [473, 476, 523].

5.1.3 Antimicrobial Peptides

Most animals and plants have innate immune systems that comprise antimicrobial peptides (AMPs) as a vital component [524, 525]. Often, a range of AMPs are secreted upon receipt of a suitable stress stimulus with a spectrum of activities to ensure a fast, effective response against any invading microbe(s) [526]. The prevalence of AMPs in nature demonstrates their effectiveness as defence compounds, and their high activity and selectivity for microbes makes them an attractive starting point for the development of novel peptide based antimicrobials [527–530].

AMPs with a range of primary and secondary structures have been isolated from a variety of natural sources, with the diverse peptides divided into groups on the basis of both these factors. These include (i) anionic peptides, (ii) cationic peptides enriched with particular amino acids (such as Pro, Arg, His, Phe, Tyr or Gly), (iii) disulfide-bond containing AMPs which have β-strand type structures, and (iv) cationic peptides which form amphipathic α-helices [526, 531–533]. The peptides studied in this work all lie in this final category of AMPs.

AMPs cause the death of microorganisms by interacting with the lipid bilayer and then causing membrane disruption, a specific but not receptor mediated process [524, 526]. Consequently, it is more difficult for an organism to develop resistance against therapeutic AMPs, compared with traditional therapies, as in many instances this would require significant alterations to membrane composition [530]. Several mechanisms have been proposed to describe this process of membrane permeabilisation. It is proposed that this may be achieved by the formation of stable pores (described by the barrel-stave and toroidal pore models), micellisation in a detergent-like fashion (described by the carpet model) or other mechanisms, such as by membrane thinning (molecular electroporation or sinking rafts models) [530, 531, 534]. Other intracellular mechanisms of toxicity have also been proposed [526].

The barrel-stave model (Figure 5.2A) [535–537] describes the mechanism by which unstructured peptides bind to the membrane and undergo a conformational change to form regular secondary structures, such as amphipathic α-helices and β-sheets. Once a threshold concentration is reached, the peptides associate to form a bundle in the membrane with

the hydrophobic surface interacting with the acyl chains of the membrane, whilst the hydrophilic portions line the internal surface of the pore [538, 539]. Often AMPs are too short to traverse the bilayer, so this pore originally forms only in the outer leaflet, and only once a significant amount of peptide has bound to the bilayer do the monomers insert further into the hydrophobic core, allowing the pore structures to traverse the entire bilayer [540].

The toroidal-pore mechanism (Figure 5.2B) [537, 541] describes a situation where the polar face of the peptide associates with the lipid head groups. In this process, the lipids of the bilayer tilt to connect the two leaflets of the membrane. These lipids intercalate with the peptide molecules which form the pore structures [526].

Finally, in the carpet model [542] (Figure 5.2C), peptides electrostatically bind to the anionic head groups of the phospholipids, covering the membrane surface as a result of interactions between the lipid head groups and the peptides polar face. Upon reaching a critical threshold concentration, the peptides may form transient toroidal holes in the membrane, increasing the quantity of peptide which can access the membrane. Subsequent disintegration and micellisation of the membrane results [543–545].

Although presented here as discrete processes, it is proposed that the three modes of membrane activity represent single points on a continuous scale of potential mechanisms [546].

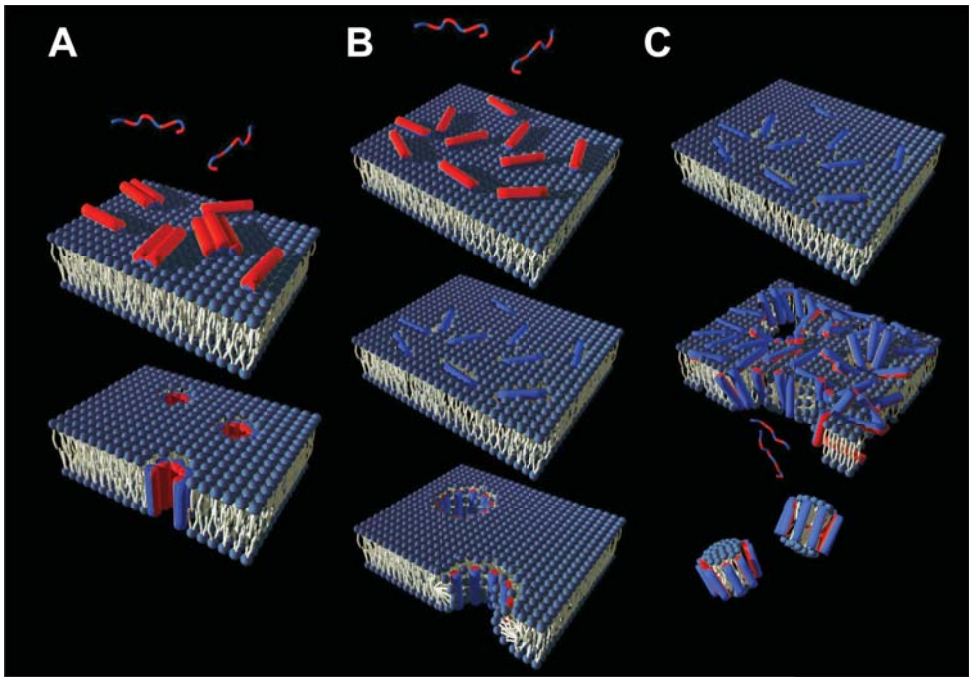


Figure 5.2: Mechanisms of membrane disruption. (A) Barrel-stave mechanism; the helical peptides aggregate on the membrane and insert into the bilayer to form hydrophilic pores, (B) toroidal-pore mechanism; the helical peptides bind to the membrane, rotate (by interacting with the lipid head groups) and insert into the bilayer to form toroidal pores, and (C) carpet mechanism; the peptides insert into the membrane which dissociates after a threshold concentration is reached. The helical peptides are represented by cylinders with hydrophobic and hydrophilic regions represented by blue and red respectively. From [547, 548].

5.1.3.1 Amyloidogenic Antimicrobial Peptides

There are several examples in the literature of AMPs that form amyloid fibrils (Table 5.3) (reviewed in [549]). In some instances, for example the temporins (Table 5.3) [550], these structures were shown to form in the presence of a lipid matrix, and it has been suggested that the membrane mediates this aggregation process [549]. It is also postulated that amyloid formation by these AMPs is directly related to their mechanism of bacterial cytotoxicity [550], and that neurodegenerative amyloidogenic proteins may also show antimicrobial activity [549, 551] (or exert their toxic effects in a similar manner to AMPs [511, 552]), as has been evidenced by Aβ [553]. This represents a novel relationship between AMPs and amyloid formation by disease related proteins, which may lead to significant advances in our understanding of amyloid diseases, and consequently lead to the development of novel therapeutics [549].

Table 5.3: Host defence peptides that form amyloid fibrils. Intramolecular disulfide bonds are indicated as lines. Adapted from [549, 554].

Peptide	Sequence
Humans	
LL-37	LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES
Amyloid-β (1-40)	DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV
Amyloid-β (1-42)	DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA
Lactoferrin*	NAGDVAFV
Eosinophil cationic protein (1-19)**	RPPQFTRAQWFAIQHISLN
Amphibians	
Magainin 2	GIGKFLHSAKKFGKAFVGEIMNS
Dermaseptin PD-3-7	LLGDLLGQTSKLVNDLTDTVGSIV
Dermaseptin S9	GLRSKIWLWVLLMIWQESNKFKKM
Temporin B	LLPIVGNLLKSLL-NH_2
Temporin L	FVQWFSKFLGRIL-NH_2
Kassinin	DVPKSDQFVGLM-NH_2
Pigs	
Protegrin 1	RGGRLCYCRRRFCVCVGR
Cow	
Indolicidin	ILPWKWPWWPWRR-NH_2
Insects	
Melittin	GIGAVLKVLTTGLPALISWIKRKRQQ
Bacteria	
Plantaracin A	KSSAYSLQMGATAIKQVKKLFKKWGW
Sakacin P	KYYGNGVHCGKHSCTVDWGTAIGNIGNNAAANWATGGNAGWNK
Synthetic	
VP1	GTAMRILGGVI

* The full length protein forms amorphous aggregates at high pH/temperature, the indicated region forms amyloid fibrils.
** The protein has been shown to form amyloid protofibrils, this region retains that activity.

5.2 Aims

The use of fibril formation inhibitors presents a promising mechanism to treat a variety of amyloidoses. Peptides are an interesting class of inhibitors, as their properties can be finely tuned to yield the required effect. Most peptide inhibitors of Aβ in the literature comprise hydrophobic residues along with β-sheet breaking residues and/or charged residues. Many amphibian AMPs have the required characteristics to inhibit fibril formation, given their overall positive charge and high degree of hydrophobic/aromatic residues. The aims of this work were to investigate a selected amphibian peptide to identify novel inhibitors of Aβ aggregation. Specifically, to examine the propensity the amphibian peptide caerin 1.8, and several synthetic modifications to inhibit fibril formation by Aβ.

This work led to the discovery that the amphibian AMP uperin 3.5 forms amyloid fibrils in solution. Consequently, this fibril formation process was investigated in some detail. The aims of this work were to gain insights into the mechanism of fibril formation by uperin 3.5, and its similarities with other disease related proteins and peptides that form amyloid structures. Studies of non-disease related amyloid fibril forming proteins and peptides can provide significant insight into the structure and function of disease related amyloids. Moreover, investigations of amyloidogenic AMPs, such as uperin 3.5, have been proposed to afford significant insight into their mechanism of antimicrobial action.

Caerin 1.8 GLFKVLGSVAKHLLPHVVPVIAEKL-NH_2
Uperin 3.5 GVGDLIRKAVSVIKNIV-NH_2

5.3 Results

The generic nature of amyloid structures mean that a variety of fundamental techniques to study their formation are well established [555]. A number of these common biophysical techniques are used in this chapter, including Thioflavin T (ThT) binding assays, transmission electron microscopy (TEM) and atomic force microscopy (AFM).

ThT is a benzothiazole dye that displays a shift in its excitation (from 385 to 440 nm) and emission maximum (from 445 nm to 480 nm), and an increase in quantum yield, by several orders of magnitude, when bound to β-sheet rich structures, such as amyloid fibrils [556–558]. Consequently, ThT is extensively used to stain and monitor the conversion of native peptide and protein structures to amyloid fibrils *in vitro*.

As ThT is an indirect method of determining fibril formation, studies are often conducted using complementary techniques, such as TEM and AFM, to determine the morphology and secondary structure of the produced species.

5.3.1 The Effect of Caerin 1.8 on Fibril Formation By Amyloid-β

The effect of caerin 1.8 on the aggregation of Aβ(1-42) was investigated by a combination of *in situ* ThT fluorescence assays and TEM (Figure 5.3). Addition of 4 molar equivalents of peptide resulted in complete inhibition of amyloid fibril formation, as can be observed by the minimal change in ThT fluorescence in Figure 5.3A. Addition of caerin 1.8 to a level below 4 molar equivalents results in an alteration to the kinetics of fibril formation, inducing a longer elongation phase.

The morphology of the species present after incubation changes from the long fibrillar species characteristic of Aβ(1-42) fibrils (Figure 5.3B) to amorphous looking aggregates (Figure 5.3E-G). The addition of a lower quantity of caerin 1.8 resulted in some attenuation of fibril formation, however, fibrillar species were still observed by TEM. From these data, it was possible to determine the percentage of protection invoked by caerin 1.8, by analysing the ThT maximum reached as a function of caerin 1.8 concentration, and fitting to a logarithmic concentration response curve (Figure 5.4). The caerin 1.8 concentration that gives 50 % inhibition of the increase in ThT fluorescence (IC_{50}) was 65 $\pm$ 7 μM.

As a control, caerin 1.8 incubated alone produced no change in ThT fluorescence over the course of the experiment and no aggregates were observed by TEM, indicating that the peptide is inherently not amyloidogenic nor likely to interact directly with free ThT (data not shown).

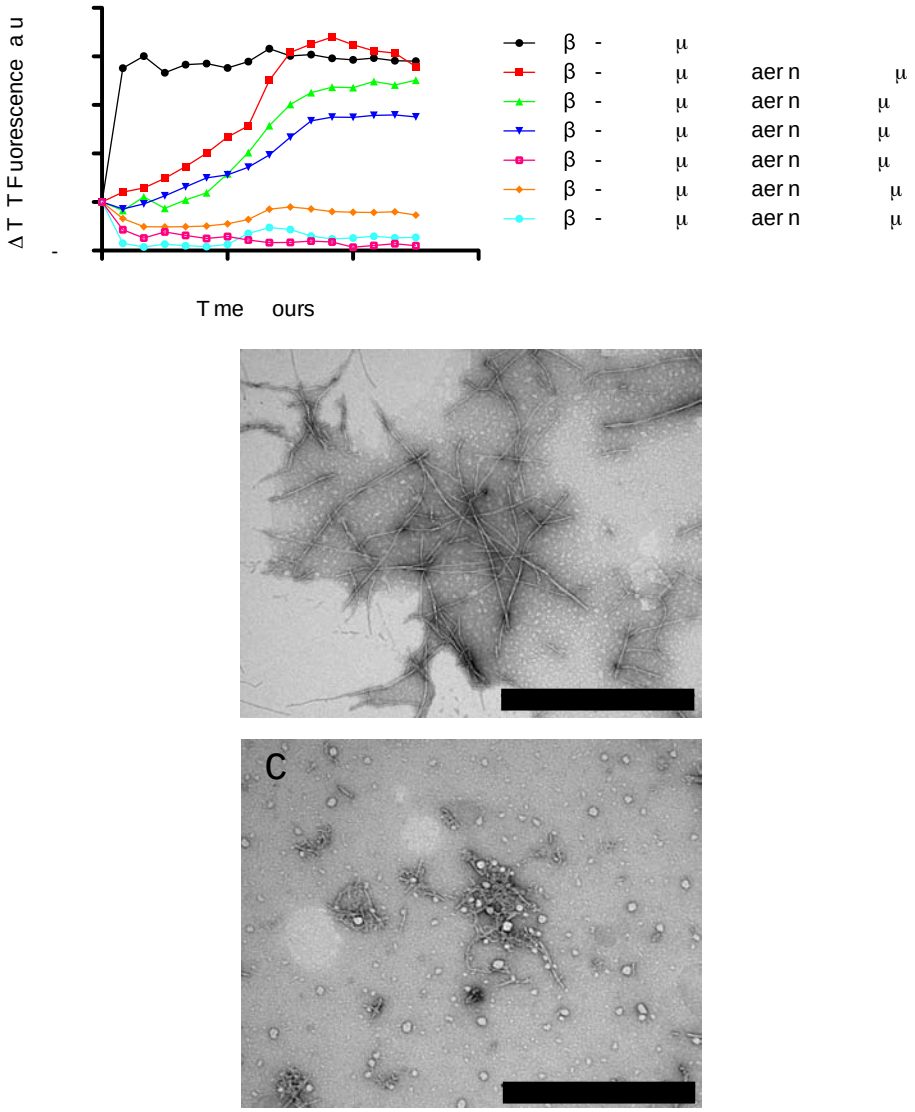


Figure 5.3: Caerin 1.8 inhibits fibril formation by Aβ(1-42). (A) *in situ* ThT fluorescence assay of Aβ(1-42) (25 μM) incubated at 37 °C in the presence and absence of caerin 1.8. Representative data are shown from three independent experiments performed in triplicate. (B and C) TEM micrographs of samples from the *in situ* ThT fluorescence assay. (B) Aβ(1-42) after incubation and (C) Aβ(1-42) after incubation in the presence of caerin 1.8 (100 μM). Scale bars represent 1 μm.

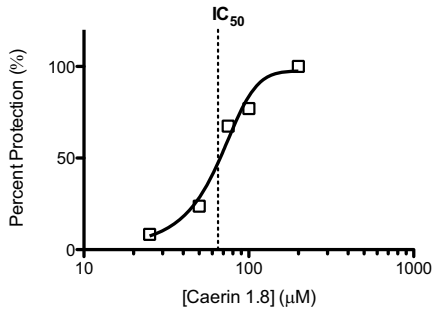


Figure 5.4: The fitted logarithmic concentration-response (percentage of protection) curve for calculating the IC_{50} for the inhibition of fibril formation by Aβ by caerin 1.8.

The effect of truncating caerin 1.8 on its ability to inhibit fibril formation by Aβ(1-42) was monitored by an *in situ* ThT fluorescence assay (Figure 5.5). Peptides are named in Figure 5.5 by adding a number to indicate the number of residues in the peptide (Table 5.4). It can be observed that reducing the length of caerin 1.8 has a significant effect on its inhibition abilities, with at least the first 11 residues required for significant inhibition of fibril formation.

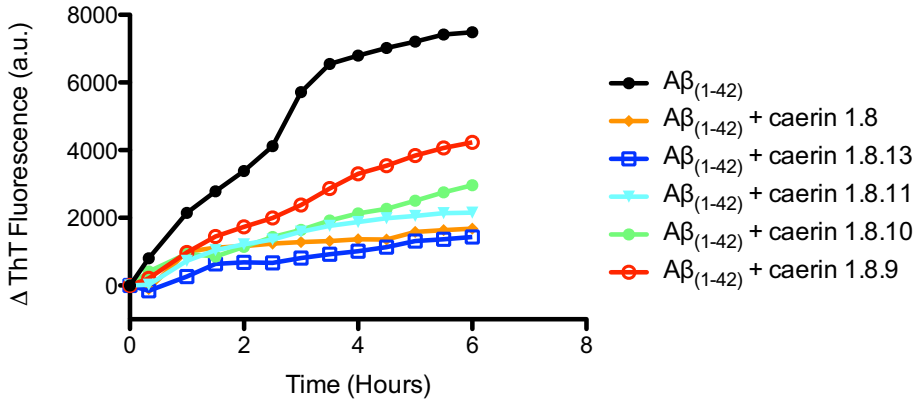


Figure 5.5: The inhibition of fibril formation by Aβ(1-42) due to caerin 1.8 and several of its truncations, as monitored by an *in situ* ThT fluorescence assay. Representative data are shown from three independent experiments performed in triplicate.

Table 5.4: Sequences of truncations of caerin 1.8 tested for their ability to inhibit amyloid fibril formation by Aβ(1-42).

Peptide	Sequence
Caerin 1.8	GLFKVLGSVAKHLLPHVVPVIAEKL-NH_2
Caerin 1.8.13	GLFKVLGSVAKHL-NH_2
Caerin 1.8.11	GLFKVLGSVAK-NH_2
Caerin 1.8.10	GLFKVLGSVA-NH_2
Caerin 1.8.9	GLFKVLGSV-NH_2

5.3.2 Amyloid Fibril Formation by Uperin 3.5

Uperin 3.5 self assembles to form amyloid fibrils in buffered solutions at physiological pH. The fibril formation process was investigated using several methods. Firstly, an *in situ* thioflavin T fluorescence assay was used to monitor the kinetics of fibril formation by uperin 3.5 (Figure 5.6A). Briefly, uperin 3.5 (100 μM, in PBS) was incubated in the presence of ThT (20 μM). An increase in ThT fluorescence is attributed to an increased quantity of β-sheet rich fibril species. The kinetics of the fibril formation process are extremely fast, with no lag phase observed, and the elongation phase being complete within approximately 5 hours.

The secondary structure content of the resultant fibrils, when compared with the native peptide was examined using far-UV CD (Figure 5.6B). It can be seen that incubation induces a structural transition from random coil structure (with a minimum at approximately 195 nm) to β-sheet structure (with a minimum at approximately 220 nm).

The morphology of the produced fibrils was analysed by TEM and AFM (Figures 5.6C and 5.6D, respectively). The fibrillar species produced are approximately 20 nm high and 400 nm in length, with characteristic coiled-coil ultrastructure.

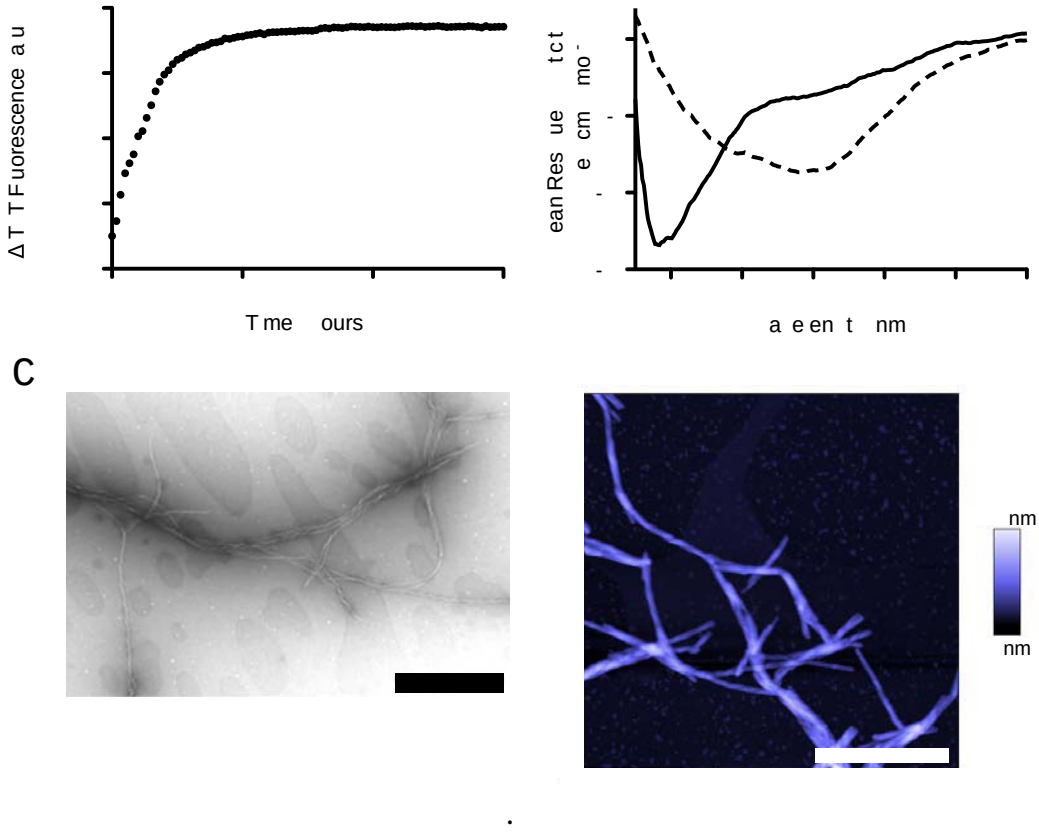


Figure 5.6: Amyloid fibril formation by uperin 3.5. (A) *in situ* ThT fluorescence assay showing the kinetics of fibril formation by uperin 3.5. Representative data are shown from three independent experiments performed in triplicate. (B) Far-UV CD spectra of a freshly prepared solution of uperin 3.5 (3 μM, solid line) and a solution incubated at 37 °C (dashed line). (C) and (D) TEM and AFM images of the fibrillar species produced by uperin 3.5 after incubation at 37 °C for 24 hr. Scale bars = 1 μm.

5.3.2.1 Ion Mobility-Mass Spectrometry

In order to study the self-assembly of uperin 3.5 by IM-MS, it was necessary to determine if the peptide was capable of forming fibrils in solutions of MS compatible buffers. This was achieved by performing an *in situ* ThT fluorescence assay using a 100 μM solution of uperin 3.5 in a 50 mM NH_4OAc solution. The kinetics of fibril formation were not significantly different from those in PBS, and the morphology of the fibrils was also similar, as determined by TEM (data not shown).

After this, the early stage oligomers were studied by IM-MS, using freshly prepared peptide solutions. Figure 5.7A shows the mass spectrum of uperin 3.5, with the oligomers (and their charge states) indicated. Oligomeric states up to a hexamer were observed in the

IM-MS spectrum. The masses of the oligomers, and their observed m/z are summarised in Table 5.5. The experimentally observed m/z values compare favourably with those expected. The oligomers observed all adopted discrete charge states distributions, indicating that they are of relatively compact, well-defined structures. IM-MS allows species of the same m/z, but different mass and charge to be separated. This is illustrated in Figure 5.7B, which shows the ATD of ions with m/z 1781. It can be seen that two different species are present, and separating them allows their respective isotopic distributions to be observed (Figure 5.7C), and the oligomers can then be characterised. Generally, the more highly charged oligomers have the highest mobility, and therefore a shorter drift time. This is observed in these data, with the trimer (+3 charge state) having a shorter drift time than the dimer (+2 charge state).

Table 5.5: Masses of uperin 3.5 oligomers, and the expected m/z for the observed charge states.

Oligomer Order	**Mass (Da)**	**Expected m/z and Charge State**
1	1781	891 (2+)
2	3562	1781 (2+)
		1188 (3+)
3	5343	1782 (3+)
	5343	1336 (4+)
4	7124	2375(3+)
5	8905	2227 (4+)
6	10686	2138 (5+)

ATDs were extracted from the spectra for the oligomeric species observed. The drift times were converted to CCS using a standard calibration protocol [43] and are displayed in Table 5.6. To minimise errors in the conversion process, CCSs were averaged over all the recorded IM-MS spectra, which were each acquired using different IM parameters (wave heights were varied between 8 - 10 V whilst maintaining the wave velocity at 300 ms^{-1}). The error associated with the measured CCSs is estimated to be approximately 8-10 %, taking into consideration errors associated with calibrant cross-sections, the calibration curve and the relative precision of replicate measurements [43].

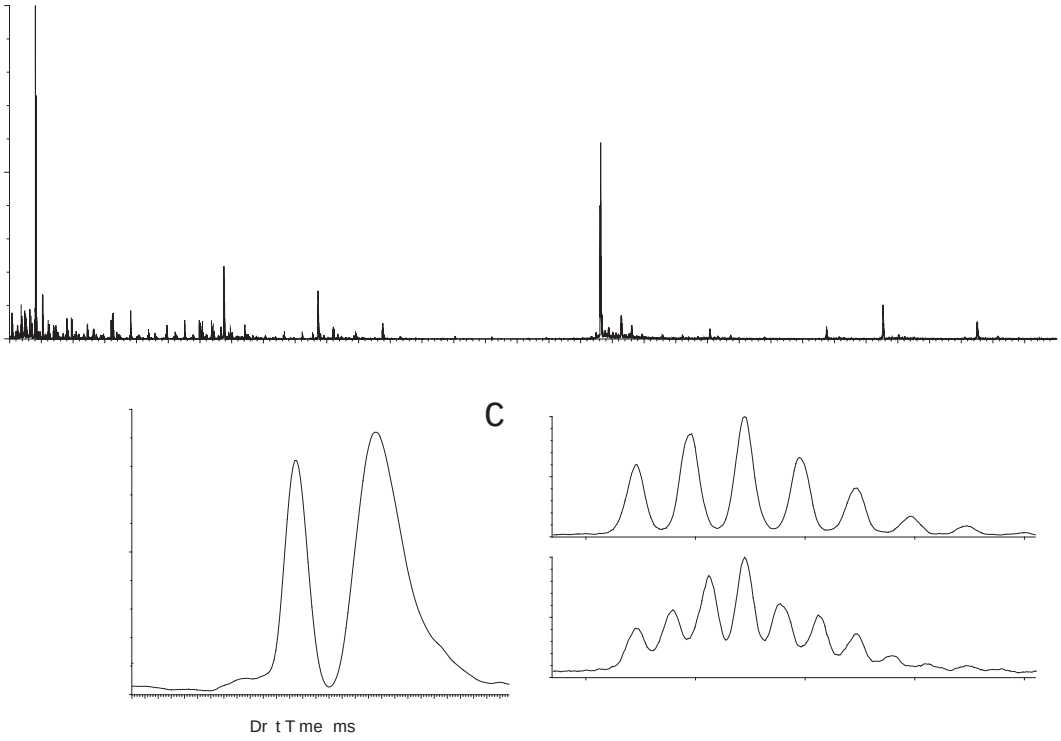


Figure 5.7: (A) Mass spectrum of uperin 3.5 (100 μM uperin 3.5, 50 mM NH_4OAc). The peaks are labelled according to the number of peptides per oligomer and their charge state. (B) The ATD of isobaric ions with m/z 1781. The peak comprises two oligomeric states which can be identified by their isotopic distributions. The 2^{2+} species has a longer drift time than the 3^{3+} species. (C) The isotopic distributions for the 2^{2+} (top panel), and 3^{3+} (lower panel) oligomers.

Table 5.6: Experimental collision cross-sections for uperin 3.5 oligomers. An estimate of the error in CCS measurement is 8 - 10 % [43].

Figure 5.8 displays the experimentally determined CCS of each oligomer plotted as a function of the number of subunits in the oligomer. Also shown is an estimate of the CCS of the oligomers calculated assuming an isotropic growth, in which the CCS of the monomer is scaled by $n^{2/3}$, where n is the number of subunits in the oligomer [559]. Additionally, a linear regression line is also shown fitted to points from oligomers comprising two to six subunits.

Figure 5.8: Measured uperin 3.5 CCSs plotted as a function of the number of subunits in the oligomer. The red line displays estimated CCSs assuming an isotropic assembly process [559]. The dotted line shows a linear regression, r^2=0.99.

5.3.2.2 Modulation of Secondary Structure and Fibril Formation by the Cosolvent 2,2,2-Trifluoroethanol

TFE is known to destabilise hydrophobic interactions within polypeptide chains and stabilise local hydrogen bonds between proximal residues in the amino acid sequence. Consequently, as a cosolvent it is capable of stabilising secondary structure where it has the propensity to form [217, 218]. Addition of TFE (0-50 % v/v) to native uperin 3.5 showed that immediately after dissolution, α-helical structure is induced, as evidenced by the far-UV CD spectra which show characteristic minima at approximately 210 and 220 nm (Figure 5.9B). Conversely in buffered aqueous solution, a far-UV CD spectrum characteristic of random coil structure was obtained.

Addition of TFE to the *in situ* ThT fluorescence assay demonstrates that addition of 10 % v/v TFE enhances aggregation significantly, 20 % TFE, only moderately enhances aggregation, whilst the addition of 50 % TFE eliminates fibril formation (Figure 5.9A).

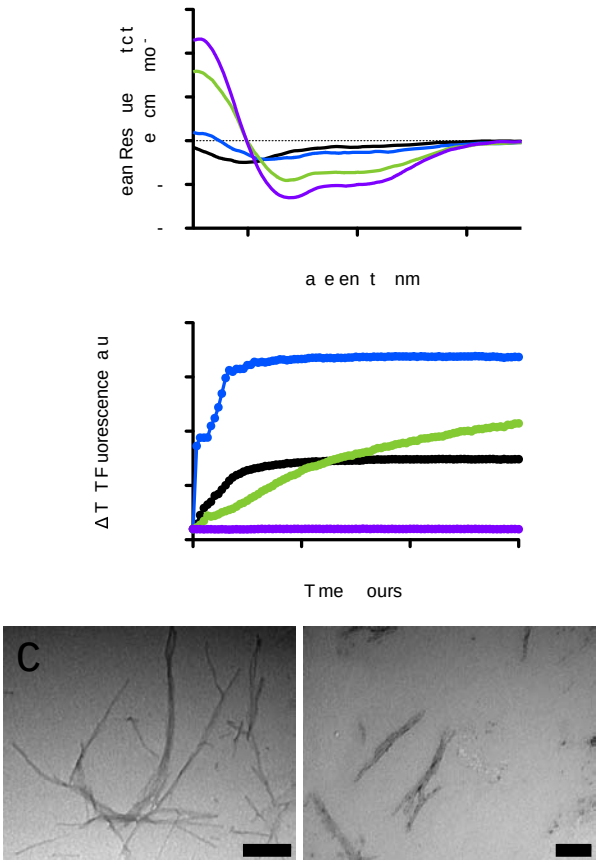


Figure 5.9: Effect of TFE on the conformation and aggregation of uperin 3.5. (A) Far-UV CD spectrum of uperin 3.5 in PBS (black), and in the presence of 10 % (blue), 20 % (green) and 50 % (purple) TFE. (B) *in situ* ThT fluorescence assay showing the kinetics of fibril formation by uperin 3.5 (black), and in the presence of 10 % (blue), 20 % (green) and 50 % (purple) TFE. Representative data are shown from three independent experiments performed in triplicate. TEM images of the fibrillar species produced by uperin 3.5 after incubation at 37 °C for 24 hr in the presence of (C) 10 % TFE and (D) 20 % TFE. Scale bars = 200 nm.

Figure 5.10 shows the far-UV CD spectra of freshly prepared uperin 3.5 and a sample incubated at 37 °C with 10 % TFE. Again, it is apparent there is a shift in the minimum of the far-UV CD spectrum consistent with a structural transition toward more β-sheet secondary structure that is present in amyloid-like fibrils. Solutions incubated in both 20 % TFE and 50 % TFE displayed far-UV spectra that did not differ significantly from the freshly prepared samples (data not shown).

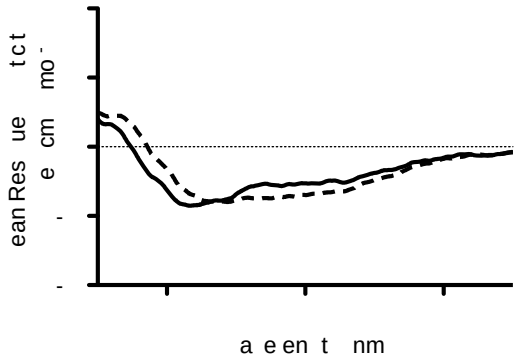


Figure 5.10: Far-UV CD spectra of freshly prepared uperin 3.5 with 10 % TFE (black line) and a solution incubated at 37 °C for 2 weeks with 10 % TFE (dashed line).

5.3.2.3 Effect of Lipids on Fibril Formation

It has been shown that membranes can accelerate, possibly even catalyse, the formation of amyloid [511]. To determine if amyloid formation by uperin 3.5 is enhanced in a lipid environment, the kinetics of aggregation were monitored using an *in situ* ThT fluorescence assay in the presence or absence of small unilamellar vesicles (SUVs) of model membranes.

Phosphocholine and cholesterol are the major constituents of mammalian membranes [560], consequently, SUVs consisting of (1) 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC); and, (2) DMPC supplemented with 20 % cholesterol (Chol), were used as mammalian membrane mimics. As the bacterial membrane is negatively charged, the mixed membrane system DMPC:1,2-dimyristoyl-*sn*-glycero-3-phospho-*rac*-(1-glycerol) (sodium salt) (DMPG) (4:1) was used as a mimic. DMPG is negatively charged, so its inclusion mimics the negatively charged components found on the outer leaflet of the bacterial membrane [561].

It is apparent from the far-UV CD spectra shown in Figure 5.11A, that addition of the neutral mammalian mimetic SUVs does not have a significant influence on the secondary structure of uperin 3.5. It is apparent by the minimum at around 195 nm in these spectra that the peptide is adopting a random coil type structure in this environment. Conversely, the DMPC:DMPG (4:1) containing SUVs cause a change in the far-UV CD spectra, resulting in two minima, and a spectrum closely resembling that of a typical α-helical peptide.

Figure 5.11B displays the results from the *in situ* ThT fluorescence assay. It is apparent that the SUVs containing mammalian mimetic lipids do not have a significant influence on the aggregation process, whilst the addition of SUVs with a net negative charge results in an enhancement of the aggregation process. TEM images showed that the morphology of the fibrils produced were not significantly different from those formed in buffered aqueous solution (data not shown).

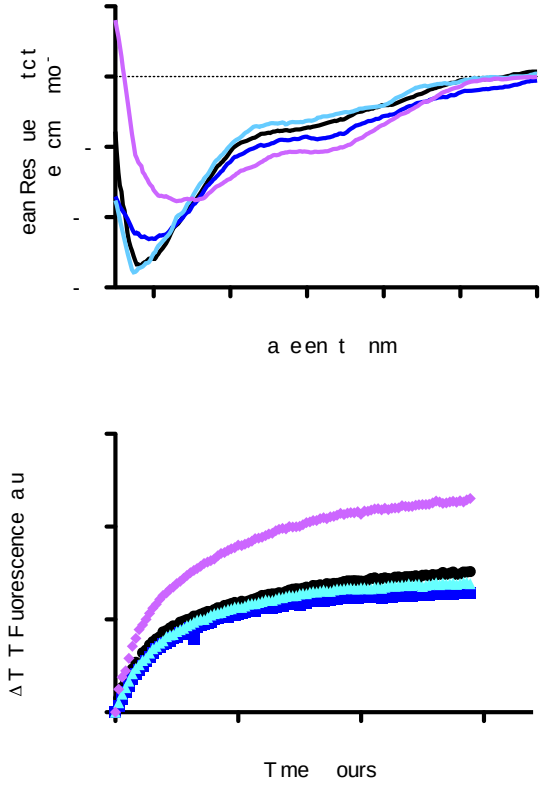


Figure 5.11: Negatively charged lipid SUVs enhance amyloid-like fibril formation by uperin 3.5. (A) Far-UV CD spectra of uperin 3.5 in aqueous buffer (black) and in the presence of SUVs comprised of DMPC (dark blue), DMPC:Chol (light blue) and DMPC:DMPG (purple). (B) *in situ* ThT fluorescence assay showing the kinetics of fibril formation by uperin 3.5 (black) and in the presence of SUVs comprised of DMPC (dark blue), DMPC:Chol (light blue) and DMPC:DMPG (purple). Representative data are shown from three independent experiments performed in triplicate.

5.3.2.4 Cytotoxicity of Uperin 3.5

It is well known that the amyloidogenic form of proteins display enhanced cytotoxicity relative to the unaggregated species [516–518]. To assess the cytotoxity of uperin 3.5, cultured pheochromocytoma-12 (PC12) cells, a neuronal cell model [562], were treated with uperin 3.5 that had been preincubated at 37 °C for 24 hr, along with freshly prepared solutions of the peptide (Figure 5.12). Following incubation of the treated cells for 48 hr, cell survival was assessed by a methylthiazolyldiphenyl-tetrazolium bromide (MTT) assay. The MTT assay measures cell viability by quantifying (by measuring absorbance at 570 nm) the amount of yellow MTT reduced to formazan, which is purple. The reaction is catalysed by succinate dehydrogenase, a mitochondrial enzyme. Consequently, the assay measures mitochondrial activity, and in turn cell viability [562].

It can be seen that cells treated with preincubated uperin 3.5, which had formed amyloid-like fibrils (as evidenced by ThT fluorescence and TEM imaging), caused a dose-dependent decrease in cell viability (Figure 5.12). Conversely, treatment with the native peptide did not result in significant cell death. Treatment with higher concentrations of native peptide (c.f. 50 μM) caused a decrease in cell viability, however, this is likely due to fibril formation in the cell culture media.

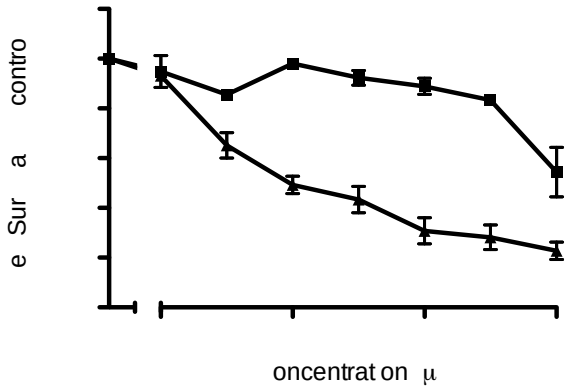


Figure 5.12: Concentration-dependent toxicity of native (square) and fibrillar (triangle) uperin 3.5. Uperin 3.5 (500 μM) was dissolved in PBS and either frozen in its native form or incubated for 24 h at 37 °C to induce fibril formation. Native (square) and fibrillar (triangle) forms of uperin 3.5 (0-50 μM final concentration) were then added to the cell culture media of the PC12 cells and the cells incubated for 48 h. Cell survival was assessed using the MTT assay. Values are presented as percentage of cell survival compared with control ± SEM of three independent experiments, each with six replicates.

5.3.2.5 Inhibition of Fibril Formation by (-)-Epigallocatechin-3-Gallate

(-)-Epigallocatechin-3-gallate (EGCG) (Figure 5.13), and other related molecules, have been the subject of much research due to their ability to inhibit amyloid formation by a variety of disease related proteins [563, 564, 564–566]. The anti-aggregation efficacy of EGCG toward uperin 3.5 was monitored using an *in situ* ThT fluorescence assay. Uperin 3.5 (100 μM) was incubated at 37 °C, in the presence of a 4 molar excess of EGCG. Incubation of uperin 3.5 with EGCG resulted in complete mitigation of any time dependent increase in ThT fluorescence (Figure 5.14A). This is indicative of EGCG preventing fibril formation by uperin 3.5. Addition of EGCG after the elongation phase resulted in a decrease in ThT fluorescence, indicative of fibril disassembly.

In addition, the change in morphology induced by EGCG was monitored using TEM. The images show a redirection of the fibril formation pathway toward amorphous aggregation (Figure 5.14B and C). Addition of EGCG after the elongation phase also resulted in the formation of amorphous type aggregates (Figure 5.14D).

Finally, to assess the cytoprotective activity of EGCG, PC12 cells were treated with uperin 3.5 preincubated at 37 °C for 24 hr with and without EGCG from the *in situ* ThT fluorescence assay. The final concentration of both uperin 3.5 and EGCG in the culture media was 1 μM. Following incubation of the treated cells for 48 hr, cell survival was assessed by MTT reduction. Treatment with EGCG alone does not affect cell survival. Significant protection from amyloid-associated toxicity is observed when uperin 3.5 was coincubated with EGCG (Figure 5.15).

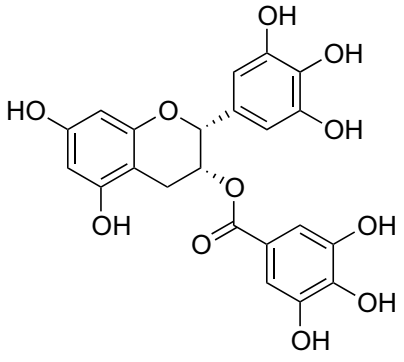


Figure 5.13: The structure of EGCG.

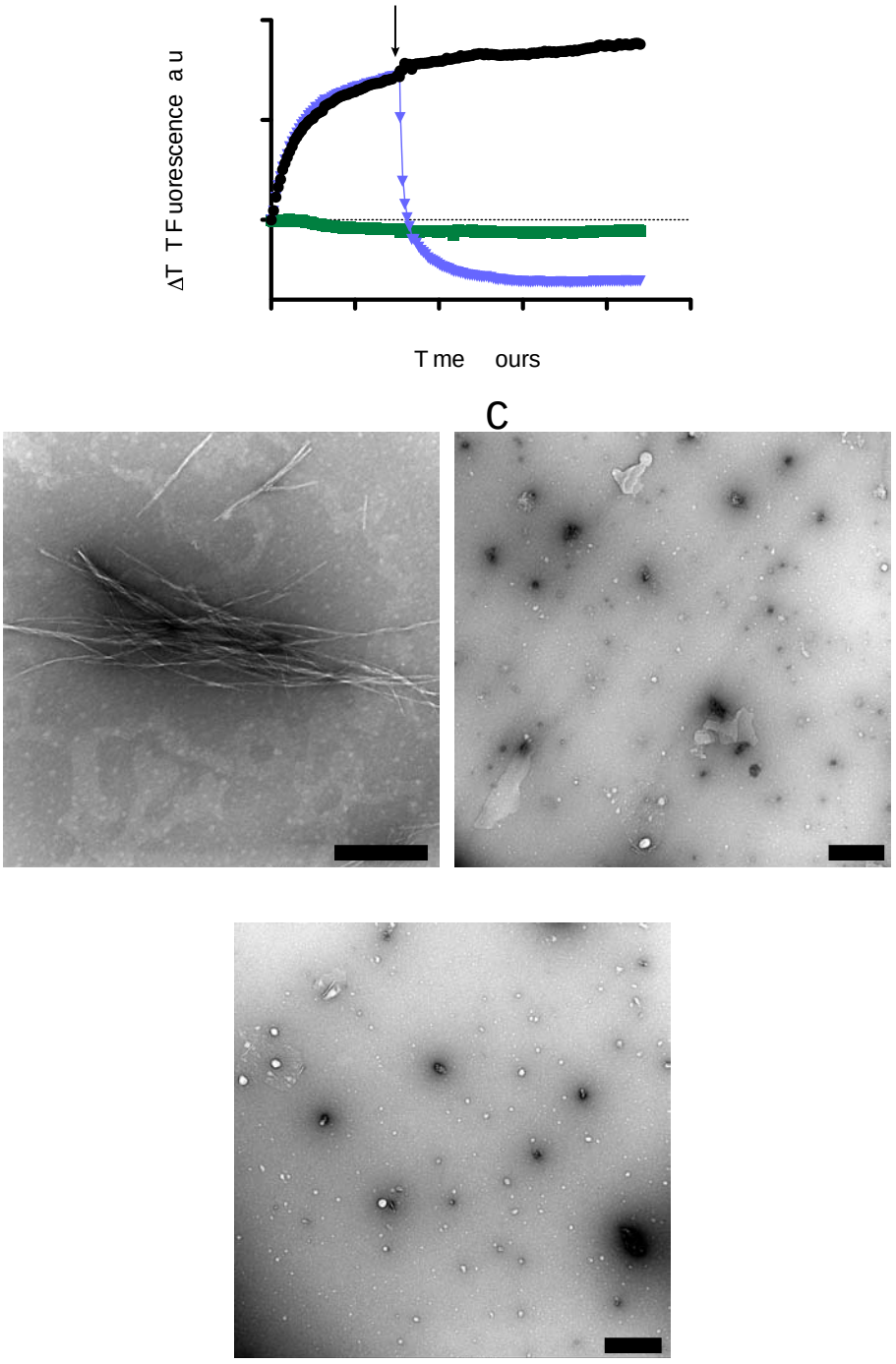


Figure 5.14: EGCG inhibits amyloid-like fibril formation by uperin 3.5. (A) *in situ* ThT fluorescence assay showing the kinetics of fibril formation by uperin 3.5 (black) and in the presence of EGCG (green, 4 molar equivalents). Addition of 4 molar equivalents of EGCG after the elongation phase (purple, addition is indicated by arrow) results in a decrease in ThT fluorescence. Representative data are shown from three independent experiments performed in triplicate. TEM images of the species produced when (B) uperin 3.5 was incubated for 24 hr at 37 °C, (C) uperin 3.5 was incubated in the presence of EGCG (4 molar equivalents) and (D) uperin 3.5 was incubated, followed by the addition of EGCG after the elongation phase. Scale bar = 500 nm.

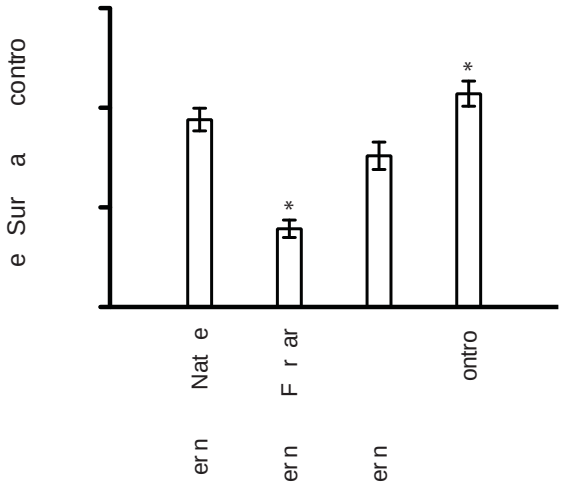


Figure 5.15: EGCG protects against cell death by preventing the formation of uperin 3.5 amyloid-like fibrils. MTT cell survival assay of PC12 cells treated with uperin 3.5 (1 μM final concentration) preincubated at 37 °C for 24 hr in the presence and in the absence of EGCG. Treatment with EGCG alone did not affect cell survival. Values are presented as percentage of cell survival compared with control ± SEM of three independent experiments, each with six replicates. * indicates P <0.05.

5.3.2.6 Effect of Mutations on Fibril Formation

The effects of mutations in the primary sequence of uperin 3.5 was probed to determine how such changes affect aggregation propensity. Several modifications were investigated which alter hydrophobicity and net charge (Table 5.7), as these factors are known to modulate aggregation propensity [567].

Firstly computational algorithms were employed to predict any change in aggregation propensity that may result. These algorithms takes into account the amino acid sequence of the peptide or protein (not any secondary structural elements) and other physio-chemical parameters, such as concentration, pH and ionic strength, to predict aggregation propensity with the assumption that the core regions are fully buried. The results from the analysis using the TANGO algorithm [568] are displayed in Figure 5.16. Figure 5.16A displays the sequence dependent aggregation score, indicating regions of the peptide sequence likely to be amyloidogenic, whilst Figure 5.16B displays the overall predicted aggregation propensity. It is evident that the mutations investigated increase the predicted aggregation propensity, some quite dramatically.

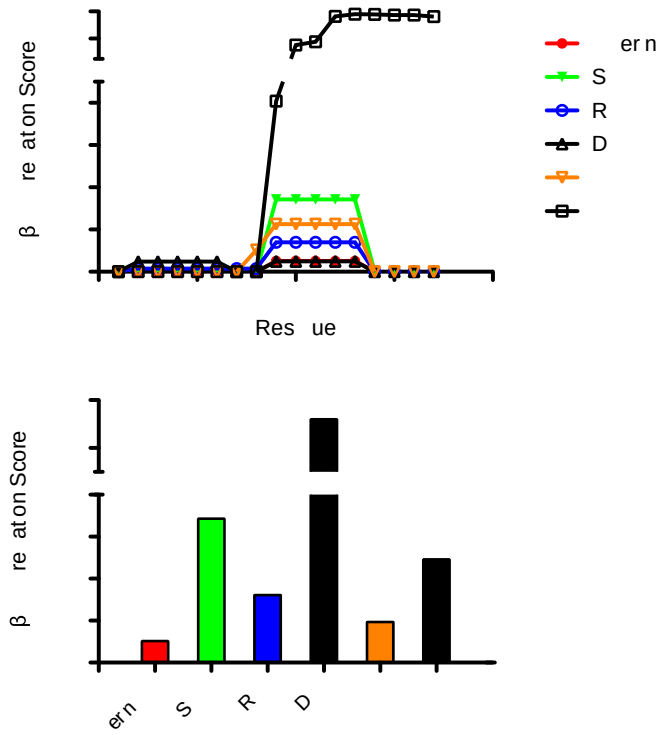


Figure 5.16: Prediction of the aggregation propensity of uperin 3.5 and several synthetic modifications by the TANGO algorithm at pH 7.0. (A) Sequence dependent aggregation propensity scores and (B) overall aggregation propensity scores. Values >0 indicate aggregation-prone segments.

The results obtained using the Zyggregator algorithm are displayed in Figures 5.17. Again, several of the mutations investigated are predicted to increase aggregation propensity. The overall aggregation propensity scores predicted using Zyggregator (Figures 5.17B) mostly match the same order as predicted by TANGO, although the S11A mutant is predicted to have a lower aggregation propensity than uperin 3.5.

To experimentally determine the effects on the aggregation propensity, *in situ* ThT fluorescence assays in combination with TEM were used. The results obtained are summarised in Table 5.7.

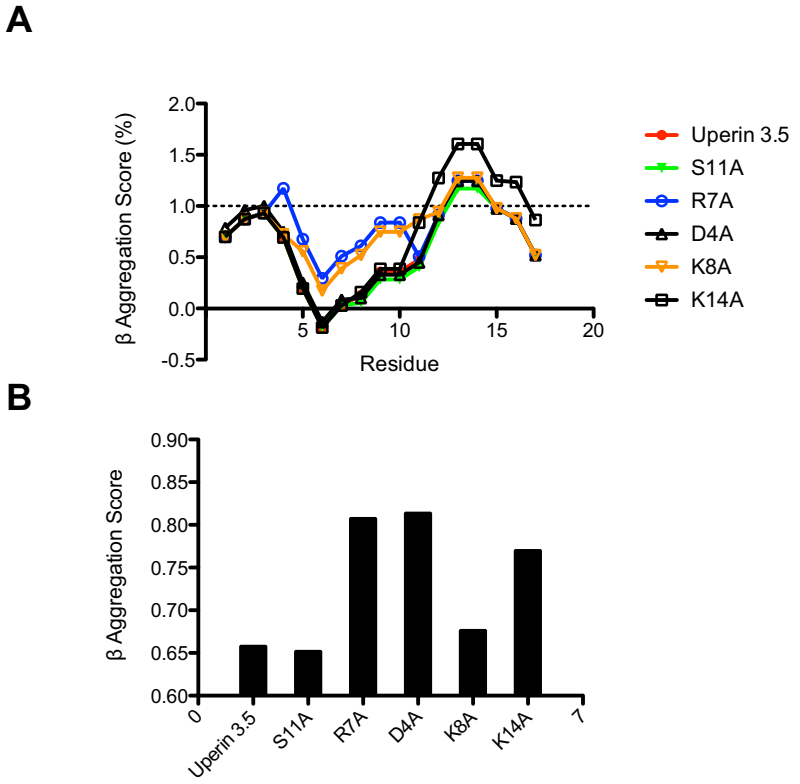


Figure 5.17: Prediction of the aggregation propensity of uperin 3.5 and several synthetic modifications by the Zyggregator algorithm at pH 7.0. (A) Sequence dependent aggregation propensity scores and (B) overall aggregation propensity scores. Values >0 indicate aggregation-prone segments.

Table 5.7: The effect on fibril formation of uperin 3.5 and mutants (100 μM), as monitored by *in situ* ThT fluorescence in PBS (pH 7.4), incubated at 37 °C. Data are indicated as relative, average maximum ThT fluorescence from three independent experiments.

Peptide	Sequence	Charge	Relative ThT Fluorescence Maximum
Uperin 3.5	GVGDLIRKAVSVIKNIV-NH_2	+3	1
G1L Uperin 3.5	LVGDLIRKAVSVIKNIV-NH_2	+3	22
K14A Uperin 3.5	GVGDLIRKAVSVIANIV-NH_2	+2	7
D4A Uperin 3.5	GVGALIRKAVSVIKNIV-NH_2	+4	9.5
K8A Uperin 3.5	GVGDLIRAAVSVIKNIV-NH_2	+2	36
R7A Uperin 3.5	GVGDLIAKAVSVIKNIV-NH_2	+2	36
S11A Uperin 3.5	GVGDLIAKAVAVIKNIV-NH_2	+2	1.9
Uperin 3.5(OH)	GVGDLIRKAVSVIKNIV-OH	+2	0.2
reverse Uperin 3.5	VINKIVSVAKRILDGVG-NH_2	+3	0.1

5.4 Discussion

5.4.1 Caerin 1.8 Inhibits Fibril Formation by Aβ(1-42)

It is hypothesised that the conversion of Aβ peptides to prefibrillar and fibrillar species contribute to the aetiology of Alzheimer's disease [8, 486, 512, 569], and that preventing formation of these species may prove a useful therapeutic tool in halting formation of amyloid deposits in the brain [473, 474, 476, 522]. A number of strategies have been developed to prevent amyloid formation by Aβ, including the use of small molecule inhibitors and also peptides [473, 476, 523]. The work presented in this chapter describes the inhibition of aggregation by Aβ(1-42) using the amphibian peptide caerin 1.8.

A number of design strategies have been employed to develop peptide inhibitors of Aβ aggregation. In particular, peptides have been developed which comprise two important structural domains. The first of these is a recognition motif, which binds to the fibrillar core region of Aβ by hydrophobic interactions. Secondly, a disrupting element is incorporated, to prevent aggregation [473, 476, 523].

The recognition domains that have been incorporated into peptide inhibitors of Aβ aggregation are short, four or five residue domains and have been designed based on the hydrophobic core region of the peptide. Notably, the sequences **KLVFF** and **KLVF** have been shown to be suitable recognition motifs [570–574]. These were identified as they correlate with a highly amyloidogenic region of Aβ. Comparison of these with the sequence of caerin 1.8 (below) shows that the amphibian peptide contains a scrambled **KLVF** motif at its N-terminus. It has been previously demonstrated that scrambling of these motifs does not significantly influence sequence recognition, indicating that only the hydrophobic characteristics of this domain are critical (although this particular scrambled motif has not been characterised) [570]. Consequently, this region may be integral for the recognition of Aβ(1-42) by caerin 1.8.

Aβ(1-42) DAEFRHDSGYEVHHQ**KLVFFAE**DVGSNKGAIIGLMVGGVVIA
Caerin 1.8 G**LFKVL**GSVAKHLLPHVVPVIAEKL-NH_2

In addition to recognition of Aβ, it is imperative that an aggregation inhibitor contains a β-sheet disrupting motif. These include, (i) a charged hydrophilic domain, which disrupts the hydrophobic interactions necessary for fibrillogenesis, (ii) β-sheet breaker domains, which contain amino acids that disrupt the β-sheet core of amyloid fibrils, (iii) bulky groups that sterically inhibit protein aggregation, or (iv) non-natural amino acids (such as α-aminoisobutyric acid and N-methylated or N-alkylated amino acids) to generate peptides which block the hydrogen bonding interactions required for amyloid formation [473, 476, 523].

In addition to the potential **FKVL** recognition element, caerin 1.8 contains several Pro residues, and has an overall charge of +3. The two Pro residues in the peptide may play a role in destabilising β-sheet structures once bound to Aβ, as these are known β-sheet breakers. Additionally, the overall charge may contribute to the effectiveness of caerin 1.8 as a disruptor of amyloid formation. Truncation of caerin 1.8 that were shorter than 11 residues had minimal anti-amyloid activity. This loss of activity is potentially attributable to the removal of Lys11 which results in a decrease in the overall charge of the peptide.

5.4.2 Amyloid-Like Fibril Formation by Uperin 3.5

AMPs and amyloidogenic proteins share many structural and functional characteristics. These include their high affinities for negatively charged membranes, and a structural conversion from random coil to α-helical secondary structure upon interacting with the a membrane [511, 575, 576]. In addition, several membrane bound AMPs have been shown to undergo conversion to amyloid-like fibrils [550]. These are in addition to those which have been shown to undergo this fibril formation in solution [577], including uperin 3.5, which forms well defined β-sheet rich amyloid-like fibrils with fast kinetics *in vitro* (Figure 5.6). Moreover, it is proposed that both amyloids and AMPs have membrane mediated mechanisms of cytotoxicity [511].

5.4.2.1 Ion Mobility-Mass Spectrometry

IM-MS has been used to gain insights into the self-assembly processes of numerous amyloidogenic proteins and peptides. The technique is ideally suited to this endeavour as it possible to characterise heterogeneous populations with relative ease, unlike many other commonly used biophysical methods [94, 95, 578]. It is apparent from the mass spectrum shown in Figure 5.7, that oligomeric species of uperin 3.5 up to a hexamer were observed under the conditions used. In most instances, especially for the higher order oligomers, one distinct charge state was observed for each complex.

Calibration of the measured drift times to give CCSs was performed and calibrated values were plotted as a function of oligomer number to give some insight into the self-assembly process for the oligomers observed (Figure 5.8). It is possible to determine if the self-assembly proceeds by a highly directional fibrillar process as the relationship between CCS and oligomer number would follow a linear correlation [559]. In this case, this is in contrast with the expected CCSs that would be observed for a spatially isotropic self assembly, as indicated in Figure 5.8. It is apparent for the small number of oligomers observed that a linear regression line can be fit to the CCS data as indicated in Figure 5.8, although, these data all lie very close to the values expected for isotropic assembly. Nevertheless, the data indicate that a highly directional, fibrillar type self-assembly process is predominating for uperin 3.5 in the early stages of aggregation.

5.4.2.2 Modulation of Fibril Formation by TFE and SUVs

Solution conditions are known to have effects on fibril formation kinetics by peptides and proteins. For example, the presence of cosolvents, high temperatures and extreme pH, have been known to effect aggregation [579–583]. TFE is a well known organic cosolvent that has the ability to induce secondary structure where it has the propensity to form [218]. Additionally, under certain solution conditions it can mimic a membrane environment [584, 585]. Several proteins and peptides have been shown to form fibrils in TFE solutions [580, 582, 586–588].

It is apparent that addition of TFE to uperin 3.5 triggers conversion from a random coil conformation to give an α-helical structure, as evidenced by the far-UV CD spectra shown in Figure 5.9A. The helical content increases continuously over the 0-50 % TFE concentrations examined. Importantly, the far-UV CD spectra demonstrate that in 10 % TFE, there is only a small amount of structured peptide (as CD spectra are a linear combination of all the secondary structural elements in the sample).

Uperin 3.5 is capable of forming amyloid-like fibrils when starting from a random coil structure in buffered aqueous solution (Figure 5.9). Upon addition of increasing quantities of TFE, the peptide adopts increasing α-helical structure, and amyloid fibrils can be formed in both 10 % and 20 % TFE. When the peptide has more α-helical structure (50 % TFE), amyloid fibril formation is inhibited. From these data, it can be concluded that the relative quantity of α-helical structure, which depends on the concentration of TFE, plays a role in fibril formation by uperin 3.5. TEM imaging shows that the fibrils formed by uperin 3.5 in TFE solutions have similar structural morphology to those prepared in buffered aqueous solution (Figure 5.9).

The effects of TFE in enhancing or preventing fibril formation can be attributed to its influences on both intra- and intermolecular hydrophobic interactions and hydrogen bonding. TFE reduces hydrophobic interactions, possibly leading to protein or peptide unfolding and decreasing intermolecular attractions [589–591]. However, adding TFE to an aqueous solution strengthens intra- and intermolecular hydrogen bonding [592, 593]. At high TFE concentrations, hydrophobic interactions are reduced significantly, preventing intermolecular association from occurring, and consequently inhibiting fibril formation [582]. At intermediate TFE concentrations, however, stabilisation of hydrogen bonding occurs, which may compensate for any any decrease in hydrophobic interactions [582]. Consequently, these conditions may lead to an enhancement of fibril formation.

It is also proposed, that in addition to these effects on inter- and intra molecular interactions, the induction of some α-helical conformation may prove to be an integral structural conversion necessary for the formation of β-sheet rich amyloid fibrils from unstructured peptides and proteins. However, when in a very rigid α-helical conformation, the structural conversion cannot occur. Helical intermediates have been proposed on the amyloid fibril formation pathways of several proteins and peptides [582, 587, 594, 595].

These data are reinforced by experiments showing the effect of SUVs of biomimetic membranes on fibril formation by uperin 3.5 (Figure 5.11). It is apparent that SUV composition plays a vital role in the ability to modulate amyloid-like fibril formation by uperin 3.5. In addition, the structural alterations induced by the SUVs correlates well with their ability to enhance fibril formation. Negatively charged DMPC:DMPG SUVs appear to have an enhancing effect on aggregation. This correlates with an induction of α-helical structure as determined by far-UV CD. Conversely neutral SUVs, DMPC or DMPC:Chol, do not appear to effect amyloid-like fibril formation by uperin 3.5. In these lipid systems the peptide retains its random coil structure. This differential effect is likely a result of electrostatic attraction between the positively charged uperin 3.5 and the negatively charged DMPG. A similar effect has been observed for other amyloid fibril forming peptides, where they have been shown to adopt α-helical structures when bound to membranes, and that this results in enhancement of fibril formation (reviewed in [511]).

These combined data indicate that induction of some α-helical structure has a positive effect on the fibril formation process by uperin 3.5, and this conformation may be a key intermediate in membrane (or TFE) mediated amyloid-like fibril formation.

5.4.2.3 Cytotoxicity of Uperin 3.5 Amyloid Fibrils

The cytotoxicities of a number of non-disease related amyloidogenic proteins and peptides have been investigated, and it has been demonstrated that an there is a correlation between the induction of cytotoxicity and formation of the fibrillar state. This is postulated to be due to the presence of prefibrillar aggregates [8, 509–515], or in some instances the fibrils themselves [516–518]. Conversely, the unaggregated forms of the protein have been shown to have a limited cytotoxic effect.

A similar distinction is observed here for uperin 3.5 using PC12 cells (Figure 5.12), a model neuronal cell line that is widely employed in the study of amyloid toxicity. Both the native and fibrillar species caused a concentration-dependent decrease in PC12 cell survival (Figure 5.12), however, fibrillar uperin 3.5 was significantly more toxic at lower concentrations. For example, at 0.5 μM, fibrillar uperin 3.5 reduced cell viability to 43 $\pm$ 5 %, compared with 92 $\pm$ 3 % for the native peptide. It has been previously demonstrated

that the fibrillar species are maintained in cell culture media, and that fibril formation does still proceed, albeit at a reduced rate [562]. The toxicity of native uperin 3.5 at high concentrations is likely due to the formation of fibrillar species in the culture media, this is due to the high aggregation propensity of uperin 3.5, which lacks a lag phase, in a similar fashion to other studies [562].

These data provide further evidence for the structural similarities between amyloid-like fibrils formed by uperin 3.5 and those of disease-related proteins. Additionally, they provide further support for the notion that the cytotoxicity of amyloid fibril forming proteins is a consequence of the general structures formed by fibril formation, and that there is perhaps a common mechanism of cytotoxicity for all amyloid forming proteins and peptides [512, 569].

5.4.2.4 Inhibition of Fibril Formation by EGCG

Flavonoids, such as EGCG, have been shown to inhibit fibrillogenesis by amyloid forming proteins both *in vitro* and *in vivo*. In particular, EGCG has been shown to inhibit fibril formation by numerous disease and non-disease related amyloid fibril forming proteins such as Aβ, α-synuclein and huntingtin, amongst others [563, 564, 564–566]. In most instances it achieves this by binding to and redirecting the folding of the protein to afford off-pathway oligomers. The mechanisms by which EGCG and other flavonoids achieve this are not well understood, and it is proposed that many different types of intermolecular contacts may mediate this inhibition, including interactions with aromatic residues and amino groups [596–599].

Uperin 3.5 does not contain any aromatic residues, indicating that inhibition of uperin 3.5 aggregation by EGCG cannot be attributed to interactions with aromatic residues. Consequently it is likely that non-specific interactions initiate these inhibition and remodelling processes. The non-specific, broad-ranging nature of EGCG interactions with proteins may explain the effectiveness of EGCG at inhibiting fibrillogenesis by numerous proteins and peptides [563, 564].

It is evident from the results presented herein that EGCG has a similar effect in preventing aggregation by uperin 3.5. When uperin 3.5 was incubated in the presence of four molar

equivalents of EGCG, there was a significant reduction in the maximum ThT fluorescence observed (Figure 5.14). These results were confirmed by TEM analysis to determine the effects on fibril formation morphology. It is evident that co-incubation with EGCG results in the formation of compact spherical oligomers, as opposed to the fibrillar species observed in the absence of EGCG. The mechanism by which this redirection occurs is still to be determined.

In addition, it has been shown that EGCG has the ability to dissagregate and remodel amyloid fibrils formed by a number of disease related proteins [600–603]. The data obtained here shows that EGCG remodels uperin 3.5, that the structures produced are not amyloid fibrils, and that they closely represent the structures produced when EGCG is added at the start of the aggregation process (Figure 5.14).

Finally, it was demonstrated that coincubation of uperin 3.5 with EGCG prevents the cyotoxicty associated with the amyloidogenic uperin 3.5 species (Figure 5.15). However, it should be noted that EGCG and other polyphenols have known beneficial properties, which may also contribute to the prevention of uperin 3.5 toxicity. These include their ability to scavenge radicals, reduce reactive oxidative species and the ability to chelate metal ions [596, 604–606]. However, many other proteins have been studied which give rise to similar morphologies upon EGCG treatment, and in all instances, the soluble spherical aggregates that are formed have been described as non-cytotoxic [563, 564, 564–566].

5.4.2.5 Effect of Mutations on Fibril Formation

A number of physicochemical factors have been shown to influence protein aggregation, including hydrophobicity and charge [567], with hydrophobic interactions arguably the most important driving force for aggregation. Specifically, it has been demonstrated that hydrophobic interactions within Aβ and other amyloidogenic proteins promote fibril formation, but only when the regions involved are incorporated into the fibrillar core [607–610]. It is thought that these hydrophobic interactions provide the initial intermolecular contacts, promoting further intermolecular associations [567]. Consequently, several mutations of uperin 3.5 were studied which altered both hydrophobicity and net charge.

All mutations of uperin 3.5 that increased hydrophobicity resulted in an enhancement of fibril formation (Table 5.7). This confirms that hydrophobic interactions are necessary in fibril formation by uperin 3.5. Notably, substitutions of Arg8 and Lys8 for Ala significantly enhance aggregation. This supports the notion that residues around the central portion of the peptide are involved in hydrophobic interactions which enhance amyloid fibril formation, and that this region of the peptide may be incorporated into the fibrillar core.

The net charge of a peptide has also been shown to have an effect on fibril formation. In the literature it is proposed that in many instances, reduction of the net charge increases aggregation propensity [567, 611–613]. Differential effects of charge are observed here for uperin 3.5 aggregation. Notably, all the amino acid substitutions which decrease net charge from +3 to +2 result in an enhancement of aggregation. Conversely, removal of the C-terminal amide and the consequent reduction of net charge to +2 reduces aggregation propensity. However, at high (mM) concentrations, this peptide still does aggregate (data not shown). This indicates that some electrostatic interactions are necessary to enhance the aggregation of uperin 3.5. The enhanced aggregation propensity that comes from a decrease in net charge, and increase in hydrophobicity, in the central portion of the peptide again supports the notion that this central region is incorporated into the fibrillar core.

5.4.2.6 Implications of Fibril Formation on the Mechanism of Antimicrobial Activity

Much work in the area of AMPs is focussed on determining their structures and modes of action, as a means of implementing and improving peptide antibiotics. It has been proposed that for some AMPs, namely the temporins B and L, their amyloidogenic nature means that soluble peptide oligomers may be responsible for their antimicrobial activity, and that the membrane may mediate amyloid formation. In other words, there is an association between the toxic function of the AMP and its ability to aggregate [550]. Consequently, the activities of amyloidogenic AMPs, such as uperin 3.5, may all follow a similar membrane mediated mechanism, as proposed in Figure 5.18.

In brief, it is proposed that the peptide, which has random coil structure in solution, is electrostatically attracted to the negatively charged membrane surface (A, in Figure 5.18).

Subsequently, intercalation into the bilayer results in a conformational change from random coil to amphipathic α-helix (B). This is followed by membrane insertion of the peptide perpendicular to the axis of the bilayer (C). To enable complete traversal of the bilayer, dimerisation may be required for short peptides, in a similar fashion to the traditional models describing AMP function. This process also requires reorientation of the anionic phospholipids as the peptides are amphipathic in nature (such that the hydrophobic side of the peptides face the lipid bilayer, as shown in Figure 5.19). The α-helices of multiple monomers may align (D), possibly by formation of an α-sheet [614] (E), and subsequently undergo a structural shift to form β-sheet conformations and β-sheet rich oligomers (F). It is possible that these oligomers may further aggregate to amyloid-like fibrils, possibly with the incorporation of lipids from the bilayer [549, 550, 615].

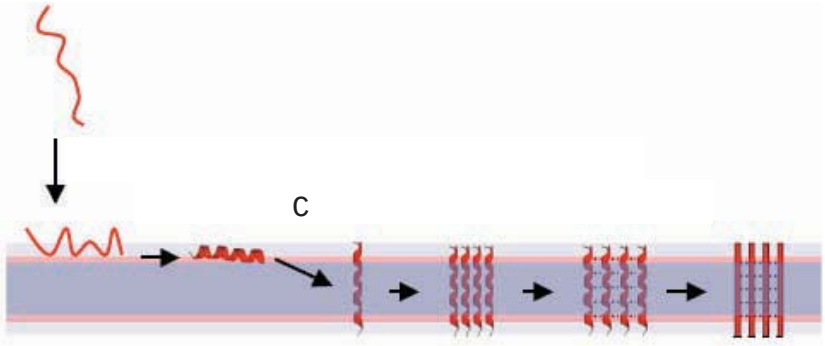


Figure 5.18: Schematic diagram showing peptide binding to a lipid bilayer, followed by insertion and oligomerisation. Adapted from [550].

The mechanism of membrane mediated oligomerisation detailed above is in many respects similar to the intermolecular associations required to undergo pore formation as detailed in Figure 5.2. However, it is proposed that the oligomers formed by amyloidogenic AMPs in a membrane may have a mechanism of membrane permeabilisation different to both the barrel-stave and toroidal pore mechanisms. This mechanism is termed the 'leaky slit' model, and is based on the formation of toxic oligomers in the bilayer, by the mechanism described above [549, 550, 615, 616]. This results in the formation of transmembrane linear, amphipathic arrays of peptides (Figure 5.19), which arrange with the hydrophobic faces interacting with the bilayer. Consequently, the hydrophilic face forces the lipids on this side to adopt a curved structure which joins the inner and outer leaflets of the membrane, resulting in membrane disruption. This model indicates that the secondary

structure of the amyloidogenic AMP is irrelevant. Consequently, this provides a generic model applicable to a variety of different scenarios.

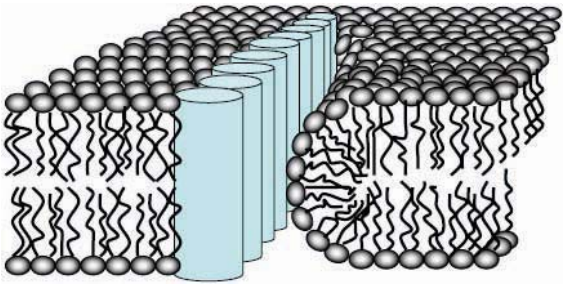


Figure 5.19: The proposed 'leaky slit' model for membrane disruption due to an amphipathic AMP oligomer. From [616].

The apparent functional link between the membrane mediated oligomerisation and fibrillisation of AMPs, such as uperin 3.5, and their mechanism of antimicrobial action requires much further research. Specifically, the membrane interactions of uperin 3.5 should be probed more closely to gain some information about the mechanistic role of its amyloidogenic nature. It is apparent that work of this nature can provide significant insights into the mechanism of disease related amyloid cytotoxicity [549].

5.5 Experimental Procedures

5.5.1 Materials

Peptides were synthesised commercially by GenicBio (Shanghai, China), using L-amino acids and standard solid-phase methods, and were typically greater than 85 % pure (by High Performance Liquid Chromatography and MS). Aβ(1-42) was purchased from Anaspec (Fremont, CA, USA).

5.5.2 *in situ* Thioflavin T Fluorescence Assay

Aβ(1-42) was prepared by dissolving in 1.0 % NH_4OH at a concentration of 2.5 mM, before dilution to 250 μM with water. Stock solutions of all other peptides were prepared by dissolving in water at a concentration of 500 μM. The peptide stock solution was divided into aliquots and stored at - 80 °C.

Samples for the *in situ* ThT assay were prepared by diluting the concentrated peptide stock solution into phosphate buffered saline (PBS) solution at the appropriate concentrations. The final concentration of ThT in the samples was 10 μM. Samples were incubated at 37 °C in black μClear 96-microwell plates (Greiner Bio-One, Stonehouse, UK) that were sealed with clear film (Sigma Aldrich, Castle Hill, NSW) to prevent evaporation. The ThT fluorescence intensities of the samples were recorded using a Fluostar Optima plate reader (BMG Labtechnologies, Offenburg, Germany) fitted with 440/490 nm excitation/emission filters. Data were normalized by plotting the change in ThT fluorescence (ΔF, arbitrary units [a.u.]) from the initial fluorescence reading.

5.5.3 Transmission Electron Microscopy

Aliquots (5 μL) were taken from samples after the ThT fluorescence assay and applied to the surface of carbon-coated 400-mesh nickel TEM grids (ProSciTech, Queensland, Australia). The grids were washed with milliQ H_2O (3 $\times$ 10 μL) and negatively stained with uranyl acetate solution (2 % w/v, 10 μL). Samples were viewed using a Philips CM100

transmission electron microscope (Philips, Eindhoven, The Netherlands) or a Tecnai G2 Spirit transmission electron microscope (FEI Company, Hillsboro, OR, USA).

5.5.4 Atomic Force Microscopy

AC mode AFM images were obtained using a JPK Nanowizard III AFM (JPK Instruments AG, Berlin, Germany) installed on a Nikon TE-2000 inverted epifluorescence microscope. Cantilevers used were Bruker NCHV model, with a nominal resonant frequency and spring constant of 320 kHz and 42 N/m respectively. Line rates of 1-3 Hz were used, and images were processed using the JPK instrumental software. Samples for AFM were placed as solution onto freshly-cleaved mica discs and dried; imaging was performed in air.

5.5.5 Ion Mobility-Mass Spectrometry

Uperin 3.5 was dissolved in 20 mM NH_4OAc at a concentration of 100 μM. IM-MS spectra were acquired on a Synapt HDMS system (Waters, UK) [48], using nanoESI in the positive ion mode. The sample was introduced using platinum-coated borosilicate capillary needles that were prepared in-house. Instrument parameters were optimised to remove adducts whilst preserving non-covalent interactions, and were typically as follows; capillary voltage, 1.7 kV; cone voltage, 100 eV; trap collision energy, 10 V; source temperature, 50 °C; backing pressure, 3.5 mBar; IMS cell pressure (N_2), 0.5 mBar; travelling wave velocity, 300 ms^{-1}; travelling wave height, 8-10 V.

Drift-time measurements obtained from the Synapt HDMS were normalised for charge state and a nonlinear correction function was applied for calibrant ions such that their relative differences mirror those previously observed for the same ions [43, 122, 456, 457]. CCSs of the reference samples were taken from the literature using values for ubiquitin, myoglobin and cytochrome c [119]. The CCS calibration procedure is described in detail elsewhere [43].

5.5.6 Circular Dichroism Spectroscopy

Far-UV CD spectra were recorded on a J-815 CD spectropolarimeter (Jasco, MD, USA) using a 0.1-cm path-length cuvette and averaging five scans. The buffer contribution was subtracted for each experiment.

5.5.7 Cell Culture

PC12 cells were cultured in RPMI 1640 medium (Invitrogen, California, USA) supplemented with 5 % v/v foetal bovine serum, L-glutamine (2 mM), and antibiotic/antimicotic mixture (100 mg/mL penicillin and 100 units/mL streptomycin), and nonessential amino acids (100 units/mL). Cells were grown in uncoated 75 cm^2 plastic flasks at 37 °C in a 5 % CO_2-humidified incubator and subcultured every 3-7 days.

5.5.8 Cytotoxicity Assay

The MTT assay was used to assess cytotoxicity of the peptides. Cells were seeded the day prior to treatment at a density of 2×10^4 cells per well in a 96-well plate. Peptides (500 μM) were dissolved in PBS and either snap-frozen on dry ice and stored at -20 °C (non-fibrillar) or incubated at 37 °C for 24 hr. After incubation the samples were diluted in PBS to the required concentrations (10 μL) and added to full-serum RPMI 1640 (100 μL). After a treatment time of 72 hr, the culture media was removed and the cells incubated with serum-free media containing MTT (0.25 mg/mL) for 3 hr. The media was removed, replaced with dimethyl sulfoxide (100 μL) and the absorbance of the resultant formazan solution was measured at 560 nm using a BMG Polarstar microplate reader (BMG Labtechnologies, Offenburg, Germany). Cell viability was assessed as percentage absorbance relative to the vehicle control (PBS only) as a mean of three independent experiments (six replicates per experiment). Where appropriate, differences between data sets were evaluated by performing analysis of variance (ANOVA) followed by DunnettâĂŹs test. A level of $P < 0.05$ was considered to be significant.

Chapter 6

Aspartic Acid Isomerisation in Amphibian Peptides

6.1 Introduction

Proteins and peptides are susceptible to many spontaneous non-enzymatic modifications under physiological conditions [617]. The presence of these modifications can be seriously detrimental to the protein/peptide, often as a result of perturbations to 3D stucture and overall stability. Alternatively, they may present few or no challenges to the ability of the biomolecule to perform its biological function [617, 618]. The occurrence of these modifications is often determined by the sequence, structure, half-life and microenvironment (pH, temperature, redox potential, osmotic conditions) of the protein, and often they occur in a time-dependent fashion [617, 619]. Consequently, their cumulative effects, mainly associated with the ageing of biomolecules, can have a significant effect on activity, a phenomenon sometimes called 'protein fatigue' [620]. Conversely, other post-translational modifications, particularly those which are enzyme catalysed, are vital for the diversification of the proteome and to ensure proteins can perform their appropriate biological function [621].

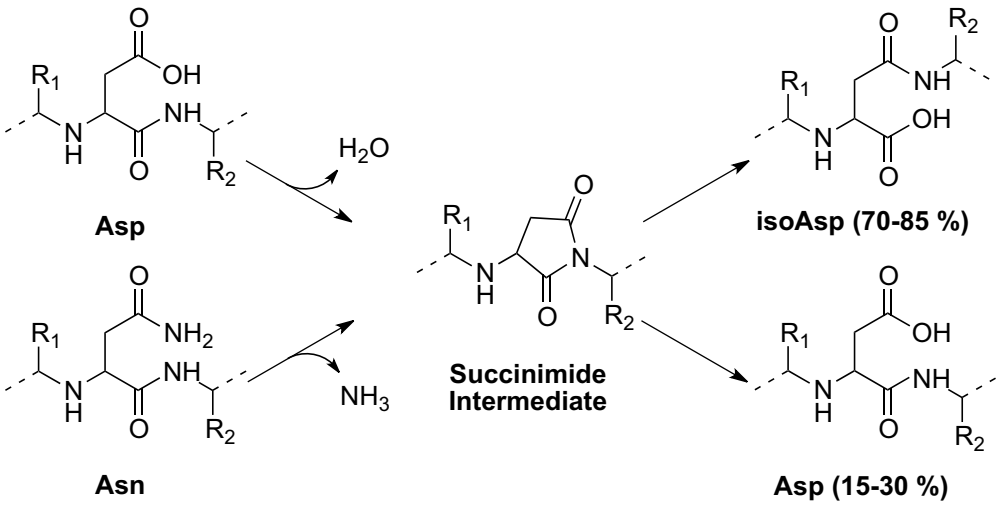
The non-enzymatic covalent modifications that a protein/peptide can undergo are varied, and include isomerisation, deamidation, racemisation, oxidation, glycation, cross-linking, β-elimination, dehydration and hydrolysis [617, 618]. Often specific amino acid residues are required to be present for any one of these reactions to occur (cf. Asp isomerisation

and Asn deamidation, Section 6.1.1), and it has been shown that some regions of structure, particularly unstructured and solvent accessible regions, may be more susceptible than others to non-enzymatic covalent modification. Additionally, the absence or inefficiency of repair mechanisms may result in the accumulation of these modifications [617, 618].

In addition to their prolific presence *in vivo*, it has been proposed that many of these modifications may be incorporated into rationally designed peptide/protein therapeutics [622, 623]. They may confer many advantages on these systems, such as minimising proteolysis, but it is imperative that their incorporation does not significantly affect the activity of any potential therapeutic [622–624].

6.1.1 isoAspartic Acid

The atypical isoAspartic acid (isoAsp, β-Asp) residue is formed as a result of the isomerisation of Asp or deamidation of Asn residues, resulting in the addition of an extra CH_2 into the peptide backbone [625–630]. The processes begins with nucleophilic attack on the carbonyl group of the side chain by the amide nitrogen in the peptide bond of the C-flanking residue, resulting in the formation of a 5-membered cyclic imide (or succinimide) ring intermediate. The succinimide ring is then hydrolytically opened at the α- or β- carboximide to form isoAsp or Asp, respectively, in an approximately 3:1 ratio (Scheme 6.1).



Scheme 6.1: Proposed mechanism for the formation of isoAsp from Asp and Asn [628].

The process of isoAsp formation is affected by the residue that is present on the C-terminal side of Asp/Asn, with Gly, Ser and His known to assist the reaction [626–630]. Conversely, the presence of large hydrophobic side chains on the residue following Asp/Asn generally disallows isoAsp formation, however an example of Asp isomerisation occurring in a cyclic peptide with Leu following Asp has recently been reported [631]. The residue preceding Asp/Asn has been found to have little effect on succinimide formation and the subsequent isomerisation (of Asp) or deamidation (of Asn) [626, 628, 630].

In addition, the local secondary and tertiary structure has a significant influence on the rate of isoAsp formation, with flexible regions of structure more susceptible to the phenomenon, whilst isoAsp is less likely to be found in regions of regular secondary structure like α-helices and β-sheets [632]. Consequently, data indicating the degree of flexibility [633] in a protein have been accurately used to predict the major sites of isoAsp formation in many proteins [634–636], and automated approaches have been developed to aid this process [637, 638].

6.1.2 Consequences of isoAspartic Acid Formation

The non-enzymatic process of isoAsp formation is a frequent outcome of age related protein degradation. This severe covalent modification may or may not affect both the tertiary/quaternary structure and activity of the polypeptide, as a consequence of the side chain CH_2 being inserted into the backbone [625–630, 639].

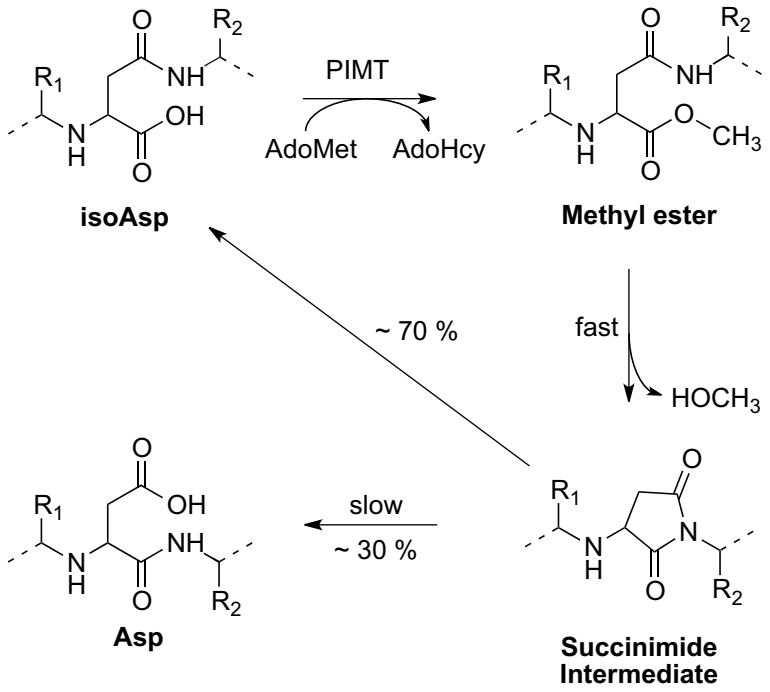
A number of examples pertaining to the disruption of bioactivity by isoAsp, and its association with a variety of disease states have been described in the literature. Notably, peptides containing isoAsp have been implicated in the amyloid diseases Type 2 diabetes mellitus [640], cataract [641–647] and Alzheimer's [638, 648–653] (Section 5.1). Specifically, deamidation of Asn in insulin [654] and isomerisation of Asp in Aβ [650, 651], have been shown to trigger amyloid formation and pathogenic aggregation. More generally, formation of isoAsp is correlated with an increased propensity to form β-sheet structures [632, 655, 656]. Additionally, the presence of isoAsp has been correlated with other diseases such as lupus [657], coeliac disease [658, 659] and cancer [660–663].

The presence of isoAsp residues perturbs protease recognition and subsequent protein degradation [628, 664–666]. In support of this observation, dipeptides of isoAsp-Xaa (where Xaa is varied) have been found in urine, possibly a consequence of incomplete lysosomal and proteasomal degradation because this unusual amino acid is not recognised by proteolytic enzymes [628]. Additionally, *in vitro* proteolytic digestion, especially prior to MS experiments, has been shown to be significantly affected by isoAsp [667–669].

6.1.3 Detection and Repair of isoAspartic Acid *in vivo*

The enzyme protein L-isoaspartyl methyl-transferase (PIMT) [670, 671] is involved in a natural isoAsp repair mechanism, and catalyses the methylation of the α-carboxyl group of isoAsp, using S-adenosyl-L-methionine (AdoMet) as a substrate (generating S-adenosyl-L-homocysteine, AdoHcy), promoting cyclisation of isoAsp to the succinimide with the release of methanol (Scheme 6.2) [627–629, 672, 673]. In nature, this cyclic intermediate can reform Asp in a 3:7 ratio with isoAsp, repairing some of the residues (only 3 in 10). This sequence of methylation, followed by succinimide formation and hydrolysis occurs repeatedly to convert many isoAsp residues back to Asp, although this is obviously a relatively inefficient repair mechanism. It is also proposed that methylation by PIMT may also be a trigger for proteasomal degradation, rather than a repair mechanism *in vivo* [628]. However, it is more favourable and efficient to undergo repair, due to the energy intensive processes required for degradation and protein resynthesis [628]. Despite this repair mechanism, should isoAsp have been formed *in vivo* as a result of Asn deamidation, the repair by PIMT is incomplete (as it affords Asp, rather than the native Asn). Consequently, even after repair the protein may prove to be non-functional.

A variety of analytical methods have been developed to determine the presence of isoAsp, including some which exploit the natural repair process by using radiolabelled AdoMet to selectively methylate isoAsp [627, 629, 634]. The radiolabelled methanol produced upon spontaneous succinimide reformation is then quantified to determine the presence, but not location, of isoAsp [627, 629, 634]. A number of other methods have been developed to infer the presence of isoAsp, including monitoring for the failure of Edman degradation at the residue during sequencing reactions [674, 675], differences in proteolytic digestion



Scheme 6.2: Proposed mechanism for the formation for the repair of isoAsp by PIMT [628].

[667–669], antibody recognition of isoAsp containing sequences [648, 651, 676], a variety of MS-based methods ([673, 677], and references therein) as well as NMR [678, 679].

6.1.4 isoAsp in Amphibian Peptides and Potential Therapeutics

Secretions of the amphibian skin contain a complex array of peptides, many with potential therapeutic properties (Section 4.1.3). To date, only two peptides isolated from anuran skin secretions have been shown to contain isoAsp; the caeridins 1.2 and 1.5, isolated from the skin secretion of *Litoria gilleni* and *L. chloris* (Table 4.1) [400, 680]. These peptides differ from the caeridins 1.1 and 1.4 by isomerisation of Asp4 to isoAsp. The role of all four of these peptides in the secretions of these animals is unknown, with their activities undetermined. Testing has shown they do not have neuropeptide activity, and show minimal or no antimicrobial/antiviral activity [400, 680]. These peptides have a Gly residue following Asp, a residue which aids the isomerisation process.

The Asp to isoAsp conversion, or the incorporation of other β-amino acids into bioactive peptides, has the effect of decreasing their susceptibility to proteolytic degradation and thereby increasing their half life *in vivo* [628, 664–666]. Consequently, there is much

research into the incorporation of isoAsp and other β-amino acids in potential peptide based therapeutics [622]. However, there is a possibility that should this substitution be performed on a natively Asp or Asn containing peptide, this may influence the bioactivity of the peptide [624]. Should this not be the case, and bioactivity is retained, or even enhanced, the increased resistance of the peptide to proteases may confer on it potentially beneficial pharmaceutical properties. Furthermore, the possibility of unwanted isoAsp formation is proving to be an integral consideration in the area of protein/peptide therapeutics, to ensure their stability over time [681, 682]. Should isoAsp formation occur (by isomerisation/deamidation) once a potential therapeutic is administered or whilst it is in storage, immunogenicity may result [683]. Alternatively, a reduction in activity may occur, or indeed an alternate activity may be acquired [684], both cases being extremely unfavourable.

6.2 Aims

The aim of research described in this chapter was to determine the effect of an Asp to isoAsp substitution in the wide spectrum antimicrobial peptide citropin 1.1, from *Litoria citropa* [402] (Table 6.1), which is active at micromolar concentrations against Gram positive pathogens. Specifically, to determine the implications of isoAsp formation on the 3D structure (determined by NMR spectroscopy in membrane mimicking solvents), antimicrobial activity and susceptibility to proteolytic cleavage. This isoAsp isomer is called isoAsp4-citropin 1.1 throughout this chapter.

The primary sequence of this peptide disfavours isoAsp formation (due to the presence of a bulky hydrophobic Val residue on the C-terminal side of Asp), consequently the isomeric peptides can be studied without concern for isomerisation occuring during the experiments. It is not possible to use the naturally occurring isoAsp-containing amphibian peptides in these studies as their activities remain unknown [400, 680]. Citropin 1.1 forms an amphipathic α-helix in membrane mimics, and this peptide was chosen so that the first high resolution structure of an isoAsp containing peptide with ordinarily helical structure could be obtained. Additionally, a number of other amphibian neuropeptides were investigated to examine the effect of isoAsp formation on their activities (Table 6.1).

Table 6.1: Sequences of the Asp containing amphibian peptides studied.

Name	**Sequence**
Citropin 1.1	GLFDVIKKVASVIGGL-NH_2
Crinia angiotensin II	APGDRIYHPF-OH
Uperin 1.1	pEADPNAFYGLM-NH_2
Uperolein	pEPDPNAFYGLM-NH_2
Physalaemin	pEADPNKFYGLM-NH_2

6.3 Results

6.3.1 NMR Spectroscopy of isoAsp4-Citropin 1.1

6.3.1.1 NMR Assignment

NMR spectra of isoAsp4-citropin 1.1 were recorded in DPC micelles at pH 6.0. The proton resonances were assigned using the sequential assignment strategy outlined by Wüthrich [160] and a combination of TOCSY, DQF-COSY and NOESY spectra. These assignments are indicated in the partial TOCSY and NOESY spectra in Figure 6.1.

There were several coincident amide resonances, creating some ambiguity in the assignment process. DQF-COSY data and NOE connectivities were used to unambiguously assign the chemical shifts. The NH region of the NOESY spectrum was utilised to sequentially assign the amide proton chemical shifts. Where sequential d_{NN} NOE cross-peaks were not present, $d_{\alpha N}$ correlations were utilised to obtain reliable resonance assignments.

There was significant overlap of the α ^{1}H/α ^{13}C region with the residual water signal in the ^{13}C HSQC spectrum, therefore a number of α ^{13}C resonances could not be reliably assigned. Consequently, these data are not presented here.

Table 6.2 summarises the assigned ^{1}H resonances.

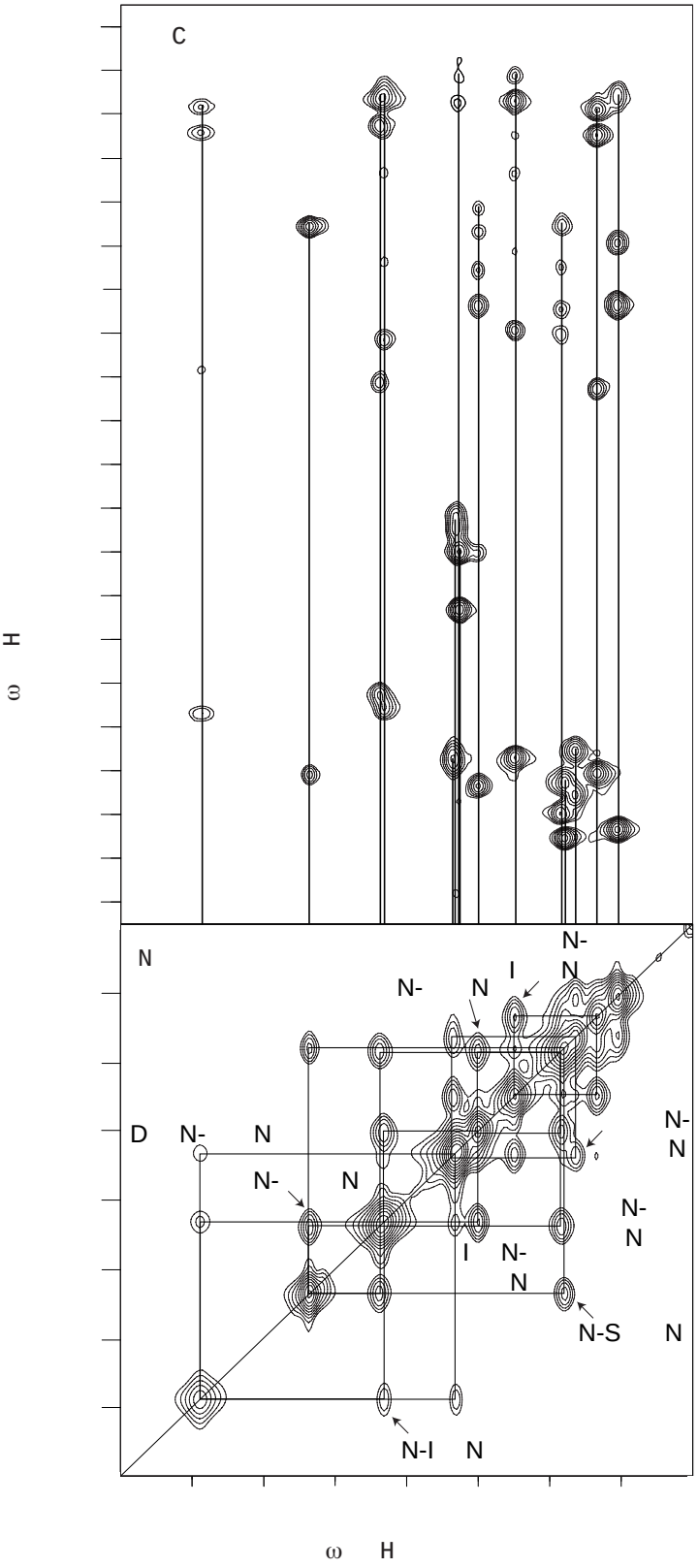


Figure 6.1: Partial NOESY and TOCSY spectra of isoAsp4-citropin 1.1 in DPC micelles. In the TOCSY spectrum, lines connect resonances of the same amino acid residue. NOEs between sequential amide protons are shown in the NOESY spectrum.

Table 6.2: Assigned ^{1}H chemical shifts for isoAsp4-citropin 1.1 in DPC micelles. n.o. indicates resonance was not observed.

	Chemical Shift (ppm)			
Residue	**NH**	**αH**	**βH**	**Other H**
Gly1	n.o.	3.87		
Leu2	8.27	4.15	1.39	γH - 1.30
				δH - 0.77, 0.84
Phe3	8.27	4.72	3.27, 3.01	δH - 7.31
isoD4	8.27	4.58	2.91, 2.85	n.o.
Val5	8.99	3.75	2.19	γH - 0.982, 1.10
Ile6	8.48	3.73	2.04	γH - 1.68, 1.28, 0.93
				δH - 0.88
Lys7	8.22	4.08	1.89	γH - 1.56, 1.45
				δH - 1.73
				ϵH - 3.02
Lys8	7.98	4.21	2.02, 1.91	γH - 1.52
				δH - 1.72
				ϵH - 2.96
Val9	8.49	3.66	2.24	γH - 1.07, 0.94
Ala10	8.69	4.03	1.52	
Ser11	7.98	4.32	4.08	
Val12	7.88	4.03	2.27	γH - 1.11, 0.99
Ile13	8.11	3.95	2.01	γH - 1.65, 1.29, 0.96
				δH - 0.84
Gly14	8.29	3.96		
Gly15	7.94	4.14, 3.93		
Leu16	7.83	4.28	1.89	γH - 1.61
				δH - 0.93
				$CONH_2$ - 7.37,7.12

6.3.1.2 Secondary Shifts

$\Delta\delta$s describe the chemical shift variations for the amino acids in the peptide from their corresponding random coil values (as determined in water [435]). The $\Delta\delta$s of the α^1H were smoothed over $\pm$ 2 residues, while the N^{1}H $\Delta\delta$s are plotted unsmoothed. All $\Delta\delta$s were plotted against the amino acid sequence to allow extraction of structural information (Figure 4.12). As discussed in Section 2.3.7.1, negative $\Delta\delta$ values indicate an upfield shift from random coil values, whilst positive values indicate a downfield shift.

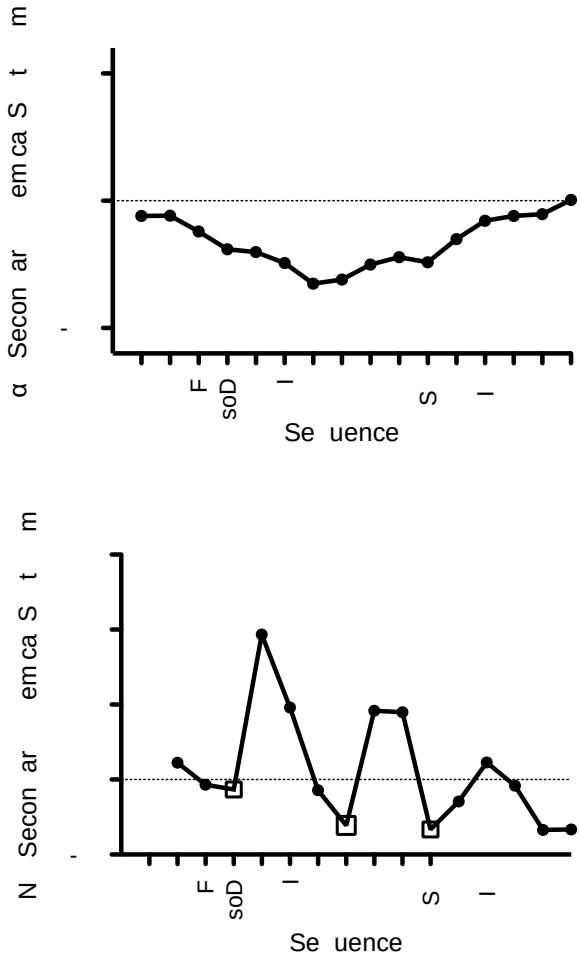


Figure 6.2: (A) α^1H secondary shifts of isoAsp4-citropin 1.1 in DPC micelles, smoothed over n ± 2 residues. (B) N^{1}H secondary shifts of isoAsp4-citropin 1.1. Negative values indicate an upfield shift from random coil values, whilst positive values indicate a downfield shift. Hydrophilic residues are indicated as open symbols.

The α^1H resonances (Figure 6.2A) display a general upfield shift from random coil values as indicated by the negative $\Delta\delta$s. This is especially true towards the centre of the sequence, with residues at the termini having $\Delta\delta$s closer to 0, indicating chemical shifts similar to those expected in a random coil structure. This indicates that the peptide is adopting a helical structure towards the centre of the peptide, with the termini being more flexible [184, 186].

The N^{1}H $\Delta\delta$s (Figure 6.2B) display a degree of periodicity across the length of the peptide, with the three hydrophilic residues having a more upfield chemical shift from random coil values when compared with hydrophobic residues. This suggests a degree of amphipathicity [184, 186] is present in the helix. Additionally, at the N-terminus the

N^1H $\Delta\delta$ values are close to zero, indicating a lack of well-defined secondary structure in this region. Together, these $\Delta\delta$ values lend support to the notion that there is a degree of helical structure present in the peptide, but this appears to be less pronounced at the termini.

6.3.1.3 NOE Connectivities

Diagnostic NOE connectivities used in structure calculations for isoAsp4-citropin 1.1 in DPC micelles are displayed in Figure 6.3. Sequential d_{NN}, $d_{\alpha N}$ and $d_{\beta N}$ NOEs are observed along most of the sequence, but are notably absent at the termini. Additionally, medium range NOEs between residues that are three or four apart are observed. Notably, $d_{\alpha N(i,i+3)}$ and $d_{\alpha N(i,i+4)}$ connectivities are not present at the N-terminal end of the peptide, but are observed throughout the remainder of the sequence.

Taken together these data suggest that isoAsp4-citropin 1.1 adopts a predominantly α-helical structure in DPC micelles. The presence of additional medium-range connectivites at the C-terminus suggests that the helical structure is more well defined in this region and that at the N-terminus a more flexible structure may be present, this is especially true due to the lack of medium range NOEs observed between the first three residues of the peptide.

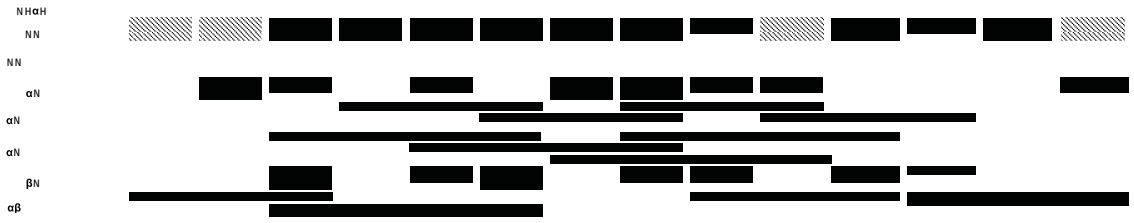


Figure 6.3: Diagnostic NOEs used in structure calculations for isoAsp4-citropin 1.1. The thickness of the bar indicates the relative strength of the NOE. $^3J_{NH\alpha H}$ values (Hz) are indicated where applicable. # indicates that signal overlap precluded assignment of the coupling constant. * indicates the coupling constant was not detected.

6.3.1.4 Structure Calculations

The NOESY spectrum for isoAsp4-citropin 1.1 was assigned and the volume of each peak converted to distance restraints using the procedures outlined by Xu *et al.* [436]. A total of 332 distance restaints were generated, which are summarised in Table 6.3. Only a small number of these distance restraints are ambiguous.

Table 6.3: Distance restraints obtained from the NOESY spectrum of isoAsp4-citropin 1.1 in DPC micelles, and utilised in structural calculations.

	Number of Restraints
Sequential NOEs	80
Medium-range NOEs	75
Long-range NOEs	0
Intra-residue NOEs	165
Ambiguous NOEs	12
Total	**332**

The amide region of a high-resolution 1D ^{1}H spectrum was used in an attempt to measure $^{3}J_{NH\alpha H}$ coupling constants (Figure 6.3). There was significant overlap in this region and many peaks were unresolved. Consequently, many coupling constants could not be measured. Attempts were also made to determine coupling constants from COSY spectra [189], but line widths were too broad to determine these values accurately. Only the $^{3}J_{NH\alpha H}$ coupling constant for the final amino acid residue (Leu16) could be measured accurately (Figure 6.3), nevertheless this indicates a degree of helical structure is present in this region (as the value is less than 6 Hz [160]).

The RMD and SA protocol of ARIA generated sixty final structures, of which the twenty with the lowest potential energy were selected for analysis. Figure 6.4 shows these twenty lowest energy structures of isoAsp4-citropin 1.1 superimposed over the well defined residues. This structure calculation result confirms the conclusions drawn from analysis of the NOE and secondary shift data from the NMR. It can be seen that isoAsp4-citropin 1.1 adopts a helical structure in the central portion of the peptide, however, it appears to be more flexible at the termini, especially the N-terminus which is closest to the isoAsp residue.

Figure 6.4: The 20 lowest energy structures of isoAsp4-citropin 1.1 in DPC micelles superimposed over the backbone atoms of well-defined residues.

The 20 lowest energy structures show considerable convergence over the entire sequence, but the RMSD values (Table 6.4) are higher than longer, previously reported peptide structures. This indicates that this structure exhibits a degree of conformational flexibility, consistent with its relatively short sequence which contains two Gly residues at the C-terminus and the unusual isoAsp residue at the N-terminus. However, the RMSD values of the well-defined residues compare favourably with those obtained for citropin 1.1 [402]. Additionally, the final structures had no violations from idealised covalent geometries, supporting the quality of the obtained structures.

Table 6.4: Structural statistics for isoAsp4-citropin 1.1 following RMD/SA calculations.

Energy (kcal.mol^{-1})	
E_{total}	21.43 ± 4.67
E_{bond}	0.56 ± 0.27
E_{angle}	8.36 ± 2.28
$E_{improper}$	2.18 ± 1.66
E_{VDW}	9.42 ± 1.74
E_{NOE}	0.91 ± 0.59
E_{cdih}	0.00
Well-defined residues	6-12
RMSD from mean geometry	**(Å)**
Backbone atoms of well-defined residues	0.55 ± 0.33
Heavy atoms of well-defined residues	1.29 ± 0.66
All backbone atoms	2.44 ± 0.67
All heavy atoms	2.72 ± 0.63

Analysis of the AOPs of the twenty lowest energy structures demonstrated that seven of the residues were well-defined (Table 6.4). A Ramachandran plot showing the average ϕ and ψ angles for isoAsp4-citropin 1.1 shows that most of the residues fall within the favoured region for α-helical structure (Figure 6.5), whilst residues at the N-terminus near the isoAsp lie outside these regions, consistent with the conclusion that isoAsp perturbs the local structure of the peptide in this region.

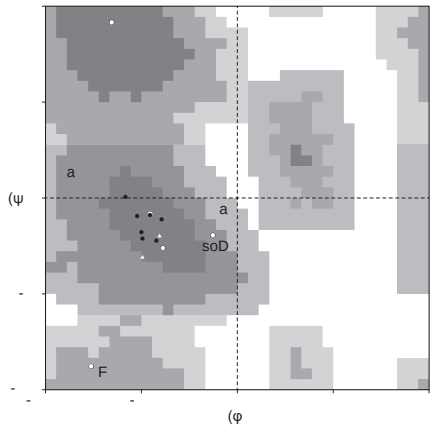


Figure 6.5: Ramachandran plot for isoAsp4-citropin 1.1 in DPC micelles. Well-defined residues are indicated by filled symbols. Favourable regions are labelled A and B for α helical and β strand structures, respectively.

The most stable calculated structure of isoAsp4-citropin 1.1 in DPC micelles is shown in Figure 6.6. The amphipathic helical structure along the length of the peptide is clearly perturbed at the N-terminal end by the presence of isoAsp.

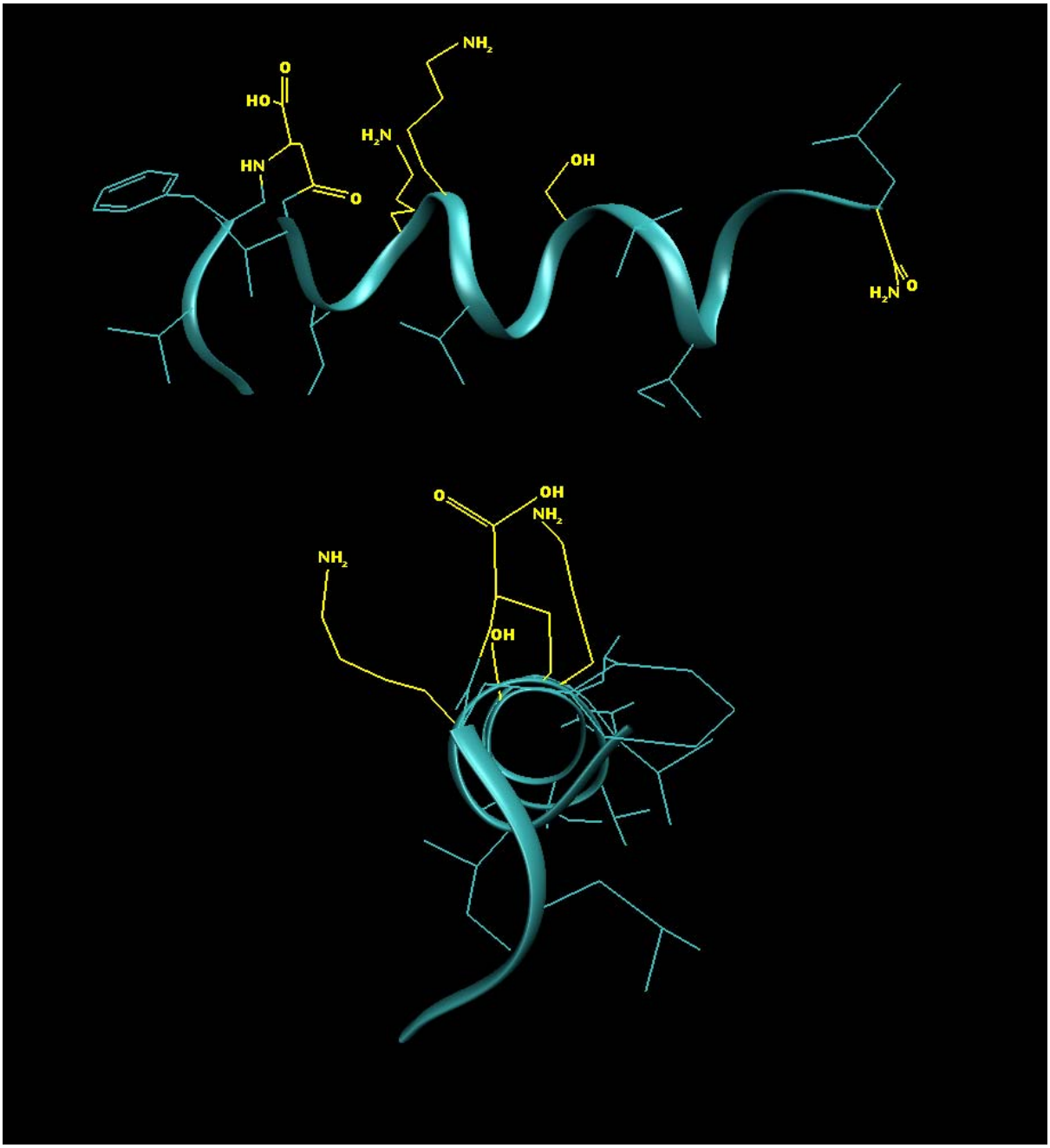


Figure 6.6: Side and axial views of the lowest calculated potential energy structure of isoAsp4-citropin 1.1 in DPC micelles. The backbone of the isoAsp residue is shown as lines. Hydrophobic groups are shown in green, hydrophilic groups are shown in yellow.

6.3.2 Antibiotic Activity

Table 6.5 summarises the antibiotic activities of citropin 1.1 and isoAsp4-citropin 1.1 against several Gram positive and Gram negative organisms. These data indicate that citropin 1.1 shows activity against Gram positive organisms, whilst isoAsp4-citropin 1.1, shows no activity below 100 μg/mL against any of the test pathogens.

Table 6.5: Antibiotic activities of citropin 1.1 and isoAsp4-citropin 1.1 against selected pathogens (Values are given as minimum inhibitory concentrations (MIC, μg/mL)).

Organism	**Citropin 1.1**	**isoAsp4-Citropin 1.1**
Mycobacterium bovis	> 100	> 100
Staphylococcus aureus	12.5	> 100
Methicillin resistant S. aureus	12.5	> 100
Bacillus subtilis	12.5	> 100
Pseudomonas aeruginosa	> 100	> 100

6.3.3 Asp Isomerisation in Other Amphibian Peptides

Several other amphibian peptides have been studied, to examine impact of Asp isomerisation on their activity. However, these have all been smooth muscle contractants, which are unstructured random coils in both aqueous solution and solvents which model a membrane environment (as determined by NMR spectroscopy, data not shown). Figure 6.7 shows that Asp isomerisation has a modest detrimental effect on the activity of *Crinia* angiotensin II, no effect on the activity of physalemin, and a concentration dependent effect on the activity of uperin 1.1 [624]. Together with the data presented here for citropin 1.1 (Section 6.3.2), it can be said that the effect of Asp isomerisation is unpredictable, and dependent on the particular system under investigation.

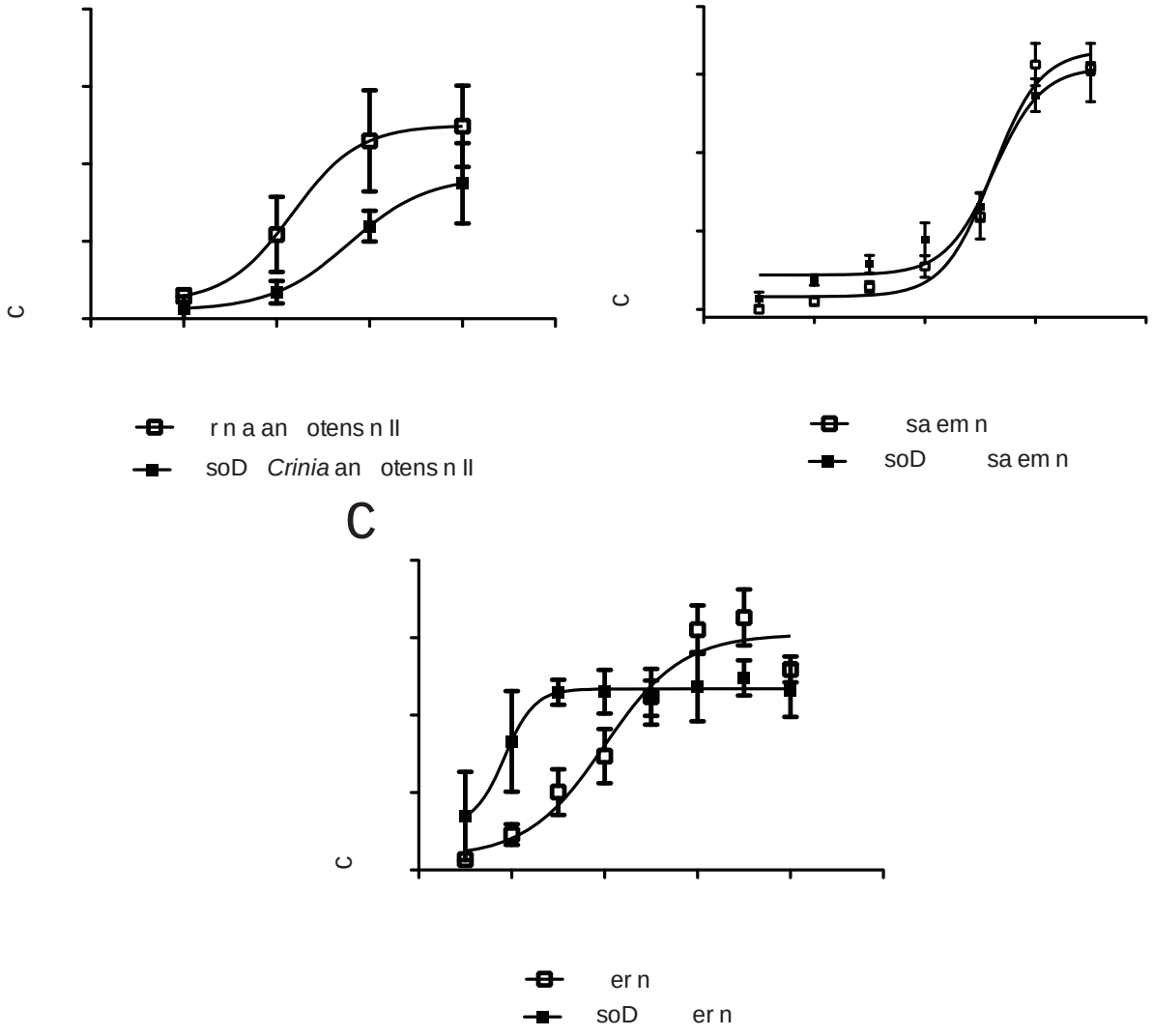


Figure 6.7: Smooth muscle contraction response curves of (A) *Crinia* angiotensin II, (B) physalaemin, (C) uperin 1.1 (Table 6.1) and their corresponding isoAsp isomers [624].

6.3.4 Effect of isoAsp on Proteolysis

The susceptibility of four bioactive amphibian peptides and their isoAsp analogues (citropin 1.1, uperin 1.1, physalaemin, *Crinia* angiotensin II) to proteolysis by Asp-N endoprotease, trypsin and α-chymotrypsin was investigated. Figure 6.8 displays the mass spectra of the Asp-N digested peptides; it is evident that cleavage is always observed at Asp, but not isoAsp.

In contrast, trypsin cleaves on the C-terminal side of Lys and Arg (Figure 6.9) in all cases to give mass spectra which are qualitatively similar for the isomeric peptides studied. Uperin 1.1 was not studied as it does not contain the required motif for protease recognition and cleavage.

α-Chymotrypsin, which cleaves the peptide bond on the C-terminal side of Tyr, Phe, Trp, and sometimes Leu, gives the same cleavage products for the Asp/isoAsp isomers of *Crinia* angiotensin II, physalaemin and uperin 1.1 (Figure 6.10). However, the data are markedly different for the two citropin 1.1 isomers (Figure 6.10). α-Chymotrypsin cleaves C-terminal of Phe3 of citropin 1.1 (adjacent to Asp4), but does not cleave next to Phe3 of isoAsp4 citropin 1.1.

These data are summarised in Table 6.6.

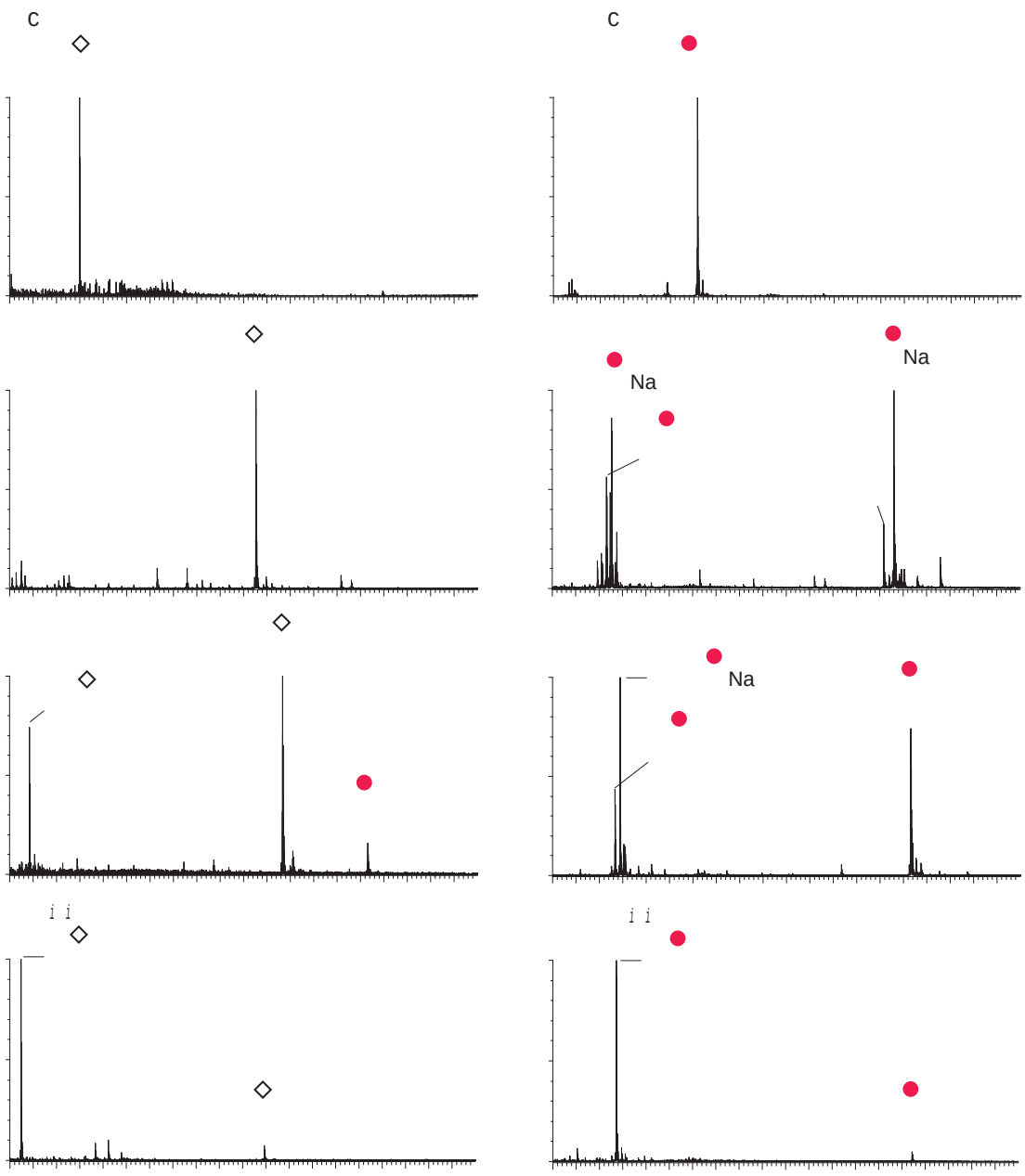


Figure 6.8: Representative mass spectra showing the proteolysis products of selected amphibian peptides and their isoAsp isomers when treated with endoprotease AspN for 2 hours. Identical spectra were obtained when the isoAsp isomers were treated overnight. Intact peptide (●) and proteolysis products are indicated (◇).

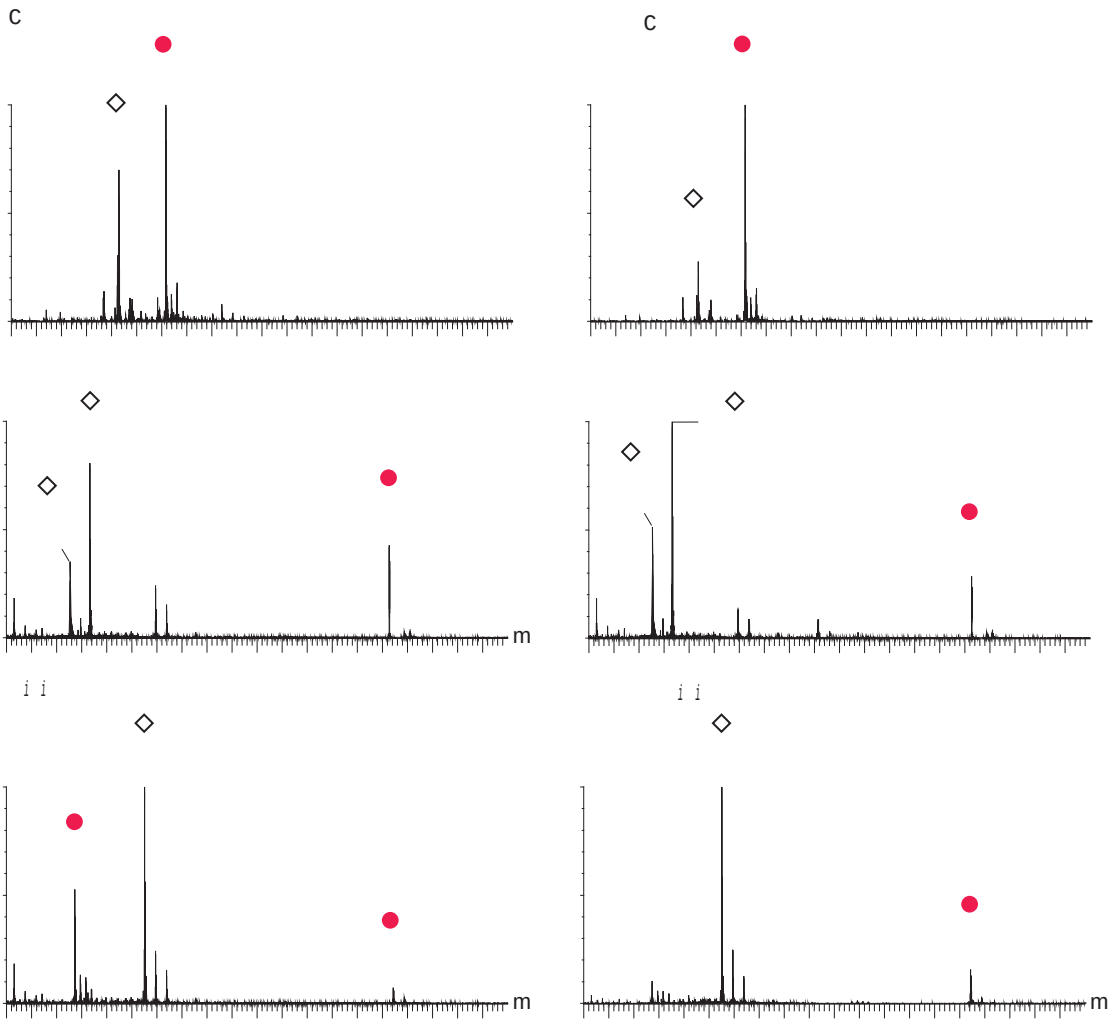


Figure 6.9: Representative mass spectra showing the proteolysis products of selected amphibian peptides and their isoAsp isomers when treated with trypsin for 2 hours. Intact peptide (●) and proteolysis products are indicated (◇).

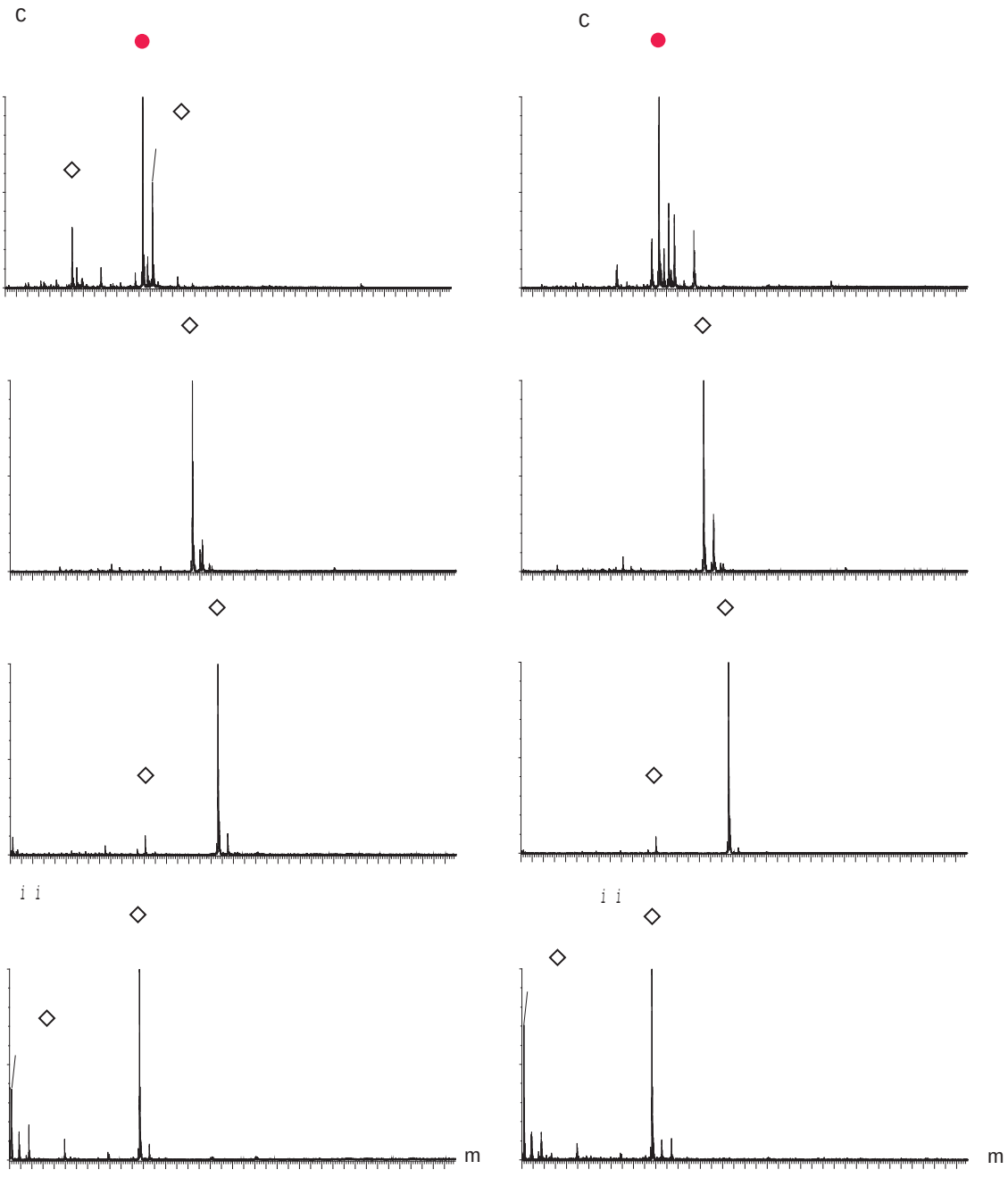


Figure 6.10: Representative mass spectra showing the proteolysis products of selected amphibian peptides and their isoAsp isomers when treated with α-chymotrypsin for 2 hours. Intact peptide (•) and proteolysis products are indicated (◇).

Table 6.6: Summary of proteolytic digestion products observed when isomeric peptides are treated with AspN, trypsin or α-chymotrypsin. * indicates the peptide is not cleaved by the protease. - indicates the peptide does not contain the required amino acid to be cleaved by the protease.

	Protease		
Precursor	**Asp N**	**Trypsin**	**Chymotrypsin**
	***Crinia* angiotensin II -** APGDRIYVHPF(OH)		
Asp	DRIYHVPF(OH)	IYHVPF(OH)	APGDRIY(OH)
		APGDR(OH)	VHPF(OH)
isoAsp	*	IYHVPF(OH)	APGisoDRIY(OH)
		APGisoDR(OH)	VHPF(OH)
	Uperin 1.1 - pEADPNAFYGLM(NH_2)		
Asp	DPNAFYGLM(OH)	-	pEADPNAFY(OH)
isoAsp	*	-	pEAisoDPNAFY(OH)
	Physalaemin - pEADPNKFYGLM(NH_2)		
Asp	DPNKFYGLM(NH_2)	pEADPNK(OH)	pEADPNKFY(OH)
		FYGLM(NH_2)	pEADPNKF(OH)
isoAsp	*	pEAisoDPNK(OH)	pEAisoDPNKFY(OH)
		FYGLM(NH_2)	pEAisoDPNKF(OH)
	Citropin 1.1 - GLFDVIKKVASVIGGL(NH_2)		
Asp	DVIKKVASVIGGL(NH_2)	GLFDVIKK(OH)	DVIKKVASVIGGL(NH_2)
		VASVIGGL(NH_2)	
isoAsp	*	GLFDVIKK(OH)	*
		VASVIGGL(NH_2)	

6.4 Discussion

6.4.1 Structure and Activity Changes in Amphibian Peptides Due To Asp Isomerisation

The 3D structure of citropin 1.1 in the membrane mimicking solvent 50 % (v/v) TFE/H_2O has been previously determined (Figure 6.11) [402]. It can be seen that the peptide displays amphipathic α-helical character along the entirety of the peptide. In contrast, the structure of isoAsp-citropin 1.1 (Figure 6.6), shows significant disruption to the helical structure around the isoAsp residue at the N-terminal end of the peptide in DPC micelles. It should be noted that the NH chemical shifts of Leu2, Phe3 and isoAsp4 in isoAsp-citropin 1.1 are identical which may prevent the detection of NOEs. However, this is unlikely to significantly influence the structures obtained, which demonstrate a significant structural perturbation upon Asp isomerisation. Other studies have been performed examining the structures of helical peptides in both 50 % TFE and DPC micelles, which show a similarity in the structures adopted in the two membrane mimicking solvent systems [437, 452].

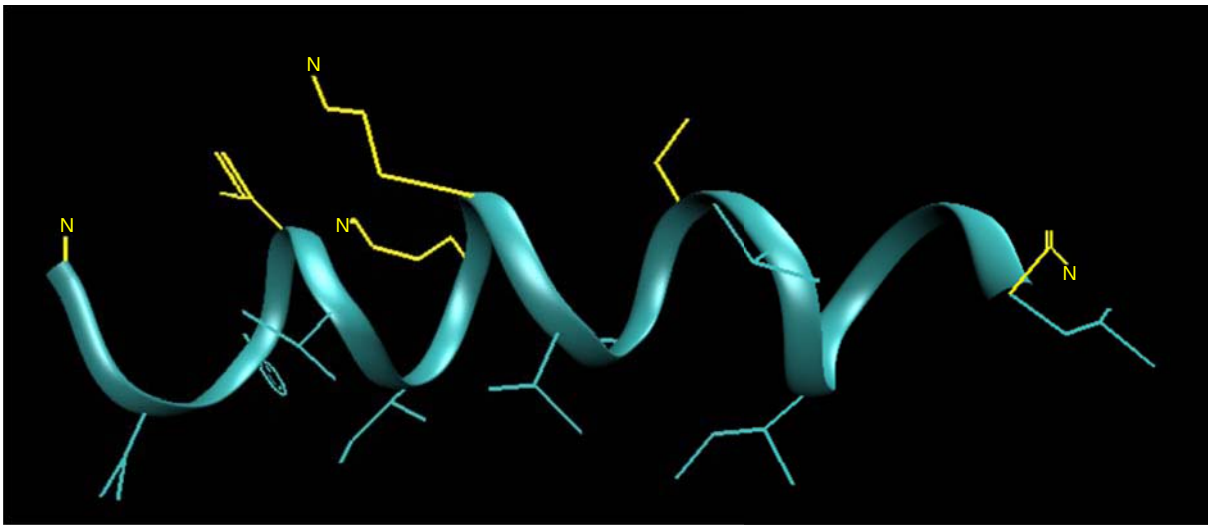


Figure 6.11: Lowest calculated potential energy structure of citropin 1.1 in TFE/H_2O (50 % v/v). Hydrophobic groups are shown in green, hydrophilic groups are shown in yellow [402].

In sodium dodecyl sulfate (SDS) micelles, citropin 1.1 exists in a different conformation, comprising two helices separated by a βIV turn at residues 8 and 9 [685]. However, this may be due to the physical properties of the SDS micelle, as SDS is a negatively charged

detergent whilst DPC is neutral. Several studies have reported differences in conformation between peptides bound to anionic and zwitterionic detergent micelles [686–689].

The antibiotic activity of citropin 1.1 and many synthetic modifications of the peptide have been studied extensively [402, 417, 685]. Like other peptides isolated from the amphibians of the genus *Litoria*, citropin 1.1 is a wide-spectrum antibiotic that is highly active against Gram positive pathogens, less so against Gram negative organisms. However, adding Lys residues to the peptide has been shown to increase its antibiotic activity against Gram negative organisms. Additionally the non-conserved replacement of Asp4 with Ala increases activity, presumably due to the increased overall positive charge of the peptide. It has also been shown that removal of the first three N-terminal residues completely eliminates the antimicrobial activity, and truncations like these have been found in amphibian skin secretions, suggesting that removal of these residues is necessary to inactivate the peptide in a biological setting.

The structural modification studied here is much more severe than any of those previously investigated, due to the significant alteration in the peptide backbone as a result of the presence of a β peptide linkage and the consequent additional rotatable bond that is present. It is apparent that this causes significant structural perturbation, as evidenced by the overlaid structures in Figure 6.12. The structure and activity alterations suggests that the isoAsp is perturbing the amphipathic α-helical structure that existed along the entirety of the peptide in the native, unmodified sequence, and as a consequence of this loss of helicity in this region, the peptide loses its antimicrobial activity. This has an effect similar to removing the N-terminal residues from the peptide sequence.

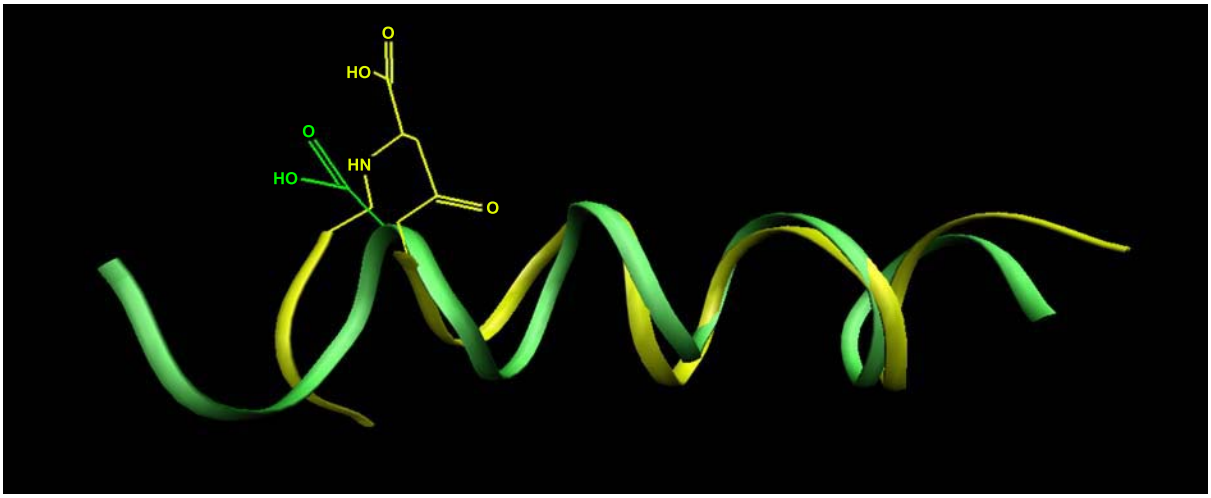


Figure 6.12: Overlaid lowest potential energy structures of citropin 1.1 (green) in TFE/H_2O (50 % v/v) [402] and isoAsp4-citropin 1.1 (yellow). The Asp/isoAsp residues are shown. The backbone around isoAsp is shown as lines for clarity.

The combination of the data from citropin 1.1 and findings from the other studied peptides (Section 6.3.3) demonstrate that the Asp isomerisation process, generally associated with peptide/protein ageing, does not necessarily result in loss of function, and in some instances it may result in an alteration of function. Most notably in the case of uperin 1.1, Asp isomerisation results in increased activity at low concentrations, and decreased activity at higher concentrations (relative to the native uperin 1.1).

6.4.2 Proteolysis Studies

The results presented in Figure 6.8 show that the AspN endoprotease always cleaves adjacent to Asp, but when isoAsp is present, this cleavage reaction does not occur. A similar effect has been observed in previous studies, and has been used as a method to identify isoAsp residues in complex mixtures, such as proteolytic digests [642, 666, 667, 673, 690–692].

The results presented for peptides treated with trypsin show a markedly different effect (Figure 6.9). Trypsin cleaves at the C-terminal side of Lys and Arg, to give mass spectra which are qualitatively similar when comparing results for the isomeric peptides studied (*Crinia* angiotensin II, physalaemin and citropin 1.1 and their isoAsp isomers). Uperin 1.1 does not contain the prototypical residues required for trypsin cleavage (i.e. Lys or Arg), so was not investigated. Again, this has been observed in other studies, demonstrating that

trypsin digestion adjacent to isoAsp residues cannot be used as a method to determine the presence of these isomeric residues with β-peptide linkages [642, 666, 667, 673, 690–692].

α-Chymotrypsin behaves the same with the isomeric pairs of *Crinia* angiotensin II, physalaemin and uperin 1.1, cleaving adjacent to either Phe or Tyr residues (Figure 6.10). In contrast, α-chymotrypsin behaves differently with the isomeric citropins 1.1, in that it cleaves at the C-terminal end of Phe3 next to Asp, but not at all for Phe3 next to isoAsp. This is an unequivocal result indicating that proteolysis of this particular backbone amide group at the C-terminal end of Phe3 (and the N-terminal end of isoAsp) is stopped by the presence of the adjacent isoAsp residue.

6.4.3 Implications of Asp Isomerisation on Peptide/Protein Structure

The secondary structure motif in which an Asp or Asn residue is located often determines the degree of isoAsp formation that occurs [632–636], and often when Asp or Asn is contained in a region of ordered secondary structure, there is decreased susceptibility to isoAsp formation. The peptide system studied here, citropin 1.1, displays amphipathic α-helical character in a membrane environment, but in aqueous solution it is extremely flexible [547], like many peptide antibiotics of this type [384, 693, 694]. This flexible structure in solution renders many peptides extremely susceptible to Asp isomerisation *in vivo* or in a pharmaceutical formulations. For use in a clinical setting, it is imperative that modification of this type in all potential medicines are prevented to ensure unwanted and potentially hazardous side-effects are minimised [695]. It is thought that this may be achieved by optimising formulation conditions, with several strategies proposed, including controlling temperature and pH. It has also been reported that it might be possible to increase the stability of these unstructured peptides by the presence of additives in formulations that increase local secondary structure content (reviewed in [682]).

In addition, the identity of the subsequent residue plays an important role in determining the occurrence of this unnatural amino acid residue [626–630]. Consequently, strategies to improve the resistance of potential peptide therapeutics to isoAsp formation may involve site directed mutagenesis to alter the succeeding amino acid residue to slow the reaction (reviewed in [682]). In the case of citropin 1.1, the bulky hydrophobic Val residue after Asp

prevents isomerisation from occurring naturally, this allows the two isomers to be studied individually, with interconversion not possible, and consequently, Asp isomerisation need not be an issue of major concern.

Several other studies have investigated the effect of isoAsp on regular secondary structural motifs in peptides and proteins. An investigation into the structural effects of deamidation on an analogue of human growth hormone releasing factor, Leu27 hGRF(1-32)NH_2 showed substitution of Asn8 for Asp8 resulted in minimal structural change, however, the presence of isoAsp8 resulted in disruption to the helical nature of the peptide, as determined by a lack of NOE connectvities in the surrounding region [696]. A similar disruption to local structure was also observed in bovine calbindin D_{9k} when isoAsp was present in place of Asp/Asn [679]. Asn deamidation in the light chain complementarity determining region 1 (CDR1) of a humanized IgG1 monoclonal antibody has also been shown to trigger an overall alteration in structure by CD [697]. Finally, the crystal structures of several proteins have been determined which contain isoAsp in loop regions. These studies demonstrate that isoAsp has a variety of effects, including a disruption of the local hydrogen bond network, an overall decrease in the rigidity of the loop and its reorientation with respect to the well-defined structural elements of the protein [698–700]. The results contained herein for isoAsp4-citropin 1.1, are consistent with these previous studies.

6.5 Experimental Procedures

6.5.1 Materials

Peptides were synthesised by GenScript (Piscataway, NJ, USA) using L-amino acids and standard solid phase methods. The peptides are of > 95 % purity, as shown by MS and HPLC. DPC and D_2O were purchased from Novachem (Melbourne, VIC, Australia). Asp-N and α-chymotrypsin were purchased from Sigma Aldrich (Sydney, NSW, Australia) and trypsin was purchased from Promega (Madison, WI, USA).

6.5.2 Sample Preparation for NMR Spectroscopy

Samples were prepared by dissolving isoAsp4-citropin 1.1 (5.7 mg, 3.55 μmol) and DPC (54.8 mg, 140.6 μmol) in an aqueous solution of 10 % D_2O buffered to pH 6 with NaH_2PO_4 (50 mM), giving a final peptide concentration of 5.05 mM, and a final DPC concentration of 0.2 M, in a volume of 0.7 mL. This concentration of lipid was selected as it greatly exceeds the CMC and it produces a solution of approximately one peptide molecule per micelle [222, 224].

6.5.3 NMR Spectroscopy

NMR spectra were recorded using a Varian Inova-600 NMR spectrometer, with a ^1H frequency of 600 MHz and a ^{13}C frequency of 150 MHz. Experiments were carried out at 25 °C and referenced to water (4.82 ppm). The water signal in both the TOCSY and NOESY experiments was suppressed by presaturation. This was achieved by application of a low power RF pulse at the frequency of the water resonance during the 1.1 s relaxation delay between scans. The water signal in the DQF-COSY spectra was suppressed using gradient methods [459].

TOCSY, DQF-COSY and NOESY experiments were acquired in the phase-sensitive mode, with time proportional phase incrementation in t_1. In each experiment, 8 to 16 time-averaged scans were acquired per increment, with a total of 200 t_1 increments for

TOCSY and NOESY spectra and 512 t_1 increments in the DQF-COSY spectrum. The FID in t_2 consisted of 2878 data points over a spectral width of 9592.3 Hz. NOESY spectra were acquired with a mixing time of 150 ms and the TOCSY pulse sequence included a spin-lock of 80 ms. HSQC experiments were recorded for the assignment of α^{13}C resonances using correlations with the attached α^{1}H. These experiments consisted of 8 time-averaged scans over a total of 128 t_1 increments. In the ^{1}H dimension (F_2), 2878 data points were collected over a spectral width of 9592.3 Hz.

The 2D spectra were processed using VNMR software (VNMRJ, Version 2.1, Revision B). The data matrices were multiplied by a Gaussian function in both dimensions then zero-filled before Fourier transformation. The final processed 2D NMR matrices consisted of 4096 × 4096 real data points.

6.5.4 Structure Calculations

The ^{1}H resonances in the NOESY, DQF-COSY and TOCSY spectra, were assigned using Sparky software (version 3.111) by means of the sequential assignment procedure [160]. The volumes of the larger of the two symmetrical cross-peaks were converted to distance restraints by the methods of Nilges *et al.* [190]. The $^{3}J_{NH\alpha H}$ coupling constants were measured from the high-resolution 1D spectra obtained. From this analysis, dihedral angles were restrained in accordance with the guidelines outlined in Section 2.3.8.3.

Structures were generated from random starting conformations, using the standard RMD and SA protocol of ARIA (Version 1.2) [461], implemented with CNS (Version 1.1) [462]. To account for the unusual amino acid residue isoAsp, structural parameters were altered to define the atypical primary structure of the residue and the dihedral and improper angles. Each ARIA run consisted of 8 iterations of structure calculations. The standard ARIA parameters were utilised, however, better convergence was achieved using parameters based on those used by Pari *et al.* [463] and Kang *et al.* [464]. The final iteration of calculations generated 60 structures from which the 20 lowest calculated potential energy structures were selected for analysis. The programs VMD (Version 1.8.2) [465] and MOLMOL (Version 2k.2) [466] were used to visualise the 3D structures.

6.5.5 Smooth Muscle Contraction

Experiments were conducted in collaboration with Dr. Ian Musgrave and Katarina Markulic, on guinea pig ileum using procedures reported in full previously [701]. Briefly, segments of ileum were suspended in 20 mL organ baths containing physiological salt solution at 310 K and were gassed with 95 % O_2 and 5 % CO_2. Segments were connected to a tissue holder and to an isometric force-displacement transducer. Tension was recorded using Powerlab v 5.0. Segments were washed by replacing the physiological salt solution twice, and were then allowed to equilibrate for a period of 30 min under 2 g of resting tension. When a stable baseline was achieved, the guinea pig ileum was pre-treated with the nitric oxide inhibitor L-NNA to inhibit underlying nitric oxide mediated contraction [702, 703]. Cumulative concentration response curves were then obtained for *Crinia* angiotensin II, isoAsp4-*Crinia* angiotensin II, uperin 1.1, isoAsp3-uperin 1.1, physalaemin, isoAsp3-physalaemin and uperolein. Responses were normalised as a percentage of the contraction produced by acetylcholine at 10^{-6} M and are shown as mean ± SEM of three independent experiments done in duplicate.

6.5.6 Antibiotic Activity Testing

Antibiotic testing was conducted by Prof. Lixin Zhang (Institute of Microbiology, Chinese Academy of Sciences). MIC (μg mL^{-1}) activities of citropin 1.1 and isoAsp4-citropin 1.1 were measured against *Mycobacterium bovis* BCG, *Staphylococcus aureus*, *methicillin resistant S. aureus*, *Bacillus subtilis* and *Pseudomonas aeruginosa* by a standard procedure [704]. Isonazid, vancomycin and ciprofioxacin were used as positive controls as described previously [705].

6.5.7 Proteolysis

To each peptide (1 nmol) in 50 mM NH_4HCO_3, the protease (1:50 peptide:protease ratio, by mass) [Asp-N, trypsin or α-chymotrypsin, as appropriate] was added and the samples incubated at 37 °C. Aliquots were taken after 5 min, at 15 min intervals for 3 hours and after 18 hr.

Nano-ESI mass spectra were acquired using a Micromass Q-ToF 2 (Waters/Micromass, Manchester, UK) orthogonal acceleration time-of-flight mass spectrometer. Samples were introduced into the spectrometer using platinum-coated borosilicate capillaries prepared in-house. Conditions were typically as follows: capillary voltage 1.4 kV, source temperature 80 °C, and cone voltage 30 V.

Chapter 7

Summary

7.1 Negative Ion MS Amenable Cross-Linking Reagents

This work details the first use of negative ion MS to identify and provide information about cross-linked peptides and proteins. It has been demonstrated that the negative ion cleavages of adducts containing disulfide cross-linking reagents can be useful in identifying intermolecular, intramolecular and dead-end cross-linked peptides amongst complex mixtures and can readily localise the cross-linking site with residue level resolution. The spectra obtained are diagnostic for the cross-link type and allow for sequencing and identification of the cross-linked species.

The experimental data presented demonstrates the utility of the disulfide containing CX-MS reagents DSA and DSP. Symmetrical cleavage of the DSA linker between the sulfur atoms yields two characteristic fragment ions that can be used to easily identify cross-linked peptides by low energy collisional activation of the precursor ion. Low energy activation of DSP cross-linked peptides yields four ions corresponding to cleavage at and around the disulfide bond. These fragmentation pathways permit peptides cross-linked by these reagents to be identified with relative ease, and each contributing peptide sequenced independently by MS^3. Consequently, these reagents are ideal candidates for this negative ion, MS identifiable CX strategy. Conversely, the longest cross-linker studied, DSB, does not yield facile fragmentation products, and higher collision energies are required to initiate cross-linker fragmentations, both between and adjacent to the disulfide atoms, together with those of the peptide backbone, making this reagent unsuitable for this approach.

The facile nature of these cross-linker fragmentation processes means they also compete with other low energy processes that are unique to negative ion MS. Of the studied side chain induced processes, it has been shown that loss of H_2O by Asp and in some instances, side chain induced backbone cleavage by Asp significantly compete with fragmentations of the cross-linkers. These competitive processes can be used as a diagnostic indicator for the presence of Asp within a cross-linked peptide, and could be used to aid identification of the peptide involved when studying a protein of unknown structure. Additionally the loss of CH_2O from the Ser side chain also competes with DSP fragmentation, however the relative abundance of these peaks are significantly lower than those due to loss of H_2O from Asp.

Negative ion MS provides a clear advantage over conventional positive ion MS for the identification of these disulfide containing cross-links due to their facile, characteristic fragmentation behaviour. These are the most facile negative ion cleavages to have been studied to date. Now that the characteristic fragmentations of the modified peptides are understood, it is possible to use this negative ion approach to develop methods for high-throughput analysis of unknown protein complex structures, so that identification of cross-linked peptides and sequencing can occur simultaneously.

7.2 Amphibian Peptides That Inhibit Neuronal Nitric Oxide Synthase

Most amphibians of the genera *Litoria* and *Crinia* that have been studied thus far produce one or more peptides that prevent NO formation by nNOS as part of their glandular secretion. Regulation of nNOS is achieved by the protein CaM, which is activated by binding Ca^{2+} ions resulting in a conformational change necessary for it to upregulate NO production. It has been well-documented that these bioactive amphibian peptides bind to Ca_4^{2+}CaM inducing a further conformational change which prevents the protein from activating nNOS. In this work, the Ca_4^{2+}CaM complexes that are formed with the amphibian peptides caerin 1.8.11, dahlein 5.6 and citropin 1.1 (mod. 13) have been comprehensively studied using a variety of biophysical techniques.

ITC analysis of the complexation process has shown that the peptides have a very high affinity for Ca_4^{2+}CaM, with K_d values in the nM range. Repetition of the ITC experiments at several temperatures, and analysis of the change in ΔH as a function of temperature

allowed ΔC_p values to be determined. These demonstrated that the binding process results in an overall reorganisation of the protein structure, and that this process results in a large decrease in solvent accessible surface area, indicating the formation of a canonical collapsed conformation.

IM-MS studies have confirmed that a structural rearrangement takes place upon CaM binding Ca^{2+}. It was demonstrated that binding of two Ca^{2+} ions to each lobe of CaM is necessary to result in a significant global structural reorganisation, leading to a more compact structure. Additionally, the complexes that are formed when Ca_4^{2+}CaM binds the amphibian peptides show generally smaller CCSs than the extended complex that forms with the C20W peptide. These CCSs compare favourably with those calculated theoretically from model compact complex structures.

NMR analysis of the truncated caerin 1.8.11 peptide has shown that it is capable of forming an amphipathic α-helix, with hydrophobic residues at positions in the sequence which are known to affect Ca_4^{2+}CaM binding. It has been confirmed using CD that this helical structure is adopted by caerin 1.8.11 and the other amphibian peptides when they bind Ca_4^{2+}CaM.

The Ca_4^{2+}CaM:caerin 1.8.11 complex was further studied using NMR spectroscopy, with large CSPs observed across the whole protein sequence. However, CSI analysis showed that the α-helical secondary structural elements of Ca_4^{2+}CaM remain relatively intact. This is again consistent with an overall reorganisation of the protein structure upon binding. Notably, CSPs were clustered around the hydrophobic domains of Ca_4^{2+}CaM, which are known to be involved in target binding. Finally, the many Met residues in this region exhibited significant CSPs, and it is well-known that these play an integral role in stabilising CaM complexes.

The combination of low resolution structural information presented in this chapter, gathered using a variety of biophysical techniques, provides significant insight into the structures of both CaM and its complexes with several amphibian peptides. This evidence indicates that binding of the peptides results in a significant structural reorganisation of Ca_4^{2+}CaM, with the amphibian peptides engulfed in a channel between the N- and C-terminal domains. These data, along with additional structural information (possibly 3D

structure data), may further aid in the development of more efficacious peptide inhibitors of NO production.

7.3 Amphibian Peptides That Inhibit Fibril Formation and Self-Assemble

It has been demonstrated that caerin 1.8 and several of its truncations have the ability to prevent fibril formation by the Alzheimer's disease related protein, Aβ(1-42). Caerin 1.8 redirects the aggregation of Aβ(1-42), leading to the formation of amorphous aggregates. There are several key features of caerin 1.8 that are shared with rationally designed peptide based inhibitors of Aβ fibrillogenesis, including a structurally similar recognition motif and several charged residues.

Many AMPs have the inherent ability to form amyloid structures. In this work, the amyloidogenicity of uperin 3.5 was investigated, with an emphasis on comparing its ability to form amyloid fibrils with other disease related proteins and peptides. It is apparent that uperin 3.5 forms amyloid-like fibrils with rapid kinetics, as the process lacks a lag phase. The resultant fibrils are of defined morphology as studied by TEM and AFM, with typical coiled-coil ultrastructure. CD has also shown that the fibrils are β-sheet rich, similar to other amyloids. The self-assembly was monitored using IM-MS, with data supporting an ordered, fibrillar type assembly, as opposed to an isotropic process.

Fibril formation in the presence of the membrane mimicking co-solvent TFE was enhanced when small quantities were added, but inhibited at higher amounts. This is likely due to the flexible α-helical structure induced at low TFE concentrations which can readily convert to β-sheet rich amyloid. However, at higher concentrations the rigid well-defined secondary structures are unable to convert. Additionally, the effects of TFE on hydrophobic interactions and hydrogen bonding may also have influenced the kinetics of fibril formation. The presence of negatively charged SUVs also enhanced fibril formation. This again correlated with an increase in α-helical content. In the future, this phenomenon should be further explored, along with the effect of both native and fibrillar uperin 3.5 on

bacterial and mammalian membranes. This will afford significant insight into uperin 3.5's specific membrane mediated mechanism of toxicity.

Fibrillar species of uperin 3.5 were shown to be cytotoxic to cultured PC12 cells, in a similar fashion to other disease related amyloids. This supports the notion that the generic amyloid structure is toxic, and that the effect is sequence independent. Furthermore, EGCG, a flavonoid from green tea, was shown to inhibit fibril formation, remodel mature fibrils and prevent the cytotoxicity associated with uperin 3.5 fibril formation.

These data provide significant evidence for the structural similarities between amyloid-like fibrils formed by uperin 3.5 and those of disease-related proteins. Additionally, they provide further support for the notion that fibrils formed by AMPs posses properties similar to those of disease related amyloidogenic proteins. This work adds to the growing body of literature that suggests mechanistic investigations of AMPs which form amyloid structures can provide significant insight into the means by which disease related amyloids elicit their sometimes devastating effects.

7.4 Aspartic Acid Isomerisation in Amphibian Peptides

Proteins and peptides undergo many spontaneous and non-spontaneous modifications, including many that are a consequence of ageing. The formation of isoAsp from both native Asp and Asn residues is one of these ageing related processes. In this work, it has been demonstrated that isoAsp formation severely impacts on the structure and function of citropin 1.1, and has moderate effects on the activities of the other amphibian peptides studied.

In the peptide systems studied, there is no clear trend that changing Asp to isoAsp will alter the activity of the peptide in a particular direction (i.e. either enhance or reduce the activity). This is evidenced by: (i) replacing Asp4 of citropin 1.1 with isoAsp destroys the antimicrobial activity against the test pathogens *S. aureus* and *B. subtilis*, (ii) the smooth muscle contraction by isoAsp4-*Crinia* angiotensin II is less than that of the Asp isomer, and (iii) the smooth muscle activity of isoAsp3-uperin 1.1 differs dramatically from uperin 1.1 (at low concentrations the isoAsp isomer is the more active, but at higher concentrations

the reverse is true). In essence, these results are in accord with those for other isoAsp systems already studied, in that replacement of Asp with isoAsp can change the activity of that system, but the way the activity is changed depends entirely on the particular system under investigation.

Replacing Asp with isoAsp in these small peptides does not change the extent of degradation of the isomeric peptides by the proteolytic enzyme trypsin. Additionally, as previously observed, cleavage by Asp-N is suppressed completely by the presence of isoAsp. However, α-chymotrypsin does not cleave next to Phe3 of isoAsp4-citropin 1.1. This observation needs further investigation for cognate peptides, but is beyond the scope of the present study.

The structural alteration in citropin 1.1 as a result of Asp isomerisation is quite severe, with a kink induced in the helix at the N-terminus, when studied in DPC micelles. This perturbation of the helical structure can be clearly seen when looking at an axial view of the peptide structure. The loss of helicity in this region can account for the disruption of antibiotic activity.

Structural modifications of this type in peptide antibiotics are particularly relevant. This is because in solution peptide therapeutics mainly adopt flexible conformations, making them highly susceptible to isoAsp formation given the appropriate primary structure. However, they may adopt a more ordered structure *in vivo* to perform their biological function, like membrane active AMPs. Consequently, it is of fundamental importance to minimise any spontaneous non-enzymatic post translational modifications (like isoAsp formation) in pharmaceuticals for human use, as these alterations in primary structure may have significant, potentially unexpected, influences on their structure and activity.

7.5 Conclusion

The work presented in this thesis demonstrates the utility of a vast array of well-established and emerging biophysical techniques in structural biology. Low resolution studies, which combine structural data from a variety of techniques in an integrative approach, are becoming commonplace to characterise many biochemical process. The application and continued refinement of existing methods, along with the development of novel analytical techniques, will prove essential in our quest to further characterise the basis of complex biological processes at the molecular level.

Bibliography

[1] F. H. C. Crick. On protein synthesis. *Symp. Soc. Exp. Biol.*, 12:138–163, 1958.

[2] F. H. C. Crick. Central dogma of molecular biology. *Nature*, 227:561–563, 1970.

[3] W. J. Wedemeyer and H. A. Scheraga. *Protein folding: Overview of pathways*, chapter in Encyclopedia of life sciences (www. els. net). Nature Publishing Group, London, 2001.

[4] U. Langel, B. F. Cravatt, A. Graslund, G. von Heijne, T. Land, S. Niessen, and M. Zorko. *Introduction to Peptides and Proteins*. Taylor and Francis, Boca Raton, 2010.

[5] C. Branden and J. Tooze. *Introduction to Protein Structure*. Garland Publishing, New York, USA, 1991.

[6] D. L. Nelson and M. M. Cox. *Lehninger Principles of Biochemistry*. W. H. Freeman and Co., New York, 5th edition, 2008.

[7] A. J. Cozzone. *Protein: Fundamental chemical properties*, chapter in Encyclopedia of life sciences (www. els. net). Nature Publishing Group, London, 2002.

[8] F. Chiti and C. M. Dobson. Protein misfolding, functional amyloid and human disease. *Annu. Rev. Biochem.*, 75:333–366, 2006.

[9] C. M. Dobson. Principles of protein folding, misfolding and aggregation. *Semin. Cell Dev. Biol.*, 15:3–16, 2004.

[10] C. M. Dobson. The structural basis of protein folding and its links with human disease. *Phil. Trans. R. Soc. Lond. B*, 356:133–145, 2001.

[11] C. V. Robinson, A. Sali, and W. Baumeister. The molecular sociology of the cell. *Nature*, 450: 973–982, 2007.

[12] K. Teilum, J. G. Olsen, and B. B. Kragelund. Protein stability, flexibility and function. *Biochim. Biophys. Acta*, 1814:969–976, 2011.

[13] C. B. Anfinsen. Principles that govern the folding of protein chains. *Science*, 181:223–230, 1973.

[14] F. U Hartl and M. Hayer-Hartl. Converging concepts of protein folding *in vitro* and *in vivo*. *Nat. Struct. Mol. Biol.*, 16:574–581, 2009.

[15] T. R. Jahn and S. E. Radford. The yin and yang of protein folding. *FEBS J.*, 272:5962–5970, 2005.

[16] S. E. Jackson. How do small single-domain proteins fold? *Fold. Des.*, 3:R81–R91, 1998.

[17] D. J. Brockwell and S. E. Radford. Intermediates: Ubiquitous species on folding energy landscapes? *Curr. Opin. Struct. Biol.*, 17:30–37, 2007.

[18] H. Ecroyd and J. A. Carver. Unravelling the mysteries of protein folding and misfolding. *IUBMB Life*, 60:769–774, 2008.

[19] F. Ulrich Hartl, A. Bracher, and M. Hayer-Hartl. Molecular chaperones in protein folding and proteostasis. *Nature*, 475:324–332, 2011.

[20] R. N. Rambaran and L. C. Serpell. Amyloid fibrils: Abnormal protein assembly. *Prion*, 2: 112–117, 2008.

[21] J. Tyedmers, A. Mogk, and B. Bukau. Cellular strategies for controlling protein aggregation. *Nat. Rev. Mol. Cell. Biol.*, 11:777–788, 2010.

[22] S. Walter and J. Buchner. Molecular chaperones - cellular machines for protein folding. *Angew. Chem. Int. Ed.*, 41:1098–1113, 2002.

[23] S. Charbonnier, O. Gallego, and A. C. Gavin. The social network of a cell: Recent advances in interactome mapping. *Biotechnol. Annu. Rev.*, 14:1–28, 2008.

[24] S. Bader, S. Kuhner, and A. C. Gavin. Interaction networks for systems biology. *FEBS Lett.*, 582:1220–1224, 2008.

[25] D. Brown and G. Superti-Furga. Rediscovering the sweet spot in drug discovery. *Drug Discov. Today*, 8:1067–1077, 2003.

[26] H. Ruffner, A. Bauer, and T. Bouwmeester. Human protein-protein interaction networks and the value for drug discovery. *Drug Discov. Today*, 12:709–716, 2007.

[27] T. Pawson, M. Raina, and P. Nash. Interaction domains: From simple binding events to complex cellular behavior. *FEBS Lett.*, 513:2–10, 2002.

[28] T. L. Pukala. Mass spectrometry for structural biology: Determining the composition and architecture of protein complexes. *Aust. J. Chem.*, 64:681–691, 2011.

[29] S. E. A. Ozbabacan, H. B. Engin, A. Gursoy, and O. Keskin. Transient protein-protein interactions. *Protein Eng. Des. Sel.*, 24:635–648, 2011.

[30] J. P. Glusker. *Structural Crystallography in Chemistry and Biology*, volume 4 of *Benchmark papers in physical chemistry and chemical physics*. Hutchinson Ross Pub. Co., Stroudsburg, 1981.

[31] J. Cavanagh, W. J. Fairbrother, A. G. Palmer, N. J. Skelton, and M. Rance. *Protein NMR Spectroscopy: Principles and Practice*. Academic Press, Boston, 2007.

[32] S. M. Kelly, T. J. Jess, and N. D. Price. How to study proteins by circular dichroism. *Biochim. Biophys. Acta*, 1715:119–139, 2005.

[33] W. Baumeister and A. C. Steven. Macromolecular electron microscopy in the era of structural genomics. *Trends Biochem. Sci.*, 25:624–631, 2000.

[34] J. Lipfert and S. Doniach. Small-angle X-ray scattering from RNA, proteins, and protein complexes. *Annu. Rev. Biophys. Biomol. Struct.*, 36:307–327, 2007.

[35] T. L. Pukala, B. T. Ruotolo, M. Zhou, A. Politis, R. Stefanescu, J. A. Leary, and C. V. Robinson. Subunit architecture of multiprotein assemblies determined using restraints from gas-phase measurements. *Structure*, 17:1235–1243, 2009.

[36] T. Wiseman, S. Williston, J. F. Brandts, and L.-N. Lin. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.*, 179: 131–137, 1989.

[37] F. Stengel, R. Aebersold, and C. V. Robinson. Joining forces: Integrating proteomics and crosslinking with the mass spectrometry of intact complexes. *Mol. Cell. Proteomics*, doi: 10. 1074/mcp. R111. 014027, 2011.

[38] J. J. Thomson. *Rays of Positive Electricity and Their Application to Chemical Analyses*. Longmans, Green and Co., London, 1913.

[39] R. Aebersold and D. R. Goodlett. Mass spectrometry in protemics. *Chem. Rev.*, 101:269–295, 2001.

[40] V. H. Wysocki, K. A. Resing, Q. Zhang, and G. Chen. Mass spectrometry of peptides and proteins. *Methods*, 35:211–222, 2005.

[41] C. T. Mant and R. S. Hodges. *High Performance Liquid Chromatography of Peptides and Proteins. Separation, Analysis and Conformation*. CRC Press, Boca Raton, 1991.

[42] C. Uetrecht, R. J. Rose, E. van Duijin, K. Lorenzen, and A. J. R. Heck. Ion mobility mass spectrometry of proteins and protein assemblies. *Chem. Soc. Rev.*, 39:1633–1655, 2010.

[43] B. T. Ruotolo, J. L. P. Benesch, A. M. Sandercock, S.-J. Hyung, and C. V. Robinson. Ion mobility-mass spectrometry analysis of large protein complexes. *Nat. Protocols*, 3: 1139–1151, 2008.

[44] G. E. Verbeck, B. T. Ruotolo, H. A. Sawyer, K. J. Gillig, and D. H. Russell. A fundamental introduction to ion mobility mass spectrometry applied to the analysis of biomolecules. *J. Biomol. Tech.*, 13:56–61, 2002.

[45] A. Patriksson, E. Marklund, and D. van der Spoel. Protein structures under electrospray conditions. *Biochem.*, 46:933–945, 2007.

[46] J. L. P. Benesch and C. V. Robinson. Dehydrated but unharmed. *Nature*, 462:576–577, 2009.

[47] A. B. Kanu, P. Dwivedi, M. Tam, L. Matz, and H. H. Hill Jr. Ion mobility-mass spectrometry. *J. Mass Spectrom.*, 43:1–22, 2008.

[48] S. D. Pringle, K. Giles, J. L. Wildgoose, J. P. Williams, S. E. Slade, K. Thalassinos, R. H. Bateman, M. T. Bowers, and J. H Scrivens. An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-Tof instrument. *Int. J. Mass Spectrom.*, 261:1–12, 2007.

[49] M. Barber, R. S. Bordoli, G. J. Elliot, R. D. Sedgwick, and A. N. Tyler. Fast atom bombardment of solids (FAB): A new ion source for mass spectrometry. *J. Chem. Soc. Chem. Commun.*, 7: 325–327, 1981.

[50] M. Barber, R. S. Bordoli, G. J. Elliot, R. D. Sedgwick, and A. N. Tyler. Fast atom bombardment mass spectrometry. *Anal. Chem.*, 54:645A–657A, 1982.

[51] M. Karas and F. Hillenkamp. Laser desorption ionization of proteins with molecular masses exceeding 10 000 daltons. *Anal. Chem.*, 60:2299A–2301A, 1988.

[52] S. D. Fuerstenau, W. H. Benner, J. J. Thomas, C. Brugidou, B. Bothner, and G. Siuzdak. Mass spectrometry of an intact virus. *Angew. Chem. Int. Ed.*, 40:541–544, 2001.

[53] M. Dole, L. L. Mach, R. L. Hines, R. C. Mobley, L. D. Ferguson, and M. B. Alice. Molecular beams of macroions. *J. Chem. Phys.*, 49:2210–2247, 1968.

[54] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246:64–71, 1989.

[55] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse. Electrospray ionization - principles and practice. *Mass Spectrom. Rev.*, 9:37–70, 1990.

[56] M. Yamashita and J. B. Fenn. Electrospray ion source. Another variation on the free-jet theme. *J. Phys. Chem.*, 88:4451–4459, 1984.

[57] M. Yamashita and J. B. Fenn. Negative ion production with the electrospray ion source. *J. Phys. Chem.*, 88:4671–4675, 1984.

[58] C. M. Whitehouse, R. N. Dreyer, M. Yamashita, and J. B. Fenn. Electrospray interface for liquid chromatographs and mass spectrometers. *Anal. Chem.*, 57:675–679, 1985.

[59] J. B. Fenn. Electrospray wings for molecular elephants (Nobel lecture). *Angew. Chem. Int. Ed.*, 42:3871–3894, 2003.

[60] J. A. Loo. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.*, 16:1–23, 1997.

[61] B. N. Pramanik, P. L. Bartner, U. A. Mirza, Y Liu, and A. K. Ganguly. Electrospray ionization mass spectrometry for the study of non-covalent complexes: An emerging technology. *J. Mass Spectrom.*, 33:911–920, 1998.

[62] H. Hernandez and C. V. Robinson. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nature Protocols*, 2:715–726, 2007.

[63] A. J. R. Heck. Native mass spectrometry: A bridge between interactomics and structural biology. *Nat. Methods*, 5:927–933, 2008.

[64] J. B. Fenn, C. K. Meng, and M. Mann. Multiply charged ions and a method for determining the molecular weight of large molecules, especially biopolymers, by electrospray mass spectrometry. USA Patent No. 9014148, 1990.

[65] A. G. Ferrige, M. J. Seddon, B. N. Green, S. A. Jarvis, J. Skilling, and J. Staunton. Disentangling electrospray spectra with maximum entropy. *Rapid Commun. Mass Spectrom.*, 6:707–711, 1992.

[66] M. Mann, C. K. Meng, and J. B. Fenn. Interpreting mass spectra of multiply charged ions. *Anal. Chem.*, 61:1702–1708, 1989.

[67] W. J. Griffiths, A. P. Jonsson, S. Liu, D. K. Rai, and Y. Wang. Electrospray and tandem mass spectrometry in biochemistry. *Biochem. J.*, 355:545–561, 2001.

[68] C. S. Lane. Mass spectrometry-based proteomics in the life sciences. *Cell. Mol. Life Sci.*, 62: 848–869, 2005.

[69] G. I. Taylor. The stability of horizontal fluid interface in a vertical electric field. *J. Fluid Mech.*, 2:1–15, 1965.

[70] M. Sharon and C. V. Robinson. The role of mass spectrometry in the structure elucidation of dynamic protein complexes. *Annu. Rev. Biochem.*, 76:167–193, 2007.

[71] D. Williams and I. Fleming. *Spectroscopic Methods in Organic Chemistry*. McGraw Hill, Berkshire, England, 5th edition, 1995.

[72] P. Kebarle and U. H. Verkerk. Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.*, 28:898–917, 2009.

[73] J. W. S Rayleigh. On the equilibrium of liquid conducting masses charged with electricity. *Philos. Mag.*, 44:184–186, 1882.

[74] P. Kebarle. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *J. Mass Spectrom.*, 35:804–817, 2000.

[75] P. Kebarle and Y. Ho. On the mechanism of electrospray mass spectrometry. In R. B. Cole, editor, *Electrospray Ionization Mass Spectrometry*, pages 3–63. John Wiley and Sons, New York, 1997.

[76] D. C. Taflin, T. L. Ward, and E. J. Davis. Electrified droplet fission and the Rayleigh limit. *Langmuir*, 5:376 – 384, 1988.

[77] J. Fernandez de la Mora. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta*, 406:93–104, 2000.

[78] J. V. Iribarne and B. A Thomson. On the evaporation of small ions from charged droplets. *J. Chem. Phys.*, 64:2287–2294, 1976.

[79] J. L. P. Benesch, B. T. Ruotolo, D. A. Simmons, and C. V. Robinson. Protein complexes in the gas phase: Technology for structural genomics and proteomics. *Chem. Rev.*, 107:3544–3567, 2007.

[80] M. Gamero-Castaño and J. Fernández de la Mora. Mechanisms of electrospray ionization of singly and multiply charged salt clusters. *Anal. Chim. Acta*, 406:67–91, 2000.

[81] A. E. Ashcroft. Recent developments in electrospray ionisation mass spectrometry: Noncovalently bound protein complexes. *Nat. Prod. Rep.*, 22:452–464, 2005.

[82] N. B. Cech and C. G. Enke. Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrom. Rev.*, 20:362–387, 2001.

[83] T. T. Herskovitis, B. Gadegbeku, and H. Jaillet. On the structural stability and solvent denaturation of proteins. 1. Denaturation by the alcohols and glycols. *J. Biol. Chem.*, 245: 2588–2598, 1970.

[84] K. A. Dill and D. Shortle. Denatured states of proteins. *Annu. Rev. Biochem.*, 60:795–825, 1991.

[85] A. L. Fink and B. Painter. Characterization of the unfolding of ribonuclease A in aqueous methanol solvents. *Biochem.*, 26:1665–1671, 1987.

[86] K. R. Babu and D. J. Douglas. Methanol-induced conformations of myoglobin at pH 4. 0. *Biochem.*, 39:14702–14710, 2000.

[87] R. D. Smith, J. A. Loo, R. R. Ogorzalek Loo, M. Busman, and H. R. Udseth. Principles and practice of electrospray ionization - mass spectrometry for large polypeptides and proteins. *Mass Spectrom. Rev.*, 10:359–451, 1991.

[88] J. C. Y. Le Blanc, J. Wang, R. Guevremont, and K. W. M. Siu. Electrospray mass spectra of protein cations formed in basic solutions. *Org. Mass Spectrom.*, 29:587–593, 1994.

[89] N. Felitsyn, M. Peschke, and P. Kebarle. Origin and number of charges observed on multiply-protonated native proteins produced by ESI. *Int. J. Mass Spectrom.*, 219:39–62, 2002.

[90] M. S. Wilm and M. Mann. Electrospray and Taylor-Cone theory theory, Dole's beam of macromolecules at last? *Int. J. Mass Spectrom. Ion Processes*, 136:167–180, 1994.

[91] M. Wilm and M. Mann. Analytical properties of the nanoelectrospray ion source. *Anal. Chem.*, 68:1–8, 1996.

[92] E. W. Chung, D. A. Henriques, D. Renzoni, C. J. Morton, T. D. Mulhern, M. C. Pitkeathly, J. E. Ladbury, and C. V. Robinson. Probing the nature of interactions in SH2 binding interfaces - evidence from electrospray ionization mass spectrometry. *Protein Sci.*, 8:1962–1970, 1999.

[93] R. Juraschek, T. Dulcks, and M. Karas. Nanoelectrospray - more than just a minimized-flow electrospray ionization source. *J. Am. Soc. Mass Spectrom.*, 10:300–308, 1999.

[94] D. M. Williams and T. L. Pukala. Novel insights into protein misfolding diseases revealed by ion mobility-mass spectrometry. *Mass Spectrom. Rev.*, in press:DOI: 10.1002/mas.21358, 2013.

[95] L. A. Woods, S. E. Radford, and A. E. Ashcroft. Advances in ion mobility spectrometry - mass spectrometry reveal key insights into amyloid assembly. *Biochim. Biophys. Acta*, in press:DOI: 10.1016/j.bbapap.2012.10.002, 2013.

[96] P. H. Dawson. *Quadrupole Mass Spectrometry and its Applications*. Elsevier, Amsterdam, 1976.

[97] W. Paul. Electromagnetic traps for charged and neutral particles. *Angew. Chem. Int. Ed.*, 29: 739–748, 1990.

[98] A. G. Marshall, C. L. Hendrickson, and G. S. Jackson. Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.*, 17:1–35, 1998.

[99] Q. Hu, R. J. Noll, H. Li, A. Makarov, M. Hardman, and R. G. Cooks. The Orbitrap: A new mass spectrometer. *J. Mass Spectrom.*, 40:430–443, 2005.

[100] R. J. Cotter. Time-of-flight mass spectrometry. In R. J. Cotter, editor, *Time-of-Flight Mass Spectrometry*, pages 16–48. American Chemical Society, Washington, 1994.

[101] G. R. Hilton and J. L. P. Benesch. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *J. R. Soc. Interface*, 9: 801–816, 2012.

[102] C. V. Robinson. Mass spectrometry characterization of multiprotein complexes. *Cold Spring Harb. Protoc.*, doi:10.1101/pdb.prot5180, 2009.

[103] W. Paul and H. S. Steinwedel. Ein neues massenspektrometer ohne magnetfeld. *Z. Naturforsch*, 69:448, 1957.

[104] M. E. Bier. Coupling ESI and MALDI sources to the quadrupole mass filter, quadrupole ion trap, linear quadrupole ion trap, and Orbitrap mass analyzers. In R. B. Cole, editor, *Electrospray and MALDI Mass Spectrometry*, pages 265–344. John Wiley and Sons, Hoboken, New Jersey, 2010.

[105] C. Steel and M. Henchman. Understanding the quadrupole mass filter through computer simulation. *J. Chem. Ed.*, 75:1049–1054, 1998.

[106] A. K. Shukla and J. H. Futrell. Collisional activation and dissociation of polyatomic ions. *Mass Spectrom. Rev.*, 12:211–255, 1993.

[107] A. K. Shukla and J. H. Futrell. Tandem mass spectrometry: Dissociation of ions by collisional activation. *J. Mass Spectrom.*, 35:1069–1090, 2000.

[108] R. Haynes and M. L. Gross. Collision-induced dissociation. *Methods Enzymol.*, 193:237–263, 1990.

[109] R. G. Cooks. Collision-induced dissociation: Readings and commentary. *J. Mass Spectrom.*, 30:1215–1221, 2005.

[110] E. de Hoffmann. Tandem mass spectrometry: A primer. *J. Mass Spectrom.*, 31:129–137, 1996.

[111] W. E. Stephens. A pulsed mass spectrometer with time dispersion. *Phys. Rev.*, 69:691, 1946.

[112] B. A. Mamyrin, V. I. Krataev, D. V. Shmikk, and V. A. Zagulin. The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Sov. Phys. JETP*, 37: 45–48, 1973.

[113] B. A. Mamyrin. Laser assisted reflectron time-of-flight mass spectrometry. *Int. J. Mass Spectrom. Ion Processes*, 131:1–19, 1994.

[114] D. M Murphy and K. Mauersberger. Operation of a microchannel plate counting system in a mass spectrometer. *Rev. Sci. Instrum.*, 56:220–226, 1985.

[115] J. L. Wiza. Microchannel plate detectors. *Nucl. Instrum. Methods*, 162:587–601, 1979.

[116] I. V. Chernushevich, A. V. Loboda, and B. A. Thompson. An introduction to quadrupole-time-of-flight mass spectrometry. *J. Mass Spectrom.*, 36:849–865, 2001.

[117] B. C. Bohrer, S. I. Merenbloom, S. L. Koeniger, A. E. Hilderbrand, and D. E. Clemmer. Biomolecule analysis by ion mobility spectrometry. *Annu. Rev. Anal. Chem.*, 1:10.1–10.35, 2008.

[118] K. Giles, S. D. Pringle, K. R. Worthington, D. Little, J. L. Wildgoose, and R. H. Bateman. Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun. Mass Spectrom.*, 18:2401–2414, 2004.

[119] D. E. Clemmer. *Cross section database, URL: http://www. indiana. edu/ clemmer/.*

[120] M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz, and M. F. Jarrold. Structural information from ion mobility measurements: Effects of the long-range potential. *J. Phys. Chem.*, 100:16082–16086, 1996.

[121] A. A. Shvartsburg and M. F. Jarrold. An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chem. Phys. Lett.*, 261:86–91, 1996.

[122] C. A. Scarff, K. Thalassinos, G. R. Hilton, and J. H Scrivens. Travelling wave ion mobility mass spectrometry studies of protein structure: Biological significance and comparison with X-ray crystallography and nuclear magnetic resonance spectroscopy measurements. *Rapid Commun. Mass Spectrom.*, 22:3297–3304, 2008.

[123] M. F. Jarrold. Unfolding, rehydration, and hydration of proteins in the gas phase. *Acc. Chem. Res.*, 32:360–367, 1999.

[124] K. Biemann and S. A. Martin. Mass spectrometric determination of the amino acid sequence of peptides and proteins. *Mass Spectrom. Rev.*, 6:1–76, 1987.

[125] D. H. Williams, C. V. Bradley, S. Santikarn, and G. Bojesen. Fast-atom-bombardment mass spectrometry. A new technique for the determination of molecular weights and amino acid sequences of peptides. *Biochem. J.*, 201:105–117, 1982.

[126] D. R. Mueller, M. Eckersley, and W. J. Richter. Hydrogen transfer reactions in the formation of 'Y+2' sequence ions from protonated peptides. *Org. Mass Spectrom.*, 23:217–222, 1988.

[127] B. Paizs and S. Sándor. Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.*, 24:508–548, 2005.

[128] R. A. J. O'Hair. The role of nucleophile-electrophile interactions in the unimolecular and bimolecular gas phase ion chemistry of peptides and related systems. *J. Mass Spectrom.*, 35: 1377–1381, 2000.

[129] A. Schlosser and W. D. Lehmann. Five-membered ring formation in unimolecular reactions of peptides: A key structural element controlling low-energy collision-induced dissociation of peptides. *J. Mass Spectrom.*, 35:1382–1390, 2000.

[130] M. J. Polce, D. Ren, and C. Wesdemiotis. Dissociation of the peptide bond in protonated peptides. *J. Mass Spectrom.*, 35:1391–1398, 2000.

[131] V. H. Wysocki, G. Tsaprailis, L. L. Smith, and L. A. Breci. Mobile and localized protons: A framework for understanding peptide dissociation. *J. Mass Spectrom.*, 35:1399–1406, 2000.

[132] A. G. Harrison, I. G. Csizmadia, and T. H. Tang. Structure and fragmentation of b2 ions in peptide mass spectra. *J. Am. Soc. Mass Spectrom.*, 11:427–436, 2000.

[133] J. M. Farrugia, R. A. J. O'Hair, and G. E. Reid. Do all b2 ions have oxazolone structures? Multistage mass spectrometry and *ab initio* studies on protonated N-acyl amino acid methyl ester model systems. *Int. J. Mass Spectrom.*, 210-211:71–81, 2001.

[134] J. M. Farrugia, T. Tarerner, and R. A. J. O'Hair. Side-chain involvement in the fragmentation reactions of the protonated methyl esters of histidine and its peptides. *Int. J. Mass Spectrom.*, 209:99–112, 2001.

[135] P. Edman and G. Begg. A protein sequenator. *Eur. J. Biochem.*, 1:80–91, 1967.

[136] J. H. Bowie, C. S. Brinkworth, and S. Dua. Collision-induced fragmentations of the $(M-H)^-$ parent anions of underivatized peptides: An aid to structure determination and some unusual negative ion cleavages. *Mass Spectrom. Rev.*, 21:87–107, 2002.

[137] G. A. Chass, C. N. J. Marai, D. H. Setiada, I. G. Csizmadia, and A. G. Harrison. A Hartree-Fock, MP2 and DFT computational study of the structures and energies of b2 ions derived from deprotonated peptides. A comparison of method and basis set used on relative product stabilities. *Theochem.*, 675:149–162, 2004.

[138] C. S Brinkworth, S. Dua, and J. H. Bowie. Backbone cleavages of $[M-H]^-$ anions of peptides. Cyclisation of citropin 1 peptides involving reactions between the C-terminal $[CONH]^-$ residue and backbone amide carbonyl groups. A new type of b cleavage: A joint experimental and theoretical study. *Rapid Commun. Mass Spectrom.*, 16:713–721, 2002.

[139] T. Wang, T. T. N. Tran, A. N. Calabrese, and J. H. Bowie. Backbone fragmentation of $[M-H]^-$ anions from peptides. Reinvestigation of the mechanism of the beta prime cleavage. *Rapid Commun. Mass Spectrom.*, 26:1832–1840, 2012.

[140] D. Bilusich and J. H. Bowie. Fragmentations of $[M-H]^-$ ions of underivatised peptides. Part 2. Characteristic cleavages of Ser and Cys and disulfides and other post translational modifications, together with some unusual rearrangements. *Mass Spectrom. Rev.*, 28:20, 2009.

[141] S. T. Steinborner and J. H. Bowie. The negative ion mass spectra of $[M-H]^-$ ions derived from caeridin and dynastin peptides. Internal backbone cleavages directed through Asp and Asn residues. *Rapid Commun. Mass Spectrom.*, 11:253–258, 1997.

[142] C. S Brinkworth, S. Dua, A. M. McAnoy, and J. H. Bowie. Negative ion fragmentations of deprotonated peptides: Backbone cleavages directed through both Asp and Glu. *Rapid Commun. Mass Spectrom.*, 15:1965–1973, 2001.

[143] P. A. Wabnitz, J. H. Bowie, J. C. Wallace, and M. J. Tyler. Peptides from the skin glands of the Australian buzzing tree frog *Litoria electrica*. Comparison with the skin peptides of the red tree frog *Litoria rubella*. *Aust. J. Chem.*, 52:639–645, 1999.

[144] C. S Brinkworth, S. Dua, and J. H. Bowie. Backbone cleavages of $[M\text{-}H]^-$ anions of peptides. New backbone cleavages following cyclisation reactions of a C-terminal $[CONH]^-$ group with Ser residues and the use of the γ backbone cleavage initiated by Gln to differentiate between Lys and Gln residues. *Eur. J. Mass Spectrom.*, 8:53–66, 2002.

[145] M. Saunders, A. Wishnia, and J. G. Kirkwood. The nuclear magnetic resonance spectrum of ribonuclease. *J. Am. Chem. Soc.*, 70:3289–3290, 1957.

[146] F. Bloch. Dynamical theory of nuclear induction. *Phys. Rev.*, 102:104–135, 1956.

[147] E. M. Purcell, H. C. Torrey, and R. V. Pound. Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.*, 69:37–38, 1946.

[148] M. P. Williamson, T. F. Havel, and K. Wüthrich. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ^{1}H nuclear magnetic resonance and distance geometry. *J. Mol. Biol.*, 182:295–315, 1985.

[149] A. Bax. Two-dimensional NMR and protein structure. *Annu. Rev. Biochem.*, 58:223–256, 1989.

[150] A. E. Ferentz and G. Wagner. NMR spectroscopy: A multifaceted approach to macromolecular structure. *Quart. Rev. Biophys.*, 33:29–65, 2000.

[151] M. Bieri, A. H. Kwan, M. Mobli, G. F. King, J. P. Mackay, and P. R. Gooley. Macromolecular NMR spectroscopy for the non-spectroscopist: Beyond macromolecular solution structure determination. *FEBS J.*, 278:704–715, 2011.

[152] A. H. Kwan, M. Mobli, P. R. Gooley, G. F. King, and J. P. Mackay. Macromolecular NMR spectroscopy for the non-spectroscopist. *FEBS J.*, 278:687–703, 2011.

[153] G. Wider. Structure determination of biological macromolecules in solution using nuclear magnetic resonance spectroscopy. *Biotechniques*, 29:1278–1294, 2000.

[154] J. N. S. Evans. *Biomolecular NMR Spectroscopy*. Biomolecular NMR Spectroscopy. Oxford University Press, Oxford, 1995.

[155] R. M. Silverstein and F. X. Webster. *Spectrometric Identification of Organic Compounds*. John Wiley and Sons, New York, 7th edition, 2005.

[156] J. B. Lambert and E. P. Mazzola. *Nuclear Magnetic Resonance: An Introduction to Principles, Applications and Experimental Methods*. Pearson Education, New Jersey, 2004.

[157] M. Ikura, G. M. Clore, A. M. Gronenborn, G. Zhu, C. B. Klee, and A. Bax. Solution structure of a calmodulin-target peptide complex by multidimensional NMR. *Science*, 256:632–638., 1992.

[158] C. Zwahlen, P. Legault, S. J. F. Vincent, J. Greenblatt, R. Konrat, and L. E. Kay. Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: Application to a bacteriophage λ N-peptide/*boxB* RNA complex. *J. Am. Chem. Soc.*, 119:6711–6721, 1997.

[159] A. M. Gronenborn and G. M. Clore. Protein structure determination in solution by two-dimensional and three-dimensional nuclear magnetic resonance spectroscopy. *Anal. Chem.*, 62:2–15, 1990.

[160] K. Wüthrich. *NMR of Proteins and Nucleic Acids*. John Wiley and Sons, New York, 1986.

[161] G. M. Clore and A. M. Gronenborn. Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. *Crit. Rev. Biochem. Mol. Biol.*, 24:479–564, 1989.

[162] D. G. Davis and A. Bax. Assignment of complex ^{1}H NMR spectra via two-dimensional homonuclear Hartmann-Hahn spectroscopy. *J. Am. Chem. Soc.*, 107:2820–2821, 1985.

[163] L. Braunschweiler and R. R. Ernst. Coherance transfer by isotropic mixing: Application to proton correlation spectroscopy. *J. Magn. Reson.*, 53:521–528, 1983.

[164] S. R. Hartmann and E. L. Hahn. Nuclear double resonance in the rotating frame. *Phys. Rev.*, 128:2042–2053, 1962.

[165] M. H. Levitt, D. Suter, and R. R. Ernst. Spin dynamics and thermodynamics in solid-state NMR cross polarisation. *J. Chem. Phys.*, 84:4243–4255, 1986.

[166] W. P. Aue, E. Bartholdi, and R. R. Ernst. Two dimensional spectroscopy. Application to nuclear magnetic resonance. *J. Chem. Phys.*, 64:2229–2246, 1976.

[167] O. W. Sørensen, G. W. Eich, M. H. Levitt, G Bodenhausen, and R. R. Ernst. Product operator formalism for the description of NMR pulse experiments. *Progr. NMR. Spectrosc.*, 16:163–192, 1983.

[168] R. W. Kriwacki and T. P. Pitner. Current aspects of practical two-dimensional (2D) nuclear magnetic resonance (NMR) spectroscopy: Applications to structure elucidation. *Pharm. Res.*, 6:531–554, 1989.

[169] H. Friebolin. *Basic One- and Two-Dimensional NMR Spectroscopy*. VCH Publishers, Weinheim, 1991.

[170] D. Neuhaus and M. P. Williamson. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*. VCH Publishers, New York, 1989.

[171] A. A. Maudsley and R. R. Ernst. Indirect magnetisation of magnetic resonance by heteronuclear two-dimensional spectroscopy. *Chem. Phys. Lett.*, 50:368–372, 1977.

[172] A. Bax. An improved method for heteronuclear chemical shift correlation by two-dimensional nmr. *J. Magn. Reson.*, 42:501–505, 1981.

[173] J. Qin, O. Vinogradova, and A. M. Gronenborn. Protein-protein interactions probed by nuclear magnetic resonance spectroscopy. *Meth. Enzymol.*, 339:377–389, 2001.

[174] O. Schon, A. Friedler, S. Freund, and A. R. Fersht. Binding of p53-derived ligands to MDM2 induces a variety of long range conformational changes. *J. Mol. Biol.*, 336:197–202, 2004.

[175] S. Grzesiek and A. Bax. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J. Mag. Res.*, 99:201–207, 1992.

[176] M Wittekind and L. Mueller. HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J. Magn. Res. Series B*, 101:201–205, 1993.

[177] S. Grzesiek and A. Bax. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.*, 114:6291–6293, 1992.

[178] S. W. Englander and A. J. Wand. Main-chain-directed strategy for the assignment of ^{1}H NMR spectra of proteins. *Biochem.*, 26:5953–5958, 1987.

[179] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, and B. D. Sykes. ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR*, 5:67–81, 1995.

[180] D. S. Wishart and B. D. Sykes. Chemical shifts as a tool for structure determination. *Methods Enzymol.*, 239:363–392, 1994.

[181] D. L. Di Stefano and A. J. Wand. Two dimensional ^{1}H NMR study of human ubiquitin: A main chain directed assignment and structure analysis. *Biochem.*, 26:7272–7281, 1987.

[182] D. S. Wishart, B. D. Sykes, and F. M. Richards. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.*, 222:311–333, 1991.

[183] A. Pastore and V. Saudek. The relationship between chemical shift and secondary structure in proteins. *J. Magn. Res.*, 90:165–176, 1990.

[184] I. D. Kuntz, P. A. Kosen, and E. C. Craig. Amide chemical shifts in many helices in peptides and proteins are periodic. *J. Am. Chem. Soc.*, 113:1406–1408, 1991.

[185] D. C. Dalgarno, B. A. Levine, and R. J. P. Williams. Structural information from NMR secondary chemical shifts of peptide a C-H protons in proteins. *Biosci. Rep.*, 3:443–452, 1983.

[186] N. E. Zhou, B. Zhu, B. D. Sykes, and R. S. Hodges. Relationship between amide proton chemical shifts and hydrogen bonding in amphipathic α-helical peptides. *J. Am. Chem. Soc.*, 114:4320–4326, 1992.

[187] K. Wüthrich, M. Billeter, and W. Braun. Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton distances. *J. Mol. Biol.*, 180: 715–740, 1984.

[188] R. S. Macomber. *A Complete Introduction to Modern NMR Spectroscopy*. John Wiley and Sons, New York, 1998.

[189] Y. Kim and J. H. Prestegard. Measurement of vicinal couplings from cross peaks in COSY spectra. *J. Magn. Reson.*, 84:9–13, 1989.

[190] M. Nilges, M. J. Macias, S. I. O'Donoghue, and H. Oschkinat. Automated NOESY interpretation with ambiguous distance restraints: The refined NMR solution structure of the pleckstrin homology domain from β-spectrin. *J. Mol. Biol.*, 269:408–422, 1997.

[191] M. Nilges and S. I. O'Donoghue. Ambiguous NOEs and automated NOE assignment. *Prog. Nucl. Magn. Reson. Spectrosc.*, 32:107–139, 1998.

[192] A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. Jiang, J. Kuszewski, M. Nilges, N. S Pannu, R. J. Read, L. M. Rice, T. Simpson, and G. L. Warren. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst. D*, 54:905–921, 1989.

[193] M. Habeck, W. Rieping, J. P. Linge, and M. Nilges. NOE assignment with ARIA 2. 0 - The nuts and bolts. In A. K. Downing, editor, *Protein NMR Techniques, Methods in Molecular Biology*, pages 379–402. Humana Press, Totowa, 2004.

[194] M. Nilges. Calculation of protein structures with ambiguous distance restraints. automated assignment of ambiguous NOE crosspeaks and disulphide connectivities. *J. Mol. Biol.*, 245: 645–660, 1995.

[195] P. Güntert, L. D. Berndt, and K. Wüthrich. The program ASNO for computer-supported collection of NOE upper distance constraints as input for protein structure determination. *J. Biomol. NMR*, 3:601–606, 1993.

[196] C. M. Fletcher, D. N. M Jones, R. Diamond, and D. Neuhaus. Treatment of NOE constraints involving equivalent or nonstereoassigned protons in calculations of biomacromolecular structures constraints involving equivalent or nonstereoassigned protons in calculations of biomacromolecular structures. *J. Biomol. NMR*, 8:292–310, 1996.

[197] A. Pardi, M. Billeter, and K. Wüthrich. Calibration of the angular dependence of the amide proton-Cα coupling constants, $3J_{HN\alpha H}$, in a globular protein. *J. Mol. Biol.*, 180:741–751, 1984.

[198] T. D. Mulhern, G. J. Howlett, G. E. Reid, R. J. Simpson, D. J. McColl, R. F. Anders, and R. S. Norton. Solution structure of a polypeptide containing four heptad repeat units from a merozoite surface antigen of *Plasmodium falciparum*. *Biochem.*, 34:3479–3491, 1995.

[199] C. S. Brinkworth, J. A. Carver, K. L. Wegner, J. Doyle, L. E. Llewellyn, and J. H. Bowie. The solution structure of frenatin 3, a neuronal nitric oxide synthase inhibitor from the giant tree frog, *Litoria infrafrenata*. *Biopolymers*, 70:424–434, 2003.

[200] T. F. Havel. An evaluation of computational strategies for use in the determination of protein structure from distance constraints obtained by nuclear magnetic resonance. *Prog. Biophys. Mol. Biol.*, 56:43–78, 1991.

[201] M. Nilges, G. M. Clore, and A. M. Gronenborn. Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms - circumventing problems associated with folding. *FEBS Lett.*, 239: 129–136, 1988.

[202] A. T. Brünger, G. M. Clore, A. M. Gronenborn, and M. Karplus. Three-dimensional structure of proteins determined by molecular dynamics with interproton distance restraints: Application to crambin. *Proc. Natl. Acad. Sci.*, 83:3801–3805, 1986.

[203] G. M. Clore, A. M. Gronenborn, A. T. Brunger, and M. Karplus. Solution conformation of a heptadecapeptide comprising the DNA binding helix F of the cyclic AMP receptor protein of *Escherichia coli*: Combined use of ^{1}H nuclear magnetic resonance and restrained molecular dynamics. *J. Mol. Biol.*, 186:435–455, 1985.

[204] R. Kaptein, E. R. P. Zuiderweg, R. M. Scheek, R. Boelens, and W. F. van Gunsteren. A protein structure from nuclear magnetic resonance data: *lac* repressor headpiece. *J. Mol. Biol.*, 182: 179–182, 1985.

[205] G. M. Clore, A. T. Brünger, M. Karplus, and A. M. Gronenborn. Application of molecular dynamics with interproton distance restraints to three-dimensional protein structure determination. A model study of Crambin. *J. Mol. Biol.*, 191:523–551, 1986.

[206] R. Kaptein, R. Boelens, R. M. Scheek, and W. F. van Gunsteren. Protein structures from NMR. *Biochem.*, 27:5389–5395, 1988.

[207] M. Karplus and G. A. Petsko. Molecular dynamics simulations in biology. *Nature*, 347: 631–639, 1990.

[208] A. T. Brünger, P. D. Adams, and L. M. Rice. New applications of simulated annealing in X-ray crystallography and solution NMR. *Structure*, 5:325–336, 1997.

[209] S. B. Nabburs, C. A. E. M. Spronk, G. Vriend, and G. W. Vuister. Concepts and tools for NMR restraint analysis and validation. *Concepts Mag. Reson. A.*, 22:90–105, 2004.

[210] P. Güntert. Structure calculation of biological macromolecules from NMR data. *Q. Rev. Biophys.*, 31:145–237, 1998.

[211] A. L. Morris, M. W. MacArthur, E. G. Hutchinson, and J. M. Thornton. Stereochemical quality of protein structure coordinates. *Proteins: Struct. Funct. Genet.*, 12:345–364, 1992.

[212] S. G. Hyberts, M. S. Goldberg, T. F. Havel, and G. Wagner. The solution structure of eglin C based on measurements of many NOEs and coupling constants and its comparison with X-ray structures. *Protein Sci.*, 1:736–751, 1992.

[213] G. N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.*, 7:95–99, 1963.

[214] M. Pellegrini and D. F. Mierke. Molecular complex of cholecystokinin-8 and N-terminus of the cholecystokinin A receptor by NMR spectroscopy. *Biochem.*, 38:14775–14783, 1999.

[215] L. Moroder. On the mechanism of hormone recognition and binding by the CCK-B/Gastrin receptor. *J. Pept. Sci.*, 3:1–14, 1997.

[216] R. Bader and O. Zerbe. Are hormones from the neuropeptide y family recognized by their receptors from the membrane-bound state? *ChemBioChem*, 6:1520–1534, 2005.

[217] D. P. Hong, M. Hoshino, R. Kuboi, and Y. Goto. Clustering of fluorine-substituted alcohols as a factor responsible for their marked effects on proteins and peptides. *J. Am. Chem. Soc.*, 121:8427–8433, 1999.

[218] M. Buck. Trifluoroethanol and colleagues: Cosolvents come of age. Recent studies with peptides and proteins. *Q. Rev. Biophys.*, 31:297–355, 1998.

[219] P. Damberg, J. Jarvet, and A. Gräslund. Micellar systems as solvents in peptide and protein structure determination. *Methods Enzymol.*, 339:271–285, 2001.

[220] L. K. Tamm and B. Liang. NMR of membrane proteins in solution. *Prog. Nucl. Magn. Reson. Spectrosc.*, 48:201–210, 2006.

[221] R. Rajan and P. Balaram. A model for the interaction of trifluoroethanol with peptides and proteins. *Int. J. Pept. Protein Res.*, 48:328–336, 1996.

[222] J. Lauterwein, C. Bösch, L. B. Brown, and K. Wüthrich. Physicochemical studies of the protein-lipid interactions in melittin-containing micelles. *Biochim. Biophys. Acta.*, 556: 244–264, 1979.

[223] P. A. McDonnell and S. J. Opella. Effect of detergent concentration on multidimesional solution NMR spectra of membrane proteins in micelles. *J. Magn. Res. B.*, 102:120–125, 1993.

[224] C. Bösch, L. B. Brown, and K. Wüthrich. Physicochemical characterization of glucagon-containing lipid micelles. *Biochim. Biophys. Acta.*, 603:298–312, 1980.

[225] J. Gesell, M. Zasloff, and S. J. Opella. Two-dimesional ^{1}H NMR experiments show that the 23-residue magainin antibiotic peptide is an α-helix in dodecylphosphocoline micelles, sodium dodecylsulfate micelles, and trifluoroethanol/water solution. *J. Biomol. NMR*, 9: 127–135, 1997.

[226] R. R. Vold, R. S. Prosser, and A. J. Deese. Isotropic solutions of phospholipid bicelles - a new membrane mimetic for high-resolution nmr studies of polypeptides. *J. Biomol. NMR*, 9: 329–335, 1997.

[227] L. Indyk and H. F. Fisher. [17] Theoretical aspects of isothermal titration calorimetry. In M. L. Johnson G. K. Ackers, editor, *Energetics of Biological Macromolecules Part B*, volume 295 of *Methods in Enzymology*, pages 350 – 364. Academic Press, 1998.

[228] H. F. Fisher and N. Singh. [9] Calorimetric methods for interpreting protein-ligand interactions. In G. K. Ackers M. L. Johnson, editor, *Energetics of Biological Macromolecules*, volume 259 of *Methods in Enzymology*, pages 194 – 221. Academic Press, 1995.

[229] M. M. Pierce, C. S. Raman, and B. T. Nall. Isothermal titration calorimetry of protein-protein interactions. *Methods*, 19:213–221, 1999.

[230] M. W. Freyer and E. A. Lewis. Isothermal titration calorimetry: Experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. In J. J. Correia and H. W. Detrich III, editors, *Biophysical Tools for Biologists, Volume One: In Vitro Techniques*, volume 84 of *Methods in Cell Biology*, pages 79 – 113. Academic Press, 2008.

[231] E. Freire, O. L. Mayorga, and M. Straume. Isothermal titration calorimetry. *Anal. Chem*, 62: 950A–959A, 1990.

[232] S. Leavitt and E. Freire. Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr. Opin. Struct. Biol.*, 11:560–566, 2001.

[233] E. A. Lewis and K. P. Murphy. Isothermal titration calorimetry. *Methods Mol. Biol*, 305:1–16, 2005.

[234] G. A. Holdgate and W. H. J. Ward. Measurements of binding thermodynamics in drug discovery. *Drug Discov. Today*, 10:1543–1550, 2005.

[235] Z. Lin, F. P. Schwarz, and E. Eisenstein. The hydrophobic nature of GroEL-substrate binding. *J. Biol. Chem.*, 270:1011–1014, 1995.

[236] V. R. Srinivas, R. G. Bhanuprakash, and A. Surolia. A predominantly hydrophobic recognition of H-antigenic sugars by winged bean acidic lectin: A thermodynamic study. *FEBS lett.*, 450: 181–185, 1999.

[237] T. E. Creighton. *Proteins, Structures and Molecular Properties*. W. H. Freeman and Co., New York, 2nd edition, 1993.

[238] K. E. Van Holde, W. C. Johnson, and P. S. Ho. *Principles of Physical Biochemistry*. Pearson Prentice Hall, New Jersey, USA, 2nd edition, 2006.

[239] S. M. Kelly and N. C. Price. The use of circular dichroism in the investigation of protein structure and function. *Curr. Protein Pep. Sci.*, 1:349–384, 2000.

[240] B. M. Bulheller, A. Rodger, and J. D. Hirst. Circular and linear dichroism of proteins. *Phys. Chem. Chem. Phys.*, 9:2020–2035, 2007.

[241] N. Sreerama and R. W. Woody. Computation and analysis of protein circular dichroism spectra. *Meth. Enzymol.*, 383:318–351, 2004.

[242] N. Greenfield and G. D. Fasman. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochem.*, 8:4108–4116, 1969.

[243] T. T Sun, A. Bollen, L. Kahan, and R. R. Traut. Topography of ribosomal proteins of the *Escherichia coli* 30S subunit as studied with the reversible cross-linking reagent methyl 4-mercaptobutyrimidate. *Biochem.*, 13:2334–2340, 1974.

[244] C. Clegg and D. Hayes. Identification of neighbouring proteins in the ribosomes of *Escherichia coli*. A topographical study with the cross-linking reagent dimethyl suberimidate. *Eur. J. Biochem.*, 42:21–28, 1974.

[245] A. Sinz. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev.*, 25:663–82, 2006.

[246] A. Sinz. Investigation of protein-protein interactions in living cells by chemical crosslinking and mass spectrometry. *Anal. Bioanal. Chem.*, 397:3433–40, 2010.

[247] A. Leitner, T. Walzthoeni, A. Kahraman, F. Herzog, O. Rinner, M. Beck, and R. Aebersold. Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell Proteomics*, 9:1634–1649, 2010.

[248] J. Rappsilber. The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.*, 173:530–40, 2011.

[249] E. V. Petrotchenko and C. H. Borchers. Crosslinking combined with mass spectrometry for structural proteomics. *Mass Spectrom. Rev.*, 29:862–876, 2010.

[250] P. Singh, A. Panchaud, and D. R. Goodlett. Chemical cross-linking and mass spectrometry as a low-resolution protein structure determination technique. *Anal. Chem.*, 82:2636–2642, 2010.

[251] D. Mouradov, G. King, I. L. Ross, J. K. Forwood, D. A. Hume, A. Sinz, J. L. Martin, B. Kobe, and T. Huber. Protein structure determination using a combination of cross-linking, mass spectrometry, and molecular modeling. *Methods Mol. Biol.*, 426:459–474, 2008.

[252] J. W. Back, L. de Jong, A. O. Muijsers, and C. G. de Koster. Chemical cross-linking and mass spectrometry for protein structural modeling. *J. Mol. Biol.*, 331:303–313, 2003.

[253] I. A. Kaltashov and S. J. Eyles. *Mass Spectrometry in Structural Biology and Biophysics: Architecture, Dynamics, and Interaction of Biomolecules*. John Wiley & Sons, Hoboken, New Jersey, 2012.

[254] M. M Young, N. Tang, J. C. Hempel, C. M. Oshiro, E. W. Taylor, I. D. Kuntz, B. W. Gibson, and G Dollinger. High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc. Natl. Acad. Sci. USA*, 97:5802–5806, 2000.

[255] A. Sinz. Investigation of protein-protein interactions in living cells by chemical crosslinking and mass spectrometry. *Anal. Bioanal. Chem.*, 397:3433–3440, 2010.

[256] A. Sinz. Chemical cross-linking and mass spectrometry for investigation of protein-protein interactions. In K. M. Downard, editor, *Mass Spectrometry of Protein Interactions*. John Wiley & Sons, Hoboken, New Jersey, 2007.

[257] S. S. Wong. *Chemistry of Protein Conjugation and Cross-Linking*. CRC Press, Boca Raton, 1991.

[258] G. T. Hermanson. *Bioconjugate Techniques*. Academic Press, London, 2nd edition, 2008.

[259] B. Schilling, R. H. Row, B. W. Gibson, X. Guo, and M. M. Young. MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides. *J. Am. Soc. Mass Spectrom.*, 14:834–850, 2003.

[260] A. Maiolica, D. Cittaro, D. Borsotti, L. Sennels, C. Ciferri, C. Tarricone, A. Musacchio, and J. Rappsilber. Structural analysis of multiprotein complexes by cross-linking, mass spectrometry, and database searching. *Mol. Cell Proteomics*, 6:2200–11, 2007.

[261] V. Kulik, M. Weyand, R. Seidel, D. Niks, D. Arac, M. F. Dunn, and I. Schlichting. On the role of αThr183 in the allosteric regulation and catalytic mechanism of tryptophan synthase. *J. Mol. Biol.*, 324:677–690, 2002.

[262] P. D. Bragg and C. Hou. Subunit composition, function, and spatial arrangement in Ca^{2+}-activated and Mg^{2+}-activated adenosine triphosphatases of *Escherichia coli* and *Salmonella typhimurium*. *Arch. Biochem. Biophys.*, 167:311–321, 1975.

[263] A. J. Lomant and G. Faribanks. Chemical probes of extended biological structures - synthesis and properties of cleavable protein cross-linking reagent dithiobis(succinimidyl-S-35 propionate). *J. Mol. Biol.*, 104:243–261, 1976.

[264] J. V. Staros. N-hydroxysulfosuccinimide active esters: Bis(N-hydroxysulfosuccinimide) esters of two dicarboxylic acids are hydrophilic, membrane impermeant, protein cross-linkers. *Biochem.*, 21:3950–3955, 1982.

[265] S. Kalkhof and A. Sinz. Chances and pitfalls of chemical cross-linking with amine reactive N-hydroxysuccinimide esters. *Anal. Bioanal. Chem.*, 392:305–312, 2008.

[266] S. Madler, C. Bich, D. Touboul, and R. Zenobi. Chemical cross-linking with NHS esters: A systematic study on amino acid reactivities. *J. Mass Spectrom.*, 44:694–706, 2009.

[267] G. Bitan and D. B. Teplow. Rapid photochemical cross-linking - a new tool for studies of metastable, amyloidogenic protein assemblies. *Acc. Chem. Res.*, 37:357–364, 2004.

[268] G. Bitan, A. Lomakin, and D. B. Teplow. Amyloid β-protein oligomerization. Prenucleation interactions revealed by photo-induced cross-linking of unmodified proteins. *J. Biol. Chem.*, 276:35176–35184, 2001.

[269] G. Bitan, M. D. Kirkitadze, A. Lomakin, S. S. Vollers, Benedek G. B., and D. B. Teplow. Amyloid β-protein (Aβ) assembly: Aβ40 and Aβ42 oligomerize through distinct pathways. *Proc. Natl. Acad. Sci.*, 100:330–335, 2003.

[270] A. Sinz, S. Kalkhof, and C. Ihling. Mapping protein interfaces by a trifunctional cross-linker combined with MALDI-TOF and ESI-FTICR mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 16:1921–31, 2005.

[271] G. B. Hurst, T. K. Lankford, and S. J. Kennel. Mass spectrometric detection of affinity purified crosslinked peptides. *J. Am. Soc. Mass Spectrom.*, 15:832–9, 2004.

[272] M. A. Nessen, G. Kramer, J. Back, J. M. Baskin, L. E. Smeenk, L. J. de Koning, J. H. van Maarseveen, L. de Jong, C. R. Bertozzi, H. Hiemstra, and C. G. de Koster. Selective enrichment of azide-containing peptides from complex mixtures. *J. Proteome Res.*, 8: 3702–11, 2009.

[273] P. T. Kasper, J. W. Back, M. Vitale, A. F. Hartog, W. Roseboom, L. J. de Koning, J. H. van Maarseveen, A. O. Muijsers, C. G. de Koster, and L. de Jong. An aptly positioned azido group in the spacer of a protein cross-linker for facile mapping of lysines in close proximity. *ChemBioChem*, 8:1281–92, 2007.

[274] S. M. Chowdhury, X. Du, N. Tolic, S. Wu, R. J. Moore, M. U. Mayer, R. D. Smith, and J. N. Adkins. Identification of cross-linked peptides after click-based enrichment using sequential collision-induced dissociation and electron transfer dissociation tandem mass spectrometry. *Anal. Chem.*, 81:5524–32, 2009.

[275] X. Tang, G. R. Munske, W. F. Siems, and J. E. Bruce. Mass spectrometry identifiable cross-linking strategy for studying protein-protein interactions. *Anal. Chem.*, 77:311–8, 2005.

[276] M. Trester-Zedlitz, K. Kamada, S. K. Burley, D. Fenyo, B. T. Chait, and T. W. Muir. A modular cross-linking approach for exploring protein interactions. *J. Am. Chem. Soc.*, 125:2416–25, 2003.

[277] F. Chu, S. Mahrus, C. S. Craik, and A. L. Burlingame. Isotope-coded and affinity-tagged cross-linking (ICATXL): An efficient strategy to probe protein interaction surfaces. *J. Am. Chem. Soc.*, 128:10362–10363, 2006.

[278] T. Taverner, N. E. Hall, R. A. O'Hair, and R. J. Simpson. Characterization of an antagonist interleukin-6 dimer by stable isotope labeling, cross-linking, and mass spectrometry. *J. Biol. Chem.*, 277:46487–46492, 2002.

[279] E. V. Petrotchenko, V. K. Olkhovik, and C. H. Borchers. Isotopically coded cleavable cross-linker for studying protein-protein interaction and protein complexes. *Mol. Cell Proteomics*, 4:1167–1179, 2005.

[280] J. W. Back, V. Notenboom, L. J. de Koning, A. O. Muijsers, T. K. Sixma, C. G. de Koster, and L. de Jong. Identification of cross-linked peptides for protein interaction studies using mass spectrometry and ^{18}O labeling. *Anal. Chem.*, 74:4417–4422, 2002.

[281] D. M. Schulz, S. Kalkhof, A. Schmidt, C. Ihling, C. Stingl, K. Mechtler, O. Zschornig, and A. Sinz. Annexin A2/P11 interaction: New insights into annexin A2 tetramer structure by chemical crosslinking, high-resolution mass spectrometry, and computational modeling. *Proteins*, 69:254–69, 2007.

[282] C. Ihling, A. Schmidt, S. Kalthof, and D. M. Schulz. Isotope-labelled cross-linkers and fourier transform ion cyclotron resonance mass spectrometry for structural analysis of a protein/peptide complex. *J. Am. Soc. Mass Spectrom*, 17:1100–1113, 2006.

[283] M. Q. Muller, F. Dreiocker, C. H. Ihling, M. Schafer, and A. Sinz. Cleavable cross-linker for protein structure analysis: Reliable identification of cross-linking products by tandem MS. *Anal. Chem.*, 82:6958–68, 2010.

[284] A. Kao, C. L. Chiu, D. Vellucci, Y. Yang, V. R. Patel, S. Guan, A. Randall, P. Baldi, S. D. Rychnovsky, and L. Huang. Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Mol. Cell Proteomics*, 10:M110 002212, 2011.

[285] J. W. Back, A. F. Hartog, H. L. Dekker, A. O. Muijsers, L. J. de Koning, and L. de Jong. A new crosslinker for mass spectrometric analysis of the quaternary structure of protein complexes. *J. Am. Soc. Mass Spectrom.*, 12:222–7, 2001.

[286] E. J. Soderblom and M. B. Goshe. Collision-induced dissociative chemical cross-linking reagents and methodology: Applications to protein structural characterization using tandem mass spectrometry analysis. *Anal. Chem.*, 78:8059–68, 2006.

[287] E. J. Soderblom, B. G. Bobay, J. Cavanagh, and M. B. Goshe. Tandem mass spectrometry acquisition approaches to enhance identification of protein-protein interactions using low-energy collision-induced dissociative chemical crosslinking reagents. *Rapid Commun. Mass Spectrom.*, 21:3395–408, 2007.

[288] G. J. King, A. Jones, B. Kobe, T. Huber, D. Mouradov, D. A. Hume, and I. L. Ross. Identification of disulfide-containing chemical cross-links in proteins using MALDI-TOF/TOF-mass spectrometry. *Anal. Chem.*, 80:5036–5043, 2008.

[289] Y. Lu, M. Tanasova, B. Borhan, and G. E. Reid. Ionic reagent for controlling the gas-phase fragmentation reactions of cross-linked peptides. *Anal. Chem.*, 80:9279–9287, 2008.

[290] M. W. Gardner, L. A. Vasicek, S. Shabbir, E. V. Anslyn, and J. S. Brodbelt. Chromogenic cross-linker for the characterization of protein structure by infrared multiphoton dissociation mass spectrometry. *Anal. Chem.*, 80:4807–4819, 2008.

[291] H. Zhang, X. Tang, G. R. Munske, N. Tolic, G. A. Anderson, and J. E. Bruce. Identification of protein-protein interactions and topologies in living cells with chemical cross-linking and mass spectrometry. *Mol. Cell Proteomics*, 8:409–420, 2009.

[292] H. Rink. Solid-phase synthesis of protected peptide fragments using a trialkoxy-diphenyl-methylester resin. *Tetrahedron Lett.*, 28:3787–3790, 1987.

[293] D. Bilusich, V. M. Maselli, C. S. Brinkworth, T. Samguina, A. T. Lebedev, and J. H. Bowie. Direct identification of intramolecular disulfide links in peptides using negative ion electrospray mass spectra of underivatised peptides. A joint experimental and theoretical study. *Rapid Commun. Mass Spectrom.*, 19:3063–3074, 2005.

[294] D. Bilusich and J. H. Bowie. Identification of intermolecular disulfide linkages in underivatised peptides using negative ion electrospray mass spectrometry. A joint experimental and theoretical study. *Rapid Commun. Mass Spectrom.*, 21:619–628, 2007.

[295] H. J. Andreazza and J. H. Bowie. The application of negative ion electrospray mass spectrometry for the sequencing of underivatized disulfide-containing proteins: Insulin and lysozyme. *Phys. Chem. Chem. Phys.*, 12:13400–13407, 2010.

[296] M. Zhang and I. A. Kaltashov. Mapping of protein disulfide bonds using negative ion fragmentation with a broadband precursor selection. *Anal. Chem.*, 78:4820–9, 2006.

[297] D. Vellucci, A. Kao, R. M. Kaake, S. D. Rychnovsky, and L. Huang. Selective enrichment and identification of azide-tagged cross-linked peptides using chemical ligation and mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 21:1432–1445, 2010.

[298] R. J. Waugh, J. H. Bowie, and R. N. Hayes. Collision-induced dissociations of deprotonated peptides. Dipeptides containing Aspartic or Glutamic acids. *Org. Mass Spectrom.*, 26: 250–256, 1991.

[299] T. Wang and J. H. Bowie. unpublished observations.

[300] C. Cheng and M. L. Gross. Applications and mechanisms of charge-remote fragmentation. *Mass Spectrom. Rev.*, 19:398–420, 2000.

[301] A. N. Calabrese, N. J. Good, T. Wang, J. He, J. H. Bowie, and T. L. Pukala. A negative ion mass spectrometry approach to identify cross-linked peptides utilizing characteristic disulfide fragmentations. *J. Am. Soc. Mass Spectrom*, 23:1364–1375, 2012.

[302] A. N. Calabrese, T. Wang, J. H. Bowie, and T. L. Pukala. Negative ion fragmentations of disulfide-containing cross-linking reagents are competitive with aspartic acid side-chain induced cleavages. *Rapid Commun. Mass Spectrom.*, 27:238–248, 2013.

[303] V. Schnaible, S. Wefing, A. Resemann, D. Suckau, A. Bucker, S. Wolf-Kummeth, and D. Hoffman. Screening for disulfide bonds in proteins by MALDI in-source decay and LIFT-TOF/TOF-MS. *Anal. Chem.*, 74:4980–4988, 2002.

[304] M. D. Jones, S. D. Patterson, and H. S. Lu. Determination of disulfide bonds in highly bridged disulfide-linked peptides by matrix-assisted laser desorption/ionization mass spectrometry with postsource decay. *Anal. Chem.*, 70:136–143, 1998.

[305] S. P. Gaucher, M. Z. Hadi, and M. M. Young. Influence of crosslinker identity and position on gas-phase dissociation of Lys-Lys crosslinked peptides. *J. Am. Soc. Mass Spectrom.*, 17: 395–405, 2006.

[306] J. M. Wells, J. L. Stephenson, and S. A. McLuckey. Charge dependence of protonated insulin decompositions. *Int. J. Mass Spectrom.*, 203:A1–A9, 2000.

[307] A. M. Palumbo, S. A. Smith, C. L. Kalcic, M. Dantus, P. M. Stemmer, and G. E. Reid. Tandem mass spectrometry strategies for phosphoproteome analysis. *Mass Spectrom. Rev.*, 30:600–625, 2011.

[308] R. B. Merrifield. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.*, 85:2149–2154, 1963.

[309] L. A. Carpino and A. El-Faham. Effect of tertiary bases on O-benzotriazolyluronium salt-induced peptide segment coupling. *J. Org. Chem.*, 59:695–698, 1994.

[310] W. S. Hancock and J. E. Battersby. A new micro-test for the detection of incomplete coupling reactions in solid-phase peptide synthesis using 2,4,6-trinitrobenzenesulphonic acid. *Anal. Biochem.*, 71:260–264, 1976.

[311] T. M. Dawson and S. H. Snyder. Gases as biological messenger: Nitric oxide and carbon monoxide in the brain. *J. Neurosci.*, 14:5147–5159, 1994.

[312] J. F. Kerwin. Nitric oxide: A new paradigm for second messengers. *J. Med. Chem.*, 38: 4343–4362, 1995.

[313] H. M. Abu-Soud, L. L. Yoho, and D. J. Stuehr. Calmodulin controls neuronal nitric-oxide synthase by a dual mechanism. Activation of intra- and interdomain electron transfer. *J. Biol. Chem.*, 269:32047–32050, 1994.

[314] J. Lincoln, C. H. V. Hoyle, and G. Burnstock. Synthesis and properties of nitric oxide. In *Nitric Oxide in Health and Disease*, pages 12–26. Cambridge University Press, Cambridge, 1997.

[315] P. J. Andrew and B. Mayer. Enzymatic function of nitric oxide synthases. *Cardiovascular Res.*, 43:521–531, 1999.

[316] S. Moncada and E. A. Higgs. The discovery of nitric oxide and its role in vascular biology. *Br. J. Pharmacol.*, 147:5193–5201, 2006.

[317] S. Moncada and E. A. Higgs. Molecular mechanisms and therapeutic strategies related to nitric oxide. *FASEB J.*, 9:1319–1330, 1995.

[318] L. J. Ignarro. Nitric oxide. A novel signal transduction mechanism for transcellular communication. *Hypertension*, 16:477–483, 1990.

[319] J. R. Lancaster. A tutorial on the diffusibility and reactivity of free nitric oxide. *Nitric Oxide*, 1:18–30, 1997.

[320] N. Toda and M. Nakanishi-Toda. Nitric oxide: Ocular blood flow, glaucoma, and diabetic retinopathy. *Prog. Retin. Eye Res.*, 26:205–238, 2007.

[321] C. Czabó. Peroxynitrite: Biochemistry, pathophysiology and development of therapeutics. *Nat. Rev. Drug Discov.*, 6:662–680, 2007.

[322] R. G. Knowles and S. Moncada. Nitric oxide synthases in mammals. *Biochem. J.*, 298: 249–258, 1994.

[323] L. J. Baek, B. A. Thiel, S. Lucus, and D. J. Stuehr. Macrophage nitric oxide subunits. purification, characterization, and the role of prosthetic groups and substrate in regulating their association into a dimeric enzyme. *J. Biol. Chem.*, 268:21120–21129, 1993.

[324] T. R. Billiar. Nitric oxide, novel biology with clinical relevance. *Ann. Surg.*, 221:339–349, 1995.

[325] M. A. Marletta. Nitric oxide synthase structure and mechanism. *J. Biol. Chem.*, 268: 12231–12234, 1993.

[326] D. J. Stuehr and S. Ghosh. Enzymology of nitric oxide synthases. In B. Mayer, editor, *Nitric Oxide: Handbook of Experimental Pharmacology*, pages 33–70. Springer, New York, 2000.

[327] J. Lincoln, C. H. V. Hoyle, and G. Burnstock. Nitric oxide: Introduction and historical background. In *Nitric Oxide in Health and Disease*, pages 12–26. Cambridge University Press, Cambridge, 1997.

[328] J. Pfeilschifter, W. Eberhardt, R. Hummel, D. Kunz, H. Muhl, D. Nitsch, C. Pluss, and G. Walker. Therapeutic strategies for the inhibition of inducible nitric oxide synthase - potential for a novel class of anti-inflammatory agents. *Cell Biol. Int.*, 20:51–58, 1996.

[329] W. K. Alderton, C. E. Copper, and R. G. Knowles. Nitric oxide synthases: Structure, function and inhibition. *Biochem. J.*, 357:593–615, 2001.

[330] D. J. Stuehr. Mammalian nitric oxide synthases. *Biochim. Biophys. Acta*, 1411:217–230, 1999.

[331] U. Siddhanta, A. Presta, B. Fan, D. Wolan, D. L. Rousseau, and D. J. Stuehr. Domain swapping in inducible nitric-oxide synthase. Electron transfer occurs between flavin and heme groups located on adjacent subunits in the dimer. *J. Biol. Chem.*, 273:18950–18958, 1998.

[332] R. Gachhui, H. M. Abu-Soud, D. K. Ghosh, A. Presta, M. A. Blazing, B. Mayer, S. E. George, and D. J. Stuehr. Neuronal nitric-oxide synthase interaction with calmodulin-troponin c chimeras. *J. Biol. Chem.*, 273:5451–5454, 1998.

[333] R. Gachhui, A. Presta, D. F. Bentley, H. M. Abu-Soud, R. McArthur, G. Brudvig, D. K. Ghosh, and D. J. Stuehr. Characterization of the reductase domain of rat neuronal nitric oxide synthase generated in the methylotrophic yeast *Pichia pastoris*. *J. Biol. Chem.*, 271: 20594–20602, 1996.

[334] C. Galli, R. MacArthur, H. M. Abu-Soud, P. Clark, D. J. Stuehr, and G. W. Brudvig. EPR spectroscopic characterization of neuronal NO synthase. *Biochemistry*, 35:2804–2810, 1996.

[335] C. Nathan and Q. W. Xie. Regulation of biosynthesis of nitric oxide. *J. Biol. Chem.*, 269: 13725–13728, 1994.

[336] J. Lincoln, C. H. V. Hoyle, and G. Burnstock. Nitric oxide in the central nervous system. In *Nitric Oxide in Health and Disease*, pages 42–60. Cambridge University Press, Cambridge, 1997.

[337] A. Crivici and M. Ikura. Molecular and structural basis of target recognition by calmodulin. *Annu. Rev. Biophys. Biomol. Struct.*, 24:85–116, 1995.

[338] K. T. O'Neil and W. F. DeGrado. How calmodulin binds its targets: Sequence independent recognition of amphiphilic α-helices. *Trends Biochem. Sci.*, 15:59–64, 1990.

[339] P. B. Stathopulos, J. B. Ames, and M. Ikura. Structural aspects of calcium-binding proteins and their interactions with targets. In J. Krebs and M. Michalak, editors, *Calcium A Matter of Life or Death*, volume 41 of *New Comprehensive Biochemistry*, pages 95 – 123. Elsevier, 2007.

[340] J. Krebs and C. W. Heizmann. Calcium-binding proteins and the EF-hand principle. In Joachim Krebs and Marek Michalak, editors, *Calcium A Matter of Life or Death*, volume 41 of *New Comprehensive Biochemistry*, pages 51 – 93. Elsevier, 2007.

[341] A. R. Rhoads and F. Friedberg. Sequence motifs for calmodulin recognition. *FASEB J.*, 11: 331–340, 1997.

[342] T. Yuan, K. L. Yap, and M. Ikura. Calmodulin target recognition: Common mechanism and structural diversity. In E. Carafoli and J. Krebs, editors, *Calcium Homeostasis*, volume 3. Springer, Berlin, 2000.

[343] H. J. Vogel and M. Zhang. Protein engineering and NMR studies of calmodulin. *Mol. Cell. Biochem.*, 149/150:3–15, 1995.

[344] X. Shen, C. A. Valencia, J. W. Szostak, B. Dong, and R. Liu. Scanning the human proteome for calmodulin-binding proteins. *Proc. Natl. Acad. Sci.*, 102:5969–5974, 2005.

[345] M. S. Marlow and A. J. Wand. Conformational dynamics of calmodulin in complex with the calmodulin-dependent kinase-kinase α calmodulin-binding domain. *Biochem.*, 45: 8732–8741, 2006.

[346] D. Chin and A. R. Means. Calmodulin: A prototypical calcium sensor. *Cell Biol.*, 10:322–328, 2000.

[347] A. R. Means, M. F. A. Van Berkum, I. Bagchi, K. P. Lu, and C. D. Rasmussen. Regulatory functions of calmodulin. *Pharmacol. Ther.*, 50:255–270, 1991.

[348] R. J. P. Williams. Calcium and calmodulin. *Cell Calcium*, 13:355–362, 1992.

[349] H. Kuboniwa, N. Tjandra, S. Grazesiek, H. Ren, C. B. Klee, and A. Bax. Solution structure of calcium-free calmodulin. *Nat. Struct. Biol.*, 2:768–776, 1995.

[350] R. Chattopadhyaya, W. E. Meador, A. R. Means, and F. A. Quiocho. Calmodulin structure refined at 1. 7 Å resolution. *J. Mol. Biol.*, 228:1177–1192, 1992.

[351] J. L. Fallon and F. A. Quiocho. A closed compact structure of native Ca^{2+}-calmodulin. *Structure*, 11:1303–1307, 2003.

[352] Y. S. Babu, J. S. Sack, T. J. Greenhorgh, C. E. Bugg, A. R. Means, and W. J. Cook. Three-dimensional structure of calmodulin. *Nature*, 315:37–40, 1985.

[353] D. A. Taylor, J. S. Sack, J. F. Maune, K. Beckingham, and F. A. Quiocho. Structure of a recombinant calmodulin from drosophila melanogaster refined at 2.2-Å resolution. *J. Biol. Chem.*, 266:21375–21380, 1991.

[354] S. T. Rao, S. Wu, K. A. Satyshur, K-Y. Ling, C. Kung, and M. Sundaralingam. Structure of *Paramecium tetraurelia* calmodulin at 1.8 Å resolution. *Prot. Sci.*, 2:436–447, 1993.

[355] Y. S. Babu, C. E. Bugg, and W. J. Cook. Structure of calmodulin refined at 2.2 Å resolution. *J. Mol. Biol.*, 204:191–204, 1988.

[356] J. J. Chou, S. Li, C. B. Klee, and A. Bax. Solution structure of Ca^{2+}-calmodulin reveals flexible hand-like properties of its domains. *Nat. Struct. Biol.*, 8:990–997, 2001.

[357] G. Barbato, M. Ikura, L. E. Kay, R. W. Pastor, and A. Bax. Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy: The central helix is flexible. *Biochem.*, 31:5269–5278, 1992.

[358] D. B. Heidom and J. Trewhella. Comparison of crystal and solution structures of calmodulin and troponin c. *Biochem.*, 27:909–915, 1988.

[359] M. Ikura, S. Spera, G. Barbato, L. E. Kay, M. Krinks, and A. Bax. Secondary structure and side-chain ^{1}H and ^{13}C resonace assignments of calmodulin in solution by heteronuclear multidimensional NMR spectroscopy. *Biochem.*, 30:9216–9228, 1991.

[360] M. Walsh and F. C. Stevens. Characterization of tryptic fragments obtained from bovine brain protein modulator of cyclic nucleotide phosphodiesterase. *J. Biol. Chem.*, 252:7440–7443, 1977.

[361] S. Spera, M. Ikura, and A. Bax. Measurement of the exchange rates of rapidly exchanging amide protons: Application to the study of calmodulin and its complex with a myosin light chain kinase fragment. *J. Biomol. NMR*, 1:155–165, 1991.

[362] W. Drabikowski, J. Kuznicki, and Z. Grabarek. Similarity in Ca^{2+}-induced changes between troponin-C and protein activator of 3':5'-cyclic nucleotide phosphodiesterase and their tryptic fragments. *Biochim. Biophys. Acta*, 485:123–133, 1977.

[363] J. Hu and L. J. Van Eldik. *Regulation of nitric oxide synthase by calmodulin*, pages 287–345. Academic Press, New York, USA, 1998.

[364] Y. Watanabe, Y. Hu, and H. Hidaka. Identification of a specific amino acid cluster in the calmodulin-binding domain of the neuronal nitric oxide synthase. *FEBS Lett.*, 403:75–78, 1996.

[365] H. M. Abu-Soud and D. J. Stuehr. Nitric oxide synthases reveal a role for calmodulin in controlling electron transfer. *Proc. Natl. Acad. Sci.*, 90:10769–10772, 1993.

[366] S. W. Vetter and E. Leclerc. Novel aspects of calmodulin target recognition and activation. *Eur. J. Biochem.*, 270:404–414, 2003.

[367] J. A. Cox, M Comte, J. E. Fitton, and W. F. DeGrado. The interaction of calmodulin with amphiphilic peptides. *J. Biol. Chem.*, 260:2527–2534, 1985.

[368] B. Elshorst, M. Hennig, H. Forsterling, A. Diener, M. Maurer, P. Schulte, H. Schwalbe, C. Griesinger, J. Krebs, H. Schmid, T. Vorherr, and E. Carafoli. NMR solution structure of a complex of calmodulin with a binding peptide of the Ca^{2+} pump. *Biochem.*, 38: 12320–12332, 1999.

[369] G. M. Contessa, M. Orsale, S. Melino, V. Torre, M. Paci, A. Desider, and D. O. Cicero. Structure of calmodulin complexed with an olfactory CNG channel fragment and role of the central linker: Residual dipolar couplings to evaluate calmodulin binding modes outside the kinase family. *J. Biomol. NMR*, 31:185–199, 2005.

[370] P. L. Wintrode and P. L. Privalov. Energetics of target peptide recognition by calmodulin: A calorimetric study. *J. Mol. Biol.*, 266:1050–1062, 1997.

[371] M. R. Nelson and W. J. Chazin. Calmodulin as a calcium sensor. In L. J. Van Eldik and D. M. Watterson, editors, *Calmodulin and Signal Transduction*, pages 17–64. Academic Press, New York, USA, 1998.

[372] W. E. Meador, A. R. Means, and F. A. Quiocho. Modulation of calmodulin plasticity in molecular recogntition on the basis of X-ray structures. *Science*, 262:1718–1721, 1993.

[373] W. E. Meador, A. R. Means, and F. A. Quiocho. Target enzyme recognition by calmodulin: 2.4 Å structure of a calmodulin peptide complex. *Science*, 257:1251–1255, 1992.

[374] M. Osawa, H. Tokumits, M. B. Swindells, H. Kurihara, M. Orita, T. Shibanuma, T. Furuya, and M. Ikura. A novel target recognition revealed by calmodulin in complex with Ca^{2+}-calmodulin-dependent kinase kinase. *Nat. Struct. Biol.*, 6:819–824, 1999.

[375] B. T. Clarke. The natural history of amphibian skin secretions, their normal functioning and potential medical applications. *Biol. Rev.*, 72:365–379, 1997.

[376] L. H. Lazarus and M. Attila. The toad, ugly and venomous, wears yet a precious jewel in his skin. *Prog. Neurobiol.*, 41:473–507, 1993.

[377] A. C. C. Nascimento, W. Fontes, A. Sebben, and M. S. Castro. Antimicrobial peptides from anurans skin secretions. *Protein Pept. Lett.*, 10:227–238, 2003.

[378] V. Erspamer. Bioactive secretions of the amphibian integument. In H. Heatwole and G. T. Barthalmus, editors, *Amphibian Biology: The Integument*, pages 178–350. Surrey, Beatty and Sons, Chipping Norton, 1994.

[379] D. Barra and M. Simmaco. Amphibian skin - a promising resource for antimicrobial peptides. *Trends Biotechnol.*, 13:205–209, 1995.

[380] J. M. Cei, V. Erspamer, and M. Rosechini. Taxonomic and evolutionary significance of biogenic amines and polypeptides occurring in amphibian skin. I. Neotropical leptodactylid frogs. *Syst. Zool.*, 16:328–342, 1967.

[381] G. A. Cunha Filho, C. A. Schwartz, I. S. Resck, M. M. Murta, S. S. Lemos, M. S. Castro, C. Kyaw, O. R. Pires, Jr., and J. R. S. Leite. Antimicrobial activity of the bufadienolides marinobufagin and telocinobufagin isolated as major components from skin secretion of the toad *Bufo rubescens*. *Toxicon*, 45:777–782, 2005.

[382] J. W. Daly. Thirty years of discovering arthropod alkaloids in amphibian skin. *J. Nat. Prod.*, 61:162–172, 1998.

[383] V. Erspamer and P. Melchiorri. Active polypeptides: From amphibian skin to the gastrointestinal tract and brain of mammals. *Trends Pharmacol. Sci.*, 1:391–395, 1980.

[384] T. L Pukala, J. H. Bowie, V. M. Maselli, I. F. Musgrave, and M. J. Tyler. Host-defence peptides from the glandular secretions of amphibians: Structure and activity. *Nat. Prod. Rep.*, 23: 368–393, 2006.

[385] J. H. Bowie, B. C. S. Chia, and M. J. Tyler. Host defence peptides from the skin glands of Australian amphibians: A powerful chemical arsenal. *Pharmacol. News*, 5:16–21, 1998.

[386] S. Grenard. *Frogs and toads*, pages 2–127. Reptiles and Amphibian Magazine, Pottsville, 1994.

[387] A. W. Bannon, M. W. Decker, M. W. Holladay, P. Curzon, D. Donnelly-Roberts, P. S. Puttfarcken, R. S. Bitner, A. Diaz, A. H. Dickenson, R. D. Porsolt, M. Williams, and S. P. Arneric. Broad-spectrum, non-opioid analgesic activity by selective modulation of neuronal nicotinic acetylcholine receptors. *Science*, 279:77–81, 1998.

[388] G. Bertaccini. Active polypeptides of nonmammalian origins. *Pharmacol. Rev.*, 28:127–177, 1976.

[389] L. S. Jacob and M. Zasloff. Potential therapeutic applications of magainins and other antimicrobial agents of animal origin. In J. Marsh and J. A. Goode, editors, *Antimicrobial Peptides.*, volume 186, pages 197–223. John Wiley and Sons, London, 1994.

[390] J. M. Conlon. The therapeutic potential of antimicrobial peptides from frog skin. *Rev. Med. Microbiol.*, 15:17–25, 2004.

[391] V. Erspamer, G. F. Erspamer, G. Mazzanti, and R. Endean. Active peptides in the skins of one hundred amphibian species from Australia and Papua New Guinea. *Comp. Biochem. Physiol. C*, 77:99–108, 1984.

[392] M. A. Apponyi, T. L Pukala, C. S Brinkworth, V. M. Maselli, J. H. Bowie, M. J. Tyler, G. W. Booker, J. C. Wallace, J. A. Carver, F. Separovic, J. Doyle, and L. E. Llewellyn. Host-defence peptides of Australian anurans: Structure, mechanism of action and evolutionary significance. *Peptides*, 25:1035–1054, 2004.

[393] R. J. Jackway, T. L. Pukala, S. C. Donnellan, P. J. Sherman, M. J. Tyler, and J. H. Bowie. Skin peptide and cDNA profiling of Australian anurans: Genus and species identification and evolutionary trends. *Peptides*, 32:161–172, 2011.

[394] J. H. Bowie, F. Separovic, and M. J. Tyler. Host-defense peptides of Australian anurans. Part 2. Structure, activity, mechanism of action, and evolutionary significance. *Peptides*, 37: 174–188, 2012.

[395] T. Rozek, J. H. Bowie, J. C. Wallace, and M. J. Tyler. The antibiotic and anticancer active aurein peptides from the Australian bell frogs *Litoria aurea* and *Litoria raniformis*. Part 2. Sequence determination using electrospray mass spectrometry. *Rapid Commun. Mass Spectrom.*, 14:2002–2011, 2000.

[396] P. A. Wabnitz, J. H. Bowie, M. J. Tyler, J. C. Wallace, and B. P. Smith. Differences in the skin peptides of the male and female Australian tree frog *Litoria splendida* - The discovery of the aquatic male sex pheromone splendipherin, together with Phe8 caerulein and a new antibiotic peptide caerin 1.10. *Eur. J. Biochem.*, 267:269–275, 2000.

[397] D. J. M. Stone, R. J. Waugh, J. H. Bowie, J. C. Wallace, and M. J. Tyler. Peptides from Australian frogs. The structures of the caerins from *Litoria caerulea*. *J. Chem. Res.*, 138: 910–936, 1993.

[398] R. J. Waugh, D. J. M. Stone, J. H. Bowie, J. C. Wallace, and M. J. Tyler. Peptides from Australian frogs. The structures of the caerins and caeridins from *Litoria gilleni*. *J. Chem. Res.*, 139:937–961, 1993.

[399] S. T. Steinborner, R. J. Waugh, J. H. Bowie, J. C. Wallace, M. J. Tyler, and S. L. Ramsay. New caerin antibacterial peptides from the skin glands of the Australian tree frog *Litoria xanthomera*. *J. Pept. Sci.*, 3:181–5, 1997.

[400] S. T. Steinborner, G. J. Currie, J. H. Bowie, J. C. Wallace, and M. J. Tyler. New antibiotic caerin 1 peptides from the skin secretion of the Australian tree frog *Litoria chloris* - Comparison of the activities of the caerin 1 peptides from the genus *Litoria*. *J. Pept. Res.*, 51:121–126, 1998.

[401] A. Anastasi, V. Erspamer, and R. Endean. Isolation and amino acid sequence of caerulein, the active decapeptide of the skin of *Hyla caerulea*. *Arch. Biochem. Biophys.*, 125:57–68, 1968.

[402] K. L. Wegener, P. A. Wabnitz, J. A. Carver, J. H. Bowie, B. C. S. Chia, J. C. Wallace, and M. J. Tyler. Host defence peptides from the skin glands of the Australian Blue Mountains tree frog *Litoria citropa*. Solution structure of the antibacterial peptide citropin 1.1. *Eur. J. Biochem.*, 265:627–637, 1999.

[403] K. L. Wegener, C. S. Brinkworth, J. H. Bowie, J. C. Wallace, and M. J. Tyler. Bioactive dahlein peptides from the skin secretions of the Australian aquatic frog *Litoria dahlii*: Sequence determination by electrospray mass spectrometry. *Rapid Commun. Mass Spectrom.*, 15: 1726–1734, 2001.

[404] R. J. Waugh, M. J. Raftery, J. H. Bowie, J. C. Wallace, and M. J. Tyler. The structures of the frenatin peptides from the skin secretions of the giant tree frog, *Litoria infrafrenata*. *J. Pept. Sci.*, 2:117–124, 1996.

[405] J. Doyle, L. E. Llewellyn, C. S. Brinkworth, J. H. Bowie, K. L. Wegener, T. Rozek, P. A. Wabnitz, J. C. Wallace, and M. J. Tyler. Amphibian peptides that inhibit neuronal nitric oxide synthase: The isolation of lesueurin from the skin secretion of the Australian stony creek frog *Litoria lesueuri*. *Eur. J. Biochem.*, 269:100–109, 2002.

[406] T. Rozek, R. J. Waugh, S. T. Steinborner, J. H. Bowie, M. J. Tyler, and J. C. Wallace. The maculatin peptides from the skin glands of the tree frog *Litoria genimaculata* - A comparison of the structures and antibacterial activities of maculatin 1. 1 and caerin 1. 1. *J. Peptide Sci.*, 4:111–115, 1998.

[407] C. S. Brinkworth, J. H. Bowie, M. J. Tyler, and J. C. Wallace. A comparison of the host defence skin peptides of the New Guinea tree frog (*Litoria genimaculata*) and the fringed tree frog

(*Litoria eucnemis*). The link between the caerin and the maculatin antimicrobial peptides. *Aust. J. Chem.*, 55:605–610, 2002.

[408] V. M. Maselli, D. Bilusich, J. H. Bowie, and M. J. Tyler. Host-defence skin peptides of the Australian streambank froglet *Crinia riparia*: Isolation and sequence determination by positive and negative ion electrospray mass spectrometry. *Rapid Commun. Mass Spectrom.*, 20:797–803, 2006.

[409] V. M. Maselli, C. S Brinkworth, J. H. Bowie, and M. J. Tyler. Host-defence skin peptides of the Australian common froglet *Crinia signifera*: Sequence determination and negative ion electrospray mass spectra. *Rapid Commun. Mass Spectrom.*, 18:2155–2161, 2004.

[410] S. T. Steinborner, C. W. Gao, M. J. Raftery, R. J. Waugh, T. Blumenthal, J. H. Bowie, J. C. Wallace, and M. J. Tyler. The structures of four tryptophyllin and three rubellidin peptides from the Australian red tree frog *Litoria rubella*. *Aust. J. Chem.*, 47:2099–2108, 1994.

[411] A. M. Bradford, M. J. Raftery, J. H. Bowie, M. J. Tyler, J. C. Wallace, G. W. Adams, and C. Severini. Novel uperin peptides from the dorsal glands of the Australian floodplain toadlet *Uperoleia inundata*. *Aust. J. Chem.*, 49:475–484, 1996.

[412] A. M. Bradford, J. H. Bowie, M. J. Tyler, and J. C. Wallace. New antibiotic uperin peptides from the dorsal glands of the Australian toadlet *Uperoleia mjobergii*. *Aust. J. Chem.*, 49: 1325–1331, 1996.

[413] J. R. Doyle, J. H. Bowie, R. J. Jackway, L. E. Llewellyn, T. L Pukala, M. A. Apponyi, and G. W. Booker. Anuran host-defense peptides that complex with Ca^{2+} calmodulin and inhibit the synthesis of the cell signaling agent nitric oxide by neuronal nitric oxide synthase. In J. Howl and S. Jones, editors, *Bioactive Peptides*. CRC Press, London, 2009.

[414] R. C. Renteria and M. Constantine-Paton. Nitric oxide in the retinotectal system: A signal but not a retrograde messenger during map refinement and segregation. *J. Neurosci.*, 19: 7066–7076, 1999.

[415] A. Gobbetti and M. Zeran. Hormonal and cellular brain mechanisms regulating the amplexus of male and female water frog (*Rana esculenta*). *J. Neuroendocrinol.*, 11:589–596, 1999.

[416] M. Molero, I. M. Hernandez, P. Lobo, P. Cardenas, R. Romero, and J. Chacin. Modulation by nitric oxide of gastric acid secretion in toads. *Acta Physiol. Scand.*, 164:229–236, 1998.

[417] J. Doyle, C. S Brinkworth, K. L. Wegener, J. A. Carver, L. E. Llewellyn, I. N. Olver, J. H. Bowie, P. A. Wabnitz, and M. J. Tyler. nNOS inhibition, antimicrobial and anticancer activity of the amphibian skin peptide, citropin 1. 1 and synthetic modification: The solution structure of a modified citropin 1. 1. *Eur. J. Biochem.*, 270:1141–1153, 2003.

[418] A. Fersht. *Enzyme Structure and Mechanism*. W. H. Freeman, New York, NY, USA, 1987.

[419] T. L. Pukala, T. Urathamakul, S. J. Watt, J. L. Beck, R. J. Jackway, and J. H. Bowie. Binding studies of nNOS-active amphibian peptides and Ca^{2+} calmodulin, using negative ion ESI mass spectrometry. *Rapid Commun. Mass Spectrom.*, 22:3501–3509, 2008.

[420] P. Kursula and V. Majava. A structural insight into lead neurotoxicity and calmodulin activation by heavy metals. *Acta Cryst.*, F63:653–656, 2007.

[421] V. Majava, M. V. Petoukhov, N. Hayashi, P. Pirilä, D. I. Svergun, and P. Kursula. Interaction between the C-terminal region of human myelin basic protein and calmodulin: Analysis of complex formation and solution structure. *BMC Struct. Biol.*, 8:doi:10.1186/1472–6807–8–10, 2008.

[422] D. C. LaPorte, B. M. Wierman, and D. R. Storm. Calcium-induced exposure of a hydrophobic surface on calmodulin. *Biochem.*, 19:3814–3819, 1980.

[423] K. P. Hoeflich and M. Ikura. Calmodulin in action: Diversity in target recognition and activation mechanisms. *Cell*, 108:739–742, 2002.

[424] L. Carlier, C. Byrne, E. Miclet, S. Bourgoin-Voillard, M. Nicaise, J.-C. Tabet, M. Desmadril, G. Leclercq, O. Lequin, and Y. Jacquot. Biophysical studies of the interaction between calmodulin and the R^{287}-T^{311} region of human estrogen receptor α reveals an atypical binding process. *Biochem. Biophys. Res. Commun.*, 419:356–361, 2012.

[425] Q. Shi, X. Wang, and J. Ren. Biophysical characterization of the interaction of p21 with calmodulin: A mechanistic study. *Biophys. Chem.*, 138:138–143, 2008.

[426] M. Myllykoski, K. Kuczera, and P. Kursula. Complex formation between calmodulin and a peptide from the intracellular loop of the gap junction protein connexin43: Molecular conformation and energetics of binding. *Biophys. Chem.*, 144:130–135, 2009.

[427] R. D. Brokx, M. M Lopez, H. J. Vogel, and G. I. Makhatadze. Energetics of target peptide binding by calmodulin reveals different modes of binding. *J. Biol. Chem.*, 276:14083–14091, 2001.

[428] K. Y. Joyce, A. M. Giannetti, and J. M. Bradshaw. Thermodynamics of calmodulin trapping by Ca^{2+}/calmodulin-dependent protein kinase II: Subpicomolar k_d determined using competition titration calorimetry. *Biochem.*, 46:4017–4027, 2007.

[429] T. Wyttenbach, M. Grabenauer, K. Thalassinos, J. H. Scrivens, and M. T. Bowers. The effect of calcium ions and peptide ligands on the relative stabilities of the calmodulin dumbbell and compact structures. *J. Phys. Chem. B*, 114:437–447, 2009.

[430] O. Nemirovskiy, R. Ramanthan, and M. L. Gross. Investigation of calcium-induced, noncovalent association of calmodulin with melittin by electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom*, 8:809–812, 1997.

[431] J. Pan and L. Konermann. Calcium-induced structural transitions of the calmodulin- melittin system studied by electrospray mass spectrometry: Conformational subpopulations and metal-unsaturated intermediates. *Biochem.*, 49:3477–3486, 2010.

[432] M. Zhang, T. Tanaka, and M. Ikura. Calcium-induced conformational transition revealed by the solution structure of apo calmodulin. *Nat. Struct. Biol.*, 2:758–767, 1995.

[433] Z. A. Ataman, K. Gakhar, B. R. Sorensen, J. W. Hell, and M. A. Shea. The NMDA receptor NR1 C1 region bound to calmodulin: Structural insights into functional differences between homologous domains. *Structure*, 15:1603–1617, 2007.

[434] N. Juranic, E. Atanasova, A. G. Filoteo, S. Macura, F. G. Prendergast, J. T. Penniston, and E.E. Strehler. Calmodulin wraps around its binding domain in the plasma membrane Ca^{2+} pump anchored by a novel 18-1 motif. *J. Biol. Chem.*, 285:4015–4024, 2010.

[435] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, and B. D. Sykes. ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR*, 5:67–81, 1995.

[436] R. X. Xu, J. M. Word, D. G. Davis, M. J. Rink, D. H. Willard, and R. T. Gampe. Solution structure of the human pp60c-src SH2 domain complexed with a phosphorylated tyrosine pentapeptide. *Biochem.*, 34:2107–2121, 1995.

[437] K. L Wegener, J. A. Carver, and J. H. Bowie. The solution structures and activity of caerin 1.1 and caerin 1.4 in aqueous trifluoroethanol and dodecylphosphocholine micelles. *Biopolymers*, 69:42–59, 2003.

[438] M. Ikura, L. E. Kay, M. Krinks, and A. Bax. Triple-resonance multidimensional NMR study of calmodulin complexed with the binding domain of skeletal muscle myosin light-chain kinase: Indication of conformational change in the central helix. *Biochem.*, 50:5498–5504, 1991.

[439] M. Ikura, L. E. Kay, and A. Bax. A novel approach for sequential assignment of ^{1}H, ^{13}C and ^{15}N spectra of larger proteins: Heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochem.*, 29:4659–4667, 1990.

[440] 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. *Biochem.*, 31:1647–1651, 1992.

[441] D. S. Wishart and B. D. Sykes. The ^{13}C chemical-shift index: A simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol. NMR*, 4:171–180, 1994.

[442] M. A. Apponyi. Amphibian Skin Peptides Which Inhibit nNOS: Structure and Binding Studies Using Heteronuclear NMR. Chemistry Ph. D. Thesis, The University of Adelaide, 2006.

[443] P. Hu, Y. Qi-Zhuang, and J. A. Loo. Calcium stoichiometry determination for calcium binding proteins by electrospray ionization mass spectrometry. *Anal. Chem.*, 66:4190–4194, 1994.

[444] J. Pan, K. Xu, X. Yang, W.-Y. Choy, and L. Konermann. Solution-phase chelators for suppressing nonspecific protein- metal interactions in electrospray mass spectrometry. *Anal. Chem.*, 81:5008–5015, 2009.

[445] A. N. Calabrese, L. Speechley, and T. L. Pukala. Characterisation of calmodulin structural transitions by ion mobility mass spectrometry. *Aust. J. Chem.*, 65:504–511, 2012.

[446] S. J. Watt, A. Oakley, M. Shiel, and J. Beck. Comparison of negative and positive ion electrospray ionization mass spectra of calmodulin and its complex with trifluoperazine. *Rapid Commun. Mass Spectrom.*, 19:2123–2130, 2005.

[447] I. A. Kaltashov and A. Mohimen. Estimates of protein surface areas in solution by electropray ionization mass spectrometry. *Anal. Chem.*, 77:5370–5379, 2005.

[448] I. A. Kaltashov and R. R. Abzalimov. Do ionic charges in ESI MS provide useful information on macromolecular structure? *J. Am. Soc. Mass Spectrom.*, 19:1239–1246, 2008.

[449] U. A. Mirza, S. L. Cohen, and B. T. Chait. Heat-induced conformational changes in proteins studied by electrospray ionization mass spectrometry. *Anal. Chem.*, 65:1–6, 1993.

[450] P. A. Faull, K. E. Korkeila, J. M. Kalapothakis, A. Gray, B. J. McCullough, and P. E. Barran. Gas-phase metalloprotein complexes interrogated by ion mobility-mass spectrometry. *Int. J. Mass Spectrom.*, 283:140–148, 2009.

[451] T. L Pukala. Structural and Mechanistic Studies of Bioactive Peptides. Chemistry Ph. D. Thesis, The University of Adelaide, 2006.

[452] R. J. Jackway. Biologically Active Peptides from Australian Amphibians. Chemistry Ph. D. Thesis, The University of Adelaide, 2008.

[453] F. H. Schumann, H. Riepl, T. Maurer, W. Gronwald, K.-P. Neidig, and H. R. Kalbitzer. Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions. *J. Biomol. NMR*, 39:275–289, 2007.

[454] R. Gopalakrishna and W. B. Anderson. Ca^{2+}-induced hydrophobic site on calmodulin: Application for purification of calmodulin by phenyl-sepharose affinity chromatography. *Biochem. Biophys. Res. Commun.*, 104:830–836, 1982.

[455] T. H. Crouch and C. B. Klee. Positive cooperative binding of calcium to bovine brain calmodulin. *Biochem.*, 19:3692–3698, 1980.

[456] A. A. Shvartsburg and R. D. Smith. Fundamentals of traveling wave ion mobility spectrometry. *Anal. Chem.*, 80:9689–9699, 2008.

[457] K. Thalassinos, M. Grabenauer, S. E. Slade, G. R. Hilton, M. T. Bowers, and J. H. Scrivens. Characterization of phosphorylated peptides using traveling wave-based and drift cell ion mobility mass spectrometry. *Anal. Chem.*, 81:248–254, 2008.

[458] L. R. Brown. Use of fully deuterated micelles for conformational studies of membrane proteins by high resolution ^{1}H nuclear magnetic resonance. *Biochim. Biophys. Acta*, 557: 135–148, 1979.

[459] B. K. John, D. Plant, P. Webb, and R. E. Hurd. Effective combination of gradients and crafted RF pulses for water suppression in biological samples. *J. Magn. Res.*, 98:200–206, 1992.

[460] D. Marion and K. Wüthrich. Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^{1}H-^{1}H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.*, 113:967–974, 1983.

[461] J. P. Linge, W. Habeck, W. Rieping, and M. Nilges. ARIA: Automated NOE assignment and NMR structure calculation. *Bioinformatics*, 19:315–316, 2003.

[462] A. T. Brünger. Crystallographic refinement by simulated annealing - Application to a 2.8 Å resolution structure of aspartate aminotransferase. *J. Mol. Biol.*, 203:803–816, 1988.

[463] K. Pari, G. A. Mueller, E. F. DeRose, T. W. Kirby, and R. E. London. Solution structure of the RNase H domain of the HIV-1 reverse transcriptase in the presence of magnesium. *Biochem.*, 42:639–650, 2003.

[464] R. S. Kang, C. M. Daniels, S. A. Francis, S. C. Shih, W. J. Salerno, L. Hicke, and I. Radhakrishnan. Solution structure of a CUE-ubiquitin complex reveals a conserved mode of ubiquitin binding. *Cell*, 113:621–630, 2003.

[465] W. Humphrey, A. Dalke, and K. Schulten. VMD: Visual molecular dynamics. *J. Mol. Graphics*, 14:33–38, 1996.

[466] R. Koradi, M. Billeter, and K. Wüthrich. MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics*, 14:51–55, 1996.

[467] L. E. Kay, P. Keifer, and T. Saarinen. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.*, 114: 10663–10665, 1992.

[468] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, 6: 277–293, 1995.

[469] 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*, 59:687–696, 2005.

[470] D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, and B. D. Sykes. ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR*, 6:135–140, 1995.

[471] D. Eisenberg and M. Jucker. The amyloid state of proteins in human diseases. *Cell*, 148: 1188–1203, 2012.

[472] A. Abbott. Dementia: A problem for our age. *Nature*, 475:S2–S4, 2011.

[473] T. Härd and C. Lendel. Inhibition of amyloid formation. *J. Mol. Biol.*, 421:441–465, 2012.

[474] J. M. Mason, N. Kokkoni, K. Stott, and A. J. Doig. Design strategies for anti-amyloid agents. *Curr. Opin. Struct. Biol.*, 13:526–532, 2003.

[475] I. W. Hamley. The amyloid beta peptide: A chemist's perspective. Role in Alzheimer's and fibrillization. *Chem. Rev.*, 112:5147–5192, 2012.

[476] K. A. DaSilva, J. E. Shaw, and J. McLaurin. Amyloid-β fibrillogenesis: Structural insight and therapeutic intervention. *Exp. Neurol.*, 223:311–321, 2010.

[477] P. K. Auluck, G. Caraveo, and S. Lindquist. α-Synuclein: Membrane interactions and toxicity in Parkinson's disease. *Annu. Rev. Cell Dev. Biol.*, 26:211–233, 2010.

[478] H. A. Lashuel, C. R. Overk, A. Oueslati, and E. Masliah. The many faces of α-synuclein: From structure and toxicity to therapeutic target. *Nat. Rev. Neurosci.*, 14:38–48, 2012.

[479] P. Westermark, A. Andersson, and G. T. Westermark. Islet amyloid polypeptide, islet amyloid, and diabetes mellitus. *Physiol. Rev.*, 91:795–826, 2011.

[480] P. Cao, A. Abedini, and D. P. Raleigh. Aggregation of islet amyloid polypeptide: From physical chemistry to cell biology. *Curr. Opin. Struct. Biol.*, 23:82–89, 2013.

[481] E. Ahmad, A. Ahmad, S. Singh, M. Arshad, A. H. Khan, and R. H. Khan. A mechanistic approach for islet amyloid polypeptide aggregation to develop anti-amyloidogenic agents for type-2 diabetes. *Biochimie*, 93:793, 2011.

[482] D. P. Smith, A. E. Ashcroft, and S. E. Radford. Hemodialysis-related amyloidosis. In M. Ramirez-Alvarado, J. W. Kelly, and C. M. Dobson, editors, *Protein Misfolding Diseases: Current and Emerging Principles and Therapies*, pages 347–380. Wiley, 2010.

[483] D. B. Corlin and N. H. H. Heegaard. β2-microglobulin amyloidosis. In J. R. Harris, editor, *Protein Aggregation and Fibrillogenesis in Cerebral and Systemic Amyloid Disease*, pages 517–540. Springer, 2012.

[484] T. Eichner and S. E. Radford. Understanding the complex mechanisms of β2-microglobulin amyloid assembly. *FEBS J.*, 278:3868–3883, 2011.

[485] V. N. Uversky and A. L. Fink. Conformational constraints for amyloid fibrillation: the importance of being unfolded. *Biochim. Biophys. Acta.*, 1698:131–153, 2004.

[486] M. Stefani and C. M. Dobson. Protein aggregation and aggregate toxicity: New insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med.*, 81:678–699, 2003.

[487] S. L. Gras. Amyloid fibrils: From disease to design. New biomaterial applications for self-assembling cross-β fibrils. *Aust. J. Chem.*, 60:333–342, 2007.

[488] S. L. Gras, A. K. Tickler, A. M. Squires, G. L. Devlin, M. A. Horton, C. M. Dobson, and C. E. MacPhee. Functionalised amyloid fibrils for roles in cell adhesion. *Biomaterials*, 29: 1553–1562, 2008.

[489] K. Rajagopal and J. P. Schneider. Self-assembling peptides and proteins for nanotechnological applications. *Curr. Opin. Struct. Biol.*, 14:480–486, 2004.

[490] L. Yu, I. A. Banerjee, X. Gao, N. Nuraje, and H. Matsui. Fabrication and application of enzyme-incorporated peptide nanotubes. *Bioconjugate Chem.*, 16:1484–1487, 2005.

[491] S. Zhang. Fabrication of novel biomaterials through molecular self-assembly. *Nat. Biotechnol.*, 21:1171–1178, 2003.

[492] H. Ecroyd, T. Koudelka, D. C. Thorn, D. M. Williams, G. Devlin, P. Hoffmann, and J. A. Carver. Dissociation from the oligomeric state is the rate-limiting step in fibril formation by κ-casein. *J. Biol. Chem.*, 283:9012–9022, 2008.

[493] M. Stefani. Structural polymorphism of amyloid oligomers and fibrils underlies different fibrillization pathways: Immunogenicity and cytotoxicity. *Curr. Protein Pep. Sci.*, 11:343–354, 2010.

[494] J. T. Jarrett and P. T. Lansbury Jr. Seeding "one-dimensional crystallization" of amyloid: A pathogenic mechanism in Alzheimer's disease and scrapie? *Cell*, 73:1055–1058, 1993.

[495] T. X. Hoang, L. Marsella, A. Trovato, F. Seno, J. R. Banavar, and A. Maritan. Common attributes of native-state structures of proteins, disordered proteins, and amyloid. *Proc. Natl. Acad. Sci.*, 103:6883–6888, 2006.

[496] L. C. Serpell, M. Sunde, M. D. Benson, G. A. Tennent, M. B. Pepys, and P. E. Fraser. The protofilament substructure of amyloid fibrils. *J. Mol. Biol.*, 300:1033–1039, 2000.

[497] N. Carulla, G. L. Caddy, D. R. Hall, J. Zurdo, M. Gairí, M. Feliz, E. Giralt, C. V. Robinson, and C. M. Dobson. Molecular recycling within amyloid fibrils. *Nature*, 436:554–558, 2005.

[498] R. Nelson, M. R. Sawaya, M. Balbirnie, A. Ø. Madsen, C. Riekel, R. Grothe, and D. Eisenberg. Structure of the cross-β spine of amyloid-like fibrils. *Nature*, 435:773–778, 2005.

[499] T. P. Knowles, A. W. Fitzpatrick, S. Meehan, H. R. Mott, M. Vendruscolo, C. M. Dobson, and M. E. Welland. Role of intermolecular forces in defining material properties of protein nanofibrils. *Science*, 318:1900–1903, 2007.

[500] J. Zurdo, J. I. Guijarro, and C. M. Dobson. Preparation and characterization of purified amyloid fibrils. *J. Am. Chem. Soc.*, 123:8141–8142, 2001.

[501] M. Sunde, L. C. Serpell, M. Bartlam, P. E. Fraser, M. B Pepys, and C. C. F. Blake. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.*, 273:729–739, 1997.

[502] J. L. Jiménez, E. J. Nettleton, M. Bouchard, C. V. Robinson, C. M. Dobson, and H. R. Saibil. The protofilament structure of insulin amyloid fibrils. *Proc. Natl. Acad. Sci.*, 99:9196–9201, 2002.

[503] H. A. Lashuel, S. R. LaBrenz, L. Woo, L. C. Serpell, and J. W. Kelly. Protofilaments, filaments, ribbons, and fibrils from peptidomimetic self-assembly: Implications for amyloid fibril formation and materials science. *J. Am. Chem. Soc.*, 122:5262–5277, 2000.

[504] J. L. Jimenez, J. I. Naki Guijarro, E. Orlova, J. Zurdo, C. M. Dobson, M. Sunde, and H. R. Saibil. Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. *EMBO J.*, 18:815–821, 1999.

[505] J.-C. Rochet and P. T. Lansbury. Amyloid fibrillogenesis: Themes and variations. *Curr. Opin. Struct. Biol.*, 10:60–68, 2000.

[506] C. M. Dobson. Protein folding and misfolding. *Nature*, 426:884–890, 2003.

[507] C. Nordstedt, J. Näslund, L. O. Tjernberg, A. R. Karlström, J. Thyberg, and L. Terenius. The Alzheimer's Aβ-peptide develops protease resistance in association with its polymerization into fibrils. *J. Biol. Chem.*, 269:30773–30776, 1994.

[508] M. Bély and Á. Apáthy. Histochemical and immunohistochemical differential diagnosis of amyloidosis - A brief illustrated essay and personal experience with Romhányi's method. *Amyloid*, 7:212–217, 2000.

[509] M. Stefani. Structural features and cytotoxicity of amyloid oligomers: Implications in Alzheimer's disease and other diseases with amyloid deposits. *Prog. Neurobiol.*, 99, 2012.

[510] R. Kayed and C. Lasagna-Reeves. Molecular mechanisms of amyloid oligomers toxicity. *J. Alzheimer's Dis.*, 33:S67–S78, 2012.

[511] S. M. Butterfield and H. A. Lashuel. Amyloidogenic protein-membrane interactions: Mechanistic insight from model systems. *Angew. Chem. Int. Ed.*, 49:5628–5654, 2010.

[512] M. Bucciantini, E. Giannoni, F. Chiti, F. Baroni, L. Formigli, J Zurdo, N. Taddei, G. Ramponi, C. M. Dobson, and M. Stefani. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*, 416:507–511, 2002.

[513] M. Sakono and T. Zako. Amyloid oligomers: Formation and toxicity of Aβ oligomers. *FEBS J.*, 277:1348–1358, 2010.

[514] M. Fändrich. Oligomeric intermediates in amyloid formation: Structure determination and mechanisms of toxicity. *J. Mol. Biol.*, 421:427–440, 2012.

[515] K. Berthelot, C. Cullin, and S. Lecomte. What does make an amyloid toxic: Morphology, structure or interaction with membrane? *Biochimie*, 95:12–19, 2013.

[516] M. Bucciantini, D. Nosi, M. Forzan, E. Russo, M. Calamai, L. Pieri, L. Formigli, F. Quercioli, S. Soria, and F. Pavone. Toxic effects of amyloid fibrils on cell membranes: The importance of ganglioside GM1. *FASEB J.*, 26:818–831, 2012.

[517] A. L. Gharibyan, V. Zamotin, K. Yanamandra, O. S. Moskaleva, B. A. Margulis, I. A Kostanyan, and L. A. Morozova-Roche. Lysozyme amyloid oligomers and fibrils induce cellular death via different apoptotic/necrotic pathways. *J. Mol. Biol.*, 365:1337–1349, 2007.

[518] V. Novitskaya, O. V. Bocharova, I. Bronstein, and I. V. Baskakov. Amyloid fibrils of mammalian prion protein are highly toxic to cultured cells and primary neurons. *J. Biol. Chem.*, 281: 13828–13836, 2006.

[519] M. F. M. Engel, L. Khemtémourian, C. C. Kleijer, H. J. D. Meeldijk, J. Jacobs, A. J. Verkleij, B. De Kruijff, J. A. Killian, and J. W. M. Höppener. Membrane damage by human islet amyloid polypeptide through fibril growth at the membrane. *Proc. Natl. Acad. Sci.*, 105:6033–6038, 2008.

[520] S. Campioni, B. Mannini, M. Zampagni, A. Pensalfini, C. Parrini, E. Evangelisti, A. Relini, M. Stefani, C. M. Dobson, and C. Cecchi. A causative link between the structure of aberrant protein oligomers and their toxicity. *Nat. Chem. Biol.*, 6:140–147, 2010.

[521] D. A. Butterfield, J. Drake, C. Pocernich, and A. Castegna. Evidence of oxidative damage in Alzheimer's disease brain: Central role for amyloid β-peptide. *Trends Mol. Med.*, 7:548–554, 2001.

[522] M. S. Wolfe. Therapeutic strategies for Alzheimer's disease. *Nat. Rev. Drug Discov.*, 1: 859–866, 2002.

[523] K. L. Sciarretta, D. J. Gordon, and S. C. Meredith. Peptide-based inhibitors of amyloid assembly. *Meth. Enzymol.*, 413:273–312, 2006.

[524] M. Zasloff. Antimicrobial peptides of multicellular organisms. *Nature*, 415:389–395, 2002.

[525] A. Cederlund, G. H. Gudmundsson, and B. Agerberth. Antimicrobial peptides important in innate immunity. *FEBS J.*, 278:3942–3951, 2011.

[526] K. A. Brogden. Antimicrobial peptides: Pore formers or metabolic inhibitors in bacteria? *Nat. Rev. Microbiol.*, 3:238–250, 2005.

[527] P. Vlieghe, V. Lisowski, J. Martinez, and M. Khrestchatisky. Synthetic therapeutic peptides: Science and market. *Drug Discov. Today*, 15:40–56, 2010.

[528] C. D. Fjell, J. A. Hiss, R. E. W. Hancock, and G. Schneider. Designing antimicrobial peptides: Form follows function. *Nat. Rev. Drug Discov.*, 11:37–51, 2011.

[529] N. Y. Yount and M. R. Yeaman. Emerging themes and therapeutic prospects for anti-infective peptides. *Annu. Rev. Pharmacol. Toxicol.*, 52:337–360, 2012.

[530] B. M. Peters, M. E. Shirtliff, and M. A. Jabra-Rizk. Antimicrobial peptides: Primeval molecules or future drugs? *PLoS Pathogens*, 6:e1001067, 2010.

[531] G. Diamond, N. Beckloff, A. Weinberg, and K. O. Kisich. The roles of antimicrobial peptides in innate host defense. *Curr. Pharm. Des.*, 15:2377–2392, 2009.

[532] J. Vizioli and M. Salzet. Antimicrobial peptides from animals: Focus on invertebrates. *Trends Pharmacol. Sci.*, 23:494–496, 2002.

[533] H. G Boman. Peptide antibiotics and their role in innate immunity. *Annu. Rev. Immunol.*, 13: 61–92, 1995.

[534] L. T. Nguyen, E. F. Haney, and H. J. Vogel. The expanding scope of antimicrobial peptide structures and their modes of action. *Trends Biotech.*, 29:464–472, 2011.

[535] L. Yang, T. A. Harroun, T. M. Weiss, L. Dong, and H. W. Huang. Barrel-stave model or toroidal model? A case study on melittin pores. *Biophys. J.*, 81:1475–1485, 2001.

[536] G. Ehrenstein and H. Lecar. Electrically gated ionic channels in lipid bilayers. *Quart. Rev. Biophys.*, 10:1–34, 1977.

[537] K. Matsuzaki. Magainins as a paradigm for the mode of action of pore forming polypeptides. *Biochim. Biophys. Acta*, 10:391–400, 1998.

[538] H. W. Huang. Action of antimicrobial peptides: Two-state model. *Biochem.*, 39:8347–8352, 2000.

[539] R. M. Epand and H. J. Vogel. Diversity of antimicrobial peptides and their mechanisms of action. *Biochim. Biophys. Acta*, 1462:11–28, 1999.

[540] D. Andreu, J. Ubach, A. Boman, B. Wahlin, D. Wade, R. B. Merrifield, and H. G. Boman. Shortened cecropin A-melittin hybrids - significant size reduction retains potent antibiotic activity. *FEBS Lett.*, 296:190–194, 1992.

[541] S. J. Ludtke, K. He, W. T. Heller, T. A. Harroun, L. Yang, and H. W. Huang. Membrane pores induced by magainin. *Biochem.*, 35:13723–13728, 1996.

[542] Y. Pouny, D. Rapaport, A. Mor, P. Nicolas, and Y. Shai. Interaction of antimicrobial dermaseptin and its fluorescently labeled analogues with phospholipid membranes. *Biochem.*, 31:12416–12423, 1992.

[543] Y. Shai. Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by α-helical antimicrobial and cell non-selective membrane-lytic peptides. *Biochim. Biophys. Acta*, 1462:55–70, 1999.

[544] B. Bechinger. The structure, dynamics and orientation of antimicrobial peptides in membranes by multidimensional solid-state NMR spectroscopy. *Biochim. Biophys. Acta*, 1462: 157–183, 1999.

[545] Z. Oren and Y. Shai. Mode of action of linear amphipathic α-helical antimicrobial peptides. *Biopolymers*, 47:451–463, 1998.

[546] M. Dathe and T. Wieprecht. Structural features of helical antimicrobial peptides: Their potential to modulate activity on model membranes and biological cells. *Biochim. Biophys. Acta*, 1462:71–87, 1999.

[547] K. L. Wegener. Amphibian peptides: Their structures and bioactivities. Chemistry Ph. D. Thesis, The University of Adelaide, 2001.

[548] J. H. Bowie, R. J. Jackway, F. Separovic, J. A. Carver, and M. J. Tyler. Host-defense peptides from the secretions of the skin glands of frogs and toads: Membrane-active peptides from the genera *Litoria*, *Uperoleia*, and *Crinia*. In J. Howl and S. Jones, editors, *Bioactive Peptides*, pages 333–355. CRC Press, London, 2009.

[549] F. Harris, S. R. Dennison, and D. A. Phoenix. Aberrant action of amyloidogenic host defense peptides: A new paradigm to investigate neurodegenerative disorders? *FASEB J.*, 26: 1776–1781, 2012.

[550] A. K. Mahalka and P. K. J. Kinnunen. Binding of amphipathic α-helical antimicrobial peptides to lipid membranes: Lessons from temporins B and L. *Biochim. Biophys. Acta. Biomembr.*, 1788:1600–1609, 2009.

[551] B. L. Kagan, H. Jang, R. Capone, F. Teran Arce, S. Ramachandran, R. Lal, and R. Nussinov. Antimicrobial properties of amyloid peptides. *Mol. Pharm.*, 9:708–717, 2011.

[552] H. A. Lashuel and P. T. Lansbury. Are amyloid diseases caused by protein aggregates that mimic bacterial pore-forming toxins? *Q. Rev. Biophys.*, 39:167–201, 2006.

[553] S. J. Soscia, J. E. Kirby, K. J. Washicosky, S. M. Tucker, M. Ingelsson, B. Hyman, M. A. Burton, L. E. Goldstein, S. Duong, R. E. Tanzi, and R. D. Moir. The Alzheimer's disease-associated amyloid β-protein is an antimicrobial peptide. *PLoS One*, 5:e9505, 2010.

[554] P. K. Singh and S. K. Maji. Amyloid-like fibril formation by tachykinin neuropeptides and its relevance to Amyloid β-protein aggregation and toxicity. *Cell Biochem. Biophys.*, 64:29–44, 2012.

[555] M. R. Nilsson. Techniques to study amyloid fibril formation *in vitro*. *Methods*, 34:151–160, 2004.

[556] M. Biancalana and S. Koide. Molecular mechanism of thioflavin-T binding to amyloid fibrils. *Biochim. Biophys. Acta*, 1804:1405–1412, 2010.

[557] 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. *J. Struct. Biol.*, 151:229–238, 2005.

[558] L. S. Wolfe, M. F. Calabrese, A. Nath, D. V. Blaho, A. D. Miranker, and Y. Xiong. Protein-induced photophysical changes to the amyloid indicator dye thioflavin T. *Proc. Natl. Acad. Sci.*, 107:16863–16868, 2010.

[559] C. Bleiholder, N. F. Dupuis, T. Wyttenbach, and M. T. Bowers. Ion mobility-mass spectrometry reveals a conformational conversion from random assembly to β-sheet in amyloid fibril formation. *Nat. Chem.*, 3:172–177, 2010.

[560] G. Van Meer, D. R. Voelker, and G. W. Feigenson. Membrane lipids: Where they are and how they behave. *Nat. Rev. Mol. Cell. Biol.*, 9:112–124, 2008.

[561] B. C. S. Chia, J. Torres, M. A. Cooper, I. T. Arkin, and J. H. Bowie. The orientation of the antibiotic peptide maculatin 1. 1 in DMPG and DMPC lipid bilayers. Support for a pore forming mechanism. *FEBS Lett.*, 512:47–51, 2002.

[562] F. C. Dehle, H. Ecroyd, I. F. Musgrave, and J. A. Carver. αB-crystallin inhibits the cell toxicity associated with amyloid fibril formation by κ-casein and the amyloid-β peptide. *Cell Stress Chaperon.*, 15:1013–1026, 2010.

[563] D. E. Ehrnhoefer, M. Duennwald, P. Markovic, J. L. Wacker, S. Engemann, M. Roark, J. Legleiter, J. L. Marsh, L. M. Thompson, S. Lindquist, P. J. Muchowski, and E. E. Wanker. Green tea (-)-epigallocatechin-gallate modulates early events in huntingtin misfolding and reduces toxicity in Huntington's disease models. *Hum. Mol. Genet.*, 15:2743–2751, 2006.

[564] D. E. Ehrnhoefer, J. Bieschke, A. Boeddrich, M. Herbst, L. Masino, R. Lurz, S. Engemann, A. Pastore, and E. E. Wanker. EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nat. Struct. Mol. Biol.*, 15:558–566, 2008.

[565] S. A. Mandel, T. Amit, O. Weinreb, L. Reznichenko, and M. B. H. Youdim. Simultaneous manipulation of multiple brain targets by green tea catechins: a potential neuroprotective strategy for Alzheimer and Parkinson diseases. *CNS Neurosci. Ther.*, 14:352–365, 2008.

[566] S. A. Hudson, H. Ecroyd, F. C. Dehle, I. F. Musgrave, and J. A. Carver. (-)-Epigallocatechin-3-gallate (EGCG) maintains κ-casein in its pre-fibrillar state without redirecting its aggregation pathway. *J. Mol. Biol.*, 392:689–700, 2009.

[567] F. Chiti. Relative importance of hydrophobicity, net charge, and secondary structure propensities in protein aggregation. In V. N. Uversky and A. L. Fink, editors, *Protein Misfolding, Aggregation, and Conformational Diseases*, volume 4 of *Protein Reviews*, pages 43–59. Springer, 2006.

[568] A.-M. Fernandez-Escamilla, F. Rousseau, J. Schymkowitz, and L. Serrano. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.*, 22:1302–1306, 2004.

[569] M. Bucciantini, G. Calloni, F. Chiti, L. Formigli, D. Nosi, C. M. Dobson, and M. Stefani. Prefibrillar amyloid protein aggregates share common features of cytotoxicity. *J. Biol. Chem.*, 279:31374–31382, 2004.

[570] M. M. Pallitto, J. Ghanta, P. Heinzelman, L. L. Kiessling, and R. M. Murphy. Recognition sequence design for peptidyl modulators of β-amyloid aggregation and toxicity. *Biochem.*, 38:3570–3578, 1999.

[571] K. Watanabe, K. Nakamura, S. Akikusa, T. Okada, M. Kodaka, T. Konakahara, and H. Okuno. Inhibitors of fibril formation and cytotoxicity of β-amyloid peptide composed of KLVFF recognition element and flexible hydrophilic disrupting element. *Biochem. Biophys. Res. Commun.*, 290:121–124, 2002.

[572] C. Hetényi, Z. Szabó, É. Klement, Z. Datki, T. Körtvélyesi, M. Zarándi, and B. Penke. Pentapeptide amides interfere with the aggregation of β-amyloid peptide of Alzheimer's disease. *Biochem. Biophy. Res. Commun.*, 292:931–936, 2002.

[573] L. O. Tjernberg, C. Lilliehöök, D. J. E. Callaway, J. Näslund, S. Hahne, J. Thyberg, L. Terenius, and C. Nordstedt. Controlling amyloid β-peptide fibril formation with protease-stable ligands. *J. Biol. Chem.*, 272:12601–12605, 1997.

[574] L. O. Tjernberg, J. Näslund, F. Lindqvist, J. Johansson, A. R. Karlström, J. Thyberg, L. Terenius, and C. Nordstedt. Arrest of-amyloid fibril formation by a pentapeptide ligand. *J. Biol. Chem.*, 271:8545–8548, 1996.

[575] Y. Shai. Mode of action of membrane active antimicrobial peptides. *Biopolymers*, 66: 236–248, 2002.

[576] K. Matsuzaki. Physicochemical interactions of amyloid β-peptide with lipid bilayers. *Biochim. Biophys. Acta Biomembr.*, 1768:1935–1942, 2007.

[577] R. Gößler-Schöfberger, G. Hesser, M. Muik, C. Wechselberger, and A. Jilek. An orphan dermaseptin from frog skin reversibly assembles to amyloid-like aggregates in a pH-dependent fashion. *FEBS J.*, 276:5849–5859, 2009.

[578] A. E. Ashcroft. Mass spectrometry and the amyloid problem - How far can we go in the gas phase? *J. Am. Soc. Mass Spectrom.*, 21:1087–1096, 2010.

[579] S. Goda, K. Takano, K. Yutani, Y. Yamagata, R. Nagata, H. Akutsu, S. Maki, and K. Namba. Amyloid protofilament formation of hen egg lysozyme in highly concentrated ethanol solution. *Protein Sci.*, 9:369–375, 2000.

[580] 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. *J. Mol. Biol.*, 300:541–549, 2000.

[581] L. Nielsen, R. Khurana, A. Coats, S. Frokjaer, J. Brange, S. Vyas, V. N. Uversky, and A. L. Fink. Effect of environmental factors on the kinetics of insulin fibril formation: Elucidation of the molecular mechanism. *Biochem.*, 40:6036–6046, 2001.

[582] W. Liu, J. M. Prausnitz, and H. W. Blanch. Amyloid fibril formation by peptide LYS (11-36) in aqueous trifluoroethanol. *Biomacromol.*, 5:1818–1823, 2004.

[583] D. E. Otzen. Amyloid formation in surfactants and alcohols: Membrane mimetics or structural switchers? *Curr. Protein Pep. Sci.*, 11:355–371, 2010.

[584] L. Chaloin, P. Vidal, A. Heitz, N. Vanmau, J. Mery, G. Divita, and F. Heitz. Conformations of primary amphipathic carrier peptides in membrane mimicking environments. *Biochem.*, 36: 11179–11187, 1997.

[585] S.-C. Li and C. M. Deber. Peptide environment specifies conformation. Helicity of hydrophobic segments compared in aqueous, organic, and membrane environments. *J. Biol. Chem.*, 268:22975–22978, 1993.

[586] F. Chiti, P. Webster, N. Taddei, A. Clark, M. Stefani, G. Ramponi, and C. M. Dobson. Designing conditions for *in vitro* formation of amyloid protofilaments and fibrils. *Proc. Natl. Acad. Sci.*, 96:3590–3594, 1999.

[587] V. L. Anderson, T. F. Ramlall, C. C. Rospigliosi, W. W. Webb, and D. Eliezer. Identification of a helical intermediate in trifluoroethanol-induced alpha-synuclein aggregation. *Proc. Natl. Acad. Sci.*, 107:18850–18855, 2010.

[588] Y. Fezoui and D. B. Teplow. Kinetic studies of amyloid β-protein fibril assembly. Differential effects of α-helix stabilization. *J. Biol. Chem.*, 277:36948–36954, 2002.

[589] Y. Yang, S. Barker, M. J. Chen, and K. H. Mayo. Effect of low molecular weight aliphatic alcohols and related compounds on platelet factor 4 subunit association. *J. Biol. Chem.*, 268: 9223–9229, 1993.

[590] J. S. Albert and A. D. Hamilton. Stabilization of helical domains in short peptides using hydrophobic interactions. *Biochem.*, 34:984–990, 1995.

[591] W. Liu, D. Bratko, J. M. Prausnitz, and H. W. Blanch. Effect of alcohols on aqueous lysozyme-lysozyme interactions from static light-scattering measurements. *Biophys. Chem.*, 107:289–298, 2004.

[592] M. Llinas and M. P. Klein. Solution conformation of the ferrichromes. VI. Charge relay at the peptide bond. Proton magnetic resonance study of solvation effects on the amide electron density distribution. *J. Am. Chem. Soc.*, 97:4731–4737, 1975.

[593] P. D. Thomas and K. A. Dill. Local and nonlocal interactions in globular proteins and mechanisms of alcohol denaturation. *Protein Sci.*, 2:2050–2065, 1993.

[594] A. Abedini and D. P. Raleigh. A role for helical intermediates in amyloid formation by natively unfolded polypeptides? *Phys. Biol.*, 6:015005, 2009.

[595] M. D. Kirkitadze, M. M. Condron, and D. B. Teplow. Identification and characterization of key kinetic intermediates in amyloid β-protein fibrillogenesis. *J. Mol. Biol.*, 312:1103–1119, 2001.

[596] Y. Porat, A. Abramowitz, and E. Gazit. Inhibition of amyloid fibril formation by polyphenols: Structural similarity and aromatic interactions as a common inhibition mechanism. *Chem. Biol. Drug Des.*, 67:27–37, 2006.

[597] Y. Porat, Y. Mazor, S. Efrat, and E. Gazit. Inhibition of islet amyloid polypeptide fibril formation: A potential role for heteroaromatic interactions. *Biochemistry*, 43:14454–14462, 2004.

[598] D. Cao, Y. Zhang, H. Zhang, L. Zhong, and X. Qian. Systematic characterization of the covalent interactions between (-)-epigallocatechin gallate and peptides under physiological conditions by mass spectrometry. *Rapid Commun. Mass Spectrom.*, 23:1147–1157, 2009.

[599] R. Zorilla, L. Liang, G. Remondetto, and M. Subirade. Interaction of epigallocatechin-3-gallate with β-lactoglobulin: Molecular characterization and biological implication. *Dairy Sci. Technol.*, 91:629–644, 2011.

[600] J. Bieschke, J. Russ, R. P. Friedrich, D. E. Ehrnhoefer, H. Wobst, K. Neugebauer, and E. E. Wanker. EGCG remodels mature α-synuclein and amyloid-β fibrils and reduces cellular toxicity. *Proc. Natl. Acad. Sci.*, 107:7710–7715, 2010.

[601] F. Meng, A. Abedini, A. Plesner, C. B. Verchere, and D. P. Raleigh. The flavanol (-)-epigallocatechin 3-gallate inhibits amyloid formation by islet amyloid polypeptide, disaggregates amyloid fibrils and protects cultured cells against IAPP induced toxicity. *Biochem.*, 49:8127–8133, 2010.

[602] I. R. Chandrashekaran, C. G. Adda, C. A. MacRaild, R. F. Anders, and R. S. Norton. EGCG disaggregates amyloid-like fibrils formed by *Plasmodium falciparum* merozoite surface protein 2. *Arch. Biochem. Biophys.*, 513:153–157, 2011.

[603] P. Cao and D. P. Raleigh. Analysis of the inhibition and remodeling of islet amyloid polypeptide amyloid fibers by flavanols. *Biochem.*, 51:2670–2683, 2012.

[604] D. Chen, V. Milacic, M. S. Chen, S. B. Wan, W. H. Lam, C. Huo, K. R. Landis-Piwowar, Q. C. Cui, A. Wali, T. H. Chan, and Q. P. Dou. Tea polyphenols, their biological effects and potential molecular targets. *Histol. Histopathol.*, 23:487–496, 2008.

[605] C. S. Yang, J. D. Lambert, and S. Sang. Antioxidative and anti-carcinogenic activities of tea polyphenols. *Arch. Toxicol.*, 83:11–21, 2009.

[606] H. Shoval, D. Lichtenberg, and E. Gazit. The molecular mechanisms of the anti-amyloid effects of phenols. *Amyloid*, 14:73–87, 2007.

[607] A. T. Petkova, Y. Ishii, J. J. Balbach, O. N. Antzutkin, R. D. Leapman, F. Delaglio, and R. Tycko. A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Acad. Sci.*, 99:16742–16747, 2002.

[608] M. Török, S. Milton, R. Kayed, P. Wu, T. McIntire, C. G. Glabe, and R. Langen. Structural and dynamic features of Alzheimer's Aβ peptide in amyloid fibrils studied by site-directed spin labeling. *J. Biol. Chem.*, 277:40810–40815, 2002.

[609] A. D. Williams, E. Portelius, I. Kheterpal, J.-T. Guo, K. D. Cook, Y. Xu, and R. Wetzel. Mapping Aβ amyloid fibril secondary structure using scanning proline mutagenesis. *J. Mol. Biol.*, 335: 833–842, 2004.

[610] A. Der-Sarkissian, C. C. Jao, J. Chen, and R. Langen. Structural organization of α-synuclein fibrils studied by site-directed spin labeling. *J. Biol. Chem.*, 278:37530–37535, 2003.

[611] F. Chiti, M. Calamai, N. Taddei, M. Stefani, G. Ramponi, and C. M. Dobson. Studies of the aggregation of mutant proteins *in vitro* provide insights into the genetics of amyloid diseases. *Proc. Natl. Acad. Sci.*, 99:16419–16426, 2002.

[612] M. Fändrich and C. M. Dobson. The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *EMBO J.*, 21:5682–5690, 2002.

[613] M. López de La Paz, K. Goldie, J. Zurdo, E. Lacroix, C. M. Dobson, A. Hoenger, and L. Serrano. De novo designed peptide-based amyloid fibrils. *Proc. Natl. Acad. Sci.*, 99:16052–16057, 2002.

[614] V. Daggett. α-Sheet: The toxic conformer in amyloid diseases? *Acc. Chem. Res.*, 39:594–602, 2006.

[615] P. K. J. Kinnunen. Amyloid formation on lipid membrane surfaces. *Open Biol. J.*, 2:163–175, 2009.

[616] H. Zhao, R. Sood, A. Jutila, S. Bose, G. Fimland, J. Nissen-Meyer, and P. K. J. Kinnunen. Interaction of the antimicrobial peptide pheromone Plantaricin A with model membranes: Implications for a novel mechanism of action. *Biochim. Biophys. Acta Biomembr.*, 1758: 1461–1474, 2006.

[617] P. A. C. Cloos and S. Christgau. Non-enzymatic covalent modifications of proteins: Mechanisms, physiological consequences and clinical applications. *Matrix Biol.*, 21:39–52, 2002.

[618] P. A. C. Cloos and S. Christgau. Post-translation modifications of proteins: Implications for aging, antigen recognition and autoimmunity. *Biogerontol.*, 5:139–158, 2004.

[619] J. E. Visick and S. Clarke. Repair, refold, recycle: How bacteria can deal with spontaneous and environmental damage to proteins. *Mol. Microbiol.*, 15:835–845, 1995.

[620] P. Galletti, D. Ingrosso, C. Manna, G. Clemente, and V. Zappia. Protein damage and methylation-mediated repair in the erythrocyte. *Biochem. J.*, 306:313–325, 1995.

[621] C. T. Walsh, S. Garneau-Tsodikova, and G. J. Gatto Jr. Protein posttranslational modifications: The chemistry of proteome diversifications. *Angew. Chem. Int. Ed.*, 44:7342–7372, 2005.

[622] M. I. Aguilar, A. W. Purcell, R. Devi, R. Lew, J. Rossjohn, A. I. Smith, and P. Perlmutter. β-amino acid-containing hybrid peptides - new opportunities in peptidomimetics. *Org. Biomol. Chem.*, 5:2884–2890, 2007.

[623] G. Walsh and R. Jefferis. Post-translational modifications in the context of therapeutic proteins. *Nat. Biotechnol.*, 24:1241–1252, 2006.

[624] A. N. Calabrese, K. Markulic, I. F. Musgrave, H. Guo, L. Zhang, and J. H. Bowie. Structural and activity changes in three bioactive anuran peptides when Asp is replaced by isoAsp. *Peptides*, 38:427–436, 2012.

[625] L. W. Dick, D. Qiu, and K. C. Cheng. Identification and measurement of isoaspartic acid formation in the complementarity determining region of a fully human monoclonal antibody. *J. Chromatog. B.*, 877:3841–3849, 2009.

[626] J. J. Cournoyer, J. L. Pittman, V. B. Ivleva, E. Fallows, L. Waskell, C. E. Costello, and P. B. O'Connor. Deamidation: Differentiation of aspartyl form isoaspartyl products in peptides by electron capture dissociation. *Protein Sci.*, 14:452–463, 2005.

[627] D. W Aswad, M. V Paranandi, and B. T Schurter. Isoaspartate in peptides and proteins: Formation, significance and analysis. *J. Pharm. Biomed. Anal.*, 21:1129–1136, 2000.

[628] K. J. Reissner and D. W. Aswad. Deamidation and isoaspartate formation in proteins: Unwanted alterations or surreptitious signals? *Cell Mol. Life Sci.*, 60:1281–1295, 2003.

[629] V. Schirch, S. Delle Fratte, and M. di Salvo. Detection of isoaspartate residues as a posttranslational modification of proteins and peptides. *Methods Mol. Biol.*, 194:269–276, 2002.

[630] T. Geiger and S. Clarke. Deamidation, isomerization and racemization at asparaginyl and aspartyl residues in peptides. *J. Biol. Chem.*, 262:785–794, 1987.

[631] D. Lacroix, S. Prado, D. Kamoga, J. Kasenene, S. Zirah, and B. Bodo. Unprecedented occurrence of isoaspartic acid in a plant cyclopeptide. *Org. Lett.*, 14:576–579, 2012.

[632] M. Xie and R. L. Schowen. Secondary structure and protein deamidation. *J. Pharm. Sci.*, 88: 8–13, 1999.

[633] R. Ragone, A. Facchiano, A. M. Facchiano, and G. Colonna. Flexibility plots of proteins. *Protein Eng.*, 2:497–504, 1989.

[634] S. M. Potter, W. J. Henzel, and D. W. Aswad. *In vitro* aging of calmodulin generates isoaspartate at multiple Asn-Gly and Asp-Gly sites in calcium binding domains II, III and IV. *Protein Sci.*, 2:1648–1663, 1993.

[635] B. A. Johnson, J. M. Shirokawa, W. S. Hancock, W. Spellman, M., L. J. Basa, and D. W. Aswad. Formation of isoaspartate at two distinct sites during *in vitro* aging of human growth hormone. *J. Biol. Chem.*, 264:14262–14271, 1989.

[636] M. V. Paranandi, A. W. Guzzetta, W. S. Hancock, and D. W. Aswad. Deamidation and isoaspartate formation during *in vitro* aging of recombinant tissue plasminogen activator. *J. Biol. Chem.*, 269:243–253, 1994.

[637] N. E. Robinson and A. B. Robinson. Prediction of protein deamidation rates from primary and three-dimensional structure. *Proc. Natl. Acad. Sci. USA*, 98:4367–4372, 2001.

[638] N. E. Robinson. Protein deamidation. *Proc. Natl. Acad. Sci. USA*, 99:5283–5288, 2002.

[639] G. I. Szendrei, H. Fabian, H. H. Mantsch, S. Lovas, O. Nyeki, I. Schon, and L. Otvos Jr. Aspartate-bond isomerization affects the major conformation of synthetic peptides. *Eur. J. Biochem.*, 226:917–924, 1994.

[640] M. R. Nilsson, M. Driscoll, and D. P. Raleigh. Low levels of asparagine deamidation can have a dramatic effect on aggregation of amyloidogenic peptides: Implications for the study of amyloid formation. *Protein Sci.*, 11:342–349, 2002.

[641] Y. H. Kim, D. M. Kapfer, J. Boekhorst, N. H. Lubsen, H. P. Bachinger, T. R. Shearer, L. L. David, J. B Feix, and K. J. Lampi. Deamidation, but not truncation, decreases the urea stability of a lens structural protein, beta b1-crystallin. *Biochem.*, 41:14076–14084, 2002.

[642] V. N. Lapko, A. G. Purkiss, D. L. Smith, and J. B. Smith. Deamidation in human gamma s-crystallin from caractous lenses is influenced by surface exposure. *Biochem.*, 41: 8638–8648, 2002.

[643] R. J. Kapphahan, E. M. Ethen, E. A. Peters, L. Higgins, and D. A. Ferrington. Modified α-a crystallin in the retina: Altered expression and truncation with aging. *Biochem.*, 42: 15310–15325, 2003.

[644] M. J. Harms, P. A. Wilmarth, D. M. Kapfer, E. A. Steek, L. L. David, H. P. Bachinger, and K. J. Lampi. Laser light-scattering evidence for an altered association of β b1-crystallin deamidated in the connecting peptide. *Protein Sci.*, 13:678–686, 2004.

[645] K. Amyx P. Ahmann and K. J. Lampi. Deamidation at the N-terminal domain interface (Q70E), but not the homologue C-terminal domain interface (Q162E) of beta b2 crystallin altered the protein stability. *Invest. Ophthalmol. Vis. Sci.*, 46:E3487, 2005.

[646] K. J. Lampi, M. Harms, and D. M. Kapfer. Deamidation at asparagine 157 in human betaB1-crystallin. *Invest. Ophthalmol. Vis. Sci.*, 44:E2358, 2003.

[647] K. J. Lampi, J. T. Oxford, H. P. Bachinger, T. R. Shearer, L. L. David, and D. M. Kapfer. Deamidation of human beta b1 alters the elongated structure of the dimer. *Exp. Eye Res.*, 72: 279–288, 2001.

[648] T. Shimizu, A. Watanabe, M. Ogawara, H. Mori, and T. Shirasawa. Isoaspartate formation and neurodegeneration in Alzheimer's disease. *Arch. Biochem. Biophys.*, 381:225–234, 2000.

[649] H. Yang, Y. Lyutvinskiy, H. Soininen, and R. A. Zubarev. Alzheimer's disease and mild cognitive impairment are associated with elevated levels of isoaspartyl residues in blood plasma proteins. *J. Alzheimer's Dis.*, 27:113–118, 2011.

[650] H. Fabian, G. I. Szendrei, H. H. Mantsch, B. D. Greenberg, and L. Otvos Jr. Synthetic post-translationally modified human Aβ peptide exhibits a markedly increased tendency to form β-pleated sheets *in vitro*. *Eur. J. Biochem.*, 221:959–964, 1994.

[651] T. Shimizu, H. Fukuda, S. Murayama, N. Izumiyama, and T. Shirasawa. Isoaspartate formation at position 23 of amyloid beta peptide enhanced fibril formation and deposited onto senile plaques and vascular amyloids in Alzheimer's disease. *J. Neurosci. Res.*, 70: 451–461, 2002.

[652] L. Bohme, T. Hoffmann, S. Manhard, R. Wolf, and H. U. Demuth. Isoaspartate-containing amyloid precursor protein-derived peptides alter efficacy and specificity of potential β-secretases. *Biol. Chem.*, 389:1055–1066, 2008.

[653] J. Orpizewski, N. Schormann, , B. Kluve-Beckermann, J. J. Liepni, and M. D. Benson. Protein aging hypothesis of Alzheimer disease. *FASEB J.*, 14:1255–1263, 2000.

[654] M. R. Nilsson and C. M. Dobson. Chemical modification of insulin in amyloid fibrils. *Protein Sci.*, 12:2637–2641, 2003.

[655] M. L. Xie, Z. Shahrokh, M. Kadkodayan, W. J. Henzel, M. F. Powell, R. T. Borchart, and R. L. Schowen. Asparagine deamidation in recombinant human lymphotoxin: Hindrance by three-dimensional structures. *J. Pharm. Sci.*, 92:869–880, 2003.

[656] M. Xie, J. Aube, R. T. Borchart, M. Morton, E. M. Topp, D. Vander Velde, and R. L. Schowen. Reactivity toward deamidation of asparagine residues in beta-turn structures. *J. Pept. Res.*, 56:165–171, 2000.

[657] S. N. McAdam, B. Fleckenstein, I. B. Rasmussen, D. G. Schmid, I. Sandlie, B. Bogen, N. J. Viner, and L. M. Sollid. T cell recognition of the dominant i-a^k-restricted hen egg lysozyme epitope: Critial role for asparagine deamidation. *J. Exp. Med.*, 193:1239–1246, 2001.

[658] M. F. Mazzeo, B. De Giulio, S. Senger, M. Rossi, A. Malorni, and R. A. Siciliano. Identification of trans glutaminase-mediated deamidation sites in a recombinant alpha-gliadin by advanced mass spectrometric methodologies. *Protein Sci.*, 12:2434–2442, 2003.

[659] H. Sjostrom, K. E. A. Lundin, O. Molberg, R. Korner, S. N. McAdam, D. Anthonsen, H. Quarsten, O. Noren, P. Roepstorff, E. Thorsby, and L. M. Sollid. Identification of a gliadin T-cell epitope in coeliac disease: General importance of gliadin deamidation for intestinal T-cell recognition. *Scand. J. Immunol.*, 48:111–115, 1998.

[660] S. J. Weintraub and S. R. Manson. Asparagine deamidation: A regulatory hourglass. *Mech. Ageing Dev.*, 125:255–257, 2004.

[661] B. L. Cook B. E. Deverman, S. R. Manson, R. A. Niederhoff, E. M. Langer, I. Rosova, L. A. Kulans, X. Y. Fu, J. S. Weinberg, J. W. Heinecke, K. A. Roth, and S. J. Weintraub. Bcl-X-l deamidation is a critical switch in the regulation of the response to DNA damage. *Cell*, 111: 51–62, 2002.

[662] T. Takehara and H. Takahashi. Asparagine deamidation as a novel posttranslation modification of bcl-xl. *Gastroenterology*, 63:A443, 2000.

[663] T. Takehara and H. Takahashi. Suppression of bcl-xl deamidation in human hepatocellular carcinomas. *Cancer Res.*, 63:3054–3057, 2003.

[664] F. E. Dorer, E. E. Haley, and D. L. Buchanan. The hydrolysis of β-aspartyl peptides by rat tissue. *Arch. Biochem. Biophys.*, 127:490–495, 1968.

[665] B. A. Johnson, E. D. Murray, S. Clarke, D. B. Glass, and D. W. Aswad. Protein carboxyl methyltransferase facilitates conversion of atypical L-isoaspartyl peptides to normal L-aspartyl peptides. *J. Biol. Chem.*, 262:5622–5629, 1987.

[666] L. Bohme, J. W. Bar, T. Hoffmann, S. Manhart, H.-H. Ludwig, F. Rosche, and H. U. Demuth. Isoaspartate residues dramatically influence substrate recognition and turnover by proteases. *Biol. Chem.*, 389:1043–1053, 2008.

[667] D. Kameoka, T. Ueda, and T. Imoto. A method for the detection of asparagine deamidation and aspartate isomerization of proteins by MALDI/TOF-mass spectrometry using endoproteinase Asp-N. *J. Biochem.*, 134:129–135, 2003.

[668] A. Di Donato, P. Galletti, and G. D'Alessio. Selective deamidation and enzymatic methylation of seminal ribonuclease. *Biochem.*, 25:8361–8363, 1986.

[669] A. Di Donato, M. A. Ciardiello, M. Denigris, R. Piccoli, L. Mazzarella, and G. D'Alessio. Selective deamidation of ribonuclease-A: Isolation and characterization of the resulting isoaspartyl and aspartyl derivatives. *J. Biol. Chem.*, 268:4754–5751, 1993.

[670] D. W. Aswad. Stoichiometric methylation of porcine adrenocorticotropin by carboxyl methyltansferase requires deamidation of asparagine 25. *J. Biol. Chem.*, 259:10714–10721, 1984.

[671] E. D. Murray Jr. and S. Clarke. Synthetic peptide substrates for the erythrocyte protein carboxyl methyltransferase. Detection of a new site of methylation at isomerized L-aspartyl residues. *J. Biol. Chem.*, 259:10722–10732, 1984.

[672] T. Shimizu, Y. Matsuoka, and T. Shirasawa. Biological significance of isoaspartate and its repair system. *Biol. Pharm. Bull.*, 28:1590–1596, 2005.

[673] J. J. Cournoyer and P. B. O'Connor. Analysis of deamidation in proteins. In J. Whitelegge, editor, *Mass Spectrometry Analysis of Proteins*. Elsevier, Netherlands, 2009.

[674] D. G. Smyth, W. H. Stein, and S. Moore. On the sequence of residues 11 to 18 in bovine pancreatic ribonuclease. *J. Biol. Chem.*, 237:1845–1850, 1962.

[675] D. G. Smyth, W. H. Stein, and S. Moore. The sequence of amino acid residues in bovine pancreatic ribonuclease: Revisions and confirmations. *J. Biol. Chem.*, 238:227–234, 1963.

[676] Y. Shin, H. S. Cho, H. Fukumoto, T. Shimizu, T. Shirasawa, S. M. Greenberg, and G. W. Rebeck. Aβ species, including isoAsp23 Aβ, in Iowa-type familial cerebral amyloid angiopathy. *Acta Neuropathol.*, 105:252–258, 2003.

[677] W. D. Lehmann, A. Schlosser, G. Erben, R. Pipkorn, D. Bossemeyer, and V. Kinzel. Analysis of isoaspartate in peptides by electrospray tandem mass spectrometry. *Protein Sci.*, 9: 2260–2268, 2000.

[678] M. J. Mamula, R. J. Gee, J. I. Elliott, A. Sette, S. Southwood, P. Jones, and P. R. Blier. Isoaspartyl post-translational modification triggers autoimmune responses to self-proteins. *J. Biol. Chem.*, 274:22321–22327, 1999.

[679] W. J. Chazin, J. Kördel, E. Thulin, E. Hofmann, T. Drakenberg, and S. Forsén. Identification of an isoaspartyl linkage formed upon deamidation of bovine calbindin D_{9k} and structural characterization by 2D ^{1}H NMR. *Biochem.*, 28:8646–8653, 1989.

[680] R. J. Waugh, S. T. Steinborner, J. H. Bowie, J. C. Wallace, M. J. Tyler, P. F. Hu, and M. L. Gross. Two isomeric alpha and beta aspartyl dodecapeptides and their cyclic amino succinyl analogue from the Australian green tree frog *Litoria gilleni*. *Aust. J. Chem.*, 48:1981–1987, 1995.

[681] M. C. Manning, D. K. Chou, B. M. Murphy, R. W. Payne, and D. S. Katayama. Stability of protein pharmaceutics: An update. *Pharm. Res.*, 27:544–575, 2010.

[682] A. A. Wakankar and R. T. Borchardt. Formulation considerations for protein susceptible to asparagine deamidation and aspartate isomerization. *J. Pharm. Sci.*, 95:2321–2336, 2006.

[683] G. Chen, B. M. Warrack, A. K. Goodenough, H. Wei, D. B. Wang-Iverson, and A. A. Tymiak. Characterization of protein therapeutics by mass spectrometry: Recent developments and future directions. *Drug Discov. Today*, 16:58–64, 2011.

[684] F. Curnis, R. Longhi, L. Crippa, A. Cattaneo, E. Dondossola, A. Bachi, and A. Corti. Spontaneous formation of L-isoaspartate and gain of function in fibronectin. *J. Biol. Chem.*, 281:36466–36476, 2006.

[685] E. Sikorska, K. Greber, S. Rodziewicz-Motowidło, Ł. Szultka, J. Łukasiak, and W. Kamysz. Synthesis and antimicrobial activity of truncated fragments and analogs of citropin 1.1: The solution structure of the SDS micelle-bound citropin-like peptides. *J. Struct. Biol.*, 168: 250–258, 2009.

[686] A. Rozek, C. L. Friedrich, and R. E. W. Hancock. Structure of the bovine antimicrobial peptide indolicidin bound to dodecylphosphocholine and sodium dodecyl sulfate micelles. *Biochem.*, 39:15765–15774, 2000.

[687] R. P. Hicks, E. Mones, H. Kim, B. W. Koser, D. A. Nichols, and A. K. Bhattacharjee. Comparision of the conformation and electrostatic surface properties of magainin peptides bound to sodium dodecyl sulfate and dodecylphosphocholine micelles. *Biopolymers*, 68:459–470, 2003.

[688] K. J. Marcinowski, H. Shao, E. L. Clancy, and M. G. Zagorski. Solution structure model of residues 1-28 of the amyloid β-peptide when bound to micelles. *J. Am. Chem. Soc.*, 120: 11082–11091, 1998.

[689] S. Kosol and K. Zangger. Dynamics and orientation of a cationic antimicrobial peptide in two membrane-mimetic systems. *J. Struct. Biol.*, 170:172–179, 2010.

[690] W. Zhang, M. J. Czupryn, P. T. Boyle, and J. Amari. Characterization of asparagine deamidation and aspartate isomerization in recombinant human interleukin-11. *Pharm. Res.*, 19:1223–1231, 2002.

[691] W. Ni, S. Dai, B. L. Karger, and Z. S. Zhou. Analysis of isoaspartic acid by selective proteolysis with Asp-N and electron transfer dissociation mass spectrometry. *Anal. Chem.*, 82:7485–7491, 2010.

[692] D. S. Rehder, D. Chelius, A. McAuley, T. M. Dillon, G. Xiao, J. Crouse-Zeineddini, L. Vardanyan, N. Perico, V. Mukku, D. N. Brems, M. Matsumura, and P. V. Bondarenko. Isomerization of a single aspartyl residue of anti-epidermal growth factor receptor immunoglobulin $\gamma 2$ antibody highlights the role avidity plays in antibody activity. *Biochem.*, 47:2518–2530, 2008.

[693] C. M. Deber and S. C. Li. Peptides in membranes: Helicity and hydrophobicity. *Biopolymers*, 37:295–318, 1995.

[694] E. F. Haney, H. N. Hunter, K. Matsuzaki, and H. J. Vogel. Solution NMR studies of amphibian antimicrobial peptides: Linking structure to function? *Biochim. Biophys. Acta*, 1788: 1639–1655, 2009.

[695] W. Chen, N. J. Ede, D. C. Jackson, J. McCluskey, and A. W. Purcell. CTL recognition of an altered peptide associated with asparagine bond rearrangement. Implications for immunity and vaccine design. *J. Immunol.*, 157:1000–1005, 1996.

[696] C. L. Stevenson, M. E. Donlan, A. R. Friedman, and R. T. Borchardt. Solution conformation of Leu^{27} hGRF(1-32)NH_2 and its deamidation products by 2D NMR. *Int. J. Pept. Protein Res.*, 42:24–32, 1993.

[697] J. Vlasak, M. C. Bussat, S. Wang, E. Wagner-Rousset, M. Schaefer, C. Klinguer-Hamour, M. Kirchmeier, N. Corvaïa, R. Ionescu, and A. Beck. Identification and characterization of asparagine deamidation in the light chain CDR1 of a humanized IgG1 antibody. *Anal. Biochem.*, 392:145–154, 2009.

[698] S. Capasso, A. Di Donato, L. Esposito, F. Sica, G. Sorrentino, L. Vitagliano, A. Zagari, and L. Mazzarella. Deamidation in proteins: The crystal structure of bovine pancreatic ribonuclease with an isoaspartyl residue at position 67. *J. Mol. Biol.*, 257:492–496, 1996.

[699] S. Noguchi, K. Miyawaki, and Y. Satow. Succinimide and isoaspartate residues in the crystal structures of hen egg-white lysozyme complexed with tri-N-acetylchitotriose. *J. Mol. Biol.*, 278:231–238, 1998.

[700] S. Noguchi, Y. Satow, T. Uchida, C. Sasaki, and T. Matsuzaki. Crystal structure of *Ustilago sphaerogena* ribonuclease U_2 at 1.8 Å resolution. *Biochem.*, 34:15583–15591, 1995.

[701] R. V. Baudinette, P. Boontheung, I. F. Musgrave, P. A. Wabnitz, V. M. Maselli, J. Skinner, P. F. Alewood, C. S. Brinkworth, and J. H. Bowie. An immunomodulator used to protect young in the pouch of the Tammar wallaby, *Macropus eugenii*. *FEBS J.*, 272:433–443, 2005.

[702] E. S. Furfine, M. F. Harmon, J. E. Paith, and E. P. Garvey. Selective inhibition of constitutive nitric oxide synthase by L-N^G-nitroarginine. *Biochem.*, 32:8512–8517, 1993.

[703] L. E. Lambert, J. P. Whitten, B. M. Baron, H. C. Cheng, N. S. Doherty, and I. A. McDonald. Nitric oxide synthesis in the CNS, endothelium and macrophages differs in its sensitivity to inhibition by arginine analogues. *Life Sci.*, 48:69–75, 1991.

[704] J. H. Jorgensen, W. A. Cleeland, G. Craig, M. Doern, J. Ferraro, C. M. Finegold, S. L. Hansen, S. G. Jenkins, W. J. Novick, M. S. Pfaller, D. A. Preston, L. B. Reller, and J. M. Swanson. Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically, 3rd approved standard. *National Committee for Clinical Laboratory Standards*, 33:1–12, 1993.

[705] L. Zhang, K. Yan, Y. Zhang, R. Huang, J. Bian, C. Zheng, H. Sun, Z. Chen, N. Sun, R. An, F. Min, W. Zhao, Y. Zhuo, J. You, Y. Song, Z. Yu, Z. Liu, K. Yang, H. Gao, H. Dai, X. Zhang, J. Wang, C. Fu, G. Pei, J. Liu, S. Zhang, M. Goodfellow, Y. Jiang, J. Kuai, G. Zhou, and X. Chen. High-throughput synergy screening identifies microbial metabolites as combination agents for the treatment of fungal infections. *Proc. Natl. Acad. Sci. USA*, 104:4606–4611, 2007.

Appendix A
Assigned Chemical Shifts of Calmodulin Bound to Caerin 1.8

Table 1: Assigned N, HN, Cα and Cβ chemical shifts for Ca_4^{2+}CaM bound to caerin 1.8.11 at pH 6.3, 25 °C, 31 mM KCl, 6.3 mM $CaCl_2$. Where no resonance is indicated it could not be reliably assigned.

		Chemical shift (ppm)			
Residue	**Amino Acid**	**HN**	**N**	**αC**	**βC**
1	A				
2	D				
3	Q	8.40	120.0	55.61	29.87
4	L	8.35	123.6	54.54	43.68
5	T	8.75	113.1	60.59	71.14
6	E	9.04	120.5	60.23	29.35
7	E	8.73	119.7	60.08	29.24
8	Q	7.76	119.8	58.84	29.50
9	I	8.46	119.9	66.15	38.00
10	A	8.05	121.5	55.50	18.11
11	E	7.87	120.7	59.31	29.53
12	F	8.46	118.6	60.48	38.40
13	K	9.25	123.7	60.26	31.90
14	E	7.90	120.6	59.30	29.30
15	A	8.06	120.9	55.18	18.43
16	F	8.97	119.6	62.23	39.41
17	S	8.00	112.7	61.52	63.20
18	L	7.41	120.1	57.13	41.59
19	F	7.45	115.3	58.39	40.53
20	D	7.98	117.5	52.53	39.18
21	K	7.61	124.8	58.47	32.68
22	D	8.33	114.3	53.12	39.76
23	G	7.76	109.2	47.20	-
24	D	8.40	120.7	53.90	40.55
25	G	10.65	113.2	45.87	-
26	T	8.15	113.2	59.93	72.68
27	I	9.89	127.1	60.46	39.78
28	T	8.48	116.9	59.68	72.62
29	T	9.23	112.5	66.49	67.75
30	K	7.65	121.3	59.37	32.67
31	E	7.68	121.6	59.42	30.00
32	L	8.87	120.8	58.05	42.98
33	G	8.75	105.5	48.52	-
34	T	7.91	117.8	66.88	68.74
35	V	7.37	122.2	66.80	31.49
36	M	8.15	120.3	60.08	30.17

Continued on next page

Table 1 – *Continued from previous page*

		Chemical shift (ppm)			
Residue	**Amino Acid**	**HN**	**N**	**αC**	**βC**
37	R	8.69	118.3	59.92	29.27
38	S	7.83	117.9	61.83	63.02
39	L	7.36	119.4	53.86	41.07
40	G	7.52	136.0	46.05	-
41	Q	7.94	118.3	54.20	30.57
42	N	8.76	116.2	51.46	39.32
43	P	-		62.45	32.08
44	T	8.83	113.2	60.57	71.17
45	E	8.83	120.6	60.12	29.07
46	A	8.28	120.8	55.21	-
47	E	7.97	119.3	59.57	29.58
48	L	8.21	120.7	58.00	42.64
49	Q	8.28	118.5	58.66	29.58
50	D	8.22	120.2	57.74	40.36
51	M	7.87	119.5	59.26	33.35
52	I	7.62	118.0	65.16	37.49
53	N	8.78	118.1	56.15	38.01
54	E	7.62	116.3	59.11	30.59
55	V	7.19	106.5	60.24	32.68
56	D	7.55	122.4	53.87	40.22
57	A	8.62	131.9	54.33	19.61
58	D	8.14	113.7	52.65	39.92
59	G	7.58	108.3	47.40	-
60	N	8.06	118.4	52.65	37.82
61	G	10.66	113.4	45.89	-
62	T	7.71	109.0	59.68	72.02
63	I	8.94	123.4	60.00	39.99
64	D	8.83	128.1	52.35	42.23
65	F	8.95	118.7	58.04	39.76
66	P	-		66.13	29.77
67	E	8.04	121.9	59.39	29.50
68	F	8.94	123.4	60.97	39.99
69	L			58.10	40.78
70	T	7.42	116.3	66.40	68.28
71	M	8.11	122.1	59.26	34.04
72	M				
73	A			55.77	18.04
74	R	8.21	118.3	59.29	30.58
75	K	7.80	118.3	58.07	32.41
76	M	8.03	116.4	57.26	33.24
77	K	7.42	118.6	57.08	33.17
78	D	7.82	120.1	54.40	41.19
79	T	8.13	115.0	61.51	69.39
80	D	8.42	122.5	54.59	41.75
81	S	8.30	115.3	56.99	64.55
82	E	8.98	126.1	59.67	29.58
83	E	8.17	116.1	59.03	31.20
84	E	8.59	118.2	58.99	30.67
85	I	7.87	119.3	65.12	36.89
86	R	8.69	122.2	60.03	29.53
87	E	8.01	118.3	59.38	29.46
88	A	7.78	121.3	55.24	18.51
89	F	8.47	117.9	62.15	39.62
90	R	7.90	115.9	58.70	30.56
91	V	7.34	118.0	65.85	30.93
92	F	6.72	113.6	59.64	41.67
93	D	7.78	116.9	52.35	38.91
94	K	7.62	124.9	58.36	32.83

Continued on next page

Table 1 – *Continued from previous page*

		Chemical shift (ppm)			
Residue	**Amino Acid**	**HN**	**N**	**αC**	**βC**
95	D	8.03	113.8	52.80	39.58
96	G	7.73	109.3	47.41	-
97	N	8.43	119.9	52.89	38.36
98	G	10.52	112.5	45.23	-
99	Y	7.77	116.6	56.52	42.61
100	I	10.17	127.4	61.30	39.31
101	S	9.09	124.2	55.78	67.16
102	A	9.41	123.3	55.82	18.07
103	A	8.27	118.5	55.34	18.33
104	E	7.75	118.8	59.21	29.89
105	L	8.22	119.5	57.68	40.45
106	R	7.97	120.3	59.41	28.22
107	H			59.97	29.03
108	V	7.85	119.5	66.97	32.19
109	M	8.21	115.7	56.94	29.49
110	T	8.74	118.0	66.72	68.61
111	N	8.02	123.5	55.87	37.69
112	L	7.72	119.0	55.84	42.70
113	G	7.88	106.5	45.57	-
114	E	8.06	120.7	54.95	31.53
115	K	8.63	125.2	55.45	32.23
116	L	8.10	124.4	54.47	44.33
117	T	9.18	114.4	60.79	71.13
118	D	8.95	121.2	58.09	39.62
119	E	8.72	119.4	60.07	29.26
120	E	7.73	120.3	59.29	30.52
121	V			66.06	29.70
122	D	8.19	120.9	58.10	41.81
123	E	8.90	118.9	60.23	30.19
124	M	8.14	120.3	60.11	30.16
125	I				
126	R			59.36	29.29
127	E	7.80	118.7	59.46	30.35
128	A	7.38	116.0	50.92	23.75
129	D	8.20	122.7	55.26	41.13
130	I	8.46	128.0	63.03	39.52
131	D	8.31	117.3	53.98	39.91
132	G	7.75	108.5	47.45	-
133	D	8.36	120.3	53.78	40.10
134	G	9.93	112.0	46.11	-
135	Q	8.07	115.7	53.35	32.95
136	V	9.23	125.5	61.26	34.26
137	N	9.58	128.9	51.17	38.67
138	Y	8.32	118.7	62.42	38.41
139	E	8.19	118.7	60.48	28.93
140	E	8.82	120.1	58.91	29.35
141	F	8.66	124.2	62.60	40.44
142	V	8.87	119.7	67.30	31.49
143	Q	7.96	120.5	59.50	28.28
144	M	7.59	118.8	58.76	31.85
145	M	7.88	114.4	55.69	30.74
146	T	7.82	108.7	62.04	70.60
147	A	7.47	127.1	53.15	19.16
148	K	8.05	126.8	57.69	33.73

Publications

Publications Related to Ph.D. Thesis

Antonio N. Calabrese, Nikki J. Good, Tianfang Wang, Jingjia He, John H. Bowie and Tara L. Pukala. A negative ion mass spectrometry approach to identify cross-linked peptides utilizing characteristic disulfide fragmentations, *J. Am. Soc. Mass Spectrom.*, 23:1364-1375, 2012

Antonio N. Calabrese, Katarina Markulic, Ian F. Musgrave, Hui Guo, Lixin Zhan and John H. Bowie. Structural and activity changes in three bioactive amphibian peptides when Asp is replaced by *iso*Asp, *Peptides*, 38:427-436, 2012

Antonio N. Calabrese, Tianfang Wang, John H. Bowie and Tara L. Pukala. Negative ion fragmentations of disulfide-containing cross-linking reagents are competitive with aspartic acid side-chain induced cleavages, *Rapid Comm. Mass Spectrom.*, 27:238-248, 2013

Additional Publications

Tianfang Wang, Hayley J. Andreazza, Tara L. Pukala, Patrick J. Sherman, Antonio N. Calabrese, John H. Bowie, Histidine-containing host-defence skin peptides of anurans bind Cu^{2+}. An electrospray ionisation mass spectrometry and computational modelling study, *Rapid Comm. Mass Spectrom.*, 25:1209-1221, 2011

Tianfang Wang, T. T. Nha Tran, Antonio N. Calabrese, John H. Bowie, Backbone fragmentations of $[M\text{-}H]^-$ anions from peptides. Reinvestigation of the mechanism of the beta prime cleavage, *Rapid Comm. Mass Spectrom.*, 26:1832-1840, 2012

Antonio N. Calabrese, Lauren A. Speechley and Tara L. Pukala, Characterisation of Calmodulin structural transitions by ion mobility mass spectrometry, *Aust. J. Chem.*, 65:504-511, 2012

Calabrese, A.N., Good, N.J., Wang, T., He, J., Bowie, J.H. & Pukala, T.L. (2012) A Negative Ion mass spectrometry approach to identify cross-linked peptides utilizing characteristic disulfide fragmentations.
Journal of the American Society for Mass Spectrometry, v. 23(8), pp. 1364-1375

Calabrese, A.N., Markulic, K., Musgrave, I.F., Guo, H., Zhang, L. & Bowie, J.H. (2012) Structural and activity changes in three bioactive anuran peptides when Aspis replaced by *iso*Asp.
Peptides, v. 38(2), pp. 427-436

Calabrese, A.N., Wang, T., Bowie, J.H. & Pukala, T.L. (2013) Negative ion fragmentations of disulfide-containing cross-linking reagents are competitive with aspartic acid side-chain-induced cleavages.
Rapid Communications in Mass Spectrometry, v. 27(1), pp. 238-248